Mild Cognitive Impairment Detection from EEG Signals Using Combination of EMD Decomposition and Machine Learning

Majid Aljalal
Electrical Engineering
King Saud University
Riyadh, Saudi Arabia
maljalal@ksu.edu.sa

Saeed A. Aldosari
Electrical Engineering
King Saud University
Riyadh, Saudi Arabia
dosari@ksu.edu.sa

Marta Molinas
Engineering Cybernetics
Norwegian University of Science and Technology
Trondheim, Norway
marta.molinas@ntnu.no

Khalil AlSharabi
Electrical Engineering
King Saud University
Riyadh, Saudi Arabia
kabdulghani@ksu.edu.sa

Fahd A. Alturki
Electrical Engineering
King Saud University
Riyadh, Saudi Arabia
falturki@ksu.edu.sa

Abstract— **Mild cognitive impairment (MCI) is the earliest stage of dementia, and its detection is crucial for disease management. Electroencephalography (EEG) has gained popularity as a tool for identifying brain disorders. This article presents methods for diagnosing MCI from EEG signals using empirical mode decomposition (EMD). The proposed methods are tested using EEG data from 32 healthy control (HC) cases and 29 MCI cases. First, major artifacts are removed through preprocessing. EMD is then used to decompose signals into intrinsic mode functions (IMFs), from which features are extracted using one of the following measures: energy, log energy entropy, sure entropy, threshold entropy, Shannon entropy, or band power. Multiple machine learning approaches are then used to classify whether the resulting features are for MCI or HC. Results reveal that the proposed methods, especially the one that combines EMD and entropy, produce competitive results. For instance, the achieved results for a KNN classifier using sure entropy in terms of classification accuracy, specificity, sensitivity, and F-score are 97.60%, 98.59%, 96.49%, and 97.44%, respectively. The findings demonstrate accurate detection of MCI using the proposed methods, indicating potential for clinical diagnosis applications.**

***Keywords*— EEG, empirical mode decomposition (EMD), machine learning, mild cognation impairment (MCI).**

I. Introduction

Dementia encompasses the progressive decline in mental abilities such as memory, speech, and cognition, leading to a hindrance in daily functioning [1]. This condition is more commonly observed in individuals aged sixty and above, with one of the initial symptoms being difficulty in recalling recent events [2]. Mild Cognitive Impairment (MCI) signifies the early stage of Alzheimer's disease (AD) and other forms of dementia. It is characterized by noticeable cognitive changes in affected individuals that do not significantly impede their daily activities [3, 4]. However, individuals with MCI face an elevated risk of developing AD or other types of dementia in the future, with approximately 15-20% of MCI patients progressing to AD each year [5]. The diagnosis of MCI or AD typically involves a time-consuming and multi-step process, incorporating procedures such as magnetic resonance imaging (MRI), computed tomography (CT), mini-mental state examinations (MMSE), blood tests, neurological examinations, positron emission tomography (PET), and spinal fluid analyses [6]. Therefore, to facilitate early intervention for MCI, it is crucial to develop MCI detection methods that are safe, objective, and easy to implement.

Several studies have investigated the diagnosis of MCI using different modalities: MRI-based methods [7, 8], PET-based methods [9], CT-based methods [10], and multi-modal-based methods [11]. Another modality, electroencephalography (EEG), has been recently used to automatically diagnose several brain disorders. EEG is a non-invasive method that captures changes in brain activity by detecting the electrical signals produced by thousands of neurons. This is achieved through the placement of electrodes, or channels, on the scalp [12]. In contrast to other imaging techniques like MRI, CT, and PET, EEG recording systems possess several advantageous characteristics, including high temporal resolution, portability, cost-effectiveness, and efficiency [12]. Due to these advantages, EEG analysis, in conjunction with machine learning approaches, has been widely employed to identify various neurological disorders. For instance, it has been utilized in the detection of autism spectrum disorder (ASD) [13, 14], epilepsy [13, 15], AD [16, 17], schizophrenia [18], Parkinson's disease (PD) [19, 20], major depressive disorder [21], as well as tasks such as emotion recognition [22]. In recent times, there have been a growing number of studies that utilize EEG signals and machine learning techniques to automatically detect MCI. These studies have explored both task-state and resting-state EEG paradigms [23–34]. In task-state EEG experiments, participants are given specific tasks to perform. For instance, in one study [23], subjects were exposed to a series of five contrasting speech sounds played sequentially, and they were required to identify these sounds while their EEG data was recorded from a single channel. On the other hand, resting-state EEG recordings involve capturing individuals' brain activity when they have their eyes closed or open, without any specific task assigned. Resting-state EEG data collection is considered more convenient, practical, and comfortable, especially for elderly individuals, as it mimics real-world circumstances. A recent systematic review [35] focusing on the utilization of resting-state EEG for diagnosing AD and evaluating its progression revealed a substantial body of literature (48 studies) that primarily focused on distinguishing AD patients from healthy controls (HC). However, there has been relatively less emphasis on discerning the differences between individuals with MCI and HC.

2023 3rd International Conference on Emerging Smart Technologies and Applications (eSmarTA) | 979-8-3503-0533-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/ESMARTA59349.2023.10293374

In the context of MCI diagnosis in resting-state EEGs, M. Kashefpoor et al. [24] proposed a technique to distinguish between individuals with MCI and HC utilizing basic spectral EEG features. With an accuracy of 88.89%, a neurofuzzy system and a K-nearest neighbor (KNN) classifier are combined to classify certain features from 11 MCI and 16 HC participants. Three years later, the same group [25] added more cases (18 MCI and 16 HC), bringing the total to 29 MCI and 32 HC, and proposed a method called correlation-based label-consistent K-SVD (CLC-KSVD). This method is based on supervised dictionary-learning-based analysis of EEG signals, which had an accuracy of 80%. The dataset in [24] was also employed by S. Hadiyoso et al. [26], along with power spectral characteristics and a KNN classifier, which led to an 81.5% classification accuracy. In [27], Hsiao, Y.-T., et al. proposed an MCI vs. HC classifier using the between-electrode kernel eigen-relative-power (KERP). Support vector machines (SVM) were then used to classify the generated KERP features with an accuracy of 90.20%. The dataset utilized in [24] was also used by Yin, Jiao, et al. in [28]. They excluded five HC cases from the initial dataset in order to make the number of MCI and HC cases equal. They used stationary wavelet transformation (SWT) to reduce noise before extracting nine statistical features, such as median, standard deviation, mean, and mode. SVM was then used to classify the extracted features with an accuracy of 96.94%. The dataset used in [24] was also used by S. Siuly et al. in [29]. In their study, an extreme learning machine (ELM) was used to extract and classify features from the auto-regressive and permutation entropy models with an accuracy of 98.78%. To distinguish MCI from HC, A. M. Alvi et al. [30] developed a deep learning-based framework based on the long-short-term memory (LSTM) model. To determine the best one, they created 20 different LSTM models and tested them on the same dataset as in [24]. The best model has an accuracy rate of 96.41%. In a study conducted by K. Lee et al. [31], several features were extracted from EEG signals, including absolute and relative power spectrum density, differential and rational asymmetry, phase-amplitude coupling, Shannon entropy, Hjorth parameters, Lyapunov exponent, Hurst exponent, and Kolmogorov complexity. The study utilized a dataset of 21 participants with MCI and 21 HC participants. By employing the SVM classifier and selecting eight channels, the highest classification accuracy achieved was 86.85%. Another study by R.A. Movahed et al. [32] extracted a total of 425 features, including spectral, functional connectivity, and nonlinear features, from a dataset of 19-channel EEGs of 18 MCI patients and 16 HC participants. By employing the linear SVM classifier, an impressive accuracy of 99.4% was achieved. More recently, A. Said and H. Göker [33] utilized the discrete wavelet transform leader (DWT) leader to extract MCI and HC features from the same dataset utilized in [32]. Multiple ensemble learning algorithms were examined for classification, and the AdaBoostM1 algorithm yielded the highest accuracy of 93.50%.

It is important to highlight that the proper design of the feature extraction stage can significantly enhance classification accuracy. Previous attempts of feature extraction approaches for the diagnosis of MCI have been insufficient, and there are still effective and well-known methods, such as empirical mode decomposition (EMD), that have not yet been explored. Additionally, there is a need to investigate more distinct resting-state EEG biomarkers that can differentiate between MCI and HC. Since EEG signals are non-stationary and nonlinear, decomposing an EEG signal into sub-signals with different frequency bands and extracting nonlinear features from each band may yield valuable biomarkers for MCI. In a previous study [20], the use of DWT signal decomposition was used in the context of Parkinson's disease classification. Results indicated a clear improvement in classification accuracy when the EEG signals were decomposed. This finding motivated further exploration of whether decomposing EEG signals and utilizing nonlinear features could lead to effective biomarkers for distinguishing MCI patients' resting-state EEGs from those of healthy individuals.

The objective of the current study is to introduce simple yet accurate models for MCI. The study contributes in several aspects. First, it presents novel and effective approaches based on EMD and machine learning for detecting MCI from resting-state EEG signals. To the best of our knowledge, no previous study has utilized EMD for decomposing signals before extracting EEG features for MCI diagnosis. Second, the study explores the use of various nonlinear measures to extract features from the decomposed signals. Specifically, features are extracted by computing energy, log energy entropy, sure entropy, threshold entropy, Shannon entropy, or band power. The combination of EMD signal decomposition and those measures is expected to produce effective biomarkers for MCI. The resulting features are then classified using different machine learning techniques, including quadratic discriminant analysis, random forest, k-nearest neighbor, and support vector machines. To ensure a balanced and larger number of subjects for analysis, similar to [25], the performance of the proposed methods is evaluated on a public dataset used in [24, 26, 28–30] (consisting of 11 MCI and 16 HC subjects) combined with the public dataset used in [32, 33] (comprising 18 MCI and 16 HC subjects).

The remainder of this paper is organized as follows. Section II outlines the EEG dataset utilized in this work as well as the pre-processing, feature extraction strategies, and classification methods used. Section III displays the classification results of the proposed strategies, while the discussion is included in Section IV. Section V presents the conclusions and some ideas for future research.

II. Methods

In this section, the proposed methods for processing EEG signals, encompassing preprocessing, feature extraction, and classification techniques, are presented. An overview of the various steps involved in EEG analysis and classification, conducted on both healthy volunteers and MCI patients, is summarized in Fig. 1. The raw EEG signals are initially read and segmented into non-overlapping segments. To identify the relevant frequency range, these segments undergo band-pass filtering. Following this, the EMD algorithm is applied for signal decomposition (see Fig. 2 for more details). Subsequently, several measures are employed to extract MCI and HC features from the decomposed signals. These measures include energy, band power, log energy entropy, Shannon entropy, sure entropy, and threshold entropy. To differentiate the features of MCI from those of healthy controls, multiple classifiers are utilized, such as random forest (RF), quadratic discriminant analysis (QDA), KNN, and SVM. The following subsections provide detailed information about each stage depicted in Fig. 1.

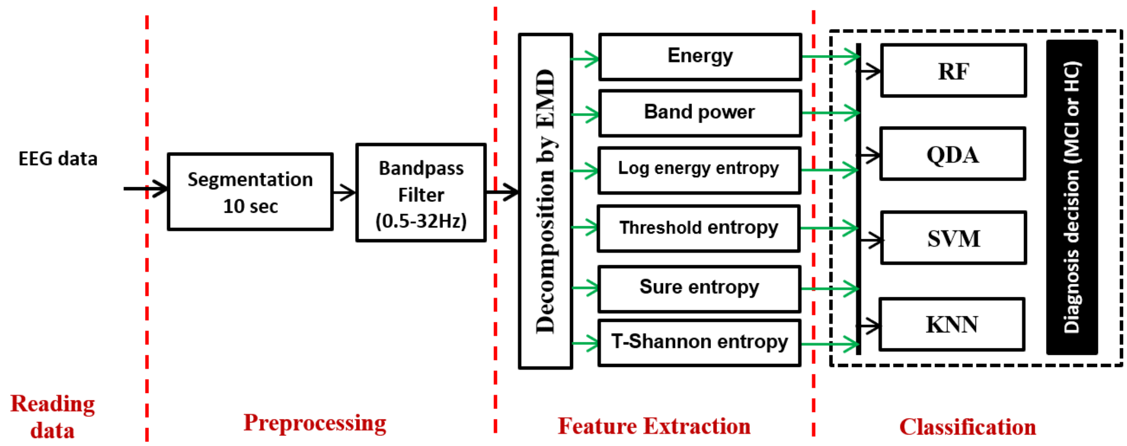

Fig. 1 Block diagram of the proposed methods.

A. Data used and pre-processing

Two freely available datasets from biosegdata.com are used to test the approaches employed in this research. [36]. The first dataset comprises 11 MCI patients and 16 normal cases, while the second dataset consists of 18 MCI patients and 16 normal cases. In order to create a unified dataset that is larger and balanced, similar to the approach taken by the authors of [25], the two datasets are combined in the present study. The combined dataset contains 61 people who are older than 55, for which 29 subjects are diagnosed with MCI and 32 normal subjects. Every participant, at the very least, had their elementary schooling completed. This study was conducted at Noor Hospital in Isfahan, Iran, and involved subject selection, cognitive evaluations, EEG recordings, and other experimental techniques. Each participant was given a comprehensive explanation of the research procedure prior to providing their informed consent. Each participant completed a neuropsychiatric interview in accordance with Peterson's MCI criteria [37]. Scores on the MMSE ranging from 21 to 26 were classified as indicative of MCI, while scores above 27 were considered normal. Additional support for the diagnosis of MCI was found through the use of the Neuropsychiatry Unit Cognitive Assessment Tool, which produced results between 75 and 86.5 [38]. All subjects conducted morning EEG recordings in a quiet room while lying down and keeping their eyes closed. The 19 EEG electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2) were positioned according to the 10-20 International System. To record the EEG, a Galileo NT (EBneuro) 32-channel digital EEG device was employed, and the electrodes' skin resistance was lower than 5 k Ω. Each participant had his or her EEG data captured for 30 minutes.

In order to mitigate participant fatigue resulting from prolonged recordings, the present study utilized only the initial 10 minutes of each recording. To minimize low-frequency drift, a high-pass filter at 0.5 Hz was applied. Additionally, manual inspection was performed to identify and remove artifacts caused by eye blinks, movements, muscle activity, electrical noise, and other types of interference. To enhance the quality of the data, additional preprocessing steps were undertaken. Specifically, a fifth-order band-pass Butterworth filter ranging from 0.5 Hz to 32 Hz was applied to each EEG recording. This filtering process effectively eliminates baseline artifacts and AC power-line noise, ensuring cleaner and more reliable EEG signals for subsequent analysis. The EEG signals are then divided into non-overlapping segments, with each segment consisting of ch×N samples. Here, ch represents the number of channels (which is 19 in the dataset used for this study), and N is the number of EEG samples per channel within a given 10-second time window (T).

B. Feature Extraction (FE)

Empirical mode decomposition (EMD):

The focus of this study is to decompose EEG signals using EMD into sub-signals. EMD is a well-known technique used to decompose a complex signal into a set of intrinsic mode functions (IMFs) corresponding to different frequency components. It was developed by Huang et al. in 1998 as a method for analyzing nonlinear and non-stationary data [39]. The decomposition is achieved by iteratively extracting the local extrema and calculating the mean of the upper and lower envelopes of the signal. This process is repeated until the resulting IMFs satisfy certain criteria, such as being well-behaved and having a mean of zero. Readers are recommended to read [39] for additional information on the EMD algorithm.

The number of IMFs is adjusted to 5 in order to acquire the 5 commonly adopted EEG sub-bands: delta (<4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (>30 Hz) [12]. Figure 2 shows an example of feature extraction from a 10-second segment using EMD and sure entropy. As shown in the figure, EMD decomposes each pre-processed and segmented signal from each channel into six sub-signals: five IMFs and one residual.

Features:

Following the EMD decompositions, features are computed from the resulting IMF signal in addition to the original signal (see Fig. 2). This study explores the use of various measures to extract the features, including energy (Eng), band power (LBP), and entropy. Entropy is a measure frequently utilized to evaluate the statistical quantification, regularity, and complexity of time series. Several studies have demonstrated the usefulness of entropy in analyzing and identifying biomarkers for various mental disorders, such as ASD [40], attention deficit

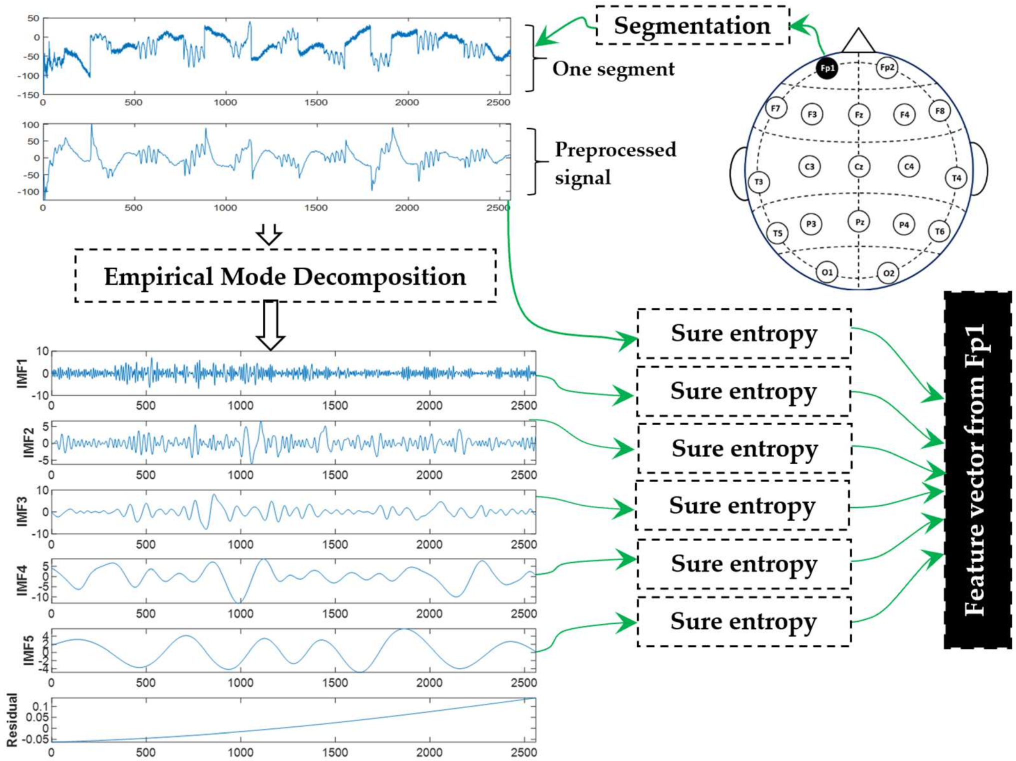

Fig. 2 Example of feature extraction from 10 sec segment using EMD and sure entropy.

hyperactivity disorder [41], and epilepsy [42]. This motivate us to investigate the effectiveness of entropy as a marker for identifying MCI. We propose computing entropy from the generated IMFs rather than directly from EEG data, which may help in the creation of effective biomarkers for MCI detection. To achieve this aim, several entropy measures, including sure entropy (SuEn), log energy entropy (LogEn), transformation-Shannon entropy (T-ShEn), and threshold entropy (ThEn), are investigated. These entropy measures, as well as energy and band power measures, are all defined in our prior works [19, 20]. Six features—five from IMF signals and one from the original signal segment—are computed using one of these measures. Each feature vector contains 6×ch features, where ch is the number of channels. In order to generate all feature vectors (feature matrix), this process is performed over all segments of all EEG recordings. The feature matrix is then fed into one of the classifiers.

C. Classification and Performance Evaluation

In this study, several classification methods are compared: RF, QDA, SVM, and KNN, to identify which method produces the best MCI versus HC classification accuracy. The set of classifier parameters shown in Table I was determined through a manual search for the optimal settings of each classifier. References [43–46] provide a comprehensive description of the considered classification methods.

Table I: Parameters used in the classifiers

Classifier	Parameters
SVM	kernel = linear, method=' least square', C = 2e-1
KNN	no. of neighbors = 3, distance ='euclidean', rule=' nearest'
RF	learner type = 'decision tree', ensemble = 'bag', no. of learners = '30'

To assess the performance of each classifier, 10-fold cross-validation is used, in which the dataset is split randomly into ten equal subsets. One of the subsets is for validation, while the remaining nine are used for training [47]. The subset assignment (validation/testing) is changed 10 times (10-fold) to compute the average performance. Using (1-4), the average classification accuracy (*CA*), specificity (Spec), sensitivity (Sen), and F-score (Fs) are calculated.

$$CA = \left(\frac{N_{correct}}{N_{total}}\right)\times 100\% \tag{1}$$

$$\text{Spec} = \left(\frac{TN}{TN+FP}\right)\times 100\% \tag{2}$$

$$\text{Sens} = \left(\frac{Tp}{Tp+FN}\right)\times 100\% \tag{3}$$

$$\text{F-score} = \left(2\times\frac{\Pr ec\times Sens}{\Pr ec + Sens}\right)\times 100\% \tag{4}$$

In equation (1), N_{total} represents the total number of feature vectors that require classification, while $N_{correct}$ denotes the number of correctly classified feature vectors. *TN* refers to true negatives, *TP* represents true positives, *FP* is false positives, and *FN* indicates false negatives. *Sensitivity* provides an indication of how well a model is able to identify individuals with the disease. In contrast, the *specificity* indicates whether the model can accurately identify individuals without the disease [48]. The *precision (pre)* metric, which is defined as counting the number of precise positive predictions produced.

$$\text{Prec} = \left(\frac{Tp}{Tp + FP}\right) \times 100\% \quad (5)$$

III. RESULTS

This section provides the classification results of the proposed methods for distinguishing MCI patients from healthy control group. In order to demonstrate the effectiveness of decomposing EEG signals on classification accuracy, results will be presented through two experiments:

Experiment I: classification results without signal decomposition.

Experiment II: classification results with the use of EMD for signal decomposition.

A. Experiment I Results (without using EMD)

In this experiment, the signal from each channel undergoes a bandpass filter ranging from 0.5 Hz to 32 Hz. Subsequently, the signals are divided into 3660 individual segments, each lasting 10 seconds (designated as M = 3660). MCI individuals contribute a total of 1740 segments, while HC participants contribute 1920 segments. From each segment, one feature (one element) is directly computed using one of the six proposed measures (energy, band power, log energy entropy, sure entropy, threshold entropy, or Shannon entropy). This process is repeated for all 19 channels, resulting in a feature vector comprised of 19 elements. Ultimately, the KNN classifier receives a 3660×19 feature matrix, derived from all the segments.

Table II contains a list of the six measures together with the corresponding classification accuracy, specificity, sensitivity, and F-score findings for each measure. Using 10-fold cross-validation, ten values are generated for every evaluation metric (for example accuracy). The mean performance and standard deviation (mean±st) of the ten values are computed. The table illustrates that the use of log energy entropy leads to the highest classification performance with 90.98% accuracy, 92.29% specificity, 89.54% sensitivity, and 90.44% F-score. On the other hand, the poorest performance is observed when Shannon entropy is employed. Table II also shows a clear difference between the values of the specificity and sensitivity in most measures, especially ThEn, SuEn, and T-ShEn.

B. Experiment II Results (with EMD)

This experiment applies pre-filtering and segmentation steps similar to Experiment I to obtain 3660 segments (M = 3660). But unlike Experiment I, EMD signal decomposition is applied before feature extraction. For each segment, a feature vector consisting of 114 elements is generated by aggregating data from 19 different channels using the adopted FE methods, as described in the FE section. Consequently, the KNN classifier receives a 3660×114 feature matrix.

Table III shows the KNN classification results using features extracted from IMF signals. 10-fold cross-validation is also used to compute the mean value of every evaluation metric for every FE method. As shown in the table, the best results are achieved with EMD+LogEn and EMD+SuEn FE methods, which have mean classification accuracy rates of 97.40% and 97.60%, respectively. The classification accuracy of other FE methods like EMD+ThEn and EMD+LBP is around 95%. Table III also shows that the EMD+Eng FE method performs the worst, with an accuracy of around 90%. The table also shows a small difference between the values of specificity and sensitivity for the entropy measures.

The above results were obtained using the KNN classifier. For additional analysis of the proposed FE methods, several other classification methods are examined. The classification accuracy rates of the RF, QDA, and SVM methods are shown in Fig. 3. KNN generates the best classification accuracy ratings

TABLE I. KNN CLASSIFICATION PERFORMANCE OF MCI VS. HC (**WITHOUT EMD**)

FE methods	**Accuracy (%)**	**Specificity (%)**	**Sensitivity (%)**	**F-score (%)**
	mean±st	**mean±st**	**mean±st**	**mean±st**
Eng	87.70±0.77	89.69±1.45	85.52±2.39	86.85±0.89
LBP	89.23±1.37	90.73±2.49	87.59±3.10	88.54±1.51
LogEn	90.98±1.61	92.29±3.18	89.54±1.67	90.44±1.57
ThEn	79.81±2.32	84.22±2.56	74.94±2.60	77.92±2.53
SuEn	88.14±2.22	90.16±2.97	85.92±2.29	87.33±2.30
TShEn	84.18±2.29	87.66±2.71	80.34±3.36	82.83±2.57

TABLE II. KNN CLASSIFICATION PERFORMANCE OF MCI VS. HC (**WITH EMD**)

FE methods	**Accuracy (%)**	**Specificity (%)**	**Sensitivity (%)**	**F-score (%)**
	mean±st	**mean±st**	**mean±st**	**mean±st**
EMD+Eng	90.30±0.89	93.18±1.56	87.13±1.54	89.52±0.96
EMD+LBP	95.00±0.66	97.03±1.04	92.76±1.39	94.63±0.72
EMD+LogEn	97.40±0.89	98.33±0.84	96.38±1.33	97.24±0.95
EMD+ThEn	95.33±0.69	95.83±1.43	94.77±2.02	95.07±0.75
EMD+SuEn	**97.60±0.92**	**98.59±0.85**	**96.49±1.42**	**97.44±0.98**
EMD+TShEn	97.16±0.63	98.33±0.98	95.86±1.04	96.98±0.67

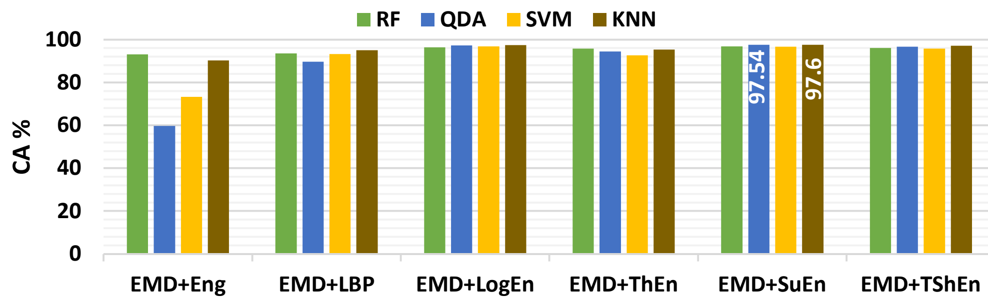

Figure 3. The classification accuracy of MCI vs. HC using different classifiers.

with most FE methods, as shown in the figure, whereas SVM achieves the poorest. Referring back to Fig. 3, results indicate that the highest accuracies of 97.60% and 97.54% are achieved when the features are extracted using EMD+SuEn and classified by KNN and QDA, respectively.

IV. DISSCUSSION

In this study, the problem of accurate identification of MCI from resting state EEG signals is addressed. We attempt to prove the effectiveness of decomposing EEG signals into sub-signals, with each having a different frequency component, before extracting the features. EMD has been employed to decompose EEG signals into a set of IMFs. In addition, several measures, including entropy, energy, and band power, have been investigated to compute the features from the decomposed signals. The extracted features are then classified using a number of classifiers that are evaluated using a public dataset. Table III shows the KNN classification performance of the proposed feature extraction methods in terms of classification accuracy, specificity, sensitivity, and F-score findings for each method. To demonstrate the effectiveness of the proposed signal decomposition method, the classification performance is re-evaluated without using EMD decomposition (results are in Table II).

Results show that when EMD is implemented, a significant improvement in classification performance is observed across all feature extraction measures. For example, in the case of the T-Shen measure, the mean classification accuracy increases from 84.18% to 97.16% when EMD is used. A similar improvement is observed for SuEn, where the mean classification accuracy increases from 88.14% to 97.60%. It can also be noted from Tables II and III that the improvement is not only limited to the classification accuracy. There is also a clear improvement in specificity, sensitivity, and F-score. Moreover, Table III demonstrates smaller standard deviation values compared to those observed in Table II. By comparing Tables II and III, there is also a clear improvement in the level of difference between the values of specificity and sensitivity, especially for entropy measures. To examine the potential of the proposed methods for MCI identification, the confusion matrices for nine methods are shown in Fig. 4. The figure shows that there are generally no considerable distinctions between sensitivity and specificity across most of the methods. For instance, in the first method (EMD+LogEn+QDA), 43 of the 1740 vectors that belong to MCI were identified as normal, but 59 of the 1920 vectors that belong to HC were identified as MCI.

These improvements in accuracy, specificity, and sensitivity may be due to the fact that when the signal is decomposed into different-frequency sub-signals, more information is obtained at the level of each frequency. In other words, when EMD decomposes an EEG signal into IMFs that represent different oscillatory components within the signal, each IMF captures a specific frequency or timescale present in the original signal. By decomposing the signal into these IMFs, EMD can uncover hidden patterns, transient behaviors, and variations at different frequencies, revealing more detailed information about the signal. This is in line with the study [20], which demonstrated that decomposing the EEG signals using DWT led to an improvement in the classification performance for diagnosing PD. The utilization of DWT decomposition facilitates the extraction and representation of signals in a multi-resolution or multi-frequency fashion, thereby generating a more comprehensive representation that encompasses additional information and details compared to the original signal [49]. DWT has been widely employed in numerous research studies to investigate various brain disorders [13, 17, 20].

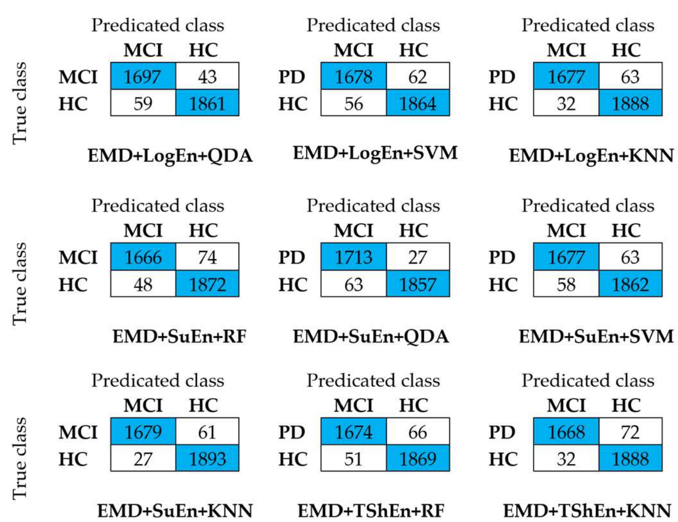

Figure 4: Confusion matrices for selected methods.

Table IV summarizes the existing state-of-the-art methods [24–33] for classifying MCI versus HC along with the corresponding classification scores. Studies [28, 33] investigated combining DWT with statistical measures. To the best of our knowledge, no prior study has examined the advantages of EMD for the

TABLE IV: COMPARISONS OF OUR RESULTS WITH THE RESULTS OF PREVIOUS STUDIES FOR MCI VS. HC CLASSIFICATION

Reference	FE methods	Classifier(s)	Dataset	CA (%)
M. Kashefpoor et al. [24, 2016]	Power, relative power, ratio power for different bands	Neurofuzzy+ KNN	11 MCI/ 16 HC	88.89%
M. Kashefpoor et al. [25, 2019]	Correlation-based Label Consistent K-SVD with spectral features		29 MCI/ 32 HC	88.9%
S. Hadiyoso et al. [26, 2019]	Power spectral features	KNN	11 MCI/ 16 HC	81.5%
Y.T. Hsiao et al. [27, 2021]	kernel Eigen-relative-power	SVM	24 MCI/ 27 HC	90.20%
Yin, Jiao et al. [28, 2019]	SWT+ statistical features	SVM	11 MCI/ 11 HC	96.94%
S. Siuly et al. [29, 2020]	Permutation entropy and auto-regressive	ELM	11 MCI/ 16 HC	98.78%
A. M. Alvi et al. [30, 2022]	--	LSTM	11 MCI/ 16 HC	96.41%
K. Lee et al. [31, 2022]	Several features using 10 measures	SVM	21 MCI/ 21 HC	86085%
R. A. Movahed et al. [32, 2022]	Spectral, functional connectivity, and nonlinear features	SVM	18 MCI/ 16 HC	99.40%
A. Said and H. Göker [33, 2023]	DWT leader + statistical features	AdaBoostM1	18 MCI/ 16 HC	93.50%
Present study	EMD+LogEn	QDA	(11 MCI/ 16 HC combined with 18 MCI/ 16 HC)	97.21%
	EMD+SuEn			97.54.%
	EMD+LogEn	KNN		97.40%
	EMD+SuEn			97.60%

identification of MCI. There are several studies investigating spectral, temporal, and non-linear measures as biomarkers for different brain disorders [14, 18, 40–42]. For MCI, most studies in Table IV also investigated those measures as protentional biomarkers for identifying MCI. The gap in Table IV is the lack of a study that extracts spectral, temporal, and non-linear features from decomposed signals. In the present study, various entropy measures, in addition to energy and band power, are explored. Unlike previous studies, we examined the computation of these measures not directly from EEG signals but from the decomposed IMF signals. This aims to derive efficient biomarkers for the identification of MCI. Table IV also includes our proposed methods and their corresponding results. The table shows that our proposed methodologies yield comparable results to previous studies, despite being evaluated using larger datasets. Moreover, in the present study, a single measure (energy, log energy entropy, sure entropy, threshold entropy, Shannon entropy, or band power) is employed for feature computation, in contrast to previous methods [29, 31, 32] that utilized multiple measures for the same purpose.

V. CONCLUSION AND FUTURE WORK

For classifying MCI versus health controls, this study provides efficient EMD-based methods where features are extracted from decomposed EEG signals (IMFs). Several measures, including entropy, energy, and band power, have been used to compute the features. The obtained features are then classified using a number of classifiers. Results show that a significant performance improvement is achieved when combining EMD with those measures (especially sure entropy (SuEn)). Results show that the proposed methods can attain comparable classification performance when compared to existing techniques in the literature. The combination of EMD and sure entropy yielded classification accuracy, specificity, and sensitivity, and F-score results of 97.60%, 98.59%, 96.49%, and 97.44%, respectively. In general, results show that decomposing the EEG signal before extracting the features improves the classification performance.

Future work includes the use of other decomposition methods, such as variational mode decomposition, to confirm the findings. The authors also plan to investigate more measures and use feature selection methods to improve the results.

REFERENCES

[1] World Health Organization, "Dementia," World Health Organization, Mar. 15, 2023.

[2] A. Burns and S. Iliffe, "Alzheimer's disease," BMJ, vol. 338, no. feb05 1, pp. b158–b158, Feb. 2009.

[3] M. Prince, E. Albanese, M. Guerchet, and A.M. Prina, "World Alzheimer Report Dementia and Risk Reduction: An Analysis of Protective and Modifiable Factors" 2014.

[4] Alzheimer's Association, "2015 Alzheimer's disease facts and figures," Alzheimer's & Dementia, vol. 11, no. 3, pp. 332–384, Mar. 2015.

[5] Alzheimer's Association, "Treatments and Research," Alzheimer's Disease and Dementia, 2019.

[6] Brooker D, La Fontaine J, Evans S, Bray J, Saad K. Public health guidance to facilitate timely diagnosis of dementia: ALzheimer's COoperative Valuation in Europe recommendations. Int J Geriatr Psychiatry. 2014 Jul;29(7):682-93.

[7] R. C. Petersen et al., "Memory and MRI-based hippocampal volumes in aging and AD," Neurology, vol. 54, no. 3, pp. 581–581, Feb. 2000.

[8] S. E. Rose et al., "Diffusion indices on magnetic resonance imaging and neuropsychological performance in amnestic mild cognitive impairment," Journal of Neurology, Neurosurgery, and Psychiatry, vol. 77, no. 10, pp. 1122–1128, Oct. 2006.

[9] L. Teng et al., "Predicting MCI progression with FDG-PET and cognitive scores: a longitudinal study," BMC Neurology, vol. 20, no. 1, Apr. 2020.

[10] H. Wu et al., "Computed Tomography Density and β-Amyloid Deposition of Intraorbital Optic Nerve May Assist in Diagnosing Mild Cognitive Impairment and Alzheimer's Disease: A 18F-Flutemetamol Positron Emission Tomography/Computed Tomography Study," Frontiers in Aging Neuroscience, vol. 14, p. 836568, 2022.

[11] Westman E, Muehlboeck JS, Simmons A. E. Westman, J-Sebastian. Muehlboeck, and A. Simmons, "Combining MRI and CSF measures for classification of Alzheimer's disease and prediction of mild cognitive

impairment conversion," NeuroImage, vol. 62, no. 1, pp. 229–238, Aug. 2012.
[12] M. Aljalal, S. Ibrahim, R. Djemal, and W. Ko, "Comprehensive review on brain-controlled mobile robots and robotic arms based on electroencephalography signals," Intelligent Service Robotics, vol. 13, no. 4, pp. 539–563, Jun. 2020.
[13] F. A. Alturki, K. AlSharabi, A. M. Abdurraqeeb, and M. Aljalal, "EEG Signal Analysis for Diagnosing Neurological Disorders Using Discrete Wavelet Transform and Intelligent Techniques," Sensors, vol. 20, no. 9, p. 2505, Apr. 2020.
[14] F. A. Alturki, M. Aljalal, A. M. Abdurraqeeb, K. Alsharabi, and A. A. Al-Shamma'a, "Common Spatial Pattern Technique With EEG Signals for Diagnosis of Autism and Epilepsy Disorders," IEEE Access, vol. 9, pp. 24334–24349, 2021.
[15] L. A. Moctezuma and M. Molinas, "EEG Channel-Selection Method for Epileptic-Seizure Classification Based on Multi-Objective Optimization," Frontiers in Neuroscience, vol. 14, Jun. 2020.
[16] J. Sheng et al., "A novel joint HCPMMP method for automatically classifying Alzheimer's and different stage MCI patients," vol. 365, pp. 210–221, Jun. 2019.
[17] K. AlSharabi, Y. Bin Salamah, A. M. Abdurraqeeb, M. Aljalal and F. A. Alturki, "EEG Signal Processing for Alzheimer's Disorders Using Discrete Wavelet Transform and Machine Learning Approaches," in IEEE Access, vol. 10, pp. 89781-89797, 2022.
[18] V. Jahmunah et al., "Automated detection of schizophrenia using nonlinear signal processing methods," Artificial Intelligence in Medicine, vol. 100, p. 101698, Sep. 2019.
[19] M. Aljalal, S. A. Aldosari, K. AlSharabi, A. M. Abdurraqeeb, and F. A. Alturki, "Parkinson's Disease Detection from Resting-State EEG Signals Using Common Spatial Pattern, Entropy, and Machine Learning Techniques," Diagnostics, vol. 12, no. 5, p. 1033, Apr. 2022.
[20] Majid Aljalal, S. A. Aldosari, M. Molinas, Khalil AlSharabi, and F. A. Alturki, "Detection of Parkinson's disease from EEG signals using discrete wavelet transform, different entropy measures, and machine learning techniques," vol. 12, no. 1, Dec. 2022.
[21] Y. Chen et al., "Automatic Sleep Stage Classification Based on Subthalamic Local Field Potentials," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 27, no. 2, pp. 118–128, Feb. 2019.
[22] L. A. Moctezuma, T. Abe, and M. Molinas, "Two-dimensional CNN-based distinction of human emotions from EEG channels selected by multi-objective evolutionary algorithm," Scientific Reports, vol. 12, no. 1, p. 3523, Mar. 2022.
[23] S. Khatun, B. I. Morshed, and G. M. Bidelman, "A Single-Channel EEG-Based Approach to Detect Mild Cognitive Impairment via Speech-Evoked Brain Responses," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 27, no. 5, pp. 1063–1070, May 2019.
[24] M. Kashefpoor, H. Rabbani, and M. Barekatain, "Automatic Diagnosis of Mild Cognitive Impairment Using Electroencephalogram Spectral Features," Journal of medical signals and sensors, vol. 6, no. 1, pp. 25–32, 2016.
[25] M. Kashefpoor, H. Rabbani, and M. Barekatain, "Supervised dictionary learning of EEG signals for mild cognitive impairment diagnosis," Biomedical Signal Processing and Control, vol. 53, p. 101559, Aug. 2019.
[26] S. Hadiyoso, C. La Febry Andira Rose Cynthia, M. Tati Latifah E. R., and H. Zakaria, "Early Detection of Mild Cognitive Impairment Using Quantitative Analysis of EEG Signals," IEEE Xplore, Sep. 01, 2019.
[27] Y.-T. Hsiao, C.-F. Tsai, C.-T. Wu, T.-T. Trinh, C.-Y. Lee, and Y.-H. Liu, "MCI Detection Using Kernel Eigen-Relative-Power Features of EEG Signals," Actuators, vol. 10, no. 7, p. 152, Jul. 2021.
[28] J. Yin, J. Cao, S. Siuly, and H. Wang, "An Integrated MCI Detection Framework Based on Spectral-temporal Analysis," International Journal of Automation and Computing, vol. 16, no. 6, pp. 786–799, Oct. 2019.
[29] 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.
[30] A. M. Alvi, S. Siuly, and H. Wang, "A Long Short-Term Memory Based Framework for Early Detection of Mild Cognitive Impairment From EEG Signals," IEEE Transactions on Emerging Topics in Computational Intelligence, pp. 1–14, 2022.
[31] K. Lee, K.-M. Choi, S. Park, Seung Hwan Lee, and C.-H. Im, "Selection of the optimal channel configuration for implementing wearable EEG devices for the diagnosis of mild cognitive impairment," Alzheimer's Research & Therapy, vol. 14, no. 1, Nov. 2022.
[32] R. A. Movahed and M. Rezaeian, "Automatic Diagnosis of Mild Cognitive Impairment Based on Spectral, Functional Connectivity, and Nonlinear EEG-Based Features," Computational and Mathematical Methods in Medicine, vol. 2022, pp. 1–17, Aug. 2022.
[33] A. Said and H. Göker, "Automatic Detection of Mild Cognitive Impairment from EEG Recordings Using Discrete Wavelet Transform Leader and Ensemble Learning Methods," DÜMF Mühendislik Dergisi, Mar. 2023.
[34] M. Aljalal, M. Molinas, S. A. Aldosari, K. AlSharabi, A. M. Abdurraqeeb, and F. A. Alturki, "Mild cognitive impairment detection with optimally selected EEG channels based on variational mode decomposition and supervised machine learning," Biomedical Signal Processing and Control, vol. 87, p. 105462, Jan. 2024.
[35] R. Cassani, M. Estarellas, R. San-Martin, F. J. Fraga, and T. H. Falk, "Systematic Review on Resting-State EEG for Alzheimer's Disease Diagnosis and Progression Assessment," Disease Markers, vol. 2018, pp. 1–26, Oct. 2018.
[36] "EEG Signals from Normal and MCI (Mild Cognitive Impairment) Cases," Available: https://misp.mui.ac.ir/en/eeg-data-0.
[37] Textbook of Alzheimer Disease and Other Dementias, American Psychiatric Publishing, Inc., 2009.
[38] M. Barekatain et al., "The relationship between regional brain volumes and the extent of coronary artery disease in mild cognitive impairment," J Res Med Sci, vol. 19, no. 8, pp. 739–45, 2014
[39] N. E. Huang et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, vol. 454, pp. 903–995, Mar. 1998.
[40] W. Bosl, A. Tierney, H. Tager-Flusberg, and C. Nelson, "EEG complexity as a biomarker for autism spectrum disorder risk," BMC Medicine, vol. 9, no. 1, Feb. 2011.
[41] N. Kannathal, M. L. Choo, U. R. Acharya, and P. K. Sadasivan, "Entropies for detection of epilepsy in EEG," Computer Methods and Programs in Biomedicine, vol. 80, no. 3, pp. 187–194, Dec. 2005.
[42] R. Catherine Joy, S. Thomas George, A. Albert Rajan, and M. S. P. Subathra, "Detection of ADHD From EEG Signals Using Different Entropy Measures and ANN," Clinical EEG and Neuroscience, p. 155005942110367, Aug. 2021.
[43] Breiman, L, "Random Forests," Machine Learning 45, 5–32, 2001.
[44] R. O. Duda, P. E. Hart, and D. G. Stork, Pattern classification.
[45] C. J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," Data Mining and Knowledge Discovery, vol. 2, no. 2, pp. 121–167, 1998, doi: https://doi.org/10.1023/a:1009715923555.
[46] K. Q. Weinberger and L. K. Saul, "Distance Metric Learning for Large Margin Nearest Neighbor Classification," Journal of Machine Learning Research, vol. 10, no. 9, pp. 207–244, 2009.
[47] P. Refaeilzadeh, L. Tang, and H. Liu, "Cross-Validation," Encyclopedia of Database Systems, pp. 532–538, 2009.
[48] A. Swift, R. Heale, and A. Twycross, "What are sensitivity and specificity?," Evidence Based Nursing, vol. 23, Nov. 2019.
[49] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 11, no. 7, pp. 674–693, Jul. 1989.