



## Research article

# A tensor decomposition scheme for EEG-based diagnosis of mild cognitive impairment

Alireza Faghfour<sup>a</sup>, Vahid Shalchyan<sup>a,\*</sup>, Hamza Ghazanfar Toor<sup>b</sup>, Imran Amjad<sup>b,c</sup>,  
Imran Khan Niazi<sup>c,d,e</sup>

<sup>a</sup> Neuroscience & Neuroengineering Research Lab, Biomedical Engineering Department, School of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran

<sup>b</sup> Riphah International University, Islamabad, Pakistan

<sup>c</sup> Centre for Chiropractic Research, New Zealand College of Chiropractic, Auckland, New Zealand

<sup>d</sup> Faculty of Health & Environmental Sciences, Health & Rehabilitation Research Institute, AUT University, Auckland, New Zealand

<sup>e</sup> Department of Health Science and Technology, Aalborg University, Aalborg, Denmark

## ARTICLE INFO

**Keywords:**

Mild cognitive impairment (MCI)  
Tensor decomposition  
Electroencephalogram (EEG)  
Alzheimer's disease  
Parallel factor analysis (PARAFAC)

## ABSTRACT

Mild Cognitive Impairment (MCI) is the primary stage of acute Alzheimer's disease, and early detection is crucial for the person and those around him. It is difficult to recognize since this mild stage does not have clear clinical signs, and its symptoms are between normal aging and severe dementia. Here, we propose a tensor decomposition-based scheme for automatically diagnosing MCI using Electroencephalogram (EEG) signals. A new projection is proposed, which preserves the spatial information of the electrodes to construct a data tensor. Then, using parallel factor analysis (PARAFAC) tensor decomposition, the features are extracted, and a support vector machine (SVM) is used to discriminate MCI from normal subjects. The proposed scheme was tested on two different datasets. The results showed that the tensor-based method outperformed conventional methods in diagnosing MCI with an average classification accuracy of 93.96% and 78.65% for the first and second datasets, respectively. Therefore, it seems that maintaining the spatial topology of the signals plays a vital role in the processing of EEG signals.

## 1. Introduction

Alzheimer's disease involves the weakening or destruction of neurons in different brain parts, often beginning in the mild to the acute stages. Since the nature of this disease is progressive, in the early stages, some mild cognitive symptoms appear; in many cases, these symptoms are perceived by the patient and those around him as a normal state related to aging and symptoms are hard to distinguish from the other types of dementia [1]. The progression of Alzheimer's disease can be divided into four general categories: The first stage is known as Mild Cognitive Impairment (MCI), which is the mediator between the normal decline of cognitive abilities due to old age and acute dementia [1]. The second and third stages are related to mild Alzheimer's disease and moderate Alzheimer's disease, which are characterized by increasing cognitive deficits as well as problems and disorders related to daily activities. The last stage is acute Alzheimer's, in which patients lose their cognitive skills and memory over time, do not feel for their surroundings, and with symptoms of mobility impairment, are not able to move and swallow food and the like. Therefore, people at this stage need a

\* Corresponding author.

E-mail address: [shalchyan@iust.ac.ir](mailto:shalchyan@iust.ac.ir) (V. Shalchyan).

<https://doi.org/10.1016/j.heliyon.2024.e26365>

Received 6 November 2023; Received in revised form 8 February 2024; Accepted 12 February 2024

Available online 15 February 2024

2405-8440/Â© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

nurse to perform their essential activities [1,2]. There is currently no definitive cure for Alzheimer's disease, and the only way to deal with it is to diagnose it early and prevent it from progressing to more acute stages. It should be noted that this is not an easy task because the symptoms of the disease, especially in the early stages, are indistinguishable from other diseases and normal symptoms. The prevalence and spread of MCI in the elderly population (50–95 years) have been reported to be between 5.8% and 18.5% [3]. Moreover, the incidence of patients with MCI with acute dementia is about 50%, a very high rate. Therefore, MCI as a transient stage between normal aging and dementia is known as one of the critical goals for treating dementia [4–7]. Diagnosis of MCI is one of the main challenges in preventing the progression of the disease because this disorder has no clear symptoms, and the abilities of individuals in daily activities are not affected [8].

Clinical and neurophysiological analyses of memory impairment have led to discovering Alzheimer's subtypes, disease staging, and prognosis. Despite new criteria for early disease prediction, including biomarkers derived from cerebrospinal fluid or hippocampal volume analysis, neurophysiological tests are still the mainstay of prognosis and segregation of disease stages [9]. Of course, these tests are very time-consuming, dependent on the individual, and require skilled test personnel [10]. On the other hand, various techniques, including Electroencephalogram (EEG), Magnetoencephalogram (MEG), and functional Magnetic Resonance Imaging (fMRI), are used to study the mechanisms of mild cognitive impairment. Due to the slow and silent changes of MCI, patients with it should be investigated for a long time to develop a system that can diagnose MCI. Such a system must necessarily be non-invasive, accessible, and inexpensive. Therefore, an EEG-based system is a good option [11]. EEG involves recording the electrical potentials generated by the brain through a set of electrodes located on the scalp.

Previous studies in the diagnosis of MCI based on EEG have shown that in people with MCI, the power of conventional frequency bands has been changed (e.g., increased gamma-band power and decreased alpha band power) [12–14], and slowing occurs in rhythms [12,15–17]. Also, because in these patients, connectivity between different parts of the brain is impaired [10,11,18,19], some studies on reducing the statistical complexity of EEG signals [7,9,20,21] and anxiety in the synchronous activity of different parts of the brain of people with mild cognitive impairment compared to healthy older people [1,2,22,23]. In recent years, studies have been conducted in this field. In 2019, Kashefpour et al. used the dictionary learning approach to classify MCI and healthy individuals [24]. In 2019, Loracitano et al. used a time-frequency information approach and continuous wavelet transform to diagnose Alzheimer's patients with mild cognitive impairment and healthy individuals [25]. Furthermore, in 2020, Duan Feng et al. used the functional connectivity features of brain topology to diagnose patients with MCI for use in their deep learning model [8]. In 2012, Latchumane et al. used the EEG spectral information and the tensor decomposition approach to analyze Alzheimer's disease. However, in their work, all EEG channels were placed together in a 1-dimensional vector; thus, the spatial information about the channels has not been preserved [26].

Since a dense number of electrodes are often used to record EEG signals, their spatial arrangement can contain helpful information to distinguish between patients with MCI and healthy individuals. Additionally, according to the results of previous studies, spectral features are the most popular for MCI detection. This study combined spatial and spectral features and considered their interactions using a tensor decomposition approach. Tensor decompositions originated with Hitchcock in 1927 [27], and the idea of a multiway model is attributed to Cattell in 1944 [28]. These concepts received scant attention until the work of Tucker in the 1960s [29], all of which appeared in psychometrics literature. Appellof and Davidson [30] are generally credited as being the first to use tensor decomposition in chemometrics, and tensors have since become extremely popular in that field. In the last ten years, interest in tensor decompositions has expanded to other fields, including signal processing, numerical linear algebra, computer vision, numerical analysis, data mining, graph analysis, neuroscience, and more [31].

In this work, by using a proper projection and a spatial arrangement of channels as well as spectral information of each electrode, the tensor dimensions of each person can be formed. In addition to maintaining spatial information about the channels, the advantages of the tensor-based approach can be utilized. Considering the obtained results and comparing the conventional methods with the method of using tensor decomposition, it seems that tensor concepts, as they have achieved good results in other applications, have a high potential in the analysis of EEG time series related to MCI. For this purpose, we used the tensor decomposition method of parallel factor analysis (PARAFAC) to analyze the multi-dimensional array and reveal the effects and characteristics of MCI.

The rest of this manuscript is organized as follows: Section 2 introduces the database used in this research and the set of methods used. Section 3 deals with the results and discussion in section 4, and finally, this article is summarized in section 5.

## 2. Materials and methods

An overview of the proposed method is shown in Figure (1). This method includes the sections of preprocessing, data tensor construction, feature extraction, and classification, and each detail is described in the following sections of the manuscript.

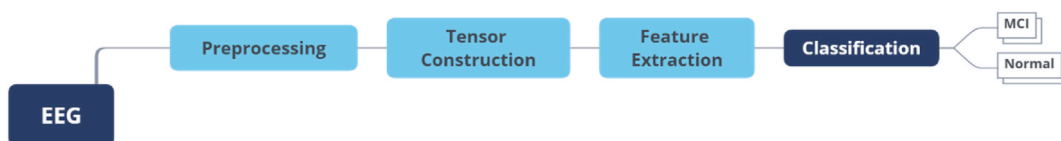


Fig. 1. Block diagram of the proposed method.

## 2.1. Participants

To evaluate the effectiveness and robustness of the proposed scheme, we used two datasets that include EEG signals from different individuals, the first of which was recorded according to the 10–20 standard and the second of which was recorded by an Emotiv Epoc device.

### 2.1.1. Dataset I

In this study, the dataset in Ref. [24] has been used, which was recorded in Noor Hospital of Isfahan Province. The dataset includes 61 people over the age of 55, 29 of them had MCI, and 32 of them were healthy. All participants had primary or higher education. According to the psychiatric interview using Patterson criteria, all subjects were scored, and those with a Mini-mental state examination (MMSE) score between 21 and 26 were classified as MCI patients. Furthermore, people who scored above 26 are considered healthy [24]. In addition, to ensure the labeling of individuals, the Neuropsychiatry unit cognitive assessment tool (NUCOG) was used as the MCI diagnosis confirmation test, with scores between 75 and 86.5 being considered as MCI patients. The EEG signals were recorded in the early morning for each person with their eyes closed and in a quiet room for 30 min. Moreover, a 19-channel EB-Neuro system whose electrodes were arranged according to a 10–20 standard was used to record the signal. Figure (2) shows how the channels are arranged in a 10–20 standard placement.

The sampling rate in signal recording was 256 Hz, and the impedance between the electrode and the scalp was kept at less than 5 K $\Omega$ . For more descriptions of the recording details, please see Ref. [24].

### 2.1.2. Dataset II

The second dataset includes EEG signals with mild, normal, and severe cognitive impairment as well as healthy individuals [32]. Throughout the signal recording process, subjects were asked to sit in a chair and refrain from any hand, finger, or face contraction for 2 min during the recording. After 30 s, upon hearing a distinct sound, the subjects have to open their eyes, and therefore, during the 2 min of signal recording, 1 min corresponds to the closed state and 1 min to the open state. The MMSE clinical criterion was used to divide individuals into different stages of the disease in order to diagnose the severity of the disease. For this purpose, similar to dataset I, individuals with MMSE scores greater than 26 were classified as healthy individuals and individuals with a score between 21 and 26 were identified as MCI [32]. The number of healthy people in this dataset was 21 people with an average age of 61 years, and the number of MCI people was equal to 22 people with an average age of 64 years. This dataset has not been used for classifying and recognition of healthy individuals and MCI, and it has been used for other applications. Since some of the signals were much shorter than 2 min, two persons were removed from the data related to healthy individuals, and one person was excluded from the data related to MCI individuals from the signal processing process. Therefore, the number of healthy people present in the next stages will be equal to 20 people, and the number of people with mild cognitive impairment will be 20 people. In this dataset, to record the EEG signals, a 14-channel Emotiv Epoc was used, and the sampling frequency was 128 Hz. Figure (3) shows how the channels are arranged in the Emotiv Epoc device.

## 2.2. Preprocessing

In order to prepare the data, for Dataset I, the EEG data underwent a two-step preprocessing procedure. Firstly, the data was normalized to ensure consistent scaling across all channels. Thus the mean of all channels is zero, and their variance is equal to one.

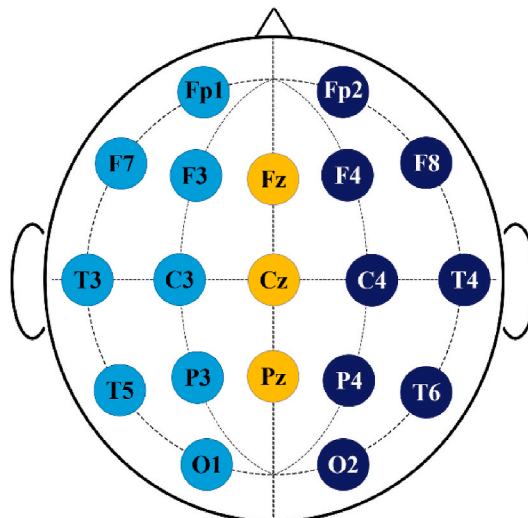


Fig. 2. Name and arrangement of electrodes on the scalp according to a 10–20 standard.

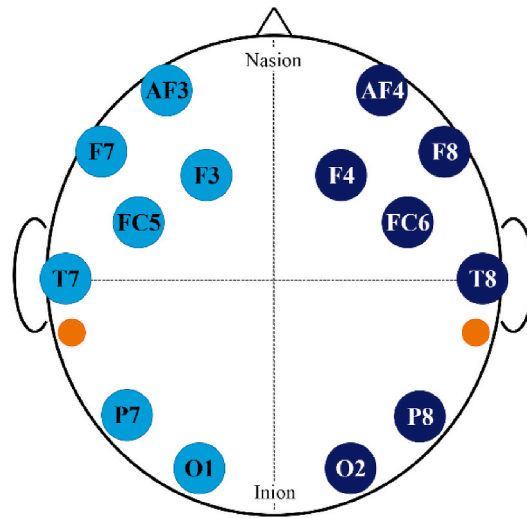


Fig. 3. Emotiv electrodes. The Emotiv Epoc headset is composed of 14 different electrodes.

Secondly, a 4th order Butterworth band-stop filter was applied to remove power line 50Hz interference. Also, during the data processing stages, the data are divided into short time patches with different lengths, including 0.5, 1, 1.5, 5, 10, 15, and 30 min to identify the most appropriate length of patches for diagnosis in the analysis. To divide the data into training and test groups in cross-validation, three distinct groups are randomly formed, and 41 subjects (20 MCI and 21 healthy) for training process and 20 subjects (9 MCI and 11 healthy) for test, are randomly selected each time. In each group, the training and testing process is performed separately, and to determine the diagnostic label of the test subjects, voting on the labels of the timepieces of each test subject is used. To increase the stability of the answers, this process is performed on three random groups and the average of the results will be reported as the test result. In the case of Dataset II, the sampling rate in the signal recording is 128 Hz, and to begin with, Similar to Dataset I, the EEG data for Dataset II also underwent a two-step preprocessing procedure. Firstly, the data was normalized to achieve consistent scaling across all channels. Secondly, a 4th order Butterworth band-stop filter was utilized to eliminate power line interference. Subsequently, the signals are divided into short time patches with different lengths, including 2, 4, 6, 10, 12, 20, 24, 40, 60, and 120 s. The difference in the time division of the two datasets is the time of data acquisition, which was 2 min in Dataset II and 30 min in Dataset I. Then, similar to dataset I, to divide the dataset II into training and test groups, three different groups are randomly formed and each time, 67% of the people are randomly selected for the training process and 33% for testing. So, 27 subjects (13 MCI and 14 healthy) for training process and 13 subjects (7 MCI and 6 healthy) for test, are randomly selected.

2.3. Construction of Data tensor

For this purpose, first using the azimuthal equidistant projection method (also called polar projection) [33], the arrangement of the

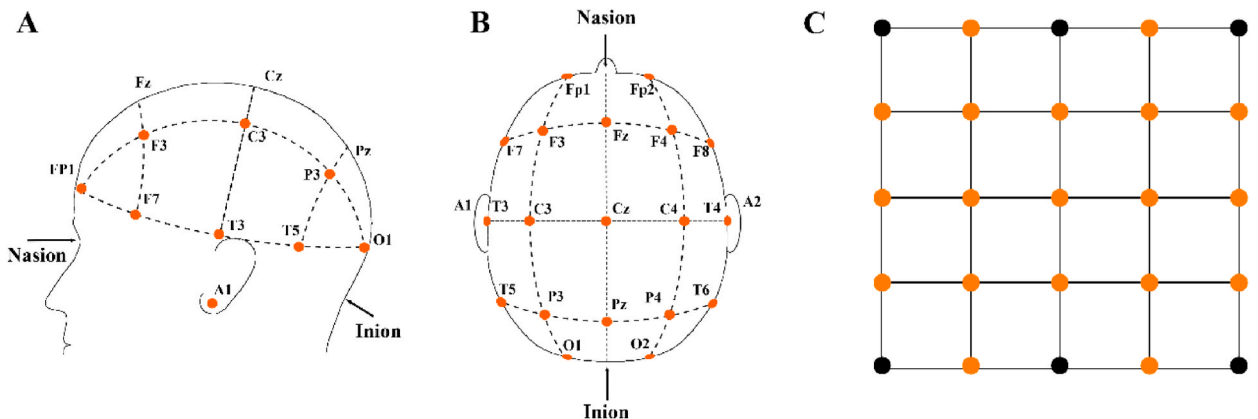
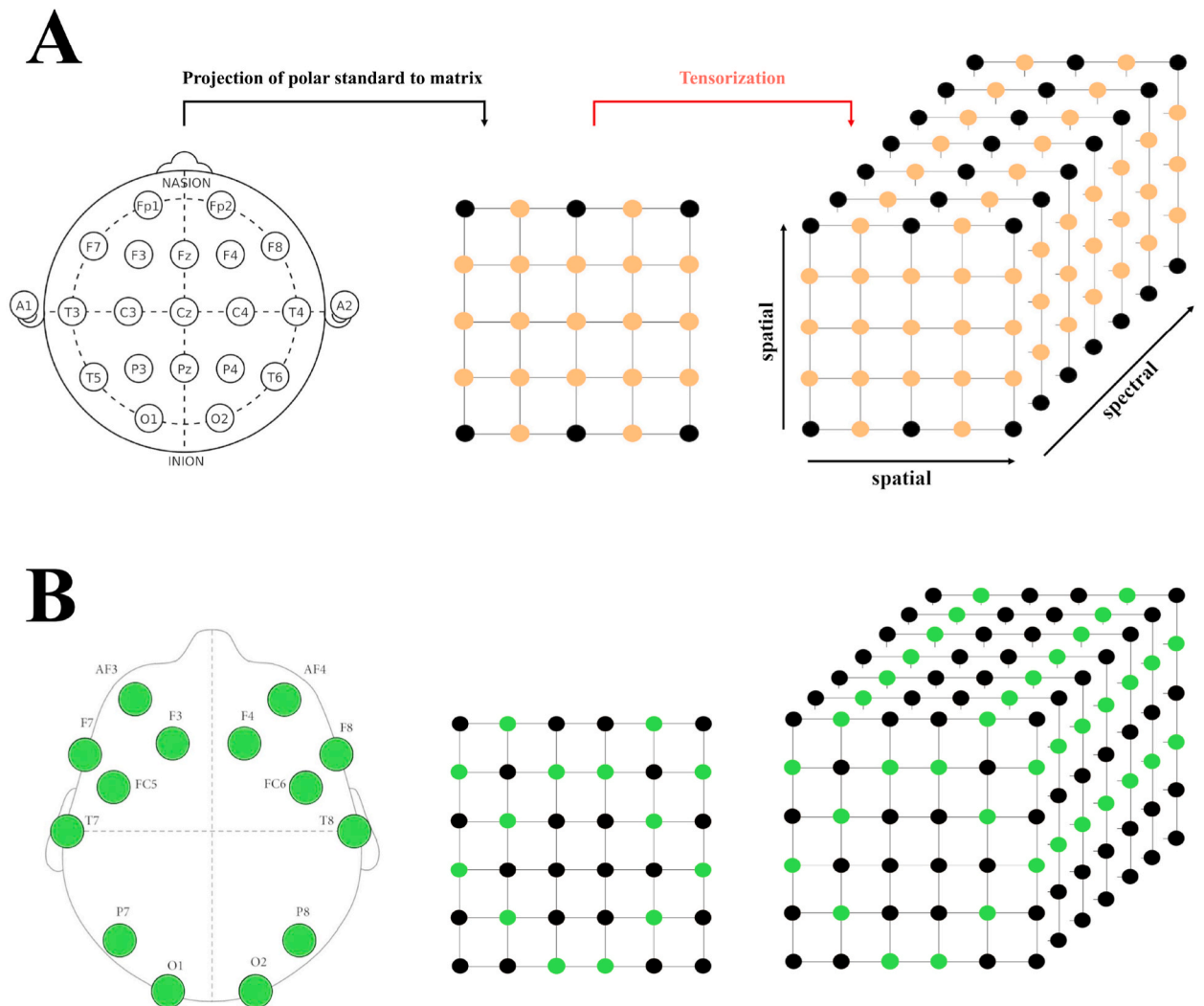


Fig. 4. Converting the coordinates and polar arrangement of the electrodes on the head to a matrix. A: The side view shows the position of the signal recording channels in polar mode. B: It is related to the front view of the location of the electrodes. C: It is an estimate of the position of the electrodes in the matrix state, which can be defined in Cartesian coordinates.

electrodes on the head is converted to the Cartesian mode. Then, for example, by placing the spectral values of each electrode in the desired matrix, a three-dimensional array (tensor) is formed. Tensors are multi-dimensional arrays. A third-order tensor has three faces. The first-order tensor is a vector, the second-order tensor is the same as the matrix, and tensors with dimensions three and higher are called high-order tensors. Moreover, Figure (4 A-C) demonstrates the transformation of electrode coordinates and their polar arrangement on the head into a matrix.

EEG signals are usually displayed as vectors or matrices by conventional methods such as time-series, spectral, and matrix analysis to facilitate data processing. However, the EEG signal inherently has more dimensions, such as time, frequency, space, and subject. Therefore, the need to represent the dataset with higher dimensions is felt more. According to figure (4), the data matrix is mapped to the five-by-five dimensions of the 2D plane. Unlike the conventional case where the data are stacked one after the other, in the proposed method, the data from each channel are placed in orange dots in order to preserve their spatial information. Therefore, a 3-dimensional tensor will be formed for each subject, and then, by stacking these three-dimensional tensors for all subjects, a four-dimensional tensor consisting of an arrangement of spatial locations of channels, spectral information, and subjects will be formed. In this case, the spatial and spectral information of everyone can be used at once, and there is no need to consider them independently. This idea will preserve the electrodes' spatial arrangement and allow individuals' spatial-frequency information to be used simultaneously for classification.

In dataset II, considering that the number of electrodes is equal to 14 channels, in order to construct the data tensor, according to the spatial arrangement of electrodes in the Emotiv Epoc device, first, the topographic map of the electrode arrangement related to the Emotiv Epoc is modeled as a six-by-six matrix. As shown in Figure (5 - B), the six-by-six data matrix is mapped to the Cartesian plane,



**Fig. 5.** Construction of a data tensor for a specific subject for dataset I (A) and dataset II (B). In the figure on the left, the standard location related to the recording device can be seen. In the middle figure, a polar projection is made to the Cartesian plane, and the desired matrix is formed. In the figure on the right, a three-dimensional tensor is formed for a specific subject by placing each of the matrices in a row.

features are located in green dots, and their spatial information is preserved. Figure (5) shows how to form a three-dimensional tensor for a specific subject for dataset I (Fig. 5 - A) and dataset II (Fig. 5 - B).

#### 2.4. Feature extraction and tensor decomposition

To learn the classification model, useful features must be extracted that can distinguish between healthy and MCI groups properly. Therefore, one of the most important parts of machine learning problems is the feature extraction part. In connection with studies related to the diagnosis of MCI, the use of signal spectral properties and the increase and decrease of frequency band power, as well as the reduction of irregularity and statistical complexity of signals due to brain disorders, have been proven. Accordingly, in the following, we will try to use other types of features to form tensors in addition to spectral features.

##### 2.4.1. Spectral features

In many studies related to the processing of EEG signals, and especially in studies related to MCI, the spectral features of each channel have useful information and have led to acceptable performance. In previous studies, the use of frequency band power in diagnosing MCI has been considered. It showed that the power in low-frequency bands increases, and in high-frequency bands decreases [34]. Classification using power ratios in frequency bands has also been considered [35,36]. In this study, two approaches have been used to implement spectral properties. The first approach is related to the traditional use of the power of conventional frequency bands in EEG signals, which include delta, theta, alpha, beta, and gamma bands. To do this, we filter the signals using a 4th degree Butterworth band pass filter and perform a Fast Fourier Transform (FFT) on these signals, and then the abs of signal power is calculated. In a more comprehensive classification, 19 spectral features related to these bands can be used, which are presented in Table (1). Furthermore, in Table (1), concepts such as relative powers and the variables in the equations in R1-R3 are included. These concepts generally pertain to the division of power values across different frequency bands, and for further elaboration, readers can refer to Ref. [37].

Another approach used in this study is inspired by the filter bank problem, which divides the signal into, for example, 40, 60, or 80 equal frequency bands and uses the power of each micro bands as a feature.

##### 2.4.2. Entropy features and statistical complexity

Researchers have also favored entropy as a measure of information theory in diagnosing patients with MCI, especially through EEG and MEG signals. Many studies have acknowledged the reduction in entropy and complexity of information in patients. Because in patients with Alzheimer's disease, as well as its mild levels, including MCI, the connections of different parts of the brain are damaged and sometimes even lost, so in the degree of information disorder in distinctive EEG signal channels, changes occur compared to normal; Furthermore, information complexity has been reported in patients less than healthy individuals.

**2.4.2.1. Shannon's entropy.** Entropy in information theory is a numerical measure of the amount of information or randomness of a random variable. To be more precise, the entropy of a random variable is the average value (expected value) of the information obtained from observing it (in other words, the higher the entropy of a random variable, the more ambiguity we have about that random variable; Therefore, by observing the definite result of that random variable, more information is obtained. So, the higher the entropy of a random variable, the more information we can get from its definitive observation). Furthermore, the entropy of an information source is the low expectation of the best compression rate without data loss for that source.

The information obtained from observing an event is defined as equal to its negative logarithm probability function; naturally,

**Table 1**  
19 Spectral features associated with conventional frequency bands [37].

Row	Name	Description
1	delta	Delta band power (0.5–3.5 Hz)
2	r-delta	The relative power of the delta band
3	theta	Theta band power (3.5–7.5 Hz)
4	r-theta	The relative power of the theta band
5	alpha1	Alpha 1 band power (7.5–9.5 Hz)
6	r-alpha1	Relative power of alpha band 1
7	alpha2	Alpha 2 band power (9.5–12.5 Hz)
8	r-alpha2	Relative power of alpha band 2
9	beta1	Beta 1 band power (12.5–17.5 Hz)
10	r-beta1	Relative power of beta band 1
11	beta2	Beta 2 band power (17.5–25 Hz)
12	r-beta2	Relative power of beta band 2
13	gamma	Gamma band power (25–40 Hz)
14	r-gamma	The relative power of the gamma band
15	total	Total EEG signal band power (0.5–40 Hz)
16	R1	$R_1 = \theta/\alpha_1 + \alpha_2 + \beta_1$
17	R2	$R_2 = (\delta + \theta)/\alpha_1 + \alpha_2 + \beta_1 + \beta_2$
18	R3	$R_3 = \theta/\alpha_1 + \alpha_2$
19	PAF	Peak Alpha Frequency



there is an expectation from any function suitable for measuring the amount of information in an observation, including that the information from observation is non-negative. The information obtained from observing a definite event (i.e., with a probability of one) is zero, and most importantly, the information obtained from two independent observations is equal to the sum of the information obtained from observing each of them. It can be shown that just a function that satisfies the above three properties is the negative of the logarithm probability function. The amount of information with different bases of the logarithm differs only by a constant coefficient. The most common logarithm basis in information computation is two, which calculates information in units of bit or Shannon. In science and engineering in general, entropy is a measure of the degree of ambiguity. In 1948, Claude Shannon in his paper introduced Shannon’s entropy and became the founder of information theory [38].

Assume that  $X$  is the random variable with the values of  $X_1, X_2, \dots, X_M$  with the probabilities in order  $P_1, P_2, \dots, P_M$ . In this case, entropy is defined as follows:

$$H(P_1, P_2, \dots, P_M) = - \sum_{i=1}^M P_i \log P_i. \tag{1}$$

Equation (1) expresses the entropy in this context.

If the base of the logarithm is 2, the unit entropy will be bits, and generally, the highest entropy for a random variable occurs in a uniform distribution, and the lowest entropy occurs in a distribution with a definite event.

**2.4.2.2. Multi-scale entropy.** Multi-Scale entropy is generally a measure of the complexity and order of a signal or time series, as well as a measure of the rate at which information is produced. In more studies, different methods have been developed to calculate entropy. For example, the Kolenogrof-Sinai entropy measure was introduced in 1983 by Grasberger [39], and a modified version was introduced in 1985 by Eckmann & Rauler [40]. Following research on how to calculate information entropy, in 1991, a measure called the approximate entropy was developed by Pincus to find a specific pattern of time series [41]. The high bias of this method, as well as the strong dependence of this method on the signal length, were the main drawbacks of this method, which is a major challenge, especially in vital signals, which are usually short and noisy. Therefore, in 2000, a method called sample entropy was developed to eliminate the defects of the previous version [42]. This method, in addition to being less dependent on the signal length, also has fewer mathematical calculations and, therefore, has been considered by researchers. In general, the sample entropy is negatively defined as the mean of the natural logarithm of the conditional probability. Thus, when two sequences at the same point  $m$  are the same, with the tolerance value  $r$ , they are the same for the next point, provided that the self-comparison is not included in statistical calculations [42]. In other words, this type of entropy is defined as a logarithmic difference between the two probabilities of having a vector  $v$  with tolerance  $r$  in the dimension  $m$  and the existence of this vector with the same tolerance in the range  $m + 1$ . Therefore, when this value is less, it indicates more similarity in the time series. The Multi-Scale Entropy (MSE) criterion uses the same sample theory and is a new method of complexity calculation that focuses on determining the information expressed by signals at different time scales. MSE analysis is based on calculating the sample entropy of several coarse-grained sequences that show the system’s dynamics at different time scales [43]. To construct the coarse-grained time series based on the scale factor, the main time series is divided into non-overlapping time windows, each with a specified length, and then the average value of each of the values inside each window is calculated. So, we have Equation (2):

$$\{y^\epsilon(j)\} = \frac{1}{\epsilon} \sum_{i=(j-1)\epsilon+1}^{j\epsilon} x(i), 1 \leq j \leq \frac{N}{\epsilon}. \tag{2}$$

After the coarse-grained time series is obtained, its sample entropy criterion is calculated.

**2.4.2.3. Lempel-Ziv complexity criterion.** In 1976, Abraham Lempel and Jacob Ziv designed an algorithm that dealt with the complexity of binary time series. In many texts related to Alzheimer’s disease, the Lempel-Ziv criterion, which actually examines the degree of randomness of signals, has been used to distinguish patients from control subjects [17,20,21,44]. To calculate the LZ complexity criterion, the signal  $\{x(n)\}$  must be converted to a finite (usually binary) sequence. Using a  $T_d$  threshold value, the original signal samples are converted to 0 and 1, which we have Equation (3):

$$\{s(i)\} = s(1), s(2), \dots, s(N); \begin{cases} s(i) = 0, x(i) < T_d \\ s(i) = 1, x(i) \geq T_d \end{cases}. \tag{3}$$

The sequence  $\{x(n)\}$  is then scanned from left to right, and the complexity counter  $c(N)$  increases by one unit each time a sequence change is observed. Furthermore, to eliminate the dependence of complexity on signal length, the obtained values must be normalized [44]. In general, the upper limit  $c(N)$  is equal to  $\frac{N}{\log_\alpha N}$ . The value of  $\alpha$  (base of the logarithm) will be the number of variations of the states, which in the binary state is equal to 2. So, we have Equation (4):

$$\lim_{N \rightarrow \infty} c(N) = b(N) \equiv \frac{N}{\log_\alpha N}. \tag{4}$$

Therefore, the value of  $c(N)$  can be normalized by division into  $b(N)$ . Finally, The formula is as follows:

$$C(N) = \frac{c(N)}{b(N)}. \quad (5)$$

Equation (5) represents the LZ complexity criterion, where higher values indicate increased complexity.

### 2.4.3. Tensor decomposition

Given that today EEG signal data, especially due to the use of dense multi-channel recordings, most data have large dimensions and big volumes, another task of the feature extraction stage is to reduce the size and summary of the data. Multi-way array analysis is a method that can extract the interactions between different dimensions of tensors, such as spatial, spectral, subjects, etc., as well as summarize them into separate components. These components are considered basic components, and a matrix of these components is called the base factor [31]. The issue of tensor decomposition was raised by Hitchcock [27] in 1927 and seventeen years later by Cattell [28] in 1944; he introduced multi-dimensional models. One of the most important tensor decomposition methods is the PARAFAC or Canonical Polyadic Decomposition (CPD) method, which was developed in 1970 by Harshman [45]. The following is also named as PARAFAC tensor decomposition method.

**2.4.3.1. PARAFAC tensor decomposition.** PARAFAC tensor decomposition actually decomposes an N-dimensional tensor into the sum of rank-1 tensors. For a hypothetical three-dimensional tensor  $\underline{X}$ , the R-component PARAFAC decomposition will be as Equation (6) [46]:

$$\begin{aligned} \underline{X} &= \mathbf{a}_1 \circ \mathbf{b}_1 \circ \mathbf{c}_1 + \mathbf{a}_2 \circ \mathbf{b}_2 \circ \mathbf{c}_2 + \dots + \mathbf{a}_R \circ \mathbf{b}_R \circ \mathbf{c}_R + \underline{\mathbf{E}} \\ &\approx \mathbf{a}_1 \circ \mathbf{b}_1 \circ \mathbf{c}_1 + \mathbf{a}_2 \circ \mathbf{b}_2 \circ \mathbf{c}_2 + \dots + \mathbf{a}_R \circ \mathbf{b}_R \circ \mathbf{c}_R \\ &\approx \underline{\mathbf{X}}_1 + \underline{\mathbf{X}}_2 + \dots + \underline{\mathbf{X}}_R. \end{aligned} \quad (6)$$

In Equation (6), the symbol "o" is the symbol related to the outer product of vectors, and all tensors  $\underline{\mathbf{X}}_1$  up to  $\underline{\mathbf{X}}_R$  are rank one. The sum of these rank-1 tensors will actually be an estimate of the original tensor. The first components  $\mathbf{a}_1$ ,  $\mathbf{b}_1$ , and  $\mathbf{c}_1$  are associated with one another, and their outer product produces rank-one tensor  $\underline{\mathbf{X}}_1$ . The second components  $\mathbf{a}_2$ ,  $\mathbf{b}_2$ , and  $\mathbf{c}_2$  are associated with one another, and their outer product generates rank-one tensor  $\underline{\mathbf{X}}_2$  and so on. Therefore, PARAFAC decomposition is the sum of rank-one tensors plus error tensor. Figure (6) shows how PARAFAC decomposes a three-dimensional array [46].

In general, to decompose N-dimensional tensor using PARAFAC methods, we have  $\underline{\mathbf{X}} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ ; we will have Equation (7):

$$\underline{\mathbf{X}} = \sum_{r=1}^R \mathbf{u}_r^{(1)} \circ \mathbf{u}_r^{(2)} \circ \dots \circ \mathbf{u}_r^{(N)} + \underline{\mathbf{E}} = \sum_{r=1}^R \underline{\mathbf{X}}_r + \underline{\mathbf{E}} = \widehat{\underline{\mathbf{X}}} + \underline{\mathbf{E}} \approx \widehat{\underline{\mathbf{X}}}. \quad (7)$$

Which we have in the above relation:

$$\underline{\mathbf{E}} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}, \widehat{\underline{\mathbf{X}}} \text{ is an estimate of the tensor } \underline{\mathbf{X}}, r = 1, 2, \dots, R, \underline{\mathbf{X}}_r = \mathbf{u}_r^{(1)} \circ \mathbf{u}_r^{(2)} \circ \dots \circ \mathbf{u}_r^{(N)}.$$

Also, after applying the PARAFAC tensor decomposition technique to the N-dimensional data tensor, the component matrices (loading factors) for the n-th dimension are obtained as shown in Equation (8):

$$\mathbf{U}^{(N)} = [\mathbf{u}_1^{(n)}, \mathbf{u}_2^{(n)}, \dots, \mathbf{u}_R^{(n)}] \in \mathbb{R}^{I_n \times R}. \quad (8)$$

Equation (7) can also be written in the form of a matrix-tensor product, which is shown in Equation (9) as follows:

$$\underline{\mathbf{X}} = \mathbf{I} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \dots \times_N \mathbf{U}^{(N)} + \underline{\mathbf{E}} = \widehat{\underline{\mathbf{X}}} + \underline{\mathbf{E}}. \quad (9)$$

In Equation (9),  $\mathbf{I}$  is the identity tensor that is actually a diagonal tensor, and all the elements on its main diagonal are equal to one [46]. " $\times_1, \times_2, \dots, \times_N$ " is the N-mode product. The n-mode (matrix) product of a tensor  $\underline{\mathbf{X}} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$  with a matrix  $\mathbf{U} \in \mathbb{R}^{J \times I_n}$  is denoted by  $\underline{\mathbf{X}} \times_n \mathbf{U}$  and is of size  $I_1 \times \dots \times I_{n-1} \times J \times I_{n+1} \times \dots \times I_N$ . Elementwise, we have Equation (10):

$$(\underline{\mathbf{X}} \times_n \mathbf{U})_{i_1 \dots i_{n-1} j i_{n+1} \dots i_N} = \sum_{i_n=1}^{I_n} x_{i_1 i_2 \dots i_N} u_{i_n j}. \quad (10)$$

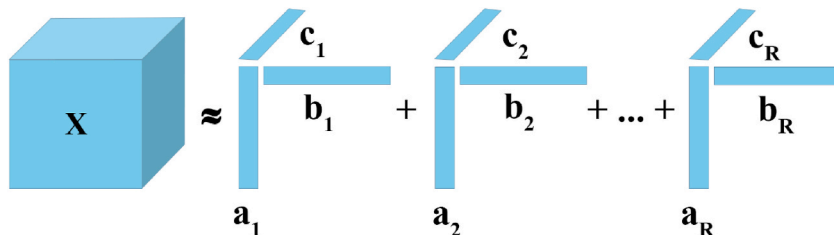


Fig. 6. PARAFAC decomposition of a 3-dimensional array.



It should be noted that for tensor decomposition, depending on the nature of each dimension, restrictive constraints can be applied to some or even all dimensions of the tensor. For example, if the data is non-negative, the non-negative clause can be used. This type of restriction has been used in many cases [47].

## 2.5. Classification

Two-class SVM has been used to classify patients with MCI and healthy individuals [48]. This method distinguishes between the two classes by creating a hyperplane in the multi-dimensional feature space. To achieve this, the Radial Basis Function (RBF) kernel was chosen due to its effectiveness in capturing complex and nonlinear relationships within the data. In this study, the RBF kernel with a kernel scale of 4 was utilized.

## 3. Results

Given that the main question of this study is related to the effect of using spatial information of electrodes with a tensor analysis approach in improving the classification results of patients with mild cognitive impairment and healthy individuals, it is necessary to compare the process with the conventional approach of previous research and without considering the concepts of tensor analysis and finally compare them. Therefore, in the following, the suitable frequency segmentation will be determined first. Afterward, using a variety of time patches, the finest time patch is determined. Next, by determining the best result of classification accuracy in the normal case, the proposed approach based on tensor analysis is applied, and using the time patch and suitable frequency bands, a data tensor will be formed and finally; with tensor analysis, the classification accuracy of the proposed method will be presented.

### 3.1. Select the finest frequency segmentation

In order to determine the suitable frequency band segmentation in feature extraction, first, the conventional method of using frequency band power without tensor analysis is applied and the classification accuracy for different features is reported in Figure (7). For this purpose, the signal of each channel is filtered using a 4th degree Butterworth band pass filter and then the absolute power of the desired frequency band is extracted. To extract the power of 60 frequency micro bands, the frequency band of the signal (0.5–40 Hz) is divided into 60 equal parts, and the power is calculated for each part. For 40 and 80-frequency micro bands, filtering and feature extraction was done in the same way. Furthermore, for the conventional bands and 19 spectral features, according to Table (1), feature extraction has been performed. Figure (7) shows the classification accuracy of the results related to the use of 40 frequency bands, 60 frequency bands, 80 frequency bands, seven traditional frequency bands and also 19 spectral features in conventional bands. The best result was related to the use of the 60 frequency micro band feature and the average accuracy was 89.54%. After that, the use of 80 frequency bands is in second place. Finally, the weakest accuracy is the use of seven traditional frequency bands.

Evaluation metric, including the average classification accuracy, sensitivity and specificity of each method, are shown in Table (2) and according to it, the best classification accuracy is related to 60 frequency bands. Therefore, the power feature of 60 frequency bands has been selected in both datasets to continue the research.

### 3.2. Select the suitable time patch

For dataset I, given that the signals were 30 min long, in order to select the optimal time patch, divide each signal into 0.5, 1, 1.5, 5,

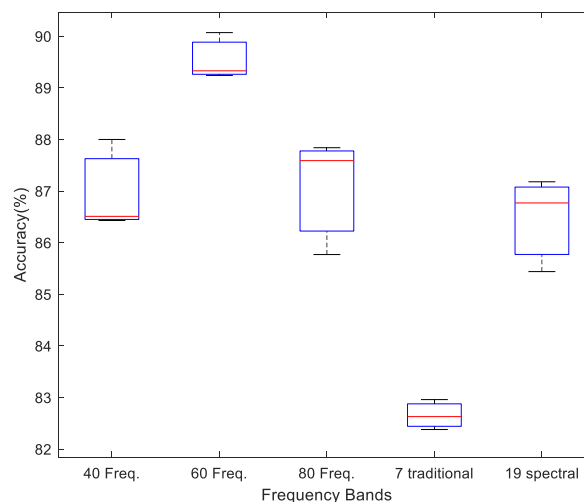


Fig. 7. Comparison of classification accuracy for power of different frequency bands (Dataset I).

**Table 2**

Evaluation metric for each feature for accuracy for power of different frequency bands.

Features/evaluation metric	Accuracy (%)	Sensitivity (%)	Specificity (%)
40 Frequency power	86.98	88.73	87.14
60 Frequency power	89.54	90.18	88.29
80 Frequency power	87.06	89.77	85.75
7 traditional bands power	82.65	85.10	80.55
19 spectral features	86.46	86.37	84.53

10, 15, and 30 min time patches, and for each case, the power feature in the 60 extracted frequency micro bands, the classification accuracy in the conventional state (without using the concept of our tensor-based scheme) was calculated, which can be seen in [Table \(3\)](#). The left column of the table corresponds to the different time patches, and the right column corresponds to the percentage of classification accuracy corresponding to each of them. According to the results, the best case was related to the time patch of 1.5 min and equal to 89.54%, which is the same division to continue the steps.

Furthermore, for the second dataset, since the recording time is different from the first dataset, in a similar way, the 2-min signal of each person at short intervals, including 2, 4, 6, 10, 12, 20, 24, 40, 60 And 120 s are divided. To determine the finest time patch, the classification accuracy for the normal state (without tensor analysis) with 60 frequency band features (similar to the first dataset) was calculated for different time patches. According to [Table \(4\)](#), for dataset II, the best time patch is 12 s, and its classification accuracy is 64.12%.

### 3.3. Our tensor-based scheme and comparison of different features

According to the results of the previous sections, the finest time patch was related to 1.5 min, and the best frequency spectrum segmentation was related to distances equal to 60 frequency bands. To continue the route, first divide the signals into short time intervals of 1.5 min, and therefore, each of the 30-min signals will be divided into 20 short signal patch.

In order to tensor decomposition of signals, a training tensor must be formed first. At this stage, for the signals of 67% of people (40 people) who were randomly selected for the training process, different features related to different frequency band power, Lempel-Ziv complexity, Shannon's entropy, and multi-scale sample entropy are stored in the training tensor while retaining spatial information. In the next step, to form the final features, the PARAFAC tensor decomposition was performed for the training tensor, and as a result; four matrices were obtained. Then, using the component matrix (the fourth matrix that belongs to subjects), the SVM was trained. For the model evaluation phase, component vector obtained from production of test tensor and pseudo-inverse of the Khatri-rao production of load matrices (first, second, and third matrices) obtained from the tensor decomposition of the training data. The results of accurate classification of control subjects and patients with MCI are obtained according to [Figure \(8\)](#).

Furthermore, according to [Table \(5\)](#), for dataset I, the best case is related to using the power feature of 60 frequency bands and the average classification accuracy is 93.96%, but according to the diagram of the third box in [Figure \(8\)](#), this criterion in one of the cases with an accuracy of 96.52% has also arrived.

[Table \(6\)](#) also shows the dimensions corresponding to each of the above methods in the training and test phase. In this table, the value of the variable related to the number of components (R) in the tensor analysis is selected using the classification accuracy assessment and in accordance with the best accuracy in each of the methods.

In relation to the second dataset, by forming the training data tensor, it will include the power of different frequency bands with equal segments. In the subsequent step, to extract the final features, PARAFAC tensor decomposition was performed for the training tensor, and as a result; four matrices were obtained. In the next step, support vector machine was learned using the component matrix. For the validation phase, the test tensor was inverted on the operating matrices obtained from the tensor analysis of the training data, and thus the component matrix for the test tensor was obtained. The results of classification accuracy were obtained for three different categories and after averaging, the classification accuracy is computed. In order to examine other features, the tensor decomposition process is used for the state of features related to the power of 60 equal frequency segments, the power of traditional bands (delta, theta, alpha 1, alpha 2, beta 1, beta two and gamma bands) and also, 19 spectral features listed in [Table \(1\)](#) were examined in dataset II to construct the tensor. [Table \(7\)](#) summarizes the results of the second dataset. The highest classification accuracy, according to the table below, is related to the power features of 60 frequency bands.

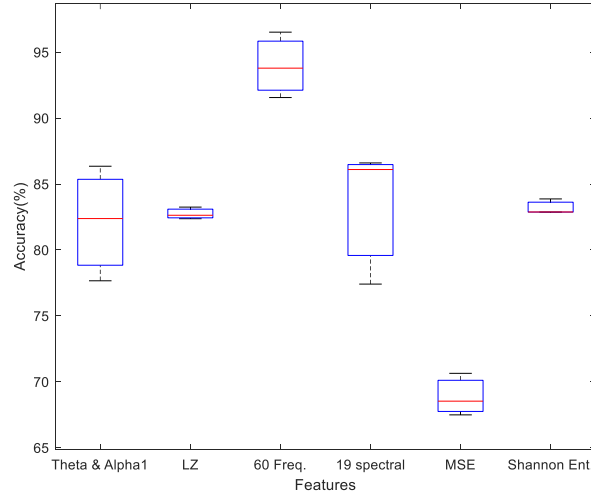
**Table 3**

Accuracy of classification in different time patches (Dataset I).

Time patch (minutes)	Classification accuracy (%)
0.5	87.34
1	89.16
1.5	89.54
5	84.78
10	78.23
15	67.37
30	53.12

**Table 4**  
Accuracy of classification in different time patches (Dataset II).

Time patch (seconds)	Classification accuracy (%)
2	56.38
4	58.35
6	59.16
10	63.76
12	64.12
20	62.86
24	61.45
40	57.64
60	41.18
120	43.24



**Fig. 8.** Classification accuracy for different methods and features (Dataset I).

**Table 5**  
Evaluation metric for each feature for the tensor-based mode (Dataset I).

Features/evaluation metric	Accuracy (%)	Sensitivity (%)	Specificity (%)
Band power of theta and alpha1	82.13	85.78	78.89
19 Spectral features	83.37	89.70	82.41
Power of 60 frequency bands	93.96	96.07	96.98
Lempel-Ziv	82.75	82.31	84.62
Shannon's entropy	83.21	84.80	80.90
Multi-scale sample entropy	68.89	70.86	68.27

**Table 6**  
Feature space of each method for the 90-s signal (Dataset I).

Features/Dimension	Train Tensor	R-component	Train Vector	Test Tensor	Test Vector
Band power of theta and alpha1	$5 \times 5 \times 2 \times 820$	50	$820 \times 50$	$5 \times 5 \times 2 \times 400$	$400 \times 50$
19 Spectral features	$5 \times 5 \times 19 \times 820$	60	$820 \times 60$	$5 \times 5 \times 19 \times 400$	$400 \times 60$
Power of 60 frequency bands	$5 \times 5 \times 60 \times 820$	65	$820 \times 65$	$5 \times 5 \times 60 \times 400$	$400 \times 65$
Lempel-Ziv	$5 \times 5 \times 7 \times 820$	45	$820 \times 45$	$5 \times 5 \times 7 \times 400$	$400 \times 45$
Shannon's entropy	$5 \times 5 \times 7 \times 820$	85	$820 \times 85$	$5 \times 5 \times 7 \times 400$	$400 \times 85$
Multi-scale sample entropy	$5 \times 5 \times 7 \times 820$	85	$820 \times 85$	$5 \times 5 \times 7 \times 400$	$400 \times 85$

### 3.4. Comparison of tensor decomposition results and traditional approach

According to the results, to compare the use of our tensor-based scheme and traditional approaches, the classification accuracy criteria for each of these cases are compared in Table (8). For this purpose, the average classification accuracy of each method has been

**Table 7**  
Feature space and accuracy of different features (Dataset II).

Features/Dimension & Accuracy	Train Tensor	R-component	Test Tensor	Accuracy (%)
19 Spectral features	$6 \times 6 \times 19 \times 270$	65	$6 \times 6 \times 19 \times 130$	52.24
Power of 60 frequency bands	$6 \times 6 \times 60 \times 270$	75	$6 \times 6 \times 60 \times 130$	78.65
Power of 7 traditional bands	$6 \times 6 \times 7 \times 270$	65	$6 \times 6 \times 7 \times 130$	69.72

**Table 8**  
Classification accuracy between traditional and proposed methods based on tensor decomposition.

Evaluation method	Accuracy (%)	
Traditional method	Dataset I	89.54
	Dataset II	64.12
Our Tensor-based Method	Dataset I	93.96
	Dataset II	78.65

used. According to the table below, the proposed tensor-based method had better results than the traditional approach and has improved the accuracy of distinguishing MCI patients from control subjects in both datasets. Also, according to Tables (6) and (7), the number of components is actually smaller than the number of subjects in the datasets.

#### 4. Discussion

In this research, our primary inquiry revolved around the question of whether considering the spatial location of electrodes as opposed to the conventional approach leads to an improvement in problem-solving. To investigate this, we employed tensor decomposition as one of the tools to delve into this matter. The ultimate findings on dataset I demonstrated a relative enhancement in performance, validating our hypothesis. For further validation, we applied this new approach to another dataset, and the same observation was noted in dataset II.

On the subject of diagnosing MCI, many studies have been performed so far, none of which used the concept of tensor and tensor decomposition. In this study, in addition to using the concept of tensor, in order to form a tensor, contrary to conventional methods, we tried to preserve the spatial information about the channels so that more useful and valuable information could be extracted. According to the results obtained in this study, which are shown in Table (8), the use of this innovation and tensor decomposition method led to the accuracy of classification in classification patients with MCI, and Healthy people recovered well; In this study, without using the tensor decomposition method, the average classification accuracy was 89.54%, which reached 93.96% using our proposed method in dataset I. To be more sure, we also tested our method on the second dataset, in which case the classification accuracy increased from 64.12% to 78.65% compared to the traditional method and therefore, our initial theory was proved, and it was shown that maintaining the spatial position of the electrodes in the processing stage contains more important information.

Dataset I that was used in this study belonged to the study [24] that the best classification accuracy reported in the result was related to the channels of the left temporal zone and equal to 80%; however, by using the same dataset and the equivalent method of cross-evaluation, by using the method proposed in our study, which preserves spatial and frequency information simultaneously and uses tensor decomposition to extract the feature, the amount of classification accuracy has greatly improved.

In the case of MCI, tensor decomposition has not been used to date; however, this concept in many applications has increased and improved performance as well as the accuracy of classification in machine learning problems. For example, in the application of Brain-Computer Interfaces (BCIs), Rob Zink et al., using tensor decomposition, caused the dependence of the system on the calibration phase to be significantly reduced for subjects, and this process became unnecessary [49]. Jingyangli et al. also used the adaptive method based on tensor decomposition to improve the classification of motor imagery in EEG signals [50]. Lucianus Spiro et al. also used the concept of tensor in the application of functional brain connectivity, which led to a better estimate of brain connections [51]. Tensor decomposition in compression-related problems also improved performance, and in a study by Dowels et al., tensor decomposition was achieved using near-lossless compression [52]. In studies related to the extraction of different stages of sleep and also the study of children's development, tensor concepts and particularly tensor decomposition has increased performance and improved the outcome compared to normal [53,54]. Thus, the use of the advantages of tensor decomposition and tensor-related concepts in many previous studies has promised to improve system performance in other applications, including the diagnosis of diseases, especially the diagnosis of MCI using tensor decomposition. The problem is quite evident in the results of our study and somehow confirms the good results of other studies related to the concept of tensors. Of course, this is not the end of the matter, and due to the limitations of this study, it is possible to generalize and expand it along the path, using different methods of tensor decomposition to extract features or also can implement a variety of other methods to form a tensor. Furthermore, in the dataset used in this study, 19 channels were used to record the signal, and it seems that in the next studies, better results can be achieved by increasing the number and density on the electrodes. Following the path and in further studies, the efficiency of this method can be measured for other paradigms of EEG signal recording from MCI patients, such as signal recording with open eyes. It also seems that good results can be obtained by fusion of signal and image modalities, such as EEG signal and fMRI image, and using tensor decomposition.

## 5. Conclusions

The diagnosis of MCI, which is the first step in the acute conditions of Alzheimer's disease and dementia, significantly reduces the progression of the disease to severe stages. Therefore, due to the mild and ambiguous clinical symptoms, it is difficult and expensive to diagnose MCI. The use of cheap and available EEG signals has made it possible to improve this by using signal processing techniques. In this study, it was found that in addition to temporal and spectral features, maintaining the spatial position of the electrodes also has valuable information. Accordingly, a method based on tensor analysis with the ability to maintain the location of the EEG electrodes in the formation of the tensor was developed to diagnose MCI. The results showed a good advantage over conventional and classical methods.

## Data Availability Statement

The authors do not have permission to share data.

## CRediT authorship contribution statement

**Alireza Faghfouri:** Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation. **Vahid Shalchyan:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Data curation, Conceptualization. **Hamza Ghazanfar Toor:** Software, Investigation, Data curation. **Imran Amjad:** Software, Investigation, Data curation. **Imran Khan Niazi:** Writing – review & editing, Resources, Funding acquisition, Data curation, Conceptualization.

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Imran Khan Niazi reports administrative support, article publishing charges, equipment, drugs, or supplies, and statistical analysis were provided by The Hamblin Trust, New Zealand. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgment

This work was supported by a grant from “The Hamblin Trust, New Zealand”, and also received co-funding from the Hamblin chiropractic research fund trust and through donations to the Centre for Chiropractic Research Supporters Program at the New Zealand College of Chiropractic.

## References

- [1] N. Handayani, Coherence and phase synchrony analyses of EEG signals in Mild Cognitive Impairment (MCI): a study of functional brain connectivity, *Pol. J. Med. Phys. Eng.* 24 (1) (2018) 1–9.
- [2] J. Dauwels, EEG synchrony analysis for early diagnosis of Alzheimer's disease: a study with several synchrony measures and EEG data sets, in: 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2009.
- [3] A.P. Collie, Maruff, The neuropsychology of preclinical Alzheimer's disease and mild cognitive impairment, *Neurosci. Biobehav. Rev.* 24 (3) (2000) 365–374.
- [4] B.A. Ally, Preserved frontal memorial processing for pictures in patients with mild cognitive impairment, *Neuropsychologia* 47 (10) (2009) 2044–2055.
- [5] E. Choi, Clinical characteristics of patients in a dementia prevention center, *J. Neurol. Sci.* 283 (1) (2009) 291–292.
- [6] S. Villeneuve, Episodic memory deficits in vascular and non vascular mild cognitive impairment, *J. Neurol. Sci.* 283 (1) (2009) 291.
- [7] M. Ahmadlou, Complexity of functional connectivity networks in mild cognitive impairment subjects during a working memory task, *Clin. Neurophysiol.* 125 (4) (2014) 694–702.
- [8] F. Duan, et al., Topological network analysis of early Alzheimer's disease based on resting-state EEG, *IEEE Trans. Neural Syst. Rehabil. Eng.* 28 (10) (2020) 2164–2172.
- [9] F. Vecchio, et al., Cortical connectivity and memory performance in cognitive decline: a study via graph theory from EEG data, *Neuroscience* 316 (2016) 143–150.
- [10] C.S. Musaeus, M.S. Nielsen, P. Høgh, Altered low-frequency EEG connectivity in mild cognitive impairment as a sign of clinical progression, *J. Alzheim. Dis.* 68 (3) (2019) 947–960.
- [11] N. Mammone, et al., Permutation disalignment index as an indirect, EEG-based, measure of brain connectivity in MCI and AD patients, *Int. J. Neural Syst.* 27 (5) (2017) 1750020.
- [12] J. McBride, et al., Discrimination of mild cognitive impairment and Alzheimer's disease using transfer entropy measures of scalp EEG, *Journal of healthcare engineering* 6 (1) (2015) 55–70.
- [13] K. Van der Hiele, EEG correlates in the spectrum of cognitive decline, *Clin. Neurophysiol.* 118 (9) (2007) 1931–1939.
- [14] C. Herrmann, T. Demiralp, Human EEG gamma oscillations in neuropsychiatric disorders, *Clin. Neurophysiol.* 116 (12) (2005) 2719–2733.
- [15] J. Dauwels, Slowing and loss of complexity in Alzheimer's EEG: two sides of the same coin? *Int. J. Alzheimer's Dis.* (2011) 2011.
- [16] C. Latchoumane, et al., Multiway analysis of Alzheimer's disease: classification based on space-frequency characteristics of EEG time series, in: *Proceedings of the World Congress on Engineering*, 2008.
- [17] R. Hornero, Nonlinear analysis of electroencephalogram and magnetoencephalogram recordings in patients with Alzheimer's disease, *Phil. Trans. Math. Phys. Eng. Sci.* 367 (2009) 317–336, 1887.
- [18] Z.-y. Jiang, Study on EEG power and coherence in patients with mild cognitive impairment during working memory task, *J. Zhejiang Univ. - Sci. B* 6 (12) (2005) 1213–1219.
- [19] G. Adler, S. Brassen, A. Jajcevic, EEG coherence in Alzheimer's dementia, *Journal of neural transmission* 110 (9) (2003) 1051–1058.
- [20] A. Fernandez, et al., Complexity analysis of spontaneous brain activity in Alzheimer disease and mild cognitive impairment: an MEG study, *Alzheimer Disease & Associated Disorders* 24 (2) (2010) 182–189.

- [21] B. Zhu, et al., Analysis of EEG complexity in patients with mild cognitive impairment, *Journal of Neurological Disorders* 5 (4) (2017).
- [22] C. Babiloni, et al., Directionality of EEG synchronization in Alzheimer's disease subjects, *Neurobiol. Aging* 30 (1) (2009) 93–102.
- [23] E. Gallego-Jutglà, et al., Diagnosis of Alzheimer's disease from EEG by means of synchrony measures in optimized frequency bands, in: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2012.
- [24] M. Kashefpoor, H. Rabbani, M. Barekatin, Supervised dictionary learning of EEG signals for mild cognitive impairment diagnosis, *Biomed. Signal Process Control* 53 (2019) 101559.
- [25] C. Ieracitano, et al., A novel multi-modal machine learning based approach for automatic classification of EEG recordings in dementia, *Neural Network*. 123 (2020) 176–190.
- [26] C.-F.V. Latchoumane, et al., Multiway array decomposition analysis of EEGs in Alzheimer's disease, *J. Neurosci. Methods* 207 (1) (2012) 41–50.
- [27] F.L. Hitchcock, The expression of a tensor or a polyadic as a sum of products, *J. Math. Phys.* 6 (1–4) (1927) 164–189.
- [28] R.B. Cattell, "Parallel proportional profiles" and other principles for determining the choice of factors by rotation, *Psychometrika* 9 (4) (1944) 267–283.
- [29] L.R. Tucker, Implications of factor analysis of three-way matrices for measurement of change, *Problems in measuring change* 15 (122–137) (1963) 3.
- [30] C.J. Appellhof, E.R. Davidson, Strategies for analyzing data from video fluorometric monitoring of liquid chromatographic effluents, *Anal. Chem.* 53 (13) (1981) 2053–2056.
- [31] T.G. Kolda, B.W. Bader, Tensor decompositions and applications, *SIAM Rev.* 51 (3) (2009) 455–500.
- [32] N. Khan, Xbox 360 Kinect cognitive games improve slowness, complexity of EEG, and cognitive functions in subjects with mild cognitive impairment: a randomized control trial, *Game. Health J.* (2019).
- [33] W. Qiao, X. Bi, Ternary-task convolutional bidirectional neural turing machine for assessment of EEG-based cognitive workload, *Biomed. Signal Process Control* 57 (2020) 101745.
- [34] H. Berendse, et al., Magnetoencephalographic analysis of cortical activity in Alzheimer's disease: a pilot study, *Clin. Neurophysiol.* 111 (4) (2000) 604–612.
- [35] D. Osipova, et al., Altered generation of spontaneous oscillations in Alzheimer's disease, *Neuroimage* 27 (4) (2005) 835–841.
- [36] J. Poza, et al., Analysis of spontaneous MEG activity in patients with Alzheimer's disease using spectral entropies, in: 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2007.
- [37] M. Kashefpoor, H. Rabbani, M. Barekatin, Automatic diagnosis of mild cognitive impairment using electroencephalogram spectral features, *Journal of medical signals and sensors* 6 (1) (2016) 25.
- [38] C.E. Shannon, A mathematical theory of communication, *The Bell system technical journal* 27 (3) (1948) 379–423.
- [39] P. Grassberger, I. Procaccia, Estimation of the Kolmogorov entropy from a chaotic signal, *Phys. Rev.* 28 (4) (1983) 2591.
- [40] J.P. Eckmann, D. Ruelle, Ergodic Theory of Chaos and Strange Attractors, *The theory of chaotic attractors*, 1985, pp. 273–312.
- [41] S.M. Pincus, Approximate entropy as a measure of system complexity, *Proc. Natl. Acad. Sci. USA* 88 (6) (1991) 2297–2301.
- [42] J.S. Richman, J.R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, *Am. J. Physiol. Heart Circ. Physiol.* (2000).
- [43] M. Costa, A.L. Goldberger, C.-K. Peng, Multiscale entropy analysis of biological signals, *Phys. Rev.* 71 (2) (2005) 021906.
- [44] X.-S. Zhang, R.J. Roy, E.W. Jensen, EEG complexity as a measure of depth of anesthesia for patients, *IEEE Trans. Biomed. Eng.* 48 (12) (2001) 1424–1433.
- [45] R.A. Harshman, Foundations of the PARAFAC Procedure: Models and Conditions for an "Explanatory" Multimodal Factor Analysis, 1970.
- [46] F. Cong, et al., Tensor decomposition of EEG signals: a brief review, *J. Neurosci. Methods* 248 (2015) 59–69.
- [47] A. Cichocki, Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation, John Wiley & Sons, 2009.
- [48] M.A. Hearst, et al., Support vector machines, *IEEE Intell. Syst. Their Appl.* 13 (4) (1998) 18–28.
- [49] R. Zink, et al., Tensor-based classification of an auditory mobile BCI without a subject-specific calibration phase, *J. Neural. Eng.* 13 (2) (2016) 026005.
- [50] X. Li, et al., Adaptation of motor imagery EEG classification model based on tensor decomposition, *J. Neural. Eng.* 11 (5) (2014) 056020.
- [51] L. Spyrou, M. Parra, J. Escudero, Complex tensor factorization with PARAFAC2 for the estimation of brain connectivity from the EEG, *IEEE Trans. Neural Syst. Rehabil. Eng.* 27 (1) (2018) 1–12.
- [52] J. Dauwels, et al., Near-lossless multichannel EEG compression based on matrix and tensor decompositions, *IEEE journal of biomedical and health informatics* 17 (3) (2012) 708–714.
- [53] E. Kinney-Lang, et al., Tensor-driven extraction of developmental features from varying paediatric EEG datasets, *J. Neural. Eng.* 15 (4) (2018) 046024.
- [54] S. Kouchaki, et al., Tensor based singular spectrum analysis for automatic scoring of sleep EEG, *IEEE Trans. Neural Syst. Rehabil. Eng.* 23 (1) (2014) 1–9.