

# EEG Signal Analysis for Mild Alzheimer's Disease Diagnosis by Means of Spectral- and Complexity-Based Features and Machine Learning Techniques



**Nilesh Kulkarni**

**Abstract** Alzheimer's disease (AD) is one of the common and fastest growing neurological diseases in the modern society. Biomarker techniques for diagnosis of Alzheimer's disease and its progression in early stage are key issues for development. Electroencephalogram is one of the powerful techniques which can be used for screening of Alzheimer's disease and dementia in early stage. The objective of this paper is to analyze the EEG signal by means of spectral and complexity features to serve EEG as a biomarker for Alzheimer's diagnosis. The research is carried on experimental database obtained from hospital. EEG relative power, spectral entropy, spectral flux, and spectral centroid are analyzed, compared, and classified for separating the data between two groups by means of support vector machine (SVM) classifier and K-nearest neighbor (KNN) classifier. The obtained results indicate severity observed in AD patients reflected in EEG signals which can be treated as benchmark for Alzheimer's diagnosis.

**Keywords** Alzheimer's disease · Electroencephalogram · Classifier · Machine learning

## 1 Introduction

Alzheimer's disease (AD) is basically characterized by impaired state of memory leading toward severe Alzheimer's in which medications are not helpful for saving the life of an individual. Statistical studies have reported that around 6–25% of mild cognitive impairment (MCI) patients are transformed toward Alzheimer's and 0.2–4% from normal to mild Alzheimer's every year [1, 2]. Due to this, early diagnosis of Alzheimer's disease and its progression is key challenges in Alzheimer's diagnosis. It is expected that strength of individuals with Alzheimer's disease is about to increase in the future. To search a computationally efficient technique for early detection of

---

N. Kulkarni (✉)  
STES's Sou. Venutai Chavan Polytechnic, Pune, India  
e-mail: [nileshkulkarni992@gmail.com](mailto:nileshkulkarni992@gmail.com)

patients who are in progress toward Alzheimer's disease but do not show any clinical symptoms of Alzheimer's is an important as well as new challenge. Although neuroimaging techniques are helpful in screening of Alzheimer's disease, they are much expensive and time consuming. EEG on other side is cheap and a promising tool for detection of several neurological disorders such as Alzheimer's, epilepsy, brain strokes [3, 4]. Review of previous research highlights that developing efficient system for automated Alzheimer's disease and dementia diagnosis using various biomarkers such as electroencephalography (EEG), magnetic resonance imaging (MRI), positron emission tomography (PET) has become difficult as well as challenging task and it has increased interests of researchers, clinicians, and scientists worldwide in this field. Earlier research findings indicate that feature extraction techniques and classification accuracy are main issues in Alzheimer's diagnosis. This paper presents EEG signal analysis in time and frequency domain for detection of Alzheimer's disease in early stage using spectral- and complexity-based features by use of suitable machine learning algorithms. Abnormalities in EEG signals of Alzheimer's disease patients not only reflect the anatomical deficits but also reflect the functional deficits of the cerebral cortex damaged by the disease. Nonlinear dynamic analysis (NDA) of EEG reveals out the loss of complexity of EEG signals and reduction in functional connections in Alzheimer's patients [3–5]. Thus, NDA of EEG signals provides important information about the progress of the disease as compared to other conventional techniques [3, 5].

The paper is organized as follows: Materials and Methods are described in Sect. 2. Database details and performance analysis of EEG data are explored in Sect. 3. Machine Learning and Classification Techniques are discussed in Sect. 4, and Conclusion is summarized in Sect. 5.

## 2 Materials and Methods

### 2.1 Participants Information

EEG data in present research work was taken from Smt. Kashibai Navale Medical College and General Hospital comprising of both mild Alzheimer's disease and age-matched healthy patients termed as normal subjects under the supervision of experienced neurologists. Alzheimer's patients were diagnosed by experienced neurologists according to NINCDS-ADRDA criteria, based on Indian version of Mini-Mental State Examination (MMSE) and Clinical Dementia Rating (CDR). Multi-channel EEG signal data was collected from 100 patients classified into 2 groups. Fifty subjects in group 1 consisted of: 40 males and 10 females (mean age: 65 years) indicating functional as well as behavioral decline. Similarly, 50 subjects in group 2 consisted of: 30 male patients and 20 female patients (mean age: 62.5 years) giving no indication of functional as well as behavioral decline. Patients belonging to the abnormal group also underwent single-photon emission computed tomography

(SPECT) scanning which was followed for 14–15 months since only functional tests such as MMSE and CDR tests are not enough to set Alzheimer's diagnosis. An additional criterion used was the presence of functional and cognitive decline for last 10 months based on interview with knowledgeable informants. The patients were also tested for different medical disorders such as diabetes, kidney disease, thyroid tests, and vitamin B12 deficiency since it can also result in cognitive decline. EEG data recordings and the study details were approved by ethical committee of the hospital and participants.

## ***2.2 EEG Data Acquisition and Recordings***

EEG recording was collected according to international 10–20 electrode placement system as recommended by American EEG society. Recorders and Medicare Systems (RMS) EEG machine was used for recording. Twenty-four-channel EEG signals were acquired with participants awake, relaxed, and their eyes closed for 15–20 min. EEG signals were recorded using RMS, India EEG machine with 12-bit resolutions and 200 Hz sampling rate. Impedance of the EEG machine was maintained below 10 Mohms. EEG signals were filtered using third-order Butterworth band-pass filter between 0.5 and 30 Hz. After successful recording, EEG data has been successfully inspected by clinical technician. EEG recordings are susceptible to certain artifacts such as electronic smog, head movement, and muscular activity. For each subject, one EEG segment of 20 s (termed as “epochs”) was extracted for analysis. These epochs are further used in the study for analysis.

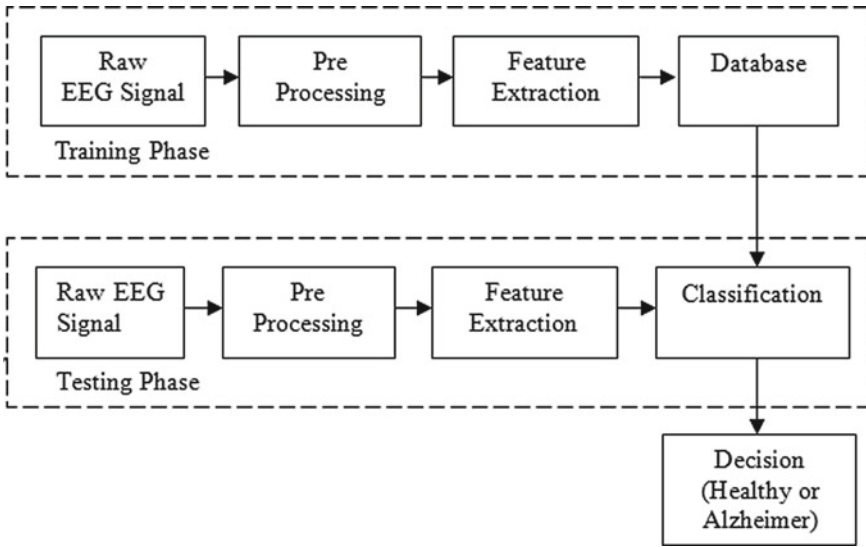
## **3 Analysis of EEG Data**

EEG signal is a nonstationary signal, and it is quite difficult to analyze it clinically. Brain rhythms exist in time as well as frequency domain. In present research, EEG signal is explored in time, frequency, and time–frequency domain [5, 6]. MATLAB (2013b version) software is used for implementing the algorithms proposed in present research. In this section, spectral- and complexity-based features are explored in detail. These features provide better performance results for distinguishing the subjects between two groups.

### ***3.1 Feature Extraction and Proposed Features***

The block diagram of the system methodology is shown in the following Fig. 1.

Alzheimer's disease diagnosis system is processed as (i) preprocessing of raw EEG signal, (ii) feature extraction of EEG data, and (iii) classification. Initially, raw



**Fig. 1** Proposed system methodology used in present research work

EEG signal is preprocessed to eliminate an artifact which gets added during signal acquisition. In preprocessing, artifacts of the signal are removed. In preprocessing stage, the EEG signal is filtered using third-order Butterworth band-pass filter between 0.5 and 30 Hz. In feature extraction stage, spectral and complexity features have been computed and it is given as an input to classifier.

### 3.2 Spectral-Based Features

Spectrum of EEG signal is affected by neurodegenerative diseases such as mild cognitive impairment and Alzheimer's. Recent research works have reported that Alzheimer's and mild cognitive impairment (MCI) cause EEG signal to slow down. Slowing effect in EEG signal of AD patients is observed by computing power spectral density (PSD) in different EEG frequency bands. EEG signal frequency bands are helpful in obtaining important information of the patients. From the results obtained in the study, it is seen that relative power (RP) is high in low-frequency bands (delta and theta bands), i.e., frequency range between 0.5 and 8 Hz. This effect of increase in power of low-frequency bands and decrease of power in high-frequency bands of EEG signal is observed in AD patients [5–8]. This irregularity of EEG signal is also quantified by using various standard measures; one of those is Lempel–Ziv complexity discussed in [6, 7, 9]. Spectrum of EEG is useful in understanding brain activity. For extracting spectral features, i.e., relative power, EEG signal is separated into four frequency bands, namely: 0.5–4 Hz (delta), 4–8 Hz (theta), 8–12 Hz (alpha),

and 30–100 Hz (gamma) using *wavedec* function. In present study, decomposition of EEG signal into five frequency sub-bands was done by use of “Daubechies” wavelet. Daubechies wavelet (db2, level 5) was used since (i) it has wide smoothing characteristics, (ii) easy understanding of the nature of the signal, and (iii) changes in the EEG signals are easily observed [10]. EEG signals are decomposed using “db2” wavelet at level of 5. In present work, relative power of EEG signal of four electrodes, namely frontal (F3 and F4), central, parietal, and temporal, is computed for all five EEG sub-bands.

The power spectral density (PSD) function helps in assessment of spectral characteristics of EEG activity of each epoch. Each sub-band power of EEG signal is calculated as Fourier transform of its autocorrelation function [11]. The normalized PSD in frequency range of 0.1–40 Hz is given as

$$RP = \sum_{f_{low}}^{f_{high}} PSD_n(f) \tag{1}$$

### 3.3 Complexity-Based Features

As benchmark, several features have been reported in the literature for analyzing the EEG signals in Alzheimer’s diagnosis. For each EEG electrode, namely central (C3), frontal (F3 and F4), parietal (P4), different complexity features such as spectral centroid, spectral entropy, and spectral flux are computed. In previous studies, Staudinger et al. [12] used some of these features for Alzheimer’s diagnosis for severe Alzheimer’s disease patients using event-related potentials (ERP) EEG signals. But, in this study some of those features are computed and tested for non-ERP EEG recordings to increase the performance of the system for getting better classification as well as diagnostic accuracy. These features depict the nonlinear changes introduced in the brain activity in case of Alzheimer’s disease patients. Let us discuss these features in more detail.

**Spectral Entropy (SE):** Spectral entropy indicates the amount of irregularity and disorder in spectrum of EEG. Higher complexity is achieved if higher amount of spectral entropy is observed [12, 13]. It is computed in the following manner:

Spectral entropy is computed by using formula

$$\text{Spectral Entropy} = \sum_{f=0.5}^{30} S(f) * \ln \frac{1}{S(f)} \tag{2}$$

where  $S(f)$  is the power spectral density (PSD) of the signal  $x(t)$  in given frequency band between 0.5 and 30 Hz.

**Spectral Centroid (SC):** Spectral centroid measures shape and position of the spectrum of EEG signals [12, 13]. Spectral centroid is calculated as

$$\text{Spectral Centroid} = \frac{\sum_{k=0}^{N-1} X(f_k) f_k}{\sum_{k=0}^{N-1} X(f_k)} \tag{3}$$

where  $X(f_k)$  is spectral magnitude of  $k$ th sample and  $f_k$  is frequency corresponding to each magnitude element.

**Spectral Flux (SF):** Spectral flux counts the change in spectral information between two successive frames [13]. It is computed in the following manner

$$\text{Spectral Flux} = \left( \sum_{k=2}^k |X(f_k) - X(f_{k-1})| \right) \tag{4}$$

### 4 Machine Learning and Classification Techniques

Machine learning is technique of programming to optimize a performance criterion based on past experience. The performance of the system is analyzed by means of various machine learning algorithms. For testing the performance of system, different classifiers are available in machine learning and pattern recognition field. Waikato environment for knowledge analysis (WEKA) is also one of the techniques for classifying the data using JAVA platform. But, in present research work, MATLAB pattern recognition toolbox is used for classification. During classification process, 70% of the data was used for the purpose of training the data using tenfold cross-validation technique and 30% data was left for testing to check the system performance. Confusion matrix is created from these machine learning algorithms which can be helpful in calculating different parameters such as accuracy, sensitivity, and specificity [14]. In present research work, supervised learning approach is used, which requires a large database. Figure 2 shows the supervised recognition flow.

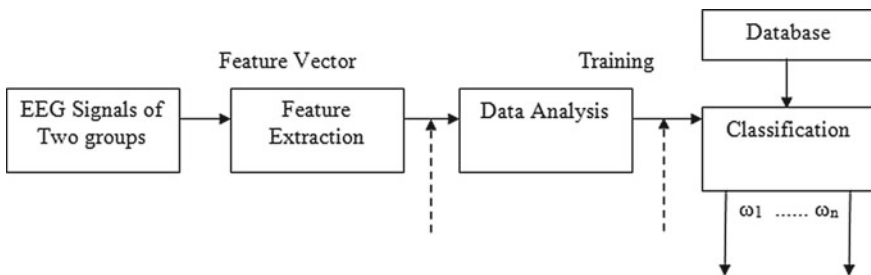


Fig. 2 Statistical recognition flow for supervised classification

### 4.1 Support Vector Machine (SVM) Classifier

This classifier relies on the principle of statistical learning. Classification and regression of data are typically carried out by SVM classifier. It is robust, efficient, and very effective in case of classification of high dimensionality data. In this classifier, an optimal hyperplane linearly separates data points belonging to two or more classes in case of higher dimensionality data. Researchers and scientists used SVM classifier for classification of data since it is simple and efficient [15]. Leave-one-subject-out (LOSO) cross-validation method is used for classification performance calculation. LOSO cross-validation technique is helpful since it avoids overfitting problem during separation of data and ensures the ease of the classifier to unseen data.

### 4.2 K-Nearest Neighbor (KNN) Classifier

Like SVM classifier, KNN classifier is also simple, efficient, and robust. It searches neighborhood of  $K$  in the training data, and a class is assigned that appears in the neighborhood of  $K$ . The value of  $K$  can be changed for each testing phase to find the match class between training and testing data, which is also used to obtain better classification rates. Basically, “1” is taken as the default value of  $K$ . But, the  $K$  value can be varied from 1 to 10 for each testing phase. “Euclidean” and “nearest” are default neighborhood setting values. For searching the object similarity in the neighborhood of  $K$ , Euclidean distance is normally used [15].

### 4.3 Performance Analysis of the EEG Data

The spectral and complexity algorithms are analyzed and evaluated successfully on EEG data. Results in terms of classification rates as well as diagnostic accuracy are compared with each spectral and complexity features using SVM and KNN classifiers. The obtained results are shown in Table 1.

**Table 1** Classification accuracy obtained for different features

Features	Accuracy (SVM)	Accuracy (KNN)
Spectral	88	82
Spectral entropy	79	78.33
Spectral centroid	74	73
Spectral flux	76	75
Combination of all features	96	94

## 5 Conclusion

In this paper, spectral- and complexity-based features are explored briefly for Alzheimer's disease diagnosis. The objective of current study was to provide better results in terms of classification rates and/or diagnostic accuracy. Although the proposed features in this paper are not novel, some improvements in results in terms of classification rates are observed. Each individual feature was computed and classified by use of both the classifiers discussed. It is also to note that EEG data analysis is mainly done for four brain regions such as frontal, parietal, central, and temporal as these are the regions where changes in the brain tend to occur at initial stage in case of Alzheimer's. EEG spectrum provides us useful information clearly stating the changes occurring in the Alzheimer's patients. Spectral features such as relative power (RP) are useful in discriminating the two subjects for classification in terms of accuracy obtained. Relative power is computed in each sub-band of EEG signal to highlight the difference between spectrum of EEG signal of normal and Alzheimer's affected patients. From this EEG spectrum, we also computed different nonlinear features which also provided satisfactory results in terms of diagnostic accuracy. Appearance of plaques and neurofibrillary tangles in the cortex and decrease in volume of hippocampus slows down the EEG of Alzheimer's disease patients. In present study, SPECT scans of Alzheimer's patients were also studied with expert clinicians since only functional tests such as MMSE and CDR are not enough to validate the results for diagnosis obtained in present research. It is also to highlight that each band of EEG signal equally contributes for analysis. Support vector machine and K-nearest neighbor classifier also helped to obtain satisfactory diagnostic results on the features selected for diagnosis purpose.

It is to conclude that when we combine all features together, classification rate as well as diagnostic accuracy obtained is more and it provides comparatively better results in terms of accuracy. Future work in this study involves making analysis of each frequency bands in depth to observe whether they carry any other significant information for better diagnosis by means of various signal processing algorithms and features extraction techniques. Our future work also includes implementing present algorithms on hardware devices such as DSP processors (application-specific integrated circuits), field-programmable gate array (FPGA) devices to make a stand-alone device for diagnosis which might be useful for doctors for correctly diagnosing the patients in early stage [16]. This is a difficult task but possible to implement, and it may come out with some new technology in the market since there is no such device available in the market for Alzheimer's diagnosis. This work explores new tool for Alzheimer's disease diagnosis, and further research using some more of these features can report remarkable achievements in this field.

**Ethical Statement** The database in present research work was collected from Smt. Kashibai Navale Medical College and General Hospital, Pune, under the supervision of Dr. Nilima Bhalerao, Neurosurgeon. The EEG data set and its details were approved by ethical committee of the hospital and participants.



## References

1. Mattson M (2004) Pathways towards and away from Alzheimer's disease. *Nature* 430:631–639
2. Meek PD, McKeithan K, Shumock GT (1998) Economics considerations of Alzheimer's disease. *Pharmacotherapy* 18:68–73
3. Wan J, Zhang Z, Rao BD, Fang S, Yan J, Saykin AJ, Shen L (2014) Identifying the neuroanatomical basis of cognitive impairment in Alzheimer's disease by correlation and nonlinearity-aware sparse bayesian learning. *IEEE Trans Med Imaging* 33(7):1475–1487
4. Kulkarni Kulkarni, Rathod PP, Nanavare VV (2017) The Role of Neuroimaging and Electroencephalogram in diagnosis of Alzheimer disease. *Int J Comput Appl (IJCA)* 3(4): 40–46
5. Jeong J (2004) EEG dynamics in patients with Alzheimer's disease. *Clin Neurophysiol* 15(7):1490–1505
6. Dauwels J, Srinivasan K, Ramasubba Reddy M, Musha T, Vialatte F-B, Latchoumane C, Jeong J, Cichocki A (2011) Slowing and loss of complexity in Alzheimer's EEG: two sides of the same coin? *Int J Alzheimers Dis* 2011:539621
7. Cassani Raymundo, Falk Tiago H, Fraga Francisco J, Kanda PAM, Anghinah R (2014) The effects of automated artifact removal algorithms on electroencephalography-based Alzheimer's disease diagnosis. *Frontiers in Aging Neuroscience* 6:1–13
8. Van der Hiele K, Vein AA, Reijntjes RH, Westendorp RG, Bollen EL, van Buchem MA, van Dijk JG, Middelkoop HA (2007) EEG correlates in the spectrum of cognitive decline. *Clin Neurophysiol* 118(9):1931–1939
9. Czigler B, Csikos D, Hidasi Z, Anna Gaal Z, Csibri E, Kiss E, Salacz P, Molnar M (2008) Quantitative EEG in early Alzheimer's disease patients—power spectrum and complexity features. *Int J Psychophysiol* 68(1):75–80
10. Daubechies I (1992) Ten lectures on wavelets. Society for Industrial and Applied Mathematics, Philadelphia, PA
11. Kang Yue, Escudero Javier, Shin Dae (2015) Principal dynamic mode analysis of EEG data for assisting the diagnosis of Alzheimer's disease. *IEEE J of Trans Eng Health Med* 3:1–10
12. Staudinger T, Polikar R (2011) Analysis of complexity based eeg features for diagnosis of alzheimer disease. In: *Proceedings International Conference of the IEEE-EMBC*. Boston, USA pp 2033–2036
13. Giannakopoulos T, Pikrakis A (2014) Introduction to audio analysis: a MATLAB approach. Elsevier
14. Rueda Andrea, Gonzalez Fabio A (2014) Extracting salient brain patterns for imaging based classification of neurodegenerative diseases. *IEEE Trans Med Imaging* 33(6):1262–1274
15. Suresh M, Ravikumar M (2013) Dimensionality reduction and classification of color features data using SVM and KNN. *Int J Image Process Visual Commun* 1(4):2319–1724
16. Kulkarni N (2017) *Int J Inf Technol*. <https://doi.org/10.1007/s41870-017-0057-0>