

Diagnosis of Alzheimer's disease by means of Wavelet & Synchrony based features & Machine Learning algorithms

¹Nilesh N Kulkarni, ²Saurabh V Parhad

¹Assistant Professor, ²Lecturer

¹Department of E & TC Engineering,

¹Smt. Kashibai Navale College of Engineering, Pune, India

Abstract: Previous studies have highlighted that EEG signal of Alzheimer's disease patients tends to be less complex and have low synchronization as compared to that of healthy and normal subjects. These changes in EEG signals of Alzheimer's disease patients start at early stage but are not clinically observed and detected. To detect these abnormalities, three synchrony measures and wavelet based features have been computed and studied on experimental database. After computing these synchrony measures and wavelet features, it is observed that Phase Synchrony and Coherence based features are able to distinguish between Alzheimer's disease patients and healthy subjects. Combining these features, healthy subjects and Alzheimer's disease patients are classified by use of different machine learning algorithms such as Support Vector Machine, Deep Learning and Naive Bayes Classifier. Combining, these synchrony features and other such relevant features can yield a reliable system for diagnosing the Alzheimer's disease.

Index Terms – Alzheimer's disease, Dementia, EEG, Complexity features, Support Vector Machine Classifier.

I. INTRODUCTION

Alzheimer's disease (AD) is one of the neurodegenerative diseases, characterized by progressive impairment that gradually destroys brain cells and other cognitive functions. Basically, it progressively leads to total dependency. It is most common neurodegenerative disorder particularly in western countries such as United States (US), Canada and many more. It is third most expensive disease and sixth leading cause of death in United States and modern society all over the world. Approximately, 50-60% of patients with dementia and Mild Cognitive impairment (MCI) over 65 years of age progresses towards Alzheimer every year [1] [2]. Therefore, diagnosis & effective treatment of patients in early stage are critical and important issues in Alzheimer's research. Alzheimer's disease is basically characterized by wide spread neuronal cell loss, neurofibrillary tangles & senile plaques in the various brain regions such as hippocampus, entorhinal cortex, neocortex and much more. As number of individuals with Alzheimer's disease is increasing and about to increase in future, diagnosis and effective treatment of disease in early stage are important research challenges. To search for a computational technique for early diagnosis of patients who are progressing towards Alzheimer's, but do not exhibit any clinical symptom of AD during medical tests is an important challenge. Despite of this, an early diagnosis screening method must be inexpensive, in order to allow screening of elderly patient. EEG is one of the most promising tools in such a case.

Neuroimaging techniques play significant role in Alzheimer's diagnosis but they possess several drawbacks as discussed in [12]. Different imaging techniques such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT), Single Photon Emission Computed Tomography (SPECT) and many more are specifically used for diagnosis of various diseases. But, the use of these techniques is also reported in literature. Although these tools are used for Alzheimer's diagnosis, EEG technique is specifically emphasized in current research for Alzheimer's diagnosis. EEG is one of the well established and well-known tools for measuring the electrical activity generated by population of neurons of the cerebral cortex. The bioelectric cells are measured non-invasively by set of electrodes placed over the head of the patients. Basically, EEG is a multivariate signal acquired over various channels (about 20-128) with different sampling rates (from 200 – 1000 Hz); depending on the clinical applications [4] [5]. This technique requires much time since it generates huge data which is to be stored and transmitted. EEG is one of the key diagnostic tools for neurologists, clinicians and doctors. Along with these applications, EEG signals are widely in Brain Computer Interface (BCI) applications [4]. Various research findings have identified the potential of EEG for diagnosing dementia and Alzheimer's disease in recent years. Since, EEG recording systems are relatively cheap, inexpensive and portable. In future, EEG technique can be used as a tool for screening large population for the risk of different neurodegenerative disorders such as Alzheimer's disease, Epilepsy, Huntington disease and many more. There are various dynamical changes of EEG signals related to normal ageing and this reflection can be easily observed by use of EEG [4]. Recent research findings have proved the EEG complexity analysis can be used to diagnose Alzheimer's disease in early stage [5].

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

II. MATERIALS & METHODS

A. MATERIALS

100 patients were selected from Smt. Kashibai Navale Medical College and General Hospital, Neurology Unit Pune. Diagnosis was made by experienced neurologists, neurosurgeons according to NINCDS-ADRDA [5] criteria and classified on Indian version of MMSE. Participants were classified into two groups. The first group (Normal, N) consisted of 50 cognitively healthy & normal patients (30 men, 20 women, mean age 65 years and 8.2 sd); the second group (mild Alzheimer & dementia) consisted of 50 patients with cognitive decline (35 men, 15 women, mean age 75, 6.3 sd). Inclusion criteria for Normal group included CDR score of 0 and MMSE score ≥ 25 and no indication of

cognitive decline in their behavior. Inclusion criteria for AD group included $0.5 \leq \text{CDR} \leq 1$ and $\text{MMSE} \leq 24$. An additional criteria used was the presence of functional and cognitive decline over last 10 months based on interview with knowledgeable informants. The patients were also tested for diabetes, kidney disease, thyroid tests and vitamin B12 deficiency as they can also cause cognitive decline. The EEG recordings and the study was approved from Ethical committee of the hospital and participants.

B. MATERIALS

Twenty one channel EEG signals were acquired with participants awake, relaxed and their eyes closed for 15 to 20 minutes. EEG signals were recorded using RMS, India EEG machine with 12 bits resolutions and sampling rate of 1024Hz. Impedance of the EEG machine was maintained below 10Mohms and the electrodes were placed according to the 10/20 International electrode placement system as recommended by American EEG society. EEG signals were Bandpass filtered using third order Butterworth filter between 0.5-30Hz. After successful recording, EEG data has been successfully inspected by Clinical technician. EEG recordings are prone to certain artifacts such as electronic smog, head movement and muscular activity etc. For each subject, one EEG segment of 20 seconds (termed as 'epochs') was extracted for analysis. These epochs are further used in the study for analysis.

III. METHODOLOGY

Due to non-stationary characteristic of EEG signal; 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, we discuss the preprocessing and feature extraction process in detail.

I. Block Diagram of the Proposed System

Alzheimer's Disease diagnosis system works in different steps, i) Preprocessing of raw EEG signal, ii) Feature extraction of EEG data, and iii) Classification between two subjects. Initially, raw EEG signal is pre-processed to eliminate an artifact which gets added during signal acquisition. In preprocessing stage, the EEG signal is filtered using third order Butterworth Band pass filter between 0.5-30Hz. In feature extraction stage, synchrony based different features are computed and a training feature matrix is computed as an input to classifier.

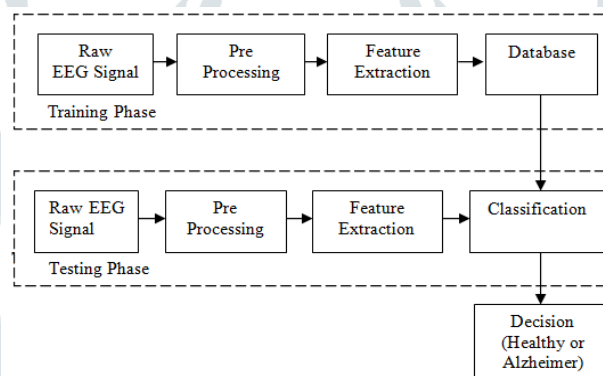


Fig. 1. Block Diagram of the Proposed System

II. Feature Extraction using Synchrony based features

In the present research study, we emphasize on synchrony based features. Synchrony justifies simultaneous appearance of rhythmic distinct patterns over different regions of head; bilaterally or unilaterally. It is observed that statistical independence seems to be low in case of AD patients around different regions of brain [5] [6] [7]. To observe this loss of synchrony, we have selected two synchrony measures which comparatively provide better results in terms of classification accuracy. The following features are used in present research work:

1. Magnitude Squared Coherence
2. Phase Synchrony

Let us discuss these features in more detail.

1) Magnitude Squared Coherence

Spectral Coherence refers to measurement of synchrony between two independent signals (Let us say x and y) and it can be mathematically defined as,

$$C_{xy} = \frac{|P_{xy}(f)|^2}{|P_{xx}(f)|P_{yy}(f)} \quad (1)$$

Where, f is frequency in hertz, P_{xy} denotes cross spectral density of x and y . P_{xx} and P_{yy} represents the power spectral densities of x and y respectively. For each pair of channels namely F3, F4, C3, C4 and O1, O2; coherence features were computed and averaged over the five sub-bands. Coherence highlights the synchronicity of activation in various brain regions. EEG coherence helps to access functional connectivity between various channels over occipital regions of brain. In total, we have computed total (6 channels * 4 frequency bands = 24 features) for 50 subjects consisting of healthy as well as Alzheimer affected in training mode. So, it consists of total $50 * 24 = 1200$ features.

2) Phase Synchrony

The interdependence among two phases ϕ_x and ϕ_y of two signal x and y is termed as Phase synchrony. Although the amplitudes of x and y are statistically dependent, the instantaneous phases can be strongly synchronized. The instantaneous phases can be synchronized strongly even when the amplitudes of x and y are statistically independent. It is given by following formula [8]:

$$\phi_{m,n} = m\phi_x(t) - n\phi_y(t) = \text{constant} \quad (2)$$

Where m and n represents integers indicating frequency locking ratio and $\phi_{m,n}$ is its relative or phase difference. For computing these phase synchronization, it is important for us to know the instantaneous phase of two signals. This can be analyzed by using Hilbert transform which is given as:

$$z(t) = x(t) + \tilde{x}(t) \quad (3)$$

In above equation, $z(t)$ is complex term with $x(t)$ real term and $\tilde{x}(t)$ represents its Hilbert transform. Hilbert Transform is computed as,

$$\tilde{x}(t) = \frac{1}{\pi} PV \int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau} dt \quad (4)$$

Here, PV represents Cauchy Principle value. Instantaneous phase $\phi_x(t)$ and $\phi_y(t)$ for signals x and y are calculated as:

$$\phi(t) = \arctan\left(\frac{\tilde{x}(t)}{x(t)}\right) \quad (5)$$

In present study, we applied synchrony measurement techniques on each pair of channel of two opposite regions for all frequency bands and then averaged. For example, we applied phase synchrony feature measurement on each pair of channel of right and left temporal regions (say (F3-T3), (T3-T4)) and so on. In present study, we computed phase synchrony for frontal, temporal, parietal and central electrode of EEG signal. Accordingly, 150 such features were computed and out of it 100 features were selected as per Wilcoxon tests on our dataset [8] [9].

IV. CLASSIFICATION TECHNIQUES

Machine learning is a method of programming to optimize a performance criterion based on past experience. The performance of the system can be analyzed by means various machine learning algorithms. For testing the performance of system, different classifiers are available in machine learning and pattern recognition subject. Waikato environment for knowledge analysis (WEKA) is also one the technique for classifying the data using JAVA platform. But, in present study, MATLAB Pattern Recognition toolbox was used for classification purpose between two groups. During classification process, 70% data was used for training using 10-fold cross-validation technique and 30% data was left for testing to test the performance of the system for classification. Confusion matrix is created from these machine learning algorithms which can be helpful in calculating different parameters such as accuracy, sensitivity and specificity [12]. In present research work, semi-supervised learning approach is used, which requires a large database. Fig. 2 shows the supervised recognition flow.

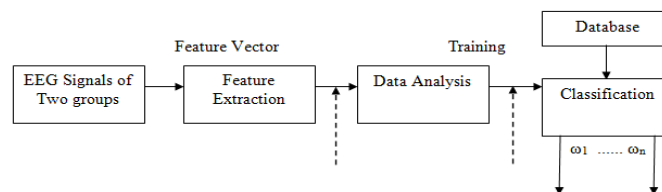


Fig. 2. Supervised Recognition flow for Supervised Classification

A. Support Vector Machine Classifier

Statistical Learning theory is basically used for formulating this classifier. It is used for classification as well as regression of data. This classifier is robust, efficient and very effective in case of classification of high dimensionality data. In this classifier, an optimal hyperplane separates data points belonging to two classes in case of higher dimensionality data linearly. Researchers and scientists used SVM classifier for classification of data since it is simple and efficient [10] [12]. Leave-one-subject-out (LOSO) cross-validation approach is used for classification performance calculation. LOSO cross-validation paradigm is helpful since it avoids over fitting problem during separation of data and ensures the generality of the classifier to unobserved data.

B. Performance Analysis of the EEG Data

The synchrony based features such as Magnitude Squared Coherence (MSC), Phase Synchrony and Cross - correlation are analyzed and evaluated successfully on EEG data. For selecting training data, computed features were selected from any of the four selected electrodes. Results in terms of classification rates as well as diagnostic accuracy are compared with synchrony features using SVM classifier. The obtained results are shown in Table.1 Based on database available for computation, 50 % of the data was trained & remaining 50 % data was left out for testing purpose. The computed values for different features used in different electrodes are used for training & testing purpose. This means 50 patients database was used for training while 50 patients database was left for testing purpose. The reason behind this is that each feature is tested individually during classification and on the similar basis the results were obtained. Based on the features calculated & classifier used, we have calculated the accuracy of classification based on following terminology [9] [10].

Based on the features calculated and classifier used, accuracy of the system is computed based on following terminology [5].

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (6)$$

$$Sensitivity = \frac{TP}{(TP+FN)} \quad (7)$$

$$Specificity = \frac{TN}{(FP+TN)} \quad (8)$$

Where, TP stands for True Positive (Correctly classified AD individuals), TN stands for True Negative (Correctly classified N individuals), FP stands for False Positives (Misclassified N individuals), FN stands for False Negative (Misclassified AD individuals) [7] [8].

Table 1. Classification Accuracy obtained for different features

Features	Accuracy (in %)
Magnitude Squared Coherence (MSC)	88
Phase Synchrony	92
Combination of MSC and Phase Synchrony	96

V. CONCLUSION & RESEARCH CHALLENGES

In this paper, Synchrony 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 based on the features used. Although the proposed features in this paper are not novel, but some improvements in results in terms of classification rates are observed. Each individual feature was computed and classified by use of SVM classifier 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 tends to occurs at initial stage in case of Alzheimer's. EEG spectrum provides us useful information clearly stating the changes occurring in the Alzheimer patients. In [11] [12], it is highlighted that appearance of plaques and neurofibrillary tangles in the cortex and decrease in volume of hippocampus slows down the EEG of Alzheimer disease patients. In present study, SPECT scans of Alzheimer 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.

In previous research work, Synchrony based features were investigated using several features as discussed in [13] [14] for identifying brain dynamics in Alzheimer's disease and healthy subjects. After computations of these features, it is observed that Coherence feature shows significant effect in EEG signals of Alzheimer disease patients in resting conditions in high frequency ranges (13 to 30Hz). But, in gamma frequency range, coherence tends to decrease with Alzheimer disease; but this observation is not included in present work since EEG signal is filtered from 0.5 – 30Hz. The synchrony based features were computed individually and classified by use of semi-supervised classifier; Support vector machine (SVM) giving better classification accuracy for distinguishing the subjects between two groups.

From the features computed as discussed in section 3, it is observed that lower dysfunctional connectivity is observed in central, parietal and occipital brain regions. A lower synchronization in these areas for alpha and beta band identifies disturbance occurred in perception and cognitive disorder. Thus, the observations computed in present research work are in agreement with several previous studies which reports on decrease of neuronal synchrony in Alzheimer's disease patients as observed on EEG signals. This loss of synchrony appears due to addition of non – consistent background activity and presence of neurofibrillary tangles in hippocampus area.

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), FPGA (Field Programmable Gate Array) devices to make a standalone device for diagnosis which might be useful for doctors for correctly diagnosing the patients in early stage. This work explores new tool for Alzheimer disease diagnosis and further research using some more of these features can report remarkable achievements in this field.

Acknowledgment

Authors gratefully acknowledge entire staff of E & TC department for their support & guidance in the present work. Author would also like to acknowledge staff of Medicine Dept. of Smt. Kashibai Navale Medical College & General Hospital for providing necessary database & constant help & guidance in present work.

REFERENCES

- [1] Mattson M, Pathways towards and away from Alzheimer's disease, *Nature*, Vol. 430, pp. 631–639, 2004.
- [2] Meek PD, McKeithan K, Shumock GT: Economics considerations of Alzheimer's disease. *Pharmacotherapy*. Vol. 18, pp. 68 – 73, 1998.

- [3] Wan J, Zhang Z, Rao BD., Fang S, Yan J, Saykin A J., Shen L: Identifying the Neuroanatomical Basis of Cognitive Impairment in Alzheimer's disease by Correlation and Nonlinearity-Aware Sparse Bayesian Learning. *IEEE Trans Med Imaging*. Vol. 33, No. 7, pp. 1475-1487, 2014.
- [4] N. N. Kulkarni, V. K. Bairagi, Electroencephalogram based diagnosis of Alzheimer disease, in *Proc. IEEE Conference on Intelligent Systems and Control*, pp. 1-5, 2015.
- [5] J. Jeong. "EEG dynamics in patients with Alzheimer's disease", *Clin. Neurophysiol.* Vol. 15, No. 7, pp. 1490–1505, 2004.
- [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] Raymundo Cassani, Tiago H. Falk, Francisco J. Fraga, Paulo A. M. Kanda and Renato Anghinah, "The effects of automated artifact removal algorithms on electroencephalography-based Alzheimer's disease diagnosis", *Frontiers in Aging Neuroscience*, Vol. 6, pp. 1 – 13, 2014.
- [8] Dhiya Al-Jumeily, Shamaila Iram, Francois-Benois Vialatte, Paul Fergus, and Abir Hussain, "A Novel Method of Early Diagnosis of Alzheimer's Disease Based on EEG Signals," *The Scientific World Journal*, vol. 2015, Article ID 931387, 11 pages, 2015. doi:10.1155/2015/931387.
- [9] Andrea Rueda, Fabio A. Gonzalez, "Extracting Salient Brain Patterns for imaging based Classification of Neurodegenerative Diseases", *IEEE Trans. Med. Imaging*, Vol. 33, No. 6, pp. 1262-1274, 2014.
- [10] Suresh, M., Ravikumar, M.: Dimensionality Reduction and Classification of Color Features data using SVM and KNN. *International Journal of Image Processing and Visual Communication* Vol. 1, No. 4, pp. 2319–1724, February 2013.
- [11] N. N. Kulkarni, S. V. Parhad and Y. P. Shaikh, "Use of Non-linear and Complexity Features for EEG Based Dementia & Alzheimer Disease Diagnosis," *2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, Pune, 2017, pp. 1-3.
- [12] N. N. Kulkarni, V. K. Bairagi, "Extracting Salient features for EEG based diagnosis of Alzheimer disease Using Support Vector machine Classifier", *IETE Journal of Research*, Vol . 63, No. 1, January 2017.
- [13] Dauwels J, Vialatte F, "A Comparative Study of Synchrony Measures for Early Diagnosis of Alzheimer's Disease Based on EEG", *Journal of Neuroimage*, 49(1):668-693, (2010).
- [14] Justin Dauwels, Francois-Benoit Vialatte, and Andrzej Cichocki, "On the Early Diagnosis of Alzheimer's disease from EEG Signals: A Mini-Review", *Advances in cognitive Neurodynamics*, 1 : (709 -716), 2010.

