Analysis of Complexity Based EEG Features for the Diagnosis of Alzheimer's Disease

Tyler Staudinger, Robi Polikar*

Abstract—As life expectancy increases, particularly in the developed world, so does the prevalence of Alzheimer's Disease (AD). AD is a neurodegenerative disorder characterized by neurofibrillary plaques and tangles in the brain that leads to neuronal death and dementia. Early diagnosis of AD is still a major unresolved health concern: several biomarkers are being investigated, among which the electroencephalogram (EEG) provides the only option for an electrophysiological information. In this study, EEG signals obtained from 161 subjects - 79 with AD, and 82 age-matched controls (CN) - are analyzed using several non-linear signal complexity measures. These measures include: Higuchi fractal dimension (HFD), spectral entropy (SE), spectral centroid (SC), spectral roll-off (SR), and zero-crossing rate (ZCR). HFD is a quantitative measure of time series complexity derived from fractal theory. Among spectral measures, SE measures the level of disorder in the spectrum, SC is a measure of spectral shape, and SR is frequency sample below which a specified percent of the spectral magnitude distribution is contained. Lastly, ZCR is simply the rate at which the signal changes signs. A *t*-test was first applied to determine those features that provide significant differences between the groups. Those features were then used to train a neural network. The classification accuracies ranged from 60-66%, suggesting they contain some discriminatory information; however, not enough to be clinically useful alone. Combining these features and training a support vector machine (SVM) resulted in a diagnostic accuracy of 78%, indicating that these feature carry complementary information.

Index Terms—Alzheimer's disease, EEG, Higuchi Fractal Dimension, Spectral Entropy, Spectral Centroid, Spectral Roll-Off, Zero-Crossing Rate

I. INTRODUCTION

Alzheimer's disease, the most common form of dementia, is a neurodegenerative disorder characterized by a progressive and severe loss of memory with cognitive impairment. It has a profound socio-economic impact with 18 million patients worldwide, and 5.4 million in the United States (U.S.) alone. The U.S. spends approximately $180 billion annually on research, treatment and care [1]. The exact causes of the disease are unknown, though abnormal proteins that form neurofibrillary tangles and plaques in the brain that ultimately kill neurons are implicated. There is no cure, however, limited treatment with recent pharmacologically targeted medications can be used to treat the symptoms and slow down the progression of the disease. As with most disorders, the effectiveness of these drugs decrease at the later stages of the disease, hence early diagnosis is essential. Early diagnosis is very difficult, however, and a definitive diagnosis is only possible with an autopsy. Currently, AD is diagnosed by neuropsychologists as a result of a battery of cognitive tests and a series of interviews with the subject and their caregivers. At dementia specific clinics, diagnostic accuracy rates reach 90%. However, at community healthcare clinics, where most subjects are evaluated, the diagnostic accuracy is around 75% [2]. Anatomical (using MRI) and metabolic (using PET imaging) biomarkers are currently being investigated, unfortunately these approaches remain costly. EEG, however, provides a neurophysiological biomarker that is inexpensive and can be made available at community clinics.

Our goal is to investigate and identify EEG based biomarkers that provide clinically meaningful diagnostic accuracy. In our previous work, we showed that wavelet coefficients obtained from event related potentials can provide 75-80% diagnostic accuracy. In this study, we look at signal complexity and spectral based features. Our hypothesis is that patients with AD have less signal complexity due to neuronal loss.

II. DATA COLLECTION

The data used in this study is a subset of a larger dataset originally intended to analyze event related potentials (ERPs) generated using an auditory oddball paradigm. The diagnosis was made by a team of neuropsychologists based on a comprehensive set of memory and cognitive tests, as well as interviews with the subjects and their caregivers.

The dataset used in this study consisted of 161 subjects, 79 with AD and 82 cognitively normal (CN). The subjects were age matched with mean ages $\mu_{AGE}(AD)=76$ and $\mu_{AGE}(CN)=77$, with mini mental state exam scores of $\mu_{MMSE}(AD)=21$ and $\mu_{MMSE}(CN)=28$. EEG data were collected through the oddball paradigm, modified from [3], where subjects heard three types of auditory stimuli: 1 kHz standard tones (about 65% of all stimuli), 2 kHz target tones (20%), and novel sound clips (15%). Each were presented randomly with an average 1.3 seconds of inter stimulus interval. Subjects were asked to press a button every time they heard a target tone. Each recording session was approximately 30 minutes in duration.

EEG signals were collected using a 16-electrode cap with the 10-20 standard electrode montage. Signals were sampled at 256 Hz. Using a sharp, 50th order digital Butterworth filter, the raw signals were bandpass filtered with cutoff frequencies of 0.5 Hz and 40 Hz. The data from the two frontal electrodes of Fp1 and Fp2 were not used due to high occurrence of ocular

artifacts. Rare ocular and muscular artifacts were removed from the remaining data using a 20th order derivative based thresholding. The data was originally acquired for wavelet based event related potential (ERP) analysis. In previous studies data from target and novel sounds were analyzed, the results of which can be found in [4-6]. Because we are not directly interested in ERPs in this analysis (which are obtained by time-locked averaging of signals in response to target and novel sounds), we used the standard tones since they provided substantially longer data records. Each subject possessed EEG data records of approximately 50,000 or more samples in response to the standard tones after the filtering and thresholding.

III. FEATURE EXTRACTION

Following filtering and artifact removal, EEG data from each electrode were divided into 5 second epochs of 1280 samples each. Each epoch was then tested for normality using the *Jarque-Bera* test [7]. This test has the null hypothesis that the samples originate from a normal distribution with unknown mean and variance. The test statistic for this test is

$$JB = \frac{n}{6}\left(s^2 + \frac{(k-3)^2}{4}\right) \qquad (1)$$

where n is the sample size, s is the sample skewness, and k is the sample kurtosis. If the null hypothesis is not rejected (indicating normality), the epoch was used to compute the features described below. Statistical significance was tested for each feature set, the p-values of which are provided below.

A. Higuchi Fractal Dimension (HFD)

The term fractal dimension, when used with time series data shows correlation at different scales. When applied to biomedical signals, fractal dimension can be interpreted as a quantitative measure of signal complexity [8]. The Higuchi method of fractal dimension measures the change in signal amplitude, while sampling the signal at increasingly longer intervals [8]. The original time series of length N is used to construct k subset time series as follows [9]:

$$x_m^k = \left\{x(m), x(m+k), x(m+2k), \dots x\left(m + \left[\frac{N-m}{k}\right]k\right)\right\} \quad (2)$$

where $m = 1, 2, \dots, k$ is the initial time index, and k is the delay between the samples. For each x_m^k, the average length is computed as:

$$L_m(k) = \frac{(N-1)\sum_{i=1}^{\left[\frac{N-m}{k}\right]}|(x(m+ik)-x(m+(i-1)k))|}{\left[\frac{N-m}{k}\right]k} \qquad (3)$$

The total average length $L(k)$ is computed for all time series having the same delay k but differing m values.

$$L(k) = \sum_{m=1}^{k} L_m(k) \qquad (4)$$

This is repeated for every k from 1 to the value of k_{max}. The slope of the least squares linear fit of the curve $\ln(L(k))$ *vs.* $\ln(1/k)$ is the estimate of the HFD [9]. The choice of the k_{max} parameter is obtained as the value k for which the HFD saturates. For this dataset, k_{max} was computed as 40; a value also confirmed by others [8]. The HFD was calculated for

every 1280 sample epoch, which was then averaged across all epochs to obtain the final HFD value for that subject. Table 1 shows the electrodes for which the final HFD values differed significantly between AD and CN groups.

TABLE 1. ELECTRODES WITH SIGNIFICANTLY DIFFERENT HFD VALUES

Electrode	P-Value	μ CN	σ CN	μ AD	σ AD
F3	.002	1.72	.04	1.69	.06
F4	.005	1.72	.05	1.7	.06
P3	.012	1.72	.04	1.7	.05
T7	.0001	1.76	.04	1.73	.05
T8	.008	1.78	.04	1.75	.06
P7	.0015	1.75	.03	1.72	.05
Fz	.0406	1.69	.05	1.67	.06

B. Spectral Entropy (SE)

Spectral entropy is a measure of unpredictability and disorder associated with the spectrum of a signal. Hence, higher SE indicates higher complexity. It is a modified version of Shannon's Entropy and can be obtained as follows [10]:

1. Given the signal $x(t)$, compute $S(f)$, the power spectral density (PSD), as the Fourier transform of the autocorrelation function of the signal $x(t)$;
2. Extract the spectral band corresponding to the frequency interval of interest, 0.5-40Hz, in this case.
3. Normalize the PSD to unit total power.
4. Compute the spectral entropy as

$$SE = \sum_{f=0.5}^{40} S(f) * \ln\frac{1}{S(f)} \qquad (5)$$

Spectral entropy was computed for each consecutive epoch of 1280 points (5 second long signal), which were then averaged across all epochs to obtain the final *SE* value. Table 2 shows the electrodes for which the final *SE* values differed significantly between AD and CN groups.

TABLE 2. ELECTRODES WITH SIGNIFICANTLY DIFFERENT SE VALUES

Electrode	P-Value	μ CN	σ CN	μ AD	σ AD
F3	.0066	5.26	.15	5.18	.22
T7	.0038	5.38	.14	5.29	.2
T8	.0399	5.36	.14	5.3	.22

The remaining three features - spectral centroid, spectral rolloff, and zero-crossing rate - were all frame based features with the following common preprocessing:
1) Break each epoch into frames of 12 samples in length, corresponding to approximately 20ms.
2) Calculate the corresponding feature value for each frame.
3) Compute the mean and standard deviation of the values across all frames for that epoch

C. Spectral Centroid

Spectral centroid (SC) is a measure of the shape of the spectrum. Higher values of SC correspond to more energy of the

signal being concentrated within higher frequencies, and is computed as follows [11]:

1. Let $x_i(n), n = 0, 1, \ldots, N-1$ be the samples of the ith frame, with $X_i(k), k = 0, 1, \ldots, N-1$ being the discrete Fourier transform coefficients of this sequence
2. Compute the spectral centroid of the ith frame as:

$$C(i) = \frac{\sum_{k=0}^{N-1} k|X_i(k)|}{\sum_{k=0}^{N-1}|X_i(k)|} \tag{6}$$

The mean value of the spectral centroid across all frames is then used as the *SC* feature for each epoch. Table 3 shows the electrodes for which the final *SC* values differed significantly between AD and CN groups.

TABLE 3. ELECTRODES WITH SIGNIFICANTLY DIFFERENT SC VALUES

Electrode	P-Value	μ CN	σ CN	μ AD	σ AD
F3	.0114	.22	.008	.21	.01
F4	.0239	.22	.009	.21	.01
T7	.0173	.23	.01	.22	.01

D. Spectral Roll-off

Spectral roll-off indicates the frequencies where most of the spectral content is focused [11]. The spectral roll-off is defined as the frequency sample, $k_c^r(i)$ below which a preset amount (P %) of the magnitude DFT coefficients reside. It is computed as follows [9]:

1. Let $x_i(n)$, $n = 0, 1, \ldots, N-1$ be the samples of ith frame, and $X_i(k), k = 0, 1, \ldots, N-1$ be the corresponding DFT coefficients;
2. Compute the spectral roll-off as the sample $k_P^r(i)$ that satisfies

$$\sum_{k=0}^{k_P^r(i)}|X_i(k)| = \frac{P}{100}\sum_{k=0}^{N-1}|X_i(k)| \tag{7}$$

where the P parameter is typically chosen as 80. The standard deviation of the spectral roll-off across all frames is computed to serve as the spectral roll-off for each epoch. This is the average variability in spectral skewness across frames, which is another measure of signal complexity. Table 4 shows the electrodes for which the final spectral roll-off values differed significantly between AD and CN groups.

TABLE 4. ELECTRODES WITH SIGNIFICANTLY DIFFERENT SPECTRAL ROLL-OFF VALUES

Electrode	P-Value	μ CN	σ CN	μ AD	σ AD
F3	.0002	.2302	.0082	.2235	.0136
F4	.0162	.2301	.0108	.2250	.0150
C3	.0010	.2324	.0077	.2271	.0117
C4	.0005	.2334	.0075	.2279	.0113
P3	.0053	.2319	.0092	.2269	.0116
P4	.0033	.2322	.0079	.2277	.0104
T7	.0001	.2346	.0075	.2286	.0106
T8	.0077	.2357	.0066	.2316	.0115
P7	.0143	.2336	.0106	.2288	.0128

E. Zero-crossing Rate

Zero-crossing rate (ZCR) is the rate at which the signal changes its sign, which can be interpreted as a time domain measure of signal complexity [11]. The zero-crossing rate is obtained as follows:

1. Let $x_i(n) = 0, 1, \ldots, N-1$ be the samples of the ith frame.
2. Compute the zero crossing rate for the ith frame as

$$Z(i) = \frac{1}{2N}\sum_{n=0}^{N-1}|sgn[x_i(n)] - sgn[x_i(n-1)]| \tag{8}$$

where

$$sgn[x_i(n)] = \begin{cases} 1, & x_i(n) \geq 0 \\ -1, & x_i(n) < 0 \end{cases} \tag{9}$$

The mean value of ZCR across all frames is first calculated for each epoch, and is then averaged across all epochs to obtain the final ZCR value. Table 5 shows the electrodes for which the final ZCR values differed significantly between AD and CN groups.

TABLE 5. ELECTRODES WITH SIGNIFICANTLY DIFFERENT ZCR VALUES

Electrode	P-Value	μ CN	σ CN	μ AD	σ AD
F3	.0191	.1114	.0123	.1064	.0149
F4	.0254	.1114	.0142	.1062	.0158
T7	.0069	.1222	.0171	.1148	.0170

IV. DIAGNOSTIC ACCURACY RESULTS

A. Features Evaluated Individually

To evaluate the diagnostic value of each of these set of features, a multilayer perceptron with 20 hidden layer nodes and an error goal of 0.05 was trained with 10 independent trials. Each trial consisted of a 6-fold cross validation. Scaled-conjugate gradient was used as the learning algorithm. The preliminary results are summarized in Table 6, which shows that the diagnostic accuracy of these features lie around the low to mid 60% range.

TABLE 6. INDIVIDUAL FEATURE SET PERFORMANCE (WITH MLP)

Feature Set	Performance
Spectral Centroid	62.61% +/-1.67%
Spectral Roll-Off	62.97%+/-1.14%
Spectral Entropy	60.28%+/-1.7%
Zero-Crossing Rate	61.75%+/-2.18%
Higutchi Fractal Dimension	66.84+/-1.84%

B. Feature Concatenation Based Data Fusion

These results indicate that these features carry only limited diagnostic value when used alone. To determine whether the features contain complementary information, they were concatenated into one feature vector of length five (each feature is the mean of each feature group across all statistically significant electrodes). This feature vector was then used to train a single MLP with 50 hidden layer neurons and an error goal of

.05. The increased amount of neurons was to account for the larger number of input patterns and expected increase in decision boundaries that the MLP was required to learn. The concatenated feature vector yielded a correct classification accuracy of 72.65% +/-1.88% using a six-fold cross validation. This is a significant improvement over HFD, the best individually performing feature. To ascertain how the features perform on a different classifier, the concatenated feature vector was also used to train a support vector machine that used a radial basis function kernel and a box constraint of 10,000. Several sigma values were chosen, ranging from 10 to 40. The results of a six-fold cross validation for the various values of sigma can be found in Table 7.

TABLE 7. COMBINED FEATURE SET PERFORMANCE (SVM)

Sigma Value	Performance
10	73.29 +/- 1.52%
20	76.45% +/-1.33%
30	77.1% +/- 1.4%
40	77.61% +/-1.69%

C. Ensemble of Classifiers Based Decision Fusion

Due to the previous success of ensemble based decision fusion approaches with AD diagnosis, [5,6], a simple ensemble of five MLPs was also examined. Three such ensembles were generated, each with a varying number of hidden layer nodes in the component classifiers. The classifiers in each ensemble were then combined through the sum rule.

$$\mu_j(x) = \sum_{t=1}^{T} d_{t,j}(x) \qquad (10)$$

where $\mu_j(x)$ is the total support received by class j and $d_{t,j}$ is the continuous valued output of classifier t for class j. Again, a six-fold cross validation was employed. The results for different numbers of hidden layer nodes can be seen in Table 8.

TABLE 8. DIAGNOSTIC ACCURACY OF THE ENSEMBLE SYSTEM (MLP)

Number of Hidden Layer Nodes	Performance
20	62.54% +/- 1.54%
30	67.34% +/-1.34%
50	71.35% +/-1.6%

V. CONCLUSIONS

We evaluated five different EEG based signal complexity measures to determine if they carry any diagnostically useful discriminatory information. Since AD affects neuronal integrity, the underlying hypothesis in this study is that EEG signals from AD subjects have less signal complexity than those from CN subjects. Each of the features showed decreased complexity values for AD cohort, confirming our hypothesis. The differences in complexity values between the cohorts were small, but statistically significant for electrodes listed in the above tables.

The results indicate that the HFD appears to carry relevant information, primarily in the parietal and temporal areas. The AD group had lower values of HFD, suggesting that AD subjects generate less complex signals. This decreased complexity could be attributed to the presence of neurofibrillary plaques and tangles. Spectral entropy values were also lower for the AD group in the frontal and temporal lobes. In the frontal and temporal lobes, there was more spectral content in the higher frequencies for CN subjects. When examining the standard deviation of the spectral roll-off, we observed in nearly every electrode that the CN group possessed a higher variation in spectral skewness across frames. This again can be interpreted as a higher level of complexity in the CN group. Finally, the ZCR of the CN group was also consistently higher than that of the AD group.

When used individually to train a single classifier, these features provided very limited diagnostic benefit. However, combining them into a feature vector significantly improved classification accuracy from the mid 60% to high 70% range. We tried both feature level concatenation and ensemble based decision fusion. The feature based fusion provided the better results.

To the best knowledge of the authors, this is the first study evaluating the above described features on ERP based EEG data for AD diagnosis. We note that these are very preliminary results obtained with classifiers whose parameters were not optimized. The results indicate that signal complexity based features, when combined with one another, can provide clinically meaningful diagnostic information. Our future work includes optimizing the classifiers as well as the feature sets (such as frame and epoch length, overlap between frames, etc.), and combining them with our previously reported wavelet based features.

REFERENCES

[1] Alzheimer's Assoc., "Alzheimer's disease statistics," Available at : http://www.alz.org/AboutAD/statistics.asp

[2] A. Lim, D. Tsuang, et al. "Clinico-neuropathological correlation ofAlzheimer's disease in a community-based case series," J. American Geriatrics Soc. vol. 47, no. 5, pp. 564-569, 1999.

[3] S. Yamaguchi, H. Tsuchiya, S. Yamagata, G. Toyoda, S. Kobayashi, "Event-related brain potentials in response to novel sounds in dementia," *Clinical Neurophysiology*, vol. 112, no. 2, pp. 195-203, 2000.

[4] M. Ahiskali, D.Green, J. Kounios, C. Clark, R. Polikar, "ERP Based Decision Fusion for AD Diagnosis across Cohorts," IEEE Engineering in Medicine and Biology Conference (EMBC 2009) , pp. 2494 - 2497 Minneapolis, MN, September 2009

[5] Polikar R., Topalis A., Green D., Kounios J., Clark C.M., Comparative multiresolution analysis and ensemble of classifiers approach for early diagnosis of Alzheimer's disease, Computers in Biology and Medicine vol. 37, no. 4, pp. 542-558, 2007.

[6] R. Polikar, A.Topalis, D. Parikh, D. Green, J. Frymiare, J. Kounios, C. Clark, "An ensemble based data fusion approach for early diagnosis of Alzheimer's disease", Information Fusion vol. 9, pp 83-95, 2008.

[7] C. Jarque and A. Bera, "A test for normality of observations and regression residuals." *International Statistical Review*. vol. 55, no. 2, pp. 163–172, 1987.

[8] M. Sabeti, S. Katebi, R. Boostani, "Entropy and complexity measures for EEG signal classification of schizophrenic and control participants", *Artificial Intelligence in Medicine*, volume 47, no. 3, pp. 263-274, November 2009.

[9] T. Higuchi, "Approach to an irregular time series on the basis of the fractal theory", *Physica,* vol. 31, pp. 277–283, 1988

[10] C. Gómez and R. Hornero, "Entropy and Complexity Analyses in Alzheimer's Disease: An MEG Study", *The Open Biomedical Engineering Journal*,vol. 4,pg 223-235, 2010

[11] T. Sergios and K. Koutroumbas. *Pattern Recognition*. 4th ed. Amsterdam: Academic Press, 2009.