Normality Evaluation of EEG Signals Based on Amplitude Level and Entropy Values

Hong Lin
Department of Computer Science & Engineering Technology, University of Houston-Downtown, Houston, Texas, USA
Email: linh@uhd.edu

V. Rajinikanth
Department of Electronics and Instrumentation Engineering, St. Joseph’s College of Engineering, Chennai, Tamilnadu, India
Email: rajinikanthv@stjosephs.ac.in

Abstract—Investigating the mental fitness is a key research field in medical domain because of its significance. The state and the activity of brain can be examined based on the brain signals and the brain images. The existing brain imaging procedures are costlier than the brain signal evaluating approach. Hence, during the preliminary examination, the condition of the brain is normally evaluated based on its ElectroEncephaloGram (EEG) recorded through dedicated electrodes. In medical discipline, EEG is widely considered to record the Electrical Activity (EA) of brain. After recording the EEG, its pattern is to be examined physically or by means of an automated procedure to categorize the mental health of patient into normal and abnormal condition. The work proposed in this paper considers a benchmark EEG dataset to classify the signal into normal and abnormal. During the experimental investigation, the amplitude and the entropy based features are extracted for the benchmark EEG signals and the classification is implemented using the random forest (RF) classifier. The experimental investigation of this paper confirms that, the amplitude and entropy based evaluation offers better result on the benchmark dataset.

Index Terms—brain state, EEG signal, amplitude level, entropy value, classification

I. INTRODUCTION

In recent years, a substantial amount of research works are proposed by the researchers to examine the brain activity based on the brain images recorded with various imaging schemes [1]-[4] and the single and multi-channel brain signals [5]-[9]. The imaging procedures, such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT) and Position Emission Tomography (PET) are normally suggested by the doctors to record the internals parts of the brain [10]. These imaging procedures are performed in a controlled environment (magnetically secured space), the imaging tools are bulky and costly. The CT imaging techniques applies radiation and the PET requires the injection of radio-pharmaceutical material into the human body in order to have a clear image. The main advantage of these methods is, visual examinations of the brain regions are possible. However, analysing the healthiness/activity of brain based on its signal is quite simple and it involves only external/wearable electrodes which is cheaper than the imaging techniques. Therefore, most of the initial brain screening practice is performed only with the Brain Signals (BS) recorded using the external scalp electrodes. The BS is mainly due to the electrical activity of the brain and the recorded signal is clinically known as ElectroEncephaloGram (EEG). It is a non-invasive technique and will not require induced radiation or the radioactive chemicals. The main advantage of BS is, based on the requirement as well as the abnormality is to be examined, a single-channel or a multi-channel signals can be recorded using the electrodes, which further can be evaluated by the doctor/computerised system to locate the nature and place of abnormality [11].

This paper focuses mainly on the examination of BS recoded with single and multi-channel electrodes. Earlier works on the BS examination procedure confirmed that, dementia, Alzheimer's disease, sleep disorder, epilepsy, mental stress level and meditation phase can be effectively examined [1], [5].

The work of Palani et al. [9] presents a detailed examination of the EEG dataset with the help of Fuzzy Entropy (FE). The other work of Palani et al. [11] also confirms the measurability of the epilepsy based on a hybrid mapping of the EEG signal and MRI image. The recent work of Lin and Li [12] implemented an experimental investigation to measure the meditation level of a meditator based on the multi-channel EEG signal recorded with the wireless electrode array known as Emotiv EPOC headset. This test was performed on a well-trained meditator under a controlled and calm environment. Similar work by Lin et al. [13] proposed an experiment to justify the measurability of the meditation level. During this experimental investigation, the multi-channel EEGs of a volunteer is recorded during the idle, speaking and the meditation stages. Later, the entropy assisted examination and classification is implemented to classify the considered EEG signals in to three categories, such as idle, speaking and meditation. This experiment also confirms that, the classification based on the
Random Forest (RF) approach outperform the other classifiers adopted in their study and offered a classification accuracy > 99.92%.

This paper proposes a computation method based on the EEG signal’s peak-to-peak (PP) value and the entropy function. The amplitude of the EEG signal depends mainly on the activity of the person and also its abnormality. The literature confirms that, the amplitude level of the electrical potential of the abnormal brain (brain associated with stroke/seizure) is normally high compared with the normal brain. During the meditation process, the meditator will be in relaxed state, which will diminish the electrical activity of the brain. Hence the amplitude level of the EEG signal during this process will be lesser than the normal and the abnormal conditions. By simply measuring the PP value of the amplitude signals, it is possible to predict the activity of the brain.

The main task considered in this paper is measuring the brain activity during the meditation process. In order to obtain the EEG dataset, an experienced meditator is chosen as the volunteer. During this experimental work, the EEG pattern is recorded in deep meditation state, at rest and during a conversation. All the three activities are separately recorded using a personal computer and the signal patterns are then scrutinized after implementing the possible signal pre-processing. This examination reveals that, the EEG recorded during the meditation is very smooth compared with the EEGs recorded during the conversation and idle condition. From this data, it is clear that, if we can measure the amplitude and the frequency level of these EEGs, it is possible to build soft-computing models which can classify the EEG dataset into idle, talking and meditation. The work of Lin et al. [13] also revealed that, based on the activities, the amplitude and frequency level of the EEG changes significantly. If the pattern of the EEG is measured, then the condition/state of the brain can be easily predicted.

During the experimental investigation, initially, the bench mark EEG dataset with the normal and abnormal (focal) signal pattern is considered. Later, the EEG database used in Lin & Li [12] is adopted to examine the formation in EEG pattern during normal (idle), abnormal (discuss, and inexperienced meditation) and relaxed (meditation 1, and meditation 2) conditions. EEG signal feature extraction is performed with the help of entropy value and the PP amplitude level and categorization is implemented with the Random Forest (RF). The superiority of the RF classifier is confirmed against the Nearest Neighbour (NN) and Decision Trees (DT) existing in the literature.

II. METHODOLOGY

This section of the paper provides the details regarding the database considered in the research, pre-processing scheme, feature extraction procedure and classification methods implemented to examine the single and multi-channel EEG signal.

A. EEG Database

The main aim of this paper is to propose an approach base on the electrical potential and the entropy of the EEG signal. To implement the task, this work considers two EEG databases as follows; (i) Bern-Barcelona EEG database; a benchmark EEG dataset consist of the normal and abnormal (Focal) EEG signals (https://www.upf.edu/web/mdm-dticl/-/lst-test-dataset?inheritRedirect=true#.WnPuFK6WZ0w) and (ii) Multi-channel EEG signal recorded during the meditation process with the help of Emotiv EPOC (http://grid.ahd.edu).

The Bern-Barcelona is already considered by the researchers to test their automated/semi-automated computer assisted feature extraction and classification procedures [6, 8]. The Multi-channel EEG is a new class, recently recorded to support the brain wave modelling experiments. This database is created with a low cost personal brain-computer interface unit known as the EPOC wearable sensor headset; developed by Emotiv Corporation (EMOTIV). EPOC is designed to give a 14 channel EEG and is constructed with 14 high quality primary electrodes and 2 reference electrodes. The primary are labelled as AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4 and the secondary are marked as P3 and P4. This EPOC tool is also supplied with some fundamental software tools for signal pre-processing, training, testing and classification of EEG signals [14]. This database is created with the volunteers and is separately stored as various classes as idle, talking, inexperienced meditation, deep meditation, etc.

B. Signal Preprocessing

The signal processing is considered only for the multi-channel EPOC data. During the EEG recording process, initial electrical signal coming out of the electrode needs pre-processing steps to smooth the measured variables by eliminating unwanted noise. Normally, the pre-processing task involves in discovery and elimination of outliers, missing information and EEG with low Contact Quality (CQ) value. Firstly, it was guessed that values with low CQ would result in poor data quality which would need to be disconnected earlier to model evaluation. Based on the quality, the range for the CQ is fixed from 0 to 4 and the electrodes signal with low CQ (ie. CQ < 2) is removed from the dataset. This procedure is implemented in all the three (normal, abnormal and meditation) cases. The pre-processing is implemented based on the assumption that, lower CQ values would represent more spread in the data and outliers than the data with higher CQ. Also, the CQ values did not relate to data quality, it is closely related with the outliers, which is to be identified and isolated. The pre-processed data is then considered to build the classifier unit.

C. Signal Feature Extraction

In the literature, a number of procedures are considered and implemented by the researchers to extract the vital information from the single channel/multi-channel EEG signals [5]-[9]. The pattern of the EEG signal is very complex and random in nature and requires a well-defined feature extraction method. In the proposed work, feature extraction is implemented based on its entropy function and the Peak – to Peak (PP) value of its electrode potential.
Generally, the concept of entropy is adopted to investigate the uncertainty existing in the test signals [15]. In EEG, the entropy value can be used to detect the level of chaos in the signal based on the non-linear measure to quantifying the degree of complexity. In this work, Sample entropy (SE) and Fuzzy entropy (FE) are considered to evaluate the chaos [16], [17].

1) Sample entropy:

Let there is M-point processed time series value of \( [t(i):1 \leq i \leq M] \) with zero mean and unity standard deviation, the vector string will be as follows.

\[
S^n_i = \{ n(i), n(i+1), ..., n(i+n-1) \} \quad \text{for} \quad 1 \leq i \leq M-n+1
\]

where \( S^n_i \) denotes n repeated t values beginning with the \( i^{th} \) point with n embedding dimensions. In a given signal, distance \( d^n_q \) between points \( S^n_i \) and \( S^n_j \) is defined as:

\[
d^n_q = d(S^n_i, S^n_j) = \max_{k=i+1}^{n-1} \left| t(i+k) - t(j+k) \right|
\]

Then, the computed SE for this time series (for the limit \( M \to \infty \)) is denoted as:

\[
\text{SampleEntropy}(n, r, M) = -\ln \left( \frac{A^n_r}{B^n_r} \right)
\]

where, \( A^n_r \) and \( B^n_r \) are the probability values. Other discussion on SE can be found in [16], [17].

2) Fuzzy entropy:

In real world, limitations among classes may be vague, and it is hard to decide whether an input pattern belongs entirely to a given class. The idea of fuzzy sets supplies a means of distinguishing such input–output relations in setting of vagueness. The discussion on FE can be found in [16], [17].

The FE estimation of random signal is defined as:

\[
\text{FuzzyEntropy}(n, r) = \lim_{M \to \infty} (\ln B^n_r - \ln A^n_r)
\]

\[
\text{FuzzyEntropy}(n, r, M) = -\ln \left( \frac{A^n_r}{B^n_r} \right)
\]

The Peak to Peak amplitude level of the EEG is also computed based on the peak detection procedure. The vital features such as the PP value and the mean of the PP values are then computed to train the classifier unit.

D. Classification

Classification system is used to develop an automated assessment of considered EEG. Firstly, the classifier is to be trained based on the constraint and later, it is to be tested to authenticate its effectiveness. In the proposed work, classifier system is used to distinguish the meditation EEG dataset in to normal and focal for the work, classifier system is used to distinguish the tested to authenticate its effectiveness. In the proposed be trained based on the constraint and later, it is to be assessment of considered EEG. Firstly, the classifier is to

1) Nearest Neighbor (NN)

In image processing, the K-nearest neighbor (KNN) procedure is largely is widely adopted to solve the classification problems [18]. The NN classifier calculates the dataset’s class based on the K-training trials with respect to nearest neighbors to the test sample, and relates it to a group which has the principal category probability.

2) Decision trees

Decision Tree (DT) approach considers a tree like structure with a series of test questions. DT approach considers the attribute test conditions as the root and internal nodes and the class label (Yes/No) form the terminal node. Once the DT structure has been constructed, classification is achieved easily based on the decision taken in the each branch of the tree. More details regarding the EEG classification based on the DT can be found in the following articles [19, 20].

3) Random forest

In recent years, Random Forest (RF) technique is widely considered to classify the complex datasets. RF was initially proposed by Breiman [21] and its detailed justification can be found in the paper by Chen et al. [22].

Let us consider, \( T \) is the training set with \( (x_1, y_1), (x_2, y_2), \ldots (x_n, y_n) \). \( N_{tree} \) is the number of tree to be built, \( M_p\) is the number of variables chosen for splitting at each node, then the classification is performed based on majority vote among the \( N_{tree} \). Detailed theory on random forest can be found in [22].

E. Execution

Proposed technique is implemented as shown in Fig. 1. Initially, the EEG signal to be examined is collected from the given datasets and pre-processed if necessary. Later, a feature extraction procedure is implemented to extract the signal features, such as the sample entropy, fuzzy entropy and the PP amplitude level, to train and test the signal classifier unit. After exacting the classification procedure, finally the performance of the proposed technique (entropy and amplitude supported classification) is validated against the Support Vector Machine (SVM) based classifier existing in the literature [8].

```plaintext
Consider the EEG signal to be examined (Single channel/ Multi channel)
Implement the pre-processing procedure (Locate and eliminate outliers, missing and low CQ data)
Extract the signal feature, such as the entropy value and the amplitude value.
Train and test the classifier system based on the extracted data.
Validate the performance of the proposed system.
Stop.
```

Figure 1. Flow diagram of the proposed EEG examination scheme.
III. EXPERIMENTAL RESULTS AND DISCUSSIONS

This division of the paper presents the outcome achieved with the experimentation and its estimation. This division also presents the particulars regarding the EEG database considered, features extraction task and the classification procedures.

Initially, the benchmark Bern-Barcelona EEG database (BBED) is considered for the evaluation. This database consists of normal and abnormal EEG signals recorded using dedicated electrodes. Fig. 2 & 3 depicts the sample EEG test signals of Normal and Focal EEGs. Fig. 2 presents the two normal sample test signals and Fig. 3 shows the focal test signals. From the visual assessment of Fig. 2 & 3, one can detect that, the amplitude level of the normal signal is comparatively lesser than the focal signal.

The feature extraction practice is implemented on the considered dataset with the Sample entropy and Fuzzy entropy and the corresponding outcomes are presented in Table I. This table also provides the maximum PP voltage (μV) of an individual peak and valley and also the mean amplitude level of the individual signal (Normal1, Normal2, Focal1 and Focal2) and also the overall average value of the normal EEG signal and the average of the focal EEG signal. The result of this table also confirms that, there is a considerable variation in the normal and the focal EEG signals. The extracted features are then considered to train and test the Random Forest (RF) classifier systems adopted in this study.

Superiority of the proposed technique is confirmed against the research work of Palani and Parvathavarthini [8]. Their work implemented an S-transform assisted feature mining and Support Vector Machine (SVM) based classification. This work also implemented three SVM classifier units and the performance of SVMs are classified with the well-known signal similarity measures, such as accuracy, sensitivity, specificity, positive predictive value and negative predictive value. The results presented in Table II confirms that, proposed approach offers better signal similarity measures compared with the SVMs based results. The graphical representation of the result is also depicted in Fig. 4. From this result, it can be confirmed that, proposed approach outperforms the previous result existing for the BBED.

<table>
<thead>
<tr>
<th>EEG data</th>
<th>Fuzzy entropy</th>
<th>Sample entropy</th>
<th>Maximum PP voltage (μV)</th>
<th>Mean PP voltage (μV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal1</td>
<td>0.021±0.0007</td>
<td>0.023±0.0042</td>
<td>203</td>
<td>118 ± 0.5264</td>
</tr>
<tr>
<td>Normal2</td>
<td>0.032±0.0001</td>
<td>0.014±0.0073</td>
<td>384</td>
<td>78 ± 0.1432</td>
</tr>
<tr>
<td>Averages</td>
<td>0.018±0.0133</td>
<td>0.017±0.0153</td>
<td>238 ± 0.8753</td>
<td>86.962 ± 0.5462</td>
</tr>
<tr>
<td>Focal1</td>
<td>0.025±0.0056</td>
<td>0.034±0.0032</td>
<td>228</td>
<td>266 ± 3.5852</td>
</tr>
<tr>
<td>Focal2</td>
<td>0.004±0.0011</td>
<td>0.060±0.0046</td>
<td>374</td>
<td>221 ± 0.7440</td>
</tr>
<tr>
<td>Averages</td>
<td>0.008±0.012</td>
<td>0.053±0.0184</td>
<td>216 ± 0.2985</td>
<td>218 ± 0.2274</td>
</tr>
</tbody>
</table>

Figure 2. Pattern of the normal test signal of the benchmark database.

Figure 3. Pattern of the focal EEG test signal of the benchmark database.

Table II. Experimental appraisal of proposed technique with existing approaches

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Accuracy (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Positive predictive value (%)</th>
<th>Negative predictive value (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear [8]</td>
<td>76</td>
<td>65</td>
<td>64</td>
<td>74</td>
<td>68</td>
</tr>
<tr>
<td>Polynomial [8]</td>
<td>82</td>
<td>66</td>
<td>67</td>
<td>76</td>
<td>74</td>
</tr>
<tr>
<td>RBF [9]</td>
<td>86</td>
<td>72</td>
<td>71</td>
<td>80</td>
<td>81</td>
</tr>
<tr>
<td>Random Forest</td>
<td>94.27</td>
<td>85.93</td>
<td>81.35</td>
<td>85.72</td>
<td>88.69</td>
</tr>
</tbody>
</table>

Figure 4. Graphical representation of performance measure values.
IV CONCLUSIONS

This paper proposes an approach to classify the EEG signal into normal and abnormal class. This work initially implements a feature extraction procedure with the fuzzy entropy, sample entropy and the Peak-to-Peak amplitude recognition procedure. The extracted features, such as entropy and PP amplitude of electrode potential are then considered to train and test the classifier system considered in this practice. During the experimental investigation process, two types of EEG datasets, like BBED and EPOC are considered. Initial investigation is implemented on the BBED dataset and the classification is performed using the Random Forest (RF) classifier unit. The result of RF is then compared and verified against the SVM based result existing in the literature. The comparative study, confirms that, for the BBED database, RF offered superior result than the SVM. This experimental result also confirms that, the brain activity of a person can be easily detected by measuring its electrode potential. The Electrode potential will be moderate during normal behavior and very high during abnormal condition.

ACKNOWLEDGMENT

This work is partially supported by Department of Defense – DOD instrumentation grant “High Performance Computing for Faculty and Students at University of Houston-Downtown”, the National Computational Science Institute Blue Waters Student Internship program, the Integrating High-Impact Practices at UHD grant from the Center for Teaching and Learning Excellence, and the Service Learning project from the Center for Community Engagement and Service Learning, the University of Houston-Downtown.

REFERENCES


Hong Lin received his PhD in Computer Science and was a postdoctoral research associate at Purdue University, an assistant research officer at the National Research Council, Canada, and a software engineer at Nokia. Inc. Dr. Lin is currently a professor with UHD. His research interests include parallel/distributed computing, data analytics, and human-centered computing. He is a senior member of the ACM.

V. Rajnikanth earned his Ph.D. degree in Electrical Engineering from Anna University, Chennai, India. Currently, a Professor in the Department of EEE, St. Joseph’s College of Engineering, Chennai-600119, India. He has published 55 papers in several international conferences and journals. Currently, he is an associate editor of International Journal of Rough Sets and Data Analysis (IJRSDA), IGI Global.