EEG-Based Screening Tool for School Student's Brain Disorders Using Machine Learning Algorithms

Abdelrahman A. Ramzy, Bassel S. Abdallah, Mohamed E. Bahgat, Sarah M. Abdelkader, Sherif H. ElGohary

Abstract—Attention-Deficit/Hyperactivity Disorder (ADHD), epilepsy, and autism affect millions of children worldwide, many of which are undiagnosed despite the fact that all of these disorders are detectable in early childhood. Late diagnosis can cause severe problems due to the late treatment and to the misconceptions and lack of awareness as a whole towards these disorders. Moreover, electroencephalography (EEG) has played a vital role in the assessment of neural function in children. Therefore, quantitative EEG measurement will be utilized as a tool for use in the evaluation of patients who may have ADHD, epilepsy, and autism. We propose a screening tool that uses EEG signals and machine learning algorithms to detect these disorders at an early age in an automated manner. The proposed classifiers used with epilepsy as a step taken for the work done so far, provided an accuracy of approximately 97% using SVM, Naïve Bayes and Decision tree, while 98% using KNN, which gives hope for the work yet to be conducted.

Keywords—ADHD, autism, epilepsy, EEG, SVM.

I. INTRODUCTION

AUTISM is a lifelong developmental disorder that affects how children interact with other people and causes communication problems, ADHD is a behavioral disorder that causes inattentiveness, hyperactivity and impulsiveness and epilepsy is a condition that causes frequent electrical surges in the brain that affects how it functions [1]. The common ground among these three disorders is that they all affect children, and ADHD and autism manifest at a young age. It is estimated by the United Nations International Children's Emergency Fund (UNICEF) that we currently have 2.2 billion children worldwide, and that one in every 160 children from that figure has autism, according to the World Health Organization (WHO), while 7.2% are estimated to have ADHD and more than 50 million people worldwide have epilepsy with an estimate that 70% of them could live epileptic-free if they are early and properly diagnosed [2]-[4].

Due to the lack of awareness and the misconceptions regarding all three, a lot of children are misdiagnosed with autism, while a lot more go undiagnosed due to the lack of a definitive diagnostic method for ADHD and autism, while children with epilepsy face discrimination among their peers.

ADHD and autism are mainly diagnosed by an assessment based on monitoring the child over a period of weeks and putting him/her in some situations or games to see how the child responds, and by using questionnaires answered by the

A. A. Ramzy, B. S. Abdallah, M. E. Bahgat, S. M. Abdelkader, and S, H. ElGohary are with the Biomedical Engineering Department, Cairo University, Giza 12613, Egypt (e-mail: aaramzy95@gmail.com, basselsamer9556@gmail.com, m.essam201435@gmail.com, sarahhamed2495@gmail.com, sh.elgohary@eng1.cu.edu.eg).

parents in the case of Autism and by both the parents and the teachers in the case of ADHD [4]. Due to the observation-based assessment, the responsibility of a proper diagnosis falls on both the parents and the teachers to notice that the child is behaving in an abnormal manner compared with children his/her age, this makes the diagnosis in Egypt difficult, as a diagnosis needs a certain level of awareness and that is a drawback here in this country due to the high level of illiteracy, as mentioned by the Social Solidarity Ministry of Egypt in 2017.

Epilepsy is diagnosed using EEG signals to detect the brain activity of the patient. The most commonly known symptoms are sudden muscle jerking and loss of consciousness. However, not all epileptic patients suffer from these visible symptoms, some patients suffer from a sudden stop of the activity at hand without them knowing what happened, and when they regain conciseness it feels to them as if they had a blackout, which appears to others as being lost in thought, eyelids fluttering quickly and other subtle signs that occur on a daily basis and can happen multiple times in the same day. It needs a certain level of attentiveness and awareness to notice these symptoms and not ignore them as they are caused by what is called absence seizures, which is a type of epilepsy, and they are often noticed in children between the ages of 8-14 years and they can affect concentration and school performance which is why early diagnosis is critical [5].

The proposed solution is to diagnose these disorders using EEG signals and machine learning algorithms to find unique detectable patterns that can be used for the diagnosis in an automated manner that negates the human factor in the diagnostic process. It is divided into two stages, the first stage it will be used as a screening tool to see whether the child is healthy or not and if he/she has any of the disorders from the EEG signals. In the case of either ADHD or autism being detected, we will go to stage two. This stage includes doing an fMRI scan, as the images will help to localize the areas of activity and thus give a definitive diagnosis. After the initial stage when the results of both the EEG and the fMRI are concise, we will eliminate the fMRI from the process and solely depend on the EEG signals.

Due to autism and ADHD not being diagnosed using EEG signals, we started to focus on autism as an initial step and found that it affects areas in the frontal and temporal lobes most notably the amygdala, which gives us an initial area of interest and where the abnormalities might originate from [6]. Previous studies have also started to investigate the relation between EEG signals and autism, including a study published in 2018 by Bosl et al. and another study by Grossi et al., in

World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:13, No:6, 2019

which both were successful in detecting autism by using EEG signals [7], [8].

This paper discusses the work done so far in the process of finding a product able to detect all of the three above mentioned disorders to help children be correctly treated and improve their overall lives and educational efficiency.

II. MATERIALS AND METHODS

A. EEG

fMRI images have been used to diagnose and detect brain disorders such as Autism and ADHD, but we found that EEG signals are easier and cheaper. EEG is a procedure that consists of tracking and recording brain wave patterns through electrodes that are attached to the scalp. EEGs are wavy lines with peaks and valleys that allow a quick assessment of whether there are any irregularities that may be a sign for either of ADHD, autism or epilepsy.

EEG signals consist of five bands as shown in Table I [9]. Each band changes according to the mental state of the patient; so, we used this property as changing events in the EEG sessions to help in detecting any irregularities.

A sample showing an example of the difference between normal and abnormal EEG signals can be seen in Fig. 1.

TABLE I
EEG FREQUENCY BANDS IN HZ

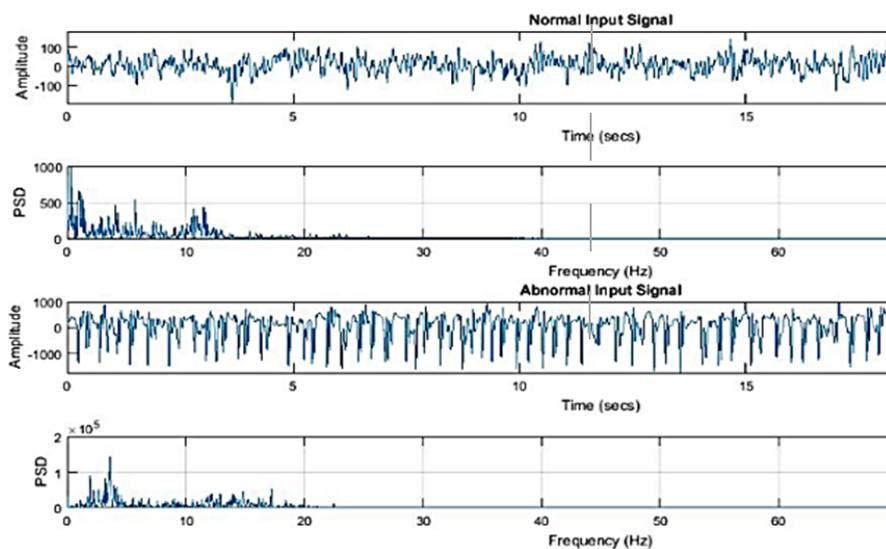| Bands	GAMMA	BETA	ALPHA	THETA	DELTA
FREQUENCY	30<	14-30	8-14	4-8	< 4

Fig. 1 Top: Normal EEG, Bottom: Abnormal EEG

B. EEG Device

We chose EPOC+ (Emotive Inc.), which is a wireless device consisting of 14 electrodes that do not need any gel to be applied on the scalp. These specifications are well suited for the needs of this study, as being wireless and gel-free is less scary for the children and gives them the ability to move freely during the session, while the electrodes cover the areas of interest in the brain.

C. EEG Datasets

The EEG datasets used for these results were taken from Bonn University [10], and consists of five sets classified as:
- Set A and Set B contained the normal EEG readings.
- Set C and Set D contained the interictal readings, which represent the period between seizures [11].
- Set E contained the ictal readings, which represent the period of a seizure [11].

Each set consisted of 100 txt files, as well as readings from five patients taken from 100 single channels of an EEG segment each with duration of 23.6 sec. The sampling rate of the data was 173.61 Hz and a 40 Hz low-pass filter was applied on the data. The data were divided into 70% as training data and 30% as testing data. Moreover, we applied various classifiers with other datasets acquired from Temple University Hospital [12]. It consists of two sets classified as:
- Normal signals recorded from 104 patients.
- Epileptic signals recorded from 133 patients.

The software systems used in the analysis of the data were:
- MATLAB (Release 2016a, The MathWorks, Inc., Natick, Massachusetts, United States).
- Python (Python Software Foundation. Python Language Reference, version 3.7.2).

D. Feature Extraction

The features used in the classification process are shown in Table II.

E. Classification

For the classification process we used:
- k-Nearest Neighbor (k-NN) which depends on the distance between points. We identified a testing point's class according to the majority class of the k-nearest training points, and we used Euclidean distance and chose K=5 as a result of an iteration process.

World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:13, No:6, 2019

TABLE II
TIME DOMAIN FEATURES USED IN THE CLASSIFICATION

Feature	Equation
Mean (μ)	$\frac{1}{N} \sum Xi$
Standard Deviation (σ)	$\sqrt{\frac{1}{N-1} \sum (Xi - \mu)^2}$
Variance	σ^2
Median (Odd)	X at $\left(\frac{N+1}{2}\right)$
Median (Even)	$\frac{1}{2} * \left(X \text{ at } \left(\frac{N}{2}\right) + X \text{ at } \left(\frac{N}{2} + 1\right) \right)$
Kurtosis	$\frac{\sum (Xi - \mu)^4}{\sigma^4 * N} - 3$
Energy	$\sum |Xi|^2$
Entropy	$\sum \log(Xi^2)$
Skewness	$\frac{\sum (Xi - \mu)^3}{\sigma^3 * N}$

- Support Vector Machine (SVM) which differentiates between classes by uses non-linear boundaries that are constructed using hyperplanes. It works by finding the optimum hyperplane, which is the one that puts the maximum distance between the nearest two points from both classes.
- Naïve Bayes (NB): It is a probabilistic classification method based on Bayes' theorem, which assumes feature independency. The NB model uses a maximum probability algorithm to determine the class of earlier probabilities and a features probability distribution from a training dataset, then maximized posteriori decision tree to find the specific class label for a new test instance.
- Decision Tree Classifier: It is based on the division of a complex problem into several sub problems. Each leaf node receives a class label, non-terminal nodes, which include the root node and other internal nodes, contain attribute testing conditions to separate records that have different characteristics.

After that, we used a confusion matrix in each case to know which one was more accurate.

F. Visualization

As the proposed system is considered, initially, to be a screening tool using EEG signals, it was found that to provide a better diagnosis that it is better to visualize the brain activity.

- We used the EEGLAB toolbox [13] on MATLAB as a first step in determining the areas of the brain that are active and visualize the EEG signals acquired to be able to minimize the channels used, whether with the autism-related datasets, epileptic datasets or the ADHD datasets [13]. We plotted 3D event-related potential maps (ERP) as shown in Fig. 2.
- We also used the FMRLAB toolbox on MATLAB [14] as an initial step for the diagnosis to determine the affected areas of the brain and visualize BOLD images to find the active centers of the brain, as shown in Fig. 3.

We started to visualize the EEG signals, to see exactly where abnormalities are occurred. Moreover, the overall brain activities will be visualized during the region of interest.

G. Database

In Egypt, there are no accurate statistics or any EEG datasets available for epilepsy, autism or ADHD. Thus, as part of this study, we started to create a database for children that undergo the proposed screening system.

We used MYSQL v8.0 (MYSQL, Inc.) in building our database. The ordinary method used for the diagnosis of brain disorders such as autism and ADHD is by screening the child and checking a questionnaire or a check list like the Modified Checklist for Autism in Toddlers, Revised, with Follow-Up (M-CHAT-R/F) and the National Institute for Children's Health Quality Vanderbilt scale. We also added the scoring result of each questionnaire in the database to use it as a part of the classification process and compare the results.

The overall system is shown in the block diagram in Fig. 4.
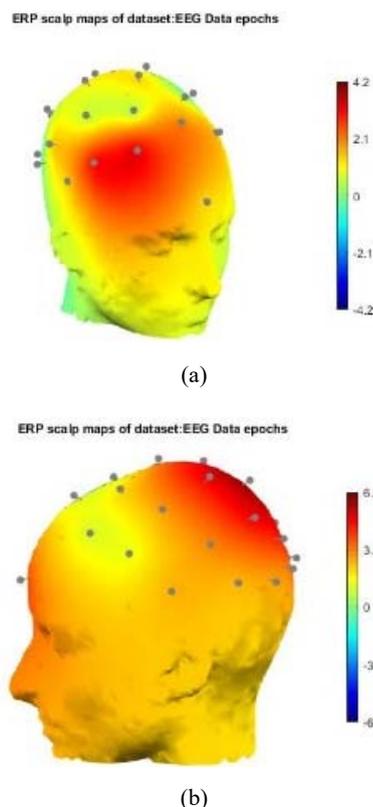

Fig. 2 (a) 3D ERP at 1000 ms, (b) 3D ERP at 1992 ms
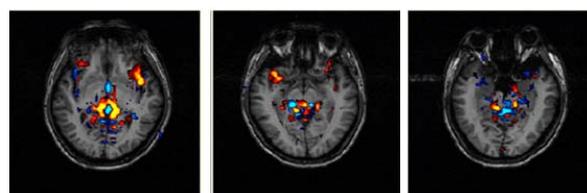

Fig. 3 BOLD image at different components

III. RESULTS

In our attempt to find a portable easy-to-use screening tool to detect any abnormalities in EEG signals to diagnose autism, ADHD and epilepsy, we started by applying classifiers on epileptic EEG signals and by using visualization toolboxes represented in EEGLAB and FMRLAB as a first step for both

World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:13, No:6, 2019

autism and ADHD.

The results of each of the previously mentioned classifiers are shown in Fig. 5 and a sample of one of the EEG sessions that was recorded can be seen in Fig. 6.

The software developed as part of the proposed system can be divided into two integrated user interfaces (UIs); the first UI is intended to be used by the doctor to view access the database and review the patient's records, EEG signals and analysis, while the second UI is utilized during recording the EEG signals. The whole screening process will be monitored by the physician and/or trained personnel conducting the session alongside the child, to simulate interested regions of the child's brain using built-in activities and games. Validating the results with a questionnaire that given for the parents to answer regarding their child's behavior. The UI for child can be seen in Fig. 7.

IV. DISCUSSION

The proposed system is a screening and a diagnostic tool that uses machine learning algorithms to differentiate between normal and abnormal EEG signals. As a first stage, we used both SVM and KNN classifiers to seek the best possible results in discriminating between normal and epileptic signals. KNN produced a higher accuracy which makes it more efficient in the process at hand, while the SVM classifier was also very efficient as the difference between the two is not that

drastic. We used both EEGLAB and FMRLAB to locate and visualize the active parts in the brain during specific events and stimulations. EEGLAB was also very efficient in plotting the signals and showing the brain activity, which makes it a very valuable tool to narrow down our channels of interest.
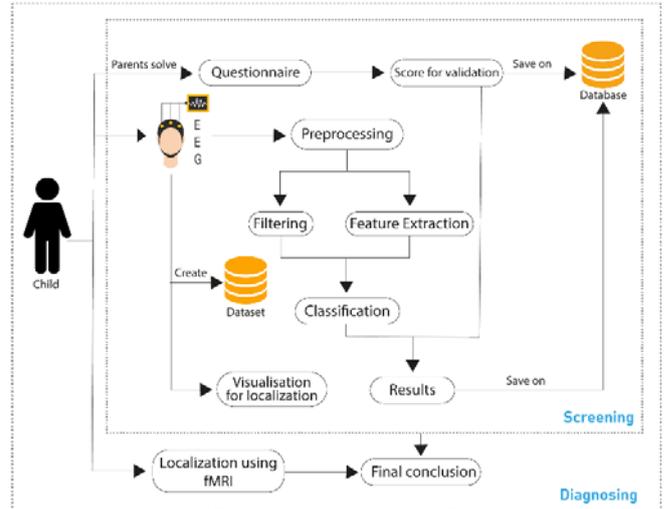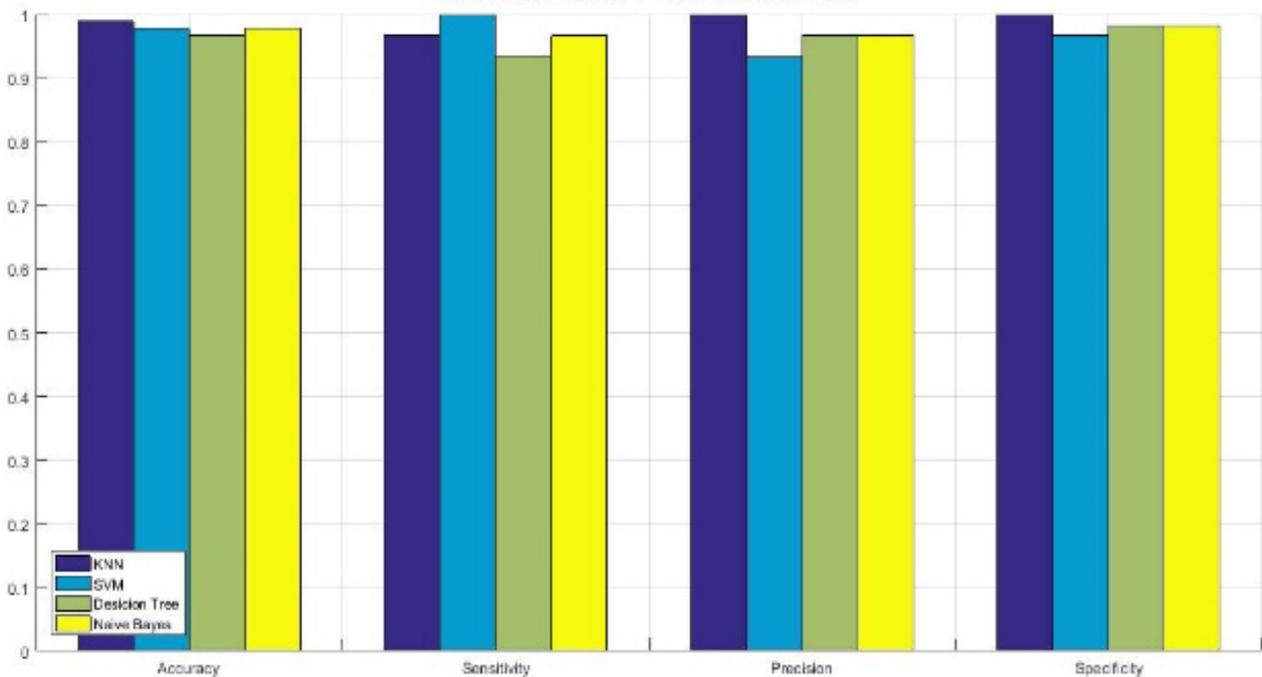

Fig. 4 Block Diagram of the proposed system

Fig. 5 Accuracy, Sensitivity, Precision and Specificity for KNN, SVM, Decision Tree and Naïve Bayes

World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
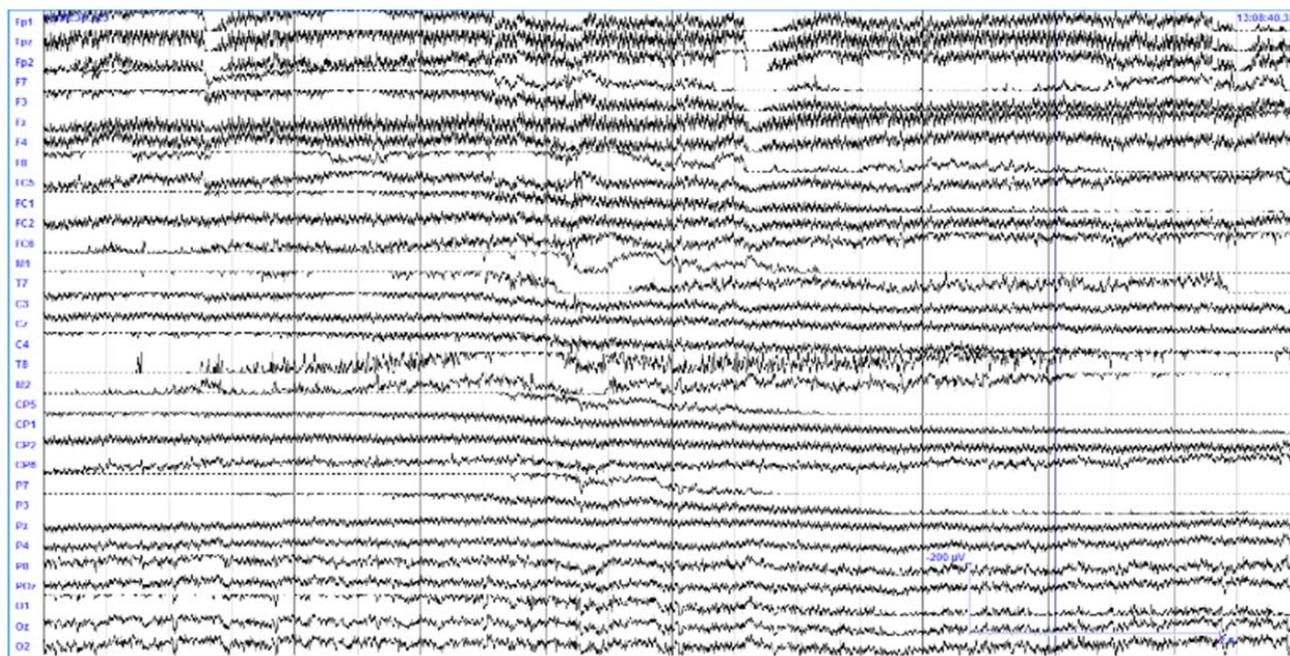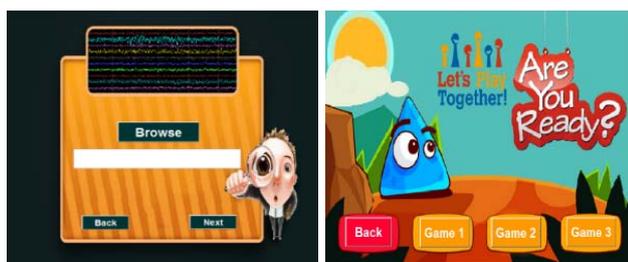Vol:13, No:6, 2019

Fig. 6 EEG Sample recorded for a normal patient

(a)

(b)

Fig. 7 (a) and (b) Implementation of the GUI done using MATLAB

V. CONCLUSION

The results reached show the accuracy in detecting epilepsy using SVM and KNN. The work to be done in the coming months will be to achieve similar results with ADHD and autism using various classifiers, as well as to provide a machine-based solution to eliminate the human factor in diagnosing any of these disorders.

ACKNOWLEDGMENT

REFERENCES

[1] National Health Service. (Online). Available: https://www.nhs.uk/.
[2] Nations International Children's Emergency Fund (UNICEF). (Online). Available: https://www.unicef.org/sowc05/english/poverty.html.
[3] Children and Adults with Attention-Deficit/Hyperactivity Disorder (CHADD). (Online). Available: https://chadd.org/about-adhd/general-prevalence/.
[4] World Health Organization (WHO). (Online). Available: https://www.who.int/.
[5] Laura M. Guilhoto, "Absence epilepsy: Continuum of clinical presentation and epigenetics", Seizure European Journal of Epilepsy.
[6] Hye Ran Park, Jae Meen Lee, Hyo Eun Moon, Dong Soo Lee, Bung-Nyun Kim, Jinhyun Kim, Dong Gyu Kim & Sun Ha Paek, "A Short Review on the Current Understanding of Autism Spectrum Disorders", Experimental Neurobiology journal.
[7] William J. Bosl, Helen Tager-Flusberg & Charles A. Nelson, "EEG Analytics for Early Detection of Autism Spectrum Disorder: A data-driven approach", Nature journal of science.
[8] Enzo Grossi, Chiara Olivieri & Massimo Buscema, "Diagnosis of autism through EEG processed by advanced computational algorithms: A pilot study".
[9] Abo-Zahhad, M & Ahmed, Sabah & Seha, Sherif Nagib. (2015). "*A New EEG Acquisition Protocol for Biometric Identification Using Eye Blinking Signals*". International Journal of Intelligent Systems and Applications (IJISA). 07. 48-54. 10.5815/ijisa.2015.06.05.
[10] Ralph G. Andrzejak, Klaus Lehnertz, Florian Mormann, Christoph Rieke, Peter David & Christian E. Elger "*Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state*", Phys. Rev. E, 64, 061907.
[11] Jennifer L. DeWolfe & Beth Ann Malow, "*Approach to Sleep-Related Seizure Identification and Management*", Therapy in Sleep Medicine, 2012.
[12] Veloso, L., McHugh, J. R., von Weltin, E., Obeid, I., & Picone, J. (2017). Big Data Resources for EEGs: Enabling Deep Learning Research. In I. Obeid & J. Picone (Eds.), Proceedings of the IEEE Signal Processing in Medicine and Biology Symposium (p. 1). Philadelphia,

Pennsylvania, USA: IEEE.

[13] Arnaud Delorme & Scott Makeig, "*EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis*", Journal of Neuroscience Methods 134:9-21.

[14] Duann, Jeng-Ren, Jung, Tzyy-Ping, and Makeig, Scott "*FMRLAB: A MATLAB toolbox for functional imaging analysis*", WWW Site, Swartz Center for Computational Neurobiology, Institute for Neural Computation, University of California San Diego, sccn.ucsd.edu/fmrlab/, (World Wide Web Publication), 2002.