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 the brain that affects how it functions [1]. The common

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 the brain that affects how it functions [1]. The common

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 the brain that affects how it functions [1]. The common

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 the brain that affects how it functions [1]. The common

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 the brain that affects how it functions [1]. The common

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 the brain that affects how it functions [1]. The common
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>
<thead>
<tr>
<th>Bands</th>
<th>GAMMA</th>
<th>BETA</th>
<th>ALPHA</th>
<th>THETA</th>
<th>DELTA</th>
</tr>
</thead>
<tbody>
<tr>
<td>FREQUENCY</td>
<td>30&lt;</td>
<td>14-30</td>
<td>8-14</td>
<td>4-8</td>
<td>&lt; 4</td>
</tr>
</tbody>
</table>

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:
- 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.
TABLE II
TIME DOMAIN FEATURES USED IN THE CLASSIFICATION

<table>
<thead>
<tr>
<th>Feature</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (μ)</td>
<td>( \frac{1}{N} \sum X_i )</td>
</tr>
<tr>
<td>Standard Deviation (σ)</td>
<td>( \sqrt{\frac{1}{N-1} \sum (X_i - \mu)^2} )</td>
</tr>
<tr>
<td>Variance</td>
<td>( \sigma^2 )</td>
</tr>
<tr>
<td>Median (Odd)</td>
<td>( X ) at ( \left( \frac{N+1}{2} \right) )</td>
</tr>
<tr>
<td>Median (Even)</td>
<td>( \frac{1}{2} \left( X ) at ( \frac{N}{2} ) + ( X ) at ( \frac{N}{2} + 1 ) )</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>( \frac{1}{N^2} \frac{\sum (X_i - \mu)^4}{\sigma^4} - 3 )</td>
</tr>
<tr>
<td>Energy</td>
<td>( \frac{\sum</td>
</tr>
<tr>
<td>Entropy</td>
<td>( \sum \log(\sigma_i^2) )</td>
</tr>
<tr>
<td>Skewness</td>
<td>( \frac{\sum (X_i - \mu)^3}{\sigma^3} )</td>
</tr>
</tbody>
</table>

- 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.

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
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.
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

We would like to thank the physicians at El-Demerdash Hospital which is part of the faculty of medicine at Ain Shams University for their help in providing us with the medical validation and background knowledge needed to make sure that our research is conducted in a correct manner.

REFERENCES

Pennsylvania, USA: IEEE.
