ORIGINAL ARTICLE

Detection of ADHD Disorder in Children Using Layer-Wise Relevance Propagation and Convolutional Neural Network: An EEG Analysis

Ali Nouri * , **Zahra Tabanfar**

Department of Biomedical Engineering, Amirkabir University of Technology, Tehran, Iran

*Corresponding Author: Ali Nouri
Email: ali.nouri@aut.ac.ir

Received: 07 June 2022 / Accepted: 20 July 2022

Abstract

Purpose: Attention-Deficit-Hyperactivity-Disorder (ADHD) is a neurodevelopmental disorder that begins in early childhood and often persists into adulthood, causing personality issues and social behavior problems. Thus, detecting ADHD in its early stages and developing an effective therapy is of tremendous interest. This study presents a deep learning-based model for ADHD diagnosis in children.

Materials and Methods: The 'First-National-EEG-Data-Analysis-Competition-with-Clinical-Application' dataset is used for this purpose. Following preprocessing, data is segmented into 3-second epochs, and frequency features are extracted from these epochs. The Fourier transform is applied to each channel separately, and the resulting two-dimensional matrix (channel×frequency) for each epoch is used as the Convolutional Neural Network's (CNN) input. The CNN is made up of two convolutional layers, two max pooling layers and two fully connected layers as well as the output layer (a total of 9 layers) for classification. To improve the method's performance, the output of the classification of each input variable is analyzed. In other words, the role of each channel/frequency in the final classification is being investigated using the Layer-wise Relevance Propagation (LRP) algorithm.

Results: According to the results of the LRP algorithm, only efficient channels are employed as Convolutional Neural Network (CNN) inputs in the following stage. This method yields a final accuracy of 94.52% for validation data. In this study, the feature space is visualized, useful channels are selected, and deep structure capabilities are exploited to diagnose ADHD disorder.

Conclusion: The findings suggest that the proposed technique can be used to effectively diagnose ADHD in children.

Keywords: Attention Deficit Hyperactivity Disorder; Convolutional Neural Network; Layer-Wise Relevance Propagation Algorithm; Electroencephalogram Signal Processing.

Copyright © 2024 Tehran University of Medical Sciences.
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International license (https://creativecommons.org/licenses/by-nc/4.0/). Noncommercial uses of the work are permitted, provided the original work is properly cited.
DOI: https://doi.org/10.18502/fbt.v11i1.14507

Frontiers in BIOMEDICAL TECHNOLOGIES

1. Introduction

Attention Deficit Hyperactivity Disorder (ADHD) is one of the most common mental disorders in school-age children (with a prevalence of 3 to 5%). Children with ADHD typically show distractibility, low focus, excessive activity, poor self-control, and other symptoms, which are generally divided into 3 major dimensions: inattention, impulsivity, and hyperactivity. While ADHD is recognized as a lifelong disorder, it can result in academic failure and an unbalanced lifestyle [1-3]. Given that approximately 65% diagnosed in childhood appear to suffer from ADHD into adulthood [4], early diagnosis and development of therapy in childhood is critical.

The clinical diagnosis of ADHD is primarily based on behavioral rating scales reported by teachers, parents, or caregivers, the most common of which is the Diagnostic and Statistical Manual of Mental Disorders (DSM) criteria [1, 3]. Although these ADHD criteria have been improved over time, they are still based on an individual's thoughts. Thus, the development of neuroimaging techniques that are independent of individual thinking and effective classification models can help to distinguish ADHD from healthy controls more reliably and robustly. Various neuroimaging modalities, such as fNIRS [5, 6], fMRI [1, 7], and EEG [8, 9] have been used to detect ADHD. While EEG is an easy-to-use and non-invasive neurophysiological test without radiation exposure, it appears to be more appropriate for investigating children with cognitive or neurodevelopmental disorders [10].

Artificial Neural Networks (ANNs) are currently a promising area of Artificial Intelligence (AI) in the medical field. Deep learning, a specialized branch of machine learning, and machine learning as a whole have been used more and more in clinical research with remarkable results [11]. The use of deep neural network models today has allowed researchers to decode the signal information and patterns with simpler features (sometimes the raw signal itself), and thus achieve their goals, including the diagnosis of diseases and disorders. So far, several studies [12-16] have presented deep learning-based models in order to recognize ADHD disorders using EEG signals. However, society has yet to receive a reliable psychiatrist/psychologist assistant system for diagnosing ADHD. Providing a neuroscientifically interpretable recognition model for this disorder can be particularly valuable to the psychiatrist/psychologist in finalizing the diagnosis of ADHD disorder in children.

Aside from the importance of high decoding accuracy, some research focuses on understanding what deep neural network models learn, or providing an interpretable model. This line of research may be fascinating for neuroscientists interested in using these models because they want to know what features in the brain signal distinguish the under study classes [17]. In [18], Convolutional Neural Network (CNN) was used along with the Gradient-weighted Class Activation Mapping (Grad-CAM) visualization algorithm to decode EEG personalized spatial-frequency irregularities in ADHD children. The accuracy of 90.29 ± 0.58% as well as the ability to produce interpretable visualized results from the process of decision making demonstrate the effectiveness of the proposed model. In [19], a Convolutional Neural Network (CNN) was trained using event-related spectrograms of patients with ADHD and healthy controls during the Flanker task. Their approach achieved a classification accuracy of 88 ± 1.12 % for the identification of this disorder. They also used the DeepDream feature visualization method, the results of which showed a power reduction in the alpha band and also power increment in the delta-theta band at about 100 ms post-stimulus for ADHD participants in comparison with healthy ones.

The frequency domain of ADHD appears to provide a great deal of information, as stated in the survey of the research literature [14, 16]. In this research, a new structure of the convolutional neural network is introduced for ADHD diagnosis using frequency domain features. To this end, the frequency domain EEG signals of children with ADHD and their healthy peers are used as network inputs. The Layer-wise Relevance Propagation technique is utilized in this study to provide an interpretable model, analyze whether model decisions reflect meaningful patterns in the input, and improve the generalizability of the proposed model. The main contribution of this research is to provide a feature selection (channel selection) for better identification of ADHD disorder by visualizing the hidden information in the CNN structure. In addition to its straightforwardness, the presented method is interpretable, and it is possible to explicate the information hidden in the structure of the network with existing physiological knowledge about the ADHD disorder as well as the type of task.

2. Materials and Methods

The block diagram of the introduced method is shown in Figure 1. Each part is described in the following sections.

2.1. EEG Dataset

In this research, the dataset of the 'First National EEG Data Analysis Competition with Clinical Application' [20] was utilized which consists of 19 channels of EEG signals with a sampling frequency of 512 Hz recorded from 31 children suffering from ADHD and 30 healthy controls (boys and girls, ages 7-12). The children with ADHD were diagnosed according to DSM-IV standards by an experienced psychiatrist and took Ritalin for up to 6 months. None of the healthy group children had a history of mental disorders, seizures, or another history of high-risk behaviors. In the task, the children were shown a series of photographs of cartoon characters, and they were asked to count the characters. Between 5 and 16, the number of characters in each image was randomly chosen, and the size of the photographs was large enough for the children to be clearly recognizable and counted. Every image was shown immediately and uninterrupted after the child's response, in order to provide continuous stimulation during the signal recording. In this research, 196 thirty-second epochs belonging to the ADHD group and 132 thirty-second epochs belonging to the healthy group were available and used in the processing process.

2.2. Pre-Processing

EEG signals are generally brain activity that is contaminated with physiological and non-physiological noises and artifacts [21]. The aim of pre-processing is usually to reduce the noises and artifacts of the signal. In this study, the typical stages of pre-processing such as re-referencing (to the average of T7 and T8 channels), baseline rejection, filtering, downsampling, and signal segmentation were used. Since the signals did not have high-frequency noise, only a high-pass filter with a 1 Hz cut-off frequency and a 50 Hz notch filter (to remove line noise) were utilized. Both are of the 5-order elliptic filters and the Forward/Backward filtering method was used to compensate for the phase shift of the IIR filters (MATLAB filtfilt command). Previous research, such as [22], demonstrated that the elliptic filter is an appropriate choice for EEG signals. As a result, the filter type was determined accordingly. Based on the nature of the data and trial and error, the filters` order and cut-off frequency of the high-pass filter were set to 5 and 1 Hz, respectively. After sampling, in order to artificially increase the data, each thirty-second epoch of the signal is transformed to 10 separate three-

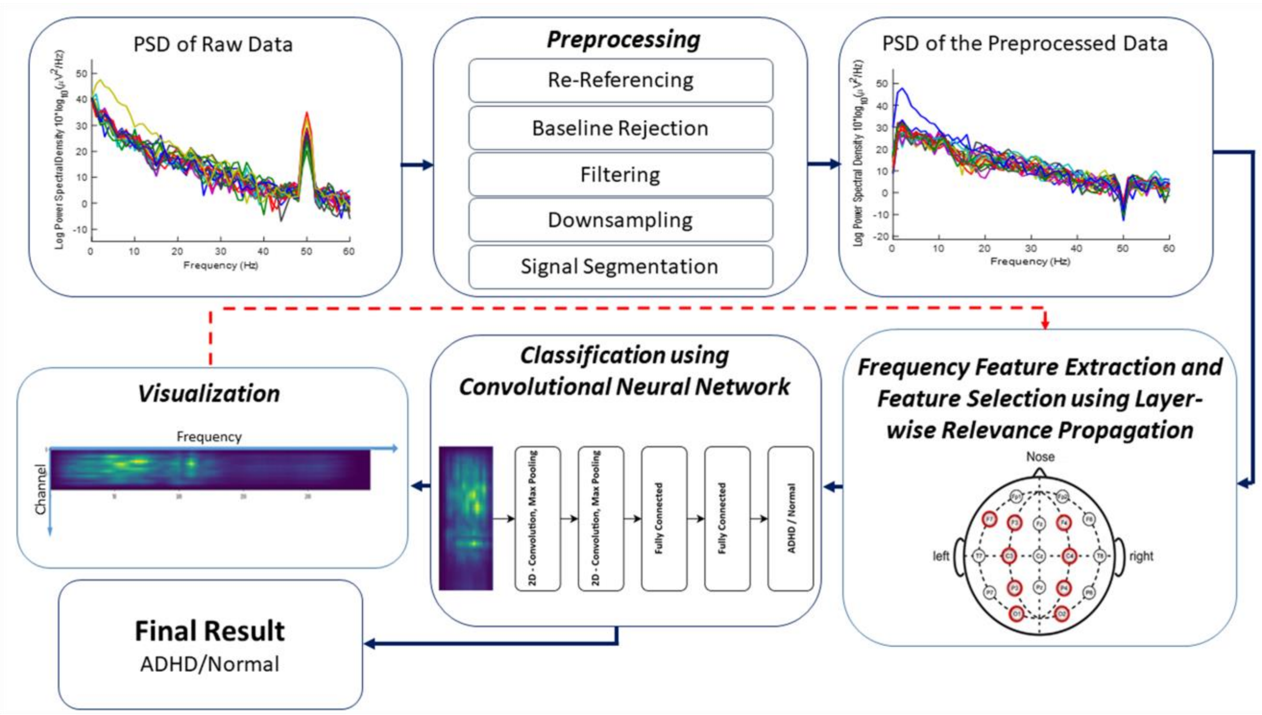


Figure 1. Block diagram of the proposed method

second segments by windowing and is considered to be independent data for further analysis. It is worth noting that in the test process, each thirty-second signal is segmented into 10 signals. The trained classifier shall be applied to each of these 10 signals and, ultimately, the maximum vote of the 10 signals is considered as the final label (ADHD/Normal) related to the whole thirty-second signal.

2.3. Feature Extraction

The features of the EEG signal in the frequency domain have been used in the present research. Fourier Transform is applied to extract the frequency domain features from the 3-second segments of the signal. The initial data consisted of 19 channels, and two channels were used for re-referencing during the pre-processing stage. Therefore, for final processing, 17 channels were used. The Fourier transform is applied to each channel separately and the results are finally stored as a two-dimensional matrix. In Figure 2, the final input domain to the neural network is shown.

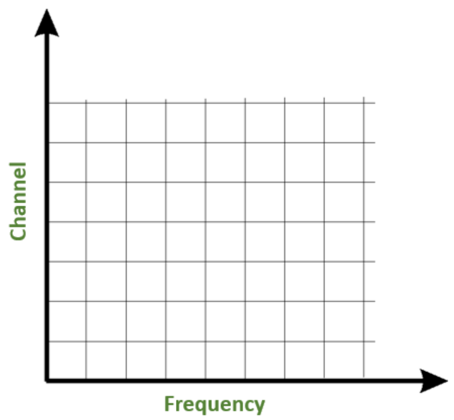


Figure 2. The final input domain to the neural network

2.4. Channel Selection Using Layer-Wise Relevance Propagation

It is important to determine whether the model decisions obey meaningful patterns in the input in order to improve the generalizability of a machine learning model. In deep neural networks, the Layer-wise Relevance Propagation algorithm is a method for obtaining this form of description. This method works by analyzing the effect of changes in network inputs on its prediction and thus by finding more effective inputs. The output of this approach is a heat map showing the role of individual input elements in the output construction.

2.5. Classification

In this research, a convolutional neural network-based classification of ADHD and healthy children using their EEG signal during a mathematical task is proposed. EEG signal processing has traditionally relied on handcrafted features and manual feature engineering. During the training process, CNNs can learn to learn appropriate features that will reduce the loss function and help it accomplish the intended goal. When data is convolved with convolutional filters in a CNN, the ultimate classification is performed using a Softmax activation function. Categorical Cross Entropy (CCE) is the network's loss function, which calculates the distance between the distribution of network output and labels [23].

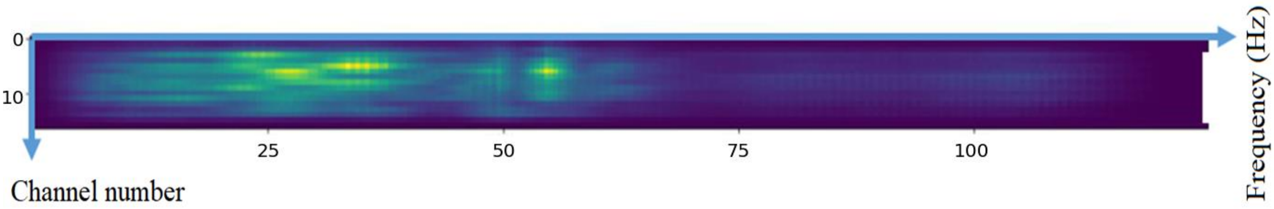
3. Results

First and foremost, an initial classification was performed on the data utilizing all 17 channels and a convolutional neural network, which yielded unsatisfactory results. To improve system performance, the LRP technique was used to analyze the effect of each input element on output classification. In other words, a channel selection was performed based on the role of each channel in the final classification of the data. The LRP approach produces a heat map, which displays the contribution of each input element in the construction of the output (Figure 3).

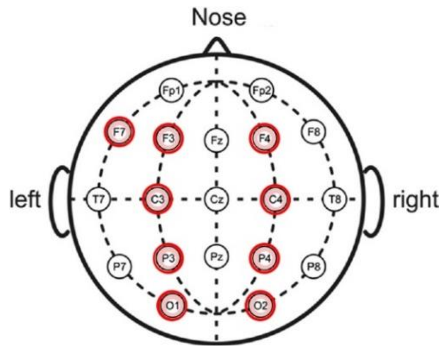
As demonstrated in Figure 3a, channels 3 to 11 are more important in output classification than other channels. As a result, only these channels were used as network input in the next stage. Figure 3b depicts these nine channels.

In this stage, the extracted features are classified using a 9-layer convolutional neural network (2 convolution layers, 2 max-pooling layers, 2 fully connected layers, and 1 output layer). Table 1 shows the design details of the neural network structure.

K-fold cross-validation (K = 5) was utilized to evaluate the neural network performance. The final classification results are shown in Table 2. The total number of data samples, training samples, and test samples were 328, 263, and 65, respectively.



a



b

Figure 3. (a) Average LRP image for all data; and (b) The resultant selected channels

Table 1. The Structure of the Proposed Convolutional Neural Network

Layer order	Layer type	Parameters	Output shape
1	2D Convolution	Number of filters = 6 Size of each filter = (4, 10) Activation = ReLu Stride = (1,1)	(6, 366, 6)
2	2D Max Pooling	Pool size = (2, 2) Stride = (1,1)	(3, 183, 6)
3	2D Convolution	Number of filters = 16 Size of each filter = (1, 7) Activation = Sigmoid Stride = (1,1)	(3, 177, 16)
4	2D Max Pooling	Pool size = (2, 2) Stride = (1,1)	(3, 44, 16)
5	Flatten	-	(1, 2112)
6	Dense	Number of neurons = 80 Activation = ReLu	(1, 80)
7	Dense	Number of neurons = 30 Activation = Sigmoid	(1, 30)
8	Dropout	Keep probability = 0.6	(1, 30)
9	Dense	Number of neurons = 2 Activation = Softmax	(1, 2)

Table 2. The Results of ADHD and Normal Groups Classification Using the Proposed Method

	Accuracy (%)	F1-score (%)	Sensitivity (%)	Specificity (%)
Train	92.45 ± 4.2	95.41 ± 4.4	93.06 ± 5.2	98.1 ± 1.5
Test	89.05 ± 5.9	92.8±4.3	90.7±6.1	96.97±2.4

After training, the proposed model is evaluated on validation data which consists of 72 EEG recording samples. The accuracy obtained for validation data is 94.52%.

4. Discussion and Conclusion

In this study, a deep learning-based model for detecting ADHD in children was proposed. For this purpose, the EEG signals of 31 ADHD children and 30 healthy peers recorded while performing a task of counting the cartoon characters shown to them were used. Following pre-processing, the power spectrum of the clean signals was calculated and used as the convolutional neural network's input. Then, the Layer-wise Relevance Propagation algorithm was utilized in order to find the most important channels as well as providing visualization for the final model interpretability. The accuracy obtained with this method is 94.52 % for validation data.

Although deep learning techniques are appealing, one major drawback of using them in clinical settings is that the models remain a "black box" for specialists or doctors. It is difficult to persuade doctors to utilize such a tool unless it can be interpreted clinically. Model visualization is one possible solution, but not all approaches are appropriate for EEG data or have clear clinical implications. In this study, the LRP method was used to visualize the feature space and select effective channels in addition to using deep structure capabilities to diagnose ADHD disorder. The results show that the proposed strategy works well with EEG data and has clear clinical implications. For example, according to Figure 3, it is shown that channels F3, F4, C3, and C4 in the gamma II frequency range have a notable effect on network output, demonstrating the importance of brain activity in the gamma II band as well as in the frontal and central regions. This finding is consistent with previous studies that found differences in brain segregation [24] and integration [25] in the gamma frequency band between ADHD children and their healthy peers.

Furthermore, the relevance of activity in the gamma II band, as well as in the frontal and central regions, when performing a mathematical task is consistent with the study presented in [26].

By playing a vital role in the functional organization of extended cortical networks, gamma-band activity is associated with higher-order neurocognitive processes [27]. According to [28], children with ADHD exhibit lower levels of resting-state gamma-band activity than their healthy peers. Additionally, it has been indicated in [29] that resting gamma power in both frequency ranges (gamma I and gamma II) is inversely correlated with the severity of ADHD symptoms. In [24], a comparison of the EEG signals of ADHD children with healthy peers while performing an emotional face recognition task also revealed a decrease in gamma-band activity in ADHD children. The current study's findings also demonstrate the significance of gamma-band oscillations in ADHD diagnosis while performing a mathematical task. Given the strong relationship between neurocognitive functions and gamma-band EEG activity, a decrease in gamma-band activity in ADHD children in different states (resting state, emotional face recognition, counting) may be related to the neurocognitive malfunctioning in this disorder, which appears to be task-unrelated.

Table 3 compares the current study to some state-of-the-art studies related to the diagnosis of ADHD using deep learning methods that either have the same dataset as this study or used the visualization of the hidden information of the network structure for interpretability. It can be seen that the accuracy obtained in the current study is either very close to or greater than that of other studies except [14]. [14] addresses the recognition of ADHD disorder using a database similar to the current article's database and a deep learning model. The accuracy of the segment classification reported in this article is 97.81 %, which is marginally better than the accuracy obtained in the current study (94.52 %). However, the current study's method is superior in terms of visualization of hidden information in network structure as well as interpretability.

This study's limitation was the non-availability of distinct data for each individual, which made subject-based classification unfeasible.

Table 3. Comparison of the model accuracy with some state-of-the-art studies in this field

Study	Dataset	Method	Accuracy	Interpretability
This study 2022	31 ADHD children and 30 healthy controls	Layer-wise Relevance Propagation technique and CNN classifier	94.52%	The relevance of activity in the gamma II band, as well as in the frontal and central regions with ADHD discrimination
Moghaddari, M. *et al.* [14] 2020	Same as this study	Frequency band separation, Making RGB images and CNN classifier	97.81%	No Visualization
Dubreuil-Vall *et al.* [19] 2020	20 ADHD adult and 20 healthy controls	Event-related potentials (ERP) during the Flanker Task, ERP spectrograms using Wavelet and CNN Classifier	88%	Decreased power in the alpha band and an increased power in the delta-theta band around 100 ms for ADHD patients compared to healthy controls
Chen, He, Yan Song, and Xiaoli Li [18] 2019	50 ADHD children and 57 healthy controls	Gradient-weighted class activation mapping (Grad-CAM) and CNN classifier	90.29%	The abnormalities of ADHD group were mostly found globally distributed in theta and beta frequency bands.
Chen, He, Yan Song, and Xiaoli Li [12] 2019	50 ADHD children and 51 healthy controls	Mutual information (MI) and CNN classifier	94.67%	Some deep features showed significant between-group differences, and had significant correlations with hand-crafted measures

References

1- Tao Zhang *et al.*, "Separated Channel Attention Convolutional Neural Network (SC-CNN-Attention) to Identify ADHD in Multi-Site Rs-fMRI Dataset." *Entropy,* Vol. 22 (No. 8), p. 893, (2020).

2- Mario Muñoz-Organero, Lauren Powell, Ben Heller, Val Harpin, and Jack Parker, "Using recurrent neural networks to compare movement patterns in ADHD and normally developing children based on acceleration signals from the wrist and ankle." *Sensors,* Vol. 19 (No. 13), p. 2935, (2019).

3- Mark L Wolraich *et al.*, "ADHD diagnosis and treatment guidelines: a historical perspective." *Pediatrics,* Vol. 144 (No. 4), p. e20191682, (2019).

4- Marta Ribasés *et al.*, "Meta-Analysis of Genome-Wide Association Studies On Adult Attention-Deficit and Hyperactivity Disorder." *European Neuropsychopharmacology,* Vol. 29, pp. S758-S59, (2019).

5- Maddalena Mauri *et al.*, "Hemodynamic and behavioral peculiarities in response to emotional stimuli in children with attention deficit hyperactivity disorder: An fNIRS study." *Journal of Affective Disorders,* Vol. 277pp. 671-80, (2020).

6- Austin Schimmel, "ADHD and Fidget Spinners: Using fNIRs to detect changes in Relative Neural Efficiency during the Purdue Pegboard Test." *University of Delaware,* (2020).

7- Atif Riaz, Muhammad Asad, Eduardo Alonso, and Greg Slabaugh, "DeepFMRI: End-to-end deep learning for functional connectivity and classification of ADHD using fMRI." *Journal of Neuroscience Methods,* Vol. 335p. 108506, (2020).

8- Hossein Dini, Farnaz Ghassemi, and Mohammad SE Sendi, "Investigation of Brain Functional Networks in Children Suffering from Attention Deficit Hyperactivity Disorder." *Brain Topography,* pp. 1-18, (2020).

9- Matthew A Jarrett, Philip A Gable, Ana T Rondon, Lauren B Neal, Hannah F Price, and Dane C Hilton, "An EEG study of children with and without ADHD symptoms: between-group differences and associations with sluggish cognitive tempo symptoms." *Journal of attention disorders,* Vol. 24 (No. 7), pp. 1002-10, (2020).

10- Y.-J.; Chang Chu, C.-F.; Shieh, J.-S.; Lee, W.-T., " The Potential Application of Multiscale Entropy Analysis of Electroencephalography in Children with Neurological and Neuropsychiatric Disorders." *Entropy,* Vol. 19 (No. 8), p. 428, (2017).

11- Laura Dubreuil-Vall, Giulio Ruffini, and Joan A Camprodon, "Deep learning convolutional neural networks discriminate adult ADHD from healthy individuals on the basis of event-related spectral EEG." *Frontiers in neuroscience,* Vol. 14p. 251, (2020).

12- He Chen, Yan Song, and Xiaoli Li, "A deep learning framework for identifying children with ADHD using an EEG-based brain network." *Neurocomputing,* Vol. 356pp. 83-96, (2019).

13- Amirali Vahid, Annet Bluschke, Veit Roessner, Sebastian Stober, and Christian Beste, "Deep learning based on event-related EEG differentiates children with ADHD from healthy controls." *Journal of clinical medicine,* Vol. 8 (No. 7), p. 1055, (2019).

14- Majid Moghaddari, Mina Zolfy Lighvan, and Sebelan Danishvar, "Diagnose ADHD disorder in children using convolutional neural network based on continuous mental task EEG." *Computer Methods and Programs in Biomedicine,* Vol. 197p. 105738, (2020).

15- Yang Chang, Cory Stevenson, I-Chun Chen, Dar-Shong Lin, and Li-Wei Ko, "Neurological state changes indicative of ADHD in children learned via EEG-based LSTM networks." *Journal of neural engineering,* Vol. 19 (No. 1), p. 016021, (2022).

16- Amirmasoud Ahmadi, Mehrdad Kashefi, Hassan Shahrokhi, and Mohammad Ali Nazari, "Computer aided diagnosis system using deep convolutional neural networks for ADHD subtypes." *Biomedical Signal Processing and Control,* Vol. 63p. 102227.

17- Robin Tibor Schirrmeister *et al.*, "Deep learning with convolutional neural networks for EEG decoding and visualization." *Human brain mapping,* Vol. 38 (No. 11), pp. 5391-420, (2017).

18- He Chen, Yan Song, and Xiaoli Li, "Use of deep learning to detect personalized spatial-frequency abnormalities in EEGs of children with ADHD." *Journal of neural engineering,* Vol. 16 (No. 6), p. 066046, (2019).

19- Laura Dubreuil-Vall, Giulio Ruffini, and Joan A Camprodon, "Deep learning convolutional neural networks discriminate adult ADHD from healthy individuals on the basis of event-related spectral EEG." *Frontiers in neuroscience,* Vol. 14(2020).

20- Armin Allahverdy Ali Motie Nasrabadi, Mehdi Samavati, Mohammad Reza Mohammadi, "EEG data for ADHD/Control Children ("EEG Competition Data")." National Brain Mapping Laboratory of Iran, Ed., ed. IEEE Dataport, (2020).

21- Zahra Tabanfar, Farzaneh Yousefipoor, Mohammad Firoozabadi, Zeynab Khodakarami, and Zeinab Shankayi, "Recognition of two emotional states of joy and sadness using phase lag index and SVM classifier." in *2016 23rd Iranian Conference on Biomedical Engineering and 2016 1st International Iranian Conference on Biomedical Engineering (ICBME)*, (2016): *IEEE*, pp. 327-30.

22- Amin Behzadnia, Farnaz Ghassemi, Soghra A Chermahini, Zahra Tabanfar, and Athena Taymourtash, "The neural correlation of sustained attention in performing conjunctive continuous performance task: an event-related potential study." *Neuroreport,* Vol. 29 (No. 11), pp. 954-61, (2018).

23- Ali Nouri and Kyanoosh Azizi, "Introducing a Convolutional Neural Network and Visualization of its Filters for Classification of EEG Signal for SSVEP Task." *Frontiers in Biomedical Technologies,* (2020).

24- Mahdiyeh Sarraf Razavi, Mehdi Tehranidoost, Farnaz Ghassemi, Parivash Purabassi, and Athena Taymourtash, "Emotional face recognition in children with attention deficit/hyperactivity disorder: Evidence from event related gamma oscillation." *Basic and Clinical Neuroscience,* Vol. 8 (No. 5), p. 419, (2017).

25- Ansarinasab, Sheida, *et al*. "Investigation of Phase Synchronization in Functional Brain Networks of Children with ADHD using Nonlinear Recurrence Measure." *Journal of Theoretical Biology*: 111381, (2022).

26- Vangelis Sakkalis, Michalis Zervakis, and Sifis Micheloyannis, "Significant EEG features involved in mathematical reasoning: evidence from wavelet analysis." *Brain Topography,* Vol. 19 (No. 1), pp. 53-60, (2006).

27- Peter J Uhlhaas and Wolf Singer, "The development of neural synchrony and large-scale cortical networks during adolescence: relevance for the pathophysiology of schizophrenia and neurodevelopmental hypothesis." *Schizophrenia bulletin,* Vol. 37 (No. 3), pp. 514-23, (2011).

28- Robert J Barry, Adam R Clarke, Mihaly Hajos, Rory McCarthy, Mark Selikowitz, and Jason M Bruggemann, "Acute atomoxetine effects on the EEG of children with attention-deficit/hyperactivity disorder." *Neuropharmacology,* Vol. 57 (No. 7-8), pp. 702-07, (2009).

29- László Tombor, Brigitta Kakuszi, Szilvia Papp, János Réthelyi, István Bitter, and Pál Czobor, "Decreased resting gamma activity in adult attention deficit/hyperactivity disorder." *The World Journal of Biological Psychiatry,* (2018).