ADHD Classification Using Convolution Neural Network

Eman salah[1], Mona Shokair[2,3], Ismail Mohammed Hafez[4,5], Wafaa A. Shalaby[6], Ahmed Elkorany[7], and M. Mokhtar Zayed[8]

(1) Higher Institute of Engineering and Modern Technology-El Marg, Egypt.
eman.salah090@gmail.com
(2) Communication Department, Faculty of Electronic Engineering, Menoufia University, Egypt.
Mona.sabry@el-eng.menofia.edu.eg
(3) Electrical Engineering Department, Faculty of Engineering-October 6 University, Egypt.
Mona.Mohamed.Eng@o6u.edu.eg
(4) Prof Electronics and Communications O6U, Dean College of Engineering, imhafez.eng@O6U.edu.eg
(5) Prof Electronics and Communications ASU, ismail_hafez@eng.asu.edu.eg
(6) Communication Department, Faculty of Electronic Engineering, Menoufia University, Egypt.
Engwafaaahmed88@yahoo.com
(7) Communication Department, Faculty of Electronic Engineering, Menoufia University, Egypt.
elkoranyahmed@el-eng.menoufia.edu.eg
(8) Department of Communications and Computers Engineering, Higher Institute of Engineering, El-Shorouk Academy, El-Shorouk City, Egypt.
mohammed.mokhtar.zayed@gmail.com

Abstract – attention deficit hyperactivity disorder (ADHD) is one of the most common neurodevelopmental disorders in childhood. Some people with ADHD primarily experience symptoms of inattention. Others mostly have symptoms of hyperactivity and impulsivity; some people experience both types of symptoms. This was further confirmed by preprocessing the fMRI raw data and extracting optimal feature methodology. Moreover, the CNN model that was used as a learning model was able to drive an accurate model with preprocessed fMRI data and features extracted. Stochastic gradient ratios with momentum (SGDM) optimizers of the fMRI datasets. Using this optimization technique for adapting the classification system of ADHD cases, it was concluded that the accuracy of PROP 1 is 97.5%, accuracy for PROP 2 is 95%, and accuracy for PROP 3 is 98.75. Finally, it’s found that PROP 3 is the best because of its high accuracy, so the system is improved.

Keywords: ADHD, fMRI, CNN, and Deep Learning.

I. Introduction

Attention deficit hyperactivity disorder (ADHD) is characterized by a persistent pattern of inattention and/or hyperactivity and impulsivity that interferes with functioning or development [1]. Primary care providers for ADHD can show affected people to mental health professionals such as psychiatrists and psychologists, where they have the ability to diagnose the health condition of the injured and provide appropriate treatment for them. In order for the condition to be diagnosed, the symptoms must appear on the infected person for a long period of time. Therefore, a comprehensive evaluation is necessary to determine the cause of the symptoms. Most children with ADHD are diagnosed during their school years [2]. In particular, these symptoms appear in the stage from 3 to 6 and may continue into adulthood and adolescence. [3] When a child reaches elementary school, symptoms of inattention may become more noticeable and cause the child to struggle academically. In adolescence, hyperactivity seems to diminish, and symptoms likely include feeling restless or fidgety, but inattention and impulsiveness may persist. Any teen with ADHD also struggles with antisocial relationships and behaviors. Inattention, anxiety, and impulsiveness tend to persist into adulthood [4].

People with symptoms of inattention often [5]:

- Ignore or miss details and make seemingly careless mistakes in homework, at work, or during other activities.
- He doesn't seem to listen when spoken to directly.
- People with ADHD always find it difficult to organize tasks, not stick to appointments, and are late in studying, as they find it difficult to understand and memorize.
- Avoiding tasks that require sustained mental effort.
- Be forgetful of household chores and tasks and not keeping appointments.

The following reasons indicate that the person has hyperactivity [6]:

- Excess physical movement.
- Act without thinking.
- Inability to focus on tasks.
- Interrupt conversations.

Because hyperactivity disorder with attention deficit hyperactivity disorder leads to a delay in development [7], some people have noticed that this disorder cannot appear in adulthood all at once, but it appears gradually in childhood first, and it may continue until it reaches adulthood. The way inattention, hyperactivity, and impulsivity affect adults can be very different from the way it affects children [8]. As with ADHD in children and teens, ADHD in adults can occur along with several related problems or conditions. Behavioural problems related to ADHD can cause the end of many relationships, a lack of adjustment with people, and an inability to interact in family meetings [9]. Often, ADHD must be treated with psychotherapy and medication. Enjoying good health is very important for all children with hyperactivity disorder and attention deficit hyperactivity disorder. In addition to psychological treatment and medication, good health behavior should be followed to make it easier for children to deal with their lives normally, so that it is easier for them to deal with this disease [10]. These symptoms can cause major problems in a child's life, such as poor achievement in school, poor social interaction with other children and adults, and discipline problems [11].

Categorization or classification is done by using a convolutional neural network. Recent research [12, 13] has also shown that abnormalities in different brain regions, such as the anterior cingulate, posterior cingulate cortex, and ventromedial prefrontal cortex, are major drivers of this disorder. A diagnosis of ADHD shows that these regions are critical when analyzing functional magnetic resonance imaging (fMRI) data. Several relevant studies [14, 15] for the classification of ADHD rely on electroencephalography (EEG), fMRI, and eye movement data using different classification approaches. This study addresses the identification of ADHD based on a seed association approach using fMRI data.

This study addresses the identification of ADHD using fMRI data. This classification process is based on a convolutional neural network (CNN), and relevant features are extracted based on the seed correlations. The high incidence of ADHD in children with clinical symptoms will continue into adulthood by showing destructive elements due to the lack of appropriate treatments [16]. ADHD can improve when children receive treatment, eat healthy food, get enough sleep and exercise, and have supportive parents who know how to respond to ADHD.

The risky factor is that these children are often involved in learning difficulties, which tend to become frustrating as they reach adulthood. Children with this disease suffer from a loss of self-confidence, and this may lead to psychological problems and severe depression. The pervasive symptoms of this type of disorder may lead to social, productivity, and learning problems compared to normal people. Victims may also be affected by other non-psychiatric disorders, depending on their mental state and the severity of the disorder [12]. Functional MRI is an effective method for quantifying brain activity by detecting changes in the blood oxygen level in the brain that occur in response to several neural activities [16]. Correct treatment helps improve ADHD. Parents and educators can teach younger children to better manage their attention, behavior, and emotions. As they get older, children must learn to improve their attention and self-control.

Many studies have used neural networks for fMRI image classification along with other feature reduction methods such as applying masks and feature selection methods such as PCA and ICA for unsupervised learning [17].

A convolutional neural network (CNN) consists of an input layer, hidden layers, and an output layer. It was used as a classifier for feature extraction and pattern recognition for classification between ADHD and other control groups. Then the features were extracted and selected from the pre-prepared data set. Convolutional layers wrap the input and pass its result to the next layer. This is similar to the response of a neuron in the visual cortex to a particular stimulus. [18] CNNs are regular versions of layered cognition. Layered cognition usually means fully connected networks, that is, every neuron in one layer is connected to all neurons in the next layer. The "full connectivity" of these networks makes them vulnerable to data overfitting [19].

Convolutional neural networks are a type of artificial neural networks that use some mathematical operations called convolution instead and are specially designed to process pixel data and are used for image recognition and processing.

CNNs take a different approach to organization: they take advantage of the hierarchical pattern in the data and combine patterns of increasing complexity using smaller, simpler patterns inscribed in their filters.

II. Related Work

A study [20] introduced a new based on granular computing for fMRI data that also explored the fusion of both the patio-temporal details of the fMRI images. A 3D convolutional neural network (CNN) was used with spatial feature data from the fMRI images, and another model of recurrent neural network and feature grouping for comparison was developed. Of the results. The proposed solution method was then extended to learn spatial and temporal data. At the same time, using the 4D CNN approach, which got high performance. Many studies have used fALFF and ReHo map features and Independent Component Analysis (ICA). For effective understanding of neuronal image data [21]. In addition, there are various groups from ReHo, Resting-State Networks (RSN), and Tried the ALFF Foundation [22]. Researchers are interested in the detection and classification of ADHD in healthy individuals and cases of ADHD disturbance. Differences are ranked based on diagnostic accuracy in the protected and exit test data sets. In another study [23] addressed various groups of ALFF, RSN, and ReHo. They showed that ALFF and ReHo provide more classification information with limited accuracy and that a combination of ALFF, ReHo, and RSN was able to achieve an accuracy of 67%. Another study [24] used Deep Bayesian Network (BN), which uses both Bayesian Network and Deep Belief Network (DBN) for implementation. Reduce dimensions to extract attributes from relationships and obtain precision; 66.3% implement an SVM classifier. Early research on ADHD classification using machine learning approaches involved the use of Support Vector Machine (SVM) classifiers and extreme machine learning methods. Using EEG data fed into an SVM classifier for adult ADHD classification, an accuracy of 73% was obtained by considering different scenarios at rest [25].

Deep learning approaches have been directed to classify ADHD in many different domains, spanning deep belief networks (DBN), deep Bayesian networks (BN), convolutional neural networks (CNN), and artificial neural networks (ANN) [12]. Also, they identified improved classification accuracy with the number of hidden layers applied. Another method was proposed by Hao.

Neural network approaches to image classification are emerging rapidly. Deshpande et al. [26] proposed a deep artificial neural network (ANN) for the classification of ADHD and its subtypes. They were able to obtain an accuracy for ADHD classification of 90% and 95% for the subtypes. A fully connected consecutive artificial neural network (FCC ANN) was selected as a well-performing classifier. Another recurrent neural network model has been developed and feature grouping for comparison purposes between results. The proposed solution method was then extended to learn spatial and temporal data at the same time using the 4D CNN method, which obtained high performance [19]. The proposed methodology was well suited to the fMRI data of connectivity patterns characterizing perturbations because it does not need a priori information about spatial patterns in the source signals.

In addition, there are various groups from ReHo, Resting-State Networks (RSN), and I tried the ALFF Foundation [22]. It has been determined that ReHo and ALFF give more. ADHD recognition still uses subjective measures where manual clinical procedures are used and policy is applied. Hence, well-defined and comprehensive objective approaches are motivated in the research to achieve high accuracy and, specifically, high sensitivity in its results. Many psychophysiological measures, such as eye movement data, electroencephalography (EEG), electromyography (EMG), magnetic radiography (MRI), and functional magnetic resonance imaging (fMRI), have been used in the identification of ADHD [27,28]. This review is only concerned with studies published between January 2000 and December 2022, and it could continue until now. The eligibility criteria thus constitute the study published in a peer-reviewed journal from 2000 to 2022. Participants were diagnosed with ADHD according to the Diagnostic and Statistical Manual criteria for mental disorders (DSM) by the American Psychiatric Association (APA); to have children and adolescents and their emotional recognition skills investigated. In general, MRI can identify specific pathologies by evaluating different parts of the cervix. Several studies have analyzed fMRI data to reveal the relationship between the tasks being performed by the subject during examination and brain activation.

III. System Design

The ADHD profile was modeled with functional magnetic resonance imaging (FMRI) data known as ADHD Care V1, as shown in Figure 7. ADHD 200-Global Competition FMRI data was used. We have developed an efficient CNN architecture that leads to optimization of the ability to learn without using classic CNN structures such as Lenet or AlexNet.

We have designed 2 models, namely:

A. CNN PROP 1: In Figure 1, the neural network was created with a convolution layer of 64 filters with a 7*7 kernel size and with the activation function of ‘Relu’. The second layer is max pooling with a size of 5*5, and the third layer is also a convolution layer, which has a filter of 128 with a 5*5 kernel size, and the last layer is average pooling that has a size of 14*14.

B. CNN PROP 2: In Figure 2, the neural network was created with a convolution layer of 64 filters with a 7*7 kernel size and the activation function of ‘Relu’. The second layer was created by 64 filters with 5*5 convolutions and with the same specification as the first layer. Then the third layer of 128 filters with 3*3 kernel sizes, as the fourth layer is also a convolution layer with 128 filters and 3*3 convolutions with an average pooling of 11*11.

Finally, a fully connected with the activation function of Softmax was added to enable the classification.

If the two models are merged together, high accuracy will be obtained, and the system will be greatly improved.

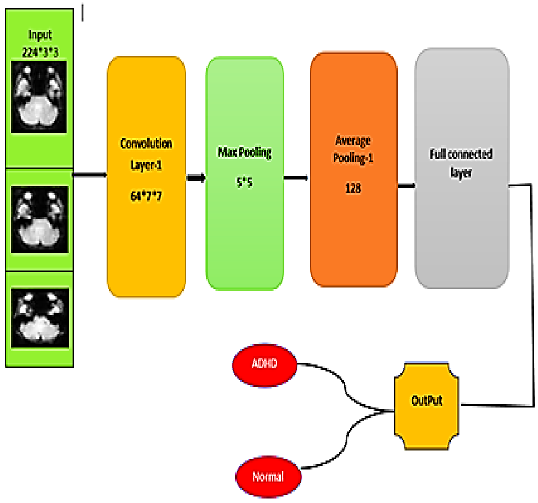

Figure 1: CNN PROP1

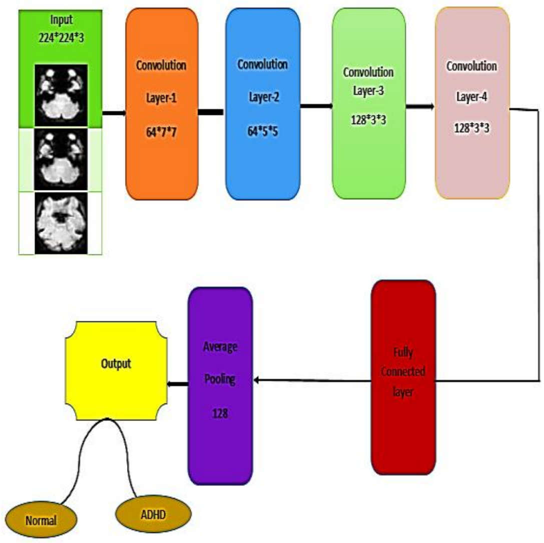

Figure 2: CNN PROP 2

IV. Results and Discussions

This part discusses the performance of CNN ADHD detection. The performance of this system is evaluated by knowing the accuracy, decision, recall, and f1-score. In order to improve the performance of the system, the accuracy of the two models is compared, and the greater the accuracy of the model, the more the system will perform. The ADHD-200 model is a model used to accelerate the scientific community's understanding of the neuropsychological basis of ADHD by implementing discovery-based science sharing. To check and achieve binary classification between all pictures associated with normally developing children's ADHD and ADHD were combined into one category, vs. normal category. The training dataset and test dataset are 20% and 80%, respectively. There is only one optimization algorithm that has been used for the proposed CNN, namely SGDM. There is only one optimization algorithm that was used for the proposed CNN, and that is SGDM. It was found that the best results were obtained at epoch 30 as follows by using stochastic gradient descent with momentum optimizer (SGDM), and it is used to show the relationship between the epochs and the percentages (%).

Through Figure 3, which is approximately equal to 97.5%, precision is 100%, recall is 98.7342%, and F-score is 97.4359%. Fig. 6 for the SGDM optimization algorithm is an extension of random gradient ratios that has recently seen wider adoption for deep learning applications in computer vision and natural language processing. And at the same parameters, the accuracy is approximately equal to 98.75%, precision is 93.55%, recall is 97.5%, and F-score is 95.08%. And also, it shows the RMS prop optimizer is similar to the momentum gradient descent algorithm, and the enhancer constrains the oscillations in the vertical direction. If the two models are merged together, high accuracy will be obtained, and the system will be greatly improved.

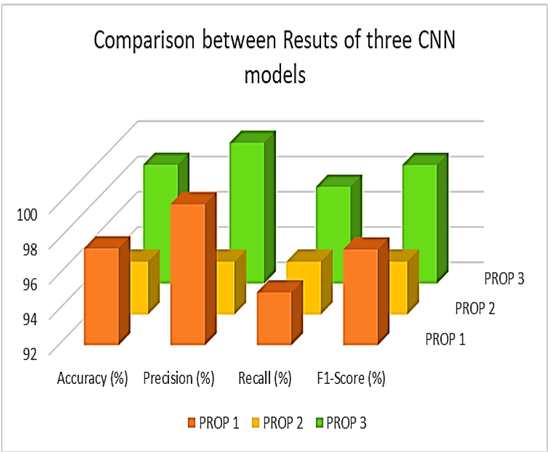

Figure 3: Performance of the Proposed CNN for fMRI Dataset under SGDM optimization algorithm, Mini-Batch size 64, and Learning.

Looking at a convolutional neural network (CNN), the confusion matrix shows where the model gets confused, i.e., which classes the model predicts correctly and which classes the model predicts incorrectly. Confusion matrices are considered as if they were a color-coded heat map using the seaborne library. Sometimes losing training, validation, and accuracy is not enough; we need to know the performance of the validation data. One way is to visualize with a confusion matrix. Figures 4,5 and 6 Confusion matrix visualization and graphical comparison of the confusion matrix of GoogleNet, ResNet, and proposed CNN.

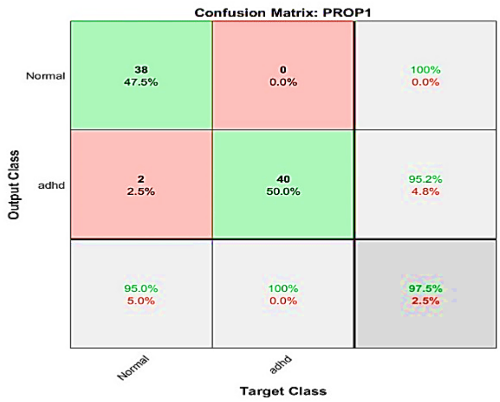

Figure 4: Confusion Matrix for PROP 1.

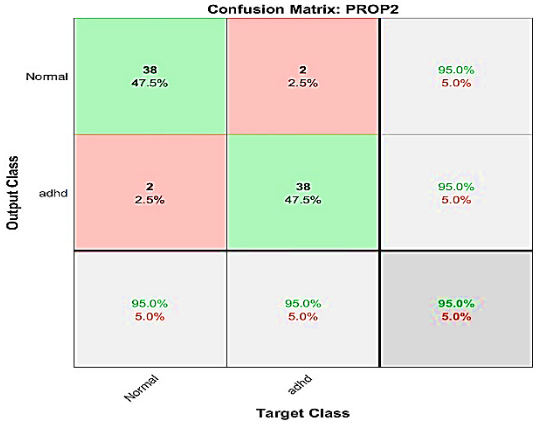

Figure 5: Confusion Matrix for PROP 2.

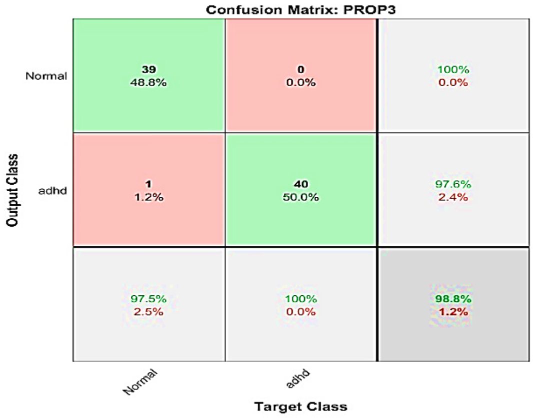

Figure 6: Confusion Matrix for PROP 3.

There are two expected possible categories: "yes" and "no.". If we suspect a disease, for example, then "yes" means they have the disease, and "no" means they don't have the disease. And the dataset used for ADHD is 200 images, and for normal cases it is 200 images. matrix is only used for testing processes. There are two possible categories expected: "yes" and "no.". If we suspect a disease, for example, "yes" means they have the disease, and "no." means they don't have the disease. The data set used for ADHD is 200 images, and for normal cases it is 200 images. The matrix is only used for testing the accuracy for PROP 1.

Figure 7 illustrates the Roc curve, a graph showing the performance of a rating or rating model across all ratings. Fortunately, there is an efficient algorithm based on sorting that can do this.

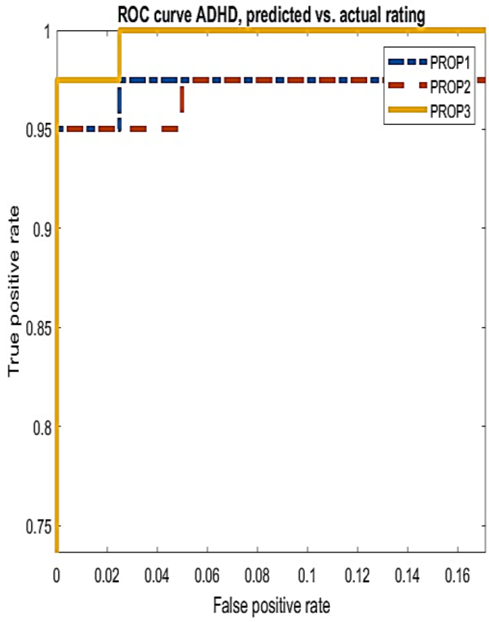

Figure 7: Roc Curve.

V. Conclusions

This research sheds light on the experiences of children and adults with ADHD. It is important for doctors, and practitioners should be aware of the positive and negative effects of the disorder and how it may affect patients' lives. More research is needed in these areas to explore patients' attitudes toward receiving a formal diagnosis, including opinions related to late diagnosis or delayed treatment. The use of fMRI in attention deficit hyperactivity disorder research has an impact. Try to shed light on the neural mechanisms that underlie the behaviors. This paper dealt with fMRI image processing and feature extraction and a classification approach that enables accurate subject classification of ADHD. Basically, this study is based on selecting optimal data processing methods and extracting features based on seed correlation, ReHo, and fALFF of a CNN-based learning model that considers the entire DMN region coordinates.

References

[1] Kooij, Sandra JJ, et al. "European consensus statement on diagnosis and treatment of adult ADHD: The European Network Adult ADHD." *BMC psychiatry* 10 (2010): 1-24.

[2] Eloyan, Ani, et al. "Automated diagnoses of attention deficit hyperactive disorder using magnetic resonance imaging." *Frontiers in systems neuroscience* 6 (2012): 61.

[3] Wolf, Robert C., et al. "Regional brain activation changes and abnormal functional connectivity of the ventrolateral prefrontal cortex during working memory processing in adults with attention-deficit/hyperactivity disorder." *Human brain mapping* 30.7 (2009): 2252-2266.

[4] Beaudry, Olivia, et al. "Featural processing in recognition of emotional facial expressions." *Cognition & emotion* 28.3 (2014): 416-432.

[5] Schwenck, Christina, et al. "Empathy in children with autism and conduct disorder: Group-specific profiles and developmental aspects." *Journal of Child Psychology and Psychiatry* 53.6 (2012): 651-659.

[6] Mrug, Sylvie, et al. "Peer rejection and friendships in children with attention-deficit/hyperactivity disorder: Contributions to long-term outcomes." *Journal of abnormal child psychology* 40 (2012): 1013-1026.

[7] Bergwerff, Catharina Elisabeth, et al. "Neurocognitive profiles in children with ADHD and their predictive value for functional outcomes." *Journal of attention disorders* 23.13 (2019): 1567-1577.

[8] Twilhaar, E. Sabrina, et al. "Social adjustment in adolescents born very preterm: evidence for a cognitive basis of social problems." *The Journal of pediatrics* 213 (2019): 66-73.

[9] Tehrani-Doost, Mehdi, et al. "Is emotion recognition related to core symptoms of childhood ADHD?" *Journal of the Canadian Academy of child and Adolescent Psychiatry* 26.1 (2017): 31.

[10] ADHD-200 consortium. "The ADHD-200 consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience." *Frontiers in systems neuroscience* 6 (2012): 62.

[11] Wolf, Robert C., et al. "Regional brain activation changes and abnormal functional connectivity of the ventrolateral prefrontal cortex during working memory processing in adults with attention-deficit/hyperactivity disorder." *Human brain mapping* 30.7 (2009): 2252-2266.

[12] Hong, Soon-Beom, et al. "Subthreshold attention-deficit/hyperactivity disorder is associated with functional impairments across domains: a comprehensive analysis in a large-scale community study." *European child & adolescent psychiatry* 23 (2014): 627-636.

[13] Polanczyk, Guilherme, et al. "The worldwide prevalence of ADHD: a systematic review and metaregression analysis." *American journal of psychiatry* 164.6 (2007): 942-948.

[14] Kirova, Anna-Mariya, et al. "Are subsyndromal manifestations of attention deficit hyperactivity disorder morbid in children? A systematic qualitative review of the literature with meta-analysis." *Psychiatry research* 274 (2019): 75-90.

[15] McConaughy, Stephanie H., et al. "Academic and social impairments of elementary school children with attention deficit hyperactivity disorder." *School Psychology Review* 40.2 (2011): 200-225.

[16] Epstein, Jeffery N., and Richard EA Loren. "Changes in the definition of ADHD in DSM-5: subtle but important." *Neuropsychiatry* 3.5 (2013): 455.

[17] Nijmeijer, Judith S., et al. "Attention-deficit/hyperactivity disorder and social dysfunctioning." *Clinical psychology review* 28.4 (2008): 692-708.

[18] Zou, Liang, et al. "3D CNN based automatic diagnosis of attention deficit hyperactivity disorder using functional and structural MRI." *Ieee Access* 5 (2017): 23626-23636.

[19] Takahashi, Naoya, et al. "Deep convolutional neural networks and data augmentation for acoustic event detection." *arXiv preprint arXiv:1604.07160* (2016).

[20] Mao, Wei, et al. "Learning trajectory dependencies for human motion prediction." *Proceedings of the IEEE/CVF international conference on computer vision*. 2019.

[21] Zou, Liang, et al. "3D CNN based automatic diagnosis of attention deficit hyperactivity disorder using functional and structural MRI." *Ieee Access* 5 (2017): 23626-23636.

[22] Miao, Bo, and Yulin Zhang. "A feature selection method for classification of ADHD." *2017 4th international conference on information, cybernetics and computational social systems (ICCSS)*. IEEE, 2017.

[23] Nigg, Joel T., et al. "Toward a revised nosology for attention-deficit/hyperactivity disorder heterogeneity." *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 5.8 (2020): 726-737.

[24] Hao, A. Junyu, B. Lianghua He, and C. Hongyi Yin. "Discrimination of ADHD children based on Deep Bayesian Network." *2015 iet international conference on biomedical image and signal processing (icbisp 2015)*. IET, 2015.

[25] TaghiBeyglou, Behrad, et al. "Detection of ADHD cases using CNN and classical classifiers of raw EEG." *Computer Methods and Programs in Biomedicine Update* 2 (2022): 100080.

[26] Deshpande, Anjali D., Marcie Harris-Hayes, and Mario Schootman. "Epidemiology of diabetes and diabetes-related complications." *Physical therapy* 88.11 (2008): 1254-1264.

[27] De Silva, Senuri, et al. "A survey of attention deficit hyperactivity disorder identification using psychophysiological data." (2019): 61-76.

[28] Kaur, Amandeep, and Karanjeet Singh Kahlon. "Accurate identification of ADHD among adults using real-time activity data." *Brain sciences* 12.7 (2022): 831.