Detection of Attention Deficit Hyperactivity Disorder by Using EEG Feature Maps and Deep Learning

Burak Akbugday and Ozge Ada Bozbas
Dept. of Electrical and Electronics Eng.
Izmir University of Economics
Izmir, TURKEY
burak.akbugday@ieu.edu.tr

Ozlem Karabiber Cura
Dept. of Biomedical Eng.
Izmir Katip Celebi University
Izmir, TURKEY
ozlem.karabiber@ikcu.edu.tr

Sude Pehlivan and Aydin Akan*
Dept. of Electrical and Electronics Eng.
Izmir University of Economics
Izmir, TURKEY
akan.aydin@ieu.edu.tr

***Abstract*—Attention deficit hyperactivity disorder (ADHD) is a mental disorder that affects the behavior of the persons, and usually onsets in childhood. ADHD generally causes impulsivity, hyperactivity, and inattention which impairs day-to-day life even in the adulthood if left undiagnosed and untreated. Although various guidelines for diagnosis of ADHD exist, a universally accepted objective diagnostic procedure is not established. Since current diagnosis of ADHD heavily relies on the expertise of healthcare providers, an EEG Topographic Feature Map (EEG-FM) based method is proposed in this study which aims to objectively diagnose ADHD. 6 different features extracted from EEG recordings acquired from 33 participants, 15 ADHD patients and 18 control subjects, converted into EEG-FM images and fed into a convolutional neural network (CNN) based classifier. Results indicate that the proposed method can accurately classify ADHD patients with up to 99% accuracy, precision, and recall.**

***Index Terms*—Attention Deficit Hyperactivity Disorder (ADHD) detection, EEG feature maps, deep learning, CNN.**

I. Introduction

Attention Deficit Hyperactivity Disorder (ADHD) is a neuro-developmental disorder which affects children and adults and has long-term consequences [1], [2]. ADHD can include problems with attention deficit, which includes symptoms such as difficulty sustaining attention. Defining ADHD disease and determining an appropriate treatment method for the patient play an important role in increasing the patient's quality of life [1]–[4].

ADHD is often diagnosed based on the information provided by the patient, their teachers, parents, and questionnaires. This subjective diagnosis is impacted by the doctor's training. Diagnosing ADHD is a difficult task, and misdiagnosis is likely to occur [1], [2]. Numerous researchers have concentrated on recording neural activity using invasive and non-invasive techniques to diagnose ADHD since it is intimately associated with brain function. The majority of research uses non-invasive techniques to diagnose ADHD, including electroencephalography (EEG), functional near-infrared spectroscopy (fNIRS), functional magnetic resonance imaging (fMRI), and magnetoencephalography (MEG). Additionally, using magnetic resonance imaging and functional neuroimaging methods may also need costly and drawn-out procedures [1], [2]. On the other hand, EEG signals have been widely utilized as an effective, inexpensive, and non-invasive method for the identification of brain activity [1], [2], [5].

Studies of ADHD have been based on recording the EEG signal under resting conditions and other cognitive conditions. There are many studies investigating the sum, absolute and relative power of EEG sub-frequency bands such as delta, theta and alpha with the aim of defining ADHD with their EEG features. [4]–[10]. At the same time, the decrease in fast wave activity in the beta sub-band is one of the important indicators in the diagnosis of ADHD.

Several nonlinear measurement techniques have been used effectively to extract important information for ADHD diagnosis from EEG signals. Complexity measures of EEG, such as entropy, have been used more and more to define ADHD. For complexity analysis of EEG activity, approximate entropy, sample entropy, Kolmogorov-Sinai entropy, Tsali entropy, fuzzy entropy, permutation entropy and wavelet entropy were used [4], [6], [11]–[15]. Another complexity criterion used to evaluate the nature of the EEG signal is the fractal dimension, and different FD measurements were calculated based on Higuchi, Katz, Sevcik, and Petrosian techniques [4], [11], [13], [16]. The complexity of the EEG signals of ADHD patients and successful results have been reported.

In recent studies [2], [17]–[22], deep learning models have been used for the diagnosis of ADHD, which can distinguish between patients who are sick and those who are healthy, using EEG data to classify very well. Some researchers have focused on using images obtained from EEG signals as input to CNN architectures [18], [20], [22] rather than using direct EEG signal as input [17], [19], [21]. In the aforementioned ADHD studies, different approaches have been proposed to generate input images to deep networks using EEG signals.

In the aforementioned ADHD investigations, many methods for producing input pictures for deep networks utilizing EEG data have been presented. EEG Topographic Feature Maps

*This study was partially supported by Izmir University of Economics, Scientific Research Projects Coordination Unit. Project number: 2022-07.

(EEG-FM) visualize the activity levels of different parts of the brain. Imaging techniques such as brain mapping can be used to link the connections and functions of the brain. Finding a functionally integrated link between physically different brain areas is made possible by brain functional connectivity [23]. EEG-FM, a novel technique developed utilizing conventional EEG features, successfully identifies emotions in experiments involving emotion identification. It represents the spatial and temporospatial information of an EEG segment [24]–[26].

In this study, an innovative, fast, and automatic classification method for the diagnosis of ADHD based on EEG-FM and CNN is proposed. EEG-FMs generated from 6 different nonlinear features extracted from each EEG channel. These features are namely Approximate Entropy (ApEn) [6], [11], [13], [15], Correlation Dimension (CD) [13], [14], Hurst Exponent (HE) [12], Higuchi's Fractal Dimension (HFD) [11]–[13], [16], Katz's Fractal Dimension (KFD) [11]–[13], [16], and Largest Lyapunov Exponent (LLE) [11]–[16]. Although these features themselves provide valuable information regarding the EEG recordings, EEG-FMs allow spatial information to be kept, enabling a more accurate discrimination of the ADHD and normal recordings [24]–[26]. Thus they were mapped onto images on a channel-by-channel basis. All six features were fed into two custom CNN structures to investigate their classification performances in terms of different parameters.

II. Material and Method

In conventional studies that uses features extracted from segmented EEG signals, electrode topography is not taken into consideration. However, the positioning of the EEG electrodes and the information acquired from each electrode is significant [24]. The change of activity in different regions of the brain is reflected on the EEG electrodes, providing valuable spatial information. By representing both temporal and spatial information in terms of topographic images, it was made easier to distinguish the levels of activity change via a 2-layer CNN.

In this study, various non-linear EEG features extracted from the each segmented EEG signals were mapped onto images. These images are then used as an input to two CNN-based classifiers to distinguish ADHD and control subjects.

A. Experimental ADHD Dataset

A 4 minute spontaneous EEG was recorded from 15 ADHD patients and 18 healthy persons while they were at rest and had their eyes open. ADHD patient population consisted of 8 female and 7 male children with mean age of 12; whereas healthy children population consisted of 14 female and 4 male children with a mean age of 13. The data was recorded at Izmir Katip Celebi University using Brain Vision EEG system. International 10-20 system is used for electrode placement and acquisition of 30 channels which were Fp1, Fp2, F7, F8, F3, F4, Fz, FT7, FT8, FC3, FC4, FCz, T3, T4, C3, C4, Cz, TP7, TP8, CP3, CP4, CPz, T5, T6, P3, P4, O1, O2, Oz. The study was carried out under the approval of Izmir Katip Celebi University Non-Interventional Clinical Research Ethics Committee numbered 76, and dated 11.07.2019.

The sampling frequency of the recordings are 1 kHz, and they are filtered with a Butterworth band-pass filter with a pass band of $[0.5-50]$ Hz to remove various sources of noise such as power-line. After denoising, data of each EEG channel is segmented into 5 s epochs for further processing.

B. The Construction of EEG Topographical Feature Maps

The construction of EEG-FMs goes through three stages. 15 ADHD patients and 18 control participants provided the 30-channel raw EEG readings, each lasting for 4 minutes. EEG segments of 5 s in length are used to extract the time domain and nonlinear features, namely Higuchi's Fractal Dimension (HFD), Katz's Fractal Dimension (KFD), Correlation Dimension (CD), Largest Lyapunov Exponent (LLE), Approximate Entropy (ApEn), and Hurst Exponent (HE), for all 30 channels' $\times$ 48 (4 min/5 s) segments after the necessary denoising and segmentation steps. For each EEG channel and each segment of ADHD patients and control individuals, 6 characterizing features are acquired. Then, the feature vector is scaled between 0 and 1, as shown in equation (1) which includes features calculated from EEG segments of both ADHD patients and healthy controls:

$$Y_i^{'} = \frac{Y_i - Y_{min}}{Y_{max} - Y_{min}} \tag{1}$$

where $Y_i^{'}$ denotes the normalized feature value of i^{th} subject, Y_{max} and Y_{min} are the maximum and minimum values of the total feature vector. By arranging the normalized feature from one subject, one segment, and all channels on a matrix as described in [24]–[26], EEG-FM matrix of any feature is constructed. For each of the 6 features of control individuals and ADHD patients.

The normalized feature values of the 30 EEG channels are mapped onto the matrix as shown in Fig.1 top-right using the same 10-20 electrode placements system as is used in the recording stage (shown in Fig.1 top-left). Some researchers recently suggested a mapping method that precisely places each electrode into a matrix with nine rows and nine columns [24]–[26]. This mapping is appropriate for all EEG acquisition systems with fewer than 81 electrodes. Therefore, using this matrix mapping, EEG-FMs were generated. The normalized feature values are inserted into dots labeled with channel names in the feature matrix. A visual representation of the mapping process is given in Fig. 1, top-right corner. Eq. 2 can be used to show how the values of the missing points which are not labeled with channels relate to the values of the surrounding labeled points.

$$W_{(i,j)} = \frac{W_{(i+1,j)}^{'} + W_{(i-,j)}^{'} + W_{(i,j+1)}^{'} + W_{(i,j-1)}^{'}}{M}, \tag{2}$$

$0 \leq i, j \leq 8; \quad (i,j) \in N$ where W indicates the normalized feature value of the gray point, $W^{'}$ is the normalized value of the point neighboring this point. The default value of K is 1, and it refers to the number of non-zero components in the numerator.

In total, 4380 and 5196 EEG-FMs were constructed for the control and ADHD group, labeled as 0 and 1, respectively.

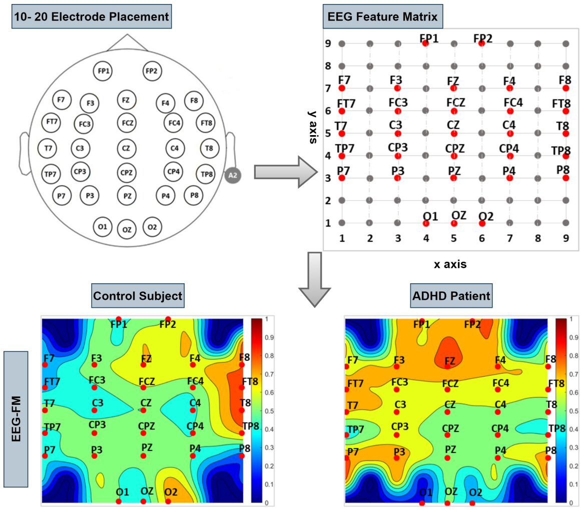

Fig. 1: 10-20 electrode mapping on the 9×9 matrix and generation of EEG-FMs.

C. Convolutional Neural Network

Deep learning has become an intense interest of researchers in recent years. Convolutional Neural Networks (CNN) is the most popular and used algorithm in deep learning studies [27]. Inspired by the brain neural networks of animals, CNN is used especially for image recognition and classification. Furthermore, CNN is also very good at classifying audio, time series and signal data [28]. The CNN algorithm, which uses images as input, consists of many interconnected layers of neurons, each of which has a specific purpose. CNN basically consists of three main layers: convolution, pooling, and fully connected layers [27].

Two different CNN models are developed and tested in order to get the more optimal results: "CNN Model #1" and "CNN Model #2" are given in Fig. 2. In both models, 681×541 sized EEG-FM images are utilized. These EEG-FMs contain EEG data gathered from subjects with and without ADHD, as mentioned in the previous sections. The CNN model was used as a binary classifier and EEG-FM images given as input are labeled as 0 and 1 indicating individuals without and with ADHD, respectively. The data was split into training and test sets using 80% and 20% of the total data, respectively. Classification results of these 2 models were compared in terms of accuracy, precision, and recall as well as their learning performance was compared with loss functions.

III. Results and Discussion

In this study, a CNN-based model in which EEG-FMs are used as input has been proposed in order to understand whether individuals have ADHD.

EEG-FMs are generated using nonlinear EEG features, and then used for the training of 2 different CNN models. EEG-FMs belonging to all 6 features were given as input to CNN, labeled as "0" and "1", indicating control and ADHD subjects. Both of the CNN models were trained for 10 epochs. Results of the training and validation processes were given in Table I.

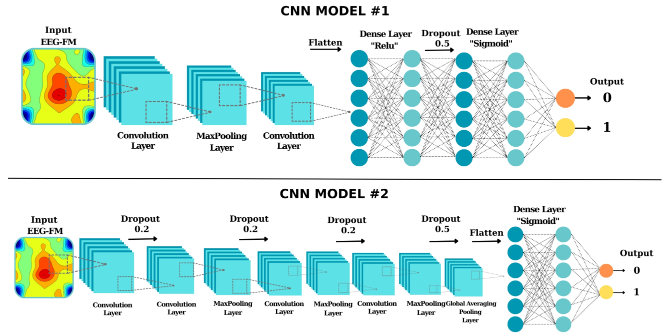

Fig. 2: (a) Layer structure of the first CNN model used in classification (b) Layer structure of the second CNN model used in classification

CNN Model #1 and CNN Model #2 were given the same EEG-FM images and both models were trained for a total of 10 epochs. Comparing the 2 models, CNN Model #1 has fewer drop outs and layers than CNN Model #2. As seen in the results in Table I, the train accuracy of the first and last epoch in CNN Model #1 is % 96.69 and % 99.9, respectively. In CNN Model #2, the train accuracy of the first and last epoch is %79.1 and %95.88, respectively. When CNN Model #1 and CNN Model #2 are compared, although CNN Model #1 has a higher accuracy result, we can conclude that CNN Model #2 learned better when looking at the first and last epochs.

Due to time constraints, it was not possible to optimize the parameters and epoch numbers. Therefore loss function behavior of both models are not ideal. However, it can be argued that the models did not overfit and were able to generalize well considering loss and corresponding accuracy, precision, and recall values. To better visualize the results, Fig. 3, was plotted to depict the accuracy and loss of CNN Models #1 and #2.

IV. Conclusion

In this study, EEG-FMs generated from 6 nonlinear features were used as input of 2 different CNN models to classify ADHD. By mapping numerical features extracted from EEG data into matrices to form EEG-FMs, both temporal and spatial information is represented on the same image. This enables preserving topographical layout of the EEG acquisition via electrodes, which is important when it comes to highlighting and making use of which areas of the brain is more active.

Preliminary results indicate that the proposed method provides encouraging classification performance. Although there are not many studies published in this field that use a comparable method to classify ADHD that correlates directly to ours, the proposed method outperforms similar studies such as [29] and [24]. While it is not possible to directly compare the outcome of this study with the existing literature, it can be argued that the method proposed can compete with previous studies while introducing significant novelties to the field of interest.

TABLE I: Classification results for both CNN Model #1 and CNN Model #2 over 10 epochs. "Num. of Epochs" indicates the number of epoch that the result was acquired. "Train." indicates data related to training of the CNN, and "Val." indicates data related to the validation of the CNN model.

Model #1								
Num. of Epochs	**Train. Loss**	**Train. Accuracy**	**Train. Precision**	**Train. Recall**	**Val. Loss**	**Val. Accuracy**	**Val. Precision**	**Val. Recall**
# 1	0.1635	0.9669	0.9734	0.9654	0.2941	0.8703	0.9087	0.8456
# 2	0.0732	0.9776	0.9786	0.9802	0.2941	0.8661	0.8794	0.8726
# 3	0.0579	0.9821	0.9835	0.9835	0.4638	0.8787	0.8941	0.8726
# 4	0.02	0.9946	0.9967	0.9934	0.5638	0.8787	0.8941	0.8803
# 5	0.0108	0.9991	0.9984	1	0.7093	0.8849	0.9016	0.8842
# 6	0.007	0.9973	0.9967	0.9984	0.7236	0.8828	0.9044	0.8764
# 7	0.0076	0.9973	0.9983	0.9967	0.7851	0.8724	0.9057	0.8533
# 8	0.0052	0.9991	0.9984	1	0.6723	0.887	0.902	0.888
# 9	0.002	1	1	1	0.8033	0.8849	0.908	0.8764
# 10	0.0048	0.9991	1	0.9984	0.6292	0.8849	0.8893	0.8996
Model #2								
Num. of Epochs	**Train. Loss**	**Train. Accuracy**	**Train. Precision**	**Train. Recall**	**Val. Loss**	**Val. Accuracy**	**Val. Precision**	**Val. Recall**
# 1	0.5272	0.791	0.7468	0.8216	0.4115	0.8332	0.7695	0.9072
# 2	0.4333	0.8173	0.7646	0.8676	0.4067	0.8311	0.7615	0.9186
# 3	0.3704	0.8529	0.8291	0.8545	0.3906	0.8377	0.7721	0.9155
# 4	0.3112	0.8826	0.8835	0.8562	0.2509	0.8868	0.8394	0.9307
# 5	0.2462	0.913	0.9148	0.893	0.2331	0.9039	0.8794	0.9155
# 6	0.1902	0.9287	0.9276	0.9155	0.2328	0.9004	0.8579	0.9376
# 7	0.1574	0.9415	0.9384	0.9335	0.2628	0.8966	0.8237	0.9848
# 8	0.1321	0.955	0.9528	0.9485	0.1926	0.9133	0.8732	0.9482
# 9	0.1268	0.9536	0.948	0.9508	0.2199	0.9077	0.8421	0.9825
# 10	0.1179	0.9588	0.9538	0.9563	0.1563	0.9251	0.8851	0.9612

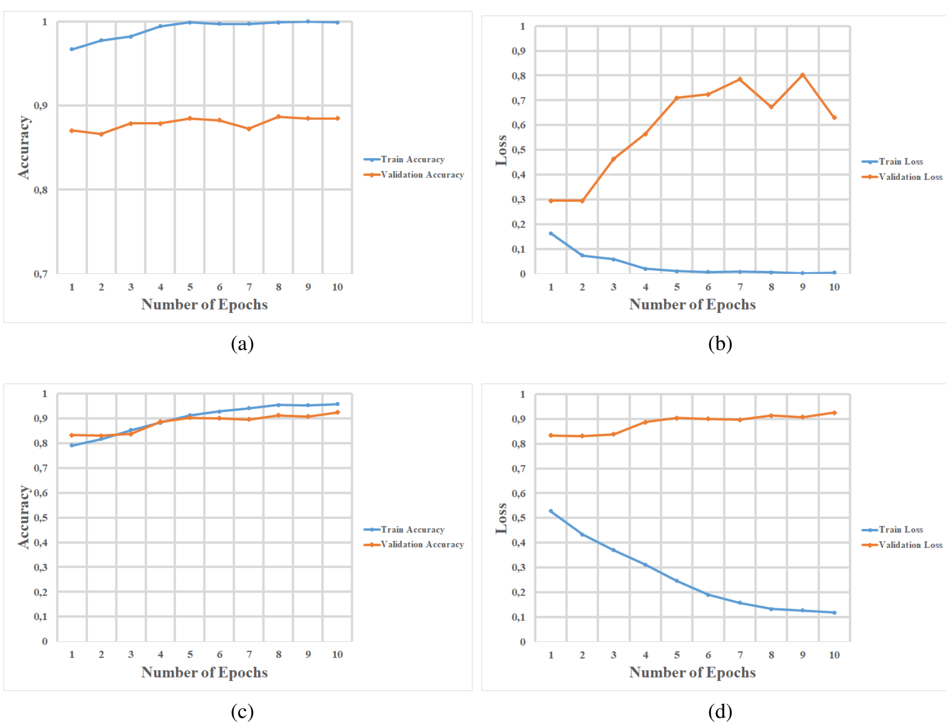

Fig. 3: Training and validation over 10 epochs of the CNN Model #1 and #2. a) Accuracy of Model #1 b) Loss of Model #1 c) Accuracy of Model #2 d) Loss of Model #2

In the future studies, 6 nonlinear features will be classified individually, and the results will be compared with this study. This may help highlighting more prominent features better, thus enabling shorter duration for the training of the CNN models. Furthermore, different CNN models will be trained to investigate the possibility of finding more suitable networks for the data used and the performance of proposed CNNs will be compared with the conventional machine learning algorithm classifier performances. Finally, EEG-FMs of the linear features that can be extracted from the dataset will be used to compare and contrast better performing feature sets in classification of ADHD.

REFERENCES

[1] P. Ghaderyan, F. Moghaddam, S. Khoshnoud, and M. Shamsi, "New interdependence feature of EEG signals as a biomarker of timing deficits evaluated in attention-deficit/hyperactivity disorder detection," *Measurement*, vol. 199, p. 111468, 2022.

[2] M. Bakhtyari and S. Mirzaei, "ADHD detection using dynamic connectivity patterns of EEG data and ConvLSTM with attention framework," *Biomedical Signal Processing and Control*, vol. 76, p. 103708, 2022.

[3] V. Férat, M. Arns, M.-P. Deiber, R. Hasler, N. Perroud, C. M. Michel, and T. Ros, "Electroencephalographic microstates as novel functional biomarkers for adult attention-deficit/hyperactivity disorder," *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, vol. 7, no. 8, pp. 814–823, 2022.

[4] M. Altınkaynak, N. Dolu, A. Güven, F. Pektaş, S. Özmen, E. Demirci, and M. İzzetoğlu, "Diagnosis of attention deficit hyperactivity disorder with combined time and frequency features," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 3, pp. 927–937, 2020.

[5] H. Chen, W. Chen, Y. Song, L. Sun, and X. Li, "EEG characteristics of children with attention-deficit/hyperactivity disorder," *Neuroscience*, vol. 406, pp. 444–456, 2019.

[6] J. J. González, L. D. Méndez, S. Mañas, M. R. Duque, E. Pereda, and L. De Vera, "Performance analysis of univariate and multivariate EEG measurements in the diagnosis of ADHD," *Clinical Neurophysiology*, vol. 124, no. 6, pp. 1139–1150, 2013.

[7] S. J. Johnstone, L. Parrish, H. Jiang, D.-W. Zhang, V. Williams, and S. Li, "Aiding diagnosis of childhood attention-deficit/hyperactivity disorder of the inattentive presentation: Discriminant function analysis of multi-domain measures including EEG," *Biological Psychology*, vol. 161, p. 108080, 2021.

[8] A. Singh, C. J. Yeh, N. Verma, and A. K. Das, "Overview of attention deficit hyperactivity disorder in young children," *Health psychology research*, vol. 3, no. 2, 2015.

[9] A. Tenev, S. Markovska-Simoska, L. Kocarev, J. Pop-Jordanov, A. Müller, and G. Candrian, "Machine learning approach for classification of ADHD adults," *International Journal of Psychophysiology*, vol. 93, no. 1, pp. 162–166, 2014.

[10] I. Buyck and J. R. Wiersema, "State-related electroencephalographic deviances in attention deficit hyperactivity disorder," *Research in developmental disabilities*, vol. 35, no. 12, pp. 3217–3225, 2014.

[11] M. R. Mohammadi, A. Khaleghi, A. M. Nasrabadi, S. Rafieivand, M. Begol, and H. Zarafshan, "EEG classification of ADHD and normal children using non-linear features and neural network," *Biomedical Engineering Letters*, vol. 6, no. 2, pp. 66–73, 2016.

[12] H. T. Tor, C. P. Ooi, N. S. Lim-Ashworth, J. K. E. Wei, V. Jahmunah, S. L. Oh, U. R. Acharya, and D. S. S. Fung, "Automated detection of conduct disorder and attention deficit hyperactivity disorder using decomposition and nonlinear techniques with EEG signals," *Computer Methods and Programs in Biomedicine*, vol. 200, p. 105941, 2021.

[13] Y. K. Boroujeni, A. A. Rastegari, and H. Khodadadi, "Diagnosis of attention deficit hyperactivity disorder using non-linear analysis of the EEG signal," *IET systems biology*, vol. 13, no. 5, pp. 260–266, 2019.

[14] F. Ghassemi, M. Hassan_Moradi, M. Tehrani-Doost, and V. Abootalebi, "Using non-linear features of EEG for ADHD/normal participants' classification," *Procedia-Social and Behavioral Sciences*, vol. 32, pp. 148–152, 2012.

[15] S. Khoshnoud, M. A. Nazari, and M. Shamsi, "Functional brain dynamic analysis of ADHD and control children using nonlinear dynamical features of EEG signals," *Journal of integrative neuroscience*, vol. 17, no. 1, pp. 17–30, 2018.

[16] A. Allahverdy, A. K. Moghadam, M. R. Mohammadi, and A. M. Nasrabadi, "Detecting ADHD children using the attention continuity as nonlinear feature of EEG," *Frontiers in Biomedical Technologies*, vol. 3, no. 1-2, pp. 28–33, 2016.

[17] H. Chen, Y. Song, and X. Li, "A deep learning framework for identifying children with ADHD using an EEG-based brain network," *Neurocomputing*, vol. 356, pp. 83–96, 2019.

[18] A. Vahid, A. Bluschke, V. Roessner, S. Stober, and C. 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.

[19] M. Moghaddari, M. Z. Lighvan, and S. Danishvar, "Diagnose ADHD disorder in children using convolutional neural network based on continuous mental task EEG," *Computer Methods and Programs in Biomedicine*, vol. 197, p. 105738, 2020.

[20] A. Ahmadi, M. Kashefi, H. Shahrokhi, and M. A. Nazari, "Computer aided diagnosis system using deep convolutional neural networks for ADHD subtypes," *Biomedical Signal Processing and Control*, vol. 63, p. 102227, 2021.

[21] L. Dubreuil-Vall, G. Ruffini, and J. 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, p. 251, 2020.

[22] M. Tosun, "Effects of spectral features of EEG signals recorded with different channels and recording statuses on ADHD classification with deep learning," *Physical and Engineering Sciences in Medicine*, vol. 44, no. 3, pp. 693–702, 2021.

[23] L. S. Hooi, H. Nisar, and Y. V. Voon, "Tracking of EEG activity using topographic maps," in *2015 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*. IEEE, 2015, pp. 287–291.

[24] A. Topic and M. Russo, "Emotion recognition based on EEG feature maps through deep learning network," *Engineering Science and Technology, an International Journal*, vol. 24, no. 6, pp. 1442–1454, 2021.

[25] H. Chao, L. Dong, Y. Liu, and B. Lu, "Emotion recognition from multiband EEG signals using capsnet," *Sensors*, vol. 19, no. 9, p. 2212, 2019.

[26] Y. Li, J. Huang, H. Zhou, and N. Zhong, "Human emotion recognition with electroencephalographic multidimensional features by hybrid deep neural networks," *Applied Sciences*, vol. 7, no. 10, p. 1060, 2017.

[27] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M. A. Fadhel, M. Al-Amidie, and L. Farhan, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, no. 1, p. 53, 2021.

[28] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[29] B. TaghiBeyglou, A. Shahbazi, F. Bagheri, S. Akbarian, and M. Jahed, "Detection of ADHD cases using CNN and classical classifiers of raw EEG," *Comput. Methods Programs Biomed. Update*, vol. 2, no. 100080, p. 100080, 2022.