Deep Neural Network Technique for Automated Detection of ADHD and CD Using ECG Signal

(Article begins on next page)

28 April 2024

Deep Neural Network Technique for Automated Detection of ADHD and CD Using ECG Signal

Hui Wen Loh[a], Chui Ping Ooi[a], Shu Lih Oh[b], Prabal Datta Barua[b,c,d,e,f,g,h,i,j], Yi Ren Tan[k], Filippo Molinari[l], Sonja March[m], U. Rajendra Acharya[n,*], Daniel Shuen Sheng Fung[k,o]

[a] School of Science and Technology, Singapore University of Social Sciences, Singapore
[b] Cogninet Australia, Sydney, NSW 2010, Australia
[c] Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, NSW 2007, Australia
[d] School of Business (Information System), University of Southern Queensland, Australia
[e] Australian International Institute of Higher Education, Sydney, NSW 2000, Australia
[f] School of Science & Technology, University of New England, Australia
[g] School of Biosciences, Taylor's University, Malaysia
[h] School of Computing, SRM Institute of Science and Technology, India
[i] School of Science and Technology, Kumamoto University, Japan
[j] Sydney School of Education and Social work, University of Sydney, Australia
[k] Developmental Psychiatry, Institute of Mental Health, Singapore
[l] Biolab, Department of Electronics and Telecommunications, Politecnico di Torino, 10129 Torino, Italy
[m] Centre for Health Research and School of Psychology and Wellbeing, University of Southern Queensland, Springfield, Australia.
[n] School of Mathematics, Physics, and Computing, University of Southern Queensland, Springfield, Australia
[o] Lee Kong Chian School of Medicine, Nanyang Technological University; DUKE NUS Medical School, National University of Singapore; Yong Loo Lin School of Medicine, National University of Singapore

* Correspondence: rajendra.acharya@usq.edu.au (U. Rajendra Acharya)

Abstract

Background and Objective: Attention Deficit Hyperactivity problem (ADHD) is a common neurodevelopment problem in children and adolescents that can lead to long-term challenges in life outcomes if left untreated. Also, ADHD is frequently associated with Conduct Disorder (CD), and multiple research have found similarities in clinical signs and behavioral symptoms between both diseases, making differentiation between ADHD, ADHD comorbid with CD (ADHD+CD), and CD a subjective diagnosis. Therefore, the goal of this pilot study is to create the first explainable deep learning (DL) model for objective ECG-based ADHD/CD diagnosis as having an objective biomarker may improve diagnostic accuracy.

Methods: The dataset used in this study consist of ECG data collected from 45 ADHD, 62 ADHD+CD, and 16 CD patients at the Child Guidance Clinic in Singapore. The ECG data were segmented into 2s epochs and directly used to train our 1-dimensional (1D) convolutional neural network (CNN) model.

Results: The proposed model yielded 96.04% classification accuracy, 96.26% precision, 95.99% sensitivity, and 96.11% F1-score. The Gradient-weighted class activation mapping (Grad-CAM) function was also used to highlight the important ECG characteristics at specific time points that most impact the classification score.

Conclusion: In addition to achieving model performance results with our suggested DL method, Grad-CAM's implementation also offers vital temporal data that clinicians and other mental healthcare professionals can use to make wise medical judgments. We hope that by conducting this pilot study, we will be able to encourage larger-scale research with a larger biosignal dataset. Hence allowing biosignal-based computer-aided diagnostic (CAD) tools to be implemented in healthcare and ambulatory settings, as ECG can be easily obtained via wearable devices such as smartwatches.

1. Introduction

Attention Deficit Hyperactivity Disorder (ADHD) is a neurodevelopment disorder that can affect children as young as four years old [1]. ADHD symptoms include disorganization, forgetfulness, difficulty completing tasks, and a lack of attention and concentration [2]. Additionally, those with ADHD have a lack of emotional control and poor decision-making skills, which negatively affects many facets of life, including academics, employment, and interpersonal relationships [2]. Therefore, it is critical for ADHD to be diagnosed early in order to receive timely treatment and management, as well as to improve outcome for those with ADHD and prevent further development of comorbidities [2]. CD is a behavioral disorder in which the affected individual shows disregard for basic societal standards and rules, resulting in rule-breaking behaviors such as theft, rule violation, property destruction, and aggression toward people and animals [3]. As a result, CD is strongly linked to later delinquency and criminal activity, particularly when combined with ADHD [4].

The "gold standard" for ADHD and CD diagnosis includes a comprehensive interview with the child's caregiver, a mental status assessment, a medical examination, cognitive testing, parent and teacher rating scales, and school reports [5]. Mental health professionals and clinicians would administer these assessment tools and interviews to confirm if the child meets the diagnostic criteria based on a classification system such as the Diagnostic and Statistical Manual of Mental Disorders-Fifth Edition (DSM-5) [2], [6]. Hence, diagnosing ADHD and CD is a time-consuming, costly, and inconvenient process that typically occurs after symptom behaviours are well established, thus impeding early detection of these disorders. Furthermore, these "gold standard" ADHD and CD diagnoses are subjective, particularly when relying on teacher, parent or caregiver reports, making them susceptible to misdiagnosis when bias is introduced during interviews and administration of rating scales. In addition, the clinical symptoms of ADHD and CD are very similar, further increasing the possibility of misdiagnosis. Therefore, this study looks into an objective diagnosis of ADHD

and CD using Electrocardiogram (ECG) signals which can overcome such diagnostic limitations.

While there is currently no standard biomarker for the diagnosis of ADHD or CD, some studies have suggested that the brain-heart interaction mediated by the autonomic nervous system exhibits changes in neurological and behavioral disorders such as ADHD and CD [7]–[9]. As a result, it has been reported that ADHD [10], [11], CD [12]–[14], and ADHD comorbid with CD (ADHD+CD) [12] patients have a higher heart rate upon triggered interaction such as performing tasks like aerobic exercises and watching film clips. ADHD children and adolescents' ECG characteristics, in particular, are significantly associated with early repolarization anomalies [13]. Furthermore, Berg et al. [15] found that ADHD medications, specifically CNS stimulants and atomoxetine, have an increased risk of sudden cardiac death and arrhythmia syndromes, so prescribing these medications to ADHD patients should be done with caution, especially if the patient has problematic cardiac arrhythmias or other cardiac disease [15], [16]. This highlights the significance of discovering ECG biomarkers in ADHD patients.

With these ECG characteristics that may be unique to ADHD or CD patients, Artificial Intelligence (AI) techniques such as Machine Learning (ML) and Deep Learning (DL) can be used to read the patients' ECG signals and classify them into respective ADHD and CD groups [14]. The recent advancement of artificial intelligence in healthcare applications has also demonstrated the strong growing potential of implementing ECG-based AI models in healthcare, such as in cardiovascular disease management, such as continuous bedside monitoring [17] and as a screening tool for cardiac disease such as cardiomyopathy, ischemia, and aberrant arrhythmia [18]. Apart from cardiovascular disease, ECG-based AI models are also widely studied in detection of sleep apnea [19], anxiety, stress and depression [20], [21]. Furthermore, explainable artificial intelligence (XAI) has been created in recent years to address the 'black box' issue that AI models face due to the lack of interpretability of their predicted outcomes. As a result, XAI adds another layer of explanations to AI models in the form of feature importance, saliency maps, or rule-based predictions, among other things, which offers essential information to physicians as they validate the outcome predicted by the AI model [22].

The detection of ADHD or CD with ECG-based AI model, on the other hand, has not been thoroughly explored and is still in its infancy unlike that of cardiovascular diseases. To date, only Koh et al. [23] have used ML to detect ADHD and CD in ECG where empirical wavelet transform (EWT) were used to extract nonlinear features of ECG which was fed into bagged-tree classifier to perform the classification task. We anticipate that, in addition to cardiovascular disease, the expanding potential of using ECG as a diagnostic and monitoring tool will extend to other areas such as ADHD and CD. Therefore, this will be the first study to employ DL to detect ADHD, CD, and ADHD+CD with ECG data. Furthermore, we will incorporate XAI into our DL model so that the predicted outcome may be understood by physicians and other mental healthcare professionals. As a result, our model outperforms prior work by Koh et al. [23], and XAI has revealed key ECG patterns that our model judged to have a tight relationship to ADHD or CD. If such models can be shown to accurately detect

2. Methods

The deep learning system proposed in this study is depicted in Fig 1. Section 2.1 contains information on the dataset used in this study. Section 2.2 describes the preprocessing of the ECG signals. Section 2.3 expands on the model architecture of proposed deep learning system, and Section 2.4 introduces Grad-CAM, which is implemented to provide time localization of ECG characteristics which are important for diagnosis of ADHD, CD, and ADHD+CD.
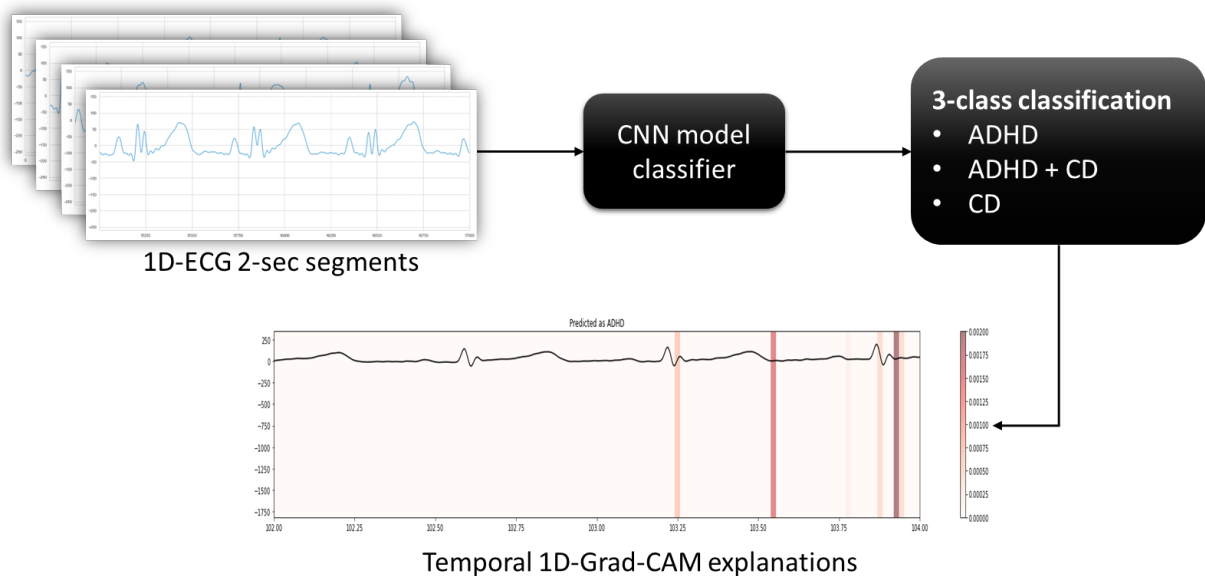
Fig 1. Experimental setup for CNN-based ADHD/CD detection using ECG signals.

2.1 Data acquisition

The Domain Specific Review Board (DSRB) of Singapore's National Healthcare Group approved a clinical trial (NCT00819429), also known as Supplement and social skills interventions (SASSI) trial, that provided the ECG data for this investigation (NHG) [24]. The objective of the SASSI trial is to investigate whether omega-3 supplementation with and without social skills training affects post-treatment efficacy for lowering reactive aggression. As a result, the data demonstrate that omega-3 supplements are more effective than standard therapy, such as medication and parent training, in lowering reactive aggression. The trial also collected ECG data at baseline time point from its participants in addition to examining the efficacy of omega supplements. As a result, the data set includes ECG recordings of 123 participants (7-16 years old) from Singapore's Child Guidance Clinic, a child and adolescent psychiatric outpatient clinic. These children were diagnosed using the Diagnostic and Statistical Manual of Mental Disorders fourth edition Text Revision (DSM-IV-TR), and their parents also completed a computerized Diagnostic Interview Schedule for Children (DISC), a standard diagnostic tool that is frequently used in ADHD research and assessment [25]. ECG

data from 123 subjects were recorded axially on the left and right ribs at the level of the heart using disposable silver or silver chloride adhesive electrodes, which could reduce noise artifacts during recording. The ECG signal was recorded for approximately 3-minutes at sampling frequency of 1,000 Hz while participants were instructed to be completely relaxed with their eyes open. The ECG was recorded using a BIOPAC ECG100C electrocardiogram amplifier connected to a proprietary MP150 data acquisition and AcqKnowledge analysis software. To reduce noise artifacts further, a Blackman Window -61dB with 0.5 Hz to 35 Hz high-frequency bandpass finite impulse response filter was used in the AcqKnowledge software, along with a notch filter of 60 Hz in MATLAB.

To protect patients' privacy, their ECG data were de-identified and anonymized. These participants were split into three groups according to DSM-IV-TR: CD only (16 children), ADHD only (45 participants), and ADHD + CD (62 participants).

2.2 Preprocessing

There was no signal preprocessing stage to extract ECG features or ECG signal decomposition. Instead, the proposed deep learning system was fed with raw ECG signals. Also, as the ECG recordings of some samples were varying between 170-sec to 180-sec, we are unable to utilize the full 3-min for this study as it will exclude some samples. Hence, we decided to select the middle 170-sec portion of every sample. For instance, to extract the middle 170-sec segment of a 178-sec ECG recording, we first divide 178 by 2 to get 89. Then, we take 89 ± 85 to get a range between 4-sec to 174-sec, resulting in length of 170-sec.

Finally, each 170-sec ECG signal was divided into 85 chunks of 2-sec epochs, yielding 3,825 ADHD samples, 5,270 ADHD+CD samples, and 1,360 CD samples.

2.3 Deep 1D-CNN model

In this study, the ECG segments were classified into three categories: CD, ADHD + CD, and CD using 1-Dimensional (1D) Convolutional Neural Network (CNN) models. The CNN model is well-known for its ability to classify images, and as a result, it has been used in applications like face and object identification [26], [27], satellite forecasting [28], [29] and analysis of medical images such as MRI, CT, X-RAY, and PET [30], [31]. Besides 2-dimensional images, CNN models have also been applied to 1-dimensional biosignals, such as ECG to identify arrhythmias [32], [33]. CNN models are also commonly employed in conjunction with other Explainable artificial intelligence (XAI) techniques for working with high-dimensional data such as medical imaging and biosignals, hence eliminating the 'black box' model issue that causes a lack of interpretation in model prediction outcomes [22], [34]. Therefore, we will be proposing 1D-CNN model in our study to classify our ECG segments into CD, ADHD + CD, and CD classes. The model architecture of our proposed deep learning model is shown in Table 1 and Fig 2.

Table 1
Summary of model layer parameters of proposed 1D-CNN model.

No.	Layer	Filter no.	Kernel size	Unit	Output

1	Conv1D	8	3	-	1998 x 8
2	MaxPooling1D	8	-	2	999 x 8
3	Conv1D	16	3	-	997 x 16
4	MaxPooling1D	16	-	2	498 x 16
5	Conv1D	32	3	-	496 x 32
6	MaxPooling1D	32	-	2	248 x 32
7	Conv1D	64	3	-	246 x 64
8	MaxPooling1D	64	-	2	123 x 64
9	Conv1D	128	3	-	121 x 128
10	GlobalMaxPooling1D	128	-	-	128
11	Dropout (0.2)	-	-	-	128
12	Dense	-	-	32	32
13	BatchNormalization	-	-	-	32
14	Dense (Softmax)	-	-	3	3
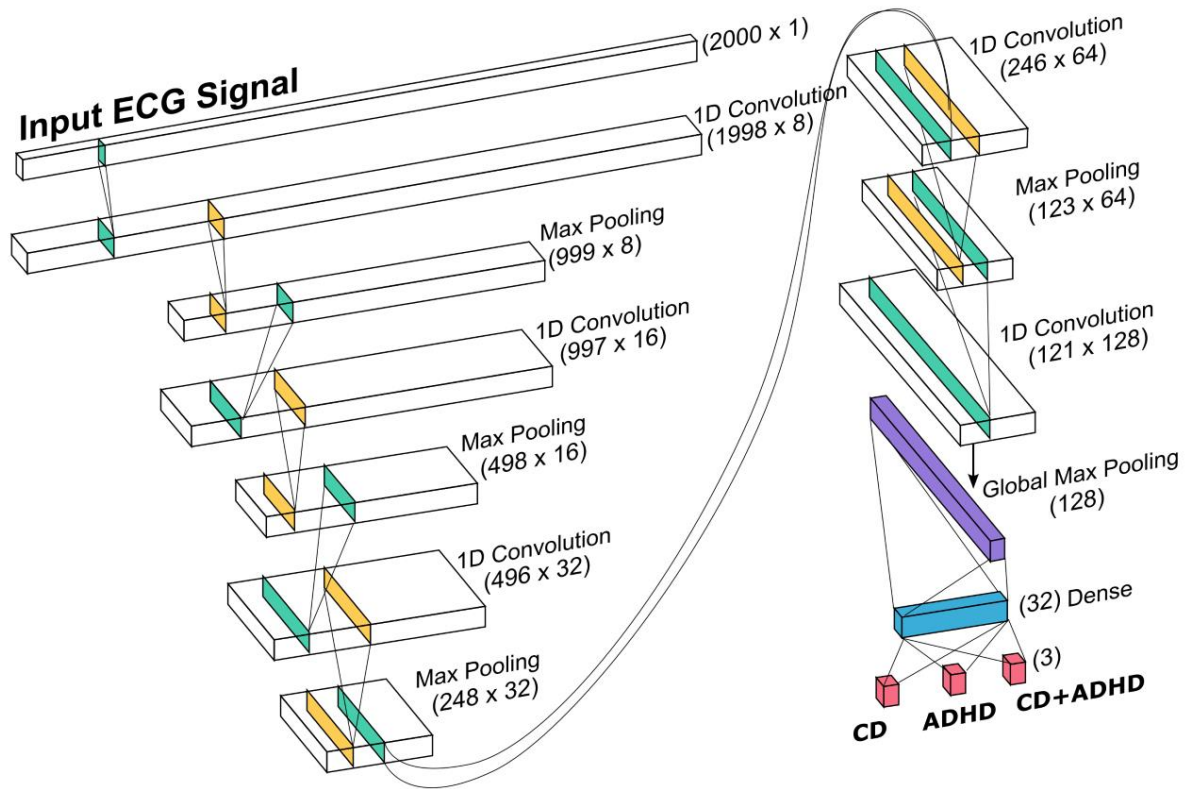
Fig 2. Model architecture of proposed deep learning model.

The CNN model is made up of three major components: a convolutional layer, a pooling layer, and a fully-connected layer. The convolutional layer and pooling layer both increase the saliency of important features while decreasing the complexity of the input, preventing model overfitting [35]. The working mechanism of the convolutional and pooling layer is shown in Eq. 1 and 2 respectively, where S is the input, $*$ is the discrete convolution operation, W is the

convolution kernel, i is the length of the input signal, and O is the resulting feature map from the convolutional operation in Eq. 1 [36], [37].

$$S * W_n = \sum_{i=1}^{|W|} WiSi + n - 1 \qquad (1)$$

$$O_n^l = S_{Wi} * Wi_n \qquad (2)$$

As a result, the convolution and max pooling operations reduced the input ECG segment from 2,000 to 121, which is then reduced to a 1D array of length 128 after the global max pooling layer. The global max pooling layer covers the entire feature map, rather than being restricted to a kernel size W like in Eq. 1 and 2. Finally, we have the fully-connected layers, which are made up of two neural network layers, the first of which contains 32 neurons and the second of which contains 3 neurons acting as the classifier layer with the SoftMax activation function. We have a dropout layer of 0.2 between these neural network layers, which also acts as a measure to prevent model overfitting. Additionally, rectified linear unit (ReLU) activation function and l2 kernel regularizer of 0.01 were inserted in each convolutional layer to further lessen the effects of overfitting; ReLU activation functions require less computational work [38] and l2 kernel regularizers impose restrictions on the model weights, lowering the model's complexity [39]. Our model was trained using 100 epochs and a 64-batch size, with Adam optimizer (learning rate = 0.001, decay rate = 0.0005) and sparse categorical cross entropy as the loss function. The weighted loss was also implemented during model training to correct the imbalance in the dataset brought on by CD samples being much smaller than ADHD and ADHD+CD samples. Tensorflow was used to build the proposed model in Python (v2.9.1). The computer's specs for training the model are: Intel Core i9-12900F processor, Nvidia QuadroA2000 12GB graphics card, 128GB of RAM, and a 1.0 TB 2.5 Inch SATA SSD.

2.4 Gradient-weighted Class Activation Mapping (Grad-CAM)

Grad-CAM is a popular XAI technique for deciphering 'black box' CNN models, where it emphasizes the key characteristics of the input data used by the CNN model to make a prediction [40], [41]. As the CNN performs the convolution operation repeatedly on the feature map, the distinctive features will become more apparent over time. The final convolutional layer is considered to hold the most distinctive feature which the CNN model uses to make predictions. Hence, the final convolutional layer is the most suitable layer to apply Grad-CAM. By analyzing the gradient data provided by the neurons in the final convolutional layer, which assigns importance value to the region of interest on the feature map, the Grad-CAM generates a heatmap with the important regions highlighted in red [40], [41]. Although Grad-CAM is often used with 2D-CNN, it can also be used for 1D-CNN. Here, we have applied Grad-CAM to our ECG data and observed temporal localization of the critical ECG features supporting the diagnosis of ADHD, CD, and ADHD+CD.
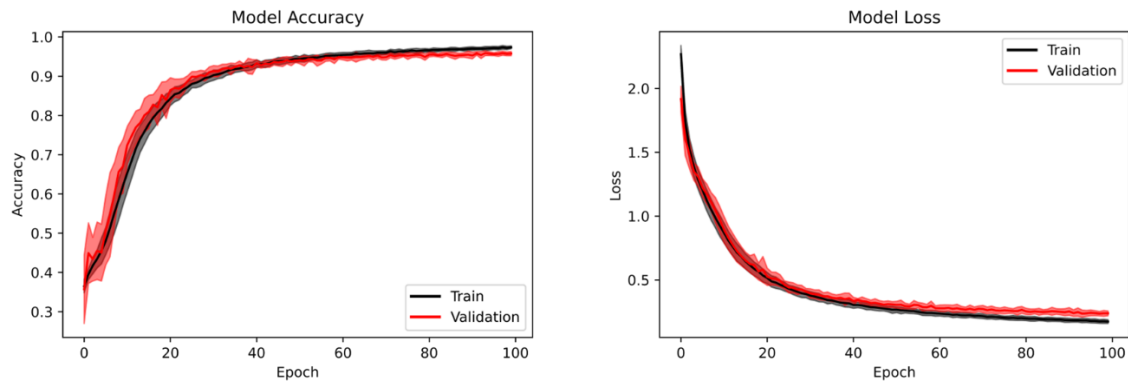
3. Results

The model performance was evaluated using ten-fold cross validation (10-fold CV) using the following performance measures (Eq. 3-6), where TP, TN, FP, and FN stands for True Positive, True Negative, False Positive, and False Negative samples. We included a callback function during the 10-fold CV that saves the best performing model during model training, whose performance will be evaluated with the test fold. ==Also, before running the next fold, we clear each session to guarantee that the memory or weights from previous model training are deleted so that we can run the next iteration with a clean slate.== Fig 3. shows the performance graphs of our model during 10-fold CV. The graphs show that the model is not overfitted because of the small gap between the training and validation curves for both the model accuracy and loss plot.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{5}$$

$$F1 - Score = \frac{2TP}{2TP + FN + FP} \tag{6}$$

Fig 3. Performance graph of proposed 1D-CNN during 10-fold CV.

We computed the performance metrics of the model using classification report from scikit-learn [42] and the resulting model's performance for each fold in the 10-fold CV is summarized in Table 2. The proposed model achieved a high average classification accuracy of 96.04%, where the proposed model correctly identified 3655/3825 ADHD, 5080/5270 ADHD+CD, and 1306/1360 CD samples as shown in the normalized confusion matrix in Fig 4. The model also demonstrated good performances in the three metrics: 96.26% precision, 95.99% sensitivity, and 96.11% F1-score (Table 2). From Eq. 4, model precision is the proportion of correctly identified positive samples (TP) among the predicted positive samples. In other words, it measures the number of FP errors the model makes. On the other hand, model sensitivity is the proportion of positive samples that are correctly identified (Eq. 5), hence it estimates the number of FN errors. When the classes are imbalanced, like in this case,

there is usually a trade-off between precision and sensitivity. Yet, our suggested model was able to attain high model precision and sensitivity, demonstrating that the model did not incorrectly categorize the data as either FP or FN samples. Another metric called F1-score was used to assess the balance between the model's precision and sensitivity. As a result, our model has a high F1-score, demonstrating its ability to balance the trade-off between model precision and sensitivity.

Table 2
Performance parameter of proposed 1D-CNN model during 10-fold cross-validation.

| Fold	Accuracy (%)	Precision (%)	Sensitivity (%)	F1-Score (%)
1	95.60	95.58	96.16	95.86
2	94.84	95.45	94.84	95.13
3	96.46	97.09	95.71	96.36
4	97.32	96.92	97.28	97.08
5	95.41	96.30	94.90	95.57
6	95.02	95.14	95.12	95.13
7	95.50	95.69	95.95	95.82
8	96.36	96.75	96.53	96.63
9	96.65	97.17	96.30	96.71
10	97.22	96.50	97.15	96.82
Mean ± SD	96.04 ± 0.84	96.26 ± 0.70	95.99 ± 0.82	96.11 ± 0.67

Fig 4. Normalized confusion matrix of Proposed 1D-CNN model after 10-fold cross-validation.

3.2 Grad-CAM explanation for 1D-ECG signals

In order to interpret the prediction made by the proposed model, we conducted another experiment by selecting one subject each from the ADHD, ADHD + CD, and CD class at

random as the test set and used the remaining subjects as the training set for the proposed model. Following that, we used Grad-CAM on the three subjects in the test set to highlight regions on the ECG segments that the model considered significant for prediction.

Each subject's ECG signal was divided into 85 chunks, as described in section 2.2. The confusion matrix obtained is shown in Fig 5. which displays the number of chunks per subject that were properly identified as belonging to the assigned class. There were 70/85 and 83/85 ECG portions that were correctly classified as ADHD and ADHD+CD class, respectively, while the CD subject had all chunks correctly classified as the CD class. The remaining 15 ADHD chunks were misclassified as ADHD+CD, while the remaining 2 ADHD+CD chunks were misclassified as ADHD, indicating the similarity in ECG characteristics between these two classes. The Grad-CAM applied to all of the ECG chunks, as shown in Fig 6., was able to highlight the regions that were thought to be important for prediction. Furthermore, time information was provided in the visualization map generated. Fig 6a., for example, shows the small ECG peaks between 44.25s and 44.50s and 45.00s to 45.25s for predicted ADHD, Fig 6b. shows the larger ECG deflection between 33.00s and 33.75s to 34.00s for ADHD+CD prediction, and Fig 6c. shows the ECG deflection between 6.50s and 6.75s and around 7.50s for CD prediction. The complete Grad-CAM analysis for the ADHD, ADHD+CD and CD subject is provided as supplementary material.
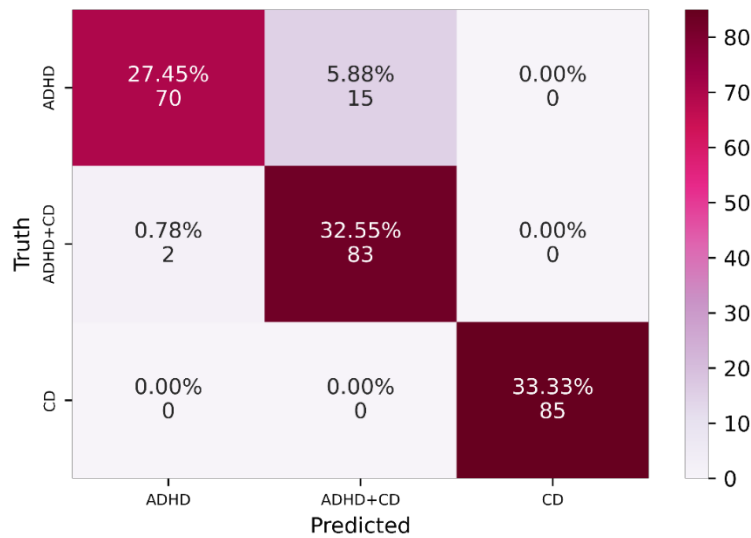
Fig 5. Confusion matrix of Proposed 1D-CNN model on test set of 3 subjects from each disorder category.
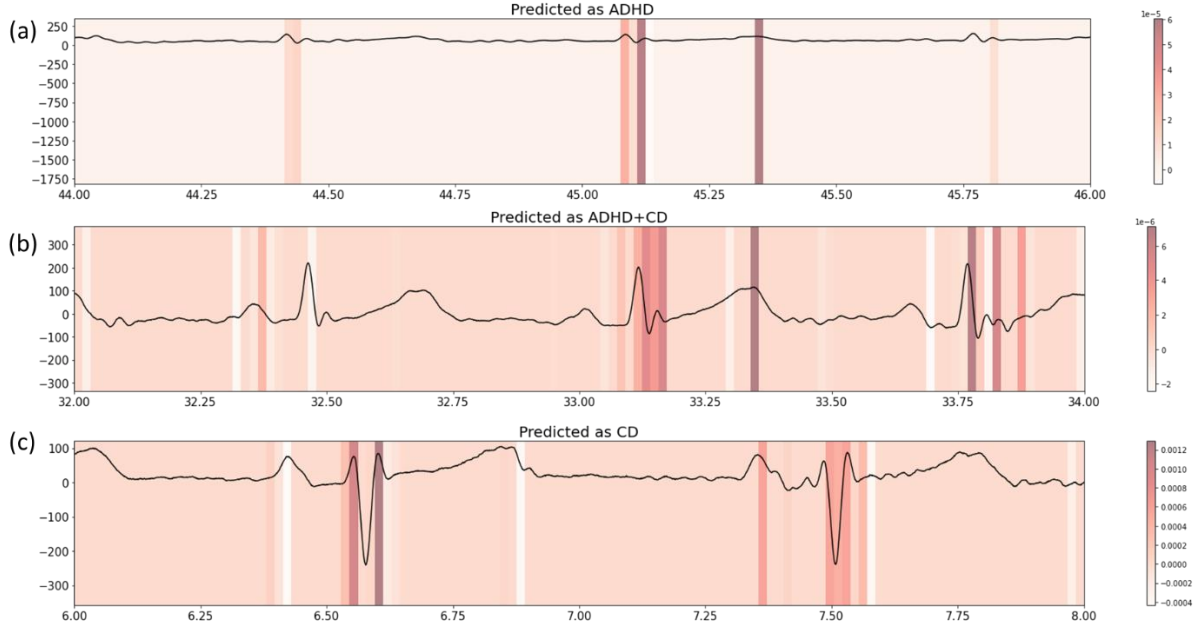
Fig 6. Example of 1D-Grad-CAM heatmap produced for each ECG segment of (a) ADHD subject, (b) ADHD+CD subject, and (c) CD subjects. Red highlighted regions indicate the ECG characteristic recognized by proposed 1D-CNN model as important for prediction. Y-axis represents the amplitude of the ECG signal, X-axis represent time in seconds, and plot title shows the predicted class by the proposed model.

4. Discussion

The aim of this pilot study is to use DL on ECG signals to distinguish the three classes (ADHD, ADHD+CD, and CD) and Grad-CAM to explain the model's prediction. To date, there are no publications on the use of ECG signals for ADHD/CD diagnosis, apart from one study published by Koh et al [23]. In their study, they proposed ML approach to distinguish between the 3 classes: ADHD, ADHD+CD, and CD. Because ML classifiers are unable to digest high dimensional data such as raw ECG signals, additional steps to decompose the signal for feature extraction were required [43]. As a result, Koh et al. [23] decomposed the signal using empirical wavelet transform (EWT) to extract nonlinear features, which were then used to train the bagged-tree classifier to perform the classification task. In addition, feature selection was included to select the top significant features for classification, resulting in information loss as not all extracted features were used.

This study, on the other hand, proposes using a DL approach, which simplifies the process by directly feeding segmented raw ECG signals into the CNN classifier. Grad-CAM was used to interpret the prediction result by highlighting important regions of interest that the CNN model deemed useful when making a prediction. As a result, this study is an improvement of the previous work, where classification accuracy increased from 87.19% to 96.04% (Table 3). Introducing XAI such as Grad-CAM also improves interpretability of our proposed model, which is frequently compromised in ML and DL models [22]. Additionally, we have programmed our Grad-CAM implementation so that temporal information is provided to

show the time when the ECG characteristics that are crucial for prediction occurred. This will provide clinicians and mental health professionals the additional critical information they need to make informed decisions.

Table 3
List of studies that used the same dataset to distinguish ADHD, ADHD+CD, and CD.

Study	Features	Classifier	Accuracy (%)
Tor [44], **2021**	EEG (Nonlinear features)	ML (kNN)	97.88
Koh [23], **2021**	ECG (EWT entropies)	ML (Bagged tree)	87.19
This work	ECG (Segmented raw signals)	DL (CNN+Grad-CAM)	96.04

Tor et al. had also performed EEG-based ADHD/CD identification on the same dataset used here [44]. Similarly, they proposed an ML approach for collecting nonlinear EEG properties and performing classification using a k-nearest neighbor (kNN) classifier. They were able to attain a high classification accuracy of 97.88% as a result (Table 3). However, there are drawbacks to EEG signals, including tedious recording procedure, noisy signals and uncomfortable headgear [45]. Hence, this study also advances on the work of Tor et al [44], as it utilized ECG signals rather than EEG signals. ECG signals can be more easily captured using devices such as smart watches, or less intrusive chest monitors, making this detection approach much more translatable in healthcare settings. As such, the significant aspects of our study can be summarized as follows:

- We are the first study to use the DL model to distinguish between ADHD, ADHD+CD, and CD using ECG signals.
- We obtained high classification accuracy of 96.04%, surpassing our previous work by Koh et al. [23].
- Grad-CAM was also implemented to highlight important ECG characteristics at specific time points.
- Our workflow process is simple and straightforward, eliminating the need for complicated signal processing procedures.

Our study also has its limitations, such as small sample size and class imbalance especially for the CD classes. As a result, the model's ability to generalize is compromised. Furthermore, because ADHD is a complex neurodevelopmental disorder with numerous etiology, there are currently no established ECG biomarkers for it [46]. Because of the lack of standardization in ECG-based ADHD diagnosis, we are unable to employ a single biomarker to diagnose ADHD patients. Nonetheless, this motivates us to incorporate XAI into our model since it identifies prospective ECG biomarkers that physicians and mental health professionals can use as considerations for future research in the discovery of ECG biomarkers for ADHD and CD. The current study also did not include a control or non-ADHD/CD sample and thus did not test the fit of the model in predicting absence of the condition. However, it was first important

to conduct such pilot examinations to identify prospective ECG biomarkers before testing the full utility of the model in clinical and non-clinical samples. We hope that by conducting these pilot studies, we will be able to demonstrate the feasibility of using biosignals to objectively diagnose ADHD and CD, thus encouraging the collection of larger datasets to validate the results obtained in our studies.

As mentioned earlier, the research by Berg et al. [15] observed that medications used for ADHD treatment, including central nervous system stimulants and atomoxetine, carry a heightened potential for sudden cardiac death and irregular heartbeat conditions. As a result, the administration of these medications to individuals with ADHD requires careful consideration, particularly when the patient exhibits cardiac arrhythmias or other heart-related ailments [15], [16]. Consequently, it is advisable that children at an elevated risk receive electrocardiogram (ECG) screening before initiating ADHD therapies [15]. Thus, this pilot study shows the potential for ECG monitoring of ADHD patients during treatment, particularly those with cardiovascular problems. Therefore, for the future of ADHD/CD diagnosis and management, we hope that a larger ADHD and CD biosignal database will one day be made available to the public, allowing future research to develop a computer-aided diagnostic (CAD) tool using biosignals such as ECG and EEG. Future research could also then examine the full utility of this approach in discriminating between clinical and non-clinical samples. Such a tool would undoubtedly benefit mental healthcare facilities, as it would speed up diagnosis procedures and allow for remote patient monitoring during treatment.

5. Conclusion

A small dataset of 45 ADHD, 62 ADHD+CD, and 16 CD patients was used to develop a 1D-CNN model to perform the classification task in this pilot study. The entire experimental setup has a simple workflow process; ECG signals were segmented into 2s length segments and were directly used to train the proposed model. As a result, the model scored highly with a classification accuracy of 96.04%. Furthermore, we implemented Grad-CAM, a XAI technique that can explain the 'black box' CNN models, and it not only highlights the important region of interest on the ECG signals, but also provides temporal information by indicating when the important ECG characteristic occurred. Therefore, this pilot study proposed an explainable DL approach for ADHD and CD diagnosis, which can provide clinicians and mental health professionals with valuable clinical information to help them make informed decisions. In the future, we hope to validate our model with a larger dataset, demonstrating the efficacy of DL models as a useful indispensable CAD tool for the healthcare community.

References

[1]   B. B. LAHEY *et al.*, "Validity of DSM-IV Attention-Deficit/Hyperactivity Disorder for Younger Children," *J. Am. Acad. Child Adolesc. Psychiatry*, vol. 37, no. 7, pp. 695–702, Jul. 1998, doi: 10.1097/00004583-199807000-00008.

[2]   W. Magnus, S. Nazir, A. C. Anilkumar, and K. Shaban, *Attention Deficit Hyperactivity Disorder*. 2021. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/28722868

[3]   M. Lillig, "Conduct Disorder: Recognition and Management.," *Am. Fam. Physician*, vol. 98, no. 10, pp. 584–592, Nov. 2018, [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/30365289

[4]   M. Mordre, B. Groholt, E. Kjelsberg, B. Sandstad, and A. M. Myhre, "The impact of ADHD and conduct disorder in childhood on adult delinquency: A 30 years follow-up study using official crime records," *BMC Psychiatry*, vol. 11, no. 1, p. 57, Dec. 2011, doi: 10.1186/1471-244X-11-57.

[5]   L. S. Goldman, "Diagnosis and Treatment of Attention-Deficit/Hyperactivity Disorder in Children and Adolescents," *JAMA*, vol. 279, no. 14, p. 1100, Apr. 1998, doi: 10.1001/jama.279.14.1100.

[6]   R. Sagar, B. N. Patra, and V. Patil, "Clinical Practice Guidelines for the management of conduct disorder.," *Indian J. Psychiatry*, vol. 61, no. Suppl 2, pp. 270–276, Jan. 2019, doi: 10.4103/psychiatry.IndianJPsychiatry_539_18.

[7]   J. F. Thayer and R. D. Lane, "Claude Bernard and the heart–brain connection: Further elaboration of a model of neurovisceral integration," *Neurosci. Biobehav. Rev.*, vol. 33, no. 2, pp. 81–88, Feb. 2009, doi: 10.1016/j.neubiorev.2008.08.004.

[8]   A. Silvani, G. Calandra-Buonaura, R. A. L. Dampney, and P. Cortelli, "Brain–heart interactions: physiology and clinical implications," *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, vol. 374, no. 2067, p. 20150181, May 2016, doi: 10.1098/rsta.2015.0181.

[9]   A. Sgoifo, N. Montano, C. Shively, J. Thayer, and A. Steptoe, "The inevitable link between heart and behavior," *Neurosci. Biobehav. Rev.*, vol. 33, no. 2, pp. 61–62, Feb. 2009, doi: 10.1016/j.neubiorev.2008.10.007.

[10]   S. Ludyga *et al.*, "The Acute Effects of Aerobic Exercise on Cognitive Flexibility and Task-Related Heart Rate Variability in Children With ADHD and Healthy Controls," *J. Atten. Disord.*, vol. 24, no. 5, pp. 693–703, Mar. 2020, doi: 10.1177/1087054718757647.

[11]   M. De Wied, A. Van Boxtel, J. A. Posthumus, P. P. Goudena, and W. Matthys, "Facial EMG and heart rate responses to emotion-inducing film clips in boys with disruptive behavior disorders," *Psychophysiology*, vol. 46, no. 5, pp. 996–1004, Sep. 2009, doi: 10.1111/j.1469-8986.2009.00851.x.

[12]   N. D. J. van Lang, J. H. M. Tulen, V. L. Kallen, B. Rosbergen, G. Dieleman, and R. F. Ferdinand, "Autonomic reactivity in clinically referred children attention-deficit/hyperactivity disorder versus anxiety disorder," *Eur. Child Adolesc. Psychiatry*, vol. 16, no. 2, pp. 71–78, Mar. 2007, doi: 10.1007/s00787-006-0575-y.

[13]   F. A. Isart, F. G. Ramos, and F. Isart-Infante, "Cardiac Early Repolarization Pattern Anomalies Among Children and Adolescents With and Without Attention-Deficit

Hyperactivity Disorder: A Community Observational Study," *Glob. Pediatr. Heal.*, vol. 6, p. 2333794X1982831, Jan. 2019, doi: 10.1177/2333794X19828311.

[14] C. Sridhar, S. Bhat, U. R. Acharya, H. Adeli, and G. M. Bairy, "Diagnosis of attention deficit hyperactivity disorder using imaging and signal processing techniques," *Comput. Biol. Med.*, vol. 88, pp. 93–99, Sep. 2017, doi: 10.1016/j.compbiomed.2017.07.009.

[15] A. Berg, E. Bråtane, H. H. Odland, C. Brudvik, B. Rosland, and A. Hirth, "[Cardiovascular risk assessment for the use of ADHD drugs in children].," *Tidsskr. Nor. Laegeforen.*, vol. 134, no. 7, pp. 710–4, Apr. 2014, doi: 10.4045/tidsskr.13.0565.

[16] C.-C. Topriceanu, J. C. Moon, G. Captur, and B. Perera, "The use of attention-deficit hyperactivity disorder medications in cardiac disease.," *Front. Neurosci.*, vol. 16, p. 1020961, 2022, doi: 10.3389/fnins.2022.1020961.

[17] O. J. Monfredi *et al.*, "Continuous ECG monitoring should be the heart of bedside AI-based predictive analytics monitoring for early detection of clinical deterioration," *J. Electrocardiol.*, vol. 76, pp. 35–38, Jan. 2023, doi: 10.1016/j.jelectrocard.2022.10.011.

[18] S. Somani *et al.*, "Deep learning and the electrocardiogram: review of the current state-of-the-art," *EP Eur.*, vol. 23, no. 8, pp. 1179–1191, Aug. 2021, doi: 10.1093/europace/euaa377.

[19] N. Salari *et al.*, "Detection of sleep apnea using Machine learning algorithms based on ECG Signals: A comprehensive systematic review," *Expert Syst. Appl.*, vol. 187, p. 115950, Jan. 2022, doi: 10.1016/j.eswa.2021.115950.

[20] A. Singh and D. Kumar, "Computer assisted identification of stress, anxiety, depression (SAD) in students: A state-of-the-art review," *Med. Eng. Phys.*, vol. 110, p. 103900, Dec. 2022, doi: 10.1016/j.medengphy.2022.103900.

[21] S. K. Khare, S. March, P. D. Barua, V. M. Gadre, and U. R. Acharya, "Application of data fusion for automated detection of children with developmental and mental disorders: A systematic review of the last decade," *Inf. Fusion*, vol. 99, p. 101898, Nov. 2023, doi: 10.1016/j.inffus.2023.101898.

[22] H. W. Loh, C. P. Ooi, S. Seoni, P. D. Barua, F. Molinari, and U. R. Acharya, "Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011–2022)," *Comput. Methods Programs Biomed.*, vol. 226, p. 107161, Nov. 2022, doi: 10.1016/j.cmpb.2022.107161.

[23] J. E. W. Koh *et al.*, "Automated classification of attention deficit hyperactivity disorder and conduct disorder using entropy features with ECG signals," *Comput. Biol. Med.*, vol. 140, p. 105120, Jan. 2022, doi: 10.1016/j.compbiomed.2021.105120.

[24] A. Raine *et al.*, "Omega-3 (ω -3) and social skills interventions for reactive aggression and childhood externalizing behavior problems: a randomized, stratified, double-blind, placebo-controlled, factorial trial," *Psychol. Med.*, vol. 49, no. 2, pp. 335–344, Jan. 2019, doi: 10.1017/S0033291718000983.

[25] A. B. Lewin *et al.*, "Utility of the Diagnostic Interview Schedule for Children for

Assessing Tourette Syndrome in Children," *J. Child Adolesc. Psychopharmacol.*, vol. 24, no. 5, pp. 275–284, Jun. 2014, doi: 10.1089/cap.2013.0128.

[26] H. Ge, Z. Zhu, Y. Dai, B. Wang, and X. Wu, "Facial expression recognition based on deep learning," *Comput. Methods Programs Biomed.*, vol. 215, p. 106621, Mar. 2022, doi: 10.1016/j.cmpb.2022.106621.

[27] Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, "Object Detection With Deep Learning: A Review," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 30, no. 11, pp. 3212–3232, Nov. 2019, doi: 10.1109/TNNLS.2018.2876865.

[28] P. Gudžius, O. Kurasova, V. Darulis, and E. Filatovas, "Deep learning-based object recognition in multispectral satellite imagery for real-time applications," *Mach. Vis. Appl.*, vol. 32, no. 4, p. 98, Jul. 2021, doi: 10.1007/s00138-021-01209-2.

[29] P. Lu *et al.*, "A time series image prediction method combining a CNN and LSTM and its application in typhoon track prediction," *Math. Biosci. Eng.*, vol. 19, no. 12, pp. 12260–12278, 2022, doi: 10.3934/mbe.2022571.

[30] S. Soffer, A. Ben-Cohen, O. Shimon, M. M. Amitai, H. Greenspan, and E. Klang, "Convolutional Neural Networks for Radiologic Images: A Radiologist's Guide," *Radiology*, vol. 290, no. 3, pp. 590–606, Mar. 2019, doi: 10.1148/radiol.2018180547.

[31] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology.," *Insights Imaging*, vol. 9, no. 4, pp. 611–629, Aug. 2018, doi: 10.1007/s13244-018-0639-9.

[32] M. Wu, Y. Lu, W. Yang, and S. Y. Wong, "A Study on Arrhythmia via ECG Signal Classification Using the Convolutional Neural Network," *Front. Comput. Neurosci.*, vol. 14, Jan. 2021, doi: 10.3389/fncom.2020.564015.

[33] Y. Zhang, S. Liu, Z. He, Y. Zhang, and C. Wang, "A CNN Model for Cardiac Arrhythmias Classification Based on Individual ECG Signals," *Cardiovasc. Eng. Technol.*, vol. 13, no. 4, pp. 548–557, Aug. 2022, doi: 10.1007/s13239-021-00599-8.

[34] S. K. Khare and U. R. Acharya, "An explainable and interpretable model for attention deficit hyperactivity disorder in children using EEG signals," *Comput. Biol. Med.*, vol. 155, p. 106676, Mar. 2023, doi: 10.1016/j.compbiomed.2023.106676.

[35] L. G. Hafemann, R. Sabourin, and L. S. Oliveira, "Learning features for offline handwritten signature verification using deep convolutional neural networks," *Pattern Recognit.*, vol. 70, pp. 163–176, Oct. 2017, doi: 10.1016/j.patcog.2017.05.012.

[36] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *2017 International Conference on Engineering and Technology (ICET)*, Aug. 2017, pp. 1–6. doi: 10.1109/ICEngTechnol.2017.8308186.

[37] O. Yildirim, U. B. Baloglu, and U. R. Acharya, "A Deep Learning Model for Automated Sleep Stages Classification Using PSG Signals.," *Int. J. Environ. Res. Public Health*, vol. 16, no. 4, 2019, doi: 10.3390/ijerph16040599.

[38] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, no. 1, p. 53, Mar. 2021, doi:

10.1186/s40537-021-00444-8.

[39] H. Mukhtar, S. M. Qaisar, and A. Zaguia, "Deep Convolutional Neural Network Regularization for Alcoholism Detection Using EEG Signals," *Sensors*, vol. 21, no. 16, p. 5456, Aug. 2021, doi: 10.3390/s21165456.

[40] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization," Oct. 2016, doi: 10.1007/s11263-019-01228-7.

[41] V. Jahmunah, E. Y. K. Ng, R.-S. Tan, S. L. Oh, and U. R. Acharya, "Explainable detection of myocardial infarction using deep learning models with Grad-CAM technique on ECG signals," *Comput. Biol. Med.*, vol. 146, p. 105550, Jul. 2022, doi: 10.1016/j.compbiomed.2022.105550.

[42] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," Jan. 2012, [Online]. Available: http://arxiv.org/abs/1201.0490

[43] B. Mirza, W. Wang, J. Wang, H. Choi, N. C. Chung, and P. Ping, "Machine Learning and Integrative Analysis of Biomedical Big Data," *Genes (Basel).*, vol. 10, no. 2, p. 87, Jan. 2019, doi: 10.3390/genes10020087.

[44] H. T. Tor *et al.*, "Automated detection of conduct disorder and attention deficit hyperactivity disorder using decomposition and nonlinear techniques with EEG signals," *Comput. Methods Programs Biomed.*, vol. 200, p. 105941, Mar. 2021, doi: 10.1016/j.cmpb.2021.105941.

[45] P. D. Barua *et al.*, "TMP19: A Novel Ternary Motif Pattern-Based ADHD Detection Model Using EEG Signals," *Diagnostics*, vol. 12, no. 10, p. 2544, Oct. 2022, doi: 10.3390/diagnostics12102544.

[46] H. W. Loh, C. P. Ooi, P. D. Barua, E. E. Palmer, F. Molinari, and U. R. Acharya, "Automated detection of ADHD: Current trends and future perspective," *Comput. Biol. Med.*, vol. 146, p. 105525, Jul. 2022, doi: 10.1016/j.compbiomed.2022.105525.