Developing Multimodal Deep Learning Framework Models for the Diagnosis and Analysis of Attention-Deficit/Hyperactivity Disorder (ADHD) using physiological data

ZHIVKO PARAPANOV

THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
BACHELOR OF SCIENCE IN COGNITIVE SCIENCE AND ARTIFICIAL INTELLIGENCE
AT THE SCHOOL OF HUMANITIES AND DIGITAL SCIENCES OF TILBURG UNIVERSITY

STUDENT NUMBER
2074677

COMMITTEE
dr. Silvy Colin
dr. Eriko Fukuda

LOCATION
Tilburg University
School of Humanities and Digital Sciences
Department of Cognitive Science & Artificial Intelligence
Tilburg, The Netherlands

DATE
May 21th, 2024

WORD COUNT
7331

ACKNOWLEDGEMENTS

I want to extend my heartfelt appreciation to everyone who contributed to the success of my thesis project. In particular, I am grateful to my supervisor, Prof. Dr. Silvy Colin, for her invaluable guidance and mentorship throughout the process. Additionally, I am thankful to my family and friends for their unwavering support, encouragement, and assistance. Their presence and understanding have been essential in helping me overcome challenges and achieve my goals.

Developing Multimodal Deep Learning Framework Models for the Diagnosis and Analysis of Attention-Deficit/Hyperactivity Disorder (ADHD) using physiological data

Zhivko Parapanov

Abstract

Attention-Deficit/Hyperactivity Disorder (ADHD) is a prevalent neurodevelopmental disorder traditionally diagnosed through subjective and time-consuming assessments. This research investigates the potential of multimodal deep learning frameworks to enhance ADHD diagnosis by leveraging physiological data, specifically heart rate variability (HRV) and motor activity data from the HYPERAKTIV dataset. Three distinct frameworks have been explored: a CNN-LSTM dual-branch model, a dual-branch 1D CNN model, and a variable-length dual-branch 1D CNN model. The dual-branch 1D CNN models outperform the CNN-LSTM model, achieving a test accuracy of 85.71%. The results highlight the effectiveness of capturing unique patterns from HRV and motor activity data streams, contributing to improved diagnostic accuracy. The research's findings underscore the potential of deep learning in enhancing ADHD diagnosis by leveraging physiological data and advanced model architectures, offering a more objective and data-driven approach than traditional methods.

Keywords: *Attention-Deficit/Hyperactivity Disorder, Deep Learning, Multimodal Frameworks, Physiological Data, Diagnosis*

DATA SOURCE, ETHICS, CODE AND TECHNOLOGY STATEMENT

The HYPERAKTIV dataset has been acquired from https://www.kaggle.com/datasets/arashnic/adhd-diagnosis-data?rvi=1&select=CPT_II_ConnersContinuousPerformanceTest.csv. The obtained data is anonymised. Work on this thesis did not involve collecting data from human participants or animals. The original owner of the data and code used in this thesis retains ownership of the data and code during and after the completion of this thesis. All the figures except Figure 1 belong to the author. To use Figure 1, the author executed open-source code created by the author that can be found at https://github.com/simula/hyperaktiv. The thesis code can be accessed through the GitHub repository following the link: https://github.com/Jivko26/2024_TilburgUniversity_Thesis.
In terms of writing, AI tools like Grammarly, Gemini and ChatGPT were utilised to refine the paper's language by paraphrasing, spell-checking, and correcting grammar. No additional tools or services were used in the research.

1. Introduction

Attention-Deficit/Hyperactivity Disorder (ADHD) is a common neurodevelopmental disorder with a complex diagnostic process. Traditional diagnostic methods combine clinical interviews, behavioural assessments, and sometimes cognitive performance tests (Hicks et al., 2021). This often involves subjective evaluations and can be time-consuming. More objective, efficient, and reliable diagnostic tools are needed to streamline the process, improve patient outcomes, and reduce healthcare burdens. Recent advancements in machine learning and deep learning show promise in diagnosing and understanding complex health conditions. Within objective ADHD diagnostics, support vector machines have shown promising abilities to discriminate between children with ADHD and healthy controls in movement data from accelerometers and gyroscopes. In neurological and psychiatric settings, studies have explored these methods, using data from electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) (Riaz et al., 2018; Fernández et al., 2009). These approaches provide valuable insights, but their reliance on specialised equipment and controlled environments can limit accessibility. Moreover, misdiagnosis remains a concern within current ADHD diagnostics (Willcutt, 2012).

This research proposes to investigate a multimodal deep learning framework for ADHD diagnosis using less commonly explored physiological data: heart rate variability (HRV) and motor activity. Understanding how these dynamic physiological measures change in the presence of ADHD could unveil unique biomarkers. By harnessing the capabilities of deep learning to identify patterns in these complex data streams, this approach can potentially refine ADHD diagnosis and contribute to developing personalised and efficient diagnostic tools.

1.1 Research Questions

The thesis explores a multimodal deep learning framework's potential in improving the diagnostic accuracy of Attention-Deficit/Hyperactivity Disorder (ADHD). The central research question driving the study is:

- ***RQ:*** *How can integrating heart rate variability and motor activity data within a multimodal deep learning framework enhance the diagnostic process for ADHD, offering improvements over traditional diagnostic approaches?*

To explore this question thoroughly, the following sub-questions will be examined:

- ***RQ 1.1:*** *How does analysing time-series data from HRV and motor activity, using deep learning models, contribute to identifying biomarkers associated with ADHD?*
- ***RQ 1.2:*** *What role does the fusion of multiple physiological data streams play in increasing the accuracy and reliability of the diagnostic model?*

To answer these questions, the thesis will investigate the rich time-series data of motor activity and heart rate alongside cognitive performance insights, all found within the HYPERAKTIV dataset (Hicks et al., 2021). The potential of deep learning models in synthesising data from multiple sources within this framework will be explored to uncover insights into ADHD's physiological manifestations.

2. Related Work

Attention-deficit/hyperactivity disorder (ADHD) is a widespread neurodevelopmental disorder affecting a significant percentage of children (Wolraich et al., 2019; Danielson et al., 2018). The understanding of ADHD has evolved considerably since its initial description by George Still in 1902. Still attributed the disorder to a *"defect in moral character"* and observed hallmark behaviours such as fidgeting, inattention, and excessive alertness. Early research explored potential links to brain damage, particularly after the influenza epidemic in 1917-1918, but subsequent cases lacking evident neurological damage led to revisions in terminology. The focus gradually shifted from hyperactivity to inattention, ultimately culminating in the current classification as *"attention-deficit/hyperactivity disorder"* in the DSM-III-Revised (Douglas et al. (1974).

2.1 Diagnosis

Despite the evolving diagnostic criteria for ADHD, the methods and tools used for evaluation have remained relatively unchanged. ADHD continues to be predominantly diagnosed based on clinical judgement. Current recommendations emphasise gathering comprehensive information through a detailed history, including prenatal, perinatal, and family history, school performance, environmental influences, and a thorough physical examination (Wolraich et al., 2019). During the physical examination, particular attention should be paid to vital signs (cardiovascular, skin, thyroid, and neurologic systems, including assessment of motor coordination), and a mental health assessment should be performed to probe for comorbid conditions. Teacher- or parent-reported behaviour-rating scales started in the late 1960s (Conners, 1969). The focus nowadays is on the behavioural criteria for ADHD as described in the DSM-5 (Douglas et al. 1974). Determining significant impairment is also an important criterion. The available objective assessments for ADHD, such as neuropsychological tests, EEG and neuroimaging, have limited effectiveness in aiding diagnostic clarity (Kemper et al., 2018). Scientists have also explored the role of neurotransmitter systems in ADHD, as animal models, neuroimaging, and pharmacologic investigations suggest the involvement of dopaminergic and adrenergic disruptions in the condition (Del Campo et al., 2011). However, no evidence-based methods for assessing these neurotransmitter systems have been developed and shown to have utility in the ADHD diagnostic assessment. Despite the high heritability of ADHD (~76%), no specific genetic profile has been identified as a necessary or sufficient cause of the disorder, so identifying genes linked to ADHD has been difficult, and genetic testing is currently not functional for diagnostic purposes (Thapar et al., 2018).

2.2 Machine Learning for ADHD Diagnosis

Machine learning (ML) has emerged as a promising tool for aiding in the diagnosis of ADHD (Hicks et al., 2021). Various studies have explored the use of ML algorithms to analyse neuroimaging data, behavioural assessments, and genetic information to identify patterns that distinguish individuals with ADHD from those without the disorder. The potential benefits of ML in ADHD diagnosis include improved accuracy, objectivity, and efficiency compared to traditional diagnostic methods. Machine learning (ML) can be applied to analyse activity data to diagnose ADHD.

A study by Hicks(2021) utilised the HYPERAKTIV dataset, which includes health, activity, and heart rate data from ADHD patients and clinical controls. Four machine-learning models were trained on features extracted from the activity data. The results revealed that all four models outperformed simple prediction rules. Logistic regression emerged as the best-performing model, achieving moderate predictive ability. The study demonstrates the potential of ML for ADHD prediction using activity data, suggesting that activity patterns may contain valuable information for identifying individuals with ADHD. However, further research is needed to improve the accuracy and robustness of ML models in this domain.

Another area for applying ML in ADHD diagnosis is analysing electroencephalography (EEG) data. EEG measures electrical activity in the brain and has been used to study various neurological and psychiatric disorders, including ADHD. A study by Tenev et al. (2014) explored using support vector machines (SVM), a ML algorithm, to classify individuals with ADHD from healthy controls based on EEG data. In their research, Tenev et al. (2014) collected EEG data from participants under four different conditions: eyes closed resting state (EC), eyes open resting state (EO), continuous performance test (CPT), and reaction time test (RT). They then preprocessed the data and applied a forward selection scheme to identify the most relevant features for each condition. These features were used to train separate SVM classifiers for each condition. The outputs of the individual SVM classifiers were then combined using a logical voter function, which considers each classifier's performance to make a final prediction about whether an individual has ADHD or not. The voter function was designed to improve the overall classification accuracy by leveraging the information from multiple conditions. The study results showed that the SVM classifier combined with the voter function achieved a higher classification accuracy than any individual classifiers alone. This suggests that combining information from multiple EEG conditions can improve the ability of ML algorithms to diagnose ADHD.

Furthermore, the study found that the classifier was able to discriminate between different ADHD subtypes (ADHD II, ADHD III, and ADHD IV) with even higher accuracy than between ADHD and healthy controls. This indicates that EEG data may contain valuable information about the specific subtypes of ADHD, which could have implications for personalised treatment approaches.

The use of machine learning algorithms, such as SVM, along with EEG data, represents a promising direction for improving the diagnosis and understanding of ADHD. Future research could explore using other ML algorithms, more extensive and diverse datasets, and different feature extraction techniques to further refine these approaches' accuracy and clinical utility.

2.3 Convolutional Neural Networks and LSTM Networks for ADHD Diagnosis

Convolutional Neural Networks (CNNs) are particularly effective for image and spatial data analysis, making them suitable for neuroimaging applications in ADHD diagnosis. A recent study by Ahmadi et al. (2021) proposed a novel computer-aided diagnosis system using deep CNNs to classify EEG signals from healthy children and those with ADHD, explicitly targeting the subtypes of Combined ADHD (ADHD-C) and Inattentive ADHD (ADHD-I). The study utilised a deep convolutional neural network to extract spatial and frequency band features from raw EEG signals. The CNN model included multiple layers designed to process EEG data effectively. The text describes a deep learning model for EEG signal classification. The model consists of several layers: an input layer, a band-filter layer, a spatial-filtering layer, a batch normalisation and activation layer, an average pooling layer, a dropout layer, a temporal-filter layer, a flattened layer, and a dense layer. The model takes an input tensor of shape (1, 2500, 19) and outputs a vector of length 3, representing the probabilities of the three classes' signals. The model achieved remarkable performance, with the highest classification accuracy observed when combining β1, β2, and γ bands. The classification accuracy, recall, precision, and Cohen's kappa values were 99.46%, 99.45%, 99.48%, and 0.99, respectively. The study's novelty lies in its use of deep CNN architecture for multiclass classification of ADHD and healthy children using EEG signals, achieving high accuracy and providing valuable insights into the contributions of different EEG frequency bands.

2.4 LSTM for ADHD Diagnosis

Long Short-Term Memory Networks (LSTMs), a type of recurrent neural network (RNN), are designed to model sequential data and capture long-range dependencies, making them well-suited for analysing time-series data such as HRV and motor activity. LSTMs have shown promise in identifying temporal patterns associated with ADHD symptoms. A recent study by Saurabh and Gupta (2024) focused on the use of a modified deep learning-based bidirectional long short-term memory (BLSTM) model to classify ADHD using resting-state functional magnetic resonance imaging (rs-fMRI) data. The research analysed the functional connectivity of 40 subjects (20 ADHD and 20 healthy controls) through voxel size blood-oxygen-level-dependent (BOLD) signals, which are functionally relevant to corresponding resting state networks (RSNs). The proposed BLSTM model automated the classification of ADHD by identifying voxels within the active regions of the RSNs. Initially, the study visualised the 28 active regions of the RSNs and a time series of behavioural data for 40 subjects with 176 timestamps. The modified BLSTM was trained using a feature vector ($40 \times 261 \times 28$) for each subject and optimised with the Adam hyper-parameter. The experimental results demonstrated that the proposed model outperformed many other models, achieving a classification accuracy of

87.50%. The study also provided a detailed comparative analysis of the proposed model with various existing state-of-the-art approaches.

2.5 Multimodal Deep Learning Frameworks

Multimodal deep learning involves integrating multiple data modalities (e.g., physiological signals, neuroimaging, behavioural data) to improve the robustness and accuracy of predictive models. This approach leverages the complementary information in different data types to enhance the diagnostic process. Recent advances have demonstrated the potential of multimodal frameworks in providing comprehensive insights into complex disorders like ADHD. In ADHD diagnosis, multimodal frameworks can integrate various types of data, such as heart rate variability (HRV), motor activity, and neuroimaging, to develop comprehensive models that capture the multifaceted nature of the disorder. By combining these modalities, researchers can achieve more accurate and reliable diagnoses. A study by Wang et al. (2022) developed a CNN-LSTM model to classify ADHD by integrating EEG data with other physiological markers. The model effectively identified ADHD and its subtypes, achieving a high classification accuracy of 98.23% in a five-fold cross-validation. This research underscores the power of combining convolutional neural networks (CNNs) for spatial feature extraction with long short-term memory (LSTM) networks for temporal pattern recognition, enhancing the diagnostic process through multimodal integration. By leveraging the temporal dynamics captured through LSTM networks and the spatial hierarchies identified by CNNs, the proposed framework offers a comprehensive diagnostic tool that addresses the limitations of traditional single-modality approaches.

Integrating these advanced techniques aims to provide a more robust and holistic method for early and accurate ADHD diagnosis, contributing to the growing body of research advocating for multimodal deep learning in medical diagnostics.

2.6.Research Gap and Proposed Approach

While existing research has made significant strides in utilising machine learning for ADHD diagnosis, a gap remains in effectively harnessing the potential of multimodal data integration. While studies like Wang et al. (2022) have begun to explore this avenue, there is a lack of comprehensive investigations into the synergistic use of physiological signals like HRV and motor activity alongside other data types, such as neuroimaging or genetic markers. Furthermore, the optimal deep learning architectures for integrating these diverse modalities remain underexplored.

This thesis aims to address this research gap by developing and evaluating multiple multimodal deep learning frameworks for ADHD diagnosis, explicitly focusing on integrating HRV and motor activity data. By exploring different architectures, the research seeks to determine the most effective approach for capturing the complex patterns inherent in these physiological signals. Moreover, this study aims to investigate the specific contributions of HRV and motor activity data to the diagnostic process, potentially revealing novel biomarkers or physiological patterns associated with ADHD.

3. Methods

This study aimed to develop a multimodal deep learning framework to diagnose ADHD accurately and efficiently by leveraging physiological data such as heart rate variability (HRV) and motor activity. This approach seeks to overcome the limitations of traditional diagnostic methods, which often rely on subjective assessments and can be time-consuming and inconsistent.

3.1 Dataset Description

The HYPERAKTIV dataset (Hicks et al., 2021) is a comprehensive collection of physiological data from participants with and without ADHD. The dataset includes a time series of motor activity and heart rate, results from a neuropsychological computer test, conclusions, sum scores from various diagnostic assessment tools, and demographic information such as gender, age, and prescribed medications. Sex was recorded as zero for females and one for males. Participant ages were categorised into four age groups: 1 (17-29 years), 2 (30-39 years), 3 (40-49 years), and 4 (50-67 years). Of the 85 patients who recorded motor activity, 23 belonged to age group 1, 26 to age group 2, 24 to age group 3, and 12 to age group 4. Most participants were not on medication, with 73% of ADHD-diagnosed individuals who provided motor activity recordings being unmedicated.

The dataset includes the following components:

- ***Activity Data***: Each file contains activity data for a single participant, starting with metadata followed by detailed activity measurements.
- ***HRV Data:*** Files contain heart rate data for each participant, beginning with metadata and followed by inter-beat interval (IBI) values.

- ***CPT-II Test Results***: Individual responses from 360 CPT-II test trials, omission and commission errors, and the ADHD Confidence Index.
- ***Pre-extracted Features***: Each line in the file corresponds to features for a single participant.
- ***Participant Information*****:** Contains 32 columns with demographic and clinical information for each participant.

Motor activity was recorded using a wrist-worn Actiwatch device (Hicks et al., 2021), which captured movement intensity across three axes at a sampling frequency of 32 Hz. Heart rate data was collected using a chest-worn Actiheart device (Hicks et al., 2021), recording IBI in milliseconds. Both data types were essential for analysing physiological patterns associated with ADHD. Activity data from one day is visualised in Figure 1.

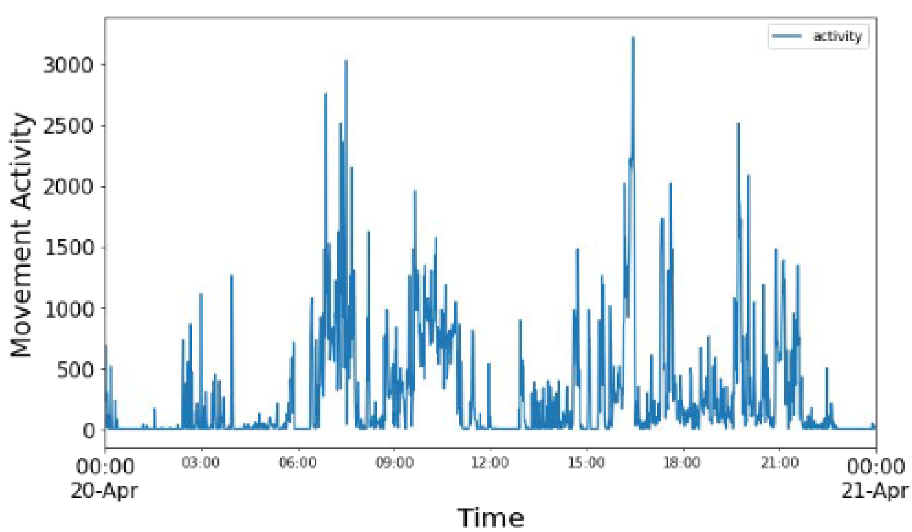

Figure 1. Example of 24 hours (from midnight to midnight) of actigraphy recordings from participant no. 57. Female, age group 17-29, diagnosed with ADHD, unipolar depression, anxiety disorder and cyclothymic temperament.

Clinical assessments included the Mini-International Neuropsychiatric Interview (MINI Plus), Adult ADHD Self-Report Scale (ASRS), Wender Utah Rating Scale (WURS), Mood Disorder Questionnaire (MDQ), Cyclothymic temperament scale (CT), Montgomery and Asberg Depression Rating Scale (MADRS), and Hospital Anxiety and Depression Scale (HADS). For example, the ASRS is an 18-item scale that evaluates current symptoms of ADHD, with higher scores indicating more severe symptoms. The WURS assesses the presence and severity of childhood ADHD symptoms, with scores ranging from 0 to 100. The MDQ, a self-reported screening instrument for bipolar disorder, includes 13 yes/no questions about hypomanic/manic symptoms. Other scales, such as the CT, MADRS, and HADS, evaluate emotional instability, depression severity, and anxiety and depression levels, respectively.

The dataset showcased a well-balanced distribution between ADHD and non-ADHD cases, as evident in Figure 2. This balance is crucial for the model's performance, preventing biases towards dominant classes. The model can effectively learn from both ADHD and non-ADHD cases, reducing the risk of overfitting to a particular class. This ensures that the model generalizes well to new data. Additionally, the diverse representation of gender, age groups, and medication status enhances the model's robustness, making it applicable to a broader population.

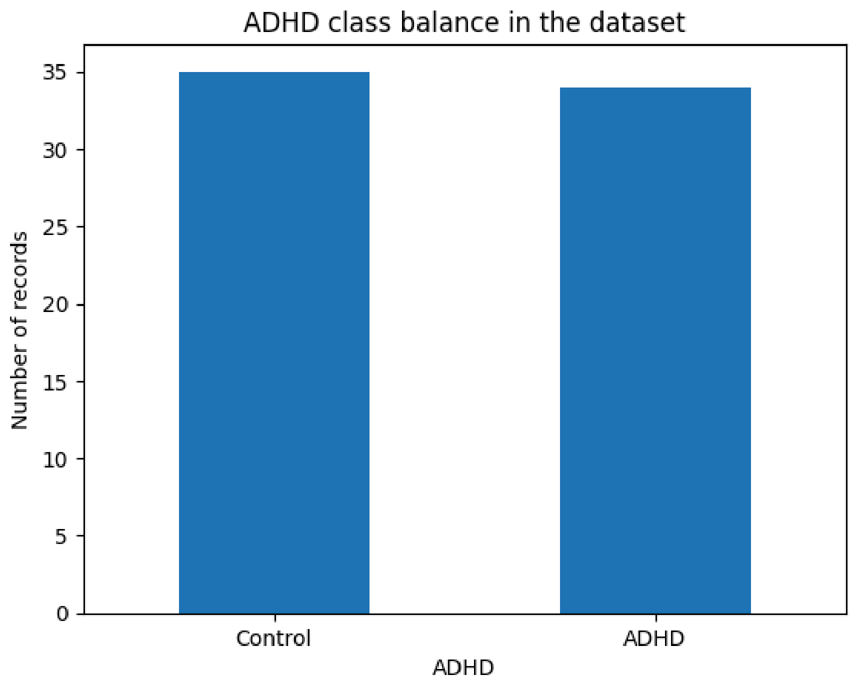

Figure 2. Balance of the target variable(ADHD) in the dataset

3.1.1 Addressing Comorbid Conditions and ADHD Subtypes

In the HYPERAKTIV dataset, participants diagnosed with ADHD were included alongside those with comorbid conditions. In this study, participants with comorbid conditions were incorporated into the ADHD group for the main reason of using the same sample size and records when recreating the experiment with the proposed architecture. This way, the comparison between both approaches would be consistent and objective. Moreover, there are other factors that influence this decision:

- ***Real-World Representation:*** Comorbid conditions are common in individuals with ADHD (Katzman et al., 2017). Excluding them would create an artificial dataset that does not reflect the clinical reality of ADHD.

- ***Potential for Biomarker Discovery:*** The interplay between ADHD and comorbid conditions might manifest in unique physiological patterns. Including these participants could reveal novel biomarkers relevant to a broader range of individuals.

However, the influence of comorbid conditions on physiological data was acknowledged. Therefore, the results were interpreted cautiously, recognising that ADHD and co-occurring conditions might influence the identified patterns. Due to the limited sample size and the dataset's structure, specific ADHD subtypes (e.g., predominantly inattentive or hyperactive-impulsive) were not analysed in this study. Future research with larger datasets should aim to investigate subtype-specific physiological patterns.

3.2 Data Preprocessing

Data preprocessing is a critical step in any machine learning project, and it is particularly crucial when dealing with physiological data like HRV and motor activity. These data types are often noisy, irregular, and contain missing values, which can significantly impact the performance and reliability of machine learning models (Kotsiantis et al., 2006). Preprocessing helps to ensure that the data is in a suitable format for analysis, enhancing the model's ability to extract meaningful patterns and make accurate predictions

It involved several stages: error handling, data cleaning, scaling, and missing values. Error handling involved identifying and correcting anomalies and outliers in the dataset to maintain data consistency. Data cleaning ensured that incomplete records and irrelevant features were discarded, enhancing the model's accuracy and efficiency. Scaling was performed using the Robust Scaler, which is more resilient to outliers than the Standard Scaler, making it suitable for the variability inherent in physiological data. The percentage of missing data varied across features in the HYPERAKTIV dataset. The specific percentages were not explicitly stated in the original dataset documentation. In contrast, 51 % of the patients have undergone activity and HRV monitoring and only they were used for the following steps. Among them, there was no missing data. Resampling and interpolation techniques were implemented to ensure reliable and accurate model training and evaluation to account for the physiological data's inherent variability and potential inconsistencies. This resulted in missing data, which was imputed with mean values. These methods facilitated the creation of a consistent and normalised dataset, a crucial step in obtaining meaningful and reliable results.

Data splitting was another essential aspect of the preprocessing stage. The dataset was split into training (80%), validation (10%), and test (10%) sets to create balanced class distributions across the sets. This stratified sampling method helped maintain a consistent proportion of each class label (ADHD and control) in all subsets, preventing bias during model training and evaluation. This fair data split is crucial for the model to learn effectively and to generalise well to unseen data. The fairness of these splits was verified through visualisations, which displayed the distribution of ADHD and control cases in each set. The graphs illustrating the class distributions in the training, validation, and test sets are shown in Figure 3, respectively.

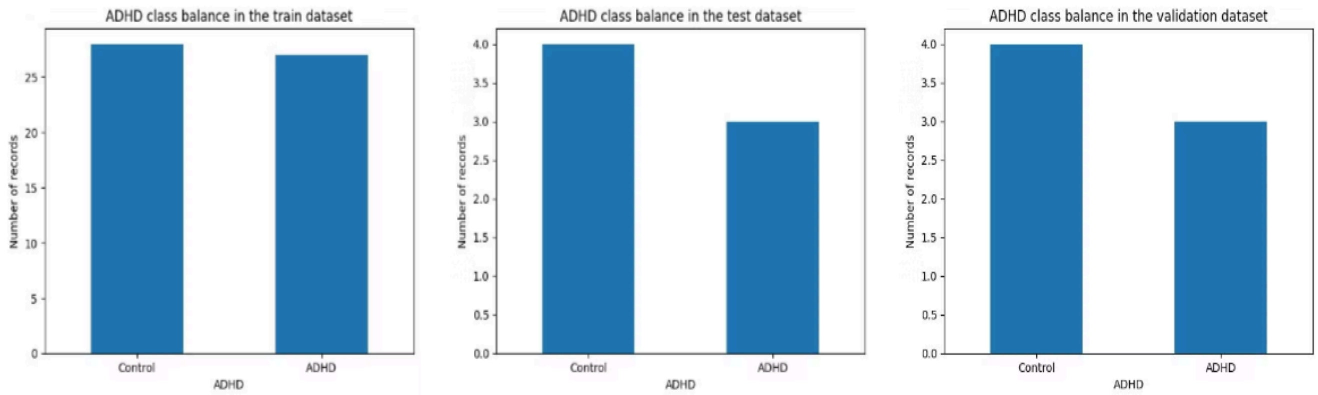

Figure 3. Target Class Balance among training, test and validation sets

3.3 Model Development

Three different multimodal deep learning frameworks were developed to explore the best approach for ADHD diagnosis. Each framework was designed to leverage different aspects of deep learning techniques to process and integrate HRV and motor activity data effectively.

3.3.1 Framework 1: CNN-LSTM Dual Branch

The CNN-LSTM framework leverages the strengths of both architectures. The CNN extracts spatial features from HRV and activity data, while the LSTM captures temporal dependencies. This combination is suitable for ADHD diagnosis because it allows the model to understand complex spatial-temporal patterns inherent in physiological data. Batch normalisation and dropout layers enhance the model's generalisation ability, and hyperparameter tuning optimises performance.

Implementing this framework involved several key steps. First, each patient's HRV and activity data are loaded and resampled to a 1-second interval. Any missing values are handled by forward-filling, ensuring data continuity. Next, a RobustScaler is applied to normalise both the HRV and activity data, which helps to stabilise the training process and improve convergence. The data is then segmented using a sliding window approach, with a window size and a step size (window size // 2), creating overlapping segments that capture temporal dynamics. The window size was also part of the hyperparameters space used for tuning. The model's architecture was designed to process the multimodal input data effectively. It consisted of two branches, one for HRV data and one for activity data. Each branch included a CNN layer for feature extraction and an LSTM layer to model temporal dependencies. The outputs of these two branches were then concatenated and passed through fully connected layers for final classification. The model's training involved experimenting with various hyperparameter configurations, including different learning rates, batch sizes, dropout rates, and optimiser types. The model's performance was evaluated using a set of standard metrics: accuracy, precision, recall, and F1-score. These metrics provided insights into the model's ability to correctly classify ADHD cases, minimise false positives and negatives, and balance precision and recall.

3.3.2 Framework 2: Dual-Branch 1D CNN

The second framework developed for the current research introduced a dual-branch 1D CNN approach to multimodal data integration. Each network branch independently extracted features from the HRV and activity data streams, effectively leveraging the strengths of CNNs in feature extraction from time-series data. The features extracted from each branch were then concatenated and processed through fully connected layers, ultimately leading to the final prediction.

This approach provided a robust model for ADHD classification by capitalising on the complementary information in HRV and activity data. The implementation details are as follows: The raw HRV and activity data were first loaded and scaled using a robust scaler to mitigate the impact of outliers. To ensure uniformity in input length, sequences were truncated or padded to a global maximum length determined during data analysis. The dataset was splitted into training and validation sets using stratified sampling to maintain balanced class distributions and enable practical model evaluation. This ensured that both sets have a representative proportion of each class label. The core of Framework 2 was its dual-branch architecture (Appendix A). Each branch consisted of multiple convolutional layers to capture the data's local patterns and temporal dependencies. These convolutional layers were interspersed with max-pooling layers, reducing the feature maps' dimensionality while retaining the most salient information. The output of each branch was then flattened and passed through fully connected layers, which learned complex relationships between the extracted features. Finally, the outputs of both branches were concatenated and fed into additional fully connected layers for final classification.

The model was trained using the Adam optimiser or stochastic gradient descent (SGD) optimiser. The choice of optimiser and the associated hyperparameters were determined through a hyperparameter tuning process to ensure optimal model performance. A comprehensive evaluation of the model's performance was conducted using standard metrics. These metrics include accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (ROC-AUC). The F1-score, in particular, provided a balanced measure of precision and recall, while the ROC-AUC assessed the model's ability to discriminate between classes across different thresholds. The dual-branch structure allowed for the independent extraction of features from HRV and activity data. This parallel processing enhanced the model's ability to learn from multimodal inputs, capturing distinct patterns and relationships that may be present in each data modality. Using convolutional layers in model architecture promoted efficient training. These layers exploited the data's inherent spatial and temporal correlations, reducing the number of parameters to be learned and enabling scalability to large datasets.

3.3.3 Framework 3: Variable Length Dual-Branch 1D CNN

Framework 3 extends the dual-branch 1D CNN to handle variable-length sequences, making it more flexible and better suited for real-world data where sequence lengths vary. The model processes each sequence independently, padding them to the maximum length within each batch, and uses adaptive average pooling to handle different sequence lengths effectively.

Regarding implementation, the data loading and splitting processes are similar to Framework 2. The data is loaded and scaled appropriately, with the added consideration of variable-length sequences, which are managed using adaptive average pooling within the model itself. The dataset is split into training and validation sets using stratified sampling to maintain balanced class distributions. The model architecture builds upon the dual-branch 1D CNN structure of Framework 2, incorporating adaptive pooling layers to address the challenge of variable-length sequences. This modification allows the model to process sequences of different lengths without sacrificing information, ultimately contributing to improved accuracy and better generalisation. The training process involves using either the Adam or SGD optimiser, with hyperparameters meticulously tuned to achieve optimal performance. The evaluation of the model's performance encompasses a range of metrics, including accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (ROC-AUC). Special attention is given to the ROC-AUC metric, as it comprehensively assesses the model's ability to discriminate between classes across various thresholds. The model summary is shown in Appendix B.

The introduction of adaptive pooling in Framework 3 addresses the limitations of fixed-length sequences encountered in previous frameworks. By accommodating variable-length sequences, the model becomes more adaptable to real-world data, where variations in sequence length are common. Furthermore, preserving information from sequences of varying lengths contributes to enhanced accuracy and improved generalisation, as the model can learn from a broader range of data patterns.

3.4 Implementation

Model implementation was done using Python 3.8 (Python Software Foundation, Wilmington, DE USA), PyTorch 1.9 (Paszke et al., 2019), and Scikit-learn (Pedregosa et al., 2011). Jupyter Notebooks and Google Colab were used for development and testing. K-fold cross-validation (K = 5) was applied to ensure robust model performance. This method divides the dataset into K subsets, training the model K times, each using a different subset as the validation set and the remaining K-1 subsets as the training set.

The chosen tools and methods were selected for their compatibility with the research goals, providing a comprehensive environment for developing and evaluating complex deep learning models. These tools ensured the models were trained efficiently, with access to necessary computational resources and a streamlined development process.

3.5 Evaluation Criteria

The performance of the models developed was assessed using several evaluation metrics. Accuracy measures the overall correctness of the model by calculating the ratio of correctly predicted instances to the total instances. Precision indicates the accuracy of optimistic predictions, representing the proportion of true positive instances among all instances predicted as positive. Recall measures the model's ability to identify all relevant instances, indicating the proportion of true positive instances among all actual positive instances. The F1-Score is the harmonic mean of precision and recall, providing a single metric that balances the two and is particularly useful for imbalanced datasets. The Area Under the Receiver Operating Characteristic Curve (ROC-AUC) evaluates the model's performance across different thresholds, reflecting its ability to distinguish between classes. These evaluation metrics provided a comprehensive understanding of each model's performance, allowing for effective comparison and validation of the multimodal frameworks.

Using multiple metrics ensured a balanced assessment, addressing different aspects of model performance and providing insights into the strengths and weaknesses of each approach. Integrating HRV and motor activity data within advanced deep learning frameworks presents a promising approach to enhancing ADHD diagnostic accuracy. The thorough preprocessing and rigorous model development methods ensure reproducibility and reliability, contributing valuable insights into the physiological markers of ADHD.

By leveraging multimodal data and sophisticated algorithms, this study aims to refine ADHD diagnosis and develop more personalised and efficient diagnostic tools. The methodologies and evaluations applied in this research underscore the potential of deep learning in medical diagnostics, highlighting the importance of integrating diverse data sources to capture the complexity of neurodevelopmental disorders like ADHD. The results from the various models highlight the potential benefits of using multimodal data and sophisticated algorithms in medical diagnostics. The methods and preprocessing steps were rigorously designed to ensure the reproducibility and robustness of the findings. This comprehensive approach aims to improve patient outcomes by providing more objective, efficient, and reliable diagnostic tools for ADHD.

4. Results

4.1 Framework 1: CNN-LSTM Dual Branch

The CNN-LSTM model aimed to capture complex spatial-temporal patterns in the data. The best-performing configuration used the SGD optimiser with a learning rate 0.001, 10 epochs, batch size 16, and dropout regularisation at 0.5. This configuration achieved a validation accuracy of 60%, precision of 0.27, recall of 0.05, and an F1-score of 0.06. However, the model's performance on the test set, as illustrated in Figure 4, indicated significant overfitting, with a test accuracy of 54%, precision of 0.20, recall of 0.05, and an F1-score of 0.08. The ROC-AUC score of 0.46 further highlighted the model's limited generalizability.

Despite promising validation results, the discrepancy between validation and test performance suggested that the model learned noise in the training data rather than the underlying patterns. Overfitting is common in deep learning, particularly with small datasets (Hawkins, 2004). Future research should focus on strategies to mitigate overfitting, such as data augmentation, regularisation techniques, or exploring different model architectures. Additionally, transfer learning from pre-trained models on larger datasets could improve the model's ability to generalise to new data (Pan and Yang, 2010).

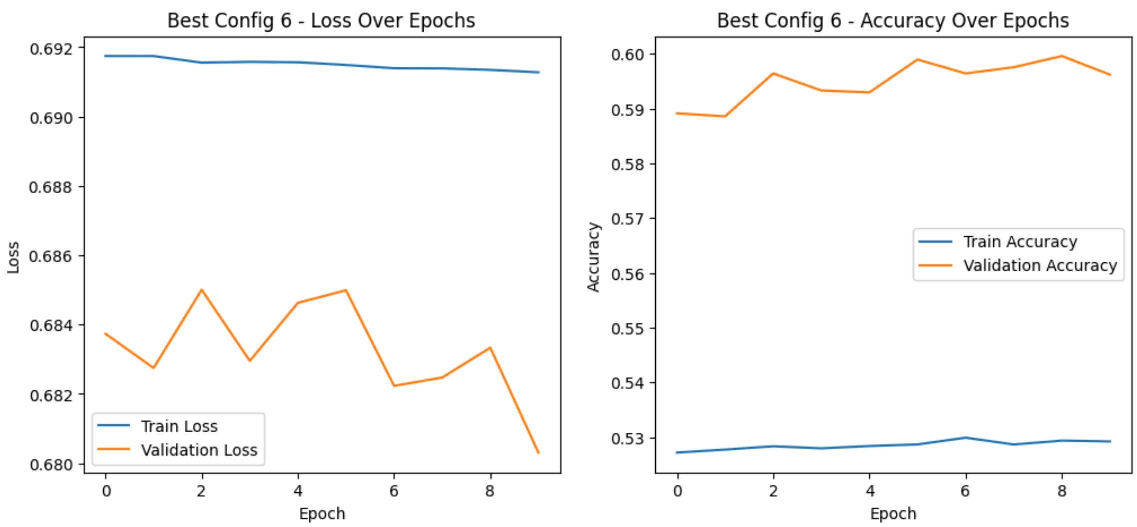

Figure 4. Training and Validation Loss Curves of the CNN-LSTM Dual Branch

4.2 Framework 2: Dual-Branch 1D CNN

Implementing the Dual-Branch 1D CNN model, Framework 2, on the HYPERAKTIV dataset, yielded promising results. Following a comprehensive hyperparameter tuning process, the best-performing model configuration utilised the SGD optimiser with a learning rate of 0.01, 20 epochs, a batch size 16, dropout regularisation with a rate of 0.2, and batch normalisation. This configuration led to a validation accuracy of 57.14%. Using SGD with momentum and dropout aligns with best practices in deep learning for optimising generalisation performance (Ruder, 2016). Batch normalisation is also widely used to accelerate training and improve model stability (Ioffe & Szegedy, 2015). The model's performance metrics can be seen in Appendix C. The training and validation loss curves are depicted in Figure 5.

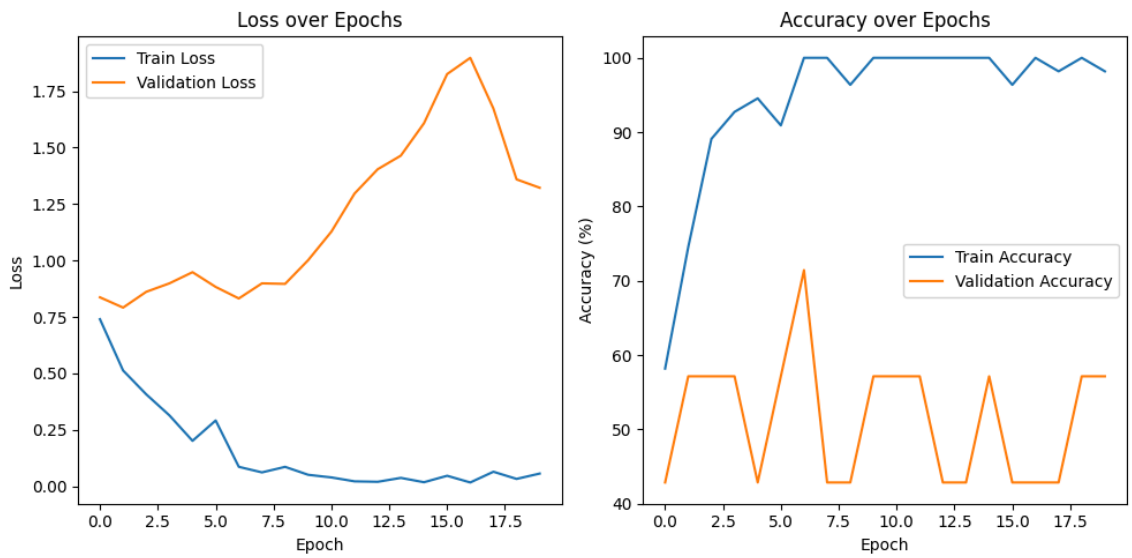

Figure 5. Training and Validation Loss Curves of the Dual-Branch 1D CNN

Remarkably, the model's performance on the unseen test set surpassed the validation metrics, achieving a test accuracy of 85.71%. This substantial improvement in validation accuracy can be attributed to several factors, including the small validation set size and potential differences in sample difficulty. The F1-score of 0.86 demonstrates a balanced performance between precision and recall. The ROC curve, depicted in Figure 6, further illustrates the model's discriminatory power. The model's precision was 0.75, and its recall was 1.00, signifying that it correctly identified all ADHD cases in the test set (see Figure 7). These results suggest that the dual-branch architecture effectively leverages the complementary information in HRV and motor activity data. The parallel processing of these modalities enables the model to capture distinct patterns and relationships within each data stream, contributing to improved classification accuracy.

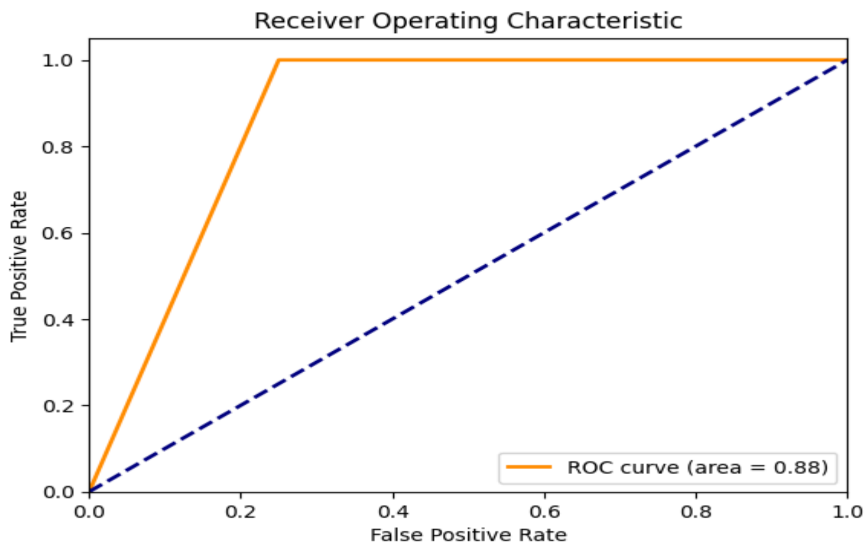

Figure 6. ROC Curve of Dual-Branch 1d CNN

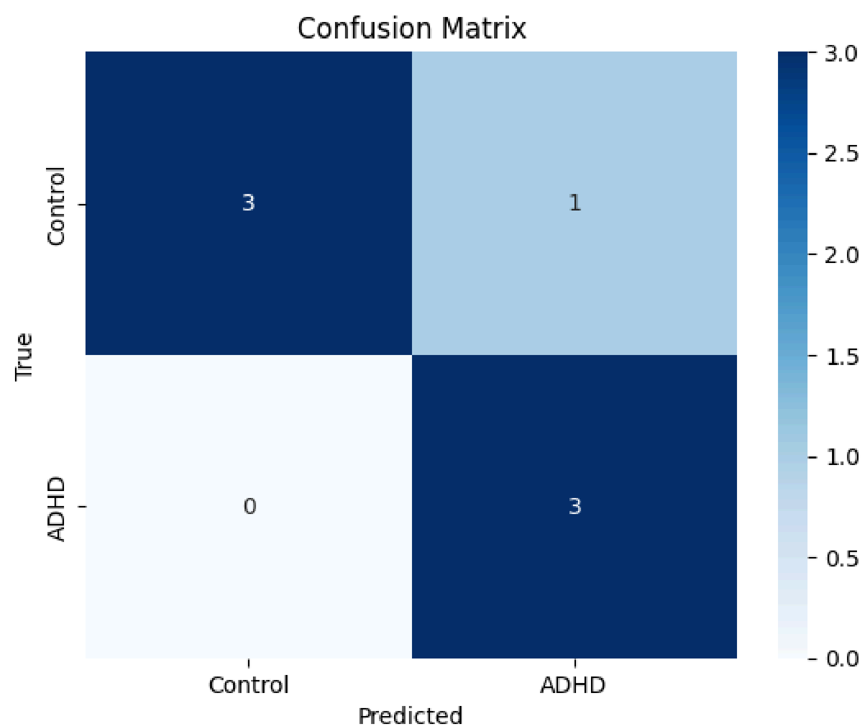

Figure 7. Confusion Matrix on Test Set for the Dual-Branch 1d CNN

4.3 Framework 3: Variable Length Dual-Branch 1D CNN

Building upon the success of Framework 2, Framework 3 sought to address the limitations of fixed sequence lengths by incorporating adaptive pooling layers. This enhancement allowed the model to accommodate the inherent variability in the length of HRV and motor activity time series data, leading to a smoother learning process, as depicted in Figure 8. The best-performing configuration for Framework 3 mirrored that of Framework 2, utilising the SGD optimiser with a learning rate of 0.01, 20 epochs, a batch size of 16, dropout regularisation with a rate of 0.2, and batch normalisation. This configuration resulted in a validation accuracy of 57.14%, consistent with Framework 2. The performance metrics for this framework are summarised in Appendix D.

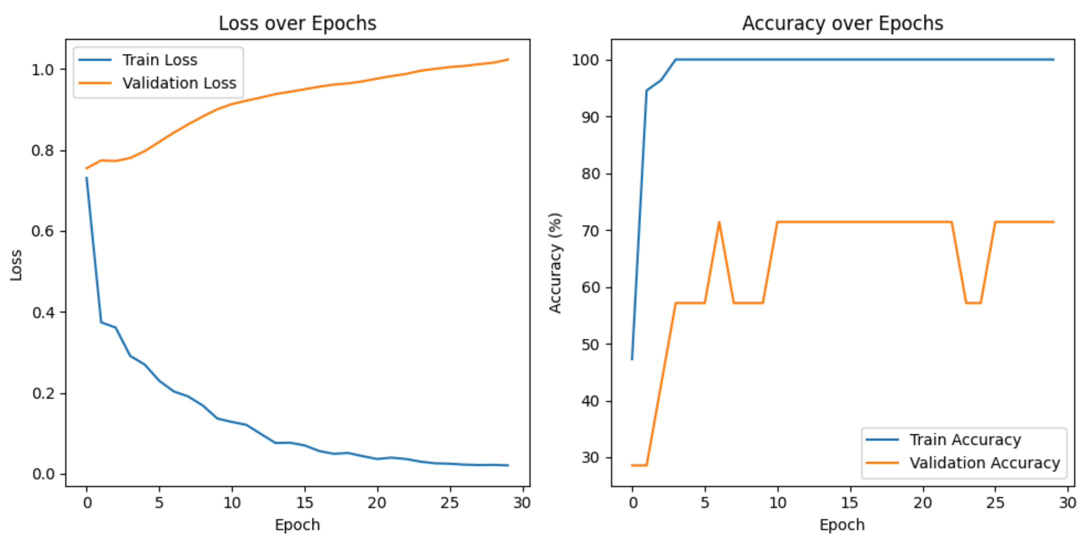

Figure 8. Training and Validation Loss Curves of the Variable Length Dual-Branch 1D CNN

The model's performance on the unseen test set was surprising, achieving a test accuracy of 85.71%, surpassing the validation metrics. This improvement in the learning process can be attributed to the adaptive pooling mechanism, which effectively handled the variable-length sequences present in the data. The model's precision of 0.75 and recall of 1.00 indicate its ability to accurately identify all ADHD cases in the test set while maintaining high precision in its predictions. The F1-score of 0.86 further demonstrates a balanced performance between precision and recall. Figures 9 and 10 visually represent the model's ROC curve and confusion matrix.

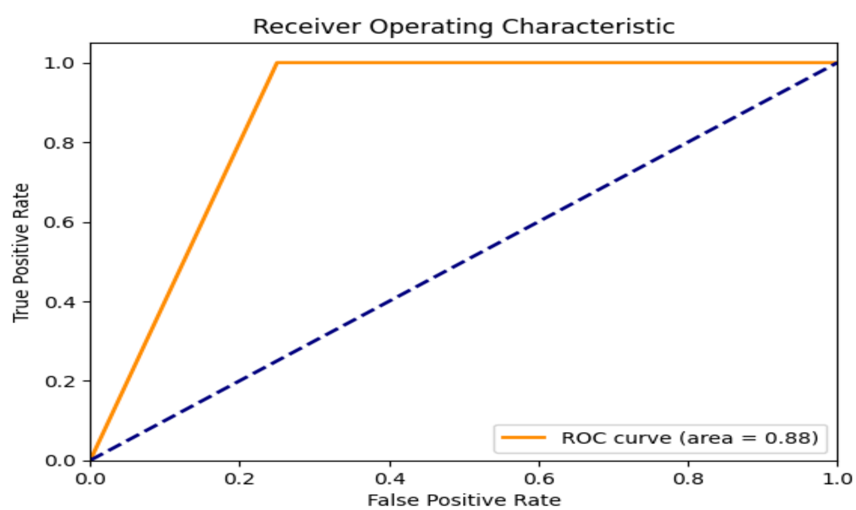

Figure 9. ROC Curve of Variable Length Dual-Branch 1d CNN

These results suggest that the dual-branch architecture and the adaptive pooling mechanism for variable-length sequences effectively leverage the complementary information in HRV and motor activity data. The parallel processing of these modalities and the model's flexibility in handling sequences of different lengths enables it to capture distinct patterns and relationships within each data stream, contributing to improved classification accuracy and learning curve.

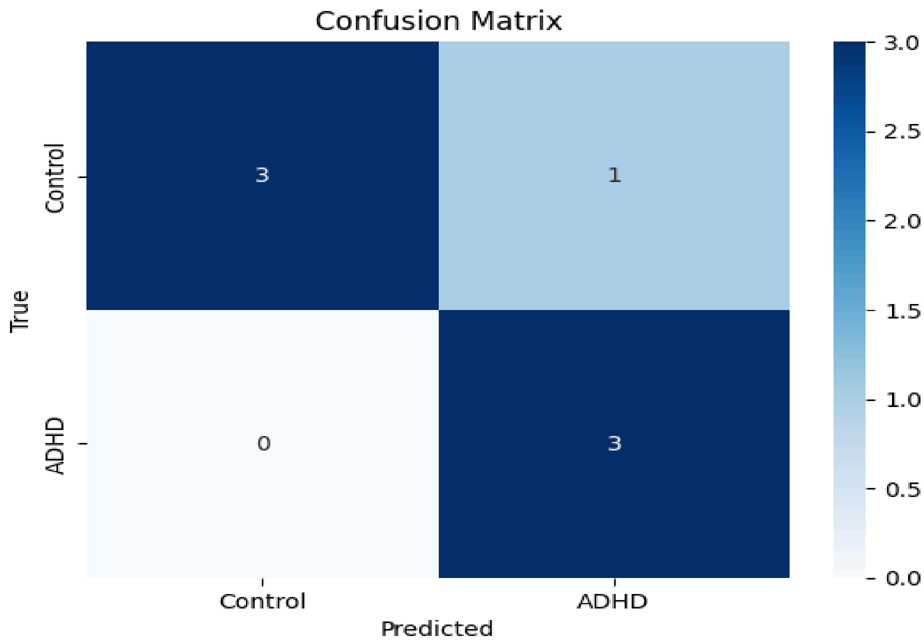

Figure 10. Confusion Matrix on Test Set for the Variable Length Dual-Branch 1d CNN

5. Discussion

The better performance of the dual-branch 1D CNN models, particularly Framework 3 with its adaptive pooling mechanism, highlights the potential of deep learning in improving ADHD diagnosis. Achieving a test accuracy of 85.71%, these models significantly outperformed the baseline study conducted on the HYPERAKTIV dataset, which utilised traditional machine learning models like logistic regression and random forests and reported a maximum accuracy of 72%. This substantial improvement aligns with the increasing recognition of deep learning's capabilities in capturing intricate patterns within complex datasets, as highlighted in recent surveys of deep learning in healthcare (Shickel et al., 2017).

The success of the dual-branch architectures in harnessing complementary information from both HRV and motor activity data is a key finding of this research. By processing these modalities independently and then integrating their features, the models were able to capture unique patterns that might have been overlooked by single-modality approaches. The results of this study are consistent with the findings reported by Wang et al. (2022), who demonstrated the enhanced diagnostic power of combining EEG data with other physiological markers in a CNN-LSTM model, albeit with a different set of modalities. The convergence of the results underscores the importance of multimodal data integration in unravelling the complex neurophysiological underpinnings of ADHD.

Delving deeper into the specific contributions of HRV and motor activity data, it is plausible that HRV reflects the autonomic nervous system's dysregulation often observed in individuals with ADHD (Rukmani et al., 2016), while motor activity captures the hallmark hyperactive and impulsive behaviours associated with the disorder. The ability of the models to leverage these distinct physiological signatures suggests a promising avenue for developing objective and reliable diagnostic tools.

The suboptimal performance of the CNN-LSTM model, despite its theoretical potential to capture spatial-temporal patterns, underscores the challenges of model complexity and overfitting, especially with limited data (Domingos, 2012). This observation is consistent with findings from other studies, such as Saurabh and Gupta (2024), who employed a modified BLSTM model for ADHD classification using rs-fMRI data and reported a similar issue of overfitting. Their solution, involving careful feature engineering and regularisation techniques, could offer valuable insights for future refinements of the CNN-LSTM framework.

The limitations of this study primarily stem from the relatively small sample size(N=69) and potential biases within the HYPERAKTIV dataset. While the dataset provides valuable insights, its limited scope raises concerns about the generalizability of the findings to larger and more diverse populations. Future research should prioritise the acquisition and analysis of more comprehensive datasets, encompassing a wider range of ages, ethnicities, socioeconomic backgrounds, and ADHD subtypes. This would not only enhance the external validity of the models but also allow for a more nuanced understanding of how individual differences and contextual factors influence the relationship between physiological data and ADHD. The inclusion of participants with comorbid conditions, while reflecting the real-world clinical landscape, necessitates a more cautious interpretation of the identified biomarkers. Future studies should aim to disentangle the effects of ADHD from those of co-occurring conditions, potentially by leveraging advanced statistical methods or by recruiting participants with more specific diagnostic profiles. This would enable a more precise identification of ADHD-specific biomarkers and contribute to the development of more targeted interventions. In addition to addressing the dataset limitations, future research should explore the incorporation of other data modalities, such as neuroimaging (e.g., fMRI, EEG), genetic data, and cognitive assessments. Integrating these diverse data sources could further enhance diagnostic accuracy and provide a more comprehensive view of ADHD's aetiology and heterogeneity. For instance, neuroimaging data could shed light on the structural and functional brain abnormalities associated with ADHD, while genetic data could reveal susceptibility genes and potential targets for therapeutic interventions. Another field for future research lies in the exploration of alternative deep learning architectures, such as recurrent neural networks (RNNs) or transformers, which may be better suited for capturing the temporal dynamics of HRV and motor activity data (Lipton et al., 2015).

Furthermore, incorporating techniques like transfer learning, where knowledge from pre-trained models on larger datasets is leveraged, could enhance model performance and address the limitations of small datasets (Pan & Yang, 2010).

6. Conclusion

The dual-branch 1D CNN models developed in this study, particularly Framework 3 with its adaptive pooling mechanism, showcase the transformative potential of deep learning and multimodal data integration in the realm of ADHD diagnosis. By achieving a test accuracy of 85.71%, surpassing the 72% accuracy of traditional machine learning models in the baseline study, this research provides compelling evidence for the efficacy of deep learning in extracting nuanced patterns from physiological data. The analysis of time-series data from HRV and motor activity has proven instrumental in identifying potential biomarkers associated with ADHD. The unique patterns captured by the dual-branch architecture, where each data stream contributes distinct information, align with the hypothesis that ADHD's underlying neurophysiological mechanisms are multifaceted and can be better understood through a multimodal approach. This approach not only answers the main research question but also opens up new areas for exploring the complex interplay between physiological signals and cognitive processes in ADHD. Furthermore, the improved performance of the dual-branch models, compared to single-modality approaches, unequivocally demonstrates the added value of fusing multiple physiological data streams. This finding, addressing RQ 1.2, underscores the importance of integrating complementary information from different sources to enhance the accuracy and reliability of diagnostic models for complex neurodevelopmental disorders.

The inclusion of additional modalities, such as neuroimaging and genetic data, could further enhance diagnostic accuracy and provide a more comprehensive understanding of ADHD's aetiology. Moreover, investigating the generalizability of these models to larger, more diverse populations is crucial for ensuring their real-world applicability and clinical utility.

By demonstrating the capabilities of multimodal data integration and advanced machine learning techniques, this study presents new possibilities for developing more accurate, efficient, and personalised diagnostic tools for a wide range of neurodevelopmental and psychiatric disorders. These advancements could lead to earlier detection, more targeted interventions, and improved outcomes for individuals struggling with these conditions.

.

References

Ahmadi, A., Kashefi, M., Shahrokhi, H., & Nazari, M. A. (2021). Computer aided diagnosis system using deep convolutional neural networks for ADHD subtypes. *Biomedical Signal Processing and Control*, *63*, 102227. https://doi.org/10.1016/j.bspc.2020.102227

Clements, S. D. (1966). *Minimal Brain Dysfunction in Children: Terminology and Identification: Phase One of a Three Phase Project*. U.S. Department of Health, Education and Welfare.

Del Campo, N., Chamberlain, S. R., Sahakian, B. J., & Robbins, T. W. (2011). The roles of dopamine and noradrenaline in the pathophysiology and treatment of Attention-Deficit/Hyperactivity Disorder. *Biological Psychiatry*, *69*(12), e145–e157. https://doi.org/10.1016/j.biopsych.2011.02.036

Danielson, M. L., Bitsko, R. H., Ghandour, R. M., Holbrook, J. R., Kogan, M. D., & Blumberg, S. J. (2018). Prevalence of Parent-Reported ADHD Diagnosis and Associated Treatment among U.S. children and Adolescents, 2016. *Journal of Clinical Child and Adolescent Psychology*, *47*(2), 199–212. https://doi.org/10.1080/15374416.2017.1417860

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, *55*(10), 78–87. https://doi.org/10.1145/2347736.2347755

Fernández, A., Quintero, J., Hornero, R., Zuluaga, P., Navas, M., Gómez, C., Escudero, J., García-Campos, N., Biederman, J., & Ortiz, T. (2009). Complexity Analysis of spontaneous brain activity in Attention-Deficit/Hyperactivity Disorder: Diagnostic implications. *Biological Psychiatry*, *65*(7), 571–577. https://doi.org/10.1016/j.biopsych.2008.10.046

Saurabh, S., & Gupta, P. K. (2023). Deep Learning-Based Modified Bidirectional LSTM network for classification of ADHD Disorder. *Arabian Journal for Science and Engineering*, *49*(3), 3009–3026. https://doi.org/10.1007/s13369-023-07786-w

Hawkins, D. M. (2003). The problem of overfitting. *Journal of Chemical Information and Computer Sciences*, *44*(1), 1–12. https://doi.org/10.1021/ci0342472

Hawkins, D. M. (2003). The problem of overfitting. *Journal of Chemical Information and Computer Sciences*, *44*(1), 1–12. https://doi.org/10.1021/ci0342472

Hicks, S. A., Stautland, A., Fasmer, O. B., Førland, W., Hammer, H. L., Halvorsen, P., Mjeldheim, K., Oedegaard, K. J., Osnes, B., Giæver Syrstad, V. E., Riegler, M. A., & Jakobsen, P. (2021). HYPERAKTIV: An Activity Dataset from Patients with Attention-Deficit/Hyperactivity Disorder (ADHD). *Proceedings of the 12th ACM Multimedia Systems Conference*, 314–319. https://doi.org/10.1145/3458305.3478454

Ioffe, S., & Szegedy, C. (2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *Proceedings of the 32nd International Conference on Machine Learning*, 448–456. https://proceedings.mlr.press/v37/ioffe15.html

Kotsiantis, S. B., Kanellopoulos, D., & Pintelas, P. E. (2006). *Data Preprocessing for Supervised Learning. 1*(1).

Lipton, Z. C., Kale, D. C., Elkan, C., & Wetzel, R. (2017). *Learning to Diagnose with LSTM Recurrent Neural Networks* (arXiv:1511.03677). arXiv. https://doi.org/10.48550/arXiv.1511.03677

Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2017). Deep learning for healthcare: review, opportunities and challenges. *Briefings in Bioinformatics*, *19*(6), 1236–1246. https://doi.org/10.1093/bib/bbx044

Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, *22*(10), 1345–1359. https://doi.org/10.1109/tkde.2009.191

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., … Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Advances in Neural Information Processing Systems*, *32*. https://proceedings.neurips.cc/paper/2019/hash/bdbca288fee7f92f2bfa9f7012727740-Abstract.html

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., & Cournapeau, D. (n.d.). Scikit-learn: Machine Learning in Python. *MACHINE LEARNING IN PYTHON*.

Riaz, A., Asad, M., Arif, S. M. M. R. A., Alonso, E., Dima, D., Corr, P., & Slabaugh, G. (2018). Deep FMRI: AN end-to-end deep network for classification of FMRI data. *AN End-to-end Deep Network for Classification of fMRI Data*. https://doi.org/10.1109/isbi.2018.8363838

Ruder, S. (2017). *An overview of gradient descent optimization algorithms* (arXiv:1609.04747). arXiv. https://doi.org/10.48550/arXiv.1609.04747

Rukmani, M. R., Seshadri, S. P., Thennarasu, K., Raju, T. R., & Sathyaprabha, T. N. (2016). Heart Rate Variability in Children with Attention-Deficit/Hyperactivity Disorder: A Pilot Study. *Annals of Neurosciences*, *23*(2), 81–88. https://doi.org/10.1159/000443574

Saurabh, S., & Gupta, P. K. (2023b). Deep Learning-Based Modified Bidirectional LSTM network for classification of ADHD Disorder. *Arabian Journal for Science and Engineering*, *49*(3), 3009–3026. https://doi.org/10.1007/s13369-023-07786-w

Shickel, B., Tighe, P., Bihorac, A., & Rashidi, P. (2018). Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. *IEEE Journal of Biomedical and Health Informatics*, *22*(5), 1589–1604. https://doi.org/10.1109/jbhi.2017.2767063

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (n.d.). *Dropout: A Simple Way to Prevent Neural Networks from Overfitting*. The journal of machine learning research, 15(1), 1929-1958. http://jmlr.org/papers/v15/srivastava14a.html

Thapar, A. (2018). Discoveries on the Genetics of ADHD in the 21st Century: New findings and their implications. *The American Journal of Psychiatry*, *175*(10), 943–950. https://doi.org/10.1176/appi.ajp.2018.18040383

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

Wang, C., Wang, X., Jing, X., Yokoi, H., Huang, W., Zhu, M., Chen, S., & Li, G. (2022). Towards high-accuracy classifying attention-deficit/hyperactivity disorders using CNN-LSTM model. *Journal of Neural Engineering*, *19*(4), 046015. https://doi.org/10.1088/1741-2552/ac7f5d

Willcutt, E. G. (2012). The Prevalence of DSM-IV Attention-Deficit/Hyperactivity Disorder: A Meta-Analytic Review. *Neurotherapeutics*, *9*(3), 490–499. https://doi.org/10.1007/s13311-012-0135-8

Wolraich, M. L., Chan, E., Froehlich, T., Lynch, R. L., Bax, A., Redwine, S. T., Ihyembe, D., & Hagan, J. F. (2019). ADHD Diagnosis and Treatment Guidelines: A Historical perspective. *Pediatrics*, *144*(4). https://doi.org/10.1542/peds.2019-1682

Zhang, J., He, T., Sra, S., & Jadbabaie, A. (2020). *Why gradient clipping accelerates training: A theoretical justification for adaptivity* (arXiv:1905.11881). arXiv. https://doi.org/10.48550/arXiv.1905.11881

Appendices

Appendix A: Model Summary of Framework 2

Layer (type)	Output Shape	Param #
Conv1d-1	[-1, 16, 48191]	96
ReLU-2	[-1, 16, 48191]	0
MaxPool1d-3	[-1, 16, 24095]	0
Conv1d-4	[-1, 32, 12046]	2,592
ReLU-5	[-1, 32, 12046]	0
MaxPool1d-6	[-1, 32, 6023]	0
Flatten-7	[-1, 192736]	0
Linear-8	[-1, 128]	24,670,336
ReLU-9	[-1, 128]	0
Conv1d-10	[-1, 16, 48191]	96
ReLU-11	[-1, 16, 48191]	0
MaxPool1d-12	[-1, 16, 24095]	0
Conv1d-13	[-1, 32, 12046]	2,592
ReLU-14	[-1, 32, 12046]	0
MaxPool1d-15	[-1, 32, 6023]	0
Flatten-16	[-1, 192736]	0
Linear-17	[-1, 128]	24,670,336
ReLU-18	[-1, 128]	0
Linear-19	[-1, 128]	32,896
ReLU-20	[-1, 128]	0
Linear-21	[-1, 1]	129

Appendix B: Model Summary of Framework 3

Layer (type)	Output Shape	Param #
Conv1d-1	[-1, 16, 48191]	96
ReLU-2	[-1, 16, 48191]	0
AvgPool1d-3	[-1, 16, 24095]	0
Conv1d-4	[-1, 32, 12046]	2,592
ReLU-5	[-1, 32, 12046]	0
AvgPool1d-6	[-1, 32, 6023]	0
Flatten-7	[-1, 192736]	0
Linear-8	[-1, 128]	24,670,336
ReLU-9	[-1, 128]	0
BatchNorm1d-10	[-1, 128]	256
Dropout-11	[-1, 128]	0
Conv1d-12	[-1, 16, 48191]	96
ReLU-13	[-1, 16, 48191]	0
AvgPool1d-14	[-1, 16, 24095]	0
Conv1d-15	[-1, 32, 12046]	2,592
ReLU-16	[-1, 32, 12046]	0
AvgPool1d-17	[-1, 32, 6023]	0
Flatten-18	[-1, 192736]	0
Linear-19	[-1, 128]	24,670,336
ReLU-20	[-1, 128]	0
BatchNorm1d-21	[-1, 128]	256
Dropout-22	[-1, 128]	0
Linear-23	[-1, 128]	32,896
ReLU-24	[-1, 128]	0
BatchNorm1d-25	[-1, 128]	256
Dropout-26	[-1, 128]	0
Linear-27	[-1, 1]	129

Appendix C: Classification Report of the Dual-Branch 1d CNN Framework

	Precision	Recall	F1-score	Support
Control	1.00	0.75	0.86	4
ADHD	0.75	1.00	0.86	3
Accuracy			0.86	7
Macro avg	0.88	0.88	0.86	7
Weighted avg	0.89	0.86	0.86	7

Appendix D: Classification Report of the Variable Length Dual-Branch 1d CNN Framework

	Precision	**Recall**	**F1-score**	**Support**
Control	1.00	0.75	0.86	4
ADHD	0.75	1.00	0.86	3
Accuracy			0.86	7
Macro avg	0.88	0.88	0.86	7
Weighted avg	0.89	0.86	0.86	7