Highlights

Graph-Based Representations of Motor Behavior Improve Machine Learning Prediction of Attention-Deficit/Hyperactivity Disorder Compared to Raw Features

Georgios Bouchouras, Konstantinos Kotis

- Graph-based representations of motor behavior outperform raw feature vectors for ADHD prediction.
- Relational modeling reveals latent structure in movement dynamics not accessible to flat feature-based learning.

Graph-Based Representations of Motor Behavior Improve Machine Learning Prediction of Attention-Deficit/Hyperactivity Disorder Compared to Raw Features

Georgios Bouchouras, Konstantinos Kotis

Intelligent Systems Lab, Department of Cultural Technology and Communication, University of the Aegean, University Hill, Mytilene, 81100, Lesvos, Greece

Abstract

Motor behavior provides a complementary and ecologically valid window into Attention-Deficit/Hyperactivity Disorder (ADHD), reflecting underlying alterations in temporal regulation, variability, and complexity of movement. Most machine learning approaches applied to movement-derived data rely on flat feature representations, treating entropy-, variability-, and complexity-based descriptors as independent inputs. Such representations neglect the semantic and relational structure through which motor behavior emerges.

In this study, we directly compare raw feature–based machine learning with a graph-based representation derived from the same movement features, under a strictly controlled experimental design. Using actigraphy-derived descriptors of variability, entropy, and temporal organization from the HYPERAKTIV dataset ($N = 85$), we construct an expert knowledge graph encoding subject–feature associations and subject–subject similarity relationships. Subject embeddings are learned using Node2Vec, and classification performance is evaluated using the same logistic regression model, feature set, and stratified cross-validation protocol across all conditions.

Results show that graph-based representations substantially outperform raw feature vectors (ROC–AUC: 0.628 vs. 0.419), despite using identical input information. Combining raw features with graph embeddings does not yield systematic additive gains, indicating partial redundancy between representations. These findings demonstrate that restructuring movement-derived information into a relational form reveals discriminative structure that remains inaccessible to flat feature-based models.

The study highlights the importance of data representation in movement-based machine learning and supports the use of knowledge graph modeling as a theoretically grounded and effective approach for studying motor dysregulation in ADHD.

Keywords: Attention-Deficit/Hyperactivity Disorder;, Motor Behavior;, Actigraphy;, Knowledge Graphs;, Graph Embeddings;, Machine Learning;, Movement Variability;, Entropy and Complexity

---

1. Introduction

Attention-Deficit/Hyperactivity Disorder (ADHD) is a neurodevelopmental condition primarily defined by persistent difficulties in attention regulation, impulsivity, and hyperactivity (Bitsko et al. (2022); El-Masry et al. (2021)) Beyond these core behavioral symptoms, growing evidence indicates that motor behavior is also systematically affected, revealing alterations in coordination, timing, and movement regulation. Such motor differences are increasingly viewed as expressions of underlying neural dysregulation rather than secondary or incidental features (Li et al. (2025); Cole et al. (2008)).

Motor behavior provides a complementary and non-invasive window into ADHD-related dysregulation, particularly when assessed under naturalistic conditions. Wearable sensors and actigraphy devices enable the continuous recording of movement over extended periods, capturing motor output as it unfolds during everyday life rather than during short, structured laboratory tasks (Bouchouras et al. (2024)). These recordings are well suited for studying temporal organization, variability, and complexity of movement, properties that are difficult to reliably assess through brief clinical evaluations.

Contemporary theories of motor control and dynamical systems emphasize that movement variability should not be interpreted solely as noise or execution error. Instead, variability reflects how the motor system regulates timing, coordination, and adaptability across time. Healthy motor behavior is characterized by structured variability, balancing stability and flexibility, whereas pathological or constrained systems tend to exhibit overly regular or overly irregular dynamics (Stergiou et al. (2011)). Within this framework, ADHD-related deficits in timing, inhibition, and self-regulation are expected to manifest as altered temporal organization and increased instability of motor output.

To capture these properties, movement analysis has increasingly relied on features that quantify variability, temporal dependency, and signal complexity rather than mean activity magnitude alone. Entropy- and complexity-based descriptors, alongside variability measures, provide a principled means of characterizing how movement unfolds over time and how adaptive the underlying control system is (Costa et al. (2005); Dhawale et al. (2017)). These descriptors are conceptually related and often reflect complementary aspects of motor regulation, motivating representations that preserve their functional relationships rather than treating them as independent variables.

Most machine learning approaches applied to ADHD-related movement data rely on flat feature representations, in which each descriptor is treated as an isolated input (Siqueira et al. (2014); Cheng et al. (2012)). Although effective in some settings, such representations neglect the semantic structure of motor behavior, where movement characteristics emerge through interacting control processes. This limitation motivates the use of structured, semantically informed representations that explicitly encode relationships among subjects and movement features.

The aim of this study is to compare two different ways of representing movement data in order to assess their effectiveness in predicting attention-deficit/hyperactivity disorder (ADHD) using machine learning. In the first approach, participants are described using raw movement features. In the second approach, the same features are organized into a graph that captures meaningful relationships between participants based on their movement patterns.

Using measures of variability, entropy, and temporal organization, we build a graph that links individuals with similar movement behavior and summarize this information using graph embeddings.

The same machine learning model, training procedure, and evaluation strategy are used for both approaches, allowing differences in performance to be attributed to the way the data are represented rather than to the learning algorithm itself.

By focusing on representation rather than algorithmic complexity, this work seeks to clarify whether semantic modeling of movement dynamics provides added value for ADHD-related motor analysis. In doing so, it offers a theoretically grounded bridge between dynamical systems theory, movement science, and machine learning, contributing to more interpretable and robust approaches for studying motor dysregulation in ADHD.

Based on the above rationale, we test the following hypotheses:

H1: Graph-based representations will outperform flat feature vectors. Using the same 51 movement-derived variables and the same classifier, we expect the knowledge graph (via Node2Vec embeddings) to yield higher ROC–AUC than the raw-feature baseline, because the graph encodes semantic structure among subjects and features that is not accessible to independent-variable learning.

H2: Combining raw features with graph embeddings will not necessarily provide additive gains. We expect that simply concatenating raw features with graph embeddings will not consistently improve performance over the graph-only model, due to partial redundancy between representations and the limited sample size.

2. Related work

2.1. Motor Variability and Movement Organization in ADHD

Although Attention-Deficit/Hyperactivity Disorder (ADHD) is primarily defined by behavioral symptoms such as inattention, impulsivity, and hyperactivity, an increasing body of evidence suggests that motor behavior is also affected. Differences in motor coordination, movement variability, and temporal organization have been reported in individuals with ADHD, indicating that altered motor control may represent a complementary behavioral dimension of the disorder Esposito et al. (2011).

Human movement does not arise from isolated processes but emerges from the interaction of neural, biomechanical, and environmental subsystems operating across multiple temporal and spatial scales. Contemporary theories of motor control emphasize that movement variability is not simply noise or error, but a functional property of an adaptive motor system. Rather than repeating identical movement patterns, healthy motor behavior exhibits structured fluctuations that allow flexibility and responsiveness to changing demands.

This perspective represents a shift from earlier views that treated variability as an undesirable deviation from an ideal movement pattern.

Research in coordination dynamics and nonlinear motor control has shown that variability reflects the balance between stability and adaptability in motor behavior (Kelso (1995); Latash (2002); Stergiou et al. (2011)). Within this framework, neither overly regular nor highly random movement patterns are considered optimal.

Stergiou and colleagues (Stergiou et al. (2011, 2006)) formalized this idea through the "loss of complexity" hypothesis, which proposes that healthy biological systems exhibit rich, structured variability, whereas pathological or constrained systems tend toward overly predictable or overly irregular behavior. From this perspective, variability becomes an informative signal of neuromotor regulation rather than a nuisance to be eliminated.

In ADHD, increased intra-individual variability has been observed across a range of behavioral and motor tasks, including reaction time, postural control, and locomotion Castellanos et al. (2005); Kofler et al. (2013). These findings do not imply that individuals with ADHD simply move more or less, but rather that their motor output may be less consistently regulated over time. Such instability aligns with broader theories of ADHD that emphasize deficits in self-regulation, timing, and executive control, which may extend beyond cognition to motor behavior.

2.2. Entropy and Complexity Measures in Movement Analysis

To capture how movement unfolds over time, researchers have increasingly relied on nonlinear descriptors such as entropy and complexity measures. These metrics aim to characterize the structure and predictability of time series, rather than focusing solely on average magnitude or dispersion.

Entropy-based measures, including approximate entropy, sample entropy, permutation entropy, and spectral entropy, quantify different aspects of temporal irregularity in movement signals (Pincus (1991); Richman & Moorman (2000); Bandt & Pompe (2002)). Lower entropy values typically indicate more regular and predictable patterns, whereas higher values reflect increased irregularity. Importantly, entropy should not be interpreted as a direct measure of impairment. Both excessively low and excessively high entropy may indicate suboptimal motor control, depending on context Stergiou et al. (2011).

Complexity measures, such as Lempel–Ziv complexity (a measure of how complex or varied a signal is over time), provide complementary information by assessing the structural richness of symbolized time series and the emergence of novel patterns over time (Lempel & Ziv (1976)). Reduced complexity has been associated with aging, neurological disease, and constrained motor states, whereas healthy systems tend to exhibit richer and more adaptable dynamics (Goldberger et al. (2002); Vaillancourt & Newell (2002)).

Illustrative examples of complexity and entropy. **Lempel–Ziv complexity.** Consider two simplified activity sequences:

- *rest* → *rest* → *rest* → *rest* (low complexity: highly repetitive pattern)
- *rest* → *move* → *rest* → *burst* → *pause* → *move* (higher complexity: richer and less repetitive structure)

Lempel–Ziv complexity quantifies how structured or repetitive a movement pattern is over time. It increases when new or less predictable subsequences appear. Importantly, it reflects the organization of movement, not how strong or fast the movement is.

Entropy-based measures. Similarly, entropy captures how predictable the temporal evolution of movement is:

- *rest* → *move* → *rest* → *move* (low entropy: regular and predictable alternation)
- *rest* → *burst* → *pause* → *move* → *rest* → *burst* (higher entropy: irregular and less predictable pattern)

Entropy-based measures quantify temporal irregularity. Lower entropy reflects more regular patterns, whereas higher entropy indicates increased unpredictability. Importantly, both excessively low and excessively high entropy may reflect suboptimal motor control, depending on context.

In ADHD research, entropy- and complexity-based descriptors have been applied to postural sway, handwriting, and activity recordings, often revealing altered temporal organization compared to typically developing controls (Zentgraf et al. (2009); Haddad et al. (2018)). These measures are best understood as descriptive indicators of movement dynamics, rather than diagnostic biomarkers. Their value lies in capturing how motor behavior is organized over time, not in identifying a single pathological signature.

2.3. Actigraphy and everyday movement behavior

Actigraphy-based movement analysis offers a high degree of ecological validity. Unlike laboratory-based gait experiments or short-duration motor tasks, actigraphy captures motor behavior continuously during daily life, integrating locomotor and non-locomotor activities in natural contexts. From a motor control perspective, the actigraph captures inertial acceleration signals comparable to those obtained from IMU systems, enabling analysis of movement dynamics across time, despite the absence of orientation-specific sensors.

This approach is particularly relevant for ADHD, where behavioral symptoms are context-dependent and may not manifest consistently during structured testing. Several studies have used actigraphy to characterize hyperactivity, circadian rhythms, and motor restlessness in ADHD, often reporting altered temporal patterns rather than simple increases in mean activity levels Porrino et al. (1983); Teicher et al. (1996); Konofal et al. (2001).

More recent work has shifted from average activity counts toward analyzing variability, intermittency, and complexity of actigraphic signals. These studies suggest that differences in temporal regulation may be more informative than overall movement intensity alone Hu et al. (2004); Wainwright (2017). Although activity signals coming from uncontrolled movement -outside of a laboratory or certain tasks- reflect a mixture of behaviors, this heterogeneity provides a broader view of motor regulation as it occurs in everyday life.

2.4. Limitations of Flat Feature Representations

Despite the richness of movement-derived descriptors, most machine learning approaches in clinical movement analysis rely on flat feature representations (Siqueira et al. (2014); Cheng et al. (2012)). In such models, features are treated as independent inputs, implicitly assuming that each variable contributes separately to classification performance.

This assumption is problematic when features are conceptually related, as is the case for entropy, complexity, and variability measures derived from the same underlying time series. For example, permutation entropy and Lempel–Ziv complexity capture overlapping but distinct aspects of temporal structure. Their joint interpretation may be more informative than either measure alone. Specifically, in motor behavior related to ADHD, where impairments are thought to arise from dysregulated coordination rather than isolated deficits, this limitation becomes particularly relevant.

2.5. Relational Modeling and Knowledge Graphs

To address these challenges, structured data representations have gained attention in biomedical research. Knowledge graphs (KGs) provide a formal framework for encoding entities and their relationships, enabling domain knowledge to be explicitly represented rather than implicitly learned from unstructured data Hogan et al. (2022).

Applied to movement analysis, a graph-based representation allows subjects and movement-derived features to be modeled as interconnected nodes. This formulation aligns with motor control theories that emphasize coordination and interaction among multiple components rather than independent factors.

Graph embedding methods, such as Node2Vec, learn low-dimensional representations that preserve relational structure and neighborhood similarity within the graph (Grover & Leskovec (2016)). These embeddings do not correspond to individual movement features. Instead, they encode how features co-occur and interact across subjects.

In the context of ADHD, relational modeling enables the integration of subject–feature associations with subject–subject similarities derived from movement dynamics. By defining similarity based on entropy- and complexity-related descriptors, individuals can be grouped according to shared temporal organization patterns rather than raw activity magnitude.

2.6. Rationale of the Present Study

Machine learning methods have been widely applied to ADHD-related data using classifiers such as support vector machines, random forests, and neural networks Eslami et al. (2021); Mostafa et al. (2019). While these approaches often achieve promising performance, they typically operate on flat feature vectors and do not explicitly model relationships among movement descriptors.

Despite growing interest in both movement-based ML and graph-based representations, there is a lack of systematic comparisons examining whether relational modeling provides meaningful advantages over traditional feature-based approaches in ADHD-related motor analysis, particularly in small clinical datasets.

The present study addresses this gap by directly comparing raw movement features and knowledge graph–based representations derived from the same feature set. By holding the machine learning model constant and varying only the data representation, we aim to determine whether performance differences arise from relational structure itself rather than increased model complexity.

3. Methods

3.1. Dataset

The activity data used in this study originate from the HYPERAKTIV dataset (Hicks et al. (2021)), a publicly available resource designed to support research on Attention-Deficit/Hyperactivity Disorder (ADHD) and related neurodevelopmental and mental health conditions. The dataset is accessible via the Open Science Framework (OSF)[1] and comprises multimodal physiological recordings, including motor activity and heart rate measurements.

Motor activity data were collected using a wrist-worn actigraph device (Actiwatch AW4, Cambridge Neurotechnology Ltd., Eastleigh, England), equipped with a piezoelectric triaxial accelerometer. The device recorded three-dimensional acceleration signals along the x, y, and z

[1] `https://osf.io/3agwr`

axes at a sampling frequency of 32 Hz. To reduce noise and ensure sensitivity to meaningful physical motion, only movements exceeding a threshold of 0.05 g (i.e., 5% of gravitational acceleration) were retained. These signals were subsequently aggregated into non-overlapping 1-minute epochs, producing integer-valued activity counts proportional to the magnitude and intensity of recorded movement.

Importantly, the dataset does not originate from controlled laboratory-based gait experiments or structured walking trials. Instead, recordings were obtained under free-living, ecologically valid conditions, capturing naturalistic motor behavior during participants' everyday activities. As such, the resulting activity signals reflect a mixture of locomotor and non-locomotor movements embedded within daily life, rather than isolated or stereotyped gait cycles.

Data collection spanned an average duration of 6.6 ± 1.3 days per participant. In the present study, we did not analyze the raw multiday actigraphy recordings directly. Instead, we relied on a derived feature-level representation provided with the dataset, where activity time series were transformed into statistical, temporal, and complexity-related movement descriptors for each participant. After preprocessing and subject alignment, the final cohort comprised $N = 85$ individuals.

The dataset was treated as a cross-sectional representation of individual movement behavior, focusing on structural relationships among subjects and movement-derived features rather than longitudinal dynamics. The binary clinical label was nearly balanced, with 40 control subjects and 45 participants diagnosed with ADHD, minimizing potential class imbalance effects.

The high temporal resolution and longitudinal nature of the recordings provide a suitable basis for investigating the temporal organization, variability, and complexity of movement dynamics. Such properties are known to be altered in ADHD and cannot be reliably captured through short-duration or laboratory-based movement assessments.

3.2. Clinical Movement Features

Feature extraction focused on clinically interpretable descriptors that capture dynamic properties of gait rather than static or mean-only characteristics. Based on prior evidence linking ADHD to impaired motor regulation, variability, and temporal instability, the extracted feature set emphasized three complementary dimensions:

- variability-based measures (e.g., standard deviation and coefficient of variation),
- entropy-based descriptors quantifying temporal irregularity,
- temporal features reflecting gait cycle organization and consistency.

Entropy-related features were computed to quantify the unpredictability and irregular structure of gait-related time series. These measures capture deviations in stride-to-stride control and rhythmic consistency, which are known to be affected in ADHD-related motor behavior.

Variability metrics were included to reflect intra-individual instability across gait cycles. Temporal descriptors were extracted to represent timing-related aspects of locomotion, capturing dynamic fluctuations that are not apparent in aggregate statistics.

Although the original HYPERAKTIV dataset provides a substantially larger pool of derived movement features, only a subset of 51 features was selected for the present study. This selection was not driven by data-driven optimization or feature ranking, but by methodological and theoretical considerations.

First, the primary aim of this work was not to maximize classification performance through exhaustive feature exploitation, but to investigate whether relational modeling of clinically meaningful movement characteristics adds value beyond flat feature representations. Consequently, feature selection was hypothesis-driven rather than performance-driven. Second, the selected features were restricted to descriptors that capture dynamic properties of movement, including variability, temporal structure, and signal complexity. Prior work on ADHD-related motor behavior consistently suggests that deficits in motor regulation are expressed primarily through increased temporal irregularity, instability, and reduced rhythmic consistency, rather than through mean activity levels alone. Features reflecting static or aggregate characteristics were therefore intentionally de-emphasized. Third, many of the additional features available in the dataset represent closely related variants or parameterizations of the same underlying constructs (e.g., multiple window sizes, aggregation functions, or redundant statistical forms). Including the full feature set would substantially increase dimensionality without a corresponding gain in physiological interpretability, thereby increasing the risk of overfitting given the moderate sample size. Finally, restricting the feature space to a compact and interpretable subset was essential to construct a meaningful knowledge graph. The selected features could be naturally organized into clinically coherent families (e.g., entropy, variability, temporal dependence), facilitating explicit modeling of relationships among movement properties and subjects. For these reasons, the final set of 51 movement-derived features was chosen as a principled compromise between expressivity, interpretability, and methodological (see Appendixes A and B for details).

All features were computed at the subject level, resulting in a compact but expressive representation of individual gait dynamics. Feature selection was hypothesis-driven and grounded in neurodevelopmental motor control theory, avoiding exhaustive automated feature mining in order to preserve interpretability and reduce overfitting risk.

3.3. Baseline: Raw Movement Features

As a baseline comparison, a traditional feature-based machine learning approach was employed using raw movement features. This baseline relied on mean values of spatiotemporal and kinematic gait descriptors, reflecting common practice in clinical gait analysis.

The baseline model treats each feature independently and represents each subject as a flat feature vector, without accounting for inter-feature relationships or higher-order biomechanical interactions. This setup serves as a reference point to assess whether relational modeling provides additional predictive value beyond isolated feature analysis.

3.4. Knowledge Graph Representation of Movement Features

To capture latent relationships among movement features, a knowledge graph representation was constructed. In this framework, subjects and movement features are modeled as nodes, while edges encode clinically meaningful relationships derived from the data.

Subject nodes were connected to feature nodes through `HAS_FEATURE` relationships, representing the presence of a specific movement characteristic. In addition, similarity relationships (`SIMILAR_TO`) were introduced between subject nodes, based on cosine similarity computed over a selected subset of entropy and variability features.

The similarity graph was constructed using a fixed k-nearest neighbor strategy, ensuring a controlled and interpretable relational structure. Similarity between subjects was defined using entropy-related gait features, which were selected to specifically capture temporal irregularity and control instability during walking. Other feature categories, such as variability and spatiotemporal descriptors, were retained in the knowledge graph through subject–feature relationships and contributed to the learned graph embeddings, but were not used to construct the subject similarity graph. This design choice was motivated by the hypothesis that temporal organization, rather than mean or amplitude-related gait characteristics, constitutes the primary discriminative dimension for inter-subject similarity in ADHD-related motor behavior.

It is important here to note that the knowledge graph does not generate new primary movement features. All 51 movement-derived variables used in the raw baseline are retained unchanged. The graph formulation introduces relational structure among subjects and features, which is subsequently encoded into latent graph embeddings via Node2Vec. These embeddings represent higher-order combinations of existing features and subject similarities, rather than novel movement measurements. They are latent representations encoding how movement-derived features co-occur and interact across subjects within the knowledge graph, rather than direct measurements of physical quantities.

The resulting graph encodes both subject–feature associations and subject–subject similarities, enabling the modeling of higher-order biomechanical interactions. Graph embeddings were generated using the Node2Vec algorithm, producing low-dimensional vector representations of subjects that preserve the structural properties of the knowledge graph.

3.5. Evaluation Protocol, Machine Learning and Metrics

All classification experiments were conducted using a logistic regression classifier. This model was selected due to its interpretability, robustness in small-to-moderate sample size settings, and widespread use as a strong baseline in clinical prediction tasks.

To ensure a fair comparison between representations, the same classifier architecture and training protocol were applied across all experimental conditions, including the raw feature baseline, the knowledge graph embeddings, and their concatenation. No model-specific tuning was performed for individual representations.

Prior to classification, feature vectors were standardized using z-score normalization. Missing values were handled using median imputation, which is robust to outliers and appropriate for clinical data with heterogeneous distributions. Regularization was applied using an ℓ_2 penalty, and the optimization was performed using a deterministic solver to ensure reproducibility across runs.

Hyperparameter tuning was intentionally avoided in order to isolate the effect of the input representation on classification performance. This design choice ensures that any observed performance differences can be attributed to the representational capacity of the raw features versus the graph-based embeddings, rather than to differences in model complexity or optimization.

Model performance was evaluated using a stratified k-fold cross-validation protocol at the subject level to prevent data leakage across folds. Three models were compared: (i) the raw feature baseline, (ii) the knowledge graph embedding model, and (iii) a combined model using concatenated raw features and graph embeddings.

Classification performance was assessed using the Area Under the Receiver Operating Characteristic Curve (AUC), which is robust to class imbalance and widely used in clinical prediction tasks. Mean AUC and standard deviation across folds are reported to reflect both performance and stability.

This evaluation framework allows direct comparison between traditional feature-based learning and graph-based representations, enabling assessment of the added value of relational modeling in ADHD gait analysis.

3.6. Feature Computation and Representation

Movement-derived features were computed from wrist-worn actigraphy signals using a standardized time-series feature extraction pipeline. The original activity signals were segmented into minute-level epochs and transformed into a set of statistical, temporal, and complexity-related descriptors at the subject level.

Feature computation focused on capturing three fundamental aspects of motor behavior:

(a) **Variability-related features**, quantifying the magnitude and dispersion of activity fluctuations across time (e.g., variance, standard deviation, coefficient of variation).

(b) **Temporal dependency features**, capturing short- and long-range correlations in movement dynamics (e.g., autocorrelation and partial autocorrelation at multiple lags).

(c) **Complexity and irregularity features**, characterizing the unpredictability, nonlinearity, and temporal organization of activity patterns (e.g., approximate entropy, sample entropy, permutation entropy, Fourier entropy, and Lempel–Ziv complexity).

These feature families were selected based on prior evidence linking ADHD to impaired motor control, increased temporal variability, and altered rhythmic organization of movement. Feature extraction was hypothesis-driven rather than exhaustive, aiming to preserve clinical interpretability while avoiding excessive dimensionality.

All features were computed identically for both the raw baseline model and the knowledge graph construction. Consequently, differences between models cannot be attributed to feature availability or preprocessing, but solely to the representational structure imposed by the graph-based formulation.

Reproducibility and Code Availability. To ensure full transparency and reproducibility, all code used in this study is publicly available. The complete implementation of the raw feature–based pipeline, the knowledge graph construction, the graph embedding procedure, and the machine learning evaluation framework can be found at:

https://github.com/GiorgosBouh/raw_vs_KG_ADHD

This repository accompanies the present manuscript and contains all scripts required to reproduce the reported results, including data preprocessing, knowledge graph generation, Node2Vec embedding, classification, and permutation-based feature importance analyses. All experiments can be reproduced using the same feature set, machine learning model, and evaluation protocol described in the Methods section, allowing independent verification of the findings and extension to other movement-based datasets.

4. Results

4.1. Classification Results

Table 1 summarizes the classification performance obtained using (i) raw movement-derived features, (ii) graph-based representations learned from the Expert Knowledge Graph, and (iii) their direct concatenation. All models were evaluated using a stratified five-fold cross-validation protocol and identical feature sets, ensuring a fair and controlled comparison.

The raw feature baseline achieved a mean ROC–AUC of 0.419, indicating limited discriminative capacity when movement descriptors are treated as independent variables. In contrast, the Expert Knowledge Graph representation substantially improved classification performance, yielding a mean ROC–AUC of 0.628. This improvement was achieved without introducing additional features, but solely by restructuring the same movement-derived information into a relational graph and learning subject embeddings via Node2Vec.

The combined model, which concatenates raw features with graph embeddings, achieved intermediate performance (AUC = 0.569). This result suggests that the graph embeddings capture complementary but partially overlapping information with the raw features, and that naïve feature fusion does not necessarily lead to additive performance gains.

Overall, these findings indicate that modeling latent relationships among movement features and subjects through a knowledge graph provides a more informative representation than isolated feature-based analysis, supporting the hypothesis that ADHD-related motor behavior is better characterized through structured relational models.

Table 1: Classification performance (ROC–AUC) for raw movement features and graph-based representations under a fair comparison protocol.

Model	**AUC (mean)**	**Std**
Raw features (baseline)	0.419	0.108
Expert Knowledge Graph (Node2Vec)	0.628	0.112
Raw + ExpertKG (concatenation)	0.569	0.078

4.2. Feature Importance and Embedding Contribution

To further investigate which movement-derived characteristics drive classification performance, permutation-based feature importance analysis was applied to the logistic regression models trained on (i) raw movement features, (ii) knowledge graph embeddings, and (iii) their concatenation.

Raw movement features.. For the raw feature baseline, the mean cross-validated classification performance was AUC $= 0.42 \pm 0.11$. Permutation analysis revealed that the most influential features were consistently related to *movement complexity* and *temporal irregularity*. Specifically, complexity-related descriptors (e.g., Lempel–Ziv complexity measures) exhibited the highest average importance (ΔAUC ≈ 0.06), followed by additional complexity variants (ΔAUC ≈ 0.02–0.03).

Entropy-based measures, including Fourier and permutation entropy, contributed moderately to performance, with average importance values in the range ΔAUC ≈ 0.01–0.02. Variability-related metrics, such as variance and standard deviation, showed smaller but non-negligible contributions (ΔAUC ≈ 0.01–0.02), indicating that dispersion of activity levels alone is less informative than the temporal structure of movement.

Overall, the raw feature analysis suggests that ADHD-related motor differences are primarily reflected in disrupted temporal organization and reduced structural regularity of movement, rather than in mean activity magnitude.

Knowledge graph embeddings.. Unlike the raw feature model, no single embedding dimension dominated the prediction. Instead, discriminative information was distributed across multiple embedding components, each capturing complementary aspects of the movement-derived relational structure encoded in the graph.

Importantly, the subject similarity graph was constructed exclusively using entropy- and complexity-related movement descriptors, which were identified as the most influential factors in the raw feature analysis. As a result, the learned embeddings primarily reflect higher-order representations of temporal irregularity, movement complexity, and dynamic instability, rather than isolated statistical properties.

The distributed contribution of embedding dimensions indicates that these factors are not expressed independently, but emerge through their interaction within the graph structure. In this formulation, temporal irregularity, complexity, and variability are jointly encoded via subject–feature and subject–subject relationships, allowing the model to capture latent biomechanical interactions that are inaccessible to flat feature vectors.

Consequently, the knowledge graph embeddings do not correspond to individual movement features, but to composite representations of coordinated motor control patterns, explaining their superior and more stable classification performance.

Combined representation.. The combined model, using both raw features and graph embeddings, achieved an intermediate performance of $\mathrm{AUC} = 0.57 \pm 0.08$. Permutation analysis showed that the most important predictors in this setting were predominantly embedding-derived, while only a small subset of raw temporal features retained moderate importance ($\Delta\mathrm{AUC} \approx 0.01$–$0.02$).

This result indicates that the knowledge graph embeddings subsume much of the information contained in the raw features, while additionally encoding relational and contextual structure that is not accessible to flat feature representations.

Summary. Taken together, the analyses indicate that temporal irregularity and movement complexity constitute the primary discriminative factors in ADHD-related motor behavior. While these properties emerge as individual features in the raw baseline, they are transformed into a distributed, relational representation within the knowledge graph, where no single dimension dominates, but their joint interaction drives classification performance.

5. Discussion

Overview of Main Findings. This study examined whether representing movement-derived features through a knowledge graph improves the prediction of Attention-Deficit/Hyperactivity Disorder compared to a conventional feature-based approach. Using the same set of 51 movement-derived variables and the same classification model, we observed a clear and consistent performance advantage for the graph-based representation.

The machine learning model trained on knowledge graph embeddings achieved substantially higher mean ROC–AUC values than the same model trained on raw movement features. This improvement was stable across cross-validation folds and was obtained without introducing any additional features or modifying the learning algorithm. The only difference for the model lay in how the same information was represented.

These results directly support Hypothesis H1. Encoding movement information within a relational structure allowed the model to capture dependencies among features and similarities among subjects that are not accessible when features are treated independently. This finding aligns with prior work suggesting that ADHD-related motor differences are expressed through altered temporal organization and coordination rather than isolated movement characteristics Castellanos et al. (2005); Kofler et al. (2013).

The results also support Hypothesis H2. The combined model, which merged raw movement features with graph embeddings, did not outperform the graph-only representation and instead showed intermediate performance. This indicates that the embeddings already capture much of the discriminative information contained in the original features. Simply combining representations therefore introduces redundancy rather than systematic performance gains, particularly in small-sample clinical settings.

Taken together, these findings indicate that the primary contribution of the knowledge graph lies in restructuring existing movement-derived information into a relational form. This restructuring exposes patterns of motor organization that remain obscured in conventional feature-based representations. Importantly, the results demonstrate that such gains can be achieved without increasing model complexity, which is especially relevant for clinically constrained datasets.

Interpretation of Raw Feature Performance. The raw feature baseline showed limited classification performance. Permutation analysis revealed that features related to movement complexity and temporal irregularity contributed most strongly, including Lempel–Ziv complexity and entropy-based measures.

This pattern is consistent with previous studies reporting that ADHD-related motor differences are more strongly reflected in temporal structure and irregularity than in mean activity level or movement magnitude (Esposito et al. (2011); Hu et al. (2004); Wainwright (2017)). Variability measures describing overall dispersion contributed less, suggesting that amplitude alone is a weak indicator of motor dysregulation.

Taken together, the raw feature analysis supports the view that ADHD-related movement differences are expressed through altered temporal organization rather than isolated statistical descriptors, but that such features remain limited when evaluated independently.

Why Knowledge Graph Embeddings Improve Performance. The improved performance of the knowledge graph model does not result from introducing new physiological measurements. Instead, the embeddings encode relationships among movement-derived features and similarities among subjects.

In this formulation, each subject is influenced not only by their own feature values but also by their position within a network of subjects with similar movement dynamics. This relational context allows the model to capture shared patterns of temporal organization that are difficult to detect using flat feature representations.

This approach aligns with motor control theories that emphasize coordination and interaction among system components rather than independent variables (Kelso (1995); Latash (2002)). It also resonates with the loss-of-complexity framework, which highlights the importance of structured variability as a marker of healthy motor regulation (Stergiou et al. (2011)).

In addition, relational modeling reduces sensitivity to noise, which is particularly relevant for free-living actigraphy data characterized by high variability and contextual heterogeneity (Konofal et al. (2001); Hu et al. (2004)). As a result, classification performance becomes more stable across folds.

Contribution of Embedding Dimensions. No single embedding dimension dominated the prediction. Instead, discriminative information was distributed across multiple dimensions, reflecting composite relationships encoded in the graph.

This distributed structure is expected, as graph embeddings do not correspond to individual movement features. Rather, they capture how entropy-, complexity-, and variability-related descriptors co-occur across subjects Grover & Leskovec (2016). Consequently, predictive information emerges from collective structure rather than isolated measurements.

Combined Representation. The combined model achieved intermediate performance. Permutation analysis showed that embedding-derived variables contributed most to prediction, while only a small subset of raw temporal features retained moderate importance.

This result indicates partial redundancy between raw features and graph embeddings. Much of the information carried by individual features appears to be implicitly encoded within the relational structure of the graph, limiting the benefit of naïve combination.

Conceptual Example. In the raw feature representation, each participant is described by a set of independent numerical descriptors, such as variability, entropy, or complexity. Each value is treated separately by the model.

In contrast, the graph-based representation links participants who show similar patterns in how their movement fluctuates over time. For example, two individuals may differ in how much they move overall, yet exhibit comparable cycles of rest and brief activity bursts during daily life. The graph connects such individuals based on shared temporal organization rather than absolute activity level.

The resulting embedding summarizes this shared motor regulation pattern into a compact representation. Instead of relying on isolated measurements, it reflects how movement behavior is structured over time. Empirically, this semantic organization leads to higher and more stable classification performance.

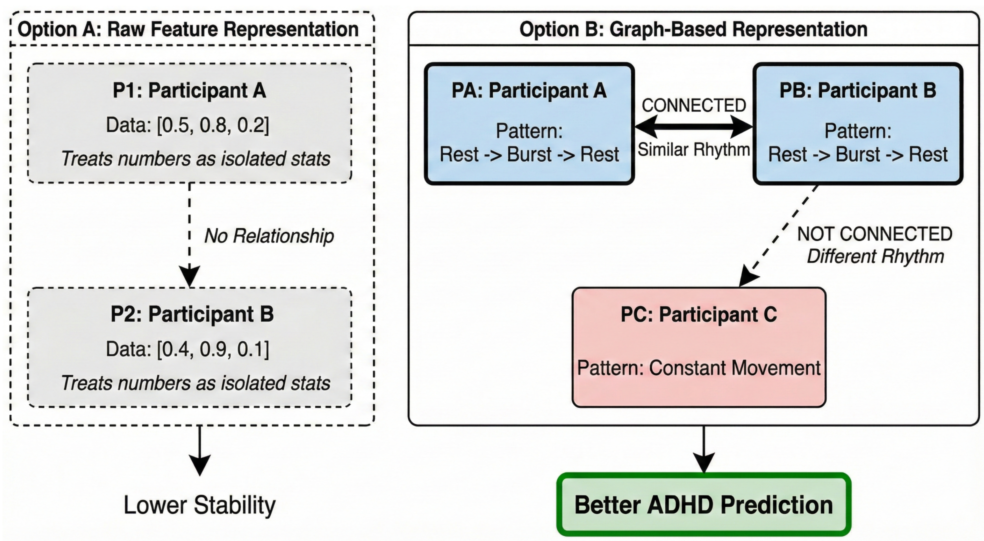

Figure 1: Conceptual illustration of raw versus graph-based representation of motor behavior. In the raw representation, participants are described by independent numerical features. In the graph-based representation, participants are linked based on shared temporal organization of movement, and embeddings summarize these relational patterns for machine learning.

Limitations. Several limitations should be noted. First, the sample size is modest, and larger cohorts are required to confirm generalizability. Second, the data reflect free-living activity rather than controlled gait experiments, which enhances ecological validity but limits biomechanical specificity. Third, graph embeddings are not directly interpretable in physiological terms, as they encode latent structure rather than measurable quantities.

Implications and Future Directions. The findings support the view that ADHD is associated with altered organization of movement rather than simple changes in activity level. This perspective aligns with contemporary theories of neurodevelopmental disorders emphasizing distributed dysregulation.

Future work should validate this framework in larger and longitudinal datasets, and explore extensions that integrate cognitive or clinical variables into multimodal graphs. Such approaches may further clarify how motor behavior relates to broader regulatory processes in ADHD.

Appendix A

Table 2: Movement-derived feature families and corresponding feature names used in the study.

Feature Family	**Feature Names (as used in** `features_entropy_variability.txt`**)**
Variance-based descriptors Measures overall dispersion and amplitude of movement activity. These features quantify how much activity fluctuates over time, reflecting motor variability and intensity regulation.	`ACC__standard_deviation`, `ACC__variance`, `ACC__variance_larger_than_standard_deviation`, `ACC__variation_coefficient`

Note. All movement-derived features were used identically in both the raw baseline and the knowledge graph models. Consequently, any observed performance differences cannot be attributed to feature selection, but solely to the structural and relational representation introduced by the graph-based formulation.

Feature Family	Feature Names (as used in `features_entropy_variability.txt`)
Autocorrelation features Capture linear temporal dependencies in the activity signal, indicating rhythmicity and persistence of movement patterns.	`ACC__autocorrelation__lag_0`, `ACC__autocorrelation__lag_1`, `ACC__autocorrelation__lag_2`, `ACC__autocorrelation__lag_3`, `ACC__autocorrelation__lag_4`, `ACC__autocorrelation__lag_5`, `ACC__autocorrelation__lag_6`, `ACC__autocorrelation__lag_7`, `ACC__autocorrelation__lag_8`, `ACC__autocorrelation__lag_9`
Aggregated autocorrelation Summary statistics of autocorrelation across multiple lags, providing compact descriptors of temporal structure.	`ACC__agg_autocorrelation__f_agg_"mean"__maxlag_40`, `ACC__agg_autocorrelation__f_agg_"median"__maxlag_40`, `ACC__agg_autocorrelation__f_agg_"var"__maxlag_40`
Partial autocorrelation Quantifies direct temporal dependencies by removing indirect lag effects, highlighting short-term motor control dynamics.	`ACC__partial_autocorrelation__lag_0`, `ACC__partial_autocorrelation__lag_1`, `ACC__partial_autocorrelation__lag_2`, `ACC__partial_autocorrelation__lag_3`, `ACC__partial_autocorrelation__lag_4`, `ACC__partial_autocorrelation__lag_5`, `ACC__partial_autocorrelation__lag_6`, `ACC__partial_autocorrelation__lag_7`, `ACC__partial_autocorrelation__lag_8`, `ACC__partial_autocorrelation__lag_9`

Note. All movement-derived features were used identically in both the raw baseline and the knowledge graph models. Consequently, any observed performance differences cannot be attributed to feature selection, but solely to the structural and relational representation introduced by the graph-based formulation.

Feature Family	Feature Names (as used in features_entropy_variability.txt)
Entropy-based complexity measures Assess temporal irregularity, unpredictability, and complexity of movement patterns, which are central to ADHD-related motor dysregulation hypotheses.	ACC__sample_entropy, ACC__approximate_entropy__m_2__r_0.1, ACC__approximate_entropy__m_2__r_0.3, ACC__approximate_entropy__m_2__r_0.5, ACC__approximate_entropy__m_2__r_0.7, ACC__approximate_entropy__m_2__r_0.9, ACC__permutation_entropy__dimension_3__tau_1, ACC__permutation_entropy__dimension_4__tau_1, ACC__permutation_entropy__dimension_5__tau_1, ACC__permutation_entropy__dimension_6__tau_1, ACC__permutation_entropy__dimension_7__tau_1
Spectral entropy Frequency-domain entropy descriptors capturing distribution of movement energy across frequencies.	ACC__fourier_entropy__bins_2, ACC__fourier_entropy__bins_3, ACC__fourier_entropy__bins_5, ACC__fourier_entropy__bins_10, ACC__fourier_entropy__bins_100
Symbolic complexity (Lempel–Ziv) Measures algorithmic complexity and repetitiveness of discretized activity sequences.	ACC__lempel_ziv_complexity__bins_2, ACC__lempel_ziv_complexity__bins_3, ACC__lempel_ziv_complexity__bins_5, ACC__lempel_ziv_complexity__bins_10, ACC__lempel_ziv_complexity__bins_100

Note. All movement-derived features were used identically in both the raw baseline and the knowledge graph models. Consequently, any observed performance differences cannot be attributed to feature selection, but solely to the structural and relational representation introduced by the graph-based formulation.

Feature Family	Feature Names (as used in `features_entropy_variability.txt`)
Time-reversal asymmetry Captures non-linear temporal asymmetry in movement dynamics, indicative of directional motor control properties.	`ACC__time_reversal_asymmetry_statistic__lag_1`, `ACC__time_reversal_asymmetry_statistic__lag_2`, `ACC__time_reversal_asymmetry_statistic__lag_3`

Appendix B: Movement-Derived Feature Definitions and Computation

This appendix provides a concise but comprehensive description of the movement-derived feature families used in this study. The complete feature set consists of 51 variables, all computed from wrist-worn accelerometry signals and provided with the original HYPERAKTIV dataset. Rather than representing 51 conceptually distinct measures, these variables correspond to multiple parameterizations of a smaller number of fundamental feature families, each designed to capture a specific aspect of movement dynamics.

All features were extracted at the subject level from long-duration activity time series and were used identically in both the raw baseline model and the knowledge graph construction.

B.1 Variability-Based Features

Variability features quantify the dispersion and instability of movement intensity over time. These measures reflect the consistency of motor output and are commonly used in clinical movement analysis.

- `ACC__variance`: Statistical variance of acceleration magnitude, capturing overall dispersion of activity levels.
- `ACC__standard_deviation`: Square root of variance, expressing variability in the original measurement units.

- `ACC__variation_coefficient`: Ratio of standard deviation to mean activity, providing a scale-invariant measure of variability.
- `ACC__variance_larger_than_standard_deviation`: Indicator of whether variability exceeds mean-related dispersion.

B.2 Autocorrelation and Partial Autocorrelation Features

Autocorrelation-based features capture temporal dependencies in movement signals by measuring the similarity of the time series to delayed versions of itself.

- `ACC__autocorrelation__lag_`k, $k = 0, \ldots, 9$: Linear correlation between the signal and its lagged version, reflecting rhythmicity and temporal persistence.
- `ACC__partial_autocorrelation__lag_`k, $k = 0, \ldots, 9$: Correlation at lag k after removing contributions from shorter lags, isolating direct temporal dependencies.
- `ACC__agg_autocorrelation__f_agg_{mean, median, var}__maxlag_40`: Aggregated autocorrelation statistics summarizing temporal structure across multiple lags.

B.3 Entropy-Based Features

Entropy measures quantify the irregularity and unpredictability of movement time series, capturing deviations from periodic or highly regular motor patterns.

- `ACC__sample_entropy`: Estimates the likelihood that similar movement patterns remain similar at the next time step, with lower values indicating more regular behavior.
- `ACC__approximate_entropy__m_2__r_`r, $r = 0.1, \ldots, 0.9$: Measures signal regularity under different tolerance thresholds r, with higher values reflecting increased irregularity.
- `ACC__permutation_entropy__dimension_`d`__tau_1`, $d = 3, \ldots, 7$: Quantifies complexity based on ordinal patterns, capturing the diversity of local temporal arrangements.
- `ACC__fourier_entropy__bins_`b, $b = 2, 3, 5, 10, 100$: Entropy of the frequency-domain representation, reflecting spectral complexity of movement dynamics.

B.4 Complexity Features

Complexity measures assess the structural richness and compressibility of movement sequences, capturing information beyond simple variability or randomness.

- `ACC__lempel_ziv_complexity__bins_`b, $b = 2, 3, 5, 10, 100$: Estimates the algorithmic complexity of symbolized movement sequences at different discretization levels. Higher values indicate less repetitive and more complex motor patterns.

B.5 Temporal Asymmetry Features

Temporal asymmetry features capture directional properties of movement fluctuations, reflecting whether increases and decreases in activity occur symmetrically over time.

- `ACC__time_reversal_asymmetry_statistic__lag_`k, $k = 1, 2, 3$: Quantifies differences between forward and backward temporal dynamics, indicating non-linear and asymmetric motor control.

B.6 Multi-Scale Parameterization

For several feature families, multiple parameter values (e.g., lags, tolerance thresholds, embedding dimensions, discretization bins) were employed. These parameterizations do not represent distinct conceptual features but rather complementary views of the same underlying property across temporal and amplitude scales. All parameter variants were retained to preserve sensitivity to movement dynamics occurring at different resolutions.

B.7 Summary

In total, the 51 movement-derived features used in this study represent a structured and clinically grounded set of variability, entropy, complexity, and temporal descriptors. All features were used identically in both the raw baseline and the knowledge graph models. Consequently, any observed performance differences arise from the relational and structural representation introduced by the graph formulation, rather than from differences in feature selection or availability.

References

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Fournier, K. A., Hass, C. J., Naik, S. K., Lodha, N., and Cauraugh, J. H. (2010). Motor coordination in autism spectrum disorders: A synthesis and meta-analysis. *Journal of Autism and Developmental Disorders*, 40(10), 1227–1240. https://doi.org/10.1007/s10803-010-0981-3

Kerry, T. O., et al. (2019). The importance of assessing multidimensional gait profiles in children with ASD. *Journal of Autism and Developmental Disorders*, 49(8), 3357–3366. https://doi.org/10.1007/s10803-019-04168-w

Kindregan, D., Gallagher, L., and Gormley, J. (2015). Gait deviations in children with autism spectrum disorders: A review. *Autism Research and Treatment*, 2015, Article ID 741480. https://doi.org/10.1155/2015/741480

Mendes, J. O. H., Cordeiro, L. B., Santos, G. N., Duarte, N. A. C., Grecco, L. A. C., Moraes, P. M., and Oliveira, C. S. (2023). Evaluation of mental maturity and right-left orientation and correlations with gait and balance in children with autism spectrum disorder: Preliminary study. *Revista Conexão*, 16(9), Article e120. https://doi.org/10.55905/revconv.16n.9-120

Gligorijević, V., Renfrew, P. D., Kosciolek, T., Koehler Leman, J., Berenberg, D., Vatanen, T., Chandler, C., Taylor, B. C., Fisk, I. M., Vlamakis, H., Xavier, R. J., Knight, R., Cho, K., and Bonneau, R. (2021). Structure-based protein function prediction using graph convolutional networks. *Nature Communications*, 12, 3168. https://doi.org/10.1038/s41467-021-23303-9

Zhang, L., Zhang, M., Tang, J., Ma, J., Duan, X., Sun, J., Hu, X., and Xu, S. (2022). Analysis of traffic accident based on knowledge graph. *Journal of Advanced Transportation*, 2022, Article ID 3915467. https://doi.org/10.1155/2022/3915467

Kipf, T. N. and Welling, M. (2017). Semi-supervised classification with graph convolutional networks. In *Proceedings of the International Conference on Learning Representations (ICLR 2017)*. https://doi.org/10.48550/arXiv.1609.02907

Esposito, G., Venuti, P., Apicella, F., and Muratori, F. (2011). Analysis of unsupported gait in toddlers with autism. *Brain and Development*, 33(5), 367–373. https://doi.org/10.1016/j.braindev.2010.07.006

Weiss, M. J., Moran, M. F., Parker, M. E., and Foley, J. T. (2013). Gait analysis of teenagers and young adults diagnosed with autism and severe verbal communication disorders. *Frontiers in Integrative Neuroscience*, 7, Article 33. https://doi.org/10.3389/fnint.2013.00033

Eslami, T., Almuqhim, F., Raiker, J. S., and Saeed, F. (2021). Machine learning methods for diagnosing autism spectrum disorder and attention-deficit/hyperactivity disorder using functional and structural MRI: A survey. *Frontiers in Neuroinformatics*, 14, 575999. https://doi.org/10.3389/fninf.2020.575999

Heinsfeld, A. S., Franco, A. R., Craddock, R. C., Buchweitz, A., and Meneguzzi, F. (2018). Identification of autism spectrum disorder using deep learning and the ABIDE dataset. *NeuroImage: Clinical*, 17, 16–23. https://doi.org/10.1016/j.nicl.2017.08.017

Mostafa, S., Tang, L., and Wu, F.-X. (2019). Diagnosis of autism spectrum disorder based on eigenvalues of brain networks. *IEEE Access*, 7, 128474–128486. https://doi.org/10.1109/access.2019.2940198

Lim, H. C. and Al-Hababi, A. Y. S. (2024). A novel hybrid Unet-RBF and CNN-RBF algorithm for autism spectrum disorder classification. *Journal of Cognitive Sciences and Human Development*, 10(1), 87–102. https://doi.org/10.33736/jcshd.6778.2024

Takahashi, Y., Murata, S., Ueki, M., Tomita, H., and Yamashita, Y. (2022). Interaction between functional connectivity and neural excitability in autism: A novel framework for computational modeling and application to biological data. *PsyArXiv Preprint*. https://doi.org/10.31234/osf.io/ut39r

Xu, L., Geng, X., He, X., Li, J., and Yu, J. (2019). Prediction in autism by deep learning short-time spontaneous hemodynamic fluctuations. *Frontiers in Neuroscience*, 13, 1120. https://doi.org/10.3389/fnins.2019.01120

Niu, K., Guo, J., Pan, Y., Gao, X., Peng, X., Li, N., and Li, H. (2020). Multichannel deep attention neural networks for the classification of autism spectrum disorder using neuroimaging and personal characteristic data. *Complexity*, 2020, 1357853. https://doi.org/10.1155/2020/1357853

Raja, K. C. and Kannimuthu, S. (2023). Deep learning-based feature selection and prediction system for autism spectrum disorder using a hybrid meta-heuristics approach. *Journal of Intelligent & Fuzzy Systems*, 45(1), 797–807. https://doi.org/10.3233/jifs-223694

Sewani, H. and Kashef, R. (2020). An autoencoder-based deep learning classifier for efficient diagnosis of autism. *Children*, 7(10), 182. https://doi.org/10.3390/children7100182

Radhakrishnan, M., Karthik, R., Choudhury, K., Won, D., and Manoharan, T. A. (2021). Performance analysis of deep learning models for detection of autism spectrum disorder from EEG signals. *Traitement du Signal*, 38(3), 853–863. https://doi.org/10.18280/ts.380332

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A. Q., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., and Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8, 53. https://doi.org/10.1186/s40537-021-00444-8

Ahmad, I., Rashid, J., Faheem, M., Akram, A., Khan, N. A., and Amin, R. U. (2024). Autism spectrum disorder detection using facial images: A performance comparison of pretrained convolutional neural networks. *Healthcare Technology Letters*, 11(4), 227–239. https://doi.org/10.1049/htl2.12073

Elshoky, B. R. G., Younis, E. M. G., Ali, A. A., and Ibrahim, O. A. S. (2022). Comparing automated and non-automated machine learning for autism spectrum disorders classification using facial images. *ETRI Journal*, 44(4), 613–623. https://doi.org/10.4218/etrij.2021-0097

Hogan, A., Gutierrez, C., Cochez, M., de Melo, G., Kirrane, S., Polleres, A., Navigli, R., Ngonga Ngomo, A.-C., Rashid, S. M., Schmelzeisen, L., Staab, S., Blomqvist, E., d'Amato, C., Labra Gayo, J. E., Neumaier, S., Rula, A., Sequeda, J., and Zimmermann, A. (2022). *Knowledge Graphs* (1st ed.). Cham, Switzerland: Springer Nature.

Zhang, C., Li, Q., and Song, D. (2019). Aspect-based sentiment classification with aspect-specific graph convolutional networks. In K. Inui, J. Jiang, V. Ng, and X. Wan (Eds.), *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 4568–4578). Hong Kong, China: Association for Computational Linguistics. https://doi.org/10.18653/v1/D19-1464

Bitsko, R. H., Claussen, A. H., Lichstein, J. C., et al. (2022). Mental health surveillance among children — United States, 2013–2019. *MMWR Supplements*, 71(2), 1–42. doi:10.15585/mmwr.su7102a1.

El-Masry, H. M. A., Ibrahim, M. F. M., & Elsayed, M. F. F. (2021). Study of attention deficit hyperactivity disorder in children with idiopathic generalized epilepsy. *The Egyptian Journal of Hospital Medicine*, 85(1), 1–6. doi:10.21608/ejhm.2021.162851.

Li, D., Luo, X., Dang, C., Kong, Y., Li, Y., Zhu, Y., Huang, J., Wang, Y., Guo, J., Sun, L., & Song, Y. (2025). Impaired sensory–motor integration in children with attention-deficit/hyperactivity disorder during visual search. *Journal of Child Psychology and Psychiatry*. doi:10.1111/jcpp.70004.

Cole, W. R., Mostofsky, S. H., Gidley Larson, J. C., Denckla, M. B., & Mahone, E. M. (2008). Age-related changes in motor subtle signs among girls and boys with ADHD. *Neurology*, 71(17), 1335–1341. doi:10.1212/01.wnl.0000334275.57734.5f.

Bouchouras, G., Sofianidis, G., & Kotis, K. (2024). Temporal anomaly detection in attention-deficit/hyperactivity disorder using recurrent neural networks. *Cureus*, 16(12), e76496. doi:10.7759/cureus.76496.

Stergiou, N., Harbourne, R. T., & Cavanaugh, J. T. (2011). Optimal movement variability: A new theoretical perspective for neurologic physical therapy. *Neurologic Physical Therapy*, 35(3), 120–129. doi:10.1097/NPT.0b013e31821a2935.

Costa, M. D., Goldberger, A. L., & Peng, C.-K. (2005). Multiscale entropy analysis of biological signals. *Physical Review E*, 71(2), 021906. doi:10.1103/PhysRevE.71.021906.

Dhawale, A. K., Smith, M. A., & Ölveczky, B. P. (2017). The role of variability in motor learning. *Annual Review of Neuroscience*, 40, 479–498. doi:10.1146/annurev-neuro-072116-031548.

Cheng, W., Ji, X., Zhang, J., & Feng, J. (2012). Individual classification of ADHD patients by integrating multiscale neuroimaging markers and advanced pattern recognition techniques. *Frontiers in Systems Neuroscience*, 6, 58. doi:10.3389/fnsys.2012.00058.

Siqueira, A. d. S., Biazoli, C. E., Comfort, W. E., Rohde, L. A., & Sato, J. R. (2014). Abnormal functional resting-state networks in ADHD: Graph theory and pattern recognition analysis of fMRI data. *BioMed Research International*, 2014, 380531. doi:10.1155/2014/380531.

Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: MIT Press.

Latash, M. L. (2002). Functional redundancy in the motor system: The uncontrolled manifold hypothesis. *Motor Control*, 6(2), 123–151.

Stergiou, N., Harbourne, R. T., & Cavanaugh, J. T. (2006). Optimal movement variability: A new theoretical perspective for neurologic physical therapy. *Neurologic Physical Therapy*, 30(3), 120–129.

Castellanos, F. X., Sonuga-Barke, E. J. S., Milham, M. P., & Tannock, R. (2005). Characterizing cognition in ADHD: Beyond executive dysfunction. *Trends in Cognitive Sciences*, 9(3), 117–123. doi:10.1016/j.tics.2005.01.011.

Kofler, M. J., Rapport, M. D., Bolden, J., Sarver, D. E., & Raiker, J. S. (2013). ADHD and working memory: The impact of central executive deficits and time-based processes. *Journal of Abnormal Child Psychology*, 41(4), 535–547. doi:10.1007/s10802-012-9707-8.

Pincus, S. M. (1991). Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences of the United States of America*, 88(6), 2297–2301. doi:10.1073/pnas.88.6.2297.

Richman, J. S., & Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology–Heart and Circulatory Physiology*, 278(6), H2039–H2049. doi:10.1152/ajpheart.2000.278.6.H2039.

Bandt, C., & Pompe, B. (2002). Permutation entropy: A natural complexity measure for time series. *Physical Review Letters*, 88(17), 174102. doi:10.1103/PhysRevLett.88.174102.

Lempel, A., & Ziv, J. (1976). On the complexity of finite sequences. *IEEE Transactions on Information Theory*, 22(1), 75–81. doi:10.1109/TIT.1976.1055501.

Goldberger, A. L., Peng, C.-K., & Lipsitz, L. A. (2002). What is physiologic complexity and how does it change with aging and disease? *Neurobiology of Aging*, 23(1), 23–26. doi:10.1016/S0197-4580(01)00266-4.

Vaillancourt, D. E., & Newell, K. M. (2002). Changing complexity in human behavior and physiology through aging and disease. *Neurobiology of Aging*, 23(1), 27–31. doi:10.1016/S0197-4580(01)00247-0.

Zentgraf, K., Stark, R., Reiser, M., Kunz, M., Schienle, A., Kirsch, P., & Vaitl, D. (2009). Differential activation of prefrontal and sensorimotor cortices during motor imagery of everyday actions. *NeuroImage*, 47(4), 1584–1593. doi:10.1016/j.neuroimage.2009.05.065.

Haddad, J. M., Rietdyk, S., Claxton, L. J., & Huber, J. E. (2018). Task-dependent postural control strategies and the role of variability in motor performance. *Human Movement Science*, 59, 19–31. doi:10.1016/j.humov.2018.03.001.

Porrino, L. J., Lucignani, G., Dow-Edwards, D., & Sokoloff, L. (1983). Correlations between local brain glucose utilization and movement in the rat. *Science*, 219(4582), 269–271. doi:10.1126/science.6849134.

Teicher, M. H., Andersen, S. L., Polcari, A., Anderson, C. M., Navalta, C. P., & Kim, D. M. (1996). The neurobiological basis of childhood hyperactivity. *Behavioral and Brain Research*, 77(1–2), 1–12. doi:10.1016/0166-4328(95)00237-0.

Konofal, E., Lecendreux, M., & Cortese, S. (2001). Sleep and ADHD: A review. *Sleep*, 24(5), 1–10.

Hu, K., Ivanov, P. Ch., Chen, Z., Carpena, P., & Stanley, H. E. (2004). Effect of trends on detrended fluctuation analysis. *Physical Review E*, 69(4), 041107. doi:10.1103/PhysRevE.69.041107.

Wainwright, M. J. (2017). Statistical learning theory and optimization. *Foundations and Trends in Machine Learning*, 10(3–4), 261–412. doi:10.1561/2200000058.

Grover, A., & Leskovec, J. (2016). node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 855–864). doi:10.1145/2939672.2939754.

Hicks, S., Stautland, A., Fasmer, O. B., Førland, W., Hammer, H. L., Halvorsen, P., Mjeldheim, K., Oedegaard, K. J., Osnes, B., Syrstad, V. E. G., Riegler, M., & Jakobsen, P. (2021). HYPERAKTIV: An activity dataset from adult patients with attention-deficit/hyperactivity disorder (ADHD). In *Proceedings of the 12th ACM Multimedia Systems Conference*. doi:10.1145/3458305.3478454.