

AI and Network Theory Approaches for Studying and Diagnosing Brain Disorders

Alexander E. Hramov

Baltic Center for Neurotechnology and Artificial Intelligence,

Immanuel Kant Baltic Federal University,

Kaliningrad, Russia

Research Institute of Applied Artificial Intelligence and Digital Solutions,

Plekhanov Russian University of Economics,

Moscow, Russia

aekhramov@kantiana.ru

Abstract—Neurological and mental disorders represent a significant and growing global health challenge, characterized by late detection, diagnostic subjectivity, and a lack of objective biomarkers. This article explores the convergence of network theory and artificial intelligence (AI) as a powerful framework to address these issues. Network theory provides quantitative, interpretable metrics to describe the brain's structural and functional organization, framing many disorders as "dysconnectivity syndromes." AI and machine learning techniques, particularly graph neural networks (GNNs) and contrastive methods, are capable of identifying complex, multidimensional patterns in these network metrics, revealing disease subtypes and building predictive diagnostic models. We present case studies on Major Depressive Disorder and Autism Spectrum Disorder, demonstrating the superior performance of GNNs in classification tasks and the ability of contrastive learning to isolate disorder-specific functional connectivity patterns. The integration of network neuroscience and AI emerges as the most promising path toward developing reliable, objective biomarkers for early diagnosis and personalized treatment of brain disorders.

Index Terms—artificial intelligence, neurotechnology, legal regulatory standards, ethical dilemmas

I. INTRODUCTION

The diagnosis of neurological and psychiatric diseases faces several fundamental challenges. Statistics from the World Health Organization (2024) indicate a dementia diagnosis every three seconds globally. A critical issue is late detection, with the first symptoms of conditions like Alzheimer's disease appearing 10–15 years before clinical diagnosis [1]. Furthermore, diagnostic accuracy without advanced imaging is only 70–80%, highlighting significant subjectivity [2]. Projections show a 204% increase in Alzheimer's incidence by 2050 compared to 2020 [3].

Three intertwined problems exacerbate this situation:

- 1) *Heterogeneity*: Mental and neurological disorders are highly heterogeneous. Averaging data across a diagnostically homogeneous group can obscure significant differences, as different patient subgroups may exhibit distinct brain alterations.

This research was supported by the Russian Science Foundation (Project 23-71-30010).

- 2) *Lack of Objective Biomarkers*: The absence of quantifiable biomarkers leads to inaccurate diagnostics, trial-and-error treatment approaches, and hinders the development of new pharmaceuticals.

- 3) *The Replication Crisis in Neuroscience*: Small sample sizes, low statistical power, and heterogeneity contribute to difficulties in replicating scientific findings.

This lecture argues that the combined application of network theory and AI presents a viable solution to these challenges simultaneously.

II. NETWORK THEORY: A NEW LANGUAGE FOR THE BRAIN

Network theory offers a mathematical framework for describing the brain's complex organization [4]. The brain is represented as a graph $G = (\mathcal{N}, \mathcal{L})$, consisting of a set $\mathcal{N}$ of nodes (brain regions) and a set $\mathcal{L}$ of edges (connections between them) [5].

A. Building Brain Networks

The construction of a brain network is a foundational step that requires precise mathematical representation of neural architecture [6]. This process hinges on two core definitions: nodes and edges. Nodes correspond to distinct brain regions, effectively discretizing the continuous neural tissue into analyzable units. At the macroscale, this is achieved using standardized brain atlases, such as the Automated Anatomical Labeling (AAL) template, which parcellates the brain into a set of regions, each serving as a node in the graph. Edges represent the relationships or interactions between these nodes. The definition of an edge is not unitary and depends on the neurobiological question of interest, leading to two primary types of connectivity [7].

Structural connectivity refers to the physical wiring of the brain. It is reconstructed in vivo using diffusion-weighted imaging (DWI), a technique that traces the diffusion of water molecules to map the large-scale white matter fiber tracts linking anatomical regions. In contrast, functional connectivity captures statistical dependencies between the dynamic activity of brain regions, measured using functional magnetic

resonance imaging (fMRI) or electroencephalography (EEG). It quantifies how synchronized or correlated the neural signals are between regions, revealing functional ensembles that may communicate regardless of whether a direct anatomical pathway exists between them. This distinction is critical, as it allows researchers to investigate both the brain's physical infrastructure and its dynamic, operational state.

B. Graph Metrics as Potential Biomarkers

Graph theory provides a powerful mathematical toolkit for moving beyond visual inspection and delivering a quantitative, objective characterization of network topology. These metrics serve as potential biomarkers by capturing different aspects of how a network is organized.

Segregation describes the network's capacity for specialized processing within densely interconnected groups of nodes, often referred to as clusters or modules. A fundamental measure of segregation is the clustering coefficient, which quantifies the extent to which a node's neighbors are also connected to each other, indicating a tendency towards forming local cliques. High segregation suggests a network organized into specialized functional units.

Conversely, integration reflects the brain's ability to combine and rapidly distribute information from distributed regions. It is captured by metrics like global efficiency and characteristic path length. Global efficiency is inversely related to the average shortest path length between all pairs of nodes; a high-efficiency network has short path lengths, enabling efficient communication and integrated processing across the entire system.

Finally, node centrality measures aim to identify the most important or influential nodes within a network, often called hubs. Different metrics highlight distinct aspects of centrality: degree centrality simply counts the number of connections a node has, while betweenness centrality identifies nodes that act as bridges by lying on the shortest paths between many other nodes. Eigenvector centrality provides a more sophisticated measure by considering not just the number of connections, but also the importance of the nodes it is connected to, thus identifying nodes within influential neighborhoods. Together, these metrics paint a comprehensive picture of a network's architecture, from its local structure to its global integration and key influential elements.

Many brain disorders are considered "dysconnectivity syndromes," exhibiting alterations in these metrics. For example, schizophrenia is associated with reduced functional connectivity and a more dispersed, less integrated network structure compared to healthy controls [8].

III. AI ENTERS THE GAME: FROM METRICS TO DIAGNOSIS

The quantitative metrics provided by network theory create a rich feature space for pattern recognition. Artificial intelligence and machine learning algorithms excel at finding complex, multidimensional patterns within this high-dimensional data, transforming graph metrics from descriptive statistics into diagnostic tools in medicine [9].

A. The Challenge of Subtlety and the Need for Advanced Approaches

While classical machine learning classifiers—such as Linear Discriminant Analysis (LDA), Support Vector Machines (SVM), and Random Forests—provide a solid foundation by operating on pre-calculated global or local graph metrics, they face significant limitations. These methods rely heavily on feature engineering, where the choice of metrics dictates what the model can learn. In many brain disorders, the alterations in network organization are exceedingly subtle and distributed, meaning they are not fully captured by standard graph metrics like clustering coefficient or path length. These conventional metrics often reduce the complex architecture of a network to a single value, potentially obscuring the very patterns that differentiate health from disease or different stages of disease [10], [11]. This fundamental challenge necessitates a move beyond traditional tools towards more sophisticated modeling techniques capable of capturing the higher-order and nuanced interactions inherent in brain networks.

B. Graph Neural Networks: Learning from Topology Directly

Graph Neural Networks (GNNs) represent a paradigm shift by addressing the limitations of feature engineering. Instead of relying on pre-defined metrics, GNNs operate directly on the graph structure itself—the adjacency matrix [12]. This allows them to learn task-specific representations by propagating and transforming information across the network's connectivity pattern. By leveraging the actual topology, GNNs can capture more nuanced and complex relational features between nodes that are often invisible to manual feature engineering, making them particularly powerful for detecting the subtle and distributed dysconnectivity that characterizes many neurological and psychiatric disorders.

C. Higher-Order Networks and Contrastive Learning

The pursuit of even greater sensitivity has led to the exploration of frameworks that move beyond pairwise interactions. Hypergraphs, for instance, provide a natural way to model higher-order relationships, where an edge can connect more than two nodes simultaneously [13]. This allows for a more accurate representation of complex brain dynamics involving coordinated activity across multiple regions [14]. Combining these advanced representations with contrastive machine learning techniques offers a particularly promising path. Contrastive learning aims to learn representations by maximizing agreement between similar instances (e.g., brains with the same disorder) and minimizing agreement between dissimilar ones (e.g., healthy brains vs. diseased brains). This approach is exceptionally well-suited for isolating the salient, disorder-specific features of functional connectivity that are often hidden within the broader spectrum of neurotypical variation, thereby enhancing both the accuracy and interpretability of automated diagnosis [15].

TABLE I
ACCURACY OF CLASSIFICATION OF PATIENTS WITH MDD USING LDA ON DIFFERENT GRAPH FEATURE SETS.

Approach	Number of features	Accuracy	F1-score
GNN: bianarized full FC matrices	1350	0.99	0.95
LDA: full FC matrices	1350	0.92	0.89
LDA: Node-wise clustering coefficient	165	0.88	0.90
Global clustering coefficient	1	0.77	0.77

IV. CASE STUDIES

A. Major Depressive Disorder (MDD)

A study compared 70 MDD patients with 94 healthy controls using resting-state fMRI. The main results of this dataset analysis can be found in the Refs. [16]–[20]. Various ML approaches were applied to graph-based features extracted from the functional networks (see Table I).

While classical ML methods like LDA provided good interpretability, a GNN model applied directly to the functional connectivity matrices achieved superior classification performance [18]. The model's performance was linked to differences in the distribution of shortest path lengths between the groups, providing a potential neurobiological explanation for its decision. This demonstrates the ability of GNNs to leverage the full complex network topology for more sensitive and accurate diagnosis compared to methods relying on pre-calculated graph metrics. Classification based on only one characteristic describing the entire network as a whole is still relatively good.

B. Autism Spectrum Disorder (ASD)

Analyzing resting-state EEG data from 251 children with ASD and 253 neurotypical (NT) children revealed a significant challenge: the functional connectivity patterns in ASD showed strong similarities to NT patterns [21]. A contrastive learning approach, using a contrastive variational autoencoder, was employed to reconstruct an ASD-specific functional connectivity network [21]. This method successfully isolated salient features, primarily in the frontal lobe, leading to improved classification performance and revealing patterns entangled within the broader neurotypical connectivity. Furthermore, a hypergraph representation of multilayer brain networks was shown to enhance ASD detection compared to traditional graph approaches [22].

V. CONCLUSION AND PERSPECTIVES

The convergence of network neuroscience and AI is the most promising path toward developing objective, interpretable biomarkers for brain disorders. The key to success lies not in the blind application of ML but in a deep understanding of network construction methodology and a focus on model interpretability. Network theory provides the interpretable language to describe brain organization, and AI provides the tools to decode the complex patterns within it, paving the way for a new era in the diagnosis and understanding of brain disorders.

Despite the considerable promise of combining network neuroscience with AI, several grand challenges must be overcome to translate this potential into clinical reality.

A primary obstacle is the problem of causality. Current methodologies excel at identifying correlations and statistical associations within brain network data; however, moving beyond these descriptions to infer genuine causal mechanisms is a formidable task. Addressing this requires the integration of techniques from causal inference and perturbation modeling, which aim to distinguish mere coincidence from cause-and-effect relationships, ultimately revealing how alterations in network topology directly drive clinical symptoms.

Another critical challenge is the unification of methodological approaches. The field currently lacks universal, standardized pipelines for constructing and analyzing brain networks. Variations in data preprocessing, node definition, edge weighting, and thresholding strategies across research groups hinder the reproducibility and direct comparability of findings. The development of such standardized frameworks is essential for ensuring that results are not only accurate and explainable but also consistently replicable across different laboratories and clinical settings, a fundamental requirement for building trust in any potential biomarker.

Finally, the robust validation of AI models demands a new paradigm in data sharing and aggregation. The performance and generalizability of any diagnostic algorithm are contingent upon the quality and scale of the data it is trained on. There is a pressing need to accumulate large-scale, well-annotated, multi-site datasets that encompass the full heterogeneity of brain disorders. Building consolidated data resources, similar to those available for other medical imaging fields like radiology, would provide the necessary substrate for training more robust models, rigorously testing their efficacy, and ultimately paving the way for their successful adoption in clinical practice.

REFERENCES

- [1] C. R. Jack Jr, D. A. Bennett, K. Blennow, M. C. Carrillo, B. Dunn, S. B. Haeberlein, D. M. Holtzman, W. Jagust, F. Jessen, J. Karlawish *et al.*, "Nia-aa research framework: toward a biological definition of alzheimer's disease," *Alzheimer's & dementia*, vol. 14, no. 4, pp. 535–562, 2018.
- [2] J. Graff-Radford, K. X. Yong, L. G. Apostolova, F. H. Bouwman, M. Carrillo, B. C. Dickerson, G. D. Rabinovici, J. M. Schott, D. T. Jones, and M. E. Murray, "New insights into atypical alzheimer's disease in the era of biomarkers," *The Lancet Neurology*, vol. 20, no. 3, pp. 222–234, 2021.
- [3] M. Hao and J. Chen, "Trend analysis and future predictions of global burden of alzheimer's disease and other dementias: a study based on the global burden of disease database from 1990 to 2021," *BMC medicine*, vol. 23, no. 1, p. 378, 2025.
- [4] C. Seguin, O. Sporns, and A. Zalesky, "Brain network communication: concepts, models and applications," *Nature reviews neuroscience*, vol. 24, no. 9, pp. 557–574, 2023.

- [5] A. E. Hramov, N. S. Frolov, V. A. Maksimenko, S. A. Kurkin, V. B. Kazantsev, and A. N. Pisarchik, "Functional networks of the brain: from connectivity restoration to dynamic integration," *Physics-Uspekhi*, vol. 64, no. 6, p. 584, 2021.
- [6] J. Faskowitz, R. F. Betzel, and O. Sporns, "Edges in brain networks: Contributions to models of structure and function," *Network Neuroscience*, vol. 6, no. 1, pp. 1–28, 2022.
- [7] H.-J. Park and K. Friston, "Structural and functional brain networks: from connections to cognition," *Science*, vol. 342, no. 6158, p. 1238411, 2013.
- [8] M. P. Van Den Heuvel and A. Fornito, "Brain networks in schizophrenia," *Neuropsychology review*, vol. 24, no. 1, pp. 32–48, 2014.
- [9] O. E. Karpov, E. N. Pitsik, S. A. Kurkin, V. A. Maksimenko, A. V. Gusev, N. N. Shusharina, and A. E. Hramov, "Analysis of publication activity and research trends in the field of ai medical applications: Network approach," *International Journal of Environmental Research and Public Health*, vol. 20, no. 7, p. 5335, 2023.
- [10] V. Khorev, S. Kurkin, E. Pitsik, M. Radutnaya, E. Bondar, L. Mayorova, and A. Hramov, "A synergistic approach for identifying disrupted functional brain subnetworks in patients with chronic disorders of consciousness due to anoxic brain damage," *The European Physical Journal Special Topics*, pp. 1–12, 2025.
- [11] S. Kurkin, L. Mayorova, V. Khorev, E. Pitsik, M. Radutnaya, E. Bondar, and A. Hramov, "Multiscale fmri analysis reveals hierarchical network disruptions underlying disorders of consciousness," *Chaos, Solitons & Fractals*, vol. 200, p. 117008, 2025.
- [12] P. Pham, Q.-T. Bui, N. T. Nguyen, R. Kozma, P. S. Yu, and B. Vo, "Topological data analysis in graph neural networks: Surveys and perspectives," *IEEE Transactions on Neural Networks and Learning Systems*, 2025.
- [13] S. Boccaletti, P. De Lellis, C. Del Genio, K. Alfaro-Bittner, R. Criado, S. Jalan, and M. Romance, "The structure and dynamics of networks with higher order interactions," *Physics Reports*, vol. 1018, pp. 1–64, 2023.
- [14] S. Zhang, H. Zhao, W. Wang, Z. Wang, X. Luo, A. Hramov, and J. Kurths, "Edge-centric effective connection network based on multi-modal mri for the diagnosis of alzheimer's disease," *Neurocomputing*, vol. 552, p. 126512, 2023.
- [15] Y. Zong, Q. Zuo, M. K.-P. Ng, B. Lei, and S. Wang, "A new brain network construction paradigm for brain disorder via diffusion-based graph contrastive learning," *IEEE transactions on pattern analysis and machine intelligence*, 2024.
- [16] S. A. Kurkin, N. M. Smirnov, R. Paunova, S. Kandilarova, D. Stoyanov, L. Mayorova, and A. E. Hramov, "Beyond pairwise interactions: higher-order q-analysis of fmri-based brain functional networks in patients with major depressive disorder," *IEEE Access*, vol. 12, p. 197168, 2024.
- [17] V. S. Khorev, S. A. Kurkin, G. Zlateva, R. Paunova, S. Kandilarova, M. Maes, D. Stoyanov, and A. E. Hramov, "Disruptions in segregation mechanisms in fmri-based brain functional network predict the major depressive disorder condition," *Chaos, Solitons & Fractals*, vol. 188, p. 115566, 2024.
- [18] E. N. Pitsik, V. A. Maksimenko, S. A. Kurkin, A. P. Sergeev, D. Stoyanov, R. Paunova, S. Kandilarova, D. Simeonova, and A. E. Hramov, "The topology of fmri-based networks defines the performance of a graph neural network for the classification of patients with major depressive disorder," *Chaos, Solitons & Fractals*, vol. 167, p. 113041, 2023.
- [19] A. V. Andreev, S. A. Kurkin, D. Stoyanov, A. A. Badarin, R. Paunova, and A. E. Hramov, "Toward interpretability of machine learning methods for the classification of patients with major depressive disorder based on functional network measures," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 33, no. 6, 2023.
- [20] D. Stoyanov, V. Khorev, R. Paunova, S. Kandilarova, D. Simeonova, A. Badarin, A. Hramov, and S. Kurkin, "Resting-state functional connectivity impairment in patients with major depressive episode," *International Journal of Environmental Research and Public Health*, vol. 19, no. 21, p. 14045, 2022.
- [21] M. S. Kabir, S. Kurkin, G. Portnova, O. Martynova, Z. Wang, and A. Hramov, "Contrastive machine learning reveals in eeg resting-state network salient features specific to autism spectrum disorder," *Chaos, Solitons & Fractals*, vol. 185, p. 115123, 2024.
- [22] E. Pitsik, S. Kurkin, O. Martynova, G. Portnova, and A. E. Hramov, "Hypergraph representation of multilayer brain network enhances autism spectrum disorder detection," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 35, no. 7, 2025.