

Computers in Biology and Medicine

Exploiting Graph Convolutional Networks for Insightful Classification and Explanation of Traumatic Brain Injury

--Manuscript Draft--

Manuscript Number:	CIBM-D-24-12879
Article Type:	Full Length Article
Keywords:	Brain connectivity; Brain modelling; Diffusion-weighted imaging; Explainable AI; Grad-CAM; Graph-based machine learning; Neuroimaging.
Corresponding Author:	Tiziana Currieri University of Palermo ITALY
First Author:	Tiziana Currieri
Order of Authors:	Tiziana Currieri Joan Falcó-Roget Elham Rostami Salvatore Vitabile Alessandro Crimi
Abstract:	Traumatic brain injury (TBI) presents complex challenges in diagnosis and treatment, requiring advanced neuroimaging and machine learning techniques to improve diagnostics, prognostics and patient stratification. Indeed, the identification of neurodegeneration in chronic patients is currently of pivotal interest and difficult to perform solely with visual inspection of magnetic resonance imaging. This study employed a graph convolutional network to classify patients into acute, chronic, and healthy control groups based on structural brain connectivity data derived from diffusion-weighted imaging. The model achieved a classification accuracy of 83.67%, with precision, recall, and F1-scores of 81.6%, 78%, and 79%, respectively. To enhance model interpretability, we applied gradient-weighted class activation mapping (Grad-CAM), which identified key brain regions, such as the thalamus, anterior cingulate cortex, and frontal cortex, as critical in distinguishing patient groups. Notably, the Grad-CAM analysis also revealed a shift from widespread neural disruption in acute TBI to more localized impairments in chronic patients. Our results provide new insights into the neural mechanisms underlying TBI and disorders of consciousness, emphasizing the role of explainable artificial intelligence (XAI) in clinical neuroscience. This research also highlights the potential to predict patient trajectories and outcomes at different stages of TBI recovery. Indeed, this study highlights the promise of combining advanced machine learning techniques like graph convolutional networks with XAI methods to enhance the diagnosis, prognostication and stratification of TBI patients.
Suggested Reviewers:	Alaa Bessadok Helmholtz Artificial Intelligence Cooperation Unit a.bessadok@tue.nl expertise: GNN for brain connectivity
	Andras Jakab University Children Hospital Zürich andras.jakab@kispi.uzh.ch expertise: infant brain connectivity
	Martijn van den Heuvel VU Amsterdam Center for Neurogenomics and Cognitive Research martijn.vanden.heuvel@vu.nl expertise: brain connectivity data from MRI
	Alessandra Griffa Lausanne University Hospital alessandra.griffa@epfl.ch

	expertise: different aspects of connectomes and machine learning
	<p>Stefan Kiebel Technische Universitat Dresden stefan.kiebel@tu-dresden.de Computational neuroscience, machine learning, neuroimaging</p>
	<p>Jennifer K. Steeves York University steeves@yorku.ca Neuropsychology, traumatic brain injury, neuroimaging</p>
	<p>David Van Essen Washington University School of Medicine in Saint Louis vanessen@wustl.edu Cognitive neuroscience, neuroimaging, brain mapping</p>
Opposed Reviewers:	

**Declaration of interests**

- The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
- The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:



Direttore Prof. Giuseppe Ferraro

Prof. Huiling Chen, Prof. Feng ZHU and Prof. Quan Zou
 Editors-in-Chief
 Computers in Biology and Medicine

Dears Prof. Huiling Chen, Prof. Feng ZHU and Prof. Quan Zou,

My name is Tiziana Currieri, and I am a PhD Student with the University of Palermo.
 I am writing to submit our manuscript entitled:

“Exploiting Graph Convolutional Networks for Insightful Classification and Explanation of Traumatic Brain Injury”

by authors:

Tiziana Currieri, Joan Falcó-Roget, Elham Rostami, Salvatore Vitabile and Alessandro Crimi

as a regular contribution to Computers in Biology and Medicine.

The integration of advanced neuroimaging techniques and machine learning has proven vital in the diagnosis and classification of traumatic brain injury (TBI), a condition that affects millions of individuals worldwide. In this study, we employed graph convolutional networks (GCNs) to classify TBI patients into acute, chronic, and control groups based on structural brain connectivity derived from diffusion-weighted imaging. Our model achieved high classification accuracy and identified key brain regions, such as the thalamus and anterior cingulate cortex, as critical in distinguishing these groups. To enhance model interpretability, we applied gradient-weighted class activation mapping (Grad-CAM), allowing us to visualize and identify the brain regions most responsible for the classification decisions. This innovative approach bridges the gap between complex machine learning techniques and clinical applicability, offering new insights into the neural mechanisms underlying TBI and disorders of consciousness.

While the dataset used in this study is relatively small, our model demonstrated strong performance and the ability to generalize across different clinical conditions. This result underlines the potential of GCNs in the field, even with limited data, and highlights the importance of continuing research with larger datasets to further validate and refine the approach.



UNIVERSITÀ
DEGLI STUDI
DI PALERMO

DIPARTIMENTO DI BIOMEDICINA,
NEUROSCIENZE E DIAGNOSTICA
AVANZATA Bi.N.D.



Direttore Prof. Giuseppe Ferraro

All authors declare no conflict of interest to declare, and this manuscript is not under consideration elsewhere. We kindly request that you consider our manuscript for publication in Computers in Biology and Medicine, as we believe it offers a significant contribution to both computational neuroscience and clinical applications of artificial intelligence in healthcare.

Thank you for your time and consideration.

The contact author for this submission is:

Tiziana Currieri,

Department of Biomedicine, Neuroscience and Advanced Diagnostics (BiND)
University of Palermo
Via del Vespro, 129 – 90127 Palermo – Italy
Phone +39 091 238 65 720
E-mail: tiziana.currieri@unipa.it, tiziana.currieri@gmail.com

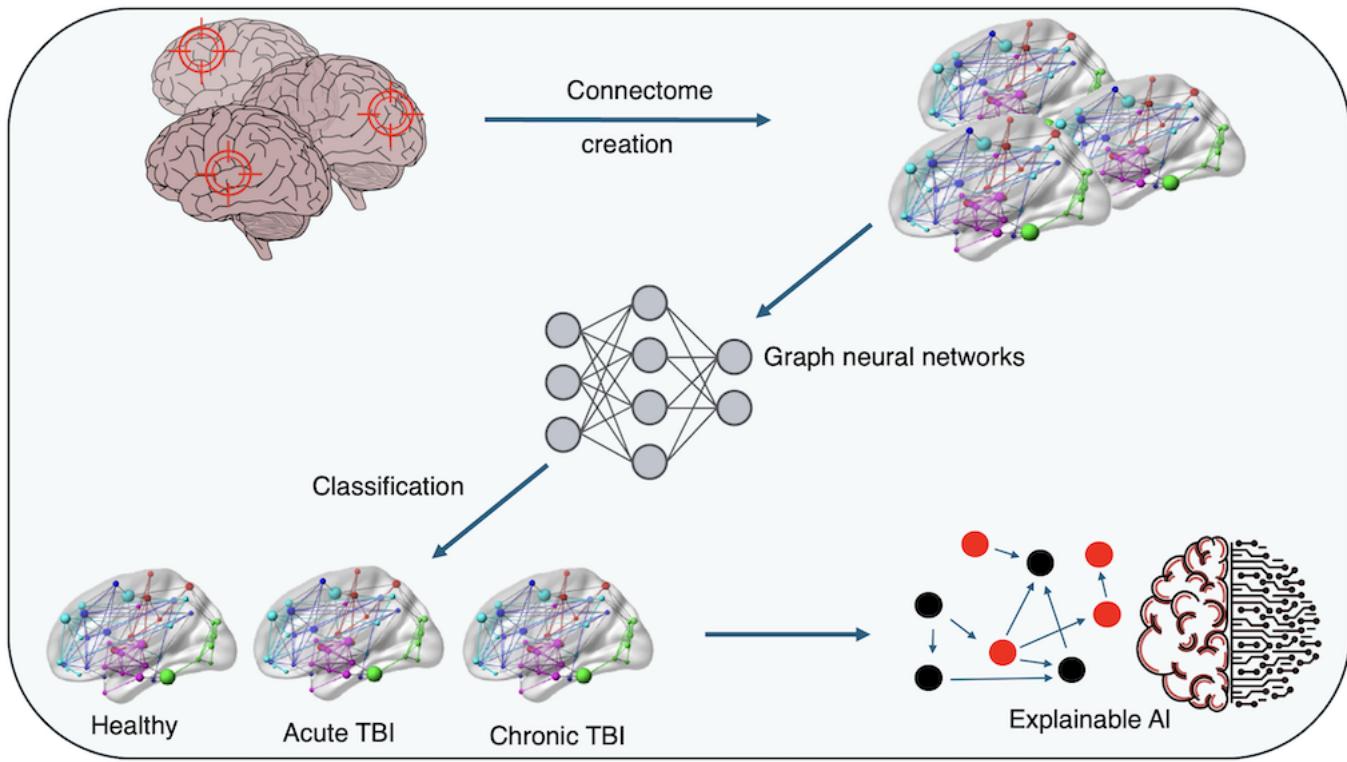
Yours sincerely
(Tiziana Currieri)



Graphical Abstract

Exploiting Graph Convolutional Networks for Insightful Classification and Explanation of Traumatic Brain Injury

Tiziana Currieri, Joan Falcó-Roget, Elham Rostami, Salvatore Vitabile, Alessandro Crimi



Highlights

Exploiting Graph Convolutional Networks for Insightful Classification and Explanation of Traumatic Brain Injury

Tiziana Currieri, Joan Falcó-Roget, Elham Rostami, Salvatore Vitabile, Alessandro Crimi

- A GCN-based model with explainability techniques is proposed to classify TBI patients into acute, chronic, and control groups.
- Grad-CAM revealed critical areas like the thalamus and frontal cortex essential for distinguishing TBI stages.
- Findings highlight a progression from broad neural damage in acute cases to localized issues in chronic TBI.

[Click here to view linked References](#)

Exploiting Graph Convolutional Networks for Insightful Classification and Explanation of Traumatic Brain Injury

Tiziana Currieri^{a,*}, Joan Falcó-Roget^b, Elham Rostami^{c,d}, Salvatore Vitabile^a and Alessandro Crimi^e

^a Department of Biomedicine, Neuroscience and Advanced Diagnostics, University of Palermo, Palermo, Italy

^b Sano Centre for Computational Medicine, Kraków, Poland

^c Department of Medical Sciences, Neurology, Uppsala University, Uppsala, Sweden

^d Department of Neuroscience, Karolinska Institutet, Stockholm, Sweden

^e AGH Science and Technology, Kraków, Poland

ARTICLE INFO

Keywords:

Brain connectivity
Brain modelling
Diffusion-weighted imaging
Explainable AI
Grad-CAM
Graph-based machine learning
Neuroimaging

ABSTRACT

Traumatic brain injury (TBI) presents complex challenges in diagnosis and treatment, requiring advanced neuroimaging and machine learning techniques to improve diagnostics, prognostics and patient stratification. Indeed, the identification of neurodegeneration in chronic patients is currently of pivotal interest and difficult to perform solely with visual inspection of magnetic resonance imaging. This study employed a graph convolutional network to classify patients into acute, chronic, and healthy control groups based on structural brain connectivity data derived from diffusion-weighted imaging. The model achieved a classification accuracy of 83.67%, with precision, recall, and F1-scores of 81.6%, 78%, and 79%, respectively. To enhance model interpretability, we applied gradient-weighted class activation mapping (Grad-CAM), which identified key brain regions, such as the thalamus, anterior cingulate cortex, and frontal cortex, as critical in distinguishing patient groups. Notably, the Grad-CAM analysis also revealed a shift from widespread neural disruption in acute TBI to more localized impairments in chronic patients. Our results provide new insights into the neural mechanisms underlying TBI and disorders of consciousness, emphasizing the role of explainable artificial intelligence (XAI) in clinical neuroscience. This research also highlights the potential to predict patient trajectories and outcomes at different stages of TBI recovery. Indeed, this study highlights the promise of combining advanced machine learning techniques like graph convolutional networks with XAI methods to enhance the diagnosis, prognostication and stratification of TBI patients.

1. Introduction

Traumatic brain injury (TBI) represents one of the most severe public health challenges, affecting more than 60 million individuals worldwide that may result in prolonged disorders of consciousness (DoC) such as the minimally conscious state (MCS) or vegetative state (VS)[1, 2]. Diagnosing and classifying patients across these various states of consciousness remains a complex task, given the wide heterogeneity in clinical presentations and neural damage patterns [3]. Traditional neuroimaging techniques, while informative, often fail to fully capture the intricacies of brain network disruptions caused by TBI [4]. As a result, the integration of machine learning methods, particularly graph convolutional networks (GCNs), has emerged as a powerful approach to analyze structural brain connectivity and provide insights into the neural underpinnings of these disorders. A crucial aspect to consider is that the incidence of TBI, especially in mild cases, is often underestimated; in fact, many patients tend to minimize their symptoms and do not go to the emergency room [5]. Neuroimaging in head trauma patients is also not always mandatory [6], especially considering the high costs and the use of valuable scanner time that could benefit patients with other conditions. Studies have shown that only a small fraction of patients

with presumed minor head injuries show positive results on computed tomography (CT) scans, and less than 1% require neurosurgical intervention. Moreover, reducing CT scans for minor head injury patients by even 10% could result in substantial annual savings [7, 8]. Distinguishing between minor and severe head injuries can be complex. The severity of TBI is typically assessed using the Glasgow Coma Scale (GCS), which evaluates a patient's level of consciousness based on eye, verbal, and motor responses. A lower GCS score indicates more severe injury. For patients with more severe TBI, as indicated by a lower GCS score or additional factors such as prolonged loss of consciousness, focal neurological deficits, or signs of skull fractures, a CT scan is typically performed to investigate the extent of the injury. [7, 9, 10]. Despite the various standards established to guide the use of imaging, no criterion is foolproof, and patients who have no obvious clinical signs and are considered low risk may still show serious abnormalities on imaging [11]. This leads to considerable diversity in TBI datasets, which are often limited and small in size, reflecting the challenges in identifying patients who would truly benefit from neuroimaging.

1.1. Related Works

Building on this context, the present study leverages GCNs in conjunction with explainable AI (XAI) techniques, such as gradient-weighted class activation mapping (Grad-CAM), to enhance the interpretability of these models in a

*Corresponding author

 tiziana.currieri@unipa.it (T. Currieri)
ORCID(s): 0009-0007-8062-4233 (T. Currieri)

clinical context [12]. Recent advancements in neuroimaging and machine learning have highlighted the efficacy of GCNs for classifying patients with neurological disorders based on brain connectivity patterns [13, 14, 15]. GCNs extend the principles of convolutional neural networks (CNNs) to graph-structured data, making them particularly suited for analyzing brain networks, where nodes represent brain regions and edges signify the strength of connections between them. Studies by [16, 17] have demonstrated the utility of GCNs in disease prediction, particularly in conditions like dementia and autism spectrum disorder, by capturing subtle patterns of connectivity that might be missed by traditional approaches. [18] explored the use of GCNs with Euclidean distance-based graph construction for classifying Parkinson's disease, achieving a high accuracy of 97.4%, which highlights the effectiveness of graph-based deep learning approaches in neurodegenerative disease diagnosis. [19] leverages advanced graph convolutional networks to categorize different clinical forms of multiple sclerosis based on brain morphological connectivity, highlighting the potential of machine learning in enhancing diagnostic precision and offering insights into the neurodegenerative processes associated with the disease. However, while GCNs have shown promising results in classifying neurological conditions, the challenge of interpreting these complex models remains significant. In clinical settings, the "black-box" nature of deep learning models can limit their utility, as clinicians require transparent and interpretable outputs to guide decision-making processes [20]. This is where XAI techniques, such as Grad-CAM, become invaluable. XAI methods aim to make machine learning models more interpretable by highlighting the regions or features most responsible for the model's decisions, thus bridging the gap between model performance and clinical trust [21]. Grad-CAM, originally developed for CNNs, has been successfully adapted to GCNs, allowing researchers to visualize the contribution of specific regions (nodes) to model predictions [22, 23]. This technique can provide critical insights into the neural mechanisms underlying TBI by identifying key brain areas involved in motor control, consciousness, and cognitive regulation, as demonstrated in studies by [24], following injury. Our work builds on these prior studies by applying GCNs to a dataset of acute and chronic TBI patients, along with healthy controls, to classify patient groups based on their brain connectivity patterns. In doing so, we employ Grad-CAM to identify the most significant brain regions driving the classification decisions, thus enhancing the interpretability of our model. One of the unique contributions of our study is the focus on explainability in the context of disorders of consciousness, a domain where understanding the underlying neural networks is crucial for personalized treatment strategies. Moreover, our approach is aligned with recent research that emphasizes the need for explainable models in medical AI, where transparency and trustworthiness are paramount. XAI, as defined by researchers like [25], offers methods to make AI models more interpretable without sacrificing performance. This is especially important in TBI,

where subtle patterns of brain disconnection can have profound clinical implications. By integrating GCNs with Grad-CAM, we not only achieve high classification accuracy but also provide a deeper understanding of the disrupted neural pathways in TBI patients. This, in turn, can inform clinical decisions, particularly in identifying which patients may benefit from targeted interventions such as neurostimulation or personalized rehabilitation.

1.2. Contribution and Organization

The main contributions are:

- Development and in-depth analysis of the optimised GCN model for the classification of TBI patients using structural connectors derived from diffusion imaging to classify patients in acute and chronic phases.
- The integration of Grad-CAM explainability into GCN for TBI classification, allowing the identification of the most relevant brain regions involved in different stages of TBI.
- Given the heterogeneity and complexity of traumatic brain injuries, this study systematically identifies and formalizes the brain regions most affected, providing a structured foundation for a deeper understanding of the injury's impact.

The subsequent structure of this paper is laid out as follows: Section 2 provides a comprehensive overview of the materials and methodologies, explaining the pre-processing steps and analytical tools used throughout the study. Section 3 elaborates on applying these methods specifically to the research objectives. Section 4 presents the experimental results, focusing on the classification performance and insights gained from the XAI techniques. Section 5 discusses these results in a clinical context, emphasizing their implications and relevance. Finally, Section 6 concludes the paper by summarizing key findings and suggesting potential future research avenues.

2. Materials and Methods

2.1. Dataset

The present study utilized a dataset publicly available on the OpenNeuro platform (<https://openneuro.org/datasets/ds003367>), comprising 40 participants divided into three distinct groups: acute, chronic, and control [26]. Diffusion-weighted imaging (DWI) data were acquired using a Siemens Skyra 3 Tesla scanner equipped with a 32-channel head coil. The imaging protocol employed High Angular Resolution Diffusion Imaging (HARDI) and utilized sixty diffusion encoding directions, supplemented by 10 b0 volumes, with a b-value of 2000 s/mm² [27].

2.1.1. Acute Group

The acute group comprised 18 subjects who had sustained severe TBI. The inclusion criteria for this group were stringent, encompassing a GCS score of 6 or lower and

no eye-opening for at least 24 hours. Notably, two subjects succumbed to their injuries during acute hospitalization; however, their data were retained in the analyses to provide a comprehensive understanding of the acute phase of severe TBI. A subgroup of 9 subjects underwent follow-up imaging approximately five months post-initial hospitalization, having regained consciousness and allowing for longitudinal evaluation of recovery.

2.1.2. Chronic Group

The chronic group included 6 patients with DoC persisting for at least five months post-injury, along with 9 patients who transitioned from the acute cohort. The group of 6 subjects consisted of individuals in VS and MCS, enabling the investigation of structural connectivity disruptions associated with long-term impaired consciousness [28]. Notably, 4 patients are in MCS and 2 are in VS, with their Coma-Revised Recovery Scale (CRS-R) scores reflecting different levels of clinical severity. VS is characterized by wakefulness without conscious awareness, where patients show no voluntary or purposeful reactions to visual, auditory, or tactile stimuli [29]. In contrast, MCS patients exhibit intentional behaviours but lack effective communication abilities [30]. MCS can be further divided into MCS+ (where patients demonstrate higher-level responses, such as following commands) and MCS- (with lower-level behaviours, including visual tracking or localized responses to pain) [31].

2.1.3. Control Group

For comparative analysis, a control group of 16 healthy subjects, matched for age and sex, was recruited to establish normative baseline data and mitigate potential confounding demographic variables.

2.2. Image Preprocessing

Diffusion-weighted images often exhibit artefacts that require thorough preprocessing to ensure data reliability. In this study, advanced software tools such as FSL [32], MRtrix3 [33], and DIPY [34] were used to perform the necessary corrections and enhance image quality. Given the absence of T1 and T2-weighted images in the dataset, the b0 volume extracted from the DWI was chosen as the anatomical reference. Initially, skull removal was performed using FSL's BET2 algorithm [35] to eliminate non-brain tissues and improve the delineation of brain structures. Subsequently, specific corrections were applied to mitigate magnetic artefacts and geometric distortions caused by magnetic susceptibilities [36]. To increase the signal-to-noise ratio and optimize overall image quality, denoising based on multivariate principal component analysis (MP-PCA) [37] was implemented, allowing for the reduction of intrinsic thermal noise. Additionally, Gibbs artefacts were corrected to eliminate errors arising from image discretization during reconstruction [38]. Furthermore, corrections were made to compensate for eddy currents, which are induced magnetic fields generated by rapid gradient changes during diffusion

image acquisition. These currents can cause geometric distortions and intensity variations in DWI images, compromising data fidelity. Eliminating eddy currents, along with correcting motion artefacts, is essential to ensure accurate reconstruction of nerve fibre trajectories [39]. At the end of preprocessing, the images were rescaled to an isotropic resolution of 1.5 mm^3 per voxel and subsequently corrected for bias field inhomogeneities [40], ensuring greater uniformity of signal intensity throughout the entire brain volume [41]. The preprocessing of DWI data involved normalizing the images to the MNI152 1.5 mm template using nonlinear registration algorithms to ensure precise alignment across subjects. The registration primarily utilized the 'mrregister' workflow from MRtrix3, which is specifically designed for DWI data. This method was chosen for its robustness in aligning DWI data with standard brain templates. Anatomical labels were extracted from the Automated Anatomical Labeling (AAL3v1) atlas [42], which divides the brain into 170 distinct regions. These regions serve as nodes within the structural network, each marked for detailed neuroanatomical analysis. During the registration process, an affine transformation was selectively applied to the AAL3v1 atlas for some subjects to align it with each individual's brain anatomy. This selective application was necessary to accommodate variations in individual anatomy, ensuring that the atlas regions accurately overlaid the MRI scans of those specific subjects. This step is crucial for the precise extraction of anatomical labels, establishing a solid foundation for further analysis. After the atlas alignment, the diffusion gradient matrix (b-matrix) was appropriately reoriented to maintain the accuracy of diffusion direction measurements post-registration [43]. This adjustment of the b-matrix, in line with the affine transformation parameters, ensured the anatomical consistency of the diffusion data, which is vital for maintaining the integrity of diffusion tensor imaging and enhancing the reliability of subsequent fibre tractography analyses. It is important also to note that in the AAL3v1 atlas, the original identifiers for the anterior cingulate cortex (ACC) and the thalamus have been replaced with new subdivisions, leading to the omission of these regions [42]. Consequently, these unassigned areas have been interpreted as isolated nodes within the brain network, lacking interconnections with other regions. This absence can influence the network's topology and the integration of information among different brain areas. Therefore, to improve the accuracy of cerebral connectivity representation by focusing on interconnected areas, future research is considering excluding these two regions. Fig 1 illustrates the brain atlas before and after the registration process, highlighting the alignment achieved and ensuring that the structural connectivity accurately reflects the underlying neuroanatomical structures.

2.3. Generation of the Structural Connectome via Streamline Tractography

Before creating the tractography and the structural connectome, fundamental intermediate steps were performed.

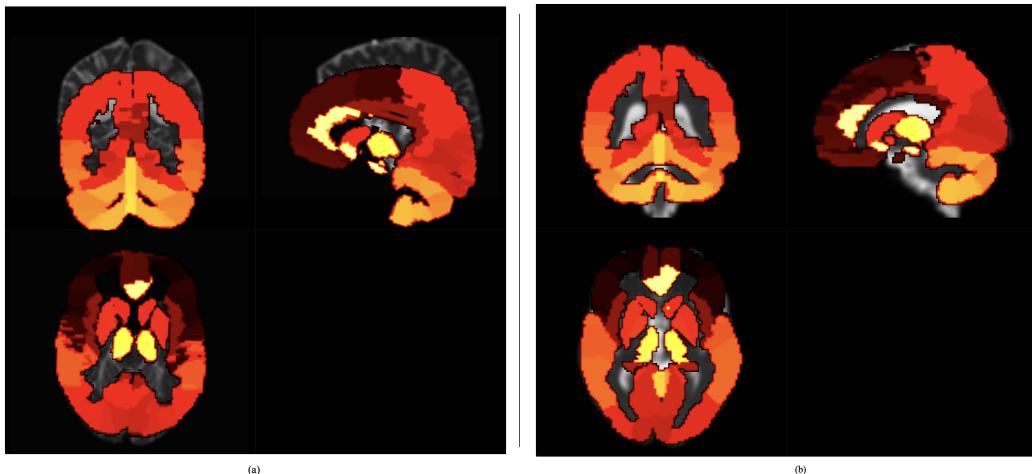


Figure 1: Panel (a) presents the pre-rotation alignment of chronic patient brain scans with the AAL3v1 atlas, shown in axial, sagittal, and coronal views. Panel (b) depicts the final aligned scans post-rotation, utilizing the same atlas. The colours represent different brain regions, facilitating spatial delineation and analysis.

Specifically, the analysis of streamlines began by placing seeds at strategic positions along the grey matter–white matter boundary. Each streamline was generated starting from a seed, tracing a path through the white matter until terminating in another region of grey matter. Some streamlines could end in anatomically implausible positions, such as at the boundary of cerebral ventricles; such anomalous trajectories were eliminated, ensuring that most connections represented valid links between distant grey matter regions. Creating an accurate boundary between grey matter and white matter was crucial. In the absence of T1- and T2-weighted acquisitions, we used the 5TTgen segmentation with FSL workflow [33] on b0 volumes to delineate grey matter, white matter, and cerebrospinal fluid boundaries. Importantly, T2-weighted and b0 images are equivalent. We then manually corrected the segmentation maps by swapping the white matter and cerebrospinal fluid labels, ensuring proper tissue contrast for accurate white matter fibre reconstruction. Subsequently, a single-shell three-tissue constrained spherical deconvolution (SS3T-CSD) [44, 45] was employed using MRtrix3 to estimate the fibre orientation distribution (FOD) in each brain voxel. This technique, adapted for single-shell DWI data, facilitates the differentiation of white matter, grey matter, and cerebrospinal fluid within a single diffusion measurement, enhancing the precision of the FOD estimations. The FOD values were then normalized to equalize intensity variations across different voxels, thereby enhancing the consistency of the data throughout the entire brain. Anatomically constrained probabilistic tractography (ACT) [46] was used, limiting the analysis to white matter through detailed anatomical segmentations that guide the propagation of streamlines. Backtracking was implemented, allowing trajectories to retrace and recalculate the path if they terminated in implausible regions, improving the accuracy of the generated connections. Seed points were placed at the junction between white matter and grey matter, representing the entry and exit points of nerve fibres. The maximum

length of streamlines was set to 250 mm to avoid generating excessively long and anatomically improbable paths, and a cutoff threshold set at 0.06 defined the minimum FOD amplitude necessary to continue trajectory propagation. To refine the connectivity model, streamline filtering was implemented using SIFT2 [47], which optimizes the selection of streamlines to ensure a more accurate and representative connectivity matrix. This approach adjusts the contribution of each streamline to match the underlying MR signal, reducing biases in connectivity estimates and improving the anatomical accuracy of the connectome. A total of 10,000,000 streamlines were generated to ensure a detailed and statistically robust representation of brain connections. Using the nerve fibre trajectories a weighted connectivity matrix of size 170×170 was constructed, corresponding to the brain regions defined in the AAL3v1 atlas. The entries in this matrix represent the total number of streamlines connecting each pair of brain regions, effectively detailing the network of connections across different areas. The values in the matrix indicate the strength and density of these neural connections, providing a comprehensive map of brain connectivity. Zero values were observed in some positions of the matrix, indicating the absence of direct connections between certain regions. Subsequent analyses revealed that patients suffering from chronic conditions exhibit a higher number of isolated nodes in their brain networks, despite a lower number of subjects in this group compared to others. This phenomenon suggests a higher level of structural disconnection, potentially indicative of disruptions in neuronal connectivity or alterations in brain communication pathways. These results highlight the importance of deepening the study of structural connectivity in chronic patients to better understand the neurobiological bases of their clinical conditions. Fig 2 delineates this analytical workflow utilized for investigating brain connectivity, encapsulating the sequence of methodologies from the initial acquisition of imaging data to the comprehensive analysis of the network.

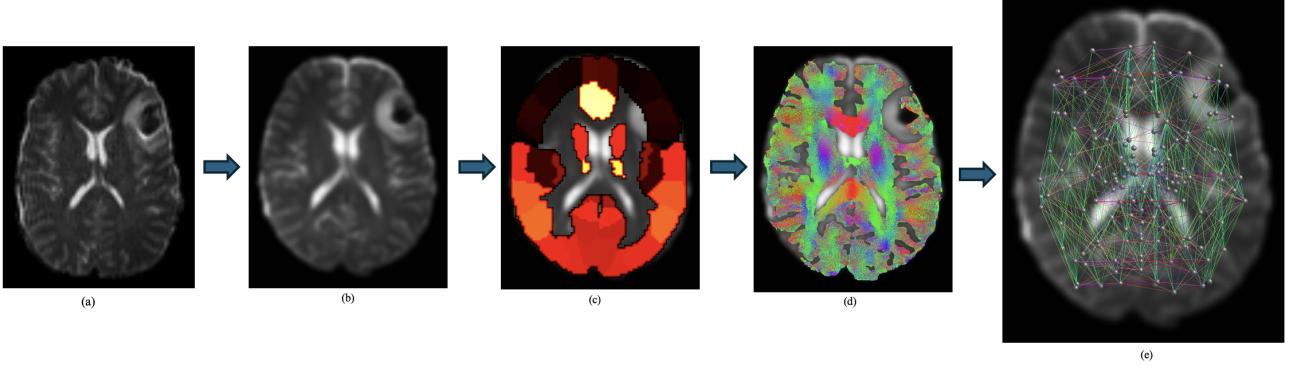


Figure 2: Sequential image processing: (a) original diffusion-weighted image, (b) pre-processed image, (c) atlas registration (d) diffusion tensor imaging representing fibre density, (e) connectome imaging with structural brain network mapping.

This illustration succinctly depicts the transition from raw diffusion-weighted imaging scans, through the enhancement of connectivity patterns, to the ultimate construction of the connectome, providing a systematic visual representation of the methodological approach.

2.4. Introduction to Graphs

A graph is a fundamental mathematical structure used to model relationships and interactions between objects or entities. Formally, a graph is defined as

$$(G = (V, E)) \quad (1)$$

where:

- V is a finite set of nodes (or vertices),
- $E \subseteq V \times V$ is a set of edges connecting pairs of nodes [48].

In the context of undirected and weighted graphs, which are often employed to represent brain networks, each edge $(i, j) \in E$ has an associated weight w_{ij} that quantifies the strength or capacity of the connection between node i and node j . The matrix representation of a graph is given by the adjacency matrix $A \in \mathbb{R}^{N \times N}$, where $N = |V|$ is the number of nodes. The elements of the matrix are defined as:

$$A_{ij} = \begin{cases} w_{ij}, & \text{if there is an edge between nodes } i \text{ and } j; \\ 0, & \text{otherwise.} \end{cases}$$

This matrix allows the application of algebraic and analytical methods to study the properties of the graph. In the field of computational neuroscience, graph theoretical approaches have become indispensable tools for modelling and analyzing complex brain networks. In such models, nodes represent discrete anatomical or functional regions of the brain, while edges denote the connections between these regions, which can be structural (e.g., white matter tracts identified via diffusion tensor imaging) or functional (e.g., statistical correlations of neural activity measured through

functional MRI) [49]. The application of graph theory enables the quantitative assessment of the brain's topological properties, facilitating a deeper understanding of how network organization relates to cognitive functions and neurological disorders [50, 51].

2.5. Graph Convolutional Networks

GCNs are a class of neural networks designed to operate directly on graph-structured data, extending the principles of CNNs to non-Euclidean domains [52]. GCNs effectively capture the complex relationships and dependencies among nodes in a graph by aggregating and transforming feature information from a node's local neighbourhood.

In traditional CNNs, convolutional operations are defined on regular grids, such as images, where the local neighbourhood of a pixel is well-defined. However, graphs have an irregular structure, making the definition of convolution operations less straightforward. GCNs address this by generalizing the convolution operation to the graph domain, enabling the network to learn representations that consider both the features of nodes and the topology of the graph. One widely adopted formulation of GCNs is based on spectral graph theory, where the convolution operation is defined in the spectral domain using the graph Laplacian. [53] proposed a more computationally efficient and scalable approach, which introduced a first-order approximation of spectral graph convolutions.

The propagation rule for a single GCN layer is defined as:

$$H^{(l+1)} = \sigma(\tilde{A}^{-1/2} \tilde{D}^{-1/2} H^{(l)} W^{(l)}), \quad (2)$$

where:

- $H^{(l)} \in \mathbb{R}^{N \times F^{(l)}}$ is the matrix of activations at layer l , with $F^{(l)}$ features per node.
- $H^{(0)} = X$ is the initial node feature matrix.
- $\tilde{A} = A + I$ is the adjacency matrix with added self-loops (where I is the identity matrix).

- \tilde{D} is the degree matrix of \tilde{A} , with $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$.
- $W^{(l)} \in \mathbb{R}^{F^{(l)} \times F^{(l+1)}}$ is the trainable weight matrix for layer l .
- $\sigma(\cdot)$ is a nonlinear activation function, such as ReLU.

This formulation performs a weighted aggregation of each node's features with those of its neighbours, normalized by the degrees of the nodes. The addition of self-loops ensures that a node's features are included in the aggregation, which is crucial for preserving the information at each node. GCNs iteratively update node representations by aggregating information from their local neighbourhoods, allowing the network to capture both local and global structural information. By stacking multiple GCN layers, nodes can incorporate information from nodes several hops away, effectively learning hierarchical feature representations. In the context of brain network analysis, GCNs are particularly powerful tools. Brain connectivity data can naturally be represented as graphs, where nodes correspond to anatomical regions of interest (ROIs) and edges represent structural connections between these regions derived from DWI.

By training the GCNs on these graphs, the network learns to generate node embeddings that capture both local connectivity patterns and global network structures. This method allows us to model complex interactions within the brain network and identify subtle connectivity changes that may indicate TBI or facilitate the discovery of patterns associated with neurological conditions. GCNs have been successfully applied in various neuroimaging studies. For example, [54] employed GCNs for disease prediction in autism spectrum disorder and Alzheimer's disease, demonstrating their effectiveness in capturing disease-specific patterns in brain connectivity. Similarly, [17] used GCNs to classify mild cognitive impairment based on functional MRI data, highlighting the potential of GCNs in early diagnosis.

2.6. Explainability with Grad-CAM

In the realm of neural network models, particularly GCNs, the interpretability of model decisions is crucial, especially when applied to sensitive domains like neuroscience and clinical diagnosis. To address this, we employed XAI methods, specifically Grad-CAM, to gain insights into the decision-making processes of our trained GCN model. Although it was originally developed for CNNs in image classification tasks [55, 56] recent advancements have adapted this technique for use with graph neural networks (GNNs), including GCNs, enabling the interpretation of model decisions in graph-structured data [57].

In our study, we employed this adapted version of Grad-CAM to elucidate the inner workings of our GCN model, which is critical given the complex and non-Euclidean nature of brain connectivity graphs. The importance of XAI in our context cannot be overstated, as it allows us to identify specific brain regions (nodes) that significantly influence the model's predictions, thereby facilitating a deeper understanding of the neurobiological underpinnings associated with different clinical conditions. By computing the

gradients of the output class scores concerning the node features or embeddings, we generated class activation maps that highlight the contribution of each node to the final prediction. Mathematically, for a given graph G and target class c , the Grad-CAM importance score for node i can be computed as:

$$\text{Importance}_i^c = \text{ReLU} \left(\sum_k \alpha_k^c h_i^k \right) \quad (3)$$

where h_i^k represents the activation of node i at layer k , and α_k^c is the importance weight of feature map k for class c , calculated by globally averaging the gradients over all nodes:

$$\alpha_k^c = \frac{1}{N} \sum_{i=1}^N \frac{\partial y_c}{\partial h_i^k} \quad (4)$$

Here, y_c is the output score (logit) for class c , N is the total number of nodes in the graph, and $\frac{\partial y_c}{\partial h_i^k}$ denotes the gradient of y_c concerning the activation h_i^k of node i at layer k . The ReLU function ensures that only positive contributions are considered, focusing on features that positively influence the class score. This approach has been validated in prior research. For instance, [22] extended Grad-CAM to GNNs, demonstrating its efficacy in interpreting models across various graph classification tasks, such as molecular property prediction and social network analysis. They adapted traditional visualization methods to graph data and showed that gradient-based techniques could successfully highlight important substructures within graphs, providing valuable insights into the model's predictions. Similarly, [58] applied Grad-CAM to graph-based models, allowing for the visualization of discriminative features across multiple brain imaging modalities. It identifies significant brain regions and illustrates their contribution to distinguishing between healthy controls, mild cognitive impairment, and Alzheimer's disease, providing valuable insights into the neural mechanisms underlying these conditions. This interpretability is particularly crucial in clinical applications, where understanding the rationale behind a model's decision can inform diagnostic processes and guide personalized treatment strategies [59]. For example, if certain brain regions are consistently identified as significant across patients with a specific condition, this could highlight potential targets for therapeutic intervention or further investigation. Therefore, employing XAI techniques like Grad-CAM bridges the gap between complex deep learning models and practical clinical utility, aligning with the broader movement in the literature toward developing interpretable and trustworthy artificial intelligence systems in healthcare [60].

Moreover, the adaptation of Grad-CAM for GCNs on graph-structured data has been explored in several studies, underscoring its relevance and applicability in our research. For instance, [61] presented a spatiotemporal graph convolutional network, STGC-GCAM, which applies

Grad-CAM to analyze fMRI data for identifying functional connectivity biomarkers in Alzheimer's disease. This model aids in highlighting crucial brain regions and connections, improving the interpretability of the network's outputs and providing insights relevant to diagnosing and understanding Alzheimer's disease.

By integrating Grad-CAM into our analysis, we enhanced the transparency of our GCN model. This allowed us to correlate the most influential nodes with known neurological functions and clinical observations, which is essential for building trust in the model's predictions and for potentially uncovering novel insights into the pathophysiology of DoC and TBI. The ability to interpret the model's decisions at the node level provides a means to validate the model against established neuroscientific knowledge and to explore new hypotheses about brain connectivity patterns associated with different clinical states.

3. Graph Analysis and Network Applications

3.1. Graph Metrics Used

In our analysis of brain network connectivity, we employed several key graph metrics to characterize the topological properties of the structural connectomes derived from diffusion-weighted imaging data. The local features selected include *degree*, *PageRank*, *betweenness centrality*, *local efficiency*, *average neighbour degree*, *weighted clustering coefficient*, *eigenvector centrality*, and *Katz centrality* [62]. Below, we provide definitions and mathematical formulations for each metric:

1. Degree (k_i):

The degree of a node i represents the total strength of its connections to other nodes in the network. In weighted graphs, it is calculated as:

$$k_i = \sum_{j=1}^N A_{ij}, \quad (5)$$

where A_{ij} is the weight of the edge between nodes i and j , and N is the total number of nodes [63].

2. PageRank (PR_i):

PageRank measures the influence of a node based on the concept of link analysis, originally developed for ranking web pages. It is defined recursively as:

$$PR_i = \frac{1-d}{N} + d \sum_{j \in M_i} \frac{PR_j}{k_j^{\text{out}}}, \quad (6)$$

where d is the damping factor (typically set to 0.85), M_i is the set of nodes linking to node i , and k_j^{out} is the out-degree of the node j [64].

3. Betweenness Centrality (BC_i):

Betweenness centrality quantifies the importance of a node in terms of the number of shortest paths passing through it. It is given by:

$$BC_i = \sum_{s \neq i \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}, \quad (7)$$

where σ_{st} is the total number of shortest paths between nodes s and t , and $\sigma_{st}(i)$ is the number of those paths that pass through node i [65].

4. Local Efficiency ($E_{\text{loc}}(i)$):

Local efficiency reflects how efficiently information is exchanged by the immediate neighbours of a node when it is removed. It is calculated as:

$$E_{\text{loc}}(i) = \frac{1}{k_i(k_i - 1)} \sum_{j,h \in N_i} ((A_{jh} + A_{hj}) \times l_{jh}^{-1}), \quad (8)$$

where N_i is the set of neighbors of node i , and l_{jh} is the shortest path length between nodes j and h [66].

5. Average Neighbor Degree (AND_i):

This metric computes the average degree of the neighbours of node i :

$$AND_i = \frac{1}{k_i} \sum_{j \in N_i} k_j, \quad (9)$$

providing insight into the connectivity of a node's immediate network [67].

6. Weighted Clustering Coefficient (C_i):

The weighted clustering coefficient assesses the tendency of a node's neighbours to form tightly knit groups, considering the weights of the connections:

$$C_i = \frac{1}{s_i(k_i - 1)} \sum_{j,h \in N_i} \frac{(w_{ij} + w_{ih})}{2} \times w_{jh}, \quad (10)$$

where $s_i = \sum_{j \in N_i} w_{ij}$ is the strength of node i , w_{ij} is the weight of the edge between nodes i and j , and the sums are over all pairs of neighbors j and h of node i . [68]

7. Eigenvector Centrality (EC_i):

Eigenvector centrality assigns relative scores to all nodes in the network based on the principle that connections to high-scoring nodes contribute more to the score of the node:

$$EC_i = \frac{1}{\lambda} \sum_{j=1}^N A_{ij} EC_j, \quad (11)$$

where λ is the largest eigenvalue of the adjacency matrix A [69].

8. Katz Centrality (KC_i):

Katz centrality measures the relative influence of a node within a network by considering the total number of walks between nodes, penalized by the length of the walks:



Figure 3: Spearman correlation matrix displaying relationships between various node features in a network analysis.

$$KC_i = \alpha \sum_{j=1}^N A_{ij} KC_j + \beta, \quad (12)$$

where α is a damping factor (must be less than the reciprocal of the largest eigenvalue of A), and β is a constant representing the initial centrality [63, 70].

Fig 3 illustrates the Spearman correlation matrix, highlighting the pairwise correlations between various graph metrics employed in our analysis. Despite the presence of significant correlations between certain metrics, such as Degree and Average Neighbor Distance, and Weighted Clustering and Local Efficiency, we opted to retain these features in our analysis. The decision to include these metrics stems from their ability to offer complementary insights into different structural and functional aspects of the network. Furthermore, maintaining a diverse set of metrics enhances the robustness of our predictive models, enabling them to capture a broader range of underlying patterns and dynamics within the brain network. Additionally, certain correlated features are identified as statistically significant predictors in our models, providing unique and valuable contributions to our understanding of the network's complexities.

3.2. Graph Convolutional Network Implementation

In our study, we utilize GCNs to analyze the structural connectomes constructed from DWI data. Each node in the graph represents a brain region defined by the AAL3v1 atlas, and edges are weighted by the strength of the structural connections between regions derived from tractography. For each node within the graph, we associated a feature vector comprising the graph metrics calculated and explained in

the previous section. Specifically, each node was characterized by an 8-dimensional feature vector. By assigning this vector of features to each node, we provided the GCN with rich information capturing both the local properties of nodes and their roles within the global network topology. The model consisted of three graph convolutional layers that progressively extracted high-level representations of node features while considering the graph structure. The specific architecture included a first graph convolutional layer that accepts the 8-dimensional node feature vector as input and produces 32 output features per node. The second graph convolutional layer takes the 32-dimensional node features from the previous layer and outputs 64 features per node. The third graph convolutional layer processes the 64-dimensional node features and outputs 128 features per node, as depicted in the attached figures.

Fig 4 illustrates the overall workflow from the initial graph input through the convolution and pooling stages to the final classification output. Fig 5 provides a detailed view of the convolution and activation process, highlighting the transformation of node features through successive layers with Leaky ReLU activation and dropout, culminating in global average pooling (GAP). After each convolutional layer, the Leaky ReLU activation function with a negative slope of 0.01 was applied to introduce non-linearity. It is an activation function that allows a small, non-zero gradient when the unit is not active (i.e., when the input is negative), which helps mitigate the "dying ReLU" problem where neurons become inactive and stop learning. Batch normalization was performed after each activation to stabilize and accelerate the training process, and dropout layers with a rate of 0.5 were incorporated to prevent overfitting. Following the convolutional layers, a GAP operation was used to aggregate node-level features into a graph-level representation, resulting in a 128-dimensional graph-level feature vector that captures the most prominent signals in the graph. The pooled graph representation was then passed through a series of fully connected layers: the first fully connected layer transformed the 128-dimensional input into 64 units, the second reduced the 64 units to 32 units, the third further reduced the 32 units to 16 units, and the output layer mapped the 16 units to 3 output units corresponding to the three classes (acute patients, chronic patients, and healthy controls). Importantly, this layer did not include an activation function, outputting raw logits. The model was trained using the Adam optimizer with an initial learning rate of 1×10^{-5} and a weight decay of 1×10^{-4} to enhance performance and prevent overfitting. The loss function used was a cross-entropy loss, suitable for multiclass classification problems which expects raw, unnormalized scores (logits) as input and internally applies the LogSoftmax operation during loss computation. The training process involved 450 epochs, providing sufficient iterations for the model to converge.

To evaluate the model's performance and generalization capability, we conducted leave-one-out cross-validation (LOOCV). In this approach, the model was trained on all subjects except one and tested on the excluded subject. This

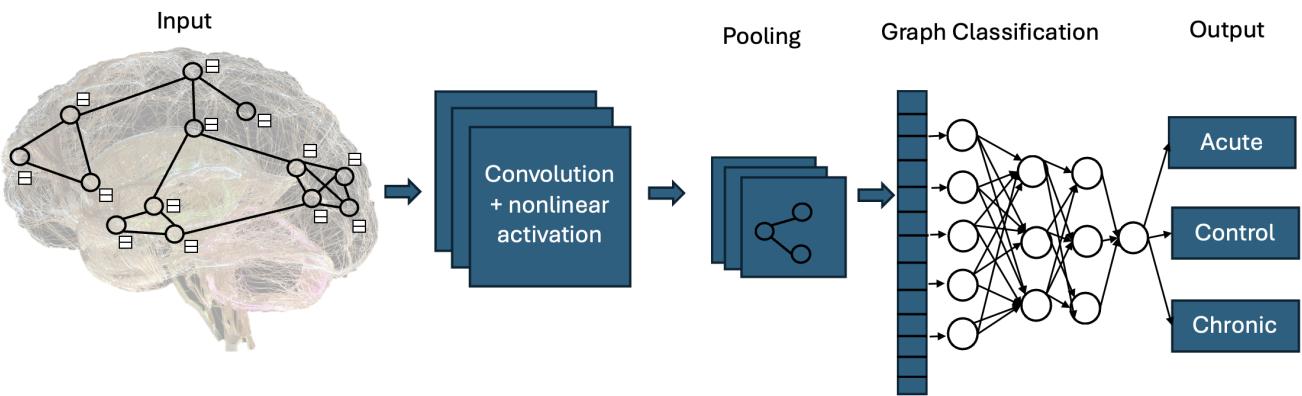


Figure 4: Overview of a graph convolutional network applied to brain connectivity data, illustrating the workflow from input graph through feature extraction and pooling, to final classification into three categories.

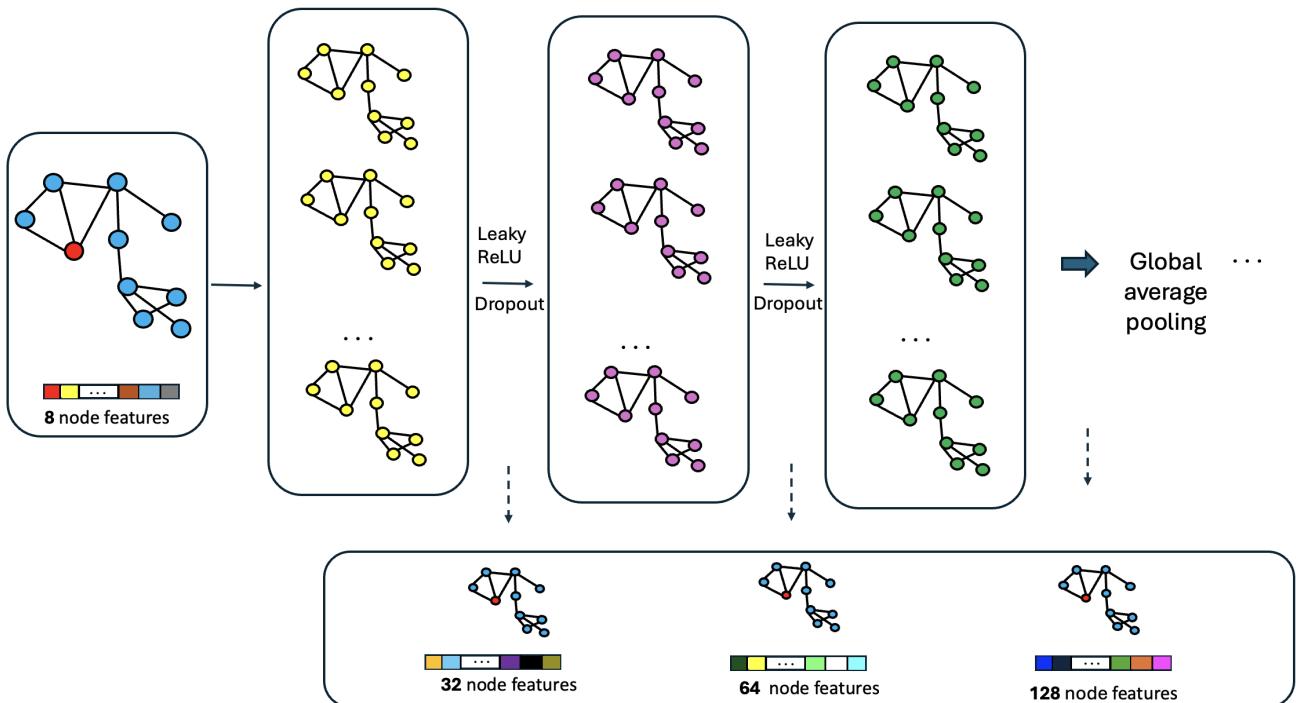


Figure 5: Detailed view of the convolution and nonlinear activation stages in a graph neural network, showing feature transformation from 8 node features to 128 node features through successive layers with Leaky ReLU activation and dropout, culminating in global average pooling.

process was repeated for each subject in the dataset, resulting in a comprehensive assessment of the model's ability to generalize to unseen data. Model performance was evaluated in terms of accuracy, precision, recall, and F1 Score. The computational experiments were conducted using Python 3.9 and PyTorch 2.2. The classification was performed on an iMac equipped with a 3.1 GHz Intel Core i5 processor (6 cores) and an AMD Radeon Pro 5300 graphics card with 4 GB of VRAM.

3.3. Unveiling Key Brain Regions with Grad-CAM

To apply Grad-CAM to our trained GCN model and identify the most significant nodes for each patient and

subsequent group, we implemented a method that leverages the gradients of the output class scores concerning the input node features, as outlined in the previous section. For each patient's brain connectivity graph, we performed a forward pass through the GCN model to obtain the raw output logits for all classes. Focusing on the target class corresponding to the patient's true label, we computed the gradients of the class score for the input node features by backpropagating through the network. These gradients capture the sensitivity of the model's predictions to changes in each node's features, effectively highlighting which nodes have the most significant impact on the classification outcome. To aggregate this

information into a single importance score per node, we averaged the gradients across all feature dimensions for each node. This process resulted in a scalar value for each node, representing its overall contribution to the model's prediction for the target class. By applying the Rectified Linear Unit (ReLU) activation function to these scores, we focused on nodes that positively influence the class score, aligning with the original Grad-CAM methodology. After obtaining the importance scores, we ranked all nodes in descending order of their scores to identify the most significant ones.

After calculating the Grad-CAM importance scores, we identified the most significant nodes and edges that influenced our GCN model's predictions for each patient and group. To achieve this, we selected the top 30 nodes and edges, which constitute approximately 17.6% of the total 170 nodes in each brain connectivity graph. This threshold was determined through methodological experimentation and domain expertise. In our exploratory analysis, reducing the number of selected nodes below this threshold failed to provide a clear separation between patients within the same clinical group, hindering the identification of the most significant brain regions. Conversely, selecting more than 30 nodes often included regions of less relevance, which diluted the meaningful patterns and increased the risk of overfitting. Therefore, selecting the top 30 nodes achieved an optimal balance, providing sufficient detail to discern impactful patterns and maintaining the analysis within a manageable scope, thereby optimizing computational efficiency and analytical precision. From a neuroscientific perspective indeed, focusing on this subset allowed us to relate our findings to established brain networks and functions without oversimplifying the inherently complex connectivity patterns of the brain. This approach is supported by practices in the neuroimaging field, where identifying and examining key nodes with high centrality measures has been shown to enhance the understanding of network dynamics [71, 72]. For instance, [71] emphasized the importance of a subset of regions in understanding the brain's modular organization, while [72] demonstrated that concentrating on highly interconnected and central nodes, or the "rich-club," provides valuable insights into the brain's global communication efficiency. By mapping these significant nodes back to their corresponding anatomical regions using the AAL3v1 atlas, we were able to correlate the model's predictions with known neuroanatomical and functional properties, thereby providing a deeper understanding of the neural mechanisms underlying DoC and TBI. By strategically selecting and analyzing the top 30 nodes, we enhanced the interpretability of our findings, facilitating the identification of key brain regions that contribute to the classification of clinical conditions, and thereby bridging the gap between complex deep learning models and meaningful neuroscientific insights. For visualization, we employed NetworkX and Matplotlib to generate graphical representations of the brain connectivity networks. Nodes were coloured based on their importance scores, and labels were assigned according to their anatomical regions. This

Table 1
Classification Performance Metrics of the GCN Model

Metric	Average
Accuracy	83.67%
Precision	81.6%
Recall	78%
F1-Score	79%

visual representation aided in illustrating the spatial distribution of significant brain regions and their connections, making it easier to identify patterns and differences across patient groups.

4. Results

4.1. GCN Classification Performance

The performance of our GCN model in classifying different patient groups was evaluated using key metrics, including accuracy, precision, recall and F1 score. In all iterations of the LOOCV procedure, the model achieved an average accuracy of 83.67%, demonstrating a high level of correctness in its classification predictions. In addition, the model achieved an average accuracy of 81.6% of positive predictions, ensuring a low false-positive rate, which is crucial in the clinical setting to avoid misclassifying healthy individuals as patients. The average recall was 78%, reflecting the model's ability to correctly identify true positives, ensuring that a substantial percentage of real patient conditions are accurately recognised. The F1 score, which harmonises accuracy and recall, was 79%, emphasising the model's balanced performance in identifying true positives and minimising false negatives.

As illustrated in Table 1, the average metrics indicate the robust overall performance of the GCN model in distinguishing between different patient groups. The consistently high average metrics demonstrate the model's overall reliability and effectiveness in leveraging complex brain connectivity patterns to distinguish between patient groups. The implications of these results are significant for clinical diagnosis and patient stratification. The high accuracy and precision indicate that the GCN model can serve as a reliable tool for classifying patients based on their brain connectivity patterns, which is essential for developing accurate diagnostic tools, personalized treatment plans and patient trajectories. The balanced recall ensures that the model effectively identifies a substantial proportion of true positive cases, enhancing its utility in clinical environments where comprehensive identification of patient conditions is paramount. Furthermore, the identification of the worst-performing fold provides valuable insights into the model's limitations and areas for potential improvement. Analyzing the specific characteristics of the misclassified case can inform strategies to enhance the model's robustness, such as incorporating additional data augmentation techniques, refining feature engineering processes, or exploring architectural modifications to better handle diverse patient profiles.

4.2. Grad-CAM Highlights: Significant Brain Regions by Patient Group

Our Grad-CAM analysis was performed to identify the most significant nodes and edges in the brain networks of patients, focusing on how specific brain regions contribute to the classification of TBI. The analysis primarily focused on identifying the top 30 most influential nodes for each patient, rather than their corresponding edges. This approach was chosen to emphasize the most critical brain regions, which are more directly relevant to understanding neural dysfunction in the context of TBI and DoC.

Although the analysis primarily focuses on identifying the most significant nodes, Fig 6 provides a complementary visualization, showing the most important edges connecting these 170 brain regions for the acute patient depicted in Fig 2. The edges, colour-coded by weight, highlight the strength of the connections between these regions. Warmer colours indicate stronger connections, reflecting higher importance as determined by Equation 4. Notably, there is a strong edge between Raphe_M (node 170) and VTA_R (node 160), a connection between two critical areas involved in modulating arousal and autonomic functions, which are often disrupted in DoC [73]. Similarly, the edge between Thal_AV_R (node 122) and Thal_MDm_R (node 136) reflects robust communication within thalamic nuclei that are essential for sensory integration and consciousness regulation. Additionally, a prominent connection exists between Raphe_D (node 169) and Raphe_M (node 170), further emphasizing the role of brainstem structures in maintaining wakefulness and alertness [74, 75]. This visualization illustrates how the key regions identified in the analysis interact within the broader network, offering insights into the neural pathways that may be involved in consciousness disorders. While the focus of this study remains on nodes, the role of these critical edges will be explored in future work.

Fig 7(a) and (b) provide an example of a detailed visualization of Grad-CAM analysis applied to the same acute patient, showing the significant nodes in the network. Panel (a) displays a focused representation of the network, excluding the isolated nodes, thus emphasizing the broader connectivity and highlighting the most influential regions within a well-defined anatomical context. Each node is labelled using the AAL3v1 atlas, showcasing regions consistently identified across multiple analyses as pivotal in understanding the neuropathology associated with TBI and DoC. Conversely, panel (b) presents a general visualization that includes isolated nodes. Notably, nodes such as Cingulate_Ant_L (35), Cingulate_Ant_R (36), Thalamus_L (81), and Thalamus_R (82), though highly significant according to the Grad-CAM results, were not labelled by the atlas but were divided into many other areas [42]. These nodes appeared as isolated regions in the network but had a significant impact on the model's predictions, suggesting that while these nodes played a role in classification, they do not correspond to anatomically recognized areas. These isolated nodes are excluded from functional analyses due to their lack of anatomical labelling but still highlight areas where the

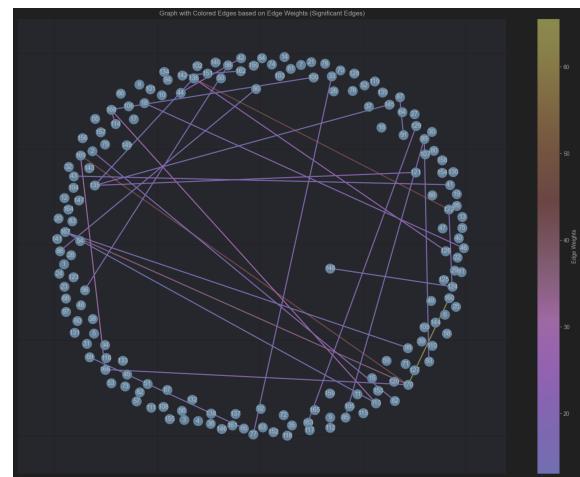


Figure 6: Graph showing all 170 nodes (representing brain regions) and the most significant edges for the acute patient shown in Fig 2. The edges are colour-coded based on their weight, with stronger connections displayed in warmer colours. This visualization emphasizes the key connections between brain regions identified through Grad-CAM analysis.

model identified critical connections. Additionally, isolated nodes such as Thal_Re_R (node 134) and Thal_Re_L (node 133) show elevated Grad-CAM values, further indicating their relevance in the network for this acute patient. The Reuniens nuclei in the thalamus coordinate communication between the prefrontal cortex and the hippocampus, which are critical for memory consolidation and cognitive functions, often disrupted in acute brain injuries [76]. The presence of these isolated thalamic nodes suggests that disruptions in these pathways could play a key role in the patient's altered consciousness and cognitive deficits. The inclusion of these isolated nodes highlights their disproportionate influence on the network, suggesting potential disruptions in critical communication pathways that could contribute to the patient's clinical symptoms.

By going into detail, in Fig 7(a) notable nodes are observed such as Thal_LP_R (node 124), Thal_PuA_R (node 148), Vermis_1_2 (node 113), LC_R (node 168), Thal_MDI_R (node 138), VTA_R (node 160) and Vermis_3 (node 114). These regions are critically involved in sensory processing, motor control, and the regulation of arousal, all of which are frequently disrupted in patients with severe TBI. From a clinical perspective, the thalamic nuclei, particularly Thal_LP_R and Thal_PuA_R, are crucial for relaying sensory information and maintaining attentional processes. Damage or dysfunction in these areas often leads to deficits in sensory integration and attention regulation [77, 78]. The Thal_MDI_R region, involved in thalamocortical signalling, is critical for higher cognitive functions, and its disruption can contribute to impaired consciousness and cognitive engagement [79]. The involvement of Vermis_1_2 and Vermis_3, regions in the cerebellum, is significant due to their role in motor coordination and balance. Damage to these regions is often associated with

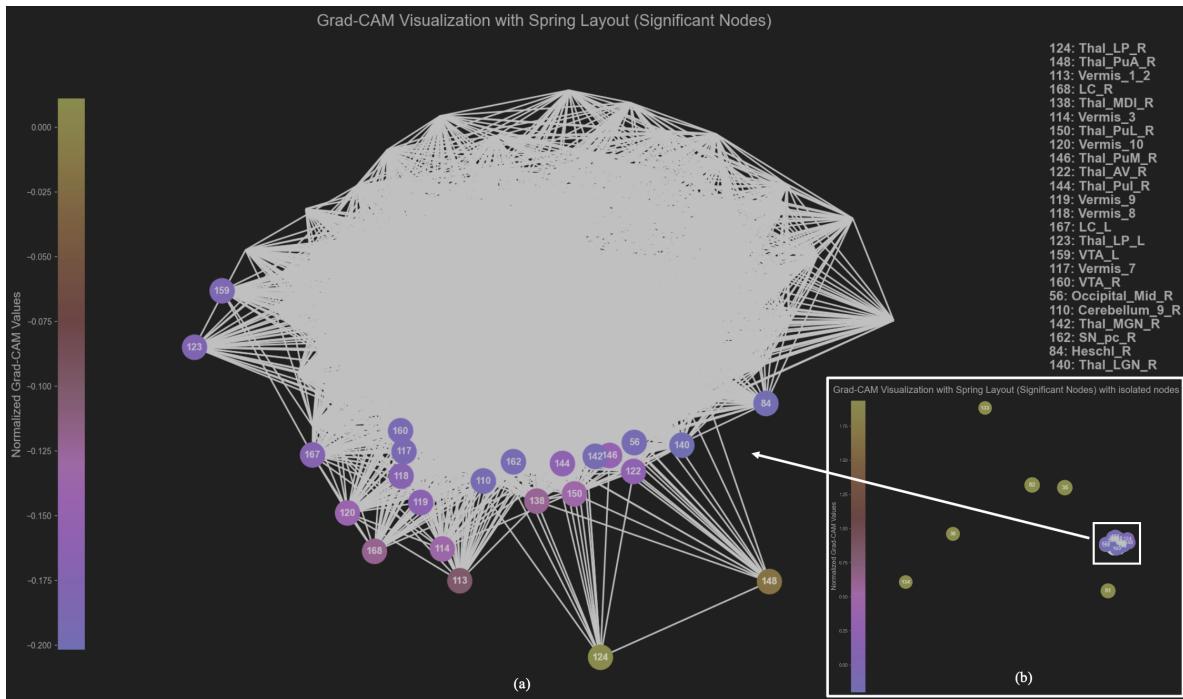


Figure 7: Panel (a) presents a focused view of the brain connectivity network, highlighting significant nodes identified using Grad-CAM, with isolated nodes excluded to emphasize the core functional interactions. Panel (b) includes the isolated nodes, marked by their high influence on network dynamics, illustrating their critical roles despite being anatomically uncharacterized.

motor impairments, such as uncoordinated movements and difficulties with posture, which are frequently seen in TBI patients. The cerebellum also plays a role in cognitive functions, and its dysfunction may contribute to cognitive deficits observed in these patients [80]. Additionally, LC_R, part of the locus coeruleus, and VTA_R, involved in the ventral tegmental area, exhibit high Grad-CAM values, indicating their importance in regulating arousal, autonomic functions, and the sleep-wake cycle. These structures are integral to maintaining wakefulness, and damage to these areas can result in decreased arousal levels or even prolonged states of unconsciousness, as commonly seen in severe TBI cases [74, 81]. The exclusion of isolated nodes from this visualization enables a more focused examination of the most significant interconnected regions, offering crucial insights into the disrupted neural pathways contributing to consciousness and motor impairments.

To visually represent our findings, we analyzed histograms depicting the distribution of node mentions across the three patient groups (acute, chronic, and controls), illustrating how specific brain regions vary in their mention frequency. These histograms show clear differentiation in the regions highlighted by each group, providing a deeper understanding of the neural patterns specific to each clinical condition. In Fig 8, we present the histogram for the acute group, showcasing the number of unique patients mentioning each brain region. The distribution shows that a significant number of regions, particularly in the frontal cortex, temporal pole, and cerebellum, are frequently mentioned across multiple patients, with several areas being highlighted

by more than five patients. This suggests a concentration of significant neural activity in these regions, potentially reflecting their involvement in acute phases of TBI. The variability in the number of mentions across different regions also indicates that certain areas are less commonly involved, possibly reflecting a more specialized or secondary role in these patients' neural network dynamics. Fig 9 presents the histogram for the control group, displaying a concentrated distribution with a higher number of mentions in a smaller set of regions, particularly within the frontal cortex. The peak, around 12 mentions, suggests that these regions are crucial for maintaining stable neural function in healthy individuals. In contrast, areas like the midbrain and temporal pole show fewer mentions, reflecting a more stable network with less variability across controls, which contrasts sharply with the widespread disruptions seen in TBI. Fig 10 illustrates the histogram for the chronic patient group, whose distribution sits between the control and acute groups. Although more widespread than in controls, the chronic group's peak values do not reach the same intensities seen in the acute group. This suggests that while regions like the frontal cortex remain affected in chronic patients, the degree of neural disruption has potentially diminished over time, likely due to compensatory mechanisms or neural adaptations. Additionally, the chronic group shows greater variability in regions such as the thalamus and ACC, indicating ongoing network reorganization that differentiates it from both the control and acute groups. This is also confirmed by the observation that included within the chronic group are both patients with DoC and patients who have

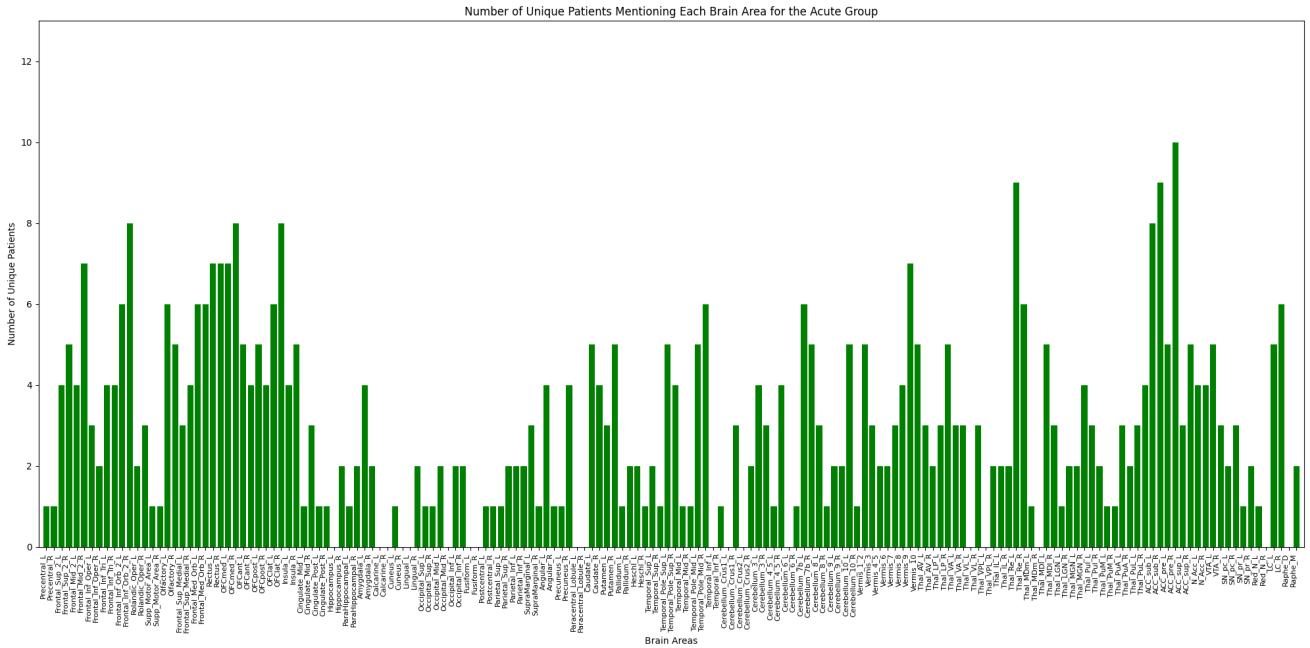


Figure 8: Histogram illustrating the distribution of significant brain regions across the acute patient group. The height of each bar represents the number of unique patients in whom each brain area was identified as significant.

recovered consciousness. By examining the nodes that Grad-CAM highlighted as significant, we aimed to identify brain regions most strongly associated with the conditions under investigation. After running the analysis across all patients, we compiled the results for each clinical group, focusing on nodes that appeared significant across multiple patients. To quantify this, we established a threshold: any node that was found to be significant in more than five patients within the same group was considered to have clinical relevance. This threshold was selected based on its ability to highlight nodes that played a consistent role in classification, striking a balance between capturing important regions and the small number of subjects.

As part of our comprehensive analysis, Fig 11 to 16 offer detailed DWI visualizations of the significant brain regions identified across different patient groups. These figures complement the respective tables, illustrating the spatial distribution of these regions and how they relate to the clinical features observed in each group. The nodes depicted in grey represent regions that, while not individually significant for that group, maintain connections with the most critical nodes. Conversely, the nodes illustrated in black do not exhibit any connections with these key regions. Each figure consistently utilizes this colour coding to depict connectivity, while additional colours in each figure highlight the most significant nodes for that specific group, emphasizing their pivotal role in the neural networks. Figure 11 corresponds to the acute patient group, represented in Table 2, where green nodes indicate the most significant regions. In acute TBI cases, disruptions in regions like the thalamus, frontal cortex, and brainstem are critical, given their roles in motor control, consciousness, and cognition. These alterations

Table 2
Grouping of significant brain regions in acute patients

Anatomical Area	Regions
Frontal Cortex	Frontal_Sup_2_R, Frontal_Mid_2_R, Frontal_Inf_Orb_2_L, Frontal_Inf_Orb_2_R, Frontal_Med_Orb_L, Frontal_Med_Orb_R, Rectus_L, Rectus_R, OFCmed_L, OFCmed_R, OFCant_L, OFCpost_L, OFClat_L, OFClat_R
Olfactory	Olfactory_L, Olfactory_R
Insula	Insula_R
Basal Ganglia	Caudate_L, Putamen_R
Temporal Pole	Temporal_Pole_Sup_L, Temporal_Pole_Mid_L, Temporal_Pole_Mid_R
Cerebellum	Cerebellum_7b_L, Cerebellum_7b_R, Cerebellum_10_L, Vermis_1_2, Vermis_9, Vermis_10
Thalamus	Thal_LP_R, Thal_Re_L, Thal_Re_R, Thal_MDI_L
Anterior Cingulate Cortex	ACC_sub_L, ACC_sub_R, ACC_pre_L, ACC_pre_R, ACC_sup_R
Midbrain	VTA_L, LC_L, LC_R

are central to the acute manifestations of TBI, reflecting widespread network disruption and offering deeper insights into the immediate neural impacts of injury. The frontal cortex, for example, is heavily involved in executive functions

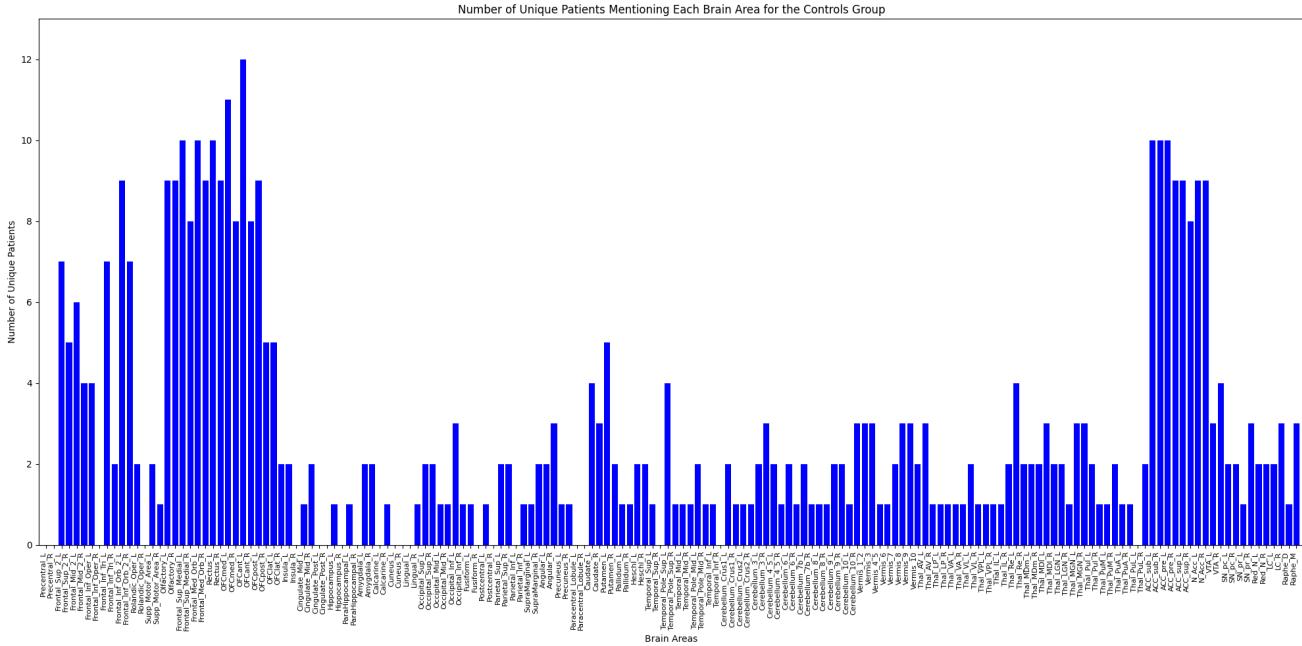


Figure 9: Histogram of brain region mentions in the controls group. The frontal cortex and cerebellum exhibit the highest concentration of mentions, suggesting these areas are most involved in maintaining normal brain function in healthy individuals.

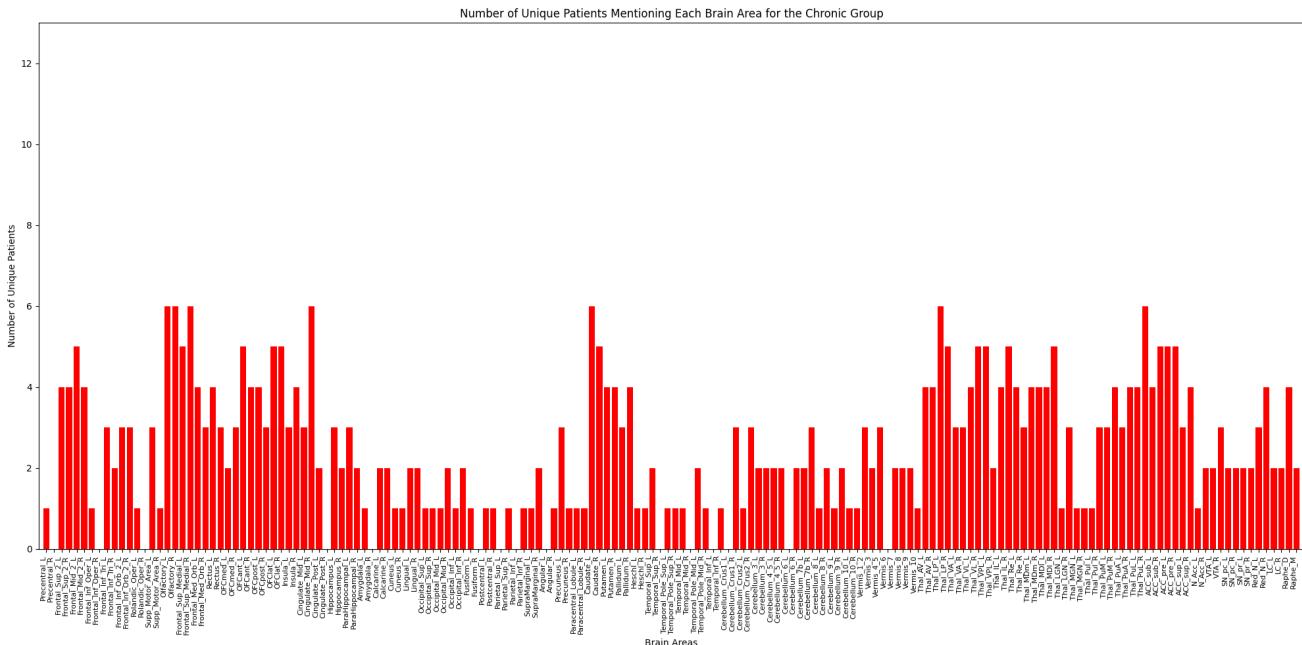


Figure 10: Histogram showing the distribution of significant brain regions in the chronic patient group. Although the overall distribution is wider than in the controls, the mentions are less concentrated than in the acute group, indicating a potential for network stabilization and adaptation over time.

and emotional regulation, while the basal ganglia and thalamus are crucial for motor control and consciousness regulation. The cerebellum, though traditionally associated with motor coordination, also plays a role in cognitive processing, further emphasizing the complexity of disruptions observed in acute TBI cases [82, 79]. In contrast, Fig 12, corresponding to Table 3, presents the control group's brain networks,

where blue nodes represent the most significant regions. These areas, primarily located within the frontal cortex and olfactory regions, reflect a stable neural network, crucial for maintaining normal cognitive, motor, and sensory functions. The high density of significant nodes in the frontal cortex underscores its role in higher-order cognitive functions, such as attention, decision-making, and emotional regulation. The

Table 3

Grouping of significant brain regions in control group

Anatomical Area	Regions
Frontal Cortex	Frontal_Sup_2_L, Frontal_Sup_2_R, Frontal_Mid_2_L, Frontal_Inf_Tri_L, Frontal_Inf_Orb_2_L, Frontal_Inf_Orb_2_R, Frontal_Sup_Medial_L, Frontal_Sup_Medial_R, Frontal_Med_Orb_L, Frontal_Med_Orb_R, Rectus_L, Rectus_R, OFCmed_L, OFCmed_R, OFCant_L, OFCant_R, OFCpost_L, OFCpost_R, OFClat_L
Olfactory	Olfactory_L, Olfactory_R
Basal Ganglia	Putamen_L
Anterior Cingulate Cortex	ACC_sub_L, ACC_sub_R, ACC_pre_L, ACC_pre_R, ACC_sup_L, ACC_sup_R
Nucleus Accumbens	N_Acc_L, N_Acc_R

Table 4

Grouping of significant brain regions in chronic patients

Anatomical Area	Regions
Frontal Cortex	Frontal_Mid_2_L, Frontal_Sup_Medial_L, Frontal_Sup_Medial_R
Olfactory	Olfactory_L, Olfactory_R
Orbitofrontal Cortex	OFCant_L, OFClat_L, OFClat_R
Cingulate Cortex	Cingulate_Mid_R
Basal Ganglia	Caudate_L, Caudate_R
Thalamus	Thal_LP_L, Thal_LP_R, Thal_VL_R, Thal_VPL_L, Thal_IL_R, Thal_MDI_R, Thal_PuM_R, Thal_PuL_R
Anterior Cingulate Cortex	ACC_sub_R, ACC_pre_R ACC_pre_L

visualization of this stable network contrasts sharply with the more disrupted and widespread network impairments seen in TBI patients. Additionally, fewer isolated regions and stronger cohesion in the control group emphasize the integrity of a healthy brain network, providing a baseline for comparison with both acute and chronic TBI patients. Fig 13 visualizes the significant brain regions identified in chronic patients, corresponding to Table 4, without subgroup distinction (including those who regained consciousness, as well as patients in MCS and VS). The red nodes highlight key areas that remain affected in the chronic phase of TBI, including the frontal cortex, thalamus, and brainstem. The chronic patient group presents a more concentrated pattern of significant regions, suggesting a stabilization or reorganization of neural networks over time. However, persistent disruptions in areas such as the ACC and thalamus continue to reflect ongoing functional impairments in consciousness

Table 5

Significant brain regions in chronic patients (Minimally Conscious State)

Anatomical Area	Regions
Frontal Cortex	Frontal_Sup_2_R, Frontal_Mid_2_R, Frontal_Sup_Medial_R
Olfactory	Olfactory_L, Olfactory_R
Orbitofrontal Cortex	OFClat_L
Insula	Insula_R
Cingulate Cortex	Cingulate_Mid_R
Basal Ganglia	Caudate_L, Caudate_R
Thalamus	Thal_AV_L, Thal_AV_R, Thal_LP_L, Thal_LP_R, Thal_VA_R, Thal_VL_R, Thal_IL_L, Thal_IL_R, Thal_Re_L, Thal_Re_R, Thal_MDM_R, Thal_MDI_L, Thal_MDI_R, Thal_LGN_R, Thal_PuM_R, Thal_PuA_R, Thal_PuL_R
Anterior Cingulate Cortex	ACC_sub_R, ACC_pre_R
Red Nucleus	Red_N_L, Red_N_R
Raphe	Raphe_D, Raphe_M

Table 6

Significant brain regions in chronic patients (Vegetative State)

Anatomical Area	Regions
Hippocampus	Hippocampus_L
Basal Ganglia	Caudate_L, Putamen_L, Pallidum_L
Thalamus	Thal_VL_L, Thal_MDM_L

Table 7

Grouping of significant brain regions in chronic patients who regained consciousness

Anatomical Area	Regions
Frontal Cortex	Frontal_Mid_2_L, Frontal_Sup_Medial_L, Frontal_Sup_Medial_R
Olfactory	Olfactory_L, Olfactory_R
Orbitofrontal Cortex	OFCant_L, OFClat_L, OFClat_R
Thalamus	Thal_VL_R, Thal_VPL_L, Thal_IL_R, Thal_PuL_R
Anterior Cingulate Cortex	ACC_sub_R, ACC_pre_L, ACC_pre_R

and cognitive processing [83]. The visualization highlights the long-term effects of TBI, where despite recovery efforts, essential networks involved in motor control, cognition, and sensory integration remain affected. In the VS subgroup, Fig 14 corresponds to Table 6, where yellow nodes mark the most significant brain regions. The visualization reveals a reduced set of significant areas, particularly concentrated in the hippocampus, basal ganglia, and thalamus, underscoring severe deficits in memory processing, motor coordination,

and sensory integration. The limited number of significant nodes reflects the widespread neural damage characteristic of VS patients, where the capacity for basic motor and sensory functions is heavily compromised. This aligns with clinical observations of the profound consciousness deficits in patients in a vegetative state. For the MCS subgroup, Fig 15 (pink nodes) represents the significant regions listed in Table 5. MCS patients exhibit a broader network engagement compared to VS patients, with key regions such as the thalamic nuclei, insula, and red nucleus showing notable involvement. The engagement of these regions suggests that despite severe impairments, MCS patients retain minimal awareness, which is consistent with their limited cognitive and sensory functions [84]. The broader engagement of thalamic nuclei reflects the partial retention of connectivity in critical regions associated with sensory integration and awareness, providing insights into the clinical presentation of MCS patients [85]. Lastly, Fig 16 visualizes the significant brain regions for chronic patients who have regained consciousness after TBI, corresponding to Table 7, with orange nodes highlighting key regions. The visualization demonstrates the ongoing recovery of neural function, with significant areas concentrated in the frontal cortex, thalamus, and ACC. These regions are vital for executive function, sensory processing, and attention regulation, and their re-engagement suggests the brain's capacity for plasticity and recovery, though residual challenges persist [86]. The ability of these patients to regain some level of consciousness reflects a partial restoration of network integrity, particularly in regions governing higher-order cognition and awareness. In all these visualizations, the calculation of the most significant nodes varies specifically for the VS, MCS, and conscious subgroups, reflecting their distinct patient characteristics. For these subgroups, regions with mentions in more than two patients per node for VS and MCS patients, and more than three patients for those who regained consciousness, were considered significant. However, for the other groups, the threshold remained as regions mentioned in more than five patients, ensuring consistency in identifying key nodes across larger patient populations. This adjusted threshold allowed for a more nuanced analysis, particularly given the smaller subgroup sizes (4 MCS patients and 2 VS patients), compared to the more stringent threshold of five patients used in Table 7. The comparison of these visualizations reveals the evolution of neural network disruptions from acute to chronic stages, with acute patients showing widespread network impairment, and chronic patients displaying more localized but persistent disruptions. Meanwhile, VS patients exhibit severe and widespread damage, while MCS patients show a broader but less severe pattern of engagement, reflecting minimal awareness. For patients who regained consciousness, the network appears to re-engage, highlighting the ongoing restoration of brain function, albeit incomplete.

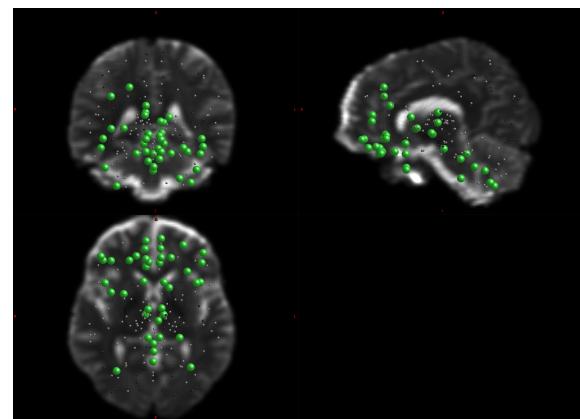


Figure 11: DWI visualization of brain networks in acute patients. The green nodes, as outlined in Table 2, represent the most significant areas, showing their spatial distribution across axial, sagittal, and coronal views. Key regions such as the thalamus, frontal cortex, and brainstem, which are involved in motor control, consciousness, and cognition, are disrupted in acute TBI.

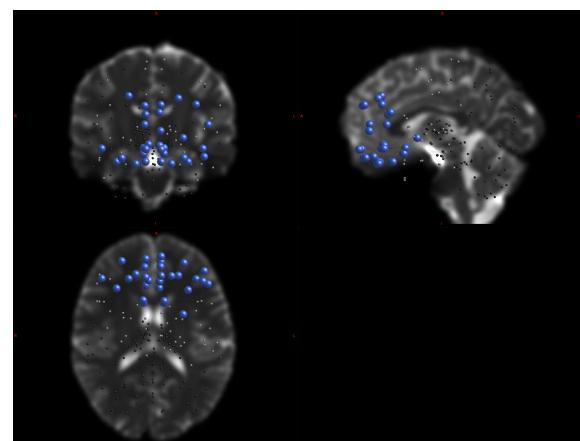


Figure 12: DWI visualization of significant brain regions in the control group. The blue nodes highlight the most important areas, with connections shown in grey. These regions, primarily in the frontal cortex and olfactory areas, reflect stable neural functions in healthy individuals, contrasting with the disruptions seen in TBI patients, as seen in Table 3.

5. Discussion

The results of the GCN model, combined with the insights derived from the Grad-CAM analysis, provide a comprehensive understanding of the neural mechanisms underlying TBI. The model's high classification accuracy, alongside the identification of significant brain regions, not only demonstrates the potential of machine learning for distinguishing between different patient groups but also offers meaningful clinical insights into the pathophysiology of these disorders. The earliest insight at the basis of this study was that chronic subjects experience some degree of neurodegeneration at a distant time from the trauma causing, and that it should be possible to quantify this for patient stratification. Indeed, the model's average accuracy

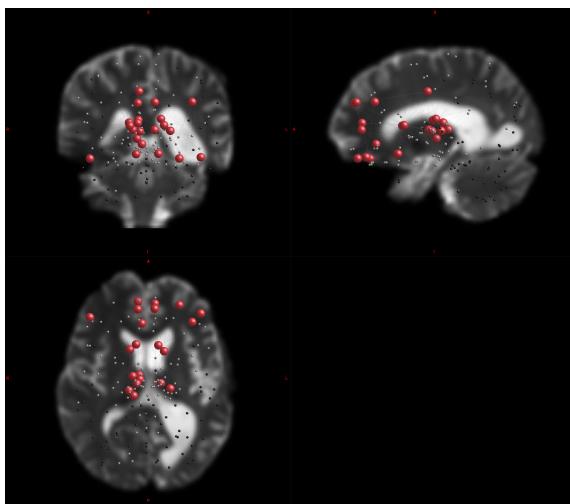


Figure 13: DWI visualization of the most significant brain regions identified in chronic patients without subgroup distinction. The red nodes represent the key regions involved across all chronic patients, highlighting areas primarily in the frontal cortex, thalamus, and brainstem, suggesting persistent disruptions in regions critical for motor control, cognition, and consciousness regulation, as indicated in Table 4.

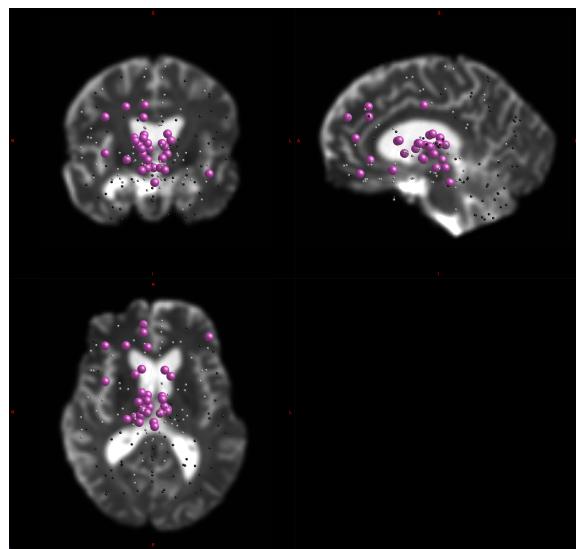


Figure 15: DWI visualization of significant brain regions in patients in a minimally conscious state (MCS). The pink nodes highlight thalamic nuclei, the insula, and brainstem regions, which are more extensively engaged compared to the VS group, reflecting the preserved but minimal awareness in MCS patients, as listed in Table 5.

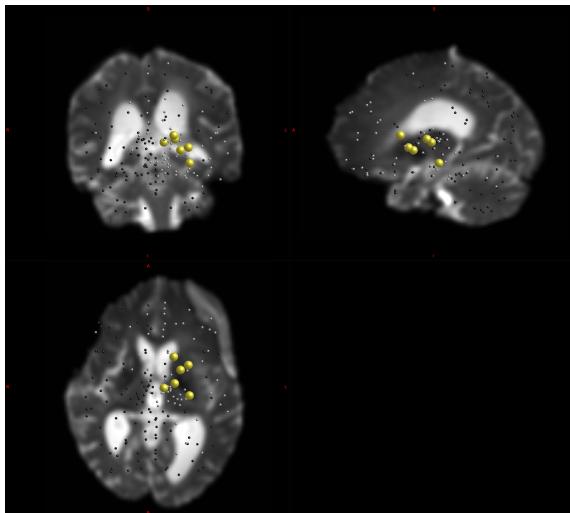


Figure 14: DWI visualization of significant brain regions in patients in a vegetative state (VS) as shown in Table 6. The yellow nodes represent the primary areas involved in sensory integration, motor coordination, and memory processing, with a strong focus on the thalamus, basal ganglia, and hippocampus, which are severely impaired in VS patients.

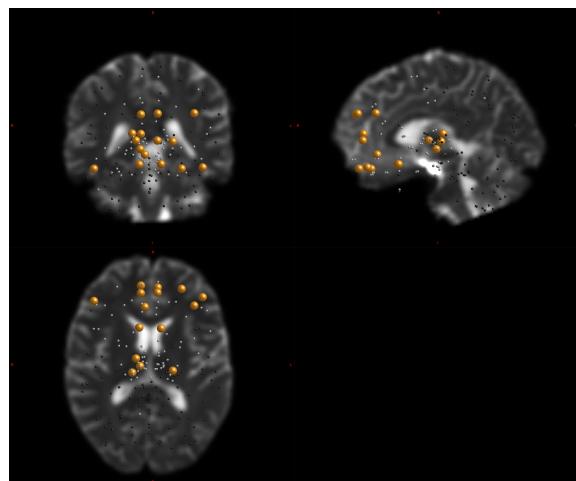


Figure 16: DWI visualization of significant brain regions in chronic patients who regained consciousness after TBI. The orange nodes show areas concentrated in the frontal cortex, thalamus, and anterior cingulate cortex, regions integral to executive functions, sensory processing, and attention regulation, indicating ongoing recovery of neural function in these patients, as outlined in Table 7.

of 83.67%, along with strong precision and recall scores, indicates a reliable ability to classify patients based on brain connectivity patterns. This has important implications for clinical practice, where accurate diagnosis and stratification of patients with TBI are critical for determining treatment pathways. The precision of 81.6% is noteworthy, especially in a clinical context where distinguishing accurately among the three patient classes is crucial. This level of precision indicates that the model effectively minimizes the number

of false positives, thus ensuring that individuals across different groups are classified with high reliability. Similarly, the recall rate of 78% suggests that the model successfully identifies true positive cases, a key factor in ensuring that patients with impaired consciousness are correctly diagnosed and managed. However, the model performance variability underscores TBI and DoC's complexity and heterogeneity. Individual cases with incorrect predictions likely represent atypical patient profiles, emphasizing the need for further

refinement of the model to account for outlier data. In clinical settings, this variability may correspond to patients with uncommon patterns of brain damage or those at the extremes of consciousness impairment, highlighting the importance of personalized treatment strategies based on specific neural profiles.

The Grad-CAM analysis enhances the interpretability of these classification results by highlighting the most influential brain regions contributing to the model's decisions. In acute TBI patients, key regions such as the thalamus, brainstem, and frontal cortex were identified as critical nodes within the brain's connectivity network. These regions play fundamental roles in motor control, consciousness, and cognitive regulation, which are often disrupted in TBI. For example, the thalamus is a central hub for sensory integration and consciousness regulation, and its involvement in acute TBI reflects the widespread sensory and cognitive deficits seen in these patients [87]. The frontal cortex's participation, particularly in executive functions and emotional regulation, also aligns with the cognitive and behavioural changes commonly observed in these patients. The inclusion of nodes such as the locus coeruleus (LC) and the ventral tegmental area (VTA) highlights the importance of brainstem structures in maintaining wakefulness and arousal, functions that are critically impaired in TBI [88]. Disruptions in these regions often correlate with prolonged states of unconsciousness or reduced arousal, emphasizing their significance in the pathophysiology of impaired consciousness. These findings suggest that targeting these regions through interventions like neurostimulation could help restore arousal levels in these patients. In chronic TBI patients, the Grad-CAM analysis revealed a more localized pattern of disruption, with key regions such as the ACC and thalamus continuing to show persistent impairment. This shift from the widespread disruptions seen in acute TBI to more focused impairments in chronic stages likely reflects the brain's attempts to reorganize and compensate for the damage over time. However, the persistence of these disruptions indicates ongoing challenges in higher-order cognitive functions, attention regulation, and sensory processing. Clinically, this highlights the potential for long-term interventions to support neural plasticity and promote recovery in these critical regions. The comparison between MCS and VS patients offers further insights into the underlying neural mechanisms of consciousness. MCS patients exhibited broader engagement of thalamic nuclei and insular regions, suggesting that, despite severe impairments, these patients retain minimal levels of awareness. This broader engagement reflects a greater capacity for connectivity in critical regions associated with sensory integration and cognitive processing. In contrast, the more limited set of significant regions identified in VS patients, particularly in the hippocampus, basal ganglia, and thalamus, underscores the profound neural damage in this group. These findings align with clinical observations of VS patients, where the capacity for basic motor and sensory functions is heavily compromised, and the potential for recovery is limited [29]. The identification of significant brain regions

in patients who regained consciousness after TBI further underscores the brain's plasticity. Key areas, including the frontal cortex, thalamus, and ACC, were identified as critical for supporting recovery. These regions are essential for executive functions, sensory processing, and attention regulation, suggesting that their re-engagement is a marker of neural recovery. From a clinical perspective, this offers hope for targeted rehabilitation strategies to enhance these networks' connectivity to support continued recovery in chronic TBI patients. Insights from the control group provide a basis for comparison with the patient groups. The significant brain regions identified in the controls, particularly in the frontal cortex, olfactory regions and ACC, represent typical patterns of neural connectivity associated with normal cognitive, motor and sensory functions. The high density of significant nodes in the frontal cortex highlights its critical role in executive functions, attention and emotional regulation. In contrast to the dysregulated and diffuse neurons observed in patients with acute and chronic TBI, the control group has fewer isolated regions and a more cohesive network [89]. This stability suggests that the brain networks of healthy individuals maintain a high degree of connectivity and organisation, in stark contrast to the chaotic neural patterns observed in patients with TBI and DoC. Clinically, this underlines the extent of neural disintegration in patients with TBI, where key regions such as the frontal cortex and thalamus show significant disruptions that impair higher-order cognitive functions.

Overall, the integration of classification performance and Grad-CAM not only provides a reliable tool for distinguishing patient groups, but also deepens our understanding of the neural networks affected by this injury. Identifying key brain regions offers a foundation for developing targeted interventions, such as neurostimulation or rehabilitation strategies, aimed at enhancing recovery in patients with impaired consciousness. This research also highlights the potential to predict patient trajectories and outcomes at different stages of TBI recovery, providing valuable insights into how the injury evolves over time. By analyzing brain connectivity patterns using machine learning models, clinicians may be able to identify key turning points in a patient's recovery process, such as transitions from acute to chronic phases of injury. This could enable more personalized treatment plans, tailored to the specific stage of recovery, and allow for earlier interventions to prevent long-term complications. Furthermore, predictive models could help in monitoring the likelihood of certain outcomes, such as cognitive or motor impairments, thus guiding both short-term clinical decisions and long-term rehabilitation strategies. Furthermore, the model's ability to highlight variability across patients suggests that personalized approaches to treatment will be critical for optimizing outcomes in this highly heterogeneous patient population. Future work will need to explore how these findings can be translated into clinical practice, with a focus on improving model robustness and understanding the long-term implications of neural network disruptions in TBI and DoC.

6. Conclusion and future works

This study demonstrates the potential of GCNs combined with Grad-CAM for classifying patients with TBI with special attention to patients with DoC, offering insights into the neural mechanisms underlying these conditions. The high classification accuracy achieved by the model, coupled with its ability to identify significant brain regions, underscores its relevance for clinical applications, particularly for diagnosis and patient stratification. By focusing on structural connectivity, this approach reveals how disruptions in key areas, such as the thalamus, frontal cortex, and brainstem, contribute to impaired consciousness and cognitive deficits in TBI and DoC patients. However, several limitations must be acknowledged. One of the primary constraints of this study is the relatively small dataset used, which limits the generalizability of the findings. Given the variability in brain injury presentations, a larger and more diverse dataset would enhance the robustness of the model and allow for more comprehensive validation. Future studies should aim to incorporate additional datasets from different clinical sources, enabling the model to account for a wider range of patient profiles and increasing its applicability in various clinical settings. Another potential avenue for future research involves conducting longitudinal studies. While this work provides insights into patients in both acute and chronic phases, the availability of follow-up imaging for some patients who regained consciousness after five months highlights the value of tracking neural changes over time. A longitudinal analysis could provide deeper insights into the mechanisms of recovery and the role of neuroplasticity in TBI, offering opportunities to refine predictive models for long-term outcomes and personalize rehabilitation strategies. Such studies could explore how brain connectivity patterns evolve as patients transition from acute to chronic stages of injury, and potentially identify markers of recovery. Additionally, while the current analysis primarily focuses on significant brain nodes identified, a future line of research could involve analyzing the most influential edges between these nodes. Understanding the connections or pathways between significant brain regions may offer a more detailed view of how neural communication is altered in TBI and DoC patients. By investigating these edges, it would be possible to gain insights into how disruptions in specific connections contribute to clinical symptoms, potentially identifying key pathways for targeted interventions. Furthermore, while Grad-CAM was effective in identifying the most influential brain regions in model predictions, future research could explore the integration of additional XAI methods. Techniques such as Layer-wise Relevance Propagation or SHAP values might complement it by providing more detailed interpretations of model behaviour at both the node and network level, offering a richer understanding of the specific pathways and interactions involved in consciousness and cognition.

In conclusion, while this work presents promising advances in the classification of TBI and DoC patients, addressing the limitations related to dataset size, conducting

longitudinal studies, and integrating additional data types and XAI techniques will be crucial for future research. These steps will help refine the model, enhance its clinical applicability, and contribute to a deeper understanding of the neural dynamics involved in brain injury and recovery.

CRediT authorship contribution statement

Tiziana Currieri: Writing – original draft, Visualization, Methodology, Formal analysis, Conceptualization, Investigation, Software. **Joan Falcó-Roget:** Software, Validation, Writing – review and editing, Investigation. **Elham Rostami:** Supervision, Validation, Writing – review and editing. **Salvatore Vitabile:** Validation, Investigation, Resources, Writing – review and editing, Supervision, Project administration. **Alessandro Crimi:** Funding acquisition, Supervision, Writing – review and editing, Validation, Conceptualization, Project administration.

Acknowledgments

This contribution is partially supported by the Ministry of University and Research (MUR) as part of the FSE REACT-EU - PON 2014-2020 "Research and Innovation" resources – Green/Innovation Action - DM MUR 1061/2021. The publication was created within the project of the Minister of Science and Higher Education "Support for the activity of Centers of Excellence established in Poland under Horizon 2020" on the basis of the contract number MEiN/2023/DIR/3796, and has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857533. This publication is supported by Sano project carried out within the International Research Agendas programme of the Foundation for Polish Science, co-financed by the European Union under the European Regional Development Fund.

References

- [1] M. C. Dewan, A. Rattani, S. Gupta, R. E. Baticulon, Y.-C. Hung, M. Punchak, A. Agrawal, A. O. Adeleye, M. G. Shrime, A. M. Rubiano, et al., Estimating the global incidence of traumatic brain injury, *Journal of neurosurgery* 130 (4) (2018) 1080–1097.
- [2] A. I. Maas, D. K. Menon, G. T. Manley, M. Abrams, C. Åkerlund, N. Andelic, M. Aries, T. Bashford, M. J. Bell, Y. G. Bodien, et al., Traumatic brain injury: progress and challenges in prevention, clinical care, and research, *The Lancet Neurology* 21 (11) (2022) 1004–1060.
- [3] S. O'Brien, K. Metcalf, J. Batchelor, An examination of the heterogeneity of cognitive outcome following severe to extremely severe traumatic brain injury, *The Clinical Neuropsychologist* 34 (1) (2020) 120–139.
- [4] A. L. Lee, Advanced imaging of traumatic brain injury, *Korean Journal of Neurotrauma* 16 (1) (2020) 3.
- [5] G. Ilie, A. Boak, E. M. Adlaf, M. Asbridge, M. D. Cusimano, Prevalence and correlates of traumatic brain injuries among adolescents, *Jama* 309 (24) (2013) 2550–2552.
- [6] K. K. Nagy, K. T. Joseph, S. M. Krosner, R. R. Roberts, C. L. Leslie, K. Dufty, R. F. Smith, J. Barrett, The utility of head computed tomography after minimal head injury, *Journal of Trauma and Acute Care Surgery* 46 (2) (1999) 268–270.
- [7] J. S. Jeret, M. Mandell, B. Anziska, M. Lipitz, A. P. Vilceus, J. A. Ware, T. A. Zesiewicz, Clinical predictors of abnormality disclosed

- by computed tomography after mild head trauma, *Neurosurgery* 32 (1) (1993) 9–16.
- [8] E. C. Miller, J. F. Holmes, R. W. Derlet, Utilizing clinical factors to reduce head ct scan ordering for minor head trauma patients, *The Journal of emergency medicine* 15 (4) (1997) 453–457.
- [9] J. E. SCHUNK, J. D. RODGERSON, G. A. WOODWARD, The utility of head computed tomographic scanning in pediatric patients with normal neurologic examination in the emergency department, *Pediatric emergency care* 12 (3) (1996) 160–165.
- [10] B. R. Duus, B. Lind, H. Christensen, O. A. Nielsen, The role of neuroimaging in the initial management of patients with minor head injury, *Annals of emergency medicine* 23 (6) (1994) 1279–1283.
- [11] K. S. Quayle, D. M. Jaffe, N. Kuppermann, B. A. Kaufman, B. C. Lee, T. Park, W. H. McAlister, Diagnostic testing for acute head injury in children: when are head computed tomography and skull radiographs indicated?, *Pediatrics* 99 (5) (1997) e11–e11.
- [12] H. Zhang, K. Ogasawara, Grad-cam-based explainable artificial intelligence related to medical text processing, *Bioengineering* 10 (9) (2023) 1070.
- [13] S. Han, Z. Sun, K. Zhao, F. Duan, C. F. Caiafa, Y. Zhang, J. Solé-Casals, Early prediction of dementia using fmri data with a graph convolutional network approach, *Journal of Neural Engineering* 21 (1) (2024) 016013.
- [14] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, G. Monfardini, The graph neural network model, *IEEE transactions on neural networks* 20 (1) (2008) 61–80.
- [15] S. Mazurek, R. Blanco, J. Falcó-Roget, A. Crimi, Explainable graph neural networks for eeg classification and seizure detection in epileptic patients, in: 2024 IEEE International Symposium on Biomedical Imaging (ISBI), IEEE, 2024.
- [16] D. Coluzzi, V. Bordin, M. W. Rivolta, I. Fortel, L. Zhang, A. Leow, G. Baselli, Biomarker investigation using multiple brain measures from mri through xai in alzheimer's disease classification, *arXiv preprint arXiv:2305.03056* (2023).
- [17] X. Li, N. C. Dvornek, Y. Zhou, J. Zhuang, P. Ventola, J. S. Duncan, Graph neural network for interpreting task-fMRI biomarkers, in: Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part V 22, Springer, 2019, pp. 485–493.
- [18] U. Saha, I. U. Ahmed, I. U. Ahmed, A.-A. Hossain, Graph convolutional network-based approach for parkinson's disease classification using euclidean distance graphs, 2024 7th International Conference on Informatics and Computational Sciences (ICICoS) (2024) 532–537.
URL <https://api.semanticscholar.org/CorpusID:271935661>
- [19] E. Chen, B. Barile, F. Durand-Dubief, T. Grenier, D. Sappéy-Marinier, Multiple sclerosis clinical forms classification with graph convolutional networks based on brain morphological connectivity, *Frontiers in Neuroscience* 17 (2024) 1268860.
- [20] F. Prinzi, T. Currieri, S. Gaglio, S. Vitabile, Shallow and deep learning classifiers in medical image analysis, *European Radiology Experimental* 8 (1) (2024) 26.
- [21] H. Zhou, L. He, B. Y. Chen, L. Shen, Y. Zhang, Multi-modal diagnosis of alzheimer's disease using interpretable graph convolutional networks, *IEEE Transactions on Medical Imaging* (2024) 1–1 doi: 10.1109/TMI.2024.3432531.
- [22] P. E. Pope, S. Kolouri, M. Rostami, C. E. Martin, H. Hoffmann, Explainability methods for graph convolutional neural networks, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 10772–10781.
- [23] F. Ferri, M. Cannariato, L. Pallante, E. A. Zizzi, M. A. Deriu, Explainable machine learning and deep learning models for predicting tas2r-bitter molecule interactions, *arXiv preprint arXiv:2406.15039* (2024).
- [24] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, M. Sun, Graph neural networks: A review of methods and applications, *AI open* 1 (2020) 57–81.
- [25] A. Holzinger, G. Langs, H. Denk, K. Zatloukal, H. Müller, Causability and explainability of artificial intelligence in medicine, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 9 (4) (2019) e1312.
- [26] S. B. Snider, Y. G. Bodien, A. Frau-Pascual, M. Bianciardi, A. S. Foulkes, B. L. Edlow, Ascending arousal network connectivity during recovery from traumatic coma, *NeuroImage: Clinical* 28 (2020) 102503.
- [27] S. B. Snider, Y. G. Bodien, M. Bianciardi, E. N. Brown, O. Wu, B. L. Edlow, Disruption of the ascending arousal network in acute traumatic disorders of consciousness, *Neurology* 93 (13) (2019) e1281–e1287.
- [28] D. Kondziella, A. Bender, K. Diserens, W. van Erp, A. Estraneo, R. Formisano, S. Laureys, L. Naccache, S. Ozturk, B. Rohaut, et al., European academy of neurology guideline on the diagnosis of coma and other disorders of consciousness, *European journal of neurology* 27 (5) (2020) 741–756.
- [29] A. Magliacano, P. Liuzzi, R. Formisano, A. Grippo, E. Angelakis, A. Thibaut, O. Gosseries, G. Lamberti, E. Noé, S. Bagnato, et al., Predicting long-term recovery of consciousness in prolonged disorders of consciousness based on coma recovery scale-revised subscores: validation of a machine learning-based prognostic index, *Brain Sciences* 13 (1) (2022) 51.
- [30] E. Landsness, M.-A. Bruno, Q. Noirhomme, B. Riedner, O. Gosseries, C. Schnakers, M. Massimini, S. Laureys, G. Tononi, M. Boly, Electrophysiological correlates of behavioural changes in vigilance in vegetative state and minimally conscious state, *Brain* 134 (8) (2011) 2222–2232.
- [31] M.-A. Bruno, A. Vanhaudenhuyse, A. Thibaut, G. Moonen, S. Laureys, From unresponsive wakefulness to minimally conscious plus and functional locked-in syndromes: recent advances in our understanding of disorders of consciousness, *Journal of neurology* 258 (2011) 1373–1384.
- [32] S. M. Smith, M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Dronjak, D. E. Flitney, et al., Advances in functional and structural mr image analysis and implementation as fsl, *NeuroImage* 23 (2004) S208–S219.
- [33] J.-D. Tournier, R. Smith, D. Raffelt, R. Tabbara, T. Dhollander, M. Pietsch, D. Christiaens, B. Jeurissen, C.-H. Yeh, A. Connolly, Mrtrix3: A fast, flexible and open software framework for medical image processing and visualisation, *NeuroImage* 202 (2019) 116137.
- [34] E. Garyfallidis, M. Brett, B. Amirbekian, A. Rokem, S. Van Der Walt, M. Descoteaux, I. Nimmo-Smith, D. Contributors, Dipy, a library for the analysis of diffusion mri data, *Frontiers in neuroinformatics* 8 (2014) 8.
- [35] M. Jenkinson, Bet2: Mr-based estimation of brain, skull and scalp surfaces, in: Eleventh Annual Meeting of the Organization for Human Brain Mapping, 2005, 2005.
- [36] J. L. Andersson, S. Skare, J. Ashburner, How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging, *NeuroImage* 20 (2) (2003) 870–888.
- [37] J. Veraart, E. Fieremans, D. S. Novikov, Diffusion mri noise mapping using random matrix theory, *Magnetic resonance in medicine* 76 (5) (2016) 1582–1593.
- [38] E. Kellner, B. Dhital, V. G. Kiselev, M. Reisert, Gibbs-ringing artifact removal based on local subvoxel-shifts, *Magnetic resonance in medicine* 76 (5) (2016) 1574–1581.
- [39] J. L. Andersson, S. N. Sotiroopoulos, An integrated approach to correction for off-resonance effects and subject movement in diffusion mr imaging, *NeuroImage* 125 (2016) 1063–1078.
- [40] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, J. C. Gee, N4itk: improved n3 bias correction, *IEEE transactions on medical imaging* 29 (6) (2010) 1310–1320.
- [41] J. Falcó-Roget, A. Cacciola, F. Sambataro, A. Crimi, Functional and structural reorganization in brain tumors: a machine learning approach using desynchronized functional oscillations, *Communications Biology* 7 (1) (2024) 419.

- [42] E. T. Rolls, C.-C. Huang, C.-P. Lin, J. Feng, M. Joliot, Automated anatomical labelling atlas 3, *NeuroImage* 206 (2020) 116189. doi:<https://doi.org/10.1016/j.neuroimage.2019.116189>. URL <https://www.sciencedirect.com/science/article/pii/S1053811919307803>
- [43] A. Leemans, D. K. Jones, The b-matrix must be rotated when correcting for subject motion in dti data, *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* 61 (6) (2009) 1336–1349.
- [44] T. Dhollander, D. Raffelt, A. Connelly, Unsupervised 3-tissue response function estimation from single-shell or multi-shell diffusion mr data without a co-registered t1 image, in: ISMRM workshop on breaking the barriers of diffusion MRI, Vol. 5, Lisbon, Portugal, 2016.
- [45] T. Dhollander, R. Mito, A. Connelly, 3-tissue compositional data analysis of developing hcp (dhcp) diffusion mri data, *Hum. Brain Mapp* 25 (2019) T498.
- [46] R. E. Smith, J.-D. Tournier, F. Calamante, A. Connelly, Anatomically-constrained tractography: improved diffusion mri streamlines tractography through effective use of anatomical information, *Neuroimage* 62 (3) (2012) 1924–1938.
- [47] R. E. Smith, J.-D. Tournier, F. Calamante, A. Connelly, Sift2: Enabling dense quantitative assessment of brain white matter connectivity using streamlines tractography, *Neuroimage* 119 (2015) 338–351.
- [48] J. A. Bondy, U. S. R. Murty, Graph theory, Springer Publishing Company, Incorporated, 2008.
- [49] E. Bullmore, O. Sporns, Complex brain networks: graph theoretical analysis of structural and functional systems, *Nature reviews neuroscience* 10 (3) (2009) 186–198.
- [50] O. Sporns, Structure and function of complex brain networks, *Dialogues in clinical neuroscience* 15 (3) (2013) 247–262.
- [51] A. Crimi, L. Giancardo, F. Sambataro, A. Gozzi, V. Murino, D. Sona, Multilink analysis: Brain network comparison via sparse connectivity analysis, *Scientific reports* 9 (1) (2019) 65.
- [52] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, P. Vandergheynst, Geometric deep learning: going beyond euclidean data, *IEEE Signal Processing Magazine* 34 (4) (2017) 18–42.
- [53] T. N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, *arXiv preprint arXiv:1609.02907* (2016).
- [54] S. Parisot, S. I. Ktena, E. Ferrante, M. Lee, R. Guerrero, B. Glocker, D. Rueckert, Disease prediction using graph convolutional networks: application to autism spectrum disorder and alzheimer's disease, *Medical image analysis* 48 (2018) 117–130.
- [55] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad-cam: Visual explanations from deep networks via gradient-based localization, in: Proceedings of the IEEE international conference on computer vision, 2017, pp. 618–626.
- [56] E. Dührkoop, L. Malihi, C. Erfurt-Berge, G. Heidemann, M. Przy sucha, D. Busch, U. Hübner, Automatic classification of wound images showing healing complications: Towards an optimised approach for detecting maceration., *Studies in health technology and informatics* 317 (2024) 347–355.
- URL <https://api.semanticscholar.org/CorpusID:272398996>
- [57] A. Chattopadhyay, A. Sarkar, P. Howlader, V. N. Balasubramanian, Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks, in: 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), 2018, pp. 839–847. doi: [10.1109/WACV.2018.00097](https://doi.org/10.1109/WACV.2018.00097).
- [58] H. Zhou, L. He, Y. Zhang, L. Shen, B. Chen, Interpretable graph convolutional network of multi-modality brain imaging for alzheimer's disease diagnosis, *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)* (2022) 1–5.
- URL <https://api.semanticscholar.org/CorpusID:248407604>
- [59] Y. Zhang, Y. Weng, J. Lund, Applications of explainable artificial intelligence in diagnosis and surgery, *Diagnostics* 12 (2) (2022) 237.
- [60] A. Holzinger, C. Biemann, C. S. Pattichis, D. B. Kell, What do we need to build explainable ai systems for the medical domain?, *arXiv preprint arXiv:1712.09923* (2017).
- [61] Y. Zhang, L. Xue, S. Zhang, J. Yang, Q. Zhang, M. Wang, L. Wang, M. Zhang, J. Jiang, Y. Li, et al., A novel spatiotemporal graph convolutional network framework for functional connectivity biomarkers identification of alzheimer's disease, *Alzheimer's research & therapy* 16 (1) (2024) 60.
- [62] M. Rubinov, O. Sporns, Complex network measures of brain connectivity: Uses and interpretations, *NeuroImage* 52 (3) (2010) 1059–1069, computational Models of the Brain. doi:<https://doi.org/10.1016/j.neuroimage.2009.10.003>. URL <https://www.sciencedirect.com/science/article/pii/S105381190901074X>
- [63] M. Newman, Networks: An Introduction, Oxford University Press, 2010. doi:[10.1093/acprof:oso/9780199206650.001.0001](https://doi.org/10.1093/acprof:oso/9780199206650.001.0001). URL <https://doi.org/10.1093/acprof:oso/9780199206650.001.0001>
- [64] O. Tanglay, I. M. Young, N. B. Dadario, H. Taylor, P. Nicholas, S. Doyen, M. E. Sughrue, Eigenvector pagerank difference as a measure to reveal topological characteristics of the brain connectome for neurosurgery, *Journal of Neuro-Oncology* 157 (2022) 49–61.
- URL <https://api.semanticscholar.org/CorpusID:246492842>
- [65] L. Freeman, A set of measures of centrality based on betweenness, *Sociometry* (1977).
- [66] V. Latora, M. Marchiori, Efficient behavior of small-world networks, *Physical review letters* 87 (19) (2001) 198701.
- [67] D. Chen, H. Su, Identification of influential nodes in complex networks with degree and average neighbor degree, *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* 13 (2023) 734–742.
- URL <https://api.semanticscholar.org/CorpusID:259825335>
- [68] A. Barrat, M. Barthelemy, R. Pastor-Satorras, A. Vespignani, The architecture of complex weighted networks, *Proceedings of the national academy of sciences* 101 (11) (2004) 3747–3752.
- [69] P. Bonacich, Power and centrality: A family of measures, *American Journal of Sociology* 92 (1987) 1170–1182.
- URL <https://api.semanticscholar.org/CorpusID:145392072>
- [70] L. Katz, A new status index derived from sociometric analysis, *Psychometrika* 18 (1) (1953) 39–43.
- [71] D. S. Bassett, E. T. Bullmore, Human brain networks in health and disease, *Current opinion in neurology* 22 (4) (2009) 340–347.
- [72] M. P. Van Den Heuvel, O. Sporns, Rich-club organization of the human connectome, *Journal of Neuroscience* 31 (44) (2011) 15775–15786.
- [73] R. P. Vertes, S. B. Linley, Efferent and afferent connections of the dorsal and median raphe nuclei in the rat, in: Serotonin and sleep: molecular, functional and clinical aspects, Springer, 2008, pp. 69–102.
- [74] S. J. Sara, S. Bouret, Orienting and reorienting: the locus coeruleus mediates cognition through arousal, *Neuron* 76 (1) (2012) 130–141.
- [75] B. L. Jacobs, E. C. Azmitia, Structure and function of the brain serotonin system, *Physiological reviews* 72 (1) (1992) 165–229.
- [76] C. Varela, S. Kumar, J. Yang, M. A. Wilson, Anatomical substrates for direct interactions between hippocampus, medial prefrontal cortex, and the thalamic nucleus reuniens, *Brain Structure and Function* 219 (2014) 911–929.
- [77] S. M. Sherman, Thalamus plays a central role in ongoing cortical functioning, *Nature neuroscience* 19 (4) (2016) 533–541.
- [78] Y. B. Saalmann, M. A. Pinsk, L. Wang, X. Li, S. Kastner, The pulvinar regulates information transmission between cortical areas based on attention demands, *science* 337 (6095) (2012) 753–756.
- [79] A. S. Mitchell, S. Chakraborty, What does the mediodorsal thalamus do?, *Frontiers in systems neuroscience* 7 (2013) 37.
- [80] H. Baillieux, H. J. De Smet, P. F. Paquier, P. P. De Deyn, P. Mariën, Cerebellar neurocognition: insights into the bottom of the brain, *Clinical neurology and neurosurgery* 110 (8) (2008) 763–773.
- [81] C. B. Saper, T. E. Scammell, J. Lu, Hypothalamic regulation of sleep and circadian rhythms, *Nature* 437 (7063) (2005) 1257–1263.
- [82] E. K. Miller, J. D. Cohen, An integrative theory of prefrontal cortex function, *Annual review of neuroscience* 24 (1) (2001) 167–202.

- [83] E. D. Bigler, Anterior and middle cranial fossa in traumatic brain injury: relevant neuroanatomy and neuropathology in the study of neuropsychological outcome., *Neuropsychology* 21 (5) (2007) 515.
- [84] J. T. Giacino, S. Ashwal, N. Childs, R. Cranford, B. Jennett, D. I. Katz, J. P. Kelly, J. H. Rosenberg, J. Whyte, R. D. Zafonte, et al., The minimally conscious state: definition and diagnostic criteria, *Neurology* 58 (3) (2002) 349–353.
- [85] R. E. Cranford, Dialogue on end-of-life decision making: What is a minimally conscious state?, *Western Journal of Medicine* 176 (2) (2002) 129.
- [86] B. L. Edlow, J. Claassen, N. D. Schiff, D. M. Greer, Recovery from disorders of consciousness: mechanisms, prognosis and emerging therapies, *Nature Reviews Neurology* 17 (3) (2021) 135–156.
- [87] S. M. Sherman, R. W. Guillery, Functional connections of cortical areas: a new view from the thalamus, MIT press, 2013.
- [88] B. L. Edlow, M. Olchanyi, H. J. Freeman, J. Li, C. Maffei, S. B. Snider, L. Zöllei, J. E. Iglesias, J. Augustinack, Y. G. Bodien, et al., Sustaining wakefulness: brainstem connectivity in human consciousness, *bioRxiv* (2023).
- [89] M. P. Van den Heuvel, O. Sporns, Network hubs in the human brain, *Trends in cognitive sciences* 17 (12) (2013) 683–696.

Tiziana Currieri received the M.S. degrees in Computer Engineering from the University of Palermo, Italy, in 2021. Now she is currently pursuing her PhD in Biomedicine, Neuroscience and Advanced Diagnostics at the BiND Department. She was a Visiting Ph.D. Student at the Brain \& More Lab - Sano Centre for Computational Medicine, Krakow, Poland. Her research interests focus on the development and use of models in the field of machine learning and explainable artificial intelligence, with a special focus on medical image analysis methods. The main research activities are devoted to the analysis and processing of biomedical images to support clinicians' decision-making activities.

Joan Falcó-Roget studied physics and biophysics at the University of Barcelona and at the Autonomous University of Madrid. He is a PhD Student at the Sano Centre for Computational doing research in functional, diffusion, and effective connectivity in clinically oriented frameworks. He is also active in the field of theoretical neuroscience and has done research in normative models of animal behaviour and in quantum computing solutions applied to computational neuroscience as well as medicine.

Elham Rostami MD, PhD is a neurosurgeon and associate professor at Uppsala University and Karolinska Institute. She specializes in traumatic brain injury (TBI) with a focus on understanding the underlying pathology to develop personalized treatment approaches. Dr. Rostami has extensive experience in both clinical TBI management and research, contributing significantly to clinical guidelines and innovative therapeutic strategies. Her work spans from clinical trials to advanced neuroimaging techniques, with the aim of improving patient outcomes and bridging gaps in TBI care. She is also a key figure in international collaborations addressing brain injury research."

Salvatore Vitabile is a Full Professor with the Department of Biomedicine, Neuroscience and Advanced Diagnostics at the University of Palermo, Italy. He is co-author of more than 200 scientific papers in referred journals and conferences. He has chaired, organized, and served as member of the organizing committee of several international conferences and workshops. He is an Associate Editor of the Human-centric Computing and Information Sciences journal and an Editorial Board Member of Electronics. His research interests include medical data processing and analysis, clinical decision support systems, specialized architecture design and prototyping, and machine and deep learning applications.

Alessandro Crimi after completing his studies in engineering at the university of Palermo, obtained a PhD in machine learning applied for medical imaging by the University of Copenhagen, and an MBA in healthcare management by the University of Basel. Alessandro worked as post-doctoral researcher at the French Institute for Research in Computer Science (INRIA), Technical School of Switzerland (ETH-Zurich), Italian Institute for Technology (IIT), and University Hospital of Zurich. He is currently a professor at AGH Krakow.



Click here to access/download
LaTeX Source Files
Latex Manuscript.zip