



Politecnico di Milano
Department of Electronics, Information and Bioengineering
Doctoral Programme in Bioengineering

Methods for Uncertainty Assessment and Interpretability of Brain Connectivity Biomarkers in Localized and Widespread Degeneration

Candidate: **Davide Coluzzi**

Advisors:

Prof. Giuseppe Baselli

Prof. Anna Maria Bianchi

Coordinator of the PhD program:

Prof. Gabriele Angelo Dubini

XXXV Cycle 2020 – 2023

Acknowledgements

This thesis would not have been possible without the trust of my supervisor Prof. Giuseppe Baselli. I am very grateful for your ideas and guidance, the passionate discussions and for your personal continuous support, being on my side despite everything.

I would like to express my most sincere gratitude to my co-advisor Prof. Anna Maria Bianchi for all the opportunities and trust.

Then, I would like to thank Valentina Bordin and Massimo Walter Rivolta, it was such a pleasure working with you that it did not even feel like work. You are great researchers and even better friends.

My sincere thanks to all the professionals involved in my PhD research projects. Your passionate participation and advice made the PhD a great possibility of growth. I would like to acknowledge Francesca Baglio, Maria Marcella Laganà, Laura Pelizzari, Monia Cabinio and Alice Pirastru for your continuous support during SPIDER-NET development. I want to express my appreciation to Martin Oswaldo Mendez Garcia and all the collaborators in the SIDERA^B project for the support. I am very grateful to Alex Leow and all the people from the University of Illinois for supporting me in the last period of my PhD. Finally, I would like to express my sincere gratitude to Roberto Sassi, whose teaching provided me with the necessary skills to approach my PhD with success.

I am grateful to all the students who cooperated with me (Andrea M., Vanja, Davide, Mattia, Andrea B., Letizia, Cesare, Carlo) and whose collaboration enriched me so much.

I would like to thank my Milanese family of B³ Lab and all building 21, who always made me feel at home and helped me in the most difficult moments. My sincere thanks also go to my other second families of la Birroteca and Afropolis. I thank all the friends I met in this experience, who was with me from the first moment, who arrived later and who was just passing through the roads of Scotland. I would like to thank Peppe and Piero for the Covid-19 concerts on the balcony and being the perfect flatmates during quarantine, Tonino for his unfailing support, the complicity during Korobé music and being the best musician I have ever met, Lupe for helping me to close the pubs and going on without using the brakes and my “adoptive” brothers Donnie, Matthew, Stefanuccio, Giulietto and Cifa for really being with me despite any challenges that might have caused us to drift apart.

I thank Sara for being the flag of Plaza de la Revolución and standing close to me despite the distance “with the great tenacity which is typical of ancient things”.

I must express my very profound gratitude to my family. I would like to give a special thanks to my mother, “my rejuvenated and invincible” role model, and to Elisa and Anna, my “adult and incorruptible” nephews.

Abstract

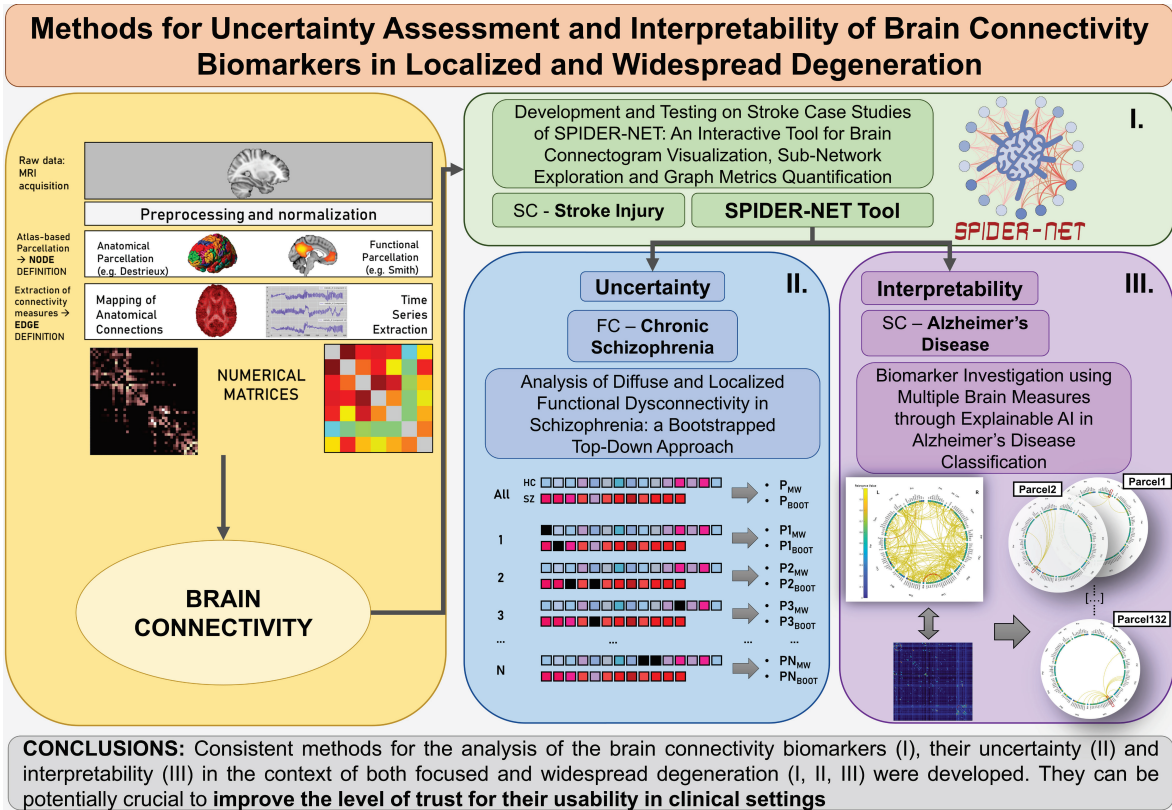
A number of different pathologies, ranging from mental disorder to neurodegeneration, share the characteristics of having both widespread effect throughout the whole brain and in specific sub-networks or regions. The analysis of the connectivity on different levels is a powerful tool to investigate impaired regions and global deficits, study cause and effects of the pathologies, support diagnoses and tailor the rehabilitative treatments. Nevertheless, conducting a multi-level assessment of brain connectivity outside the research settings is not a simple process due to three primary concerns: i) many tools are available, but a user-friendly, interactive and flexible environment allowing automatic qualitative and quantitative assessment at all levels is missing; ii) both structural and functional connectivity measures for edge-weighting lack gold-standard methodologies, with a number of uncertainty sources, resulting in noisy data; iii) the possible biomarkers which can be highlighted from huge amount of data, reducing uncertainty and using artificial intelligence (AI) methodologies, are not always adherent to domain knowledge and difficult to be interpreted. In this PhD work, methods for improving usability of the brain connectivity biomarkers were proposed.

More specifically, the aforementioned general aspects were addressed in three studies. First, an interactive and user-friendly tool called SPIDER-NET to allow qualitative and quantitative analysis of brain networks and sub-networks was developed. The tool was validated on the structural connectivity of 2 hemorrhagic stroke case studies and 17 healthy controls (HC). Second, a multi-level bootstrapping approach was applied to enable robust abnormalities detection. This approach was experimented on the functional connectivity of 12 schizophrenic patients and 15 HC. Finally, convolutional neural networks employing structural connectivity data and 3D T1-weighted volumes were developed and analyzed by Explainable Artificial Intelligence (XAI). The last study addressed the classification of Alzheimer's disease subjects (135 Magnetic Resonance sessions) and HCs (557 sessions). First, SPIDER-NET resulted an effective tool to represent the expected (dis)connectivity pattern due to a stroke lesion, in testing a-priori hypothesis by extracting a sub-network of interest and in investigating graph-based topological indexes. Furthermore, it allowed to better interpret complex networks and compare the results from two processing pipelines (Diffusion Tensor Imaging - DTI vs Constrained Spherical Deconvolution - CSD), having different uncertainty causes. Second, the bootstrapped top-down approach revealed different abnormalities of the schizophrenic group on different levels, which resulted to be more stable and robust compared to direct testing and having a trend towards results with greater number of data and subjects. Third, as evaluated through a statistical test ($p < 0.05$) and ranking of the most relevant parcels (first 15%), XAI analysis of interpretability

revealed the involvement of target brain areas for both models employing 3D T1-weighted volumes and structural connectivity. These anatomical targets were the medial temporal lobe and the default mode network, respectively. Although the obtained findings had limitations, results suggested that combining different imaging modalities may lead to increased model performance, interpretability, and, thus, reliability.

Although the great potential of measures extracted from brain connectivity, the open issues of uncertainty and interpretability limited their trust and, thus, their usability within clinical settings. Consistent methods to address these issues have a direct connection to the understanding of the relationships between localized affection and widespread degeneration. Improved reliability and interpretation are fundamental in the study of the brain both in health, to map the nervous system and comprehend the mechanisms underlying the brain processes, and in disease, to support clinicians in the early detection and during rehabilitation.

Graphical Abstract



Contents

| | |
|---|-----------|
| Acknowledgements | 2 |
| Abstract | 3 |
| Graphical Abstract | 5 |
| Contents | 6 |
| List of Publications | 8 |
| Author’s Contribution | 9 |
| List of Tables | 11 |
| List of Figures | 12 |
| 1 Extended Summary | 15 |
| 1.1 Introduction..... | 15 |
| 1.2 Methods | 16 |
| 1.3 Results..... | 27 |
| 1.4 Discussion..... | 34 |
| 2 Cover Essay | 40 |
| 2.1 Motivation: Localized and Widespread Degeneration | 40 |
| 2.2 Background: Brain Networks from dMRI and fMRI | 45 |
| 2.2.1 Imaging-Based Parcellation of the Human Brain | 46 |
| 2.2.2 Methods for the Extraction of Macroscale Connectivity | 48 |
| 2.2.3 Limitations of Macroscale Connectivity | 55 |
| 2.2.4 Analysis of Complex Networks | 59 |

| | | |
|----------|--|------------|
| 2.2.5 | Connectomics | 63 |
| 2.3 | State of the Art and Open Issues in Brain Connectivity..... | 69 |
| 2.3.1 | Mapping, Visualization and Analysis of Brain Networks and Sub-Networks | 70 |
| 2.3.2 | Robustness of the Connectivity Measures | 74 |
| 2.3.3 | Explainable Artificial Intelligence in Brain Imaging and Connectomics 76 | |
| 2.3.4 | Multi-Level Degenerations | 80 |
| 2.4 | Objectives and Outline | 83 |
| 2.5 | Summary of the Publications..... | 84 |
| 3 | Discussion and Conclusions..... | 90 |
| | Annex | 96 |
| | Scientific Publications | 161 |
| | List of Abbreviations | 163 |
| | References..... | 165 |

List of Publications

This thesis consists of a cover essay and of the following publications which are referred in the text by their Roman numerals.

I: Coluzzi, D.; Pirastru, A.; Pelizzari, L.; Cabinio, M.; Laganà, M.M.; Baselli, G.; Baglio, F.; “Development and Testing of SPIDER-NET: An Interactive Tool for Brain Connectogram Visualization, Sub-Network Exploration and Graph Metrics Quantification”, *Frontiers in Neuroscience*, <https://doi.org/10.3389/fnins.2022.818385>, Special Issue Advanced Computational Tools for Mapping the Multidimensional Architecture of the Brain, 2022

II: Coluzzi, D.; Baselli, G.; “Diffuse and Localized Functional Dysconnectivity in Schizophrenia: a Bootstrapped Top-Down Approach”, *Fundamenta Informaticae*, <https://doi.org/10.48550/arXiv.2305.02369>, Special Issue Tomography and Applications, 2022
(Accepted, Pending Publication)

III: Coluzzi, D.; Bordin, V.; Rivolta, M.; Fortel I.; Leow A.; Baselli, G.; “Biomarker Investigation using Multiple Brain Measures from MRI through XAI in Alzheimer’s Disease Classification”, <https://doi.org/10.48550/arXiv.2305.03056>
(Under Submission)

Author's Contribution

Publication I: “Development and Testing of SPIDER-NET: An Interactive Tool for Brain Connectogram Visualization, Sub-Network Exploration and Graph Metrics Quantification”

Based on an idea of F. Baglio and G. Baselli, the author selected the graph analysis methods to be implemented and their effective visual and numerical representation. Next, he developed the software tool and validated it on two stroke cases studies in comparison to a control group. The tool was implemented based on discussions with all the co-authors, with the dataset used for validation that was processed by A. Pirastru, L. Pelizzari, M. Cabinio, M. M. Laganà. The author wrote the computational codes, made the data analysis and wrote the first draft of the manuscript, that was improved with the comments of the co-authors.

Publication II: “Diffuse and Localized Functional Dysconnectivity in Schizophrenia: a Bootstrapped Top-Down Approach”

The author designed the study, developed the second release of the software tool, performed the statistical analyses, wrote the computational codes, derived the results and wrote the manuscript. The idea of the paper was based on discussions with G. Baselli, who improved the paper with his comments.

Publication III: “Biomarker Investigation using Multiple Brain Measures from MRI through XAI in Alzheimer’s Disease Classification”

The author designed the study, performed the analyses, wrote the computational codes, derived the results and wrote the manuscript in cooperation with V. Bordin. The author focused on the Artificial Intelligence models training and testing, and implementation of the explainability method. The idea of the paper was based on discussions with G. Baselli and V. Bordin. The author wrote the first draft of the paper in cooperation with V. Bordin, and all co-authors improved the paper with their comments.

List of Tables

| | |
|--|----|
| Table 1: Characteristics of the datasets employed, and pre-conditioning steps performed. SZ: Chronic Schizophrenia; HC: Healthy Control; AD: Alzheimer’s Disease; AAL: Automated Anatomical Labeling; HOA+AAL: FSL Harvard-Oxford maximum likelihood cortical Atlas combined AAL; DTI: Diffusion Tensor Imaging; CSD: Constrained Spherical Deconvolution; MNI: Montreal Neurological Institute; FOV: Field Of View; | 20 |
| Table 2: Results of the statistical tests performed on global indexes. The global version of the degree represents the mean of all edges associated to all nodes, whereas the density is the ratio between the number of present edges and the number of possible edges in the network. For this reason, the result coincides, and it is reported only once. HC: Healthy Controls; SZ: Schizophrenic Patients; RST: Robust Statistical Test; MW: Mann Whitney Test; BOOT: Bootstrap Hypothesis Testing; #Sub: number of subjects removed; CC: Clustering Coefficient; CPL: Characteristic Path Length; Eff: Efficiency. | 30 |
| Table 3: Normalized confusion matrix for the ResNet18 and BC-GCN-SE model performance, reported as median value and interquartile range..... | 32 |
| Table 4: Most widespread DTI metrics, with acronyms and formulation..... | 51 |
| Table 5: Graph-based topological properties with graphical representation and formulation divided according to the level of analysis..... | 63 |
| Table 6: Intepretability methods in neuroimaging studies. *Applied for Autism Spectrum Disorder classification task..... | 79 |

List of Figures

- Figure 1: Flowchart for SPIDER-NET usage. First, the Atlas and Label input files are browsed and loaded (blue box). Then, the Connectivity Matrix file is loaded and the selection of the sub-network of interest is performed (red box). Optionally (dashed lines), it is possible to compute and visualize topological properties of the selected brain network (green box). Finally, the connectogram is generated according to the selection made and to the chosen visualization settings (yellow box)..... 21
- Figure 2: Pipeline of the top-down approach proposed. The functional connectivity matrixes of the two populations were analyzed at the global level, extracting the DMN as sub-network of interest and investigating the distributions of the local indexes to the edge level, assessing correlations and anticorrelations..... 23
- Figure 3: Schema representing the robustness statistical test randomly removing two subjects (black squares) at each iteration (RST2). In this case, N is equal to 350. 25
- Figure 4: Implemented workflow for the AD classification and explainability assessment performed on ResNet18 and BC-GCN-SE. On the top right, the processing steps used to derive the SC data (displayed as matrixes and connectograms in the yellow rectangle) are shown. The models' architecture is shown in the middle (blue box) for both models comprising convolutional layers, Squeeze-and-Excitation blocks (SE), Global Average or Edge/Node Pooling (GAP, EP/NP) layer, Fully Connected layers and sigmoid activation function for the binary classification. The outputs derived from the convolutional layers were processed using Grad-CAM (bottom, green box) and then averaged. The final heatmap was then multiplied for the binary masks underlying the HOA + AAL atlas parcels in the 3D T1-weighted case and averaged across rows/columns (node) in the SC data case. 27
- Figure 5: Investigation of connectograms and local topological properties of HCs, Case 1, Case 2. On the left, the connectograms (B) derived by both DTI (left) and CSD (right) processing, of the sub-network extracted using as seeds all parcels which are overlapped with the stroke lesion of either Case 1 or Case 2, are reported for HCs (top panel), Case

1 (middle panel) and Case 2 (bottom panel). Specifically, seeds were defined according to lesions and all the parcels of the brain were considered as targets for connectivity analysis. On the right, local node degree computed for HCs, Case 1 and Case 2. The X-axis represents the 165 brain parcels, even if only lobe labels are reported (e.g., Fro). The left hemisphere is represented in the left half of the graph, while the right hemisphere is represented in the right half. Vertical colored stripes represent different lobes (e.g., the frontal lobe is represented in pink). The same lobe in left and right hemispheres is shown with the same color. Each dot in the graph represents the local node degree of a brain parcel. The 17 parcels (10% of 165) exhibiting the highest local node degree are represented as red dots. The 17 parcels (10% of 165) with the lowest local node degree are represented as yellow dots. All the other parcels are represented as blue dots. The right Putamen is highlighted by an arrow. L-left hemisphere, R-right hemisphere, Fro-frontal, Ins-insular, Tem-temporal, Par-parietal, Occ-occipital, Sbc-subcortical, CeB-cerebellum, Bst-brainstem, PrCG-precentral gyrus, LoInG/CInS-long insular gyrus and central insular sulcus, ShoInG-short insular gyri, Pal-pallidum, Pu-putamen, CaN-caudate nucleus, Tha-thalamus..... 28

Figure 6: Distributions of the local topological properties and connectograms showing anti-correlations. On the left top, distributions of the local degree values in the two populations divided according to nodes of DMN in left and right hemisphere (yellow) or not (gray) are shown. On the left bottom, distributions of the local strength values in the two populations. Strength of nodes which are statistically significant different ($p < 0.05$) between the two populations are shown in different colors according to the different methods used: pale brown for bootstrap hypothesis testing, yellow for Mann-Whitney and red for statistically significantly different nodes found by both. On the right, connectograms of negative average group networks. a) and b) are the results on the whole-brain, c) and d) on only the DMN of SZ and HC groups respectively. 5% density thresholding is applied on the shown connectograms for graphical clarity only. The order and colors of the lobes is the same as the rectangles on the left bottom. 31

Figure 7: Mean RV of all parcels, for the ResNet18 (panel A) and BC-GCN-SE (panel B) models. The AD and HC classes are reported separately. The lobe belonging is indicated through rectangle colors. The RV of each parcel is labeled though dot colors according to the following criterion: red indicates the 15% of parcels characterized by

highest RV, yellow indicates the 15% of parcels characterized by lowest RV, grey indicates the remaining parcels. Plots were created using the SPIDER-NET tool. Fro: Frontal Lobe; Ins: Insular Cortex; Lim: Limbic Lobe; Tem: Temporal Lobe; Par: Parietal Lobe; Occ: Occipital Lobe; SbC: Subcortical Structures; Ceb: Cerebellum; BSt: Brainstem. 33

Figure 8: Sagittal, coronal and axial views of the Harvard-Oxford atlas. 47

Figure 9: From tensor to tractography [59]. Examples of different levels of anisotropy resulting in a diverse shape of the diffusion ellipsoid, resulting DTI ellipsoids from axial slice of the brain, zoomed in one area, and the resulting tractography visualization are shown. 50

Figure 10: Static and dynamic functional connectivity, which identify states characterized by communities of nodes and edges. The image is taken from a previous study by Patil and colleagues examining static and dynamic configuration of FC brain networks associated with creative cognition in young adults [78]. 58

Figure 11: Schema of the phases of processing for the definition of brain structural and functional network. 65

Figure 12: Reconstruction of Phineas Gage's damage from the work made by Van Horn and colleagues [122]. The skull on display at the Warren Anatomical Museum at Harvard Medical School, computed tomography image volumes, rendering and view of the interior of the skull and the connectogram are shown. 71

Figure 13: Connectograms of the sub-network composed by the ROIs overlapping the stroke lesions (reported in B) of a group of HC and two stroke case studies, together with their five axial slices. 73

Figure 14: Results obtained from research made on the Scopus Database. The number of documents published per year is reported in green, using as keyword “connectivity”, and in blue, using “images”, together with Deep Learning and MRI or fMRI. 77

1 Extended Summary

1.1 Introduction

The human brain is one of the most complicated systems in nature, with a highly intricate, dense, multi-scale, and multi-level network architecture [1]. From a macroscale perspective, it can be analyzed from an anatomical or functional point of view. Indeed, the brain can be represented as a complex network depicting gray matter (GM) regions as nodes and structural (white matter fiber bundles) or functional (Blood Oxygenation Level Dependent - BOLD signal activations) characteristics as edges. Once defined the network, it can be examined in its segregation and integration properties, thus in its global configuration, on sub-networks or communities' existence, in local regions of interest and in the single edges. This is possible through the well-established topological properties from graph theory [2]. Different abnormalities for several pathologies were thus found in a number of studies using these measures [3]–[10]. More specifically, a number of different pathologies, ranging from mental disorders to neurodegeneration, share the characteristics of having both widespread effect throughout the whole brain and in specific sub-network or regions resulting in diverse (dis)connection patterns. In this sense, the analysis of the connectivity on different levels could be of great interest to study cause and effects of the pathologies, support diagnoses and tailor the rehabilitative treatments. However, this parallel multi-level evaluation is not that easy considering three main issues: i) many tools are available, but a unique, accessible, user-friendly, interactive and flexible environment allowing the qualitative and quantitative assessment on all the levels in group studies is missing; ii) the connectivity measures extracted from both structural (SC) and functional (FC) connectivity for edge-weighting lack a gold-standard methodologies, with a number of uncertainty sources, resulting in noisy data; iii) the possible biomarkers which can be

highlighted are not always adherent to domain knowledge and machine and deep learning methodologies used for diverse tasks are difficult to be interpreted.

These three aspects strongly limited the diffusion of the brain connectivity biomarkers out of the research context and the usability within clinical settings. For this reason, we aimed with this PhD thesis:

- to fill the gaps of uncertainty and interpretability of existing approaches for the mapping, visualization, extraction and analysis of brain connectivity;
 - developing a novel interactive and flexible software tool called SPIDER-NET that allows for the user-friendly qualitative and quantitative analysis of brain networks and sub-networks and validating it on multimodal structural data (i.e., DTI and CSD) (study I);
 - proposing a robust multi-level investigation through bootstrapping of the functional brain connectivity topological measures in the context of a group study (study II);
 - employing XAI in Alzheimer’s Disease (AD) classification task of the structural brain data (3D T1-weighted volumes and structural connectivity) to assess and compare the interpretability of known brain biomarkers from multiple brain measures (study III);
- to demonstrate the potential of the proposed methods in the context of pathologies characterized by both focused disruptions and widespread degeneration, such as stroke, schizophrenia (SZ) and AD, with the purpose of improving the usability of the brain connectivity biomarkers and improve the understanding of brain functioning in health and disease.

1.2 Methods

Data Acquisition, Study Population and Pre-Conditioning

For each of the three studies a different dataset was used:

- I. the first dataset comprises two patient case studies with stroke (males, 44 and 37 years old, referred to as Case 1 and Case 2, respectively) who had right hemisphere lesions with prominent subcortical expression, as well as 17 HC

participants (7 males and 10 females; mean \pm σ age: 52.5 ± 8.3 years). Each subject provided a written informed permission and was enrolled at the IRCCS Fondazione Don Carlo Gnocchi in Milan. Six months after hemorrhagic stroke the patients underwent the Magnetic Resonance Imaging (MRI) examination. All the participants performed it on a 1.5 T Siemens Magnetom Avanto scanner equipped with a 12-channels head coil.

DTI-based and CSD-based techniques were employed to construct SC matrixes for both stroke case studies and HCs. Specifically, the edges of the matrixes were established based on the reconstructed fibers count (NF) of white matter tracts connecting the 165 parcels of the Destrieux Atlas [11]. Furthermore, to account for differences in brain volumes, NF was normalized by the sum of the volumes of the respective connected parcels. To obtain a probabilistic group matrix that represented the HC group as a whole, proportional thresholding was employed retaining only connections shared by at least half of HCs and resulting in a matrix with density that is the same of the initial median population density. Thus, the HC group mean matrix was obtained. The acquisition protocol and other details are reported in article I.

- II. the second dataset was collected as part of the research undertaken by Zalesky and colleagues [12]. The dataset comprises of 15 HC with a mean age of 33.3 years ($\sigma = 9.2$ years), 14 of whom were male, and 12 individuals with chronic schizophrenia with a mean age of 32.8 years ($\sigma = 9.2$ years), 10 of whom were male. The patients were diagnosed based on the standard operational criteria outlined in the Diagnostic and Statistical Manual of Mental Disorders IV. The HC and SZ groups were matched with regards to age, pre-onset IQ, and years of education. T2*-weighted echo-planar images depicting blood oxygenation level dependent contrast were obtained using a 1.5 Tesla scanner (GE Signa, General Electric, Milwaukee, WI). The construction of the matrix involved defining nodes based on a specific set of areas within the Automated Anatomical Labeling (AAL) atlas [13], while the edges were generated by calculating the correlation between preprocessed time series. Due to poor coverage in certain subjects, it was not possible to accurately estimate the node-averaged time series for certain brain regions defined by the AAL atlas. Therefore, these brain

regions were excluded resulting in connectivity matrixes with dimensions of 74 x 74. According to the aim of this study, implying the assessment of the robustness of the bootstrapping-based method regardless the thresholding method, original data with no thresholding was employed. However, a threshold in 0 was applied as pre-conditioning step to extract positive and negative matrixes, allowing their separate analysis. Prior to data collection, all participants provided written informed consent. All the other details are reported in article II and in [12].

- III. The third dataset used is the third release of the Open Access Series of Imaging Studies (OASIS-3), which is a publicly available longitudinal collection of data focused on the effects of normal aging and early-stage AD [14]. The dataset consisted of 1098 participants, including 605 HCs and 493 individuals at varying stages of cognitive decline who underwent neuroimaging and clinical assessments. Matching MRI and clinical data within a 3-month time span yielded 1076 sessions, of which 874 had available T1-weighted and BOLD scans. This latter selection criteria of having both T1-weighted and BOLD scans was chosen in agreement to a previous study that extracted 1326 SC and FC matrixes from the OASIS-3 imaging sessions [15]. Matching these 1326 sessions with the 874 from the former procedure resulted in a subset of 692 sessions from 543 participants, with an age range of 42-95 years and a mean age of 70.06 ± 8.85 years (388 female and 304 male). Each session was associated with a T1-weighted scan and a SC and FC matrix. The presented study does not involve functional data. The reason is that, in this first approach, the primary objective was to compare biomarkers relevant to the structural/anatomical brain damage as seen in the T1 anatomical volumes and in the SC data derived from DTI. Sessions were classified as either HC or AD based on the Clinical Dementia Rating (CDR) Scale, with a CDR score of 0 representing normal cognitive function and scores of 0.5, 1, or 2 indicating varying degrees of impairment. A small percentage of the subjects (2.76%) have multiple sessions assigned to different CDR class that were independently analyzed. The final ratio of HC to AD sessions was 557:135. Of all the available T1-weighted scans, 135 were acquired with a 3T Siemens

Biograph_mMR scanner, while the remaining 557 with a pair of 3T Siemens TimTrio scanners (Siemens Medical Solutions USA, Inc). All acquisition details are reported in article III. As the processing steps, the T1-weighted volumes were divided into 132 brain-covering regions, including 91 cortical and 15 subcortical parcels from the FSL Harvard-Oxford maximum likelihood cortical atlas (HOA) [16], and 26 cerebellar parcels from the Automated Anatomical Labelling atlas (AAL) [13]. The resulting combination of the HOA and AAL atlas is referred to as HOA + AAL. On the one hand, these gray matter regions were combined with DTI white matter fiber tracking to generate 692 undirected graphs. Further details on the minimal pre-conditioning are reported in article III, in the study by Amodeo and colleagues [15] and in a diagram of the SC matrix extraction process (Figure 4, top right). For this reason, and the interest in analyzing the result of XAI on single and possible spurious edges, no thresholding method was applied. On the other hand, the pre-processing of the available T1-weighted scans using the FSL v.6.0 tool [17] was carried out to create a suitable dataset, as better detailed in article III. Importantly, registration to the Montreal Neurological Institute space was carried out for all subjects both for the T1-weighted volumes and for the DTI parcellation. The main characteristics of the datasets are summarized in Table 1.

| Study | Population | Data | Atlas (#nodes) | Edge-Weighting | Pre-Conditioning |
|-------|-------------------------------|---|---------------------|---------------------|---|
| I | 2 Stroke Cases vs 17 HC | Structural Connectivity | Destrieux (165×165) | DTI and CSD | Normalization by volumes, HC mixed proportional-consistency thresholding to create an HC group matrix |
| II | 12 SZ vs 15 HC | Functional Connectivity | AAL (74×74) | Pearson Correlation | Exclusion of poor coverage AAL nodes, separate analysis of correlations and anticorrelations |
| III | 135 AD vs 557 HC MRI sessions | Structural Connectivity, 3D T1-Weighted Volumes | HOA+ AAL (132×132) | DTI | On 3D T1-weighted scans: skull-stripping, bias field correction, registration to standard MNI template, cropping of the images FOV, normalization of intensity values using variance scaling, resize to 115x144x118 |

Table 1: Characteristics of the datasets employed, and pre-conditioning steps performed. SZ: Chronic Schizophrenia; HC: Healthy Control; AD: Alzheimer’s Disease; AAL: Automated Anatomical Labeling; HOA+AAL: FSL Harvard-Oxford maximum likelihood cortical Atlas combined AAL; DTI: Diffusion Tensor Imaging; CSD: Constrained Spherical Deconvolution; MNI: Montreal Neurological Institute; FOV: Field Of View;

SPIDER-NET Tool and Stroke Case Studies Investigation

SPIDER-NET (Software Package Ideal for Deriving Enhanced Representations of brain NETWORKS) is developed in Matlab and delivered as standalone software. It is a tool that enables users to select, visualize, and quantify partial connectograms of brain networks in a flexible and user-friendly manner. The SPIDER-NET Graphical User Interface (GUI) allows for rapid exploration of networks and interactive real-time sub-network definition. The software automatically generates connectivity figures based on user selections and offers additional features for matrix thresholding, computation of network indices, and interactive visualization preferences. Three input files representing an Atlas file, a Label

file and a Connectivity Matrix file are required. After uploading the inputs, the GUI (Figure 1, red box) allows the user to choose between selecting individual parcels, entire groups of parcels, or attributes specified in the “Atlas” file. This selection feature is enabled in the GUI following input upload (Figure 1, blue box).

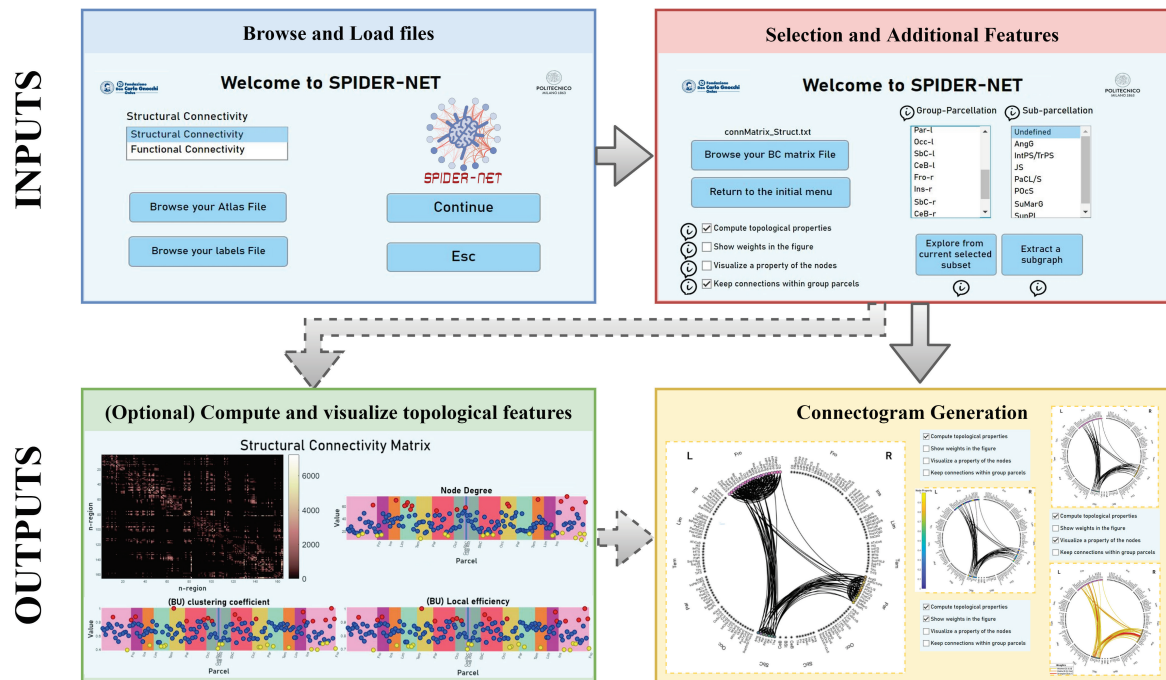


Figure 1: Flowchart for SPIDER-NET usage. First, the Atlas and Label input files are browsed and loaded (blue box). Then, the Connectivity Matrix file is loaded and the selection of the sub-network of interest is performed (red box). Optionally (dashed lines), it is possible to compute and visualize topological properties of the selected brain network (green box). Finally, the connectogram is generated according to the selection made and to the chosen visualization settings (yellow box).

SPIDER-NET provides two complementary logic options for defining a partial network, as represented by the input Connectivity matrix (shown in the red box in Figure 1). These options are “Explore from current selected subset” (Option 1) and “Extract a subgraph” (Option 2). Option 1 allows to define a “seed” and a “target”, which can be a parcel, or a group of parcels defined by group-parcellation, or attribute depicted in the atlas file. The target can also be all the parcels of the brain. Option 2 involves choosing a single subset of nodes, which can be done on individual parcels or by groups. Then, a number of additional features are present, such as the exclusion of within-edges (e.g., intra-lobe connections), color-coding node properties and thresholding. SPIDER-NET also has an optional feature

(shown in the green box in Figure 1) that allows for the computation of graph-based topological properties. These properties are useful for a quantitative evaluation of the network's topology and can be interactively explored on a visualization panel. The most widespread graph-based indexes automatically computed are reported in Table 5 of Par. 2.2.4.2.

On the available data, DTI-based connectograms of the right hemisphere were generated for Case 1 and Case 2 to firstly visualize the effect of the lesion. A subgraph analysis was then performed by selecting brain regions overlapping with the lesions as seeds and the whole brain as target. The same sub-network analysis was performed for HC as well, allowing for a comparison. Local and global graph analysis indexes were extracted from both the weighted and binary connectivity matrixes to describe network topology. Additionally, the study compared the connectivity results derived from DTI and CSD processing techniques.

Bootstrapped Multi-Level Approach in Schizophrenia

SPIDER-NET was then improved and upgraded to have the possibility of uploading more than one matrix at once allowing group studies. Furthermore, a top-down bootstrapping approach was proposed to enable a robust multi-level assessment. This approach was used to compare the HC and SZ groups involving several levels of analysis. The investigation began with the highest level, which involved examining and comparing the overall characteristics of the matrixes from the two populations, and proceeded to the lowest level, which involved analyzing individual connections. To ensure the accuracy of the investigation, bootstrapping was performed on both the global topological indexes and the local indexes of the Default Mode Network (DMN). Community analysis was performed before obtaining the group mean matrixes which were studied to identify the strongest activations and deactivations, with a focus on correlations and anticorrelations. The whole pipeline is summarized in Figure 2.

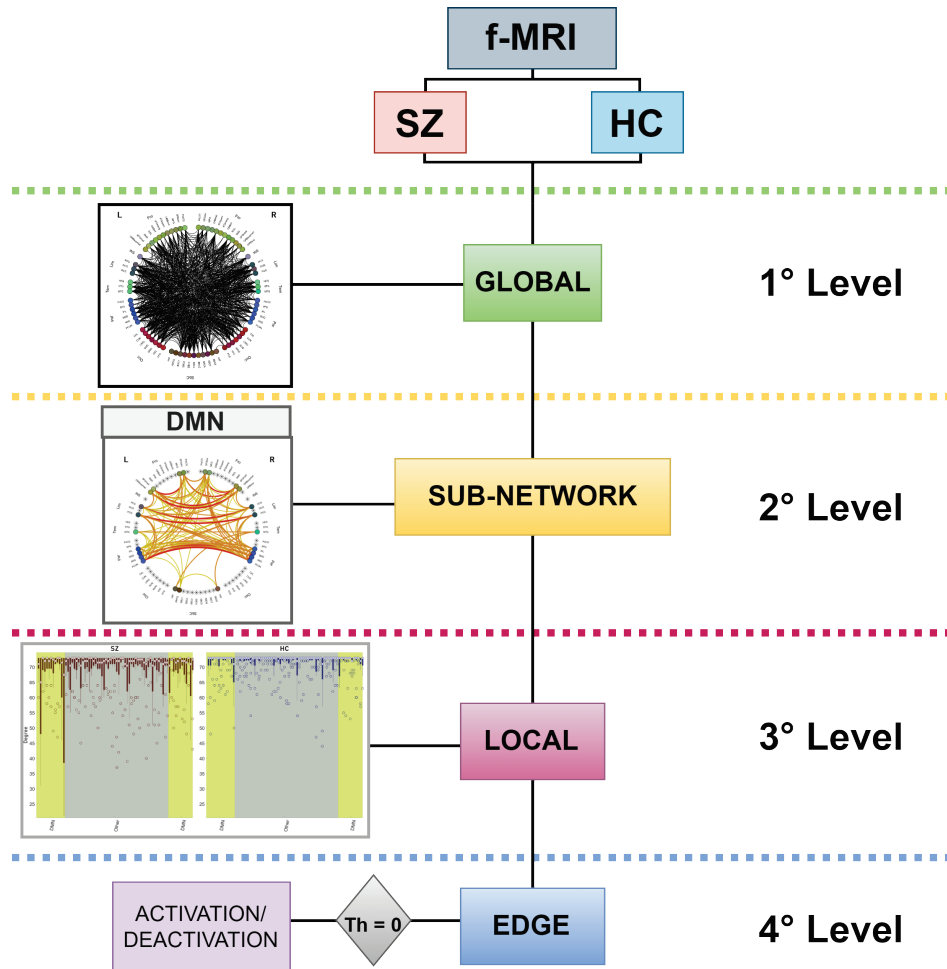


Figure 2: Pipeline of the top-down approach proposed. The functional connectivity matrixes of the two populations were analyzed at the global level, extracting the DMN as sub-network of interest and investigating the distributions of the local indexes to the edge level, assessing correlations and anticorrelations.

Mann-Whitney testing (MW) was used to determine whether there were statistically significant differences in global indexes between two groups. However, due to factors such as the sample size ($\#HC = 15$; $\#SZ = 12$), the non-stationarity of the data [25], and the possibility of spurious connections caused by limitations in functional Magnetic Resonance Imaging (fMRI) processing [29], the reliability of significant group differences was also assessed. To address this, the study also used bootstrap hypothesis testing (BOOT). BOOT provides a more accurate estimate of the null distribution of network measures, which allows for the evaluation of confidence intervals and the uncertainty of statistics, resulting in a more robust detection of abnormalities [41]. The procedure to test bootstrap hypothesis was the following:

1. calculation of the test statistic $\hat{\theta} = |\tilde{x} - \tilde{y}|$, given x_1, \dots, x_N a random sample from distribution F with median \tilde{x} and y_1, \dots, y_M another independent random sample from distribution G with median \tilde{y} ;
2. Bootstrapping: extraction of B sets of random samples x^* (size N) and y^* (size M) with replacement from x and y , respectively;
3. Calculation of the test statistic $\hat{\theta}_b^* = |\tilde{x}_b^* - \tilde{y}_b^*|$ for each resample;
4. These B resampled test statistics are then made into a null distribution by $\hat{\theta}'_b = \hat{\theta}_b^* - \hat{\theta}$
5. Estimate of the p-value as

$$p = \frac{\sum_{b=1}^B C\{\hat{\theta}'_b \geq \hat{\theta}\} + 1}{B + 1}$$

Where $C\{condition\} = 1$ when the condition is true and 0 otherwise.

F and G represent the distributions of a global index in HC and SZ, thus N and M representing the number of HC and SZ, 15 and 12 respectively. The resampling from these two sets is carried out 5000 (B - number of resamples) times for all bootstrap hypothesis tests and the test statistic is the difference in medians [42].

To evaluate the robustness of the measures, a leave-n-subject-out approach was employed with n equal to 1 and 2. Both methods were used to analyze variations in the statistics computed, by changing the population samples according to the two procedures. The first procedure, referred to as Robust Statistical Testing with leave 1 subject out - RST1, involved evaluating the variation of p-values obtained by removing one at a time all subjects from the population for performing MW tests and creating new resamples for BOOT. The second procedure, referred to as RST2, involved the random removal of pairs of subjects from the population for performing both tests. In total, 350 random extractions of pairs of subjects were performed, representing about 50% of the possible pairs ($27 \cdot 26 = 702$). RST2 is summarized in Figure 3. Mean and standard deviation were evaluated across all tests for both procedures.

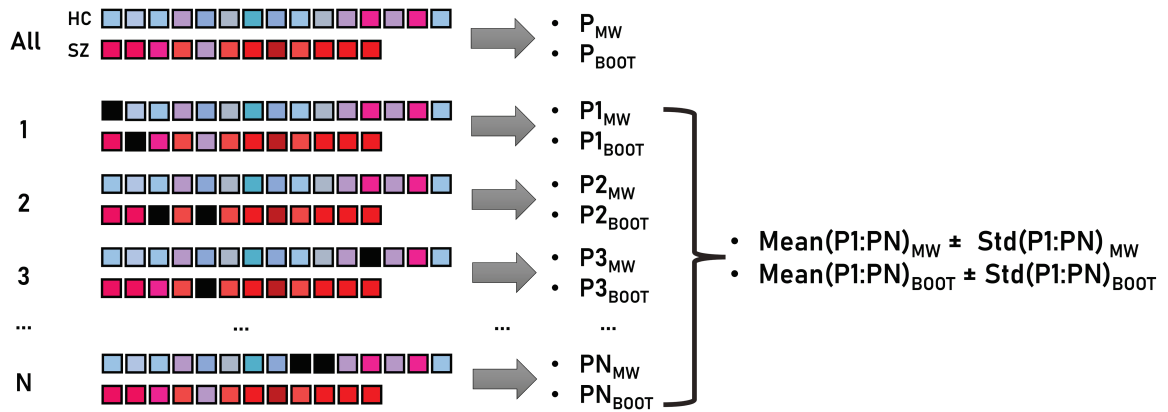


Figure 3: Schema representing the robustness statistical test randomly removing two subjects (black squares) at each iteration (RST2). In this case, N is equal to 350.

XAI Approach for Alzheimer’s Disease Classification

In order to improve the interpretability of connectivity networks in a broader study involving normal aging and early-stage AD, XAI methods were applied to the dataset III, as described in Section 1.2. These methods aimed to analyze anatomical biomarkers and provide a clearer understanding of data and its features. First, Deep Learning (DL) methods were employed on both SC data and 3D T1-weighted volumes. After pre-processing, the dataset was split into training and validation using a 10-fold cross-validation strategy, with an attentive split to avoid data leakage. Performances were evaluated by computing the median and the interquartile range of the true positive (TPR) and true negative (TNR) rates, beyond that accuracy, across all 10 folds. Finally, a new dataset split maintaining hyperparameters and proportion between HC and AD was carried out to assess the explainability of the two DL models. Resnet18, a pre-trained DL model, was utilized for the AD recognition with 3D T1-weighted volumes. The model architecture is characterized by 18 layers including a 3D convolutional layer and four sets of residual blocks, each containing two 3D convolutional layers. More details on the architectures are reported in Figure 4 and hyperparameters details in the article III. This model was originally trained for multi-class classification of images from the ImageNet database [18], which was adapted to 3D input [19]. The selection of Resnet18 was based on previous research, which has demonstrated its superior performance compared to other pre-trained models for AD recognition [40]. As the DL model using graph-structured data, several

graph convolutional networks (GCNs) have been proposed for various applications, but these models are typically designed for sparse graph-structured data, resulting in unsuitable models for brain connectivity. Brain connectivity data yields dense graphs, and the DTI-derived metrics result in graphs with a high number of connections, making both direct and indirect connections essential for communication. Recently, an edge based GCN was proposed and used successfully in FC tasks. Indeed, the BC-GCN model, with the potential extension of a Squeeze-and-Excitation block (BC-GCN-SE), performed well in regression and classification tasks [20]. We adapted this model to AD classification task with SC data. The resulting BC-GCN-SE is mainly composed of five major units: the graph path convolution (GPC), which allow the extraction of the feature maps, the edge (EP) and node pooling (NP), the Squeeze-and-Excitation (SE), to emphasize or suppress them, and the fully connected block for classification. More details on the architectures are reported in Figure 4 and hyperparameters details are reported in article III. The explainability analysis was based on GRAD-CAM method [21], and the mean heatmap G_c of different convolutional layers in both models was used to address both class-discriminative features with low spatial extent and fine-grained details. Afterwards, we related these mean heatmaps obtained from both ResNet18 and BC-GCN-SE to the HOA + AAL atlas to assess which brain regions were most involved in the classification and comprehend the decisions made by the models. Thus, we defined a Relevance Value (RV) measure for each parcel p in both 3D T1-weighted volumes and SC matrixes:

$$RV_{p,c} = \frac{\sum_{x=1}^{115} \sum_{y=1}^{144} \sum_{z=1}^{118} M_p(x, y, z) G_c(x, y, z)}{\sum_{x=1}^{115} \sum_{y=1}^{144} \sum_{z=1}^{118} M_p(x, y, z)} \quad (1)$$

$$RV_{p,c} = \frac{\sum_{q=1, q \neq p}^{132} G_c(p, q)}{131} \quad (2)$$

Where c is the considered class, G_c is the resulting mean heatmap from Grad-CAM, and $M_p(x, y, z)$ is a binary mask obtained from the HOA + AAL atlas to depict each parcel p . A diagram summarizing the XAI steps for 3D T1-weighted volumes and SC are reported in the bottom parts of Figure 4 (green boxes). XAI assessment was performed in comparison to domain-knowledge, addressing anatomical targets of the two models. As for the ResNet18 model, the medial temporal lobe (MTL), was selected as 3D T1-weighted volume anatomical target, whereas for BC-GCN-SE model, the DMN was selected as SC

target. According to previous studies [22], [23], patterns of atrophy in the MTL represent a well-established structural MRI biomarker for AD and is often used a diagnostic criterion for individuals displaying early symptoms, while DTI and SC abnormalities are found in the DMN [24], [25]. These anatomical targets were assessed evaluating statistical differences between AD and HC and then ranking the RV for each of the parcels. More details on the statistical evaluation are reported in article III.

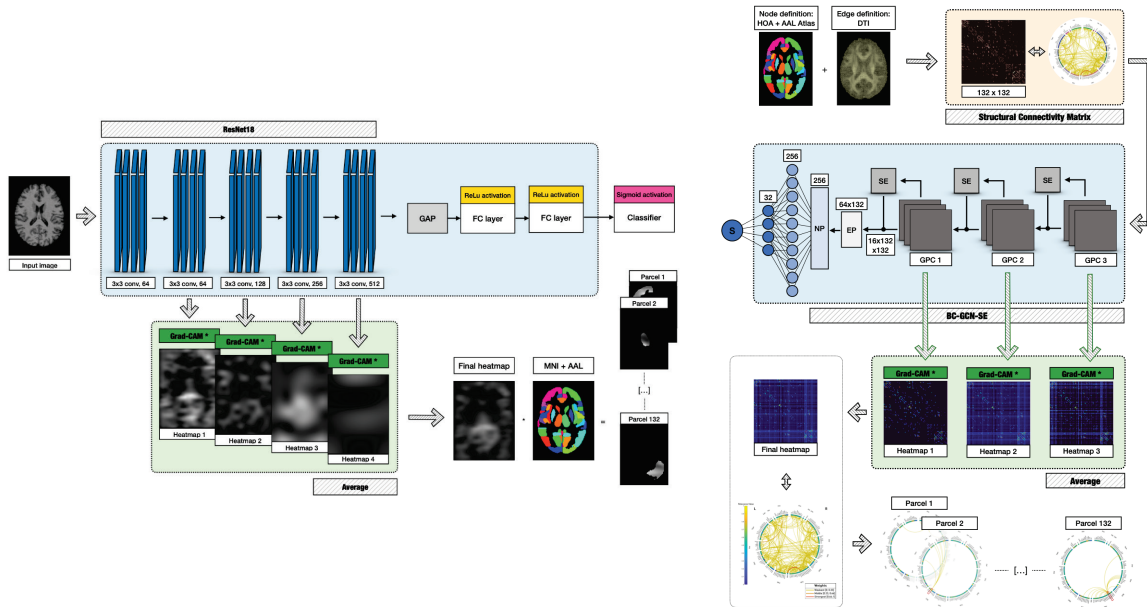


Figure 4: Implemented workflow for the AD classification and explainability assessment performed on ResNet18 and BC-GCN-SE. On the top right, the processing steps used to derive the SC data (displayed as matrixes and connectograms in the yellow rectangle) are shown. The models' architecture is shown in the middle (blue box) for both models comprising convolutional layers, Squeeze-and-Excitation blocks (SE), Global Average or Edge/Node Pooling (GAP, EP/NP) layer, Fully Connected layers and sigmoid activation function for the binary classification. The outputs derived from the convolutional layers were processed using Grad-CAM (bottom, green box) and then averaged. The final heatmap was then multiplied for the binary masks underlying the HOA + AAL atlas parcels in the 3D T1-weighted case and averaged across rows/columns (node) in the SC data case.

1.3 Results

Stroke Case Studies Testing (Study I)

The connectograms generated with SPIDER-NET to explore the DTI and CSD-based connectivity between gray matter parcels intersecting the lesions of patients with stroke injury and the whole brain are shown in Figure 5B. Upon visual inspection, the sub-network connectivity pattern of both patients with stroke injury looks altered compared to

HCs. In addition, it is worth noting that differences can be qualitatively observed both in the right hemisphere (where the stroke lesions are present) and in the contralateral one. Case 1 displays a less dense right hemisphere connectivity pattern compared to Case 2. At visual inspection, the connectograms derived from DTI and CSD processing generally preserve the same connectivity patterns, highlighting the difference between Case 1 and Case 2, besides denser results from CSD.

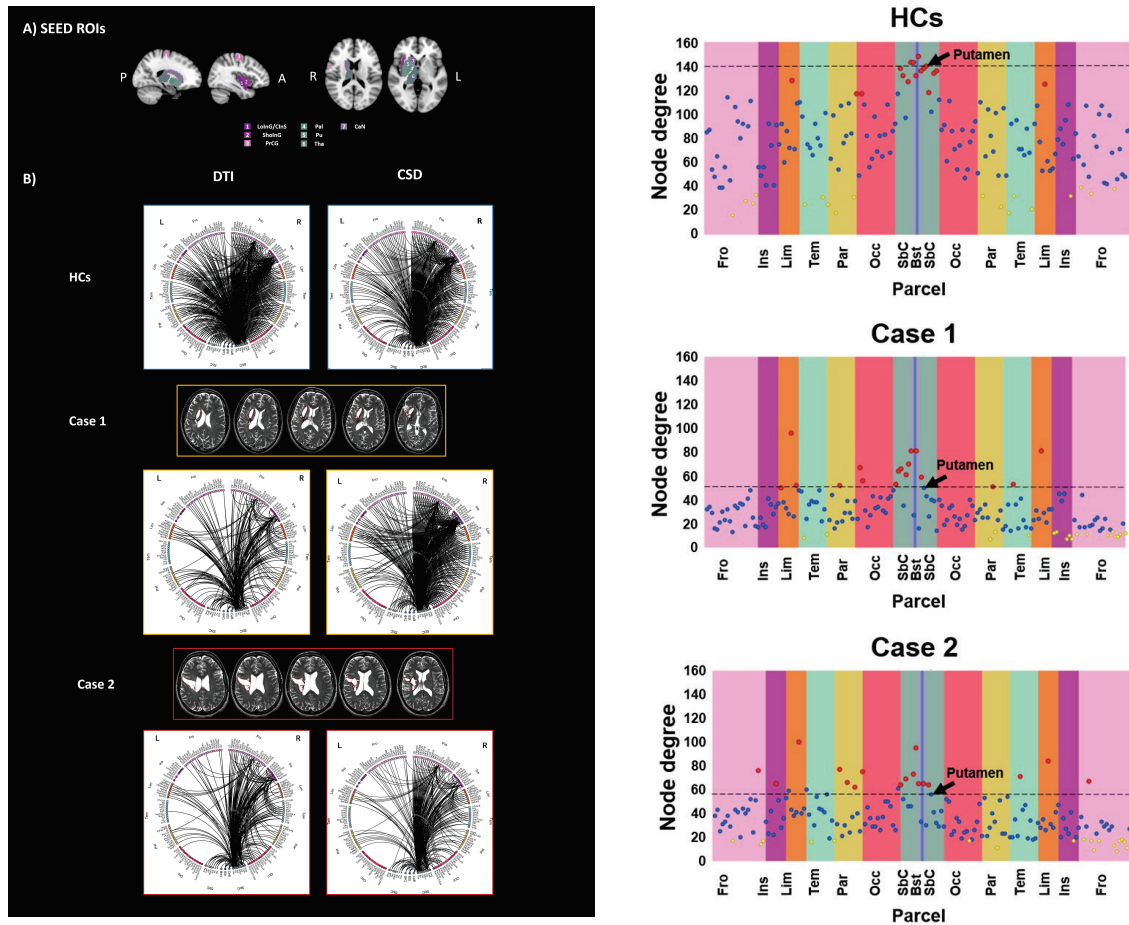


Figure 5: Investigation of connectograms and local topological properties of HCs, Case 1, Case 2. On the left, the connectograms (B) derived by both DTI (left) and CSD (right) processing, of the sub-network extracted using as seeds all parcels which are overlapped with the stroke lesion of either Case 1 or Case 2, are reported for HCs (top panel), Case 1 (middle panel) and Case 2 (bottom panel). Specifically, seeds were defined according to lesions and all the parcels of the brain were considered as targets for connectivity analysis. On the right, local node degree computed for HCs, Case 1 and Case 2. The X-axis represents the 165 brain parcels, even if only lobe labels are reported (e.g., Fro). The left hemisphere is represented in the left half of the graph, while the right hemisphere is represented in the right half. Vertical colored stripes represent different lobes (e.g., the frontal lobe is represented in pink). The same lobe in left and right hemispheres is shown with the same color. Each dot in the graph represents the local node degree of a brain parcel. The 17 parcels (10% of 165) exhibiting the highest local node degree are represented as red dots. The 17 parcels (10% of 165) with the lowest local node degree are represented as blue dots. All the other parcels are represented as blue dots. The right Putamen is highlighted by an arrow. L-left hemisphere, R-right hemisphere, Fro-frontal, Ins-insular, Tem-temporal, Par-parietal, Occ-occipital, Sbc-subcortical, CeB-cerebellum, Bst-brainstem, PrCG-precentral gyrus, LoInG/CInS-long insular gyrus and central insular sulcus, ShoInG-short insular gyri, Pal-pallidum, Pu-putamen, CaN-caudate nucleus, Tha-thalamus.

Regarding the local topological properties, patients with stroke injury presented lower node degrees when compared with HCs. In HCs, regions with the highest node degrees, represented as red dots in Figure 5 (right panel), were mostly located in the dark-green vertical stripe, representing subcortical regions. Conversely, in the patients with stroke injury, both characterized by a right hemisphere lesion with prevalent subcortical expression, the regions showing the highest local node degree are more distributed across the cortical lobes. In addition, caudate nucleus, pallidum, putamen, and thalamus, which were classified by SPIDER-NET as regions with the highest node degree in HCs (red dots), were not classified as nodes with high node degree in both Case 1 and Case 2. In Figure 5, putamen node degree values are highlighted for HCs, Case 1 and Case 2. Regarding the global topological properties, both Case 1 and Case 2 presented differences for all the indexes when compared to HCs. Percentage differences ranged from 10.5 to 96.9% for Case 1, and from 14.3 to 81.9% for Case 2.

Robust Multi-Level Assessment of Schizophrenia (Study II)

Results from the comparison between HC and SZ of the overall indexes using nonparametric test, bootstrap hypothesis testing and of the robustness statistical tests are summarized in Table 2. Differences between MW and BOOT are highlighted. Specifically, higher p-values in BOOT are found for the weighted characteristic path length and modularity indexes. The latter highlighted a statistically significant ($p < 0.05$) difference between HC and SZ modularity that was not confirmed by using BOOT. The robustness assessment highlighted more stable results for BOOT. Indeed, it is worth noting that the mean of the p-values, considering all indexes and removing both one and two subjects, remains closer with BOOT than the MW test in almost all indexes considered. Furthermore, the variability was assessed, resulting in lower standard deviation employing BOOT with respect to MW in almost all indexes.

| Graph Index | Whole Comparison | | | | RST1 (#Sub = 1) | | RST2 (#Sub = 2) | |
|------------------|--------------------|--------------------|-----------------|-------------------|----------------------------|------------------------------|----------------------------|------------------------------|
| | HC data (mean±std) | SZ data (mean±std) | P _{MW} | P _{BOOT} | P _{MW} (mean±std) | P _{BOOT} (mean±std) | P _{MW} (mean±std) | P _{BOOT} (mean±std) |
| Degree (Density) | 71.746 ± 2.178 | 69.207 ± 5.149 | 0.130 | 0.086 | 0.146 ± 0.05 | 0.089 ± 0.044 | 0.162 ± 0.079 | 0.103 ± 0.063 |
| Strength | 37.053 ± 7.362 | 31.680 ± 10.246 | 0.124 | 0.079 | 0.143 ± 0.047 | 0.096 ± 0.029 | 0.155 ± 0.071 | 0.108 ± 0.045 |
| Bin CC | 0.986 ± 0.025 | 0.960 ± 0.057 | 0.150 | 0.124 | 0.150 ± 0.052 | 0.130 ± 0.038 | 0.184 ± 0.089 | 0.138 ± 0.057 |
| Wei CC | 0.489 ± 0.105 | 0.415 ± 0.144 | 0.150 | 0.105 | 0.155 ± 0.052 | 0.119 ± 0.027 | 0.183 ± 0.082 | 0.128 ± 0.050 |
| Bin CPL | 1.017 ± 0.030 | 1.052 ± 0.071 | 0.130 | 0.084 | 0.153 ± 0.052 | 0.089 ± 0.044 | 0.162 ± 0.079 | 0.102 ± 0.063 |
| Wei CPL | 2.216 ± 0.610 | 2.618 ± 0.892 | 0.124 | 0.129 | 0.151 ± 0.061 | 0.150 ± 0.054 | 0.155 ± 0.072 | 0.159 ± 0.072 |
| Bin Eff | 0.991 ± 0.015 | 0.974 ± 0.035 | 0.130 | 0.086 | 0.150 ± 0.052 | 0.089 ± 0.044 | 0.162 ± 0.079 | 0.102 ± 0.063 |
| Wei Eff | 0.525 ± 0.085 | 0.466 ± 0.113 | 0.113 | 0.078 | 0.147 ± 0.051 | 0.099 ± 0.023 | 0.143 ± 0.065 | 0.111 ± 0.041 |
| Modularity | 0.004 ± 0.009 | 0.016 ± 0.026 | 0.043 | 0.133 | 0.053 ± 0.018 | 0.128 ± 0.019 | 0.059 ± 0.030 | 0.144 ± 0.036 |
| Coreness | 0.018 ± 0.022 | 0.037 ± 0.036 | 0.164 | 0.122 | 0.141 ± 0.060 | 0.122 ± 0.039 | 0.199 ± 0.091 | 0.138 ± 0.061 |
| Small-Worldness | 1.003 ± 0.005 | 1.015 ± 0.019 | 0.178 | 0.085 | 0.146 ± 0.062 | 0.09 ± 0.013 | 0.214 ± 0.098 | 0.096 ± 0.030 |

Table 2: Results of the statistical tests performed on global indexes. The global version of the degree represents the mean of all edges associated to all nodes, whereas the density is the ratio between the number of present edges and the number of possible edges in the network. For this reason, the result coincides, and it is reported only once. HC: Healthy Controls; SZ: Schizophrenic Patients; RST: Robust Statistical Test; MW: Mann Whitney Test; BOOT: Bootstrap Hypothesis Testing; #Sub: number of subjects removed; CC: Clustering Coefficient; CPL: Characteristic Path Length; Eff: Efficiency.

Afterwards, we analyzed the local values of the indexes in the DMN and across the lobes, showing the distributions of the values of the degree across the same group in Figure 6. Differences between DMN regions and the other areas in SZ and HC, because of the different influence of the negative connections, were highlighted. Specifically, the most variable indexes resulted to be located within the DMN for the SZ group. On the other hand, they are distributed more across all the brain regions in HC group, even resulting more variable in the areas not included in the DMN. More precisely, one of the most variable values of the degree in both groups was the one assumed by the amygdala, located at the end of the first yellow box in Figure 6, representing the left hemisphere nodes of the DMN. Afterwards, the strength was analyzed, and the resulting distributions are shown in Figure 6. More specifically, the most variable local indexes among their populations were extracted from frontal, parietal lobes and subcortical structures. Furthermore, the

distributions of the strengths highlighted statistically significant differences ($p < 0.05$) found through both MW and BOOT or by only one of the two. Most of the differences were found in regions of the frontal lobe of both hemispheres. Finally, anticorrelations analysis through the creation of probabilistic averaged networks in whole-brain and DMN (shown in Figure 6) highlighted more negative connections in SZ than HC and less modular structure (3 communities found in SZ, 4 in HC). Main great differences were found in the frontal and parietal lobes, resulting to have reduced inter-hemispheric connectivity both in number of edges and weights and different configurations of the communities found (more details on community analysis are reported in article III).

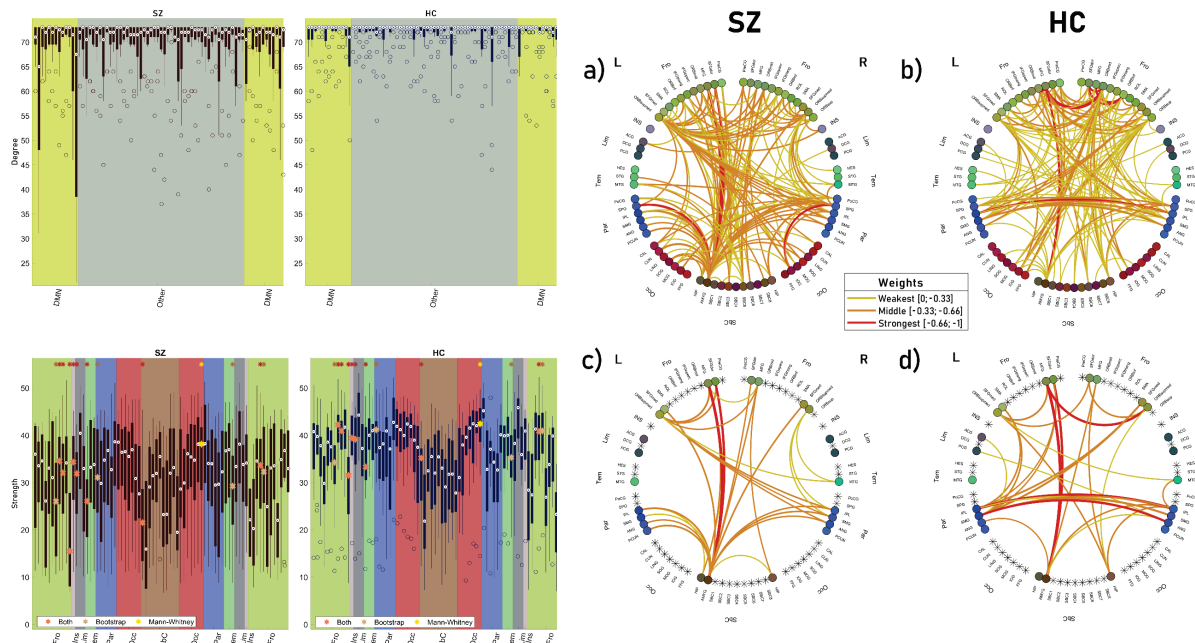


Figure 6: Distributions of the local topological properties and connectograms showing anti-correlations. On the left top, distributions of the local degree values in the two populations divided according to nodes of DMN in left and right hemisphere (yellow) or not (gray) are shown. On the left bottom, distributions of the local strength values in the two populations. Strength of nodes which are statistically significant different ($p < 0.05$) between the two populations are shown in different colors according to the different methods used: pale brown for bootstrap hypothesis testing, yellow for Mann-Whitney and red for statistically significantly different nodes found by both. On the right, connectograms of negative average group networks. a) and b) are the results on the whole-brain, c) and d) on only the DMN of SZ and HC groups respectively. 5% density thresholding is applied on the shown connectograms for graphical clarity only. The order and colors of the lobes is the same as the rectangles on the left bottom.

Interpretability of AD Classification DL Models Using Multiple Brain Measures (Study III)

Table 3 summarizes the outcomes of the 10-fold cross-validation performed on the DL models ResNet18, which used 3D T1-weighted volumes, and BC-GCN-SE, which used SC data.

| Cross-Validation Normalized Confusion Matrix | | Predicted | |
|--|---|--------------------------|-------------------------|
| ResNet18 | | P | N |
| Actual | P | 0.8167 [0.7734, 0.8462] | 0.1833 [0.1539, 0.2266] |
| | N | 0.1835 [0.1674, 0.2332] | 0.8165 [0.7668, 0.8326] |
| BC-GCN-SE | | P | N |
| Actual | P | 0.7033 [0.6719, 0.7692] | 0.2967 [0.2308, 0.3281] |
| | N | 0.2614 [0.24189, 0.3023] | 0.7386 [0.6977, 0.7581] |

Table 3: Normalized confusion matrix for the ResNet18 and BC-GCN-SE model performance, reported as median value and interquartile range.

Both models achieved good performance, with slightly superior performance of ResNet18. Regarding XAI, RVs of 70 and 46 parcels of AD in contrast to HC for ResNet18 and BC-GCN-SE, respectively, were determined to be statistically significant ($p < 0.05$) among all the parcels of the atlas employed. 22 parcels, in particular, were in agreement between the two approaches, mainly coming from the cortex. More specifically, in the case of ResNet18, 7 out of 8 total target MTL parcels were shown to be significant. In the case of BC-GCN-SE, instead, 12 out of 17 target DMN parcels were found to be statistically significant ($p < 0.05$) in at least one hemisphere (70.59%). Considering lateralization, the number was 16 out of 31 total DMN parcels (51.61%). Afterwards, the 20 most relevant parcels for the AD classification from both models were analyzed. First, it was noted that 11 out the 20 most relevant parcels for ResNet18 were also found to be statistically significant. The anterior division of both right and left parahippocampal gyrus from target MTL were among these 11 parcels. Considering the remaining 9, 7 were nonetheless found to be among the most relevant for the HC classification. Second, 14 out of the 20 most relevant parcels for BC-GCN-SE showed a significant AD vs. HC contrast. Of these 14, 5 parcels belong to target DMN. Among the remaining 6 most relevant parcels 4 were found

to be among the most relevant ones for the classification of HC. To summarize, the diagrams indicating the mean RV of all correctly classified sessions of both classes and models per each parcel and divided in lobes are shown in Figure 7.

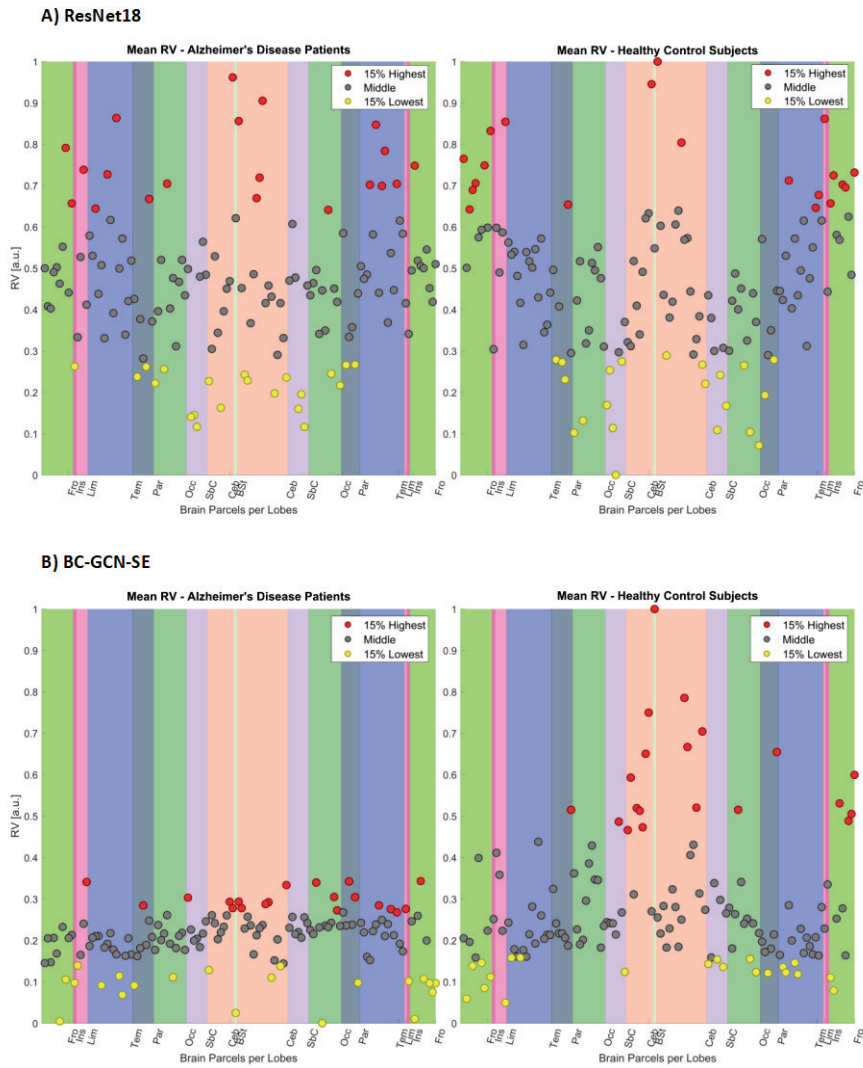


Figure 7: Mean RV of all parcels, for the ResNet18 (panel A) and BC-GCN-SE (panel B) models. The AD and HC classes are reported separately. The lobe belonging is indicated through rectangle colors. The RV of each parcel is labeled though dot colors according to the following criterion: red indicates the 15% of parcels characterized by highest RV, yellow indicates the 15% of parcels characterized by lowest RV, grey indicates the remaining parcels. Plots were created using the SPIDER-NET tool. Fro: Frontal Lobe; Ins: Insular Cortex; Lim: Limbic Lobe; Tem: Temporal Lobe; Par: Parietal Lobe; Occ: Occipital Lobe; Sbc: Subcortical Structures; Ceb: Cerebellum; BS: Brainstem.

1.4 Discussion

In the present PhD work, the aims of enhancing the reliability and the interpretability of the brain biomarkers were addressed. More specifically, this investigation was carried out on three different datasets formed of different number of subjects and patients, affected by diverse pathologies sharing the characteristics of both affecting specific regions of the brain and leading to diffused degeneration of the whole brain.

Network and Sub-Network Investigation through SPIDER-NET

First, a flexible and user-friendly tool called SPIDER-NET was proposed to both qualitatively and quantitatively investigate brain connectivity in health and disease. More precisely, SPIDER-NET was validated resulting an effective tool to represent the expected (dis)connectivity pattern due to a stroke lesion, in testing a-priori hypothesis by extracting a sub-network of interest and in computing topological indexes. The main findings found from the explorations on the two stroke case studies through SPIDER-NET were first that the pattern of connectomics of case 1 is significantly different from case 2, mirroring the greater clinical severity of the former. Second, the pattern of disconnection involved both the right hemisphere, where the stroke lesions were present, and the contralateral one. Third, the quantitative results mirrored the differences qualitatively observed with connectograms, providing a comprehensive description of brain connectivity highlighting the impact of graph-based metrics in line with the literature [8], [26]. Both local and global metrics derived from the whole-brain networks of Case 1 and Case 2 differed from HC one, as expected [8], [26], [27]. Fourth, the impairment of the cortical areas of interest determined a decrease in both short-range and long-range connections within the hemisphere ipsilateral to the stroke lesion. Fifth, interhemispheric connectivity was severely hampered, probably because subcortical nuclei (such as the putamen), which are integration hubs of extrapyramidal systems, were affected by the lesions. Finally, DTI-based and CSD-based sub-network connectograms presented comparable connectivity patterns, highlighting that valuable information is provided by both the processing techniques. CSD processing pipeline yielded to reconstruct denser connectograms, in agreement to the known CSD ability to better deal with the problem of the crossing fibers

when compared to DTI [28]. This is in line with differences between DTI and CSD that were observed in terms of interhemispheric connections, that were particularly evident for Case 1. Hence, it was shown that the SPIDER-NET tool can be helpful in clinical research settings, in tailoring the rehabilitative treatments and can greatly contribute to improve the understanding of the brain networks through a better interpretation and reliability of the measures. Normally, networks are composed of hundreds of nodes and thousands of links, thus the analysis and the interpretation are often tough. Automatic and interactive tools to assess brain network would be of great support in research and potentially in clinics. Indeed, analyzing sub-networks of interest, from a-priori hypotheses or through other tools, is of great importance in the understanding of the mechanisms underlying the cognitive processes or the circuit mapping of the brain. In addition, generating SPIDER-NET connectograms could be a good general strategy to test the robustness of the processing pipeline, including the connectivity metrics, further conditioning (e.g., thresholding or binarization), and global or local graph indices. As one of the main limitations of connectomics, so far, is the lack of standardized procedures for network construction and edge weighting [29], [30], SPIDER-NET may be applied as a flexible and easy tool for calibrating connectomics analyses. Specifically, it could allow to quickly identify the expected pattern of disconnection and to easily highlight major errors if present. This quality check may offer a benchmark before addressing less trivial connectivity alterations, as the ones induced by diffused neurodegeneration.

Improving the Robustness of the Multi-Level Connectivity in Group Studies

As the uncertainty assessment, a bootstrapped top-down approach to assess reliable abnormalities found on different levels was proposed. In particular, the assessment from global indexes to single edges (brain activations and deactivations) was carried out on the FC of a population characterized by SZ, thought to cause both diffused and focalized dysconnectivity patterns. The multi-level analysis resulted to potentially favor more robust results in contexts where the focus is not well-known and statistical tests can be easily biased by uncertainty. Increased global index values were found in HC for all segregation properties and efficiency, whereas path length was longer in SZ, as expected, as well as small-worldness and coreness statistics. Discrepant significance of modularity was found between MW ($p=0.043$) and BOOT ($p=0.133$). The robustness assessment tests (both

RST1 and RST2) highlighted more stable results for BOOT than the direct testing. Significant results were also found at lower levels. The analysis of the DMN highlighted a higher variability, reduced connectivity and strength and increased deactivation in the SZ group. At local level, 13 areas were found to be significantly different ($p < 0.05$) in the groups, highlighting a greater divergence in the frontal lobe. The community detection analysis of the anticorrelations shows as, in SZ, most nodes from frontal, limbic and temporal lobes are clustered together in a community. This deactivation organization may be relevant in the study of failed activity inhibition of SZ characterized by selective disruption of an automatic inhibitory process, and failure to limit the current contents of consciousness [31], [32]. In the parietal lobes configuration other differences were noticed, especially regarding the clustering with contralateral occipital lobe in SZ. In other studies [33], [34], these lobes were investigated in relation to SZ highlighting important functions. For example, it was hypothesized that cognitive deficits and delusions may be related to malfunctions in the parietal lobe [35] or that the maintenance of visuospatial information is associated to a network of occipital cortex regions [33]. These results were confirmed by analyzing the mean matrix of single negative edges of the groups in the whole-brain and DMN, suggesting an inverted connectivity among prefronto-temporal areas which can reflect abnormal inhibition of the regions' activity involved. In general, the results highlighted a higher stability and robustness of the BOOT in RST1 and RST2, providing support to its possible use when dealing with a small dataset which is also affected by uncertainty of the measures. Furthermore, the results highlighted a trend towards results obtained with a greater number of subjects and data, and appeared to be in line with other studies where the major finding in SZ is that the connectivity is not systematically increased/reduced but generally different, due to the functional reorganization of SZ [36], [37]. However, it remains unclear whether SZ abnormalities are the result of a localized dysconnection exerting widespread effects throughout the brain, or a whole-brain dysfunction that affects certain regions more. The robust approach for assessing the differences in the network topology turned out to be of great interest, especially when small datasets are available. Furthermore, the possibility of performing automatic top-down investigations may improve our understanding of both diffuse and localized dysconnections typical of some pathologies.

Comparing and Combining Explainable AI

As the interpretability issue, the comparison between two XAI methods trained on different types of structural MRI data, namely 3D T1-weighted scans and SC matrixes extracted from DTI was performed. Overall, the DL models highlighted good classification results. On the one hand, the BC-GCN-SE obtained acceptable performance with respect to the existing literature [38]. On the other hand, the ResNet18 model achieved slightly superior accuracies with respect to the BC-GCN-SE when classifying both AD and HC subjects (TPR_{median} = 0.8167; TNR_{median} = 0.8165). The obtained results were in line with the existing literature [39]. To the best of our knowledge, no other studies compared DL models trained using 3D T1-weighted volumes and SC data neither in terms of accuracy nor explainability. Moreover, this is the first work testing BC-GCN-SE model on SC with the aim of classifying diverse AD subjects of the OASIS-3 dataset. These models are relatively recent, especially when compared to more mature Convolutional Neural Network (CNN) architectures working on images such as the ResNet18 model. This latter is a widely tested pre-trained model with weights obtained from more than a million images, probably resulting in a better ability to generalize. In addition, it is worth noting that the SC data obtained from DTI have those above-mentioned inherent limitations related to the processing pipelines that may result in noisy connections [30]. Given this premises, both results can be considered promising, even considering the great variability of severities of the AD subjects, sessions parameters, the presence of multiple acquisition session without a predefined design setting in the OASIS-3 dataset. Moreover, the division into classes according to the CDR results in 15 subjects having sessions with different labels and in a majority of sessions from subjects with very mild impairment (n. 97 sessions with CDR 0.5 out of 135 total AD sessions). Clearly, being the range of dementia effects wide and of difficult definition, it results in a more difficult task. The use of DL models on brain connectivity graphs and their assessment in comparison to more established approaches would be of great importance. The processing effort of extracting connectivity data, providing relevant data, would be thus enhanced if casting new light on the exploration of brain biomarkers and computational cost in parallel to optimal results of AI methods. As the XAI analysis, both models highlighted advantages and drawbacks. The statistical tests and the ranking of the most relevant parcels highlighted a good degree of agreement with respect to domain knowledge. Thus, ResNet18 model indicated a good

portion of the MTL and BC-GCN-SE model showed several parcels composing the DMN. On the one hand, 7/8 parcels of the MTL were highlighted by the statistical test ($p < 0.05$) with the specific case of the parahippocampal gyri of the two hemispheres also being among the 20 most relevant parcels for AD identification with 3D T1-weighted volumes. 70.59% of the DMN parcels were instead found to be significant ($p < 0.05$), with 5 parcels being among the most 14 explainable and significant parcels using SC data. However, some limitations need to be considered. One may argue that the number of parcels displaying a difference by the solely statistical test in the AD and HC levels of RV is high when compared to the total number of parcels we investigated. This may increase the possibility to detect the target structures. The investigation of which parcels – among the significant ones – were also included in the 20 most relevant (i.e., highest RV values) for the AD case allowed to reinforce the interpretation of the results. This served as a strong confirmation that the algorithm is leveraging them to identify the presence of AD. In addition, it is worth noting that most of the parcels identified were cortical. The population was made up of subjects at various AD stages and severities that could simply suggest different involvement of the disease in different MRI sessions lead to general extensive importance of a considerable portion of the cortex. This is particularly true for the ResNet18 model, that points out more significant parcels. In addition, it is a deeper model with respect to BC-GCN-SE, possibly focusing on more complex features. As the BC-GCN-SE model, instead, a limitation can be given by differences found in DMN across the hemispheres. However, we acknowledge that there are considerable discussions and increasing evidence among researchers as to abnormality of topological asymmetry between hemispheric brain White Matter (WM) in AD and Mild Cognitive Impairment (MCI) [40], [41]. More specifically, all parcels which were found to exhibit asymmetry in the study by Yang and colleagues agreed to the parcels of only one hemisphere used by our model [40]. This can be a valuable indication supporting for hemispheric lateralization and aberration possibly due to the long-range connection loss that might be further investigated. In general, valuable conclusions can be reached from the investigation of the two models. First, it was found that some DMN regions have good importance in the classification from SC matrixes. Second, it was also shown as the regions characterizing the MTL, found as replicated hallmark in structural MRI studies, contain inherent important features for the ResNet18 model employing 3D T1-weighted volumes. In this

context, it is also of great interest to underline how some important regions of the MTL such as amygdala, parahippocampal gyri, and hippocampus that were also found to be different in AD connectivity data [42], were not underlined by the XAI analysis of BC-GCN-SE. Indeed, apart from right amygdala, that was found to be statistically significant, all the other regions of the MTL did not result from either statistical test or examining most relevant parcels. This finding may highlight a limitation of BC-GCN-SE model interpretability or the effect of some noise sources inherently present in the connectivity data. At the same time, this complementary relevance of parcels of interest appears promising in the perspective of developing superior and more trustworthy models. Indeed, the use of well-known hallmarks from multiple measures may offer the opportunity of focusing on different information that would be of great interest and significance if used concurrently. In this context, only few studies considered the combination of morphological features of regions from 3D T1-weighted volumes and interregional properties obtained through structural connectivity data, but that may potentially lead to more accurate results and to a better interpretation [43]. Nowadays, there is indeed the potential to easily collect multiple data from multiple modalities, and in this sense, efforts to assure even greater use of the whole potentiality of DL models, exploiting their peculiarities would be of unvaluable interest for diagnosis and rehabilitation of the pathology. The development of better and more interpretable models can represent accurate and robust solutions to the well-known problem of trust in “black-box models”, which limited their diffusion within real settings so far.

2 Cover Essay

2.1 Motivation: Localized and Widespread Degeneration

The human brain is an incredibly complex network which boasts a highly intricate and dense neuronal architecture that is considered one of the most sophisticated systems in nature [1]. Comprising billions of neurons linked by fibers and synapses, it is organized across multiple spatial scales and functionally interacts over multiple temporal scales [2]. In this huge scenario of neurosciences, brain imaging is yielding a wealth of information relevant to the structure and function of the brain, though limited to the macroscale. At a such level, different brain regions, typically referring to patches of GM that are spatially adjacent and functionally coherent, can be thought of as connected through fiber bundles that reflect the anatomical structure or through simultaneous functional activations during rest or while performing a task.

Moreover, the segregation/integration paradigm plays an important role. It is indeed well-known that specific brain areas and circuits are specialized for some functions (segregation) such as motor, language, somatosensory, speech etc. and their cooperation is essential for the joint processing of information from diverse sources allowing complex cognitive functions (integration) [44]. Different cognitive processes set different demands on locally segregated and globally integrated brain activity. However, it remains an open question how the brain configures its functional organization to balance the demands on network segregation and integration to best serve cognition, as well as the connection to the structural mapping.

Therefore, the analysis of the brain organization at different levels can be of great interest to both improve the understanding of its functioning in health and investigate possible changes in disease. In addition, a possible multi-level analysis can be helpful in the

investigation of the relations between focused deficit and global abnormalities. Different brain and mental disorders such as stroke, AD, SZ and many other pathologies, indeed, often present with both localized and widespread degeneration, heavily affecting the functioning of specific anatomical areas or functional circuits, but also leading to global changes. The stroke is a prime example, since it manifests as significant damage to specific areas of the brain tissue, but also results in a strong impact on the entire nervous system. However, in many other cases such as SZ, it remains unclear whether specific disruptions in frontal lobe or in the DMN, found in different studies [36], [45], [46], exert widespread effects throughout the brain, or a whole brain dysfunction affects certain regions more [47]. For these reasons, it is of great interest to investigate the correlates between the changes at the different levels, which are not fully explained yet, in the perspective for example of tailoring rehabilitative treatments.

The diverse affection of the pathological conditions results in significant changes of the brain connectivity, which is known to play a crucial role in various cognitive and behavioral functions. In general, brain connectivity or connectomics refers to the complex network of connections between units of the brain which determine how it elaborates information and functions. As all disciplines commonly referred to as “-omics”, connectomics is characterized by the integration of large amounts of data to map a complex system, which can be formalized using graph theory. Graph theory is a powerful mathematical approach which was effectively applied in many scientific fields for understanding and modeling structures characterized by a specific architecture and topology. A graph (network) is a mathematical representation of a system composed of vertices (nodes) and edges (links), describing any kind of relationship between pairs of vertices. In the case of human connectome, the brain can be modeled as a network on the different scales with the goal of mapping its structure to function and behavior. Connectomics reconfigures the study of brain structure and function by outlining the entire nervous system in terms of neural units and their connections, but also in the macro scale perspective we are addressing. Indeed, the concepts represented by nodes and edges of a brain network can vary depending on the technique or imaging modality used to extract connectivity data. In the context of this PhD project, macro-scale SC from diffusion Magnetic Resonance Imaging (dMRI) and FC from fMRI were considered.

In this representation of the brain at a macroscopic level, the nodes typically represent GM parcels defined by well-known atlases [11], [13], [16], [48]. These atlases segment the brain into sets of regions based on anatomical or functional criteria. The edges, representing the relationship between nodes, can also depict either SC or FC. The former refers to the anatomical associations given by WM pathways between brain regions, while the latter represents the temporal correlations between the signals produced by pairs of brain regions. These measures can be quantified using various indices depending on the imaging modality employed. For example, the number of streamlines from deterministic WM tractography can be used as weights in an dMRI-derived structural network, while correlation between BOLD time series can be used to define edges of fMRI-derived functional network. Networks can be represented as n -by- n association matrixes where n is the number of nodes of the network, and each element (e_{ij}) represents the link connecting the nodes i and j .

The quantitative examination of brain connectivity patterns can provide insight on both healthy and pathological states for figuring out how the brain works and processes information. Typical graph-based metrics can be extracted from human connectomes highlighting segregation and integration dysfunctions and abnormalities. Analyzing the changes in the organization and architecture of the networks in health and disease can highlight brain connectivity biomarkers providing unvaluable insights on the underlying mechanisms of the pathologies. For example, important abnormalities have already been found in a number of studies revealing changes of topological properties of different brain disorders, such as stroke [8], schizophrenia [9], Alzheimer's disease [6], mild cognitive impairment [5], Parkinson's disease [4], epilepsy [10] and autism [3]. These findings may represent an important asset to support correct diagnoses and helping to develop effective treatments. However, studying brain connectivity can be challenging due to the inherent uncertainty and complexity of the data.

Indeed, despite the large evidence produced in the last years, gold-standard methodologies to process the data, standardized procedures to extract correct connectivity measures, edge weighting and, thus, formalize the brain as a graph, are still lacking. In fact, inherent sources of error persist in producing inaccurate or missing connections. For example, the deterministic approach of streamline tracking from DTI, which allows to produce SC measures and data, have still limitations. dMRI does not account for either systematic or

stochastic errors in predicting fiber orientation. Systematic errors arise from complex fiber geometries such as crossing, twisting, bending and kissing fibers, which cannot be adequately captured using a tensor model and the principal eigenvector orientation. Additionally, the technique of dMRI is inherently noisy, and susceptibility gradients, head motion, and eddy currents can lead to artifacts. Also, FC analysis represents a powerful tool for understanding brain organization, but it is subject to uncertainty due to the limitations of fMRI techniques and the dynamic nature of brain activity. The Pearson correlation coefficient, a commonly used metric for FC, can change over time [49]–[51], and this variability must be taken into account. Furthermore, FC is affected by factors such as low signal-to-noise ratio, non-neural noise and hardware instability, and there are still controversies related to the definition and interpretation of the results. Hence, both SC and FC studies can be often biased by a huge number of factors, leading to the false presence/absence of edges and inaccurate weights thus resulting in connectivity data affected by uncertainty. In this context, changes and abnormalities which can be found in disease by analyzing the topological measures of the networks must be considered cautiously and in relation to their degree of robustness. The term robustness in network theory refers to the ability to withstand perturbations and failures. Hereafter, it will be referred to the withstanding of the connectivity networks results regardless possible absence/presence of edges or issues in edge-weighting due to biasing factors such as those ones related to the processing pipelines or the limited amount of data. More robust detection of abnormalities leads to more reliable results, which can be more widely usable. Furthermore, the resulting data are complex networks composed of hundreds of nodes and thousands of edges which are difficult to be interpreted. The general concept of interpretability refers to the possibility of determining the salience of network components. Hereafter, this term will be used to describe the objective of establishing a relationship between a condition (such as a pathology) and its impact on the connectivity data, or vice versa, to improve their understanding. In fact, despite the ability of entire brain connectivity matrixes to exhaustively and quantitatively depict the human connectome, this representation does not always provide a clear and intuitive visualization and investigation of the connectivity pattern. The matrixes are often too large to be easily analyzed, which can hide important information. The interpretation of brain connectivity is particularly important for exploratory analyses, with the goal of identifying distinctive patterns that

may differentiate pathological conditions from physiological ones, identify circuit or sub-network of interest, or to evaluate changes following a pharmacological treatment or rehabilitation. As a consequence, together with the difficulty of analyzing data and their representation, extracted brain connectivity biomarkers can also be tricky to be interpreted. Also, latest methods for classification from DL approaches such as CNN or Graph Neural Networks (GNN) [52], [53], and their underlying processes, were only partially interpreted. This implies that, although these complex methods allow for correct classification and recognition of the pathologies, the understanding of the reasons behind these choices is often tough. This interpretability issue strongly limited their trust within clinical settings. Recently, novel methods were developed with the objective of enhancing our understanding of the representation and significance of model features, as well as their contribution to the performance of the model. This concept is the so-called explainability, which allows to investigate the “explanations” behind the choice of the DL models and therefore to compare them to the domain knowledge with the final aim of having more reliable and accepted tools.

In light of these challenges, there is a need for methods that can accurately assess the uncertainty and interpretability of brain connectivity biomarkers. These aspects limited their diffusion within real settings and usability in clinical contexts. Such methods can help to better understand the underlying mechanisms of the disorders and ultimately lead to the development of more effective treatments. In the context of pathologies sharing the characteristics of presenting both focalized and widespread degeneration, the thesis addresses the open issues of uncertainty and interpretation through: I) the development of a novel tool for visualizing and studying brain connectivity that was validated on emorrhagic stroke patient cases, II) the investigation of the robustness of connectivity indexes together with a multi-level analysis in schizophrenic patients and III) the use of DL and explainability techniques in the context of AD classification task.

2.2 Background: Brain Networks from dMRI and fMRI

MRI is a non-invasive, safe and widely used imaging technique that uses a magnetic field and radio waves to produce detailed images of the brain. One of the main advantages of MRI is its high-resolution and high-contrast capabilities, which allows for a clear and detailed view of the brain. This high spatial resolution allows for accurate diagnosis and monitoring of a wide range of medical conditions. It has several other advantages such as multi-planar capability and ability to provide rich information about both structural and functional information of the human brain.

Structural MRI is used to create images of the brain's anatomy, while fMRI is used to create images that show how the brain's tissues and organs are functioning. On one hand, MRI is an important diagnostic tool in medical imaging as it can provide detailed images of internal organs, bones, and soft tissues without the use of ionizing radiation. On the other hand, fMRI provides images of brain activity through the measurement of changes in blood flow and oxygenation in response to neural activity, allowing to study the brain in action. Both techniques can be employed to extract valuable data in the study of the interactions between the different brain regions, which can be structurally connected or functionally activated/deactivated.

Hence, the objective of this chapter is to provide an overview of the most important methods and approaches employed to extract, analyze and interpret connectivity data from both structural and functional MRI, far from being exhaustive on the research topic of MRI. More specifically, the Section 2.2.1 is focused on the imaging-based parcellation of the human brain and the most widespread atlases, which were employed. Section 2.2.2 discusses methods for the extraction of macroscale connectivity, including both SC and FC measures for edge-weighting. Section 2.2.3 discusses the limitations of macroscale connectivity. Section 2.2.4 covers the methods for the analysis of complex networks, including graph theory-based measures for the extraction of topological properties of the networks and graph models. Section 2.2.5 discusses the research field of connectomics, which put together the MRI methodologies and graph theory to define and study the human connectome with the final aim of mapping structure and function of the brain in health and disease.

2.2.1 Imaging-Based Parcellation of the Human Brain

The development of dMRI and fMRI techniques over the past five decades led to a significant advancement in the field of human brain mapping, which is referred as the "golden age" of brain mapping. In order to study how different regions of the brain interact and work together, it is necessary to subdivide the brain into a set of distinct regions - i.e., parcellation of the brain. Therefore, different brain parcellation atlases representing reference maps based on an average of multiple individuals' brain scans were established.

It is worth noting that the first fundamental contributions can be traced back to the early 20th century, as it played a crucial role in the comprehension of the structural and functional subdivisions of the brain, specifically in the cerebral cortex. The foundational knowledge and achievements accomplished during this time laid the foundation for current advancements in brain mapping [54]. The German physician Korbinian Brodmann, who made substantial contribution to the field of neuroscience, was in the vanguard of that pioneering effort. In his seminal work [55], Brodmann proposed a system for subdividing the cerebral cortex into distinct regions, which were identified and designated by numerical labels, based on the characteristics of their cellular structure and laminar organization. To this day, Brodmann's system of regional designation continues to be widely utilized in the field of neurology.

Parallel to Brodmann's work, other brain mappers produced a number of early maps of the parcellated cerebral cortex based on either cytoarchitectonics or myeloarchitectonics utilizing postmortem, hand-drawn images [56]. This process of cortical parcellation pioneered by Brodmann and the other early brain mappers a century ago continues until the present time. In the pre-tomographic imaging era, progress in brain atlas development was made in several key areas. Examples were: the shift from a few maps of the brain to more clinically applicable brain atlases, including the whole brain and specific parts such as the cerebellum and brainstem, the use of multiple specimens to account for anatomic variability, and the integration of structure and function [57].

The development of computerized electronic brain atlases was a natural progression in brain atlas development, as it aims to overcome limitations of static contents and difficulty in mapping atlas content to an individual brain scan. These efforts were directed to several areas, including the direct digitalization of existing print atlases, their extension into 3D,

the creation of improved atlases through postprocessing, enhancements, and extensions, and the development of new electronic atlases.

Recent advancements in imaging, brain mapping, and computing led to the development of new human electronic brain atlases. Criteria such as parcellation, modality, plurality, quality, abnormality, lifespan, extendibility, ethnicity, spatial and temporal scales, integration, and techniques of creation can be used to identify and classify the various directions in the evolution of these atlases.

Nowadays, among all the brain connectivity parcellation atlases, some of the most widely used are the Desikan-Killiany atlas, the Destrieux atlas, the AAL atlas, and the HOA atlas [11], [13], [16]. These atlases are commonly used in neuroimaging research, and also employed in this PhD thesis, as well as in clinical applications such as diagnostics and treatment planning. More in general, based on different criteria and methods, these parcellations allow to map and compare brain structure and function across individuals, groups, and conditions. The Desikan-Killiany atlas and the Destrieux atlas, for example, are both based on the cytoarchitectonic features of the cerebral cortex, which are the histological characteristics of the different layers and cell types of the brain. The AAL atlas is based on the functional properties of 90 brain regions, as inferred from meta-analyses of neuroimaging data on different cognitive and behavioral tasks. The HOA atlas is based on the structural connectivity of the brain regions, as inferred from DTI data. It provides a high-resolution parcellation of the cerebral cortex and subcortical structures. The HOA atlas is shown in the Figure 8.

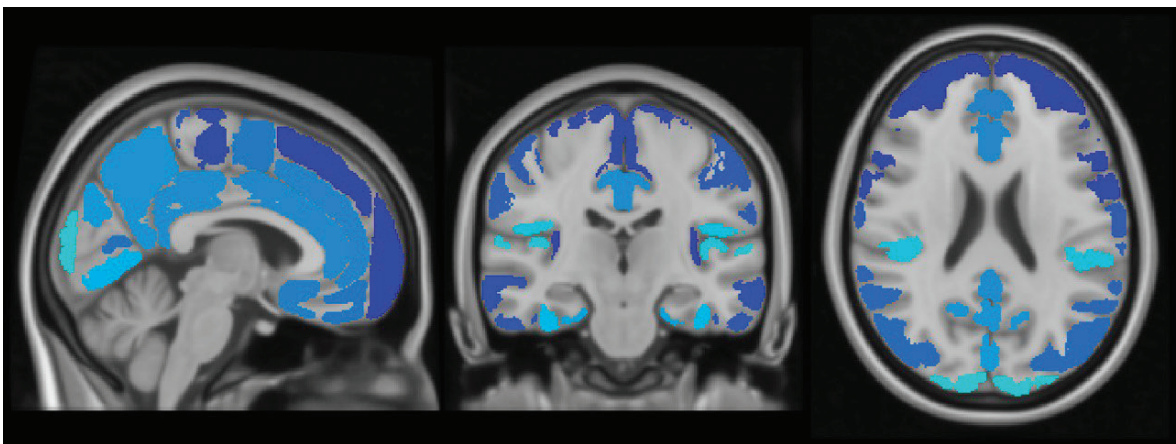


Figure 8: Sagittal, coronal and axial views of the Harvard-Oxford atlas.

Independently from the use of specific brain parcellations and the related recent innovations in brain parcellation, there remains the potential for different granularities by relating brain regions to anatomical or functional circuits. For example, specific parcels of the atlas can be grouped by large anatomical systems such as the hemispheres or the distinct lobes of the brain. Likewise, functional networks, like the Resting-State Networks (RSNs) can be identified and linked to specific parcels of the atlas (see the next Section 2.2.2.2). This division allows for a more comprehensive understanding of the brain's organization, putting together structure and function, and facilitating the identification and study of specific circuits within the brain.

2.2.2 Methods for the Extraction of Macroscale Connectivity

Once parcels are obtained, relations between them can be created. Different methods and approaches can be used to establish the connections, each with their own advantages and limitations, representing macroscale structural or functional characteristics. Given the breadth of this area of research and the aim of the work of examining the uncertainty of the measurement, attention will be given to the specific approaches related to the scope of this thesis.

2.2.2.1 Methods for the Extraction of Structural Connectivity

Primary method for studying macroscale SC of the brain is dMRI. Through this approach, the diffusion of water molecules in biological tissues is used to provide information about the microstructure of the tissue. The underlying principle is that the diffusion of water molecules in biological tissue is restricted by the presence of cell membranes, other microstructural features, and the geometry of the tissue. It is thus possible to obtain unique information about the microstructure of biological tissue that is not available from other imaging modalities.

In this context, Diffusion Weighted Imaging (DWI) is a technique that measures the apparent diffusion coefficient (ADC) of water molecules in tissue, which is a measure of the degree of diffusion restriction. By measuring the ADC, DWI can detect changes in the

microstructure of tissue, such as changes in cell density, the presence of fibrous structures, or the presence of a tumor. DTI is a technique based on DWI that uses multiple diffusion-weighted images acquired at different gradient directions to estimate the diffusion tensor, which characterizes the directionality and anisotropy of the diffusion. The diffusion tensor is a 3x3 matrix that describes the diffusion in the three spatial directions and can be visualized as an ellipsoid (diffusion ellipsoid), which represents the shape and orientation of the diffusion. Processing different parameters of the diffusion ellipsoid in each voxel, such as direction of greatest diffusivity, it is then possible to reconstruct streamlines, which estimate the trajectories of WM fibers in the brain. This post-processing technique is called fiber tractography [58].

2.2.2.1.1 DTI-Derived Indices for Edge-Weighting

The macroscale structural connections can be thus determined by different DTI-derived indices. Connections mirrors the anatomical structure, depicting WM pathways (or fiber bundles), or related characteristics. Once the DTI tensor D has been fully characterized, it is possible to calculate quantities that provide a description of the diffusion process for each individual voxel under consideration. The tensor D can be diagonalized into its three eigenvectors, v_1 , v_2 , and v_3 , and corresponding eigenvalues, λ_1 , λ_2 , and λ_3 , where λ_1 is the largest eigenvalue and λ_3 is the smallest. At this point, D can be represented by an ellipsoid with axes of length λ_1 , λ_2 , and λ_3 . This ellipsoid can resemble a sphere if all three eigenvalues are equal, which indicates isotropic diffusion. Conversely, in WM, the presence in most voxels of fibers with a common direction renders causes anisotropic diffusion. This is revealed by an elongated shape of the ellipsoid with one eigenvalue significantly larger than the others. Considering all the voxels, a complete fiber tractography map can be obtained as a set of lines, also known as streamlines, that represent the fibers in the brain. An illustration of this concept can be found in Figure 9.

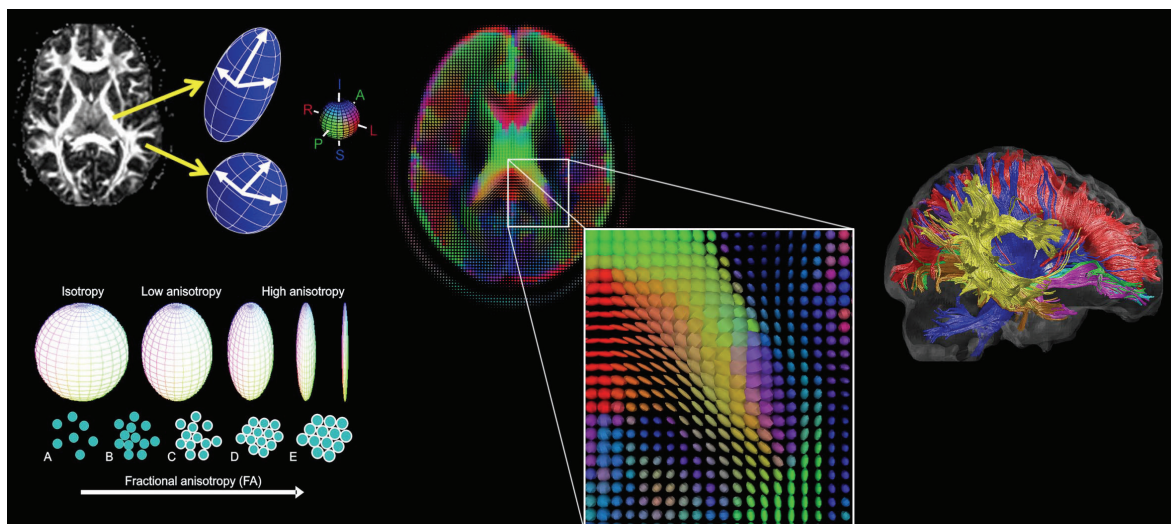


Figure 9: From tensor to tractography [59]. Examples of different levels of anisotropy resulting in a diverse shape of the diffusion ellipsoid, resulting DTI ellipsoids from axial slice of the brain, zoomed in one area, and the resulting tractography visualization are shown.

There are several DTI-derived metrics that are commonly used, including and number of reconstructed fibers (NF), mean diffusivity (MD), fractional anisotropy (FA), volume ratio (VR), radial diffusivity (RD), axial diffusivity (AxD).

NF is a simple metric that represents the number of reconstructed streamlines connecting two parcelled brain areas. This purely quantitative criterion can be combined with several qualitative indexes referred to the voxels crossed by the streamlines. MD is a measure of the movement of water molecules (i.e., the ADC) averaged in all directions. FA is a measure of the degree of anisotropy in the diffusion of water molecules within a voxel, ranging between 0 and 1, with higher values indicating greater anisotropy and therefore more ordered WM structure. VR is a metric that measures the ratio of the volume of the voxel to the volume of the ellipsoid that has the same eigenvalues and eigenvectors as the diffusion tensor of the voxel. RD is a measure of the diffusivity of water molecules perpendicular to the main eigenvector. AD is a measure of the diffusivity of water molecules parallel to the main axis of the diffusion tensor. The Table 4 summarizes these metrics and their formalization. Overall, the major qualitative index in SC is the value of FA averaged over the NF connecting two areas.

| DTI-metrics | Acronym | Formula |
|-----------------------|---------|--|
| Mean diffusivity | MD | $\frac{\lambda_1 + \lambda_2 + \lambda_3}{3}$ |
| Fractional anisotropy | FA | $\frac{\sqrt{\frac{1}{2} \sqrt{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}$ |
| Volume ratio | VR | $\frac{\lambda_1 \lambda_2 \lambda_3}{\left(\frac{\lambda_1 + \lambda_2 + \lambda_3}{3}\right)^3}$ |
| Axial diffusivity | AxD | λ_1 |
| Radial diffusivity | RD | $\frac{\lambda_2 + \lambda_3}{2}$ |

Table 4: Most widespread DTI metrics, with acronyms and formulation.

2.2.2.1.2 Beyond the Tensor Model

The tensor model, while simple in nature, has the limitation of not being able to identify multiple fiber populations present within a single voxel. Specifically, DTI models the diffusion of water molecules in the brain using a single tensor, which assumes a trivariate Gaussian distribution for the spatial displacements of these molecules. However, this model may not accurately capture the complex and heterogeneous structures found in the brain, such as the presence of multiple fibers crossing, diverging, or converging [28]. In these cases, a multicompartmental modeling approach may be more appropriate to accurately describe the diffusion of water molecules [60].

Various techniques were proposed to address the issue of fiber reconstruction in DTI [60], [61]. One of the most commonly used methods is High Angular Resolution Diffusion Imaging (HARDI), which utilizes a spherical sampling scheme and samples the diffusivity signal along a uniform set of (hundreds) directions on a sphere. This approach provides more information on complex diffusion processes than traditional DTI methods [60]. Among the different approaches for recovering fiber orientations using HARDI, one of the most efficient is based on spherical deconvolution [62]–[64]. This method models the HARDI signal as a convolution of the spatial orientation of fibers and a smoothing kernel, which describes the signal from a single fiber. The use of a deconvolution approach (such

as CSD) was shown to yield interesting results in terms of both high angular resolution and efficient computational time [63], [64].

In the context of this thesis, DTI techniques and measures were used. Furthermore, HARDI data were compared to those from DTI to analyze the capabilities of the latter in comparison to the former, which provides a more reliable benchmark and better ability in fibers reconstruction.

2.2.2.2 Methods for the Extraction of Functional Connectivity

FC at the macroscale can be measured using fMRI. It is thus possible to study the activity of the brain regions, their communication and interactions, and it can be used to identify networks of brain regions that are functionally connected. fMRI is a widely used non-invasive technique for investigating brain function by measuring changes in BOLD contrast. In general, functional imaging is based on the cerebrovascular response to local activation of brain areas. For this reason, it has the advantage of admitting non-invasive imaging as BOLD and provides a good spatial resolution. Conversely, severe limitations are found in time resolution since the slow hemodynamic response function which shows a time to peak of 4-5 seconds and a total duration of 10-12 seconds [65], [66].

The BOLD contrast is based on two principles: i) the magnetic properties of hemoglobin vary depending on its level of oxygenation (oxyhemoglobin is diamagnetic, while deoxyhemoglobin is paramagnetic) [67]; ii) blood oxygenation levels in brain regions vary in response to neural activity. These principles allow for the indirect measurement of brain activity [68]. Specifically, the activation of a cortical area leads to an increase in the inflow of oxygenated blood that exceeds the metabolic demand of the area. Hence, the BOLD signal provides a T2*-weighted contrast which has quite low changes with functional activity, i.e., 4-5% of the total contrast. Also, functional dynamics is mixed with other physiological hemodynamics related to the heartbeat, respiration and Mayer waves. For this reason, several preprocessing steps are required. Limiting to the main ones: 1. Spatial smoothing with kernels of 4-8 millimeters; 2. Time-filtering reducing the physiological hemodynamic components [69].

fMRI studies can be broadly divided into two approaches: task-induced and task-independent protocols. On the one hand, in standard task-induced fMRI, an experimental

task is presented alternately with a control condition, and the BOLD signal during the experimental task is compared to the BOLD signal during the control condition. On the other hand, the study of ongoing brain activity is expressed by structured BOLD fluctuations when subjects are not performing any specific task. This is known as Resting State fMRI (RS-fMRI) and task-independent approach. In RS-fMRI studies, subjects are asked to rest quietly for several minutes while brain images are acquired. Since no external time reference, detection of significant activity relies only on the mutual dependence of ongoing activity in different areas.

FC is a concept that describes the temporal dependency between neurophysiological events that occur in spatially distant brain regions, reflecting their level of communication [66]. In the context of functional neuroimaging, FC is typically assessed in RS-fMRI. The first study to use RS-fMRI for assessing FC was conducted by Biswal and colleagues [70], who observed high correlation between the fMRI BOLD time series of the left and right hemispheric regions of the primary motor network during rest, suggesting ongoing information processing. It was thus observed that, at rest, the brain is organized into networks, called RSNs, consistent across subjects and highly similar to networks of task-induced activations and deactivations [71]. They are believed to belong to distinct networks serving different functions such as vision, language, etc, as abovementioned in the Section 2.2.1. The most studied RSNs are: the DMN, the sensory motor network, the right and the left lateral networks, the salience network, the ventral stream network, the task positive network, the primary, the medial and the lateral visual networks and the auditory network.

The absence of a preconceived hypothesis regarding brain activation during the acquisition process in RS-fMRI makes the data analysis more difficult in comparison to task-based fMRI. Despite this challenge, various techniques have been developed for analyzing RS-fMRI functional connectivity. Beyond that network analysis methods, also seed-based and model-free methods are active lines of research.

Seed-based or voxel-based technique was the first method employed for FC analyses in RS-fMRI. In seed-based approaches, one or more regions of interest (ROIs) are selected a priori to evaluate the similarity of their average time course with every other area or single voxel in the brain. The outcome is a map of brain voxels significantly correlating with the chosen seed ROI.

Data-driven methods, such as principal component analysis and independent component analysis (ICA), were introduced as FC analysis methods to investigate general connectivity patterns across brain regions. These methods aim to discover the underlying structure of the data rather than impose a priori knowledge on the model, through a blind separation of meaningful sources.

As regards the network analysis approach, the first step consists of identifying a set of functional nodes. They are often defined as spatial regions of interest obtained from brain atlases, as mentioned in the Section “Imaging-Based Parcellation of the Human Brain”, task-fMRI activation, or through data-driven clustering techniques such as hierarchical clustering or ICA. Once the nodes are defined, the connections or "edges" between them are estimated based on the correlation of their associated time series. More details about this approach, for both structural and functional cases, are reported in the Section “Analysis of Complex Networks”.

2.2.2.2.1 fMRI-Derived Indices for Edge-Weighting

Full correlation is a simple measure for estimating connections between nodes that is most typically computed as Pearson correlation between the BOLD time series. Being the most widespread technique, it was the method addressed in the specific thesis being discussed. However, this measure does not take into account possible influence of other regions, assumes linearity and does not imply causality or the direction of information flow.

To address these limitations, other methods such as partial correlation, mutual information and Granger causality were proposed. Partial correlation can be used to estimate direct connections [72]. Partial correlation is a method that is used to control the effects of other regions in the brain, by removing the linear relationship between the seed region and each other. This allows for the identification of only those specifically correlated with the seed region, cancelling the indirect connections. Mutual information measures the amount of information that one brain region provides about another [72]. It is based on the probability of observing a certain pattern of activity in one region given a pattern of activity in another region and it can be used to identify non-linear relationships between regions. Granger causality is a method that is used to identify causal relationships between different regions in the brain [72]. This is done by analyzing the temporal dependencies between different

regions and determining whether one region's activity can be used to predict the activity of another region. If this is true, then it is considered to have a causal influence on that region. On the other hand, these methods can be computationally demanding, sensitive to certain choices (such as the control regions for partial correlation or the model order for Granger causality) and can make the interpretation of results more difficult. Hence, the great use of the Pearson correlation can be reconducted to its simplicity, providing information about the parallel activations of the regions at low computational cost and easy interpretation. It is essential to evaluate the advantages and disadvantages of Pearson correlation as it remains a crucial method for extracting valuable functional information, pending refinements of more complex methods.

2.2.3 Limitations of Macroscale Connectivity

Despite the utilization of various methodologies, numerous limitations remain in both the identification of nodes through brain parcellation atlases and the determination of connections through measures of SC and FC. Limitations persist regardless of the techniques and processing employed, due to the lack of established gold standard methodologies.

On the one hand, the most widespread brain parcellations, although crucial for generating neurobiologically-significant brain atlases, may not fully capture the intrinsic organization of the brain or encompass the functional variability that is inherent in individual brains, due to factors such as maturation or injury [73]. Furthermore, these parcellations are typically generated on a single or small set of individuals, which can lead to bias and inaccurate representation of population variability, resulting in ill-defined nodes in the constructed network. Several connectivity-driven parcellation methods were proposed to overcome the limitations of parcellation approaches. These methods include hierarchical clustering, k-means, Gaussian mixture models, spectral graph theory, edge detection, region growing, ICA, Bayesian modelling [73]. However, the major challenge is still the evaluation of these methods because of the absence of a universally accepted parcellation as a reference [74].

On the other hand, brain connections are also noisy.

First, structural data collected through DTI is inherently noisy and sensitive to motion. Even small movements during the scan, such as those caused by respiration, blood flow, eye movement, or cardiac pulsation, can greatly corrupt the data. This is the main reason why single-shot Echo Planar Imaging sequences, characterized by a high temporal resolution are used. However, these sequences are also susceptible to other forms of artifacts, such as eddy currents. Eddy currents are distortions of the images caused by changes in the magnetic field, which can result in misalignment of structures through shears, stretches, and translations. This can negatively impact the orientation of the eigenvectors and the accuracy of DTI tractography results.

Second, DTI fiber tracking algorithms used to trace the pathways of WM fibers in the brain have limitations. These algorithms can be divided into two main categories: deterministic and probabilistic. Deterministic algorithms are based on the assumption that there is only one fiber orientation per voxel and the most likely fiber orientation are chosen based on the data. Utilizing a single estimation of fiber orientation at each point in the imaging space, primary eigenvector from voxel to voxel is followed. These algorithms typically begin with user-defined seed voxels and explore the angular relationships between eigenvectors in a given voxel and its neighboring voxels, forming connections based on the smallest angles. Deterministic tractography employs stopping criteria to prevent streamline propagation through regions of high uncertainty and reduce the number of false positives. These criteria, which affects the results, may include an angular threshold between eigenvector directions and a threshold for FA, which serves to ensure that streamlines continue in regions where the direction of maximal diffusion is well defined [75]. Furthermore, deterministic approach does not take into account either systematic or stochastic errors in predicting the orientation of fibers. Systematic errors are a result of complex fiber geometries with varying orientations, which cannot be adequately captured by using the tensor model and the principal eigenvector orientation, as introduced in the Section 2.2.2.1.2. Probabilistic algorithms, on the other hand, take into account the uncertainty in fiber orientations. They propagate a large number of streamlines, each with its own direction of propagation, and are based on a collection of fiber orientations in which there is a quantifiable trust. Probabilistic tractography aims to integrate uncertainty into the tracking algorithm by generating a set of multiple pathways or streamlines that pass through the reached voxel. This allows for mapping of fiber paths in areas where

deterministic tractography methods would have to stop. These methods can be more robust to noise and motion artifacts, but they also require more computational resources and are more complex to implement and interpret. Deterministic and probabilistic approaches have distinct advantages and drawbacks and are typically employed for different purposes. By following the paths of WM bundles, deterministic algorithms have the advantage of providing a sharp evaluation of the streamline number. This is particularly useful when the focus is on the shape or properties of individual tracts. Conversely, they are limited by the adopted diffusion model, which, in the case of DTI, permits a single direction per voxel. Deterministic methods also have the benefit of computational speed. Stochastic methods, conversely, explore probabilistically the extension of the streamline based directly on HARDI data, e.g., by Monte Carlo simulations. These methods are more versatile and allow to incorporate the noise inherent in DTI acquisitions into the modeling of water diffusion. However, the tendency of a fuzzy filling of WM volumes from a seed region makes more difficult the determination of the weight of connections [76], [77]. Clearly, studies about the impact of the two approaches on connection weights and in turn graph properties would be of high interest. In addition, it is also worth noting that the estimation uncertainty of fiber orientation is influenced by factors such as the signal-to-noise ratio in the data and the total number of measurements used in estimation. Finally, the outputs are directly dependent on the parameters of the dMRI experiment, and the accuracy can decrease with factors such as pathway length, shape and size of the reference region, and shape of the tract in question [29].

Third, functional measures extracted from fMRI approach also have limitations, arising from both the technique itself and the specific methods used for analysis. Increasing the spatial and temporal resolution of fMRI data is necessary to improve the localization of activations and to conduct a more detailed analysis of the BOLD signal. Indeed, in FC analysis, it is often assumed that the FC metrics, such as the Pearson correlation coefficient, remain constant over time. However, given the dynamic and condition-dependent nature of brain activity, it has been shown that the FC metrics can vary over time [49]–[51]. This variability necessitates the consideration of dynamic FC analysis, which cannot be ignored as it can also vary within the same subject and even within the same session. In addition, it was demonstrated in dynamic FC studies [78] that different patterns of brain activity, reflecting functional interactions between the brain regions over

time, called states, are identified. A representation of dynamic FC extraction and states recognition is shown in Figure 10.

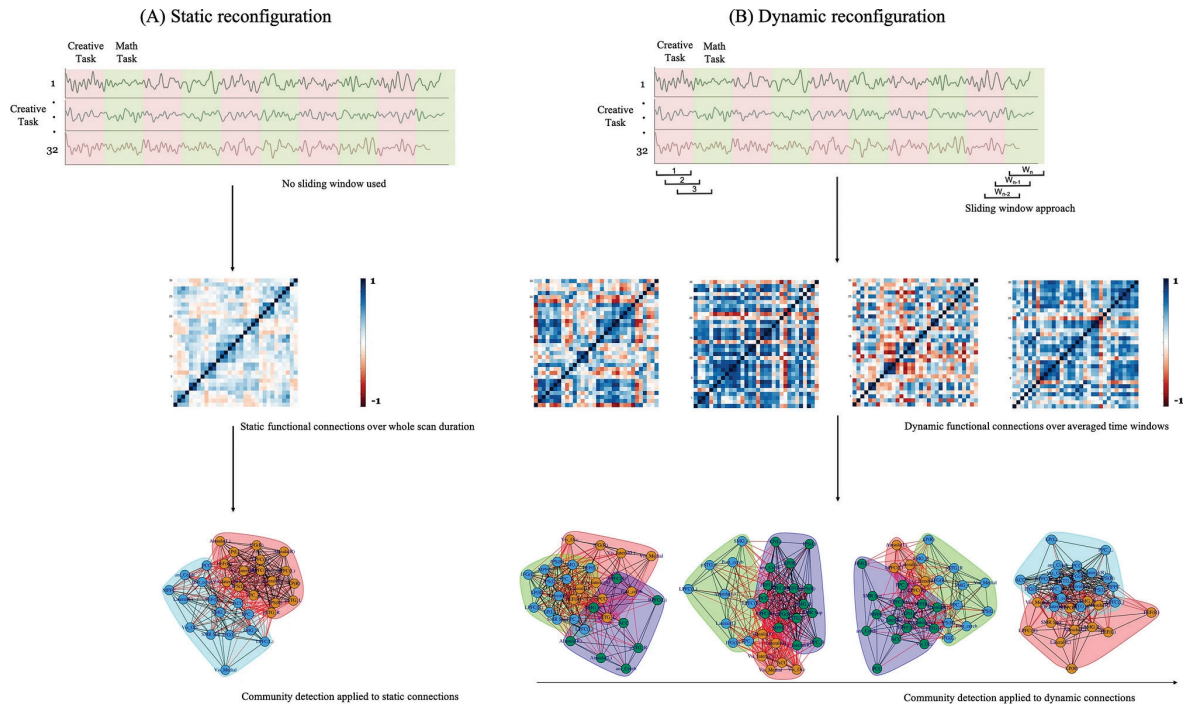


Figure 10: Static and dynamic functional connectivity, which identify states characterized by communities of nodes and edges. The image is taken from a previous study by Patil and colleagues examining static and dynamic configuration of FC brain networks associated with creative cognition in young adults [78].

As such, the evaluation of static FC is subject to uncertainty, that must always be taken into account. Additionally, FC acquisitions are characterized by low signal-to-noise ratio and non-neural noise related to cardiac and respiratory processes, movement and hardware instability [79]. Furthermore, there are ongoing debates regarding issues such as the window length and other confounding factors [80]–[82]. Developing effective methods for identifying and removing artifacts while retaining as much neuronally-related signal as possible is crucial for accurate analysis.

2.2.4 Analysis of Complex Networks

Nowadays, networks can be found everywhere. Social interactions and personal relationships, financial transactions, the spreading of rumors, or physical transport and travel, increasingly occur within networks that are evolving over time. All these networks are examples of complex systems, with highly structured connectivity patterns, multiscale organization, nonlinear dynamics, resilient responses to external challenges, and the capacity for self-organization that gives rise to collective or group phenomena. At its core, network theory is focused on the examination of complex systems in manifold applications that can be represented as networks, where nodes represent individual components of the system and edges represent the interactions between those components [1]. Complex networks then refer to networks that possess non-trivial characteristics such as high levels of interconnectedness, non-random structures, and dynamic behavior. These networks are characterized by a large number of nodes and edges, as well as by the presence of patterns and regularities that are not easily predictable from the properties of individual components. Modern developments in graph theory and complex systems delivered important insights into the structure and function of these diverse networks, as well as quantitative models that can both explain and predict network phenomena [2]. The objective is to comprehend the properties and behavior of these systems through the analysis of the properties and patterns that represent them. One of the fundamental concepts in network theory is the concept of network structure, which pertains to the patterns and organization of the nodes and edges that comprise a network. Another crucial concept is network dynamics, which pertains to the way networks evolve over time. Overall, network theory offers a robust framework for understanding the properties and behavior of complex systems. It enables to analyze the patterns and organization of the systems and make predictions about the dynamics of the systems.

2.2.4.1 Network Definition

The adjacency matrix is the fundamental mathematical representation of the network. This matrix is utilized to create a graph-based representation of a network, in which each row and column corresponds to a node. When the number of nodes in a network is equal to n , the connectivity matrix is n^2 in size. The diagonal elements of the matrix may be used to indicate the connectedness of each node with itself or to express a node's inherent feature. However, many graph theoretical metrics used to analyze brain networks neglect the matrix diagonal by forcing its values to zero. The element at the intersection of the i -th row and the j -th column stores information about the link between nodes i and j , otherwise known as the edge between the two nodes. Hence, the off-diagonal elements are strongly dependent on the method used for connection estimation. Edges can be differentiated based on weight and directionality. In the case of undirected edges, information flows in both directions, the upper and lower triangles are identical resulting in a symmetric matrix. Conversely, directed edges are restricted in their flow of information, leading to an asymmetric matrix. Unweighted or binary networks provide a simplification of the weighted case, in which edges can take either value of zero or one, depending on the presence of the link between two nodes. Whether the edge between two nodes is present, the nodes are said to be neighbors. Despite the widespread use of binary analysis in past studies and correspondingly developed networks metrics, binary graph models are considered far from an optimal strategy. Nonetheless, evaluating the binary topology of a connection matrix may frequently provide useful insights into network organization.

2.2.4.2 Topological Properties

Graph-based indexes are a powerful tool for analyzing the structure and organization of complex networks. These indexes can take on a variety of forms, including local, global, and intermediate. Local indexes, such as path length and clustering coefficient, focus on individual nodes within a network. Intermediate indexes, such as modularity and rich-club coefficient, analyze intermediate structures such as communities or families within a network. Global indexes, on the other hand, focus on the overall organization of the network, including integration and segregation properties such as global efficiency or

global clustering coefficient, and specific characteristics that can be linked to well-known graph models such as small-world and scale-free networks [1], [2].

Among all local graph-based metrics, the degree/strength represents the sum of edges/weights connected to the node. The clustering coefficient quantifies local edge density by counting the triangles average. A triangle occurs if a neighbor of the node is also a neighbor of another neighbor of the node. In the weighted case, the number of triangles is replaced with the geometric mean of its weights. The average shortest path length or characteristic path length expresses the average shortest paths between a node and all the others. The shortest path is a path with the minimum number of edges connecting the node i to node j . It can be expressed as the number of edges passing through, or, in the weighted case, as the sum of these edge weights. The efficiency of a node is defined as the inverse of the average shortest paths connecting all neighbors to it and it is a measure of the integration of a node in the network.

Mesoscale analysis reveals how much the network presents a particular structure. Specifically, the modularity quantifies to what extent the intra-/inter-community link densities are anomalous in comparison to chance. Large values typically reveal significant community structures and fundamental sub-networks. Maximized coreness statistics is a measure of how much the network follows the core/periphery paradigm, which is a partition of the network into two groups, where the number and weight of the edges is maximized in the core and minimized within the periphery.

All the local segregation and integration metrics can also be global, simply by averaging the values across all nodes (global degree, clustering coefficient etc.). Global efficiency is the average inverse of the characteristic path length (the average shortest path in the whole network), and it represents how efficiently the information travel through the network. The density is a measure of sparsity of the network. It is the ratio between the number of actual connections and the maximum number of possible connections. The global average degree is proportional to the density.

An alternative approach to quantify the networks is to classify them based on their underlying structure, which can provide valuable insights into their spatial and topological organization. One method of classification is through the similarity to well-known simulated network models.

One of these models formalizes the concept of small-world phenomenon. It is a surprisingly universal macroscopic behavior in complex systems, first described for large social networks. Watts and Strogatz in [83] investigated how path length and clustering behaved in a basic generative model (henceforth the WS model). The WS model iteratively rewires a binary network of n nodes, each connected to the same number of nearest neighbors via edges of similar weight (simple lattice network). Next, a rewiring process is performed by randomly deleting an existing edge between nodes i and j , and subsequently replacing it with a new edge between node i and any node $k \neq j$. They discovered that as the probability of random rewiring increased incrementally from zero, the initial lattice is progressively organized, despite the random procedure. Sparsely rewired networks show both strong clustering (similar to a lattice) and short path length (like a random network). These algorithmically created graphs were called small-world networks. The small-world organization of the brain demonstrates that it is possible to combine apparently conflicting tendencies of functional integration and segregation within a single architecture. Another model of complex network is known as scale-free [84]. These networks possess a very broad and nonhomogeneous degree distribution, which results in the presence of nodes with far higher degrees. Scale-free networks are characterized by a degree distribution that follows a power law. The term "scale-free" refers to the fact that a power-law distribution does not have a characteristic scale, meaning that "zooming in" on any segment of the distribution has no effect on its shape. Power law degree distributions can be generated by a process known as preferential attachment, which means that new nodes preferentially connect to nodes that already have high degrees, resulting in the "rich nodes getting richer." This process results in the existence of a small number of highly connected nodes, also known as hubs, and many less connected nodes [1].

The similarity to these graph models can highlight characteristics of graphs' topology. For example, one way to quantify the small-worldness property is through the ratio of the clustering coefficient to the path length, after normalizing these metrics by comparing them to the values obtained for a random graph with the same number of nodes and edges. This characteristic highlights a structure in which most nodes are not neighbors of one another. Conversely, the neighbors of any given node are likely to be neighbors of each other, resulting in an easy access by most nodes to every other node with a small number of steps. Main graph-based indexes are summarized in the Table 5.

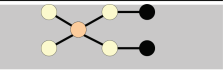
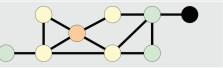






| Analysis | Graph Properties | Graphical Representation | Mathematical Expression |
|-----------------------------------|---|---|---|
| Segregation (local and global) | Degree Strength |  | $k = \text{avg}(\sum_j a_{ij}) ; k^w = \text{avg}(\sum_j w_{ij})$ where i represents the nodes and j the neighbors and weights, respectively |
| | Average clustering coefficient (binary and weighted) |  | $C = \frac{1}{n} \sum_{i \in N} \frac{2t_i}{k_i(k_i-1)} ; C^w = \frac{1}{n} \sum_{i \in N} \frac{2t_i^w}{k_i(k_i-1)}$ where t_i represent the triangles of the node i |
| Integration (local and global) | Characteristic path length (binary and weighted) |  | $L = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} d_{ij}}{n-1} ; L^w = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} d_{ij}^w}{n-1}$ where d_{ij} represent the distances between nodes, and n is the maximum number of possible connections |
| | Efficiency (binary and weighted) |  | $E = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} (d_{ij})^{-1}}{n-1} ; E^w = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} (d_{ij}^w)^{-1}}{n-1}$ |
| Mesoscale | Modularity |  | $Q = \frac{1}{2m} \sum_{i,j \in c_k} [a_{ij} - \frac{k_i k_j}{2m}]$ where m is the total number of edges and c_k are the communities |
| | Coreness statistics |  | $N = \frac{1}{v_s} \left(\sum_{i,j \in V_c} (w_{ij} - \bar{w}) - \sum_{i,j \in V_p} (w_{ij} - \bar{w}) \right)$ where V_c, V_p are sets of nodes in the core, periphery, \bar{w} is the average edge weight and v_s is a normalization constant |
| Global structure | Density |  | $D = \frac{X}{(n^2 - n)/2}$ where X is the number of connections |
| | Small-worldness |  | $S = \frac{C / C_{rand}}{L / L_{rand}}$ where $rand$ refers to measures of random network |

Table 5: Graph-based topological properties with graphical representation and formulation divided according to the level of analysis.

2.2.5 Connectomics

The term “connectomics” refers to the mapping of the human brain in terms of its neural units and connections [85]. The term comes from the definition of the recent “-omics” disciplines which share the characteristics of integrating big data for the mapping of complex systems. In several scientific fields, large amount of data can be comprehensively and flexibly analyzed through network theory and graph-based indexes. In this context, the study of the human connectome revolutionized the study of brain structure and function [86]. Moreover, beyond the connection between brain anatomy, activity and behavior, the brain can be modeled as a network on various scales [87]. Multi-scale and level data is possible to be extracted thanks to the significant advancements which have been made in the field of neuroimaging. A variety of techniques were developed to extract brain parcels and structural and functional connections within the brain, as introduced in the previous Section 2.2. From data, as all “-omics” disciplines, it is possible to use graph theory and its derived graph-based metrics to identify abnormalities. Several biomarkers for a wide range of neurological and psychiatric conditions were already found. These analyses provided

new insights into the underlying neural mechanisms of different conditions, opening new avenues for diagnosis and treatment. The easy ability to study brain connectivity in a non-invasive way provided researchers and clinicians with valuable information about the brain's structural and functional organization which can potentially lead to the identification of new diagnostic biomarkers and targets for therapeutic intervention.

2.2.5.1 Interest in Mapping the Human Connectome

Advancements in imaging networks in the living human brain and improved computational power for data processing and storage have fueled unprecedented growth in the field. The connectome approach represents the culmination of various research areas coming together, which evolved over centuries. The widespread belief is that connectomics and network analysis holds the potential to advance our comprehension of the real working of the human brain and its associated pathologies [88].

This was strongly evidenced by recent initiations of some multicentre research initiatives, which reflected the global scientific community's dedication to this field. The Human Connectome Project (<http://www.humanconnectome.org>) for example was funded research from the National Institutes of Health with the goal of mapping the connections within the human brain in 1200 healthy individuals through advanced functional and structural imaging techniques. The Human Brain Project (<http://www.humanbrainproject.eu>), is another major research initiative backed by the European Commission for the purpose of simulating the connectivity of the human brain and its computational abilities at a neuronal level. The goal was also to establish a research infrastructure for decoding the human brain and developing information technology influenced by the brain [89]. The Brain Research through Advancing Innovative Neurotechnologies – BRAIN (<http://www.braininitiative.org>) Initiative also drove development in this field aiming to advance technology to catalyze neuroscience discovery [90].

These brain-related great initiatives propelled an unprecedented support to the whole neuroscience, cognitive sciences, neurology and bioengineering communities. Indeed, we witnessed in recent years a tremendous explosion of human atlas and connectivity brain projects with various goals, scopes, and sizes [91]. The Big Brain project for example aimed to achieve ultra-high resolution neuroimaging [92], the CONNECT project to

combine macro- and micro-structural knowledge [93], the Brainnetome project to comprehend the brain and its related disorders and create a comprehensive brain network atlas [94] and others [95], [96]. Therefore, with the introduction of advanced acquisition techniques and the generation of vast amounts of data at ease, the concept, role, and understanding of a human brain atlas continue to evolve and expand.

2.2.5.2 Human Brain Networks Definition

In the framework of the human connectome, adjacency matrixes depicted in the previous Section 2.2.4 are brain connectivity matrixes. As said, the nodes and edges of a brain network may vary based on the method utilized to obtain the connectivity data. When constructing brain connectivity matrixes from MRI data, nodes represent GM parcels from well-established brain atlases based on either anatomical or functional criteria and the edges describe either SC or FC features (as said in the Section 2.2.2). SC and FC generally depict a non-causal and symmetric relationship between brain regions, resulting in a symmetrical connectivity matrix composed of $n(n-1)/2$ pairwise connections between n nodes. A schema representing the pipeline of brain connectivity network extraction is shown in the following Figure 11.

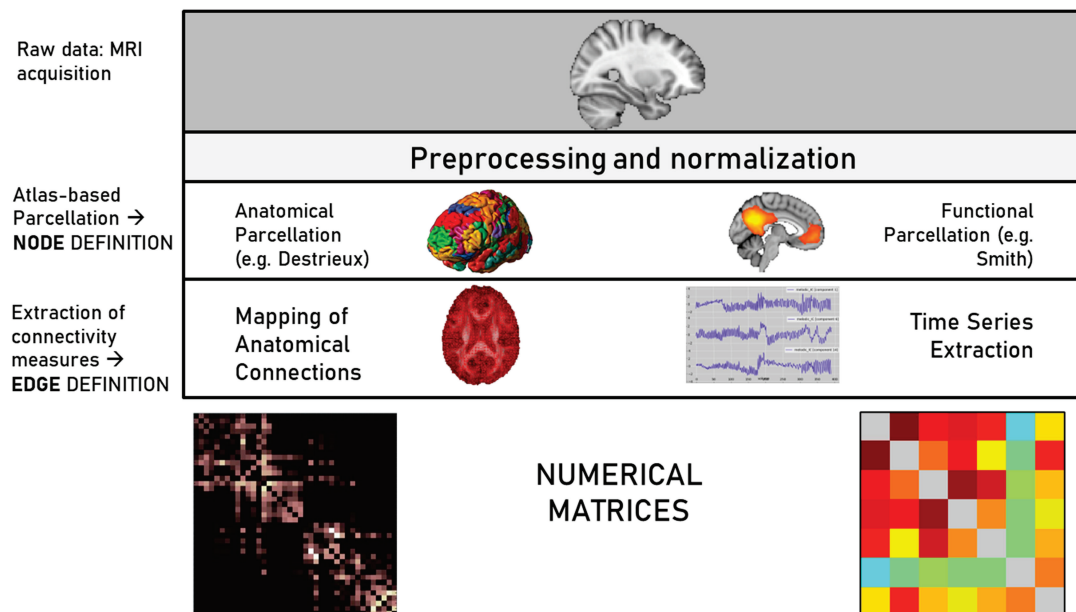


Figure 11: Schema of the phases of processing for the definition of brain structural and functional network.

The range of weights observed and the way they should be interpreted is determined by the method used for connectivity estimation. It is worth noting that in the specific case of FC being obtained from Pearson's correlation of BOLD signals it is possible to have negative edges. These inverse connectivity patterns were not often considered, being thresholded from the resulting matrixes [97], [98]. However, these anti-correlations were proposed to play a fundamental role in the pathological mechanisms underlying some neuropsychiatric disorders [98], [99]. In addition, the anticorrelations could indicate inhibition of one region over another, which is a possible physiological mechanism. For this reason, separate analysis of correlations and anti-correlations and many other more complex methods were proposed to analyze them [100].

2.2.5.3 Pre-Conditioning Methods and Limitations

In the field of brain research, it has been widely acknowledged that traditional methods for both structural and functional edge weighting have limitations in accurately determining the strength of connections (as also reported in the Section 2.2.3). As a result, numerous studies disregarded connection strengths and instead focus on graph-theoretical analysis of binary connectomes. The goal is to capture the topology of connections, but not their quantification. However, this approach can result in weak connections being indistinguishable from strong connections [101]. On the one hand, using a binary graph model is an inappropriate approach for network analysis of tract tracing data [86]. The biological diversity of fiber connection strengths, spanning over six orders of magnitude in mammalian cortical networks, would be removed. On the other hand, in FC studies, the obtained weights of the connections reveal the nature of the brain regions activations, mirroring positive and negative correlations. As a result, the application of a threshold to maintain all edges above it leads to the loss of the relevant information of anticorrelations that, for example, was proved to be especially related to automatic inhibitory mechanisms [31], [32]. Only recently the trend to suppress them was reversed with the result of more and more studies aiming to separately analyze positive and negative networks or extracting ad hoc topological measures mirroring segregation and integration of activations and deactivations [100]. Furthermore, different values of connection weights drive different fundamental information to the whole topology [102].

Thresholding Procedures

Advancements in the connectivity definition methodologies facilitated the calculation of dense weighted connectomes also for SC, providing a more reliable measurement of connection strengths [101], [102]. This progress allowed for the direct use of weight information in graph theoretical analysis. However, the probabilistic nature of the tracking process and, more in general, the measurement noise in SC, as well as the considered errors in fMRI-derived indexes also result in an undetermined number of false edges [79], [81], [103]. As a result, network thresholding techniques are frequently used to remove suspected spurious connections and emphasize topological features. However, it is yet to be precisely determined how different thresholding methods impact basic network features. Nevertheless, different thresholding strategies were proposed in various studies for different aims.

Among all, absolute thresholding result to be the one of the most common in connectivity studies [104]. It uniformly applies a single threshold across the network to retain only connections with weights higher than the predefined threshold. However, this approach can result in different numbers of network edges across datasets and varied levels of network density. Many graph metrics have been shown to be influenced by network density [105], potentially leading to statistical differences in network metrics between patient and control populations. These differences should be attributed to the varying number of links in the network rather than differences in network topology related to the disease. As a result, absolute thresholding is considered less suitable for case-control investigations [97].

To address the issue of varying numbers of connections between groups, a density-based thresholding approach was proposed [97], [105]. This strategy aims to maintain a constant number of connections across all individuals and reduce the impact of network density on graph metric computation and comparison between groups. In this thresholding method, a specified percentage of the strongest connections in each individual network is selected and retained. This strategy is one of the most widespread and, for this reason, it was implemented in the software tool developed in article I. It is a procedure especially used in FC analysis with the goal of producing more sparse matrixes, which are generally complete because of correlation result different from zero. In this case, the potential differences in graph metrics between groups are believed to be a result of differences in the topological

organization of edges, rather than differences in the number of edges. However, the inclusion of lower edges to reach the preset density can also imply the potential inclusion of less reliable connections. As such, more random connections result in the creation of shortcuts and for example the decrease of the overall shortest path length. This results in an improvement in global efficiency, a reduction in clustering, and a topology that is closer to a random network configuration.

To summarize, it is acknowledged that the topological properties of a network are contingent on the number of edges, thereby necessitating the use of density-based thresholding. Conversely, variations in connection density between individuals or groups may provide valuable information, indicating that any effect these differences have on network topology should be regarded as real and not eliminated through density thresholding [106]. In general, absolute and density-based thresholds can introduce a confound on graph-based indexes in the context of a group study. On one hand, different levels of sparsity can be obtained by applying the same threshold to all matrixes. On the other hand, matrixes can be matched according to sparsity, hence selecting a distinct threshold value. However, in many cases, the weight distribution of one of the two groups can be reduced, strongly affecting weighted indexes results [107], [108]. Therefore, it is crucial to exercise caution when applying these thresholding methods.

Another technique commonly used is proportional thresholding in which connections are retained based on their prevalence across a predefined proportion of subjects [7], [109]. The connections that meet or exceed the group threshold in terms of detection count are then kept. On the one hand, low group threshold allows for the inclusion of connections that are challenging to reconstruct and helps prevent false negatives. On the other hand, a high group threshold demands the presence of a connection in a significant number of participants, thereby enhancing the accuracy of reported connections and reducing the occurrence of false positives. For example, in the study [110], it was demonstrated that the threshold value has a significant effect on commonly used network metrics, emphasizing the importance of methodological understanding in the selection of this parameter. However, the selection of specific group threshold is often based on empirical evidence, but it can introduce bias and affect the comparability of results between different studies.

Another approach known as consistency thresholding was proposed [111], [112]. This technique involves retaining connections with consistent weight values across individuals,

with the assumption that connections exhibiting excessive inter-subject variability are likely to be erroneous. The ultimate goal of this technique is to construct a group-representative network, which serves as a single representation of the connection data obtained from multiple subjects. A mixed proportional and consistency thresholding method was used in article I to obtain a HC template matrix to be compared with two case studies. The method allowed the thresholding according to the original HC population density and the creation of an HC group mean matrix.

Recently, new more advanced methods such as percolation analysis or spanning tree approaches were proposed [47], [101], [113], [114]. The former for example was widely employed in a number of studies to find a potential optimal deletion of the spurious edges maintaining brain network structure and connectedness [114]–[116]. In [116], the percolation analysis was performed with respect to community analysis. The size of the giant component of the network upon weakest edges deletion to find the optimal sparsification threshold preserving structure and connectedness while removing potentially spurious correlations was measured. Many novel thresholding procedures were proposed [117], [118], but they must be cautiously chosen according to the specific case and application.

2.3 State of the Art and Open Issues in Brain Connectivity

The proposed research focused on developing methods for improving the interpretability and the uncertainty of brain connectivity biomarkers in healthy and pathological populations with the final aim of contributing to an improved usability. The lack of methods designed for this purpose and the need for more robust and generalizable approaches are evaluated.

In Section 2.3.1, the interpretation of the brain connectivity is assessed considering previous tools and methods for the mapping and the investigation of brain networks and sub-networks of interest. In Section 2.3.2, possible approaches to the problem of the uncertainty assessment and robustness control are analyzed in both brain connectivity applications and others, addressing the limitations of the most widespread connectivity

measures. AI methodologies applied on brain data are then considered for the classification task of AD in Section 2.3.3. XAI approaches are then discussed to assess interpretability of the models for a better comprehension of the DL underlying processes and the possible biomarkers obtained from peculiar sub-networks or regions. The applicative background integrating localized disruptions and widespread degeneration, such as stroke and schizophrenia, is examined in Section 2.3.4. The field of connectivity is vast, and as such, the "Open Issues" outlined in this paragraph are not an exhaustive list of all possible research lines. Nonetheless, they encompass the problems addressed in the present thesis work, which will be summarized in Section 2.4.

2.3.1 Mapping, Visualization and Analysis of Brain Networks and Sub-Networks

The representation of the human connectome through brain connectivity matrixes is extensive and quantifiable, but it does not always provide a clear and intuitive visualization of the connectivity pattern. The large size of the brain connectivity matrixes makes it difficult to comprehend the information, leading to concealed information. As a result, methodologies for mapping and visualization of connectivity data are important to enhance the interpretation of connectivity measures. In particular, they are essential for exploratory analyses with the aim of identifying characteristic patterns that distinguish pathological conditions from physiological ones, to evaluate changes after pharmacological treatment or rehabilitation, or also to make a quality check of the processing pipelines.

The connectogram offers a solution to this issue by providing a graphical representation that bridges the gap between quantitative connectivity analyses and intuitive visualization. Irimia and colleagues [119] introduced connectograms as circular graphs in which nodes of a network are represented along the perimeter of the circle and edges of the network are depicted as arcs connecting pairs of nodes. The use of connectograms in brain connectivity mapping was inspired by the usage of similar graphical representations used in genomics. After this work, several connectivity studies confirmed the possibilities offered by the connectogram representation [7], [120], [121]. A notable example was given by the work provided by Van Horn and colleagues [122], which tackled one of the most popular and

studied clinical neuroanatomy case of Phineas P. Gage. Gage was a foreman in the field of railway construction who gained notoriety for surviving a remarkably traumatic accident in which a massive iron rod penetrated his skull, resulting in extensive damage to his left frontal lobe. Despite his injury, Gage lived for an additional 12 years, during which his personality and behavior underwent significant changes. As a consequence, the case was widely discussed for the prefrontal area involvement and function and its relation to personality and behaviour. In the work by Van Horn and colleagues, authors attempted to reconstruct the trajectory of the rod to determine the extent of the injury's impact and represent it on the connectogram, as shown in Figure 12.

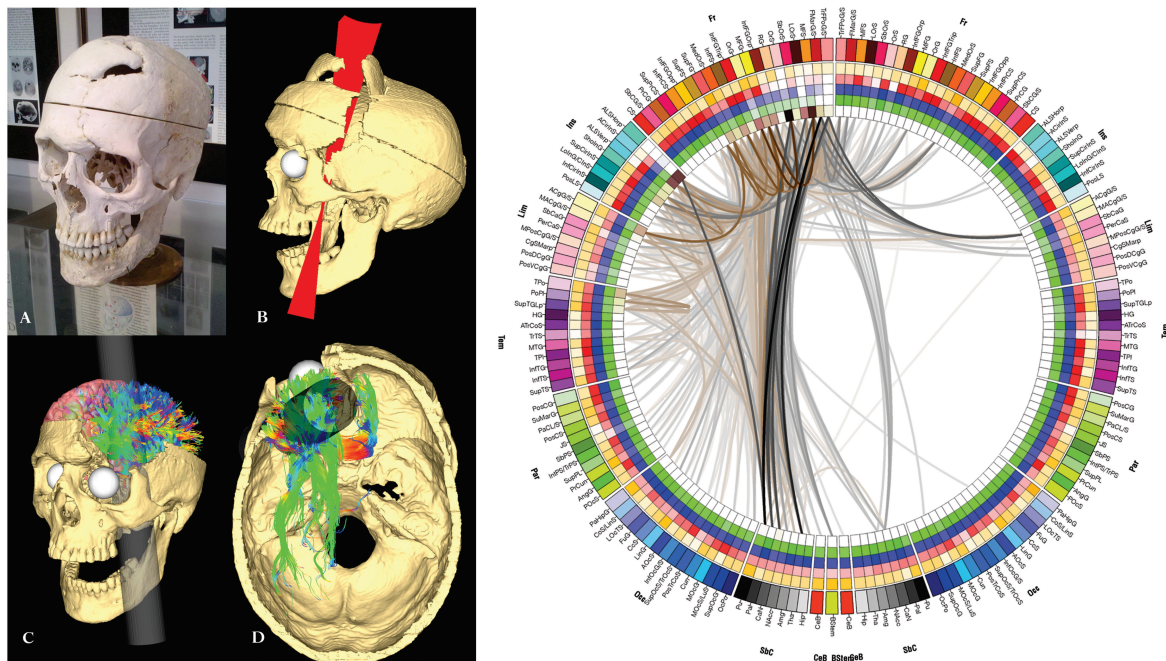


Figure 12: Reconstruction of Phineas Gage's damage from the work made by Van Horn and colleagues [122]. The skull on display at the Warren Anatomical Museum at Harvard Medical School, computed tomography image volumes, rendering and view of the interior of the skull and the connectogram are shown.

The production of connectograms can be done using the powerful software package Circos [123], which is designed for visualizing data, exploring relationships between objects and creating publication-quality illustrations. Circos is highly flexible and can be utilized in various fields, but it lacks a user-friendly interface and must be run through command-lines, which may pose a challenge to researchers without programming experience. A more accessible user-friendly interface for researchers with knowledge and interest in brain

connectivity, but not computer science, is missing and may broaden the availability of connectograms.

Despite the intuitive visualization provided by connectograms, it is worth noting that the huge number of connections typical of healthy subjects can still make the interpretation of the networks tough. The human brain is indeed composed of a vast network of connections, which are linked to high-level cognitive functions. To gain a deeper understanding of the brain's connectivity, it is a common approach to focus on sub-networks [12], [120], [124], which are characterized by unique patterns of brain activation revealing specific domains of behavior and cognition [124]. Focusing on sub-networks can also lead to easier data interpretation driven by the addressed physiological and/or pathological problem. In this regard, also a tool called Network Based Statistics (NBS) was implemented to find one or more sub-networks of interest identifying the differences between two groups [12]. Indeed, the brain is composed of several interacting lower-scale specialized circuits, as mentioned in the Section 2.2.1, which can be identified as unvaluable biomarkers revealing helpful information. The specialized analysis of network of interest can be also of great importance for the pathological cases having damage in localized areas of the brain with the purpose of investigating its effect, such as the work by Van Horn and colleagues or in stroke investigation [8], [122]. In this regard, the Figure 13 shows the regions affected by a stroke lesion case and its related sub-network of interest, composed by the areas overlapping with the lesion.

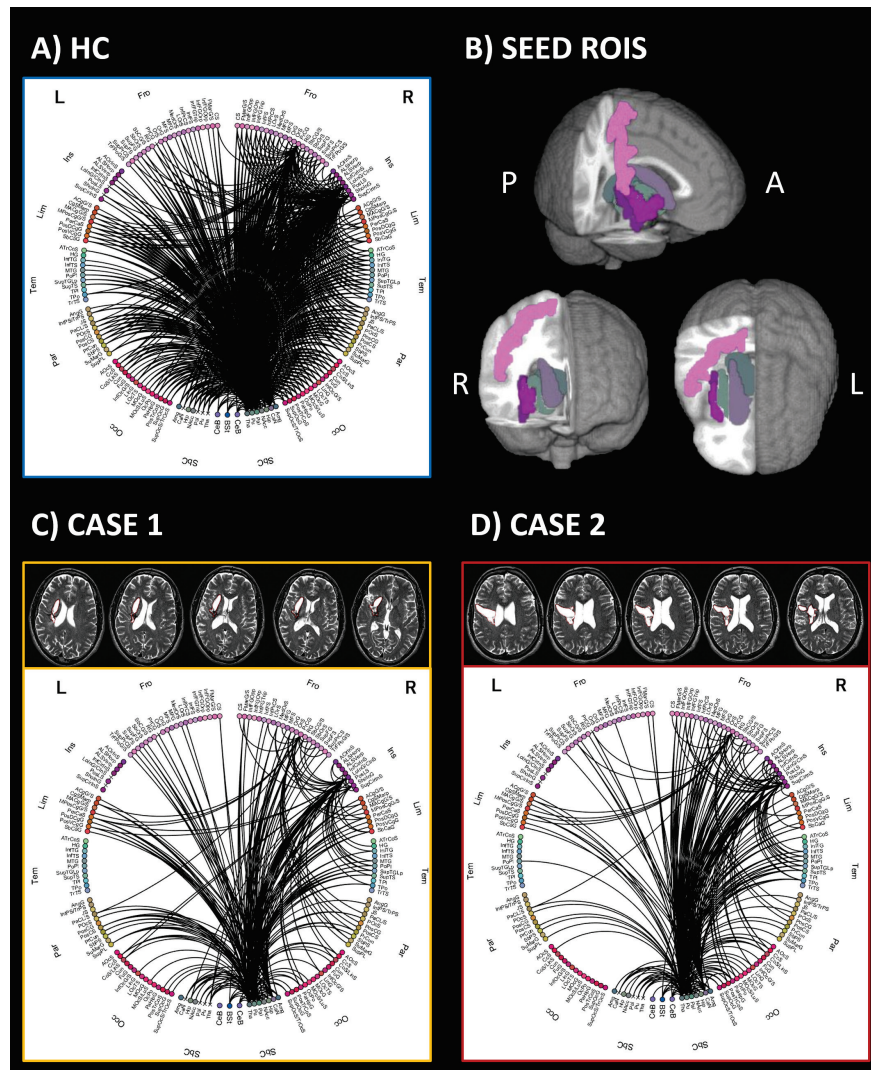


Figure 13: Connectograms of the sub-network composed by the ROIs overlapping the stroke lesions (reported in B) of a group of HC and two stroke case studies, together with their five axial slices.

However, current software for visualizing connectivity patterns, such as Circos [123], BrainNet Viewer [125], BRAPH [126] or NBS [12], although assessing modules in the networks, do not provide the capability of interactively and flexibly selecting individual nodes/sub-networks within a whole-brain network for the qualitative and quantitative evaluation. This is because direct upload of the specific sub-network of interest is usually necessary, or it is automatically computed.

Analyzing sub-networks can be done both qualitatively, through connectograms, but also quantitatively, through the calculation of various graph-based network properties. Such properties, including node degree, small-worldness, modularity, clustering, and central

hubs, can provide valuable insights into the network's topology and architecture. Through these organizational indexes, changes of brain networks in various neurological and neurodegenerative disorders were revealed [3], [5]–[8], [47]. Despite this and the possibility of computing them through different tools, gold-standard methodologies for extracting brain connectivity are still missing. The evaluation of these alterations of the brain networks in these conditions thus remains an open issue.

To further advance our knowledge of brain disorders and the effects of treatments, such as disease modifying therapies and rehabilitation, accessible tools for assessing the architecture of brain networks, considering the uncertainty of the measures, are needed. These tools would provide a larger body of evidence and promote a deeper understanding of the human brain.

2.3.2 Robustness of the Connectivity Measures

A major challenge in connectomics is related to the limitations of methods for the extraction of the connectivity, as detailed in the Section 2.2.3. Indeed, although the analysis of changes in topological properties is a powerful tool to analyze brain organization, its statistical investigation is often subject to uncertainty. Different statistical analyses can be used to allow for a more robust evaluation.

Among all statistical methodologies, bootstrapping was already used in other imaging problems, such as in the work conducted by Lazar and Alexander [127]. This approach enables the robust estimation of statistical features of a population from a limited number of measurement samples, without making assumptions about the distribution of the initial data. In contrast to traditional hypothesis testing, bootstrapping does not require a test statistic that satisfies specific assumptions dependent on the experimental design or knowledge of the data's properties. Thus, the main advantage of bootstrapping is the ability to quantify the uncertainty variability of the estimator, characterizing errors and dispersion in the null hypothesis [128], [129]. The number of bootstrap samples used results in a statistic representing smaller-sized random sample with replacements from the initial distribution. It is important to note that obtaining thousands of bootstrap observations from the initial data does not equate to collecting new data. The approach is based on a simulation of data (surrogates), and its usefulness lies in the quantification of statistical

quantities such as the standard error, potential bias, and confidence interval of a particular sample of data.

In the context of investigating graph-structured data, the quantification of inherent uncertainty is crucial for their scientific usefulness. The study by Green and Shalizi [130] demonstrated the application of bootstrapping on random graphs, where resampling of data was capable of approximating the distributions of motif densities, such as the number of times fixed subgraphs appear in the network. This study also utilized bootstrapping to quantify the uncertainty of network metrics. Another example is provided by the research conducted by Gel and colleagues, where bootstrapping was used to estimate the uncertainty in graph degree distributions [129]. As the number of resampled data increases, applying bootstrapping to a data distribution results in a more refined definition of variability centered around the initial mean of the same distribution [131].

However, the bootstrapping approach was only partially applied in brain connectivity studies. For example, in the study conducted by Wei and colleagues this approach was used to perform connectivity matrix feature selection in a regression task cognitive traits prediction [132]. Spearman correlation analysis was indeed performed between connectivity and cognitive measures in each resample subset to extract a feature vector. The generation of surrogate graphs reduced the uncertainty in the feature selection method applied to four different models for predicting cognitive abilities. In another study [133], the authors suggested using a bootstrapping approach on FC data, similar to a previous study on Parkinson's disease patients. Since the subjects' sample might not be fully representative, a way to reliably test the significant group differences in the graph-based indexes was this resampling methodology. Bootstrap allowed for a better estimation of the null distribution of mean FC. However, a drawback of this approach is the computational cost, as bootstrapping requires multiple resampling of the original dataset which can be computationally intensive, particularly for large datasets.

To the best of our knowledge, there is great potential for implementing bootstrap approaches on connectivity data effectively and more quantitatively, but underlying concerns still need to be resolved. It is evident that diverse kinds of noise may have a significant influence on connectivity measures, especially with small cohorts. Moreover, the static evaluation of FC has limitations [49], [81]. In order to make brain connectivity

biomarkers more understandable and useful, it is critical to assess their stability and robustness.

In addition, as discussed in the Section 2.2.5.3, most widespread thresholding methods are arbitrary and can heavily affect the results [104], [105], [110]. Although used for the deletion of spurious connections, they can strongly influence the network topology and organization. For this reason, different novel approaches appearing to be less arbitrary than typical density or absolute thresholding and which may potentially improve even more the estimates. such as percolation or spanning tree [47], [101], [109], [113], [134], were proposed. However, agreement on standard methodology is still missing.

2.3.3 Explainable Artificial Intelligence in Brain Imaging and Connectomics

Recent advancements in neuroimaging and the possibility of collecting large amount of data in a much easier way and at minor cost opened up new opportunities of reducing the uncertainty of the biomarkers and robustly assess the measures. In addition, through AI methods is possible to develop intelligent healthcare systems and clinical decision support. In particular, with the growing global incidence of neurodegenerative disorders, such as AD, there was a heightened interest among researchers to advance in areas such as diagnosis, treatment, prevention, drug discovery, and provision of improved healthcare services. Most of the research in the domain of AD and more in general neurodegeneration is centered on using brain imaging. Traditional Machine Learning (ML) methods with classification methods like Logistic Regression, Decision Tree, Random Forest, Support Vector Machines and Multiple Kernel Learning were widely utilized both from MRI (mainly T1, but also Fluid Attenuated Inversion Recovery - FLAIR and T2), fMRI images, related features and connectivity data [43], [135], [136].

In addition, the availability of significant computational resources and the advancements of DL algorithms enabled the wide application of these techniques to improve the accuracy of computer-assisted diagnoses. Also, deep networks were widely applied for different tasks from both brain images (both 2D and 3D volumes) and connectivity data. Regarding models using brain images, many approaches have been widely investigated, becoming a well-established area of research. In this context, CNNs have particular importance, since

it was recently demonstrated to have impressive performance in medical analysis, both having as input 2D slices and 3D scans [137]–[140]. In addition, different CNN-based pre-trained models, such as ResNet18, EfficientNet-B0 and VGG etc., were widely employed in neuroimaging research, providing state-of-the-art performance for different tasks [140]–[142]. ResNet18, for example, is a model characterized by 18 layers including a 2D convolutional layer and four sets of residual blocks, each containing two 2D convolutional layers, which was proved to be suitable in AD recognition [143]. Furthermore, its appropriateness for handling 3D volumes [39] further motivated its selection, use and transfer learning within the article III. Subsequently, other approaches, mainly adapted CNNs, GNNs and Autoencoders, were started to be used feeding into the models the brain connectivity matrixes, as also summarized by the research made in the Scopus Database and reported in the Figure 14.

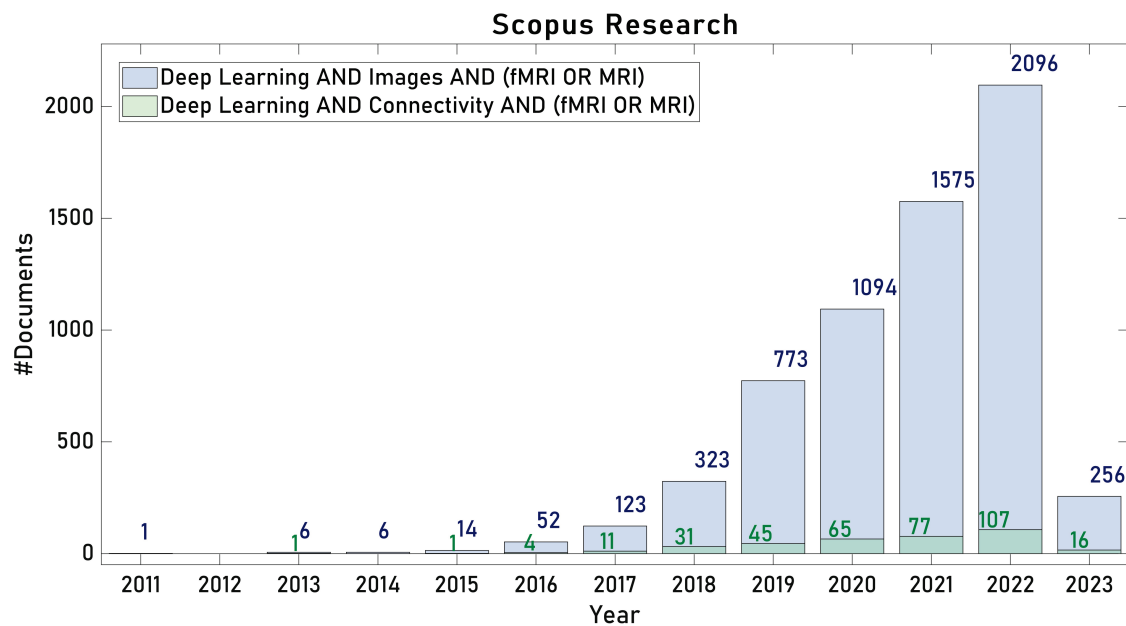


Figure 14: Results obtained from research made on the Scopus Database. The number of documents published per year is reported in green, using as keyword “connectivity”, and in blue, using “images”, together with Deep Learning and MRI or fMRI.

Clearly, the use of DL on connectivity is not as mature as using images, probably because of the cost of the processing, but there has been an increasing interest in the last years in this regard. Meszlényi and colleagues proposed a connectome-based CNN architecture for MCI classification, using FC. The convolution was applied in two layers, first row-by-row and then by column [144]. Similarly, the “BrainNetCNN” is a deeper convolutional

network also adapted for brain networks that was validated on SC matrixes from DTI images for regression task [145]. Subsequently, this model was tested and compared to different ML models with respect to the AD classification task [38]. Alorf and Khan obtained good performance in the AD stages classification by employing BCG-CN (a GNN adapted for brain connectivity data) [20] and Stacked Sparse Autoencoders [146]. An extended and adapted version of the former was employed in article III. Its ability in dealing with dense matrixes, focusing on direct and indirect connections resulted in the best performance during AD recognition from SC data. Also, graph variational autoencoder employing both SC and FC was employed on AD dataset to find a unified embedding via a classification task [15].

Despite the effectiveness of DL algorithms in various classification tasks, their widespread adoption and trust in clinical settings is limited due to their well-known "black box" nature. Their architecture is understood but the internal processes used for classification remain inaccessible to humans. The recent development of XAI methods aims to bridge the gap between the high performance of DL networks and the need for human comprehension of their processes. XAI methods enabling visualization and interpretation of the results of these networks were applied across a wide range of applications. Also in the neuroimaging field, together with the rise in the usage of DL techniques, the use of XAI methods strongly diffused. In this perspective, different methods were applied in different contexts, adapted to different types of data, and tailored for diverse purposes. Although the increased interest on DL brain connectivity, it is worth noting that many approaches were adapted to graph structured data, but validation on human brain networks was not widely performed yet. Illustrative neuroimaging studies, with regard to AD classification, using popular interpretability methods are reported in Table 6. Among all approaches, the most widespread methods in neuroimaging can be roughly classified into three categories. These are gradient/feature-based methods, perturbation-based methods, and distillation methods.

| Interpretability Method | Category | | | Studies | |
|--|--------------|--------------|----------|---------------------------|---------------|
| | Perturbation | Distillation | Gradient | Images | Connectivity |
| Occlusion's Sensitivity | X | | | [39], [139], [141], [147] | [148] |
| LIME | | X | | [149], [150] | [151]* |
| SHAP | | X | | [152], [153] | [151], [154]* |
| Cam and Grad-Cam | | | X | [152], [155], [156] | [157], [158] |
| Guided Backpropagation Or Layer-wise Relevance Propagation | | | X | [139], [159], [160] | [158] |

Table 6: Interpretability methods in neuroimaging studies. *Applied for Autism Spectrum Disorder classification task

Gradient/feature-based methods use gradients or hidden feature maps to determine the importance of different input features in the model's predictions [161]. Perturbation-based methods evaluate the effect of changes in input information on the model's predictions. The importance of input features is measured by monitoring the variation of the model's output for different input perturbations [53]. If important input information is retained, the predictions should be similar to the original prediction. Other methods which can be considered perturbation-based methods are the so-called distillation methods. They are based on local perturbation of the initial input to create separate models and usually considered model-agnostic (training surrogate models or computing the marginal contribution) [162]. The category of perturbation-based methods, as it is defined, is not model independent that is the reason why they fall under a separate category.

The lack of interpretability of DL models made it imperative to develop ways to gain a deeper understanding of the problem at hand. In fact, these methodologies can also be helpful for the reinforcement of brain biomarkers. In addition, the cost of processing for the extraction of the brain connectivity networks can be evaluated in relation to the different interpretation and accuracy provided by this data with respect to more typical brain images employment. This assessment can bring to the development of novel interpretable models able to exploit both morphological and interregional features from multiple data.

2.3.4 Multi-Level Degenerations

The analysis of the connectivity networks is complex, concerning different scales, characteristics and layers. This is particularly important considering that many pathological conditions exert either widespread effects throughout the brain or localized disruptions, or even both, affecting either anatomy or function or both. In this thesis, three different conditions sharing the characteristics of affecting the brain at different levels were assessed.

First, stroke is a leading cause of significant disability globally, with over 50 million people experiencing a stroke each year [163]. A stroke can result in loss of motor function, speech, and cognitive abilities, severely affecting a survivor's quality of life and ability to carry out daily activities [164]. This can result in a reduction of healthy life and substantial costs for post-stroke care. The onset of stroke occurs when a blood vessel supplying the brain is blocked/burst (ischemic/hemorrhagic), causing a cluster of brain cells to become damaged, or lesioned. Stroke has extensive impacts on the brain that go beyond the lesioned area. Therefore, the changes to the neural structure induced by stroke extend throughout the brain, including GM and WM structures that are remote from the injury site. Research revealed that WM degeneration and GM atrophy occur beyond the direct injury site [165], [166], and that significant abnormalities can be observed in regions which are structurally connected to primary lesion location, such as the parietal lobe or superior frontal gyrus [167]. In addition, Sommer and colleagues found decreased perfusion in the cerebellum, which is caused by reduced excitatory signals from the cortex [168]. In addition, these changes could play a crucial role in functional recovery after stroke as it was demonstrated that the FC of the brain may be compensated by indirect connections [169]. It is thus clear that the extent of the stroke lesion heavily affects the whole functioning of the brain thus reporting abnormalities not only in the lesioned area but also in the lesioned hemisphere and in the contralateral one.

Second, SZ is a mental disorder characterized by detached mind's processes which may be depicted as a disorder of brain connectivity. This is evident in the hallmark symptoms of the pathology, such as diminished cognitive and emotional abilities, which demonstrate the disconnection of various mental processes [47]. The specific clinical manifestations of the

disorder include hallucinations, agitation, social isolation, anhedonia, and apathy. Being a disorder of brain connectivity and considering the functional dysfunctions, several studies explored the relations between the symptoms and functional associations [46], [134]. Studies found connectivity deficits in SZ on multiple levels. At the global level, a reduction in connectivity was reported in various studies [45], [107], [134]. At the sub-network level, fronto-temporal and occipito-temporal dysconnections were found through NBS [12]. Liu and colleagues observed a significant alteration in the pattern of small-world topological properties in several brain regions of the prefrontal, parietal, and temporal lobes. Skudlarski and colleagues found differences in the connectivity pattern originating from the posterior cingulate cortex within the DMN [36]. The DMN is considered important in the context of SZ because it is involved in social behavior and emotional state control, which were found to be altered in patients with the disorder [170]–[172]. Although, the observations regarding the DMN remain controversial, with no average connectivity changing according to the dynamic mental state and with individual connections subject to reorganization [36]. Amongst all brain regions, the prefrontal cortex was identified as one of the most affected regions [45], [46], [107].

Third, as populations continue to age, the number of individuals affected by neurodegenerative diseases has become a significant public health concern, imposing a heavy social and economic burden on many nations. A major form both for severity and diffusion is AD, leading to impaired memory and dementia [173]. In recent times, research focused extensively on AD and neuroimaging techniques emerged as critical tools in the diagnosis and examination of neurodegenerative diseases. AD structural brain changes visible through MRI revealed the early involvement of the MTL, particularly in terms of hippocampal and parahippocampal atrophy [174], [175]. Numerous studies have also connected the entorhinal cortex to modifications in the cognitive function of patients. [23]. This condition does have a significant impact on the GM of the brain, causing substantial cortical and hippocampal atrophy, but it also has an impact on the WM [176]. WM fiber tracts lose axons and myelin degenerates in AD [177]. In this regards, T2-weighted or FLAIR imaging techniques are frequently employed to assess the presence of WM hyperintensities, which serve as a hallmark of the condition [141]. There is growing empirical support for the notion that the disruption of the brain's fiber networks may play a role in exacerbating symptoms as the disease progresses [6]. Additionally, the

neurodegenerative process that underlies AD involves the cerebral cortex. Both the temporal and parietal areas exhibit atrophic alterations that are consistent with the condition's early stages [178]. On the other hand, the symptomatic progression is often accompanied by a more extensive brain cortical thinning alongside ventricular enlargement [174], [179]. Due to the aforementioned dendritic, myelin, and axonal loss that comes along with atrophy, dMRI and fiber tracking can be used to support AD assessments [23]. It has been well demonstrated that AD causes coherent alterations in the connectivity structure across many dMRI measures used for edge-weighting [24]. Beyond changes in structure, also differences in function occur across a range of spatial scales and levels [180], [181]. At both the FC and SC global level, AD was depicted as a disconnection syndrome which can be characterized by the connectome degeneration, affecting network topology, that is governed by long-range connections [182]–[184]. At sub-network level, studies demonstrated that the DMN, which is involved in memory processes, is vulnerable to atrophy, amyloid protein deposition and WM microstructure alterations resulting in a disrupted SC configuration [24], [25], [185]. AD abnormalities found in SC and dMRI studies also involved the temporal lobe, whose disruptions contribute to memory impairment [6], [24], [186], [187], and, consistently to morphological data, some regions of the MTL [42], which are also often reported in the DMN. RS-fMRI consistently revealed a decreased FC between the posterior and anterior portions of the DMN, including the precuneus, posterior cingulate cortex, anterior cingulate, medial prefrontal cortex, and MTL structures [188], [189]. DMN was found to be replicable hallmark for AD also in SC analysis, highlighting many common regions [190].

It is thus worth noting that all the considered conditions are different from each other. However, the brain is characterized by a strong affection at different levels, with consequences on both segregation and integration. In this sense, the study of connectivity can provide many insights for a better understanding of their relations. As such, there is the need for a multi-scale approach and assessment of the brain connectivity and derived biomarkers to address both global and local structures in different brain disorders.

2.4 Objectives and Outline

Hence, the proposed PhD thesis has the overarching: i) methodological aim to fill the gaps of uncertainty and interpretability of existing approaches for the mapping, visualization, extraction and analysis of brain connectivity biomarkers; ii) translational aim to demonstrate the potential of the proposed methods in the context of pathologies characterized by both focused disruptions and widespread degeneration, with the purpose of improving usability.

Specifically, the objectives of this PhD thesis were:

- To provide a review of the connectomics field, considering the most widespread and employed methods for extraction, analysis and interpretation of brain networks and sub-networks (Chapter 1)
- To discuss the open issues limiting the use of obtained biomarkers in clinical research settings and practice (Chapter 1)
- To analyze functional and structural brain connectomics multi-level findings characterizing the domain knowledge of the different neurological and mental disorders considered (stroke, SZ and AD) (Chapter 1)
- To develop a novel software tool called SPIDER-NET that allows for the user-friendly qualitative and quantitative analysis of brain networks (Study 1)
- To provide an interactive and flexible management of connectogram visualization and network extraction achieved through SPIDER-NET according to user-defined and well-known human brain atlases to allow for an easy investigation (Study 1)
- To validate SPIDER-NET on multimodal structural data (DTI and CSD) from two stroke case studies compared to a group of HCs (Study 1)
- To investigate robustness of the brain connectivity topological measures and the incidence of the uncertainty in the statistical methodologies (Study 2)
- To compare FC data of schizophrenic patients and HCs with the novel multi-level approach for the extraction of the abnormalities proposed within SPIDER-NET (Study 2)
- To assess the accuracy of DL models (CNNs and GNNs) for AD classification using 3D brain volumes versus brain connectivity data. (Study 3)

- To employ XAI to study interpretability of the DL models and analyze the brain regions whose morphological and interregional features may represent possible brain biomarkers of AD (Study 3)
- To validate the multi-level investigation capability of SPIDER-NET and the utilization of the tool with different population size (case and group studies) of three both focused and widespread degenerations, such as stroke (2 hemorrhagic stroke case studies vs 17 HC), schizophrenia (12 SZ patients vs 15 HC) and Alzheimer's Disease (135 AD vs 557 HC MRI sessions). This PhD work also validated the use of the novel tool SPIDER-NET on different connectivity data, SC in the first and third studies, FC in the second study. (Studies 1-2-3)

2.5 Summary of the Publications

- (I) **Coluzzi, D.;** Pirastru, A.; Pelizzari, L.; Cabinio, M.; Laganà, M.M.; Baselli, G.; Baglio, F.; “Development and Testing of SPIDER-NET: An Interactive Tool for Brain Connectogram Visualization, Sub-Network Exploration and Graph Metrics Quantification”, *Frontiers in Neuroscience*, <https://doi.org/10.3389/fnins.2022.818385>, Special Issue Advanced Computational Tools for Mapping the Multidimensional Architecture of the Brain, 2022.
- In the first work, we developed SPIDER-NET (Software Package Ideal for Deriving Enhanced Representations of brain NETworks), a software package that provides a flexible and user-friendly tool for the selection of partial connectograms, their visualization, and their quantification. The SPIDER-NET Graphical User Interface intuitively allows rapid network exploration and interactive real-time sub-network definition. Figures for connectivity studies are automatically generated, based on the user selections. Furthermore, the toolbox provides additional features to apply matrix thresholding, to easily and automatically compute topological network indices and to interactively define visualization preferences. The potential benefits of using SPIDER-NET in clinical research case studies were tested on a SC dataset composed of two patients with emorrhagic stroke injury characterized by a right hemisphere

lesion with prevalent subcortical expression (males, age 44 and 37 years old, referred to as Case 1 and Case 2, respectively) and 17 HCs subjects (7 males and 10 females; mean age \pm SD: 52.5 ± 8.3 years). Destrieux Atlas was used to define the nodes of the network. The connectios were depicted using different processing pipelines, such as CSD and DTI, which were tested and compared. Specifically, the following aspects were tested: a) providing an effective representation of brain connectivity patterns, b) interactively extracting sub-networks to test a priori hypotheses, specifically evaluating the connectivity between the GM regions affected by the stroke lesions and c) deriving whole-brain quantitative connectivity metrics mirroring local and global topological properties of the graphs.

The main findings found from these qualitative explorations and quantitative analyses through SPIDER-NET were:

- In the case 1 pattern of connectomics is significantly different from the case 2, mirroring the greater clinical severity of the former.
- The pattern of disconnection involved both the right hemisphere, where the stroke lesions were present, and the contralateral one.
- The impairment of the cortical areas of interest determined a decrease in both short-range and long-range connections within the hemisphere ipsilateral to the stroke lesion.
- Interhemispheric connectivity was particularly compromised, probably because subcortical nuclei, which are integration hubs of extrapyramidal systems, were affected by the lesions.
- At a visual inspection, DTI-based and CSD-based sub-network connectograms presented comparable connectivity patterns, highlighting that valuable information provided by both the processing techniques. Furthermore, CSD processing pipeline yielded denser connectograms, as expected. Indeed, CSD ability to better deal with the problem of the crossing fibers when compared to DTI is well-established [28]. This is in line with differences between DTI and CSD that were observed in terms of interhemispheric connections, that were particularly evident for Case 1.

SPIDER-NET resulted an effective and flexible tool to represent and examine the expected (dis)connectivity pattern due to a stroke lesion, in testing a-priori hypothesis by extracting a sub-network of interest and in computing topological indexes. In addition, because of the absence of defined protocols for network creation and edge weighting SPIDER-NET, was demonstrated to be potentially employed as a flexible and simple tool for calibrating connectomics investigations. In particular, it would make it possible to immediately spot anticipated patterns of disconnections and, if any are present, flag significant errors. Before addressing more minor connection changes, this quality check may provide a benchmark. The quantitative results mirrored the differences qualitatively observed with connectograms, regardless the employed modality, providing a comprehensive description of brain connectivity in line with the literature [8], [26]. The tool may also be a useful tool to investigate and interpret individual connectivity patterns during rehabilitation practice.

SPIDER-NET is available online and freely downloadable at <https://caditer.dongnocchi.it/spidernet/>.

- (II) **Coluzzi, D.;** Baselli, G.; “Diffuse and Localized Functional Dysconnectivity in Schizophrenia: a Bootstrapped Top-Down Approach” (*Under Submission*)
- The second work focuses on the comparison of connectivity in small cohorts trying to enhance robustness against uncertainty and noise through the bootstrapping procedure. Data used in this study is composed by the FC matrixes of 15 HC and 12 SZ subjects [12]. These matrixes were formed by 74 AAL regions. Starting from top, the global topological indexes were evaluated to assess the abnormalities in the network topology. In order to enable a robust investigation, these properties were bootstrapped and the stability of the method randomly removing one (RST1) or two subjects (RST2) from the entire pool was evaluated. Second, a sub-network of interest for SZ (DMN) and local indexes were evaluated to visualize significant changes between regions. Third, the connection weights of the group mean matrixes were analyzed to enhance common strongest activations/deactivations and community structure. The

multi-level approach was included in SPIDER-NET, to also perform group study and automatically evaluate abnormalities.

Increased global index values were found in HC for all segregation properties and efficiency, whereas path length is longer in SZ, as expected, as well as small-worldness and coreness statistics. The significance of modularity ($p=0.043$) was not confirmed by BOOT ($p=0.133$). The robustness assessment tests (both RST1 and RST2) highlighted more stable results for BOOT than the direct testing. Furthermore, the use of BOOT highlighted a trend towards studies with a greater number of subjects [191], or of dynamic FC [192]. Significant results were also found at lower levels. First, the analysis of the DMN highlighted a higher variability, reduced connectivity and strength and increased deactivation in the SZ group. At local level, 13 areas were found to be significantly different ($p<0.05$) in the groups, highlighting a greater divergence in the frontal lobe. These results were confirmed analyzing the mean matrix of single negative edges of the groups in the whole-brain, the DMN and the communities detected, suggesting an inverted connectivity between prefronto-temporal areas which can reflect an abnormal inhibition of the activity of these regions. The results highlighted a trend towards results obtained with a greater number of subjects and data appearing to be in line with other studies where the major finding in SZ is that the connectivity is not systematically increased/reduced but generally different, due to the functional reorganization of SZ [36], [47], [191]. However, it remains unclear whether SZ abnormalities are the result of a localized dysconnection exerting widespread effects throughout the brain, or a whole-brain dysfunction that affects certain regions more than others. New investigations would allow to gain insight into the causes and effects related to the onset of the condition. The robust approach for assessing the differences in the network topology resulted to be of great interest, especially when small datasets and static FC are available, and inherent uncertainty can jeopardize the results. Furthermore, the possibility of performing automatic top-down investigations may improve our understanding of both diffuse and localized dysconnections typical of some pathologies.

(III) **Coluzzi, D.;** Bordin, V.; Rivolta, M.W.; Fortel I.; Leow A.; Baselli, G.; “Biomarker Investigation using Multiple Brain Measures from MRI through XAI in Alzheimer’s Disease Classification” (*Under Submission*)

- The third study focused on the 3D T1-weighted scans and extracted SC of 692 selected subjects among all those contained in OASIS-3 dataset [14]. This dataset is a longitudinal collection of data focusing on the effects of normal aging and early-stage AD. In this regard, the work proposes two advanced DL models, which were considered to be suitable for the task, employing the different data: ResNet18 and BC-GCN-SE [20], [39]. The models are evaluated in terms of their classification accuracy, through 10-fold cross-validation and interpretability using a XAI approach called Grad-CAM [21]. The evaluation is conducted across 132 brain parcels extracted from combined HOA and AAL brain atlases and compared to well-known pathological regions to measure adherence to domain knowledge. Indeed, most replicated hallmarks from T1 data and SC (anatomical targets) were investigated. For the former, the involvement of the MTL and, for the latter, the DMN were assessed [23], [24]. More specifically, a statistical test ($p < 0.05$) and a ranking of the most relevant parcels (first 15%) were performed to comprehend the model’s decision.

The results of the evaluation show that both models achieve acceptable classification performance, in comparison to literature. ResNet18 shows a median true positive rate (TPR) of 0.8167 and a median true negative rate (TNR) of 0.8165, whereas BC-GCN-SE a median TPR of 0.7033 and a median TNR of 0.7385. The XAI approach also reveals the involvement of the MTL (7/8 significant parcels) and the DMN (more than 70% significant parcels) in the models, indicating their potential explainable diagnostic relevance. This result was also confirmed by analyzing the most relevant parcels for AD and HC, which show some anatomical targets. However, the study acknowledges that the interpretability of the models has limitations, and further improvements are necessary. Nonetheless, the results suggest that combining different imaging modalities and data can enhance the classification performance and reliability of DL models. Indeed, complementary relevant regions and related features were found by the two models. Several studies have shown that utilizing

prominent characteristics from various measures can be advantageous as it allows for the examination of different information simultaneously, which may be valuable and meaningful [43]. However, there has been limited exploration of the integration of morphological traits of specific brain regions from 3D T1-weighted volumes and interregional properties derived from SC data. This can potentially increase confidence in DL models and promote their widespread use as diagnostic aids. Overall, the study highlights the potential of DL models in advancing the field of medical diagnosis and underscores the importance of interpretability in ensuring their efficacy and applicability.

3 Discussion and Conclusions

In the present PhD work, the aims of enhancing the reliability and the interpretability of the brain connectivity biomarkers were addressed. More specifically, the investigation was carried out on three different datasets formed of different number of subjects and patients, affected by diverse pathologies sharing the characteristics of both affecting specific regions of the brain and leading to diffused degeneration of the whole brain. In the presented work, methods aimed at achieving the ultimate objective of enhancing usability are proposed.

First, a flexible and user-friendly tool called SPIDER-NET was proposed to both qualitatively and quantitatively investigate brain connectivity in health and disease. The validation of SPIDER-NET has demonstrated its effectiveness as a tool for representing the pattern of connectivity changes resulting from a stroke lesion. This was achieved through the testing of a-priori hypotheses, extracting a sub-network of interest, computing relevant topological indexes and comparing the results of DTI to CSD processing pipelines, which were not jeopardized, regardless the different ability in the reconstruction of the crossing fibers and thus denser networks [28]. It was shown that the SPIDER-NET tool can be helpful in clinical research settings, in tailoring the rehabilitative treatments and can greatly contribute to improve the understanding of the brain networks through a better interpretation and reliability of the measures. Normally, networks are composed of hundreds of nodes and thousands of links, thus their analysis and interpretation are often tough. Automatic and interactive tools to assess brain network would be of great support in clinical research. Indeed, analyzing sub-networks of interest, from a-priori hypotheses or through other tools, is of great importance in the understanding of the mechanisms underlying the cognitive processes [193] or the circuit mapping of the brain [194]. SPIDER-NET offers a flexible sub-network extraction method which relies on manual selection of parcels/group-parcels and attributes based on a priori hypothesis testing. However, there are alternative approaches available for automatically identifying relevant sub-networks [12], [195], particularly when there are no apparent gross brain abnormalities. An intriguing avenue for future research would be to incorporate automatic

and data-driven algorithms for sub-network extraction, allowing for comparison with hypothesis-driven selection. Furthermore, it is essential to explore SPIDER-NET's sensitivity in detecting brain connectivity changes by investigating neurological diseases beyond stroke and assessing the impact of treatments such as drugs or rehabilitation. Another promising application could involve examining FC using SPIDER-NET and integrating structural and functional information, leveraging its flexibility in extracting sub-networks. Additionally, there is promising potential for future development in utilizing the SPIDER-NET application to analyze brain connectivity matrices obtained from alternative modalities like EEG, MEG, and NIRS. Although SPIDER-NET was originally presented and tested using MRI datasets, its versatile nature enables its application in diverse contexts, including various -omics disciplines. For example, in the field of rehabilomics [196], which integrates transdisciplinary biomarker evaluation, SPIDER-NET could contribute to defining personalized rehabilitative treatments for patients. In addition, the generation of SPIDER-NET connectograms can serve as a useful approach to evaluate the strength of the processing pipeline, as well as the connectivity metrics, conditioning methods (such as thresholding or binarization), and global or local graph indexes. The absence of standardized protocols for network construction and edge weighting remains one of the primary challenges in connectomics [29], [30], [194]. Therefore, SPIDER-NET can be utilized as a versatile and straightforward method to fine-tune connectomics analyses. Specifically, it could allow to quickly identify the expected pattern of disconnection and to easily highlight major errors if present. This quality check may offer a benchmark before addressing less trivial connectivity alterations, as the ones induced by diffused neurodegeneration, which might be another application field of SPIDER-NET.

In this regard, a bootstrapped top-down approach to reliably assess abnormalities found on different levels considering the inherent uncertainty was proposed. In particular, the assessment of the connectivity from global indexes to single edges (brain activations/deactivations) was carried out on a population characterized by chronic schizophrenia, thought to cause both diffused and focalized dysconnectivity patterns. The multi-level analysis resulted to potentially favor more robust results in contexts where the focus is not well-known and statistical tests can be easily biased by uncertainty. The results obtained through BOOT were more stable in comparison to standard nonparametric tests, being more robust to noise. At the same time, it is also worth noting that is paramount to

focus and apply new and different thresholding methods to assess their influence. Alternative methods such as percolation or spanning tree techniques [47], [101], [113], [114], [134], [197] seem to be less arbitrary compared to conventional density or absolute thresholding approaches. These methods have the potential to further enhance the accuracy of the estimates. The former, for example, was employed in several studies to potentially obtain an optimal deletion of the spurious edges maintaining brain network structure and connectedness [114]–[116], [198]. In this context, it may be worthwhile to explore the relationship between percolation analysis and the findings obtained through bootstrap analysis. Another approach could involve the use of bootstrapping techniques to identify potential errors, generate reliable templates for comparison, and estimate the measure's uncertainty at the connection level. Additionally, the outcomes of this study revealed a trend towards findings obtained with larger sample size and data volume. It is important to acknowledge that a limitation of this study is the small size of the healthy HC and SZ groups available for analysis. To address this, it would be desirable to validate the bootstrap method on graph-based indexes on a larger FC dataset. This validation would provide more robust results and align with findings from larger studies. Additionally, it should be noted that the study lacks discrimination among different types and severity levels of schizophrenia, which is another limitation identified [170]. Furthermore, the statistical analyses conducted using the leave-k-subject-out procedure could be extended to dynamical connectivity data. This would enable an assessment of the influence of the different states assumed during a resting-state functional MRI, thereby strengthening the usability and providing a better understanding of the functional activations and deactivations associated with the approach.

The results were also consistent with prior research indicating that individuals with schizophrenia exhibit distinct patterns of connectivity, as opposed to increased or reduced connectivity, likely due to functional reorganization within the condition [36], [47], [191]. Uncertainty also exists regarding the cause of SZ abnormalities, namely whether they stem from a localized dysconnection that has wide-ranging consequences across the brain or a whole-brain malfunction that disproportionately affects some regions. New investigations would allow to learn more about the factors related to the onset of the disorder and their impact. The robust approach for assessing the differences in the network topology resulted to be of great interest, especially when small datasets and static FC are available. The

possibility of performing automatic top-down investigations in manifold of applications and on different data may improve our understanding of both diffuse and localized dysconnections typical of some pathologies.

Finally, an XAI investigation was performed to compare different MRI data: 3D T1-weighted scans and SC matrixes. Both models showed good classification performance compared to existing literature [38]. However, ResNet18 model using 3D T1-weighted volumes achieved slightly better accuracies than the BC-GCN-SE employing connectivity data when classifying AD and HC sessions. The BC-GCN-SE and, more in general, models adapted to graph-structured data are more recent models than CNNs architectures using images or volumes, such as the well-known pre-trained ResNet18 model employed. Moreover, it should be noted that the SC data obtained from DTI have inherent limitations related to the processing pipelines [30], as abovementioned. Despite this, the variability in the severity of Alzheimer's disease among subjects, the lack of a predefined design setting for multiple acquisition sessions, and the majority of subjects having mild impairments, both results can be seen as promising. In addition, it is essential to evaluate DL models on brain connectivity networks and compare them with more established approaches to assess computational cost of processing. XAI analysis of the models revealed both advantages and disadvantages. The statistical tests and the most relevant parcels ranking showed a good level of agreement with domain knowledge. Hence, ResNet18 and BC-GCN-SE pointed out MTL and several parcels composing the DMN, respectively. These sub-networks represented the anatomical targets addressed. However, one may argue that the number of significances found by the statistical test is high if compared to the number of parcels investigated, increasing the possibility to detect the target regions. For this reason, we investigated which significant parcels were also among the 20 most relevant for the AD case reinforcing the interpretation of the results. Moreover, this high number of parcels was found to be composed of most cortical parcels, that could be traced back to the different characteristics of the population. The presence of subjects at various AD stages and severities could lead to general extensive importance of a considerable portion of the cortex. Then, results underlined by BC-GCN-SE model appear to be in agreement to an increasing evidence of topological asymmetry between hemispheric brain WM in AD found in two studies [40], [41]. This may represent a useful insight into hemispheric lateralization and aberration possibly resulting from long-range connection loss. The

present study is not without limitations. Firstly, it is important to recognize that different approaches for XAI may not always be effective in all scenarios [199]. Therefore, it would be valuable to incorporate and compare alternative methods, such as perturbation or distillate techniques, to further validate the obtained results. Additionally, this study primarily focuses on the mean relevance value extracted from whole parcels, enabling direct comparison to 3D T1-weighted volumes. However, in the case of SC data, it would be intriguing to investigate the most explainable connections, which could shed light on crucial long-range connections and their disruption in AD [183], [184]. Moreover, beyond the analysis of the DMN, exploring other RSNs could provide novel insights into the employed models. Further investigations into WM could also be conducted using the 3D T1-weighted volumes to visualize similarities and differences concerning WM metrics associated with SC data. A particular consideration is the presence of WM Hyperintensities, which can serve as important biomarkers of the AD condition [200]. In a previous preliminary study with a limited number of subjects [141], their relevance was confirmed using a DL model employing FLAIR images. It would be worthwhile to explore their effects within the connectivity data as well. For instance, employing tools like NeMo it would be of great interest to extract the impact of these lesions within the connectivity data [201]. In general, the concurrent investigation of the models led to valuable indications. While both models showed agreement in identifying anatomical targets, there were some key regions in the MTL that were not highlighted in the XAI analysis of BC-GCN-SE, despite being important in AD connectivity data as well [42]. With the exception of the right amygdala, other regions of the MTL were not identified through statistical testing or examination of relevant parcels. These findings could mirror a limitation of BC-GCN-SE interpretability or the effect of noise sources inherently present in these DTI-derived data. At the same time, this complementary importance of different parcels appears promising in the perspective of yielding more accurate and reliable models. In this context, previous studies have not extensively explored the idea of combining morphological features with interregional properties, although the potential demonstrated in the study by Liu and colleagues [43]. This may have significant impacts on the diagnosis of various pathologies. Indeed, by incorporating both types of data, a more comprehensive understanding of the underlying mechanisms can be gained leading to the development of

better and more interpretable models. This is particularly important given the limited trust in "black-box models" in real-world settings.

In general, numerous valuable brain biomarkers with specific reference to those extracted from connectivity data were detected, investigated, and employed through different methods, DL approaches included. Although some limitations such as the use of different brain atlases among the three studies or the population size, they revealed a great potential for several purposes and in a manifold of applications ranging from diagnostics to rehabilitation. However, they appear to be still lacking reliability and interpretation, which are the issues that limited their diffusion within clinical settings. The proposed PhD thesis addressed these problems developing and validating novel methods in the context of pathologies whose affection is both global and local. However, it will be paramount to improve them with the purpose of increasing the level of trust of brain connectivity biomarkers for enhanced usability. The robust extraction of the abnormalities has a direct importance and connection to the understanding of the relationships between the localized affection and widespread degeneration. Novel and reliable methods for a consistent analysis of the brain connectivity biomarkers can indeed be of unvaluable importance in the study of the brain both in health, to map the nervous system and comprehend the mechanisms underlying the mental processes, and in disease, to support clinicians in the early detection and during recovery.

Annex



Development and Testing of SPIDER-NET: An Interactive Tool for Brain Connectogram Visualization, Sub-Network Exploration and Graph Metrics Quantification

OPEN ACCESS

Edited by:

Jesús Poza,
University of Valladolid, Spain

Reviewed by:

Thomas Bolton,
Advanced Telecommunications
Research Institute International (ATR),
Japan
Emin Serin,
Charité University Medicine Berlin,
Germany
Chuyang Ye,
Beijing Institute of Technology, China

*Correspondence:

Davide Coluzzi
davide.coluzzi@polimi.it
Alice Pirastru
apirastru@dongnocchi.it

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 19 November 2021

Accepted: 17 February 2022

Published: 17 March 2022

Citation:

Coluzzi D, Pirastru A, Pelizzari L,
Cabinio M, Laganà MM, Baselli G and
Baglio F (2022) Development
and Testing of SPIDER-NET: An
Interactive Tool for Brain
Connectogram Visualization,
Sub-Network Exploration and Graph
Metrics Quantification.
Front. Neurosci. 16:818385.
doi: 10.3389/fnins.2022.818385

**Davide Coluzzi^{1*}, Alice Pirastru^{2*}, Laura Pelizzari², Monia Cabinio²,
Maria Marcella Laganà², Giuseppe Baselli¹ and Francesca Baglio²**

¹ Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy, ² IRCCS Fondazione Don Carlo Gnocchi Onlus, Milan, Italy

Brain connectomics consists in the modeling of human brain as networks, mathematically represented as numerical connectivity matrices. However, this representation may result in difficult interpretation of the data. To overcome this limitation, graphical representation by connectograms is currently used via open-source tools, which, however, lack user-friendly interfaces and options to explore specific sub-networks. In this context, we developed SPIDER-NET (Software Package Ideal for Deriving Enhanced Representations of brain NETWORKS), an easy-to-use, flexible, and interactive tool for connectograms generation and sub-network exploration. This study aims to present SPIDER-NET and to test its potential impact on pilot cases. As a working example, structural connectivity (SC) was investigated with SPIDER-NET in a group of 17 healthy controls (HCs) and in two subjects with stroke injury (Case 1 and Case 2, both with a focal lesion affecting part of the right frontal lobe, insular cortex and subcortical structures). 165 parcels were determined from individual structural magnetic resonance imaging data by using the Destrieux atlas, and defined as nodes. SC matrices were derived with Diffusion Tensor Imaging tractography. SC matrices of HCs were averaged to obtain a single group matrix. SC matrices were then used as input for SPIDER-NET. First, SPIDER-NET was used to derive the connectogram of the right hemisphere of Case 1 and Case 2. Then, a sub-network of interest (i.e., including gray matter regions affected by the stroke lesions) was interactively selected and the associated connectograms were derived for Case 1, Case 2 and HCs. Finally, graph-based metrics were derived for whole-brain SC matrices of Case 1, Case 2 and HCs. The software resulted effective in representing the expected (dis) connectivity pattern in the hemisphere affected by the stroke lesion in Cases 1 and 2. Furthermore, SPIDER-NET allowed to test an *a priori* hypothesis by interactively extracting a sub-network of interest: Case 1 showed a sub-network connectivity pattern different from Case 2,

reflecting the different clinical severity. Global and local graph-based metrics derived with SPIDER-NET were different between cases with stroke injury and HCs. The tool proved to be accessible, intuitive, and interactive in brain connectivity investigation and provided both qualitative and quantitative evidence.

Keywords: MRI, brain networks, connectograms, brain connectivity, graph analysis, stroke

INTRODUCTION

In the last decades, the emergence of -omics disciplines led to the development of flexible and comprehensive methods to effortlessly analyze big data sets. Graph theory is a suitable means for leveraging big data and for modeling complex real-world systems, characterized by specific architecture and topology. This mathematical approach has been effectively applied in several scientific fields. One of the most impressive and popular application is the so called “human connectome” (Bullmore and Sporns, 2009), namely modeling the human brain as a network on many different scales (Rubinov and Sporns, 2010), aiming to connect its structure to function and behavior. Similar to network genomics, which models the influence of genes in a larger biomolecular system, brain connectomics reconfigures the study of brain structure and function by mapping the whole brain in terms of neural units and their connections (Sporns et al., 2005). Indeed, brain regions are strongly connected through neuroanatomical white matter (WM) pathways, intuitively determining a complex system. In parallel to structural connectivity (SC), synchronous and asynchronous activity of specific brain regions results in related complex cognitive functions, which can be investigated in terms of functional connectivity (FC). Exploring SC and FC patterns can provide insight of brain function both in physiological and pathological conditions.

A network is a mathematical representation of a complex system that is defined by a collection of nodes (vertices) and links (edges), describing any kind of relationship between pairs of nodes, at different scales. Networks can be easily represented as n -by- n association matrices, where n is the number of nodes composing the network, while each element e_{ij} represents the link connecting the nodes i and j . The elements of the matrix can be either binary (i.e., describing the presence/absence of links between pairs of nodes) or weighted (i.e., describing the strength of the links between pairs of nodes). In the framework of the human connectome, association matrices are brain connectivity matrices. According to the technique or imaging modality employed to extract the connectivity data, nodes and edges of a brain network can represent different concepts. When constructing brain connectivity matrices from magnetic resonance imaging (MRI) dataset, nodes usually represent gray matter parcels, defined according to well-known atlases (Tzourio-Mazoyer et al., 2002; Desikan et al., 2006; Smith et al., 2009; Yeo et al., 2011). Brain atlases segment the brain into sets of voxels (i.e., parcels), based either on anatomical or functional criteria. Similarly, the edges of a brain network, describing a relationship between nodes, can depict either SC or FC features. SC refers to anatomical associations between neural elements or

brain regions, while FC represents the magnitude of temporal correlations between the signal produced by pairs of brain regions. SC and FC can be quantified with various indices, depending on the imaging modality which is used to investigate the connectivity pattern. For instance, the number of streamlines derived with deterministic WM tractography can be used as weights in an MRI-derived SC matrix, while correlation between blood oxygenation level-dependent (BOLD) time series can be used to define edges of MRI-derived FC matrices. SC and FC generally mirror an undirected relationship between brain regions (i.e., non-causal), resulting in a symmetric connectivity matrix [i.e., $(N \times (N-1))/2$ pairwise connections between N nodes].

Although brain connectivity matrices can exhaustively and quantitatively outline the human connectome, this representation does not always provide an intuitive and direct visualization of the connectivity pattern. Brain connectivity matrices are generally too large to be visually interpreted, thus important information might remain concealed. For this reason, conceiving new methods for the visualization of connectivity data is needed to aid the interpretation of brain connectivity measures. This is important especially for explorative analyses, with the aim of identifying characteristic patterns that may allow to distinguish the pathological condition from the physiological one, or to assess changes after a pharmacological treatment or rehabilitation.

Connectograms are graphical representations that meet these needs, bridging the gap between quantitative connectivity analyses and intuitive visualization. Connectograms are circular graphs in which all the nodes of a network are represented along the perimeter of the circle, while the edges of the network are shown as arcs connecting pairs of nodes. This layout was previously used in other fields (e.g., genomics) and it was introduced for brain connectivity mapping by Irimia et al. (2012b) about 10 years ago. Connectograms can be produced using Circos (Kryzwiniski et al., 2009), which is a powerful software package designed for visualizing data, for exploring relationships between objects and for creating publication-quality illustrations. It is extremely flexible, and it can be used in several diverse fields. However, Circos has no interface, and it has to be run by command-lines. This approach does not create any problem to brain connectivity researchers who are UNIX users, but it may be uncomfortable to researchers who do not have any programming experience. A user-friendly interface that could be used by people having knowledge and interest in brain connectivity but not in computer science may broaden the accessibility to connectograms.

A complete whole-brain network can be made of thousands of links, and it is well-known that this large-scale network is associated with high-level cognitive functions. However,

the brain is composed of several interacting lower-scale sub-networks, which are characterized by distinct patterns of brain activation, identifying specific domains of behavior and cognition (Bassett and Sporns, 2017). Therefore, extracting sub-networks is common practice in explorative studies of brain connectivity, both in physiological and pathological conditions (Zalesky et al., 2010; Bassett and Sporns, 2017; Berron et al., 2020; Isernia et al., 2021). Indeed, focusing on sub-networks can lead to easier data interpretation driven by the addressed physiological and/or pathological problem. Sub-network should be analyzed both qualitatively, by a reduced connectogram, and quantitatively by local and global subgraph indexes. Nevertheless, software that are currently available for connectivity pattern visualization (e.g., Circos, BrainNet Viewer, Xia et al., 2013) does not allow for the interactive selection of some specific nodes within a whole-brain network, as the direct upload of the sub-network of interest is generally required. Beside outlining the brain connectome through an association matrix and finding an intuitive way to effectively represent it, graph-based network properties can be calculated to depict a complete picture of the architecture of the whole network and of each sub-network of interest. Indices such as node degree, small-worldness, modularity, clustering, and central hubs, can add meaningful information about the network topology. These are valuable pieces of information in a brain connectivity analysis as many studies revealed changes of organizational and topological properties in a number of brain disorders, such as Alzheimer's Disease (Daianu et al., 2013), mild cognitive impairment (Baggio et al., 2014), Parkinson's Disease (Göttlich et al., 2013), epilepsy (Ji et al., 2017), autism (Barttfeld et al., 2012) and borderline intellectual functioning (Blasi et al., 2020). Despite the large evidence produced in the last years, gold-standard methodologies are not established yet, and connectivity alterations in neurological and neurodegenerative diseases are an open issue, so far. Therefore, further investigations on human brain networks are warranted. Accessible tools for easily assessing the topology and architecture of brain networks would provide larger amount of evidence that may promote deeper knowledge of brain disorders and of the effect of treatments (e.g., disease modifying therapies, rehabilitation) on brain networks.

In this framework, we developed SPIDER-NET (Software Package Ideal for Deriving Enhanced Representations of brain NETworks), a software package that provides a very flexible and user-friendly tool for the selection of partial connectograms, their visualization, and their quantification. The SPIDER-NET Graphical User Interface (GUI) intuitively allows rapid network exploration and interactive real-time sub-network definition. Figures for connectivity studies are automatically generated, based on the user selections. Furthermore, the toolbox provides additional features to apply matrix thresholding, to easily and automatically compute topological network indices and to interactively define visualization preferences. The aims of this study were: (1) presenting SPIDER-NET, and (2) testing the potential benefits of using SPIDER-NET in clinical research case studies. Specifically, the following aspects were tested: (2a) providing an effective representation of brain connectivity patterns, (2b) interactively extracting sub-networks to test *a priori* hypotheses and (2c) deriving whole-brain

quantitative connectivity metrics mirroring local and global topological properties.

MATERIALS AND METHODS

SPIDER-NET Overview

SPIDER-NET was developed in Matlab but is delivered as a standalone software (.exe in Windows, .app in macOS, .sh in UNIX). The tool allows flexible and effective representation of brain networks through connectograms. It enables the exploration of network architecture and topology and, optionally, the extraction of topological properties describing the network architecture and nodes properties. A schematic flowchart is shown in **Figure 1**.

SPIDER-NET Inputs

SPIDER-NET requires 3 input files, which are an Atlas file, a Label file and a Connectivity Matrix file.

- 1) The Atlas file is an XSL/XSLX Excel worksheet that provides information on the atlas the user adopts to define the network nodes. The list of the Atlas parcels is reported in a column of the worksheet. All the parcels listed in the Atlas file are reported as nodes in the connectogram generated by SPIDER-NET. The sorting of the parcels in the Atlas file determines the positions of the parcels in the connectogram obtained with SPIDER-NET (i.e., the first parcel is represented on the top of the circle). A short legend has to be associated to each parcel in the Atlas file. The legend is shown in the interface to help interactive node selection. Reporting optional parcel grouping (i.e., Group Parcellation) is also allowed in the Atlas file (e.g., brain lobes, resting state networks). In additional columns (i.e., Attribute), the Atlas file can enclose additional optional attributes associated with each parcel (e.g., functional attribute). Both Group Parcellation and Attributes can be used to rapidly select entire groups of parcels (i.e., parcels sharing Group Parcellation tag or Attribute tag) for sub-network extraction. Several Atlas files, for the most used structural and functional atlases, are provided as templates together with the software. Moreover, it is possible to customize or create new Atlas files according to the user's preferences and aims, by simply composing new worksheets.
- 2) The Label file (ASCII text file, .txt) is a list of the parcel names. The order of parcels in this file must strictly repeat the order of rows and columns in the Connectivity Matrix, which is the third input for SPIDER-NET. Therefore, the order of the parcels in the Label file is tied to the Connectivity Matrix generation. The design choice of repeating the same list of parcels in both the Atlas file and in the Label file permits great flexibility for connectogram generation, as the same Atlas file can be used with many different Label files (and Connectivity matrices). Indeed, all the parcels being equal, the order of parcels can vary across Label files (according to the associated Connectivity

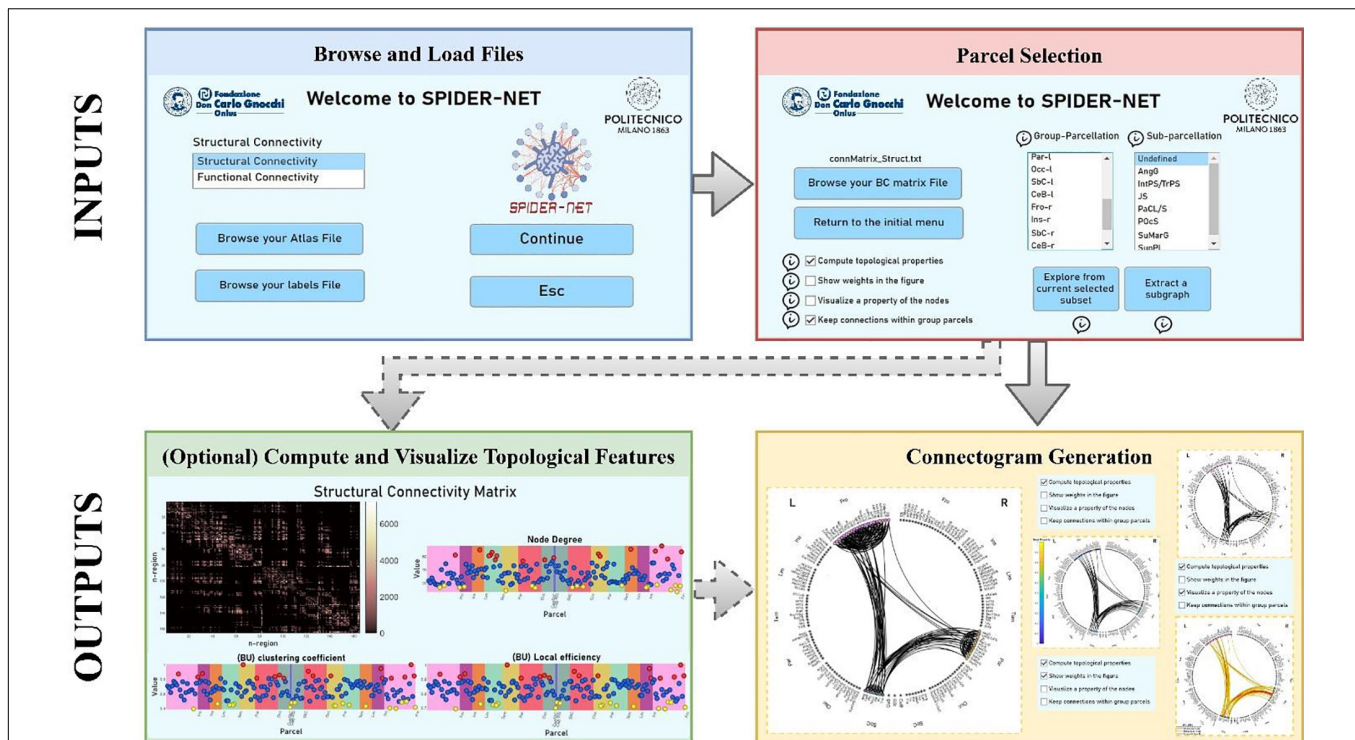


FIGURE 1 | Flowchart for SPIDER-NET usage. First, the Atlas and Label input files are browsed and loaded (blue box). Then, the Connectivity Matrix file is loaded and the selection of the sub-network of interest is performed (red box). Optionally (dashed lines), it is possible to compute and visualize topological properties of the selected brain network (green box). Finally, the connectogram is generated according to the selection made and to the chosen visualization settings (yellow box).

matrices), while the sorting of the parcels in the circular representation remains the same if the same Atlas file is used for generating the connectograms.

- 3) The Connectivity matrix file is an ASCII text file, containing the matrix of association weights with row and column order matching with the Label file. Any measure of SC or FC derived with MRI is applicable. The input connectivity matrix must be square ($N \times N$) and symmetric. Conventionally, the main diagonal must be set to zero.

Once the inputs have been uploaded (**Figure 1**, blue box), the selection of either single parcels, entire group-parcels or attributes defined in the “Atlas” file, is enabled in the SPIDER-NET GUI (**Figure 1**, red box).

Parcel Selection and Connectogram Generation

Two complementary logics are offered by SPIDER-NET in the interactive definition of a partial network out of the global one represented by the input Connectivity matrix (**Figure 1**, red box): “Explore from current selected subset” (Option 1) and “Extract a subgraph” (Option 2).

When the interactive definition of the addressed sub-network is completed on the GUI, with either Option 1 or Option 2, SPIDER-NET generates the partial connectogram (**Figure 1**, yellow box). The figure of the connectogram was designed by reengineering and extending the circular graph package

developed for Matlab (GitHub. Retrieved July 28, 2021).¹ This library allows to draw nodes along a circumference, and their connections, whose shape is defined by the Poincaré hyperbolic disk (Gao et al., 2020).

The connectogram figure generated by SPIDER-NET is automatically saved in a folder created at run-time for each execution. However, interactive changes to the resulting connectogram are also possible within the software figure-management GUI. Specifically, changes improving readability of too crowded diagrams are available: (i) single nodes can be selected to hide/show the respective labels; (ii) connections related to specific nodes can be temporarily removed from the connectogram. The modified figures can be saved in addition to the original one.

Option 1 – Explore From Current Selected Subset

One or more “seed” parcels are defined. Next, the set of target parcels is defined, which can be either all the parcels of the brain, just some other parcels or a specific group of parcels (i.e., defined by Group Parcellation or any additional Attribute defined in the Atlas file). It is worth noting that in the set of target parcels is also possible to include the parcels already selected as “seed.” Only the edges between each seed-target pair are represented in the resulting connectogram. SPIDER-NET deals with non-directional graphs, so “seed” and “target” are fully conventional

¹<https://github.com/paul-kassebaum-mathworks/circularGraph>

names. This option can be useful for the analysis of alterations due to focal lesions or for qualitative pilot quality control of the processing pipeline, to check for major errors in the connectivity matrix generation. For instance, this option can be used in a preliminary quality check of structural connectivity data verifying the presence of connections between a chosen seed and all the other parcels which are linked to it by existing WM tracts, basing upon anatomical knowledge. This can be particularly valuable given the complex image acquisition and upstream processing.

Option 2—Extract a Subgraph

This option is based on the selection of a single subset of nodes (i.e., brain parcels). The user can select either parcels one by one, or entire groups of parcels, defined according to Group Parcellation or any additional Attribute reported in the Atlas file. Therefore, after the selection, a subgraph is defined with respect to the original Connectivity matrix, and connections between all the possible pairs of the selected parcels are shown in the resulting connectogram. This kind of exploration is useful when specific cerebral circuits are addressed or deeper verification of single well-known connections in the quality control of the pipeline are requested.

Additional Features

Although SPIDER-NET has been designed as an easy-to-use software GUI for simple connectogram generation, additional optional features were developed and can be set simultaneously to parcel selection (**Figure 1**, red box).

First, if the parcels are grouped according to a higher-level classification (e.g., brain lobes) in the Atlas file, links between pairs of parcels belonging to the same group (e.g., Group Parcellation) can be optionally omitted from the connectogram. Excluding within-group links from the connectogram enables a clearer visualization of long-range connections, especially when many connections are displayed.

Second, SPIDER-NET allows to optionally visualize color-coded properties of the nodes, which can be either local graph-theory based (e.g., node degree) or representing other properties of the parcels (e.g., cortical thickness, parcel volume, classification according to functional circuits). Edge properties (i.e., strength of connection) may also be color-coded.

Furthermore, density-based thresholding is commonly applied to matrices in brain connectivity studies, to either remove spurious connections and/or to binarize a weighted matrix. Although the user can provide SPIDER-NET with an already thresholded connectivity matrix, the software is designed to allow for optional density-based thresholding at run-time. In particular, once the user has selected the desired density, the software iteratively searches for the best threshold to approximate the selected density, starting from zero. The thresholded matrix is then used to draw the connectogram.

Compute and Visualize Topological Features

Another feature optionally implemented by SPIDER-NET (**Figure 1**, green box) is the computation of graph-based topological indices for a quantitative assessment. Local, global,

and intermediate structure (i.e., community detection, core-periphery analysis, rich-clubs) analyses are performed for a total number of 20 computed indices, computed basing on Brain Connectivity Toolbox (Rubinov and Sporns, 2010).² Importantly, the computed indices refer to both the original Connectivity Matrix and to the currently selected subset of nodes, thus providing targeted quantification of cerebral circuits. Besides the connectogram, the following quantitative information is graphically represented when graph-based topological indices are optionally computed:

1. Connectivity weights, shown as a color-coded connectivity matrix.
2. Local indexes (i.e., node degree, clustering coefficient, local efficiency for both the binary and weighted case), shown in plots. The horizontal axis reports the parcel names in the same order as around the connectogram, while the vertical axis scales the local index values. The local index value for each parcel is represented with dots. The index value is color-coded to highlight the most and the least influential nodes. Namely, the top 10% parcels and the least 10% ones are highlighted in red and yellow, respectively. These plots can be effective in pinpointing network hubs or, conversely, lesion related drops.
3. Values of the global indices, listed below the plots of local indices.

These graph-based outputs can be interactively explored, selecting specific elements of the connectivity matrix or dots of local indexes to obtain additional information (i.e., weights, corresponding higher-level classification). It is worth remarking that results of the interactive subgraph analyses are always shown on the screen in parallel to results of the analysis of the original complete connectivity matrix. Once the interactive process is fulfilled, whole graph and subgraph results are saved.

SPIDER-NET Application on Case Studies

Participants

The dataset consists of two patients with stroke injury characterized by a right hemisphere lesion with prevalent subcortical expression (males, age 44 and 37 years old, referred to as Case 1 and Case 2, respectively) and 17 healthy control (HCs) subjects (7 males and 10 females; mean age \pm SD: 52.5 \pm 8.3 years). All the subjects were enrolled at IRCCS Fondazione Don Carlo Gnocchi in Milan and signed a written informed consent.

Magnetic Resonance Imaging Acquisition and Matrix Construction

All the participants underwent a MRI examination performed on a 1.5 T Siemens Magnetom Avanto scanner equipped with a 12-channels head coil. Both patients with stroke injury were scanned six months after hemorrhagic stroke.

The acquisition protocol included:

²<https://sites.google.com/site/bctnet/>

1. a high-resolution 3D T1-weighted Magnetization Prepared Rapid Gradient-Echo (MPRAGE) image, (repetition time (TR)/echo time (TE) = 1,900/3.37 ms, Field of View (FoV) = $192 \times 256 \text{ mm}^2$, resolution = $1 \times 1 \times 1 \text{ mm}^3$, 176 axial slices);
2. a diffusion-weighted echo planar images (EPI) image along 64 directions (b -value $1,500 \text{ s/mm}^2$, TR/TE 7,800/109 ms, matrix size = $102 \times 102 \times 46$, resolution = $2.5 \times 2.5 \times 2.5 \text{ mm}^3$) and 3 b0 images (2 with AP, and 1 with PA encoding direction);
3. a dual-echo turbo spin echo proton density PD/T2-weighted image (TR = 4,540 ms, TE = 28/112 ms, matrix size = $320 \times 320 \times 60$, resolution = $0.75 \times 0.75 \times 2 \text{ mm}^3$).

After standard preprocessing, 3D T1-weighted volumes were parcellated, at subject-level, and automatically labeled into 75 cortical parcels for each hemisphere (150 in total) according to the Destrieux atlas (Destrieux et al., 2010) using FreeSurfer (version 6). Seven subcortical regions per hemisphere (thalamus, caudate, putamen, pallidum, nucleus accumbens, amygdala and hippocampus) and the brainstem were also segmented using the FreeSurfer automatic labeling process (Fischl et al., 2002) for a total of 165 parcels.

Diffusion-weighted images were preprocessed using the FMRIB's Software Library (FSL) tools with a standard pipeline (i.e., correction for susceptibility-induced geometric distortions, for eddy current distortion and head movements) (Pelizzari et al., 2019) and diffusion tensor imaging (DTI) was estimated for each voxel using the FSL DTIFIT toolbox (Andersson et al., 2003; Behrens et al., 2003; Andersson and Sotiropoulos, 2016). Then, DTI-derived whole brain tract was generated. In addition, diffusion weighted data were processed also with Constrained Spherical Deconvolution (CSD) approach. Specifically, StarTrack³ was used both to estimate the fiber orientation distribution function and to perform subsequent deterministic whole brain tractography, according to high angular resolution diffusion imaging (HARDI) processing (Dell'Acqua et al., 2010).

Cortical and subcortical parcels, obtained from the 3D T1-weighted images, were registered to the respective diffusion-weighted space using the FSL flirt toolbox (Jenkinson et al., 2002). Then, for each subject, WM tracts connecting each pair of registered parcels were reconstructed with TrackVis software,⁴ basing both on DTI-derived whole brain tract and CSD-derived whole brain tract.

In both patients, the stroke lesions were segmented by an experienced operator on the PD/T2 volumes with Jim software.⁵

DTI-based and CSD-based SC matrices were derived for patients with stroke injury and HCs, by computing the edges as the number of the reconstructed fiber (NF) of each WM tract connecting each pair of the 165 parcels. In order to account for differences in brain volumes, NF was normalized by the sum of the volumes of the pair of respective connected parcels

(Blasi et al., 2020). A probabilistic group matrix was computed to represent the HC group as a whole, retaining only the connections shared by at least 53% of the HCs subjects (Blasi et al., 2020). Therefore, three matrices were finally obtained for both DTI and CSD approach: one for each patient with stroke injury and one HC group matrix. The obtained matrices were then normalized by the respective maximum edge value, so that the matrix elements ranged from 0 to 1.

Running SPIDER-NET

SPIDER-NET inputs were defined as follows. The Atlas file was constructed based on the Destrieux atlas.

The Label file, matching the connectivity matrices, reported the Destrieux atlas parcel names. The normalized connectivity matrices for Case 1, Case 2 and HCs (Connectivity matrix files) were uploaded one at a time, for each separate analysis.

First, for the two patients with stroke injury, DTI-based connectograms showing the connectivity pattern in the right hemisphere, where the lesions of both subjects are located, were generated to test the ability of SPIDER-NET to show altered connectivity patterns due to a focal lesion.

Second, for each patient with stroke injury, a DTI-based subgraph focused on expected altered circuit was extracted. Specifically, all the parcels overlapping with the stroke lesions were selected as seeds, while all the remaining brain parcels were set as target for this sub-network analysis. The same sub-network was investigated for the two stroke cases and for the HC group to enable a comparison. The following regions of interest were considered as seeds: right precentral gyrus, right long insular gyrus and central insular sulcus, right short insular gyri, right caudate nucleus, right pallidum, right putamen, and right thalamus. This analysis was performed to test the ability of SPIDER-NET to visually explore sub-networks of interest. Furthermore, the same sub-network analysis was performed for connectivity matrices derived with CSD processing, to compare the connectivity results based on DTI and CSD processing techniques.

Third, the main local and global graph-analysis indices describing network topology were extracted both for the weighted and for the binary connectivity DTI-based matrices. The node degree was the main focus relevant to local graph index analysis.

RESULTS

Connectogram Visualization of the Connectivity Pattern Altered by Stroke Lesions

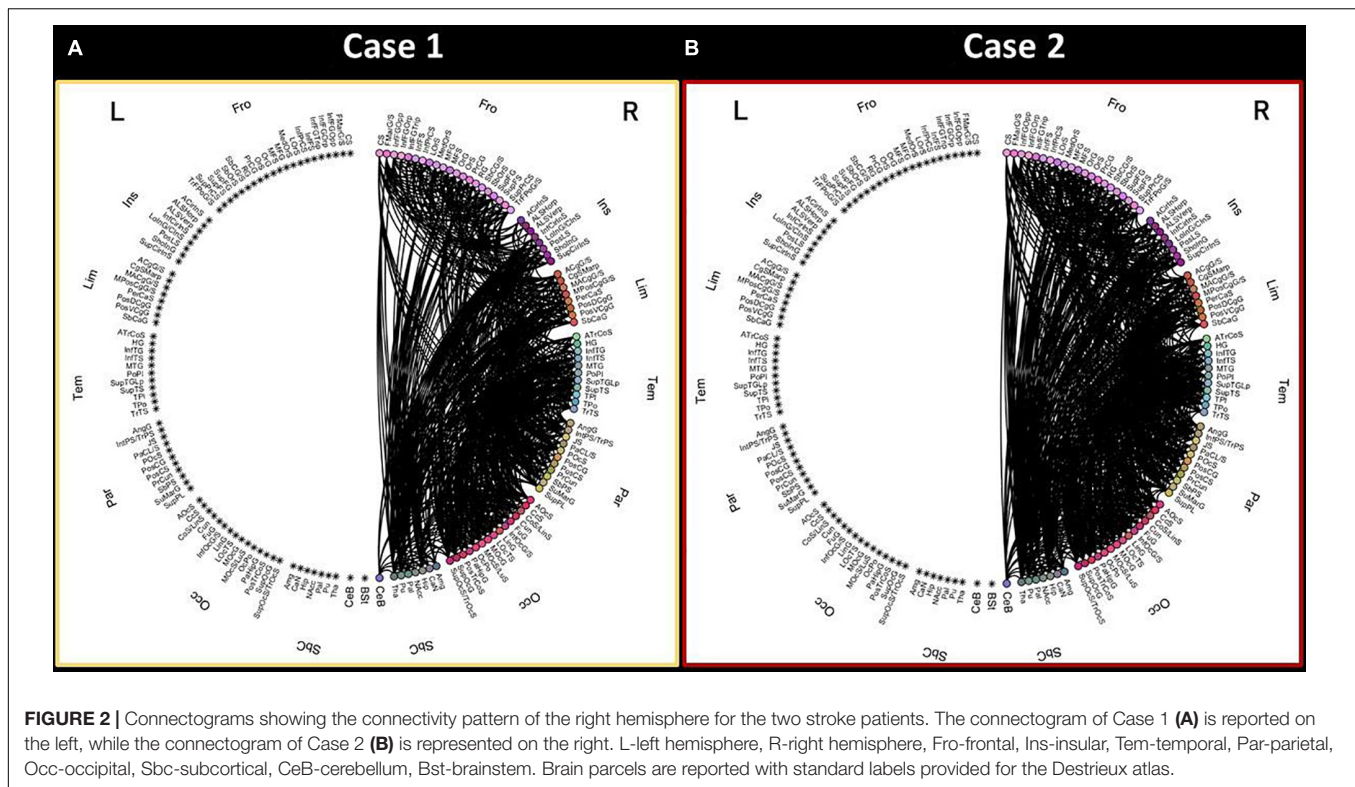
The connectograms showing the DTI-based connectivity pattern of the right hemisphere of the two patients with stroke injury are shown in **Figure 2**. All the 165 parcels (cortical parcels of Destrieux atlas and subcortical regions) are reported in the circular representation and divided in 8 anatomical lobes (i.e., Group Parcellation defined in the Atlas file) per hemisphere.

Upon visual inspection, the connectivity pattern of the right hemisphere is different in Case 1 compared to Case 2. Specifically,

³www.natbrainlab.co.uk

⁴<http://trackvis.org/>

⁵<http://www.xinapse.com/>



Case 1 shows a less dense connectivity pattern in the right hemisphere, especially in terms of connections among the frontal lobe, insular cortex and subcortical structures.

Connectograms Visualization for Sub-Network Analysis

The connectograms generated with SPIDER-NET to explore the DTI-based connectivity between gray matter parcels intersecting the lesions of patients with stroke injury (**Figure 3A**) and the rest of the brain are shown in **Figure 3B**. The same connectograms obtained with Circos software⁶ are reported in **Supplementary Material** to allow for comparison.

Upon visual inspection, the sub-network connectivity pattern of both patients with stroke injury looks altered compared to HCs. In addition, it is worth noting that differences can be qualitatively observed both in the right hemisphere (where the stroke lesions are present) and in the contralateral one. Case 1 displays a less dense right hemisphere connectivity pattern compared to Case 2. On the other hand, the sub-network connectogram of Case 2 qualitatively shows less connections in the left frontal-insular area and in the left parietal-occipital lobes compared to Case 1.

Diffusion Tensor Imaging-Based and Constrained Spherical Deconvolution-Based Connectivity: Visual Comparison

The connectivity between regions overlapping the lesions of stroke patients and the whole brain was also investigated

from matrices derived with CSD processing. The resulting connectograms are shown in **Figure 3B**.

The sub-network connectograms of both Case 1 and Case 2 looks different when compared to the HC group one, as for the DTI-based connectograms. At visual inspection, the connectograms derived from DTI and CSD processing generally preserve the same connectivity patterns. CSD-based connectograms highlight the difference between Case 1 and Case 2 connectivity pattern.

Local and Global Topological Properties Analysis

Node degree explorative figures produced by SPIDER-NET for Case 1, Case 2 and HCs are shown in **Figure 4**. In general, patients with stroke injury presented lower node degrees when compared with HCs. In HCs, regions with the highest node degrees, represented as red dots in **Figure 4** (upper panel), were mostly located in the dark-green vertical stripe, representing subcortical regions. Conversely, in the patients with stroke injury, both characterized by a right hemisphere lesion with prevalent subcortical expression, the regions showing the highest local node degree are more distributed across the cortical lobes. In addition, caudate nucleus, pallidum, putamen, and thalamus, which were classified by SPIDER-NET as regions with the highest node degree in HCs (red dots), were not classified as nodes with high node degree in both Case 1 and Case 2. In **Figure 4**, putamen node degree values are highlighted for HCs, Case 1 and Case 2. Patients with stroke injury presented lower network density (18.65 and 22.32%, respectively, for Case 1 and Case 2) with

⁶<http://circos.ca/>

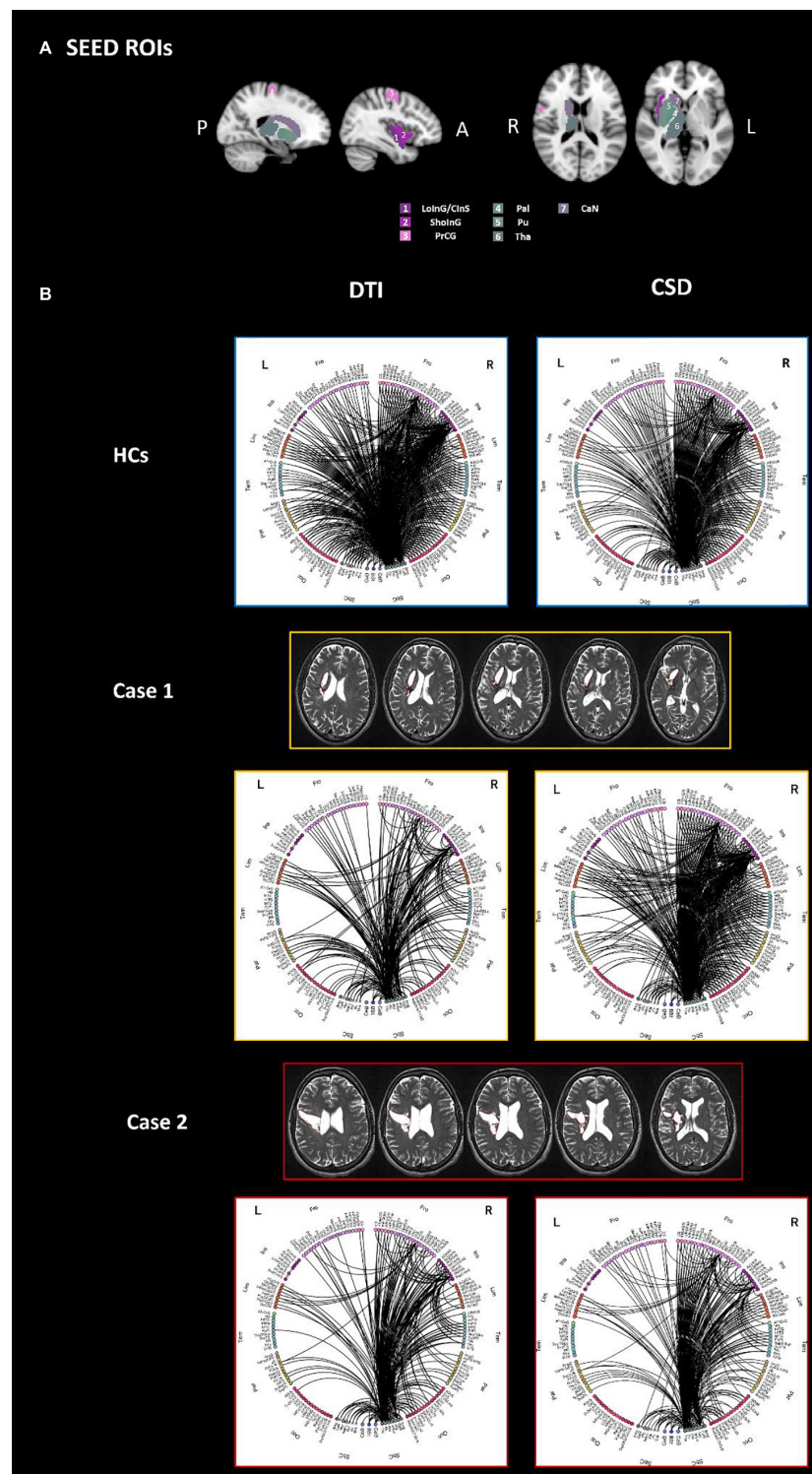


FIGURE 3 | The connectograms (**B**) derived by both DTI (left) and CSD (right) processing, of the sub-network extracted using as seeds all parcels which are overlapped with the stroke lesion of either Case 1 or Case 2, are reported for HCs (top panel), Case 1 (middle panel) and Case 2 (bottom panel). Specifically, seeds were defined as follows: right LoinG/ClnS, right SholnG, right Pal, right Pu, right CaN, right PrCG, right Tha (**A**). All the other parcels of the brain were considered as target for the connectivity analysis. L-left hemisphere, R-right hemisphere, Fro-frontal, Ins-insular, Tem-temporal, Par-parietal, Occ-occipital, Sbc-subcortical, CeB-cerebellum, Bst-brainstem, PrCG-precentral gyrus, LoinG/ClnS-long insular gyrus and central insular sulcus, SholnG-short insular gyri, Pal-pallidum, Pu-putamen, CaN-caudate nucleus, Tha-thalamus.

respect to HC (47.51%), resulting in differences of 28.9 and 25.2%. Global topological binary and weighted indices extracted from the whole-brain network of HC, Case 1 and Case 2 are shown in **Table 1**.

Both Case 1 and Case 2 presented differences for all the global topological indices when compared to HCs. Percentage differences ranged from 10.5 to 96.9% for Case 1, and from 14.3 to 81.9% for Case 2. Except for clustering coefficient, Case 1 presented with greater percentage difference with HCs when compared to Case 2.

DISCUSSION

In this study we presented SPIDER-NET, an innovative tool for exploring and visualizing brain connectivity through full and partial connectograms. The tests on two readily interpretable cases with stroke injury proved that this tool is capable of producing meaningful connectograms and of interactively extracting and analyzing focused sub-networks.

As previously mentioned, connectograms were introduced by Irimia et al. (2012a; 2012b) to provide intuitive and clear visualization of neuroconnectivity relationships, alternatively to large numerical matrices which do not allow prompt inference or hypothesis testing about either network properties or pathological damage. Although a variety of tools exploiting connectograms for studying connectomics already exist (Krzywinski et al., 2009; Whitfield-Gabrieli and Nieto-Castanon, 2012; Nieto-Castanon, 2020), so far, none of them allowed both interactive network exploration and the selection of sub-networks, while also providing a user-friendly interface. The interactivity allows for a faster execution of the tool and its use at ease, without the need of recompilation and re-uploading of the files. The user is also guided through the selection of different parameters providing a description for each feature and preventing from possible incidental choices. Therefore, SPIDER-NET could broaden the access to connectivity investigation to any interested user in the neuroscience field (e.g., neuropsychologists, physicians), other than computer scientists.

In the presented application examples, SPIDER-NET allowed to highlight different connectivity patterns between two patients with stroke injury. Interestingly, although both patients were characterized by a right hemisphere stroke lesion with prevalent subcortical expression, the disconnection patterns of the whole right hemisphere looked different between them. The different connectivity pattern of the right hemisphere was highlighted thanks to appropriate subgraph selections, interactively allowed by the tool. As expected, after considering the lesion patterns, right hemisphere connectivity differences qualitatively observed in the connectograms included connections with frontal lobe, insular cortex, and subcortical structures. Of note, although the connectograms presented in the figures were very dense, we chose to maintain the original density of the networks (19.65 and 22.32%, respectively, for Case 1 and 2) before any selection. This was carried out to avoid the possible introduction of thresholding biases, thus reducing the capability to capture the main differences between the connectivity patterns of the

two patients, especially considering the low original values of density. Although dense connectograms could result in poor readability, SPIDER-NET solves this issue allowing both interactive exploration of the network and stringent selection as successively performed by analyzing a sub-network of interest.

Indeed, in addition to a first explorative visual investigation of the right hemisphere, SPIDER-NET was used to perform a more focused analysis on a sub-network of interest, interactively testing an *a priori* hypothesis. The comparison of HCs, Case 1, and Case 2 sub-network connectograms generated by SPIDER-NET confirmed the *a priori* hypothesis and provided additional information about the disconnectivity pattern of Case 1 and Case 2, which may be used to improve the understanding of clinical manifestations and to drive personalized treatment. Indeed, both patients presented facio-brachio-crural hemisyndrome, with main brachial expression and severe functional limitation of movements of the left upper limb, especially of the hand. However, this limitation was more severe in Case 1 than in Case 2. This was mirrored by different residual connectivity patterns showed by SPIDER-NET connectograms. The following additional aspects were highlighted thanks to the circular diagrams. First, the pattern of disconnection involved both the right hemisphere, where the stroke lesions were present, and the contralateral one. Second, for both Case 1 and Case 2, the impairment of the cortical areas of interest determined a decrease in both short-range (within lobe) and long-range (between lobes) connections within the hemisphere ipsilateral to the stroke lesion. Third, in both patients with stroke injury the pattern of interhemispheric connectivity was also compromised, probably because subcortical nuclei, which are integration hubs of extrapyramidal systems, were extensively affected by the lesions. Therefore, producing connectograms on focused sub-networks with SPIDER-NET allowed to overcome the difficulty of visualizing the large number of edges that would be present in the whole-brain connectograms.

At a visual inspection, DTI-based and CSD-based sub-network connectograms presented comparable connectivity patterns, highlighting that valuable information is provided by both the processing techniques. Furthermore, CSD processing pipeline yielded to reconstruct denser connectograms, as expected. Indeed, CSD ability to better deal with the problem of the crossing fibers when compared to DTI is well-established (Dell'Acqua et al., 2007). This is in line with differences between DTI and CSD that were observed in terms of interhemispheric connections, that were particularly evident for Case 1.

In this study Case 1 and Case 2 were compared with a HC template obtained with the same method of Blasi et al. (2020). Although subjects included in the group allowed a good age-match with Case 1 and 2, SC is dependent on age. Therefore, defining an even more homogeneous HC template group is warranted for future studies using SPIDER-NET.

The connectivity patterns of pathological cases with focal lesions were here chosen for test purposes of a novel tool.

Nonetheless, generating SPIDER-NET connectograms could be a good general strategy to test the robustness of the processing pipeline, including the connectivity metrics, further

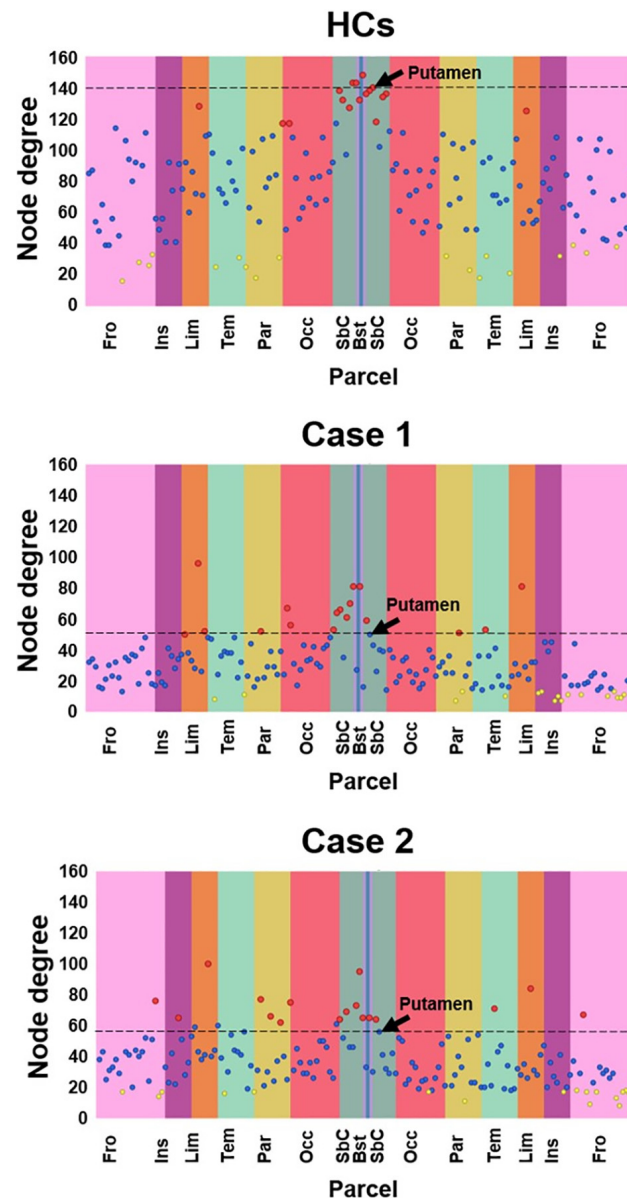


FIGURE 4 | Local node degree computed for HCs, Case 1 and Case 2. The X-axis represent the 165 brain parcels, even if only lobe labels are reported (e.g., Fro). The left hemisphere is represented in the left half of the graph, while the right hemisphere is represented in the right half. Vertical colored stripes represent different lobes (e.g., the frontal lobe is represented in pink). The same lobe in left and right hemispheres is shown with the same color. Each dot in the graph represents the local node degree of a brain parcel. The 17 parcels (10% of 165) exhibiting the highest local node degree are represented as red dots. The 17 parcels (10% of 165) with the lowest local node degree are represented as yellow dots. All the other parcels are represented as blue dots. The SPIDER-NET interactive interface allows to visualize information about each dot, navigating on them. The right Putamen is highlighted by an arrow.

conditioning (e.g., thresholding or binarization), and global or local graph indices. As one of the main limitation of connectomics, so far, is the lack of standardized procedures for network construction and edge weighting (Campbell and Pike, 2014; Maier-Hein et al., 2017), SPIDER-NET may be applied as a flexible and easy tool for calibrating connectomics analyses. Specifically, it could allow to quickly identify expected patterns of disconnection and to easily highlight major errors if present. This quality check may offer a benchmark before addressing

less trivial connectivity alterations, as the ones induced by diffused neurodegeneration, which might be another application field of SPIDER-NET.

The last step of graph analysis usually involves the computation of a set of different indices describing network topology and architecture, and dedicated software packages are generally employed. Beyond connectograms generation, SPIDER-NET allows to derive quantitative connectivity metrics, representing global and local (i.e., node level) network properties

TABLE 1 | Global graph-based topological properties of the HC template and the two stroke cases. Reported Delta values were computed as (HC-Case)*100/HC.

| Graph-based indexes | HCs | Case 1 | HCs-Case1 Delta (%) | Case 2 | HCs-Case2 Delta (%) |
|--------------------------------|--------|--------|---------------------|--------|---------------------|
| Average degree | 77.915 | 30.582 | 60.7% | 36.606 | 53.0% |
| Average strength (W) | 3.048 | 1.233 | 59.5% | 1.396 | 54.2% |
| Clustering coefficient | 0.76 | 0.614 | 19.2% | 0.604 | 20.5% |
| Clustering coefficient (W) | 0.019 | 0.017 | 10.5% | 0.015 | 21.1% |
| Characteristic path length | 1.542 | 2.092 | -35.7% | 1.927 | -25.0% |
| Characteristic path length (W) | 18.302 | 35.211 | -92.4% | 28.544 | -56.0% |
| Global efficiency | 0.735 | 0.549 | 25.3% | 0.587 | 20.1% |
| Global efficiency (W) | 0.068 | 0.039 | 42.6% | 0.045 | 33.8% |
| Small-worldness | 1.591 | 2.881 | -81.1% | 2.531 | -59.1% |
| Modularity | 0.193 | 0.38 | -96.9% | 0.351 | -81.9% |
| Coreness statistic | 0.321 | 0.402 | -25.2% | 0.367 | -14.3% |

(Rubinov and Sporns, 2010). For instance, characteristic path length is a global index mirroring communication efficiency within the network, while clustering coefficient is a global measure of network segregation. Among the several graph-based indices, the node degree is a basic local property of a network node, representing the number of connections with other nodes. Characterizing node degree distribution is an important component to identify putative hubs, namely nodes with high node degree, which significantly impact on the network topology. The example of application presented in this study highlighted the impact of graph-based metrics in connectivity analysis. Both local and global metrics derived from the whole-brain networks of Case 1 and Case 2 differed from HC one, as expected (Crofts et al., 2011; Cheng et al., 2019; Li et al., 2021). This quantitative result mirrored the differences qualitatively observed with connectograms. Node degree graphs produced by SPIDER-NET provided an intuitive tool to interactively explore network local properties. A first general visual comparison of Case 1, Case 2, and HCs node degree distribution highlighted that the patients with stroke injury were characterized by lower node degree across the whole brain (Sotelo et al., 2020). Therefore, although the two stroke lesions were limited to a portion of the right hemisphere, an alteration of the whole connectivity pattern was induced (Crofts et al., 2011; Cheng et al., 2019). In addition, it is noteworthy that SPIDER-NET graph-analysis confirmed that subcortical gray matter regions (e.g., the putamen) presented high node degrees in HCs, while these brain areas had lower node degrees in Case 1 and Case 2. This result reflected the prevalent subcortical expression of the two stroke lesions. Also, global graph metrics of segregation and integration were derived with SPIDER-NET, emphasizing the differences between cases with stroke injury and HCs (e.g., a drop in density of 28.9 and 25.2%, respectively, vs. HCs). Reduced connectivity in Case 1 and Case 2 compared to HCs was also numerically paralleled by large differences in all the parameters describing network topology and architecture. Specifically, patients with stroke injury were characterized by lower network integration, segregation, and efficiency. It is remarkable a greater difference in the characteristic path length (-92.4%, -56%) rather than in the clustering coefficient (10.5%, 21.1%) between the HCs and the two cases. The strong effect of the stroke lesion seems to

lead to a much more reduced integration than segregation in the contralesional hemisphere as shown in Crofts et al. (2011). Furthermore, Case 1 presented larger differences with HCs than Case 2, mirroring the greater clinical severity of the former. Therefore, SPIDER-NET automatically and easily provided useful metrics to quantitatively describe the impairment of the stroke patients included in this study.

Currently, the major limitation of SPIDER-NET is that it allows the analysis of one connectivity matrix at a time. The upload of more than one matrix to extract sub-matrices based on the user selection will be implemented in future SPIDER-NET versions. Furthermore, pre-conditioning operations are currently limited to density thresholding, as this is the most widespread thresholding method in connectivity studies (Wang et al., 2009; van Wijk et al., 2010; Beare et al., 2017). However, at present, other customized approaches can be used prior to the employment of SPIDER-NET by directly uploading already processed matrices. SPIDER-NET offers a flexible sub-network extraction method which relies on *a priori* hypothesis testing by manual selection of parcels/group-parcels and attributes. However, different approaches exist to automatically identify sub-graphs of interest (Hopcroft and Tarjan, 1973; Zalesky et al., 2010), especially in cases in which gross brain abnormalities may not be present. An interesting perspective may be to include automatic and data-driven algorithms for sub-network extraction and comparison with hypothesis-driven selection. In future works, investigating neurological diseases other than stroke and assessing changes associated with treatments (e.g., drugs or rehabilitation) is warranted to test SPIDER-NET sensitivity in detecting brain connectivity changes. Another interesting application might be the investigation of FC with SPIDER-NET and the integration of structural and functional information thanks to the flexibility in extracting sub-networks. In addition, SPIDER-NET application to brain connectivity matrices derived with other modalities (e.g., EEG, MEG, NIRS) could be a further future development. Although SPIDER-NET was presented and tested in this study for MRI datasets, its broad flexibility would actually allow applications even in several other diverse contexts, including all -omics disciplines. For instance, in the framework of rehabiliomics (Wagner and Sowa, 2014), which integrates evaluation of transdisciplinary biomarkers, SPIDER-NET may help in the definition of patient-tailored rehabilitative treatments.

CONCLUSION

In this work, we proposed a new freely available software package called SPIDER-NET⁷ and we tested it for deriving qualitative and quantitative valuable information of brain connectivity. First, the tool provided a facilitated, interactive, and real-time visualization of connectograms, based on flexible investigation of brain sub-networks. In addition, the automatic computation of topological properties of the networks completed the assessment with quantitative metrics. In conclusion, SPIDER-NET proved to be an accessible and useful tool for human brain connectome investigation in both physiological and pathological conditions.

DATA AVAILABILITY STATEMENT

The data presented in this study are available on request from the corresponding author. The MRI data are not publicly available due to privacy concerns.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by IRCCS Fondazione Don Carlo Gnocchi Ethics

⁷<https://caditer.dongnocchi.it>

REFERENCES

- Andersson, J. L., Skare, S., and Ashburner, J. (2003). How to Correct Susceptibility Distortions in Spin-Echo Echo-Planar Images: application to Diffusion Tensor Imaging. *NeuroImage* 20, 870–888. doi: 10.1016/S1053-8119(03)00336-7
- Andersson, J. L. R., and Sotiropoulos, S. N. (2016). An Integrated Approach to Correction for Off-Resonance Effects and Subject Movement in Diffusion MR Imaging. *NeuroImage* 125, 1063–1078. doi: 10.1016/j.neuroimage.2015.10.019
- Barttfeld, P., Wicker, B., Cukier, S., Navarta, S., Lew, S., Leiguarda, R., et al. (2012). State-Dependent Changes of Connectivity Patterns and Functional Brain Network Topology in Autism Spectrum Disorder. *Neuropsychologia* 50, 3653–3662. doi: 10.1016/j.neuropsychologia.2012.09.047
- Baggio, H. C., Sala-Llonch, R., Segura, B., Marti, M. J., Valldeoriola, F., Compta, Y., et al. (2014). Functional brain networks and cognitive deficits in Parkinson's disease. *Hum. brain mapp.* 35, 4620–4634. doi: 10.1002/hbm.22499
- Bassett, D. S., and Sporns, O. (2017). Network neuroscience. *Nat. Neurosci.* 20, 353–364.
- Beare, R., Adamson, C., Bellgrove, M. A., Vilgis, V., Vance, A., Seal, M. L., et al. (2017). Altered Structural Connectivity in ADHD: a Network Based Analysis. *Brain Imag. Behav.* 11, 846–858. doi: 10.1007/s11682-016-9559-9
- Behrens, T. E. J., Woolrich, M. W., Jenkinson, M., Johansen-Berg, H., Nunes, R. G., Clare, S., et al. (2003). Characterization and Propagation of Uncertainty in Diffusion-Weighted MR Imaging. *Magn. Reson. Med.* 50, 1077–1088. doi: 10.1002/mrm.10609
- Berron, D., van Westen, D., Ossenkoppele, R., Strandberg, O., and Hansson, O. (2020). Medial temporal lobe connectivity and its associations with cognition in early Alzheimer's disease. *Brain* 143, 1233–1248.
- Blasi, V., Alice, P., Cabinio, M., Di Tella, S., Laganà, M. M., Giangiacomo, A., et al. (2020). Early Life Adversities and Borderline Intellectual Functioning

Committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

DC, GB, and FB contributed to the conception and design of the study. DC developed SPIDER-NET, performed the associated analyses, and wrote the first draft of the manuscript. AP, LP, MC, and ML performed the image processing to derive connectivity matrices. DC, AP, and LP drafted the first version of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This work was supported by Lombardy Region (Announcement POR-FESR 2014–2020—Azione I.1.B.1.3), within the project named Smart&TouchID and by the Italian Ministry of Health (“Ricerca Corrente”).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2022.818385/full#supplementary-material>

- Negatively Impact Limbic System Connectivity in Childhood: a Connectomics-Based Study. *Front. Psychiatry* 11:497116. doi: 10.3389/fpsy.2020.497116
- Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198. doi: 10.1038/nrn2575
- Campbell, J. S. W., and Pike, G. B. (2014). Potential and Limitations of Diffusion MRI Tractography for the Study of Language. *Brain Lang.* 131, 65–73. doi: 10.1016/j.bandl.2013.06.007
- Cheng, B., Schlemm, E., Schulz, R., Boenstrup, M., Messé, A., Hilgetag, C., et al. (2019). Altered Topology of Large-Scale Structural Brain Networks in Chronic. *Stroke Brain Commun.* 1:fcz020. doi: 10.1093/braincomms/fcz020
- Crofts, J. J., Higham, D. J., Bosnell, R., Jbabdi, S., Matthews, P. M., Behrens, T. E. J., et al. (2011). Network Analysis Detects Changes in the Contralateral Hemisphere Following Stroke. *NeuroImage* 54, 161–169. doi: 10.1016/j.neuroimage.2010.08.032
- Daianu, M., Dennis, E. L., Jahanshad, N., Nir, T. M., Toga, A. W., Jack, C. R., et al. (2013). *Alzheimer's Disease Disrupts Rich Club Organization in Brain Connectivity Networks.* *IEEE 10th International Symposium on Biomedical Imaging.* San Francisco: IEEE, 266–269.
- Dell'Acqua, F., Rizzo, G., Scifo, P., Clarke, R. A., Scotti, G., and Fazio, F. (2007). A model-based deconvolution approach to solve fiber crossing in diffusion-weighted MR imaging. *IEEE Trans. Biomed. Eng.* 54, 462–472. doi: 10.1109/TBME.2006.888830
- Dell'Acqua, F., Scifo, P., Rizzo, G., Catani, M., Simmons, A., Scotti, G., et al. (2010). A modified damped Richardson-Lucy algorithm to reduce isotropic background effects in spherical deconvolution. *Neuroimage* 49, 1446–1458. doi: 10.1016/j.neuroimage.2009.09.033
- Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31, 968–980. doi: 10.1016/j.neuroimage.2006.01.021

- Destrieux, C., Fischl, B., Dale, A., and Hagren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage* 53, 1–15. doi: 10.1016/j.neuroimage.2010.06.010
- Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., et al. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 33, 341–355. doi: 10.1016/s0896-6273(02)00569-x
- Gao, S., Mishne, G., and Scheinost, D. (2020). “Poincaré Embedding Reveals Edge-Based Functional Networks of the Brain” In *Medical Image Computing and Computer Assisted Intervention – MICCAI*. L. M. Anne, A. Purang, S. Danaïl, M. Diana, A. Z., Maria, Z. S., Kevin, et al., New York: Springer Cham. 448–457.
- Göttlich, M., Münte, T. F., Heldmann, M., Kasten, M., Hagenah, J., and Krämer, U. M. (2013). Altered Resting State Brain Networks in Parkinson’s Disease. *PLoS One* 8:e77336. doi: 10.1371/journal.pone.0077336
- Hopcroft, J., and Tarjan, R. (1973). Algorithm 447: efficient Algorithms for Graph Manipulation. *Commun. ACM* 16, 372–378.
- Irimia, A., Chambers, M. C., Torgerson, C. M., and Van Horn, J. D. (2012b). Circular representation of human cortical networks for subject and population-level connectomic visualization. *NeuroImage* 60, 1340–1351. doi: 10.1016/j.neuroimage.2012.01.107
- Irimia, A., Chambers, M., Torgerson, C., Filippou, M., Hovda, D., Alger, J., et al. (2012a). Patient-Tailored Connectomics Visualization for the Assessment of White Matter Atrophy in Traumatic Brain Injury. *Front. Neurol.* 3:10. doi: 10.3389/fneur.2012.00010
- Isernia, S., Pirastru, A., Massaro, D., Rovaris, M., Marchetti, A., and Baglio, F. (2021). Resting-State Functional Brain Connectivity for Human Mentalizing: biobehavioral Mechanisms of Theory of Mind in Multiple Sclerosis. *Soc. Cogn. Affect. Neurosci.* nsab120. [Epub online ahead of print]. doi: 10.1093/scan/nsab120.
- Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage* 17, 825–841. doi: 10.1016/s1053-8119(02)91132-8
- Ji, G. J., Yu, Y., Miao, H. H., Wang, Z. J., Tang, Y. L., and Liao, W. (2017). Decreased Network Efficiency in Benign Epilepsy with Centrottemporal Spikes. *Radiology* 283, 186–194. doi: 10.1148/radiol.2016160422
- Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: an Information Aesthetic for Comparative Genomics. *Genom. Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109
- Li, Y., Yu, Z., Wu, P., and Chen, J. (2021). The Disrupted Topological Properties of Structural Networks Showed Recovery in Ischemic Stroke Patients: a Longitudinal Design Study. *BMC Neurosci.* 22:47. doi: 10.1186/s12868-021-00652-1
- Maier-Hein, K. H., Neher, P. F., Houde, J. C., Côté, M. A., Garyfallidis, E., Zhong, J., et al. (2017). The Challenge of Mapping the Human Connectome Based on Diffusion Tractography. *Nat. Commun.* 8:1349.
- Nieto-Castanon, A. (2020). *Handbook of Functional Connectivity Magnetic Resonance Imaging Methods in CONN*. Stockholm: Hilbert Press.
- Pelizzari, L., Laganà, M. M., Di Tella, S., Rossetto, F., Bergsland, N., Nemni, R., et al. (2019). Combined Assessment of Diffusion Parameters and Cerebral Blood Flow Within Basal Ganglia in Early Parkinson’s Disease. *Front. Aging Neurosci.* 11:134. doi: 10.3389/fnagi.2019.00134
- Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *NeuroImage* 52, 1059–1069. doi: 10.1016/j.neuroimage.2009.10.003
- Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., et al. (2009). Correspondence of the brain’s functional architecture during activation and rest. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13040–13045. doi: 10.1073/pnas.0905267106
- Sotelo, M. R., Kalinosky, B. T., Goodfriend, K., Hyngstrom, A. S., and Schmit, B. D. (2020). Indirect Structural Connectivity Identifies Changes in Brain Networks After Stroke. *Brain Connect.* 10, 399–410. doi: 10.1089/brain.2019.0725
- Sporns, O., Tononi, G., and Kötter, R. (2005). The human connectome: a structural description of the human brain. *PLoS Comput. Biol.* 1:e42. doi: 10.1371/journal.pcbi.0010042
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289.
- van Wijk, B. C., Stam, C. J., and Daffertshofer, A. (2010). Comparing Brain Networks of Different Size and Connectivity Density Using Graph Theory. *PLoS One* 5:e13701. doi: 10.1371/journal.pone.0013701
- Wagner, A. K., and Sowa, G. (2014). Rehabilomics Research: a Model for Translational Rehabilitation and Comparative Effectiveness Rehabilitation Research. *Am. J. Phys. Med. Rehabil.* 93, 913–916. doi: 10.1097/PHM.0000000000000114
- Wang, J., Wang, L., Zang, Y., Yang, H., Tang, H., Gong, Q., et al. (2009). Parcellation-Dependent Small-World Brain Functional Networks: a Resting-State fMRI Study. *Hum. Brain Mapp.* 30, 1511–1523. doi: 10.1002/hbm.20623
- Whitfield-Gabrieli, S., and Nieto-Castanon, A. (2012). Conn: a Functional Connectivity Toolbox for Correlated and Anticorrelated Brain Networks. *Brain Connect.* 2, 125–141. doi: 10.1089/brain.2012.0073
- Xia, M., Wang, J., and He, Y. (2013). BrainNet Viewer: a Network Visualization Tool for Human Brain Connectomics. *PLoS One* 8:e68910. doi: 10.1371/journal.pone.0068910
- Yeo, B. T., Sepulcre, J. F. M., Sabuncu, M. R., Lashkari, D., Hollinshead, M., and Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* 106, 1125–1165.
- Zalesky, A., Fornito, A., and Bullmore, E. T. (2010). Network-based statistic: identifying differences in brain networks. *Neuroimage* 53, 1197–1207. doi: 10.1016/j.neuroimage.2010.06.041

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Coluzzi, Pirastru, Pelizzari, Cabinio, Laganà, Baselli and Baglio. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Diffuse and Localized Functional Dysconnectivity in Schizophrenia: a Bootstrapped Top-Down Approach

Davide Coluzzi^C

Dipartimento di Elettronica, Informazione e Bioingegneria
Politecnico di Milano
davide.coluzzi@polimi.it

Giuseppe Baselli

Dipartimento di Elettronica, Informazione e Bioingegneria
Politecnico di Milano
giuseppe.baselli@polimi.it

Abstract. Schizophrenia is a brain disorder leading to detached mind's normally integrated processes. Hence, the exploration of the symptoms in relation to functional connectivity (FC) had great relevance in the field. Connectivity can be investigated on different levels, going from global features to single edges between pairs of regions, revealing diffuse and localized dysconnection patterns. In this context, schizophrenia is characterized by a different global integration with reduced connectivity in specific areas of the brain, part of the Default Mode Network (DMN). However, the assessment of FC presents various sources of uncertainty. This study proposes a multi-level approach for more robust group-comparison.

FC data between 74 AAL brain areas of 15 healthy controls (HC) and 12 subjects with chronic schizophrenia (SZ) were used. Multi-level analyses were carried out by the previously published SPIDER-NET tool. Graph topological indexes were evaluated to assess global abnormalities. Robustness was augmented by bootstrapped (BOOT) data and the stability was evaluated by removing one (RST1) or two subjects (RST2). The DMN subgraph was extracted and specifically evaluated. Changes relevant to the overall local indexes were also analyzed. Finally, the connection weights were explored to enhance common strongest activations/deactivations.

At a global level, expected trends of the indexes were found and the significance of modularity ($p = 0.043$) was not confirmed by BOOT ($p = 0.133$). The robustness assessment tests (both RST1

^CCorresponding author

Address for correspondence: Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano

and RST2) highlighted more stable results for BOOT compared to the direct data testing. Conversely, significant results were found in the analysis at lower levels. The DMN highlighted reduced connectivity and strength as well as increased deactivation in the SZ group. At local level, 13 areas were found to be significantly different ($p < 0.05$) in the groups, highlighting a greater divergence in the frontal lobe. These results were confirmed analyzing the single negative edges, suggesting inverted connectivity between prefronto-temporal areas.

In conclusion, multi-level analysis supported by BOOT is highly recommended when analyzing FC, especially when diffuse and localized dysconnections must be investigated in limited samples.

1. Introduction

The brain is the most tangled network known in nature. Indeed, it comprises billions of neurons which form trillions of synapses between each other. Conversely, at a macroscale extent, brain regions can be thought as connected through fiber bundles, mirroring the anatomical structure, or functional activations of the areas, in resting-state or during a task. All this collective activity makes behaviour, thought, memory and consciousness possible. In this context, clinical disorders are characterized by alterations of the connections' paths. For this reason, it is paramount to explore these abnormalities to enforce our understanding of the brain in health and disease.

Considering the methods used to investigate brain networks, Magnetic Resonance Imaging (MRI) is the dominant technique for macroscale analysis mainly because of its safety, spatial resolution and availability. On one hand, Diffusion Tensor Imaging (DTI) or other methods allow to visualize and examine the organization of structural connectivity by white matter (WM) tracts. On the other hand, functional MRI (fMRI) inspects the dynamics of activity in each gray matter (GM) area. The functional activations of the GM areas are based on the Blood Oxygen Level Dependent (BOLD) response.

Through these methodologies it is possible to obtain connectivity matrices, which can be investigated in their organization and function. Thus, a matrix (graph) can be defined as a collection of nodes (brain regions) and edges (anatomical or functional connections) between pairs of nodes [1]. The nodes are obtained by mapping the parcels of the network according to well-known brain atlases defined from the structural or functional point of view [2, 3]. Accordingly, the edges represent the links detected between the regions describing the weights of the connections. Indeed, the strength of each connection is weighted according to the addressed connectivity measure such as the number of fibers in the structural case, or the Pearson's correlation in the functional case.

The comprehension of these networks is a complex field of research named connectomics, which can potentially address the brain at all its scales, levels and features [4]. Focusing on the macroscale of brain GM areas, it is first possible to analyze and compare different networks through the investigation of their graph-based global (top level) topological features such as node degree, node strength, clustering coefficient, efficiency, path length or modularity [1, 5].

Second, the brain is also known to be divided into interacting specialized sub-networks, devoted to specific domains of behavior and cognition [6]. Extracting sub-networks is an increasingly common practice in explorative studies to easily interpret the brain connectivity data in physiological and pathological conditions. [6, 7, 8, 9, 10]. In this regard, we presented in a previous study a novel software tool allowing the interactive and flexible analysis of anatomical/functional networks and sub-networks called SPIDER-NET [11].

Third, the above-mentioned graph indexes can also be computed at local level providing information regarding the configuration of the single nodes. Valuable information is extracted about hubs of the networks and regions associated to specific disruptions, such as lesions [11, 12].

These three levels of analysis (network, sub-network and node-level) together with the analysis of the

single edges resulted helpful in the study of many pathological conditions. Indeed, global changes of relevant properties, different sub-network configurations, disrupted node connectivity and missing relevant connections were already found in stroke, Alzheimer's disease and other pathologies [13, 14]. However, the integration and the interpretation of the results extracted at all three levels is still difficult. The parallel multi-level analysis can be particularly helpful in case of pathologies, such as the schizophrenia which the localized/diffuse origin is unknown yet [15].

Schizophrenia may be characterized as a disorder of brain connectivity since it leads to detached mind's processes. The main symptoms of the pathology such as reduced cognitive and emotional deficits acknowledge this breakdown [15]. More specifically, the clinical manifestations of the disorder can be hallucinations, disorganized thinking and agitation, social isolation, emotional flattening, anhedonia and apathy. Often, these manifestations may be preceded by a series of prodromal symptoms, such as inability to perform one's job or neglect of personal hygiene.

Many studies [16, 17] explored the connections between these symptoms and functional associations. Specifically, connectivity deficits were found on all abovementioned levels [15]. First, a global connectivity reduction was discovered in different studies [17, 18, 19]. Second, at interconnected sub-network level, fronto-temporal and occipito-temporal dysconnections were found through the use of the network-based statistics [7]. Third, amongst all brain regions, the prefrontal cortex resulted to be one of the most affected nodes [16, 18, 19]. More generally, Liu and colleagues found a significant alteration of the pattern of small-world topological properties in many brain regions of the prefrontal, parietal and temporal lobes. Four, Skudlarski and colleagues analyzed single connectivity edges, finding differences in the connectivity pattern originating from posterior cingulate cortex within the Default Mode Network (DMN) suggesting to have the most affected connections by the functional reorganization of the schizophrenia [20].

DMN is reasonably important in the context of schizophrenia since it is deeply involved in social behavior, control of the emotional state of the individual [21, 22], which characterize altered integrated processes by the pathology. From different studies it resulted that connections within this circuit were different than healthy controls, being related to emotion control and memory [21, 23]. However, the observations for the DMN remained controversial, since the average connectivity of the patients showed higher or lower values than healthy controls according to the dynamic mental state and the individual connections were subject to reorganization [20].

Indeed, although the analysis of changes in functional connectivity is a powerful tool to analyze brain organization, due to the limitations of the fMRI techniques, the statistical investigation of brain connectivity datasets is often subject to uncertainty. Indeed, stationarity is often assumed in the interpretation of the results coming from this technique. However, considering the known dynamic and condition-dependent nature of brain activity, it is obvious that the functional connectivity metrics such as the Pearson correlation coefficient will change over time [24, 25, 26]. This variability allows to define the paradigm of dynamic functional connectivity analysis, which cannot be ignored, since it also varies within the same subject and even between time windows within the same session. In this context, the evaluation of functional static connectivity is subjected to uncertainty. In addition, functional connectivity acquisitions are characterized by low signal-to-noise ratio (SNR) and non-neural noise related to cardiac and respiratory processes and hardware instability. Finally, other definition issues such as the window length and other confounds remain controversial [27, 28, 29, 30].

Among all statistical methodologies, bootstrapping was already used in some imaging problems, such as in the work conducted by Lazar and Alexander [31], since it allows to robustly estimate the statistical features of a population from a limited number of measurement samples, without any assumptions about the distribution assumed by the initial data. This approach also represents an alternative to traditional hypothesis testing, since it does not require to have a test statistic satisfying certain assumptions largely dependent on the experimental design and to know the properties of the data. As a result, the main advantage of the method is that the uncertainty variability of the estimator

can be quantified, characterizing the dispersion and other errors in the null hypothesis [32, 33]. Considering the number of bootstrap samples used, the resulting statistics represent a random sample with replacements from the initial distribution characterize by a smaller size. It is worth remarking that obtaining thousands of bootstrap observations from the initial data is not the same as collecting new data. Indeed, the approach is based on an ensemble of simulated data (surrogates) and the usefulness of bootstrapping is related to the quantification of statistical quantities such as the standard error, a possible bias and a confidence interval of particular sample of data.

In the context of investigating graph-structured data, the quantification of the uncertainty intrinsic to the data is essential for their scientific usefulness. In the recent probabilistic study conducted by Green and Shalizi [34] the bootstrap was applied on simulated random graphs. It was seen that the resampling of the synthetic data was able to approximate the distributions of motif densities, such as, the number of times fixed subgraph appear in the random network. In this work, the bootstrap was also applied to quantify the uncertainty of the network metrics. Another example is given by the study conducted by Gel and colleagues, where bootstrap was applied for the quantification of uncertainty in graph degree distributions of collaboration networks, where for example articles are represented as nodes and cross-references as edges [33]. Indeed, when applying the bootstrap over a distribution of data, the expected result is a better definition of the variability centered with respect to the initial mean of the same distribution at the increasing of the number of resampled data [35]. However, a limitation of this approach is related to computational costs since the bootstrapping methods are based on multiple resampling of the original dataset that can be computationally expensive, especially for large datasets.

However, the bootstrapping approach was only partially applied in brain connectivity studies. For example, in the study conducted by Wei and colleagues this approach was used to perform connectivity matrix feature selection in a regression task cognitive traits prediction [36]. Spearman correlation analysis was indeed performed between connectivity and cognitive measures in each resample subset to extract a feature vector. In a previous study [37], the results from mean functional connectivity of Parkinson's disease patients were qualitatively analyzed through bootstrap.

To the best of our knowledge, there is significant potential for effectively and more quantitatively applying bootstrap techniques on connectivity data, although fundamental issues still need to be addressed. Indeed, connectivity measurements are well-known to be affected by different sources of noise, which can have a strong impact especially with limited population size. Also, the static evaluation of functional connectivity is limited with respect to dynamic functional connectivity [24, 28]. In this sense, it is crucial to evaluate the robustness and the stability of potential biomarkers of brain connectivity to improve their usability and understanding. Moreover, there is the need to evaluate abnormalities on multiple levels, especially in conditions characterized by both localized and diffuse degeneration, such as schizophrenia.

Hence, the main aim of this study was the robust assessment of connectivity indexes on different levels through the bootstrap procedure to ensure a reliable detection of abnormalities in schizophrenia. To investigate both diffuse and localized alterations, a top-down analysis of: i) global connectivity deficits, ii) sub-network disruptions, extracting the DMN, iii) dysconnectivity in individual regions and its inter-subject variability, and iv) abnormalities in positive/negative connections resulting in activation/deactivation circuits or communities is proposed. The methodology was developed in the context of the SPIDER-NET tool [11] to allow automatic multi-level investigations in group studies. Afterwards, the stability of results is assessed by comparing bootstrapping to direct testing via a leave-n-subject-out approach to evaluate the impact of sources of uncertainty in functional connectivity. The whole pipeline of multi-level analysis was validated on a dataset acquired from healthy and schizophrenic subjects highlighting several abnormalities at the examined levels.

2. Methods

2.1. Data Acquisition and Study Population

Data used in this study were collected in the context of the work conducted by Zalesky and colleagues [7]. Regarding the written informed consent and ethical approval, reference is made to the original work. The anonymized connectivity dataset is publicly available at <https://nitrc.org/projects/nbs/>. The dataset is composed by 15 healthy controls - HC (mean age 33.3 years, $\sigma = 9.2$ years, 14 males) and 12 subjects with chronic schizophrenia - SZ (mean age 32.8 years, $\sigma = 9.2$ years, 10 males). The patients were diagnosed according to standard operational criteria in the Diagnostic and Statistical Manual of Mental Disorders IV (American Psychiatric Association, 2000). The matching of the two populations was done according to age, pre-onset IQ and years of education. SZ subjects did not receive medication on the day of acquisition to reduce acute drug effects on the data. T2*-weighted echo-planar images depicting blood oxygenation level dependent contrast were acquired through a 1.5 Tesla scanner (GE Signa, General Electric, Milwaukee, WI). As regards the matrix construction: the nodes were defined according to a subset of areas of the AAL atlas, whereas the edges were obtained computing the correlation between times series which were previously preprocessed. More specifically, scale 3 of the wavelet transform ($0.03 < f < 0.06$ Hz) was considered and the filtered time series were corrected for fluctuations of signals through linear regression against reference time courses extracted from seed regions. Some AAL nodes were excluded since it was not possible to correctly estimate the node-averaged time series, due to poor coverage in some subjects. The resulting connectivity matrices for each subject have dimensions 74×74 . More details on ethical approval, imaging parameters and processing are reported in the cited work [7].

Generally, after composing the matrix, a threshold is defined to emphasize its topological features by removing spurious associations. Among all methods, absolute and density thresholding are the most common in connectivity studies [5, 38]. In this case, we decided to not threshold the matrices since it introduces a confound on properties of the graphs in the context of a group study. On one hand, different levels of sparsity can be obtained by applying the same threshold to all matrices. In this case, it is thus not possible to rule out systematic density variations as the main reason of group abnormalities of graph-based metrics. On the other hand, matrices can be matched according to sparsity using density thresholding, hence selecting a distinct threshold value. In this case, whether the weight distribution in SZ is reduced, the application of density thresholding can strongly affect weighted indexes results [19, 39]. Furthermore, no thresholding allowed to work on all raw connectivity measures, avoiding the loss of information of the original data and assessing the results of the bootstrap methodology in relation to all sources of uncertainty. On one hand, all weakest and thus possible spurious connections were maintained to account them in the robust evaluation of the indexes through bootstrapping. On the other hand, the density differences are only given by the deletion of the negative connections, also allowing the investigation of binary topological indexes. Hence, two versions of same matrix were obtained to analyze separately positive/negative edges and, thus, activations and deactivations.

2.2. Top-Down Bootstrapping Approach for Group Comparison

The top-down bootstrapping approach to enable the HC vs SZ group comparison consists of different steps of analysis, done on different levels. The analysis pipeline was built starting from the top level, related to the global characteristics of the matrices coming from the two populations, until the bottom level, related to the single connections. The procedure is summarized in Figure 1.

First, the global topological indexes were obtained to evaluate the abnormalities in the network topology and global deficits. In order to enable a robust investigation, the values of these properties were

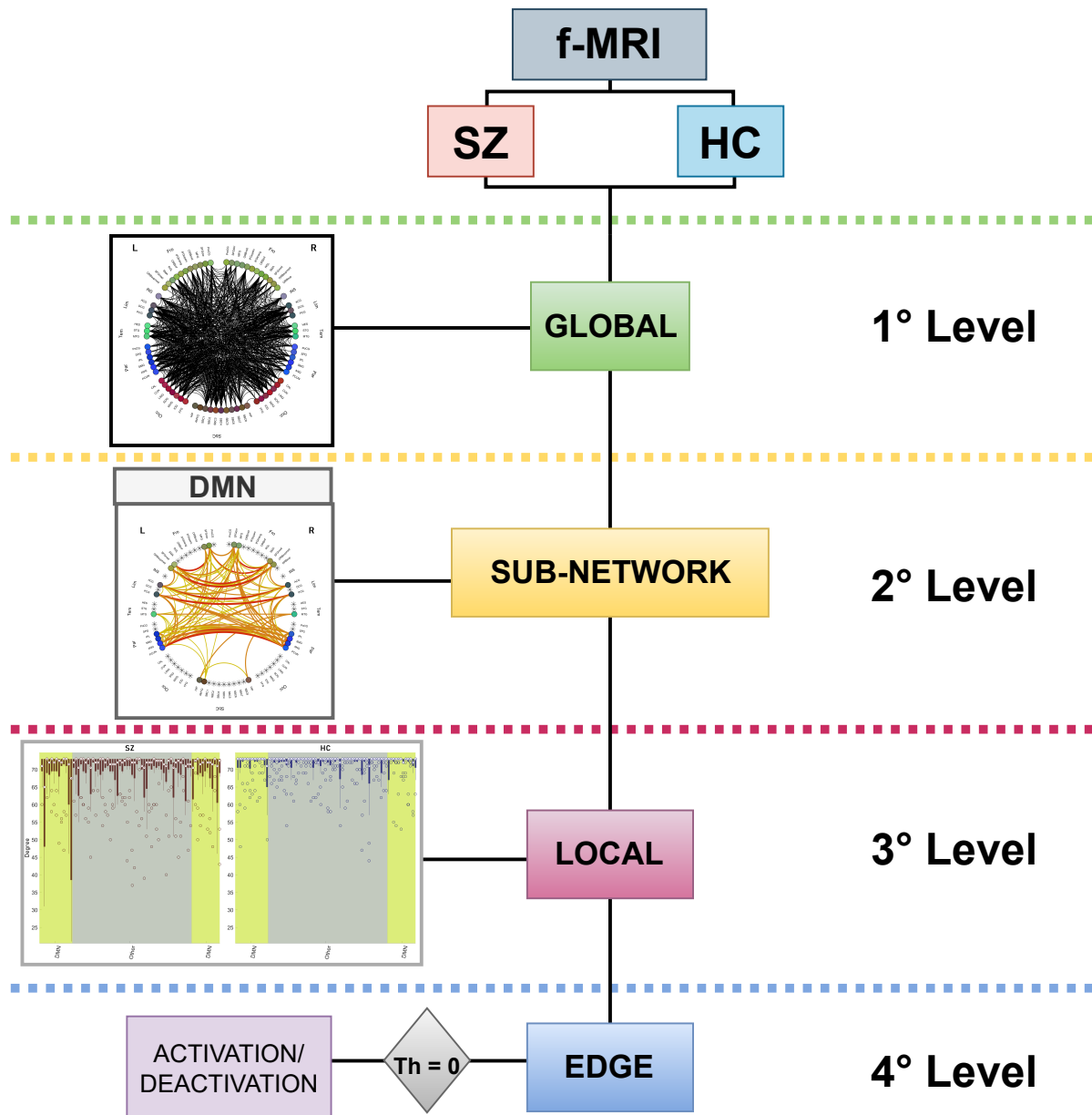


Figure 1: Pipeline of the top-down approach proposed. The functional connectivity matrices of the two populations were analyzed from the global level, extracting the DMN as sub-network of interest and investigating the distributions of the local indexes to the edge level, assessing correlations and anticorrelations.

bootstrapped, as reported in Section 2.2.1. Second, bootstrapping was also applied on the extracted sub-network of interest (DMN) and local indexes to visualize significant changes between groups and different subjects of the same groups (see Section 2.2.2). Finally, the negative connections of the group mean matrixes were analyzed in Section 2.2.3. In this section, the analysis of the communities was performed to enhance common deactivations. Afterwards, a qualitative investigation of strongest and most frequent negative connections reflecting anticorrelations was performed.

The group comparison functions required by the present study were implemented as upgrade of our previously developed SPIDER-NET tool, freely available to researchers [11] The current upgrades will be included in a new release.

2.2.1. Global Connectivity

The global topological indexes from the graph theory which were considered are reported in the Figure 2.

The indexes are divided according to integration/segregation indexes (e.g. path length, clustering

| Analysis | Graph Properties | Mathematical Expression |
|-----------------------------------|---|---|
| Segregation (local and global) | Degree Strength | $k = \text{avg}(\sum_j a_{ij}) ; k^w = \text{avg}(\sum_j w_{ij})$ <small>where i represents the nodes and j the neighbors and weights, respectively</small> |
| | Average clustering coefficient (binary and weighted) | $C = \frac{1}{n} \sum_{i \in N} \frac{2t_i}{k_i(k_i-1)} ; C^w = \frac{1}{n} \sum_{i \in N} \frac{2t_i^w}{k_i(k_i-1)}$ where t_i represent the triangles of the node i |
| Integration (local and global) | Characteristic path length (binary and weighted) | $L = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} d_{ij}}{n-1} ; L^w = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} d_{ij}^w}{n-1}$ <small>where d_{ij} represent the distances between nodes, and n is the maximum number of possible connections</small> |
| | Efficiency (binary and weighted) | $E = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} (d_{ij})^{-1}}{n-1} ; E^w = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} (d_{ij}^w)^{-1}}{n-1}$ |
| Mesoscale | Modularity | $Q = \frac{1}{2m} \sum_{c_n} \sum_{i,j \in c_n} \left[a_{ij} - \frac{k_i k_j}{2m} \right]$ where <small>m is the total number of edges and c_n are the communities</small> |
| | Coreness statistics | $N = \frac{1}{v_s} \left(\sum_{i,j \in V_c} (w_{ij} - \bar{w}) - \sum_{i,j \in V_p} (w_{ij} - \bar{w}) \right)$ where <small>V_c, V_p are sets of nodes in the core, periphery, \bar{w} is the average edge weight and v_s is a normalization constant</small> |
| Global structure | Density | $D = \frac{X}{(n^2 - n)/2}$ <small>where X is the number of connections</small> |
| | Small-worldness | $S = \frac{C/C_{rand}}{L/L_{rand}}$ <small>where $rand$ refers to measures of random network</small> |

Figure 2: Table of all graph properties considered, divided according to the level of analysis. The segregation and integration properties can be globally assessed through the average of all local values. The mathematical expressions reported indicate global case formulation.

coefficient), mesoscale analysis (e.g. modularity, rich-club coefficient) and related to whole network structure organization (e.g. small-worldness).

The degree/strength represent the sum of edges/weights connected to the node. The clustering coefficient quantifies local edge density by counting the triangles average. A triangle occurs if a neighbor of the node is also a neighbor of another neighbor of the node. In the weighted case, the number

of triangles is replaced with the geometric mean of its weights. The characteristic path length is a measure of integration which expresses the average shortest path between node pairs. The global efficiency is the average inverse of the characteristic path length, and it represents how efficiently the information travels through the network. All the segregation and integration metrics can be local and global, simply by averaging the values across all nodes. Mesoscale analysis reveals how much the network present a particular structure. Specifically, the modularity computed through Newman's approach [40] quantifies to what extent the intra-/inter-community link densities are anomalous in comparison to chance. Large values typically reveal significant community structures. Maximized coreness statistic is a measure of how much the network follows the core/periphery paradigm, which is a partition of the network into two groups, where the number and weight of the edges is maximized in the core and minimized within the periphery.

The density is a measure of sparsity of the network. It is the ratio between the number of actual connections and the maximum number of possible connections. A small-world network is a structure in which most nodes are not neighbors of one another. Conversely, the neighbors of any given node are likely to be neighbors of each other, resulting in an easy access by most nodes to every other node with a small number of steps. Small-world networks associate short path length and high clustering coefficient.

Statistical Tests and Robustness Assessment Mann–Whitney tests (MW) were performed to assess statistically significant differences in the global indexes between the two groups. Because of the dimensions of the sample (#HC = 15; #SZ = 12), the intrinsic uncertainty given by the non-stationarity of the data [24] and the possible presence of spurious connections due to the limitations in the fMRI processing [28], which can significantly affect the results, the reliability of significant group differences was also tested. First, bootstrap hypothesis testing (BOOT) was performed on all global indexes and compared to MW testing. BOOT allows for a better estimation of the null distribution of network measures, providing confidence intervals to evaluate the uncertainty of the statistics and more robust abnormalities detection [41]. The procedure to test bootstrap hypothesis was the following: The procedure to test bootstrap hypothesis was the following:

1. Calculation of the test statistic $\hat{\theta} = |\tilde{x} - \tilde{y}|$, given x_1, \dots, x_N a random sample from distribution F with median \tilde{x} and y_1, \dots, y_N another independent random sample from distribution G with median \tilde{y} .
2. Bootstrapping: extraction of B sets of random samples x^* (size N) and y^* (size M) with replacement from x and y , respectively.
3. Calculation of the test statistic $\hat{\theta}_b^* = |\tilde{x}_b^* - \tilde{y}_b^*|$ for each resample.
4. These B resampled test statistics are then made into a null distribution by $\hat{\theta}'_b = \hat{\theta}_b^* - \hat{\theta}$.
5. Estimate of the p-value as

$$p = \frac{\sum_{b=1}^B [C(\hat{\theta}'_b \geq \hat{\theta}) + 1]}{B + 1}$$

where $C\{condition\} = 1$ when the condition is true and 0 otherwise.

F and G represent the distributions of a global index in HC and SZ, thus N and M representing the number of HC and SZ, 15 and 12 respectively. The resampling from these two sets is carried out 5000 (B - number of resamples) times for all bootstrap hypothesis tests and the test statistic is the difference in medians [42].

Second, the robustness of the measures was assessed through a leave-n-subject-out approach, with n equal to 1 and 2. The variations of the statistics computed through both methods were analyzed

changing the population samples according to the two procedures.

The first one, hereafter referred as RST1, evaluates the variation of the p-values obtained removing all subjects one at a time from the populations values to perform MW tests and to create new resamples for BOOT. The second one, hereafter referred as RST2, evaluates the variation of the p-values obtained randomly removing pairs of subjects from the populations values to perform both tests. In this case, 350 random extractions of pairs of subjects were performed. The total possible pairs which can be extracted are $27 * 26 = 702$, thus we decided to run the test for about 50% of the possibilities. RST2 is summarized in Figure 3.

Then, mean and standard deviation across all tests were evaluated for both procedures.

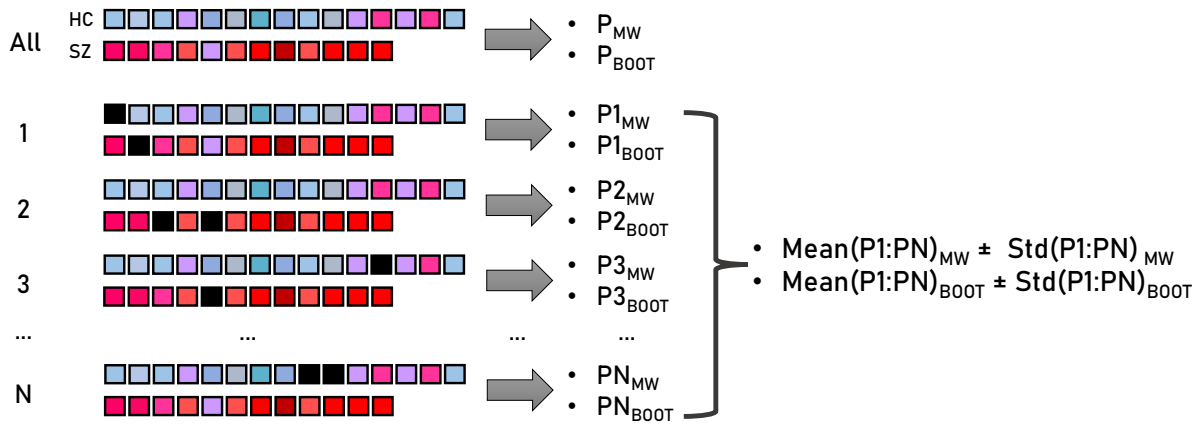


Figure 3: Schema representing the robustness statistical test randomly removing two subjects (black squares) at each iteration (RST2). In this case, N is equal to 350.

2.2.2. Local Topological Properties and Sub-Network Extraction

Afterwards, local topological properties related to the specific nodes of the network across all subjects in the two groups were analyzed. These indexes are the local counterpart of the global ones, as reported in the Section 2.2.1 and Figure 2. In this regard, we obtained the distributions of these values assessing differences in the local structure and in the variability across the two groups. Specifically, we focused on degree and strength. The former helps to represent the influence of negative connections, being the original matrixes complete and divided by using a threshold in 0. The latter instead identifies node or set of nodes whose weighted values differ in the two populations. In this case, we performed MW and BOOT, as reported in the previous section 2.2.1, on the local index values. The resampling was performed with 5000 samples for all BOOT tests. Furthermore, these differences were investigated across the nodes of the different lobes and the nodes within/outside the DMN to obtain the most deviating local values between the two groups. The distributions per areas and the inter-subject variability, which can be related to the different intra-group characteristics, were highlighted.

In order to perform these analyses, we divided the nodes according to the lobes and we extracted the DMN from all matrices of both groups. Specifically, we defined the DMN according to previous studies highlighting the regions characterizing the network [43, 44, 45, 46, 47]. Generally, the brain regions found to be activated within the DMN comprise the medial frontal cortex, the medial temporal lobe, the posterior cingulate cortex, the ventral precuneus and the inferior parietal cortical

regions. Apart from these typical core areas, lateral temporal cortex, hippocampal formation and amygdala are also often reported as parts of the network [44, 46] Thus, from the typical Broadmann areas comprising the DMN (BA: 9, 10m, 10r, 10p, 23/31, 24, 29/30, 32a, 32c, 39, 49) and these additional regions, a list of 13 AAL parcels from the left hemisphere and 12 AAL parcels from the right hemisphere was obtained. The parcel which is present only in the left hemisphere is the amygdala since the right counterpart was not present in the data, due to the previously described issue of poor coverage in some subjects (see Section 2.1) [7]. The sub-networks were then composed of these nodes and the edges existing between pairs of DMN areas (“Extract” mode in SPIDER-NET), maintaining the intra-lobe connections.

2.2.3. Connection-Level Investigation

The examination of the single connections was performed computing the mean matrix of the groups which were compared to visualize differences in weights and deactivation. Thus, the average networks composed by all negative contributions were obtained. Then, community detection analysis through the Newman’s method [40] implemented in the Brain Connectivity Toolbox [48] and embedded into SPIDER-NET [11] was performed. The resulting negative sub-networks were obtained and analyzed. In addition, connectograms were drawn considering the entire network with no thresholding apart the separation of positive and negative edges. The rationale was to analyze graphs formed by all brain regions and edges, or all the DMN areas and edges. Thresholding was applied only for graphical purposes in some connectogram figures where density obscured the main connections, as highlighted in the captions.

3. Results

First, the global topological properties were assessed. Table 1 summarizes the results obtained from the comparison between HC and SZ of the overall indexes using nonparametric test and bootstrap hypothesis testing. The results from the robustness tests are also summarized in Table 1.

It is possible to notice that HC presented increased values for all segregation indexes and efficiency. Conversely, concerning the characteristic path length, mesoscale analysis, and small-worldness SZ resulted to have higher values. Differences are then enhanced in the comparison between MW and BOOT. In particular, higher p-values in BOOT are found for the weighted characteristic path length and modularity indexes. Specifically, the latter highlighted a statistically significant ($p < 0.05$) difference between HC and SZ modularity that was not confirmed by using BOOT. The robustness assessment highlighted more stable results for BOOT. Indeed, it is worth noting that the mean of the p-values, considering all indexes and removing both one and two subjects, remains closer with BOOT than the MW test in almost all indexes considered. In few cases (binary and weighted clustering coefficient in RST1) the difference between the mean of the robust assessment tests and initial tests is slightly minor in MW or comparable. Furthermore, the variability was assessed, resulting in lower standard deviation employing BOOT with respect to MW in almost all indexes. In few cases (e.g. weighted characteristic path length and modularity in RST2) this reduction is not enhanced with comparable values. Second, DMN was extracted from all subjects of the two populations. In Figure 4 the connectograms of this sub-network obtained averaging all positive connections across the whole populations are shown. It is worth noting that from a visual inspection a greater number of weak connections is highlighted in the SZ group, represented by the yellow edges in Figure 4.

Table 1: Results of the statistical tests performed on global indexes.

| Graph Index | Whole Comparison | | | | RST1 (#Sub=1) | | RST2 (#Sub=2) | |
|-------------------------|--------------------|--------------------|-----------------|-------------------|----------------------------|------------------------------|----------------------------|------------------------------|
| | HC data (mean±std) | SZ data (mean±std) | P _{MW} | P _{BOOT} | P _{MW} (mean±std) | P _{BOOT} (mean±std) | P _{MW} (mean±std) | P _{BOOT} (mean±std) |
| Degree (Density) | 71.746±2.178 | 69.207±5.149 | 0.130 | 0.086 | 0.146±0.05 | 0.089±0.044 | 0.162±0.079 | 0.103±0.063 |
| Strength | 37.053±7.362 | 31.680±10.246 | 0.124 | 0.079 | 0.143±0.047 | 0.096±0.029 | 0.155±0.071 | 0.108±0.045 |
| Bin CC | 0.986±0.025 | 0.960±0.057 | 0.150 | 0.124 | 0.150±0.052 | 0.130±0.038 | 0.184±0.089 | 0.138±0.057 |
| Wei CC | 0.489±0.105 | 0.415±0.144 | 0.150 | 0.105 | 0.155±0.052 | 0.119±0.027 | 0.183±0.082 | 0.128±0.050 |
| Bin CPL | 1.017±0.030 | 1.052±0.071 | 0.130 | 0.084 | 0.153±0.052 | 0.089±0.044 | 0.162±0.079 | 0.102±0.063 |
| Wei CPL | 2.216±0.610 | 2.618±0.892 | 0.124 | 0.129 | 0.151±0.061 | 0.150±0.054 | 0.155±0.072 | 0.159±0.072 |
| Bin Eff | 0.991±0.015 | 0.974±0.035 | 0.130 | 0.086 | 0.150±0.052 | 0.089±0.044 | 0.162±0.079 | 0.102±0.063 |
| Wei Eff | 0.525±0.085 | 0.466±0.113 | 0.113 | 0.078 | 0.147±0.051 | 0.099±0.023 | 0.143±0.065 | 0.111±0.041 |
| Modularity | 0.004±0.009 | 0.016±0.026 | 0.043 | 0.133 | 0.053±0.018 | 0.128±0.019 | 0.059±0.030 | 0.144±0.036 |
| Coreness | 0.018±0.022 | 0.037±0.036 | 0.164 | 0.122 | 0.141±0.060 | 0.122±0.039 | 0.199±0.091 | 0.138±0.061 |
| Small-Worldness | 1.003±0.005 | 1.015±0.019 | 0.178 | 0.085 | 0.146±0.062 | 0.090±0.013 | 0.214±0.098 | 0.096±0.030 |

The global version of the degree represents the mean of all edges associated to all nodes, whereas the density is the ratio between the number of present edges and the number of possible edges in the network. For this reason, the result coincides, and it is reported only once. HC: Healthy Controls; SZ: Schizophrenic Patients; RST: Robust Statistical Test; MW: Mann Whitney Test; BOOT: Bootstrap Hypothesis Test; #Sub: number of subjects removed; CC: Clustering Coefficient; CPL: Characteristic Path Length; Eff: Efficiency.

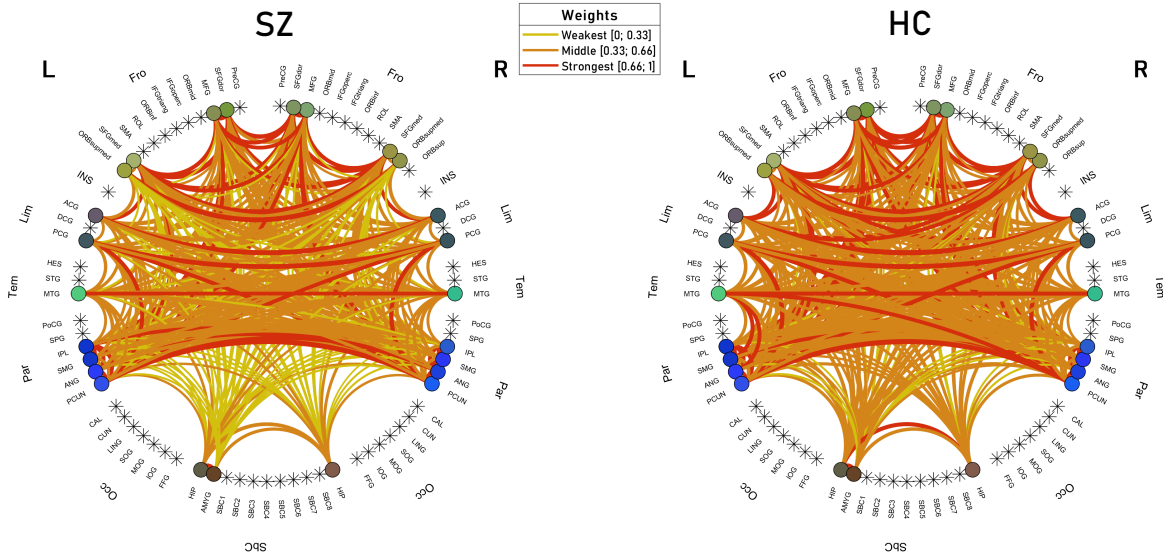


Figure 4: Connectograms composed by only nodes of the DMN and positive connections between pairs of these nodes.

Second, we analyzed the local values of the indexes in the DMN and across the lobes, showing the distributions of the values of the degree across the same group in Figure 5. Differences between DMN regions and the other areas in SZ and HC, because of the different influence of the negative connections, were highlighted. Specifically, the most variable indexes resulted to be located within the DMN for the SZ group. On the other hand, they are more distributed across all the brain regions in HC group, even resulting more variable in the areas not included in the DMN. More precisely, one of the most variable values of the degree in both groups was the one assumed by the amygdala, located at the end of the first yellow box in Figure 5, representing the left hemisphere nodes of the DMN.

Afterwards, the strength was analyzed, and the resulting distributions are shown in Figure 6. First, it was noticed that the distributions of the SZ have a slightly higher variability in the majority of the nodes with respect to HC. More specifically, the most variable local indexes among their populations were extracted from frontal, parietal lobes and subcortical structures. On the other hand, general slightly higher values of strength are present in the HC group. Furthermore, the distributions of the strengths are shown in Figure 6 highlighting statistically significant differences ($p < 0.05$) found through both MW and BOOT (in red). Most of the differences were found in regions of the frontal lobe of both hemispheres. Significance of BOOT only (in pale brown) resulted in the left orbital part of inferior frontal gyrus, postcentral gyrus, Heschl's gyrus, and right Rolandic operculum nodes. Significance of MW only (in yellow) was found in the in right calcarine fissure and the surrounding cortex node. The complete list of the significant different node strengths is reported in the caption of Figure 6.

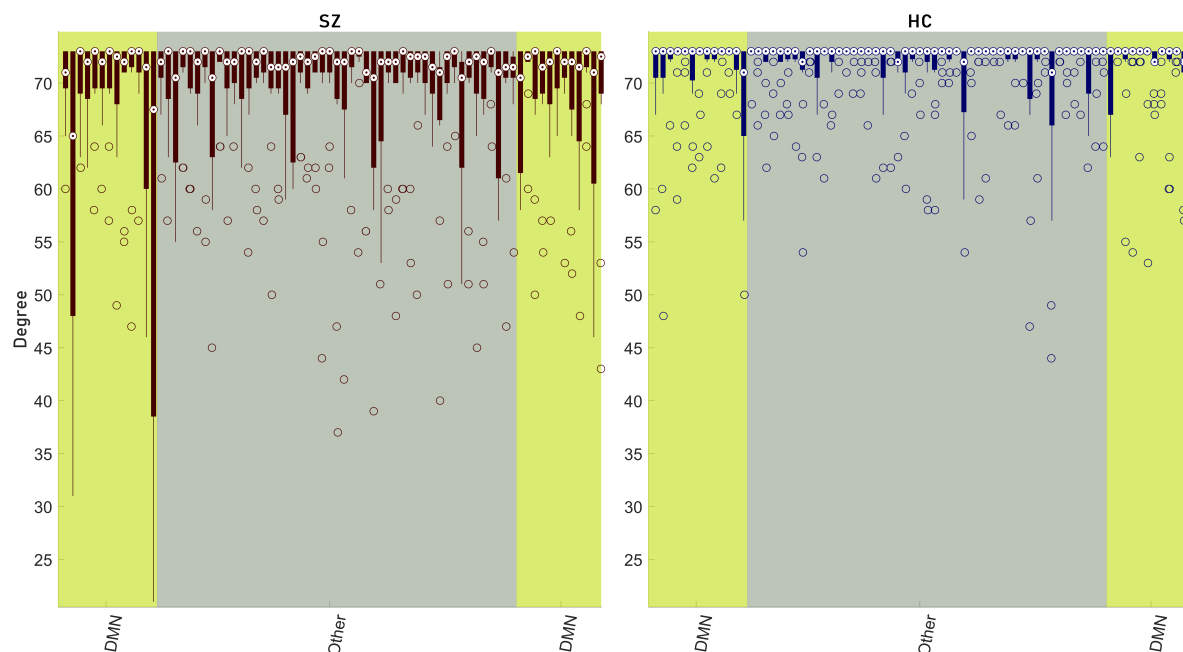


Figure 5: Distributions of the local degree values in the two populations divided according to nodes of DMN in left and right hemisphere (yellow) or not (gray).

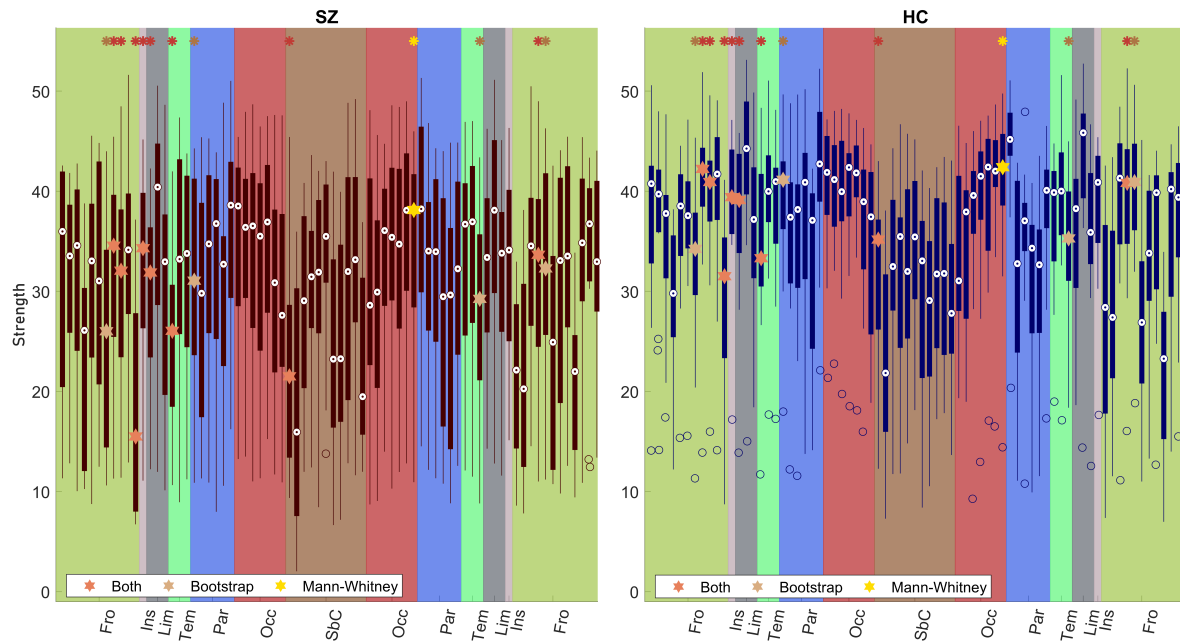


Figure 6: Distributions of the local strength values in the two populations divided according to the lobes. The order and colors of the lobes is the same as the connectograms (Figures 4 and 7) from the left to the right hemisphere. Strength of nodes which are statistically significant different ($p < 0.05$) between the two populations are shown in different colors according to the different methods used: pale brown for bootstrap hypothesis testing, yellow for Mann-Whitney and red for statistically significantly different nodes found by both. More specifically, both methods identify differences in the left Rolandic operculum, supplementary motor area, medial orbital superior frontal gyrus, insula, anterior cingulate and paracingulate gyri, hippocampus, Heschl's gyrus and right supplementary motor area, bootstrap hypothesis testing in left orbital part of the inferior frontal gyrus, postcentral gyrus, right Rolandic operculum and right Heschl's gyrus and Mann-Whitney test in the right calcarine fissure and surrounding cortex.

Third, we analyzed the negative connections from the group average matrixes. Results from community detection analysis are shown in Figure 7. Three communities, a) b) and c) on top of the Figure, composed of 32, 22, 20 brain regions respectively were identified for the negative SZ group. Four communities, a) b) c) and d) on bottom, composed of 20, 15, 20 and 19 regions respectively were identified for the negative HC group. A difference between the formed communities is the inclusion of most frontal, limbic and temporal regions in the community a) of SZ group, which is not observable in any HC community. Furthermore, the configuration of parietal regions from both hemispheres, mostly included in the community a) in HC, is not noticeable in SZ, where these regions are divided according to hemisphere b) and c) of SZ, also appearing to be clustered with the contralateral occipital lobe.

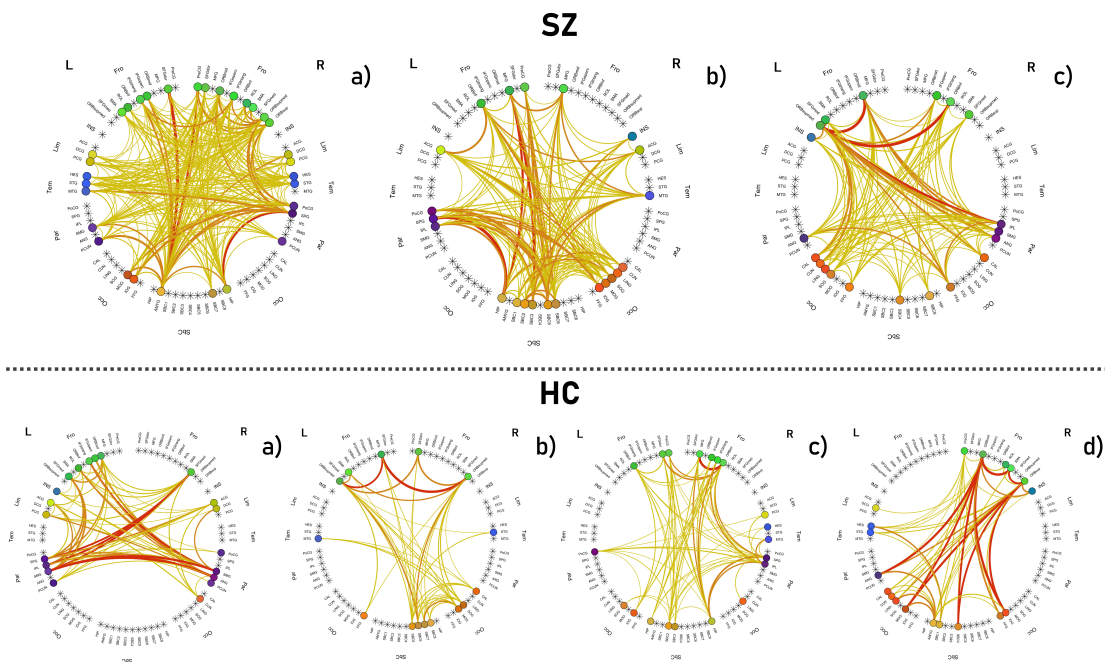


Figure 7: Connectograms of the communities detected in the negative average group networks. On top, the three communities found in the SZ group (modularity=0.079), whereas, on bottom, the four communities of HC group (modularity=0.165). The edges are color-coded in red, orange, and yellow to represent the strongest, middle, and weakest negative connections within the community, respectively. No thresholding was applied.

Then we visualize the strongest and frequent negative edges in the whole brain and in the DMN, as shown in Figure 8. Specifically, the connectograms formed by the average of all matrixes composed by the negative edges in the whole brain (7a and 7b) and in the DMN (7c and 7d) are shown applying a threshold to keep 5% strongest negative edges. First, SZ average network resulted to have more negative connections in the whole brain with respect to HC. Another great difference noticed both in the number of preserved negative edges and in their weights was in the inter-hemispheric connectivity of the frontal lobe, as well as the parietal. In addition, it is possible to notice in SZ a greater number of negative connections in the occipital lobe with respect to HC. The configurations given by the only strongest connections in DMN highlighted similarities and differences. Among

all, connections between amygdala and some frontal regions were in common, whereas strong edges between the parietal lobes of the two hemispheres were found only in HC. In SZ parietal regions are instead connected to amygdala and superior frontal gyrus. Finally, HC average matrix revealed an important edge between the latter, but in the right hemisphere, and the middle frontal gyrus, that was not found in SZ.

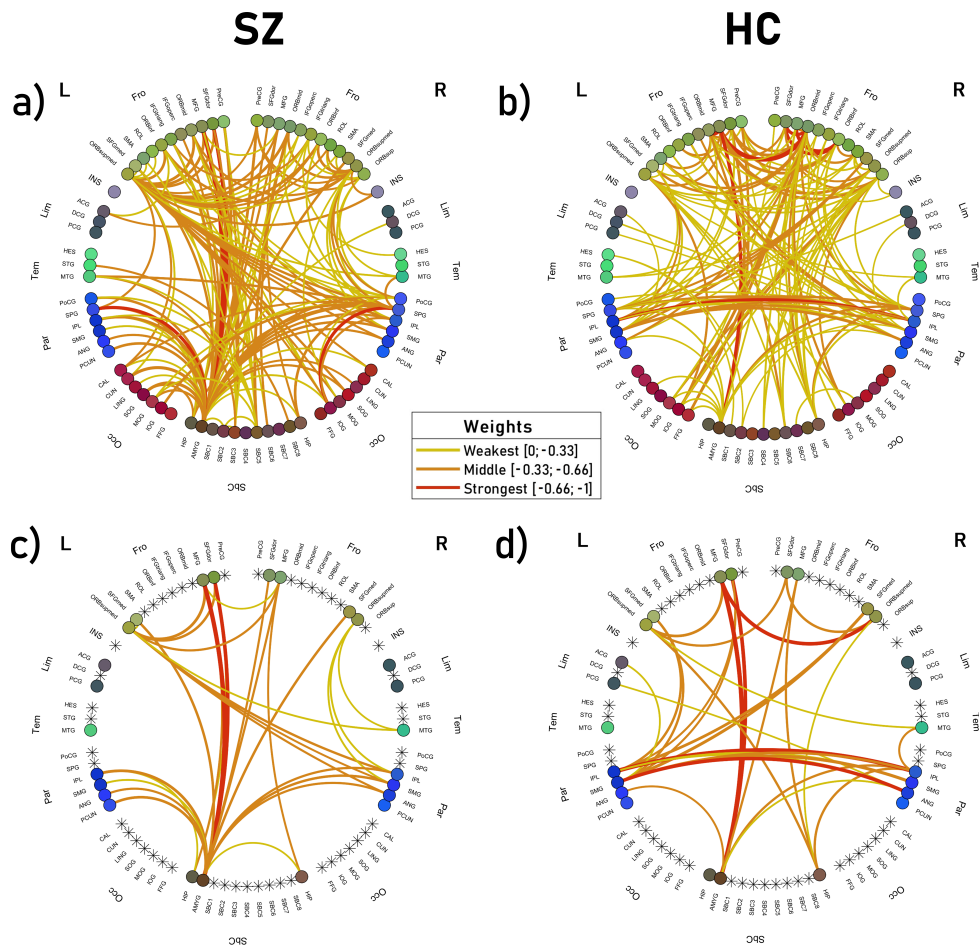


Figure 8: Connectograms of negative average group networks. a) and b) are the results on the whole-brain, c) and d) on only the DMN of SZ and HC groups respectively. 5% density thresholding is applied on the shown connectograms for graphical clarity only.

4. Discussion

In this work, we proposed a multi-level approach for the case-control analysis of connectivity data structured on different levels of investigation. In particular, the assessment of the connectivity from global indexes to single edges (brain activations/deactivations) was carried out on a population characterized by chronic schizophrenia, thought to cause both diffused and focalized dysconnectivity patterns. The multi-level analysis resulted to potentially favor more robust results in contexts where

the focus is not well-known and statistical tests can be easily biased by uncertainty.

First, global-level results were analyzed. As previously mentioned, this level was examined according to three main investigations. All experiments comprised the evaluation of 11 graph-based indexes, either binary or weighted, which are amongst the most widespread in connectivity studies [1, 49, 50, 51]. The first analysis involved the comparison between HC and SZ groups, compared through two statistical tests (MW vs BOOT). In Table 1 we showed the values of the indexes along with the p-values obtained from both tests. The increased segregation indexes and efficiency in HC with respect to SZ is in line to what reported from previous studies [18, 16], as well as reduced path length, mesoscale analysis and small-worldness. More specifically, MW highlighted one statistically significant difference between the two population in the modularity ($p_{MW} = 0.043$). The modularity resulted to be significantly higher in SZ providing evidence that the deletion of negative connections, in this case analyzed separately, tend to create families in SZ that in HC are less visible. However, it is worth noting that similar studies presented different results for this particular index and also uncertainty for clustering coefficient and characteristic path length, indicating either increased or reduced values in SZ [15, 16, 17, 52, 53, 54, 55, 56], probably dependent on the thresholding method. For these reasons, as previously mentioned, the reliability of these results was assessed through BOOT because of the reduced size of the dataset, the presence of the uncertainty due to the non-stationarity of the functional connectivity and the possible presence of spurious connections. The results of the bootstrap hypothesis testing highlights differences with respect to the standard nonparametric approach. On one hand, the modularity did not present significant difference ($p_{BOOT} = 0.133$). On the other hand, a trend towards lower p-values can be noticed for all other indexes, except the weighted path length ($p_{MW} = 0.124$; $p_{BOOT} = 0.129$) and the above-mentioned modularity. One of the most replicated results is the significant difference found in the efficiency between HC and SZ. According to different studies, functional connectivity fluctuations appear to be coordinated throughout the brain so as to realize global variations in network efficiency, which could represent a balance between optimizing information processing and minimizing metabolic expenditure [15, 57, 58]. Furthermore, examining studies with a greater number of subjects [59], or of dynamic connectivity [57], the results obtained through BOOT seem to be closer than MW. At this stage of understanding, we may speculate that BOOT seems to provide an alternative method to test more robustly the differences between functional connectivity of two groups, especially when composed by low number of data and subjects.

However, we acknowledge that there is considerable debate among researchers as to uncertainty influence and integration of dynamical functional connectivity data. For this reason, we extended the analysis comparing MW and BOOT through robustness assessment tests. In particular, the variability of the statistical estimate was evaluated at the removal of one or two subjects from the whole population, just as in leave-n-subject-out cross validation. In Table 1 the results of these experiments are summarized, finding clear support for the hypothesis of more stability and reliability of the BOOT results. In particular, the index mean values obtained by BOOT across the random extractions (27 for leave one out and 350 for leave two out) of subjects to be removed remain much more stable with respect to the values obtained with entire sample (e.g., degree - RST1: $p_{MW=0.130}$; $mean(p_{MW}) = 0.146$; $p_{BOOT} = 0.086$; $mean(p_{BOOT}) = 0.089$; RST2: $mean(p_{MW}) = 0.162$; $mean(p_{BOOT}) = 0.103$). Also, the standard deviation across all iterations resulted to be much lower by using BOOT than MW (e.g., small-worldness - RST1: $std(p_{MW}) = 0.062$; $std(p_{BOOT}) = 0.013$; RST2: $std(p_{MW}) = 0.098$; $std(p_{BOOT}) = 0.030$). The observations regarding mean and standard deviation can be generalized to all indexes considered, resulting in less variable results or, in few cases remaining similar to MW. More specifically, the cases of the resulting mean of binary and weighted clustering coefficient for RST1, or the standard deviation of the weighted characteristic path length and modularity for RST2, which anyway did not result to be less stable with BOOT. In general, the results highlighted a higher stability and robustness of the BOOT in RST1 and RST2,

providing support to its possible use when dealing with a small dataset which is also affected by uncertainty of the measures.

It is worth discussing that the results revealed by the robustness tests may answer the question of evaluating the results of the functional connectivity in a more reliable way. Although the needs for dynamic functional acquisition [24, 27, 60] and of more attractive methods for removing spurious connections are well-known [15, 61, 62, 63], having a reliable method to evaluate functional connectivity of small datasets, which can potentially approximate to results from greater datasets and avoid the loss of information, can help in the interpretation and understanding.

Second, focusing on the results of the second level, comprising sub-network (DMN) and local indexes, interesting findings were noticed. It is worth noting that a first visual difference between the mean matrix of the groups was visible. Indeed, in Figure 4, the connectograms highlight weaker connections in the DMN of SZ with respect to HC, in agreement to previous studies [21, 59]. However, no significant differences between the indexes of the sub-networks were found. On the other hand, the evaluation of the local indexes allowed to identify some notable patterns. From Figure 5, it is indeed possible to notice that the most varying local degrees are those referred to nodes of the DMN, especially in the SZ group. Analyzing, instead, the results in Figure 6 on the strength index, the differences are less visible although a slightly greater value is noticed across all nodes of the HC group, together with a confirmed greater variability distributed across the whole brain in the SZ group. On one hand, these findings report a greater presence of negative connections in the DMN of SZ, which can be seen as an inverted connectivity nature between prefronto-temporal areas, in agreement to diversity found in a previous study [64]. This is particularly indicative considering the involvement of the DMN in SZ [21] and that many symptoms of the pathology mirrors a failure in the integration between the generated behaviour and concurrent perceptive phenomena [65]. On the other hand, this variability can also be influenced by the limitations of the data. It is indeed well-known as one of the states identified by performing dynamical connectivity experiments can be traced to the DMN [60].

Furthermore, we analyzed the nodes which had the greatest variability of the degree index and that mostly differed in the strength value. Amygdala, medial orbital part of the superior and inferior frontal gyrus and part of the cingulate cortex resulted to be the nodes which were most affected by the presence of negative connections in the SZ group. Interestingly, amygdala, which is also a well-known hub of the brain [66], part of the superior frontal gyrus and cingulate cortex are all part of the DMN, whereas the inferior frontal gyrus not. However, the inferior frontal gyrus was widely studied because of its divergent characteristics of anatomical/functional connectivity in SZ and its relation to semantic processing [67]. The enhanced variability may be indeed related to specific characteristics and deficiencies of the patients, which may be potentially investigated through correlation studies with demographic and severity variables, or, in this case, semantic function tests, as already widely done in literature [59, 49, 55]. Afterwards, we also computed MW and BOOT tests on the local values of the strength, highlighting the nodes mostly differing in the value of this index between the two populations. In general, it is possible to notice that the most different local indexes are those referred to frontal, parietal lobes and subcortical structures, in agreement to previous study [7, 46, 68]. In particular, significance of left orbital part of inferior frontal gyrus, postcentral gyrus, Heschl's gyrus and right Rolandic operculum nodes were found by BOOT and right calcarine fissure and surrounding cortex node by MW. The results on these areas appear to be in line with previous studies, where these nodes were found to be different, although the case of Heschl's and postcentral gyri were found in the specific case of SZ patients with auditory hallucinations [55, 69, 70].

The community detection analysis of the anticorrelations highlighted differences between the two groups. First, in SZ, most nodes from frontal, limbic and temporal lobes are clustered together in the first community. This deactivation organization may be relevant in the study of failed activity inhibition of schizophrenia characterized by selective disruption of an automatic inhibitory process,

and failure to limit the current contents of consciousness [71, 72]. In the parietal lobes configuration other differences were noticed, especially regarding the clustering with contralateral occipital lobe in SZ. In other studies [73, 74, 75], these lobes were investigated in relation to SZ highlighting important functions. For example, it was hypothesized that cognitive deficits and delusions may be related to malfunctions in the parietal lobe [75] or that the maintenance of visuospatial information is associated to a network of occipital cortex regions [73]. It would be of great interest to study correlations between symptoms and activity in these communities, as the contralateral negative cooperation may reinforce the hypothesis of an involvement of parieto-occipital sub-network in auditory hallucination, such as the positive correlation found in an EEG and MEG study [74], or the relations to other disorganization symptoms.

Third, the connectogram pattern of negative edges obtained by averaging all the negative contributions in the whole brain and DMN of the groups highlighted interesting indications. Primarily, SZ resulted to have more and strongest negative connections in the whole brain with respect to HC. Second, an increased deactivation in the occipital lobe appear to confirm the results previously obtained through network based statistics, which found dysconnections in the same areas [7]. Then, the number and the weight of the preserved negative correlations are generally greater in SZ in the frontal, which includes prefrontal as well, and parietal lobes, in agreement to what previously shown on sub-network and nodes. This finding appears to be in line with previous studies where different organization of the connectivity was found in prefrontal and frontal lobes [76, 23]. As said, the anti-correlations could indicate an abnormal inhibition of some regions or lobes activity. In the context of DMN, except the connections from amygdala to the frontal regions expected to be common to both groups being a well-known hub of the brain networks, as abovementioned [66], a different configuration of the edges was noticed. First, in the HC group, the parietal lobes of the two hemispheres are deactivated with strong interactions that are not present in the SZ group, in agreement to community detection analysis. In this case, parietal regions of DMN mostly communicate with amygdala and superior frontal gyrus. Furthermore, HC group is characterized by a very strong negative connection between middle frontal gyrus and the medial orbital part of the superior frontal gyrus that is not even present in the SZ group. These findings may be related to significant small-world topology abnormalities found in regions of the prefrontal and parietal lobes in a previous study [16]. Indeed, the difference of negative network involving the DMN regions of the parietal lobe is the most evident, both for the number of edges and their interactions with the DMN regions of the frontal lobe. From previous works [21, 77, 78], it is known as intrinsic task-positive and task-negative networks exist in HC during rest, comprising some areas of these lobes. It can be speculated that this organization may be responsible for external environmental and self-referential processing respectively and serves mental processes. In individuals with SZ or paranoid tendencies, abnormal negative connectivity within and between these two networks could cause them to be oversensitive to both internal thoughts and external stimuli. In this sense, abnormal presence of anticorrelations, together with a low integration between the two networks, may reflect abnormal inhibition of the other circuit and that extrospective and introspective thinking is switched back and forth in an overly excessive manner [21, 79]. In this work, differences were also found in the temporal lobe that, however, in deactivation analysis of our dataset remained comparable. Minor differences given by the presence of some negative weak connections in HC which are not present in SZ were found in the regions of the DMN of the limbic lobe, confirming the affection of the functional reorganization of the SZ originating from posterior cingulate cortex [20]. This resulted to be in line with previous findings on cingulate gyrus in SZ and its importance in functional integration [80].

However, it is worth noting that the small size of the HC and SZ groups available for this study is a limitation. A subsequent validation of the bootstrap method on graph-based indexes obtained from a wider functional connectivity dataset would be attractive. Although the greater stability to changes of the measures and a trend towards results of larger studies is promising, validation with

more subjects is needed. A related limitation concerns the lack of discrimination among different types and severity levels of schizophrenia [21]. Also, the leave-k-subject-out procedure to test the statistical analyses may be performed on dynamical connectivity data. Thus, the assessment of the influence of the different states which are assumed during a resting-state fMRI would reinforce the potential of the approach in underpinning usability and give better understanding of the functional activations and deactivations.

Second, one of the main aims of the study was the validation of the approach through bootstrapping of the indexes accounting all connections, which include uncertainty as well, but without any loss of information or arbitrary thresholding of spurious edges and very low correlations. The results obtained through BOOT were more stable in comparison to standard nonparametric tests, being robust to noise, but at the same time it is paramount to focus and apply new and different thresholding methods to assess their influence, as briefly mentioned. Examples are percolation or spanning tree approaches [15, 17, 61, 62, 63] appearing to be less arbitrary than typical density or absolute thresholding and which may potentially improve even more the estimates. The former, for example, was widely employed in a number of studies to find a potential optimal deletion of the spurious edges maintaining brain network structure and connectedness [81, 82, 10, 83]. In this perspective, results from bootstrap analysis may be investigated in relation to percolation analysis. Another possibility could be to apply bootstrapping of the groups at connection-level to obtain more robust templates for comparison, identify errors and quantify the uncertainty of the measure.

Third, the top-down approach was implemented in the SPIDER-NET tool which automatically allows the investigation of topological properties that are the most widespread in brain connectivity studies [11]. However, some of the indexes analyzed have intrinsic limitations, such as the weakness of Newman's modularity to detect small modules [84]. In addition, thresholding methods can strongly affect the results of community analyses and topological indexes. For this reason, further investigations and comparisons to other methods, such as random walk-based approaches for community detection, would be of great interest.

Finally, the robust multi-level approach helped to identify differences in the connectivity of SZ patients. More specifically, it did not result in increased or decreased connectivity with respect to HC but in a different configuration, highlighted by the bootstrapping. As said, the most replicated graph-feature is the efficiency, whereas other features appear to be more affected by the pre-processing and there is no strict agreement. Also, the great intra-group variability, particularly relevant for SZ, plays an important role that may be investigated in future considering the clinical condition and severity of the patients or other specific characteristics and deficits. From these considerations, the results confirmed that the connectivity in SZ is not systematically increased/decreased but generally different, as already reported in another study [20]. Nonetheless, it is clear and replicated that some areas are affected more in SZ [15, 54]. However, it remains unclear whether SZ abnormalities are the result of a localized dysconnection exerting widespread effects throughout the brain, or a whole-brain dysfunction that affects certain regions more. New investigations would be certainly needed to further investigate the causes and effects related to the onset of the condition.

5. Conclusion

In this work, we described a multi-level robust approach for the investigation of the brain connectivity. Starting from the measures of global connectivity, which can be robustly and more stably explored through bootstrapping, general abnormalities in case-control groups can be identified. In some cases where the connectivity is not evidently increased/reduced, or particularly different in some areas, going towards lower levels can be helpful. In this case, we analyzed separately the functional connectivity in the DMN, beyond those indexes related to specific nodes. Finally, the ac-

tivations and the deactivations were investigated. The whole procedure is developed in the SPIDER-NET software tool (<https://caditer.dongnocchi.it/spidernet/>) to automatically help in the different level investigations of various pathology to better understand their nature and effects.

References

- [1] Mikail Rubinov and Olaf Sporns. Weight-conserving characterization of complex functional brain networks. *NeuroImage*, 56(4):2068–2079, 2011.
- [2] Rahul S. Desikan, Florent Ségonne, Bruce Fischl, Brian T. Quinn, Bradford C. Dickerson, Deborah Blacker, Randy L. Buckner, Anders M. Dale, R. Paul Maguire, Bradley T. Hyman, Marilyn S. Albert, and Ronald J. Killiany. An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. *NeuroImage*, 31(3):968–980, 2006.
- [3] Christophe Destrieux, Bruce Fischl, Anders Dale, and Eric Halgren. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage*, 53(1):1–15, 2010.
- [4] Alex Fornito, Andrew Zalesky, and Edward T. Bullmore. *Fundamentals of Brain Network Analysis*. Elsevier/Academic Press, 2016.
- [5] Luis M. Colon-Perez, Michelle Couret, William Triplett, Catherine C. Price, and Thomas H. Mareci. Small worldness in dense and weighted connectomes. *Frontiers in Physics*, 4(MAY):1–15, 2016.
- [6] Danielle S. Bassett and Olaf Sporns. Network neuroscience. *Nature Neuroscience*, 20(3):353–364, 3 2017.
- [7] Andrew Zalesky, Alex Fornito, and Edward T. Bullmore. Network-based statistic: Identifying differences in brain networks. *NeuroImage*, 53(4):1197–1207, 12 2010.
- [8] David Berron, Danielle van Westen, Rik Ossenkoppele, Olof Strandberg, and Oskar Hansson. Medial temporal lobe connectivity and its associations with cognition in early alzheimer’s disease. *Brain*, 143(4):1233–1248, 4 2020.
- [9] Sara Isernia, Alice Pirastru, Davide Massaro, Marco Rovaris, Antonella Marchetti, and Francesca Baglio. Resting-state functional brain connectivity for human mentalizing: Biobehavioral mechanisms of theory of mind in multiple sclerosis. *Social Cognitive and Affective Neuroscience*, 11 2021.
- [10] Cécile Bordier, Carlo Nicolini, Giulia Forcellini, and Angelo Bifone. Disrupted modular organization of primary sensory brain areas in schizophrenia. *NeuroImage: Clinical*, 18:682–693, 1 2018.
- [11] Davide Coluzzi, Alice Pirastru, Laura Pelizzari, Monia Cabinio, Maria Marcella Laganà, Giuseppe Baselli, and Francesca Baglio. Development and testing of spider-net: An interactive tool for brain connectogram visualization, sub-network exploration and graph metrics quantification. *Frontiers in Neuroscience*, 16, 2022.
- [12] Bastian Cheng, Eckhard Schlemm, Robert Schulz, Marlene Boenstrup, Arnaud Messé, Claus Hilgetag, Christian Gerloff, and Götz Thomalla. Altered topology of large-scale structural brain networks in chronic stroke. *Brain Communications*, 1(1), 1 2019.
- [13] Madelaine Daianu, Emily L. Dennis, Neda Jahanshad, Talia M. Nir, Arthur W. Toga, Clifford R. Jack, Michael W. Weiner, and Paul M. Thompson. Alzheimer’s disease disrupts rich club organization in brain connectivity networks. *Proceedings - International Symposium on Biomedical Imaging*, pages 266–269, 2013.

- [14] Jonathan J Crofts, Desmond J Higham, Rose Bosnell, Saâd Jbabdi, Paul M Matthews, TEJ Behrens, and Heidi Johansen-Berg. Network analysis detects changes in the contralesional hemisphere following stroke. *NeuroImage*, 54(1):161–169, 1 2011.
- [15] Alex Fornito, Andrew Zalesky, Christos Pantelis, and Edward T. Bullmore. Schizophrenia, neuroimaging and connectomics. *NeuroImage*, 62(4):2296–2314, 10 2012.
- [16] Yong Liu, Meng Liang, Yuan Zhou, Yong He, Yihui Hao, Ming Song, Chunshui Yu, Haihong Liu, Zhening Liu, and Tianzi Jiang. Disrupted small-world networks in schizophrenia. *Brain*, 131(4):945–961, 2008.
- [17] Aaron F. Alexander-Bloch, Nitin Gogtay, David Meunier, Rasmus Birn, Liv Clasen, Francois Lalonde, Rhoshel Lenroot, Jay Giedd, and Edward T. Bullmore. Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia. *Frontiers in Systems Neuroscience*, 4(October):1–16, 2010.
- [18] Mary-Ellen Lynall, Danielle S Bassett, Robert Kerwin, Peter J McKenna, Manfred Kitzbichler, Ulrich Muller, and Ed Bullmore. Functional connectivity and brain networks in schizophrenia. *Journal of Neuroscience*, 30(28):9477–9487, 7 2010.
- [19] Alex Fornito, Jong Yoon, Andrew Zalesky, Edward T. Bullmore, and Cameron S. Carter. General and specific functional connectivity disturbances in first-episode schizophrenia during cognitive control performance. *Biological Psychiatry*, 70(1):64–72, 7 2011.
- [20] Pawel Skudlarski, Kanchana Jagannathan, Karen Anderson, Michael C. Stevens, Vince D. Calhoun, Beata A. Skudlarska, and Godfrey Pearlson. Brain connectivity is not only lower but different in schizophrenia: A combined anatomical and functional approach. *Biological Psychiatry*, 68(1):61–69, 7 2010.
- [21] Mao Lin Hu, Xiao Fen Zong, J. John Mann, Jun Jie Zheng, Yan Hui Liao, Zong Chang Li, Ying He, Xiao Gang Chen, and Jin Song Tang. A review of the functional and anatomical default mode network in schizophrenia. *Neuroscience Bulletin*, 33(1):73–84, 2017.
- [22] Douglass Godwin, Andrew Ji, Sridhar Kandala, and Daniel Mamah. Functional connectivity of cognitive brain networks in schizophrenia during a working memory task. *Frontiers in Psychiatry*, 8, 2017.
- [23] Yuan Zhou, Peter Zeidman, Shihao Wu, Adeel Razi, Cheng Chen, Liuqing Yang, Jilin Zou, Gaohua Wang, Huiling Wang, and Karl J. Friston. Altered intrinsic and extrinsic connectivity in schizophrenia. *NeuroImage: Clinical*, 17:704–716, 1 2018.
- [24] R. Matthew Hutchison, Thilo Womelsdorf, Elena A. Allen, Peter A. Bandettini, Vince D. Calhoun, Maurizio Corbetta, Stefania Della Penna, Jeff H. Duyn, Gary H. Glover, Javier Gonzalez-Castillo, Daniel A. Handwerker, Shella Keilholz, Vesa Kiviniemi, David A. Leopold, Francesco de Pasquale, Olaf Sporns, Martin Walter, and Catie Chang. Dynamic functional connectivity: Promise, issues, and interpretations. *NeuroImage*, 80:360–378, 10 2013.
- [25] Akhil Kottaram, Leigh A. Johnston, Luca Cocchi, Eleni P. Ganella, Ian Everall, Christos Pantelis, Ramamohanarao Kotagiri, and Andrew Zalesky. Brain network dynamics in schizophrenia: Reduced dynamism of the default mode network. *Human Brain Mapping*, 40(7):2212–2228, 2019.
- [26] Raphaël Liégeois, Timothy O. Laumann, Abraham Z. Snyder, Juan Zhou, and B. T. Thomas Yeo. Interpreting temporal fluctuations in resting-state functional connectivity mri. *NeuroImage*, 163:437–455, 12 2017.
- [27] Rikkert Hindriks, Mohit H Adhikari, Yusuke Murayama, Marco Ganzetti, Dante Mantini, Nikos K Logothetis, and Gustavo Deco. Can sliding-window correlations reveal dynamic functional connectivity in resting-state fmri? *NeuroImage*, 127:242–256, 2 2016.

- [28] Nora Leonardi and Dimitri Van De Ville. On spurious and real fluctuations of dynamic functional connectivity during rest. *NeuroImage*, 104:430–436, 1 2015.
- [29] Martin A. Lindquist, Yuting Xu, Mary Beth Nebel, and Brain S. Caffo. Evaluating dynamic bivariate correlations in resting-state fmri: A comparison study and a new approach. *NeuroImage*, 101:531–546, 11 2014.
- [30] Andrew Zalesky and Michael Breakspear. Towards a statistical test for functional connectivity dynamics. *NeuroImage*, 114:466–470, 7 2015.
- [31] Mariana Lazar and Andrew L. Alexander. Bootstrap white matter tractography (boot-trac). *NeuroImage*, 24(2):524–532, 2005.
- [32] Anthony Kulesa, Martin Krzywinski, Paul Blainey, and Naomi Altman. Sampling distributions and the bootstrap. *Nature methods*, 12(6):477–478, 6 2015.
- [33] Yulia R. Gel, Vyacheslav Lyubchich, and L. Leticia Ramirez Ramirez. Bootstrap quantification of estimation uncertainties in network degree distributions. *Scientific Reports*, 7(1):5807, 7 2017.
- [34] Alden Green and Cosma Rohilla Shalizi. Bootstrapping exchangeable random graphs. *Electronic Journal of Statistics*, 16(1):1058–1095, 1 2022.
- [35] Victor Picheny, Nam Ho Kim, and Raphael T. Haftka. Application of bootstrap method in conservative estimation of reliability with limited samples. *Structural and Multidisciplinary Optimization*, 41(2):205–217, 3 2010.
- [36] Lijiang Wei, Bin Jing, and Haiyun Li. Bootstrapping promotes the rsfc-behavior associations: An application of individual cognitive traits prediction. *Human Brain Mapping*, 41(9):2302–2316, 2020.
- [37] Brian D. Berman, Jason Smucny, Corey P. Wylie, Erika Shelton, Eugene Kronberg, Maureen Leehey, and Jason R. Tregellas. Levodopa modulates small-world architecture of functional brain networks in parkinson’s disease. *Movement Disorders*, 31(11):1676–1684, 2016.
- [38] Martijn P. van den Heuvel, Siemon C. de Lange, Andrew Zalesky, Caio Seguin, B. T. Thomas Yeo, and Ruben Schmidt. Proportional thresholding in resting-state fmri functional connectivity networks and consequences for patient-control connectome studies: Issues and recommendations. *NeuroImage*, 152:437–449, 5 2017.
- [39] Julia M. Sheffield and Deanna M. Barch. Cognition and resting-state functional connectivity in schizophrenia. *Neuroscience & Biobehavioral Reviews*, 61:108–120, 2 2016.
- [40] Mark E.J. Newman. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582, 6 2006.
- [41] Bradley Efron and Robert Tibshirani. *An introduction to the bootstrap*. Monographs on statistics and applied probability. Chapman & Hall, New York, 1993.
- [42] Matthew G. Johnston and Christine Faulkner. A bootstrap approach is a superior statistical method for the comparison of non-normal data with differing variances. *New Phytologist*, 230(1):23–26, 2021.
- [43] Katell Mevel, Gaël Chételat, Francis Eustache, and Béatrice Desgranges. The default mode network in healthy aging and alzheimer’s disease. *International Journal of Alzheimer’s Disease*, 2011:535816, 6 2011.
- [44] Wanqing Li, Xiaoqin Mai, and Chao Liu. The default mode network and social understanding of others: what do brain connectivity studies tell us. *Frontiers in Human Neuroscience*, 8, 2014.

- [45] Kevin J. Manning and David C. Steffens. Systems neuroscience in late-life depression. In *Systems Neuroscience in Depression*, pages 325–340. Academic Press, San Diego, 1 2016.
- [46] Pedro Nascimento Alves, Chris Foulon, Vyacheslav Karolis, Danilo Bzdok, Daniel S. Margulies, Emmanuelle Volle, and Michel Thiebaut de Schotten. An improved neuroanatomical model of the default-mode network reconciles previous neuroimaging and neuropathological findings. *Communications Biology*, 2(1):1–14, 10 2019.
- [47] Martijn P. van den Heuvel and Hilleke E. Hulshoff Pol. Exploring the brain network: A review on resting-state fmri functional connectivity. *European Neuropsychopharmacology*, 20(8):519–534, 8 2010.
- [48] Mikail Rubinov and Olaf Sporns. Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, 52(3):1059–1069, 9 2010.
- [49] Valeria Blasi, Alice Pirastru, Monia Cabini, Sonia Di Tella, Maria Marcella Laganà, Alice Giangiacomo, Gisella Baglio, Michela Zanette, Maria Paola Canevini, Mauro Walder, Mario Clerici, and Francesca Baglio. Early life adversities and borderline intellectual functioning negatively impact limbic system connectivity in childhood: A connectomics-based study. *Frontiers in Psychiatry*, 11, 2020.
- [50] Danielle S. Bassett and Edward T. Bullmore. Small-world brain networks revisited. *The Neuroscientist*, 23(5):499–516, 10 2017. publisher: SAGE Publications Inc STM.
- [51] Isaura Oliver, Jaroslav Hlinka, Jakub Kopal, and Jörn Davidsen. Quantifying the variability in resting-state networks. *Entropy*, 21(9):882, 9 2019. number: 9 publisher: Multidisciplinary Digital Publishing Institute.
- [52] Suzanne N. Avery, Baxter P. Rogers, and Stephan Heckers. Hippocampal network modularity is associated with relational memory dysfunction in schizophrenia. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 3(5):423–432, 5 2018.
- [53] Daan van den Berg, Pulin Gong, Michael Breakspear, and Cees van Leeuwen. Fragmentation: loss of global coherence or breakdown of modularity in functional brain architecture? *Frontiers in Systems Neuroscience*, 6, 2012.
- [54] Martijn P. van den Heuvel and Alex Fornito. Brain networks in schizophrenia. *Neuropsychology Review*, 24(1):32–48, 3 2014.
- [55] Qingbao Yu, Jing Sui, Srinivas Rachakonda, Hao He, William Gruner, Godfrey Pearlson, Kent A. Kiehl, and Vince D. Calhoun. Altered topological properties of functional network connectivity in schizophrenia during resting state: A small-world brain network study. *PLOS ONE*, 6(9):e25423, 2011. Public Library of Science.
- [56] Liang Wang, Paul D. Metzak, William G. Honer, and Todd S. Woodward. Impaired efficiency of functional networks underlying episodic memory-for-context in schizophrenia. *Journal of Neuroscience*, 30(39):13171–13179, 9 2010.
- [57] Yu Sun, Simon L Collinson, John Suckling, and Kang Sim. Dynamic reorganization of functional connectivity reveals abnormal temporal efficiency in schizophrenia. *Schizophrenia Bulletin*, 45(3):659–669, 4 2019.
- [58] Andrew Zalesky, Alex Fornito, Luca Cocchi, Leonardo L. Gollo, and Michael Breakspear. Time-resolved resting-state brain networks. *Proceedings of the National Academy of Sciences*, 111(28):10341–10346, 7 2014.
- [59] Paul G. Unschuld, Alison S. Buchholz, Mark Varvaris, Peter C. M. van Zijl, Christopher A. Ross, James J. Pekar, Christoph Hock, John A. Sweeney, Carol A. Tamminga, Matcheri S. Keshavan, Godfrey D. Pearlson, Guntav K. Thaker, and David J. Schretlen. Prefrontal brain network connectivity indicates degree of both schizophrenia risk and cognitive dysfunction. *Schizophrenia Bulletin*, 40(3):653–664, 5 2014.

- [60] Eswar Damaraju, Elena A Allen, Aysenil Belger, Judith M Ford, S McEwen, DH Mathalon, BA Mueller, GD Pearlson, SG Potkin, A Preda, et al. Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia. *NeuroImage: Clinical*, 5:298–308, 1 2014.
- [61] Oren Civier, Robert Elton Smith, Chun-Hung Yeh, Alan Connelly, and Fernando Calamante. Is removal of weak connections necessary for graph-theoretical analysis of dense weighted structural connectomes from diffusion mri? *NeuroImage*, 194:68–81, 7 2019.
- [62] Patric Hagmann, Leila Cammoun, Xavier Gigandet, Reto Meuli, Christopher J. Honey, Van J. Wedeen, and Olaf Sporns. Mapping the structural core of human cerebral cortex. *PLOS Biology*, 6(7):e159, 2008.
- [63] Prejaas Tewarie, Edwin van Dellen, Arjan Hillebrand, and Cornelis J Stam. The minimum spanning tree: An unbiased method for brain network analysis. *NeuroImage*, 104:177–188, 1 2015.
- [64] Danielle S. Bassett, Brent G. Nelson, Bryon A. Mueller, Jazmin Camchong, and Kelvin O. Lim. Altered resting state complexity in schizophrenia. *NeuroImage*, 59(3):2196–2207, 2 2012.
- [65] David R. Hemsley. Perceptual and cognitive abnormalities as the bases for schizophrenic symptoms. In *The Neuropsychology Of Schizophrenia*. Psychology Press, 2019. number-of-pages: 20.
- [66] Dardo Tomasi and Nora D. Volkow. Association between functional connectivity hubs and brain networks. *Cerebral Cortex*, 21(9):2003–2013, 9 2011.
- [67] Bumseok Jeong, Cynthia G. Wible, Ryu-Ichiro Hashimoto, and Marek Kubicki. Functional and anatomical connectivity abnormalities in left inferior frontal gyrus in schizophrenia. *Human Brain Mapping*, 30(12):4138–4151, 2009.
- [68] Mladen Sormaz, Charlotte Murphy, Hao-ting Wang, Mark Hymers, Theodoros Karapanagiotidis, Giulia Poerio, Daniel S. Margulies, Elizabeth Jefferies, and Jonathan Smallwood. Default mode network can support the level of detail in experience during active task states. *Proceedings of the National Academy of Sciences of the United States of America*, 115(37):9318–9323, 9 2018.
- [69] Ann K. Shinn, Justin T. Baker, Bruce M. Cohen, and Dost Öngür. Functional connectivity of left heschl’s gyrus in vulnerability to auditory hallucinations in schizophrenia. *Schizophrenia Research*, 143(2):260–268, 2 2013.
- [70] Daisuke Koshiyama, Makoto Miyakoshi, Kumiko Tanaka-Koshiyama, Yash B. Joshi, Juan L. Molina, Joyce Sprock, David L. Braff, and Gregory A. Light. Neurophysiologic characterization of resting state connectivity abnormalities in schizophrenia patients. *Frontiers in Psychiatry*, 11, 2020.
- [71] Noa Peskin, Dan Koren, and Shai Gabay. Subcortical neural tracks play an important role in executive function in schizophrenia: An experimental study among patients with schizophrenia and healthy comparisons. *Schizophrenia Research: Cognition*, 22:100185, 12 2020.
- [72] Anthony Beech, Trevor Powell, Jonathan McWilliam, and Gordon Claridge. Evidence of reduced ‘cognitive inhibition’ in schizophrenia. *British Journal of Clinical Psychology*, 28(2):109–116, 1989.
- [73] Ilona Henseler, Peter Falkai, and Oliver Gruber. Disturbed functional connectivity within brain networks subserving domain-specific subcomponents of working memory in schizophrenia: Relation to performance and clinical symptoms. *Journal of Psychiatric Research*, 44(6):364–372, 4 2010.

- [74] Toshiro Fujimoto, Eiichi Okumura, Kouzou Takeuchi, Atsushi Kodabashi, Toshiaki Otsubo, Katsumi Nakamura, Shinichiro Kamiya, Yuji Higashi, Tadahiko Yuji, Kenichi Honda, Susumu Shimooki, and Toshiyo Tamura. Dysfunctional cortical connectivity during the auditory odd-ball task in patients with schizophrenia. *The Open Neuroimaging Journal*, 7:15–26, 4 2013.
- [75] Murat Yildiz, Stefan J. Borgwardt, and Gregor E. Berger. Parietal lobes in schizophrenia: Do they matter? *Schizophrenia Research and Treatment*, 2011:e581686, 10 2011.
- [76] William Pettersson-Yeo, Paul Allen, Stefania Benetti, Philip McGuire, and Andrea Mechelli. Dysconnectivity in schizophrenia: Where are we now? *Neuroscience & Biobehavioral Reviews*, 35(5):1110–1124, 4 2011.
- [77] Michael D. Fox, Abraham Z. Snyder, Justin L. Vincent, Maurizio Corbetta, David C. Van Essen, and Marcus E. Raichle. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences*, 102(27):9673–9678, 7 2005.
- [78] Peter Fransson. Spontaneous low-frequency bold signal fluctuations: An fmri investigation of the resting-state default mode of brain function hypothesis. *Human Brain Mapping*, 26(1):15–29, 2005.
- [79] Yuan Zhou, Meng Liang, Lixia Tian, Kun Wang, Yihui Hao, Haihong Liu, Zhening Liu, and Tianzi Jiang. Functional disintegration in paranoid schizophrenia using resting-state fmri. *Schizophrenia Research*, 97(1):194–205, 12 2007.
- [80] Devorah Segal, M. Mehmet Haznedar, Erin A. Hazlett, Jonathan J. Entis, Randall E. Newmark, Yuliya Torosjan, Jason S. Schneiderman, Joseph Friedman, King-Wai Chu, Cheuk Y. Tang, Monte S. Buchsbaum, and Patrick R. Hof. Diffusion tensor anisotropy in the cingulate gyrus in schizophrenia. *NeuroImage*, 50(2):357–365, 4 2010.
- [81] Cécile Bordier, Carlo Nicolini, and Angelo Bifone. Graph analysis and modularity of brain functional connectivity networks: Searching for the optimal threshold. *Frontiers in Neuroscience*, 11, 2017.
- [82] Carlo Nicolini, Cécile Bordier, and Angelo Bifone. Community detection in weighted brain connectivity networks beyond the resolution limit. *NeuroImage*, 146:28–39, 2 2017.
- [83] Shu Guo, Xiaoqi Chen, Yimeng Liu, Rui Kang, Tao Liu, and Daqing Li. Percolation analysis of brain structural network. *Frontiers in Physics*, 9, 2021.

Biomarker Investigation using Multiple Brain Measures from MRI through XAI in Alzheimer's Disease Classification

D. Coluzzi^{1,*}, V. Bordin^{1,*}, M. W. Rivolta², I. Fortel³, L. Zhan⁴, A. Leow^{3,5}, G. Baselli¹

¹ Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy

² Dipartimento di Informatica, Università degli Studi di Milano, Milan, Italy

³ Department of Biomedical Engineering, University of Illinois Chicago, Chicago, IL, USA

⁴ Department of Electrical and Computer Engineering, University of Pittsburgh, Pittsburgh, PA, USA

⁵ Department of Psychiatry, University of Illinois Chicago, Chicago, IL, USA

* Equal contributions

Corresponding author:

Davide Coluzzi

Politecnico di Milano

Department of Electronics, Information and Bioengineering

Via Camillo Golgi 39, 20133 Milan

davide.coluzzi@polimi.it

ABSTRACT

Alzheimer's Disease (AD) is the world leading cause of dementia, a progressively impairing condition leading to high hospitalization rates and mortality. To optimize the diagnostic process, numerous efforts have been directed towards the development of deep learning approaches (DL) for the automatic AD classification. However, their typical black box outline has led to low trust and scarce usage within clinical frameworks. In this work, we propose two state-of-the-art DL models, trained respectively on structural MRI (ResNet18) and brain connectivity matrixes (BC-GCN-SE) derived from diffusion data. The models were initially evaluated in terms of classification accuracy. Then, results were analyzed using an Explainable Artificial Intelligence (XAI) approach (Grad-CAM) to measure the level of interpretability of both models. The XAI assessment was conducted across 132 brain parcels, extracted from a combination of the Harvard-Oxford and AAL brain atlases, and compared to well-known pathological regions to measure adherence to domain knowledge. Results highlighted acceptable classification performance as compared to the existing literature (ResNet18: $TPR_{median} = 0.817$, $TNR_{median} = 0.816$; BC-GCN-SE: $TPR_{median} = 0.703$, $TNR_{median} = 0.738$). As evaluated through a statistical test ($p < 0.05$) and ranking of the most relevant parcels (first 15%), Grad-CAM revealed the involvement of target brain areas for both the ResNet18 and BC-GCN-SE models: the medial temporal lobe and the default mode network. The obtained interpretabilities were not without limitations. Nevertheless, results suggested that combining different imaging modalities may result in increased classification performance and model reliability. This could potentially boost the confidence laid in DL models and favor their wide applicability as aid diagnostic tools.

Table of Acronyms

| | | | | | |
|-----|------------------------|--------|--|-----|---------|
| FP | Frontal Pole | pTFusC | Temporal Fusiform Cortex, posterior division | CaN | Caudate |
| SFG | Superior Frontal Gyrus | CO | Central Opercular Cortex | Pu | Putamen |

| | | | | | |
|----------|---|--------|---|---------|------------------|
| MidFG | Middle Frontal Gyrus | PP | Planum Polare | Pal | Pallidum |
| IFG_tri | Inferior Frontal Gyrus, pars triangularis | HG | Heschls Gyrus | Hip | Hippocampus |
| IFG_oper | Inferior Frontal Gyrus, pars opercularis | PT | Planum Temporale | Amg | Amygdala |
| PreCG | Precentral Gyrus | PostCG | Postcentral Gyrus | NAcc | Accumbens |
| SMA | Juxtapositional Lobule Cortex - formerly Supplementary Motor Cortex | SPL | Superior Parietal Lobule | Cereb1 | Cerebellum Crus1 |
| FOrb | Frontal Orbital Cortex | aSMG | Supramarginal Gyrus, anterior division | Cereb2 | Cerebellum Crus2 |
| FO | Frontal Operculum Cortex | pSMG | Supramarginal Gyrus, posterior division | Cereb3 | Cerebellum 3 |
| MedFC | Frontal Medial Cortex | AG | Angular Gyrus | Cereb45 | Cerebellum 4 5 |
| IC | Insular Cortex | PO | Parietal Operculum Cortex | Cereb6 | Cerebellum 6 |
| AC | Cingulate Gyrus, anterior division | PrCun | Precuneous Cortex | Cereb7 | Cerebellum 7b |
| PC | Cingulate Gyrus, posterior division | sLOC | Lateral Occipital Cortex, superior division | Cereb8 | Cerebellum 8 |
| SubCalC | Subcallosal Cortex | iLOC | Lateral Occipital Cortex, inferior division | Cereb9 | Cerebellum 9 |
| PaCiG | Paracingulate Gyrus | ICC | Intracalcarine Cortex | Cereb10 | Cerebellum 10 |
| TP | Temporal Pole | Cuneal | Cuneal Cortex | Ver12 | Vermis 1 2 |
| aSTG | Superior Temporal Gyrus, anterior division | aPaHC | Parahippocampal Gyrus, anterior division | Ver3 | Vermis 3 |
| pSTG | Superior Temporal Gyrus, posterior division | pPaHC | Parahippocampal Gyrus, posterior division | Ver45 | Vermis 4 5 |
| aMTG | Middle Temporal Gyrus, anterior division | LG | Lingual Gyrus | Ver6 | Vermis 6 |
| pMTG | Middle Temporal Gyrus, posterior division | TOFusC | Temporal Occipital Fusiform Cortex | Ver7 | Vermis 7 |
| toMTG | Middle Temporal Gyrus, temporooccipital part | OFusG | Occipital Fusiform Gyrus | Ver8 | Vermis 8 |
| aITG | Inferior Temporal Gyrus, anterior division | SCC | Supracalcarine Cortex | Ver9 | Vermis 9 |
| pITG | Inferior Temporal Gyrus, posterior division | OP | Occipital Pole | Ver10 | Vermis 10 |
| tolTG | Inferior Temporal Gyrus, temporooccipital part | Tha | Thalamus | BSt | Brain-Stem |
| aTFusC | Temporal Fusiform Cortex, anterior division | | | | |

Brain parcel acronyms of the Harvard-Oxford combined with AAL atlas.

1. INTRODUCTION

As populations continue to age, dementia cases are on the rise, which pose a serious public health risk and place a huge social and economic burden on many countries. The most prevalent cause of dementia is Alzheimer’s Disease (AD), a neurodegenerative pathology occurring when nerve cells die in the brain, which initially manifests with impaired memory [1]. Dementia is a common disorder among the elderly population which is associated with significant disability, increased hospitalization, and mortality. In recent times, research focused extensively on AD and neuroimaging techniques emerged as a critical tool for diagnosing and monitoring the disease progressions. Structural brain changes, such as those visible on magnetic resonance imaging (MRI), can reveal the early involvement of the medial temporal lobe (MTL) in AD, particularly in terms of hippocampal and parahippocampal atrophy [2], [3]. Additionally, numerous studies linked the entorhinal cortex to changes in the cognitive performance of diseased individuals [4]. The brain cortex is also involved in the neurodegenerative process underlying AD. Both temporal and parietal regions present atrophic changes in correspondence of the early stages

of the condition [5]. On the other hand, the symptomatic progression is often accompanied by a more extensive brain cortical thinning alongside ventricular enlargement [2], [6].

In addition, AD assessment can be supplemented with complementary information through diffusion MRI (dMRI) and fiber tracking due to dendritic, myelin and axonal loss which accompanies atrophy [4]. From dMRI measures, it is possible to extract structural connectivity data, depicting the brain as a graph, parcels as nodes, and derived dMRI metrics (such as the number of reconstructed streamlines) as edges. Coherent changes across different dMRI metrics used for edge-weighting in the connectivity structure due to AD were widely documented, occurring across a range of spatial scales and levels [7]. At the global level, AD was depicted as a disconnection syndrome which can be characterized by the connectome degeneration, affecting graph topology, that is governed by long-range connections [8], [9]. At sub-graph level, studies demonstrated that the Default Mode Network (DMN), which is involved in memory processes, is vulnerable to atrophy, amyloid protein deposition, and white matter microstructure alterations resulting in a disrupted structural connectivity configuration [7], [10], [11]. AD abnormalities in structural connectivity and dMRI studies also found the temporal lobe, whose disruptions contribute to memory impairment [7], [12], [13], and some regions of the MTL [14], which are also often reported in the DMN. These findings were corroborated by functional connectivity studies which consistently revealed a decreased functional connectivity between the posterior and anterior portions of the DMN [15], [16]. DMN in general was found to be replicable hallmark and an important area for AD in either structural and functional connectivity analysis, highlighting many common regions [17].

With the growing global incidence of neurodegenerative disorders, such as AD, a heightened interest to advance in areas such as diagnosis, treatment, prevention, drug discovery, and provision of improved healthcare services was noticed. Moreover, clinical decision support systems are possible to be developed through AI methods. Most of the research in the domain of AD, and more in general, neurodegeneration focused on using brain imaging. Traditional Machine Learning methods with well-known classification methods like Logistic Regression, Random Forest and Support Vector Machines were widely utilized both from MRI (mainly T1-weighted, but also Fluid Attenuated Inversion Recovery - FLAIR and T2-weighted contrasts), fMRI images, related features and connectivity data [18], [19].

In addition, the availability of significant computational resources and the advancements of DL algorithms enabled the application of these techniques to improve the accuracy of computer-assisted diagnoses. Also, Deep Learning (DL) models were widely applied for different tasks from both 3D brain volumes, images and connectivity data. As regards models using brain images or volumes, many approaches were investigated, becoming a well-established area of research. In this context, Convolutional Neural Networks (CNNs) have particular importance, since they were recently demonstrated to have remarkable performance in medical analysis, both employing 2D slices and 3D volumes [20]–[22]. In addition, different CNN-based pre-trained models, such as ResNet18, EfficientNet-B0 and VGG etc., were largely employed in neuroimaging research, providing state-of-the-art performance for different tasks [22]–[24]. Subsequently, other approaches, such as Graph Neural Networks (GNNs), were developed. The capabilities of GNNs, which are neural networks that operate on graph data, have recently risen after a decade of development and advancements [25]. These models, along with CNNs and Autoencoders, were adapted and employed to use the brain connectivity matrixes as input.

For example, a connectome-based CNN architecture for classifying Mild Cognitive Impairment was proposed. Two layers of convolution were used, first row by row and then column by column [26]. Similarly, BrainNetCNN is a deeper convolutional network that was also tailored for brain graphs. It uses a CNN-like kernel to compute the convolution of the connections and treats each graph edge as a pixel in an image. The model was validated using structural connectivity matrixes from DTI images for regression task [27]. Subsequently, this model was tested and compared to different ML models with respect to the AD classification task [28]. Alorf and Khan obtained good performance in the AD stages classification employing BC-GCN (Brain Connectivity Graph Convolutional Network), an adapted GNN for brain connectivity data that was previously tested on a regression task [29], and Stacked

Sparse Autoencoders [30]. Also, graph variational autoencoder employing both structural and functional connectivity was employed on AD dataset to find a unified embedding via a classification task [31].

Despite the effectiveness of DL algorithms in various classification tasks, their widespread adoption and trust in clinical settings is limited due to their well-known "black box" nature. The latest advancements in Explainable Artificial Intelligence (XAI) techniques aim to bridge the gap between the performance of DL models and the need for human comprehension of their processes. However, validation on human brain graphs was not widely performed yet, as for T1-weighted, FLAIR, T2-weighted images and volumes.

Among all approaches, the most widespread XAI methods in neuroimaging can be roughly classified into two categories. First, gradient/feature-based methods use gradients or hidden feature maps to determine the importance of different input features in the model's predictions. Examples are Class Activation Mapping (CAM), Grad-CAM with different variations or guided back propagation [32]. They were used in recent imaging studies [21], [33], but also to interpret decision of recent models employing connectivity data [34], [35]. Second, perturbation-based methods evaluate the effect of changes in input information. The importance of input features is measured by monitoring the variation of the model's output for different input perturbations [32]. These methods were widely employed in many studies using images [21], [23], [36] and connectivity data [37].

The lack of interpretability of DL models necessitated to develop ways for a deeper comprehension of the problem at hand. In fact, XAI methodologies may highlight different features of interest for a possible reinforcement of their roles as brain biomarkers of AD. Additionally, the cost of processing for the extraction of the brain connectivity graphs can be evaluated in relation to the different interpretation and accuracy provided by these data in comparison to the use of more conventional brain images and volumes. To the best of our knowledge, no studies compared AI approaches using different sMRI (structural MRI) data to evaluate and address these issues.

In this work, we employed both brain 3D T1-weighted volumes and structural connectivity data from the OASIS-3 dataset [38] for AD classification task. Two state-of-the-art DL models, as well as the inherent characteristics of the data, were compared in terms of accuracy and explainability. More specifically: i) we evaluated the performance of DL models such as ResNet18 and BC-GCN-SE for AD classification using multiple brain data: 3D T1-weighted volumes and structural connectivity data, respectively; ii) we employed an XAI method, Grad-CAM, to assess interpretability of the two DL models. The investigation of XAI was made at different levels highlighting morphological and interregional features of interest in agreement to domain-knowledge and possibly reinforcing their roles as brain biomarkers of AD. Consistent and divergent information employed by the two DL models were analyzed highlighting the advantages of the two methods and data and the potential for the development of superior and more trustworthy DL models.

2. MATERIALS AND METHODS

2.1. Study population

The dataset used in this study is the third release of the Open Access Series of Imaging Studies (OASIS-3), a longitudinal collection of data focused on the effects of normal aging and early-stage AD [38]. The dataset, released in 2019, includes 1098 participants among which 605 cognitively normal adults (i.e., healthy controls – HC) and 493 subjects at various stages of cognitive decline. Each participant underwent both neuroimaging and clinical assessments, that were conducted independently throughout the study.

The final dataset was obtained with the following steps: i) matching the MRI and clinical data by selecting the closest records within a 3-month time span; ii) matching with the dataset of Amodeo and colleagues [31], that extracted the structural connectivity matrixes from OASIS-3; This resulting dataset was composed of 692 sessions relative to 543 participants (age range 42-95 years, mean age 70.06 ± 8.85 years, F:M = 388:304). Each session was associated to a T1-weighted scan and a structural connectivity matrix.

Patients at each session were eventually differentiated in HC and AD according to the Clinical Dementia Rating (CDR) Scale (available in the dataset), with normal cognitive functions being represented by a CDR equal to 0, and diseased conditions by a score of either 0.5 (very mild impairment), 1 (mild impairment), or 2 (moderate dementia). The same subject could have been assigned to both classes because different CDRs were quantified from different imaging sessions. These data (2.76% of the overall subjects) were retained within the dataset but properly handled during the training/validation process of DL models (see Sec. 2.3 for further details). The final ratio between HC and AD sessions was 557:135.

2.2. Data acquisition and data processing

Of all the available T1-weighted scans, 135 were acquired with a 3T Siemens Biograph_mMR scanner, while the remaining 557 with a pair of 3T Siemens TimTrio scanners (Siemens Medical Solutions USA, Inc). Three different imaging sequences were used, as detailed in Table 1.

The imaging scans were initially processed by Amodeo and colleagues [31] to derive the structural connectivity matrixes used to feed the BC-GCN-SE DL model (see Sec. 2.6 for further details). The original T1-weighted volumes were divided in 132 brain-covering regions. Of these, 91 cortical and 15 subcortical parcels were derived from the FSL Harvard-Oxford maximum likelihood cortical atlas (HOA) [39], while the remaining 26 cerebellar parcels were derived from the Automated Anatomical Labelling atlas (AAL) [40]. Henceforth, the combined HOA and AAL atlas will be referred as HOA + AAL. The corresponding DTI volumes were extracted from T1-weighted scans as detailed in [31]. The combination of the outlined gray matter regions with the DTI white matter fiber tracking resulted in 692 undirected graphs (i.e., positively weighted connectivity matrixes) that underwent a minimal data processing step according to procedures described in [31]. A diagram summarizing the extraction process for structural connectivity matrixes is reported in the orange box of Fig. 2.

Then, all the available T1-weighted scans were pre-processed using the FSL v.6.0 tool [41] to create a suitable dataset for the ResNet18 DL model (see Sec. 2.5 for further details). First, images were skull-stripped and bias field corrected using the *fsl_anat* script (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/fsl_anat). Then, they were registered to the standard Montreal Neurological Institute template (1x1x1 mm²) by applying the non-linear warp transformation provided by *fsl_anat*. The images field of view was cropped to a dimension of 148x180x144 voxels to focus on brain tissues while the intensity values were normalised using variance scaling. Eventually, the dimensionality was resized to 115x144x118.

Table 1. Acquisition details for the three sequences involved in our study.

| | 3T Siemens Biograph_mMR | 3T Siemens Biograph_mMR | 3T Siemens TimTrio |
|-------------------------------|----------------------------|-------------------------|--------------------|
| Sequence | T1 (MPRAGE_GRAPPA2) | T1 (MPRAGE isoWU) | T1 (MPRAGE) |
| TR (ms) | 2300 | 2400 | 2400 |
| TE (ms) | 2.95 | 2.13 | 3.16 |
| Flip angle (degrees) | 8 | 8 | 8 |
| Voxel size (mm ³) | 1.20x1.05x1.05 | 1.00x1.00x1.00 | 1.00x1.00x1.00 |
| FOV (mm ²) | 176x240 | 176x256 | 176x256 |
| Slices per slab | 256 | 256 | 256 |
| TI (ms) | 900 | 1000 | 1000 |

| Orientation | Sagittal | Sagittal | Sagittal |
|-------------|----------|----------|----------|
|-------------|----------|----------|----------|

Legend: MPRAGE, Magnetization Prepared Rapid Acquisition Gradient Echo; GRAPPA2, Generalized Autocalibrating Partially Parallel Acquisition version 2; isoWU, isotropic Weighted Undersampling; TR, repetition time; TE, echo time; FOV, field of view; TI, inversion time.

2.3. Training-validation strategy and evaluation

After pre-processing, the dataset was split into training and validation using a stratified 10-fold cross-validation strategy ensuring that each subject was included in only one of the folds. This approach ensures to preserve the ratio between AD and HC sessions.

Having highly imbalanced classes within the dataset, before training the BC-GCN-SE model (see Sec. 2.6 for further details), we applied the synthetic minority over-sampling technique (SMOTE) [42] to the training set. This data augmentation method allows to generate synthetic samples through a linear interpolation of the real neighboring ones, identified through a k-nearest neighbor approach. This methodology has already been applied to connectivity data in previous studies, though with different target tasks [43], [44]. No data augmentation strategy was instead carried out before training the ResNet18 model (see Sec. 2.7 for further details).

For both models, the obtained performances were evaluated by calculating, across all 10 folds, the median and interquartile range of the true positive (TPR) and true negative (TNR) rates and the median and interquartile range of the classification accuracy.

Once the performance was confirmed to be satisfactory through the k-fold cross validation, the final model was trained by performing a new split of the original dataset while maintaining the same proportion of the cross-validation step (i.e., 90% and 10% of the entire dataset) and the same hyperparameters. This was performed to assess the level of explainability of the DL models under evaluation (see Sec 2.6 and 2.7 for further details). This configuration was used to derive the final AD/HC classification for the entire dataset. Results were evaluated in terms of TPR, TNR and classification accuracy while considering the union of both training and validation data.

2.4. ResNet18: Deep Learning model for 3D T1-weighted volumes

The recognition of AD from 3D T1-weighted volumes was conducted using Resnet18, a pre-trained DL model that was trained for multi-class classification of images of the ImageNet database [45], which was adapted for 3D inputs [46]. ResNet18 was selected based on previous studies which demonstrated its good performance in AD recognition compared to other pre-trained models [36]. The model is characterized by 18 layers and its general structure includes a 3D convolutional layer and four sets of residual blocks, each containing two 3D convolutional layers, with a shortcut connection that bypasses the convolutional layers and adds the input directly to the output of the second convolutional layer. Transfer learning was employed on the pre-trained model by adding a set of layers consisting of a Global Average Pooling (GAP) layer, two fully connected layers (128 and 32 neurons) with Rectified Linear Unit (ReLU) activation, and a sigmoid activation unit, as depicted at the top of the Fig. 1.

The binary cross-entropy loss function was used for training, with a batch size of 16 and class weighting. The best model was chosen by applying the early stopping criterion to the validation loss, with a patience of 8. The model with the highest Area Under the Curve (AUC) value across training epochs was retained.

2.5. BC-GCN-SE: Deep Learning model for structural connectivity graphs

As mentioned in the Introduction, many GCNs were recently proposed with different purposes and in different applications. However, GCN is a node-based DL method based on features related to the different nodes. These kinds of models focus the pattern extraction on nodes and features, learning the messages passing through the connections. This structure was demonstrated to be very suitable for sparse graph-structured data, containing only

a few edges connecting nodes. Conversely, the brain connectivity data are known to yield dense graphs. In the structural case, the DTI-derived metrics used after fiber tractography, such as the number of reconstructed streamlines or the fractional anisotropy, result in graphs with a high number of connections, far from being sparse. Also, latest advancements in the fiber tracking and connectivity definition methodologies facilitated the calculation of dense weighted connectomes also for structural connectivity [47]. As a result, both direct and indirect connections between different brain areas result to be crucial for brain communication. For these reasons, common GCNs usually employed with success in other research fields can be unsuitable for the study of structural and functional connectivity. At the same time, an edge-based GCN, the so-called BC-GCN, was recently proposed and specifically employed in functional connectivity tasks [29], [30]. The BC-GCN model with the possible extension of a Squeeze-and-Excitation block (BC-GCN-SE) resulted in performing well in regression and classification tasks. For this reason, we modified this model to be adapted to our AD classification task with structural connectivity data.

BC-GCN-SE is mainly composed of five major units: the graph path convolution (GPC), which allow the extraction of the feature maps, the edge (EP) and node pooling (NP), the Squeeze-and-Excitation (SE), to emphasize or suppress them, and the fully connected block for classification.

As said, the communication between different brain areas is achieved through a combination of direct and indirect connections. To extract significant information from these pathways and represent the multi-order information, GPC layers are utilized. This module represents the counterpart of convolutional layers in CNNs leveraging meaningful characteristics from high-order paths by stacking multiple layers. EP and NP pooling layers are employed to aggregate information from nodes and edges downsampling the feature maps, which are the output of the convolutional layers. EP and NP are the counterpart of pooling layers in typical CNNs. The SE block is constructed as the typical SE model [29], but the convolution is modified to correspond to the GPC outlined above. SE layer is positioned after each of the three GPC layers. Finally, the classification part, composed of fully connected layers, was composed of two fully connected blocks after the NP and before the final sigmoid activation, adapting the network to a binary classification task. The final employed architecture of the BC-GCN-SE model is summarized in the green block of the Fig. 2.

The binary cross-entropy loss function and batch size of 64 were employed during training. The model selection process involved utilizing the early stopping. Whether a minimum validation loss was not achieved within a certain number of epochs, thus stop decreasing, the training was halted. The model with greater AUC value was retained.

2.6. Grad-CAM

To perform XAI on the model's predictions, the Grad-CAM method was employed, which is a technique that generates heatmaps to visualize the important regions of an input image that the model uses for classification [48]. Unlike standard CAM, Grad-CAM is a more general approach that does not require any changes to the architecture of the CNN. CAM, on the other hand, requires the removal of the fully connected block of the network and the addition of a GAP layer followed by a single fully connected layer to obtain a relevance heatmap. Using Grad-CAM, it is possible to employ the original CNN architecture, regardless the application. Indeed, the greater flexibility is given by the fact that it is not only limited to classification tasks and allows to obtain relevance heatmaps from any layer of the network, representing features at different granularities [48]. Grad-CAM, thus, produces a heatmap $g \in R^{x \times y}$ to identify the regions of an input image $i \in R^{X \times Y}$, having the greatest influence on the classifier score in support of a given class c . In the context of our application, i represents the structural connectivity matrix, where X and Y are the same, referring to the number of rows/columns, or the 3D T1-weighted volume extending what previously reported to the third dimension, with X , Y and Z equal to the 3D voxel position (3D scan $i \in R^{X \times Y \times Z}$; 3D heatmap: $g \in R^{x \times y \times z}$).

The score for class c is denoted by s_c , while $a^k \in R^{x \times y}$, $k = 1, \dots, K$ represents the activation maps that correspond to the k -th filter of the selected convolutional layer. The Grad-CAM for class c is achieved by computing a weighted average of a^k , $k = 1, \dots, K$, followed by a Rectified Linear Unit (ReLU) activation function to consider only the positive contributions:

$$g_c = \text{ReLU} \sum_k w_c^k a^k \quad (1)$$

Where:

$$w_c^k = \frac{1}{xy} \sum_x \sum_y \frac{\partial s_c}{\partial a^k(x, y)} \quad (2)$$

w_c^k are the average derivatives of s_c with respect to each element (x, y) of the input matrix in the activation a^k and are indicated as *relevance weights*. The same is achieved for 3D T1-weighted volumes simply computing the average derivatives with respect to elements (x, y, z) and dividing by xyz .

Typically, following the implementation of CAM, the last convolutional layer is chosen. The resulting heatmap g_c generated by Grad-CAM exhibits low resolution due to the architecture of the model. Consequently, it is necessary to up-sample the heatmap to match the input image size, using bicubic interpolation. This procedure facilitates the overlay of Grad-CAM onto the input image, which enhances the interpretation of model decisions. An alternative method to increase the resolution of Grad-CAM involves selecting activations from previous convolutional layers of the network that exhibit higher spatial resolution.

This trade-off between the identification of class-discriminative features with low spatial extent and fine-grained details and thus the resolution of the Grad-CAM was addressed in several works [49]–[51]. Since the goal of this work was to analyze the contribution to the final classification of features of diverse scale (see Par 2.7 for further details) and especially last layer of ResNet18 produced coarse-grained ones due to inherent architecture, we computed the heatmaps at different layers for every session. Such heatmaps were then averaged for each structural connectivity matrix or 3D T1-weighted volume - as performed in [50] and implemented in [52] - according to the following equations:

$$G_c(x, y) = \frac{\sum_{l=1}^L \sum_{x=1}^{132} \sum_{y=1}^{132} g_c^l(x, y)}{L} \quad (3)$$

$$G_c(x, y, z) = \frac{\sum_{l=1}^L \sum_{x=1}^{115} \sum_{y=1}^{144} \sum_{z=1}^{118} g_c^l(x, y, z)}{L} \quad (4)$$

Where G_c is the resulting mean heatmap and L is the number of layers considered for each model. More specifically, the output layers of the four stages dividing ResNet18 and the three GPC-SE blocks in BC-GCN-SE model were considered.

2.7. Explainability assessment

In order to determine the most important information leveraged by the models, we assessed the contributions of each parcel. More specifically, we related the mean heatmaps of AD and HC sessions obtained from both ResNet18

and BC-GCN-SE to the HOA + AAL atlas to assess which brain regions were mostly involved in the classification. This same parcellation allowed for a direct comparison between the level of explainability of the two models. The involved atlas consisted of the combination between 106 regions of the HOA and 26 cerebellar parcels of the AAL. More precisely, it was used to highlight either nodes from the structural connectivity matrixes or volumetric regions from the 3D T1-weighted scans. Thus, from the mean heatmap of one session, the contributions highlighted by each connection of each node (i.e., parcel) of connectivity matrixes or in each voxel of each parcel p of 3D T1-weighted scans were averaged to extract a quantity hereafter called relevance value (RV):

$$RV_{p,c} = \frac{\sum_{x=1}^{115} \sum_{y=1}^{144} \sum_{z=1}^{118} M_p(x, y, z) G_c(x, y, z)}{\sum_{x=1}^{115} \sum_{y=1}^{144} \sum_{z=1}^{118} M_p(x, y, z)} \quad (5)$$

$$RV_{p,c} = \frac{\sum_{q=1, q \neq p}^{132} G_c(p, q)}{131} \quad (6)$$

where $M_p(x, y, z)$ is a binary mask obtained from the HOA + AAL atlas to define each parcel p . A diagram summarizing the XAI steps for 3D T1-weighted volumes is reported in the bottom part of Fig. 1, whereas for structural connectivity is reported in the yellow block of the Fig. 2.

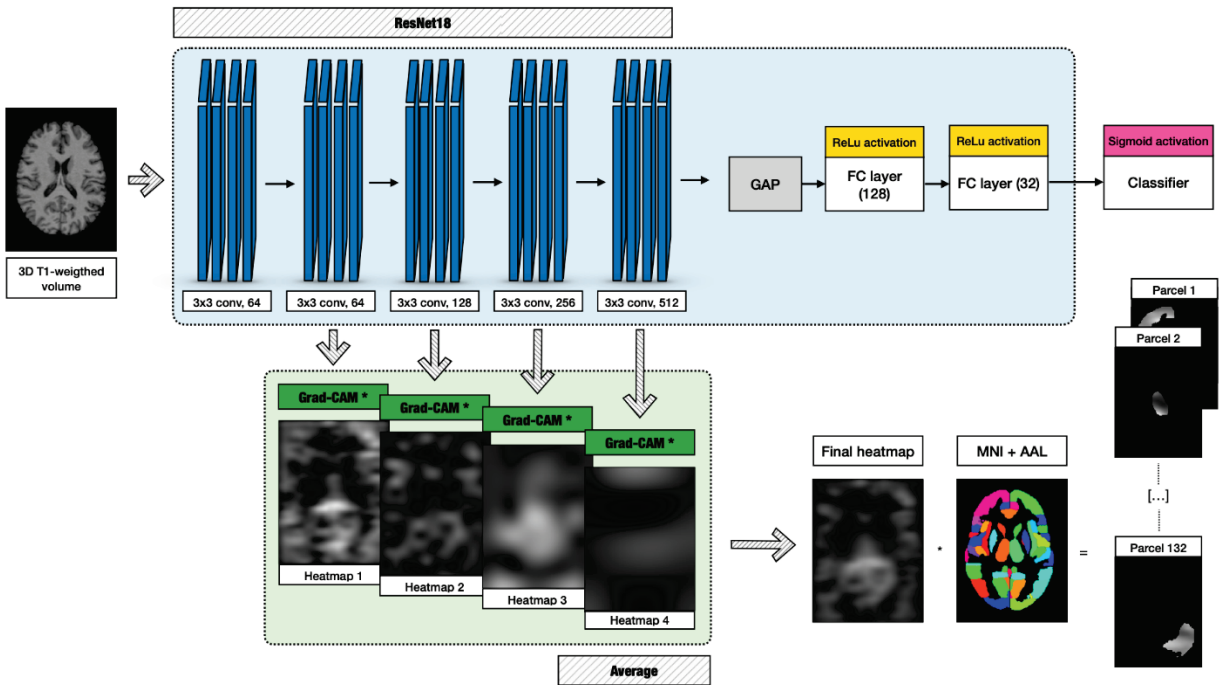


Figure 1. Implemented workflow for the AD classification and explainability assessment performed on ResNet18. The architecture of the implemented model (blue panel) comprised 5 convolutional layers, a Global Average Pooling (GAP) layer, three Fully Connected (FC) dense layers with Rectified Linear Unit (ReLU) activation and a sigmoid activation function for the binary classification. The outputs derived from the last 4 convolutional layers were processed using Grad-CAM (green panel) and then averaged. The final heatmap was then multiplied for the binary masks underlying the HOA + AAL atlas parcels.

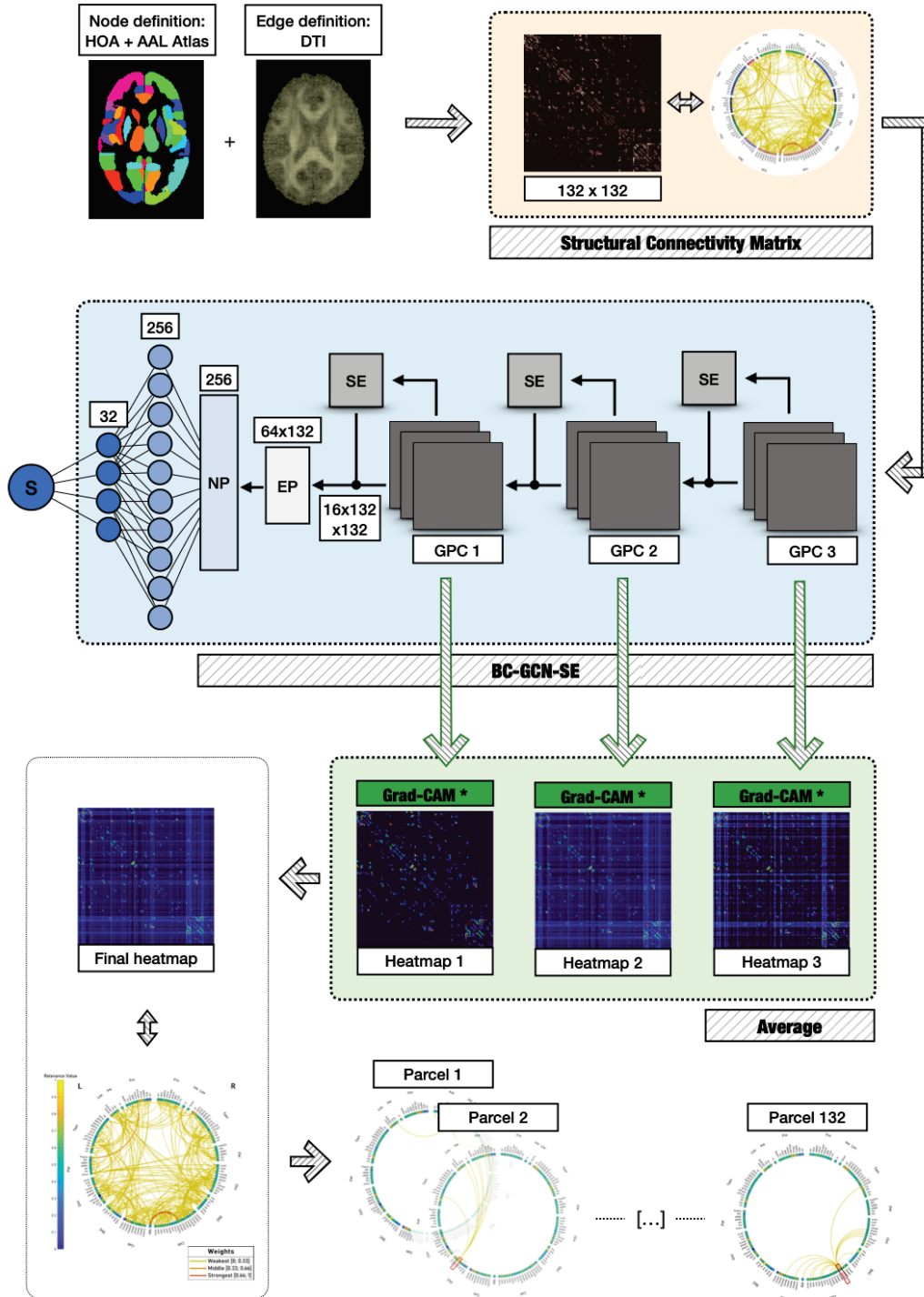


Figure 2. Implemented workflow for the connectivity data extraction, AD classification and explainability assessment performed on BC-GCN-SE. The processing steps used to derive the structural connectivity data are displayed in the orange panel. The architecture of the implemented model (blue panel) comprised three Graph Path Convolution layers (GPC), an Edge Pooling (EP) layer, a Node Pooling (NP) layers and two fully connected layers. The outputs derived from the 3 convolutional layers of the model were processed using Grad-CAM (green panel) and then averaged. From the final heatmap the contributions highlighted by each connection of each node of the HOA + AAL atlas were averaged to extract the RV measure of each parcel. A

connectogram representing both connectivity edges of the heatmap and color-coded RV values (circle perimeter) is displayed. These plots were created using the SPIDER-NET tool [53].

2.7. Statistical Evaluation

2.7.1 Anatomical targets

After evaluating the DL models (see Sec. 2.3 for further details) results of the XAI assessment were analyzed in comparison to domain knowledge.

As for the BC-GCN-SE model, we selected as structural connectivity target a hallmark of AD: the DMN [7], [10], [11]. The DMN was defined by 31 parcels of the HOA according to the indications provided in the “Cortical & White Matter Parcellation Manual” (<https://cma.mgh.harvard.edu/>) and to different previous studies [54]–[57]. Specifically, we defined the DMN according to previous studies highlighting the regions characterizing the network. The medial prefrontal cortex, the posterior cingulate cortex, a portion of the temporal lobe, the precuneus, and the inferior parietal areas like the angular and supramarginal gyri often make up these brain regions. In addition to these usual core regions, the network frequently include the lateral temporal cortex, hippocampus, and amygdala [54], [55], [58]. In this regard, no unique definition exists. While prior research offers preliminary indications of a more comprehensive characterization of the DMN system, additional studies are required to define the anatomical scope of certain parcel contribution [58].

As for the ResNet18 model, a different cluster of parcels was selected as anatomical target. According to previous studies [59], [60], patterns of atrophy in the medial temporal lobe (MTL) represent a well-established sMRI biomarker for AD and is often used a diagnostic criterion for individuals displaying early symptoms. Thus, while conducting the XAI assessment, 4 structures were considered bilaterally, for a total of 8 regions of interest: hippocampus, anterior and posterior parahippocampal gyri and amygdala.

2.7.2 Statistical tests

The investigation of the processes underlying the decision of the classifiers was initially carried out by performing a statistical test to identify the different contributions to AD/HC classification of the parcels taken from the HOA + AAL atlas for both models. To predispose the statistical comparison, some operations were carried out. First, we obtained two sets from the RV associated to each parcel of each subject. These sets were created by: (i) removing the misclassified sessions; (ii) aggregating the RV values across single parcels, for subjects with multiple MRI sessions and same class, (iii) Removing the sessions associated to the same subjects but with different classes (removing records); (iv) separating the HC and AD sessions; (v) normalizing between 0 and 1 using the minimum and the maximum values among all RV values obtained from the corresponding models; for example, all RV values obtained with ResNet18 were normalized according to the maximum and minimum obtained from all RV related to each parcel and each correctly classified subject.

The resulting sets were independent and of different size according to the class and the number of sessions misclassified by ResNet18 and BC-GCN-SE.

In order to evaluate the regions and inherent characteristics that were particularly used among all parcels for the classification task of AD, two analyses were performed. First, the Kolmogorov-Smirnov test was performed on all the parcel samples to evaluate normality. The two sets were compared between each other through the Mann-Whitney or independent samples t-test using the Benjamini-Yekutieli correction in both cases [61]. The significance level was set to 0.05 and the correction was applied on the result to adjust p-values accounting for multiple comparisons. In this way, it is possible to obtain the parcels that were used differently by the algorithms to classify one or the other class in relation to all parcels.

Second, most relevant parcels for the classification of AD and HC were investigated separately. These subsets were obtained according to the 15th percentile of the highest RV for both classes (20 parcels out of the 132 total HOA + AAL parcels) and both models were analyzed. Indeed, these most relevant parcels are the most employed by the models for the classification.

From these subsets (i.e., most relevant parcels), we analyzed the following subgroups: i) parcels that were also statistically significant. These were strongly influencing the model while also providing a greater contribution towards one of the two classes; ii) parcels that were not statistically significant. Of these, particular attention was paid to the ones in common between the AD and HC groups as they were strongly influencing the classification of both classes without leveraging one.

3. RESULTS

The results of the 10-fold cross-validation from the DL models ResNet18 using 3D T1-weighted volumes and BC-GCN-SE using structural connectivity data are summarized in Table 1.

Table 1. Normalized confusion matrix for the ResNet18 and BC-GCN-SE model performance, reported as median value and interquartile range.

| Cross-Validation Normalized Confusion Matrix | | Predicted | |
|--|---|----------------------|----------------------|
| ResNet18 | | P | N |
| Actual | P | 0.817 [0.773, 0.846] | 0.183 [0.154, 0.227] |
| | N | 0.183 [0.167, 0.233] | 0.816 [0.767, 0.833] |
| BC-GCN-SE | | P | N |
| Actual | P | 0.703 [0.672, 0.769] | 0.297 [0.231, 0.328] |
| | N | 0.261 [0.242, 0.302] | 0.739 [0.698, 0.758] |

Both models achieved good performance. Specifically, ResNet18 achieved balanced performance for AD and HC during cross-validation ($TPR_{median} = 0.817$; $TPR_{IQR} = 0.073$; $TNR_{median} = 0.816$; $TNR_{IQR} = 0.066$). The median accuracy was 0.811 and the interquartile range was 0.073. On the new split, the performance of ResNet18, obtained merging both the training and validation sets, were $TPR = 0.985$ and $TNR = 0.989$, with an accuracy of 0.988.

BC-GCN-SE provided slightly inferior performance with respect to ResNet18 ($TPR_{median} = 0.703$; $TNR_{median} = 0.738$). More precisely, the model showed slightly higher results and also less variability for HC during cross-validation ($TPR_{IQR} = 0.097$; $TNR_{IQR} = 0.060$). The median accuracy is 0.742 and the interquartile range is 0.058. Finally, the results of BC-GCN-SE on training and validation sets on the new split highlighted $TPR = 0.956$ and $TNR = 0.788$, with a total accuracy of 0.821.

Table 2. Brain parcels displaying a significant difference between the AD and HC level of relevance.

| ResNet18 (3D T1-weighted volumes) | | | | | | | | | BC-GCN-SE (Structural connectivity) | | | | | |
|-----------------------------------|---------|------------|----|--------|------------|----|----------|------------|-------------------------------------|---------|------------|----|--------|------------|
| N. | Parcel* | Adjusted p | N. | Parcel | Adjusted p | N. | Parcel | Adjusted p | N. | Parcel* | Adjusted p | N. | Parcel | Adjusted p |
| 1 | AG_r | < 0.001 | 23 | AC | < 0.001 | 47 | Pa_l | 0.006 | 1 | AG_r | 0.025 | 23 | NAcc_r | 0.015 |
| 2 | aITG_r | < 0.001 | 24 | Amg_l | < 0.001 | 48 | PostCG_l | 0.003 | 2 | aITG_r | < 0.001 | 24 | aMTG_l | < 0.001 |
| 3 | Amg_r | < 0.001 | 25 | CaN_l | < 0.001 | 49 | PrCun | < 0.001 | 3 | Amg_r | 0.017 | 25 | aSMG_r | 0.031 |

| | | | | | | | | | | | | | | |
|----|-----------|---------|----|------------|---------|----|----------|---------|----|-----------|---------|----|----------|---------|
| 4 | aMTG_r | 0.002 | 26 | Cereb1_l | < 0.001 | 50 | PreCG_r | 0.024 | 4 | aMTG_r | < 0.001 | 26 | aSTG_l | 0.038 |
| 5 | aTFusC_r | < 0.001 | 27 | Cereb6_l | < 0.001 | 51 | SFG_r | 0.010 | 5 | aTFusC_r | 0.014 | 27 | Cereb1_r | 0.028 |
| 6 | BSt | 0.024 | 28 | Cereb6_r | 0.001 | 52 | SubCalC | < 0.001 | 6 | BSt | < 0.001 | 28 | HG_r | 0.016 |
| 7 | Cereb2_r | < 0.001 | 29 | Cereb7_l | < 0.001 | 53 | TOFusC_l | < 0.001 | 7 | Cereb2_r | 0.034 | 29 | ICC_r | 0.031 |
| 8 | Cuneal_r | < 0.001 | 30 | Cereb8_l | < 0.001 | 54 | Tha_r | < 0.001 | 8 | Cuneal_r | 0.016 | 30 | iLOC_r | < 0.001 |
| 9 | FO_r | < 0.001 | 31 | Cereb8_r | < 0.001 | 55 | Ver10 | < 0.001 | 9 | FO_r | 0.001 | 31 | LG_r | 0.044 |
| 10 | FP_r | < 0.001 | 32 | Cereb9_l | < 0.001 | 56 | Ver45 | 0.005 | 10 | FP_r | 0.018 | 32 | OFusG_r | 0.018 |
| 11 | IFG_tri_l | < 0.001 | 33 | Cereb9_r | < 0.001 | 57 | Ver8 | < 0.001 | 11 | IFG_tri_l | 0.018 | 33 | PaI_r | 0.028 |
| 12 | IFG_tri_r | < 0.001 | 34 | Cuneal_l | < 0.001 | 58 | aITG_l | < 0.001 | 12 | IFG_tri_r | 0.044 | 34 | pMTG_r | 0.018 |
| 13 | MedFC | < 0.001 | 35 | FO_l | < 0.001 | 59 | aPaHC_l | < 0.001 | 13 | MedFC | 0.006 | 35 | PO_l | 0.034 |
| 14 | PaCiG_l | < 0.001 | 36 | FP_l | < 0.001 | 60 | aPaHC_r | < 0.001 | 14 | PaCiG_l | < 0.001 | 36 | PO_r | 0.021 |
| 15 | PaCiG_r | < 0.001 | 37 | Hip_l | < 0.001 | 61 | aTFusC_l | < 0.001 | 15 | PaCiG_r | 0.020 | 37 | PostCG_r | 0.010 |
| 16 | PreCG_l | < 0.001 | 38 | Hip_r | < 0.001 | 62 | pITG_l | 0.006 | 16 | PreCG_l | 0.024 | 38 | pSMG_l | 0.032 |
| 17 | SFG_l | < 0.001 | 39 | IC_r | < 0.001 | 63 | pITG_r | < 0.001 | 17 | SFG_l | < 0.001 | 39 | pSMG_r | < 0.001 |
| 18 | Tha_l | < 0.001 | 40 | IFG_oper_l | < 0.001 | 64 | pMTG_l | 0.011 | 18 | Tha_l | 0.031 | 40 | pSTG_r | 0.028 |
| 19 | toMTG_r | 0.045 | 41 | IFG_oper_r | < 0.001 | 65 | pPaHC_r | 0.009 | 19 | toMTG_r | 0.044 | 41 | PT_r | 0.019 |
| 20 | Ver9_r | < 0.001 | 42 | MidFG_l | < 0.001 | 66 | pSTG_l | 0.010 | 20 | Ver9 | 0.024 | 42 | SMA_l | 0.018 |
| 21 | NAcc_l | < 0.001 | 43 | MidFG_r | < 0.001 | 67 | pTFusC_r | < 0.001 | 21 | NAcc_l | 0.003 | 43 | SPL_r | 0.018 |
| 22 | CaN_r | < 0.001 | 44 | OFusG_l | 0.026 | 68 | sLOC_l | < 0.001 | 22 | CaN_r | 0.007 | 44 | toITG_r | < 0.001 |
| | | | 45 | OP_l | < 0.001 | 69 | sLOC_r | < 0.001 | | | | 45 | Ver12 | 0.044 |
| | | | 46 | OP_r | < 0.001 | 70 | toITG_l | < 0.001 | | | | 46 | Ver7 | 0.018 |

Results relative to the Mann-Whitney or independent samples *t*-tests are reported in terms of *p*-value, after performing the Benjamini-Yekutieli correction. The first column of the ResNet18 and BC-GCN-SE model (from N.1 to N.22) is marked with an asterisk (i.e., Parcel*), to highlight the significant parcels common to both.

The results from the statistical tests on the parcel XAI heatmaps are summarized in Table 2. Among all the parcels of the atlas used, the relevance values of 70 and 46 parcels of AD in comparison to HC for ResNet18 and BC-GCN-SE respectively were found to be statistically significant ($p < 0.05$). In particular, 22 parcels were in agreement between the two methods (first column of Table 2, highlighted with asterisk), mainly belonging to cortical parcels (except two cerebellar parcels). More specifically, in the case of ResNet18, 7 out of 8 total target MTL and hippocampus parcels, with the only exception of the anterior division of the left parahippocampal gyrus were found to be significant. In the case of BC-GCN-SE, instead, all the parcels belonging to target DMN except the frontal orbital cortex, the posterior part of the cingulate gyrus, the temporal pole, the precuneus and the hippocampus (12 out of 17 parcels) were found to be statistically significant ($p < 0.05$) in at least one hemisphere (70.59%). Considering lateralization, the number was 16 out of 31 total DMN parcels (51.61%). The distributions of the relevance values from both models of these target parcels are shown in Fig. 3 and Fig. 4.

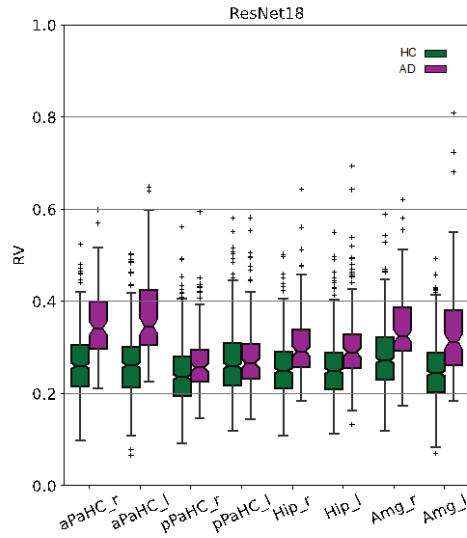


Figure 3. Boxplots of the RV values relative to the target parcels (specified on the x axis) for both the BC-GCN-SE model. Results for the AD (in purple) and HC (in green) subjects are resented using different colors.

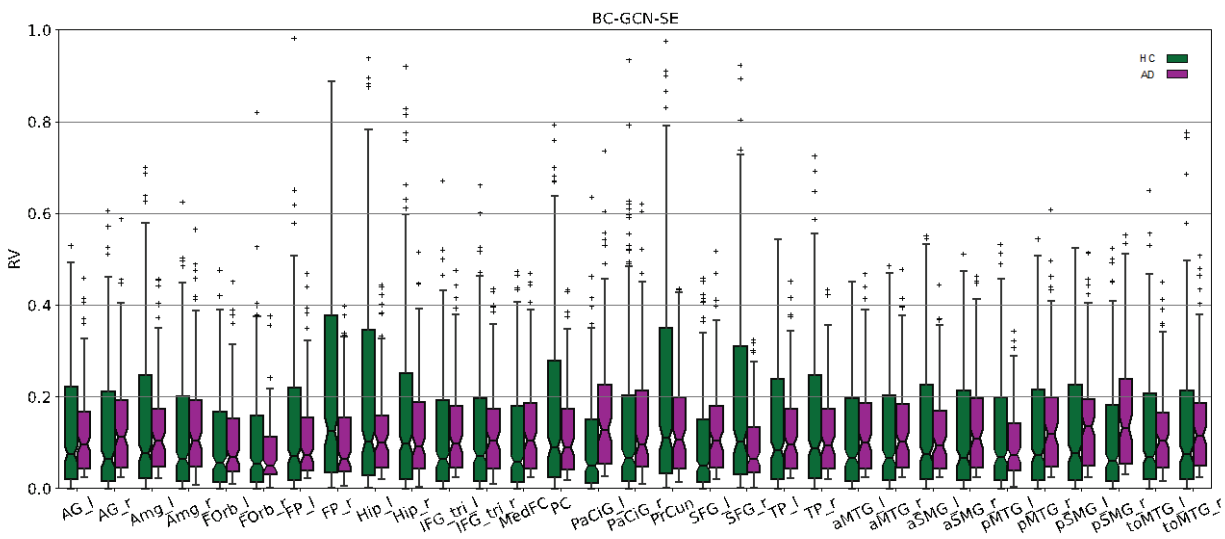


Figure 4. Boxplots of the RV values relative to the target parcels (specified on the x axis) for both the BC-GCN-SE model. Results for the AD (in purple) and HC (in green) subjects are resented using different colors.

It is worth noting that most of the regions, with the exceptions of *pPaHC_l* for the ResNet18 model and *FP_r* for the BC-GCN-SE model, resulted to be significantly different with greater relevance values for the AD case.

Afterwards, the most relevant parcels for the classification of AD from both models were analyzed. First, it was noted that 11 out the 20 most relevant parcels for ResNet18 (*Ver10*, *aTFusC_l*, *aTFusC_r*, *aITG_r*, *SubCalc*, *aITG_l*, *aPaHC_l*, *pITG_r*, *MedFC*, *pSTG_l*, *aPaHC_r*), were also found to be statistically significant. The anterior division of both right and left parahippocampal gyrus from target MTL were among these 11 parcels. The remaining 9 most relevant parcels that were not statistically significant were *Cereb10_l*, *Cereb10_r*, *Forb_l*, *Forb_r*, *CO_r*, *pSTG_r*, *PO*, *Ver12*, and *Ver3* whose first 7 were also found to be among the most relevant for the HC classification.

Second, 14 out of the 20 most relevant parcels for BC-GCN-SE (*Cereb1_r*, *iLOC_r*, *pMTG_r*, *pSMG_l*, *pSMG_r*, *pSTG_r*, *toITG_r*, *ICC_r*, *OFusG_r*, *PaCiG_l*, *PaCiG_r*, *SPL_r*, *Tha_l*, *Ver9*), were also found to be statistically different. Of these 14, 5 parcels belong to target DMN. The remaining 6 most relevant parcels that were not statistically significant were *Cereb8_r*, *Cereb9_l*, *Cereb9_r*, *PreCG_r*, *Ver10*, *Cereb10_l*, whose first 4 were also found to be among the most relevant for the classification of HC.

To summarize, a diagram indicating the mean RV of all correctly classified sessions of both classes and models per each parcel and divided in lobes are shown in Fig. 5.

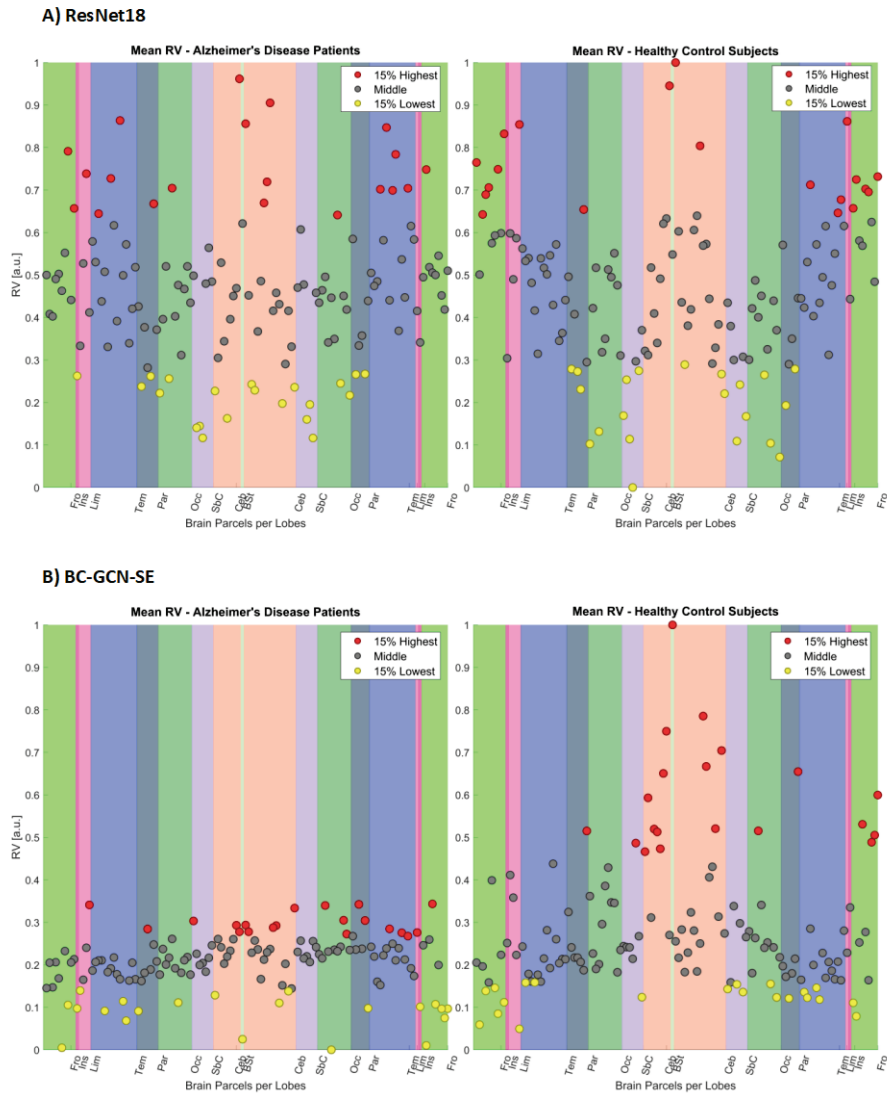


Figure 5. Mean RV of all parcels, for the ResNet18 (panel A) and BC-GCN-SE (panel B) models. The AD and HC classes are reported separately. The lobe's belonging is indicated through rectangle colors. The rectangles have different size according to the number of parcels contained in the lobes. The RV of each parcel is labeled by dot colors according to the following criterion: red indicates the 15% of parcels characterized by highest RV, yellow indicates the 15% of parcels characterized by lowest RV, grey indicates the remaining parcels. Plots were created using the SPIDER-NET tool [53]. Fro: Frontal Lobe; Ins: Insular Cortex; Lim: Limbic Lobe; Tem: Temporal Lobe; Par: Parietal Lobe; Occ: Occipital Lobe; SbC: Subcortical Structures; Ceb: Cerebellum; BS: Brainstem.

4. DISCUSSION

In this work, we compared two AI methods trained on different types of sMRI data, namely 3D T1-weighted scans and structural connectivity matrixes extracted from dMRI data. Both approaches were applied to an AD classification task and evaluated first in terms of model accuracy and then in terms of explainability. The population considered was a subset of the OASIS-3 dataset. In particular, we focused on 543 subjects comprising both healthy controls and individuals at various stages of AD cognitive decline. Since a group of subjects underwent multiple imaging sessions that we treated as separate records, the final dataset consisted of 692 MRI volumes and structural connectivity matrixes, with an HC:AD ratio of 557:135.

4.1 DL model performance

Overall, the DL models highlighted good classification results. On the one hand, the BC-GCN-SE obtained acceptable performance with respect to the existing literature [28]. Higher accuracy ($TNR_{\text{median}} = 0.738$) and slightly lower cross-validation variability ($TNR_{\text{IQR}} = 0.060$) was found for the HC classification as compared to the AD case ($TPR_{\text{median}} = 0.703$; $TPR_{\text{IQR}} = 0.097$). On the other hand, the ResNet18 model achieved slightly superior accuracies with respect to the BC-GCN-SE when classifying both AD and HC subjects ($TPR_{\text{median}} = 0.817$; $TNR_{\text{median}} = 0.816$). The obtained results were in line with the existing literature [36]. The AD and HC variabilities seemed comparable ($TPR_{\text{IQR}} = 0.073$; $TNR_{\text{IQR}} = 0.066$). To the best of our knowledge, no other studies compared DL models trained using 3D T1-weighted volumes and structural connectivity data neither in terms of accuracy nor explainability. Moreover, this is the first work testing BC-GCN-SE model on structural connectivity with the aim of classifying diverse AD subjects of the OASIS-3 dataset. Few other AD classification studies were conducted on alternative populations, such as the well-known ADNI dataset [28], [30], highlighting good performance for the DL models adapted to graph-structured data - mostly functional connectivity. However, these models are relatively recent, especially when compared to more mature CNNs architectures working on images such as the ResNet18 model. This latter is a widely tested pre-trained model with weights obtained from more than a million images from the ImageNet database, probably resulting in a better ability to generalize.

In addition, it is worth noting that the structural connectivity data obtained from DTI have inherent limitations related to the processing pipelines that may result in noisy connections. The absence of gold-standard methods for the creation of connectivity graphs and for edge weighting is indeed one of connectomics' current biggest shortcomings [62], [63]. DTI fiber tracking algorithms suffer, for example, from the assumption of a single fiber orientation per voxel, resulting in systematic errors for complex fiber geometries (such as crossing, kissing, twisting fibers etc.). In addition, other issues can strongly affect the final extracted data since they are directly dependent on the parameters of the dMRI experiment, and the accuracy can decrease with factors such as pathway length, shape and size of the reference region, and shape of the tract in question [62]. Usually, thresholding methods are widely applied on the raw connectivity matrixes to remove spurious edges, although the optimal approach is yet to be found. This results in altered brain graph structure and connectedness. Future studies could fruitfully explore this issue further by applying some of these most recent methods to assess the performance after connectivity data cleaning [64].

Given this premises, both results can be considered promising, even considering the great variability of severities of the AD subjects, sessions parameters, the presence of multiple acquisition sessions without a predefined design setting (such as scheduled acquisitions) in the OASIS-3 dataset. Moreover, the division into classes is made according to the CDR values, also resulting in 15 subjects having sessions with different labels and in a majority of sessions from subjects with very mild impairment (n. 97 sessions with CDR 0.5 out of 135 total AD sessions). Clearly, being the range of dementia effects wide and of difficult definition, it results in a more difficult task.

As said, the use of DL models on brain connectivity graphs is quite recent, thus assessing them in comparison to more established approaches would be of great importance. The processing effort of extracting connectivity data,

providing new information and relevant data would be thus enhanced if casting new light on the exploration of brain biomarkers and computational cost in parallel to optimal results of AI. Future research should further develop and confirm these first findings, aiming to improve the performance since demonstrated higher potential. This would be of great relevance in context of the definition of prodromal symptoms of the pathologies, diagnosis and rehabilitation.

4.2 ResNet18 explainability results

After using the DL models to classify our data, this work focused on their explainability. The aim was to validate the functioning of AI methodologies employing MRI volumes and structural connectivity matrixes and to assess their agreement with respect to domain knowledge. As we applied Grad-CAM on both classification models, results were evaluated across 132 parcels derived from the HOA + AAL brain atlas. At first, we extracted the levels of relevance characterizing each parcel in both the AD and HC subjects and compared results, to determine if the regional information was used primarily to identify either one of the two classes. Then, we investigated the most relevant parcels for the AD classification task according to a ranking criterion applied on the RV measures.

As for the imaging classification task, we saw a significant difference between AD and HC in 70 brain areas. Among these, particular attention was paid to the regions deemed as more relevant to identify neurodegenerative processes from sMRI data, as indicated by the existing literature. Numerous studies highlighted the involvement of the MTL in the pathogenesis of AD, and pointed at its volumetric loss as an early sign of the disease progression [2], [65], [66]. Additionally, the medial temporal atrophy has been indicated among the diagnostic criteria for AD by a revised version of the traditional National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer’s Disease and related Disorders association criteria [60]. This version removed the original requirement for dementia onset to classify subjects as AD, thus making the MTL a reliable biomarker to detect the disease presence even before the disability phase occurs (i.e., dementia onset and progression). As mentioned in the Methods section (see Sec. 2.7.1 for further details), the brain structures involved in the MTL are the hippocampus, amygdala and parahippocampal regions, that are all key for the episodic and spatial memory [59]. However, among these, the most validated and established sMRI biomarker for AD is the hippocampus. A number of studies have linked its volumetric loss to the memory decline stages [2], [67], [68]. Additionally, Hall and colleagues proved that the amount of time required by cognitively normal subjects to develop dementia was shorter in the presence of hippocampal atrophy [68]. Based on these considerations we assessed the difference between the amount of repeatability characterising AD and HC subjects in the following parcels: *Hip_r*, *Hip_l*, *Amg_r*, *Amg_l*, *aPaHC_r*, *aPaHC_l*, *pPaHC_r*, *pPaHC_l*. The presence of a significant difference between the AD and HC groups in 7 regions – all except the pPaHC – suggests that the algorithm is leveraging them to perform the classification. Additionally, the higher levels of relevance of the AD class as compared to HC in all 7 structures (see Fig. 3), indicates that the algorithm is mainly utilizing them to identify the presence of AD, rather than a physiological condition. From this we may infer a good level of explainability for the implemented ResNet18 model, that seems to be positively influenced by almost all the anatomical regions that are usually involved in the AD progression process. However, one may argue that the number of parcels displaying a difference in the AD and HC levels of RV is high when compared to the total number of parcels we investigated. This may increase the possibility to detect the target structures belonging to the MTL. To overcome this limitation, we investigated which parcels – among the significant ones were also included in the 20 most relevant (i.e., parcels with the highest RV values) for the AD case. We identified a total of 11 parcels, 2 of which belonged to the MTL: *aPaHC_r*, *aPaHC_l*. This serves as a strong confirmation that the algorithm is leveraging them to identify the presence of AD. Additionally, it is noticeable that almost all the remaining regions – the only exception being represented by Vermis10, a parcel of the human cerebellum – are part of the brain cerebral cortex, that is known to be involved in the process of disease progression [66]. In particular, 6 parcels belonged to the temporal lobe (see Fig. 5), whose atrophy pattern

has been linked to neuronal loss, visuospatial, language and behavioral impairment by a number of studies [60], [69], [70]. Furthermore, the presence of *MedFC* and *SubCalC* – belonging to the frontal and limbic cortex, respectively – among the most relevant features could be explained by the fact that, during the latest stages of the disease, atrophy spreads to the remaining cortical areas, sparing only the visual and sensory motor cortex, as indicated by Eskildsen and colleagues [2]. Overall, the obtained results indicate that the designed classification approach is correctly using the structural information provided by the 3D T1-weighted images to recognize diseased individuals. Even though having all of the 8 target regions among the most relevant ones would have proven the achievement of higher levels of explainability, the outlined findings seem promising and open the way for further investigations on the processes underlying the AD classification. Additionally, it is worth noting that the 11 parcels discussed so far – besides aiding the classification – seem also able to avoid the introduction of potential biases or confounding factors. Indeed, only 2 out of them were among the most explainable features for both AD and HC, while the remaining ones were favoring the sole pathological class, thus proving their role as an sMRI biomarker for AD. However, in this regard, it is important to mention that 7 among the remaining AD most explainable features had very high and comparable RV measures with respect to the HC class. Among these, we found 2 cerebellar and 5 cortical regions (either frontal, temporal or parietal). As for the cortical regions we may speculate that the DL model is using them to extract opposite information with respect to the ones generally used to classify AD. The dichotomy between atrophic and physiological cortical volume could play a significant role in the AD/HC distinction. This consideration could still hold for the cerebellar parcels, since the volumetric loss of their molecular and granular layers, despite not being among the most established AD biomarkers, have been linked to the pathology presence by different studies [71], [72]. However, further investigation is certainly needed to confirm the outlined hypothesis or, more generally, to shed light on how these specific regions are used by the DL model to discriminate between healthy and pathological individuals.

4.3 BC-GCN-SE explainability results

Regarding the relevance of structural connectivity data and its use within BC-GCN-SE model to classify AD and HC, the DMN was used as target. Indeed, according to different studies [7], [11], [12], changes in the DMN are well-known markers, since age and Alzheimer's pathology can alter the WM and disrupt the DMN normal functioning. In a DTI study was also found that core parcels of the DMN, such as cingulate structure and hippocampus, are even increasingly gaining attention in the AD field since they are also recognized as key structure of the memory system, confirming sMRI findings [7]. We assessed the relevance of these target regions in the task through BC-GCN-SE model (see Sec. 3 for further details), as previously done for the ResNet18 model.

The GRAD-CAM results assessment highlighted 46 parcels which differently contribute to the classification of the two classes with respect to all parcels. Among these, it was found that more than the half of the total parcels belonging to the DMN were statistically significant ($p < 0.05$) with majority of values in favor to AD. It is also worth noting that an even higher percentage of the target parcels (more than 70%) were relevantly used by the model if not considering lateralization. However, we acknowledge that there are considerable discussions and increasing evidence among researchers as to abnormality of topological asymmetry between hemispheric brain WM in AD and Mild Cognitive Impairment [73], [74]. More specifically, when comparing our results to those of other studies, it must be pointed out that all parcels which were found to exhibit asymmetry in the study by Yang and colleagues agreed to the parcels of only one hemisphere used by our model [73]. Examples were angular gyrus and amygdala, belonging to DMN, or inferior temporal gyrus, cuneal cortex, supplementary motor area and lingual gyrus. In addition, several of these parcels were also found to be among the most relevant for the identification of AD. This can be a valuable indication supporting for hemispheric lateralization and aberration possibly due to the long-range connection loss that would be of great interest to be further investigated.

Also, the relevance of the changes within DMN was confirmed by analyzing the statistically significant and at the same time most explainable parcels for AD class. Indeed, among these 14 parcels, 5 belonged to DMN. However, whether on the one hand the results appear to provide an important involvement of the DMN, on the other hand, important parcels such as hippocampus, posterior cingulate cortex or precuneus, were not accounted by the BC-GCN-SE. Beyond that the case of these missing important DMN parcels, also the MTL and, hippocampus included, appeared to exert a small influence on the classification process. Apart from the right amygdala that was found to be statistically significant, other regions such as parahippocampal cortex appeared to not be relevantly used ($p \geq 0.05$). This may represent a limitation in terms of explainability of the model, since some of these regions are also often included in the DMN [7] and MTL was found to be relevant in dMRI studies [14], although less replicated than DMN core areas [75]. At the same time, it is possible to notice from figure Fig. 5 a wide presence of most explainable relevant parcels focused on the right temporal and parietal lobes. In this regard, other studies highlighted general temporal and parietal lobe disruptions contributing to memory impairment [7], [10], [12]. Other particular cases were related to brainstem, the nucleus accumbens and the cerebellum. These regions were not considered as targets, but their analysis appears to be of interest. Indeed, the brainstem was the parcel mostly differing between the two cases in favour to HC, with the highest RV as shown in figure Fig. 5. At this stage of understanding, we believe that this result may highlight an important influence of the corticospinal tract in healthy subject [76]. In addition, the brainstem was found to be associated to apolipoprotein status, resulting in an altered radial diffusivity [77], [78]. Alteration in AD was also found in a study by Nie and colleagues in the accumbens nucleus, that resulted to be statistically different in AD/HC classification made by BC-GCN-SE in both hemispheres [79]. Moreover, analyzing the most explainable relevant parcels that were in common between the two classes it was found that many cerebellum areas aided the model to recognize both classes. This finding can support the hypothesis of a strong involvement of this area in AD that was only recently reported by some studies [80]. Alternatively, studies of whole-brain graph organization revealed subregions of the cerebellum connected with the cortical regions of the DMN [58]. Although the important role of the cerebellum that could be investigated in future works, it is worth noting that the uncertainty of the connectivity data can have a great influence of all these results, especially if referring to cerebellar parts that can be massively affected by noise [81]. In general, these unexpected parcels and less replicated results need further investigation which may potentially bring to new findings regarding the involvement of certain areas in AD or to highlight limitation of the model both in terms of interpretability and accuracy of the results. It is indeed paramount to focus on the interpretation of graph-structured based DL models to provide new evidence of their reliability and trust, and to improve their performance through possible less arbitrary data cleaning and thresholding.

4.4 Limitations and perspectives

The present study is not without limitations. For example, different approaches for XAI exist and they may not always be effective in all scenarios [32]. The use and comparison to other methods such as perturbation or distillate methods would be of great interest to further validate the results. In addition, this work only focuses on the mean relevance value extracted from a whole parcel to allow straightforward comparison to 3D T1-weighted volumes. It would also be appealing, in the case of structural connectivity data, to investigate most explainable connections which could highlight most important long-range connections and their loss in AD [8], [9]. In addition, beyond that the DMN analysis, the investigation of other Resting State Networks (such as Frontoparietal network in relation to fronto-temporal dementia etc.) may highlight new insights on the employed models. Further investigations on the WM might also be performed in the 3D T1-weighted volumes to visualize similarities and differences with respect to WM metrics related to structural connectivity data. A particular case is related to the presence of White Matter Hyperintensities which can represent important biomarkers of AD condition [82]. In a previous preliminary study, we confirmed their relevance within a DL model employing FLAIR images, although

with a low number of subjects [23]. A possible perspective would be to further investigate their effects in regard to connectivity data as well. For example, it would be of great interest to employ tools such as NeMo to extract the effect of these lesions within connectivity data structure [83].

4.5 Comparing and combining explanations

The agreement, peculiarities and limitations of the two DL models employing sMRI and dMRI data emerge as key findings. First, 22 regions were found to be significantly used by both models to distinguish the pathological condition. Furthermore, among all the other parcels that did not agree, it was still noticed a large agreement on specific regions without considering hemisphere's belonging. Some other parcels were found to be important, both including the target ones and not. Indeed, although adherence to known markers from 3D T1-weighted volumes and structural connectivity data are given, other unexpected parcels, mainly from the cortex, resulted in being important, reducing the strength of the explanation. It is worth also noting that the population was composed of subjects at various AD stages, severities, and affections that might simply suggest that different involvement of the disease in different MRI sessions lead to general extensive importance of a considerable portion of the cortex. This is particularly true for the ResNet18 model, that points out more statistically significant parcels. In addition, it is a deeper model with respect to BC-GCN-SE, possibly focusing on more complex features. Together with the comparison between AD and HC most relevant regions were investigated. Among them, the two DL models employing different data highlighted a notable agreement to target. First, it was found that some DMN regions have good importance in the classification from structural connectivity matrixes. Second, it was also shown as the regions characterizing the MTL, found as replicated hallmark in sMRI studies, contain inherent important features for the ResNet18 model employing 3D T1-weighted volumes. In this context, it is also of great interest to underline how some important regions of the MTL such as amygdala and parahippocampal gyri, and hippocampus that were also found to be different in AD connectivity data [14], were not underlined by the XAI analysis of BC-GCN-SE. Indeed, apart from right amygdala, that was found to be statistically significant, all the other regions of the MTL did not result from either statistical test and examining most relevant parcels. This finding may highlight a limitation of BC-GCN-SE model interpretability or the effect of some noise sources inherently present in the connectivity data. At the same time, this complementary relevance of parcels of interest appears promising in the perspective of developing superior and more trustworthy models. Indeed, the use of well-known hallmarks from multiple measures may offer the opportunity of focusing on different information that would be of great interest and significance if used concurrently. In this context, only few studies considered the combination of morphological features of regions from 3D T1-weighted volumes and interregional properties obtained through structural connectivity data, but that may potentially lead to more accurate results and to a better interpretation [84]. Nowadays, there is indeed the potential to easily collect multiple data from multiple modalities, and in this sense, efforts to assure even greater use of the whole potentiality of DL models, exploiting their peculiarities would be of unvaluable interest for diagnosis and rehabilitation of the pathology. The development of better and more interpretable models can represent accurate and robust solutions to the well-known problem of trust in "black-box models", which limited their diffusion within real settings so far.

5. CONCLUSION

In this work, we assessed two DL models working on data from multiple MRI acquisitions performed on a subset of AD and HC subjects from the OASIS-3 dataset. Specifically, we employed 3D T1-weighted sMRI volumes (ResNet18) and structural connectivity matrixes (BC-GCN-SE) defined according to the HOA + AAL atlas and to dMRI metrics. In this perspective, we compared the models according to their accuracy and explainability. The method employed was GRAD-CAM, which pointed out some target regions, found to highlight markers of the pathology in sMRI and dMRI. More precisely, the expected involvement of the MTL was found using ResNet18, whereas the DMN

relevance was found to be important in the decision made by the BC-GCN-SE. Even though an important involvement of these targets in the classification decision, part of these expected regions was missing in the analysis of one of the two models, highlighting complementary explanations. For example, important areas such as hippocampus or parahippocampal gyri were identified by the ResNet18, whereas excluded by BC-GCN-SE. This work emphasized the potential held by imaging and connectivity data as for the creation of better and more reliable models. The opportunity to focus on different information may be provided by well-known hallmarks from multiple measures, if employed concurrently. In this context, combining interregional properties found by brain connectivity graphs with the morphological characteristics of regions from 3D T1-weighted scans would result in more accurate findings and better interpretation. In summary, the comparison of models using multiple data points highlights the strengths of the two modalities that might help in the creation of more understandable models and, in the long-term, that may result in a rise of the confidence and trust laid in AI models.

BIBLIOGRAPHY

- [1] Z. Cui, Z. Gao, J. Leng, T. Zhang, P. Quan, and W. Zhao, "Alzheimer's Disease Diagnosis Using Enhanced Inception Network Based on Brain Magnetic Resonance Image," in *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Nov. 2019, pp. 2324–2330. doi: 10.1109/BIBM47256.2019.8983046.
- [2] S. F. Eskildsen, P. Coupé, V. S. Fonov, J. C. Pruessner, and D. L. Collins, "Structural imaging biomarkers of Alzheimer's disease: predicting disease progression," *Neurobiol. Aging*, vol. 36, pp. S23–S31, Jan. 2015, doi: 10.1016/j.neurobiolaging.2014.04.034.
- [3] A. M. Hall, R. Y. Moore, O. L. Lopez, L. Kuller, and J. T. Becker, "Basal forebrain atrophy is a presymptomatic marker for Alzheimer's disease," *Alzheimers Dement.*, vol. 4, no. 4, pp. 271–279, 2008, doi: 10.1016/j.jalz.2008.04.005.
- [4] G. B. Frisoni, N. C. Fox, C. R. Jack, P. Scheltens, and P. M. Thompson, "The clinical use of structural MRI in Alzheimer disease," *Nat. Rev. Neurol.*, vol. 6, no. 2, Art. no. 2, Feb. 2010, doi: 10.1038/nrneurol.2009.215.
- [5] E. M. Reiman and W. J. Jagust, "Brain imaging in the study of Alzheimer's disease," *NeuroImage*, vol. 61, no. 2, pp. 505–516, Jun. 2012, doi: 10.1016/j.neuroimage.2011.11.075.
- [6] S. M. Nestor *et al.*, "Ventricular enlargement as a possible measure of Alzheimer's disease progression validated using the Alzheimer's disease neuroimaging initiative database," *Brain*, vol. 131, no. 9, pp. 2443–2454, Aug. 2008, doi: 10.1093/brain/awn146.
- [7] M. Weiler, B. M. de Campos, M. H. Nogueira, B. Pereira Damasceno, F. Cendes, and M. L. F. Balthazar, "Structural connectivity of the default mode network and cognition in Alzheimer's disease," *Psychiatry Res. Neuroimaging*, vol. 223, no. 1, pp. 15–22, Jul. 2014, doi: 10.1016/j.psychresns.2014.04.008.
- [8] P. M. Matthews, N. Filippini, and G. Douaud, "Brain Structural and Functional Connectivity and the Progression of Neuropathology in Alzheimer's Disease," *J. Alzheimers Dis.*, vol. 33, no. s1, pp. S163–S172, Jan. 2013, doi: 10.3233/JAD-2012-129012.
- [9] J. Zimmermann *et al.*, "Differentiation of Alzheimer's disease based on local and global parameters in personalized Virtual Brain models," *NeuroImage Clin.*, vol. 19, pp. 240–251, Jan. 2018, doi: 10.1016/j.nicl.2018.04.017.
- [10] R. L. Buckner *et al.*, "Molecular, Structural, and Functional Characterization of Alzheimer's Disease: Evidence for a Relationship between Default Activity, Amyloid, and Memory," *J. Neurosci.*, vol. 25, no. 34, pp. 7709–7717, Aug. 2005, doi: 10.1523/JNEUROSCI.2177-05.2005.
- [11] C. A. Brown, Y. Jiang, C. D. Smith, and B. T. Gold, "Age and Alzheimer's pathology disrupt default mode network functioning via alterations in white matter microstructure but not hyperintensities," *Cortex*, vol. 104, pp. 58–74, Jul. 2018, doi: 10.1016/j.cortex.2018.04.006.

- [12] D. Xiao, K. Wang, L. Theriault, E. Charbel, and A. D. N. Initiative, "White matter integrity and key structures affected in Alzheimer's disease characterized by diffusion tensor imaging," *Eur. J. Neurosci.*, vol. n/a, no. n/a, Jan. 2022, doi: 10.1111/ejn.15815.
- [13] Y. Zhang *et al.*, "White matter damage in frontotemporal dementia and Alzheimer's disease measured by diffusion MRI," *Brain*, vol. 132, no. 9, pp. 2579–2592, Sep. 2009, doi: 10.1093/brain/awp071.
- [14] L. Clerx, P. J. Visser, F. Verhey, and P. Aalten, "New MRI Markers for Alzheimer's Disease: A Meta-Analysis of Diffusion Tensor Imaging and a Comparison with Medial Temporal Lobe Measurements," *J. Alzheimers Dis.*, vol. 29, no. 2, pp. 405–429, Jan. 2012, doi: 10.3233/JAD-2011-110797.
- [15] A. Hafkemeijer, J. van der Grond, and S. A. R. B. Rombouts, "Imaging the default mode network in aging and dementia," *Biochim. Biophys. Acta BBA - Mol. Basis Dis.*, vol. 1822, no. 3, pp. 431–441, Mar. 2012, doi: 10.1016/j.bbadis.2011.07.008.
- [16] M. R. Brier *et al.*, "Loss of Intranetwork and Internetwork Resting State Functional Connections with Alzheimer's Disease Progression," *J. Neurosci.*, vol. 32, no. 26, pp. 8890–8899, Jun. 2012, doi: 10.1523/JNEUROSCI.5698-11.2012.
- [17] B. Zhou *et al.*, "Structural and functional connectivity abnormalities of the default mode network in patients with Alzheimer's disease and mild cognitive impairment within two independent datasets," *Methods*, vol. 205, pp. 29–38, 2022, doi: 10.1016/j.ymeth.2022.06.001.
- [18] J. Venugopalan, L. Tong, H. R. Hassanzadeh, and M. D. Wang, "Multimodal deep learning models for early detection of Alzheimer's disease stage," *Sci. Rep.*, vol. 11, no. 1, Art. no. 1, Feb. 2021, doi: 10.1038/s41598-020-74399-w.
- [19] H. Jia, Y. Wang, Y. Duan, and H. Xiao, "Alzheimer's Disease Classification Based on Image Transformation and Features Fusion," *Comput. Math. Methods Med.*, vol. 2021, p. e9624269, Dec. 2021, doi: 10.1155/2021/9624269.
- [20] H. Shahamat and M. Saniee Abadeh, "Brain MRI analysis using a deep learning based evolutionary approach," *Neural Netw.*, vol. 126, pp. 218–234, Jun. 2020, doi: 10.1016/j.neunet.2020.03.017.
- [21] J. Rieke, F. Eitel, M. Weygandt, J.-D. Haynes, and K. Ritter, "Visualizing Convolutional Networks for MRI-Based Diagnosis of Alzheimer's Disease," in *Understanding and Interpreting Machine Learning in Medical Image Computing Applications*, D. Stoyanov, Z. Taylor, S. M. Kia, I. Oguz, M. Reyes, A. Martel, L. Maier-Hein, A. F. Marquand, E. Duchesnay, T. Löfstedt, B. Landman, M. J. Cardoso, C. A. Silva, S. Pereira, and R. Meier, Eds., in *Lecture Notes in Computer Science*. Cham: Springer International Publishing, 2018, pp. 24–31. doi: 10.1007/978-3-030-02628-8_3.
- [22] Y. Huang, J. Xu, Y. Zhou, T. Tong, X. Zhuang, and the Alzheimer's Disease Neuroimaging Initiative (ADNI), "Diagnosis of Alzheimer's Disease via Multi-Modality 3D Convolutional Neural Network," *Front. Neurosci.*, vol. 13, p. 509, May 2019, doi: 10.3389/fnins.2019.00509.
- [23] V. Bordin, D. Coluzzi, M. W. Rivolta, and G. Baselli, "Explainable AI Points to White Matter Hyperintensities for Alzheimer's Disease Identification: a Preliminary Study," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Jul. 2022, pp. 484–487. doi: 10.1109/EMBC48229.2022.9871306.
- [24] C. Yang, A. Rangarajan, and S. Ranka, "Visual Explanations From Deep 3D Convolutional Neural Networks for Alzheimer's Disease Classification," *AMIA. Annu. Symp. Proc.*, vol. 2018, pp. 1571–1580, Dec. 2018.
- [25] J. Zhou *et al.*, "Graph neural networks: A review of methods and applications," *AI Open*, vol. 1, pp. 57–81, Jan. 2020, doi: 10.1016/j.aiopen.2021.01.001.
- [26] R. J. Meszlényi, K. Buza, and Z. Vidnyánszky, "Resting State fMRI Functional Connectivity-Based Classification Using a Convolutional Neural Network Architecture," *Front. Neuroinformatics*, vol. 11, 2017, Accessed: Jan. 04, 2023. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fninf.2017.00061>

- [27] J. Kawahara *et al.*, “BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment,” *NeuroImage*, vol. 146, pp. 1038–1049, Feb. 2017, doi: 10.1016/j.neuroimage.2016.09.046.
- [28] H. Jiang, P. Cao, M. Xu, J. Yang, and O. Zaiane, “Hi-GCN: A hierarchical graph convolution network for graph embedding learning of brain network and brain disorders prediction,” *Comput. Biol. Med.*, vol. 127, p. 104096, Dec. 2020, doi: 10.1016/j.combiomed.2020.104096.
- [29] Y. Li *et al.*, “Brain Connectivity Based Graph Convolutional Networks and Its Application to Infant Age Prediction,” *IEEE Trans. Med. Imaging*, vol. 41, no. 10, pp. 2764–2776, Oct. 2022, doi: 10.1109/TMI.2022.3171778.
- [30] A. Alorf and M. U. G. Khan, “Multi-label classification of Alzheimer’s disease stages from resting-state fMRI-based correlation connectivity data and deep learning,” *Comput. Biol. Med.*, vol. 151, p. 106240, Dec. 2022, doi: 10.1016/j.combiomed.2022.106240.
- [31] C. Amodeo, I. Fortel, O. Ajilore, L. Zhan, A. Leow, and T. Tulabandhula, “Unified Embeddings of Structural and Functional Connectome via a Function-Constrained Structural Graph Variational Auto-Encoder,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*, L. Wang, Q. Dou, P. T. Fletcher, S. Speidel, and S. Li, Eds., in Lecture Notes in Computer Science, vol. 13431. Cham: Springer Nature Switzerland, 2022, pp. 406–415. doi: 10.1007/978-3-031-16431-6_39.
- [32] K. Aggarwal, M. M. Jimeno, K. S. Ravi, and G. Gonzalez, “Developing and deploying deep learning models in brain MRI: a review,” Mar. 2023.
- [33] K. Oh, Y.-C. Chung, K. W. Kim, W.-S. Kim, and I.-S. Oh, “Classification and Visualization of Alzheimer’s Disease using Volumetric Convolutional Neural Network and Transfer Learning,” *Sci. Rep.*, vol. 9, no. 1, p. 18150, Dec. 2019, doi: 10.1038/s41598-019-54548-6.
- [34] M. Leming, J. M. Górriz, and J. Suckling, “Ensemble Deep Learning on Large, Mixed-Site fMRI Datasets in Autism and Other Tasks,” *Int. J. Neural Syst.*, vol. 30, no. 07, p. 2050012, Jul. 2020, doi: 10.1142/S0129065720500124.
- [35] B.-H. Kim and J. C. Ye, “Understanding Graph Isomorphism Network for rs-fMRI Functional Connectivity Analysis,” *Front. Neurosci.*, vol. 14, 2020, Accessed: Feb. 13, 2023. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fnins.2020.00630>
- [36] A. Abrol, M. Bhattarai, A. Fedorov, Y. Du, S. Plis, and V. Calhoun, “Deep residual learning for neuroimaging: An application to predict progression to Alzheimer’s disease,” *J. Neurosci. Methods*, vol. 339, p. 108701, Jun. 2020, doi: 10.1016/j.jneumeth.2020.108701.
- [37] A. Essemlali, E. St-Onge, M. Descoteaux, and P.-M. Jodoin, “Understanding Alzheimer disease’s structural connectivity through explainable AI,” in *Proceedings of the Third Conference on Medical Imaging with Deep Learning*, PMLR, Sep. 2020, pp. 217–229. Accessed: Feb. 15, 2023. [Online]. Available: <https://proceedings.mlr.press/v121/essemlali20a.html>
- [38] P. J. LaMontagne *et al.*, “OASIS-3: Longitudinal Neuroimaging, Clinical, and Cognitive Dataset for Normal Aging and Alzheimer Disease,” *Radiology and Imaging*, preprint, Dec. 2019. doi: 10.1101/2019.12.13.19014902.
- [39] R. S. Desikan *et al.*, “An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest,” *NeuroImage*, vol. 31, no. 3, pp. 968–980, Jul. 2006, doi: 10.1016/j.neuroimage.2006.01.021.
- [40] N. Tzourio-Mazoyer *et al.*, “Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain,” *NeuroImage*, vol. 15, no. 1, pp. 273–289, Jan. 2002, doi: 10.1006/nimg.2001.0978.
- [41] M. Jenkinson, C. F. Beckmann, T. E. J. Behrens, M. W. Woolrich, and S. M. Smith, “FSL,” *NeuroImage*, vol. 62, no. 2, pp. 782–790, Aug. 2012, doi: 10.1016/j.neuroimage.2011.09.015.

- [42] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, Jun. 2002, doi: 10.1613/jair.953.
- [43] C. Li, Y. Wei, X. Chen, and C.-B. Schönlieb, "BrainNetGAN: Data Augmentation of Brain Connectivity Using Generative Adversarial Network for Dementia Classification," in *Deep Generative Models, and Data Augmentation, Labelling, and Imperfections*, S. Engelhardt, I. Oksuz, D. Zhu, Y. Yuan, A. Mukhopadhyay, N. Heller, S. X. Huang, H. Nguyen, R. Sznitman, and Y. Xue, Eds., in Lecture Notes in Computer Science. Cham: Springer International Publishing, 2021, pp. 103–111. doi: 10.1007/978-3-030-88210-5_9.
- [44] S. Sarraf and G. Tofghi, "Deep learning-based pipeline to recognize Alzheimer's disease using fMRI data," in *2016 Future Technologies Conference (FTC)*, San Francisco, CA, USA: IEEE, Dec. 2016, pp. 816–820. doi: 10.1109/FTC.2016.7821697.
- [45] O. Russakovsky *et al.*, "ImageNet Large Scale Visual Recognition Challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015, doi: 10.1007/s11263-015-0816-y.
- [46] R. Solovyev, A. A. Kalinin, and T. Gabruseva, "3D convolutional neural networks for stalled brain capillary detection," *Comput. Biol. Med.*, vol. 141, p. 105089, Feb. 2022, doi: 10.1016/j.compbiomed.2021.105089.
- [47] D. S. Bassett and E. T. Bullmore, "Small-World Brain Networks Revisited," *The Neuroscientist*, vol. 23, no. 5, pp. 499–516, Oct. 2017, doi: 10.1177/1073858416667720.
- [48] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization," *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 336–359, Feb. 2020, doi: 10.1007/s11263-019-01228-7.
- [49] X. Shi, S. Khademi, Y. Li, and J. van Gemert, "Zoom-CAM: Generating Fine-grained Pixel Annotations from Image Labels." arXiv, Oct. 16, 2020. Accessed: Mar. 26, 2023. [Online]. Available: <http://arxiv.org/abs/2010.08644>
- [50] P. Morbidelli, D. Carrera, B. Rossi, P. Fragneto, and G. Boracchi, "Augmented Grad-CAM: Heat-Maps Super Resolution Through Augmentation," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2020, pp. 4067–4071. doi: 10.1109/ICASSP40776.2020.9054416.
- [51] X. Lei, Y. Fan, and X.-L. Luo, "On fine-grained visual explanation in convolutional neural networks," *Digit. Commun. Netw.*, Dec. 2022, doi: 10.1016/j.dcan.2022.12.012.
- [52] J. Gildenblat and contributors, "PyTorch library for CAM methods." GitHub, 2021. [Online]. Available: <https://github.com/jacobgil/pytorch-grad-cam>
- [53] D. Coluzzi *et al.*, "Development and Testing of SPIDER-NET: An Interactive Tool for Brain Connectogram Visualization, Sub-Network Exploration and Graph Metrics Quantification," *Front. Neurosci.*, vol. 16, 2022, Accessed: May 19, 2022. [Online]. Available: <https://www.frontiersin.org/article/10.3389/fnins.2022.818385>
- [54] P. Finotelli, C. Piccardi, E. Miglio, and P. Dulio, "A Graphlet-Based Topological Characterization of the Resting-State Network in Healthy People," *Front. Neurosci.*, vol. 15, p. 665544, Apr. 2021, doi: 10.3389/fnins.2021.665544.
- [55] W. Li, X. Mai, and C. Liu, "The default mode network and social understanding of others: what do brain connectivity studies tell us," *Front. Hum. Neurosci.*, vol. 8, 2014, Accessed: Oct. 20, 2022. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fnhum.2014.00074>
- [56] M. E. Raichle, "The Brain's Default Mode Network," *Annu. Rev. Neurosci.*, vol. 38, no. 1, pp. 433–447, 2015, doi: 10.1146/annurev-neuro-071013-014030.
- [57] K. Çiftçi, "Minimum Spanning Tree Reflects the Alterations of the Default Mode Network During Alzheimer's Disease," *Ann. Biomed. Eng.*, vol. 39, no. 5, pp. 1493–1504, May 2011, doi: 10.1007/s10439-011-0258-9.
- [58] P. N. Alves *et al.*, "An improved neuroanatomical model of the default-mode network reconciles previous neuroimaging and neuropathological findings," *Commun. Biol.*, vol. 2, no. 1, Art. no. 1, Oct. 2019, doi: 10.1038/s42003-019-0611-3.

- [59] V. Cutsuridis and M. Yoshida, "Editorial: Memory Processes in Medial Temporal Lobe: Experimental, Theoretical and Computational Approaches," *Front. Syst. Neurosci.*, vol. 11, Apr. 2017, doi: 10.3389/fnsys.2017.00019.
- [60] G. B. Frisoni, N. C. Fox, C. R. Jack, P. Scheltens, and P. M. Thompson, "The clinical use of structural MRI in Alzheimer disease," *Nat. Rev. Neurol.*, vol. 6, no. 2, pp. 67–77, Feb. 2010, doi: 10.1038/nrneuro.2009.215.
- [61] Y. Benjamini and D. Yekutieli, "The control of the false discovery rate in multiple testing under dependency," *Ann. Stat.*, vol. 29, no. 4, pp. 1165–1188, Aug. 2001, doi: 10.1214/aos/1013699998.
- [62] J. S. W. Campbell and G. B. Pike, "Potential and limitations of diffusion MRI tractography for the study of language," *Brain Lang.*, vol. 131, pp. 65–73, Apr. 2014, doi: 10.1016/j.bandl.2013.06.007.
- [63] K. H. Maier-Hein *et al.*, "The challenge of mapping the human connectome based on diffusion tractography," *Nat. Commun.*, vol. 8, no. 1, Art. no. 1, Nov. 2017, doi: 10.1038/s41467-017-01285-x.
- [64] C. Bordier, C. Nicolini, and A. Bifone, "Graph Analysis and Modularity of Brain Functional Connectivity Networks: Searching for the Optimal Threshold," *Front. Neurosci.*, vol. 11, 2017, Accessed: Feb. 22, 2023. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fnins.2017.00441>
- [65] P. Coupé *et al.*, "Scoring by nonlocal image patch estimator for early detection of Alzheimer's disease," *NeuroImage Clin.*, vol. 1, no. 1, pp. 141–152, Jan. 2012, doi: 10.1016/j.nicl.2012.10.002.
- [66] H. Zetterberg and B. B. Bendlin, "Biomarkers for Alzheimer's disease—preparing for a new era of disease-modifying therapies," *Mol. Psychiatry*, vol. 26, no. 1, pp. 296–308, Jan. 2021, doi: 10.1038/s41380-020-0721-9.
- [67] M. De Leon, A. George, L. Stylopoulos, G. Smith, and D. Miller, "EARLY MARKER FOR ALZHEIMER'S DISEASE: THE ATROPHIC HIPPOCAMPUS," *The Lancet*, vol. 334, no. 8664, pp. 672–673, Sep. 1989, doi: 10.1016/S0140-6736(89)90911-2.
- [68] A. M. Hall, R. Y. Moore, O. L. Lopez, L. Kuller, and J. T. Becker, "Basal forebrain atrophy is a presymptomatic marker for Alzheimer's disease," *Alzheimers Dement.*, vol. 4, no. 4, pp. 271–279, Jul. 2008, doi: 10.1016/j.jalz.2008.04.005.
- [69] G. B. Frisoni *et al.*, "The pilot European Alzheimer's Disease Neuroimaging Initiative of the European Alzheimer's Disease Consortium," *Alzheimers Dement.*, vol. 4, no. 4, pp. 255–264, Jul. 2008, doi: 10.1016/j.jalz.2008.04.009.
- [70] C. R. McDonald *et al.*, "Regional rates of neocortical atrophy from normal aging to early Alzheimer disease," *Neurology*, vol. 73, no. 6, pp. 457–465, Aug. 2009, doi: 10.1212/WNL.0b013e3181b16431.
- [71] J. Wegiel *et al.*, "Cerebellar atrophy in Alzheimer's disease—clinicopathological correlations," *Brain Res.*, vol. 818, no. 1, pp. 41–50, Feb. 1999, doi: 10.1016/S0006-8993(98)01279-7.
- [72] M. Sjöbeck and E. Englund, "Alzheimer's Disease and the Cerebellum: A Morphologic Study on Neuronal and Glial Changes," *Dement. Geriatr. Cogn. Disord.*, vol. 12, no. 3, pp. 211–218, 2001, doi: 10.1159/000051260.
- [73] C. Yang, S. Zhong, X. Zhou, L. Wei, L. Wang, and S. Nie, "The Abnormality of Topological Asymmetry between Hemispheric Brain White Matter Networks in Alzheimer's Disease and Mild Cognitive Impairment," *Front. Aging Neurosci.*, vol. 9, 2017, Accessed: Feb. 13, 2023. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fnagi.2017.00261>
- [74] M. Daianu *et al.*, "Breakdown of Brain Connectivity Between Normal Aging and Alzheimer's Disease: A Structural k-Core Network Analysis," *Brain Connect.*, vol. 3, no. 4, pp. 407–422, Aug. 2013, doi: 10.1089/brain.2012.0137.
- [75] C. E. Sexton, U. G. Kalu, N. Filippini, C. E. Mackay, and K. P. Ebmeier, "A meta-analysis of diffusion tensor imaging in mild cognitive impairment and Alzheimer's disease," *Neurobiol. Aging*, vol. 32, no. 12, p. 2322.e5-2322.e18, Dec. 2011, doi: 10.1016/j.neurobiolaging.2010.05.019.
- [76] F. Agosta *et al.*, "White Matter Damage in Alzheimer Disease and Its Relationship to Gray Matter Atrophy," *Radiology*, vol. 258, no. 3, pp. 853–863, Mar. 2011, doi: 10.1148/radiol.10101284.

- [77] J. R. Harrison *et al.*, “Imaging Alzheimer’s genetic risk using diffusion MRI: A systematic review,” *NeuroImage Clin.*, vol. 27, p. 102359, Jan. 2020, doi: 10.1016/j.nicl.2020.102359.
- [78] L. T. Westlye, I. Reinvang, H. Rootwelt, and T. Espeseth, “Effects of APOE on brain white matter microstructure in healthy adults,” *Neurology*, vol. 79, no. 19, pp. 1961–1969, Nov. 2012, doi: 10.1212/WNL.0b013e3182735c9c.
- [79] X. Nie *et al.*, “Subregional Structural Alterations in Hippocampus and Nucleus Accumbens Correlate with the Clinical Impairment in Patients with Alzheimer’s Disease Clinical Spectrum: Parallel Combining Volume and Vertex-Based Approach,” *Front. Neurol.*, vol. 8, p. 399, Aug. 2017, doi: 10.3389/fneur.2017.00399.
- [80] S. Toniolo *et al.*, “Cerebellar White Matter Disruption in Alzheimer’s Disease Patients: A Diffusion Tensor Imaging Study,” *J. Alzheimers Dis.*, vol. 74, no. 2, pp. 615–624, Jan. 2020, doi: 10.3233/JAD-191125.
- [81] L. Walker *et al.*, “Effects of physiological noise in population analysis of diffusion tensor MRI data,” *NeuroImage*, vol. 54, no. 2, pp. 1168–1177, Jan. 2011, doi: 10.1016/j.neuroimage.2010.08.048.
- [82] J. M. Wardlaw, M. C. Valdés Hernández, and S. Muñoz-Maniega, “What are White Matter Hyperintensities Made of?: Relevance to Vascular Cognitive Impairment,” *J. Am. Heart Assoc.*, vol. 4, no. 6, Jun. 2015, doi: 10.1161/JAHA.114.001140.
- [83] A. Kuceyeski, J. Maruta, N. Relkin, and A. Raj, “The Network Modification (NeMo) Tool: Elucidating the Effect of White Matter Integrity Changes on Cortical and Subcortical Structural Connectivity,” *Brain Connect.*, vol. 3, no. 5, pp. 451–463, Oct. 2013, doi: 10.1089/brain.2013.0147.
- [84] J. Liu, J. Wang, Z. Tang, B. Hu, F.-X. Wu, and Y. Pan, “Improving Alzheimer’s Disease Classification by Combining Multiple Measures,” *IEEE/ACM Trans. Comput. Biol. Bioinform.*, vol. 15, no. 5, pp. 1649–1659, Sep. 2018, doi: 10.1109/TCBB.2017.2731849.

Scientific Publications

Papers

Articles part of this PhD Thesis:

- **Coluzzi, D.***; Bordin, V.*; Rivolta M.W.; Fortel I.; Zhan L.; Leow A.; Baselli, G. (*equal contributions); Biomarker Investigation using Multiple Brain Measures from MRI through XAI in Alzheimer's Disease Classification, <https://doi.org/10.48550/arXiv.2305.03056> (*Under Submission*)
- **Coluzzi, D.**; Baselli, G., Diffuse and Localized Functional Dysconnectivity in Schizophrenia: a Bootstrapped Top-Down Approach, *Fundamenta Informaticae* 2022, <https://doi.org/10.48550/arXiv.2305.02369>, Special Issue Tomography and Applications, 2022 (*Accepted, Pending Publication*)
- **Coluzzi, D.**; Pirastru, A.; Pelizzari, L.; Cabinio, M.; Laganà, M.M.; Baselli, G.; Baglio, F., Development and Testing of SPIDER-NET: An Interactive Tool for Brain Connectogram Visualization, Sub-Network Exploration and Graph Metrics Quantification, *Frontiers in Neuroscience* 2022, 16: 818385-818385, <https://doi.org/10.3389/fnins.2022.818385>, Special Issue Advanced Computational Tools for Mapping the Multidimensional Architecture of the Brain

Other articles published during the PhD, but not included in the PhD Thesis:

- **Coluzzi, D.**; Baselli, G.; Bianchi, A.M.; Guerrero-Mora, G.; Kortelainen, J.M.; Tenhunen, M.L.; Mendez, M.O. Multi-Scale Evaluation of Sleep Quality Based on Motion Signal from Unobtrusive Device, *Sensors* 2022, 22, 5295. <https://doi.org/10.3390/s22145295>, Special Issue Biomedical Sensors for Functional Mapping: Techniques, Methods, Experimental and Medical Applications
- **Coluzzi, D.**; Rivolta, M.W.; Mastropietro, A.; Porcelli, S.; Mauri, M.L.; Civiello, M.T.L.; Denna, E.; Rizzo, G.; Sassi, R., Design and Validation of a Minimal Complexity Algorithm for Stair Step Counting, *Computers* 2020, 9, 31. <https://doi.org/10.3390/computers9020031>, Special Issue Methods and Applications for Imaging, Simulation, and Modelling in Biology and Medicine: Artificial Intelligence, Current Research, New Trends

Conference Contributions

- 05/16/2023 – Invited speaker at 17th Edition Meeting on Tomography and Applications Discrete Tomography, Neuroscience and Image Reconstruction (<https://www.mate.polimi.it/events/TAIR2023/>) with the seminar “Biomarker Investigation using Multiple Brain Measures from MRI through XAI in Alzheimer’s Disease Classification”, Dipartimento di Matematica – Politecnico di Milano, Milano (IT)
- Bordin, V.*; **Coluzzi, D.***; Rivolta, M.W; and Baselli, G. (*equal contributions); Explainable AI Reveals White Matter Hyperintensities for Alzheimer’s Disease Identification: a Preliminary Study, *44th Annual International Conference of the IEEE Engineering in Medicine and Biology Society 2022*, <https://doi.org/10.1109/EMBC48229.2022.9871306>, Glasgow (SC)
- **Coluzzi, D.**; Pirastru, A.; Pelizzari, L.; Cabinio, M.; Baselli, G.; and Baglio F., A Tool for Brain Networks Visualization and Investigation Finalized to Patient-Tailored Rehabilitation: SPIDER-NET, *28th OHBM Annual Meeting 2022*, Glasgow (SC)
- 05/03/2022 – Invited speaker at 16th Edition Meeting on Tomography and Applications Discrete Tomography, Neuroscience and Image Reconstruction (<https://www.mate.polimi.it/events/TAIR2022/>) with the seminar “Brain Connectivity through Graph Theory: SPIDER-NET a New Tool to Explore Sub-Networks”, Dipartimento di Matematica – Politecnico di Milano, Milano (IT)

Other

Preliminary works were also proposed as MSc Theses at Politecnico di Milano, for which the author was a co-advisor:

- Amodeo C.; Baselli G. (advisor); Leow A. (external advisor); Coluzzi D. (co-advisor); Human Connectome joint analysis: a unified embedding of Structural and Functional connectivity via a variational graph deep learning approach. Master of Science in Biomedical Engineering, Politecnico di Milano.
- Rocco C.; Baselli G. (advisor); Coluzzi D. (co-advisor); Brain connectivity assessment of graph and sub-graph indices: uncertainty evaluation by bootstrap and sensitivity analysis. Master of Science in Biomedical Engineering, Politecnico di Milano, <http://hdl.handle.net/10589/195725>.
- Curcic V.; Baselli G. (advisor); Pelizzari L.; Pirastru A.; Coluzzi D.; (co-advisors); Analysis of structural connectivity in brain disorders and validation of a novel tool for the investigation of brain connectivity. Master of Science in Biomedical Engineering, Politecnico di Milano, <http://hdl.handle.net/10589/177230>.

List of Abbreviations

| | |
|--------------|--|
| AAL | Automated Anatomical Labeling |
| AD | Alzheimer's Disease |
| AxD | Axial Diffusivity |
| ADC | Apparent Diffusion Coefficient |
| AI | Artificial Intelligence |
| BOLD | Blood Oxygenation Level Dependent |
| BOOT | Bootstrap Hypothesis Testing |
| CDR | Clinical Dementia Rating |
| CNN | Convolutional Neural Network |
| CSD | Constrained Spherical Deconvolution |
| DL | Deep Learning |
| DMN | Default Mode Network |
| dMRI | Diffusion Magnetic Resonance Imaging |
| DTI | Diffusion Tensor Imaging |
| DWI | Diffusion Weighted Imaging |
| FA | Fractional Anisotropy |
| FC | Functional Connectivity |
| FLAIR | Fluid Attenuated Inversion Recovery |
| fMRI | functional Magnetic Resonance Imaging |
| GCN | Graph Convolutional Networks |
| GM | Gray Matter |
| GNN | Graph Neural Network |
| GUI | Graphical User Interface |
| HARDI | High Angular Resolution Diffusion Imaging |
| HC | Healthy Control |
| HOA | FSL Harvard-Oxford maximum likelihood cortical Atlas |
| ICA | Independent Component Analysis |
| MCI | Mild Cognitive Impairment |
| MD | Mean Diffusivity |
| ML | Machine Learning |
| MRI | Magnetic Resonance Imaging |
| MTL | Medial Temporal Lobe |
| MW | Mann-Whitney Testing |
| NBS | Network Based Statistics tool |
| NF | Number of Fibers |
| OASIS | Open Access Series of Imaging Studies |

| | |
|------------------------|---|
| RD | Radial Diffusivity |
| ROI | Region Of Interest |
| RS-fMRI | Resting State functional Magnetic Resonance Imagining |
| RSN | Resting-State Network |
| RST_n | Robust Statistical Testing with leave n subject out |
| RV | Relevance Value |
| SC | Structural Connectivity |
| SZ | Schizophrenia |
| VR | Volume Ratio |
| WM | White Matter |
| XAI | Explainable Artificial Intelligence |

References

- [1] A. Fornito, A. Zalesky, and E. T. Bullmore, *Fundamentals of brain network analysis*. Amsterdam ; Boston: Elsevier/Academic Press, 2016.
- [2] O. Sporns, *Networks of the brain*. Cambridge, Mass: MIT Press, 2011.
- [3] P. Barttfeld *et al.*, “State-dependent changes of connectivity patterns and functional brain network topology in autism spectrum disorder,” *Neuropsychologia*, vol. 50, no. 14, pp. 3653–3662, Dec. 2012, doi: 10.1016/j.neuropsychologia.2012.09.047.
- [4] M. Göttlich, T. F. Münte, M. Heldmann, M. Kasten, J. Hagenah, and U. M. Krämer, “Altered Resting State Brain Networks in Parkinson’s Disease,” *PLOS ONE*, vol. 8, no. 10, p. e77336, Oct 2013, doi: 10.1371/journal.pone.0077336.
- [5] H.-C. Baggio *et al.*, “Functional brain networks and cognitive deficits in Parkinson’s disease,” *Hum. Brain Mapp.*, vol. 35, no. 9, pp. 4620–4634, 2014, doi: 10.1002/hbm.22499.
- [6] M. Daianu *et al.*, “Alzheimer’s disease disrupts rich club organization in brain connectivity networks,” in *2013 IEEE 10th International Symposium on Biomedical Imaging*, Apr. 2013, pp. 266–269. doi: 10.1109/ISBI.2013.6556463.
- [7] V. Blasi *et al.*, “Early Life Adversities and Borderline Intellectual Functioning Negatively Impact Limbic System Connectivity in Childhood: A Connectomics-Based Study,” *Front. Psychiatry*, vol. 11, 2020, doi: 10.3389/fpsy.2020.497116.
- [8] J. J. Crofts *et al.*, “Network analysis detects changes in the contralesional hemisphere following stroke,” *NeuroImage*, vol. 54, no. 1, pp. 161–169, Jan. 2011, doi: 10.1016/j.neuroimage.2010.08.032.
- [9] D. S. Bassett, B. G. Nelson, B. A. Mueller, J. Camchong, and K. O. Lim, “Altered resting state complexity in schizophrenia,” *NeuroImage*, vol. 59, no. 3, pp. 2196–2207, Feb. 2012, doi: 10.1016/j.neuroimage.2011.10.002.

- [10] G.-J. Ji, Y. Yu, H.-H. Miao, Z.-J. Wang, Y.-L. Tang, and W. Liao, “Decreased Network Efficiency in Benign Epilepsy with Centrottemporal Spikes,” *Radiology*, vol. 283, no. 1, pp. 186–194, Apr. 2017, doi: 10.1148/radiol.2016160422.
- [11] C. Destrieux, B. Fischl, A. Dale, and E. Halgren, “Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature,” *NeuroImage*, vol. 53, no. 1, pp. 1–15, Oct. 2010, doi: 10.1016/j.neuroimage.2010.06.010.
- [12] A. Zalesky, A. Fornito, and E. T. Bullmore, “Network-based statistic: Identifying differences in brain networks,” *NeuroImage*, vol. 53, no. 4, pp. 1197–1207, Dec. 2010, doi: 10.1016/j.neuroimage.2010.06.041.
- [13] N. Tzourio-Mazoyer *et al.*, “Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain,” *NeuroImage*, vol. 15, no. 1, pp. 273–289, Jan. 2002, doi: 10.1006/nimg.2001.0978.
- [14] P. J. LaMontagne *et al.*, “OASIS-3: Longitudinal Neuroimaging, Clinical, and Cognitive Dataset for Normal Aging and Alzheimer Disease,” *Radiology and Imaging*, preprint, Dec. 2019. doi: 10.1101/2019.12.13.19014902.
- [15] C. Amodeo, I. Fortel, O. Ajilore, L. Zhan, A. Leow, and T. Tulabandhula, “Unified Embeddings of Structural and Functional Connectome via a Function-Constrained Structural Graph Variational Auto-Encoder,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*, L. Wang, Q. Dou, P. T. Fletcher, S. Speidel, and S. Li, Eds., in Lecture Notes in Computer Science. Cham: Springer Nature Switzerland, 2022, pp. 406–415. doi: 10.1007/978-3-031-16431-6_39.
- [16] R. S. Desikan *et al.*, “An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest,” *NeuroImage*, vol. 31, no. 3, pp. 968–980, Jul. 2006, doi: 10.1016/j.neuroimage.2006.01.021.
- [17] M. Jenkinson, C. F. Beckmann, T. E. J. Behrens, M. W. Woolrich, and S. M. Smith, “FSL,” *NeuroImage*, vol. 62, no. 2, pp. 782–790, Aug. 2012, doi: 10.1016/j.neuroimage.2011.09.015.
- [18] O. Russakovsky *et al.*, “ImageNet Large Scale Visual Recognition Challenge,” *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015, doi: 10.1007/s11263-015-0816-y.

- [19] R. Solovyev, A. A. Kalinin, and T. Gabruseva, “3D convolutional neural networks for stalled brain capillary detection,” *Comput. Biol. Med.*, vol. 141, p. 105089, Feb. 2022, doi: 10.1016/j.combiomed.2021.105089.
- [20] Y. Li *et al.*, “Brain Connectivity Based Graph Convolutional Networks and Its Application to Infant Age Prediction,” *IEEE Trans. Med. Imaging*, vol. 41, no. 10, pp. 2764–2776, Oct. 2022, doi: 10.1109/TMI.2022.3171778.
- [21] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization,” *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 336–359, Feb. 2020, doi: 10.1007/s11263-019-01228-7.
- [22] V. Cutsuridis and M. Yoshida, “Editorial: Memory Processes in Medial Temporal Lobe: Experimental, Theoretical and Computational Approaches,” *Front. Syst. Neurosci.*, vol. 11, Apr. 2017, doi: 10.3389/fnsys.2017.00019.
- [23] G. B. Frisoni, N. C. Fox, C. R. Jack, P. Scheltens, and P. M. Thompson, “The clinical use of structural MRI in Alzheimer disease,” *Nat. Rev. Neurol.*, vol. 6, no. 2, Art. no. 2, Feb. 2010, doi: 10.1038/nrneurol.2009.215.
- [24] M. Weiler, B. M. de Campos, M. H. Nogueira, B. Pereira Damasceno, F. Cendes, and M. L. F. Balthazar, “Structural connectivity of the default mode network and cognition in Alzheimer’s disease,” *Psychiatry Res. Neuroimaging*, vol. 223, no. 1, pp. 15–22, Jul. 2014, doi: 10.1016/j.psychresns.2014.04.008.
- [25] R. L. Buckner *et al.*, “Molecular, Structural, and Functional Characterization of Alzheimer’s Disease: Evidence for a Relationship between Default Activity, Amyloid, and Memory,” *J. Neurosci.*, vol. 25, no. 34, pp. 7709–7717, Aug. 2005, doi: 10.1523/JNEUROSCI.2177-05.2005.
- [26] Y. Li, Z. Yu, P. Wu, and J. Chen, “The disrupted topological properties of structural networks showed recovery in ischemic stroke patients: a longitudinal design study,” *BMC Neurosci.*, vol. 22, no. 1, p. 47, Aug. 2021, doi: 10.1186/s12868-021-00652-1.
- [27] B. Cheng *et al.*, “Altered topology of large-scale structural brain networks in chronic stroke,” *Brain Commun.*, vol. 1, no. 1, Jan. 2019, doi: 10.1093/braincomms/fcz020.

- [28] F. Dell'Acqua, G. Rizzo, P. Scifo, R. A. Clarke, G. Scotti, and F. Fazio, "A Model-Based Deconvolution Approach to Solve Fiber Crossing in Diffusion-Weighted MR Imaging," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 3, pp. 462–472, Mar. 2007, doi: 10.1109/TBME.2006.888830.
- [29] J. S. W. Campbell and G. B. Pike, "Potential and limitations of diffusion MRI tractography for the study of language," *Brain Lang.*, vol. 131, pp. 65–73, Apr. 2014, doi: 10.1016/j.bandl.2013.06.007.
- [30] K. H. Maier-Hein *et al.*, "The challenge of mapping the human connectome based on diffusion tractography," *Nat. Commun.*, vol. 8, no. 1, Art. no. 1, Nov. 2017, doi: 10.1038/s41467-017-01285-x.
- [31] N. Peskin, D. Koren, and S. Gabay, "Subcortical neural tracks play an important role in executive function in schizophrenia: An experimental study among patients with schizophrenia and healthy comparisons," *Schizophr. Res. Cogn.*, vol. 22, p. 100185, Dec. 2020, doi: 10.1016/j.scog.2020.100185.
- [32] A. Beech, T. Powell, J. McWilliam, and G. Claridge, "Evidence of reduced 'cognitive inhibition' in schizophrenia," *Br. J. Clin. Psychol.*, vol. 28, no. 2, pp. 109–116, 1989, doi: 10.1111/j.2044-8260.1989.tb00821.x.
- [33] I. Henseler, P. Falkai, and O. Gruber, "Disturbed functional connectivity within brain networks subserving domain-specific subcomponents of working memory in schizophrenia: Relation to performance and clinical symptoms," *J. Psychiatr. Res.*, vol. 44, no. 6, pp. 364–372, Apr. 2010, doi: 10.1016/j.jpsychires.2009.09.003.
- [34] T. Fujimoto *et al.*, "Dysfunctional Cortical Connectivity During the Auditory Oddball Task in Patients with Schizophrenia," *Open Neuroimaging J.*, vol. 7, pp. 15–26, Apr. 2013, doi: 10.2174/18744440001307010015.
- [35] M. Yildiz, S. J. Borgwardt, and G. E. Berger, "Parietal Lobes in Schizophrenia: Do They Matter?," *Schizophr. Res. Treat.*, vol. 2011, p. e581686, Oct. 2011, doi: 10.1155/2011/581686.
- [36] P. Skudlarski *et al.*, "Brain Connectivity Is Not Only Lower but Different in Schizophrenia: A Combined Anatomical and Functional Approach," *Biol. Psychiatry*, vol. 68, no. 1, pp. 61–69, Jul. 2010, doi: 10.1016/j.biopsych.2010.03.035.

- [37] M. P. van den Heuvel and A. Fornito, “Brain Networks in Schizophrenia,” *Neuropsychol. Rev.*, vol. 24, no. 1, pp. 32–48, Mar. 2014, doi: 10.1007/s11065-014-9248-7.
- [38] H. Jiang, P. Cao, M. Xu, J. Yang, and O. Zaiane, “Hi-GCN: A hierarchical graph convolution network for graph embedding learning of brain network and brain disorders prediction,” *Comput. Biol. Med.*, vol. 127, p. 104096, Dec. 2020, doi: 10.1016/j.compbiomed.2020.104096.
- [39] A. Abrol, M. Bhattarai, A. Fedorov, Y. Du, S. Plis, and V. Calhoun, “Deep residual learning for neuroimaging: An application to predict progression to Alzheimer’s disease,” *J. Neurosci. Methods*, vol. 339, p. 108701, Jun. 2020, doi: 10.1016/j.jneumeth.2020.108701.
- [40] C. Yang, S. Zhong, X. Zhou, L. Wei, L. Wang, and S. Nie, “The Abnormality of Topological Asymmetry between Hemispheric Brain White Matter Networks in Alzheimer’s Disease and Mild Cognitive Impairment,” *Front. Aging Neurosci.*, vol. 9, 2017, <https://www.frontiersin.org/articles/10.3389/fnagi.2017.00261>
- [41] M. Daianu *et al.*, “Breakdown of Brain Connectivity Between Normal Aging and Alzheimer’s Disease: A Structural k-Core Network Analysis,” *Brain Connect.*, vol. 3, no. 4, pp. 407–422, Aug. 2013, doi: 10.1089/brain.2012.0137.
- [42] L. Clerx, P. J. Visser, F. Verhey, and P. Aalten, “New MRI Markers for Alzheimer’s Disease: A Meta-Analysis of Diffusion Tensor Imaging and a Comparison with Medial Temporal Lobe Measurements,” *J. Alzheimers Dis.*, vol. 29, no. 2, pp. 405–429, Jan. 2012, doi: 10.3233/JAD-2011-110797.
- [43] J. Liu, J. Wang, Z. Tang, B. Hu, F.-X. Wu, and Y. Pan, “Improving Alzheimer’s Disease Classification by Combining Multiple Measures,” *IEEE/ACM Trans. Comput. Biol. Bioinform.*, vol. 15, no. 5, pp. 1649–1659, Sep. 2018, doi: 10.1109/TCBB.2017.2731849.
- [44] O. Sporns, “Network attributes for segregation and integration in the human brain,” *Curr. Opin. Neurobiol.*, vol. 23, no. 2, pp. 162–171, Apr. 2013, doi: 10.1016/j.conb.2012.11.015.

- [45] M.-E. Lynall *et al.*, “Functional Connectivity and Brain Networks in Schizophrenia,” *J. Neurosci.*, vol. 30, no. 28, pp. 9477–9487, Jul. 2010, doi: 10.1523/JNEUROSCI.0333-10.2010.
- [46] Y. Liu *et al.*, “Disrupted small-world networks in schizophrenia,” *Brain*, vol. 131, no. 4, pp. 945–961, Apr. 2008, doi: 10.1093/brain/awn018.
- [47] A. Fornito, A. Zalesky, C. Pantelis, and E. T. Bullmore, “Schizophrenia, neuroimaging and connectomics,” *NeuroImage*, vol. 62, no. 4, pp. 2296–2314, Oct. 2012, doi: 10.1016/j.neuroimage.2011.12.090.
- [48] B. T. Thomas Yeo *et al.*, “The organization of the human cerebral cortex estimated by intrinsic functional connectivity,” *J. Neurophysiol.*, vol. 106, no. 3, pp. 1125–1165, Jun. 2011, doi: 10.1152/jn.00338.2011.
- [49] R. M. Hutchison *et al.*, “Dynamic functional connectivity: Promise, issues, and interpretations,” *NeuroImage*, vol. 80, pp. 360–378, Oct. 2013, doi: 10.1016/j.neuroimage.2013.05.079.
- [50] A. Kottaram *et al.*, “Brain network dynamics in schizophrenia: Reduced dynamism of the default mode network,” *Hum. Brain Mapp.*, vol. 40, no. 7, pp. 2212–2228, 2019, doi: 10.1002/hbm.24519.
- [51] R. Liégeois, T. O. Laumann, A. Z. Snyder, J. Zhou, and B. T. T. Yeo, “Interpreting temporal fluctuations in resting-state functional connectivity MRI,” *NeuroImage*, vol. 163, pp. 437–455, Dec. 2017, doi: 10.1016/j.neuroimage.2017.09.012.
- [52] J. Zhou *et al.*, “Graph neural networks: A review of methods and applications,” *AI Open*, vol. 1, pp. 57–81, Jan. 2020, doi: 10.1016/j.aiopen.2021.01.001.
- [53] A. Singh, S. Sengupta, and V. Lakshminarayanan, “Explainable Deep Learning Models in Medical Image Analysis,” *J. Imaging*, vol. 6, no. 6, Art. no. 6, Jun. 2020, doi: 10.3390/jimaging6060052.
- [54] J. Annese, “In Retrospect: Brodmann’s brain map,” *Nature*, vol. 461, no. 7266, Art. no. 7266, Oct. 2009, doi: 10.1038/461884a.
- [55] K. Brodmann, “Comparative localization studies in the brain cortex, its fundamentals represented on the basis of its cellular architecture,” *Leipz. JA Barth*, 1909.

- [56] W. L. Nowinski, “Evolution of Human Brain Atlases in Terms of Content, Applications, Functionality, and Availability,” *Neuroinformatics*, vol. 19, no. 1, pp. 1–22, Jan. 2021, doi: 10.1007/s12021-020-09481-9.
- [57] G. Schaltenbrand, “Atlas for stereotaxy of the human brain,” *Georg Thieme*, 1977.
- [58] D. Le Bihan *et al.*, “Diffusion tensor imaging: Concepts and applications,” *J. Magn. Reson. Imaging*, vol. 13, no. 4, pp. 534–546, 2001, doi: 10.1002/jmri.1076.
- [59] “Multimodal NeuroImaging Analysis: Image.”
<https://multimodalneuroimaging.files.wordpress.com/2015/04/dwi.jpg>.
- [60] D. S. Tuch, T. G. Reese, M. R. Wiegell, N. Makris, J. W. Belliveau, and V. J. Wedeen, “High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity,” *Magn. Reson. Med.*, vol. 48, no. 4, pp. 577–582, 2002, doi: 10.1002/mrm.10268.
- [61] K. Pannek, J. L. Mathias, E. D. Bigler, G. Brown, J. D. Taylor, and S. Rose, “An automated strategy for the delineation and parcellation of commissural pathways suitable for clinical populations utilising high angular resolution diffusion imaging tractography,” *NeuroImage*, vol. 50, no. 3, pp. 1044–1053, Apr. 2010, doi: 10.1016/j.neuroimage.2010.01.020.
- [62] T. e. j. Behrens *et al.*, “Characterization and propagation of uncertainty in diffusion-weighted MR imaging,” *Magn. Reson. Med.*, vol. 50, no. 5, pp. 1077–1088, 2003, doi: 10.1002/mrm.10609.
- [63] J.-D. Tournier, F. Calamante, D. G. Gadian, and A. Connelly, “Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution,” *NeuroImage*, vol. 23, no. 3, pp. 1176–1185, Nov. 2004, doi: 10.1016/j.neuroimage.2004.07.037.
- [64] A. W. Anderson, “Measurement of fiber orientation distributions using high angular resolution diffusion imaging,” *Magn. Reson. Med.*, vol. 54, no. 5, pp. 1194–1206, 2005, doi: 10.1002/mrm.20667.
- [65] W. D. Penny, K. J. Friston, J. T. Ashburner, S. J. Kiebel, and T. E. Nichols, *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Elsevier, 2011.

- [66] K. J. Friston, “Functional and Effective Connectivity: A Review,” *Brain Connect.*, vol. 1, no. 1, pp. 13–36, Jan. 2011, doi: 10.1089/brain.2011.0008.
- [67] S.-G. Kim and S. Ogawa, “Biophysical and Physiological Origins of Blood Oxygenation Level-Dependent fMRI Signals,” *J. Cereb. Blood Flow Metab.*, vol. 32, no. 7, pp. 1188–1206, Jul. 2012, doi: 10.1038/jcbfm.2012.23.
- [68] E. Amaro and G. J. Barker, “Study design in fMRI: Basic principles,” *Brain Cogn.*, vol. 60, no. 3, pp. 220–232, Apr. 2006, doi: 10.1016/j.bandc.2005.11.009.
- [69] S. C. Strother, “Evaluating fMRI preprocessing pipelines,” *IEEE Eng. Med. Biol. Mag.*, vol. 25, no. 2, pp. 27–41, Mar. 2006, doi: 10.1109/MEMB.2006.1607667.
- [70] B. Biswal, F. Zerrin Yetkin, V. M. Haughton, and J. S. Hyde, “Functional connectivity in the motor cortex of resting human brain using echo-planar mri,” *Magn. Reson. Med.*, vol. 34, no. 4, pp. 537–541, 1995, doi: 10.1002/mrm.1910340409.
- [71] M. P. van den Heuvel and H. E. Hulshoff Pol, “Exploring the brain network: A review on resting-state fMRI functional connectivity,” *Eur. Neuropsychopharmacol.*, vol. 20, no. 8, Art. no. 8, Aug. 2010, doi: 10.1016/j.euroneuro.2010.03.008.
- [72] S. M. Smith *et al.*, “Network modelling methods for FMRI,” *NeuroImage*, vol. 54, no. 2, pp. 875–891, Jan. 2011, doi: 10.1016/j.neuroimage.2010.08.063.
- [73] S. Arslan, S. I. Ktena, A. Makropoulos, E. C. Robinson, D. Rueckert, and S. Parisot, “Human brain mapping: A systematic comparison of parcellation methods for the human cerebral cortex,” *NeuroImage*, vol. 170, pp. 5–30, Apr. 2018, doi: 10.1016/j.neuroimage.2017.04.014.
- [74] S. B. Eickhoff, B. Thirion, G. Varoquaux, and D. Bzdok, “Connectivity-based parcellation: Critique and implications,” *Hum. Brain Mapp.*, vol. 36, no. 12, pp. 4771–4792, 2015, doi: 10.1002/hbm.22933.
- [75] O. Ciccarelli, M. Catani, H. Johansen-Berg, C. Clark, and A. Thompson, “Diffusion-based tractography in neurological disorders: concepts, applications, and future developments,” *Lancet Neurol.*, vol. 7, no. 8, pp. 715–727, Aug. 2008, doi: 10.1016/S1474-4422(08)70163-7.
- [76] B. Jeurissen, M. Descoteaux, S. Mori, and A. Leemans, “Diffusion MRI fiber tractography of the brain,” p. 22.

- [77] S. Jbabdi and H. Johansen-Berg, “Tractography: Where Do We Go from Here?,” *Brain Connect.*, vol. 1, no. 3, pp. 169–183, Sep. 2011, doi: 10.1089/brain.2011.0033.
- [78] A. U. Patil, S. Ghate, D. Madathil, O. J. L. Tzeng, H.-W. Huang, and C.-M. Huang, “Static and dynamic functional connectivity supports the configuration of brain networks associated with creative cognition,” *Sci. Rep.*, vol. 11, no. 1, Art. no. 1, Jan. 2021, doi: 10.1038/s41598-020-80293-2.
- [79] J. D. Power, K. A. Barnes, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen, “Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion,” *NeuroImage*, vol. 59, no. 3, pp. 2142–2154, Feb. 2012, doi: 10.1016/j.neuroimage.2011.10.018.
- [80] R. Hindriks *et al.*, “Can sliding-window correlations reveal dynamic functional connectivity in resting-state fMRI?,” *NeuroImage*, vol. 127, pp. 242–256, Feb. 2016, doi: 10.1016/j.neuroimage.2015.11.055.
- [81] N. Leonardi and D. Van De Ville, “On spurious and real fluctuations of dynamic functional connectivity during rest,” *NeuroImage*, vol. 104, pp. 430–436, Jan. 2015, doi: 10.1016/j.neuroimage.2014.09.007.
- [82] M. A. Lindquist, Y. Xu, M. B. Nebel, and B. S. Caffo, “Evaluating dynamic bivariate correlations in resting-state fMRI: A comparison study and a new approach,” *NeuroImage*, vol. 101, pp. 531–546, Nov. 2014, doi: 10.1016/j.neuroimage.2014.06.052.
- [83] D. J. Watts and S. H. Strogatz, “Collective dynamics of ‘small-world’ networks,” *Nature*, vol. 393, no. 6684, Art. no. 6684, Jun. 1998, doi: 10.1038/30918.
- [84] A. L. Barabasi and R. Albert, “Emergence of scaling in random networks,” *Science*, vol. 286, no. 5439, pp. 509–512, Oct. 1999, doi: 10.1126/science.286.5439.509.
- [85] O. Sporns, G. Tononi, and R. Kötter, “The Human Connectome: A Structural Description of the Human Brain,” *PLOS Comput. Biol.*, vol. 1, no. 4, p. e42, set 2005, doi: 10.1371/journal.pcbi.0010042.
- [86] E. Bullmore and O. Sporns, “Complex brain networks: graph theoretical analysis of structural and functional systems,” *Nat. Rev. Neurosci.*, vol. 10, no. 3, pp. 186–198, Mar. 2009, doi: 10.1038/nrn2575.

- [87] M. Rubinov and O. Sporns, “Complex network measures of brain connectivity: Uses and interpretations,” *NeuroImage*, vol. 52, no. 3, Art. no. 3, Sep. 2010, doi: 10.1016/j.neuroimage.2009.10.003.
- [88] M. Catani, M. Thiebaut de Schotten, D. Slater, and F. Dell’Acqua, “Connectomic approaches before the connectome,” *NeuroImage*, vol. 80, pp. 2–13, Oct. 2013, doi: 10.1016/j.neuroimage.2013.05.109.
- [89] K. Amunts, C. Ebell, J. Muller, M. Telefont, A. Knoll, and T. Lippert, “The Human Brain Project: Creating a European Research Infrastructure to Decode the Human Brain,” *Neuron*, vol. 92, no. 3, pp. 574–581, Nov. 2016, doi: 10.1016/j.neuron.2016.10.046.
- [90] L. A. Jorgenson *et al.*, “The BRAIN Initiative: developing technology to catalyse neuroscience discovery,” *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 370, no. 1668, p. 20140164, May 2015, doi: 10.1098/rstb.2014.0164.
- [91] W. L. Nowinski, “Human brain atlas: past, present and future,” *Neuroradiol. J.*, vol. 30, no. 6, pp. 504–519, Dec. 2017, doi: 10.1177/1971400917739274.
- [92] K. Amunts *et al.*, “BigBrain: an ultrahigh-resolution 3D human brain model,” *science*, vol. 340, no. 6139, pp. 1472–1475, 2013.
- [93] Y. Assaf *et al.*, “The CONNECT project: Combining macro- and micro-structure,” *NeuroImage*, vol. 80, pp. 273–282, Oct. 2013, doi: 10.1016/j.neuroimage.2013.05.055.
- [94] T. Jiang, “Brainnetome: A new -ome to understand the brain and its disorders,” *NeuroImage*, vol. 80, pp. 263–272, Oct. 2013, doi: 10.1016/j.neuroimage.2013.04.002.
- [95] X.-N. Zuo, Y. He, R. F. Betzel, S. Colcombe, O. Sporns, and M. P. Milham, “Human Connectomics across the Life Span,” *Trends Cogn. Sci.*, vol. 21, no. 1, pp. 32–45, Jan. 2017, doi: 10.1016/j.tics.2016.10.005.
- [96] N. Sadato *et al.*, “Neuroethical Issues of the Brain/MINDS Project of Japan,” *Neuron*, vol. 101, no. 3, pp. 385–389, Feb. 2019, doi: 10.1016/j.neuron.2019.01.006.
- [97] M. P. van den Heuvel, S. C. de Lange, A. Zalesky, C. Seguin, B. T. T. Yeo, and R. Schmidt, “Proportional thresholding in resting-state fMRI functional connectivity networks and consequences for patient-control connectome studies: Issues and recommendations,” *NeuroImage*, vol. 152, pp. 437–449, May 2017, doi: 10.1016/j.neuroimage.2017.02.005.

- [98] A. J. Schwarz and J. McGonigle, “Negative edges and soft thresholding in complex network analysis of resting state functional connectivity data,” *NeuroImage*, vol. 55, no. 3, pp. 1132–1146, Apr. 2011, doi: 10.1016/j.neuroimage.2010.12.047.
- [99] Y. Zhou *et al.*, “Functional disintegration in paranoid schizophrenia using resting-state fMRI,” *Schizophr. Res.*, vol. 97, no. 1, pp. 194–205, Dec. 2007, doi: 10.1016/j.schres.2007.05.029.
- [100] L. Zhan *et al.*, “The significance of negative correlations in brain connectivity,” *J. Comp. Neurol.*, vol. 525, no. 15, pp. 3251–3265, 2017, doi: 10.1002/cne.24274.
- [101] O. Civier, R. E. Smith, C.-H. Yeh, A. Connelly, and F. Calamante, “Is removal of weak connections necessary for graph-theoretical analysis of dense weighted structural connectomes from diffusion MRI?,” *NeuroImage*, vol. 194, pp. 68–81, Jul. 2019, doi: 10.1016/j.neuroimage.2019.02.039.
- [102] D. S. Bassett and E. T. Bullmore, “Small-World Brain Networks Revisited,” *The Neuroscientist*, vol. 23, no. 5, pp. 499–516, Oct. 2017, doi: 10.1177/1073858416667720.
- [103] C. R. Buchanan *et al.*, “The effect of network thresholding and weighting on structural brain networks in the UK Biobank,” *NeuroImage*, vol. 211, p. 116443, May 2020, doi: 10.1016/j.neuroimage.2019.116443.
- [104] L. M. Colon-Perez, M. Couret, W. Triplett, C. C. Price, and T. H. Mareci, “Small Worldness in Dense and Weighted Connectomes,” *Front. Phys.*, vol. 4, 2016, doi: 10.3389/fphy.2016.00014.
- [105] B. C. M. van Wijk, C. J. Stam, and A. Daffertshofer, “Comparing Brain Networks of Different Size and Connectivity Density Using Graph Theory,” *PLoS ONE*, vol. 5, no. 10, p. e13701, Oct. 2010, doi: 10.1371/journal.pone.0013701.
- [106] A. Fornito, A. Zalesky, and E. T. Bullmore, *Fundamentals of brain network analysis*. Amsterdam ; Boston: Elsevier/Academic Press, 2016.
- [107] A. Fornito, J. Yoon, A. Zalesky, E. T. Bullmore, and C. S. Carter, “General and Specific Functional Connectivity Disturbances in First-Episode Schizophrenia During Cognitive Control Performance,” *Biol. Psychiatry*, vol. 70, no. 1, pp. 64–72, Jul. 2011, doi: 10.1016/j.biopsych.2011.02.019.

- [108] J. M. Sheffield and D. M. Barch, “Cognition and resting-state functional connectivity in schizophrenia,” *Neurosci. Biobehav. Rev.*, vol. 61, pp. 108–120, Feb. 2016, doi: 10.1016/j.neubiorev.2015.12.007.
- [109] P. Hagmann *et al.*, “Mapping Human Whole-Brain Structural Networks with Diffusion MRI,” *PLOS ONE*, vol. 2, no. 7, p. e597, lug 2007, doi: 10.1371/journal.pone.0000597.
- [110] M. A. de Reus and M. P. van den Heuvel, “Estimating false positives and negatives in brain networks,” *NeuroImage*, vol. 70, pp. 402–409, Apr. 2013, doi: 10.1016/j.neuroimage.2012.12.066.
- [111] J. A. Roberts, A. Perry, G. Roberts, P. B. Mitchell, and M. Breakspear, “Consistency-based thresholding of the human connectome,” *NeuroImage*, vol. 145, pp. 118–129, Jan. 2017, doi: 10.1016/j.neuroimage.2016.09.053.
- [112] R. F. Betzel, A. Griffa, P. Hagmann, and B. Mišić, “Distance-dependent consensus thresholds for generating group-representative structural brain networks,” *Netw. Neurosci.*, vol. 3, no. 2, pp. 475–496, Mar. 2019, doi: 10.1162/netn_a_00075.
- [113] P. Tewarie, E. van Dellen, A. Hillebrand, and C. J. Stam, “The minimum spanning tree: An unbiased method for brain network analysis,” *NeuroImage*, vol. 104, pp. 177–188, Jan. 2015, doi: 10.1016/j.neuroimage.2014.10.015.
- [114] C. Bordier, C. Nicolini, and A. Bifone, “Graph Analysis and Modularity of Brain Functional Connectivity Networks: Searching for the Optimal Threshold,” *Front. Neurosci.*, vol. 11, 2017, doi: 10.3389/fnins.2017.00441.
- [115] C. Nicolini, C. Bordier, and A. Bifone, “Community detection in weighted brain connectivity networks beyond the resolution limit,” *NeuroImage*, vol. 146, pp. 28–39, Feb. 2017, doi: 10.1016/j.neuroimage.2016.11.026.
- [116] C. Bordier, C. Nicolini, G. Forcellini, and A. Bifone, “Disrupted modular organization of primary sensory brain areas in schizophrenia,” *NeuroImage Clin.*, vol. 18, pp. 682–693, Jan. 2018, doi: 10.1016/j.nicl.2018.02.035.
- [117] T. Samantaray, J. Saini, and C. N. Gupta, “Subgrouping and structural brain connectivity of Parkinson’s disease – past studies and future directions,” *Neurosci. Inform.*, vol. 2, no. 4, p. 100100, Dec. 2022, doi: 10.1016/j.neuri.2022.100100.

- [118] N. Theis *et al.*, “Evaluating Network Threshold Selection for Structural and Functional Brain Connectomes”.
- [119] A. Irimia, M. C. Chambers, C. M. Torgerson, and J. D. Van Horn, “Circular representation of human cortical networks for subject and population-level connectomic visualization,” *NeuroImage*, vol. 60, no. 2, Art. no. 2, Apr. 2012, doi: 10.1016/j.neuroimage.2012.01.107.
- [120] S. Isernia, A. Pirastru, D. Massaro, M. Rovaris, A. Marchetti, and F. Baglio, “Resting-State Functional Brain Connectivity for Human Mentalizing: Biobehavioral Mechanisms of Theory of Mind in Multiple Sclerosis,” *Soc. Cogn. Affect. Neurosci.*, p. nsab120, Nov. 2021, doi: 10.1093/scan/nsab120.
- [121] C. P. Chen *et al.*, “Diagnostic classification of intrinsic functional connectivity highlights somatosensory, default mode, and visual regions in autism,” *NeuroImage Clin.*, vol. 8, pp. 238–245, Jan. 2015, doi: 10.1016/j.nicl.2015.04.002.
- [122] J. D. V. Horn, A. Irimia, C. M. Torgerson, M. C. Chambers, R. Kikinis, and A. W. Toga, “Mapping Connectivity Damage in the Case of Phineas Gage,” *PLOS ONE*, vol. 7, no. 5, p. e37454, mag 2012, doi: 10.1371/journal.pone.0037454.
- [123] M. Krzywinski *et al.*, “Circos: An information aesthetic for comparative genomics,” *Genome Res.*, vol. 19, no. 9, Art. no. 9, Sep. 2009, doi: 10.1101/gr.092759.109.
- [124] D. S. Bassett and O. Sporns, “Network neuroscience,” *Nat. Neurosci.*, vol. 20, no. 3, pp. 353–364, Mar. 2017, doi: 10.1038/nn.4502.
- [125] M. Xia, J. Wang, and Y. He, “BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics,” *PLoS ONE*, vol. 8, no. 7, p. e68910, Jul. 2013, doi: 10.1371/journal.pone.0068910.
- [126] M. Mijalkov, E. Kakaei, J. B. Pereira, E. Westman, G. Volpe, and for the A. D. N. Initiative, “BRAPH: A graph theory software for the analysis of brain connectivity,” *PLOS ONE*, vol. 12, no. 8, p. e0178798, ago 2017, doi: 10.1371/journal.pone.0178798.
- [127] M. Lazar and A. L. Alexander, “Bootstrap white matter tractography (BOOT-TRAC),” *NeuroImage*, vol. 24, no. 2, pp. 524–532, Gennaio 2005, doi: 10.1016/j.neuroimage.2004.08.050.

- [128] A. Kulesa, M. Krzywinski, P. Blainey, and N. Altman, “Sampling distributions and the bootstrap,” *Nat. Methods*, vol. 12, no. 6, pp. 477–478, Jun. 2015.
- [129] Y. R. Gel, V. Lyubchich, and L. L. Ramirez Ramirez, “Bootstrap quantification of estimation uncertainties in network degree distributions,” *Sci. Rep.*, vol. 7, no. 1, Art. no. 1, Jul. 2017, doi: 10.1038/s41598-017-05885-x.
- [130] A. Green and C. R. Shalizi, “Bootstrapping exchangeable random graphs,” *Electron. J. Stat.*, vol. 16, no. 1, pp. 1058–1095, Jan. 2022, doi: 10.1214/21-EJS1896.
- [131] V. Picheny, N. H. Kim, and R. T. Haftka, “Application of bootstrap method in conservative estimation of reliability with limited samples,” *Struct. Multidiscip. Optim.*, vol. 41, no. 2, pp. 205–217, Mar. 2010, doi: 10.1007/s00158-009-0419-8.
- [132] L. Wei, B. Jing, and H. Li, “Bootstrapping promotes the RSFC-behavior associations: An application of individual cognitive traits prediction,” *Hum. Brain Mapp.*, vol. 41, no. 9, pp. 2302–2316, 2020, doi: 10.1002/hbm.24947.
- [133] B. D. Berman *et al.*, “Levodopa modulates small-world architecture of functional brain networks in Parkinson’s disease,” *Mov. Disord.*, vol. 31, no. 11, pp. 1676–1684, 2016, doi: 10.1002/mds.26713.
- [134] A. Alexander-Bloch *et al.*, “Disrupted Modularity and Local Connectivity of Brain Functional Networks in Childhood-Onset Schizophrenia,” *Front. Syst. Neurosci.*, vol. 4, 2010, <https://www.frontiersin.org/articles/10.3389/fnsys.2010.00147>
- [135] R. Mohtasib *et al.*, “MRI biomarkers for Alzheimer’s disease: the impact of functional connectivity in the default mode network and structural connectivity between lobes on diagnostic accuracy,” *Heliyon*, vol. 8, no. 2, p. e08901, Feb. 2022, doi: 10.1016/j.heliyon.2022.e08901.
- [136] Y. Wang *et al.*, “Diagnosis and prognosis of Alzheimer’s disease using brain morphometry and white matter connectomes,” *NeuroImage Clin.*, vol. 23, p. 101859, Jan. 2019, doi: 10.1016/j.nicl.2019.101859.
- [137] H. Shahamat and M. Saniee Abadeh, “Brain MRI analysis using a deep learning based evolutionary approach,” *Neural Netw.*, vol. 126, pp. 218–234, Jun. 2020, doi: 10.1016/j.neunet.2020.03.017.

- [138] F. Li, L. Tran, K.-H. Thung, S. Ji, D. Shen, and J. Li, “A Robust Deep Model for Improved Classification of AD/MCI Patients,” *IEEE J. Biomed. Health Inform.*, vol. 19, no. 5, pp. 1610–1616, Sep. 2015, doi: 10.1109/JBHI.2015.2429556.
- [139] J. Rieke, F. Eitel, M. Weygandt, J.-D. Haynes, and K. Ritter, “Visualizing Convolutional Networks for MRI-Based Diagnosis of Alzheimer’s Disease,” in *Understanding and Interpreting Machine Learning in Medical Image Computing Applications*, D. Stoyanov, Z. Taylor, S. M. Kia, I. Oguz, M. Reyes, A. Martel, L. Maier-Hein, A. F. Marquand, E. Duchesnay, T. Löfstedt, B. Landman, M. J. Cardoso, C. A. Silva, S. Pereira, and R. Meier, Eds., in *Lecture Notes in Computer Science*. Cham: Springer International Publishing, 2018, pp. 24–31. doi: 10.1007/978-3-030-02628-8_3.
- [140] Y. Huang, J. Xu, Y. Zhou, T. Tong, X. Zhuang, and the Alzheimer’s Disease Neuroimaging Initiative (ADNI), “Diagnosis of Alzheimer’s Disease via Multi-Modality 3D Convolutional Neural Network,” *Front. Neurosci.*, vol. 13, p. 509, May 2019, doi: 10.3389/fnins.2019.00509.
- [141] V. Bordin, D. Coluzzi, M. W. Rivolta, and G. Baselli, “Explainable AI Points to White Matter Hyperintensities for Alzheimer’s Disease Identification: a Preliminary Study,” in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Jul. 2022, pp. 484–487. doi: 10.1109/EMBC48229.2022.9871306.
- [142] C. Yang, A. Rangarajan, and S. Ranka, “Visual Explanations From Deep 3D Convolutional Neural Networks for Alzheimer’s Disease Classification,” *AMIA. Annu. Symp. Proc.*, vol. 2018, pp. 1571–1580, Dec. 2018.
- [143] J. V. Shanmugam, B. Duraisamy, B. C. Simon, and P. Bhaskaran, “Alzheimer’s disease classification using pre-trained deep networks,” *Biomed. Signal Process. Control*, vol. 71, p. 103217, Jan. 2022, doi: 10.1016/j.bspc.2021.103217.
- [144] R. J. Meszlényi, K. Buza, and Z. Vidnyánszky, “Resting State fMRI Functional Connectivity-Based Classification Using a Convolutional Neural Network Architecture,” *Front. Neuroinformatics*, vol. 11, 2017, <https://www.frontiersin.org/articles/10.3389/fninf.2017.00061>

- [145] J. Kawahara *et al.*, “BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment,” *NeuroImage*, vol. 146, pp. 1038–1049, Feb. 2017, doi: 10.1016/j.neuroimage.2016.09.046.
- [146] A. Alorf and M. U. G. Khan, “Multi-label classification of Alzheimer’s disease stages from resting-state fMRI-based correlation connectivity data and deep learning,” *Comput. Biol. Med.*, vol. 151, p. 106240, Dec. 2022, doi: 10.1016/j.compbiomed.2022.106240.
- [147] J. Bae *et al.*, “Transfer learning for predicting conversion from mild cognitive impairment to dementia of Alzheimer’s type based on a three-dimensional convolutional neural network,” *Neurobiol. Aging*, vol. 99, pp. 53–64, Mar. 2021, doi: 10.1016/j.neurobiolaging.2020.12.005.
- [148] A. Essemli, E. St-Onge, M. Descoteaux, and P.-M. Jodoin, “Understanding Alzheimer disease’s structural connectivity through explainable AI,” in *Proceedings of the Third Conference on Medical Imaging with Deep Learning*, PMLR, Sep. 2020, pp. 217–229. <https://proceedings.mlr.press/v121/essemli20a.html>
- [149] K. V. Saboo *et al.*, “Deep learning identifies brain structures that predict cognition and explain heterogeneity in cognitive aging,” *NeuroImage*, vol. 251, p. 119020, May 2022, doi: 10.1016/j.neuroimage.2022.119020.
- [150] Md. S. Kamal, A. Northcote, L. Chowdhury, N. Dey, R. G. Crespo, and E. Herrera-Viedma, “Alzheimer’s Patient Analysis Using Image and Gene Expression Data and Explainable-AI to Present Associated Genes,” *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–7, 2021, doi: 10.1109/TIM.2021.3107056.
- [151] C. Abrate and F. Bonchi, “Counterfactual Graphs for Explainable Classification of Brain Networks,” in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, in KDD ’21. New York, NY, USA: Association for Computing Machinery, Agosto 2021, pp. 2495–2504. doi: 10.1145/3447548.3467154.
- [152] Y. Turkan and F. B. Tek, “Convolutional Attention Network for MRI-based Alzheimer’s Disease Classification and its Interpretability Analysis,” in *2021 6th International Conference on Computer Science and Engineering (UBMK)*, Sep. 2021, pp. 1–6. doi: 10.1109/UBMK52708.2021.9558882.

- [153] S. El-Sappagh, J. M. Alonso, S. M. R. Islam, A. M. Sultan, and K. S. Kwak, “A multilayer multimodal detection and prediction model based on explainable artificial intelligence for Alzheimer’s disease,” *Sci. Rep.*, vol. 11, no. 1, p. 2660, Jan. 2021, doi: 10.1038/s41598-021-82098-3.
- [154] A. Perotti, P. Bajardi, F. Bonchi, and A. Panisson, “GRAPHSHAP: Motif-based Explanations for Black-box Graph Classifiers.” arXiv, Feb. 17, 2022. <http://arxiv.org/abs/2202.08815>
- [155] X. Zhang, L. Han, W. Zhu, L. Sun, and D. Zhang, “An Explainable 3D Residual Self-Attention Deep Neural Network for Joint Atrophy Localization and Alzheimer’s Disease Diagnosis Using Structural MRI,” *IEEE J. Biomed. Health Inform.*, vol. 26, no. 11, pp. 5289–5297, Nov. 2022, doi: 10.1109/JBHI.2021.3066832.
- [156] M. Dyrba, A. H. Pallath, and E. N. Marzban, “Comparison of CNN Visualization Methods to Aid Model Interpretability for Detecting Alzheimer’s Disease,” in *Bildverarbeitung für die Medizin 2020*, T. Tolxdorff, T. M. Deserno, H. Handels, A. Maier, K. H. Maier-Hein, and C. Palm, Eds., in Informatik aktuell. Wiesbaden: Springer Fachmedien, 2020, pp. 307–312. doi: 10.1007/978-3-658-29267-6_68.
- [157] M. Leming, J. M. Górriz, and J. Suckling, “Ensemble Deep Learning on Large, Mixed-Site fMRI Datasets in Autism and Other Tasks,” *Int. J. Neural Syst.*, vol. 30, no. 07, p. 2050012, Jul. 2020, doi: 10.1142/S0129065720500124.
- [158] S. Srivishagan, L. Kumaralingam, K. Thanikasalam, U. A. J. Pinidiyaarachchi, and N. Ratnarajah, “Discriminative patterns of white matter changes in Alzheimer’s,” *Psychiatry Res. Neuroimaging*, vol. 328, p. 111576, Jan. 2023, doi: 10.1016/j.psychresns.2022.111576.
- [159] K. Oh, Y.-C. Chung, K. W. Kim, W.-S. Kim, and I.-S. Oh, “Classification and Visualization of Alzheimer’s Disease using Volumetric Convolutional Neural Network and Transfer Learning,” *Sci. Rep.*, vol. 9, no. 1, p. 18150, Dec. 2019, doi: 10.1038/s41598-019-54548-6.
- [160] M. Böhle, F. Eitel, M. Weygandt, and K. Ritter, “Layer-Wise Relevance Propagation for Explaining Deep Neural Network Decisions in MRI-Based Alzheimer’s Disease Classification,” *Front. Aging Neurosci.*, vol. 11, 2019, <https://www.frontiersin.org/articles/10.3389/fnagi.2019.00194>

- [161] G. Montavon, W. Samek, and K.-R. Müller, “Methods for interpreting and understanding deep neural networks,” *Digit. Signal Process.*, vol. 73, pp. 1–15, Feb. 2018, doi: 10.1016/j.dsp.2017.10.011.
- [162] E. Thibeau-Sutre, S. Collin, N. Burgos, and O. Colliot, “Interpretability of Machine Learning Methods Applied to Neuroimaging.” arXiv, Apr. 14, 2022. <http://arxiv.org/abs/2204.07005>
- [163] C. O. Johnson *et al.*, “Global, regional, and national burden of stroke, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016,” *Lancet Neurol.*, vol. 18, no. 5, pp. 439–458, May 2019, doi: 10.1016/S1474-4422(19)30034-1.
- [164] S. Rajsic *et al.*, “Economic burden of stroke: a systematic review on post-stroke care,” *Eur. J. Health Econ.*, vol. 20, no. 1, pp. 107–134, Feb. 2019, doi: 10.1007/s10198-018-0984-0.
- [165] M. B. Jirjis, A. Vedantam, M. D. Budde, B. Kalinosky, S. N. Kurpad, and B. D. Schmit, “Severity of spinal cord injury influences diffusion tensor imaging of the brain,” *J. Magn. Reson. Imaging*, vol. 43, no. 1, pp. 63–74, 2016, doi: 10.1002/jmri.24964.
- [166] L. E. Wang *et al.*, “Degeneration of corpus callosum and recovery of motor function after stroke: A multimodal magnetic resonance imaging study,” *Hum. Brain Mapp.*, vol. 33, no. 12, pp. 2941–2956, 2012, doi: 10.1002/hbm.21417.
- [167] M. R. Sotelo, B. T. Kalinosky, K. Goodfriend, A. S. Hyingstrom, and B. D. Schmit, “Indirect Structural Connectivity Identifies Changes in Brain Networks After Stroke,” *Brain Connect.*, vol. 10, no. 8, pp. 399–410, Oct. 2020, doi: 10.1089/brain.2019.0725.
- [168] W. H. Sommer *et al.*, “Crossed cerebellar diaschisis in patients with acute middle cerebral artery infarction: Occurrence and perfusion characteristics,” *J. Cereb. Blood Flow Metab.*, vol. 36, no. 4, pp. 743–754, Apr. 2016, doi: 10.1177/0271678X15617953.
- [169] W. Li, Y. Li, W. Zhu, and X. Chen, “Changes in brain functional network connectivity after stroke,” *Neural Regen. Res.*, vol. 9, no. 1, pp. 51–60, Jan. 2014, doi: 10.4103/1673-5374.125330.
- [170] M.-L. Hu *et al.*, “A Review of the Functional and Anatomical Default Mode Network in Schizophrenia,” *Neurosci. Bull.*, vol. 33, no. 1, pp. 73–84, Feb. 2017, doi: 10.1007/s12264-016-0090-1.

- [171] D. Godwin, A. Ji, S. Kandala, and D. Mamah, “Functional Connectivity of Cognitive Brain Networks in Schizophrenia during a Working Memory Task,” *Front. Psychiatry*, vol. 8, 2017, <https://www.frontiersin.org/articles/10.3389/fpsy.2017.00294>
- [172] Y. Zhou *et al.*, “Altered intrinsic and extrinsic connectivity in schizophrenia,” *NeuroImage Clin.*, vol. 17, pp. 704–716, Jan. 2018, doi: 10.1016/j.nicl.2017.12.006.
- [173] Z. Cui, Z. Gao, J. Leng, T. Zhang, P. Quan, and W. Zhao, “Alzheimer’s Disease Diagnosis Using Enhanced Inception Network Based on Brain Magnetic Resonance Image,” in *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Nov. 2019, pp. 2324–2330. doi: 10.1109/BIBM47256.2019.8983046.
- [174] S. F. Eskildsen, P. Coupé, V. S. Fonov, J. C. Pruessner, and D. L. Collins, “Structural imaging biomarkers of Alzheimer’s disease: predicting disease progression,” *Neurobiol. Aging*, vol. 36, pp. S23–S31, Jan. 2015, doi: 10.1016/j.neurobiolaging.2014.04.034.
- [175] A. M. Hall, R. Y. Moore, O. L. Lopez, L. Kuller, and J. T. Becker, “Basal forebrain atrophy is a presymptomatic marker for Alzheimer’s disease,” *Alzheimers Dement.*, vol. 4, no. 4, pp. 271–279, 2008, doi: 10.1016/j.jalz.2008.04.005.
- [176] L. Clerx, P. J. Visser, F. Verhey, and P. Aalten, “New MRI Markers for Alzheimer’s Disease: A Meta-Analysis of Diffusion Tensor Imaging and a Comparison with Medial Temporal Lobe Measurements,” *J. Alzheimers Dis.*, vol. 29, no. 2, pp. 405–429, Jan. 2012, doi: 10.3233/JAD-2011-110797.
- [177] H. Braak and E. Braak, “Development of Alzheimer-related neurofibrillary changes in the neocortex inversely recapitulates cortical myelogenesis,” *Acta Neuropathol. (Berl.)*, vol. 92, no. 2, pp. 197–201, Jul. 1996, doi: 10.1007/s004010050508.
- [178] E. M. Reiman and W. J. Jagust, “Brain imaging in the study of Alzheimer’s disease,” *NeuroImage*, vol. 61, no. 2, pp. 505–516, Jun. 2012, doi: 10.1016/j.neuroimage.2011.11.075.
- [179] S. M. Nestor *et al.*, “Ventricular enlargement as a possible measure of Alzheimer’s disease progression validated using the Alzheimer’s disease neuroimaging initiative database,” *Brain*, vol. 131, no. 9, pp. 2443–2454, Aug. 2008, doi: 10.1093/brain/awn146.

- [180] B. T. Hyman *et al.*, “National Institute on Aging–Alzheimer’s Association guidelines for the neuropathologic assessment of Alzheimer’s disease,” *Alzheimers Dement.*, vol. 8, no. 1, pp. 1–13, Jan. 2012, doi: 10.1016/j.jalz.2011.10.007.
- [181] M. W. Weiner *et al.*, “The Alzheimer’s Disease Neuroimaging Initiative: A review of papers published since its inception,” *Alzheimers Dement.*, vol. 9, no. 5, pp. e111–e194, Sep. 2013, doi: 10.1016/j.jalz.2013.05.1769.
- [182] Y. Sun *et al.*, “Disrupted Functional Brain Connectivity and Its Association to Structural Connectivity in Amnesic Mild Cognitive Impairment and Alzheimer’s Disease,” *PLOS ONE*, vol. 9, no. 5, p. e96505, mag 2014, doi: 10.1371/journal.pone.0096505.
- [183] P. M. Matthews, N. Filippini, and G. Douaud, “Brain Structural and Functional Connectivity and the Progression of Neuropathology in Alzheimer’s Disease,” *J. Alzheimers Dis.*, vol. 33, no. s1, pp. S163–S172, Jan. 2013, doi: 10.3233/JAD-2012-129012.
- [184] J. Zimmermann *et al.*, “Differentiation of Alzheimer’s disease based on local and global parameters in personalized Virtual Brain models,” *NeuroImage Clin.*, vol. 19, pp. 240–251, Jan. 2018, doi: 10.1016/j.nicl.2018.04.017.
- [185] C. A. Brown, Y. Jiang, C. D. Smith, and B. T. Gold, “Age and Alzheimer’s pathology disrupt default mode network functioning via alterations in white matter microstructure but not hyperintensities,” *Cortex*, vol. 104, pp. 58–74, Jul. 2018, doi: 10.1016/j.cortex.2018.04.006.
- [186] D. Xiao, K. Wang, L. Theriault, E. Charbel, and A. D. N. Initiative, “White matter integrity and key structures affected in Alzheimer’s disease characterized by diffusion tensor imaging,” *Eur. J. Neurosci.*, vol. n/a, no. n/a, Jan. 2022, doi: 10.1111/ejn.15815.
- [187] Y. Zhang *et al.*, “White matter damage in frontotemporal dementia and Alzheimer’s disease measured by diffusion MRI,” *Brain*, vol. 132, no. 9, pp. 2579–2592, Sep. 2009, doi: 10.1093/brain/awp071.
- [188] A. Hafkemeijer, J. van der Grond, and S. A. R. B. Rombouts, “Imaging the default mode network in aging and dementia,” *Biochim. Biophys. Acta BBA - Mol. Basis Dis.*, vol. 1822, no. 3, pp. 431–441, Mar. 2012, doi: 10.1016/j.bbadis.2011.07.008.

- [189] M. R. Brier *et al.*, “Loss of Intranetwork and Internetwork Resting State Functional Connections with Alzheimer’s Disease Progression,” *J. Neurosci.*, vol. 32, no. 26, pp. 8890–8899, Jun. 2012, doi: 10.1523/JNEUROSCI.5698-11.2012.
- [190] B. Zhou *et al.*, “Structural and functional connectivity abnormalities of the default mode network in patients with Alzheimer’s disease and mild cognitive impairment within two independent datasets,” *Methods*, vol. 205, pp. 29–38, 2022, doi: 10.1016/j.ymeth.2022.06.001.
- [191] Y. Sun, S. L. Collinson, J. Suckling, and K. Sim, “Dynamic Reorganization of Functional Connectivity Reveals Abnormal Temporal Efficiency in Schizophrenia,” *Schizophr. Bull.*, vol. 45, no. 3, pp. 659–669, Apr. 2019, doi: 10.1093/schbul/sby077.
- [192] P. G. Unschuld *et al.*, “Prefrontal Brain Network Connectivity Indicates Degree of Both Schizophrenia Risk and Cognitive Dysfunction,” *Schizophr. Bull.*, vol. 40, no. 3, pp. 653–664, May 2014, doi: 10.1093/schbul/sbt077.
- [193] H. Kurashige *et al.*, “Revealing Relationships Among Cognitive Functions Using Functional Connectivity and a Large-Scale Meta-Analysis Database,” *Front. Hum. Neurosci.*, vol. 13, 2020, <https://www.frontiersin.org/articles/10.3389/fnhum.2019.00457>
- [194] A. Fornito, A. Zalesky, and M. Breakspear, “Graph analysis of the human connectome: Promise, progress, and pitfalls,” *NeuroImage*, vol. 80, pp. 426–444, Oct. 2013, doi: 10.1016/j.neuroimage.2013.04.087.
- [195] J. Hopcroft and R. Tarjan, “Algorithm 447: efficient algorithms for graph manipulation,” *Commun. ACM*, vol. 16, no. 6, pp. 372–378, Giugno 1973, doi: 10.1145/362248.362272.
- [196] A. K. Wagner and G. Sowa, “Rehabilomics research: a model for translational rehabilitation and comparative effectiveness rehabilitation research,” *Am. J. Phys. Med. Rehabil.*, vol. 93, no. 10, pp. 913–916, Oct. 2014, doi: 10.1097/PHM.000000000000114.
- [197] P. Hagmann *et al.*, “Mapping the Structural Core of Human Cerebral Cortex,” *PLOS Biol.*, vol. 6, no. 7, p. e159, lug 2008, doi: 10.1371/journal.pbio.0060159.
- [198] S. Guo, X. Chen, Y. Liu, R. Kang, T. Liu, and D. Li, “Percolation Analysis of Brain Structural Network,” *Front. Phys.*, vol. 9, 2021, <https://www.frontiersin.org/articles/10.3389/fphy.2021.698077>

- [199] K. Aggarwal, M. M. Jimeno, K. S. Ravi, and G. Gonzalez, “Developing and deploying deep learning models in brain MRI: a review,” Mar. 2023.
- [200] J. M. Wardlaw, M. C. Valdés Hernández, and S. Muñoz-Maniega, “What are White Matter Hyperintensities Made of?: Relevance to Vascular Cognitive Impairment,” *J. Am. Heart Assoc.*, vol. 4, no. 6, Jun. 2015, doi: 10.1161/JAHA.114.001140.
- [201] A. Kuceyeski, J. Maruta, N. Relkin, and A. Raj, “The Network Modification (NeMo) Tool: Elucidating the Effect of White Matter Integrity Changes on Cortical and Subcortical Structural Connectivity,” *Brain Connect.*, vol. 3, no. 5, pp. 451–463, Oct. 2013, doi: 10.1089/brain.2013.0147.