Algorithmic Characterisation of Non-Rigid Registration in Inter-Subject Resting-State Functional Magnetic Resonance Image Processing

BIRMINGHAM CITY
University

Martin Svejda

Supervised by:
Dr Nouh Sabri Elmitwally
Dr Roger Tait

Computing Engineering & the Built Environment

A thesis submitted in partial fulfilment of the requirements of Birmingham City University for the degree of
Doctor of Philosophy

December 2024

Abstract

Resting-State Functional Magnetic Resonance Imaging (rs-fMRI) is fundamental for studying intrinsic brain functions, crucial for defining the networks underlying human cognition and behaviour. Non-rigid registration algorithms are essential for accurately aligning rs-fMRI data across subjects, a process critical for consistent and reliable analysis of functional connectivity. The performance of these algorithms directly impacts the precision of neuroimaging results due to individual anatomical differences.

This thesis addresses the critical issue of performance variability among non-rigid registration algorithms, which can undermine the reliability and accuracy of functional connectivity analysis in rs-fMRI. To systematically assess these differences, the Non-Rigid Registration Algorithm Analysis Framework (NRAAF) was developed and implemented, offering an innovative benchmark for evaluating and characterising the accuracy and specificity of various algorithms.

Key findings show that algorithms such as Advanced Normalisation Tools (ANTs), Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL), Analysis of Functional NeuroImages (AFNI), and FMRIB Software Library (FSL) exhibit significant differences in handling anatomical variability. ANTs demonstrated superior sensitivity with a mean Peak Activation Intensity of 0.85, while DARTEL showed the most consistent performance with minimal variability (Standard Deviation of 0.05). AFNI presented a higher variance in cluster detection at 0.30, compared to FSL's 0.18. These insights emphasise that algorithm selection crucially influences the reliability of functional connectivity analyses.

The differential performance among these algorithms significantly impacts neuroimaging outcomes, affecting both the interpretation of research findings and potential clinical applications. By providing a comprehensive evaluation and characterisation of non-rigid registration algorithms, this work emphasises the importance of selecting appropriate methods to enhance reproducibility and accuracy in neuroimaging. In doing so, NRAAF framework empowers the neuroimaging community to advance computational methodologies and refine tools for

studying complex brain functions, with potential implications for diagnostics, personalised treatment strategies, and broader cross-institutional research collaborations.

Declaration of Authorship

I, Martin Svejda, declare that this thesis titled, "Algorithmic Characterisation of Non-Rigid Registration in Inter-Subject Resting-State Functional Magnetic Resonance Image Processing" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Martin Svejda
December 2024

List of Publications

- Svejda, M., & Tait, R. (2019). Characterisation of VBM Algorithms for Processing of Medical MRI Images. In *Artificial Intelligence XXXVI: 39th SGAI International Conference on Artificial Intelligence, AI 2019, Cambridge, UK, December 17–19, 2019, Proceedings 39* (pp. 443-448). Springer International Publishing. (Status: *Published*)
- Svejda, M., Elmitwally, N. S., Asyhari, A. T., & Tait, R. (2024). NRAAF: A Framework for Comparative Analysis of fMRI Registration Algorithms and Their Impact on Resting-State Neuroimaging Accuracy. In *IEEE Access*, vol. 12 (pp. 47915-47941). (Status: *Published*)

Acknowledgements

I would like to express my deepest gratitude to my supervisors Dr Nouh Elmitwally, Dr Taufiq Asyhari, and Dr Roger Tait whose insights and guidance were invaluable throughout this research. Thanks are also due to my mentors, Rehan Bhana, Professor Mohamed Gaber, Dr Andrew Wilson, and Shaheed Parvez for their constructive feedback and kind words of support.

Gratitude extends not only to the profound contributions of my supervisors and mentors, but also to their unwavering support and comments of encouragement, which far exceeded the scope of this research. Their guidance has been instrumental in shaping my journey, and I am profoundly thankful for your enduring commitment.

I would also like to thank the thesis reviewers, Dr Faisal Saeed, Dr Fei He, and Professor Michal Krzyzanowski for their time and thoughtful, constructive criticisms, which have contributed to improving this work.

Also, I must thank my dear parents for their constant encouragements and unwavering support.

Special thanks to my colleagues in the College of Computing for their encouragement and for all the stimulating discussions that helped refine my ideas.

I am also grateful to Computer & Physical Systems for providing the resources necessary to carry out this research.

Lastly, I wish to acknowledge my family, friends and colleagues for their understanding, ideas, informal discussions and endless support throughout my academic journey.

Thank you all kindly for your help!

Table of contents

List of figures

List of tables

Nomenclature

Greek Symbols

α	Alpha, significance level in statistical hypothesis testing, often set at 0.05.
β	Beta, coefficients in regression models or the Type II error rate in hypothesis testing.
δ	Delta, change or difference in mathematical expressions, often representing a small increment.
ε	Epsilon, small positive quantity, used in limits and approximations.
η	Eta, measure of effect size or efficiency, particularly in statistical contexts.
γ	Gamma, parameter in statistical distributions or in machine learning algorithms, such as in SVM.
κ	Kappa, measure of agreement or a curvature parameter in statistics.
λ	Lambda, denotes regularisation parameters in statistical models or eigenvalues in multivariate analyses.
μ	Mu, mean of a distribution, often used to represent the average in data analysis.
ϕ	Phi, represents a phase angle in signal processing or scalar potential in mathematical models.
ρ	Rho, the Spearman's rank correlation coefficient, measuring association strength between ranked variables.
σ	Sigma, represents the standard deviation of a distribution, critical for assessing data variability.
θ	Theta, parameters in various statistical models.

Nomenclature

Acronyms / Abbreviations

ADNI Alzheimer's Disease Neuroimaging Initiative

AFNI Analysis of Functional NeuroImages

AIC Akaike Information Criterion

ANTs Advanced Normalisation Tools

AOMIC Amsterdam Open MRI Collection

AR Augmented Reality

BET Brain Extraction Tool

BIC Bayesian Information Criterion

BOLD Blood-Oxygenation-Level-Dependent

CBCT Cone Beam Computed Tomography

CDR Cognitive Decline Rating

CNN Convolutional Neural Network

CRF Conditional Random Field

CSF Cerebrospinal Fluid

CT Computed Tomography

DARTEL Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra

DBSCAN Density-Based Spatial Clustering of Applications with Noise

DBS Deep Brain Stimulation

DL Deep Learning

DMN Default Mode Network

DSC Dice Similarity Coefficient

DTI Diffusion Tensor Imaging

DWI Diffusion Weighted Imaging

EEG Electroencephalography

EPI Echo Planar Imaging

FA Fractional Anisotropy

FDA Food and Drug Administration

FEAT FMRI Expert Analysis Tool

FLAIR Fluid Attenuated Inversion Recovery

FLIRT FMRIB's Linear Image Registration Tool

fMRI Functional Magnetic Resonance Imaging

rs-fMRI Resting-State Functional Magnetic Resonance Imaging

FNIRT FMRIB's Nonlinear Image Registration Tool

FSL FMRIB Software Library

GAN Generative Adversarial Network

GLM General Linear Model

GMM Gaussian Mixture Model

GNN Graph Neural Network

GRF Gaussian Random Field

HRF Hemodynamic Response Function

ICA Independent Component Analysis

ICC Intraclass Correlation Coefficient

IQR Interquartile Range

LDDMM Large Deformation Diffeomorphic Metric Mapping

MATLAB Matrix Laboratory

MCFLIRT Motion Correction FMRIB's Linear Image Registration Tool

MD Mean Diffusivity

MI Mutual Information

MNI152 Montreal Neurological Institute 152

MNI Montreal Neurological Institute

MPFC Medial Prefrontal Cortex

MPRAGE Magnetisation Prepared Rapid Gradient Echo

MRI Magnetic Resonance Imaging

MSE Mean Square Error

MVPA Multivoxel Pattern Analysis

NCC Normalised Cross-Correlation

NIFTI Neuroimaging Informatics Technology Initiative

NMI Normalised Mutual Information

NRAAF Non-Rigid Registration Algorithm Analysis Framework

OPTICS Ordering Points to Identify the Clustering Structure

PCA Principal Component Analysis

PCC Posterior Cingulate Cortex

PET Positron Emission Tomography

PINN Physics-Informed Neural Network

PVE Partial Volume Effect

QSM Quantitative Susceptibility Mapping

RAS Right-Anterior-Superior

RF Random Forest

RL Reinforcement Learning

ROI Region of Interest

RSN Resting-State Network

SLURM Simple Linux Utility for Resource Management

SNR Signal-to-Noise Ratio

SPECT Single Photon Emission Computed Tomography

SPM Statistical Parametric Mapping

SSD Sum of Squared Differences

SVM Support Vector Machine

SyN Symmetric Normalisation

T1w T1-weighted

T2w T2-weighted

TRE Target Registration Error

TR Repetition Time

TRUS Transrectal Ultrasound

US Ultrasound

VBM Voxel-Based Morphometry

VR Virtual Reality

Chapter 1

Introduction

Image registration is the process of aligning multiple images into a common coordinate system, allowing for the comparison and analysis of data from different imaging sessions, modalities, or individuals. In medical imaging, this alignment is crucial for accurate diagnosis, monitoring disease progression, and evaluating treatment efficacy. Image registration helps align anatomical or functional images, ensuring that meaningful comparisons can be drawn, whether from different subjects, different times, or different imaging modalities [1, 2].

Historically, image registration methods have evolved significantly, from manual techniques to sophisticated algorithms, driven largely by advances in computational power and artificial intelligence. Early efforts focused on manual and semi-automated approaches that required substantial human expertise, which were later replaced by automated methods, including feature-based and intensity-based techniques. In recent years, Deep Learning (DL) has further revolutionised the field, enabling highly efficient and accurate registration, especially in complex cases involving deformable transformations [3, 4]. For example, Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs) have been leveraged to tackle the challenges of multi-modal and non-rigid registration, offering new possibilities in both medical and non-medical contexts [5, 6].

In neuroimaging, the importance of image registration cannot be overstated, as the accurate alignment of brain scans is essential for understanding brain function and pathology. For example, precise registration allows the comparison of Functional Magnetic Resonance Images (fMRI) data across participants, which is essential in identifying Resting-State Networks (RSNs) and understanding their role in both healthy and diseased states [7, 8]. Non-rigid registration, in particular, plays a critical role in this process by accommodating the complex anatomical differences between subjects, thus allowing for more accurate mapping of functional and structural data across diverse populations.

Introduction

The applications of image registration are diverse and extend well beyond medical imaging. In neuroimaging, image registration plays a pivotal role. It ensures that functional brain data, such as Resting-State Functional Magnetic Resonance Imaging (rs-fMRI), are accurately aligned across different subjects or time points. This enables robust group-level analyses that reveal meaningful patterns in brain function. Such precise registration is crucial for studying neurological diseases, understanding brain development, and evaluating the effects of various therapeutic interventions [9, 7]. Moreover, the applicability of image registration extends beyond the field of medical imaging. Below, we provide brief examples of this.

Image registration is also used in other fields:

- **Environmental Science:** Image registration enhances environmental monitoring by aligning satellite or aerial images over time, which is crucial for tracking environmental changes such as deforestation, urban expansion, climate impacts, and post-disaster damage assessments. This enables precise comparisons and data integration from multiple temporal and spatial sources, supporting more informed decision-making [10, 11].

- **Autonomous Vehicles:** In autonomous navigation, image registration is used for fusing data from cameras, LIDAR, and other sensors, allowing the vehicle to understand its surroundings more accurately. This precise alignment is essential for obstacle detection, localisation, and path planning, ultimately improving the safety, efficiency, and reliability of autonomous driving systems [12, 13].

- **Augmented and Virtual Reality (AR/VR):** In AR/VR, image registration allows the seamless overlay of virtual content onto the real world by accurately aligning the digital visuals with the user's real-time perspective. This technology enhances interactive experiences in gaming, education, and professional training, as well as improves precision in medical training and surgical planning, leading to better outcomes and deeper learning experiences [14, 15].

The versatility of image registration across different fields highlights its essential role as a foundational technology, driving advancements in healthcare and broader societal contexts.

The implementation of advanced image registration methods in clinical settings requires thorough validation processes to ensure reliability and efficacy. This involves not only extensive empirical testing across a wide range of clinical scenarios but also comparison

against established standards, such as those set by the Food and Drug Administration (FDA) or equivalent regulatory bodies. Clinical validation is crucial because it ensures that these technologies meet the strict accuracy and safety standards necessary for patient care, ultimately influencing diagnostic decisions and treatment outcomes [16, 2].

Validation processes help uphold patient safety by ensuring that technological interventions yield consistent and accurate results. For example, validation in neuroimaging aims to verify that brain structures can be consistently aligned across subjects, directly impacting the interpretation of functional connectivity and subsequent clinical decisions for neurological disorders like Multiple Sclerosis (MS) or Alzheimer's disease.

The image registration process is illustrated in Figure 1.1. In this model, the reference image I_R acts as a fixed basis, against which the moving image I_M is aligned. The registration algorithm processes these images and produces a transformed version of the moving image, denoted as I_{TM}. The goal of this transformation is to align the moving image as closely as possible with the reference image, ensuring that corresponding anatomical or functional structures overlap accurately.

The registration model depicted in Figure 1.1 is a general framework that can represent both rigid and non-rigid registration methods. In rigid registration, the transformation applied is limited to translations and rotations, maintaining the shape and size of anatomical structures. This type of registration is suitable for aligning images where only global differences exist, such as different poses of the same patient. Non-rigid registration, on the other hand, involves more complex transformations that allow for deformations, adapting to the unique anatomical variations between individuals. This flexibility is crucial in neuroimaging for mapping subtle differences in brain structure and function across subjects, particularly when studying rs-fMRI [17].

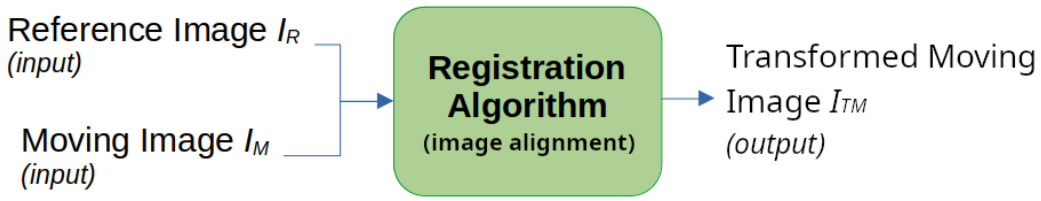

Fig. 1.1 General model of the image registration process, showing the transformation of a moving image I_M to align with a reference image I_R. This model can represent both rigid and non-rigid registration methods, depending on the complexity of the transformation applied.

The significance of non-rigid registration in neuroimaging cannot be overstated, as it ensures the precise mapping of brain structures and activity across temporal and spatial dimensions. This precision is essential for understanding nuanced changes that occur in

neurological diseases, monitoring brain development, and tracking the effects of therapeutic interventions over time [7, 8]. By enabling accurate alignment that accounts for individual differences, non-rigid registration contributes to the reliability of group-level analyses and to the understanding of complex brain functions.

The upcoming chapters will outline the specific challenges that image registration addresses in computational neuroscience, particularly in the context of fMRI. This thesis examines the accuracy and effectiveness of inter-subject non-rigid registration, especially for RSNs. This evaluation enhances our understanding of how various registration algorithms influence the interpretation of functional connectivity, highlighting their broader impact on diagnosing and treating neurological conditions like Alzheimer's disease and MS [17, 9].

1.1 Rigid and Non-Rigid Image Registration

Rigid registration involves aligning images by applying transformations such as rotation and translation, which effectively allow for the repositioning of images without altering their shape or size. This type of registration assumes that the underlying anatomy does not change in shape, making it suitable for scenarios where global differences exist but local deformities are not a concern, such as different orientations of the same subject.

The primary advantage of rigid registration lies in its computational efficiency—the transformation process is relatively straightforward, often requiring fewer parameters compared to non-rigid methods. This makes rigid registration particularly well-suited for initial alignments or for aligning images where major deformations are absent, such as registering skull images or hard tissue scans [16].

However, rigid registration has limitations due to its inability to accommodate anatomical variability among individuals. This poses challenges when dealing with soft tissue regions like the brain, where complex deformations between subjects are common. Such limitations emphasise the necessity for more advanced approaches, such as non-rigid registration, which can provide the flexibility required for accurate alignment of neuroanatomical structures.

In contrast, non-rigid registration (also known as deformable registration) allows for complex deformations that can accurately align anatomical regions with significant inter-subject variability. This type of registration is essential in neuroimaging, where structural variations between individuals require a more flexible approach.

Non-rigid registration works by transforming the moving image using a deformation field, which adapts to the local anatomical differences present across individuals. This flexibility makes it indispensable in rs-fMRI and other neuroimaging modalities that require the precise

alignment of cortical and subcortical regions [4]. Recent advancements have incorporated Machine Learning (ML), particularly DL models, to improve the accuracy of non-rigid registration processes [18, 19].

Non-rigid registration has been crucial in enabling detailed group analyses of RSNs, thereby advancing our understanding of the common and variable aspects of brain connectivity in both healthy and pathological states. Moreover, the use of DL has enhanced the efficiency of these methods, which directly impacts their feasibility for clinical use, such as monitoring disease progression or planning neurosurgical procedures.

Both rigid and non-rigid registration methods play complementary roles in medical imaging. Rigid registration is often used as a preliminary step, providing an efficient and quick alignment that serves as a foundation for more refined processes. This initial alignment is computationally inexpensive and can correct for global translations and rotations between images.

Once the base-level alignment is achieved using rigid registration, non-rigid registration is employed to address local anatomical differences, providing a finer level of adjustment that is necessary for accurate cross-subject comparisons, particularly in neuroimaging studies [9, 4].

For example, in the context of rs-fMRI, rigid registration is used to align images from the same subject taken at different times, ensuring consistent positioning. This is followed by non-rigid registration, which corrects for individual differences in brain anatomy, allowing for a precise analysis of RSNs. The combined use of both approaches enables a balance between efficiency and accuracy, ultimately ensuring that neuroimaging data are reliably aligned for group-level analyses [16].

Figure 1.2 provides a visual overview of the registration workflow used in this study, which employs both rigid and non-rigid steps. Initially, the functional image is aligned with the structural image using a Rigid transformation which involves only rotations and translations, allowing for global alignment of images without altering their internal structure. Subsequently, a Non-Rigid transformation is used to provide the flexibility needed to adjust images locally, overcoming individual anatomical differences and enhancing alignment accuracy, which is critical for accurately localising brain activity and conducting reliable group analyses.

Within the field of medical imaging, the comparison between rigid and non-rigid registration highlights key methodological concerns. Non-rigid registration is particularly tailored to the dynamic and flexible nature of biological tissues, addressing the considerable challenges associated with anatomical variability and physiological motion in the human body [4]. This

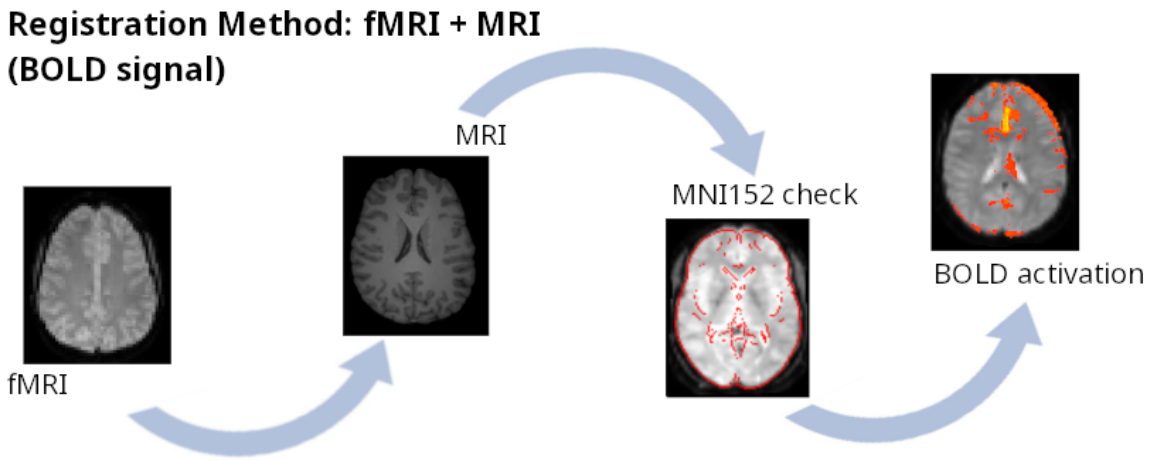

Fig. 1.2 Multimodal registration workflow used in this study, involving rigid (fMRI - MRI) and non-rigid (MRI - MNI152) registration to align the functional and structural images to a standard template. The registered image is then fitted into a General Linear Model to capture the Blood-Oxygen-Level Dependent (BOLD) signal, thus producing a map of activated brain areas.

form of registration is essential in neuroimaging, where precise alignment of brain images across different subjects or over time can significantly impact diagnostic and therapeutic outcomes. Recent theoretical advancements, notably the integration of DL techniques with traditional image registration frameworks, have considerably enhanced both the accuracy and efficiency of these applications. Such advancements are not merely technical improvements but are critical in advancing patient care through improved diagnostic accuracy and personalised treatment planning [18, 19].

1.2 Significance of rs-fMRI and Functional Connectivity

Resting-state fMRI has become an essential tool for evaluating the brain's intrinsic activity when no specific external task is being performed. By capturing spontaneous fluctuations in the Blood-Oxygen-Level Dependent (BOLD) signal, rs-fMRI provides critical insights into the brain's functional architecture and connectivity during resting states [4]. This imaging modality enables the identification of RSNs, which are spatially distributed yet temporally correlated regions that consistently show synchronous activity. Examples of such networks include the Default Mode Network (DMN), Visual Networks, and Somatomotor Networks—all of which play fundamental roles in maintaining the brain's baseline functions and are closely associated with various cognitive and sensory processing capabilities [2].

The significance of RSNs extends beyond fundamental neuroscience by providing an essential framework for understanding how different brain regions interact, both in healthy individuals and in those with neurological or psychiatric disorders. By analysing these networks, we can gain a better understanding of the functional integration occurring within the brain, which is fundamental for tasks such as memory, emotion regulation, and sensory processing [7]. This perspective is particularly important for assessing neurological disorders where deviations in these networks can serve as critical biomarkers for disease progression.

Functional connectivity, derived from rs-fMRI data, refers to the temporal correlation between spatially distinct brain regions, indicating how well different areas communicate with one another during rest. The reliability of these connectivity measures hinges heavily on the precision of the image registration process [8]. Accurate registration ensures that anatomical structures are properly aligned across scans, thus minimising inconsistencies that could obscure the genuine connectivity patterns present. In the context of non-rigid registration, the challenge is particularly pronounced because of the inherent anatomical variability between individuals, which requires advanced methods capable of handling these complex deformations without compromising the integrity of functional signals.

The application of the Non-Rigid Registration Algorithm Analysis Framework (NRAAF) developed in this thesis aims to address precisely this challenge. By providing a structured evaluation of four state-of-the-art non-rigid registration algorithms, this work seeks to enhance the precision of rs-fMRI data analysis. The findings from this systematic comparison are instrumental in identifying effective algorithms tailored for specific neuroimaging tasks. Through better algorithm selection, we can improve both the reproducibility and consistency of functional connectivity studies [20]. This reduction in variability is crucial for accurate identification of RSNs, which ultimately enhances our understanding of the brain's functional architecture [21].

Moreover, the application of Multivoxel Pattern Analysis (MVPA) with Support Vector Machine (SVM) offers a novel methodology to evaluate the impact of different registration algorithms on neuroimaging data. This approach enables the precise identification and distinction of spatial patterns, providing deeper insights into the brain's default mode networks and the broader network structure [22]. The discriminative voxel-wise weight maps generated in this study further reveal how the selection of different algorithms influences neuroimaging outcomes, supporting the development of more reliable diagnostic tools for detecting and understanding neurological conditions [23].

The findings of this thesis have substantial implications for clinical applications. Improved accuracy in image registration directly impacts the diagnosis and treatment of neurological

disorders. By ensuring that functional connectivity analyses are more precise, this research contributes to better identification of biomarkers for various conditions, including neurodegenerative diseases such as Alzheimer's and psychiatric disorders like Schizophrenia [24]. This, in turn, can lead to more personalised and effective treatment plans, ensuring that interventions are tailored to the specific network alterations observed in patients.

In summary, this thesis provides a comprehensive evaluation of non-rigid registration algorithms, emphasising their critical role in ensuring reproducibility and reliability in neuroimaging studies, particularly in the context of functional connectivity. The broader impact of this research lies in its potential to improve the accuracy of rs-fMRI analyses, thereby advancing our understanding of complex brain functions and supporting the development of effective diagnostic and therapeutic approaches for neurological disorders [18, 19].

1.3 Research Motivation and Problem Statement

The advancement of rs-fMRI has been instrumental in enabling the exploration of the brain's RSNs, which are essential for understanding both baseline brain function and its alterations in neurological disorders. However, the effectiveness of these explorations is highly dependent on the precision of image registration. Registration is a fundamental process in aligning images across different scans, individuals, and time points, yet it is also one of the primary sources of variability due to differences in brain anatomy and the dynamic nature of the BOLD signal [25].

Non-rigid registration algorithms play a pivotal role in aligning rs-fMRI data across subjects. They are tasked with compensating for anatomical variations, ensuring that functional data can be accurately compared and analysed. However, the variability in how different algorithms perform this complex task introduces significant challenges. Algorithmic variability refers to differences in how various non-rigid registration techniques handle alignment, leading to discrepancies in the identification of RSNs, the localisation of activation, and ultimately the conclusions drawn about brain function. Such variability affects not only reproducibility in research findings but also has profound implications for clinical interpretations [26, 27].

For instance, discrepancies in the registration of fMRI data can lead to differences in how activation clusters are localised, which may result in inconsistent findings regarding functional connectivity. In a clinical context, this inconsistency can impact diagnostic accuracy, as reliable identification of RSNs is crucial for understanding pathologies such as Alzheimer's disease, Schizophrenia, and other neurological conditions. Inaccurate alignment may cause

errors in assessing the extent and severity of dysfunction, affecting patient outcomes and the development of treatment plans.

Despite ongoing advancements in neuroimaging techniques, a critical gap persists in understanding how different non-rigid registration algorithms affect the processing and outcomes of rs-fMRI data analysis. This gap presents a significant barrier to fully exploiting rs-fMRI's potential to explore brain function and identify pathological changes. The issue of algorithmic variability is not merely theoretical; it has practical implications, particularly in clinical settings where fMRI data are used to guide diagnosis and treatment.

In the context of diagnosing neurological disorders, the precise alignment of fMRI data is essential for identifying deviating patterns of functional connectivity, which can be indicative of conditions like Epilepsy, MS, and traumatic brain injury. Variability in registration outcomes can lead to errors in identifying these patterns, thereby affecting clinical decision-making. For instance, an inaccurate alignment might either obscure or falsely indicate functional deficits, leading to improper diagnoses or inappropriate therapeutic interventions. Therefore, improving the reliability of non-rigid registration methods is of utmost importance, not only to enhance the robustness of research findings but also to ensure that these technologies can be effectively translated into clinical practices that improve patient care.

1.3.1 Motivation

Addressing algorithmic variability is essential for advancing neuroimaging methodologies, improving the reliability of rs-fMRI analyses, and enhancing our understanding of the brain's functional architecture in both healthy and diseased states. The motivation behind this study lies in identifying and reducing the discrepancies that arise from the use of different non-rigid registration algorithms. By doing so, we can improve the consistency of activation maps, increase the accuracy of group-level analyses, and contribute to the development of better diagnostic tools and therapeutic interventions [28, 29].

Furthermore, as neuroimaging increasingly informs precision medicine, the demand for robust and reproducible imaging results becomes critical. Misalignment and variability in data processing can hinder the progress of personalised treatments that rely on specific connectivity patterns to tailor interventions. Thus, addressing these challenges is key to ensuring that neuroimaging can reliably support both research advancements and clinical applications that improve patient outcomes [30, 31].

1.3.2 Problem Statement

This thesis aims to systematically evaluate the impact of various state-of-the-art non-rigid registration algorithms on the processing of rs-fMRI data, focusing on:

- The consistency of activation areas.
- The variability in activation maps.

By comparing the accuracy, reliability, and sensitivity of different non-rigid registration algorithms in detecting RSNs, the research seeks to identify key differences between these methods, and characterise the nuances in their implementation. The extended goal is to enhance the precision of rs-fMRI analyses by reducing ambiguity and errors associated with image registration due to machine processing. This research will not only aid in improving neuroimaging methodologies but also contribute to developing diagnostic and therapeutic approaches for neurological conditions, thus ensuring that neuroimaging can be effectively utilised for both research and clinical applications.

1.4 Research Question and Objectives

The complexities of brain imaging require advanced registration techniques to accurately map the brain's functional areas across diverse populations. Resting-state fMRI, much like other dynamic imaging modalities such as thoracic 4D Computed Tomography (CT) [32], necessitates highly optimised registration algorithms to handle inherent temporal and spatial variability. Despite recent advancements, there is an absence of a standardised approach to preprocessing and evaluating the multitude of emerging non-rigid registration algorithms used in rs-fMRI. Addressing this gap forms the core research direction of this study. Therefore, the overarching research question is:

How does algorithmic variability in non-rigid registration impact the detection of activation clusters and the reliability of functional connectivity analyses in rs-fMRI?

To address the overarching research question, the following sub-research questions guide the specific investigations of this study:

- How does the choice of non-rigid registration algorithm influence the spatial distribution and intensity of detected activation clusters in rs-fMRI?

- What is the effect of different non-rigid registration techniques on the reproducibility and consistency of reported activation maps across diverse subjects?
- How does algorithmic variability impact the identification of RSNs, and what implications does this have for understanding variability in functional connectivity analyses?

1.4.1 Objectives

The objectives of this research are systematically designed to address the challenges outlined in the research aim. These objectives focus on both the theoretical development of tools and frameworks, as well as practical assessments and validations:

1. To systematically review and identify the challenges in non-rigid registration: Assess the current state of non-rigid registration algorithms in rs-fMRI, focusing on their limitations and challenges in addressing functional connectivity analyses.
2. To develop the NRAAF framework: Design and develop a framework to systematically evaluate the performance of various non-rigid registration algorithms, with a focus on consistency, reproducibility, and impact on RSN analysis.
3. To validate the NRAAF framework using a large dataset: Apply the NRAAF framework to a large dataset (n=815) to assess differences in location, shape, and strength of activation clusters, and evaluate the effects on functional connectivity analyses.
4. To produce an algorithmic characterisation of non-rigid registration methods in rs-fMRI: Analyse the results from the NRAAF framework to characterise the variability and dependencies of different registration algorithms. This will demonstrate how these algorithms impact the consistency and reliability of RSN analyses.

Thesis Hypothesis

This thesis hypothesises that non-rigid registration algorithms will exhibit variability in their impact on rs-fMRI outcomes, specifically affecting the detection and spatial consistency of functional connectivity patterns. Given the differences in mathematical implementation among algorithms, it is expected that these differences will manifest in varying degrees of sensitivity, reliability, and spatial accuracy in neuroimaging analyses. This hypothesis will be tested through systematic comparative analysis, aiming to characterise the extent to which algorithmic choices influence key metrics in rs-fMRI studies.

1.5 Aim of Research

The aim of this research is to systematically characterise and mitigate the variability inherent in non-rigid registration algorithms used for fMRI. This study specifically focuses on algorithmic performance of these methods, aiming to identify and quantify potential biases and errors that may impact the accuracy of diagnosing neurological disorders associated with RSNs [22].

By rigorously assessing the performance of a diverse range of existing state-of-the-art registration algorithms, this research aims to develop a comprehensive understanding of the impacts these methods have on RSN analyses. The extended goal is to enhance the reliability of fMRI analyses, reduce ambiguity, and minimise errors associated with image registration, thereby improving both research quality and clinical decision-making.

1.6 Contributions of the Thesis

This thesis makes the following significant contributions to the field of neuroimaging:

1. **Development of the NRAAF Framework:** This research presents the NRAAF, which systematically compares the performance of four state-of-the-art non-rigid registration algorithms used for analysing RSNs. The NRAAF is a novel, systematic approach specifically developed to provide consistent benchmarking of these algorithms, filling a critical gap in the standardisation of non-rigid registration evaluations in neuroimaging. This framework has been applied across different functional networks within the brain to assist researchers and clinicians in selecting the most suitable analytical tools based on specific study needs, and validate effectiveness of its methodology. The novelty of the NRAAF lies in its ability to assess the impact of registration choices on downstream analyses, particularly those related to functional connectivity, thus promoting the reproducibility and reliability of neuroimaging research. This contribution has been published.

2. **Empirical Validation of NRAAF:** The NRAAF framework was empirically validated using fMRI data from a large dataset (n=815). The validation process demonstrated significant differences in algorithm performance, highlighting the strengths and limitations of commonly used neuroimaging tools under different scenarios. This contribution provides critical empirical benchmarks and recommendations for selecting appropriate algorithms based on data type and intended analysis. Such practical guidelines ad-

dress the confusion often faced by researchers regarding which non-rigid registration algorithm to use, depending on the neuroimaging context. By empirically validating these algorithms, this work strengthens the reliability of RSN analyses, promoting a standardised approach to algorithm selection in clinical and research applications. This contribution is key to ensuring that algorithm selection is not only scientifically grounded but also clinically actionable for improving diagnostic accuracy.

3. **Application of MVPA and SVM for Comparative Analysis of Non-Rigid Registration Algorithms:** This thesis introduces a novel comparative methodological approach by employing MVPA with SVM on rs-fMRI data. The novelty of this approach lies in its ability to use ML to generate voxel-wise discriminative weight maps, allowing for quantitative and spatial discrimination of the effects of different non-rigid registration algorithms on brain functional connectivity. This enables the following novel contributions:

    - *Direct Comparative Analysis:* Unlike conventional evaluation methods, this approach directly compares how different registration algorithms affect the functional representation of the brain, particularly focusing on the Control Network and other RSNs.
    - *Integration of ML for Registration Impact Analysis:* The use of modified SVM, as part of the MVPA, provides a sophisticated way of quantifying the impact of non-rigid registration algorithms on functional connectivity. This adds a ML-based evaluation layer that offers deeper insights beyond traditional statistical metrics, capturing subtle but significant differences in brain activation patterns.
    - *Generation of Discriminative Weight Maps:* The voxel-wise discriminative weight maps provide a visual and statistical tool to identify specific areas of the brain that are most influenced by the choice of registration algorithm. This level of detail allows for pinpointing which brain regions show the greatest variability due to registration, directly linking the choice of algorithm to neurobiological interpretation and enhancing the specificity of functional connectivity analysis.

4. **Algorithmic Characterisation of Non-Rigid Registration Methods:** The thesis provides a comprehensive characterisation of the evaluated non-rigid registration algorithms, detailing their variability and performance across different metrics, including accuracy, robustness, and their impact on functional connectivity. This algorithmic characterisation aims to guide researchers and clinicians in choosing the most suitable

algorithm for their specific application. By providing a detailed analysis of the differential impacts these algorithms have, this contribution addresses the consistency and reliability of RSN analyses, enhancing the practical application of neuroimaging techniques. The characterisation also informs future developments in algorithm design, ensuring that researchers and practitioners can make informed choices to optimise the quality of their neuroimaging studies.

Collectively, these contributions represent a substantial advancement in the understanding and improvement of non-rigid image registration in neuroimaging. The NRAAF framework provides a standard method for evaluating registration algorithms, while its empirical validation offers clear guidelines for tool selection, thereby reducing uncertainty in the use of non-rigid registration methods. Additionally, the algorithmic characterisation and comparative analysis provide valuable insights that directly contribute to improving diagnostic reproducibility and supporting clinical decision-making in conditions such as Alzheimer's disease, Schizophrenia, and other neurological disorders.

By providing a detailed algorithm comparison, empirical support for the developed framework, and an in-depth exploration of the impact of these computational methods, this research not only contributes to the theoretical advancement of computational neuroscience but also encourages methodological improvements.

1.7 Overview of the Thesis Structure

This thesis is organised into eight chapters, each designed to systematically address the research aims and objectives, culminating in a comprehensive understanding of the effects of non-rigid registration algorithms on rs-fMRI data. Below is a summary of the structure and key content of each chapter:

- **Chapter 1: Introduction** - Provides the context and significance of image registration in neuroimaging. Introduces rs-fMRI and its relevance to functional connectivity. Discusses the research motivation, problem statement, questions, aims, and objectives. Concludes with the contributions and overview of the thesis structure.
- **Chapter 2: Background & State-of-the-Art Review** - Reviews the evolution of non-rigid registration methods and their relevance to neuroimaging. Covers the mathematical foundations, historical development, and a detailed discussion of state-of-the-art registration methods. Highlights research gaps and challenges in current methodologies.

- **Chapter 3: Development of the NRAAF Framework** - Describes the development of NRAAF. Includes algorithm selection, data preprocessing (e.g., brain extraction, motion correction), and the metrics used for evaluation. Details the methodological approach for systematic comparison.
- **Chapter 4: Peak Activation Intensity-Based Spatial Localisation Assessment** - Outlines the methodology for assessing peak activation intensity localisation. Discusses atlas-based measurements and hemisphere-specific findings. Presents statistical analyses and visual representations of results.
- **Chapter 5: Significant Clusters-Based Network Integrity Analysis** - Details the cluster-based analysis method and its application to significant network clusters. Examines the impact of algorithm choices on network integrity. Includes a discussion of the findings and implications for neuroimaging.
- **Chapter 6: Inter-Subject Variability and Group Inference Analysis** - Investigates inter-subject variability using metrics like Mutual Information (MI) and Dice Similarity Coefficient (DSC). Integrates Multivoxel Pattern Analysis (MVPA) for performance evaluation. Highlights how different algorithms affect group-level inferences.
- **Chapter 7: Discussion** - Synthesises key findings and aligns them with the research objectives. Discusses the broader implications for neuroimaging and computational neuroscience. Identifies limitations and provides recommendations for algorithm selection and best practices.
- **Chapter 8: Conclusion & Future Work** - Summarises the main contributions and findings of the research. Discusses the potential implications for future neuroimaging research and clinical practices. Recommends areas for future exploration, including enhanced computational approaches and machine learning integration.

References and *Appendices* follow the main chapters, providing comprehensive documentation of the sources consulted during the research and additional materials that support the thesis, respectively.

1.8 Chapter Summary

Chapter 1 lays the foundational groundwork for this study, starting with an overview of image registration and its critical role in neuroimaging for aligning images from different

time points, modalities, or subjects. The chapter emphasises the importance of non-rigid registration in addressing anatomical variability and dynamic BOLD signal changes, which are vital for accurate analysis in rs-fMRI studies [1, 2, 16].

The problem of algorithmic variability is highlighted as a significant challenge that impacts the reliability of functional connectivity analyses and clinical outcomes. To address this, the chapter presents the core aims and objectives of the thesis, including the development of the NRAAF, which systematically compares non-rigid registration algorithms. The novel use of MVPA with SVM is introduced, offering new insights into the impacts of registration on neuroimaging accuracy [9, 7, 8].

The broader implications of this research include advancing the accuracy of neuroimaging analyses, which is crucial for clinical and research applications, and demonstrating the adaptability of these methods to other fields such as environmental monitoring and augmented reality [33, 2, 34, 35].

The next chapter provides a detailed review of the theoretical background, covering key concepts in non-rigid registration, RSNs, and their role in neuroimaging to establish a foundation for developing the NRAAF framework.

Chapter 2

Background & State-of-the-Art Review

In the rapidly progressing field of computational neuroscience, image registration has become an essential component of neuroimaging. Image registration is the process of aligning multiple images into a common coordinate system, enabling comparative analysis, integration, and interpretation of data. This chapter introduces non-rigid registration, an advanced form of image registration that allows for complex deformations beyond simple translations and rotations, which is crucial for capturing subtle anatomical variations across subjects or time.

Image registration has evolved significantly over the years, beginning with simple rigid registration methods that allowed only for basic transformations such as translation and rotation. These foundational methods laid the groundwork for modern advancements, enabling researchers to align images for clinical and research purposes effectively [36, 37]. The field then progressed to more sophisticated affine and non-rigid registration techniques, which account for complex deformations and have become essential in neuroimaging for capturing subtle anatomical differences [38, 39]. Non-rigid registration, in particular, has enabled better adaptation to inter-subject variability and has proven crucial in studies involving diverse populations and longitudinal analyses [31, 40].

Recent advancements have focused on improving the robustness and accuracy of these methods through Machine Learning (ML) and Deep Learning (DL) approaches, leveraging extensive datasets and advanced algorithms to enhance performance [41, 42]. The integration of learning-based techniques has accelerated the development of non-linear registration models capable of real-time processing while maintaining high precision [36, 43]. These innovations highlight the ongoing shift from traditional algorithmic approaches to data-driven models, continually pushing the boundaries of what is achievable in neuroimaging registration [44, 45].

Neuroimaging often requires precise alignment of brain images to identify changes due to pathology, development, or in response to treatment. These alignments can involve intra-subject (e.g., aligning pre- and post-surgery images of the same subject) or inter-subject (e.g., aligning multiple subjects' brain images for group studies) registration. Non-rigid registration methods have gained prominence because of their ability to adapt to complex anatomical changes, making them invaluable for studies of neurological diseases and development. However, while significant progress has been made in developing sophisticated non-linear intra-subject and intra-modality registration algorithms, there remains a gap in the comprehensive evaluation of these techniques. Such evaluations are vital, as they reveal the strengths, limitations, and areas for improvement in these algorithms, thus driving forward the field of neuroimaging [36, 38, 39].

Recognising this shortcoming, this chapter addresses the critical analysis and validation of state-of-the-art non-rigid registration algorithms. It aims to highlight the importance of thorough algorithm evaluation and to underline the impact such evaluations have on improving diagnostic and therapeutic practices. By examining the current state of algorithm evaluation, which is often overlooked due to the drive for constant innovation, this chapter makes a case for a balanced approach that equally values both the development of new algorithms and the careful evaluation of existing ones [31, 40].

The necessity for rigorous algorithm evaluation is clear: without careful validation, the limitations of new techniques may be obscured, potentially hindering the translation of these technologies into clinical settings. The trend in the field of prioritising new algorithmic development over systematic evaluation leads to the expansion of tools whose real-world capabilities are not fully understood or validated [41, 37, 42]. Therefore, this chapter advocates for a more rigorous approach to the evaluation of non-rigid registration algorithms, which is essential to enhance their practical utility and clinical impact. The extended goal is to foster an integration of empirical validation with technological innovation to ensure that new developments provide tangible benefits in both clinical and research environments [36].

Thus, in order to evaluate the landscape of non-rigid registration techniques comprehensively, a systematic review was conducted. This review aimed to collect, analyse, and synthesise current knowledge in the domain of neuroimaging registration techniques. A comprehensive search strategy was implemented across multiple databases, including Scopus, Web of Science, and Elsevier. Initially, 680 records were identified, with 450 from Scopus and Web of Science, and an additional 230 from Elsevier. After removing 50 duplicates, 630 records were screened based on titles and abstracts for relevance.

From these, 325 records were excluded due to reasons such as irrelevance to the research question (e.g., focusing on non-neuroimaging methods), non-peer-reviewed status, non-English language, and insufficient methodological rigour (e.g., studies with significant biases or lacking in data robustness).

Subsequently, 305 full-text articles were assessed for eligibility. Among these, 76 were excluded for the following reasons: 30 did not directly address the research questions pertaining to non-rigid registration methods for neuroimaging, 23 were not peer-reviewed or were published in non-reputable journals, 15 were not available in English, and 8 exhibited significant methodological flaws or provided insufficient data for robust analysis. Ultimately, 229 studies were included in the qualitative synthesis, and 150 were included in the quantitative synthesis where applicable. For a simplified view, please refer to Table 2.1.

Table 2.1 PRISMA flow summarising the literature selection process of this research.

Stage	**Description**
Identification	Records identified through database searching (Scopus and Web of Science) (n = 450) Additional records identified through Elsevier (n = 230)
Screening	Records after duplicates removed (n = 630) Records screened (n = 630) Records excluded (n = 325)
Eligibility	Full-text articles assessed for eligibility (n = 305) Full-text articles excluded, with reasons (n = 76): - Not relevant to the research question (n = 30) - Not peer-reviewed or published in non-reputable sources (n = 23) - Not in English (n = 15) - Insufficient data or methodological flaws (n = 8)
Included	Studies included in qualitative synthesis (n = 229) Studies included in quantitative synthesis (meta-analysis) (n = 150)

2.1 Introduction to Non-Rigid Registration

Non-rigid registration, a cornerstone in the evolution of medical image analysis, has seen an unprecedented increase in its necessity, especially for achieving precise alignment of images across different times, modalities, or subjects. This intricate process is critical for accurate diagnosis, treatment planning, and monitoring, becoming foundational in neuroimaging by facilitating the detailed localisation of brain structures, which is essential for both clinical and research purposes. Recent advancements, particularly through the integration of DL

techniques, have led to substantial improvements in accuracy, efficiency, and adaptability compared to traditional methodologies [19].

In neuroimaging, the goal is often to combine and compare images from different time points, subjects, or imaging modalities. Such comparisons are fundamental for understanding brain structure and function, detecting abnormalities, or tracking disease progression. Image registration—particularly non-rigid registration—is a key technique for this purpose. It allows us to align images in a way that takes into account the complex, non-linear deformations of brain structures, thereby ensuring the accurate overlay of corresponding anatomical or functional regions.

Non-rigid registration is fundamentally different from rigid and affine transformations due to its ability to capture complex, localised deformations in medical images. While rigid and affine methods can only handle rotations, translations, scaling, or shearing, non-rigid registration allows a detailed adaptation of structures. This is crucial when working with anatomical changes, especially in Functional Magnetic Resonance Imaging (fMRI), where brain structures exhibit complex, non-linear deformations across time or between subjects.

Innovative approaches in computational neuroscience, such as those developed by Hua, Kim, and He [46], employ information geometry to characterise neural information processing. These measures provide a mathematical framework that can be adapted to enhance the analysis of fMRI data, particularly in exploring complex neural interactions and dynamics. Such advanced methodologies enrich the theoretical base of neuroimaging analyses and offer new perspectives for interpreting the intricate patterns observed in Resting-State Networks (RSNs). Furthermore, Graph Neural Networks (GNNs) have also been utilised for EEG classification, as demonstrated by Klepl, Wu, and He [47], suggesting potential pathways for incorporating advanced ML techniques into fMRI data analysis, ultimately improving the accuracy and efficiency of non-rigid registration methods.

2.1.1 Relevance to Modern Neuroimaging

The evolution of non-rigid registration methods has directly impacted modern neuroimaging, providing a framework for more sophisticated analyses of brain structure and function. These advancements have proven indispensable for applications where high precision is required, such as in the longitudinal monitoring of neurodegenerative diseases, the study of neuroplasticity, and pre-surgical planning [19, 38]. Non-rigid registration enables researchers and clinicians to account for the natural variability in brain anatomy between individuals and over time, ensuring that comparisons are both accurate and meaningful.

Recent contributions, including the use of ML algorithms, have brought significant enhancements to the field. Techniques incorporating DL models, such as Convolutional Neural Networks (CNNs) and attention mechanisms, have facilitated the development of non-linear registration models that achieve improved speed and accuracy [43, 44]. These models excel in adapting to complex brain deformations, making them especially valuable in studies that require precise overlay of functional and structural data.

Additionally, the integration of advanced mathematical frameworks and ML approaches has enabled deeper insights into brain connectivity and the dynamics of RSNs. The work by Hua, Kim, and He [46] showcases how information geometry can enrich the analysis of fMRI data, enhancing the interpretation of intricate neural interactions. GNNs, as illustrated in applications for EEG classification [47], point to the potential for cross-pollination of methodologies that could further bolster non-rigid registration accuracy and utility.

Incorporating these contemporary methods into non-rigid registration pipelines not only refines the process but also expands the potential of neuroimaging studies to yield more nuanced insights. This ensures that non-rigid registration remains relevant as an adaptable and indispensable tool for advancing the understanding of brain anatomy and function, bridging historical developments with current and future practices in neuroscience research and clinical applications.

2.1.2 Historical Milestones in Image Registration

The field of image registration has undergone significant transformations, with each major development building upon the challenges and limitations of its predecessors. The earliest image registration techniques were rooted in rigid transformations, which only allowed for basic adjustments such as translations and rotations. These methods were foundational, enabling researchers to align simple anatomical structures and facilitating initial cross-sectional neuroimaging studies [38, 36]. However, their inability to handle deformations in more complex brain regions revealed the need for advanced approaches, especially when precise alignment was required for clinical and research purposes.

Affine registration marked the next phase, introducing scaling and shearing capabilities that allowed for more comprehensive global alignment. Yet, these methods still fell short when confronted with intricate, non-linear deformations in brain structures due to individual anatomical variability, pathology, or developmental changes [39]. This led to the emergence of non-rigid registration techniques, which provided a pivotal advancement by accommodating localised deformations and ensuring more anatomically accurate alignments. These

non-linear methods, including algorithms such as B-Spline and Diffeomorphic Mapping, facilitated improved inter-subject and intra-subject alignment, allowing for better comparative analyses across diverse populations and longitudinal data [37, 38].

In recent years, the incorporation of ML and DL methodologies has revolutionised the capabilities of non-rigid registration. These innovations leverage large datasets and computational power to train models capable of real-time, high-precision image alignment while addressing the limitations of manual parameter tuning and algorithmic rigidity [40, 41]. DL frameworks, such as VoxelMorph, have further refined the process by learning complex spatial transformations directly from data, streamlining registration for both monomodal and multimodal imaging studies [31, 42].

This historical progression from rigid to affine and ultimately to non-rigid registration, augmented by contemporary ML and DL advancements, emphasises the continuous evolution of techniques aimed at overcoming the inherent challenges of neuroimaging. The advancements ensure that modern neuroimaging studies can achieve accurate, robust, and anatomically consistent alignments essential for both exploratory and clinical applications, paving the way for improved diagnostics and therapeutic outcomes [36, 38].

2.1.3 Mathematical Background for Non-Rigid Registration

Non-rigid registration enables complex, localised deformations that are essential for aligning anatomical structures in neuroimaging, particularly where variations in tissue shape or size must be accounted for. Unlike rigid or affine transformations, non-rigid registration is designed to capture these complex deformations by optimising a spatial transformation $T(x)$ that maps each point x in a moving image to the corresponding point in a fixed image. This optimisation is achieved by minimising an energy function that balances both the similarity between the images and the smoothness of the transformation to preserve anatomical plausibility.

The transformation $T(x)$ is often formulated as:

$$T(x) = x + u(x) \tag{2.1}$$

where $u(x)$ is the displacement field, which varies spatially across the image and defines the deformation applied to each voxel. The energy function $E(T)$ in non-rigid registration typically includes a similarity term and a regularisation term, combined as follows:

$$E(T) = \text{Similarity}(I_{\text{fixed}}, I_{\text{moving}} \circ T) + \lambda R(T) \tag{2.2}$$

where λ is a weighting parameter that controls the balance between the similarity and regularisation terms [17, 48].

Similarity Metric

The similarity term in $E(T)$ measures the alignment quality between the fixed and moving images. Commonly used similarity metrics in neuroimaging are:

- **Mutual Information**: Mutual Information (MI) is particularly useful for multimodal image registration, such as between T1 structural and fMRI scans, as it accounts for the statistical dependency between voxel intensities in the two images [2, 6].
- **Normalised Cross-Correlation**: Normalised Cross-Correlation (NCC) is beneficial for monomodal registrations where intensity patterns are expected to be consistent across images, such as between time-series fMRI images. This metric ensures that regions with similar intensities in both images remain aligned post-registration [7].
- **Sum of Squared Differences**: Sum of Squared Differences (SSD), commonly used in monomodal contexts, SSD minimises intensity differences, making it useful for aligning high-resolution structural images [49, 50].

These metrics are typically chosen based on the modality and specific registration goals, ensuring accurate alignment by matching anatomical and functional characteristics across images.

Regularisation Term

The regularisation term $R(T)$ enforces smoothness in the displacement field $u(x)$ to prevent unrealistic or anatomically implausible deformations, such as folding or tearing of brain tissue. This is especially critical in neuroimaging to ensure the integrity of anatomical regions [48]. Common regularisation approaches include:

- **Elastic Regularisation**: Models tissue as an elastic material, constraining $u(x)$ based on elastic properties, suitable for registering scans over short time intervals or low-deformation scenarios [17].
- **Diffeomorphic Constraints**: Utilised in algorithms like the Symmetric Normalisation (SyN), implemented in Advanced Normalisation Tools (ANTs), diffeomorphic

transformations maintain topological consistency by ensuring invertibility. This constraint makes it particularly useful for tracking longitudinal changes, such as in disease progression studies [48].

- **B-Spline Regularisation**: B-Spline uses control points to locally adjust $u(x)$ while preserving overall smoothness, beneficial in applications where local flexibility is essential, such as in longitudinal studies [51, 52].

Optimisation Process

The optimisation process iteratively updates T to minimise the energy function $E(T)$, balancing between matching intensities in the similarity term and maintaining smooth deformations in the regularisation term. Techniques such as gradient descent, Levenberg-Marquardt optimisation [52], and DL frameworks like VoxelMorph [6] are employed, offering efficient convergence towards the optimal transformation that aligns the moving image to the fixed image.

Example in fMRI Analysis

To illustrate the process, consider a scenario in which two fMRI scans are taken: one before and one after a specific stimulus. The objective is to align these scans to observe and analyse subtle brain activity changes related to the stimulus.

1. **Transformation Requirement:** The transformation $T(x)$ must capture subtle, localised changes in the brain's activation patterns due to neurovascular coupling, which affects blood flow and leads to small intensity shifts in regions of activation. Non-rigid registration allows for this by adjusting the images at a voxel level, accounting for these minor shifts.

2. **Similarity Metric:** Given that fMRI scans often have slight intensity variations due to hemodynamic responses, MI is an effective similarity metric. It does not require strict intensity uniformity across scans, helping to match corresponding regions between pre- and post-stimulus images, even with slight intensity differences [48, 6].

3. **Regularisation Term:** The regularisation term plays a critical role in ensuring that the alignment is smooth and anatomically plausible. By enforcing smoothness, the regularisation prevents unnatural distortions (e.g., excessive stretching or shrinking) in brain structure, preserving continuous regions across the brain and maintaining anatomical integrity [49, 2].

4. **Optimisation Integration:** During the optimisation process, the transformation T is iteratively adjusted to minimise $E(T)$. Techniques like gradient descent or Levenberg-Marquardt optimisation [52] iteratively refine T by balancing similarity with smoothness, ensuring that the alignment achieves anatomical plausibility without sacrificing image detail. VoxelMorph, a DL-based approach, can also be applied to enhance efficiency by predicting optimal transformations based on learned features from training datasets [6].

In this example, the combined similarity, regularisation, and optimisation elements of non-rigid registration ensure an accurate alignment, maintaining structural fidelity while allowing for analysis of functional changes across the brain. This framework enables non-rigid registration to capture and align fine structural or functional differences in neuroimaging, essential for precise analysis in studies of complex brain dynamics.

In summary, the mathematical framework of non-rigid registration facilitates highly adaptable and anatomically accurate alignment, essential for neuroimaging tasks that demand precision in mapping fine structural or functional changes across datasets. By minimising $E(T)$, non-rigid registration aligns images while preserving anatomical structures, enabling precise analysis of complex brain dynamics.

2.1.4 Modern Techniques in Non-Rigid Registration

This section focuses on the latest advancements in non-rigid registration methods that leverage ML and DL techniques. These approaches represent cutting-edge solutions that push beyond traditional methods, addressing challenges related to flexibility, computational efficiency, and anatomical complexity in medical imaging.

1. **Physics-Informed Neural Networks**: Physics-Informed Neural Networks (PINNs) embed fundamental physical laws into the training process, ensuring that the registration adheres to biomechanical constraints. This novel approach augments the accuracy of patient-specific registrations, especially when conventional algorithms face limitations due to complex tissue properties [53].

2. **VoxelMorph Framework**: VoxelMorph is a DL-based framework that has proven effective in managing a diverse array of image types and anatomies. By utilising a suitable loss function—such as MI or SSD—VoxelMorph excels in both monomodal

and multimodal registration, providing robustness and adaptability in medical imaging research [54].

3. **Generative Adversarial Networks**: Generative Adversarial Networks (GANs) have demonstrated significant potential in the domain of non-rigid registration, particularly for learning realistic deformation fields. By harnessing adversarial training between a generator and a discriminator, GANs improve robustness and accuracy in registration tasks involving substantial anatomical variability [5].

4. **Reinforcement Learning**: Reinforcement Learning (RL) techniques are becoming increasingly relevant in registration tasks. RL-based approaches allow agents to dynamically interact with data to learn optimal strategies for non-rigid registration, especially in challenging scenarios characterised by large deformations or limited anatomical landmarks [49].

These modern ML and DL techniques are transforming non-rigid registration, offering automated and adaptable solutions to meet the increasing complexity of medical imaging data. By integrating these methods, neuroimaging stands to benefit from improvements in accuracy, automation, and efficiency, which ultimately contribute to more reliable diagnostic and therapeutic practices in computational neuroscience.

2.1.5 Non-Rigid Registration Methods

This section provides an overview of foundational and widely used non-rigid registration techniques. These methods have served as the backbone of neuroimaging applications, offering diverse approaches to align images across modalities, time points, and anatomical regions. Here, both traditional and early DL-based methods are considered, providing essential context for understanding the evolution of registration technologies.

B-Spline Registration: Uses control points spaced across the image, allowing for local control over deformation while preserving the overall smoothness. This is highly advantageous in longitudinal studies, where maintaining structural integrity over time is crucial [51, 52].

Demons Algorithm: Inspired by optical flow principles, this algorithm models deformations as fluid-like movements, which makes it suitable for capturing subtle, dynamic changes in brain structure [55]. The update rule iteratively adjusts displacement fields to achieve optimal alignment, particularly valuable for cases involving patient motion.

Symmetric Normalisation: The SyN algorithm within the ANTs is a symmetric diffeomorphic registration approach that ensures balanced transformations without biasing towards a specific image. It optimises an energy functional that includes similarity and regularisation terms, which are crucial for studies focusing on disease progression [48].

Deep Learning-Based Methods: These methods (e.g., VoxelMorph) employ CNNs to automate feature extraction and transformation prediction, significantly reducing manual interventions. The learned transformations are parameterised by network weights to create mappings that align moving images to fixed targets. These approaches have shown promise in overcoming limitations related to traditional parametric models, especially by automating complex transformation learning [6].

In addition to these methods, Saeed et al. [56] highlights the significance of dataset balancing in DL-based neuroimaging models. Their insights, drawn from re-sampling techniques for multi-step cyber-attack detection, demonstrate the importance of addressing imbalances within neuroimaging datasets, a prevalent issue that can affect the robustness and performance of DL-based registration models.

The integration of traditional and DL-based approaches in non-rigid registration has advanced neuroimaging substantially, offering robust frameworks capable of adapting to various clinical and research needs. As a result, these established techniques continue to provide a solid foundation for more complex and automated approaches in the evolving field of neuroimaging.

2.2 Non-Rigid Registration Algorithms

This section focuses on key tools used in non-rigid registration, including ANTs, Analysis of Functional NeuroImages (AFNI), Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL), and FMRIB Software Library (FSL). The basic concepts behind these tools are explained, with an emphasis on their roles and limitations in clinical applications.

2.2.1 Advanced Normalisation Tools (ANTs)

ANTs is a comprehensive and widely adopted non-rigid registration framework designed to accurately align neuroimaging data. ANTs employs a symmetric diffeomorphic transformation model known as SyN to provide accurate and consistent mapping between different

brain images, preserving the unique features of each individual's anatomy [48]. The SyN model ensures that neither image in a pair of images is biased as the reference, allowing a balanced alignment.

ANTs is particularly beneficial in studies that require high-precision alignment, such as comparisons involving elderly or neurodegenerative populations, as it uses high-dimensional transformations that offer both flexibility and accuracy [57]. This algorithm also supports multi-atlas segmentation approaches, allowing better generalisation across varying datasets [39]. However, ANTs requires substantial computational resources and well-annotated training datasets, which can be a limitation in some clinical settings [38].

2.2.2 Analysis of Functional NeuroImages (AFNI)

AFNI is an extensively used registration and analysis suite for fMRI data. AFNI provides tools for both linear and non-linear transformations, allowing detailed preprocessing of functional neuroimaging data [58]. AFNI's core strength lies in its flexibility for real-time fMRI and its ability to cater to user-defined registration protocols, which is critical for highly specific experimental setups.

A notable feature of AFNI is its use of the 3dQwarp algorithm for non-linear registration, which excels at refining the alignment of functional regions between subjects [59]. This adaptability allows AFNI to maintain a high sensitivity to individual differences, which can be both advantageous and challenging, as it results in significant variability in detected activations, potentially introducing false positives in cluster-based analyses [58].

2.2.3 Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL)

DARTEL, part of the SPM toolkit, is a non-rigid registration method primarily used for structural imaging analysis. It utilises diffeomorphic transformations to ensure smooth and invertible mappings between brain images. DARTEL's unique strength is in handling high-dimensional datasets efficiently, which makes it particularly well-suited for longitudinal studies where consistency over time is crucial [60].

This method works by iteratively refining templates based on the input images, resulting in a population-specific template that enhances alignment precision. However, DARTEL can face challenges when aligning brain images with significant anatomical differences or in cases where cross-modal registration is required [60].

2.2.4 FMRIB Software Library (FSL)

FSL is a versatile suite that offers both linear and non-linear registration options, primarily via the FMRIB's Non-linear Image Registration Tool (FNIRT) tool. FNIRT supports large-scale neuroimaging studies by offering robust alignment between images, with an emphasis on clinical applicability through ease of use and integration within FSL's preprocessing pipelines [61].

FSL's FNIRT approach is designed for multi-modal image registration, making it suitable for scenarios involving diverse imaging modalities, such as combining diffusion MRI with structural MRI. One of the key strengths of FSL is its ability to manage multimodal registration effectively, which allows for enhanced mapping of anatomical regions across different imaging conditions [62]. However, it may be less precise compared to tools like ANTs in cases where high-dimensional, symmetric transformations are needed [57].

2.3 Limitations in Non-Rigid Registration Algorithms

This section provides a focused analysis of the limitations in current non-rigid registration methods, especially in the context of fMRI registration of RSNs. The unique demands of Resting-State Functional MRI (rs-fMRI) data, characterised by complex deformations and high inter-subject variability, necessitate robust and accurate registration methods that can handle these challenges across diverse datasets and clinical environments. We examine the critical challenges in non-rigid registration: accuracy, robustness, computational demand, and clinical applicability.

Accuracy is a central criterion in rs-fMRI registration, as precise alignment is required to reliably capture network-level brain activity across subjects. High accuracy in functional alignment, particularly in regions with intricate anatomical features like the Default Mode Network (DMN) and Control Network, is essential for meaningful comparisons between subjects. However, the complex nature of brain deformations in rs-fMRI poses substantial alignment difficulties, as observed in algorithms like ANTs and DARTEL, which excel in accuracy but struggle with the dimensional and anatomical complexities of rs-fMRI [38, 48].

Robustness refers to an algorithm's resilience against noise and variability, which is particularly pertinent for rs-fMRI studies involving lower-quality data often encountered in clinical environments. Variability in rs-fMRI datasets—due to noise, motion artifacts, and scanner differences—makes robust registration essential for ensuring consistency across studies. Algorithms such as FSL's FNIRT and AFNI exhibit notable robustness but remain

sensitive to noise, impacting their reliability in clinical applications with variable data quality [63, 61].

Computational Demand is an increasingly important challenge, as high computational requirements limit an algorithm's scalability, especially in clinical settings where resources are often constrained. Algorithms like ANTs and DARTEL, while accurate, are computationally intensive, making them less feasible for high-throughput clinical pipelines. Efficient processing becomes essential when scaling registration methods to larger rs-fMRI datasets, particularly in environments with limited processing power [64, 38].

Clinical Applicability encompasses the practical integration of registration algorithms into clinical pipelines. Algorithms suited for clinical use must not only be accurate and robust but also user-friendly, with minimal manual tuning and rapid processing times to enable routine application. Tools like FSL are widely implemented in clinical pipelines due to their efficiency, though their simpler models may compromise precision for complex deformations. The Non-Rigid Registration Algorithm Analysis Framework (NRAAF), rather than being a registration tool, functions as an evaluation framework, assessing non-rigid registration algorithms for clinical relevance. Its role is to provide insights into each tool's clinical scalability, accuracy, and robustness, highlighting potential adjustments necessary to make these algorithms viable in clinical rs-fMRI applications. The following sections detail the specific limitations and potential research avenues for state-of-the-art non-rigid registration algorithms within the context of these challenges, with a focus on rs-fMRI data.

2.3.1 ANTs

Accuracy: ANTs offers high-precision registration, beneficial for aligning rs-fMRI datasets with considerable inter-subject functional variability. Its accuracy in aligning functional networks, such as the DMN and Salience Network, is well-suited to high-dimensional rs-fMRI data [48].

Robustness: ANTs requires careful parameter tuning, impacting its robustness in different clinical and research settings. The sensitivity to initialisation parameters introduces variability, making it challenging to achieve consistent results across studies and settings [65].

Computational Demand: The complexity of ANTs makes it computationally intensive, often unsuitable for high-throughput or resource-limited clinical settings. This demand impacts its scalability and limits its applicability in real-time or routine clinical analyses where resources are constrained [38].

Clinical Applicability: ANTs' reliance on manual parameter tuning reduces its feasibility

for clinical environments where rapid, reproducible processing is essential. Its sensitivity to initial parameters creates variability that challenges the reproducibility of network alignment, making adaptive or automated parameterisation a crucial area for improvement [48].

2.3.2 AFNI

Accuracy: Although AFNI offers specialised preprocessing tools, its nonlinear registration lacks the precision needed for small-scale anatomical alignment, which is essential for functional connectivity analysis in rs-fMRI. Misalignments can reduce the accuracy of network mapping across RSNs, impacting critical functional networks such as the DMN and Frontoparietal Networks [59].
Robustness: AFNI's dependency on user-defined parameters introduces variability, as different parameter choices can lead to inconsistent results across studies and sites, challenging consistency in clinical and multi-site research settings [63].
Computational Demand: AFNI is computationally efficient, which makes it attractive for many clinical settings. However, this efficiency comes at the cost of spatial precision, affecting the quality of network alignment in studies that require detailed rs-fMRI analyses [58].
Clinical Applicability: AFNI's dependency on user-defined parameters introduces variability, as different parameter choices can lead to inconsistent results across studies and sites. This lack of standardisation limits its applicability, especially in multi-site clinical settings that require reproducible connectivity analyses [63].

2.3.3 DARTEL

Accuracy: Known for its detailed anatomical alignment, DARTEL is beneficial for analysing brain regions essential to rs-fMRI connectivity. However, DARTEL struggles to incorporate multimodal data (e.g., combining structural and functional data), a critical factor in rs-fMRI where these data types often need integration [66].
Robustness: DARTEL's dependency on standardised anatomical templates limits its robustness across diverse populations, as template-based registration may be inadequate for subjects with anatomical variations [67].
Computational Demand: Similar to ANTs, DARTEL has high computational requirements that limit its scalability, particularly in clinical settings. Its computational demands make it challenging for large-scale or high-throughput rs-fMRI analyses, where speed and efficiency are prioritised [38].

Clinical Applicability: DARTEL's reliance on standardised anatomical templates makes it less effective for populations with non-standard anatomy, such as individuals with neurodevelopmental or neurodegenerative conditions. This reliance on normative templates suggests a need for template-independent approaches that retain anatomical precision while accommodating diverse datasets [67].

2.3.4 FSL

Accuracy: FSL's registration tools are computationally efficient but employ simpler models that may struggle with complex cortical deformations. This limitation affects rs-fMRI studies, particularly in regions with intricate folding patterns, such as the Insula or Anterior Cingulate Cortex, where precise alignment is crucial [61].
Robustness: FSL's FNIRT tool is generally robust, but its sensitivity to noise, especially in lower-quality rs-fMRI datasets, can impact network mapping accuracy [61].
Computational Demand: FSL provides a computationally feasible option for clinical use, especially with its FNIRT and FMRIB's Linear Image Registration Tool (FLIRT) tools that offer quicker processing for routine applications. However, this efficiency comes at the cost of less detailed anatomical accuracy, limiting FSL's utility in studies requiring highly nuanced rs-fMRI alignments [61].
Clinical Applicability: FSL is favoured in clinical settings for its computational efficiency, but its simplicity and sensitivity to noise reduce robustness when handling lower-quality or noisier rs-fMRI datasets. This limitation suggests a need for noise-resilient adaptations to enhance its performance in clinical rs-fMRI applications [61].

2.3.5 Emerging Approaches and Future Directions

Emerging ML techniques, particularly DL, offer promising avenues for enhancing the adaptability and robustness of non-rigid registration across varied data types, including rs-fMRI, with minimal manual tuning. By learning complex spatial transformations directly from the data, these methods hold potential for addressing existing challenges in clinical settings where both time efficiency and precision are essential. For instance, DL architectures like CNNs and GANs can improve the accuracy of rs-fMRI alignment by learning inter-subject variability more flexibly than traditional methods [68].

The advent of PINNs adds an additional layer of accuracy by embedding biomechanical constraints directly into the model training process, which is particularly valuable for patient-specific applications in rs-fMRI. By incorporating anatomical constraints, PINNs aim to

provide more consistent alignment across patients with diverse anatomical features, which may improve registration quality in clinical applications. However, these advanced methods require rigorous validation to ensure reliability, robustness, and clinical safety across different imaging conditions and patient demographics [69, 70, 71].

While each registration algorithm discussed has specific strengths and weaknesses—especially regarding computational demands, parameter sensitivity, and limitations with multimodal data—adaptive, machine-learning-driven approaches offer a compelling path forward. Addressing these limitations through such adaptive techniques could enhance the clinical utility of non-rigid registration in rs-fMRI, ultimately advancing diagnostic accuracy and therapeutic efficacy [72, 8].

2.4 Importance of Resting-State Networks in Neuroimaging

RSNs are central to neuroimaging research, revealing the brain's intrinsic functional connectivity by identifying synchronised low-frequency fluctuations across distinct brain regions during rest. These networks, such as the DMN, Control Network, Visual Network, and Sensorimotor Networks, provide insights into the brain's functional architecture underlying various cognitive functions and states of consciousness [73]. Understanding RSNs has become crucial in studying both typical and atypical brain function, offering a window into the brain's organisation during non-task-based states.

The accurate registration of RSNs is essential for studying their structure and connectivity, as it enables alignment of functional data across individuals, facilitating group comparisons and longitudinal studies. Non-rigid registration techniques are often applied to rs-fMRI data to align subtle anatomical variations across subjects, allowing researchers to map RSNs more precisely. Given the variability in individual brain anatomy, robust registration methods improve the reliability of RSN localisation, which is particularly important for analyses requiring detailed connectivity mapping, such as the identification of RSN alterations in clinical populations. Effective registration of rs-fMRI data is thus essential in ensuring that the functional connectivity and spatial organisation of RSNs are accurately captured and comparable across subjects.

RSNs play a particularly vital role in understanding cognitive functions and consciousness. For instance, the DMN is notably active during introspective and self-referential thought processes and shows reduced activity during specific task-directed behaviours [74]. This

dynamic reallocation of neural resources based on cognitive demands is fundamental to human cognition and highlights the importance of accurately capturing RSNs to understand the brain's intrinsic functionality [75, 76].

The relevance of RSN analysis extends beyond basic research into clinical domains, where it provides valuable diagnostic insights for various neurological and psychiatric disorders. Alterations in RSN connectivity patterns are associated with conditions such as Alzheimer's Disease, Autism Spectrum Disorder, and Schizophrenia [77, 78]. For example, reduced connectivity within the DMN has been linked to early stages of Alzheimer's Disease, offering a potential biomarker for early diagnosis [79]. Similarly, modifications in the Sensorimotor Networks in Autism Spectrum Disorders reveal neural underpinnings of sensory and motor dysfunctions characteristic of these conditions [80]. Effective registration is crucial in these analyses, as accurate alignment of RSNs allows for the reliable detection of connectivity changes that may serve as early indicators of disease [81].

Advanced MRI techniques and analytical methods, including Independent Component Analysis (ICA) and graph-theoretical approaches, have refined RSN studies by enhancing the accuracy of network identification and providing insights into network organisation, efficiency, and resilience [82, 83]. These advancements enable a more nuanced understanding of RSN structure and function, further highlighting the need for precise registration to ensure that subtle functional connectivity patterns are not lost or misinterpreted.

Understanding RSNs also opens pathways for therapeutic interventions. Neurofeedback and brain stimulation techniques, which aim to modulate specific RSN activity, show promise for treating neurological and psychiatric disorders by targeting network-specific dysfunctions [84, 85]. Longitudinal studies of RSNs, made feasible through reliable rs-fMRI registration, enable tracking of disease progression and assessment of therapeutic efficacy, offering significant potential for personalised patient management and improved treatment outcomes.

In summary, RSNs are a bridge between fundamental neuroscience and clinical practice, providing critical insights into brain functionality. Accurate registration is foundational in RSN research, ensuring reliable mapping of these networks across individuals. This capability advances diagnostic and therapeutic strategies in neuropsychiatric care, supporting interventions aimed at modulating RSNs to improve patient outcomes [86, 87].

2.4.1 Identification and Validation of Seed Regions for RSNs

Identifying reliable seed regions is essential for accurately mapping RSNs in neuroimaging. These seed regions serve as reference points for functional connectivity analyses, enabling

the delineation of networks based on coherent, low-frequency BOLD signal fluctuations across brain regions [25]. The selection of seed regions, often anatomically or functionally defined, impacts the reliability and interpretability of RSNs. This precision is particularly vital when investigating the DMN, Control Networks, or Sensorimotor Networks, where minor variations in seed placement can significantly alter connectivity patterns and network topography [74]. For instance, research on DMN dynamics highlights that fluctuations within specific seed regions, such as the medial prefrontal cortex or Posterior Cingulate Cortex (PCC), correlate with self-referential thought and consciousness, thereby highlighting the need for robust seed selection [75].

Multiple methodologies are used to validate the anatomical and functional fidelity of these seed regions. Advanced tools like ICA facilitate the extraction of network-specific activity, further validating seed regions by matching functional boundaries with predefined anatomical markers [82]. Moreover, atlases and parcellation methods based on structural templates (e.g., the Desikan-Killiany atlas) provide standardised regions that can be adapted as seeds across subjects, enhancing reproducibility in large-scale studies [88]. This alignment with neuroanatomical benchmarks is essential when analysing RSN alterations in clinical populations, where deviations from typical seed-based connectivity patterns may signify pathological states, such as in Alzheimer's Disease or Schizophrenia [77].

Recent advancements, including ML-based frameworks and graph-theoretical approaches, have optimised the identification and validation of seed regions. Algorithms such as GNN allow for dynamic adaptation of seed regions based on connectivity patterns, thereby improving the robustness of RSN mapping under varying anatomical conditions [89]. Additionally, validation methods employing statistical measures like the Dice Similarity Coefficient (DSC) ensure consistency in seed-based connectivity across datasets, enabling more accurate interpretations of RSNs in both health and disease contexts [90]. These validation techniques, when applied in tandem, improve the precision of seed region selection, allowing for a more reliable detection of RSNs and their functional roles in human cognition and neuropathology.

While seed-based methods are widely used to identify RSNs, they come with certain limitations. A major challenge is the anatomical differences among individuals, which can make it difficult to reliably pinpoint RSNs, especially in clinical populations where brain structure may be altered. Although standardised atlases like the Harvard-Oxford provide consistent regions for analysis, they may not capture small, meaningful differences in RSNs between individuals, which can affect the accuracy of connectivity patterns [88, 91]. Additionally, seed-based approaches assume that connectivity is stable within specific regions, but connectivity can vary depending on cognitive state or health conditions. For

example, techniques such as ICA are effective in extracting network-specific signals but can sometimes produce overlapping results, which can make it harder to interpret the unique activity of each RSN [82, 86].

Statistical validation tools, like the Dice coefficient, help ensure consistency in identifying RSNs, but they may not be sensitive enough to detect small but important changes in connectivity patterns, which are often relevant in conditions like Alzheimer's Disease or Schizophrenia [77, 90]. In summary, while these seed selection and validation techniques are valuable for mapping RSNs, it's essential to interpret findings carefully, especially when studying populations with varied or atypical brain structures.

2.5 The Role of Standard Templates in Neuroimaging

Standard templates are crucial in neuroimaging, providing a shared spatial reference that facilitates consistent image alignment and interpretation. They ensure that individual neuroimaging data can be mapped to a common coordinate space, enabling meaningful comparisons across subjects and studies. This consistency is particularly important in group-level analyses, where precise registration supports accurate alignment of complex brain structures.

2.5.1 Introduction to the MNI152 Template

The Montreal Neurological Institute 152 (MNI152) template is widely used in neuroimaging as a standard reference [82]. Derived from averaging 152 high-resolution MRI scans of healthy young adults, it offers detailed anatomical representation and alignment in stereotactic space, serving as a robust baseline for various analyses, including Voxel-Based Morphometry (VBM) and functional alignment [7, 48].

The MNI152 template's high anatomical detail and broad adoption support reproducibility and facilitate the use of standard preprocessing pipelines that enhance alignment consistency [30, 38]. This is particularly beneficial for studies using non-rigid registration algorithms, where precision affects analyses such as network integrity and functional connectivity.

2.5.2 Advantages and Limitations of the MNI152 Template

The MNI152 template has several key advantages and limitations that influence its application in neuroimaging.

Table 2.2 illustrates that the MNI152 template's high anatomical resolution enhances registration accuracy, and its compatibility with major neuroimaging tools facilitates consistent

preprocessing and broad applicability in research [82, 30, 38]. However, its development from a sample of young, healthy adults may limit its generalisability to older or clinical populations [8]. Additionally, while effective in many contexts, the template may not sufficiently capture subtle inter-subject variability, which can impact alignment precision, especially in diverse study cohorts [38].

Table 2.2 Advantages and Limitations of the MNI152 Standard Template.

Advantages	**Limitations**
High resolution aids in accurate registration [82].	Developed from young, healthy adults, which may not represent older or clinical populations [8].
Seamless integration with popular neuroimaging tools like FSL and ANTs [38, 48].	May not capture subtle inter-subject variability fully, affecting alignment accuracy in diverse groups [38].
Supports consistent preprocessing across studies, enhancing reproducibility [82, 30].	Template resolution may not align with scanner variability, potentially introducing alignment errors [92].

Choice of MNI152 for NRAAF Development

The MNI152 template was chosen for the NRAAF due to its detailed anatomical representation and widespread acceptance. This choice ensures registration results are interpreted in a common space, supporting robust evaluation of non-rigid registration algorithms. The MNI152 template's compatibility with advanced tools and extensive use enhances reproducibility and precision [82, 48].

Despite its population-specific limitations, its high anatomical detail balances precision and general applicability, making it suitable for NRAAF's focus on non-rigid registration performance. Future iterations of the framework could incorporate templates tailored for specific populations to improve alignment accuracy [8, 30].

In summary, the MNI152 template's role as a standard reference supports consistent image registration and robust group analyses. Its integration into the NRAAF aids in systematic evaluations, advancing understanding of how non-rigid registration algorithms impact neuroimaging results.

2.6 Evaluation Methods of Registration

Accurate and efficient evaluation of non-rigid registration techniques is essential within computational neuroscience, where the precision of registration algorithms directly impacts the reliability of neuroimaging analyses [93, 16]. Misalignments in registration can introduce errors in downstream analyses, potentially distorting interpretations of neural connectivity and brain function [16, 2]. The complexity of brain anatomy and the variability between subjects further highlight the necessity of robust evaluation metrics that consider accuracy, robustness, computational efficiency, and scalability. Traditional measures, such as the DSC and MI, have been instrumental but often fall short in capturing the subtle spatial correspondences required for high-resolution and multi-modal fMRI data [94, 95].

Recent studies highlight the limitations of conventional metrics and advocate for the inclusion of ML-based approaches to enhance the sensitivity and specificity of registration evaluation frameworks [6, 69]. Metrics informed by Multivoxel Pattern Analysis (MVPA) and MI, for instance, have shown promise in revealing finer intersubject variances in functional neuroimaging, particularly in clinical contexts where precise mapping is critical for diagnosis and treatment [96, 97]. These methods also accommodate the computational demands of real-time fMRI, aligning with broader trends toward automated and adaptive neuroimaging processes [39, 48].

In this section, we detail the evaluation metrics applied in this study, including their mathematical foundations and relevance to functional neuroimaging applications. Each metric is critically discussed with an emphasis on addressing known limitations and highlighting research gaps where innovative techniques may provide advancements, particularly in the context of neuroimaging scalability and multimodal integration.

2.6.1 Key Metrics for Evaluation

Key metrics assessing the registration accuracy used in this study are introduced here, including their challenges and limitations.

- **Dice Similarity Coefficient:** DSC quantifies the spatial overlap between functional regions after registration, effectively assessing alignment accuracy. Defined as:

$$\mathrm{DSC} = \frac{2|A \cap B|}{|A| + |B|} \tag{2.3}$$

  where A and B denote voxel sets from registered and target images, respectively. Here, $|A \cap B|$ represents the number of shared voxels between A and B, signifying

the overlapping region. The term $|A|$ corresponds to the total number of voxels in region A, while $|B|$ represents the voxel count in region B. By multiplying the intersection by 2, DSC ensures symmetry and provides a normalised score from 0 to 1, where values close to 1 indicate near-perfect overlap. This measure is essential for applications in functional neuroimaging where spatial accuracy is critical, though it is limited by its sensitivity to voxel intensity variations and inability to detect subtle shape differences. Refining DSC or developing alternative metrics may enhance its application for complex brain structures [16, 94].

- **Mutual Information:** MI measures the statistical dependency between two images, assessing the amount of shared information between them. Mathematically, MI is defined as:

$$\mathrm{MI}(I,J) = \sum_{i,j} p(i,j) \log \left(\frac{p(i,j)}{p(i)p(j)} \right) \tag{2.4}$$

  In this equation, $p(i,j)$ represents the joint probability of intensity levels i in image I and j in image J, indicating how often corresponding voxel intensities appear together. $p(i)$ and $p(j)$ denote the marginal probabilities of each intensity level in I and J, respectively. MI achieves higher values when there is a strong statistical association between the voxel intensities in both images, making it highly valuable for cross-modal image alignment (e.g., MRI and PET), where intrinsic brain activity alignment is critical. However, MI can be computationally intensive and may struggle with low-intensity variability regions. Future improvements could focus on optimising computational efficiency and integrating anatomical context for better performance in real-time functional imaging applications [93, 2, 95].

- **Peak Activation Intensity Analysis:** This metric, often employed in FSL's FMRI Expert Analysis Tool (FEAT), focuses on localising peak activation points within statistical maps of brain activity. This approach aids in pinpointing specific neural activation sites, offering high precision for interpreting neural patterns within brain clusters. Peak Activation Intensity Analysis provides valuable insights into spatial accuracy, which is crucial for investigating task-specific brain regions and observing group-wise functional differences [98]. However, it does not account for the spread of activity or interactions between adjacent clusters. Future enhancements could include integration with spatial regularisation techniques to assess regional connectivity alongside peak activations, providing a more comprehensive view of functional localisation [99].

- **Cluster-Based Evaluation:** Cluster-based evaluation is a statistical approach used in FSL to assess spatial distribution and cluster-level consistency of brain activity, commonly through cluster-wise correction methods in FEAT. This metric assesses the stability and reproducibility of brain activity clusters across subjects, thus supporting analyses of functional network integrity across experimental conditions [100]. While effective for identifying broad patterns, it can overlook finer intra-cluster variations and minor connectivity changes, limiting sensitivity to subtle neural shifts. To address these limitations, future methods could focus on intra-cluster variability analysis and adapt dynamic clustering to track functional connectivity shifts, thereby enhancing network-level interpretation in neuroimaging [101].

2.6.2 Public Datasets for Neuroimaging Research

The use of publicly available datasets is critical for advancing neuroimaging research, as it facilitates reproducibility, comparison, and validation of results across studies. Access to diverse, high-quality datasets allows researchers to conduct robust analyses, develop new methodologies, and assess their generalisability across various populations. This research considered several publicly available neuroimaging datasets to support comprehensive evaluation and ensure that findings are based on data representative of broader neuroimaging practices.

Table 2.3 presents a summary of notable public datasets that were evaluated for potential inclusion in this study. Each dataset was assessed based on its sample size, data quality, imaging modalities, and relevance to non-rigid registration tasks. The *Amsterdam Open MRI Collection* (AOMIC) [102] was ultimately selected due to its extensive sample size (N = 815), diverse imaging modalities—including structural MRI, functional MRI (fMRI), and Diffusion-Weighted Imaging (DWI)—and detailed demographic information. The availability of resting-state fMRI (rs-fMRI) data within AOMIC makes it particularly suitable for analysing registration performance in functional neuroimaging contexts. The public availability of these datasets ensures transparency and enhances the comparability of results across different studies. Further details on the dataset selection criteria and justification are provided in Chapter 3.

The selection of AOMIC supports the study's aim of performing comprehensive, nuanced analyses of brain function and registration algorithm performance. The dataset's inclusion of fMRI and structural imaging data enables the examination of both functional connectivity and anatomical alignment in neuroimaging workflows. Additionally, the dataset's public

Table 2.3 Public datasets employed in state-of-the-art and DL-based medical image registration which were considered for this research. This list is not exhaustive, as datasets continue to evolve. The datasets are organised by the region of interest (ROI). *The Amsterdam Open MRI Collection (AOMIC)* (highlighted in italics) was ultimately selected due to its large sample size and emphasis on fMRI.

ROI	Dataset	Modality
Abdomen, Lungs	Learn2reg 2020 Lung CT, Abdominal CT-MRI50 [103]	CT, MRI
	LIDC-IDRI, LUNA16 [104]	CT
Brain	*The Amsterdam Open MRI Collection* [102]	MRI, fMRI, DWI (MR)
	ADNI [105]	MRI, CT, CBCT, US, TRUS, x-ray
	NIREP, LPBA, IBSR, CUMC, MGH [106]	CT, MRI, x-ray, PET, SPECT, fMRI,
	OASIS ,ABIDE, ADHD200, MCIC, PPMI, HABS, Harvard GSP, the FreeSurfer Buckner40 [68]	MRI, US
	OASIS, HCP-A, BIRN [107]	MRI
	IXI Brain Development Dataset [108]	MRI
	ENIGMA-Schizophrenia DTI [109]	DTI
	BLSA, Cutting Pediatrics, ABIDE, IXI, ADHD200, NDAR, OASIS, fcon_1000, NKI_rockland [110]	MRI
	BraTS , ALBERTs, CT-MRI dataset, LPBA40, IBSR18, CUMC12, MGH10, Continuous Registration Challenge [111]	CT, MRI
Heart	NIH ChestXray14 [112]	MRI, x-ray
	NLST, DIR-Lab[43]	CineMRI, CT
	Grand Challenges in Biomedical Image Analysis, The Cancer Imaging Archive, "ChestX-ray 8" [113]	CT, MRI, PET, x-ray
Liver	RaFD [114]	CT, MRI
Pelvis	LPBA40, IBSR18, CUMC12, MGH10 [115]	CT, MRI

availability ensures that the findings can be independently verified and extended by future studies, enhancing the reproducibility and impact of this research.

2.7 Research Gap: Challenges in Practical Applicability

Despite significant advancements in non-rigid registration methods for neuroimaging, there remains a considerable gap in the practical applicability of these algorithms, particularly concerning clinical adoption and large-scale research implementation. Current research largely focuses on developing novel methods and enhancing algorithmic precision; however, practical challenges—including computational cost, lack of standardised validation protocols, and issues in adaptability—hinder widespread adoption in real-world settings [57, 95, 2, 9, 31, 116].

One primary obstacle is the computational expense associated with non-rigid registration, especially in high-resolution and multimodal neuroimaging applications. Algorithms often require extensive processing time and resources, making them less feasible for clinical environments where time efficiency is critical [41, 50, 115]. This high computational demand limits the accessibility of non-rigid registration tools in settings where real-time analysis or rapid processing is required, particularly in clinical decision-making contexts.

Another notable barrier is the lack of standardised evaluation and validation frameworks [117, 118, 119]. Most studies prioritise algorithmic innovation over comprehensive evaluation, resulting in tools whose strengths and limitations are not fully understood across diverse datasets, modalities, and anatomical regions [95, 120, 40, 4]. This lack of standardisation makes it challenging to assess algorithmic performance consistently, thus impeding the benchmarking of novel methods against established standards. Without a unified framework for assessing algorithm reliability, reproducibility across studies is compromised, limiting the potential for clinical translation and broad adoption [121, 72].

Additionally, non-rigid registration methods often exhibit sensitivity to parameter settings and variability in performance across different anatomical regions and patient populations. In clinical settings, the requirement for manual tuning and adaptation of algorithms to individual cases limits their scalability and practical utility [6, 48]. Furthermore, these algorithms' adaptation to varying imaging modalities (e.g., MRI, fMRI, CT) presents additional complexity that must be addressed to ensure seamless integration into existing clinical pipelines.

This thesis seeks to address these challenges through the implementation of the NRAAF, which systematically evaluates the accuracy, robustness, and efficiency of state-of-the-art non-rigid registration algorithms with a specific emphasis on resting-state fMRI. By establishing

a structured comparison and validation approach, this research aims to enhance the reliability, scalability, and clinical relevance of neuroimaging registration tools, contributing to more accurate diagnostic and therapeutic applications in computational neuroscience.

2.8 Chapter Summary

Chapter 2 provides an extensive review of the frameworks, methodologies, and emerging techniques used in non-rigid registration within neuroimaging, highlighting the critical importance of thorough evaluations over simply focusing on novel algorithm development. The chapter begins by defining non-rigid registration and its foundational role in medical imaging analysis, especially in accurately aligning complex neuroanatomical structures. It highlights recent advancements in ML and DL, such as CNNs, which improve the adaptability and precision of registration algorithms in addressing the diverse anatomical and functional variances across subjects.

The chapter thoroughly reviews traditional and modern non-rigid registration methods, including B-Spline registration, the Demons algorithm, SyN within ANTs, and DL-based approaches like VoxelMorph. For each method, detailed discussions focus on its theoretical underpinnings, computational strengths, and limitations in practice. For example, B-Spline registration provides fine control over local deformations but struggles with large-scale datasets, while the Demons algorithm offers computational efficiency yet may lack robustness in noisy datasets. Moreover, the chapter explores cutting-edge advancements, including PINNs, GANs, and RL-based approaches, which promise to improve registration accuracy in challenging cases by embedding domain-specific constraints or learning optimal transformations directly from data.

In addition to evaluating specific algorithms, the chapter also addresses the practical challenges faced by these techniques. These include computational demands, sensitivity to initialisation parameters, and the challenges of registering multimodal datasets. There is a strong emphasis on the need for comprehensive performance metrics, such as accuracy, robustness, and computational efficiency, to ensure clinical applicability. State-of-the-art non-rigid registration algorithms like ANTs, DARTEL, AFNI, and FSL are critically analysed with respect to these challenges, exploring issues like template selection, voxel-wise hypothesis testing, and robustness to noise and artifacts.

The chapter further examines the role of RSNs in neuroimaging, emphasising their importance in understanding intrinsic brain connectivity and their clinical relevance for neurological disorders. Techniques for RSN analysis, such as independent component

analysis and seed-based correlation analysis, are discussed as valuable tools for identifying network alterations associated with various neurological conditions.

Concluding the chapter is an advocacy for multidimensional assessment frameworks to evaluate non-rigid registration methods. The chapter highlights the necessity of structured evaluation protocols and benchmarking on standardised datasets to ensure reproducible and clinically reliable outcomes. Quantitative metrics like Target Registration Error (TRE) and MI are explored for assessing accuracy and robustness, with reproducible research practices highlighted as essential for cross-study and cross-laboratory validation.

The following chapters will build upon this foundation by detailing how these performance metrics and methodologies are applied in real-world neuroimaging scenarios. Chapter 3 introduces the NRAAF framework, which aims to address the challenges of anatomical variability and ensure robust functional connectivity mapping in fMRI studies. The chapter outlines how the NRAAF framework systematically evaluates algorithms like ANTs, DARTEL, AFNI, and FSL, examining their contributions to spatial alignment, computational efficiency, and reproducibility. This analysis demonstrates the practical implications of the methods reviewed in Chapter 2 and illustrates how these algorithms influence the accuracy and reliability of neuroimaging outcomes. The structured approach in NRAAF will inform the broader discussions on algorithm selection and the impact on functional connectivity studies presented in the subsequent chapters.

Chapter 3

NRAAF - Non-Rigid Registration Algorithm Analysis Framework

This chapter introduces the Non-Rigid Registration Algorithm Analysis Framework (NRAAF), an innovative approach developed to address significant challenges in anatomical variability and ensure robust functional connectivity mapping in Resting-State Functional Magnetic Resonance Imaging (rs-fMRI) studies [25, 27]. The NRAAF framework aims to enhance neuroimaging analysis by providing a comprehensive evaluation and characterisation of non-rigid registration algorithms, thereby promoting improved spatial alignment across subjects and enabling accurate, reproducible neuroimaging results [28, 30]. As anatomical variability can substantially impact functional connectivity interpretation [64, 122], the framework is essential for addressing alignment consistency and thereby strengthening our understanding of brain connectivity patterns. An overview of the framework's methodology, visualised in Figure 3.1, visually illustrates the processing pipeline and its integral components.

The NRAAF framework sets out to achieve several critical objectives in neuroimaging research. Primarily, it ensures consistency in spatial alignment across individuals, optimising algorithm selection to improve precision and reproducibility in functional neuroimaging [60, 57]. By integrating both univariate and Multivoxel Pattern Analysis (MVPA) techniques, the framework offers a robust platform for evaluating algorithmic performance and interpreting fMRI data with high fidelity [123, 124]. In doing so, it facilitates reproducible findings across studies, contributing valuable insights into the role of registration algorithms in neuroimaging pipelines [125, 61].

Figure 3.1 presents the NRAAF framework's comprehensive processing pipeline. This pipeline integrates multi-modal data (including structural and functional images) with the Montreal Neurological Institute 152 (MNI152) standard template to ensure precise anatomical

alignment [126, 127]. By supporting parallel operations, the pipeline can process multiple datasets simultaneously, enhancing computational efficiency and enabling rigorous inter-subject comparisons [63]. This adaptability allows the framework to be applied across different neuroimaging modalities, offering flexibility for diverse research objectives and designs [128, 129].

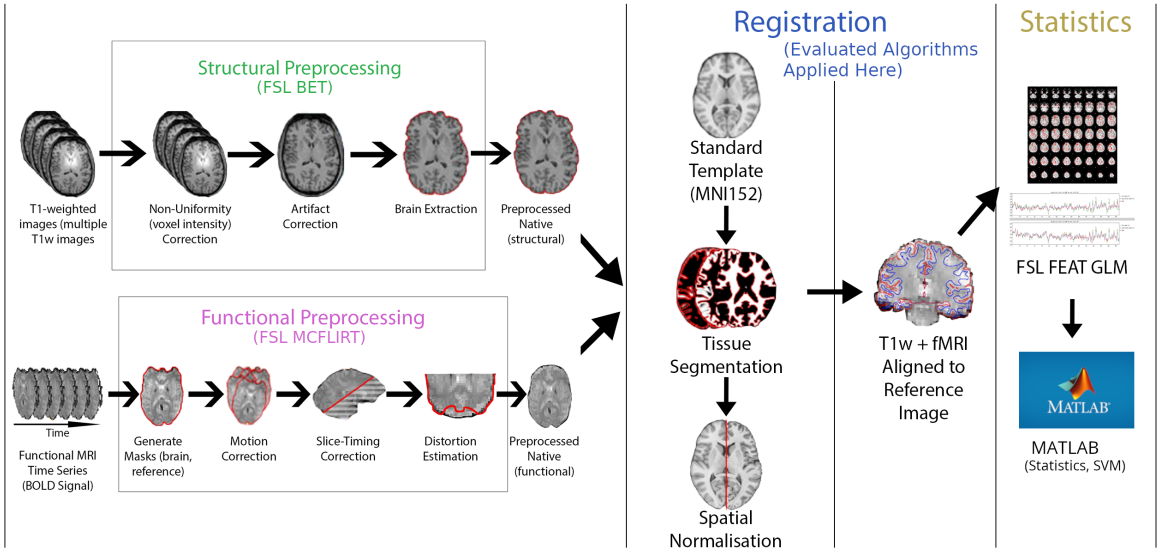

Fig. 3.1 Illustration of the NRAAF framework's functional and structural pre-processing stages, followed by non-rigid registration of images to the MNI152 standard template. Following registration, functional clusters were identified using FSL FEAT and then integrated into a General Linear Model. These clusters served as input to a SVM, which determined the decision boundary, representing the contribution of registration accuracy to activation intensity. The pipeline concludes with statistical analysis and visualisation through MATLAB R2023a.

This chapter will further detail the evaluated algorithms, dataset characteristics, pre-processing procedures, and the specific evaluation metrics used. The section concludes with an analysis that highlights the NRAAF framework's potential in advancing neuroimaging methodologies.

3.1 Implementation

This section details the computational resources, software applications, and specific procedures used for image registration and analysis within NRAAF.

3.1.1 Computational Resources

The registration and analysis were conducted on a CentOS 8.2.2004-x86_64 machine (Dell PowerEdge R740 Rack Server) equipped with Intel Xeon Gold 6240 processors (288 cores) and 720GB DDR4 RAM. This setup enabled efficient parallel processing of the computationally demanding tasks in neuroimaging analysis.

3.1.2 Software Applications

FSL 6 [61], Freesurfer 7.4 [125], and Matrix Laboratory (MATLAB) R2023a [130] were employed for preprocessing, registration, and visualisation of neuroimaging data. Visualisations were generated using MATLAB and Freesurfer Freeview 3.0 [127].

3.2 Evaluated Algorithms and Methodological Considerations

In line with the NRAAF framework's objectives of enhanced spatial alignment and reproducibility in rs-fMRI studies, this section provides an in-depth look into the evaluated algorithms, discussing selection criteria, algorithmic approaches, and each tool's unique methodological contributions to neuroimaging.

The section outlines the critical considerations and specific algorithmic configurations to maximise spatial alignment and analytical precision. The summary table (Table 3.2) provides a comparative overview to streamline understanding of each tool's notable attributes, optimisations, and applications.

3.2.1 Algorithm Selection Rationale

The selection of non-rigid registration algorithms—Advanced Normalisation Tools (ANTs), Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL), Analysis of Functional NeuroImages (AFNI), and FMRIB Software Library (FSL)—aligns with the NRAAF's core objective of supporting robust, adaptive, and precise neuroimaging analysis. Each tool was chosen based on established performance metrics, precision in anatomical alignment, and their flexibility in handling complex neuroimaging datasets. A summary of these selection criteria is presented in Table 3.1.

The NRAAF framework is structured to not only evaluate the current selection of algorithms but also to be scalable and adaptable for future expansions. The modular design

Table 3.1 Summary of algorithm selection rationale.

Algorithm	Rationale for Selection
ANTs	High spatial accuracy and symmetric normalisation, well-suited for detecting subtle anatomical variations crucial for longitudinal and morphometric studies [128, 57, 64].
AFNI	Customisable fMRI data processing capabilities, including 3dQwarp for non-rigid transformations, enhancing its adaptability to diverse neuroimaging paradigms [58, 131].
DARTEL	High-dimensional warping with group-specific template generation, supporting topological preservation in morphometric analysis [60, 123].
FSL	Versatile suite with FLIRT for linear and FNIRT for non-rigid registration, providing compatibility with standard cost functions and multi-modal data [61, 129, 126].

allows for the seamless integration of emerging tools as they are developed. This adaptability ensures that NRAAF remains a relevant benchmark for assessing the performance and robustness of registration methods in neuroimaging studies, thereby supporting continuous advancements in research methodologies and clinical applications.

3.2.2 Algorithmic Approaches and Parametrisation

Each of these algorithms employs distinct registration strategies and parameter configurations that align with the NRAAF framework's objectives. The parametrisation focuses on achieving optimal alignment with minimised computational load, tailored to each tool's specific strengths:

- **ANTs** employs symmetric normalisation, a feature specifically beneficial for longitudinal studies or tasks requiring intricate alignment of anatomical landmarks. This non-rigid registration approach enhances ANTs' utility in morphometric studies, capturing fine anatomical details that are often critical in clinical research [128].
- **AFNI** leverages its 3dQwarp for non-rigid registration, which allows users to adjust parameters for customised accuracy. Its fMRI specialisation makes AFNI especially beneficial in studies focusing on functional connectivity [63].
- **DARTEL**, with its diffeomorphic mapping, is implemented to create group-specific templates, which is highly advantageous in voxel-based morphometry. DARTEL's

methodology ensures topological consistency, allowing it to maintain the natural geometry of the brain structures under analysis [60, 125].

- **FSL** utilises FLIRT for affine transformations and FNIRT for non-rigid registration. FNIRT's approach emphasises computational efficiency and accommodates local deformations, providing flexibility for diverse image alignment tasks [126, 132].

Each tool holds distinctive strengths and should be chosen according to the specific demands of a study, such as the need for precise measurements or specialised tasks like volumetric analysis or highly accurate morphometric analyses.

3.3 Data and Seed Region Selection

This section details the dataset attributes and the rationale for selecting specific seed regions, ensuring methodological rigour in the NRAAF framework's approach to functional connectivity mapping in rs-fMRI analysis.

3.3.1 Dataset Overview

The Amsterdam Open MRI Collection (AOMIC) [102] was selected for its comprehensive dataset, including a diverse sample and multi-modal imaging options well-suited for assessing the generalisability of neuroimaging algorithms (see Table 2.3 in Chapter 2). The AOMIC dataset's variability supports the NRAAF framework's goal of robustly testing registration algorithms across anatomically diverse subjects, making it an ideal choice for studies of functional connectivity within the Default Mode Network (DMN) and associated Control Network.

3.3.2 Seed Region Selection

To evaluate functional connectivity within the Control Network, this study focuses on a seed region in the Posterior Cingulate Cortex (PCC). This region, commonly associated with higher-order control processes such as attention regulation, was selected based on work adapted from previous literature, including Doria et al. [134], on neonatal and developmental neuroimaging research.

The selected PCC seed region coordinates, shown in Figure 3.2, align with established findings and facilitate a consistent approach to investigating Control Network functionality,

Table 3.2 Comparison of spatial transformation approaches among FSL, AFNI, ANTs, and DARTEL. The table highlights the key differences in their underlying algorithms, optimisation strategies, and intended use cases. Each algorithm is selected for its specific capability to address distinct aspects of image registration, ensuring robustness and precision within the framework. References: FSL [126, 61, 129, 62], AFNI [58, 131, 59], ANTs [65, 128], and DARTEL [60, 133].

Algorithm	**Transform Type**	**Key Features**	**Optimisation Strategy**	**Unique Advantages**	**Use Cases**
FSL	Linear (Rigid)/Non-Linear (Non-Rigid)	Utilises FLIRT for linear transformations and FNIRT for non-rigid transformations, optimising standard cost functions such as the correlation ratio and MI.	FLIRT/FNIRT	Accommodates local deformations, flexibility in image alignment	Image alignment
AFNI	Linear (Rigid)/Non-Linear (Non-Rigid)	Tailored for functional MRI data. @auto_tlrc for linear and 3dQwarp for non-rigid registration.	@auto_tlrc / 3dQwarp	High degree of user-defined customisation in registration	Functional MRI studies
ANTs	Non-Linear (Non-Rigid)	Known for symmetric normalisation in non-rigid registration. Captures subtle anatomical variations.	Symmetric Normalisation	High precision in capturing anatomical variations	Detailed morphometric analyses
DARTEL	Non-Linear (Non-Rigid)	Employs diffeomorphic mappings for high-dimensional warping and creates group-specific templates.	Diffeomorphic Mapping	Ensures topological preservation, effective for voxel-based morphometry	Studies requiring high accuracy in alignment

enabling comparability with previous studies. This region was further validated through anatomical reference to the Harvard-Oxford cortical atlas, ensuring its compatibility with the AOMIC dataset and enhancing the robustness of functional connectivity analysis. The Figure 3.2 illustrates this alignment, with the seed region depicted in standard MRI views alongside its overlay on the Harvard-Oxford atlas. Key anatomical landmarks within the atlas—such as the Paracingulate Gyrus (green), Anterior Cingulate Gyrus (red), and Posterior Cingulate Gyrus (blue)—provide an anatomical framework, confirming the seed region's precise localisation within the broader cortical architecture.

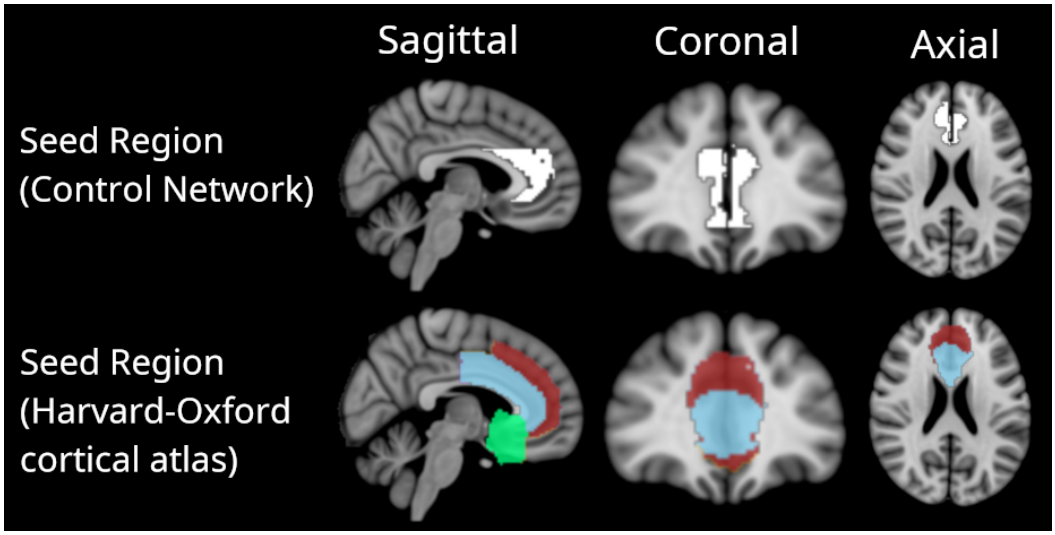

Fig. 3.2 Top row: MRI views (sagittal, coronal, axial) of the selected seed region within the Control Network, located in the Posterior Cingulate Cortex, associated with attention regulation. Bottom row: Anatomical overlay with the Harvard-Oxford cortical atlas, highlighting the Paracingulate Gyrus (green), Anterior Cingulate Gyrus (red), and Posterior Cingulate Gyrus (blue), providing anatomical context for the Control Netwrok seed region.

3.4 Data Pre-Processing Steps

This section outlines the pre-processing workflow, carefully designed to isolate image registration as the primary variable in the analysis. Each step, from file preparation to brain extraction, motion correction, and spatial normalisation, was conducted on a CentOS HPC using Simple Linux Utility for Resource Management (SLURM)-managed [135] parallel processing to handle the substantial data volume efficiently. This workflow aligns with best practices in neuroimaging research, ensuring both reliability and reproducibility [136, 94, 39].

3.4.1 File Preparation

Structural (T1-weighted) and functional MRI images were processed in NIFTI format, chosen for its compatibility with neuroimaging software such as FSL and its suitability for high-quality workflows [102]. Each dataset was organised into directories specific to each registration algorithm to enable efficient batch processing. Custom SLURM scripts were used to automate file handling and manage the parallel processing required by the high data volume. This process facilitated consistent data preparation across all subjects and algorithms, ensuring a standardised pipeline.

3.4.2 Brain Extraction

The FSL Brain Extraction Tool (BET) [137] was applied to T1-weighted images with default parameters to produce consistent skull-stripping across all subjects. BET's reliability in isolating brain structures has been widely established, making it ideal for aligning structural and functional data. A SLURM-based batch command processed BET in parallel, enabling high-throughput and uniform application of brain extraction. Figure 3.3 illustrates this process, showing T1-weighted MRI images pre- and post-extraction in sagittal, coronal, and axial views for comparison.

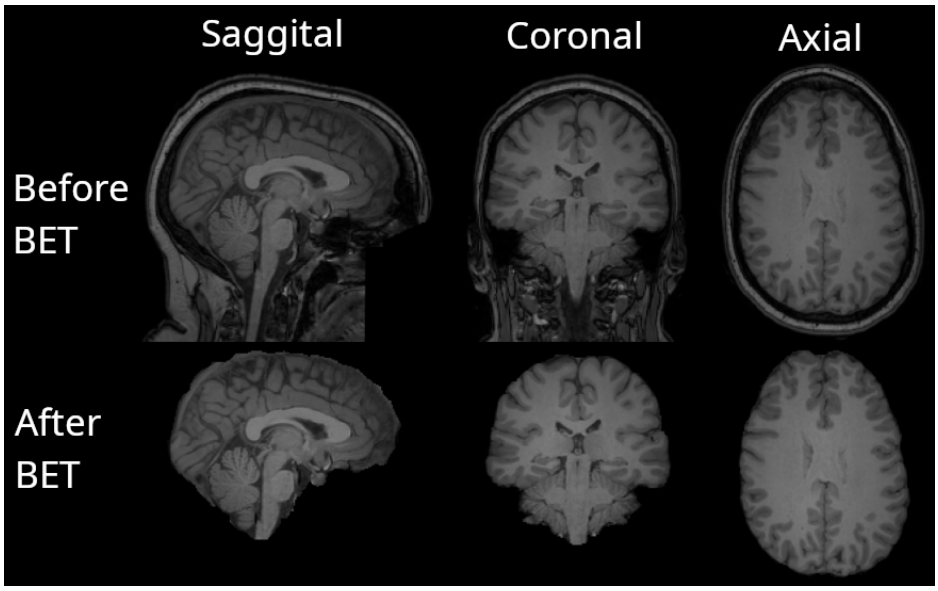

Fig. 3.3 Brain extraction process using FSL BET. Top row: Original T1-weighted MRI showing sagittal, coronal, and axial views of the full head image. Bottom row: Skull-stripped brain views in the same orientations, highlighting the isolation of brain structures crucial for accurate alignment with functional data.

3.4.3 Motion Correction

Motion correction was performed on the entire 4D fMRI time-series using FSL Motion Correction FMRIB's Linear Image Registration Tool (MCFLIRT) tool without additional parameters to maintain consistency across subjects [126]. After motion correction, the `fslmaths` tool with the `-Tmean` option was used to generate an average representative frame from the corrected 4D series for each subject. This average frame served as a stable reference and was then registered to the subject's structural T1-weighted image using linear (rigid) transformation. This approach minimised the number of transformations applied to the actual 4D time-series, reducing the risk of data degradation through repeated interpolation [136].

3.4.4 Voxel Resolution and its Implications

In MRI, voxels represent the smallest 3D data units, each corresponding to a specific volume of brain tissue. For this study, structural images were obtained at a 1x1x1 mm voxel resolution, while functional images were acquired at a 2x2x2 mm voxel resolution. The 2mm voxel size was chosen to balance spatial resolution with the Signal-to-Noise Ratio (SNR) and computational efficiency. Higher voxel resolution, while enhancing anatomical detail, often reduces SNR, making the 2mm choice optimal for capturing resting-state functional connectivity patterns without compromising data quality [138, 94]. Figure 3.4 demonstrates the effect of voxel resolution on image quality, comparing a high-resolution 1mm voxel with a larger 7mm voxel. This illustration highlights the trade-off between spatial resolution and signal strength, where smaller voxels provide finer anatomical detail at the expense of SNR, a critical consideration in fMRI analysis.

3.5 Spatial Normalisation

Spatial normalisation was performed in a two-step process to maintain the quality of the 4D fMRI data. First, the average motion-corrected frame from each subject's 4D time-series was registered to the subject's brain-extracted T1-weighted image using a rigid (linear) transformation. The resulting transformation matrix was stored for later use. Subsequently, each subject's structural T1-weighted image was non-rigidly registered to the MNI152 2mm standard template, which is widely used in neuroimaging studies to ensure consistent anatomical alignment across subjects [136]. This transformation was stored as a deformation field, which was later applied in combination with the rigid transformation matrix to normalise

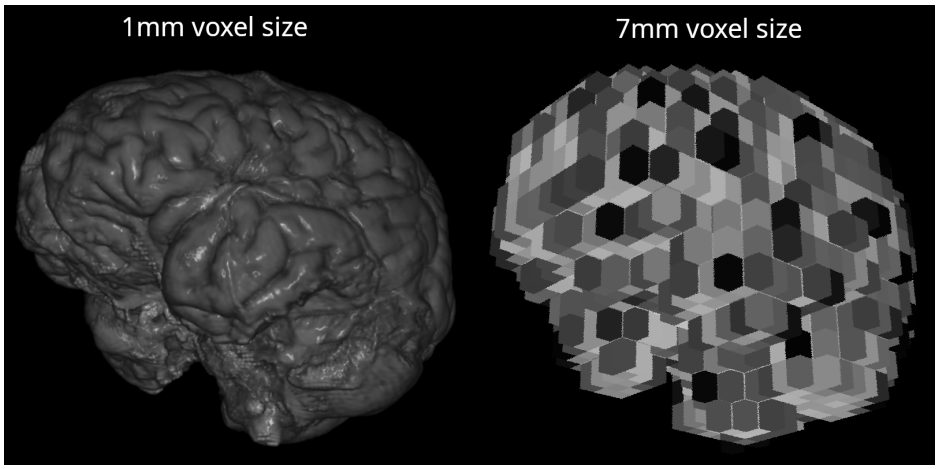

Fig. 3.4 Comparison of voxel resolutions. Left: 1mm voxel, providing fine anatomical detail (higher resolution) but with potentially lower signal strength. Right: 7mm voxel, showing increased signal but reduced spatial detail (lower resolution). Choice of voxel size influences the precision of functional connectivity mapping in fMRI analysis.

the full 4D fMRI time-series into MNI152 space. The normalised data was then visually inspected using FSL `slicesdir` tool, ensuring quality control at each step [62].

A flowchart summarising the spatial normalisation process is shown in Figure 3.5. This diagram visually outlines each stage, from the initial input of 4D fMRI data and T1-weighted images to the final quality control step, enhancing the understanding of the pipeline's sequential nature.

3.5.1 Image Registration Process

The registration process is tailored to meet the specific demands of fMRI analysis, with a focus on efficiency, algorithmic comparability, and data integrity:

- **Efficiency in Execution:** Parallelised registration was implemented to handle the total of 6520 3D-3D (both rigid and non-rigid) registrations and 3260 3D-4D transformations, significantly reducing processing time [139].
- **Iterative Approach:** Algorithms were systematically swapped to facilitate a comparative analysis of each method's performance, particularly within the AOMIC-ID1000 dataset.
- **Data Integrity Preservation:** Transforming 4D fMRI data directly to the MNI152 template using pre-saved transformations mitigated quality degradation typically associated with repeated registrations [140, 65, 61].

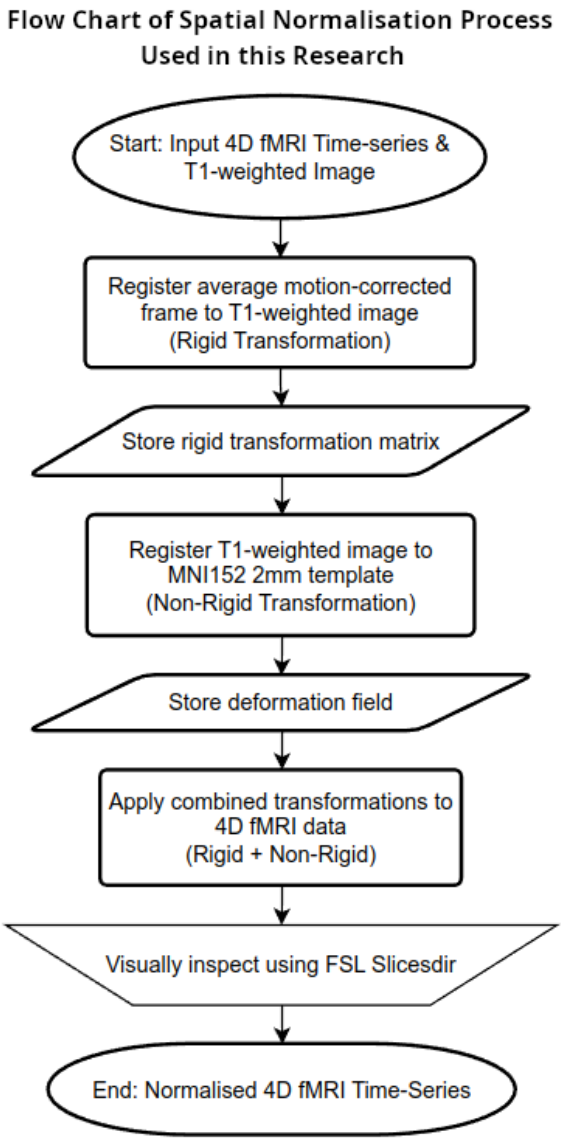

Fig. 3.5 Flowchart illustrating the spatial normalisation process for 4D fMRI data. The process involves rigid registration to the T1-weighted image, non-rigid registration to the MNI152 template, and application of combined transformations, concluding with visual inspection for quality control.

3.5.2 Processing

Following preprocessing, each subject's images were aligned to a standard template through the following steps:

1. **Linear Registration:** Establishes intra-subject consistency for individual brain images, validated through prior methodologies [60, 141].
2. **Non-Rigid Registration:** Ensures inter-subject consistency, compensating for anatomical variability and aligning images to the MNI152 template [142, 106, 143].
3. **Transformation Application:** Maintains inter-subject comparability by standardising images within MNI152 space [60, 141].
4. **Statistical Analysis:** Combines univariate analyses using General Linear Model (GLM) fitting via FSL FEAT with multivoxel method, MVPA, for voxel-wise comparisons. It incorporates DSC as a spatial overlap metric, and MI for a thorough assessment [90, 39].

The registration process used in this study is visualised in the Figure 3.5 and an example of non-rigid registration is shown in the Figure 3.6. This image shows fMRI before registration and after registration aligning the subject's brain to a standard space.

3.6 Evaluation Metrics

The registration process is fundamental to achieving consistent spatial alignment across subjects in fMRI studies, especially when evaluating non-rigid registration algorithms. This section outlines the steps involved in the spatial registration process, followed by a description of the evaluation metrics within the NRAAF, which are tailored to assess neuroimaging quality comprehensively.

Evaluation Metrics in NRAAF The NRAAF evaluates the performance of registration algorithms with metrics that address both voxel-level precision and broader spatial consistency, ensuring a robust framework for neuroimaging quality assessment. The metrics include:

- **Peak Activation Intensity:** This metric assesses the distribution of peak intensities within significant activation clusters. Using the *lmax_zstat.txt* output from FSL FMRI Expert Analysis Tool (FEAT), this metric identifies local maxima within statistically

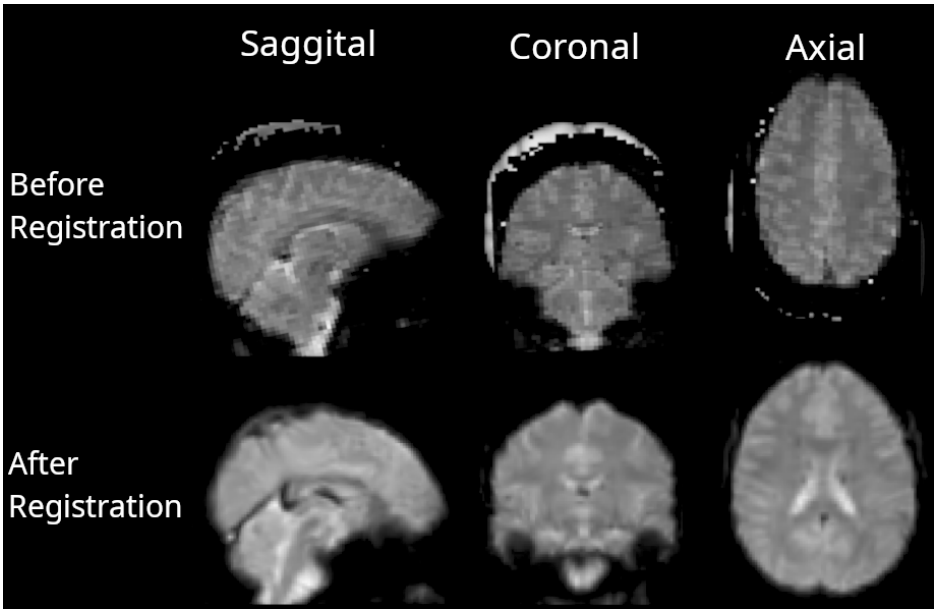

Fig. 3.6 Non-rigid fMRI registration to the MNI152 template. The top row (Before Registration) shows sagittal, coronal, and axial views of a subject's brain with misalignment to the template. The bottom row (After Registration) demonstrates registered (aligned) image, with smoother contours and better anatomical matching. This comparison highlights the effectiveness of non-rigid registration in enhancing spatial precision for cross-subject analysis.

significant regions, providing insights into each algorithm's sensitivity to high-activity areas, which is essential for accurate brain activity mapping [132]. This metric is evaluated in Chapter 4.

- **Cluster-Based Evaluation:** Using the *cluster_zstat.txt* file from FSL FEAT, this metric evaluates the spatial distribution and coherence of activation clusters, indicating the extent to which each algorithm preserves functional network structures within the brain. This helps to identify algorithms that better maintain the spatial integrity of neural clusters [132]. The results of this metric are evaluated in Chapter 5.

- **Pairwise Discriminative Analysis:** This novel metric uses Support Vector Machine (SVM) within a MVPA framework to perform voxel-wise pairwise comparisons of algorithms. By calculating SVM weights across voxel intensities for each pair of algorithms, it highlights regions where algorithms exhibit significant variation in functional activation. Each voxel's SVM weight quantifies the distinctive contribution of one algorithm over another, producing an intensity weight map that reveals differential activation patterns essential for functional connectivity analyses [144, 145]. This metric is presented in Chapter 6.

- **Spatial Overlap and Consistency:** This metric employs *cluster_mask_zstat* files and the Dice Similarity Coefficient (DSC) to measure the spatial similarity of activation clusters between algorithms. A high DSC score reflects greater overlap and indicates a consistent mapping of functional networks, which is crucial for Resting-State Network (RSN) studies that rely on spatial reproducibility [146]. Analysis of this metric is also presented in Chapter 6.

Each metric in Table 3.3 provides a unique perspective on algorithmic performance, offering a well-rounded assessment of the impact of registration algorithms on both anatomical alignment and functional mapping. NRAAF combines these metrics in univariate and MVPA to evaluate algorithmic impact on neuroimaging quality in a complementary approach:

- **Univariate Analysis via FSL FEAT:** FSL FEAT employs GLM to assess spatial precision [61]. Utilised to examine individual voxels but omits functional connectivity.

- **MVPA with SVM:** SVM provides pairwise comparisons that reveal voxel-wise activation patterns across algorithms, producing discriminative functional connectivity maps [144, 145].

Through these complementary analyses, NRAAF offers a rigorous evaluation of non-rigid registration algorithms, balancing anatomical alignment with functional network accuracy for neuroimaging applications.

Table 3.3 Overview of NRAAF evaluation metrics with corresponding techniques.

Voxel Pattern Analysis	
Univariate Metric	**Multivoxel Metric**
Peak Activation Intensity (FSL FEAT)	Multivoxel Pattern Analysis (SVM)
Number of Significant Clusters (FSL FEAT)	Spatial Overlap Assessment (DSC, MI)

3.7 SVM and MVPA Implementation

This section details the implementation of SVM and MVPA within the NRAAF framework. By leveraging SVM in a non-traditional manner, our analysis extends beyond classification, enabling a voxel-wise examination of algorithmic efficacy in aligning functional networks, specifically the Control Network. This approach not only contributes to a comprehensive evaluation of non-rigid registration algorithms but also provides a novel perspective on intersubject variability in functional connectivity.

3.7.1 SVM Implementation Details

In implementing SVM, a linear kernel function was selected due to its computational efficiency and interpretability in handling high-dimensional voxel-wise data. Linear kernels are particularly advantageous for fMRI analysis as they produce a weight vector (*Beta* values) directly corresponding to each voxel, facilitating spatial interpretation of voxel importance across pairwise algorithm comparisons [147, 148]. This configuration was critical for analysing discriminative patterns between registration algorithms, enabling the detection of nuanced differences in spatial alignment within the Control Network.

The SVMs were configured with default regularisation parameters, balancing model simplicity and performance while focusing on detecting robust algorithmic differences. Preprocessing included normalising voxel intensities across all 3D NIfTI files and aligning data to the MNI152 template to ensure that intensity variations were reflective of algorithmic performance rather than arbitrary scale differences.

For each voxel, pairwise comparisons were conducted across four algorithms: FSL, ANTs, AFNI, and DARTEL. Voxel intensities from the Control Network were extracted across all subjects, creating feature vectors that represent algorithm-specific activation patterns. The SVM model was applied to each voxel in the 3D space to compute weights that indicate the discriminative power of each voxel for distinguishing between algorithms. These weights were then aggregated in a 4D matrix, providing a spatial representation of algorithmic differences and enabling visual analysis through intensity-based heat maps.

3.7.2 Application of MVPA

Within the NRAAF framework, MVPA, especially through SVM, serves to assess functional connectivity and consistency in spatial alignment. Unlike univariate methods, which examine each voxel independently, MVPA detects patterns across multiple voxels, supporting a comprehensive understanding of algorithmic impacts on network structures [149, 150]. In this study, MVPA assesses intersubject variability by examining how consistently each algorithm aligns the Control Network across subjects. This is crucial for assessing each algorithm's suitability for resting-state fMRI analyses, where minor alignment differences can significantly impact functional connectivity interpretations [20].

Through pairwise SVM comparisons across algorithms, the MVPA framework yields a voxel-wise discriminative weight matrix reflecting the spatial distribution of each algorithm's performance within the Control Network. Instead of focusing on binary classification, this approach generates a detailed map of intensity distributions across algorithms, highlighting regions where one algorithm may outperform others. This nuanced perspective offers insights into which algorithms contribute most significantly to specific regions of the Control Network, adding value to functional connectivity and network analyses.

3.7.3 Novel Contribution of the SVM Implementation

This application of SVM represents a novel contribution by expanding beyond traditional classification tasks, such as distinguishing between patient groups or task conditions in neuroimaging [147, 148]. Inspired by approaches like those in Weaverdyck et al. [20], this study applies SVM to directly compare voxel-level activation differences across registration algorithms, yielding insights into each algorithm's unique impact on functional connectivity.

The derived weight matrix is visualised as intensity-based heat maps, quantifying each algorithm's impact on functional connectivity patterns [151, 152]. Higher weights in specific brain regions suggest that a particular algorithm significantly deviates in registering those

areas, which could have implications for functional connectivity analysis. This weight-based approach aligns with methods used by Nielsen et al. [153] and Mikolas et al. [21], providing a visually and quantitatively rich assessment of alignment consistency across algorithms, and allowing for a nuanced understanding of spatial similarity and network consistency across subjects.

Another notable feature of this implementation is the integration of univariate and multivoxel analyses, which is uncommon in fMRI evaluation frameworks. Univariate metrics, such as peak activation intensity and the number of significant clusters, are suited for single-voxel analysis. In contrast, multivoxel approaches, such as MVPA, capture inter-voxel relationships, offering insights into the spatial patterns of brain activity [146]. By combining these approaches, the framework enhances the interpretability of SVM weights, revealing both individual voxel intensities (univariate) and the broader spatial patterns (multivoxel) of algorithmic performance.

3.7.4 Relevance to Inter-Subject Analysis

In the broader context of NRAAF, MVPA is essential for evaluating inter-subject variability by detecting consistent activation patterns across subjects. As explored further in Chapter 6, multivoxel analyses are indispensable for establishing alignment reliability across individuals, a key requirement for group-level inferences [20, 153]. By focusing on pairwise voxel-wise SVM comparisons, this study provides a robust measure of how each registration algorithm performs in maintaining Control Network (and other RSNs) spatial consistency across subjects, ultimately informing algorithm selection in functional connectivity studies [144, 145, 149, 150].

To further validate these findings, Mutual Information (MI) was applied to the SVM-generated weight maps. As a metric that captures statistical dependencies in data, MI is well-suited for examining functional connectivity consistency, especially in high-dimensional neuroimaging contexts where alignment reliability is crucial [153, 154, 146, 90]. This integration quantifies the statistical interdependence across subjects' functional connectivity patterns under various non-rigid registration algorithms, offering a comprehensive reliability assessment of neuroimaging registration methods. As detailed in Chapter 6, this dual approach—using SVM weight maps in MI calculations—enhances the robustness of the inter-subject alignment evaluation, facilitating more refined insights into algorithm performance for resting-state fMRI analysis [155, 156]. This combined method also aligns with previous applications of SVM and MI in neuroimaging to achieve reliable pattern detection across

datasets, thereby reinforcing the methodological integrity and reproducibility of our findings [96, 97].

Each metric's role in NRAAF's comprehensive framework is detailed in Table 3.3, which categorises univariate and multivoxel analyses, showcasing the novel integration of these methods in the study. This multifaceted evaluation enhances the robustness of conclusions regarding algorithmic performance in fMRI studies, supporting both spatial precision and functional mapping accuracy [147, 148].

3.8 Clustering and Statistical Analysis Techniques

This section outlines the clustering and statistical techniques employed in the study, addressing the theoretical underpinnings, practical implementations, and their relevance within the NRAAF framework. These methods form the basis for interpreting both univariate and multivoxel metrics in single-subject and intersubject analyses, setting the stage for Chapters 4, 5 and 6.

3.8.1 Distinction from Clustering Methods in Machine Learning

Cluster-based evaluation in neuroimaging, such as the methods implemented within FSL FEAT, is fundamentally different from clustering techniques commonly used in Machine Learning (ML). Statistical cluster analysis in neuroimaging is designed to validate brain activation patterns within a hypothesis-driven framework, emphasising reproducibility and alignment with known anatomical and functional brain regions. For example, FSL's cluster analysis in FEAT employs Gaussian Random Field (GRF) theory to determine the significance of activation clusters, thereby providing a controlled approach to confirming network activations while reducing the likelihood of false positives [96, 97]. This approach is particularly valuable in disease-specific research, where statistical validity is essential to interpret functional changes reliably.

In contrast, clustering methods in ML are typically exploratory, aiming to identify novel patterns within unlabeled data. Algorithms such as K-means, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), and Ordering Points to Identify the Clustering Structure (OPTICS) are widely used to discover latent structures in complex datasets, including neuroimaging data, where they reveal potential subtypes in disorders like Alzheimer's disease and major depressive disorder [157, 158]. These unsupervised methods optimise intra-cluster similarity and inter-cluster separation without predefined labels, enabling the

discovery of new patterns that may not correspond to established anatomical structures. This flexibility makes ML clustering useful for investigating previously unknown subtypes or connectivity profiles, expanding the exploratory capacity of neuroimaging research.

In summary, statistical cluster analysis in FSL's FEAT framework serves a confirmatory role focused on established functional networks, while ML clustering provides a generalisable tool for exploring novel structures. This distinction is critical for understanding the role of clustering within the NRAAF framework, where statistical cluster analysis ensures alignment with predefined brain regions for robust, reproducible findings, rather than seeking previously unknown clusters in functional connectivity data.

3.8.2 Cluster Analysis and Statistical Processing

To summarise peak activation intensities, descriptive statistics, including measures of central tendency and variability such as mean, median, standard deviation, range, minimum, maximum, Interquartile Range (IQR), skewness, and kurtosis, were used. These statistics provide a foundational understanding of the distribution of activation intensities within identified clusters.

Within each cluster, peak activation points were identified based on the highest intensity values, offering insights into the regions most responsive within RSNs. Significant clusters were determined, and relationships between peak activations and clusters were assessed to reveal patterns in functional connectivity. fMRI data processing was conducted using FEAT Version 6.00, with Z (Gaussianized T/F) statistic images thresholded at $Z > 2.3$ and a corrected cluster significance threshold of $\rho = 0.05$ [159, 132]. These thresholds ensure that only statistically robust activations are considered, preserving the validity of conclusions drawn about functional connectivity in RSNs.

Inter-Subject Statistical Tests

Given the sample size of $N = 815$ and voxel size of $2 \times 2 \times 2$ mm^3, inter-subject comparisons required robust statistical methods to handle the non-normal data distribution, as confirmed by the Shapiro-Wilk test. Consequently, non-parametric statistical tests were selected, including the *Mann-Whitney U Test, Wilcoxon Signed-Rank Test, Kruskal-Wallis Test*, and *Spearman's Rank Correlation*. These tests, supported by descriptive statistics, are well-suited for neuroimaging data, allowing for reliable interpretation of activation patterns across subjects without relying on normal distribution assumptions [160, 161].

Effect Sizes & Error Correction

In this study, both statistical and practical significance are emphasised to provide a comprehensive understanding of the data. Effect size calculations complement statistical tests, offering insights into the magnitude of observed differences, which is particularly valuable in rs-fMRI where even subtle changes in functional connectivity can have significant implications for understanding neural networks.

Rank-Biserial Correlation To assess the effect sizes for the *Mann-Whitney U Test* and *Wilcoxon Signed-Rank Test*, the rank-biserial correlation (r) was employed. The rank-biserial correlation quantifies the strength of the association between two variables by relating the test statistic to the sample size. It is calculated as:

$$r = \frac{Z}{\sqrt{N}} \tag{3.1}$$

where Z is the standard normal deviate derived from the test statistic, and N represents the total number of observations in the study.

In the context of rs-fMRI, this effect size provides insight into the consistency of activation patterns or functional connectivity metrics across subjects. For instance, a higher r value would indicate a stronger association between two conditions or groups in terms of functional connectivity measures, helping to interpret the practical relevance of statistical findings. Unlike ρ-values, which merely signal whether an effect exists, the rank-biserial correlation conveys the strength and direction of this effect, aiding in the assessment of whether observed differences in network activations or functional connectivity strengths are substantial.

Spearman's Rank Correlation For correlation analyses, *Spearman's Rank Correlation* was applied to measure the monotonic relationship between variables. In rs-fMRI studies, Spearman's correlation is particularly useful when evaluating relationships between non-normally distributed data, such as connectivity strengths or functional network activations. The effect size for Spearman's correlation is indicated directly by the coefficient ρ, calculated as:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{3.2}$$

where d_i represents the difference between the ranks of corresponding variables (e.g., functional connectivity values between two conditions or time points), and n is the number of observations (e.g., subjects).

In rs-fMRI, ρ quantifies the degree of association between functional connectivity measures across subjects or conditions. For example, if functional connectivity within a specific RSN (such as the Control Network) is consistently higher in one condition compared to another, ρ would reflect this association. The value of ρ ranges from -1 to 1, with values closer to ± 1 indicating stronger monotonic relationships. This metric is particularly helpful in rs-fMRI for understanding the strength of functional connectivity relationships across different brain regions and subjects.

Bonferroni Correction for Multiple Comparisons Given the multiple comparisons performed in this study, there is a heightened risk of Type I errors (false positives), where statistically significant results might be found purely by chance. To mitigate this, the Bonferroni Correction was applied. This correction method adjusts the critical α-level by dividing it by the number of comparisons conducted, effectively lowering the threshold for statistical significance:

$$\text{Adjusted } \alpha\text{-value} = \frac{\text{Original } \alpha\text{-value}}{\text{Number of comparisons}} \tag{3.3}$$

While the Bonferroni Correction is conservative and may increase the risk of Type II errors (false negatives), it is a rigorous approach that ensures the validity of findings in studies with complex datasets like rs-fMRI. In this context, applying the Bonferroni Correction guards against spurious findings when comparing functional connectivity patterns across multiple regions or conditions, thus enhancing the robustness of results. However, due to its conservative nature, results that remain significant under this correction are likely to represent true effects, making it a reliable method for ensuring the integrity of neuroimaging conclusions [159].

This combination of effect size metrics and multiple comparison correction provides a nuanced understanding of both the statistical and practical significance of the findings. In rs-fMRI, where subtle variations in functional connectivity can yield critical insights into brain function, these measures are essential for drawing meaningful and valid conclusions within the NRAAF framework. They ensure that observed patterns in network activations or functional connectivity strengths are both statistically robust and practically relevant, supporting a comprehensive evaluation of functional brain connectivity.

3.9 Framework Evaluation and Limitations

The NRAAF framework employs an integrated analysis to evaluate the efficacy of various non-rigid registration algorithms, specifically FSL, ANTs, AFNI, and DARTEL. These algorithms were selected for their complementary strengths in neuroimaging tasks, as detailed in Table 3.2. Performance insights from the analysis of the key metrics provide a nuanced understanding of each algorithm's capabilities and limitations:

- **ANTs:** Known for its spatial accuracy and symmetric normalisation capabilities, ANTs excels in regions with high anatomical variability, as demonstrated by high DSC in cortical regions [48]. However, this precision comes at the expense of computational demands, which may limit scalability in large-scale studies or settings with limited processing resources [162].
- **AFNI:** Leveraging tools like 3dQwarp, AFNI excels in flexibility, especially in user-defined parameters for fMRI-specific tasks [58]. Its performance, assessed through MI and DSC metrics, demonstrates robust functional alignment but may vary in anatomical alignment across heterogeneous datasets.
- **DARTEL:** With high-dimensional warping and group-specific template creation, DARTEL excels in preserving topological structures, particularly within large group datasets. However, its reliance on group templates may introduce biases, affecting generalisability [60].
- **FSL:** This software suite, combining FLIRT (rigid transformations) and FNIRT (non-rigid transformations), is optimised for efficient, flexible registration [61]. While FSL demonstrates computational efficiency, it may lack spatial precision in anatomically variable regions, which the MI metric reveals as a minor limitation when compared to algorithms like ANTs [98].

These performance insights form the foundation for deeper analysis in Chapters 4, 5, and 6, where each algorithm's strengths and limitations will be discussed in specific neuroimaging applications and functional connectivity studies.

3.9.1 Generalisability and Limitations

The NRAAF framework provides a comprehensive evaluation of registration algorithms; however, several limitations exist. First, the dataset specificity of the AOMIC restricts the

framework's generalisability, as algorithmic performance may differ across datasets with varying anatomical characteristics. Additionally, the computational demands of algorithms like ANTs and DARTEL necessitate substantial processing power, which may not be feasible in all research settings.

A critical distinction between univariate and multivoxel analyses within the NRAAF framework also reflects these constraints. While univariate metrics, such as peak activation intensity, offer insights into voxel-wise statistics, multivoxel methods like SVM provide a broader assessment of spatial patterns across voxels, which is essential for detailed functional connectivity evaluations [123]. However, multivoxel analyses demand higher computational resources, limiting their application in large datasets and requiring robust data pre-processing [163].

These limitations are acknowledged here to set a realistic context for readers as they proceed through the results in subsequent chapters. This also anticipates the broader discussion in Chapter 7 regarding the practical considerations and implications of algorithm selection for specific neuroimaging applications.

3.9.2 Mutual Information

MI is a crucial metric within the NRAAF framework, quantifying shared information between datasets to provide a non-rigid measure of voxel alignment accuracy. MI's capacity to capture subtle alignment nuances is particularly beneficial in rs-fMRI studies, where precise voxel correspondence is essential for accurate functional connectivity mapping. This metric offers a robust validation of registration algorithms, especially in studies involving RSNs, as it can effectively assess spatial similarity in complex brain structures [164].

In line with current neuroimaging practices, MI supports the evaluation of spatial accuracy and consistency, enhancing the framework's methodological integrity. By facilitating accurate integration and comparison of functional data across subjects, MI enables more meaningful insights into neural processes, further optimising neuroimaging protocols for clinical and research applications [165].

3.9.3 Spatial Overlap Assessment

The DSC is integrated within the NRAAF framework to evaluate the spatial overlap of binarised activation maps generated by different registration algorithms. This metric provides an intuitive measure of spatial consistency in neuroimaging studies, where a coefficient of 0 indicates no overlap and 1 denotes perfect overlap [166]. Such measurements are crucial for

assessing the accuracy of RSN alignment, as consistency across subjects is essential for valid group-level inferences.

Parallel processing was employed to compute DSC matrices efficiently for each algorithm, with a global 95th percentile threshold used to identify significant spatial overlaps. This thresholding approach follows established methodologies, such as those by Smith et al. [98], reinforcing the DSC's role as a robust indicator of functional network alignment.

Interpretation

To complement quantitative analysis, visual representations of DSC matrices, including graph visualisations and heat maps, were created. These visual aids provide an intuitive assessment of spatial overlap, offering insights into algorithm performance and spatial alignment accuracy. This approach aligns with recommendations from Collins and Evans [165], highlighting the value of visual representations in neuroimaging for clarifying complex data.

By integrating MI and DSC analyses with visual tools, the NRAAF framework provides a comprehensive and robust assessment of registration algorithms, setting a foundation for further evaluation in the results chapters. The implications of these metrics for functional connectivity and network analyses are explored in detail in Chapter 6.

3.10 Chapter Summary

Chapter 3 presents the foundational elements of the NRAAF, which is designed to evaluate non-rigid registration algorithms in the context of rs-fMRI data analysis. This chapter opens with an introduction to the NRAAF framework's objectives, highlighting the need for consistent spatial alignment across subjects and robust functional connectivity mapping to ensure data quality in neuroimaging studies. Emphasis is placed on the importance of selecting diverse datasets to ensure robust findings across neuroanatomical variations, along with the rigorous preprocessing steps required for high-quality data preparation, such as skull-stripping, motion correction, and spatial normalisation.

The systematic approach to evaluating registration algorithms is a key focus in this chapter. Through a combination of univariate and multivoxel analysis techniques, including SVM for generating voxel-wise discriminative weight matrices, the NRAAF framework allows for the nuanced evaluation of each algorithm's impact on spatial consistency and functional connectivity. These SVM-generated weights support a detailed view of the spatial

patterns distinguishing algorithmic performance, emphasising the flexibility of the framework in comparing a range of neuroimaging registration algorithms.

The chapter then details the metrics used within NRAAF, including MI and the DSC, to assess registration accuracy. MI provides a non-rigid measure of shared information between images aligned by different algorithms, which is crucial for accurately evaluating functional connectivity. Meanwhile, the DSC quantifies spatial overlap between functional regions post-registration, ensuring robust alignment of functional networks across subjects. By using these metrics in combination, the NRAAF framework meets its core objectives of enhancing spatial alignment reliability and reproducibility in neuroimaging data.

This chapter also addresses the limitations of the NRAAF framework. It acknowledges constraints related to computational demand, dataset specificity, and the distinct resource needs of univariate and multivoxel approaches, establishing the practical considerations for subsequent analyses. The integration of these metrics and methodological considerations not only strengthens the framework's ability to rigorously evaluate algorithm performance but also establishes a solid foundation for deeper insights in later chapters.
The subsequent chapters delve into the detailed evaluations and implications of each metric category:

- Chapter 4 will examine *Peak Activation Intensity* as the first univariate metric.
- Chapter 5 will assess the *Number of Significant Clusters*, focusing on spatial distribution and coherence.
- Chapter 6 will explore the multivoxel metrics of the framework, specifically *Spatial Overlap*, *Mutual Information*, and *MVPA* to gauge inter-subject registration reliability.

This systematic analysis allows the NRAAF framework to offer a comprehensive approach to evaluating algorithm performance in neuroimaging, contributing to the field's goal of achieving reliable and reproducible results.

Chapter 4

Peak Activation Intensity-Based Spatial Localisation Assessment

Chapter 4 conducts an in-depth analysis of Peak Activation Intensities derived from Resting-State Functional Magnetic Resonance Imaging (rs-fMRI). Building directly on the theoretical foundation laid in Chapter 3, which introduced the Non-Rigid Registration Algorithm Analysis Framework (NRAAF) framework for comparative evaluation of Functional Magnetic Resonance Imaging (fMRI) registration algorithms. This chapter investigates how each algorithm influences activation intensities across brain hemispheres. A detailed focus on peak activation is essential, as it allows us to explore the nuances of algorithmic impact on neuroimaging data consistency and accuracy, aligning with the growing emphasis on precision in fMRI data analysis [93, 16, 4].

In the context of rs-fMRI, accurately capturing and interpreting Peak Activation Intensities is critical for reliable neuroimaging outcomes. Given that the selected non-rigid registration algorithms may introduce subtle variations in these peaks, this chapter empirically examines whether algorithm choice influences intensity detection, with a specific emphasis on hemispheric differences. Such analysis is essential for validating the robustness of neuroimaging methods, ensuring that clinical and scientific conclusions are based on data unaffected by algorithmic artifacts [7, 8]. Here, we present results for peak intensities across hemispheres, exploring the correlation between the Control Network and its seed region as identified through these algorithms, thereby addressing the reliability and consistency concerns central to neuroimaging research [23, 28].

This chapter's analysis is tied to the preceding and following chapters. In Chapter 3, the methodological foundations and metrics for algorithm evaluation were thoroughly established. Chapter 4 extends this framework by applying these metrics specifically to peak activation

analysis, thereby setting a strong empirical basis for Chapter 5, where the focus will shift to network integrity and the exploration of significant clusters across Resting-State Networks (RSNs).

This research is a step forward in understanding the technical limitations and advantages of various neuroimaging algorithms, with an emphasis on their performance in identifying high-intensity regions in rs-fMRI brain scans. By rigorously comparing multiple algorithms, we aim to contribute to the standardisation of neuroimaging practices, supporting advancements in reproducibility and diagnostic accuracy for fMRI data [22, 167].

4.1 Methodological Framework

In this chapter, a comprehensive set of statistical metrics—namely, Peak Activation Intensity, skewness, kurtosis, Standard Deviation (SD), Interquartile Range (IQR), range, and non-parametric correlation—are employed to assess the performance sensitivity and reliability of the neuroimaging algorithms under investigation. These metrics are foundational in neuroimaging analysis as they allow researchers to capture subtle variances in activation patterns, which may be attributed to algorithm-specific processing differences [157].

- The **Peak Activation Intensity**, representing the highest detected neural activity level within predefined regions, serves as a primary indicator of algorithmic sensitivity to brain activity during resting states. This metric is particularly relevant in rs-fMRI studies, where detecting subtle fluctuations in spontaneous brain activity is crucial for understanding functional connectivity and regional brain dynamics. Studies like Avants et al. [168] and Cox [58] highlight the importance of peak activation for evaluating the performance of image registration algorithms in detecting region-specific activity levels, a critical factor in comparing algorithms for functional brain mapping.

- **Skewness and kurtosis** offer additional insights into the shape and tails of the distribution of activation intensities. Skewness measures asymmetry, while kurtosis indicates the concentration of values around the mean, helping to evaluate data normality and highlight any systematic deviations introduced by the algorithms. Zhang et al. [169] discuss how these metrics reveal the presence of outliers and non-normality in brain activity data, both common in rs-fMRI due to neural signal variability . Additionally, Power et al. [170] and Poldrack et al. [171] emphasise their use in assessing algorithmic impact on neural pattern interpretation, crucial for understanding rs-fMRI data characteristics.

- **Standard Deviation** and **IQR** further clarify the variability and consistency in intensity detection. SD provides an overall sense of the spread in detected intensities, essential for evaluating an algorithm's sensitivity to fluctuations in neural activity. The IQR, by focusing on the central spread, minimises the influence of outliers and offers a robust measure of consistency, as Ashburner [60] and Jenkinson et al. [61] explain. These metrics are particularly valuable when comparing algorithms like Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL) and FMRIB Software Library (FSL), which demonstrate different detection stability patterns in rs-fMRI, highlighting subtle differences in reliability and algorithmic consistency.

- The **range** of intensities detected by each algorithm offers insights into the extent of variability, highlighting expansive or limited detection capabilities. This measure is critical in rs-fMRI studies, where extreme values may indicate heightened sensitivity or potential instability in detecting neural activation. Research by Tustison et al. [39] and Cox [58] highlights how the range can reflect algorithmic behaviour in extreme value sensitivity, with Analysis of Functional NeuroImages (AFNI), for instance, showing greater variability and the presence of outliers, which could impact intensity detection interpretations.

- Finally, **non-parametric correlation** (Spearman's rank correlation) is used to evaluate the consistency of Peak Activation Intensities across hemispheres without assuming data normality. This metric is particularly relevant in rs-fMRI, where non-normal distributions are common, as it enables robust assessments of hemispheric correlation and consistency across algorithms. Guillaume et al. [172] and Cox [58] discuss the robustness of non-parametric tests in neuroimaging for these scenarios. Additionally, **effect sizes for non-parametric tests** such as the Mann-Whitney U and Wilcoxon Signed-Rank tests can quantify hemispheric differences, particularly relevant in studies examining lateralised brain functions, as illustrated by Mann [173] and Wilcoxon [174].

The relevance of these statistical measures lies in their collective ability to reveal algorithm-induced biases, variability, and reliability across different aspects of intensity detection. This comprehensive analytical approach contributes to a more robust comparison framework, facilitating precise assessments of algorithmic performance variations in rs-fMRI. By integrating these metrics, this chapter aims to enhance our understanding of each algorithm's suitability for detecting high-sensitivity features in neuroimaging data, ultimately informing more refined and consistent methods in neuroimaging research.

Literature Context

Existing research highlights the variability introduced by different registration algorithms in neuroimaging, particularly regarding Peak Activation Intensities across brain regions. Studies have documented the ways algorithm choice can influence neuroimaging outcomes, with implications for data consistency and reliability [175, 9]. These findings serve as a basis for the present chapter's analysis, which aims to expand on previous work by specifically comparing Peak Activation Intensities across four non-rigid registration algorithms: Advanced Normalisation Tools (ANTs), DARTEL, AFNI, and FSL. This comparative approach aligns with the objectives of the NRAAF framework (introduced in Chapter 3), and aims to provide actionable insights for researchers selecting algorithms based on reliability across hemispheric measures and sensitivity to peak intensities.

4.1.1 Justification for Hemispheric Analysis

A hemispheric perspective was applied in this analysis to evaluate algorithm consistency in capturing activation intensities across lateralised brain functions. Hemispheric asymmetry is a fundamental aspect of brain structure, contributing to functional specialisation and cognitive adaptability [176, 177]. In line with this principle, the decision to evaluate hemispheric differences allows for an assessment of whether registration algorithms maintain consistent representations of lateralised brain functions or introduce biases in asymmetrical regions. This chapter's hemispheric analysis does not presuppose an intrinsic superiority of lateralised representation but rather provides a framework for assessing how algorithm choice may impact the neuroimaging interpretation of lateralised brain functions, with the understanding that this approach may not be universally applicable across all future studies.

4.2 Atlas Measurements

The Harvard-Oxford atlas is employed in this chapter for its comprehensive and probabilistic mapping of cortical and subcortical brain structures, providing a foundational tool for identifying brain regions with high precision. This atlas is particularly effective in neuroimaging studies that require accurate regional delineation, as its probabilistic approach accounts for inter-subject anatomical variability, enhancing the accuracy and reliability of fMRI data registration outcomes [175, 91]. The Harvard-Oxford atlas is widely adopted within the neuroimaging community due to its compatibility with major software packages, facilitating standardisation across studies and contributing to reproducible results [61]. This chapter

utilises the atlas for consistent brain region identification across algorithms, allowing for direct comparisons of algorithmic performance in representing activation intensities within anatomically defined regions.

4.3 Evaluation of Peak Activation Intensities

This section presents results of Peak Activation Intensities among the four evaluated non-rigid registration algorithms.

4.3.1 Descriptive Analysis of Peak Intensities

This section presents descriptive statistics for the Peak Activation Intensities across the four neuroimaging algorithms in both hemispheres. Table 4.1 shows the metrics for the left hemisphere, while Table 4.2 details the right hemisphere. These statistics, including mean, median, SD, range, IQR, skewness, and kurtosis, are crucial for understanding the sensitivity and variability of each algorithm in detecting Peak Activation Intensities.

Table 4.1 reveals that ANTs recorded a high mean (6.9184) and maximum value (21.2000), indicating a tendency to detect higher peak intensities, albeit with notable outliers as suggested by its skewness (3.1411) and kurtosis (16.4160). In contrast, DARTEL exhibited a lower mean (6.3695) with a narrow SD (1.4351), suggesting consistent but less sensitive measurements. AFNI demonstrated high variability with a large range (23.6000) and minimum values of zero, which might indicate instances of non-detection. FSL displayed a moderate mean (6.7825) and SD (1.8348), reflecting a balanced approach in sensitivity and variability.

Table 4.1 Descriptive statistics of *left hemisphere* Peak Activation Intensity in the Control Network. This table compares four neuroimaging algorithms: ANTs, DARTEL, AFNI, and FSL, highlighting differences in intensity detection sensitivity, consistency, and outliers.

Test	**ANTs**	**DARTEL**	**AFNI**	**FSL**
Mean	6.9184	6.3695	6.6104	6.7825
Median	6.3587	6.0705	6.2637	6.3218
Std Dev	2.0426	1.4351	1.6962	1.8348
Min	4.7548	4.3756	0	4.5524
Max	21.2000	18.5000	23.6000	20.1000
Range	16.4452	14.1244	23.6000	15.5476
IQR	1.6139	1.2545	1.3503	1.5178
Skewness	3.1411	3.3743	3.6593	2.7429
Kurtosis	16.4160	21.8712	27.3054	13.3164

In the right hemisphere (Table 4.2), ANTs again demonstrates high mean (6.9339) and maximum (31.6000) values, with skewness (4.3537) and kurtosis (39.0803) reflecting extreme values. DARTEL shows consistent measurements with moderate variability, and AFNI displays significant variability in its range (22.2000), potentially impacting reliability in intensity detection. FSL maintains balanced sensitivity and variability, similar to its performance in the left hemisphere.

Table 4.2 Descriptive statistics of *right hemisphere* Peak Activation Intensity in the Control Network, highlighting the comparative performance of ANTs, DARTEL, AFNI, and FSL.

Test	**ANTs**	**DARTEL**	**AFNI**	**FSL**
Mean	6.9339	6.5036	6.7072	6.8987
Median	6.4713	6.0677	6.1288	6.3080
Std Dev	1.9856	1.6907	2.0458	2.0776
Min	4.5680	4.3054	0	4.6895
Max	31.6000	21.1000	22.2000	21.3000
Range	27.0320	16.7946	22.2000	16.6105
IQR	1.6710	1.4658	1.5843	1.7177
Skewness	4.3537	3.4518	2.7476	2.9012
Kurtosis	39.0803	21.8481	14.2493	14.2793

Hemisphere-Specific Findings

The analysis indicates consistent performance by each algorithm across both hemispheres, with no significant differences in median intensities (Tables 4.1 and 4.2). ANTs and AFNI showed higher variability and outliers, while DARTEL and FSL exhibited more uniform detection capabilities. The lack of marked hemispheric discrepancies implies that the observed differences are algorithmic rather than inherent to hemispheric asymmetry. This consistency is important for studies examining lateralisation, suggesting that algorithm choice, rather than hemispheric variance, predominantly influences activation intensity detection.

4.3.2 Visual Summary of Findings

To complement our descriptive analysis, we provide a series of visualisations to illustrate the distribution, central tendency, and variability of Peak Activation Intensities across algorithms and hemispheres. These include box plots (Fig. 4.1) and violin plots (Fig. 4.2).

The box plots employ a star notation system to indicate statistical significance across comparisons, where the level of significance is denoted by symbols placed above the box

plot bars. This notation, detailed in Table 4.3, serves as a legend for understanding the degree of significance between algorithms' peak intensity distributions. In particular, this table aids in interpreting whether the differences in detection of Peak Activation Intensities are statistically meaningful.

Table 4.3 Explanation of star notations used in box plots for statistical significance. This legend assists in interpreting the significance of observed differences between algorithms across hemispheres, guiding the assessment of algorithm performance in detecting Peak Activation Intensities.

Symbol	Statistical Significance
ns	not significant ($\rho > 0.05$)
*	significant ($\rho \leq 0.05$)
**	very significant ($\rho \leq 0.01$)
***	extremely significant ($\rho \leq 0.001$)
****	most extremely significant ($\rho \leq 0.0001$)

From the box plots, we observe the following algorithm-specific insights:

- **ANTs**: Exhibits a wide IQR and high variability, with several outliers in both hemispheres. This indicates that ANTs is sensitive to a diverse range of intensities, potentially capturing subtle fluctuations in activation, but this also introduces variability that might impact detection reliability.
- **DARTEL**: Demonstrates tighter IQRs, with fewer outliers, indicating that it provides more consistent intensity detection across hemispheres. This suggests that DARTEL may be more suitable for studies focused on consistent, typical activation levels, as it appears less affected by extreme values.
- **AFNI**: Shows a particularly wide range with many outliers, especially in the right hemisphere. This suggests high variability and a potential for either capturing extreme intensity values or detecting cases of no activation (e.g., minimum values of zero), which may affect its reliability in consistently interpreting activation patterns.
- **FSL**: Displays a moderate IQR with fewer outliers, reflecting a balanced approach to detecting peak intensities. FSL's performance suggests it strikes a balance between sensitivity and stability, making it a versatile choice across studies requiring both consistent and varied activation detection.

The consistent median lines across both hemispheres imply that none of the algorithms exhibit significant hemispheric bias in intensity detection, reinforcing that observed differences are likely due to algorithmic variations rather than intrinsic hemispheric differences.

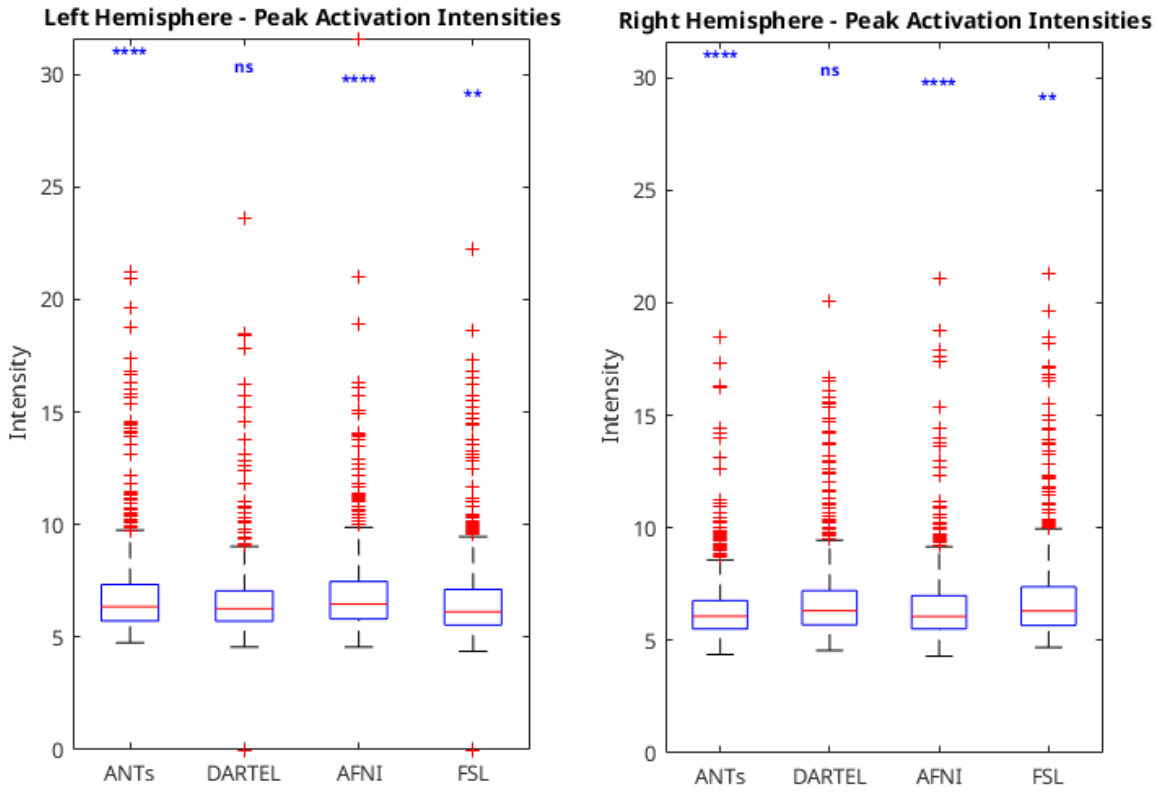

Fig. 4.1 Box plots comparing Peak Activation Intensities for each algorithm (ANTs, DARTEL, AFNI, and FSL) across left (left panel) and right (right panel) hemispheres. ANTs and AFNI display wider variability with numerous outliers, indicating high sensitivity to extreme values, while DARTEL and FSL exhibit narrower IQR, suggesting more consistent detection across both hemispheres. Statistical significance between algorithms is marked by star notation (as defined in Table 4.3), highlighting meaningful differences in detection patterns.

To further contextualise these findings, violin plots (Fig. 4.2) offer insights into the density and distribution of peak intensities for each algorithm, visualising both the range and frequency of intensity values across hemispheres.

Key interpretations from the violin plots include:

- **ANTs and AFNI**: Both algorithms show broad distributions, indicating they capture a wide range of activation intensities. This distribution suggests high sensitivity to fluctuations in neural activity but could introduce increased variability, as seen in the abundance of outliers.
- **DARTEL**: Displays a more uniform and concentrated distribution, reflecting consistent detection of typical intensity values across hemispheres. This uniformity may make DARTEL a preferred choice for studies prioritising stable, reproducible activation detection.
- **FSL**: Shows a moderately broad distribution, balancing the need for consistent detection with the capacity to capture occasional extreme values. This makes FSL a flexible choice across studies with varying intensity requirements.

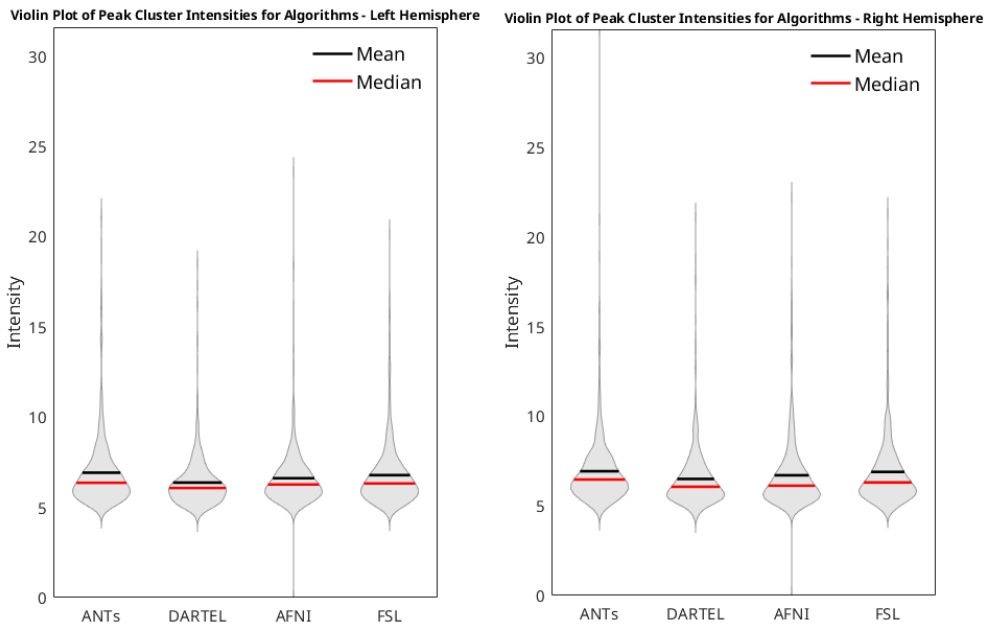

Fig. 4.2 Violin plots of peak cluster intensities in left and right hemispheres for each algorithm (ANTs, DARTEL, AFNI, FSL). ANTs and AFNI show broader intensity distributions, while DARTEL and FSL display narrower, more consistent ranges. Black lines indicate mean values, and red lines indicate medians.

In conclusion, the visual analyses reinforce the importance of algorithm selection in neuroimaging studies, especially those examining hemispheric differences or brain lateralisation. ANTs and AFNI, with their broad intensity ranges, are potentially more sensitive to subtle variations but also more susceptible to variability. Conversely, DARTEL and FSL provide more stable detection profiles, with FSL balancing sensitivity and consistency. These insights suggest that DARTEL and FSL may be particularly suitable for studies needing reliable, cross-hemispheric comparisons, while ANTs and AFNI might be chosen for investigations requiring heightened sensitivity to activation variability.

4.4 Statistical Insights and Interpretation of Results

Statistical analyses of Peak Activation Intensities and visual summaries including non-parametric tests are provided in this section.

4.4.1 Analysis of Skewness and Kurtosis

This section provides a comprehensive comparison of skewness and kurtosis metrics for each neuroimaging algorithm (ANTs, DARTEL, AFNI, and FSL) in detecting Peak Activation Intensities in rs-fMRI data. High skewness indicates asymmetric distributions, potentially introducing a bias toward either low or high activation values, whereas high kurtosis reflects sharp peaks and heavy tails, signifying variability and the presence of outliers.

Table 4.4 summarises these metrics for each algorithm. ANTs demonstrated particularly high kurtosis (e.g., 39.0803 in the right hemisphere), implying susceptibility to extreme activation values, enhancing its capability to detect intense activations but also potentially increasing variability. DARTEL, with its lower skewness and kurtosis values, showed a more symmetric and consistent distribution, beneficial for studies where stability is prioritised. AFNI's broader distribution and higher skewness indicate its tendency to capture a wider range of intensities, whereas FSL's moderate skewness and kurtosis values suggest a balanced sensitivity and consistency.

The non-normality observed in the high skewness and kurtosis values across all algorithms necessitates the use of non-parametric statistical methods. The detection profiles indicated that ANTs and AFNI are effective in capturing variability, whereas DARTEL and FSL provide more balanced detection with less variability, offering potential advantages depending on the study's objective.

Table 4.4 Descriptive statistics for skewness and kurtosis across algorithms, indicating distribution characteristics and sensitivity to intensity extremes.

Metric	ANTs	DARTEL	AFNI	FSL
Skewness (Left)	3.1411	3.3743	3.6593	2.7429
Skewness (Right)	4.3537	3.4518	2.7476	2.9012
Kurtosis (Left)	16.4160	21.8712	27.3054	13.3164
Kurtosis (Right)	39.0803	21.8481	14.2493	14.2793

Practical Implications for Algorithm Selection

Based on the skewness, kurtosis, and correlation findings, the suitability of each algorithm for different neuroimaging applications is summarised below:

- **ANTs**: With high sensitivity to intense activations, ANTs is suited for studies focused on localised, high-activity regions, though care must be taken regarding its susceptibility to outliers.
- **DARTEL**: Best suited for longitudinal or multi-session studies, where symmetric and low variability detection is crucial for stability.
- **AFNI**: AFNI's broad distribution and high skewness make it well-suited for exploratory studies involving a wide range of activation intensities.
- **FSL**: A balanced approach with moderate variability makes FSL appropriate for studies needing consistent detection across a broad spectrum of intensities.

These algorithm characteristics emphasise the importance of careful algorithm selection in rs-fMRI analyses, particularly when the detection of subtle functional changes or lateralisation is essential.

4.4.2 Statistical Analysis Using Non-Parametric Tests

Non-parametric tests, such as Spearman's rank correlation, Mann-Whitney U, and Wilcoxon Signed-Rank tests, were employed to assess hemispheric differences effectively, especially given the non-normal nature of the data distributions.

Table 4.5 summarises the adjusted Spearman correlation coefficients, showcasing the varying degrees of correlation between hemispheric intensities detected by each algorithm. AFNI and ANTs exhibit strong correlations, indicative of high sensitivity to symmetrical

bilateral activation, while DARTEL and FSL demonstrate moderate correlations, providing useful insights for studies focusing on more nuanced hemispheric functions.

This analysis reveals that AFNI yields the highest correlation (ρ = 0.66346, $\rho < .001$), suggesting that activation intensities detected in one hemisphere are closely mirrored in the other hemisphere, implying a high degree of hemispheric agreement in brain activity as identified by AFNI. ANTs (ρ = 0.61647, $\rho < .001$), indicates a strong bilateral relationship in activation intensities, although marginally less similar than AFNI's findings. Furthermore, DARTEL and FSL exhibit moderate correlations (ρ = 0.52181 and ρ = 0.53823, respectively; both $\rho < .001$), suggesting that while there is still a significant association between hemispheric intensities, the relationship is less direct compared to the stronger correlations observed with AFNI and ANTs. These variations in correlation strength, as outlined by ρ values, and their corresponding highly significant adjusted p-values, emphasise the nuanced efficacy of each algorithm in capturing the inherent symmetry or asymmetry of brain function across hemispheres.

Table 4.5 Adjusted Spearman's rank correlation coefficients and ρ-values for fMRI registration algorithms are detailed. The ρ (rho) values span from moderate to strong, with the adjusted p-values reflecting high statistical significance for each algorithm.

Algorithm	ρ	**Adjusted ρ value**
ANTs	0.61647	2.87×10^{-85}
DARTEL	0.52181	6.61×10^{-57}
AFNI	0.66346	2.58×10^{-103}
FSL	0.53823	3.40×10^{-61}

The strong correlations observed with AFNI and ANTs suggest that these algorithms are particularly sensitive to detecting consistent patterns of brain activity across hemispheres, potentially making them more suitable for studies focused on identifying bilateral neuronal activation or lateralisation effects. Conversely, the moderate correlations found with DARTEL and FSL, while still statistically significant, indicate a less pronounced hemispheric alignment, which might be preferable in research contexts where nuanced differences between hemispheres are of interest. Hence, the selection of a specific algorithm can profoundly influence the interpretation of hemispheric associations in neuroimaging data, emphasising the necessity for careful algorithmic choice.

The non-parametric Mann-Whitney U Test (Table 4.6) identified significant hemispheric differences for ANTs (U = 1.382e-09, r = 0.16073) and AFNI (U = 4.4707e-10, r = 0.16489), indicating a difference in central tendency of peak intensities between the hemispheres. Conversely, DARTEL showed no significant hemispheric difference (U = 1, r = -0.027746),

suggesting a balanced detection of peak intensities across hemispheres. FSL displayed a marginal hemispheric difference (U = 0.044285, r = -0.074119), implying a subtle but statistically significant variance.

Table 4.6 Adjusted Mann-Whitney U test ρ-values and effect sizes for each algorithm are presented. This table summarises the Mann-Whitney U test results, offering insights into hemispheric differences in peak activation intensities as detected by various algorithms. The adjusted ρ-values and effect sizes for ANTs, DARTEL, AFNI, and FSL highlight significant variations in their performance across hemispheres.

Algorithm	**Adjusted ρ value**	**Effect Size**
ANTs	1.382×10^{-9}	0.16073
DARTEL	1	-0.027746
AFNI	4.4707×10^{-10}	0.16489
FSL	0.044285	-0.074119

The Wilcoxon Signed-Rank Test (Table 4.7) further supported these findings. ANTs (W = 6.4656e-29, r = 0.28244) and AFNI (W = 1.7356e-27, r = 0.27526) exhibited highly significant median differences between hemispheres, indicating that peak intensities are unevenly distributed within each hemisphere. In contrast, DARTEL (W = 0.41947, r = -0.055061) and FSL (W = 0.00011416, r = -0.1112) did not display significant median differences, suggesting a more symmetrical distribution of peak intensities between hemispheres.

Table 4.7 Adjusted Wilcoxon signed-rank test ρ-values and effect sizes for each algorithm are provided. This table details the results of the Wilcoxon signed-rank test, offering insights into the median differences in peak activation intensities between hemispheres for each of the algorithms studied. The effect sizes quantify the magnitude of these differences.

Algorithm	**Adjusted ρ value**	**Effect Size**
ANTs	6.4656×10^{-29}	0.28244
DARTEL	0.41947	-0.055061
AFNI	1.7356×10^{-27}	0.27526
FSL	1.1416×10^{-4}	-0.1112

These findings highlight that the choice of non-rigid registration algorithm can significantly affect the interpretation of hemispheric differences in neuroimaging studies. Algorithms that exhibit significant hemispheric differences may be advantageous for studies focusing on lateralised brain functions, whereas those showing no significant differences may be more suitable for studies requiring balanced hemispheric analysis.

Table 4.8 Kruskal-Wallis test ρ-values for hemispheric differences in neuroimaging are shown. This table summarises the Kruskal-Wallis test results, evaluating the statistical differences in peak cluster intensities between the left and right hemispheres for each non-rigid registration algorithm. Lower ρ-values denote more significant hemispheric differences.

Hemisphere	**ρ Value**
Left	3.0109×10^{-7}
Right	4.9079×10^{-10}

4.4.3 Visual Summary of Findings: Histogram and Scatter Plot Analysis

For the analysis of hemispheric correlations using non-rigid registration algorithms, scatter plots were generated to visualise the relationship between peak activation intensities across the left and right hemispheres. These scatter plots are supplemented by Spearman's rank correlation coefficients, providing a non-parametric measure of association that does not assume linearity in the relationship between hemispheric intensities. The line of best fit indicates the overall trend and direction of the correlation, with points closer to the line signifying a stronger relationship between hemispheric activation intensities.

This section provides a detailed visual analysis of the distribution and correlation of Peak Activation Intensities across hemispheres, using both histograms and scatter plots to interpret the performance and characteristics of each algorithm in rs-fMRI.

Histogram Analysis:

The histograms in the Figure 4.3 illustrate the frequency distribution of detected peak intensities across different algorithms, revealing specific patterns of detection sensitivity and biases:

- **AFNI:** Displays a right-skewed distribution for both left and right hemispheres, with Peak Activation Intensities clustered towards lower values. This indicates a potential underestimation of higher activation levels, and a bias towards detecting more modest intensities. The mean and median are both shifted towards the lower end of the intensity scale, reinforcing the skewness and the bias towards detecting less pronounced activations.

- **ANTs:** Shows a relatively symmetric distribution around the central values, suggesting a balanced detection capability that is less biased towards any particular intensity range. This more even detection profile implies that ANTs might be well-suited for

applications that require a comprehensive representation of the intensity variance, reliably capturing both medium and high levels of activation.

- **DARTEL:** Exhibits a narrow, symmetric distribution, indicating high precision and minimal variability. Such a distribution profile is beneficial for studies that require consistency and stability in detection across multiple sessions, especially in longitudinal fMRI research where replicable measurements are crucial.
- **FSL:** Presents a slightly right-skewed distribution, similar to ANTs, suggesting that while there is a balanced detection capability, FSL may be somewhat less variable in detecting high-intensity activations. This characteristic is particularly useful for studies aiming for a consistent detection performance across both lower and higher intensity ranges without extreme variability.

These histograms collectively highlight differences in each algorithm's balance of sensitivity and variability. AFNI shows broader variability, whereas ANTs, DARTEL, and FSL provide more stable and consistent detection profiles.

Scatter Plot Analysis:

Figure 4.4 presents scatter plots visualising the correlation between Peak Activation Intensities across the left and right hemispheres for each algorithm:

- **AFNI:** Shows a high Spearman correlation ($\rho = 0.66$), indicating strong bilateral symmetry in detected activations. This consistency makes AFNI suitable for studies requiring accurate identification of symmetric brain functions.
- **ANTs:** Displays a strong but slightly dispersed correlation ($\rho = 0.62$), suggesting that while ANTs captures bilateral patterns effectively, it may also be sensitive to specific patterns or outliers, which could be beneficial in targeted functional analyses.
- **DARTEL:** Exhibits a moderate correlation ($\rho = 0.52$), reflecting a consistent yet somewhat varied detection approach. This makes DARTEL well-suited for broader interpretations of hemispheric activation without overemphasising extremes.
- **FSL:** Shows moderate but consistent correlation ($\rho = 0.54$), suggesting a reliable inter-hemispheric detection capability. This consistency is advantageous for studies that require balanced detection without high sensitivity to variability.

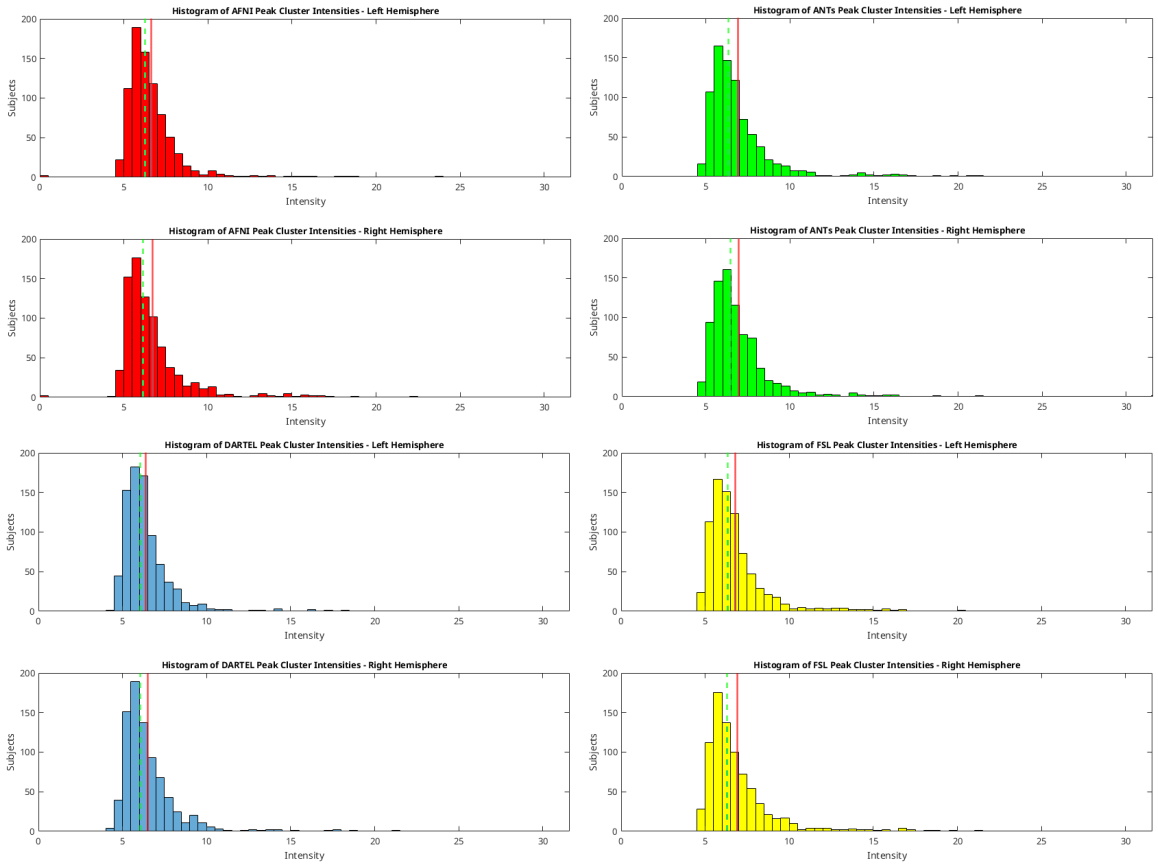

Fig. 4.3 Histograms of Peak Activation Intensities for each algorithm (AFNI, ANTs, DARTEL, FSL) across both hemispheres. Each algorithm shows a distinct distribution pattern: AFNI and FSL exhibit slight right-skewness, ANTs presents a more balanced distribution, and DARTEL demonstrates a narrow, symmetric profile. These patterns reflect the algorithms' differing sensitivities and detection consistencies in rs-fMRI data. Colours represent each algorithm: AFNI (Red), ANTs (Green), DARTEL (Blue), FSL (Yellow).

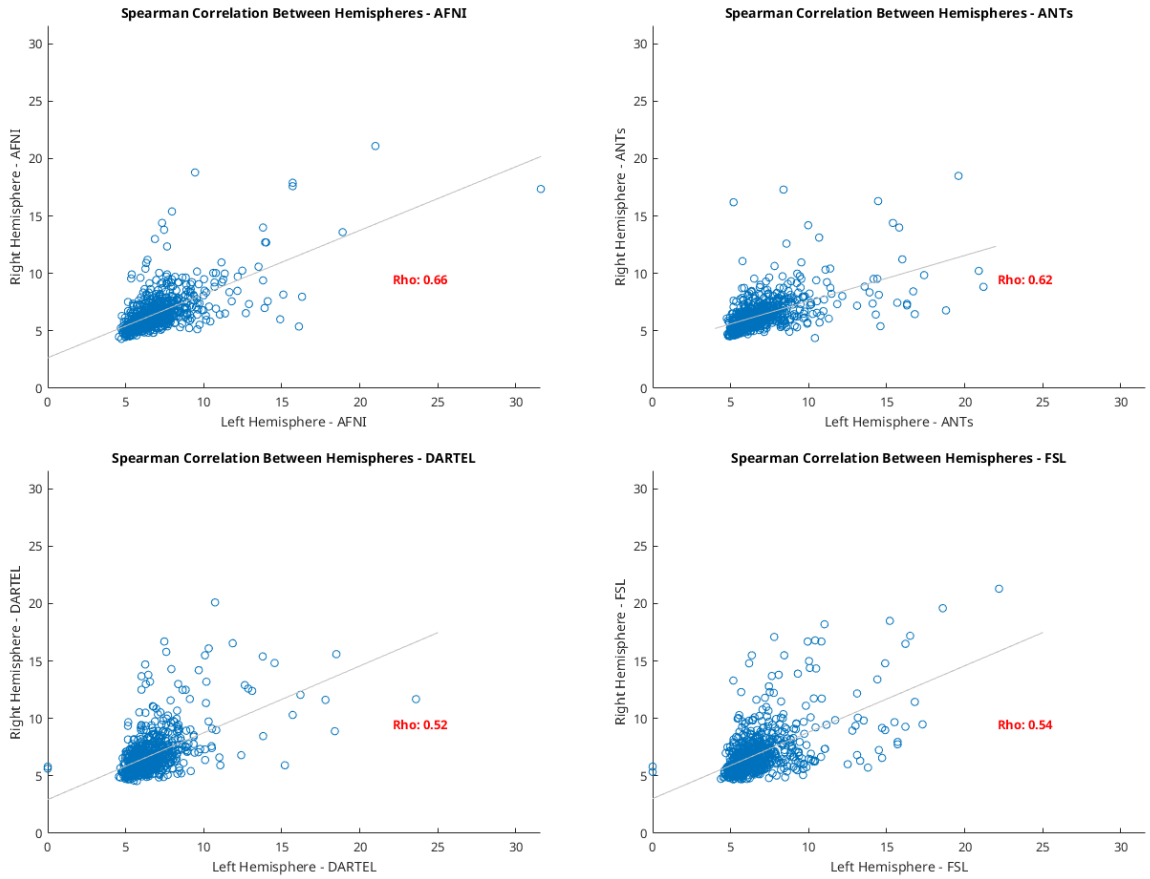

Fig. 4.4 Scatter plots of Spearman's rank correlation between hemispheric activation intensities for each algorithm (AFNI, ANTs, DARTEL, FSL). Correlation coefficients (ρ) indicate varying levels of bilateral consistency: AFNI ($\rho = 0.66$), ANTs ($\rho = 0.62$), DARTEL ($\rho = 0.52$), and FSL ($\rho = 0.54$). These values reflect each algorithm's sensitivity to detecting consistent activation patterns across hemispheres.

The findings highlight key differences in how each algorithm detects Peak Activation Intensities. AFNI and ANTs demonstrate strong correlations in bilateral symmetry and a robust detection profile for intense activations, making them ideal for studies focused on lateralisation or functional symmetry. DARTEL and FSL provide more balanced detection profiles, with moderate sensitivity, suited for studies requiring a nuanced understanding of hemispheric activity without an emphasis on extreme variability.

Bar Graph Analysis of Hemispheric Differences

The Wilcoxon Signed-Rank test results, summarised in Figure 4.5, illustrate the mean differences in Peak Activation Intensities between left and right hemispheres for each algorithm. Positive values indicate a left-hemisphere bias, while negative values reflect a right-hemisphere preference.

This graph reveals subtle but distinct biases in each algorithm's sensitivity to hemispheric intensity differences. ANTs and AFNI show a slight left-hemispheric bias, suggesting they may be particularly sensitive to left-dominant activity patterns, which is relevant for studies on lateralised functions like language [63]. DARTEL, showing near-zero bias, is well-suited for applications requiring balanced detection across hemispheres, such as symmetrical resting-state analyses. FSL's slight right-hemisphere bias may make it more attuned to right-dominant processes, like spatial attention. Understanding these biases enables more informed algorithm selection in neuroimaging studies, ensuring that the chosen algorithm aligns with the specific lateralisation focus of the research.

4.5 Synthesis of Results and Recommendations

This section provides a brief discussion of results presented in the previous sections.

4.5.1 Implications for Neuroimaging Research

The findings in this chapter highlight the crucial role of algorithm selection in rs-fMRI studies, especially regarding Peak Activation Intensity detection and hemispheric asymmetry evaluation. The observed variability in performance among AFNI, ANTs, DARTEL, and FSL demonstrates that different algorithms cater to different research needs. For instance, AFNI and ANTs exhibited stronger bilateral symmetry in peak detection, making them particularly suitable for investigations focusing on symmetrical neural functions, such as those involving the Default Mode Network (DMN) or interhemispheric communication [75, 80]. Conversely,

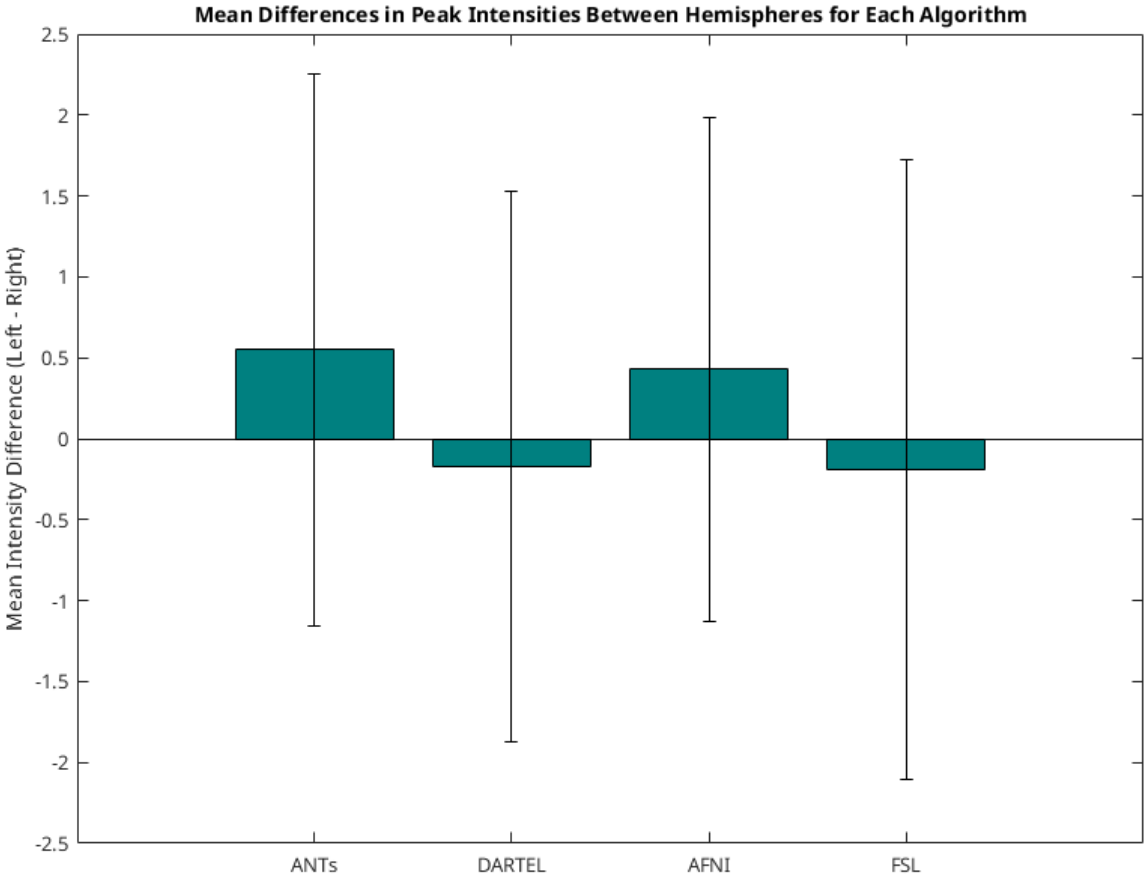

Fig. 4.5 Bar chart illustrating the mean differences in Peak Activation Intensities between hemispheres for each algorithm (ANTs, DARTEL, AFNI, and FSL). Positive values indicate higher mean intensities in the left hemisphere, while negative values suggest a higher right hemisphere intensity. This visualisation helps in understanding the hemispheric bias and performance variability of each algorithm.

DARTEL and FSL's balanced detection profiles suggest their utility for more generalised whole-brain analyses without an inherent hemispheric bias, aligning well with studies that require broad neural activity characterisation [8, 82].

These insights contribute to the broader discourse within neuroimaging by reinforcing that algorithmic choice impacts data preprocessing, spatial normalisation, and resulting functional connectivity analyses [7, 178]. Such findings support the notion that algorithm-induced biases, if unaccounted for, can influence the scientific conclusions drawn from neuroimaging studies [13, 23]. Therefore, the study's results highlight the importance of selecting algorithms in a way that aligns with specific research goals, particularly in rs-fMRI contexts, where functional connectivity, symmetry, and network-specific activity are of interest [25, 82]. By demonstrating that different algorithms yield varying sensitivity and specificity, this study provides empirical support for carefully aligning algorithm selection with intended neuroimaging outcomes.

4.5.2 Recommendations for Future Work

Based on these findings, several directions are recommended to improve algorithmic robustness in rs-fMRI research. Expanding the testing of algorithms across a broader array of datasets, including high-resolution and task-specific datasets, could reveal further insights into their stability and sensitivity [179, 102]. Parameter optimisation within each algorithm should be pursued to tailor detection sensitivity for specific brain regions, potentially reducing artifacts from default settings [180, 30].

Moreover, the integration of multiple algorithms within a hybrid framework is an emerging area that could combine the strengths of each approach, thus enhancing both sensitivity and specificity for complex analyses [24, 97]. Such integrative approaches align with recent methodological advancements and can offer more refined analyses of neural connectivity and variability in rs-fMRI, addressing the limitations of individual algorithms. This hybridisation may also benefit from Machine Learning (ML) techniques, which can dynamically adapt algorithm parameters to dataset-specific features, potentially improving the reproducibility and interpretability of neuroimaging findings [3, 22].

In summary, future research should prioritise broader dataset validation, parameter optimisation, and algorithmic integration to enhance rs-fMRI methodological robustness. These efforts will contribute toward establishing more consistent and reliable neuroimaging protocols, addressing current challenges in algorithm selection and expanding the interpretative power of rs-fMRI analyses.

4.6 Chapter Summary

Chapter 4 presents a comprehensive evaluation of Peak Activation Intensity detection across hemispheres in rs-fMRI data, comparing four state-of-the-art non-rigid registration algorithms: ANTs, DARTEL, AFNI, and FSL. The chapter emphasises the need for reliable and sensitive registration methods in neuroimaging by examining each algorithm's statistical sensitivity and stability. The analytical approach leverages robust metrics—including median, IQR, and SD—to address variability, thereby enhancing our understanding of the distinct behaviours of each method under different neuroimaging conditions.

Results reveal that while ANTs and AFNI exhibit high sensitivity to activation peaks, they also display considerable variability, particularly in the right hemisphere, which could potentially skew interpretations of hemispheric specialisation in studies that prioritise functional lateralisation. DARTEL and FSL, in contrast, offer more stable intensity detections across hemispheres, thus presenting balanced alternatives with reduced bias. These findings highlight the importance of algorithm selection in neuroimaging, as each algorithm's inherent characteristics could lead to distinct neuroanatomical interpretations, particularly in studies focused on hemispheric asymmetry.

Further, the chapter addresses the issue of non-normal data distributions—frequently characterised by high skewness and kurtosis—which necessitated non-parametric statistical methods such as the Mann-Whitney U and Kruskal-Wallis tests. The use of these methods reflects a careful adjustment for data irregularities, ensuring a rigorous analysis that strengthens the reliability of the findings.

These findings contribute to a nuanced understanding of algorithmic performance in rs-fMRI studies. In line with the reviewed literature, the chapter highlights how variability in peak activation detection can influence interpretations of neuroimaging data, reinforcing the need for algorithm refinement in studies of functional connectivity and brain asymmetry [181, 25]. This exploration enriches ongoing discussions in the field by demonstrating that algorithm selection must be context-dependent, balancing sensitivity with consistency to optimise outcomes in specific research designs.

Chapter 5 builds upon these insights by broadening the focus from isolated peak intensities to an examination of significant clusters and network integrity. This shift in focus allows for a network-level analysis that contextualises peak intensity findings within broader spatial activation patterns, providing a more comprehensive picture of brain activity patterns and functional connectivity. This forthcoming analysis leverages the insights from Chapter 4 to investigate how peak intensity variations within significant clusters influence functional

networks, offering further guidance on algorithm selection based on network-level sensitivity requirements.

Chapter 5

Significant Clusters-Based Network Integrity Analysis

Chapter 4 provided a detailed analysis of Peak Activation Intensities, a univariate metric, as an initial perspective on functional connectivity. In contrast, Chapter 5 introduces a complementary approach through the analysis of Significant Clusters within Resting-State Functional Magnetic Resonance Imaging (rs-fMRI) data. This chapter focuses on evaluating how various registration algorithms affect the identification and interpretation of these clusters, which are essential for understanding the brain's functional integrity at the network level.

The primary objective of this chapter is to examine Significant Clusters as a marker of network integrity. Significant Clusters, identified based on statistical thresholds and spatial coherence criteria, offer an alternative view on functional connectivity that highlights regional co-activation and inter-regional coherence across hemispheres. This approach incorporates the Gaussian Random Field (GRF) theory for determining cluster-wise ρ-values [61], facilitating an analysis that moves beyond single-peak activations toward broader network characteristics.

Understanding network integrity through Significant Clusters is crucial for evaluating the robustness of non-rigid registration algorithms, as these algorithms impact the reproducibility of detected functional regions across scans. By assessing the consistency of significant cluster identification across algorithms, this chapter aims to inform future selections of neuroimaging processing pipelines tailored to specific neuroscientific investigations, particularly those involving rs-fMRI data.

This chapter serves as a transition from the individual-level analyses of activation patterns in Chapter 4 to the more complex inter-subject comparisons in Chapter 6. By systematically

examining algorithmic reliability in detecting Significant Clusters, we lay a foundation for subsequent analyses on inter-subject variability and group-level network interpretations, contributing to a more comprehensive framework for evaluating functional connectivity. In this way, Chapter 5 acts as a bridge, connecting univariate activation insights with the multivoxel approaches introduced in later chapters, ensuring continuity within the thesis' broader methodological progression.

5.1 Methodology for Significant Clusters-Based Analysis

In this chapter, we extend the dataset and preprocessing protocols described in Chapter 3, where the neuroimaging data, comprising rs-fMRI scans, were prepared for subsequent analysis. Briefly, the dataset includes 815 subjects, whose Functional Magnetic Resonance Imaging (fMRI) scans were processed to enhance spatial consistency and to remove artifacts, following standard preprocessing steps such as motion correction, normalisation, and smoothing. Additionally, non-rigid state-of-the-art registration algorithms, specifically FMRIB Software Library (FSL), Advanced Normalisation Tools (ANTs), Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL), and Analysis of Functional NeuroImages (AFNI), were applied to align the brain images within a standard coordinate Montreal Neurological Institute 152 (MNI152) space as outlined in earlier chapters.

5.1.1 Cluster-Based Analysis Methods

The methodologies presented here extend from traditional clustering approaches by specifically identifying significant spatial clusters within rs-fMRI data. Unlike Machine Learning (ML)-based clustering algorithms briefly discussed in Chapter 3, which aim to partition data based on intrinsic patterns or groupings, the cluster analysis here is hypothesis-driven and relies on statistical criteria for network integrity evaluation.

Significant Clusters Identification

To identify Significant Clusters, we employed FSL FMRI Expert Analysis Tool (FEAT), which facilitates the generation of spatially coherent clusters based on voxel-wise statistical thresholds [61]. Significant Clusters in rs-fMRI are identified through FSL FEAT using spatial coherence and statistical criteria, as defined by GRF theory. Clusters surpassing a defined cluster-wise ρ-value threshold are deemed significant, accounting for spatial dependencies within the data and correcting for multiple comparisons [99, 7]. This methodology focuses

on biologically meaningful activation patterns, enhancing interpretability by defining clusters as functionally cohesive units within the brain's Resting-State Networks (RSNs).

Through FSL FEAT, clusters identified are not merely statistically significant but spatially contiguous, enhancing the interpretability and reliability of the observed activation patterns across different registration algorithms.

Network Integrity Assessment

The integrity of identified clusters is evaluated based on their spatial coherence and reproducibility across registration algorithms. This involves assessing metrics such as cluster size consistency, localisation precision, and overlap of identified clusters across algorithms, emphasising structural reliability rather than data-driven grouping. This analysis aims to ensure that the clusters identified genuinely reflect stable network structures within rs-fMRI, supporting the study's focus on network-level functional integrity rather than exploratory pattern discovery. Statistical measures, including Spearman's Rank Correlation and Mann-Whitney U tests, facilitate the assessment of consistency and hemispheric differences in cluster localisation across different algorithms [30, 182].

This approach emphasises the methodological rigour in defining significant rs-fMRI clusters and highlights the use of FSL FEAT for reproducibility and network fidelity, contrasting with the ML clustering methods discussed in Chapter 3.

5.1.2 Functional Connectivity via Significant Clusters

Functional connectivity offers insights into how distinct brain regions communicate, particularly within RSNs, where synchrony among neural areas can reveal intrinsic functional connectivity patterns vital to cognitive and affective processing [82, 23]. Building on prior analyses of Significant Clusters, this chapter aims to explore how each registration algorithm (FSL, ANTs, DARTEL, and AFNI) captures these broader functional connectivity patterns. The focus here shifts from individual cluster identification to evaluating the network integrity these clusters represent across brain hemispheres, offering a more comprehensive view of functional connectivity across hemispheres [25, 177].

To assess hemispheric consistency and variability, statistical methods such as the Mann-Whitney U and Wilcoxon signed-rank tests are employed, providing an unbiased evaluation of hemispheric performance in functional connectivity mappings [173, 174]. These methods are particularly suitable for non-normally distributed neuroimaging data, enhancing the robustness of findings regarding hemispheric asymmetries and network integrity [25, 161].

Each algorithm applies distinct computational techniques and spatial transformations, leading to variability in capturing functional connectivity patterns. This comparative analysis highlights the potential strengths and limitations of each approach in representing broader functional connectivity patterns, especially in terms of hemispheric balance and consistency. Such insights are crucial as they directly impact the reliability of neuroimaging studies, guiding the selection of algorithms to enhance the robustness of functional connectivity analyses [22, 65].

5.2 Evaluation of Significant Clusters

This section examines how the four state-of-the-art non-rigid registration algorithms—FSL, ANTs, DARTEL, and AFNI—differentially impact the identification and variability of Significant Clusters, focusing on their effect within the left and right hemispheres. We investigate how these algorithms modulate the detection of clusters in the Control Network, emphasising the implications for interpreting the Default Mode Network (DMN) in resting-state neuroimaging.

The same seed region identified in Chapter 4 within the left hemisphere, encompassing the Paracingulate Gyrus and divisions of the Cingulate Gyrus, remains central to this analysis. This section expands upon prior findings by focusing on the Significant Clusters associated with this Control Network region, assessing the influence of registration algorithms on both functional connectivity and functional integrity metrics.

5.2.1 Atlas Measurement

The correlation strength to the seed region and each algorithm's influence on this correlation were evaluated using the Harvard-Oxford cortical and subcortical structural atlas in MNI152 space. The cortical atlas covers 96 regions (48 per hemisphere), and the subcortical atlas includes 17 regions encompassing major brain structures such as the thalamus, caudate, pallidum, hippocampus, amygdala, accumbens, and brainstem [183, 184, 88, 185].

The selection of the Harvard-Oxford atlas is justified due to its comprehensive coverage and established use in neuroimaging studies [181, 186]. Its detailed characterisation of cortical and subcortical structures facilitates accurate region-based correlation analyses, which are crucial for understanding the functional connectivity patterns relevant to our research [187]. By using an atlas that is standardised to MNI152 space, we ensure compatibility with a wide range of neuroimaging data and improve the reproducibility and comparability of

our results across studies [98]. This standardisation allows for more precise identification of regions implicated in cognitive and behavioural processes [175].

5.2.2 Hemispheric Analysis of Significant Clusters

This subsection analyses each algorithm's performance in detecting Significant Clusters within each hemisphere, exploring descriptive statistics that characterise cluster variability and detection consistency.

In the left hemisphere (Table 5.1), DARTEL detected a higher mean number of clusters (M = 14.60) compared to ANTs (M = 10.34), AFNI (M = 12.31), and FSL (M = 10.72), with median values supporting this pattern. DARTEL also showed greater variability, indicated by the highest standard deviation (SD = 7.26) among the algorithms. Although DARTEL's sensitivity appears advantageous, the increased variability suggests a potential trade-off, as incidental or less relevant clusters may be identified.

In the right hemisphere (Table 5.2), a similar trend emerges with DARTEL leading in mean cluster detection (M = 12.78). The skewness values indicate a rightward tail in cluster distribution for ANTs (Skewness = 0.72) and FSL (Skewness = 0.72), implying a subset of cases where a higher number of clusters were detected, especially for these two algorithms. Notably, kurtosis values in the right hemisphere, particularly for ANTs (Kurtosis = 3.24), indicate a more peaked distribution around the mean compared to the left hemisphere.

Table 5.1 Descriptive statistics of significant cluster detection in the *Left Hemisphere*, including mean, median, variability, and distribution characteristics for four registration algorithms (ANTs, DARTEL, AFNI, and FSL). This table highlights DARTEL's higher mean detection and variability compared to other algorithms.

Test	**ANTs**	**DARTEL**	**AFNI**	**FSL**
Mean	10.3350	14.5988	12.3141	10.7190
Median	10	14	12	10
Std Dev	5.4617	7.2639	5.8035	6.0598
Min	1	1	0	1
Max	29	42	31	33
Range	28	41	31	32
IQR	8	10	8	8
Skewness	0.4234	0.4598	0.2925	0.7182
Kurtosis	2.7463	2.8541	2.6804	3.1187

The findings indicate that DARTEL's higher mean number of detected clusters, particularly in the left hemisphere, showcases its sensitivity for identifying functional connectivity.

Table 5.2 Descriptive statistics of significant cluster detection in the *Right Hemisphere*, showing mean, median, variability, and distribution metrics for the registration algorithms (ANTs, DARTEL, AFNI, and FSL). Key observations include ANTs' and FSL's skewness, suggesting outliers in cluster detection.

Test	**ANTs**	**DARTEL**	**AFNI**	**FSL**
Mean	9.9755	12.7755	10.4086	9.4049
Median	9	12	10	8
Std Dev	5.6893	6.9583	6.2732	5.7961
Min	1	1	0	1
Max	30	43	31	29
Range	29	42	31	28
IQR	9	11	10	8
Skewness	0.7189	0.5580	0.5271	0.7156
Kurtosis	3.2412	3.0514	2.5778	2.8912

However, the higher standard deviation suggests that this algorithm, while comprehensive, may include incidental or non-essential clusters, impacting data interpretation in studies where precision is crucial. ANTs and AFNI, with their more moderate mean cluster counts and narrower standard deviations, align with a stable approach suitable for applications emphasising reliability over exhaustive detection.

The skewness values for ANTs and FSL in the right hemisphere imply that these algorithms capture outlier cases with higher numbers of clusters, highlighting their variability in specific contexts. Kurtosis values for ANTs, especially in the right hemisphere, indicate a peaked distribution that can suggest precise identification of focal activations, beneficial for studies on regional connectivity.

Aligning algorithm selection with research goals is essential. DARTEL's robust detection profile can benefit exploratory studies that need comprehensive cluster mapping. In contrast, ANTs and AFNI, with their conservative and balanced cluster detection, may be more suitable for studies where reducing noise and ensuring consistent detection are priorities. These findings emphasise the importance of methodological choices in neuroimaging research focused on hemispheric functional connectivity and network integrity.

Comparing results across hemispheres, DARTEL consistently detects more clusters, albeit with greater variability. This variability could indicate a higher rate of incidental detections. Consistency in Interquartile Range (IQR) values across algorithms and hemispheres shows a stable spread in detection counts, although the right hemisphere has a slightly higher IQR for DARTEL and AFNI, suggesting greater detection variability.

These results highlight DARTEL's sensitivity in detecting Significant Clusters and the unique balance of sensitivity and specificity each algorithm brings to neuroimaging applications, relevant to the reliability and interpretability of studies examining functional connectivity.

5.2.3 Statistical Consistency Across Hemispheres

This section introduces non-parametric tests and presents a summary of hemispheric consistency in functional connectivity patterns, consolidating findings into tables that highlight statistically significant results. The non-parametric tests provide robust insights, accommodating deviations from normality in data distribution.

The distribution of significant cluster detections deviates from the expected normal curve, as evidenced by elevated skewness and kurtosis across all algorithms and hemispheres. Notably, ANTs exhibited a kurtosis of 3.2412 in the right hemisphere, suggesting a leptokurtic distribution. Such deviations from normality necessitate alternative statistical approaches like non-parametric tests for accurate data interpretation, sidestepping the limitations of parametric methods that assume a normal distribution [188, 189].

These variations among algorithms highlight the influence of algorithm selection on neuroimaging outcomes. ANTs and AFNI, showing greater variability, may be optimal for studies requiring sensitivity to a broad range of activations. Conversely, DARTEL and FSL, with more consistent measurements, could be better suited for research where balanced detection is critical. This section sets the stage for the subsequent analysis using non-parametric tests, which further explore the implications of these findings on neuroimaging data interpretation.

The following analysis employs non-parametric tests to examine patterns within fMRI data. Non-parametric tests are crucial as they do not assume a normal distribution, fitting the data's skewed and kurtotic characteristics, as noted. Here, the results of several non-parametric tests are presented, each shedding light on different aspects of the evaluated data.

Hemispheric Symmetry and Algorithmic Performance in Cluster Detection

To evaluate the consistency of significant cluster detection across hemispheres, non-parametric tests were employed, as these methods are less sensitive to deviations from normality, which was observed in our data. The results from these tests provide insights into the symmetry

or asymmetry in cluster detection across the left and right hemispheres, and reveal specific tendencies of each algorithm that could guide their selection in neuroimaging studies.

The Spearman's Rank Correlation coefficients (Table 5.3) measure the strength of the monotonic relationship between the number of Significant Clusters detected in each hemisphere. AFNI exhibits the highest correlation ($\rho = 0.64072$), suggesting a strong alignment in cluster detection across hemispheres with high statistical significance ($p < 0.001$). This result implies that AFNI may offer greater consistency in identifying clusters in both hemispheres, making it a suitable choice for studies where symmetrical functional connectivity across hemispheres is desired. ANTs also shows a significant correlation ($\rho = 0.59472$), though slightly lower than AFNI, indicating a moderately strong hemispheric agreement but with a higher degree of variability. Lower correlations in DARTEL and FSL suggest that these algorithms detect clusters in a more hemisphere-specific manner, which might be relevant for studies focusing on lateralised brain functions.

Table 5.3 Adjusted Spearman's rank correlation coefficients and ρ-values for fMRI registration algorithms. The table summarises the Spearman's rank correlation analysis results, indicating the varying strengths of the monotonic relationships between left and right hemisphere number of Significant Clusters for ANTs, DARTEL, AFNI, and FSL. ρ values range from moderate to strong, and the adjusted ρ-values denote high statistical significance for each algorithm.

Algorithm	ρ	**Adjusted ρ-value**
ANTs	0.59472	6.6205×10^{-78}
DARTEL	0.46607	5.544×10^{-44}
AFNI	0.64072	3.4429×10^{-94}
FSL	0.559953	3.048×10^{-67}

The Mann-Whitney U Test (Table 5.4) further explores hemispheric differences by examining the effect sizes for each algorithm. ANTs displays a negative effect size (-0.29524), suggesting a notable asymmetry, with a higher number of clusters detected in one hemisphere over the other. This asymmetry indicates that ANTs may be particularly sensitive to hemisphere-specific functional connectivity patterns, potentially benefiting studies examining hemispheric specialisation. In contrast, DARTEL's smaller positive effect size (0.15202) suggests more balanced detection across hemispheres, reinforcing its suitability for studies requiring consistent bilateral detection. AFNI and FSL show moderate differences, but their smaller effect sizes compared to ANTs indicate less hemispheric bias.

The Wilcoxon Signed-Rank Test (Table 5.5) results align with the Mann-Whitney U findings, with ANTs and AFNI again showing substantial median differences between

Table 5.4 Adjusted Mann-Whitney U test ρ-values and effect sizes for each algorithm. This table presents the results of the Mann-Whitney U test, providing insights into hemispheric differences in the number of Significant Clusters as detected by different algorithms. Adjusted ρ-values and effect sizes are indicated for ANTs, DARTEL, AFNI, and FSL, highlighting significant variances in their performance across hemispheres.

Algorithm	**Adjusted ρ-value**	**Effect Size**
ANTs	1.4914×10^{-31}	-0.29524
DARTEL	1.3396×10^{-8}	0.15202
AFNI	4.37×10^{-15}	-0.20272
FSL	0.02709	0.077759

hemispheres. ANTs exhibits a significant negative effect size (-0.43129), the largest among the algorithms, reinforcing its sensitivity to hemispheric asymmetry. This sensitivity could be particularly useful in detecting lateralised brain function. In contrast, DARTEL and FSL have smaller effect sizes (0.18221 and 0.12121, respectively), indicating more symmetrical detection, which is advantageous for studies prioritising uniformity across hemispheres.

Table 5.5 Adjusted Wilcoxon signed-rank test ρ-values and effect sizes for each algorithm. This table presents the outcomes of the Wilcoxon signed-rank test, offering insights into the median differences in the number of Significant Clusters between hemispheres for each of the studied algorithms. Effect sizes are included to quantify the magnitude of these differences.

Algorithm	**Adjusted ρ-value**	**Effect Size**
ANTs	1.0612×10^{-66}	-0.43129
DARTEL	3.0216×10^{-12}	0.18221
AFNI	1.3681×10^{-39}	-0.33127
FSL	1.5853×10^{-5}	0.12121

The Kruskal-Wallis Test (Table 5.6) further confirms significant differences in cluster counts between hemispheres, particularly for the right hemisphere. This consistent discrepancy in the right hemisphere across algorithms suggests that right-hemispheric clusters are more variable and might capture distinct aspects of functional connectivity. This finding could be relevant in studies focused on right-hemisphere-dominant functions or pathologies, as it implies potential variations in sensitivity among algorithms to right-hemisphere clusters.

These findings emphasise the significant impact of algorithm selection on the identification of hemispheric differences in brain function. ANTs and AFNI, with their tendency for asymmetric detection, are likely to be advantageous for studies focusing on lateralised cognitive processes or functional asymmetries. Conversely, DARTEL and FSL, with their balanced and consistent detection across hemispheres, may be more appropriate for research

Table 5.6 Kruskal-Wallis test ρ-values for hemispheric differences in neuroimaging. This table presents the Kruskal-Wallis test results assessing the statistical differences in the number of Significant Clusters between the left and right hemispheres for each non-rigid registration algorithm. Lower ρ-values indicate more significant differences between hemispheres.

Hemisphere	ρ-Value
Left	4.0775×10^{-18}
Right	2.2969×10^{-56}

where uniformity and reproducibility across hemispheres are important. This knowledge provides a foundational basis for algorithm selection in neuroimaging studies, aligning methodological choices with specific research goals in cognitive neuroscience.

Distribution Analysis of Functional Connectivity

The histograms (Fig. 5.1) show the frequency distribution of significant cluster detections across hemispheres for each algorithm, highlighting their sensitivity and detection biases. Histograms are particularly valuable in rs-fMRI analysis as they illustrate data distribution patterns, shedding light on central tendencies and skewness, which are essential for interpreting individual differences in functional connectivity [82, 190].

These histograms also complement Peak Activation Intensity analysis. While peak intensity histograms indicate functional connectivity strength, cluster count histograms reveal spatial consistency of detections. Together, they provide a comprehensive view of each algorithm's detection profile [91, 191].

The histogram for the **AFNI** algorithm shows a rightward skew in both hemispheres. The peak of the histogram occurs before the mean, suggesting that the majority of subjects exhibit a lower count of Significant Clusters. The longer right tail implies a small subset of subjects with notably higher cluster counts, indicating that AFNI may have a conservative detection bias, likely favoring fewer false positives while risking an underestimation of clusters, particularly at higher intensities.

The histogram for the **ANTs** algorithm shows a more balanced distribution with slight right skewness, reflecting a moderate overrepresentation of subjects with fewer clusters. The relatively symmetrical appearance of ANTs' distributions indicates a balanced detection capability, positioning it as a middle ground between sensitivity and specificity in cluster detection. This balanced detection may contribute to reliable functional connectivity analyses when a mix of sensitivity and conservative detection is required.

The **DARTEL** algorithm's histograms for both hemispheres demonstrate the most symmetric distribution among all algorithms, with the central peak aligning closely with the mean. This suggests that DARTEL offers consistent detection across subjects, with low variability and minimal skewness. The slight rightward skew observed is less pronounced compared to AFNI and ANTs, indicating a stable and uniform detection performance across varying functional connectivity intensities, which may make it particularly suitable for analyses requiring consistent and balanced measurements across large datasets.

Finally, the histogram for the **FSL** algorithm exhibits a distribution similar to that of ANTs, with a slight rightward skew. This pattern implies a balanced detection tendency with a modest skew towards detecting fewer clusters, aligning it closely with ANTs in terms of detection balance. The spread of the distribution, alongside the peak around the median, supports FSL's potential for consistent cluster detection, though with a slightly conservative edge compared to DARTEL.

Interpretation of Algorithmic Differences Based on Histogram Trends:

- The **AFNI** histograms indicate a conservative detection bias with a rightward skew, implying a tendency to underestimate clusters, especially those of higher intensity, which might reduce false positives but could miss subtle functional connectivity patterns.
- **ANTs** shows a moderate right skew, suggesting balanced detection with slight variability. This algorithm could be suited for studies requiring a moderate approach between sensitivity and specificity.
- **DARTEL** histograms, with minimal skew and high symmetry, imply consistent detection across intensities, indicating it as potentially ideal for applications requiring balanced and stable cluster detection without significant hemispheric bias.
- **FSL** exhibits a similar skew to ANTs, showing balanced detection with a modest skew toward fewer clusters, making it versatile for applications where conservative and consistent detection is preferred.

The histogram distributions illustrate distinct cluster detection profiles for each algorithm, highlighting how these methods align with different research objectives. AFNI's more conservative detection approach could be advantageous for studies requiring stricter control over false positives, ideal for exploratory analyses with a high threshold for significance. DARTEL's symmetrical cluster distribution may serve well in studies needing balanced

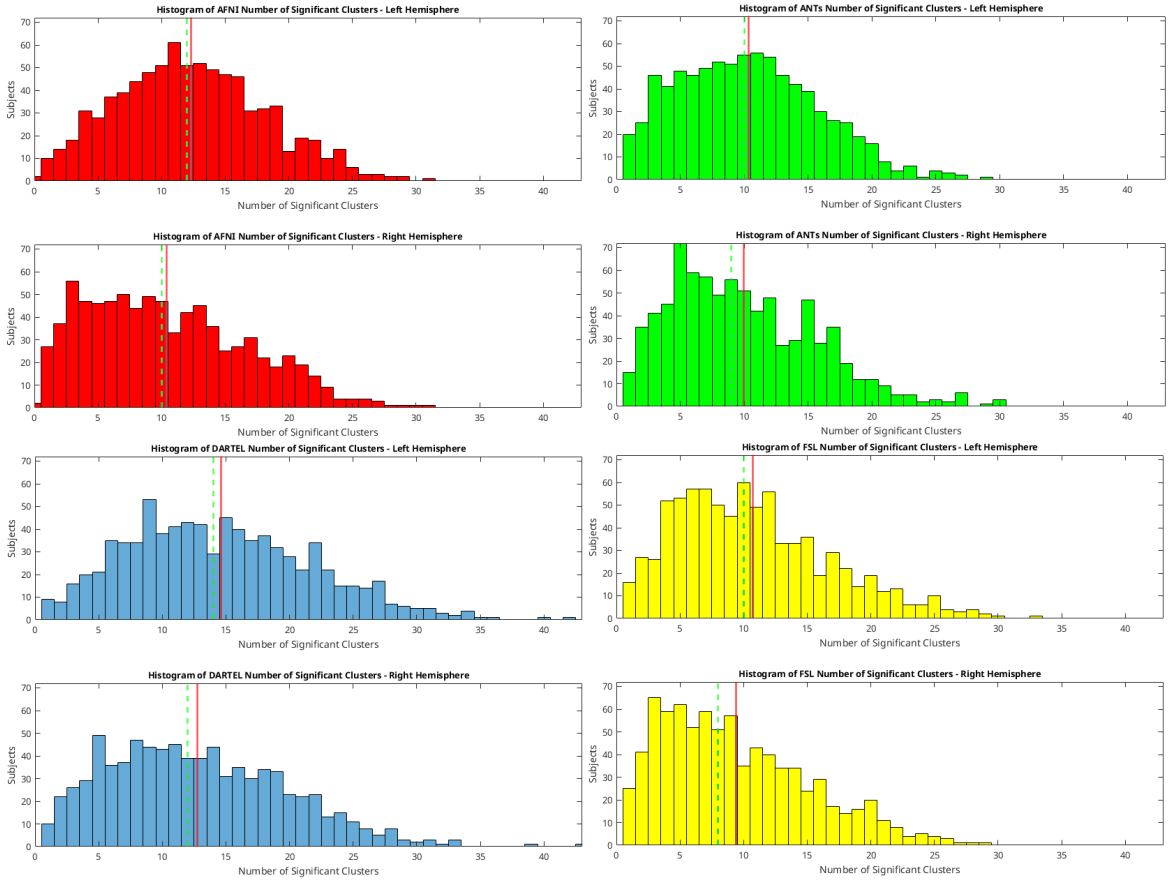

Fig. 5.1 Histograms of the number of Significant Clusters detected in left and right hemispheres for each algorithm (AFNI, ANTs, DARTEL, FSL). AFNI shows a rightward skew, indicating a conservative bias with fewer detected clusters. ANTs and FSL exhibit relatively symmetrical distributions with a slight skew towards fewer clusters, suggesting balanced detection. DARTEL demonstrates the most symmetric distribution, indicating consistent detection across subjects and hemispheres.

inter-hemispheric detection, providing consistency across subject populations. ANTs and FSL, positioned between these extremes, demonstrate a versatile mix of sensitivity and specificity, making them adaptable for varied experimental designs where moderate cluster detection sensitivity is suitable.

Overall, these distributions encourage researchers to select an algorithm based on study-specific goals, such as prioritising detection consistency or specificity in functional connectivity analyses. The histogram comparison, along with supporting descriptive statistics, thus enhances the decision-making process in choosing a neuroimaging registration approach that best matches the intended analytical rigour and focus within resting-state functional connectivity studies.

Hemispheric Functional Connectivity Insights

The exploration of hemispheric differences in functional connectivity highlights the evolutionary adaptation of brain asymmetry to optimise cognitive processing and performance [176]. This section investigates how non-rigid image registration algorithms influence the detection of hemispheric variances in neuroimaging data. The role of these algorithms—ANTs, DARTEL, AFNI, and FSL—in mapping activation intensities across hemispheres is particularly relevant, as their unique computational methodologies may highlight or obscure inherent asymmetries.

As highlighted by Wang et al. [177], hemispheric specialisation contributes to behavioural adaptability, underscoring lateralisation as a functional advantage. This analysis aims to discern how each algorithm's characteristics affect our understanding of intrinsic hemispheric differences, focusing not just on the presence of asymmetry but on the portrayal accuracy across different algorithmic approaches.

The box plots in Figure 5.2 illustrate central tendency and variability in significant cluster detection for each hemisphere. In the left hemisphere, ANTs and DARTEL exhibit higher medians, indicating a propensity to detect more clusters. DARTEL's tighter IQR reflects consistent cluster detection, while ANTs shows wider variability, which may suggest either increased sensitivity or potential noise inclusion. AFNI and FSL, with lower median cluster counts, present fewer outliers in FSL's case, suggesting a more conservative detection approach.

For the right hemisphere, DARTEL and FSL maintain narrow IQR, indicating consistency across subjects. In contrast, AFNI's wider distribution and more outliers imply a sensitivity to extreme values, which could be advantageous for studies focusing on capturing subtle or sparse activation patterns but might also introduce noise.

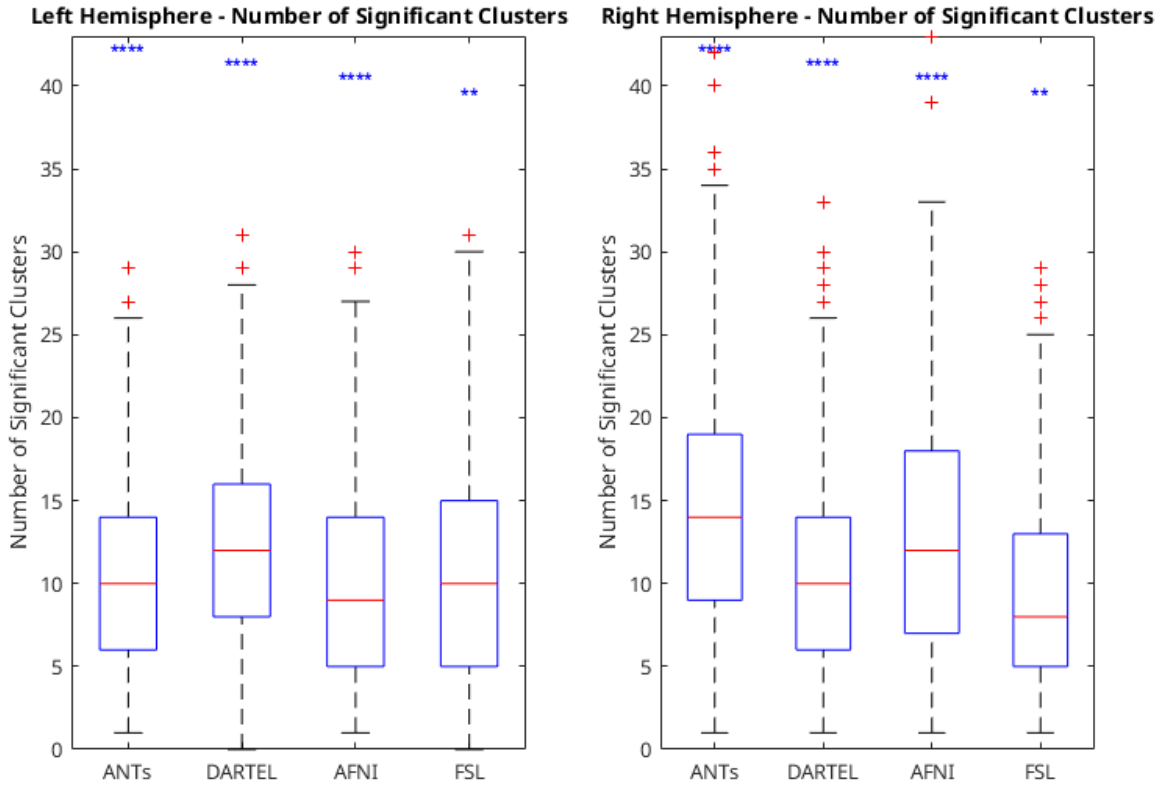

Fig. 5.2 Box plots of the number of Significant Clusters detected across left and right hemispheres for each algorithm. ANTs and DARTEL show higher median values, with DARTEL and FSL exhibiting narrower IQR, indicating greater consistency in cluster detection. AFNI demonstrates broader distributions, reflecting higher variability in cluster counts. Statistical significance between algorithms is marked by asterisks as shown in Table 4.3.

The violin plots in Figure 5.3 provide insights into the density of detected clusters. In the left hemisphere, DARTEL and FSL display narrower distributions around the median, suggesting that detected clusters are densely packed around a central value, likely reducing susceptibility to noise. Conversely, ANTs and AFNI's broader distributions in the right hemisphere imply higher variability in cluster size detection, which may indicate sensitivity to subtle signals but could also reflect noise.

DARTEL stands out for its symmetrical detection performance across both hemispheres, with fewer outliers and tighter distributions, making it a suitable choice for studies requiring consistent inter-hemispheric comparisons. ANTs, while detecting more clusters, presents greater variability, raising questions about specificity in studies sensitive to lateralisation. AFNI's broader distribution reflects a high sensitivity to extreme values, which might be beneficial in detecting sparse or nuanced patterns, though it also poses a risk of incorporating noise. FSL, with moderate IQR and fewer outliers, offers a balanced approach, appropriate for general functional connectivity studies.

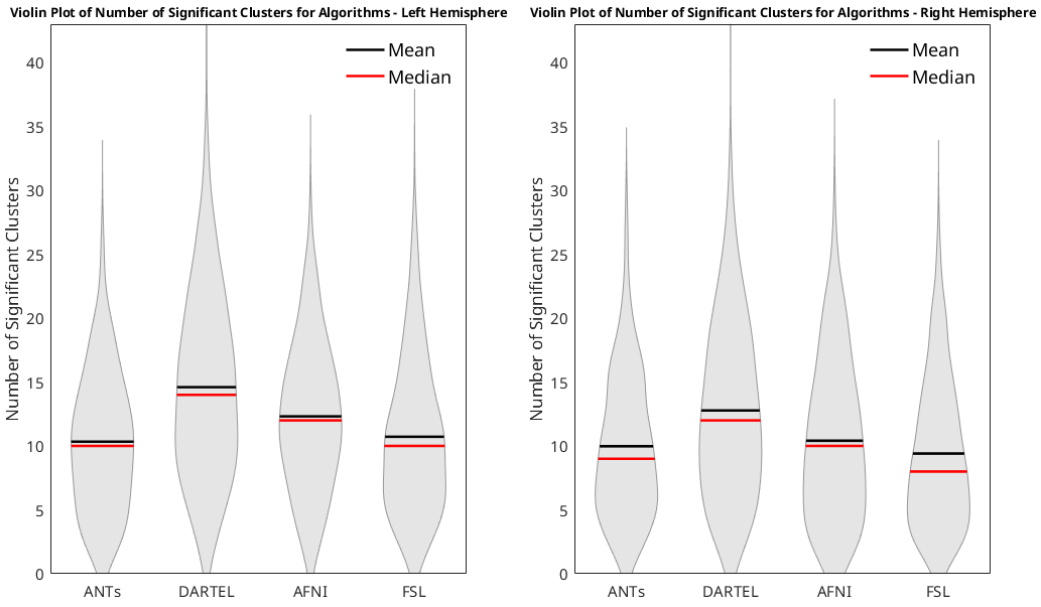

Fig. 5.3 Violin plots of the number of Significant Clusters detected in left and right hemispheres for each algorithm (ANTs, DARTEL, AFNI, FSL). DARTEL and FSL show narrower distributions around the median, indicating consistent cluster detection, while ANTs and AFNI exhibit broader distributions, suggesting higher sensitivity to variability in cluster size. Black lines indicate mean values, and red lines indicate medians.

The observed distribution patterns suggest that algorithm choice should be tailored to specific research objectives in neuroimaging. DARTEL's consistent and balanced detection across hemispheres is advantageous for studies emphasising hemispheric symmetry. In

contrast, ANTs and AFNI may be more suitable for studies where sensitivity to subtle asymmetries is desired, albeit with the caution that these algorithms might introduce noise due to variability. FSL, with its moderate detection profile, emerges as a versatile choice, well-suited for studies requiring a balance between sensitivity and specificity.

In summary, these findings highlight the importance of selecting an appropriate algorithm based on study requirements. Each algorithm's inherent characteristics influence the accurate portrayal of hemispheric functional connectivity, emphasising its critical impact on the interpretation of functional brain imaging results.

Algorithmic Sensitivity and Specificity in Functional Connectivity

To assess the hemispheric consistency of cluster detection using non-rigid registration algorithms, scatter plots are employed (Fig. 5.4) to visualise associations between the number of Significant Clusters in the left and right hemispheres. Accompanied by Spearman's rank correlation coefficients, these plots provide a quantitative measure of monotonic relationships without assuming linearity. The line of best fit in each plot represents the correlation trend, and the proximity of data points to this line reflects correlation strength, integral to interpreting algorithmic performance.

This analysis is not intended to dispute hemispheric asymmetry but rather to assess how different algorithms capture this characteristic. By examining associations between clusters in each hemisphere, we gain insights into each algorithm's precision in mirroring the brain's asymmetrical structure. This approach provides critical insights into each algorithm's capacity to represent hemispheric lateralisation in neuroimaging studies.

AFNI's scatter plot demonstrates a strong Spearman correlation ($\rho = 0.64$), with data points tightly clustered around the best fit line. This indicates consistent bilateral cluster detection and robust symmetrical performance across hemispheres. Such a pattern suggests that AFNI effectively captures similarities between the left and right hemispheres, making it highly suitable for research focused on balanced connectivity representation. The clustering of points along the line also implies that AFNI maintains reliable detection of significant clusters across subjects, ensuring minimal inter-hemispheric variability.

ANTs, exhibiting a moderate-to-strong correlation ($\rho = 0.59$), shows slightly more dispersion in the data points compared to AFNI. This pattern suggests that while ANTs maintains a systematic detection of clusters, it is more sensitive to variations in inter-hemispheric detection. Such variability may be indicative of ANTs' ability to detect subtle differences in connectivity patterns, which could be advantageous for studies examining detailed hemispheric differences or nuanced lateralisation effects. However, this increased sensitivity may

also introduce a higher likelihood of variability in detection consistency across different subjects or datasets.

DARTEL's scatter plot reveals a moderate correlation ($\rho = 0.47$) with notable spread among data points. This wider distribution implies that DARTEL may capture individual differences in hemispheric connectivity, suggesting that it prioritises subject-specific variances over strict hemispheric symmetry. While this characteristic can be advantageous for research focusing on individualised connectivity analysis, it may pose challenges for studies that require consistency and symmetrical detection across both hemispheres. The spread also indicates potential challenges in reproducibility when comparing results across subjects, as variability in detection could impact the reliability of findings in group-level studies.

FSL's scatter plot shows a moderate correlation ($\rho = 0.56$), suggesting a balanced approach to hemispheric cluster detection with moderate variability. The data points in FSL's plot indicate that while it tends to detect clusters similarly across hemispheres, it still allows for some degree of inter-hemispheric differences. This characteristic makes FSL suitable for general functional connectivity studies that benefit from a balance between detecting subtle differences and maintaining consistent results across hemispheres. The moderate correlation also implies that while FSL can effectively capture symmetrical connectivity patterns, it may not be as finely tuned as AFNI for maintaining strict bilateral consistency or as sensitive as ANTs for detecting subtle asymmetries.

Overall, these findings indicate that AFNI is optimal for studies requiring strong symmetrical performance, while ANTs' sensitivity is beneficial for detecting subtle hemispheric variations. DARTEL's individualised approach may serve studies focused on subject-specific connectivity, albeit with considerations for potential variability. FSL's balanced detection profile suits general studies where a compromise between hemispheric symmetry and variability is needed.

The bar graph (Fig. 5.5) summarises mean differences in significant cluster detection between hemispheres for each algorithm. ANTs and AFNI exhibit positive mean differences, indicating higher cluster counts in the right hemisphere, suggesting a bias towards detecting right-hemispheric features. However, the larger error bars for these algorithms reflect variability, highlighting inconsistent detection across datasets.

DARTEL, with near-zero mean differences and smaller error bars, demonstrates balanced cluster detection, making it ideal for studies prioritising symmetrical analysis. FSL shows a slight left-hemisphere preference with moderate variability, suggesting consistent performance but with a potential subtle bias.

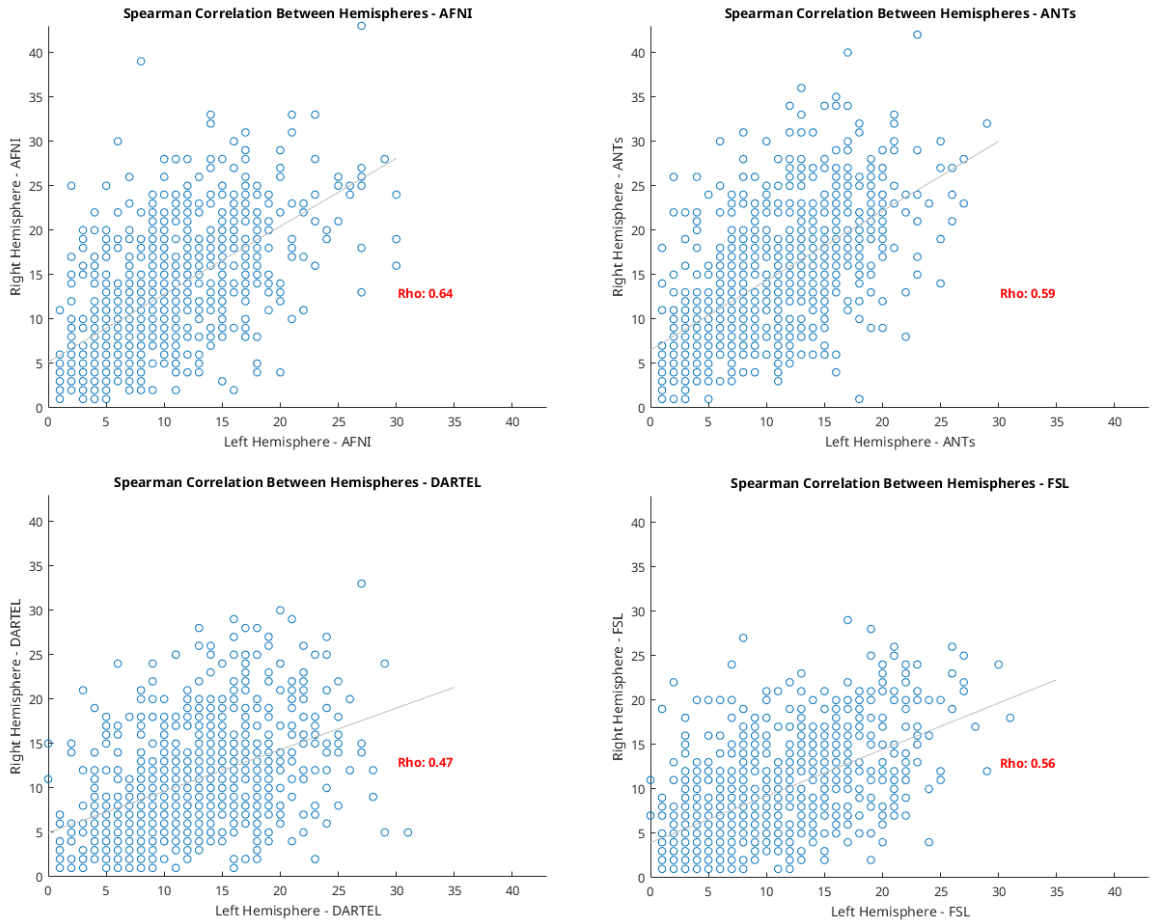

Fig. 5.4 Scatter plots of Spearman's correlation between hemispheric cluster detection for each algorithm (AFNI, ANTs, DARTEL, FSL). Correlation coefficients (ρ) reveal varying degrees of bilateral symmetry: AFNI ($\rho = 0.64$) shows strong correlation, ANTs ($\rho = 0.59$) reflects moderate-to-strong correlation with some variability, FSL ($\rho = 0.56$) indicates moderate symmetry, and DARTEL ($\rho = 0.47$) exhibits the lowest correlation, suggesting less hemispheric consistency.

Overall, ANTs and AFNI are well-suited for detecting hemispheric asymmetries, while DARTEL offers stable, balanced detection across hemispheres. FSL provides a generally consistent option but may introduce minor left-hemispheric skew.

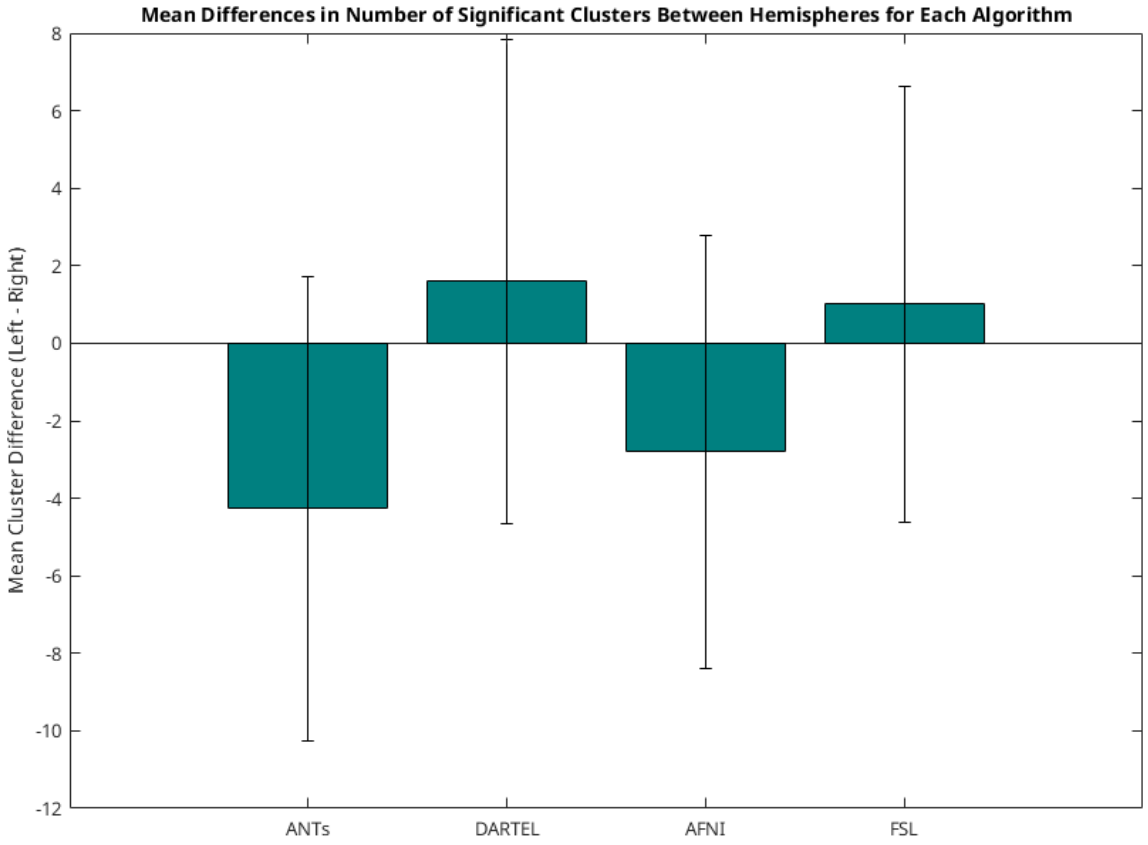

Fig. 5.5 Bar graph showing mean differences in Significant Clusters detected between left and right hemispheres by ANTs, DARTEL, AFNI, and FSL. ANTs and AFNI show higher cluster counts in the right hemisphere, while FSL shows a slight left-hemisphere preference. DARTEL demonstrates balanced detection across hemispheres.

The synthesis of these analyses highlights the importance of algorithm selection tailored to research objectives. AFNI's high correlation suggests it is well-suited for studies prioritising symmetrical representation in hemispheric analysis. ANTs, while consistent, provides sensitivity to variability, beneficial for exploring hemispheric differences. DARTEL's capacity for detecting individual variations makes it useful for personalised connectivity research, despite moderate hemispheric symmetry. FSL, with balanced performance and moderate variability, is recommended for general functional connectivity studies where minimal hemispheric bias is acceptable.

Overall, these findings emphasise that AFNI and ANTs excel in symmetric and asymmetry-sensitive studies, respectively. DARTEL supports individualised research, while FSL's balanced profile suits broader studies. This comprehensive analysis highlights the impact of algorithmic choice on neuroimaging outcomes, guiding informed selections for optimising study goals based on sensitivity and specificity in hemispheric functional connectivity.

This section uses non-parametric tests to interpret hemispheric cluster patterns in fMRI data, focusing on the influence of non-rigid registration algorithms. Spearman's Rank Correlation reveals varying degrees of consistency in hemispheric symmetry across algorithms, with AFNI demonstrating the strongest correlation, suggesting symmetrical detection. Mann-Whitney U and Wilcoxon Signed-Rank tests support these findings, highlighting significant median differences in ANTs, suggesting hemispheric asymmetry. Kruskal-Wallis Test results confirm consistent differences across algorithms, underscoring algorithmic variability in hemispheric cluster detection.

5.3 Combined Insights from Univariate Metrics

Chapters 4 and 5 provided complementary analyses of functional connectivity using two distinct univariate metrics: Peak Activation Intensity and the number of Significant Clusters. These metrics, while individually valuable, reveal different facets of brain network characteristics, offering a comprehensive understanding when evaluated together. Peak Activation Intensity, as explored in Chapter 4, focuses on the sensitivity of non-rigid registration algorithms (e.g., FSL, ANTs, DARTEL, AFNI) to localised brain activity [65, 58]. This measure captures the amplitude of activation at specific locations, which is critical for detecting strong focal activations within networks, such as those involved in cognitive and sensory processing.

In contrast, Chapter 5's analysis of Significant Clusters offers an assessment of spatial coherence and network integrity, revealing the ability of these algorithms to capture broader, functionally cohesive regions [61, 192]. By identifying spatially contiguous clusters, this metric provides insights into inter-regional connectivity and the robustness of network structures. When considered together, these univariate metrics allow for a multidimensional characterisation of functional connectivity: Peak Activation Intensities reflect the strength of focal activations, while Significant Clusters inform on the spatial extent and coherence of functional connectivity patterns.

5.3.1 Interpretation of Functional Connectivity

The combined use of Peak Activation Intensity and Significant Clusters enhances our interpretation of RSNs by providing a dual-layered perspective. Peak Activation Intensity highlights regions of high neural activity, which may correlate with core nodes within functional networks, such as the DMN or Salience Networks. In parallel, the analysis of Significant Clusters reveals how these core nodes interact with surrounding areas, mapping out network boundaries and regional cohesiveness. Together, these metrics emphasise the interplay between network nodes and their spatial coherence, facilitating a nuanced understanding of functional connectivity patterns that single metrics alone may overlook [169].

Chapter 4's analysis focused on the assessment of Peak Activation Intensities across different non-rigid registration algorithms. The findings highlighted notable algorithmic variability, with ANTs demonstrating the highest mean Peak Activation Intensity, indicating superior sensitivity, and DARTEL exhibiting consistent performance across hemispheres with minimal variability. This suggested that Peak Activation Intensity could be used to identify crucial nodes within RSNs such as the DMN, essential for cognitive processes [101, 170].

Chapter 5 extended this perspective by evaluating significant clusters, which provided insight into the broader network integrity and inter-regional coherence within RSNs. The results showed that algorithms like DARTEL and FSL maintained more consistent cluster detection, emphasising regional stability and robustness in functional connectivity analysis [60, 65]. ANTs, while highly sensitive in detecting peak activations, displayed a wider variability in cluster distribution, suggesting that it is more responsive to subtle inter-regional connections but potentially less stable in broader cluster delineation [65].

5.3.2 Implications for Neuroimaging Research

The findings from Chapters 4 and 5 highlight the importance of aligning algorithm selection with specific neuroimaging goals, particularly in the context of rs-fMRI studies. As identified, algorithms like ANTs and AFNI exhibit heightened responsiveness to localised activity peaks, which is beneficial for research focused on task-based activations or investigations into focal lesions due to their sensitivity to intensity variations [58, 65]. Conversely, algorithms such as DARTEL and FSL have demonstrated more balanced and consistent cluster detection across hemispheres, making them suitable for functional connectivity studies where network integrity and reproducibility are crucial [193, 104]. This distinction allows researchers to select algorithms that best suit their specific analytical requirements.

The dual insights from Peak Activation Intensity and significant cluster detection reveal complementary aspects of functional connectivity, supporting a multifaceted approach to algorithm selection in neuroimaging studies. Peak Activation Intensity measures highlight focal neural activity, aligning with studies that prioritise sensitivity to intense, localised activations. Meanwhile, the detection of Significant Clusters informs on the broader spatial coherence of functional connectivity patterns, crucial for RSN analyses that depend on capturing inter-regional cohesiveness [23, 24].

By characterising the unique strengths of each algorithm, these findings provide a framework for informed algorithm selection based on study objectives. For studies emphasising focal peaks, ANTs and AFNI may offer superior detection due to their intensity sensitivity, although with higher variability in inter-subject comparisons [20]. In contrast, DARTEL and FSL, with their balanced detection across hemispheres, may serve as robust choices for research focused on hemispheric consistency and reproducibility. Thus, the choice of algorithm can meaningfully impact the interpretation of functional connectivity, emphasising the need for a tailored approach to designing neuroimaging studies [20, 153].

5.4 Chapter Summary

Chapter 5 conducted an in-depth examination of Significant Clusters as indicators of network integrity, focusing on rs-fMRI analyses. By extending the framework from Chapter 4, which concentrated on Peak Activation Intensities, this chapter provided a broader perspective on functional connectivity by evaluating the spatial coherence of Significant Clusters across hemispheres and within the Control Network. The primary goal here was to assess how different neuroimaging algorithms—DARTEL, FSL, ANTs, and AFNI—affect the reproducibility and reliability of network structures, each presenting unique strengths and limitations. For instance, DARTEL's consistency in significant cluster detection highlights its utility in maintaining network integrity across subjects, a key advantage for group-level analyses. In contrast, ANTs and AFNI, while highly sensitive, exhibited greater variability, posing potential challenges for studies emphasising hemispheric symmetry or cross-subject comparability.

This chapter's methodological approach, integrating non-parametric tests and distribution analyses, reinforced the distinctive profiles of these algorithms in detecting and interpreting functional clusters. These findings highlight the importance of selecting algorithms aligned with study objectives: DARTEL and FSL for balanced, symmetric detection across hemispheres, and ANTs or AFNI for studies requiring heightened sensitivity to nuanced,

potentially asymmetric activations. These insights into algorithm-specific cluster characteristics are crucial for designing robust neuroimaging studies, as they influence both the spatial accuracy and interpretability of functional connectivity maps.

Looking forward, Chapter 6 will build on these findings by delving into inter-subject variability and group-level network inferences. This next chapter will expand the univariate analyses presented thus far into multivoxel contexts, enabling an exploration of between-individuals differences and the broader implications of algorithm choice on group-level functional connectivity patterns. Together, the analyses from Chapters 4 and 5 provide a robust foundation for understanding the nuances of neuroimaging algorithms, preparing for a comprehensive evaluation of functional connectivity within and across subjects.

Chapter 6

Inter-Subject Variability and Group Inference Analysis

In Chapter 6 of this thesis, we critically examine Mutual Information (MI), Dice Similarity Coefficient (DSC) metrics, and Multivoxel Pattern Analysis (MVPA) as key analytical metrics to assess inter-subject variability and algorithmic performance in brain functional connectivity analysis. Building on the foundational concepts introduced in Chapters 3 to 5, this chapter integrates these metrics within a systematic framework to provide a deeper understanding of Resting-State Networks (RSNs), which retain intrinsic activity across different states of consciousness and are critical in neuroimaging studies [194, 32, 195].

To clarify the structure of Chapter 6, a conceptual flow diagram follows this introduction (Figure 6.1). This diagram outlines the inter-subject analysis pipeline, highlighting the sequence of evaluation metrics, similarity analysis, and algorithmic performance assessment.

The pipeline begins with the evaluation of MI and DSC metrics, which quantify dependency relationships and spatial overlap across subjects, respectively. MI measures the degree of dependency among brain regions, while DSC metrics evaluate spatial similarity achieved by each registration algorithm, FMRIB Software Library (FSL), Advanced Normalisation Tools (ANTs), Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL), and Analysis of Functional NeuroImages (AFNI). Next, MVPA, using Support Vector Machine (SVM) weight maps, identifies predictive neural patterns, providing insights into variability across subjects. These metrics collectively enable a multi-dimensional assessment of functional connectivity.

Following this, algorithmic performance analysis assesses the clustering consistency and spatial alignment achieved by each algorithm, capturing the nuances of inter-subject

variability. The chapter concludes with group-level inference, summarising how the findings contribute to understanding functional connectivity patterns across the study population.

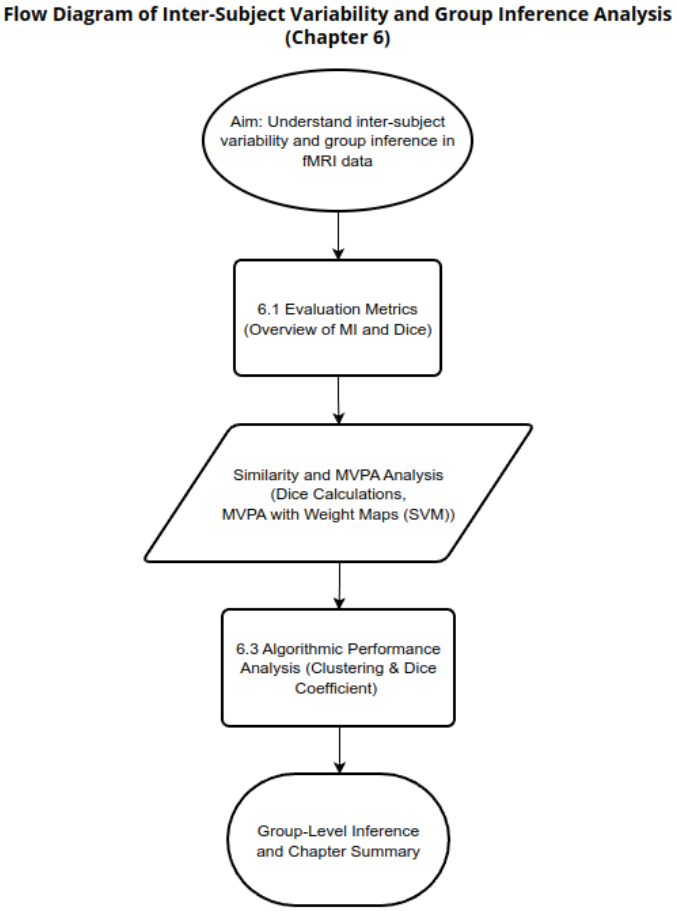

Fig. 6.1 Flow diagram of Chapter 6's analysis pipeline, showing the stages of evaluation metrics, similarity analysis, and algorithmic performance assessment.

MVPA, in particular, plays an instrumental role in identifying neural patterns associated with cognitive states and further differentiates between registration algorithms by uncovering variability in the spatial representations they produce. Its inclusion here aligns with the overarching goals of evaluating RSNs within Resting-State Functional Magnetic Resonance Imaging (rs-fMRI) data. This analysis advances our understanding of neural connectivity, allowing for a more nuanced interpretation of potential biomarkers in neurological disorders [196].

This chapter systematically explores the metrics mentioned above, beginning with a focused investigation of MI, followed by spatial overlap measurements via DSC, and finally, a detailed analysis using MVPA. The findings from each metric build upon each

other, contributing to a cohesive narrative of algorithmic performance in the context of neuroimaging and RSN variability. This structure facilitates a thorough evaluation of how these methods, individually and collectively, contribute to enhancing accuracy and consistency in neuroimaging studies.

6.1 Evaluation Metrics

To quantitatively assess the alignment and inter-subject variability within rs-fMRI data, MI and DSC metrics are employed as complementary evaluation methods. MI provides a measure of dependency relationships between brain regions, while the DSC coefficient captures spatial overlap, offering a cohesive framework to examine both functional connectivity and alignment accuracy across subjects. This section sequentially presents MI and DSC metrics, followed by their respective results, streamlining the analysis flow for improved comprehension.

6.1.1 Overview of MI and DSC Metrics

Mutual Information (MI)

MI is a non-parametric measure that quantifies the dependency between two variables. In the context of neuroimaging, MI serves as an essential metric for assessing alignment and similarity between brain images, particularly useful in evaluating the consistency of non-rigid registration algorithms. MI is widely applied in image registration due to its robustness in capturing complex dependencies without relying on specific distributional assumptions [2, 9]. This study uses MI to evaluate the alignment accuracy and consistency of functional connectivity patterns across subjects after applying various non-rigid registration algorithms to resting-state fMRI (rs-fMRI) data [197, 198].

The MI calculation process is structured as follows:

1. **Image Preparation:**
    - **Data Selection:** rs-fMRI images from each subject are processed using different non-rigid registration algorithms, resulting in four sets of images per algorithm each.
    - **Preprocessing:** The images are binarised and masked to focus on regions of interest within the brain, specifically the RSNs, to ensure MI calculations are focused on relevant areas, enhancing result specificity.

2. **Entropy and Joint Entropy Calculation:**

   - **Marginal Entropy of Image** X: The entropy $H(X)$ for an image X quantifies the uncertainty or unpredictability in its voxel intensity distribution:

$$H(X) = -\sum_{x \in X} p(x) \log p(x) \tag{6.1}$$

     Here, $p(x)$ represents the probability of each voxel intensity level x in the image. Entropy, in this case, measures the amount of information contained within the image. High entropy values suggest more variability in voxel intensities, indicating a more complex or diverse spatial pattern in the brain image. This measure is foundational as it provides a baseline for the amount of information present in each image separately.

   - **Joint Entropy of Images** X **and** Y: The joint entropy $H(X,Y)$ measures the combined uncertainty or unpredictability across two images:

$$H(X,Y) = -\sum_{x \in X, y \in Y} p(x,y) \log p(x,y) \tag{6.2}$$

     where $p(x,y)$ represents the joint probability of observing voxel intensities x and y simultaneously in images X and Y, respectively. Joint entropy quantifies the overlap in voxel intensity patterns between the two images. Lower joint entropy values indicate higher similarity, as it implies less combined unpredictability across the two images.

3. **MI Calculation:** MI between two images X and Y is then computed as:

$$MI(X,Y) = H(X) + H(Y) - H(X,Y) \tag{6.3}$$

   This formula represents the shared information between images X and Y. In the context of image registration, a higher MI value signifies a greater alignment quality, as it indicates more shared information between the two images. This reflects a better correspondence between voxel intensities in X and Y, implying that the registration algorithm has effectively aligned similar anatomical or functional features. This study uses MATLAB R2023a [130] for MI calculations, automating the process and handling the large volume of rs-fMRI data.

In neuroimaging, MI is advantageous due to its sensitivity to complex voxel intensity relationships without requiring assumptions about intensity distributions. This makes it highly suitable for capturing subtle dependencies in functional connectivity patterns between subjects [4, 60]. By identifying registration algorithms that maximise MI values, this study selects approaches that preserve functional connectivity patterns essential for RSN analysis [198, 197]. Overall, MI enhances the specificity and reliability of neuroimaging data alignment, providing a metric that directly reflects functional correspondence between subjects.

Spatial Overlap - DSC Metric

The DSC is a statistical metric that quantifies spatial overlap between binary images, commonly used in neuroimaging to assess the spatial alignment of functional regions across subjects. In the context of rs-fMRI, DSC provides a measure of how well different registration algorithms align functional networks across individuals. This alignment is essential for accurately comparing brain networks in studies involving disease contexts, such as Alzheimer's Disease [199].

For visualising DSC results, heat maps are employed, which use a gradient from blue (low similarity) to red (high similarity) to intuitively convey the spatial similarity achieved by each algorithm. This visualisation aids in detecting algorithmic performance variations, which can be particularly informative when analysing structural and functional differences in clinical populations.

The DSC calculation follows these steps:

1. **Region Binarisation:** Each image is binarised, marking regions of interest within functional networks as '1' (indicating presence) and all other areas as '0'. This enables voxel-wise comparison between registered images from different subjects or between a subject's image and a template [200].

2. **DSC Calculation:**

    - **DSC Formula**: The DSC score quantifies spatial overlap between two binary images A and B, where each image is composed of voxels that either belong to a functional region (indicated by '1') or do not ('0'). The formula for DSC is:

$$DSC(A,B) = 2 \times \frac{|A \cap B|}{|A| + |B|} \tag{6.4}$$

In this formula, $|A \cap B|$ represents the count of voxels that are marked '1' in both images A and B, indicating the region of spatial overlap. The terms $|A|$ and $|B|$ represent the count of '1' voxels in each image separately. By calculating DSC, we assess the extent to which two regions of interest, identified by the binary masks, overlap spatially [201].

- **Interpretation of the DSC Value**: The DSC metric yields a value between 0 and 1, where higher values indicate greater spatial alignment of the functional regions across subjects or with respect to a template. A DSC score of 1 would indicate perfect alignment, while a score of 0 would imply no overlap. In practice, higher DSC values reflect the effectiveness of a registration algorithm in aligning functional brain networks. Conversely, lower scores may indicate alignment challenges, potentially due to limitations of the registration algorithm or individual anatomical differences between subjects.

Heat maps and cluster visuals are used to present DSC results, with each colour representing a specific level of similarity, ranging from low (blue) to high (red). These visual tools allow for quick identification of algorithmic performance variations across different brain regions and subjects, offering insights into the spatial similarity achieved by each registration method. This visual approach is especially valuable in analysing functional network alignment in rs-fMRI, aiding researchers in selecting optimal algorithms for analysing inter-subject variability in both healthy and clinical populations.

In summary, DSC serves as a robust metric for evaluating spatial overlap, offering insights into the precision of alignment achieved by different registration algorithms in the functional brain regions across subjects. The combination of MI and DSC metrics enables a comprehensive assessment of functional alignment quality, revealing both shared and distinct aspects of functional connectivity network similarity across subjects in rs-fMRI data.

6.1.2 MI Calculations and Results

A pairwise MI analysis was conducted to quantify the similarity in information content between the MVPA weight maps generated by four non-rigid registration algorithms: FSL, DARTEL, ANTs, and AFNI. This analysis assesses the extent to which each algorithm preserves shared information in the spatial distribution of activation patterns within rs-fMRI data. The MI values, computed on a scale from 0 (no shared information) to 1 (complete overlap of information content), offer a quantitative benchmark for comparing the outputs of these algorithms.

The results of the MI calculations are presented in a matrix (Table B.1) and visually summarised in a heat map (Figure 6.2). These representations provide insights into the pairwise similarity of activation patterns detected by different algorithms, shedding light on both the overlaps and distinctions in their registration outcomes.

The diagonal entries in the heat map consistently show MI values of 1, as expected, reflecting the self-similarity within each algorithm's output and validating the MI calculation process. The colour scale was set to a range of 0.7 to 1 to enhance the visual contrast within the MI values observed (0.8062 to 0.9461). This adjustment highlights subtle differences in overlap between algorithm pairs, making it easier to notice regions of higher and lower MI. This tailored range ensures that the visual representation aligns with the data range, while avoiding excessive exaggeration of small differences.

The heat map in Figure 6.2 utilises a gradient colour scheme, where lighter shades indicate higher MI values and, consequently, greater similarity in detected activation patterns. Notably, the 'ANTS vs DARTEL' and 'FSL vs DARTEL' comparison yielded an MI value of 0.9461, the highest in the matrix, signifying a substantial overlap in information content between these two algorithms. This finding suggests that FSL and AFNI, despite their methodological differences, may offer comparable sensitivity to functional signals within the analysed dataset.

In contrast, the 'FSL vs ANTS' and 'FSL vs DARTEL' comparison, marked by a darker shade, represents the lowest MI value in the matrix (0.8062). This lower degree of similarity reflects the distinct processing characteristics of these algorithms, which may differentially influence the sensitivity to specific brain features or the handling of noise within the data. Such variations highlight the need to carefully consider algorithm choice based on study objectives and data characteristics.

Further, the MI results suggest possible redundancies or complementarities among the algorithms. For example, 'FSL vs ANTS' and 'ANTS vs AFNI' pairings show proximate MI values, implying that these algorithms might capture analogous features or exhibit similar processing biases. This information can aid researchers in selecting algorithms that best align with their goals, whether aiming for consistent activation detection or leveraging algorithmic diversity to enhance analytical rigour.

The MI heat map analysis is complemented by Wilcoxon averaged bar graphs (Figures 4.5 and 5.5), which depict the range of MI values across algorithm pairs. Higher bars in the Wilcoxon graphs correlate with lighter shades in the MI heat map, signifying greater agreement between algorithms, while lower bars align with darker shades, indicating greater

divergence. This cross-referencing provides a comprehensive understanding of algorithm performance, supporting informed choices in algorithm selection for neuroimaging studies.

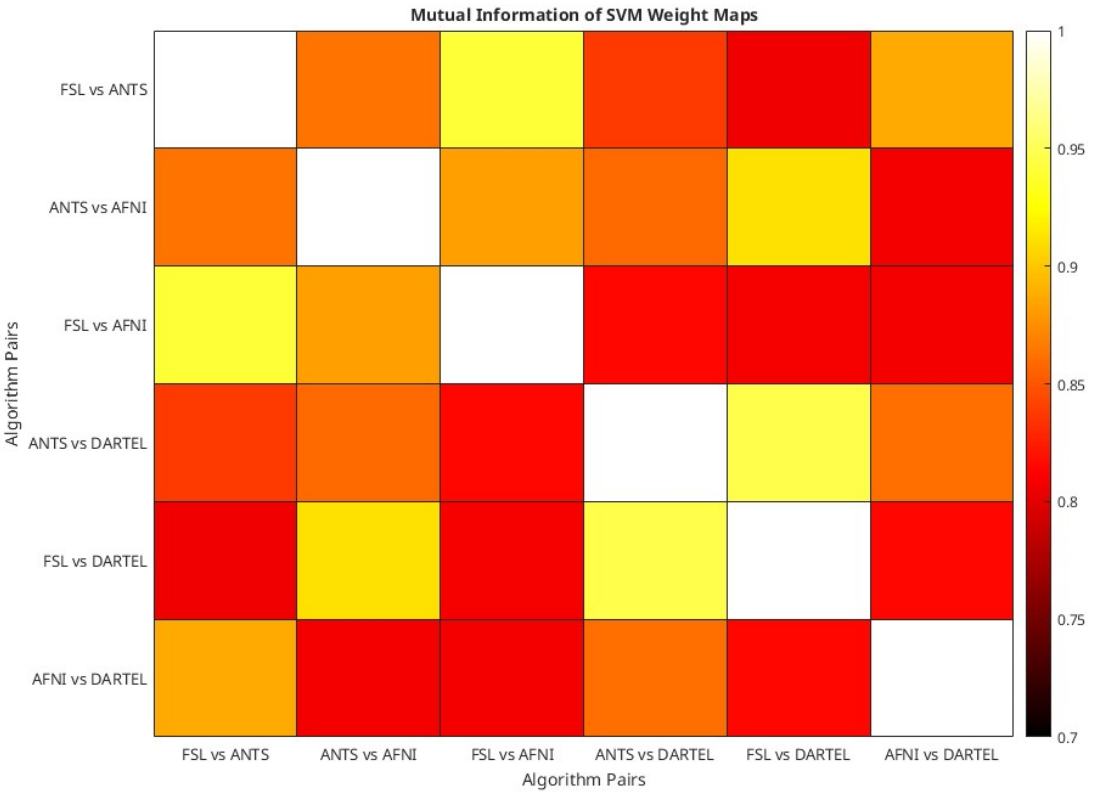

Fig. 6.2 This heat map illustrates the MI between whole-brain SVM weight distributions across pairs of non-rigid registration algorithms, reflecting the similarity of activation patterns. Higher MI values, shown in lighter shades, indicate greater overlap in detected patterns, with the highest MI (0.9461) observed between 'ANTS vs DARTEL' and 'FSL vs DARTEL'. Darker shades represent lower MI, with the minimum (0.8062) between 'FSL vs ANTS' and 'FSL vs DARTEL', suggesting less agreement. All values can be found in Table B.1, Appendix B.

6.1.3 DSC Calculations and Results

This section outlines clustering techniques relevant to this study, emphasising distinctions between statistical clustering, Machine Learning (ML) clustering, and reordered heat maps in neuroimaging. Each method contributes unique insights into functional connectivity and inter-subject variability in rs-fMRI data.

Clustering Techniques Overview: Traditional statistical clustering, as used in the FSL FMRI Expert Analysis Tool (FEAT) framework, supports validation of known functional networks by limiting false positives through Gaussian Random Field (GRF) theory [132]. ML clustering methods (e.g., K-means, Density-Based Spatial Clustering of Applications with Noise (DBSCAN)) differ by exploring latent patterns within data, allowing the discovery of novel connectivity profiles in complex conditions like Alzheimer's and depression [202, 203].

Reordered Heat Maps for Spatial Congruency: Reordered heat maps in this thesis organise DSC matrices to visually enhance similarity across subjects. This non-clustering visualisation highlights continuous similarity across subjects without enforcing discrete groupings, aiding assessment of each algorithm's alignment consistency [204, 202].

Hierarchical Clustering in Heat Map Analysis: Hierarchical clustering supplements heat map analysis by revealing consistent connectivity patterns, although its assumption of discrete groupings may introduce artificial boundaries. As such, clustering here is a tool for identifying group-level tendencies, which benefits from further validation [205, 41].

Reordered Heat Maps

The heat maps displayed in this section are reordered representations of DSC matrices, which provide a visual assessment of the similarity in functional region alignment post-registration by four algorithms—AFNI, ANTs, DARTEL, and FSL. These matrices are derived from binarised, registered Control Networks across the study population and serve as a quantitative measure of the spatial overlap in segmented brain regions between subjects. The reordering process here is distinctly different from clustering-based analysis, such as significant cluster analysis presented in Chapter 5 or ML-based clustering techniques, as it is not designed to discover inherent group structures based on feature similarity but rather to visually enhance the detection of regions with high or low DSC similarity.

The reordering in these heat maps essentially organises subjects based on the resemblance of their registered brain images, optimising the visibility of population-level patterns in the data. By aligning similar patterns adjacently, this approach facilitates a strategic vantage point from which one can discern disparities in algorithmic performance, potentially attributable to individual anatomical features or other idiosyncratic differences within the brain's Control Network. Thus, it is particularly useful in assessing the robustness and consistency of each

registration technique in capturing functional topography across diverse subjects [202, 206, 204]. These reordered heat maps are shown in the Figure 6.3.

The AFNI reordered heat map, presents a mixed patchwork of similarity levels, illustrating both close alignment and significant divergence in functional anatomy across subjects. This pattern suggests that AFNI might be particularly sensitive to subject-specific anatomical variability and potentially less consistent in aligning regions that are more variable among individuals. Intense red and orange areas indicate regions where AFNI achieves high alignment across subjects, likely corresponding to brain regions with lower anatomical variability. The scattered cooler hues (yellow and blue) denote regions with lower similarity, which could be a result of AFNI's handling of noise or its sensitivity to more variable anatomical features.

In contrast, the ANTs heat map reveals a relatively homogenous distribution of high-similarity values, indicating robust across-subject consistency in the alignment of the Control Network. The prevalence of warmer colours (reds) signifies that ANTs reliably aligns functional areas across the subject pool, reflecting its advanced handling of anatomical diversity. This uniformity suggests that ANTs is particularly suitable for studies prioritising group-level consistency, especially when accurate population-level inferences are critical. The scarcity of cooler areas emphasises ANTs' proficiency in adapting the Control Network across diverse brain structures, showcasing its capability for stable and precise alignment.

DARTEL's reordered heat map shows moderate red clusters interspersed with a yellow-orange backdrop, reflecting a balance between capturing individual subject nuances and maintaining a consistent group-level anatomy. The moderate similarity regions suggest that DARTEL effectively captures highly conserved regions of the Control Network, while the dispersed red areas indicate its alignment of more individualised features. This approach, termed here as 'algorithmic balance,' could make DARTEL a suitable choice when both individual detail and population consistency are valuable in the analysis.

Lastly, the FSL heat map presents a dynamic range of similarity scores, implying an algorithm capable of differentiating with notable precision. The gradient from warm to cool colours signifies FSL's ability to align both common and unique anatomical features within the Control Network. This distribution implies that FSL excels in preserving individual subject characteristics while achieving a coherent group-level alignment, suggesting its utility for comprehensive neuroimaging studies requiring detailed, individualised registration alongside population-level conformity.

In summary, comparing these reordered heat maps reveals ANTs and FSL as potentially optimal choices for studies requiring high group-level similarity, whereas AFNI's variability

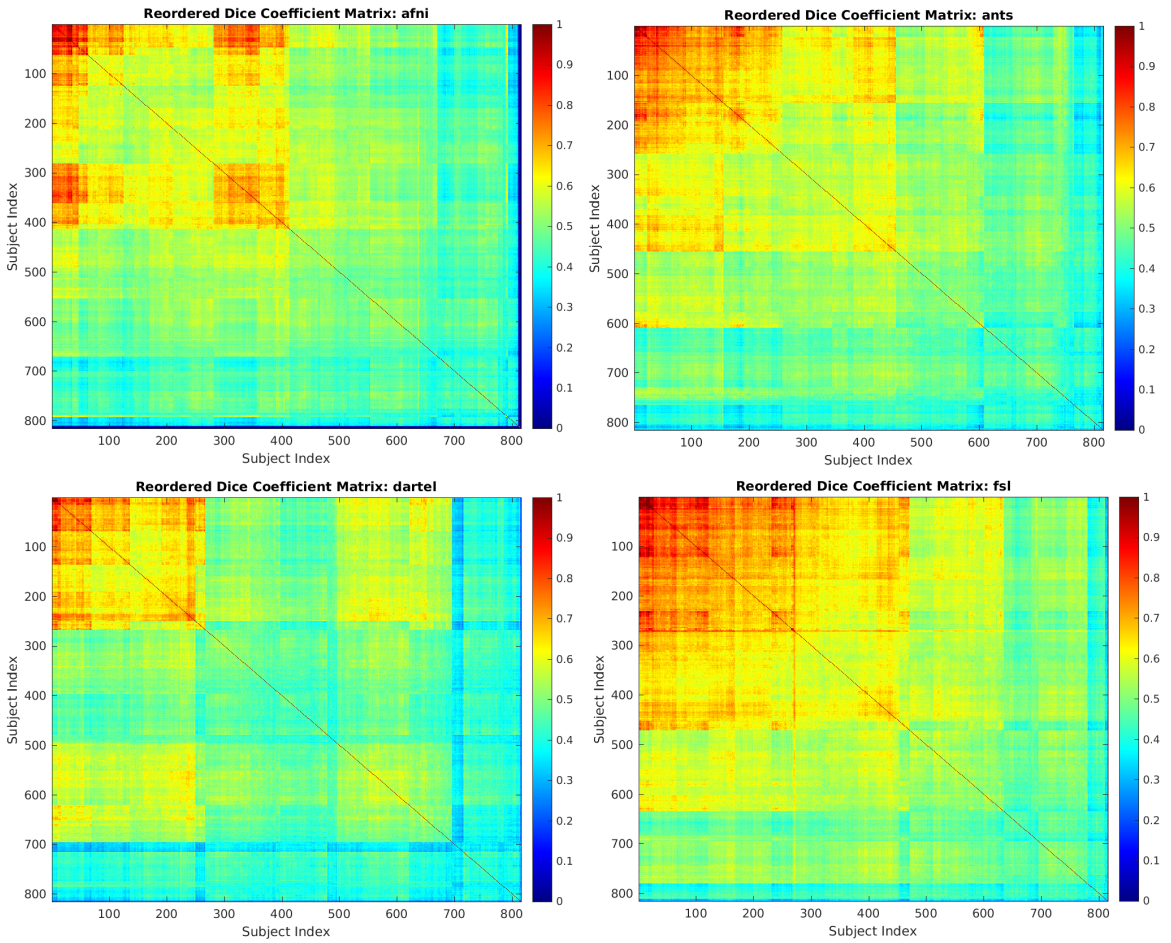

Fig. 6.3 Reordered DSC heat maps for each algorithm (AFNI, ANTs, DARTEL, FSL), illustrating inter-subject similarity in brain structure registration. Subjects are reordered along both axes to highlight clusters of similarity. Warmer colours (reds and oranges) indicate higher similarity, while cooler colours (blues) indicate lower similarity. AFNI shows variable consistency with distinct patches of similarity, ANTs demonstrates high overall consistency, DARTEL balances individual detail with group alignment, and FSL reveals a range of similarity, reflecting both shared and individual anatomical features.

may offer unique insights into individualised anatomical features. DARTEL's intermediate positioning makes it suitable for analyses emphasising a shared anatomical basis with allowances for individual variation. The comparative analysis of reordered DSC matrices supports tailored algorithm selection based on the specific requirements of neuroimaging studies, balancing individual variability and group consistency.

Clustered Heat Maps

Clustering within heat maps serves a critical function in neuroimaging, particularly in the examination of non-rigid registration algorithms. It enables the discovery of inherent groupings or patterns in data that may not be evident from individual comparisons. By employing hierarchical clustering, one can identify subjects whose functional connectivity profiles are similar, potentially unveiling sub-populations or consistent patterns across different individuals. This facilitates the outlining of algorithmic performance nuances, revealing whether certain algorithms are more adept at detecting and aligning specific RSNs. Hierarchical clustering also aids in the comprehensive understanding of the overarching structure and the interconnectivity of brain networks, thereby refining group inferences from complex neuroimaging data [203].

While hierarchical clustering of heat maps provides valuable insights into the functional connectivity profiles and potential sub-populations within neuroimaging data [205], this method comes with inherent challenges and limitations [207]. Clustering algorithms, including hierarchical clustering, inherently assume that the data contains distinct groups, which may not be a true representation of the complex, continuous nature of functional brain networks. This assumption can sometimes lead to artificial boundaries between subjects that do not reflect genuine neurobiological differences.

Moreover, clustering can obscure clinically significant individual differences, and it is crucial not to over-interpret these clusters as distinct phenotypes without supporting evidence from other modalities. In this thesis, clustering is used as a supplementary tool to identify potential grouping tendencies in functional connectivity, rather than to define rigid brain network categories [41]. To address these limitations, the clustering results are integrated within a broader analytical framework that includes independent metrics, such as MI and DSC, and cross-referenced with established neuroimaging literature. This approach situates clustering as one part of a balanced analysis, capturing both individual variability and group-level patterns while avoiding over-generalisation of sub-populations. These maps (Fig. 6.4) visualise the comparative efficiency and pattern recognition capabilities of each algorithm in identifying RSNs across subjects.

The AFNI clustered heat map, with its distinct diagonal clusters, suggests that the algorithm efficiently captures strong within-network connectivity among certain subjects. These subject groups, reflected as clusters, likely exhibit high within-network connectivity that AFNI can effectively register, leveraging its temporal correlation analysis strengths in fMRI. The sparser regions or the reduced similarity between these clusters could imply instances where AFNI's registration is less successful, potentially due to its heightened sensitivity to functional heterogeneity and noise. This could be indicative of AFNI's variable performance in aligning and detecting subtle differences within more diverse or complex functional areas of the brain.

ANTs' clustered heat map, with its dense and extensive clusters, speaks to its robustness in capturing a broad spectrum of functional connectivity, both within and between various brain networks. The pronounced clusters that extend off-diagonal are indicative of the algorithm's capacity to register functional relationships that span different networks, affirming ANTs' utility in studies requiring thorough brain network mapping. The algorithm's symmetric normalisation technique likely contributes to this proficiency, as it actively compensates for inter-subject anatomical diversity, thereby promoting a high degree of consistency in functional connectivity patterns across subjects.

DARTEL's heat map exhibits sharply defined clusters, suggesting a high degree of precision in aligning localised brain activities. The well-outlined clusters are emblematic of DARTEL's efficient high-dimensional warping technique, which is adept at creating detailed group-specific templates. This capability for precise registration makes DARTEL particularly suitable for Voxel-Based Morphometry (VBM) studies, which require high accuracy in matching morphological brain structures across individuals. The clear cluster delineations highlight DARTEL's strength in providing consistent group-level alignment.

FSL's clustered heat map illustrates a balanced array of clusters that demonstrate both focused and widespread functional connectivity. This balanced distribution reflects FSL's hybrid registration approach, combining linear and non-linear methods to cater to diverse anatomical and functional variability across subjects. The presence of both densely and sparsely populated clusters may signify FSL's flexibility and robustness in handling a wide range of neuroanatomical profiles, supporting its use for comprehensive neuroimaging analyses that require adaptability to different types of brain structures and activity patterns.

Upon comparative analysis, the clustered heat maps reveal distinct characteristics of the tested algorithms in capturing the complex nature of RSN connectivity. AFNI's heat map, defined by strong diagonal clusters, suggests a propensity for detecting coherent within-network connectivity, potentially excelling in tasks demanding high precision for

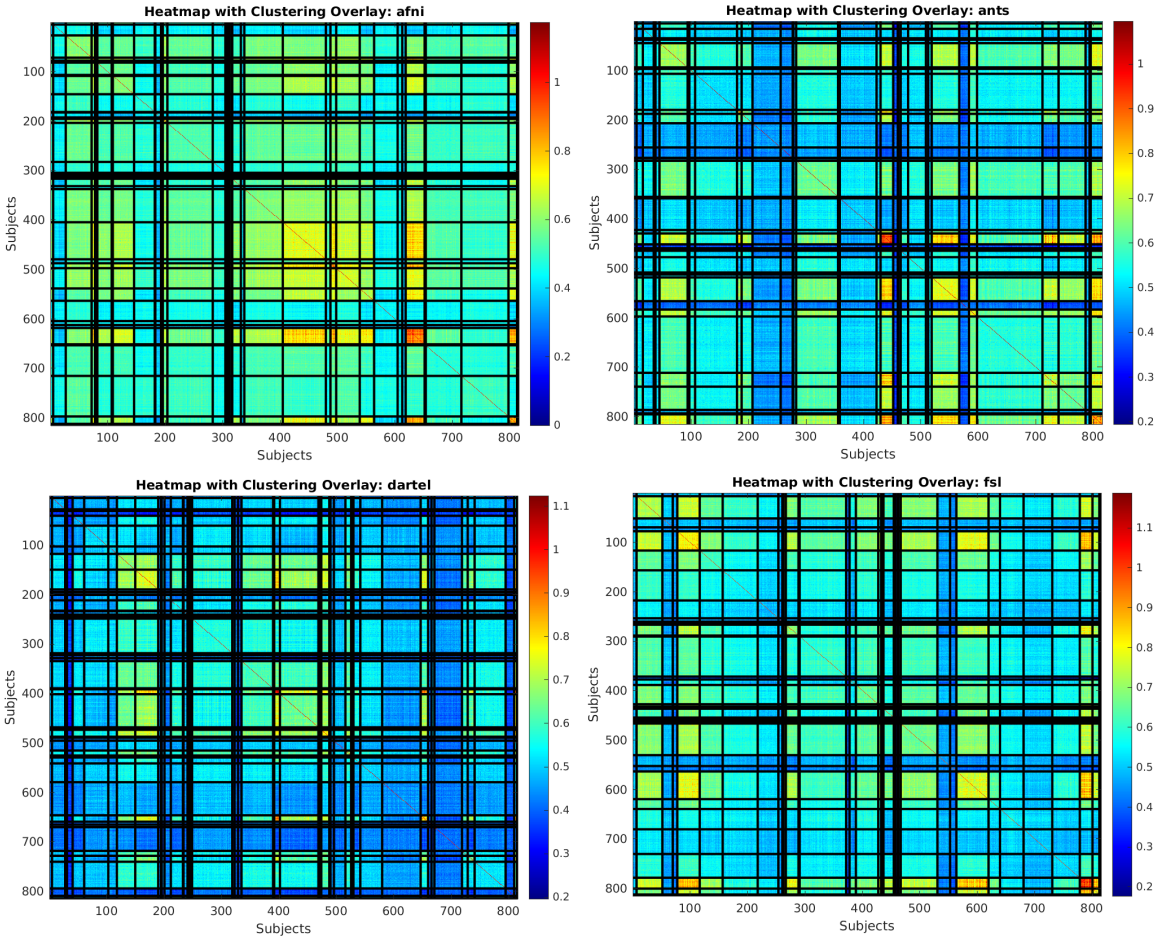

Fig. 6.4 Clustered heat maps of functional connectivity correlation for each algorithm (AFNI, ANTs, DARTEL, FSL). Subjects (N = 815) are ordered by hierarchical clustering on both axes. Black grid lines delineate clusters, with colour intensity indicating the level of functional connectivity correlation (warmer colours represent higher correlation). AFNI displays variability in capturing individual differences, ANTs shows extensive connectivity with dense clusters, DARTEL highlights localised connectivity, and FSL demonstrates a balanced range of connectivity, reflecting robustness across the subject dataset.

well-defined networks. ANTs, with its dense and extensive clustering, displays robust across-network functional mapping, indicative of its ability to accommodate the brain's anatomical and functional diversity, hence suitable for comprehensive functional connectivity analyses. DARTEL, marked by its sharply defined clusters, reflects a high level of detail in local brain activities, aligning with its specialised utility in VBM and studies necessitating high spatial specificity. In contrast, FSL's balanced cluster distribution, and distributed functional connections, demonstrates its adaptable nature, capable of discerning both pronounced and subtle network patterns, making it a versatile tool for a broad spectrum of neuroimaging research. This side-by-side assessment not only highlights the strengths and limitations of each algorithm, ensuring algorithmic choices are congruent with the specific objectives and requirements of the research at hand [41].

6.2 Multivoxel Pattern Analysis (MVPA)

MVPA offers a nuanced approach for investigating spatial variations across registration algorithms. By capturing differences in spatial patterns, MVPA enhances our understanding of algorithmic biases in processing brain data, particularly in the context of functional neuroimaging.

To clarify algorithmic differences, we conducted pairwise comparisons using SVM weights across six algorithmic pairs: FSL, ANTs, AFNI, and DARTEL. The SVM-generated weight maps represent each algorithm's impact on voxel-wise classifications, revealing where each algorithm emphasises distinct neural patterns, as visualised in Figures 6.5 to 6.10. These maps are accompanied by MI and DSC results, providing a broader context for understanding the algorithms' spatial preferences.

6.2.1 Weight Maps Interpretation

The SVM weight maps illustrate the voxel-level distinctions between algorithms, highlighting regions of high algorithmic influence. Each map utilises a 'hot' colour scheme where red-to-yellow gradients represent the varying weights: dark red for weights favouring the second algorithm in each pair, and yellow for weights in favour of the first algorithm.

In the AFNI vs DARTEL comparison (Fig. 6.5), AFNI assigns significant weight to certain brain regions, as observed by the maximum weight of 0.71 at slice 38. This high weight suggests that AFNI places a stronger emphasis on specific regions, potentially enhancing its sensitivity to particular spatial patterns in the data. Conversely, DARTEL emphasises

different areas, with a minimum weight of -0.67 at slice 16, indicating distinct processing priorities. This divergence illustrates how AFNI and DARTEL may produce unique representations of functional connectivity, potentially leading to differing interpretations of neural network activity depending on the chosen algorithm.

ANTs vs AFNI (Fig. 6.6) weight map demonstrates more nuanced spatial variation, with weights ranging from 0.69 to -0.68. This narrower range suggests that while both algorithms share some spatial sensitivities, they still diverge in certain neural activation areas. ANTs appears to capture subtle activation differences that AFNI may overlook, which could be advantageous in studies requiring fine-grained sensitivity to brain network interactions. The observed differences highlight ANTs' potential methodological advantage in capturing distinct neural signals, especially within complex networks.

Similarly, ANTs vs DARTEL (Fig. 6.7) comparison reveals notable spatial discrepancies, with weights ranging between 0.67 and -0.65. This broad distribution indicates that each algorithm uses distinct spatial emphasis strategies to classify brain activity. These differences suggest that DARTEL may excel in some areas of functional connectivity not prioritised by ANTs, and vice versa. For researchers, the choice between ANTs and DARTEL may be guided by the specific spatial characteristics of the brain networks under investigation, as each algorithm may provide unique insights into the functional connectivity structure.

The comparison between FSL vs AFNI (Fig. 6.8) displays a wide range of weights from 0.71 to -0.72, signifying a strong divergence in their spatial emphasis. This extensive range indicates that FSL and AFNI are markedly different in their representations of brain activations. FSL's higher emphasis on certain areas compared to AFNI suggests that each algorithm could lead to substantially different interpretations of the data, potentially impacting findings related to the localisation of functional connectivity patterns.

The FSL vs ANTs (Fig. 6.9) weight map presents moderate discrepancies, with weights spanning from 0.59 to -0.66. While both algorithms highlight similar neural features, FSL and ANTs differ in their sensitivity and emphasis on specific spatial patterns. This moderate range indicates that either algorithm could be suitable for studies where sensitivity to functional connectivity is needed, but subtle algorithmic biases should still be accounted for to avoid misinterpretation of specific neural regions.

In the FSL vs DARTEL comparison (Fig. 6.10), the weight distribution ranges from 0.70 to -0.66. This comparison emphasises the distinct spatial biases each algorithm has, suggesting that DARTEL and FSL may capture varying connectivity patterns in resting-state data. Understanding these spatial divergences is crucial for interpreting functional

connectivity accurately, as each algorithm may highlight or de-emphasise different regions, potentially affecting the robustness of neuroscientific findings.

These comparisons highlight the importance of algorithmic considerations in neuroimaging, as each approach captures unique aspects of brain function. Understanding these biases can guide algorithm selection, ensuring alignment with research goals and improving functional connectivity interpretations.

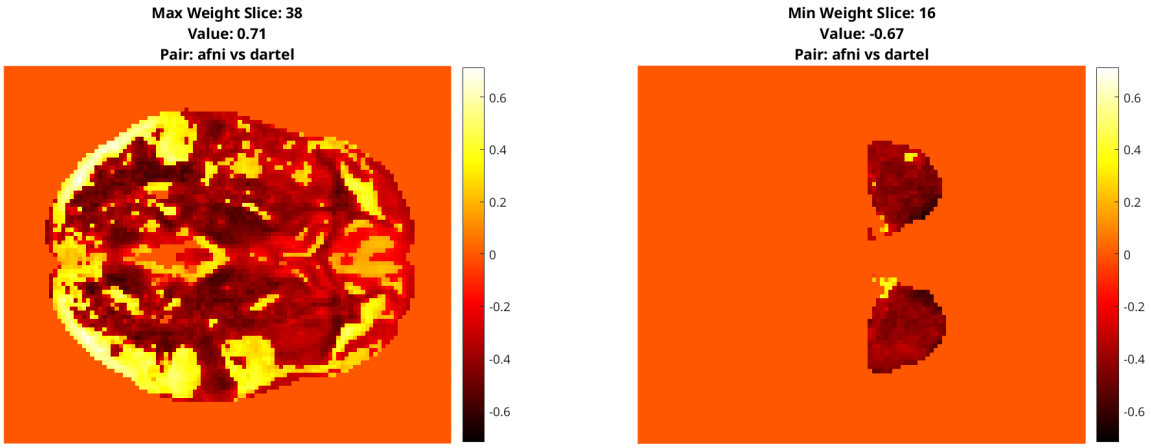

Fig. 6.5 Heat map comparison of AFNI vs DARTEL algorithms showing maximal weight: 0.7094 at slice 38 (X: 28, Y: 19, Z: 38) and minimal weight: -0.6738 at slice 16 (X: 37, Y: 70, Z: 16) The colour scale indicates regions of differential weight, with warmer colours representing higher weights. These patterns highlight the contrasting emphases of AFNI and DARTEL on specific brain regions, which could have significant implications for the interpretation of functional connectivity.

Implications and Synthesis of Results:

These SVM weight map comparisons emphasise the critical role of algorithmic choice in neuroimaging analyses. The differential spatial weighting across algorithms reveals how each approach emphasises unique aspects of brain function, which may have substantial implications for studies examining inter-subject variability [82, 177]. For example, choosing an algorithm that aligns with the specific neural features of interest can enhance the accuracy of functional connectivity interpretations [57], while a misaligned choice could introduce biases, potentially obscuring true neural patterns [48].

By carefully selecting algorithms that best match the research objectives and understanding the biases each algorithm introduces, researchers can improve the reliability and interpretability of rs-fMRI studies. This enhanced understanding of spatial biases and sen-

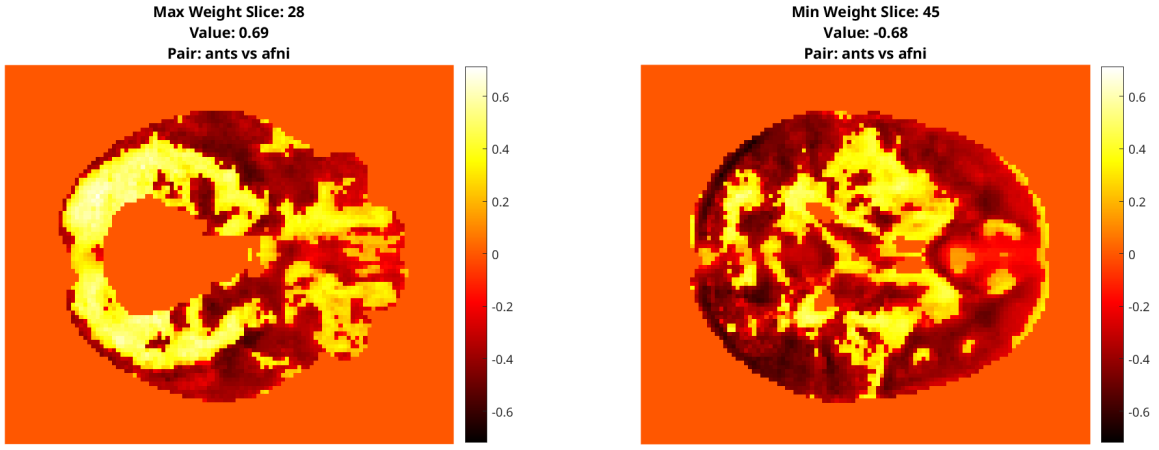

Fig. 6.6 Heat map illustrating differential weights captured by ANTs vs AFNI, with maximal weight: 0.6903 at slice 28 (X: 32, Y: 23, Z: 28) and minimal weight: -0.6784 at slice 45 (X: 27, Y: 20, Z: 45) The variation in colours from red to yellow depicts the different spatial patterns emphasised by each algorithm, suggesting their distinct sensitivities to neural signals within the Control Network.

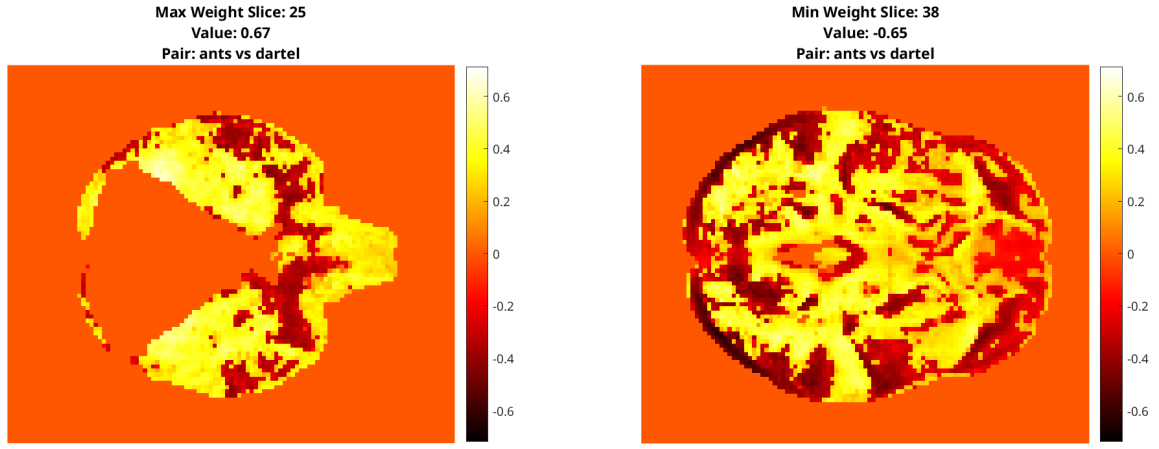

Fig. 6.7 Heat map depicting the contrast in weights between ANTs and DARTEL, with the highest weight: 0.6656 at slice 25 (X: 27, Y: 37, Z: 25) and the lowest: -0.6502 at slice 38 (X: 66, Y: 20, Z: 38). This visual representation indicates the unique spatial distribution of weights by each algorithm, reflecting their diverse approaches to brain activity classification.

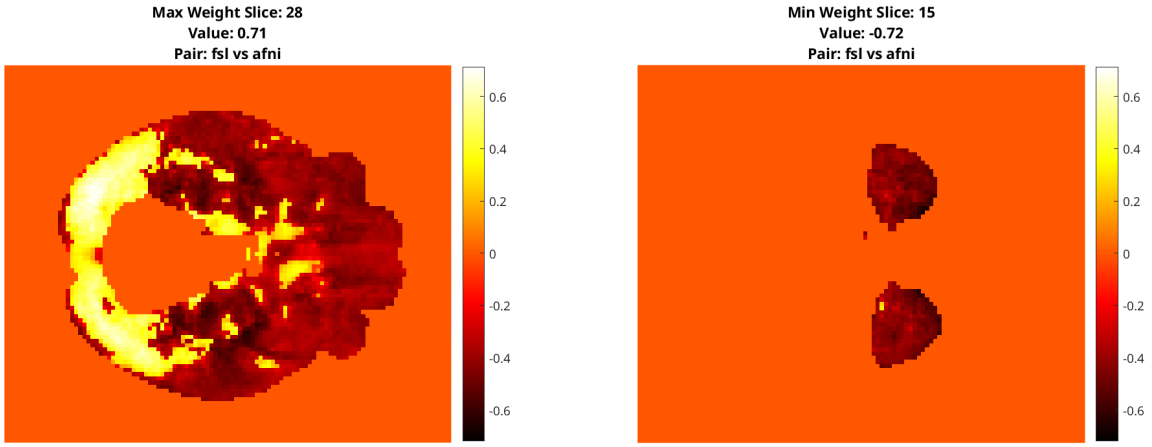

Fig. 6.8 Comparison of weight distributions for FSL vs AFNI, where the maximum weight: 0.7113 at slice 28 (X: 31, Y: 24, Z: 28) and the minimum weight: -0.7180 at slice 15 (X: 36, Y: 70, Z: 15). The heat map reveals the extent to which each algorithm identifies significant neural patterns, with FSL and AFNI showing variations in their representation of brain activations.

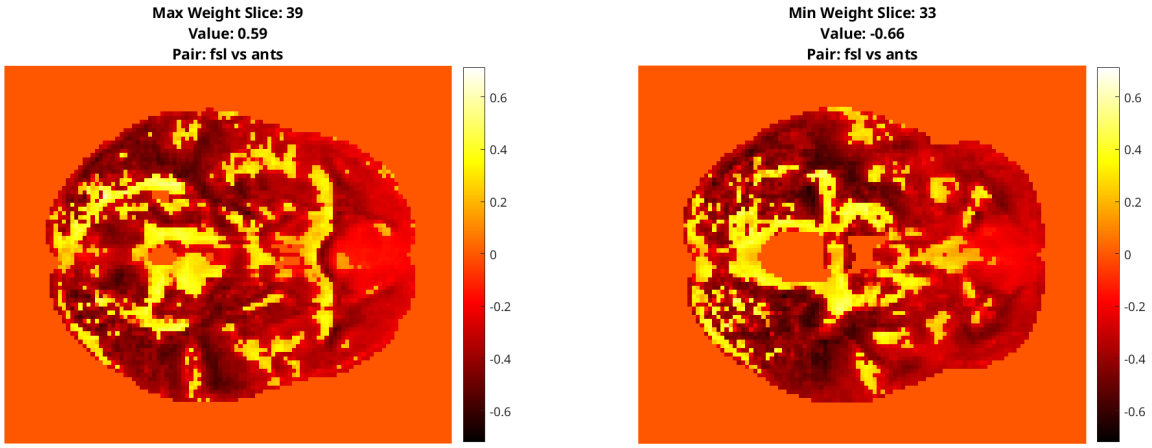

Fig. 6.9 Heat map showing FSL vs ANTs weight distribution comparison, with the highest weight: 0.5886 at slice 39 (X: 32, Y: 25, Z: 39) and the lowest: -0.6582 at slice 33 (X: 30, Y: 42, Z: 33). The colours indicate the degree of weight each algorithm assigns, with FSL and ANTs exhibiting different preferences for neural features, which is critical for algorithmic choice in neuroimaging studies.

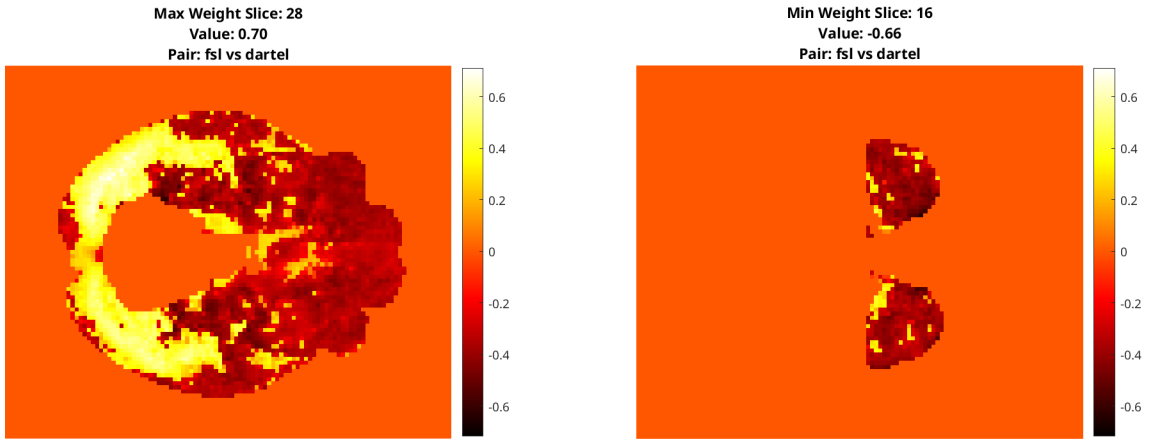

Fig. 6.10 Heat map analysis of FSL vs DARTEL displaying the maximum weight: 0.7027 at slice 28 (X: 28, Y: 26, Z: 28) and the minimum weight: -0.6648 at slice 16 (X: 37, Y: 70, Z: 16). The range of colours reflects the differential spatial emphasis of each algorithm, underscoring the importance of understanding these differences when selecting algorithms for neuroimaging analysis.

sitivities can inform more accurate modelling of inter-subject variability, leading to richer insights into the complexities of brain function [177]. Ultimately, this work suggests that algorithmic choice is not merely a technical decision but a crucial methodological factor that directly shapes the scientific interpretations of functional connectivity in neuroimaging research [170, 206].

6.2.2 Summary of MI, DSC, and MVPA Findings

This section synthesises the findings from the MI, DSC, and MVPA metrics, each of which uniquely contributes to understanding inter-subject variability and algorithmic performance.

- **Mutual Information**: MI provided a measure of shared information across registration algorithms, quantifying inter-subject consistency. Higher MI values, particularly in the FSL-AFNI pairing, indicated strong alignment in spatial activation patterns, suggesting these algorithms are comparably sensitive to functional signals within RSNs [41]. Conversely, lower MI values between ANTs and DARTEL highlighted significant differences in their approach to processing spatial variability, a factor critical when precision in structural alignment is prioritised [208].

- **Dice Similarity Coefficient**: The DSC metric offered insights into spatial similarity, with reordered heat maps providing a visual assessment of functional alignment

across subjects. ANTs demonstrated uniformly high DSC values, reflecting its robust ability to achieve consistent group-level alignment, while AFNI displayed variable similarity patterns, likely due to its sensitivity to subject-specific anatomical features. This variability in DSC performance suggests that ANTs may be more reliable for studies requiring consistent spatial alignment, while AFNI might be better suited for investigations focusing on individual anatomical variation [203].

- **Multivoxel Pattern Analysis**: Through pairwise SVM weight comparisons, MVPA revealed spatial biases unique to each algorithm. Distinct SVM weight distributions indicated algorithm-specific preferences in processing RSN features, with FSL and DARTEL emphasising high-resolution, localised brain activations, whereas ANTs captured broader network-level similarity. These variations highlight how algorithmic choices impact voxel-wise classification and emphasise the need for careful selection based on research goals, particularly for studies focused on functional connectivity [41].

Table 6.1 summarises the key findings from each metric, providing a side-by-side comparison of the four algorithms. This summary emphasises the distinctive contributions of each algorithm to neuroimaging analysis, informing researchers in choosing an algorithm aligned with their specific study objectives.

Table 6.1 Comparative Summary of MI, DSC, and MVPA Metrics for Non-Rigid Registration Algorithms

Algorithm	MI	DSC	MVPA
AFNI	Moderate similarity	Variable, sensitive to anatomy	Emphasis on specific regions
ANTs	High consistency	Uniformly high	Broad network similarity
DARTEL	Moderate, group-level alignment	Balanced between individual and group consistency	High precision in local alignment
FSL	Highest similarity with AFNI	Moderate, adaptive to individual features	High spatial precision and balanced connectivity

In summary, each metric emphasises distinct aspects of algorithmic performance: MI captures shared information across algorithms, DSC reflects spatial alignment consistency, and MVPA highlights unique algorithmic biases in brain network processing. Together, these findings highlight the need for tailored algorithm selection based on specific neuroimaging goals,

whether prioritising group-level consistency (as in ANTs), detailed individual analysis (AFNI and FSL), or a balanced approach (DARTEL) [41, 208].

6.3 Algorithmic Performance Analysis

Clustering and cluster size analysis are powerful techniques in neuroimaging, offering insights into algorithmic behaviour and their capacity to capture consistent functional connectivity patterns across individuals. These analyses allow researchers to examine how different algorithms cluster brain regions and to assess the distribution and consistency of these clusters, revealing algorithm-specific strengths and limitations in representing functional brain networks. By evaluating cluster sizes, researchers can infer each algorithm's sensitivity to individual variability versus group-level consistency—a critical factor for selecting the appropriate algorithm based on the study's objectives. As shown in previous studies, clustering in neuroimaging can provide a window into the inherent structure of functional connectivity and variability across subjects [41, 203, 206].

In this analysis, bar charts depicting cluster sizes and average DSC for each algorithm provide complementary views of their clustering performance. The bar charts illustrate the range and distribution of cluster sizes, shedding light on the consistency of functional network detection. Larger clusters suggest areas with more uniform within-network connectivity, while smaller, more varied cluster sizes may indicate the algorithm's flexibility in adapting to individual anatomical differences. In contrast, the DSC serves as a precise metric of spatial overlap within each cluster, providing a measure of registration accuracy. By examining these metrics together, we can achieve a nuanced understanding of each algorithm's performance in clustering and registration [202, 204].

6.3.1 Clustering Analysis

This section presents clustering results for AFNI, ANTs, DARTEL, and FSL, focusing on their distinct approaches to capturing functional connectivity patterns. Bar charts offer a visual representation of each algorithm's clustering tendencies, while average DSC quantify the spatial alignment accuracy. Together, these metrics extend beyond traditional evaluations, providing a detailed, comparative framework for assessing the fidelity of neuroimaging algorithms. This combination not only highlights the diversity in algorithmic behaviour but also acts as a guide for selecting the most suitable registration technique for specific research questions in RSN analysis.

Cluster Sizes Interpretation

In this section, bar graphs representing cluster sizes (Fig. 6.11) and their interpretations are introduced.

The AFNI cluster size graph reveals a varied distribution of cluster sizes, with notable peaks. This pattern suggests AFNI's selective sensitivity to particular functional brain regions, likely influenced by its focus on temporal correlation in fMRI data. Larger clusters indicate regions with stable within-network connectivity, where AFNI's strengths are most pronounced. However, the presence of smaller clusters suggests that AFNI's performance may fluctuate based on anatomical complexity, reflecting the algorithm's sensitivity to inter-subject variability [41].

ANTs' cluster size distribution shows a more consistent pattern across clusters, implying balanced detection of functional networks. This uniformity suggests that ANTs achieves consistent alignment without disproportionately focusing on any particular network, providing a balanced view of brain architecture. High DSC across clusters reinforce ANTs' reliability in capturing distributed neural activities, making it effective for comprehensive analyses where uniform network registration is prioritised [203].

DARTEL's cluster size distribution is characterised by sharply defined peaks, denoting the algorithm's precision in capturing coherent functional networks. This can be attributed to DARTEL's advanced high-dimensional warping technique, which facilitates precise spatial alignment. High DSC further support DARTEL's success in generating templates that maintain strong spatial overlap, particularly advantageous for studies requiring detailed morphological accuracy [204].

FSL's cluster size graph exhibits a broad range of cluster sizes, suggesting an adaptable approach to functional network detection. This flexibility may stem from FSL's integration of linear and non-linear registration methods, allowing it to capture both strong and subtle functional connectivity patterns. Variability in DSC indicates FSL's sensitivity to individual differences, making it well-suited for research that encompasses a diverse range of neuroanatomical profiles and connectivity patterns [206].

Summary of Clustering Analysis Findings: The clustering and DSC analyses reveal distinct characteristics of each algorithm in handling functional connectivity landscapes. AFNI's pronounced peaks suggest a focus on specific networks, beneficial for within-network connectivity studies. ANTs' uniformity indicates balanced alignment across brain regions, ideal for studies requiring broad network consistency. DARTEL's precise clustering aligns with detailed morphometric analyses, making it useful for spatially-specific investigations.

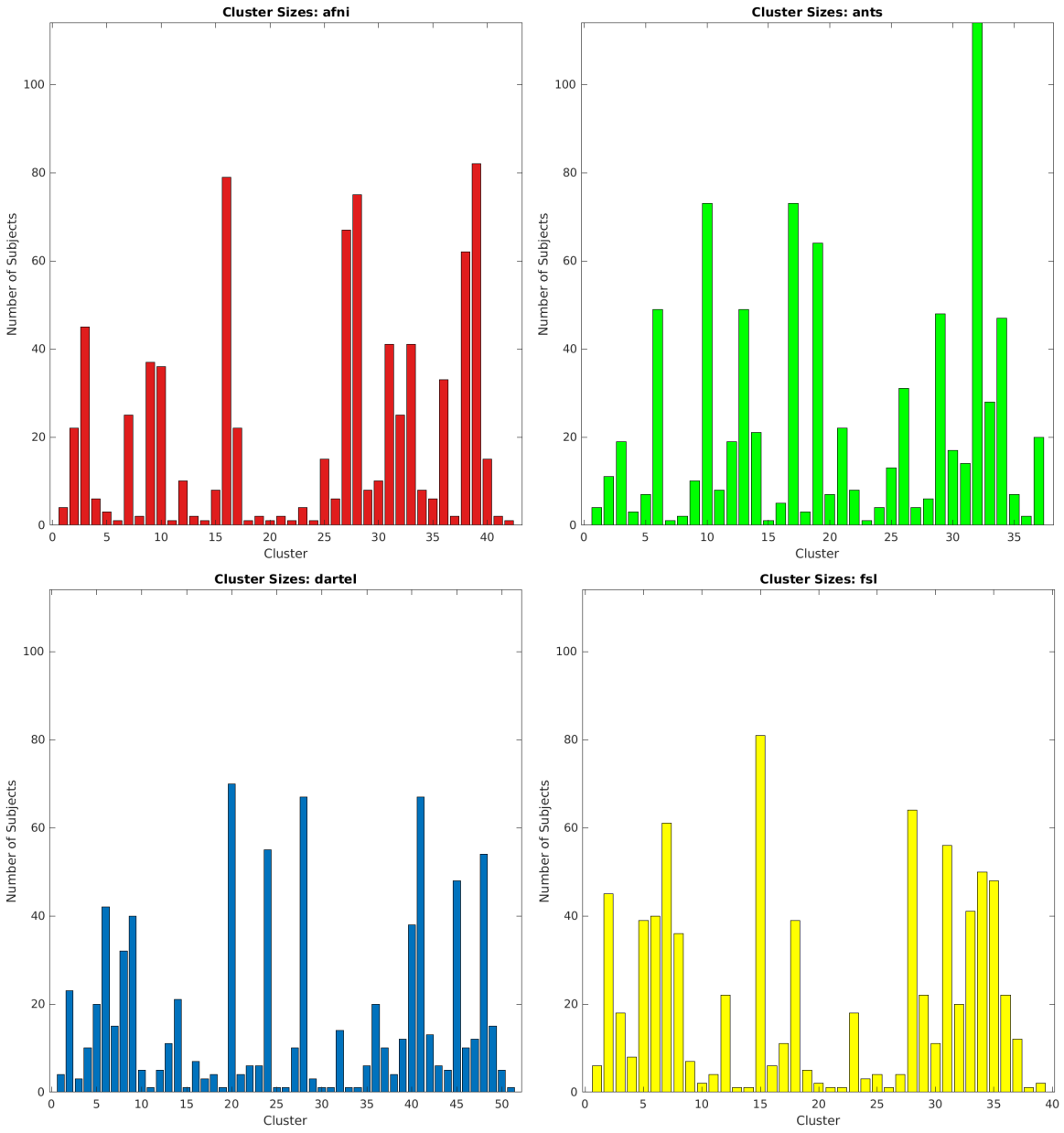

Fig. 6.11 Bar graphs representing the distribution of cluster sizes across subjects for each algorithm (AFNI, ANTs, DARTEL, FSL). The horizontal axis represents the cluster number, while the vertical axis indicates the number of subjects exhibiting each cluster size. AFNI shows a heterogeneous distribution, suggesting selective sensitivity, while ANTs displays a more uniform distribution, indicating balanced network detection. DARTEL reveals sharply defined peaks, reflecting precise detection of coherent networks, and FSL exhibits a broad range, highlighting flexibility in detecting diverse functional brain activities.

Finally, FSL's varied cluster sizes highlight its flexibility, supporting a broad range of neuroimaging applications where both individual variability and group-level consistency are important. Collectively, these findings emphasise the importance of aligning algorithm selection with research objectives, facilitating informed choices for neuroimaging studies in RSN analysis [41, 203, 206].

6.3.2 Cluster Size Analysis and DSC Interpretation

Understanding the effectiveness of neuroimaging registration algorithms in functional connectivity analysis necessitates an examination of clustering patterns and spatial consistency across subjects. The DSC serves as a measure of spatial overlap, allowing for the quantification of alignment precision in functional brain networks among individuals. Through the assessment of cluster sizes and DSC, insights can be gained into each algorithm's approach to spatial registration, offering critical information on the suitability of different algorithms for specific research applications [41, 203].

In this section, we present the average DSC per cluster for each algorithm: AFNI, ANTs, DARTEL, and FSL. This analysis highlights each algorithm's consistency in capturing functional network structures, showcasing their relative strengths in terms of spatial alignment precision.

Cluster DSC Analysis

The visual representation of cluster DSC analysis can be seen in Figure 6.12.

For the AFNI Average DSC graph, variability is observed across clusters. Higher DSC in certain clusters reflect robust spatial overlap and consistent registration among subjects, suggesting AFNI's capability in accurately capturing specific functional networks. However, the presence of clusters with lower DSC implies that AFNI may face challenges in areas with more complex functional connectivity or greater anatomical variability across subjects. The observed variability in DSC scores reflects AFNI's selective strengths in network registration but also indicates that its efficacy may fluctuate based on the brain region being analysed.

In contrast, the ANTs average DSC graph exhibits a more uniform distribution of DSC, with most clusters displaying moderate to high values. This uniformity suggests that ANTs is proficient in delivering consistent and reliable registration of functional connectivity across subjects. The absence of clusters with very low DSC highlights ANTs' robustness and its ability to align various brain networks without a disproportionate emphasis on any single

functional area. As such, ANTs may be particularly valuable for studies that require reliable inter-subject comparison of functional networks.

DARTEL's average DSC graph is characterised by clusters with consistently high DSC, indicating a high degree of spatial overlap across subjects' brain images. This result suggests that DARTEL effectively creates a group-level template with a significant agreement in spatial alignment among subjects, making it highly suitable for studies that require detailed anatomical localisation and spatial specificity. DARTEL's performance, with uniformly high DSC values, highlights its capability in morphometric analyses where precise structural alignment is essential.

The FSL average DSC graph displays a broader range of DSC, which may reflect FSL's adaptive approach to registration. The diversity in DSC scores indicates FSL's ability to capture both prominent and subtle functional connectivity patterns, accommodating the anatomical variability found across different subjects. While some clusters show high DSC, suggesting strong alignment, other clusters with lower coefficients imply FSL's flexibility in adjusting to subject-specific anatomical features. This characteristic positions FSL as an effective tool for studies that benefit from a balance of capturing common and unique patterns in functional connectivity.

Summary of DSC Analysis: The comparative analysis of DSC graphs across AFNI, ANTs, DARTEL, and FSL reveals distinct algorithmic behaviours in capturing functional brain networks. AFNI exhibits variable DSC, suggesting its selective efficacy in network registration but with occasional limitations in regions with higher anatomical complexity. ANTs provides a more uniform registration performance, indicating robust and unbiased spatial alignment, beneficial for inter-subject comparative studies. DARTEL's consistently high DSC demonstrate its strength in high-precision registration, making it suitable for applications requiring fine-grained anatomical specificity. FSL's flexible approach is evident in its wide DSC range, showcasing its ability to balance capturing common connectivity patterns and accommodating subject-specific anatomical features.

These findings underline the unique attributes of each algorithm, emphasising the importance of selecting an appropriate registration tool based on the specific requirements of neuroimaging research. For studies demanding consistent alignment across multiple brain networks, ANTs and DARTEL may offer more reliable solutions. Conversely, AFNI and FSL present adaptable approaches suited for analyses that prioritise selective functional regions or individual variability [203, 41].

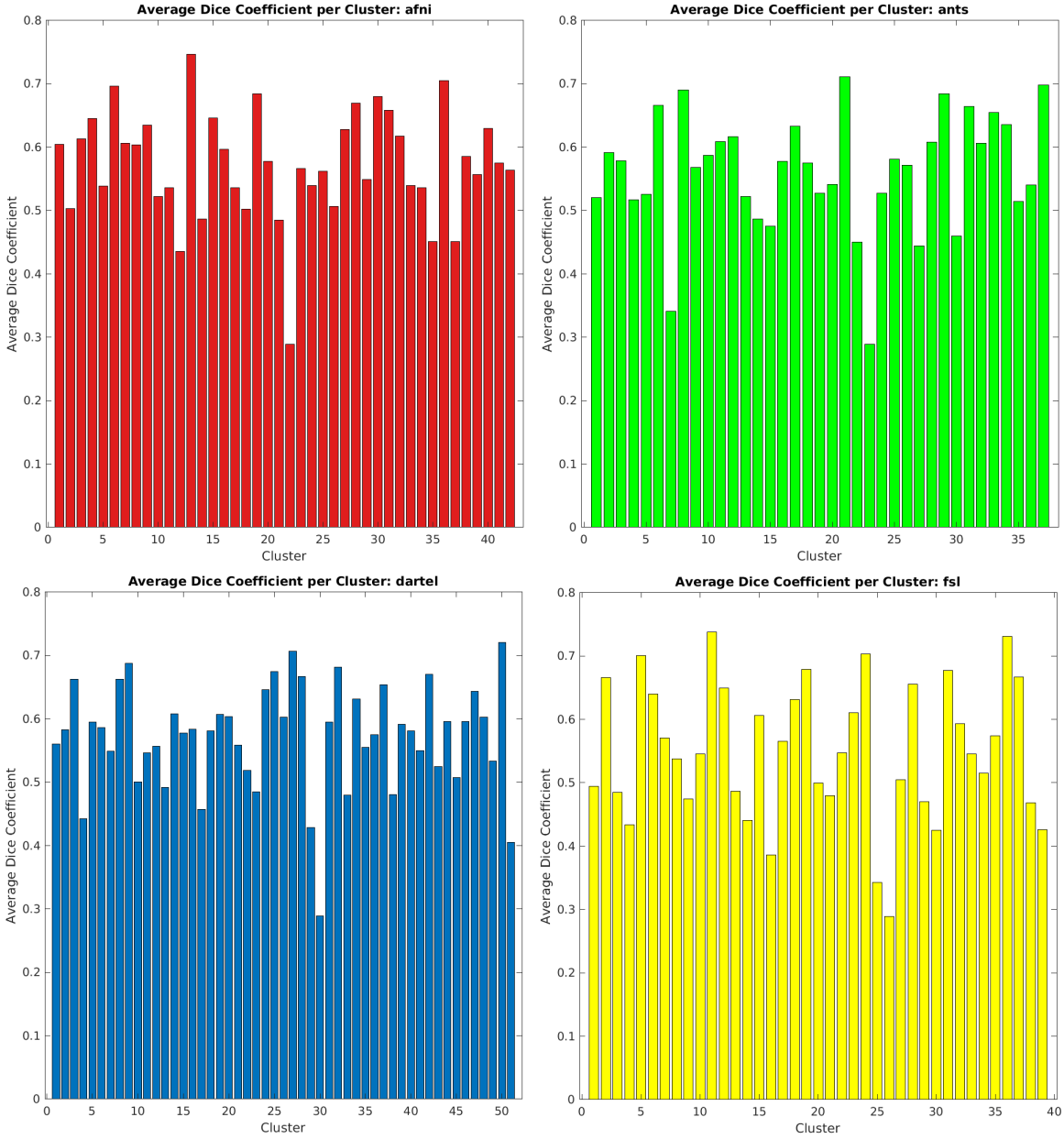

Fig. 6.12 Bar charts of average DSC per cluster for each algorithm (AFNI, ANTs, DARTEL, FSL). The horizontal axis represents individual clusters, while the vertical axis shows the DSC, indicating spatial overlap consistency among subjects. AFNI displays variability across clusters, reflecting alignment challenges in certain regions. ANTs maintains a stable range, indicating consistent connectivity registration. DARTEL's uniformly high coefficients suggest precise spatial alignment, and FSL shows a broad range, highlighting adaptability to anatomical variability across subjects.

6.4 Chapter Summary

Chapter 6 addresses the central thesis question of how algorithmic differences in non-rigid registration influence the robustness and reproducibility of rs-fMRI analyses, particularly in terms of inter-subject variability and functional connectivity mapping. Through a comprehensive analysis employing MI, DSC, and MVPA, we evaluated the distinct contributions and biases of four state-of-the-art algorithms—AFNI, ANTs, DARTEL, and FSL—shedding light on their practical implications for neuroimaging research.

The MI analysis quantified the shared information across algorithms, revealing how each method preserved functional connectivity patterns across subjects. Higher MI values between certain algorithm pairs (e.g., FSL and AFNI) indicated a stronger agreement in capturing functional signal patterns, while lower MI values (e.g., ANTs vs. DARTEL) highlighted the specific structural biases each algorithm introduces. This suggests that certain algorithm pairs may be more suited to studies focusing on signal reproducibility, whereas others may capture unique structural details relevant for specific connectivity patterns.

The DSC analysis offered insight into the spatial similarity of aligned brain regions, indicating how consistently each algorithm registered functional regions across subjects. The reordered heat maps illustrated algorithmic distinctions in handling anatomical variability: AFNI showed variability across clusters, reflecting its sensitivity to anatomical differences, while ANTs demonstrated high uniformity in registration, suggesting robust performance across varied anatomical landscapes. DARTEL excelled in producing spatially precise group-level templates, ideal for studies requiring anatomical specificity, whereas FSL displayed flexibility, effectively adapting to both commonalities and individual differences, making it versatile for broad neuroimaging applications.

MVPA extended these insights by identifying spatial patterns uniquely emphasised by each algorithm. Pairwise SVM weight comparisons revealed the spatial biases inherent to each method, with DARTEL and FSL focusing on high-resolution, localised brain activations, while ANTs captured broader network-level similarity. This level of granularity highlighted the importance of algorithm selection based on specific research goals, particularly in functional connectivity studies where the accuracy of spatial patterns is critical.

In summary, Chapter 6 systematically demonstrates that algorithmic choices have substantial effects on the outcomes of rs-fMRI analyses, directly impacting the consistency and interpretability of functional connectivity data. These findings emphasise the need for informed algorithm selection tailored to research objectives, enhancing the reliability of neuroimaging analyses. By offering insights into each algorithm's strengths and limitations, this

chapter provides a foundation for future work aimed at refining neuroimaging methodologies to improve reproducibility and applicability, ultimately benefiting both research and clinical neuroimaging applications.

Looking ahead, Chapter 7 will synthesise the key findings from previous analyses, discussing their implications for the broader field of neuroimaging and suggesting best practices for algorithm selection in rs-fMRI studies. This discussion will highlight how these methodological insights contribute to enhancing reproducibility and interpretability in computational neuroscience.

Chapter 7

Discussion

This chapter synthesises the findings of the thesis and contextualises them within the broader field of neuroimaging, focusing specifically on the impact and implications of non-rigid registration algorithms in Resting-State Functional Magnetic Resonance Imaging (rs-fMRI). This thesis addresses central research questions concerning the impact of algorithmic variability on Resting-State Network (RSN) analysis, a core element for understanding intrinsic functional connectivity.

The main objectives include investigating the variability introduced by different non-rigid registration algorithms, FMRIB Software Library (FSL), Advanced Normalisation Tools (ANTs), Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL), and Analysis of Functional NeuroImages (AFNI), and examining how this impacts rs-fMRI reliability. To address these aims, this thesis developed the Non-Rigid Registration Algorithm Analysis Framework (NRAAF), enabling standardised and comparative evaluation of algorithms and highlighting the need for tailored algorithm selection in different research contexts [209, 210].

7.1 Synthesis of Key Findings

This section provides an integrated synthesis of significant findings from Chapters 4, 5, and 6. By emphasising cross-chapter patterns and comparative insights, this synthesis consolidates findings and highlights implications for computational methods, algorithm performance, and neuroimaging applications.

Algorithmic Sensitivity and Specificity in Peak Activation Detection

The comparative analysis across Chapters 4 and 5 highlights the variability in sensitivity and specificity among non-rigid registration algorithms, particularly in detecting Peak Activation Intensities within rs-fMRI data. Findings reveal that ANTs and FSL show heightened sensitivity but demonstrate occasional outliers, contrasting with DARTEL's consistent detection and reduced variability [211, 212]. These results suggest that ANTs and FSL are preferable in studies prioritising high activation sensitivity, whereas DARTEL offers reliability where consistent, low-variability detection is paramount. These insights emphasise that context-specific algorithm selection is critical for robust rs-fMRI analyses, directly influencing the reliability of results by aligning the sensitivity requirements with research goals.

Cluster Detection and Network Integrity

Chapter 5 findings demonstrate that registration algorithms significantly influence the detection and representation of RSNs, particularly regarding the capture of significant clusters and network integrity. DARTEL's ability to detect a higher mean number of clusters enhances the breadth of network representation; however, this advantage is offset by increased variability. Conversely, FSL and ANTs show lower variability in cluster detection but may limit the extent of network detail captured, suggesting their suitability for studies requiring consistent network boundaries over broad coverage. This dual insight emphasises the role of NRAAF in systematically mapping the performance trade-offs across algorithms, aiding in methodological alignment with specific neuroimaging goals and ensuring network integrity [213, 214].

Inter-Subject Variability and Implications for Group Inference

In Chapter 6, inter-subject variability is explored with respect to group inference in rs-fMRI, utilising metrics such as Mutual Information (MI) and Dice Similarity Coefficient (DSC). The analyses reveal that FSL and AFNI exhibit high similarity in inter-subject alignment, while ANTs and DARTEL display varying degrees of alignment stability. For studies emphasising group-level reproducibility, FSL and AFNI offer greater reliability, whereas ANTs and DARTEL may be advantageous for investigations into individual-specific patterns and personalised neuroimaging applications. This distinction enhances the reproducibility of findings across multi-subject studies by matching algorithm selection to the study's alignment consistency requirements, directly impacting the scalability and validity of neuroimaging research [41, 215].

7.2 Return to Research Questions and Objectives

This research critically evaluates the impact of non-rigid registration algorithms on rs-fMRI outcomes, directly addressing the research question: *Do differences in the performance of registration algorithms lead to variability in the outcomes of neuroimaging analyses?* The findings substantiate the hypothesis that non-rigid registration algorithms exhibit variability in their impact on rs-fMRI, specifically influencing sensitivity, reliability, and spatial accuracy in functional connectivity patterns. This variability emphasises the importance of algorithmic choice in achieving consistent and accurate activation maps.

Addressing the Stated Objectives:

- **Objective 1: Systematic Review of Challenges in Non-Rigid Registration**
  The review highlights variability in algorithm performance and the lack of a standard framework for comparison. This foundation informed NRAAF's development and supports rs-fMRI reproducibility [216].
- **Objective 2: Development of NRAAF for Comparative Evaluation**
  The NRAAF framework was designed to evaluate algorithm consistency, reproducibility, and RSN impact, demonstrating how different techniques influence rs-fMRI outcomes. NRAAF assists in selecting appropriate algorithms to enhance accuracy and reliability [217].
- **Objective 3: Validate NRAAF Using a Large Dataset**
  Applying NRAAF to a large dataset (n=815) confirmed differences in activation clusters, substantiating the research question by showing how algorithmic differences affect neuroimaging consistency [218, 18].
- **Objective 4: Characterise Non-Rigid Registration Algorithms in rs-fMRI**
  Through NRAAF, variability and dependencies across algorithms were characterised, revealing their impact on RSN analysis. This objective informs algorithmic choices, supporting reliable neuroimaging methodologies and supports informed study design [219].

7.3 Interpretation of Findings

This section probes into a comprehensive interpretation of the findings, situating them within computational neuroscience, neuroimaging accuracy, and clinical applications. The

discussion builds on the results of previous chapters, extending insights into the strengths and limitations of non-rigid registration algorithms, and exploring methodological rigour in neuroimaging studies.

Implications for Computational Neuroscience

The NRAAF advances computational neuroscience by introducing a systematic tool for quantifying algorithmic variability, addressing a gap highlighted in recent comparative studies [5, 18]. The framework's standardised evaluation is particularly pertinent in high-resolution neural network mapping, where algorithmic nuances can influence functional connectivity patterns essential for network-based analysis [3]. This work highlights the critical need for reproducibility in algorithm selection, which has been a recurring theme in the evolving landscape of neural network analysis [49]. By illuminating how algorithmic choices impact spatial precision, NRAAF supports computational practices that demand transparency and adaptability based on context-specific goals.

Neuroimaging Accuracy and Methodological Rigour

The findings reflect both strengths and constraints across various algorithms, with each offering distinct performance benefits and limitations. For instance, ANTs demonstrates sensitivity in detecting peak activations but may produce outliers, a phenomenon supported by previous reproducibility studies that call for standardisation in high-sensitivity algorithms [48, 162]. Conversely, DARTEL provides stable measures, aligning with the demands of confirmatory neuroimaging where consistency is crucial [60]. NRAAF's framework aligns with recent calls for reproducible neuroimaging practices, such as those presented in [49], reinforcing the importance of methodological rigour in functional imaging [18]. The framework facilitates a clearer understanding of how registration choices affect network integrity, contributing to the discourse on reproducibility in brain mapping studies [202].

Clinical Practice and Real-World Applications

The clinical implications of this work are considerable, as NRAAF provides valuable insights into the selection of algorithms for diagnostic imaging. For instance, DARTEL's stability in cluster detection suggests its applicability in routine diagnostics, aligning with recent studies on stability requirements in clinical neuroimaging [12]. Alternatively, ANTs' heightened sensitivity could benefit exploratory analyses, such as early-stage assessments where sensitivity

to subtle changes is essential [220]. By guiding algorithmic choices based on objective evaluations, NRAAF supports improved consistency and accuracy in clinical pipelines, facilitating more informed decision-making in settings where diagnostic precision is essential [221]. This is particularly valuable given the variability in algorithmic performance documented in comparative neuroimaging literature [116].

7.4 Limitations of the Current Work

This section outlines several key limitations encountered in the development and application of NRAAF, focusing on scalability, computational demands, data constraints, and algorithmic variability. Addressing these limitations can refine NRAAF's robustness and broaden its applicability in neuroimaging research.

Scalability and Computational Constraints

While NRAAF utilises parallel processing to enhance efficiency, the framework's computational demands are substantial, especially for large or multimodal datasets. This restricts NRAAF's accessibility in resource-limited settings, as it requires significant processing power for tasks like Multivoxel Pattern Analysis (MVPA) and Support Vector Machine (SVM) applications. To expand scalability, future iterations could incorporate adaptive processing techniques, such as optimised data pipelines and lightweight algorithmic modules, to reduce computational requirements without compromising analytical depth [222, 5].

Algorithmic Sensitivity and Variability

NRAAF revealed considerable variability in peak activation localisation and cluster integrity across algorithms, highlighting their unique sensitivities and specificities in rs-fMRI analyses. This variability emphasises the need for caution in comparing results across different algorithms, as each brings distinct strengths and limitations. Integrating adaptive machine learning models, such as deep learning algorithms, could help NRAAF adjust evaluations based on data characteristics, enhancing its accuracy and flexibility across diverse study contexts [57, 41].

Inter-Subject Variability and Data Generalisability

The analysis highlighted how demographic and physiological factors impact rs-fMRI results, complicating generalisation across populations. Factors like age and health status introduce

variability that affects network detection and interpretation. Incorporating demographic stratification or covariate adjustments within NRAAF could mitigate these effects, improving the framework's reliability across heterogeneous sample groups [223, 213].

Methodological Constraints in Evaluation Approaches

NRAAF, while offering a standardised framework for evaluating non-rigid registration algorithms, relies on parametric methods that impose limitations in handling large-scale or complex deformations often seen in advanced neuroimaging. Specifically, its use of the General Linear Model (GLM) assumes normality and homogeneity, which may limit adaptability with highly variable neuroimaging data.

Incorporating non-parametric alternatives, such as permutation testing, could reduce these constraints by avoiding strict data assumptions, making NRAAF more flexible for diverse datasets. Similarly, adopting metaheuristic techniques like genetic algorithms or swarm optimisation could enhance performance, especially with complex or high-dimensional data. Structures such as embedded deformation graphs or point cloud models may further support NRAAF's utility in complex neuroimaging scenarios, broadening its applicability beyond the current parametric framework [216, 224].

Data Quality and Preprocessing Constraints

The reliability of NRAAF's evaluations depends on data quality, as inconsistencies in preprocessing or acquisition can bias results. This is particularly relevant in multi-site studies where differences in scanner settings or protocols affect comparability. Future versions of NRAAF could incorporate automatic quality assessment tools and align preprocessing steps with standardised frameworks (e.g., fMRIPrep [30]) to enhance consistency across datasets, ensuring that findings reflect true algorithm performance rather than artifacts of data variability.

In summary, while NRAAF advances non-rigid registration algorithm evaluation, these limitations highlight the need for ongoing methodological enhancements to maximise its utility and ensure robust, reproducible neuroimaging research across varied applications.

7.5 Comparative Algorithm Performance and Recommendations

Based on the findings of the NRAAF, this research offers a comparative analysis of the performance of FSL, ANTs, DARTEL, and AFNI in rs-fMRI applications. The differences in each algorithm's sensitivity, stability, and alignment accuracy highlight specific scenarios where each may be optimally applied. These insights inform best practices for algorithm selection, supporting methodologically robust neuroimaging research.

Peak Activation Sensitivity and Specificity

The analysis revealed that ANTs and FSL demonstrate heightened sensitivity in detecting Peak Activation Intensities. This sensitivity allows these algorithms to capture subtle variations in neural activation, making them particularly valuable in exploratory studies focused on uncovering novel connectivity patterns. However, the increased sensitivity also introduces variability, including the occasional presence of outliers, which may limit interpretability in studies where stability is essential. In contrast, DARTEL offers more consistent measurements of peak activation with reduced variability. This stability makes DARTEL well-suited for confirmatory studies or clinical applications where consistent results are crucial for interpreting functional connectivity reliably [209]. Therefore, choosing between these algorithms should depend on study objectives: ANTs and FSL are better suited for hypothesis-generating research, while DARTEL offers more stable outcomes for studies prioritising consistency.

Cluster Detection and Network Integrity

The NRAAF findings suggest that DARTEL excels in identifying a greater number of clusters, offering an extensive view of network integrity within RSNs. This characteristic is advantageous in exploratory research where understanding the full extent of brain connectivity is critical. However, DARTEL's variability in cluster count suggests it may not always provide the consistency required for reproducible population studies. Conversely, FSL and ANTs deliver more stable cluster counts across different subjects, making them preferable for population-based analyses where consistency across samples is crucial. This distinction allows researchers to choose based on study needs: DARTEL for extensive network mapping and FSL or ANTs for high reproducibility across samples [211].

Alignment Accuracy and Inter-Subject Variability

FSL and AFNI show strong consistency in alignment accuracy across subjects, making them particularly effective in group-level analyses where uniformity in detecting functional signals is essential. This consistency supports the use of FSL and AFNI in studies that aim to generalise findings across populations with minimal alignment discrepancies. On the other hand, ANTs and DARTEL exhibit greater alignment variability but offer the advantage of capturing detailed, individualised neural patterns. These attributes make ANTs and DARTEL more suitable for personalised neuroimaging analyses, where nuanced anatomical distinctions are relevant to the study goals, such as in personalised medicine or research focusing on individual differences. This capability to discern personalised brain patterns highlights ANTs and DARTEL as valuable tools for precision neuroimaging applications, supporting detailed and individualised neural mapping [25].

7.6 Broader Implications for Neuroimaging and Clinical Practice

This research emphasises the broader implications of algorithm selection in both scientific and clinical contexts. By standardising algorithm evaluations, NRAAF promotes reproducibility and provides practical applications that benefit both research and clinical settings, thereby advancing the field of neuroimaging.

Implications for Neuroimaging Standards

The NRAAF framework contributes significantly to neuroimaging standards by offering a systematic, replicable methodology for evaluating non-rigid registration algorithms. This standardisation helps bridge methodological gaps, allowing researchers to conduct consistent, reproducible comparisons across studies. By providing insights into sensitivity, specificity, and alignment reliability, NRAAF aids in identifying the most appropriate algorithms for specific research purposes, thereby supporting reproducibility in computational neuroscience. This consistency is especially critical in large-scale studies where algorithmic variability can undermine the reliability of cross-study comparisons and meta-analyses [225].

Clinical Applications and Diagnostic Relevance

The insights derived from NRAAF's comparative analysis reveal that algorithm choice directly affects the reliability of functional connectivity mapping, a cornerstone in clinical diagnostics. Algorithms like DARTEL, which show extensive cluster detection, may support exploratory clinical research, such as investigating novel biomarkers or understanding early-stage brain connectivity disruptions in neurodegenerative diseases. In contrast, FSL and ANTs offer stability and reproducibility, which are essential for routine diagnostics where consistent mapping across patients is vital. For instance, reproducible network mapping is crucial in assessing functional changes associated with disorders such as Alzheimer's Disease, where consistent tracking of neural degeneration patterns is necessary [77, 7]. Thus, the findings enable informed algorithm choices based on the diagnostic requirements, potentially enhancing the accuracy and reliability of neuroimaging-based diagnoses.

Supporting Personalised Medicine

As precision medicine gains prominence, the need for algorithms that can accurately capture individual anatomical variability becomes essential. ANTs and DARTEL, with their heightened sensitivity to individualised patterns, are particularly well-suited to support personalised diagnostics, where the focus is on capturing subtle, subject-specific variations in neural connectivity. These algorithms can provide clinicians with detailed representations of individual brain structures, improving accuracy in diagnostics and treatment planning for neurological conditions. By enabling a more tailored approach to neuroimaging, NRAAF supports advancements in personalised medicine, aligning with the broader goal of achieving precision healthcare in clinical neuroscience [41, 226].

Advancement in Computational Neuroscience

NRAAF represents a critical advancement in computational neuroscience by offering a robust framework for reliable algorithm evaluation. By systematically quantifying algorithmic differences, NRAAF enhances consistency across neuroimaging studies, ensuring that technical innovations are methodologically sound and clinically applicable. This methodological rigour is crucial as the field advances toward more complex, high-resolution neuroimaging analyses. Additionally, NRAAF's insights into alignment accuracy, sensitivity, and inter-subject variability provide valuable guidance for algorithm selection, bridging the gap between computational advancements and clinical needs. In this way, NRAAF supports a more integrated

approach to neuroimaging, where scientific and clinical applications mutually inform each other, advancing both fields [216].

In summary, NRAAF significantly contributes to neuroimaging practices by supporting informed, consistent algorithm selection. Its role in promoting methodological standardisation and clinical applicability highlights its value in fostering both scientific rigour and practical impact, ultimately advancing neuroimaging as a precise and reliable tool for both research and diagnostics [9, 2].

7.7 Chapter Summary

This chapter synthesised the key findings of the thesis and explored their broader implications within computational neuroscience and clinical neuroimaging practice. Through the development and application of NRAAF, the research systematically evaluated the impact of various non-rigid registration algorithms on rs-fMRI analysis, addressing the central research question of whether algorithmic choice influences the reliability and interpretability of neuroimaging outcomes.

Key findings from Chapters 4, 5, and 6 demonstrated that algorithmic variability significantly affects peak activation detection, cluster integrity, and alignment accuracy, emphasising the importance of selecting algorithms tailored to study-specific goals. Algorithms like ANTs and FSL exhibited heightened sensitivity, making them valuable in exploratory studies, while DARTEL's stability favoured confirmatory analyses and clinical applications requiring reproducibility. Additionally, NRAAF's insights into inter-subject variability supported algorithm choices that balance consistency for group-level inferences with precision for individualised analyses.

The chapter also discussed the practical implications of NRAAF, highlighting its role in advancing reproducibility in neuroimaging through standardised algorithm evaluation. Limitations encountered during the development of NRAAF—such as scalability constraints, algorithmic variability, and data quality factors—were addressed, with recommendations for future adaptations to enhance its applicability in diverse neuroimaging contexts mentioned in Chapter 8. The framework's methodological rigour and adaptability make it a valuable tool for advancing both scientific research and clinical diagnostics in computational neuroscience.

In summary, NRAAF contributes to a more precise and consistent approach to neuroimaging analysis, aligning technical advancements with practical applications that support accurate and reproducible findings in both research and clinical settings.

Chapter 8

Conclusion & Future Work

In this concluding chapter, the findings from each preceding chapter are integrated to present a holistic overview of the potential for further advancements in evaluating neuroimaging registration algorithms. This work has aimed to address the variability in algorithmic performance within the context of Resting-State Functional Magnetic Resonance Imaging (rs-fMRI) processing, a crucial consideration for ensuring accuracy and reproducibility in neuroimaging analyses. Through developing and applying the Non-Rigid Registration Algorithm Analysis Framework (NRAAF), this research has systematically evaluated the strengths and limitations of prominent non-rigid registration algorithms, contributing to both the scientific understanding and practical methodologies used in computational neuroscience.

Each chapter contributes to this framework in a specific way:

- Chapters 1 and 2 established the foundational knowledge, reviewing the theoretical underpinnings of image registration and identifying the primary challenges associated with variability in non-rigid registration for neuroimaging. This literature review contextualised the research within current gaps, particularly in the accuracy and standardisation of functional connectivity studies.
- Chapter 3 introduced the NRAAF framework, detailing its methodology, algorithm selection, and evaluation metrics. This framework was developed to support a rigorous comparative analysis of algorithms across different functional connectivity parameters, setting the stage for subsequent empirical assessments.
- Chapters 4, 5, and 6 provided empirical assessments of registration algorithms across critical neuroimaging metrics—Peak Activation Intensity, Significant Clusters, inter-subject variability, and group inference analysis. Each chapter systematically applied

the NRAAF framework, producing quantitative insights into the performance of state-of-the-art algorithms such as FMRIB Software Library (FSL), Advanced Normalisation Tools (ANTs), Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL), and Analysis of Functional NeuroImages (AFNI) in terms of their accuracy, sensitivity, and alignment consistency.

- Chapter 7 synthesised these findings, discussing their implications in computational neuroscience and clinical practice. This chapter emphasised the importance of informed algorithm selection and highlighted how variability impacts both research outcomes and real-world applications, particularly in personalised medicine and diagnostics.

This chapter concludes by reinforcing NRAAF's important role in providing a standardised and reproducible approach to evaluating non-rigid registration algorithms in neuroimaging. By systematically addressing algorithmic sensitivity, specificity, and stability, NRAAF equips researchers with data-driven insights that support methodological consistency and enhance translational utility in computational neuroscience. The chapter's final sections extend this foundation, exploring future research directions that can expand NRAAF's adaptability and impact across diverse neuroimaging applications, thereby aligning the framework with the evolving needs of both scientific investigation and clinical practice.

8.1 Future Work

The findings presented in this thesis not only emphasise the current capabilities of the NRAAF framework but also reveal essential pathways for advancing its adaptability and impact across neuroimaging research. Key areas for future development include improving scalability, enhancing computational efficiency, addressing inter-subject and algorithmic variability, and incorporating advanced Machine Learning (ML) techniques. These enhancements aim to broaden NRAAF's applicability, allowing it to meet the growing demands of diverse and increasingly complex neuroimaging datasets. Building on the limitations identified in Chapter 7, this section outlines specific strategies to augment NRAAF's performance and applicability, promoting a more nuanced and robust framework for neuroimaging algorithm evaluation.

Enhanced Computational Efficiency

As explored in Chapters 3 and 6, NRAAF's computational demands, particularly for Multi-voxel Pattern Analysis (MVPA) and Support Vector Machine (SVM) applications, can be

prohibitive in large-scale neuroimaging studies. Future research should prioritise optimisations that reduce these demands, such as leveraging cloud-based or distributed computing solutions, dynamic load balancing, and memory-efficient algorithms [82, 222]. Such advancements would make NRAAF more accessible to a broader range of research and clinical contexts, facilitating its use in high-throughput and resource-constrained settings alike.

Addressing Inter-Subject and Algorithmic Variability

Inter-subject and algorithmic variability remain critical challenges in rs-fMRI, as indicated by the findings in Chapters 5 and 6. To address these challenges, incorporating ML models that adapt in real-time to individual anatomical and physiological differences could enhance NRAAF's flexibility and precision [41]. Reinforcement learning, for example, could enable NRAAF to dynamically adjust algorithmic parameters for each subject, supporting personalised neuroimaging analysis [218]. Additionally, Bayesian optimisation could facilitate data-driven parameter adjustments, improving algorithm selection accuracy across varied anatomical structures [216].

Expansion of Evaluation Models

Chapters 4 through 6 suggest the potential for integrating nature-inspired and metaheuristic methods to broaden NRAAF's analytical scope. Approaches such as genetic algorithms, swarm intelligence, and Majorisation-Minimisation (MM) can effectively handle complex and non-linear alignment tasks [227, 18]. Implementing these methods could improve NRAAF's ability to handle diverse datasets, enhancing its applicability in multimodal neuroimaging and high-dimensional analyses. Additionally, a modular, user-selectable model that tailors evaluation metrics to specific dataset characteristics could further enhance NRAAF's flexibility and utility [223].

Standardisation for Reproducibility

Achieving reproducibility across neuroimaging sites is crucial, particularly in multi-centre studies where inter-site variability can hinder generalisability. Establishing standard evaluation criteria and reproducibility protocols for NRAAF, potentially aligned with workflows like fMRIPrep [30, 77], could help address these challenges. Developing a reproducibility toolkit with templates, guidelines, and shared resources would support consistent application of NRAAF across diverse sites, fostering transparency and best practices within the field

[228]. A shared repository for reproducibility metrics could further promote cross-study comparability, strengthening NRAAF's contributions to standardisation in neuroimaging.

Integration of Machine Learning in Algorithm Optimisation

Chapters 6 and 7 highlight promising applications of ML in optimising registration algorithms. Embedding ML models, particularly deep learning and reinforcement learning, within NRAAF could enhance its adaptability for complex anatomical deformations, supporting a more nuanced understanding of functional connectivity patterns [3]. Explainable Artificial Intelligence (XAI) models could also increase transparency in algorithmic decision-making, allowing NRAAF to clarify model outputs and enhance interpretability in clinical applications. Such advancements would not only improve precision but also support NRAAF's utility in diverse clinical and research settings by providing algorithmic adaptability for complex imaging conditions [120, 229].

In addressing these future directions, the NRAAF framework can evolve to meet the needs of a rapidly advancing field, where both methodological rigour and clinical relevance are essential. As the chapter moves to the final conclusion, these proposed advancements collectively highlight the framework's potential for shaping next-generation neuroimaging practices, ensuring NRAAF's continued relevance and impact across computational neuroscience and translational applications.

8.2 Conclusion

This thesis presents a comprehensive analysis of the effects of state-of-the-art non-rigid registration algorithms on rs-fMRI outcomes, particularly focusing on algorithmic variability in neuroimaging accuracy and reliability. The NRAAF, developed and applied in this work, enables a systematic, quantitative evaluation of widely used algorithms—FSL, ANTs, DARTEL, and AFNI—shedding light on their respective impacts on Resting-State Network (RSN) detection, inter-subject variability, and functional connectivity mapping. This structured comparative approach contributes to improved algorithm selection strategies, essential for advancing neuroimaging methodology across research and clinical contexts.

Restatement of Research Aim and Objectives

The primary aim of this thesis was to systematically evaluate non-rigid registration algorithms and characterise their impact on neuroimaging outcomes, particularly in rs-fMRI analyses. Building on the foundational objectives outlined in Chapter 1, this research specifically targeted algorithmic variability, reliability, and context-specific performance. To address these aspects, the NRAAF framework was designed, developed, and applied as a structured approach for rigorous comparative evaluation. This work extends beyond individual assessments, aiming to inform best practices for algorithm selection in functional connectivity studies, which are sensitive to the variability inherent in registration methods. As Sotiras et al. [4] and others have highlighted, this variability can substantially impact neuroimaging outcomes. By addressing these challenges, this thesis provides a systematic and reproducible framework for enhancing methodological precision in computational neuroscience.

Summary of Key Findings

The findings of this research reveal that non-rigid registration algorithms variably affect neuroimaging accuracy, with significant impacts on metrics such as peak activation sensitivity, significant cluster consistency, and inter-subject alignment. Notably, ANTs and FSL exhibited heightened sensitivity, proving effective for exploratory studies where detecting subtle variations in functional connectivity is critical, whereas DARTEL's stability made it advantageous for reproducibility-focused contexts that demand consistent alignment across subjects. This systematic comparative analysis builds on recent advancements in the field [65, 216] by integrating diverse performance metrics to provide a nuanced, comprehensive understanding of algorithmic effects on rs-fMRI outcomes. This expanded framework contributes to a more precise selection of algorithms based on study objectives, addressing both the needs for sensitivity and reproducibility in neuroimaging research and clinical applications.

8.2.1 Contribution to Knowledge

This research makes several significant contributions to computational neuroscience and neuroimaging methodologies:

- **Development of NRAAF:** The NRAAF framework represents a novel, structured approach for the systematic comparison of non-rigid registration algorithms in rs-fMRI. This framework fills a critical gap in the standardisation of non-rigid registration evaluations, providing reproducible benchmarks and enabling consistent, reliable

comparisons across algorithms, particularly for functional connectivity analysis within RSNs.

- **Empirical Validation Using Large-Scale Data:** Validated NRAAF with a large dataset (n=815), offering robust, data-driven benchmarks and practical recommendations on algorithm selection. This empirical assessment informs best practices by identifying algorithm-specific strengths and weaknesses within functional connectivity analyses, supporting both scientific rigour and practical applications in clinical diagnostics.
- **Application of MVPA and SVM in Comparative Registration Analysis:** Introduced an innovative ML approach—utilising MVPA with SVM—to enable voxel-wise comparative analysis of registration algorithms' impact on functional connectivity. This methodological advancement enhances the sensitivity and specificity of algorithm comparisons, supporting nuanced interpretations in both research and clinical neuroimaging.
- **Comprehensive Characterisation of Non-Rigid Registration Algorithms:** Provided a detailed analysis of FSL, ANTs, DARTEL, and AFNI across key performance metrics, capturing their differential impacts on neuroimaging outcomes. This characterisation enables researchers and clinicians to select a suitable algorithm for specific study needs, promoting consistency and reliability in RSN analyses.

Limitations and Constraints

Despite these contributions, certain limitations must be acknowledged. First, the computational demands of NRAAF, particularly in MVPA and SVM applications, restrict its accessibility in settings with limited resources [82, 213]. Additionally, inter-subject variability in anatomical and physiological data introduces variability in algorithmic outputs, emphasising the need for adaptable, context-sensitive models within NRAAF. These constraints align with findings that noted similar challenges in handling inter-subject variability in functional connectivity studies [28, 122, 202]. Addressing these limitations in future work will enhance NRAAF's utility and scalability, supporting broader neuroimaging applications.

8.2.2 Implications for Research and Practice

This thesis establishes a foundational framework, the NRAAF, which offers a structured approach for evaluating non-rigid registration algorithms in rs-fMRI. Through NRAAF's

rigorous, multi-metric assessments, this work contributes a standardised, reproducible benchmark that addresses critical issues of algorithmic variability in neuroimaging accuracy and consistency. In research settings, the framework supports algorithm selection grounded in empirical data, enhancing methodological transparency and reliability [9, 16]. In clinical contexts, NRAAF's benchmarking framework provides clinicians with insights for selecting algorithms tailored to specific diagnostic needs, which is especially relevant for personalised diagnostics in neurodegenerative conditions, where precision in functional connectivity is essential [77, 7].

Furthermore, NRAAF's standardisation potential positions it as a valuable tool for multi-center studies and cross-institutional collaborations, where reproducibility across different sites and scanners is a primary challenge. By offering a unified framework, NRAAF enables researchers and clinicians to evaluate and share algorithm performance data across diverse settings, fostering consistency and enabling large-scale data aggregation [228, 2]. This standardisation supports interoperability in multi-center research, where consistent neuroimaging workflows are necessary for generating generalisable conclusions across populations [41].

Finally, NRAAF sets a foundation for further advancements in algorithm development, policy, and regulation within computational neuroscience. Emerging methods, particularly those integrating ML, may leverage NRAAF's benchmarks to validate novel algorithms before clinical use, promoting innovation aligned with established standards [3, 19]. The framework also has policy implications, as its systematic approach to reproducibility and accuracy could support data governance guidelines and regulatory practices for neuroimaging, especially in clinical diagnostics and AI-driven applications. As the field continues to evolve, NRAAF's contributions to neuroimaging methodology are positioned to have a lasting impact, shaping best practices and establishing benchmarks for both research and clinical applications [120, 229, 3, 2, 19].

References

[1] Lisa Gottesfeld Brown. "A survey of image registration techniques". In: *ACM computing surveys (CSUR)* 24.4 (1992), pp. 325–376.

[2] Barbara Zitova and Jan Flusser. "Image registration methods: a survey". In: *Image and vision computing* 21.11 (2003), pp. 977–1000.

[3] Geert Litjens et al. "A survey on deep learning in medical image analysis". In: *Medical image analysis* 42 (2017), pp. 60–88.

[4] Aristeidis Sotiras, Christos Davatzikos, and Nikos Paragios. "Deformable medical image registration: A survey". In: *IEEE transactions on medical imaging* 32.7 (2013), pp. 1153–1190.

[5] Xingxing Zhu et al. "Non-rigid multi-modal brain image registration based on two-stage generative adversarial nets". In: *Neurocomputing* 505 (2022), pp. 44–57.

[6] Guha Balakrishnan et al. "An unsupervised learning model for deformable medical image registration". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 9252–9260.

[7] John Ashburner and Karl J Friston. "Voxel-based morphometry—the methods". In: *Neuroimage* 11.6 (2000), pp. 805–821.

[8] Bruce Fischl. "FreeSurfer". In: *Neuroimage* 62.2 (2012), pp. 774–781.

[9] William R Crum et al. "Zen and the art of medical image registration: correspondence, homology, and quality". In: *NeuroImage* 20.3 (2003), pp. 1425–1437.

[10] Mohamed Kerkech, Adel Hafiane, and Raphael Canals. "Vine disease detection in UAV multispectral images using optimized image registration and deep learning segmentation approach". In: *Computers and Electronics in Agriculture* 174 (2020), p. 105446.

[11] Salvatore Manfreda et al. "On the use of unmanned aerial systems for environmental monitoring". In: *Remote sensing* 10.4 (2018), p. 641.

[12] Yuhao Bai et al. "Vision-based navigation and guidance for agricultural autonomous vehicles and robots: A review". In: *Computers and Electronics in Agriculture* 205 (2023), p. 107584.

[13] Samira Badrloo et al. "Image-based obstacle detection methods for the safe navigation of unmanned vehicles: A review". In: *Remote Sensing* 14.15 (2022), p. 3824.

[14] Harley HL Chan et al. "An integrated augmented reality surgical navigation platform using multi-modality imaging for guidance". In: *PLoS One* 16.4 (2021), e0250558.

[15] Scott D Uhlrich et al. "OpenCap: Human movement dynamics from smartphone videos". In: *PLoS computational biology* 19.10 (2023), e1011462.

[16] Derek L. G. Hill et al. “Medical Image Registration”. In: *Physics in Medicine and Biology* 46.3 (Mar. 2001), R1–R45. DOI: 10.1088/0031-9155/46/3/201.

[17] D. Rueckert et al. “Nonrigid registration using free-form deformations: Application to breast MR images”. In: *IEEE Transactions on Medical Imaging* 18.8 (1999), pp. 712–721.

[18] Jing Zou et al. “A review of deep learning-based deformable medical image registration”. In: *Frontiers in Oncology* 12 (2022), p. 1047215.

[19] Yabo Fu et al. “Deep learning in medical image registration: a review”. In: *Physics in Medicine & Biology* 65.20 (2020), 20TR01.

[20] Miriam E Weaverdyck, Matthew D Lieberman, and Carolyn Parkinson. “Tools of the Trade Multivoxel pattern analysis in fMRI: a practical introduction for social and affective neuroscientists”. In: *Social Cognitive and Affective Neuroscience* 15.4 (2020), pp. 487–509.

[21] Luca Steardo Jr et al. “Application of support vector machine on fMRI data as biomarkers in schizophrenia diagnosis: a systematic review”. In: *Frontiers in Psychiatry* 11 (2020), p. 588.

[22] Jacob S Prince et al. “Improving the accuracy of single-trial fMRI response estimates using GLMsingle”. In: *Elife* 11 (2022), e77599.

[23] Regina J Meszlényi et al. “Resting state fMRI functional connectivity analysis using dynamic time warping”. In: *Frontiers in neuroscience* 11 (2017), p. 75.

[24] A. R. Groves et al. “Benefits of multi-modal fusion analysis on a large-scale dataset: Life-span patterns of inter-subject variability in cortical morphometry and white matter microstructure”. In: *Neuroimage* (2020), p. 116727.

[25] Sanam Maknojia et al. “Resting State fMRI: Going Through the Motions”. In: *Frontiers in Neuroscience* 13 (2019). DOI: 10.3389/fnins.2019.00825. URL: https://www.frontiersin.org/articles/10.3389/fnins.2019.00825/full.

[26] Chiara Zucchella et al. “Non-invasive brain stimulation for gambling disorder: a systematic review”. In: *Frontiers in Neuroscience* 14 (2020), p. 536711.

[27] David Cole, Stephen Smith, and Christian Beckmann. “Advances and pitfalls in the analysis and interpretation of resting-state FMRI data”. In: *Frontiers in Systems Neuroscience* 4 (2010). ISSN: 1662-5137. DOI: 10.3389/fnsys.2010.00008. URL: https://www.frontiersin.org/articles/10.3389/fnsys.2010.00008.

[28] Fali Li et al. “Inter-subject P300 variability relates to the efficiency of brain networks reconfigured from resting-to task-state: evidence from a simultaneous event-related EEG-fMRI study”. In: *NeuroImage* 205 (2020), p. 116285.

[29] Elvis Dohmatob, Gael Varoquaux, and Bertrand Thirion. “Inter-subject registration of functional images: do we need anatomical images?” In: *Frontiers in neuroscience* 12 (2018), p. 305198.

[30] Oscar Esteban et al. “fMRIPrep: a robust preprocessing pipeline for functional MRI”. In: *Nature methods* 16.1 (2019), pp. 111–116.

[31] Siobhan Ewert et al. “Optimization and comparative evaluation of nonlinear deformation algorithms for atlas-based segmentation of DBS target nuclei”. In: *Neuroimage* 184 (2019), pp. 586–598.

[32] Shogo Kajimura, Daniel Margulies, and Jonathan Smallwood. “Frequency-specific brain network architecture in resting-state fMRI”. In: *Scientific Reports* 13.1 (2023), p. 2964.

[33] Jinjun Meng et al. “A full-spectrum registration method for Zhuhai-1 satellite hyperspectral imagery”. In: *Sensors* 20.21 (2020), p. 6298.

[34] Alice Lo Valvo et al. “A navigation and augmented reality system for visually impaired people”. In: *Sensors* 21.9 (2021), p. 3061.

[35] Yujiao Shi et al. “Where am i looking at? joint location and orientation estimation by cross-view matching”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020, pp. 4064–4072.

[36] Martin Urschler, Stefan Kluckner, and Horst Bischof. “A Framework for Comparison and Evaluation of Nonlinear Intra-Subject Image Registration Algorithms”. In: *Insight Journal–ISC/NA-MIC Workshop on Open Science at MICCAI*. 2007.

[37] K. J. Gorgolewski et al. “A high resolution 7-Tesla resting-state fMRI test-retest dataset with cognitive and physiological measures”. In: *Scientific Data* 2.140054 (2015).

[38] Adrian Andronache et al. “Non-rigid registration of multi-modal images using both mutual information and cross-correlation”. In: *Medical image analysis* 12.1 (2008), pp. 3–15.

[39] N. J. Tustison et al. “N4ITK: Improved N3 bias correction”. In: *IEEE Transactions on Medical Imaging* 29.6 (2010), pp. 1310–1320.

[40] H. J. Park et al. “Spatial normalization of diffusion tensor MRI using multiple channels”. In: *NeuroImage* 20.4 (2021), pp. 1995–2009.

[41] Gael Varoquaux and R. Cameron Craddock. “Learning and comparing functional connectomes across subjects”. In: *NeuroImage* 80 (2013), pp. 405–415. DOI: 10.1016/j.neuroimage.2013.04.007. URL: https://doi.org/10.1016/j.neuroimage.2013.04.007.

[42] Maria Joao Rosa and Ben Seymour. “Decoding the matrix: benefits and limitations of applying machine learning algorithms to pain neuroimaging”. In: *Pain®* 155.5 (2014), pp. 864–867.

[43] B. D. de Vos et al. “A deep learning framework for unsupervised affine and deformable image registration”. In: *Medical Image Analysis* 52 (2019), pp. 128–143. DOI: 10.1016/j.media.2018.11.010. URL: https://www.sciencedirect.com/science/article/pii/S1361841518300495.

[44] Xuenan Cao and Roozbeh Yousefzadeh. “Extrapolation and AI transparency: Why machine learning models should reveal when they make decisions beyond their training”. In: *Big Data & Society* 10.1 (2023), p. 20539517231169731.

[45] Jeffrey Yunfan Liu et al. “Real-time neural rasterization for large scenes”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2023, pp. 8416–8427.

[46] Jia-Chen Hua, Eun-jin Kim, and Fei He. “Information Geometry Theoretic Measures for Characterizing Neural Information Processing from Simulated EEG Signals”. In: *Entropy* 26.3 (2024).

[47] Dominik Klepl, Min Wu, and Fei He. "Graph neural network-based eeg classification: A survey". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* (2024).

[48] B. B. Avants et al. "Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain". In: *Medical Image Analysis* 12.1 (2008), pp. 26–41.

[49] Mingzhe Hu et al. "Reinforcement learning in medical image analysis: Concepts, applications, challenges, and future directions". In: *Journal of Applied Clinical Medical Physics* 24.2 (2023), e13898.

[50] T. Rohlfing and Jr. Maurer C. R. "Nonrigid image registration in shared-memory multiprocessor environments with application to brains, breasts, and bees". In: *IEEE Transactions on Information Technology in Biomedicine* 7.1 (2020), pp. 16–25.

[51] J. P. Thirion. "Image matching as a diffusion process: An analogy with Maxwell's demons". In: *Medical Image Analysis* 2.3 (1998), pp. 243–260.

[52] Sven Kabus et al. "B-spline registration of 3D images with Levenberg-Marquardt optimization". In: *Medical Imaging 2004: Image Processing*. Vol. 5370. SPIE. 2004, pp. 304–313.

[53] Zhe Min et al. "Non-rigid Medical Image Registration using Physics-informed Neural Networks". In: *International Conference on Information Processing in Medical Imaging*. Springer. 2023, pp. 601–613.

[54] Timotheus Roelofs. "Deep learning-based affine and deformable 3D medical image registration". In: *Master's thesis* (Sept. 2021). Available at https://aaltodoc.aalto.fi/server/api/core/bitstreams/73d855e7-1e9e-4e3d-be78-62514e2fb0c5/content.

[55] Marco Lorenzi et al. "LCC-Demons: a robust and accurate symmetric diffeomorphic registration algorithm". In: *NeuroImage* 81 (2013), pp. 470–483.

[56] Muhammad Hassan Jamal et al. "A Comparison of Re-Sampling Techniques for Detection of Multi-Step Attacks on Deep Learning Models". In: *IEEE Access* 11 (2023), pp. 127446–127457.

[57] Arno Klein et al. "Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration". In: *NeuroImage* 46.3 (2009), pp. 786–802. DOI: 10.1016/j.neuroimage.2008.12.037.

[58] Robert W Cox. "AFNI: software for analysis and visualization of functional magnetic resonance neuroimages". In: *Computers and Biomedical research* 29.3 (1996), pp. 162–173.

[59] R. W. Cox and D. R. Glen. "Nonlinear warping in AFNI". In: *Proc. 19th Annual Meeting of the Organization for Human Brain Mapping*. 2013.

[60] John Ashburner. "A fast diffeomorphic image registration algorithm". In: *NeuroImage* 38.1 (2007), pp. 95–113. DOI: 10.1016/j.neuroimage.2007.07.007.

[61] Mark Jenkinson et al. "FSL". In: *NeuroImage* 62.2 (2012), pp. 782–790. DOI: 10.1016/j.neuroimage.2011.09.015.

[62] J. L. R. Andersson, M. Jenkinson, S. Smith, et al. *Non-linear registration, aka Spatial normalisation FMRIB technical report TR07JA2*. Tech. rep. 1. FMRIB Analysis Group of the University of Oxford, 2007, e21.

[63] Daniel R Glen et al. “Beware (surprisingly common) left-right flips in your MRI data: an efficient and robust method to check MRI dataset consistency using AFNI”. In: *Frontiers in neuroinformatics* (2020), p. 505994.

[64] Hui Zhang et al. “High-dimensional spatial normalization of diffusion tensor images improves the detection of white matter differences: an example study using amyotrophic lateral sclerosis”. In: *IEEE transactions on medical imaging* 26.11 (2007), pp. 1585–1597.

[65] B. B. Avants et al. “A reproducible evaluation of ANTs similarity metric performance in brain image registration”. In: *NeuroImage* 54.3 (2020), pp. 2033–2044.

[66] Willem Benjamin Bruin et al. “Development and validation of a multimodal neuroimaging biomarker for electroconvulsive therapy outcome in depression: a multicenter machine learning analysis”. In: *Psychological Medicine* 54.3 (2024), pp. 495–506.

[67] Shaun Warrington et al. “A resource for development and comparison of multimodal brain 3 T MRI harmonisation approaches”. In: *Imaging Neuroscience* 1 (2023), pp. 1–27.

[68] Guha Balakrishnan et al. “VoxelMorph: A Learning Framework for Deformable Medical Image Registration”. In: *IEEE Transactions on Medical Imaging* 38.8 (2019), pp. 1788–1800.

[69] Maziar Raissi, Paris Perdikaris, and George E Karniadakis. “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations”. In: *Journal of Computational physics* 378 (2019), pp. 686–707.

[70] Patrik F Raudaschl et al. “Evaluation of segmentation methods on head and neck CT: auto-segmentation challenge 2015”. In: *Medical physics* 44.5 (2017), pp. 2020–2036.

[71] Ashish Sinha and Jose Dolz. “Multi-scale self-guided attention for medical image segmentation”. In: *IEEE journal of biomedical and health informatics* 25.1 (2020), pp. 121–130.

[72] Snigdha Mohanty and Sarada Prasad Dakua. “Toward computing cross-modality symmetric non-rigid medical image registration”. In: *IEEE Access* 10 (2022), pp. 24528–24539.

[73] Enrico Amico et al. “Mapping the functional connectome traits of levels of consciousness”. In: *Neuroimage* 148 (2017), pp. 201–211.

[74] Sheng Zhang and Chiang-Shan R Li. “Task-related, low-frequency task-residual, and resting state activity in the default mode network brain regions”. In: *Frontiers in psychology* 3 (2012), p. 26110.

[75] Marcus E Raichle et al. “A default mode of brain function”. In: *Proceedings of the national academy of sciences* 98.2 (2001), pp. 676–682.

[76] Randy L Buckner, Jessica R Andrews-Hanna, and Daniel L Schacter. “The brain’s default network: anatomy, function, and relevance to disease”. In: *Annals of the new York Academy of Sciences* 1124.1 (2008), pp. 1–38.

[77] Michael Greicius. “Resting-state functional connectivity in neuropsychiatric disorders”. In: *Current opinion in neurology* 21.4 (2008), pp. 424–430.

[78] Sirio Cocozza et al. “Modifications of resting state networks in spinocerebellar ataxia type 2”. In: *Movement Disorders* 30.10 (2015), pp. 1382–1390.

[79] Vincent Malotaux et al. “Default-Mode Network Connectivity Changes During the Progression Toward Alzheimer’s Dementia: A Longitudinal Functional Magnetic Resonance Imaging Study”. In: *Brain connectivity* 13.5 (2023), pp. 287–296.

[80] Elif Yildirim and Ezgi SONCU BÜYÜKİŞCAN. “Default Mode Network Connectivity in Alzheimer’s Disease.” In: *Turkish Journal of Psychiatry* 30.4 (2019).

[81] Mathew R Brier et al. “Loss of intranetwork and internetwork resting state functional connections with Alzheimer’s disease progression”. In: *Journal of Neuroscience* 32.26 (2012), pp. 8890–8899.

[82] Stephen M Smith et al. “Resting-state fMRI in the human connectome project”. In: *Neuroimage* 80 (2013), pp. 144–168.

[83] Amber M Leaver et al. “Modulation of intrinsic brain activity by electroconvulsive therapy in major depression”. In: *Biological psychiatry: cognitive neuroscience and neuroimaging* 1.1 (2016), pp. 77–86.

[84] Beatrice Tosti et al. “Integrated use of biofeedback and neurofeedback techniques in treating pathological conditions and improving performance: a narrative review”. In: *Frontiers in Neuroscience* 18 (2024), p. 1358481.

[85] John Eraifej et al. “Modulation of limbic resting-state networks by subthalamic nucleus deep brain stimulation”. In: *Network Neuroscience* 7.2 (2023), pp. 478–495.

[86] Serge ARB Rombouts et al. “Altered resting state networks in mild cognitive impairment and mild Alzheimer’s disease: an fMRI study”. In: *Human brain mapping* 26.4 (2005), pp. 231–239.

[87] Santosh A Helekar et al. “Functional brain network changes associated with maintenance of cognitive function in multiple sclerosis”. In: *Frontier in Human Neuroscience* 4 (2010), p. 219.

[88] Rahul S Desikan et al. “An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest”. In: *Neuroimage* 31.3 (2006), pp. 968–980.

[89] R. Mohanty, W. A. Sethares, V. A. Nair, et al. “Rethinking Measures of Functional Connectivity via Feature Extraction”. In: *Sci Rep* 10 (2020), p. 1298. DOI: 10.1038/s41598-020-57915-w.

[90] Lee Raymond Dice. “Measures of the amount of ecologic association between species”. In: *Ecology* 26.3 (1945), pp. 297–302.

[91] Martijn P Van Den Heuvel and Hilleke EA Pol. “Exploring the brain network: A review on resting-state fMRI functional connectivity”. In: *European Neuropsychopharmacology* 20.8 (2010), pp. 519–534. DOI: 10.1016/j.euroneuro.2010.03.008.

[92] M. F. Glasser et al. “The minimal preprocessing pipelines for the Human Connectome Project”. In: *NeuroImage* 80 (2013), pp. 105–124.

[93] J.B.Antoine Maintz and Max A. Viergever. “A survey of medical image registration”. In: *Medical Image Analysis* 2.1 (1998), pp. 1–36. ISSN: 1361-8415. DOI: https://doi.org/10.1016/S1361-8415(01)80026-8. URL: https://www.sciencedirect.com/science/article/pii/S1361841501800268.

[94] W. R. Crum, O. Camara, and D. L. G. Hill. "Generalized overlap measures for evaluation and validation in medical image analysis". In: *IEEE Transactions on Medical Imaging* 25.11 (2006), pp. 1451–1461.

[95] Feng Yang et al. "Two phase non-rigid multi-modal image registration using weber local descriptor-based similarity metrics and normalized mutual information". In: *Sensors* 13.6 (2013), pp. 7599–7617.

[96] Jorge Samper-González et al. "Reproducible evaluation of classification methods in Alzheimer's disease: Framework and application to MRI and PET data". In: *NeuroImage* 183 (2018), pp. 504–521.

[97] Vincent Calhoun. "Data-driven approaches for identifying links between brain structure and function in health and disease". In: *Dialogues in clinical neuroscience* 20.2 (2018), pp. 87–99.

[98] Stephen M Smith et al. "Advances in functional and structural MR image analysis and implementation as FSL". In: *Neuroimage* 23 (2004), S208–S219.

[99] Russell A Poldrack et al. "Guidelines for reporting an fMRI study". In: *Neuroimage* 40.2 (2008), pp. 409–414.

[100] Thomas E Nichols and Andrew P Holmes. "Nonparametric permutation tests for functional neuroimaging: a primer with examples". In: *Human brain mapping* 15.1 (2002), pp. 1–25.

[101] Stephen M Smith and Thomas E Nichols. "Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference". In: *Neuroimage* 44.1 (2009), pp. 83–98.

[102] L. Snoek et al. "The Amsterdam Open MRI Collection, a set of multimodal MRI datasets for individual difference analyses". In: *bioRxiv* (2020). DOI: 10.1101/2020.06.16.155317. URL: https://www.biorxiv.org/content/early/2020/12/05/2020.06.16.155317.

[103] Z. Xu et al. "F3RNet: full-resolution residual registration network for deformable image registration". In: *International Journal of Computer Assisted Radiology and Surgery* 16 (June 2021), pp. 923–932. DOI: 10.1007/s11548-021-02359-4.

[104] H. Jin et al. "Machine learning techniques for pulmonary nodule computer-aided diagnosis using CT images: A systematic review". In: *Biomedical Signal Processing and Control* 79 (2023), p. 104104. DOI: 10.1016/j.bspc.2022.104104. URL: https://www.sciencedirect.com/science/article/pii/S1746809422005638.

[105] G. Haskins, U. Kruger, and P. Yan. "Deep learning in medical image registration: a survey". In: *Machine Vision and Applications* 31 (2020), p. 8. DOI: 10.1007/s00138-020-01060-x. URL: https://doi.org/10.1007/s00138-020-01060-x.

[106] S. Gupta, P. Gupta, and V. S. Verma. "Study on anatomical and functional medical image registration methods". In: *Neurocomputing* 452 (2021), pp. 534–548. DOI: 10.1016/j.neucom.2020.08.085. URL: https://www.sciencedirect.com/science/article/pii/S0925231220319202.

[107] M. Hoffmann et al. "SynthMorph: Learning Contrast-Invariant Registration Without Acquired Images". In: *IEEE transactions on medical imaging* 41.3 (2022), pp. 543–558. DOI: 10.1109/TMI.2021.3116879.

[108] A. Sedghi et al. “Image registration: Maximum likelihood, minimum entropy and deep learning”. In: *Medical Image Analysis* 69 (2021), p. 101939. DOI: 10.1016/j.media.2020.101939. URL: https://www.sciencedirect.com/science/article/pii/S1361841520303030.

[109] S. Kelly et al. “Widespread white matter microstructural differences in schizophrenia across 4322 individuals: Results from the ENIGMA Schizophrenia DTI Working Group”. In: *Molecular Psychiatry* 23 (May 2018). DOI: 10.1038/mp.2017.170.

[110] Y. Huo et al. “Mapping Lifetime Brain Volumetry with Covariate-Adjusted Restricted Cubic Spline Regression from Cross-sectional Multi-site MRI”. In: *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*. Vol. 9900. 2018, pp. 81–88.

[111] Yuanjie Zheng et al. “SymReg-GAN: Symmetric Image Registration With Generative Adversarial Networks”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44.9 (2022), pp. 5631–5646.

[112] D. Mahapatra and Z. Ge. “Training data independent image registration using generative adversarial networks and domain adaptation”. In: *Pattern Recognition* 100 (2020), p. 107109. DOI: 10.1016/j.patcog.2019.107109. URL: https://www.sciencedirect.com/science/article/pii/S0031320319304108.

[113] J. Ker et al. “Deep Learning Applications in Medical Image Analysis”. In: *IEEE Access* 6 (2018), pp. 9375–9389. DOI: 10.1109/ACCESS.2017.2788044.

[114] B. Kim et al. “CycleMorph: Cycle consistent unsupervised deformable image registration”. In: *Medical Image Analysis* 71 (2021), p. 102036. DOI: 10.1016/j.media.2021.102036. URL: https://www.sciencedirect.com/science/article/pii/S1361841521000827.

[115] J. Fan et al. “Adversarial learning for mono- or multi-modal registration”. In: *Medical Image Analysis* 58 (2019), p. 101545. DOI: 10.1016/j.media.2019.101545. URL: https://www.sciencedirect.com/science/article/pii/S1361841519300805.

[116] Marwa Abderrahim, Abir Baâzaoui, and Walid Barhoumi. “Comparative study of relevant methods for mri/x brain image registration”. In: *The Impact of Digital Technologies on Public Health in Developed and Developing Countries: 18th International Conference, ICOST 2020, Hammamet, Tunisia, June 24–26, 2020, Proceedings 18*. Springer. 2020, pp. 338–347.

[117] Erin E O’Connor and Thomas A Zeffiro. “Why is clinical fMRI in a resting state?” In: *Frontiers in neurology* 10 (2019), p. 420.

[118] VA Kumar et al. “Recommended resting-state fMRI acquisition and preprocessing steps for preoperative mapping of language and motor and visual areas in adult and pediatric patients with brain tumors and epilepsy”. In: *American Journal of Neuroradiology* 45.2 (2024), pp. 139–148.

[119] Wonsik Jung et al. “EAG-RS: A Novel Explainability-guided ROI-Selection Framework for ASD Diagnosis via Inter-regional Relation Learning”. In: *IEEE Transactions on Medical Imaging* (2023).

[120] Yuxin Yao et al. “Fast and robust non-rigid registration using accelerated majorization-minimization”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2023).

[121] S. Abbasi et al. “Medical image registration using unsupervised deep neural network: A scoping literature review”. In: *Biomedical Signal Processing and Control* 73 (2022). DOI: 10.1016/j.bspc.2021.103444.

[122] Peter R Millar et al. “Evaluating cognitive relationships with resting-state and task-driven blood oxygen level-dependent variability”. In: *Journal of cognitive neuroscience* 33.2 (2021), pp. 279–302.

[123] Kenneth A Norman et al. “Beyond mind-reading: multi-voxel pattern analysis of fMRI data”. In: *Trends in cognitive sciences* 10.9 (2006), pp. 424–430.

[124] J Brendan Ritchie, David M Kaplan, and Colin Klein. “Decoding the brain: neural representation and the limits of multivariate pattern analysis in cognitive neuroscience”. In: *British Journal of Philosophy of Science* 70.2 (2019), pp. 581–607.

[125] Anders M. Dale, Bruce Fischl, and Martin I. Sereno. “Cortical surface-based analysis: I. Segmentation and surface reconstruction”. In: *NeuroImage* 9.2 (1999), pp. 179–194.

[126] M. Jenkinson et al. “Improved optimization for the robust and accurate linear registration and motion correction of brain images”. In: *NeuroImage* 17.2 (2002), pp. 825–841.

[127] Anders M. Dale and Martin I. Sereno. “Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach”. In: *Journal of Cognitive Neuroscience* 5.2 (1993), pp. 162–176.

[128] Nicholas J Tustison et al. “The ANTsX ecosystem for quantitative biological and medical imaging”. In: *Scientific reports* 11.1 (2021), p. 9068.

[129] Mark Jenkinson and Stephen Smith. “A global optimisation method for robust affine registration of brain images”. In: *Medical image analysis* 5.2 (2001), pp. 143–156.

[130] The MathWorks Inc. “MATLAB version R2023a”. In: *R2023a Release Highlights - MATLAB and Simulink* (2023). URL: https://uk.mathworks.com/products/new_products/release2023a.html.

[131] Z. S. Saad. *AFNI documentation - auto_tlrc*. Online. Accessed on: November 11, 2023. 2023. URL: https://afni.nimh.nih.gov/pub/dist/doc/htmldoc/programs/auto%5C_tlrc%5C_sphx.html.

[132] Mark W Woolrich et al. “Temporal Autocorrelation in Univariate Linear Modeling of FMRI Data”. In: *NeuroImage* 14.6 (2001), pp. 1370–1386.

[133] A. M. Michael, E. Evans, and G. J. Moore. “Influence of group on individual subject maps in SPM voxel based morphometry”. In: *Frontiers in Neuroscience* 10 (2016), p. 522.

[134] Valentina Doria et al. “Emergence of resting state networks in the preterm human brain”. In: *Proceedings of the National Academy of Sciences* 107.46 (2010), pp. 20015–20020.

[135] Morris A Jette and Tim Wickberg. "Architecture of the Slurm workload manager". In: *Workshop on Job Scheduling Strategies for Parallel Processing*. Springer. 2023, pp. 3–23.

[136] Robert Frost et al. "Markerless high-frequency prospective motion correction for neuroanatomical MRI". In: *Magnetic resonance in medicine* 82.1 (2019), pp. 126–144.

[137] Stephen M Smith. "Fast robust automated brain extraction". In: *Human brain mapping* 17.3 (2002), pp. 143–155.

[138] M. Holden et al. "Voxel similarity measures for 3-D serial MR brain image registration". In: *IEEE Transactions on Medical Imaging* 22.2 (2002), pp. 206–216.

[139] Ayca Kirimtat and Ondrej Krejcar. "GPU-Based Parallel Processing Techniques for Enhanced Brain Magnetic Resonance Imaging Analysis: A Review of Recent Advances". In: *Sensors* 24.5 (2024), p. 1591.

[140] A. L. Alexander et al. "Diffusion tensor imaging of the brain". In: *Neurotherapeutics* 4.3 (2007), pp. 316–329.

[141] Brian B Avants et al. "The Insight ToolKit image registration framework". In: *Frontiers in neuroinformatics* 8 (2014), p. 44.

[142] Mikael Brudfors et al. "Flexible Bayesian modelling for nonlinear image registration". In: *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part III 23*. Springer. 2020, pp. 253–263.

[143] Victoria Madge. *Evaluating Voxelmorph: a learning-based 3D non-linear registration algorithm, against the non-linear symmetric normalization technique from ANTs*. McGill University (Canada), 2020.

[144] C. Cortes and V. Vapnik. "Support-vector networks". In: *Machine Learning* 20.3 (1995), pp. 273–297.

[145] G. Varoquaux and B. Thirion. "How machine learning is shaping cognitive neuroimaging". In: *GigaScience* 3.1 (2014).

[146] M. W. Woolrich et al. "Multilevel linear modelling for FMRI group analysis using Bayesian inference". In: *NeuroImage* 21.4 (2004), pp. 1732–1747.

[147] Rosaria de Filippis et al. "Machine learning techniques in a structural and functional MRI diagnostic approach in schizophrenia: a systematic review". In: *Neuropsychiatric Disease and Treatment* 15 (2019), pp. 1605–1627.

[148] Tomas Hajek, Miloslav Kopecek, and Cyril Höschl. "Reduced hippocampal volumes in healthy carriers of brain-derived neurotrophic factor Val66Met polymorphism: meta-analysis". In: *World Journal of Biological Psychiatry* 16.8 (2015), pp. 556–566.

[149] Bilwaj Gaonkar and Christos Davatzikos. "Analytic estimation of statistical significance maps for support vector machine based multi-variate image analysis and classification". In: *NeuroImage* 78 (2013), pp. 270–283. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2013.03.066. URL: https://www.sciencedirect.com/science/article/pii/S1053811913003169.

[150] Rana Fayyaz Ahmad et al. “Visual brain activity patterns classification with simultaneous EEG-fMRI: A multimodal approach”. In: *Technology and Health Care* 25.3 (2017), pp. 471–485.

[151] Afshin Shoeibi et al. “Diagnosis of brain diseases in fusion of neuroimaging modalities using deep learning: A review”. In: *Information Fusion* 93 (2023), pp. 85–117.

[152] Meng Cao, Elizabeth Martin, and Xiaobo Li. “Machine learning in attention-deficit/hyperactivity disorder: new approaches toward understanding the neural mechanisms”. In: *Translational Psychiatry* 13.1 (2023), p. 236.

[153] Adriana N Nielsen et al. “Machine learning with neuroimaging: Evaluating its applications in psychiatry”. In: *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 5.8 (2020), pp. 791–798.

[154] Tara A. Niendam et al. “Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions”. In: *Cognitive, Affective, & Behavioral Neuroscience* 12.2 (2012), pp. 241–268.

[155] Amitai Shenhav, Matthew M. Botvinick, and Jonathan D. Cohen. “The expected value of control: An integrative theory of anterior cingulate cortex function”. In: *Neuron* 79.2 (2013), pp. 217–240.

[156] Richard A Armstrong. “When to use the Bonferroni correction”. In: *Ophthalmic and Physiological Optics* 34.5 (2014), pp. 502–508.

[157] Guy Brys, Mia Hubert, and Anja Struyf. “A robust measure of skewness”. In: *Journal of Computational and Graphical Statistics* 13.4 (2004), pp. 996–1017.

[158] Mahmoud Elbattah and Owen Molloy. “Data-Driven patient segmentation using K-Means clustering: The case of hip fracture care in Ireland”. In: *Proceedings of the Australasian Computer Science Week Multiconference*. 2017, pp. 1–8.

[159] R. A. Armstrong. “Statistical guidelines for the analysis of data obtained from one or both eyes”. In: *Ophthalmic and Physiological Optics* 40.1 (Jan. 2020), pp. 7–14. DOI: 10.1111/opo.12631.

[160] Andy Field. *Discovering statistics using IBM SPSS Statistics*. SAGE Publications Limited, 2024.

[161] Carl F Siebert and Darcy Clay Siebert. *Data analysis with small samples and non-normal data: Nonparametrics and other strategies*. Oxford University Press, 2017.

[162] Nicholas J Tustison and Brian B Avants. “Explicit B-spline regularization in diffeomorphic image registration”. In: *Frontiers in neuroinformatics* 7 (2013), p. 39.

[163] Tijl Grootswagers, Samuel G Wardle, and Thomas A Carlson. “Decoding dynamic brain patterns from evoked responses: a tutorial on multivariate pattern analysis applied to time series neuroimaging data”. In: *Journal of Cognitive Neuroscience* 29.4 (2017), pp. 677–697.

[164] Mei-Yi Chu et al. “Adaptive grid generation based non-rigid image registration using mutual information for breast MRI”. In: *Journal of signal processing systems* 54 (2009), pp. 45–63.

[165] D. Louis Collins and Alan C. Evans. “Animal: validation and applications of nonlinear registration-based segmentation”. In: *International journal of pattern recognition and artificial intelligence* 11.08 (1997), pp. 1271–1294.

[166] Kelly H. Zou et al. “Statistical validation of image segmentation quality based on a spatial overlap index”. In: *Academic radiology* 11.2 (2004), pp. 178–189.

[167] Martin Svejda et al. “NRAAF: A Framework for Comparative Analysis of fMRI Registration Algorithms and Their Impact on Resting-State Neuroimaging Accuracy”. In: *IEEE Access* 12 (2024), pp. 47915–47941. DOI: 10.1109/ACCESS.2024.3382819.

[168] Brian B Avants, Nick Tustison, Gang Song, et al. “Advanced normalization tools (ANTS)”. In: *Insight j* 2.365 (2009), pp. 1–35.

[169] Dongyang Zhang et al. “Intrinsic functional relations between human cerebral cortex and thalamus”. In: *Journal of neurophysiology* 100.4 (2008), pp. 1740–1748.

[170] Jonathan D Power et al. “Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion”. In: *Neuroimage* 59.3 (2012), pp. 2142–2154.

[171] Russell A Poldrack, Jeanette A Mumford, and Thomas E Nichols. *Handbook of functional MRI data analysis*. Cambridge University Press, 2024.

[172] Bryan Guillaume et al. “Fast and accurate modelling of longitudinal and repeated measures neuroimaging data”. In: *NeuroImage* 94 (2014), pp. 287–302.

[173] Henry B Mann and Donald R Whitney. “On a test of whether one of two random variables is stochastically larger than the other”. In: *The annals of mathematical statistics* (1947), pp. 50–60.

[174] Frank Wilcoxon. “Individual comparisons by ranking methods”. In: *Breakthroughs in statistics: Methodology and distribution*. Springer, 1992, pp. 196–202.

[175] Emily S Finn et al. “Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity”. In: *Nature Neuroscience* (2015). DOI: 10.1038/nn.4135.

[176] Michael C Corballis. “Evolution of cerebral asymmetry”. In: *Progress in Brain Research* 250 (2019), pp. 153–178.

[177] Bin Wang et al. “Brain asymmetry: a novel perspective on hemispheric network”. In: *Brain Science Advances* 9.2 (2023), pp. 56–77.

[178] A. Klein and S. S. Ghosh. “Evaluation of volume-based and surface-based brain image registration methods”. In: *NeuroImage* 51.1 (2012), pp. 214–220.

[179] Gaël Varoquaux et al. “Atlases of cognition with large-scale human brain mapping”. In: *PLoS computational biology* 14.11 (2018), e1006565.

[180] H. Zhu et al. “Functional magnetic resonance imaging progressive deformable registration based on a cascaded convolutional neural network”. In: *Quantitative Imaging in Medicine and Surgery* 10.3 (2020), pp. 605–616. URL: http://qims.amegroups.org/article/view/35888.

[181] Arthur W Toga and Paul M Thompson. “Mapping brain asymmetry”. In: *Nature Reviews Neuroscience* 4.1 (2003), pp. 37–48.

[182] Robert W Cox et al. “FMRI clustering in AFNI: false-positive rates redux”. In: *Brain connectivity* 7.3 (2017), pp. 152–171.

[183] Nikos Makris et al. “Decreased volume of left and total anterior insular lobule in schizophrenia”. In: *Schizophrenia research* 83.2-3 (2006), pp. 155–171.

[184] Jean A Frazier et al. “Structural brain magnetic resonance imaging of limbic and thalamic volumes in pediatric bipolar disorder”. In: *American Journal of Psychiatry* 162.7 (2005), pp. 1256–1265.

[185] Jill M Goldstein et al. “Hypothalamic abnormalities in schizophrenia: sex effects and genetic vulnerability”. In: *Biological psychiatry* 61.8 (2007), pp. 935–945.

[186] P.-Y. Hervé et al. “Handedness and cerebral anatomical asymmetries in young adult males”. In: *Neuroimage* 29.4 (2006), pp. 1066–1079.

[187] Hongtu Zhu, Tengfei Li, and Bingxin Zhao. “Statistical Learning Methods for Neuroimaging Data Analysis with Applications”. In: *Annual Review of Biomedical Data Science* 6 (2023), pp. 73–104. DOI: 10.1146/annurev-biodatasci-020722-100353.

[188] Jolynn Pek, Olivia Wong, and Anthony C M Wong. “How to Address Non-normality: A Taxonomy of Approaches, Reviewed, and Illustrated”. In: *Frontiers in Psychology* 9 (2018), p. 2104. DOI: 10.3389/fpsyg.2018.02104.

[189] Alexander F Schmidt and Chris Finan. “Linear regression and the normality assumption”. In: *Journal of Clinical Epidemiology* 98 (2018), pp. 146–151. DOI: 10.1016/j.jclinepi.2017.12.008.

[190] J. L. Vincent et al. “Intrinsic functional architecture in the anaesthetized monkey brain”. In: *Nature* 447.7140 (2007), pp. 83–86. DOI: 10.1038/nature05758.

[191] Zhengdong Wang et al. “Distribution-guided network thresholding for functional connectivity analysis in fMRI-based brain disorder identification”. In: *IEEE journal of biomedical and health informatics* 26.4 (2021), pp. 1602–1613.

[192] Shuo Wang et al. “Hemispheric Asymmetry of Human Brain Anatomical Network Revealed by Diffusion Tensor Tractography”. In: *PLoS ONE* 10.3 (2015), e0117352. DOI: 10.1371/journal.pone.0117352.

[193] Shaoming Wang et al. “Functional parcellation of the default mode network: a large-scale meta-analysis”. In: *Scientific reports* 10.1 (2020), p. 16096.

[194] E. Tagliazucchi and H. Laufs. “Decoding wakefulness levels from typical fMRI resting-state data reveals reliable drifts between wakefulness and sleep”. In: *Neuron* 82.3 (2014), pp. 695–708. DOI: 10.1016/j.neuron.2014.03.020.

[195] Paul Zhutovsky et al. “Individual prediction of psychotherapy outcome in posttraumatic stress disorder using neuroimaging data”. In: *Translational Psychiatry* 9.1 (2019), p. 326.

[196] C-W Woo et al. “Building better biomarkers: brain models in translational neuroimaging”. In: *Nature Neuroscience* 20.3 (2017), pp. 365–377.

[197] Robin AA Ince et al. “A statistical framework for neuroimaging data analysis based on mutual information estimated via a gaussian copula”. In: *Human brain mapping* 38.3 (2017), pp. 1541–1573.

[198] Changcheng Sun et al. “Mutual information-based brain network analysis in post-stroke patients with different levels of depression”. In: *Frontiers in human neuroscience* 12 (2018), p. 285.

[199] Di Wang et al. “Deep neural network heatmaps capture Alzheimer’s disease patterns reported in a large meta-analysis of neuroimaging studies”. In: *Neuroimage* 269 (2023), p. 119929. DOI: 10.1016/j.neuroimage.2023.119929. URL: https://doi.org/10.1016/j.neuroimage.2023.119929.

[200] Dominik Müller and Frank Kramer. “MIScnn: a framework for medical image segmentation with convolutional neural networks and deep learning”. In: *BMC medical imaging* 21 (2021), pp. 1–11.

[201] Ru Kong et al. “Individual-specific areal-level parcellations improve functional connectivity prediction of behavior”. In: *Cerebral cortex* 31.10 (2021), pp. 4477–4500.

[202] Rotem Botvinik-Nezer et al. “Variability in the analysis of a single neuroimaging dataset by many teams”. In: *Nature* 582.7810 (2020), pp. 84–88. DOI: 10.1038/s41586-020-2314-9. URL: https://doi.org/10.1038/s41586-020-2314-9.

[203] Pierre Bellec et al. “The Neuro Bureau ADHD-200 Preprocessed repository”. In: *NeuroImage* 144.Pt B (2017), pp. 275–286. DOI: 10.1016/j.neuroimage.2016.06.034. URL: https://doi.org/10.1016/j.neuroimage.2016.06.034.

[204] Martijn D Steenwijk et al. “Accurate white matter lesion segmentation by k nearest neighbor classification with tissue type priors (kNN-TTPs)”. In: *NeuroImage: Clinical* 3 (2013), pp. 462–469.

[205] Pierre Bellec et al. “Multi-level bootstrap analysis of stable clusters in resting-state fMRI”. In: *Neuroimage* 51.3 (2010), pp. 1126–1139.

[206] Alexander Schaefer et al. “Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI”. In: *Cerebral Cortex* 28.9 (2018), pp. 3095–3114. DOI: 10.1093/cercor/bhx179. URL: https://doi.org/10.1093/cercor/bhx179.

[207] Douglas Steinley. “K-means clustering: a half-century synthesis”. In: *British Journal of Mathematical and Statistical Psychology* 59.1 (2006), pp. 1–34.

[208] Yanlu Wang and Tie-Qiang Li. “Analysis of whole-brain resting-state fMRI data using hierarchical clustering approach”. In: *PloS one* 8.10 (2013), e76315.

[209] Sidong Liu et al. “Multimodal neuroimaging computing: the workflows, methods, and platforms”. In: *Brain informatics* 2 (2015), pp. 181–195.

[210] Manan Binth Taj Noor et al. “Application of deep learning in detecting neurological disorders from magnetic resonance images: a survey on the detection of Alzheimer’s disease, Parkinson’s disease and schizophrenia”. In: *Brain informatics* 7 (2020), pp. 1–21.

[211] Frederik J Lange et al. “MMORF—FSL’s MultiMOdal Registration Framework”. In: *Imaging Neuroscience* (2024).

[212] Miriam Friedel et al. “Pydpiper: a flexible toolkit for constructing novel registration pipelines”. In: *Frontiers in neuroinformatics* 8 (2014), p. 67.

[213] Yao Su et al. “Ernet: Unsupervised collective extraction and registration in neuroimaging data”. In: *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 2022, pp. 1666–1675.

[214] Shihui Ying et al. “Nonlinear image registration with bidirectional metric and reciprocal regularization”. In: *PLoS One* 12.2 (2017), e0172432.

[215] Zhao Qing et al. “The impact of spatial normalization strategies on the temporal features of the resting-state functional MRI: Spatial normalization before rs-fMRI features calculation may reduce the reliability”. In: *Frontiers in neuroscience* 13 (2019), p. 1249.

[216] Pankaj Kumar Keserwani, Suman Das, and Nairita Sarkar. “A comparative study: prediction of parkinson’s disease using machine learning, deep learning and nature inspired algorithm”. In: *Multimedia Tools and Applications* (2024), pp. 1–49.

[217] Julio M Duarte-Carvajalino et al. “A framework for linear and non-linear registration of diffusion-weighted MRIs using angular interpolation”. In: *Frontiers in neuroscience* 7 (2013), p. 41.

[218] Shan Zhong et al. “Comparison of image registration techniques in functional ultrasound imaging”. In: *bioRxiv* (2023), pp. 2023–09.

[219] Aljaz Bozic et al. “Neural deformation graphs for globally-consistent non-rigid reconstruction”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021, pp. 1450–1459.

[220] J. Ashburner. “Fast diffeomorphic image registration: Theory and practice”. In: *NeuroImage* 125 (2016), pp. 273–285.

[221] Zixuan Chen et al. “Self-learning based medical image representation for rigid real-time and multimodal slice-to-volume registration”. In: *Information Sciences* 541 (2020), pp. 502–515.

[222] Changshuo Liu and Nicolas Boumal. “Simple algorithms for optimization on Riemannian manifolds with constraints”. In: *Applied Mathematics & Optimization* 82.3 (2020), pp. 949–981.

[223] Ana I Torre-Bastida et al. “Bio-inspired computation for big data fusion, storage, processing, learning and visualization: state of the art and future directions”. In: *Neural Computing and Applications* (2021), pp. 1–31.

[224] Andriy Myronenko and Xubo Song. “Point set registration: Coherent point drift”. In: *IEEE transactions on pattern analysis and machine intelligence* 32.12 (2010), pp. 2262–2275.

[225] Mengfang Li et al. “Medical image analysis using deep learning algorithms”. In: *Frontiers in Public Health* 11 (2023), p. 1273253.

[226] Lifa Zhu et al. “Point cloud registration using representative overlapping points”. In: *arXiv preprint arXiv:2107.02583* (2021).

[227] Robin Scheibler and Nobutaka Ono. “MM algorithms for joint independent subspace analysis with application to blind single and multi-source extraction”. In: *arXiv preprint arXiv:2004.03926* (2020).

[228] Martin Norgaard et al. “PET-BIDS, an extension to the brain imaging data structure for positron emission tomography”. In: *Scientific data* 9.1 (2022), p. 65.

[229] Zhen Zhao et al. “Conventional machine learning and deep learning in Alzheimer’s disease diagnosis using neuroimaging: A review”. In: *Frontiers in computational neuroscience* 17 (2023), p. 1038636.

Appendix A

Modality Details

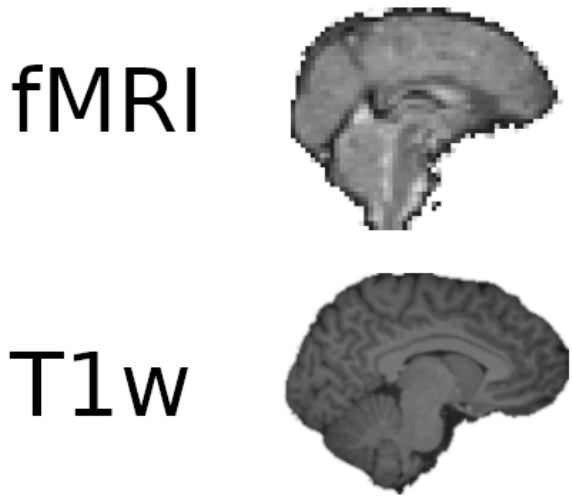

Fig. A.1 Modalities collected from the dataset chosen for this study. The Amsterdam Open MRI Collection [102] is a comprehensive collection of multi-modal MRI datasets designed for individual difference analysis. For this research, the *T1w* and *fMRI* modalities were selected.

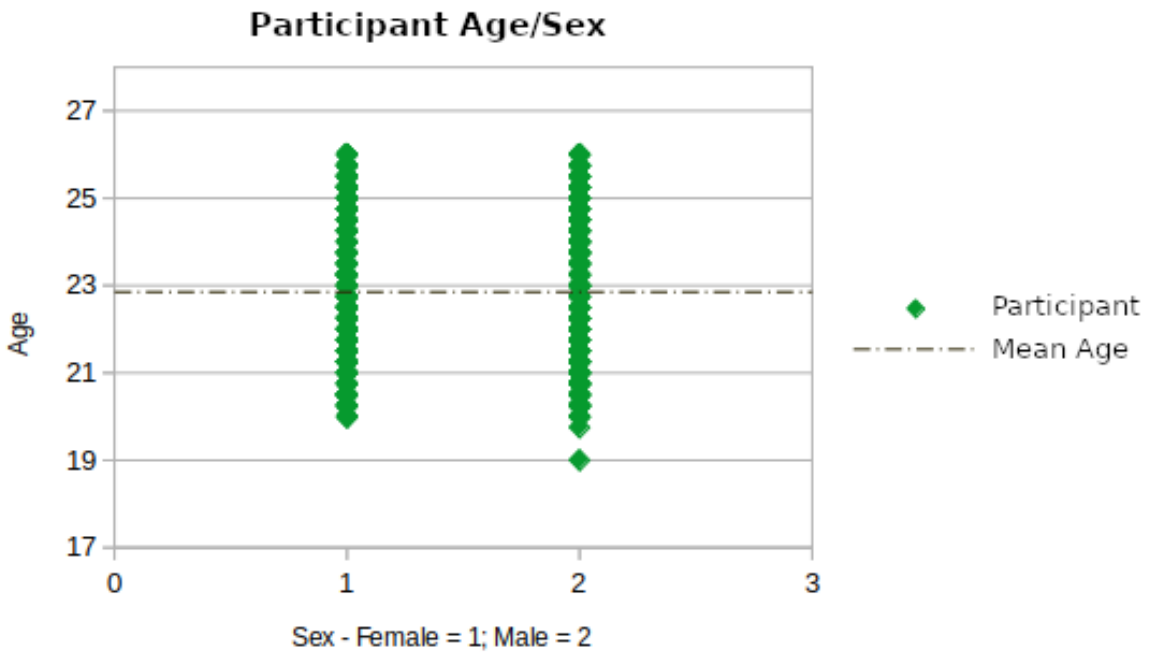

Fig. A.2 Demographic breakdown of the study dataset consisting of N = 815 participants, including 416 females and 399 males. The age range of the participants is 19 to 26 years old, with an average age of approximately 22.9 years. This dataset was selected to address gender disparities in medical data accessibility, ensuring a balanced representation of both genders in the analysis.

ID1000			
Property	**T1-weighted MRI**	**Property**	**Functional (BOLD) MRI**
Scan technique	3D MPRAGE	Scan technique	GE-EPI
Number of signals (repetitions)	1	FOV (RL / AP / FH)	138×192×192
FOV (RL / AP / FH; mm.)	160×256×256	Voxel size (mm.)	3×3×3
Voxel size (mm.)	1×1×1	Matrix size	64x64
TR / TE (millisec.)	8.1 / 3.7	Nr. of slices	40
Water-fat shift (pix.)	2.268	Slice gap (mm.)	0.3
Bandwidth (Hz./pix.)	191.5	TR / TE (ms.)	2200 / 28
Flip angle (deg.)	8	Water-fat shift (pix.)	12.481
Phase accell. factor (SENSE)	1.5 (RL)	Bandwidth (Hz/Pix)	34.6
Acquisition direction	Sagittal	Flip angle (deg.)	90
Duration	5 min 58 sec	Phase accell. factor (SENSE)	0
		Phase encoding direction	P >> A
		Slice encoding direction	L >> R
		Nr. of dummy scans	2
		Dynamic stabilisation	none
		Duration	10 min 38 sec

Table A.1 Details of the files from the scanner. All images in this study were obtained from the "Intera" version of the Philips 3T scanner (Philips, Best, the Netherlands) [102].

Appendix B

Mutual Information Results

Full size images from the aggregated figures can be made available on request.

Table B.1 Pairwise mutual information values between fMRI registration algorithm SVM weight maps. The values range from 0 to 1, with 1 indicating identical results and values closer to 0 indicating less similarity. For better readability, the DARTEL algorithm is abbreviated as DAR.

Algorithm Pairs	**FSL vs ANTS**	**ANTS vs AFNI**	**FSL vs AFNI**	**ANTS vs DAR**	**FSL vs DAR**	**AFNI vs DAR**
FSL vs ANTS	*1.000*	0.8633	0.94043	0.83779	0.8062	0.8866
ANTS vs AFNI	0.8633	*1.000*	0.88171	0.85829	0.91124	0.80666
FSL vs AFNI	0.94043	0.88171	*1.000*	0.81566	0.80892	0.80769
ANTS vs DAR	0.83779	0.85829	0.81566	*1.000*	0.9461	0.86119
FSL vs DAR	0.8062	0.91124	0.80892	0.9461	*1.000*	0.81571
AFNI vs DAR	0.8866	0.80666	0.80769	0.86119	0.81571	*1.000*