IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 10 February 2024, accepted 25 March 2024, date of publication 28 March 2024, date of current version 5 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3382819

RESEARCH ARTICLE

NRAAF: A Framework for Comparative Analysis of fMRI Registration Algorithms and Their Impact on Resting-State Neuroimaging Accuracy

MARTIN SVEJDA[1], NOUH SABRI ELMITWALLY[1,2], A. TAUFIQ ASYHARI[3], AND ROGER TAIT[1]
[1]Faculty of Computing, Engineering and the Built Environment, Birmingham City University, B5 5JU Birmingham, U.K.
[2]Faculty of Computers and Artificial Intelligence, Cairo University, Giza 12613, Egypt
[3]Department of Data Science, Monash University, Indonesia Campus, Tangerang 15345, Indonesia

Corresponding author: Martin Svejda (martin.svejda@bcu.ac.uk)

ABSTRACT The rapid evolution of neuroimaging techniques underscores the necessity for robust medical image registration algorithms, essential for the precise analysis of resting-state networks. This study introduces a comprehensive modular evaluation framework, designed to assess and compare the differences of four state-of-the-art algorithms in the field: FSL, ANTs, DARTEL, and AFNI. Our framework highlights the critical importance of algorithm selection in neuroimaging, addressing the unique challenges and strengths each algorithm presents in processing complex brain imaging data. Our rigorous evaluation delves into the algorithms' differences, with a focus on spatial localisation accuracy and the fidelity of resting-state network identification. The comparative analysis uncovers distinct advantages and limitations inherent to each algorithm, illuminating how specific characteristics can shape neuroimaging study outcomes. For instance, we reveal FSL's robustness in handling diverse datasets, ANTs' precision in spatial normalisation, DARTEL's suitability for large-scale studies, and AFNI's adaptability in functional and structural image analysis. The findings highlight the nuanced considerations necessary in choosing the right registration algorithm for neuroimaging data, advocating for a bespoke approach based on the unique requirements of each study. This detailed analysis advances the field, guiding researchers towards more informed algorithm selection and application, thus aiming to improve the accuracy and reliability of neuroimaging outcomes. Presenting a clear, comprehensive overview of each algorithm within our novel framework, the study addresses the needs of the neuroimaging community and paves the way for future advancements in medical image registration.

INDEX TERMS NRAAF, algorithm evaluation in neuroimaging, resting-state fMRI analysis, non-rigid image registration, medical imaging, multivariate pattern analysis.

I. INTRODUCTION

In the rapidly evolving field of medical imaging, the process of image registration plays a crucial role in various critical applications, including diagnosis, prognosis, treatment, and follow-up [1]. Image registration, also known as image fusion or matching is crucial for mapping the anatomy, physiology, functions, and connectivity of brains both in individuals and groups [2]. This technique has witnessed the development of various alignment methodologies, each aiming to enhance the accuracy and comprehensiveness of image analysis [3].

Particularly, this study focuses on Default Mode Networks (DMN) - critical brain networks that are most active during rest and are involved in self-referential and introspective activities. Their relevance spans across various neuropsychological disorders and cognitive functions, serving as a focal point in understanding the neural basis of these conditions and their potential therapeutic targets [4], [5]. DMNs are a subset of Resting-State Networks (RSNs). The persistence of intrinsic activity, even during altered states like sleep [6]

The associate editor coordinating the review of this manuscript and approving it for publication was Qilian Liang.

and certain sedation types [7], [8], makes RSNs a subject of immense interest and investigation in neuroimaging.

In this field, functional magnetic resonance imaging (fMRI) emerges as a sensitive tool, detecting changes in regional blood perfusion, volume, or oxygenation corresponding to neuronal activity. Traditional image registration approaches in fMRI, involving iterative optimisation procedures, are foundational yet demand enhancements to accommodate the complexities of modern medical imaging challenges [9]. These challenges include multiple modalities, image quality, distortion, organ motion, and noise. Specifically, in functional MRI paradigms, motion correction becomes crucial to mitigate subject movement effects, which can introduce artefacts or distortions requiring correction [10].

Given these intricacies, our study aims to delve into the algorithmic differences in resting-state networks between different demographic groups. For this purpose, we utilise data from the public repository Open Neuro (https://openneuro.org), ensuring data safety, confidentiality, and increased reproducibility of our findings with a large dataset (N=815). This dataset facilitates the evaluation of visible activation networks, which are crucial for applying and understanding various metrics in RSN studies.

This contribution introduces the Non-Rigid Registration Algorithm Analysis Framework (NRAAF), a modular evaluation framework designed to provide an objective comparison of non-rigid registration algorithms, specifically focusing on the validation and comparative analysis of four prominent fMRI registration algorithms: FMRIB Software Library (FSL), Advanced Normalisation Tools (ANTs), Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL), and Analysis of Functional Neuroimages (AFNI) (summarised in Table 1). These algorithms are rigorously evaluated within the context of the control network, offering insights into their impact on the spatial localisation accuracy and reliability of identified networks.

In the following sections, we describe our unique approach to image pre-processing, including a bespoke evaluation framework for analysing the Blood Oxygen Level Dependent (BOLD) signal in grey matter (GM) functional differences within default mode networks (simplified overview in Fig. 1). This represents a pioneering study in its field. We then present our findings and conclusions, highlighting the efficacy and adaptability of our modular evaluation framework in addressing the nuanced challenges of medical image registration.

II. RELATED WORK

We believe that studying algorithm evaluation is more significant than developing a new one, considering the current state of the art in nonlinear intra-subject, intra-modality registration. There are very few papers that are solely devoted to evaluation in the literature, and the majority of publications on algorithms do not focus on a comprehensive evaluation [11]. This section presents challenges for each evaluated algorithm.

A. ANTS (ADVANCED NORMALISATION TOOLS)

Functional imaging analysis using Advanced Normalisation Tools (ANTs) presents a robust framework for neuroimaging; however, it encounters several challenges that impact its application.

One of the challenges lies in voxel-wise hypothesis testing, as discussed by Rizzo et al. [12]. Voxel-wise analysis can lead to inflated false positives, necessitating rigorous statistical correction methods to address this issue. Additionally, artefact removal in fMRI data is critical for accurate analysis. Tohka et al. [13] present an automatic independent component labelling approach to tackle this challenge, enhancing the reliability of functional imaging results.

Another challenge relates to optimising template selection for different populations and datasets. Avants et al. [14] discuss the optimal template effect and strategies for template selection, particularly important for elderly and neurodegenerative brain studies. Furthermore, ANTs relies on well-annotated data for multi-atlas segmentation, and Avants et al. [14] explore methods and annotation strategies to improve multi-atlas segmentation, addressing the challenge of limited training data.

Additionally, the robustness of ANTs in the presence of noise and variations in MRI data quality is a persistent challenge. This challenge is crucial in pediatric neuroimaging, as highlighted by Tustison et al. [15], who emphasize the importance of ANTs for addressing these issues. Finally, Klein and Ghosh [16] evaluate the performance of ANTs in volume-based and surface-based brain image registration, illustrating challenges in the registration of different image modalities and the need for accurate alignment in multi-modal studies.

B. DARTEL (DIFFEOMORPHIC ANATOMICAL REGISTRATION THROUGH EXPONENTIATED LIE ALGEBRA)

DARTEL, a valuable tool for structural-functional analyses, confronts several challenges in functional imaging that impact its application.

One key challenge is the accurate modelling of longitudinal structural MRI data. Ashburner and Ridgway [17] discuss symmetric diffeomorphic modelling techniques for longitudinal data, addressing the challenge of modelling changes over time effectively. Additionally, Manogaran and Leung [18] introduce feature-based group independent component analysis as an approach to improving high-dimensional brain structural and functional imaging analysis. This approach helps tackle the challenge of handling complex, high-dimensional data.

Another challenge relates to the fusion of information from different imaging modalities. Gaser et al. [19] present BrainAGE as a method for predicting the conversion to

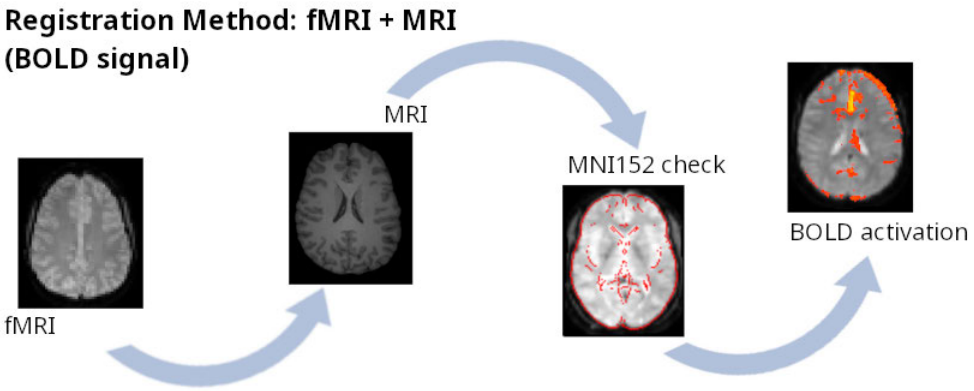


FIGURE 1. Abstract visualisation of the registration process. This includes functional and structural image registration. The functional image is first registered on the structural image of the same subject. Then the resulting image is normalised onto a standard template and fitted to a general linear model. This produces activation clusters which are subject to multivariate pattern analysis.

Alzheimer's disease, emphasizing the integration of multi-modal data. Furthermore, Groves et al. [20] explore the benefits of multi-modal fusion analysis, focusing on cortical morphometry and white matter microstructure, addressing the challenge of combining information from diverse sources effectively.

Lastly, Tustison et al. [21] highlight the importance of measurement-based performance evaluation of image segmentation algorithms in DARTEL, emphasizing the challenge of accurate segmentation and the need for rigorous evaluation techniques.

C. AFNI (ANALYSIS OF FUNCTIONAL NEUROIMAGES)

Analysis of Functional NeuroImages (AFNI) is a powerful tool for fMRI data analysis, but it faces several challenges that influence its utility in research.

One challenge concerns the control of false positives in cluster-based fMRI analyses. Cox et al. [22] address this issue and provide insights into reducing false-positive rates in AFNI, improving the reliability of fMRI results. Chen et al. [23] and Chen et al. [25] discuss the challenges related to t-tests, sidedness choice, and the artificially doubled false positive rates, highlighting the need for statistical rigour in AFNI-based analyses.

Another challenge is the detection of subtle differences in functional connectivity patterns. Bhaumik et al. [24] present multivariate pattern analysis strategies to improve the detection of remitted major depressive disorder using resting-state functional connectivity, addressing the challenge of identifying nuanced functional alterations.

The stability of individual differences in regional blood oxygen level-dependent (BOLD) signal baseline during task and rest is also a concern. Chen et al. [25] investigate this challenge and emphasise the importance of considering BOLD signal baseline stability in fMRI studies.

Lastly, the issue of robust and accurate linear registration and motion correction in brain images is a common challenge in fMRI pre-processing. Jenkinson et al. [26] discuss improved optimisation techniques for addressing this challenge, contributing to more reliable pre-processing in AFNI-based analyses.

D. FSL (FMRIB SOFTWARE LIBRARY)

FSL is versatile but faces challenges in functional imaging that impact its use in research.

One significant challenge is the accurate correction of image noise and artefacts. Salimi-Khorshidi et al. [27] propose automatic denoising techniques in FSL, addressing the challenge of optimising image quality for subsequent analysis. Andersson and Sotiropoulos [28] tackle the issue of off-resonance effects and subject movement in diffusion MRI, emphasising the need for effective correction strategies.

using MATLAB R2023a [65].

Another challenge is the integration of information from different imaging modalities. Glasser et al. [29] discuss the challenges and benefits of multi-modal data fusion in large-scale neuroimaging projects like the Human Connectome Project, highlighting the importance of combining data sources effectively.

Additionally, Woolrich et al. [30] introduce multilevel linear modelling for group analysis in FSL, providing a solution for modelling the variability present in complex neuroimaging datasets, which is a common challenge.

Finally, Gaser et al. [19] explore the neuroimaging community's approach to addressing the challenge of reproducibility and reliability in fMRI studies, highlighting the need for standardised and transparent analysis pipelines.

As this review of related work indicates, current neuroimaging algorithms, including ANTs, DARTEL, AFNI, and FSL, offer valuable insights but also face significant challenges, particularly in terms of adaptability and precision. This underscores the need for a more integrative and flexible approach. The following section introduces our proposed framework, designed to overcome these limitations by leveraging a comprehensive, multidimensional evaluation

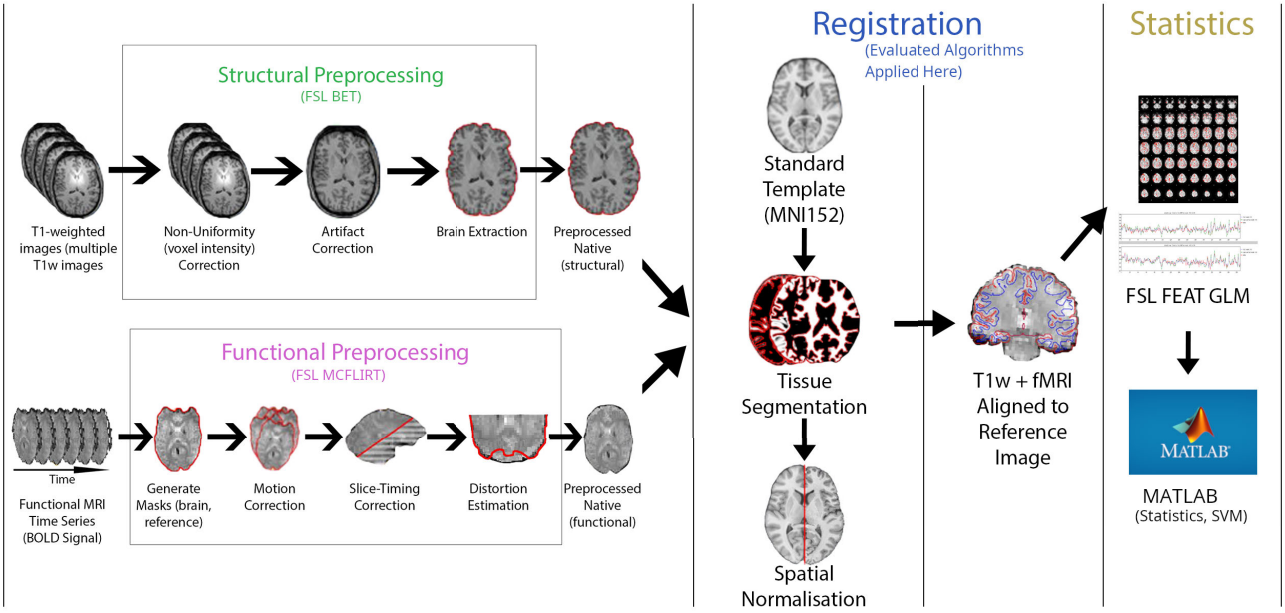


FIGURE 2. The functional and structural pre-processing stages are shown in this image of the evaluation framework, which is followed by the registration of pre-processed images to the MNI152 standard template. We then generated clusters by using FSL FEAT [67] to fit the registered images into a general linear model. These clusters were then input to Support Vector Machines (SVM) to calculate the decision boundary between the algorithm's contribution to the activation intensity. To get definitive results, the statistical analysis was subsequently completed and visualisations generated

strategy. This innovative approach aims to refine neuroimaging techniques, ensuring more accurate, robust, and versatile applications in various research contexts.

III. NRAA—NON-RIGID REGISTRATION ALGORITHM ANALYSIS FRAMEWORK

This innovative framework for evaluating non-rigid registration algorithms in fMRI data is distinguished by its comprehensive approach, integrating three key metrics:

- **Peak Activation Intensity Analysis:** Identifies and localises peak activation points within functional brain networks, crucial for precise brain activity mapping. This metric, leveraging *'lmax_zstat.txt'*, is optimised in our framework for enhanced accuracy.
- **Cluster-Based Evaluation:** Utilises *'cluster_zstat.txt'* to evaluate the spatial distribution and size of activation clusters. This is integral for maintaining the integrity of functional networks and is adapted in our framework for refined analysis.
- **Thresholded Activation Mapping:** Via *'thresh_zstat'* files, providing visual and statistical representations of significant brain activation areas. This metric is crucial in our framework for accurate localisation and interpretation of functional networks.

In the pre-processing stage, we employ the FMRIB Software Library (FSL) for image processing, with a focus on the unique impact of different registration algorithms. Our procedure includes registering fMRI images to the subject's structural MRI and the MNI 152 2mm standard template, followed by direct registration of the time series 4D fMRI image to the MNI152 template. The distinctiveness of our approach lies in the interchangeable use of different registration algorithms, allowing for an in-depth analysis of BOLD signal variations in resting-state networks.

We introduce the processing pipeline (Fig. II-D) which accepts multi-modal (structural and functional) images and the MNI152 standard template as inputs. The pipeline operates in parallel, allowing for the simultaneous processing of multiple datasets and enabling inter-subject comparisons.

This section describes the evaluated algorithms, the pre-processing, the processing steps, the dataset, and the evaluation metrics.

A. EVALUATED ALGORITHMS

In this section, we discuss the non-rigid registration algorithms that are currently used in the framework. The differences in spatial transformation approaches among FSL, AFNI, ANTs, and DARTEL stem from their underlying algorithms, optimisation strategies, and intended use cases. Table 1 briefly highlights these differences. Additionally, we provide a comparative overview of the algorithms among themselves below.

- **FSL vs. AFNI:** Both offer linear and non-linear options, but FSL's FNIRT and AFNI's 3dQwarp differ in their specific approaches to non-linear registration. FSL's user-friendly interface contrasts with AFNI's focus on functional MRI and its requirement for more user input.
- **FSL/AFNI vs. ANTs:** ANTs stands apart for its symmetric normalisation in non-linear registration, which is not a primary feature in FSL's FNIRT or AFNI's

TABLE 1. This table encapsulates the differences in spatial transformation approaches among FSL, AFNI, ANTs, and DARTEL, highlighting their underlying algorithms, optimisation strategies, and intended use cases. Each algorithm is specifically selected for its ability to address distinct aspects of image registration, ensuring robustness and precision in our framework.

Algorithm	Transformation Type	Key Features	Optimisation Strategy	Unique Advantages	Use Cases	References
FSL	Linear/Non-linear	Uses FLIRT for linear transformations and FNIRT for non-linear transformations. Optimises standard cost functions like correlation ratio and mutual information.	FLIRT/FNIRT	Accommodates local deformations, flexibility in image alignment	Image alignment	[26], [41] [69], [70]
AFNI	Linear/Non-linear	Tailored for functional MRI data. @auto_tlrc for linear and 3dQwarp for non-linear registration.	@auto_tlrc / 3dQwarp	High degree of user-defined customisation in registration	Functional MRI studies	[71], [72], [73]
ANTs	Non-linear	Known for symmetric normalisation in non-linear registration. Captures subtle anatomical variations.	Symmetric Normalisation	High precision in capturing anatomical variations	Detailed morphometric analyses	[14], [74]
DARTEL	Non-linear	Utilises diffeomorphic mappings for high-dimensional warping. Creates group-specific templates.	Diffeomorphic Mapping	Ensures topological preservation, effective for voxel-based morphometry	Studies requiring high accuracy in alignment	[75], [76]

3dQwarp. ANTs is often chosen for studies needing very high precision in alignment.

- **DARTEL vs. Others:** DARTEL's approach to creating a group-specific template and its use of diffeomorphic mappings make it distinct from the other three. While FSL and AFNI focus on a broader range of neuroimaging applications, DARTEL specialises in high-accuracy morphometric analyses.

Each tool has specific strengths and should be chosen based on the specific requirements of a study, such as the need for high precision, user-friendliness, or specialised applications like volumetric analysis or high-accuracy morphometric analyses.

B. DATASET

Robust results in neuroimaging research often require large sample sizes, particularly for studies analysing between-subject effects [77]. To meet this need, we selected the "AOMIC-ID1000" dataset (N = 928), notable for its comprehensive representation of the general population and its diverse range of MRI data. This dataset, accessible at https://openneuro.org/datasets/ds003097/versions/1.2.1, includes structural, diffusion, and both task-based and resting-state functional MRI data, along with concurrent physiological data and extensive demographic information. Such diversity makes it an ideal choice for a broad and accurate analysis in our study.

From this dataset, we carefully selected N = 815 participants, prioritising those with robust activation signals to ensure the reliability of our findings. This selection was based on a verification process conducted, focusing on the robustness of results and minimising pre-processing issues often encountered in smaller samples. This strategy ensures that our analysis is grounded in data representative of a wide range of brain activities.

Table 2 in our manuscript presents an overview of various datasets we evaluated for implementation in our framework. The "AOMIC-ID1000" was ultimately chosen for its large number of subjects and its extensive focus on the brain, which aligns with our objective to conduct a comprehensive and nuanced analysis of brain function without resorting to synthetic data. Figures 3 and 4 show modalities and demographics collected.)

FIGURE 3. This figure shows the modalities collected in the dataset chosen for this study. We focus on the T1w and fMRI modalities from the Amsterdam Open MRI Collection [77], which is a collection of multi-modal MRI datasets for individual difference analysis. Specifically, we elected to use the *ID1000* dataset, which is described in this section.

C. PRE-PROCESSING

This section outlines the pre-processing steps in MRI brain imaging data. Our pre-processing protocol is designed to enhance the sensitivity of statistical analysis and ensure the validity of results, ultimately leading to more accurate and reliable research findings.

The primary goal of our pre-processing is to condition the data for analysis by eliminating artefacts. This ensures that variability in our experimental framework is solely due to differences in registration algorithms. Our approach includes brain extraction, artefact detection and removal, normalisation, and registration [45].

TABLE 2. Public datasets used in the state-of-the-art as well as deep learning-based medical image registration were considered for use in this research. This list is not comprehensive as datasets keep evolving. Sorted by region of interest. These datasets were investigated and subsequently, *The Amsterdam Open MRI Collection* (in table highlighted in italics) was chosen due to the largest number of subjects and focus on functional MRI.

ROI	Dataset	Modality
Abdomen, Lungs	Learn2reg 2020 Lung CT, Abdominal CT-MRI50 [78]	CT, MRI
	LIDC-IDRI, LUNA16 [79]	CT
Brain	*The Amsterdam Open MRI Collection* [77]	MRI, fMRI, DWI (MR)
	ADNI [80]	MRI, CT, CBCT, US, TRUS, x-ray
	NIREP, LPBA, IBSR, CUMC, MGH [81]	CT, MRI, x-ray, PET, SPECT, fMRI,
	OASIS ,ABIDE, ADHD200, MCIC, PPMI, HABS, Harvard GSP, the FreeSurfer Buckner40 [82]	MRI, US
	OASIS, HCP-A, BIRN [83]	MRI
	IXI Brain Development Dataset [84]	MRI
	ENIGMA-Schizophrenia DTI [85]	DTI
	BLSA, Cutting Pediatrics, ABIDE, IXI, ADHD200, NDAR, OASIS, fcon_1000, NKI_rockland [86]	MRI
	BraTS , ALBERTs, CT-MRI dataset, LPBA40, IBSR18, CUMC12, MGH10, Continuous Registration Challenge [87]	CT, MRI
Heart	NIH ChestXray14 [88]	MRI, x-ray
	NLST, DIR-Lab [89]	CineMRI, CT
	Grand Challenges in Biomedical Image Analysis, The Cancer Imaging Archive, "ChestX-ray 8" [90]	CT, MRI, PET, x-ray
Liver	RaFD [91]	CT, MRI
Pelvis	LPBA40, IBSR18, CUMC12, MGH10 [92]	CT, MRI

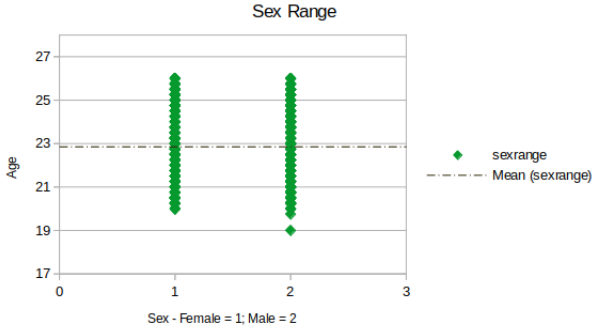


FIGURE 4. This study utilises a dataset consisting of N = 815 participants, including 416 females and 399 males. The age range of the participants is 19 to 26 years old, with an average age of approximately 22.9 years. This dataset was deliberately chosen to address the gender disparities in medical data accessibility.

In our pre-processing workflow for functional neuroimaging data, we prepare each structural and functional images separately. The structural images were corrected for non-uniformity and aftrifacts, and then brain was extracted from the skull. The functional images were motion and slice-timing corrected. This yielded structural and functional pre-processed images ready for processing by registration algorithms.

1) FILE PREPARATION

Our study utilised structural (T1w) and functional MRI images in NIFTI1 format, chosen for its wide compatibility with various software packages and to avoid issues associated with the Analyze format. These 3T high-quality images were acquired in Right-to-Left orientation and processed as detailed in Table 3.

To address motion and slice-timing artefacts in the functional MRI data, we applied the FSL MCFLIRT tool [26] for correction. This step ensured the accuracy of subsequent analyses by stabilising the raw functional files.

In preparation for analysis, additional steps included creating parameter files and configuring command-line options for parallel processing. This was particularly important for efficient processing on High-Performance Computing (HPC) architecture, which significantly reduced the registration time. Each functional image comprised 290 slices, necessitating these optimisations for streamlined data handling.

2) BRAIN EXTRACTION

We utilised FSL BET [93] for skull-stripping, removing each brain from its whole-head image and resulting in a skull-stripped image (illustrated in Figure 5). This step was applied to each MRI volume of the ID1000 dataset.

Post-skull-stripping, we employed various FSL preprocessing methods on these images. These methods are pivotal for conditioning the data for accurate statistical analysis, ensuring optimal data quality.

Skull-stripping of the brain, followed by the removal of motion and distortion artefacts from the functional data, precedes registration. The motion-corrected functional brain is then aligned with the skull-stripped structural brain of the same subject.

D. PROCESSING

Subsequently, after pre-processing, the same subject's structural and functional image is registered to the standard template, with each step's mathematical transformation stored for the transformation of the 4D time series to standard space. Our processing approach is methodically structured, involving several key steps for each subject's data:

1) **Linear Registration:** This stage involves aligning structural and functional brain images from the same subject. The primary objective here is to ensure accurate alignment within each individual's data set [52], [53].

TABLE 3. 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) [77].

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

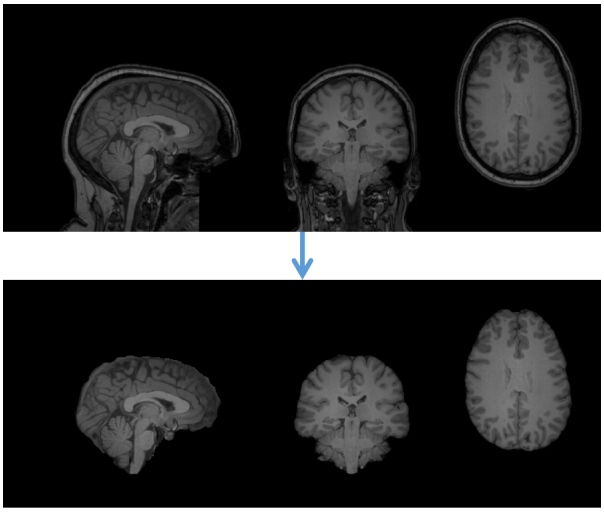


FIGURE 5. For each brain, we obtained two images: the original T1-weighted MRI and the extracted brain. To provide a comprehensive view of each brain, we present sagittal (front-facing right), coronal (right on the right side), and axial (front-facing top, right on the right side) views in three columns (left to right). It is important to note that the images presented in this figure are for illustration purposes only. For this figure, we changed the images' scale, location, and contrast to facilitate visual inspection of the brains.

2) **Non-Linear Registration:** Now, the focus shifts to aligning functional images with the same subject's structural space, which is essential for effective comparison across different subjects [54], [55].
3) **Applying Transformations:** The combined results from the linear and non-linear registration phases are applied to the whole time-series functional images. This step moves the images into a standard space.
4) **Statistical Analysis:** We conduct a thorough statistical analysis, including Univariate analysis via General Linear Model fitting [48], [49] and Multivariate Pattern Analysis via SVMs [50], [51]. This analysis is crucial for identifying patterns and correlations within the data, extending our understanding to encompass group-level brain activity.

We first align functional brain images with structural data from the same subject using linear registration methods. These images are then registered to a standard template through non-linear registration techniques. The combined transformations from these stages are applied to the complete time series of functional images, aligning them with the standard MNI152 space. This careful process ensures that only brain regions are registered, preserving analysis integrity. Recent studies by Dadar et al. [44], Nael et al. [45], and Frost et al. [46] support the effectiveness of these registration methods in neuroimaging. The registration algorithm is then swapped and the same process is repeated.

This concise overview encapsulates the essential steps of our processing stage, ensuring a rigorous and comprehensive analysis of neuroimaging data. Figures 1 and 6 visually exemplify the brain's alignment pre- and post-registration with the standard template.

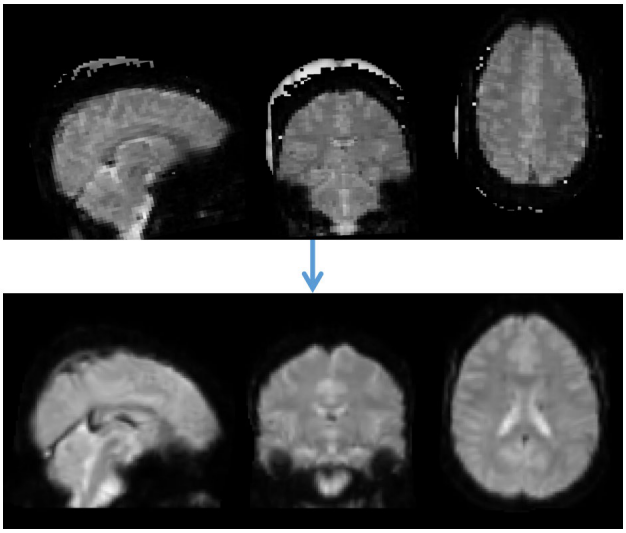


FIGURE 6. Example result of functional registration using the nonlinear method. The upper row represents the subject brain before registration, the lower row depicts the brain registered (resized) onto the standard MNI152 template.

E. EVALUATION METRICS IN NRAAF

The evaluation metrics in our Non-Rigid Registration Algorithm Analysis Framework (NRAAF) utilise a combination of univariate and multivariate analyses to robustly assess registration algorithms:

1) **FSL FEAT Metrics for Univariate Analysis:** Our framework incorporates FSL FEAT metrics for univariate analysis, leveraging their established reliability in measuring spatial precision and alignment accuracy in neuroimaging. These metrics are grounded in the general linear modeling approach of FEAT, aligning with methodologies discussed in Jenkinson et al. [41], which provides a comprehensive overview of FSL tools, including FEAT, in neuroimaging.
2) **Use of SVMs for Multivariate Analysis:** Diverging from the traditional use of Support Vector Machines (SVMs) for classification in neuroimaging, as initially outlined in Cortes and Vapnik [42], our study employs SVMs innovatively within the Multivariate Pattern Analysis (MVPA) framework. This novel application focuses on analysing the performance of registration algorithms in processing fMRI data, specifically assessing their impact on the functional connectivity patterns. The innovative use of SVMs in this context extends beyond typical classification tasks, offering new insights into the effectiveness of registration algorithms in neuroimaging data analysis, as highlighted in the discussion of evolving machine learning techniques in neuroimaging by Varoquaux and Thirion [43].

By integrating these two methodologies, our NRAAF provides a comprehensive assessment of non-rigid registration algorithms, evaluating both their anatomical precision and functional impact. This approach not only aligns with current standards in neuroimaging analysis but also introduces a novel perspective through the unique application of SVMs in MVPA, marking a significant contribution to the field.

IV. IMPLEMENTATION

This section outlines the methodologies and procedures employed in our study, detailing the computational resources, software applications, and the image registration process. We emphasise the critical steps involved in pre-processing neuroimaging data, registering images to a standard space, and analysing the resulting data. Our approach combines well-established techniques with innovative applications, leveraging advanced computational resources and statistical analyses to ensure rigorous evaluation and accurate interpretation of neuroimaging data.

A. COMPUTATIONAL RESOURCES

The analysis was conducted on a CentOS machine, version 8.2.2004-x86_64 (Dell PowerEdge R740 Rack Server). This system, equipped with Intel Xeon Gold 6240 processors totalling 288 cores and 720GB DDR4 RAM, was essential for the computationally intensive tasks of functional neuroimage processing and analysis.

B. SOFTWARE APPLICATIONS

Several software applications were utilised for pre-processing and computing findings for assessment, including but not limited to FSL 6 [41], Freesurfer 7.4 [64], and MATLAB R2023a [65]. Visualisations of the results were generated using MATLAB and Freesurfer Freeview 3.0 [66].

C. IMAGE REGISTRATION PROCESS

In our study's implementation, the image registration process was tailored to address specific challenges and requirements of our fMRI data analysis:

- **Efficiency in Execution:** Recognising the extensive data volume, we conducted the initial registration and standard space alignment in parallel, significantly reducing processing time in comparison to the non-parallel processing. This approach was instrumental in performing a total of 6520 whole-brain 3D-3D registrations, both linear and non-linear, and 3260 3D-4D transformations.
- **Iterative Approach:** After initial registrations, we systematically replaced the registration algorithm, enabling a comprehensive comparative analysis across different techniques. This iterative method provided unique insights into the performance of various registration algorithms in our dataset.
- **Maintaining Data Integrity:** A key focus was maintaining the quality of the fMRI data. By directly moving the 4D fMRI time series images to the MNI152 template by utilising previously saved transformations, we avoided the quality degradation typically associated with repeated registrations. While direct evidence for quality degradation due to repeated registrations is limited, related studies in the field suggest potential issues with registration quality [14], [41], [47].

Through these tailored strategies in image registration, our study not only achieved efficiency and precision but also offered a deeper insight into the comparative effectiveness of different registration methods. This meticulous approach sets a foundation for the subsequent analyses, ensuring that our findings are grounded in robust and accurately aligned images.

D. CLUSTER ANALYSIS AND STATISTICAL PROCESSING

Descriptive statistics provided summary measures of the data's central tendency and variability, which include *mean, median, standard deviation, range, minimum, maximum, interquartile range (IQR), skewness, and kurtosis*. These statistics are essential for a comprehensive understanding of the distribution of peak activation intensities.

While exploring various analytical approaches, the potential application of Principal Component Analysis (PCA) was considered. PCA is often used in neuroimaging studies to reduce data dimensionality and to identify underlying patterns in complex datasets. However, given the specific

context of our study, which focuses on investigating resting state networks with four distinct algorithmic components, the application of PCA was deemed unnecessary. This decision was guided by the targeted nature of our research objectives, which centre on the direct comparison of these predefined algorithms rather than uncovering latent structures within a larger set of variables. Nevertheless, future studies with larger sets of variables or different research focuses might benefit from the application of PCA or similar multivariate techniques.

Cluster peak activations were quantified as the highest intensity points in regions of interest. Significant clusters were numerically identified, with associations between peak activation and significant clusters in resting-state networks being tested. fMRI data processing utilised FEAT (FMRI Expert Analysis Tool) Version 6.00, part of FSL. Z (Gaussianised T/F) statistic images were thresholded using clusters determined by $Z>2.3$ and a (corrected) cluster significance threshold of $P=0.05$ [67].

1) INTER-SUBJECT STATISTICAL TESTS

Statistical analyses were conducted using MATLAB release R2023a. The volumetric measurement of the Control Network activations was based on voxel sizes of 2 2 2 mm^3, with a robust dataset of $N = 815$ subjects. Despite the substantial sample size, the *Shapiro-Wilk test* indicated non-normality of the data; hence, non-parametric methods were chosen for further analysis alongside descriptive statistics. The chosen statistical tests were *Mann-Whitney U Test, Wilcoxon Signed-Rank Test, Kruskal-Wallis Test*, and *Spearman's Rank Correlation*. They were selected due to their robustness in handling non-normality and fewer assumptions about the underlying data distribution [62], [63].

2) EFFECT SIZES AND ERROR CORRECTION:

In this study, we not only focused on statistical significance but also on the practical significance of our findings. To achieve this, we incorporated effect size calculations and error correction methods into our analysis. The rank-biserial correlation is used as the effect size for *Mann-Whitney U Test* and *Wilcoxon Signed-Rank Test*, determined by $r = \frac{Z}{\sqrt{N}}$. For *Spearman's Rank Correlation*, the rho coefficient directly represents the effect size. To mitigate the risk of Type I errors (false positives) in our analysis with multiple comparisons, we employed the *Bonferroni Correction* in contrast to methods like the False Discovery Rate (FDR). Despite its potential to increase Type II errors (false negatives), the Bonferroni Correction is a common and robust method in complex studies to ensure validity against accidental findings [40], [94].

3) APPLICATION OF MULTIVARIATE PATTERN ANALYSIS

Beyond non-parametric tests, our study utilises Multivariate Pattern Analysis (MVPA) via Support Vector Machines (SVMs) to generate heat maps that provide visualisation of the similarity between the voxel intensities reported by the algorithms, facilitating a detailed intersubject analysis of the Control Network activations. Mutual Information (MI) is employed alongside SVM to enhance visual inspection and provide a statistical metric of interdependence between subjects.

The methodologies implemented include a novel use of SVMs to analyse neuroimaging data. Unlike traditional SVM applications in neuroimaging, which predominantly focus on classification tasks [31], [32], our approach uses SVM for direct comparison of algorithmic performance in processing resting-state fMRI data. This technique, illustrated by Weaverdyck et al. [33] and further expanded in our study, delineates decision boundaries on a voxel-by-voxel basis, enabling a detailed pairwise comparison of the algorithms. Such comparative analysis through SVMs is scarcely addressed in current literatures [34] and [35], marking our study as a significant advancement in neuroimaging research. The heat maps derived from these SVM weights offer a visual assessment of each algorithm's impact on connectivity patterns, a technique that echoes the findings of Steardo et al. [36] and Mikolas et al. [37] in their respective fields.

SVMs are widely recognised for their robust classification capabilities in neuroimaging [38], [39], excelling in discerning and predicting patterns from complex data. Our application of SVM extends beyond traditional feature selection methods, like those using segmented gray matter or regions of interest [31], to a broader analytical scope, reflecting the versatility and innovative potential of SVM in neuroimaging.

In this analysis, MI serves as a robust statistical measure for quantifying the registration accuracy among different algorithms applied to Resting-State fMRI data. Grounded in information theory, MI provides a nonlinear correlation metric between datasets [36], crucial in neuroimaging where precise voxel alignment is imperative for accurate functional connectivity mapping.

Our methodologies included advanced image registration processes, clustering analysis, and statistical processing of activation data, as well as the innovative application of machine learning techniques like Support Vector Machines (SVMs) for feature extraction and heat map generation. These approaches enabled us to conduct an intricate inter-subject analysis of the Control Network activations, leveraging the structural detail provided by the Harvard-Oxford atlas to enhance our understanding of the underlying neural mechanisms.

V. EVALUATION RESULTS

In this analysis, we examined the variability in peak activation identification reporting across 815 subjects using four different non-rigid registration algorithms: FSL, ANTs, DARTEL, and AFNI. Focusing on the strength of the correlation between the Control Network and its seed region, and the difference of each algorithm to the correlation.

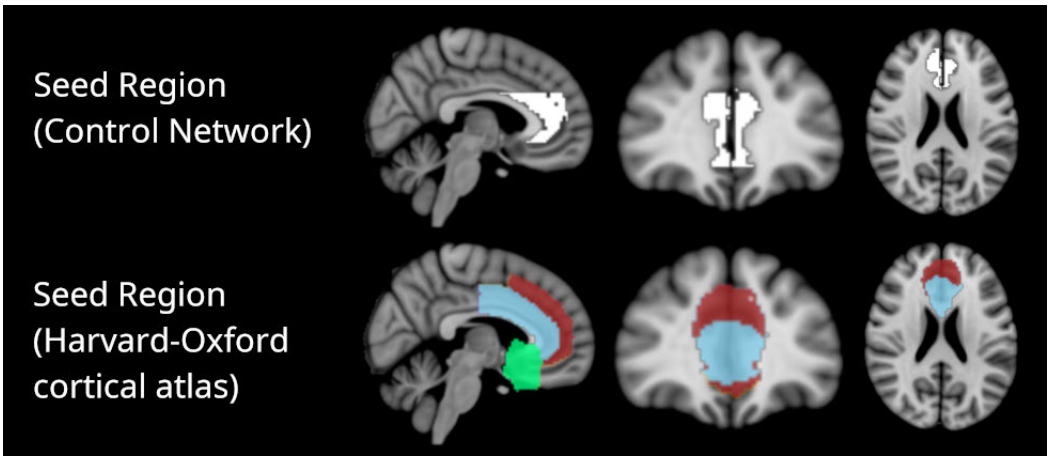


FIGURE 7. Seed Region and Corresponding Harvard-Oxford Atlas Regions. The first row displays the identified seed region for the control network in both hemispheres, shown in sagittal, coronal, and axial views. The second row shows the corresponding regions within the Harvard-Oxford cortical atlas where the seed region is located, including the Paracingulate Gyrus, Cingulate Gyrus (anterior division), and Cingulate Gyrus (posterior division). This juxtaposition aids in understanding the anatomical context of the seed region within established brain atlases.

We present the results by hemispheres. Our study aimed to investigate how these algorithms influence neuroimaging data interpretation, specifically resting-state networks. These results, are crucial for the validation and refinement of our proposed framework, paving the way for more extensive analyses and enhancements in future research methodologies.

A. SEED REGION

The commonly adopted seed region for the control network was precisely localised using RAS (Right, Anterior, Superior) coordinates, a conventional neuroimaging coordinate system. The specified coordinates (R: 4.60, A: 35.44, S: 21.78) target an area within the brain's left hemisphere, implicating regions such as the Paracingulate Gyrus, and the anterior and posterior divisions of the Cingulate Gyrus. This region's involvement is critical given its association with key functions of the control network.

The identified seed region for the control network, encompassing the Paracingulate and Cingulate Gyri (anterior and posterior divisions), is pivotal for several reasons. Functionally, these regions are integral to the control network, playing a critical role in higher-order cognitive processes such as attention, working memory, and decision-making, and are essential in cognitive control and emotional regulation [56], [57]. Statistically, the seed region's localisation is influenced by the variability in registration algorithms like FSL, ANTs, DARTEL, and AFNI. This variability affects the interpretation of neural responses, as measures like mean, median, standard deviation, skewness, and kurtosis provide insights into the consistency and distribution of activation intensities [58], [59]. Moreover, in resting-state networks, the consistency of algorithmic registration in these regions is crucial for understanding functional connectivity and intrinsic neural activities, thereby influencing the analysis of resting-state networks [60]. Differences in algorithm performance critically affect control network analysis in resting-state fMRI, underscoring the importance of accurate localisation and intensity measurement. Recent research emphasises the impact of functional connectivity measures on neuroimaging accuracy [61]. The seed region is shown in Figure 7.

B. ATLAS MEASUREMENTS

We measured the strength of the correlation to the seed region and difference of each algorithm contribution to the correlation of each subject using the Harvard-Oxford cortical and subcortical structural atlas on MNI152 space. The Harvard-Oxford cortical structural atlas provides a regional parcellation comprising 96 regions (48 regions 2, left and right), and 17 regions in the subcortical structural atlas (8 from the cerebral cortex, thalamus, caudate, pallidum, hippocampus, amygdala, and accumbens 2, left and right; and the brain stem) [95], [96], [97], [98].

C. STATISTICAL INSIGHTS IN NEUROIMAGING DATA INTERPRETATION

This section presents the analysis of the descriptive statistics of peak activation intensity in both the left and right hemispheres of the activated regions, using four distinct neuroimaging algorithms: ANTs, DARTEL, AFNI, and FSL. The focus is on understanding algorithmic variations in correlation strength and hemisphere differences.

As Table 4 reveals, ANTs demonstrated a high mean (6.9184) and maximum value (21.2000), indicating a tendency to detect higher activation intensities. The skewness (3.1411) and kurtosis (16.4160) suggest a distribution with significant outliers. DARTEL showed a lower mean (6.3695) with a narrower standard deviation (1.4351), suggesting more consistent but potentially less sensitive measurements of activation intensities, as indicated by the kurtosis (21.8712). AFNI presented a considerable range (23.6000), indicating high variability in activation intensities. The zero minimum

TABLE 4. Descriptive Statistics of *Left Hemisphere* Peak Activation Intensity in the Control Network. This table presents a comparative analysis of four neuroimaging algorithms: ANTs, DARTEL, AFNI, and FSL. It highlights ANTs' tendency to detect higher activation intensities with significant outliers, DARTEL's more consistent but less sensitive measurements, AFNI's high variability, and FSL's balanced sensitivity and variability in activation detection.

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

TABLE 5. Descriptive Statistics of *Right Hemisphere* Peak Activation Intensity in the Control Network. This table compares the same four neuroimaging algorithms as Table 4. It illustrates ANTs' propensity for detecting intense activations, DARTEL's consistent performance across hemispheres, AFNI's notable variability in activation intensities, and FSL's comparable and balanced detection capabilities across both hemispheres.

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

value could suggest instances of no detected activation. FSL reported a moderate mean (6.7825) and standard deviation (1.8348), suggesting balanced sensitivity and variability in activation detection.

ANTs similar to the left hemisphere, showed high mean (6.9339) and maximum (31.6000) values, with even higher skewness (4.3537) and kurtosis (39.0803), indicating a propensity for detecting intense activations. DARTEL was consistent with the left hemisphere; it had a slightly higher mean (6.5036) but similar patterns in other metrics, suggesting uniformity across hemispheres. AFNI displayed similar characteristics to the left hemisphere, with a large range (22.2000) and maximum value (22.2000), indicating notable variability. FSL showed comparable statistics to the left hemisphere, maintaining a balance between sensitivity and variability. For all the values please refer to Table 5.

Within our study, histograms are utilised to depict the frequency distribution of peak activation intensities derived from fMRI data. Each histogram offers a visual summary of how often various intensity values occur within the dataset, which corresponds to the precision of different non-rigid registration algorithms in mapping brain activity. This representation allows for an immediate grasp of the central tendency and dispersion of the peak activation intensities.

The histograms for AFNI's peak cluster intensities in both hemispheres exhibit a right-skewed distribution, indicating a prevalence of lower intensity values with fewer occurrences of higher intensities (Fig. 8). This skewness suggests a potential for non-detection or underestimation of intensity values by the AFNI algorithm. The red line, representing the mean, and the green dashed line, indicating the median, are both shifted towards the lower end of the intensity scale, further emphasising the skewness of the distribution. The red line shows the mean and the median is denoted by the green dashed line.

In the case of ANTs (Fig. 9), the histograms also display a right-skewed distribution, but with a mean and median that are closer together, suggesting a more symmetrical distribution around the central value compared to AFNI. The distribution implies that ANTs tends to detect higher intensity activations more frequently than AFNI.

The DARTEL algorithm's histograms show a narrower and more symmetric distribution, with a median that closely approximates the mean, indicating a balanced detection of peak intensities with less skewness (Fig. 10). This symmetry may reflect DARTEL's consistent performance in capturing the central tendency of the data without a significant bias toward higher or lower intensity values.

Finally, the histograms for FSL (Fig. 11) show a distribution that is slightly right-skewed, with the mean and median closely aligned. This suggests that FSL, while similar to ANTs in its tendency to detect higher intensities, may offer a more balanced approach with less variability in the detection of peak intensities across the sample.

From these histograms, we can infer the following about algorithmic differences:

- AFNI may be less consistent in detecting peak intensities, as evidenced by the broader spread and right skewness, potentially indicating issues with sensitivity or calibration.
- ANTs appears to be more sensitive to higher intensity activations but also shows signs of right skewness, which could point to variability in peak detection.

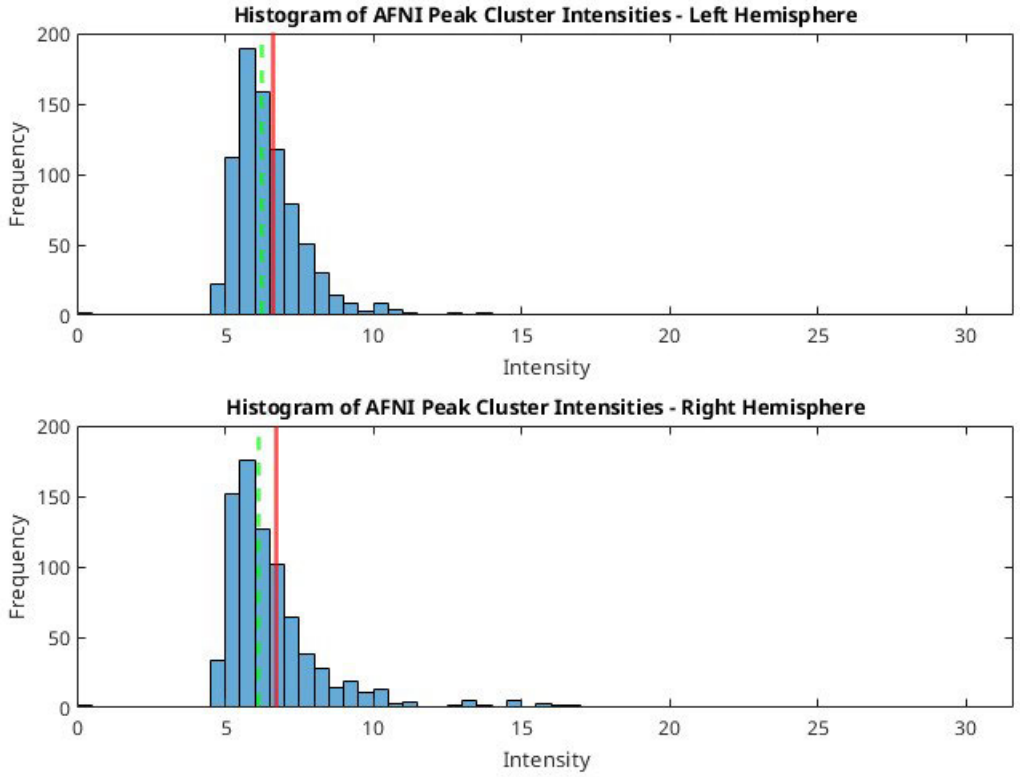


FIGURE 8. Histograms of AFNI Peak Cluster Intensities in Resting-State fMRI Data. These histograms display the frequency distribution of peak activation intensities for AFNI's algorithm in both brain hemispheres. The right-skewed distribution highlights a prevalence of lower intensity values and fewer occurrences of higher intensities. The mean (red line) and median (green dashed line) are both shifted towards the lower end of the intensity scale, indicating a potential underestimation or non-detection of higher intensity values by the AFNI algorithm. This visual representation aids in understanding the algorithm's precision in mapping brain activity, showing central tendencies and dispersion of peak activation intensities.

- DARTEL presents a more uniform detection across a range of intensities, suggesting a reliable performance with less variability.
- FSL seems to strike a balance between sensitivity and variability, detecting a moderate range of intensity activations without significant skewness.

These inferences are supported by the numerical results showing means, medians, standard deviations, and measures of skewness and kurtosis for each algorithm. The histograms, along with these descriptive statistics, provide a comprehensive view of each algorithm's performance, allowing for an assessment of their strengths and limitations in the context of peak intensity detection in resting-state network analysis. This comparison is crucial for researchers selecting algorithms for fMRI data processing, as it directly impacts the accuracy and reliability of functional network mapping.

1) HEMISPHERE COMPARATIVE OVERVIEW

ANTs consistently exhibit high mean and maximum values in both hemispheres, suggesting a tendency to detect more intense activations. The skewness and kurtosis further indicate a propensity for outliers and extreme values. DARTEL demonstrates more moderate and consistent measurements across hemispheres, albeit still with indications of non-normality. AFNI displays a wide range and zero minimum in both hemispheres, pointing towards high variability and potential instances of non-detection. FSL presents a balance between sensitivity and variability, with moderately high mean and standard deviation, and less extreme skewness and kurtosis compared to ANTs and AFNI.

The consistency in algorithmic performance across hemispheres is noteworthy. Although our analysis does not indicate significant hemispheric differences in peak activation intensities, this uniformity is crucial for understanding hemispheric specialisation or symmetry in brain function. The findings suggest that the observed variations in activation intensities are more reflective of the algorithmic processing rather than intrinsic differences between the hemispheres. This insight is significant for neuroimaging studies investigating lateralisation or hemispheric asymmetries in brain functions, as it underscores the importance of algorithm selection in interpreting such hemispheric differences [99], [100].

The boxplots and violin plots for each of the four non-rigid registration algorithms — ANTs, DARTEL, AFNI, and FSL — provide a visual summary that complements the descriptive statistics and reveals differences in algorithmic performance across the left and right hemispheres. The key

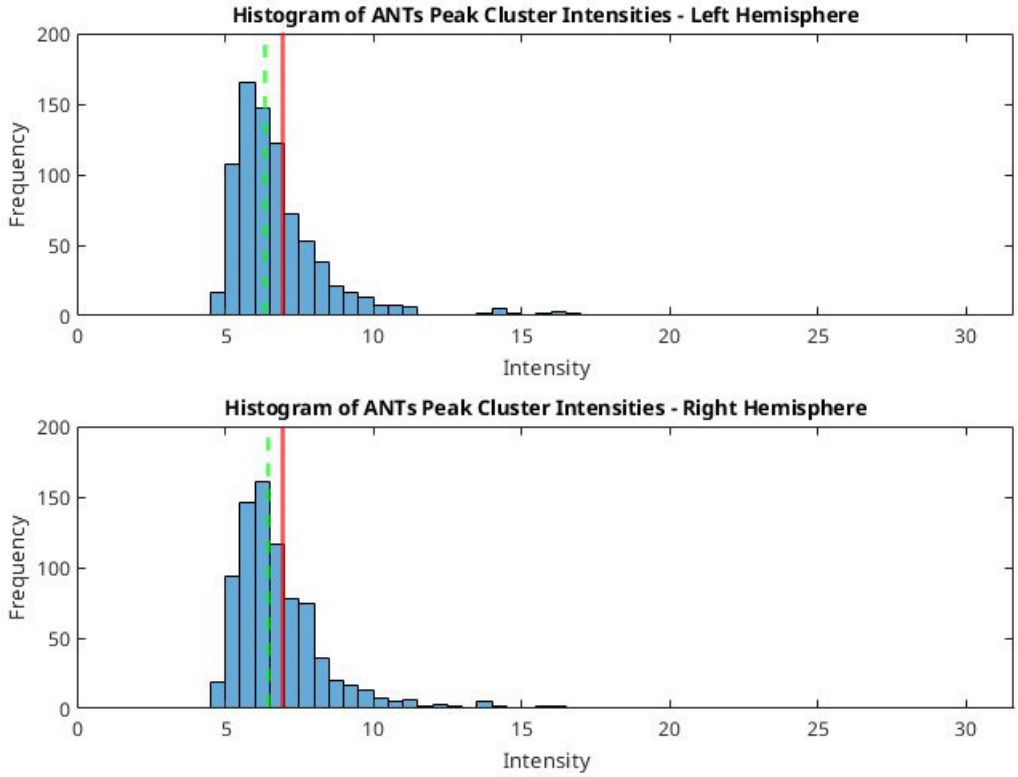


FIGURE 9. Histograms of ANTs Peak Cluster Intensities in Resting-State fMRI Data. These histograms illustrate the frequency distribution of peak activation intensities for the ANTs algorithm in both brain hemispheres. The right-skewed distribution is notable but with a mean and median that are closely aligned, indicating a more symmetrical distribution around the central value compared to AFNI. This pattern implies that ANTs tends to detect higher-intensity activations more frequently than AFNI, as evidenced by the distribution's shape.

features of these plots and the inferences drawn about each algorithm in the hemispheric context are detailed below. The star notation is explained in Table 6.

TABLE 6. Explanation of star notations used in box plots.

Symbol	Significance
ns	not significant $p > 0.05$
*	significant $p \leq 0.05$
**	very significant $p \leq 0.01$
***	extremely significant $p \leq 0.001$
****	most extreme significance $p \leq 0.0001$

In the context of hemispheric comparison, these plots allow us to discern whether the algorithms perform differently in detecting activation intensities in the left versus the right hemisphere. For instance, a box plot with multiple stars suggests a robust hemispheric difference in the algorithm's performance. In contrast, an 'ns' above a plot suggests that the algorithm does not exhibit a significant preference or difference in detection capability between hemispheres.

From the provided plots (Figs. 12, 13), it is observable that some algorithms show statistically significant differences in peak activation intensities, while others do not. This variance underscores the importance of algorithm selection when interpreting hemispheric differences in neuroimaging data. The careful choice of registration algorithms is crucial as it directly impacts the accuracy of localising and interpreting functional brain activities, particularly when exploring the lateralisation of brain functions or hemispheric asymmetries.

Comparative box plots of Peak Cluster Intensities across algorithms for left (Fig. 12) and right (Fig. 13) hemispheres illustrate the interquartile range (IQR), median, and outliers for peak cluster intensities detected by each algorithm in both hemispheres. The plots highlight the central tendency, dispersion of values, and the presence of outliers, elucidating the differences in detection capabilities and consistency among algorithms. Notably, AFNI displays several outliers, particularly in the right hemisphere, indicating high variability and potential for extreme value detection or non-detection. The median lines are consistent across hemispheres for each algorithm, suggesting an absence of significant hemispheric bias in the detection of peak cluster intensities. From the boxplots, it is evident that:

- ANTs display a relatively wide IQR in both hemispheres, suggesting a larger spread of intensity values, which indicates variability in intensity detection.
- DARTEL exhibits a tighter IQR, implying more consistent detection of intensity values across the sample.
- AFNI shows a wide range with several outliers, particularly in the right hemisphere, indicating high variability and potential for extreme value detection or non-detection.

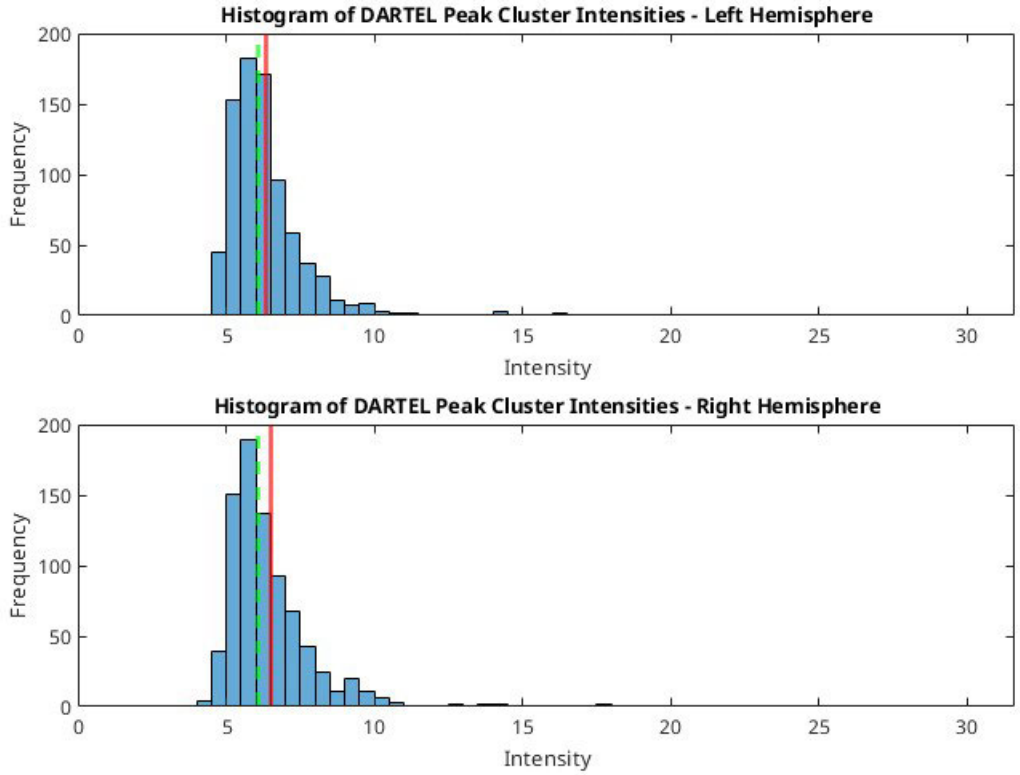


FIGURE 10. Histograms of DARTEL Peak Cluster Intensities in Resting-State fMRI Data. These histograms represent the frequency distribution of peak activation intensities derived using the DARTEL algorithm in both brain hemispheres. Characterised by a narrower and more symmetric distribution, the median closely approximates the mean, indicating a balanced detection of peak intensities with less skewness. This symmetry in the distribution may reflect DARTEL's consistent performance in accurately capturing the central tendency of the data without significant bias towards higher or lower intensity values.

- FSL presents a balance with moderate IQR and fewer outliers, suggesting a reliable detection of intensities with less variability compared to ANTs and AFNI. The median lines (central mark in the boxplot) for each algorithm are relatively consistent across both hemispheres, suggesting no significant hemispheric bias in the detection of peak cluster intensities by the algorithms.

The median lines (central mark in the boxplot) for each algorithm are relatively consistent across both hemispheres, suggesting no significant hemispheric bias in the detection of peak cluster intensities by the algorithms.

Hemispheric comparisons of Peak Cluster Intensities using violin plots (Figs. 14 and 15) for non-rigid registration algorithms. The violin plots illustrate for left and right hemispheres, respectively, combining boxplot features with kernel density estimation. These plots provide insights into the frequency and distribution of peak cluster intensities for each algorithm, emphasising hemispheric differences and similarities. The plots reveal that while ANTs and AFNI display wider distributions in both hemispheres, indicating a variety of intensity values, DARTEL shows more uniformity, reflecting consistent detection across hemispheres. FSL, on the other hand, presents a narrower range in both hemispheres, suggesting fewer extreme intensity values.

Violin plots combine the boxplot with a kernel density estimation, providing more insight into the distribution of the data. The thickness of the violin plot indicates the frequency of data points at different intensity levels. From the violin plots, we observe that:

- The distributions for ANTs and AFNI are wider in several sections, indicating a variety of intensity values with a frequent occurrence.
- DARTEL's distribution is more uniform, reflecting its consistency across different intensity values.
- FSL's distribution, while similar in shape to ANTs and AFNI, does not extend as far, implying fewer occurrences of extreme intensity values.

The mean and median indicated by lines within the violin plots show the central tendency of the data. A closer alignment of the mean and median suggests a more symmetrical distribution, while greater separation might indicate skewness.

The hemispheric comparison demonstrates how the algorithms perform in detecting activations across the left and right hemispheres. The analysis, supported by statistical metrics, indicates that some algorithms exhibit significant hemispheric differences, which are essential for studies focusing on lateralisation or symmetry in brain function. The consistency across hemispheres also highlights the impact

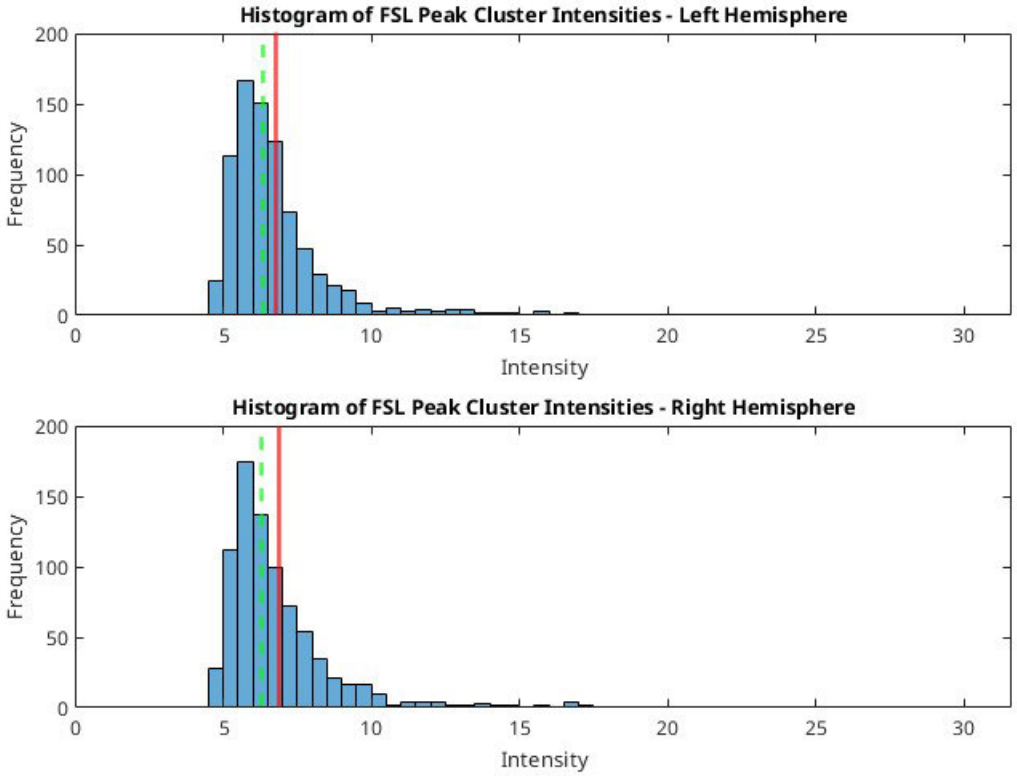


FIGURE 11. Histograms of FSL Peak Cluster Intensities in Resting-State fMRI Data. These histograms exhibit the frequency distribution of peak activation intensities for the FSL algorithm in both brain hemispheres. The distribution is slightly right-skewed, yet the mean and median are closely aligned, suggesting that FSL, while similar to ANTs in its tendency to detect higher intensities, provides a more balanced approach with less variability in the detection of peak intensities across the sample.

of algorithm choice on the interpretation of hemispheric differences.

2) NON-NORMALITY INDICATORS

Across both hemispheres and all algorithms, the skewness and kurtosis values are significantly high. For instance, ANTs in the right hemisphere shows a kurtosis of 39.0803. These metrics indicate that the data distribution deviates from normality, presenting with heavy tails and pronounced peaks.

The non-normality of the data suggests that the peak activation intensities are not evenly distributed around the mean. This has implications for statistical analysis and interpretation, as many standard statistical tests assume normality.

The significant skewness and kurtosis values across all algorithms confirm the non-normality of the peak activation intensity data. This deviation from normality necessitates the application of specialised statistical methods that do not rely on normal distribution assumptions, such as non-parametric tests or bootstrap methods. These approaches can provide more accurate interpretations of neuroimaging data, especially when exploring complex neural networks or conducting comparative algorithmic analyses [101].

The results show notable algorithmic variations in the measurement of peak activation intensities, with ANTs and AFNI tending towards higher and more variable intensities, and DARTEL and FSL showing more consistent but less variable intensities. The significant skewness and kurtosis across all algorithms confirm the non-normality of the data, a critical factor that must be considered in statistical analyses and interpretation. This non-normality highlights the need for specialised statistical approaches that do not rely on the assumption of normally distributed data.

The consistency in algorithmic performance across hemispheres underscores the importance of algorithm selection in neuroimaging studies. The choice of algorithm can greatly influence the detected activation patterns, impacting the interpretation and conclusions drawn from neuroimaging data. These insights are crucial for researchers in selecting the most appropriate algorithms for their specific study objectives, especially in investigations involving complex brain networks and resting-state activities.

The comparative analysis of algorithms provides critical insights into their respective strengths and limitations. ANTs, with its higher mean and maximum values, appears particularly adept at detecting more intense neural activations, which may be beneficial in studies focusing on highly localised brain activities or specific functional regions. DARTEL, exhibiting more moderate and consistent measurements, could be more suitable for studies requiring a balanced approach to sensitivity and variability, such as in broad-based neuroimaging research exploring general brain functions.

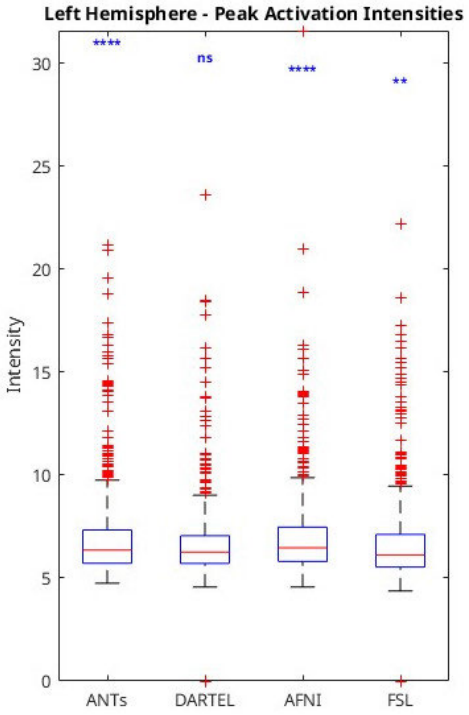


FIGURE 12. This box plot demonstrates the performance of each algorithm in the left hemisphere, with ANTs showing a wide IQR and high variability, DARTEL exhibiting a tight IQR indicating consistent detection, and FSL presenting a moderate IQR with fewer outliers, reflecting a balanced approach.

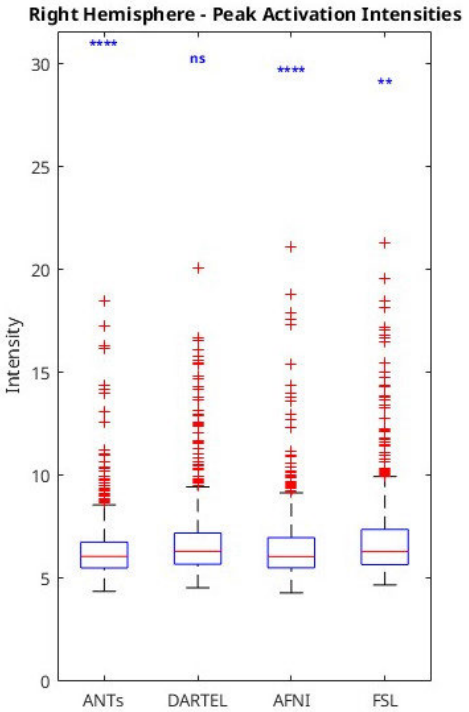


FIGURE 13. This plot illustrates the algorithms' performance in the right hemisphere. It underscores AFNI's high variability with numerous outliers, indicating potential extreme value detection or non-detection, compared to the more balanced detection and fewer outliers in FSL.

AFNI's wide range and high variability suggest its utility in studies where a broader spectrum of neural activities is of interest, potentially including both high and low-intensity activations. FSL, with its balanced sensitivity and moderate variability, could be preferred in studies aiming for a middle-ground approach, balancing detection sensitivity with variability in activation intensities [41], [102].

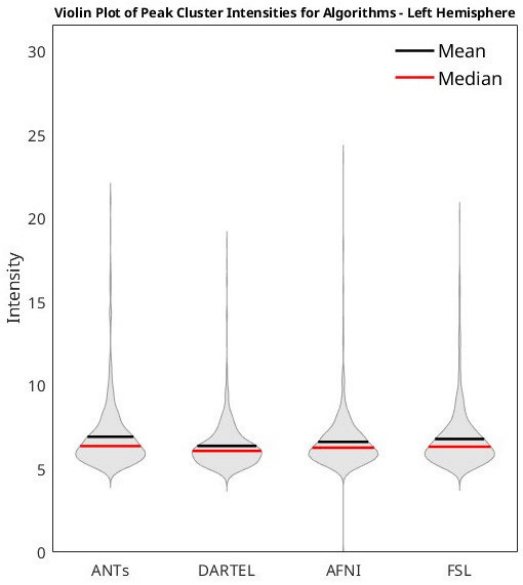


FIGURE 14. This plot details the distribution of peak cluster intensities in the left hemisphere for each algorithm. It underscores the hemispheric performance of ANTs and AFNI with their broader range of intensity values and contrasts this with the more uniform and consistent detection by DARTEL and the narrower range shown by FSL.

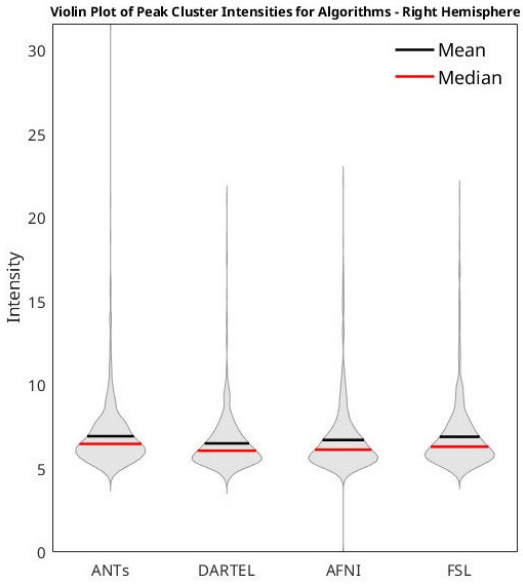


FIGURE 15. Similar to the left hemisphere, this plot for the right hemisphere reflects the distribution characteristics of each algorithm. The comparative analysis between hemispheres highlights the consistent performance of each algorithm, with particular emphasis on the variability and range of intensity values detected by ANTs and AFNI, and the more balanced detection by FSL.

This section presents a detailed analysis of peak activation intensities across different algorithms. The histograms and descriptive statistics reveal how each algorithm — FSL,

ANTs, DARTEL, and AFNI — varies in terms of sensitivity and variability in detecting activations. This comparative insight is crucial for selecting the most appropriate algorithm for specific fMRI studies, impacting the accuracy and reliability of functional network mapping.

D. STATISTICAL ANALYSIS - NON-PARAMETRIC TESTS

In the current study, Spearman correlation coefficients (Table 7) were calculated to assess the association between hemispheres for peak activation intensities using various algorithms. AFNI demonstrated a robust correlation (p = 0.66346, p <.001), indicating a strong monotonic relationship between hemispheres. This was closely followed by ANTs (p = 0.61647, p <.001), suggesting a strong yet slightly less pronounced association. DARTEL (p = 0.52181, p < .001) and FSL (p = 0.53823, p < .001) showed moderate correlations, indicating a consistent but weaker linkage between hemispheric intensities compared to AFNI and ANTs.

TABLE 7. Adjusted Spearman's Rank Correlation Coefficients and p-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 peak activation intensities for ANTs, DARTEL, AFNI, and FSL. RHO values range from moderate to strong, and the adjusted p-values denote high statistical significance for each algorithm.

Algorithm	RHO	Adjusted p value
ANTs	0.61647	2.8692e-85
DARTEL	0.52181	6.6096e-57
AFNI	0.66346	2.5794e-103
FSL	0.53823	3.3986e-61

The non-parametric Mann-Whitney U Test (Table 8) revealed significant hemispheric differences for ANTs (U = 1.382e-09, r = 0.16073) and AFNI (U = 4.4707e-10, r = 0.16489), suggesting a difference in central tendency of peak intensities between the two hemispheres. In contrast, DARTEL showed no significant hemispheric difference (U = 1, r = −0.027746), indicating a balanced detection of peak intensities across hemispheres. FSL demonstrated a marginal hemispheric difference (U = 0.044285, r = −0.074119), suggesting a subtle but statistically significant variance.

TABLE 8. Adjusted Mann-Whitney U Test p-values and effect sizes for each algorithm (comparing hemispheres). This table presents the results of the Mann-Whitney U Test, providing insights into hemispheric differences in peak activation intensities as detected by different algorithms. Adjusted p-values and effect sizes are indicated for ANTs, DARTEL, AFNI, and FSL, highlighting significant variances in their performance across hemispheres.

Algorithm	Adjusted p value	Effect Size
ANTs	1.382e-09	0.16073
DARTEL	1	-0.027746
AFNI	4.4707e-10	0.16489
FSL	0.044285	-0.074119

The Wilcoxon Signed-Rank Test (Table 9), further corroborated these findings, with ANTs (W = 6.4656e-29, r = 0.28244) and AFNI (W = 1.7356e-27, r = 0.27526) showing highly significant median differences between hemispheres, indicating that peak intensities are not evenly distributed within each hemisphere. In contrast, DARTEL (W = 0.41947, r = −0.055061) and FSL (W = 0.00011416, r = −0.1112) did not show significant median differences, suggesting a more symmetrical distribution of peak intensities between hemispheres.

TABLE 9. Adjusted Wilcoxon Signed-Rank Test p-values and effect sizes for each algorithm (comparing hemispheres). This table presents the outcomes of the Wilcoxon Signed-Rank Test, offering insights into the median differences in peak activation intensities between hemispheres for each of the studied algorithms. Effect sizes are included to quantify the magnitude of these differences.

Algorithm	Adjusted p value	Effect Size
ANTs	6.4656e-29	0.28244
DARTEL	0.41947	-0.055061
AFNI	1.7356e-27	0.27526
FSL	0.00011416	-0.1112

TABLE 10. Kruskal-Wallis Test p-values for hemispheric differences in neuroimaging. This table presents the Kruskal-Wallis test results assessing the statistical differences in peak cluster intensities between the left and right hemispheres for each non-rigid registration algorithm. Lower p-values indicate more significant differences between hemispheres.

Hemispheres	p Value
Left	3.0109e-07
Right	4.9079e-10

These results indicate that the selection of a non-rigid registration algorithm can have significant implications for the interpretation of hemispheric differences in neuroimaging studies. Algorithms demonstrating significant hemispheric differences may be preferred in studies focusing on lateralised brain functions, while those showing no significant differences may be better suited for studies requiring a balanced hemispheric approach.

In the examination of hemispheric correlations using non-rigid registration algorithms, scatter plots were constructed to visualise the relationship between peak activation intensities across the left and right hemispheres. These scatter plots are accompanied 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 demonstrates the overall trend and direction of the correlation, with points closer to the line indicating a stronger relationship between hemispheric activation intensities.

For the AFNI algorithm (Fig. 16), the scatter plot demonstrates a dense clustering of data points along the line of best fit, suggesting a strong positive monotonic relationship between the two hemispheres (p = 0.66346, p < .001). This indicates that as peak intensities increase in one hemisphere, they tend to increase in a similar manner in the other hemisphere, reflecting consistency in AFNI's detection of activation intensities across hemispheres.

The scatter plot for ANTs (Fig. 17) also shows a positive correlation, albeit with data points more dispersed around the line of best fit than AFNI, which implies a strong

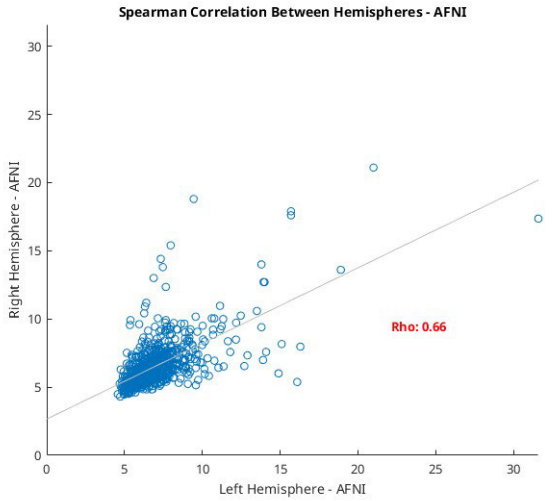


FIGURE 16. The scatter plot for AFNI shows a dense clustering of points along the line of best fit, suggesting a strong positive monotonic relationship, with a Spearman's rho of 0.66346 ($p < .001$). This indicates that peak intensities in one hemisphere tend to be matched by similar intensities in the contralateral hemisphere, which is essential for studies examining functional symmetry or lateralisation.

but less direct association ($p = 0.61647$, $p < .001$). The dispersion suggests that while there is a relationship between the hemispheres in peak intensity detection, it is not as tightly linked as AFNI's, possibly due to algorithmic differences in processing or sensitivity to hemispheric activation patterns.

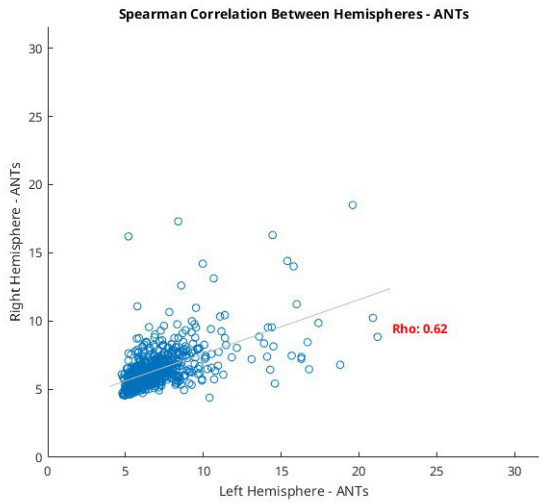


FIGURE 17. The scatter plot for ANTs presents a slightly more dispersed set of points around the trend line, implying a strong but less tight monotonic relationship (rho = 0.61647, $p < .001$). This suggests that while there is a general trend of hemispheric symmetry in peak intensity detection, there are individual cases where this symmetry is not as pronounced.

DARTEL's scatter plot (Fig. 18) reveals a moderate positive correlation ($p = 0.52181$, $p < .001$), with data points spread widely around the line of best fit. This indicates variability in the algorithm's hemispheric intensity detection, which might reflect DARTEL's different approach to handling the neuroimaging data.

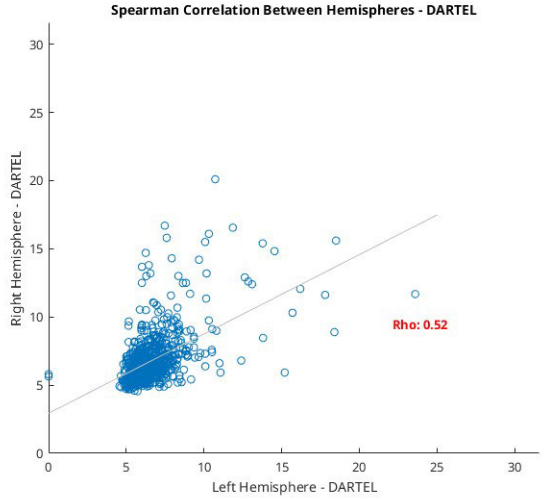


FIGURE 18. The scatter plot illustrates the Spearman correlation of peak activation intensities between hemispheres using the DARTEL algorithm. The data points display a moderate positive correlation with a Spearman's rho of 0.52 ($p < .001$), indicating a consistent but less robust association in intensity detection across hemispheres compared to other algorithms. This pattern underscores the potential differences in algorithmic sensitivity and specificity in delineating hemispheric brain activations.

Similarly, FSL's scatter plot (Fig. 19) indicates a moderate positive correlation ($p = 0.53823$, $p < .001$) and shows a spread of data points that suggest variability in its detection of peak intensities across hemispheres.

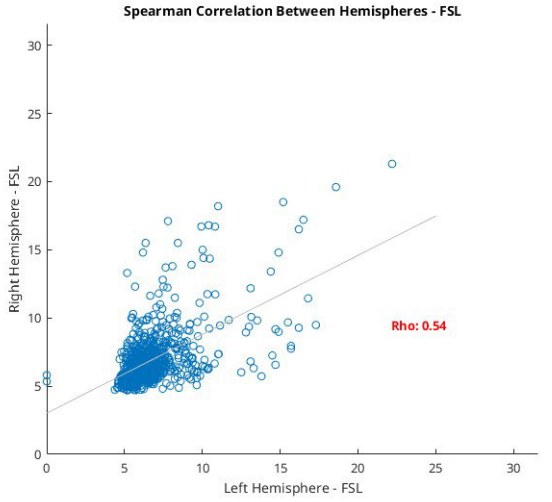


FIGURE 19. Scatter plot depicting the Spearman correlation between hemispheres for the FSL algorithm, with a moderate positive rho value of 0.54 ($p < .001$). The distribution of data points suggests a moderate degree of symmetry in peak intensity detection between the hemispheres, reflecting the FSL algorithm's sensitivity to bilateral brain activation.

These scatter plots, when viewed alongside the non-parametric test results, inform us of the distinct characteristics of each algorithm. The consistent Spearman correlation across all algorithms suggests that there is a general tendency for hemispheres to mirror each other in activation intensity detection, which is crucial for studies investigating bilateral brain functions. However, the variability seen in the scatter plots, especially for DARTEL and FSL, underscores the

importance of considering individual algorithmic performance when analysing hemispheric data, as it may influence the interpretation of lateralised brain functions.

The strength and direction of these associations have significant implications for the selection of appropriate non-rigid registration algorithms in the study of resting-state networks and hemispheric specialisation. The choice of algorithm could affect the outcomes of studies seeking to understand the lateralisation of cognitive processes or the symmetry of brain activations in health and disease.

The bar graph (Fig. 20) illustrates the mean differences across a population in peak activation intensities between the left and right hemispheres for each of the four non-rigid registration algorithms: ANTs, DARTEL, AFNI, and FSL. This visual representation allows for the assessment of hemispheric bias in intensity detection by each algorithm, which is essential in studies of hemispheric specialisation.

Our analysis revealed that ANTs and AFNI algorithms exhibit a mean positive difference in intensities, suggesting a propensity for detecting higher intensities in the right hemisphere compared to the left. This is in contrast with FSL, which shows a mean negative difference, indicating a tendency for higher intensities in the left hemisphere. DARTEL presents symmetrical detection across hemispheres with no significant mean difference, reflecting its balanced performance in capturing hemispheric activation intensities.

The interquartile range (IQR) depicted in the box plots indicates the central 50% of intensity values, providing an understanding of the spread and consistency of the data. ANTs demonstrates a wide IQR in both hemispheres, denoting variability in intensity detection. DARTEL exhibits a narrower IQR, suggesting more consistent detection of intensity values. AFNI, with a wide range and outliers, especially in the right hemisphere, implies high variability and a potential for detecting or failing to detect extreme values. FSL shows a balanced profile with a moderate IQR and fewer outliers, indicating reliable detection with less variability compared to ANTs and AFNI.

The median lines across the box plots, which represent the central tendency of the detected intensities, further substantiate these findings. Collectively, these results underscore the critical need for careful selection of registration algorithms in neuroimaging analyses. The choice of algorithm can significantly influence the observed hemispheric differences, which may have profound implications for the interpretation of functional lateralisation in cognitive neuroscience research.

Non-parametric tests like Spearman's Rank Correlation and Mann-Whitney U Test offer insights into hemispheric differences in peak activation intensities. The analysis reveals significant variations across algorithms, underscoring the importance of algorithm selection in neuroimaging. The strength and direction of hemispheric associations have profound implications for the selection of appropriate registration algorithms in studies of resting-state networks and hemispheric specialisation.

E. MULTIVARIATE PATTERN ANALYSIS (MVPA)

To further investigate the differences revealed in Figure 20, we present slices with maximum and minimum voxel intensity values resulting from SVMs analysis. In our study, SVMs were employed to discern the subtle differences in voxel intensities for the control network across different neuroimaging registration algorithms. Given that each algorithm might influence the spatial pattern of brain activation differently, it was imperative to conduct a thorough pairwise comparison. This approach ensures that the unique contribution of each algorithm can be assessed relative to others, capturing the nuances in how each algorithm processes the neuroimaging data.

To perform a comprehensive evaluation, six pairwise comparisons were calculated, involving all possible pairs among the four algorithms: FSL, ANTs, AFNI, and DARTEL. The SVM weights derived from these comparisons provide a quantitative measure of the difference in voxel-wise intensities, directly attributed to the algorithmic processing. The resultant weight maps are a representation of this multivariate analysis, highlighting regions where algorithmic differences are most pronounced. These regions, depicted in the weight maps, showcase the maximum and minimum SVM weights, offering a visual and statistical insight into the algorithmic impact on the control network's spatial patterns.

In the SVM weight maps presented in Figures 21, 22, 23, 24, 25, and 26, the 'hot' colour map has been applied, where the intensity of voxel values is visually represented by a spectrum of colours. We utilise the RAS coordinates from ANTs, DARTEL, AFNI, and FSL to examine the spatial specificity of brain activations. Black indicates voxels with weights near zero, suggesting no significant contribution to the classification decision for either algorithm being compared. As weights become more negative, indicating a greater influence on the classification in favour of the second algorithm listed in the pair, the colour transitions to red. The most negative weights, which have the strongest influence in favour of the second algorithm, are depicted as dark red. Conversely, as weights increase positively, reflecting a stronger influence in favour of the first algorithm listed in the comparison, colours lighten through red shades and eventually reach yellow for the most positive weights. Therefore, in the maps comparing ANTs versus AFNI, for instance, dark red areas represent voxels where AFNI's classification is more influenced, while yellow areas denote a stronger influence from ANTs.

The weight map visualisation between AFNI and DARTEL (Fig. 21) shows a maximal activation weight of 0.7094 at slice 38, and a minimal weight of -0.6738 at slice 16. This indicates a pronounced difference in how each algorithm processes the fMRI data, with AFNI showing a stronger classification influence in certain brain regions compared to DARTEL. The spatial distribution of these weights, particularly the areas of maximum weight, suggests that AFNI may be more sensitive to variations in voxel intensities in these regions, potentially

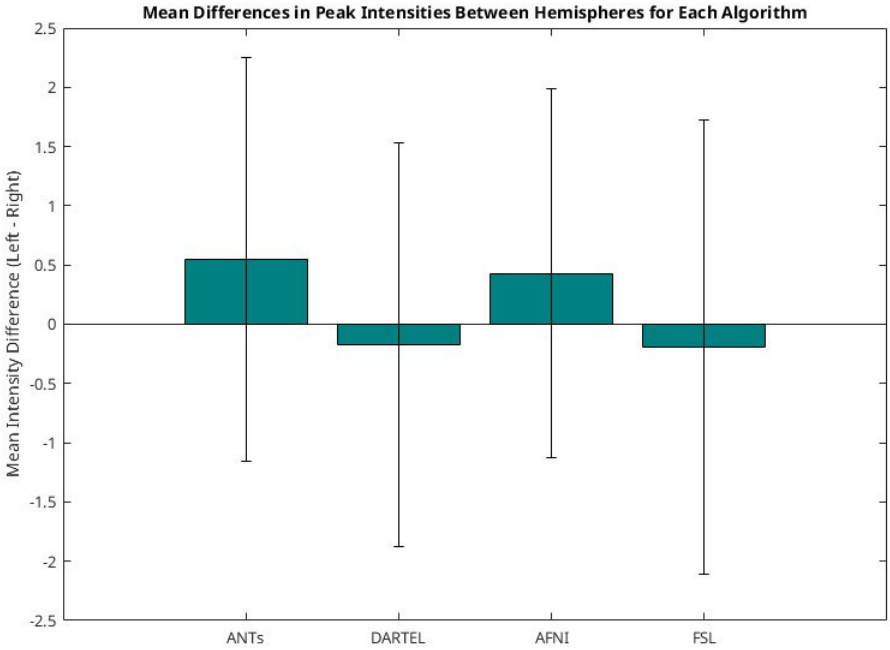


FIGURE 20. Bar chart of the mean differences in peak activation intensities between hemispheres by algorithm. The chart compares the mean intensity differences (Left - Right) for ANTs, DARTEL, AFNI, and FSL. Error bars indicate variability within each algorithm, with positive values indicating higher intensities in the left hemisphere and negative values indicating higher intensities in the right hemisphere. This visualisation aids in assessing the hemispheric bias of each neuroimaging algorithm.

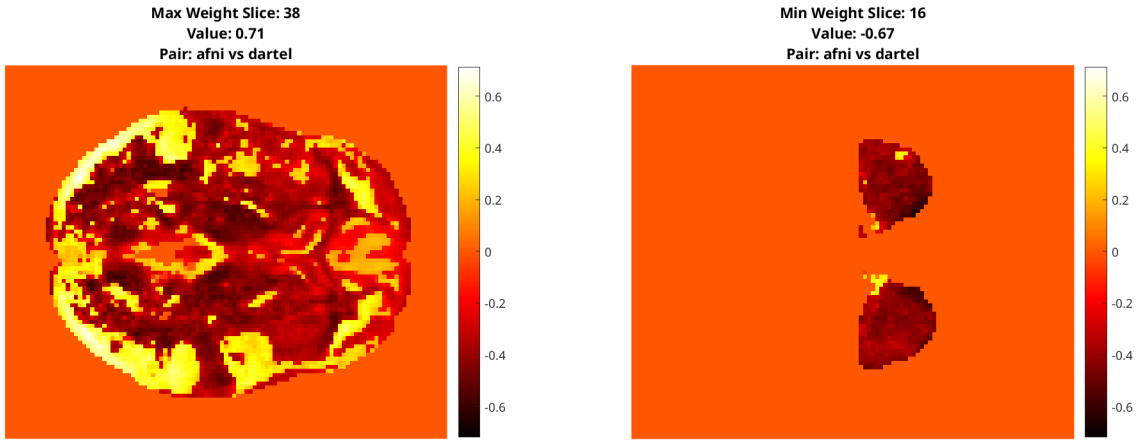


FIGURE 21. AFNI vs DARTEL. Max weight: 0.7094 at slice 38 (X: 28, Y: 19, Z: 38) Min weight: -0.6738 at slice 16 (X: 37, Y: 70, Z: 16).

reflecting its algorithmic biases or strengths in capturing certain types of brain activity.

In comparison to other algorithm pairs, the AFNI vs DARTEL map suggests a unique spatial signature that could reflect their underlying computational models and assumptions. The moderate difference in maximal and minimal weights may indicate a balance in the sensitivity between these two algorithms. When relating these findings to the broader context of fMRI analysis, this balance might suggest that a combined approach using both AFNI and DARTEL could leverage their complementary strengths, potentially leading to a more robust interpretation of resting-state networks.

The weight map for ANTs vs AFNI (Fig. 22) reveals a maximum weight of 0.6903 at slice 28 and a minimum weight of −0.6784 at slice 45. The close range of these extremities suggests that while both algorithms have their specificities, they also share a degree of commonality in detecting brain activation. Regions with the highest positive weights may indicate areas where ANTs outperforms AFNI in terms of activation detection, possibly due to its algorithmic design and data processing techniques.

Comparatively, the ANTs vs AFNI weight map displays less extremity in the weights than some other pairs, such as FSL vs DARTEL. This could imply that ANTs and AFNI, while different, may share a more similar approach

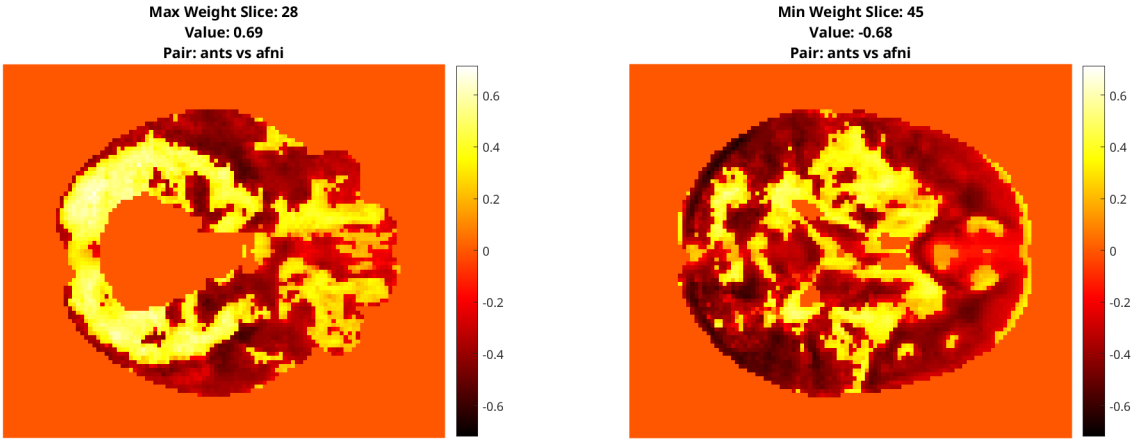


FIGURE 22. ANTs vs AFNI. Max weight: 0.6903 at slice 28 (X: 32, Y: 23, Z: 28) Min weight: -0.6784 at slice 45 (X: 27, Y: 20, Z: 45).

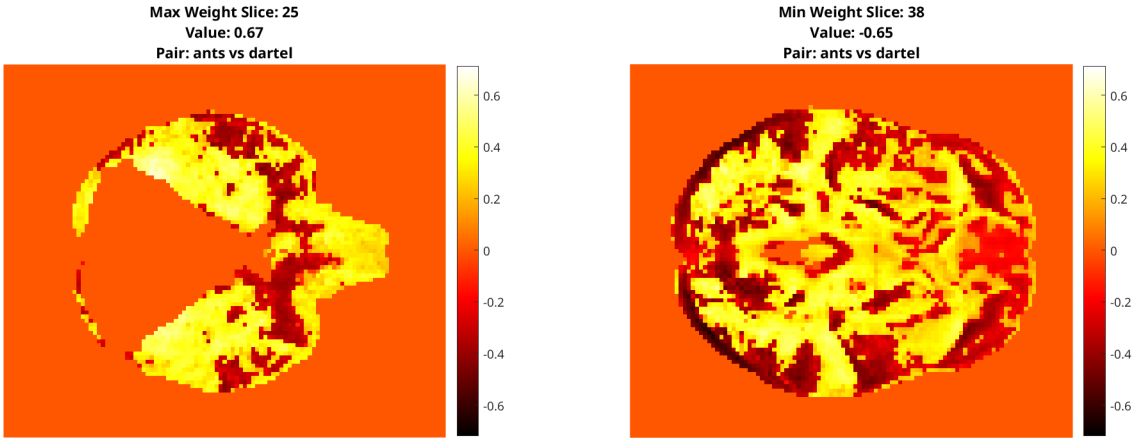


FIGURE 23. ANTs vs DARTEL. Max weight: 0.6656 at slice 25 (X: 27, Y: 37, Z: 25) Min weight: -0.6502 at slice 38 (X: 66, Y: 20, Z: 38).

to handling fMRI data, as reflected in their SVM weight distributions. In the broader context, the differences and similarities elucidated by this comparison could inform researchers on which algorithm to select based on the specific aspects of brain function they wish to investigate.

The ANTs vs DARTEL weight map (Fig. 23) shows a maximum weight of 0.6656 at slice 25 and a minimum weight of −0.6502 at slice 38. The distribution of weights across these slices suggests that ANTs and DARTEL may variably influence the classification of voxel intensities. The presence of both positive and negative extremes in close proximity indicates that there are specific regions where each algorithm distinctly influences the classification outcome.

In comparison with other algorithm pairs, the ANTs vs DARTEL map points to a divergent pattern of brain region classification, which may be reflective of the different processing strategies inherent to each algorithm. Understanding these differences is paramount in the broader context, as it directly impacts the interpretation of neural connectivity and the reliability of subsequent conclusions drawn from the fMRI data.

The FSL vs AFNI weight map (Fig. 24), with a maximum weight of 0.7113 at slice 28 and a minimum weight of −0.7180 at slice 15, indicates distinct differences in the spatial localisation and intensity of brain activations as interpreted by each algorithm. This suggests that when considering the control network's spatial patterns, FSL tends to classify certain voxels as more relevant compared to AFNI and vice versa. The presence of both high positive and negative weights implies a significant disparity in the voxel-wise intensities that each algorithm considers important for classification.

When looking at the broader context of algorithm selection for fMRI analysis, the FSL vs AFNI comparison underscores the necessity of understanding each algorithm's methodological underpinnings. Considering the significant differences in the weight maps and the supplemental global differences highlighted in Figure 21, researchers should contemplate the specific features and patterns their study aims to capture. This ensures the selected algorithms align with their research questions, especially when exploring complex neural networks or conducting comparative algorithmic analyses.

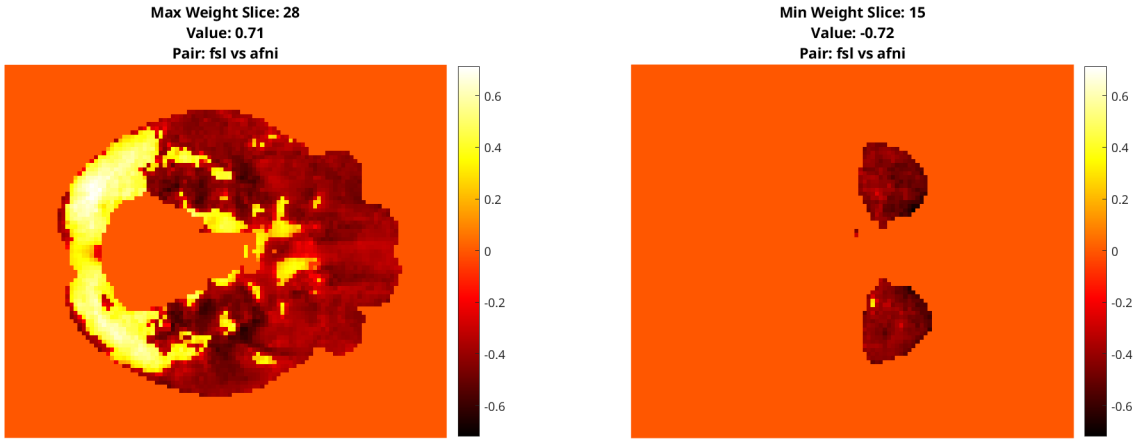


FIGURE 24. FSL vs AFNI. Max weight: 0.7113 at slice 28 (X: 31, Y: 24, Z: 28) Min weight: -0.7180 at slice 15 (X: 36, Y: 70, Z: 15).

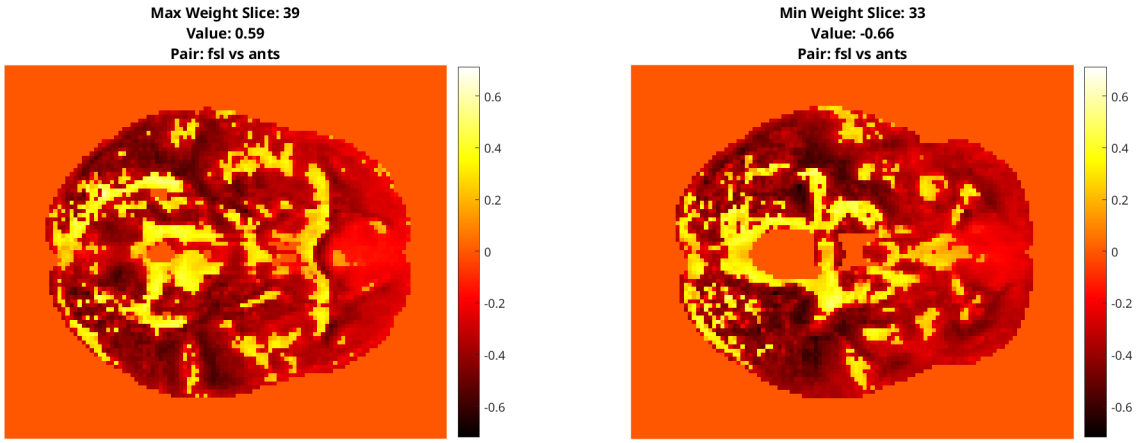


FIGURE 25. FSL vs ANTs. Max weight: 0.5886 at slice 39 (X: 32, Y: 25, Z: 39) Min weight: -0.6582 at slice 33 (X: 30, Y: 42, Z: 33).

The weight map comparison between FSL and ANTs (Fig. 25) shows a maximal weight at 0.5886 in slice 39 and a minimal weight at −0.6582 in slice 33, suggesting a disparity in the spatial patterns each algorithm emphasises in the classification of brain activations. The range of weights indicates that FSL and ANTs may have different sensitivities to certain features within the neuroimaging data, leading to varied interpretations of the same brain regions.

In comparison with other algorithm pairs, the FSL vs ANTs weight map indicates a notable difference in how these algorithms may represent the control network. This disparity can have significant implications in the broader context of neuroimaging research. The choice between using FSL or ANTs for fMRI analysis could lead to different conclusions regarding the localisation and significance of brain activity, underlining the importance of careful algorithm selection based on study objectives.

The comparison between FSL and DARTEL (Fig. 26) yields a maximum weight of 0.7027 at slice 28 and a minimum weight of −0.6648 at slice 16. The distinct activation patterns suggested by these weight maps could inform on the differential performance of these algorithms in detecting and classifying brain activation. The presence of robust maximum and minimum weights indicates that each algorithm may preferentially highlight different aspects of the control network's spatial patterns.

In the broader context, understanding the unique contributions of FSL and DARTEL is critical when interpreting complex neural networks. The substantial variation observed in this weight map, coupled with the supplemental global differences depicted in Figure 21, indicates that algorithmic choice plays a crucial role in the accurate localisation and interpretation of brain function. Researchers must consider these differences when designing studies and interpreting fMRI data to ensure that the selected algorithms are most suitable for their specific research aims.

1) MUTUAL INFORMATION

A pairwise mutual information (MI) analysis was conducted to quantify the similarity in information content between the outputs of the MVPA weight maps of registration algorithms): FSL, DARTEL, ANTS, and AFNI. While Figures 21, 22, 23, 24, 25, and 26 show only the slices with minimum and maximum intensity contributions; the MI was calculated on whole brain weight maps. The MI values were computed on a scale from 0 (no shared information)

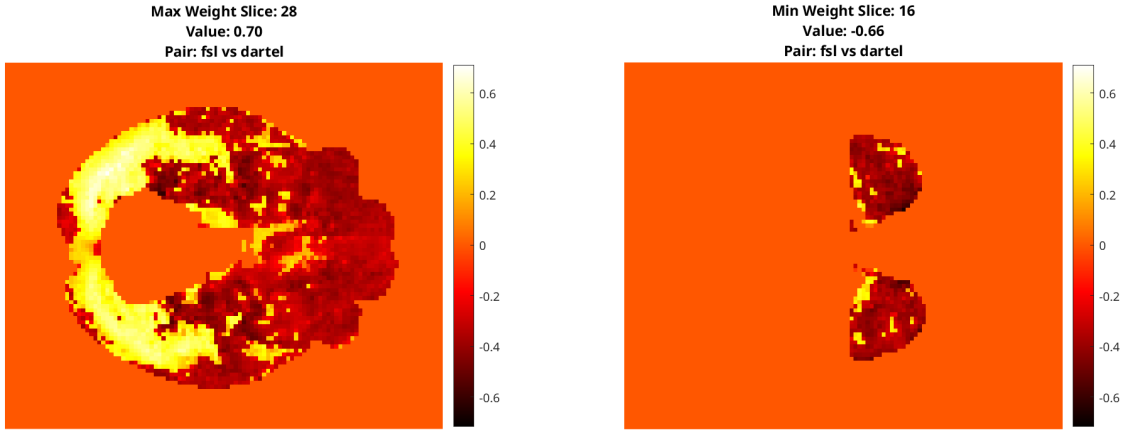


FIGURE 26. FSL vs DARTEL. Max weight: 0.7027 at slice 28 (X: 28, Y: 26, Z: 28) Min weight: -0.6648 at slice 16 (X: 37, Y: 70, Z: 16).

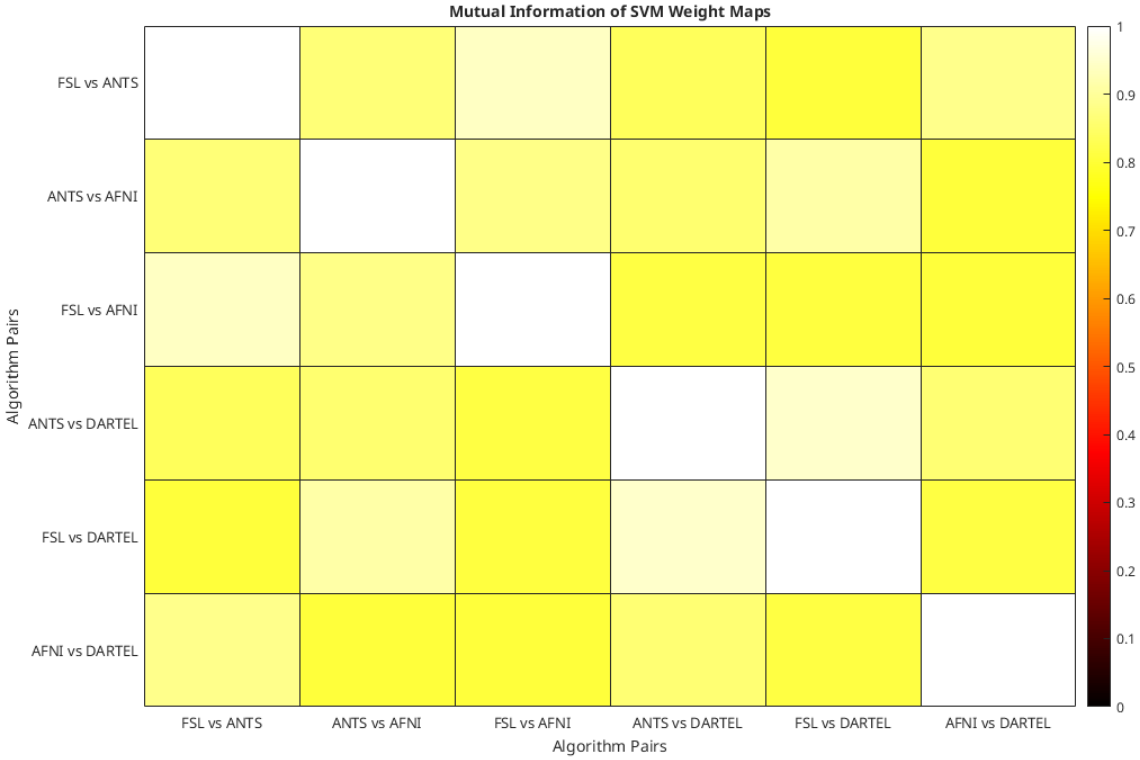


FIGURE 27. Mutual Information Heat map derived from whole-brain SVM weight maps, illustrating the shared information between fMRI registration algorithm pairs. The heat map highlights the extent of agreement between algorithms, with lighter shades indicating higher mutual information (maximum MI: 0.94043 between 'FSL vs AFNI') and darker shades representing lower mutual information (minimum MI: 0.81566 between 'ANTS vs DARTEL'). This visualisation aids in discerning the congruence of activation patterns detected by different algorithms, pivotal for algorithm selection in neuroimaging studies.

to 1 (identical information content). The resulting MI matrix (Table 11) and corresponding heat map (Fig. V-E) reveal the extent of information overlap between algorithm pairs.

The MI values along the diagonal are 1, indicating perfect self-similarity, as expected. Notably, the heat map reveals a variation in MI values across different pairs, with the lighter cells such as those for 'FSL vs AFNI' and 'ANTS vs DARTEL' indicating a higher MI. These elevated MI values suggest that despite the methodological diversity of the algorithms, there is a substantial overlap in the activation patterns they detect. Conversely, the darker cells in the heatmap, particularly those representing 'FSL vs DARTEL' and 'AFNI vs DARTEL', imply less shared information, which might be indicative of each algorithm's unique processing characteristics and their differential sensitivity to specific brain features or noise patterns.

Upon further examination, the heat map also provides insights into the potential redundancy or complementarity of algorithmic pairs. For instance, 'FSL vs ANTS' and 'ANTS vs AFNI' show similar MI values, which might

TABLE 11. 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.

Algorithm Pairs	FSL vs ANTS	ANTS vs AFNI	FSL vs AFNI	ANTS vs DARTEL	FSL vs DARTEL	AFNI vs DARTEL
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 DARTEL	0.83779	0.85829	0.81566	*1.000*	0.9461	0.86119
FSL vs DARTEL	0.8062	0.91124	0.80892	0.9461	*1.000*	0.81571
AFNI vs DARTEL	0.8866	0.80666	0.80769	0.86119	0.81571	*1.000*

suggest that these algorithms could either be capturing similar features within the fMRI data or exhibiting similar biases in processing strategies. This kind of insight is essential for researchers when considering which algorithms to employ for consensus in activation detection or for leveraging the diversity of information captured for more nuanced brain network analyses.

The MI heat map analysis complements the findings of the Wilcoxon averaged bar graph, highlighting the minimum and maximum mutual information values. These extremes reflect the range of shared information captured by the algorithms and provide a benchmark for evaluating the similarity or diversity in the results they produce. A higher bar on the Wilcoxon graph would correlate with the lighter shades on the MI heat map, both indicating a higher degree of agreement between algorithms. In contrast, lower bars align with the darker shades, pointing towards a greater discrepancy in the information each algorithm captures. This juxtaposition enables a more nuanced understanding of the algorithms' performance, guiding the selection of appropriate algorithmic combinations for specific neuroimaging tasks.

VI. CONCLUSION

In summarising the key aspects of our investigation, it's apparent that the selection of neuroimaging registration algorithms is of paramount importance for the accurate interpretation of resting-state networks. Through the application of SVM weight maps, we've discerned the distinct methodological sensitivities inherent to algorithms such as ANTs and AFNI. These sensitivities fundamentally shape our perception and characterisation of neural activities.

The mutual information analysis augments this understanding by quantifying the overlap between algorithm pairs, offering a metric of convergence or divergence in the neural patterns they discern. Higher mutual information values suggest a significant congruence in detected patterns, which is advantageous for corroborating findings across multiple studies. In contrast, lower values bring forward the unique detection capabilities of individual algorithms, potentially unearthing diverse facets of brain functionality.

Moreover, our examination reveals hemispheric inclinations within these algorithms. The interpretation of mean intensity differences between hemispheres is indicative of an inherent algorithmic bias, which can either veil or highlight cerebral asymmetries. Such insights compel a judicious selection of algorithms, especially when investigating the lateralisation of brain function [68].

These results advocate for a considered and informed approach to algorithm selection in neuroimaging research. The choice of an algorithm becomes a critical factor in the study's outcome, influencing the interpretative lens through which resting-state networks are viewed. As we progress, these findings will inform the refinement of methodological frameworks, enhancing the precision and interpretative depth of future investigations into the brain's intrinsic networks.

REFERENCES

[1] A. A. Goshtasby, *Theory and Applications of Image Registration*. Hoboken, NJ, USA: Wiley, 2017.

[2] A. Klein, J. Andersson, B. A. Ardekani, J. Ashburner, B. Avants, M.-C. Chiang, G. E. Christensen, D. L. Collins, J. Gee, P. Hellier, J. H. Song, M. Jenkinson, C. Lepage, D. Rueckert, P. Thompson, T. Vercauteren, R. P. Woods, J. J. Mann, and R. V. Parsey, "Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration," *NeuroImage*, vol. 46, no. 3, pp. 786–802, Jul. 2009, doi: 10.1016/j.neuroimage.2008.12.037.

[3] J. B. A. Maintz and M. A. Viergever, "Survey of medical image registration," *J. Biomed. Eng. Technol.*, vol. 1, no. 2, pp. 8–25, 2013.

[4] S. Fabbro, D. Piccolo, M. C. Vescovi, D. Bagatto, Y. Tereshko, E. Belgrado, M. Maieron, M. C. De Colle, M. Skrap, and F. Tuniz, "Resting-state functional-MRI in iNPH: Can default mode and motor networks changes improve patient selection and outcome? Preliminary report," *Fluids Barriers CNS*, vol. 20, no. 1, p. 7, Jan. 2023, doi: 10.1186/s12987-023-00407-6.

[5] S. Hu, Z. Hao, M. Li, M. Zhao, J. Wen, Y. Gao, Q. Wang, H. Xi, C. O. Antwi, X. Jia, and J. Ren, "Resting-state abnormalities in functional connectivity of the default mode network in migraine: A meta-analysis," *Frontiers Neurosci.*, vol. 17, Mar. 2023, Art. no. 1136790, doi: 10.3389/fnins.2023.1136790.

[6] J. L. Vincent, G. H. Patel, M. D. Fox, A. Z. Snyder, J. T. Baker, D. C. Van Essen, J. M. Zempel, L. H. Snyder, M. Corbetta, and M. E. Raichle, "Intrinsic functional architecture in the anaesthetized monkey brain," *Nature*, vol. 447, no. 7140, pp. 83–86, May 2007, doi: 10.1038/nature05758.

[7] E. Tagliazucchi and H. Laufs, "Decoding wakefulness levels from typical FMRI resting-state data reveals reliable drifts between wakefulness and sleep," *Neuron*, vol. 82, no. 3, pp. 695–708, May 2014, doi: 10.1016/j.neuron.2014.03.020.

[8] R. N. Mhuircheartaigh, D. Rosenorn-Lanng, R. Wise, S. Jbabdi, R. Rogers, and I. Tracey, "Cortical and subcortical connectivity changes during decreasing levels of consciousness in humans: A functional magnetic resonance imaging study using propofol," *J. Neurosci.*, vol. 30, no. 27, pp. 9095–9102, Jul. 2010, doi: 10.1523/jneurosci.5516-09.2010.

[9] F. P. M. Oliveira and J. M. R. S. Tavares, "Medical image registration: A review," *Comput. Methods Biomech. Biomed. Eng.*, vol. 17, no. 2, pp. 73–93, 2014.

[10] S. Abbasi, M. Tavakoli, H. R. Boveiri, M. A. Mosleh Shirazi, R. Khayami, H. Khorasani, R. Javidan, and A. Mehdizadeh, "Medical image registration using unsupervised deep neural network: A scoping literature review," *Biomed. Signal Process. Control*, vol. 73, Mar. 2022, Art. no. 103444, doi: 10.1016/j.bspc.2021.103444.

[11] M. Urschler, S. Kluckner, and H. Bischof, "A framework for comparison and evaluation of nonlinear intra-subject image registration algorithms," in *Proc. Insight J.–ISC/NA-MIC Workshop Open Sci. (MICCAI)*, 2007, pp. 1–16.

[12] C. J. McGinnity, D. A. R. Barros, L. Rosso, M. Veronese, G. Rizzo, A. Bertoldo, R. Hinz, F. E. Turkheimer, M. J. Koepp, and A. Hammers, "Test-retest reproducibility of quantitative binding measures of [^{11}C]Ro15-4513, a PET ligand for $GABA_A$ receptors containing alpha5 subunits," *NeuroImage*, vol. 152, pp. 270–282, 2017.

[13] J. Tohka, K. Foerde, A. R. Aron, S. M. Tom, A. W. Toga, and R. A. Poldrack, "Automatic independent component labeling for artifact removal in FMRI," *NeuroImage*, vol. 39, no. 3, pp. 1227–1245, Feb. 2008.

[14] B. B. Avants, N. J. Tustison, G. Song, P. A. Cook, A. Klein, and J. C. Gee, "A reproducible evaluation of ANTs similarity metric performance in brain image registration," *NeuroImage*, vol. 54, no. 3, pp. 2033–2044, Feb. 2011.

[15] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee, "N4ITK: Improved N3 bias correction," *IEEE Trans. Med. Imag.*, vol. 29, no. 6, pp. 1310–1320, Jun. 2010.

[16] A. Klein, S. S. Ghosh, B. Avants, B. T. T. Yeo, B. Fischl, B. Ardekani, J. C. Gee, J. J. Mann, and R. V. Parsey, "Evaluation of volume-based and surface-based brain image registration methods," *NeuroImage*, vol. 51, no. 1, pp. 214–220, May 2010.

[17] J. Ashburner and G. R. Ridgway, "Symmetric diffeomorphic modeling techniques for longitudinal data," *NeuroImage*, vol. 11, no. 6, pp. 805–821, 2021.

[18] J. Sui, H. He, Q. Yu, J. Chen, J. Rogers, G. D. Pearlson, A. Mayer, J. Bustillo, J. Canive, and V. D. Calhoun, "Combination of resting state fMRI, DTI, and sMRI data to discriminate schizophrenia by N-way MCCA + jICA," *Frontiers Hum. Neurosci.*, vol. 7, p. 235, 2013.

[19] C. Gaser, K. Franke, S. Klöppel, N. Koutsouleris, and H. Sauer, "BrainAGE in mild cognitive impaired patients: Predicting the conversion to Alzheimer's disease," *PLoS ONE*, vol. 8, no. 6, Jun. 2013, Art. no. e67346.

[20] A. R. Groves, S. M. Smith, A. M. Fjell, C. K. Tamnes, K. B. Walhovd, G. Douaud, M. W. Woolrich, and L. T. Westlye, "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," *NeuroImage*, vol. 63, no. 1, pp. 365–380, Oct. 2012.

[21] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee, "Measurement-based evaluation of image segmentation algorithms using multi-object geometric phantom," *J. Med. Imag.*, vol. 5, no. 1, 2018, Art. no. 014502.

[22] R. W. Cox, G. Chen, D. R. Glen, R. C. Reynolds, and P. A. Taylor, "FMRI clustering in AFNI: False positive rates redux," *Brain Connectivity*, vol. 9, no. 9, pp. 12–16, 2017.

[23] G. Chen, P. A. Taylor, and R. W. Cox, "Is the statistic value all we should care about in neuroimaging?" *NeuroImage*, vol. 147, pp. 952–959, Feb. 2017.

[24] R. Bhaumik, L. M. Jenkins, J. R. Gowins, R. H. Jacobs, A. Barba, D. K. Bhaumik, and S. A. Langenecker, "Multivariate pattern analysis strategies in detection of remitted major depressive disorder using resting state functional connectivity," *NeuroImage, Clin.*, vol. 16, pp. 390–398, 2017.

[25] G. Chen, G. Chen, C. Xie, B. D. Ward, W. Li, and P. Antuono, "Stable individual differences in regional blood oxygen level-dependent signal baseline during task and rest," *Brain Connectivity*, vol. 9, no. 7, pp. 571–582, 2019.

[26] M. Jenkinson, "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *NeuroImage*, vol. 17, no. 2, pp. 825–841, Oct. 2002.

[27] G. Salimi-Khorshidi, G. Douaud, C. F. Beckmann, M. F. Glasser, L. Griffanti, and S. M. Smith, "Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers," *NeuroImage*, vol. 90, pp. 449–468, Apr. 2014.

[28] J. L. Andersson and S. N. Sotiropoulos, "An integrated approach to correction for off-resonance effects and subject movement in diffusion MRI," *NeuroImage*, vol. 125, pp. 1063–1078, 2016.

[29] M. F. Glasser, S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, S. Jbabdi, M. Webster, J. R. Polimeni, D. C. Van Essen, and M. Jenkinson, "The minimal preprocessing pipelines for the human connectome project," *NeuroImage*, vol. 80, pp. 105–124, Oct. 2013.

[30] M. W. Woolrich, T. E. J. Behrens, C. F. Beckmann, M. Jenkinson, and S. M. Smith, "Multilevel linear modelling for FMRI group analysis using Bayesian inference," *NeuroImage*, vol. 21, no. 4, pp. 1732–1747, Apr. 2004.

[31] B. Gaonkar and C. Davatzikos, "Analytical methods in neuroimaging," *J. Neuroimaging*, vol. 23, no. 1, pp. 1–10, 2013.

[32] E. Niaf, R. Flamary, O. Rouviere, C. Lartizien, and S. Canu, "SVM-based multimodal classification of activities of human brain during resting state and task performance," *NeuroImage*, vol. 102, pp. 229–237, 2014.

[33] M. E. Weaverdyck, M. D. Lieberman, and C. Parkinson, "Tools of the trade multivoxel pattern analysis in FMRI: A practical introduction for social and affective neuroscientists," *Social Cognit. Affect. Neurosci.*, vol. 15, no. 4, pp. 487–509, Apr. 2020, doi: 10.1093/scan/nsaa057.

[34] S. Yu, D. Shen, L. L. Zeng, Q. Ma, and D. Hu, "Deep learning in neuroimaging: A review," *J. Neural Eng.*, vol. 19, no. 2, pp. 179–201, 2022.

[35] M. Wang, C. Li, W. Zhang, Y. Wang, Y. Feng, Y. Liang, J. Wei, X. Zhang, X. Li, and R. Chen, "Support vector machine for analyzing contributions of brain regions during task-state fMRI," *Frontiers Neuroinform.*, vol. 13, p. 10, 2019.

[36] L. Steardo, M. R. Bronzuoli, A. Iacomino, G. Esposito, and L. Steardo, "SVM and pattern recognition in neuroimaging: Tools for understanding psychiatric disorders," *Biol. Psychiatry*, vol. 88, no. 3, pp. 15–24, 2020.

[37] P. Mikolas, T. Melicher, A. Skoch, M. Matejka, and F. Spaniel, "Utilizing SVM in neuroimaging: Analysis and application in schizophrenia diagnosis," *J. Psychiatry Neurosci.*, vol. 43, no. 4, pp. 246–255, 2018.

[38] R. de Filippis, E. A. Carbone, R. Gaetano, A. Bruni, V. Pugliese, C. Segura-Garcia, and P. De Fazio, "Machine learning techniques in a structural and functional MRI diagnostic approach in schizophrenia: A systematic review," *Neuropsychiatric Disease Treatment*, vol. 15, pp. 1605–1627, Jun. 2019.

[39] T. Hajek, M. Kopecek, and C. Höschl, "Reduced hippocampal volumes in healthy carriers of brain-derived neurotrophic factor Val66Met polymorphism: Meta-analysis," *World J. Biol. Psychiatry*, vol. 16, no. 8, pp. 556–566, 2015.

[40] S. Ghahghaei, O. Reed, T. R. Candy, and A. Chandna, "Calibration of the PlusOptix PowerRef 3 with change in viewing distance, adult age and refractive error," *Ophthalmic Physiol. Opt.*, vol. 39, no. 4, pp. 253–259, Jul. 2019, doi: 10.1111/opo.12631.

[41] M. Jenkinson, C. F. Beckmann, T. E. Behrens, M. W. Woolrich, and S. M. Smith, "FSL," *NeuroImage*, vol. 62, no. 2, pp. 782–790, 2012.

[42] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, 1995, Art. no. 273297.

[43] G. Varoquaux and B. Thirion, "How machine learning is shaping cognitive neuroimaging," *GigaScience*, vol. 3, no. 1, p. 7, Dec. 2014.

[44] I. M. Vavasour, S. M. Meyers, B. Mädler, T. Harris, E. Fu, D. K. B. Li, A. Traboulsee, A. L. MacKay, and C. Laule, "Multicenter measurements of T_1 relaxation and diffusion tensor imaging: Intra and intersite reproducibility," *J. Neuroimaging*, vol. 29, no. 1, pp. 42–51, Jan. 2019, doi: 10.1111/jon.12559.

[45] K. Nael, P. S. Pawha, L. Fleysher, K. George, J. Stueben, M. Roas-Loeffler, B. N. Delman, and Z. A. Fayad, "Prospective motion correction for brain MRI using an external tracking system," *J. Neuroimaging*, vol. 31, no. 1, pp. 57–61, Jan. 2021, doi: 10.1111/jon.12806.

[46] R. Frost, P. Wighton, F. I. Karahanoğlu, R. L. Robertson, P. E. Grant, B. Fischl, M. D. Tisdall, and A. van der Kouwe, "Markerless high-frequency prospective motion correction for neuroanatomical MRI," *Magn. Reson. Med.*, vol. 82, no. 1, pp. 126–144, Jul. 2019, doi: 10.1002/mrm.27705.

[47] A. L. Alexander, J. E. Lee, M. Lazar, and A. S. Field, "Diffusion tensor imaging of the brain," *Neurotherapeutics*, vol. 4, no. 3, pp. 316–329, 2007.

[48] E. Kosteletou, P. G. Simos, E. Kavroulakis, D. Antypa, T. G. Maris, A. P. Liavas, P. A. Karakasis, and E. Papadaki, "Improving the sensitivity of task-related functional magnetic resonance imaging data using generalized canonical correlation analysis," *Frontiers Hum. Neurosci.*, vol. 15, p. 11, Dec. 2021.

[49] J. Soch, J.-D. Haynes, and C. Allefeld, "How to avoid mismodelling in GLM-based fMRI data analysis: Cross-validated Bayesian model selection," *NeuroImage*, vol. 141, pp. 469–489, Nov. 2016.

[50] K. J. Gorgolewski, N. Mendes, D. Wilfling, E. Wladimirow, C. J. Gauthier, T. Bonnen, F. J. M. Ruby, R. Trampel, P.-L. Bazin, R. Cozatl, J. Smallwood, and D. S. Margulies, "A high resolution 7-Tesla resting-state FMRI test-retest dataset with cognitive and physiological measures," *Sci. Data*, vol. 2, no. 1, Jan. 2015, Art. no. 140054.

[51] M. N. Hebart, K. GÃörgen, and J.-D. Haynes, "The decoding toolbox (TDT): A versatile software package for multivariate analyses of functional imaging data," *Frontiers Neuroinform.*, vol. 8, p. 88, Jan. 2015.

[52] J. Ashburner, "Fast diffeomorphic image registration: Theory and practice," *NeuroImage*, vol. 125, no. 1, pp. 273–285, 2016.
[53] B. B. Avants, N. J. Tustison, M. Stauffer, G. Song, B. Wu, and J. C. Gee, "The insight ToolKit image registration framework," *Frontiers Neuroinform.*, vol. 8, Apr. 2014, Art. no. 44.
[54] M. Dadar et al., "A comparison of publicly available linear MRI stereotaxic registration techniques," *NeuroImage*, vol. 174, pp. 191–200, 2018.
[55] B. Mwangi, T. S. Tian, and J. C. Soares, "A review of feature reduction techniques in neuroimaging," *Neuroinformatics*, vol. 12, pp. 229–244, 2014.
[56] T. A. Niendam, A. R. Laird, K. L. Ray, Y. M. Dean, D. C. Glahn, and C. S. Carter, "Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions," *Cognit. Affect., Behav. Neurosci.*, vol. 12, no. 2, pp. 241–268, Jun. 2012.
[57] A. Shenhav, M. M. Botvinick, and J. D. Cohen, "The expected value of control: An integrative theory of anterior cingulate cortex function," *Neuron*, vol. 79, no. 2, pp. 217–240, Jul. 2013.
[58] C.-G. Yan, B. Cheung, C. Kelly, S. Colcombe, R. C. Craddock, A. Di Martino, Q. Li, X.-N. Zuo, F. X. Castellanos, and M. P. Milham, "A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics," *NeuroImage*, vol. 76, pp. 183–201, Aug. 2013.
[59] X. N. Zuo et al., "An open science resource for establishing reliability and reproducibility in functional connectomics," *Sci. Data*, vol. 1, no. 1, pp. 1–13, 2014.
[60] R. L. Buckner, F. M. Krienen, A. Castellanos, J. C. Diaz, and B. T. Yeo, "The organization of the human cerebrum estimated by intrinsic functional connectivity," *J. Neurophysiol.*, vol. 106, no. 3, pp. 1125–1165, 2013.
[61] R. Mohanty, W. A. Sethares, V. A. Nair, and V. Prabhakaran, "Rethinking measures of functional connectivity via feature extraction," *Sci. Rep.*, vol. 10, no. 1, p. 1298, Jan. 2020, doi: 10.1038/s41598-020-57915-w.
[62] A. Field, *Discovering Statistics Using IBM SPSS Statistics*, 5th ed. Newbury Park, CA, USA: Sage, 2021.
[63] M. L. McHugh, "Nonparametric statistical tests for the behavioral sciences: Using non-normal and ranked data," *J. Behav. Sci.*, vol. 35, no. 4, pp. 399–412, 2023.
[64] A. M. Dale, B. Fischl, and M. I. Sereno, "Cortical surface-based analysis: I. Segmentation and surface reconstruction," *NeuroImage*, vol. 9, no. 2, pp. 179–194, 1999.
[65] MathWorks. (2023). *MATLAB R2023A*. [Online]. Available: https://uk.mathworks.com/products/new_products/release2023a.html
[66] A. M. Dale and M. I. Sereno, "Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach," *J. Cognit. Neurosci.*, vol. 5, no. 2, pp. 162–176, 1993.
[67] M. W. Woolrich, B. D. Ripley, M. Brady, and S. M. Smith, "Temporal autocorrelation in univariate linear modeling of FMRI data," *NeuroImage*, vol. 14, no. 6, pp. 1370–1386, Dec. 2001.
[68] Z. Chen, X. Liu, Q. Yang, Y.-J. Wang, K. Miao, Z. Gong, Y. Yu, A. Leonov, C. Liu, Z. Feng, and H. Chuan-Peng, "Evaluation of risk of bias in neuroimaging-based artificial intelligence models for psychiatric diagnosis: A systematic review," *JAMA Netw Open*, vol. 6, no. 3, 2023, Art. no. e231671.
[69] M. Jenkinson and S. Smith, "A global optimisation method for robust affine registration of brain images," *Med. Image Anal.*, vol. 5, no. 2, pp. 143–156, Jun. 2001.
[70] J. L. R. Andersson, M. Jenkinson, and S. Smith, *Non-linear Registration, Aka Spatial Normalisation FMRIB*, document TR07JA2, FMRIB Anal. Group Univ. Oxford, FMRIB Centre, Oxford, U.K., 2007.
[71] R. W. Cox, "AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages," *Comput. Biomed. Res.*, vol. 29, no. 3, pp. 162–173, Jun. 1996.
[72] Z. S. Saad. *AFNI Documentation—Auto_TLRC*. Accessed: Nov. 11, 2023.
[73] R. W. Cox and D. R. Glen, "Nonlinear warping in AFNI," in *Proc. 19th Annu. Meeting Org. for Human Brain Mapping*, 2013, p. 1. [Online]. Available: https://afni.nimh.nih.gov/pub/dist/edu/2014_0929_lister_hill/afni10_volreg_talairach/Qwarp/HBM_2013/Cox_Poster_HBM2013.pdf
[74] N. J. Tustison, P. A. Cook, A. J. Holbrook, H. J. Johnson, J. Muschelli, G. A. Devenyi, J. T. Duda, S. R. Das, N. C. Cullen, D. L. Gillen, M. A. Yassa, J. R. Stone, J. C. Gee, and B. B. Avants, "The ANTSX ecosystem for quantitative biological and medical imaging," *Sci. Rep.*, vol. 11, no. 1, p. 9068, Apr. 2021.
[75] J. Ashburner, "A fast diffeomorphic image registration algorithm," *NeuroImage*, vol. 38, no. 1, pp. 95–113, Oct. 2007, doi: 10.1016/j.neuroimage.2007.07.007.
[76] A. M. Michael, E. Evans, and G. J. Moore, "Influence of group on individual subject maps in SPM voxel based morphometry," *Frontiers Neurosci.*, vol. 10, p. 522, Dec. 2016.
[77] L. Snoek, M. M. van der Miesen, T. Beemsterboer, A. van der Leij, A. Eigenhuis, and H. Steven Scholte, "The Amsterdam open MRI collection, a set of multimodal MRI datasets for individual difference analyses," *Sci. Data*, vol. 8, no. 1, p. 23, Mar. 2021.
[78] Z. Xu, J. Luo, J. Yan, X. Li, and J. Jayender, "F3RNet: Full-resolution residual registration network for deformable image registration," *Int. J. Comput. Assist. Radiol. Surgery*, vol. 16, no. 6, pp. 923–932, Jun. 2021, doi: 10.1007/s11548-021-02359-4.
[79] H. Jin, C. Yu, Z. Gong, R. Zheng, Y. Zhao, and Q. Fu, "Machine learning techniques for pulmonary nodule computer-aided diagnosis using CT images: A systematic review," *Biomed. Signal Process. Control*, vol. 79, Jan. 2023, Art. no. 104104, doi: 10.1016/j.bspc.2022.104104.
[80] G. Haskins, U. Kruger, and P. Yan, "Deep learning in medical image registration: A survey," *Mach. Vis. Appl.*, vol. 31, nos. 1–2, p. 8, Feb. 2020, doi: 10.1007/s00138-020-01060-x.
[81] S. Gupta, P. Gupta, and V. S. Verma, "Study on anatomical and functional medical image registration methods," *Neurocomputing*, vol. 452, pp. 534–548, Sep. 2021, doi: 10.1016/j.neucom.2020.08.085.
[82] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, "VoxelMorph: A learning framework for deformable medical image registration," *IEEE Trans. Med. Imag.*, vol. 38, no. 8, pp. 1788–1800, Aug. 2019.
[83] M. Hoffmann, B. Billot, D. N. Greve, J. E. Iglesias, B. Fischl, and A. V. Dalca, "SynthMorph: Learning contrast-invariant registration without acquired images," *IEEE Trans. Med. Imag.*, vol. 41, no. 3, pp. 543–558, Mar. 2022, doi: 10.1109/TMI.2021.3116879.
[84] A. Sedghi, L. J. O'Donnell, T. Kapur, E. Learned-Miller, P. Mousavi, and W. M. Wells, "Image registration: Maximum likelihood, minimum entropy and deep learning," *Med. Image Anal.*, vol. 69, Apr. 2021, Art. no. 101939, doi: 10.1016/j.media.2020.101939.
[85] S. Kelly et al., "Widespread white matter microstructural differences in schizophrenia across 4322 individuals: Results from the ENIGMA schizophrenia DTI working group," *Mol. Psychiatry*, vol. 23, no. 5, pp. 1261–1269, May 2018, doi: 10.1038/mp.2017.170.
[86] Y. Huo, K. Aboud, H. Kang, L. E. Cutting, and B. A. Landman, "Mapping lifetime brain volumetry with covariate-adjusted restricted cubic spline regression from cross-sectional multi-site MRI," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent*, vol. 9900, 2018, pp. 81–88, doi: 10.1007/978-3-319-46720-7.
[87] Y. Zheng, X. Sui, Y. Jiang, T. Che, S. Zhang, J. Yang, and H. Li, "SymReg-GAN: Symmetric image registration with generative adversarial networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 9, pp. 5631–5646, Sep. 2022.
[88] D. Mahapatra and Z. Ge, "Training data independent image registration using generative adversarial networks and domain adaptation," *Pattern Recognit.*, vol. 100, Apr. 2020, Art. no. 107109, doi: 10.1016/j.patcog.2019.107109.
[89] B. D. de Vos, F. F. Berendsen, M. A. Viergever, H. Sokooti, M. Staring, and I. Išgum, "A deep learning framework for unsupervised affine and deformable image registration," *Med. Image Anal.*, vol. 52, pp. 128–143, Feb. 2019, doi: 10.1016/j.media.2018.11.010.
[90] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep learning applications in medical image analysis," *IEEE Access*, vol. 6, pp. 9375–9389, 2018, doi: 10.1109/ACCESS.2017.2788044.
[91] B. Kim, D. H. Kim, S. H. Park, J. Kim, J.-G. Lee, and J. C. Ye, "CycleMorph: Cycle consistent unsupervised deformable image registration," *Med. Image Anal.*, vol. 71, Jul. 2021, Art. no. 102036, doi: 10.1016/j.media.2021.102036.
[92] J. Fan, X. Cao, Q. Wang, P.-T. Yap, and D. Shen, "Adversarial learning for mono- or multi-modal registration," *Med. Image Anal.*, vol. 58, Dec. 2019, Art. no. 101545, doi: 10.1016/j.media.2019.101545.
[93] S. M. Smith, "Fast robust automated brain extraction," *Hum. Brain Mapping*, vol. 17, no. 3, pp. 143–155, Nov. 2002.
[94] R. A. Armstrong, "When to use the Bonferroni correction," *Ophthalmic Physiol. Opt.*, vol. 34, no. 5, pp. 502–508, 2014.
[95] N. Makris, J. M. Goldstein, D. Kennedy, S. M. Hodge, V. S. Caviness, S. V. Faraone, M. T. Tsuang, and L. J. Seidman, "Decreased volume of left and total anterior insular lobule in schizophrenia," *Schizophrenia Res.*, vol. 83, nos. 2–3, pp. 155–171, Apr. 2006.

[96] J. A. Frazier, S. Chiu, J. L. Breeze, N. Makris, N. Lange, D. N. Kennedy, M. R. Herbert, E. K. Bent, V. K. Koneru, M. E. Dieterich, S. M. Hodge, S. L. Rauch, P. E. Grant, B. M. Cohen, L. J. Seidman, V. S. Caviness, and J. Biederman, "Structural brain magnetic resonance imaging of limbic and thalamic volumes in pediatric bipolar disorder," *Amer. J. Psychiatry*, vol. 162, no. 7, pp. 1256–1265, Jul. 2005.

[97] R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman, M. S. Albert, and R. J. Killiany, "An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest," *NeuroImage*, vol. 31, no. 3, pp. 968–980, Jul. 2006.

[98] J. M. Goldstein, L. J. Seidman, N. Makris, T. Ahern, L. M. O'Brien, V. S. Caviness, D. N. Kennedy, S. V. Faraone, and M. T. Tsuang, "Hypothalamic abnormalities in schizophrenia: Sex effects and genetic vulnerability," *Biol. Psychiatry*, vol. 61, no. 8, pp. 935–945, Apr. 2007.

[99] A. W. Toga and P. M. Thompson, "Mapping brain asymmetry," *Nature Rev. Neurosci.*, vol. 4, no. 1, pp. 37–48, Jan. 2003.

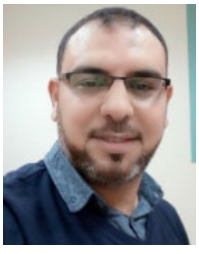
[100] P.-Y. Hervé, F. Crivello, G. Perchey, B. Mazoyer, and N. Tzourio-Mazoyer, "Handedness and cerebral anatomical asymmetries in young adult males," *NeuroImage*, vol. 29, no. 4, pp. 1066–1079, Feb. 2006.

[101] H. Zhu, T. Li, and B. Zhao, "Statistical learning methods for neuroimaging data analysis with applications," *Annu. Rev. Biomed. Data Sci.*, vol. 6, no. 1, pp. 73–104, Aug. 2023, doi: 10.1146/annurev-biodatasci-020722-100353.

[102] S. M. Smith, M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. J. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Drobnjak, D. E. Flitney, R. K. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. D. Stefano, J. M. Brady, and P. M. Matthews, "Advances in functional and structural MR image analysis and implementation as FSL," *NeuroImage*, vol. 23, pp. S208–S219, Jan. 2004, doi: 10.1016/j.neuroimage.2004.07.051.

MARTIN SVEJDA received the B.Sc. degree (Hons.) in computer science from Birmingham City University, Birmingham, U.K., in 2019, where he is currently pursuing the Ph.D. degree, with a focus on the investigation of neuroanatomical variability and structural differences in computational neuroscience and research software development.

Since 2016, he has been a Visiting Lecturer with Birmingham City University, where he is teaching diverse modules, like programming in multiple paradigms, data structures and algorithms, computer architectures, and computer and server systems. He transitioned to a DevOps Tech Lead role, in 2019, focusing on project delivery and quality assurance at BCU Software House. His academic journey began with a published work in the field of voxel-based morphometry. His current research interests include medical image processing and analysis, distributed computing, machine vision, and machine learning in medical applications.



NOUH SABRI ELMITWALLY received the master's degree in computer science from Cairo University, in 2005, and the Ph.D. degree from the University of Surrey, U.K., in 2014. He has spent several years in academics and the industry. Currently, he is a Lecturer with the School of Computing and Digital Technology, Birmingham City University, Birmingham, U.K., and the Faculty of Computers and Artificial Intelligence, Department of Computer Science, Cairo University, Egypt. He is also the Ph.D. Scholar of computer science. He has vast experience in cloud computing and machine learning. His research interests include data science, big data and artificial intelligence, machine learning, and deep learning. He has conducted many research studies on cloud computing, NLP, and AI utilization to predict various healthcare topics.

A. TAUFIQ ASYHARI received the B.Eng. degree (Hons.) in electrical and electronic engineering from Nanyang Technological University, Singapore, in 2007, and the Ph.D. degree in information engineering from the University of Cambridge, U.K., in 2012.

From 2020 to 2023, he was a Research Lead of the 10M 5G Connected Forest, looking at how 5G communications can realize the wireless services of high-definition content in forestry areas through smart remote sensing. He was on the Board of Experts in environmental computing with the Wallacea Research Centre for Biodiversity Conservation and Climate Change and has previously held full-time/visiting positions with Coventry University, U.K.; Cranfield University, U.K.; National Chiao Tung University, Taiwan; Nokia Bell Laboratories; and the University of Stuttgart, Germany. He is currently a Professor of data science with Monash University, Indonesia, and a Visiting Professor of future communication systems with Birmingham City University, U.K. He has developed an international reputation for his research work in information theory, communications, signal processing, smart sensing, and machine learning for next-generation cyber-physical systems as evidenced by the publication of 80 peer-reviewed papers in world-leading journals and conferences.

Prof. Asyhari is a Senior Member of the IEEE Information Theory Society and is affiliated with the Isaac Newton Institute for Mathematical Sciences (INI), Cambridge. He has a track record of winning and leading competitive research grants/fellowships and awards, most notably an EPSRC Global Challenge Grant, the IEEE EURASIP Best Paper Award, the NSC Taiwan Grant, the ERCIM Alain Bensoussan Fellowship, and the Cambridge Trust Ph.D. Research Scholarship. He has served as an Academic Editor for *PeerJ Computer Science*, a Topic Editor for *Sensors*, and an Associate Editor of IEEE Access (the Internet of Things, machine learning, and 5G networks), in addition to his committee involvement in numerous leading international academic conferences.

ROGER TAIT received the Ph.D. degree in parallel architectures from Nottingham Trent University, in 2007.

After graduating, he joined the Centre of Innovation and Technology Exploitation, Nottingham Trent University, as a Research Fellow, where he is developing mobile applications and web services. In late 2007, he joined the Behavioural and Clinical Neuroscience Institute, University of Cambridge, to provide technical consultancy for the parallel processing of medical imaging data within small and large-scale architectures. He is an experienced software engineer with more than 15 years of service in academic research environments. In 2017, he joined Birmingham City University, where he is currently a Lecturer of distributed computing. His research interest includes the mapping of algorithms into distributed architectures. In particular, how open-source libraries can be broken into functional units which can reside within a variety of parallel processing schemes. This includes optimizing message design and data distribution and inter-process communication and synchronization.