


Tissue Probability Based Registration of Diffusion-Weighted Magnetic Resonance Imaging

Cfir Malovani, MSc,¹ Naama Friedman, MSc,² Noam Ben-Eliezer, PhD,^{3,4,5} and Ido Tavor, PhD^{2,3,6*} 

Background: Current registration methods for diffusion-MRI (dMRI) data mostly focus on white matter (WM) areas. Recently, dMRI has been employed for the characterization of gray matter (GM) microstructure, emphasizing the need for registration methods that consider all tissue types.

Purpose: To develop a dMRI registration method based on GM, WM, and cerebrospinal fluid (CSF) tissue probability maps (TPMs).

Study Type: Retrospective longitudinal study.

Population: Thirty-two healthy participants were scanned twice (legacy data), divided into a training-set ($n = 16$) and a test-set ($n = 16$), and 35 randomly-selected participants from the Human Connectome Project.

Field Strength/Sequence: 3.0T, diffusion-weighted spin-echo echo-planar sequence; T1-weighted spoiled gradient-recalled echo (SPGR) sequence.

Assessment: A joint segmentation-registration approach was implemented: Diffusion tensor imaging (DTI) maps were classified into TPMs using machine-learning approaches. The resulting GM, WM, and CSF probability maps were employed as features for image alignment. Validation was performed on the test dataset and the HCP dataset. Registration performance was compared with current mainstream registration tools.

Statistical Tests: Classifiers used for segmentation were evaluated using leave-one-out cross-validation and scored using Dice-index. Registration success was evaluated by voxel-wise variance, normalized cross-correlation of registered DTI maps, intra- and inter-subject similarity of the registered TPMs, and region-based intra-subject similarity using an anatomical atlas. One-way ANOVAs were performed to compare between our method and other registration tools.

Results: The proposed method outperformed mainstream registration tools as indicated by lower voxel-wise variance of registered DTI maps (SD decrease of 10%) and higher similarity between registered TPMs within and across participants, for all tissue types (Dice increase of 0.1–0.2; $P < 0.05$).

Data Conclusion: A joint segmentation-registration approach based on diffusion-driven TPMs provides a more accurate registration of dMRI data, outperforming other registration tools. Our method offers a “translation” of diffusion data into structural information in the form of TPMs, allowing to directly align diffusion and structural images.

Level of Evidence: 1

Technical Efficacy Stage: 1

J. MAGN. RESON. IMAGING 2021;54:1066–1076.

The registration between two MRI scans is a fundamental step in many types of neuroimaging studies, allowing for an anatomical comparison between different individuals. Accurate transformation from each individual's native space

to a common one, while reducing the intersubject variability, is essential for group-level analysis.¹ Registration is also crucial for longitudinal studies, in which the same participant is scanned multiple times, as the comparison between

View this article online at wileyonlinelibrary.com. DOI: 10.1002/jmri.27654

Received Feb 9, 2021, Accepted for publication Apr 2, 2021.

*Address reprint requests to: Ido Tavor, Ramat Aviv 6997801, Israel. E-mail address: idotavor@tauex.tau.ac.il
Noam Ben-Eliezer and Ido Tavor contributed equally to this work.

From the ¹School of Electrical Engineering, Faculty of Engineering, Tel Aviv University, Tel Aviv, Israel; ²Department of Anatomy and Anthropology, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel; ³Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel; ⁴Department of Bio-Medical Engineering, Faculty of Engineering, Tel Aviv University, Tel Aviv, Israel; ⁵Center for Advanced Imaging Innovation and Research (CAI2R), New-York University Langone Medical Center, New York, New York, USA; and ⁶Strauss Center for Computational Neuroimaging, Tel Aviv University, Tel Aviv, Israel

Additional supporting information may be found in the online version of this article

timepoints depends on accurate alignment. While several multimodal registration techniques have been suggested in recent years,^{2–5} most default MRI registration tools are based on unimodal deformable methods, which perform transformations based on voxels' intensity values, usually in T1-weighted (T1w) or T2-weighted (T2w) images.

The registration of diffusion MRI (dMRI) images is particularly challenging, due to echo-planar imaging (EPI) sensitivity to magnetic field inhomogeneity, leading to distortions in the image and relatively low contrast between tissue types.¹ The brain's white matter (WM) is highly visible in some diffusion-driven maps, such as the fractional anisotropy (FA), while the cerebrospinal fluid (CSF) is highly visible in others, such as mean diffusivity (MD) images.⁶ Therefore, the boundaries between gray matter (GM) and CSF are difficult to detect in FA images, while the boundaries between GM and WM are difficult to detect in MD maps.

Registration methods for diffusion MRI data that rely on one contrast alone can only account for what is visible in that contrast and thus may not be sensitive to the distinction between all tissue types. In an attempt to address this problem, the recommended registration pipelines in some of the most common image-processing toolboxes, including FMRIB Software Library (FSL⁷); Statistical Parametric Mapping (SPM⁸); and Diffusion Imaging in Python (DIPY⁹);, utilize non-EPI anatomical scans (e.g. T1w, T2w) as an intermediate space for registering dMRI images.^{10,11} For example, FSL's boundary-based registration¹² has been used to align diffusion images to the T1w space (see also¹³).

Recently, several alternative approaches for dMRI registration have been proposed^{14–18}: some of these approaches rely on registration of the tensor components¹⁵ or FA maps,¹⁸ thus favoring areas of high anisotropy, i.e. the WM tracts. A multicomponent registration approach based on the different elements of the tensor has shown accurate registration of FA and WM tracts images, but its performance over non-WM parts of the brain was not demonstrated.¹⁵ Other approaches^{16,17} applied different registration methods exclusively over the WM tracts, possibly missing valuable information.

Registration procedures that focus on the WM are not adequate for an increasing body of research that uses diffusion MRI for the characterization of GM microstructure. This includes measurements of neurite distribution, myelin content, and glial cells arrangement across the cortex,¹⁸ as well as investigations of GM neuroplasticity following learning.^{19–22} Importantly, these studies rely on the comparison of GM structures between individuals or between timepoints, emphasizing the need for better registration methods considering all brain tissues rather than WM only.

We suggest that while a single dMRI contrast (e.g., FA or MD) cannot distinguish accurately between all three tissue types (i.e., GM, WM, and CSF), a combination of several

dMRI indices can. dMRI is commonly analyzed using the diffusion tensor imaging model (DTI), which enables the calculation of FA and MD as well as the three primary directions of the diffusivity tensor (denoted here as L1, L2, and L3). As each tissue displays different diffusion characteristics, multiple diffusion indices may be employed to segment the brain into GM, WM, and CSF.^{23–26} We hypothesize that using machine learning approaches, GM, WM, and CSF tissue probability maps can be generated, and that these TPMs could then be used to perform registration of dMRI images.

A joint segmentation-registration procedure has been previously suggested for T1w voxel-based morphometry (VBM) purposes and has become the basis for SPM's registration algorithm for T1w scans.²⁷ A similar approach was implemented in,²⁸ also for T1w scans.

Thus, the purpose of this study was to develop a joint segmentation-registration approach for diffusion MRI data, considering the complete anatomical composition of the brain and treating all tissue types as features to be matched during optimizations. In this approach, scans are registered according to data-driven GM, WM, and CSF probability maps rather than the original scans' intensity values.

Methods

Overview

Assuming there is available prior knowledge of tissue segmentation of the common space in the form of tissue probability maps, we applied the following steps¹: Calculated DTI parameters from the subject's diffusion-weighted scan: FA and the three perpendicular diffusion eigenvalues L1, L2, and L3.² Used a pretrained classifier to predict voxel-wise tissue probability maps based on the DTI parameters.³ Calculated a suitable transformation field from the subject's native space to the common space by registering the calculated probability maps to the supplied TPMs—in our case, the MNI TPMs.^{29,4} Applied the resultant transformation field to the subject—either on the original diffusion-weighted scan or directly on the DTI parameters. Figure 1 presents the proposed analysis pipeline.

Data for Algorithm Testing

The proposed algorithm was tested on three datasets: a DTI Shepp-Logan phantom,³⁰ an experimental legacy dataset,²² and the Human Connectome Project (HCP) dataset.³¹ The DTI phantom is described in the Supplementary Methods and Figs. S1 and S2.

Experimental data: The previously acquired experimental dataset²² consisted of 32 healthy volunteers (mean age 25.7; SD 3.3, 13 males, all right-handed), with no history of neurological diseases, psychological disorders, drug or alcohol abuse, or use of neuropsychiatric medication. All participants

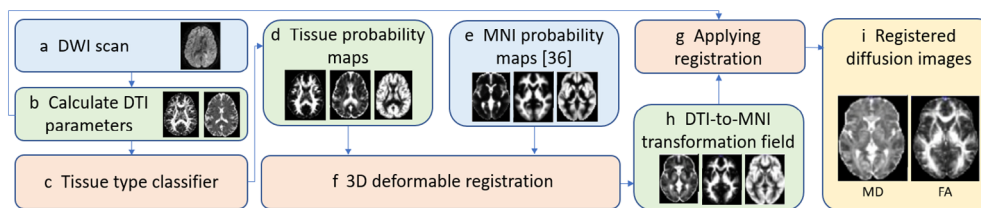


FIGURE 1: The proposed method pipeline. Given a DWI scan of a participant (a), the DTI parameters are calculated (b) and inserted into our classifier (c), resulting in voxel-wise tissue probability maps (TPMs) for that participant (d). Using the MNI pre-existing TPMs (e) and the proposed probability-based registration method (f), we calculate a transformation field from the participant's native space to the MNI common space (g). Applying this transformation (h) on the participant's DTI images (b) yields diffusion scans in the MNI common space (i). Inputs for the process are presented in blue, outputs in yellow, major stages of the proposed method in red, and intermediate calculations in green.

signed an informed consent form; the research protocol was approved by the Institutional Review Board.

Data were acquired in 2015 on a GE Signa 3 T scanner (GE Healthcare, Milwaukee, WI, USA) using a then-standard protocol. Participants underwent MRI scans at two or three timepoints approximately an hour apart (only two were used in this study). The MRI protocol included DTI and conventional anatomical sequences for radiological screening, all acquired with an eight-channel head-coil. Spin-echo diffusion weighted echo-planar imaging (DW-EPI) sequences were performed with up to 70 axial slices to cover the whole brain and resolution of $2.1 \text{ mm} \times 2.1 \text{ mm} \times 2.1 \text{ mm}$ reconstructed to $1.58 \text{ mm} \times 1.58 \text{ mm} \times 2.1 \text{ mm}$ (field of view = 202; acquisition matrix = 96×96 reconstructed to 128×128). Diffusion parameters were: $\Delta/\delta = 33/26 \text{ ms}$, a b-value of 1000 s/mm^2 was acquired with 30 gradient directions together with an additional no diffusion weighting (b0) image. T1-weighted images were acquired with a three-dimensional spoiled gradient-recalled echo sequence with a resolution of $1 \times 1 \times 1 \text{ mm}^3$.

All scans were preprocessed using FSL 6.0.3. The preprocessing included head motion correction, skull stripping, and extraction of the DTI parameters to be used for segmentation: fractional anisotropy (FA) and the three eigenvalues of the diffusion tensor (L1, L2, and L3). Parameters were separately demeaned and scaled to unit variance across all scans (approximating normal distribution). They were then concatenated to create 4D images of size $[W, H, D, 4]$, where W, H, D are the spatial dimensions of the original scan and the fourth dimension contains the four diffusion indices (FA, L1, L2, and L3).

HCP Data: To further explore the performance of our method on a state-of-the-art dataset rather than on the clinical, legacy data described above, we used dMRI data from 35 randomly-selected participants from the Human Connectome Project.³¹

Diffusion-Based Segmentation

The first step in the proposed registration method (Fig. 1c) is to create probability maps for each tissue type (GM, WM,

and CSF) based on the DTI parameters derived from the scan. We propose a machine learning based classifier trained on a manually labeled dataset, randomly sampled from the experimental legacy data. The target outputs were 4D images of size $[W, H, D, 4]$ where W, H, D are the spatial dimensions of the original scan and the fourth dimension consists of the different probability maps.

The tissue probability maps (TPM) contain four different modalities: GM, WM, CSF, and background (BG; non-brain voxels). Each modality represents the probability of each voxel's affinity to that tissue type, with the background image being used to separate the brain from the rest of the image, as well as to preserve unitarity of the probability space:

$$P_{GM}(x) + P_{WM}(x) + P_{CSF}(x) + P_{BG}(x) = 1 \quad \forall x \in \Omega \quad (1)$$

where P_{GM} , P_{WM} , P_{CSF} and P_{BG} are the probabilities of the voxel representing GM, WM, CSF, and non-brain (BG), respectively. This equation holds for every voxel x in the MRI image (denoted here as the set Ω).

MANUALLY LABELLED DATASET. Based on the previously-acquired experimental data,²² a dataset with manually labeled voxels was created by six neuroanatomists (Table SI), using an in-house dedicated MATLAB 9.3.0 (MathWorks, Natick, MA) script for visualizing and labeling of MRI scans (Figs. S3 and S4). A total of 32 scans (the first scan of each of the 32 participants²²) were split into a training set ($N = 16$), which was used for manual labeling, and a test set ($N = 16$) previously unseen by the developed classifier. The resulting dataset was therefore composed of 16 partially labeled scans, with 1–2 k voxels from each tissue type labeled in each scan, totaling approximately 68 k different voxels assigned ground-truth labels for the classifier's training; 40.2% of these voxels were labeled as GM, 27.2% as WM, and 32.6% as CSF. This ratio is reflected in the weighted loss function of the different classifiers.

CLASSIFIERS. Segmentation was performed using traditional machine learning techniques, classifying each voxel as an individual object. The following methods were examined: Logistic

Regression (**LR**); Gradient Boosting (**GB**); and Random Forests (**RF**). These techniques were used after testing known segmentation tools such as FSL’s FAST³² and DIPY Markov Random Field based segmentation,⁹ which yielded unsatisfying performance on dMRI data, emphasizing the need for a dedicated dMRI tissue classifier.

The classifiers were trained on the manually labeled dataset, each voxel having the DTI parameters FA, L1, L2, and L3 as features. Each classifier was trained as a multi-class problem, in which the classifier was trained to distinguish between the three different labels, with the last label (background) determined by the skull stripping. The TPMs were constructed from the classifier’s raw output, with no application of arg-max for hard-labeling. Smoothing was applied to the TPMs in order to minimize possible effects of single misclassified voxels.

For classifiers’ training, we used available Python libraries: scikit-learn 0.22³³ for LR and RF and XGBoost 0.9³⁴ for GB, using weighted categorical cross-entropy as the target loss function. Each voxel’s classification was determined individually in order to reduce masking of anatomical features such as thin WM fibers.

SEGMENTATION VALIDATION. Segmentation performance was evaluated on images from three sources: a DTI phantom (see Supplementary Methods), the experimental legacy data, and the HCP data.

Experimental data: Classifiers’ performance was evaluated on the experimental data using Leave-one-out Cross-Validation (CV) as follows: Given N manually labeled participants (in our case, $N = 16$), all classifiers were trained N times, each time on a different subset of all voxels from $N - 1$ participants, and then tested on the voxels’ subset of the left-out participant. All classifiers were scored using the Dice (Sørensen–Dice) index over each tissue type separately as well as over the dataset as a whole, with a final score for each classifier type given as an average of all N leave-one-out training results for that classifier.

HCP data: Of the three classifiers tested on the phantom and experimental data, the one showing the best performance was further tested on the HCP data. This classifier’s segmentation performance was compared to the segmentation of diffusion images that are already in the T1-space, provided by the HCP. These can serve as a good approximation to the DWI segmentation; however, since there is no actual ground truth, Dice scores between the TPMs generated by our classifier and those from the HCP were calculated over thresholded maps, with a threshold of 0.66 to ensure evaluation over areas of high confidence in both methods.

Spatial Registration

The second major step in the proposed algorithm (Fig. 1f) was applying a spatial registration based on the TPMs generated in the segmentation step.

REGISTRATION IMPLEMENTATION. Implementation of the proposed registration¹ was done by modifying an existing Python neuroimaging tool—DIPY’s⁹ implementation of Symmetric Normalization (**SyN**)³⁵ based registration using normalized cross-correlation (**NCC**) with the implementation of³⁶ as a proof-of-concept for our hypothesis. The motivation for choosing the SyN-NCC implementation derives from its high ranking in several comparative reviews^{37,38} relative to other widely used deformable registration methods over T1w volumes. Specifically, the use of cross-correlation, which is calculated over the neighborhood of each voxel, is suited to our goal of matching prominent spatially similar areas of the brain in each tissue type. Our modifications are made in two stages of the implemented algorithm: the cost function and the transformation application.

MULTIVARIATE COST FUNCTION. The cost function, which originally handled cases of unimodal 3D scans, was extended to apply to 4D multivariate images, with the fourth dimension representing the different TPMs (not unlike the third dimension in RGB images)—effectively treating the different TPMs as different modalities. The modified NCC cost function was calculated for each modality separately, with the result being a summation over all modalities:

$$NCC(\mathbf{x}, v) = \frac{\left[\sum_{\mathbf{z} \in W_{\mathbf{x}}} (T(\mathbf{z}, v) - \mu_{\mathbf{x}}^T(v)) (M(\phi(\mathbf{z}), v) - \mu_{\mathbf{x}}^M(v)) \right]^2}{\left[\sum_{\mathbf{z} \in W_{\mathbf{x}}} (T(\mathbf{z}, v) - \mu_{\mathbf{x}}^T(v))^2 \right] \left[\sum_{\mathbf{z} \in W_{\mathbf{x}}} (M(\phi(\mathbf{z}), v) - \mu_{\mathbf{x}}^M(v))^2 \right]} \quad (2)$$

$$mNCC(\mathbf{x}) = \sum_{v \in V} NCC(\mathbf{x}, v) \quad (3)$$

$$mNCC_{Total}(M, T, \phi) = \sum_{x \in \Omega} mNCC(x) \quad (4)$$

where x denotes the current voxel; Ω is the voxel space (all voxels in the image); $\mathbf{z} \in W_{\mathbf{x}}$ is the voxel in neighborhood of x ; $W_{\mathbf{x}}$ denotes the neighborhood of x —a box centered at x ; v is the current modality (in this study, the current probability map); V represents the modalities’ space (in this study, the complete TPM); M is the moving image; T is the target image; ϕ is the transformation fields; $\mu_{\mathbf{x}}^M(v)$ denotes the local means of M around x for modality v , and $\mu_{\mathbf{x}}^T(v)$ denotes the local means of T around x for modality v .

As each modality now produces a separate NCC, and by extension a separate gradient value, the output gradient

¹Code is available on: https://github.com/cfimalovani/tpm_registration_for_dmri.

was modified as well to become a summation of the different modalities' yielded gradients, thus giving equal weights in the normalization cost evaluation to the different tissues in the MRI scan. Based on the work of,³⁶ the final gradients were calculated as follows:

$$\nabla_{\phi(\mathbf{x})} NCC(\mathbf{x}, v) = \frac{2A}{BC} \left(\bar{T} - \frac{A}{C} \bar{M} \right) (\nabla \phi) (\nabla \bar{M}) \quad (5)$$

$$\nabla_{\phi(\mathbf{x})} NCC_{modified}(\mathbf{x}) = \sum_{v \in V} \nabla_{\phi(\mathbf{x})} NCC(\mathbf{x}, v) \quad (6)$$

with the following values defined for ease of writing:

$$A(\mathbf{x}, v) = \sum_{\mathbf{z} \in \mathbb{W}_x} (T(\mathbf{z}, v) - \mu_{\mathbf{x}}^T(v)) (M(\phi(\mathbf{z}), v) - \mu_{\mathbf{x}}^M(v)) \quad (7)$$

$$B(\mathbf{x}, v) = \sum_{\mathbf{z} \in \mathbb{W}_x} (T(\mathbf{z}, v) - \mu_{\mathbf{x}}^T(v))^2 \quad (8)$$

$$C(\mathbf{x}, v) = \sum_{\mathbf{z} \in \mathbb{W}_x} (M(\phi(\mathbf{z}), v) - \mu_{\mathbf{x}}^M(v))^2 \quad (9)$$

$$\bar{T}(\mathbf{x}, v) = T(\mathbf{x}, v) - \mu_{\mathbf{x}}^T(v) \quad (10)$$

$$\bar{M}(\mathbf{x}, v) = M(\phi(\mathbf{x}), v) - \mu_{\mathbf{x}}^M(v) \quad (11)$$

APPLYING THE TRANSFORMATION FIELDS. The transformation fields update remains the same under our new cost calculation, but the application of those fields is also modified: As the different probability maps all reside in the same space, the transformation was applied to each modality separately, and the transformed maps were rejoined to the updated 4D scan data. Then, as the transformations do not necessarily uphold the need for unitarity (the sum of all probabilities in each voxel must be equal to one), the transformed scan was normalized to validate the probability values, with the new background threshold determined as $P(\text{background}) > 0.5$.

$$M_{warped}(\mathbf{x}, v) = \frac{M(\phi(\mathbf{x}), v)}{\sum_{v \in V} M(\phi(\mathbf{x}), v)} \quad (12)$$

VALIDATING IMAGE REGISTRATION. To test the performance of the proposed registration method, we compared its results with widely-used registration methods implemented in FSL (6.0.3), SPM,¹² and DIPY (1.0, original unmodified version). For each of these methods, we used its recommended target image in the MNI space (since the MNI TPMs can only be used as targets for our proposed method and not for current methods that are not based on TPMs). We used FSL in a pipeline similar to that suggested by the HCP for dMRI registration.³⁹ It applies a Boundary-Based Registration (BBR) method¹² to register diffusion scans to the counterpart

T1w scans of the same participant, which are then used as an intermediate step toward the goal registration. The resulting dual transformation fields were applied over the different DTI maps. For dMRI registration with SPM and DIPY (SyN implementation), we used the recommended pipelines.^{10,11} Both SPM and DIPY registration procedures use deformable registration from each participant's b0 image to the MNI-T2w scan, applying the resulting transformation fields on the different DTI parameters maps. In order to explore only the differences arising from registration processes, all scans were preprocessed by FSL to extract brain masks as well as the DTI parameters maps, which were then fed into the different registration processes.

The evaluation of the registration success is not trivial, as there is a need to test registration performance not on the probability maps—which do not exist for the other methods—but on the different DTI parameters maps—specifically on the FA and MD maps. We used three approaches for assessing registration performance¹: Calculation of the voxel-wise variance of each registered participant to the average scan—which could be used as a de-facto common space: the lower the overall voxel-wise variance, the better each participant is aligned. The ideal registration will not yield zero variance across the transformed scans, since inter-subject variability is still expected due to differences in brain structure between participants. However, high variance reflects registration error, as it indicates that different tissues were registered to the same voxel. Calculating this variance per voxel also allows determination and analysis of 'weak spots' in the different registration methods, meaning anatomical areas where registration is more susceptible to errors.² Calculation of the voxel-wise normalized cross-correlation (NCC) of each registered participant from the average space. A higher NCC score (which varies between 0 and 1) indicates a stronger similarity of the registered image to the MNI image. This evaluation used a window of $9 \times 9 \times 9$ voxels, with the averaged NCC score used as the evaluation score.³ Calculation of the similarity of the registered probability maps—both intra-subject (comparing each participant's two scans) and intersubject (comparing each participant to all other participants). A successful registration process would yield high similarity between registered scans, reflected by high Dice scores. The higher the score, the better the process is in transforming different probability maps—and the scans themselves—to the target common space. Generating probability maps for all available 64 scans (32 subjects \times 2 scans/subject) and setting the probability threshold at 0.5 for both GM and WM maps, the Dice scores were calculated between each scan and all other 63 scans.

An additional test was performed to determine that the success of our method rise not from the use of a multivariate registration in general, but from the strength of the TPM basis: We implemented a multivariate registration process

based directly on the DTI eigenvalues using the DIPY framework, and compared its performance to those of our method (see Supplementary Fig. S11 and Table SVII).

It has been previously suggested⁴⁰ that whole-brain image similarity and tissue overlap are not the optimal surrogate measurements of registration accuracy, advocating for region-based comparison as a possible evaluation method. Therefore, in addition to the three evaluation methods described above, we used an existing anatomical atlas to test registration performance on a regional basis (see supplementary Fig. S10 and Table SV).

Finally, we compared the registration performance of the HCP recommended pipeline, as implemented in the DIPY toolbox, with that of our proposed method. Both pipelines were run on 35 randomly-chosen participants from the HCP, and we measured the voxel-wise and averaged standard deviation from the common space for both FA and MD images.

STATISTICAL ANALYSIS. To compare the registration performance of the proposed method with those of the other tools, we performed one-way analyses of variance (ANOVA) for each registration performance measure separately, with the performance measure (e.g., voxel-wise variance from the averaged image) as the dependent variable and the registration method (i.e., FSL, SPM, DIPY, and the proposed method) as the independent variable. Results were considered statistically significant at $P < 0.05$.

Results

Testing the proposed method required validation of both segmentation and registration processes. We first report the performance of the different classifiers tested for the segmentation step on the experimental data (see Supplementary Results for segmentation of the DTI phantom and the HCP data—Figs. S5–S8). We then describe the results of the registration step, based on the segmentation outputs.

Diffusion-Based Segmentation

SEGMENTATION OF THE EXPERIMENTAL DATA. Table 1 shows the Dice results of all classifiers, calculated over all cross-validation models in order to avoid single-subject dataset bias. All classifiers exhibited good performance on these experimental data, with Dice > 0.9 for all tissue types. The Gradient-Boosting based classifier displayed the best performance between the tested models. Comparison of the different classifiers' performance is shown in Fig. S6, and multiple examples of the Gradient-Boosting classifier performance are shown in Fig. 2 and Fig. S7.

TABLE 1. Tissue Segmentation Dice Scores of the Experimental Data

Classifier	GM	WM	CSF	Overall
LR	0.9034	0.9423	0.9164	0.9213
RF	0.9011	0.9437	0.9097	0.9186
GB	0.9084	0.9473	0.9162	0.9245

Spatial Registration

REGISTRATION OF THE EXPERIMENTAL DATA. Despite the relatively low resolution of our experimental data, our method yielded anatomically valid registered tissue maps as shown in Fig. 3, which presents these maps on a 0.5 mm MNI T1w scan.

Figure 4 and Table 2 present the voxel-wise and averaged standard deviation from the common space in each registration method (ours, FSL, SPM, and DIPY) for FA and MD images. Our method provided the least average deviation ($P < 0.05$) from the common space and showed the smoothest SD maps, reflecting no anatomical areas sensitive to registration error. Of note is the improvement in registration over the original DIPY registration process on which our proof of concept is based—especially in MD images, where the average deviation is reduced by 28%.

The registration process will ideally display no anatomical deviation from the common space due to misalignment, but only differences in FA and MD that arise from inter-subject differences in brain microstructure. Therefore, a successful registration would show minimal anatomical outlines in these SD maps, since high values of SD along certain anatomical regions indicate a difficulty in registering them between all participants. Figure 4 displays such difficulty in registering the corpus callosum in all other tested methods, as well as some difficulty in defining the brain borders in FSL and SPM.

Our method provided the highest NCC scores in both FA and MD images relative to DIPY, FSL, and SPM ($P < 0.05$), reflecting better alignment of the registered scans to the common space (Fig. S9 and Table SIV).

The Dice scores for the registered maps created by our classifier and other registration methods are shown in Fig. 5 and Table 3. The matrices in Fig. 5 are arranged such that the first scans of all participants appear first, followed by the second scans in the same participants' order. This is reflected by strong sub-diagonals in the matrices, indicating that each post-registration map is most similar to the second map from the same participant. Our method provided the highest Dice score for both intra- and intersubject registration ($P < 0.05$), with the respective average Dice scores being 0.957 and 0.897 for GM, and 0.940 and 0.855 for WM. Of note is the

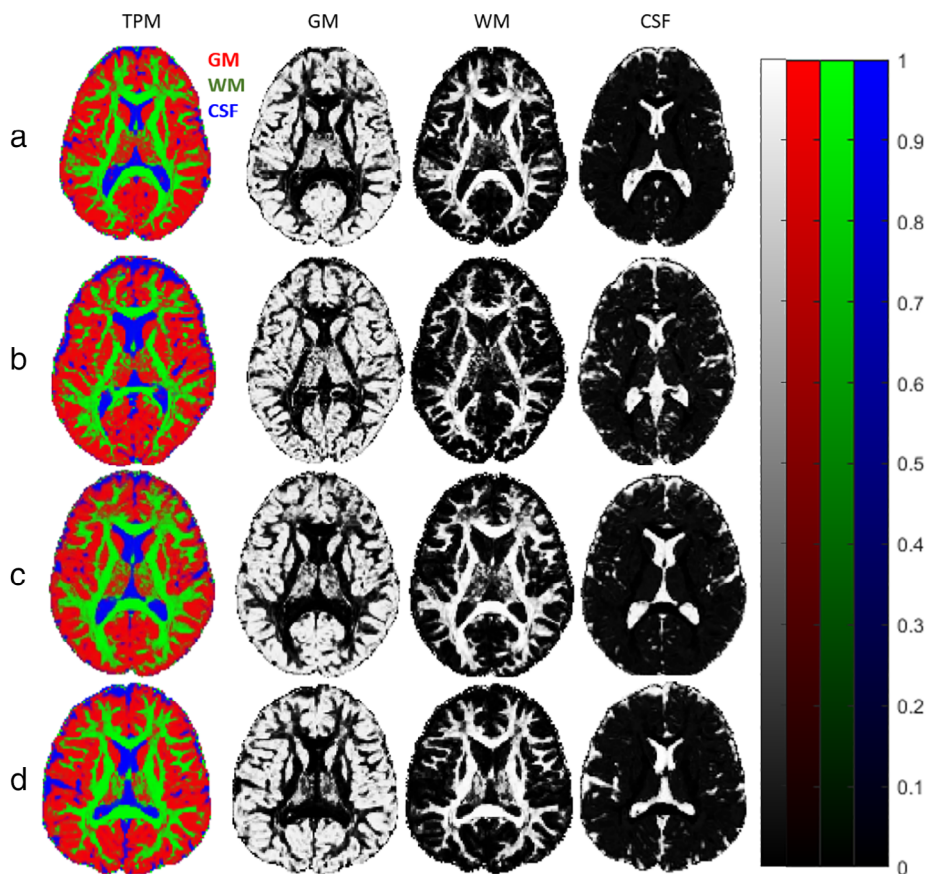


FIGURE 2: Results of the gradient-boosting based classifier on experimental diffusion MRI data. Segmentation results in four participants (i-iv) are presented as RGB probability maps and separated into different tissue types (representing probabilities as well): GM, WM, and CSF.

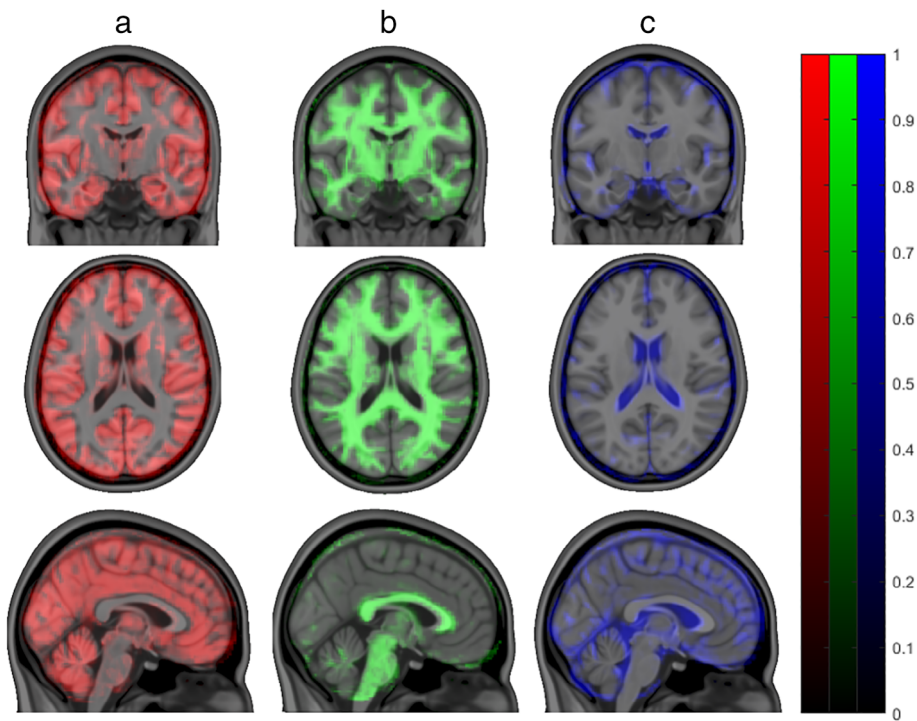


FIGURE 3: Tissue registration into MNI space. Sample probability maps for one subject, registered to MNI space, and separated to GM, WM, and CSF. Although registration was performed to MNI 2 mm atlas, the results here are displayed on 0.5 mm atlas to show similarity to brain structures.

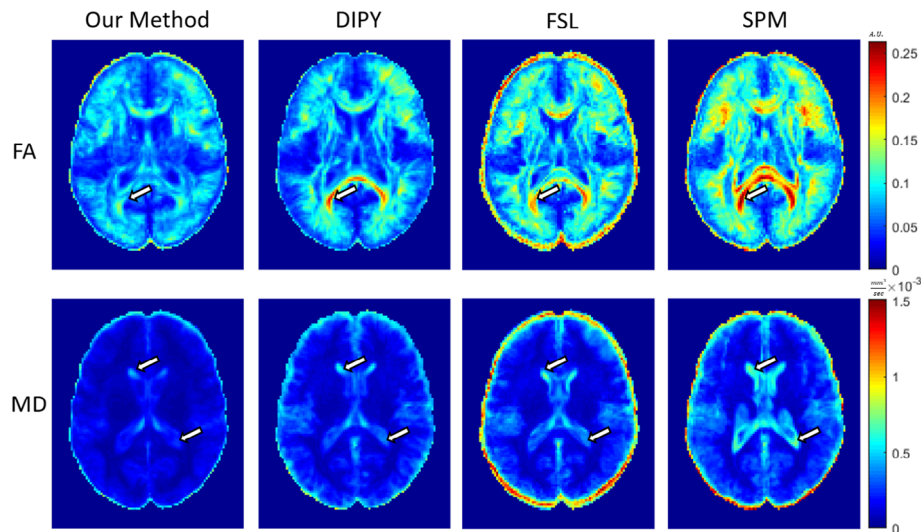


FIGURE 4: Comparison between different registration methods using deviation of the registered images from the common MNI space. The figure presents the voxel-wise standard deviation (SD) of registered images vs. common space estimated across 32 participants. Average SD values for FA and MD maps are shown in Table 2. The proposed method outperforms DIPY, FSL, and SPM procedures ($P < 0.05$). Arrows indicate areas of mis-registration.

Scan	Our Method	DIPY	FSL	SPM
FA [A.U.] $\times 10^2$	6.79	7.52	10.18	10.45
MD [mm ² /s] $\times 10^4$	1.96	2.73	3.84	3.82

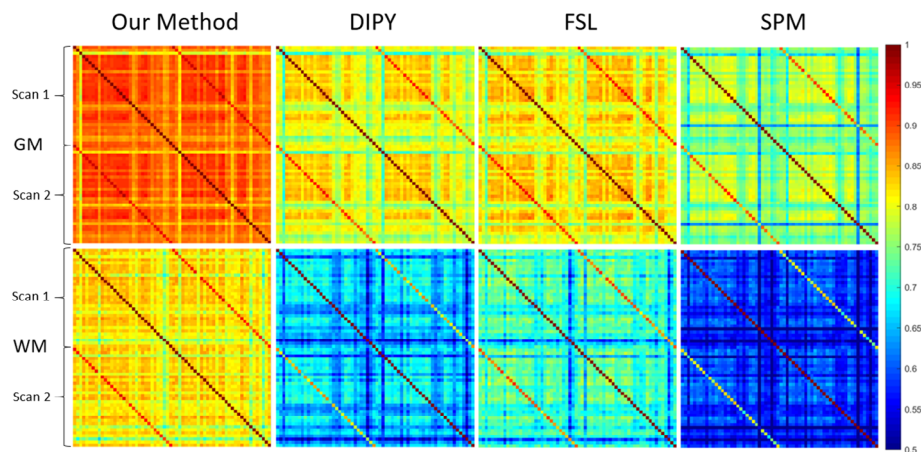


FIGURE 5: Comparison between different registration methods using similarities between the registered segmentation maps. The figure shows pairwise Dice indices of all participants’ segmentation maps after registration using our method, DIPY, FSL, and SPM, for GM (top) and WM (bottom). Each matrix contains two consecutive scans per participants. The proposed method outperforms the other procedures in both intra- and intersubject correlation ($P < 0.05$).

improved WM registration using our method compared to FSL, considering that BBR relies on WM-based registration of the DWI scan to T1w native space.

Finally, the DTI-based multivariate approach improved registration performance relative to the univariate method in several areas, most noticeable in the corpus callosum

(Fig. S11). However, it created additional misalignments in other parts of the brain, likely due to low contrast presentation in the eigenvalues’ parameter maps. Table SVII shows that the multivariate method achieved an overall better performance over the univariate one—while still being inferior to our proposed method.

TABLE 3. Comparison Between Different Registration Methods: Average Dice Scores Between Registered Segmentation Maps

Tissue Map	Dice Score	Our Method	DIPY	FSL	SPM
GM	Intra-subject	0.9568	0.9026	0.9164	0.8791
	Intersubject	0.8967	0.7975	0.8111	0.7585
WM	Intra-subject	0.9395	0.8184	0.8750	0.7785
	Intersubject	0.8553	0.6507	0.7070	0.5639

REGISTRATION OF THE HCP DATA. Both our method and DIPY (chosen as the lead method from previous tests) performed better on HCP scans than on our legacy data. However, the proposed method performed significantly better than DIPY on the HCP data ($P < 0.05$; Fig. S12 and Table SVIII).

Discussion

This study presents a novel method of tissue probability-based registration, in which diffusion MRI scans are registered according to data-driven probability maps rather than the original scans' intensity values. The proposed method is grounded on two premises. The first premise is the ability to segment brain tissues based on DTI parameters derived from dMRI scans. As the diffusivity of water molecules differs between tissue types, diffusion profiles may form a valid basis for an accurate tissue segmentation method. While each DTI parameter is partially susceptible to differences between brain tissues, multiple parameters are needed to achieve complete segmentation of the brain. The second premise is that a univariate registration cannot be optimal for dMRI images, as a single contrast does not contain enough information to accurately describe the entire brain. Basing the registration process on a single contrast will yield a good fit for the visible portions of the brain in that contrast, but is likely to overlook other areas in the process, leading to a deformed result. Our proposed method makes use of multiple DTI indices to create three tissue probability maps. Using these TPMs (i.e., GM, WM, and CSF) as the basis for registration not only provides more comprehensive information about the brain due to its multivariate nature, but also enacts a strong biological constraint on the registration procedure, allowing a better match of the overall process. This approach is similar to the procedure presented in Reference 27 for T1w scans, but is now applicable to dMRI scans as well. While other multimodal methods for dMRI registration may rely on the diffusion parameters themselves, here we "translate" the diffusion data into meaningful structural information in the form of TPMs, thus making it possible to directly align diffusion and structural images.

The proposed method was implemented in a fully automated pipeline for both tissue segmentation of dMRI scans and their registration to the common MNI space. This pipeline consists of a DTI-based tissue classifier, and a DIPY-based multivariate registration process. The registration to the MNI space was achieved by the use of a-priori existing probability maps of the MNI standard space. We showed that this proposed pipeline outperformed existing registration methods, supporting the idea that considering all brain tissues during the registration process, rather than white matter alone, not only improves results in the previously disregarded tissues (e.g., GM structures), but also in the WM. Furthermore, the relative success of our method in the registration of the brain's outer borders (especially the separation of the cortex from the subarachnoid space) compared to SPM and FSL unimodal methods reflects the higher accuracy of GM and CSF separation. As most existing methods rely on WM-based registration, our method may be more suitable for studies focusing on gray matter microstructure.^{20,22}

Limitations

The proposed registration process depends on the existence of TPMs for both the registered and target (e.g., MNI) scans, or the ability to accurately calculate such probability maps via an external classifier. In the second case, the quality of the registration is directly proportional to that of the tissue classifier. Here, we have demonstrated the proposed registration process using pre-existing MNI TPMs and a tissue classifier for diffusion MRI scans. Applying the process for other types of scans or different registration targets will require modifications to the existing pipeline. Nevertheless, high-quality TPMs in standard space are publicly available, and multiple methods for tissue segmentation exist, so such modifications should not necessarily be complex. As the TPMs themselves are modality-invariant, our method is applicable regardless of the original data modality (e.g., TPMs computed from T1w images can be registered to TPMs computed from diffusion images).

The tissue-classifier was trained on a manually labeled dataset. Such training dataset requires time and resources from expert neuroanatomist; heavily depends on the labelers'

level of expertise; and lacks inter-labeler correlations. Future studies may aim to increase the size of the training-set in terms of the number of labeled participants, the number of voxels labeled for each, the number of labelers, and an assessment of the inter-labeler variance.

As both the training dataset and target TPMs used in this study had spatial resolution of about 2 mm^3 (specifically, 2 mm^3 for the TPMs and $1.58 \times 1.58 \times 2.1 \text{ mm}$ for the diffusion images), registration of higher-resolution scans will require more accurate classifiers. Although all trained classifiers in this study achieved high scores on the labeled dataset, they all featured a prominent disadvantage of visible bias toward GM classifications (labeling anatomically known, but unlabeled, non-GM areas as GM). This is especially prominent in the GM-CSF borders, where all classifiers favored labeling such voxels as GM. A possible cause for this bias is the existence of partial volume errors in the low-resolution scans, and thus using higher-resolution scans may be expected to resolve such biases. This problem may also arise from the nature of the manual labels used as the training dataset for the classifier: the neuroanatomists were instructed to extract segments of tissue-specific areas in the brain, rather than to focus on the ‘difficult’ regions and tissue borders. Increasing the size of the manually-labeled dataset, as well as labeling ill-defined voxels, may assist in reducing partial volume biases. Importantly though, the fact that registration was successful even when using a previously acquired, clinical dataset rather than a state-of-the-art one (e.g., the HCP database³¹) highlights the potential applicability of our method for regular datasets that do not meet the HCP standard.

Notably, the current method treats all tissue types equally. For studies focusing on gray matter areas of the brain, a possible improvement of the procedure would be adding a weighting factor to the registration cost function, e.g., greater penalizing of errors in the GM than WM and CSF.

Conclusion

A two-stage approach for dMRI registration based on tissue probability maps has been demonstrated on human participants, outperforming other widely-used registration methods. Furthermore, the probabilistic nature of our registration process—ultimately independent of the diffusion parameters—makes it highly compatible with other registration pipelines, regardless of scan resolution or study goals. We thus provide a promising method for diffusion MRI registration, which considers all tissue types and provides an accurate, anatomically-valid registration.

Acknowledgments

Contract grant sponsor: Israel Science Foundation (Grant Nos. 1603/18, 2009/17).

References

1. Jones DK, Cercignani M. Twenty-five pitfalls in the analysis of diffusion MRI data. *NMR Biomed* 2010;23:803-820.
2. Alegro M, Amaro E, Loring B, et al. Multimodal whole brain registration: MRI and high resolution histology. *IEEE Comput Soc Conf Comput Vis Pattern Recognit Work* 2016;634-642. <https://doi.org/10.1109/CVPRW.2016.85>.
3. Simonovsky M, Gutiérrez-Becker B, Mateus D, Navab N, Komodakis N. A deep metric for multimodal registration. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)* 2016; 9902 LNCS (October): pp. 10–18.
4. Goubran M, Leuze C, Hsueh B, et al. Multimodal image registration and connectivity analysis for integration of connectomic data from microscopy to MRI. *Nat Commun* 2019;10:1-17.
5. Yang H, Qian P, Fan C. An indirect multimodal image registration and completion method guided by image synthesis. *Comput Math Methods Med* 2020;2020:1-10.
6. Assaf Y, Pasternak O. Diffusion tensor imaging (DTI)-based white matter mapping in brain research: A review. *J Mol Neurosci* 2008;34:51-61.
7. Smith SM, Jenkinson M, Woolrich MW, et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 2004;23(SUPPL 1):208-219.
8. Berry AS, Zanto TP, Clapp WC, et al. Statistical parametric mapping: An annotated bibliography. *PLoS One* 2001;5:e11537.
9. Garyfallidis E, Brett M, Amirbekian B, et al. Dipy, a library for the analysis of diffusion MRI data. *Front Neuroinform* 2014;8:1-17.
10. Liu G, Yao L, Zhao X. Optimization of image preprocessing for diffusion tensor magnetic resonance imaging. Paper presented at: Proc - 2010 6th Int Conf Nat Comput ICNC 2010; 8(Icnc): pp. 4139–4142.
11. Soares JM, Marques P, Alves V, Sousa N. A hitchhiker's guide to diffusion tensor imaging. *Front Neurosci* 2013;7:1-14.
12. Greve DN, Fischl B. Accurate and robust brain image alignment using boundary-based registration. *Neuroimage* 2010;48:63-72.
13. Jacobacci F, Jovicich J, Lerner G, et al. Improving spatial normalization of brain diffusion MRI to measure longitudinal changes of tissue microstructure in the cortex and white matter. *J Magn Reson Imaging* 2020; 52:766-775.
14. Smith SM, Jenkinson M, Johansen-Berg H, et al. Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *Neuroimage* 2006;31:1487-1505.
15. Khader M, Schiavi E, Hamza A. Ben: A multicomponent approach to nonrigid registration of diffusion tensor images. *Appl Intell* 2017;46: 241-253.
16. Ziyang U, Sabuncu MR, Grimson WEL, Westin CF. Consistency clustering: A robust algorithm for group-wise registration, segmentation and automatic atlas construction in diffusion MRI. *Int J Comput Vis* 2009;85: 279-290.
17. Wang Y, Gupta A, Liu Z, et al. DTI registration in atlas based fiber analysis of infantile Krabbe disease. *Neuroimage* 2011;55:1577-1586.
18. Fukutomi H, Glasser MF, Zhang H, et al. Neurite imaging reveals microstructural variations in human cerebral cortical gray matter. *Neuroimage* 2018;182(May 2017):488-499.
19. Sagi Y, Tavor I, Hofstetter S, Tzur-Moryosef S, Blumenfeld-Katzir T, Assaf Y. Learning in the Fast lane: New insights into neuroplasticity. *Neuron* 2012;73:1195-1203.
20. Brodt S, Gais S, Beck J, Erb M, Scheffler K, Schönauer M. Fast track to the neocortex: A memory engram in the posterior parietal cortex. *Science* 80 2018;362:1045-1048.
21. Tavor I, Hofstetter S, Assaf Y. Micro-structural assessment of short term plasticity dynamics. *Neuroimage* 2013;81:1-7.
22. Tavor I, Botvinik-Nezer R, Bernstein-Eliav M, Tsarfay G, Assaf Y. Short-term plasticity following motor sequence learning revealed by diffusion magnetic resonance imaging. *Hum Brain Mapp* 2020;41:442-452.

23. Hasan KM, Halphen C, Sankar A, et al. Diffusion tensor imaging-based tissue segmentation: Validation and application to the developing child and adolescent brain. *Neuroimage* 2007;34:1497-1505.
24. Liu T, Li H, Wong K, Tarokh A, Guo L, Wong STC. Brain tissue segmentation based on DTI data. *Neuroimage* 2007;38:114-123.
25. Cheng H, Newman S, Afzali M, Fadnavis SS, Garyfallidis E. Segmentation of the brain using direction-averaged signal of DWI images. *Magn Reson Imaging* 2020;69:1-7.
26. Ciritsis A, Boss A, Rossi C. Automated pixel-wise brain tissue segmentation of diffusion-weighted images via machine learning. *NMR Biomed* 2018;31:1-9.
27. Ashburner J, Friston KJ. Unified segmentation. *Neuroimage* 2005;26:839-851.
28. Heckemann RA, Keihaninejad S, Aljabar P, Rueckert D, Hajnal JV, Hammers A. Improving intersubject image registration using tissue-class information benefits robustness and accuracy of multi-atlas based anatomical segmentation. *Neuroimage* 2010;51:221-227.
29. Fonov V, Evans AC, Botteron K, Almli CR, McKinstry RC, Collins DL. Unbiased average age-appropriate atlases for pediatric studies. *Neuroimage* 2011;54:313-327.
30. Shepp LA, Logan BF. Fourier reconstruction of a head section. *IEEE Trans Nucl Sci* 1974;NS-21:21-43.
31. Van Essen DC, Ugurbil K, Auerbach E, et al. The human connectome project: A data acquisition perspective. *Neuroimage* 2012;62:2222-2231.
32. Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging* 2001;20:45-57.
33. Varoquaux G, Buitinck L, Louppe G, Grisel O, Pedregosa F, Mueller A. Scikit-learn: Machine learning in python. *GetMobile Mob Comput Commun* 2015;19:29-33.
34. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. Paper presented at: *Proc ACM SIGKDD Int Conf Knowl Discov Data Min* 2016; 13-17 Aug, pp. 785-794.
35. Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal* 2008;12:26-41.
36. Ocegueda O, Dalmau O, Garyfallidis E, Descoteaux M, Rivera M. On the computation of integrals over fixed-size rectangles of arbitrary dimension. *Pattern Recognit Lett* 2016;79(November 2017):68-72.
37. Ou Y, Akbari H, Bilello M, Da X, Davatzikos C. Comparative evaluation of registration algorithms in different brain databases with varying difficulty: Results and insights. *IEEE Trans Med Imaging* 2014;33:2039-2065.
38. Ewert S, Horn A, Finkel F, Li N, Kühn AA, Herrington TM. Optimization and comparative evaluation of nonlinear deformation algorithms for atlas-based segmentation of DBS target nuclei. *Neuroimage* 2019;184 (August 2018):586-598.
39. Glasser MF, Sotiropoulos SN, Wilson JA, et al. The minimal preprocessing pipelines for the human connectome project. *Neuroimage* 2013;80:105-124.
40. Rohlfing T. Image similarity and tissue overlaps as surrogates for image registration accuracy: Widely used but unreliable. *IEEE Trans Med Imaging* 2012;31:153-163.