Digital Analysis of Smart Registration Methods for Magnetic Resonance Images in Public Healthcare

Tao Chen[1], Mengxue Yuan[2], Jiajie Tang[2]* and Long Lu[2,3]*

[1] School of Information Technology, Shangqiu Normal University, Shangqiu, China, [2] School of Information Management, Wuhan University, Wuhan, China, [3] Institute of Pediatrics, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China

Brain development and atrophy accompany people's life. Brain development diseases, such as autism and Alzheimer's disease, affect a large part of the population. Analyzing brain development is very important in public healthcare, and image registration is essential in medical brain image analysis. Many previous studies investigate registration accuracy by the "ground truth" dataset, marker-based similarity calculation, and expert check to find the best registration algorithms. But the evaluation of image registration technology only at the accuracy level is not comprehensive. Here, we compare the performance of three publicly available registration techniques in brain magnetic resonance imaging (MRI) analysis based on some key features widely used in previous MRI studies for classification and detection tasks. According to the analysis results, SPM12 has a stable speed and success rate, and it always works as a guiding tool for newcomers to medical image analysis. It can preserve maximum contrast information, which will facilitate studies such as tumor diagnosis. FSL is a mature and widely applicable toolkit for users, with a relatively stable success rate and good performance. It has complete functions and its function-based integrated toolbox can meet the requirements of different researchers. AFNI is a flexible and complex tool that is more suitable for professional researchers. It retains most details in medical image analysis, which makes it useful in fine-grained analysis such as volume estimation. Our study provides a new idea for comparing registration tools, where tool selection strategy mainly depends on the research task in which the selected tool can leverage its unique advantages.

Keywords: MRI, registration technology, feature comparison, SPM, FSL, AFNI

INTRODUCTION

Image registration plays an essential role in image fusion, pattern recognition, and voxel-wise group analysis. Establishing correspondences across brains for comparison and group analysis is almost universally done by registering images to one another directly or *via* a template. Image registration is generally categorized into two groups, rigid and non-rigid registration. Rigid registration only consists of rotation and translation (1). Non-rigid registration is more complex and would take deformation methods, such as affine transformation and spline transform. There are a large number of optimized non-rigid registration algorithms and available registration tools, including

advanced normalization tools (ANTS), analysis of functional neuroimages (AFNI), automated image registration (AIR), Drop, FMRIB's software library (FSL), and statistical parametric mapping (SPM). Thus, a comprehensive evaluation of different registration methods has become a research topic of interest. It is the basis for users to choose the most suitable methods for the specific research problem and for algorithm developers to be better informed theoretically (2). However, it is hard to assess image registration algorithms under different research contexts accurately. According to literature and survey, the main obstacles to building a unified evaluation standard of registration algorithm are as follows:

(1) No ground truth results are supplied in practice. In reality, the registration results are mainly assessed by an experienced expert. The lack of golden truth makes the evaluation research less convincing in terms of data. Also, it is hard to be flexible and straightforward to evaluate the performance of registration methods in a complex reality.

(2) Various parameter configurations for registration algorithms may lead to registered images of different quality. Experienced investigators can set appropriate parameters based on images' information. For the same input image, the difference in parameter setting would cause different registered outcomes (3).

(3) The same image data may be processed with different registration techniques to complete various tasks. Therefore, the data sets should be classified first and then preprocessed by other data registration methods according to different data characteristics.

(4) Data quality and raw data quality substantially impact registration results, which means that the data received by the registration program are quite different. The same algorithm may make a profound difference in the registration outcome when dealing with data with a significant quality difference.

Some researchers have done relevant research about the evaluation and comparison of registration algorithms to investigate this topic. Hellier et al. compared five different fully automated non-linear brain image registration software programs using the quantitative measures (4–6). Klein et al. evaluated 14 non-linear deformation algorithms applied to human brain MRI registration on four datasets and ranked these algorithms (7). Rajagopalan and Pioro investigated disparate voxel-based morphometry (VBM) results between SPM and FSL software in patients with ALS to determine which tool has the best performance of VBM in the ALS disease setting (8). Most recently, Dadar et al. compared publicly available linear MRI stereotaxic registration techniques by viewing the registered images by an expert rater to assess the quality of the registrations (9). At the same time, Viergever et al. made a retrospective view on the past two decades of medical image registration. They mentioned that "validation of registration methods and translation of image registration research results to the clinical practice still is the highlighted research and be more urgent than two decades ago (10)." Meanwhile, Rohlfing provides experimental evidence that registration accuracy

measures such as tissue label overlap scores, image similarity, image difference, or transformation inverse consistency error, even when used in combination, cannot distinguish accurate from inaccurate registrations (11). Some previous accuracy standards are no more extended evaluation standards. Simulated data and a database with expert landmark annotations have been employed for comparison in the last decade to measure accuracy.

The registration of brain MRI with common templates is a long-standing problem. Different individuals' brain shapes and cortical topology are very different, especially in diseases affecting brain morphology and structure, such as brain atrophy. Previous researches mainly focus on registration accuracies, such as overlap, volume similarity, and registration error. However, the choice of registration technology has a more significant impact on image registration results. Few studies have used different registration tools to process image features. This means that the existing studies on the selection of image registration technology are insufficient. The image registration methods used in many studies are not standardized, and the accuracy level after registration is unknown and difficult to be quantified. Such problems will significantly affect the results of follow-up research. Therefore, the evaluation of image registration technology only at the accuracy level is not comprehensive. When facing different image processing projects, researchers need a flexible selection scheme to help them select the appropriate registration technology according to data sets' characteristics and personalized needs. Here, we investigate the performance of three publicly registration tools in two typical different disease cases, aiming to find a suitable registration tool for different disease types and research focus. Our research uses ABIDE and TCGA date sets to conduct experiments and deeply compares the registration performance of FSL, SPM12, and AFNI. In addition, we compare the three mainstream image registration technologies in terms of user experience and recommend them to different groups.

Our research investigates the registered results in two disease types, autism spectrum disorder (ASD), and glioma. Glioma may cause deformation in the morphology of the brain, and physicians can use high-resolution structural MRI images to detect the focus of infection and glioma grades. By contrast, ASD is connected with brain function and emotion and is often a pervasive developmental disorder. The pathogenesis of ASD is related to heredity, so the patient would have ASD at a very young age. The structural MRI and functional MRI images can be used to diagnose ASD in the clinic. We collect the MRI image data from public databases to construct our disease dataset. We do not intend to evaluate accuracy and error by manual labeling and check which are not available in reality. We plan to compare the features and characteristics of different tools' results because registration techniques are designed for follow-up analysis and application. We focus on the difference in registered results under various analyses, such as texture and edge-gradient features, voxel-based morphometry, and the impacts of this inter-method discrepancy on prediction and detection. From the application's perspective, we investigate the difference between three mature and powerful software to provide instructions for

TABLE 1 | Dataset information.

Dataset	ABIDE II-ETH	TCGA-LGG
Slice thickness (mm)	0.9	2.5
No. of slices	180	69–80
Scan matrix	256 ×256	256 × 256
Repetition time (ms)	8.4	10,002
Echo time(ms)	22	125
No. of scans	31	27

different researchers to select a suitable tool and more quickly get better achievements.

The structure of our study is as follows: in Materials and Methods, we describe the data acquisition and registration tools, then the strategy for data processing, and the ABIDE II and TCGA-LGG data set used to assess different registration tools. We focus on the difference in registered results under various analyses, such as texture and edge-gradient features, voxel-based morphometry, and the impacts of these inter-method discrepancies on prediction and detection. As a result, we instruct different researchers to select a suitable tool and more quickly get better achievements by investigating the difference between three mature and powerful software. Finally, in the discussion, we summarize the current work, emphasize the importance of selecting appropriate registration methods, and discuss the limitations and future improvement ideas.

MATERIALS AND METHODS

Data Acquisition and Quality Control

This section introduces the basic information of ABIDE II-ETH and TCGA-LGG. **Table 1** presents the basic information of the database.

(1) ABIDE II-ETH. Autism Brain Imaging Data Exchange II was established to further promote discovery science on the brain connectome in ASD (12). It is a multi-center and multi-scanner study involving 19 sites, ten charter institutions and seven new members, overall donating 1,114 datasets from 521 individuals with ASD and 593 controls (age range: 5–64 years). We select the ETH dataset to finish our research. There are 37 samples in the ETH set (age range 14–31 years), 13 for ASD and 24 for typical control (TC). Then, we manually check the ETH dataset to discard the bad quality image with serious artifacts.

(2) TCGA-LGG. The Cancer Genome Atlas Low-Grade Glioma (TCGA-LGG) data collection is part of a more significant effort to build a research community focused on connecting cancer phenotypes to genotypes by providing clinical images matched to subjects from The Cancer Genome Atlas (TCGA) (13). It contains CT and MRI data collected from 199 patients, and the number of studies in TCGA-LGG is 224. To gain the low-level glioma data with the same scan sequence and quality similar to the autism data set, we manually select

a small part of the subject containing T2-w and T2-Flair sequence MRI data to finish our research.

Registration Tools

SPM12

The SPM has several versions, and the latest is the SPM12 (14, 15). SPM12 is designed to work with MATLAB to run on Windows systems, while other similar tools are always supported by Linux or Mac and are easy to install. If someone is new to imaging, SPM12 is a recommended choice because plenty of tutorials give practical instructions on how to implement the various methodologies.

The spatial normalization method in SPM12 is a unified model based on a probabilistic approach. It combines the functions of image registration, tissue classification, and bias correction in the same generative model. The model is based on a Gaussian mixture and is extended to incorporate a smooth intensity variation and non-linear registration with tissue probability maps (15).

FSL

FSL is a software package developed by the Oxford Center for Functional MRI of the Brain (Oxford University), composed of image analysis and statistical tools for neuroimage data study. Flirt or Fnirt performed image registration in FSL. Flirt uses a multi-start, multiresolution global optimization method (16, 17), tailored explicitly for volumetric registration of brain images to give a reliable estimate of the worldwide minimum given some time restriction. The optimization algorithm uses four resolution levels: $n = 8, 4, 2, 1$ mm. Initially, the procedure starts with a large-scale search at 8 mm resolution (e.g., applying a set of initial rotations). Following this, various local optimizations are performed with multiple starting points in the local neighborhood of the best issues identified in the search. Then a series of multi-start local optimizations at 4 mm resolution is completed.

AFNI

AFNI is a leading software suite of C, Python, R programs, and shell scripts primarily developed to analyze and display anatomical and functional MRI (fMRI) data (18). The software is made to run on virtually a Unix system with X11 and Motif displays and is a widely used tool for image preprocessing. There are plenty of functions to perform registration, e.g., @auto_tlrc, 3dWarp, 3dWarpDrive, 3dAllineate, @toMNI_Awarp, @toMNI_Qwarp. The common function for registration is @auto_tlrc. This function is run by default parameters. Here, we use @auto_tlrc to execute registration (19).

Data Processing

We download data from the website and also check the quality and image label. After eliminating unqualified data, there are 33 for autism data sets (ASD + TC) and 45 for glioma data sets (glioma + TC) as a study data source. It is learned from **Table 2** that the available data formats for each tool are not entirely the same. Firstly, we convert images of DICOM format

TABLE 2 | Format requirements.

Public tools data format	DICOM	ANALYZE	NIFTI_PAIR	NIfTI-1	NIFTI_GZ	HEAD/BRIK	MINC	CTF	ASCII
SPM	✓	✓	✓	✓					
FSL		✓	✓	✓	✓				
AFNI	✓	✓	✓	✓	✓	✓	✓	✓	✓

FIGURE 1 | Comparison framework.
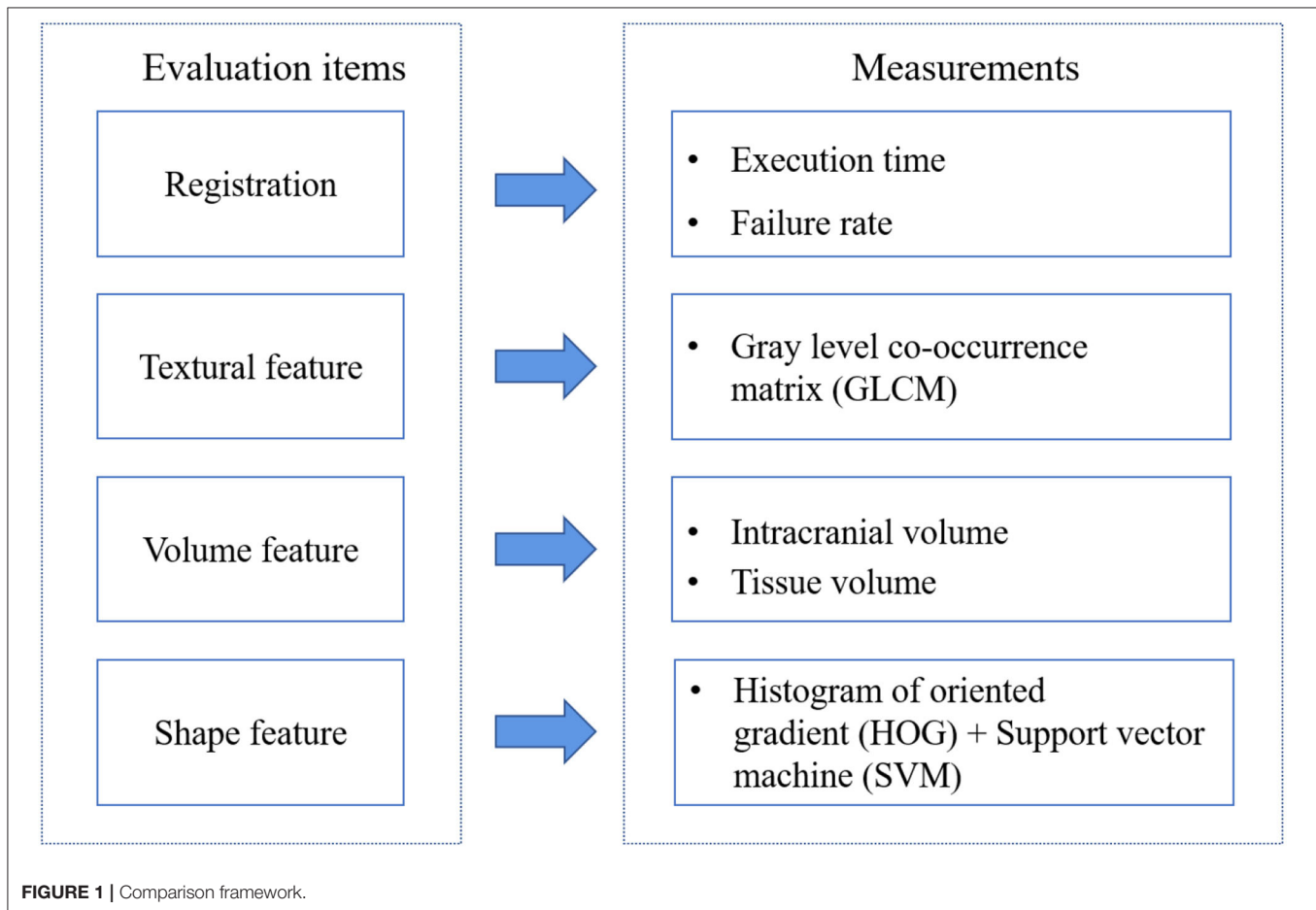
into NIFTI-1 format. NIFTI-1 format is also the most commonly used international MRI format.

To explore the features difference among the three tools registered results, we construct the comparison framework shown in **Figure 1**. After organizing the dataset, we execute the registration step by AFNI, SPM, and FSL. T1-w and T2-FLAIR images were all registered to the standard space. SPM12 operations are finished on MATLAB R2012a. FSL and AFNI are installed and run in Ubuntu14 on a virtual machine.

There are three primary features for images: color, texture, and shape. Here, MRI images are gray images without color features. We consider volume as one of the evaluation items. Then we will introduce each item and its measurements.

Textural Feature
A textural feature is a diagnostic tool for physicians when reading MRI images. Learning from this, Zacharaki et al. extract MRI texture and shape features, constructing the feature vector to train a support vector machine (SVM) classifier to classify brain tumor type and grade [20]. A textural feature is usually employed for tumor grading in tumor studies because it provides fine-grained, repeatable information. Besides, for autism, commonly diagnosed by behavioral, chemical, clinical, structural, and functional changes in the brain, MRI texture represents a new image feature, which can supplement the traditional image features (such as volume measurement). Vidhusha and Anandhan focused on extracting texture features for autistic and control subjects, and validated them using neural classifiers to achieve automatic detection of autism [21]. Textural features are extracted from original MRI ROI images rather than registered results, but here, we explore the textural difference between results suffering different processing pipelines, and we will do a single statistical analysis to discuss the significant difference.

The textural feature has several extraction methods, including statistical, model, signal processing, and geometric forms. Here, we choose the gray level co-occurrence matrix (GLCM) proposed by Haralick et al. regarded as one of the most promising texture analysis methods (22, 23). It estimates the image properties related to second-order statistics. GLCM is a joint distribution describing the gray level of two pixels with a particular spatial positional relationship. Five derived variables were calculated from the GLCM matrix, contrast, correlation, energy, homogeneity, and IDM to investigate the difference between textural features. Then these computations will be described in detail next.

Contrast. The deeper the texture groove, the greater the contrast and the clearer the visual effect; conversely, the contrast is small, the groove is shallow, and the effect is blurred.

$$contrast = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i-j)^2 p(i,j) \tag{1}$$

Correlation. It represents the consistency of the image texture. If there is a horizontal direction texture in the image, the COR of the horizontal direction matrix is larger than the COR value of the remaining matrix.

$$correlation = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j)ij - \mu_x\mu_y}{\sigma_x(i)\sigma_y(j)} \tag{2}$$

Energy. It is the sum of the squares of the gray level co-occurrence matrix values, so it is also called angular second moment, which reflects the uniformity of image gray distribution and texture thickness. A large energy value indicates a textured pattern that is more uniform and regular.

$$energy = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (p(i,j))^2 \tag{3}$$

Homogeneity. It reflects the density of elements in GLCM relative to the diagonal distribution of GLCM.

$$homogeneity\ 1 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i,j)}{1 + |i-j|} \tag{4}$$

Inverse Difference Moment. IDM measures the local homogeneity of an image and reports the inverse difference moment of an image. IDM weights are the inverse of contrast weights. It has the value that determines whether the print is textured or non-textured.

$$IDM = \sum_{k=0}^{N_g-1} \frac{p_{x-y}(k)}{1 + k^2} \tag{5}$$

Volume Feature

Volume could provide significant information on pathological study and can capture volume change in the brain. Quantitatively detecting the density and volume of brain tissue at the voxel level can reflect the differences in brain tissue components and characteristics in different brain regions. The latest study explores that the inter-method discrepancies in brain volume estimation may drive inconsistent findings in autism (3). In this part, we employ SPM12 as the brain volume estimation tool. The images are segmented into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) volume in the native space to estimate the volume by using the New Segment tool of SPM12 (6). Spm_get_volumes script is used to calculate the tissue volumes using c1, c2, and c3 images corresponding to native space tissue maps of GM, WM, and CSF, respectively. Native space volumes are selected to minimize volume changes due to spatial transformations. Total intracranial volume (TIV) is calculated as the sum of the GM, WM, and CSF volumes in the native space of the structural MRI images.

Shape Feature

We calculate textural matrix to understand the textural feature and compare the brain tissue volume estimation, but it is hard to compare shape differences for shape features. Thus, we employ machine learning algorithms using shape features as the input and record the prediction accuracy to evaluate the shape feature performance. Based on the shape features extracted from MRI, machine learning algorithms could find the standard classifier function to distinguish the healthy and disease cases. To capture shape features in MRI images, we extracted the histogram of oriented gradient [HOG; (24)], a prominent local image feature widely used in computer vision applications, from MRI images to establish the feature sets. The HOG technique counts occurrences of gradient orientation in localized portions of an image. When studying the brain tumor, the HOG could describe local object appearance and shape within an image by distributing intensity gradients or edge directions. Thus, this descriptor can detect the abnormal edge of the tumor issues and help physicians find the lesions. About ASD, the HOG descriptor may also find some markers to distinguish the healthy and disease cases.

Then, the classification method is the support-vector machine (SVM), the supervised learning model. The combination of HOG and SVM is the classical strategy in object detection, especially in human detection (24). In the model, the cross-validation (CV) method is used to evaluate the performance of classifiers trained by different MRI data sets registered by three tools. The widely used methods in brain image analysis include leave-one-out cross-validation (25, 26), leave-two-out cross-validation (27–29), k-fold cross-validation (1), and stratified k-fold cross-validation (30, 31). The cross-validation results are conducted from the strategy that divides the raw data into K groups (K-fold), and each subset of data is used as a verification set. The remaining K-1 subset data is used as a training set. Then the K models, respectively, evaluate the results in the verification set, which effectively utilizes limited data, and the results are as close as possible to the model's performance on the test set. According

to review studies and previous research, we use stratified ten-fold cross-validation to measure classification performance. All operations are performed in python3.7 in Ubuntu14 on a virtual machine.

RESULTS

Time and Failure Rate

Execution time is a crucial factor that needs to consider, which can reflect the effectiveness of tools. During the registration step, we calculate the time in each registration pipeline. **Table 3** and **Figure 2** show that AFNI gets the best speed performance and then SPM12 while FSL is the slowest.

Registration is also regarded as an optimization process, which means there are failures in this step. We also collect the failure rate in the registration step. From **Table 3**, the failure rate in AFNI is higher than in FSL, and in FSL it is higher than in SPM. Auto-registration in SPM12 is based on prior knowledge using the tissue probability map to find the optimum transformation to achieve high and accurate registration. What is more, in SPM12, there is no feedback mechanism, and the final transformed results will be returned, whether good or bad. By contrast, FSL and AFNI both have the mechanism that if the cost function cannot be optimized, they will return the failure message

and some additional suggestions to help the user optimize the registration command and get better results. So, AFNI and FSL have improvable abilities while SPM12 is poor.
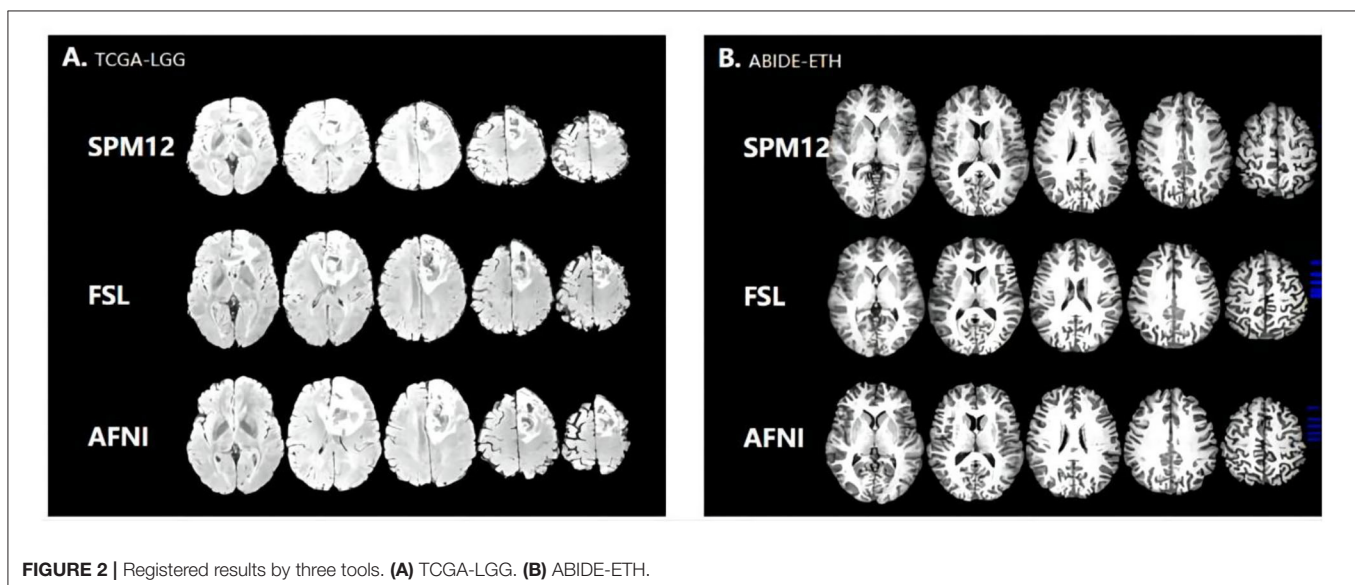
The purpose of our study is not to attempt to find the most accurate one that has been investigated in many previous studies (7, 9, 32–34). We compare the registered results in the image to explore the degree of change in the features and structures made by the three registration tools and find the difference between these changes to help the user select the registration tool based on their registration results difference and research purpose.

Texture Feature

We extract valuable features from the registered images by different tools to find the difference. Brain regions were extracted using the brain extraction tool [BET; (35)]. We extracted textural features from disease data sets to study the difference between registration tools and disease types (36, 37). Input images are brain-extracted MRI images to gain whole-brain texture GLCM matrix and calculate the five features, i.e., contrast, correlation, energy, homogeneity, and inverse difference moment (IDM).

Figure 3 shows the results of the comparison of texture features. The value distribution is similar among SPM, FSL, and AFNI and shows no significant difference. But FSL gets a higher value, implying that FSL is more inclined to retain more uniform and regular texture patterns. AFNI has a higher correlation value corresponding to its biggest size of MRI holding more pixels and details that could be measured more similar pattern.

In the statistical analysis, SPM12 values shows significant difference with AFNI ($p = 0.017314065 < 0.05$, 95%) and FSL ($p = 0.046368617 < 0.05$, 95%), indicating that SPM12 registered images have more sharp textural changes. A larger contrast value correlates with a greater disparity in intensity values among neighboring voxels, suggesting that SPM12 processed results have more sharp intensity changes and deeper texture grooves. Next, the inverse difference moment, also known as

TABLE 3 | Time acquisition and failure rate.

	ABIDE II-ETH		TCGA-LGG	
	Time (Average)	Failure rate (%)	Time (Average)	Failure rate (%)
SPM12	94 s	0	100 s	0
FSL	110 s	0	4 min	16
AFNI	78 s	0	86 s	20

FIGURE 2 | Registered results by three tools. **(A)** TCGA-LGG. **(B)** ABIDE-ETH.
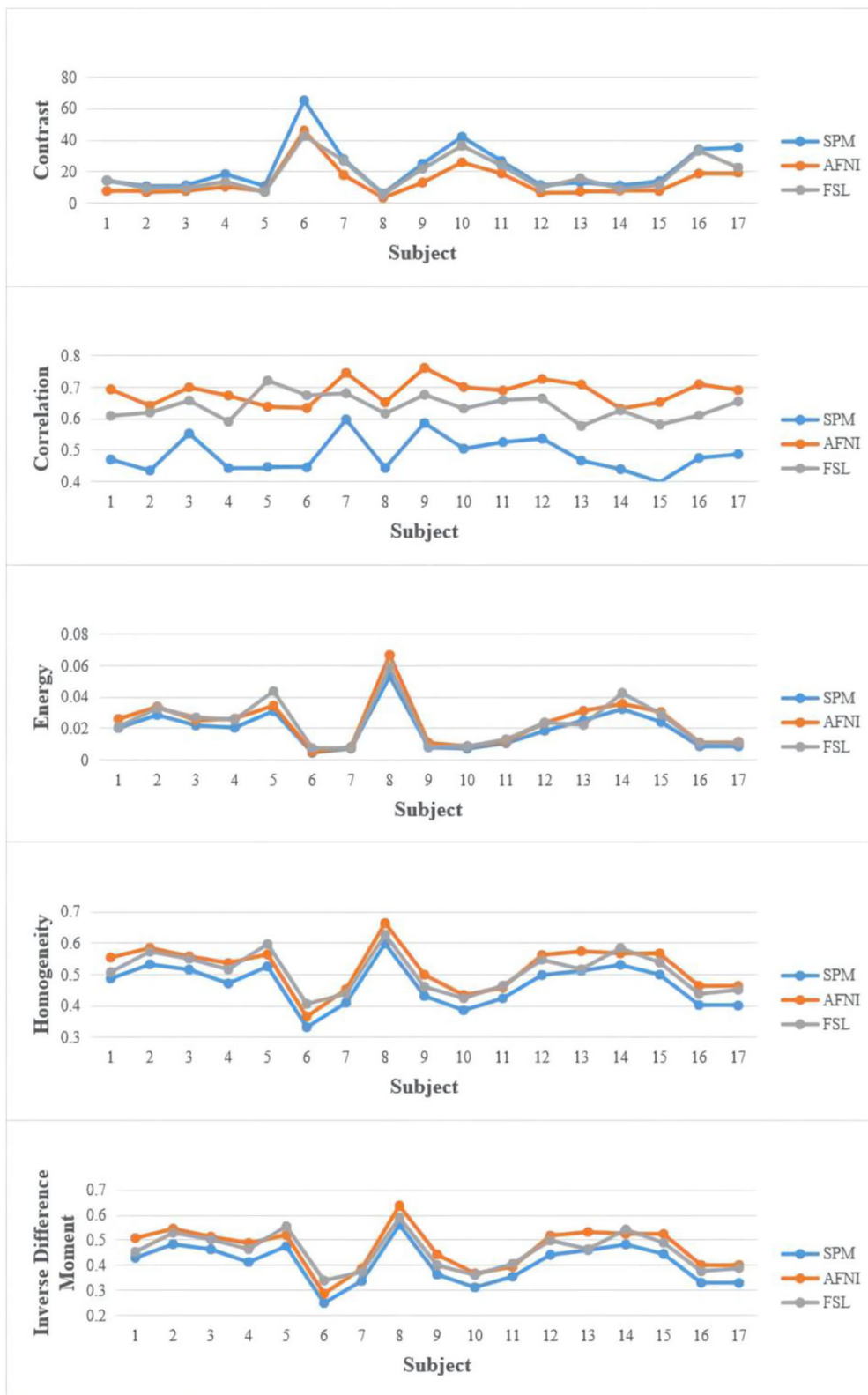
FIGURE 3 | Texture feature extracted from the dataset.

TABLE 4 | Brain volume estimation results.

	TIV (mean)	GM% (mean)	WM% (mean)	CSF% (mean)
TCGA-LGG	1.5364	39.89	45.17	14.94
SPM (TCGA)	1.9012	35.41	41.61	22.97
AFNI (TCGA)	1.8936	36.08	45.13	18.79
FSL (TCGA)	1.5123	33.62	43.89	22.49
ETH	1.2960	58.79	32	9.21
SPM (ETH)	1.8776	54.21	30.17	15.02
AFNI (ETH)	1.8977	55.63	29.63	14.74
FSL (ETH)	1.5641	55.81	28.01	16.19

ETH = ABIDE II-ETH.

the inverse variance, reflects the clarity and regularity of texture. The images with clear, regular, easy to describe texture would have a larger value demonstrating that these images are more in line with human feelings. The AFNI is clearer than the other two data sets in **Figure 2**, which is consistent with our calculated values. SPM12 also has significant difference with AFNI ($p = 0.018590934 < 0.05$, 95%) and FSL ($p = 0.044412676 < 0.05$, 95%) as well. We guess the reason is that AFNI retains more detailed information, including uniform and sharp textural change, while SPM12 abandons these details and hold sharp textural feature.

Brain Volume Estimation

Some brain diseases may give rise to the volume change in the brain, such as edema, which is obvious in the image. We calculate the TIV and the percentage of gray matter, white matter, and CSF in the original and registered structural MRI images. We set the volume estimation from the MRI in the original space as the standard result. From **Table 4** and **Figure 4**, AFNI shows the closest results with similar GM, WM, and CSF percentages, and sometimes overestimates the total volume. The whole-brain volume estimated by SPM was always higher than the standard TIV. FSL estimates of CSF are the highest.

Shape Feature-Based Classifier

Finally, we compare the evaluated accuracy of the cross-validation. For the ASD dataset, the number of healthy and disease cases is 22 and 11, respectively. For the glioma dataset, the number of healthy and disease cases is 22 and 17, respectively, and the k that we use in k-fold is 10.

From **Table 5**, we can observe the difference between the registration tools. The performance of SPM shows the highest accuracy, while AFNI is better than FSL. Since the HOG feature represents the shape information by edge directions, which collects contrast information, this result is consistent to compare results found in the texture feature part. MRI images registered by SPM retain richer contrast sources, which may mean that SPM is a suitable preprocessing tool when dealing with structural problems. The SPM performs best in our structural HOG feature model.

DISCUSSION

To estimate the performance of registration algorithms of SPM, FSL, and AFNI, we calculate the registration time, compare the texture feature extracted from skull-striped structural MRI images, and investigate the inter-methods discrepancies in brain volume estimate. Finally, we compare these registered images' applied performance in the classification model and we get the conclusion that SPM12 is usable for newcomers or employed when dealing with basic and normal processing workflow because it is like a black box that users cannot know the registration process details. FSL is a powerful toolkit for research. Its function-based integrated toolbox facilitates researchers for the different hierarchy of requirements. However, FSL's semi-integrated design cannot meet scholars' requirements for function module combinations and workflow model optimization. In contrast, AFNI is cumbersome and complex to use. However, it can provide some personalized customization and allow some threshold adjustments, which is very useful for many professional researchers, but it may be difficult for novices to use.

According to our analysis conclusion, tool selection mainly depends on research tasks. If the focus of studies is only MRI images' single feature, it is wise to choose the feature-targeted tools, for example, the contrast feature of texture feature for SPM, and if studies require multiple features or detailed features, AFNI will be a good choice. When involving multi-modal MRI sequences, like DWT, PET, and DTI, FSL would become the first choice. For glioma and autism spectrum disorder, the glioma is more sensitive to contrast features, so SPM12 may perform better registration results when studying the glioma dataset. When using the shape features like HOG, SIFT, or LBA, SPM12 may have the best performance on comparing AFNI and FSL. Tumor detection, prediction, and classification are likely to employ these structural and edge-related features, and so SPM12 will be an advisable process tool with stable performance and feature-sensitive advantages, facilitating tumor study.

Image registration research is still progressing steadily, and new research results are emerging. At present, there is still no unified evaluation criterion. Experience and habits are still the main factors affecting people's choice of methods and tools. Our study provides a new idea for the comparison of registration tools. Using data sets from different diseases, we study the differences in texture features, structure, and classification performance of the three registration tools, discuss their characteristics, and analyze the causal correlation between features and classification models. Our study is not without limitations. The focus of this study is only on structural images, while the registration between functional images and other modal data has not been investigated in detail. Secondly, only three kinds of public registration tools are selected, and some tools in current research are omitted. In addition, the default parameter settings are adopted in the study. Still, in the specific analysis, different parameter settings will significantly impact the results, and the appropriate parameter configuration can optimize the processing results. There is some research on MRI registration
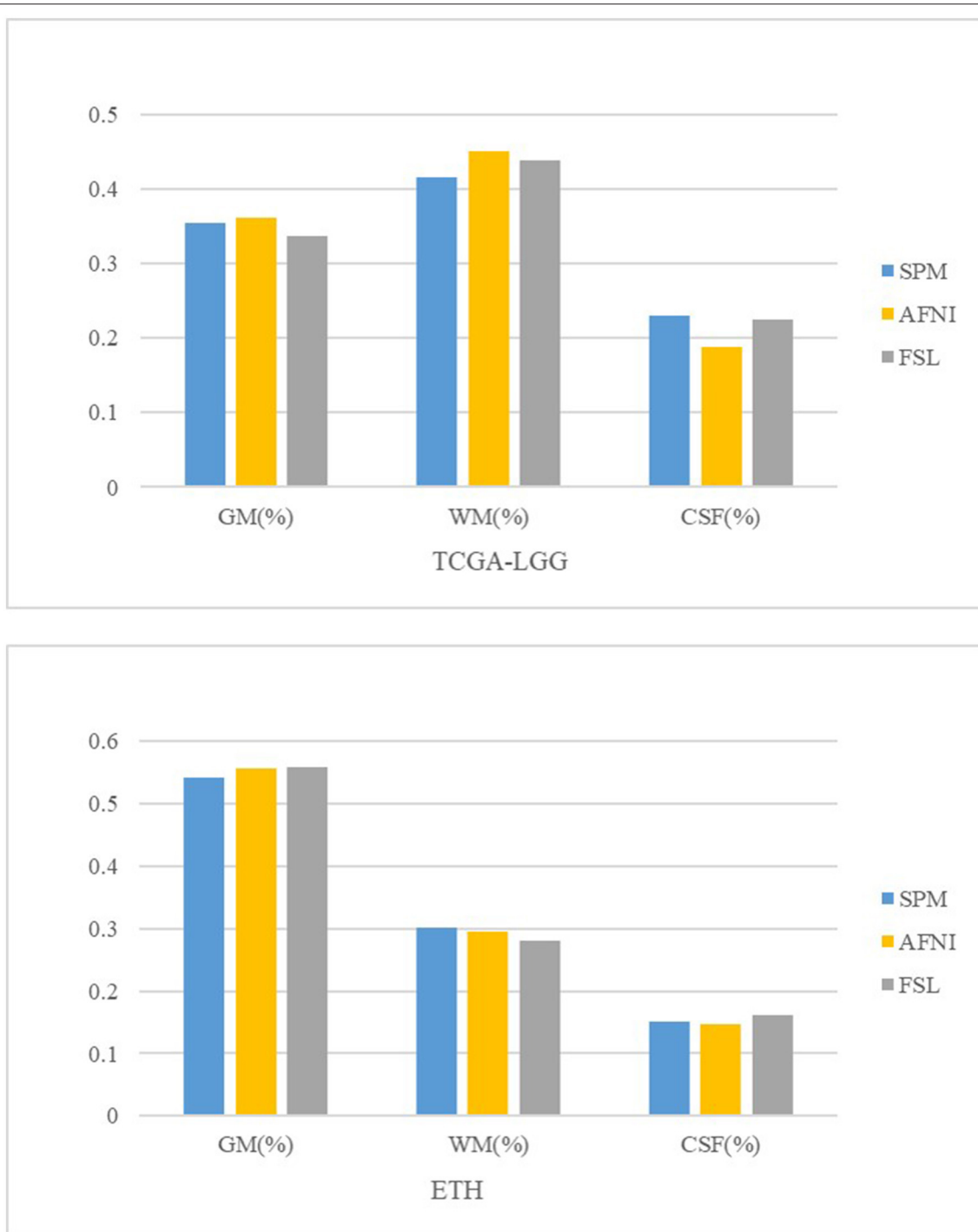
FIGURE 4 | Brain tissue volume comparison.

using a deep learning algorithm, which may achieve better performance than traditional MRI registration methods. In the next step, we need to compare and analyze the registration results of these methods (38, 39).

CONCLUSIONS

Although SPM has the most stable performance, it also produces poor registered results when dealing with low-quality data

TABLE 5 | Prediction results on two datasets.

ABIDE II-ETH dataset	SPM %	AFNI %	FSL %
Accuracy (high)	78.79	72.73	63.64
Accuracy (low)	69.70	63.64	54.55
TCGA-LGG dataset	**SPM %**	**AFNI %**	**FSL %**
Accuracy (high)	92.31	88.83	81.17
Accuracy (low)	86.24	80.52	72.08

sets. According to its advantages and drawbacks, SPM12 has a stable speed and success rate but does not guarantee the processed results, especially when handling low-quality images. FSL is a mature and widely applicable toolkit for employers, with relatively stable performance, a success rate, and complete functions. Its graphical user interface provides basic processing flow operation, while multi-parameter settings and failure feedback mechanisms can realize complex user processing and meet the optimization requirements. It is an excellent toolbox for beginners and proficient hands. Last, AFNI, with quite a fast speed and complex modules, can meet various requirements. But it is not friendly to newcomers, even though it offers an integrated processing function-similar function and modules, unique data format, and command-line operation, all confusing to beginners.

The texture feature is significant in image registration. SPM registration enhances the contrast feature of the texture feature, giving up other detailed information. FSL registration contains enough contrast information and most of the details. AFNI registration is apparent to view details in the brain structure. Due to plentiful details, too much information may make finding key features to differentiate brain structure by single textual features. Therefore, SPM registered image is easy to distinguish the different and distinct tissue types in the images. MRI images processed by AFNI and FSL are more suitable for complex and fine-grained investigations.

In conclusion, AFNI will be a good product application for its immediate and quick response with processed results but requires extra steps to ensure quality. It also plays an essential role in academics and research with its excellent flexibility, which could quickly test various combined processing steps in research. It is also the only MRI registration tool that supports all formats. SPM12 always works as a guide tool for newcomers to medical imaging. A rich online teaching resource and simple and easy-to-learn operations make everyone accomplish the basic processing flow and get stable results. Its support for the Windows environment is the main reason for its popularity. FSL combined advantages of both, providing GUI for basic and simple processing flow and customized parameters setting for flexible and optimized procedure. In addition, FSL is a suitable tool for other MRI sequences like DWT and has been widely used in numerous studies.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: https://www.nitrc.org/frs/downloadlink.php/9090 ABIDE II-ETH (39); https://wiki.cancerimagingarchive.net/display/Public/TCGA-LGG TCGA-LGG (224).

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

REFERENCES

1. Zhe L, Deng D, Guang-Zhi W. Accuracy validation for medical image registration algorithms: a review. *Chin Med Sci J.* (2012) 27:176–81. doi: 10.1016/S1001-9294(14)60052-3

2. 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) 33:2039–65. doi: 10.1109/TMI.2014.2330355

3. Katuwal GJ, Baum SA, Cahill ND, Dougherty CC, Evans E, Evans DW, et al. Inter-Method discrepancies in brain volume estimation may drive inconsistent findings in autism. *Front Neurosci.* (2016) 10:439. doi: 10.3389/fnins.2016.00439

4. Hellier P, Barillot C, Corouge I, Gibaud B, Goualher GL, Collins DL, et al. Retrospective evaluation of inter-subject brain registration. In: *Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Utrecht: Springer-Verlag (2001). p. 258–65. doi: 10.1007/3-540-45468-3_31

5. Hellier P, Barillot C, Corouge I, Gibaud B, Le Goualher G, Collins DL, et al. Retrospective evaluation of intersubject brain registration. *IEEE Trans Med Imaging.* (2003) 22:1120–30. doi: 10.1109/TMI.2003.816961

6. Hellier P, Ashburner J, Corouge I, Barillot C, Friston KJ. Inter subject registration of functional and anatomical data using Spm. In: *Proceedings of the 5th International Conference on Medical Image Computing and Computer-Assisted Intervention-Part II*. Springer-Verlag (2002). p. 590–7. doi: 10.1007/3-540-45787-9_74

7. Klein A, Andersson J, Ardekani BA, Ashburner J, Avants B, Chiang MC, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain Mri registration. *NeuroImage.* (2009) 46:786–802. doi: 10.1016/j.neuroimage.2008.12.037

8. Rajagopalan V, Pioro EP. Disparate voxel based morphometry (Vbm) results between Spm and Fsl Softwares in Als patients with frontotemporal dementia: which Vbm results to consider? *BMC Neurol.* (2015) 15:32. doi: 10.1186/s12883-015-0274-8

9. Dadar M, Fonov VS, Collins DL. A comparison of publicly available linear Mri stereotaxic registration techniques. *NeuroImage.* (2018) 174:191–200. doi: 10.1016/j.neuroimage.2018.03.025

10. Viergever MA, Maintz JBA, Klein S, Murphy K, Staring M, Pluim JPW. A survey of medical image registration - under review. *Med Image Anal.* (2016) 33:140–4. doi: 10.1016/j.media.2016.06.030

11. Rohlfing T. Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unreliable. *IEEE Trans Med Imaging.* Pune (2012) 31:153–63. doi: 10.1109/TMI.2011.2163944

12. Di Martino A, Yan CG, Li Q, Denio E, Castellanos FX, Alaerts K, et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol Psychiatry.* (2014) 19:659–67. doi: 10.1038/mp.2013.78

13. Tomczak K, Czerwińska P, Wiznerowicz M. The cancer genome atlas (Tcga): an immeasurable source of knowledge. *Contemp Oncol.* (2015) 19:A68–77. doi: 10.5114/wo.2014.47136

14. Ashburner J, Neelin P, Collins DL, Evans A, Friston K. Incorporating prior knowledge into image registration. *NeuroImage.* (1997) 6:344–52. doi: 10.1006/nimg.1997.0299

15. Ashburner J, Friston KJ. Nonlinear spatial normalization using basis functions. *Hum Brain Map.* (1999) 7:254–66. doi: 10.1002/(SICI)1097-0193(1999)7:4<254::AID-HBM4>3.0.CO;2-G

16. Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage.* (2002) 17:825–41. doi: 10.1006/nimg.2002.1132

17. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. Fsl. *NeuroImage.* (2012) 62:782–90. doi: 10.1016/j.neuroimage.2011.09.015

18. Cox RW. Afni: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res Int J.* (1996) 29:162–73. doi: 10.1006/cbmr.1996.0014

19. Cox RW. Afni: what a long strange trip it's been. *NeuroImage.* (2012) 62:743–7. doi: 10.1016/j.neuroimage.2011.08.056

20. Zacharaki EI, Wang S, Chawla S, Soo Yoo D, Wolf R, Melhem ER, et al. Classification of brain tumor type and grade using mri texture and shape in a machine learning scheme. *Magnet Resonan Med.* (2009) 62:1609–18. doi: 10.1002/mrm.22147

21. Vidhusha S, Anandhan K, editors. Analysis and evaluation of autistic brain Mr images using learning vector quantization and support vector machines. In: *International Conference on Industrial Instrumentation & Control.* (2015). doi: 10.1109/IIC.2015.7150873

22. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *Stud Media Commun.* (1973) 3:610–21. doi: 10.1109/TSMC.1973.4309314

23. Haralick RM. Statistical and structural approaches to texture. *Proc IEEE.* (2005) 67:786–804. doi: 10.1109/PROC.1979.11328

24. Dalal N, Triggs, B, editors. Histograms of oriented gradients for human detection. In: *IEEE Computer Society Conference on Computer Vision & Pattern Recognition.* (2005). doi: 10.1109/CVPR.2005.177

25. Toews M, Wells W III, Collins DL, Arbel T. Feature-Based morphometry: discovering group-related anatomical patterns. *NeuroImage.* San Diego, CA (2010) 49:2318–27. doi: 10.1016/j.neuroimage.2009.10.032

26. Castellani U, Rossato E, Murino V, Bellani M, Rambaldelli G, Perlini C, et al. Classification of schizophrenia using feature-based morphometry. *J Neural Transm.* (2012) 119:395–404. doi: 10.1007/s00702-011-0693-7

27. Ecker C, Rocha-Rego V, Johnston P, Mourao-Miranda J, Marquand A, Daly EM, et al. Investigating the predictive value of whole-brain structural Mr scans in autism: a pattern classification approach. *NeuroImage.* (2010) 49:44–56. doi: 10.1016/j.neuroimage.2009.08.024

28. Lao Z, Shen D, Xue Z, Karacali B, Resnick SM, Davatzikos C. Morphological classification of brains via high-dimensional shape transformations and machine learning methods. *NeuroImage.* (2004) 21:46–57. doi: 10.1016/j.neuroimage.2003.09.027

29. Etzel JA, Valchev N, Keysers C. The impact of certain methodological choices on multivariate analysis of fmri data with support vector machines. *NeuroImage.* (2011) 54:1159–67. doi: 10.1016/j.neuroimage.2010.08.050

30. Richiardi J, Eryilmaz H, Schwartz S, Vuilleumier P, Van De Ville D. Decoding brain states from Fmri connectivity graphs. *NeuroImage.* (2011) 56:616–26. doi: 10.1016/j.neuroimage.2010.05.081

31. Acharya UR, Ng EY, Tan JH, Sree SV. Thermography based breast cancer detection using texture features and support vector machine. *J Med Syst.* (2012) 36:1503–10. doi: 10.1007/s10916-010-9611-z

32. Klein A, Ghosh SS, Avants B, Yeo BT, Fischl B, Ardekani B, et al. Evaluation of volume-based and surface-based brain image registration methods. *NeuroImage.* (2010) 51:214–20. doi: 10.1016/j.neuroimage.2010.01.091

33. Yahaya K, Zawiah S, Zadry H. Medical image registration: Comparison and evaluation of nonlinear transformation algorithms. In: *2010 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES).* (2010). doi: 10.1109/IECBES.2010.5742209

34. Robbins S, Evans AC, Collins DL, Whitesides S. Tuning and comparing spatial normalization methods. *Med Image Anal.* (2004) 8:311–23. doi: 10.1016/j.media.2004.06.009

35. Smith SM. Fast robust automated brain extraction. *Hum Brain Map.* (2002) 17:143–55. doi: 10.1002/hbm.10062

36. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* (2017) 77:e104–7. doi: 10.1158/0008-5472.CAN-17-0339

37. Cavedo E, Suppa P, Lista S, Lange C, Hampel H, editors. Impact of the Mri acquisition protocol on the hippocampus volume obtained by two methodologically distinct methods for fully automated volumetry: Spm12 and Fsl-first. *Alzheimers Dement.* (2015) 11:P689. doi: 10.1016/j.jalz.2015.06.1530

38. Zhu Z, Cao Y, Qin C, Rao Y, Ni D, Wang Y. Unsupervised 3d end-to-end deformable network for brain Mri registration. *Ann Int Conf IEEE Eng Med Biol Soc.* (2020) 2020:1355–9. doi: 10.1109/EMBC44109.2020.9176475

39. Zhang F, Wells WM, O'Donnell LJ. Deep diffusion mri registration (Ddmreg): a deep learning method for diffusion Mri registration. *IEEE Trans Med Imaging.* (2021). doi: 10.1109/TMI.2021.3139507