Pattern Recognition 100 (2020) 107171

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/patcog

# Deep morphological simplification network (MS-Net) for guided registration of brain magnetic resonance images

Dongming Wei[a], Lichi Zhang[a], Zhengwang Wu[b], Xiaohuan Cao[c], Gang Li[b], Dinggang Shen[b,d,*], Qian Wang[a,*]

[a] Institute for Medical Imaging Technology, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200030, China
[b] Department of Radiology and BRIC, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA
[c] Shanghai United Imaging Intelligence Co., Ltd., Shanghai, China
[d] Department of Brain and Cognitive Engineering, Korea University, Seoul 02841, Republic of Korea

ARTICLE INFO

*Article history:*
Received 3 May 2019
Revised 27 November 2019
Accepted 15 December 2019
Available online 24 December 2019

*Keywords:*
Deformable image registration
Deep learning
Anatomical complexity

ABSTRACT

Deformable brain MR image registration is challenging due to large inter-subject anatomical variation. For example, the highly complex cortical folding pattern makes it hard to accurately align corresponding cortical structures of individual images. In this paper, we propose a novel deep learning way to simplify the difficult registration problem of brain MR images. Specifically, we train a morphological simplification network (MS-Net), which can generate a *simple* image with less anatomical details based on the *complex* input. With MS-Net, the complexity of the fixed image or the moving image under registration can be reduced gradually, thus building an individual (simplification) trajectory represented by MS-Net outputs. Since the generated images at the ends of the two trajectories (of the fixed and moving images) are so simple and very similar in appearance, they are easy to register. Thus, the two trajectories can act as a bridge to link the fixed and the moving images, and guide their registration. Our experiments show that the proposed method can achieve highly accurate registration performance on different datasets (*i.e.*, NIREP, LPBA, IBSR, CUMC, and MGH). Moreover, the method can be also easily transferred across diverse image datasets and obtain superior accuracy on surface alignment. We propose MS-Net as a powerful and flexible tool to simplify brain MR images and their registration. To our knowledge, this is the first work to simplify brain MR image registration by deep learning, instead of estimating deformation field directly.



## 1. Introduction

Deformable image registration [1,2] aims to estimate the deformation field, following which the moving image can be warped to the space of the fixed image. This technique plays an important role in medical image analysis, as it can help build anatomical correspondences across images and facilitate the subsequent analysis. Whereas registration is often perceived as an optimization problem [3], the deformation field needs to be optimized iteratively, with certain smoothness regularization, to maximize the similarity [4,5] between the fixed and the moving images. Commonly used methods for brain magnetic resonance (MR) image registration include AIR [6], ART [7], SyN [8], HAMMER [9,10], Demons [11–13], SPM [14], DRAMMS [15], DROP [16], CC/MI/SSD-FFD [17], FNIRT [18], LDDMM [19], *etc.* Although comprehensive comparisons of these methods are reported in [20,21], it is still difficult to assert the best algorithm for a certain application especially when dealing with diverse datasets.

The large anatomical variation across different images is a great challenge to image registration. In brain MR images, the cortical folding patterns are known to be complex with high inter-subject variation. Whereas imaging-based studies require highly accurate alignment of the corresponding neuroanatomies across different subjects, most existing methods struggle in estimating deformation fields to register tiny structures (*e.g.*, cortical areas) precisely. For example, one may evaluate the overlap ratio between the same anatomical structures of the fixed image and the warped moving image as a metric of the registration quality [22]. Although a reduced smoothness constraint of the deformation field may increase

* Corresponding authors at: Shanghai Jiao Tong University and University of North Carolina at Chapel Hill.
*E-mail addresses:* dongming.wei@sjtu.edu.cn (D. Wei), dgshen@unc.edu.cn (D. Shen), wang.qian@sjtu.edu.cn (Q. Wang).

https://doi.org/10.1016/j.patcog.2019.107171
0031-3203/

the overlap metric, the topology-preserving property of the deformation fields would then be at high risk to be destroyed, leading to a possible failure to the entire registration task. Thus, a high-performance registration method, which could consistently work well for different datasets and tasks with minimal parameter tuning, is of great interest to the community.

To address the concern of large anatomical variation, several works have introduced intermediate images into the deformation pathway between the fixed and the moving images, including mono-modal registration [23,24] and multi-modal registration [25,26]. A manifold is often instantiated to account for the distribution of the imaging data. Then, the very long pathway connecting two images that are far away on the manifold is divided into several short segments by the intermediate images, each of which corresponds to an easier-to-estimate deformation field. However, it is non-trivial to create the manifold. The imaging data are high-dimensional, implying that a sufficient number of (intermediate) images is necessary to model the complex distribution of the image population. Meanwhile, a global image similarity metric is needed by the manifold, yet the metric is challenging to design and often fails to describe local anatomical variation effectively [27].

In last decades, machine learning has become a frequently used tool to image registration. In particular, convolution neural network (CNN) is recently employed to directly predict the deformation field from a pair of fixed/moving images in [28], where the ground truth for training is acquired by SyN and Demons. The initial momenta for LDDMM can also be predicted by CNN as in [29]. Moreover, it is shown that the deformation field can be predicted from the input images through a deep network trained without supervision [30–33]. That is, the image similarity metric and the smoothness regularization can jointly guide the training of the network in back-propagation. The adversarial strategy is also used to impose the regularization such that the local minima in registration can be better avoided [34]. Although registration can be solved in black-box by the powerful computation capability of deep learning, most previous works fail to consider the high complexity of the image manifold. The issue of large anatomical variation is still challenging to many brain MR image registration applications.

In this paper, we propose a novel deep learning way to simplify brain MR image registration. Specifically, we train CNNs to reduce the anatomical complexity of the fixed and the moving images (*i.e.*, cortical folding pattern), which is a bottleneck to the registration task. We derive a trajectory from fixed/moving image, which consists of a series of images with gradually-reduced anatomical complexity. As the two trajectories approach to the ends, the anatomical complexity of the simplified fixed/moving images becomes low, implying that the two images become similar with each other in the simplified morphological space. The two simplified images are then easier to be registered compared to the case of directly registering the original fixed and moving images. In this way, the two trajectories generated by deep learning can act as a bridge to link the fixed/moving images. By composing multiple deformable registration tasks along the two trajectories, the moving image can finally be registered with the fixed image accurately and reliably.

Our method is unique as it breaks the barrier of the complex image manifold in deformable registration of brain MR images, and provides intermediate guidance through deep learning for the first time. To effectively simplify the complexity of brain MR images, we train the morphological simplification networks (MS-Nets) particularly. MS-Nets are able to generate a set of T1 images given an input image. The generated images are gradually simplified in terms of the cortical folding patterns, while no segmentation, parcellation, or cortical surface reconstruction is needed for a test image. To our knowledge, our method is the first to simplify brain MR image registration via deep learning, instead of estimating the deformation field directly as a black-box. Moreover, our experiments show that the MS-Nets trained with a certain dataset are robust to transfer to other new datasets, making our method highly adaptable to many clinical applications.

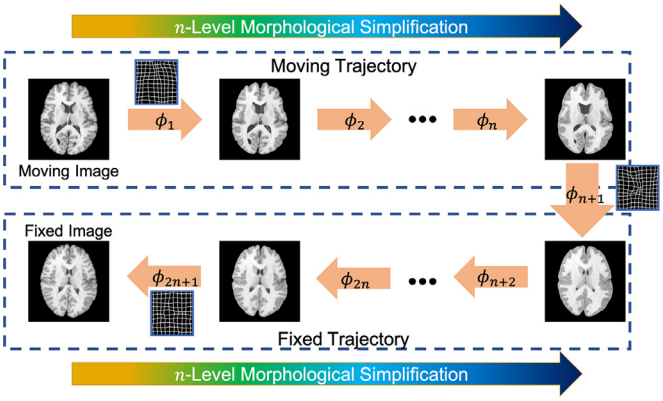


**Fig. 1.** Illustration of the pipeline of the proposed method. The fixed and the moving images reduce their anatomical complexity gradually through $n$-level morphological simplification process by deep learning, which results in the fixed/moving (image) trajectories, respectively. At the ends of the two trajectories, the fixed and the moving images become similar in appearance, such that their registration can be easily completed. Finally, the registration between the moving and the fixed images can be attained by concatenating multiple deformation fields (*i.e.*, $\{\phi_i | i = 1, \ldots, 2n+1\}$ each of which is denoted by an arrow.

## 2. Method

We propose to simplify the registration of brain MR images by deep learning. The pipeline of our method is shown in Fig. 1. In particular, we train the MS-Nets to reduce the anatomical complexity, and generate the trajectories for the fixed/moving images. The anatomical complexity is gradually reduced along each trajectory, while the images at the ends of the fixed/moving trajectories become *simple* and similar, implying that they are easy to be registered in the simplified morphological space. In this way, we can follow the fixed/moving (image) trajectories and decompose the original complex registration problem into several easy ones. Thus, our method consists of two steps: (1) generating the trajectories for the fixed and moving images by $n$-level morphological simplification networks (MS-Nets); (2) performing image registration with the guidance of the trajectories.

### 2.1. Morphological simplification network (MS-Net)

The key point of our method relies on the gradual reduction of the anatomical complexity in brain MR images. It is known that the cortex is highly folded in human brain. To acquire more accurate alignment of the anatomical structures, high-order features and sophisticated constraints derived from brain tissue segmentation are shown to be effective [9,35]. Recently, Zhang et al. proposed to use the smoothed cortical surface with reduced complexity to guide the registration of the 3D brain volumes [36]. However, it is non-trivial to get high-quality tissue segmentation especially when multi-center data is considered sometimes even expert editing of tissue segmentation is necessary. To this end, we propose to complete morphological simplification of brain MR images in the intensity space by deep learning, without need of any segmentation or surface reconstruction to the test image.

*Training data preparation.* We aim to train a morphological simplification network (MS-Net) to reduce the complexity in brain MR

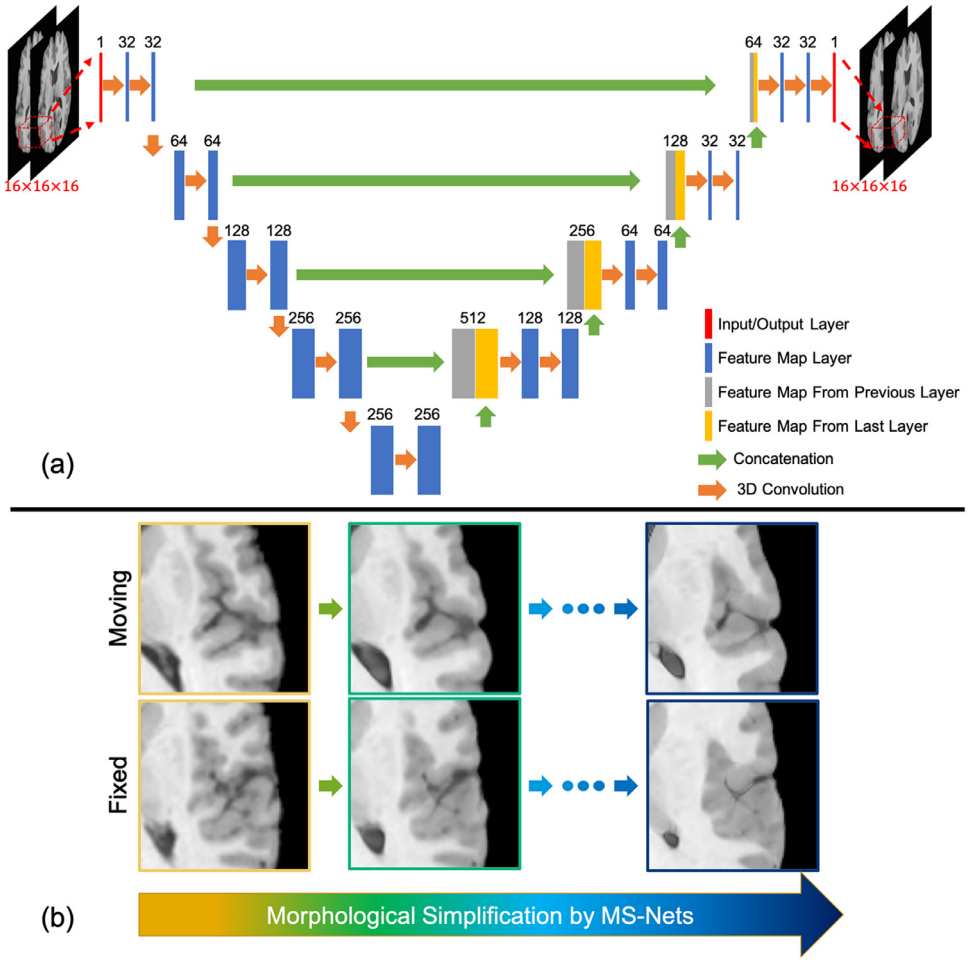


**Fig. 2.** (a) The architecture of MS-Net. The input and output of MS-Net are image patches by the red blocks. Each blue block indicates a multi-channel feature map, while the number of the channels is denoted with the block. The kernel size is $3 \times 3 \times 3$ for all layers. Grey and yellow blocks represent the feature maps copied for concatenation. (b) The example fixed/moving images and their simplified outputs by MS-Nets. From left to right, the complexity of the original fixed/moving image is reduced gradually. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

images. The training data consists of original MR images and the corresponding *simplified* images. We get the *simplified* MR images by: (1) reconstructing the inner/outer cortical surfaces from the segmented tissue map of the intensity image, including the topology correction [37]; (2) applying Laplacian smoothing on the reconstructed surfaces and converting the smoothed surfaces into tissue segmentation maps again by the strategy in [38], which ensures that the segmented tissues are not shrunk after surface smoothing; (3) registering the original and the smoothed tissue segmentation maps using diffeomorphic Demons [12] to get the deformation field; (4) warping the original MR intensity images by applying the aforementioned deformation fields to get the *simplified* intensity images. The above steps are repeated iteratively to generate MR images of fewer and fewer morphological details. The two corresponding images, before and after each smoothing operation, compose of a *complex-simple* pair for training the MS-Net. Specifically, $n$ *complex-simple* pairs $\{(I_s^{i-1}, I_s^i)|i = 1, \ldots, n\}$ can be prepared from subject $s$, where $I_s^0$ represents the original intensity MR image of the subject. Specifically, Laplacian smoothing is operating upon the cortical surface meshes. With $x_i$ representing the location of the $i$th vertex of the mesh of the inner/outer cortical surface, its new coordinate after smoothing is

$$x_p = x_p + \sum_{i \in N_p} w_i (x_i - x_p) \tag{1}$$

where $x_i$ ($i \in N_p$) is the neighbor of the vertex $x_p$, $w_i = 1/m$ indicates the weight, and $m$ is the size of the neighborhood $N_p$.

*Network configuration and training.* The architecture of MS-Net and the detailed configurations can be found in Fig. 2(a) and Table 3 in Appendix B. For easy illustration, we show the 3D layers inside the network by 2D boxes in the figure, while the number above each box indicates the channel number after convolution and concatenation. Each layer in the MS-Net is a 3D layer without pooling. The kernel size is $3 \times 3 \times 3$ and the stride is 1. Zero padding is adopted to keep the sizes of the feature maps and also make the output the same as the input through the MS-Net.

We train the MS-Net in a patch-by-patch way. Particularly, we sample 3D cubic patches sized $16 \times 16 \times 16$ from the training images. The sampling complies with the probability calculated at the center of each potential patch (denoted by $u$):

$$p(u) = \frac{|\nabla g_u^x| + |\nabla g_u^y| + |\nabla g_u^z|}{\|\nabla g\|} \tag{2}$$

where $\nabla g_u^x$, $\nabla g_u^y$, and $\nabla g_u^z$ are the gradients at $u$ in three directions, and $\|\nabla g\|$ is the gradient norm. In this way, the sampled patches cover the entire brain volumes and pay more attention to the regions of abundant appearance information [28]. To train the MS-Net, we usually extract 20,000 patch samples from each pair of the prepared *complex* and *simple* images. The network is trained

on an Nvidia Titan X GPU by Keras. The optimizer is Adam with 0.001 as the initial learning rate. For the loss function, we use the sum of the squared differences between the output patch and the ground-truth patch.

*Application of the MS-Net.* In the application stage, the trained network can be directly applied to generate the 3D *simple* output image from a *complex* input in the end-to-end way. Concerning the capacity of GPU, we implement to process every 16 axial slices ($256 \times 256 \times 16$), which are partially overlapped. Each $256 \times 256 \times 16$ patch is inputted into the MS-Net. Finally, the whole test image can then generate its *simple* version by averaging the results of all 31 test tasks, with two neighboring tasks sharing 8 overlapped slices. In this way, a typical test image sized $256 \times 256 \times 256$ is simplified by the MS-Net within ~ 3.3 s. Examples of the fixed/moving images and their simplified outputs are available in Fig. 2(b). Note that several MS-Nets are applied to derive the trajectories for the fixed/moving images in the figure. For $n$-level morphological simplification, we have $n$ MS-Nets, each one is input with its *complex* intensity image and output its *simple* version.

### 2.2. Trajectory and its guidance to registration

To generate the trajectory where a brain MR image is gradually simplified, we train a sequence of MS-Nets one by one. In particular, there are 7 levels of morphological simplification in our implementation ($n = 7$ as in Fig. 1), corresponding to 7 different MS-Nets. Each MS-Net is assigned to generate a corresponding intensity image with adequate smoothing scale, to ensure that the two consecutive images in the trajectory can be similar enough and then easily registered.

After trajectory construction, here we need to calculate the accurate deformation field between the fixed and the moving images. The trajectory guided registration is implemented by: (1) registering (using diffeomorphic Demons) the current moving image $I_m^0(\phi_0 \circ \ldots \circ \phi_k)$ with its corresponding *simplified* image $I_m^{k+1}$ $(k = 0, \ldots, n-1)$ or $I_f^{2n-k}$ $(k = n, \ldots, 2n)$ to obtain the deformation field $\phi_{k+1}$, where $\phi_o$ represents the identity transformation; (2) warping the moving image to obtain $I_m^0(\phi_0 \circ \ldots \circ \phi_k \circ \phi_{k+1})$. The above two steps are iteratively executed from $k = 0$ to $k = 2n$. The *simplified* fixed image $I_f^n$ and the moving image $I_m^0(\phi_0 \circ \ldots \circ \phi_n)$ at the ends of the trajectories are registered together to bridge the link between the fixed and moving trajectories. Specifically, $\phi_{n+1}$ is easier to be recovered than the direct deformation field between the original fixed and moving images. Subtle anatomical structural difference is encoded into $\{\phi_i | i = 1, \ldots, n, n+2, \ldots, 2n+1\}$, and thus only large-scale shape variation need to be registered at the ends of trajectories. Thus, the final deformation field that warps the original moving image to the fixed image is estimated by composing the intermediate deformation fields along the two trajectories as follows:

$$\phi = \phi_0 \circ \phi_1 \circ \ldots \circ \phi_{2n+1} \tag{3}$$

## 3. Experiments

We conduct comprehensive validation of the proposed method by using five public datasets, *i.e.*, NIREP NA0, LONI LPBA40, IBSR18, CUMC12 and MGH10. All datasets are widely adopted in the literature to evaluate the performance of brain MR image registration algorithms [20,22]. In pre-processing, skull-stripping has been applied to all images, which are later resampled to the same resolution ($1 \times 1 \times 1$ mm$^3$). All images are also processed through bias correction and linearly registered to the MNI152 space by FLIRT in FSL. Note that our pre-processing is consistent with the report in [20] for fair comparison.

Our method is mainly compared to Diffeomorphic Demons [12] and SyN [8], both of which are highly recommended in the large-scale validation in [20]. To quantitatively evaluate the registration performance, we adopt three metrics, *i.e.*, Dice similarity coefficient (DSC), target overlap (TO), and average symmetric surface distance (ASSD) of the annotated tissues or regions-of-interest (ROIs). These three metrics are widely used to quantify registration performance – a higher DSC/TO or lower ASSD usually indicates better registration quality. We note that DSC and ASSD are frequently adopted when evaluated upon large tissues (*i.e.*, GM/WM). For small ROIs, we adopt TO in order to keep consistent with [20].

### 3.1. NIREP Dataset

The NIREP dataset consists of 16 brain MR images, each of which contains GM/WM labeling and 32 small ROIs. In particular, we train all 7 MS-Nets with NIREP only. The trained networks are applied to other datasets in subsequent experiments. To this end, for the evaluation upon the NIREP dataset, we use the leave-two-out strategy. In every test case, there are 14 subjects for training and the rest 2 images act as the pair of the fixed/moving images. Thus, there are 240 test cases for the NIREP dataset in total, from which the evaluation metrics are computed.

The evaluation results are summarized in Table 1. The DSC scores for our method are 87.06 ± 0.91 (GM) and 89.60 ± 0.72 (WM), both of which are significantly higher than Demons (GM: 81.69 ± 1.02; WM: 84.20 ± 0.80) and SyN (GM: 81.59 ± 2.26; WM: 83.91 ± 2.10). The results of ASSD are similar with DSC, as our method yields significantly superior performances compared to Demons and SyN. Regarding the TO scores of 32 small ROIs, the overall average TO is 70.78 ± 5.00 for our method, compared to 67.39 ± 6.26 for Demons and 66.92 ± 6.96 for SyN. Paired $t$-tests indicate that, on 27/32 ROIs (except for Left Cingulate Gyrus, Right Cingulate Gyrus, Right Insula Gyrus, Left Temporal Pole and Right Parahippocampal Gyrus), our method performs better than the other two methods ($p < .01$). The visualization of the registration results in Fig. 3 shows that our proposed method achieves more accurate surface alignment especially in the regions indicated by red arrows.

The length of the trajectory, or the number of the MS-Nets used, is a critical parameter in our method. We particularly cascade 7 MS-Nets to simplify brain complexity gradually, which is also verified by the experiments on the NIREP dataset. That is, we randomly draw a pair of images from the NIREP dataset for 15 times. Given each drawn image pair, we use the other 14 images in the dataset to train the sequence of MS-Nets. Then, we evaluate the registration quality on the drawn image pair, while different numbers of MS-Nets are used and thus the length of the sequence/trajectory is altered. The results of the DSC scores of GM/WM over the 15 randomly selected pairs for testing are show in Fig. 4. Note that the DSC scores increase rapidly for a short trajectory (*i.e.*, $n \leq 3$) and become mostly stable after $n = 7$. To this end, we choose 7 as the optimal number of MS-Nets in all of our experiments.

### 3.2. Other datasets

With all MS-Nets trained with the NIREP dataset, we apply them directly to the other four datasets. The quantitative results are summarized in Table 2.

- The LPBA dataset consists of 40 brain MR images, each of which contains GM/WM labeling and 56 ROIs. In particular, we draw $40 \times 39$ pairs of the fixed and moving images from the dataset, which lead to 1560 testing cases in total. The 7 MS-Nets are directly transferred from

**Table 1**
Registration accuracy evaluated on the NIREP dataset.

| Evaluation on GM/WM Tissue Labels | | | |
|---|---|---|---|
| | Proposed | Demons | SyN |
| GM (DSC: %) | 87.06 ± 0.91 | 81.69 ± 1.02 | 81.59 ± 2.26 |
| WM (DSC: %) | 89.60 ± 0.72 | 84.20 ± 0.80 | 83.91 ± 2.10 |
| GM (ASSD: mm) | 0.30 ± 0.05 | 0.43 ± 0.06 | 0.40 ± 0.08 |
| WM (ASSD: mm) | 0.41 ± 0.09 | 0.53 ± 0.09 | 0.51 ± 0.12 |

| Evaluation on 32 small ROIs (TO: %) | | | |
|---|---|---|---|
| | Proposed | Demons | SyN |
| Overall | 70.78 ± 5.00 | 67.39 ± 6.26 | 66.92 ± 6.96 |

| | Left Hemisphere | | | Right Hemisphere | | |
|---|---|---|---|---|---|---|
| | Proposed | Demons | SyN | Proposed | Demons | SyN |
| Occipital Lobe | 71.81 ± 8.61 | 65.74 ± 8.51 | 66.87 ± 7.83 | 74.23 ± 6.08 | 66.73 ± 7.31 | 69.34 ± 5.88 |
| Cingulate Gyrus | 68.26 ± 9.01# | 68.94 ± 8.60 | 68.00 ± 8.28 | 69.36 ± 7.89# | 69.42 ± 8.00 | 68.67 ± 7.72 |
| Insula Gyrus | 76.74 ± 4.14 | 76.69 ± 4.64 | 75.85 ± 4.68 | 77.66 ± 3.62# | 78.68 ± 4.97 | 78.03 ± 3.39 |
| Temporal Pole | 75.63 ± 11.97# | 76.09 ± 9.90 | 74.88 ± 11.04 | 78.49 ± 8.04 | 77.70 ± 8.02 | 76.59 ± 8.24 |
| Superior Temporal Gyrus | 68.11 ± 7.73 | 66.00 ± 8.12 | 65.69 ± 7.63 | 67.85 ± 8.07 | 64.62 ± 9.03 | 64.50 ± 8.02 |
| Infero Temporal Region | 76.15 ± 5.07 | 72.09 ± 5.04 | 72.36 ± 5.23 | 77.19 ± 5.45 | 71.85 ± 6.46 | 72.65 ± 5.83 |
| Parahippocampal Gyrus | 73.22 ± 5.45 | 72.86 ± 5.74 | 72.01 ± 5.51 | 75.29 ± 5.72# | 75.29 ± 5.70 | 74.46 ± 5.34 |
| Frontal Pole | 74.15 ± 9.55 | 71.49 ± 10.19 | 72.43 ± 9.65 | 72.91 ± 10.56 | 69.84 ± 10.34 | 70.94 ± 9.85 |
| Superior Frontal Gyrus | 71.86 ± 7.18 | 68.20 ± 8.43 | 68.01 ± 7.27 | 72.24 ± 8.48 | 67.20 ± 9.67 | 67.51 ± 8.29 |
| Middle Frontal Gyrus | 70.84 ± 8.95 | 64.49 ± 9.10 | 65.67 ± 8.71 | 67.27 ± 8.64 | 60.17 ± 8.69 | 61.66 ± 8.09 |
| Inferior Gyrus | 66.10 ± 4.28 | 60.44 ± 14.09 | 61.04 ± 14.00 | 67.41 ± 9.98 | 60.03 ± 11.49 | 62.26 ± 9.77 |
| Orbital Frontal Gyrus | 75.97 ± 6.34 | 74.67 ± 6.84 | 73.70 ± 6.29 | 75.16 ± 5.88 | 73.58 ± 6.17 | 72.66 ± 5.57 |
| Precentral Gyrus | 66.31 ± 8.13 | 61.22 ± 8.48 | 62.46 ± 7.28 | 64.11 ± 7.42 | 58.89 ± 8.44 | 59.99 ± 7.33 |
| Superior Parietal Lobule | 67.33 ± 9.30 | 59.07 ± 10.9 | 61.23 ± 9.27 | 66.28 ± 7.85 | 57.94 ± 7.87 | 60.81 ± 7.25 |
| Inferior Parietal Lobule | 69.30 ± 8.16 | 61.73 ± 8.57 | 64.28 ± 7.78 | 68.71 ± 8.95 | 60.18 ± 9.42 | 63.30 ± 8.82 |
| Postcentral Gyrus | 60.70 ± 12.24 | 56.37 ± 12.46 | 55.94 ± 11.79 | 58.17 ± 9.71 | 52.94 ± 9.28 | 52.42 ± 9.34 |

# indicates the ROIs (*i.e.*, Left Cingulate Gyrus, Right Cingulate Gyrus, Right Insula Gyrus, Left Temporal Pole and Right Paraphippocampal Gyrus) where our method does not outperform both Demons and SyN significantly at the same time (*i.e.*, $p > .01$ in paired $t$-tests against either Demons or SyN).

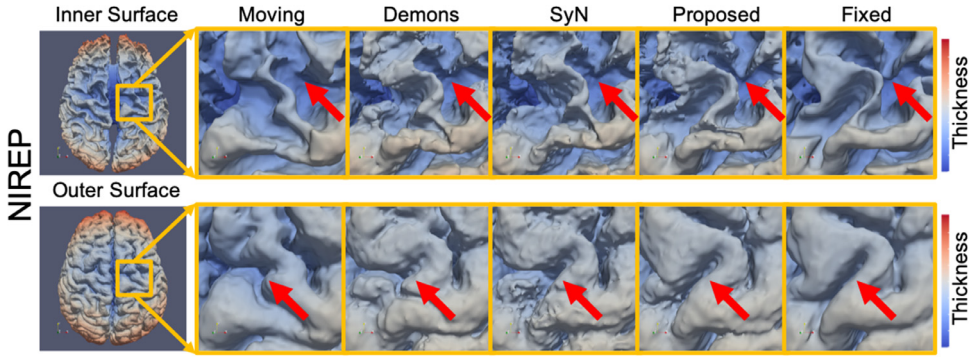


**Fig. 3.** Visualization of the registration results of the NIREP dataset by Demons, SyN and our proposed method. The images are shown in reconstructed inner (top row) and outer (bottom row) cortical surfaces, and colored in accordance to cortical thickness. Our method shows more accurate surface alignment especially in the regions highlighted by red arrows. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the NIREP dataset. According to the results in Table 2, the DSC scores for our method are 82.00 ± 1.50 (GM) and 87.86 ± 0.61 (WM), both of which are significantly higher than Demons (GM: 76.53 ± 1.88; WM: 82.55 ± 0.76) and SyN (GM: 77.74 ± 1.90; WM: 84.23 ± 0.91). The results of ASSD are similar, as our method performs significantly better than Demons and SyN. The average TO scores of 56 small ROIs, however, is not improved (our method: 69.48 ± 5.84; Demons: 71.12 ± 5.28; SyN: 72.26 ± 5.23). A detailed discussion of TO scores will be provided in the next.

- CUMC consists of 12 brain MR images with 128 ROIs. For evaluation, we conduct 12 × 11 pairs of registration tasks. While the MS-Nets are directly transferred from the NIREP dataset, the DSC scores for our method are 78.62 ± 1.59 (GM) and 85.98 ± 0.81 (WM), both of which are significantly higher than Demons (GM: 73.03 ± 1.58; WM: 80.45 ± 0.80) and SyN (GM: 75.15 ± 1.63; WM: 82.42 ± 1.11). The results of ASSD are also similar. Regarding the TO scores of the small ROIs, the average TO of our method is mostly comparable (our method: 52.08 ± 14.58; Demons: 51.59 ± 15.06; SyN: 52.17 ± 15.11).
- IBSR consists of 18 brain MR images and 84 ROIs. For the 17 × 18 pairs of registration tasks, the DSC scores for our method are 84.59 ± 3.52 (GM) and 81.14 ± 3.88 (WM), both of which are significantly higher than Demons (GM: 83.26 ± 2.14; WM: 78.76 ± 2.61) and SyN (GM:

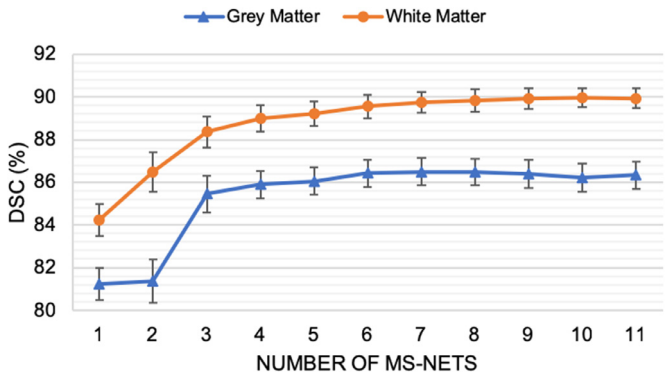


**Fig. 4.** The mean DSC of GM/WM based on 15 image pairs drawn from the NIREP dataset, with respect to the gradually increasing number of MS-Nets used to derive the fixed/moving sequences.

84.41 ± 2.40; WM: 80.28 ± 2.91). The results of ASSD are similar, as our method yields significantly better performance in GM and comparable performance in WM, compared to Demons and SyN. Regarding the TO scores, the average TO of all 84 ROIs is not improved (our method: 49.68 ± 14.58; Demons: 51.59 ± 10.51; SyN: 53.96 ± 10.76).

- MGH consists of 10 brain images and 74 ROIs. There are thus 10 × 9 pairs of registration tasks to evaluate. The DSC scores for our method are 81.12 ± 2.12 (GM) and 86.15 ± 2.25 (WM), both of which are significantly higher than Demons (GM: 78.26 ± 1.30; WM: 81.43 ± 1.29) and SyN (GM: 80.35 ± 1.39; WM: 83.77 ± 1.04). The results of ASSD are similar, as our method yields significantly better performance in GM and comparable performance in WM, compared to Demons and SyN. Regarding the TO scores of 74 ROIs, the average TO of 130 ROIs is not improved (our method: 53.00 ± 12.05; Demons: 56.64 ± 12.45; SyN: 57.04 ± 13.30).

In general, the above quantitative evaluation supports the conclusion that our method can improve the registration accuracy of brain MR images. The disagreement largely comes from the metric of TO scores, while a detailed analysis will be provided later. Moreover, we show the visualization of the typical registration results in Fig. 5. All images in the figure are rendered through their inner and outer cortical surfaces, which are reconstructed from the labeling of GM/WM. We observe that, after being warped through the deformation fields of different methods, the moving image becomes similar with the fixed image. Particularly, the proposed method results in the most accurate alignment of the cortical surfaces, especially in the regions that are highlighted by the red arrows. To this end, we argue that the proposed method can achieve superior registration performance when it is generalized to diverse datasets.

## 4. Discussion

In this paper, we have proposed a novel deep learning based method to guide deformable registration of brain MR images. The MS-Nets simplify morphological complexity of the fixed and the moving images, such that these two images become similar with each other and easy to be registered eventually. Our experiments show superior alignment performance, especially near the cortical surface, attained by our method, compared with the state-of-the-art methods. Moreover, our method has demonstrated its promising generalization capability. While the MS-Nets are trained with a certain dataset (*i.e.*, NIREP), the registration quality on other four datasets are mostly satisfactory.

Note that our method is highly different from deep learning based registration methods, such as VoxelMorph [39]. First, MS-Net is a fully convolution network without pooling, which aims to infer an image with simplified morphology. The output of MS-Net then acts as the target, toward which the input moving image is registered and deformed. Therefore, although the output image of MS-Net reduces appearance complexity with respect to the input image, the actual registration above and the deformed image preserves all information that is inherited from the input image prior to simplification or deformation since the anatomical details are encoded into the estimated deformation field. Second, MS-Net is not designed to estimate the deformation field directly, which is commonly produced through a black-box of deep learning in several recent works. Admittedly, in the future work, we intend to integrate all MS-Nets into a unified network, such that the simplified intermediate images and the deformation pathway between the input images can be generated simultaneously. In this way, we will have a more efficient implementation, while the deformation pathway is clearly tractable.

Although our method has shown its superior performance by visual inspection and by DSC/ASSD scores, one may note that the TO scores of small ROIs of our method are often short of Demons and SyN. A possible reason is because of the inconsistent quality in labeling ROIs, especially in reference to GM/WM boundaries. For example, given an LPBA subject, one may refine the ROIs by inter-

**Table 2**
Registration accuracy evaluated on the four datasets of LPBA, CUMC, IBSR, and MGH.

| Dataset | Method | DSC (%) | | ASSD (mm) | | TO (%) |
|---|---|---|---|---|---|---|
| | | GM | WM | GM | WM | ROIs |
| LPBA | Demons | 76.53 ± 1.88 | 82.55 ± 0.76 | 0.42 ± 0.04 | 0.50 ± 0.09 | 71.12 ± 5.28 (68.93 ± 6.10†) |
| | SyN | 77.74 ± 1.90 | 84.23 ± 0.91 | 0.38 ± 0.04 | 0.45 ± 0.09 | 72.26 ± 5.23 (71.46 ± 5.75†) |
| | Proposed | 82.00 ± 1.50 | 87.86 ± 0.61 | 0.33 ± 0.04 | 0.41 ± 0.09 | 69.48 ± 5.84 |
| CUMC | Demons | 73.03 ± 1.58 | 80.45 ± 0.80 | 0.41 ± 0.05 | 0.50 ± 0.09 | 51.59 ± 15.06 (46.46 ± 15.56†) |
| | SyN | 75.15 ± 1.63 | 82.42 ± 1.11 | 0.37 ± 0.06 | 0.47 ± 0.10 | 52.17 ± 15.11 (51.63 ± 14.88†) |
| | Proposed | 78.62 ± 1.59 | 85.98 ± 0.81 | 0.31 ± 0.04 | 0.42 ± 0.10 | 52.08 ± 14.58 |
| IBSR | Demons | 83.26 ± 2.14 | 78.76 ± 2.61 | 0.47 ± 0.15 | 0.59 ± 0.20 | 51.59 ± 10.51 (46.82 ± 9.89†) |
| | SyN | 84.41 ± 2.40 | 80.28 ± 2.91 | 0.44 ± 0.15 | 0.55 ± 0.19 | 53.96 ± 10.76 (52.81 ± 10.45†) |
| | Proposed | 84.59 ± 3.52 | 81.14 ± 3.88 | 0.35 ± 0.13 | 0.56 ± 0.29 | 49.68 ± 14.58 |
| MGH | Demons | 78.26 ± 1.30 | 81.43 ± 1.29 | 0.40 ± 0.07 | 0.50 ± 0.11 | 56.64 ± 12.45 (52.28 ± 12.97†) |
| | SyN | 80.35 ± 1.39 | 83.77 ± 1.04 | 0.37 ± 0.08 | 0.44 ± 0.08 | 57.04 ± 13.30 (56.83 ± 13.02†) |
| | Proposed | 81.12 ± 2.12 | 86.15 ± 2.25 | 0.29 ± 0.07 | 0.45 ± 0.19 | 53.00 ± 12.05 |

† indicates the results copied from Klein et al. [20]. Note that our reproduced results are clearly better, partially due to the continuous improvement of Demons and SyN in past years.

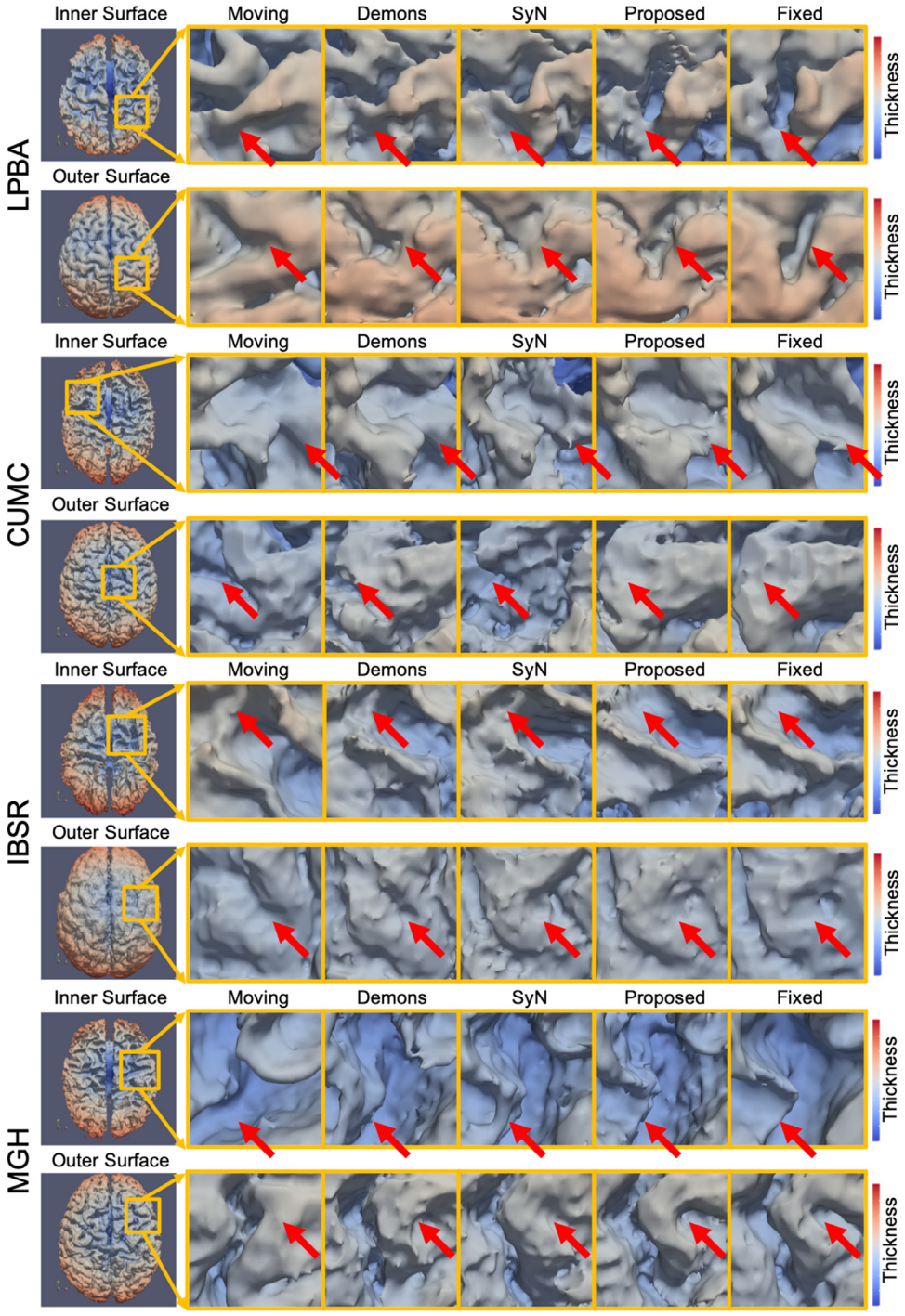


**Fig. 5.** Visualization of typical registration results by Demons, SyN and our proposed method on the four datasets of LPBA, CUMC, IBSR, and MGH. Our method shows more accurate surface alignment in the regions indicated by the red arrows. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

secting with GM/WM tissue labels [40]. In this way, each ROI can be split into ROIGM and ROIWM. We further investigate the registration quality by computing TO scores in ROIGM and ROIWM, respectively. The average scores for the four datasets (LBPA, CUMC, IBSR, and MGH) are compared in Fig. 6.

While the TO scores on the entire ROIs of our method may be slightly lower than the other methods under comparison, it is clear that our method performs better (or in a comparable way) by referring to the splitted ROIs. For example, regarding ROIsGM, the average TO score for our method on LPBA is 57.91 $\pm$ 13.41, compared to 55.83 $\pm$ 13.25 (Demons) and 41.15 $\pm$ 12.32 (SyN). For CUMC, the scores are 51.44 $\pm$ 13.85 (our method), 46.96 $\pm$ 15.12 (Demons), and 47.88 $\pm$ 14.94 (SyN), respectively. Meanwhile, although our methods are often better in ROIsWM, we argue that the results of the four datasets might be strongly biased if the ROI labeling within WM is counted in. In particular, the TO scores on ROIsWM produced by Demons and SyN are very low (usually about 10–30), such that one may challenge whether the boundary of the

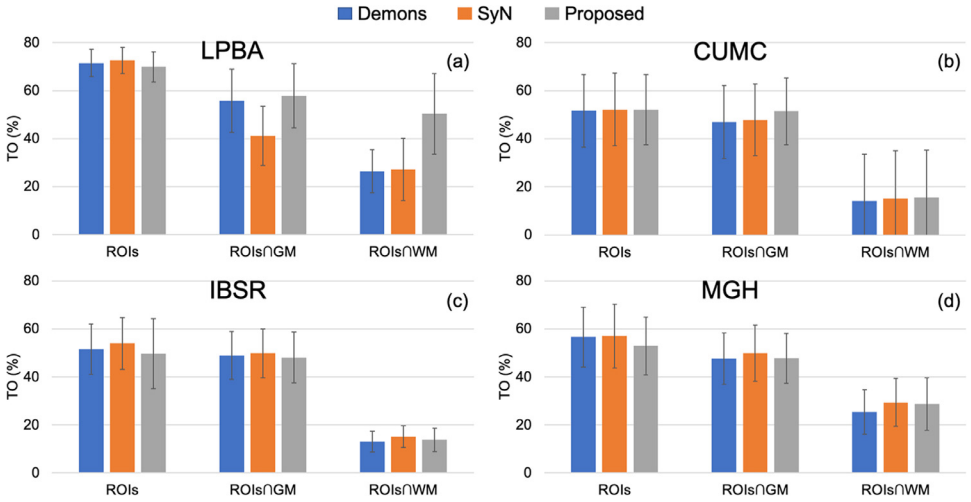


**Fig. 6.** The TO scores measured from the labeled ROIs and the ROIs intersected by GM/WM. From (a)–(d), the results on four datasets (LPBA, CUMC, IBSR, and MGH) are reported.

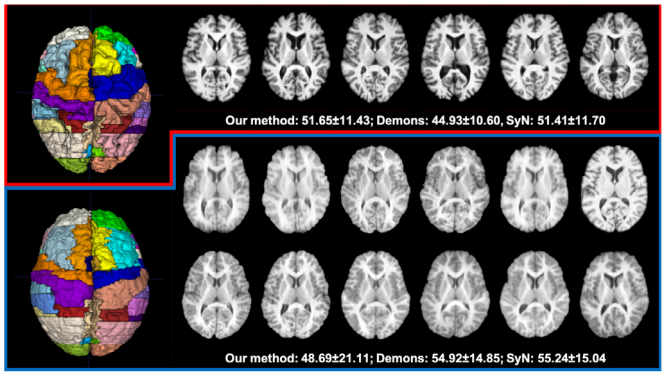


**Fig. 7.** Visualization of all 18 subjects in the IBSR dataset, as well as the outer cortical surface rendering of the ROIs of two exemplar subjects. The dataset is clearly divided into two groups by appearance. The red group (top) includes 6 subjects with high contrast, while the quality of the 12 subjects in the blue group (bottom) is relatively low. The ROI labeling of the blue group often fails to reveal the subtle gyral and sulcal structures, resulting in an unexpectedly smooth cortical surface labeled by the ROIs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ROI is determined properly in WM or near the GM/WM interface. In this case, the scores on the small ROIs may not be the proper indicators of registration quality. In addition to the inner cortical surface between GM and WM, the ambiguity near the outer cortical surface that is partially due to low imaging quality also challenges the reliability of the TO scores. By referring to the visualization of all images in the IBSR dataset in Fig. 7, it is clear to observe that the two groups of images are significantly different in their appearance. The quality of the red group (in the top of the figure) appears better, *e.g.*, with fewer artefacts and clearer details. For image pairs in the red group, the registration performance of our method is much better than other two methods. However, our method fails to compete with other two methods in the blue group by TO scores. We argue that the labeling quality of the ROIs is directly related with the TO scores in this case. Particularly, an example of the blue group shows barely details of the labeled ROIs at the outer cortical surface, which is clearly caused by the low quality of the image itself. A similar observation can also be acquired from the MGH dataset, where the quality of the ROIs labeling might be questionable.

## 5. Conclusion

In conclusion, this study establishes MS-Net as a powerful and flexible tool to simplify the MR images for deformable registration. The MS-Net provides morphologically simplified images as intermediate guidance, which is also robust to transfer to a new dataset. Our proposed unique method divides the highly complex inter-subject registration task into several easy tasks. Experimental results show superior alignment performance especially near cortical surface compared with state-of-the-art methods on multiple datasets.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

This work was partially supported by the National Key Research and Development Program of China (2018YFC0116400), STCSM (19QC1400600, 17411953300, 19PJ1406800), and Medical-Engineering Cross Research Foundation of Shanghai Jiao Tong University.

## Appendix A. Details of parameter settings

Here we provide the details of the parameters [8,12] involved in each publicly available registration method used.

1. Demons:
   DemonsRegistration -f < fixed_image.nii > -m < moving_image.nii > -O < output.mha > -e -s 2 -i 30×20×10
   DemonsWarp -m < labeled_moving.nii > -b < output.mha > -o < output_labels.nii > -I
2. SyN:
   ANTS 3 -m PR[ < fixed_image.nii > , < moving_image.nii > , 1, 2] -O < output > -i 30 × 99 × 11 -t SyN[0.5] -r Gauss[2,0] -use-Histogram-Matching
   WarpImageMultiTransform 3 < labeled_moving.nii > < output_label.nii > -R < fixed_image.nii > < outputWarp.nii.gz > < outputAffine.txt > -use-NN

## Appendix B. Details of MS-Net

**Table 3**
The detailed layer configurations of MS-Net. Conv3D represents the 3D convolution layer with the specified filter size, stride and number of filters (Num. F.). Zero padding is adopted to keep the sizes of the feature maps and also make the output the same with the input through MS-Net. The BatchNorm indicates whether the Conv3D layer is followed by a Batch Normalization layer. Skip connection is marked by the prefix Concat, which concatenates the previous feature maps with the current Conv3D feature maps.

| Layer Name | Filter Size | Num. F. | Stride | BatchNorm | Nonlinearity |
|---|---|---|---|---|---|
| Input layer | | | | N | |
| Conv3D_1 | $3 \times 3 \times 3$ | 32 | 1 | N | ReLU |
| Conv3D_2 | $3 \times 3 \times 3$ | 32 | 1 | Y | ReLU |
| Conv3D_3 | $3 \times 3 \times 3$ | 64 | 1 | Y | ReLU |
| Conv3D_4 | $3 \times 3 \times 3$ | 64 | 1 | Y | ReLU |
| Conv3D_5 | $3 \times 3 \times 3$ | 128 | 1 | Y | ReLU |
| Conv3D_6 | $3 \times 3 \times 3$ | 128 | 1 | Y | ReLU |
| Conv3D_7 | $3 \times 3 \times 3$ | 256 | 1 | Y | ReLU |
| Conv3D_8 | $3 \times 3 \times 3$ | 256 | 1 | Y | ReLU |
| Conv3D_9 | $3 \times 3 \times 3$ | 256 | 1 | Y | ReLU |
| Conv3D_10 | $3 \times 3 \times 3$ | 256 | 1 | Y | ReLU |
| Concat_1(Conv3D_8) | | 256+256 | | N | |
| Conv3D_11 | $3 \times 3 \times 3$ | 128 | 1 | Y | ReLU |
| Conv3D_12 | $3 \times 3 \times 3$ | 128 | 1 | Y | ReLU |
| Concat_2(Conv3D_6) | | 128+128 | | N | |
| Conv3D_13 | $3 \times 3 \times 3$ | 64 | 1 | Y | ReLU |
| Conv3D_14 | $3 \times 3 \times 3$ | 64 | 1 | Y | ReLU |
| Concat_3(Conv3D_4) | | 64+64 | | N | |
| Conv3D_15 | $3 \times 3 \times 3$ | 32 | 1 | Y | ReLU |
| Conv3D_16 | $3 \times 3 \times 3$ | 32 | 1 | Y | ReLU |
| Concat_4(Conv3D_2) | | 32+32 | | N | |
| Conv3D_17 | $3 \times 3 \times 3$ | 32 | 1 | Y | ReLU |
| Conv3D_18 | $3 \times 3 \times 3$ | 32 | 1 | Y | ReLU |
| Conv3D_19 | $3 \times 3 \times 3$ | 1 | 1 | N | ReLU |

## References


[1] A. Sotiras, C. Davatzikos, N. Paragios, Deformable medical image registration: a survey, IEEE Trans. Med. Imaging 32 (7) (2013) 1153–1190.
[2] H. Lester, S.R. Arridge, A survey of hierarchical non-linear medical image registration, Pattern Recognit. 32 (1) (1999) 129–149.
[3] R.W. So, T.W. Tang, A.C. Chung, Non-rigid image registration of brain magnetic resonance images using graph-cuts, Pattern Recognit. 44 (10–11) (2011) 2450–2467.
[4] D. Shen, Image registration by local histogram matching, Pattern Recognit. 40 (4) (2007) 1161–1172.
[5] R.W. So, A.C. Chung, A novel learning-based dissimilarity metric for rigid and non-rigid medical image registration by using Bhattacharyya Distances, Pattern Recognit. 62 (2017) 161–174.
[6] R.P. Woods, S.T. Grafton, C.J. Holmes, S.R. Cherry, J.C. Mazziotta, Automated image registration: I. General methods and intrasubject, intramodality validation, J. Comput. Assist. Tomogr. 22 (1) (1998) 139–152.
[7] B.A. Ardekani, S. Guckemus, A. Bachman, M.J. Hoptman, M. Wojtaszek, J. Nierenberg, Quantitative comparison of algorithms for inter-subject registration of 3d volumetric brain mri scans, J. Neurosci. Methods 142 (1) (2005) 67–76.
[8] B.B. Avants, C.L. Epstein, M. Grossman, J.C. Gee, Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain, Med. Image Anal. 12 (1) (2008) 26–41.
[9] D. Shen, C. Davatzikos, Hammer: hierarchical attribute matching mechanism for elastic registration, IEEE Trans. Med. Imaging 21 (11) (2002) 1421–1439.
[10] D. Shen, Fast image registration by hierarchical soft correspondence detection, Pattern Recognit. 42 (5) (2009) 954–961.
[11] M. Lorenzi, N. Ayache, G.B. Frisoni, X. Pennec, A.D.N.I. (ADNI), et al., LCC-Demons: a robust and accurate symmetric diffeomorphic registration algorithm, Neuroimage 81 (2013) 470–483.
[12] T. Vercauteren, X. Pennec, A. Perchant, N. Ayache, Diffeomorphic demons: efficient non-parametric image registration, Neuroimage 45 (1) (2009) S61–S72.
[13] J.-P. Thirion, Image matching as a diffusion process: an analogy with Maxwell's demons, Med. Image Anal. 2 (3) (1998) 243–260.
[14] P. Hellier, J. Ashburner, I. Corouge, C. Barillot, K.J. Friston, Inter-subject registration of functional and anatomical data using SPM, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2002, pp. 590–597.
[15] Y. Ou, A. Sotiras, N. Paragios, C. Davatzikos, DRAMMS: deformable registration via attribute matching and mutual-saliency weighting, Med. Image Anal. 15 (4) (2011) 622–639.
[16] B. Glocker, N. Komodakis, G. Tziritas, N. Navab, N. Paragios, Dense image registration through MRFs and efficient linear programming, Med. Image Anal. 12 (6) (2008) 731–741.
[17] D. Rueckert, L.I. Sonoda, C. Hayes, D.L. Hill, M.O. Leach, D.J. Hawkes, Nonrigid registration using free-form deformations: application to breast MR images, IEEE Trans. Med. Imaging 18 (8) (1999) 712–721.
[18] J. Andersson, S. Smith, M. Jenkinson, FNIRT–FMRIB's non-linear image registration tool, Hum. Brain Mapp. 2008 (2008).
[19] M.F. Beg, M.I. Miller, A. Trouvé, L. Younes, Computing large deformation metric mappings via geodesic flows of diffeomorphisms, Int. J. Comput. Vis. 61 (2) (2005) 139–157.
[20] A. Klein, J. Andersson, B.A. Ardekani, J. Ashburner, B. Avants, M.-C. Chiang, G.E. Christensen, D.L. Collins, J. Gee, P. Hellier, et al., Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration, Neuroimage 46 (3) (2009) 786–802.
[21] Y. Ou, H. Akbari, M. Bilello, X. Da, C. Davatzikos, Comparative evaluation of registration algorithms in different brain databases with varying difficulty: results and insights, IEEE Trans. Med. Imaging 33 (10) (2014) 2039–2065.
[22] T. Rohlfing, Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unreliable, IEEE Trans. Med. Imaging 31 (2) (2012) 153–163.
[23] J. Hamm, D.H. Ye, R. Verma, C. Davatzikos, GRAM: a framework for geodesic registration on anatomical manifolds, Med. Image Anal. 14 (5) (2010) 633–642.
[24] D.H. Ye, J. Hamm, D. Kwon, C. Davatzikos, K.M. Pohl, Regional manifold learning for deformable registration of brain MR images, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2012, pp. 131–138.
[25] D. Wei, S. Ahmad, J. Huo, W. Peng, Y. Ge, Z. Xue, P.-T. Yap, W. Li, D. Shen, Q. Wang, Synthesis and inpainting-based mr-ct registration for image-guided thermal ablation of liver tumors, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2019, pp. 512–520.
[26] N.P. Castellanos, P. Angel, V. Medina, Nonrigid medical image registration technique as a composition of local warpings, Pattern Recognit. 37 (11) (2004) 2141–2154.

[27] Q. Wang, M. Kim, Y. Shi, G. Wu, D. Shen, A.D.N. Initiative, et al., Predict brain MR image registration via sparse learning of appearance and transformation, Med. Image Anal. 20 (1) (2015) 61–75.
[28] X. Cao, J. Yang, J. Zhang, Q. Wang, P.-T. Yap, D. Shen, Deformable image registration using a cue-aware deep regression network, IEEE Trans. Biomed. Eng. 65 (9) (2018) 1900–1911.
[29] X. Yang, R. Kwitt, M. Styner, M. Niethammer, Quicksilver: fast predictive image registration–a deep learning approach, Neuroimage 158 (2017) 378–396.
[30] D. Mahapatra, Z. Ge, Training data independent image registration using generative adversarial networks and domain adaptation, Pattern Recognit. (2019) 107109.
[31] H. Li, Y. Fan, Non-rigid image registration using self-supervised fully convolutional networks without training data, in: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), IEEE, 2018, pp. 1075–1078.
[32] R. Liao, S. Miao, P. de Tournemire, S. Grbic, A. Kamen, T. Mansi, D. Comaniciu, An artificial agent for robust image registration, in: Thirty-First AAAI Conference on Artificial Intelligence, 2017.
[33] A.V. Dalca, G. Balakrishnan, J. Guttag, M.R. Sabuncu, Unsupervised learning for fast probabilistic diffeomorphic registration, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2018, pp. 729–738.
[34] Y. Hu, E. Gibson, N. Ghavami, E. Bonmati, C.M. Moore, M. Emberton, T. Vercauteren, J.A. Noble, D.C. Barratt, Adversarial deformation regularization for training image registration neural networks, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2018, pp. 774–782.
[35] A.A. Joshi, D.W. Shattuck, P.M. Thompson, R.M. Leahy, Surface-constrained volumetric brain registration using harmonic mappings, IEEE Trans. Med. Imaging 26 (12) (2007) 1657–1669.
[36] J. Zhang, Q. Wang, G. Wu, D. Shen, Cross-manifold guidance in deformable registration of brain MR images, in: International Conference on Medical Imaging and Augmented Reality, Springer, 2016, pp. 415–424.
[37] G. Li, J. Nie, G. Wu, Y. Wang, D. Shen, A.D.N. Initiative, et al., Consistent reconstruction of cortical surfaces from longitudinal brain MR images, Neuroimage 59 (4) (2012) 3805–3820.
[38] G. Taubin, Curve and surface smoothing without shrinkage, in: Proceedings of IEEE International Conference on Computer Vision, IEEE, 1995, pp. 852–857.
[39] G. Balakrishnan, A. Zhao, M.R. Sabuncu, J. Guttag, A.V. Dalca, An unsupervised learning model for deformable medical image registration, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 9252–9260.
[40] G.E. Christensen, X. Geng, J.G. Kuhl, J. Bruss, T.J. Grabowski, I.A. Pirwani, M.W. Vannier, J.S. Allen, H. Damasio, Introduction to the non-rigid image registration evaluation project (NIREP), in: International Workshop on Biomedical Image Registration, Springer, 2006, pp. 128–135.

**Dongming Wei** received the BE degree from the School of Instrument and Optoelectronic Engineering, Beihang University, China in 2016. He is now pursuing the Ph.D. degree from the school of biomedical engineering, Shanghai Jiao Tong Univeristy, China. He is also a visiting student in The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. His research interests include pattern recognition and medical image analysis.

**Lichi Zhang** received the BE degree from the School of Computer Science, Beijing University of Posts and Telecommunications, China in 2008, and the Ph.D. degree from Department of Computer Science, University of York, UK, in 2014. He was working as a postdoc researcher in Med-X Research Institute, Shanghai Jiao Tong University, China and University of North Carolina at Chapel Hill, US from 2014 to 2017. He is now an assistant professor in Shanghai Jiao Tong University. His research interests include 3D shape reconstruction, machine learning and medical image analysis.

**Zhengwang Wu** is a postdoc researcher at the University of North Carolina at Chapel Hill. He got his Ph.D. degree on pattern recognition and intelligent system from Xian Jiaotong University, China. His research focuses on computer vision, machine learning, pattern recognition methods and their applications on medical image segmentation, classification and visualization. He is currently working on infant cortical surface atlas construction and computational pipeline development, as well as using state-of-the-art machine learning methods for infant MRI data analysis.

**Xiaohuan Cao** received the bachelors, masters and Ph.D. degrees from the School of Automation, Northwestern Polytechnical University. She is researcher in Shanghai United Imaging Intelligence Co., Ltd., Shanghai, China. Her research interests include machine learning and medical image analysis.

**Gang Li** is an Assistant Professor in the University of North Carolina at Chapel Hill. He completed his Ph.D. in 2010 from Northwestern Polytechnical University. He was a Research Fellow at The Methodist Hospital Research Institute, Weill Medical College of Cornell University (2007–2008) and a Research Assistant at Harvard Medical School (2005–2007).

**Dinggang Shen** is Jeffrey Houpt Distinguished Investigator, and a Professor of Radiology, Biomedical Research Imaging Center (BRIC), Computer Science, and Biomedical Engineering in the University of North Carolina at Chapel Hill (UNC-CH). He is currently directing the Center for Image Analysis and Informatics, the Image Display, Enhancement, and Analysis (IDEA) Lab in the Department of Radiology, and also the medical image analysis core in the BRIC. He was a tenure-track assistant professor in the University of Pennsylvanian (UPenn), and a faculty member in the Johns Hopkins University. Dr. Shen's research interests include medical image analysis, computer vision, and pattern recognition. He has published more than 1000 papers in the international journals and conference proceedings, with H-index 95. He serves as an editorial board member for eight international journals. He has also served in the Board of Directors, The Medical Image Computing and Computer Assisted Intervention (MICCAI) Society, in 2012-2015, and is General Chair for MICCAI 2019. He is Fellow of IEEE, Fellow of The American Institute for Medical and Biological Engineering (AIMBE), and also Fellow of The International Association for Pattern Recognition (IAPR).

**Qian Wang** received the B.S. and M.S. degrees in electronic engineering from Shanghai Jiao Tong University, in 2006 and 2009, respectively, and the Ph.D. degree in computer science from The University of North Carolina at Chapel Hill, in 2013. He is currently the Director of the Medical Image Computing Laboratory, Institute of Medical Imaging Technology, School of Biomedical Engineering, Shanghai Jiao Tong University. His re- search interests include medical image analysis, computer vision, machine learning, artificial intelligence, and translational medical studies.