

RESEARCH

Open Access



# Deformable registration of magnetic resonance images using unsupervised deep learning in neuro-/radiation oncology

Alexander F. I. Osman<sup>1\*</sup>, Kholoud S. Al-Mugren<sup>2</sup>, Nissren M. Tamam<sup>2</sup> and Bilal Shahine<sup>3</sup>

## Abstract

**Purpose** Accurate deformable registration of magnetic resonance imaging (MRI) scans containing pathologies is challenging due to changes in tissue appearance. In this paper, we developed a novel automated three-dimensional (3D) convolutional U-Net based deformable image registration (ConvUNet-DIR) method using unsupervised learning to establish correspondence between baseline pre-operative and follow-up MRI scans of patients with brain glioma.

**Methods** This study involved multi-parametric brain MRI scans (T1, T1-contrast enhanced, T2, FLAIR) acquired at pre-operative and follow-up time for 160 patients diagnosed with glioma, representing the BraTS-Reg 2022 challenge dataset. ConvUNet-DIR, a deep learning-based deformable registration workflow using 3D U-Net style architecture as a core, was developed to establish correspondence between the MRI scans. The workflow consists of three components: (1) the U-Net learns features from pairs of MRI scans and estimates a mapping between them, (2) the grid generator computes the sampling grid based on the derived transformation parameters, and (3) the spatial transformation layer generates a warped image by applying the sampling operation using interpolation. A similarity measure was used as a loss function for the network with a regularization parameter limiting the deformation. The model was trained via unsupervised learning using pairs of MRI scans on a training data set ( $n = 102$ ) and validated on a validation data set ( $n = 26$ ) to assess its generalizability. Its performance was evaluated on a test set ( $n = 32$ ) by computing the Dice score and structural similarity index (SSIM) quantitative metrics. The model's performance also was compared with the baseline state-of-the-art VoxelMorph (VM1 and VM2) learning-based algorithms.

**Results** The ConvUNet-DIR model showed promising competency in performing accurate 3D deformable registration. It achieved a mean Dice score of  $0.975 \pm 0.003$  and SSIM of  $0.908 \pm 0.011$  on the test set ( $n = 32$ ). Experimental results also demonstrated that ConvUNet-DIR outperformed the VoxelMorph algorithms concerning Dice (VM1:  $0.969 \pm 0.006$  and VM2:  $0.957 \pm 0.008$ ) and SSIM (VM1:  $0.893 \pm 0.012$  and VM2:  $0.857 \pm 0.017$ ) metrics. The time required to perform a registration for a pair of MRI scans is about 1 s on the CPU.

**Conclusions** The developed deep learning-based model can perform an end-to-end deformable registration of a pair of 3D MRI scans for glioma patients without human intervention. The model could provide accurate, efficient, and robust deformable registration without needing pre-alignment and labeling. It outperformed the state-of-the-art

\*Correspondence:  
Alexander F. I. Osman  
alexanderfadul@yahoo.com

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

VoxelMorph learning-based deformable registration algorithms and other supervised/unsupervised deep learning-based methods reported in the literature.

**Keywords** Deformable image registration, MRI, Convolutional neural networks, Unsupervised learning, Glioma, Neuro-/radiation oncology

## Introduction

The current standard conventional deformable registration algorithms and toolboxes, including Syn [1], Elastix [2], advanced normalization tools (ANTs) [3], and demons [4] involve solving a numerical optimization problem independently for each volumetric pair of images by applying geometric methods. This process is usually computationally expensive and time-consuming due to its iterative nature. In addition, these algorithms are non-learning from previous registration and frequently reiterate and optimize. Although the numerical optimization-based image registration methods perform reasonably well, they are restricted by their slow registration speeds.

Deep learning-based techniques have recently received significant attention in medical imaging and cancer treatment e.g. neuro-/radiation oncology [5]. Their capability for the deformable registration task has been increasingly investigated in different medical image modalities [6]. In these methods, in contrast to classical ones, deformable registration is defined as a parametric function. The optimization process is carried out by tuning the learning parameters given a set of fixed and moving images. When the deep learning-based network is trained, it can register a pair of three-dimensional (3D) medical images significantly faster than current standard algorithms [6]. In addition to a significant reduction in the processing time, the recently published deep learning registration methods such as VoxelMorph [7], DLIR [8], and FAIM [9] have demonstrated comparable performance to the standard ones. Based on the way of training the model and the presence/absence of ground-truth data, deep learning-based image registration can be categorized into fully supervised [10], unsupervised [7], and weakly-supervised learning-based methods [11].

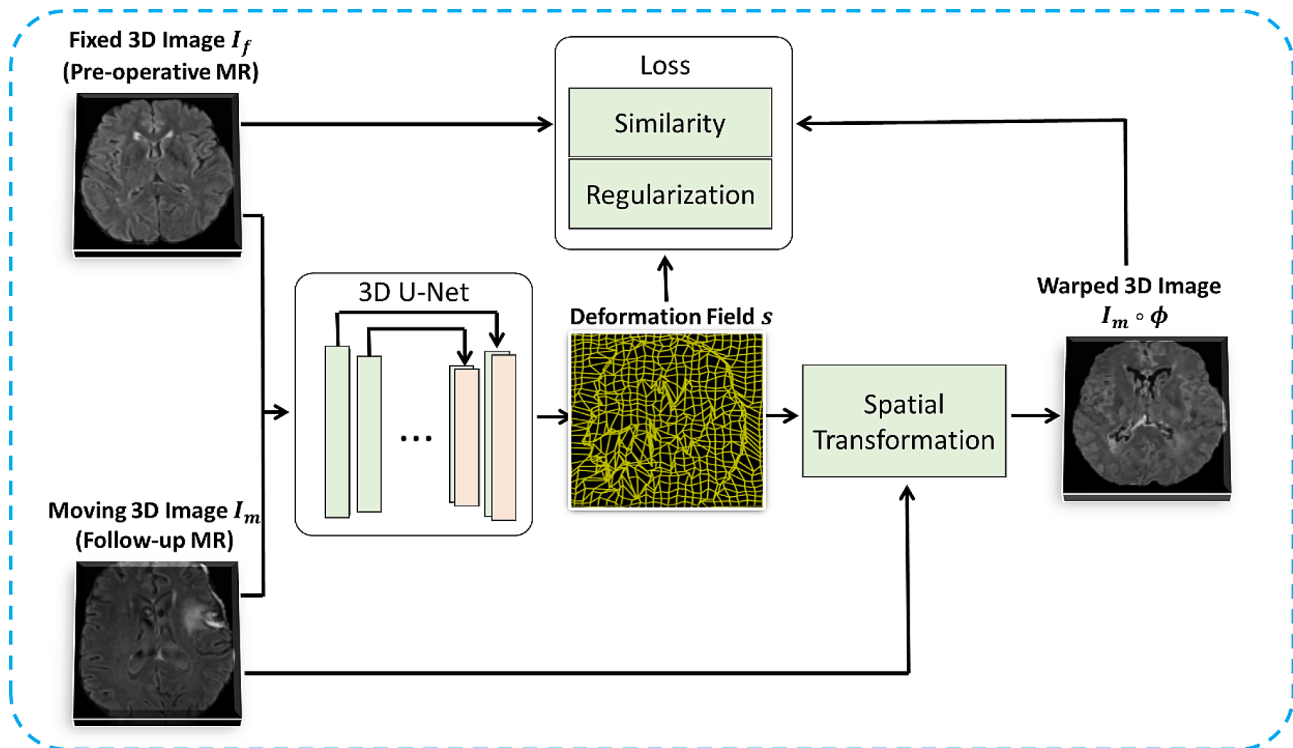
Deformable registration of baseline pre-surgical and follow-up volumetric MRI scans is essential to the treatment plan and diagnosis of brain gliomas [12, 13] in neuro-/radiation oncology. Although the decent performance of modern deep learning-based deformable registration algorithms concerning computation time and accuracy, their performance on the brain MR images containing pathologies is far from perfect. This issue remains clinically unresolved. One of the reasons is the high deformation of the brain tissues induced by the resected tumor after the surgery. Some heavy deformations are not restricted to the lesion area and can affect the entire brain. Another reason is that the intensity profiles of the

pre-surgical and follow-up scans are inconsistent. A third reason is the lack of correspondence between the pre- and post-operative images [14].

Considering the above points, establishing spatial correspondences between MRI scans acquired at two time-point (e.g., pre- and post-operation) from glioma patients can further help in the understanding of the mechanisms of these tumors. Precisely, for potential tumor recurrence and tumor infiltration, it can further contribute to developing predictive modeling for related pathophysiological processes. It can also aid in understanding the biophysical dynamic and plasticity characteristic of brain tissues besides neurosurgical planning. Few studies have addressed this correspondence utilizing different convolutional neural networks (CNNs) for deformable registration. These methods include proposing a one-stage [15–19], two-stage [20] or three-stage [21] registration pipeline for this purpose. Despite the inspiring performance of these models, obtaining accurate results may remain challenging for deep learning-based deformable registration because of the large deformation of the healthy images. In addition, most of these methods need initial rigid registration performed separately before deformable registration. They involve more than a single stage (two- and three-stage methods) and cannot be fully automated. In this paper, we develop an unsupervised 3D convolutional U-Net-based deformable image registration (ConvUNet-DIR) framework to estimate the correspondences between pre-surgical and follow-up 3D MRI scans for glioma patients. Our method can perform an end-to-end deformable registration (i.e., without human intervention) and does not need supervision (e.g., data such as ground-truth registration fields or anatomical landmarks) during the network training. The network resembles a multi-scale U-Net style architecture [22] to capture the feature maps, and its parameters are updated during the training by minimizing the dissimilarity between the baseline and warped images.

## Materials and methods

The ConvUNet-DIR framework (Fig. 1) proposed in this study is to estimate the optimal parameterized mapping function  $\phi$  between a baseline pre-operative MR image (fixed image,  $I_f$ ) and the follow-up MR image (moving image,  $I_m$ ). The  $\phi$  is a nonlinear voxel-wise correspondence between  $I_f$  and  $I_m$ . The deformed/warped image ( $I_m \circ \phi$ ) from  $I_m$  can be registered to  $I_f$ . Here, the global mapping function  $\phi(x) = x + s(x)$  is formed



**Fig. 1** The framework of our unsupervised convolutional U-Net based deformable image registration (ConvUNet-DIR) of pre-operative and follow-up magnetic resonance images of glioma patients

by an identity spatial transform and deformation field  $s$ . Once the framework has been trained, the deformation field  $s$  can be obtained from a new pair of MRI scans.

#### Dataset

A multi-institutional dataset ( $n=160$ ) of multi-parametric MRI scans, representing the training and validation sets of the Brain Tumor Sequence Registration (BraTS-Reg 2022) challenge, was used in this study [23]. The patients were diagnosed with glioma and clinically scanned with multi-parametric MRI acquisition protocol. The multi-parametric MRI scans of each patient included T1-weighted (T1), contrast-enhanced T1 (T1-ce), T2, and T2 Fluid Attenuated Inversion Recovery (T2-FLAIR or simply FLAIR). The images were acquired at two time-point: pre-operative (treatment-naïve) and follow-up. The follow-up scans range from 27 days to 37 weeks. The images have been manipulated already. All scans were first transformed into the same coordinate system and then rigidly co-registered to the same anatomical coordinate using the greedy diffeomorphic registration algorithm [24]. Then, the images were skull-stripped by extracting the brain tissue and sampled down to  $240 \times 240 \times 155$  dimensions with  $1 \text{ mm}^3$  spatial resolution. The brain extraction was performed using the Brain Mask Generator, a deep learning-based algorithm [25].

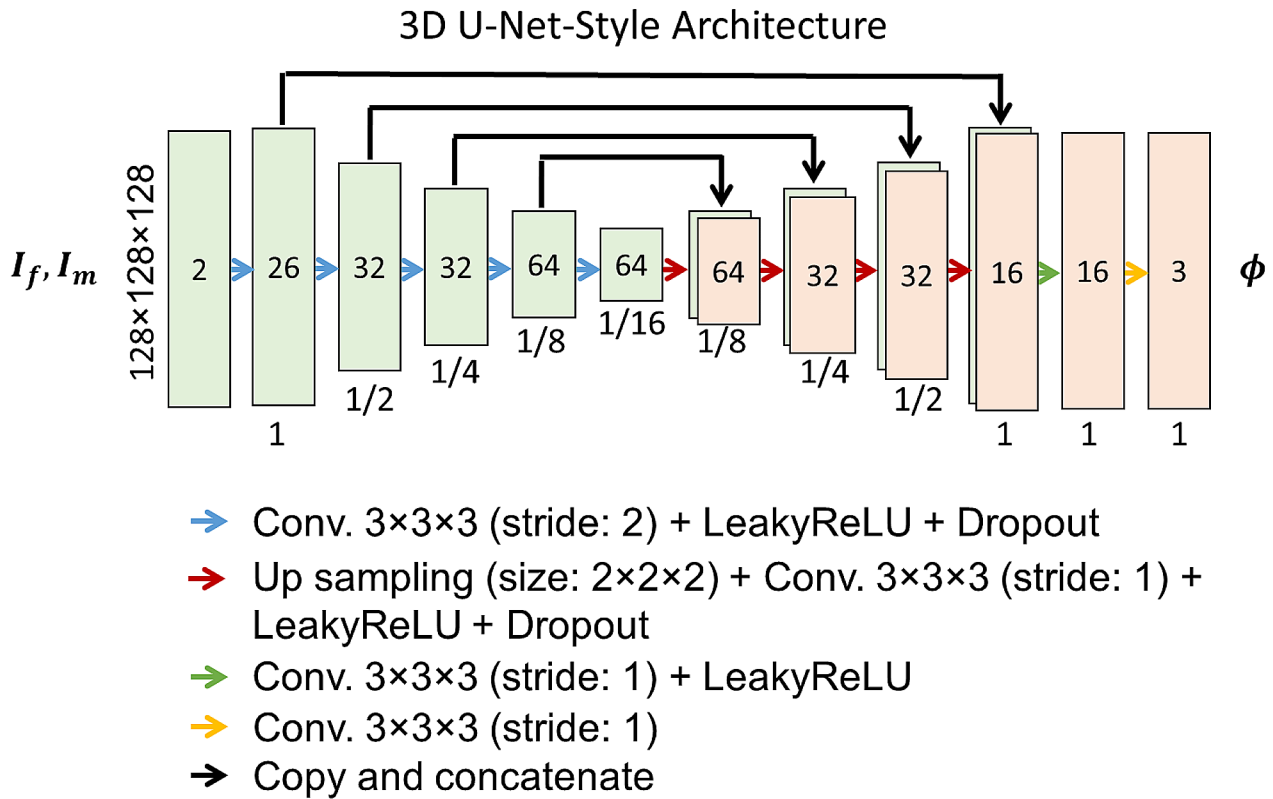
Precisely, non-cerebral tissues such as the skull, scalp, and dura were removed from all MRI scans.

#### Preprocessing

We implemented some preprocessing to the multi-parametric MRI data before being utilized to train the proposed model. *First*, we cropped the multi-parametric MRI data into smaller sizes of  $224 \times 224 \times 155$  dimensions by excluding peripheral voxels with no information. *Then*, we resized the data into  $128 \times 128 \times 128$  dimensions. *Next*, the image data were normalized using the zero mean and unit variance technique and scaled the data to the  $[0, 1]$  range. Accurate registration requires the input MRI scans to be normalized (i.e., voxel intensities range from 0 to 1) to produce consistent results for the images acquired with different scanners/imaging protocols. *Finally*, we randomly split the data into 64% ( $n=102$ ) as a training set, 16% ( $n=26$ ) as a validation set, and 20% ( $n=32$ ) as a test set.

#### CNN architecture

Our proposed ConvUNet-DIR framework uses U-Net [22] as a core (Fig. 2). It takes a two-channel 3D MR image representing the concatenation of  $I_f$  and  $I_m$  as input. The input size of the network is  $128 \times 128 \times 128 \times 2$ . The convolutional network consists of an encoder with four downsampling layers, a bottleneck or bridge, and



**Fig. 2** The convolutional U-Net-style architecture. Numbers at the bottom of the blocks represent the spatial resolution ratio of each volume to the input volume. Numbers inside the block indicate the extracted features. Arrows donate different operations

a decoder network of four upsampling layers. In the encoder, 3D convolutional layers ( $3 \times 3 \times 3$  kernel size and  $2 \times 2 \times 2$  stride) and the LeakyReLU activation are used to extract features and down-sample/reduce the spatial resolution of the volumes by a factor of  $1/2$  at different spatial resolution levels. The base level of the model or bottleneck is composed of two convolutional layers ( $3 \times 3 \times 3$  kernel size and  $1 \times 1 \times 1$  stride). The decoder network is made up of four upsampling convolutional layers ( $3 \times 3 \times 3$  kernel size and  $1 \times 1 \times 1$  stride) and the LeakyReLU activation function to reconstruct the deformation field. The encoder and decoder networks are concatenated via skip connections to detect and combine information at different spatial levels to produce the deformation field. Three additional convolutional layers ( $3 \times 3 \times 3$  kernel size and  $1 \times 1 \times 1$  stride) with LeakyReLU activation function are added after the decoder network. The final layer is a convolutional layer ( $3 \times 3 \times 3$  kernel size and  $1 \times 1 \times 1$  stride) with a linear activation used for a regression output. The linear activation in this layer enables positive and negative values as the output of the deformation field. The output layer of the network is a deformation field  $s$  with a size of  $128 \times 128 \times 128 \times 3$ .

#### Spatial transformation layer

The spatial transformation layer computes  $I_m \circ \phi$ . The position of individual voxels in  $I_m$  is determined in the space of  $I_f$ . In this layer,  $\phi$  is used to warp  $I_m$  and then obtain  $I_m \circ \phi$ . To ensure that the spatial transformation layer is differentiable, we used linear interpolation to estimate the voxel value of  $I_m$  in the  $x$ ,  $y$ , and  $z$  coordinates. As a result, back-propagation of the errors through the network could be implemented during the training.

#### Loss function

We implemented an unsupervised loss function ( $Loss$ ) to evaluate the model using only the  $I_f$  and  $I_m \circ \phi$ . The loss function consists of a similarity loss ( $L_{sim}$ ) and a regularization term ( $L_{smooth}$ ) that penalizes large deformations to produce smooth registration fields.  $L_{sim}$  penalizes the difference in appearance, whereas  $L_{smooth}$  penalizes the first derivative of  $s$  to produce a fine registration field. The  $Loss$  is defined as:

$$Loss(I_f, I_m \circ \phi) = L_{sim}(I_f, I_m \circ \phi) + \lambda \bullet L_{smooth}(I_f, I_m \circ \phi)$$

where  $\lambda$  is a regularization parameter.



The  $L_{sim}$  was set to the negative local normalized cross-correlation (NCC) coefficient [26] of  $I_f$  and  $I_m \circ \phi$ . This intensity-based similarity measure was found to be optimal for single/mono-modality image registration where the image pair shares similar intensity distribution [2]. The NCC coefficients were calculated over a volume of  $9 \times 9 \times 9$ . The  $L_{sim}$  is defined as:

$$L_{sim}(I_f, I_m \circ \phi) = -NCC(I_f, I_m \circ \phi)$$

Minimizing the  $L_{sim}$  without applying constraints can lead to  $I_m \circ \phi$  with unrealistic organ appearances. Obtaining a smooth  $s$  requires implementing a diffusion regularizer on the spatial gradients ( $\nabla$ ) of deformation  $s$  as follows:

$$L_{smooth}(I_f, I_m \circ \phi) = \sum \|\nabla s(I_f, I_m \circ \phi)\|^2$$

The spatial gradients were approximated using differences between neighboring voxels.

### Training and validation

The ConvUNet-DIR model was trained using an unsupervised manner for deformable registration of a pair of MRI scans ( $I_f, I_m$ ) with volumes of  $128 \times 128 \times 128$  on a training data set ( $n=102$ ). Adam optimizer, implementing the gradient descent approach, was set to optimize the learning model with a 0.0001 learning rate. The regularization parameter,  $\lambda$ , was set to 1. We assessed different settings for  $\lambda$ , including 0 (no regularization), 0.1, 0.2, 0.5, and 1. The  $\lambda=1$  found to work best with our task. The model was trained for 150 epochs using a batch size of 1. Our GPU memory does not permit us to use a larger batch size. During the network training, each pair of MRI scans is concatenated into a 2-channel 3D image and fed into the 3D U-Net. The deformation field  $s$  was computed through the convolutional layers of U-Net. The spatial transformation layer was used to warp  $I_m$  into  $I_m \circ \phi$  by using linear resampling. The network parameters are regularly tuned during the training by minimizing the dissimilarity between the  $I_f$  and  $I_m \circ \phi$ . The  $s$  is punished by regularization terms to encourage smoothness (i.e., regularize the predicted deformation). During the training, the model was validated on the validation set ( $n=26$ ) to assess its generalizability and to update the hyper-parameters.

The training was performed using Keras API (version 2.10) with a Tensorflow (version 2.10) platform as the backend in Python (version 3.10, Python Software Foundation, Wilmington, DE, USA) by using an NVIDIA Quadro M1200 4 GB GPU. The trained model takes less than 2 s to register a pair of MRI scans for a new patient, making its deployment to the clinical practice feasible.

### Evaluation

The model was assessed to evaluate its performance on the registration of pairs of  $I_f$  and  $I_m$  on a test set ( $n=32$ ). One of the utilized metrics is the Dice similarity coefficient. It is used to estimate the volume overlap of the brain fields which was determined using the generated brain masks. The Dice score of the warped mask ( $A_m \circ \phi$ ) and the fixed mask ( $A_f$ ) of the  $I_f$  is calculated as follows:

$$Dice(A_m \circ \phi, A_f) = \frac{2|A_m \circ \phi \cap A_f|}{|A_m \circ \phi| + |A_f|}$$

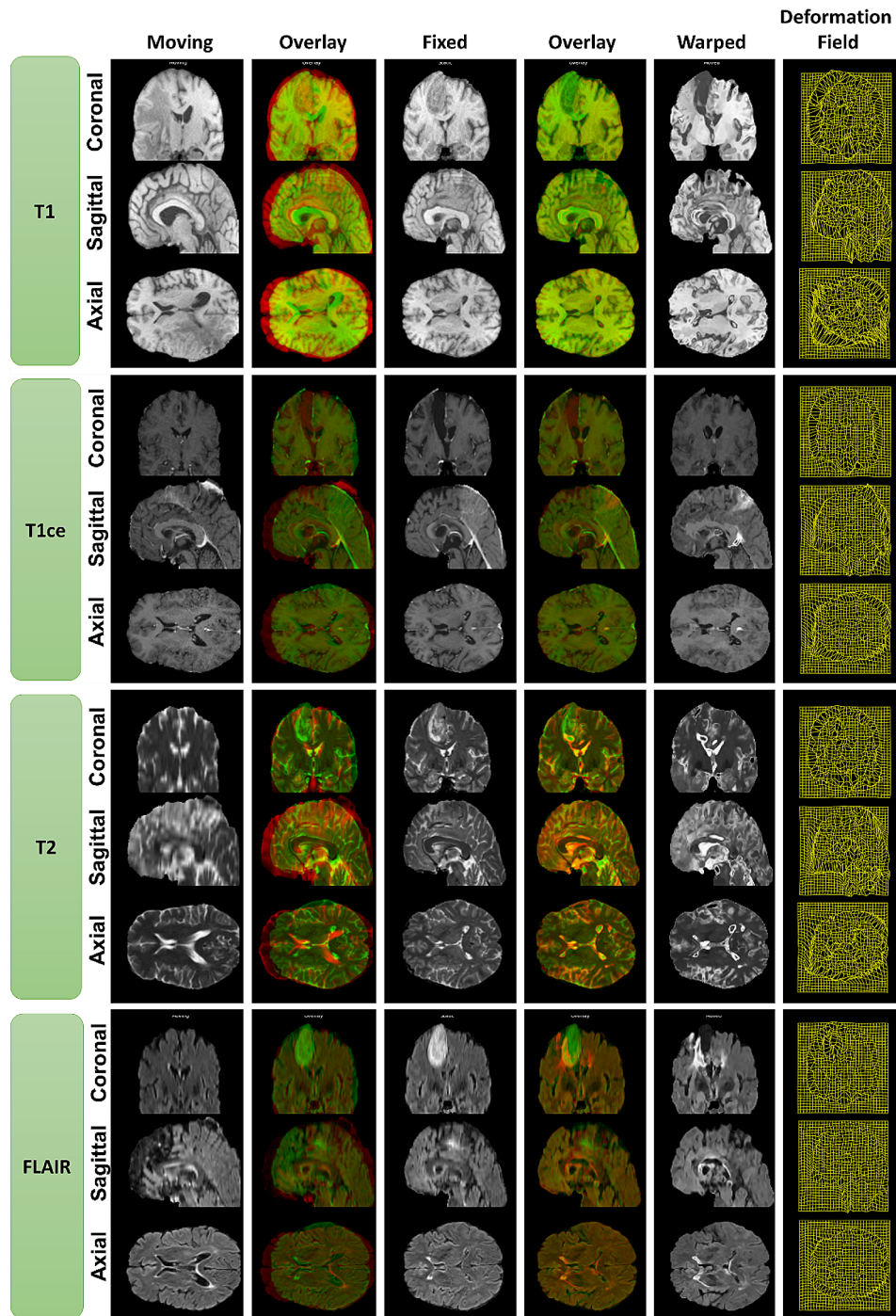
Another metric used to evaluate our proposed method is the structural similarity index (SSIM). This metric simulates the human-perceived quality of images by comparing two images. Mathematically, it is defined as:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)},$$

where  $\mu$  is the mean image intensity,  $\sigma^2$  is the variance of the image,  $\sigma_{xy}$  is the covariance of the fixed ( $x$ ) and moving ( $y$ ) images, and  $C_1$  and  $C_2$  are constants added to stabilize the division with a weak denominator. The performance of the ConvUNet-DIR model was also compared with the open-source VoxelMorph (VM1 and VM2) methods [7], deep learning-based algorithms. We trained the two versions of the VoxelMorph algorithms from scratch on the BraTS-Reg 2021 dataset for a fair comparison.

### Results

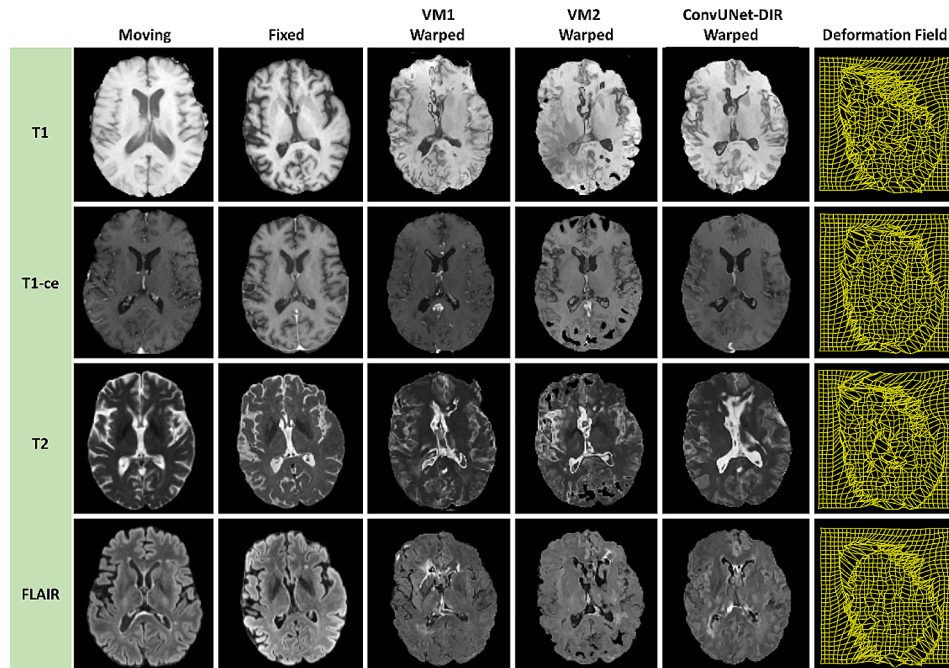
The results of the ConvUNet-DIR model are reported for pairs of T1, T1-ce, T2, and FLAIR MRI scans on the test set ( $n=32$ ). Figure 3 shows representative registration results of aligning two images for one patient on axial, coronal, and sagittal views. The figure also displays the overlay of the fixed (pre-operative) and moving (follow-up) images, the overlay of the fixed and warped (deformed) images, and the deformation field. The overlay of the fixed and moving images shows variable degrees of deformation, whereas the warped image is nearly overlapping on the fixed image. The figure shows the fixed image (in green color) on top of the moving image (in red color) and the warped image (in green color) on top of the fixed image (in red color). It looks like the deformed image has a slightly lower spatial resolution. The results of the deformable registration of the four MRI scans with ConvUNet-DIR and VoxelMorph (VM1 and VM2) models are shown in Fig. 4. It is clear that the ConvUNet-DIR model provides better registration results compared to both VM1 and VM2 algorithms.



**Fig. 3** Example registration results of pairs of T1, T1-ce, T2, and FLAIR 3D MRI scans on three planes (axial, coronal, and sagittal) by our proposed ConvUNet-DIR model for one patient in the test set. Overlays of the fixed image (green color) and moving/warped image (red color) on top of each other and the deformation field are also displayed

Figure 5 shows the effects of the regularization parameter,  $\lambda$ , on the warped image (T1 as an example) generated by the ConvUNet-DIR model. The results demonstrate that the best registration is obtained with the regularization weight set to 1, which agrees with that reported by Chen et al. [27].

Table 1 presents a quantitative summary of all registration results with the ConvUNet-DIR, VM1, and VM2 models on the test set ( $n=32$ ). The mean SSIM and Dice scores were reported for all correspondences. The SSIM and the Dice score were calculated for the brain fields of a 3D MR image. From the results, we can observe that



**Fig. 4** Example registration results of pairs of T1, T1-ce, T2, and FLAIR 3D MRI scans with the ConvUNet-DIR and VoxelMorph (VM1 and VM2) models

the performance of the ConvUNet-DIR method is consistently better compared to VM1 and VM2. The table also reports the CPU computation time required by each model to perform deformable registration on a pair of volumetric MRI scans for a new patient. The computation time of all models was about 1 s, which is significantly shorter than the traditional methods. The learning parameters of all models also were reported in the table.

## Discussion

Deformable registration of MRI scans of patients with pathologies is challenging. In this study, we developed an unsupervised deep learning-based deformable registration algorithm to establish complex correspondences between the pre- and post-operative 3D MRI scans of patients with glioma. Given a pair of MRI scans at two time-point as input, the ConvUNet-DIR model computes the voxel-wise deformation between the two images. In addition to its 3D nature that accounts for adjacent slices in the volumetric image data, ConvUNet-DIR does not need supervision during its training.

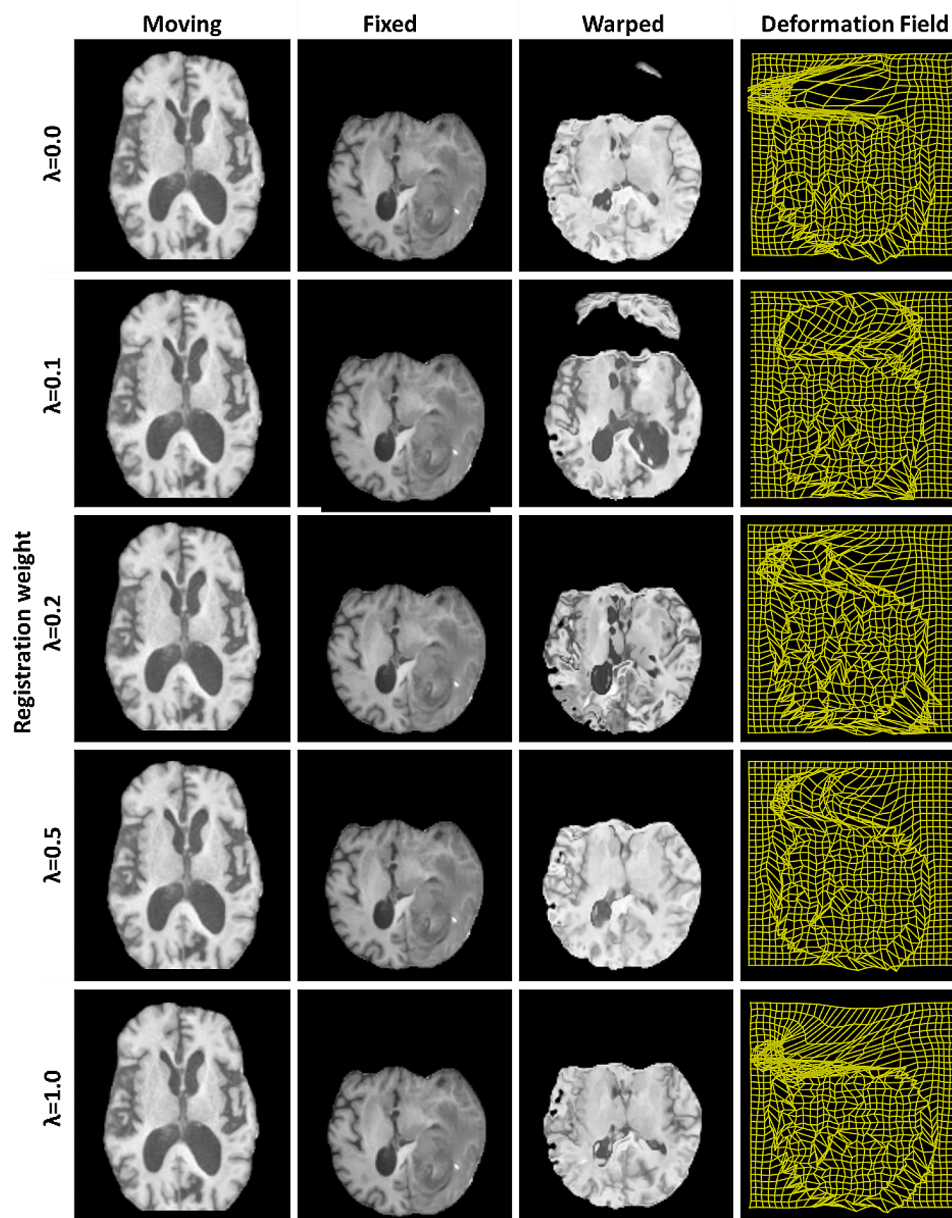
The qualitative results are illustrated in Figs. 3 and 4 for an example patient in the test set. The ConvUNet-DIR method demonstrated registration with high accuracy while preserving deformation smoothness. It also seemed to preserve the original intensity distribution better, while VM1 and VM2 appeared to have an impact on the intensity values. The quantitative registration results are presented in Table 1. The table shows impressive results achieved by our proposed model with a mean SSIM of 0.908 and a mean Dice score of 0.975. The execution

time of the ConvUNet-DIR model was about 1 s, which is significantly shorter than the conventional methods. This advantage signifies its clinical deployment for critical time applications in neuro-oncology and radiation oncology.

This study also compared the performance of the ConvUNet-DIR with the VoxelMorph models [7], as shown in Table 1. Both methods are deep learning-based algorithms that allow unsupervised 3D deformable registration. The results achieved by ConvUNet-DIR (SSIM=0.908, Dice=0.975) were superior to VM1 (SSIM=0.893, Dice=0.969) and VM2 (SSIM=0.857, Dice=0.957). Our method produces a more competitive deformation field than VoxelMorph. We also compared the performance of the ConvUNet-DIR with the current state-of-the-art supervised and unsupervised deep learning-based approaches reported in the literature utilized public brain MRI datasets. Our method (Dice=0.980) outperformed Han et al. [19] (Dice=0.839), Martin et al. [28] (Dice=0.756), Wu et al. [29] (Dice=0.873), Meng et al. [30] (Dice=0.654), Kuang and Schmah [9] (Dice=0.533), Mok and Chung [31] (Dice=0.770), Huang et al. [32] (Dice=0.707), Xu et al. [33] (Dice=0.830), Dey et al. [34] (Dice=0.781), Fan et al. [35] (Dice=0.788), Liu et al. [36] (Dice=0.909), Chen et al. [37] (Dice=0.873), Wang et al. [38] (Dice=0.731) methods using unsupervised approach; and Zhu et al. [39] (Dice=0.637) method using supervised approach.

This study has demonstrated that the proposed method can perform an end-to-end deformable registration of a pair of volumetric brain MR images without human





**Fig. 5** Results of effects of the regularization parameter,  $\lambda$ , on the warped image (T1 as an example) produced by the ConvUNet-DIR model. Rows show different regularization weights. Columns show the moving image, fixed image, warped image, and the deformation field, respectively

intervention. It does not require pre-alignment compared to similar algorithms that require at least two steps to perform the registration [20, 21]. Compared to supervised registration networks, our method does not require additional segmentation masks or anatomical marks for network training. ConvUNet-DIR could also support multimodality image registration by substituting NCC with mutual information. Combining T1 pre- and post-contrast, T2, and FLAIR might achieve even better registration. The reason is that different scans are implicitly co-registered within the pre- or post-operative data (just some rigid registration), barring image distortion from

the acquisition). It is worth noting that ConvUNet-DIR works for a broad window of time (27 days to 37 months follow-up).

We can briefly discuss the limitations of this study in three points. First, our model requires resizing all images, which may cause information loss. Employing a tri-linear interpolation could minimize this issue. Second, due to the restriction of the GPU memory, we set the batch to only 1 sample. Using a small batch size may cause registration errors. Third, this work is a multi-institution study with a relatively small dataset. To assess the performance of this model in a more general way in clinical practice,

**Table 1** Comparison results of the proposed model (ConvUNet-DIR), VM1, and VM2 models for deformable registration of pairs of 3D MRI scans evaluated on a testing data set ( $n = 32$ ). Results were reported in the form of mean  $\pm$  1 standard deviation. Values in bold indicate the best results

Modality	Model	SSIM	Dice Score	CPU (s)	Parameters
T1	VM1	0.876 $\pm$ 0.052	0.965 $\pm$ 0.007	0.83 $\pm$ 0.17	259,675
	VM2	0.818 $\pm$ 0.120	0.943 $\pm$ 0.015	1.34 $\pm$ 0.22	300,547
	ConvUNet-DIR	<b>0.898 <math>\pm</math> 0.041</b>	<b>0.976 <math>\pm</math> 0.005</b>	1.32 $\pm$ 0.14	328,227
T1-ce	VM1	0.906 $\pm$ 0.039	0.974 $\pm$ 0.008	0.83 $\pm$ 0.17	259,675
	VM2	0.848 $\pm$ 0.104	0.959 $\pm$ 0.012	1.34 $\pm$ 0.22	300,547
	ConvUNet-DIR	<b>0.919 <math>\pm</math> 0.035</b>	<b>0.980 <math>\pm</math> 0.008</b>	1.32 $\pm$ 0.14	328,227
T2	VM1	0.875 $\pm$ 0.053	0.962 $\pm$ 0.010	0.83 $\pm$ 0.17	259,675
	VM2	0.816 $\pm$ 0.118	0.937 $\pm$ 0.018	1.34 $\pm$ 0.22	300,547
	ConvUNet-DIR	<b>0.894 <math>\pm</math> 0.043</b>	<b>0.971 <math>\pm</math> 0.008</b>	1.32 $\pm$ 0.14	328,227
FLAIR	VM1	0.878 $\pm$ 0.046	0.964 $\pm$ 0.010	0.83 $\pm$ 0.17	259,675
	VM2	0.821 $\pm$ 0.113	0.945 $\pm$ 0.022	1.34 $\pm$ 0.22	300,547
	Ours	<b>0.896 <math>\pm</math> 0.041</b>	<b>0.973 <math>\pm</math> 0.008</b>	1.32 $\pm$ 0.14	328,227
Overall (T1, T1ce, T2, & FLAIR)	VM1	0.884 $\pm$ 0.011	0.966 $\pm$ 0.004	0.83 $\pm$ 0.17	259,675
	VM2	0.826 $\pm$ 0.011	0.946 $\pm$ 0.006	1.34 $\pm$ 0.22	300,547
	ConvUNet-DIR	<b>0.902 <math>\pm</math> 0.009</b>	<b>0.975 <math>\pm</math> 0.004</b>	1.32 $\pm$ 0.14	328,227

SSIM: structural similarity index; VM1: voxelmorph1; VM2: voxelmorph2

we recommend training the model using more data samples from several institutions for better generalizability.

## Conclusions

We developed a ConvUNet-DIR framework based on unsupervised learning to establish correspondence between a pair of 3D MRI scans acquired at two time-point from patients with glioma. The proposed method demonstrated registration accuracy superior to the state-of-the-art VoxelMorph (VM1 and VM2) registration tools (open-sourced learning-based registration algorithms) and other supervised/unsupervised deep learning-based algorithms reported in the literature. It can perform an automated deformable registration of a pair of 3D MRI scans for glioma patients. The model could provide accurate, efficient, and robust deformable registration without needing pre-alignment and labeling, resulting in a significantly shorter registration time. This method has the potential for application in clinical practice in neuro-/radiation oncology.

## Abbreviations

3D Three	dimensional
ANTS	Advanced normalization tools
BraTS-Reg 2022	Brain Tumor Sequence Registration 2022 challenge
CNNs	Convolutional neural networks
ConvUNet-DIR	Convolutional U-Net-based deformable image registration
MRI	Magnetic resonance imaging
NCC	Normalized cross-correlation
SSIM	Structural similarity index
T1	T1-weighted
T1-ce	T1-weighted with contrast enhancement
T2	T2-weighted
T2-FLAIR	T2-weighted Fluid Attenuated Inversion Recovery
VM1	VoxelMorph1
VM2	VoxelMorph2

## Acknowledgements

The authors express their gratitude to Princess Nourah bint Abdulrahman University Researchers Supporting Project (Grant No. PNURSP2024R10), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

## Author contributions

AO contributed to the conception and design of the study, methodology, technical aspects, drafted the manuscript, and revised it for important intellectual content. KA, NT, and BS contributed to revising the manuscript for important intellectual content. All authors contributed to manuscript revision and approved the submitted version.

## Funding

This work is funded by Researchers Supporting Project (Grant No. PNURSP PNURSP2024R10), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

## Data availability

The datasets generated during and/or analyzed during the current study are available at the Brain Tumor Sequence Registration (BraTS-Reg) Challenge repository (<https://www.med.upenn.edu/cbica/brats-reg-challenge/>).

## Declarations

### Ethical approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

### Author details

<sup>1</sup>Department of Medical Physics, Al-Neelain University, Khartoum 11121, Sudan

<sup>2</sup>Department of Physics, College of Science, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

<sup>3</sup>Department of Radiation Oncology, American University of Beirut Medical Center, Beirut, Lebanon

Received: 26 September 2023 / Accepted: 13 May 2024

Published online: 21 May 2024



## References

- Durrleman S, Prastawa M, Charon N, et al. Morphometry of anatomical shape complexes with dense deformations and sparse parameters. *NeuroImage*. 2014;101:35–49. <https://doi.org/10.1016/j.neuroimage.2014.06.043>.
- Klein S, Staring M, Murphy K, et al. Elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imag*. 2009;29(1):196–205. <https://doi.org/10.1109/TMI.2009.2035616>.
- Avants BB, Tustison NJ, Song G, et al. A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*. 2011;54(3):2033–44. <https://doi.org/10.1016/j.neuroimage.2010.09.025>.
- Vercauteren T, Pennec X, Perchant A, et al. Diffeomorphic demons: efficient nonparametric image registration. *NeuroImage*. 2009;45(1):S61–72. <https://doi.org/10.1016/j.neuroimage.2008.10.040>.
- Osman AFI. Radiation Oncology in the era of Big Data and Machine Learning for Precision Medicine. In: Aceves-Fernandez MA, editor. *Book: Artificial Intelligence - Applications in Medicine and Biology*. Volume 1. London, UK: IntechOpen; 2019. pp. 41–70. <https://doi.org/10.5772/intechopen.84629>.
- Zou J, Gao B, Song Y, Qin J. A review of deep learning-based deformable medical image registration. *Front Oncol*. 2022;12:1047215. <https://doi.org/10.3389/fonc.2022.1047215>.
- Balakrishnan G, Zhao A, Sabuncu MR, Gutttag J, Dalca AV. Voxel-Morph: a learning framework for deformable medical image registration. *IEEE Trans Med Imaging*. 2019;38(8):1788–800. <https://doi.org/10.1109/TMI.2019.2897538>.
- de Vos BD, Berendsen FF, Viergever MA, Sokooti H, Staring M, Išgum I. A deep learning framework for unsupervised affine and deformable image registration. *Med Image Anal*. 2019;52:128–43. <https://doi.org/10.1016/j.media.2018.11.010>.
- Kuang D, Schmah T. FAIM – a ConvNet Method for unsupervised 3D Medical Image Registration. In: Suk HI, Liu M, Yan P, Lian C, editors. *Machine Learning in Medical Imaging*. MLMI 2019. Lecture Notes in Computer Science. Volume 11861. Cham: Springer; 2019. [https://doi.org/10.1007/978-3-030-32692-0\\_74](https://doi.org/10.1007/978-3-030-32692-0_74).
- Yang X, Kwitt R, Styner M, Niethammer M. Quicksilver: Fast predictive image registration—a deep learning approach. *NeuroImage*. 2017;158:378–96. <https://doi.org/10.1016/j.neuroimage.2017.07.008>.
- Xu Z, Niethammer M, DeepAtlas. Joint semi-supervised learning of image registration and segmentation, in: *International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI 2019* Springer Nature Switzerland AG:420–9, 2019. [https://doi.org/10.1007/978-3-030-32245-8\\_47](https://doi.org/10.1007/978-3-030-32245-8_47).
- Heiss WD, Raab P, Lanfermann H. Multimodality assessment of brain tumors and tumor recurrence. *J Nucl Med*. 2011;52(10):1585–600.
- Price SJ, Jena R, Burnet NG, Carpenter TA, Pickard JD, Gillard JH. Predicting patterns of glioma recurrence using diffusion tensor imaging. *Eur Radiol*. 2007;17(7):1675–84.
- Dean BL, Drayer BP, Bird CR, Flom RA, Hodak JA, Coons SW, Carey RG. Gliomas: classification with mr imaging. *Radiol*. 1990;174(2):411–5.
- Zeineldin RA, Karar ME, Mathis-Ullrich F, Burgert O. Self-supervised iRegNet for the Registration of Longitudinal Brain MRI of Diffuse Glioma Patients. [preprint] arXiv:2211.11025v1, 2022. <https://arxiv.org/abs/2211.11025>.
- Meng M, Bi L, Feng D, Kim J. Brain Tumor Sequence Registration with Non-iterative Coarse-to-fine Networks and Dual Deep Supervision. [preprint] arXiv:2211.07876, 2022. <https://arxiv.org/abs/2211.07876>.
- Mok TCW, Chung ACS. Unsupervised Deformable Image Registration with Absent Correspondences in Pre-operative and Post-Recurrence Brain Tumor MRI Scans. [preprint] arXiv:2206.03900v1, 2022. <https://arxiv.org/abs/2206.03900>.
- Wodzinski M, Jurgas A, Marini N, Atzori M, Muller H. Unsupervised Method for Intra-patient Registration of Brain Magnetic Resonance Images based on Objective Function Weighting by Inverse Consistency: Contribution to the BraTS-Reg Challenge. [preprint] arXiv:2211.07386v1, 2022. <https://arxiv.org/abs/2211.07386>.
- Han L, Dou H, Huang Y, Yap PT. Deformable Registration of Brain MR images via a hybrid loss. In: Auberville M, Zimmerman D, Heinrich M, editors. *Biomedical Image Registration, Domain Generalisation and out-of-distribution analysis*. MICCAI 2021. Lecture notes in Computer Science(). Volume 13166. Cham: Springer; 2022. [https://doi.org/10.1007/978-3-030-97281-3\\_20](https://doi.org/10.1007/978-3-030-97281-3_20).
- Abderezai J, Pionteck A, Chopra A, Kurt M. 3D Inception-Based TransMorph: Pre- and Post-operative Multi-contrast MRI Registration in Brain Tumors. [preprint] arXiv:2212.04579, 2022. <https://arxiv.org/abs/2212.04579>.
- Mok TCW, Chung ACS. Robust Image Registration with Absent Correspondences in Pre-operative and Follow-up Brain MRI Scans of Diffuse Glioma Patients. [preprint] arXiv:2210.11045v2, 2022. <https://arxiv.org/abs/2210.11045>.
- Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Springer; 2015. pp. 234–41.
- Baheti B, Waldmannstetter D, Chakrabarty S et al. The Brain Tumor Sequence Registration Challenge: Establishing Correspondence between Pre-Operative and Follow-up MRI scans of diffuse glioma patients. [preprint] arXiv:2112.06979, 2021. <https://arxiv.org/abs/2112.06979>.
- Joshi S, Davis B, Jomier M, Gerig G. Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage*. 2004;23:151–60.
- Thakur S, Doshi J, Pati S, Rathore S, Sako C, Bilello M, Ha SM, Shukla G, Flanders A, Kotrotsou, et al. Brain extraction on mri scans in presence of diffuse glioma: multi-institutional performance evaluation of deep learning methods and robust modality-agnostic training. *NeuroImage*. 2020;220:117081.
- Penney GP, Weese J, Little JA, Desmedt P, Hill D, Hawkes D. A comparison of similarity measures for use in 2-D-3-D medical image registration. *IEEE Trans Med Imaging* 1998;17:586–95. <https://doi.org/10.1109/42.730403>.
- Chen J, Li Y, Du Y, Frey EC. Generating anthropomorphic phantoms using fully unsupervised deformable image registration with convolutional neural networks. *Med Phys*. 2020;47(12):6366–80. <https://doi.org/10.1002/mp.14545>.
- Martin AD, Kim B, Ye JC. Unsupervised deformable image Registration using polyphase UNet for 3D brain MRI volumes. *Investig Magn Reson Imaging*. 2020;24(4):223–31.
- Wu G, Kim M, Wang Q, Gao Y, Liao S, Shen D. Unsupervised deep feature learning for deformable registration of MR brain images. *Med Image Comput Comput Assist Interv*. 2013;16(Pt 2):649–56. [https://doi.org/10.1007/978-3-642-40763-5\\_80](https://doi.org/10.1007/978-3-642-40763-5_80).
- Meng M, Bi L, Feng D, Kim J. Non-iterative coarse-to-fine Registration based on single-pass deep cumulative learning. In: Wang L, Dou Q, Fletcher PT, Speidel S, Li S, editors. *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*. MICCAI 2022. Lecture Notes in Computer Science. Volume 13436. Cham: Springer; 2022. [https://doi.org/10.1007/978-3-031-16446-0\\_9](https://doi.org/10.1007/978-3-031-16446-0_9).
- Mok TCW, Chung ACS. Conditional deformable image Registration with Convolutional Neural Network. In: de Buijnne M, et al. editors. *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*. MICCAI 2021. Lecture notes in Computer Science(). Volume 12904. Cham: Springer; 2021. [https://doi.org/10.1007/978-3-030-87202-1\\_4](https://doi.org/10.1007/978-3-030-87202-1_4).
- Huang M, Ren G, Zhang S, Zheng Q, Niu H. An unsupervised 3D image Registration Network for Brain MRI Deformable Registration. *Comput Math Methods Med*. 2022;2022(10):9246378. <https://doi.org/10.1155/2022/9246378>.
- Xu J, Chen EZ, Chen X, Chen T, Sun S. Multi-scale Neural ODEs for 3D Medical Image Registration. [preprint] arXiv:2106.08493, 2021. <https://arxiv.org/abs/2106.08493>.
- Dey N, Schlemper J, Salehi S, Zhou B, Gerig G, Sofka M, ContraReg. Contrastive Learning of Multi-modality Unsupervised Deformable Image Registration. [preprint] arXiv:2206.13434, 2022. <https://arxiv.org/abs/2206.13434>.
- Fan X, Li Z, Li Z, Wang X, Liu R, Luo Z, Huang H. Automated Learning for Deformable Medical Image Registration by Jointly Optimizing Network Architectures and Objective Functions. [preprint] arXiv:2203.06810, 2022. <https://arxiv.org/abs/2203.06810>.
- Liu Y, Zuo L, Han S, Xue Y, Prince JL, Carass A. Coordinate Translator for Learning Deformable Medical Image Registration. [preprint] arXiv:2203.03626, 2022. <https://arxiv.org/abs/2203.03626>.
- Chen J, Liu Y, He Y, Du Y. Deformable Cross-Attention Transformer for Medical Image Registration. [preprint] arXiv:2303.06179, 2023. <https://arxiv.org/abs/2303.06179>.
- Wang Z, Wang H, Wang Y. Pyramid Attention Network for Medical Image Registration. [preprint] arXiv:2402.09016, 2024. <https://arxiv.org/abs/2402.09016>.
- Zhu Z, Cao Y, Qin C, Rao Y, Lin D, Dou Q, Ni D, Wang Y. Joint affine and deformable three-dimensional networks for brain MRI registration. *Med Phys*. 2021;48(3):1182–96. <https://doi.org/10.1002/mp.14674>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.