MetaMorph: Learning Metamorphic Image Transformation with Appearance Changes

Jian Wang[1(✉)], Jiarui Xing[2], Jason Druzgal[3], William M. Wells III[4,5], and Miaomiao Zhang[1,2]

[1] Computer Science, University of Virginia, Charlottesville, VA, USA
`jw4hv@virginia.edu`
[2] Electrical and Computer Engineering, University of Virginia, Charlottesville, VA, USA
[3] Radiology and Medical Imaging, University of Virginia, Charlottesville, VA, USA
[4] Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
[5] Computer Science and Artificial Intelligence Laboratory, MIT, Cambridge, MA, USA

Abstract. This paper presents a novel predictive model, MetaMorph, for metamorphic registration of images with appearance changes (i.e., caused by brain tumors). In contrast to previous learning-based registration methods that have little or no control over appearance-changes, our model introduces a new regularization that can effectively suppress the negative effects of appearance changing areas. In particular, we develop a piecewise regularization on the tangent space of diffeomorphic transformations (also known as initial velocity fields) via learned segmentation maps of abnormal regions. The geometric transformation and appearance changes are treated as joint tasks that are mutually beneficial. Our model MetaMorph is more robust and accurate when searching for an optimal registration solution under the guidance of segmentation, which in turn improves the segmentation performance by providing appropriately augmented training labels. We validate MetaMorph on real 3D human brain tumor magnetic resonance imaging (MRI) scans. Experimental results show that our model outperforms the state-of-the-art learning-based registration models. The proposed MetaMorph has great potential in various image-guided clinical interventions, e.g., real-time image-guided navigation systems for tumor removal surgery.

1 Introduction

Deformable image registration is an important tool in a variety of medical image analysis tasks, such as multi-modality image alignment [12,18,25], statistical analysis for population image studies [26,32,35], atlas-guided image segmentation or classification [27,30,33], and object tracking with anomaly detection [11,24]. In many clinical applications, it is desirable that the estimated transformations are diffeomorphisms (i.e., bijective, smooth, and inverse smooth mappings) because they produce anatomically plausible images [7]. Despite recent

achievements in treating the problem of diffeomorphic image registration as a fast learning task, current approaches oftentimes have an assumption that the topology of objects presented in images is intact [6,10,17,31]. Existing algorithms fail badly in cases where appearance changes occur (e.g., missing data caused by pathology, such as tumors, myocardial scars, multiple sclerosis, and etc.) because they have little to no control over these unknown variables.

To address this issue, a few algorithms of image metamorphosis have been developed to incorporate the modeling of appearance changes in registration functions [8,14,16,21,23]. Existing metamorphic image registration methods mainly fall into two categories: (i) exclude appearance changes via manually delineated segmentations of abnormal regions [21,23], and (ii) treat the appearance changes as unknown variables estimated out from images [8,14]. These approaches either heavily depend on manually segmented labels of 3D volumetric data that are time and labor-consuming, or struggle with balancing between the effects of appearance vs. geometric changes. A recent work [8] has developed a metamorphic autoencoder that estimates the deformation and appearance variations by decoupling the geometric and appearance representations in latent spaces. However, such a model is highly sensitive to parameter-tuning due to its difficulty in differentiating changes caused by geometric transformations vs. appearances.

In this paper, we develop a novel learning-based model of metamorphic image registration, named as MetaMorph, that provides more robust and accurate registration results in images with appearance changes. In contrast to previous approaches [8,14,21,23], we incorporate a new *appearance-aware regularization* in the network loss function that enforces a piecewise constraint on geometric transformation fields. Such a constraint will be learned simultaneously from a jointly optimized segmentation task. In addition, we effectively augment the segmentation labels by utilizing the learned transformations in the training process. This not only substantially improves the segmentation performance, but also reduces the requirement for massive ground truth segmentation labels. The main contributions of our proposed MetaMorph are summarized in three folds:

– To the best of our knowledge, MetaMorph is the first predictive registration algorithm that utilizes jointly learned segmentation maps to model appearance changes.
– MetaMorph learns a new appearance-aware regularization that piecewisely constrains the variations of image intensities caused by geometric transformations separately from appearance changes.
– The joint learning scheme of MetaMorph maximizes the mutual benefits of metamorphic image registration and segmentation.

To demonstrate the effectiveness of our model, we validate MetaMorph on real 3D human brain tumor MRIs. Experimental results show that MetaMorph outperforms the state-of-the-art learning-based registration models [6,8] with substantially increased accuracy. The developed MetaMorph has great potential in various image-guided clinical interventions, e.g., real-time image-guided navigation systems for tumor removal surgery.

2   Background: Diffeomorphic Image Registration

In this section, we briefly review the concept of the diffeomorphic image registration in the setting of large deformation diffeomorphic metric mapping (LDDMM) with a geodesic shooting algorithm [7, 20, 29].

Let S be the source image and T be the target image defined on a d-dimensional torus domain $\Gamma = \mathbb{R}^d/\mathbb{Z}^d$ $(S(x), T(x) : \Gamma \to \mathbb{R})$. The problem of diffeomorphic image registration is to find the geodesic (a.k.a. shortest path) to generate time-varying diffeomorphisms $\{\psi_t(x)\} : t \in [0, 1]$ such that $S \circ \psi_1$ is similar to T, where $\circ$ is an interpolation operation that deforms S by the smooth deformation field ψ_1. This is typically formulated as an optimization problem by minimizing an explicit energy function over the transformation fields ψ_t as

$$\mathrm{E}(v_t) = \mathrm{Dist}[S \circ \psi_1(v_t), T] + \mathrm{Reg}[\psi_t(v_t)], \tag{1}$$

where the distance function $\mathrm{Dist}(\cdot, \cdot)$ measures the image dissimilarity between the source and the deformed image. Commonly used distance functions include a sum-of-squared difference of image intensities [7], normalized cross correlation [4], and mutual information [34, 36]. The regularization term $\mathrm{Reg}(\cdot)$ is a constraint that enforces the spatial smoothness of transformations, arising from a distance metric on the tangent space V of diffeomorphisms, i.e., an integral over the norm of time-dependent velocity fields $\{v_t(x)\} \in V$,

$$\mathrm{Reg}(\psi_t) = \int_0^1 (Lv_t, v_t)\, dt, \quad \text{with} \quad \frac{d\psi_t}{dt} = -D\psi_t \cdot v_t, \tag{2}$$

where $L : V \to V^*$ is a symmetric, positive-definite differential operator that maps a tangent vector $v_t \in V$ into its dual space as a momentum vector $m_t \in V^*$. We typically write $m_t = Lv_t$, or $v_t = Km_t$, with K being an inverse operator of L. The notation $(\cdot, \cdot)$ denotes the pairing of a momentum vector with a tangent vector, which is similar to an inner product. Here, the operator D denotes a Jacobian matrix and $\cdot$ represents element-wise matrix multiplication.

A geodesic curve with a fixed endpoint is characterized by an extremum of the energy function (2) that satisfies the Euler-Poincaré differential (EPDiff) equation [2, 20],

$$\frac{\partial v_t}{\partial t} = -K\left[(Dv_t)^T \cdot m_t + Dm_t \cdot v_t + m_t \cdot \mathrm{div}\, v_t\right], \tag{3}$$

where div is the divergence. This process in Eq. (3) is known as *geodesic shooting*, stating that the geodesic path $\{\psi_t\}$ can be uniquely determined by integrating a given initial velocity v_0 forward in time by using the rule (3).

Therefore, we rewrite the optimization of Eq. (1) equivalently as

$$\mathrm{E}(v_0) = \mathrm{Dist}[S \circ \psi_1(v_0), T] + (Lv_0, v_0), \text{ s.t. Eq. (2)\&Eq. (3).} \tag{4}$$

3   Our Model: MetaMorph

The objective function of diffeomorphic image registration in Eq. (4) works well under the condition that images are ideally of good quality with preserved topology. This assumption breaks when corruptions such as appearance changes or occlusions occur. In this section, we first define an objective function of the metamorphic image registration that considers the modeling of appearance changes. An appearance-aware regularization is developed to effectively suppress the negative influences of appearance changes in typical diffeomorphic image registration algorithms. We then develop a joint learning framework that includes i) a segmentation network for appearance change detection, and ii) a metamorphic registration network incorporating the newly formulated objective function as part of the network loss.

Appearance-Aware Regularization. The purpose of metamorphic image registration is to find an optimal transformation $\psi(v_0, \delta)$ that is composed of two variables: the optimal initial velocity field v_0, and the appearance change δ. A recent work proposed to learn these variables via disentangled latent representations in an encoder-decoder neural network [8]. However, it is extremely challenging for this algorithm to differentiate the variations of image intensities caused by geometric transformations from appearance changes since they unavoidably compensate for each other. The ambiguity introduced by optimizing two compensating variables without any guidance fails to search for accurate registration solutions. Additionally, this makes the algorithm highly sensitive to network parameters with an increased risk of poor convergence. To alleviate this issue, we introduce an appearance-aware regularization in the registration framework, guided by learned segmentations of the appearance-changing areas.

Assume U is a union of the learned segmentations of appearance-changing areas from the source image S and the target image T. Analogous to Eq. (4), we define the appearance-aware regularization $\mathbf{Reg}^*(\cdot)$ in the space of initial velocity fields. To suppress the effects of appearance variations, we piecewise constrain the initial velocity fields through a segmentation indicator, i.e.,

$$\mathbf{Reg}^*(v_0) = \left(L(v_0 \odot (1 - U)), v_0 \odot (1 - U)\right), \text{ s.t. Eq. (3)}, \qquad (5)$$

where $\odot$ represents an element-wise multiplication between a vector field and a scalar field. For the purpose of notation simplicity, we define $\hat{v}_0 \overset{\Delta}{=} v_0 \odot (1 - U)$ in the following sections.

With the newly defined regularization in Eq. (5), we arrive at the objective function of metamorphic image registration as

$$\mathbf{E}^*[\hat{\psi}(\hat{v}_0)] = \mathbf{Dist}^*[\hat{S} \circ \hat{\psi}_1(\hat{v}_0), \hat{T}] + \mathbf{Reg}^*(\hat{v}_0), \qquad (6)$$

where $\hat{S}$ and $\hat{T}$ denotes the source and target images with appearance changes masked out, i.e., $\hat{S} = S \odot (1 - U)$, and $\hat{T} = T \odot (1 - U)$. Here, the $\mathbf{Dist}^*[\cdot, \cdot]$ is the image dissimilarity term that measures the dissimilarity between the consistent area between the deformed image and target.

3.1   Predictive Metamorphic Image Registration

We develop a deep learning framework to jointly learn the segmentation for appearance change and the masked-out velocity field $\hat{v}_0$. An overview of our proposed MetaMorph architecture is shown in Fig. 1.

Appearance change can be masked by a fixed foreground segmentation via pre-running image segmentation algorithms [21,23]. However, performing manual annotations of segmentation labels is time and labor-consuming. In this work, instead of using a fixed mask, we treat the appearance change as a variable from the segmentation network and jointly optimize with the optimal registration solution. We utilize an encoder-decoder based neural network to learn the segmentation masks and then apply them to the associate image pairs for masking out the appearance change. Although we adopt UNet-based architecture for segmentation in this work [28], other networks such as recurrent residual neural networks [1], transformer-based networks [9,15] can also be easily plugged into the proposed method.

With the developed segmentation network, now we are ready to formulate the loss function of MetaMorph,

$$\ell = \mathbf{Dist}^*[\hat{S} \circ \hat{\psi}_1(\hat{v}_0), \hat{T}] + \mathbf{Reg}^*(\hat{v}_0) + \gamma \cdot \ell_{seg}, \;\; \text{s.t. Eq. (5)}. \qquad (7)$$

Here, γ is a weighting parameter that balances the segmentation and registration loss, ℓ_{seg} is a segmentation loss that maximizes the Sørensen-Dice coefficient [13] between ground truth y and the predicted $\hat{y}$,

$$\ell_{seg} = 1 - \text{Dice}(y, \hat{y}), \qquad (8)$$

where $\text{Dice}(y, \hat{y}) = 2(|y| \cap |\hat{y}|)/(|y| + |\hat{y}|)$.

We adopt an approximated region-based mutual information (RMI) [36], which is a broadly-used distance metric for images from different domains. For simplicity, we let $\hat{S}_\psi$ denote the deformed image. Let $f(\hat{S}_\psi)$ and $f(\hat{T})$ denote the probability density functions for the deformed image and target respectively, and their joint probability density function is $f(\hat{S}_\psi, \hat{T})$. The image dissimilarity with RMI can be formulated as

$$\mathbf{Dist}^*[\hat{S}_\psi, \hat{T}] = \text{RMI}(\hat{S}_\psi, \hat{T}) = \int_{\hat{S}_\psi} \int_{\hat{T}} f(\hat{S}_\psi, \hat{T}) \log \frac{f(\hat{S}_\psi, \hat{T})}{f(\hat{S}_\psi) f(\hat{T})}$$

$$\approx l_{ce}(\hat{S}_\psi, \hat{T}) - \frac{1}{B} \sum_{b=1}^{B} I_b(\hat{T}; \hat{S}_\psi), \qquad (9)$$

where $L_{ce}(\cdot, \cdot)$ is a cross entropy loss between two images. The $I_b(\cdot; \cdot)$ is a batch-wise lower bound that $I_b(\hat{T}; \hat{S}_\psi) = \frac{1}{2} \log[\det(\Sigma_{\hat{T}|\hat{S}_\psi})]$, where $\Sigma_{\hat{T}|\hat{S}_\psi}$ is the posterior covariance matrix of $\hat{T}$ (a symmetric positive semi-definite matrix), given $\hat{S}_\psi$. Here B denotes the number of images in a mini-batch b. Please refer to [36] for more derivation details.
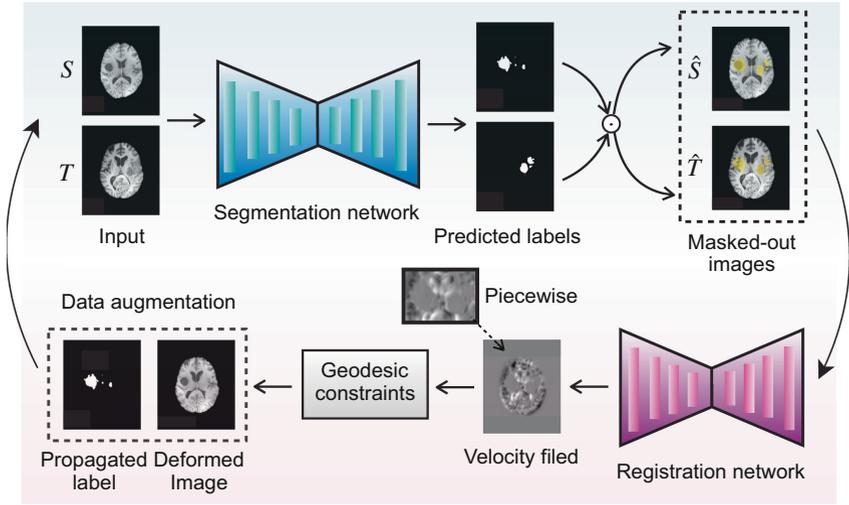
Fig. 1. An illustration of the network architecture for MetaMorph. Top left to right: input a pair of images into a segmentation network, and apply predicted labels onto images to mask out the appearance change. Bottom right to left: input a pair of images (with masked-out appearance change) to the registration network and predict a piecewise velocity field, integrate geodesic constraints, and produce a deformed image and transformation-propagated segmentation. The deformed images and labels are circulated into the segmentation network as augmented data.

We develop an alternating optimization scheme [22] to minimize the network loss defined in Eq. (7). All network parameters are optimized jointly by alternating between the training of segmentation and image registration. A summary of our joint learning of MetaMorph is in Algorithm 1.

4   Experimental Evaluation

To demonstrate the effectiveness of the proposed model, we compare both segmentation and registration tasks with state-of-the-arts.

Data. For 3D brain tumor MRI scans with tumor segmentation labels, we include 100 public T1-weighted brain scans of different subjects from Brain Tumor Segmentation (BraTS) [5,19] challenge 2021. We also include 28 landmarks (16 for brain ventricle and 12 for corpus callosum) that are annotated by clinicians to better evaluate the image registration performance. All MRIs are $155 \times 240 \times 240$, $1.25\,\mathrm{mm}^3$ isotropic voxels. As a preprocessing step, we run affine registration, intensity normalization, and bias field correction on all images.

Experiments. We compare our metamorphic image registration method with two registration baselines, an unsupervised predictive diffeomorphic registration method (VoxelMorph as VM) [6], and a metamorphic autoencoder (MAE) [8]

Algorithm 1: Joint learning of MetaMorph.

Input : Source and target images, the number of iterations q.
Output: Segmentation labels, the deformed image, and the transformation.

1 for $i = 1$ to q **do**
    /* Train image segmentation network   */
2 | Minimize the segmentation loss in Eq. (8);
3 | Output the predicted segmentations and adopt both labels to mask
       | appearance change in images;
       | /* Train appearance-aware registration network   */
4 | Minimizing the metamorphic loss in Eq. (6) with appearance-aware geodesic
       | constraints;
5 | Output the predicted velocity field and the deformed image;
6 end
7 Until convergence

---

that learns disentangled appearance and shape representations. To better visualize the deformations, we show predicted transformation grids and deformed images with transformation-propagated landmarks for all methods. Quantitatively, we compute the L_2 distance of landmarks as registration error between the propagated and the target frames over 60 pairs.

We evaluate the brain tumor segmentation via computing Dice score [13] by comparing MetaMorph with three segmentation backbones, U-Net architecture [28], U-Net based on recurrent residual convolutional neural network (R2-Unet) [1], and transformer-based Unet (UnetR) [15]. We also show the performance of MetaMorph by replacing the segmentation module in our model with all backbones (named MetaMorph:Unet, MetaMorph:R2-Unet, and Meta-Morph:UnetR). We visualize the predicted segmentations overlaid with testing images across all methods.

Parameter Settings. We set parameter $\alpha = 3$ for the operator L, the number of time steps for Euler integration in EPDiff (Eq. (3)) as 10. We set the weight parameter $\gamma = 0.5$ and the batch size as 4. We use an adaptive cosine annealing learning rate scheduler that starts from an initial value at $\eta = 5e - 4$ for network training. We run all models for 100 epochs with Adam optimizer and save the networks with the best validation performance. The training and prediction procedure of all learning-based methods are performed on two Nvidia GTX 2070Ti GPUs. We run five-fold cross validation and split the images by using 70% as training images, 20% as validation images, and 10% as testing images.

Results. Figure 2 visualizes the image registration prediction of two 3D brain MRIs of study across all methods. It shows MetaMorph significantly outperforms both VM and MAE. General diffeomorphic registration models (e.g., VM) without an appearance-control mechanism may fail and produce less satisfied deformed images without sufficient deformations. MAE offers accurate deformations to a certain level while it produces artifacts. By excluding the appearance change, MetaMorph more accurately deforms all regions (e.g., ventricles and
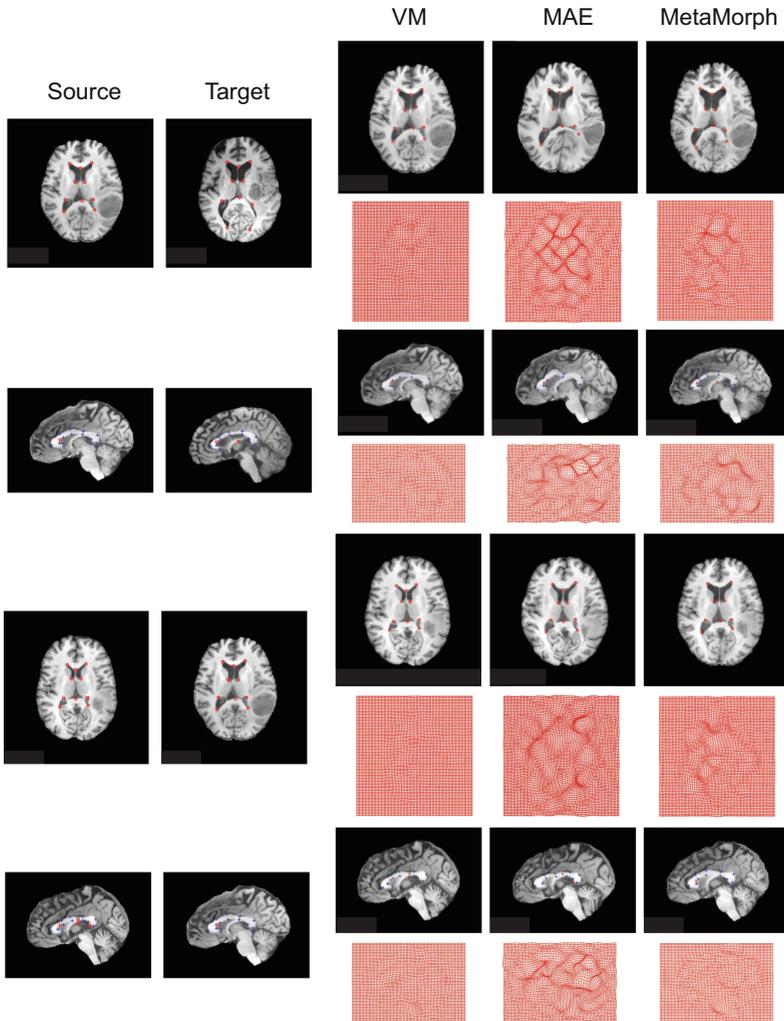
Fig. 2. Image registration performance comparison for all methods. From left to right, source, target, deformed images by VoxelMorph (VM), metamorphic autoencoder (MAE), and our method. All images are overlaid with annotated landmarks (red circle for ventricle and blue cross for corpus callosum). (Color figure online)

corpus callosum). It also shows that our propagated landmarks align best with the target.

Figure 3 shows two examples of image segmentation performance comparison for all methods. It indicates that MetaMorph-based models predict better segmentation labels (closer to ground truth) than original backbones. The predicted labels by MetaMorph have slightly better segmentations of the brain tumor boundary. This is because we use deformed images and labels that are
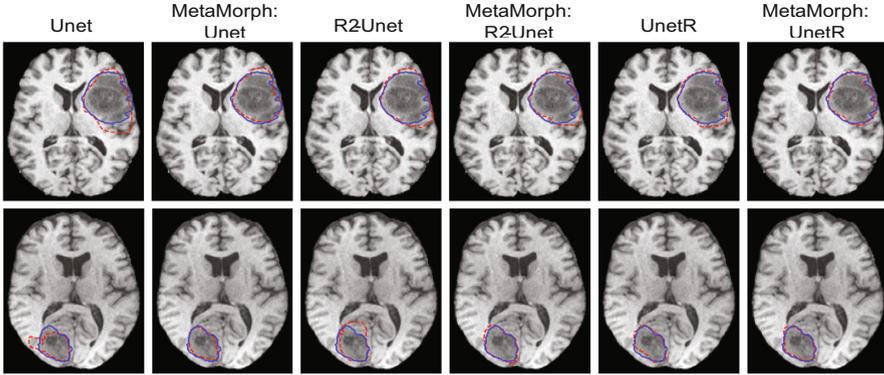
Fig. 3. Image segmentation visualization for all methods. Left to right: overlaid segmentation map comparison between the predicted label (red) and the ground truth (blue) for Unet, MetaMorph: Unet, R2-Unet, MetaMorph: R2-Unet, UnetR and MetaMorph: UnetR. (Color figure online)
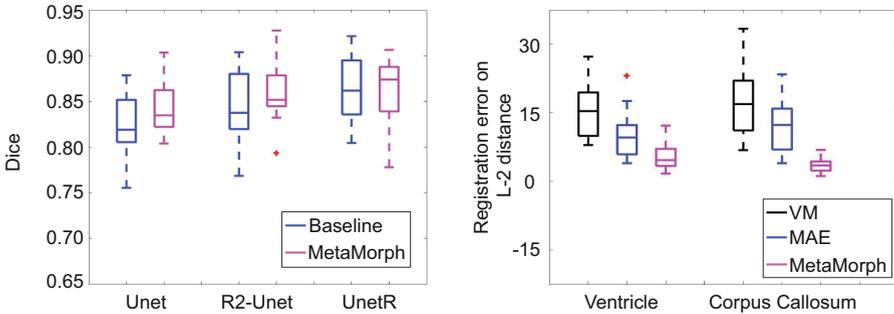

Fig. 4. Left: Dice comparison on brain tumor segmentation across all methods over images. The means of baseline vs. **our method** are 0.815/**0.834**, 0.835/**0.856**, 0.861/**0.874**; Right: registration error (computed on L_2 distance) of two anatomical landmarks for 60 brain pairs. The means of errors for VM vs. MAE vs. **our method** are 15.02/10.53/**4.64**, 16.48/13.59/**4.10**.

produced by a joint registration framework as augmented data for each subject; thus learning a broader spectrum for appearance variation in data and offering more accurate prediction when new testing data arrives.

Figure 4 (left panel) statistical reports the Dice coefficient comparison. It shows that MetaMorph consistently achieves a higher segmentation accuracy than backbones. Transformer-based methods (UnetR-based) produce the highest Dice for all methods. Figure 4 (right panel) reports the landmark-based registration error between the target image and the deformed image. MetaMorph outperforms other methods with the lowest error, indicating our proposed method finishes the metamorphic image registration task with higher accuracy.

5   Conclusion

We present a predictive metamorphic image registration model, MetaMorph, via deep neural networks in this paper. Different from existing models that have limited control over appearance change, we develop a joint learning framework that adopts a segmentation module to accurately guide the registration network to learn diffeomorphic transformation fields. The developed segmentation module maximally excludes the disadvantageous effect caused by appearance change for learned deformations; thus enabling more precise correspondence alignment between deformed and target frames. Experimental results on 3D brain MRIs with real tumors show that our proposed framework yields a better registration as well as a segmentation model. While our algorithm is presented in the setting of LDDMM with geodesic shooting, the theoretical development is generic to other deformation models, e.g., stationary velocity fields [3]. Our model has great clinical potential on solving one of the most challenging registration problems, e.g., real-time brain shift estimation between preoperative and intraoperative MRI scans with missing data values. Interesting future works of MetaMorph will be i) building a probabilistic model to quantify the registration uncertainty along the boundary of tumor areas and ii) extending the proposed method to more advanced clinical scenarios that appearance changes are difficult to detect, e.g., real-time automated image registration for ultrasound images.

References

1. Alom, M.Z., Hasan, M., Yakopcic, C., Taha, T.M., Asari, V.K.: Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation. arXiv preprint arXiv:1802.06955 (2018)
2. Arnold, V.: Sur la géométrie différentielle des groupes de lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. In: Annales de l'institut Fourier, vol. 16, pp. 319–361 (1966)
3. Arsigny, V., Commowick, O., Pennec, X., Ayache, N.: A log-Euclidean framework for statistics on diffeomorphisms. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4190, pp. 924–931. Springer, Heidelberg (2006). https://doi.org/10.1007/11866565_113
4. Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C.: Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med. Image Anal. **12**(1), 26–41 (2008)
5. Baid, U., et al.: The RSNA-ASNR-MICCAI brats 2021 benchmark on brain tumor segmentation and radiogenomic classification. arXiv preprint arXiv:2107.02314 (2021)
6. Balakrishnan, G., Zhao, A., Sabuncu, M.R., Guttag, J., Dalca, A.V.: VoxelMorph: a learning framework for deformable medical image registration. IEEE Trans. Med. Imaging **38**(8), 1788–1800 (2019)
7. Beg, M.F., Miller, M.I., Trouvé, A., Younes, L.: Computing large deformation metric mappings via geodesic flows of diffeomorphisms. Int. J. Comput. Vis. **61**(2), 139–157 (2005)

8. Bône, A., Vernhet, P., Colliot, O., Durrleman, S.: Learning joint shape and appearance representations with metamorphic auto-encoders. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12261, pp. 202–211. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59710-8_20

9. Chen, J., et al.: TransUNet: transformers make strong encoders for medical image segmentation. arXiv preprint arXiv:2102.04306 (2021)

10. Chen, J., Frey, E.C., He, Y., Segars, W.P., Li, Y., Du, Y.: TransMorph: transformer for unsupervised medical image registration. Med. Image Anal. **82**, 102615 (2022)

11. Chen, M., Kanade, T., Pomerleau, D., Rowley, H.A.: Anomaly detection through registration. Pattern Recogn. **32**(1), 113–128 (1999)

12. Chung, A.C.S., Wells, W.M., Norbash, A., Grimson, W.E.L.: Multi-modal image registration by minimising Kullback-Leibler distance. In: Dohi, T., Kikinis, R. (eds.) MICCAI 2002. LNCS, vol. 2489, pp. 525–532. Springer, Heidelberg (2002). https://doi.org/10.1007/3-540-45787-9_66

13. Dice, L.R.: Measures of the amount of ecologic association between species. Ecology **26**(3), 297–302 (1945)

14. François, A., Gori, P., Glaunès, J.: Metamorphic image registration using a semi-Lagrangian scheme. In: Nielsen, F., Barbaresco, F. (eds.) GSI 2021. LNCS, vol. 12829, pp. 781–788. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-80209-7_84

15. Hatamizadeh, A., et al.: UNETR: transformers for 3D medical image segmentation. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 574–584 (2022)

16. Holm, D., Trouvé, A., Younes, L.: The Euler-Poincaré theory of metamorphosis. Q. Appl. Math. **67**(4), 661–685 (2009)

17. Kim, B., Han, I., Ye, J.C.: DiffuseMorph: unsupervised deformable image registration along continuous trajectory using diffusion models. arXiv preprint arXiv:2112.05149 (2021)

18. Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P.: Multimodality image registration by maximization of mutual information. IEEE Trans. Med. Imaging **16**(2), 187–198 (1997)

19. Menze, B.H., et al.: The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans. Med. Imaging **34**(10), 1993–2024 (2014)

20. Miller, M.I., Trouvé, A., Younes, L.: Geodesic shooting for computational anatomy. J. Math. Imaging Vis. **24**(2), 209–228 (2006)

21. Niethammer, M., et al.: Geometric metamorphosis. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011. LNCS, vol. 6892, pp. 639–646. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-23629-7_78

22. Nocedal, J., Wright, S.J.: Numerical Optimization. Springer, Heidelberg (1999)

23. Patenaude, B., Smith, S.M., Kennedy, D.N., Jenkinson, M.: A Bayesian model of shape and appearance for subcortical brain segmentation. Neuroimage **56**(3), 907–922 (2011)

24. Prastawa, M., Bullitt, E., Ho, S., Gerig, G.: A brain tumor segmentation framework based on outlier detection. Med. Image Anal. **8**(3), 275–283 (2004)

25. Qin, C., Shi, B., Liao, R., Mansi, T., Rueckert, D., Kamen, A.: Unsupervised deformable registration for multi-modal images via disentangled representations. In: Chung, A.C.S., Gee, J.C., Yushkevich, P.A., Bao, S. (eds.) IPMI 2019. LNCS, vol. 11492, pp. 249–261. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-20351-1_19

26. Rao, A., Aljabar, P., Rueckert, D.: Hierarchical statistical shape analysis and prediction of sub-cortical brain structures. Med. Image Anal. **12**(1), 55–68 (2008)

27. Riklin-Raviv, T., Van Leemput, K., Menze, B.H., Wells, W.M., III., Golland, P.: Segmentation of image ensembles via latent atlases. Med. Image Anal. **14**(5), 654–665 (2010)

28. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28

29. Vialard, F.X., Risser, L., Rueckert, D., Cotter, C.J.: Diffeomorphic 3D image registration via geodesic shooting using an efficient adjoint calculation. Int. J. Comput. Vis. **97**(2), 229–241 (2012)

30. Wachinger, C., Golland, P.: Atlas-based under-segmentation. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) MICCAI 2014. LNCS, vol. 8673, pp. 315–322. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10404-1_40

31. Wang, J., Zhang, M.: DeepFLASH: an efficient network for learning-based medical image registration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4444–4452 (2020)

32. Wang, J., Zhang, M.: Bayesian atlas building with hierarchical priors for subject-specific regularization. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12904, pp. 76–86. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87202-1_8

33. Wang, J., Zhang, M.: Geo-SIC: learning deformable geometric shapes in deep image classifiers. In: The Conference on Neural Information Processing Systems (2022)

34. Wells, W.M., III., Viola, P., Atsumi, H., Nakajima, S., Kikinis, R.: Multi-modal volume registration by maximization of mutual information. Med. Image Anal. **1**(1), 35–51 (1996)

35. Zhang, M., Wells, W.M., Golland, P.: Low-dimensional statistics of anatomical variability via compact representation of image deformations. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9902, pp. 166–173. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46726-9_20

36. Zhao, S., Wang, Y., Yang, Z., Cai, D.: Region mutual information loss for semantic segmentation. In: Advances in Neural Information Processing Systems, vol. 32 (2019)