Anatomy-specific acquisition-agnostic affine registration learned from fictitious images

Malte Hoffmann, Andrew Hoopes, Bruce Fischl, Adrian Dalca

Malte Hoffmann, Andrew Hoopes, Bruce Fischl, Adrian V. Dalca, "Anatomy-specific acquisition-agnostic affine registration learned from fictitious images," Proc. SPIE 12464, Medical Imaging 2023: Image Processing, 1246402 (3 April 2023); doi: 10.1117/12.2653251

SPIE.

Event: SPIE Medical Imaging, 2023, San Diego, California, United States

Anatomy-specific acquisition-agnostic affine registration learned from fictitious images

Malte Hoffmann[a–d], Andrew Hoopes[a,b,e], Bruce Fischl[a–e,*], and Adrian V. Dalca[a–c,e,*]

[a]Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, MA, USA
[b]Department of Radiology, Massachusetts General Hospital, Boston, MA, USA
[c]Department of Radiology, Harvard Medical School, Boston, MA, USA
[d]Division of Health Sciences and Technology, MIT, Cambridge, MA, USA
[e]Computer Science & Artificial Intelligence Laboratory, MIT, Cambridge, MA, USA

ABSTRACT

Affine image registration is a cornerstone of medical image processing backed by decades of development. While classical algorithms can achieve excellent accuracy, they solve a time-consuming optimization for every new image pair. In contrast, deep-learning (DL) methods learn a function that maps an image pair to an output transform. Evaluating the function is fast, but these methods tend to be susceptible to domain shift. A network trained on a specific image type may perform poorly if an image characteristic changes at test time, such as the imaging contrast or resolution. Secondly, many classical and DL registration algorithms cannot distinguish between relevant and irrelevant anatomy: the global nature of the linear registration problem means that accuracy will suffer if parts of the image deform independently. This is why neuroimage processing, for example, often starts with brain extraction, to enhance the accuracy of brain-specific registration. We address these shortcomings of linear registration by training deep neural networks using a generative strategy that synthesizes wildly varying images from label maps. Optimizing label overlap decouples the loss from the image appearance, encouraging network invariance to acquisition specifics. It also enables the registration model to distinguish between anatomy of interest and irrelevant structures, which alleviates the need for segmenting images prior to registration to remove distracting content. We test brain-specific registration across a variety of magnetic resonance imaging protocols that approximate the diversity of real-world data, demonstrating consistent and improved accuracy relative to state-of-the-art baselines. We freely distribute our easy-to-use tool at https://w3id.org/synthmorph.

Keywords: affine image registration, deep learning, domain shift, contrast invariance, anatomical specificity

1. DESCRIPTION OF PURPOSE

Affine image registration is a cornerstone of medical image processing and analysis that estimates a linear mapping from the anatomy in one image to the space of another.[1–4] While classical registration has been extensively studied and can achieve excellent accuracy, these algorithms solve a time-consuming optimization problem for every new image pair.[5–8] In contrast, deep learning (DL) methods learn a function that maps an input image pair to an output transform. Evaluating this function is fast, but DL algorithms are typically limited to registering images similar to the type of data seen at training.[9–13] A network trained with magnetic resonance imaging (MRI) data of T1-weighted (T1w) contrast, for example, will not accurately register T2-weighted (T2w) images to proton-density-weighted (PDw) scans. Although one could include these data types in the training set, the network would still not generalize to new contrast pairings unseen at training. Worse, even at similar contrast, the domain shift caused by unseen noise or smoothness levels alone can reduce accuracy at test time.

Accurate alignment of specific anatomy of interest requires ignoring or down-weighting irrelevant image content. For example, the presence of neck and tongue tissue in MRI scans can reduce the accuracy of brain registration, because these structures can move independently of the brain and even deform non-linearly.[14–16] However, many existing classical and DL algorithms share an inability to focus the registration on specific anatomical features,[17] and therefore expect—and require—that irrelevant image content be removed from the input data for

*These authors contributed equally. Send correspondence to Malte Hoffmann, mhoffmann@mgh.harvard.edu.

Medical Imaging 2023: Image Processing, edited by Olivier Colliot, Ivana Išgum,
Proc. of SPIE Vol. 12464, 1246402 · © 2023 SPIE
1605-7422 · doi: 10.1117/12.2653251

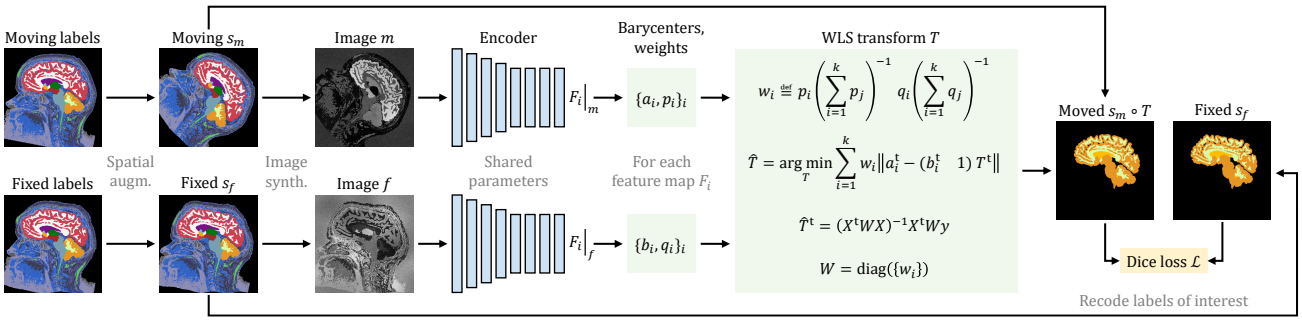

Figure 1. Training strategy. At each iteration, we augment a pair of moving and fixed label maps $\{s_m, s_f\}$ and synthesize images $\{m, f\}$ from them. A 3D convolutional encoder predicts $k = 64$ ReLU-activated feature maps F_i $(i \in \{1, 2, ..., k\})$ for m, and separately for f. We compute their barycenters a_i and b_i, and total masses p_i and q_i, to fit the affine transform $\hat{T}$ that aligns these point sets in a weighted least-squares (WLS) sense, where X and y are the matrices whose i^{th} rows are $(a_i^{\text{t}} \quad 1)$ and b_i^{t}, respectively. The Dice loss $\mathcal{L}$ recodes the labels in $\{s_m, s_f\}$ to optimize the overlap of WM, GM, and CSF. All convolutions except the last use 256 filters, LeakyReLU activation ($\alpha = 0.2$), and kernel size $3 \times 3 \times 3$. Blue blocks smaller than their predecessor indicate subsampling by a factor of 2 via max pooling.

optimal performance.[1,2,4,11,18,19] An exception are classical methods down-weighting the contribution of image regions that do not transform linearly, by using iteratively re-weighted least squares (LS).[4,6,20–23]

The purpose of our work is to address the data-type dependency of affine registration, to alleviate the need for retraining affine networks for new image types and to remove the need for segmentation as a pre-processing step to achieve anatomy-specific registration, such as skull-stripping[24–27] in neuroimaging applications.

2. NEW OR BREAKTHROUGH WORK TO BE PRESENTED

Leveraging a recent strategy[27–29] that trains networks with wildly variable synthetic data only, we present an approach to producing powerful models for affine registration of real medical images. The synthesis of training data from label maps enables use of a Dice-based loss decoupled from the image appearance, promoting network invariance to acquisition specifics, for example MRI contrast and resolution. Importantly, optimizing the spatial overlap of select anatomical labels makes the proposed network aware of the anatomy of interest, such that it will focus on aligning brain tissue, keeping any non-brain signal from reducing brain-specific registration accuracy.

We freely distribute our easy-to-use registration tool and code at https://w3id.org/synthmorph and as the command-line utility `mri_synthmorph` within the upcoming FreeSurfer[30] 7.3.4 release. The research presented is original and has not been submitted elsewhere.

3. METHOD

Let m and f be a moving and a fixed gray-scale image defined in N-dimensional (ND) space, respectively. We train a deep neural registration network h_θ with learnable parameters θ to predict an affine transform $T_\theta = h_\theta(m, f)$ aligning $\{m, f\}$ as shown in Figure 1.

We extend a recent strategy[28,29] that synthesizes intentionally unrealistic images $\{m, f\}$ from corresponding label maps $\{s_m, s_f\}$. Let K be the complete set of labels in $\{s_m, s_f\}$. While we synthesize $\{m, f\}$ using all labels K, we encourage the network to register specific anatomy while ignoring irrelevant image content: we recode $\{s_m, s_f\}$ such that they include only select anatomical labels $J \subset K$, and optimize their overlap using a (soft) Dice[31] loss $\mathcal{L}$, independent of image appearance:

$$\mathcal{L}(T_\theta, s_m, s_f) = -\frac{2}{|J|} \sum_{j \in J} \sum_{x \in \Omega} \frac{(s_m\big|_j \circ T_\theta)(x) \times s_f\big|_j(x)}{(s_m\big|_j \circ T_\theta)(x) + s_f\big|_j(x)}, \tag{1}$$

where $s\big|_j$ is the one-hot encoded label $j \in J$ of label map s defined at voxel locations $x \in \Omega$, and the transform $T_\theta : \Omega \to \mathbb{R}^N$ maps the discrete spatial domain Ω of image f onto m.

3.1 Training-Data Synthesis

At each training iteration, we spatially augment a randomly drawn segmentation map s_m by applying a non-linear deformation including translation, rotation, scaling, and shear. The remainder of the image synthesis builds heavily on our recent work on *deformable* registration:[29] from s_m, we generate gray-scale image m by drawing a separate intensity for each label $j \in \{1, 2, ..., K\}$ of s_m, and we randomly corrupt the image with noise, blurring, cropping, downsampling, a spatially varying intensity bias field, and gamma exponentiation. Similarly, we synthesize image f from label map s_f.

We sample the parameters governing the affine augmentation from noise distributions, within the ranges defined in Table 1. We adapt all other synthesis hyperparameters from prior work,[27,29] which thoroughly analyzed their impact on registration accuracy.[29] Figure 2 shows examples of typical synthetic training images.

3.2 Anatomy-Specific Registration

Our focus is on exploiting the synthesis strategy for anatomy-specific and acquisition-agnostic affine registration. While other networks may be equally suitable, we chose an architecture[19,32] that derives k learned features $F_i\big|_m$ ($i \in \{1, 2, ..., k\}$) from image m, and separately for image f, to fit a transform that aligns these features in a LS sense.[33–37] We extend prior work to predict a full affine transform[19] *and* perform a weighted LS fit[32] for robustness. Briefly, we compute barycenter a_i and channel mass p_i for each predicted feature map $F_i\big|_m$ derived from m,

$$a_i = p_i^{-1} \sum_{x \in \Omega} x F_i\big|_m(x) \quad \text{and} \quad p_i = \sum_{x \in \Omega} F_i\big|_m(x), \tag{2}$$

and similarly barycenter b_i and mass q_i for each $F_i\big|_f$ of image f. Introducing the normalized weights

$$w_i = p_i \Big(\sum_{j=1}^{k} p_j \Big)^{-1} q_i \Big(\sum_{j=1}^{k} q_j \Big)^{-1}, \tag{3}$$

we interpret the sets $\{a_i\}$ and $\{b_i\}$ as corresponding moving and fixed point clouds and compute the transform $T_\theta \in \mathbb{R}^{N \times (N+1)}$ that aligns $\{a_i\}$ and $\{b_i\}$ in a LS sense, subject to

$$\hat{T}_\theta = \arg\min_T \sum_{i=1}^{k} w_i \left\| a_i^{\mathrm{t}} - \begin{pmatrix} b_i^{\mathrm{t}} & 1 \end{pmatrix} T^{\mathrm{t}} \right\|^2, \tag{4}$$

where T^{t} is the matrix transpose of T. Denoting $W = \mathrm{diag}(\{w_i\})$, and by X and y the matrices whose i^{th} rows are $\begin{pmatrix} a_i^{\mathrm{t}} & 1 \end{pmatrix}$ and b_i^{t}, respectively, the closed-form solution $\hat{T}_\theta$ of Equation (4) is

$$\hat{T}_\theta^{\mathrm{t}} = (X^{\mathrm{t}} W X)^{-1} X^{\mathrm{t}} W y. \tag{5}$$

3.3 Implementation Details

The network h_θ implements a fully convolutional feature detector that operates on a single image as shown in Figure 1. This detector consists of eight 3D convolutional blocks with $w = 256$ filters each, whose outputs we activate with LeakyReLU (parameter $\alpha = 0.2$). A ninth convolutional block outputs $k = 64$ ReLU-activated feature maps F_i ($i \in \{1, 2, ..., k\}$). We downsample the output of each of the first four blocks by a factor of 2 using max-pooling. For computational efficiency, we also downsample the network inputs $\{m, f\}$ by a factor of 2, all while evaluating the loss on full-size segmentation maps $\{s_m, s_f\}$. All kernels are of size $3 \times 3 \times 3$. We min-max normalize the input images such that their intensities fall in the interval $[0, 1]$. We fit model parameters with stochastic gradient descent using a learning rate of $l = 10^{-5}$ and batch size 1 until the loss visually converges.

4. EXPERIMENT

While we train our network with intentionally unrealistic images generated from label maps, all tests use real acquired 3D brain MRI scans. We compare classical and readily available deep-learning baselines trained by their respective authors, to assess their generalization capabilities and to evaluate what level of accuracy the interested reader can expect from off-the-shelf methods without retraining. Our network trained following the synthesis strategy surpasses baseline accuracy even though it does not sample any real data at training.

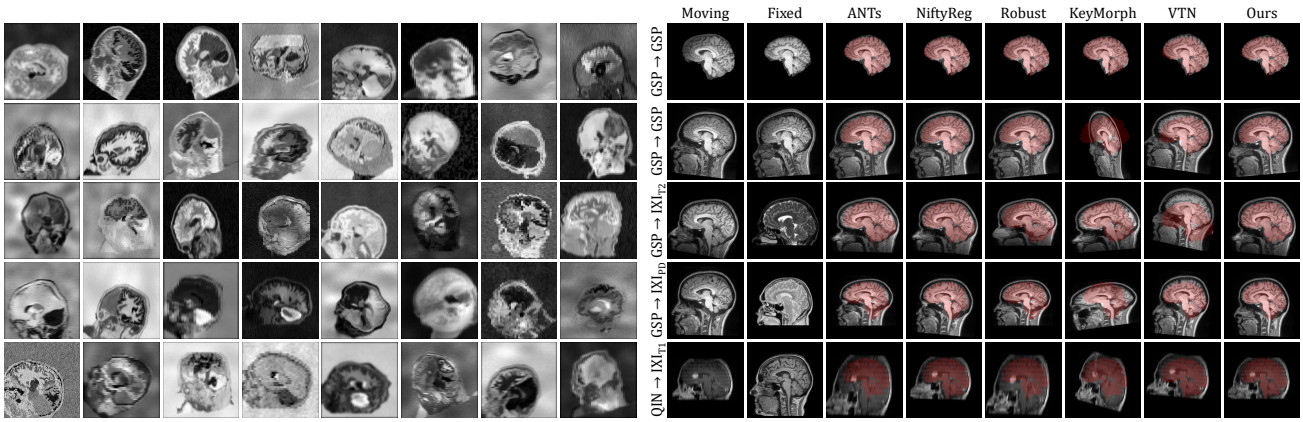

Figure 2. Left: synthetic training data generated from brain label maps. The images exceed the realistic range to promote network generalization across acquisition protocols. All examples are based on the same label map. In practice, we use data from several subjects. Right: representative 3D affine registration pairs showing the image moved by each method overlaid with the fixed brain mask. Each row is an example from a different dataset. Subscripts indicate MRI contrast.

4.1 Data

We create a training set of 100 anatomical label maps derived from T1w FSM[38] and OASIS[39] data. The evaluation set includes T1w, T2w, and PDw images from the GSP[40] and IXI[41] datasets. As all these images have near-isotropic ∼1-mm voxels, we also add contrast-enhanced clinical stacks of T1w 6-mm slices from QIN[42–44] subjects with glioblastoma. While users of our registration tool do not need to preprocess their data, we resample and symmetrically crop and zero-pad all images to obtain volumes of $256 \times 256 \times 256$ isotropic 1-mm voxels, rearranged to produce left-inferior-anterior orientation of the brain relative to the volume axes. We skull-strip a subset of these data with SynthStrip[27] to compare registration performance after common preprocessing steps.

We derive brain labels for training and evaluation from each individual image using SynthSeg.[28] To synthesize diverse training-image content outside the labeled anatomy, we add non-brain labels using a simple thresholding procedure.[27] We sort non-zero image voxels outside the brain into one of six intensity bins, equalizing bin sizes on a per-image basis. These six added labels do not necessarily represent meaningful anatomical structures but expose the network to non-brain image content.

4.2 Setup

As illustrated in Figure 1, each training iteration generates input images from all $K = 38$ brain and non-brain labels, including background, whereas training optimizes the overlap of three merged brain-tissue classes $J \subset K$ only: white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF). These large labels ensure that small structures like the caudate do not have a disproportionate influence on brain alignment. We test cross-subject registration on 30 held-out image pairs for each of several data-type pairings and compare our strategy to popular baselines. For each image pair, we measure brain registration accuracy by computing the mean Dice overlap D between the moved label map $s_m \circ T_\theta$ and the fixed label map s_f across classes J.

4.3 Baselines

We test classical affine registration with ANTs,[7] using recommended parameters[45] for the NCC metric within and MI across MRI contrasts, and NiftyReg[4] with the NMI metric. As an example of a method down-weighting image regions which do not transform linearly, we run Robust Registration[6] with its robust cost functions but highlight that the cross-modal robust-entropy metric is deemed experimental. For a thorough comparison, we initialize affine Robust Registration with a rigid run and use up to 100 iterations.

As DL baselines, we test original models pre-trained by the respective authors using skull-stripped T1-w MRI. This analysis assesses the accuracy achievable without retraining. For reasons including the need for recreating complex development environments and hardware availability, retraining is generally challenging for

users, limiting the reusability of DL methods. We test KeyMorph,[19] which uses an architecture similar to ours, and the affine cascade of the pair-wise 10-cascade Volume Tweening Network[11] (VTN). We choose the most recent version of each package, ensuring that network inputs have the expected orientation and intensity normalization.

4.4 Results

Figure 3 compares registration accuracy measured as mean Dice scores over tissue classes J, and Figure 2 shows representative examples for each method and dataset. Despite not having access to real training images, our strategy produces a network matching the best-performing baseline NiftyReg on skull-stripped data and outperforming all baselines by 4.5–7.3 Dice points across all other data-type pairings. The NiftyReg performance follows ours most closely, whereas the other baselines struggle at cross-contrast registration.

As expected, the baselines perform best on skull-stripped T1w images, because most cannot distinguish between brain and non-brain tissue. The learning baselines break down for full-head images, which they did not see at training. However, VTN's accuracy is not ideal even for skull-stripped T1w scans, possibly due to a noise or MR-contrast shift relative to the training distribution.

Average single-threaded runtimes for ANTs, NiftyReg, and Robust Registration are approx. 10.9, 3.5, and 18.5 minutes on an AMD EPYC 7452 2.35-GHz CPU, with little difference between cost functions. In contrast, the DL methods KeyMorph, VTN, and ours require only 21.9, 28.9, and 49.6 seconds on the CPU, respectively. On an NVIDIA V100 GPU, our method takes less than 4 seconds per image pair for registration, IO, and resampling combined. One-time model initialization requires an additional 25.3 seconds, after which the user could register any number of image pairs.

5. CONCLUSIONS

We demonstrate the feasibility of training accurate affine registration networks that generalize to unseen image types and outperform well-established baselines over a range of image contrasts and resolutions. In an experiment involving diverse real-world data, we show that our network achieves invariance to acquisition specifics while training solely with wildly variable images generated from label maps. While tuning DL methods during inference is challenging, users could always fall back to a classical method in case of a problematic registration pair. Alternatively, our model could serve as initialization of an iterative follow-up registration step.[46] Although the presented experiment focuses on neuroimaging data, our general strategy is applicable to other anatomy as long as label maps are available for training—there is no need for these during inference. Crucially, optimizing the spatial overlap of select anatomical labels enables anatomy-specific registration without the need for segmentation, which we believe has great potential for applications such as real-time motion correction of MRI.[47–49]

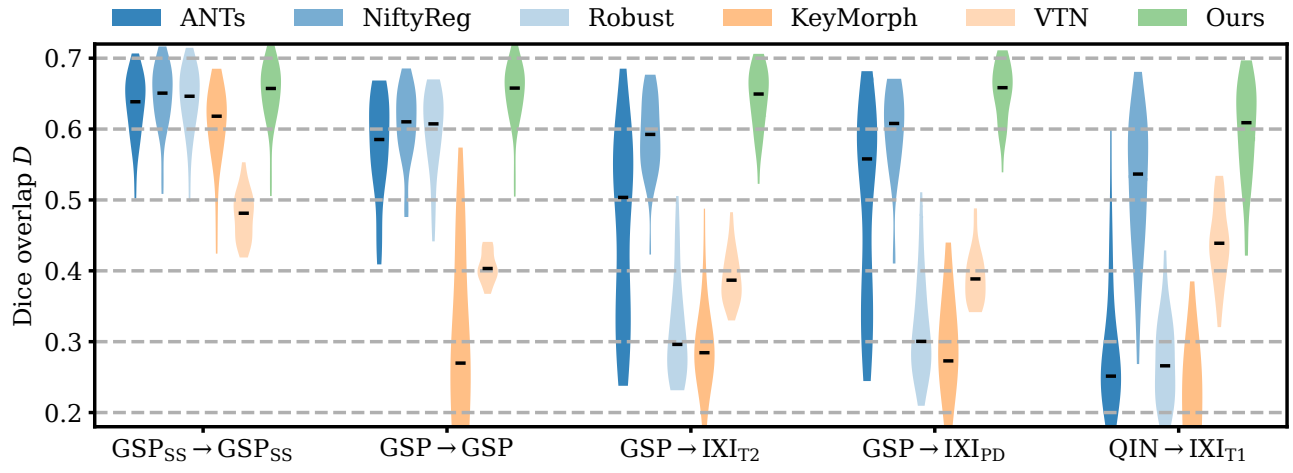

Figure 3. Affine 3D registration accuracy (mean Dice scores over white matter, gray matter, and cerebrospinal fluid). Each bar shows the distribution across 30 separate subject pairs. Subscripts indicate MRI contrast; SS denotes skull-stripping.

Table 1. Uniform hyperparameter sampling ranges $[a, b]$ for synthesizing wildly variable training images from label maps. We abbreviate standard deviation (SD), full width at half maximum (FWHM), and field of view (FOV).

Hyperparameter	**Unit**	a	b
Translation	mm	−30	30
Rotation	°	−45	45
Scaling	%	90	110
Shear	%	90	110
Warp sampling SD	mm	0	4
Warp blurring FWHM	mm	8	32
Label intensity mean	a.u.	0	1
Noise intensity SD	%	3	10
Image blurring FWHM	mm	0	8
Bias field sampling SD	%	0	10
Bias field blurring FWHM	mm	48	64
FOV cropping	%	0	20
Downsampling factor	%	1	8
Gamma exponent	–	0.5	1.5

ACKNOWLEDGMENTS

The authors thank Douglas N. Greve and OASIS Cross-Sectional (PIs D. Marcus, R. Buckner, J. Csernansky, J. Morris; NIH grants P50 AG05681, P01 AG03991, P01 AG026276, R01 AG021910, P20 MH071616, U24 R021382) for sharing data. Support for this research was provided in part by the BRAIN Initiative Cell Census Network (U01 MH117023), the National Institute of Biomedical Imaging and Bioengineering (P41 EB015896, R01 EB023281, R21 EB018907, R01 EB019956, P41 EB030006), the National Institute of Child Health and Human Development (K99 HD101553), the National Institute on Aging (R56 AG064027, R01 AG016495, R01 AG070988), the National Institute of Mental Health (RF1 MH121885, RF1 MH123195), and the National Institute of Neurological Disorders and Stroke (R01 NS070963, R01 NS083534, R01 NS105820). Additional support was provided by the NIH Blueprint for Neuroscience Research (U01 MH093765), part of the multi-institutional Human Connectome Project. The project was made possible by the resources provided by Shared Instrumentation Grants (S10 RR023401, S10 RR019307, S10 RR023043) and by computational hardware generously provided by the Massachusetts Life Sciences Center (https://www.masslifesciences.com).

Bruce Fischl has a financial interest in CorticoMetrics, a company whose medical pursuits focus on brain imaging and measurement technologies. Massachusetts General Hospital and Mass General Brigham review and manage this interest in accordance with their conflict of interest policies. The authors have no other known financial and personal interests that could have inappropriately influenced this work.

REFERENCES

[1] Friston, K. J., Ashburner, J., Frith, C. D., Poline, J.-B., Heather, J. D., and Frackowiak, R. S., "Spatial registration and normalization of images," *Human Brain Mapping* **3**(3), 165–189 (1995).

[2] Cox, R. W., "AFNI: software for analysis and visualization of functional magnetic resonance neuroimages," *Computers and Biomedical Research* **29**(3), 162–173 (1996).

[3] Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., and Smith, S. M., "Fsl," *Neuroimage* **62**(2), 782–790 (2012).

[4] Modat, M., Cash, D. M., Daga, P., Winston, G. P., Duncan, J. S., and Ourselin, S., "Global image registration using a symmetric block-matching approach," *Journal of Medical Imaging* **1**(2), 024003 (2014).

[5] Jiang, A., Kennedy, D. N., Baker, J. R., Weisskoff, R. M., Tootell, R. B., Woods, R. P., Benson, R. R., Kwong, K. K., Brady, T. J., Rosen, B. R., et al., "Motion detection and correction in functional MR imaging," *Human Brain Mapping* **3**(3), 224–235 (1995).

[6] Reuter, M., Rosas, H. D., and Fischl, B., "Highly accurate inverse consistent registration: A robust approach," *Neuroimage* **53**(4), 1181–1196 (2010).

[7] Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., and Gee, J. C., "A reproducible evaluation of ANTs similarity metric performance in brain image registration," *Neuroimage* **54**(3), 2033–2044 (2011).

[8] Hoffmann, M., Carpenter, T. A., Williams, G. B., and Sawiak, S. J., "A survey of patient motion in disorders of consciousness and optimization of its retrospective correction," *Magnetic Resonance Imaging* **33**(3), 346–350 (2015).

[9] De Vos, B. D., Berendsen, F. F., Viergever, M. A., et al., "A deep learning framework for unsupervised affine and deformable image registration," *Medical Image Analysis* **52**, 128–143 (2019).

[10] Zhao, S., Dong, Y., Chang, E. I., Xu, Y., et al., "Recursive cascaded networks for unsupervised medical image registration," in [*Proceedings of the IEEE/CVF International Conference on Computer Vision*], 10600–10610 (2019).

[11] Zhao, S., Lau, T., Luo, J., Eric, I., Chang, C., and Xu, Y., "Unsupervised 3D end-to-end medical image registration with volume tweening network," *IEEE Journal of Biomedical and Health Informatics* **24**(5), 1394–1404 (2019).

[12] Shen, Z., Han, X., Xu, Z., and Niethammer, M., "Networks for joint affine and non-parametric image registration," in [*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*], 4224–4233 (2019).

[13] Zhu, Z., Cao, Y., Qin, C., Rao, Y., Lin, D., Dou, Q., Ni, D., and Wang, Y., "Joint affine and deformable three-dimensional networks for brain MRI registration," *Medical Physics* **48**(3), 1182–1196 (2021).

[14] Fischmeister, F. P. S., Höllinger, I., Klinger, N., Geissler, A., Wurnig, M., Matt, E., Rath, J., Robinson, S., Trattnig, S., and Beisteiner, R., "The benefits of skull stripping in the normalization of clinical fMRI data," *Neuroimage: Clinical* **3**, 369–380 (2013).

[15] Andrade, N., Faria, F. A., and Cappabianco, F. A. M., "A practical review on medical image registration: From rigid to deep learning based approaches," in [*2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*], 463–470, IEEE (2018).

[16] Hoffmann, M., Frost, R., Salat, D., Tisdall, M. D., Polimeni, J., and van der Kouwe, A., "Real-time brain masking algorithm improves motion tracking accuracy in scans with volumetric navigators (vNavs)," in [*International Society for Magnetic Resonance in Medicine*], 3367, ISMRM (2020).

[17] Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., Van Der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., et al., "Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain," *Neuron* **33**(3), 341–355 (2002).

[18] Jenkinson, M. and Smith, S., "A global optimisation method for robust affine registration of brain images," *Medical Image Analysis* **5**(2), 143–156 (2001).

[19] Yu, E. M., Wang, A. Q., Dalca, A. V., and Sabuncu, M. R., "KeyMorph: Robust Multi-modal Affine Registration via Unsupervised Keypoint Detection," in [*Medical Imaging with Deep Learning*], (2022).

[20] Nestares, O. and Heeger, D. J., "Robust multiresolution alignment of MRI brain volumes," *Magnetic Resonance in Medicine* **43**(5), 705–715 (2000).

[21] Gelfand, N., Mitra, N. J., Guibas, L. J., and Pottmann, H., "Robust global registration," in [*Symposium on Geometry Processing*], **2**(3), 5, Vienna, Austria (2005).

[22] Puglisi, G. and Battiato, S., "A robust image alignment algorithm for video stabilization purposes," *IEEE Transactions on Circuits and Systems for Video Technology* **21**(10), 1390–1400 (2011).

[23] Billings, S. D., Boctor, E. M., and Taylor, R. H., "Iterative most-likely point registration (IMLP): A robust algorithm for computing optimal shape alignment," *PLOS One* **10**(3), e0117688 (2015).

[24] Smith, S. M., "Fast robust automated brain extraction," *Human Brain Mapping* **17**(3), 143–155 (2002).

[25] Iglesias, J. E., Liu, C.-Y., Thompson, P. M., and Tu, Z., "Robust brain extraction across datasets and comparison with publicly available methods," *IEEE Transactions on Medical Imaging* **30**(9), 1617–1634 (2011).

[26] Eskildsen, S. F., Coupé, P., Fonov, V., Manjón, J. V., Leung, K. K., Guizard, N., Wassef, S. N., Østergaard, L. R., Collins, D. L., Initiative, A. D. N., et al., "BEaST: brain extraction based on nonlocal segmentation technique," *Neuroimage* **59**(3), 2362–2373 (2012).

[27] Hoopes, A., Mora, J. S., Dalca, A. V., Fischl, B., and Hoffmann, M., "SynthStrip: skull-stripping for any brain image," *Neuroimage* **260**, 119474 (2022).
[28] Billot, B., Greve, D. N., Van Leemput, K., Fischl, B., Iglesias, J. E., and Dalca, A., "A Learning Strategy for Contrast-agnostic MRI Segmentation," in [*Medical Imaging with Deep Learning*], 75–93, PMLR (2020).
[29] Hoffmann, M., Billot, B., Greve, D. N., Iglesias, J. E., Fischl, B., and Dalca, A. V., "SynthMorph: learning contrast-invariant registration without acquired images," *IEEE Transactions on Medical Imaging* **41**(3), 543–558 (2022).
[30] Fischl, B., "FreeSurfer," *Neuroimage* **62**(2), 774–781 (2012). 20 YEARS OF fMRI.
[31] Milletari, F., Navab, N., and Ahmadi, S.-A., "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in [*3DV*], 565–571 (2016).
[32] Moyer, D., Abaci Turk, E., Grant, P. E., Wells, W. M., and Golland, P., "Equivariant Filters for Efficient Tracking in 3D Imaging," in [*MICCAI*], 193–202, Springer (2021).
[33] Besl, P. J. and McKay, N. D., "Method for registration of 3-D shapes," in [*Sensor fusion IV: control paradigms and data structures*], **1611**, 586–606, Spie (1992).
[34] Meyer, C. R., Leichtman, G. S., Brunberg, J. A., Wahl, R. L., and Quint, L. E., "Simultaneous usage of homologous points, lines, and planes for optimal, 3-D, linear registration of multimodality imaging data," *IEEE Transactions on Medical Imaging* **14**(1), 1–11 (1995).
[35] Myronenko, A. and Song, X., "Point set registration: Coherent point drift," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **32**(12), 2262–2275 (2010).
[36] Toews, M. and Wells III, W. M., "Efficient and robust model-to-image alignment using 3D scale-invariant features," *Medical Image Analysis* **17**(3), 271–282 (2013).
[37] Rister, B., Horowitz, M. A., and Rubin, D. L., "Volumetric image registration from invariant keypoints," *IEEE Transactions on Image Processing* **26**(10), 4900–4910 (2017).
[38] Greve, D. N., Billot, B., Cordero, D., et al., "A deep learning toolbox for automatic segmentation of subcortical limbic structures from MRI images," *Neuroimage* **244**, 118610 (2021).
[39] Marcus, D. S., Wang, T. H., et al., "Open access series of imaging studies (oasis): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults," *Journal of Cognitive Neuroscience* **19**(9), 1498–507 (2007).
[40] Holmes, A. J., Hollinshead, M. O., O'keefe, T. M., et al., "Brain Genomics Superstruct Project initial data release with structural, functional, and behavioral measures," *Scientific Data* **2**(1), 1–16 (2015).
[41] Imperial College London, "IXI Dataset," (2015).
[42] Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., et al., "The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository," *Journal of Digital Imaging* **26**(6), 1045–1057 (2013).
[43] Mamonov, A. and Kalpathy-Cramer, J., "Data From QIN GBM Treatment Response," (2016).
[44] Prah, M., Stufflebeam, S., Paulson, E., Kalpathy-Cramer, J., Gerstner, E., Batchelor, T., Barboriak, D., Rosen, B., and Schmainda, K., "Repeatability of standardized and normalized relative CBV in patients with newly diagnosed glioblastoma," *American Journal of Neuroradiology* **36**(9), 1654–1661 (2015).
[45] Pustina, D. and Cook, P., "Anatomy of an antsRegistration call," (2017).
[46] Balakrishnan, G., Zhao, A., Sabuncu, M., Guttag, J., and Dalca, A. V., "VoxelMorph: A Learning Framework for Deformable Medical Image Registration," *IEEE Transactions on Medical Imaging* **38**(8), 1788–1800 (2019).
[47] Thesen, S., Heid, O., Mueller, E., and Schad, L. R., "Prospective acquisition correction for head motion with image-based tracking for real-time fMRI," *Magnetic Resonance in Medicine* **44**(3), 457–465 (2000).
[48] White, N., Roddey, C., Shankaranarayanan, A., Han, E., Rettmann, D., Santos, J., Kuperman, J., and Dale, A., "PROMO: real-time prospective motion correction in MRI using image-based tracking," *Magnetic Resonance in Medicine* **63**(1), 91–105 (2010).
[49] Tisdall, M. D., Hess, A. T., Reuter, M., Meintjes, E. M., Fischl, B., and van der Kouwe, A. J., "Volumetric navigators for prospective motion correction and selective reacquisition in neuroanatomical MRI," *Magnetic Resonance in Medicine* **68**(2), 389–399 (2012).