Unsupervised fetal brain MR segmentation using multi-atlas deep learning registration

Valentin Comte[1] Mireia Alenya[1] Andrea Urru[1] Ayako Nakaki[2]
Francesca Crovetto[2] Oscar Camara[1] Elisenda Eixarch[2] Fàtima Crispi[2]
Gemma Piella Fenoy[1] Mario Ceresa[1] Miguel A. González Ballester[1,2]

[1] BCN MedTech, Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain
[2] Maternal Fetal Medicine, BCNatal, Center for Maternal Fetal and Neonatal Medicine (Hospital Clínic and Hospital Sant Joan de Déu), Barcelona, Spain
[3] ICREA, Barcelona, Spain

Abstract

Deep Learning is now well established as the most efficient method for medical image segmentation. Yet, it requires large training sets and ground-truth labels, annotated by clinicians in a time-consuming process. We propose an unsupervised segmentation method using multi-atlas registration. The architecture of our registration model is composed of cascaded networks that produce small amounts of displacement to warp progressively the moving image towards the fixed image. Once the networks are trained, multiple annotated magnetic resonance (MR) fetal brain images and their labels are registered with the image to segment, the resulting warped labels are then combined to form a refined segmentation. Experiments show that our cascaded architecture outperforms the state-of-the-art registration methods by a significant margin. Furthermore, the derived multi-atlas segmentation method obtains similar results as one of the most robust state-of-the-art segmentation methods, without using any labels during training.

1 Introduction

Deformable image registration is an essential procedure in many image analysis applications. It aims to compute a non-linear deformation field which aligns a pair of images together. In the scope of medical imaging, the alignment of anatomical structures finds applications for intraoperative guidance [9, 10, 5], 3D reconstruction [6], or segmentation [11]. Classical registration methods use intensity-based similarity metrics, such as cross-correlation (CC), mutual information (MI), or sum of square distance (SSD). During the past decade, Deep Learning (DL) has emerged as a powerful alternative, achieving similar results in a fraction of the time needed by classical approaches. Those methods are mostly based on UNet-like networks and use the similarity metrics listed above as loss function [3]. Although DL-based registration has achieved remarkable results, its performances in the context of atlas-based segmentation are still significantly lower than most DL networks for segmentation [8, 7].

36th Conference on Neural Information Processing Systems (NeurIPS 2022).

2 Method

2.1 Cascaded registration

We present a cascaded registration model, which consists of several networks that warp progressively the moving image into the fixed image. As illustrated by Figure 1, the first network takes as an input the moving and fixed images, $\mathcal{X}_{\mathrm{mv}}$ and $\mathcal{X}_{\mathrm{fx}}$, and outputs a dense deformation field φ_0 which warps the moving image to align it with the fixed image as $\mathcal{X}_{\mathrm{mv}} \circ \varphi_0 \approx \mathcal{X}_{\mathrm{fx}}$. But only a fraction k_0 of φ_0 is applied to the moving image, producing a partially warped image $\mathcal{X}_{\mathrm{wp,0}} = \mathcal{X}_{\mathrm{mv}} \circ (k_0 \cdot \varphi_0)$. The second network takes $\mathcal{X}_{\mathrm{wp,0}}$ and $\mathcal{X}_{\mathrm{fx}}$ as inputs and outputs φ_1, which is scaled by k_1 summed with $k_0 \cdot \varphi_0$ to warp $\mathcal{X}_{\mathrm{mv}}$ into $\mathcal{X}_{\mathrm{wp,1}}$. The rest of the networks acts repeatedly with the same pattern, hence the general expression for the n-th warped image is:

$$\mathcal{X}_{\mathrm{wp,n}} = \left(\sum_{i=0}^{n} k_i \varphi_i \right) \circ \mathcal{X}_{\mathrm{mv}} \tag{1}$$

Where k_i are the scaling factors of the deformation fields φ_i. The loss function is composed of two terms, an image similarity loss based on normalized cross-correlation (NCC) and a regularization loss of the deformation field, as presented in [3]. Both are computed on the final outputs of the networks, $\mathcal{X}_{\mathrm{wp,n}}$ and $\varphi = \varphi_n \circ \varphi_{n-1} \circ \cdots \varphi_0$ respectively, ensuring a collaborative behaviour of the cascades. The two major major differences with [12] are: 1) the scaling factors k_i are not constrained to $k_i = 1/n$ or any other value, so that the cascaded networks find the set of scaling factors yielding the best alignment, 2) the deformation fields are accumulated and only applied to the moving image, so that there is no loss of information by applying multiple interpolations to the same image.

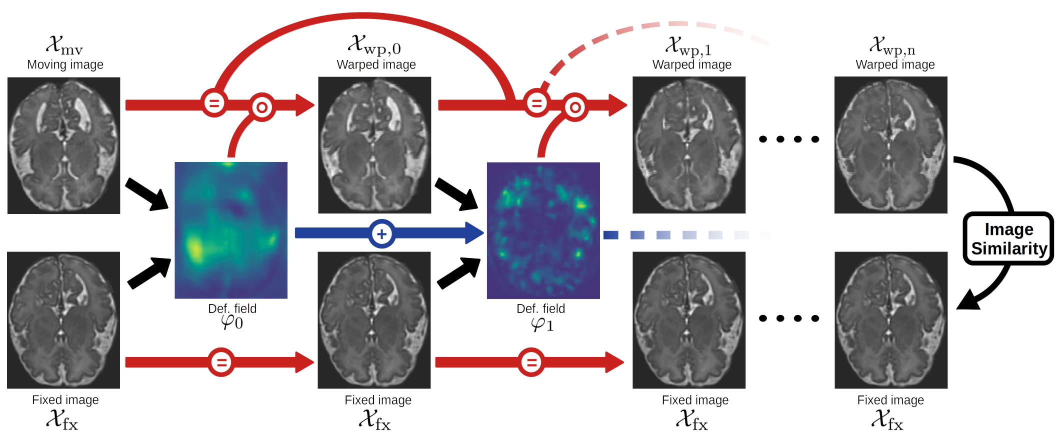

Figure 1: Overview of the cascaded registration model.

2.2 Multi-atlas segmentation

Relying on the cascaded registration method presented in Section 2.1, we propose a segmentation method based on multi-atlas registration. Multiple moving images and their ground-truth labels are registered to the image to segment. The best aligned volumes are selected based on the average local cross-correlation with the fixed images. The corresponding warped labels are then combined to form a refined segmentation using a local weighted voting approach [1], which consists of propagating the labels of the warped images based on a weighting strategy, which gives more weight to the labels corresponding to the highest local similarity. The weights assigned to each voxel k of the atlas i can be expressed as:

$$w_{k,i} = |m(i \in \Omega)|^p \tag{2}$$

where Ω is a region of size $d \times d \times d$ around the voxel i, m is the average local similarity metrics in the region Ω, and p is the gain factor, which can take different values depending on the similarity metrics used. Here we adopted a NCC metrics and a gain factor $p = 1$.

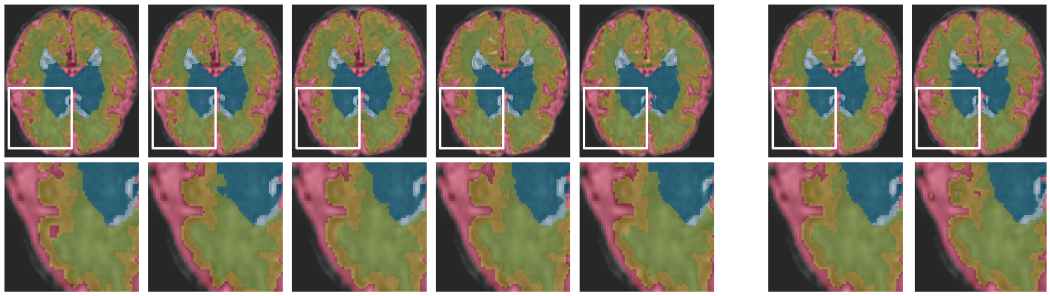

Figure 2: Example of the label selection by local weighted voting.

3 Experimental settings

The model was developed using Python 3.8, Pytorch 1.11, and a GPU Nvidia Titan Xp. As a baseline architecture we used a modified version of VoxelMorph [3] with smaller layer sizes to limit memory usage. We used a datasets of 170 fetal brain MRIs between 32 and 37 gestational weeks. The scans we aquired in Hospital San Joan de Déu and Hospital Clinic of Barcelona. Single-shot fast spin-echo T2-weighted was performed using two 3.0 T MR scanners, Philips Ingenia (repetition/echo time : 1570/150 ms, slice thickness : 3 mm, field of view 290×250 mm, voxel spacing $0.7 \times 0.7 \times 3.0$ mm), and Siemens Magneton Vida (repetition/echo time : 1390/160 ms, slice thickness : 3 mm, field of view 230×230 mm, voxel spacing $1.2 \times 1.2 \times 3.0$ mm). All fetuses included in this study did not have any major malformation. The dataset is split into 140-10-20 for train, validation and test sets, respectively. As a preprocessing step, the images are cropped to remove the non brain regions, resized to $128 \times 128 \times 128$ voxels, and normalized between 0 and 1.

4 Results

The results of our cascaded registration method are given in Table 1. It shows the average Dice scores between the warped and ground-truth labels of the fixed images for all anatomical regions. The average Dice scores after rigid registration are given as a reference. The results obtained with VoxelMorph [3] and TransMorph [4] are also presented in this table, showing that our method outperforms them for all the anatomical labels.

Table 1: Average Dice scores obtained using ANTs rigid registration [2], VoxelMorph[3], TransMorph[4] and the proposed method.

Label	Rigid	VoxelMorph[3]	TransMorph[4]	Ours
CSF	0.516 ± 0.008	0.802 ± 0.007	0.796 ± 0.009	$\mathbf{0.855 \pm 0.008}$
Grey Matter	0.501 ± 0.003	0.724 ± 0.009	0.725 ± 0.011	$\mathbf{0.781 \pm 0.006}$
White Matter	0.663 ± 0.003	0.782 ± 0.005	0.780 ± 0.007	$\mathbf{0.850 \pm 0.004}$
Ventricules	0.495 ± 0.011	0.761 ± 0.017	0.751 ± 0.016	$\mathbf{0.812 \pm 0.012}$
Cerebellum	0.784 ± 0.010	0.881 ± 0.009	0.873 ± 0.014	$\mathbf{0.932 \pm 0.005}$
Thalamus	0.812 ± 0.004	0.871 ± 0.008	0.871 ± 0.007	$\mathbf{0.915 \pm 0.004}$
Brain Stem	0.767 ± 0.005	0.856 ± 0.011	0.853 ± 0.012	$\mathbf{0.919 \pm 0.004}$
Average	0.648 ± 0.054	0.811 ± 0.023	0.807 ± 0.023	$\mathbf{0.866 \pm 0.022}$

Table 2 presents the results of the segmentation using multi-atlas registration. Each of the 20 images of our training set was segmented using the method described in Section 2.2. The warped segmentations of the 10 best aligned volumes were used to produce the refined segmentation in a multi-atlas approach. The Dice scores obtained by training nnUnet on the same training set are given, showing that our method achieves similar performances, while avoiding the need of large annotated datasets for training.

Table 2: Average Dice scores obtained using multi-atlas registration and nnUnet [7]

Label	Ours	NnUNet[7]
CSF	$\mathbf{0.923 \pm 0.006}$	0.897 ± 0.011
Grey Matter	$\mathbf{0.877 \pm 0.006}$	0.871 ± 0.012
White Matter	0.915 ± 0.006	$\mathbf{0.919 \pm 0.006}$
Ventricules	$\mathbf{0.902 \pm 0.005}$	0.879 ± 0.011
Cerebellum	0.964 ± 0.004	$\mathbf{0.965 \pm 0.004}$
Thalamus	0.950 ± 0.002	$\mathbf{0.957 \pm 0.004}$
Brain Stem	$\mathbf{0.955 \pm 0.003}$	0.953 ± 0.006
Average	$\mathbf{0.926 \pm 0.012}$	0.920 ± 0.014

5 Conclusion

We present a registration model based on cascaded networks. Experiments on our fetal brain MR dataset show that this architecture performs significantly better than the state-of-the-art methods tested [3, 4]. The derived multi atlas segmentation method achieves similar performances as one of the most robust state-of-the-art segmentation methods [7], without requiring any labels for training.

Potential negative societal impact

This work aims at helping clinicians with automatic brain segmentation, and to the best of our knowledge, does not represent any potential negative societal impact.

Acknowledgments

This publication is part of the project PCI2021-122044-2A, funded by the project ERA-NET NEURON Cofund2, by MCIN/AEI/10.13039/501100011033/ and by the European Union "NextGenerationEU"/PRTR. G. Piella is supported by ICREA under the ICREA Academia programme.

References

[1] Xabier Artaechevarria, Arrate Munoz-Barrutia, and Carlos Ortiz-de Solorzano. Combination strategies in multi-atlas image segmentation: application to brain mr data. *IEEE transactions on medical imaging*, 28(8):1266–1277, 2009.

[2] Brian B Avants, Nicholas J Tustison, Gang Song, Philip A Cook, Arno Klein, and James C Gee. A reproducible evaluation of ants similarity metric performance in brain image registration. *Neuroimage*, 54(3):2033–2044, 2011.

[3] Guha Balakrishnan, Amy Zhao, Mert R. Sabuncu, John V. Guttag, and Adrian V. Dalca. Voxelmorph: A learning framework for deformable medical image registration. *CoRR*, abs/1809.05231, 2018.

[4] Junyu Chen, Eric C Frey, Yufan He, William P Segars, Ye Li, and Yong Du. Transmorph: Transformer for unsupervised medical image registration. *Medical Image Analysis*, page 102615, 2022.

[5] Langston T Holly, Orin Bloch, and J Patrick Johnson. Evaluation of registration techniques for spinal image guidance. *Journal of Neurosurgery: Spine*, 4(4):323–328, 2006.

[6] BF Hutton, AZ Kyme, YH Lau, DW Skerrett, and RR Fulton. A hybrid 3-d reconstruction/registration algorithm for correction of head motion in emission tomography. *IEEE Transactions on Nuclear Science*, 49(1):188–194, 2002.

[7] Fabian Isensee, Jens Petersen, Andre Klein, David Zimmerer, Paul F. Jaeger, Simon Kohl, Jakob Wasserthal, Gregor Koehler, Tobias Norajitra, Sebastian Wirkert, and Klaus H. Maier-Hein. nnu-net: Self-adapting framework for u-net-based medical image segmentation, 2018.

[8] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[9] Guy-Anne Turgeon, Glen Lehmann, Gerard Guiraudon, Maria Drangova, David Holdsworth, and Terry Peters. 2d-3d registration of coronary angiograms for cardiac procedure planning and guidance. *Medical physics*, 32(12):3737–3749, 2005.

[10] Osamu Ukimura, Mihir M Desai, Suzanne Palmer, Samuel Valencerina, Mitchell Gross, Andre L Abreu, Monish Aron, and Inderbir S Gill. 3-dimensional elastic registration system of prostate biopsy location by real-time 3-dimensional transrectal ultrasound guidance with magnetic resonance/transrectal ultrasound image fusion. *The Journal of urology*, 187(3):1080–1086, 2012.

[11] Xiaofeng Yang and Baowei Fei. 3d prostate segmentation of ultrasound images combining longitudinal image registration and machine learning. In *Medical Imaging 2012: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 8316, page 83162O. International Society for Optics and Photonics, 2012.

[12] Shengyu Zhao, Yue Dong, Eric I-Chao Chang, and Yan Xu. Recursive cascaded networks for unsupervised medical image registration. *CoRR*, abs/1907.12353, 2019.