Bi-level Learning of Task-Specific Decoders for Joint Registration and One-Shot Medical Image Segmentation

Supplementary Material

In this document, we supply some experiments of our proposed Bi-JROS. We first provide visual results for different tasks among state-of-the-art approaches in Sec. 1 to show the excellence of our algorithm. To further validate the rapid adaptability of Bi-JROS, Sec. 2 presents a detailed performance comparison of naive alternating training, BRBS, and our method on new datasets.

1. Comparison results:

Fig. 2 illustrates the segmentation performance of the JRS method across various anatomical structures. For the registration task, our approach exhibited the best segmentation performance in nine out of thirteen structures, matched the performance of BRBS in two structures, and secured the second rank in two structures. Through a comprehensive analysis of Figs. 2 and 3 (in the manuscript), we observe a distinct trend of synchronous increase and decrease between registration and segmentation tasks. This phenomenon not only demonstrates the coordinated performance variations of these two tasks but also further highlights the intrinsic mutual enhancement relationship between them. The discovery of this trend aligns well with our motivation.

To more vividly demonstrate the visualization outcomes, we further present a comparative display of segmentation

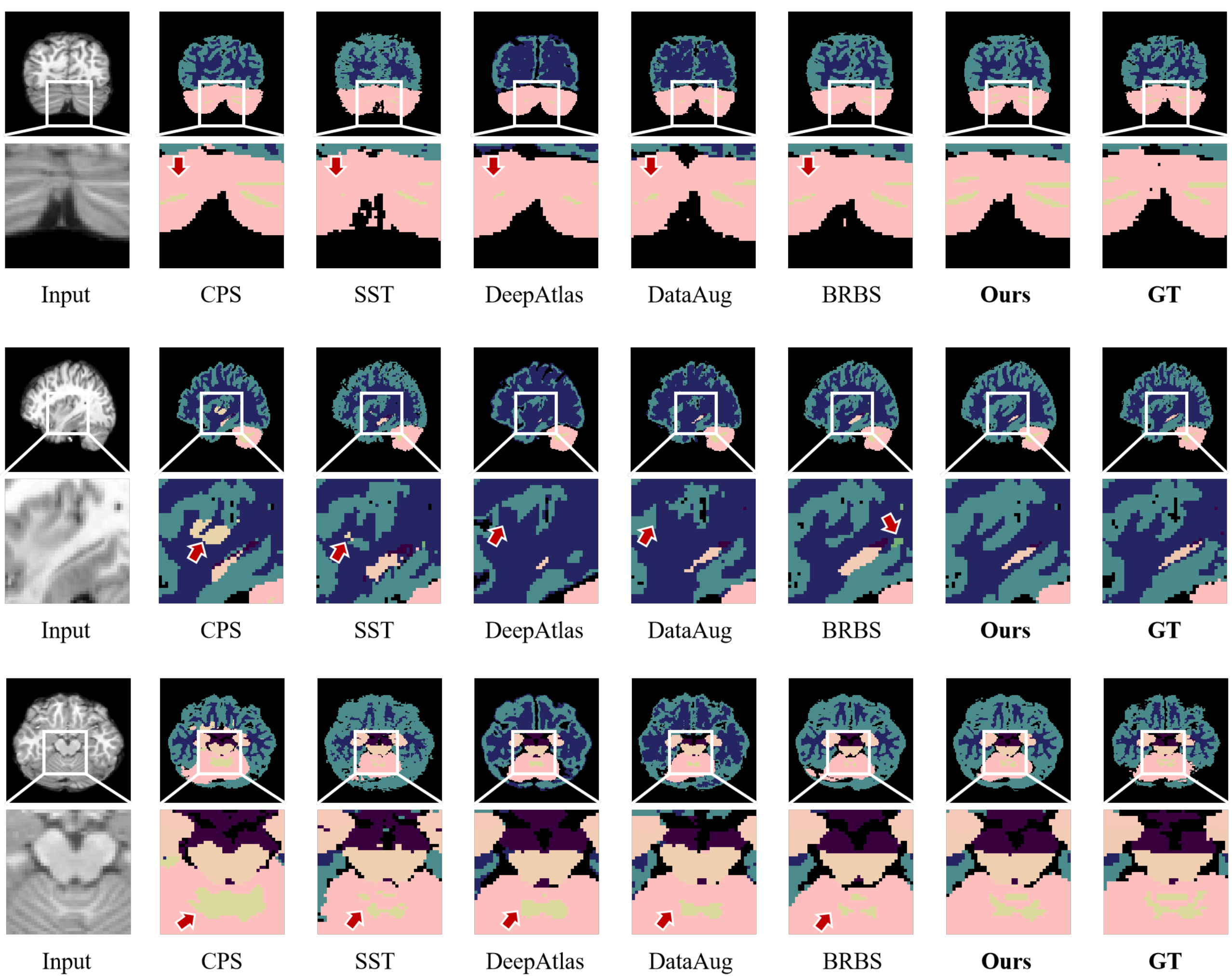

Figure 1. Segmentation visualization results of different methods on different structures. The red arrows point to the segmentation errors.

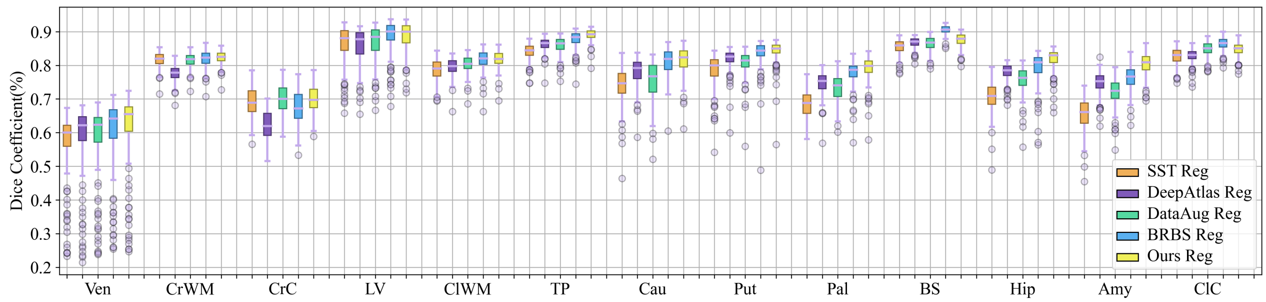

Figure 2. Boxplots of performance comparison towards various registration methods of Dice scores with 13 categories of brain anatomical structures.

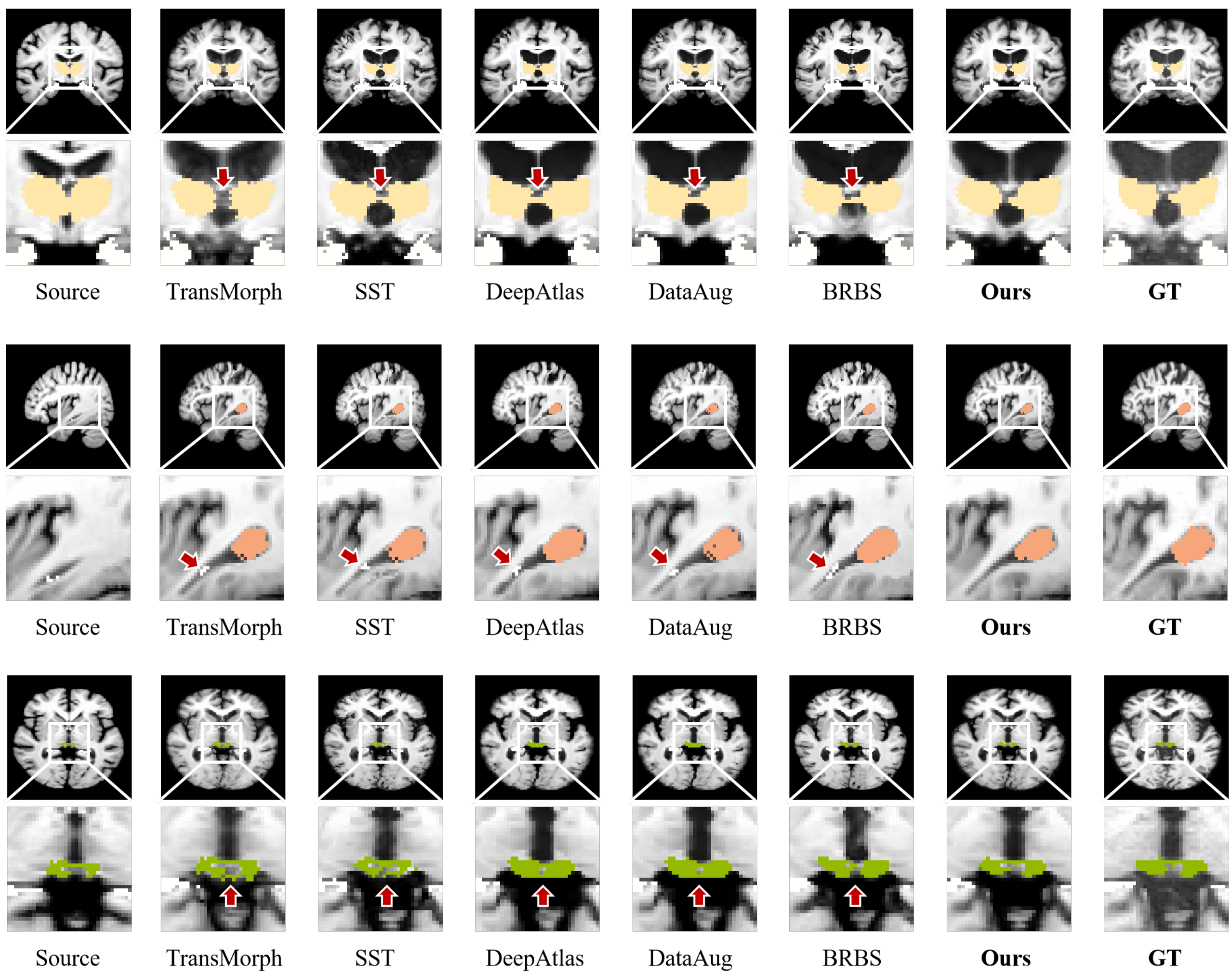

Figure 3. Registration visualization results of different methods on different structures. The red arrows point to the registration errors.

and registration results of six representative methods across different individuals and anatomical structures in Figs. 1 and 3. In Fig. 1, our approach achieves segmentation results with the highest overlap with the GT on the Cerebral Cortex and Cerebellar Cortex. In Fig. 3, our method exhibits the most similar registration results to the target image, particularly in the Thalamus Proper (yellow area), Lateral Ventricle (orange area), and Brain Stem (green area), closely aligning with the GT. This multi-perspective comparison distinctly highlights the accuracy and effectiveness of our approach.

2. Ablation Experiments:

We conducted a validation of the proposed Bi-JROS method's adaptability across various datasets. The adaptability of the method on OASIS [2] has been elaborately discussed in the main text. In this section, we further present the experimental results on the HCP [3]. We divide the data into 6, and 6 for training and testing, respectively. The segmentation labels of the HCP dataset are consistent with those mentioned in the manuscript.

The data comparison presented in Tab. 1 showcases the adaptability of naive alternating training, the BRBS [1], and our proposed Bi-JROS in scenarios with extremely limited data volume. The BRBS method, through multifaceted design optimizations addressing alignment distortions and lack of data diversity in the registration process, demonstrates a slight advantage over simple alternating training. However, Tab. 1 distinctly indicates that even under limited data conditions, the performance of BRBS significantly falls short of our method. This outcome further validates the superior adaptability of our proposed approach in dealing with small volumes of new data.

Methods	Seg	Reg	
	Dice(%)	Dice(%)	Ncc
Initial	-	71.5 ± 1.4	0.145 ± 0.007
Naive Alter	80.6 ± 0.9	77.9 ± 1.2	0.311 ± 0.009
BRBS	81.2 ± 0.9	78.1 ± 1.4	0.264 ± 0.006
Ours	$\mathbf{81.7 \pm 0.9}$	$\mathbf{80.3 \pm 1.0}$	$\mathbf{0.357 \pm 0.009}$

Table 1. Results among BRBS, naive alternative training and our Bi-JROS for registration (Reg) and segmentation (Seg) tasks. The top-ranked method is highlighted in **bolded** form.

References

[1] Yuting He, Rongjun Ge, Xiaoming Qi, Yang Chen, Jiasong Wu, Jean-Louis Coatrieux, Guanyu Yang, and Shuo Li. Learning better registration to learn better few-shot medical image segmentation: Authenticity, diversity, and robustness. *IEEE Trans. Neural Networks Learn. Syst.*, 35(2):2588–2601, 2024. 3

[2] Daniel S Marcus, Anthony F Fotenos, John G Csernansky, John C Morris, and Randy L Buckner. Open access series of imaging studies: longitudinal mri data in nondemented and demented older adults. *J. Cogn. Neurosci.*, 22(12):2677–2684, 2010. 3

[3] David C Van Essen, Stephen M Smith, Deanna M Barch, Timothy EJ Behrens, Essa Yacoub, Kamil Ugurbil, Wu-Minn HCP Consortium, et al. The wu-minn human connectome project: an overview. *Neuroimage*, 80:62–79, 2013. 3