

SACB-Net: Spatial-awareness Convolutions for Medical Image Registration

Supplementary Material

1. The Encoder Architecture

Figure 1 illustrates the architecture of the shared encoder in SACB-Net. The multi-scale features extracted from each convolutional block are used for pyramid flow estimation.

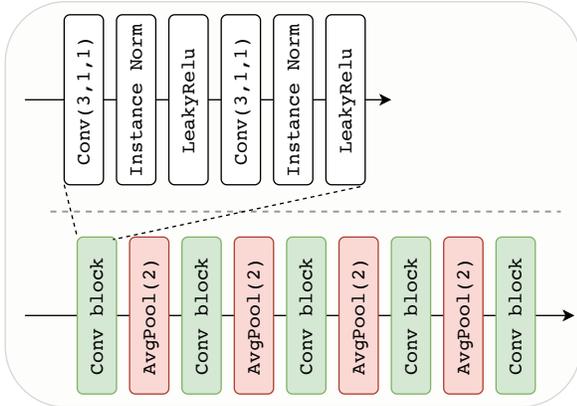


Figure 1. Diagram of the shared encoder architecture, featuring five convolutional blocks to extract multi-scale feature maps and four average pooling layers for downsampling.

2. Normalized Cross-Correlation Loss

The normalized cross-correlation loss denotes as

$$\mathcal{L}_{NCC}(I_f, I_m \circ \phi) = - \sum_{p \in \Omega} \frac{\sum_{p_i} (I_f(p_i) - \overline{I_f(p)}) (I_m \circ \phi(p_i) - \overline{I_m \circ \phi(p)})}{\sqrt{\sum_{p_i} (I_f(p_i) - \overline{I_f(p)})^2 \sum_{p_i} (I_m \circ \phi(p_i) - \overline{I_m \circ \phi(p)})^2}}, \quad (1)$$

where $\overline{I_f(p)}$ and $\overline{I_m \circ \phi(p)}$ denote the local mean intensity values of the images. Here, p_i represents the positions within a local w^3 window centered at p . During training, we set the window size w to 9.

3. Evaluation on SSIM metric

Table 1 presents the results of Structural Similarity Index Measure (SSIM) for the comparison methods. However, it has been highlighted by [10] that a higher degree of image similarity does not always indicate improved registration; anatomical structures are more reliable measures.

4. Discussion on LPBA dataset

Figure 2 presents the boxplot of Dice scores for different organs. It is clear that organ size significantly influences reg-

Table 1. SSIM \uparrow results.

Method	IXI	LPBA
Affine	0.680 \pm 0.012	0.716 \pm 0.027
VM-1 [1]	0.896 \pm 0.012	0.940 \pm 0.012
VM-2 [1]	0.900 \pm 0.012	0.944 \pm 0.012
NCA-Morph [9]	0.880 \pm 0.016	0.922 \pm 0.014
LKU [3]	0.860 \pm 0.015	0.949 \pm 0.012
B-Spline-Diff [8]	0.858 \pm 0.015	0.887 \pm 0.023
Fourier-Net [4]	0.841 \pm 0.016	0.908 \pm 0.017
LapIRN [7]	0.898 \pm 0.013	0.940 \pm 0.013
PRNet++ [5]	0.929 \pm 0.011	0.959 \pm 0.012
ModeT [11]	0.922 \pm 0.012	0.960 \pm 0.010
Im2Grid [6]	0.890 \pm 0.015	0.953 \pm 0.011
RDN [2]	0.906 \pm 0.011	0.950 \pm 0.011
Ours	0.915 \pm 0.012	0.965 \pm 0.011

istration performance, with the gallbladder being the smallest and most challenging to register, while the liver is the easiest.

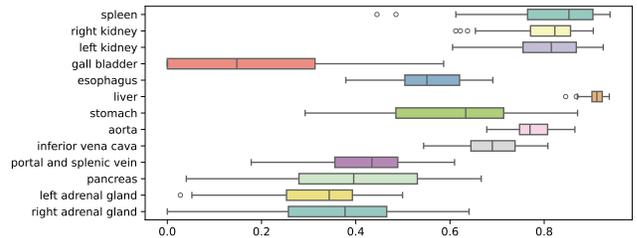


Figure 2. The boxplot of Dice scores for 13 labeled organs.

We selected the case with the lowest average Dice (< 0.5) as a failure case and presented in Figure 3. As shown, small organs are prone to mismatches, significantly impacting registration accuracy.



Figure 3. Illustration of a failure case registration with Dice (< 0.5).

5. Additional Results

Figures 4, 5, and 6 present additional visualization results for the LPBA, IXI and Abdomen CT datasets, respectively, as shown on the following pages.

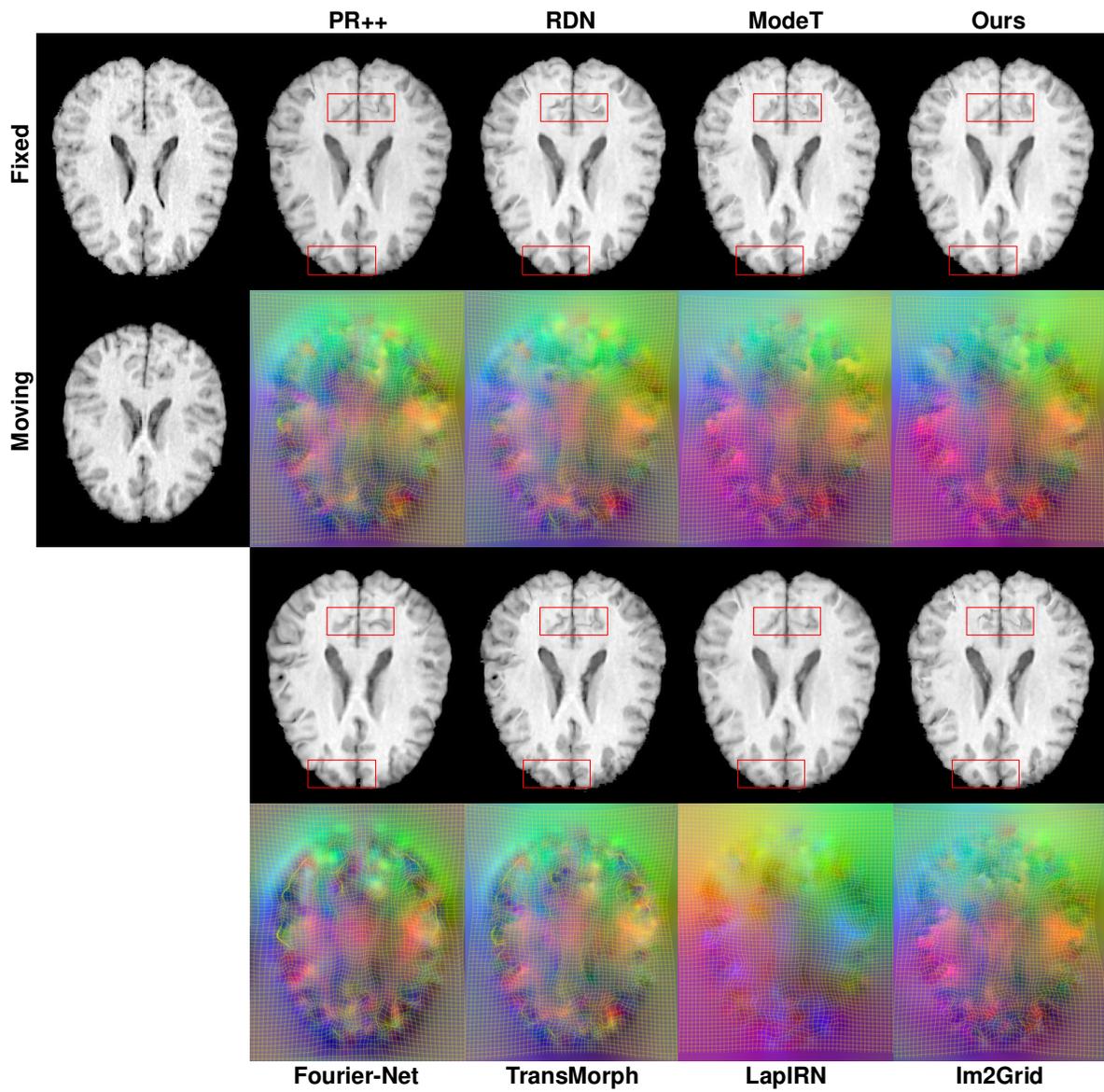


Figure 4. Visual comparisons on LPBA dataset. Columns 2-5: warped moving images (top), displacement fields as RGB images (bottom).

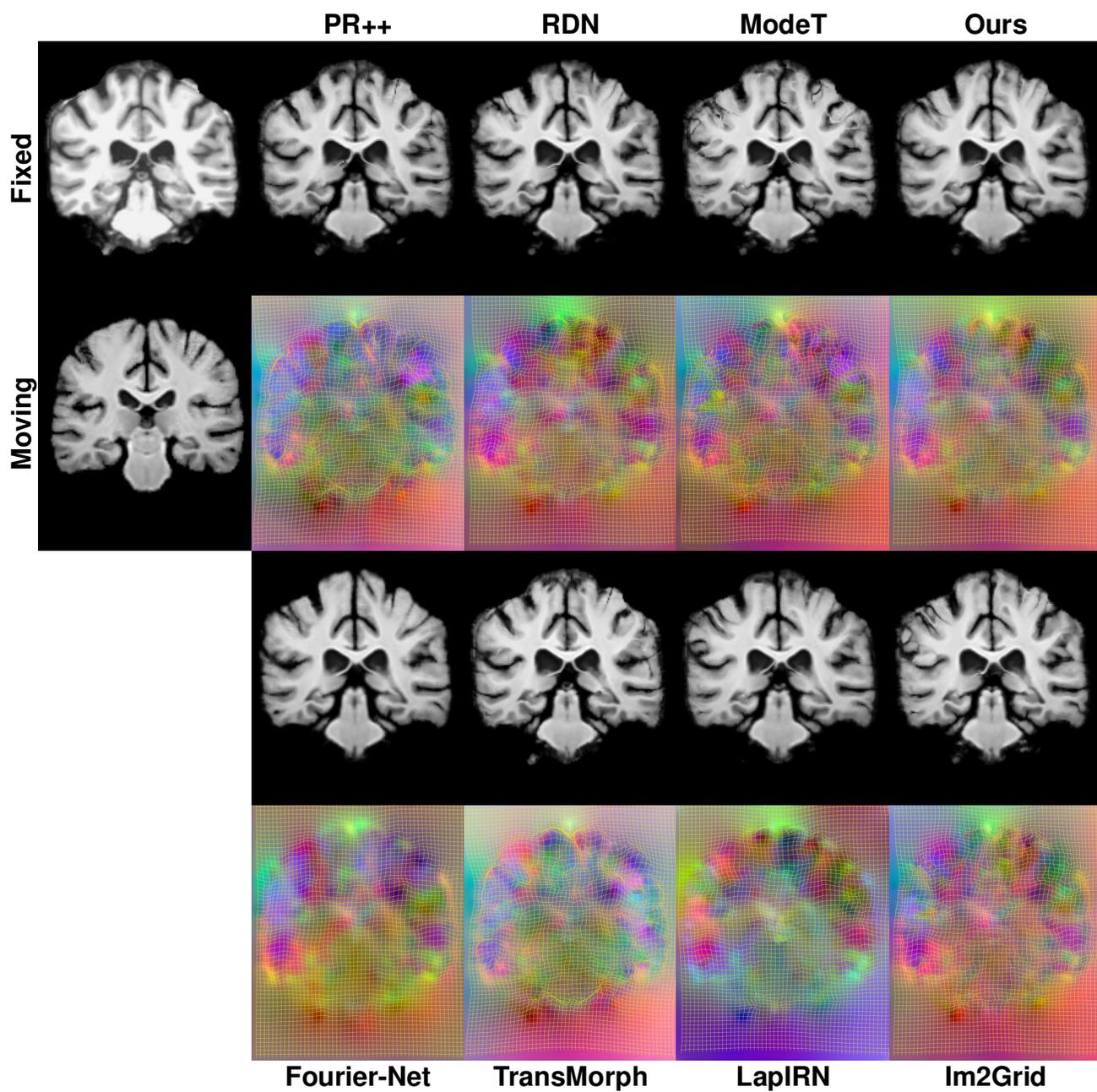


Figure 5. Visual comparisons on IXI dataset. Columns 2-5: warped moving images (top), displacement fields as RGB images (bottom).

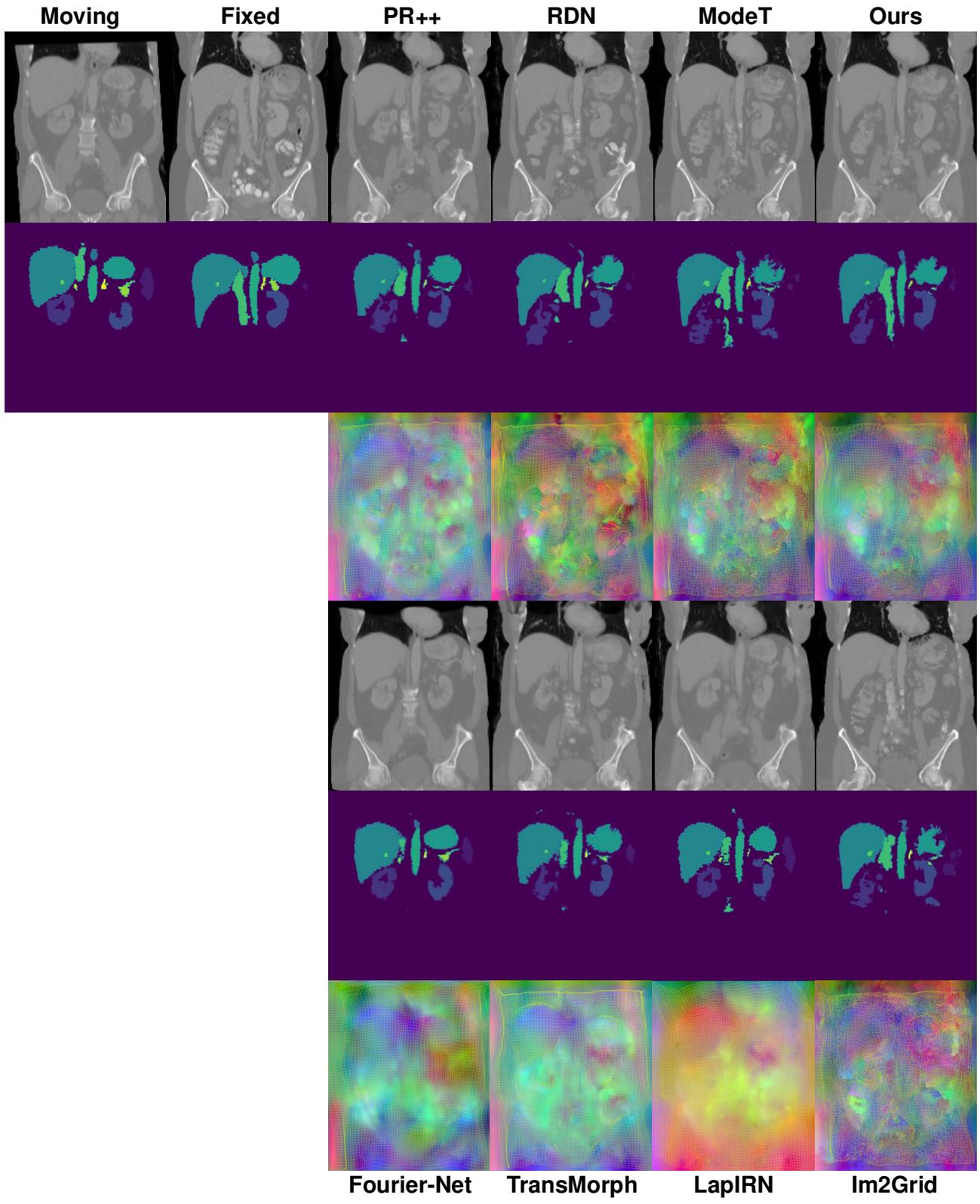


Figure 6. Visual comparisons on Abdomen CT dataset. Columns 3-6: warped moving images (top), warped moving segmentation masks (middle) and displacement fields as RGB (bottom).

References

- [1] Guha Balakrishnan, Amy Zhao, Mert R Sabuncu, John Guttag, and Adrian V Dalca. Voxelmorph: a learning framework for deformable medical image registration. *IEEE transactions on medical imaging*, 38(8):1788–1800, 2019. [1](#)
- [2] Bo Hu, Shenglong Zhou, Zhiwei Xiong, and Feng Wu. Recursive decomposition network for deformable image registration. *IEEE Journal of Biomedical and Health Informatics*, 26(10):5130–5141, 2022. [1](#)
- [3] Xi Jia, Joseph Bartlett, Tianyang Zhang, Wenqi Lu, Zhaowen Qiu, and Jinming Duan. U-net vs transformer: Is u-net outdated in medical image registration? In *International Workshop on Machine Learning in Medical Imaging*, pages 151–160. Springer, 2022. [1](#)
- [4] Xi Jia, Joseph Bartlett, Wei Chen, Siyang Song, Tianyang Zhang, Xinxing Cheng, Wenqi Lu, Zhaowen Qiu, and Jinming Duan. Fourier-net: Fast image registration with band-limited deformation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 1015–1023, 2023. [1](#)
- [5] Miao Kang, Xiaojun Hu, Weilin Huang, Matthew R Scott, and Mauricio Reyes. Dual-stream pyramid registration network. *Medical image analysis*, 78:102379, 2022. [1](#)
- [6] Yihao Liu, Lianrui Zuo, Shuo Han, Yuan Xue, Jerry L Prince, and Aaron Carass. Coordinate translator for learning deformable medical image registration. In *International workshop on multiscale multimodal medical imaging*, pages 98–109. Springer, 2022. [1](#)
- [7] Tony CW Mok and Albert CS Chung. Large deformation diffeomorphic image registration with laplacian pyramid networks. In *Medical Image Computing and Computer Assisted Intervention—MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part III 23*, pages 211–221. Springer, 2020. [1](#)
- [8] Huaqi Qiu, Chen Qin, Andreas Schuh, Kerstin Hammernik, and Daniel Rueckert. Learning diffeomorphic and modality-invariant registration using b-splines. In *Medical Imaging with Deep Learning*, 2021. [1](#)
- [9] Amin Ranem, John Kalkhof, and Anirban Mukhopadhyay. Nca-morph: Medical image registration with neural cellular automata. *arXiv preprint arXiv:2410.22265*, 2024. [1](#)
- [10] Torsten Rohlfing. Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unreliable. *IEEE transactions on medical imaging*, 31(2):153–163, 2011. [1](#)
- [11] Haiqiao Wang and Yi Ni, Dongand Wang. Modet: Learning deformable image registration via motion decomposition transformer. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023*, pages 740–749, 2023. [1](#)