IIRP-Net: Iterative Inference Residual Pyramid Network for Enhanced Image Registration

Supplementary Material

In the supplementary material, we provide additional experimental results regarding the performance of IIRP-Net and the extensibility of the iterative inference strategy. Sec. 6 offers results of extending the iterative inference strategy to more pyramid-based registration methods. Sec. 7 reports the registration results of IIRP-Net's IterStop mechanism using different similarity metrics. Sec. 8 displays the registration results and deformation fields of RP-Net, IIRP-Net, and 10 comparative methods on the FLARE and Mindboggle datasets. Sec. 9 shows the results of each iteration in the iterative inference process of IIRP-Net.

6. Iterative Inference Strategy Applied to More Pyramid Registration Methods

Table 4. The registration results of DualPRNet++, NICE-Net, PIViT, Image2Grid and RP-Net on the FLARE and Mindboggle datasets without and with the iterative inference strategy.

Method	FLARE			Mindboggle		
	Dice(%) ↑	$\lvert J_s\rvert_{\leq 0}$ ↓	t_{infer} ↓	Dice(%) ↑	$\lvert J_s\rvert_{\leq 0}$ ↓	t_{infer} ↓
DualPRNet++	58.8±13.8	<2.6%	0.13s	61.5±3.0	<0.5%	0.46s
IIDualPRNet++	68.1±9.8	<4.1%	0.30s	61.7±1.8	<0.1%	0.77s
NICE-Net	64.1±13.1	<2.9%	0.03s	61.4±1.5	<0.9%	0.10s
IINICE-Net	65.8±13.1	<5.1%	0.07s	63.2±1.4	<1.3%	0.23s
PIViT	67.3±12.8	<0.4%	**0.02**s	61.8±1.2	<0.1%	0.06s
IIPIViT	70.0±12.0	<0.5%	0.04s	62.1±1.3	<**0.03**%	0.09s
Image2Grid	67.9±10.8	<**0.2**%	0.11s	62.0±1.4	<**0.03**%	0.66s
IIImage2Grid	72.7±9.6	<0.4%	0.22s	63.9±1.2	<0.05%	0.87s
RP-Net	68.7±12.8	<1.9%	**0.02**s	64.7±1.2	<0.4%	**0.05**s
IIRP-Net	**73.3**±**11.5**	<1.5%	0.06s	**65.8**±**1.2**	<0.08%	0.12s

In the manuscript, the iterative inference strategy demonstrates the potential to enhance registration effectiveness when applied to other pyramid registration methods. Therefore, in this experiment, we implement this strategy in four pyramid registration methods for further validation: DualPRNet++, NICE-Net, PIViT, and Image2Grid. Table 4 presents the registration results of the four pyramid methods and RP-Net without and with the iterative inference strategy. Methods prefixed with "II" indicate the incorporation of the iterative inference strategy. The results demonstrate that the accuracy of each pyramid registration method improves after employing the iterative inference strategy. IIDualPRNet++ shows the most notable improvement among them on the FLARE dataset, with an increase of 9.3%, while IIImage2Grid shows the largest improvement among them on the Mindboggle dataset, with an increase of 1.9%. Notably, implementing the iterative inference strategy does not require retraining the model; it can be achieved simply by incorporating the iterative inference strategy into the inference code.

7. More Similarity Metrics in the IterStop Mechanism

Table 5. The Dice score of IIRP-Net using MAE, MSE, and PSNR as the similarity metrics for the IterStop mechanism. The numbers in parentheses indicate the differences compared to RP-Net.

Method		FLARE Dice	MindBoggle Dice
RP-Net		68.68 (-)	64.67 (-)
MAE	δ=0.005	72.14 (+3.46)	65.73 (+1.06)
	δ=0.001	73.40 (+4.72)	65.81 (+1.14)
	δ=0.0005	73.44 (+4.76)	**65.85 (+1.18)**
	δ=0.0001	73.19 (+4.51)	65.84 (+1.17)
MSE	δ=0.001	72.93 (+4.25)	65.78 (+1.11)
	δ=0.0005	73.29 (+4.61)	65.80 (+1.13)
	δ=0.0001	73.44 (+4.76)	65.84 (+1.17)
	δ=0.00005	73.36 (+4.68)	65.83 (+1.16)
NCC	δ=0.01	73.05 (+4.37)	65.61 (+0.94)
	δ=0.005	73.34 (+4.66)	65.84 (+1.17)
	δ=0.001	**73.47 (+4.79)**	65.74 (+1.07)
	δ=0.0005	73.44 (+4.76)	65.58 (+0.91)
PNSR	δ=0.5	72.18 (+3.50)	65.84 (+1.17)
	δ=0.1	73.28 (+4.60)	65.84 (+1.17)
	δ=0.05	73.33 (+4.65)	**65.85 (+1.18)**
	δ=0.01	73.20 (+4.52)	65.82 (+1.15)

In the IterStop mechanism of the iterative inference strategy, many similarity metrics can be used to determine when to stop the iteration. In our manuscript, we employ Normalized Cross-Correlation (NCC) as the similarity metric. Additionally, we explore other similarity metrics: Mean Absolute Error (MAE), Mean Squared Error (MSE), and Peak Signal-to-Noise Ratio (PSNR). The experimental results are presented in Table 5. The results show that as the threshold δ decreases, the Dice score gradually increases, but when δ is too small, the Dice score decreases. Using different similarity metrics and thresholds δ has a relatively small impact on the accuracy of IIRP-Net.

8. More Comparative Experiment Visualizations

In this section, to provide a clearer comparison, we present the registration results, post-registration labels, and defor-

mation fields of RP-Net, IIRP-Net, and all comparison methods. Figures 6 and 7 show the visualization results on the FLARE and Mindboggle datasets, respectively. Red, green, and blue labels in Figure 6 represent the liver, kidney, and spleen, respectively. In Figure 6, registration methods other than IIRP-Net fail to correctly align anatomical structures within a specific region. Specifically, in the areas marked by yellow boxes, these methods do not accurately match the spleen in the moving image to the corresponding position in the fixed image but mistakenly align it with the left kidney, resulting in suboptimal registration accuracy within that region. IIRP-Net, to some extent, corrects this misalignment. In the red boxed area of Figure 7, IIRP-Net also achieves precise alignment. In addition, the deformation field of IIRP-Net is refined based on RP-Net and remains relatively smooth.

To further illustrate the variations in registration results on anatomical structures using different methods, Figure 8 presents a box plot of registration accuracy for all methods on the FLARE dataset. IIRP-Net achieves the best results across all organs, especially in the challenging pancreas alignment. Compared to the base network RP-Net, IIRP-Net shows enhanced accuracy and robustness.

9. Iterative Inference Process Visualization

In this section, we present the visualizations of each iteration during the iterative inference process of IIRP-Net. Figures 9 and 10 illustrate the residual deformation fields $\Delta\phi_t^i$, overall deformation fields $\widetilde{\phi}_t^i$, and warped images $I_m \circ \widetilde{\phi}_t^i$ on a pair of images from the FLARE and Mindboggle datasets, respectively. "Stage i" in the figures indicates the application of the residual flow estimator f_θ^i during the i_{th} stage of the iterative inference process. For the FLARE image pair, IIRP-Net automatically determines the iteration counts for f_θ^i following the IterStop mechanism, resulting in 1, 3, 3, and 2 iterations respectively ($i \in [4, \cdots, 1]$). For the Mindboggle image pair, the corresponding iteration counts are 1, 3, 2, and 2, respectively. In the yellow and red boxed areas of Figure 9, the upper and lower edges of the spleen in I_m are better aligned during the iterative inference with f_θ^3 and f_θ^2, respectively. In the red boxed area of Figure 10, the alignment of fine structures improves with the iterative inference of f_θ^3. The visualizations in Figures 9 and 10 illustrate that iterative inference can enhance registration in areas inadequately aligned with a single inference of f_θ^i.

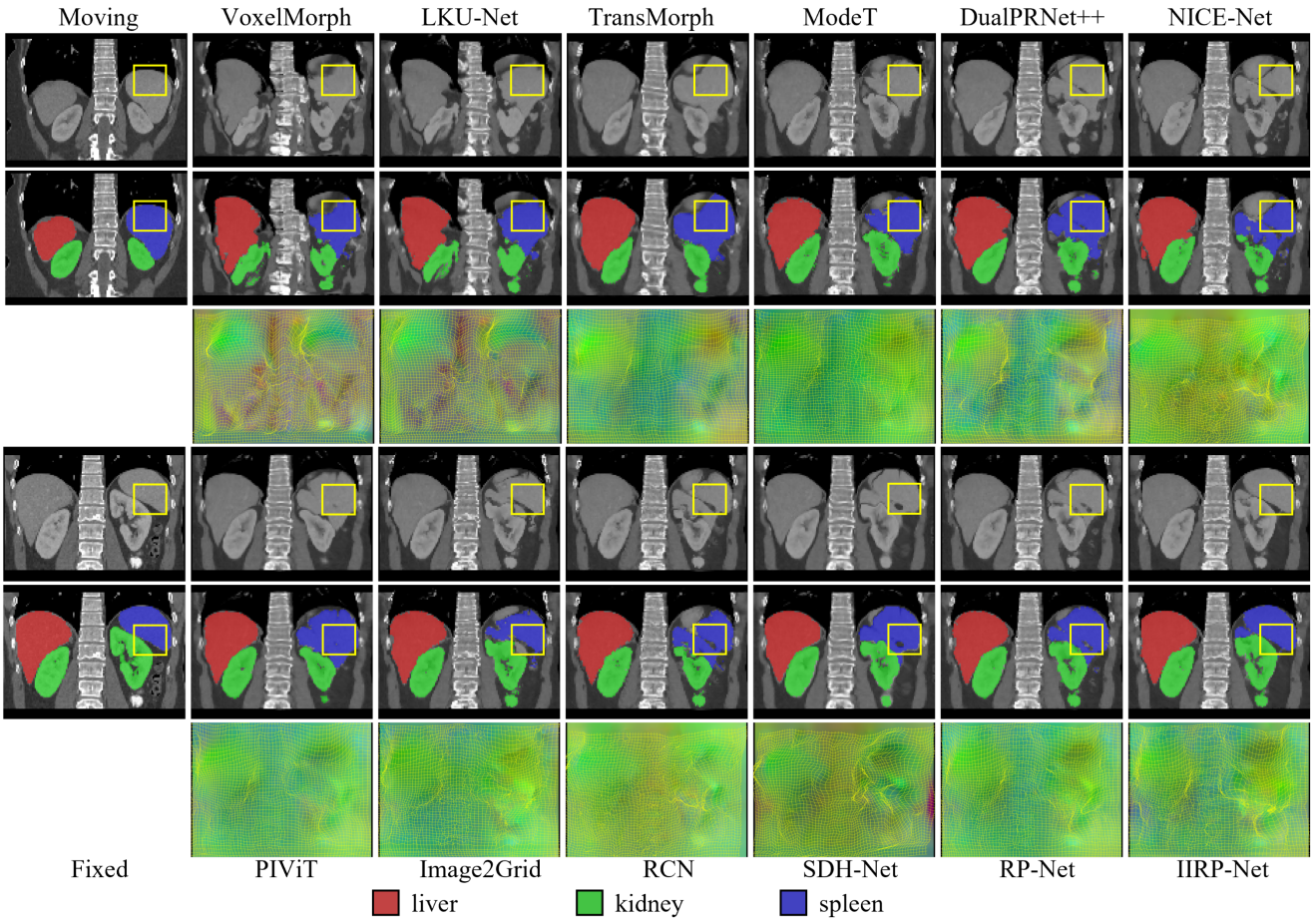

Figure 6. Example slices from the fixed images, moving images, and warped images by VoxelMorph, LKU-Net, TransMorph, ModeT, DualPRNet++, NICE-Net, PIViT, Image2Grid, RCN, SDH-Net, RP-Net and IIRP-Net on FLARE datasets. The red, green, and blue labels indicate liver, kidney, and spleen, respectively.

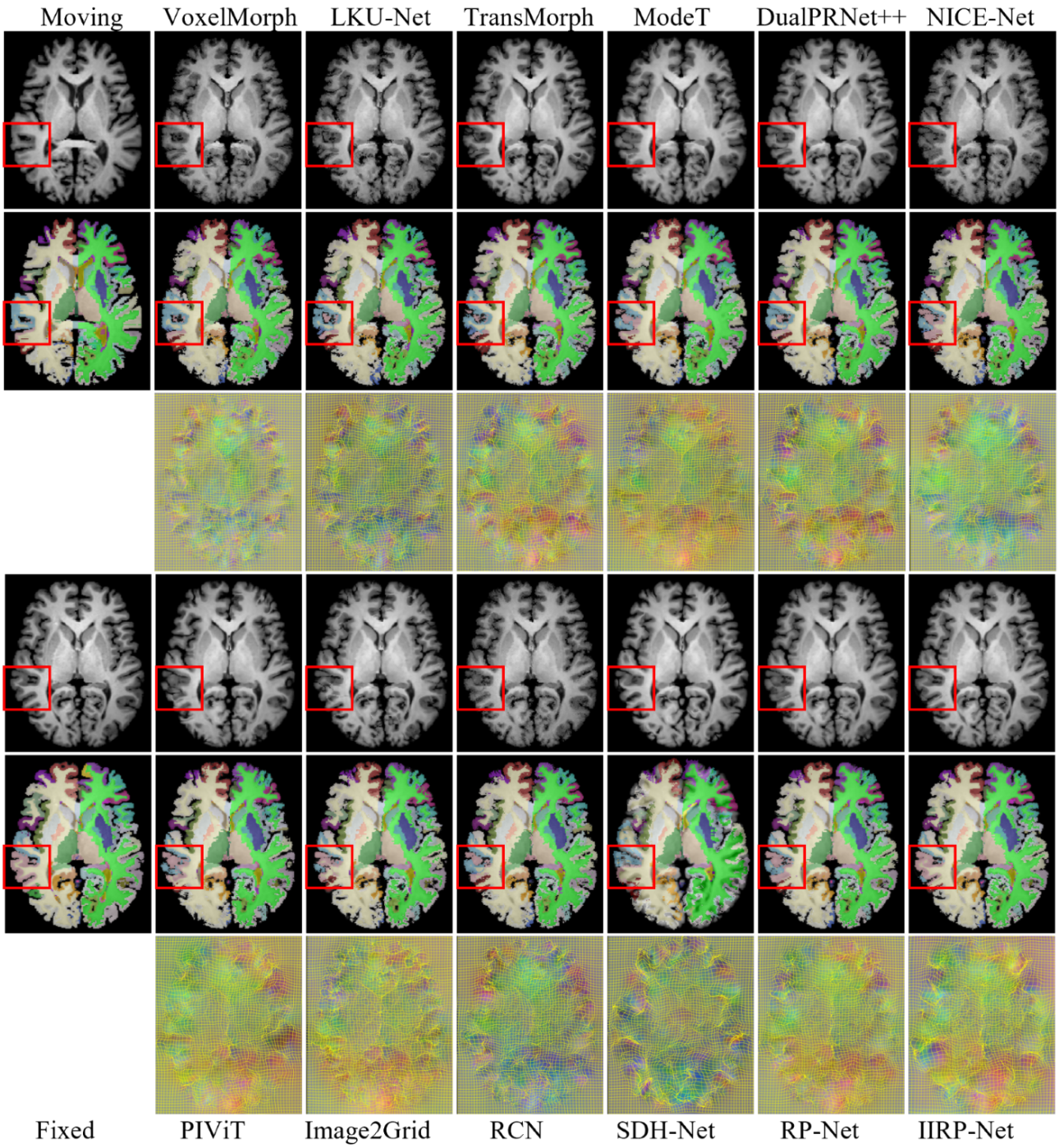

Figure 7. Example slices from the fixed images, moving images, and warped images by VoxelMorph, LKU-Net, TransMorph, ModeT, DualPRNet++, NICE-Net, PIViT, Image2Grid, RCN, SDH-Net, RP-Net and IIRP-Net on Mindboggle datasets. Red boxes highlight regions where IIRP-Net achieves precise alignment.

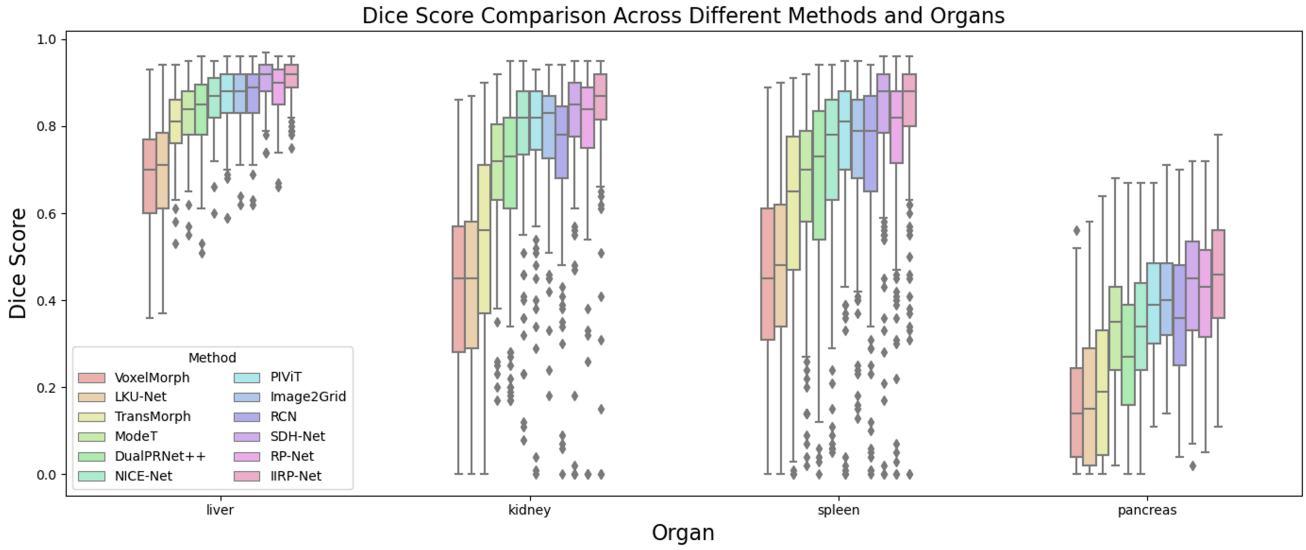

Figure 8. Box plots of Dice scores for VoxelMorph, LKU-Net, TransMorph, ModeT, DualPRNet++, NICE-Net, PIViT, Image2Grid, RCN, SDH-Net, RP-Net, and IIRP-Net on the FLARE dataset. The FLARE dataset includes four organs: liver, kidney, spleen, and pancreas.

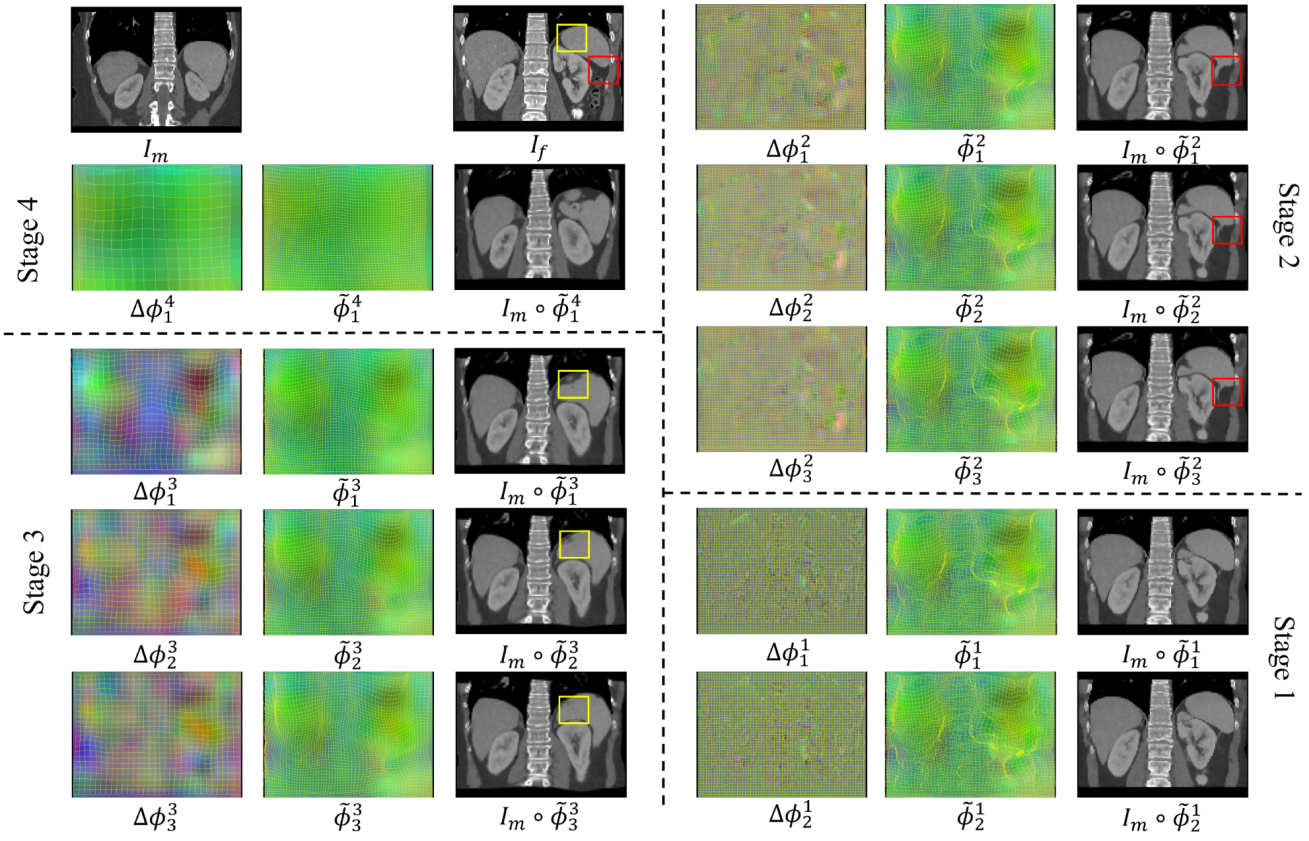

Figure 9. The process of iterative inference registration for a pair of images on the FLARE dataset using IIRP-Net.

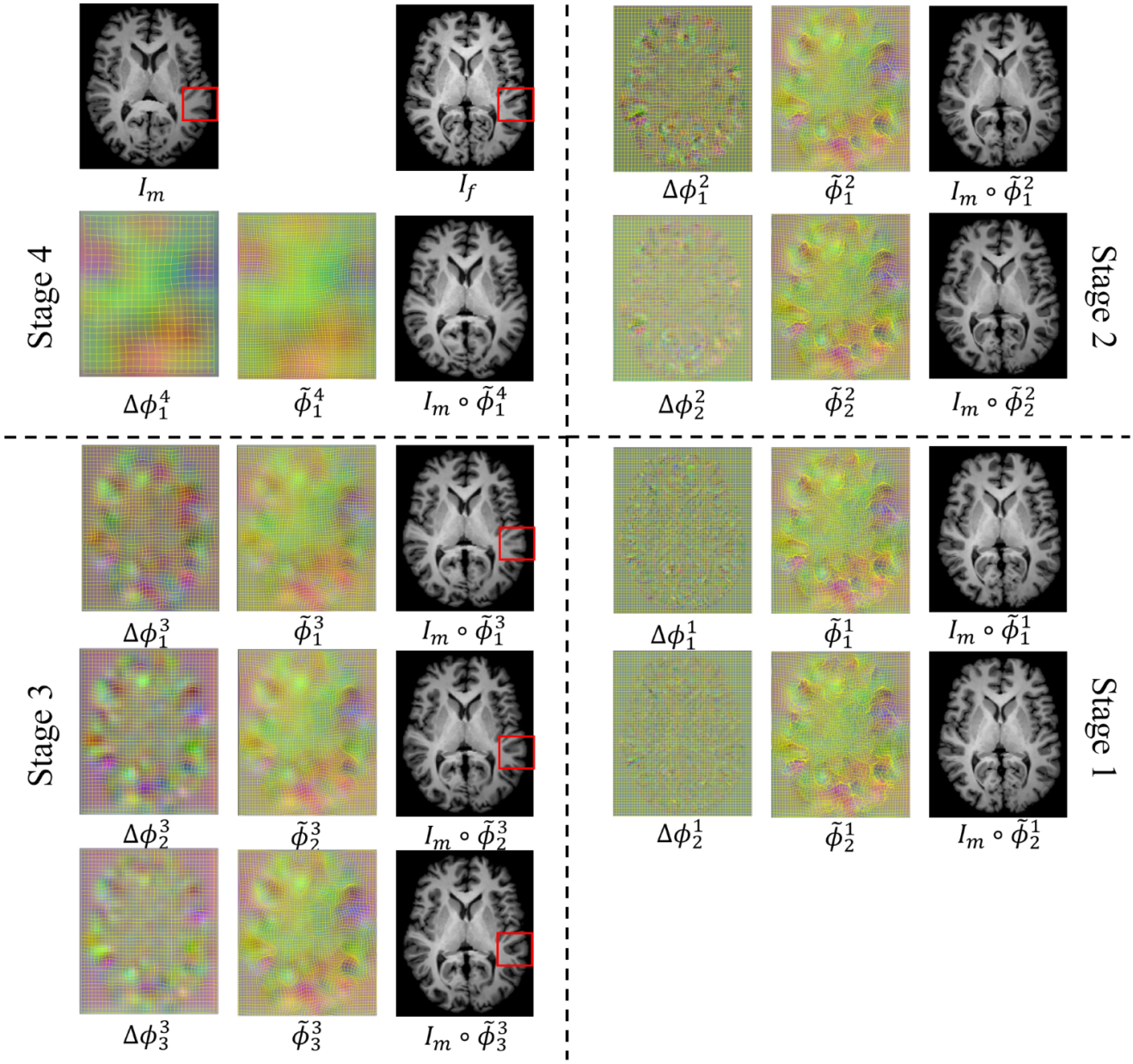

Figure 10. The process of iterative inference registration for a pair of images on the Mindboggle dataset using IIRP-Net.