Progressive and Coarse-to-Fine Network for Medical Image Registration Across Phases, Modalities and Patients

Sheng Wang[1,2], Jinxin Lv[1,2], Hongkuan Shi[1,2], Yilang Wang[1,2], Yuanhuai Liang[1,2], Zihui Ouyang[1,2], Zhiwei Wang[1,2(✉)], and Qiang Li[1,2(✉)]

[1] Britton Chance Center for Biomedical Photonics, Wuhan National Laboratory for Optoelectronics-Huazhong University of Science and Technology, Wuhan 430074, Hubei, China
{zwwang,liqiang8}@hust.edu.com

[2] MoE Key Laboratory for Biomedical Photonics, Collaborative Innovation Center for Biomedical Engineering, School of Engineering Sciences, Huazhong University of Science and Technology, Wuhan 430074, Hubei, China

Abstract. In this paper, we apply our proposed PCNet [12] on three different registration tasks assigned by the Learn2Reg challenge 2021 [1,5], i.e., CT-MR thorax-abdomen registration [3,11,14], lung inspiration-expiration registration [6] and whole brain registration [4,13], well covering three key demands in clinical practice, i.e., registration across modalities, across phases, and across patients. In these tasks, an accurate and reasonable deformation field plays a crucial role while it is often difficult to estimate in large misalignments. The core conception of our PCNet is to decompose the target deformation field into multiple sub-fields in both progressive and coarse-to-fine manners, which dramatically simplifies the direct estimation of deformation field and thus leads to a robust registration performance. The evaluation results on the three tasks demonstrate a competitive performance of PCNet and its great scalability to meet various registration demands.

Keywords: Medical image registration · Deep learning · Learn2Reg

1 Introduction

Given an image pair consisting of a fixed image and a moving image, registration requires to solve a deformation field to spatially align the fixed-moving image pair, which is a key enabling technique for varied clinical usages. The Learn2Reg challenge 2021 recently initiated three tasks representing typical clinical scenarios, i.e., CT-MR thorax-abdomen intra-patient registration (registration across modalities), CT lung inspiration-expiration registration (registration across phases) and MR whole brain registration (registration across patients). Shared by these tasks, an accurate and fast estimation of deformation field is a crucial step but challenges the academic and industrial circles.

S. Wang and J. Lv are the co-first authors.

M. Aubreville et al. (Eds.): MIDOG 2021/MOOD 2021/L2R 2021, LNCS 13166, pp. 180–185, 2022.
https://doi.org/10.1007/978-3-030-97281-3_26

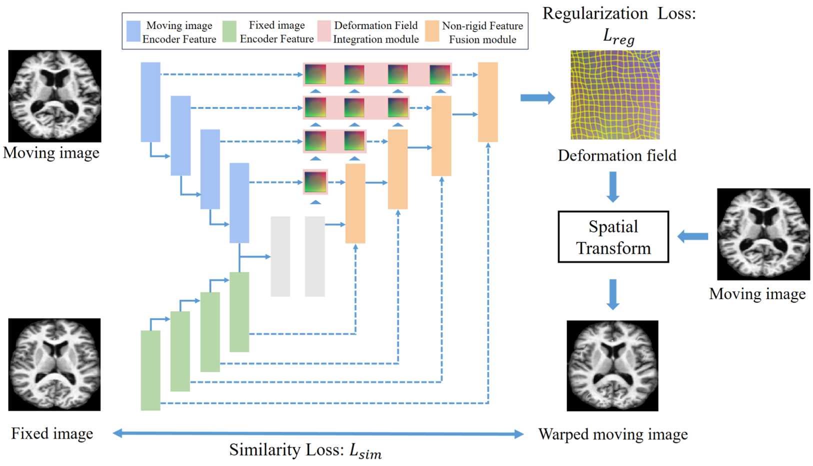

Fig. 1. The structure of proposed registration network which consist of a dual-encoder U-Net backbone, and a set of DFI and NFF modules in each decoding block.

In our recent work [12], we argued that decomposition of the target deformation field is a potential and promising solution and proposed PCNet combining the strengths of progressive registration method [15] and coarse-to-fine registration method [7]. Specifically, PCNet is built upon the backbone of dual-encoder U-net [10] and involves two key modules, i.e., deformation field integration (DFI) and non-rigid feature fusion (NFF), so as to decompose the target deformation field into several sub-fields in both progressive and coarse-to-fine manners simultaneously. By doing so, PCNet can handily predict more accurate and reasonable deformation fields with fewer parameters, and achieves competitive performance in all tasks of the Learn2Reg challenge.

2 Method

In this section, we introduce our PCNet to make this paper self-included (Sect. 2.1), and implementation details to address the three specific tasks (Sect. 2.2).

2.1 Progressive and Coarse-to-Fine Network

As shown in Fig. 1, given an unregistered pair of moving image (I_m) and fixed image (I_f), PCNet first utilizes two separate encoders to extract multi-scale features from I_m and I_f respectively. Then these multi-scale feature maps in decoding blocks is used to estimate sub-fields from coarse to fine until full-size deformation field, i.e., target deformation field are obtained. Specifically, the coarsest feature maps of I_m and I_f (gray block in Fig. 1) are combined to produce the coarsest deformation field through convolution layers. The estimation of higher level's deformation field is quite different with the lowest level. In the higher level, each decoding block sets up two key modules, named Deformation Filed Integration (DFI) and Non-rigid Feature Fusion (NFF) module to take advantage of the effort of all previous estimations and help with the next estimation.

In each decoding block, DFI module integrates sub-fields from all previous decoding blocks through warping the sub-fields progressively, and the integrated field is used to warp its corresponding feature map of I_m. We believe that the estimation of deformation field can be eased through first spatially align the feature maps of I_m and I_f. In views of this, the NFF module dynamically fuses the feature maps from three pathways, including the warped feature map of I_m, the feature map of the I_f and the up-scaled feature maps from last previous decoding block. Through NFF module, the inference of deformation field can be eased in the next decoding block. More details about these two modules can be found in [12].

2.2 Implementation Details

Learn2Reg 2021 challenge consists of 3 sub-tasks: abdominal CT-MR registration (task1), inspiration-expiration lung CT registration (task2) and whole brain MR registration (task3). According to the data characteristics and task requirements, we employ different loss functions and data processing schemes for each specific sub-task.

Loss Function. For each task, the total loss (L_{total}) basically consists of image similarity loss (L_{sim}), deformation regularization loss (L_{reg}) and weakly supervised loss (L_{weakly}) if additional segmentation labels or landmarks are provided:

$$L_{total} = \alpha L_{sim} + \beta L_{reg} + \gamma L_{weakly}, \tag{1}$$

where α, β and γ represent the weight parameters of corresponding loss respectively.

The image similarity loss (L_{sim}) calculates the similarity between the warped moving image and fixed image. We use Normalization Local Correlation Coefficient (L_{LNCC}) [12] for single-modality registration as similarity loss, and Normalization Mutual Information (L_{NMI}) [4] for multi-modality registration.

The regularization loss (L_{reg}) ensures the continuity of the deformation field based on its spatial gradients [12].

If the segmentation labels or landmarks are provided for training data, we use them as weakly supervised loss (L_{weakly}) to help the training. We construct the segmentation loss (L_{seg}) based on Dice coefficient and Cross Entropy (CE) as in [9].

Target registration error (L_{TRE}) of landmarks [8] are used as an additional weakly supervise for better alignment of small structures if landmarks are provided.

We use the same regularization loss in all tasks, while similarity loss and weakly supervised loss for each task are different. Task1 is a multi-modality registration task, thus we chose L_{NMI} as the similarity loss. The weakly supervised loss is the segmentation loss L_{seg} that calculates the misalignment of correspond region in fixed and warped moving image. For task2, we use L_{LNCC} as the similarity loss. As this task focus on the registration of the internal structure in lung, we only calculate the similarity for internal part of lungs. The TRE loss L_{TRE} is utilized for better alignment of the small structure such as pulmonary vessels. We use L_{LNCC} as the similarity loss and L_{seg} for weakly supervised loss in task3. The weight parameters $\{\alpha, \beta, \gamma\}$ are set to $\{1, 1, 1\}$ in task1, $\{1, 1, 10\}$ in task2, and $\{1, 5, 1\}$ in task3.

Data Processing and Data Augmentation. The data processing methods include twice down-sampling, window adjusting and min-max normalization. Besides, we implement random rigid and elastic deformation for some tasks to augment the limited data. Table 1 shows the implementation of data processing and data augmentation for specific sub-task.

Table 1. Data processing and data augmentation for specific sub-task.

	2× Down-sampling	Window adjusting	Min-max normalization	Random rigid & elastic deformation
Task1	√	[−170, 230]	√	√
Task2	–	[0, 1100]	√	√
Task2	√	–	√	–

3 Results

Table 2 shows the result of our method for each sub-task in Learn2Reg 2021 challenge. As shown in this table, we compared the initial score with the score after registration via our method, and the results show a great improvement in all sub-tasks. According to the official final results [2], our method ranks third

Table 2. The results of Learn2Reg 2021 challenge. This table exhibits metrics including the Dice similarity coefficient of segmentation, 95% percentile of Hausdorff distance (HD) of segmentation, target registration error (TRE) of landmarks and the standard deviation of log Jacobian determinant (SDlogJ) of the deformation field. We also exhibit the time of predicting a deformation field. The last column shows the score which integrates all these metrics.

	Dice	HD	TRE	SDlogJ	Time	Initial score	Score
Task1	0.76	17.20	–	0.13	1.90	0.26	0.78
Task2	–	–	2.70	0.10	2.70	0.25	0.59
Task3	0.80	2.00	–	0.08	2.00	0.17	0.80

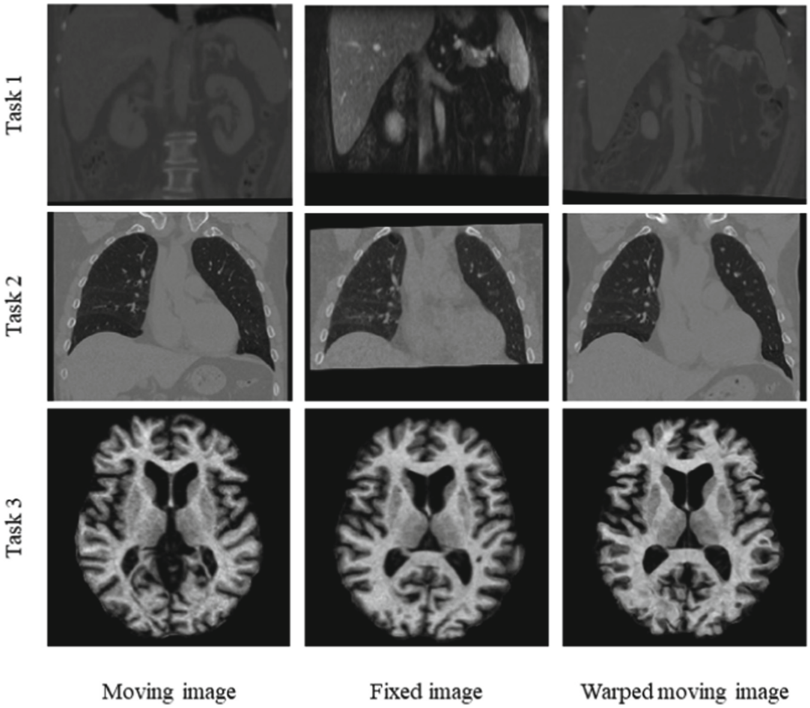

Fig. 2. Examples of registration results for task1, task2, and task3.

among 13 unofficial participating teams. Comparing with other top performance teams, our method obtains deformation field in relatively less time for all tasks. Figure 2 shows some visualization results of our method in three tasks. As shown in the figure, the warped moving image is able to match the fixed image with high accuracy in all tasks. These results indicate that our method works robustly in multiple situations, such as multi-modality registration, large misalignment registration, tiny and intricate structure registration.

4 Conclusion

In this paper, we introduce our learning-based registration network and evaluate its effectiveness on Learn2Reg 2021 challenge. Our method decomposes the target deformation field in both progressive and coarse-to-fine manners simultaneously, which enables the network accomplish accurate spatial alignment even when the displacement is quite large such as in the task2, or when the structure is complex such as in task3. The results of Learn2Reg 2021 challenge confirm that our method can achieve relatively satisfactory results in a variety of scenarios.

References

1. MICCAI registration challenge. https://learn2reg.grand-challenge.org/
2. Workshop-grand challenge. https://learn2reg.grand-challenge.org/Workshop/
3. Clark, K., et al.: The cancer imaging archive (TCIA): maintaining and operating a public information repository. J. Digit. Imaging **26**(6), 1045–1057 (2013)
4. Hoopes, A., Hoffmann, M., Fischl, B., Guttag, J., Dalca, A.V.: HyperMorph: amortized hyperparameter learning for image registration. In: Feragen, A., Sommer, S., Schnabel, J., Nielsen, M. (eds.) IPMI 2021. LNCS, vol. 12729, pp. 3–17. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-78191-0_1
5. Hering, A., et al.: Learn2Reg: comprehensive multi-task medical image registration challenge, dataset and evaluation in the era of deep learning. arXiv preprint arXiv:2112.04489 (2021)
6. Hering, A., Murphy, K., van Ginneken, B.: Learn2Reg challenge: CT lung registration - training data, May 2020. https://doi.org/10.5281/zenodo.3835682
7. Hu, X., Kang, M., Huang, W., Scott, M.R., Wiest, R., Reyes, M.: Dual-stream pyramid registration network. In: Shen, D., et al. (eds.) MICCAI 2019. LNCS, vol. 11765, pp. 382–390. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32245-8_43
8. Hu, X., Yang, J., Yang, J.: A CNN-based approach for lung 3D-CT registration. IEEE Access **8**, 192835–192843 (2020). https://doi.org/10.1109/ACCESS.2020.3032612
9. Isensee, F., et al.: NNU-Net: self-adapting framework for u-net-based medical image segmentation. arXiv preprint arXiv:1809.10486 (2018)
10. Jaderberg, M., Simonyan, K., Zisserman, A., et al.: Spatial transformer networks. Adv. Neural. Inf. Process. Syst. **28**, 2017–2025 (2015)
11. Kavur, A.E., Selver, M.A., Dicle, O., Barış, M., Gezer, N.S.: CHAOS - combined (CT-MR) healthy abdominal organ segmentation challenge data, April 2019. https://doi.org/10.5281/zenodo.3362844
12. Lv, J., et al.: Joint progressive and coarse-to-fine registration of brain MRI via deformation field integration and non-rigid feature fusion. arXiv preprint arXiv:2109.12384 (2021)
13. Marcus, D.S., Wang, T.H., Parker, J., Csernansky, J.G., Morris, J.C., Buckner, R.L.: Open access series of imaging studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. J. Cogn. Neurosci. **19**(9), 1498–1507 (2007)
14. Xu, Z., et al.: Evaluation of six registration methods for the human abdomen on clinically acquired CT. IEEE Trans. Biomed. Eng. **63**(8), 1563–1572 (2016)
15. Zhao, S., Lau, T., Luo, J., Eric, I., Chang, C., Xu, Y.: Unsupervised 3D end-to-end medical image registration with volume Tweening network. IEEE J. Biomed. Health Inform. **24**(5), 1394–1404 (2019)