A Dataset and implementation details

Dataset	Anatom. region	# of patients	# per patient	# of pairs	Type	Modality
1. COPDGene [28]	Lung	899	2	899	Intra-pat.	CT
2. OAI [26]	Knee	2532	1	3,205,512	Inter-pat.	MRI
3. HCP [34]	Brain	1076	1	578,888	Inter-pat.	MRI
4. L2R-Abdomen [35]	Abdomen	30	1	450	Inter-pat.	CT
5. Dirlab-COPDGene [5]	Lung	10	2	10	Intra-pat.	CT
6. OAI-test [26]	Knee	301	1	301	Inter-pat.	MRI
7. HCP-test [34]	Brain	32	1	100	Inter-pat.	MRI
8. L2R-NLST-val[9] [31,7]	Lung	10	2	10	Intra-pat.	CT
9. L2R-OASIS-val [21,16]	Brain	20	1	19	Inter-pat.	MRI
10. IXI-test[10]	Brain	115	1	115	Atlas-pat.	MRI
11. L2R-CBCT-val [17,18]	Lung	3	3	6	Intra-pat.	CT/CBCT
12. L2R-CTMR-val [8,1,20,11]	Abdomen	3	2	3	Intra-pat.	CT/MRI
13. L2R-CBCT-train [17,18]	Lung	3	11	22	Intra-pat.	CT/CBCT

Table 6. Listing of datasets used for training and evaluation. Datasets 1-4 are used for training `uniGradICON`. Datasets 5-7 are used for in-distribution evaluation. Datasets 8-13 are used for out-of-distribution zero-shot evaluation and assessment of fine-tuning results. We follow the official `Learn2Reg` (https://learn2reg.grand-challenge.org/) dataset split so that our results can be evaluated on the official website. According to this split, the images in the L2R-Abdomen validation set are included in the training set.

Methods	Transformation Models	Similarity Measure	Official default hyper-parameters
`SyN` [2]	Affine + SVF	Mutual Information	✓
`VoxelMorph` [3]	SVF	Mean Squared Error	✓
`LapIRN` [24]	SVF	LNCC	✓
`uniGradICON`	DVF	LNCC	✓

Table 7. Settings of the methods in the experiments. We train the universal `VoxelMorph` using MSE instead of LNCC because the model faces difficulty converging properly when using LNCC. `VoxelMorph` requires an affine pre-alignment. However, for a large training dataset, it is not feasible to compute the pre-alignment without a universal affine registration network.

B Some example registration results

Fig. 2 shows example `uniGradICON` registration results.

[9] https://www.cancerimagingarchive.net/collection/nlst/

[10] https://brain-development.org/ixi-dataset/

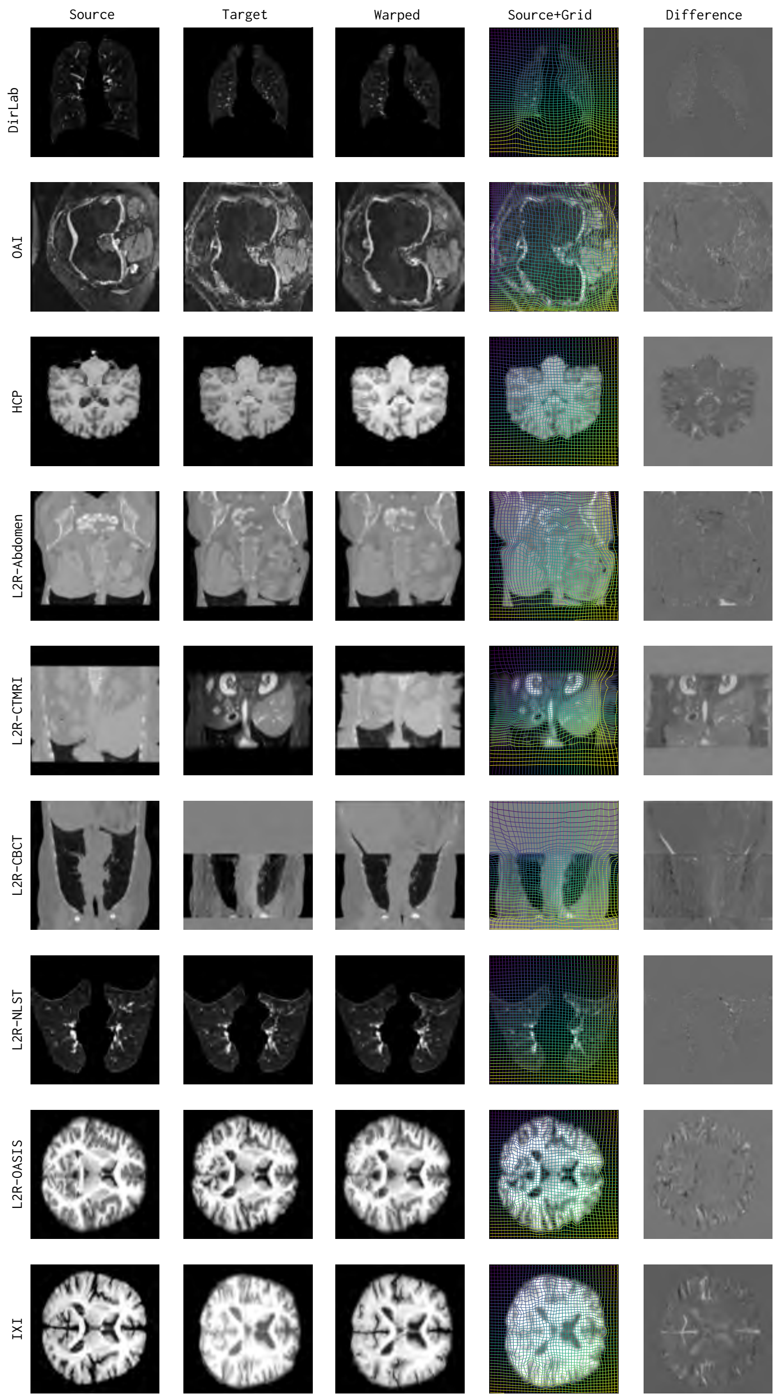

Fig. 2. Visualization of `uniGradICON` registration results for *zero-shot inference*. We display images as they are presented to our `uniGradICON` foundation model, i.e., not necessarily based on the typical anatomical convention. Note that while for a task-specific network it is important to use consistent image orientations it is much less clear that this is the case for a universal registration network, as such a network is ideally expected to be able to handle images of any orientation. In future work this could be further studied, for example, by exploring specific orientation-changing data augmentation strategies.