Variable Fraunhofer MEVIS RegLib
Comprehensively Applied to Learn2Reg Challenge

Stephanie Häger, Stefan Heldmann, Alessa Hering, Sven Kuckertz, and Annkristin Lange
Fraunhofer Institute for Digital Medicine MEVIS, Lübeck, Germany
alessa.hering@mevis.fraunhofer.de

Abstract. In this paper, we present our contribution to the learn2reg challenge. We applied the Fraunhofer MEVIS registration library RegLib comprehensively to all 4 tasks of the challenge. For tasks 1–3, we used a classic iterative registration method with NGF distance measure, second order curvature regularizer, and a multi-level optimization scheme. For task 4, a deep learning approach with a weakly supervised trained U-Net was applied using the same cost function as in the iterative approach.

Keywords: Image registration · Registration challenge · Learn2Reg

1 Introduction

Image registration is a key task in medical image analysis to estimate deformations between images and to obtain spatial correspondences. The goal of image registration is to find a reasonable deformation for a pair of fixed and moving image so that the transformed moving image and the fixed image are similar. Image registration is typically formulated as an optimization problem where a suitable cost function is minimized through iterative optimization schemes. Over time, a variety of image registration models and approaches have been developed. Therefore comparison possibilities are needed. In order to ensure comparability, challenges are created in which the different registration procedures are evaluated on the same image data and under the same computation conditions. One such challenge is the Learn2Reg: 2020 MICCAI Registration Challenge [1,3]. It consists of 4 different registration tasks that cover both intra- and inter-patient alignment, CT, ultrasound and MRI modalities, neuro-, thorax and abdominal anatomies. In this paper we present our solutions to all 4 tasks of the challenge.

2 Method and Results

Task 1–3 is solved by classical iterative methods. Task 4 is tackled by an U-Net that has been weakly supervised trained for end-to-end registration. All tasks...
build on cost functions and losses made up from several terms that are selected for the specific task. Common to all is the use of normalized gradient fields (NGF) \[5\] image similarity for fixed and moving images \(\mathbb{F}, \mathbb{M} : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}\)

\[
\text{NGF}(\mathbb{F}, \mathbb{M}) = \frac{1}{2} \int_{\Omega} 1 - \frac{\langle \nabla \mathbb{F}, \nabla \mathbb{M} \rangle_{\epsilon_F, \epsilon_M}^2}{\|\nabla \mathbb{F}\|_{\epsilon_F}^2 \|\nabla \mathbb{M}\|_{\epsilon_M}^2} \, dx
\] (1)

with parameters \(\epsilon_F, \epsilon_M > 0\), \(\langle x, y \rangle_\epsilon := x^\top y + \epsilon\) and \(\|x\|_\epsilon = \sqrt{\langle x, y \rangle_\epsilon}\) and 2nd order curvature (CURV) regularization \[4\] of displacement vector fields \(u : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3\)

\[
\text{CURV}(u) = \frac{1}{2} \int_{\Omega} \sum_{\ell=1}^{3} \|\Delta u_\ell\|^2 \, dx.
\] (2)

Furthermore, the methods for task 1–3 use a coarse-to-fine multi-level iterative registration scheme where a Gaussian image pyramid is generated for both images to obtain downsampled and smoothed images. Then, a registration is performed on the lowest resolution level and the resulting deformation field serves as the starting point for the following registration on the next highest level. This proceeds till the finest level with quasi-Newton L-BFGS optimization at each level. This procedure allows to align larger structures on the lower levels and helps to avoid local minima, to reduce topological changes or foldings, and to speed up run times.

Metrics for accuracy (TRE, DICE, Hausdorff95), robustness (DICE30) and plausibility of the deformation field (LogJacDetStd) are computed for evaluation of the challenge. More details can be found in [2]. Table 1 shows the results of our methods for all 4 tasks. The runtimes for our methods were not measured during the challenge evaluation itself, but measured afterwards performing our methods on the challenge data with a NVIDIA GeForce RTX 2070 with 8 GB memory and an Intel Core i7-9700K.

**Task 1**
The first task deals with the challenge of multimodal MRI vs. US registration of the brain. The provided data was preprocessed including a resampling to the size of 256 × 256 × 288 voxels at an isotropic 0.5 mm resolution. We focused on a parametric rigid registration due to the fact that we assume only minor local deformations inside the brain. The registration consists of two stages: an initial fast translational alignment and a multi-level registration restricted to rigid deformations. In both stages, the NGF distance measure with parameters \(\epsilon_F = 3\) and \(\epsilon_M = 2\) is minimized. In contrast to the task definition, we use the US images as fixed and the FLAIR MRIs as moving images, respectively. We changed the roles of the images in order to mask the distance measure to the significantly smaller US volume using a threshold. Afterwards, the resulting registration matrix is inverted and transformed to a corresponding displacement field. While our method successfully accomplishes all registrations on the validation set (improving the TRE from 7.02 mm to 2.75 mm on average for 7 cases), we had difficulties in some cases of the test set. This can be seen in Table 1, where a
Table 1. Results of our methods in the Learn2Reg Challenge. For the challenge the target registration error (TRE), DICE score, robustness score (30% lowest DICE of all cases, DICE30), 95% percentile of the Hausdorff distance (Hausdorff95) and the standard deviation of log Jacobian determinant of the deformation field (LogJacDetStd) were measured. We additionally measured the runtimes for our methods on a local machine.

|                      | Task 1       | Task 2       | Task 3 | Task 4 |
|----------------------|--------------|--------------|--------|--------|
| TRE before reg. [mm]  | 6.37 ± 4.34  | 10.24 ± 6.57 | -      | -      |
| TRE after reg. [mm]   | 7.02 ± 5.95  | 1.72 ± 0.59  | -      | -      |
| DICE before reg.      | -            | -            | 0.23 ± 0.21 | 0.55 ± 0.15 |
| DICE after reg.       | -            | -            | 0.47 ± 0.29 | 0.85 ± 0.04 |
| DICE30 before reg.    | -            | -            | 0.01    | 0.36   |
| DICE30 after reg.     | -            | -            | 0.21    | 0.84   |
| Hausdorff95 before reg. [mm] | - | - | 46.07 | 3.91 |
| Hausdorff95 after reg. [mm] | - | - | 43.32 | 1.55 |
| LogJacDetStd          | 0.00         | 0.07         | 0.14   | 0.05   |
| runtime [s]           | 14.74 ± 5.84 | 92.71 ± 17.22 | 0.97 ± 0.23 | 0.39 ± 0.13 |

slightly increasing TRE is shown. However, in 5 out of 10 test cases we improve the TRE from 5.84 mm to 2.86 mm on average. The partial inexact registrations could be the result of the difficult parameterization and therefore generalization of the challenging task of MRI vs. US registration and need further investigation.

Task 2
The aim of the second task was the registration of expiration to inspiration CT scans of the lung. The provided data consists of 20 training scan pairs [8] and 10 test scan pairs [7]. All scan pairs were resampled to a image size of 192 × 192 × 208 and were affine pre-registered. The main challenges are the large deformation due to breathing and that the lungs in the expiration scans are not fully visible.

Our submitted method based on our previous work [10]. First, a graph-based matching of a large number of keypoints for the estimation of robust large-motion correspondences is performed. Then, this is followed by a continuous, deformable image registration incorporating both image intensities and keypoint information. Herefore, we used the NGF distance measure with edge parameter $\epsilon = 0.1$. For a smooth deformation field we selected the curvature regularizer with weight parameter $\alpha = 1$. In contrast to [10], we are not integrating the lung mask into a cost term to enforce lung boundary alignment, because the expiration lung is not fully visible. However, we mask the NGF distance measure with the expiration lung mask. A coarse-to-fine multilevel scheme with 3 levels was applied. With a target registration error of $1.72 \pm 0.59$ mm, we archived the highest accuracy of all submissions in the challenge. The whole registration pipeline takes about 92.7 s which includes the keypoint detection with 86.4 s
Fig. 1. Example coronal slices extracted from an exemplary case for task 2: a) The expiration image, b) inspiration image, c) the difference image before the registration and d) the difference image after registration. For a better visualization, we only show the image inside the lung, however, the full thorax scan was used.

and the actual registration with 6 s. All results are summarized in Table 1. To illustrate the registration results, we show the difference images $\mathcal{F} - \mathcal{M}(y)$ before and after registration in Fig. 1. The breathing motion was successfully recovered and inner lung structures are well aligned.

Task 3
The aim of the third task was the inter-patient registration of abdominal CT scans to transfer organ segmentations. The 30 training and 20 test CT scans were provided with preprocessing such as same isotropic voxel resolutions (2 mm) and spatial dimensions $(192 \times 160 \times 256)$ as well as affine preregistration. Therefore we only use a non-parametric registration approach with the normalized gradient fields as a distance measure. The edge parameter $\epsilon$ was set to $\epsilon_F = \epsilon_M = 5$. To achieve a smooth deformation field we selected the curvature regularizer weighted with parameter $\alpha = 10$. A multilevel scheme with 4 levels was applied, where the resolution of the deformation field was one level lower than the image resolution. For all specified parameters different values were tested and the ones chosen that led to the best results regarding the DICE score on the training images. To solve the optimization problem efficiently, we used an implementation on the graphics card. This reduced the average runtime for a single registration from 25 s to less than 1 s. As shown in Table 1 the DICE score could be improved from 0.23 to 0.47 on the test CT scans. In Fig. 2 an exemplary registration result on the validation data with the overlayed segmentations is shown. The example illustrates the difficulty of interpatient registration due to the large anatomical differences. Partially the registration can compensate for this, but especially if the anatomical differences are large but the intensity changes are small, a good registration result is difficult to achieve.
Fig. 2. Exemplary registration result for task 3. The original fixed and moving images are shown in axial direction with their according labels (a, b). In (c) the fixed CT is overlayed by the deformed moving labels.

Task 4
The data for the fourth task consists of 394 MRI scans covering the hippocampus formation in the brain, divided into 263 and 131 scans for training and testing, respectively. The aim of this task is the alignment of the hippocampus head and body using inter-patient registration. For the training cases segmentations of these two structures are available and the data is additionally provided preprocessed to same voxel resolutions and spatial dimensions. In order to accomplish this task, we train a U-Net with four levels in weakly supervised manner, aiming to optimize the NGF image similarity with $\epsilon_F = \epsilon_M = 1$ and a curvature deformation regularizer including a penalty for implausible grid foldings. Additionally, we measure the alignment of the given segmentations using a sum of squared differences [6,9]. Our method receives only a fixed and a moving image as input, computes the displacement field at the same resolution and includes corresponding segmentations exclusively during training (weak supervision). Furthermore, our training is not depending on difficult to access ground-truth displacement fields. After the training of our network, only a single pass through the network is required for registration of unseen image pairs, computing displacement fields in a fraction of a second (on a GPU). As shown in Table 1, the weakly supervised training enables our method to improve the DICE score on average from 0.55 to 0.85, while maintaining physically plausible results.

3 Conclusion
We showed that the Fraunhofer MEVIS RegLib is successfully applicable to all 4 tasks of the Learn2Reg challenge that differ greatly and cover both intra- and inter-patient alignment, various modalities and anatomies. We chose the iterative method for tasks 1–3 due to the limited amount of data available and a deep learning approach for task 4. Our methods achieved the forth place in the challenge without consideration of our fast runtime. Furthermore, we achieved the overall highest registration accuracy with our method in task 2.
References

1. Learn2Reg: 2020 MICCAI registration challenge. https://learn2reg.grand-challenge.org/
2. Learn2Reg challenge, metrics and evaluation. https://learn2reg.grand-challenge.org/Submission/
3. Dalca, A., et al.: Learn2Reg - the challenge, March 2020. https://doi.org/10.5281/zenodo.3715652
4. Fischer, B., Modersitzki, J.: Curvature based image registration. J. Math. Imaging Vis. 18(1), 81–85 (2003)
5. Haber, E., Modersitzki, J.: Intensity gradient based registration and fusion of multimodal images. In: Medical Image Computing and Computer-Assisted Intervention - MICCAI 2006, vol. 3216, pp. 591–598 (2006)
6. Hering, A., Kuckertz, S., Heldmann, S., Heinrich, M.P.: Enhancing label-driven deep deformable image registration with local distance metrics for state-of-the-art cardiac motion tracking. Bildverarbeitung für die Medizin 2019. I, pp. 309–314. Springer, Wiesbaden (2019). https://doi.org/10.1007/978-3-658-25326-4_69
7. Hering, A., Murphy, K., van Ginneken, B.: Learn2Reg Challenge: CT Lung Registration - Test Data, September 2020. https://doi.org/10.5281/zenodo.4048761
8. Hering, A., Murphy, K., van Ginneken, B.: Learn2Reg Challenge: CT Lung Registration - Training Data, May 2020. https://doi.org/10.5281/zenodo.3835682
9. Kuckertz, S., Papenberg, N., Honegger, J., Morgas, T., Haas, B., Heldmann, S.: Deep learning based CT-CBCT image registration for adaptive radio therapy. In: Medical Imaging 2020: Image Processing, vol. 11313, pp. 149–154. International Society for Optics and Photonics, SPIE (2020)
10. Rühaak, J., Polzin, T., Heldmann, S., Simpson, I.J., Handels, H., Modersitzki, J., Heinrich, M.P.: Estimation of large motion in lung CT by integrating regularized keypoint correspondences into dense deformable registration. IEEE Trans. Med. Imaging 36(8), 1746–1757 (2017)