Geometric Deep Learning and Heatmap Prediction for Large Deformation Registration of Abdominal and Thoracic CT

In Young Ha¹
Lasse Hansen¹ Matthias Wilms² Mattias P. Heinrich¹
¹ Institute of Medical Informatics, University of Luebeck
² Department of Radiology, University of Calgary, Calgary, Canada

Editors: Under Review for MIDL 2019

Abstract

We propose a novel concept for supervised registration learning of large deformations. Based on ideas from human pose estimation, we design a network architecture that learns to predict discrete heatmaps for the relative displacement of a number of sparse keypoints between two scans. Graph convolutions are used to model a globally smooth transformation and learn a suitable metric to estimate sparse displacements based on dense feature representations of the volumetric scans in an end-to-end manner. An extension to weakly-supervised label-driven registration demonstrates an improvement of 19% points of overlap accuracy compared to a state-of-the-art deep learning approach.

1. Introduction

A large number of recent works have addressed medical image registration with deep convolutional networks (Balakrishnan et al., 2019; Hu et al., 2018), which enable rapid inference times and the potential of optimising challenging alignment tasks with regards to expert supervision labels. Yet, large and complex deformations remain problematic for classical feed-forward regression networks, which led to errors of >2.5 mm for inhale-exhale lung CT registration in (Eppenhof et al., 2018; de Vos et al., 2019; Sentker et al., 2018), which is inferior to conventional methods with errors of less than 1 mm (Rühaak et al., 2017).

Heatmap regression is often used for human joint detection (Bulat and Tzimiropoulos, 2016) and enables a probabilistic estimate of global keypoint locations, which can be extended to the regression of relative displacement vectors for sparse correspondences. Graph convolution networks (Defferrard et al., 2016; Bronstein et al., 2017) learn filters on irregular graphs and leverage their structural information by propagating messages along the connected edges. It is especially advantageous for tasks, which can benefit from geometric spatial relations as shown in recent works (Cucurull et al., 2018; Hansen et al., 2018).

2. Methods

Our proposed deep geometric registration framework is trained with heatmaps that are generated from sparse correspondences using conventional graphical optimisation (Heinrich et al., 2015). The network architecture consists of the following four parts:
Figure 1: Concept of proposed geometric deep learning architecture for heatmap regression of large displacements. A CNN extracts low-level 16-dim. features from images, which are sparsely sampled at keypoint locations in the fixed scan. A quantised grid of displacement is considered around each control point for the moving scan, e.g. using $17 \times 17 \times 17$ offsets. For the graph convolution, fixed and moving features are concatenated for each keypoint and each displacement. We learn to predict heatmaps that yield a globally smooth transform and are locally similar to the ground truth ones. The network may use skip connections for metric learning.

1. a convolutional feature extraction part that learns a mapping from the input intensities to 16-dimensional feature descriptors;
2. a set of keypoints that are first used to define control points for which the local similarity over all considered displacements between fixed and moving features is calculated;
3. a graph convolutional network based on a kNN-graph (between keypoints) that is used for learning a suitable metric for alignment and regularising the transformation globally using graphical message passing; and
4. two loss terms that optimise the accuracy of the predicted displacement heatmaps and the regularity based on the weighted kNN-graph using a mean-squared error.

These steps are trained together in end-to-end fashion within one network, as shown in detail in Fig. 1. Many different algorithmic choices are possible within this presented framework and we restricted our initial experiments to Obelisk feature extractors (Heinrich et al., 2019) and diffusion graph CNNs (Defferrard et al., 2016). To improve the regularity of the predicted deformations, we include a few non-trainable smoothing filters that act on the dimensions of the 3D displacement grid, which was shown to improve label compatibility in conventional graphical optimisation (Felzenszwalb and Huttenlocher, 2006; Krähenbühl and Koltun, 2011). To extend the concept of heatmap regression for tasks without one-to-one correspondences to weakly label-driven supervision, we compute pseudo heatmaps based on one-hot representations of segmentation labels.
Table 1: Quantitative cross-validation for 10 scans of the VISCERAL anatomy 3 dataset. The initial alignment for the 9 labels was 36.1%.

<table>
<thead>
<tr>
<th>Method</th>
<th>liver</th>
<th>spleen</th>
<th>pancreas</th>
<th>gallbladder</th>
<th>bladder</th>
<th>r. kidney</th>
<th>l. kidney</th>
<th>r. psoas muscle (psoas)</th>
<th>l. psoas</th>
<th>∅ (avg of labels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label-Reg</td>
<td>75.7</td>
<td>50.8</td>
<td>14.7</td>
<td>7.8</td>
<td>48.6</td>
<td>59.1</td>
<td>60.0</td>
<td>55.5</td>
<td>59.1</td>
<td>47.9%</td>
</tr>
<tr>
<td>corrField</td>
<td>73.2</td>
<td>67.1</td>
<td>17.7</td>
<td>5.9</td>
<td>48.5</td>
<td>54.3</td>
<td>70.9</td>
<td>47.2</td>
<td>59.7</td>
<td>49.4%</td>
</tr>
<tr>
<td>proposed</td>
<td>88.9</td>
<td>73.6</td>
<td>30.7</td>
<td>35.7</td>
<td>69.3</td>
<td>79.8</td>
<td>79.2</td>
<td>72.3</td>
<td>72.4</td>
<td>66.9%</td>
</tr>
</tbody>
</table>

Figure 2: Left: CT scan with ground truth labels. Middle: Same CT with segmentation of moving scan. Right: Visual outcome of our proposed deep geometric heatmap registration used to transfer the moving segmentation to match the target scan.

3. Results and Conclusion

Quantitative experiments were performed for a three-fold cross-validation for 10 scans of the VISCERAL anatomy 3 dataset (Jimenez-del Toro et al., 2016) with 9 anatomical segmentation labels: liver, spleen, pancreas, gallbladder, bladder, r. kidney, l. kidney, r. psoas muscle (psoas) and l. psoas. We consider the task of atlas-based alignment of abdominal anatomies across patients with large deformations. That means the keypoints of one scan are sampled based on the class distribution of segmented organs, while this information is unknown for the second scan. We compare our proposed approach to Label-Reg, a label-driven deep learning registration tool (Hu et al., 2018) and corrField, an MRF-based conventional registration method for keypoints (Heinrich et al., 2015). A dense deformation field was estimated for our method using a least-squares fitting based on the sparse displacement probabilities. Table 1 reports the average Dice scores across 66 registrations with a quantitative gain of more than 17% points Dice for our method. A visual example is shown in Fig. 2. The average 3D inference time is around 3 sec.

Preliminary results were also obtained for inhale-exhale COPD lung CT registration as visually shown as 3D vector rendering in Fig. 1. While a full quantitative evaluation is still ongoing, the first results demonstrate substantially reduced correspondence errors on par with conventional MRF optimisation.

Conclusion: We have presented a new concept for large deformation estimation using graph convolutions and heatmap regression. The initial results for abdominal CT are very encouraging, outperforming previous approaches by a large margin. Thoracic lung registration is considered in our ongoing research, which also includes model-based augmentation to deal with the small number of training scans (Uzunova et al., 2017).
References


Yipeng Hu, Marc Modat, Eli Gibson, Nooshin Ghavami, Ester Bonmati, Caroline M Moore, Mark Emberton, J Alison Noble, Dean C Barratt, and Tom Vercauteren. Label-driven weakly-supervised learning for multimodal deformable image registration. In *Biomedical*
Deep Geometric Heatmap Registration


