Learning Registration Models with Differentiable Gauss-Newton Optimisation

Mattias P. Heinrich[1] HEINRICH@IMI.UNI-LUEBECK.DE
[1] *Institute of Medical Informatics, University of Luebeck, Germany*

Editors: Under Review for MIDL 2022

Abstract

We propose to capture large deformations in few iterations by learning a registration model with differentiable Gauss-Newton and compact CNNs that predict displacement gradients and a suitable residual function. By incorporating a sparse Laplacian regulariser, structural / semantic representations and weak label-supervision we achieve state-of-the-art performance for abdominal CT registration.

Keywords: Learn to optimise, medical image registration, second order descent

1. Motivation

Hybrid DL-registration models successfully combine supervised networks for semantic / structural feature extraction and conventional optimisation (Hering et al., 2021). Many DL-registration models employ differentiable steps for diffeomorphic transformations and thereby incorporate spatial regularisation into the learning (De Vos et al., 2019; Dalca et al., 2019). In our SUITS framework (Blendowski et al., 2021), we combined a differentiable iterative optical flow solution with a compact CNN for semantic feature extraction. By employing a least-squares solution with a sparse Laplacian regulariser that approximates a second order descent, it can robustly capture displacements even in homogeneous regions with less than a dozen iterations. In contrast, first order optimisation e.g. Adam used in (Sandkühler et al., 2018; Mok and Chung, 2021) requires often one hundred or more iterations and is not differentiable or useable in end-to-end learning.

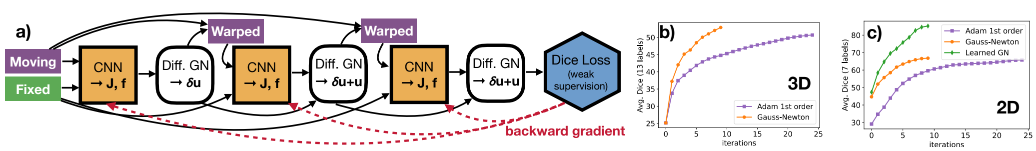

Figure 1: a) Given (structural) representations of fixed and moving scan, CNN blocks learn to predict Jacobian and function values as input for optimisation modules with differentiable Gauss-Newton (using a sparse-matrix Laplacian regulariser). b) Few iterations / warps can capture large and smooth deformations (3D validation), and c) further improved when trained with a cross-entropy loss (2D Dice).

The aforementioned approaches are limited by their reliance on pre-defined deformation and/or optimisation models. We explore a new concept to learn deformation model and

© 2022 M.P. Heinrich.

gradient computations for iterative least-square solvers. Our second order Gauss-Newton optimisation can reach performance comparable to state-of-the-art for abdominal registration in few iterations when using pre-trained semantic features and a carefully hand-tuned gradient and deformation model. In addition, we show that we can improve the performance with an automatically learned registration model in an end-to-end fashion using the proposed differentiable Gauss-Newton optimisation and reasonable initial constraints. Code and data is available at https://github.com/mattiaspaul/LearnGN.

2. Gauss-Newton for DL-based registration

Here, we aim to learn an optimisation model for pre-defined features, which are obtained from an nnUNet (Isensee et al., 2021) but could without loss of generalisation be any mapping φ from input scans to (multichannel) tensors: structural representations, corner distinctiveness, vesselness filters, contrast-invariant self-similarities, semantic features, distance maps, and more. We convert multi-label segmentations into a single signed Euclidean distance map that only distinguishes fore- and background. Next, we aim to register two abdominal scans across subjects as laid out for Task 3 of Learn2Reg 2020 (Hering et al., 2021). Following Gauss-Newton least-squares optimisation[1] we may find an approximate and unique minimiser for $\frac{1}{2}\mathbf{f}(\mathbf{x})^T\mathbf{f}(\mathbf{x})$ by solving for $(\mathbf{J}^T\mathbf{J})\delta\mathbf{u} = -\mathbf{J}^T\mathbf{f}$ in few iterations with $\mathbf{u} := \mathbf{u} + \alpha\delta\mathbf{u}$. Where $\mathbf{J}$ is the Jacobian (displacement grid gradients), $\mathbf{f}$ the function value (difference between warped and fixed representation) and α=1. A diffusion regulariser based on the sparse Laplacian is employed, which requires the solution of a sparse linear system to avoid inverting $(\mathbf{J}^T\mathbf{J})$ and enables the filling in of displacements into homogeneous regions. In our experiments, we perform 45 inter-subject registrations using a coarse grid with stride 4 and the same automatic distance maps. Adam optimiser with 25 iterations[2] reaches 50.7% validation Dice across 13 (small) anatomies and is outperformed in only 10 iterations by the proposed second-order Gauss-Newton with D =52.9% (4th rank /13 in challenge based on overlap). Both improve over supervised VoxelMorph (D =43.9%).

3. Differentiable GN for learning registration models

As detailed in (Blendowski et al., 2021) the derivative of each Gauss-Newton step can also be computed by solving another sparse equation system, which can be drastically speeded-up using GPUs: `cupyx.scipy.sparse.linalg.cg`, leaving us with a choice for defining (spatial) gradients $\mathbf{J}$ of the warped features with respect to a (coarse) deformation and the (residual) function values $\mathbf{f}$. A reasonable choice are finite difference stencils and smoothing kernels but our end-to-end training enables us to further fine-tune them using a compact trainable CNN. For this proof-of-concept, we restricted ourselves to two-layer CNNs and used only 2D coronal CT slices of the abdomen and their respective semantic features (with 7 visible labels). Architectural details can be found in our source-code. Nevertheless, the extension to 3D is straightforward and the run time for each training iteration would be less than a second. As shown in Fig. 1 multiple iterative warps and CNN-based gradient predictions can be concatenated - with differentiable Gauss-Newton

1. http://www2.imm.dtu.dk/pubdb/edoc/imm3215.pdf
2. https://github.com/multimodallearning/convexAdam

optimisation steps interleaved and being trained in an end-to-end fashion. We employ Euclidean distance maps as inputs, a weighted CE-loss for supervision (training for about 20 epochs) and initialised the network with the same kernels that are used for our conventional Gauss-Newton optimisation. As before, Gauss-Newton requires fewer iterations as first-order descent (Adam) for comparable accuracy. Fig. 1 c) demonstrates a clear advantage of our fine-tuning - raising the Dice score from below 70% to over 80% for the same hold-out validation cases. Further loss terms could be easily included.

4. Conclusion and Outlook

We have presented a promising new concept for learning gradients for second-order optimisation steps in medical image registration. Our results for 3D and 2D multi-organ abdominal CT registration demonstrate the usefulness of Gauss-Newton optimisation in general and the end-to-end fine-tuning CNN predictors when backpropagating through several iterations. In future work, we will perform more extensive validation, e.g. including the use of unsupervised feature representations and other registration tasks where fast instance optimisation has great benefit, such as inspiration-exhale lung alignment.

References

Max Blendowski, Lasse Hansen, and Mattias P Heinrich. Weakly-supervised learning of multi-modal features for regularised iterative descent in 3d image registration. *Medical image analysis*, 67:101822, 2021.

Adrian V Dalca, Guha Balakrishnan, John Guttag, and Mert R Sabuncu. Unsupervised learning of probabilistic diffeomorphic registration for images and surfaces. *Medical image analysis*, 57:226–236, 2019.

Bob D De Vos, Floris F Berendsen, Max A Viergever, Hessam Sokooti, Marius Staring, and Ivana Išgum. A deep learning framework for unsupervised affine and deformable image registration. *Medical image analysis*, 52:128–143, 2019.

Alessa Hering, Lasse Hansen, Tony CW Mok, Albert Chung, Hanna Siebert, Stephanie Häger, Annkristin Lange, Sven Kuckertz, Stefan Heldmann, Wei Shao, 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.

Fabian Isensee, Paul F Jaeger, Simon AA Kohl, Jens Petersen, and Klaus H Maier-Hein. nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature methods*, 18(2):203–211, 2021.

Tony CW Mok and Albert Chung. Conditional deep laplacian pyramid image registration network in learn2reg challenge. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 161–167. Springer, 2021.

Robin Sandkühler, Christoph Jud, Simon Andermatt, and Philippe C Cattin. Airlab: autograd image registration laboratory. *arXiv preprint arXiv:1806.09907*, 2018.