Fraunhofer MEVIS

OPTIMIZATION OF DEFORMABLE IMAGE REGISTRATION WITH CONVOLUTIONAL NEURAL NETWORKS

Hering A.[1], Heldmann S.[1]

[1]Fraunhofer Institute for Medical Image Computing MEVIS, Lübeck

Abstract

We propose a CNN-based registration method in the context of follow-up analysis for thoracic CT scans. In contrast to conventional registration approaches, we do not employ iterative optimization schemes at runtime on each image pair. Instead, we invest computational time to suitably pretrain a CNN. Therefore, at inference, given a new pair of images, only one forward pass through the network yields the desired displacement field much faster.

Material and Methods

Deformable image registration

A variational image registration objective function was used as the loss $\mathcal{L}$ of a deep learning approach for deformable thoracic CT registration:

$$\mathcal{L} = \mathcal{D}(\mathcal{F}, \mathcal{M}(y)) + \alpha \mathcal{S}(y) \to \min$$

with sum of squared differences [1] as image distance measure

$$\mathcal{D}(\mathcal{F}, \mathcal{M}(y)) := \int_\Omega (\mathcal{M}(y(\mathbf{x})) - \mathcal{F}(\mathbf{x}))^2 d\mathbf{x}$$

for fixed image $\mathcal{F}$ and moving image $\mathcal{M}$ and diffusion regularization [2]

$$\mathcal{S}(y) := \frac{1}{2} \sum_{l=1}^{3} \int_\Omega \|\nabla u_l(\mathbf{x})\|^2 \, d\mathbf{x}$$

yielding the deformation $y(x)$ with $y(x) = x + u(x)$.

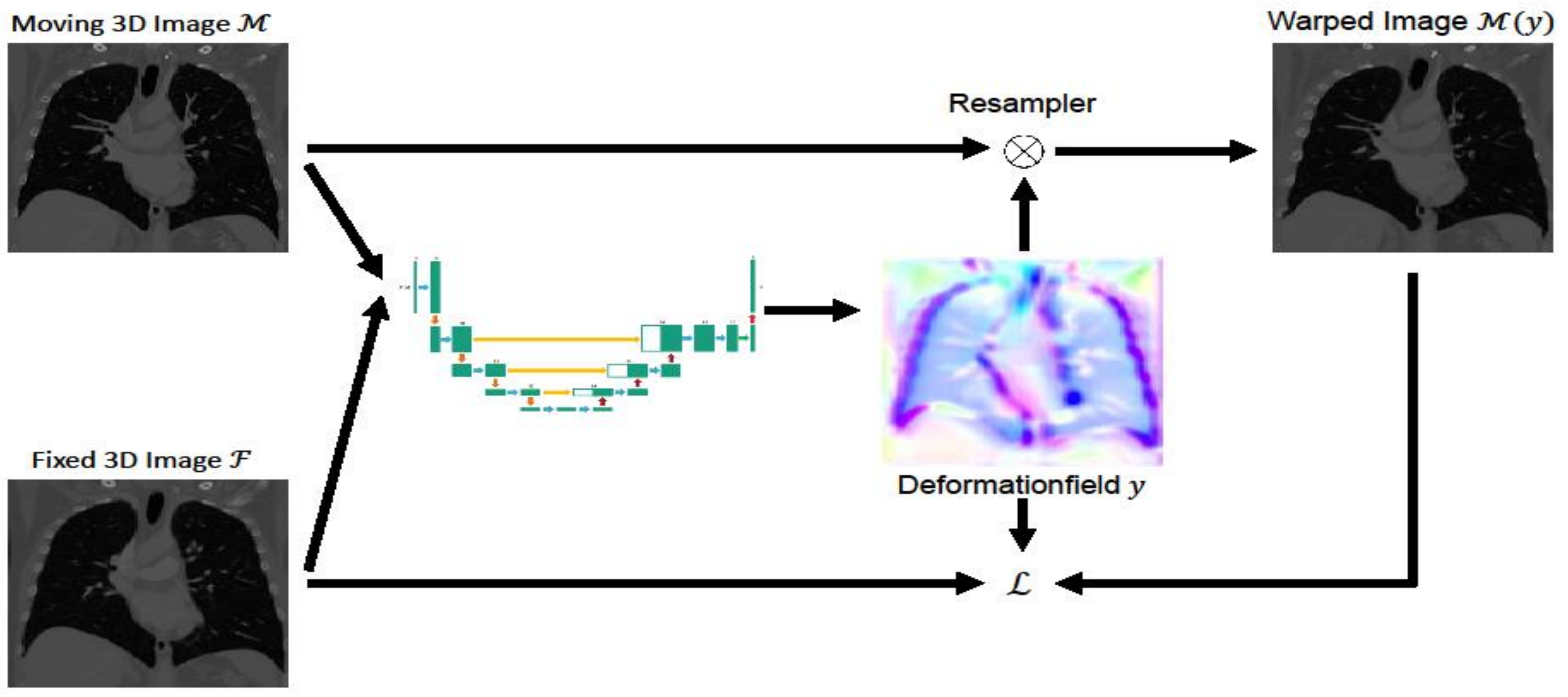

Figure 1: Our proposed registration pipeline for deformable thorax registration.

Dataset

Our dataset consists of 968 thorax CT follow-up image pairs from 487 different patients. For each image a segmentation of the lung is available. We divided the dataset into training, validation and test set with 652, 163 and 153 images, respectively. All scans were resampled to a 128x128x256 grid.

Network Architecture

The employed CNN is based on the U-Net architecture [3] having the following features:

- Receptive field of 80^3 voxels
- Five resolution levels comparable to multi-resolution image registration

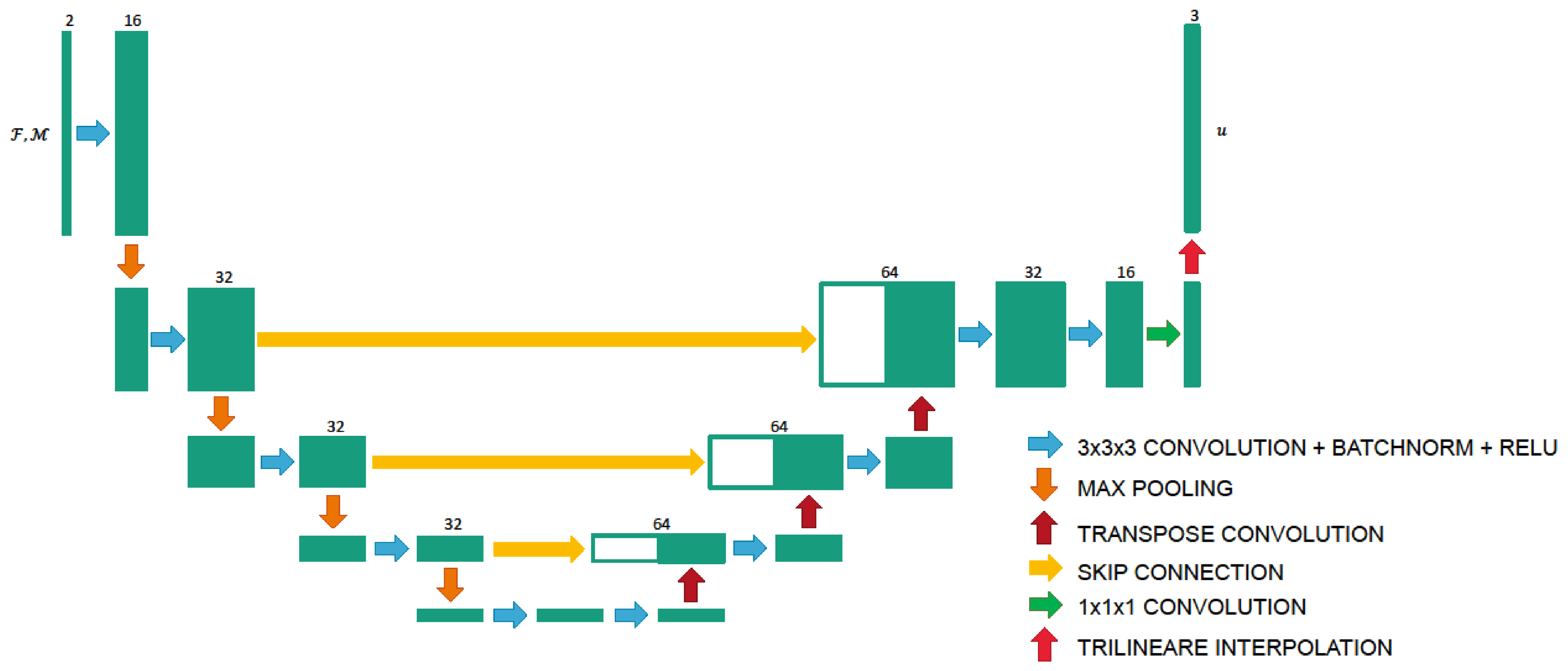

Figure 2: Overview of the U-Net based architecture with five levels. The numbers on top of layers correspond to the channel count.

Preliminary Results

Our method reduced the distance measure on average from $1.2 \cdot 10^5$ to $0.3 \cdot 10^5$ and improved the dice coefficient of the lung segmentation from 0.82 to 0.93.

The method needs on average 0.17s per case for a deformable image registration.

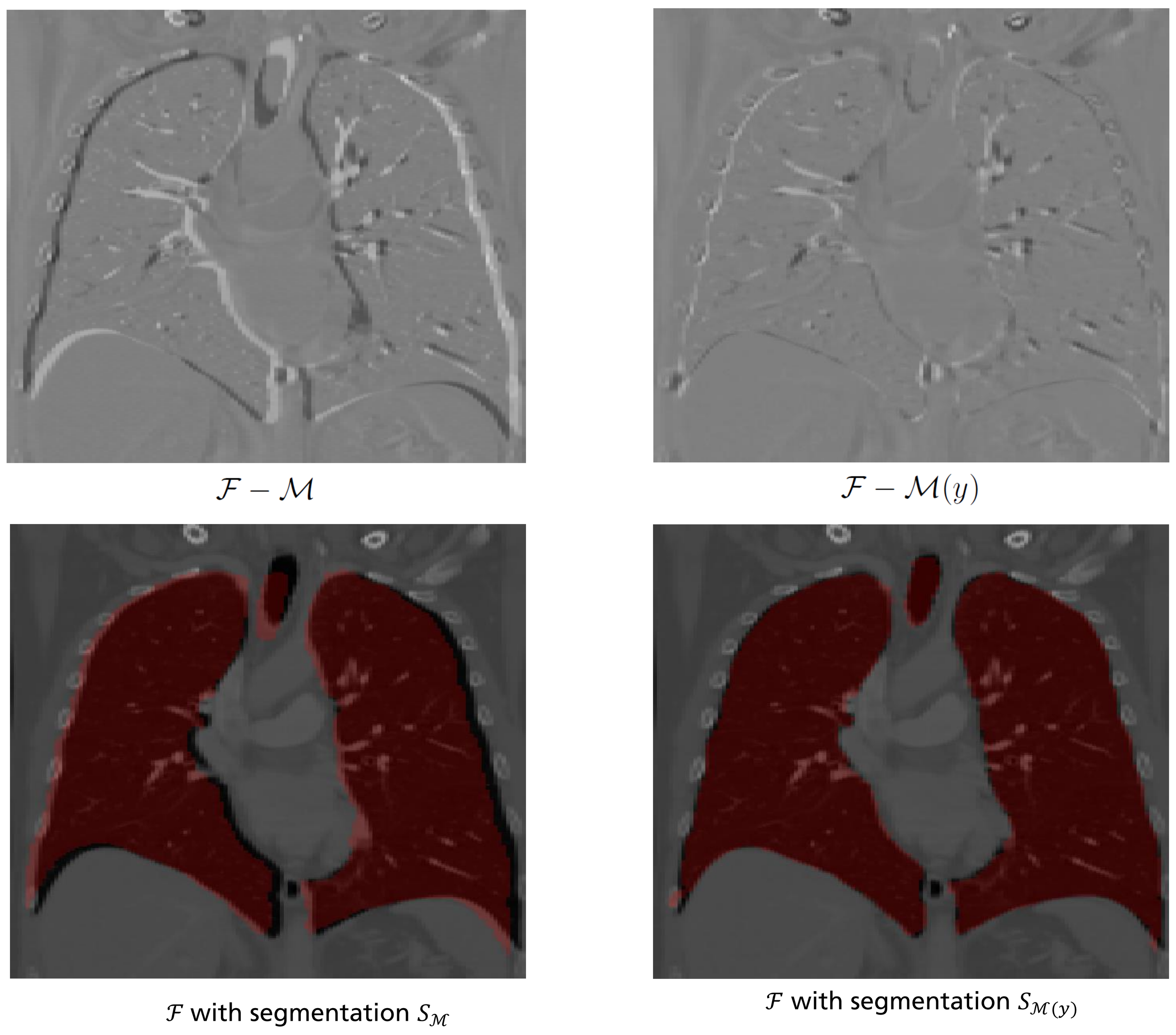

Figure 3: Exemplary results of our registration method. First row: difference of fixed with moving and moved image. Second row: Fixed image with the segmentation of the moving and moved image.

References

[1] Rühaak, J et al.(2017). A matrix-free approach to efficient affine-linear image registration on CPU and GPU. *Journal of Real-Time Image Processing*, *13*(1), 205-225.
[2] Maintz, J. A., & Viergever, M. A. (1998). A survey of medical image registration. Medical image analysis, 2(1), 1-36.
[3] Fischer, B., & Modersitzki, J. (2002). Fast diffusion registration. Contemporary Mathematics, 313, 117-128.
[4] O. Ronneberger et al. "U-net: Convolutional Networks for Biomedical Image Segmentation." In Proc. of MICCAI (2015).

Contact: alessa.hering@mevis.fraunhofer.de
www.mevis.fraunhofer.de