Deep Patient Motion Estimation: Pretraining, Overfitting, or Pretraining and Overfitting?

Timothy Jay Herbst[1,2], Joscha Maier[1], Marcel Arheit[3], Pascal Paysan[3], and Marc Kachelrieß[1,2]

[1]Divison of X-Ray Imaging and CT, German Cancer Research Center (DKFZ), Heidelberg, Germany

[2]Ruprecht-Karls-University Heidelberg, Heidelberg, Germany

[3]Varian Medical Systems imaging Lab, Gmbh, Baden-Dätwill, Switzerland

Corresponding author: marc.kachelriess@dkfz.de

dkfz. GERMAN CANCER RESEARCH CENTER IN THE HELMHOLTZ ASSOCIATION

Research for a Life without Cancer

Introduction

- During a CBCT scan, which can take up to a minute, a patient is likely to move causing motion artifacts (e.g. blurring).
- If the motion is known, one can compensate for the motion artifacts.
- Several approaches, such as Demons [1] or DEEDS [2], [3] algorithms, exist to determine the deformation vector field (DVF) between two time points.
- A neural network approach called VoxelMorph [4], had been proposed to register 3D magnetic resonance brain scans onto each other.
- We propose StdRegNet, an adaptation of VoxelMorph, to register one breathing phase of a patient onto another (see Fig. 1).

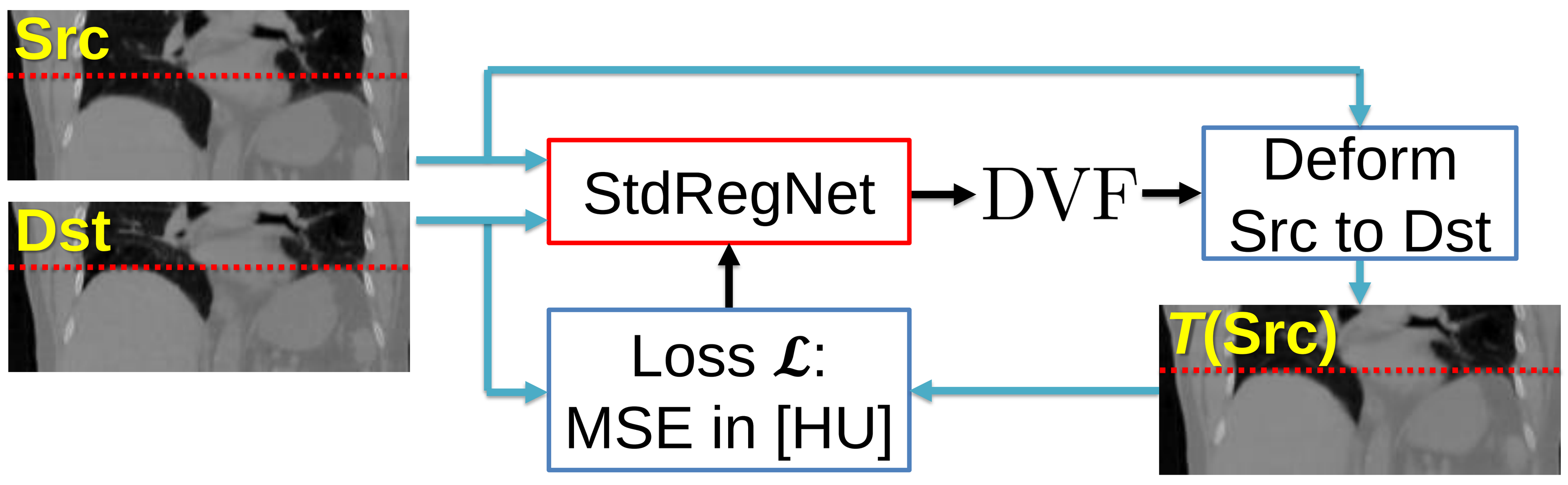

Fig. 1: Method for training StdRegNet.

- StdRegNet is trained to minimize the difference between the transformed volume T(Src) and the destination volume Dst in a voxel-wise manner while penalizing strong gradients of the DVF:

$$\mathcal{L} = ||\mathrm{Dst} - T(\mathrm{Src})||_2^2 + \lambda^2 ||\nabla \mathrm{DVF}||_2^2,$$

with λ = 410 HU being a regularization parameter.

Material and Methods

- 77 CT patient datasets (63 training, 14 test datasets) containing ten volumes of the patient in different phases of the respiratory cycle.
- Simulate CBCT data from these 770 CT volumes based on Varian TrueBeam™ geometry (see Fig. 2)
- Detector pixels were downsampled by a factor of four to accommodate the GPU memory requirements.
- Reconstructed into volume of size 224×224×128 voxels of size 2×2×2 mm^3.

Figure 2: Varian TrueBeam™ geometry: R_F = 100 cm, R_{FD} = 150 cm, N_{360} = 656 projections per full rotation, shifted detector, 1024×768 detector pixels of size 0.388×0.388 mm^2, t_{rot} = 60 s rotation time

StdRegNet was trained in three different ways:

- Pretrained on the training data before being applied to the test data (two phases of same test patient)
- Overfitted (no pretraining) onto the test data
- Pretrained on training data before being overfitted to the test data

Results

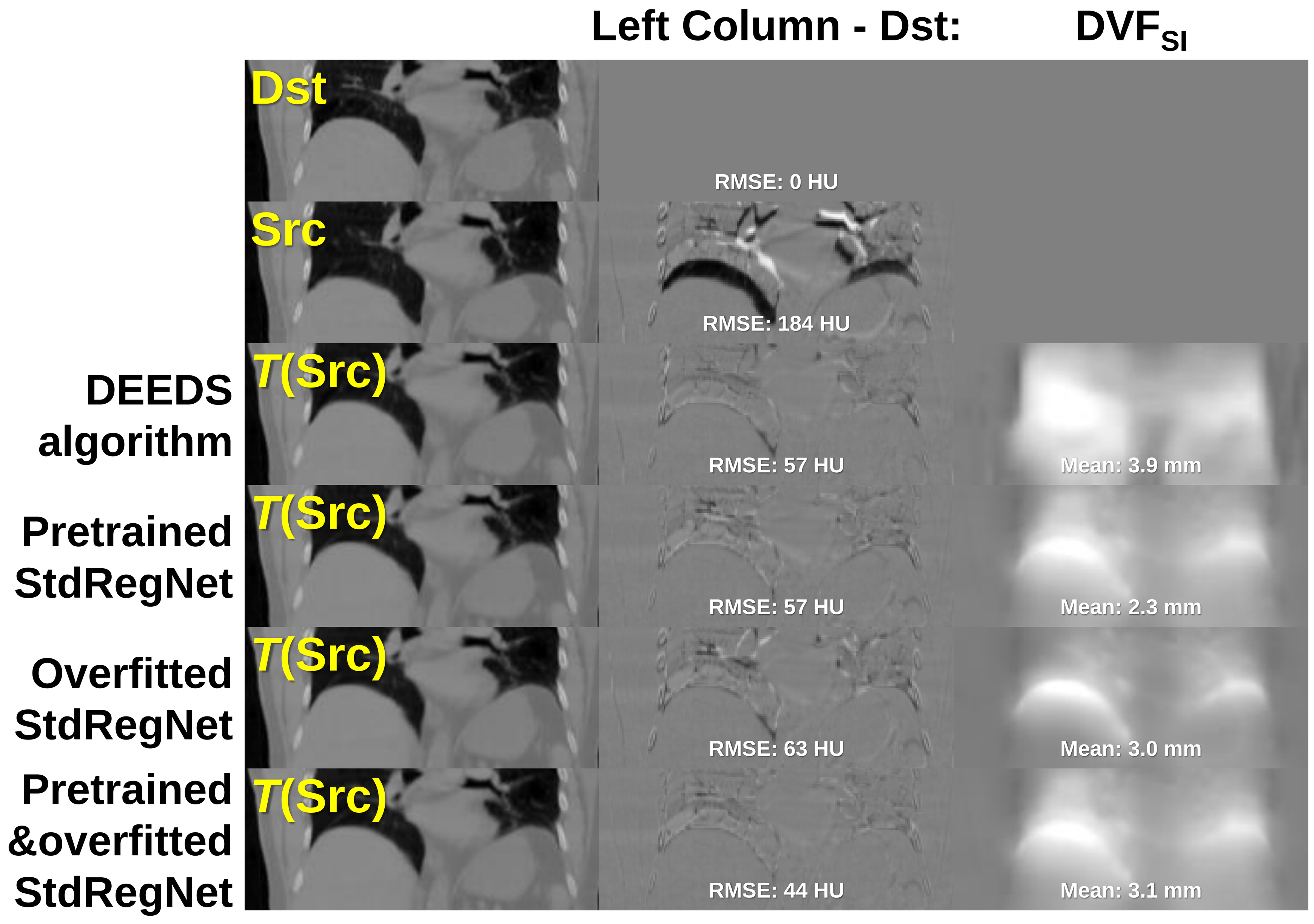

Fig. 3: Registering end-inhale (Src) volume onto end-exhale (Dst) volume of a test patient using DEEDS and StdRegNet trained in three different ways. The left column shows the transformed Src, the middle column shows the difference to the Dst and the right shows the estimated DVF along the superior-inferior axis. Left and middle: C = 0 HU, W = 2000 HU, right: C = 0 mm, W = 40 mm.

The difference between the transformed Src and the Dst and mean DVF is measured in an ROI within the patient volume (as determined by a segmentation algorithm). Table 1 shows the differences for registering end-inhale onto end-exhale averaged over all test patients.

RMSE Difference of T(Src) with Dst using	DVF=0, T(Src)= Src	DEEDS	Pretrained	Overfitted	Pretrained Overfitted
Patient Fig. 3	184 HU	57 HU	57 HU	63 HU	44 HU
Average over all test patients	(134 ± 20) HU	(48 ± 4) HU	(41 ± 6) HU	(55 ± 6) HU	(35 ± 4) HU

Table 1: RMSE averaged over all 14 test patients for registering end-inhale onto end-exhale. All RMSE are measured only within the patient volume as determined by a simple segmentation algorithm.

Conclusions

StdRegNet works to register two different phases of the same patient onto each other with a quality comparable to DEEDS. Pretrained and overfitted StdRegNet produce results of comparable quality to DEEDS. StdRegNet performs best if pretrained before overfitting the target data.

Acknowledgment

This study was supported in parts by Varian Medical Systems, a Siemens Healthineers Company, and by the Society of High Performance Computational Imaging (SHPCI) e.V. Parts of the reconstruction software were provided by RayConStruct® GmbH, Nürnberg, Germany.

[1] J.-P. Thirion, Medical Image Analysis, vol. 2, no. 3, pp. 143-260, 1998.
[2] M. P. Heinrich, et al., IEEE Transactions on Medical Imaging, vol. 32, no. 7, pp. 1239-1248, 2013.
[3] M. P. Heinrich, et al., ISBI, New York, NY, USA, 2015.
[4] G. Balakrishnan, et al., IEEE Transactions On Medical Imaging, vol. 38, no. 8, pp. 1788-1800, 2019.