AAPM 2022
JULY 10–14 | WASHINGTON, DC
64TH ANNUAL MEETING & EXHIBITION

# Deep Learning-Based Simultaneous Multi-Phase Deformable Image Registration of Sparse 4D-CBCT

**I. HERZIG**[1], P. PAYSAN[2], S. SCHEIB[2], A. ZUEST[1], F-P. SCHILLING[3], J. MONTOYA[3], M. AMIRIAN[3], T. STADELMANN[3], P. EGGENBERGER[1], R. FÜCHSLIN[1], L. LICHTENSTEIGER[1]
1 Zurich University of Applied Sciences ZHAW, Institute for Applied Mathematics and Physics IAMP, Switzerland
2 Varian Medical Systems Imaging Laboratory, Daettwil AG, CH
3 Zurich University of Applied Sciences ZHAW, Centre for Artificial Intelligence CAI, Switzerland

Zurich University of Applied Sciences
zh aw
varian
A Siemens Healthineers Company

## INTRODUCTION

Respiratory gated 4D-CBCT suffers from sparseness artefacts caused by the limited number of projections available for each respiratory phase. These artefacts severely impact traditional deformable image registration methods used to extract motion information. We use a supervised deep learning method that is able to predict displacement vector-fields (DVF) from sparse 4D-CBCT despite the presence of artefacts.

## MATERIALS & METHODS

### Motion Simulation for Data Generation

Training data is generated by a motion simulation framework based on Ref. [1] but extended to use multiple phases. The simulation uses respiratory phase gated 4D-CT scans and a collection of recorded breathing curves.

*Preprocessing 4D-CT:* Deformable image registration (Deeds [3]) between the 10 consecutive breathing phases creates a set of 10 patient-specific cyclic DVFs.

*Simulated 4D-CBCT Scan:*

1. Select a combination of breathing curve and 4D-CT scan as inputs to the simulation. This enables data-augmentation by producing multiple simulated scans based on the same 4D-CT using different breathing curves.
2. For each x-ray acquisition, deform the max-exhale 4D-CT volume according to the curves' amplitude at this point using the previously generated DVFs.
3. Forward project the deformed volumes to simulate the x-ray acquisition.

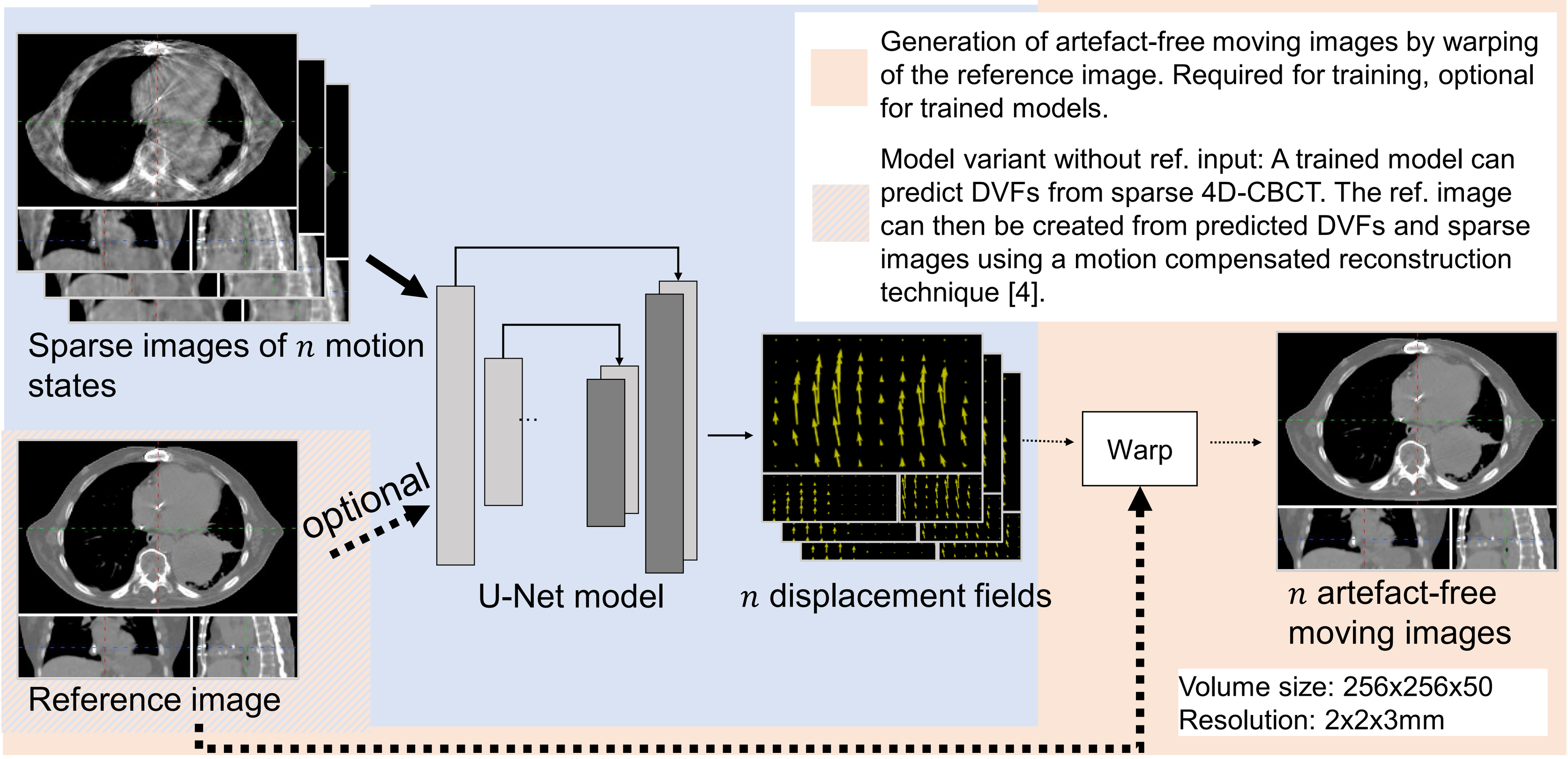
This yields a full set of 4D-CBCT projections where each projection corresponds to a different respiratory state and a set of ground truth volumes for all phases. An iterative reconstruction algorithm is then used to create the sparse (phase gated) training volumes from these .

### Datasets

The training dataset consists of 560 simulated 4D-CBCT of 56 different patients, each scan using a different breathing curve; the generated data include fully sampled ground-truth images that are used to train the network. For validation and calculation of the reported metrics we use a new set of 120 scans of 12 unseen patients.

### Models & Implementation

We trained U-Net-type convolutional neural network models to predict multiple (10) DVFs in a single forward pass given multiple sparse, gated CBCT and an optional artefact-free reference image as inputs. The predicted DVFs are used to warp the reference image to the different motion states, resulting in an artefact-free image for each state. The models are trained in a supervised way by comparing the resulting images to the ground truth. The trained models are four layers deep, the image size is halved at each downward step and nearest-neighbour upsampling is used in the upward branch. All activations are leaky ReLu (slope=0.2).




## CONCLUSIONS

To the best of our knowledge, this is the first time CNNs are used to predict multi-phase DVFs in a single forward pass. This enables novel applications such as 4D-auto-segmentation, motion compensated image reconstruction, motion analyses, and patient motion modeling.

## RESULTS

Our method clearly outperforms pairwise registration using the Deeds algorithm alone. PSNR improved from 28.6 to 38.6. In addition, the runtime of our learning-based method is orders of magnitude shorter (2 seconds instead of over 10 minutes). Our results also indicate slightly improved performance compared to pairwise registration with VoxelMorph (delta-PSNR=1.6). We also trained a model that does not require the artefact-free reference image during inference demonstrating only marginally compromised results (delta-PSNR=-0.7).

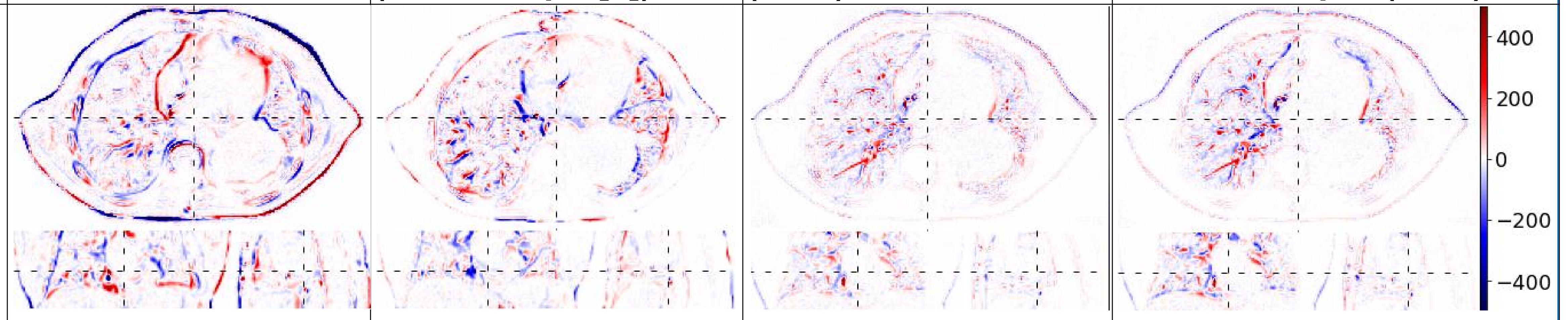
| | Pairwise Deeds [3] | Pairwise U-Net (VoxelMorph [2]) | Multiphase U-Net, (ours) | Multiphase U-Net, no reference input (ours) |
|---|---|---|---|---|
| Diff. to ground truth at max. amplitude |  | | | |
| PSNR | 28.6 | 37.0 | 38.6 | 37.9 |
| Runtime* (mm:ss) | 10:40 | 00:04** | 00:02 | 00:02 |



*Table 1:* Summary of results for our proposed registration methods compared to deeds and pairwise VoxelMorph registration algorithms. All metrics are averaged over a validation set of 12 unseen patients. All plots use level/window 0/1000 HU.
* Average runtime to register a single scan of 10 phases on a machine with 2xAMD EPYC 7513 32-Core CPUs (Deeds) and using a single Nvidia RTX A6000 (other methods)
** Pairwise registration of all phases requires 10 forward passes through the network

## REFERENCES

[1] Paysan P et al. CT Based Simulation Framework for Motion Artifact and Ground Truth Generation of Cone-Beam CT. AAPM Annual Meeting 2019.
[2] Balakrishnan G et al. VoxelMorph: A Learning Framework for Deformable Medical Image Registration. IEEE Transactions on Medical Imaging. 2019;38(8):1788-1800. doi:10.1109/TMI.2019.2897538
[3] Heinrich MP et al. MRF-based deformable registration and ventilation estimation of lung CT. IEEE Trans Med Imaging. 2013;32(7):1239-1248. doi:10.1109/TMI.2013.2246577
[4] Brehm M et al. Self-adapting cyclic registration for motion-compensated cone-beam CT in image-guided radiation therapy. Med Phys. 2012;39(12):7603-7618. doi:10.1118/1.4766435

## ACKNOWLEDGEMENTS

Co-financed by Innosuisse, grant no. 35244.1 IP-LS.
S.S. and P.P. are full-time employees of Varian Medical Systems Imaging Laboratory GmbH

## CONTACT INFORMATION

Lukas Lichtensteiger, ZHAW: licn@zhaw.ch