Deep Learning-based Deformable Registration of MR Cine Sequences for Automated and Patient-specific Target Tracking on MR-Linacs

Brady Hunt, PhD, Thayer School of Engineering at Dartmouth; Gregory A. Russo, MD; Bassem I. Zaki, MD; Gobind Gill, MD; Zhong Han; David J. Gladstone, ScD; Rongxiao Zhang, PhD

Introduction
MR-Linacs utilize real-time MR cine imaging to track the motion of internal targets and gate beam delivery. However, due to the time constraints of fast MR registration to support real-time gating, current approaches do not robustly account for deformable motion, which is common in the abdomen and thorax. Furthermore, conventional approaches are susceptible to imaging noise and cannot readily be optimized online to improve performance on a patient-to-patient basis.

Hypothesis
Deep neural networks can be trained to accurately predict deformation vector fields between two MR cine images. Such a model may better account for complex motion of deformable targets and support real-time deformable motion tracking on MR-Linacs.

Methods
We utilized Voxelmorph, a previously developed deep convolutional neural network, which takes two images (a referenced image and a moving image) and outputs the deformation field between the two. The network was trained using MR cine frames from 10 patients who underwent MR-guided radiation therapy using the Viewray MRIdian MR-Linac. MR cine sequences for each patient were partitioned temporally with the first half of the cine sequence used for training and the second half used to test the accuracy of the trained network. Two conventional methods, a fast affine and nonrigid b-spline, were also utilized on the test sequences to compare speed and accuracy of registration results. Accuracy was measured using the root mean square error (RMSE) between the transformed and reference images.

Results
A total of 2,000 MR cine frames (a 25-second MR cine sequence from each patient) were used to train the deep learning deformable image registration method, and an additional 2,000 frames were withheld as a test set. Deep learning-based deformable image registration reduced the average registration error compared to conventional affine and b-spline methods (RMSE deep neural network vs affine: 34.4 vs 53.0, paired t-test, p < 0.001; RMSE deep neural network vs b-spline: 34.4 vs 48.7, paired t-test, p < 0.001). Average computational time required to register a single frame was 31.9 ms, 44.4 ms, and 445 ms for neural network, affine, and b-spline methods respectively.

Conclusion
Deep learning-based deformable image registration was demonstrated using MR cine data acquired during MR-guided radiation therapy. The proposed approach better accounted for deformable patient motion while still supporting real-time tracking at high frame rates.

Statement of Impact
Accurate tracking of deformable patient motion can improve the precision of MR-guided radiation therapy. The proposed deep learning method is well-suited to overcome the tradeoffs in speed/accuracy of conventional registration approaches.
Overview of the proposed deep learning-based deformable image registration method. The approach uses an encoder decoder network which takes a reference and moving frame as input and outputs a dense vector field estimation of the deformable motion. The motion field can then be applied to the moving image to register it to the reference image, or otherwise utilized to transform a target contour to track its position over time.


Keywords
deep learning; radiation therapy; image registration