CoRX-NET: a Novel Unsupervised Deep Image Registration Network with Feature Sharing for Accurate Voxel-based Dosimetry

Keon Min Kim[a,b] , Min Seok Suh [c], Teik Hin Tan [d] ,
Haniff Shazwan Muhd Safwan Selvam [d], Gi Jeong Cheon [c], Keon Wook Kang [c]
and Jae Sung Lee [c,e*]

[a] *Interdisciplinary Program in Bioengineering, College of Engineering, Seoul National Graduate School, Seoul, Korea (Rep. of)*
[b] *Integrated Major in Innovative Medical Science, Seoul National Graduate School, Seoul, Korea (Rep. of)*
[c] *Department of Nuclear Medicine, College of Medicine, Seoul National University, Seoul, Korea (Rep. of)*
[d] *Nuclear Medicine Centre, Sunway Medical Centre, Selangor, Malaysia*
[e] *Department of Biomedical Sciences, College of Medicine, Seoul National University, Seoul, Korea (Rep. of)*
* *Corresponding author: jaes@snu.ac.kr*

Purpose– Generating 3D time-integrated activity (TIA) maps in the voxel-based dosimetry for radionuclide therapy requires non-linear registration of SPECT/CT images acquired at different time points. Most widely used approach for SPECT/CT registration and TIA map generation is to apply an affine transformation matrix and nonlinear deformation fields generated from a pair of CT images to the SPECT images. However, in this conventional approach, the accuracy of SPECT image registration and TIA map generation is not high enough for tumor regions with insufficient morphological information in CT images. Therefore, in this study, we propose CoRX-NET, a novel unsupervised deep-learning based approach that simultaneously uses SPECT and CT images for accurate image registration required for voxel-based dosimetry.

Methods– For network training, 22 SPECT/CT image sets from 7 patients treated with ^{177}Lu-DOTATATE at Seoul National University Hospital were used. Each image set consists of 3 or 4 SPECT/CT images acquired at different time points. 468 pairs sampled from the image sets were divided into 408, 48, and 12 pairs for network training, evaluation, and internal test, respectively. Each training pair consists of two SPECT/CT images (i.e. fixed and moving SPECT/CT images) acquired at different times. The ^{177}Lu-DOTATATE SPECT/CT images acquired from 14 patients treated at Malaysia Sunway Medical Centre were used as the external test sets and the number of image pairs was 28. Swin-Transformer (CVPR 2021) was used as a backbone network and cross-stitch layer (CVPR 2016) was applied to share features generated from SPECT and CT images. LNCC (local normalized cross-correlation) loss was used for the similarity measurement of CT and SPECT, respectively. In addition, a bending energy regularization loss term was added to avoid image folding artifacts. The network performance was compared to Elastix and TransMorph (MIA 2022) using L1 loss and SSIM of CT, SSIM of normalized SPECT, and LNCC of SPECT as metrics. Furthermore, the voxel-wise RMSE of the TIA map fitted with mono-exponential function was calculated for the comparison of the goodness of fitting.

Results– The comparison of metrics included in an ablation study is summarized in Table 1. The ablation study has shown that attaching a cross-stitch layer to the network results in better registration with less CT L1 loss and SPECT LNCC and higher SSIM of SPECT and CT. CoRX-NET showed more accurate registration performance compared to the Elastix and TransMorph, except for CT registration of external test sets. Although Elastix showed better registration performance for CT than CoRX-NET on external test sets, the deformed CT is not used directly for TIA map generation. The mean RMSEs of TIA maps in tumor and kidney were 178.60 and 57.13 for Elastix and 100.32 and 43.54 for CoRX-NET. CoRX-Net takes 6 seconds for registration whereas Elastix takes 50 seconds.

Conclusion– The proposed network showed better performance in sequential SPECT/CT registration compared to the conventional iterative registration method (Elastix) and TransMorph, which has shown high performance in medical image registration. As a result, a more reliable TIA maps could be generated using CoRX-NET, mitigating errors in voxel-based dosimetry due to mis-registration of SPECT images. Practical voxel-based dosimetry will be also possible with reduced registration time and a convenient end-to-end registration process.

Table 1. The comparison of metrics (L1 loss and SSIM of CT and LNCC loss and SSIM of normalized SPECT) between Elastix, TransMorph and CoRX-NET

		CT L1 loss	CT SSIM	Normalized SPECT SSIM	LNCC Loss
Elastix	Internal	0.0173	0.9024	0.9520	0.3772
	External	**0.025**	**0.9594**	0.9073	0.2923
TransMorph (MIA 2022)	Internal	0.0099	0.9073	0.9544	0.3584
	External	0.0284	0.9371	0.9179	0.3477
CoRX w/o cross-stitch	Internal	0.0098	0.9124	0.967	0.211
	External	0.0278	0.9425	0.9202	0.1544
CoRX with cross-stitch	Internal	**0.0090**	**0.9226**	**0.9681**	**0.1959**
	External	0.0259	0.9482	**0.9222**	**0.1466**

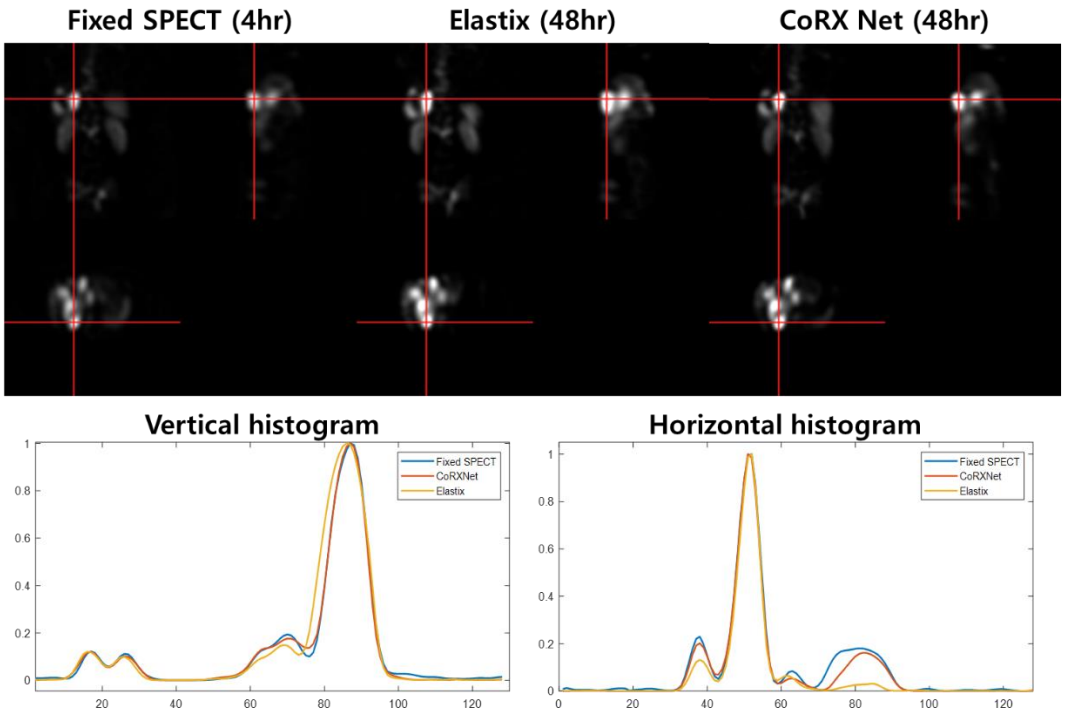

Fig. 1. The vertical and horizontal line profiles of fixed and deformed SPECT acquired using Elastix and CoRX-NET