

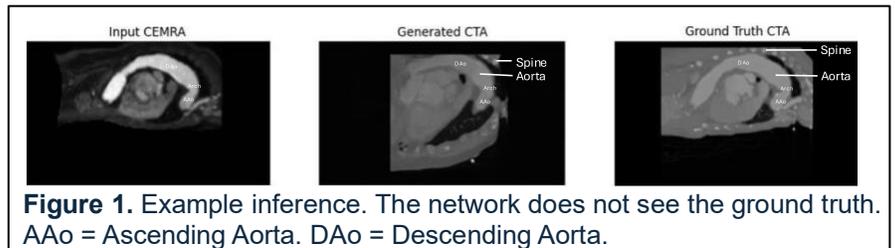
Deep Learning Image-to-Image Translation from CEMRA to CTA via Conditioned Diffusion

Sebastian Cohn¹, William Dong¹, Bradley Allen¹, Michael Markl¹

¹Northwestern University, Department of Radiology

Purpose: Generative AI has potentially powerful applications in healthcare, ranging from data augmentation to cross-modality image synthesis. In this paper, we explore the use of a diffusion deep learning model to create a synthetic computed tomography angiography (CTA). We guide the diffusion process by conditioning the network on a patient's contrast-enhanced magnetic resonance angiography (CEMRA) to preserve patient-specific anatomy.

Methods: The 2D Denoising Diffusion Probabilistic Model (DDPM) implementation in Project MONAI¹ was originally used to generate synthetic hand X-ray images from noise. In this project, the model architecture was modified to generate synthetic CTA images by adding a second input channel to the model to accept a conditional input, the CEMRA. We used pairs of chest CEMRA and CTA data for $n = 53$ patients. CEMRA data were used as input data to the AI model and CTA images served as ground truth data for the CEMRA to CTA image translation. All CEMRA and CTA images were first resampled to have a voxel size of 2.5 mm x 2.5 mm x 3 mm using MATLAB. The CTA images were cropped to only include the thoracic aorta region and zero-filled to a matrix size of 100 x 104 x 175. The resampled CEMRA images were registered to corresponding CTA images using the ANTs package in 3D Slicer (the CTA served as the fixed image while the CEMRA was the moving image). We used a combination of 3 registration algorithms (called QuickSyN in 3D Slicer): rigid, then affine, then SyN. Rigid transformation only involves translation and rotation; affine transformation is rigid with the addition of shear and scaling (12 parameters); symmetric normalization (SyN) is affine plus a deformable transformation, which is iteratively optimized based on mutual information between the CTA and CEMRA images. After the registration process, all images were normalized to have an intensity range of $[-1, 1]$ to allow the network to better compare CEMRA and CTA values. Additionally, sagittal slices toward the edge (15 on each side of every volume) were discarded as they included mostly blank images (resulting in a total of 70 slices per image). Finally, the 3D images were broken up into 2D sagittal slices with matrix sizes of 104 x 175 to be inputted into the training model. The results of the registration process were visually inspected and rated by 2 researchers to be either sufficient or poor. Registrations with poor performance were repeated until improved results were obtained, poor results were still yielded after 10 repetitions. 13 of 53 CTA/CEMRA pairs failed to yield satisfactory registration results. Overall, we had 40 sets of image pairs * 70 imaging slices/volume. We used 38 patients (2660 slices) for model training and 2 patients (140 slices) for testing. The model was trained for 75 epochs, using a batch size of 8, a timestep T set to 1000, the ADAM optimizer, and a mean square error loss function. To evaluate the performance of our network, we calculate the signal intensity difference in average intensity of a small (7 mm diameter circle) region of interest in the aorta and spine for both the AI-generated and original ground-truth CTA images.



Results: Out of the 140 generated CTA imaging slices for testing (Fig. 1), 27 were noisy and 85 do not contain both the aorta and spine for intensity evaluation. This left 28 images suitable for analysis. As seen in Table 1, the mean intensities of the aorta and spine were similar between the AI-generated and ground truth CTA images, but the intensity variance was higher in the AI-generated images compared to ground truth CTA.

AI Generated CTA mean +- STD intensity			Ground Truth CTA mean +- STD intensity			Paired t-test p-value		
Aorta SI	Spine SI	Aorta - Spine SI	Aorta SI	Spine SI	Aorta - Spine SI	Aorta SI	Spine SI	Aorta - Spine SI
-0.38 +- 0.10	-0.29 +- 0.13	-0.08 +- 0.07	-0.33 +- 0.01	-0.27 +- 0.06	-0.06 +- 0.06	0.02	0.38	0.18

Table 1. Statistics for generated and ground truth CTA intensity mean and standard deviation. SI = signal intensity.

Discussion: Our work shows the potential for implementing deep learning based cross-modality synthesis in chest imaging using a CEMRA to generate a CTA for a patient. A limitation of our work is our small training and testing set. Another limitation is that the paired data from each patient was not aligned fully (different fields of view and time of acquisition) resulting in failed image registration for some patients. In addition, we do not include a quantitative assessment of our registration error. Next steps will include training the model on a larger dataset and implementing further evaluation metrics. If successfully developed, this technology could provide physicians with CTA-equivalent insights without requiring patients to undergo an additional scan if they already have a CEMRA.

References: [1] Pinaya WH, Graham MS, Kerfoot E, Tudosi PD, Dafflon J, Fernandez V, Sanchez P, Wolleb J, Da Costa PF, Patel A, Chung H. Generative AI for Medical Imaging: extending the MONAI Framework. arXiv preprint arXiv:2307.15208. 2023 Jul 27.