Paired MR-to-sCT Translation using Conditional GANs
An Application to MR-guided Radiotherapy

Alexandra Alain-Beaudoin, Laurence Savard and Silvain Bériault
Advanced Development Engineering, Elekta Ltd., Montréal, Canada

Abstract

We present a method for MR-to-sCT image translation using paired training data. The method is based on the Pix2Pix conditional GAN architecture. A multi-channel (2.5D) approach is used to improve translation results thru-plane in comparison to applying a 2D model independently on each slice, while keeping inference time small in comparison to a full 3D approach. Separate models were trained for both brain (T1-weighted) and pelvis (T1- and T2-weighted) using already paired data as provided by SynthRAD2023 challenge. Models were validated using 60 validation subjects provided by the challenge. Image similarity metrics obtained during the validation phase are: mean absolute error (MAE) of 64.27 ± 14.15, peak signal-to-noise ratio (PSNR) of 28.64 ± 1.77, structure similarity index (SSIM) of 0.872 ± 0.032.
New: Results of the test phase have been added to this poster. Image similarity metrics are comparable to those of validation phase (no degradation). Dosimetric evaluation results are also provided.

Data

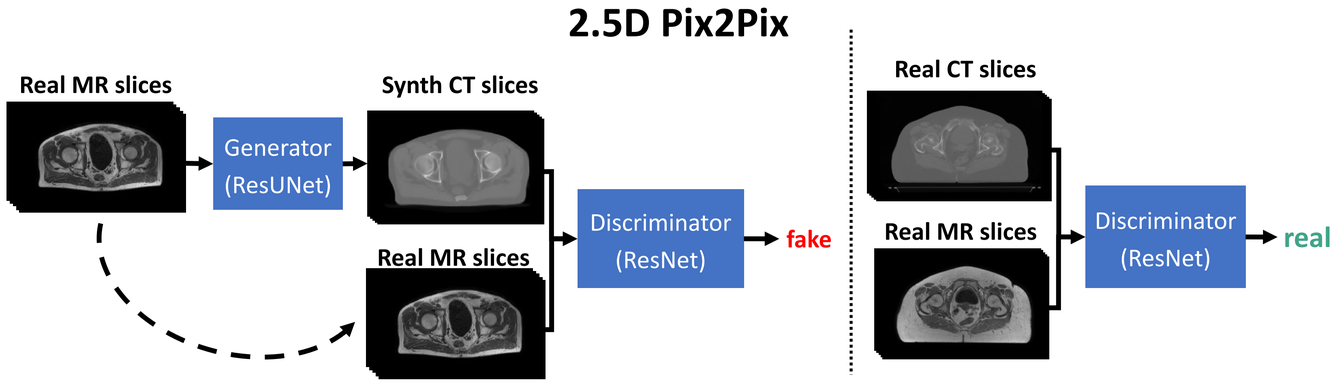
Challenge Data (3 centers)

Tumor Site	Paired Data*	Training	Validation	Test
Pelvis	CT + MR (T1w or T2w)	180	30	60
Brain	CT + MR (T1w)	180	30	60

*Rigidly aligned CT and MR

Method

2.5D Pix2Pix

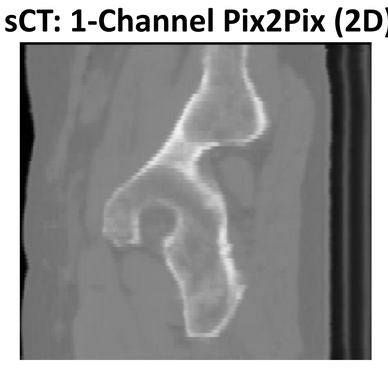


Main Hyperparameters

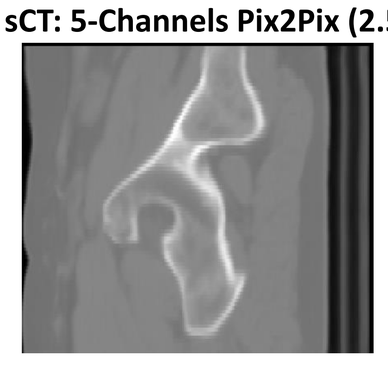
- Loss function: $\mathcal{L}(G,D) = \mathcal{L}_{GAN}(G,D) + \lambda\mathcal{L}_R(G)$
  - $\mathcal{L}_{GAN}$ is a least-square loss term
  - $\mathcal{L}_R$ is a L1 loss term($\lambda = 50$)
- Data augmentation:
  - geometric (MR and CT)
  - intensity (MR-only)
- 2500 iter. | 300 batches/iter. | batch size = 16

Less Staircase Effects

sCT: 1-Channel Pix2Pix (2D)



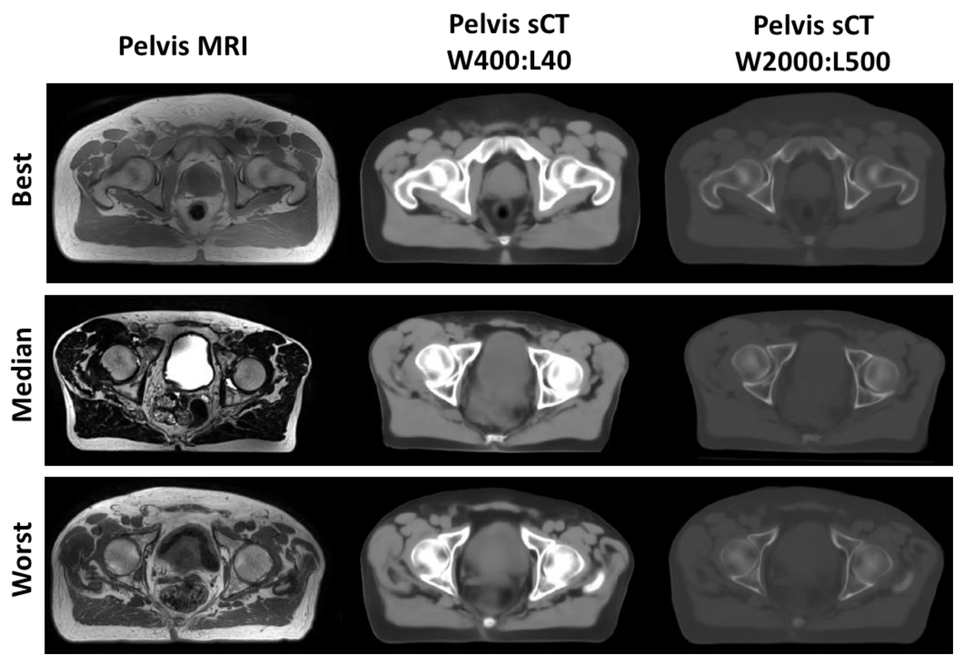
sCT: 5-Channels Pix2Pix (2.5D)

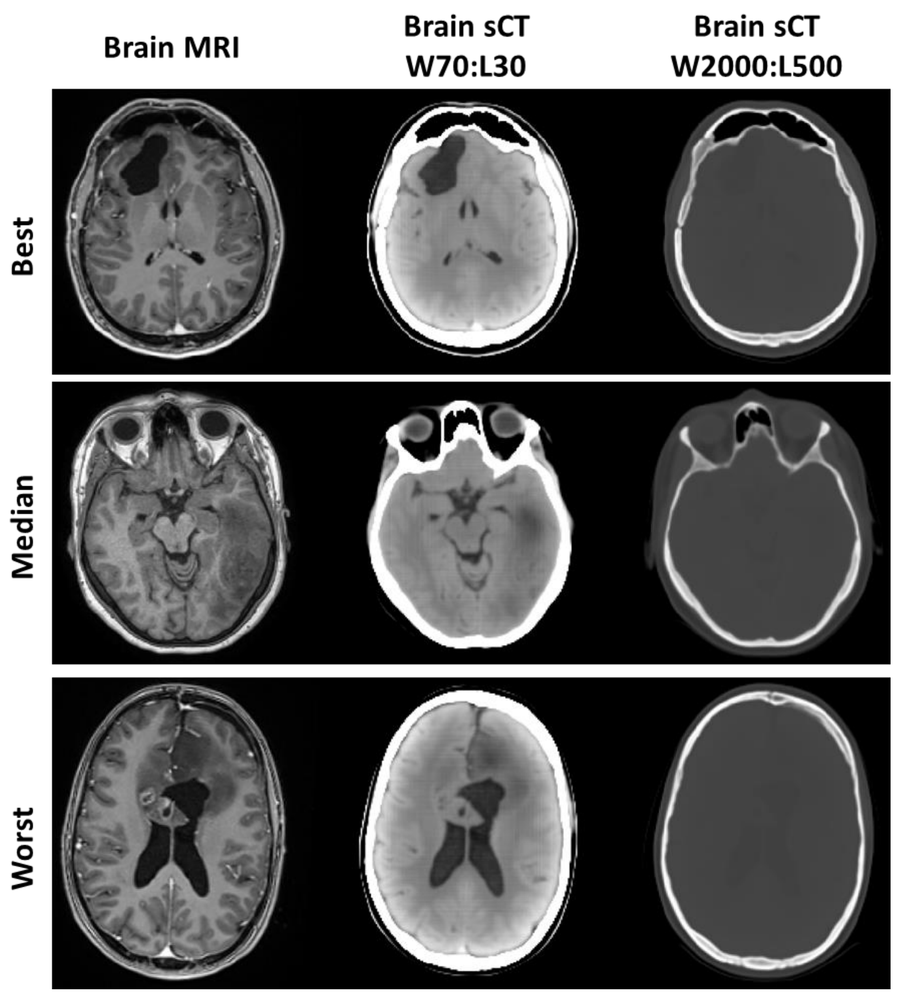


Full List of Hyperparameters

Hyperparameter	Description	Recommended value
$\mathcal{L}_{GAN}$	Type of adversarial loss function to use (either Classic, Least Squares or Hinge)	Least Squares
Learning rate	Learning rate for the generator (G) and discriminator(D). The learning rate will linearly decrease to zero starting when half the number of iterations has been completed.	G: 1e-4 D: 5e-5
Num filters	Number of filters per layer in the generator (G) and discriminator (D)	G: 64,128,256,512 D: 64,128, 256,512,512
Num disc updates	Number of discriminator updates per generator update	2
Spectral norm	Specify if spectral normalization layers are used in both the discriminator and generator networks	True
L1 weight (λ)	Weight of L1 loss term	50
Voxel size	Voxel size in mm	Brain: 1x1x1 mm Pelvis: 1x1x2.5 mm
Sample size	The size of a sample (in voxels) used during training and inference. For a 2D network, the 3rd dimension is the channel dimension (c.f. Section 2.4)	192x192x5
Batch size	Number of samples per batch	16
Num batches per iteration	Number of batches per iteration	300
Num iterations	Total number of iterations	2500
Data augmentation	Option to turn on data augmentation	True

Representative Examples





Test Results

Final Results of the Test Phase

	MAE	PSNR	SSIM	DVH (photon)	Gamma (photon)	Dose MAE (photon)	DVH (proton)	Gamma (proton)	Dose MAE (proton)
Mean	62.76	28.80	0.875	0.0286	98.15	0.00398	0.227	97.27	0.0332
Std Dev	13.06	1.60	0.030	0.0525	4.21	0.00359	0.276	2.50	0.0220
Min	29.59	24.56	0.784	0.0013	78.12	0.00039	0.007	88.33	0.0060
25pc	54.10	27.80	0.861	0.0091	99.24	0.00193	0.040	96.01	0.0173
50pc	60.82	28.81	0.882	0.0156	99.88	0.00280	0.104	97.85	0.0292
75pc	70.73	29.74	0.891	0.0313	99.99	0.00493	0.256	99.25	0.0409
Max	108.64	32.52	0.974	0.4278	100.00	0.02759	1.295	100.00	0.1333

Observations:

- Image similarity metrics are comparable to those of the validation phase (no degradation)
- High Std Dev observed in dosimetric evaluation results due to outlier cases (to be further investigated)

Conclusion

Pix2Pix remains a good baseline model for MR-to-sCT translation

Inference time is important for online workflow
- 20 sec for brain model
- 30 sec for pelvis model

Model ensembling
- Small effect on image similarity metric
- Inference time increases linearly

Other considerations/Future work
- Improving image registration
- Gas pockets in pelvis

SynthRAD2023
Synthesizing computed tomography for radiotherapy

MICCAI 2023 Vancouver CANADA