SYNTHETIC CT GENERATION USING MRI WITH DEEP LEARNING: HOW DOES THE SELECTION OF INPUT IMAGES AFFECT THE RESULTING SYNTHETIC CT?

Andrew P. Leynes[⋆†] *Peder E. Z. Larson*[⋆†]

[⋆] UC Berkeley - UC San Francisco Joint Graduate Program in Bioengineering
[†] Department of Radiology and Biomedical Imaging, University of California San Francisco

ABSTRACT

Synthetic x-ray computed tomography (CT) images derived from magnetic resonance imaging (MRI) is a recent area of focus for medical imaging researchers for applications in attenuation correction in simultaneous PET/MRI systems and MRI-guided radiotherapy planning. Several research groups have demonstrated the potential of deep learning to generate the synthetic CT images, however, there are several major open questions that remain with this approach. We investigated how the selection of MRI inputs affect the resulting output using a fixed network. We found that Dixon MRI may be sufficient for quantitatively accurate synthetic CT images and ZTE MRI may provide additional information to capture bowel air distributions.

Index Terms— synthetic ct, deep learning, mri, Dixon mri, zero echo-time mri

1. INTRODUCTION

Synthetic CT images are used for applications where CT images are required but are unavailable, such as in positron emission tomography (PET) attenuation correction in simultaneous PET/MRI systems and MRI-guided radiotherapy planning.

Four different MRI protocols are typically used for synthetic CT generation: conventional T1- or T2-weighted MRI [1, 2], Dixon MRI [3], UTE MRI (RESOLUTE) [4], and ZTE MRI [5]. The major advantage of using Dixon MRI, UTE MRI, and ZTE MRI over conventional MRI is that they have been found to have signal intensity correlations with Hounsfield units: Dixon MRI fractional fat/water maps have a linear relationship in fat compartments; UTE MRI R2* mapping has a non-linear relationship in soft tissues and bone; and ZTE MRI proton-density signal intensity has log-linear relationships in bone.

Synthetic CT images must be both quantitatively and geometrically accurate. MRI radiotherapy planning requires precise bone geometry measurements to accurately register planning imaging results and in-room imaging during treatment. In PET attenuation correction, the quantitative accuracy is more important than anything else, since this is used for correcting for radiotracer uptake. Any errors in the synthetic CT will propagate to the PET images.

Various image processing and machine learning methods have been used to turn these MRI images into synthetic CT images [6]. Most recently, synthetic CT generation methods have been utilizing deep learning. The deep learning task is essentially image transformation: MRI images are converted into synthetic CT images. Results from prior work have demonstrated the potential for deep learning to produce synthetic CT images [1, 2, 7]. However, several major open questions remain with this approach: it is not clear what MRI images would produce the best synthetic CT images; it is not known how the selection of the inputs affect the resulting output; and it is not clear what kinds of networks would do the best job. To begin to answer these questions, this paper uses different combinations of input images with deep learning and assesses the resulting synthetic CT images. This work is focused on the pelvis, where synthetic CTs are useful for evaluation of prostate cancer and other pelvic malignancies.

2. METHODOLOGY

Patients with pelvis lesions were scanned using an integrated 3 Tesla time-of-flight (TOF) PET/MRI system [8] (SIGNA PET/MR, GE Healthcare, Waukesha, WI, USA). The patient population consisted of 26 patients (Age = 58.1 ± 14.2 years old, 15 males, 11 females): ten (10) patients were used for training and sixteen (16 patients) were used for validation.

The following MRI sequences were acquired: Dixon MRI (FOV = $500 \times 500 \times 312$ mm, resolution = 1.95×1.95 mm, slice thickness = 5.2 mm, slice spacing = 2.6 mm, scan time = 18 s) and ZTE MRI (cubical FOV = $340 \times 340 \times 340$ mm, isotropic resolution = $2 \times 2 \times 2$ mm, 1.36 ms readout duration, FA = 0.6^o, 4μs hard RF pulse, scan time = 123s). The Dixon fat and water maps were produced by the scanner reconstruction software.

Helical CT images of the patients were acquired separately on different machines and were co-registered to the MR images using the ANTS [9] registration package using the

The authors of this work received research funding from GE Healthcare.

SyN diffeomorphic deformation model with combined mutual information and cross-correlation metrics. Multiple CT protocols were used with variable parameter settings (110 - 130 kVp, 30 - 494 mA, rotation time = 0.5 s, pitch = 0.6 - 1.375, 11.5 - 55 mm/rotation, axial FOV = 500 - 700 mm, slice thickness = 3 - 5 mm, matrix size = 512×512).

The same methodology in our previous work was used for MRI and CT image pre-processing and co-registration [5]. This included N4 bias correction [10] and soft-tissue normalization for ZTE images. A similar pre-processing was done for Dixon T1-weighted images: N4 bias correction was applied and the signal intensities were normalized to the fat peak in the image histogram.

Figure 1 shows the different MRI inputs used for this work: bias-corrected soft-tissue normalized proton-density-weighted ZTE, bias-corrected and fat-peak normalized Dixon T1-weighted image, Dixon fractional fat map, and Dixon fractional water map.

A previously-published deep convolutional neural network for synthetic CT generation was used [7] and the same training methodology and synthetic CT generation for the network was performed. The combinations of MRI inputs used were as follows:

1. ZTE + Dixon fractional fat + Dixon fractional water (ZeDD) [7]
2. ZTE only
3. Dixon T1-w + Dixon fractional fat + Dixon fractional water (Dixon-all)
4. Dixon T1-w only
5. Dixon fractional fat + Dixon fractional water (Dixon fat/water)
6. ZTE + Dixon T1-w + Dixon fractional fat + Dixon fractional water (MRI-all)

Qualitative analysis of the training curves of the different combinations of MRI inputs was performed. Visual inspection of the synthetic CT images was performed to observe any qualitative differences. The quantitative analysis of the synthetic CT images was performed only in body voxels (synthetic CT $>$ -120 HU AND CT $>$ -120 HU) to eliminate any errors due to air. The mean absolute error was measured over the whole body, in soft-tissues (-120 to 100 HU), in spongeous bone (100 to 300 HU), and cortical bone ($>$ 300 HU) across the validation dataset. Additionally, the synthetic CT images were down-sampled to a resolution of $4.69 \times 4.69 \times 2.81$ mm and filtered with a 10 mm Gaussian kernel in the quantitative analysis to simulate the pre-processing step used for producing PET/MRI attenuation coefficient maps.

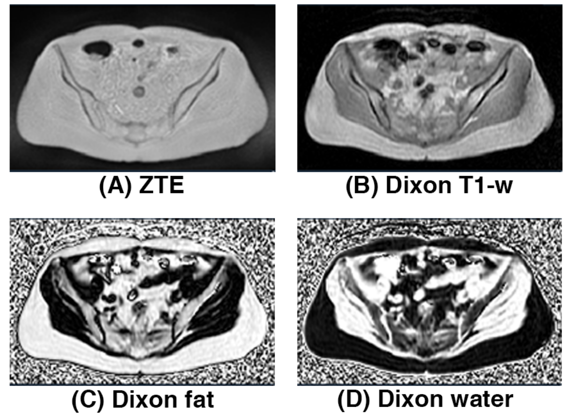

Fig. 1. Different MRI input images used for generating the synthetic CT images.

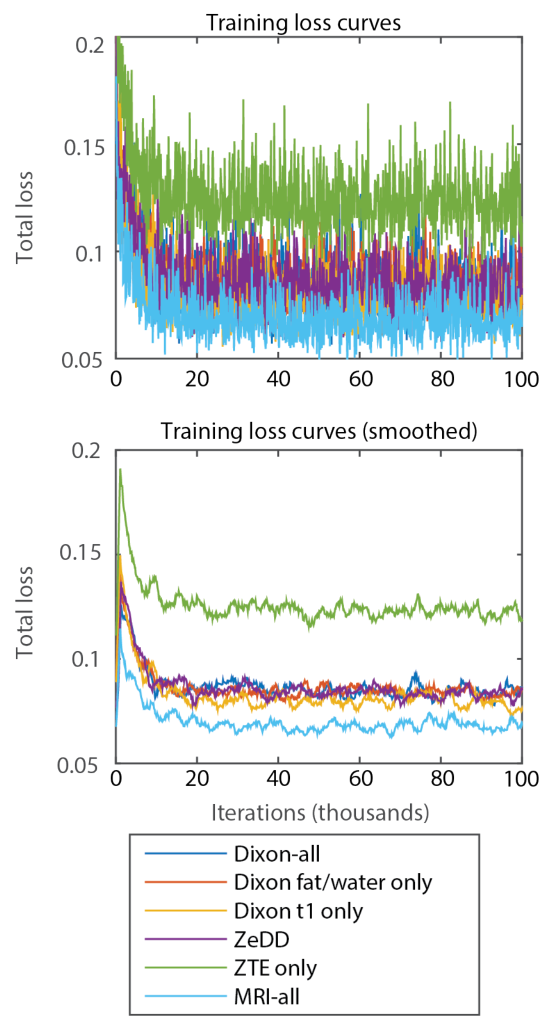

Fig. 2. Training loss curves for each combination of MRI inputs.

3. RESULTS

Figure 2 shows the training curves for each combination of MRI inputs. Each training converged at approximately 20 thousand iterations. ZTE only inputs achieved the highest training loss, MRI-all achieved lowest training loss, while all other inputs provided very similar training behaviors at about 0.7 total training loss.

Figure 3 shows the synthetic CT images produced with the different combinations of MRI inputs and the ground-truth CT for one patient in the validation set. Similar to the training curves shown in Figure 2, the ZTE-only synthetic CT appeared significantly different, qualitatively, compared to the other synthetic CT images. All the other synthetic CT methods provided excellent depiction of soft-tissues and bone. However, only the ZTE-based methods were able to produce bowel air in the synthetic CT images. Difference images for each synthetic CT method with ground truth CT are shown in Figure 4.

Figure 5 shows the mean absolute error of each synthetic CT compared to ground-truth CT. A summary of the errors is shown in Table 1. As expected from previous figures, ZTE-only synthetic CT produced the most error across the whole body. In each specific tissue compartment, ZTE-only had the largest errors. For the other synthetic CT images, ZeDD, Dixon-all, Dixon T1-w only, Dixon fat/water only, and MRI-all were comparable in the soft-tissue compartments. Dixon-all, Dixon T1-w, and MRI-all were comparable in bone compartments and had less error than ZeDD and Dixon fat/water. Although MRI-all provided the lowest training loss, the mean absolute error did not differ by more than 5 HU from Dixon-all and Dixon T1-w only.

4. DISCUSSION

Several groups have now demonstrated that deep learning can effectively produce synthetic CT images from MRI. This work extends on this idea, evaluating what types of MRI inputs are required. We found that, in the pelvis, using only T1-weighted MRI as the input was effective at generating synthetic CT. This was a surprising result, as we expected ZTE MRI would provide additional bone information that could improve the synthetic CT. One possible explanation is that the bone regions on T1-weighted MRI in the pelvis are well-defined by the lack of signal, and thus the non-zero signal in bone with ZTE MRI is not required. The ZTE MRI also experiences blurring due to fat chemical shift, which may create some blurring around bone marrow and intra-abdominal fat in the resulting synthetic CT.

One limitation of this study is that we did not explore network architecture optimizations for the different combinations of input images. We also could not examine the quantitative accuracy of bowel air. This is because there is no gold-standard CT for training, as the bowel air will shift between CT and MRI scans. Qualitatively we observed that the synthetic CT that utilized ZTE as an input would have some bowel air in certain patients. All other synthetic CT methods did not produce any bowel air, which suggests that ZTE provides additional information for distinguishing bowel air.

These results in the pelvis may not generalize to other anatomical regions. For the brain in particular, the major challenge is to distinguish bone and air in the sinuses. This may particularly benefit from ZTE MRI data.

Based on this work, we could explore further optimizations of the MRI inputs, such as spatial resolution, degree of T1-weighted contrast and scan time to further improve synthetic CT generation.

5. CONCLUSION

We investigated the effects of having different combinations of MRI inputs to generate synthetic CT images with a deep convolutional neural network model. We found that, in the pelvis, Dixon MRI may be sufficient to produce quantitatively accurate synthetic CT images.

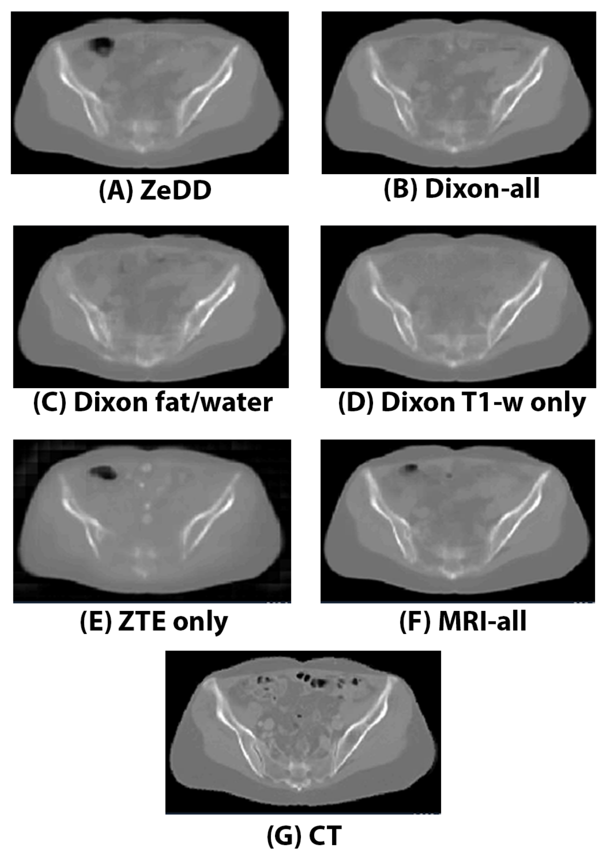

Fig. 3. Resulting synthetic CT images using different combinations of MRI inputs. Note that the methods that utilized ZTE were able to produce bowel air while the others filled it with soft-tissue HU values.

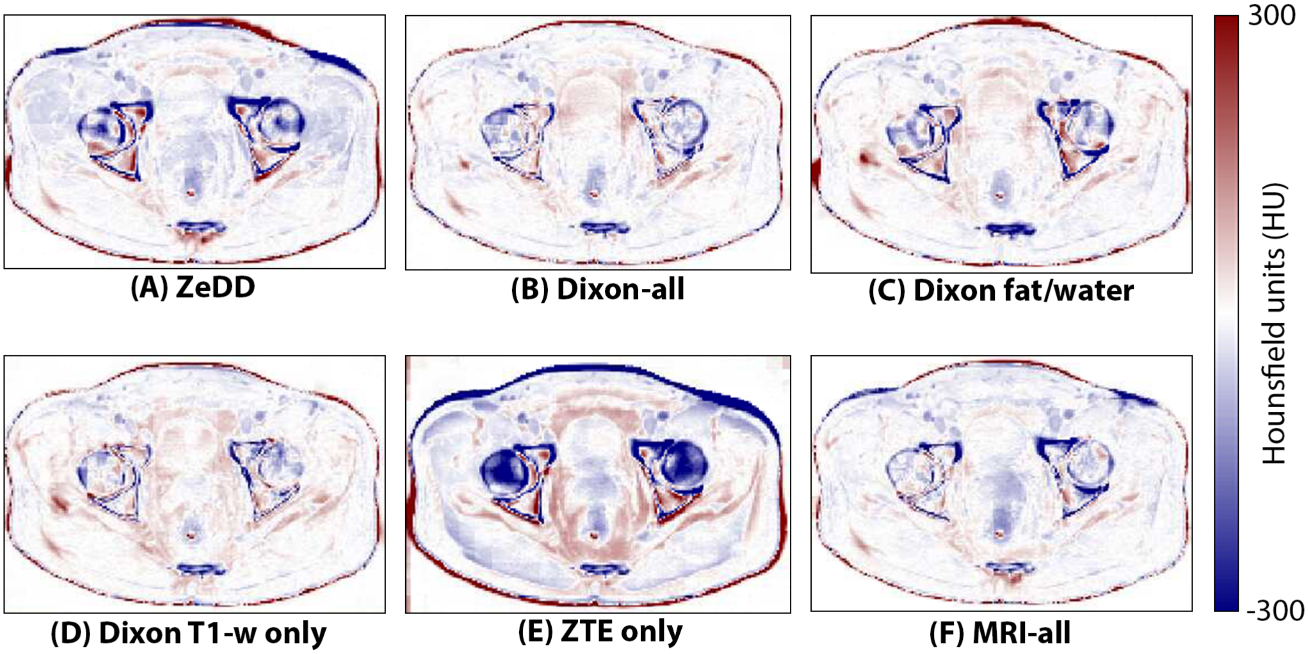

Fig. 4. Difference images of (A) ZeDD, (B) Dixon-all, (C) Dixon fat/water, (D) Dixon T1-w only, (E) ZTE only, and (F) MRI-all with ground truth CT for one slice in the pelvis for one patient. The difference images show good quantitative accuracy of Dixon-all, Dixon T1-w only, and MRI-all compared to the other synthetic CT methods.

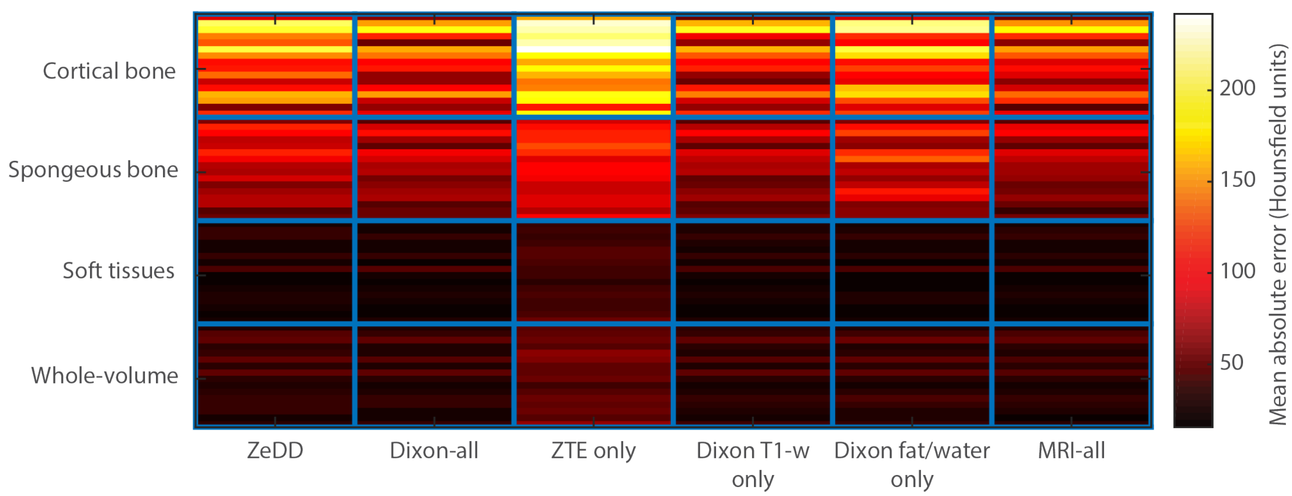

Fig. 5. Mean absolute error for each synthetic CT generated from different MRI inputs (columns) in the whole body and different tissue compartments (rows). Each line in a cell corresponds to one patient. The models saved at iteration 45,000 were used.

Table 1. Mean absolute error (mean $\pm$ standard deviation) for each synthetic CT across patients for each tissue type.

	ZeDD	Dixon-all	ZTE only	Dixon T1-w only	Dixon fat/water only	MRI-all
Whole volume	31.85 ± 7.86	28.80 ± 8.07	47.15 ± 6.58	29.81 ± 7.57	31.65 ± 7.89	29.05 ± 7.76
Soft tissues	23.19 ± 6.32	21.93 ± 6.53	36.44 ± 4.59	23.33 ± 5.88	22.11 ± 6.06	22.46 ± 6.21
Spongeous bone	75.73 ± 19.65	65.66 ± 19.00	95.56 ± 13.67	62.72 ± 17.79	81.81 ± 24.85	62.44 ± 18.55
Cortical bone	131.23 ± 43.26	104.05 ± 41.36	181.41 ± 37.55	104.07 ± 42.51	135.65 ± 46.08	101.9 ± 39.65

6. REFERENCES

[1] Dong Nie, Xiaohuan Cao, Yaozong Gao, Li Wang, and Dinggang Shen, "Estimating CT Image from MRI Data Using 3d Fully Convolutional Networks," in *Deep Learning and Data Labeling for Medical Applications*. Oct. 2016, pp. 170–178, Springer, Cham, DOI: 10.1007/978-3-319-46976-8_18.

[2] Fang Liu, Hyungseok Jang, Richard Kijowski, Tyler Bradshaw, and Alan B. McMillan, "Deep Learning MR Imagingbased Attenuation Correction for PET/MR Imaging," *Radiology*, p. 170700, Sept. 2017.

[3] S.D. Wollenweber, S. Ambwani, A.H.R. Lonn, D.D. Shanbhag, S. Thiruvenkadam, S. Kaushik, R. Mullick, H. Qian, G. Delso, and F. Wiesinger, "Comparison of 4-Class and Continuous Fat/Water Methods for Whole-Body, MR-Based PET Attenuation Correction," *IEEE Transactions on Nuclear Science*, vol. 60, no. 5, pp. 3391–3398, Oct. 2013.

[4] Claes N. Ladefoged, Didier Benoit, Ian Law, Sren Holm, Andreas Kjr, Liselotte Hjgaard, Adam E. Hansen, and Flemming L. Andersen, "Region specific optimization of continuous linear attenuation coefficients based on UTE (RESOLUTE): application to PET/MR brain imaging," *Physics in Medicine & Biology*, vol. 60, no. 20, pp. 8047, 2015.

[5] Andrew P. Leynes, Jaewon Yang, Dattesh D. Shanbhag, Sandeep S. Kaushik, Youngho Seo, Thomas A. Hope, Florian Wiesinger, and Peder E. Z. Larson, "Hybrid ZTE/Dixon MR-based attenuation correction for quantitative uptake estimation of pelvic lesions in PET/MRI," *Medical Physics*, vol. 44, no. 3, pp. 902–913, Mar. 2017.

[6] Claes N. Ladefoged, Ian Law, Udunna Anazodo, Keith St. Lawrence, David Izquierdo-Garcia, Ciprian Catana, Ninon Burgos, M. Jorge Cardoso, Sebastien Ourselin, Brian Hutton, Ins Mrida, Nicolas Costes, Alexander Hammers, Didier Benoit, Sren Holm, Meher Juttukonda, Hongyu An, Jorge Cabello, Mathias Lukas, Stephan Nekolla, Sibylle Ziegler, Matthias Fenchel, Bjoern Jakoby, Michael E. Casey, Tammie Benzinger, Liselotte Hjgaard, Adam E. Hansen, and Flemming L. Andersen, "A multi-centre evaluation of eleven clinically feasible brain PET/MRI attenuation correction techniques using a large cohort of patients," *NeuroImage*, vol. 147, pp. 346–359, Feb. 2017.

[7] Andrew Palmera Leynes, Jaewon Yang, Florian Wiesinger, Sandeep S. Kaushik, Dattesh D. Shanbhag, Youngho Seo, Thomas Armstrong Hope, and Peder Eric Zufall Larson, "Direct PseudoCT Generation for Pelvis PET/MRI Attenuation Correction using Deep Convolutional Neural Networks with Multi-parametric MRI: Zero Echo-time and Dixon Deep PseudoCT (ZeDD-CT)," *Journal of Nuclear Medicine*, Oct. 2017, Accepted.

[8] C. S. Levin, S. H. Maramraju, M. M. Khalighi, T. W. Deller, G. Delso, and F. Jansen, "Design Features and Mutual Compatibility Studies of the Time-of-Flight PET Capable GE SIGNA PET/MR System," *IEEE Transactions on Medical Imaging*, vol. 35, no. 8, pp. 1907–1914, Aug. 2016.

[9] B.B. Avants, N.J. Tustison, M. Stauffer, G. Song, B. Wu, and J.C. Gee, "The Insight ToolKit image registration framework," *Frontiers in Neuroinformatics*, vol. 8, no. 44, Apr. 2014.

[10] N.J. Tustison, B.B. Avants, P.A. Cook, Yuanjie Zheng, A. Egan, P.A. Yushkevich, and J.C. Gee, "N4itk: Improved N3 Bias Correction," *IEEE Transactions on Medical Imaging*, vol. 29, no. 6, pp. 1310–1320, June 2010.