Optimizing the generation of pseudo-CT from MRI based on a highly efficient 3D neural network

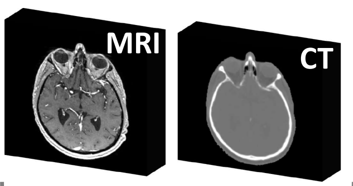
Emilie Alvarez Andres[1,2,3], Lucas Fidon[2,4], Maria Vakalopoulou[4], Marvin Lerousseau[1,3,4], Alexandre Carré[1,3], Roger Sun[1,3,4], Anne Beaudré[3], Eric Deutsch[1,3], Nikos Paragios[2], Charlotte Robert[1,3]

[1]Molecular radiotherapy and innovative therapeutics, INSERM UMR1030, Gustave Roussy Cancer Campus, Université Paris Saclay, Villejuif, France; [2]TheraPanacea, Paris, France; [3]Department of radiation oncology, Gustave Roussy Cancer Campus, Villejuif, France; [4]MICS Laboratory, CentraleSupélec, Paris-Saclay University, 91190, Gif-sur-Yvette, France

Introduction

Magnetic Resonance Imaging (MRI) has become increasingly popular these last few years, since it offers an excellent soft tissue contrast. This imaging modality is required during a brain tumor radiotherapy treatment to complement Computed Tomography (CT) imaging. Yet, this dual acquisition implies an image registration, process which was proved to induce errors up to 2mm (1) and thus to increase margins. As a result, the generation of pseudo Computed Tomography (pCT) from MRI appears to be of most interest for an optimized patient safety. This study aimed at evaluating key parameters: the training set size, MRI input sequences namely T1 weighted MRI (T1) or contrast enhanced T1 MRI (T1-Gd), MRI standardizations namely Histogram-Based (HB) (2), Zero Mean-Unit Variance (ZMUV) (3), White Stripe (WS) (4), No Standardization (NS), N4 bias field filter application (5) and network architectures, i.e. HighResNet (default) (6) and 3D Unet (7), impacts on the pCT quality.

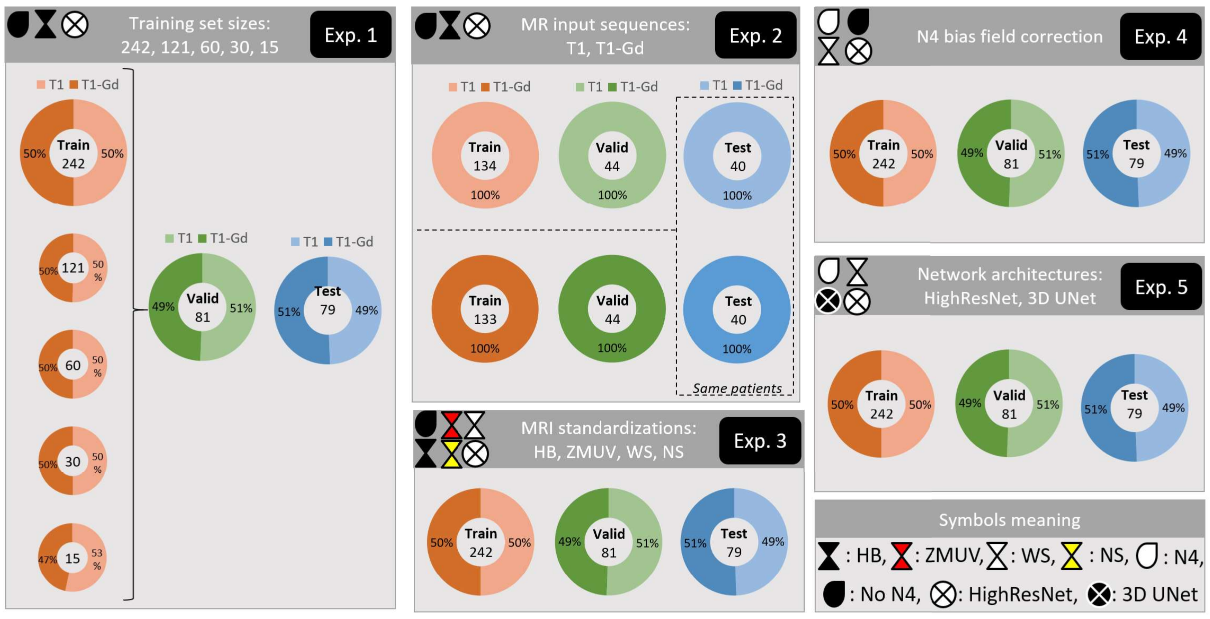
Material and Methods

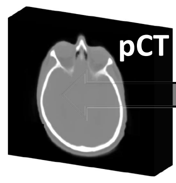


201 couples of CT/T1 and 201 CT/T1-Gd

Preprocessing:

- Rigid registration
- Linear resampling to 1mmx1mmx1mm voxels
- Hounsfield Units (HU) and MRI intensities clip





- Mean Absolute Error (MAE) defined as: $\frac{1}{N} \times \sum_{i=1}^{N} |Intensity\ CT(i) - Intensity\ pCT(i)|$

With N the total number of voxels. MAE was computed in the whole head, air, bone and water areas.

- Dosimetry: Transfer of the initial plan on the pCT. Pencil beam dose calculation performed with iPlan (BrainLab). Calculations of differences in Dose Volume Histograms (DVH) of the planning target volume and 3D global gamma indexes (1%/1mm, 2%/2mm, 3%/3mm) without dose threshold.
- Wilcoxon tests to assess the significance of the observed differences. Threshold set to 0.05.

Results and Discussion

- Experiment 1 proved that generating pCT with all the available cases for training resulted in higher pCT quality (Figure 1).
- Regarding experiment 3, obtained pCT with the HB standardization, are presented in Figure 2. WS slightly outperformed HB, ZMUV and NS in terms of MAE only (Table 1).
- Experiments 2, 4, 5 led to similar performances regardless of the adopted approach (Table 1).

	Head MAE	Head MAE p-values	3%/3mm gamma index	3%/3mm gamma index p-values
Experiment 2	T1 : 84HU +/- 25HU T1-Gd : 87HU +/- 28HU	**p-value = 0.0047**	T1 : 99.84% +/- 0.18% T1-Gd : 99.85% +/- 0.18%	**p-value = 0.044**
Experiment 3	HB : 92HU +/- 23HU ZMUV : 83HU +/- 22HU WS : 78HU +/- 22HU NS : 96HU +/- 23HU	**p-values WS/HB, WS/ZMUV, WS/NS < 0.0001**	HB : 99.86% +/- 0.16% ZMUV : 99.83% +/- 0.19% WS : 99.85% +/- 0.17% NS : 99.86% +/- 0.18%	p-values WS/HB, WS/ZMUV, WS/NS > 0.14
Experiment 4	Without N4 : 78HU +/- 22HU With N4 : 81HU +/- 22HU	**p-value < 0.0001**	Without N4 : 99.85% +/- 0.17% With N4 : 99.83% +/- 0.19%	**p-value = 0.012**
Experiment 5	HighResNet : 81HU +/- 22HU 3D UNet: 90HU +/- 21HU	**p-value < 0.0001**	HighResNet : 99.83% +/- 0.19% 3D UNet: 99.74% +/- 0.24%	**p-value < 0.0001**

Table 1: Head Mean Absolute Error (MAE), 3%/3mm gamma indexes and p-values for experiments 2, 3, 4 and 5.

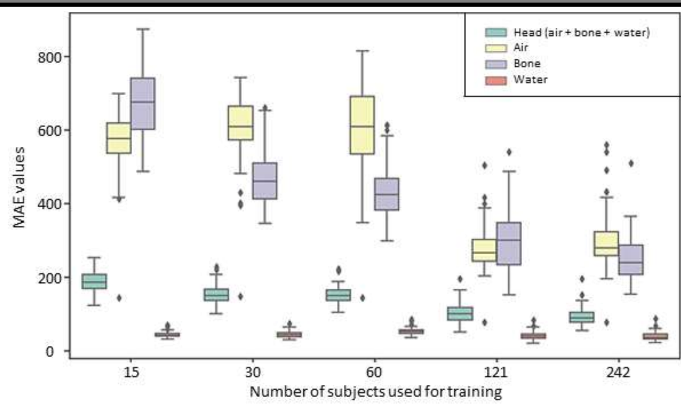


Figure 1: Evolution of the MAE when modifying the number of subjects in the training set during experiment 1.

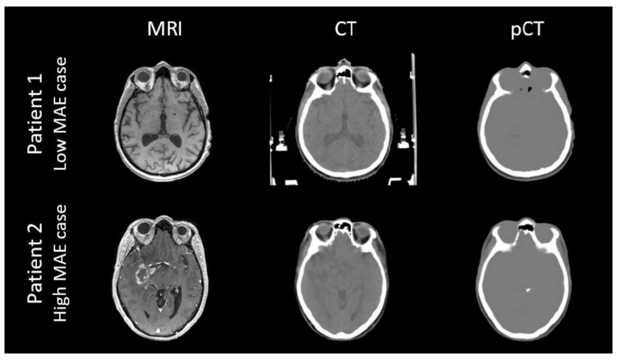


Figure 2: MRI, original CT and pCT with soft tissues windows and levels respectively for Patient 1 (Head MAE = 64HU) and Patient 2 (Head MAE = 110HU) derived from experiment 3 with HB standardization.

Conclusion

The goal of the study was to evaluate the impact of key generation parameters on pCT quality. Results, derived from experiments based on a large cohort composed of 402 patients, proved that the largest training set size provided the best quality pCT. Regarding the MR input sequence, standardization, bias field application and network architecture, no major dosimetry differences were obtained suggesting the clinical equivalence of the studied approaches. Future work will include the generation of pCT combined with a segmentation task and the extension of the presented model to another anatomical site, such as pelvis.

Acknowledgement and References

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 880314.

1. Ulin K, Urie MM, Cherlow JM. Results of a multi-institutional benchmark test for cranial CT/MR image registration. Int J Radiat Oncol Biol Phys. 1 août 2010;77(5):1584-9.
2. Nyúl LG, Udupa JK. On standardizing the MR image intensity scale. Magn Reson Med. déc 1999;42(6):1072-81.
3. Reinhold JC, Dewey BE, Carass A, Prince JL. Evaluating the impact of intensity normalization on MR image synthesis. Med Imaging 2019 Image Process. 2019;10949:109493H.
4. Shinohara R, Sweeney E, Goldsmith J, Shiee N, Mateen F, Calabresi P, et al. Statistical normalization techniques for magnetic resonance imaging. Neuroimage Clin. 15 août 2014;6:9-19.
5. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, et al. N4ITK: Improved N3 Bias Correction - IEEE Journals & Magazine. IEEE Trans Med Imaging. juin 2010;29(6):1310-20.
6. Li W, Wang G, Fidon L, Ourselin S, Cardoso MJ, Vercauteren T. On the Compactness, Efficiency, and Representation of 3D Convolutional Networks: Brain Parcellation as a Pretext Task. In: Niethammer M, Styner M, Aylward S, Zhu H, Oguz I, Yap P-T, et al., éditeurs. Information Processing in Medical Imaging. Springer International Publishing; 2017. p. 348-60. (Lecture Notes in Computer Science).
7. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. In: Ourselin S, Joskowicz L, Sabuncu MR, Unal G, Wells W, éditeurs. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016. Cham: Springer International Publishing; 2016. p. 424-32. (Lecture Notes in Computer Science).