Assessing the impact of key preprocessing concepts on the pseudo CT generation

Emilie Alvarez Andres[#1,2,3], Lucas Fidon[#2,4], Maria Vakalopoulou[4], Marvin Lerousseau[1,3,4], Alexandre Carré[1,3], Roger Sun[1,3,4,5], Guillaume Klausner[1,5], Samy Ammari[6], Nathan Benzazon[1,3], Sylvain Reuzé[1,3], Théo Estienne[1,3,4], Stéphane Niyoteka[1,3], Enzo Battistella[1,3,4], Angéla Rouyar[1,3], Ben Glocker[7], Georges Noël[8], Anne Beaudre[3], Frédéric Dhermain[5], Eric Deutsch[1,5], Nikos Paragios[2], Charlotte Robert[1,3]

[1]U1030 Molecular Radiotherapy, Paris-Sud University - Gustave Roussy - Inserm - Paris-Saclay University, Villejuif, France
[2]TheraPanacea, Paris, France
[3]Department of Medical Physics, Gustave Roussy - Paris-Saclay University, Villejuif, France
[4]MICS Laboratory, CentraleSupélec, Paris-Saclay University, 91190, Gif-sur-Yvette, France
[5]Department of Radiotherapy, Gustave Roussy - Paris-Saclay University, Villejuif, France
[6]Department of Radiology, Gustave Roussy - Paris-Saclay University, Villejuif, France
[7]Biomedical Image Analysis Group, Imperial College, London, United Kingdom
[8]Department of Radiotherapy, Paul Strauss Institute, Strasbourg, France
#: Equal contribution

Introduction: Magnetic Resonance Imaging (MRI) only radiotherapy workflows have been increasingly appealing as they ensure a reduction of the dose received by the healthy tissues and thus a higher patient safety. We used a modified version of the 3D neural network presented in (1) to map any high-resolution T1-weighted MRI (T1) or contrast enhanced T1 weighted MRI (T1Gd) to a pseudo Computed Tomography (pCT). The goal of this study was to quantify the impact of critical preprocessing steps on the 3D pCT generation.

Methods: The global dataset was composed of 402 rigidly-registered pairs of brain CT and T1 or T1Gd. Experiment 1 consisted in evaluating the impact of the MRI input sequence. To this aim, two datasets composed of 222 and 220 T1-only and T1Gd-only MR images respectively were constituted and distributed between the training, validation and test sets. A second experiment was conducted to quantify the impact of the training set size. Five models were used, with 242, 121, 60, 30 and 15 patients respectively in the training set. The validation and test sets remained the same and were composed of 81 and 79 patients respectively. A last experiment to assess the role of MR standardization was performed. 242, 81 and 79 patients were included in the training, validation and test sets respectively. Three different approaches were investigated: no standardization (NS), histogram-based (HB) (2) and zero mean-unit variance (ZMUV) standardizations. To evaluate the quality of all generated pCT, the Mean Absolute Error (MAE) was computed within four areas (whole head, air, bone and water components). T-tests were conducted for statistical evaluation.

Results: For the first experiment, a p-value of 0.72 was achieved between the T1-only and T1Gd-only models suggesting that the two sequences can be used without any distinction. The MAE obtained for the second experiment are presented in Figure 1. We showed that more than 100 patients must be used in the training set to reach optimal performances. The last experiment led to head MAE of 92HU +/- 23HU, 83HU +/- 22HU and 96HU +/- 23HU for the HB, ZMUV, NS approaches respectively. Differences between HB/ZMUV and ZMUV/NS methods were found to be significantly different, with p-values of 0.013 and <0.001 respectively.

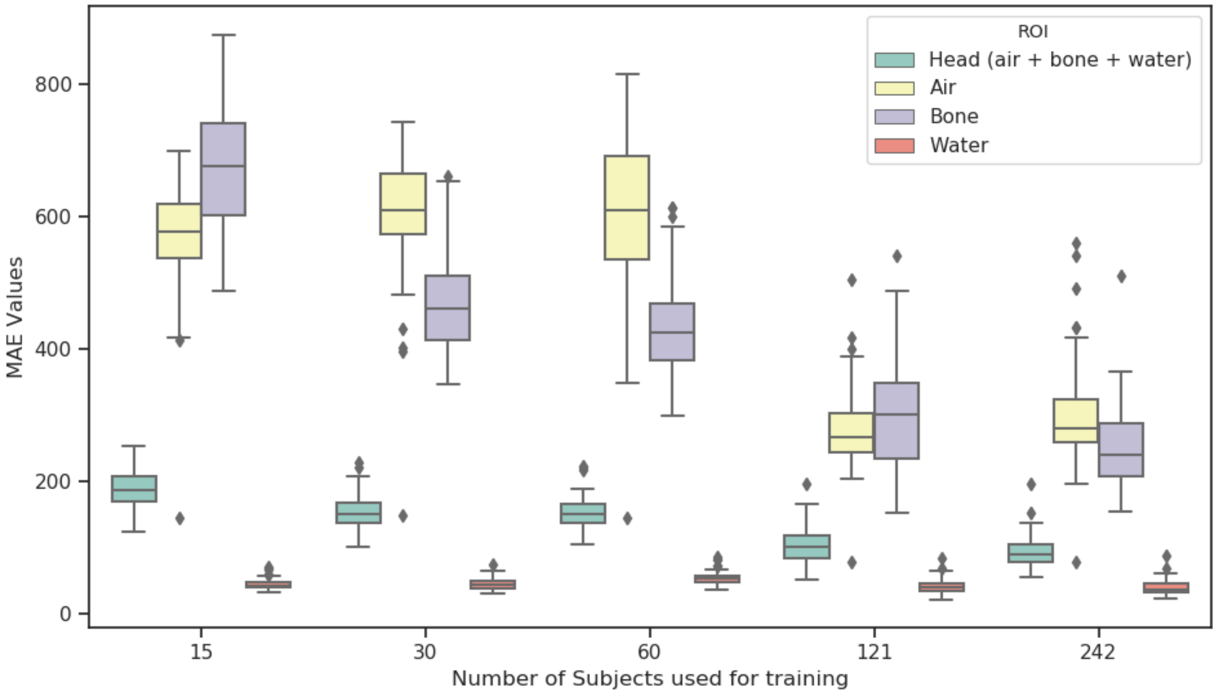

Figure 1: Evolution of the MAE distribution when modifying the number of training subjects.

Conclusion: In this study, pCT were generated from a large cohort (402 patients) with a deep learning 3D architecture, ensuring reliable and accurate images. Based on our patient data, it appeared that ZMUV standardization should be used on T1-weighted images, regardless the sequence, and that the training set should be composed of more than one hundred of patients. These novel guidelines will have a significant impact on the pCT generation optimization.

References

1. 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).

2. Nyúl LG, Udupa JK, Zhang X. New variants of a method of MRI scale standardization. IEEE Trans Med Imaging. févr 2000;19(2):143-50.