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

Authors

Emilie Alvarez Andres[#], Lucas Fidon[#], Maria Vakalopoulou, Marvin Lerousseau, Alexandre Carré, Roger Sun, Anne Beaudre, Eric Deutsch, Nikos Paragios, Charlotte Robert
[#]: Equal contribution

Purpose/Objective

Brain pseudo Computed Tomography (pCT) were generated from Magnetic Resonance Imaging (MRI) using a 3D convolutional neural network.
The contributions of this study are threefold: assessing the best suited MRI input sequence, benchmarking different MRI standardization methods and inferring the minimal number of patients of the training set for a high quality pCT.

Material and Methods

402 institutional brain tumor patients were retrieved yielding to associations of 182 CT/T1 weighted MRI (T1), 181 CT/contrast enhanced T1 weighted MRI (T1-Gd) and 39 CT/T1/T1Gd. These data were used to train, validate and test a modified version of the 3D neural network HighResNet (Li et al., 2017).
First, to assess the most informative MR input sequence, two models were developed either based on T1 MRI sequence only (218 patients) or T1-Gd only (217 patients) cohorts. Then, three standardization strategies, namely Histogram Based (HB), Zero Mean Unit Variance (ZMUV) and No Standardization (NS), were compared based on training, validation and testing sets composed of 242, 81 and 79 patients respectively. Finally, further models were trained on subsets of the training set (242, 121, 60, 30, 15 patients) and compared based on fixed validation and testing sets (81 and 79 patients respectively) to assess the behavior of the subsequent models performance in function of the input size.
Comparisons between the ground-truth CT and pCT were conducted computing the Mean Absolute Error (MAE) within four areas (whole head, air, water and bone), global 1%/1mm, 2%/2mm, 3%/3mm gamma indexes and Dose Volume Histograms (DVH) differences based on planning target volume. Paired samples Wilcoxon tests were performed as statistical analysis.

Results

Figure 1 presents qualitative results. Reported results are presented such as [median, interquartile range]. Head MAE of [80Hounsfield Units (HU), 27HU] and [84HU, 28HU] were achieved for the T1 only and T1-Gd only based experiments respectively. Gamma indexes differences were not found significant. Regarding the three standardization strategies, head MAE equal to [89HU, 27HU], [81HU, 26HU], [92HU, 27HU] were obtained for the HB, ZMUV and NS respectively, proving the significant superiority of the ZMUV approach (p-values≤0.0001). All DVH differences medians were below 0.25%. Finally, Figure 2 presents the MAE distribution computed from the training set size experiment, and suggests the use of at least 121 patients in the training set for this study.

Conclusion

Competitive pCT were generated when combining ZMUV standardization with a training set containing at least about 120 patients. Using T1 only or T1-Gd only MR sequences did not impact the quality of dosimetric maps calculated from pCT.

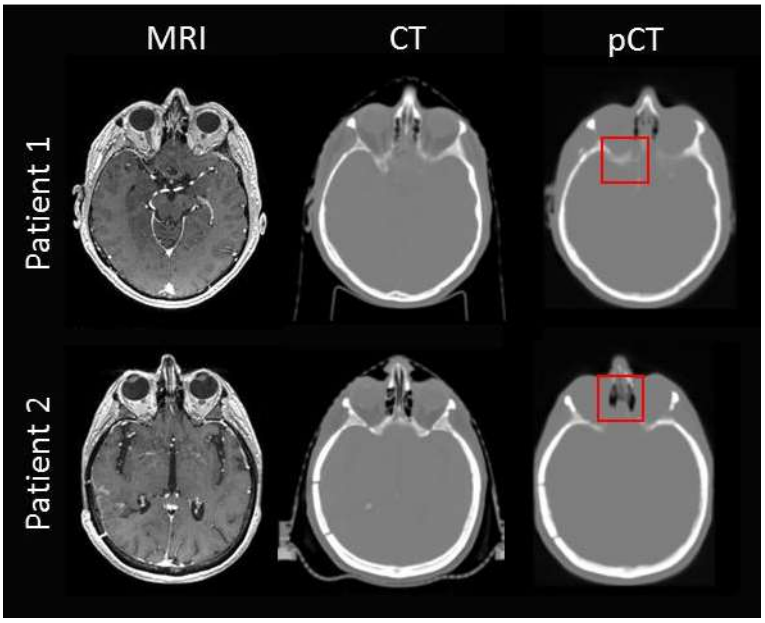

Fig. 1. MRI, ground-truth CT and pCT. The red squares highlight incorrect reconstructed areas.

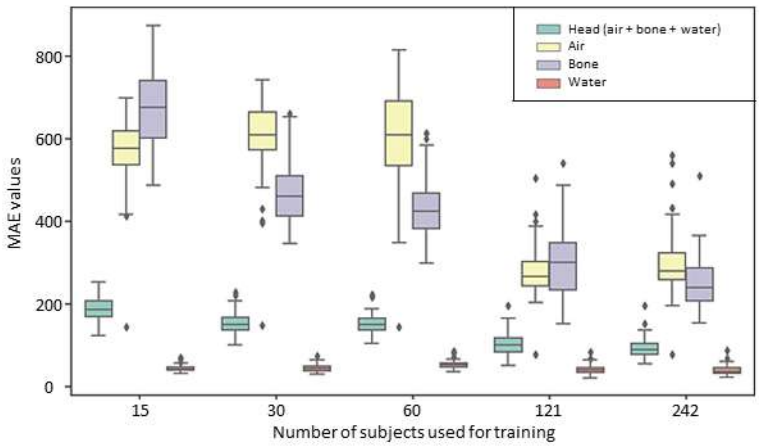

Fig. 2. MAE distribution evolution when varying the number of patients in the training set.