Comparison of synthetic CT generation algorithms for MRI-only radiation planning in the pelvic region

Hossein Arabi, Jason A. Dowling, Ninon Burgos, Xiao Han, Peter B. Greer, Nikolaos Koutsouvelis, and Habib Zaidi†, *IEEE Fellow*

***Abstract*– Accurate radiation dose calculation is a major challenge in magnetic resonance imaging (MRI)-only radiation therapy (RT) treatment planning as the required electron density map is not readily provided by this modality. In this work, a number of state-of-the-art synthetic-CT (sCT) generation methods, exhibited promising results in the literature, were evaluated based on common quantitative metrics and patients dataset. This includes four atlas-based approaches, specifically median of atlas images (A-Median) [1], atlas-based voxel-wise weighting (A-VW) [2], bone enhanced atlas-based voxel-wise weighting (A-Bone) [3], iterative atlas-based voxel-wise weighting (A-Iter) [4], and a method based on deep learning convolutional neural network (DL-CNN) [5]. Automatic organ delineation was performed for bladder, rectum and bone. Overall, A-VW, A-Bone, A-Iter and A-VWexhibited comparable performance while DL-CNN showed slightly better segmentation performance resulting in Dice metrics of 0.93, 0.90, and 0.93, respectively. The dosimetric evaluation demonstrated that A-Median, A-VW, A-Bone, A-Iter and DL-CNN resulted in comparable mean dose errors within organs at risk and target volumes showing less than 1% dose difference against the CT-based RT planning. The two-dimensional gamma analysis performed at 1%/1 mm criterion demonstrated comparable pass rates of 94.99±5.15%, 94.59±5.65%, 93.68±5.53% and 93.10±5.99% for A-Bone, DL-CNN, A-Median and A-Iter, respectively. Whereas A-VW and water-only resulted in pass rates of 86.91±13.50% and 80.77±12.10%, respectively. DL-CNN and advanced atlas-based approaches showed promising dosimetric and segmentation accuracy (DL-CNN is slightly better) suggesting that these methods are able to resolve the challenge of synthetic-CT generation from MR images with clinically acceptable errors.**

I. Introduction

Magnetic resonance imaging (MRI) is gaining momentum nowadays and is increasingly used in radiation therapy (RT) planning due to its ability to resolve soft-tissues in fine details and non-ionizing radiation. However, the major barrier to establish MRI-alone RT planning workflows is the lack of electron density map as MR images do not directly provide photon attenuation coefficients. So far, many approaches have been proposed in the literature that reported only few per cent of dose discrepancy compared to standard CT-based RT planning. In general, the MR-guided synthetic-CT generation methods reported in the literature could be split into three generic categories [6-8]; tissue segmentation: this involves bulk segmentation of the MR images into major tissue classes. Atlas-based method: this approach relies on the alignment of CT atlas images or templates to the target MR image. Machine learning method: this approach attempts to estimate tissues electron densities directly from intensity of MR images. Even though each of these methods exhibited promising results in the literature, there is no comprehensive comparison of these state-of-the-art approaches putting into perspective their key performance parameters. , As such, in this work, we aimed to assess the performance of six promising synthetic-CT generation methods in the pelvic region for the purpose of MRI-only RT planning. An attempt was made to include at least one representative approach from each of the above-mentioned three generic MR-guided sCT generation methods. Four state-of-the-art atlas-based methods were chosen including the median of atlas images (A-Median) [1], atlas-based voxel-wise weighting (A-VW) [2], bone enhanced atlas-based voxel-wise weighting (A-Bone) [3], iterative atlas-based voxel-wise weighting (A-Iter) [4]. The fifth method relies on deep learning convolutional neural network (DL-CNN) [5] to establish direct conversion between MR and CT images. The sixth method is based on tissue segmentation approach using a 2 tissue-class synthetic-CT taking into account only water and air. A unified comparison between these six MRI-based synthetic-CT generation methods was performed based on the same cohort of patients (containing aligned MR and CT images) and common quantitative metrics taking the CT-based RT planning as the reference. These methods are also capable of performing automatic organ contouring (auto-contouring) from the target MR images. Therefore, the accuracy of the organ auto-contouring was evaluated using the standard metrics of segmentation assessment versus manually defined organ contours.

II. Materials and Methods

A. Patient data acquisition

The cohort contains 39 patients that have each one a CT image with 2.5 mm slice thickness and a whole-pelvis T2 weighted MR image aligned to the CT image. The ground-truth contours for rectum, bone and bladder were generated manually

H. Arabi is with the Division of Nuclear Medicine & Molecular Imaging, Geneva University Hospital, CH-1211, Geneva, Switzerland (e-mail: hossein.arabi@unige.ch).
J.A. Dowling is with the CSIRO Australian e-Health Research Centre, Herston, Queensland, Australia (jason.dowling@csiro.au).
N. Burgos is with Inria Paris, Aramis project-team, Institut du Cerveau et de la Moelle épinière, ICM, Inserm U 1127, CNRS UMR (ninon.burgos@inria.fr).
X. Han is with Elekta Inc., Maryland Heights, MO 63043, USA (Xiao.Han@elekta.com).
P. B. Greer is with Calvary Mater Newcastle Hospital, Waratah, New South Wales, Australia (peter.greer@newcastle.edu.au).
N. Koutsouvelis is with the Division of Radiation Oncology, Geneva University Hospital, CH-1211 Geneva, Switzerland (Nikolaos.Koutsouvelis@hcuge.ch).
†H. Zaidi is with the Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, CH-1211 Geneva, Switzerland, Geneva University Neurocenter, CH-1205 Geneva, Switzerland. He is also with the Department of Nuclear Medicine and Molecular Imaging, University of Groningen, University Medical Center Groningen, Groningen, Netherlands, and the Department of Nuclear Medicine, University of Southern Denmark, DK-500, Odense, Denmark (e-mail: habib.zaidi@hcuge.ch).

978-1-5386-8494-8/18/$31.00 ©2018 IEEE

on the MR images by three experienced observers and the majority voting scheme was used to combine the decisions.

B. Synthetic-CT generation methods

(I) Median of atlas images (A-Median) [1]: Given a number of CT-MRI pairs registered to the target MR image, A-Median sCT is generated through calculation of the median voxels value across the entire atlas CT images. The calculation of median value was shown to result in a more accurate CT value estimation compared to average value.

(II) Atlas-based voxel-wise weighting (A-VW) [2]: In this method the CT-MRI atlas dataset is aligned to the test (target) MR image followed by local weighted voting based on the morphological similarity between MR images in the atlas dataset and the target MR image to combine aligned atlas CT images to estimate the synthetic-CT image.

(III) Bone enhanced atlas-based voxel-wise weighting (A-Bone) [3]: Synthetic-CT is generated through a two-phase (bone extraction and atlas fusion) atlas-based approach attempting to maximize bone identification accuracy. Phase congruency map (providing local morphologic features) was exmloyed to compute the image morphology likelihood between the atlas and target MR images. This approach exhibited robustness to noise and inter-subject MR intenstiy non-uniformity.

(IV) Iterative atlas-based voxel-wise weighting (A-Iter) [4]: An iterative multi-atlas information propagation scheme is used to jointly estimate a synthetic-CT and segment organs from structural MR images. This method is intended to combine the organ auto-contouring and synthetic-CT generation in a single pipeline to take advantage of the existing correlation between them.

(V) Deep Learning Convolutional Neural Network (DL-CNN) [5]: Synthetic-CT and automated organ segmentation are generated using a particular deep convolutional neural network architecture having 27 convolutional layers and 35 million free parameters. The organ auto-contouring is performed separately (from the synthetic-CT generation) by the DL-CNN approach, as such, the training of the DL-CNN was repeated uniquely for each organ.

(VI) Water-only: The contour of the body was first identified on the target MR image then, the Hounsfield Unit of HU=0 (soft-tissue or water) and HU= -1000 (background air) were assigned to all the voxels within and outside the body contour, respectively.

C. Evaluation

Automatic organ delineation (A) was evaluated against the ground-truth (manual) organ contours *(R)* using DSC and the mean absolute surface distance (MASD). Moreover, the mean absolute error (MAE) and mean error (ME) were computed between the reference CT images (R_{CT}) and sCTs (A_{CT}) within each organ.

$$DSC(A,R) = \frac{2|A \cap R|}{|A|+|R|}$$

$$MASD(A,R) = \frac{d_{ave}\ (S_A S_R) + d_{ave}\ (S_R S_A)}{2}$$ ($d_{av.}$:Euclidean distance, S: Surface)

$$ME_{CT} = \frac{1}{P}\sum_{i=1}^{P} A_{CT}(i) - R_{CT}(i) \qquad MAE_{CT} = \frac{1}{P}\sum_{i=1}^{P} |A_{CT}(i) - R_{CT}(i)|$$

Treatment planning was performed based on the volumetric-modulated arc therapy intended to give the total dose of 36.25 Gy to the target volumes (planning target volume (PTV)). The difference of absorbed dose estimated on the ground truth CT and the sCT images was calculated using mean (ME) and absolute mean errors (MAE) indices within organs at risk (OARs) and target volumes.

$$ME_{Dose}\ (\%) = 100 \times \frac{Dose_{sCT} - Dose_{CT}}{Dose_{CT}}$$

$$MAE_{Dose}\ (\%) = 100 \times \frac{|Dose_{sCT} - Dose_{CT}|}{Dose_{CT}}$$

Moreover, a two-dimensional Gamma analysis using the 3%/3 mm, 2%/2 mm and 1%/1 mm dose difference/distance to agreement criterion was performed between dose maps calculated on the reference CT and sCT images.

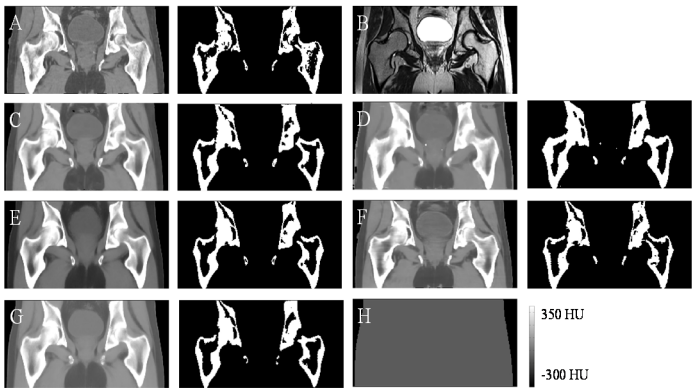

Fig 1. A) Reference CT, B) Target MRI, C) Pseudo-CT of A-Iter, D) A-VW, E) A-Bone, F) DL-CNN, G) A-Median and H) 2 tissue-class (Water-only).

III. Results and Discussion

Figure 1 depicts the representative views of the various sCT images along with the ground truth CT and target MR images. In addition, next to each CT image the binary bone maps extracted by operating intensity thresholding (HU > 140 HU) on the sCT images are shown.

The results of the quantitative evaluation of the organ auto-contouring are presented in Table 1. In general, DCNN approach exhibited slightly more accurate automated organ delineation as well as CT value estimation.

TABLE 1. Accuracy of organ contouring and synthetic CT estimation.

Bladder	**A-Median**	**A-VW**	**A-Bone**	**A-Iter**	**DL-CNN**
DSC	0.82±0.20	0.86±0.12	0.86±0.13	0.87±0.10	0.93±0.17
MASD(mm)	7.01±4.17	5.10±4.57	7.88±4.78	7.56±4.42	2.36±2.44
ME (HU)	-1.5±20.2	-2.9±18.7	8.1±17.6	7.7±16.1	-1.8±12.9
MAE (HU)	30.0±17.6	24.1±13.6	26.4±12.7	25.2±10.1	18.4±6.6

Rectum	**A-Median**	**A-VW**	**A-Bone**	**A-Iter**	**DL-CNN**
DSC	0.81±0.08	0.84±0.06	0.84±0.70	0.84±0.06	0.90±0.04
MASD(mm)	5.03±2.72	2.37±1.34	4.95±2.39	4.81±2.22	2.09±1.11
ME (HU)	37.6±84.9	6.9±81.7	27.6±90.5	-30.3±94.6	22.7±84.8
MAE (HU)	93.5±71.2	88.1±60.8	100.0±62.0	114.8±63.6	78.3±69.2

Bone	**A-Median**	**A-VW**	**A-Bone**	**A-Iter**	**DL-CNN**
DSC	0.88±0.04	0.91±0.03	0.92±0.02	0.92±0.02	0.93±0.02
MASD(mm)	3.73±0.58	1.45±0.47	1.94±0.45	2.07±0.43	3.51±3.92
ME (HU)	-32.9±55.4	-6.4±46.5	26.6±56.7	19.5±46.3	-4.1±40.7
MAE (HU)	161.1±30.0	134.2±24.0	163.8±25.0	130.2±23.4	119.9±22.6

Table 2 shows the mean absorbed dose errors measured within the OARs and target organs. The estimated mean absorbed doses for the target volumes and OARs showed insignificant difference between different methods resulting in less than 1% error compared to the CT-based RT planning.

TABLE 2. Relative error (mean(%) ±SD) of mean absorbed dose.

	A-Median ME±Std MAE±Std	**A-VW** ME±Std MAE±Std	**A-Bone** ME±Std MAE±Std	**A-Iter** ME±Std MAE±Std	**DL-CNN** ME±Std MAE±Std	**Water-only** ME±Std MAE±Std
Bladder	0.09±0.66 0.52±0.41	-0.55±0.76 0.75±0.56	-0.05±0.6 0.47±0.37	0.14±0.68 0.53±0.44	-0.02±0.57 0.41±0.39	-0.06±1.10 0.82±0.73
Rectum	0.28±0.78 0.66±0.49	-0.59±0.74 0.70±0.63	0.09±0.72 0.56±0.46	0.15±0.82 0.63±0.53	0.06±0.73 0.50±0.52	0.47±1.45 0.98±1.16
Left HOF	0.06±0.49 0.40±0.27	-0.67±0.56 0.72±0.50	-0.14±0.48 0.40±0.30	0.07±0.48 0.38±0.31	-0.07±0.43 0.32±0.30	1.27±0.69 1.27±0.69
Right HOF	0.09±0.48 0.40±0.28	-0.63±0.58 0.69±0.51	-0.14±0.48 0.39±0.31	0.04±0.49 0.39±0.30	-0.08±0.42 0.32±0.28	1.29±0.67 1.29±0.67
CTV	0.22±0.73 0.62±0.42	-0.73±0.73 0.83±0.61	-0.05±0.69 0.55±0.42	0.24±0.74 0.63±0.43	-0.02±0.67 0.51±0.43	0.86±1.21 1.14±0.95
PTV	0.23±0.72 0.61±0.44	-0.72±0.71 0.81±0.60	-0.04±0.68 0.54±0.41	0.18±0.72 0.60±0.42	-0.01±0.64 0.47±0.43	0.82±1.22 1.10±0.97

The results of two-dimensional gamma analysis of iso-center dose distributions are summarized in table 3. At the 3%/3 mm criterion, all the methods exhibited similar performance; however, considering the criteria of 1%/1 mm, significant differences appeared among the evaluated methods. A-Bone, DL-CNN, A-Median and A-Iter methods led to more than 93% of pass rate at 1%/1 mm threshold.

TABLE 3. GAMMA ANALYSIS COMPARING THE CT AND SYNTHETIC RT PLANS.

	A-Median (mean±Std)	**A-VW** (mean±Std)	**A-Bone** (mean±Std)	**A-Iter** (mean±Std)	**DL-CNN** (mean±Std)	**Water-only** (mean±Std)
3%/3 mm	98.96±0.78	98.41±1.56	99.51±0.32	98.96±0.57	99.22±0.46	98.22±1.75
2%/2 mm	97.92±1.49	96.93±2.69	98.84±0.48	97.99±1.02	98.47±0.68	95.38±5.17
1%/1 mm	93.68±5.53	86.91±13.50	94.99±5.15	93.10±5.99	94.59±5.65	80.77±12.10

IV. CONCLUSIONS

In this study, a collection of state-of-the-art MR-guided synthetic CT estimation methods were evaluated using a common cohort of patients for the use of MR-only radiation planning. Overall, DCNN appeared to have slightly better performance owing to accurate organ segmentation and small dosimetric errors. Given the results obtained in this study, the challenge of synthetic-CT generation from MR images could be sufficiently resolved generating clinically tolerable diametric errors.

ACKNOWLEDGEMENTS

This work was supported by the Swiss National Science Foundation under grant SNSF 320030_176052 and the Swiss Cancer Research Foundation under Grant KFS-3855-02-2016.

REFERENCES

[1] J. Sjölund, D. Forsberg, M. Andersson, and H. Knutsson, "Generating patient specific pseudo-CT of the head from MR using atlas-based regression," *Phys Med Biol,* vol. 60, p. 825-839, 2015.

[2] J. A. Dowling, J. Sun, P. Pichler, D. Rivest-Hénault, S. Ghose, H. Richardson*, et al.*, "Automatic substitute computed tomography generation and contouring for Magnetic Resonance Imaging (MRI)-alone external beam radiation therapy from standard MRI sequences," *Int J Radiat Oncol Biol Phys,* vol. 93, p. 1144-1153, 2015.

[3] H. Arabi, N. Koutsouvelis, M. Rouzaud, R. Miralbell, and H. Zaidi, "Atlas-guided generation of pseudo-CT images for MRI-only and hybrid PET–MRI-guided radiotherapy treatment planning," *Phys Med Biol,* vol. 61, p. 6531-6552, 2016.

[4] N. Burgos, F. Guerreiro, J. McClelland, B. Presles, M. Modat, S. Nill*, et al.*, "Iterative framework for the joint segmentation and CT synthesis of MR images: application to MRI-only radiotherapy treatment planning," *Phys Med Biol,* vol. 62, p. 4237-4253, 2017.

[5] X. Han, "MR-based synthetic CT generation using a deep convolutional neural network method," *Med Phys,* vol. 44, p. 1408-1419, 2017.

[6] A. Mehranian, H. Arabi, and H. Zaidi, "Vision 20/20: Magnetic resonance imaging-guided attenuation correction in PET/MRI: Challenges, solutions, and opportunities," *Med Phys,* vol. 43, pp. 1130-1155, 2016.

[7] E. Johnstone, J. J. Wyatt, A. M. Henry, S. C. Short, D. Sebag-Montefiore, L. Murray*, et al.*, "Systematic review of synthetic computed tomography generation methodologies for use in magnetic resonance imaging–only radiation therapy," *Int J Radiat Oncol Biol Phys,* vol. 100, p. 199-217, 2018.

[8] J. M. Edmund and T. Nyholm, "A review of substitute CT generation for MRI-only radiation therapy," *Radiat Oncol,* vol. 12, p. 28, 2017.