Generalizable cone-beam CT esophagus segmentation using in-silico data augmentation

JULY 12–16 2020 VIRTUAL JOINT AAPM | COMP MEETING EASTERN TIME [GMT-4]

S R Alam[1*], T Li[1], S Zhang[2], D. Lee[1], P Zhang[1], S Nadeem[1]
[1]Department of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, USA
[2]Peking University Cancer Hospital & Institute, Beijing, China

Memorial Sloan Kettering Cancer Center

INTRODUCTION/PURPOSE

- Sparing esophagus as the major organ at risk (OAR) is critical in radiotherapy (RT) of lung cancer patients to minimize radiation-induced toxicities such as acute esophagitis (develops in 50% of the patients)
- Cross-modality automated segmentation of esophagus in cone-beam CT (CBCT) and planning CT (pCT) using 3D convolutional neural networks for image guided/adaptive RT
- Semantic in-silico modeling (image-driven simulation) of data augmentation by inducing the noise/scatter artifacts from CBCTs to their pCTs

METHOD

- 60 lung patients treated via IMRT and had weekly CBCTs.
- 7 variations of scatter artifacts/noise were extracted from the week 1 CBCTs using different parameters of Power-Law Adaptive Histogram Equalization that contained the highest to the smoothest frequency components. The parameters were i)α=0.5 β=1 ii)α=1 β=0.5 iii)α=0.5 β=0.5 iv)α=1 β=0 v)α=0.5 β=0 vi)α=0 β=1 and vii)α=0 β=0.5.
- Extracted CBCT artifacts were added to their corresponding pCT and were reconstructed using iterative ordered-subset simultaneous algebraic reconstruction technique to generate **pseudo-CBCTs (ps-CBCT)**. (Fig.1)
- The ps-CBCTs were quantitatively evaluated against their ground-truth CBCT using structure similarity index (SSIM), root mean square error (RMSE), cross correlation (CC) and universal quality index (UQI).
- 3D-UNet models with a multi-objective loss function (dice coefficient and binary cross-entropy) were fed using i)ps-CBCTs ii)pCT and iii)w1 CBCT images by 5 fold splitting of training/testing cases (80/20) for esophagus segmentation using planning esophagus contours as ground-truth. (Fig.2)
- The models were externally validated on the weekly CBCTs and pCTs using dice coefficient (DSC) and Sensitivity between the physician-contoured and UNet-segmented esophagus.

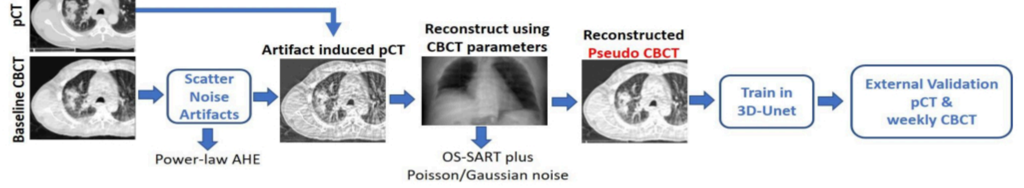

Figure 1. Entire workflow for generating pseudo-CBCTs trained with 3D-UNet.

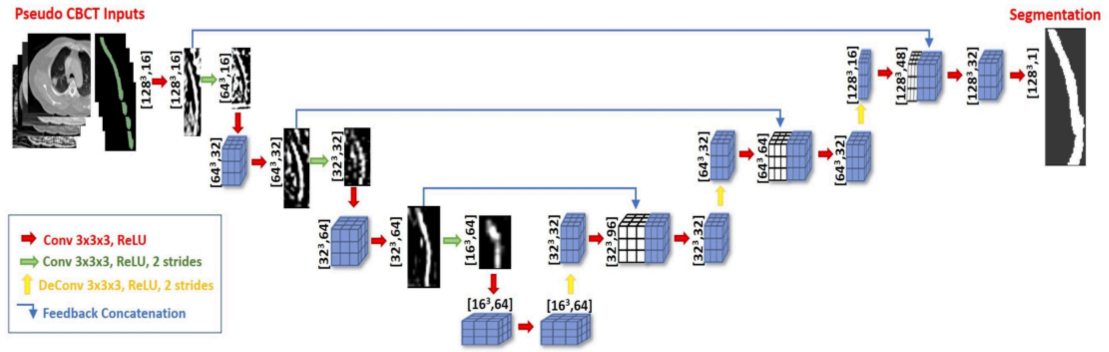

Figure 2. Proposed 3D-UNet architecture.

RESULTS

- The best reconstructed ps-CBCTs (α=0.5 β=1) had average SSIM=0.89, RMSE=0.05, CC=0.97 and UQI=0.95 in the cohort and the worst (α=0 β=0.5) had SSIM=0.44, RMSE=0.14, CC=0.81 and UQI=0.74 (Fig. 3)
- pCT, week1 and week2 CBCTs along with their best ps-CBCT are shown in Fig.4 for two typical cases. Validation results for all three different models are presented in Table 1.
- The proposed model trained using ps-CBCT images segmented esophagus on ps-CBCTs, weekly CBCTs and pCT images with high DSC of 0.74±0.03, 0.72±0.05, and 0.77±0.04, respectively. (Table 1)
- The pCT model could segment esophagus on pCT images with high DSC 0.76 but failed to segment on w1 CBCT (DSC of 0.63).

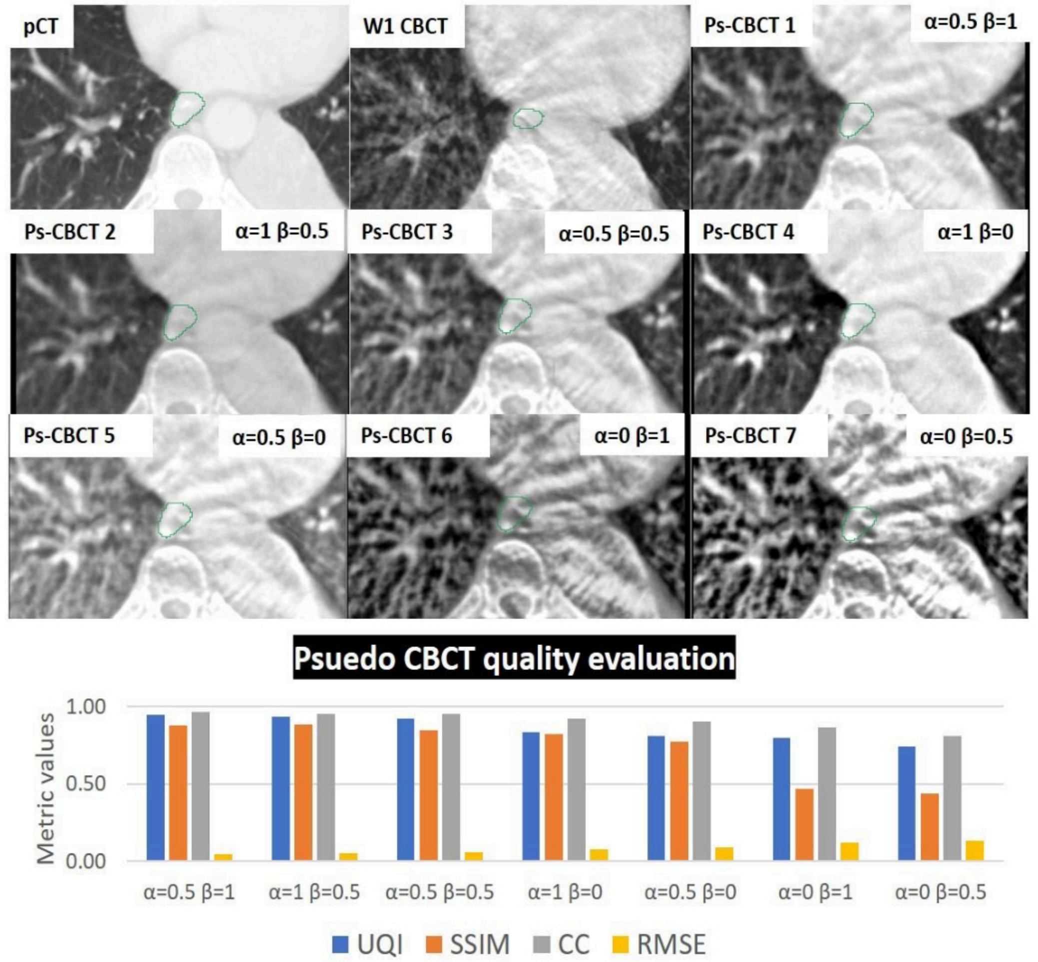

Figure 3. Variation of ps-CBCTs images along with their pCT and the ground-truth week1 CBCT. ps-CBCTs are shown in the order of the highest to the lowest similarity with the week1 CBCT. Green contours are ground-truth esophagus contours. Quantitative ps-CBCT quality evaluation is presented at the bottom.

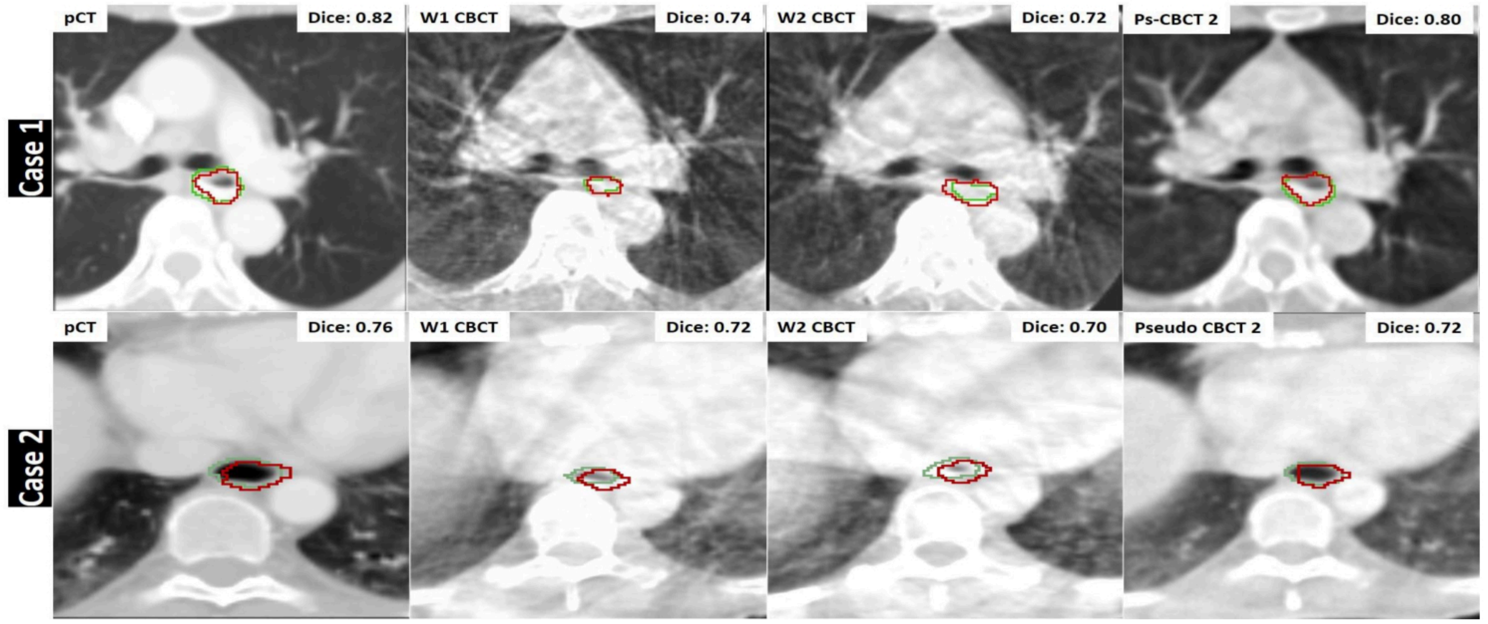

Figure 4. pCT, week1 and week2 CBCTs along with their best ps-CBCT are shown for two typical cases. For each case, green and red contours are ground-truth and segmented esophagus contours, respectively.

Table 1. Dice similarity coefficient and sensitivity results.

	Models/Test data	Ps-CBCT	pCT	w1 CBCT	w2 CBCT	w3 CBCT	w4 CBCT	w5 CBCT	w6 CBCT
DSC	***Ps-CBCT***	0.74±0.03	0.77±0.04	0.72±0.05	0.71±0.04	0.70±0.05	0.69±0.06	0.71±0.05	0.71±0.06
	w1 CBCT	-	0.68±0.07	0.70±0.06	0.68±0.07	0.68±0.08	0.68±0.07	0.68±0.07	0.68±0.09
	pCT	-	0.76±0.05	0.63±0.07	-	-	-	-	-
Sensitivity	***ps-CBCT***	0.79±0.07	0.78±0.08	0.83±0.07	0.82±0.08	0.81±0.09	0.83±0.08	0.83±0.07	0.81±0.07
	w1 CBCT		0.64±0.11	0.73±0.07	0.72±0.09	0.70±0.12	0.70±0.12	0.69±0.10	0.68±0.13
	pCT		0.80±0.08	0.69±0.09	-	-	-	-	-

CONCLUSIONS

- Our image driven in-silico data augmentation spans the realistic noise/artifact spectrum across patient CBCT/pCT data and can generalize well across modalities, eventually improving the accuracy of treatment setup and response analysis.
- 3D-UNet model trained on the pseudo-CBCTs was robust and generalizable enough to produce results on the pCTs, achieving 0.77 dice overlap against the previous best of 0.72 [1].
- 3D-UNet model trained on more realistic artifact-induced pCTs, could segment esophagus on both weekly CBCTs and pCTs with high accuracy for longitudinal imaging studies. The model has a potential to segment any OAR and therefore, can be used as a cross-modality segmentation tool to provide image guidance.

REFERENCES

1. Yang, J., Veeraraghavan, H., Armato III, et al.: Autosegmentation for thoracic radiation treatment planning: A grand challenge at aapm 2017. Medical physics 45(10), 4568–4581 (2018)

CONTACT INFORMATION

*Sadegh R Alam, Ph.D., riyahiam@mskcc.org