Evaluation of Deep Learning-Based Auto-Segmentation of Target Volumes and Organs-at-Risk in Breast Cancer Patients

Seung Yeun Chung[1,2], Jee Suk Chang[1,*], Min Seo Choi[1], Yongjin Chang[3], Byong Su Choi[1], Jaehee Chun[1], Ki Chang Keum[1], Jin Sung Kim[1,*], and Yong Bae Kim[1]

1Department of Radiation Oncology, Yonsei University College of Medicine, Seoul, Korea; 2Department of Radiation Oncology, Ajou University School of Medicine, Suwon, Korea; 3CorelineSoft, Co., Ltd, Seoul, Korea

*Corresponding authors

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

INTRODUCTION

- In breast cancer patients receiving radiotherapy, **accurate target delineation** and reduction of radiation doses to the nearby normal organs is important.
- However, manual clinical target volume (CTV) and organs-at-risk (OAR) segmentation for treatment planning **increases physicians' workload and inter-physician variability considerably**.

AIM

- Here, we evaluated the **potential benefits of deep learning-based auto-segmented contours** by comparing them to manually delineated contours for breast cancer patients.

METHOD

- CTVs for **bilateral breasts, regional lymph nodes, and OARs (including heart, lung, esophagus, spinal cord, and thyroid)** were manually delineated on planning computed tomography scans of 55 breast cancer patients who received breast-conserving surgery.
- Subsequently, a **two-stage convolutional neural network algorithm** was used.
- **Quantitative metrics**, including the Dice similarity coefficient (DSC) and 95% Hausdorff distance, and **qualitative scoring** by 2 panels from 10 institutions were used for analysis.
- **Inter-observer variability** and delineation time were assessed.
- Furthermore, **dose-volume histograms and dosimetric parameters** were also analyzed using another set of patient data.

RESULTS

Quantitative metrics

- Examples of deep learning-based auto-segmentation and manual contours are shown in **Figure 1.**
- **Table 1** compares the auto-segmented contours and manual contours for OARs and CTVs using mean DSC and 95% HD.

Qualitative scoring

- For qualitative scoring, two panels, an expert breast cancer radiation oncologist panel (n = 11) and a non-expert panel that included residents and radiation oncologists whose specialty is not breast cancer (n = 15), answered the following questions after watching an example video:
  1) What score would you give for the differences between manually delineated contours and auto-segmentation contours? (Difference scores)
  2) How much do you think auto-segmentation would assist you in real-world clinical practice? (Assistance scores)

  → Answer: 0 (worst) to 10 (best)
- The results were all acceptable as shown in **Figure 2.**

Inter-observer variability

- OARs and CTVs were delineated by three board-certified radiation oncologists for a randomly selected patient.
- For OAR, only the heart and lungs showed a DSC above 0.80, whereas the other organs showed DSCs lower than 0.80.
- For CTV, although breast CTV showed an acceptable mean DSC of 0.85, other CTVs such as CTVn_L1, L2, L3, CTVn_IMN, CTVn_L4, and CTVn_SCL showed poor results, with mean DSC ranging from 0.45 to 0.75.

Dosimetric analysis

- As for DVH, most manual and auto-segmented contours were similar, except for a minor difference for the spinal cord.
- As for regional LN RT plans, there was a considerable difference in the coverage for regional nodal contours such as axillary lymph node levels 1, 2, 3, and IMN.
- Various dosimetric parameters for OARs such as heart, lung, esophagus, and spinal cord were analyzed.
- The mean absolute differences for all parameters were minimal, showing the efficacy of auto-segmented contours.

Figure 1. *Example of deep learning-based auto-segmentation (green) and manual contours (red)*

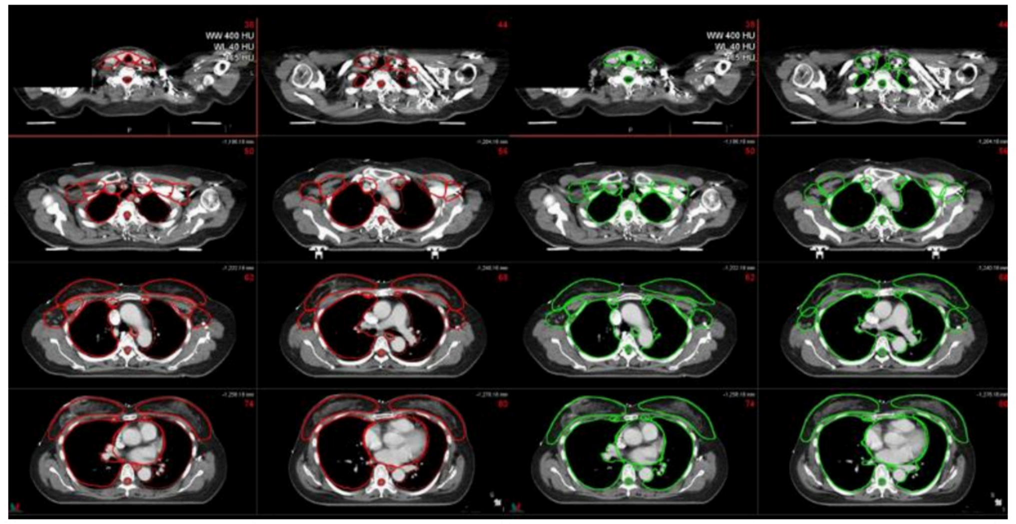
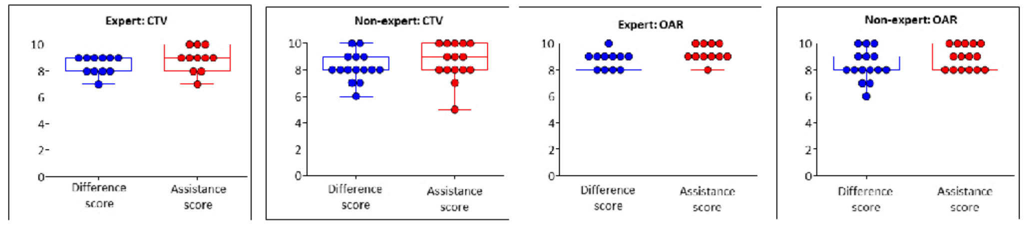


Table 1. *Comparison of deep learning auto-segmentation and manual contours of organs-at-risk and target volumes*

	DSC	STD	95% HD (mm)	STD (mm)
Organs-at-risk				
Heart	0.95	0.01	5.83	2.73
Rt Lung	0.98	0.01	5.17	2.82
Lt Lung	0.97	0.01	3.18	1.03
Thyroid	0.87	0.03	3.57	3.07
Esophagus	0.82	0.02	3.64	0.68
Spinal cord	0.82	0.05	2.59	1.01
Target				
CTVp_breast	0.91	0.02	7.93	2.44
CTVn_L1	0.79	0.05	12.62	9.27
CTVn_L2	0.81	0.04	7.51	2.28
CTVn_L3	0.72	0.07	7.02	2.30
CTVn_IMN	0.75	0.05	6.60	2.95
CTVn_L4	0.72	0.11	7.63	3.97
CTVn_SCL RTOG	0.78	0.06	7.55	3.06

** CTVs of right and left breasts (CTVp_breast); axillary levels 1, 2, and 3 (CTVn_L1, L2, L3); internal mammary chain (CTVn_IMN); lymph node level 4 by ESTRO guidelines (CTVn_L4); supraclavicular lymph node by RTOG guidelines (CTVn_SCL)*

Figure 2. *Qualitative scoring by expert (n = 11) and non-expert panel (n = 15) for difference and assistance scores*



CONCLUSIONS

- In conclusion, **the feasibility of deep learning-based auto-segmentation was demonstrated** in this study for breast cancer patients receiving RT after BCS.
- Although deep learning-based auto-segmentation cannot serve as a substitute for the experience of radiation oncologists, it has **potential to serve as a useful tool** in assisting them.

ACKNOWLEDGEMENTS

N/A

REFERENCES

1. Li XA, Tai A, Arthur DW, et al. Variability of target and normal structure delineation for breast cancer radiotherapy: an RTOG Multi-Institutional and Multiobserver Study. International Journal of Radiation Oncology* Biology* Physics 2009;73:944-951.
2. Offersen BV, Boersma LJ, Kirkove C, et al. ESTRO consensus guideline on target volume delineation for elective radiation therapy of early stage breast cancer. Radiotherapy and oncology 2015;114:3-10.
3. Tan M, Le QV. Efficientnet: Rethinking model scaling for convolutional neural networks. arXiv preprint arXiv:1905.11946 2019.
4. Angermueller C, Parnamaa T, Parts L, et al. Deep learning for computational biology. Mol Syst Biol 2016;12:878.
5. Mamoshina P, Vieira A, Putin E, et al. Applications of Deep Learning in Biomedicine. Mol Pharm 2016;13:1445-1454.
6. Men K, Zhang T, Chen X, et al. Fully automatic and robust segmentation of the clinical target volume for radiotherapy of breast cancer using big data and deep learning. Phys Med 2018;50:13-19.
7. Anders LC, Stieler F, Siebenlist K, et al. Performance of an atlas-based autosegmentation software for delineation of target volumes for radiotherapy of breast and anorectal cancer. Radiother Oncol 2012;102:68-73.
8. Comparing deep learning-based auto-segmentation of organs at risk and clinical target volumes to expert inter-observer variability in radiotherapy planning. Radiother Oncol 2020;144:152-158.

CONTACT INFORMATION

First author: Seung Yeun Chung, email: chungsy@aumc.ac.kr