Radiotherapy and Oncology xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original article

Clinical evaluation of atlas and deep learning based automatic contouring for lung cancer

Tim Lustberg [a,*], Johan van Soest [a], Mark Gooding [b], Devis Peressutti [b], Paul Aljabar [b], Judith van der Stoep [a], Wouter van Elmpt [a], Andre Dekker [a]

[a] Department of Radiation Oncology (MAASTRO), GROW School for Oncology and Developmental Biology, Maastricht University Medical Centre+, The Netherlands; [b] Mirada Medical Ltd., Oxford, United Kingdom

ARTICLE INFO

Article history:
Received 26 July 2017
Received in revised form 16 November 2017
Accepted 21 November 2017
Available online xxxx

Keywords:
Lung cancer
Organs at risk
Radiotherapy
Atlas contouring
Deep learning contouring

ABSTRACT

Background and purpose: Contouring of organs at risk (OARs) is an important but time consuming part of radiotherapy treatment planning. The aim of this study was to investigate whether using institutional created software-generated contouring will save time if used as a starting point for manual OAR contouring for lung cancer patients.

Material and methods: Twenty CT scans of stage I–III NSCLC patients were used to compare user adjusted contours after an atlas-based and deep learning contour, against manual delineation. The lungs, esophagus, spinal cord, heart and mediastinum were contoured for this study. The time to perform the manual tasks was recorded.

Results: With a median time of 20 min for manual contouring, the total median time saved was 7.8 min when using atlas-based contouring and 10 min for deep learning contouring. Both atlas based and deep learning adjustment times were significantly lower than manual contouring time for all OARs except for the left lung and esophagus of the atlas based contouring.

Conclusions: User adjustment of software generated contours is a viable strategy to reduce contouring time of OARs for lung radiotherapy while conforming to local clinical standards. In addition, deep learning contouring shows promising results compared to existing solutions.

The contouring of organs at risk (OAR) and target volumes is an important aspect of treatment planning in radiation oncology. This process is time consuming and the quality of contours depends on the skill level of the observer [1]. Automatic contouring software could potentially speed up the process and improve consistency between observers. There are a number of commercially available products but these are not frequently used in clinical practice [2].

In the last two decades a lot of effort was put into learning new ways to recognize structures in a range of different imaging modalities (CT, PET, and MRI). Approaches range from knowledge-based algorithms such as atlas-based contouring, machine learning and statistical shape and appearance models; region-based methods such as adaptive thresholding, graph cuts and watershed contouring; or a combination of the knowledge- and region-based methods [2]. In products commercially available in 2014, all vendors use a form of atlas-based contouring and approximately half complement this with a model-based method but these are generally limited to certain OARs [2]. Recently, machine learning techniques, and deep learning methods in particular, have become popular for a wider range of tasks. These approaches, based on artificial neural networks, have shown outstanding capabilities, outperforming most classification and regression methods to date. The main advantage of deep learning methods is the ability to automatically learn the most suitable data representation for the task at hand.

The clinical applicability of automatic contouring software is well-reported for regions such as head and neck, breast, and abdomen [2–6]. For the lung, there are a number of studies reporting automatic contouring [7–14]. Some focus solely on a single method and OAR or the Gross Tumor Volume (GTV) and only evaluate the accuracy of the method without any clinical implications such as time gain. Two studies address the usability of atlas-based contouring for the thorax OARs [9,14]. Dolz et al. provides a framework for a region-based contouring technique in clinical practice [13], they advise further investigation into more accurate atlas selection methods to improve the clinical usability. All these methods should be followed by manual correction of the imperfections of the software contouring with the techniques currently available

* Corresponding author.
E-mail address: tim.lustberg@maastro.nl (T. Lustberg).

https://doi.org/10.1016/j.radonc.2017.11.012
0167-8140/© 2017 The Authors. Published by Elsevier Ireland Ltd.

to assess the clinical usability of these methods. In this study, we hypothesize that using a software-generated contour created with an institution specific model as a starting point for OAR contouring will reduce contouring time for lung cancer patients in a clinical setting. For this we evaluate two contouring methods, atlas-based contouring and one novel method based on deep learning in a clinical representative scenario.

Materials and methods

Atlas-based contours

A commercial atlas-based contouring software (Mirada RTx 1.6 and Workflow Box 1.4, Mirada Medical Ltd., Oxford, United Kingdom) was used to automatically generate contours of the OARs. The atlas-based contouring employed 20 stage I NSCLC patients collected from clinical practice at our institute with minimal geometric distortions and small lesion volumes. These atlas patients were contoured by a senior radiotherapy technician specialized in the thorax region using institutional guidelines and carefully inspected by radiation oncologist for correctness.

Deep learning contours

A prototype of deep learning contouring software ("Mirada DLC Expert" prototype, Mirada Medical Ltd., Oxford, United Kingdom) was used to create contours of the OARs. Both automated contouring methods were performed on a standard desktop computer, additionally the DLC used a graphics card to perform the calculations. The prototype uses a deep learning model based on convolutional neural networks [15], a sub-class of deep learning techniques tailored to process imaging data. Convolutional neural networks employ models with a large number of degrees of freedom (in the order of millions), and are therefore able to learn complex non-linear relationships within the imaging data. Training such models requires a large amount of imaging data and uses a backpropagation algorithm based on a stochastic gradient descent to optimize the free parameters [15]. Contours of 450 lung patients were collected from clinical practice and used to train the model.

Patient selection

Twenty consecutive stage I-III NSCLC patients treated in the period January to February 2016 were selected from routine clinical practice and the mid-ventilation phase of a 4D CT scan (Siemens Biograph PET/CT or Sensation Open CT scanner) was used to contour the OARs. The OARs were defined by institutional guidelines and comprised the left lung, right lung, heart, spinal cord, esophagus, and mediastinum.

Contour methods

For this study we created 5 contour sets: A manual contour (MC); an atlas-based contour (AC); a user adjustment of the atlas-based contour (UAC), meaning the atlas-based contour was used as a starting point and adjustments by a radiotherapy technician were allowed to meet institutional guidelines; a deep learning contour (DLC); and a user adjustment of the deep learning contour (UDLC). A single radiotherapy technician performed contouring tasks to prevent any inter-observer variability. Manual contouring tasks were performed using the software used in clinical practice (Eclipse, version 11.0, Varian, Palo Alto, United States of America). For each of the contouring tasks, the time required for completion was recorded per patient and OAR.

Subjective scoring of contours

The software generated contours were subjectively scored by the technician from one to four. (1) None of the results would form a useful basis for further editing, no time is expected to be saved contouring is expected compared to manual contouring; (2) Some of the results form a useful basis for further editing, little time would be saved contouring is expected compared to manual contouring; (3) Many of the results form a useful basis for further editing, a moderate time saving is expected compared to manual contouring; (4) Most of the results form a useful basis for further editing, a significant time saving is expected compared to manual contouring.

Contour consistency measurement

All contours were exported and analyzed in Matlab 8.6 (The MathWorks Inc., Natick, MA, USA). To assess similarity, the manual contour was compared to the software generated and user adjusted contours using several metrics. To quantify the similarity between the different contour sets, the contours were projected onto a three-dimensional grid matching the dimensions of the corresponding CT scan. The Dice index, which is the volume of the union normalized by the mean of the two volumes, of each OAR in comparison with the manual contour was calculated. The distance between the surfaces of each contour was measured using a nearest neighbor Euclidean distance calculation, the maximum nearest neighbor Euclidean distance, i.e. the Hausdorff distance, were compared. To evaluate differences in the results of the Dice index, Hausdorff distances and contouring time, a ranked Wilcoxon test was performed; p-values smaller than 0.05 were assumed to be statistically significant.

Results

Contouring time and consistency

The total median time saved was 7.8 min [range −2.2 to 13 min] and 10 min [range 5.2–15 min] for the UAC and UDLC respectively with respect to the MC. This is a large reduction compared to the median time required to contour all OARs for a lung case, which was 20 min. An overview of the recorded times per OAR is given in Fig. 1. Both lungs and the spinal cord show significant time reductions for the UAC and UDLC ($p < 0.05$), except for the left lung UAC. This was the result of one large outlier in the dataset, where the time needed to contour was 0.9 min and the time to adjust the AC was 3.6 min. The AC was incorrect due to the fact the lung had collapsed (Fig. 2A). Repeating the ranked Wilcoxon test without this sample resulted in a significant time gain for the left lung ($p = 0.014$). The DLC remained robust in the case where the lung had collapsed (Fig. 2A). The UDLC performs best with a median time to evaluate and correct under one minute for the lungs and spinal cord. Comparing the Dice and Hausdorff distance for the DLC and UDLC to MC results in no significant differences ($p > 0.05$), this supports the assumption that the time needed to adjust these OARs represents the time the technician needs to decide the contour meets clinical guidelines and can be used without major adjustments. An overview of the Dice scores and Hausdorff distances can be found in Figs. 3 and 4.

The UAC of the esophagus did not result in a significant time reduction with a median time saving of 0.3 min [range −1.9 to 1.9]. By contrast, the UDLC resulted in a significant time reduction of 1.5 min [range 0–3.2]. Fig. 2B shows an example slice of the esophagus contour where the DLC matches the MC but the AC is different. The heart and mediastinum showed a significant time reductions for both UAC and UDLC when compared to the MC,

Please cite this article in press as: Lustberg T et al. Clinical evaluation of atlas and deep learning based automatic contouring for lung cancer. Radiother Oncol (2017), https://doi.org/10.1016/j.radonc.2017.11.012

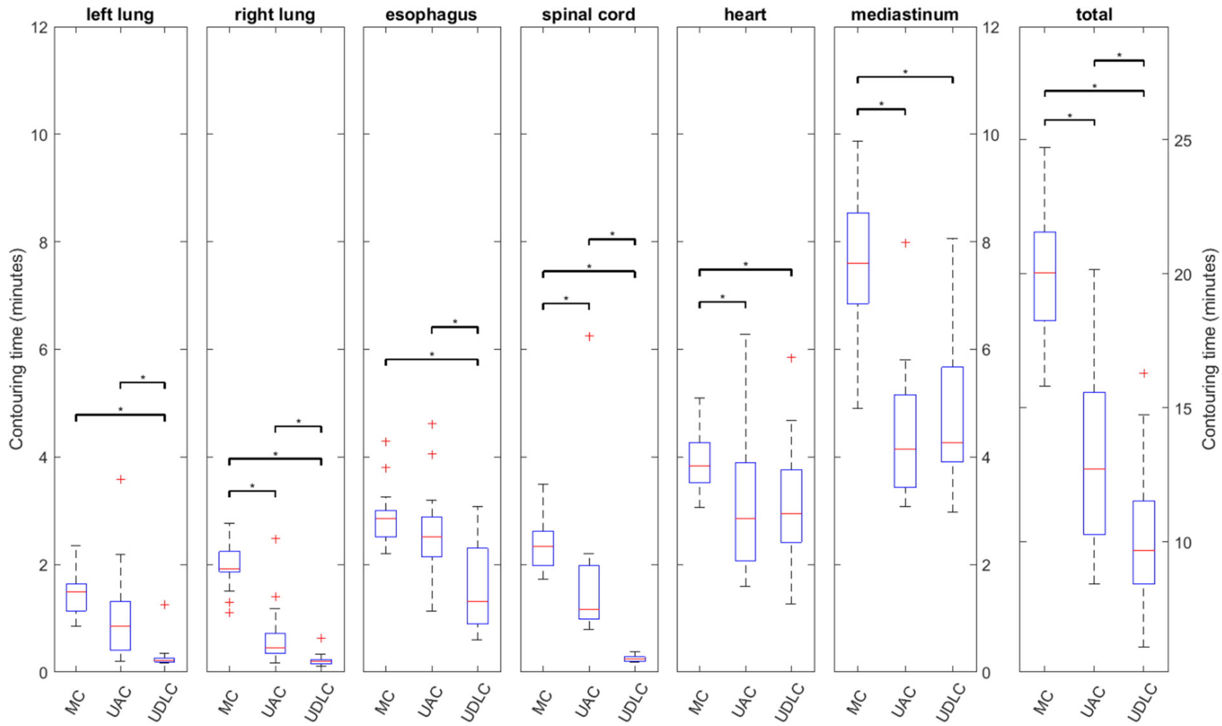

Fig. 1. Contouring time of the manual contour (MC) user adjusted atlas-based contour (UAC) and user adjusted deep learning contour (UDLC) displayed for each OAR and the total time of all OARs. * indicates significant difference between manual and adjustment timing ($p < 0.05$, ranked Wilcoxon test).

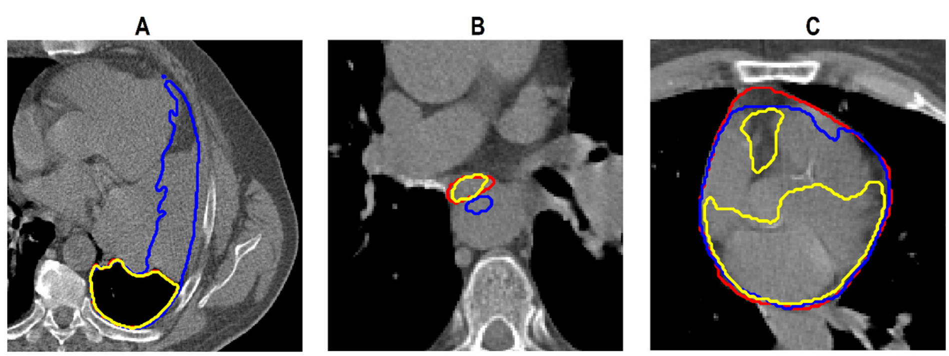

Fig. 2. Example cases showing manual contour (MC, red), atlas-based contour (AC, blue) and the deep learning contour (DLC, yellow) for left lung (A) the esophagus (B) and heart (C).

however, no significant difference was observed using UDLC over UAC ($p = 0.65$ and $p = 0.32$ for the OARs respectively). Fig. 2C shows an example of the heart, where the AC matches the MC but the DLC is different. The mediastinum had a median time reduction 3.4 min [range 0.6–5.7], which is the largest contribution to the overall time reduction. Comparing the AC and DLC to the UAC and UDLC for the esophagus, heart and mediastinum resulted in a significant different time required when compared to the MC ($p < 0.05$).

Subjective scoring of contours

For the lungs and spinal cord, AC performed well according to the subjective score of the technician, with a median score of 4 [range 2–4]. DLC performed even better for these OARs, where every contour was scored as 4. The esophagus AC performed poorly according to the technician with all contours having a score of 1. The DLC was perceived to be an improvement in most cases with a median score of 3 [range 1–4]. The AC of the heart and mediastinum were perceived to be slightly better than the DLC with respective median scores of 3 [range 2–3] and 3 [range 1 to 3]. The OARs with a median subjective score of 4 were consistent (median Dice score >90%, median Hausdorff distance <1.5 cm) when compared against the MC, however the spinal cord performed slightly worse (median Dice score 83%, median Hausdorff distance 1.6 cm).

Grouping the contours per subjective score instead of the corresponding OAR or contouring method resulted in the Dice scores, Hausdorff distances and time saved as shown in Fig. 5. Because the median time to contour an OAR varies greatly, e.g. 1.5 min for the left lung compared with 7.6 min for the mediastinum, the time saved was expressed in percentage of the time to manually contour the OAR. Comparing the manual contour to the software generated contours for OARs with a subjective score of 1 resulted in a median Dice score of 57% [range 16%–99%]; a median Hausdorff distance of 2.6 cm [range 0.4 cm–4.5 cm]; and a median time saved of 27% [range −70% to 92%]. Performing the same analysis for the OARs with a subjective score of 4 resulted in a median Dice score of 98% [range 70%–99%]; a median Hausdorff distance of 0.4

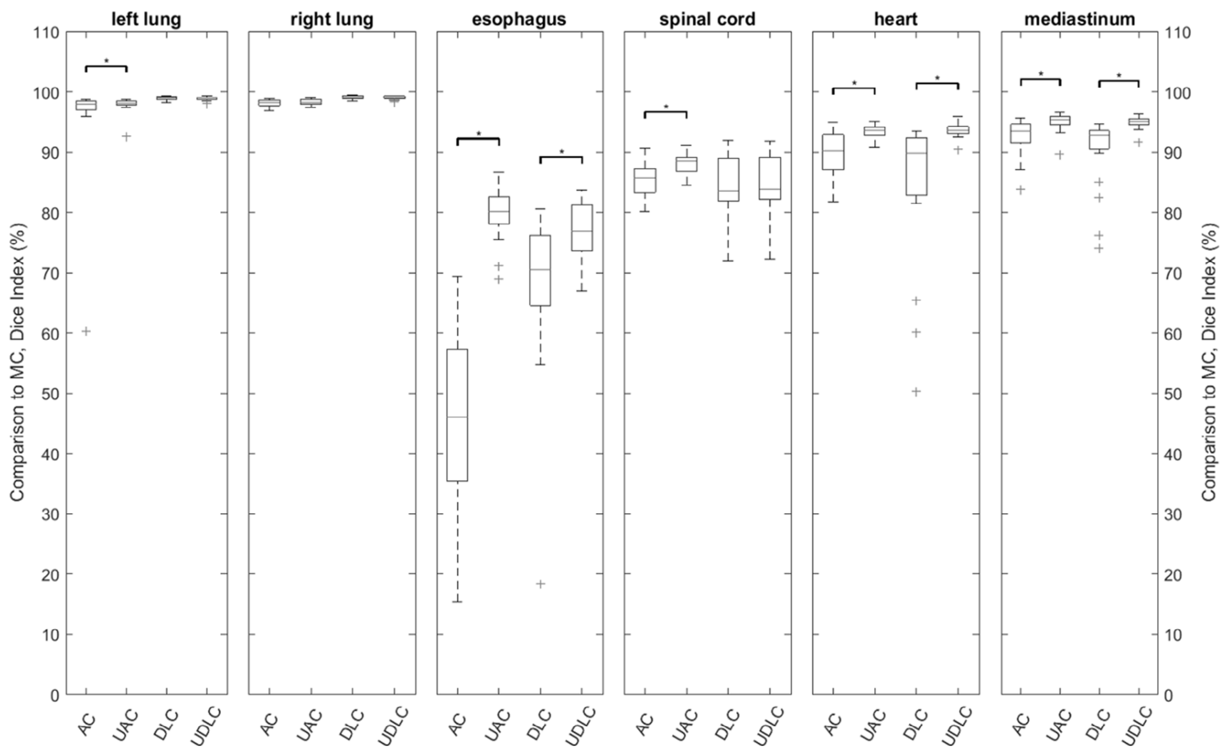

Fig. 3. Dice scores comparing the atlas contour (AC), user adjusted atlas contour (UAC), deep learning contour (DLC) and user adjusted deep learning contour (UDLC) to the manual delineation MC displayed for each OAR. * indicates significant difference between a software generated contour and user adjusted contour ($p < 0.05$, ranked Wilcoxon test).

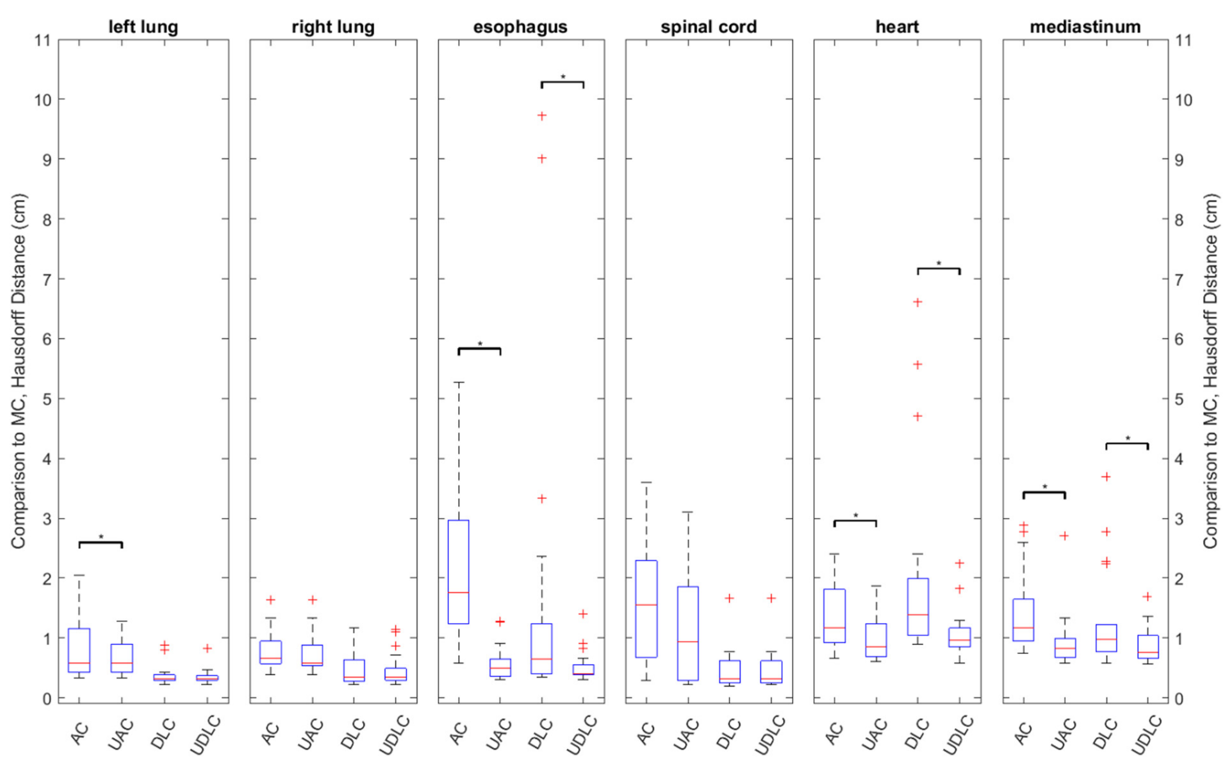

Fig. 4. Hausdorff distance comparing the atlas contour (AC), user adjusted atlas contour (UAC), deep learning contour (DLC) and user adjusted deep learning contour (UDLC) to the manual contour MC displayed for each OAR. * indicates significant difference between a software generated and user adjusted contour ($p < 0.05$, ranked Wilcoxon test).

Please cite this article in press as: Lustberg T et al. Clinical evaluation of atlas and deep learning based automatic contouring for lung cancer. Radiother Oncol (2017), https://doi.org/10.1016/j.radonc.2017.11.012

ARTICLE IN PRESS

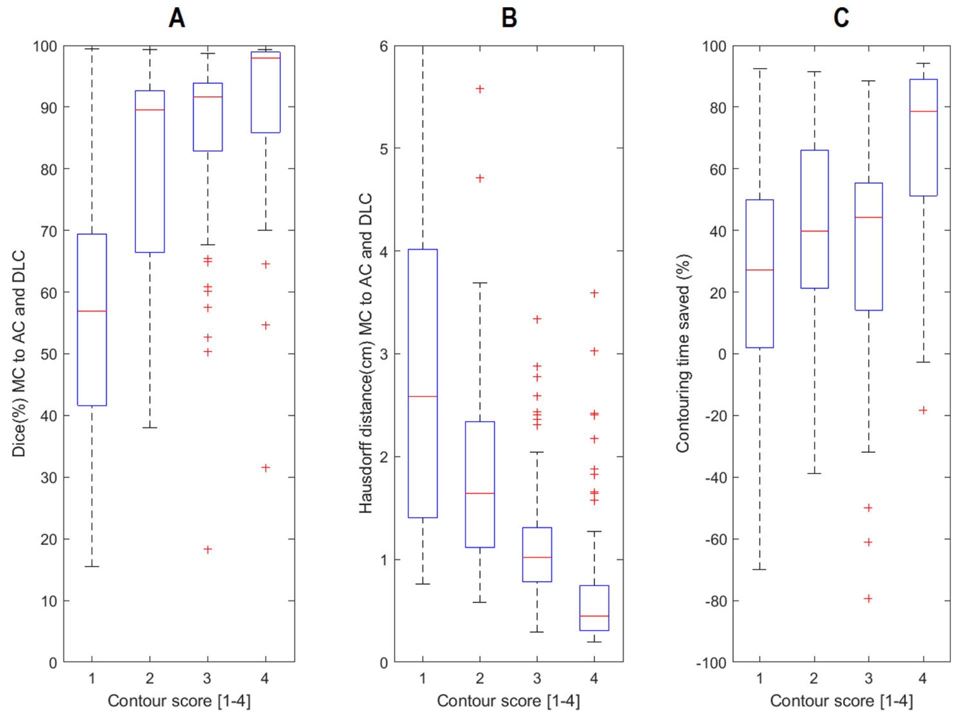

Fig. 5. The comparison results of the manual contour (MC) and the atlas based contour (AC) and deep learning contour (DLC) showing the Dice (A) and the Hausdorff distance (B) each subjective score group. Part C shows the time saved adjusting the software generated contours as a percentage of the manual contour time, displayed for each score group.

cm [range 0.2 cm–1.3 cm]; and a median time saved of 79% [range −3% to 94%].

Discussion

The methods presented in this study can be used to evaluate any commercially available auto-contouring software. Sharp et al. provide a comprehensive overview of the software that was available a few years ago (2). Evaluating contour consistency and recording time saved can be used to setup a commissioning protocol to use auto-contouring software in clinical practice and can be used to evaluate the cost-effectiveness of such products. Systematically creating a model for a treatment site, creating a manual, software generated and user adjusted contour and evaluating them will provide the necessary information to safely introduce contouring software into clinical practice. Having a low-tech alternative, i.e. user adjustment strategies as opposed to a software strategy [14], could support the clinical use of auto-contouring software in the near future.

In similar studies, it is shown that auto-contouring techniques can save time [16,17], while other studies show promising new techniques that could potentially save time if investigated [10,18]. Young et al. showed a potential time save for contouring lymph node for pelvic cancer patients [19]. Grouping the Dice scores, Hausdorff distances and time saved for each subjective score group comparing the manual to the user adjusted contours shows a relation between the consistency measurements and the subjective score. This demonstrates the ability of a radiation technician, while accounting for clinical guidelines, to judge whether contour quality is sufficient to lead to time savings compared with routine clinical practice. The technician can delete poor software generated contours and create a manual contour instead. This could from a basis of a viable and principled strategy for introducing imperfect automated contouring methods into clinical practice in order to save time while maintaining local clinical guidelines.

The results observed in the esophagus, heart and mediastinum highlight several issues with knowledge-based contouring software. First, the manual and user adjusted contours were significantly different while both were accepted as contours in compliance with local clinical guidelines by the same observer (intra-observer variability). For instance, the Dice score of the esophagus shows that even after adjusting the software generated contour to meet the clinical guidelines, the median Dice score is 78%, when comparing to the manual contour. Inter-observer variability in contours is reported to be an uncertainty that could greatly improve the quality of radiotherapy plans if reduced [1,3,20]. Van Baardwijk et al. investigated a model-based contouring method for the Gross Tumor Volume (GTV) in CT-PET scans and showed a decrease in Inter-observer variability [7]. Further investigation is needed to determine whether the automated contouring methods as described in this work could potentially reduce the intra- and inter-observer variability for OARs.

Using knowledge-based auto-contouring software, such as atlas-based and deep learning contouring, might improve the consistency of contours created by different observers because the software creates a contour which is the consensus of multiple observers when learning the model. In a future study, multiple observers should be included to determine the areas where they agree and where they do not. Comparing if the differences of the software generated contours happen in these specific parts of the contour will be a better quantification if the atlas is ready for clinical practice. This information, combined with the earlier suggested user adjustment method, could help with an important aspect of bringing automated contour techniques in the clinic, acceptance of these methods by the clinicians [2].

Deep learning is a state of the art machine learning technique that is utilized for many applications [15]. In health care specifically, there are several studies which investigate the utilization of deep learning for image contouring [21–23]. In our study, there are some promising results utilizing deep learning for the automatic contours of OARs for lung cancer patients. The deep learning contouring outperformed the atlas-based contouring for lungs and spinal cord. The deep learning performed better for the esophagus but further improvements remain necessary. In some cases, the user adjustment of the DLC did not save much time because slices near the stomach were missing and the technician judged these are clinically relevant to include. A similar case can be made for the heart example given in Fig. 2C. The clinical training data included substantial differences of opinion in the training data where some observers include or exclude the vessels at the top of the heart differently, therefore the deep learning contouring method is trying to combine these differences of opinion, leading to inaccuracy. For both cases, reaching consensus between observers prior to training DLC could provide improved performance.

As pointed out by Nelms et al. variation in contouring can have a huge impact on the dosimetric properties of a radiotherapy treatment plan for head and neck cancers [24]. In a contouring peer-review of lung cancer, the effect or contouring deviation on the resulting radiation treatment was shown [25]. The latter study discusses the dosimetric differences compared to the clinical guidelines and concludes that further investigation on the actual impact on tumor control and normal tissue toxicity is needed. In our study, we used the objective assessment of the technician to determine if contours meet clinical guidelines. Further investigation is needed to show the impact of these possible differences in contouring methods on dosimetry and finally treatment outcome.

Comparing the results of the automated contouring methods should be done with caution. The atlases are created from a highly curated set including an expert technician performing the contouring tasks that is validated by the radiation oncologist [26] while the DLC is learned on clinical data, which includes multiple observer preferences and possible imperfections. Comparing these two methods might be considered an unfair comparison, however, in this study we compared the current state of the art of both auto-contouring methods [22,26].

Time saving depends roughly on two main factors, the visualization of the boundary of the organ (e.g. lungs vs. esophagus) and the volume of the OAR to contour (e.g. mediastinal structures). High contrast edges, for instance the lungs, are easier to detect for both software and a human observer while low contrast edges, for instance the esophagus, are much harder. Automatic delineation methods typically are less accurate for small visible soft-tissue boundaries, which increases the time needed for adjustments, on top of this it is also more difficult for the human to distinguish where the contour should be, again increasing the time needed to adjust. Even if we assume that the auto-contouring techniques will reach human level contouring performance in the future, a human observer will probably still need time to evaluate difficult to contour OARs such as the esophagus.

Conclusion

Automatic contouring software as a starting point for clinical contours of OARs in lung radiation therapy allows for a significant time gain when contouring lungs, spinal cord, heart and mediastinum. DLC shows promising results with regard to the creation of institution-based models and to automatically generate high quality contours, providing a greater time saving compared to existing solutions. In addition, clinicians are able to assess if a software generated contour will potentially save time or not.

Acknowledgement

This work was funded by the Eurostars Cloud-Atlas project. (Project ID 9 297). MAASTRO clinic has research agreements with Mirada Medical Ltd.

References

[1] Vinod SK, Jameson MG, Min M, Holloway LC. Uncertainties in volume delineation in radiation oncology: A systematic review and recommendations for future studies. Radiother Oncol J Eur Soc Ther Radiol Oncol 2016 Nov;121(2):169–79.
[2] Sharp G, Fritscher KD, Pekar V, Peroni M, Shusharina N, Veeraraghavan H, et al. Vision 20/20: Perspectives on automated image segmentation for radiotherapy. Med Phys 2014;41 [Internet. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4000389/].
[3] Reed VK, Woodward WA, Zhang L, Strom EA, Perkins GH, Tereffe W, et al. Automatic segmentation of whole breast using an atlas approach and deformable image registration. Int J Radiat Oncol Biol Phys 2009 Apr 1;73(5):1493–500.
[4] Comparison and Evaluation of Methods for Liver Segmentation From CT Datasets - IEEE Xplore Document [Internet]. [cited 2017 Jan 23]. Available from: http://ieeexplore.ieee.org/document/4781564/
[5] Stapleford LJ, Lawson JD, Perkins C, Edelman S, Davis L, McDonald MW, et al. Evaluation of Automatic Atlas-Based Lymph Node Segmentation for Head-and-Neck Cancer. Int J Radiat Oncol 2010 Jul 1;77(3):959–66.
[6] Hoang Duc AK, Eminowicz G, Mendes R, Wong S-L, McClelland J, Modat M, et al. Validation of clinical acceptability of an atlas-based segmentation algorithm for the delineation of organs at risk in head and neck cancer. Med Phys 2015 Sep;42(9):5027–34.
[7] van Baardwijk A, Bosmans G, Boersma L, Buijsen J, Wanders S, Hochstenbag M, et al. PET-CT–based auto-contouring in non–small-cell lung cancer correlates with pathology and reduces interobserver variability in the delineation of the primary tumor and involved nodal volumes. Int J Radiat Oncol 2007 Jul 1;68(3):771–8.
[8] Luo Y, Liao Z, Jiang W, Gomez D, Williamson R, Court L, et al. TU-H-CAMPUS-JeP2-05: Can automatic delineation of cardiac substructures on noncontrast ct be used for cardiac toxicity analysis? Med Phys 2016 Jun;43(6):3783.
[9] Kim J, Han J, Ailawadi S, Baker J, Hsia A, Xu Z, et al. SU-F-J-113: multi-atlas based automatic organ segmentation for lung radiotherapy planning. Med Phys 2016 Jun;43(6):3433.
[10] Meng Q, Kitasaka T, Nimura Y, Oda M, Ueno J, Mori K. Automatic segmentation of airway tree based on local intensity filter and machine learning technique in 3D chest CT volume. Int J Comput Assist Radiol Surg 2017 Feb;12(2):245–61.
[11] Kopriva I, Ju W, Zhang B, Shi F, Xiang D, Yu K, et al. Single-channel sparse nonnegative blind source separation method for automatic 3D delineation of lung tumor in PET images. IEEE J Biomed Health Inform 2016.
[12] Rebouças Filho PP, Cortez PC, da Silva Barros AC, Albuquerque VHC, Tavares JMRS. Novel and powerful 3D adaptive crisp active contour method applied in the segmentation of CT lung images. Med Image Anal 2017 Jan;35:503–16.
[13] Dolz J, Kirişli HA, Fechter T, Karnitzki S, Oehlke O, Nestle U, et al. Interactive contour delineation of organs at risk in radiotherapy: clinical evaluation on NSCLC patients. Med Phys 2016 May;43(5):2569.
[14] Schreibmann E, Marcus DM, Fox T. Multiatlas segmentation of thoracic and abdominal anatomy with level set-based local search. J Appl Clin Med Phys 2014 Jul 1;15(4):22–38.
[15] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015 May 28;521 (7553):436–44.
[16] Haas B, Coradi T, Scholz M, Kunz P, Huber M, Oppitz U, et al. Automatic segmentation of thoracic and pelvic CT images for radiotherapy planning using implicit anatomic knowledge and organ-specific segmentation strategies. Phys Med Biol 2008;53(6):1751.
[17] Lim JY, Leech M. Use of auto-segmentation in the delineation of target volumes and organs at risk in head and neck. Acta Oncol 2016 Jul 2;55(7):799–806.
[18] Wang J, Guo H. Automatic approach for lung segmentation with juxta-pleural nodules from thoracic CT based on contour tracing and correction. Comput Math Methods Med 2016;2016:2962047.
[19] Young AV, Wortham A, Wernick I, Evans A, Ennis RD. Atlas-based segmentation improves consistency and decreases time required for contouring postoperative endometrial cancer nodal volumes. Int J Radiat Oncol 2011;79:943–7.
[20] Louie AV, Rodrigues G, Olsthoorn J, Palma D, Yu E, Yaremko B, et al. Inter-observer and intra-observer reliability for lung cancer target volume delineation in the 4D-CT era. Radiother Oncol J Eur Soc Ther Radiol Oncol 2010 May;95(2):166–71.
[21] Hu P, Wu F, Peng J, Liang P, Kong D. Automatic 3D liver segmentation based on deep learning and globally optimized surface evolution. Phys Med Biol 2016 Dec 21;61(24):8676–98.
[22] Mansoor A, Cerrolaza JJ, Perez G, Biggs E, Nino G, Linguraru MG. Marginal Shape Deep Learning: Applications to Pediatric Lung Field Segmentation. Proc SPIE– Int Soc. Opt Eng 2017 Feb;11:10133.
[23] Suzuki K. Overview of deep learning in medical imaging. Radiol Phys Technol 2017.
[24] Nelms BE, Tomé WA, Robinson G, Wheeler J. Variations in the contouring of organs at risk: test case from a patient with oropharyngeal cancer. Int J Radiat Oncol 2012;82:368–78.
[25] Lo AC, Liu M, Chan E, Lund C, Truong PT, Loewen S, et al. the impact of peer review of volume delineation in stereotactic body radiation therapy planning for primary lung cancer: a multicenter quality assurance study. J Thorac Oncol 2014;9:527–33.
[26] Peressutti D, Schipaanboord B, van Soest J, Lustberg T, van Elmpt W, Kadir T, et al. TU-AB-202-10: how effective are current atlas selection methods for atlas-based auto-contouring in radiotherapy planning? Med Phys 2016;43:3738–9.