6-8 (2023) 100018

Contents lists available at ScienceDirect

IPEM-Translation

journal homepage: www.elsevier.com/locate/ipemt

Artificial intelligence based auto-contouring solutions for use in radiotherapy treatment planning of head and neck cancer

Virginia Marin Anaya [1]

University College London Hospitals NHS Foundation Trust, Radiotherapy Physics, 250 Euston Road, London NW1 2PG, United Kingdom

A R T I C L E I N F O

Keywords:
Radiotherapy
Artificial intelligence based auto-contours
Geometric analysis
Dosimetric analysis
Treatment planning

A B S T R A C T

Background: Manual contouring is time-consuming and subjective. Thus, auto-segmentation methods, which can be deployed in the existing workflow, are needed. The objective of this study was to assess the feasibility of Limbus AI and AI Rad Companion auto-contours for head and neck treatment planning.

Methods: Head and neck patients treated with RapidArc were selected retrospectively. The manual contours on the planning CT were used as reference. Geometric analysis of the auto-contours was performed using several evaluation metrics such as the Dice Similarity Coefficient (DSC) and the Mean Distance to Conformity (MDC). Dosimetric analysis was performed by recalculating the original plan on the auto-contours and comparing Dose Volume Histogram (DVH) metrics to the original plan.

Results and discussion: Both AI tools tend to underestimate the volumes of brainstem and cord. For brainstem and parotids, median DSC values were ≥ 0.8. For all auto-contours, median MDC values were $\sim$ 3–6 mm. Median differences were found of up to ± 7 % in DVH points on the auto-contours relative to the planning CT contours, but these were not statistically-significant.

Conclusion: The auto-contours could be used as a starting point to assist the clinician with the manual contouring of structures on the planning and re-scanning planning CT.

Introduction

According to Cancer Research UK [1], there are around 12,400 new head and neck cancer cases in the UK every year. Radiotherapy alone, or in combination with other treatments such as surgery or chemotherapy, is used as part of the primary cancer treatment. Radiotherapy treatment planning requires delineation of target volumes and organs at risk (OARs). In current practice, contouring is performed manually by the clinician, which is tedious, time-consuming, and prone to inter- and intra-observer variability [2]. This reduces clinicians' availability for other tasks and can delay the start of patients' treatments leading to poorer tumour control probability, and for some patients, reduced probability of survival. Thus, auto-segmentation methods, which can be deployed in the existing workflow, are needed in order to improve contouring consistency, optimise patient treatment pathways and improve patient outcomes, whilst enabling effective use of staff resources.

Furthermore, radiotherapy treatment planning of head and neck can be challenging. This is because there can be significant changes in patient anatomy during the course of radiotherapy treatment [3]. These changes are mainly due to patient weight loss and tumour volume shrinkage, which can result in loose immobilization devices, increasing the patient set-up uncertainty [4]. Several studies have shown that the GTV (Gross Tumour Volume) can shrink by 1.8–3.9 % per day [5]. Similarly, the parotid glands can experience significant volume reduction; various studies have shown that they can shrink by 0.6–1.1 % per day [6,7]. These changes, if not taken into account properly, can lead to insufficient dose coverage of the target volumes and potential overdose to healthy tissue and organs at risk. Re-planning, including re-contouring of target volumes and organs at risk, may be required. Auto-contouring could potentially speed up plan adaptation which would be a great asset in radiotherapy.

Atlas-based auto-segmentation has dominated commercial packages, but is now being superseded by deep-learning methods [2]. Limbus AI and AI Rad Companion Organs RT are commercial solutions for auto-contouring based on deep learning. The objective of this study was to assess the feasibility of AI auto-contours generated by Limbus Contour version 1.5.0 (Limbus AI Inc., Canada) and AI Rad Companion Organs

E-mail address: virginia.marinanaya1@nhs.net.

[1] Proton Therapy Physicist & NIHR Innovation Fellow & Topol Digital Fellow.

https://doi.org/10.1016/j.ipemt.2023.100018
Received 28 April 2023; Received in revised form 13 October 2023; Accepted 29 October 2023
Available online 24 November 2023

RT version VA31 (Siemens Healthineers, Siemens Healthcare GmbH., Germany) for head and neck radiotherapy treatment planning. As part of the study, a geometric and dosimetric analysis of the auto-contours generated by both AI tools were performed. It is worth noting that a comparison between the two AI tools was beyond the scope of this study.

Methods

Ten head and neck patients, with no significant imaging artifacts, treated with RapidArc (65 Gy and 54 Gy in 30 fractions) were selected retrospectively. All the plans had bilateral nodal irradiation. Each patient had a planning CT with manual contours of target volumes and organs at risk. CT images were acquired with a Siemens SOMATOM Confidence scanner (Siemens Healthineers, Siemens Healthcare GmbH., Germany) and a slice thickness of 0.1 cm. The manual contours on the planning CT were used as reference. Inter-observer variability was eliminated since all the delineations were performed by the same clinical oncologist. It was not possible to eliminate intra-observer variability and the potential bias in the reference contours being from the single clinician.

Geometric analysis

The aim of the geometric analysis was to evaluate the accuracy of AI auto-contours for head and neck treatment planning. The geometric analysis (Fig. 1) was carried out using the contour comparison module of ImSimQA version 4.2 (Oncology Systems Limited, UK)

Metrics

Geometric analysis of the auto-contours was performed using several evaluation metrics such as the Dice Similarity Coefficient (DSC), the Mean Distance to Conformity (MDC) in mm, the Target Registration Error (TRE) in mm, the Inclusiveness Index and the Sensitivity Index.

The DSC is a measure of spatial overlap.

$$DSC = \frac{2(A \cap B)}{A + B} \tag{1}$$

Where A and B are the evaluation and the reference contours respectively.

The MDC is defined as the mean distance of each outlying voxel from the reference contour to the evaluation contour.

$$MDC = \frac{1}{N(A)} \sum_{i=0}^{N(A)} d(A_i, B_i) \tag{2}$$

Where *d(Ai,Bi)* is the distance from point *i* on surface *A* to the closest point on surface *B*, and *N(A)* is the total number of surface points on contour *A* [8,9].

The Target Registration Error (TRE) is defined as the average residual error between the identified points on the evaluation volume of interest and the points identified of the reference volume of interest.

The Sensitivity Index (overlapping index) measures the probability that the evaluation contours match their corresponding reference.

$$Sens.\ Index = \frac{A \cap B}{B} \tag{3}$$

Where A and B are the evaluation and the reference contours respectively [10].

The Inclusiveness Index (specificity) measures the probability that a voxel of the evaluation contour is really a voxel of the reference contour.

$$Incl.\ Index = \frac{A \cap B}{A} \tag{4}$$

Where A and B are the evaluation and the reference contours respectively [10].

Dosimetric analysis

The aim of the dosimetric analysis was to determine the dosimetric impact of using auto-contours instead of the manual contours. Dosimetric analysis (Fig. 2) was performed by re-calculating the original plan on the auto-contours and comparing Dose Volume Histogram (DVH) metrics to the original plan.

The original plans were generated by experienced dosimetrists using standard templates in Eclipse treatment planning system. All plans were checked by a senior radiotherapy physicist and in accordance with departmental protocols.

The Analytical Anisotropic Algorithm (AAA) version 16.1.0 was used for the dose calculation.

In Eclipse treatment planning system version 16 (Varian A Siemens Healthineers Company, Siemens Healthcare GmbH., Germany), a new course was created. The original treatment plan was copied and pasted, assigning the new structure set that contains the auto contours. The new treatment plan was re-calculated with the same settings, including the same number of monitor units (MUs) as in the original plan.

For the dosimetric analysis, the change in plan metrics on the manually-delineated contours by the clinician (reference) were compared against the change in plan metrics observed when using auto-contours. To assess the dose to the left and right parotid, their

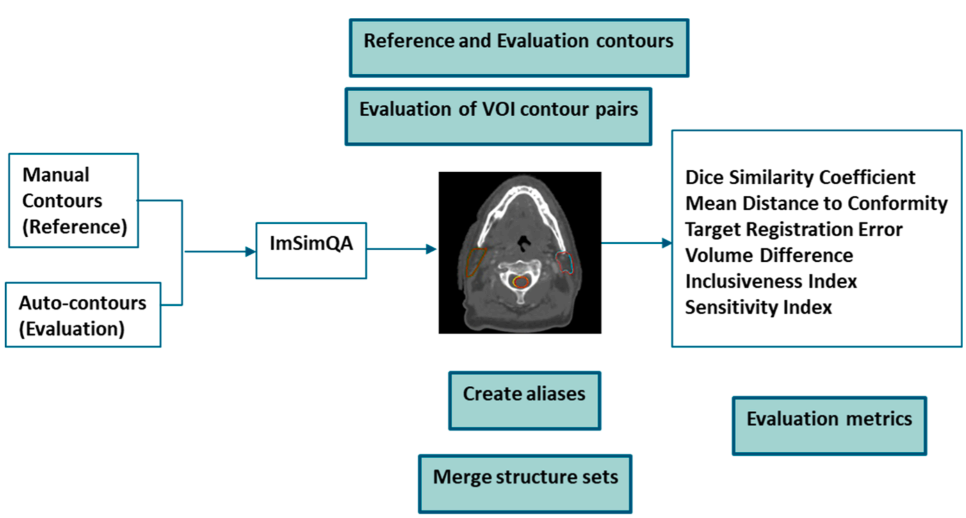

Fig. 1. Overview of the study design for the geometric analysis.

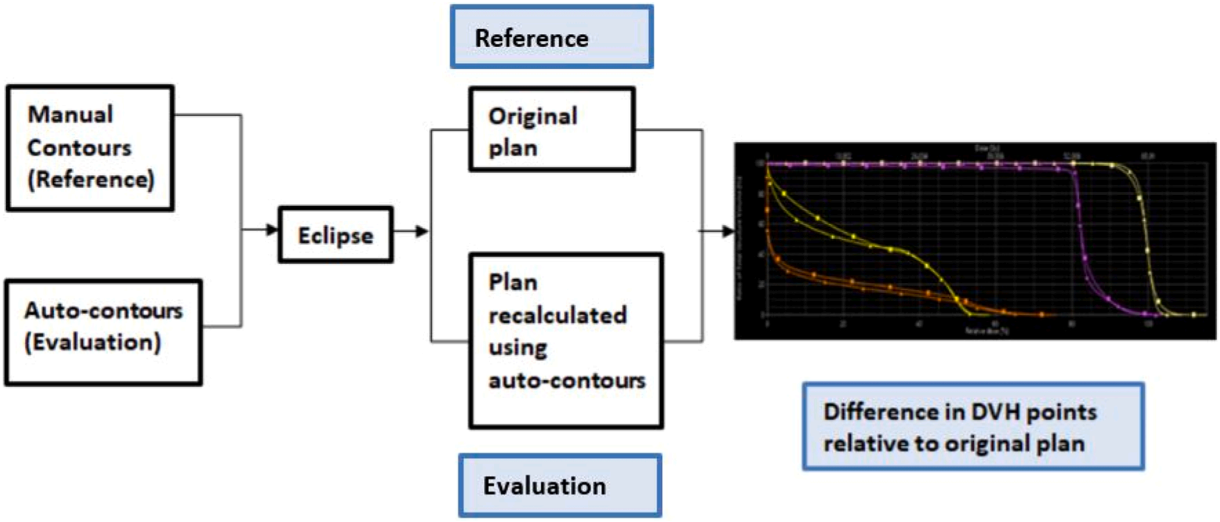

Fig. 2. Overview of the study design for the dosimetric analysis.

corresponding mean doses were recorded. For the brainstem and the spinal cord, which are serial organs at risk, the dose to 0.1cc (D0.1cc) and the maximum dose (Dmax) were determined. This is in accordance with departmental clinical guidelines.

It is worth acknowledging that there is a limitation in our method. The dosimetric difference has been quantified by recalculating the original plan, with fixed MU, on the auto-contoured set. A more meaningful comparison would have been to optimise plans on the auto-contour structure set and then recalculate these plans, with fixed MU, on the original structure set for comparison. This would have captured any optimisation compromises as a result of the auto-contours by comparing dosimetrically to the original structure set, which is considered the reference. The auto-contours would either result in differently optimised plans or be edited by the clinician to more closely match the original structure set. This might be investigated as part of future work.

Results

Geometric analysis

Both AI tools tend to underestimate the volumes of brainstem and cord when comparing the auto-contours to their corresponding manual delineations on the original planning CT (Table 1, Table 2). For brainstem and parotids, median DSC values were $\geq$ 0.8. For all auto-contours, median MDC values and median TRE values were ~ 3–6 mm and ~1–4 mm respectively. Except for the cord, the median Sensitivity and Inclusiveness Indexes values were $\geq$ 0.7.

Table 1
Geometric analysis of Limbus AI. Results are expressed as median and range between brackets.

Structure	DSC	MDC (mm)	TRE (mm)	Vol. Diff. (%)	Sens. Index	Incl. Index
Brainstem	0.87 (0.82, 0.92)	3.34 (2.52, 4.38)	1.00 (0.35, 2.06)	−8.06 (−20.73, 0.84)	0.84 (0.73, 0.92)	0.92 (0.89, 0.92)
Spinal Cord	0.71 (0.57, 0.88)	2.71 (2.14, 63.67)	1.06 (0.57, 63.86)	−18.31 (−43.99, 5.22)	0.57 (0.53, 0.91)	0.86 (0.63, 0.99)
Left Parotid	0.87 (0.79, 0.88)	4.19 (3.66, 5.30)	1.64 (1.21, 2.38)	3.94 (1.27, 10.25)	0.88 (0.79, 0.92)	0.83 (0.78, 0.86)
Right Parotid	0.85 (0.62, 0.86)	4.36 (3.33, 9.14)	1.62 (0.86, 7.19)	−3.93 (−16.31, 10.11)	0.83 (0.57, 0.91)	0.85 (0.67, 0.87)

Table 2
Geometric analysis of AI Rad Companion Organs RT. Results are expressed as median and range between brackets.

Structure	DSC	MDC (mm)	TRE (mm)	Vol. Diff. (%)	Sens. Index	Incl. Index
Brainstem	0.83 (0.78, 0.91)	3.96 (3.18, 6.39)	1.67 (1.02, 7.53)	−20.42 (−28, −8.36)	0.71 (0.70, 0.87)	0.95 (0.88, 0.99)
Spinal Cord	0.54 (0.34, 0.79)	3.37 (2.98, 21.19)	3.62 (0.67, 19.61)	−21.02 (−63.2, 53.43)	0.46 (0.36, 0.71)	0.89 (0.30, 1.00)
Left Parotid	0.80 0.69, 0.84)	5.37 (4.65, 8.95)	3.04 (1.42, 8.31)	31.68 (18.92, 45.4)	0.92 (0.83, 0.96)	0.71 (0.58, 0.77)
Right Parotid	0.79 (0.59, 0.83)	5.59 (3.77, 12.56)	3.31 (2.11, 7.92)	15.74 (−6.51, 28.71)	0.85 (0.64, 0.93)	0.74 (0.55, 0.80)

The best geometric results were attained by the brainstem, whereas the poorest geometric results were found for the cord. For example, in one patient, MDC values for spinal cord of ~64 mm (Table 1) between the auto-contour and the manual contour were recorded. It was noted that the AI tool outlines the entire spinal cord whilst the manual contour only outlines the structure within the field size.

Qualitative analysis of AI auto-contours by oncologists was beyond the scope of this study. However, clinicians showed a preference towards editing auto-contours, if necessary, rather than outlining from scratch, saving overall contouring time. Assessing the effect on contouring time is not stated in the aim of the study. This might be investigated as part of future work.

Dosimetric analysis

Median differences were found of up to ±7 % in DVH points on the auto-contours relative to the planning CT contours (Table 3).

Statistical analysis of the plan evaluation metrics was performed using IBM SPSS Statistics Version 27. A Wilcoxon Signed Rank test between the plans was carried out to determine whether the differences in DVH points for the different structures relative to the original plan were statistically-significant. The statistical significance was set at 0.05.

There was no statistically-significant difference in the DVH points of OARs between the plans.

The calculated doses, derived from re-calculation on the AI auto-contours, tend to be higher for the parotids and lower for the

Table 3
Dosimetric analysis. Results were expressed as median and range between brackets.

Structure	DVH Point	Limbus AI (%)	AI Rad Comp (%)
Brainstem	**Dmax**	−4.6 (−9.1, −0.1)	−6.9 (−20.1, 0.2)
	D0.1cc	−4.4 (−9.6, −0.1)	−6.1 (−0.3 −6.1)
Cord	**Dmax**	−1.6 (−5.1, 0.0)	−3.1 (−5.9, 0.0)
	D0.1cc	−1.9 (−4.6, −0.3)	−2.1 (−5.7, −1.3)
Left Parotid	**Dmean**	0.4 (−0.7, 6.1)	3.2 (0.8, 18.6)
Right Parotid	**Dmean**	1.6 (−10.3, 4.3)	6.9 (−2.0, 10.4)

brainstem and cord with reference to the original plan. However, it was found that these changes were not statistically significant (p-Value>0.05).

Discussion

In our study, we performed a geometric and dosimetric analysis of auto-contours generated by Limbus AI and AI-Rad Companion Organs RT for head and neck.

The evaluation of auto-contouring algorithms is commonly performed using geometric metrics [11]. Although various evaluation metrics have been proposed in the literature, there is no widely accepted method for geometrical comparison. Hanna et al. [12] carried out a review of comparison methods for geometric analysis of radiotherapy volumes and recommends using at least a volume assessment in addition to a metric to assess positional displacement such as the centre of mass shift. There is no single metric that provides a full description of the change [13].

Available algorithms in commercial software packages are mainly atlas-based methods. However, these auto-contours require significant editing. Recently, artificial intelligence approaches based on deep learning convolutional neural networks are gaining popularity. There is evidence in the literature [14–18] that suggests that deep-learning auto-contours have better accuracy when compared to manual contours than those produced by model-based or atlas-based approaches. Deep learning auto-contouring has the potential to reduce the clinical burden by reducing the time spent on producing acceptable contours.

Ibragomov et al. [19] proposed the first convolutional neural network method for delineation of OARs for head and neck cancer. The method showed more accurate results than atlas-based models. Moreover, Nikolov et al. [20] and Zhu et al. [21] investigated the use and performance of a 3D U-NET convolutional neural network (DeepMind and AnatomyNet respectively) for whole volume delineation of head and neck cancer with promising results. However, these studies did not evaluate the benefits of adopting these deep learning methods into the clinical workflow.

Van der Veen et al. [18] studied the benefits of a deep learning method (DeepVoxNet) for delineation of organs at risk in head and neck cancer in terms of geometric accuracy, efficiency and consistency compared to manual delineation. The accuracy of the automated delineation method was assessed for each OAR using the Dice Similarity Coefficient (DSC) and the Average Symmetric Surface Distance (ASSD). The efficiency of the automated delineation method was quantified by comparing the time needed for manual delineation to the time needed for correction of the auto-contours. The method showed average delineation time savings of ~33 % when auto-contours were edited instead of outlining from scratch. In addition, the method showed an increase in consistency compared to manual delineation. The study concluded the benefits could justify its implementation in clinical practice.

In a recent study, D'Aviero et al. [22] performed a geometric analysis of Limbus AI for head and neck; the auto-contours were compared with manual contours using DSC and Hausdorff Distance. The study found that Limbus AI provided acceptable head and neck OARs delineations. Their results suggested that Limbus AI could be considered as a starting point for review and edited if necessary, optimising workload and resources in radiotherapy departments. This is in line with our geometric analysis.

In another study by Wong et al. [23], time savings were estimated as at least 26 min for head and neck. As suggested by D'Aviero et al. [21] and Wong et al. [22] auto-contours could be used as a helpful tool to assist the clinician with the manual contouring of structures on the planning and re-scanning planning CT.

For the purpose of research, Microsoft InnerEye developed by Microsoft Research Lab in Cambridge is an open-source AI toolkit to train models on medical images. InnerEye is based on a machine learning model that uses a 3D convolutional neural network for automated delineation which has already been applied to train head and neck and prostate datasets to generate contours of OARs from CT images. Oktay et al. [24] performed a study on 242 head and neck and 519 male pelvic CT image datasets acquired for radiotherapy treatment at 8 different cancer centres. The head and neck model and the prostate model were trained on a subset of dataset to automatically outline OARs. The study found that Microsoft InnerEye contours achieved accuracy within inter-observer expert variability. Statistical agreement was obtained for 13 out of 15 OARs. Moreover, Oktay et al. used ten head and neck patients and found the mean clinician time to be ~87 min for manual contouring compared with ~5 min for auto-contours, representing a reduction in mean clinician time of ~93 %.

Although in the literature, much attention has been justifiably focused on the geometrical accuracy of the contours, the dosimetric implications have not been frequently studied. An analysis based on dosimetric parameters can be more relevant in clinical practice [11]. As highlighted by Pukala et al. [25], one limitation shared by all evaluation metrics is the disconnect between the quantification of the geometrical accuracy and the effect that this has on the dose distribution. Voet et al. [26] and Tsuji et al. [27] found that a high value of DSC for target volumes cannot be used as a predictor for dose coverage. Geometric and dosimetric accuracy paradigms, although related, are not equivalent [28].

Kawula et al. [11] found no statistically significant correlation between geometric and dosimetric metrics showing that both types of analysis should be included in the evaluation of auto-contours of OARs in radiotherapy. Kawula et al. performed a dosimetric analysis of deep learning-based CT auto-segmentation (3D U-NET based on the V-Net architecture) for prostate cancer. Dosimetric analysis based on clinically relevant DVH parameters of VMAT (volumetric modulated arc therapy) plans did not show statistically significant differences for rectum and bladder. This is in line with our dosimetric analysis that shows no statistically significant differences in the DVH points of plans for OARs in head and neck. Guo et al. [29] studied the dosimetric impact of deep learning-based auto-segmentation of OARs in nasopharyngeal and rectal cancer. The treatment plan was re-optimised based on the auto-contours and then used the manual contours to assess the dosimetric differences between the re-optimised and the original plans. The study found no strong correlations between the geometric metrics and dosimetric differences for OARs.

Conclusion

To the best of our knowledge this is the first study that includes a geometric and a dosimetric analysis of Limbus AI and AI Rad Companion autocontours for head and neck. The dosimetric analysis showed that there was no statistically-significant difference in the DVH points of OARs between the plan re-calculated using the auto-contours and the original plan. The results of our geometric analysis are in line with those previously published. The auto-contours produced by AI tools are not able to completely replace manual contouring by the clinician. There is evidence in the literature that suggests that reviewing and editing the auto-contours if necessary can save time and resources.

Declaration of Competing Interest

The author has been awarded a Healthcare Science Innovation Fellowship hosted by the National Institute of Health and Care Research (NIHR) Devices for Dignity MedTech and In Vitro Diagnostics Cooperative. The Fellowship is also supported by the Chief Scientific Officer team and Health Education England through the National School of Healthcare Science.

Acknowledgements

The author would like to thank her mentor Professor Wendy Tindale OBE (Scientific Director at Sheffield Teaching Hospitals NHS Foundation Trust) for her advice. She would also like to thank Radiotherapy Physics staff at University College London Hospitals NHS Foundation Trust and the organisers of IPEM Science Technology and Engineering Forum (STEF) 2023. This project has been supported by NIHR Devices for Dignity MedTech and In Vitro Diagnostics Cooperative. The views expressed are those of the author and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Ethical approval

Patients gave their informed consent to the collection and use of data for continual improvement of the service. Data was used retrospectively and the findings of the study will not affect patients' treatments.

In addition, HRA Approval was given (IRAS Project ID: 306032) for UCLH evaluation of InnerEye head and neck cancer model for radiotherapy planning using UCLH dataset.

References

[1] Cancer Research UK. Head and neck cancers incidence. Available at https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/head-and-neck-cancers#heading-Zero (accessed April 2023).
[2] K. Harrison, H. Pullen, C. Welsh, et al., Machine learning for auto-segmentation in radiotherapy planning, Clin. Oncol. 34 (2) (2022) 74–88.
[3] S. Marzi, P. Pinnarò, D. D'Alessio, L. Strigari, V. Bruzzaniti, C. Giordano, et al., Anatomical and dose changes of gross tumour volume and parotid glands for head and neck cancer patients during intensity-modulated radiotherapy: effect on the probability of xerostomia incidence, Clin. Oncol. (R. Coll. Radiol.) 24 (3) (2012) 54–62.
[4] C. Yip, C. Thomas, A. Michaelidou, D. James, R. Lynn, M. Lei, et al., Co-registration of cone beam CT and planning CT in head and neck IMRT dose estimation: a feasible adaptive radiotherapy strategy, Br. J. Radiol. 87 (1034) (2014) 20130532.
[5] J.L. Barker Jr, A.S. Garden, K.K. Ang, J.C. O'Daniel, H. Wang, L.E. Court, et al., Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system, Int. J. Radiat. Oncol. Biol. Phys. 59 (2004) 960–970.
[6] P. Castadot, X. Geets, J.A. Lee, N. Christian, V. Gregoire, Assessment by a deformable registration method of the volumetric and positional changes of target volumes and organs at risk in pharyngo-laryngeal tumours treated with concomitant chemo-radiation, Radiother. Oncol. 95 (2010) 209–217.
[7] C. Han, Y.J. Chen, A. Liu, T.E. Schultheiss, J.Y. Wong, Actual dose variation of parotid glands and spinal cord for nasopharyngeal cancer patients during radiotherapy, Int. J. Radiat. Oncol. Biol. Phys. 70 (2008) 1256–1262.
[8] R. Jena, N.F. Kirkby, K.E. Burton, A.C.F. Hoole, L.T. Tan, N.G. Burnet, A novel algorithm for the morphometric assessment of radiotherapy treatment planning volumes, Br. J. Radiol. 83 (985) (2010) 44–51.
[9] V. Marin Anaya, J. Fairfoul, Assessing the feasibility of adaptive planning for prostate radiotherapy using SmartAdapt deformable image registration, Med. Eng. Phys. 64 (2019) 65–73.
[10] M. La Macchia, F. Fellin, M. Amichetti, M. Cianchetti, S. Gianolinis, V. Paola, et al., Systematic evaluation of three different commercial software solutions for automatic segmentation for adaptive therapy in head and neck, prostate and pleural cancer, Radiat. Oncol. 7 (2012) 160.
[11] M. Kawula, D. Purice, L. Minglun, G. Vivar, S.A. Ahmadi, K. Parodi, et al., Dosimetric impact of deep learning based CT auto-segmentation on radiation therapy treatment planning for prostate cancer, Radiat. Oncol. 17 (2022) 21.
[12] G.G. Hanna, A.R. Hounsell, J.M O'Sullivan, Geometrical analysis of radiotherapy target volume delineation: a systematic review of reported comparison methods, Clin. Oncol. 22 (7) (2010) 515–525.
[13] K.K. Brock, S. Mutic, T.R. McNutt, H. Li, M.L. Kessler, Use of image registration and fusion algorithms and techniques in radiotherapy: report of the AAPM Radiation Therapy Committee Task Group No. 132, Med. Phys. 44 (7) (2017) 43–76.
[14] J. Zhu, et al., Comparison of the automatic segmentation of multiple organs at risks in CT images of lung cancer between deep convolutional neural networks –based and atlas-based techniques, Acta Oncol. 58 (2019) 257–264.
[15] T. Lustberg, et al., Clinical evaluation of atlas and deep learning based automatic contouring for lung cancer, Radiother. Oncol. 126 (2) (2018) 312–317.
[16] W.J. Zabel, et al., Clinical evaluation of deep learning and atlas-based auto-contouring of bladder and rectum for prostate radiation therapy, Pract. Radiat. Oncol. 11 (1) (2021) e80–e89.
[17] L.V. van Dijk, L. Van den Bosch, P. Aljabar, D. Peressutti, S. Both, Steenbakkers R J H M, J.A. Langendijk, M.J. Gooding, C.L Brouwer, Improving automatic delineation for head and neck organs at risk by Deep Learning Contouring, Radiother. Oncol. 142 (2020) 115–123.
[18] J. Van der Veen, et al., Benefits of deep learning for delineation of organs at risk in head and neck cancer, Radiother. Oncol. 138 (2019) 68–74.
[19] B. Ibragimov, L. Xing, Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks, Med. Phys. 44 (2017) 547–557.
[20] S. Nikolov, S. Blackwell, A. Zverovitch, R. Mendes, M. Livne, J. De Fauw, et al., Clinically applicable segmentation of head and neck anatomy for radiotherapy: deep learning algorithm development and validation study, J. Med. Internet Res. 23 (7) (2021) e26151.
[21] W. Zhu, Y. Huang, L. Zeng, X. Chen, Y. Liu, Z. Qian, et al., AnatomyNet: deep learning for fast and fully automated whole-volume segmentation of head and neck anatomy, Med. Phys. 46 (2019) 576–589.
[22] A. D'Aviero, A. Re, F. Catucci, D. Piccari, C. Votta, D. Piro, et al., Clinical validation of a deep-learning segmentation software in head and neck: an early analysis in a developing radiation oncology center, Int. J. Environ. Res. Public Health 19 (15) (2022) 9057.
[23] J. Wong, A. Fong, N. McVicar, S. Smith, J. Giambattista, D. Wells, et al., Comparing deep learning-based auto-segmentation of organs at risk and clinical target volumes to expert inter-observer variability in radiotherapy planning, Radiother. Oncol. 144 (2020) 152–158.
[24] O. Oktay, et al., Evaluation of deep learning to augment image-guided radiotherapy for head and neck and prostate cancers, JAMA Netw. Open 3 (11) (2020), e2027426.
[25] J. Pukala, P.B. Johnson, A.P. Shah, K.M. Langen, F.J. Bova, R.J. Staton, et al., Benchmarking of five commercial deformable image registration algorithms for head and neck patients, J. Appl. Clin. Med. Phys. 17 (3) (2016) 25–40.
[26] P.W.J. Voet, M.L.P. Dirkx, D.N. Teguh, M.S. Hoogeman, P.C. Levendag, B.J. M. Heijmen, Does atlas-based autosegmentation of neck levels require subsequent manual contour editing to avoid risk of severe target underdosage? A dosimetric analysis, Radiother. Oncol. 98 (3) (2011) 373–377.
[27] S.Y. Tsuji, A. Hwang, V. Weinberg, S.S. Yom, J.M. Quivey, P. Xia, Dosimetric evaluation of automatic segmentation for adaptive IMRT for head-and-neck cancer, Int. J. Radiat. Oncol. Biol. Phys. 77 (3) (2010) 707–714.
[28] C. Paganelli, G. Meschini, S. Molinelli, M. Riboldi, G. Baroni, Patient-specific validation of deformable image registration in radiation therapy: overview and caveats, Med. Phys. 45 (10) (2018) e908–e922.
[29] H. Guo, J. Wang, X. Xia, et al., The dosimetric impact of deep learning-based auto-segmentation of organs at risk on nasopharyngeal and rectal cancer, Radiat. Oncol. 16 (2021) 113.