Abstract Overview

Multicentric Comparative Analysis of 8 Commercial Deep Learning-Based Autocontouring Solutions for the Delineation of Head and Neck Organs-at-Risk

Authors: M. Westermayer, N. Delaby, A. Sottiaux, S. Kara Mohammed, MC. Biston, C. Khamphan, S. Huger, S. Chiavassa, I. Bessieres, I. Achag , M. Stadler, S. Farzam, C. Millardet-Martin , T. Lacornerie, C. Fontbonne, JM Fontbonne, L. Madec, R. Castanet, A. Coutte, J. Biau, Y. Pointreau, A.A. Serre, J. Castelli, A. Barateau, A. Badey

Purpose

- The study aims to compare the performance of 8 different commercial DL-based AS solutions for H&N on CT images.

Mat&Methods

- Data: 20 HN patients.
- AI solutions: MiradaDLCExpert, MVision , RayStation, ART-Plan, Syngo.Via, Limbus, MIM, AutoContour.
- OARs: brainstem, oral cavity, spinal canal, parotid glands, mandible, optic nerves, eyes, and submandibular glands.
- Manual contours: done by 5 rad Oncs .
- Geometric evaluation: DICE and HD95%.
- Dosimetric impact: by projecting the reference treatment plan (optimized using manual reference contours) onto the autocontours and comparing DVH points.
- Statistical investigation of relation between geometric results and dosimetric differences.

Results

- The mean vDSC values exceeded 0.8 across almost all structures and solutions.
- TheraPanacea had the highest vDICE results for all the OARs except Oral Cavity (with highest DICE was for syngo.via and MVision).
- Dosimetric differences had very few strong corellations with geometric indices, and in particular for organs located closer to the target volumes.

Conclusion

- All commercial solution provided satisfying geometric results, however since a weak correlation with dosimetric impact, case specific manual correction and approval by a rad onc are still required.

- Leading performance vs. other commercial solutions as a result of High geometric accuracy across OARs.
  Achieving the highest DICE for most OARs, demonstrating competitive performance among 8 evaluated tools.
  ART-Plan is among the solutions achieving mean DICE scores > 0.8 for most structures in Head and neck cancer.

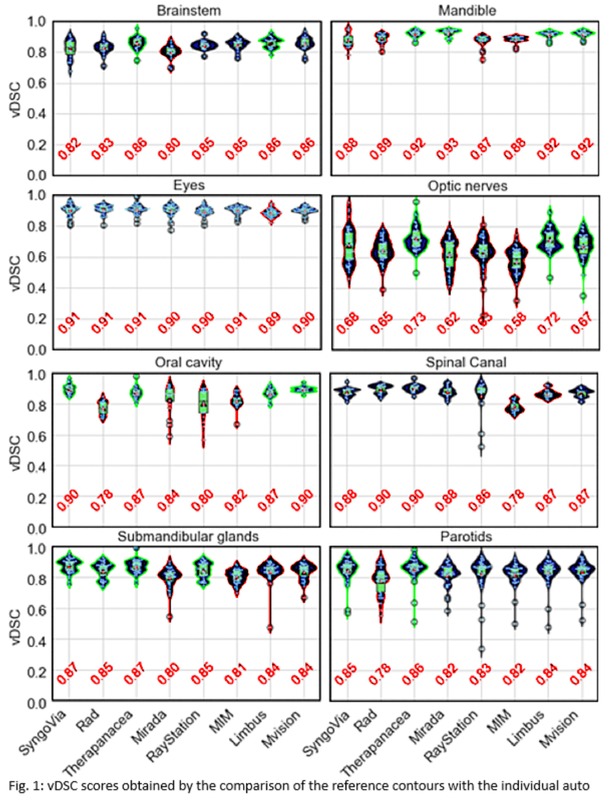


Fig. 1: vDSC scores obtained by the comparison of the reference contours with the individual auto segmentation solutions contours. Solutions that differ significantly ($p < 0.05$) from more than three other solutions are highlighted with either a green or red frame , indicating superior or inferior performance, respectively. Mean values are indicated as red numbers.

Internal