ESTRO

Abstract

Title

AI-based OAR delineation in brain T1w-MRI: Overcoming Inter- and Intra-observer variability

Authors

Thais Roque[1], Gorkem Gungor[2], Guillaume Klausner[3], Gokhan Gur[4], Ilkay Serbez[2], Bilgehan Temur[2], Alexandre Caffaro[5], Leo Hardy[5], Sanmady Kandiban[5], Ayoub Oumani[5], Basile Bertrand[5], Kumar Shreshtha[6], Banu Atalar[4], Nikos Paragios[7,5], Enis Ozyar[2]

Authors Affiliations

[1]TheraPanacea, Clinical and Partnerships Affairs, Paris, France; [2]Acibadem MAA University, Maslak Hospital, Department of Radiation Oncology, Istanbul, Turkey; [3]Université de médecine Pierre et Marie Curie, Paris Sorbonne Université, Université de médecine Pierre et Marie Curie, Paris, France; [4]Acibadem MAA University, Maslak Hospital, Department of Radiation Oncology, Istanbul, Turkey; [5]TheraPanacea, Research and Development, Paris, France; [6]TheraPanacea, AI Research Department, Paris, France; [7]CentraleSupelec, University of Paris-Saclay, Compute Sciences and Applied Mathematics, Gif-sur-Yvette, France

Purpose or Objective

Organ-at-risk (OAR) delineation is a key step for radiotherapy treatment (RT) planning. Manual delineation of OARs is a tedious process, time consuming, and prone to errors due to intra- and inter-observer variations. The management of brain tumors typically involves RT planning based on computed tomography (CT) and magnetic resonance imaging (MRI). MRI is acquired for detailed tumor localization and delineations of the target and OARs thanks to its excellent soft-tissue contrast. In this study, we propose to use an automatic artificial intelligence-based OAR segmentation to support MR-based treatment planning for brain.

Materials and Methods

ART-Net®, a CE-marked, FDA-cleared anatomically preserving deep-learning ensemble architecture for automatic contouring (AC) of OAR, was retrained on T1w-MRIs following the EPTN 2018 guidelines. For a total of 80 patients, automatic annotations of 25 OARs were performed and compared against the inter-expert variability between three expert annotators. In addition, 9 (chiasma, encephalon, left cochlea, left cornea, left eye lens, midbrain, posterior cerebellum, and right lacrimal gland) out of the 25 OARs were randomly selected and submitted to two independent observers for evaluation. Experts' contours used for RT delivery were blended with the ones delineated by ART-Net® at a 50%-50% ratio. Random blending at the patient level was performed guaranteeing that, among contours being evaluated per patient and OAR, the 50%-50% split was satisfied. Experts were asked to score ACs and the contours from clinical experts (CE) as A/acceptable, B/ acceptable after minor corrections, and C/ not acceptable for clinical use. To avoid any bias, experts were blind to the origin of the contours (manual or AI).

Results

Average Dice score coefficient (DSC) for ART-Net® on the testing set was 72.6% while that of the inter-expert variability was 69.1%. ART-Net® achieved mean dice greater than the inter-expert variability for all 25 organs (Tab.1). Encephalon reported the highest DSC 97.45 ± 0.50, while left vestibular and semi-circular canals (VSCC) reported the lowest DSC of 52.58 ± 14.55. These organs are normally not visible on T1Gd sequences, which might explain these results. 4 out of the 9 organs were evaluated as A for 100% of the cases. All 9 OARs achieved a 100% A+B score for the AI contours. A+B score for the manual delineation from experts was 100% for all but one organ (left eye lens).

Organ	Test Cases	Mean AI Dice [%]	Mean Expert Variability Dice [%]
Left retina	86	74.94 ± 9.36	66.9
Right optical nerve	89	75.36 ± 6.70	71.3
Left eye lens	82	78.16 ± 13.32	73.8
Right hypothalamus	89	62.20 ± 7.08	60.2
Chiasma	89	67.11 ± 8.44	62.8
Left cochlea	89	58.33 ± 17.38	48.6
Pons	89	91.26 ± 2.48	88.4
Spinal cord	82	85.30 ± 5.63	76.1
Right cornea	84	73.97 ± 10.94	70.8
Right vscc	87	53.39 ± 15.16	51.9
Medulla oblangata	88	84.79 ± 6.86	83.0
Midbrain	89	84.16 ± 6.63	79.6
Encephalon	89	97.45 ± 0.50	96.7
Anterior cerebellum	89	83.96 ± 7.73	78.8
Posterior cerebellum	89	93.84 ± 1.70	91.9
Left lacrimal gland	89	65.09 ± 11.90	58.8

Table 1: Number of test patients, DICE score achieved by selected OARs covering a variety of sizes and shapes for AI tool against expert intra variability dice score.

Figure 1: Examples of non-edited automatic delineations for brain MR scans using ART-Net® for OAR delineation. Left image (axial view) with ACs of lacrimal glands, retinas, encephalon, chiasma, hypothalamus, midbrain, hippocampus, anterior cerebellum; middle (sagittal view) with ACs of encephalon, chiasma, hypophyses, right hypothalamus, midbrain, pons, medulla oblongata, spinal cord, anterior cerebellum, posterior cerebellum; and right (coronal view) with ACs of hippocampus right, hippocampus left, midbrain, anterior cerebellum, posterior cerebellum, pons, medulla oblongata, spinal cord.

Conclusion

In this study, we investigated the advantages of using a CE-marked, FDA-cleared anatomically preserving deep-learning ensemble architecture for automatic segmentation of OARs for brain RT on MRI. This work illustrates the strength of using ART-Net® for segmentation to support MR-based treatment planning in the brain as it consistently generates delineations with high clinical acceptability which can lead to reduction in variability of practice. Future work will include dosimetric evaluations.