CBCT Segmentation in Head and Neck ART: A Longitudinal Deep Learning Approach

Lucía Cubero[1,2][0000−0003−3134−1433], Elena Navarro Alfaro[1], Irina Rakotoarisedy[1], Anaïs Barateau[2], Joël Castelli[2], Renaud de Crevoisier[2], Oscar Acosta[2][0000−0002−5447−1479], and Javier Pascau[1,3][0000−0003−1484−731X]

[1] Departamento de Bioingeniería, Universidad Carlos III de Madrid, Spain
[2] Université Rennes, CLCC Eugène Marquis, Inserm, LTSI—UMR 1099, F-35000 Rennes, France
[3] Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Spain

Abstract. Adaptive radiotherapy (ART) for head and neck cancer (HNC) relies on repeated cone-beam CT (CBCT) imaging to monitor anatomical changes throughout treatment course. Accurate organ-at-risk (OAR) segmentation on these images is essential for dose monitoring, yet most existing CBCT segmentation methods process each fraction independently, ignoring temporal anatomical changes. We propose a longitudinal ConvLSTM-UNet for multi-organ segmentation of CBCT-derived synthetic CT (sCT) images. The framework was evaluated under two clinically relevant settings: fully automatic autoregressive inference and replanning-guided inference that updates temporal memory using contours from replanning CTs. Compared with a nnU-Net baseline, replanning-guided inference achieved the best overall performance (mean DSC 77.4%, mean ASD 1.67 mm), significantly improving 12 of 17 OARs, while autoregressive inference remained competitive without manual updates. We further performed a longitudinal dosimetric analysis that revealed significant dose increases in the parotid and submandibular glands. 46% of patients exceeded the 26 Gy contralateral parotid mean-dose constraint during treatment despite remaining below it on the planning CT based on rigid dose propagation. These results demonstrate that longitudinal deep learning improves multi-organ CBCT segmentation while enabling automated dosimetric monitoring, supporting its potential for ART workflows.

Keywords: head and neck cancer · adaptive radiotherapy · longitudinal segmentation · CBCT · synthetic CT · dosimetric monitoring

1 Introduction

Head and neck cancer (HNC) is among the most common cancers worldwide, accounting for nearly 950,000 new cases and 500,000 deaths each year, and radiotherapy (RT) is a cornerstone of its treatment [4]. However, the close proximity of target volumes and organs at risk (OARs) requires highly accurate OAR

delineation to ensure tumor control while minimizing radiation-induced toxicities. Anatomical changes occurring throughout the treatment course, including weight loss and tumor or gland shrinkage, may lead to discrepancies between the planned and delivered dose distributions. Adaptive radiotherapy (ART) addresses these changes through repeated imaging, typically using Cone-Beam CT (CBCT), to monitor patient anatomy and guide treatment adaptation. Consequently, accurate automatic OAR segmentation on longitudinal CBCT images is a key component for dose monitoring and ART, but remains challenging due to poor image quality, artifacts, and the scarcity of annotated CBCT datasets [9].

Deep learning (DL) has become the standard approach for OAR segmentation on planning CT (pCT) images, but its application to CBCT remains more limited. To overcome the domain gap, previous studies have explored image enhancement, domain adaptation, pseudo-label generation, deformable image registration (DIR), and synthetic CT (sCT) generation. Among these strategies, sCT-based workflows are particularly attractive because they combine CT-like image quality with the daily anatomical information acquired from CBCT. Recent studies have demonstrated accurate segmentation on sCT using personalized CT models [8] and registration-guided deep learning frameworks [7]. In our previous work [1], we addressed the limited availability of annotated CBCT datasets by propagating pCT contours to CBCT-derived sCT images using DIR, and trained a 3D full-resolution nnU-Net to segment 25 OARs on CT and sCT images. The proposed framework demonstrated accurate geometric and dosimetric performance. However, as in previous sCT-based segmentation approaches, each treatment fraction was processed independently, without exploiting the temporal continuity of anatomical changes throughout RT.

Longitudinal segmentation models aim to exploit this temporal information by incorporating knowledge from previous treatment fractions as prior images, contours, or learned representations. Zhao et al. [10] introduced a recurrent 3D U-Net with embedded LSTM units that propagated anatomical information across fractions, demonstrating improved progressive OAR segmentation over conventional memory-free networks. However, their framework was limited to binary segmentation and was trained using simulated longitudinal data, leaving its applicability to multi-organ delineation on real clinical ART datasets largely unexplored.

In this work, building upon our previous segmentation framework [1], we incorporate longitudinal information into multi-organ segmentation of HNC CBCT images using a real clinical ART cohort. We further evaluate the clinical utility of the proposed method by investigating its ability to monitor dose variations throughout treatment and identify clinically meaningful deviations that may support adaptive replanning decisions. To the best of our knowledge, this is the first study to combine longitudinal DL-based multi-organ segmentation with dosimetric monitoring for HNC ART.

Table 1: Database characteristics for train and test sets.

Set	Patients	CT			CBCT
		Planning	Replanning	Total	
Train	129	129	95	224	818
Test	26	26	77	103	156
Dimensions (voxels)		$512\times512\times(79–488)$			$(270–410)^2\times(81–132)$
Spacing (mm)		$(0.578–1.367)^2\times(1–3)$			$1\times1\times2$

2 Materials and Methods

2.1 Database

Patients were retrospectively selected from our previously cohort [1], preserving the same training and test split (Table 1). All patients were adults (18–75 years) diagnosed with HNC and treated with external beam RT at CLCC Eugène Marquis (Rennes, France). All CBCT scans were acquired using the XVI imaging system (Elekta Versa HD). Patients with fewer than two CBCT scans or CBCTs containing fewer than 50 axial slices were excluded, resulting in a cohort of 155 patients. The training set comprised 129 patients (224 CT and 818 CBCT scans) treated with a standard protocol. The independent test set consisted of 26 patients enrolled in the ARTIX trial [5], all diagnosed with locally advanced oropharyngeal carcinoma (T3–T4/N2–N3) and undergoing weekly adaptive replanning, yielding 103 CT and 156 CBCT scans. sCT images were generated from each CBCT using the method proposed by Boussot et al. [2, 3] to reduce the impact of CBCT image artifacts during segmentation. sCTs preserved CBCT resolution and size, with Hounsfield Units (HU) ranging from -1024 to 2000.

2.2 Preprocessing

Pseudo-labels for the CBCT-derived sCTs were generated by pairing each sCT with the temporally closest CT from the same patient. The CT was registered to the sCT using rigid and third-order B-spline deformable registration. The resulting deformation was applied to the CT label map using nearest-neighbour interpolation, yielding pseudo-labels in the sCT space.

To ensure spatial consistency across longitudinal scans, all CT and sCT images from each patient were rigidly registered to the first available sCT. Body masks were obtained by thresholding at -700 HU to remove background voxels. Finally, images were cropped to $256 \times 256 \times 128$ voxels, intensity values were clipped to $[-700, 700]$ HU, and normalized to $[0, 1]$.

2.3 Longitudinal Model

Our model follows the LSTM-UNet architecture proposed by Zhao et al. [10], combining a 3D U-Net backbone with ConvLSTM modules embedded in each skip connection. The encoder uses a base width of 32 across five levels (32, 64,

128, 256, 256), mirrored by a four-stage decoder (128, 64, 32, 32). At each time step t, the network receives the current image I_t, the previous image I_{t-1}, and the previous segmentation M_{t-1}, and predicts the current multi-organ segmentation $\hat{M}_t$. ConvLSTM hidden states are maintained independently at each resolution level and reset at the beginning of every patient sequence. Unlike the original binary formulation, both M_{t-1} and $\hat{M}_t$ are represented as multiclass label maps.

Data were split at the patient level (80%/20% training/validation), ensuring that all fractions from the same patient remained in the same subset. Light data augmentation was applied during training, using patient-wise geometric transformations (axial rotation, translation, scaling) to preserve spatial consistency across longitudinal scans and fraction-wise intensity augmentation (random scaling, Gaussian noise, gamma augmentation). The pCT was always used as I_0, whereas I_t corresponded to either an sCT or a replanning CT. Teacher forcing was adopted during training (Fig. 1), where the model received I_t, I_{t-1}, and the ground-truth (GT) segmentation M_{t-1} to predict $\hat{M}_t$, supervised against the reference mask M_t. To reduce over-reliance on previous segmentations, previous-mask dropout ($p = 0.15$) randomly replaced M_{t-1} with zeros. Class imbalance was addressed using inverse square-root class-frequency weighting (background weight = 0.1, OAR weights clipped to $[0.5, 3.0]$). The model was optimized using an equally weighted Dice and Cross-Entropy loss with Adam (learning rate 10^{-4}, weight decay 10^{-8}), mixed precision, ReduceLROnPlateau scheduling, and early stopping (max. 200 epochs).

As a baseline, a 3D full-resolution nnU-Net v2 [6] was trained for 2500 epochs using the same training/validation split and images as the proposed longitudinal model. Rather than reusing the nnU-Net from our previous study [1], the baseline was retrained on the present dataset to ensure a fair comparison, as the longitudinal cohort differed from the previous one in both the number of images and the sCT generation pipeline. Unlike the proposed longitudinal model, each CT or sCT image was segmented independently, without exploiting temporal relationships between treatment fractions.

2.4 Inference Settings

Inference was performed sequentially for each patient, preserving the chronological order of treatment fractions. Two inference settings were evaluated to represent different clinical workflows, differing only in how the prior mask was updated when a replanning CT was available.

Autoregressive Inference. This setting represents a fully automatic clinical workflow, where manual contours beyond the pCT are unavailable. Following initialization with the pCT reference mask, each predicted segmentation $\hat{M}_{t-1}$ was used as the prior for the next fraction image, regardless of image modality.

Replanning-guided Inference. This setting simulates an ART workflow in which manual contours are available for replanning CTs. Inference remained autoregressive except when the previous image corresponded to a replanning CT,

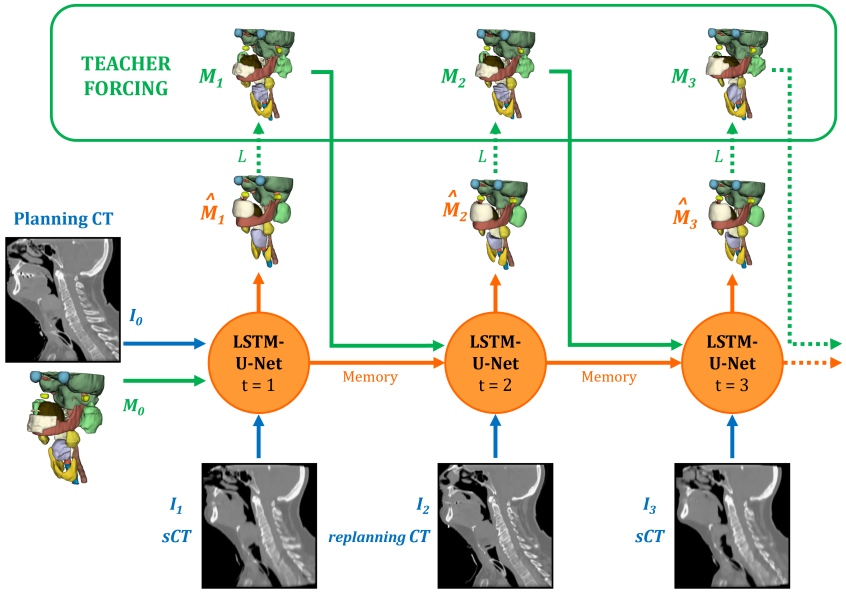

Fig. 1: Training strategy. At each time step, the previous reference mask (M_{t-1}) is provided together with the current and previous images, while the predicted segmentation ($\hat{M}_t$) is supervised against the current reference mask (M_t).

in which case its reference segmentation replaced the predicted prior mask for the next prediction. Subsequent fractions were again segmented autoregressively until another replanning CT became available. Importantly, reference masks were never used to segment the corresponding replanning CT itself, but only to update the temporal memory for future predictions.

Final segmentation maps were obtained by applying arg max to the output logits. No post-processing was applied. Performance was evaluated using the Dice Similarity Coefficient (DSC) and Average Surface Distance (ASD). Statistical comparisons between the two inference settings and the nnU-Net baseline were performed for each OAR using paired Wilcoxon signed-rank tests with Holm correction ($p_{\text{Holm}} < 0.05$).

2.5 Dosimetric Analysis

A dosimetric analysis was performed to evaluate whether automatic OAR segmentations derived from sCT could identify clinically relevant dose deviations throughout treatment. Patients received 70 Gy in 35 fractions according to the ARTIX protocol [5]. The planned dose distribution, calculated on the pCT (Pinnacle, Philips), was rigidly propagated to all sCT images. Mean dose (D_{mean}) was extracted from the manual pCT contours and the corresponding automatic sCT segmentations, analyzing ipsilateral and contralateral organs separately.

Longitudinal dose changes were quantified by fitting, for each patient and OAR, a linear regression of D_{mean} versus treatment week (week 0 corresponding to the pCT). Population-level significance was assessed using one-sample Wilcoxon signed-rank tests with Benjamini–Hochberg correction ($\alpha = 0.05$). Given its clinical relevance for xerostomia, the contralateral parotid gland was further analyzed with respect to the 26 Gy QUANTEC mean dose constraint by identifying patients who exceeded this threshold during treatment. Parotid gland volume and medial displacement were also quantified, and their association with dose changes was assessed using Pearson correlation.

3 Results

3.1 Longitudinal OAR Segmentation Performance

Table 2 summarizes the segmentation performance of the two longitudinal inference settings and the nnU-Net baseline on the sCT test set.

Replanning-guided inference consistently outperformed autoregressive inference across all 17 OARs. Mean DSC increased from $73.8 \pm 15.6\%$ to $77.4 \pm 14.9\%$, while mean ASD decreased from 2.00 ± 2.21 to 1.67 ± 2.06 mm, demonstrating the benefit of updating the temporal memory with manual contours whenever replanning CTs are available.

Compared with the nnU-Net baseline, replanning-guided inference achieved significantly higher DSC for 12 of the 17 OARs, with the largest improvements observed for the larynx (+12.6 DSC), lips (+12.6 DSC), esophagus (+7.0 DSC) and esophageal inlet (+7.2 DSC). In contrast, nnU-Net performed significantly better only for the brain (+1.0 DSC), optic nerves (+2.8 DSC) and mandible (+1.5 DSC), while no significant differences were observed for the eyeballs or thyroid. ASD showed a similar trend, with significantly lower values for most of the structures that showed an improvement in DSC.

Despite relying exclusively on its own previous predictions, autoregressive inference remained competitive with nnU-Net. Significant improvements were observed for six OARs (brainstem, inner ears, TMJs, lips, larynx and esophagus) with the autoregressive model, whereas nnU-Net outperformed it in seven. No significant differences were found for the remaining structures.

Overall, replanning-guided inference achieved DSC values above 70% for all OARs except the optic nerves (63.8%) and esophageal inlet (51.5%), High accuracy was obtained not only for large structures such as the brain (94.6%), mandible (86.8%), spinal canal (85.9%) and oral cavity (85.7%), but also for several clinically relevant smaller OARs, including the brainstem (84.8%), parotid glands (81.4%), eyeballs (80.8%), esophagus (77.5%), submandibular glands (74.0%), inner ears (73.4%) and pharyngeal constrictor muscles (70.0%). ASD remained below 2 mm for all but three structures (oral cavity, larynx and esophageal inlet).

3.2 Dosimetric Analysis

Population-level analysis identified significant increases in mean dose for both parotid glands (ipsilateral: +0.600 Gy/week; contralateral: +0.563 Gy/week)

Table 2: Dice Similarity Coefficient (DSC) and Average Surface Distance (ASD) for each organ at risk (OAR) on test, comparing nnU-Net with autoregressive and replanning-guided (Replan.) inference. Bold indicates the best result and * a statistically significant difference from nnU-Net. TMJs: Temporomandibular Joints; SMGs: Submandibular Glands; PCM: Pharyngeal Constrictor Muscles.

OAR	n	DSC (%)			ASD (mm)		
			Longitudinal Model			Longitudinal Model	
		nnU-Net	Autoregressive	Replan.	nnU-Net	Autoregressive	Replan.
Brain	156	**95.6 ± 2.0**	94.3 ± 4.5*	94.6 ± 4.6*	**0.84 ± 0.50**	1.14 ± 1.34*	1.09 ± 1.33*
Brainstem	156	81.8 ± 4.6	82.2 ± 12.3*	**84.8 ± 12.5***	1.81 ± 0.51	1.77 ± 1.35*	**1.53 ± 1.37***
Spinal canal	156	85.4 ± 5.8	83.9 ± 8.2*	**85.9 ± 8.4***	**1.19 ± 0.64**	1.47 ± 1.62*	1.31 ± 1.65
Eyeballs	153	**82.4 ± 8.2**	78.0 ± 11.3*	80.8 ± 11.2	1.47 ± 1.74	1.73 ± 2.13*	**1.47 ± 1.88**
Optic nerves	150	**66.6 ± 13.6**	62.4 ± 15.1*	63.8 ± 15.4*	**0.98 ± 0.75**	1.25 ± 1.41*	1.22 ± 1.42*
Inner ears	156	66.4 ± 12.6	70.4 ± 14.8*	**73.4 ± 14.6***	1.43 ± 0.77	1.47 ± 2.83*	**1.11 ± 0.73***
Parotid glands	156	80.6 ± 4.7	76.3 ± 9.3*	**81.4 ± 10.0***	**1.68 ± 0.32**	2.40 ± 3.94*	1.91 ± 3.96*
Mandible	156	**88.3 ± 2.1**	85.4 ± 5.5*	86.8 ± 5.5*	**0.97 ± 0.20**	1.29 ± 0.89*	1.16 ± 0.89*
TMJs	156	69.8 ± 8.0	71.1 ± 8.5*	**74.4 ± 8.5***	**1.69 ± 0.66**	2.07 ± 4.33	1.90 ± 4.43*
Oral cavity	156	82.8 ± 6.1	81.5 ± 5.3	**85.7 ± 5.4***	2.82 ± 1.10	3.17 ± 0.97*	**2.37 ± 0.93***
Lips	156	64.4 ± 12.1	70.9 ± 9.2*	**77.0 ± 8.6***	2.74 ± 1.48	2.19 ± 1.31*	**1.65 ± 1.16***
SMGs	156	71.1 ± 7.8	69.3 ± 9.0	**74.0 ± 9.1***	1.82 ± 0.48	2.08 ± 1.02*	**1.71 ± 1.01***
Larynx	156	64.7 ± 13.3	71.0 ± 13.6*	**77.3 ± 12.7***	3.45 ± 1.28	2.87 ± 1.13*	**2.07 ± 0.98***
PCM	156	65.7 ± 9.6	64.5 ± 12.9	**70.0 ± 10.6***	1.73 ± 0.90	1.84 ± 1.08	**1.45 ± 0.81***
Thyroid	156	**76.4 ± 10.0**	71.7 ± 11.5*	75.6 ± 9.9	1.46 ± 0.55	1.76 ± 0.60*	**1.46 ± 0.47**
Esophagus	156	70.5 ± 10.1	74.3 ± 8.2*	**77.5 ± 7.4***	2.21 ± 1.30	1.74 ± 0.79*	**1.55 ± 0.83***
Esophageal inlet	156	44.3 ± 23.3	46.4 ± 23.3	**51.5 ± 23.7***	3.76 ± 3.06	3.72 ± 3.50	**3.40 ± 3.49***
Mean		73.9 ± 15.6	73.8 ± 15.6	**77.4 ± 14.9**	1.89 ± 1.43	2.00 ± 2.21	**1.67 ± 2.06**

and submandibular glands (ipsilateral: +0.168 Gy/week; contralateral: +0.254 Gy/week), whereas no significant trend was observed for the remaining OARs.

The contralateral parotid gland was further analyzed (Fig. 2). Although the population median D_{mean} remained below the 26 Gy QUANTEC constraint throughout treatment, 12 of 26 patients (46%) exceeded this threshold despite remaining below it on the pCT, including six during the first treatment week. The remaining 14 patients (54%) stayed below the constraint throughout treatment. Representative examples are shown in Fig. 3: one patient exhibited a progressive increase from 25.4 Gy to 29.4 Gy by week 3, whereas another remained below the threshold (20.6–22.3 Gy) throughout treatment.

Parotid analysis revealed significant ipsilateral volume reduction (-0.59 cm^3/ week, $p = 0.004$), and a non-significant contralateral shrinkage (-0.22 cm^3/week, $p = 0.57$). In contrast, the contralateral gland showed significant medial displacement (+0.31 mm/week, $p < 0.001$), which correlated with its individual D_{mean} increase in patients exceeding the 26 Gy constraint (Pearson $r = 0.64$, $p = 0.03$), whereas volume change did not ($r = 0.12$, $p = 0.72$).

4 Discussion

This study investigated whether incorporating longitudinal information improves multi-OAR sCT delineation in HNC ART and whether the resulting segmentations can support dosimetric monitoring. Unlike previous work based on binary

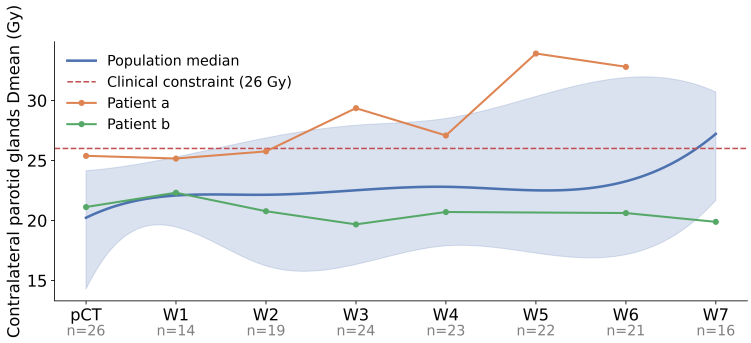

Fig. 2: Population evolution of the contralateral parotid D_{mean} from the planning CT (pCT) throughout treatment, with n denoting the number of patients at each week (W). The solid line and shaded area the interquartile range. The dashed red line indicates the 26 Gy QUANTEC constraint. Patients a and b are highlighted.

segmentation and simulated longitudinal datasets [10], we extended memory-based segmentation to 17 OARs and evaluated it on a real longitudinal CT/sCT cohort derived from our previous framework [1].

Two inference settings were evaluated to reflect common ART workflows. Replanning-guided inference consistently outperformed both autoregressive inference and the nnU-Net baseline, demonstrating the benefit of updating temporal memory whenever reference contours become available. Despite relying exclusively on previous predictions, autoregressive inference achieved performance comparable to nnU-Net, suggesting that temporal information remains beneficial even in a fully automatic workflow. Clinically, this also indicates that manually corrected sCT contours could be incorporated into the temporal memory to improve subsequent predictions without modifying the network architecture.

The longitudinal model achieved its largest improvements for anatomically variable structures such as the larynx, lips, parotid and submandibular glands, whereas nnU-Net performed better for highly stable structures including the brain and mandible, likely because it segments each image independently and is therefore unaffected by temporal error propagation. To ensure a fair comparison, the nnU-Net baseline was retrained using the same cohort, preprocessing pipeline and train/validation split as the longitudinal model, rather than reusing our previous model [1]. In addition, the previous closed-source sCT synthesis method was replaced by a publicly available state-of-the-art approach, improving the reproducibility of the complete framework.

Beyond segmentation accuracy, this work demonstrates the feasibility of longitudinal dosimetric monitoring using automatically generated OAR contours on CBCT. Significant dose increases were observed for both parotid and submandibular glands. In particular, the contralateral parotid mean dose increased from approximately 20 to 26 Gy over treatment, and 46% of patients exceeded

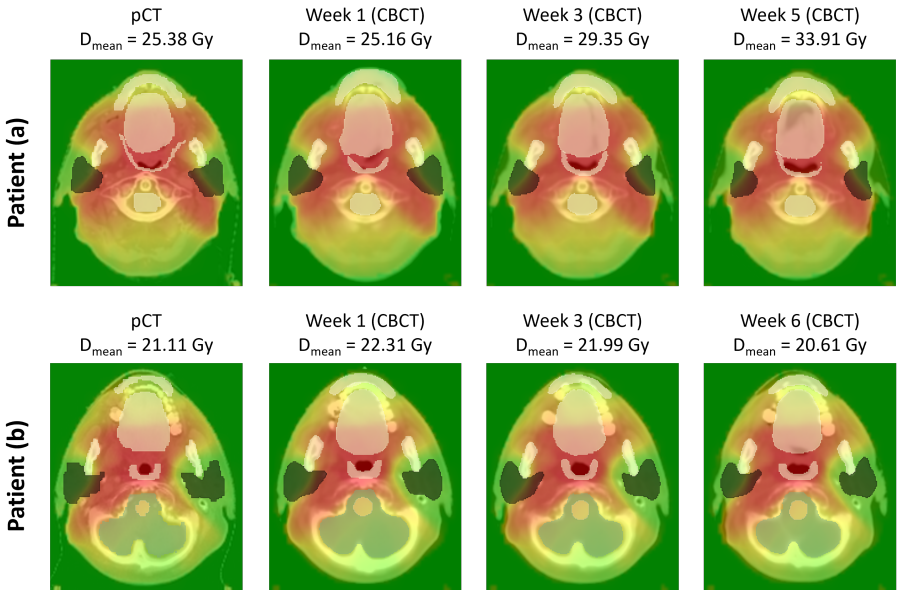

Fig. 3: Representative dose evolution in two patients. Axial dose distributions and automatic contralateral parotid contours (black contour) are shown. Patient (a): progressive dose increase exceeding the 26 Gy constraint at week 3. Patient (b): stable dose remaining below the constraint throughout treatment.

the 26 Gy planning constraint despite satisfying it initially. It should be noted that the dose was rigidly propagated rather than recalculated. Anatomical analysis further suggested that this increase was associated with a progressive medial displacement of the contralateral parotid gland rather than with gland shrinkage: only the ipsilateral gland showed significant volume loss, whereas medial displacement significantly correlated with the individual dose increase among patients exceeding the planning constraint. These findings highlight the potential of longitudinal monitoring to identify patient-specific dosimetric deviations that could support adaptive replanning.

Several limitations should be acknowledged. First, no manual annotations were available for the sCT images, requiring pseudo-label generation through CT-to-sCT deformable registration. Consequently, reported metrics reflect agreement with the DIR algorithm rather than true anatomical accuracy. Second, the study was conducted on a relatively small single-centre cohort. Third, no dose recalculation was performed on the sCT images; instead, the planned dose distribution was rigidly propagated, so global anatomical changes such as patient weight loss were not reflected in the dose distribution itself. Finally, autoregressive inference remains susceptible to temporal error accumulation. Future work should investigate uncertainty-aware or registration-guided temporal updates,

together with validation on larger multicentre cohorts and prospective clinical evaluation of the proposed framework.

In conclusion, this work extends longitudinal memory-based segmentation to multi-organ delineation on a real HNC ART cohort and demonstrates its potential clinical value for longitudinal dose monitoring. The proposed framework provides accurate automatic OAR segmentation and further improves performance when replanning information is incorporated, offering a promising foundation for personalized ART.

Acknowledgments. Research supported by INCA (French National Cancer Institute), projects SeQRTH-55 (Call SEQ-RTH 22), PID2023-149604OB-I00 (MCIU/AEI /10.13039/501100011033 and Comunidad de Madrid (MAGERIT-CM TEC-2024/COM-44). The authors have no competing interests to declare.

References

1. Removed for anonymization purposes
2. Boussot, V., Dillenseger, J.L.: Konfai: A modular and fully configurable framework for deep learning in medical imaging. arXiv preprint arXiv:2508.09823 (2025)
3. Boussot, V., Hémon, C., Nunes, J.C., Dillenseger, J.L.: Why registration quality matters: Enhancing sct synthesis with impact-based registration. arXiv preprint arXiv:2510.21358 (2025). https://doi.org/10.48550/arXiv.2510.21358
4. Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R.L., Soerjomataram, I., et al.: Global cancer statistics 2022: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians **74**(3), 229–263 (2024). https://doi.org/10.3322/caac.21834
5. Castelli, J., Thariat, J., Benezery, K., Hasbini, A., Gery, B., Berger, A., et al.: Weekly adaptive radiotherapy vs standard intensity-modulated radiotherapy for improving salivary function in patients with head and neck cancer: A phase 3 randomized clinical trial. JAMA Oncology **9**(8), 1056 (Aug 2023). https://doi.org/10.1001/jamaoncol.2023.1352
6. Isensee, F., Jaeger, P.F., Kohl, S.A.A., Petersen, J., Maier-Hein, K.H.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. Nature Methods **18**(2), 203–211 (2021). https://doi.org/10.1038/s41592-020-01008-z
7. Liang, X., Morgan, H., Bai, T., Dohopolski, M., Nguyen, D., Jiang, S.: Deep learning based direct segmentation assisted by deformable image registration for cone-beam ct based auto-segmentation for adaptive radiotherapy. Physics in Medicine & Biology **68**(4), 045012 (Feb 2023). https://doi.org/10.1088/1361-6560/acb4d7
8. Liu, Y., Yang, B., Chen, X., Zhu, J., Ji, G., Liu, Y., et al.: Efficient segmentation using domain adaptation for mri-guided and cbct-guided online adaptive radiotherapy. Radiotherapy and Oncology **188**, 109871 (Nov 2023). https://doi.org/10.1016/j.radonc.2023.109871
9. Rabe, M., Kurz, C., Thummerer, A., Landry, G.: Artificial intelligence for treatment delivery: image-guided radiotherapy. Strahlentherapie und Onkologie **201**(3), 283–297 (2025). https://doi.org/10.1007/s00066-024-02277-9
10. Zhao, H., Liang, X., Meng, B., Dohopolski, M., Choi, B., Cai, B., et al.: Progressive auto-segmentation for cone-beam computed tomography-based online adaptive radiotherapy. Physics and Imaging in Radiation Oncology **31**, 100610 (Jul 2024). https://doi.org/10.1016/j.phro.2024.100610