Patient-specific 4DCT respiratory motion synthesis using tumor-aware GANs

Yi-Heng Cao, Vincent Jaouen, Vincent Bourbonne, François Lucia, Nicolas Boussion, Ulrike Schick, Julien Bert, Dimitris Visvikis

Patient-specific 4DCT respiratory motion synthesis using tumor-aware GANs

Yi-Heng Cao[1], Vincent Jaouen[1], Vincent Bourbonne[1,2], François Lucia[1,2], Nicolas Boussion[1,2], Ulrike Schick[1,2], Julien Bert[1,2], Dimitris Visvikis[1]

[1] LaTIM, UMR Inserm 1101, Université de Bretagne Occidentale, IMT Atlantique, Brest, France

[2] CHRU Brest University Hospital, Brest, France

***Abstract*—Four-dimensional computed tomography (4DCT) is required in lung radiotherapy treatment planning to track tumor motion for more accurate dose coverage. However, such acquisitions expose the patient to more radiation than a standard CT protocol. In previous works, we demonstrated the feasibility of patient-specific 4DCT generation from static 3DCT images using generative adversarial networks (GAN) conditioned on the actual patient's respiratory amplitude. Synthetic motion was achieved globally in the lung, but with yet unsatisfactory accuracy at the tumor level. This work addresses this issue by better taking into account tumor motion through tumor awareness. We condition the image-to-image GAN architecture by a static 3DCT image, a respiratory amplitude and a further condition on the tumor segmentation mask. We train the model under a combined segmentation and weighted L_1 objective. We performed experiments on synthetic phantoms, where we demonstrate better tumor motion synthesis both qualitatively and quantitatively, and show preliminary results on clinical data.**

I. Introduction

FOUR-dimensional computed tomography (4DCT) imaging is used routinely in lung cancer radiotherapy treatment planning to accurately track tumor motion during breathing. However, such acquisitions lead to increased radiation dose delivered to the patient, up to six times a single-helical (3DCT) acquisition [1]. For this reason, alternative planning methods aiming at reducing exposure to imaging-induced radiation are an active field of research, with recent studies showing promising results using deep machine learning techniques [2].

We explored previously the potential interest of generating 4DCT phases from static 3D images using deep generative adversarial networks (GANs) [3] trained on real 4DCT data. Chang et al. [4] proposed a method combining a conditional variational encoder with a latent regressor for generating realistic 4D-XCAT phantom breathing dynamics. Romaguera et al. [5] presented an unsupervised predictive framework to generate 4D MRI volumes by learning a distribution of motion fields over a population dataset. In these approaches, the generated images do not correspond to the actual patient's respiratory dynamics but are rather a population average. To overcome this issue, we introduced recently a novel image-to-image network architecture to further condition image synthesis on respiratory motion amplitude [6]. Such value can typically be obtained in clinical routine from respiratory tracking devices. However, while the global lung synthetic motion was recovered, accuracy at smaller scales (i.e. at the tumor level) was not satisfactory.

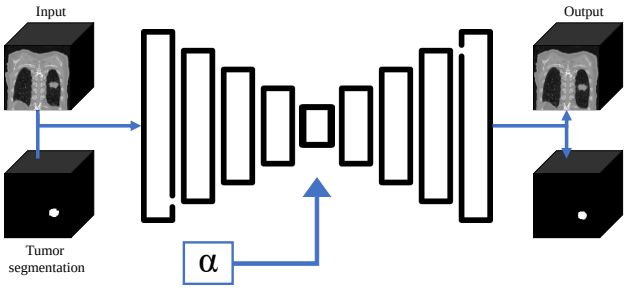


Fig. 1: Generator architecture with scalar value α and a tumor segmentation as input with image.

In this paper, we propose to further condition the generative model for improved synthesis of dynamic 4DCT images from 3D images using GANs. We force the model to account for tumor motion through a tumor aware loss [7] by giving more weight to the L_1 loss in the tumor region and combining it with a segmentation objective. We performed experiments on synthetic phantoms and show preliminary results on clinical data to study the influence of this conditioning. Results demonstrate superior tumor motion synthesis both qualitatively and quantitatively.

II. Method

We consider a reference breath-hold CT scan $\mathcal{I}(\boldsymbol{x})$, where $\boldsymbol{x} \in \mathbb{R}^3$ is the image domain, and its corresponding tumor segmentation mask $\mathcal{S}(\mathcal{I})$. Our objective is to synthesize a vector-valued phase-gated 4DCT acquisition $\mathcal{J}^* : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}$, where $\mathcal{J}^*(\boldsymbol{x}, g)$ is the image value at location $\boldsymbol{x}$ and gate $g \in \{1, 2, ..., N_g\}$. To this end, we learn a mapping $\varphi_g(\boldsymbol{x}, \alpha, \mathcal{S})$ conditioned on three terms: 1) the input image $\mathcal{I}$, 2) its tumor segmentation mask $\mathcal{S}(\mathcal{I})$, and a (scalar) respiratory amplitude α such that a generated image $\mathcal{I} \circ \varphi_g(\boldsymbol{x}, \alpha, \mathcal{S}) = \mathcal{J}^*(\boldsymbol{x}, \alpha)$ is close to the real phase $\mathcal{J}(\boldsymbol{x}, g)$. The value α can be obtained from e.g. a respiratory tracking belt.

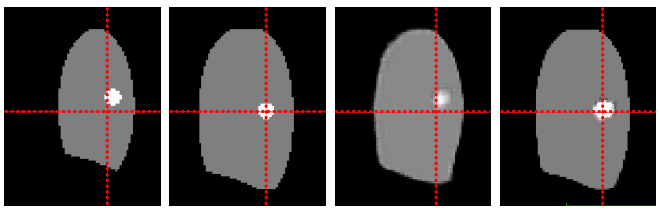
We modify our previous phase-to-phase model architecture [6] so that it can accept as input both the images and the tumor segmentation masks as two channels (Fig. 1). To further focus on the tumor region during image to image translation, we consider a weighted L_1 loss, or $L_1(\omega)$ with varying weighting depending on the region considered (tumor or remaining tissues), where this value is treated as a hyper-parameter.

During training, pairs of phases and masks to be mapped to one another are selected at random. The respiratory amplitude value α is injected into the bottleneck of the generator through an AdaIN mechanism [6]. The generator outputs an image $\mathcal{J}^*$ and a tumor segmentation corresponding to the target phase.

III. Experiments and results

We first evaluated the model on phantom images. These images represent synthetic lung-like images by using pairs of ellipsoids with an arc-shaped cut to mimic the diaphragm. We inserted a spherical hyper-signal object to simulate the presence of a tumor (Fig. 2). We generated various motion amplitudes through random morphological dilations along both the anterior-posterior and the superior-inferior directions and displaced the tumor's center of mass accordingly. We used $800/200$ images for training/testing respectively. The α value was the lung volume variation between two phases and the weights ω for the L_1 were set to 10 for the tumor region and 1 for the remaining tissues. We then evaluated our model on clinical 4DCT data. We used 79/32 images from 35/11 patients respectively for training/testing. Since no tumor segmentation was available, we used a pre-trained nnU-Net model [8] to generate tumor masks that we manually corrected if required. The α value used is the signal respiratory amplitude in mm. The models were evaluated using the Dice similarity coefficient (DSC) between the generated and the true tumor mask, as well as the Euclidean distance between their centers of masses.

Fig. 2 shows representative results on the synthetic dataset with and without tumor awareness (TA). In both cases, a global lowering of the diaphragm corresponding to the target phase was successfully achieved. However, inspecting more closely, the tumor did not move accordingly, which is a problem corrected through TA. These qualitative observations are confirmed quantitatively on the entire dataset with a sharp increase of similarity to ground truth using the proposed TA scheme Table I. The initial errors show that there are poor tumor overlaps between phases, with a DSC of 0.28 and a displacement distance of 3.73mm. When using the L_1 objective only, tumor motion was not well reproduced. However, when

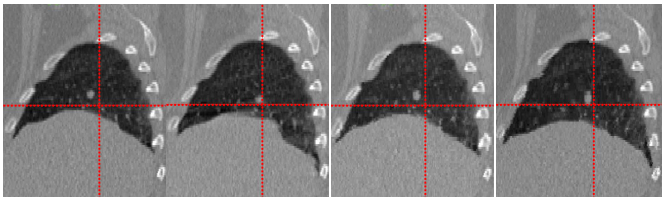


(a) Input (b) Reference (c) Without TA (d) With TA

Fig. 2: Example of phantoms results in the sagittal view.

Model	DSC	Distance
Initial error between phases	0.28	3.73
using L_1	0.29	3.36
TA using weighted $L_1(\omega)$	0.53	1.94
TA using weighted $L_1(\omega)$ + segmentation loss	0.67	1.08

TABLE I: Ablation study on the objective terms for the phantom dataset.



(a) Input (b) Reference (c) Without TA (d) With TA

Fig. 3: Example of real clinical data visual synthesis results.

considering the various tumor awareness mechanism proposed (weighted L_1 and segmentation loss), a sharp improvement was obtained, with best results achieved using a combination of both TA terms with a DSC of 0.67 and a distance error of 1.08mm on average.

Fig. 3 shows a preliminary visual result on a patient from our clinical 4DCT dataset, where we observed on average better tumor motion with TA. However, quantitative evaluation is not yet satisfactory (not shown in this summary) due to training conditions. Many patients showed little to no motion in the superior lobes. We believe this led the network to globally underestimate tumor motion amplitude.

IV. Conclusion

We have proposed a new method for patient-specific 4DCT respiratory motion synthesis from static CT images using a tumor-aware image-to-image GAN to allow for realistic modeling of the target motion. The weighting mechanism was validated on synthetic dynamic images where both weighted L_1 and segmentation losses showed beneficial effects. Further experiments will focus on designing more balanced training conditions to better compensate for the relative lack of tumor motion in the superior lobes.

References

[1] J. de Koste *et al.*, "Renal mobility during uncoached quiet respiration: An analysis of 4DCT scans," *International Journal of Radiation Oncology*Biology*Physics*, vol. 64, no. 3, pp. 799–803, 2006.

[2] A. Mylonas *et al.*, "A review of artificial intelligence applications for motion tracking in radiotherapy," *Journal of Medical Imaging and Radiation Oncology*, vol. 65, Aug. 2021.

[3] V. Jaouen *et al.*, "4D respiratory motion synchronized image synthesis from static CT images using GANs," in *IEEE NSS/MIC*, Manchester, United Kingdom, Oct. 2019.

[4] Y. Chang *et al.*, "A generative adversarial network (GAN)-based technique for synthesizing realistic respiratory motion in the extended cardiac-torso (XCAT) phantoms," *Physics in Medicine & Biology*, vol. 66, May 2021.

[5] L. V. Romaguera *et al.*, "Probabilistic 4D predictive model from in-room surrogates using conditional generative networks for image-guided radiotherapy," *Medical Image Analysis*, vol. 74, Dec. 2021.

[6] Y.-H. Cao *et al.*, "Image and volume conditioning for respiratory motion synthesis using GANs," in *IEEE NSS MIC 2021*, Oct. 2021.

[7] A. Chartsias *et al.*, "Adversarial Image Synthesis for Unpaired Multi-modal Cardiac Data," in *Simulation and Synthesis in Medical Imaging*. Springer International Publishing, 2017, vol. 10557.

[8] F. Isensee *et al.*, "nnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation," *arXiv:1809.10486 [cs]*, Sep. 2018.