2026 JULY 19–22 VANCOUVER, BC
JOINT AAPM | COMP MEETING

VoRAIN-GAN: Volumetric Attention GAN Trained with Calibrated MVCT for kVCT Synthesis in Head and Neck Cancer of Tomotherapy

Tie Lyu[1,2], Chuanbin Xie[3], Gaolong Zhang[2], Nan Li[2], Shouping Xu[*1]

[1]National Cancer Center/Cancer Hospital- Chinese Academy of Medical Sciences and Peking Union Medical College, Department of Radiation Oncology, Beijing, China

[2]Beihang University, School of Physics, Beijing, China.

[3]The First Medical Center of PLA General Hospital, Department of Radiation Oncology, Beijing, China.

ABSTRACT

Purpose: During fractionated radiotherapy, anatomical changes—such as patient setup, tumor regression, and weight loss—may impact the suitability of the initial treatment plan, necessitating the application of adaptive radiation therapy (ART). The megavoltage computed tomography (MVCT) serves as the patient positioning image, enabling dose reconstruction. However, its limited soft-tissue contrast and significant noise limit its application in ART. This study aimed to develop an image synthesis network, VoRAIN GAN, that further improved the image quality of synthetic kVCT (skVCT) based on prior work.

Methods: MVCT and corresponding kVCT data from 63 patients with head and neck cancer were collected. Among them, 41 patients were used for network training, while 22 patients were reserved for testing. This study employs a CycleGAN network incorporating attention gates, residual blocks, and volume control loss, named "Volumetric Residual Attention-Integrated GAN (VoRAIN-GAN)" to synthesize skVCT images from MVCT. The MVCT images were calibrated using the CT-electron density (CT-ED) curve for MVCT and kVCT. The VoRAIN-GAN was trained using calibrated MVCT. For test cases, contours and plans from kVCT were directly propagated to skVCT and MVCT images for dose recalculation. The evaluation indicators include mean absolute error (MAE), structural similarity index measure (SSIM), and peak signal-to-noise ratio (PSNR) for image quality and gamma passing rate for dose.

Results: Compared with MVCT, the generated skVCT demonstrated substantially improved in image quality. The MAE, PSNR, and SSIM of MVCT were respectively 102±14.9 HU, 27.3±1.4 dB, and 0.917±0.018, and those of skVCT generated by the previously published model were 67.9±13.9 HU, 31.4±2.0 dB, and 0.960±0.02, while those of skVCT in this study were 59.6±6.0 HU, 32.7±1.2 dB, and 0.965±0.017, respectively. The gamma passing rates improved from 96.39 ± 2.13% to 97.13 ± 1.81% (1 mm/ 1%), from 98.63 ± 0.95% to 98.86 ± 0.81% (2 mm/ 2%), and from 99.29 ± 0.55% to 99.43 ± 0.50% (3 mm/3%), respectively.

Conclusions: This study developed a Cal VoRAIN-GAN with deeper discriminators, attention gates, a residual network, a volume-control loss, and calibrated MVCT to generate MVCT-based skVCT images. The proposed method effectively enhances image quality and improves dose calculation accuracy, as demonstrated in the original work, indicating potential to improve treatment accuracy and enable ART implementation.

Funding: This work was supported by the National Natural Science Foundation of China (No. 12375359) and CAMS Innovation Fund for Medical Sciences (CIFMS, 2024-I2M-C&T-B-076).

[name] Tie Lyu
[organization] National Cancer Center/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China
[email] lvtie0303@buaa.edu.cn

INTRODUCTION

In recent years, the worldwide incidence of cancer has continued to rise. The radiation therapy (RT) has advanced significantly. However, anatomical changes during treatment—such as weight loss, tumor shrinkage, and positioning variations—compromise its precision. Adaptive radiation therapy (ART) addresses this by dynamically adjusting plans based on real-time patient changes. Studies demonstrate that weekly ART can reduce mean accumulated parotid dose by 5.4%, and MR-LINAC-based ART showes dose reductions in most parotid glands, underscoring its clinical value for HNC patients.

Helical tomotherapy (HT) employs a rotating megavoltage X-ray source for image-guided RT. The daily megavoltage CT (MVCT) acquired for positioning offers potential for ART due to its linear CT-to-electron density relationship, enabling dose reconstruction. However, MVCT suffers from low signal-to-noise ratio and poor contrast because high-energy photons produce minimal photoelectric effect.Therefore, improving MVCT image quality is essential for effective ART implementation.

Various strategies have been explored to improve MVCT image quality for clinical use, including noise-reduction algorithms, reconstruction techniques, and deep learning methods. Among deep learning approaches, diffusion models are computationally expensive, U-Net variants tend to over-smooth textures, and standard GANs suffer from training instability. CycleGAN and its attention-enhanced variant have shown promise for synthetic kVCT generation, yet challenges remain in loss function design and fine detail recovery. Notably, no prior study has leveraged calibrated MVCT for network training.

To address these gaps, we propose an enhanced Attention CycleGAN that integrates attention gates, residual blocks, and volume-control loss, named "Volumetric Residual Attention-Integrated GAN (VoRAIN-GAN). VoRAIN-GAN is trained using calibrated MVCTs (Cal VoRAIN-GAN).

METHODS AND MATERIALS

MVCT and corresponding kVCT data from 63 patients with head and neck cancer were collected. Among them, 41 patients were used for network training, while 22 patients were reserved for testing. This study employs a VoRAIN-GAN to synthesize skVCT images from MVCT. The MVCT images were calibrated using the CT values-electron density (CT-ED) curve of MVCT and kVCT. The VoRAIN-GAN was trained using calibrated MVCT.

Ablation Experiment: We employed the loss function developed in this study to train the original Attention CycleGAN [1], named Volumetric Residual Attention Cycle GAN (VAC-GAN). VAC-GAN trained using a calibrated MVCT is referred to as Cal VAC-GAN. Compared with VAC-GAN, VoRAIN-GAN further deepens the discriminator. VoRAIN-GAN trained using a calibrated MVCT is referred to as Cal VoRAIN-GAN.

RESULTS

Figure 1 presents a comprehensive visual comparison among kVCT, original MVCT, calibrated MVCT, and synthetic kVCT (skVCT) images generated by various methods. Compared with $skVCT_{v1}$, the overall brightness distribution of $skVCT_{v2}$ is closer to that of kVCT, indicating that training using a calibrated MVCT enables the model to effectively learn global intensity mapping relationships. $skVCT_{v3}$, synthesized by the proposed "Cal VoRAIN-GAN," demonstrated superior overall image quality, exhibiting the closest resemblance to the reference kVCT in terms of anatomical fidelity, contrast resolution, and texture preservation.

To identify optimal parameters, attention CycleGAN variants were trained and evaluated, with image quality metrics summarized in Table 1. MVCT calibration yielded substantial improvements in image quality. The VAC-GAN-generated skVCT achieved MAE and PSNR values comparable to calibrated MVCT. This suggests that VAC-GAN still has potential for further development. Performance was further enhanced in the "Cal VAC-GAN" model. Furthermore, the deepened discriminator has led to sustained performance improvements, with 'Cal VoRAIN-GAN' performing best.
The gamma passing rates improved from 96.39 ± 2.13% to 97.13 ± 1.81% (1 mm/ 1%), 98.63 ± 0.95% to 98.86 ± 0.81% (2 mm/ 2%), and 99.29 ± 0.55% to 99.43 ± 0.50% (3 mm/3%), respectively.

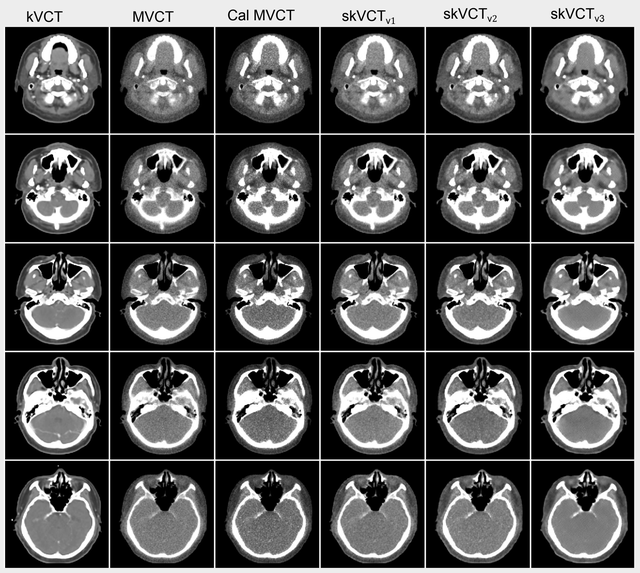

Figure 1. Comparisons between kVCT, MVCT, calibrated MVCT, and skVCT generated by different models. $skVCT_{v1}$: generated by VAC-GAN; $skVCT_{v2}$: generated by Cal VAC-GAN; $skVCT_{v3}$: generated by Cal VoRAIN-GAN.

Table 1. Comparison of MAE, SSIM, and PSNR parameters for synthetic images across.

Network	MAE (HU)	SSIM	PSNR (dB)
MVCT	102.0±14.9	0.917±0.018	27.3±1.4
MVCTcalibrated	68.7±6.8	0.924±0.021	32.0±1.2
VAC-GAN	67.9±13.9	0.960±0.020	31.4±2.0
Cal VAC-GAN	62.4±6.9	0.963±0.021	32.5±1.2
Cal VoRAIN-GAN	**59.6±6.0**	**0.965±0.017**	**32.7±1.2**

DISCUSSION

This study proposes a Cal VoRAIN-GAN that integrates attention gates, residual networks, and volume control loss with an innovative pre-calibration strategy for MVCT inputs. By increasing discriminator depth guided by loss curve analysis, the model significantly enhances MVCT image quality, producing sharper organ and soft-tissue boundaries with reduced noise, outperforming the attention CycleGAN [1].

Crucially, this work proposes, for the first time, a method for pre-calibrating MVCTs prior to generating skVCTs. Unlike previous studies, it leverages the fact that deep learning can focus exclusively on nonlinear mappings when linear transformations are handled beforehand [2,3]. Therefore, this pre-calibration approach simplified the learning space and improving synthetic image fidelity.

As shown in Figure 1, the brightness distribution of the uncalibrated $skVCT_{v1}$ is similar to that of the MVCT. The calibrated MVCT, however, is very close to the kVCT, and this advantage is directly inherited by the model trained using the calibration data.

Although many researchers equate MVCT-to-skVCT synthesis with CBCT-to-CT translation, MVCT possesses a precise linear CT-to-electron density relationship absent in CBCT [4], making pre-calibration not merely beneficial but a unique, essential, and non-negotiable step specific to MVCT-based synthetic kVCT generation.

CONCLUSIONS

This study presents Cal VoRAIN-GAN, a novel generative framework specifically engineered for generating high-quality synthetic kVCT (skVCT) images from MVCT. The architecture integrates multiple advanced components: a deepened discriminator, attention, residual networks, and a volume-control loss function. Crucially, the model incorporates pre-calibrated MVCT inputs, enabling the network to learn nonlinear tissue mappings more efficiently. Comprehensive evaluations demonstrate that the generated skVCT images exhibit significantly enhanced contrast resolution, reduced noise, and sharper soft-tissue boundaries compared to conventional methods. The proposed method effectively enhances image quality and improves dose calculation accuracy, as demonstrated in the original work, indicating potential to improve treatment accuracy and enable ART implementation.

REFERENCES

1. Lv T, Xie CB, Zhang YH, et al. A qualitative study of improving megavoltage computed tomography image quality and maintaining dose accuracy using cycleGAN-based image synthesis. Article. *Med Phys*. Jan 2024;51(1):394-406. doi:10.1002/mp.16633
2. LeCun Y, Bengio Y, Hinton G. Deep learning. Review. *Nature*. May 2015;521(7553):436-444. doi:10.1038/nature14539
3. Beardall WAV, Stan GB, Dunlop MJ. Deep Learning Concepts and Applications for Synthetic Biology. Review. *GEN Biotechnol*. Aug 2022;1(4):360-371. doi:10.1089/genbio.2022.0017
4. Yang M, Virshup G, Mohan R, Shaw CC, Zhu XR, Dong L. Improving accuracy of electron density measurement in the presence of metallic implants using orthovoltage computed tomography. Article. *Med Phys*. May 2008;35(5):1932-1941. doi:10.1118/1.2905030