Supervised deep learning–based synthetic computed tomography from kilovoltage cone-beam computed tomography images for adaptive radiation therapy in head and neck cancer

Chirasak Khamfongkhruea[1,2], Tipaporn Prakarnpilas[1], Sangutid Thongsawad[1,2], Aphisara Deeharing[2], Thananya Chanpanya[2], Thunpisit Mundee[2], Pattarakan Suwanbut[2], Kampheang Nimjaroen[1,2]

[1]Medical Physics Program, Princess Srisavangavadhana College of Medicine, Chulabhorn Royal Academy, Bangkok, Thailand
[2]Department of Radiation Oncology, Chulabhorn Hospital, Chulabhorn Royal Academy, Bangkok, Thailand

Purpose: To generate and investigate a supervised deep learning algorithm for creating synthetic computed tomography (sCT) images from kilovoltage cone-beam CT (kV-CBCT) images for adaptive radiation therapy (ART) in head and neck cancer (HNC).

Materials and Methods: This study generated the supervised U-Net deep learning model using 3,491 image pairs from planning CT (pCT) and kV-CBCT datasets obtained from 40 HNC patients. The dataset was split into 80% for training and 20% for testing. The evaluation of the sCT images compared to pCT images focused on three aspects: Hounsfield units accuracy, assessed using mean absolute error (MAE) and root mean square error (RMSE); image quality, evaluated using the peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) between sCT and pCT images; and dosimetric accuracy, encompassing 3D gamma passing rates for dose distribution and percentage dose difference.

Results: MAE, RMSE, PSNR, and SSIM showed improvements from their initial values of 53.15 ± 40.09, 153.99 ± 79.78, 47.91 ± 4.98 dB, and 0.97 ± 0.02 to 41.47 ± 30.59, 130.39 ± 78.06, 49.93 ± 6.00 dB, and 0.98 ± 0.02, respectively. Regarding dose evaluation, 3D gamma passing rates for dose distribution within sCT images under 2%/2 mm, 3%/2 mm, and 3%/3 mm criteria, yielded passing rates of 92.1% ± 3.8%, 93.8% ± 3.0%, and 96.9% ± 2.0%, respectively. The sCT images exhibited minor variations in the percentage dose distribution of the investigated target and structure volumes. However, it is worth noting that the sCT images exhibited anatomical variations when compared to the pCT images.

Conclusion: These findings highlight the potential of the supervised U-Net deep learningmodel in generating kV-CBCT-based sCT images for ART in patients with HNC.

Keywords: Cone-beam computed tomography, Synthetic CT, Deep learning, Supervised U-Net, Adaptive radiation therapy, Head and neck cancer

Introduction

Head and neck cancer (HNC) has emerged as a prominent global health issue, ranking as the sixth most prevalent cancer globally. According to estimates from the Global Cancer Observatory, the overall incidence of HNC is on an ongoing upward trajectory, with predictions indicating an increase of 30% by the end of 2030 [1,2]. Due to the presence of multiple critical organs at risk (OARs) close to the treatment volume, radiation therapy commonly utilizes techniques such as intensity-modulated radiation therapy (IMRT) and volumetric-modulated arc therapy (VMAT) [3]. These advanced techniques enable precise targeting of the treatment volume while

minimizing radiation exposure to surrounding critical structures. Nevertheless, the complexity of treatment techniques and anatomical changes can introduce uncertainty, leading to discrepancies between the planned and delivered doses. Thus, an innovative approach called adaptive RT (ART) has been proposed to address this challenge. ART involves adjusting the treatment plan to accommodate anatomical variations throughout the treatment course [4,5]. This is accomplished by assessing the patient's anatomical changes and re-calculating the dose distribution based on the updated images. Typically, these images can be obtained either through a repeat computed tomography (CT) scan or by utilizing images captured on the same day, such as kilovoltage cone-beam CT (kV-CBCT) [6]. Implementing the repeat CT scan technique would lead to an increase in the patient's radiation dose and require additional time and labor. Thus, utilizing kV-CBCT images for patient position alignment within the treatment room can effectively address these limitations. However, the direct application of kV-CBCT images for re-planning is constrained due to various factors, including inconsistent CT numbers, image quality issues related to scattered artifacts and noise, as well as limitations in the field-of-view [7-9]. Therefore, prior to utilizing kV-CBCT in ART, it is necessary to perform kV-CBCT correction.

In the existing literature, several techniques have been proposed for performing kV-CBCT dose calculation [8-10]. These methods encompass (1) the establishment of a calibration curve that correlates Hounsfield units (HU) with electron densities; this curve can be defined using either an adapted phantom or patient kV-CBCT images. However, it is essential to note that these methods are susceptible to image artifacts and patient scatter, which may impact their reliability [9]. (2) the utilization of density assignment methods, specifically bulk density override; this method involves segmenting an image into different tissue classes, such as soft tissues, air, and bones, and assigning appropriate densities to each class. When kV-CBCT images contain significant artifacts, considering alternative approaches can be a favorable option [11]. However, it relies on accurate structure segmentation and may result in an image with homogeneous tissues [9]. (3) the application of deformable image registration; the planning CT (pCT) images are deformed to kV-CBCT images, mainly including the deformation image registration and histogram matching methods. While this approach can effectively correct HU information in kV-CBCT images, it demands high accuracy in the image registration algorithm and matching method, especially in cases with substantial anatomical variations, such as tumor shrinkage or weight loss [12]. Additionally, this approach can be challenging due to inherent limitations in kV-CBCT imaging, including noise, low contrast, and a reduced field-of-view [13]. Finally, (4) the adoption of artificial intelligence

algorithms to generate a synthetic CT (sCT); machine learning algorithms, specifically deep learning techniques, can be utilized to learn the mapping between kV-CBCT images and pCT images. These algorithms can learn intricate relationships and patterns through training on a vast dataset of paired kV-CBCT and pCT scans, enabling them to generate sCT images from kV-CBCT inputs [14]. Several studies have successfully demonstrated significant enhancements in CT number accuracy, image quality, and the mitigation of artifacts, thereby increasing the potential for the clinical implementation of ART [14-19]. Typically, there are two approaches for employing deep learning methods in sCT images. The first approach utilizes supervised training with paired images, often implementing algorithms like U-Net [20,21]. Conversely, the second approach involves unsupervised training with unpaired images, leveraging techniques such as generative adversarial networks (GANs) [14,16,22]. In order to integrate into clinical practice seamlessly, our research primarily concentrates on utilizing the supervised U-Net algorithm. This choice is motivated by its straightforward implementation, stable convergence, and rapid training process. Therefore, this study aims to generate and assess the performance of a U-Net deep learning-based algorithm in converting kV-CBCT scans to sCT images in terms of HU accuracy, image quality, and dosimetric accuracy similarity in the head and neck region.

Materials and Methods

1. Patient selection and image dataset

In this study, we conducted a retrospective analysis using a total of 3,491 image pairs. The images consisted of paired kV-CBCT scans and pCT images obtained from 40 datasets of patients with HNC who underwent VMAT between January 2018 and December 2021. This study received approval from the Ethics Committee of Chulabhorn Royal Academy (No. EC 010/2565). The data used for this study were collected from the Department of Radiation Oncology at Chulabhorn Hospital. For image acquisition, a dedicated 16-slice helical Big-Bore CT simulator (Phillips Medical Systems, Andover, MA, USA) was utilized to obtain the pCT image datasets. The pCT images were obtained using a tube voltage of 120 kVp and an exposure range of 300–400 mAs. The kV-CBCT image datasets were obtained using onboard imaging functionality available on the TrueBeam linear accelerator (Varian Medical System Inc., Palo Alto, CA, USA). These kV-CBCT images were acquired with a tube voltage of 100 kVp and an exposure of 150 mAs. In order to minimize the impact of anatomical variations, only the kV-CBCT images captured during the first fraction before treatment were included in the study. The pCT images were characterized by a voxel spacing of 1.00 mm × 1.00 mm × 3.00 mm, whereas the kV-CBCT images

had a voxel spacing of 0.51 mm × 0.51 mm × 2.00 mm. In terms of image dimensions, both the pCT and kV-CBCT images were sized at 512 × 512 pixels.

The patient group included in the study received prescribed doses of 66–70 Gy for the high-risk planning target volume (PTV_HR), 59.4 Gy for the intermediate-risk PTV (PTV_IR), and 54.0 Gy for the low-risk PTV (PTV_LR). The treatment was delivered in 33 fractions. The dose prescriptions for the target volume were reported based on the minimum dose required to cover 95% of the target volume (D_{95}), as determined from the dose-volume histogram (DVH). The Eclipse Treatment Planning System version 16.1 (Varian Medical System Inc.) was employed to optimize the treatment plans. The plans were optimized using a simultaneous integrated boost-VMAT delivery technique on the pCT images. For the OARs, which aligns with the guidelines outlined in the RTOG 0225 Report: the maximum dose (D_{max}) to the brain stem was kept below 54.0 Gy. Both the left and right parotid glands had 50% of their volume receiving a dose below 30 Gy. The D_{max} to the spinal cord was limited to less than 45.0 Gy.

2. Image preparation

To align the pCT images with the kV-CBCT images, a rigid registration process was conducted using an Open-source Registration Graphical User Interface (OpenREGGUI), a MATLAB-based medical image processing software. During the registration process, the pixel dimensions of the pCT images were resampled to match the voxel spacing of the kV-CBCT images, which was set at 0.51 mm × 0.51 mm × 2.00 mm. Due to the incomplete field-of-view of kV-CBCT images in HNC, kV-CBCT images with incomplete body outlines were excluded from the image datasets. Additionally, to ensure consistency, a structure representing the area outside the body was generated and assigned an air density value of HU for both the pCT and kV-CBCT images.

3. Model architecture

The U-Net served as the basis for the proposed model in this study. The network model was developed using Keras and TensorFlow 2.9.0, utilizing Python version 3.8. The implementation involved NVIDIA CUDA Deep Neural Network Library version 8.1 and Compute Unified Device Architecture version 11.2. All experiments were performed on an NVIDIA Quadro RTX 8000 GPU with 48 GB of memory. The training process took place within a JetBrains PyCharm anaconda environment. The U-Net model, as illustrated in Fig. 1, follows a convolutional encoder–decoder network structure. The model utilizes pairs of kV-CBCT and pCT images as input and generates sCT images as output. The network architecture can be divided into two main phases: the encoding phase and the decoding phase.

The model encodes by downscaling using 2D convolutions, rectified linear unit (ReLU) activation function, and maxpooling in six blocks: initial features of 32, increasing by 2; maxpooling of 2 × 2. Encoder reduces the image to 4 × 4 × 1024 with final 3 × 3 convolutions, 2048 features. Decoding uses transposed convolutions. Each up-sample block has a transposed convolution, concatenation, and two 3 × 3 convolutions; starting at 2048, features decrease by 2; decoder convolutions of 3 × 3, 32 features. The U-Net model is optimized using mean absolute error (MAE) loss.

4. Model training

The dataset was split into training and testing sets using an 80% and 20% division. This resulted in 2,976 images from 32 patients being used for training, while the remaining 515 images from eight patients were assigned for testing. Within the training dataset, 20% of the data was further allocated for validation purposes. During the training process, various hyperparameters were employed, encompassing a 1×10^{-3} learning rate, batch size of 8, and 200 epochs. These hyperparameters were carefully selected to enhance the training procedure and attain optimal model performance.

5. Model testing

Following the completion of model training, the model was tested on eight independent patient datasets. The performance of the predictive sCT images was assessed based on criteria such as HU accuracy, image quality, and dosimetric accuracy. A comparative analysis was conducted between the pCT and sCT images to assess these aspects, employing the following evaluations.

1) HU accuracy

The image intensity was assessed by calculating the differences in HU values between the pCT and sCT images. MAE and root mean square error (RMSE) were employed as metrics in this study, and they are calculated using the following equations:

$$MAE(pCT, sCT) = \frac{1}{n} \sum_{i=1}^{n} |pCT_i - sCT_i|, \tag{1}$$

$$RMSE(pCT, sCT) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} |pCT_i - sCT_i|^2}, \tag{2}$$

where pCT and sCT represent the corresponding pixel values in the pCT and sCT images, and n is the total number of pixels in the images. These metrics were utilized to quantify the differences between the pCT and sCT images and assess the accuracy of the predictive sCT images.
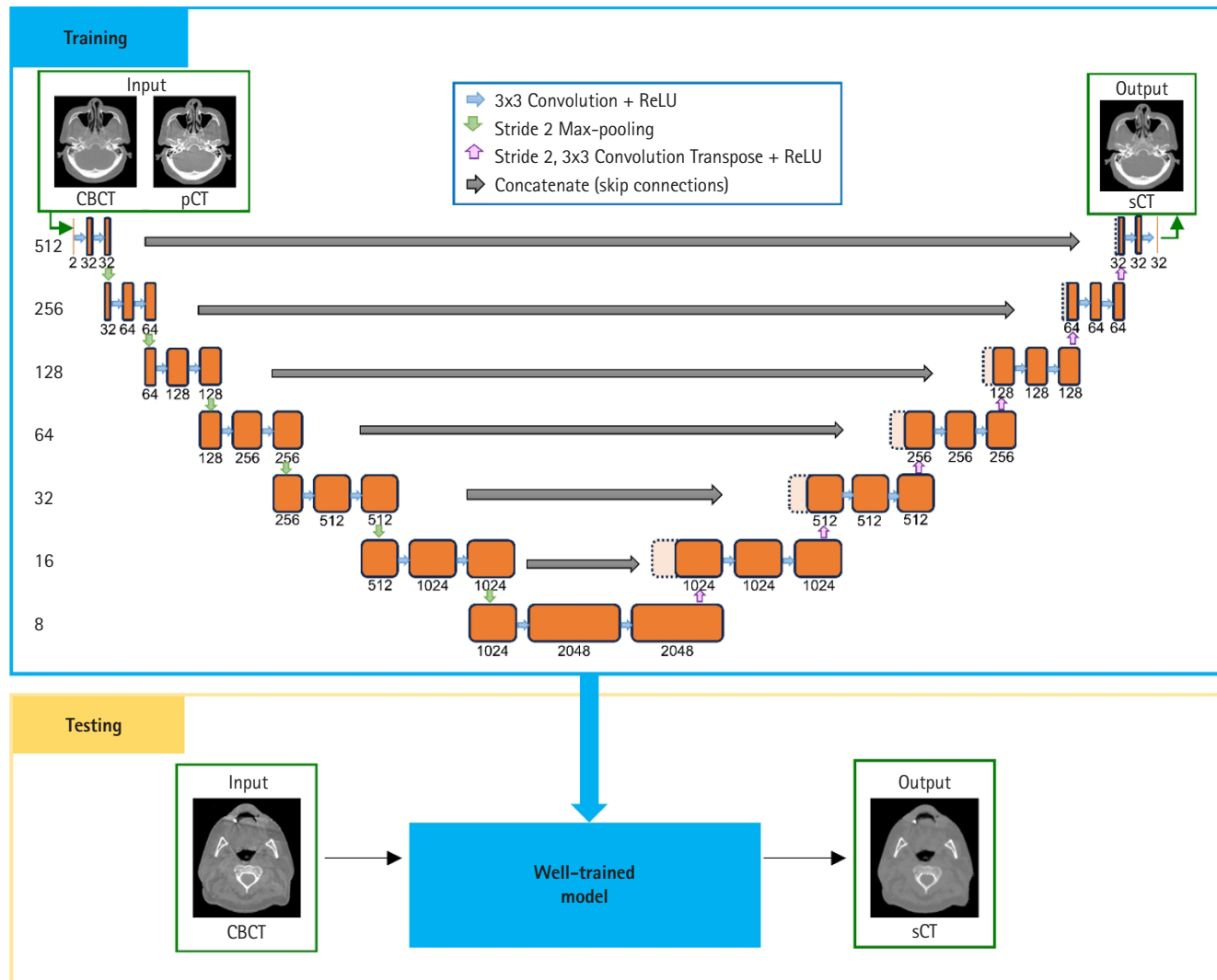
Fig. 1. U-Net network architecture: training (upper) and testing (lower). pCT, planning computed tomography; sCT, synthetic computed tomography; CBCT, cone-beam computed tomography; ReLU, rectified linear unit.

2) Image quality

The image quality of the generated sCT images was assessed using two commonly used metrics: peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM). These metrics serve as indicators of image similarity and overall quality. PSNR measures the ratio between the maximum possible signal power and the power of the noise present in the image. A higher PSNR value signifies a more remarkable resemblance between the sCT and reference pCT images. SSIM, on the other hand, gauges the structural similarity between the sCT and reference pCT images, considering factors such as luminance, contrast, and structural information. A higher SSIM value indicates a more significant similarity between the images. Both PSNR and SSIM metrics were employed in this study to provide quantitative evaluations of the image quality exhibited by the generated sCT images. The calculation of each term can be represented as follows:

$$PSNR(pCT, sCT) = 20 \cdot log_{10}\left(\frac{MAX(pCT)}{RMSE(pCT, sCT)}\right), \quad (3)$$

$$SSIM(pCT, sCT) = \frac{(2\mu_{sCT}\mu_{pCT}+C_1)+(2\sigma_{sCT,pCT}+C_2)}{(\mu_{sCT}^2+\mu_{pCT}^2+C_1)+(\sigma_{sCT}^2+\sigma_{pCT}^2+C_2)}, \quad (4)$$

where

MAX: Maximum intensity value for both the pCT and sCT images.

μsCT: Mean value of HU (Hounsfield Units) for the sCT image.

μpCT: Mean value of HU for the pCT image.

σsCT: Variance of HU values for the sCT image.

σpCT: Variance of HU values for the pCT image.

σsCTpCT: Covariance between HU values of the sCT and pCT images.

Additionally, the parameters C_1 and C_2 are used to stabilize the division when the denominators are weak. They are calculated as follows: $C_1 = (k_1 {}^* L)^2$, where $k_1 = 0.01$ and L represents the range of HU values in the CT image. $C_2 = (k_2 {}^* L)^2$, where $k_1 = 0.02$.

3) Dosimetric accuracy

All independent testing datasets of sCT images were imported into the Eclipse Treatment Planning System. Due to the field-of-view limitation, the pCT images were regenerated with the same volume as the sCT images. Image registration was then performed between the pCT and sCT images. Subsequently, all structures, including PTV_HR, PTV_IR, and OARs, were transferred from the original pCT to the sCT, excluding PTV_LR. This exclusion was necessary because the field-of-view of the kV-CBCT images did not encompass the entire tumor volume as in PTV_LR. Following the structure transfer, the dose was re-calculated. The original treatment plan was copied and re-calculated using the anisotropic analytical algorithm algorithm with preset monitor unit values from the original plan on both the sCT and matched pCT images. The passing rates of 3D gamma analysis for the dose distributions on sCT images were computed and compared against pCT images under different criteria (3%/3 mm, 3%/2 mm, and 2%/2 mm) with 10% dose threshold. Furthermor, The dose difference between the pCT and sCT images was assessed in terms of statistical evaluation by measuring metrics such as $D_{95\%}$ (the dose received by 95% of the volume) to evaluate target volume coverage, $D_{2\%}$ to assess hotspots within the target volume, as well as $D_{50\%}$ or D_{max} for the OARs.

Results

1. HU accuracy and image quality

Table 1 presents the MAE, RMSE, PSNR, and SSIM values of kV-CBCT and sCT images when compared to the reference pCT images. When comparing the metrics of the sCT to kV-CBCT images, there was an improvement in the values of MAE, RMSE, and PSNR. Specifically,

the values decreased from 53.15 ± 40.09 to 41.47 ± 30.59 for MAE, from 153.99 ± 79.78 to 130.39 ± 78.06 for RMSE, and increased from 47.91 ± 4.98 dB to 49.93 ± 6.00 dB for PSNR, respectively. The SSIM of sCT images was 0.98 ± 0.02, which was higher compared to the SSIM of kV-CBCT images, which was 0.97 ± 0.02.

Fig. 2 showcases three axial slices of pCT, kV-CBCT, and sCT images. The sCT images demonstrate enhanced HU values that closely resemble those of pCT images while maintaining the geometric information present in kV-CBCT images.

Furthermore, the sCT images effectively reduce streak artifacts observed in kV-CBCT images, as indicated by the red arrow in Fig. 2.

Fig. 3 illustrates the HU line profile across the body of pCT (orange), kV-CBCT (green), and sCT (blue) images. The line profiles indicate that the HU line of sCT images closely resembles that of pCT images, particularly at the boundaries of the body. sCT images exhibit reduced artifacts and smoother transitions compared to kV-CBCT images. Notably, at the interface between soft tissue and air, sCT images display a smoother edge compared to kV-CBCT images.

2. Dosimetric accuracy

Table 2 presents the 3D gamma passing rates concerning the dose distribution within sCT images, assessed using independent testing data. The evaluation was conducted utilizing gamma criteria of 2%/2 mm, 3%/2 mm, and 3%/3 mm. The outcomes reveal passing rates of 92.1% ± 3.8%, 93.8% ± 3.0%, and 96.9% ± 2.0%, respectively. These results highlight the remarkable performance of the 3D gamma passing rates in capturing the dose distribution within sCT images. The observed passing rate consistently maintains alignment with acceptable levels, even when subjected to slightly more stringent evaluation criteria.

In Fig. 4, box plots are depicted to showcase the percentage dose difference of target volume and OARs when comparing pCT and sCT images from eight independent patient datasets. Overall, the dose metric outcomes from the sCT-based plan closely resemble those obtained from the pCT-based plan. The percentage difference in dose for all structures experienced slight changes. The highest values fell within a 3% range, except for the parotid glands, where values exceeded 3% but remained below 5.2%. The observed deviation, which includes the presence of an anatomical change and the utilization of a bolus, can be attributed to patient #6, as depicted in Fig. 5.

Fig. 5 presents the calculated dose distribution using pCT and sCT. In patient #7, the dose distribution exhibits a relatively similar pattern between the two image sets. However, in patient #6, the dose distribution differs significantly due to anatomical changes.

Fig. 6 showcases the DVH of structures comparing the pCT and sCT-based plans for patients #6 and #7. In patient #7, a slight dose
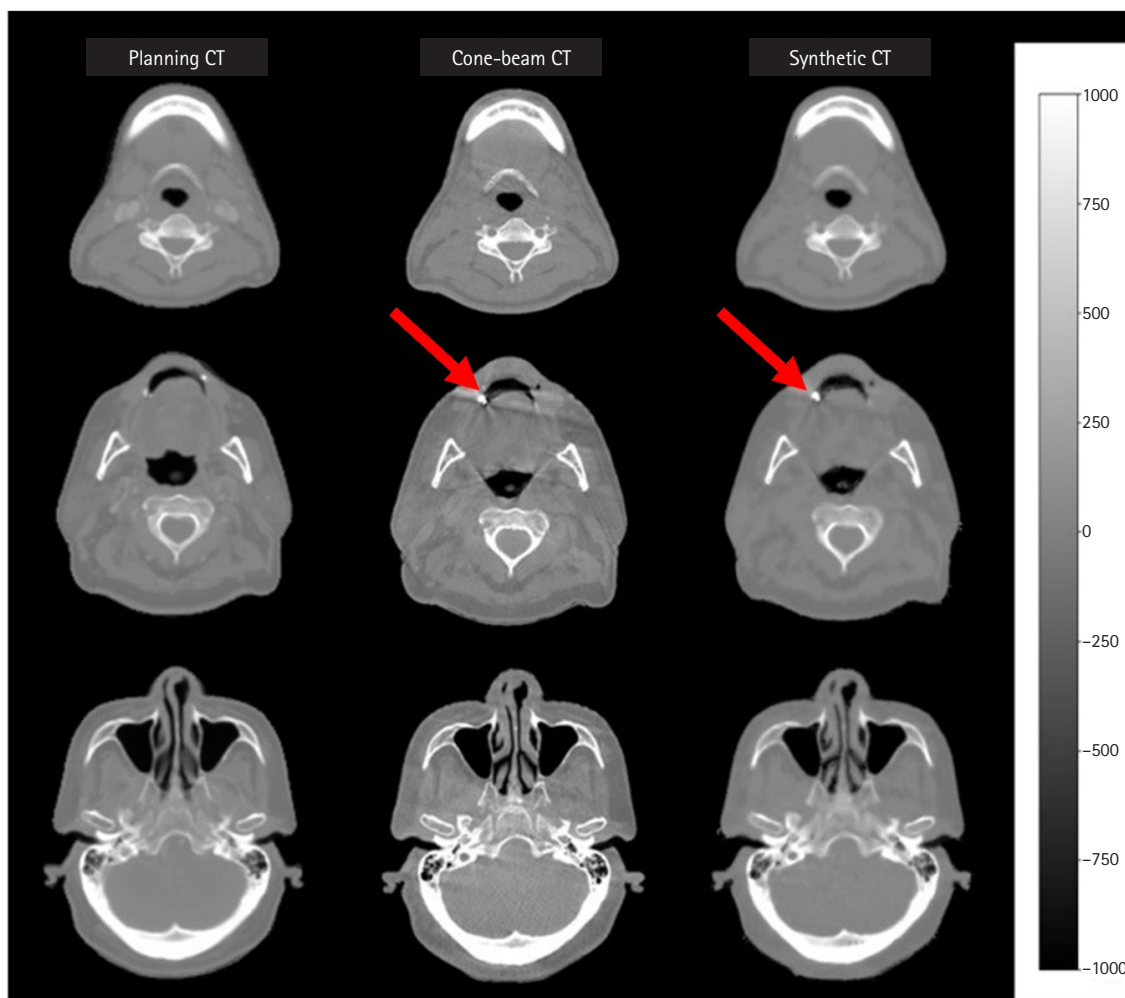
Table 1. MAE, RMSE, PSNR, and SSIM values of kV-CBCT and sCT images

Image	MAE	RMSE	PSNR (dB)	SSIM
kV-CBCT	53.15 ± 40.09	153.99 ± 79.78	47.91 ± 4.98	0.97 ± 0.02
sCT	41.47 ± 30.59	130.39 ± 78.06	49.93 ± 6.00	0.98 ± 0.02

Values are presented as mean ± standard deviation.
kV-CBCT, kilovoltage cone-beam computed tomography; sCT, synthetic computed tomography; MAE, mean absolute error; RMSE, root mean square error; SSIM, structural similarity index; PSNR, peak signal-to-noise ratio.

Fig. 2. Comparison of three axial slices from planning CT, kV-CBCT, and synthetic CT images. In the synthetic CT images, indicated by the red arrow, a reduction in streak artifact is observed compared to the kV-CBCT images. CT, computed tomography; kV-CBCT, kilovoltage cone-beam computed tomography.

difference was observed in the DVH. However, the dose difference between the pCT and sCT-based plans was more pronounced in patient #6. These findings highlight the impact of anatomical changes and the limitations of using kV-CBCT images for generating sCT, particularly in capturing accurate target volumes.

Discussion and Conclusion

This study aimed to investigate the model for generating sCT images from kV-CBCT scans using the supervised U-Net deep learning algorithm for ART treatment planning in the HNC region. This region is particularly susceptible to anatomical variations, such as weight and tumor size changes. The study focused on assessing the accuracy of HU, image quality, and dosimetric accuracy. Four evaluation metrics were employed to determine the performance of the

generated sCT images. HU accuracy was evaluated using the MAE and RMSE, while image quality was assessed using PSNR and SSIM. In addition, the dosimetric accuracy was evaluated by measuring metrics such as $D_{95\%}$ and $D_{2\%}$ to assess the target volume and $D_{50\%}$ or D_{max} for the OARs.

Our study's findings are consistent with the results reported by Kida et al. [19], Li et al. [23], and Chen et al. [20]. In line with these studies, our results likewise indicated a reduction in MAE and RMSE values, accompanied by an enhancement in PSNR and SSIM values when compared to the original kV-CBCT images. Kida et al. [19] utilized a U-Net model structure similar to ours, excluding ReLU in transpose convolution, and employed MAE as the loss function in the pelvic regions. Our study demonstrates improved image quality regarding PSNR and SSIM, with values of 50.9 dB and 0.967, respectively. Comparing our study to the work of Li et al.
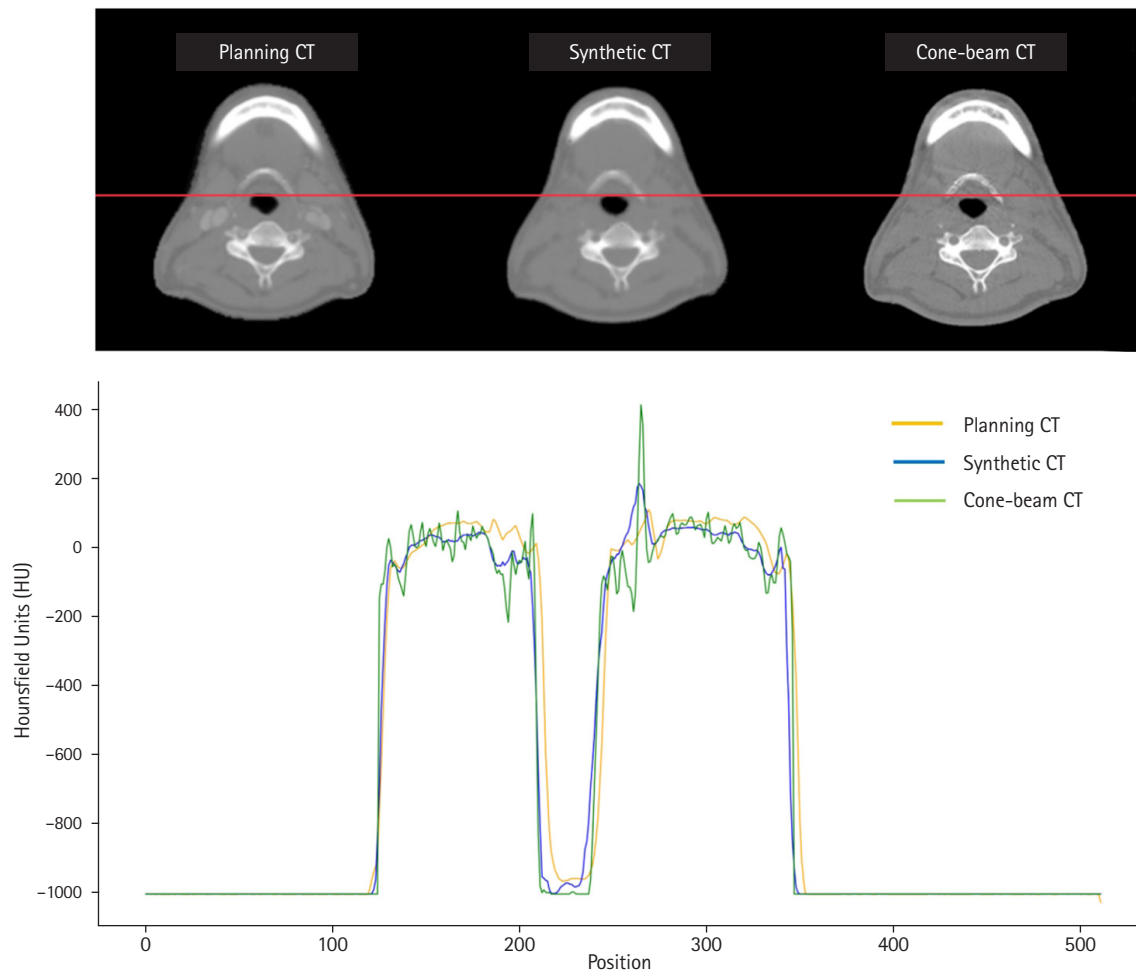
Fig. 3. Axial images (upper) and Hounsfield unit line profiles (lower) of planning CT images (in orange), kV-CBCT images (in green), and synthetic CT images (in blue) acquired from independent datasets. CT, computed tomography; kV-CBCT, kilovoltage cone-beam computed tomography.

Table 2. The 3D gamma passing rates (%) of dose distribution in sCT images for independent testing data

Gamma critera		
2%/2 mm	3%/2 mm	3%/3 mm
92.1 ± 3.8	93.8 ± 3.0	96.9 ± 2.0

[23], a 2D U-Net neural network in HNC was investigated. Our study surpasses their MAE, PSNR, and SSIM performance. However, RMSE was not considered in their work. They reported MAE, PSNR, and SSIM values of 56.89 ± 13.84, 28.80 ± 2.46 dB, and 0.71 ± 0.032, respectively. However, it is essential to note that Chen et al. [20] achieved better HU accuracy metrics results than our study by employing a loss function that combined structure dissimilarity and MAE. They reported an average MAE, RMSE, PSNR, and SSIM of 18.98, 60.16, 33.26 dB, and 0.8911, respectively. Among the observed results, the most notable finding is that our study achieved

the highest SSIM value. This indicates higher similarity and quality between our generated sCT images and the reference CT images (pCT). The results depicted in Figs. 2 and 3 demonstrated that the sCT images achieved better image quality. Notably, the presence of scatter artifacts was effectively eliminated, leading to images free from disturbances. Furthermore, the smoothness of the synthesized images was found to be comparable to that of pCT images.

According to the HU line profiles of pCT, kV-CBCT, and sCT images, the sCT images not only successfully eliminated scatter artifacts but also demonstrated a higher degree of similarity in HU values to the pCT images compared to the kV-CBCT images, especially at the body boundaries. At the thyroid cartilage region, all sCT images showed better intensity with the reduction from the peak of kV-CBCT images. The smoothness of soft tissue had improved in all sCT images. Although when the HU line crosses through the surrounding edge of the air region, such as the outside of the body and subglottic larynx, it indicates that sCT images had rounded corners
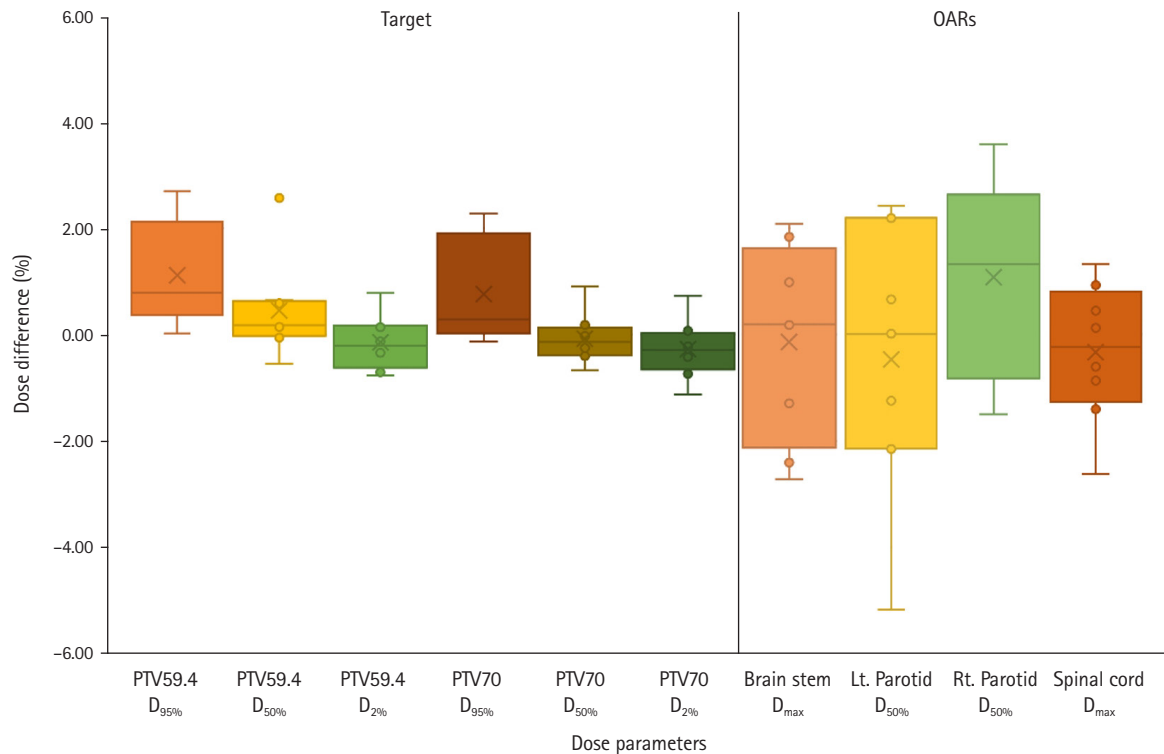
Fig. 4. The percentage dose difference between the planning CT and synthetic CT-based plans for the target and organs at risk (OARs). CT, computed tomography; PTV, planning target volume; Lt., left; Rt., right.
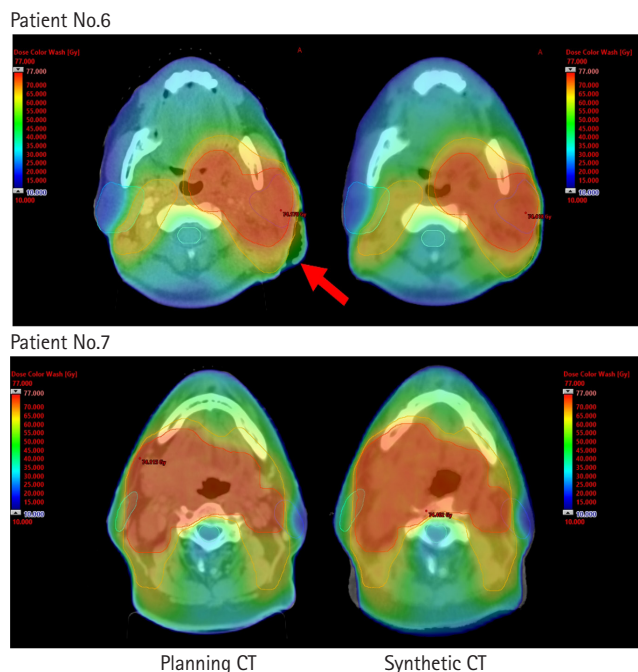

Fig. 5. Comparison of dose distribution between the planning CT and synthetic CT-based plans for patients #6 (left) and #7 (right). The PTV 70 Gy is visually emphasized in red, while the PTV 59.4 Gy is represented in orange. The left and right parotid glands are depicted in blue and cyan, respectively. CT, computed tomography; PTV, planning target volume.

close to pCT images.

In our analysis of dosimetry, we found that our study had a lower gamma passing rate of 2%, 2 mm criteria compared to another study. Specifically, our results indicated this rate to be 92.1% ± 3.8%. Jihong et al. [24] reported a rate of 95.7% ± 1.9% for sCT images with uncorrected CBCT and 97.1% ± 1.9% for sCT images with corrected CBCT. Meanwhile, Yoo et al. [25] achieved an impressive result of 99.7% ± 0.0% for sCT using a combination loss functions for model training. Both studies utilized advanced deep learning techniques, with Jihong et al. [24] employing unsupervised learning via CycleGAN with HU correction, and Yoo et al. [25] enhancing performance through the integration of perceptual loss into L1 and structural similarity loss functions during model training. These findings suggest that unsupervised deep learning and specialized loss functions can enhance the quality of sCT images, and preprocessing techniques such as HU correction can further improve outcomes.

It is important to note that the presence of registration errors in the training data can lead to inaccuracies in the model's performance as it may be trained to generate erroneous image predictions. In addition, the anatomical changes between the pCT and sCT images may influence the results of the dose evaluation. Further analysis revealed that the outlier data points were primarily associated with
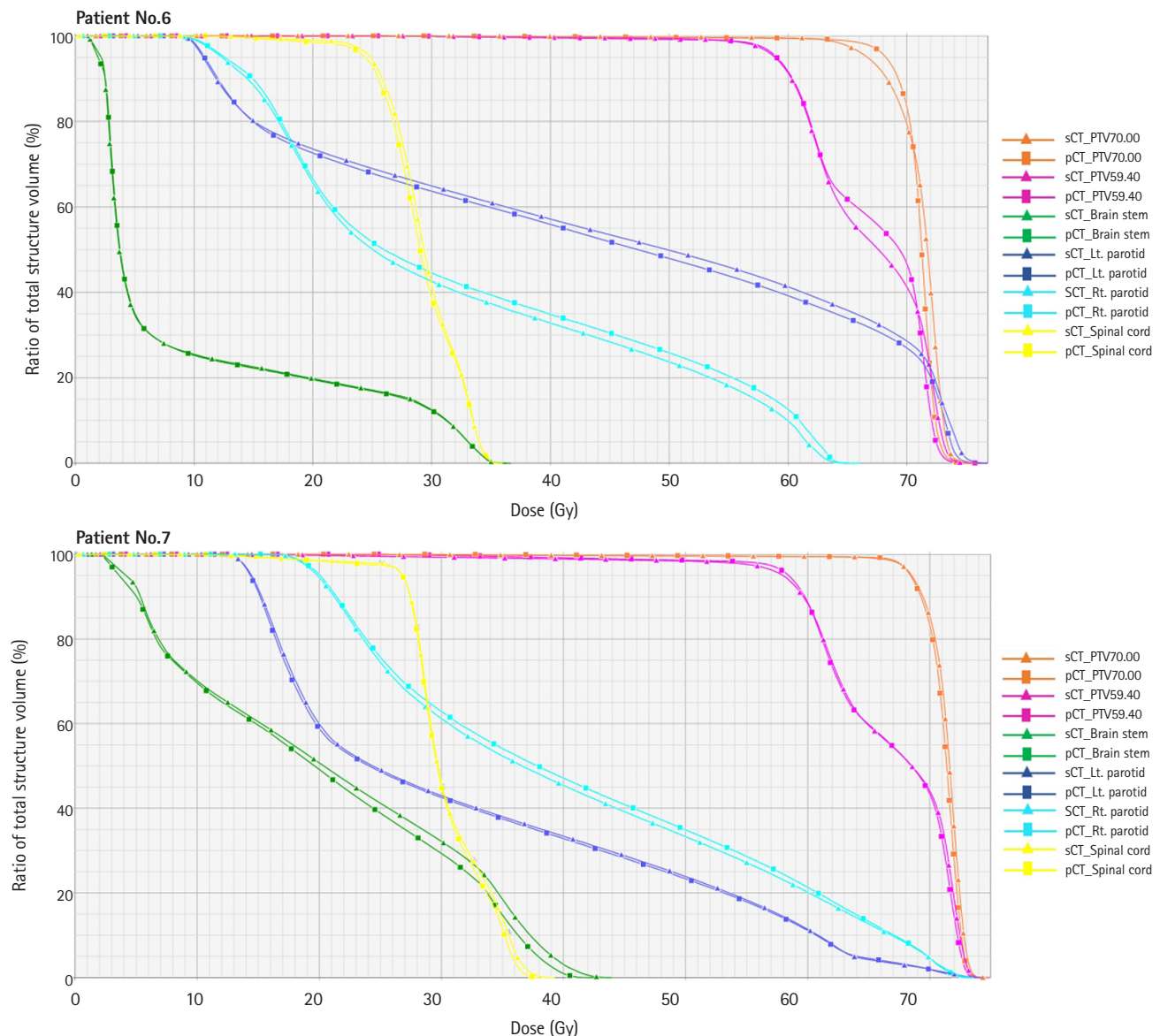
Fig. 6. The comparison of dose-volume histogram (DVH) between the pCT and sCT-based plans for patients #6 (upper) and #7 (lower). Squared lines represent the DVH of the pCT-based plan, while the triangle lines represent the DVH of the sCT-based plan. pCT, planning computed tomography; sCT, synthetic computed tomography; PTV, planning target volume; Lt., left; Rt., right.

one specific patient, precisely patient #6. This particular patient exhibited significant anatomical changes. The parotid gland of patient #6 showed higher dose difference values compared to the other patients. The observed dose uncertainties can be attributed to the fact that the anatomical changes in the target volume of patients.

The U-Net model presents numerous benefits, such as reducing global scattering and enhancing local HU accuracy by integrating global and local features within the image's spatial domain [20]. Moreover, this model offers advantages such as straightforward implementation, stable convergence, and the fastest training [26]. Nevertheless, a limitation of the model lies in determining the op-

timal number of depths in the encoder–decoder network, which depends on the complexity of the training task. Additionally, the design of skip connections between the encoder and decoder networks lacks a robust theoretical framework.

While our study demonstrated improvements in HU accuracy, image quality, and dosimetric accuracy, several limitations should be acknowledged. Firstly, the limited field-of-view of kV-CBCT images restricted the availability of anatomical information, particularly for low-risk target volumes in the shoulder region of HNC. This limitation could affect the accuracy of the generated sCT images in those areas. Secondly, our study had a relatively small image data sample

size for training and testing. A more extensive and diverse dataset could yield even better results and provide a more robust evaluation of the model's performance accuracy. Additionally, a strategic approach to dataset enhancement through image augmentation techniques involving translation, inversion, and slight rotation presents a viable avenue for further improvement. Thirdly, anatomical changes may occur between the pCT and kV-CBCT images due to the time gap between image acquisitions (usually taken on separate days). These anatomical variations could potentially impact the training of the U-Net algorithm, which relies on paired image data. Therefore, exploring unsupervised models such as GANs could enhance the model's performance, particularly when dealing with unpaired image data (intra-individual co-registration) [18,24-30]. Furthermore, the process of rigidly registering the CT and CBCT images might not have been adequate in establishing the required image similarity for network training. This concern could be mitigated by incorporating deformable image registration, which could have been essential for enhancing image similarity to a greater extent. In addition, the model developed in this study is based on a 2D approach, generating sCT images slice-by-slice. To further enhance the performance of the 2D model, prior research [20] has proposed the utilization of a volumetric neural network operating in three dimensions. This could involve extracting image features more precisely and effectively. Despite these limitations, the 2D supervised U-Net model remains a valuable tool for improving the accuracy and quality of the sCT generation from kV-CBCT images, effectively addressing the challenges posed by anatomical variations in HNC patients.

In conclusion, this study used supervised U-Net deep learning to generate sCT images from kV-CBCT scans in patients with HNC. Our evaluation focused on assessing the HU accuracy, image quality, and dosimetric accuracy of the sCT images. The sCT images produced by the U-Net model successfully reduced artifacts and noise while preserving the anatomical structure obtained from the kV-CBCT images. The HU accuracy and image quality of the sCT images were close to those of the pCT images. Furthermore, regarding dosimetric accuracy, the structures in the sCT images, including PTV, and OARs, closely resembled their counterparts in the pCT images. However, it is essential to acknowledge that anatomical changes can occur between acquiring pCT and kV-CBCT images. Our findings suggest that the U-Net model holds significant potential in generating kV-CBCT-based sCT images for ART in HNC patients. This approach offers a convenient method of obtaining updated anatomical information without subjecting the patient to additional radiation doses.

Statement of Ethics

The study received approval from the Institutional Review Board of

Chulabhorn Royal Academy (No. EC 010/2565).

Conflict of Interest

No potential conflicts of interest relevant to this article were disclosed.

Author Contributions

Conceptualization, CK, TP; Project administration, CK; Supervision, CK, ST; Investigation and methodology, TP, CK, ST, AD; Programming, TP; Writing of manuscript, CK, TP; Data analysis, TP, CK, AD, TC, TM, PS; Visualization, TP, CK, AD, TC, TM, PS; Gamma index analyse: KN.

Data Availability Statement

The supporting data for the findings of this study can be obtained from the corresponding author upon reasonable request.

References

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71:209–49.
2. Gormley M, Creaney G, Schache A, Ingarfield K, Conway DI. Reviewing the epidemiology of head and neck cancer: definitions, trends and risk factors. Br Dent J 2022;233:780–6.
3. Buciuman N, Marcu LG. Adaptive radiotherapy in head and neck cancer using volumetric modulated arc therapy. J Pers Med 2022;12:668.
4. Glide-Hurst CK, Lee P, Yock AD, et al. Adaptive radiation therapy (ART) strategies and technical considerations: a state of the ART review from NRG oncology. Int J Radiat Oncol Biol Phys 2021; 109:1054–75.
5. Badey A, Barateau A, Delaby N, et al. [Overview of adaptive radiotherapy in 2019: From implementation to clinical use]. Cancer

Radiother 2019;23:581–91.

6. Talbot A, Devos L, Dubus F, Vermandel M. Multimodal imaging in radiotherapy: focus on adaptive therapy and quality control. Cancer Radiother 2020;24:411–7.

7. Posiewnik M, Piotrowski T. A review of cone-beam CT applications for adaptive radiotherapy of prostate cancer. Phys Med 2019;59:13–21.

8. Delaby N, Barateau A, Chiavassa S, et al. Practical and technical key challenges in head and neck adaptive radiotherapy: the GORTEC point of view. Phys Med 2023;109:102568.

9. Barateau A, De Crevoisier R, Largent A, et al. Comparison of CBCT-based dose calculation methods in head and neck cancer radiotherapy: from Hounsfield unit to density calibration curve to deep learning. Med Phys 2020;47:4683–93.

10. Giacometti V, Hounsell AR, McGarry CK. A review of dose calculation approaches with cone beam CT in photon and proton therapy. Phys Med 2020;76:243–76.

11. Fotina I, Hopfgartner J, Stock M, Steininger T, Lutgendorf-Caucig C, Georg D. Feasibility of CBCT-based dose calculation: comparative analysis of HU adjustment techniques. Radiother Oncol 2012;104:249–56.

12. Kurz C, Kamp F, Park YK, et al. Investigating deformable image registration and scatter correction for CBCT-based dose calculation in adaptive IMPT. Med Phys 2016;43:5635.

13. Sun H, Fan R, Li C, et al. Imaging study of pseudo-CT synthesized from cone-beam CT based on 3D CycleGAN in radiotherapy. Front Oncol 2021;11:603844.

14. Liang X, Chen L, Nguyen D, et al. Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy. Phys Med Biol 2019;64:125002.

15. Chen L, Liang X, Shen C, Nguyen D, Jiang S, Wang J. Synthetic CT generation from CBCT images via unsupervised deep learning. Phys Med Biol 2021;66.

16. Liu J, Yan H, Cheng H, et al. CBCT-based synthetic CT generation using generative adversarial networks with disentangled representation. Quant Imaging Med Surg 2021;11:4820–34.

17. Maspero M, Houweling AC, Savenije MH, et al. A single neural network for cone-beam computed tomography-based radiotherapy of head-and-neck, lung and breast cancer. Phys Imaging Radiat Oncol 2020;14:24–31.

18. Eckl M, Hoppen L, Sarria GR, et al. Evaluation of a cycle-generative adversarial network-based cone-beam CT to synthetic CT conversion algorithm for adaptive radiation therapy. Phys Med 2020;80:308–16.

19. Kida S, Nakamoto T, Nakano M, et al. Cone beam computed tomography image quality improvement using a deep convolutional neural network. Cureus 2018;10:e2548.

20. Chen L, Liang X, Shen C, Jiang S, Wang J. Synthetic CT generation from CBCT images via deep learning. Med Phys 2020;47:1115–25.

21. Xue X, Ding Y, Shi J, et al. Cone beam CT (CBCT) based synthetic CT generation using deep learning methods for dose calculation of nasopharyngeal carcinoma radiotherapy. Technol Cancer Res Treat 2021;20:15330338211062415.

22. Zhang Y, Yue N, Su MY, et al. Improving CBCT quality to CT level using deep learning with generative adversarial network. Med Phys 2021;48:2816–26.

23. Li Y, Zhu J, Liu Z, et al. A preliminary study of using a deep convolution neural network to generate synthesized CT images based on CBCT for adaptive radiotherapy of nasopharyngeal carcinoma. Phys Med Biol 2019;64:145010.

24. Jihong C, Kerun Q, Kaiqiang C, Xiuchun Z, Yimin Z, Penggang B. CBCT-based synthetic CT generated using CycleGAN with HU correction for adaptive radiotherapy of nasopharyngeal carcinoma. Sci Rep 2023;13:6624.

25. Yoo SK, Kim H, Choi BS, Park I, Kim JS. Generation and evaluation of synthetic computed tomography (CT) from cone-beam CT (CBCT) by incorporating feature-driven loss into intensity-based loss functions in deep convolutional neural network. Cancers (Basel) 2022;14:4534.

26. Rusanov B, Hassan GM, Reynolds M, et al. Deep learning methods for enhancing cone-beam CT image quality toward adaptive radiation therapy: a systematic review. Med Phys 2022;49:6019–54.

27. Liu Y, Lei Y, Wang T, et al. CBCT-based synthetic CT generation using deep-attention cycleGAN for pancreatic adaptive radiotherapy. Med Phys 2020;47:2472–83.

28. Szmul A, Taylor S, Lim P, et al. Deep learning based synthetic CT from cone beam CT generation for abdominal paediatric radiotherapy. Phys Med Biol 2023;68:105006.

29. Suwanraksa C, Bridhikitti J, Liamsuwan T, Chaichulee S. CBCT-to-CT translation using registration-based generative adversarial networks in patients with head and neck cancer. Cancers (Basel) 2023;15:2017.

30. Zhang Y, Ding SG, Gong XC, et al. Generating synthesized computed tomography from CBCT using a conditional generative adversarial network for head and neck cancer patients. Technol Cancer Res Treat 2022;21:15330338221085358.