Semi-Supervised Deformation-Free Image-to-Image Translation for Realistic CT Synthesis from CBCT

Ji Yong Han[1,*], Su Yang[2,*], Sujeong Kim[1], Sunjung Kim[3], Sang-Heon Lim[1], Heejin Yun[1], Dahee Kim[2], Won-Jin Yi[1,2,4]

[1]Interdisciplinary Program in Bioengineering, Graduate School of Engineering, Seoul National University, Republic of Korea
[2]Applied Bioengineering, Graduate School of Convergence Science and Technology, Seoul National University, Republic of Korea
[3]Imaging RD Center, Osstem Implant Co., LTD, Republic of Korea
[4]Oral and Maxillofacial Radiology and Dental Research Institute, School of Dentistry, Seoul National University, Republic of Korea
wjyi@snu.ac.kr

Abstract. Cone-Beam Computed Tomography (CBCT) is widely used for diagnostics and treatment planning in oral and maxillofacial field due to its low radiation dose and high spatial resolution. Still, its clinical utility is limited by low contrast and incorrect Hounsfield Unit (HU) values. In contrast, multi-detector CT (CT) provides high contrast and reliable HU measurements, with a higher radiation dose. In this work, we present a novel two-stage framework for unpaired CBCT-to-CT synthesis that ensures the exact preservation of anatomical structure, maintains high resolution, and achieves accurate HU value. In the first stage, we generate pseudo-paired CT images. In the second stage, we utilize a UNet++ generator enhanced with Interpolation and Convolution Upsampling (ICUP), Edge-Conditioned Skip Connections (ECSC), and a dual discriminator strategy for a semi-supervised approach. Consequently, we generate realistic CT images using pseudo-paired CT images. Extensive quantitative and qualitative evaluations demonstrate that our method outperforms existing unpaired translation techniques, producing realistic CT images that closely match CT images in both HU accuracy and exactly preserve anatomical structure of the CBCT. The code is available at https://github.com/HANJIYONG/Semi-Supervised-Deformation-Free-I2I.

Keywords: Computed Tomography · Unpaired Image-to-Image Translation · Semi-supervised Learning

1 Introduction

Cone-Beam Computed Tomography (CBCT) is widely used for diagnostics and treatment planning in the oral and maxillofacial field [1]. CBCT has the advantages of lower radiation exposure and higher spatial resolution compared

* These authors contributed equally.

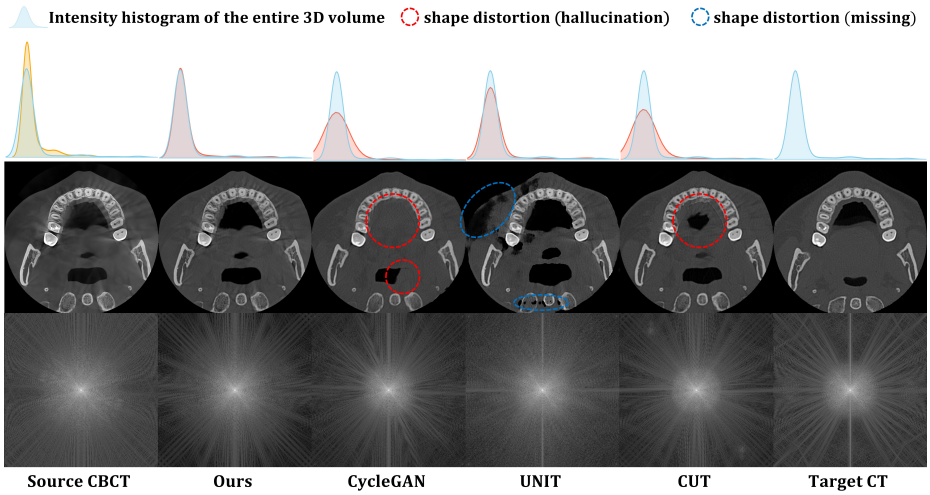


Fig. 1. Illustration of shape inconsistency and resolution degradation in unpaired CBCT-to-CT translation. Our method preserves structural integrity, aligns HU intensities with CT, and maintains fine spatial details. The Fourier-domain magnitude maps visualize spatial frequency characteristics: CBCT shows a broad, circular energy distribution indicating high-resolution content, while the target CT and other methods exhibit more centralized, low-frequency energy patterns. Our method retains a wider radial frequency distribution, highlighting better preservation of anatomical detail.

with multi-detector CT (CT), and it provides three-dimensional information on anatomical structures. However, its reduced contrast resulting from lower radiation dose limits use in diagnostics and bone density measurements [2, 3]. In contrast, CT provides smoother images with enhanced contrast while providing consistent Hounsfield Unit (HU) measurements [4]. Recent advances in AI-based solutions have demonstrated that image-to-image translation (I2I) methods can selectively preserve the benefits of both modalities [5–8]. These approaches enhance image quality with high spatial resolution and uniform HU scale. However, in clinical settings, obtaining paired CBCT and CT data is challenging due to variations in patient posture, breathing, and motion during image acquisition. Unpaired I2I methods have been explored to address these challenges. CycleGAN [9] enforces cycle consistency on unpaired datasets to map images between domains. Similarly, contrastive unpaired translation (CUT) [10] employs a contrastive learning framework to align corresponding patches between input and output images without an explicit cycle consistency constraint. UNIT [11] assumes a shared latent space between the two domains, enabling unsupervised translation via coupled generative adversarial networks. Although these methods have shown promising performance, they may cause shape distortions, which are critical in medical imaging. Moreover, since CT typically exhibits lower spatial resolution than CBCT, the high-resolution advantage of CBCT is compromised during translation (see Fig. 1). In this paper, we propose a novel

semi-supervised I2I framework that addresses the above issues. First, we develop an Edge-conditioned unpaired I2I model, inspired by the CUT approach, to generate pseudo-paired CT images. Second, we perform a paired I2I using pseudo-paired data to synthesize realistic CT images without shape distortion. We validate our method against various unpaired I2I methods, demonstrating improvements in both image quality and fidelity. Our contributions can be summarized as follows:

- We present a novel semi-supervised framework for realistic CT synthesis, employing a pseudo-paired CT generation strategy followed by a paired CT generation stage that prevents structural distortion and preserves resolution
- We design an edge-conditioned generator that robustly guides CT synthesis to avoid structural distortion and ensure HU consistency. Moreover, we propose a multi-task loss that combines several loss functions to preserve structure, reduce resolution loss, and achieve accurate HU representation.
- We validate our method using structural integrity, resolution, and HU accuracy metrics, with comprehensive qualitative and quantitative analyses confirming its effectiveness.

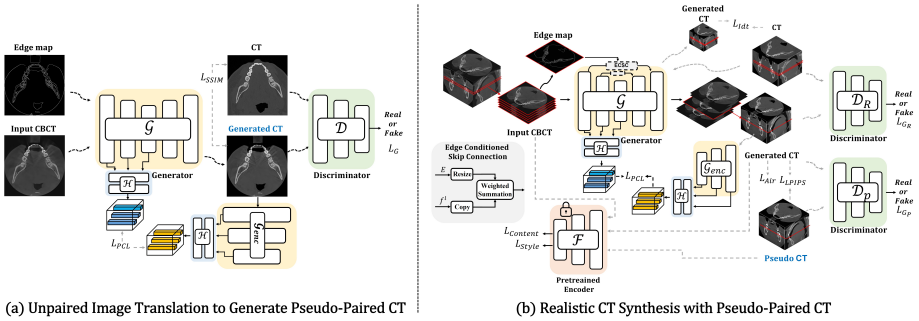
2 Method

2.1 Preliminaries: Contrastive Unpaired Translation

Contrastive Unpaired Translation (CUT) [10] is an unpaired I2I method that maximizes the mutual information between corresponding patches in the source and generated output. Unlike cycle consistency-based approaches, CUT eliminates the need for reverse mapping by using patch-wise contrastive learning to preserve local features during translation. In this process, a source image x is transformed by the generator $\mathcal{G}$ into a translated image $\hat{y} = \mathcal{G}(x)$ that aligns with the target domain. To maintain local feature consistency, CUT maximizes the mutual information between patches in x and the corresponding patches in $\hat{y}$ while distinguishing them from unrelated patches. Patches are embedded as vectors $z_k = \mathcal{H}(\mathcal{G}_{enc}(x)_k)$ and $w_k = \mathcal{H}(\mathcal{G}_{enc}(\hat{y})_k)$, where $\mathcal{H}$ is a multi-layer perceptron (MLP) and k indexes the spatial locations in $\mathcal{G}_{enc}(x)$. The patch-wise contrastive loss is defined as:

$$L_{PatchNCE} = \sum_{k \in \psi} \left[-\log \frac{exp(z_k^\top w_k/\tau)}{\sum_{i \neq k} exp(z_i^\top w_k/\tau)} \right] \tag{1}$$

where ψ is a set of spatial indices sampled from $\{1, 2, ..., S\}$ (with S being the total number of spatial locations) and τ is the temperature parameter controlling the sharpness of the distribution. In this formulation, w_k serves as anchor, (w_k, z_k) constitutes a positive pair, and (w_k, z_i) for $i \neq k$ forms a negative pair.



(a) Unpaired Image Translation to Generate Pseudo-Paired CT

(b) Realistic CT Synthesis with Pseudo-Paired CT

Fig. 2. Overview of the proposed two-stage framework. (a) Unpaired image-to-image translation with edge condition generates pseudo-CT images from CBCT data. (b) A paired translation stage generates realistic CT images, ensuring consistent HU value and preserving resolution.

2.2 Unpaired Image Translation to Generate Pseudo-Paired CT

CUT [10] preserves structural consistency by maximizing mutual information between local patches in the input and generated images, eliminating the cycle consistency loss required by CycleGAN [9] and reducing computational overhead. However, unpaired supervision can still lead to content deformation. To address this, we propose an edge-conditional generator that synthesizes CT while retaining essential content. Specifically, we extract a gradient edge map $x_{edge} \in R^{H\times W\times 1}$, to capture structural boundaries, then concatenate it with the source CBCT image from a two-channel input $x \in R^{H\times W\times 2}$. We also incorporate a multi-scale discriminator [12] that evaluates both global structure and local texture details, enhancing the quality and stability of the translation. This approach faithfully reconstructs fine textures while maintaining structural coherence, forming a crucial step in creating pseudo-paired datasets for higher-fidelity synthesis in the subsequent stage.

2.3 Realistic CT Synthesis with Pseudo-Paired CT

Semi-Supervised Framework. With the constructed pseudo-paired CT dataset, we perform a paired image-to-image translation. Since the pseudo dataset is not entirely accurate, real CT images are incorporated to better align the HU value distribution. To prevent structural deformation, we adopt a UNet++ [13] generator enhanced with Edge-Conditioned Skip Connections (ECSC) and an Interpolation and Convolution Upsampling (ICUP) strategy. ECSC fuses feature maps with the edge map through a learnable weighted summation, preserving resolution and preventing shape distortion. The ICUP method replaces transposed convolution-based upsampling, eliminating checkerboard artifacts and yielding smoother 3D reconstructions. All batch normalization layers are replaced with instance normalization for improved consistency [14]. While the pseudo-paired

dataset exhibits structural coherence and approximate HU alignment, it does not guarantee perfect HU accuracy. To address this, we employ a dual discriminator framework: D_R discriminates between synthetic CT and real unpaired CT, and D_P compares synthetic CT with pseudo-paired CT, thereby enhancing the reliability of the final translation.

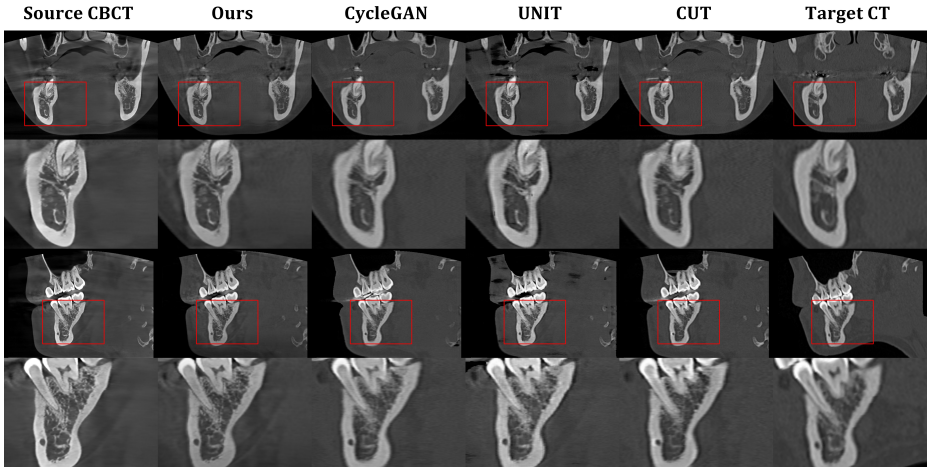


Fig. 3. Qualitative comparison with different methods for unpaired image-to-image translation. The first and third views represent the coronal and sagittal view, respectively. The second and fourth rows correspond to cropped images of red-boxed regions.

Multi-task Loss. Our overall loss function is a weighted sum of components that enforce perceptual quality, structural consistency, and HU value alignment. We include LPIPS loss (L_{LPIPS}) [15], style loss (L_{Style}) [16] to guide HU values, patch-wise contrastive loss (L_{PCL}), content loss ($L_{Content}$) [16] , identity loss (L_{Idt}) [9], and airway segmentation loss (L_{Air}) for structural consistency. Additionally, a dual adversarial loss ($L_{DualAdv}$) is introduced to enforce consistency in pseudo-paired data and to reduce HU differences with CT images. The LPIPS loss preserves perceptual similarity by comparing deep feature representations between the generated and pseudo-paired images:

$$L_{LPIPS} = E_{X\sim P(X)}[\|\phi(G(X)) - \phi(X_p)\|_2^2] \tag{2}$$

where $\phi(\cdot)$ denotes feature embeddings extracted from a frozen pre-trained network. We adopt the VGG-16 [17] pretrained on ImageNet [18]. $G(X)$ is generated image, and X_p is the pseudo-paired reference. The Style loss enforces texture consistency by aligning the Gram matrices of feature maps:

$$L_{Style} = \sum_l \|\text{Gram}(\psi_l(G(X))) - \text{Gram}(\psi_l(X_p))\|_2^2 \tag{3}$$

Table 1. Performance comparison of our method with other unpaired image-to-image translation networks."Structure" denotes the structure score, "Luminance" represents the luminance score, and "Contrast" indicates the contrast score.

Methods	Content & Resolution		Hounsfield Unit		
	Structure (↑)	GMSD (↓)	Luminance (↑)	Contrast (↑)	NCC (↑)
Cycle GAN	0.940 ± 0.049	0.196 ± 0.028	0.792 ± 0.011	0.919 ± 0.028	0.856 ± 0.051
UNIT	0.944 ± 0.028	0.234 ± 0.028	0.736 ± 0.133	0.919 ± 0.029	0.851 ± 0.069
CUT	0.965 ± 0.021	0.146 ± 0.027	0.761 ± 0.110	0.934 ± 0.016	0.863 ± 0.050
OURS	**0.988 ± 0.006**	**0.139 ± 0.015**	**0.809 ± 0.091**	**0.948 ± 0.019**	**0.888 ± 0.046**

where $\psi_l(\cdot)$ represents the feature maps at layer l of pre-trained network, and Gram$(\cdot)$ computes the corresponding Gram matrix. The Patch-wise Contrastive Loss L_{PCL} is applied as defined in Eq. (1) to enhance local feature consistency. The Content loss ensures structural fidelity by minimizing the difference between feature embeddings of the generated and pseudo-paired images:

$$L_{Content} = \sum_{l} \|\psi_l(G(X)) - \psi_l(X_p)\|_2^2 \tag{4}$$

The Identity loss encourages the generator to maintain the target domain's structure:

$$L_{Idt} = E_{X \sim P(X_R)}[\|X - G(X)\|_2^2] \tag{5}$$

where X_R represents images from the real CT dataset. To further enforce structural consistency, particularly in the air regions, we introduce an Air Segmentation Loss that minimizes the mean squared error (MSE) between the predicted air regions of the source CBCT and the generated CT:

$$L_{Air} = E_{X \sim P(X)}[\|S(G(X)) - S(X)\|_2^2] \tag{6}$$

with $S(\cdot)$ denoting the air segmentation network. Finally, to align the anatomical structure similarity of pseudo-paired CT images and the HU value with real CT images, we employ a dual adversarial loss using two discriminators, D_P, and D_R. The dual adversarial loss is defined as:

$$\begin{aligned} L_{DualAdv} =& E_{y \sim P_{real}(Y)}[\log D_R(y)] + E_{x \sim P(X)}[\log(1 - D_R(G(x)))] + \\ & E_{y \sim P_{pseudo}(Y)}[\log D_P(y)] + E_{x \sim P(X)}[\log(1 - D_P(G(x)))] \end{aligned} \tag{7}$$

The total multi-task loss is the weighted sum of these terms:

$$\begin{aligned} L_{Total} =& \lambda_{DualAdv} L_{DualAdv} + \lambda_{LPIPS} L_{LPIPS} + \lambda_{Style} L_{Style} + \\ & \lambda_{Content} L_{Content} + \lambda_{PCL} L_{PCL} + \lambda_{Idt} L_{Idt} + \lambda_{Air} L_{Air} \end{aligned} \tag{8}$$

where $\lambda_{DualAdv} = 2$, $\lambda_{LPIPS} = 1$, $\lambda_{style} = 2$, $\lambda_{content} = 1$, $\lambda_{PCL} = 1$, $\lambda_{idt} = 1$, and $\lambda_{Air} = 1$ are weights determined empirically.

Table 2. Performance comparison of HU values and mean absolute error (MAE) across hard tissue regions (enamel, dentin, cortical bone, and trabecular bone). The proposed method achieves HU values closest to CT and exhibits lower MAE compared to CBCT, CycleGAN, UNIT, and CUT in all regions except trabecular bone.

	Enamel		Dentin		Cortical Bone		Trabecular Bone	
Methods	HU	MAE	HU	MAE	HU	MAE	HU	MAE
CT	2577.67	–	1560.56	–	1355.32	–	557.84	–
CBCT	3453.14	875.46	1956.90	396.34	1953.16	597.83	1128.78	570.93
CycleGAN	2421.24	156.43	1669.25	108.94	1459.47	104.16	503.08	**54.77**
UNIT	2700.22	122.55	1973.66	413.52	1485.27	287.98	833.80	275.93
CUT	2758.98	181.25	2045.66	485.10	1732.15	339.86	398.39	163.96
Ours	2614.89	**37.21**	1620.85	**60.29**	1497.69	**67.35**	708.75	150.91

3 Experiment

3.1 Datasets and Implementation details

Dataset. We collected CBCT and corresponding CT images from 40 subjects (20 males and 20 females; age range 21-80 years), CT scans were acquired using a Somatom Definition Edge CT scanner (Siemens AG, Erlangen, Germany) at 120 kVp and 120 mA, with voxel sizes of $0.49 \times 0.49 \times 0.5mm^3$, and dimensions of 512×512 pixels. CBCT scans were acquired using a Green X 12 CBCT scanner (Vatech, Hwaseong-si, South Korea) at 94 kVp and 11.7 mA, with voxel sizes of $0.2 \times 0.2 \times 0.2mm^3$, and dimensions of 600×600 pixels. For training, 10,625 axial slices from 25 subjects were used, and 6,375 axial slices from 15 subjects were reserved for evaluation. All images were resized to 576×576 pixels, and data augmentation including horizontal flipping, shifts of 0.1%, rotations of 20, and scaling variations of 0.1% was applied to enhance generalization.
Training setup. All models were trained by Adam optimizer for 50 epochs with an initial learning rate of 2×10^{-4}. A batch size of 1 was used, and training was performed on an NVIDIA GeForce RTX 4090 GPU (24 GB RAM). All models were implemented in Python3 using the PyTorch framework and executed in identical computing environments to ensure a fair comparison.
Evaluation Metrics. To assess the performance of the generated images, we computed the structure score from the structural similarity index measure (SSIM) [19] and the gradient magnitude similarity deviation (GMSD) [20] to evaluate structural and resolution similarity between the generated output and the source CBCT. Additionally, we used luminance and contrast scores derived from SSIM, along with normalized cross-correlation (NCC), to evaluate the similarity of the HU distribution between the generated output and the CT images.

3.2 Experimental Results

Quantitative Comparison with other methods. We evaluated the performance of our proposed method against several unpaired image-to-image translation methods, including CycleGAN [9], UNIT [11], and CUT [10]. Our method

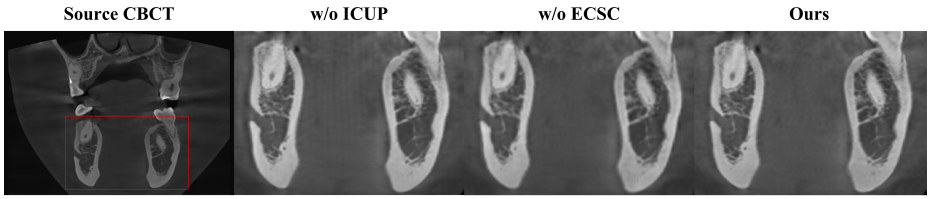


Fig. 4. Qualitative comparison of the impact of ICUP and ECSC. The red box indicates the region of interest. Without ICUP, checkerboard artifacts cause vertical line distortions, and without ECSC, resolution is degraded.

achieves the best performance across all evaluation metrics. In particular, our method outperforms CUT with the second highest performance by obtaining 0.988±0.006, 0.139±0.015, 0.809±0.091, 0.948±0.018, and 0.888±0.046 for Structure, GMSD, Luminance, Contrast, and NCC, as summarized in Table 1. Table 2 presents the analysis results for hard tissues. Among enamel, dentin, cortical bone, and trabecular bone, our method demonstrated most similar to CT images in all regions except trabecular bone by obtaining mean absolute error (MAE) 37.21, 60.29, 67.35, and 150.91.

Qualitative Comparison with other methods. Figure 3 shows that our method produces images that closely resemble the CT, preserving structural details without deformation while maintaining a resolution comparable to the source CBCT. Moreover, as demonstrated in Figure 1, our approach generates realistic CT images that closely approximate the target CT, outperforming other methods. In particular, while other models exhibit slice inconsistency, our method achieves exact slice consistency.

Effectiveness of proposed components. We conducted ablation experiments to assess the effectiveness of our proposed ICUP and ECSC modules. Table 3 shows the quantitative performance variations with and without these components. The results indicate that ICUP effectively suppresses checkerboard artifacts, eliminating vertical line artifacts in the reconstructed volumes. Additionally, ECSC robustly prevents shape deformation and reduces resolution loss by explicitly integrating edge conditions into the skip connections (see Fig. 5).

Table 3. Quantitative performance comparison of the proposed ICUP and ECSC modules.

Methods	Content & Resolution		Hounsfield Unit		
	Structure (↑)	GMSD (↓)	Luminance (↑)	Contrast (↑)	NCC (↑)
Baseline	0.968 ± 0.003	0.150 ± 0.020	0.772 ± 0.120	0.945 ± 0.016	0.878 ± 0.046
+ICUP	0.970 ± 0.004	0.143 ± 0.028	0.782 ± 0.112	0.945 ± 0.015	0.877 ± 0.046
+ECSC	0.980 ± 0.003	0.141 ± 0.018	0.788 ± 0.106	0.946 ± 0.016	0.882 ± 0.045
+ICUP, ECSC (Ours)	**0.988 ± 0.006**	**0.139 ± 0.015**	**0.809 ± 0.091**	**0.948 ± 0.019**	**0.888 ± 0.046**

4 Conclusion

In this paper, we proposed a semi-supervised framework for CBCT-to-CT synthesis that ensures the exact preservation of anatomical shape, high resolution, and accurate HU value. Our two-stage approach employs an edge-conditioned unpaired translation model to generate pseudo-paired CT images, followed by a paired translation stage using a UNet++ generator enhanced by ICUP, ECSC, and dual discriminators. Extensive evaluations demonstrate that our method outperforms existing techniques, achieving HU values closest to the CT. In future work, we plan to further explore artifact-free CT synthesis to enhance clinical applicability.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

Acknowledgements. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No.2023R1A2C2 00532611). This work was also supported by the Technology Innovation Program (or Industrial Strategic Technology Development Program-Advanced Biomaterials) (RS-2025-14322975), funded by the Ministry of Trade, Industry & Energy (MOTIE).

References

1. Choi, J.W., Lee, S.S., Choi, S.C., Heo, M.S., Huh, K.H., Yi, W.J., Kang, S.R., Han, D.H., Kim, E.K.: Relationship between physical factors and subjective image quality of cone-beam computed tomography images according to diagnostic task. Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology **119**(3), 357–365 (2015)
2. Scarfe, W.C., Farman, A.G., Sukovic, P., et al.: Clinical applications of cone-beam computed tomography in dental practice. Journal-Canadian Dental Association **72**(1), 75 (2006)
3. Kiljunen, T., Kaasalainen, T., Suomalainen, A., Kortesniemi, M.: Dental cone beam ct: A review. Physica Medica **31**(8), 844–860 (2015)
4. Ludlow, J.B., Ivanovic, M.: Comparative dosimetry of dental cbct devices and 64-slice ct for oral and maxillofacial radiology. Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology **106**(1), 106–114 (2008)
5. Dahiya, N., Alam, S.R., Zhang, P., Zhang, S.Y., Li, T., Yezzi, A., Nadeem, S.: Multitask 3d cbct-to-ct translation and organs-at-risk segmentation using physics-based data augmentation. Medical physics **48**(9), 5130–5141 (2021)
6. Kang, S.R., Shin, W., Yang, S., Kim, J.E., Huh, K.H., Lee, S.S., Heo, M.S., Yi, W.J.: Structure-preserving quality improvement of cone beam ct images using contrastive learning. Computers in Biology and Medicine **158**, 106803 (2023)
7. Yong, T.H., Yang, S., Lee, S.J., Park, C., Kim, J.E., Huh, K.H., Lee, S.S., Heo, M.S., Yi, W.J.: Qcbct-net for direct measurement of bone mineral density from quantitative cone-beam ct: A human skull phantom study. Scientific Reports **11**(1), 15083 (2021)

8. 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 **15**(7), 2017 (2023)
9. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE international conference on computer vision. pp. 2223–2232 (2017)
10. Park, T., Efros, A.A., Zhang, R., Zhu, J.Y.: Contrastive learning for unpaired image-to-image translation. In: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IX 16. pp. 319–345. Springer (2020)
11. Liu, M.Y., Breuel, T., Kautz, J.: Unsupervised image-to-image translation networks. Advances in neural information processing systems **30** (2017)
12. Kumar, K., Kumar, R., De Boissiere, T., Gestin, L., Teoh, W.Z., Sotelo, J., De Brebisson, A., Bengio, Y., Courville, A.C.: Melgan: Generative adversarial networks for conditional waveform synthesis. Advances in neural information processing systems **32** (2019)
13. Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J.: Unet++: A nested u-net architecture for medical image segmentation. In: Deep learning in medical image analysis and multimodal learning for clinical decision support: 4th international workshop, DLMIA 2018, and 8th international workshop, ML-CDS 2018, held in conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, proceedings 4. pp. 3–11. Springer (2018)
14. Ulyanov, D., Vedaldi, A., Lempitsky, V.: Instance normalization: The missing ingredient for fast stylization. arXiv preprint arXiv:1607.08022 (2016)
15. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 586–595 (2018)
16. Gatys, L.A., Ecker, A.S., Bethge, M.: Image style transfer using convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2414–2423 (2016)
17. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
18. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition. pp. 248–255. Ieee (2009)
19. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing **13**(4), 600–612 (2004)
20. Xue, W., Zhang, L., Mou, X., Bovik, A.C.: Gradient magnitude similarity deviation: A highly efficient perceptual image quality index. IEEE transactions on image processing **23**(2), 684–695 (2013)