Unpaired-paired learning for shading correction in cone-beam computed tomography

HYOUNG SUK PARK¹, KIWAN JEON¹, SANG-HWY LEE², AND JIN KEUN SEO³

¹National Institute for Mathematical Sciences, Daejeon, 34047, South Korea
²Department of Oral and Maxillofacial Surgery, Oral Science Research Center, College of Dentistry, Yonsei University, Seoul, South Korea
³School of Mathematics and Computing (Computational Science and Engineering), Yonsei University, Seoul, South Korea

Corresponding author: Jin Keun Seo (email: seoj@yonsei.ac.kr).

This work was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (HI20C0127). H.S.P. and K.J. were partially supported by the National Institute for Mathematical Sciences (NIMS) grant funded by the Korean government (No. NIMS-B22910000). J.K.S. was partially supported by Samsung Science & Technology Foundation (No. SRFC-TT1902-09).

ABSTRACT Cone-beam computed tomography (CBCT) is widely used in dental and maxillofacial imaging applications. However, CBCT suffers from shading artifacts owing to several factors, including photon scattering and data truncation. This paper presents a deep-learning-based method for eliminating the shading artifacts that interfere with the diagnostic and treatment processes. The proposed method involves a two-stage generative adversarial network (GAN)-based image-to-image translation, where it operates on unpaired CBCT and multidetector computed tomography (MDCT) images. The first stage uses a generic GAN along with the fidelity difference between the original CBCT and MDCT-like images generated by the network. Although this approach is generally effective for denoising, at times, it introduces additional artifacts that appear as bone-like structures in the output images. This is because the weak input fidelity between the two imaging modalities can make it difficult to preserve the morphological structures from complex shadowing artifacts. The second stage of the proposed model addresses this problem. In this stage, paired training data, excluding inappropriate data, were collected from the results obtained in the first stage. Subsequently, the fidelity-embedded GAN was retrained using the selected paired samples. The results obtained in this study reveal that the proposed approach substantially reduces the shadowing and secondary artifacts arising from incorrect data fidelity while preserving the morphological structures of the original CBCT image. In addition, the corrected image obtained using the proposed method facilitates accurate bone segmentation compared to the original and corrected CBCT images obtained using the unpaired method.

INDEX TERMS Computed tomography, shading correction, unpaired learning, generative adversarial network

I. INTRODUCTION

Dental cone-beam computed tomography (CBCT) is being increasingly used for diagnosis and treatment in implant, dental and maxillofacial surgery [2], [12]. The three-dimensional (3D) models of the maxillofacial bones extracted from CBCT images are commonly used during preoperative analysis and planning. Therefore, accurate bone segmentation from CBCT images is important to facilitate sophisticated treatment planning. However, most dental CBCT devices are designed to reduce the radiation dose by limiting the scan field of view (FOV) [35]. Moreover, they frequently have no pre- and post-patient collimation, which greatly reduces the amount of scattered radiation [8]. The photon scattering and data truncation due to the limited FOV of CBCT cause bright and dark shadows in the reconstructed image, significantly degrading the image quality of CBCT and interfering with accurate bone segmentation. In contrast, multidetector computed tomography (MDCT) images are considered nearly scatter-free [23] and void of data-truncation errors, as depicted in Fig. 1 (b). This suggests the possibility of developing an artifact-correction function that maps CBCT images to MDCT-like images that contain negligible shading artifacts.

Several approaches for reducing shading artifacts in CBCT images have been reported in previous studies. Scattering...
correction involves the use of methods based on Monte Carlo (MC) simulations [5], [15], beam blockers [22], [30], [56], and scatter kernels [4], [29]. Meanwhile, truncation correction involves the use of methods based on projection extrapolation [9] and iterative reconstruction using various handcrafted priors [16], [31], [34]. Although these methods have demonstrated potential to reduce the shading artifacts, their use in clinical applications is faced with several limitations, such as high computational cost, additional scan time, and additional exposure.

Recently, deep-learning techniques have been successfully employed to perform CT-image enhancements, such as noise reduction [10], [18], [24], [32]. In clinical practice, acquiring (real-world) paired data is virtually impossible owing to difficulties associated with the simultaneous attainment of the patient CT data and corresponding ground-truth image. Therefore, some researchers have adopted the generative adversarial network (GAN) [6] to facilitate the learning of a correction function using unpaired data from artifact-free and artifact-affected images. In the variational framework, the unpaired model can be viewed as an energy minimization problem of a correction function. This function includes data fidelity and data-driven regularization (referred to as fidelity-embedded GAN) therein the data fidelity refers to the difference between an input image (denoted by \( z \)) and the corresponding image (denoted by \( G(z) \)) generated by the GAN network [24], [32]. It might include additional cycle-loss [7], [10], [18].

This study adopts a fidelity-embedded GAN that uses unpaired CBCT and MDCT images to reduce the shading artifacts. In general, it is difficult to infer the desired MDCT image priors using unpaired training data from two different imaging modalities (i.e., CBCT and MDCT). Therefore, learned data-driven regularization can be used to generate plausible structures from complex shading artifacts. The roughly enforced data fidelity (i.e., the sum of the difference between the input \( z \) and generated output \( G(z) \)) could lead to the problem of preserving the morphological structures of the input \( z \). Consequently, the use of the fidelity-embedded GAN results in the unwanted distortion of some pixels. Fig. 1(c) reveals that the fidelity-embedded GAN creates unwanted bone-like artifacts in the output. These artifacts vary with the input data.

This problem is solved by retraining the fidelity-embedded GAN using a filtered paired dataset \( \{\{z_k, x_k\} : k = 1, \cdots, K\} \). In this dataset, \( z_k = G(x_k) \), \( k = 1, \cdots, K \), are generated by selecting the correct output images produced via application of the fidelity-embedded GAN to unpaired data. The resulting change in fidelity (i.e., the sum of the difference between \( G(z) \) and labeled data \( x \)) dramatically improves the reduction of the bone-like and shading artifacts (Fig. 1(d)). The proposed unpaired–paired learning approach was validated using a clinical CBCT dataset. Moreover, in this study, the feasibility of the proposed method was further investigated via application of the same to accomplish an image-segmentation task.

II. METHOD

In this paper, \( z \) represents the CBCT image and \( x \) represents the MDCT image. Let \( Z \) and \( X \) be distributions of CBCT images and MDCT images, respectively. The goal is to learn CBCT-to-MDCT transformation \( G : Z \to X \) from unpaired training samples \( S_z = \{z_i\}_{i=1}^N \) and \( S_x = \{x_j\}_{j=1}^M \), that are drawn from \( Z \) and \( X \), respectively. Note that it is impossible to know the probability density functions \( p_z \) and \( p_x \) corresponding to \( Z \) and \( X \), respectively.

Specifically, the proposed method is a two-stage GAN-based image-to-image translation method that trains the correction map \( \hat{G} : z \to x \) by minimizing the following loss model:

\[
dist(p_{G(z)}, p_x) + \lambda F_{\text{pixel-level}}(G(z))
\]

where \( p_{G(z)} \) is the density of \( G(z) \) in the distribution generated by the generator \( G \); \( dist(p_{G(z)}, p_x) \) is the metric that measures the distance between two probability density functions \( p_{G(z)} \) and \( p_x \); and \( F_{\text{pixel-level}}(G(z)) \) represents the pixel-level fidelity of \( G(z) \), which is constrained to preserve the structure of the CBCT image \( z \). Here, \( \lambda > 0 \) is the regularization parameter that controls the tradeoff between \( \phi_{\text{pixel}}(p_{G(z)}, p_x) \approx 0 \) and \( F_{\text{pixel-level}}(G(z)) \). Without paired training data like \( (z, x) \), we cannot train \( G \) with the good pixel-level fidelity \( ||G(z) - x|| \).

For the metric \( dist(p_{G(z)}, p_x) = \chi^2_{\text{Pearson}}(p_{G(z)} + \)

\[
\begin{align*}
\chi^2_{\text{Pearson}}(p_{G(z)} + p_x - 1, X) \quad &
\end{align*}
\]

(a) CBCT image

(b) MDCT image

(c) unpaired learning

(d) the proposed method

FIGURE 1. Results obtained using the proposed method that substantially reduces the shading artifacts. (a) CBCT image and (b) MDCT image (WW/WL=1500/250 for all CT images).
p_x, 2p_z), where \( \chi^2_{Pearson} \) is the Pearson \( \chi^2 \) divergence, minimizing of \( dist(p_G(z), p_x) \) in (1) can be converted to the least squares GAN (LSGAN) framework [19], [20]: \( G \) is trained simultaneously with a discriminator \( D \) in an adversarial relationship to improve their mutual abilities;

\[
\begin{align*}
\arg\min_G E_{z \sim p_z} \left[ D(G(z))^2 \right] \\
\arg\min_D E_{x \sim p_x} \left[ \left(1 - D(x)\right)^2 \right]
+ E_{z \sim p_z} \left[ \left(1 + D(G(z))\right)^2 \right].
\end{align*}
\]

where \( E[x] \) is the expectation of \( x \).

For the constraint of the pixel-level fidelity \( F_{\text{pixel-level}}(G(z)) \) in (1), we use two different fidelities. First, we obtain an unpaired generator \( G^u \) using the fidelity \( F_{\text{pixel-level}}(G(z)) = E_{z \sim p_z} \| G(z) - z \| \). This \( G^u \) is used to select a subset of \( Z \) that consists of \( z \) where \( G^u(z) \) is a MDCT-like image. Using this subset as training data, we train a paired generator \( G^p \) using the fidelity norm \( \| G(z) - G^u(z) \| \), which is equivalent to supervised learning.

Now, we are ready to explain practical details of the proposed method, where finite training samples, \( S_z \) and \( S_x \), are used to train \( G \).

A. STAGE 1: UNPAIRED MODEL

The first step of the proposed method is to generate MDCT-like images from unpaired CBCT images using the following LSGAN model:

\[
\begin{align*}
G^u &= \arg\min_G E_{z \sim S_z} \left[ D(G(z))^2 \right] \\
&+ \lambda E_{z \sim S_z} \left[ \| G(z) - z \| \right] \\
D^u &= \arg\min_D E_{x \sim S_x} \left[ \left(1 - D(x)\right)^2 \right] \\
&+ E_{z \sim S_z} \left[ \left(1 + D(G(z))\right)^2 \right]
\end{align*}
\]

where \( E_{z \sim S_z} \| (G(z) - z) \| \) corresponds to the constraint of the pixel-level fidelity \( F_{\text{pixel-level}}(G(z)) \).

With this pixel-level fidelity \( \| G(z) - z \| \), it is generally difficult to train \( G \), resulting in generating the plausible artifacts that do not exist in MDCT images. To explain the reason, suppose that \( x^* \) is a target MDCT image corresponding to the input \( z \). If \( \| x^* - z \| \) is reasonably small, then the method of (3) can generate \( G^u(z) \) that is different from the target \( x^* \). On the other hand, if \( \| x^* - z \| \) is not small, then the method of (3) may produce an unwanted \( G^u(z) \) that is different from the target \( x^* \). Hence, the performance of \( G^u \) depends heavily on the input \( z \) and the parameter \( \lambda \).

Fig. 2 depicts various \( G^u(z) \) with different values of \( \lambda \). The case corresponding to \( G^u(z) \) with \( \lambda = 0 \) in (3) exclusively reflects the overall characteristics of MDCT images, whereas it loses the detailed structures of \( z \), thereby resulting the generation of bone-like structures owing to the use of imperfect data-driven priors. In contrast, the case wherein \( G^u(z) \) is generated considering a small value of \( \lambda = 10^{-2} \) reveals reduced bone-like artifacts, albeit the morphological

FIGURE 2. Results of the unpaired learning with different regularization parameters \( \lambda \) in (3) (WW WL=1500/250 for all CT images)
structures of the \( z \) (blue arrows in Fig. 2) are nonetheless lost. The case wherein \( G^u(z) \) is generated considering \( \lambda = 10^1 \) demonstrates an effective reduction in the scattering artifacts while preserving the morphological structures of the tissues. However, in the case where \( z \) is severely corrupted by the shading artifacts (i.e., \( \| x - z \| \gg 0 \)), the bone-like artifacts may remain in \( G^u(z) \) (yellow arrows in Fig. 2). Hence, the model described in [1] may cause the problem in preserving the morphological structures of tissues and bones from the complex shading and plausible artifacts as well.

### B. STAGE 2: PAIRED MODEL

Stage 2 uses the observation from Stage 1 that the performance of \( G^u \) depends on the input \( z \). For each \( z_i \in S_z \), we examine whether \( G^u(z_i) \) is a MDCT-like image or not. We choose a subset of \( S_z \), denoted by \( S_{z,G}^{G^u} \), such that \( \{G^u(z) : z \in S_{z,G}^{G^u}\} \) is a set consisting only of MDCT-like images. The proposed method uses the set \( S_{pair} := \{(z, G^u(z)) : z \in S_{z,G}^{G^u}\} \) as paired training data for supervised learning. Now, we train \( G \) using the selected paired dataset \( S_{pair} := \{(z, G^u(z)) : z \in S_{z,G}^{G^u}\} \) through the following paired model:

\[
\begin{align*}
G^p &= \arg \min_G E_{z \sim S_{z,G}^{G^u}} [D(G(z))^2] + \lambda E_{z \sim S_{z,G}^{G^u}} [(G(z) - G^u(z))^2] \\
D^p &= \arg \min_D E_{x \sim S_z} [(1 - D(x))^2] + E_{z \sim S_{z,G}^{G^u}} [(1 + D(G(z))^2]
\end{align*}
\]

In our experiment, for the trained \( G^u \) and \( G^p \), it follows that \( dist(p_{G^u}(z), p_\theta) = 0.24 \), and \( dist(p_{G^p}(z), p_\theta) = 0.19 \). This implies that using the selected paired data, the trained \( G^p \) generates images that effectively reflect the characteristics of the MDCT domain more accurately compared to \( G^u \).

Here, for trained \( G^u \) and \( G^p \), the distances \( dist(p_{G^u}(z), p_\theta) \) and \( dist(p_{G^p}(z), p_\theta) \) are approximately calculated using finite samples \( S_z, S_{z,G}^{G^u} \) and \( S_x \) [20], [25]:

\[
dist(p_{G^u}(z), p_\theta) \approx E_{x \sim S_x} [D^u(x)^2] + E_{z \sim S_z} [D^u(G^u(z))^2]
\]

and

\[
dist(p_{G^p}(z), p_\theta) \approx E_{x \sim S_x} [D^p(x)^2] + E_{z \sim S_{z,G}^{G^u}} [D^p(G^p(z))^2],
\]

respectively.

In summary, the operation of the proposed method involves the following steps.

1) Training of \( G^u \) in [3] using the unpaired datasets \( S_z \) and \( S_x \)
2) Training of \( G^p \) in [4] using the selected paired datasets \( S_{pair} \).

The diagram of the proposed method is illustrated in Fig. 3.

### C. DATASETS

This study was approved by the institutional review board of our institution, and the requirement for informed consent was waived. We collected the CBCT and MDCT images of 20 and 28 patients, respectively. The CBCT images were acquired using a circular-trajectory CBCT scanner (Xoran CAT, Xoran Technologies; USA) considering a pixel size of 0.40 mm × 0.40 mm, image size of 400 × 400, tube voltage of 120 kVp, and tube current of 48 mA. Note that several dental CBCT scanners, such as i-CAT [3], [21], KaVo 3D eXam [28], and ProMax 3-D Max [14], reconstruct images using a pixel size of 0.4 mm × 0.4 mm. In contrast, the MDCT images were acquired using a helical-trajectory MDCT scanner (SOMATOM Definition Flash, Siemens Healthineers;
patches with a size of \(128 \times 128\) were extracted considering a Gaussian distribution efficiently. In this study, given a CT image, the patch manifold is generally characterized as described in Sections II-A and II-B. Finally, the corrected CT images were synthesized using the corrected patches. The pixel values in the overlapping patch regions were averaged.

D. NETWORK ARCHITECTURES

Throughout this paper, we used the same architectures for both \(G_u\) and \(G_p\), and for both \(D_u\) and \(D_p\), respectively. For the generator, we adopted the deep convolutional framelets [33], a multi-scale convolutional neural network (CNN). For the discriminator, we adopted the standard CNN without a fully connected layer at the end of the network. The network architectures are exactly same as those of [24], [25].

In our experiments, the update of both generator and discriminator was performed by using the Adam solver [13] with a learning rate of 0.0002, min-batch size of 20, and number of epochs of 100. The network weights were initialized following a Gaussian distribution with a mean of 0 and standard deviation (SD) of 0.01. We chose \(\lambda = 10\) in (5) and \(\lambda = 100\) in (4), respectively. Training was implemented using TensorFlow [1] on a CPU (Intel(R) Core(TM) i9-9980XE, 3.0GHZ) and a GPU (NVIDIA, Titan RTX 24GB) systems. It took approximately 29 hours to train each \(G_u\) and \(G_p\).

III. RESULTS

Fig. [5] compares the performance of the proposed method with the unpaired learning methods for the test patients. Furthermore comparisons were performed using the unpaired model with an additional cyclic loss [37] (called cycle-GAN). More specifically, we solved the following unpaired problem.

\[
\begin{align*}
G^{\text{cycle}} &= \arg\min_{D} E_{z \sim S_z} [D(G(z))^2 + \lambda \|G(z) - z\| \\
&\quad + \|T(G(z)) - z\| + \|G(T(x)) - x\|] \\
D^{\text{cycle}} &= \arg\min_{E} E_{x \sim S_x} \left[(1 - D(x))^2 \right] \\
&\quad + E_{z \sim S_z} \left[(1 + D(G(z))^2) \right] \\
\end{align*}
\]

where \(T: x \rightarrow z\) is the inverse function of \(G^{\text{cycle}}\) satisfying

\[
\begin{align*}
T &= \arg\min_{D} E_{x \sim S_x} \left[D(T(x))^2 + \lambda \|T(x) - x\| \right] \\
&\quad + \|G(T(x)) - z\| + \|G(T(x)) - x\|) \\
D^t &= \arg\min_{E} E_{z \sim S_z} \left[(1 - D(z))^2 \right] \\
&\quad + E_{x \sim S_x} \left[(1 + D(T(x))^2) \right].
\end{align*}
\]

During the experiment performed in this study, the parameter values were selected as \(\lambda = 10\). \(\mu = 10\). The parameters with the best performance (maximum Dice coefficient(DC)) regarding bone segmentation were manually selected. Numerous recently published papers [7], [10], [17], [18] have adopted cycle-GAN or cycle-GAN version variants to train unpaired datasets. As shown in Fig. [5], both \(G_u\) and \(G^{\text{cycle}}\) substantially reduce the shading artifacts while enhancing the contrast between the bone and tissue. The intensity profiles depicted in Fig. [6] confirm the contrast enhancement. However, the weak data fidelity in both unpaired-learning models causes the appearance of bone-like artifacts in the

FIGURE 4. Preprocessing of the MDCT image for training

Overall, 11000 CBCT (\(S_z\)) and 11422 MDCT (\(S_p\)) images obtained from 18 and 28 patients, respectively, were used during the unpaired training (\(G_u\) in [4]). Meanwhile, 11000 CBCT images obtained from two patients comprised the test dataset. For efficient learning, we randomly selected 1024 CBCT and MDCT images per epoch and used them to update the networks. Upon completion of the unpaired training, 6831 corrected CBCT images, which remained unaffected by the bone-like artifacts, and 6831 original CT images (\(S_{pair}\)) were selected and used for paired learning (\(G_p\) in [4]).

Instead of learning entire images at once, the proposed networks learn individual images in a patch-by-patch manner. This is because the patch manifold is generally characterized by a low-dimensional structure compared to the image manifold [26], [27]. This allows the networks to learn the generated distribution efficiently. In this study, given a CT image, patches with a size of \(128 \times 128\) were extracted considering strides of 16 along each direction in the image domain.
images generated by both \( G^u \) and \( G^{\text{cycle}} \). In contrast, the proposed method substantially reduces the bone-like artifacts while preserving the morphological structures of bones and tissues. In addition, the proposed method does not appear to introduce unexpected new artifacts. The fidelity using the paired dataset \( S_{\text{pair}} \) selected in the second stage appears to prevent the creation of new unexpected artifacts.

We computed the mean and SD for bone intensity, where the bone mask was manually selected except for teeth. Table 1 presents a summary of the computed mean and SD values for CBCT and the corrected CBCT on the test dataset. As listed in Table 1, the mean HU values of the corrected CBCT images are increased over those of the uncorrected CBCT image. This is because the deep learning methods reduce shading artifacts while enhancing the contrast between the bone and tissue.

Fig. 7 compares the 3D bone-segmentation results between the original and corrected CBCT images obtained using the unpaired and proposed methods. The 3D modeling was performed using the MIMIC software (MIMIC, Materialise; Belgium). The threshold value (tv) for bone segmentation was chosen to maximize the DC in the training dataset. Thereafter, the selected tv was used for bone segmentation in two test patients. Here, manual segmentation performed by a 3D bone-segmentation expert was considered as a ground truth. The DC value between two binary images (or sets) \( x \) and \( y \) was calculated as

\[ DC(x, y) = \frac{2|x \cap y|}{|x| + |y|}. \]

The chosen tvs for the uncorrected and corrected 3D CBCT images are illustrated in Fig. 7. The results are visualized using the Visualization Toolkit (Kitware Inc.). As depicted in Fig. 7 compared to the original CBCT image and its corrected version obtained using unpaired learning, more accurate bone segmentation can be realized in the corrected CT image obtained using the proposed method.

Moreover, the segmentation quality has been objectively assessed by evaluating the mean DC for the two test patients. Table 2 presents the corresponding results. As illustrated in Table 2, the corrected image obtained using the proposed method achieves a higher DC value compared to the original.
CBCT image and that obtained using the unpaired-learning model.

FIGURE 6. Intensity profiles of the corrected CT images along the dotted line in Fig. 5.

TABLE 2
QUANTITATIVE EVALUATIONS OF SEGMENTATION QUALITY USING THE DC FOR THE TWO TEST PATIENTS.

| Patient  | CBCT | unpaired | unpaired-cycle | proposed |
|----------|------|----------|----------------|----------|
| Patient 1 | 0.8107 | 0.8547   | 0.8401         | 0.9052   |
| Patient 2 | 0.8335 | 0.8631   | 0.8571         | 0.9098   |

IV. DISCUSSION AND CONCLUSION

This paper presents an unpaired CBCT-to-MDCT translation method to alleviate photon scattering and truncation error, one of the main factors that degrade the quality of dental CBCT images. Although fidelity-embedded GAN approaches for unpaired learning have demonstrated promising results for denoising CT images, their performance has remained limited owing to the use of naive fidelity despite significant differences between CBCT and MDCT. This naive fidelity, which constrains the output image to match the corresponding input spatially and anatomically, demonstrates inconsistent performance, which varies depending on the input image quality. This creates unwanted bone-like artifacts during image-to-image translation. For similar reasons, despite the recent advances in image-to-image translation, the application of GAN-based techniques, such as cycle-GAN, in high-quality medical-imaging applications remains a challenge.

The above-discussed problems can be resolved using paired learning. Based on the domain knowledge of the morphological structure of the human skull, a clinical expert constructed paired data by manually collecting the accurate results obtained using the unpaired-learning model. Using the selected paired data, the same GAN-based method with its fidelity term influenced by the pair difference demonstrated significantly improved performance. Fig. 5 illustrates the empirical effect of data fidelity. Note that even the input images that produced bone-like artifacts during unpaired learning did not generate them during the secondary paired learning. It is very interesting to observe that the CBCT image correction, which failed in the first stage, succeeded in the second stage with the paired learning. In addition, the proposed approach provided more accurate results for 3D bone segmentation compared to unpaired learning. It takes approximately 200 ms to acquire a single 400 × 400 corrected CBCT image. This can be reduced by using, for example, parallel computation and a multiple GPU system.

The proposed method can be improved in the following aspects. First, in the fidelity-embedded GAN, the consideration of high-accuracy data fidelity considering CT physics can improve the network performance in terms of shading correction. For example, using data fidelity that considers photon noise in low-dose CT images can be combined with GAN [25]. Second, manual selection in the second stage of the proposed network can be time-consuming and subjective. Therefore, further research to improve the robustness of the proposed approach should be undertaken in the near future.

V. APPENDIX

A. IMPACT OF DIFFERENT IMAGE RESOLUTIONS OF CBCT

In this section, we investigated the impact of the CBCT image pixel size. The training was performed using CBCT scans of 30 patients and MDCT scans of 28 patients, which were acquired using a circular-trajectory CBCT scanner (NewTom, pixel size of 0.25 mm) and a helical-trajectory MDCT scanner (SOMATOM Definition Flash, pixel size of 0.42–0.49 mm), respectively. Note that the CBCT scanner has a smaller pixel size than the MDCT scanner used as a label. Fig. 8 shows the results of the proposed approach in axial, coronal, and sagittal views. As shown in Fig. 8, the proposed method effectively reduces the shading artifacts; even for the CBCT image with a smaller pixel size, the resolution degradation is rarely seen in the corrected image.

FIGURE 8. Results of the proposed method for the CBCT with a pixel size of 0.25 mm. (WW/WL = 1500/250 for all CT images)
FIGURE 7. Comparison of the 3D bone segmentation from the original and corrected CBCT images for two test patients. Here, the threshold value (tv) was selected to maximize the Dice coefficient.

REFERENCES

[1] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, *et al.*, “TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems”, 2016, [Online] Available: https://arxiv.org/abs/1603.04467.

[2] H. M. Alamri, M. Sadrameli, M. A. Alshalhoob, and M. A. Alshehri, “Applications of CBCT in dental practice: a review of the literature”, Gen. Dent., vol. 60, pp. 390–400, 2012.

[3] S. Baunsgaeterl, J. M. Palomo, L. Palomo, and M. G. Hans, “Reliability and accuracy of cone-beam computed tomography dental measurements”, Am. J. Orthod. Dentofac. Orthop., vol. 136, pp. 19–25, 2009.

[4] J. M. Boone, and J. A. Seibert, “An analytical model of the scattered radiation distribution in diagnostic radiology”, Med. Phys., vol. 15, pp. 721–725, 1988.

[5] A. P. Colijn, and F. J. Beekman, “Accelerated simulation of cone beam X-ray scatter projections”, IEEE Trans. Med. Imaging, vol. 23, pp. 584–590, 2004.

[6] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets”. In NIPS, 2014.

[7] J. Gu, and J. C. Ye, "AdaIN-based tunable cycleGAN for efficient unsupervised low-dose CT denoising", IEEE Trans. Comput. Imaging, vol. 7, pp. 73–85, 2021.

[8] P. P. Jaju, M. Jain, A. Singh, and A. Gupta, “Artefacts in cone beam CT”, Open Journal of Stomatology, vol. 3, pp. 292, 2013.

[9] J. Hsieh, E. Chao, J. Thibault, B. Grekovicz, A. Horst, S. McOlash, and T. J. Myers, “Algorithm to extend reconstruction field-of-view.” In 2004 2nd IEEE International Symposium on Biomedical Imaging, 2004.

[10] E. Kang, H. J. Koo, D. H. Yang, J. B. Seo, and J. C. Ye, “Cycle-consistent
adversarial denoising network for multiphase coronary CT angiography", Med. Phys., vol. 46, pp. 550–562, 2019.

[11] S. H. Kang, K. Jeon, S. H. Kang, and S. H. Lee, "3D cephalometric landmark detection by multiple stage deep reinforcement learning", Sci. Rep., vol. 11, pp. 1–13, 2021.

[12] S. D. Kapila and J. M. Nervina, "CBCT in orthodontics: assessment of treatment outcomes and indications for its use", Dentomaxillofacial Radiol., vol. 44, pp. 20140282, 2015.

[13] D. P. Kingma, and J. Ba, "Adam: A Method for Stochastic Optimization", 2014, [Online] Available: https://arxiv.org/abs/1412.6980.

[14] T. Riedel, and W. A. Kalender, "Combining deterministic and Monte Carlo calculations for fast estimation of scatter intensities in CT", Med. Phys. Biol., vol. 51, pp. 4567, 2006.

[15] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, and S. P. Smolley, "Least Squares Generative Adversarial Networks", 2017, [Online] Available: https://arxiv.org/abs/1611.04076

[16] M. Lee, Y. Han, J. P. Ward, et al., "Interior tomography using 1D generalized total variation. Part II: Multiscale implementation", SIAM J. Imaging Sci., vol. 8, pp. 2452–2486, 2015.

[17] J. Lee, J. Gu, and J. C. Ye, "Unsupervised CT metal artifact learning using attention-guided beta-cyclegan", IEEE Trans. Med. Imaging, vol. 40, pp. 3932 – 3944, 2021.

[18] Z. Li, S. Zhou, J. Huang, L. Yu, and M. Jin, "Investigation of low-dose CT image denoising using unpaired deep learning methods", IEEE Trans. Radiat. Plasma Med. Sci., vol. 5, pp. 224–234, 2020.

[19] J. P. Ward, M. Lee, J. C. Ye, and M. Unser, "Interior tomography using 1D generalized total variation. Part I: Mathematical foundation", SIAM J. Imaging Sci., vol. 36, pp. 2258–2268, 2009.

[20] T. Niu, M. Sun, and J. M. Star-Lack, "Improved scatter correction using adaptive scatter kernel superposition", Phys. Med. Biol., vol. 55, pp. 6695, 2010.

[21] F. C. Wagner, A. Macovski, and D. C. Nishimura, "Dual-energy x-ray projection imaging: Two sampling schemes for the correction of scattered radiation", Med. Phys., vol. 15, pp. 732–748, 1988.

[22] J. P. Ward, M. Lee, J. C. Ye, and M. Unser, "Interior tomography using 1D generalized total variation. Part I: Mathematical foundation", SIAM J. Imaging Sci., vol. 8, pp. 226–247, 2015.

[23] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Išgum, "Generative adversarial networks for noise reduction in low-dose CT", IEEE Trans. Med. Imaging, vol. 36, pp. 2536–2545, 2017.

[24] R. Ning, X. Tang, and D. Conover, "X-ray scatter correction algorithm for cone beam CT imaging", Med. Phys., vol. 31, pp. 1195–1202, 2004.

[25] T. Niu, M. Sun, J. Star-Lack,H. Gao, Q. Fan, and L. Zhu, "Shading correction for on-board cone-beam CT in radiation therapy using planning MDCT images", Med. Phys., vol. 37, pp. 5395–5406, 2010.

[26] H.S. Park, J. Baek, S.K. You, J.K. Choi, and J.K. Seo, "Unpaired image denoising using a generative adversarial network in X-ray CT", IEEE Access, vol. 7, pp. 110414–110425, 2019.

[27] H. S. Park, K. Kim, and J. Jeon, "Low-dose CT Image Reconstruction with a Deep Learning Prior", IEEE Access, vol. 8, pp. 158647 – 158655, 2020.

[28] G. Peyré. "Image processing with nonlocal spectral bases", Multiscale Model. Sim., vol. 7, pp.703–730, 2008.

[29] G. Peyré, "Manifold models for signals and images", Comput. Vis. Image Und., vol. 113, pp.249–260, 2009.

[30] R. Schilling, and M. A. Geisel, "Assessment of the effective doses from two dental cone beam CT devices", Dentomaxillofac. Radiol., vol. 42, pp. 20120273, 2013.

[31] M. Sun, and J. M. Star-Lack, "Improved scatter correction using adaptive scatter kernel superposition", Phys. Med. Biol., vol. 55, pp. 6695, 2010.

[32] F. C. Wagner, A. Macovski, and D. C. Nishimura, "Dual-energy x-ray projection imaging: Two sampling schemes for the correction of scattered radiation", Med. Phys., vol. 15, pp. 732–748, 1988.

[33] J. P. Ward, M. Lee, J. C. Ye, and M. Unser, "Interior tomography using 1D generalized total variation. Part I: Mathematical foundation", SIAM J. Imaging Sci., vol. 8, pp. 226–247, 2015.

[34] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Išgum, "Generative adversarial networks for noise reduction in low-dose CT", IEEE Trans. Med. Imaging, vol. 36, pp. 2536–2545, 2017.

[35] J. Ye, Y. Han, and E. Cha, "Deep convolutional framelets: a general deep learning framework for inverse problems", SIAM J. Imaging Sci., vol. 11, pp. 991–1048, 2018.

[36] L. Zhu, Y. Xie, J. Wang, and L. Xing, "Scatter correction for cone-beam CT in radiation therapy", Med. Phys., vol. 36, pp. 2258–2268, 2009.

[37] J.Zhu, T.Park, P.Isola, and A.A.Efros,"Unpaired image-to-image translation using cycle-consistent adversarial networks", 2017, [Online] Available: https://arxiv.org/abs/1703.10593.