

PAPER • OPEN ACCESS

Deep learning method for reducing metal artifacts in dental cone-beam CT using supplementary information from intra-oral scan

To cite this article: Chang Min Hyun *et al* 2022 *Phys. Med. Biol.* **67** 175007

View the [article online](#) for updates and enhancements.

You may also like

- [CLIMAR: classified linear interpolation based metal artifact reduction for severe metal artifact reduction in x-ray CT imaging](#)
Huisu Yoon, Kyoung-Yong Lee and Ibrahim Bechwati
- [A new dental CBCT metal artifact reduction method based on a dual-domain processing framework](#)
Hui Tang, Yu Bing Lin, Su Dong Jiang et al.
- [C-arm orbits for metal artifact avoidance \(MAA\) in cone-beam CT](#)
P Wu, N Sheth, A Sisniega et al.



PAPER

OPEN ACCESS

RECEIVED
20 May 2022REVISED
25 July 2022ACCEPTED FOR PUBLICATION
9 August 2022PUBLISHED
25 August 2022

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Deep learning method for reducing metal artifacts in dental cone-beam CT using supplementary information from intra-oral scan

Chang Min Hyun¹ , Taigyntuya Bayaraa¹, Hye Sun Yun¹, Tae-Jun Jang¹, Hyoung Suk Park^{2,*}  and Jin Keun Seo¹

¹ School of Mathematics and Computing (Computational Science and Engineering), Yonsei University, Seoul, 03722, Republic of Korea

² National Institute for Mathematical Sciences, Daejeon, 34047, Republic of Korea

* Author to whom any correspondence should be addressed.

E-mail: hspark@nims.re.kr

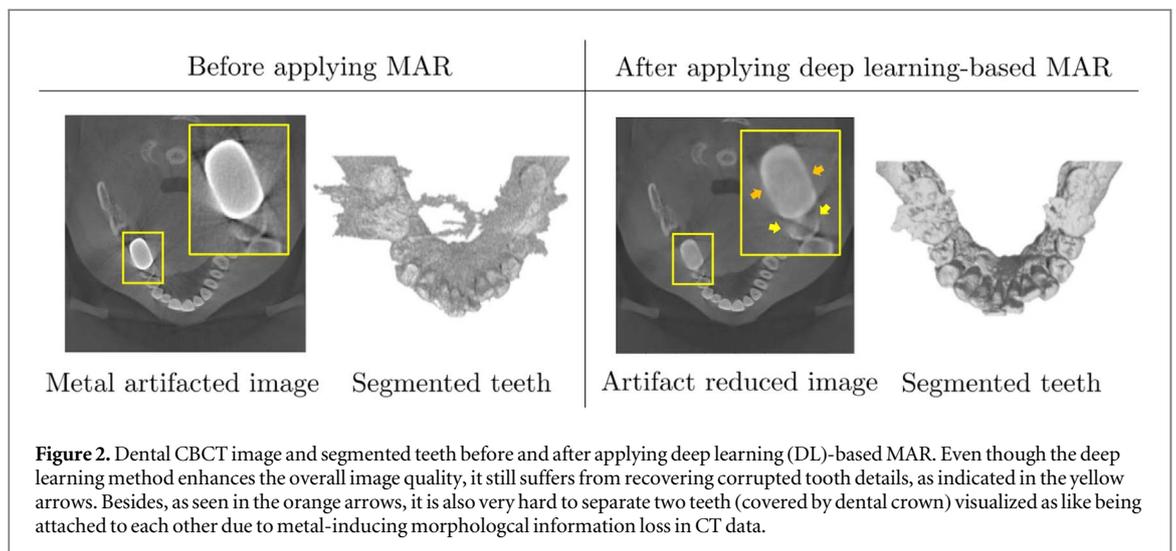
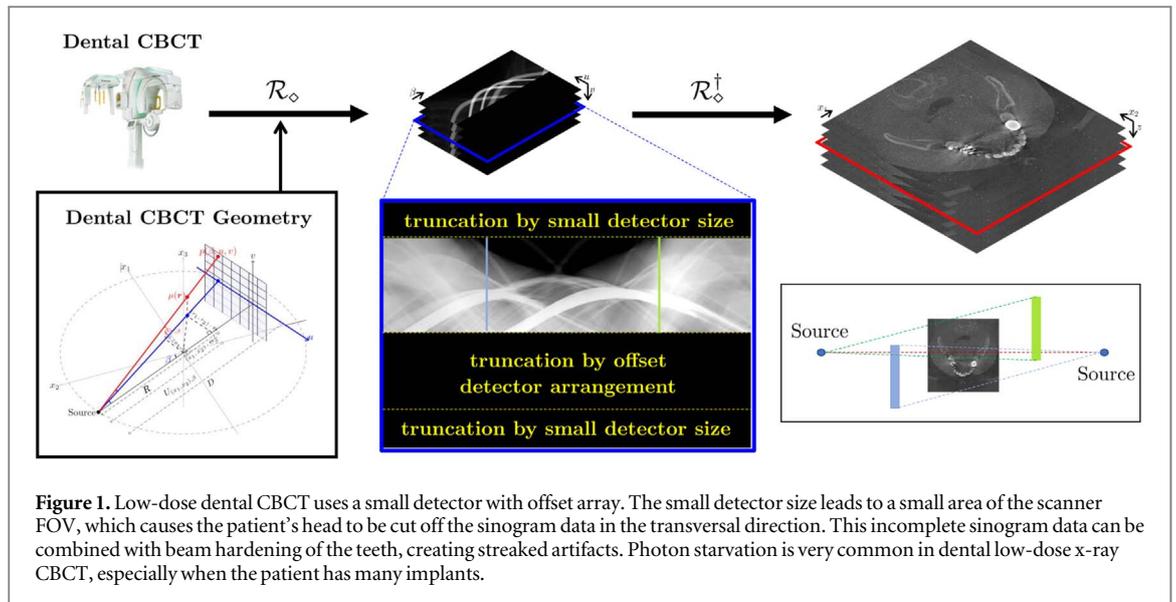
Keywords: cone beam computed tomography, metal artifact reduction, deep learning, digital dentistry, intra-oral scan

Abstract

Objective. Recently, dental cone-beam computed tomography (CBCT) methods have been improved to significantly reduce radiation dose while maintaining image resolution with minimal equipment cost. In low-dose CBCT environments, metallic inserts such as implants, crowns, and dental fillings cause severe artifacts, which result in a significant loss of morphological structures of teeth in reconstructed images. Such metal artifacts prevent accurate 3D bone-teeth-jaw modeling for diagnosis and treatment planning. However, the performance of existing metal artifact reduction (MAR) methods in handling the loss of the morphological structures of teeth in reconstructed CT images remains relatively limited. In this study, we developed an innovative MAR method to achieve optimal restoration of anatomical details. **Approach.** The proposed MAR approach is based on a two-stage deep learning-based method. In the first stage, we employ a deep learning network that utilizes intra-oral scan data as side-inputs and performs multi-task learning of auxiliary tooth segmentation. The network is designed to improve the learning ability of capturing teeth-related features effectively while mitigating metal artifacts. In the second stage, a 3D bone-teeth-jaw model is constructed with weighted thresholding, where the weighting region is determined depending on the geometry of the intra-oral scan data. **Main results.** The results of numerical simulations and clinical experiments are presented to demonstrate the feasibility of the proposed approach. **Significance.** We propose for the first time a MAR method using radiation-free intra-oral scan data as supplemental information on the tooth morphological structures of teeth, which is designed to perform accurate 3D bone-teeth-jaw modeling in low-dose CBCT environments.

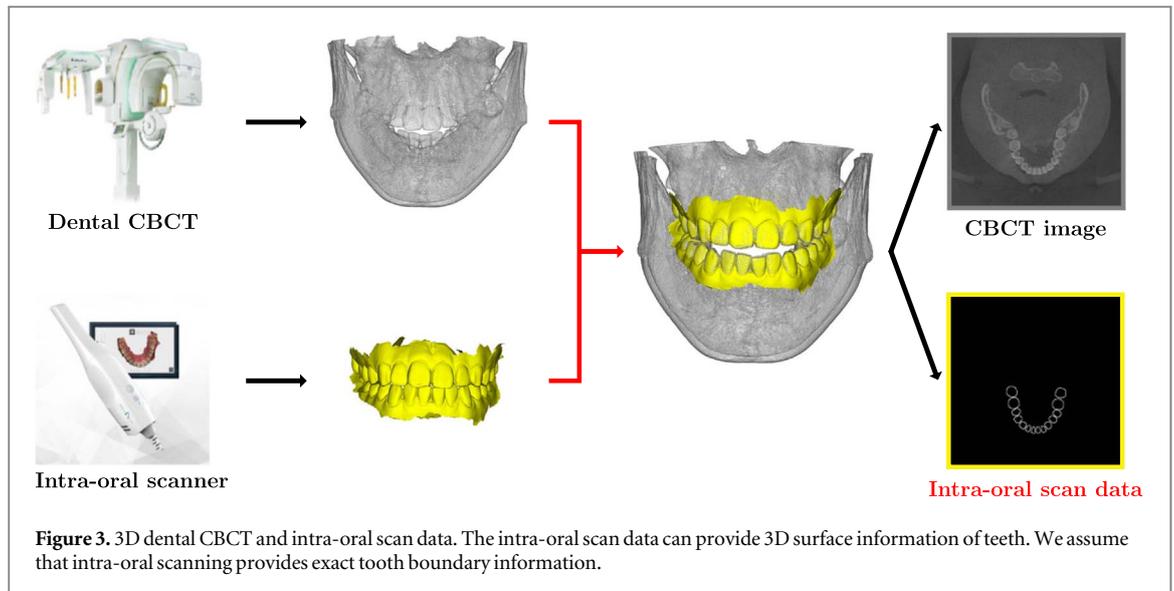
1. Introduction

Dental cone-beam computed tomography (CBCT) has been developed to significantly reduce radiation dose while maintaining image resolution with minimal equipment cost and is increasingly being used in several dental applications such as implant planning and, dental and maxillofacial surgery (Sukovic 2003, Marchetti *et al* 2007, Miracle and Mukherji 2009, Swennen *et al* 2009, 2009, Gupta and Ali 2013, Scarfe *et al* 2017, Weiss and Read-Fuller 2019). Currently, the removal of metallic object-related artifacts poses a major challenge in low-dose dental CBCT. Artifacts related to metallic objects result in severe streaking and shadowing artifacts, which cause a significant loss of the morphological structures of teeth in CT images. Consequently, such artifacts interfere with 3D bone-teeth-jaw modeling for planning diagnosis and treatment in clinical practice (Santler *et al* 1998, Gateno *et al* 2007, Schulze *et al* 2011, Nardi *et al* 2015). As the number of people with metallic oral appliances such as implants, crowns, and dental fillings continues to increase, metal artifacts have been common and their reduction has drawn increased attention (Draenert *et al* 2007, Sanders *et al* 2007, Razavi *et al* 2010, Schulze *et al* 2010, Esmaeili *et al* 2012, Pauwels *et al* 2013, Sancho-Puchades *et al* 2015).



Metal artifact reduction (MAR) is a very difficult problem because the generation of metal-induced streaking and shadowing artifacts is intricately intertwined with interactions between metal, bone, and tissue involving various factors such as beam hardening, scattering, nonlinear partial volume effects, photon starvation, and a high degree of attenuation inhomogeneity (i.e. metal, bone, tissue, air) (Schulze *et al* 2011, Gjestebj *et al* 2016). MAR is much more challenging in low-dose dental CBCT environment owing to offset detection, truncation of the field of view (FOV), low radiation dose, and 3D characteristics in image reconstruction (Bayaraa *et al* 2020). See figure 1. Moreover, when multiple and strong metallic inserts occupy a significant area, low radiation doses often cause photon starvation along metal traces. These result in severe loss or disruption of tooth structures around the inserts in the reconstructed image.

Numerous MAR methods have been developed, including dual-energy approaches (Alvarez and Macovski 1976, Lehmann *et al* 1981, Yu *et al* 2012), statistical iterative correction (De Man *et al* 2001, Elbakri and Fessler 2002, Williamson *et al* 2002, Menvielle *et al* 2005, O'Sullivan and Benac 2007), sinogram inpainting-based correction (Kalender *et al* 1987, Bazalova *et al* 2007, Abdoli *et al* 2010, Meyer *et al* 2010, Park *et al* 2013), and deep learning methods (Park *et al* 2018, Zhang and Yu 2018, Gjestebj *et al* 2019, Lin *et al* 2019, Yu *et al* 2020). Although these methods have been shown to mitigate metal-induced artifacts, their performance in dental applications remain unsatisfactory, and they involve limitations in low-dose dental CBCT environments. Figure 2 highlights that the improvement of corrupted and missing details remains an arduous task even with state-of-art deep learning methods. There seems to be a fundamental limitation in accurately restoring the morphological structure of teeth using only sinogram data severely damaged by metal inserts.



In this study, we propose a deep learning-based MAR method using radiation-free intra-oral scan data as supplemental information for tooth morphological structure, as shown in figure 3. The proposed MAR method is a two-stage approach, which is designed to perform accurate 3D bone-teeth-jaw modeling. In the first stage, we employ a deep learning network that utilizes intra-oral scan data as side-inputs and performs multi-task learning of auxiliary tooth segmentation. The network is designed to improve the learning ability of capturing teeth-related features effectively while mitigating metal artifacts. The suitable incorporation of explicit shape-prior information from intra-oral scan data with deep learning models can provide significant benefits in terms of accuracy, learnability, feature extraction, and so forth (Sun *et al* 2019, Liu *et al* 2021). In the second stage, a 3D bone-teeth-jaw model is constructed with weighted thresholding, where the weighting region is determined depending on the geometry of the intra-oral scan data. We also adopted a simulation approach to train the proposed deep learning network. For each metal-free CBCT scan, the corresponding metal-free CBCT and intra-oral scans are generated using a self-developed data generation tool that does not involve any time-consuming and labor-intensive manual processes.

We conducted numerical simulations and clinical experiments to investigate the potential impact of the use of intra-oral scan data in MAR and bone-teeth-jaw modeling. The results of the experiments demonstrate the feasibility of the proposed method and the benefits of using intra-oral scan data in low-dose dental CBCT environments.

2. Method

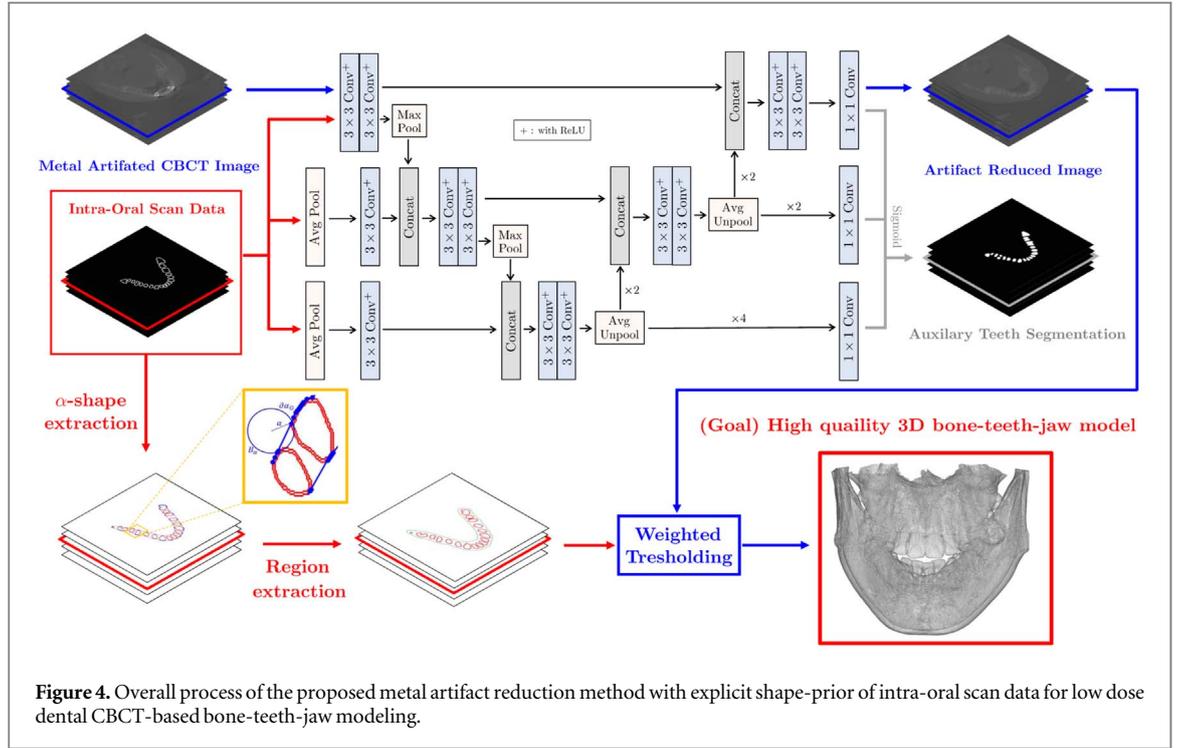
In low-dose dental CBCT, the measured sinogram data \mathbf{P} can be expressed as

$$\mathbf{P} = \mathcal{T}(-\ln \int_E \eta(E) \exp(-\mathcal{R}_\diamond \mu_E) dE + \mathbf{n}). \quad (1)$$

Here, μ_E is the attenuation coefficient distribution of a 3D human body to be scanned at an energy E , η is the normalized energy distribution of the x-ray source, \mathcal{R}_\diamond is a cone beam projection, \mathbf{n} is the CT noise, and \mathcal{T} is truncation caused by the size and arrangement of the detector (typically, small and offset). See figure 1. In the presence of metallic objects inside the FOV, the standard FDK algorithm (Feldkamp *et al* 1984) produces severe streaking and shadowing artifacts that cause the image quality of maxillofacial structures to deteriorate. Hence, high-quality 3D bone-teeth-jaw modeling is arduous only with the image.

The goal of the proposed method is to provide a high-quality 3D bone-teeth-jaw (or maxillofacial) model from metal-affected sinogram data \mathbf{P} by leveraging intra-oral scan data \mathbf{O} . The output should be competitive with a 'gold-standard' bone-teeth-jaw model \mathbf{y}_{mf} acquired from an artifact-free CT image \mathbf{y} that is reconstructed by \mathbf{P}_* , where \mathbf{P}_* represents the artifact-free sinogram data corresponding to \mathbf{P} . The intra-oral scan data \mathbf{O} provide a 3D tooth surfaces that can be useful as prior information about tooth geometry. It is assumed that intra-oral scan data \mathbf{O} provides exact tooth boundary information.

The proposed method is based on the image-to-image learning approach and weighted thresholding that leverages intra-oral scan data as explicit shape prior information of tooth geometry for MAR. The reconstruction map f can be expressed as



$$f = f_{\alpha\text{-WT}} \circ f_{\text{IE}} \circ \mathcal{R}_{\text{FDK}}^{\dagger}, \quad (2)$$

where

- $\mathcal{R}_{\text{FDK}}^{\dagger}$ is the weighted FDK algorithm involving the sinogram extrapolation method for addressing offset detector arrangement and FOV truncation.
- f_{IE} is the tooth geometry prior information-based-image-enhancing network f_{IE} , which mitigates metal-related artifacts.
- $f_{\alpha\text{-WT}}$ is a weighted thresholding, wherein the weighting region is determined in basis of the α -shape from intra-oral scan data. This procedure is used for further removing the remaining streaking artifacts around the teeth in constructing a bone-teeth-jaw model.

Here, the input of f is a pair of metal-affected data \mathbf{P} and intra-oral scan data \mathbf{O} (i.e. $f: (\mathbf{P}, \mathbf{O}) \mapsto f(\mathbf{P}, \mathbf{O}) \approx \mathbf{y}_{\text{mt}}$). The overall process is illustrated in figure 4.

Stage 1. Image-enhancing network f_{IE}

In our experience, an image domain-learning-based approach can mitigate metal-related artifacts effectively, whereas it tends to have weakness in recovering tooth shape, especially when being destroyed by severe artifacts or when being missing. To compensate for this weakness, we attempt to take advantage of supplemental shape information from intra-oral scan data. We emphasize that data acquisition by the intra-oral scanner does not increase the total amount of radiation exposure to a patient.

Let \mathbf{x} be a 3D CBCT image reconstructed using the FDK algorithm (i.e. $\mathbf{x} = \mathcal{R}_{\text{FDK}}^{-1}(\mathbf{P})$). The image-enhancing network f_{IE} aims to provide

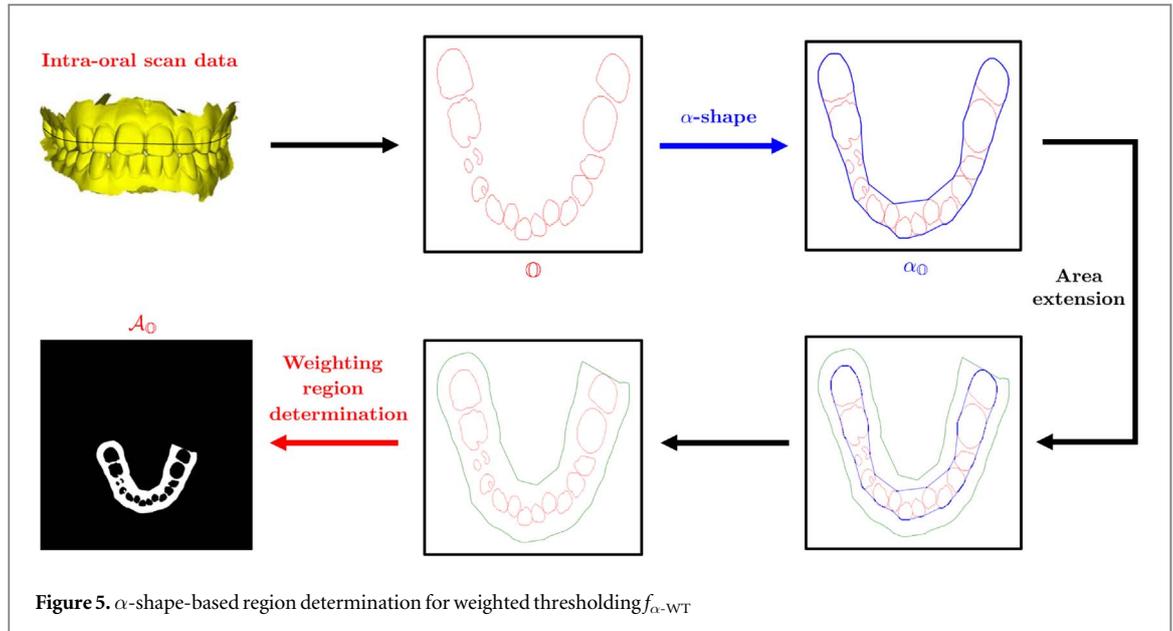
$$f_{\text{IE}}(\mathbf{x}_j, \mathbf{O}_j) \approx \mathbf{y}_j, \quad (3)$$

where \mathbf{y}_j is the j th slice of a metal-artifact-free image (i.e. $\mathbf{y} = \mathcal{R}_{\text{FDK}}^{-1}(\mathbf{P}_*)$). It is also desirable that the output satisfies

$$\partial f_{\text{IE}}(\mathbf{x}_j, \mathbf{O}_j)|_{\text{teeth}} \approx \mathbf{O}_j, \quad (4)$$

where $\partial f_{\text{IE}}(\mathbf{x}_j, \mathbf{O}_j)|_{\text{teeth}}$ is a binary mask of the tooth surface region on the output image $f_{\text{IE}}(\mathbf{x}_j, \mathbf{O}_j)$.

To accomplish these goals, two strategies are adopted; side-input layer and multi-task learning. First, additional information of intra-oral scan data is repeatedly enriched during feature extraction in an encoding path. These side inputs can help the network extract tooth shape while compensating for missing or severely distorted structures through high quality shape information provided by intra-oral scan data. Second, multi-task learning is applied, which learns image reconstruction and auxiliary tooth segmentation in a parallel fashion. In the medical imaging field, it has been reported that deep learning-based image reconstruction ability



can be boosted by learning other image-related tasks, such as segmentation and registration (Sun *et al* 2019, Liu *et al* 2021). In terms of image recovery, the auxiliary tooth segmentation is expected to reveal the shapes of the teeth in the decoding path and the interference of tooth features, which are joint domain information of the interrelated tasks, through the shared parameters.

Figure 4 shows the overall procedure of the proposed image-enhancing network f_{IE} . Inspired by M-net (Mehta and Sivaswamy 2017), the proposed network has side-input and side-output layers. In the side-input layers, intra-oral scan data \mathbf{O} with suitable resizing is repeatedly added to the encoding path after 3×3 convolution. In the side-output layers, tooth segmentation masks are obtained during the decoding path. The detailed backbone structure can be found in Mehta and Sivaswamy (2017).

When $(s_0)_j^{(i)}$ is the final network output of i th training data and j th slice (i.e. $(s_0)_j^{(i)} = f_{\text{IE}}(\mathbf{x}_j^{(i)}, \mathbf{O}_j^{(i)})$), the network f_{IE} is trained as follows.

$$\begin{aligned} \operatorname{argmin}_f \sum_i \mathcal{L}_{\ell_2}(((s_0)_j^{(i)})^{(1)}, \mathbf{y}_j^{(i)}) \\ + \mathcal{L}_{ce}(((s_0)_j^{(i)})^{(2)}, \mathbf{S}_j^{(i)}) + \sum_{k=1}^2 \mathcal{L}_{ce}((s_k)_j^{(i)}, \mathbf{S}_j^{(i)}), \end{aligned} \quad (5)$$

where $((s_0)_j^{(i)})^{(1)}$ and $((s_0)_j^{(i)})^{(2)}$, respectively, denote the first and second channel outputs of $(s_0)_j^{(i)}$, $\mathbf{S}_j^{(i)}$ denotes a binary segmentation mask of the reference tooth region in the metal-artifact-free image $\mathbf{y}_j^{(i)}$, $\{(s_k)_j^{(i)}\}_{k=1}^2$ is a set of side outputs in the decoding path, \mathcal{L}_{ℓ_2} is the standard ℓ_2 loss, and \mathcal{L}_{ce} is the cross-entropy loss. For convenience, the notation $f_{\text{IE}}(\mathbf{x}, \mathbf{O})$ is used to represent the output image (i.e. the first channel output).

Stage 2. α -shape-based weighted thresholding $f_{\alpha\text{-WT}}$

The next step is bone-teeth-jaw modeling from the metal-artifact-reduced CBCT image obtained in the previous stage. A final 3D bone-teeth-jaw model is obtained by weighted thresholding, which can further reduce the remaining streaking artifacts around teeth. The weighting region is determined depending on the geometry of the intra-oral scan data \mathbf{O} . To extract the geometry, the α -shape technique (Edelsbrunner and Mücke 1994) is used. It provides a family of piece-wise linear lines associated with the shape of the teeth. Figure 5 shows the overall process.

When \mathbf{y}_{dl} is $\mathbf{y}_{\text{dl}} = f_{\text{IE}}(\mathbf{x}, \mathbf{O})$, the weight thresholding $f_{\alpha\text{-WT}}$ can be expressed as

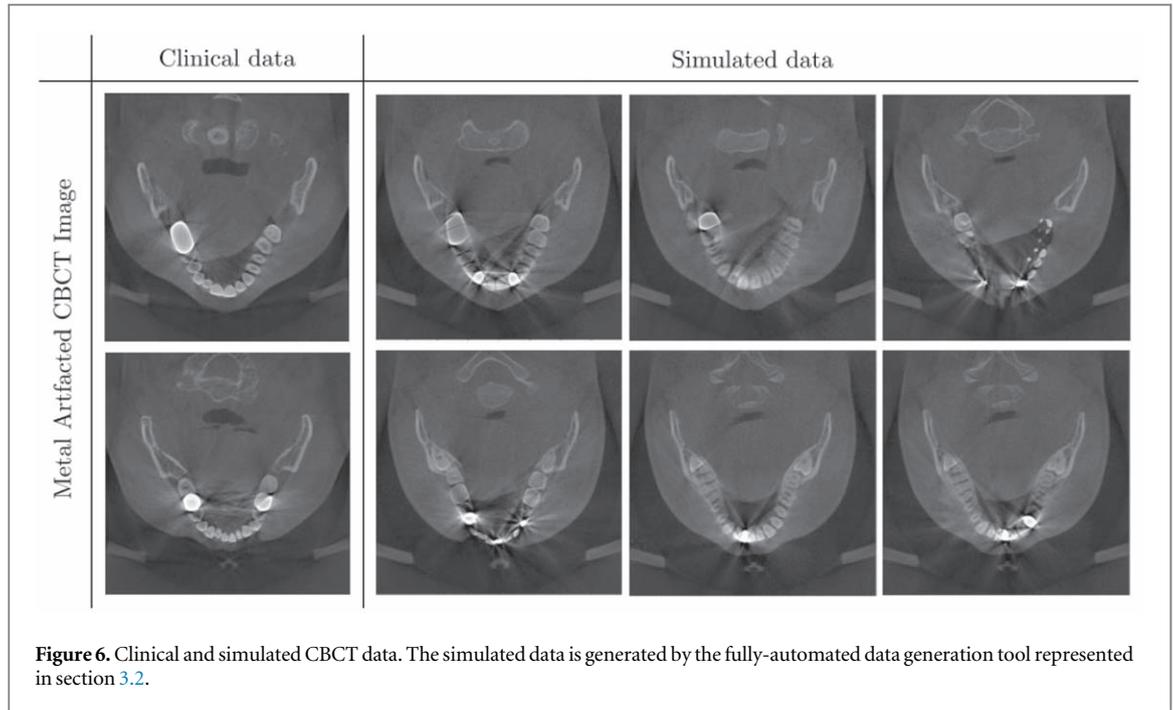
$$f_{\alpha\text{-WT}}^T(\mathbf{y}_{\text{dl}}, \mathbf{O}) = \hat{\mathbf{y}}_{\text{mf}}, \quad (6)$$

where

$$\hat{\mathbf{y}}_{\text{mf}} = \begin{cases} \hat{\mathbf{y}}_{\text{mf}}(p) = 0 & \text{if } p \in \mathcal{A}_O \text{ or } \mathbf{y}_{\text{mf}}(p) < \tau \\ \hat{\mathbf{y}}_{\text{mf}}(p) = 1 & \text{otherwise} \end{cases}. \quad (7)$$

Here, p is a point in a grid of \mathbf{y}_{mf} , τ is a thresholding constant, and \mathcal{A}_O is a thresholding region obtained using the α -shape from \mathbf{O} .

The region \mathcal{A}_O is obtained as follows. For given intra-oral scan data \mathbf{O} , \mathbb{O} is a point cloud corresponding to \mathbf{O} . Denoted by $\alpha_{\mathbb{O}}$, an α -shape of \mathbb{O} is given by a polytope with a boundary $\partial\alpha_{\mathbb{O}}$, which is defined by



$$\partial\alpha_{\mathcal{O}} = \{ \Delta_{\mathbb{T}} \mid \mathbb{T} \subset \mathcal{O}, |\mathbb{T}| \leq 3, \Delta_{\mathbb{T}} \text{ is } \alpha\text{-exposed} \}, \quad (8)$$

where $\Delta_{\mathbb{T}}$ denotes a simplex for \mathbb{T} , and $\Delta_{\mathbb{T}}$ is α -exposed if and only if there exists an open ball B_{α} with radius α such that $B_{\alpha} \cap \mathcal{O} = \emptyset$ and $\partial B_{\alpha} \cap \mathcal{O} = \mathbb{T}$. Here, ∂B_{α} is a boundary of B_{α} . After the α -shape is obtained, an extension direction on each vertex of $\alpha_{\mathcal{O}}$ is defined by taking the average of the normal vectors on the faces that contain the vertex. Along the direction, $\alpha_{\mathcal{O}}$ is extended while preserving its shape and converted into a binary mask $\alpha_{\mathcal{O}}$, where the inner regions of the shape boundary are filled with one. Finally, the region $\mathcal{A}_{\mathcal{O}}$ is determined by

$$\mathcal{A}_{\mathcal{O}} = \alpha_{\mathcal{O}} - \bar{\mathcal{O}}, \quad (9)$$

where $\bar{\mathcal{O}}$ is the binary mask where the inner part of tooth surfaces in \mathcal{O} is filled with one.

3. Experiment and result

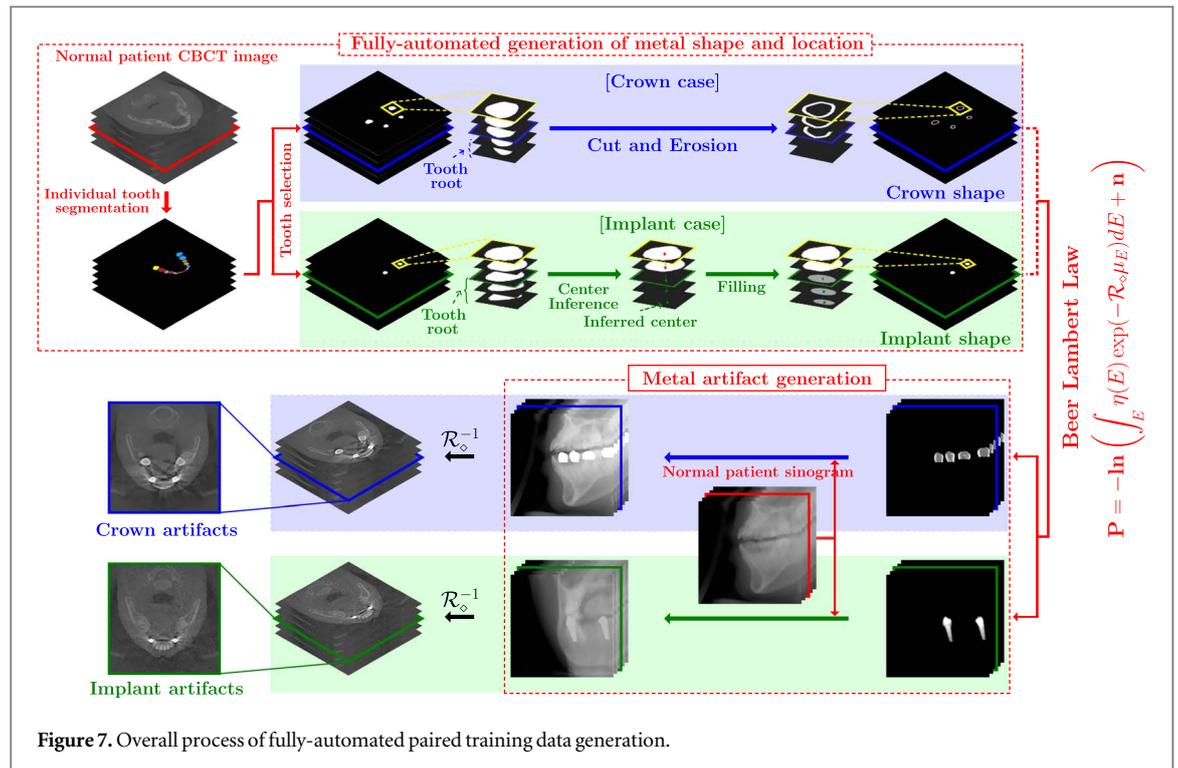
3.1. Experiment setting

The sinogram data of a real patient were obtained from a commercial CBCT machine (Q-FACE, HDXWILL). The voxel size was $1200 \times 654 \times 658$ with real scale of 0.2 mm for each axis, where 1200 is the number of uniformly sampled projection views in $[0, 2\pi)$, and 654×658 is the number of samples measured by the 2D flat detector for each projection view. CBCT images were reconstructed in a voxel size of $800 \times 800 \times 400$ with a real scale of 0.2 mm. For cone beam projection, an open-source code, known as TIGRE (Biguri *et al* 2016), was used, where the projection algorithm was implemented by a ray-driven method. The scattering was not considered in this study.

All simulated data were consistently generated to have same scale as the real data. A self-developed fully-automated paired data generation tool was used. The detailed process is described in section 3.2. Figure 6 shows several samples of the simulated data using the data generation tool.

Metal-free CBCT sinogram data were collected from 20 patients without any metallic objects. They were used for data generation. Metal-affected CBCT data were collected from nine patients. They were used for test purposes. Among the metal-affected data, real intra-oral scan data for one patient was provided. The intra-oral scan data was acquired from a scanner (i500, MEDIT), where the file format was provided by the standard triangle language (STL). A set of its vertices is a point cloud in millimeters, where the maxilla and mandible are represented by approximately 100 000 and 70 000 points, respectively. For registration into the dental CBCT system, the method described by Jang *et al* (2021) was applied.

In PyTorch environment (Paszke *et al* 2019), all deep learning experiments were conducted with a computer system equipped with two Intel Xeon CPUs E5-2630 v4, 128GB DDR4 RAM, and four NVIDIA GeForce GTX 2080ti GPUs. The optimization was conducted using Adam optimizer (Kingma and Ba 2014) and multi-GPUs. Batch normalization was applied to achieve fast convergence and minimization (Ioffe and Szegedy 2015). The network capacity (i.e. feature and network depths) was minimized as much as possible while maintaining the



backbone structure because of the huge computational cost associated with the CBCT image size of $800 \times 800 \times 400$.

For α -shape implementation, open source packages, Visualization ToolKit (VTK) and Alpha Shape Toolbox (AST), were used. The adaptive values α and τ were selected empirically.

3.2. Fully-automated paired data generation

We generated a realistic paired training dataset for MAR through the following procedure, which do not involve any time-consuming and labor-intensive manual process (see figure 7 for overall workflow). As a first step, fully-automated individual tooth segmentation was performed on metal-free CBCT data by using the technique reported by Jang *et al* (2021). Several tooth positions were chosen randomly in which virtual metal implants could be placed. For the crown case, a crown mask was constructed by cutting the roots of chosen teeth based on crown height information for each tooth (Nelson 2014), and then by the erosion process. The crown thickness was randomly set from 0.6 to 1.4 mm. For an implant case, instead of erosion, another process was applied to create an implant screw bar. A line was defined for each tooth that passed through two points of the tooth center in the lowest and middle slices, except those containing a tooth root. Then, the root parts were filled with circles whose center was located at the line, and the radius was empirically set. Using the generated dental crown or tooth implant mask, metal-affected sinogram data was artificially synthesized using the Beer–Lambert law (1) and combined with metal-free sinogram.

The simulation projection data was generated at a tube voltage of 85 keV. A metal attenuation coefficient was randomly assigned from {Au, Pd, Ni, Cr, Zr, Al}. For the numerical simulation, the energy distribution of the x-ray source and attenuation coefficient values were those described elsewhere (Hubbell and Seltzer 1995, Mahesh 2013). Poisson and Gaussian noise were added to take account of the CT noise. A total of 20 metal-free scans were split into two disjoint sets (i.e. 15 and 5 scans) and used for training and testing, respectively. There is no common ground-truth (i.e. metal-free scan) between the two sets. From 15 scans, total 60 paired data (4 data from each scan) were generated and only used for training purpose. From 5 scans, 10 test data (2 data from each scan) were generated. Here, the number of inserted metal implants was randomly set from two to five.

The intra-oral scan data was simulated as a boundary mask of teeth with inserted metal objects. The boundary mask was obtained by applying the erosion process to the segmented teeth and inserted metal objects.

3.3. Experimental results

To investigate the advantages of the proposed network, performance comparisons were conducted with various MAR methods. The experiments were based on three test sets: synthesized CBCT data + simulated intra-oral scan data, clinical CBCT data + simulated intra-oral scan data, and clinical CBCT data + real intra-oral scan data. Qualitative and quantitative evaluations were conducted on the synthesized CBCT dataset in which the

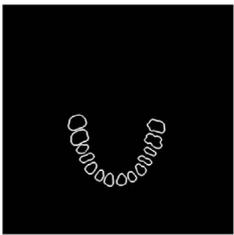
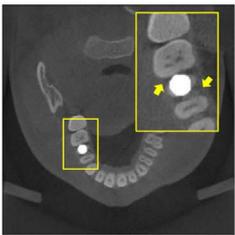
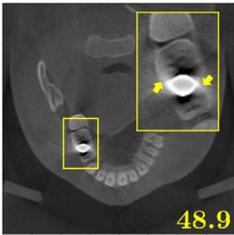
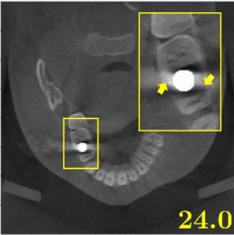
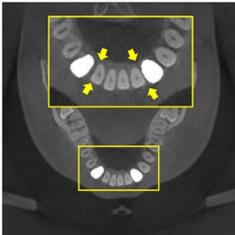
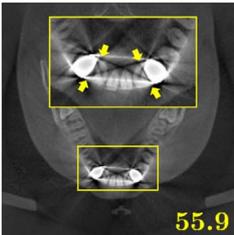
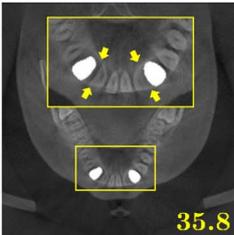
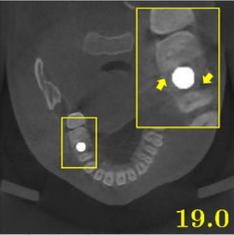
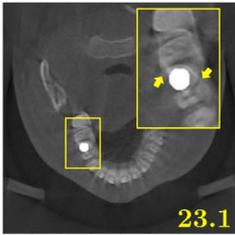
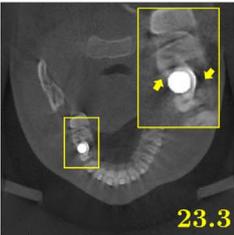
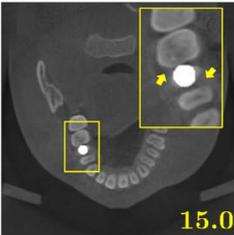
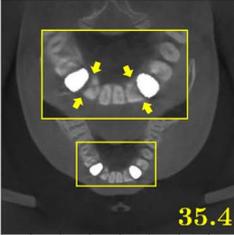
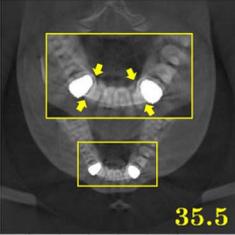
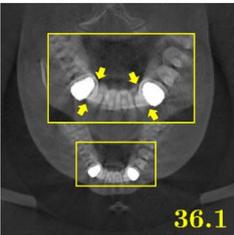
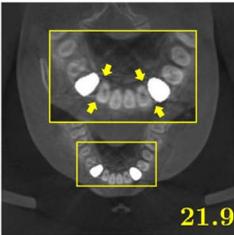
	Intra-Oral Scan	Ground Truth	Uncorrected	LI
Case 1			 48.9	 24.0
Case 2			 55.9	 35.8
	Img DL	Sino DL	Sino Inpaint DL	Proposed Network
Case 1	 19.0	 23.1	 23.3	 15.0
Case 2	 35.4	 35.5	 36.1	 21.9

Figure 8. Comparison of metal artifact reduction over simulated data with various MAR methods; linear interpolation (LI), image domain learning (Img DL), sinogram domain learning (Sino DL), sinogram inpainting learning (Sino Inpaint DL), and the proposed network. Case 1 is the best MAR case and Case 2 is the worst MAR case. The Hausdorff distance between tooth boundary segmented manually from a CT image and the corresponding intra-oral scan data is provided as a yellow value. The distance was computed in the region of a yellow box.

corresponding ground-truth images are given. For clinical CBCT data, qualitative evaluations were performed. For a quantitative comparison of tooth shape restoration near the metallic objects, we computed the Hausdorff distance (Huttenlocher *et al* 1993) between the tooth boundary segmented manually from a CT image and the corresponding intra-oral scan data. Here, the Hausdorff distance was computed on a region around the metal.

It should be mentioned that comparison with other methods is unfair, because the proposed method takes advantage of additional information from intra-oral scan data.

3.3.1. Test on synthesized CBCT and simulated intra-oral scan data

Figure 8 and table 1 show qualitative and quantitative performance comparisons of the proposed network with linear interpolation, an image domain network, a sinogram domain network, and a sinogram inpainting network. For the linear interpolation, the sinogram reflection technique reported by Bayaraa *et al* (2020) was applied to deal with metal trace truncation. Image thresholding was used to extract metal traces. For the image domain network, U-net (Ronneberger *et al* 2015) was trained, which directly maps from an uncorrected image to the corresponding ground truth image. For the sinogram domain network, U-net was trained, which directly maps from an uncorrected sinogram to the corresponding ground truth sinogram. For the sinogram inpainting network, U-net was trained such that only the metal traces in the sinogram were corrected by a network output.

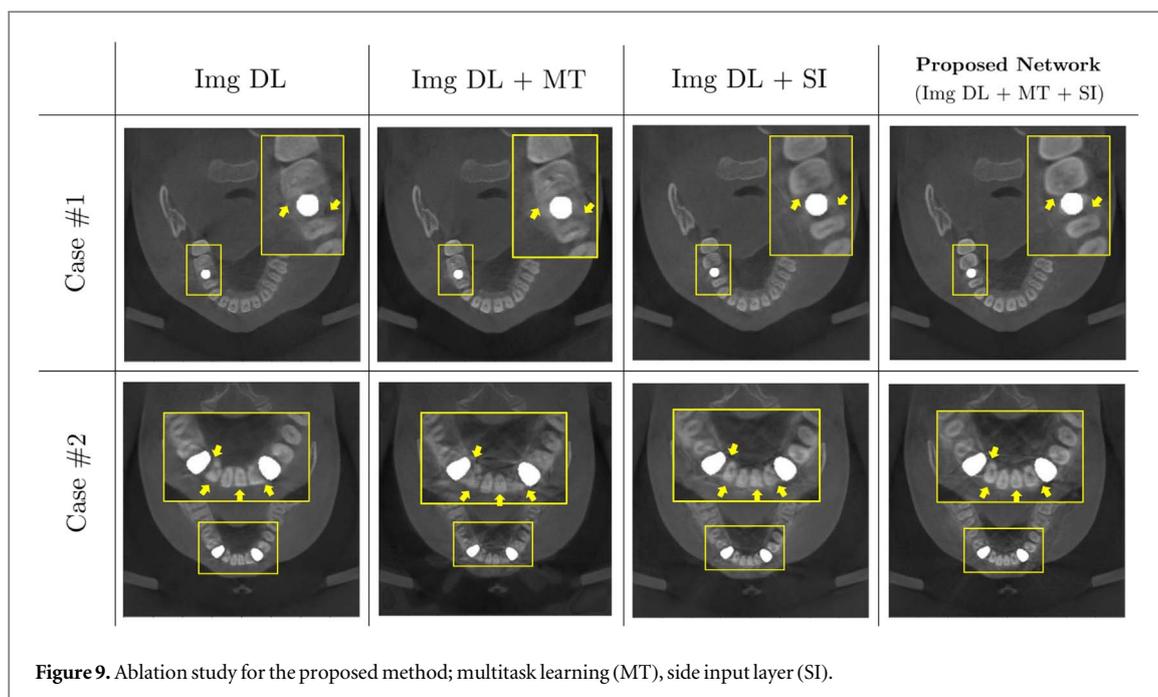


Figure 9. Ablation study for the proposed method; multitask learning (MT), side input layer (SI).

Table 1. Quantitative comparison of deep learning-based MAR results for simulated patient data in terms of normalized mean square error (NMSE), structural similarity index (SSIM), and peak signal-to-noise ratio (PSNR).

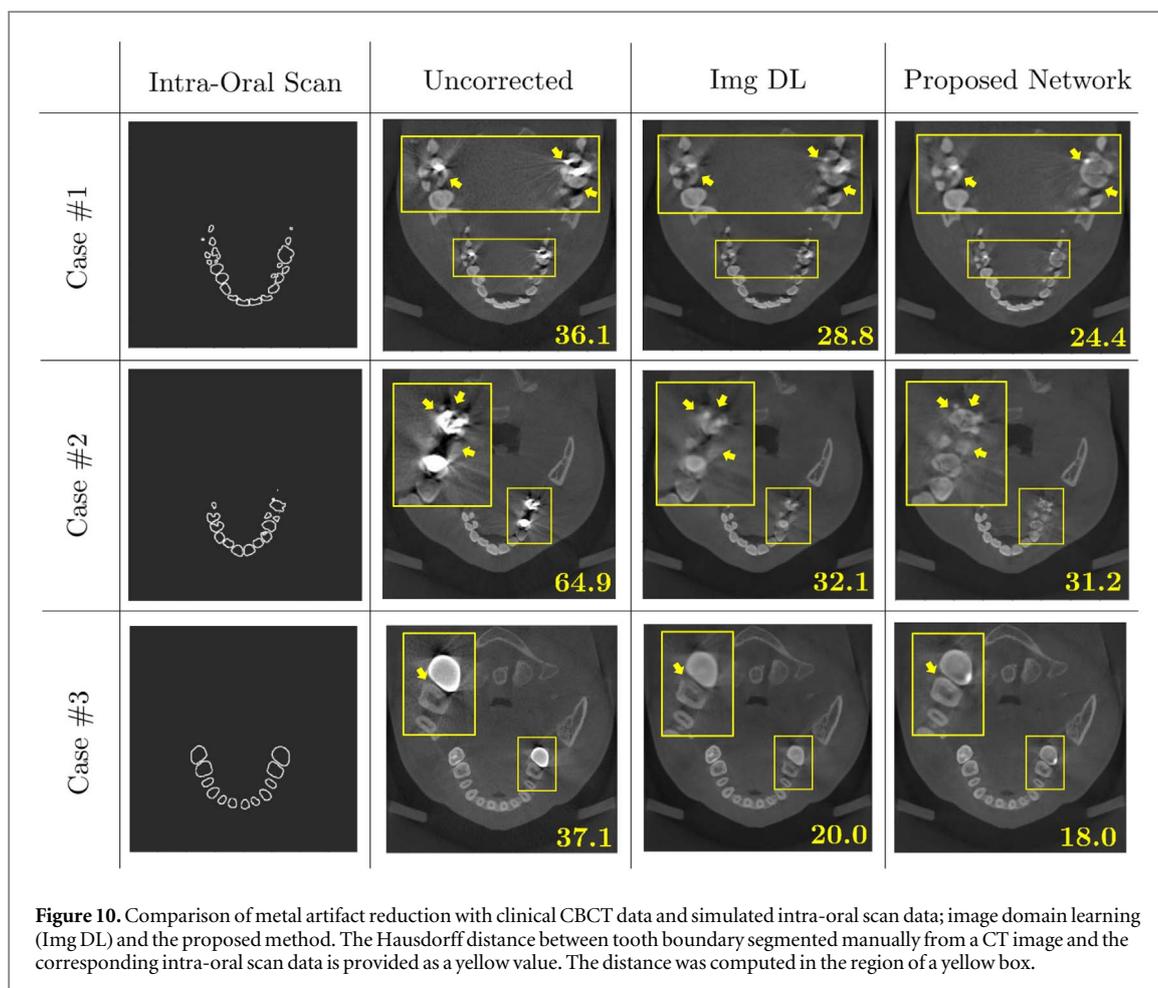
Metric	Uncorrected	LI	Sino DL	Sino Inp DL
NMSE	0.6298	0.4158	0.5445	0.7017
SSIM	0.9908	0.9948	0.9882	0.9846
PSNR	52.32	55.74	53.40	51.20
Metric	Img DL	Img DL+MT	Img DL+SI	Proposed Network
NMSE	0.3577	0.3885	0.3458	0.3421
SSIM	0.9954	0.9924	0.9963	0.9965
PNSR	57.06	56.36	57.35	57.44

In the experiments, the proposed network exhibited the best performance, significantly improving the shape quality of teeth and bone associated with bone-teeth-jaw modeling. In particular, the proposed network appears to have an outstanding ability to recover the tooth shape, even if it is fairly disrupted or missing because of metal-related artifacts. The performance of the proposed method was validated as well via 4-fold cross validation. See [appendix](#) for details.

As shown in figure 9 and table 1, an ablation study for multi-task learning (MT) and side input layer (SI) in the proposed network was conducted qualitatively and quantitatively. The single use of MT did not provide any advantage in the sense of improving the reconstruction ability in the quantitative and qualitative sense. Either SI or a combination of SI and MT enhances the reconstruction performance both qualitatively and quantitatively. The combination of SI and MT appears to provide an optimal result owing to the synergistic effect.

3.3.2. Test on clinical CBCT and simulated intra-oral scan data

Figure 10 shows a comparison of the test set of real clinical CBCT data and simulated intra-oral scan data, where the intra-oral scan data were obtained by tooth segmentation from the clinical CBCT data. Here, the method of Jang *et al* (2021) was utilized, which provides considerably accurate tooth segmentation, even in the presence of metal-related artifacts. Several simulated intra-oral scan data are listed in the first column of figure 10. In three cases from different patients, the proposed network successfully reduced metal artifacts while recovering the boundary of the teeth effectively, whereas the image domain network tended to suffer from loss, blurring, or disruption of the tooth boundary around metal objects.



3.3.3. Test on clinical CBCT and real intra-oral scan data

Figure 11 shows reconstructed results using clinical CBCT and real intra-oral scan data. The proposed method consistently preserves or recovers the boundary of the teeth around metal objects compared with the image domain network. See regions highlighted by yellow arrows in figure 11.

The performance of the proposed method was compared as well when using simulated and real intra-oral scan data for the same clinical CBCT data. There was some performance degradation in the case of real intra-oral scan data relative to the simulated intra-oral scan case. See the region indicated by the orange arrows in figure 11.

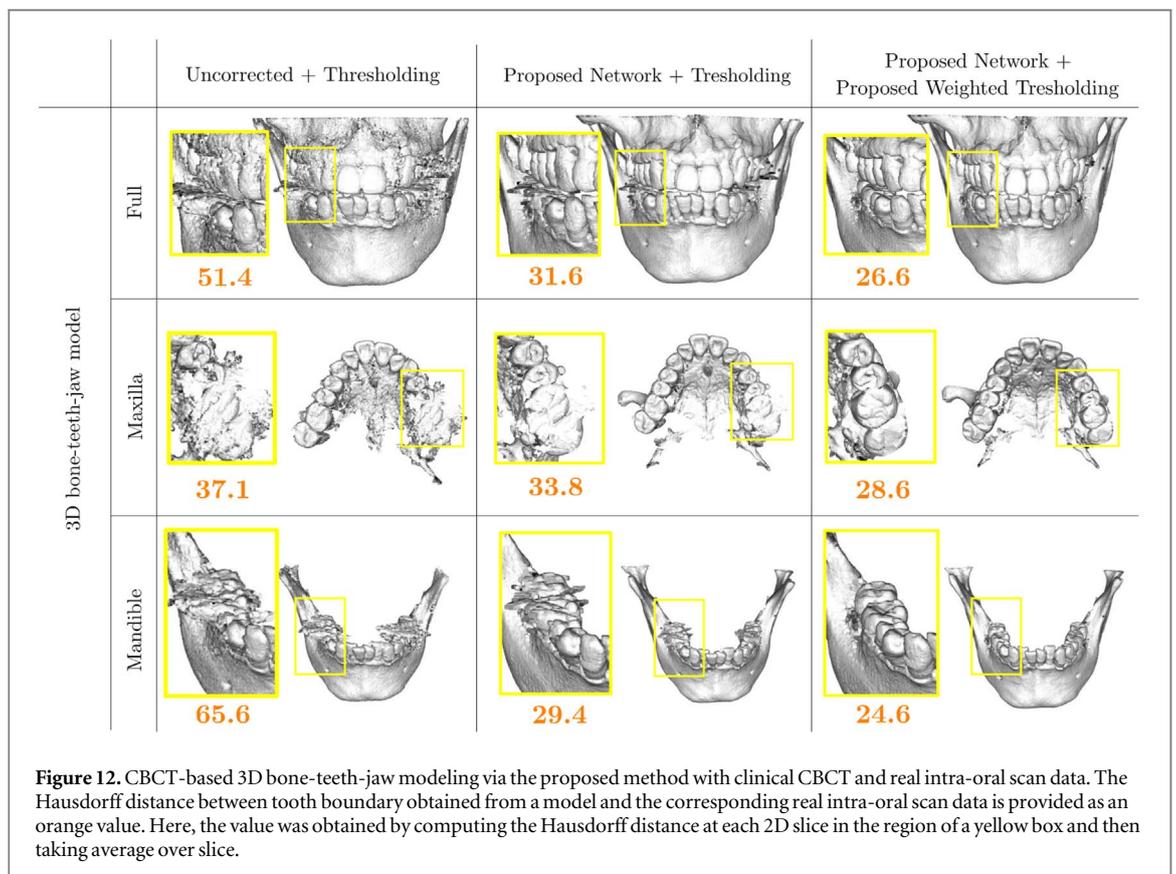
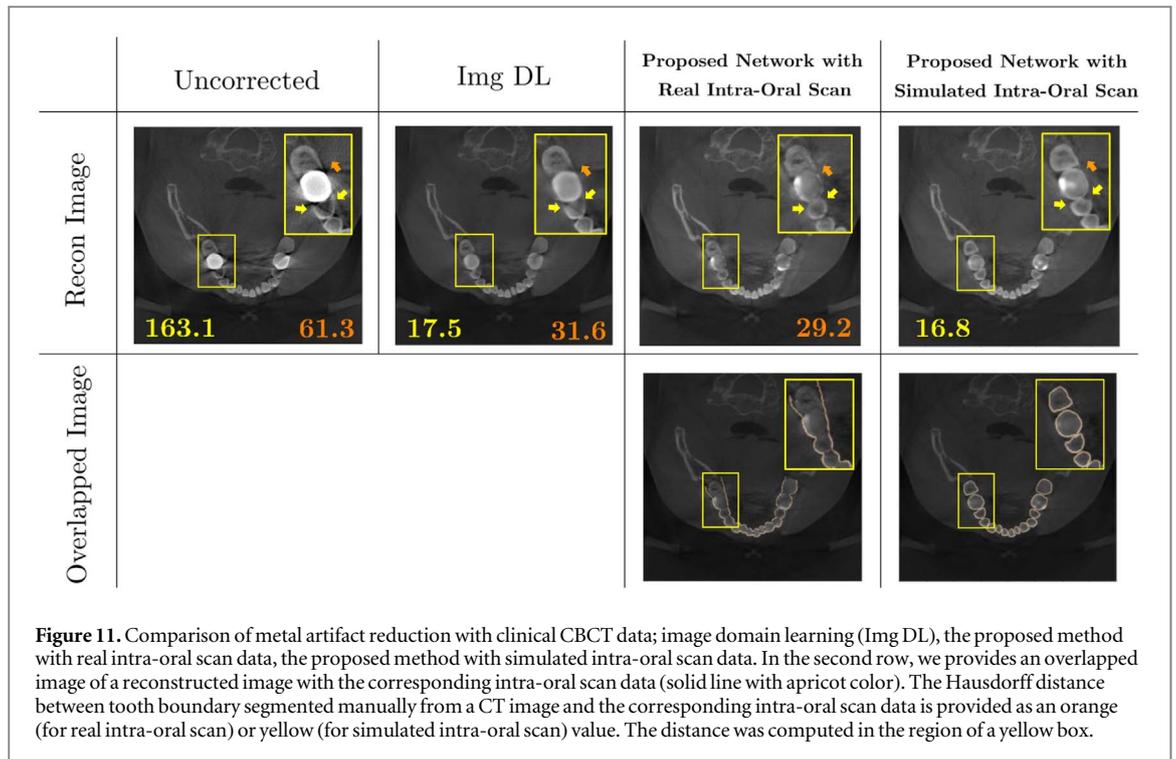
3.3.4. 3D bone-teeth-jaw model construction

Figure 12 shows 3D segmented bone-teeth-jaw models by uncorrected image + image thresholding, the proposed network + image thresholding, and the proposed method (the proposed network + the proposed α -shape-based weighted thresholding). The result was obtained using clinical CBCT data and real intra-oral scan data. The proposed method clearly enhanced the quality of a 3D bone-teeth-jaw model so that it precisely depicted the tooth and bone structures. The α -shape-based weighted thresholding was found to be powerful in real intra-oral scan data for high quality bone-teeth-jaw modeling.

4. Conclusion and discussion

This study is a first attempt to pave the way toward MAR utilizing the shape prior from intra-oral scan data. The utilization of radiation-free intra-oral scan data is meaningful in the trend of that dental CBCT has been being developed toward the direction of minimizing radiation exposure while maintaining diagnostic image quality. Our experiments demonstrated the tremendous potential of the intra-oral scan data to have a significant positive effect on the restoration of tooth shape loss by metal-related artifacts.

To train the proposed network, a paired dataset of metal-artifacted data, metal-artifact-free data, and intra-oral scan data is required, but data accessibility is limited in clinical practice. Hence, the data generation tool was utilized to provide a realistic paired dataset, where the intra-oral scan data for training was simulated as a set of boundaries of individual teeth segmented in an artifact-free CBCT image. However, the simulation did not fully



reflect the real scanning environment, such as scanning protocol, condition, and performance. The difference between the training and test domains brought the performance degradation. The performance of the proposed network on real intra-oral scan data can be improved if more realistic simulated intra-oral scan data or a sufficient number of real oral scan data for training can be obtained (Hyun *et al* 2021).

Intra-oral scan is very accurate for small area scans, but its accuracy gradually decreases as the scan moves away from the start of the scan due to cumulative stitching errors (Nagy *et al* 2020). Recent advances in intra-oral

scan technology have significantly reduced stitching errors in full-arch description (Winkler and Gkantidis 2020). Specifically, an average full arch description error of 0.00883 ± 0.01088 mm in *vivo* analysis was reported by Kwon *et al* (2021) for the intra-oral scanner (i500, MEDIT, Seoul, South Korea) used in this study. In our CBCT imaging setup with a spatial resolution of 0.2 mm, the errors can produce variations of 1 pixel or 2 pixels. This error can be effectively addressed using the stitching error correction method proposed by Jang *et al* (2021) to mitigate the possible influence on MAR performance caused by error-related shape variation.

The ability of the proposed MAR method can be further improved through complex network architectures and a large-scale training dataset. However, there is a trade-off with the total computational cost for learning that can be critical, especially in high-dimensional data applications (Hyun *et al* 2020). Even for the simple M-net architecture shown in figure 4, at least 10 d are required for training of 300 epochs with a dataset of 60 image voxels under the computational resources used in this study. Even though the use of sophisticated networks or large training datasets can potentially enhance MAR capability, associated hurdles involving high dimensionality should be addressed for practical dental CBCT applications.

Acknowledgments

This work was supported by Samsung Science & Technology Foundation (No. SRF-C-IT1902-09). H S Park was partially supported by the National Institute for Mathematical Sciences (NIMS) grant funded by the Korean government (No. NIMS-B22920000). We are deeply grateful to Dr H Jung and Dr S M Lee of HDXWILL for their help and collaboration.

Appendix. Cross validation for MAR performance comparison

A total of twenty metal-free scans were equally split into four non-overlapping folds. Here, one fold was retained and used for testing, and the remaining three folds were used for training. This validation process was then repeated 4 times. From each scan of folds, we generated four (for a training fold) or two (for a test fold) realistic metal-artifacted data using the method described in section 3.2. In each iteration (or partition), a total of 60 paired data ($= 4 \text{ syntheses} \times 5 \text{ scans} \times 3 \text{ folds}$) were used for training, and a total of 10 paired data ($= 2 \text{ syntheses} \times 5 \text{ scans} \times 1 \text{ fold}$) were used for test. Figure A1 illustrates the 4-fold cross-validation process. Table A1 shows the normalized mean square error for the test data in four different partitions. The result shows that the proposed method outperforms the img DL method on all partitions.

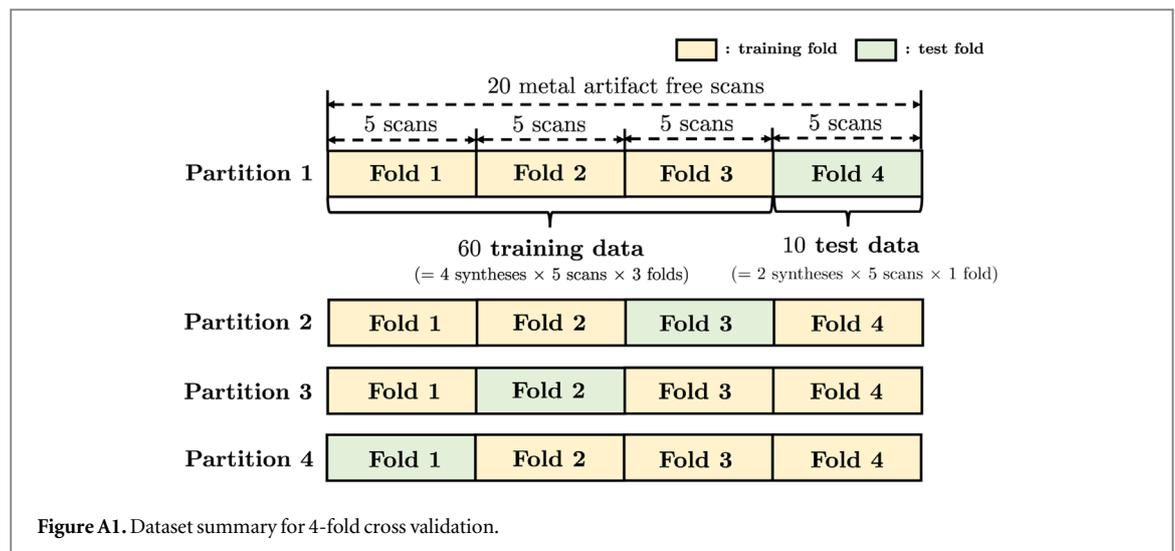


Table A1. Cross validation results of the img DL and proposed method.

NMSE	Partition 1	Partition 2	Partition 3	Partition 4
Img DL	0.3577	0.3821	0.3169	0.3626
Proposed Method	0.3421	0.3738	0.3009	0.3459

ORCID iDs

Chang Min Hyun  <https://orcid.org/0000-0002-7072-7489>

Hyoung Suk Park  <https://orcid.org/0000-0003-0032-4630>

References

- Abdoli M, Ay M R, Ahmadian A, Dierckx R A J O and Zaidi H 2010 Reduction of dental filling metallic artifacts in CT-based attenuation correction of PET data using weighted virtual sinograms optimized by a genetic algorithm *Med. Phys.* **37** 6166–77
- Alvarez R E and Macovski A 1976 Energy-selective reconstructions in X-ray computerized tomography *Phys. Med. Biol.* **21** 733–44
- Bazalova M, Beaulieu L, Palefsky S and Verhaegen F 2007 Correction of CT artifacts and its influence on monte carlo dose calculations *Med. Phys.* **34** 2119–32
- Biguri A, Dosanjh M, Hancock S and Soleimani M 2016 TIGRE: a MATLAB-GPU toolbox for CBCT image reconstruction *Biomed. Phys. Eng. Express* **2** 055010
- Bayaraa T, Hyun C M, Jang T J, Lee S M and Seo J K 2020 A two-stage approach for beam hardening artifact reduction in low-dose dental CBCT *IEEE Access* **8** 225981
- De Man B, Nuyts J, Dupont P, Marchal G and Suetens P 2001 An iterative maximum-likelihood polychromatic algorithm for CT *IEEE Trans. Med. Imaging* **20** 999–1008
- Draenert F, Coppenrath E, Herzog P, Muller S and Mueller-Lisse U 2007 Beam hardening artefacts occur in dental implant scans with the NewTom cone beam CT but not with the dental 4-row multidetector CT *Dentomaxillofacial Radiol.* **36** 198–203
- Edelsbrunner H and Mucke E P 1994 Three-dimensional alpha shapes *ACM Trans. Graph.* **13** 43–72
- Elbakri I A and Fessler J A 2002 Statistical image reconstruction for polyenergetic X-ray computed tomography *IEEE Trans. Med. Imaging* **21** 89–99
- Esmacilli F, Johari M, Haddadi P and Vatankhah M 2012 Beam hardening artifacts: comparison between two cone beam computed tomography scanners *J. Dental Res., Dental Clin., Dental Prospects* **6** 49
- Feldkamp L A, Davis L C and Kress J W 1984 Practical cone-beam algorithm *J. Opt. Soc. Am. A* **1** 612–9
- Gateno J et al 2007 Clinical feasibility of computer-aided surgical simulation (CASS) in the treatment of complex cranio-maxillofacial deformities *J. Oral Maxillof. Surg.* **65** 728–34
- Gupta J and Ali S P 2013 Cone beam computed tomography in oral implants *Natl J. Maxillofacial Surg.* **4** 2
- Gjesteby L, De Man B, Jin Y, Paganetti H, Verburg J, Giantsoudi D and Wang G 2016 Metal artifact reduction in CT: where are we after four decades? *IEEE Access* **4** 5826–49
- Gjesteby L, Shan H, Yang Q, Xi Y, Jin Y, Giantsoudi D, Paganetti H, Man B D and Wang G 2019 A dual-stream deep convolutional network for reducing metal streak artifacts in CT images *Phys. Bio. Med.* **64** 235003
- Huttenlocher D P, Klanderma G A and Rucklidge W J 1993 Comparing images using the Hausdorff distance *IEEE Trans. Pattern Anal. Mach. Intell.* **15** 850–63
- Hubbell J H and Seltzer S M 1995 Tables of X-ray mass attenuation coefficients and mass energy-absorption coefficients 1 keV to 20 MeV for elements Z=1 to 92 and 48 additional substances of dosimetric interest <http://physics.nist.gov/PhysRefData/XrayMassCoef/cover.html>
- Hyun C M, Kim K C, Cho H C, Choi J K and Seo J K 2020 Framelet pooling aided deep learning network: the method to process high dimensional medical data *Mach. Learn.: Sci. Technol.* **1** 015009
- Hyun C M, Baek S H, Lee M, Lee S M and Seo J K 2021 Deep learning-based solvability of underdetermined inverse problems in medical imaging *Med. Image Anal.* **69** 101967
- Ioffe S and Szegedy C 2015 Batch normalization : accelerating deep network training by reducing internal covariate shift *International conference on machine learning, PMLR* **37**, 448–56
- Jang T J, Kim K C, Cho H C and Seo J K 2021 A fully automated method for 3D individual tooth identification and segmentation in dental CBCT *IEEE Trans. Pattern Anal. Mach. Intell.* (<https://doi.org/10.1109/TPAMI.2021.3086072>)
- Jang T J, Yun H S and Seo J K 2021 Fully automatic integration of dental CBCT images and full-arch intraoral impressions with stitching error correction via individual tooth segmentation and identification arXiv:2112.01784
- Kalender W A, Hebel R and Ebersberger J 1987 Reduction of CT artifacts caused by metallic implants *Radiology* **164** 576–7
- Kingma D and Ba J 2014 Adam: A method for stochastic optimization arXiv:1412.6980
- Kwon M, Cho Y, Kim D W, Kim M S, Kim Y J and Chang M 2021 Full-arch accuracy of five intraoral scanners: In vivo analysis of trueness and precision *Korean J. Orthod.* **51** 95–104
- Lin W-A, Liao H, Peng C, Sun X, Zhang J, Luo J, Chellappa R and Zhou S K 2019 DUDONET: Dual domain network for ct metal artifact reduction *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* pp 10504–13
- Liu J, Aviles-Rivero A I, Ji H and Schonlieb C B 2021 Rethinking medical image reconstruction via shape prior, going deeper and faster: Deep joint indirect registration and reconstruction *Med. Image Anal.* **68** 101930
- Lehmann L, Alvarez R, Macovski A, Brody W, Pelc N, Riederer S and Hall A 1981 Generalized image combinations in dual kVp digital radiography *Med. Phys.* **8** 659–67
- Marchetti C, Bianchi A, Bassi M, Gori R, Lamberti C and Sarti A 2007 Mathematical modeling and numerical simulation in maxillofacial virtual surgery *J. Craniofacial Surg.* **18** 826–32
- Miracle A and Mukherji S 2009 Conebeam CT of the head and neck, part 2: clinical applications *Am. J. Neuroradiol.* **30** 1285–92
- Mahesh M 2013 The essential physics of medical imaging *Med. Phys.* **40** 077301
- Mehta R and Sivaswamy J 2017 M-net: a convolutional neural network for deep brain structure segmentation 2017 *IEEE 14th Int. Symp. on Biomedical Imaging (ISBI 2017)*, *IEEE* pp 437–40
- Menvielle N, Goussard Y, Orban D and Soulez G 2005 Reduction of beam-hardening artifacts in X-ray CT *27th Annual Int. Conf. of the Engineering in Medicine and Biology Society, IEEE* pp 1865–8
- Meyer E, Raupach R, Lell M, Schmidt B and Kachelrieß M 2010 Normalized metal artifact reduction (NMAR) in computed tomography *Med. Phys.* **37** 5482–93
- Nelson S J 2014 *Wheeler's Dental Anatomy, Physiology and Occlusion* (Amsterdam: Elsevier)
- Nardi C, Borri C, Regini F, Calistri L, Castellani A, Lorini C and Colagrande S 2015 Metal and motion artifacts by cone beam computed tomography (CBCT) in dental and maxillofacial study *La Radiol. Med.* **120** 618–26

- Nagy Z, Simon B, Mennito A, Evans Z, Renne W and Vag J 2020 Comparing the trueness of seven intraoral scanners and a physical impression on dentate human maxilla by a novel method *BMC Oral Health* (<https://doi.org/10.1186/s12903-020-01090-x>)
- O'Sullivan J A and Benac J 2007 Alternating minimization algorithms for transmission tomography *IEEE Trans. Med. Imaging* **26** 283–97
- Pauwels R, Stamatakis H, Bosmans H, Bogaerts R, Jacobs R, Horner K, Tsiklakis K and SEDENTEXCT Project Consortium 2013 Quantification of metal artifacts on cone beam computed tomography images *Clin. Oral Implants Res.* **24** 94–9
- Park H S, Choi J K, Park K R, Kim K S, Lee S H, Ye J C and Seo J K 2013 Metal artifact reduction in CT by identifying missing data hidden in metals *J. X-Ray Sci. Technol.* **21** 357–72
- Park H S, Lee S M, Kim H P, Seo J K and Chung Y E 2018 CT sinogram consistency learning for metal-induced beam hardening correction *Med. Phys.* **45** 5376–84
- Paszke A et al 2019 PyTorch: An imperative style, high-performance deep learning library *Adv. Neural Inf. Process. Syst.* **32** 8026–37 arXiv:1912.01703
- Razavi T, Palmer R M, Davies J, Wilson R and Palmer P J 2010 Accuracy of measuring the cortical bone thickness adjacent to dental implants using cone beam computed tomography *Clin. Oral Implants Res.* **21** 718–25
- Ronneberger O, Fischer P and Brox T 2015 U-net: Convolutional networks for biomedical image segmentation *Int. Conf. on Medical Image Computing and Computer-assisted Intervention* pp 234–41
- Santler G, Karcher H and Ruda C 1998 Indications and limitations of three-dimensional models in cranio-maxillofacial surgery *J. Cranio Maxillof. Surg.* **26** 11–6
- Sukovic P 2003 Cone Beam Computed Tomography in craniofacial imaging *Orthodontics Craniofacial Res.* **6** 31–6
- Sanders M, Hoyjberg C, Chu C, Leggitt V and Kim J 2007 Common orthodontic appliances cause artifacts that degrade the diagnostic quality of CBCT images *J. California Dental Assoc.* **35** 850–7
- Swennen G R J et al 2009 A cone-beam CT based technique to augment the 3D virtual skull model with a detailed dental surface *Int. J. Oral Maxillofacial Surg.* **38** 48–57
- Swennen G R J et al 2009 A cone-beam computed tomography triple scan procedure to obtain a three-dimensional augmented virtual skull model appropriate for orthognathic surgery planning *J. Craniofacial Surg.* **20** 297–307
- Schulze R K W, Berndt D and D'Hoedt B 2010 On cone-beam computed tomography artifacts induced by titanium implants *Clin. Oral Implants Res.* **21** 100–7
- Schulze R, Heil U, Grob D, Bruellmann D, Dranischnikow E, Schwanecke U and Schoemer E 2011 Artefacts in CBCT: a review *Dentomaxillofacial Radiol.* **40** 265–73
- Sancho-Puchades M, Hammerle C H and Benic G I 2015 In vitro assessment of artifacts induced by titanium, titanium-zirconium and zirconium dioxide implants in cone-beam computed tomography *Clin. Oral Implants Res.* **26** 1222–8
- Scarfe W, Azevedo B, Toghiani S and Farman A 2017 Cone beam computed tomographic imaging in orthodontics *Australian Dental J.* **62** 33–50
- Sun L, Fan Z, Ding X, Huang Y and Paisley J 2019 Joint CS-MRI reconstruction and segmentation with a unified deep network *Int. Conf. on Information Processing in Medical Imaging, Springer* pp 492–504
- Weiss R and Read-Fuller A 2019 Cone beam computed tomography in oral and maxillofacial surgery: an evidence-based review *Dentistry J.* **7** 52
- Williamson J F, Whiting B R, Benac J, Murphy R J, Blaine G J, O'Sullivan J A, Politte D G and Snyder D L 2002 Prospects for quantitative computed tomography imaging in the presence of foreign metal bodies using statistical image reconstruction *Med. Phys.* **29** 2404–18
- Winkler J and Gkantidis N 2020 Trueness and precision of intraoral scanners in the maxillary dental arch: an in vivo analysis *Sci. Rep.* **10** 1–11
- Yu L, Leng S and McCollough C H 2012 Dual-energy CT-based monochromatic imaging *Am. J. Roentgenol.* **199** S9–15
- Yu L, Zhang Z, Li X and Xing L 2020 Deep Sinogram Completion with Image Prior for Metal Artifact Reduction in CT Images *IEEE Trans. Med. Imaging* **40** 228–38
- Zhang Y and Yu H 2018 Convolutional neural network based metal artifact reduction in x-ray computed tomography *IEEE Trans. Med. Imaging* **37** 1370–81