PAPER • OPEN ACCESS

Generative adversarial networks improve interior computed tomography angiography reconstruction

Biomedical Physics & Engineering Express

PAPER

Generative adversarial networks improve interior computed tomography angiography reconstruction

Juuso H J Ketola[1,2], Helinä Heino[1], Mikael A K Juntunen[1,3], Miika T Nieminen[1,3,4], Samuli Siltanen[5] and Satu I Inkinen[1]

[1] Research Unit of Medical Imaging, Physics and Technology, University of Oulu, FI-90014, Finland
[2] The South Savo Social and Health Care Authority, Mikkeli Central Hospital, FI-50100, Finland
[3] Department of Diagnostic Radiology, Oulu University Hospital, FI-90029, Finland
[4] Medical Research Center Oulu, University of Oulu and Oulu University Hospital, FI-90014, Finland
[5] Department of Mathematics and Statistics, University of Helsinki, Helsinki, FI-00014, Finland

E-mail: juuso.ketola@oulu.fi

Keywords: computed tomography, convolutional neural networks, generative adversarial networks, image reconstruction, interior tomography, sinogram extension

OPEN ACCESS

RECEIVED
27 July 2021

REVISED
29 September 2021

ACCEPTED FOR PUBLICATION
21 October 2021

PUBLISHED
29 October 2021

Abstract

In interior computed tomography (CT), the x-ray beam is collimated to a limited field-of-view (FOV) (e.g. the volume of the heart) to decrease exposure to adjacent organs, but the resulting image has a severe truncation artifact when reconstructed with traditional filtered back-projection (FBP) type algorithms. In some examinations, such as cardiac or dentomaxillofacial imaging, interior CT could be used to achieve further dose reductions. In this work, we describe a deep learning (DL) method to obtain artifact-free images from interior CT angiography. Our method employs the Pix2Pix generative adversarial network (GAN) in a two-stage process: (1) An extended sinogram is computed from a truncated sinogram with one GAN model, and (2) the FBP reconstruction obtained from that extended sinogram is used as an input to another GAN model that improves the quality of the interior reconstruction. Our double GAN (DGAN) model was trained with 10 000 truncated sinograms simulated from real computed tomography angiography slice images. Truncated sinograms (input) were used with original slice images (target) in training to yield an improved reconstruction (output). DGAN performance was compared with the adaptive de-truncation method, total variation regularization, and two reference DL methods: FBPConvNet, and U-Net-based sinogram extension (ES-UNet). Our DGAN method and ES-UNet yielded the best root-mean-squared error (RMSE) (0.03 ± 0.01), and structural similarity index (SSIM) (0.92 ± 0.02) values, and reference DL methods also yielded good results. Furthermore, we performed an extended FOV analysis by increasing the reconstruction area by 10% and 20%. In both cases, the DGAN approach yielded best results at RMSE (0.03 ± 0.01 and 0.04 ± 0.01 for the 10% and 20% cases, respectively), peak signal-to-noise ratio (PSNR) (30.5 ± 2.6 dB and 28.6 ± 2.6 dB), and SSIM (0.90 ± 0.02 and 0.87 ± 0.02). In conclusion, our method was able to not only reconstruct the interior region with improved image quality, but also extend the reconstructed FOV by 20%.

Introduction

Computed tomography (CT) is widely used in medical diagnostics to produce a three-dimensional representation of internal organs, bones, and soft tissue. Computed tomography angiography (CTA) is a specific CT technique, in which a contrast agent is intravenously injected into the bloodstream to allow good visualization of blood vessels in the body. Coronary CTA is used to determine the extent of stenosis in the coronary arteries. In CT scans, x-ray attenuation data are measured from several projection angles, exposing the patients to adverse stochastic effects of ionizing radiation. Utilization of CT increased significantly in the early 2000's, and it has been a major contributor to population dose since [1, 2]. Significant variations in dose levels between countries have been reported [3]. In line with the

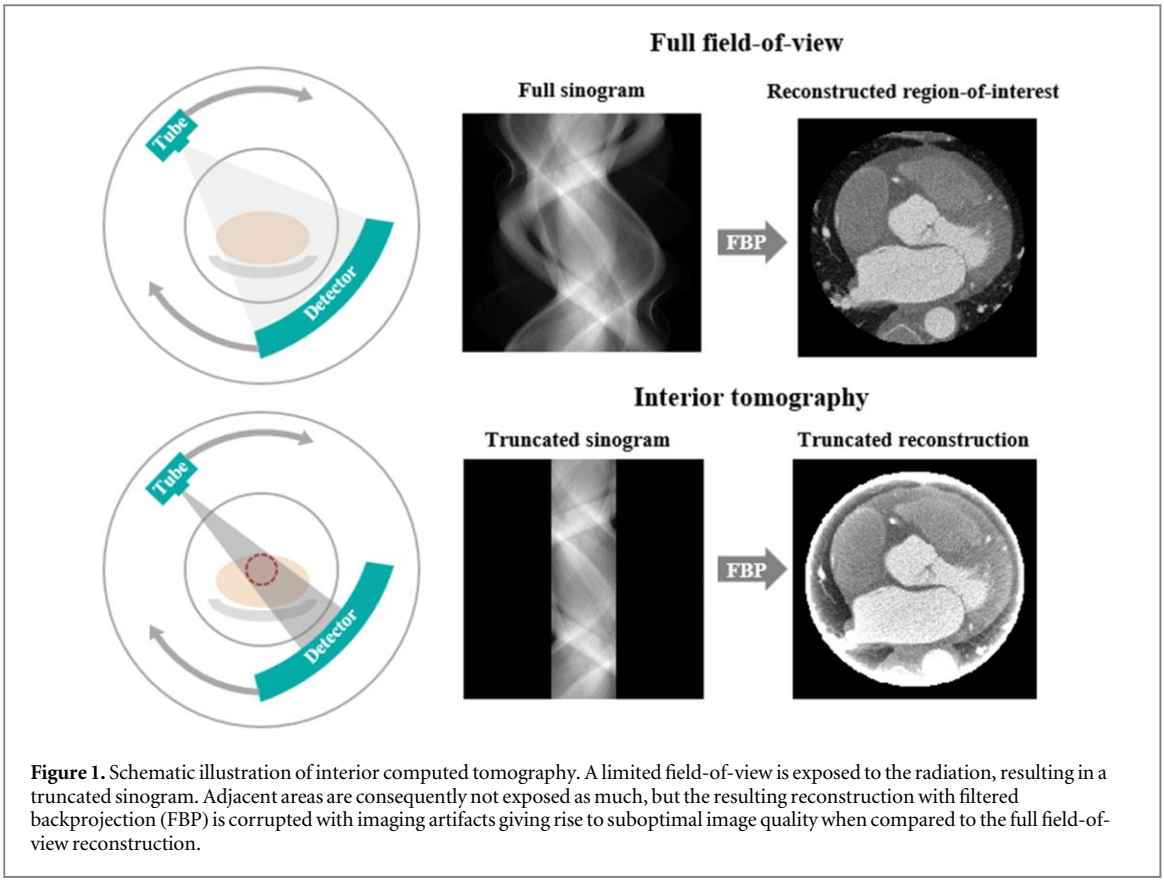

Figure 1. Schematic illustration of interior computed tomography. A limited field-of-view is exposed to the radiation, resulting in a truncated sinogram. Adjacent areas are consequently not exposed as much, but the resulting reconstruction with filtered backprojection (FBP) is corrupted with imaging artifacts giving rise to suboptimal image quality when compared to the full field-of-view reconstruction.

ALARA (as low as reasonably achievable) principle, many innovative strategies have been developed for dose reduction in the past years, such as automatic exposure control, beam filtering methods, and electrocardiography (ECG) and respiratory gating. Techniques to improve suboptimal image quality of low-dose CT scans are mainly focused on iterative reconstruction [4–6] and deep learning reconstruction or image restoration techniques [7, 8].

Interior tomography is another novel approach for CT dose reduction [9]. It refers to the collection and reconstruction of attenuation data from a limited field-of-view (FoV) by directing x-ray exposure to an internal region-of-interest (RoI) (figure 1). For example, in cardiac studies, the heart could be imaged by restricting the x-ray beam to irradiate an RoI bounded by the heart. It has been estimated that a dose reduction as high as 58% could be achieved with interior tomography [10]. A recent study demonstrated that doses in organs next to the heart could be reduced by up to 60.5% in cardiac interior tomography [11]. However, the major drawback of this acquisition strategy is the prominent imaging artifact seen in reconstructed images arising from partially measured data outside the FoV [12] (figure 1). This truncation artifact produces a 'halo' or 'cupping' effect that may hide diagnostically important features near the RoI edges, rendering image quality suboptimal.

Mathematically, these artifacts arise from the ill-posedness of the associated inverse problem as it has no unique solution [13]. However, a substantial amount of research efforts has been made to solve it. Compressed sensing (CS) methods have been widely used in ill-posed inverse problems, such as low-dose and interior CT reconstruction. CS is based on applying a suitable transform on the data to form a sparse representation of the data, allowing recovery of signals sampled below the Nyquist criterion. Total variation (TV) regularization, based on minimizing the L_1-norm of the image gradient, is a widely used and well-known CS method for CT reconstruction. In interior tomography, TV-based methods have been successfully used when a small sub-region inside the RoI is known *a priori* [12, 14–16], when the RoI is piecewise constant [17], and when the RoI is piecewise polynomial [18, 19]. Another popular alternative is the differentiated back-projection algorithm that exploits the properties of the Hilbert transform, and that has been demonstrated to produce exact reconstructions given that a sub-region is known within the RoI [20–23]. Furthermore, in the local tomography approach, an edge-enhanced reconstruction is produced but attenuation values are lost [24, 25]. As these methods include heavy restrictions that should be known *a priori*, they are limited when such information is not available.

Another fundamentally different approach to the interior problem is to pre-process the truncated projection data (or sinograms in a two-dimensional setting) before applying a reconstruction algorithm. In principle, this is done by extrapolation of raw data in the truncated areas. Strategies of such sinogram

detruncation (sometimes also referred to as inpainting) include, but are not limited to, extrapolation by symmetric mirroring [26], convex hull estimation followed by extrapolation [27, 28], extrapolation by the average value of near-boundary pixels [29], estimation of best-fitting water cylinders at the truncation edge [30], and sinogram decomposition followed by curve approximation and reprojection [31]. While able to improve image quality with relatively few computational resources, these pre-processing methods are often limited to approximating the sinogram shape in the truncated area with elementary functions, that can lead to remaining artefacts in the reconstructions as the projection data is not realistic. Furthermore, details beyond the interior sampling FoV cannot be recovered using these methods.

Deep learning (DL) can be used as an alternative approach for low-dose CT reconstruction. 'Deep reconstruction' not only has the potential to reduce radiation dose by improving low-dose image quality but also to speed up reconstruction times, as after network training is complete the output of the convolutional neural network (CNN) is fast to compute [32–34]. One strategy to apply DL in low-dose image reconstruction is by performing image-to-image transformation in the reconstruction domain, like in the FBPConvNet, where low-dose filtered back-projection (FBP) reconstructions are improved with a U-Net architecture [35]. Another branch of deep reconstruction research focuses on combining iterative reconstruction methods with DL, such as the LEARN network, in which the regularization term of iterative reconstruction is learned from data [36], or the learned primal-dual network, that unrolls proximal primal-dual optimization in the form of a CNN [37]. Tomographic data can also be processed with DL in the projection domain before applying reconstruction by, for example, up-sampling sparse-view data [38, 39]. Specifically in interior tomography, recent deep reconstruction approaches include removal of cupping artifacts with U-Net [40], linear sinogram extrapolation followed by reconstruction and U-Net post-processing [41], using a deep-learning prior obtained with U-Net in iterative reconstruction [42], and correcting truncation artifacts in projection space by generating extrapolated projections with generative adversarial networks (GAN) [43].

GAN is a specific type of CNN, where a generative CNN, *G*, is trained to output images *G*(x) that a discriminative network, *D*, tries to identify as fake images [44]. The networks are trained simultaneously to achieve two goals: (1) *G* is trained to produce such good images that *D* would confuse them as real images, while (2) *D* is trained to distinguish between real and fake images as well as possible. GANs have been studied in many medical applications, including classification, detection, segmentation, and reconstruction [44]. Other examples of GAN approaches for CT reconstruction include low-dose image denoising [45–47], sinogram up-sampling in sparse-view CT [48], sinogram in-painting in limited angle tomography [49], and metal artifact reduction [50].

In the present work, we applied GANs for interior tomography using a two-staged process: (1) an extended sinogram is computed from a truncated sinogram with one GAN model in projection space, and (2) the resulting FBP reconstruction from the extended sinogram is improved with another GAN model in reconstruction space, aiming to generate images with similar image quality compared to ground truth. As interior CT reconstruction is essentially a de-truncation problem, and many examples in the literature attempt to solve it with various sinogram inpainting techniques, GAN was selected owing to its ability to generate realistic data to missing regions [43]. Our aim was to investigate how well a GAN can generate the missing data in the sinogram domain, resulting in improved image quality in the reconstruction space, as well as to study if any remaining 'halo-like' truncation artefacts or other image artefacts due to the interior sampling geometry could be diminished after the reconstruction step with a secondary GAN. Furthermore, our aim was to study if the reconstructed FOV could be increased beyond the limits of the interior sampling FOV. We trained our models with real CTA data. Image quality was compared between our method and several state-of-the-art reconstruction techniques for interior tomography, including TV regularization, adaptive sinogram de-truncation, FBPConvNet, and U-Net-based sinogram extension.

Materials and methods

Data collection and preprocessing

CTA scans covering the entire heart craniocaudally (a total of 500 scans) were collected for this study (Oulu University Hospital register permission no. 104/2018). Imaging was performed with SOMATOM Definition Flash CT scanner (Siemens Healthineers, Erlangen, Germany) with a CTA protocol (80–120 kVp, 256 mean mAs, 0.73 × 0.73 mm^2 mean pixel size, 2-mm slice thickness, 512 × 512 image size). CTA images were preprocessed as follows (figure 2). First, the cardiac region was centered on the image stacks using zero-padding. Subsequently, the size of the slice images was unified according to the largest zero-padded slice image yielding an image size of 768 × 768. Finally, the image stacks were manually cropped to cover only the cardiac region in the z-direction. A total of 10 500 CTA slice images from the full data (30 834 images) were randomly chosen for this study. These data were split to training ($N = 9500$), validation ($N = 500$), and test ($N = 500$) sets. Training and validation data were obtained from 475 subjects, and test data were obtained from separate 25 subjects.

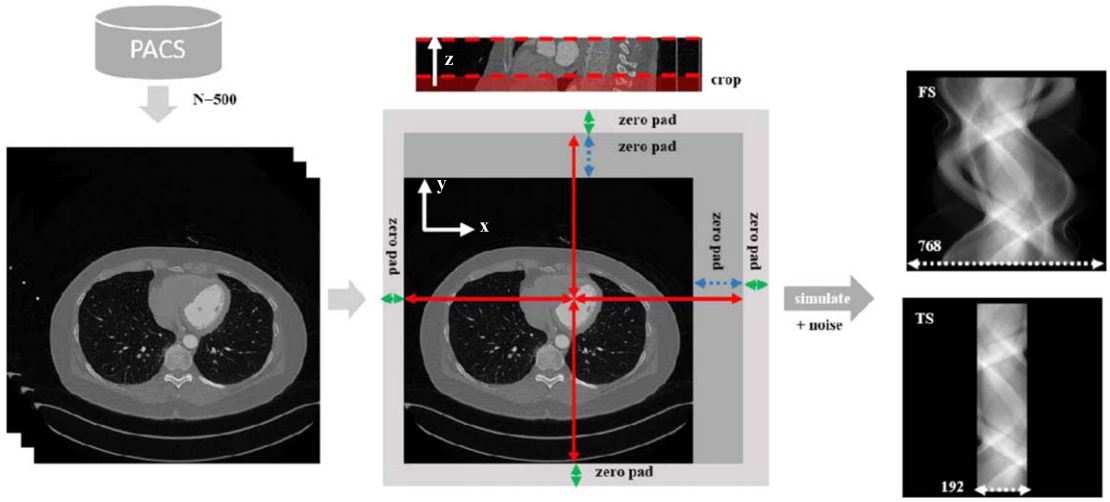

Figure 2. Data preprocessing pipeline. Computed tomography angiography volumes (N = 500) were collected from the picture archiving and communication systems (PACS). Each volume was centered with zero padding such that the heart was in the center. Further zero padding of slices to 768 × 768 was done to make all the images of same size. Unwanted slices were cropped out of the volumes. Sinograms were simulated using the ASTRA tomography toolbox to yield full sinograms (FS, 768 pixels wide) and truncated sinograms (TS, 192 pixels wide) simulating interior tomography measurements. The resulting sinograms were zero padded to 768 × 768 so that they could be fed in the deep learning networks.

Sinogram measurements for full FoV and interior cases were simulated with the ASTRA tomography toolbox (v1.8, imec-Vision Lab, University of Antwerp, CWI, Amsterdam, the Netherlands) [51, 52]. In the case of full FoV, a 768-pixel (56 cm) wide detector with a pixel-pitch of 0.73 mm was used in the simulation, and in the interior case, the detector width was decreased to 192 pixels (14 cm) since it should cover only the entire heart in a CTA scan (figure 1). Detector dimensions for simulation were determined from the dataset. Both measurements were simulated with the parallel-beam geometry using 720 projection angles evenly spaced in a full 360° rotation. Additional Gaussian noise (amplitude of 1% of the maximum value in the simulated full FOV sinogram) was added to the sinograms to simulate measurement noise and avoid inverse crime. Truncated sinograms were also reconstructed with FBP (Ram-Lak filter) with the ASTRA toolbox.

GAN architecture and training

We employed the Pix2Pix network architecture as our base GAN model that was used in both the pre- and post-processing networks (figure 3). Pix2Pix is a conditional GAN architecture that has been used in image-to-image mapping tasks, where the discriminator network input is a pair of images containing a 'fake' image synthesized by the generator network and a 'true' image that is the ground truth label [53]. We used same GAN architectures with input sizes of 768 × 768 × 1 and 512 × 512 × 1 for the sinogram extension and reconstruction post-processing models, respectively. We chose our U-Net as our generator model, and PatchGAN as our discriminator model, as in the work by Isola *et al* [53]. As suggested in the work by Isola *et al*, we included dropout layers in the U-Net bottleneck convolution blocks (dropout rate of 50%) instead of providing noise input to the network to introduce randomness to the network, as well as leaky ReLU activation functions (slope = −0.2) instead of conventional ReLUs (figure 3). PatchGAN classifies 70 × 70-pixel patches of the input image into 'real' or 'fake' classes (1 or 0, respectively). The final class is the average of these patchwise predictions.

Training the GAN was done as follows (figure 4). During every optimization step, the discriminator network (D) was trained first with input consisting of pairs of truncated and full sinograms. The concept of real and fake losses was utilized in training D. Real loss was defined as binary cross entropy (BCE) between output of D and a matrix of ones ('real' labels). In this case, the inputs of D were a truncated sinogram, x, and its corresponding ground truth label sinogram, y. The real loss was given by

$$\mathcal{L}_{Real} = \mathrm{BCE}(D(x, y), 1) = -\sum_i \log d_i,$$

where $d_i \in \{0, 1\}$ are the predictions in the output of D for patches i. On the other hand, the fake loss was defined as BCE between output of D and matrix of zeroes ('fake' labels). In this case, the inputs of D were a truncated sinogram, x, and the corresponding sinogram extended by the generator network (G), $G(x)$. The fake loss was given by

$$\mathcal{L}_{Fake} = \mathrm{BCE}(D(x, G(x)), 0) = -\sum_i \log(1 - d_i).$$

The discriminator loss, that was finally used to optimize D, was then obtained by averaging real and fake losses:

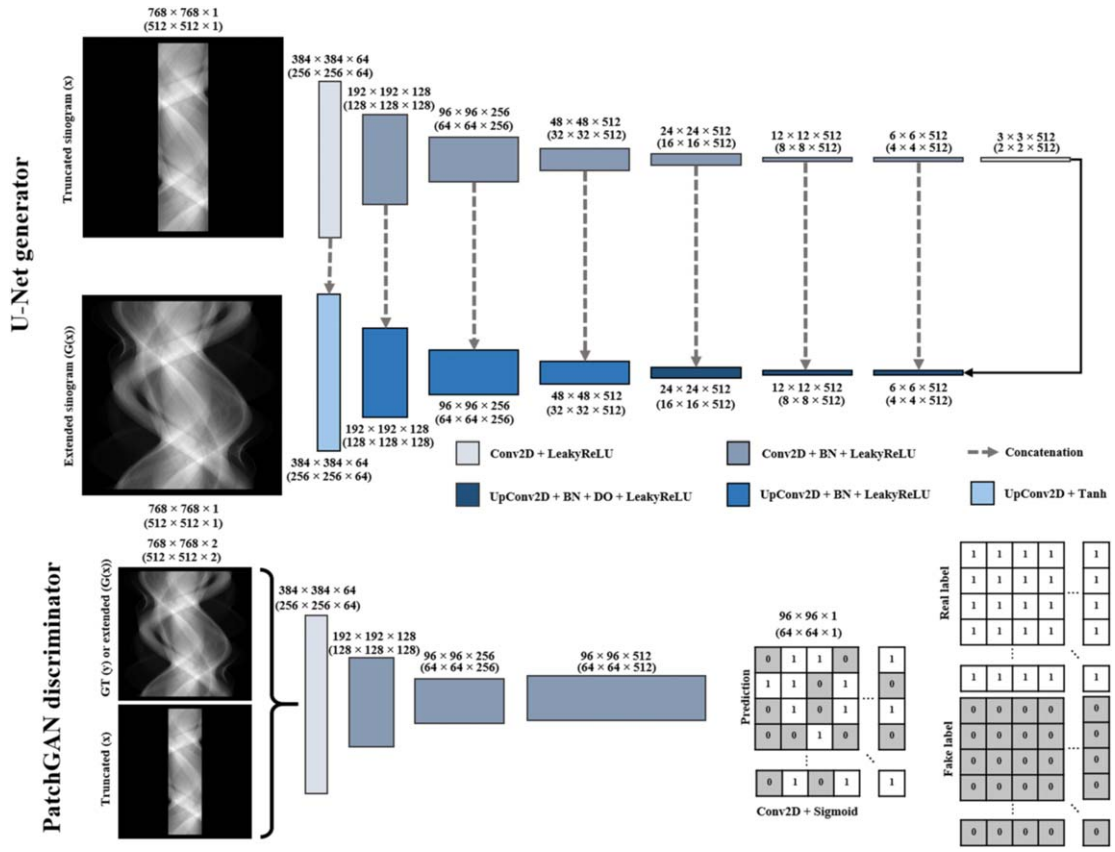

Figure 3. Pix2Pix generative adversarial network structure used in this study. The network consists of two subnetworks, a U-Net generator network and a PatchGAN discriminator network. The U-Net consists of encoding, bottleneck and decoding parts with skip connections in between (dashed arrows). The PatchGAN decodes image pairs of truncated and full sinograms (or reconstructions) to patches from which it predicts whether the pair of patches is real (input is the ground truth label) or fake (input is generated by U-Net). Filter tensor dimensions corresponding to the convolution blocks are listed above them. Numbers in brackets refer to tensor sizes for when reconstructions (512 × 512) are used as inputs instead of sinograms. Abbreviations: Conv2D, two-dimensional convolution; LeakyReLU, leaky rectified linear unit activation function; BN, batch normalization; DO, dropout; Tanh, hyperbolic tangent activation function.

$$\mathcal{L}_D = \frac{1}{2}(\mathcal{L}_{Real} + \mathcal{L}_{Fake}).$$

After the discriminator optimization step, the generator was optimized by summing two different loss functions. First, the L_1-loss between the sinogram extended by G, $G(x)$, and the corresponding ground truth label sinogram, y, was computed with

$$\mathcal{L}_{L_1} = \|G(x) - y\|_1 = \sum_j |g_j - y_j|,$$

where g_j and y_j refer to individual pixels. Next, the GAN loss, defined as the real loss using $G(x)$ and x as input, was computed with

$$\mathcal{L}_{GAN} = \mathrm{BCE}(D(x, G(x)), 1).$$

The GAN loss was computed to optimize G to generate such good quality images that D makes a wrong prediction. The final loss for G was obtained by weighting the two loss terms with a factor $\lambda = 0.01$

$$\mathcal{L}_G = \mathcal{L}_{L1} + \lambda\mathcal{L}_{GAN}.$$

Training the reconstruction space GAN model was done otherwise identically as described above, but the sinograms were replaced with the corresponding FBP reconstructions.

D and G were trained simultaneously for 200 epochs (figure 5, table 1). The two-timepoint update rule was used in the training process by using a higher learning rate for D to discourage it from becoming over-confident [54]. Label smoothing ($\alpha = 0.2$) was used for the same reason, scaling the output values of PatchGAN to 0.1 and 0.9 instead of 0 and 1 [55]. Separate networks were trained from scratch for the sinogram extension GAN and reconstruction post-processing GAN. A GPU with 24 GB memory was used to train these networks (NVIDIA Quadro P6000, NVIDIA, Santa Clara, CA). Training the networks took approximately 50 h per GAN.

Double GAN reconstruction algorithm

The proposed double GAN (DGAN) approach for interior CT image reconstruction works as follows (figure 6):

1. Extend truncated sinogram (TS) with the sinogram extension network (ExtGAN)

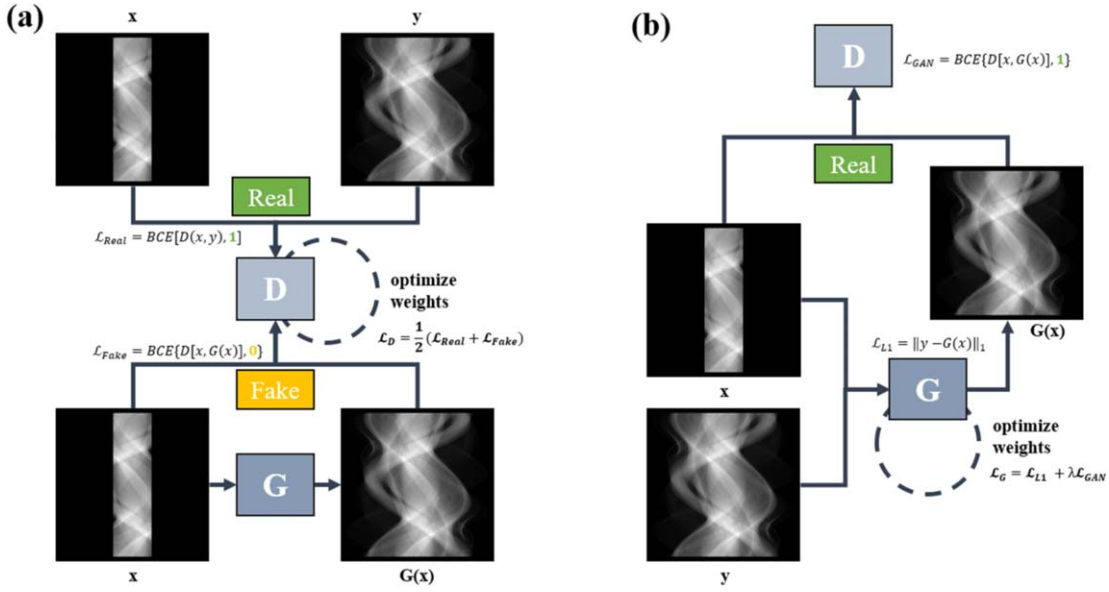

Figure 4. Training the GANs. A. During every optimization step, the discriminator D is trained first by feeding image pairs of truncated and ground truth sinograms (labeled real) and truncated and generated sinograms (labeled as fake). For the reconstruction domain GAN, the sinograms are replaced with the corresponding reconstructions. The loss used for training D is the mean of binary cross entropy (BCE) loss values corresponding to these real and fake cases. B. After D has been updated, G is trained by combining loss arising from the L1 loss between generated and ground truth sinograms and BCE loss of D when pairs of truncated and real sinograms are fed labeled as true (G attempts to make D guess wrong).

2. Superimpose the truncated sinogram on the extended sinogram to ensure any measured data were not changed by GAN
3. Perform FBP to obtain initial reconstruction
4. Post-process the initial reconstruction with the reconstruction post-processing network (RecGAN)
5. Forward project the resulting reconstruction and superimpose the original truncated sinogram in the improved extended sinogram
6. Perform FBP to obtain the final reconstruction.

Reference methods

To compare how our method performs against well-known reference methods, we used truncated FBP (TS-FBP) and truncated TV regularization (with Barzilai-Borwein optimization [56]) (TS-TV) as reference reconstruction methods on truncated sinograms. Furthermore, as a sinogram extension strategy, we utilized adaptive detruncation, and subsequently reconstructed the extended sinogram with FBP (ES-ADT) [57]. Two DL models were also used as reference methods: Image-to-image transform from truncated FBP reconstruction to full FBP reconstruction with FBPConvNet [35], and U-Net-based sinogram extrapolation followed with FBP reconstruction (ES-UNet) [58]. The same training data were used to train these networks for 200 epochs.

Image quality analysis

The image quality of the different methods was determined using both quantitative and visual evaluation. The quantitative image quality analyses consisted of root-mean-squared error (RMSE), peak signal-to-noise ratio (PSNR), and structural similarity index (SSIM) and were performed on the test data. Comparisons were done using the original FOV size (radius = 96 voxels) and FOVs that were increased by 10% (r = 106 voxels), 20% (r = 115 voxels), and 50% (r = 144 voxels). The increased FOV analysis was included to explore the possibility of enlarging the FOV area with DL methods.

Results

Our DGAN approach and ES-UNet yielded the best root-mean-squared error (RMSE) (0.03 ± 0.01) and structural similarity index (SSIM) (0.92 ± 0.02) values, and the best PSNR value was given by ES-UNet (32.6 ± 2.9 dB) in the original FOV (table 2). Example reconstructions show that the truncation artifact was removed from ES-ADT, FBPConvNet, ES-UNet, and DGAN (figure 7). When examining line profiles in the horizontal and vertical directions all the studied DL algorithms showed excellent agreement with the GT data even in the extended FOV, whereas TS-TV, and ES-ADT resulted in inferior agreement (figure 8).

When increasing the FOV, the DGAN approach yielded best results for RMSE (0.03 ± 0.01, 0.04 ± 0.01, and 0.04 ± 0.01 for the 10%, 20%, and 50% FOV increase, respectively), PSNR

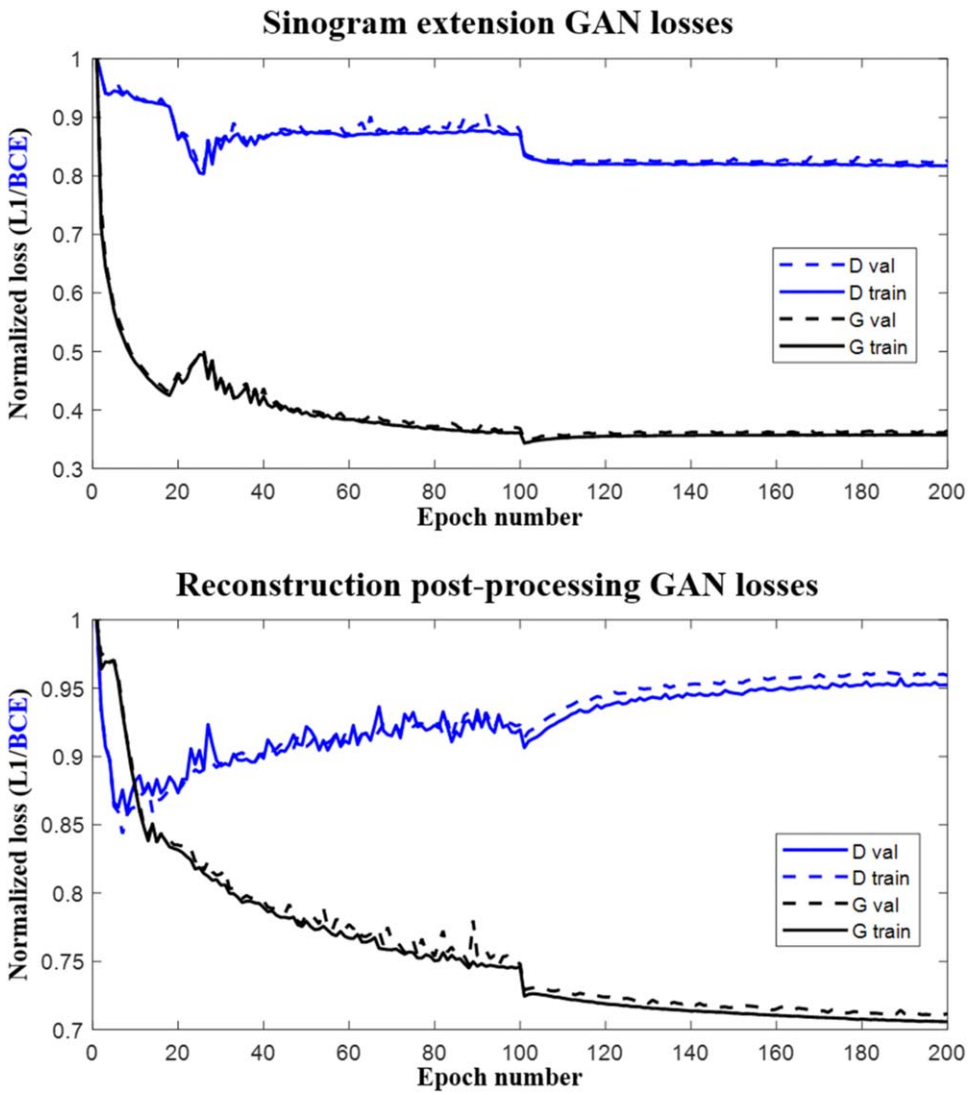

Figure 5. Training curves of the two generative adversarial network (GAN) models. D, discriminator; G, generator; BCE, binary cross entropy. After 100 epochs, the learning rate was adjusted from 0.001 to 0.0001 for G and from 0.004 to 0.0004 for D.

Table 1. Training parameters for the GAN models trained for this study. GAN, generative adversarial network; BCE, binary cross entropy.

	Generator	PatchGAN
# epochs	200	200
Loss	L1	BCE
Learning rate	0.001*	0.004*
Optimization	Adam**	Adam**
Batch size	4	4

*decayed by a factor of 10 after 100 epochs
**with default parameters $\beta_1 = 0.9$, $\beta_2 = 0.999$

(30.5 ± 2.6 dB, 28.6 ± 2.6 dB, 26.1 ± 2.2), and SSIM (0.90 ± 0.02, 0.87 ± 0.02, 0.82 ± 0.03) values (table 2). Visual inspection shows close resemblance between ground truth and FBPConvNet, ES-UNet, and DGAN in the extended areas (figure 9).

In addition, comparing the error metrics given by DGAN in the original FOV as a function of weight between each case in the test set where weight information was available ($N = 12$) showed a linear trend where the results slightly worsened in heavier patients (figure 10). Linear regression fit predicted an approximate increase from 2.4% RMSE to 3.1% RMSE (0.7% increase) between the heaviest (130 kg) and lightest (58 kg) cases. Correspondingly, the PSNR value decreased approximately from 32.5 dB to 30.5 dB (2 dB decrease), and the SSIM value decreased approximately from 0.94 to 0.89 (0.05 units decrease).

Discussion

In this study, we used a DGAN approach to reconstruct simulated truncated sinogram data from interior CTA scans to remove the truncation artifact and to extend the reconstruction FOV by 10%–20%. Our DGAN approach consisted of extending the sinogram with one GAN model and applying another GAN model to improve the extended sinogram

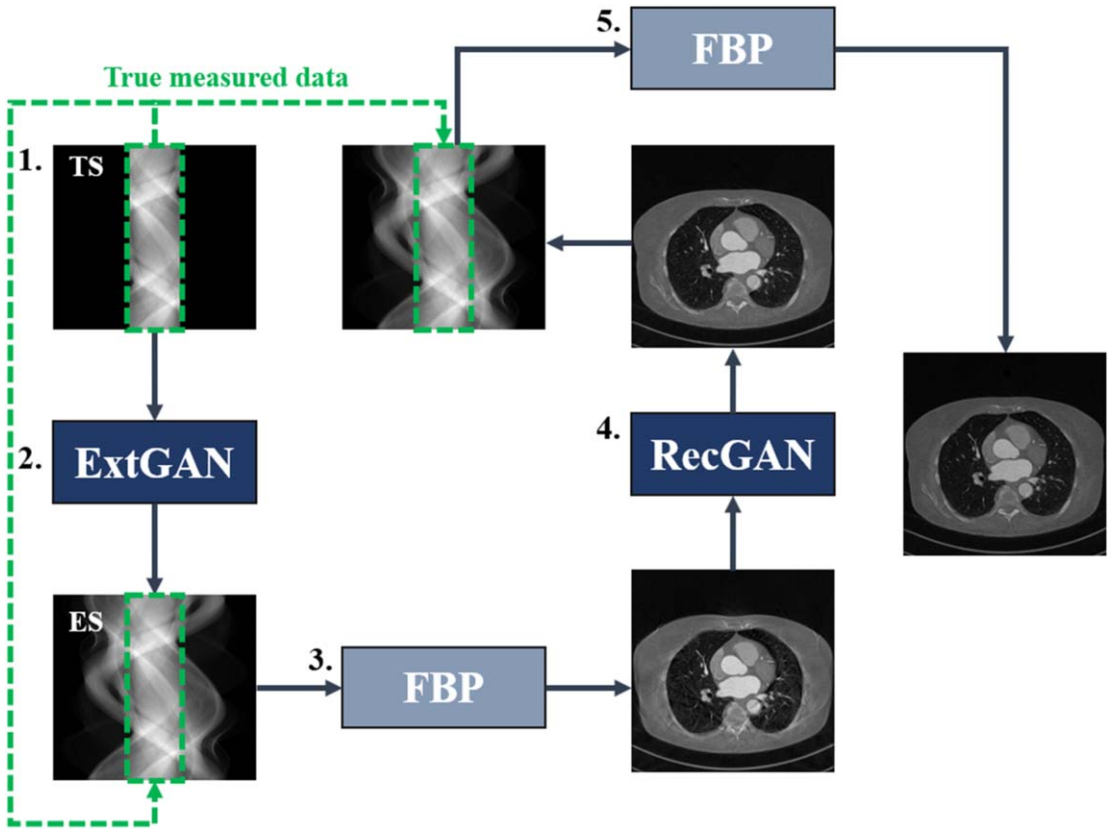

Figure 6. Block chart of the proposed DGAN method. A truncated sinogram (1) is first extended with sinogram extension GAN (ExtGAN) (2), and the original measured data is superimposed in it. After an initial reconstruction has been computed (3), reconstruction post-processing GAN is used to yield an improved reconstruction (4). To further enhance image quality, the improved reconstruction is forward projected and the original data is superimposed in the sinogram before final application of filtered backprojection (5).

Table 2. Image quality metrics for the studied algorithms with original and increased fields-of-view (FOV). The number in brackets denotes the radius of the circular FOV drawn in the image. RMSE, root mean squared error; PSNR, peak signal-to-noise ratio; SSIM, structural similarity index; TS-FBP, truncated sinogram filtered backprojection; TS-TV, truncated sinogram total variation regularization; ES-ADT, extended sinogram with adaptive detruncation followed with filtered backprojection; ES-UNet, extended sinogram with U-Net followed with filtered backprojection; DGAN, double generative adversarial networks for interior tomography.

	Original FOV (96 voxels)					
	TS-FBP	TS-TV	ES-ADT	FBPConvNet	ES-UNet	DGAN
RMSE	0.40 ± 0.10	0.14 ± 0.06	0.05 ± 0.03	0.07 ± 0.02	0.03 ± 0.01	0.03 ± 0.01
PSNR	8.3 ± 2.5	17.8 ± 4.0	26.3 ± 3.5	24.0 ± 2.9	32.6 ± 2.9	32.2 ± 2.3
SSIM	0.22 ± 0.09	0.75 ± 0.07	0.92 ± 0.06	0.85 ± 0.03	0.92 ± 0.02	0.92 ± 0.02
	10% increased FOV (106 voxels)					
	TS-FBP	TS-TV	ES-ADT	FBPConvNet	ES-UNet	DGAN
RMSE	0.36 ± 0.08	0.14 ± 0.05	0.07 ± 0.03	0.07 ± 0.02	0.03 ± 0.02	0.03 ± 0.01
PSNR	9.2 ± 2.4	18.0 ± 3.5	23.5 ± 2.8	24.2 ± 2.8	30.0 ± 3.3	30.5 ± 2.6
SSIM	0.23 ± 0.09	0.72 ± 0.06	0.89 ± 0.06	0.82 ± 0.03	0.90 ± 0.03	0.90 ± 0.02
	20% increased FOV (115 voxels)					
	TS-FBP	TS-TV	ES-ADT	FBPConvNet	ES-UNet	DGAN
RMSE	0.32 ± 0.07	0.13 ± 0.04	0.09 ± 0.03	0.06 ± 0.02	0.05 ± 0.02	0.04 ± 0.01
PSNR	10.3 ± 2.3	18.4 ± 3.0	21.0 ± 2.3	24.2 ± 2.7	27.4 ± 3.2	28.6 ± 2.6
SSIM	0.27 ± 0.08	0.69 ± 0.05	0.83 ± 0.07	0.80 ± 0.03	0.88 ± 0.03	0.87 ± 0.02
	50% increased FOV (144 voxels)					
	TS-FBP	TS-TV	ES-ADT	FBPConvNet	ES-UNet	DGAN
RMSE	0.28 ± 0.06	0.13 ± 0.04	0.13 ± 0.03	0.06 ± 0.02	0.06 ± 0.02	0.05 ± 0.01
PSNR	11.4 ± 2.1	18.2 ± 2.5	18.1 ± 2.5	24.1 ± 2.2	24.9 ± 2.6	26.1 ± 2.2
SSIM	0.31 ± 0.08	0.63 ± 0.05	0.72 ± 0.09	0.76 ± 0.03	0.82 ± 0.03	0.82 ± 0.03

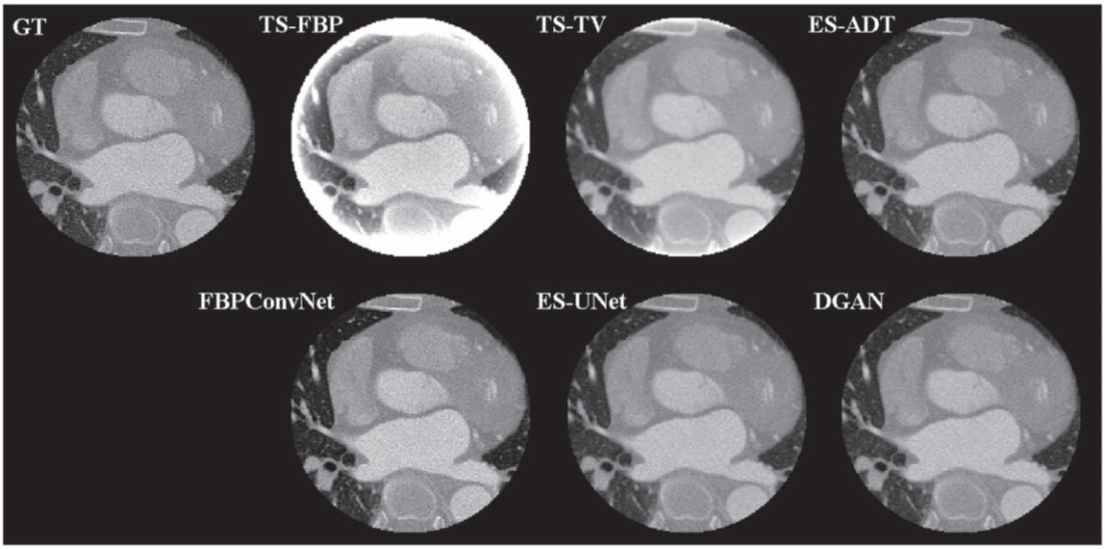

Figure 7. Example reconstructions given by the studied algorithms. Images have been normalized and windowed. GT, ground truth [Windowing: 0,1]; TS-FBP, truncated sinogram filtered backprojection [0,0.1]; TS-TV, truncated sinogram total variation regularization [0,0.75]; ES-ADT, extended sinogram with adaptive detruncation followed by filtered backprojection [0,1.1]; FBPConvNet [0,1]; ES-UNet, extended sinogram with U-Net followed by filtered backprojection [0,1.1]; DGAN, double generative adversarial networks for interior tomography [0,0.95].

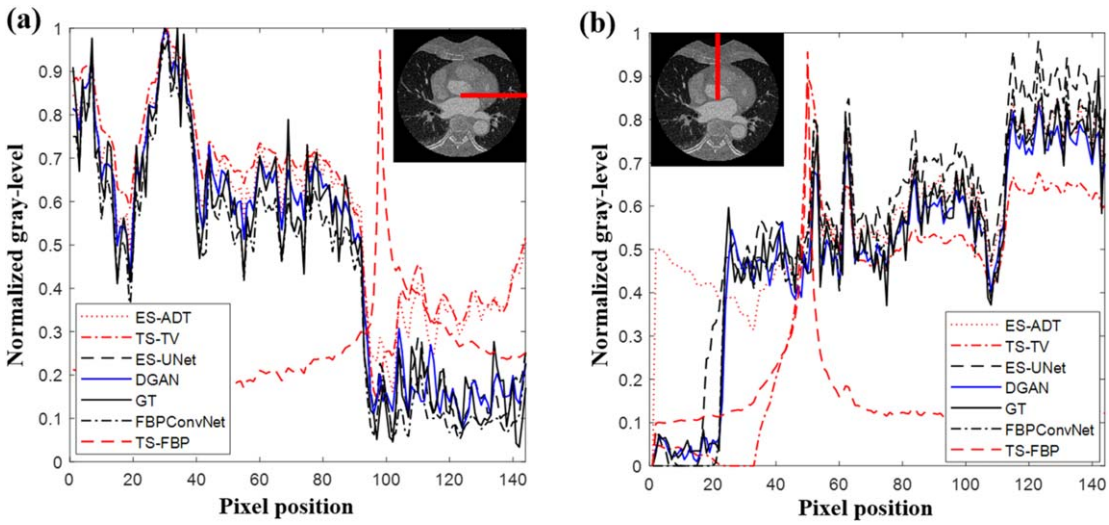

Figure 8. Partial normalized line profiles in the horizontal (A) and vertical (B) directions along the red line in the 50% increased FOV image. GT, ground truth; TS-FBP, truncated sinogram filtered backprojection; TS-TV, truncated sinogram total variation regularization; ES-ADT, extended sinogram with adaptive detruncation followed by filtered backprojection; ES-UNet, extended sinogram with U-Net followed by filtered backprojection; DGAN, double generative adversarial networks for interior tomography.

reconstruction. We employed the Pix2Pix GAN model in our network design. Results were compared between DGAN and several state-of-the-art reconstruction methods for interior tomography.

In the original FOV of 96 voxels, DGAN and ES-UNet exhibited the best numerical image quality across the studied algorithms. All algorithms except for TS-FBP and TS-TV were able to remove the truncation artifact in the images. Only minor differences were present in the horizontal and vertical line profiles between ground truth and the DL algorithms. In addition, DGAN was robust to varying patient sizes in the test set, and only minor deviation in the error metrics were observed between the lighter and heavier cases.

In the increased FOVs, the DL methods were able to recover incompletely sampled structures near the fully sampled VOI. In the 50% increased case, the DGAN approach was able to recover shapes and texture when inspected visually. However, increasing FOV by 50% is not likely to be a very realistic approach in clinical imaging. Therefore, we think that increasing the FOV by 10% to 20% would be feasible with DGAN. The DL methods also preserved similar noise structure and sharpness that is present in the ground truth. In addition to removing the truncation artifact,

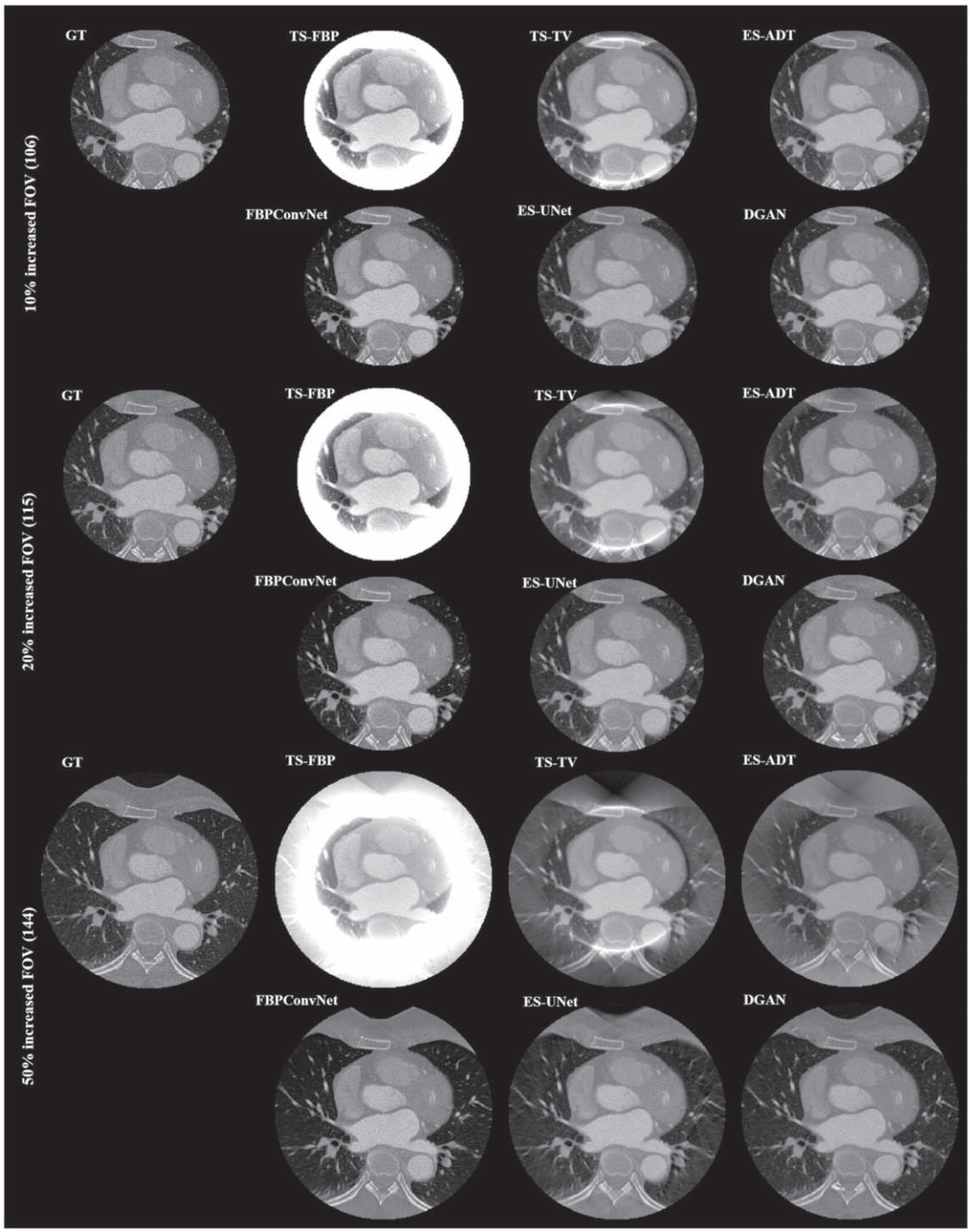

Figure 9. Example reconstructions with enlarged FOVs. Images have been normalized and windowed. GT, ground truth [Windowing: 0,]; TS-FBP, truncated sinogram filtered backprojection [0,0.1]; TS-TV, truncated sinogram total variation regularization [0,0.75]; ES-ADT, extended sinogram with adaptive detruncation followed by filtered backprojection [0,1.1]; FBPConvNet [0,1]; ES-UNet, extended sinogram with U-Net followed by filtered backprojection [0,1.1]; DGAN, double generative adversarial networks for interior tomography [0,0.95].

increasing the reconstructed FOV slightly could be beneficial in many imaging studies and promote dose efficiency in CT imaging. We hypothesize that a similar methodology could be useful in dentomaxillofacial and extremity imaging as well.

Compared to CS methods, our DGAN approach does not include any restrictions based on what prior information is available. Furthermore, compared to sinogram extension methods, our method improves the extension approach by letting it be learnt from data, and additionally, it improves the resulting reconstruction with another CNN processing step. To achieve better correspondence to measured data, we superimposed the true measured truncated data on the extended sinograms to remove any generated or removed detail in that area. This makes our network stand out from pure image-to-image transformations such as FBPConvNet. The DGAN approach includes

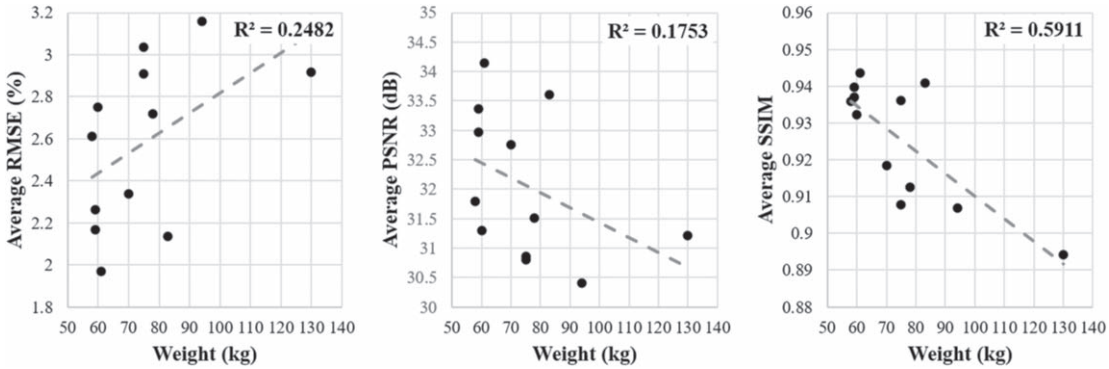

Figure 10. Root mean squared error (RMSE), peak signal-to-noise ratio (PSRN), and structural similarity index (SSIM) as a function of the weight of the patient in DGAN reconstructions (N = 12) using the original FOV of 96 pixels radius. Not all patients in the test set had weight information available. Average error was calculated as mean of error metrics in all the slices in individual cases. Linear regression predictions were plotted to visualize trendlines in the data.

several improvements to our previous U-Net-based sinogram extrapolation approach [58]. Recently, a method for processing truncated cone-beam CT data with GANs was presented [43]. Compared to this work, our method includes an additional post-processing step and is used on more severely truncated data with the extended FOV analysis, but on the other hand, our work is based on 2D simulated data.

Some limitations need to be discussed. As we did not have access to the raw data from the CT scanner, we simulated the CT measurements using a computational monochromatic forward model. In the interest of computational and time resources we simulated two-dimensional axial measurements instead of a helical scanning trajectory that is employed in most modern scanners. Thus, the measurement setting is not entirely realistic. We did, however, add extra noise to our simulations to introduce statistical uncertainty that would be present when measuring x-ray data. In addition, the voxel size of the CTA images was quite large, because they were the only data clinically available that also included the torso to simulate the interior problem. The data were also quite noisy limiting the quality of our ground truth images.

In the future, we aim to apply this extrapolation methodology in other areas of interest, such as cone-beam computed tomography of the maxillofacial region and imaging of obese patients where truncation artifacts are often present. In addition, it would be interesting to apply these techniques to real projection data, which are rarely available from clinical CT scanners.

Acknowledgments

The authors gratefully acknowledge financial support from Business Finland (project no. 1392/31/2016), Technology Industries of Finland Centennial Foundation, Jane & Aatos Erkko Foundation, Academy of Finland (project no. 316899), and Tauno Tönning Foundation (grants no. 20180083 and 20190204). The authors thank NVIDIA Corporation for donating the Quadro P6000 GPU used in this study. Pasi Sepponen is gratefully acknowledged for collecting the data used in this study.

Data availability statement

The data generated and/or analysed during the current study are not publicly available for legal/ethical reasons but are available from the corresponding author on reasonable request.

ORCID iDs

Juuso H J Ketola https://orcid.org/0000-0002-7760-6241
Mikael A K Juntunen https://orcid.org/0000-0003-4646-6714
Miika T Nieminen https://orcid.org/0000-0002-2300-2848
Samuli Siltanen https://orcid.org/0002-5988-5232
Satu I Inkinen https://orcid.org/0000-0002-9774-8925

References

[1] Bellolio F *et al* 2017 Increased computed tomography utilization in the emergency department and its association with hospital admission *West J Emerg Med.* **18** 835–45
[2] Mettler F A 2019 Medical radiation exposure in the United States: 2006–2016 trends *Health Phys.* **116** 126–8
[3] Smith-Bindman R *et al* 2019 International variation in radiation dose for computed tomography examinations: Prospective cohort study *Brit. Med. J.* **364** 1–12
[4] Beister M, Kolditz D and Kalender W A 2012 Iterative reconstruction methods in x-ray CT *Phys Medica.* **28** 94–108
[5] Willemink M J *et al* 2013 Iterative reconstruction techniques for computed tomography: II. Initial results in dose reduction and image quality *Eur. Radiol.* **23** 1632–42
[6] Willemink M J *et al* 2013 Iterative reconstruction techniques for computed tomography: I. Technical principles *Eur Radiol.* **23** 1623–31

[7] Hsieh J, Liu E, Nett B, Tang J, Thibault J-B and Sahney S 2019 A new era of image reconstruction: TrueFidelity ™ White Paper *GE Healthcare* **1** 1–15 (JB68676XX)

[8] Boedeker K 2019 AiCE deep learning reconstruction : bringing the power of ultra-high resolution CT to routine imaging *Canon Medical Systems Corporation*

[9] McCollough C H *et al* 2012 Achieving routine submillisievert CT scanning: report from the summit on management of radiation dose in CT *Radiology* **264** 567–80

[10] Bharkhada D *et al* 2010 Demonstration of dose and scatter reductions for interior computed tomography *J. Comput. Assist. Tomogr.* **33** 967–72

[11] Juntunen M A K *et al* 2020 Interior photon counting computed tomography for quantification of coronary artery calcium: Pre-clinical phantom study *Biomed. Phys. Eng. Express* **6** 055011

[12] Yu H and Wang G 2009 Compressed sensing based interior tomography *Phys. Med. Biol.* **54** 2791–805

[13] Natterer F 2001 *The Mathematics of Computerized Tomography*. (Philadelphia, PA: SIAM)

[14] Ye Y, Yu H, Wei Y, Wang G, Division B I and Engineering B 2007 A general local reconstruction approach based on a truncated hilbert transform *Int. J. Biomed. Imaging* **2007** 063634

[15] Jin X, Katsevich A, Yu H, Wang G, Li L and Chen Z 2012 Interior tomography with continuous singular value decomposition *IEEE Trans. Med. Imaging* **31** 2108–19

[16] Tang Q, Tsui B M W, Taguchi K, Cammin J, Xu J and Srivastava S 2011 Interior region-of-interest reconstruction using a small, nearly piecewise constant subregion *Med. Phys.* **38** 1307–12

[17] Cong W, Yang J and Wang G 2012 Differential phase-contrast interior tomography *Phys. Med. Biol.* **57** 2905–14

[18] Yang J, Yu H, Jiang M and Wang G 2010 High-order total variation minimization for interior tomography High-order total variation minimization for interior tomography *Inverse Probl.* **26** 035013

[19] Ward J P, Lee M, Ye J C and Unser M 2015 Interior tomography using 1D generalized total variation. part i: mathematical foundation *SIAM J Imaging Sci.* **8** 226–47

[20] Noo F, Clackdoyle R and Pack J D 2004 A two-step Hilbert transform method for 2D image reconstruction *Phys. Med. Biol.* **49** 3903

[21] Pan X, Zou Y and Xia D 2005 Image reconstruction in peripheral and central regions-of-interest and data redundancy *Med. Phys.* **32** 673–84

[22] Defrise M, Noo F, Clackdoyle R and Kudo H 2006 Truncated Hilbert transform and image reconstruction from limited tomographic data *Inverse Probl.* **22** 1037–53

[23] Kudo H, Courdurier M, Noo F and Defrise M 2008 Tiny *a priori* knowledge solves the interior problem in computed tomography *Phys. Med. Biol.* **53** 2207–31

[24] Faridani A, Ritman E L and Smith K T 1992 Local tomography *SIAM J. Appl. Math.* **52** 459–84

[25] Faridani A, Finch D V, Ritman E L and Smith K T 1997 Local tomography II *SIAM J. Appl. Math.* **57** 1095–127

[26] Ohnesorge B, Flohr T, Schwarz K, Heiken J P and Bae K T 2000 Efficient correction for CT image artifacts caused by objects extending outside the scan field of view *Med. Phys.* **27** 39–46

[27] Sourbelle K, Kachelrieß M and Kalender W A 2005 Reconstruction from truncated projections in CT using adaptive detruncation *Eur Radiol.* **15** 1008–14

[28] Kolditz D, Meyer M, Kyriakou Y and Kalender W A 2011 Comparison of extended field-of-view reconstructions in C-arm flat-detector CT using patient size, shape or attenuation information *Phys. Med. Biol.* **56** 39–56

[29] Kyrieleis A, Titarenko V, Ibison M, Connolley T and Withers P J 2011 Region-of-interest tomography using filtered backprojection: Assessing the practical limits *J. Microsc.* **241** 69–82

[30] Hsieh J *et al* 2004 A novel reconstruction algorithm to extend the CT scan field-of-view *Med. Phys.* **31** 2385–91

[31] Zamyatin A A and Nakanishi S 2007 Extension of the reconstruction field of view and truncation correction using sinogram decomposition *Med. Phys.* **34** 1593–604

[32] Wang G, Ye J C, Member S, Mueller K, Member S and Fessler J A 2018 Image reconstruction is a new frontier of machine learning *IEEE Trans. Med. Imaging* **37** 1289–96

[33] Wang G 2016 A perspective on deep imaging *IEEE Access.* **4** 8914–24

[34] Willemink M J and Noël P B 2019 The evolution of image reconstruction for CT—from filtered back projection to artificial intelligence *Eur Radiol.* **29** 2185–95

[35] Jin K H, McCann M T, Froustey E and Unser M 2016 *Deep Convolutional Neural Network for Inverse Problems in Imaging.* **26** 4509–22

[36] Chen H *et al* 2018 LEARN: learned experts' assessment-based reconstruction network for sparse-data CT *IEEE Trans. Med. Imaging* **37** 1333–47

[37] Adler J and Öktem O 2018 Learned primal-dual reconstruction *IEEE Trans. Med. Imaging* **37** 1322–32

[38] Lee H, Lee J, Kim H, Cho B and Cho S 2019 Deep-neural-network-based sinogram synthesis for sparse-view CT Image reconstruction *IEEE Trans Radiat Plasma Med Sci.* **3** 109–19

[39] Yuan H and Zhu Z 2018 Sipid: a deep learning framework for sinogram interpolation and image denoising in low-dose ct reconstruction *IEEE 15th Int Symp Biomed Imaging (ISBI 2018)* 1521–4

[40] Han Y and Ye J C 2019 One network to solve all ROIs: deep learning CT for any ROI using differentiated backprojection *Med. Phys.* **46** e855–72

[41] Fournié É, Baer-beck M and Stierstorfer K C T 2019 CT field of view extension using combined channels extension and deep learning methods arXiv:1908.09529

[42] Huang Y, Preuhs A and Maier A 2020 Field of view extension in computed tomography using deep learning prior *Bildverarbeitung für die Medizin.* 186–91

[43] Huang Y, Preuhs A, Manhart M, Lauritsch G and Maier A 2021 Data extrapolation from learned prior images for truncation correction in computed tomography *IEEE Trans. Med. Imaging* (https://doi.org/10.1109/TMI.2021.3072568)

[44] Yi X, Walia E and Babyn P 2019 Generative adversarial network in medical imaging: a review *Med. Image Anal.* **58** 101552

[45] Wolterink J M, Leiner T, Viergever M A and Išgum I 2017 Generative adversarial networks for noise reduction in low-dose CT *IEEE Trans. Med. Imaging* **36** 2536–45

[46] Yang Q, Yan P, Member S and Zhang Y 2018 Low-dose CT image denoising using a generative adversarial network with wasserstein distance and perceptual loss *IEEE Trans. Med. Imaging* **37** 1348–57

[47] Tatsugami F *et al* 2019 Deep learning—based image restoration algorithm for coronary CT angiography *Eur Radiol.* **29** 5322–9

[48] Ghani M U and Karl W C 2018 Deep learning-based sinogram completion for low-dose CT *2018 IEEE 13th Image, Video, Multidimensional Signal Processing Workshop (IVMSP).* 1–5

[49] Li Z *et al* 2019 Promising generative adversarial network based sinogram inpainting method for ultra-limited-angle computed tomography imaging *Sensors* **19** 3941

[50] Ghani M U and Karl W C 2018 Deep learning based sinogram correction for metal artifact reduction *Electron. Imaging* **15** 4721–8

[51] van Aarle W *et al* 2016 Fast and flexible x-ray tomography using the ASTRA toolbox *Opt. Express* **24** 25129

[52] van Aarle W *et al* 2015 The ASTRA Toolbox: a platform for advanced algorithm development in electron tomography *Ultramicroscopy* **157** 35–47

[53] Isola P, Zhu J Y, Zhou T and Efros A A 2017 Image-to-image translation with conditional adversarial networks *Proc. of the IEEE conference on computer vision and pattern recognition* 1125–34

[54] Heusel M, Ramsauer H, Unterthiner T, Nessler B and Hochreiter S 2017 GANs trained by a two time-scale update rule converge to a local nash equilibrium arXiv:1706.08500

[55] Müller R, Kornblith S and Hinton G 2019 When does label smoothing help? arXiv:1906.02629
[56] Hämäläinen K, Harhanen L, Hauptmann A, Kallonen A, Niemi E and Siltanen S 2014 Total variation regularization for large-scale X-ray tomography *Int J Tomogr Simul.* **25** 1–25
[57] Sourbelle K, Kachelriess M and Kalender W A 2005 Reconstruction from truncated projections in CT using adaptive detruncation *Eur Radiol.* **15** 1008–14
[58] Ketola J H, Heino H, Juntunen M A K, Nieminen M T and Inkinen S I 2021 Deep learning-based sinogram extension method for interior computed tomography *SPIE Medical Imaging 2021: Physics of Medical Imaging.* **11595** 115953Q