Physical and Engineering Sciences in Medicine (2023) 46:1399–1410
https://doi.org/10.1007/s13246-023-01307-7

SCIENTIFIC PAPER

Deep learning-based ultrasound transducer induced CT metal artifact reduction using generative adversarial networks for ultrasound-guided cardiac radioablation

Sathyathas Puvanasunthararajah[1,2] · Saskia M. Camps[3] · Marie-Luise Wille[2,4,5] · Davide Fontanarosa[1,2]

Received: 20 March 2023 / Accepted: 20 July 2023 / Published online: 7 August 2023
© The Author(s) 2023

Abstract
In US-guided cardiac radioablation, a possible workflow includes simultaneous US and planning CT acquisitions, which can result in US transducer-induced metal artifacts on the planning CT scans. To reduce the impact of these artifacts, a metal artifact reduction (MAR) algorithm has been developed based on a deep learning Generative Adversarial Network called Cycle-MAR, and compared with iMAR (Siemens), O-MAR (Philips) and MDT (ReVision Radiology), and CCS-MAR (Combined Clustered Scan-based MAR). Cycle-MAR was trained with a supervised learning scheme using sets of paired clinical CT scans with and without simulated artifacts. It was then evaluated on CT scans with real artifacts of an anthropomorphic phantom, and on sets of clinical CT scans with simulated artifacts which were not used for Cycle-MAR training. Image quality metrics and HU value-based analysis were used to evaluate the performance of Cycle-MAR compared to the other algorithms. The proposed Cycle-MAR network effectively reduces the negative impact of the metal artifacts. For example, the calculated HU value improvement percentage for the cardiac structures in the clinical CT scans was 59.58%, 62.22%, and 72.84% after MDT, CCS-MAR, and Cycle-MAR application, respectively. The application of MAR algorithms reduces the impact of US transducer-induced metal artifacts on CT scans. In comparison to iMAR, O-MAR, MDT, and CCS-MAR, the application of developed Cycle-MAR network on CT scans performs better in reducing these metal artifacts.

Keywords Cardiac radioablation · Ultrasound guidance · Metal artifact reduction · Deep learning · Computed tomography

Introduction

Cardiac radioablation is a new non-invasive modality for the treatment of cardiac arrhythmias. This treatment method is based on delivering a radiation dose to the arrhythmogenic tissues using external beam radiation therapy [1–3]. A typical treatment workflow includes acquiring a planning computed tomography (CT) used to delineate the arrhythmogenic tissue (target) and the organs-at-risk (OARs) for the Hounsfield Units (HU) derived foreseen radiation dose calculation. At treatment, the dose is then delivered to the target while sparing the OARs as much as possible. However, the complex cardiorespiratory motion may impact the accuracy of dose delivery [4, 5]. This makes real-time monitoring of the cardiorespiratory motion of paramount importance to achieve safe and effective treatment delivery.

A possible candidate image modality for real-time guidance in cardiac radioablation is transthoracic ultrasound (US) imaging [6, 7]. This approach relies on identifying the cardiac tissue position using US imaging at both the simulation and the treatment delivery stages. Comparing the two allows to compensate for possible displacements of cardiac structures at treatment. To minimize clinical workflow steps and ease treatment planning, it may be favourable to acquire the US scans simultaneously with the planning

✉ Sathyathas Puvanasunthararajah
sathyathas.puvanasunthararajah@hdr.qut.edu.au

1 School of Clinical Sciences, Queensland University of Technology, Brisbane, QLD, Australia

2 Centre for Biomedical Technologies, Queensland University of Technology, Brisbane, QLD, Australia

3 EBAMed SA, Geneva, Switzerland

4 School of Mechanical, Medical & Process Engineering, Faculty of Engineering, Queensland University of Technology, Brisbane, QLD, Australia

5 ARC ITTC for Multiscale 3D Imaging, Modelling, and Manufacturing, Queensland University of Technology, Brisbane, QLD, Australia

CT scan at the simulation stage. However, this approach is prone to creating transducer-induced metal artifacts on the CT scans, caused by the internal metal components of the US transducers [8, 9].

Metal artifacts may generate improper representations of anatomical structures and incorrect HU values, resulting in potentially inaccurate radiation dose calculation [10–12]. Several algorithms for metal artifact reduction (MAR) have been developed, mainly focusing on the artifacts generated by implanted metal structures. The majority of these MAR algorithms, both commercially available and research-based, follow a conventional analytical approach. Recently, new works based on deep learning have been proposed [11, 12]. Commercially available MAR algorithms use an iterative approach of correction of CT projection data [12]. Among them, the Orthopaedics Metal Artifact Reduction (O-MAR, Philips Health System) and the iterative Metal Artifact Reduction (iMAR, Siemens Healthcare) algorithms were widely investigated for the improvement of radiation therapy planning [13–21]. Among the research-based MAR algorithms, Metal Deletion Technique (MDT, ReVision Radiology) [22] was the one most often compared to the commercially available algorithms [23–26]. All these algorithms typically suffer from improper restoration of HU values, from distortion of anatomical structures, and even from creation of secondary artifacts [11]. Recently, our group [27] investigated a newly developed MAR algorithm, Combined Clustered Scan-based MAR (CCS-MAR), which followed the traditional analytical approach, and compared its performance to commonly used commercial and research-based MAR algorithms. The results of this study revealed that further development or improvements were needed to reduce residual artifacts and further improve HU value restoration capabilities.

In recent years, the development of deep learning-based algorithms for metal artifact reduction in CT imaging gained significant interest [28–38]. In general, to learn in a supervised manner [39, 40] the complex metal artifact patterns and propagation, these algorithms require data including CT scans with artifacts (CT_{art}), and corresponding artifact-free (CT_{ref}) scans. In the absence of adequate paired data, typical metal artifacts which resemble real clinical scenarios can be simulated [28, 30, 31, 38]. For our work, we have used generative adversarial networks (GAN) [41]. Typically, these architectures consist of a generator and a discriminator, which are types of convolutional neural networks (CNN) [41, 42]. The generator and the discriminator are trained in an adversarial manner to perform the transformation. We have focused in particular on their extension CycleGAN [43], which utilizes two GAN architectures, and has already been studied in literature for the transformation of CT_{art} scans into artifact-reduced CT (CT_{cor}) scans for radiation therapy applications [28, 29, 31].

This work aims to develop a deep learning-based MAR algorithm for the reduction of US transducer-induced metal artifacts on CT scans. The proposed algorithm was designed following a supervised learning scheme with a CycleGAN architecture using paired clinical CT scans. Then it was evaluated on phantom CT scans with real artifacts and clinical CT scans with simulated artifacts. In addition, the performance of CycleGAN has been compared with the performance of iMAR, O-MAR, MDT and CCS-MAR.

Materials and methods

CT data preparation

Clinical CT scans

Paired clinical CT scans with and without US transducer-induced metal artifacts were used. In particular, DICOM CT scans of the thoracic region were utilized from the "COVID19-CT-Dataset" online database [44]. Initially, CT scans from 180 patients were downloaded and visually inspected for suitability. CT scans or CT slices with large COVID19 induced density changes in the lungs region, with suboptimal quality due to low resolution, and/or with presence of metal artifacts resulting from external foreign bodies were excluded. This resulted in CT scans from 84 patients, and from each of these CT scans, axial CT slices composed of 22–100 slices per patient covering the cardiac structure from the apex to the base were selected. These CT scans consisted of 512 × 512-pixel slices with 1.5 or 3 mm slice thickness [45].

To simulate the metal artifacts on the CT scans, Sakamoto et al. [46]. developed a MatLab package based on the study by Zhang et al. [38]. This package was modified in our work for the specific simulation of US transducer-induced metal artifacts on the selected COVID19-CT-Dataset (See Fig. 1 a). In our procedure, initially the pixels identifying the US transducers on the phantom CT_{art} scans were manually segmented and stored separately. Then, the segmented US transducers were copied to the clinical CT scans and positioned on the scans using rigid rotation-translations to be imaging the cardiac structures. A threshold value of 2000 HU was applied on the imported US transducer to extract metal components, which were then saved as binary images. The metal extraction threshold value of 2000 HU was chosen according to the previous research published in the literature [27, 47, 48]. Then, pre-defined HU thresholds were used to segment the bone, lung, and water-equivalent tissues on the clinical CT_{ref} slices. To convert the HU values of the pixels into linear attenuation coefficient for varying X-ray energies, corresponding mass attenuation coefficients were used from the NIST [49] database. Subsequently, the polychromatic

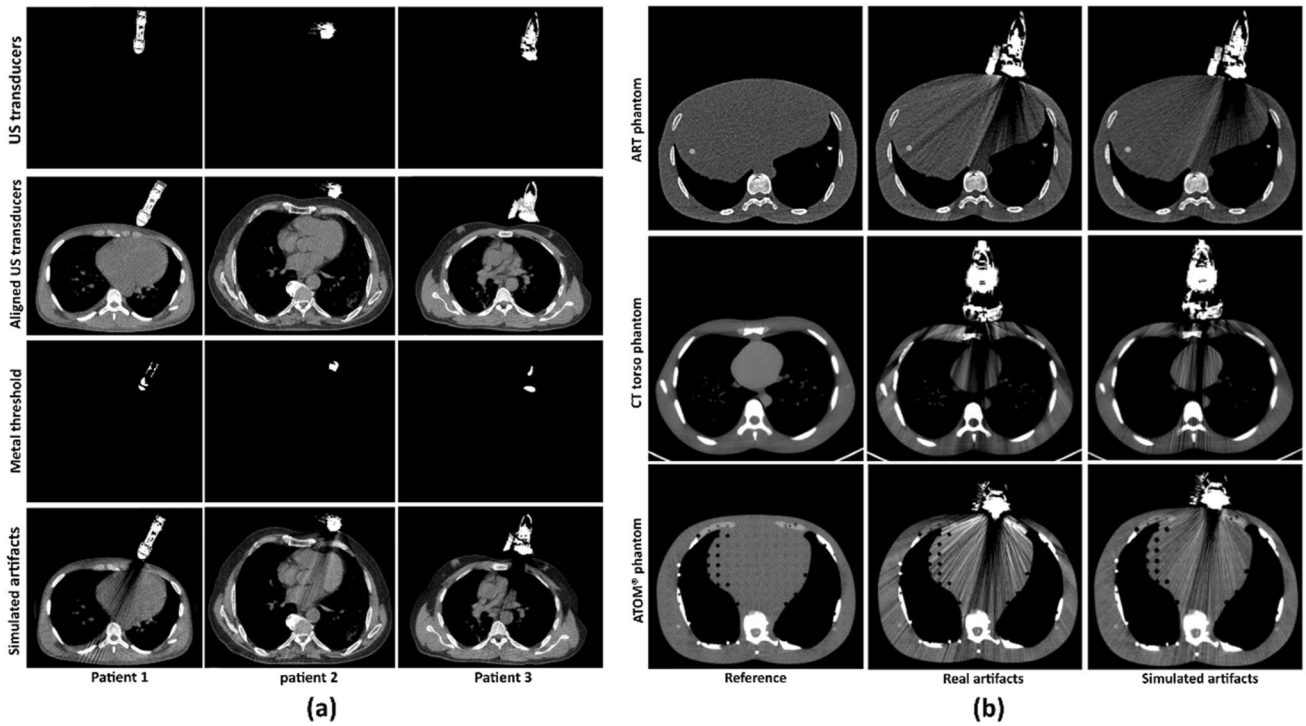


Fig. 1 **a** The procedure for simulation of US transducer-induced metal artifact on clinical CT scans. The first and second row show the segmented US transducers from the phantom CT scans and the aligned US transducers with the suitable clinical CT slices, respectively. Pixels containing metal parts of a US transducer and the CT slices consisting of simulated artifacts are shown in the last two rows, respectively. **b** Visual representation of the real artifacts and the simulated artifacts on the phantom scans. [Window level/ width: 50/ 350]

projection data for corresponding X-ray energies were simulated from the segmented bone, lung, water equivalent tissue and from the metal binary image. As the metal components of a US transducer consist of lead, zirconate, and titanate [50], the average mass attenuation coefficient value of these metals was used to generate the projection data. Consequently, metal-containing projection data was created from those simulated projections with Poisson distribution for the reconstruction of a CT_{art} slice.

To check the correctness of this artifact simulation method, phantom CT scans were utilized (See Fig. 1 b). From a particular phantom CT scan, initially, CT_{ref} and the corresponding CT_{art} slices were selected. Then, the US transducer-induced metal artifacts were simulated on the phantom CT_{ref} slices based on the procedure described above. The simulated CT_{art} slices were visually validated against the corresponding phantom CT_{art} slices comparing them to the real US transducer-induced metal artifacts.

Phantom CT scans

Table 1 shows the combinations of CT scanners, anthropomorphic phantoms, and US transducers used in this work. In particular, three types of adult anthropomorphic phantoms were used to scan with and without a total of four types of US transducers. The utilized anthropomorphic phantoms were an ART-211 male phantom (ART, Radiology Support Devices, Long Beach, CA, USA); an ATOM® male phantom (CIRS, Model-701, Norfolk, VA, USA); and a CT torso phantom (CT Torso, Model CTU-41, Kyoto Kagaku Ltd, Japan). These anthropomorphic phantoms were constructed using tissue-equivalent epoxy materials that mimics the density and attenuation characteristics of human tissues. They include a range of components, such as cardiac structures, air-equivalent materials for simulating lungs, and bone-like materials with simulated air pockets. To obtain the paired CT scans, each phantom was CT scanned with and without a US transducer, resulting in a CT_{art} scan and the corresponding CT_{ref} scan, respectively (See Fig. 2 for an example of the procedure). The US transducers were positioned on the phantoms at various angles to be suitable for proper imaging of the heart. As the dimensions of US transducers, including the size and width of their metal components, can have an impact on the creation of metal artifacts on CT scans. Among the US transducers used, the linear volume array transducer was the largest and had the widest metal component, which was measured to be 6 cm using CT

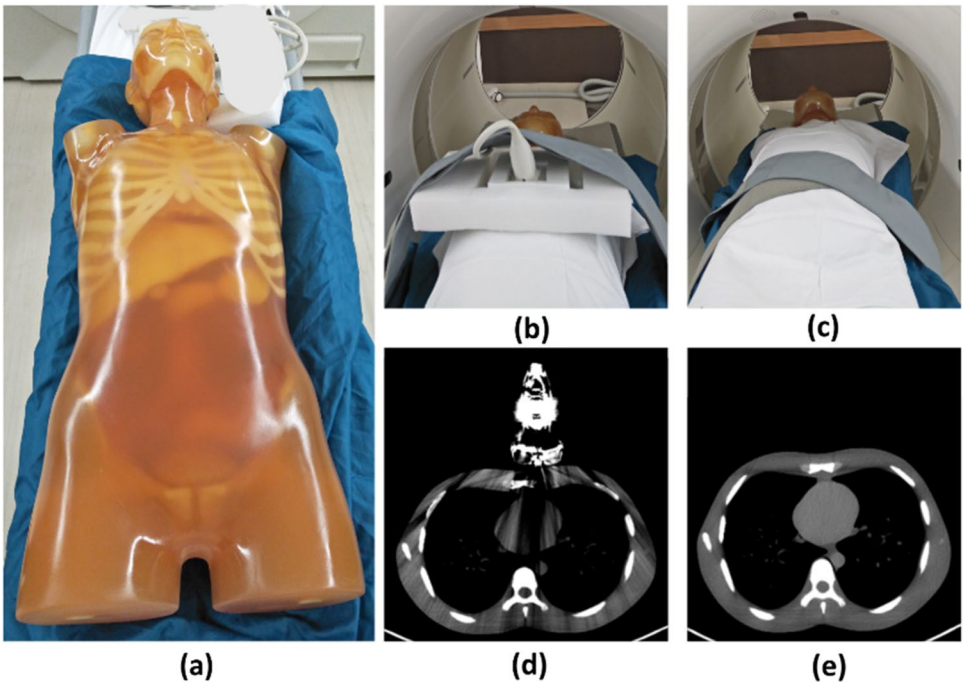


Fig. 2 Workflow to obtain the dataset pairs: in (**a**) the CT torso phantom is shown. The positioning of the phantom with (**b**) and without (**c**) the linear volume array transducer resulted in CT scans as shown in (**d**) and (**e**)

Table 1 The details of utilized CT scanners, anthropomorphic phantoms, and the US transducers

CT scanner (Model)	Clinical site	Phantom	US transducers
Siemens CT (SOMATOM Definition AS)	Kantonnspital Aarau, Aarau, Switzerland	ART-211	Single-plane phased array, Bi-plane phased array
Philips CT (Brilliance Big Bore)	Geneva University Hospital, Geneva, Switzerland	ATOM®	Single-plane phased array, Bi-plane phased array
Siemens PET-CT (Biograph 128)	Herston Imaging Research Facility (HIRF), Brisbane, Australia	CT torso	Linear array, Linear volume array

Single-plane phased array: Telemed P5-1S15-A6 from Telemed (UAB, Vilnius, Lithuania); bi-plane phased array: Terason XY mini from Teratech Corporation (Burlington, MA, USA); linear array: Philips L7-4 Philips Healthcare (Andover, MA, USA) and linear volumetric array: Philips VL13-5

scans, and induced a greater quantity of metal artifacts compared to the other transducers.

Cycle-MAR network

The CycleGAN model [43] has been proposed in literature for unpaired training data. However, we used paired data in this work to enforce the restoration accuracy of anatomical structures and HU values [51]. The workflow of the developed Cycle-MAR network is illustrated in Fig. 3.

The CycleGAN translates the metal artifact domain (X) into an artifact-free domain (Y) by using adversarial loss ($\mathcal{L}_{adv}$), cycle consistency loss $\mathcal{L}_{(cycle)}$, and identity loss $\mathcal{L}_{(identity)}$. This process includes two mapping functions, $G_X : X \rightarrow Y$ and $G_Y : Y \rightarrow X$. The mapping function G_X translates a CT_{art} slice into a CT_{cor} slice, whereas G_Y translates a CT_{ref} slice into CT_{art} slice. The network also consists of two adversarial discriminators D_X and D_Y which aim to distinguish the translated domain as fake. D_X aims to distinguish between x from $G_X(y)$ and D_Y aims to distinguish y from $G_Y(x)$. For G_Y, the $(\mathcal{L}_{adv})$ is the mean squared error (MSE) between output $G_Y(x)$ and target domain Y. The $\mathcal{L}_{(cycle)}$ calculates the translation error between $x$and $G_X\big(G_Y(x)\big)$ through the translation of $X \rightarrow Y \rightarrow X$ and in vice versa of $Y \rightarrow X \rightarrow Y$. The $\mathcal{L}_{(identity)}$ was introduced to regularize the G_X and G_Y to not induce any changes when x and y were the input for them, respectively.

In this study, regularization parameters values of $\lambda_{cycle} = 10$ for, and$\lambda_{identity} = 15$ were chosen among several examined parameter sets. To implement Cycle-MAR, the ResNet [52] and the PatchGAN [42] architectures were used as the generator and the discriminator, respectively. The network was trained using the Adam optimizer [53]

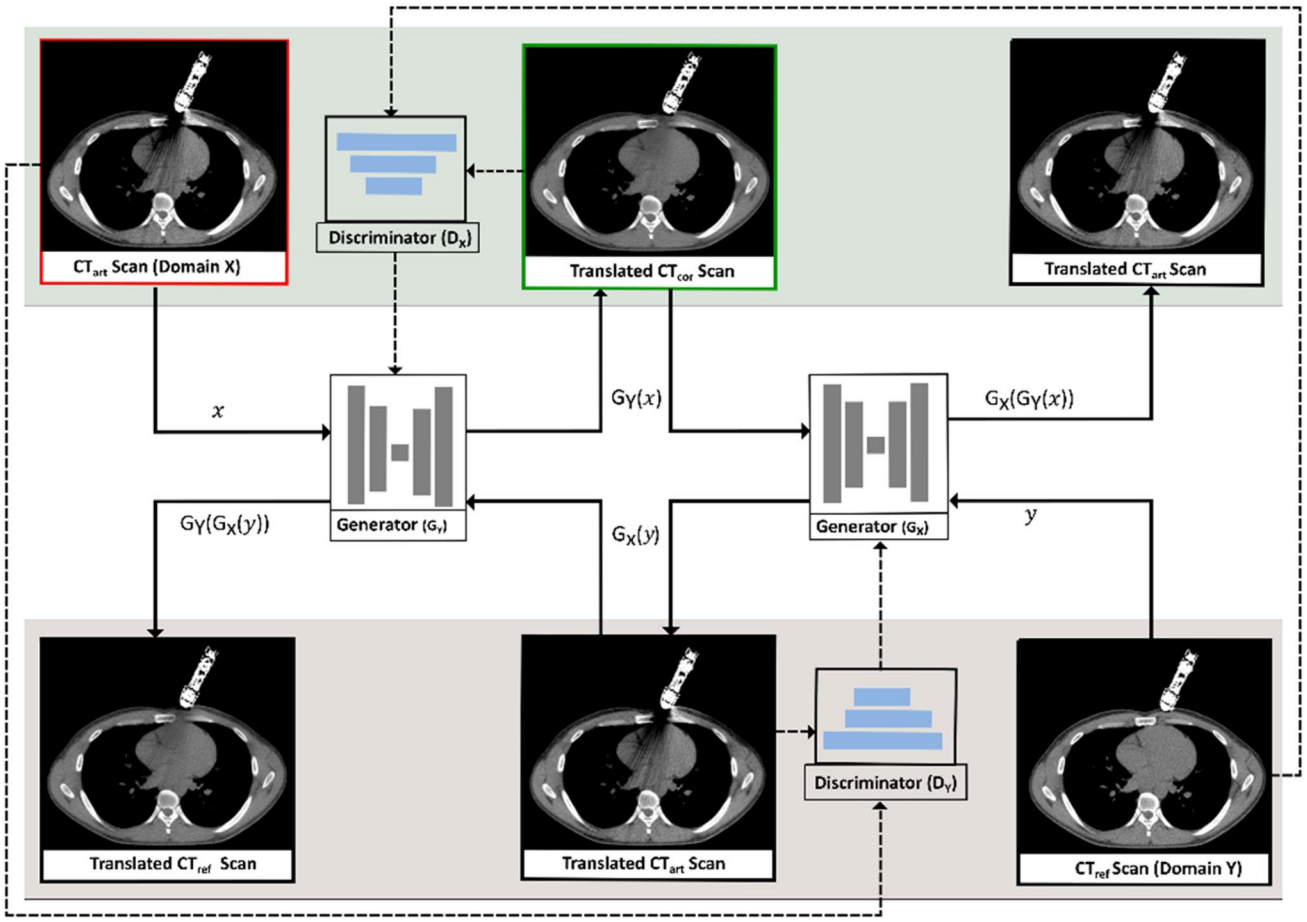


Fig. 3 Training workflow of the Cycle-MAR network for the reduction of the US transducer-induced metal artifacts. It has two mapping functions: generator (G_Y) transforms the CT_{art} scan (domain X) into CT_{cor} scan, while generator (G_X) transforms the CT_{ref} scan (domain Y) into an CT_{art} scan. Discriminator (D_X) and discriminator (D_Y) aim to distinguish translated CT_{ref} scan from the CT_{ref} (domain Y) and distinguish the translated CT_{art} scan from the CT_{art} scan (domain X), respectively

for 500 epochs using a batch size of 1 which produced the best results and validated with five-fold cross-validation. The cross-validation with five-fold was chosen based on the previous studies [29, 33, 54]. To avoid overfitting, data augmentation methods of rotation, horizontal flip, resized crop and image perspective were utilized during the training session. The Cycle-MAR network was implemented using Python 3.8.8, with the PyTorch (version 1.10.2) framework. A NVIDIA (TITAN RTX) GPU with 25GB was used throughout the experiments.

Table 2 Data split strategy for training and testing of the Cycle-MAR network

Data	Data split	Patient/phantom	Number of CT slices (%)
Clinical CT scans	Training (paired CT_{ref} & CT_{art})	70	3515 (85)
	Testing (CT_{art})	14	650 (15)
Phantom CT scans	Testing (CT_{art})	ART	20 (41.1)
		ATOM®	23 (31.5)
		CT torso	30 (27.4)

Data split for training and testing

Table 2 shows the data split strategy for the training and testing of Cycle-MAR. The network was trained using randomly selected paired clinical CT scans consisting of CT_{ref} scans and the corresponding simulated CT_{art} scans. Then it was tested on the phantom CT scans with real artifacts, and on the clinical CT scans with simulated artifacts. The HU

values on the clinical CT scans were clipped between − 1000 and 1000, and the remaining pixel values were normalized between − 1 and 1 to improve the training efficiency [28].

Comparison with commercial and research-based MAR algorithms

The performance of the developed Cycle-MAR network was compared with commercially available MAR algorithms and research-based MAR algorithms. iMAR and O-MAR were directly applied during the reconstruction of the phantom CT scans by the scanners. iMAR was not applied to the Siemens PET-CT scans, because it was not available on this particular scanner. In addition to this, MDT, CCS-MAR, and Cycle-MAR were also applied to all phantom and clinical CT scans.

Image quality metrics analysis

Structural similarity (SSIM) index, root mean square error (RMSE) of the HU values, and peak signal-to-noise ratio (PSNR) [38, 55] were calculated to evaluate the performance of the Cycle-MAR network for metal artifact reduction and image quality improvement. These metrics were calculated for the CT_{art} and the CT_{cor} scans compared to the CT_{ref} scans. The analysis was performed using the overall mean values of these image quality metrics calculated for all CT_{art} and CT_{cor} scans.

HU value restoration evaluation

For the clinical CT scans, HU value measurements on specific regions were performed on the CT_{art} scans, CT_{cor} scans and the CT_{ref} scan. The contour-based mean HU values and standard deviation (STD) were calculated for the entire heart, lungs, and bone regions from all CT slices using MatLab (The MathWorks Inc, USA) (See Fig. 5). The percentage of mean HU value improvement for CT_{cor} scans was calculated using the following equation,

$$\text{Mean HU value improvement}\,(\%) = \frac{|\Delta HU|_{CT_{ref},CT_{art}} - |\Delta HU|_{CT_{ref},CT_{cor}}}{|\Delta HU|_{CT_{ref},CT_{art}}} \times 100\% \quad (1)$$

where $|\Delta HU|_{CT_{ref},CT_{art}}$ and $|\Delta HU|_{CT_{ref},CT_{cor}}$ indicate the absolute difference in the contour-based mean HU values between the CT_{ref} scans and the corresponding CT_{art} and CT_{cor} scans, respectively.

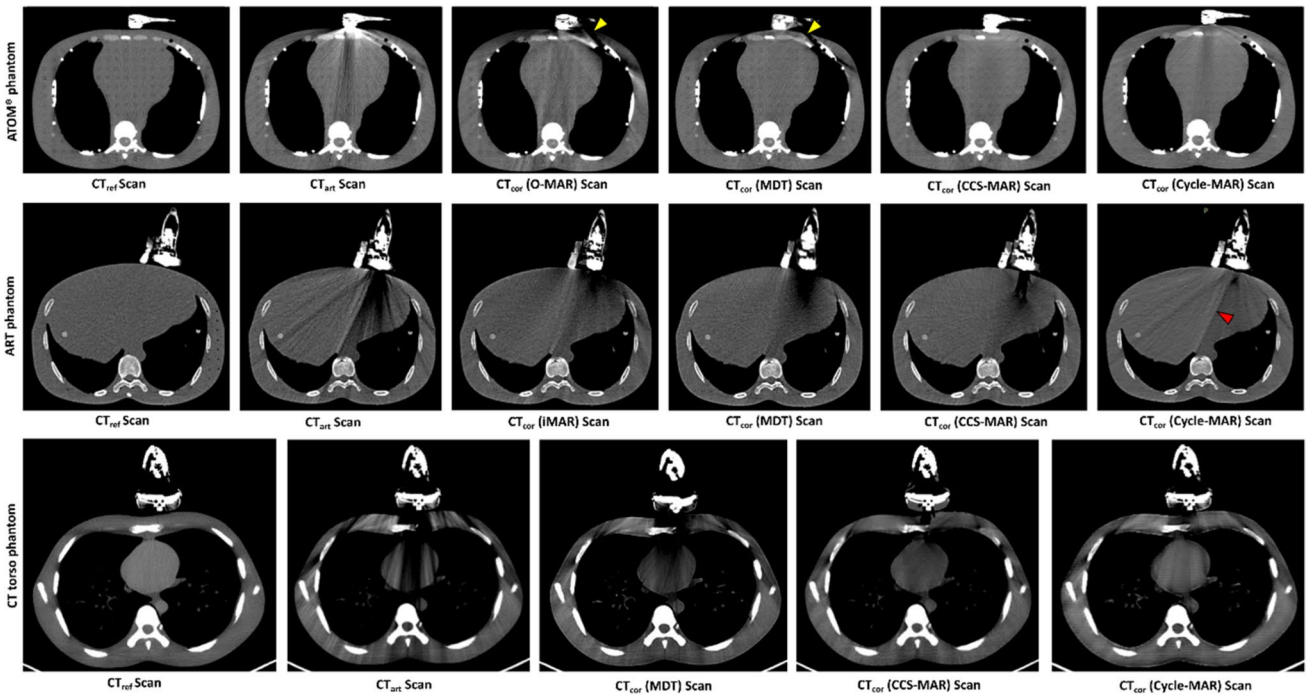


Fig. 4 From top row to bottom: CT scans of the ATOM®, ART, and CT torso phantoms. The images from left to right show: the CT_{ref} scan with the US transducer details, CT_{art} scan, and the CT_{cor} scans after application of the MAR algorithms [Window level/width: 50/350]. The yellow marks indicate induced secondary artifacts, while the red mark indicates residual streaks artifacts, respectively

Results

Phantom scans analysis

Figure 4 shows the CT_{ref} scans, and the corresponding CT_{art} and CT_{cor} scans from the ART, ATOM®, and CT torso phantoms. In general, Cycle-MAR outperformed other MAR algorithms to reduce the intense dark or bright regions near the US transducer during the visual inspection. Residual streak artifacts were observed on CT_{cor} after the Cycle-MAR application, especially in the ART phantom scans (red mark in Fig. 4). In the ATOM® phantom scans, O-MAR and MDT applications induced secondary dark streak artifacts (yellow marks in Fig. 4). The Cycle-MAR application on the CT_{art} scans generally improved the calculated SSIM and PSNR values, while the RMSE values were decreased (Table 3).

Clinical scans analysis

An example of a CT_{art} scans from three randomly selected patients and the effect of MDT, CCS-MAR, and CycleGAN applications on them for metal artifact reduction is shown in Fig. 5. Based on the visual inspection, Cycle-MAR application on CT_{art} scans restored the soft tissue and bone details better than the MDT and CCS-MAR algorithms (red and green arrows in Fig. 5). The overall mean values of SSIM, PSNR, and RMSE for all clinical CT scans from 14 patients are shown in Table 4. Cycle-MAR application on CT_{art} scans generally resulted in higher mean SSIM and PSNR values and lower mean RMSE values.

Table 5 shows the overall mean (±STD) HU values for the CT_{ref} Scans, and the calculated absolute differences between the overall mean and the differences of standard deviation (STD) of HU values for the CT_{art} Scans, and CT_{cor} Scans compared to CT_{ref} scans for the heart, lungs, and bone regions on the clinical CT scans. The application of MAR algorithms improved the HU value measurements across all regions. Cycle-MAR application restored the mean HU values for the heart and lung region better than MDT and CCS-MAR.

The percentage of mean HU value improvement for the heart, lungs and bone regions is shown in Fig. 6. For the heart region, in which typically the target is located, the improved HU value percentage after MDT, CCS-MAR, and Cycle-MAR applications was 59.58%, 62.22%, and 72.84%, respectively. The regions of lungs and bone were considered as OARs in this study, for these regions, the highest improvement percentage was found after the application of Cycle-MAR and MDT, respectively.

Discussion

In this study, a Cycle-MAR algorithm which used paired CT scans for training purposes was proposed to reduce the US transducer-induced metal artifacts on planning CT scans for US-guided cardiac radioablation. Cycle-MAR was evaluated for the improvement of image quality and HU value restoration compared to the commonly used commercial and research-based MAR algorithms.

Overall, the proposed model effectively reduced the metal artifacts on the clinical CT scans more than on the phantom CT scans. For the Cycle-MAR training, only clinical CT scans, and no phantom CT scans were used. This might be a reason for the noticeable residual streaks on the phantom CT_{cor} scans after the Cycle-MAR application (See Fig. 4). This potentially can be improved by adding separate sets of phantom scans to the algorithm training set. On the other hand, in the end, the algorithm will be used in the clinic

Table 3 Mean values of SSIM, PSNR and RMSE for the CT scans of the ATOM®, ART and CT torso phantoms

Image quality metrics	Phantom CT scans	CT_{art} scan	CT_{cor} Scans				
			iMAR	O-MAR	MDT	CCS-MAR	Cycle-MAR
Mean (±STD)							
SSIM	**ATOM®**	0.71 (±0.12)	*N/A*	0.86 (±0.05)	0.95 (±0.07)	0.95 (±0.08)	**0.95 (±0.04)**
	ART	0.65 (±0.19)	0.85 (±0.08)	*N/A*	0.85 (±0.04)	0.86 (±0.06)	**0.87 (±0.05)**
	CT torso	0.60 (±0.27)	*N/A*	*N/A*	0.81 (±0.14)	0.84 (±0.17)	**0.88 (±0.11)**
PSNR (dB)	**ATOM®**	19.02 (±0.43)	*N/A*	29.02 (±1.40)	31.03 (±1.29)	31.38 (±1.02)	**32.48 (±0.26)**
	ART	16.41 (±2.50)	24.99 (±1.83)	*N/A*	25.11 (±1.57)	**25.6 (±1.65)**	25.58 (±1.31)
	CT torso	18.56 (±3.70)	*N/A*	*N/A*	26.31 (±1.70)	26.75 (±1.93)	**27.39 (±1.10)**
RMSE	**ATOM®**	73.8 (±9.90)	*N/A*	42.52 (±5.83)	32.33 (±6.24)	29.96 (±4.84)	**28.11 (±5.70)**
	ART	87.94 (±6.72)	33.48 (±3.36)	*N/A*	32.97 (±3.38)	31.21 (±4.21)	**30.59 (±2.65)**
	CT torso	114.13 (±9.71)	*N/A*	*N/A*	63.73 (±5.43)	55.67 (±4.43)	**53.33 (±3.40)**

The best performance is indicated with bold numbers

STD Standard Deviation

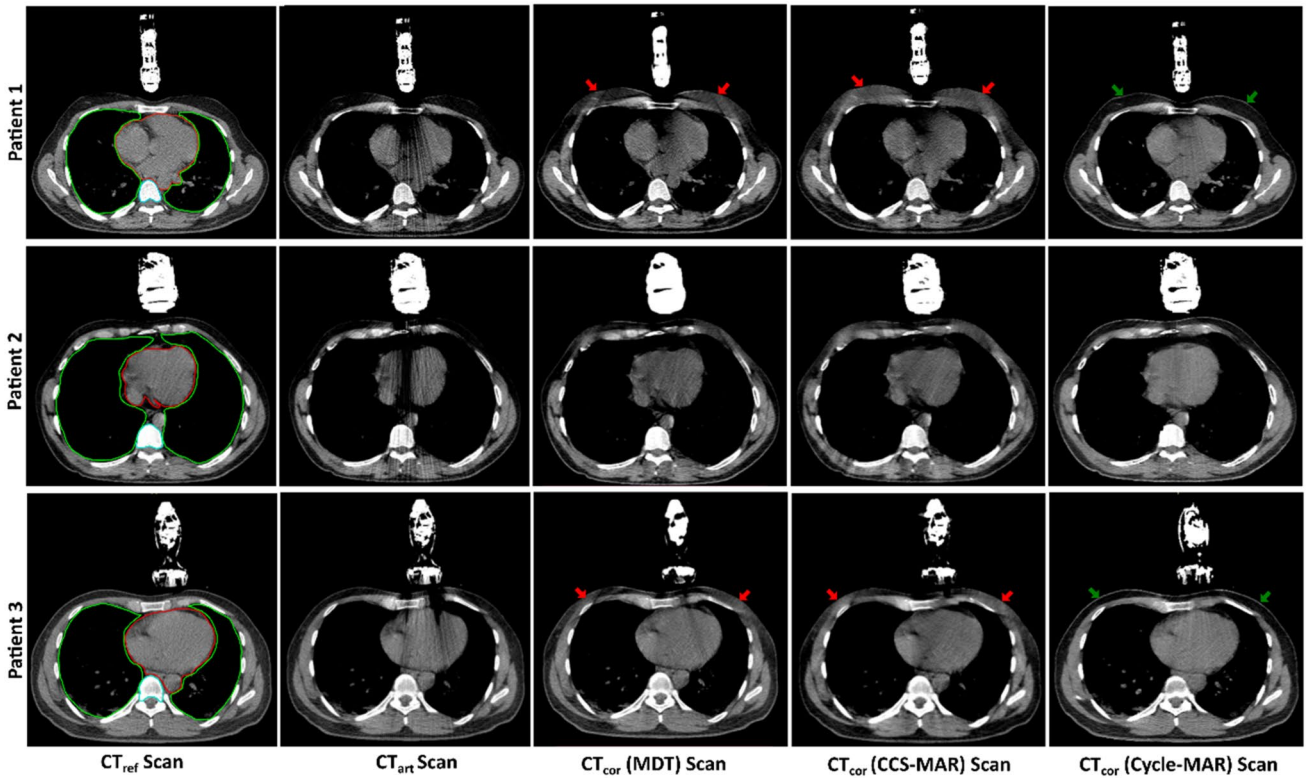


Fig. 5 CT scans from three patients. The images from left to right show: the CT_{ref} scan with the US transducer details, CT_{art} scan, and the CT_{cor} scans after MDT, CCS-MAR, and Cycle-MAR applications [Window level/width: 50/350]. The red and green arrows indicate the changes of soft tissue density. The counters for the HU value measurements were shown on the CT_{ref} scans: the heart (red), bone (blue) and lungs (green)

Table 4 The overall calculated mean values of SSIM, PSNR and RMSE values for all clinical CT scans from 14 patients

Image quality metrics	CT_{art}scan	CT_{cor} Scans		
		MDT	CCS-MAR	Cycle-MAR
Mean (±STD)				
SSIM	0.75 (±0.21)	0.91 (±0.08)	0.91 (±0.09)	**0.92 (±0.19)**
PSNR (dB)	17.57 (±3.08)	25.86 (±2.10)	25.21 (±2.04)	**26.32 (±1.80)**
RMSE	84.83 (±11.25)	33.85 (±10.05)	32.63 (±9.09))	**30.86 (±8.80)**

The best performance is indicated with bold numbers

STD Standard Deviation

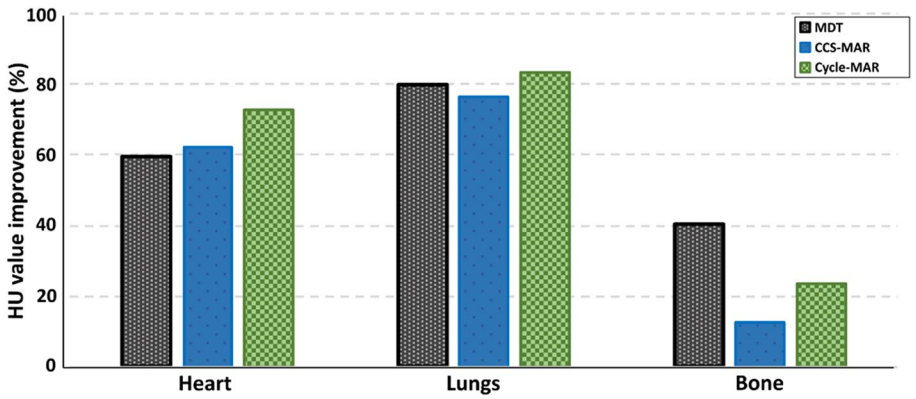


Fig. 6 HU values improvement percentage for the heart, lungs and bone regions on the CT_{cor} scans after MDT, CCS-MAR and Cycle-MAR application

Table 5 The overall region-based Hounsfield unit (HU) value measurements and calculation using all clinical CT scans from 14 patients

Region	CT_{ref}Scans	CT_{art}Scans	CT_{cor}Scans		
			MDT	CCS-MAR	Cycle-MAR
	Mean (±STD) HU	|Δ mean| (Δ STD) HU			
Heart	33.34 (±37)	46.55 (15)	20.10 (3)	20.32 (2)	9.18 (7)
Lungs	− 675.00 (±174)	30.01 (36)	6.00 (11)	7.01 (16)	5.00 (6)
Bone	233.00 (±64)	40.69 (9)	29.00 (1)	35.53 (1)	31.04 (2)

|Δ mean| (± Δ STD) HU represents the absolute differences of overall mean and the differences of STD between CT_{ref}Scans and the corresponding CTart $_{art}$ Scans, and as well as between CT_{ref} Scans and CT_{cor} Scans, respectively

STD Standard Deviation

and therefore good performance on phantom scans is less important. To further reduce the residual streaks, the CT_{cor} scans which resulted from the training of Cycle-MAR may be added to the training data set and this is under consideration for future work.

All MAR algorithms restored the measured HU values for the cardiac structures within 21 HU, which is well below the tolerance accuracy of 30 HU for the waterlike material recommended by the American Association of Physicists in Medicine (AAPM) guidelines [56] for image-guided radiation therapy. Remarkably, the same guidelines recommend that the HU value deviation for the lung and bone be within 50 HU.

The application of conventional MAR algorithms on CT_{art} scans modified the anatomical structures and induced a number of secondary artifacts (See Figs. 4 and 5). These MAR algorithms apply their correction on the projection data, therefore, small errors in local corrections in the projection data can affect the reconstructed CT scan globally [12]. However, Cycle-MAR works in the image space and does not require any projection data for artifact corrections. This means that the local changes are applied only to a specific area on the CT scan.

To the best of our knowledge, this is the first study to investigate the application of a deep learning model, CycleGAN, for the reduction of US transducer-induced metal artifact on CT scans which has been compared to state-of-the-art MAR algorithms. Even though Cycle-MAR generally well reduced the metal artifacts compared to other MAR algorithms, especially in the clinical CT scans, a reduction in image contrast was observed on the CT_{cor} scans after Cycle-MAR application. A possible reason for this is the inherent limitation of the generator in the conversion and/or reassignment accuracy of pixel values while performing feature extraction and the image translation process. In addition, the direct optimization in pixel differences through the loss function may also result in reduced image contrast or blurry appearance on CT scans [31, 57]. Therefore, investigating a different generator, especially DenseNet (Densely Connected Network) [58, 59] instead of ResNet, and also examining appropriate loss functions may solve this issue.

This work has a few limitations: the performance of Cycle-MAR was evaluated using clinical CT scans with the simulated US-transducer-induced metal artifacts. To draw final conclusions regarding the performance of the proposed MAR algorithm, further evaluation using clinical data with real artifacts is necessary. However, this can be a challenging task due to ethical justification of acquiring an additional CT scan [one with the probe in place (CT_{art}) and another one without it (CT_{ref})].

In addition to the evaluation of image quality improvement and HU value restoration, the dosimetric impact of the metal artifact reduction including the accuracy in contouring, and the calculation of dose distribution for the arrhythmogenic tissue (target) and OARs during the treatment planning is also crucial. Future work will therefore include an evaluation of the dosimetric impacts of the application of the Cycle-MAR network.

Conclusion

This work developed a MAR network based on a deep learning CycleGAN which can be used to reduce metal artifacts resulting from the presence of a US transducer during CT scan acquisition. The performance of the proposed algorithm was evaluated for the metal artifact reduction abilities on phantom and clinical CT scans in comparison with commonly used commercial and research-based MAR algorithms. The results of the study have shown that the proposed Cycle-MAR considerably reduces the metal artifacts, while preserving the bone density and soft tissue details. Future challenges and analysis include exploring appropriate loss functions for the improvement of adversarial training, and dosimetric evaluations using clinical CT scans.

Acknowledgements The authors would like to thank F. Edward Boas, MD, Ph.D. for providing the MDT algorithm, and Giovanna Dipasquale and Nikolaos Koutsouvelis from the Division of Radiation Oncology at

the Geneva University Hospital as well as Nicoletta Lomax and Roger Hälg from Aarau Kantonspital for their support during the data acquisition of the ART and ATOM® phantoms for this study. The authors are grateful to Peta Gray from the Herston Medical Research Institute (HIRF) for providing support to collect data from the CT torso phantom. Finally, the authors also acknowledge the important contribution of Damjan Vukovic from the Queensland University of Technology for providing assistance to access the GPU system.

Author contributions Study conception and design: SP, SC, DF. Statistical analysis and interpretation of data: SP, SC, MLW, DF. Drafting the article: SP. Paper review and editing: SC, MLW, DF.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. This publication is part of a project that has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 954783. The data acquisition at the Geneva University Hospital, Geneva, Switzerland, was supported by Innosuisse project no. 40854.1 INNO-LS. Data collection at HIRF, Queensland, Australia was funded by a grant under the HIRF Project Support Scheme 2020.

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

Ethical approval Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. van der Ree MH, Visser J, Planken RN, Dieleman EMT, Boekholdt SM, Balgobind BV et al (2022) Standardizing the cardiac radioablation targeting workflow: enabling semi-automated angulation and segmentation of the heart according to the American heart association segmented model. Adv Radiat Oncol 7(4):100928. https://doi.org/10.1016/j.adro.2022.100928
2. Siedow M, Brownstein J, Prasad RN, Loccoh E, Harfi TT, Okabe T et al (2021) Cardiac radioablation in the treatment of ventricular tachycardia. Clin Transl Radiat Oncol 31:71–79. https://doi.org/10.1016/j.ctro.2021.02.005
3. Suzanne Lydiard PGD, Blanck O, Hugo G, O'Brien R, Keall P (2021) A review of cardiac radioablation (CR) for arrhythmias: procedures, technology, and future opportunities. Int J Radiat Oncol Biol Phys 109(3):783–800. https://doi.org/10.1016/j.ijrobp.2020.10.036
4. Lydiard S, Pontré B, Lowe BS, Ball H, Sasso G, Keall P (2021) Cardiac radioablation for atrial fibrillation: target motion characterization and treatment delivery considerations. Med Phys 48(3):931–941
5. Calkins H, Hindricks G, Cappato R, Kim YH, Saad EB, Aguinaga L et al (2018) 2017 HRS/EHRA/ECAS/APHRS/SOLAECE expert consensus statement on catheter and surgical ablation of atrial fibrillation: executive summary. Eur Eur pacing, arrhythmias, Card Electrophysiol J Work groups Card pacing, arrhythmias, Card Cell Electrophysiol Eur Soc Cardiol 20(1):157–208
6. Casula M, Dusi V, Camps S, Gringet J, Benoit T, Garonna A et al (2022) Feasibility of an automatic ultrasonographic image acquisition system associated with an artificial intelligence algorithm for real-time monitoring of cardiac motion during cardiac radioablation. Front Cardiovas Med. https://doi.org/10.3389/fcvm.2022.849234
7. Perrin R, Maguire P, Garonna A, Weidlich G, Bulling S, Fargier-Voiron M et al (2022) Case report: treatment planning study to demonstrate feasibility of transthoracic ultrasound guidance to facilitate ventricular tachycardia ablation with protons. Front Cardiovasc Med. https://doi.org/10.3389/fcvm.2022.849247
8. Ipsen S, Bruder R, García-Vázquez V, Schweikard A, Ernst F (2019) Assessment of 4D ultrasound systems for image-guided radiation therapy - image quality, framerates and CT artifacts. Curr Dir Biomed Eng 5(1):245–248
9. Schlosser J, Hristov D (2016) Radiolucent 4D ultrasound imaging: system design and application to radiotherapy guidance. IEEE Trans Med Imaging 35(10):2292–2300
10. Giantsoudi D, De Man B, Verburg J, Trofimov A, Jin Y, Wang G et al (2017) Metal artifacts in computed tomography for radiation therapy planning: dosimetric effects and impact of metal artifact reduction. Phys Med Biol 62(8):R49–R80
11. Puvanasunthararajah S, Fontanarosa D, Wille ML, Camps SM (2021) The application of metal artifact reduction methods on computed tomography scans for radiotherapy applications: a literature review. J Appl Clin Med Phys 22(6):198–223
12. Katsura M, Sato J, Akahane M, Kunimatsu A, Abe O (2018) Current and novel techniques for metal artifact reduction at CT: practical guide for radiologists. Radiographics 38(2):450–461
13. Andersson KM, Nowik P, Persliden J, Thunberg P, Norrman E (2015) Metal artefact reduction in CT imaging of hip prostheses—an evaluation of commercial techniques provided by four vendors. Br J Radiol 88(1052):20140473
14. Andersson KM, Dahlgren CV, Reizenstein J, Cao Y, Ahnesjö A, Thunberg P (2018) Evaluation of two commercial CT metal artifact reduction algorithms for use in proton radiotherapy treatment planning in the head and neck area. Med Phys 45(10):4329–4344
15. Axente M, Paidi A, Von Eyben R, Zeng C, Bani-Hashemi A, Krauss A et al (2015) Clinical evaluation of the iterative metal artifact reduction algorithm for CT simulation in radiotherapy. Med Phys 42(3):1170–1183
16. Kwon H, Kim KS, Chun YM, Wu HG, Carlson JNK, Park JM et al (2015) Evaluation of a commercial orthopaedic metal artefact reduction tool in radiation therapy of patients with head and neck cancer. Br J Radiol 88(1052):13–22
17. Bär E, Schwahofer A, Kuchenbecker S, Häring P (2015) Improving radiotherapy planning in patients with metallic implants using the iterative metal artifact reduction (iMAR) algorithm. Biomed Phys Eng Express 1(2):025206
18. Maerz M, Mittermair P, Krauss A, Koelbl O, Dobler B (2016) Iterative metallartefakt reduktion verbessert die Genauigkeit der Dosisberechnung: Phantomstudie mit Zahnimplantaten. Strahlentherapie und Onkol 192(6):403–413
19. Huang JY, Followill DS, Howell RM, Liu X, Mirkovic D, Stingo FC et al (2016) Approaches to reducing photon dose calculation errors near metal implants. Med Phys 43(9):5117–5130

20. Righetto R, Clemens LP, Lorentini S, Fracchiolla F, Algranati C, Tommasino F et al (2020) Accurate proton treatment planning for pencil beam crossing titanium fixation implants. Phys Medica 70:28–38. https://doi.org/10.1016/j.ejmp.2020.01.003
21. Akdeniz Y, Yegingil I, Yegingil Z (2019) Effects of metal implants and a metal artifact reduction tool on calculation accuracy of AAA and Acuros XB algorithms in small fields. Med Phys 46(11):5326–5335
22. Boas FE, Fleischmann D (2011) Evaluation of two iterative techniques for reducing metal artifacts in computed tomography. Radiology 259(3):894–902
23. Odlozilikova A, Kurzyukova A, Sepsi M, Pospisil D, Slampa P (2018) Použití metody metal deletion technique pro odstraňování artefaktů při plánování radioterapie u pacientů s implantovanými kardiostimulátory. Klin Onkol 31(6):434–438
24. Borges FK, Bhandari M, Guerra-Farfan E, Patel A, Sigamani A, Umer M et al (2020) Accelerated surgery versus standard care in hip fracture (HIP ATTACK): an international, randomised, controlled trial. Lancet 395(10225):698–708
25. Wagenaar D, Van Der Graaf ER, Van Der Schaaf A, Greuter MJW (2015) Quantitative comparison of commercial and non-commercial metal artifact reduction techniques in computed tomography. PLoS One 10(6):1–9. https://doi.org/10.1371/journal.pone.0127932
26. Wellenberg RHH, Hakvoort ET, Slump CH, Boomsma MF, Maas M, Streekstra GJ (2018) Metal artifact reduction techniques in musculoskeletal CT-imaging. Eur J Radiol 107:60–69. https://doi.org/10.1016/j.ejrad.2018.08.010
27. Puvanasunthararajah S, Camps SM, Wille ML, Fontanarosa D (2022) Combined clustered scan-based metal artifact reduction algorithm (CCS-MAR) for ultrasound-guided cardiac radioablation. Phys Eng Sci Med 45(4):1273–1287. https://doi.org/10.1007/s13246-022-01192-6
28. Nakao M, Imanishi K, Ueda N, Imai Y, Kirita T, Matsuda T (2020) Regularized three-dimensional generative adversarial nets for unsupervised metal artifact reduction in head and neck CT images. IEEE Access 8:109453–109465
29. Koike Y, Anetai Y, Takegawa H, Ohira S, Nakamura S, Tanigawa N (2020) Deep learning-based metal artifact reduction using cycle-consistent adversarial network for intensity-modulated head and neck radiation therapy treatment planning. Phys Medica 78:8–14. https://doi.org/10.1016/j.ejmp.2020.08.018
30. Liao H, Lin WA, Zhou SK, Luo J (2020) ADN: Artifact Disentanglement Network for unsupervised metal artifact reduction. IEEE Trans Med Imaging 39(3):634–643
31. Nakamura M, Nakao M, Imanishi K, Hirashima H, Tsuruta Y (2021) Geometric and dosimetric impact of 3D generative adversarial network-based metal artifact reduction algorithm on VMAT and IMPT for the head and neck region. Radiat Oncol 16(1):96. https://doi.org/10.1186/s13014-021-01827-0
32. Zhu L, Han Y, Xi X, Li L, Yan B (2021) Completion of metal-damaged traces based on deep learning in sinogram domain for metal artifacts reduction in CT images. Sensors 21(24):8164
33. Xu L, Zhou S, Guo J, Tian W, Tang W, Yi Z (2022) Metal artifact reduction for oral and maxillofacial computed tomography images by a generative adversarial network. Appl Intell 52(11):13184–13194. https://doi.org/10.1007/s10489-021-02905-2
34. Wang T, Xia W, Huang Y, Sun H, Liu Y, Chen H, et al. (2021) Dual-Domain Adaptive-Scaling Non-local Network for CT Metal Artifact Reduction. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics). 12906 LNCS:243–253.
35. Shi Z, Wang N, Kong F, Cao H, Cao Q (2022) A semi-supervised learning method of latent features based on convolutional neural networks for CT metal artifact reduction. Med Phys 49(6):3845–3859. https://doi.org/10.1002/mp.15633
36. Huang Z, Zhang G, Lin J, Pang Y, Wang H, Bai T et al (2022) Multi-modal feature-fusion for CT metal artifact reduction using edge-enhanced generative adversarial networks. Comput Methods Programs Biomed 217:106700
37. Lin WA, Liao H, Peng C, Sun X, Zhang J, Luo J et al (2019) DuDoNet: dual domain network for CT metal artifact reduction. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit 2019–June:10504–10513
38. Zhang Y, Yu H (2018) Convolutional neural network based metal artifact reduction in X-Ray computed Tomography. IEEE Trans Med Imaging 37(6):1370–1381
39. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O et al (2021) Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data 8(1):53. https://doi.org/10.1186/s40537-021-00444-8
40. Sahiner B, Pezeshk A, Hadjiiski LM, Wang X, Drukker K, Cha KH et al (2019) Deep learning in medical imaging and radiation therapy. Med Phys 46(1):e1–e36
41. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S et al (2020) Generative adversarial networks. Commun ACM 63(11):139–144
42. Isola P, Zhu JY, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. Proc - 30th IEEE Conf Comput Vis Pattern Recognition, CVPR 2017. 2017:5967–5976.
43. Zhu JY, Park T, Isola P, Efros AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. Proc IEEE Int Conf Comput Vis 2017–Octob:2242–2251
44. Mostafavi SM. COVID19-CT-Dataset: An Open-Access Chest CT Image Repository of 1000+ Patients with Confirmed COVID-19 Diagnosis. V1 ed. Harvard Dataverse https://doi.org/10.7910/DVN/6ACUZJ
45. Shakouri S, Bakhshali MA, Layegh P, Kiani B, Masoumi F, Ataei Nakhaei S et al (2021) COVID19-CT-dataset: an open-access chest CT image repository of 1000+ patients with confirmed COVID-19 diagnosis. BMC Res Notes 14(1):1–3. https://doi.org/10.1186/s13104-021-05592-x
46. Sakamoto M, Hiasa Y, Otake Y, Takao M, Suzuki Y, Sugano N et al (2020) Bayesian segmentation of hip and thigh muscles in metal artifact-contaminated CT using convolutional neural network-enhanced normalized metal artifact reduction. J Signal Process Syst 92(3):335–344
47. Hahn A, Knaup M, Brehm M, Sauppe S, Kachelrieß M (2018) Two methods for reducing moving metal artifacts in cone-beam CT. Med Phys 45(8):3671–3680
48. Meyer E, Raupach R, Lell M, Schmidt B, Kachelrieß M (2012) Frequency split metal artifact reduction (FSMAR) in computed tomography. Med Phys 39(4):1904–1916. https://doi.org/10.1118/1.3691902
49. Hubbell J, Seltzer S (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;
50. Safari A, (2008) Piezoelectric and Acoustic Materials for Transducer Applications. Springer, New York https://link.springer.com/content/pdf/10.1007%2F978-0-387-76540-2.pdf
51. Harms J, Lei Y, Wang T, Zhang R, Zhou J, Tang X et al (2019) Paired cycle-GAN-based image correction for quantitative cone-beam computed tomography. Med Phys 46(9):3998–4009
52. He K, Zhang X, Ren S, Sun J. (2016) Deep residual learning for image recognition. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit. 2016: 770–778
53. Kingma DP, Ba JL, Adam (2015) A method for stochastic optimization. 3rd Int Conf Learn Represent ICLR 2015-Conf Track Proc. 1–15

54. Rohleder M, Gottschalk TM, Maier A, Kreher B (2022) Cross-domain metal segmentation for CBCT metal artifact reduction. 2022:4
55. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612
56. Bissonnette JP, Balter PA, Dong L, Langen KM, Lovelock DM, Miften M et al (2012) Quality assurance for image-guided radiation therapy utilizing CT-based technologies: a report of the AAPM TG-179. Med Phys 39(4):1946–1963
57. Yi X, Walia E, Babyn P (2019) Generative adversarial network in medical imaging: a review. Med Image Anal 58:101552
58. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. (2017) Densely connected convolutional networks. Proc-30th IEEE Conf Comput Vis Pattern Recognition, CVPR 2017. 2017: 2261–2269
59. Wang H, Lu X, Deng F. (2022) Improving CycleGAN for Image-to-Image Style Transfer by DenseNet. 2022 7th Int Conf Comput Commun Syst ICCCS 2022: 326–330

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.