Learning Metal Artifact Reduction in Cardiac CT Images with Moving Pacemakers

T. Lossau (née Elss)[a,b], H. Nickisch[a], T. Wissel[a], M. Morlock[b], M. Grass[a]

[a] *Philips Research, Hamburg, Germany*
[b] *Hamburg University of Technology, Germany*

Abstract

Metal objects in the human heart such as implanted pacemakers frequently lead to heavy artifacts in reconstructed CT image volumes. Due to cardiac motion, common metal artifact reduction methods which assume a static object during CT acquisition are not applicable.

We propose a fully automatic *Dynamic Pacemaker Artifact Reduction* (DyPAR+) pipeline which is built of three convolutional neural network (CNN) ensembles. In a first step, pacemaker metal shadows are segmented directly in the raw projection data by the SegmentationNets. Second, resulting metal shadow masks are passed to the InpaintingNets which replace metal-affected line integrals in the sinogram for subsequent reconstruction of a metal-free image volume. Third, the metal locations in a pre-selected motion state are predicted by the ModelingNets based on a stack of partial angle back-projections generated from the segmented metal shadow mask. We generate the data required for the supervised learning processes by introducing synthetic, moving pacemaker leads into 14 clinical cases without pacemakers.

The SegmentationNets and the ModelingNets achieve average Dice coefficients of $94.16\% \pm 2.01\%$ and $55.60\% \pm 4.79\%$ during testing on clinical data with synthetic metal leads. With a mean absolute reconstruction error of 11.54 HU $\pm$ 2.49 HU in the image domain, the InpaintingNets outperform the hand-crafted approaches PatchMatch and inverse distance weighting. Application of the proposed DyPAR+ pipeline to 9 clinical test cases with real pacemakers leads to significant reduction of metal artifacts and demonstrates the transferability to clinical practice. Especially the SegmentationNets and InpaintingNets generalize well to unseen acquisition modes and contrast protocols.

Keywords: Cardiac CT, Metal Artifact Reduction, Convolutional Neural Network

1. Introduction

Metal devices like implanted pacemakers lead to streak-shaped artifacts in reconstructed CT image volumes which degrade the image quality and diagnostic value. In particular, the evaluation of neighboring anatomy e.g. with regard to inflammations or calcifications is frequently precluded (Mak and Truong, 2012). Beam hardening, photon starvation, scattered radiation and the partial volume effect are identified as potential causes of such metal artifacts (De Man et al., 1999).

Beside acquisition improvement strategies like dual-energy protocols for material decomposition (Bamberg et al., 2011), several software-based solutions for CT metal artifact reduction (MAR) have been developed in the last decades (Mouton et al., 2013; Gjesteby et al., 2016). An exhaustive collection of relevant related papers is listed and compared in Table 1. The literature on MAR approaches is generally grouped into three major approaches (see first columns of Table 1): sinogram completion, iterative reconstruction and image-to-image transfer.

Sinogram completion The generation of consistent projection data by replacement of metal-affected line integrals commonly comprises the following steps:

1. Reconstruction of an initial image volume
2. Metal segmentation in the initial image volume e.g. by thresholding
3. Forward projection of the metal mask yielding the metal shadow in the originally acquired projection geometry
4. Replacement of metal-affected line integrals e.g. by interpolation or by incorporating prior knowledge
5. Reconstruction of a metal-free image volume using filtered back-projection (FBP)
6. Metal reinsertion using the metal-only image obtained in step 2.

A multitude of variants of this method is known which include different types of inpainting, pre- and post-processing procedures (Kalender et al., 1987; Meyer et al., 2010, 2012; Hahn et al., 2018).

Iterative Reconstruction An alternative to the replacement of metal-affected line integrals and subsequent FBP is the use of iterative reconstruction techniques for the generation of CT image volumes directly from incomplete or inconsistent projection data, e.g. by means of constrained optimization (Zhang et al., 2011). While the FBP assumes the projection data to be consistent and complete, i.e. uniformly sampled, iterative approaches seek

Table 1: Comparative review of CT metal artifact reduction methods in terms of Approach, Properties and Data in chronological order. Our proposed approach is a first pass method for *Dynamic Pacemaker Artifact Removal* (DyPAR+) which is robust to cardiac motion.

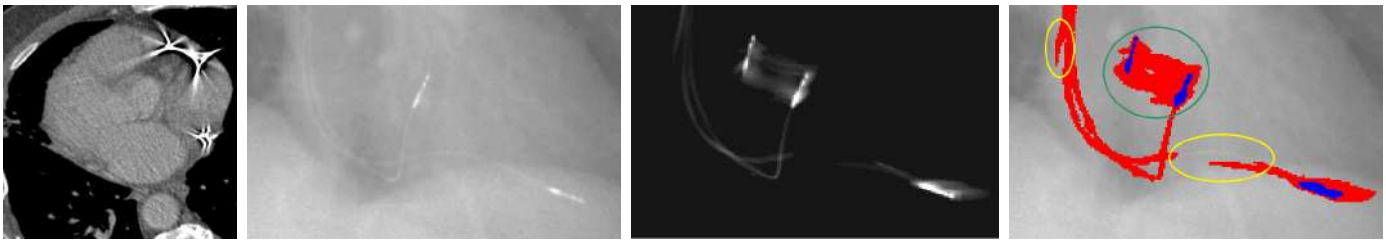
Paper	Approach			Properties				Data Source				Keywords
Symbols ✓: applicable ∼: not examined T : test data L : learning data	Sinogram Completion	Iterative Reconstruction	Image-To-Image Transfer	Motion-Robust	Multi-Pass	Hand-Crafted	Data-Driven	Hardware Phantom	Digital Phantom	Computer-Simulated Metal	Clinical Data	**Abbreviations** LI linear interpolation; EM expectation maximization ART algebraic reconstruction technique ML maximum likelihood NMAR normalized metal artifact reduction FBP filtered back-projection BHC beam hardening correction; DLB deep-learning-based CNN convolutional neural network MoCo motion compensation
Kalender 1987	✓				✓	✓		T			T	semi-automatic metal boundary determination; LI
Wang 1996		✓			✓	✓			T			iterative deblurring; incomplete projections; EM; ART
Oehler 2007	✓	✓			✓	✓					T	directional interpolation; weighted ML-EM
Meyer 2010	✓				✓	✓			T		T	NMAR; prior generation; projection normalization
Zhang 2011		✓			✓	✓		T	T			incomplete projections; constrained optimization
Slambrouck 2012		✓		∼	✓	✓		T	T			block-iterative scheme; local models of varying complexity
Meyer 2012	✓				✓	✓			T		T	initial NMAR; frequency split; spatial weighting
Toftegaard 2014	✓			✓		✓					T	marker segm. in projections; trajectory estimation
Gjesteby 2017a	✓		✓		✓	✓	✓		L			initial NMAR; image-based MAR
Gjesteby 2017b	✓				✓	✓	✓		L			initial NMAR; DLB sinogram correction
Zhang 2018	✓		✓		✓	✓	✓			L	T	initial FBP; initial BHC; initial LI; DLB prior generation
Xu 2018			✓	∼			✓	T	L			deep residual CNN; metal artifact prediction
Huang 2018			✓	∼			✓			L	T	deep residual CNN; metal artifact prediction
Park 2018	✓				✓		✓			L		image-based metal segm.; DLB sinogram correction
Hahn 2018	✓			✓	✓	✓			T		T	metal shadow refinement; LI; NMAR after MoCo
Lossau 2019	✓			✓			✓			L	T	DyPAR; DLB; metal shadow segmentation
This manuscript	✓			✓			✓			L	T	DyPAR+; DLB; metal shadow segm.; inpainting; modeling

the optimal fit to the measured data and therefore exhibit a higher robustness regarding inexact projection data (Wang et al., 1996). As these approaches are more time-consuming, van Slambrouck and Nuyts (2012) introduced an image block-iterative scheme where metal-regions are reconstructed with a fully polychromatic model whereas non-metal regions are reconstructed with a model of reduced complexity. Also hybrid approaches using both, sinogram completion and iterative reconstruction, have been investigated (Oehler and Buzug, 2007).

Image-to-Image Transfer Direct mapping from metal-affected to metal-free CT images e.g. by means of deep residual convolutional neural networks (CNNs) allows for artifact suppression without considering the corresponding raw projection data (Xu and Dang, 2018; Huang et al., 2018). However, these approaches are essentially restricted by the information content of the metal-affected input patches and therefore often combined with existing hand-crafted MAR algorithms (Gjesteby et al., 2017a; Zhang and Yu, 2018)

Most of the existing MAR methods are second pass approaches which fail in the presence of motion as they are based on metal segmentation in an initially reconstructed image volume. In the projection data, each recorded projection view corresponds to a specific motion state, i.e. metal positions are well-defined in each projection view but not consistent across multiple views. During reconstruction motion states are mixed, thus precluding the extraction of 3D metal models. As already stated by Toftegaard et al. (2014) and illustrated in Figure 1, MAR approaches based on metal segmentation in the image domain mainly suffer from two problems: (1) Due to the motion blur, hounsfield units (HU) of metal objects might get below the segmentation threshold resulting in incomplete metal removal. (2) Metal object sizes might be increased by the range of motion resulting in overestimated metal shadow areas. The first pass moving metal artifact reduction (MMAR) method of Toftegaard et al. (2014) avoids these problems by automatically segmenting cylindrical gold markers directly in the projection domain. Two methods for MMAR have been introduced by Hahn et al. (2018). The first approach utilizes image-based metal segmentation for coarse metal shadow determination. Inside the coarse mask, metal shadows are than refined by exploiting edge information in the projection data. The second approach assumes respiratory-gated CT data for the application of normalized MAR (NMAR) on a time series of previously motion compensated CT image volumes.

We aim for a pacemaker artifact reduction method which works completely rawdata-based and is applicable to gated as well as ungated CT scans. Furthermore, the method has to be robust to extremely low deviations of



(a) (b) (c) (d)

Figure 1: Metal artifact reduction (MAR) approaches which assume that the object is static during CT acquisition are not applicable for moving metal objects like pacemakers. (a) The initially reconstructed image volume exhibits severe metal and motion artifacts. (b) Each projection view contains the metal shadow at a specific motion state. (c) Image-based metal segmentation and subsequent forward projection yields metal shadows of mixed motion states. (d) Clear disagreements of real and predicted pacemaker shadow are visible. Blurred electrode shadows (green circle) as well as shifted and interrupted lead shadows (yellow circle) preclude conventional sinogram inpainting.

metal shadow and background line integrals as occurring in case of pacemaker leads. The literature in Table 1 is sorted chronologically. An increased incidence of data-driven and especially deep-learning-based approaches can be observed in the recent literature. Over the past few years, CNNs have been driving advances in many image-related tasks such as classification, inpainting, segmentation, generation and style transfer (Krizhevsky et al., 2012; Chen et al., 2018; Gatys et al., 2016). In previous work (Lossau et al., 2019), we developed a dynamic pacemaker artifact reduction (DyPAR) pipeline which is based on CNNs trained on clinical data with synthetic pacemaker leads for the task of metal shadow segmentation. As the method does not rely on initially reconstructed image volumes which are potentially motion-perturbed, superiority over standard sinogram completion was demonstrated. Here, DyPAR+ is introduced comprising the following adaptations and extensions compared to DyPAR:

1. The existing forward model for synthetic lead insertion is extended by a motion model which takes concomitant ECG data into account, i.e. dynamic leads instead of static ones are simulated (see Section 2.1)
2. The learning setup for the SegmentationNets is slightly adapted and post-processing of the output metal shadow masks by largest connected component extraction is omitted (see Section 3.3).
3. An additional network ensemble is trained for inpainting of metal-affected line integrals. The so-called InpaintingNets replace inverse distance weighting (IDW) in the DyPAR+ pipeline (see Section 3.3).
4. An additional network ensemble is trained for 3D lead modeling based on segmented metal shadow masks. The so-called ModelingNets are integrated as novel component into the DyPAR+ pipeline (see Section 3.4).

Generalization capabilities of DyPAR+ are investigated based on 9 clinical CT cases with real pacemakers (see Section 4.2). For comparison, the previous DyPAR pipeline and a second pass approach which comprises image-based metal segmentation and IDW are considered.

2. Data

The generation process of the synthetic learning data is detailed in Section 2.1. Simulated leads are inserted into the *target data without pacemakers*, whereby reasonable lead positions and pathways are extracted from the *reference data with pacemakers*. To evaluate the trained neural networks in the DyPAR+ pipeline, clinical test data with real pacemakers was acquired as described in Section 2.2.

2.1. Synthetic learning data

Target cases without pacemakers The raw projection data of 14 contrast-enhanced cardiac CT data sets without pacemakers is collected for synthetic lead insertion. In all target cases, acquisition was performed with a 256-slice CT scanner (Brilliance iCT, Philips Healthcare, Cleveland, Ohio , USA) using a retrospective gating protocol with helical trajectory. Details on the acquisition settings are summarized in Table 2.

Reference cases with pacemakers Seven reconstructed CT image volumes with pacemakers are collected for the extraction of pacemaker lead positions and pathways with respect to the cardiac anatomy. Dual as well as triple chamber pacemakers are included, i.e. synthesis of right atrial, right ventricular and coronary sinus leads is aimed for.

Dynamic forward model The data generation process is visualized in Figure 2. The forward model takes one reference case (i.e. a reconstructed image volume with pacemaker) and one target case (i.e. one ECG-gated sinogram without pacemaker P_{org}) as input and delivers synthetic data required for the supervised learning processes as output. First, a set of ten target image volumes is obtained by multi-phase reconstruction with a temporal distance of 10% cardiac cycle using aperture-weighted cardiac reconstruction (AWCR) (Koken and Grass, 2006). For the resulting target phase-volumes and the reference image volume, corresponding heart meshes are determined by model-based heart segmentation according to Ecabert et al. (2008). Along each pacemaker lead in the reference case, at least ten B-spline knots are manually selected. It has to be noted, that the definition of these landmarks

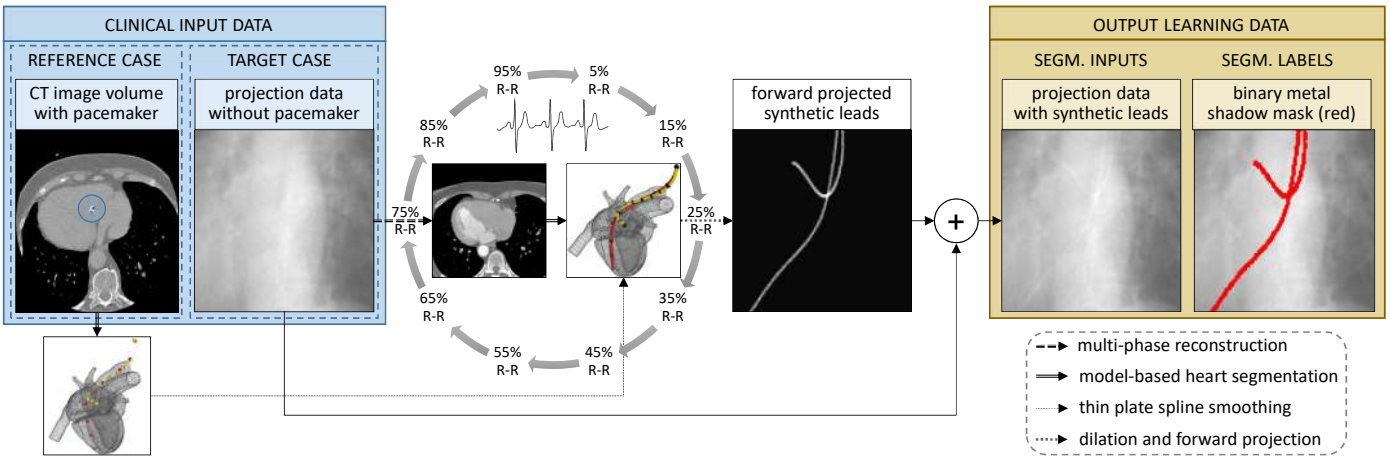


Figure 2: The data required for supervised learning is generated by a forward model which introduces synthetic pacemaker leads into the image and projection data of clinical cases without pacemakers. Surface meshes delineating the segmented heart during multi-phase reconstruction allow for sensible insertion positions and motion trajectories.

Table 2: Comparison of clinical test database and synthetic learning database with regard to pacemaker type, acquisition settings and scanner type (iCT: Brilliance iCT / B64: Brilliance 64, Philips Healthcare). The scanner type and the helical pitch determine the scan trajectory and thus also the reconstruction geometry. The rotation time [sec] and the number of recorded projection views per gantry turn define the temporal distance within the projection data.

	Clinical Test Data									Synthetic Learning Data			
	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 1	Case 2-4	Case 5-7	Case 8-14
Defibrillator	yes	no	no	no	no	yes	no	no	no	no	no	no	no
Gated	yes	yes	yes	yes	yes	yes	no	no	no	yes	yes	yes	yes
Dose Modulation	no	no	yes	yes	yes	no	no	no	no	no	no	no	no
Scanner Type	iCT	iCT	iCT	B64	iCT	iCT	iCT	iCT	iCT	iCT	iCT	iCT	iCT
Helical Pitch	0.160	0.160	0.165	0.200	0.180	0.180	0.664	0.664	0.993	0.160	0.160	0.180	0.180
Rotation Time / Turn	0.272	0.272	0.272	0.420	0.330	0.330	0.330	0.330	0.750	0.272	0.272	0.272	0.272
Number Views / Turn	1800	1800	2400	2320	2400	2400	2400	2400	2400	2400	1800	1800	2400

represents the only manual processing step within the dynamic forward model. Thin plate spline smoothing based on point-to-point correspondences in the segmented heart meshes allows the transformation of the B-spline knots from the reference case into each phase volume of the target case.

Synthetic metal shadows in the originally acquired projection geometry of the target case P_{metal} are gradually filled by the following procedure. For each projection view in the target sinogram, the corresponding cardiac phase point t is determined. Landmark positions in this specific motion state are calculated by linear interpolation of the B-spline knots associated with the two neighboring phase points within $\{5\%, 15\%, \ldots, 95\%\}$ cardiac cycle. This approach ensures continuous movements across various projection views. The corresponding B-spline curve is determined by cubic B-spline interpolation for each pacemaker lead, separately. Dilation of the resulting lines with a chosen lead diameter of 2 millimeters and an attenuation value of 4500 HU yields the binary image volume $I_{\text{metal}}(t)$ in the target image geometry. Subsequent forward projection delivers the metal shadow for the currently processed projection view. Clinical projection data with synthetic leads $P_{\text{input}} = P_{\text{org}} + P_{\text{metal}}$ is obtained by summation of the original projection data and the forward projected lead mask. Thresholding with zero defines the corresponding target segmentation mask $P_{\text{mask}} = \mathbb{1}_{P_{\text{metal}}>0}$.

The dynamic forward model is applied two times per reference case as twice as many target cases are available. With an iCT detector shape of 128×672 and an average number of 9 500 projection views per target case, a total amount of $14 \cdot 128 \cdot 672 \cdot 9\,500 \approx 1.15 \cdot 10^{10}$ labeled line integrals is collected. The database required for the segmentation, inpainting and modeling learning tasks comprises for each provided target case:

P_{input}	projection data with synthetic leads,
P_{mask}	binary mask of metal-affected line integrals,
P_{org}	original metal-free projection data,
$I_{\text{metal}}(t)$	time-dependent metal mask (image domain).

The learning data is case-wise separated into training, validation and testing subsets with a ratio of 8:4:2, or rather 4:2:1 with respect to the corresponding reference

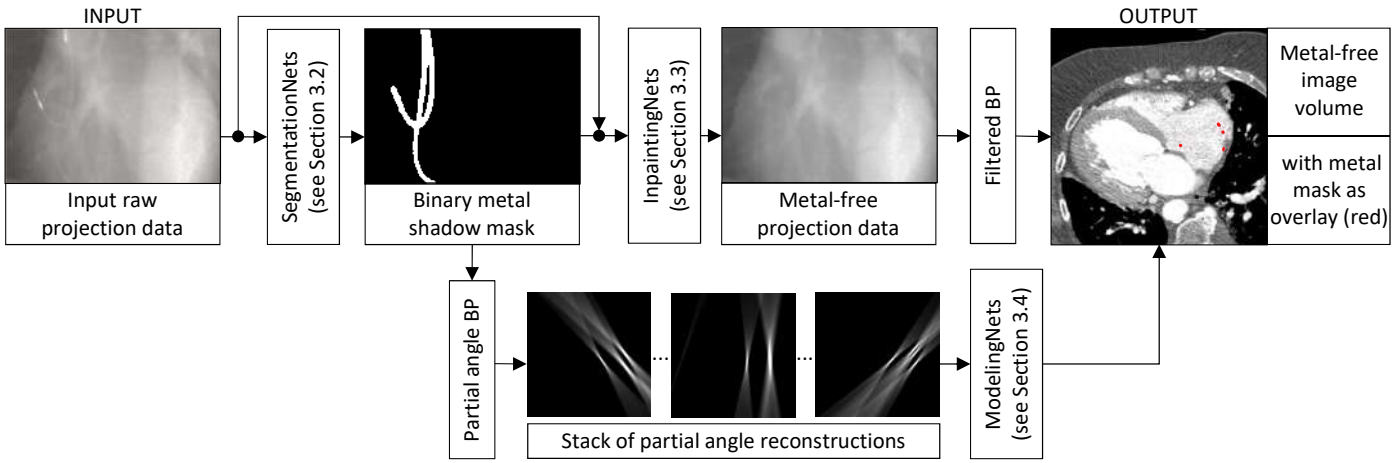


Figure 3: Neural networks trained for the tasks of metal shadow segmentation, sinogram inpainting and metal reinsertion represent the basic elements of the proposed dynamic pacemaker artifact reduction (DyPAR+) pipeline. The back-projection operator is abbreviated to BP.

cases in order to ensure disjoint pacemaker geometries and background line integrals among the subsets.

2.2. Clinical test data

In order to furthermore investigate generalization capabilities of the proposed DyPAR+ approach in clinical practice, the raw projection data of 9 additional cardiac CT data sets with real pacemakers is collected. The clinical test data allows one to evaluate the networks behavior in the presence of unseen features like electrodes or defibrillators and the robustness with regard to variations in contrast-enhancement, motion levels and acquisition settings. ECG-gated as well as ungated contrast-enhanced CT scans with helical acquisition trajectories are included (see Table 2). The ECG-gated test cases are reconstructed with a cardiac gating window around 75% R-R using AWCR, i.e. at mid-diastolic quiescent phase. For the ungated test cases which exhibit lower contrast agent densities, simple FBP is applied. Evaluation results on the clinical test data are presented in Section 4.2.

3. Method

The proposed DyPAR+ pipeline is build of three CNN ensembles and takes the raw projection data of a metal-affected CT scan as input. The SegmentationNets identify metal-affected line integrals directly in the projection domain, i.e. independent of motion. The InpaintingNets treat metal-affected values as missing data and refill the projection data based on surrounding line integrals. Subsequent reconstruction of the inpainted sinogram delivers the CT image volume without metal. The ModelingNets finally determine metal positions in the image domain based on the segmented metal shadow mask. The resulting metal mask can optionally be visualized as overlay. The DyPAR+ processing pipeline is illustrated in Figure 3.

3.1. Shared learning framework

This Section details components of the learning framework including network architecture and hyper-parameter settings which are shared across the different tasks of metal shadow segmentation, sinogram inpainting and metal modeling. Information on task-specific learning setups are provided in the following Sections 3.2, 3.3 and 3.4.

Network architecture During training, the neural networks take patches of size $c \times n \times n \times k$ as input and deliver patches of size $n \times n$ as output. The number of channels c, the number of slices k and the plane size $n \times n$ are task-specific parameters. Figure 4 illustrates the utilized U-Net architecture which is adapted from (Ronneberger et al., 2015). In case of multi-slice inputs ($k > 1$), feature extraction in the contracting path is performed for each slice separately using shared weights. Slice features are joint in the bottleneck and merely feature maps of the center slice are copied from the contracting to the expanding path in the skip connections. The network exhibits a receptive field size of $81 \times 81 \times k$. In contrast to (Lossau et al., 2019), partial convolution based padding according to (Liu et al., 2018b) is performed to keep in-plane input and output sizes equal. In general, arbitrary output shapes are enabled by the fully convolutional network, therefore, during validation and testing, metal shadow segmentation and inpainting is performed over the full detector size of 128×672 in a single step.

Bagging approach The stochastic gradient descent solver Adam (Kingma and Ba, 2015) with an initial learning rate of 0.01 and a momentum of 0.8 is used for network optimization. The learning rate decreases with a factor of two after 33% and 66% of the overall training time and L2 regularization with a weight of 0.0002 is used. One training epoch is defined by 10^5 processed samples. For each task, an ensemble of seven CNNs is trained by the following bagging approach:

1. Select test data from a single reference case.

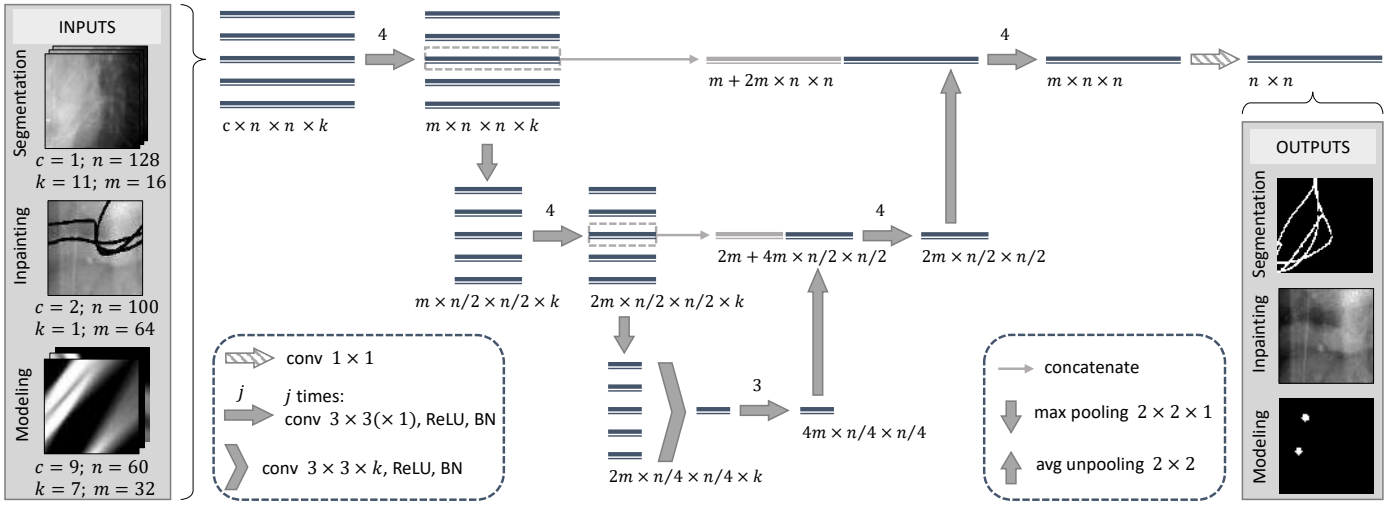


Figure 4: The fundamental network architecture is shared across the tasks of metal shadow segmentation, sinogram inpainting and metal modeling. The adapted U-Net design enables multi-channel and multi-slice input patches. Task-specific parameters include the number of input channels c, the number of input slices k, the plane size $n \times n$ and the number of feature maps at each layer determined by m. Each double line corresponds to a multi-channel feature map. A slice number of $k = 5$ is illustrated for clarity. The actual shape of each feature map is denoted at the lower edge of the line block. The arrows represent the different operations.

2. Randomly sample validation data belonging to two reference cases from the remaining cases.
3. Train networks training on the remaining data corresponding to four reference cases.
4. After every epoch, assess the network's generalization capabilities by a validation metric on the validation data.
5. Select the model with the highest validation metric within all training epochs for performance evaluation and application in DyPAR+.
6. Repeat steps 1.-5. seven times; every synthetic learning case is selected one time for testing.

The chosen case-wise subsets for training, validation and testing remain unchanged across the tasks.

3.2. SegmentationNets

SegmentationNets are trained to map projection data with metal leads P_{input} to corresponding binary masks P_{mask} of metal-affected line integrals. On average, 1.37% of the labeled projection data contain metal shadow voxels. To compensate for this foreground-background class imbalance, a patch-based learning approach is applied.

Patch sampling During training, the SegmentationNets take patches of size $1 \times 128 \times 128 \times 11$ as input and deliver patches of size 128×128 as output. The first dimension of the networks input corresponds to the number of channels (here $c = 1$). The second and the third dimension contain the information of the detector row and column. The forth dimension indicates projection views which are equidistantly sampled with respect to the number of views per gantry turn so that 12 degrees gantry rotation are captured. The SegmentationNets target the segmentation mask of the middle sixth view. By including neighboring projection views, the networks get additional information on the rotation velocity, i.e. the distance of supposed pacemaker leads to the rotation center. The sampling process is controlled such that 75% of the target output patches contain at least one object voxel while the remaining 25% are randomly sampled.

Learning framework In the contracting path of the SegmentationNets, 2D lead features are extracted for each view, separately and joint in the bottleneck to exploit the temporal information. In the expanding path, merely location information of the center slice to be segmented are copied from the contracting path. In the last network layer, the soft-max function is used as activation. The SegmentationNets architecture with its shared weights in the contracting path has 423 730 learned parameters. Training is performed over 30 epochs using a mini-batch size of 32. The learning process is driven by the focal loss (Lin et al., 2018) which is well suited for imbalanced segmentation tasks using a focusing parameter of $\gamma = 2$. The Dice coefficient with a probability threshold of $\tau = 0.5$ is used as validation metric in the bagging approach.

Application in DyPAR+ In order to increase the robustness of the metal shadow segmentation, the entire ensemble of seven SegmentationNets yielded by the bagging approach is applied on the input raw projection data. The output probability maps are averaged across the ensemble and contain values in $[0, 1]$. The binary metal shadow mask (1: metal-affected, 0: background) is obtained by thresholding with $\tau = 0.15$. The choice of the relatively low threshold is motivated by the fact that incompletely segmented metal shadows may lead to newly introduced artifacts after inpainting and reconstruction. Therefore, we judge sensitivity as more important performance measure than precision. The threshold of $\tau = 0.15$ corresponds to a maximal false negative rate of 1% during testing on

synthetic learning data.

During DyPAR+, InpaintingNets and ModelingNets rely on the outputs of the SegmentationNets, i.e. they have to deal with false positive and false negative metal shadow segmentations. For each data split in the bagging approach, the model selected in step 5. is applied on the test cases. Binary segmentation masks resulting by thresholding with $\tau = 0.15$ are stored in P_{segm}. As described in the following Sections, InpaintingNets and ModelingNets are trained on P_{segm}, i.e. the testing output masks with slight inaccuracies, rather than on the ground truth masks P_{mask}.

3.3. InpaintingNets

InpaintingNets are trained to map metal-afflicted projection data P_{input} masked by $M = \neg P_{\text{segm}}$ to corresponding metal-free line integrals from the original projection data P_{org}. Line integrals with a mask value of zero are treated as missing data.

Patch sampling During training, the network takes patches of size $2 \times 100 \times 100 \times 1$ as input, whereby channel information are sampled from the projection data P_{input} and the corresponding mask M. Online data augmentation is performed by randomly treating up to ten additional line integrals in every tenth input patch as unknown. The sampling process is controlled in such a way that each patch contains at least one missing line integral.

Learning framework Several deep-learning-based approaches dealing with free-form masks and non-blind image inpainting (i.e. regions to be inpainted are known a priori) have been presented in the last years (Nazeri et al., 2019; **?**). We adapted the network architecture illustrated in Figure 4 for the inpainting task by replacing all convolutional layers with partial convolutions as suggested by Liu et al. (2018a). In a partial convolution only valid pixels are taken into account and the layers output is re-normalized according to the ratio of kernel size and masking area. The mask is updated after every layer and passed as additional single-channel feature map through the network. Within the skip connections, merely common feature maps are copied and concatenated, i.e. upsampled masks from deeper layers are utilized in the expanding path. After the last convolution layer with linear activation function, inpainted areas of the networks output $\text{NN}_{\text{inpaint}}$ are combined with the original input by $P_{\text{clean}} = M \odot P_{\text{input}} + (1-M) \odot \text{NN}_{\text{inpaint}}$. The network has an increased number of feature maps compared to the SegmentationNets and 3 805 313 learned parameters in total. Training is performed over 60 epochs using a mini-batch size of 64. The learning is driven by the combined loss function:

$$\begin{aligned} L_{\text{inpaint}} = \frac{1}{2N_P}(&\|P_{\text{clean}} - P_{\text{org}}\|_1 \\ &+ \|S_x * P_{\text{clean}} - S_x * P_{\text{org}}\|_1 \\ &+ \|S_y * P_{\text{clean}} - S_y * P_{\text{org}}\|_1) \end{aligned}$$

N_P denotes the number of elements in P_{org} and $S_{x/y}$ represents Sobel convolution kernels for vertical and horizontal derivative approximation. The loss function therefore penalizes differences in edge information. The validation metric utilized in step 4) of the bagging approach is replaced by a per-pixel reconstruction accuracy

$$\text{nMAE} = \frac{1}{\|P_{\text{segm}}\|_1} \|P_{\text{clean}} - P_{\text{org}}\|_1 \tag{1}$$

normalized by the number of line integrals.

Application in DyPAR+ InpaintingNets take the raw projection data P_{input} and the corresponding metal shadow mask P_{segm} predicted by the SegmentationNets as input. View-wise processing and subsequent ensemble averaging yields the inpainted projection data P_{clean}. The metal-free image volume is obtained by reconstruction of P_{clean} using either FBP for ungated cases or AWCR when concomitant ECG data is available.

3.4. ModelingNets

ModelingNets are trained to predict metal positions in the reconstructed image volume based on segmented binary metal shadow masks. Since the SegmentationNets might produce slight inaccuracies, the networks have to be robust regarding false positives and false negatives. Furthermore, cardiac motion needs to be compensated by the network to produce metal masks without blurring artifacts. Our metal reinsertion method is inspired by existing motion compensation approaches which exploit the increased temporal resolution of partial angle reconstructions (PARs) (Kim et al., 2015; Grass et al., 2016; Hahn et al., 2017).

Partial Angle Reconstruction Each projection view P_j is associated with a specific gantry rotation angle γ_j and acquisition time point $t_j \in [0\ \%\ \text{R-R}, 100\ \%\ \text{R-R})$ within the cardiac cycle. Given a center projection view P_c, we use the back-projection operator $\mathcal{B}$ without high-pass filtering to reconstruct 9 partial angle volumes of disjoint 20° angle segments

$$\begin{aligned} A_k = \sum_{j \in \Gamma_k} \mathcal{B}(P_{\text{segm},j}), \quad &\text{for } k \in \{-4, -3, \ldots, 4\} \\ \text{whereby } \Gamma_k = \{j : |\gamma_j - 20°k - \gamma_c| \ &< 10°\}. \end{aligned} \tag{2}$$

By this procedure 180° gantry rotation are covered in total as illustrated in Figure 5. The target metal mask $I_{\text{metal}}(t_c)$ which corresponds to the motion state at acquisition time point t_c is highlighted in red. Depending on the center index c, a partial field of view (pFOV) is defined as

$$\Omega_c = \begin{cases} 1, & \text{if voxel is part of the FOV over} \\ & \text{the full angle range of } 180°. \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

For each target case, four stacks of PARs $S_c = \{A_{-4}, A_{-3}, \ldots, A_4\}$ with varying center index c and corresponding target metal masks $I_{\text{metal}}(t_c)$ are created. Solely

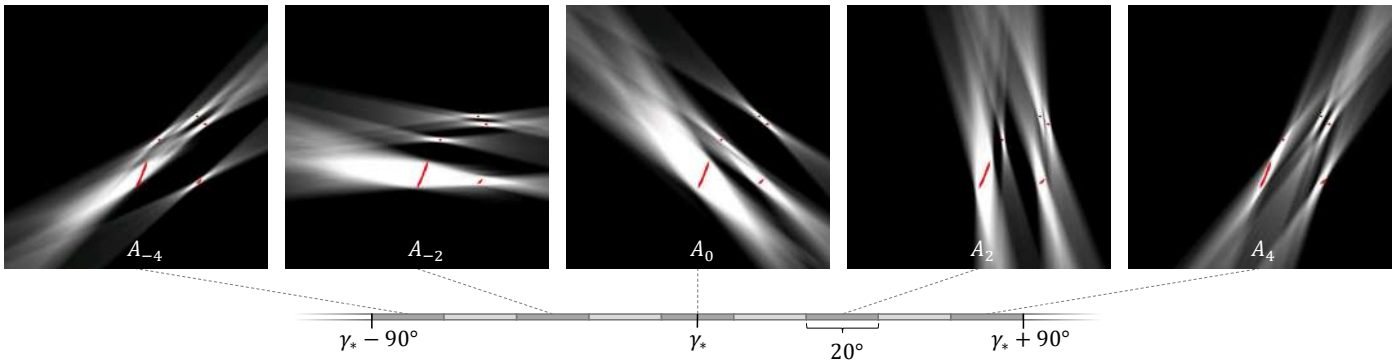


Figure 5: The ModelingNets take a stack of partial angle reconstructions $A_{-4}, A_{-3}, \ldots, A_4$ obtained from the segmented metal shadow mask as input and deliver the corresponding metal mask in the image domain $I_{\text{metal}}(t_c)$ as output (highlighted in red). Projection beams in A_0 are closest in time to the target motion state determined by t_c. With increasing temporal distance, slight shifts between back-projected beams and target metal mask can be observed.

image regions which are part of the pFOV are included in the learning data. By this procedure, a total amount of $14 \cdot 512 \cdot 512 \cdot 4 \cdot 170 \approx 2.5 \cdot 10^9$ labeled voxels is collected whereby on average 0.024% of the image data contains object voxels. In order to compensate this class imbalance, a patch-based approach is applied again.

Patch sampling During training, the ModelingNets take patches of size $9 \times 60 \times 60 \times 7$ as input and deliver patches of size 60×60 as output. The PARs $A_{-4}, \ldots, A_4$ which belong to different angular segments are provided as channel information. The second and third dimension of the networks input contain the information of the axial plane. Seven neighboring axial slices are included, whereby the ModelingNets target the segmentation mask $I_{\text{metal}}(t_c)$ of the middle forth view. By including neighboring axial views, the networks get additional information on the lead pathways. Online data augmentation is performed by random axial rotation of the input and target patches by 0°, 90°, 180° or 270°. Furthermore, mirroring along the scanners x- and z-axis increases the training data base. To enforce a clockwise system rotation direction, the channel order is inverted in case of mirroring along the x-axis. The sampling process is controlled in such a way that 90% of the target output patches contain at least one object voxel while the remaining 10% are randomly sampled.

Learning framework The number of feature maps is doubled compared to the SegmentationNets, resulting in a total number of 1 398 114 learned parameters. The learning setup including loss function, validation metric and hyper-parameter settings remains unchanged compared to Section 3.2. Case sampling is performed with regard to the reference cases, i.e. 8 testing, 16 validation and 32 training volumes are selected for each split in the bagging approach.

Application in DyPAR+ Gated CT scans are reconstructed by AWCR whereas simple FBP is used for ungated test cases. Therefore, application of the ModelingNets is adapted depending on the availability of ECG data.

ECG-gated test case: A specific heart phase t_j is associated with each projection view P_j by means of the ECG data. During AWCR a reference heart phase r needs to be specified which determines the center of the cardiac gating window. For each recorded cardiac cycle $\mathcal{H}$ a corresponding subset of views is assigned by $\Lambda(\mathcal{H}) = \{j | P_j \text{ is aquired within } \mathcal{H}\}$. The projection view $P_{c(\mathcal{H})}$ with heart phase closest to the reference phase is identified by $c(\mathcal{H}) = \underset{j \in \Lambda(\mathcal{H})}{\operatorname{argmin}} |t_j - r|$. The stack of corresponding partial angle volumes $S_{c(\mathcal{H})}$ is generated according to equation (2) and fed into the ModelingNets. The networks output NN_{model} includes averaging across the ensemble. Under consideration of the pFOV $\Omega_{c(\mathcal{H})}$, the output probability map is calculated by

$$I_{\text{metal}} = \frac{\sum_{\mathcal{H}} \text{NN}_{\text{model}}(S_{c(\mathcal{H})}) \odot \Omega_{c(\mathcal{H})}}{\sum_{\mathcal{H}} \Omega_{c(\mathcal{H})}}. \tag{4}$$

Thresholding with 0.5 finally delivers the binary metal mask in image domain.

Ungated test case: In case of ungated CT data, projection views P_j can not be associated with a specific time point t_j. But, for each axial image slice I_z, a corresponding nearest projection view P_j can be calculated by

$$j(z) = \frac{\text{NVPT}}{\text{Pitch}} \left(\text{res}_z \cdot z + T_z - \text{Source}_z\right), \tag{5}$$

whereby NVPT denotes the number of views per gantry turn and res_z is the image resolution in z. The expression $T_z - \text{Source}_z$ specifies the distance in z of the first axial slice in the image FOV to the center of rotation for the first projection view. In order to avoid blending of different motion phases in the metal image, the metal probability map is block-wise filled, i.e. $I_{\text{metal},Z} = \text{NN}_{\text{model}}(S_{c(Z)})$, whereby Z denotes a subset of axial slice indices. The center index $c(Z) = j(z^*)$ is calculated according to equation (5) whereby z^* is defined as the center slice of Z. The axial block size $|Z|$ is an adjustable parameter. A small block size leads to smoother output probability maps, but requires a longer runtime. It has to be noted that the metal reinsertion process for ungated cases can be significantly

Table 3: Test results on the synthetic learning data including mean and standard deviation of selected performance metrics for segmentation and inpainting tasks. The threshold τ defines the metal-background class separation, whereby $\tau = 0$ would correspond to classifying all pixels/voxels as metal-affected. Except for the mean absolute error (MAE) in the projection and in the image domain, all scores are expressed in percent.

Approach	**Metric**	**Score**
SegmentationNets (threshold $\tau = 0.5$)	Dice coefficient	94.16±1.49
	Sensitivity	93.88±1.37
	Precision	94.57±2.34
SegmentationNets (threshold $\tau = 0.15$)	Dice coefficient	88.05±2.17
	Dice coefficient (D1)	97.88±2.42
	Sensitivity	99.02±0.79
	Sensitivity (D1)	99.53±0.53
	Precision	79.44±2.94
	Precision (D1)	96.51±4.03
InpaintingNets	nMAE (projection)	6.040±0.88
	MAE (image) [HU]	11.54±2.49
Inverse Distance Weighting	nMAE (projection)	6.337±0.87
	MAE (image) [HU]	12.17±2.57
PatchMatch	nMAE (projection)	6.912±1.16
	MAE (image) [HU]	12.72±2.71
ModelingNets (threshold $\tau = 0.5$)	Dice coefficient	55.60±4.79
	Dice coefficient (D1)	76.02±6.98
	Sensitivity	53.35±5.07
	Sensitivity (D1)	73.24±7.34
	Precision	59.15±5.03
	Precision (D1)	80.96±6.70

*D1-corrected: false positives and false negatives within the dilated true positive area (using a $3 \times 3 \times 3$ structure element) are ignored

accelerated by reusing partial image volume under consideration of the table movement during 20° gantry rotation. Furthermore, PARs should only be back-projected to the relevant image area defined by Z. Thresholding of $I_{\text{metal},Z}$ with 0.5 finally delivers the binary metal mask in image domain.

4. Experiments and Results

For all experiments, the Microsoft Cognitive Toolkit (CNTK v2.5+, Microsoft Research, Redmond, WA, USA) is used as deep learning framework. Section 4.1 deals with quantitative and qualitative evaluation of the network's performance on the synthetic learning data. The network's generalization capabilities to clinical test data with real pacemakers are examined in 4.2.

4.1. Evaluation on synthetic learning data

Performance measurements of the network ensembles achieved on the testing subsets are summarized in Table 3. As most segmentation errors produced by the SegmentationNets and the ModelingNets occur at the boundaries of the metal mask, also D1-corrected performance measures are considered. Despite thin line-shaped object masks, remarkably high dice coefficients are achieved by SegmentationNets. The ModelingNets have to deal with a more extreme class imbalance and segmentation errors of the SegmentationNets. Furthermore, predicting the exact position and diameter of the pacemaker leads, based on the PARs is indeed a difficult task. The image volumes in the learning data are reconstructed with a voxel resolution between 0.4 mm and 0.5 mm. Therefore, lead pathways shifted by few voxels are tolerable.

The InpaintingNets are compared with the hand-crafted approaches PatchMatch (Barnes et al., 2009) and inverse distance weighting (Shepard, 1968). For PatchMatch we used a third-party implementation[1]. The IDW is performed in the following experiments by interpolating metal-affected line integrals based on border pixels (defined in a 8-neighborhood around the segmentation mask) using the L_∞ metric as distance function. All inpainting approaches are tested by view-wise processing using the projection data P_{input} and the ground-truth metal shadow masks P_{mask} as input. Beside the normalized mean absolute error in the projection domain (nMAE) introduced in Equation (1), the mean absolute deviation from I_{org} is regarded, as the image quality after reconstruction is most crucial. In both domains, the deep-learning-based approach outperforms the hand-crafted ones. In Figure 6, example reconstruction results after inpainting are compared. The visual impression coincides with the performance scores. The InpaintingNets induced least streak-shaped artifacts and seem to fill the projection data with higher consistency across the projection views.

4.2. Evaluation on clinical test data

DyPAR+ is applied to 9 clinical test cases with real pacemakers described in Section 2.2. Figure 7 provides qualitative evaluation results of the networks outputs in the projection domain. As already stated in (Lossau et al., 2019), the SegmentationNets object-background separation also generalizes to electrodes and defibrillators despite the lag of dedicated learning data. False negatives occur especially at the pacemaker leads due to a low deviation of metal shadow and background line integrals (see Figure 7d,f). ECG-leads and pacemaker leads are visually hard to distinguish based on a single projection view. Apart from few exceptions (see Figure 7h), the SegmentationNets are remarkably successful in their separation and seem also to consider rotation velocities (see Figure 7b,c,e). The InpaintingNets are able to fill metal-affected line integrals. However, inpainted areas exhibit removed noise patterns and reduced edge preservation, e.g. in the case of interrupted ECG-leads (see Figure 7b,e). This is a known effect of many data-driven inpainting approaches that use the MAE as loss function.

Maximum intensity projections visualized in Figure 8 compare between real and predicted lead pathways. The

[1] https://github.com/younesse-cv/patchmatch

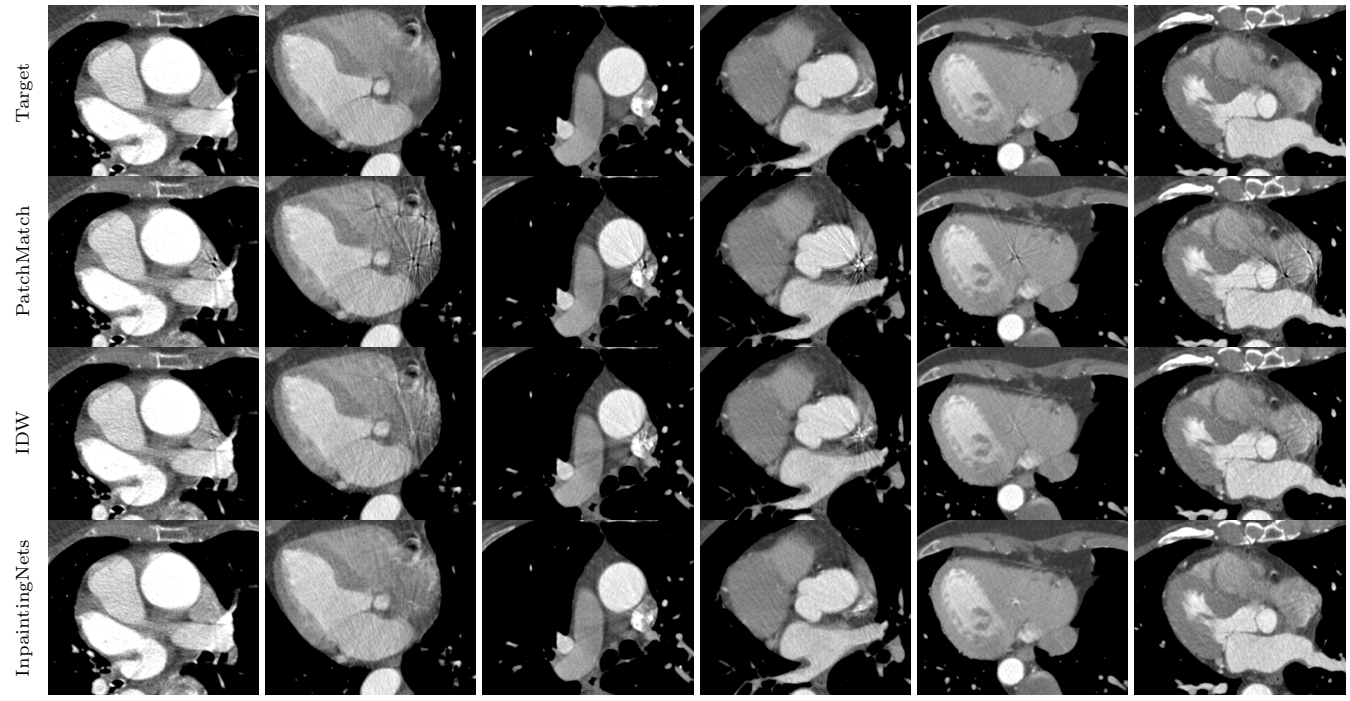


Figure 6: Comparison of reconstructed axial image slices after inpainting of synthetic metal shadows using PatchMatch, inverse distance weighting (IDW) and the InpaintingNets with the original target image volume without synthetic leads.

output of the ModelingNets highly depends on the quality of the previous metal shadow segmentation step. In all test cases, hardly false positive activations are present. In the gated test cases, edges at the boundaries of the partial FOVs $\Omega_{c(\mathcal{H})}$ are visible in the probability maps (see Figure 8a-c). However, extracted binary metal masks after thresholding are coherent and inserted at the correct positions. The redundancy in the projection data due to the low pitches and averaging across different cardiac cycles increases the robustness of metal modeling for gated cases. In contrast, the predicted metal probability map of the ungated case in Figure 8d exhibits stack transition artifacts and increased blurring. One fifth of the axial slices is selected as block size $|Z|$. Interrupted leads occur in the extracted binary metal mask after thresholding. In this case it might be more sensible to use the probability map without binarization as heatmap overlay. It has to be mentioned that especially in the ungated test cases, contrast enhancement and acquisition settings like pitch and gantry rotation speed vary from the learning data (see Table 2). Nevertheless, a significant metal artifact reduction can be observed in the DyPAR+ output.

We compare the proposed pacemaker artifact reduction method with a common second pass approach and the previous DyPAR pipeline. In the second pass approach, metal masks are segmented in the image domain using 3D hysteresis thresholds of 1000 HU and 1500 HU. The metal shadow areas are yielded by forward projection and thresholding with zero. IDW is applied as inpainting strategy. The previous DyPAR pipeline proposed in (Lossau et al., 2019) comprises deep-learning-based metal shadow segmentation and IDW. For each clinical test case, two axial slices are depicted in Figure 9 which exhibit severe metal artifacts after conventional reconstruction without MAR. Image slices after MAR are ranked by visual comparison of blurring and streak-shaped artifact levels, whereby a ranking of 1 corresponds to the highest image quality. In case of similar artifact levels, MAR approaches yield the same score.

Mean observer rankings of $2.\overline{5}$ and $1.9\overline{4}$ are achieved by the second pass and the DyPAR approach. The second pass approach is not robust regarding cardiac motion and leads to increased blurring in the neighboring anatomy, incomplete metal removal and introduction of new severe artifacts in several slices (see Figure 9b,c,g,h,i). For visual inspection of motion perturbations in metal shadow masks resulting from image-based segmentation we refer to Figure 1 and (Lossau et al., 2019). During DyPAR, metal shadow masks are post-processed by largest connected component extraction in order to reduce the number of false positives. This post-processing step is not applied in DyPAR+, as it might lead to incorrect removal of true positives in case of gaps in the metal shadow masks (see Figure 9f). Furthermore, no metal reinsertion is performed in contrast to second pass MAR and DyPAR+.

With a mean observer ranking of 1.0, axial image slices after DyPAR+ exhibit least artifacts, i.e. it benefits from the deep-learning-based metal shadow segmentation and inpainting. However, partial angle artifacts due to incomplete metal shadow segmentations (see Figure 9d,g) and introduction of streak artifacts due to inconsistencies among the 2D projections occurred after the inpainting

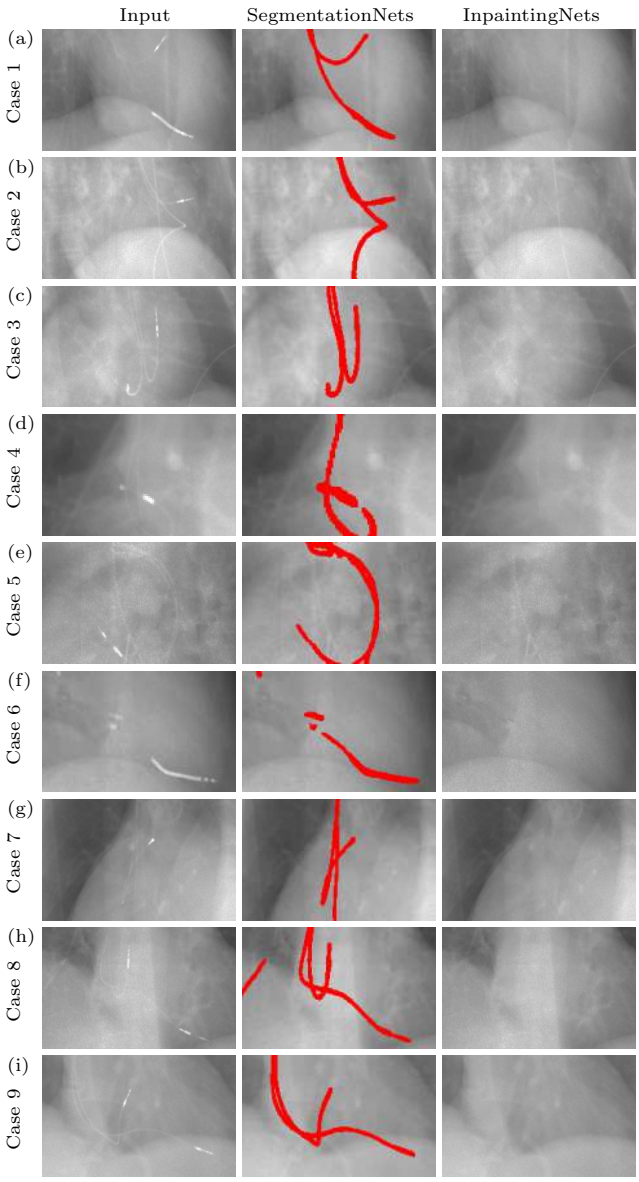


Figure 7: For each clinical test case, one example view of the input projection data P_{input} with corresponding outputs of the SegmentationNets P_{segm} and InpaintingNets P_{clean} is depicted.

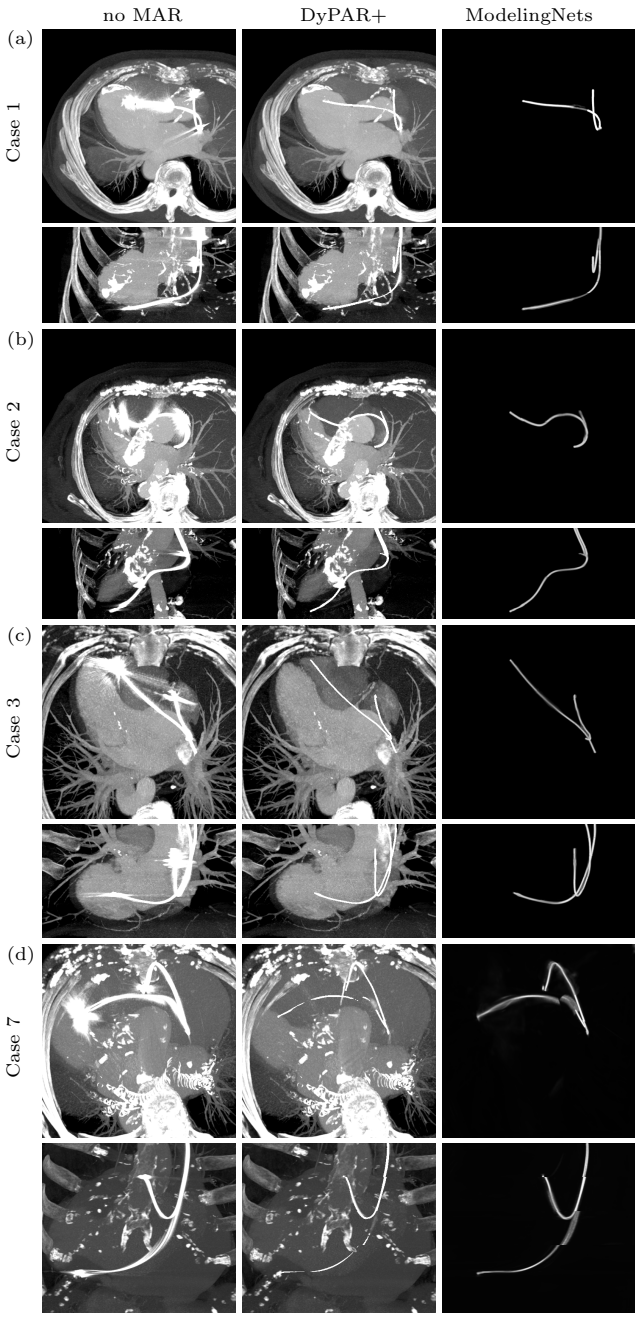


Figure 8: Maximum intensity projections of axial (top) and sagittal (bottom) slices are compared for image volumes reconstructed without MAR, image volumes after DyPAR+ and the output probability maps of the ModelingNets I_{metal}.

step (see Figure 9a) are still the main sources of image quality degradation. Nevertheless, metal artifacts are successfully reduced by DyPAR+ and the evaluation of neighboring anatomy is facilitated in most cases. In Figure 9h number and position of pacemaker leads can be identified without cardiac motion blur and in Figure 9a-c, metal artifacts are removed which hamper evaluation of portions of the coronary arteries. Therefore, DyPAR+ might be used to improve procedure planning from cardiac CT data for minimal invasive pacemaker lead extraction and examination of other pathologies. It shows a high robustness to different noise levels, contrast agent densities and motion velocities. The experiments demonstrate the feasibility of pacemaker artifact reduction without the need of an initial image reconstruction and the transferability from synthetic leads to real pacemakers.

5. Discussion

The proposed approach for dynamic MAR offers a lot of potential for further research and exhibits several tunable parameters. By reducing the angular range of the PARs in Section 3.4 to less than 20°, the temporal resolution could be further increased. During parameter tuning we observed that especially the number of input slices k in

Figure 9: A selection of axial image slices without MAR, after 2nd pass MAR, after DyPAR and after DyPAR+ are visualized using a level/window setting of 150/750 HU. Reinserted metal is highlighted in red. Observer rankings of the MAR approaches are provided in the upper left corner of each slice whereby a score of 1 corresponds to least artifacts.

the SegmentationNets and the ModelingNets significantly influences the network performances. Except for simple thresholding, no post-processing of the network outputs is performed so far. Application of morphological operations and frequency splitting (Meyer et al., 2012) might further increase the resulting image quality.

Establishment and maintenance of consistent projection data is most crucial for artifact removal and avoidance, but the current view-wise processing precludes the consideration of long time dependencies. Approaches for spatial propagation, e.g. by means of recurrent neural networks (Chen et al., 2016) or spatio-temporal adversarial objectives (Wang et al., 2018), could provide solutions here. The extension of the dynamic forward model for introduction of synthetic electrodes and defibrillators with additional simulation of beam hardening and Poisson noise as performed by Zhang and Yu (2018) might enable the combination of the segmentation and the inpainting step by directly learning to predict residual metal shadows. By this, also information behind the metal shadow are exploitable.

The methodology of first transforming reference device models by means of model-based heart segmentation and thin plate spline smoothing into clinical target cases for subsequent device-specific network training is, in principle, not restricted to pacemakers. Transferability to projection-based detection and removal of other high-intensity objects like artificial valves, electrodes, or left ventricular assist devices is part of future research. Beside the device-specific learning, also protocol-specific MAR is feasible. Cardiac CT images are acquired with a wide variety of imaging protocols. By providing a set of sinograms without metal implants, DyPAR+ is on-site trainable on data of arbitrary scanner type, acquisition mode and contrast protocol.

6. Conclusion

We propose a fully automatic, first-pass metal artifact reduction method for cardiac CT data with implanted pacemakers. The method is a substantial extension to a previous algorithm in several respects: dynamic leads instead of static ones, improved learning setup, inpainting for improved line integral correction, and 3D lead modeling. We successfully demonstrate the generalization capabilities of convolutional neural networks (CNNs) trained on clinical data with synthetic pacemaker leads to clinical cases with real metal-perturbations. In order to bring the very promising results with our DyPAR+ pipeline into clinical practice, a qualitative validation studies is required.

Acknowledgments

We thank Samer Hakmi (University Heart Center) and Clemens Spink (Department of Diagnostic and Interventional Radiology and Nuclear Medicine) both from University Medical Center Hamburg-Eppendorf (UKE), Germany, for the cardiac CT data with implanted pacemakers.

References

Bamberg, F., Dierks, A., Nikolaou, K., Reiser, M.F., Becker, C.R., Johnson, T.R., 2011. Metal artifact reduction by dual energy computed tomography using monoenergetic extrapolation. European Radiology 21, 1424–1429.

Barnes, C., Shechtman, E., Finkelstein, A., Goldman, D.B., 2009. Patchmatch: A randomized correspondence algorithm for structural image editing, in: ACM Transactions on Graphics (ToG), ACM. p. 24.

Chen, J., Yang, L., Zhang, Y., Alber, M., Chen, D.Z., 2016. Combining fully convolutional and recurrent neural networks for 3D biomedical image segmentation, in: Advances in Neural Information Processing Systems, pp. 3036–3044.

Chen, L., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2018. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence 40, 834–848.

De Man, B., Nuyts, J., Dupont, P., Marchal, G., Suetens, P., 1999. Metal streak artifacts in X-ray computed tomography: a simulation study. IEEE Transactions on Nuclear Science 46, 691–696.

Ecabert, O., Peters, J., Schramm, H., Lorenz, C., von Berg, J., Walker, M.J., Vembar, M., Olszewski, M.E., Subramanyan, K., Lavi, G., et al., 2008. Automatic model-based segmentation of the heart in CT images. IEEE Transactions on Medical Imaging 27, 1189–1201.

Gatys, L.A., Ecker, A.S., Bethge, M., 2016. Image style transfer using convolutional neural networks, in: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (CVPR).

Gjesteby, L., De Man, B., Jin, Y., Paganetti, H., Verburg, J., Giantsoudi, D., Wang, G., 2016. Metal artifact reduction in CT: where are we after four decades? IEEE Access 4, 5826–5849.

Gjesteby, L., Yang, Q., Xi, Y., Shan, H., Claus, B., Jin, Y., De Man, B., Wang, G., 2017a. Deep learning methods for CT image-domain metal artifact reduction, in: Developments in X-ray Tomography XI, International Society for Optics and Photonics. p. 103910W.

Gjesteby, L., Yang, Q., Xi, Y., Zhou, Y., Zhang, J., Wang, G., 2017b. Deep learning methods to guide CT image reconstruction and reduce metal artifacts, in: Medical Imaging 2017: Physics of Medical Imaging, International Society for Optics and Photonics. p. 101322W.

Grass, M., Thran, A., Bippus, R., Kabus, S., Wiemker, R., Vembar, M., Schmitt, H., 2016. Fully automatic cardiac motion compensation using vessel enhancement, in: Abstracts of the 11th Annual Scientific Meeting of the Society of Cardiovascular Computed Tomography, JCCT.

Hahn, A., Knaup, M., Brehm, M., Sauppe, S., Kachelrieß, M., 2018. Two methods for reducing moving metal artifacts in cone-beam CT. Medical Physics 45, 3671–3680.

Hahn, J., Bruder, H., Rohkohl, C., Allmendinger, T., Stierstorfer, K., Flohr, T., Kachelrieß, M., 2017. Motion compensation in the region of the coronary arteries based on partial angle reconstructions from short-scan CT data. Medical Physics 44, 5795–5813.

Huang, X., Wang, J., Tang, F., Zhong, T., Zhang, Y., 2018. Metal artifact reduction on cervical CT images by deep residual learning. Biomedical Engineering Online 17, 175.

Kalender, W.A., Hebel, R., Ebersberger, J., 1987. Reduction of CT artifacts caused by metallic implants. Radiology 164, 576–577.

Kim, S., Chang, Y., Ra, J.B., 2015. Cardiac motion correction based on partial angle reconstructed images in X-ray CT. Medical Physics 42, 2560–2571.

Kingma, D., Ba, J., 2015. Adam: A method for stochastic optimization.

Koken, P., Grass, M., 2006. Aperture weighted cardiac reconstruction for cone-beam CT. Physics in Medicine and Biology 51, 3433.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems (NIPS), pp. 1097–1105.

Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P., 2018. Focal loss for dense object detection, in: Proceedings of the IEEE International Conference on Computer Vision, pp. 2980–2988.

Liu, G., Reda, F.A., Shih, K.J., Wang, T.C., Tao, A., Catanzaro, B., 2018a. Image inpainting for irregular holes using partial convolutions, in: Proceedings of the European Conference on Computer

Vision (ECCV), pp. 85–100.
Liu, G., Shih, K.J., Wang, Ting-Chun andReda, F.A., Sapra, K., Yu, Z., Tao, A., Catanzaro, B., 2018b. Partial convolution based padding. arXiv preprint arXiv:1811.11718 .
Lossau, T., Nickisch, H., Wissel, T., Hakmi, S., Spink, C., Morlock, M., Grass, M., 2019. Dynamic pacemaker artifact removal (DyPAR) from CT data using CNNs. Medical Imaging with Deep Learning (MIDL), https://openreview.net/forum?id=rkx5InjA1N .
Mak, G.S., Truong, Q.A., 2012. Cardiac CT: imaging of and through cardiac devices. Current Cardiovascular Imaging Reports 5, 328–336.
Meyer, E., Raupach, R., Lell, M., Schmidt, B., Kachelrieß, M., 2010. Normalized metal artifact reduction (NMAR) in computed tomography. Medical Physics 37, 5482–5493.
Meyer, E., Raupach, R., Lell, M., Schmidt, B., Kachelrieß, M., 2012. Frequency split metal artifact reduction (FSMAR) in computed tomography. Medical Physics 39, 1904–1916.
Mouton, A., Megherbi, N., van Slambrouck, K., Nuyts, J., Breckon, T.P., 2013. An experimental survey of metal artefact reduction in computed tomography. Journal of X-ray Science and Technology 21, 193–226.
Nazeri, K., Ng, E., Joseph, T., Qureshi, F., Ebrahimi, M., 2019. Edgeconnect: Generative image inpainting with adversarial edge learning. arXiv preprint arXiv:1901.00212 .
Oehler, M., Buzug, T.M., 2007. The λ-mlem algorithm: An iterative reconstruction technique for metal artifact reduction in CT images, in: Advances in Medical Engineering. Springer, pp. 42–47.
Park, H.S., Lee, S.M., Kim, H.P., Seo, J.K., Chung, Y.E., 2018. CT sinogram-consistency learning for metal-induced beam hardening correction. Medical Physics 45, 5376–5384.
Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer. pp. 234–241.
Shepard, D., 1968. A two-dimensional interpolation function for irregularly-spaced data, in: Proceedings of the 1968 23rd ACM National Conference, ACM. pp. 517–524.
van Slambrouck, K., Nuyts, J., 2012. Metal artifact reduction in computed tomography using local models in an image block-iterative scheme. Medical Physics 39, 7080–7093.
Toftegaard, J., Fledelius, W., Seghers, D., Huber, M., Brehm, M., Worm, E.S., Elstrøm, U.V., Poulsen, P.R., 2014. Moving metal artifact reduction in cone-beam CT scans with implanted cylindrical gold markers. Medical Physics 41, 121710.
Wang, G., Snyder, D.L., O'Sullivan, J.A., Vannier, M.W., 1996. Iterative deblurring for CT metal artifact reduction. IEEE Transactions on Medical Imaging 15, 657–664.
Wang, T.C., Liu, M.Y., Zhu, J.Y., Liu, G., Tao, A., Kautz, J., Catanzaro, B., 2018. Video-to-video synthesis. arXiv preprint arXiv:1808.06601 .
Xu, S., Dang, H., 2018. Deep residual learning enabled metal artifact reduction in ct, in: Medical Imaging 2018: Physics of Medical Imaging, International Society for Optics and Photonics. p. 105733O.
Zhang, X., Wang, J., Xing, L., 2011. Metal artifact reduction in X-ray computed tomography (CT) by constrained optimization. Medical Physics 38, 701–711.
Zhang, Y., Yu, H., 2018. Convolutional neural network based metal artifact reduction in X-ray computed tomography. IEEE Transactions on Medical Imaging 37, 1370–1381.