Self-contained deep learning-based boosting of 4D cone-beam CT reconstruction

Frederic Madesta
Department of Computational Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg 20246, Germany

Thilo Sentker
Department of Computational Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg 20246, Germany
Department of Radiotherapy and Radio-Oncology, University Medical Center Hamburg-Eppendorf, Hamburg 20246, Germany

Tobias Gauer
Department of Radiotherapy and Radio-Oncology, University Medical Center Hamburg-Eppendorf, Hamburg 20246, Germany

René Werner
Department of Computational Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg 20246, Germany

(Received 21 January 2020; revised 2 June 2020; accepted for publication 16 July 2020; published 15 October 2020)

Purpose: Four-dimensional cone-beam computed tomography (4D CBCT) imaging has been suggested as a solution to account for interfraction motion variability of moving targets like lung and liver during radiotherapy (RT) of moving targets. However, due to severe sparse view sampling artifacts, current 4D CBCT data lack sufficient image quality for accurate motion quantification. In the present paper, we introduce a deep learning-based framework for boosting the image quality of 4D CBCT image data that can be combined with any CBCT reconstruction approach and clinical 4D CBCT workflow.

Methods: Boosting is achieved by learning the relationship between so-called sparse view pseudo-time-average CBCT images obtained by a projection selection scheme introduced to mimic phase image sparse view artifact characteristics and corresponding time-average CBCT images obtained by full view reconstruction. The employed convolutional neural network architecture is the residual dense network (RDN). The underlying hypothesis is that the RDN learns the appearance of the streaking artifacts that is typical for 4D CBCT phase images — and removes them without influencing the anatomical image information. After training the RDN, it can be applied to the 4D CBCT phase images to enhance the image quality without affecting the contained temporal and motion information. Different to existing approaches, no patient-specific prior knowledge about anatomy or motion characteristics is needed, that is, the proposed approach is self-contained.

Results: Application of the trained network to reconstructed phase images of an external (SPARE challenge) as well as in-house 4D CBCT patient and motion phantom data set reduces the phase image streak artifacts consistently for all patients and state-of-the-art reconstruction approaches. Using the SPARE data set, we show that the root mean squared error compared to ground truth data provided by the challenge is reduced by approximately 50% while normalized cross correlation of reconstruction and ground truth is improved up to 10%. Compared to direct deep learning-based 4D CBCT to 4D CT mapping, our proposed method performs better because inappropriate prior knowledge about the patient anatomy and physiology is taken into account. Moreover, the image quality enhancement leads to more plausible motion fields estimated by deformable image registration (DIR) in the 4D CBCT image sequences.

Conclusions: The presented framework enables significantly boosting of 4D CBCT image quality as well as improved DIR and motion field consistency. Thus, the proposed method facilitates extraction of motion information from severely artifact-affected images, which is one of the key challenges of integrating 4D CBCT imaging into RT workflows. © 2020 The Authors. Medical Physics published by Wiley Periodicals LLC on behalf of American Association of Physicists in Medicine. [https://doi.org/10.1002/mp.14441]

Key words: 4D cone-beam computed tomography, artifact reduction, deep learning, image reconstruction, image registration, image quality

1. INTRODUCTION

Modern complex radiation therapy (RT) techniques like hypofractionated volumetric modulated arc therapy (HVMAT) enable highly precise target irradiation, but also rely on highly accurate target localization. To minimize uncertainties during treatment delivery, different imaging systems are employed before and during the course of the treatment to frequently...
control the target position, leading to the term of image-guided RT (IGRT).\textsuperscript{2–3} For many lesion sites, it is meanwhile clinical standard to acquire a three-dimensional cone-beam computed tomography (3D CBCT; using imaging devices integrated into treatment machines) directly before the individual treatment fractions to evaluate residual patient setup errors and target position shifts by comparison of the 3D CBCT images to treatment planning 3D CT images. Three-dimensional image guidance is, however, insufficient to counter uncertainties inherent to thoracic and abdominal lesions, which are known to exhibit breathing-induced motion amplitudes during treatment up to several centimeters\textsuperscript{6,7} and significant interfraction motion variability.\textsuperscript{7} To increase dose application accuracy and precision for moving targets, it has been recommended to acquire time-resolved, that is, 4D (= 3D + t; instead of aforementioned 3D CBCT data) CBCT images in order to verify that tumor and organ motion patterns at the time of treatment correspond to the treatment planning situation and the related treatment planning four-dimensional (4D) CT.\textsuperscript{8} However, integration of 4D CBCT into clinical 4D RT workflows is hampered by low quality of the individual phase images, that is, the 3D CBCT images that represent the anatomy of the patient at different breathing phases, which leads to unreliable error-prone, target localization, and motion quantification.\textsuperscript{6}

The low image quality of current 4D CBCT phase images is inherent to the CBCT data acquisition procedure and its combination with standard reconstruction approaches. While a typical patient breathing cycle length is about 4 s,\textsuperscript{9} CBCT projection data acquisition time is defined by the duration of a full rotation of the linac gantry, which is in the order of 60 s. For 4D CBCT image reconstruction, the acquired projection data are binned according to associated breathing phase information derived from a simultaneously recorded breathing signal. Reconstruction of a 3D CBCT at a specific breathing phase is then based on only the projection data of the desired phase. Thus, assuming $N_b$ phase bins to be given, this approach discards $1 - 1/N_b$ or 90\% for a typical value of $N_b = 10$ of the whole set of acquired projection data during reconstruction of a CBCT phase image. This, in turn, results in images severely affected by sparse view sampling artifacts, also known as streak artifacts.\textsuperscript{10}

Motivated by the clinical relevance of the problem, various approaches have already been suggested to minimize the artifacts and, thus, to increase 4D CBCT image quality. Straightforward concepts are to adapt and modify CBCT imaging parameters, that is, to increase acquisition time, the total number of projections and imaging dose to reduce CBCT image artifacts. This, however, leads to higher imaging dose, does therefore not conform to the ALARA standard for imaging (As Low as Reasonable Achievable), and the overall image quality is only slightly improved for 4D reconstruction.\textsuperscript{11,12} An alternative approach is to integrate prior information into the 4D CBCT reconstruction process, for example, by integrating a motion model and/or image data derived from the patient’s original treatment planning 4D CT.\textsuperscript{13,14} Yet, utilizing prior information may lead to biased image appearance and motion quantification. Most sophisticated approaches are further utilizing image (post-)processing of either the acquired CBCT raw data (e.g., for projection data interpolation\textsuperscript{15}) or already reconstructed images in conjunction with deformable image registration (DIR) between the CBCT phase images to extract motion fields and to utilize the fields to generate CBCT phase images of higher quality.\textsuperscript{16,17} Although respective results are promising, major drawbacks of these approaches are relatively long computation times and potential uncertainties introduced by DIR of the artifact-affected data.

Addressing above-mentioned issues of existing 4D CBCT reconstruction approaches, the contributions of the present work to the body of knowledge are:

1. We propose a novel self-contained 4D CBCT reconstruction boosting method to improve the image quality of the CBCT phase images that is independent of prior information and assumptions about the patient-specific breathing patterns.
2. The proposed boosting approach does not require adjusting imaging parameters or increasing imaging dose and can be combined with any CB CT reconstruction approaches and clinical 4D CBCT workflows.
3. The boosting approach is computationally efficient during application (application time of only a few seconds).
4. The boosting approach is shown to significantly improve 4D CBCT image quality even if combining it with highly specialized state-of-the-art 4D CBCT reconstruction approaches.
5. To foster reproducibility of our research, (i) the evaluation is based on a publicly available dataset, (ii) motion phantom measurements with known ground truth that are provided as open data alongside with this publication, and (iii) the source code and the boosting models are made also provided as open source.

Methodically, the proposed boosting method is based on deep learning using a convolutional neural network (CNN) architecture inspired by top performance in the context of super-resolution and image restoration. A preliminary version of this study has been presented as part of a conference contribution,\textsuperscript{18} which now has been extended to a comprehensive framework. Source code, trained boosting models, and the in-house phantom data set are provided via https://github.com/IPMI-ICNS-UKE/4d-cbct-boosting.

2. MATERIALS AND METHODS

The main idea of the boosting framework is to learn the relationship between low-quality, artifact-containing 3D CBCT images and high-quality 3D CBCT images of exactly the same patient geometry and breathing state, that is, image pairs not affected by breathing phase differences and with an image appearance not being biased toward a specific breathing phase. Moreover, the low-quality images are supposed to
closely resemble 4D CBCT phase image characteristics in terms of image quality and contained streak artifacts. Therefore, we propose a specific projection selection scheme to reconstruct pseudo-time-averaged CBCTs of phase image-inherent image quality, serving as low-quality images; high-quality images are corresponding 3D time-average CBCTs reconstructed using all available projection data. After training a CNN to transform low-quality into high-quality average CBCTs, the network can be applied to boost reconstructed 4D CBCT phase images in order to reduce the streaking artifacts while maintaining the temporal and motion information represented by the phase images. This idea is subsequently formalized. For better understanding, an overview of the framework workflow is outlined in Fig. 1.

2.A. Cone-beam computed tomography image acquisition and reconstruction

In CBCT imaging, the x-ray source moves along a circular axial curve \( f : I \subseteq [0,2\pi) \rightarrow \mathbb{R}^3 \) around the patient and projects the density function

\[
f : \mathbb{R}^3 \rightarrow \mathbb{R}, \quad \text{supp}(f) = \Omega \subset \mathbb{R}^3
\]

(1)
of the imaged objects onto the two-dimensional (2D) flat-panel detector.

Using a x-ray projection operator \( X_\Gamma \), the projection can be described by

\[
X_\Gamma f = p.
\]

(2)

where by \( p \) is a function in the projection space

\[
p : \xi \times I \rightarrow \mathbb{R}, \quad \xi = \xi_{\text{col}} \times \xi_{\text{row}},
\]

(3)

with \( \xi_{\text{col}}, \xi_{\text{row}} \subset \mathbb{R} \) being coordinates along the detector column and row axis. The inverse transformation, that is, 3D image reconstruction, is symbolically denoted as

\[
X_\Gamma^{-1} p = \tilde{f}, \quad \tilde{f} : \Omega \rightarrow \mathbb{R}
\]

(4)

and results in a reconstructed image \( \tilde{f} \). In general, \( \tilde{f} \) is only an approximation of \( f \), caused by practical limitations; especially a limited number of projections renders CBCT image reconstruction an ill-posed problem. Hence, the objective of \( X_\Gamma^{-1} \) is to approximate the actual \( f \) as closely as possible.

During 4D CBCT imaging, each 2D projection \( p|_i, i \in I \) is associated with a respiratory amplitude value \( A : I \rightarrow \mathbb{R} \), extracted from a breathing signal that is recorded in parallel to the projection data acquisition. Typically, the amplitude values and projection data are re-assigned to a respiratory phase \( \phi : I \rightarrow [0,2\pi) \) that is calculated for every projection location \( i \in I \), assuming a sinusoidal-like breathing curve \( A(i) \propto \sin(\phi(i) + \phi_0) \). That is, a linearly increasing phase with a fixed offset \( \phi_0 \). Based on the phase information, each \( p|_i \) is finally assigned to one of the \( N_b \) bins \( b_j = 1, \ldots, N_b \),

\[
b_j = \left\{ i \middle| \frac{2\pi}{N_b} (j - 1) - \frac{\pi}{N_b} \leq \phi(i) < \frac{2\pi}{N_b} j - \frac{\pi}{N_b} \right\}
\]

(5)

these bins and corresponding projection data define the basis of the reconstruction of the individual 3D phase images that eventually form the 4D CBCT image series, that is,

\[
X_\Gamma^{-1} p|_{b_j} = \tilde{f}_{b_j}
\]

(6)

represents the reconstruction of a 3D CBCT image that corresponds to the \( j \)-th bin and breathing phase, respectively.

2.B. Projection selection for pseudo-average cone-beam computed tomography generation

Let \( X_{\Gamma\text{bin}}, X_{\Gamma\text{full}} \) be arbitrary reconstruction operators for sparse view and full projection reconstruction, respectively. The objective of the projection selection for pseudo-average...
CBCT reconstruction is then to gather a proper projection subset $p_{pa}^j$, such that the following conditions are fulfilled:

1. $\mathcal{X}_{\Gamma_{bin}}^{-1}p_{pa}^j$ has the same artifact characteristics as $\mathcal{X}_{\Gamma_{bin}}^{-1}p_{bj}^j$; in particular, angular spacing and manifestation of the streak artifacts are almost identical.

2. The reconstruction using all projections, that is, $\mathcal{X}_{\Gamma_{bin}}^{-1}p$, is a high-quality and streak-free representation of $\mathcal{X}_{\Gamma_{bin}}^{-1}p_{pa}^j$.

**Algorithm 1** Pseudo-average projection selection

```plaintext
1: procedure SELECT(p, φ, N_b, j)

2:  $j_0 \leftarrow j$; $p_{pa}^j \leftarrow \{\} \quad \triangleright$ initialize

3:  for $c = 0, \ldots, \max(C(i))$ do \quad \triangleright loop over all cycles

4:     $I_c = \{i \in I | C(i) = c\}$ \quad \triangleright select cycle indices

5:     $p_{c,j} = p_b(i_1,j_1)$ \quad \triangleright select bin $j$ in cycle $c$

6:     $p_{pa}^j \leftarrow p_{pa}^j \cup p_{c,j}$ \quad \triangleright add to projection set

7:     $j \leftarrow (j + 1) \mod N_b \quad \triangleright$ increment $j$

8: end for

9: return $p_{pa}^j \quad \triangleright$ return pseudo-average projection set

10: end procedure
```

To define an appropriate projection selection scheme, we introduce a mapping $C : I \rightarrow \mathbb{N}$ between the x-ray source position index and the corresponding breathing cycle\(^8\)

$$C(i) = \sum_{(k,l) \in I \times I} \left[ \left[ \exists \ell \in I : k < \ell < l \right] \wedge (\phi(l) < \phi(k)) \right]. \quad (7)$$

$[\cdot]$ is the Iverson bracket notation, counting the occurrences of starting respiratory cycles, defined by a jump discontinuity in $\phi$ regarding the interval boundaries of $[0,2\pi)$. In order to select a proper pseudo-average projection set $p_{pa}^j$, we apply Algorithm 1, with $j \in [1, N_b]$ being the starting respiratory bin index. For instance, let $j = 1$; then, for breathing cycle $c = 0$ (and $c = N_b$, $c = 2N_b$, etc.), projections belonging to the first breathing phase interval ($j = 1$) are selected; for $c = 1$ (and $c = N_b + 1$, $c = 2N_b + 1$, etc.), data corresponding to breathing phase bin $j = 2$ is added to $p_{pa}^j$. This procedure is conducted for the entire respiratory curve to obtain the final pseudo-average projection subset $p_{pa}^j$. Algorithm 1 is carried out for all $j \in [1, N_b]$ yielding

$$p_{pa}^j \cup p_{pa}^{j+1} \cup \cdots \cup p_{pa}^{j_{N_b}} = p,$$ \quad (8)

with the breathing phases corresponding to the contained data being almost equally distributed for each individual subset. A schematic visualization of this projection selection algorithm can be found in the supplementary material (Fig. S2).

### 2.C. Image reconstruction

Based on the introduced binning/selection approaches described by Eq. (5) and Algorithm 1, we subsequently differentiate between the following three reconstruction types and images:

$$\mathcal{X}_{\Gamma_{bin}}^{-1}p_{bj}^j : \quad \text{phase image},$$

$$\mathcal{X}_{\Gamma_{bin}}^{-1}p_{pa}^j : \quad \text{pseudo-time average image},$$

$$\mathcal{X}_{\Gamma_{bin}}^{-1}p = \tilde{f} : \quad \text{real time average image}.$$ 

Image quality and artifact characteristics of the reconstructed images are illustrated in Fig. 2; 4D animations of both a phase image as well as the corresponding pseudo-time-average image can be found in the supplementary material (Movie S1 & S2). Image reconstruction for the $N_b$ projection data subsets $p_{pa}^j$ by an $\mathcal{X}_{\Gamma_{bin}}^{-1}$ operator results in $N_b$ artifact-affected image representations $f_{pa}^j$ of the real average image reconstruction $\mathcal{X}_{\Gamma_{bin}}^{-1}p$, which exploits the entire set of available projection data. The figure also shows that, due to the proposed projection selection scheme, artifact characteristics of $f_{bj}$ and $f_{pa}^j$ are similar; however, different to $f_{bj}$, $f_{pa}^j$ does not correspond to a specific breathing state but features a temporally averaged representation of the patient.

### 2.D. Deep learning-based boosting framework

To approximate the sought boosting function

$$B : C(\Omega, \mathbb{R}) \rightarrow C(\Omega, \mathbb{R}), \quad f_{pa}^j \rightarrow \tilde{f} \quad \forall j,$$ \quad (9)

with $C(\Omega, \mathbb{R})$ as the space of continuous functions mapping $\Omega$ onto $\mathbb{R}$, in principle any image-to-image mapping CNN with matching input and output shapes can be applied. Here,

![Fig. 2. Comparison of a phase image $\tilde{f}_{bj}$ (left, here: $j = 1$), a pseudo-average image $\tilde{f}_{pa}^j$ (middle, here: $j = 1$), and the corresponding full projection-reconstructed average image $f$ (right). As desired, artifact characteristics of the pseudo-average closely match those of the phase image. In particular, the angular spacing of the streak artifacts as well as the overall visual image impression are similar. The red contour depicts the patient’s tumor area. [Color figure can be viewed at wileyonlinelibrary.com]](image-url)
the employed CNN is based on the residual dense network (RDN), which is one of the top-performing models in image super-resolution and restoration tasks.\textsuperscript{19,20} The main RDN component is a residual dense block (RDB), which is illustrated in Fig. 3 and is characterized by following features:

1. \textit{Densely connected convolutional layers:} Introduced with DenseNet (dense convolutional network),\textsuperscript{21} the idea is to use concatenated feature maps of all preceding layers as input to subsequent layers. Extracted features are therefore globally available to subsequent network parts. This, in turn, facilitates feature reuse and allows for a considerable reduction of network parameters while still achieving top performance on, for example, object recognition benchmark data sets.\textsuperscript{21}

2. \textit{Local feature fusion:} To control the RDB output features, a $1 \times 1$ convolution is applied to the concatenation of the preceding RDB’s output and the resulting feature maps of the considered RDB. By using a fixed number of RDB output features, deeper networks remain trainable.

3. \textit{Local residual learning:} Since it is difficult for neural networks, especially with increasing depth, to learn the identity function,\textsuperscript{22} a local skip connection is implemented, which is added to the output tensor yielded by local feature fusion. As a consequence, the convolutional branch has to learn only the residual.

Due to the large image size of $\tilde{f}_{\text{pa}}^0$ and $\tilde{f}$, the RDN is composed of only two RDBs with four densely connected convolutional layers each to keep both graphics processing unit (GPU) memory requirements manageable as well as to restrict the total trainable parameters (in this case to approximately 275,000), allowing fitting the model with limited data sets. Analogous to local residual skip connections, the low-quality representation $\tilde{f}_{\text{pa}}^0$ is directly passed to the output layer, where it is added to the output of the RDB cascade. The final network (Fig. 4) is implemented using TensorFlow v2.0.\textsuperscript{23} In contrast to common networks employing residual connections, in particular RDN, all convolutional layers and especially nonlinear activation functions after the summation layer are removed, so that the total network function $B$ can be split into two parts,

$$B = B_{\text{res}} + B_{\text{skip}},$$

with $B_{\text{res}}$ being the nonlinear residual mapping and $B_{\text{skip}}$ the linear skip connection

$$B_{\text{skip}} : \tilde{f}_{\text{pa}}^0 \rightarrow \tilde{f}_{\text{pa}}^0, \quad \alpha \in (0, 1] \quad \forall j.$$ \hspace{1cm} (11)

In order to illustrate the information flow through the neural network, in particular with focus on the two branches, we express — with no loss of generality — the low-quality representation $\tilde{f}_{\text{pa}}^0$ as the sum of the high-quality average image $\bar{f}$ and an image-specific artifact representation $\Delta \tilde{f}_{\text{pa}}^0$:

$$\tilde{f}_{\text{pa}}^0 = \bar{f} + \Delta \tilde{f}_{\text{pa}}^0.$$ \hspace{1cm} (12)

Combining Eqs. (9)–(12) results in

\begin{equation}
B(\tilde{f}_{\text{pa}}^0) = (B_{\text{res}} + B_{\text{skip}})(\tilde{f}_{\text{pa}}^0) \\
= B_{\text{res}}(\tilde{f}_{\text{pa}}^0) + B_{\text{skip}}(\tilde{f}_{\text{pa}}^0) \\
= B_{\text{res}}(\tilde{f}_{\text{pa}}^0) + \alpha \bar{f} + \alpha \Delta \tilde{f}_{\text{pa}}^0 = \tilde{f}
\end{equation}

and therefore

$$B_{\text{res}}(\tilde{f}_{\text{pa}}^0) = (1 - \alpha)\bar{f} - \alpha \Delta \tilde{f}_{\text{pa}}^0.$$ \hspace{1cm} (14)
Thus, using \( x < 1 \), the residual branch of the network is forced to learn features of both the high-quality average image \( f \) and the artifact image \( \Delta f_p \). This eventuates applying the CNN-based boosting function \( B \) that was trained on (pseudo-)average CBCT images to the real 4D CBCT phase images \( f_{b_j} \) if they have similar artifact characteristics to the training data. This, however, was postulated by condition 1 in Section 2.B and satisfied by the proposed projection selection scheme for pseudo-average CBCT image generation.

3. EXPERIMENTS

3.A. 4D CBCT data sets

Evaluation of the proposed boosting scheme was based on three different 4D CBCT data sets: two in-house data sets, consisting of clinical image data as well as phantom measurements, and a subset of the SPARE challenge data set.

3.A.1. In-house patient data set

The in-house data set comprised 20 liver and lung tumor patients that underwent radiotherapy treatment at the University Medical Center Hamburg-Eppendorf (retrospective data analysis; ethics committee waiver WF-82/18). For each patient, a 4D CBCT scan was performed directly before each of the five treatment sessions. The CBCT scans were performed using the built-in CBCT imaging system of the Varian TrueBeam v2.5 linear accelerator. To assign breathing phase information to the projection data, breathing records were acquired in parallel to CBCT projection data acquisition using the Varian real-time position management system. Key imaging parameters are summarized in Table I. As our clinical imaging protocol is directly derived from a 60 s half-fan beam 3D CBCT protocol (900 projections), reconstructing a 4D image data set of reasonable image quality is considerably challenging. Therefore, the state-of-the-art ROOSTER algorithm\(^{24} \) is used for \( \lambda_{\Gamma, \text{bin}} \); the FDK algorithm\(^{25} \) for \( \lambda_{\Gamma, \text{full}} \). For reconstruction, the RTK (The Reconstruction Toolkit\(^{26} \)) toolkit was applied and the GPU used to speed up the reconstruction process. All images were reconstructed using a \( 1 \times 1 \times 1 \) mm\(^3 \) voxel spacing and a total image size defined by the field of view given by the particular imaging geometry. To fit the network’s input and output shape, all axial image slices were then cropped to \((448 \times 448)\) pixel; the image volumes consisted of typically 220 axial image slices. No additional pre- or postprocessing steps such as correction for scatter or polyenergetic spectrum were carried out. In total, 1000 (= 20 patients \( \times 5 \) treatment fractions \( \times 10 \) breathing phases) CBCT phase images \( f_{b_j} \) were available for subsequent experiments.

3.A.2. Phantom data set

To analyze both the image characteristics of the pseudo-average images as well as the final image quality improvement after boosting using a well-defined setup, we additionally performed phantom measurements using the CIRS dynamic thorax phantom (Computerized Imaging Reference Systems, Inc., VA, USA). A spherical insert of 30 mm diameter was inserted and continuously moved along z-direction (inferior–superior) during the scan to simulate the tumor trajectory

\[
\tau_j(t) = A_0 \cdot \cos^4 \left( \frac{\pi}{T} t \right),
\]

with periods of \( T = 2.5 s; 5.0 s, 7.5 s \) and amplitude \( A_0 = 20 \) mm. The applied 4D CBCT scan protocol was the same as for the patient data. High-quality and time-resolved reference images were acquired by a Siemens SOMATOM 4D CT scanner (Siemens Healthcare GmbH, Forchheim, Germany).

3.A.3. SPARE challenge data set

For the in-house patient data set, no ground truth 4D CBCT reconstructions were available that allowed for quantitative analysis of image quality before and after boosting. Therefore, we utilized the clinical Varian data subset of the SPARE challenge data set\(^{27,28} \) as an external data set, comprising five patients with six 4D CBCT scans each. For each scan, the following data relevant for evaluation purposes were available:

1. A full half-fan beam projection set \( p_{\text{full}} \) with about 2400 projections and corresponding respiratory phase \( \phi \) and phase bins \( b_j \),
2. a subsampled version \( p_{\text{sub}} \) of this full projection set with 680 projections and corresponding respiratory phase \( \phi \) and phase bins \( b_j \),
3. ten reference 4D CBCT bins \( f_{b_i, \text{set}} \) reconstructed using the full projection set via the FDK algorithm, and
4. binary masks \( \Omega_m \subset \Omega \) for body, lungs, and planned target volume (PTV), that is, tumor region.

The subsampled projections were reconstructed using a conjugate gradient (CG)-based variant of the McKinnon-Bates (MKB) algorithm\(^{29,30} \) for both the phase bin and pseudo-average images. To match the reference image size defined by SPARE, all images were reconstructed to an in-plane (= axial) image size of \((450 \times 450)\) pixel using an isotropic voxel size of \( 1 \) mm\(^3 \). The axial slices were cropped in the same way as for the in-house data set to match the network’s input size of \((448 \times 448)\) pixel.

| Table I. Key four-dimensional cone-beam computed tomography (4D CBCT) imaging parameters used to acquire the in-house data set: x-ray tube voltage \((U)\) and current \((I)\), pulse length \((PL)\), number of projections \((#P)\), overall scan duration \((SD)\), trajectory \((\text{Traj.)}\) and the mode and filtration setting \((\text{Mode/filt.)})\). |
|---|---|---|---|---|---|
| \( U \) (kV) | \( I \) (mA) | \( PL \) (ms) | \#P | SD (s) | Traj. | Mode/filt. |
| 125 | 15 | 20 | 900 | 60 | 2\pi | Half fan/half bowtie |

Medical Physics, 47 (11), November 2020
3.B. Network training

Network training was based on the axial slices (with respect to the CBCT rotation axis) of the (pseudo-)time-average CBCT data, since these contain continuous streak artifacts that are in the focus of the intended image quality boosting. As the reconstructed images were not normalized to a bounded scale like the Hounsfield scale but reconstructed directly from projection data, the voxel intensity values were normalized to facilitate network training. To this end, the image intensities were clipped using the 1st/99th (in-house data sets) and 5th/95th percentile (SPARE data set) of the image intensity distribution, respectively, followed by intensity rescaling to [0,1]. The parameter \(x\) weighting the skip connection of the applied neural network was set to 0.9. This value was found by an extensive hyperparameter optimization and resulted both in the most stable training process as well as top performing training/validation losses.

Due to the fact that the image acquisition starts and ends with arbitrary respiratory phases, the \(N_b\) breathing phases are in general not included with exactly equal frequency in the pseudo-average projection data. Thus, \(f_{\text{proj}}\) may be slightly biased toward the real respiratory phase bin \(j\). To account for this, all \(N_b\) artifact-affected slices corresponding to one high-quality image slice were considered an individual mini-batch during stochastic gradient descent (Adam optimizer\(^31\)), ensuring an unbiased gradient calculation and update of the network parameters during training. Furthermore, the training scheme included a learning rate decay defined by

\[
\lambda_{\text{ep}}(i_{\text{ep}}) = 10^{-3} \cdot 0.9^{i_{\text{ep}}}
\]

(16)
to improve convergence. Here, the learning rate \(\lambda_{\text{ep}}\) was multiplied with 0.9 every five epochs \(i_{\text{ep}} = 5, 10, \ldots\), with the initial learning rate set to \(\lambda_{\text{ep}}(1) = 10^{-3}\) according to.\(^31\) The whole network was trained for 50 epochs until convergence. Corresponding plots for all tested values of \(x\) are available in the supplementary material (Fig. S1).

3.C. Boosting/network application

To boost a reconstructed phase image \(\tilde{f}_{\text{b}}\), the trained network \(B\) was applied axial slice-wise until the whole volume was processed, that is, about 220 forward passes were needed to cover the whole scan range.

3.D. Evaluation

3.D.1. Pseudo-average image characteristics

To validate the proposed projection selection (cf. Algorithm 1) and compare the corresponding reconstructed pseudo-average images with real phase images, we analyzed the streak frequency as well as the angular streak spacing for said image types. We used the in-house phantom data set (as described in Section 3.A.2) to perform this evaluation in a well-defined setting. As streak artifacts are characterized by linear intensity edges, we employed the Hough transform as a feature extractor to each axial slice to find these streak edges. We kept the transformation parameters and thresholds fixed in favor of comparability between pseudo-average and phase images. Sought streak locations and angular orientations can directly be obtained from the resulting Hough space. The exact same procedure was applied to the clinical in-house patient data set and compared to the phantom results.

3.D.2. Image quality

Due to different imaging conditions and x-ray projection normalization, separate instances of the boosting model were trained for the in-house and the SPARE 4D CBCT data sets. Using the SPARE challenge data, quantitative evaluation of the image quality before and after boosting was carried out as follows: For each patient, the subsampled projection data set was used to reconstruct \(N_b = 10\) phase images \(f_{\text{b,j}}\),

\[
\mathcal{X}_{\text{proj,sub}}^{\text{b,j}} = \tilde{f}_{\text{b,j}}, \quad j = 1, \ldots, N_b,
\]

(17)
that were to be boosted. Thereafter, the trained network \(B\) was applied to each of the ten respiratory phase images, yielding the boosted representations \(B(\tilde{f}_{\text{b,j}})\). The \(B(\tilde{f}_{\text{b,j}})\) were compared to the reference 4D CBCT phase images \(\tilde{f}_{\text{b,j,ref}}\) provided as ground truth. Following,\(^32,33\) quantitative evaluation measures were root mean squared error (RMSE)

\[
\text{RMSE}(B(\tilde{f}_{\text{b,j}}),\tilde{f}_{\text{b,j,ref}}) = \sqrt{\frac{1}{\Omega_x} \int_{\Omega_x} (B(\tilde{f}_{\text{b,j}})(x) - \tilde{f}_{\text{b,j,ref}}(x))^2 \, dx}
\]

(18)
and normalized cross correlation (NCC)

\[
\text{NCC}(B(\tilde{f}_{\text{b,j}}),\tilde{f}_{\text{b,j,ref}}) = \frac{\langle B(\tilde{f}_{\text{b,j}}) - B(\tilde{f}_{\text{b,j}}), \tilde{f}_{\text{b,j,ref}} - \tilde{f}_{\text{b,j,ref}} \rangle}{\left\| B(\tilde{f}_{\text{b,j}}) - B(\tilde{f}_{\text{b,j}}) \right\|_2 \left\| \tilde{f}_{\text{b,j,ref}} - \tilde{f}_{\text{b,j,ref}} \right\|_2},
\]

(19)
separately calculated for each mask \(\Omega_m\) and structure of interest, respectively. Here, \(\langle \cdot \rangle\) denotes the inner product and \(\| \cdot \|_2\) the \(L^2\)-norm.

Evaluation for the SPARE and in-house data set was performed by fivefold cross validation (CV): First, the respective folds were created for each data sets individually, followed by a fold-wise union. Thus, each combined training (test) set comprises 4 (1) SPARE and 16 (4) in-house patients. The phantom data set was not used for training in any step of our proposed framework.

3.D.3. Comparison to direct U-Net-based 4D CBCT-to-4D CT translation

To compare our results with state-of-the-art deep learning-based methods for enhancing the image quality of sparse-
view CT directly, we trained a framing U-Net using a fivefold CV on the SPARE data set. Here, we used the 4D CBCT phase images as input and the corresponding pretreatment 4D CT phase images as the model's output. Affine image registration was performed on the average images and subsequently applied to the 4D CT phase images after visual validation of the registration result. Along with the training U-Net, we applied our proposed method and trained the model on the same fivefold.

3.D.4. Deformation vector fields

To analyze effects of streak artifacts and the proposed boosting method on motion estimation by deformable image registration (DIR), as often performed during dose accumulation in RT context, we performed a cyclic DIR for each 4D CBCT of the in-house data set for both the nonboosted as well as the boosted images. Here, we utilize the open source toolbox VariationalRegistration to calculate deformation vector fields (DVF) \( \phi^{j+1} : \Omega \rightarrow \Omega \) between consecutive respiratory phase images \( \phi^{j+1} \):

- computation of \( \phi_1^2 \) : DIR of \( \tilde{f}_{b_{j+1}} \) → \( \tilde{f}_{b_{j+2}} \)
- computation of \( \phi_2^2 \) : DIR of \( \tilde{f}_{b_{j+2}} \) → \( \tilde{f}_{b_{j+3}} \)

...computation of \( \phi_{N_p}^2 \) : DIR of \( \tilde{f}_{b_{N_p-1}} \) → \( \tilde{f}_{b_{N_p}} \)
- computation of \( \phi_1^1 \) : DIR of \( \tilde{f}_{b_j} \) → \( \tilde{f}_{b_{j+1}} \)

The transformation \( \phi^{j+1} \) is found by minimizing the functional

\[
J[\phi^{j+1}] = D[\tilde{f}_{b_{j+1}}, \tilde{f}_{b_j} \circ \phi^{j+1}] + \alpha S[\phi^{j+1}] 
\]

with \( D \) and \( S \) being the distance measure and a regularization term, respectively. We set \( S \) to a Gaussian smoother with a variance of \( \sigma^2 = 2 \) to reduce high DVF gradients caused by moving streak artifacts. Demon forces were used for iteratively updating the vector field. The maximum number of iterations was set to 800 for each of the four resolutions of an applied image pyramid.

To quantify the cyclic consistency of the calculated DVF sequence, we computed the residual DVF \( \Delta \phi \) obtained by cyclic concatenation

\[
\Delta \phi = \phi_{N_p}^1 \circ \phi_{N_p-1}^2 \circ \cdots \circ \phi_1^2 = \text{id} + \varepsilon. 
\]

Assuming a cyclic average breathing cycle with respect to the scan duration, the residual DVF and thus the mean residual error \( \varepsilon \) are supposed to vanish for all three spatial dimensions

\[
\varepsilon = \frac{1}{|\Omega_m|} \int \int \int_{\Omega_m} ||\varepsilon(x,y,z)||_2 \, dx \, dy \, dz \rightarrow 0. 
\]

Here, the evaluation region \( \Omega_m \) is set to the lung mask of the image to be evaluated.

4. RESULTS

4.A. Pseudo-average image characteristics

The Hough transform consistently yielded prominent streak artifact intensity edges for both the pseudo-average images and the phase images. Averaged over all images of the phantom data set, the Hough transform extracted 22.9 ± 4.1 and 25.4 ± 4.4 streaks per slice with a mean angular spacing of 7.6 ± 1.9 and 6.6 ± 1.3 for the pseudo-average and phase images, respectively. The distance between selected phase bins is 11 for the pseudo-average images due to the introduced bin drift (cf. line 7 of Algorithm 1). This aligns with the theoretically estimated difference of approximately 10 % for both the streak frequency as well as angular spacing. Corresponding results for the in-house patient scans show that this 10 % discrepancy becomes negligible as it averages out over the whole data set mainly due to existing breathing period irregularity: For the patient data set, the mean number of streaks is 24.1 ± 2.9 and 24.6 ± 2.9 with an angular spacing of 6.9 ± 1.1 and 6.8 ± 1.0 for the pseudo-average and phase images, respectively.

4.B. Image quality

Exemplary results for our in-house data set are shown in Fig. 5. As can be seen from the axial slices, streak artifacts are reduced while air/tissue boundaries as well as bony structures are sharpened. Lung vessels surrounding the tumor (cf. Fig. 5, middle row) are substantially enhanced after boosting of the ROOSTER reconstructed phase image, despite the spatial extent of the structures being in the order of only a few voxels. Here, it should be stressed out that a variant of ROOSTER is one the top-performing reconstruction algorithms of the SPARE challenge. Furthermore, the overall visual image impression of the boosted converges toward the quality of the real average images reconstructed utilizing the full projection data set. When visualizing a full breathing cycle (see Fig. 6), it can be seen that our proposed method does not affect the breathing motion. In particular, the total motion amplitude remains unchanged.

Similarly, we also evaluated the image quality enhancement for the 4D CBCT phantom data set with a 4D CT image acting as ground truth. As can be seen in Fig. 7, the streak artifacts and blurriness are suppressed resulting in a significantly improved image quality. This was the case for all of the three tested breathing periods as well as all corresponding phase images (see Table SI of the supplementary material for detailed RMSE und NCC metrics). The overall motion range of the insert is not affected by our proposed method (cf. Fig. S3 of the supplementary material).

Similar observations regarding the image quality hold true for the SPARE 4D CBCT images. In Fig. 8, an unprocessed, CG-MKB-reconstructed phase image (Fig. 8, left) as well as the boosted representation (Fig. 8, middle) are depicted and compared to the corresponding reference reconstruction (Fig. 8, right). In the unprocessed image, the overall voxel
intensity appears to be too high, especially along streak artifacts. Such areas and effects are corrected after boosting, resulting in a more harmonized intensity distribution throughout the image. Furthermore, streak artifacts are strikingly suppressed especially inside the lungs, illustrating the capability of the model to distinguish between anatomical lung structures and sparse view sampling-related artifacts.

The qualitative impression is further supported by the quantitative analysis using the SPARE data set, see Table II: Image quality after boosting is consistently higher regarding both RMSE and NCC for all five leave-one-patient-out CV sets. In particular, the RMSE is reduced by about half after boosting. It should be noted that the absolute performance of CV 5, that is, patient P5 of SPARE, is worse compared to the
other CV folds (cf. body of CV 5 in Table II). This is mainly due to the fact that the patient’s body partially extends beyond the axial field of view resulting in truncation artifacts, that is, much brighter image intensities. This is the case for reconstruction algorithms using multiple forward- and backprojection steps, in particular, the applied CG-based MKB algorithm. The very same was observed and reported by the authors of SPARE.28

4.C. Comparison to direct U-Net-based 4D CBCT-to-4D CT translation

The results for the framing U-Net indicate that the overall performance is strongly dependent on the correlation between the pretreatment 4D CT image as prior knowledge and the actual 4D CBCT. In particular, patient P2 of the Varian subset of SPARE features considerable differences in one of its lungs. This directly leads to the fact that 25% of the training data is severely corrupted for four of five CV folds. As depicted in Fig. 9 this results in more blurred images for the corresponding test cases. The metrics of this comparison are listed in the supplementary material in detail in Table SII.

4.D. Deformation vector fields

For all 20 patients of the in-house data set, the DVFs calculated based on the nonboosted phase images are severely affected by streak artifacts, despite the fact that ROOSTER as

![Fig. 7. Resulting boosted image (right) of a ROOSTER reconstruction of the applied CIRS dynamic thorax phantom. The round insert is highlighted with the red arrow. Streak artifacts are significantly reduced inside the whole phantom while the insert keeps its circular shape. [Color figure can be viewed at wileyonlinelibrary.com]](image1)

![Fig. 8. Conjugate gradient (CG-MKB) reconstruction of a phase image of patient P1 (CV 1) of the SPARE challenge data set without (left) and with deep learning-based boosting (middle); the corresponding reference image is shown on the right. It is evident that the overall degradation of the image due to streak artifacts is considerably reduced, improving visibility of smaller inner-lung structures and details.](image2)

**Table II.** Results of the leave-one-patient-out cross validation (CV), using the Varian subset of the SPARE challenge data.

|        | CV 1       | CV 2       | CV 3       | CV 4       | CV 5       | Mean       |
|--------|------------|------------|------------|------------|------------|------------|
|        | CG-MKB     | Boosted    | CG-MKB     | Boosted    | CG-MKB     | Boosted    | CG-MKB     | Boosted    |
| RMSE (10^{-3} mm\(^{-1}\)) |            |            |            |            |            |            |            |            |
| **Body** | 3.9(5)     | 2.1(1)     | 3.7(3)     | 2.4(6)     | 5.7(4)     | 2.2(2)     | 3.6(1)     | 2.1(2)     | 9.3(9)     | 5.1(6)     | 5.2        | 2.8        |
| **Lungs** | 3.9(4)     | 2.0(1)     | 3.9(4)     | 2.3(4)     | 4.2(1)     | 2.1(2)     | 3.7(2)     | 1.9(1)     | 4.4(1)     | 1.9(1)     | 4.0        | 2.0        |
| **PTV**  | 3.1(5)     | 2.0(2)     | 3.4(4)     | 2.3(7)     | 4.0(2)     | 2.4(3)     | 3.4(2)     | 2.6(7)     | 3.8(2)     | 1.7(1)     | 3.5        | 2.2        |
| NCC     |            |            |            |            |            |            |            |            |            |            |            |            |
| **Body** | 0.878(19)  | 0.905(10)  | 0.873(44)  | 0.890(42)  | 0.702(64)  | 0.802(32)  | 0.906(16)  | 0.926(13)  | 0.102(58)  | 0.410(68)  | 0.692      | 0.787      |
| **Lungs** | 0.624(78)  | 0.673(65)  | 0.673(98)  | 0.693(81)  | 0.725(31)  | 0.747(31)  | 0.775(34)  | 0.808(26)  | 0.612(48)  | 0.659(50)  | 0.682      | 0.716      |
| **PTV**  | 0.655(68)  | 0.731(41)  | 0.527(127) | 0.544(111) | 0.635(116) | 0.680(78)  | 0.665(82)  | 0.697(88)  | 0.696(64)  | 0.723(68)  | 0.636      | 0.675      |

For each computed tomography (CT) set, namely CV 1 to CV 5, root mean squared error (RMSE) and normalized cross correlation (NCC) were calculated with respect to body, lungs, and planned target volume (PTV) as regions of interest. Mean and standard deviation calculated over the respective volume are stated using the concise notation, such that 0.123 ± 0.045 is equivalent to 0.123(45). Better results, that is, lower RMSE or higher NCC, are highlighted by bold font.

Medical Physics, 47 (11), November 2020
Fig. 9. Conjugate gradient (CG-MKB) reconstruction of a phase image of patient P1 (CV 1) of the SPARE challenge data set after boosting with the framing U-Net (left) and our proposed method (middle); the corresponding reference image is shown on the right. It is clearly visible, that the framing U-Net trained to directly map the 4D CBCT to the pretreatment 4D CT produces severely blurred images. In particular, most of the small lung structures completely vanished.

**TABLE III.** Results of the cyclic deformable image registration for the in-house data set.

| Recon. algo. | Boosting | x        | y        | z        | Total     |
|-------------|----------|----------|----------|----------|-----------|
| ROOSTER     | No       | 3.7 ± 3.1| 3.4 ± 1.8| 1.2 ± 0.5| 5.9 ± 3.7**|
|             | Yes      | 2.5 ± 2.0**| 2.3 ± 0.8**| 1.2 ± 0.5| 4.2 ± 2.2**|
| CG-MKB      | No       | 6.5 ± 4.1| 7.0 ± 1.4| 1.9 ± 0.6| 11.1 ± 4.4**|
|             | Yes      | 5.0 ± 2.1*| 5.4 ± 0.6**| 1.8 ± 0.5| 8.6 ± 2.2**|

The mean residual error is given for all three directions (x: right–left, y: posterior–anterior, z: inferior–superior); “total” refers to the error magnitude. Boosted values marked with one or more asterisks (*) signal significant differences between errors obtained after and without boosting (*P < 0.05/6; **P < 0.01/6; paired t test with Bonferroni correction).

a state-of-the-art 4D CBCT reconstruction algorithm utilizes total variation regularization in space and time. Since these artifacts rotate within the axial plane, that is, (x, y)-plane, when iterating through the Nₘ phases images, the main DVF contributions originate from the streak’s motion. As a result, the displacement in x- and y-direction accumulates over the full breathing cycle. Thus, as can be seen in Table III, the mean residual error [cf. Eq. (22)] evaluated over the lung region is largest in x- and y-direction even though the respective respiratory motion amplitude is much smaller compared to the z-direction. After application of the trained boosting model, these errors are reduced consistently for all 20 patients for said directions. As the CG-MKB algorithm does not utilize regularization at all, streak artifacts are even (much) more pronounced, resulting in larger residual DVF errors. Nevertheless, these errors are significantly reduced in x- and y-direction and in total after boosting (paired sample one-sided t-test with Bonferroni correction).

The error of the z-direction, which is nearly perpendicular to the rotating streak artifacts, remains unchanged. Figure 10 illustrates the clearly visible effect of boosting on the magnitude ∥ε(x, y, z)∥₂ of the residual DVF for one of the in-house data set patients.

5. DISCUSSION

We proposed a deep learning-based 4D CBCT boosting framework that improves the image quality of the respiratory phase images of a 4D CBCT image sequence, especially with regard to common streak artifacts that are due to sparse sampling of the full CBCT projection set when generating a time-resolved 3D image sequence. The boosting approach is independent of the particular reconstruction approach applied to reconstruct the phase and pseudo-average CBCT image. This was illustrated using state-of-the-art reconstruction approaches like CG-MKB and the ROOSTER algorithm employing both patient as well as phantom data. Moreover, after training the boosting CNN (training time in the order of approximately 1 min per epoch and CBCT scan in the training set), application is very fast (here: 3 s for 220 axial slices; Nvidia GeForce RTX 2080 Ti).

The quantitative image quality evaluation using the freely available SPARE challenge data set demonstrated that the presented boosting consistently enabled enhancement of the phase images for all CV sets and regions of interest. In particular, our proposed method seems to be robust against unseen artifacts, like it was the case for truncation artifacts in CV 5 test set. Moreover, it should be noted that the reference SPARE reconstructions f̂₀ ref still only represent an approximation of the actual unknown ground truth and contain streak artifacts (see Fig. 8, right). From that perspective, we assume the actual gain in image quality still to be underestimated by the applied evaluation procedure and the metrics to be very likely biased toward more streaky images. When comparing the results of the proposed method to the framing U-Net trained to map 4D CBCT directly onto the corresponding
pretreatment 4D CT, it is clearly evident that not using prior knowledge is advantageous if non-neglectable image differences are present, which is the case for SPARE.

Based on an in-house 4D CBCT data set comprising 100 60 s CBCT scans of 20 patients, we furthermore showed that the influence of rotating streak artifacts on DIR is significantly reduced, resulting in a cyclically more consistent image registration. Thus, we believe the improved image quality of the respiratory phase images to be of great value especially in the context of 4D IGRT and, for example, DIR-based tumor trajectory verification and motion field computation for motion modeling and dose calculation purposes. This will be studied in detail as part of our future work.

It is important to note that our study mainly focusses on streak artifacts caused by the heavily ill-posed 4D CBCT reconstruction problem. As for any binning-based 4D reconstruction approach, breathing irregularities and residual motion inside the phase bins itself result in blurring of regions affected by respiratory motion. However, when comparing these so-called motion artifacts with streak artifacts, it becomes clear that the former have a less pronounced impact on the overall 4D CBCT image quality. Nevertheless, reducing motion artifacts in 4D CBCT is the next logical step to improve the image quality and will be the subject of our future work.

6. CONCLUSIONS

The present paper demonstrates feasibility of deep learning-based improvement of the quality of reconstructed phase images of 4D CBCT image sequences without using any a-priori information (like planning 4D CT images, motion vector fields or respiratory trajectories). It is implemented as a self-contained boosting method to reduce image artifacts, in particular streak artifacts due to sparse view sampling. To train a boosting model, only projection raw data are required that is recorded anyway during clinical practice, usually to reconstruct a temporally averaged 3D CBCT. Thus, no additional data availability and collection requirements have to be met. Boosted CBCT phase images show a suppressed manifestation of streak artifacts, enabling calculation of cyclically more consistent DVF s and therefore an overall more plausible DIR.

The presented boosting framework is applicable to any 4D CBCT reconstruction method that makes use of projections binned by respiratory phase or amplitude. Therefore, it is not in competition with existing reconstruction methods, but rather an additional postprocessing step applied to the respective reconstruction. This facilitates easy integration into existing reconstruction pipelines.

ACKNOWLEDGMENT

This work was funded by DFG research grant WE 6197/2-1 and supported by Forschungszentrum Medizintechnik Hamburg, grant 02fmthh2017. We further thank NVIDIA for donating the applied GPU. RW and TG received funding from Siemens Healthcare, Erlangen, Germany, but declare that there is no conflict to disclose regarding this project. Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

FM and TS have no conflict to disclose.

REFERENCES

1. Teoh M, Clark CH, Wood K, Whitaker S, Nisbet A. Volumetric modulated arc therapy: a review of current literature and clinical use in practice. Br J Radiol. 2011;84:967–996.
2. Grills IS, Hugo G, Kestin LL, et al. Image-guided radiotherapy via daily online cone-beam CT substantially reduces margin requirements for stereotactic lung radiotherapy. Int J Radiat Oncol Biol Phys. 2008;70:1045–1056.
3. Sothmann T, Blanck O, Poels K, Werner R, Gauer T. Real time tracking in liver SBRT: comparison of Cyberknife and Vero by planning structure-based γ-evaluation and dose-area histograms. Phys Med Biol. 2016;61:1677–1691.
4. Wortel RC, Incrocci L, Pos FJ, et al. Acute toxicity after image-guided intensity modulated radiation therapy compared to 3D conformal
radiation therapy in prostate cancer patients. Int J Radiat Oncol Biol Phys. 2015;91:737–744.

5. Zelensky MJ, Kollmeier M, Cox B, et al. Improved clinical outcomes with high-dose image guided radiotherapy compared with non-IGRT for the treatment of clinically localized prostate cancer. Int J Radiat Oncol Biol Phys. 2012;84:125–129.

6. Keall PJ, Mageras GS, Balter JM, et al. The management of respiratory motion in radiation oncology report of AAPM Task Group 76. Med Phys. 2006;33:3874–3900.

7. Worm ES, Hansen AT, Petersen JB, Muren LP, Prastegaard LH, Hoyer M. Inter- and infratractical localisation errors in cone-beam CT guided stereotactic radiation therapy of tumours in the liver and lung. Acta Oncol. 2010;49:1177–1183.

8. Sweeney RA, Seubert B, Stark S et al. Accuracy and inter-observer variability of 3D versus 4D cone-beam CT based image-guidance in SBRT for lung tumors. Radiat Oncol. 2012;7:81.

9. Suh Y, Dieterich S, Cho B, Keall PJ. An analysis of thoracic and abdominal tumour motion for stereotactic body radiotherapy patients. Phys Med Biol. 2008;53:3623–3640.

10. Bian J, Sieversdjen JH, Han X, et al. Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT. Phys Med Biol. 2010;55:6575–6599.

11. Santoso AP, Song KH, Qin Y, et al. Evaluation of gantry speed on image quality and imaging dose for 4d cone-beam CT acquisition. Radiat Oncol. 2016;11:98.

12. Shimosugi Y, Araki F, Maruyama M, et al. Image quality of four-dimensional cone-beam computed tomography obtained at various gantry rotation speeds for liver stereotactic body radiation therapy with fiducial markers. Physica Medica. 2018;45:19–24.

13. Rit S, Woltfus JWH, van Herk M, Sonke JJ. On-the-fly motion-compensated cone-beam CT using an a priori model of the respiratory motion. Med Phys. 2009;36:2283–2296.

14. Zhang Y. Yang J, Zhang L, Court LE, Gao S, Balter PA, Dong L. Digital reconstruction of high-quality daily 4d cone-beam CT images using prior knowledge of anatomy and respiratory motion. Comput Med Imag Grap. 2015;40:30–38.

15. Zhang H, Knuis M, Sonke JJ. Directional sinogram interpolation for motion weighted 4d cone-beam CT reconstruction. Phys Med Biol. 2017;62:2254–2275.

16. Brehm M, Paysan P, Oelhafen M, Kachelries M. Artifact-resistant motion estimation with a patient-specific artifact model for motion-compensated cone-beam CT. Med Phys. 2013;40:101913.

17. Sappee S, Kuhm J, Brehm M, Paysan P, Seghers D, Kachelries M. Motion vector field phase-to-amplitude resampling for 4d motion-compensated cone-beam CT. Phys Med Biol. 2018;63:035032.

18. Frederic M, Tobias G, Thilo S, Werner R. Self-consistent deep learning-based boosting of 4d cone-beam computed tomography reconstruction, In: Medical Imaging 2019: Image Processing. International Society for Optics and Photonics; 2019;10949:1094902.

19. Zhang Y, Tian YY, Kong BZ, Raymond YF. Residual dense network for image restoration. arXiv, 2018.

20. Zhong Y, Tian YY, Kong BZ, Raymond YF. Residual dense network for image super-resolution. 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2018:2472–2481.

21. Huang G, Liu Z, van der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); IEEE; 2017.

22. Szegedy C, Loffe S, vanhoucke V. Inception-v4, inception-resnet and dense connections. In: 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16); 2016:265–283.

23. Mory C, Avrav V, Zhang B. Cardiac c-arm computed tomography using a 3D + time ROI reconstruction method with spatial and temporal regularization. Med Phys. 2014;41:021903.

24. Feldkamp LA, Davis LC, Kress JW. Practical cone-beam algorithm. J Opt Soc Am A. 1984;1:612.

25. Rit S, Oliva VM, Brousmiche S, Labarre R, Sarrut D, Sharp GC. The reconstruction toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the insight toolkit (ITK). J Phys Conf Ser. 2014;489:012079.