

TOPICAL REVIEW

Artifact Reduction in 3D and 4D Cone-Beam Computed Tomography Images With Deep Learning: A Review

MOHAMMADREZA AMIRIAN¹, DANIEL BARCO¹, IVO HERZIG²,
AND FRANK-PETER SCHILLING¹

¹Centre for AI (CAD), Zurich University of Applied Sciences (ZHAW), 8400 Winterthur, Switzerland

²Institute of Applied Mathematics and Physics (IAMP), Zurich University of Applied Sciences (ZHAW), 8400 Winterthur, Switzerland

Corresponding author: Frank-Peter Schilling (scik@zhaw.ch)

This work was supported in part by Innosuisse under Grant 56768.1 IP-LS.

ABSTRACT Deep learning based approaches have been used to improve image quality in cone-beam computed tomography (CBCT), a medical imaging technique often used in applications such as image-guided radiation therapy, implant dentistry or orthopaedics. While deep learning methods have been applied to reduce various types of CBCT image artifacts arising from motion, metal objects, or low-dose acquisition, a comprehensive review summarizing the successes and shortcomings of these approaches, with a primary focus on the type of artifacts rather than the architecture of neural networks, is lacking in the literature. In this review, the data generation and simulation pipelines, as well as artifact reduction techniques are specifically investigated for each type of artifact. We provide an overview of deep learning techniques that have successfully been shown to reduce artifacts in 3D, as well as in time-resolved (4D) CBCT through the use of projection- and/or volume-domain optimizations, or by introducing neural networks directly within the CBCT reconstruction algorithms. Research gaps are identified to suggest avenues for future exploration. One of the key findings of this work is an observed trend towards the use of generative models including GANs and score-based or diffusion models, accompanied with the need for more diverse and open training datasets and simulations.

INDEX TERMS Cone-beam computed tomography (CBCT), deep learning, artifacts.

I. INTRODUCTION

Cone-beam computed tomography (CBCT) is an imaging technique to acquire volumetric scans in medical domains such as implant dentistry, orthopaedics, or image-guided radiation therapy (IGRT). In particular, in the case of IGRT, on-board imaging mounted directly on radiotherapy machines is used to assess a patient's current anatomy before radiation treatment sessions. Changes in anatomy during the treatment period and since the acquisition of the planning CT (pCT) can lead to inefficiencies in the treatment process. Recent research has demonstrated that utilizing 3D or 4D (volumetric data with additional time dimension to

track motion) CBCT scans in IGRT [2] improves patient positioning and dose calculation for radiotherapy sessions.

The quality of CBCT scans suffers from similar types of artifacts as for spiral/helical CT scans, including those arising from beam hardening and scatter effects, metal implants, and patient motion. In addition, new artifacts arise due to the cone-beam geometry. Further, minimizing the radiation dose in radiotherapy is important for the safety of the patients. However, reducing the imaging dose per scan, acquiring fewer X-ray projections, or acquiring projection data from a limited angle can result in streak artifacts.

This paper provides an overview of the current body of research on artifact reduction in 3D and 4D CBCT with applications including, but not limited to, IGRT, aiming to improve scan quality while also minimizing the imaging radiation dose. The significant variation in the methods

The associate editor coordinating the review of this manuscript and approving it for publication was Gustavo Olague¹.

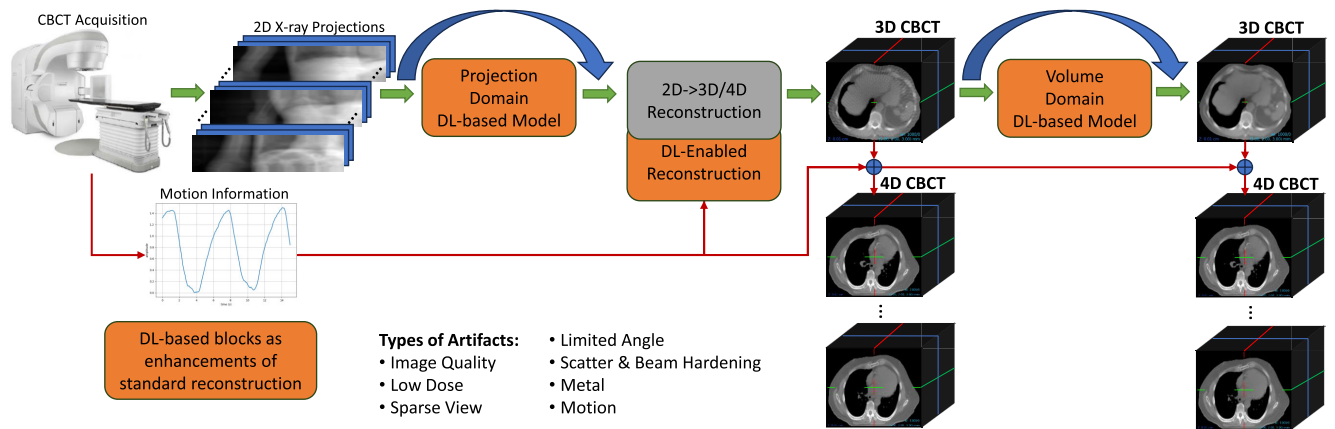


FIGURE 1. Visual Abstract: An illustration of the CBCT acquisition process in IGRT for lung CBCT and the application of deep learning for artifact correction. The diagram depicts the acquisition of 2D projections (initial corrections such as scatter corrections have already been applied), including (optionally) time- and motion-related information (e.g. breathing amplitude signal), standard CBCT reconstruction (typically 2D→3D), and DL-based components for image enhancement. Incorporating acquired temporal and motion information provides the opportunity to apply a projection binning which can be used to reconstruct 4D CBCT images (3D images at various states of motion). During the course of CBCT reconstruction, several types of artifacts (e.g. arising from cone-beam geometry, low dose, sparse view or limited angle scans, scatter, metal or beam hardening) can be mitigated through DL-based optimization in the projection and/or volume domain, or by improving (parts of) the reconstruction algorithm itself using neural networks. The illustration of a commercial radiotherapy system is adapted from [1].

and techniques used to mitigate different types of artifacts suggests to organize the literature based on the type of artifact. For instance, sparse-view artifacts can be addressed in the projection domain by interpolating new projections, but refining the original projections is not beneficial; however, motion artifact mitigation is possible through projection refinement. Furthermore, the survey aims to present a clear picture of all necessary steps in the artifact mitigation process for all relevant types of artifacts individually.

In particular, we review the current state-of-the-art research which uses deep learning (DL) [3] to reduce various artifacts in CBCT scans, and we categorize the research based on the types of artifacts they address. While Ref. [4] focuses broadly on the use of DL methods in IGRT, the closest literature reviews to our work are presented in references [5], [6], [7]. The first survey [5] is focused on synthetic CT generation from various types of input scans, including CBCT, with the aim to enhance the scan quality. Its content partially overlaps with what we present in Section III. However, it does not cover all the other artifacts which can degrade CBCT image quality as discussed after Section III. Ref. [6] discusses supervised, self-supervised, and unsupervised techniques for artifact reduction in CT scans, and it covers unrolling the reconstruction, as well as optimization methods in both the projection (raw 2D X-ray images) and volume (reconstructed 3D images) domains. However, it is essential to note that Ref. [6] primarily focuses on CT scans, which differs from the main focus of this work, namely CBCT scans. The third survey [7] provides an in-depth literature analysis, considering criteria such as anatomy, loss functions, model architectures, and training methods for supervised learning specifically applied to CBCT scans. In our work, instead of dividing the literature based on the deep learning methods, we group the research based on the type of artifacts, discussing results employing projection- and/or volume-

domain optimization, dividing the methods based on the type of supervision, and also including research addressing time-resolved 4D CBCT reconstruction.

Artifacts in CBCT images can principally be reduced by optimizations in the projection, volume, or dual-domain (both projections and volumes), as well as by DL-enabled reconstruction (see Figure 1). This survey presents an overview of deep learning techniques able to reduce artifacts in 3D as well as time-resolved 4D CBCT using optimizations in the above domains, and through novel CBCT reconstruction methods. Furthermore, it addresses the challenges and limitations associated with these approaches and provides recommendations for future research directions.

This survey organizes the literature according to the type of artifacts which is addressed, and presents and contrasts the methodologies used within each specific artifact group (see Figure 2). The remainder of this paper is organized as follows: Section II briefly summarizes the basic aspects of CBCT acquisition and the assessment of scan quality. Thereafter, the literature is discussed based on different types of artifacts (as outlined in [8], [9]) as follows: Section III presents methods attempting to improve CBCT image quality by reducing artifacts generated because of the cone-beam geometry and by bringing the CBCT quality closer to the one of CT scans. The subsequent sections focus on various methods to address artifacts resulting from reduced acquisition dose. Firstly, Section IV discusses techniques that lower the dose per X-ray projection to achieve dose reduction. This is followed by Section V, which explains methods for artifact reduction when acquiring fewer projections by uniformly dropping some of them (sparse-view reconstruction). Section VI explores artifact reduction methods specifically for CBCT scans acquired from a limited angular range. The paper then proceeds to discuss methods targeting scatter and beam hardening artifacts in Section VII. Section VIII is dedicated

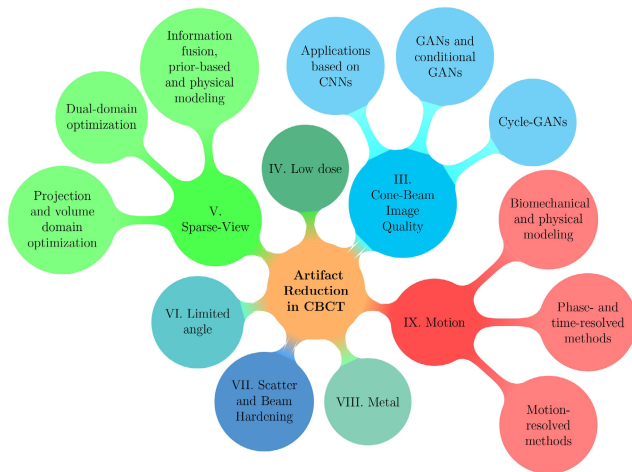


FIGURE 2. Visualisation of the content of this survey and the literature covered.

to research on reducing metal artifacts. Section IX focuses on motion compensation techniques for 3D and 4D CBCT. The main trends in the recent literature on using deep learning-based architectures for CBCT artifact mitigation are presented in Section X, complemented with a discussion concerning the connections amongst the methods used for various types of artifacts and recommendations for future work. Finally, the paper concludes with Section XI.

II. PRELIMINARIES

This section briefly reviews the basics of CBCT reconstruction and evaluation methods employed in artifact reduction and scan quality assessment.

A. CONE-BEAM GEOMETRY RECONSTRUCTION AND DEEP LEARNING

CBCT scans are acquired by means of an imaging system consisting of an X-ray source and a flat-panel (2D) detector mounted on a gantry system which rotates around the body region of interest. Several hundred 2D X-ray images are acquired at various angles. These projections can be acquired from a limited angular range (so-called short scan) or a full 360° trajectory (full scan). Following the acquisition, a volumetric 3D image is reconstructed from the 2D projection images. Several methods exist to solve this ill-posed inverse problem. The most popular one is based on an analytic method developed by Feldkamp, Davis, and Kress (FDK [10]) which provides a fast and reliable approximation of the inverse Radon transform. Alternatively, iterative algebraic reconstruction techniques (ART [11]) have become popular as well. Moreover, by tracking the patients' motion, e.g. by capturing an external or internal breathing signal, and dividing the projections based on the motion state, it is possible to reconstruct 4D (motion-resolved) volumetric images. 4D scans include both the 3D volumetric information as well as their temporal dynamics.

In a nutshell, deep learning based approaches can be deployed at various stages of the CBCT reconstruction

process. Firstly, deep neural networks can be trained to correct the acquired 2D projections (projection domain correction); secondly, they can be used to correct the reconstructed CBCT volumetric images (volume domain correction); and thirdly, the two approaches can be combined into a dual-domain correction. Another approach is to augment or replace (parts of) the 2D-3D CBCT reconstruction itself with deep learning based components. The components of the FDK algorithm were mapped into a deep neural network by means of a novel deep learning enabled cone beam back-projection layer [12], [13]. The backward pass of the layer is computed as a forward projection operation. This approach thus permits joint optimization of correction steps in both volume and projection domain. An open source implementation of differentiable reconstruction functions is available [14]. The networks are often trained in a supervised fashion by comparing reconstructed CBCT images with an artifact-free ground truth. Unsupervised [15], [16] and self-supervised [17], [18] learning approaches have been employed as well.

While datasets of 3D or 4D CBCT scans obtained from phantoms, animals or human subjects are available for training, they generally lack ground truth information required for deep learning based artifact mitigation employing supervised learning. To overcome this, artificial or simulated CBCT data is often used, obtained e.g. by means of forward projecting existing CT scans in a CBCT setup and manual incorporation of artifacts. For example, motion artifacts can be included by sampling CBCT projections at scan angles and time steps matching interpolated phases of a given 4D CT scan.

The general acquisition and reconstruction process of CBCT scans, including deep learning based corrections, is summarized in the visual abstract in Figure 1.

B. EVALUATION METRICS

Several metrics have been utilized in the literature to evaluate the quality of CBCT scans enhanced by deep learning-based techniques. The main qualitative evaluation metrics, computed between a reconstructed volume (with artifacts) and the ground truth reference, can be divided into two main groups as follows, according to [7]:

- **Image Similarity Metrics:** These metrics compute the similarity between scans and include (mean) absolute error (ME and MAE), (root) mean squared error (MSE and RMSE), (peak) signal-to-noise ratio (SNR and PSNR), structural similarity (SSIM) [19], and Dice coefficient [20].
- **Dosimetric Similarity Metrics:** These metrics measure the consistency in dosimetry using a pair of scans, such as dose difference pass rate (DPR); dose-volume histogram (DVH), and gamma pass rate (GPR).

In addition to the metrics mentioned above, metal artifact index (MAI [21]), and streak index (SI [22]) have been used in the literature to measure the level of specific artifacts in CT and CBCT scans. For motion, visual information

fidelity (VIF) [23] or autofocus (sharpness) metrics have been employed, among others.

C. CLINICAL EVALUATION

The numerical evaluation metrics mentioned above compute the similarity of the improved CBCT compared with a reference, or report the level of the presence of artifacts, scan sharpness, or other quality criteria. Ideally, these metrics should reflect the scan quality; hence, they should correspond to the preference of the experts in using the scans in clinical routine. However, it is essential to note likely inconsistencies between simulated (where ground truth references exist) and real-world clinical data, so clinical evaluations are necessary to ensure the applicability of the presented methods for practical applications. A clinical evaluation can be conducted by completing surveys with experts such as medical doctors or radiation physicists to directly assess the level of artifacts and the performance of the artifact reduction techniques, and the applicability of the improved images in various clinical tasks such as dose calculation, soft-tissue segmentation, and patient positioning [24].

III. CONE-BEAM IMAGE QUALITY

Cone-beam geometry and the size of the flat-panel detector result in the coverage of larger body areas but at lower resolution and degradation in scan quality compared to fan-beam CT scan acquisition. Consequently, significant attention and extensive research has been directed at improving the quality of CBCT scans, often referred to as removing cone-beam or geometry artifacts in the literature. One of the initial approaches to enhance CBCT quality involves employing supervised learning and training a 39-layer deep convolutional neural network (CNN) to map input CBCT scans to the corresponding planning CT as ground truth (reference) volumes [25]. This mapping of CBCT images to match corresponding CT images is often called synthetic CT (sCT) from CBCT.

A. APPLICATIONS BASED ON CNNs

Researchers have explored several CNN-based architectures with various supervised training objectives to enhance CBCT quality. For instance, denoising has been targeted through solving the multi-agent consensus equilibrium (MACE) problem and multi-slice information fusion techniques [26]. CNN models have demonstrated the ability to reduce ring artifacts from flat-panel CBCT scans using pre-corrected and artifact-free scans as ground truth [27]. Geometric artifacts caused by misalignment of the CBCT system were reduced using a modified fully convolutional neural network (M-FCNN), without using any pooling layers [28]. A further approach used a 3D block-based residual encoder-decoder convolutional neural network (RED-CNN) architecture coupled with a bilateral 3D filter and a 2D-based Landweber iteration to successfully remove Poisson noise while preserving the

image structure at tissue edges [29]. Training 3D models using a multi-task learning objective improved the quality of CBCTs by producing high-quality synthetic CT (sCT) scans from noisy and artifact-ridden scans for segmenting organs-at-risk (OARs) [30]. Lately, using InceptionV3 [31] as a backbone has proven beneficial in reducing the artifacts observed in CBCT short scans due to the misalignment of the detection plane around the z-axis [32].

B. GANS AND CONDITIONAL GANS

Researchers have used self-supervised and unsupervised techniques to eliminate the need for paired CBCT and CT scans in supervised learning and to consider anatomical changes between the acquisition of planning CT (pCT) and CBCT. These techniques mainly involve training auto-encoders, (conditional) generative adversarial networks (GANs [33]), and cycle-consistent generative adversarial networks (Cycle-GANs [34]). Combining auto-encoders and GANs as a complementary approach to reweighting in analytical and iterative reconstruction methods has improved the quality of CBCT scans [35]. Training conditional GANs has shown promising results in enhancing the quality of CBCT through style transfer, effectively removing artifacts and discrepancies between CBCT and pCT for average tumor localization [36] and adaptive therapy [37]. Moreover, a more advanced GAN variant called temporal coherent generative adversarial network (TecoGAN) also improves the quality of simulated 4D CBCT scans by considering the time dependencies and motion for quality enhancement [38], [39].

C. CYCLE-GANS

Using Cycle-GANs for unpaired translation from CBCT to pCT has received significant attention among researchers. Notably, Cycle-GANs have successfully generated high-quality synthetic CT scans from CBCT for various organs, including prostate [40], lung [41], and abdominal scans [42]. A novel architecture inspired by contrastive unpaired translation (CUT [43]), trained in an unsupervised manner, improves the quality of CBCT scans by addressing fringe artifacts and noise degradation for dose calculation in adaptive radiotherapy [15]. The combination of binary cross-entropy, gradient difference, and identity losses with Cycle-GANs has further improved the quality of head and neck CBCT scans [44]. Introducing the residual block concept in the implementation of Res-Cycle-GAN has demonstrated advancements in the quality of sCT scans [45]. Moreover, researchers have explored the combination of a Cycle-GAN with classical image processing techniques [46] and U-Net [47] architectures [16] in two-step approaches. These approaches aim to initially reduce artifacts and subsequently generate sCT scans to improve the quality. Ultimately, researchers demonstrated that trained Cycle-GANs enhance the quality of CBCT scans and achieve high accuracy in volumetric-modulated arc photon therapy (VMAT) [48].

D. ALTERNATIVE METHODS

In addition to adopting mainstream trends and computer vision architectures for artifact reduction in CBCT scans, researchers have explored creative methods specifically tailored to CBCT reconstruction using deep learning and neural networks. For instance, U-Nets have been optimized for spectral blending of independently reconstructed sagittal and coronal views to enhance the CBCT quality [49]. Neural networks have also been integrated into the core of the reconstruction algorithms in the Feldkamp, Davis and Kress (FDK) technique to introduce the NN-FDK technique for CBCT quality improvement [50]. Another novel architecture, known as the iterative reconstruction network (AirNet), incorporates several variants in selecting projections based on random-phase (RP), prior-guided (PG), and all-phases (AP) for reconstruction [51]. Geometry-guided deep learning (GDL [52]), and its multi-beamlet-based approach (GMDL [53]) are additional examples of leveraging deep learning to enhance the reconstruction geometry effectively. Finally, CNNs have been employed to predict the quality of the scans and accordingly dynamically adapt the C-arm source trajectory in the imaging acquisition process to avoid generating artifacts in the final scans [54].

IV. LOW DOSE

The reduction of the acquisition dose in CBCT scans, which leads to the increased presence of artifacts, has been addressed through various approaches such as adjusting the radiation dose per X-ray projection [55], increasing the acquisition speed or collecting fewer projections [56]. Early research focused on low-dose artifact reduction primarily by removing artifacts in the volume domain using deep CNNs with U-Net architectures. The studies demonstrated the potential of decreasing the overall radiation dose through both dose reduction methods mentioned above [55], [56]. Moreover, a combination of 2D and 3D concatenating convolutional encoder-decoder (CCE-3D) with a structural sensitive loss (SSL) was employed to denoise low-dose CBCT scans and remove artifacts in both projection and volume domains. This approach showed promising results in improving the quality of CBCT scans based on several metrics, such as PSNR and SSIM, and with greater improvements reported in the projection domain compared with the volume domain [57]. In addition, a CNN-based iterative reconstruction framework was integrated with a plug-and-play proximal gradient descent framework to leverage DL-based denoising algorithms and enhance CBCT reconstruction [56]. Training models inspired by self-supervised learning approaches for inpainting and denoising Poisson and Gaussian noise have shown promising results in removing low-dose artifacts [58]. Similarly, models optimized for removing Gaussian noise and addressing view aliasing artifacts through 2D iterations with 3D kernels have been developed [59]. Furthermore, researchers combined a non-subsampled contourlet transform (NSCT) and a Sobel filter with U-Net architectures,

referred to as NCS-Unet, to improve the quality of low-dose CBCT scans by enhancing both low- and high-frequency components [60].

V. SPARSE-VIEW

This section summarizes research aiming at reducing artifacts in CBCT reconstruction occurring from using uniformly downsampled full-scan (360°) projections, primarily with the goal of dose reduction. Sparse-view artifact reduction is closely related to mitigation of artifacts caused by limited angle acquisition and breathing-phase-correlated 4D reconstruction, which will be reviewed in the upcoming sections VI and IX, respectively. While the underlying motivations for sparse-view (acquisition dose reduction), limited angle (geometric constraints), and 4D (time resolved imaging) acquisition are different, in all cases artifacts are created due to the lack of projections from various angles. Decreasing the number of projections and the resulting data insufficiency for the reconstruction algorithm results in artifacts appearing in the shape of symmetric and uniform streaks, as depicted in Figure 3.

A. PROJECTION AND VOLUME DOMAIN OPTIMIZATION

The body of literature on sparse-view artifact reduction using deep learning has been consistently growing since 2019, when initial research demonstrated the opportunity to reproduce the original image quality with using as few as one-seventh of the projections with symmetric CNN's as post-processing operation in the volume domain [61]. Similarly, using a multi-scale residual dense network (MS-RDN) successfully improved the quality of CBCTs reconstructed from one-third of the projections [62]. In addition to training in the volume domain, the intensities of under-sampled projections can be corrected using deformation vector fields (DVF) to match the original data, resulting in negligible streak artifacts after reconstruction [63]. Similarly, symmetric residual CNN's (SR-CNN) can enhance the sharpness of the edges in anatomical structures reconstructed from sparse-view projections with total variation (TV) regularization in half-fan scans [61]. Furthermore, a counter-based total variational CBCT reconstruction using a U-Net architecture enhances the smoothed edges in lung CT reconstructed scans from half-fan projections [64]. In Ref. [65], a Reconstruction-Friendly Interpolation Network (RFI-Net) is developed, which uses a 3D-2D attention network to learn inter-projection relations for synthesizing missing projections, and then introduces a novel Ramp-Filter loss to constrain a frequency consistency between the synthesized and real projections. The authors of [66] developed a dual-domain attention-guided network framework (Dual-AGNet) which works in both projection and reconstruction domains, featuring spatial attention modules and a joint loss function.

B. DUAL-DOMAIN OPTIMIZATION

Though interpolating missing data in the projections and removing artifacts in the volume domain are straightforward

approaches to sparse-view artifact reduction, combining both and backpropagating the error through the reconstruction algorithm is not trivial. Despite the complexity involved, researchers attempted to unroll the proximal gradient descent algorithm for reconstruction and backpropagate the gradient through a U-Net architecture to reduce streak artifacts in [67]. Since optimization in the volume domain and projection interpolation are regression problems with different or the same data channels as input and output, autoencoder-decoder architectures have also gained popularity for artifact reduction [68]. To avoid complications regarding backpropagation through the reconstruction (back-projection) algorithm, DEER is introduced as an efficient end-to-end model for directly reconstructing CBCT scans from few-view projections [69]. Furthermore, DeepOrganNet could fine-tune the lung mesh by skipping the reconstruction step and avoiding sparse-view artifacts appearing on organ mesh [70]. Furthermore, the recent deep intensity field network (DIF-Net) model uses the latent representation (feature maps) of the 2D projections coupled with a view-specific query for extracting information from the projections. This information is then fed through cross-view fusion and intensity regression models to reconstruct a volume without artifacts [71].

C. INFORMATION FUSION, PRIOR-BASED AND PHYSICAL MODELING

Recent research trends seek to minimize sparse-view artifacts by incorporating multi-slice [72] and scale [73] information fusion techniques, as well as combining information from different scan views (coronal, axial, and sagittal) [74]. As the computational resources have become more powerful, deep learning for sparse-view artifact reduction has extended from 2D models for single slice processing to 3D models and processing of 4D CBCT scans [72]. The use of prior (planing) CT and CBCT volumes to enhance the trained models, such as regularized iterative optimization reconstruction (PRIOR-Net [75]) and merge-encoder CNN (MeCNN [73]) have recently become popular for sparse-view artifact reduction. Researchers have also investigated using perception-aware [76] and physics-based [75] methods. The learning paradigm has expanded beyond purely supervised learning to different tasks, such as denoising (DRUNet [77]), artifact reduction [78], self-supervised by dropping projections [18] and unsupervised learning through training conditional and generative adversarial networks (GANs) [79].

VI. LIMITED ANGLE

Besides lowering the imaging dose through uniformly downsampled projections, another approach to reducing the number of acquired projections and scanning dose is scanning the body from a limited angle. Such scan settings are especially common when using a full-fan acquisition technique in a short-scan, where reconstruction is performed using projections from an angular range covering less than 360 degrees. Although Parker weights [80] can be utilized

to compensate for the loss of mass in the resulting CBCT scans, artifacts still appear due to the smaller number of acquired projections when scans are acquired from limited angles. One of the initial attempts used learnable Parker weights in the projection domain to address the mass loss in the angular range from $180^\circ + \theta$ to 360° (θ being the fan angle) [12]. A subsequent study optimized a deep artifact correction model (DAC) using a 3D-ResUNet architecture to create high-quality scans and improve artifacts in limited-angle circular tomosynthesis (cTS), confirming the potential for quality enhancement in the volume domain [81]. Further research demonstrated that combining FDK-based reconstruction with a neural network can achieve outstanding performance in 3D CBCT reconstruction from projections acquired from only 145° [82].

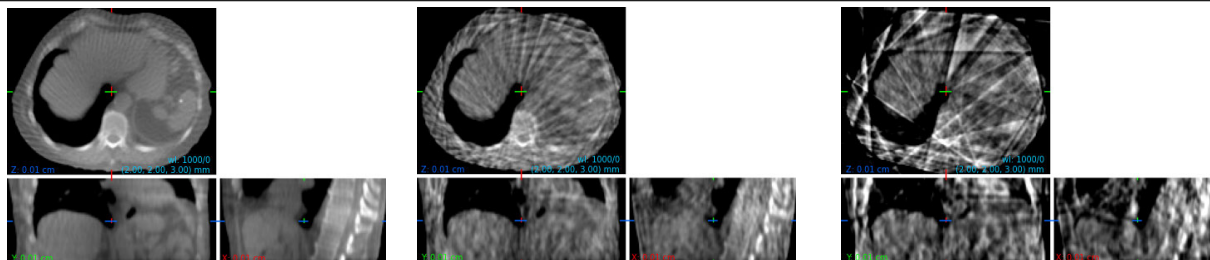
Supervised learning, frequently implemented through training U-Net architectures, for shading corrections in CBCT volumes with a narrow field of view (FOV) notably improved the quality of reconstructed CBCT scans, using CT scans as ground truth [83]. Another approach involves using a prior based on a fully sampled CT or CBCT and training a 2D3D-RegNet, which demonstrates the effectiveness of using a patient-specific prior for limited-angle sparseness artifact reduction [84]. A conventional method for 4D CBCT reconstruction is dividing the projections based on the breathing phases and then reconstructing the body volume in those phases. As a result of using only a subset of the projections for each motion state, sparseness artifacts are prevalent for this special case of limited angle acquisition. These artifacts have been addressed in the projection domain by interpolating the projections from different breathing phases [85]. In the volume domain, transfer learning, layer freezing, and fine-tuning have been employed to adapt the trained DL models to individual patients and mitigate sparseness artifacts [86].

VII. SCATTER AND BEAM HARDENING

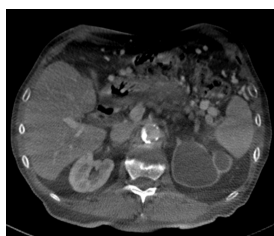
Large cone angles within the CBCT geometry setup have been observed to contribute to scatter artifacts, which have been addressed in the projection domain by leveraging Monte Carlo photon transport simulations to compute ground truth projections for supervised learning [89]. A CNN-based deep scatter estimation (DSE [89]) architecture, as well as a scatter correction network (ScatterNet [87]) are the results of research endeavors using supervised learning for artifact correction in the projection domain. The DSE model has demonstrated the potential to accurately emulate scatter artifacts and reduce the computational burden of using Monte-Carlo simulations while being orders of magnitude faster [90]. ScatterNet is considerably faster than the classical methods and might allow for on-the-fly shading correction [87]. ScatterNet, in combination with shading correction, also showed satisfactory results for dose calculation using volumetric modulated arc radiation therapy (VMAT), but yielded unsatisfactory outcomes for



Simulated 4D CBCT scan at three distinct motion phases, without significant motion artifacts



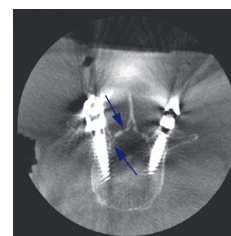
Sparse-view artifacts at various sub-sampling rates (from left to right: 1/6, 1/18 and 1/48)



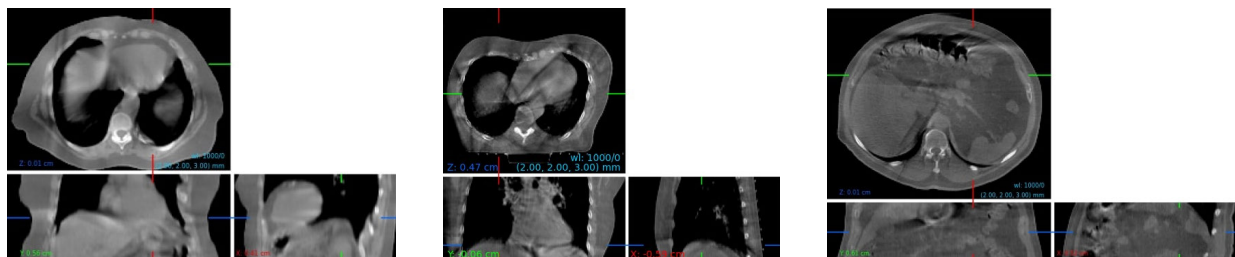
Limited angle artifacts [12]



Scatter artifacts [87]



Metal artifacts [88]



Motion artifacts in simulated (left) and real (middle and right) CBCT scans [24]

FIGURE 3. Examples of different kinds of artifacts appearing in CBCT scans. Shown are several artifact-free motion states obtained with a simulated 4D CBCT scan (1st row), sparse-view artifacts at various sub-sampling rates (2nd row), limited-angle, scatter and metal artifacts (3rd row), as well as motion artifacts (4th row).

intensity-modulated proton therapy (IMPT). Despite the abundant research work on scatter artifact corrections, studies tackling beam hardening are scarce. One such study involved training a U-Net-based architecture to predict monoenergetic X-ray projections from polyenergetic X-ray projections using supervised learning on Monte Carlo simulation-based ground truth in the projection domain [91].

Compared with the classical fast adaptive scatter kernel superposition (fASKS) scatter reduction technique [92], a U-Net-based architecture outperformed in scatter artifact reduction for both full-fan and half-fan scans based on several metrics [93]. Additionally, a U-Net-based model trained on

simulated CBCT projections has shown comparable performance to a validated empirical scatter correction technique in dose calculation for correcting the scatter artifacts in head and neck scans, computing the corrected volumes in less than 5 seconds [94]. Besides classical approaches of scatter artifact reduction, CT scans have been used as ground truth volumes for training a modified U-Net architecture with a multi-objective loss function specifically targeting scatter artifact reduction in esophagus scans [95].

Apart from supervised learning methods, researchers have also trained Cycle-GAN models to improve the quality of CBCT scans, remove scatter artifacts, and generate

sCT. In particular, Cycle-GAN has demonstrated superior performance compared to similar techniques using deep convolutional generative adversarial networks (DCGAN [96]) and progressive growing GANs (PGGAN [97]) [98].

VIII. METAL

Metal objects and implants in the patient's body result in scattered radiation reaching the detector, leading to streak artifacts. In the early research addressing metal artifacts, a CNN-based regression model has been trained to predict the detectability rank of metal implants to recommend out-of-plane angulation for C-arm source trajectories [99]. Further research in this area has proposed predicting the X-ray spectral shift after the localization of metal objects to define the optimal C-arm source-detector orbit [100]. The metal artifact avoidance (MAA) technique uses low-dose scout projections to roughly localize metal objects for the identification of a circular or non-circular orbit of C-arm source-detector to minimize variations in spectral shift and avoid metal artifacts [101].

Researchers have also employed supervised learning for reducing metal artifacts and estimating the deviation of the voxel values after inserting neuroelectrodes [102]. Self-supervised learning approaches, focused on training models for inpainting the regions affected by metal artifacts, have demonstrated improvements in simultaneously tackling metal artifact reduction while preserving the essential anatomical structures near the inserted implants [88]. In addition to supervised and self-supervised techniques, various types of GANs have been employed in the literature for metal artifact reduction. Optimized conventional GANs can reduce metal artifacts in high-resolution and physically realistic CT scans, with good generalization to clinical CBCT imaging technologies for inner-ear scans [103]. Conditional GANs, inspired by the pix2pix-GAN [104], have successfully reduced metal artifacts in spine CBCT scans, enabling precise recovery of fiducial markers located outside the C-arm's field-of-view (FOV) [105]. A Cycle-GAN has also been employed to efficiently reduce metal artifacts by generating synthetic CT (sCT) from Megavolt CBCT (MVCBCT) and improving the quality of CBCT scans [106].

IX. MOTION

Many of the state-of-the-art volumetric reconstruction techniques for CBCT rely heavily on the initial assumption that the projections are acquired from a stationary object. However, this assumption is often violated because of periodic respiratory and cardiac motions or non-voluntary and non-periodic movement of air bubbles in the abdominal area. When reconstructing CBCT volumes using projections acquired from various body states under motion, motion streak artifacts appear in the reconstructed volume, as shown in Figure 3. The severity of the resulting artifacts is positively correlated with the intensity of motion. The most common approach to tackle motion artifacts in CBCT scans is dividing the projections based on the motion

state (motion-resolved [107], [108], [109], [110], [111], [112]), periodic motion state (phase-resolved [111], [113], [114]) or acquisition time (time-resolved [115], [116]), and then reconstruct multiple volumes based on each batch of projections to generate a 4D CBCT.

A. MOTION-RESOLVED METHODS

A novel approach using CNNs to predict the missing projections in motion-resolved 4D-CBCT combined with a bin-sharing technique to accelerate the acquisition process, substantially removed streak artifacts compared with standard conjugate gradient reconstruction [107]. Training a residual U-Net also reduces the streak artifacts appearing in 4D-CBCT by addressing the sparseness of the projections acquired in each breathing phase [108]. Residual dense networks (RDNs [110]) have successfully improved sparseness artifacts using an in-house lung and liver dataset, as well as a public dataset of the SPARE challenge [117], [118]. Similar research demonstrates that combining the information of the different breathing phases to train a prior-guided CNN can effectively reduce artifacts in motion-resolved 4D-CBCT scans [109]. In addition to training single models, researchers attempted to optimize a cascade of spatial and temporal CNN models to combine spatial and temporal information for maximum artifact removal and to avoid errors in the tomographic information [112]. A dual-encoder CNN (DeCNN) architecture simultaneously processes and combines the information of 4D motion-resolved volumes and the averaged volume, thereby improving the sharpness of the edges in moving and fixed tissues in 4D-CBCT [119].

B. PHASE- AND TIME-RESOLVED METHODS

Phase-resolved CBCT is a specific case of motion-resolved CBCT, where projections are selected based on the different phases of body volume under periodic, respiratory, or cardiac motion. Motion compensation learning-induced sparse tensor constraint reconstruction (MCL-STCR) was shown to improve 4D-CBCT scans for all motion phases [120]. 3D-CNNs have shown to effectively mitigate sparse-view artifacts in motion-compensated 4D-CBCT scans reconstructed using FDK, thereby enhancing the overall quality [114]. N-Net uses the prior volume reconstructed using all projections to remove streak artifacts. CycN-Net combines the temporal correlation among the phase-resolved scans to reduce streak artifacts that are caused by sparse-view sampled motion-resolved projections [111]. Furthermore, training a patient-specific GAN-based model on phase-resolved 4D-CBCT to reproduce CT quality using CBCT scans demonstrates improvements when applied to test set projections acquired from the same patient [113]. In addition to motion- and phase-resolved methods, training a U-Net can remove sparseness artifacts from time-resolved 4D-CBCT without requiring any prior information [115]. GANs have also demonstrated the capacity of estimating sCT scans from time-resolved 4D-CBCT and the average 3D-CBCT volume, resulting in

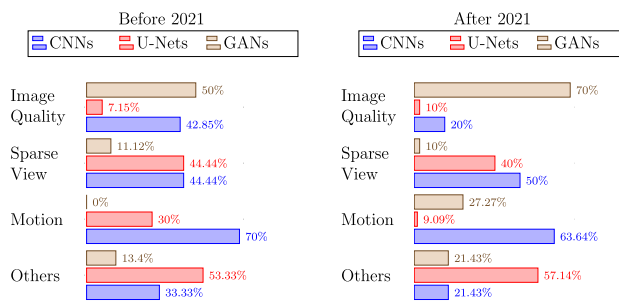
a comparable improvement in dose calculation using both strategies [116].

C. BIOMECHANICAL AND PHYSICAL MODELING

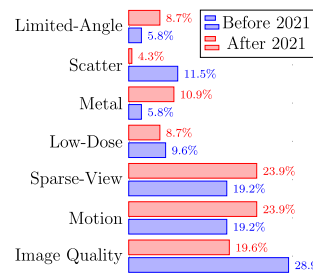
In addition to phase-, motion-, and time-resolved techniques, researchers have also explored targeting motion artifacts by physically modeling the motion using a deformation-vector-field (DVF) and by optimizing an autofocus metric (i.e., maximizing some measure of sharpness). The simultaneous motion estimation and image reconstruction (SMEIR) model, as well as its biomechanical modeling-guided version (SMEIR-Bio), are examples of models developed for motion effect prediction in lung 4D CBCT scans [121]. These models have also been enhanced using a U-Net-based DVF optimization technique, leveraging a population-based deep learning scheme to improve the accuracy of intra-lung DVF prediction (SMEIR-Unet) in the same research work. By incorporating the reference phase in 4D CBCT as an extra channel to their model, training a 4D U-Net for motion estimation, with fine-tuning the estimated DVFs, the performance of SMEIR models increases for motion artifact reduction [122]. CNN-based architectures have been optimized to estimate deformable motion and predict the motion intensity on 8×8 grids covering the axial slice, followed by a preconditioning technique to favor more likely motion intensities [123]. CNNs have also been trained for motion compensation in CBCT scans to solve the high-dimensional and no-convex problem of optimizing the autofocus metric [124].

D. ALTERNATIVE METHODS

The autofocus metric has also been replaced with the context-aware deep learning-based visual information fidelity (CADL-VIF) image similarity metric to optimize multi-resolution CNNs [125]. This approach aims to improve motion degradation and compute sharp scans while preserving the tissue structures by optimizing visual information fidelity (VIF) without requiring motion-free ground truth. An alternative to the autofocus metric is using contrastive loss to train GAN architectures to enhance the quality of 4D-CBCT scans and to reduce streak and motion artifacts [15]. To address the slow speed of reconstruction and to compensate for the errors of 4D-CBCT due to the severe intraphase undersampling, a feature-compensated deformable convolutional network (FeaCoDCN [126]) model has been proposed. It achieves nearly real-time reconstruction and accurate CBCT, outperforming the previous method applied to the SPARE Challenge [117], [118]. Besides the numerous research studies addressing motion in 4D CBCT, which requires recording the patient’s breathing curve, researchers have also simulated motion in CBCT scans based on the estimation of DVFs according to 4D CT ground truth scans [127]. They subsequently trained a dual-domain model to mitigate 3D CBCT motion artifacts in the projection and volume domains. The clinical validation



(a) Distribution based on model architecture.



(b) Distribution based on artifact type.

FIGURE 4. A visual summary of the distribution of the covered research literature in CBCT artifact mitigation using deep learning, separately for two time periods, (a) based on three generic deep learning architecture categories given a broad categorization by artifact type, and (b) based on the distribution according to the type of artifact.

on real-world CBCT images yielded positive feedback from clinical experts, demonstrating the effectiveness of their approach for motion compensation [24]. In addition to all methods to reduce motion artifacts, researchers have successfully used an artifact-driven slice sampling technique to avoid artifacts caused by moving air bubbles in the segmentation of the female pelvis [128].

X. DISCUSSION AND RECOMMENDATIONS

The previous sections have outlined the methodology and the complete workflow employed for deep learning based mitigation of artifacts in CBCT scans, addressing each specific type of artifact separately. This section presents a summary, emphasizing the central role of various deep learning approaches. The objective is to offer a comprehensive review of the architectures employed for different artifact types, highlighting both the promising aspects and the limitations in the current literature.

In general, a trend is observed in shifting from conventional supervised learning with CNNs and U-Net-type architectures to exploring more modern learning paradigms such as GANs, and investigating self-supervised and unsupervised methods, leveraging e.g. Cycle-GANs, as depicted in Figure 4. In particular, Cycle-GAN-based architectures offer the appealing feature of enabling model training without needing paired labeled data [131]. However, they come with high data requirements, rising attention toward methods and projects for data collection, synthetic data generation, dataset merging from diverse sources, and data homogenization. This trend suggests the rise of research works attempting

TABLE 1. Summary of a subset of studies per artifact type, based on recency and number of citations. The table provides details about artifact category, publication year, study title, anatomic site, model type, number of patients, GPU hardware, and whether the code was published.

Artifact type	Year	Title	Anatomic site	Model	Patients	GPU Hardware	Published code?
image quality	2019	Paired cycle-GAN-based image correction for quantitative cone-beam computed tomography [45]	brain, pelvis	cycle GAN	44	NVIDIA TITAN XP	-
	2019	CBCT correction using a cycle-consistent generative adversarial network and unpaired training to enable photon and proton dose calculation [48]	pelvis	cycle GAN	33	NVIDIA Tesla P100	-
low-dose	2019	Computationally efficient deep neural network for computed tomography image reconstruction [67]	abdomen	U-Net	10	NVIDIA GTX 1080 Ti	-
	2020	Neural networks-based regularization for large-scale medical image reconstruction [55]	cardiac	U-Net	19	-	-
sparse-view	2023	Sub-volume-based Denoising Diffusion Probabilistic Model for Cone-beam CT Reconstruction from Incomplete Data [129]	breast	diffusion model	-	128x NVIDIA Tesla V100	-
	2023	Learning Deep Intensity Field for Extremely Sparse-View CBCT Reconstruction [71]	knee	learned reconstruction	-	NVIDIA RTX 3090	yes
	2020	Self-contained deep learning-based boosting of 4D cone-beam CT reconstruction [110]	liver, lung	residual dense network	20	NVIDIA GeForce RTX 2080 Ti	yes
	2020	Deep Efficient End-to-End Reconstruction (DEER) Network for Few-View Breast CT Image Reconstruction [69]	breast	GAN	42	NVIDIA Titan RTX	yes
limited-angle	2020	C-arm orbits for metal artifact avoidance (MAA) in cone-beam CT [101]	chest phantom	U-Net	0	NVIDIA TITAN X	-
scatter	2019	Real-time scatter estimation for medical CT using the deep scatter estimation: Method and robustness analysis with respect to different anatomies, dose levels, tube voltages, and data truncation [90]	head, thorax, pelvis	U-Net	21	NVIDIA Quadro P6000	-
metal	2021	Inner-ear augmented metal artifact reduction with simulation-based 3D generative adversarial networks [130]	temporal bone images	GAN	597	11 GB GPU	-
motion	2022	Enhancement of 4-D Cone-Beam Computed Tomography (4D-CBCT) Using a Dual-Encoder Convolutional Neural Network (DeCNN) [119]	lung	CNNs	26	NVIDIA Titan RTX	-
	2022	Deep learning-based motion compensation for four-dimensional cone-beam computed tomography (4D-CBCT) reconstruction [114]	thorax	CNNs	18	NVIDIA Tesla V100S	yes

at the adaptation of generative models including GANs, Cycle-GANs, as well as scored-based models [132], [133], in upcoming research endeavors. A recent example [129], which employs denoising diffusion probabilistic models [134], [135] for sparse-view CBCT reconstruction, demonstrates a lot of potential for future research, however at the expense of tremendous compute resources (up to 128 GPUs, see also Table 1). On the other hand, less computationally intense, U-Net-based, architectures have demonstrated their merit in successfully addressing artifacts across all categories, making them a highly recommended and robust baseline approach for artifact mitigation.

In the context of this survey, the primary DL-based architectures used in the literature can be divided into four key categories: CNNs, U-Nets, GANs, and cycle-GANs. Here, we categorize architectures with multi-scale information fusion, i.e. including connections from the network's input (encoding) layers to output (decoding) layers (such as [67]) under the category U-Net, while those without such direct connections (such as autoencoders [136]) are categorized as CNNs. DL-based models generally require medium to large datasets for training, validation and testing through clinical evaluation. While medium-sized datasets, including multiple

patients, can serve as starting points for training CNNs and U-Nets [83], GANs perform better using datasets containing at least dozens of patient scans [42]. This trend generalizes to 3D and 4D reconstruction, where larger input sizes and a higher number of scans become essential, in particular for 4D [122]. A review of the studies presented in Table 1 reveals that the majority of research was conducted with fewer than 50 patients. This relatively small number of patients can pose challenges for validating the approach across a diverse population. Consequently, the robustness of these models warrants further scrutiny to ensure their ability to generalize well across various human anatomies.

CNN architectures, known for their stable convergence and versatility, demonstrate a wide range of applications for artifact reduction through adapting different vision backbones [32] and incorporating diverse architectural components such as attention blocks [24]. However, in terms of multi-scale information fusion, they are inferior to U-Nets and their variants (e.g., U-Net++ [137]), which demonstrate a fast convergence in supervised learning due to the internal architectural connections between different layers enhancing the multi-resolution information fusion [7]. Since CNNs and U-Nets are predominantly being trained in

a supervised manner, their learning technique necessitates explicitly labeled data to define the task. On the other hand, generative models (GANs), incorporating an adversarial loss, also offer potential applications in generating high-quality synthetic scans to meet the data needs of the deep learning-based architectures [36]. Moreover, Cycle-GANs compute the inverse path of artifact reduction automatically, using a cycle-consistent loss, thus being able to learn artifact reduction without the need for paired artifact-free ground truth [48].

Only four of the papers presented in Table 1 provide a public code repository to reproduce their results. This highlights a considerable shortage of open science practices, such as sharing code, to promote transparency and reproducibility in research. It is strongly recommended for researchers to share their code publicly to enhance the credibility and reproducibility of their work and accelerate scientific progress in this field.

XI. CONCLUSION

We presented a survey on the application of deep learning and convolutional neural networks to reduce various types of artifacts in CBCT scans. We categorized the existing literature based on the type of artifacts they address as well as the methodology employed. Figure 4b illustrates the amount of the recent research works based on the type of artifacts. It is observed that there has been considerable growth in artifact reduction research compared with focusing more generically on scan quality after 2021. The opportunity of reducing the imaging dose with the help of compensating for artifacts when using low-dose scans, sparse-view, and limited-angle acquisition techniques have gained substantial attention due to the ease of simulation and computing the ground truth, especially for sparse-view and limited-angle approaches.

However, metal and scatter artifacts have received less attention. This may also be due to the challenges involved in computing the ground truth for metal artifacts, or the high computational cost of Monte-Carlo simulation for scatter artifacts. We expect that the research community could profit from open-source accurate and fast artifact simulations for training models (as before with XCAT [138]). The development of such simulations could also serve as a driving force for physics-based artifact modeling or training physics-informed neural networks (PINN) [139] for artifact reduction. These simulations would benefit from GPU implementations for data generation to enable on-the-fly integration into the training pipelines with neural networks. In addition to simulations, there is a research gap for open-source data augmentation techniques, such as [140] and [141], also based on incorporating simulated artifacts into real datasets.

In addition to simulation and augmentation tools for modelling, the research community would benefit from the availability of open-source datasets. Researchers are still reporting results on phantoms and cadavers, indicating a need for more diverse and realistic publicly available

datasets. Nevertheless, despite the lack of open-source 4D CBCT datasets with raw projections and breathing curves, there is an increase of motion artifact reduction research in recent literature. The collection and sharing of up-to-date benchmark datasets on a large scale, similar to the SPARSE [117], [118] and SynthRAD [142] challenges, would enhance the quality of many research works and provide the opportunity for fair and accurate comparison of different approaches. Furthermore, many studies suffer from a lack of clinical evaluation. The availability of open-source standard clinical evaluation platforms would be of significant help in addressing this issue.

In terms of methodology, there has been a noticeable trend of moving beyond supervised learning towards self-supervised, unsupervised, and domain adaptation methods in recent years. Researchers have started incorporating more physically inspired ideas into the neural networks and utilizing prior patient knowledge to personalize the models for specific anatomies. One of the drawbacks often observed in the current literature is the absence of ablation studies. For example, in the case of approaches employing dual-domain optimization in both projection and volume domains, the performance gained in each domain should be estimated separately. Besides artifact reduction after the CBCT acquisition, adapting the acquisition process itself using neural networks, such as C-arm trajectory adjustments applied to metal artifact reduction, present a further exciting avenue for future research.

In summary, substantial progress has been made in recent years transferring state-of-the-art methods from deep learning based computer vision to the domain of CBCT imaging and in particular the amelioration of prevalent imaging artifacts, with a clear potential to improve diagnosis and treatment in clinical practice.

REFERENCES

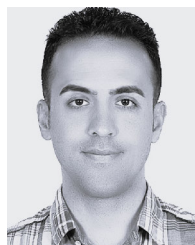
- [1] R. Shende, G. Gupta, G. Patel, and S. Kumar, "Commissioning of TrueBeam™ medical linear accelerator: Quantitative and qualitative dosimetric analysis and comparison of flattening filter (FF) and flattening filter free (FFF) beam," *Int. J. Med. Phys., Clin. Eng. Radiat. Oncol.*, vol. 5, no. 1, pp. 51–69, 2016.
- [2] D. A. Jaffray, J. H. Siewerdsen, J. W. Wong, and A. A. Martinez, "Flat-panel cone-beam computed tomography for image-guided radiation therapy," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 53, no. 5, pp. 1337–1349, Aug. 2002.
- [3] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: <http://www.deeplearningbook.org>
- [4] P. Paysan, I. Peterlík, T. Roggen, L. Zhu, C. Wessels, J. Schreier, M. Buchacek, and S. Scheib, "Deep learning methods for image guidance in radiation therapy," in *Artificial Neural Networks in Pattern Recognition* (Lecture Notes in Computer Science), vol. 12294, F.-P. Schilling and T. Stadelmann, Eds. Winterthur, Switzerland: Springer, Sep. 2020, pp. 3–22, doi: [10.1007/978-3-030-58309-5_1](https://doi.org/10.1007/978-3-030-58309-5_1).
- [5] M. F. Spadea, M. Maspero, P. Zaffino, and J. Seco, "Deep learning based synthetic-CT generation in radiotherapy and PET: A review," *Med. Phys.*, vol. 48, no. 11, pp. 6537–6566, Nov. 2021.
- [6] M. Zhang, S. Gu, and Y. Shi, "The use of deep learning methods in low-dose computed tomography image reconstruction: A systematic review," *Complex Intell. Syst.*, vol. 8, no. 6, pp. 5545–5561, Dec. 2022, doi: [10.1007/s40747-022-00724-7](https://doi.org/10.1007/s40747-022-00724-7).

- [7] B. Rusanov, G. M. Hassan, M. Reynolds, M. Sabet, J. Kendrick, P. Rowshanfarzad, and M. Ebert, "Deep learning methods for enhancing cone-beam CT image quality toward adaptive radiation therapy: A systematic review," *Med. Phys.*, vol. 49, no. 9, pp. 6019–6054, Sep. 2022.
- [8] R. Schulze, U. Heil, D. Groß, D. Bruellmann, E. Dranischnikow, U. Schwanecke, and E. Schoemer, "Artefacts in CBCT: A review," *Dentomaxillofacial Radiol.*, vol. 40, no. 5, pp. 265–273, Jul. 2011.
- [9] F. E. Boas and D. Fleischmann, "CT artifacts: Causes and reduction techniques," *Imag. Med.*, vol. 4, no. 2, pp. 229–240, Apr. 2012.
- [10] L. A. Feldkamp, L. C. Davis, and J. W. Kress, "Practical cone-beam algorithm," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 1, no. 6, pp. 612–619, 1984.
- [11] R. Gordon, R. Bender, and G. T. Herman, "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography," *J. Theor. Biol.*, vol. 29, no. 3, pp. 471–481, Dec. 1970.
- [12] T. Würfl, M. Hoffmann, V. Christlein, K. Breininger, Y. Huang, M. Unberath, and A. K. Maier, "Deep learning computed tomography: Learning projection-domain weights from image domain in limited angle problems," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1454–1463, Jun. 2018.
- [13] A. K. Maier, C. Syben, B. Stimpel, T. Würfl, M. Hoffmann, F. Schebesch, W. Fu, L. Mill, L. Kling, and S. Christiansen, "Learning with known operators reduces maximum error bounds," *Nature Mach. Intell.*, vol. 1, no. 8, pp. 373–380, Aug. 2019.
- [14] C. Syben, M. Michen, B. Stimpel, S. Seitz, S. Ploner, and A. K. Maier, "Technical note: PYRO-NN: Python reconstruction operators in neural networks," *Med. Phys.*, vol. 46, no. 11, pp. 5110–5115, Nov. 2019. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6899669/>
- [15] G. Dong, C. Zhang, L. Deng, Y. Zhu, J. Dai, L. Song, R. Meng, T. Niu, X. Liang, and Y. Xie, "A deep unsupervised learning framework for the 4D CBCT artifact correction," *Phys. Med. Biol.*, vol. 67, no. 5, Mar. 2022, Art. no. 055012.
- [16] Y. Liu, X. Chen, J. Zhu, B. Yang, R. Wei, R. Xiong, H. Quan, Y. Liu, J. Dai, and K. Men, "A two-step method to improve image quality of CBCT with phantom-based supervised and patient-based unsupervised learning strategies," *Phys. Med. Biol.*, vol. 67, no. 8, Apr. 2022, Art. no. 084001.
- [17] K. Choi, "A comparative study between image- and projection-domain self-supervised learning for ultra low-dose CBCT," in *Proc. 44th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2022, pp. 2076–2079.
- [18] Y.-J. Han and H.-J. Yu, "Self-supervised noise reduction in low-dose cone beam computed tomography (CBCT) using the randomly dropped projection strategy," *Appl. Sci.*, vol. 12, no. 3, p. 1714, Feb. 2022.
- [19] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [20] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, Jul. 1945.
- [21] L. Zhu, Y. Chen, J. Yang, X. Tao, and Y. Xi, "Evaluation of the dental spectral cone beam CT for metal artifact reduction," *Dentomaxillofacial Radiol.*, vol. 48, no. 2, Feb. 2019, Art. no. 20180044.
- [22] W. Cao, T. Sun, G. Fardell, B. Price, and W. Dewulf, "Comparative performance assessment of beam hardening correction algorithms applied on simulated data sets," *J. Microsc.*, vol. 272, no. 3, pp. 229–241, Dec. 2018.
- [23] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb. 2006.
- [24] M. Amirian, J. A. Montoya-Zegarra, I. Herzig, P. E. Hotz, L. Lichtensteiger, M. Morf, A. Züst, P. Paysan, I. Peterlik, S. Scheib, R. M. Fuchsli, T. Stadelmann, and F. Schilling, "Mitigation of motion-induced artifacts in cone beam computed tomography using deep convolutional neural networks," *Med. Phys.*, vol. 50, no. 10, pp. 6228–6242, Oct. 2023.
- [25] S. Kida, T. Nakamoto, M. Nakano, K. Nawa, A. Haga, J. Kotoku, H. Yamashita, and K. Nakagawa, "Cone beam computed tomography image quality improvement using a deep convolutional neural network," *Cureus*, vol. 10, no. 4, p. e2548, 2018.
- [26] S. Majee, T. Balke, C. A. J. Kemp, G. T. Buzzard, and C. A. Bouman, "4D X-ray CT reconstruction using multi-slice fusion," in *Proc. IEEE Int. Conf. Comput. Photography (ICCP)*, May 2019, pp. 1–8.
- [27] S. Chang, X. Chen, J. Duan, and X. Mou, "A hybrid ring artifact reduction algorithm based on CNN in CT images," *Proc. SPIE*, vol. 11072, May 2019, Art. no. 1107226.
- [28] K. Xiao, Y. Han, Y. Xu, L. Li, X. Xi, H. Bu, and B. Yan, "X-ray cone-beam computed tomography geometric artefact reduction based on a data-driven strategy," *Applied Opt.*, vol. 58, no. 17, pp. 4771–4780, Jun. 2019.
- [29] D. Choi, J. Kim, S. Chae, B. Kim, J. Baek, A. Maier, R. Fahrig, H. Park, and J. Choi, "Multidimensional noise reduction in C-arm cone-beam CT via 2D-based Landweber iteration and 3D-based deep neural networks," *Proc. SPIE*, vol. 10948, May 2019, Art. no. 1094837.
- [30] N. Dahiya, S. R. Alam, P. Zhang, S. Zhang, T. Li, A. Yezzi, and S. Nadeem, "Multitask 3D CBCT-to-CT translation and organs-at-risk segmentation using physics-based data augmentation," *Med. Phys.*, vol. 48, no. 9, pp. 5130–5141, Sep. 2021.
- [31] C. Szegegy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.
- [32] Z. Fang, B. Ye, B. Yuan, T. Wang, S. Zhong, S. Li, and J. Zheng, "Angle prediction model when the imaging plane is tilted about z-axis," *J. Supercomput.*, vol. 78, no. 17, pp. 18598–18615, Nov. 2022.
- [33] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020.
- [34] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2242–2251.
- [35] D. Clark and C. Badea, "Spectral data completion for dual-source X-ray CT," *Proc. SPIE*, vol. 10948, Mar. 2019, Art. no. 109481F.
- [36] R. Wei, B. Liu, F. Zhou, X. Bai, D. Fu, B. Liang, and Q. Wu, "A patient-independent CT intensity matching method using conditional generative adversarial networks (cGAN) for single X-ray projection-based tumor localization," *Phys. Med. Biol.*, vol. 65, no. 14, Jul. 2020, Art. no. 145009.
- [37] A. Santhanam, M. Lauria, B. Stiehl, D. Elliott, S. Seshan, S. Hsieh, M. Cao, and D. Low, "An adversarial machine learning based approach and biomechanically-guided validation for improving deformable image registration accuracy between a planning CT and cone-beam CT for adaptive prostate radiotherapy applications," *Proc. SPIE*, vol. 11313, Mar. 2021, Art. no. 113130P.
- [38] M. Chu, Y. Xie, J. Mayer, L. Leal-Taixé, and N. Thuerley, "Learning temporal coherence via self-supervision for GAN-based video generation," *ACM Trans. Graph.*, vol. 39, no. 4, pp. 1–13, Aug. 2020.
- [39] Z. Zhang, M. Huang, Z. Jiang, Y. Chang, J. Torok, F.-F. Yin, and L. Ren, "4D radiomics: Impact of 4D-CBCT image quality on radiomic analysis," *Phys. Med. Biol.*, vol. 66, no. 4, Feb. 2021, Art. no. 045023.
- [40] S. Kida, S. Kaji, K. Nawa, T. Imae, T. Nakamoto, S. Ozaki, T. Ohta, Y. Nozawa, and K. Nakagawa, "Visual enhancement of Cone-beam CT by use of CycleGAN," *Med. Phys.*, vol. 47, no. 3, pp. 998–1010, Mar. 2020.
- [41] K. Usui, K. Ogawa, M. Goto, Y. Sakano, S. Kyougoku, and H. Daida, "A cycle generative adversarial network for improving the quality of four-dimensional cone-beam computed tomography images," *Radiat. Oncol.*, vol. 17, no. 1, p. 69, Dec. 2022.
- [42] T. Hase, M. Nakao, K. Imanishi, M. Nakamura, and T. Matsuda, "Improvement of image quality of cone-beam CT images by three-dimensional generative adversarial network," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Nov. 2021, pp. 2843–2846.
- [43] T. Park, A. A. Efros, R. Zhang, and J.-Y. Zhu, "Contrastive learning for unpaired image-to-image translation," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 319–345.
- [44] J. Joseph, P. N. Pournami, and P. B. Jayaraj, "Supervised fan beam computed tomography image synthesis using 3D CycleGAN," in *Proc. IEEE Int. Conf. Signal Process., Inform., Commun. Energy Syst. (SPICES)*, vol. 1, Mar. 2022, pp. 81–86.
- [45] J. Harms, Y. Lei, T. Wang, R. Zhang, J. Zhou, X. Tang, W. J. Curran, T. Liu, and X. Yang, "Paired cycle-GAN-based image correction for quantitative cone-beam computed tomography," *Med. Phys.*, vol. 46, no. 9, pp. 3998–4009, Sep. 2019.
- [46] C. J. O'Hara, D. Bird, B. Al-Qaisieh, and R. Speight, "Assessment of CBCT-based synthetic CT generation accuracy for adaptive radiotherapy planning," *J. Appl. Clin. Med. Phys.*, vol. 23, no. 11, 2022, Art. no. e13737.
- [47] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham, Switzerland: Springer, 2015, pp. 234–241.

- [48] C. Kurz, M. Maspero, M. H. F. Savenije, G. Landry, F. Kamp, M. Pinto, M. Li, K. Parodi, C. Belka, and C. T. van den Berg, "CBCT correction using a cycle-consistent generative adversarial network and unpaired training to enable photon and proton dose calculation," *Phys. Med. Biol.*, vol. 64, no. 22, Nov. 2019, Art. no. 225004.
- [49] Y. Han, J. Kim, and J. C. Ye, "Differentiated backprojection domain deep learning for conebeam artifact removal," *IEEE Trans. Med. Imag.*, vol. 39, no. 11, pp. 3571–3582, Nov. 2020.
- [50] M. J. Lagerwerf, D. M. Pelt, W. J. Palenstijn, and K. J. Batenburg, "A computationally efficient reconstruction algorithm for circular cone-beam computed tomography using shallow neural networks," *J. Imag.*, vol. 6, no. 12, p. 135, Dec. 2020.
- [51] G. Chen, Y. Zhao, Q. Huang, and H. Gao, "4D-AirNet: A temporally-resolved CBCT slice reconstruction method synergizing analytical and iterative method with deep learning," *Phys. Med. Biol.*, vol. 65, no. 17, Sep. 2020, Art. no. 175020.
- [52] K. Lu, L. Ren, and F.-F. Yin, "A geometry-guided deep learning technique for CBCT reconstruction," *Phys. Med. Biol.*, vol. 66, no. 15, Aug. 2021, Art. no. 15LT01.
- [53] K. Lu, L. Ren, and F.-F. Yin, "A geometry-guided multi-beamlet deep learning technique for CT reconstruction," *Biomed. Phys. Eng. Exp.*, vol. 8, no. 4, Jul. 2022, Art. no. 045004.
- [54] M. Thies, J.-N. Zäch, C. Gao, R. Taylor, N. Navab, A. Maier, and M. Unberath, "A learning-based method for online adjustment of C-arm cone-beam CT source trajectories for artifact avoidance," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 15, no. 11, pp. 1787–1796, Nov. 2020.
- [55] A. Kofler, M. Haltmeier, T. Schaeffter, M. Kachelrieß, M. Dewey, C. Wald, and C. Kolbitsch, "Neural networks-based regularization for large-scale medical image reconstruction," *Phys. Med. Biol.*, vol. 65, no. 13, Jul. 2020, Art. no. 135003.
- [56] T. Kurosawa, T. Nishio, S. Moriya, M. Tsuneda, and K. Karasawa, "Feasibility of image quality improvement for high-speed CBCT imaging using deep convolutional neural network for image-guided radiotherapy in prostate cancer," *Phys. Medica*, vol. 80, pp. 84–91, Dec. 2020.
- [57] D. S.-C. Jin, L.-S. Chang, Y.-H. Wang, J.-C. Chen, S. H. Tseng, and T.-Y. Liu, "Virtual and real-world implementation of deep-learning-based image denoising model on projection domain in digital tomosynthesis and cone-beam computed tomography data," *Biomed. Phys. Eng. Exp.*, vol. 8, no. 6, Nov. 2022, Art. no. 065021.
- [58] K. Choi, "Self-supervised projection denoising for low-dose cone-beam CT," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Nov. 2021, pp. 3459–3462.
- [59] D. Choi, W. Kim, J. Lee, M. Han, J. Baek, and J.-H. Choi, "Integration of 2D iteration and a 3D CNN-based model for multi-type artifact suppression in C-arm cone-beam CT," *Mach. Vis. Appl.*, vol. 32, no. 6, p. 116, Nov. 2021.
- [60] K. Chen, L.-B. Zhang, J.-S. Liu, Y. Gao, Z. Wu, H.-C. Zhu, C.-P. Du, X.-L. Mai, C.-F. Yang, and Y. Chen, "Robust restoration of low-dose cerebral perfusion CT images using NCS-UNet," *Nucl. Sci. Techn.*, vol. 33, no. 3, Mar. 2022.
- [61] Z. Jiang, Y. Chen, Y. Zhang, Y. Ge, F.-F. Yin, and L. Ren, "Augmentation of CBCT reconstructed from under-sampled projections using deep learning," *IEEE Trans. Med. Imag.*, vol. 38, no. 11, pp. 2705–2715, Nov. 2019.
- [62] Z. Fu, H. W. Tseng, S. Vedantham, A. Karellas, and A. Bilgin, "A residual dense network assisted sparse view reconstruction for breast computed tomography," *Sci. Rep.*, vol. 10, no. 1, p. 21111, Dec. 2020.
- [63] Y. Zhang, L. Chen, B. Li, M. Folkert, X. Jia, X. Gu, and J. Wang, "Incorporating biomechanical modeling and deep learning into a deformation-driven liver CBCT reconstruction technique," *Proc. SPIE*, vol. 10948, pp. 647–656, Mar. 2019.
- [64] Y. Chen, F.-F. Yin, Z. Jiang, and L. Ren, "Daily edge deformation prediction using an unsupervised convolutional neural network model for low dose prior contour based total variation CBCT reconstruction (PCTV-CNN)," *Biomed. Phys. Eng. Exp.*, vol. 5, no. 6, Oct. 2019, Art. no. 065013.
- [65] Y. Wang, L. Chao, W. Shan, H. Zhang, Z. Wang, and Q. Li, "Improving the quality of sparse-view cone-beam computed tomography via reconstruction-friendly interpolation network," in *Computer Vision—ACCV*, L. Wang, J. Gall, T.-J. Chin, I. Sato, and R. Chellappa, Eds. Cham, Switzerland: Springer, 2023, pp. 86–100.
- [66] L. Chao, P. Zhang, Y. Wang, Z. Wang, W. Xu, and Q. Li, "Dual-domain attention-guided convolutional neural network for low-dose cone-beam computed tomography reconstruction," *Knowl.-Based Syst.*, vol. 251, Sep. 2022, Art. no. 109295.
- [67] D. Wu, K. Kim, and Q. Li, "Computationally efficient deep neural network for computed tomography image reconstruction," *Med. Phys.*, vol. 46, no. 11, pp. 4763–4776, Nov. 2019.
- [68] H. Xie, H. Shan, and G. Wang, "Deep encoder–decoder adversarial reconstruction (DEAR) network for 3D CT from few-view data," *Bioengineering*, vol. 6, no. 4, p. 111, 2019.
- [69] H. Xie, H. Shan, W. Cong, C. Liu, X. Zhang, S. Liu, R. Ning, and G. Wang, "Deep efficient end-to-end reconstruction (DEER) network for few-view breast CT image reconstruction," *IEEE Access*, vol. 8, pp. 196633–196646, 2020.
- [70] Y. Wang, Z. Zhong, and J. Hua, "DeepOrganNet: On-the-fly reconstruction and visualization of 3D/4D lung models from single-view projections by deep deformation network," *IEEE Trans. Vis. Comput. Graphics*, vol. 26, no. 1, pp. 960–970, Jan. 2020.
- [71] Y. Lin, Z. Luo, W. Zhao, and X. Li, "Learning deep intensity field for extremely sparse-view CBCT reconstruction," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI* (Lecture Notes in Computer Science), H. Greenspan, A. Madabhushi, P. Mousavi, N. Salcudean, J. Duncan, T. Syeda-Mahmood, and R. Taylor, Eds. Cham, Switzerland: Springer, 2023, pp. 13–23.
- [72] S. Majee, T. Balke, C. A. J. Kemp, G. T. Buzzard, and C. A. Bouman, "Multi-slice fusion for sparse-view and limited-angle 4D CT reconstruction," *IEEE Trans. Comput. Imag.*, vol. 7, pp. 448–462, 2021.
- [73] Z. Jiang, Z. Zhang, Y. Chang, Y. Ge, F.-F. Yin, and L. Ren, "Prior image-guided cone-beam computed tomography augmentation from under-sampled projections using a convolutional neural network," *Quantitative Imag. Med. Surg.*, vol. 11, no. 12, p. 4767, 2021.
- [74] C. A. A. Júnior, L. F. A. Pereira, G. D. C. Cavalcanti, and T. I. Ren, "Ensemble of convolutional neural networks for sparse-view cone-beam computed tomography," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2022, pp. 1–7.
- [75] D. Hu, Y. Zhang, J. Liu, Y. Zhang, J. L. Coatrieux, and Y. Chen, "PRIOR: Prior-regularized iterative optimization reconstruction for 4D CBCT," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 11, pp. 5551–5562, Nov. 2022.
- [76] S. Ghosh, P. Ernst, G. Rose, A. Nürnberger, and S. Stober, "Towards patient specific reconstruction using perception-aware CNN and planning CT as prior," in *Proc. IEEE 19th Int. Symp. Biomed. Imag. (ISBI)*, Mar. 2022, pp. 1–5.
- [77] S. K. Devalla, P. K. Renukanand, B. K. Sreedhar, G. Subramanian, L. Zhang, S. Perera, J.-M. Mari, K. S. Chin, T. A. Tun, N. G. Strouthidis, T. Aung, A. H. Thiéry, and M. J. A. Girard, "DRUNET: A dilated-residual U-Net deep learning network to segment optic nerve head tissues in optical coherence tomography images," *Biomed. Opt. Exp.*, vol. 9, no. 7, pp. 3244–3265, 2018.
- [78] L. Jiang, X. Wang, H. Jiang, X. Wang, H. Guo, and X. He, "Sparse-view CBCT reconstruction using combined DRUNet and HQS," in *Proc. 5th Int. Conf. Pattern Recognit. Artif. Intell. (PRAI)*, Aug. 2022, pp. 1051–1054.
- [79] Y. Yang, C. Fang, and L. Zhu, "Sparse-view cone-beam breast CT reconstruction via cGAN constrained by image edges," *Zhongguo Yi Liao Qi Xie Za Zhi*, vol. 46, no. 2, pp. 119–125, 2022.
- [80] D. L. Parker, "Optimal short scan convolution reconstruction for fan beam CT," *Med. Phys.*, vol. 9, no. 2, pp. 254–257, Mar. 1982.
- [81] A.-K. Schnurr, K. Chung, T. Russ, L. R. Schad, and F. G. Zöllner, "Simulation-based deep artifact correction with convolutional neural networks for limited angle artifacts," *Zeitschrift Für Medizinische Physik*, vol. 29, no. 2, pp. 150–161, May 2019.
- [82] Y. Wang, T. Yang, and W. Huang, "Limited-angle computed tomography reconstruction using combined FDK-based neural network and U-Net," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 1572–1575.
- [83] M. Rossi, G. Belotti, C. Paganelli, A. Pella, A. Barcellini, P. Cerveri, and G. Baroni, "Image-based shading correction for narrow-FOV truncated pelvic CBCT with deep convolutional neural networks and transfer learning," *Med. Phys.*, vol. 48, no. 11, pp. 7112–7126, Nov. 2021.
- [84] Y. Zhang, "An unsupervised 2D–3D deformable registration network (2D3D-RegNet) for cone-beam CT estimation," *Phys. Med. Biol.*, vol. 66, no. 7, Apr. 2021, Art. no. 074001.
- [85] S. Dweek, S. Dhou, and T. Shanableh, "In-between projection interpolation in cone-beam CT imaging using convolutional neural networks," *Proc. SPIE*, vol. 12031, Apr. 2022, Art. no. 1203129.

- [86] L. Sun, Z. Jiang, Y. Chang, and L. Ren, "Building a patient-specific model using transfer learning for four-dimensional cone beam computed tomography augmentation," *Quant. Imag. Med. Surg.*, vol. 11, no. 2, pp. 540–555, Feb. 2021.
- [87] D. C. Hansen, G. Landry, F. Kamp, M. Li, C. Belka, K. Parodi, and C. Kurz, "ScatterNet: A convolutional neural network for cone-beam CT intensity correction," *Med. Phys.*, vol. 45, no. 11, pp. 4916–4926, Nov. 2018.
- [88] T. M. Gottschalk, A. Maier, F. Kordon, and B. W. Kreher, "DL-based inpainting for metal artifact reduction for cone beam CT using metal path length information," *Med. Phys.*, vol. 50, no. 1, pp. 128–141, Jan. 2023.
- [89] J. Maier, Y. Berker, S. Sawall, and M. Kachelriess, "Deep scatter estimation (DSE): Feasibility of using a deep convolutional neural network for real-time X-ray scatter prediction in cone-beam CT," *Proc. SPIE*, vol. 10573, pp. 393–398, Mar. 2018.
- [90] J. Maier, E. Eulig, T. Vöth, M. Knaup, J. Kuntz, S. Sawall, and M. Kachelrieß, "Real-time scatter estimation for medical CT using the deep scatter estimation: Method and robustness analysis with respect to different anatomies, dose levels, tube voltages, and data truncation," *Med. Phys.*, vol. 46, no. 1, pp. 238–249, Jan. 2019.
- [91] B. van der Heyden, S. Roden, R. Dok, S. Nuyts, and E. Sterpin, "Virtual monoenergetic micro-CT imaging in mice with artificial intelligence," *Sci. Rep.*, vol. 12, no. 1, p. 2324, Feb. 2022.
- [92] M. Sun and J. M. Star-Lack, "Improved scatter correction using adaptive scatter kernel superposition," *Phys. Med. Biol.*, vol. 55, no. 22, pp. 6695–6720, Nov. 2010.
- [93] Y. Nomura, Q. Xu, H. Shirato, S. Shimizu, and L. Xing, "Projection-domain scatter correction for cone beam computed tomography using a residual convolutional neural network," *Med. Phys.*, vol. 46, no. 7, pp. 3142–3155, Jul. 2019.
- [94] A. Lalonde, B. Winey, J. Verburg, H. Paganetti, and G. C. Sharp, "Evaluation of CBCT scatter correction using deep convolutional neural networks for head and neck adaptive proton therapy," *Phys. Med. Biol.*, vol. 65, no. 24, Dec. 2020, Art. no. 245022.
- [95] S. R. Alam, T. Li, P. Zhang, S.-Y. Zhang, and S. Nadeem, "Generalizable cone beam CT esophagus segmentation using physics-based data augmentation," *Phys. Med. Biol.*, vol. 66, no. 6, Mar. 2021, Art. no. 065008.
- [96] Z. Liu, P. Luo, X. Wang, and X. Tang, "Deep learning face attributes in the wild," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 3730–3738.
- [97] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of GANs for improved quality, stability, and variation," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2018, pp. 1–26.
- [98] X. Liang, L. Chen, D. Nguyen, Z. Zhou, X. Gu, M. Yang, J. Wang, and S. Jiang, "Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy," *Phys. Med. Biol.*, vol. 64, no. 12, Jun. 2019, Art. no. 125002.
- [99] J. Zaech, C. Gao, B. Bier, R. Taylor, A. Maier, N. Navab, and M. Unberath, "Learning to avoid poor images: Towards task-aware C-arm cone-beam CT trajectories," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*, vol. 11768, D. Shen, T. Liu, T. Peters, L. Staib, C. Essert, S. Zhou, P. Yap, and A. Khan, Eds. Cham, Switzerland: Springer, 2019, pp. 11–19. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-32254-0_2#citeas
- [100] P. Wu, N. Sheth, A. Sisniega, A. Uneri, R. Han, R. Vijayan, P. Vagdari, B. Kreher, H. Kunze, G. Kleinszig, S. Vogt, S. Lo, N. Theodore, and J. Siewerdsen, "Method for metal artifact avoidance in C-arm cone-beam CT," *Proc. SPIE*, vol. 11312, pp. 522–530, Mar. 2020.
- [101] P. Wu, N. Sheth, A. Sisniega, A. Uneri, R. Han, R. Vijayan, P. Vagdari, B. Kreher, H. Kunze, G. Kleinszig, S. Vogt, S.-F. Lo, N. Theodore, and J. H. Siewerdsen, "C-arm orbits for metal artifact avoidance (MAA) in cone-beam CT," *Phys. Med. Biol.*, vol. 65, no. 16, Aug. 2020, Art. no. 165012.
- [102] A. Uneri, P. Wu, C. K. Jones, P. Vagdari, R. Han, P. A. Helm, M. G. Luciano, W. S. Anderson, and J. H. Siewerdsen, "Deformable 3D-2D registration for high-precision guidance and verification of neuroelectrode placement," *Phys. Med. Biol.*, vol. 66, no. 21, Nov. 2021, Art. no. 215014.
- [103] X. Wang, W. Jian, B. Zhang, L. Zhu, Q. He, H. Jin, G. Yang, C. Cai, H. Meng, X. Tan, F. Li, and Z. Dai, "Synthetic CT generation from cone-beam CT using deep-learning for breast adaptive radiotherapy," *J. Radiat. Res. Appl. Sci.*, vol. 15, no. 1, pp. 275–282, Mar. 2022.
- [104] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2017, pp. 5967–5976.
- [105] F. Fan, B. Kreher, H. Keil, A. Maier, and Y. Huang, "Fiducial marker recovery and detection from severely truncated data in navigation-assisted spine surgery," *Med. Phys.*, vol. 49, no. 5, pp. 2914–2930, May 2022.
- [106] Z. Cao, X. Gao, Y. Chang, G. Liu, and Y. Pei, "A novel approach for eliminating metal artifacts based on MVCBCT and CycleGAN," *Frontiers Oncol.*, vol. 12, Nov. 2022, Art. no. 1024160.
- [107] J. Beaudry, P. Esquinas, and C. Shieh, "Learning from our neighbours: A novel approach on sinogram completion using bin-sharing and deep learning to reconstruct high quality 4DCBCT," *Proc. SPIE*, vol. 10948, pp. 1025–1035, Mar. 2019.
- [108] D. Lee, K. Kim, W. Kim, S. Kang, C. Park, H. Cho, Y. Lim, G. Kim, S. Park, H. Lim, H. Lee, D. Jeon, J. Park, C. Seo, and M. Lee, "Four-dimensional CBCT reconstruction based on a residual convolutional neural network for improving image quality," *J. Korean Phys. Soc.*, vol. 75, no. 1, pp. 73–79, Jul. 2019.
- [109] S. Zhi, J. Duan, J. Cai, and X. Mou, "Artifacts reduction method for phase-resolved cone-beam CT (CBCT) images via a prior-guided CNN," *Proc. SPIE*, vol. 10948, pp. 562–567, Mar. 2019.
- [110] F. Madesta, T. Sentker, T. Gauer, and R. Werner, "Self-contained deep learning-based boosting of 4D cone-beam CT reconstruction," *Med. Phys.*, vol. 47, no. 11, pp. 5619–5631, Nov. 2020.
- [111] S. Zhi, M. Kachelrieß, F. Pan, and X. Mou, "CycN-Net: A convolutional neural network specialized for 4D CBCT images refinement," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3054–3064, Nov. 2021.
- [112] S. Zhi and X. Mou, "TN-Net: A spatiotemporal plus prior image-based convolutional neural network for 4D-CBCT reconstructions enhancement," *Proc. SPIE*, vol. 11595, pp. 474–481, Feb. 2021.
- [113] Z. Zhang, M. Huang, Z. Jiang, Y. Chang, K. Lu, F.-F. Yin, P. Tran, D. Wu, C. Beltran, and L. Ren, "Patient-specific deep learning model to enhance 4D-CBCT image for radiomics analysis," *Phys. Med. Biol.*, vol. 67, no. 8, Apr. 2022, Art. no. 085003.
- [114] Z. Zhang, J. Liu, D. Yang, U. S. Kamilov, and G. D. Hugo, "Deep learning-based motion compensation for four-dimensional cone-beam computed tomography (4D-CBCT) reconstruction," *Med. Phys.*, vol. 50, no. 2, pp. 808–820, Feb. 2023.
- [115] F. Madesta, T. Gauer, T. Sentker, and R. Werner, "Self-consistent deep learning-based boosting of 4D cone-beam computed tomography reconstruction," *Proc. SPIE*, vol. 10949, Mar. 2019, Art. no. 1094902.
- [116] A. Thummerer, C. S. Oria, P. Zaffino, S. Visser, A. Meijers, G. G. Marmitt, R. Wijsman, J. Seco, J. A. Langendijk, A. C. Knopf, M. F. Spadea, and S. Both, "Deep learning-based 4D-synthetic CTs from sparse-view CBCTs for dose calculations in adaptive proton therapy," *Med. Phys.*, vol. 49, no. 11, pp. 6824–6839, Nov. 2022.
- [117] C. Shieh, X. Jia, B. Li, Y. Gonzalez, S. Rit, and P. Keall, "AAPM grand challenge: SPARE-sparse-view reconstruction challenge for 4D cone-beam CT," in *Proc. Amer. Assoc. Physicists Med. Annu. Meeting*, 2018, pp. e120–e706. [Online]. Available: <https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.12938> and <http://amos3.aapm.org/abstracts/pdf/137-41908-452581-138082-560681006.pdf>
- [118] C. Shieh, Y. Gonzalez, B. Li, X. Jia, S. Rit, C. Mory, M. Riblett, G. Hugo, Y. Zhang, Z. Jiang, X. Liu, L. Ren, and P. Keall, "SPARE: Sparse-view reconstruction challenge for 4D cone-beam CT from a 1-min scan," *Med. Phys.*, vol. 46, no. 9, pp. 3799–3811, Sep. 2019.
- [119] Z. Jiang, Z. Zhang, Y. Chang, Y. Ge, F.-F. Yin, and L. Ren, "Enhancement of 4-D cone-beam computed tomography (4D-CBCT) using a dual-encoder convolutional neural network (DeCNN)," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 6, no. 2, pp. 222–230, Feb. 2022.
- [120] J. Liu, Y. Kang, D. Hu, and Y. Chen, "4D-CBCT reconstruction via motion compensation learning induced sparse tensor constraint," in *Proc. 12th Int. Congr. Image Signal Process., BioMedical Eng. Inform.*, 2019, pp. 1–5.
- [121] Y. Zhang, X. Huang, and J. Wang, "Advanced 4-dimensional cone-beam computed tomography reconstruction by combining motion estimation, motion-compensated reconstruction, biomechanical modeling and deep learning," *Vis. Comput. Ind., Biomed., Art.*, vol. 2, no. 1, p. 23, Dec. 2019.
- [122] X. Huang, Y. Zhang, L. Chen, and J. Wang, "U-Net-based deformation vector field estimation for motion-compensated 4D-CBCT reconstruction," *Med. Phys.*, vol. 47, no. 7, pp. 3000–3012, Jul. 2020.

- [123] A. Sisniega, S. Capostagno, W. Zbijewski, J. Stayman, C. Weiss, T. Ehtiati, and J. Siewerdsen, "Estimation of local deformable motion in image-based motion compensation for interventional cone-beam CT," *Proc. SPIE*, vol. 11312, pp. 385–390, Mar. 2020.
- [124] A. Sisniega, H. Huang, W. Zbijewski, J. Stayman, C. Weiss, T. Ehtiati, and J. Siewerdsen, "Deformable image-based motion compensation for interventional cone-beam CT with a learned autofocus metric," *Proc. SPIE*, vol. 11595, pp. 241–248, Feb. 2021.
- [125] H. Huang, J. H. Siewerdsen, W. Zbijewski, C. R. Weiss, M. Unberath, and A. Sisniega, "Context-aware, reference-free local motion metric for CBCT deformable motion compensation," *Proc. SPIE*, vol. 12304, pp. 230–235, Oct. 2022.
- [126] Z. Jiang, Y. Chang, Z. Zhang, F. Yin, and L. Ren, "Fast four-dimensional cone-beam computed tomography reconstruction using deformable convolutional networks," *Med. Phys.*, vol. 49, no. 10, pp. 6461–6476, Oct. 2022.
- [127] I. Herzig, P. Paysan, S. Scheib, A. Züst, F.-P. Schilling, J. Montoya, M. Amirian, T. Stadelmann, P. E. Hotz, and R. M. Fuchsli, "Deep learning-based simultaneous multi-phase deformable image registration of sparse 4D-CBCT," *Med. Phys.*, vol. 49, no. 6, pp. 325–326, 2022.
- [128] A. Hansch, V. Dicken, J. Klein, T. Morgasb, B. Haas, and H. Hahn, "Artifact-driven sampling schemes for robust female pelvis CBCT segmentation using deep learning," *Proc. SPIE*, vol. 10950, pp. 212–219, Mar. 2019.
- [129] W. Xia, C. Niu, W. Cong, and G. Wang, "Sub-volume-based denoising diffusion probabilistic model for cone-beam CT reconstruction from incomplete data," 2023, *arXiv:2303.12861*.
- [130] Z. Wang, C. Vandersteen, T. Demarcy, D. Gnansia, C. Raffaelli, N. Guevara, and H. Delingette, "Inner-ear augmented metal artifact reduction with simulation-based 3D generative adversarial networks," *Computerized Med. Imag. Graph.*, vol. 93, Oct. 2021, Art. no. 101990.
- [131] T. Imae, S. Kaji, S. Kida, K. Matsuda, S. Takenaka, A. Aoki, T. Nakamoto, S. Ozaki, K. Nawa, H. Yamashita, K. Nakagawa, and O. Abe, "Improvement in image quality of CBCT during treatment by cycle generative adversarial network," *Jpn. J. Radiological Technol.*, vol. 76, no. 11, pp. 1173–1184, 2020.
- [132] Y. Song, L. Shen, L. Xing, and S. Ermon, "Solving inverse problems in medical imaging with score-based generative models," 2021, *arXiv:2111.08005*.
- [133] W. Wu, Y. Wang, Q. Liu, G. Wang, and J. Zhang, "Wavelet-improved score-based generative model for medical imaging," *IEEE Trans. Med. Imag.*, Oct. 2023. [Online]. Available: <https://ieeexplore.ieee.org/document/10288274>, doi: [10.1109/TMI.2023.3325824](https://doi.org/10.1109/TMI.2023.3325824).
- [134] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," 2020, *arXiv:2006.11239*.
- [135] B. Kwar, M. Elad, S. Ermon, and J. Song, "Denoising diffusion restoration models," 2022, *arXiv:2201.11793*.
- [136] X. Dai, J. Bai, T. Liu, and L. Xie, "Limited-view cone-beam CT reconstruction based on an adversarial autoencoder network with joint loss," *IEEE Access*, vol. 7, pp. 7104–7116, 2019.
- [137] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation," in *Proc. Int. Workshop Deep Learn. Med. Image Analysis*, vol. 11045, Sep. 2018, pp. 3–11. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7329239/>
- [138] W. P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. W. Tsui, "4D XCAT phantom for multimodality imaging research," *Med. Phys.*, vol. 37, no. 9, pp. 4902–4915, Sep. 2010, doi: [10.1118/1.3480985](https://doi.org/10.1118/1.3480985).
- [139] M. Raissi, P. Perdikaris, and G. Em Karniadakis, "Physics informed deep learning (Part I): Data-driven solutions of nonlinear partial differential equations," 2017, *arXiv:1711.10561*.
- [140] F. Pérez-García, R. Sparks, and S. Ourselin, "TorchIO: A Python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning," *Comput. Methods Programs Biomed.*, vol. 208, Sep. 2021, Art. no. 106236, doi: [10.1016/j.cmpb.2021.106236](https://doi.org/10.1016/j.cmpb.2021.106236).
- [141] M. J. Cardoso, W. Li, R. Brown, N. Ma, E. Kerfoot, Y. Wang, B. Murrey, A. Myronenko, C. Zhao, and D. Yang, "MONAI: An open-source framework for deep learning in healthcare," 2022, *arXiv:2211.02701*.
- [142] A. Thummerer et al., "SynthRAD2023 challenge design: Synthesizing computed tomography for radiotherapy," *Med. Phys.*, vol. 50, pp. 4664–4674, 2023. [Online]. Available: <https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.16529>



MOHAMMADREZA AMIRIAN received the M.Sc. degree in electrical communications technology from Ulm University, Germany, in 2017, and the Ph.D. degree in computer science from the Neural Information Processing Institute, Ulm University. During the Ph.D. study, he was a Researcher with the Institute of Applied Information Technology (InIT) and the Center for Artificial Intelligence (CAI), Zurich University of Applied Sciences (ZHAW), Winterthur, Switzerland. His research interests include biophysiological signal processing for person-centered medical and affective pattern recognition. Furthermore, his research pursuits extend to interpretable deep learning algorithms for medical image processing and quality enhancement in imaging technologies.



DANIEL BARCO received the M.Sc. degree in applied information and data science from the Lucerne University of Applied Sciences and Arts, in 2020. He is currently pursuing the Ph.D. degree in advancing AI with the University of Zurich (UZH). He is also a Researcher with the Centre for Artificial Intelligence, Zurich University of Applied Sciences (ZHAW), Winterthur, Switzerland. His research pursuits revolve around pioneering novel neural architectures for computer vision, while also contributing to the development of robust and trustworthy AI solutions.



IVO HERZIG received the B.Sc. degree in mechatronics from the Zurich University of Applied Sciences (ZHAW), Switzerland. He is an engineer and a computer scientist with professional background in software development, computational geometry, computer graphics, and robotics. He is currently a Researcher with the Institute of Applied Mathematics and Physics (IAMP), ZHAW, where he is focusing on deep-learning for medical image analysis in the area of image-guided radiation therapy (IGRT).



FRANK-PETER SCHILLING received the Ph.D. degree in physics from the University of Heidelberg, Germany, in 2001. Subsequently, he spent many years in fundamental research with the physics laboratories, including CERN, Geneva, Switzerland, where he was involved in the discovery of the Higgs particle, in 2012. He joined the Zurich University of Applied Sciences (ZHAW), Winterthur, Switzerland, in 2018. He is a Senior Lecturer, a Group Leader, and the Deputy Director of the ZHAW's Centre for AI (CAI). Besides managing international scientific projects and teams, and being a top-cited author of particle physics research journal publications, he developed a strong profile in computer science, big data, statistical modeling, and machine learning. His research interests include AI and deep learning, with a focus on computer vision (in particular for Medical Imaging) and machine learning operations (MLOps). In addition, he is interested in trustworthy and certifiable AI, and in applications of deep learning in the physical sciences.