applied sciences

MDPI

Systematic Review

Artificial Intelligence for Artifact Reduction in Cone Beam Computed Tomographic Images: A Systematic Review

Parisa Soltani [1,†], Gianrico Spagnuolo [1,2,†], Francesca Angelone [3,*], Asal Rezaeiyazdi [4], Mehdi Mohammadzadeh [4], Giuseppe Maisto [1], Amirhossein Moaddabi [5], Mariangela Cernera [1], Niccolò Giuseppe Armogida [1], Francesco Amato [6,‡] and Alfonso Maria Ponsiglione [6,*,‡]

1 Department of Neurosciences, Reproductive and Odontostomatological Sciences, University of Naples 'Federico II', 80131 Naples, Italy; parisa.soltani@live.com (P.S.); gspagnuo@unina.it (G.S.); giuseppe.maisto29@gmail.com (G.M.); mariangela.cernera@icloud.com (M.C.); ng.armogida@gmail.com (N.G.A.)
2 Global Research Cell, D.Y. Patil Dental College & Hospital, D.Y. Patil Vidyapeeth, Pune 411018, India
3 Department of Engineering, University of Sannio, 82100 Benevento, Italy
4 Student Research Committee, School of Dentistry, Isfahan University of Medical Sciences, Isfahan 8174673461, Iran; rezaei.asal@gmail.com (A.R.); mehdi.mohammadzadeh80@gmail.com (M.M.)
5 Department of Oral and Maxillofacial Surgery, Dental Research Center, Mazandaran University of Medical Sciences, Sari 4815733971, Iran; a.moaddabi@mazums.ac.ir
6 Department of Electrical Engineering and Information Technology, University of Naples 'Federico II', 80125 Naples, Italy; framato@unina.it
* Correspondence: fangelone@unisannio.it (F.A.); alfonsomaria.ponsiglione@unina.it (A.M.P.)
† These authors contributed equally to this work.
‡ These authors also contributed equally to this work.

Abstract

Cone beam computed tomography (CBCT) allows for rapid and accessible acquisition of three-dimensional images with a lower radiation dose compared to conventional computed tomography (CT) scans. However, the quality of CBCT images is limited by a variety of artifacts. This systematic review attempts to explore different artificial intelligence-based solutions for enhancing the quality of CBCT scans and reducing different types of artifacts in these three-dimensional images. PubMed, Web of Science, Scopus, Embase, Cochrane, and Google Scholar were searched up to March 2025. Risk of bias of included studies was assessed using the QUADAS-II tool. Extracted data included bibliographic information, aim, imaging modality, anatomical site of interest, artificial intelligence modeling approach and details, data and dataset details, qualitative and quantitative performance metrics, and main findings. A total of 27 papers from 2018 to 2025 were included. These studies focused on five areas: metal artifact reduction, scatter correction, image reconstruction improvement, motion artifact reduction, and noise reduction. Artificial intelligence models mainly used U-Net variants, though hybrid and transformer-based models were also explored. The thoracic region was the most analyzed, and the structural similarity index measure and peak signal-to-noise-ratio were common performance metrics. Data availability was limited, with only 26% of studies providing public access and 15% sharing model source codes. Artificial intelligence-driven approaches have demonstrated promising results for CBCT artifact reduction. This review highlights a wide variability in performance assessments and that most studies have not received diagnostic validation, limiting conclusions on the true clinical impact of these artificial intelligence-based improvements.

Keywords: artifact; artificial intelligence; cone beam computed tomography

Academic Editor: Iole Vozza

Received: 24 November 2025
Revised: 24 December 2025
Accepted: 27 December 2025
Published: 30 December 2025

1. Introduction

Cone beam computed tomography (CBCT) is a valuable imaging technique widely utilized in the dental and medical fields. This imaging modality provides detailed three-dimensional images with minimal radiation exposure compared to traditional computed tomography [1]. Despite its strengths, the scanning and reconstruction steps may introduce artifacts that degrade CBCT's diagnostic accuracy [2]. Specific artifacts, such as metallic artifacts, beam hardening, and scatter, may arise from the presence of materials, implants, crowns, and screws. Additionally, images may be degraded as a result of patient motion. The impact of these inconsistencies on the quality and interpretability of images shows the necessity to explore different methodologies for artifact mitigation.

CBCT artifacts have direct clinical implications, as they can obscure anatomical landmarks, mimic or mask pathology, and compromise quantitative accuracy [3,4]. Beam hardening and metal artifacts may hinder the evaluation of adjacent bone or soft tissue structures [5,6]. Motion artifacts degrade spatial resolution and increase the risk of misinterpretation of tissue boundaries [7]. These limitations highlight the importance of artifact reduction strategies, since diagnostic accuracy and treatment decisions usually depend on precise visualization of fine structures.

Earlier CBCT imaging relied mostly on hardware-based solutions, notably Monte Carlo simulations, to model photon interactions during the imaging process [8]. These methods were effective but computationally intensive and time-consuming, limiting their practicality in real-time clinical applications [9,10]. In the search for more efficient approaches, researchers have turned to artificial intelligence (AI) for CBCT image reconstruction and artifact mitigation, due to the neural networks' ability to learn intricate patterns and enhance image fidelity [11].

In recent years, convolutional neural networks (CNNs) [12], U-Net architectures [13], Cycle Generative Adversarial Networks (GANs) [14], and other AI models have gained significant attention for their potential in optimizing CBCT images [15].

CNNs have shown promising results in enhancing CBCT image quality, leading to significant improvements in noise reduction, contrast optimization, and overall image fidelity [16]. U-Net architectures, owing to their superior performance in segmentation tasks, have been employed to precisely delineate structures in basis CBCT images and thus improve diagnostics and treatment planning [17].

Numerous studies have investigated the application of CNNs in image processing, demonstrating their efficacy in tasks such as image denoising, artifact reduction, and feature enhancement [18]. By learning hierarchical representations from data, CNNs show strong potential for overcoming the challenges introduced by CBCT artifacts [19]. Early studies have shown promising results in mitigating metal artifacts, scatter, and other distortions inherent in CBCT scans [20].

U-Net architectures, with their encoder/decoder structure, have proven successful in semantic segmentation tasks. The integration of U-Nets into the CBCT workflow has shown promise in achieving improved segmentation accuracy and artifact reduction [21]. The synergy between AI techniques like U-Nets and the challenges presented by CBCT artifacts enables researchers to seek novel advancements in image reconstruction [22].

Recent systematic reviews have examined AI-based CBCT enhancement. Rusanov et al. reviewed 40 studies and found that synthetic CT generation from CBCT accurately represented CT density, with Cycle-GAN models outperforming U-Net and GANs in image quality and anatomical preservation [23]. Amirian et al. analyzed 13 studies, noting both successes and limitations in deep learning methods for CBCT quality improvement [24]. Bibliometric studies further confirm a surge of interest in AI for medical imaging since 2020, driven by advances in reconstruction algorithms and the clinical need for effective

artifact reduction, especially with the increasing use of metallic implants. Although these studies provided valuable insights, a comprehensive scope across all artifact domains and a quantitative synthesis of performance metrics were largely absent. Therefore, this review aims to systematically analyze AI-based CBCT artifact reduction with a focus on numerical outcome metrics.

2. Materials and Methods

The reporting of this systematic review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 checklist.

2.1. Information Sources and Search Strategy

The search for this systematic review was performed in six databases and search engines including PubMed, Web of Science, Scopus, Embase, Cochrane, and Google Scholar in March 2025 using the following keywords: "artificial intelligence", "deep learning", "neural network", "machine learning", "u-net", "cone beam computed tomography", "cone beam CT", "dental CT", "CBCT", "improvement", "enhancement", "correction", "scatter", "artifact" and "artefact". The exact strategy for each database is presented in Table 1.

Table 1. Detailed queries used in the search strategy.

Database	Search Strategy
PubMed	((Cone AND Beam AND (Computed OR computerized) AND Tomography) OR (Cone AND Beam AND CT) OR (CBCT) OR (dental CT)) AND ((artificial AND intelligence) OR (deep AND learning) OR (machine AND learning) OR (neural AND network) OR (U-net)) AND ((quality AND (enhance* OR improve* OR correct*)) OR artifact* OR artefact* OR scatter*)
Scopus	TITLE-ABS-KEY(((artificial AND intelligence) OR (deep AND learning) OR (machine AND learning) OR (neural AND network) OR (U-net)) AND ((Cone AND Beam AND (Computed OR computerized) AND Tomography) OR (Cone AND Beam AND CT) OR (CBCT) OR (dental CT)) AND ((quality AND (enhance* OR improve* OR correct*)) OR artifact* OR artefact* OR scatter*))
Web of Science	TS = (((artificial AND intelligence) OR (deep AND learning) OR (machine AND learning) OR (neural AND network) OR (U-net)) AND ((Cone AND Beam AND (Computed OR computerized) AND Tomography) OR (Cone AND Beam AND CT) OR (CBCT) OR (dental CT)) AND ((quality AND (enhance* OR improve* OR correct*)) OR artifact* OR artefact* OR scatter*))
Embase	(artificial:ti,ab,kw AND intelligence:ti,ab,kw OR (deep:ti,ab,kw AND learning:ti,ab,kw) OR (machine:ti,ab,kw AND learning:ti,ab,kw) OR (neural:ti,ab,kw AND network:ti,ab,kw) OR 'u net':ti,ab,kw) AND (cone:ti,ab,kw AND beam:ti,ab,kw AND (computed:ti,ab,kw OR computerized:ti,ab,kw) AND tomography:ti,ab,kw OR (cone:ti,ab,kw AND beam:ti,ab,kw AND ct:ti,ab,kw) OR cbct:ti,ab,kw OR 'dental ct':ti,ab,kw) AND (quality:ti,ab,kw AND (enhance*:ti,ab,kw OR improve*:ti,ab,kw OR correct*:ti,ab,kw) OR artifact*:ti,ab,kw OR artefact*:ti,ab,kw OR scatter*:ti,ab,kw)
Google scholar	((artificial AND intelligence) OR (deep AND learning) OR (machine AND learning) OR (neural AND network) OR (U-net)) AND ((Cone AND Beam AND (Computed OR computerized) AND Tomography) OR (Cone AND Beam AND CT) OR (CBCT) OR (dental CT)) AND ((quality AND (enhance* OR improve* OR correct*)) OR artifact* OR artefact* OR scatter*)

2.2. Eligibility Criteria and Selection Process

The review question was "Can the use of artificial intelligence reduce the artifacts of CBCT images?" The PICOS framework was adopted for addressing the research question as follows:

- Population: in-vivo human CBCT images;
- Intervention: artifact reduction techniques based on AI;
- Comparator: no algorithm or alternative artifact reduction methods (other than AI);

- Outcome: image quality metrics;
- Study design: diagnostic accuracy studies, controlled trials, retrospective/prospective cohorts comparing AI vs. comparator, cross-sectional studies, and technical validation papers.

The inclusion criteria for this study were as follows:

- Articles in English without restrictions on time of publication;
- Studies using AI to reduce the artifacts of CBCT images;
- Studies using artificial intelligence models including human CBCT scans.

The exclusion criteria were as follows:

- Studies applying AI to any imaging modality other than CBCT (e.g., CT, 4DCBCT);
- Studies using generative AI models in order to enhance the quality of CBCT images by generating sCTs;
- Reviews and conference abstracts;
- Unavailable full-text;
- Studies employing AI models trained on CBCT images from phantoms, objects, or animal models.

2.3. Screening and Study Selection

Screening and study selection were performed by two researchers independently (PS and FA). A third researcher (AMP) double-checked and verified the included articles. The disagreements were resolved through the consensus of the involved researchers. Inter-reviewer agreement during screening and study selection was assessed using Cohen's kappa statistic, which yielded a value of 0.80, indicating substantial agreement.

2.4. Data Items

The data extraction was performed in two separate spreadsheets (Google, Mountain View, CA, USA) by two researchers independently (PS and FA). A third researcher (AMP) double-checked and verified the extracted data. In this phase, inter-reviewer agreement was assessed qualitatively by comparing independent decisions and resolving discrepancies through discussion and consensus. In addition to the basic information regarding the documents to be screened (i.e., authors, title, year of publication, DOI, and abstract), the following fields were chosen to systematically analyze and cluster the data gathered from the included studies:

- General characteristics:
  - Author, Title, Year;
  - Aim of the study: brief description of the research question of the study;
  - Anatomical region of interest: indication of the anatomical region on which the imaging is focused (e.g., dentition, pelvis, chest, etc.);
  - Main results: brief description of the main results of the study in terms of outcomes.
- Dataset characteristics and management:
  - Dataset size: brief description in terms of number of patients and images analyzed;
  - Dataset publicly available: indication of the dataset availability;
  - Simulated data: indication about the use of synthetic or artificially generated data.
- AI modeling characteristics:
  - AI model: type of model architecture adopted (e.g., CNN, recurrent neural network, U-Net, etc.);

- AI model code publicly available: indication of the availability of the AI model source code;
- Data augmentation: indication regarding the use of data augmentation techniques;
- Performance metrics: indication of quantitative or qualitative methods and metrics to assess the performance of the proposed AI-based approach.

2.5. Data Extraction

The data extraction was performed in two separate spreadsheets (Google, Mountain View, CA, USA) by two researchers independently (PS and FA). A third researcher (AMP) double-checked and verified the extracted data. The disagreements were resolved through the consensus of the involved researchers. Extracted data included authors' names, year of publication, type of publication, aim, imaging modality, anatomical site of interest, AI modeling approach, AI model details, availability of AI model code and dataset, dataset size, dataset availability, use of simulated data, data augmentation, qualitative and quantitative performance metrics, and findings.

2.6. Risk of Bias Assessment

The risk of bias assessment was used on the selected studies using the QUADAS-2 tool. The QUADAS-2 tool is a checklist for the quality assessment of diagnostic accuracy studies consisting of four domains: patient selection, index test, reference standard, and flow and timing. The risk of bias is assessed in each domain with the help of signaling questions. Two team members (NGA and GM) independently performed the risk of bias assessment. In case of disagreement, a third researcher (PS) was consulted.

2.7. Synthesis and Analysis of Results

The results from the analysis of the collected data items were graphically represented using different visual tools. In particular, bar charts, treemaps, and sunburst graphs were used to visualize categorical data and group the studies according to different factors (e.g., aim of the study, AI modeling approach, dataset characteristics, etc.). Scatter plots, histograms, and boxplots were used to represent and display continuous data (e.g., performance metric values of the implemented AI models). Tables were also used to report descriptive statistics and hypothesis test results in terms of *p*-values. More specifically, statistical analyses were performed to compare distributions of the values of the reported performance metrics across the included studies, grouped by different factors. Shapiro–Wilk tests and histograms visual inspection were carried out to assess the normality of the data. Mann–Whitney U tests and Kruskal–Wallis tests were then used to assess differences between groups. Hypothesis tests were carried out at a 95% confidence interval (significance level alpha equal to =0.05). Quantitative data analysis software, namely, Microsoft® Office 365 Excel, SPSS (SPSS Statistics ver.29, IBM Corp.), and Python (Spyder; Python ver. 3.10) were used to represent and display the data as well as to carry out all statistical analyses.

In addition, an extensive descriptive summary of the included studies, grouped according to their main and most relevant characteristics, was also provided.

3. Results

3.1. Search Results

A PRISMA flow diagram was prepared to illustrate the selection process of included papers, as shown in Figure 1. The search yielded a total of 2728 records; of these, 27 met the selection criteria and were included in the systematic review.

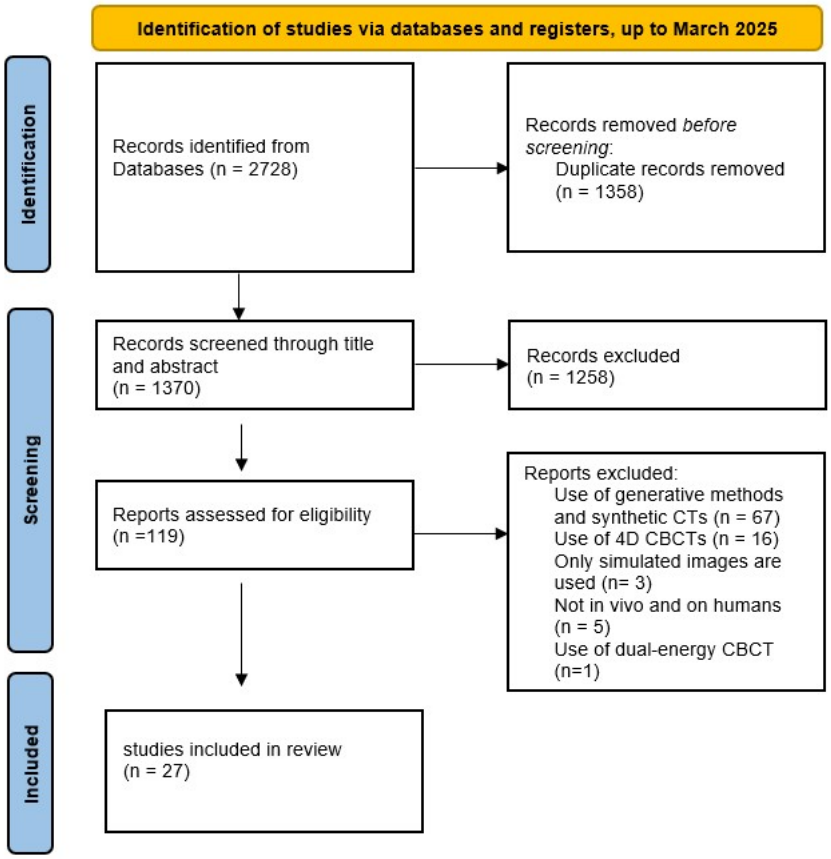

Figure 1. PRISMA flow chart.

3.2. Characteristics of the Included Studies

Table 2 presents the characteristics of the 27 studies included in this systematic review.

3.3. Temporal Distribution

The search for papers highlighted a distribution of papers concerning the topic included in the last eight years, from 2018 to 2025, with a greater interest in 2023, as shown in Figure 2.

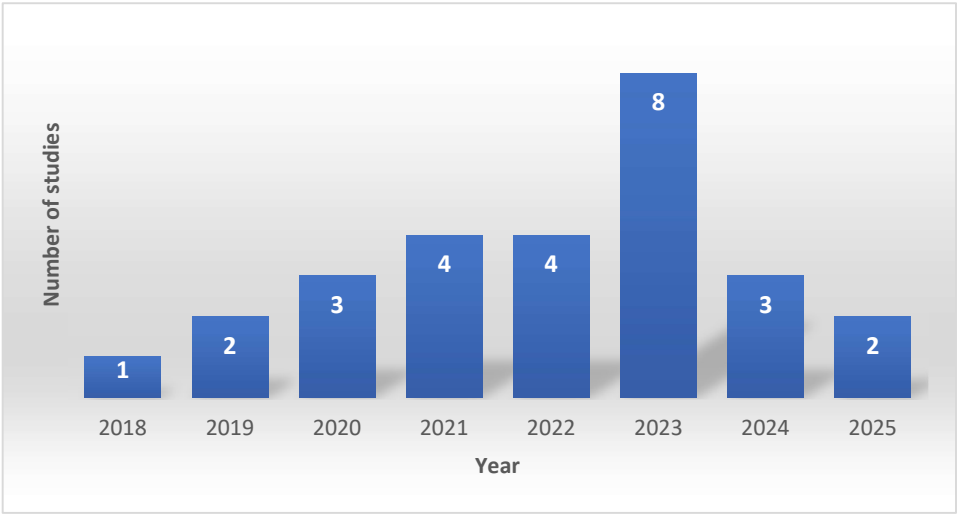

Figure 2. Distribution of articles over the years.

Table 2. Characteristics of the 27 included studies.

No.	Authors	Year	Aim	Main Task	Anatomic Region	AI Model	AI Model Availability	Dataset Availability	Dataset Size	Dataset Splitting	Simulated Data	Data Augmentation	Qualitative Metrics	Quantitative Metrics	Main Quantitative Outcomes
1	Ko et al. [25]	2021	Reducing motion artifacts	Motion artifact reduction	Dentition	AttBlocks	yes	yes	12,560 scans	FBCT teeth dataset: Training: 9660+N2:N18 Testing: 2900 CQ500 dataset: Training: 9600 Testing: 834 CBCT teeth dataset: Training: 9660 Testing: 2900 Chest dataset: Training: 5000 Testing: 2000	yes	yes	Visual evaluation	PSNR, SSIM	PSNR [dB]: 38.21 SSIM: 0.93
2	Jiang et al. [18]	2019	Performing scatter correction	Scatter correction	Pelvis	DRCNN	no	no	20 patients	Training: 15 patients Testing: 2 patients Validation: 3 patients	no	no	Visual evaluation	RMSE, SSIM	RMSE [HU]: 18.80 SSIM: 0.99
3	van der Heyden et al. [26]	2020	Performing scatter correction	Scatter correction	Head-and-neck	DCAE	no	yes	8 patients	Training: 1287 simulated projection pairs (primary and scatter) Testing: 360 simulated projections + 8 real patients	yes	yes	Visual evaluation	CNR, MAE, RMSE, PSNR, SSIM	CNR: 2.50–5.00 MAE [cm^{-1}]: 5.80×10^{-3} MSE [cm^{-2}]: 0.10×10^{-3} PSNR [dB]: 36.80 SSIM: 0.997
4	Hyun et al. [27]	2022	Reducing metal artifacts	Metal artifact reduction	Dentition	Image-enhancing network fIE	no	no	29 scans	Training: 20 patients Testing: 9 patients	yes	no	Visual evaluation	NMSE, PSNR, SSIM	NMSE [HU]: 0.34 SSIM: 0.99 PSNR [dB]: 57.44
5	Hansen et al. [28]	2018	Performing scatter correction	Scatter correction	Pelvis, Prostate	ScatterNet	yes	no	30 patients	Training: 15 patients Testing: 7 patients Validation: 8 patients	no	yes	Visual evaluation	MAE, ME	MAE [HU]: 46.00 ME [HU]: −3.00
6	Kurosawa et al. [29]	2020	Generating high-quality CBCT images by paired low-quality/high-quality CBCT images	Reconstruction improvement	Pelvis	U-Net	no	no	36 patients	Training: 30 patients Testing: 6 patients	yes	yes	Visual evaluation	MAE, SSIM, PSNR	MAE [HU]: 12.00 (ROI bone),13.90 (ROI prostate) PSNR [dB]: 52.8 SSIM: 0.98
7	Ryu et al. [30]	2023	Correcting CBCT images by paired CT images	Reconstruction improvement	Dentition	COMPUNet	no	no	30 patients		yes	yes	Visual evaluation Expert rating	NRMSE, SSIM, MAE	NRMSE: 0.14 SSIM: 0.85 MAE [HU]: 131.60
8	Gottschalk et al. [31]	2023	Reducing metal artifacts by inpainting of metal regions	Metal artifact reduction	Knees, Spines	PaintNet	no	no	55 scans	Training: 33 scans Validation: 11 scans Testing: 11 scans	no	yes	Visual evaluation	SSIM, PSNR	SSIM: 0.99 PSNR [dB]: +6.00

Table 2. *Cont.*

No.	Authors	Year	Aim	Main Task	Anatomic Region	AI Model	AI Model Availability	Dataset Availability	Dataset Size	Dataset Splitting	Simulated Data	Data Augmentation	Qualitative Metrics	Quantitative Metrics	Main Quantitative Outcomes
9	Kim et al. [16]	2022	Reducing metal artifacts and streaking artifacts	Metal artifact reduction	Dentition	U-Net	no	yes	27 XCAT phantoms and 1252 real slices		yes	no	Visual evaluation	NRMSE, SSIM	SSIM: 0.99 NRMSE: 0.03
10	Park et al. [32]	2022	Reducing metal artifacts by iterative correction	Metal artifact reduction	Dentition	Deep convolutional framelets	no	no	10 patients	Training: 49 paired data Validation: 7 paired data Testing: 14 paired data	yes	yes	Visual evaluation	RMSE, STD	RMSE [HU]: 0.39 SD [HU]: 14.97
11	Zhang et al. [33]	2023	Performing scatter correction	Scatter correction	Pelvis	FSTUNet	no	yes	44 patients	MC simulation dataset: Training: 18 patients Testing: 4 patients Validation: 4 patients Frequency split dataset: Training: 34 patients Testing: 5 patients Validation: 5 patients	yes	no	Visual evaluation	RMSE, SSIM, UQI	RMSE [HU]: 7.62 UQI: 0.99 SSIM: 0.93
12	Rusanov et al. [34]	2021	Performing scatter correction by intensity correction	Scatter correction	Head-and-neck	U-Net based	no	no	4 anthropomorphic phantoms and 2 patients	Training: 2001 projections Testing: 1000 projections	yes	yes	Visual evaluation	MAE, SSIM, CNR	MAE [HU]: 74.00 SSIM: 0.81 CNR: 13.90
13	Choi et al. [35]	2021	Reducing complex low-dose noise	Noise reduction	Thorax, Knees	REDCNN	no	no	2 phantoms and 8 patients	Training set: 5 patients, Test set: 3 patients	yes	no	Visual evaluation	SSIM, PSNR	PSNR [dB]: 37.41 SSIM: 0.99 tSNR: 12.10
14	Thies et al. [36]	2020	Reducing metal artifacts by trajectory optimization	Metal artifact reduction	Chest	ConvNet	no	yes	2739 simulated images	1: Training: 1368 Testing: 1 chest CT scan 2: Training: 1368 images Testing: 3 simulations	yes	yes	Visual evaluation	FWHM, Fourier Spectrum Intensity, SSIM	SSIM: 0.90 (no noise), 0.89 (with 4.10×10^5 noise), 0.85 (with 5.10×10^4 noise) FWHM: 6.38 Fourier spectrum intensity: 9.05
15	Hegazy et al. [17]	2019	Reducing metal artifacts by metal segmentation	Metal artifact reduction	Dentition	U-NET	no	no	5 patients		no	yes	Visual evaluation	REL, SSD, NAD	REL (%): 5.70 SSD (%): 6.80 NAD (%): 8.20
16	Ketcha et al. [37]	2021	Reducing metal artifacts and correcting downsampling	Metal artifact reduction	Thorax, Lumbar area	CNNMAR-2	no	no	25 scans	Training: 19 scans Testing: 6 scans	yes	yes	Visual evaluation	RMSE	RMSE [mm^{-1}]: 3.40×10^{-3}
17	Zhuo et al. [38]	2023	Performing scatter correction	Scatter correction	Head, Thorax, Abdomen	Dual-encoder U-Net-like network	no	no	600 projections		yes	yes	Visual evaluation	MAPE, SSIM, RMSE	MAPE (%): 4.73 RMSE [HU]: 4.28 SSIM: 0.93

Table 2. *Cont.*

No.	Authors	Year	Aim	Main Task	Anatomic Region	AI Model	AI Model Availability	Dataset Availability	Dataset Size	Dataset Splitting	Simulated Data	Data Augmentation	Qualitative Metrics	Quantitative Metrics	Main Quantitative Outcomes
18	Agrawal et al. [39]	2023	Reducing metal artifacts by metal segmentation	Metal artifact reduction	Multiple extremity anatomies (e.g., knee, wrist, foot, ankle, palm, forearm)	Modified U-Net	no	no	26 scans	Training: 10 metal-affected scans Testing: 10 metal-affected (100 projection pairs) + 6 metal-free scans (2400 projections)	yes	yes	Visual evaluation	DSC, IOU, FPR	DSC: 94.8 IOU: 90.2 FPR ~0.51×10^{-3}
19	Fan et al. [40]	2024	Reducing metal artifacts by metal segmentation	Metal artifact reduction	Knee and lower limb extremities	SwinConvUNet	yes	no	8200 projections	Training: 6600 projections Validation: 600 projections Testing: 600 projections + 10 cadaver scans (400 projections per scan) + 1 clinical scan (434 projections)	yes	yes	Visual evaluation	PSNR, SSIM	PNSR: 40.598 SSIM: 0.987
20	Hu et al. [41]	2023	Reconstructing high-quality limited-angle images	Reconstruction improvement	Head-and-neck, Pelvis, Thorax	SEA-Net	yes	no	90 patients	Training: 102,500 images Validation: 5000 images Testing: 5000 images	no	yes	Visual evaluation	RMSE, PSNR, SSIM	RMSE: 2.38×10^{-4} PSNR [dB]: 33.61 SSIM: 0.9131
21	Jiang et al. [42]	2025	Reducing metal artifacts by metal segmentation	Metal artifact reduction	Spine	HIDE-Net	no	no	21 patients	Training: 1644 images Validation: 387 images	yes	no	Visual evaluation	RMSE, PSNR, SSIM	RMSE: 24.22 PSNR: 44.800 SSIM: 0.9986
22	Piao et al. [43]	2024	Performing scatter correction	Scatter correction	Head and neck, pelvis	Scatter Kernel Deconvolution + Deep Q-Learning	no	no	3 patients	Training: 40 projections Testing: 1336 projections	yes	no	Visual evaluation	MAPE, MAE, PSNR	MAPE [%]: 6.22 MAE: 0.42 PSNR [dB]: 27.92
23	Song et al. [44]	2024	Reducing metal artifacts by jointly modeling artifact generation and elimination	Metal artifact reduction	Dentition	b-MAR framework: the artifact encoder (E), the metal artifact generator ($G_{(AFtoA)}$), and the metal artifact eliminator ($G_{(AtoAF)}$)	no	no	10,903 images	Training: 8802 images Testing: 2001 images	yes	no	Visual evaluation Expert rating	RMSE, PSNR, SSIM	RMSE [HU]: 2.3373 PSNR [dB]: 42.5753 SSIM: 0.9931
24	Tang et al. [45]	2023	Reducing metal artifacts by a dual-domain (image and projection domain) approach	Metal artifact reduction	Dentition	Prior based sinogram linearization correction + 2 U-Net-based CNNs	no	no	60 projection series	Training: 42 series Validation: 6 series Testing: 13 series	yes	no	Visual evaluation	NRMSD, SSIM	NRMSD (%): 4.0196 SSIM: 0.9924

https://doi.org/10.3390/app16010396

Table 2. *Cont.*

No.	Authors	Year	Aim	Main Task	Anatomic Region	AI Model	AI Model Availability	Dataset Availability	Dataset Size	Dataset Splitting	Simulated Data	Data Augmentation	Qualitative Metrics	Quantitative Metrics	Main Quantitative Outcomes
25	Wajer et al. [46]	2024	Reducing metal artifacts and improving image quality by noise reduction	Metal artifact reduction and Noise reduction	Dentition	ClariCT.AI	no	no	61 patients	not applicable	no	not applicable	Visual evaluation Expert rating	Voxel value difference, artifact index, CNR	Voxel value difference: 174.07 Artifact index: 158.31 CNR: 0.93
26	Yang et al. [47]	2025	Performing scatter correction	Scatter correction	-	DR-Net + FF-Net	no	no	10 patients (+2 phantom scans)	Training: 7 patients (5040 projections) Validation: 1 patient (720 projections) Testing: 2 patients (1440 projections) + 2 phantom scans	yes	no	Visual evaluation	MAE, PSNR	MAE: 3.195×10^{-4} PSNR [dB]: 35.441
27	Yun et al. [48]	2023	Improving the image quality of bowtie-filter-equipped CBCT scans by reducing specific artifacts through a dual-domain approach	Reconstruction improvement	-	Modified residual U-Net + attention U-Net	no	yes	6 patients (+11 phantoms)	Training: 3820 Validation: 1154 Testing: 240	yes	no	Visual evaluation	RMSE, SSIM, CNR	RMSE: 4.57×10^{-2} SSIM: 0.71 CNR: 18.6

MAR: metal artifact reduction; CBCT: cone beam computed tomography; CT: computed tomography; CNN: convolutional neural network; PSNR: peak signal-to-noise ratio; SSIM: structural similarity index measure; RMSE: root mean square error; NRMSE: normalized root mean square error; NMSE: normalized mean square error; MAE: mean absolute error; ME: mean error; STD/SD: standard deviation; CNR: contrast-to-noise ratio; MSE: mean square error; UQI: universal quality index; MAPE: mean absolute percentage error; NRMSD: normalized root mean square deviation; REL: relative error; SSD: sum of squared differences; NAD: normalized absolute difference; DSC: dice similarity coefficient; IOU: intersection over union; FPR: false positive rate; FWHM: full width at half maximum; tSNR: temporal signal-to-noise ratio; AI: artifact index.

3.4. Main Categories of the Included Studies

The studies were grouped into five principal categories based on their stated aims: metal artifact reduction, scatter correction, image reconstruction improvement, motion artifact reduction, and noise reduction.

As seen in Figure 3, most of the papers included in the study deal with removing metal artifacts, followed by papers that deal with scatter correction. Wajer et al. [46] included both metal artifact reduction and noise reduction.

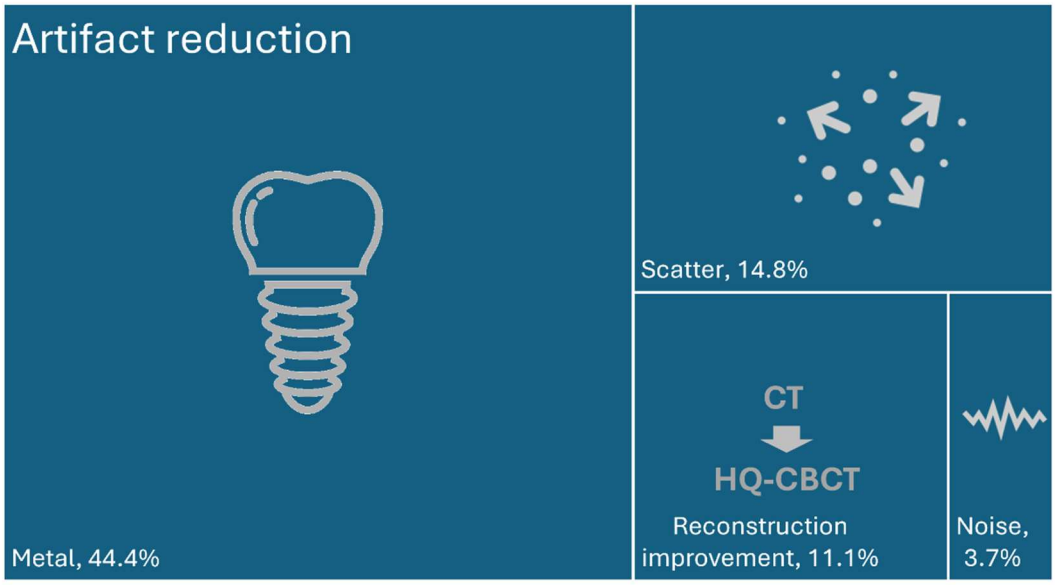

Figure 3. Distribution of the studies according to the main categories.

For metal artifact reduction, the studies provided a broad set of metrics. RMSE values in different studies varied significantly, from very low errors such as 0.39 HU and 2.34 HU in the studies of Park et al. [32] and Song et al. [44] up to 24.22 HU by Jiang et al. [42]. While the PSNR of manipulated images ranged from approximately 40 dB by SwinConvUNet employed by Fan et al. [40] to 57 dB achieved by employing an image-enhancing network (fIE) along with supplementary intra-oral scan data in the study of Hyun et al. [27]. One of the most important image quality metrics, SSIM, extended from 0.85 in noisier conditions [36] to nearly perfect values higher than 0.99 in the study of Jiang et al. [42], Hyun et al. [27], Van Der Heiden et al. [26], Gottschalk et al. [31], among others.

Scattering is a common problem in all X-ray-based diagnostic techniques and becomes more significant as the irradiated area of the object increases. Therefore, scattering is a major issue in CBCT imaging due to the conical shape of the incident X-ray beam. From this point of view, traditional interventions are made from a hardware point of view (anti-scattering grid, airgap, inter-pixel on the detector or septal collimation) and/or from a software point of view, with techniques such as Scatter Kernel Superposition or estimation with mathematical models. These methods are generally limited by low accuracy due to the difficulties in determining scattering kernels or good mathematical models [49,50]. AI-based approaches are becoming increasingly common, also introducing self-supervised approaches, which overcome the limitation related to the lack of completely scatter-free reference images in the clinical context [51]. In the domain of CBCT scattering correction, images with high fidelity were developed by different models, with an SSIM ranging from 0.81 by U-Net-based approaches in the study of Rusanov et al. [34] to 0.99 achieved by the DRCNN model in the study by Jiang et al. [18].

In the image reconstruction improvement category, the PSNR values across these studies range from 33.61 by the SEA-Net model in Hu et al. [41], who aimed at reconstructing high-quality limited-angle images, to 52.8 dB achieved by the U-Net model employed by Kurosawa et al. [29], which aimed to generate high-quality CBCT images by paired low-quality/high-quality images. The SSIM values span from 0.71 to 0.98, according to Yun et al. [48] and Kurosawa et al. [29], respectively.

Among the articles in the noise reduction domain, Choi et al. [35] focused on mitigating complex low-dose noise in C-arm cone beam CT scans. Their model, based on the REDCNN architecture, delivered a PSNR of 37.41 dB and an exceptionally high SSIM of 0.99, indicative of excellent preservation of structural details and substantial noise suppression despite low-dose imaging challenges. Wajer et al. [46] addressed noise reduction in combination with metal artifact reduction using the commercially available ClariCT.AI in dental CBCT imaging, and reported a voxel value difference of 174.07, an artifact index of 158.31, and a CNR of 0.93.

Finally, the study by Ko et al. [25] was the sole study focusing on motion artifact reduction. With a PNSR of 38.21 dB and an SSIM of 0.93, the AttBlocks model used by Ko et al. performed robustly in improving image clarity of CBCT images affected by subject motion.

To summarize, most studies addressed metal artifact reduction, showing wide RMSE variation (0.39–24 HU) but consistently high PSNR (40–57 dB) and SSIM values often above 0.95. Scatter correction achieved SSIM between 0.81–0.99 and RMSE of 4–19 HU, reflecting moderate but meaningful gains. Reconstruction improvement produced PSNR from 33–53 dB and SSIM from 0.71–0.98, with outcomes dependent on dataset quality and architecture. Noise reduction yielded stable improvements, with PSNR around 37–38 dB and SSIM up to 0.99. Motion artifact reduction (Ko et al. [25]) reported PSNR 38 dB and SSIM 0.93, confirming robust enhancement in dentition CBCT scans.

3.5. Anatomic Regions

The CBCT was adopted to image different anatomic regions, distributed as displayed in Figure 4.

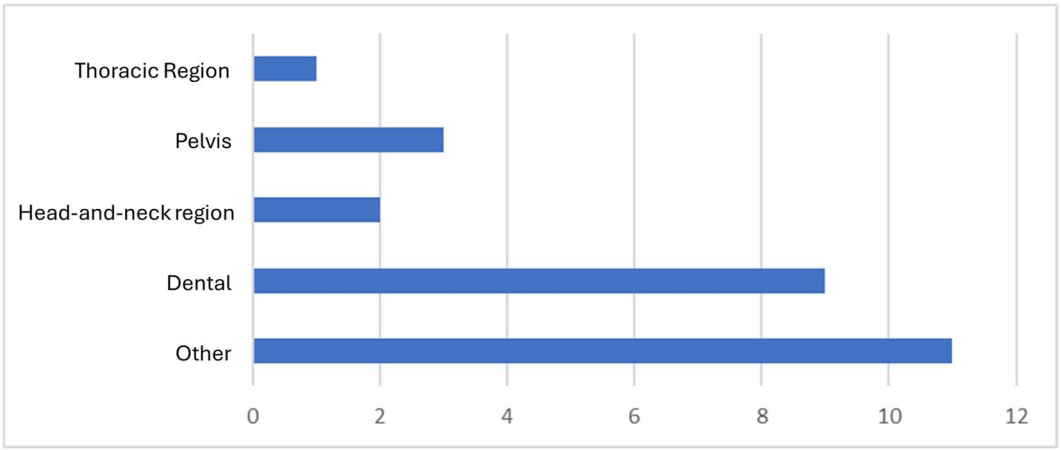

Figure 4. Number of studies distributed according to the anatomical region of interest.

It can be observed that the thoracic region (which takes into account studies focused on breast, chest, lungs, liver, esophagus) is the most investigated one, followed by pelvis, head-and-neck, dental, and other anatomic regions. A notable observation is that AI effectiveness appears to vary with tissue density and anatomical complexity. In dentition, where high-contrast structures dominate, models frequently achieved near-perfect SSIM

values (>0.98) and PSNR above 40 dB, reflecting robust artifact reduction. In contrast, pelvic and thoracic applications, which involve more heterogeneous soft-tissue densities, reported more variable quantitative outcomes (e.g., SSIM 0.81–0.93, PSNR 27–37 dB).

3.6. AI Modeling Approaches

Within the 27 studies, the AI models can be broadly categorized into several groups.

U-Net and its variants:

A prominent group centers on the U-Net architecture and its modifications. For instance, Kurosawa et al. [29] employed a standard U-Net for image reconstruction improvement, while Kim et al. [16] utilized a U-Net for reducing metal and streak artifacts in dental CBCT images. Other U-Net–based approaches include Hegazy et al. [17], who used U-Net for metal segmentation; Agrawal et al. [39], who applied a modified U-Net for projection domain metal segmentation; Tang et al. [45], who combined a prior-based sinogram linearization correction with two U-Net–based CNNs in a dual-domain framework; and Yun et al. [48], who integrated a modified residual U-Net with an attention U-Net for reconstruction improvement. Additionally, Ryu et al. [30] developed COMPUNet, a U-Net variant, further underscoring the dominance of U-Net derivatives in the artifact removal field.

Other CNN-based architectures:

Several studies employed distinct CNN designs beyond the standard U-Net family. Jiang et al. [18] used a deep residual convolutional neural network (DRCNN) for scatter correction; van der Heyden et al. [26] adopted a deep convolutional autoencoder (DCAE) for scatter removal; and Hansen et al. [28] introduced ScatterNet for intensity correction. Complementary architectures include Gottschalk et al. [31]'s PaintNet for inpainting, Park et al. [32]'s use of deep convolutional framelets (DCFs), Choi et al. [35]'s REDCNN, Thies et al. [36]'s ConvNet for trajectory optimization, and Ketcha et al. [37]'s CNNMAR-2 for metal artifact reduction.

Hybrid and dual-domain networks:

Recognizing the benefits of multi-domain learning, some studies integrated two or more branches to capture different image domains. Zhuo et al. [38] implemented a dual-encoder U-Net–like network for scatter correction. In another work, Piao et al. [43] combined scatter kernel deconvolution with deep Q-learning. Similarly, Yang et al. [47] employed a dual-domain network (DR-Net coupled with FF-Net) to mitigate scatter artifacts.

Attention and transformer-enhanced models:

Emerging approaches have started incorporating attention mechanisms and transformer principles to boost performance. Zhang et al. [33] introduced FSTUNet, a Flip Swin Transformer U-shape network, while Fan et al. [40] used SwinConvUNet, a vision transformer-based model, to effectively address metal artifact challenges. SEA-Net, or Structure-Enhanced Attention Network, employed by Hu et al. [41], is another model indicating the integration of attention mechanisms for improved limited-angle CBCT reconstruction.

Other distinct approaches:

A few studies deployed unique or proprietary network architectures. Hyun et al. [27] leveraged an image-enhancing network (fIE) that incorporated supplementary intra-oral scan data for metal artifact reduction, and Wajer et al. [46] utilized ClariCT.AI, a commercially available dedicated deep learning solution for both metal artifact and noise reduction in dental CBCT imaging.

3.7. Dataset and Availability Issues

Among the 27 studies, the reported dataset sizes vary considerably depending on whether the focus is on patient-based data, image-based data, simulated images, or combinations thereof. Sample sizes range from as few as 3 patients (Piao et al., [43]) up to 90 patients (Hu et al., [41]) and 61 patients (Wajer et al. [46]). Ko et al. [25] used a large-scale dataset of 12,560 scans; Fan et al. [40] utilized 12,324 projections; Song et al. [44] employed 10,903 images; and Thies et al. [36] reported 2739 simulated images.

In this review, the dataset was publicly available in 26% of the studies (7 out of 27). In addition, the model source code was available in only 15% of the included studies (4 out of 27). Namely, AttBlocks by Ko et al. [25], ScatterNet by Hansen et al. [28], SwinConvUNet by Fan et al. [40], and SEA-Net by Hu et al. [41].

Out of the 27 studies, 17 (approximately 63%) explicitly reported using simulated data in their experiments. The simulated data used in the included studies mainly refers to the addition of artifacts on clear images to create image pairs with and without noise for supervised approaches or to the use of anthropomorphic phantoms only for the network training phase, again to be able to use a supervised approach and have clean–noisy image pairs.

Out of the 27 studies, 15 reported using data augmentation techniques, approximately 56% of the studies. This means that more than half of the approaches incorporated data augmentation, likely to expand training set diversity and help mitigate overfitting, while the remaining 12 did not employ or report any data augmentation methods. The data augmentation techniques adopted in the included studies mainly refer to cropping, rotation, translation, horizontal and vertical flipping, patch extraction, and downsampling operations, even combined, to increase the training dataset.

3.8. Performance Metrics

Among the quantitative performance metrics reported in the studies, the most commonly used metric was the SSIM, which appeared in 19 out of 27 studies. This prevalence indicates its importance in revealing the preservation of anatomical detail and structural integrity after image correction and processing. The PSNR was reported in 12 studies, reflecting its important role as a quantifier of signal fidelity and overall noise suppression. Error-based measures, particularly RMSE or its variants (NRMSE/NRMSD), were found in 10 studies, providing a direct numerical estimation of error levels in corrected images, while MAE was used in 7 studies. The Pareto Chart of the performance metrics adopted in the included studies is shown in Figure 5.

The SSIM takes three factors into account when comparing images before and after denoising: luminance (l), contrast (c), and structure (s). Considering that, in our case, x and y are the CBCT images before and after the denoising, the formula for calculating the SSIM is therefore given by:

$$SSIM(x,y) = [l(x,y)]^{\alpha}[c(x,y)]^{\beta}[s(x,y)]^{\gamma} \quad (1)$$

where α, β, γ are positive scalars that take into account the relative importance of each metric considered. Considering that:

$$l(x,y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$$

$$c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$

$$s(x,y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$$

where μ_x e μ_y are the local means for images x and y, σ_x, and σ_y are the standard deviations, σ_{xy} is the cross-variance, and that $\alpha = \beta = \gamma = 1$ and $C_3 = C_2/2$, the formula in (1) became:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)\left(\sigma_x^2 + \sigma_y^2 + C_2\right)}$$

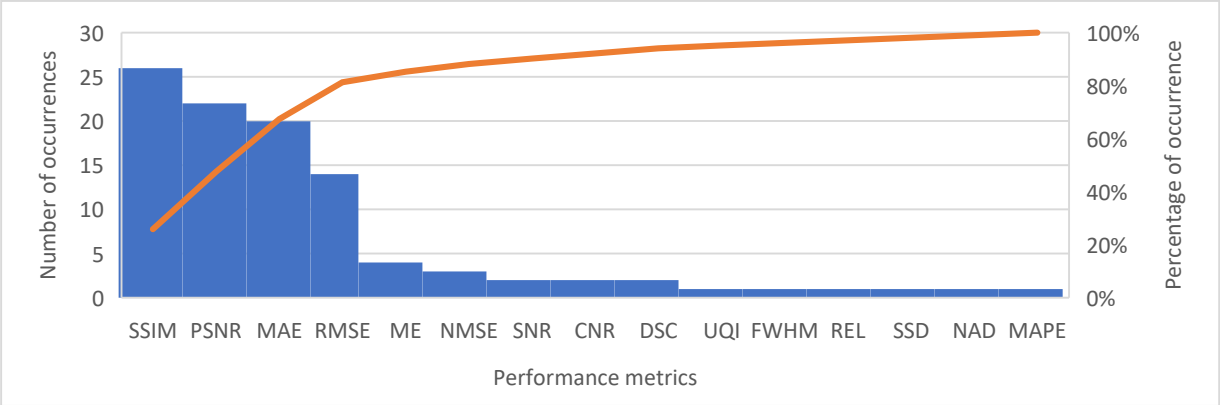

Figure 5. Pareto Chart of the performance metrics adopted in the included studies. Blue bars represent the number of occurrences of each performance metric in the reviewed studies, while the orange line represents the cumulative percentage of occurrences (Pareto distribution). Structure similarity index measure (SSIM), peak signal-to-noise ratio (PSNR), mean absolute error (MAE), root mean square error (RMSE), mean error (ME), signal-to-noise ratio (SNR), contrast-to-noise ratio (CNR), normalized cross-correlation (NCC), Dice–Sørensen coefficient (DSC), normalized mean square error (NMSE), Hausdorff distance (HD), CT number (CTnum), normalized mutual information (NMI), percent integrity uniformity (PIU), universal quality index (UQI), full width at half maximum (FWHM), relative error (REL), sum of square difference (SSD), normalized absolute difference (NAD), mean absolute percentage error (MAPE).

The PSNR is given by the ratio between the maximum power of the signal and the power of corrupting noise. Assuming that the images have size $m \times n$, we can define the formula of the MSE,

$$MSE(x,y) = \frac{1}{m\,n}\sum_{i=0}^{m-1}\sum_{i=0}^{n-1}[x(i,j) - y(i,j)]^2$$

The PSNR is defined as:

$$PSNR = 20(max_x) - 10(MSE)$$

The MAE can be defined as the mean absolute distance between the pixel in the clear image after denoising and its corresponding pixel in the noisy image. In the formula, it can be defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|x_i - y_i|$$

Finally, we can define the RMSE as a measure of the magnitude of error, and it is simply the square root of the MSE.

Only three studies (Ryu et al. [30], Song et al. [44], and Wajer et al. [46]) included expert ratings as a component of their qualitative metrics. The remainder of the studies relied solely on visual representation and comparison of the processed and unprocessed CBCT images, without detailing the diagnostic validation outcomes.

Figure 6 represents the balloon chart and histograms of the distribution of the three most used metrics in the selected studies (SSIM, PSNR, and MAE). In particular, the balloon charts allow us to visualize the performances in relation to the sample size of the dataset. It is not possible to highlight a clear correspondence between sample size and performance for any of the three metrics considered.

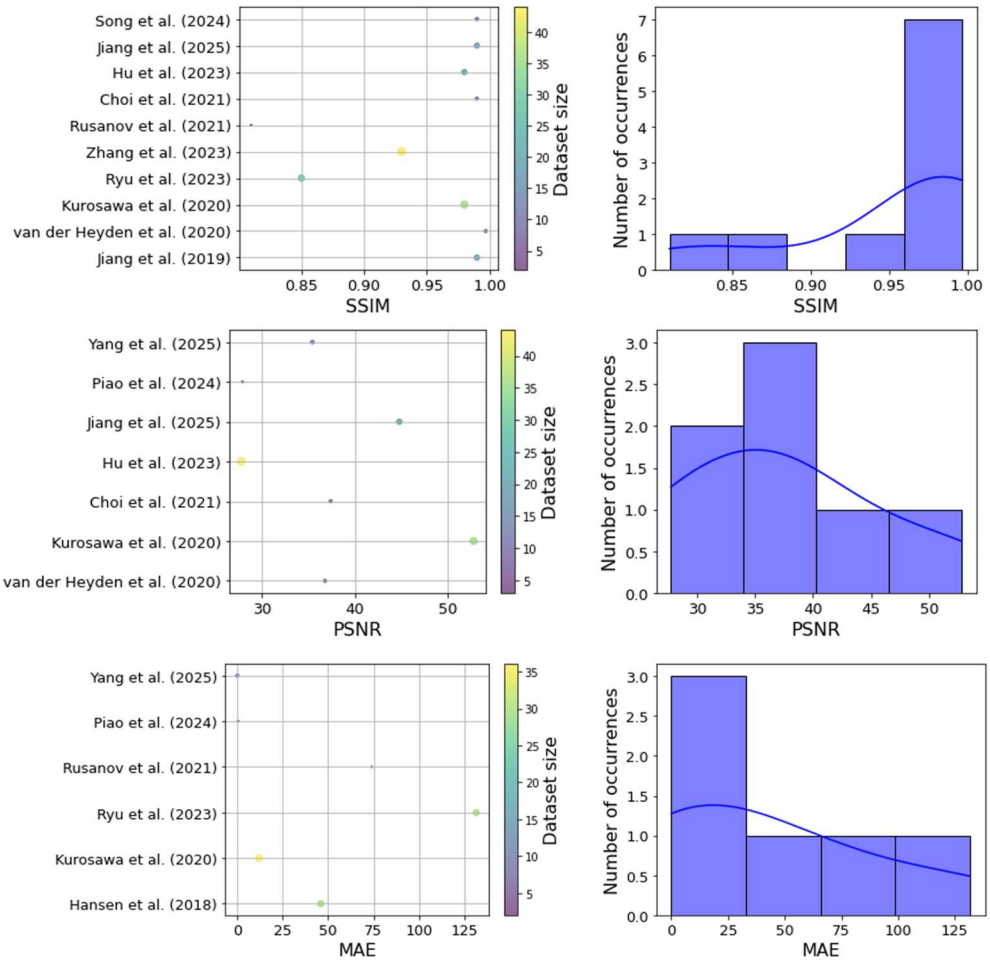

Figure 6. Distribution of performance metrics values across the studies (balloon chart and histogram) for SSIM (**top**), PSNR (**middle**), and MAE (**bottom**). The bars represent the frequency distribution (histogram) of the metric values reported in the included studies, while the solid line represents the corresponding kernel density estimate, used to highlight the general trend of the distribution [18,26,28–30,33–35,41–44,47].

The histograms, on the other hand, allow us to visualize the distribution of the values of the metrics considered in the included studies. In this case, great variability was highlighted in the studies included.

4. Discussion

In this review, the applications of AI for the reduction of different artifacts in CBCT images were explored. Through reviewing 27 studies, five principal applications were identified: metal artifact reduction, scatter correction, image reconstruction improvement, motion artifact reduction, and noise reduction. CBCT images of the thoracic region were the most commonly investigated anatomic area in the studies. AI models predominantly relied on U-Net variants, though other CNN-based and hybrid architectures, including transformer-enhanced models, were also employed.

Artifacts can affect different images, potentially limiting the diagnostic value of radiographic modalities [52]. In CT imaging, artifact refers to any systematic discrepancy between the reconstructed image's CT numbers and the object's true attenuation coefficients. CBCT, with its unique features, provides rapid and cost-beneficial imaging. CBCT imaging offers several distinct features that set it apart from conventional CT scans. Its advantages lie in its specialized design, including cone beam geometry, area scanner, and unique source-detector assembly movements, leading to specific clinical applications, rapid acquisition, and reduced radiation exposure. However, these very advantages can sometimes compromise image quality. CBCT scans are performed quickly, often taking only seconds to complete. This rapid acquisition is advantageous for patient comfort, especially in cases where immobilization is challenging (e.g., pediatric or claustrophobic patients). In addition, CBCT employs lower radiation doses compared to conventional CT. However, spatial resolution in CBCT is generally inferior to that of conventional CT due to larger voxels and limited detector arrays. Moreover, soft tissue contrast is compromised as a result of increased scatter to the area detectors of the CBCT scanner, affecting the ability to differentiate subtle structures. Additionally, CBCT is susceptible to different types of artifacts [53]. CBCT is tailored for specific anatomical regions, such as the maxillofacial area, dental structures, and extremities. Its focused field of view (FOV) allows detailed visualization of localized structures, minimizing unnecessary exposure to other areas.

Different classifications have been proposed for artifacts in CT/CBCT images [4]. One convenient way is to categorize CBCT artifacts based on their etiology. (1) Inherent artifacts are those resulting from the intrinsic characteristics of CBCT imaging, including cone beam geometry, voxel size, and area detectors. This group consists of cone beam effects, partial volume averaging, and scatter artifacts. (2) Procedure-related artifacts that arise during image acquisition and are linked with image processing, including downsampling, scanner-related artifacts (ring artifact), and misalignment of the source-detector assembly. (3) Introduced artifacts, resulting from contents of the FOV and primarily consisting of beam hardening, creating cupping (blooming) of metallic objects and metallic artifacts in the form of streaking and extinction. (4) Motion artifacts that result either from physiologic subtle motions, such as breathing and cardiac cycle, or movement of the patient's head or extremities during CBCT scanning [5,54].

Among the included 27 studies, many studies relied on U-Net and its variants, such as COMPUNet and modified U-Net architectures, due to their proven ability to maintain structural integrity and their relatively straightforward implementation with limited datasets. These models typically excel in tasks like artifact segmentation and reconstruction improvement, where preserving fine anatomical details is crucial [13,55]. On the other hand, more recent approaches have integrated advanced techniques such as dual-domain processing, attention mechanisms, and transformer-based modules (e.g., FSTUNet and SwinConvUNet) to capture both local and global features more effectively. These innovations have been particularly beneficial in more challenging scenarios like metal artifact reduction, where the complex interplay of noise, scattering, and beam hardening demands a more nuanced treatment [56–58].

Similar reviews have been conducted in the past few years. Rusanov et al. [23] performed a systematic review on 40 articles from January 2018 to April 2022 that reviewed the use of DL (U-net, GANs, deep CNN, and artifact disentanglement networks) to improve the quality of CBCT images. The included studies were divided into two categories including synthetic CT (sCT) generation and projection domain. They concluded that sCTs generated by CBCT precisely represented real CT density information for the head and neck, pelvic, and thoracic regions. Cycle-GAN models had better performance than both the U-Net and GAN approaches regarding image quality and anatomical preservation. For

photon and proton plans, the dosimetric performance of sCTs was completely validated in the pelvic and head and neck regions. Another review by Amirian et al. [24] on 13 studies explored the success and limitations of CBCT quality improvement methods based on deep learning. The search for papers was conducted without setting time limits. A higher number of papers concerning the topic were observed in recent years. Similarly, bibliometric studies indicate an increased interest in AI applications for various medical purposes starting from 2020 [59,60]. This growing interest can be attributed to several key factors. Recent years have witnessed significant developments in deep learning-based image reconstruction algorithms. These AI-driven algorithms effectively reduce image artifacts and enhance image quality. Additionally, the increasing use of metallic implants for the aging population necessitates effective artifact reduction in CBCT images. Therefore, the combination of advancements in AI and clinical needs for high-quality CBCT images has led to the growing interest in using AI-driven methods for artifact reduction.

The present review showed that the model source code is available in only 15% of the included studies. This indicates a lack of transparency and reproducibility of the performance of individual models. A similar trend was also seen in other studies. In a study surveying the scientific manuscripts published in the Radiological Society of North America journals, Venkatesh et al. reported that among original articles concerning AI from 2017 through 2021, only 34% of the studies shared their code, while 11% of all articles shared reproducible code [61]. The availability of datasets and model source codes in clinical research deeply impacts the field by enhancing transparency, reproducibility, and collaboration. Open-source models allow researchers to validate and improve the existing algorithms and compare newly developed models against established benchmark models. Researchers can examine the data quality, preprocessing steps, and model architecture. In addition, reproducing research findings becomes feasible when others can access the same data and code. It ensures that results are not isolated incidents but can be independently verified. Additionally, open-source models allow rigorous testing, and researchers can validate the performance, generalization, and limitations of existing algorithms. Finally, transparent models facilitate ethical discussions, as researchers can assess biases, fairness, and potential errors of the developed models [62].

While the use of simulated/synthetic data and augmented data in clinical research could be seen as an advancement in terms of enhancing the generalizability and reproducibility of AI models by using data that can mimic real patient characteristics and information without compromising privacy or security, it could, however, impact the model's reliability. Indeed, synthetic data could introduce biases when they do not accurately represent real-world scenarios and variability. This could also cause a drop in the model's performance in real-world or clinically relevant settings. On the other hand, data augmentation techniques, which can enrich datasets with additional information, potentially help inherently unbalanced datasets, thus providing fairer predictions.

An analysis of the main quantitative outcomes that emerged from the included studies was conducted. In particular, the most used metrics to evaluate the quality of CBCT images in the included studies were SSIM, PSNR, and error-based metrics, respectively. SSIM is a particularly important quality metric because, unlike the other two metrics, which quantify the error compared to a reference image, it quantifies the difference in structural information compared to a reference image, reproducing the human visual evaluation system. By quantifying the structural similarity between the original and distorted images, SSIM allows researchers to assess how perceptually significant these artifacts are for accurate diagnosis. SSIM is, thus, to fine-tune artifact reduction algorithms, optimizing image quality while minimizing artifacts and bridges the gap between technical metrics and clinical relevance, ensuring that CBCT images remain diagnostically reliable [63]. The limited reporting of

diagnostic validation outcomes highlights a major gap in the current literature. While most studies demonstrate technical improvements in image quality, few establish whether these enhancements translate into improved diagnostic accuracy or clinical decision-making. This scarcity restricts conclusions about the real-world utility of AI-based CBCT artifact reduction and indicates the need for future research to incorporate diagnostic endpoints, such as pathology detection, treatment planning accuracy, and inter-observer agreement.

A proper dataset size in AI studies is a much-discussed issue in the scientific community, especially in the medical field, where access to a large amount of data is not always possible. Although adequate sample dimensionality has been identified as a crucial factor in the evaluation and development of robust and generalizable models [64,65], as well as influencing performance [66], no methods for calculating the required sample size for a given model have been developed and standardized. In this review, we attempted to relate the performances with the sample size of the dataset, through appropriate balloon charts, trying to highlight how much the different image samples used in the different papers influenced the performance trend. However, no significant pattern could be highlighted for any of the three most used metrics, not highlighting in this specific case a significant influence of the sample size on performances.

Across different tasks, the performance of AI models varied substantially according to architectural choices and application contexts. In motion artifact reduction, the singular study by Ko et al. [25] using the AttBlocks model achieved a PSNR of 38.21 dB and an SSIM of 0.93, demonstrating considerable handling of motion distortions. For scatter correction, a range of models was employed, from U-Net-based approaches that reached an SSIM as low as 0.81 to dedicated models like DRCNN by Jiang et al. [18], which achieved an exceptional SSIM of 0.99. Metal artifact reduction studies showed metric variability, with RMSE values ranging from as low as 0.39 HU (by Park et al. [32]) to 24.22 HU (Jiang et al. [42]), PSNR values from approximately 40 dB with the SwinConvUNet model (Fan et al. [40]) to 57 dB using an image-enhancing network with supplementary intra-oral scan data (Hyun et al. [27]), and SSIM scores from 0.85 to nearly perfect values above 0.99. In image reconstruction improvement, the U-Net-based approach by Kurosawa et al. [29] showed outstanding performance (52.8 dB PSNR and 0.98 SSIM) compared to the SEA-Net model by Hu et al. [41], which reported a lower PSNR of 33.61 dB and a moderate SSIM of 0.9131, while Yun et al. [48] with the dual-domain network reported an SSIM as low as 0.71 in challenging conditions. Lastly, noise reduction approaches, for instance REDCNN-based method in Choi et al. [35], achieved a PSNR of 37.41 dB and an SSIM of 0.99. Collectively, these findings show that the success of CBCT artifact correction is highly dependent on the chosen model architecture, the quality and type of training data, and the specific performance metrics targeted in each study. Regardless of the study context, AI-based artifact reduction remains effective, emphasizing its broad applicability.

Despite demonstrating strong technical performance, the clinical readiness of AI-driven approaches remains limited. Three CT vendor companies were among the first to launch Food and Drug Administration-approved deep learning algorithms for image reconstruction: TrueFidelity (GE Healthcare, Wauwatosa, WI, USA), AiCE (Canon Medical Systems, Tustin, CA, USA), and Precise Image (Philips Healthcare, Amsterdam, The Netherlands). In addition, independent vendor-neutral platforms, including ClariCT.AI (ClariPi, Seoul, South Korea) and PixelShine (AlgoMedica, Heidelberg, Germany), were also approved by the Food and Drug Administration. However, CBCT applications are still largely at the proof-of-concept stage. Among the included studies, Wajer et al. [46] investigated noise reduction in combination with metal artifact reduction using ClariCT.AI in dental CBCT imaging. They reported that the method significantly enhances image

quality by reducing noise and metal artifacts, while acknowledging that further research with larger, multicenter cohorts is needed to validate these findings.

As highlighted by this review, AI-driven methods for CBCT artifact reduction show promising advances but remain far from routine clinical adoption. Although diverse architectures, from U-Net variants to attention-based models, demonstrate improvements in image quality, their integration into workflows requires rigorous validation and transparency. Outcomes are influenced not only by model design but also by dataset quality and training protocols, showing the need for standardized evaluation. A further challenge is the heterogeneity in naming and describing deep learning architectures, with many proprietary labels masking incremental modifications of established networks. This terminological fragmentation complicates comparisons and obscures the true added value of each approach. Moreover, few studies have systematically compared advanced architectures with robust baselines such as U-Net or ResNet, leaving uncertainty about whether reported gains reflect genuine architectural superiority or dataset-specific factors.

It is also important to highlight another limitation: the near absence of diagnostic or task-based evaluations. Metrics such as SSIM and PSNR quantify similarity but do not capture clinical utility. Only three studies included expert assessments, and none evaluated diagnostic performance (accuracy, sensitivity, specificity, confidence, or inter-reader agreement). Artifact reduction should ultimately be judged by its ability to support more reliable clinical decisions [67].

In light of these findings, future development should include at least three elements currently missing: (1) direct comparisons between models, including robust baselines, to isolate the contribution of advanced architectures; (2) structured clinical assessments with experts, reading tests, and task-based metrics such as detectability index or ROC analysis for specific lesions; and (3) standardized and shared reporting metrics that extend beyond pixel-wise similarity to emphasize diagnostic validation. Without these steps, algorithms risk improving images only in aesthetic or numerical terms, without a meaningful impact on diagnostic quality.

5. Conclusions

AI-driven methods show strong promise for CBCT artifact reduction, achieving PSNR values between ~28 and 57 dB and SSIM scores up to 0.99 across tasks such as metal artifact reduction, scatter correction, reconstruction, and noise suppression. U-Net variants remain reliable, while newer attention- and transformer-based models demonstrate added strength in complex scenarios. Despite these advances, variability in datasets, reliance on simulations, and inconsistent metrics emphasize the need for standardized protocols and clinical validation before routine adoption.

Author Contributions: P.S.: Conceptualization, Methodology, Data curation, Writing—Original draft preparation, Writing—review and editing; G.S.: Conceptualization, Methodology, Data curation, Writing—Original draft preparation, Writing—review and editing; F.A. (Francesca Angelone): Visualization, Investigation, Software, Data curation, Writing—original draft, Writing—review and editing; A.R.: Data Curation, Visualization, Investigation, Writing—Original draft preparation; M.M.: Data Curation, Visualization, Investigation, Writing—Original draft preparation; G.M.: Data Curation, Visualization, Investigation, Methodology, Writing—Original draft preparation; A.M.: Data Curation, Visualization, Investigation, Writing—review and editing; M.C.: Data Curation, Visualization, Investigation, Writing—review and editing; N.G.A.: Data Curation, Visualization, Investigation, Methodology, Writing—review and editing; F.A. (Francesco Amato): Supervision, Software, Validation, Methodology, Writing—review and editing; A.M.P.: Supervision, Software, Validation, Methodology, Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study are available upon request from the authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Mehdizadeh, M.; Booshehri, S.G.; Kazemzadeh, F.; Soltani, P.; Motamedi, M.R.K. Level of knowledge of dental practitioners in Isfahan, Iran about cone-beam computed tomography and digital radiography. *Imaging Sci. Dent.* **2015**, *45*, 133–135. [CrossRef]
2. Schulze, R.; Heil, U.; Groβ, D.; Bruellmann, D.D.; Dranischnikow, E.; Schwanecke, U.; Schoemer, E. Artefacts in CBCT: A review. *Dentomaxillofacial Radiol.* **2011**, *40*, 265–273. [CrossRef] [PubMed]
3. Soltani, P.; Moaddabi, A.; Mehdizadeh, M.; Bateni, M.R.; Naghdi, S.; Cernera, M.; Mirrashidi, F.; Azimipour, M.M.; Spagnuolo, G.; Valletta, A. Effect of a metal artifact reduction algorithm on cone-beam computed tomography scans of titanium and zirconia implants within and outside the field of view. *Imaging Sci. Dent.* **2024**, *54*, 313. [CrossRef]
4. Nagarajappa, A.K.; Dwivedi, N.; Tiwari, R. Artifacts: The downturn of CBCT image. *J. Int. Soc. Prev. Community Dent.* **2015**, *5*, 440–445. [CrossRef] [PubMed]
5. Shavakhi, M.; Soltani, P.; Aghababaee, G.; Patini, R.; Armogida, N.G.; Spagnuolo, G.; Valletta, A. A Quantitative Evaluation of the Effectiveness of the Metal Artifact Reduction Algorithm in Cone Beam Computed Tomographic Images with Stainless Steel Orthodontic Brackets and Arch Wires: An Ex Vivo Study. *Diagnostics* **2024**, *14*, 159. [CrossRef]
6. Merone, G.; Valletta, R.; De Santis, R.; Ambrosio, L.; Martina, R. A novel bracket base design: Biomechanical stability. *Eur. J. Orthod.* **2010**, *32*, 219–223. [CrossRef] [PubMed]
7. Peterlik, I.; Strzelecki, A.; Lehmann, M.; Messmer, P.; Munro, P.; Paysan, P.; Plamondon, M.; Seghers, D. Reducing residual-motion artifacts in iterative 3D CBCT reconstruction in image-guided radiation therapy. *Med. Phys.* **2021**, *48*, 6497–6507. [CrossRef]
8. Golosio, B.; Schoonjans, T.; Brunetti, A.; Oliva, P.; Masala, G.L. Monte Carlo simulation of X-ray imaging and spectroscopy experiments using quadric geometry and variance reduction techniques. *Comput. Phys. Commun.* **2014**, *185*, 1044–1052. [CrossRef]
9. Qin, P.; Lin, G.; Li, X.; Piao, Z.; Huang, S.; Wu, W.; Qi, M.; Ma, J.; Zhou, L.; Xu, Y. A correlated sampling-based Monte Carlo simulation for fast CBCT iterative scatter correction. *Med. Phys.* **2023**, *50*, 1466–1480. [CrossRef]
10. Atanassov, E.; Dimov, I.T. What Monte Carlo models can do and cannot do efficiently? *Appl. Math. Model.* **2008**, *32*, 1477–1500. [CrossRef]
11. Amirian, M.; Montoya-Zegarra, J.A.; Herzig, I.; Eggenberger Hotz, P.; Lichtensteiger, L.; Morf, M.; Züst, A.; Paysan, P.; Peterlik, I.; Scheib, S.; et al. Mitigation of motion-induced artifacts in cone beam computed tomography using deep convolutional neural networks. *Med. Phys.* **2023**, *50*, 6228–6242. [CrossRef]
12. Yamashita, R.; Nishio, M.; Do, R.K.G.; Togashi, K. Convolutional neural networks: An overview and application in radiology. *Insights Into Imaging* **2018**, *9*, 611–629. [CrossRef]
13. Zhang, C.; Xing, Y. CT artifact reduction via U-net CNN. In Proceedings of the Medical Imaging 2018: Image Processing, Houston, TX, USA, 11–13 February 2018; pp. 440–445.
14. Lan, L.; You, L.; Zhang, Z.; Fan, Z.; Zhao, W.; Zeng, N.; Chen, Y.; Zhou, X. Generative Adversarial Networks and Its Applications in Biomedical Informatics. *Front. Public Health* **2020**, *8*, 164. [CrossRef]
15. Dong, G.; Zhang, C.; Liang, X.; Deng, L.; Zhu, Y.; Zhu, X.; Zhou, X.; Song, L.; Zhao, X.; Xie, Y. A Deep Unsupervised Learning Model for Artifact Correction of Pelvis Cone-Beam CT. *Front. Oncol.* **2021**, *11*, 686875. [CrossRef]
16. Kim, S.; Ahn, J.; Kim, B.; Kim, C.; Baek, J. Convolutional neural network–based metal and streak artifacts reduction in dental CT images with sparse-view sampling scheme. *Med. Phys.* **2022**, *49*, 6253–6277. [CrossRef]
17. Hegazy, M.A.; Cho, M.H.; Cho, M.H.; Lee, S.Y. U-net based metal segmentation on projection domain for metal artifact reduction in dental CT. *Biomed. Eng. Lett.* **2019**, *9*, 375–385. [CrossRef]
18. Jiang, Y.; Yang, C.; Yang, P.; Hu, X.; Luo, C.; Xue, Y.; Xu, L.; Hu, X.; Zhang, L.; Wang, J. Scatter correction of cone-beam CT using a deep residual convolution neural network (DRCNN). *Phys. Med. Biol.* **2019**, *64*, 145003. [CrossRef]
19. Yang, F.; Zhang, D.; Zhang, H.; Huang, K.; Du, Y.; Teng, M. Streaking artifacts suppression for cone-beam computed tomography with the residual learning in neural network. *Neurocomputing* **2020**, *378*, 65–78. [CrossRef]
20. Huang, X.; Wang, J.; Tang, F.; Zhong, T.; Zhang, Y. Metal artifact reduction on cervical CT images by deep residual learning. *Biomed. Eng. OnLine* **2018**, *17*, 175. [CrossRef] [PubMed]
21. van der Heyden, B.; Roden, S.; Dok, R.; Nuyts, S.; Sterpin, E. Virtual monoenergetic micro-CT imaging in mice with artificial intelligence. *Sci. Rep.* **2022**, *12*, 2324. [CrossRef] [PubMed]

22. Zunair, H.; Ben Hamza, A. Sharp U-Net: Depthwise convolutional network for biomedical image segmentation. *Comput. Biol. Med.* **2021**, *136*, 104699. [CrossRef]
23. Rusanov, B.; Hassan, G.M.; Reynolds, M.; Sabet, M.; Kendrick, J.; Rowshanfarzad, P.; Ebert, M. Deep learning methods for enhancing cone-beam CT image quality toward adaptive radiation therapy: A systematic review. *Med. Phys.* **2022**, *49*, 6019–6054. [CrossRef]
24. Amirian, M.; Barco, D.; Herzig, I.; Schilling, F.-P. Artifact reduction in 3D and 4D cone-beam computed tomography images with deep learning: A review. *IEEE Access* **2024**, *12*, 10281–10295. [CrossRef]
25. Ko, Y.; Moon, S.; Baek, J.; Shim, H. Rigid and non-rigid motion artifact reduction in X-ray CT using attention module. *Med. Image Anal.* **2021**, *67*, 101883. [CrossRef]
26. Van Der Heyden, B.; Uray, M.; Fonseca, G.P.; Huber, P.; Us, D.; Messner, I.; Law, A.; Parii, A.; Reisz, N.; Rinaldi, I. A Monte Carlo based scatter removal method for non-isocentric cone-beam CT acquisitions using a deep convolutional autoencoder. *Phys. Med. Biol.* **2020**, *65*, 145002. [CrossRef]
27. Hyun, C.M.; Bayaraa, T.; Yun, H.S.; Jang, T.-J.; Park, H.S.; Seo, J.K. Deep learning method for reducing metal artifacts in dental cone-beam CT using supplementary information from intra-oral scan. *Phys. Med. Biol.* **2022**, *67*, 175007. [CrossRef] [PubMed]
28. Hansen, D.C.; Landry, G.; Kamp, F.; Li, M.; Belka, C.; Parodi, K.; Kurz, C. ScatterNet: A convolutional neural network for cone-beam CT intensity correction. *Med. Phys.* **2018**, *45*, 4916–4926. [CrossRef] [PubMed]
29. Kurosawa, T.; Nishio, T.; Moriya, S.; Tsuneda, M.; Karasawa, K. Feasibility of image quality improvement for high-speed CBCT imaging using deep convolutional neural network for image-guided radiotherapy in prostate cancer. *Phys. Medica* **2020**, *80*, 84–91. [CrossRef]
30. Ryu, K.; Lee, C.; Han, Y.; Pang, S.; Kim, Y.H.; Choi, C.; Jang, I.; Han, S.-S. Multi-planar 2.5 D U-Net for image quality enhancement of dental cone-beam CT. *PLoS ONE* **2023**, *18*, e0285608. [CrossRef]
31. Gottschalk, T.M.; Maier, A.; Kordon, F.; Kreher, B.W. DL-based inpainting for metal artifact reduction for cone beam CT using metal path length information. *Med. Phys.* **2023**, *50*, 128–141. [CrossRef]
32. Park, H.S.; Seo, J.K.; Hyun, C.M.; Lee, S.M.; Jeon, K. A fidelity-embedded learning for metal artifact reduction in dental CBCT. *Med. Phys.* **2022**, *49*, 5195–5205. [CrossRef]
33. Zhang, X.; Jiang, Y.; Luo, C.; Li, D.; Niu, T.; Yu, G. Image-based scatter correction for cone-beam CT using flip swin transformer U-shape network. *Med. Phys.* **2023**, *50*, 5002–5019. [CrossRef]
34. Rusanov, B.; Ebert, M.A.; Mukwada, G.; Hassan, G.M.; Sabet, M. A convolutional neural network for estimating cone-beam CT intensity deviations from virtual CT projections. *Phys. Med. Biol.* **2021**, *66*, 215007. [CrossRef] [PubMed]
35. Choi, D.; Kim, W.; Lee, J.; Han, M.; Baek, J.; Choi, J.-H. Integration of 2D iteration and a 3D CNN-based model for multi-type artifact suppression in C-arm cone-beam CT. *Mach. Vis. Appl.* **2021**, *32*, 116. [CrossRef]
36. Thies, M.; Zäch, J.-N.; Gao, C.; Taylor, R.; Navab, N.; Maier, A.; Unberath, M. A learning-based method for online adjustment of C-arm Cone-beam CT source trajectories for artifact avoidance. *Int. J. Comput. Assist. Radiol. Surg.* **2020**, *15*, 1787–1796. [CrossRef]
37. Ketcha, M.D.; Marrama, M.; Souza, A.; Uneri, A.; Wu, P.; Zhang, X.; Helm, P.A.; Siewerdsen, J.H. Sinogram+ image domain neural network approach for metal artifact reduction in low-dose cone-beam computed tomography. *J. Med. Imaging* **2021**, *8*, 052103. [CrossRef] [PubMed]
38. Zhuo, X.; Lu, Y.; Hua, Y.; Liu, H.; Zhang, Y.; Hao, S.; Wan, L.; Xie, Q.; Ji, X.; Chen, Y. Scatter correction for cone-beam CT via scatter kernel superposition-inspired convolutional neural network. *Phys. Med. Biol.* **2023**, *68*, 075011. [CrossRef]
39. Agrawal, H.; Hietanen, A.; Särkkä, S. Deep learning based projection domain metal segmentation for metal artifact reduction in cone beam computed tomography. *IEEE Access* **2023**, *11*, 100371–100382. [CrossRef]
40. Fan, F.; Ritschl, L.; Beister, M.; Biniazan, R.; Wagner, F.; Kreher, B.; Gottschalk, T.M.; Kappler, S.; Maier, A. Simulation-driven training of vision transformers enables metal artifact reduction of highly truncated CBCT scans. *Med. Phys.* **2024**, *51*, 3360–3375. [CrossRef] [PubMed]
41. Hu, D.; Zhang, Y.; Li, W.; Zhang, W.; Reddy, K.; Ding, Q.; Zhang, X.; Chen, Y.; Gao, H. SEA-Net: Structure-enhanced attention network for limited-angle CBCT reconstruction of clinical projection data. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 4507613. [CrossRef]
42. Jiang, C.; Lyu, T.; Ma, G.; Wu, Z.; Zhong, X.; Xi, Y.; Chen, Y.; Zhu, W. CBCT projection domain metal segmentation for metal artifact reduction using hessian-inspired dual-encoding network with guidance from segment anything model. *Med. Phys.* **2025**, *52*, 3900–3913. [CrossRef]
43. Piao, Z.; Deng, W.; Huang, S.; Lin, G.; Qin, P.; Li, X.; Wu, W.; Qi, M.; Zhou, L.; Li, B. Adaptive scatter kernel deconvolution modeling for cone-beam CT scatter correction via deep reinforcement learning. *Med. Phys.* **2024**, *51*, 1163–1177. [CrossRef] [PubMed]
44. Song, Y.; Yao, T.; Peng, S.; Zhu, M.; Meng, M.; Ma, J.; Zeng, D.; Huang, J.; Bian, Z.; Wang, Y. b-MAR: Bidirectional artifact representations learning framework for metal artifact reduction in dental CBCT. *Phys. Med. Biol.* **2024**, *69*, 145010. [CrossRef]
45. Tang, H.; Lin, Y.B.; Jiang, S.D.; Li, Y.; Li, T.; Bao, X.D. A new dental CBCT metal artifact reduction method based on a dual-domain processing framework. *Phys. Med. Biol.* **2023**, *68*, 175016. [CrossRef]

46. Wajer, R.; Wajer, A.; Kazimierczak, N.; Wilamowska, J.; Serafin, Z. The Impact of AI on Metal Artifacts in CBCT Oral Cavity Imaging. *Diagnostics* **2024**, *14*, 1280. [CrossRef]
47. Yang, S.; Wang, Z.; Chen, L.; Cheng, Y.; Wang, H.; Bai, X.; Cao, G. A dual-domain network with division residual connection and feature fusion for CBCT scatter correction. *Phys. Med. Biol.* **2025**, *70*, 045014. [CrossRef]
48. Yun, S.; Jeong, U.; Lee, D.; Kim, H.; Cho, S. Image quality improvement in bowtie-filter-equipped cone-beam CT using a dual-domain neural network. *Med. Phys.* **2023**, *50*, 7498–7512.
49. Angelone, F.; Ponsiglione, A.M.; Grassi, R.; Amato, F.; Sansone, M. A general framework for the assessment of scatter correction techniques in digital mammography. *Biomed. Signal Process. Control* **2024**, *89*, 105802. [CrossRef]
50. Sansone, M.; Ponsiglione, A.M.; Angelone, F.; Amato, F.; Grassi, R. Effect of X-ray scatter correction on the estimation of attenuation coefficient in mammography: A simulation study. In Proceedings of the 2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE), Rome, Italy, 26–28 October 2022; pp. 323–328.
51. Angelone, F.; Franco, A.; Ponsiglione, A.M.; Ricciardi, C.; Belfiore, M.P.; Gatta, G.; Grassi, R.; Sansone, M.; Amato, F. Assessment of an unsupervised denoising ap-proach based on Noise2Void in digital mammography. *Sci. Rep.* **2025**, *15*, 35712. [CrossRef] [PubMed]
52. Soltani, P.; Devlin, H.; Etemadi Sh, M.; Rengo, C.; Spagnuolo, G.; Baghaei, K. Do metal artifact reduction algorithms influence the detection of implant-related injuries to the inferior alveolar canal in CBCT images? *BMC Oral Health* **2024**, *24*, 268. [CrossRef]
53. Tang, X.; Krupinski, E.A.; Xie, H.; Stillman, A.E. On the data acquisition, image reconstruction, cone beam artifacts, and their suppression in axial MDCT and CBCT–A review. *Med. Phys.* **2018**, *45*, e761–e782. [CrossRef]
54. Lam, E.; Mallya, S. *White and Pharoah's Oral Radiology-E-BOOK: White and Pharoah's Oral Radiology-E-BOOK*; Elsevier Health Sciences: St. Louis, MI, USA, 2024.
55. Liu, J.; Zhang, T.; Kang, Y.; Qiang, J.; Hu, D.; Zhang, Y. SureUnet: Sparse autorepresentation encoder U-Net for noise artifact suppression in low-dose CT. *Neural Comput. Appl.* **2025**, *37*, 7561–7573. [CrossRef]
56. Li, Y.; Ma, C.; Li, Z.; Wang, Z.; Han, J.; Shan, H.; Liu, J. Semi-supervised spatial-frequency transformer for metal artifact reduction in maxillofacial CT and evaluation with intraoral scan. *Eur. J. Radiol.* **2025**, *187*, 112087. [CrossRef]
57. Zheng, S.; Zhang, D.; Yu, C.; Jia, L.; Zhu, L.; Huang, Z.; Zhu, D.; Yu, H. MAReraser: Metal Artifact Reduction with Image Prior Using CNN and Transformer Together. In Proceedings of the 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Lisbon, Portugal, 3–6 December 2024; pp. 4060–4065.
58. Wu, Z.; Zhong, X.; Lyu, T.; Xi, Y.; Ji, X.; Zhang, Y.; Xie, S.; Yu, H.; Chen, Y. PRAISE-Net: Deep Projection-domain Data-consistent Learning Network for CBCT Metal Artifact Reduction. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 4506113. [CrossRef]
59. Shi, J.; Bendig, D.; Vollmar, H.C.; Rasche, P. Mapping the bibliometrics landscape of AI in medicine: Methodological study. *J. Med. Internet Res.* **2023**, *25*, e45815. [CrossRef]
60. Kocak, B.; Baessler, B.; Cuocolo, R.; Mercaldo, N.; Pinto dos Santos, D. Trends and statistics of artificial intelligence and radiomics research in radiology, nuclear medicine, and medical imaging: Bibliometric analysis. *Eur. Radiol.* **2023**, *33*, 7542–7555. [CrossRef]
61. Venkatesh, K.; Santomartino, S.M.; Sulam, J.; Yi, P.H. Code and data sharing practices in the radiology artificial intelligence literature: A meta-research study. *Radiol. Artif. Intell.* **2022**, *4*, e220081. [CrossRef]
62. Tripathi, S.; Gabriel, K.; Dheer, S.; Parajuli, A.; Augustin, A.I.; Elahi, A.; Awan, O.; Dako, F. Understanding biases and disparities in radiology AI datasets: A review. *J. Am. Coll. Radiol.* **2023**, *20*, 836–841. [CrossRef]
63. Mudeng, V.; Kim, M.; Choe, S.-W. Prospects of structural similarity index for medical image analysis. *Appl. Sci.* **2022**, *12*, 3754. [CrossRef]
64. Balki, I.; Amirabadi, A.; Levman, J.; Martel, A.L.; Emersic, Z.; Meden, B.; Garcia-Pedrero, A.; Ramirez, S.C.; Kong, D.; Moody, A.R. Sample-size determination methodologies for machine learning in medical imaging research: A systematic review. *Can. Assoc. Radiol. J.* **2019**, *70*, 344–353. [CrossRef] [PubMed]
65. Obermeyer, Z.; Emanuel, E.J. Predicting the future—Big data, machine learning, and clinical medicine. *N. Engl. J. Med.* **2016**, *375*, 1216–1219. [CrossRef] [PubMed]
66. Tang, A.; Tam, R.; Cadrin-Chênevert, A.; Guest, W.; Chong, J.; Barfett, J.; Chepelev, L.; Cairns, R.; Mitchell, J.R.; Cicero, M.D. Canadian Association of Radiologists white paper on artificial intelligence in radiology. *Can. Assoc. Radiol. J.* **2018**, *69*, 120–135. [CrossRef] [PubMed]
67. Pace, M.; Cioffi, I.; D'antò, V.; Valletta, A.; Valletta, R.; Amato, M. Facial attractiveness of skeletal class I and class II malocclusion as perceived by laypeople, patients and clinicians. *Minerva Stomatol.* **2018**, *67*, 77–85. [CrossRef] [PubMed]