Co-MoCo: Collaborative Learning Enabled Motion Compensation for Weight-bearing CBCT

Tong Lin[1,2], Xi Wang[1,2], Yan Xi[3], Hao Tang[4], Yang Chen[1,2], Yikun Zhang[1,2]

[1] Laboratory of Image Science and Technology, School of Computer Science and Engineering, Southeast University, Nanjing, China
[2] Key Laboratory of New Generation Artificial Intelligence Technology and Its Interdisciplinary Applications, Southeast University, Nanjing, China
yikun@seu.edu.cn
[3] Shanghai First-Imaging Information Technology Co., Ltd., Shanghai, China
[4] Beijing Jishuitan Hospital, Capital Medical University, Beijing, China
pkuhao@163.com

Abstract. Weight-bearing cone-beam computed tomography (WB-CBCT) enables physiological load-bearing imaging but suffers from motion artifacts caused by involuntary patient tremors. Existing software-based approaches typically address image-domain restoration and motion estimation as separate problems, often leading to over-smoothed reconstructions or suboptimal convergence. To address these problems, we propose Co-MoCo, a collaborative motion compensation framework that integrates image-domain priors with physical consistency through an end-to-end differentiable pipeline. The framework employs a two-stage optimization: Stage I utilizes a Motion-Aware Dual-Decoder Network (MADN), trained via Optimal Transport Conditional Flow Matching (OT-CFM), to generate an artifact-free prior. In Stage II, this prior is integrated into an end-to-end differentiable correction loop to iteratively refine 6-degree-of-freedom (6-DoF) rigid motion trajectories. This synergistic design enables image restoration and motion estimation to mutually regularize each other. Extensive experiments on simulated datasets demonstrate that Co-MoCo significantly outperforms competing methods, achieving SSIM of 0.9849 under mild motion and 0.9152 under severe motion, with performance gains of up to 0.0528. Real-world validation on human and porcine hind limbs further confirms its superior robustness in artifact suppression and detail preservation for high-quality WB-CBCT imaging. Code is available at https://github.com/TomSHINING/CoMoCo.

Keywords: Weight-bearing CBCT · Image-domain priors · Motion artifact suppression · Differentiable pipeline.

1 Introduction

WB-CBCT [12, 16] is becoming more prevalent in orthopedic and musculoskeletal rehabilitation settings, as it enables imaging of the human body under physiological load conditions rather than in a supine position. This capability provides

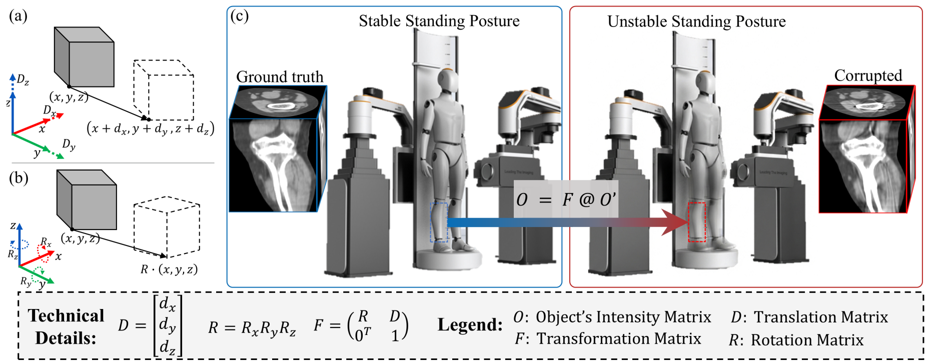

Fig. 1. Modeling of 6-DoF Motion. (a-b) define motion matrices (D, R), while (c) contrasts GT versus motion-corrupted images resulting from stable and unstable postures.

clinically relevant insights into patellofemoral tracking [2, 15], joint space narrowing [23, 4], and lower-limb alignment of the foot and ankle [13, 3], which are difficult to fully characterize using conventional CT.

However, in WB-CBCT, the temporal resolution is limited by factors such as detector frame rates and the motion speed of robotic arms or gantry components. Consequently, a typical WB-CBCT scan requires more time than multi-detector CT (MDCT) systems [12, 6, 9]. This increases the likelihood of patient motion, particularly for individuals with Parkinson's disease, nerve injuries, or muscular fatigue during standing, potentially leading to motion artifacts.

A straightforward approach to mitigating motion artifacts is to shorten the scanning time, thereby reducing the likelihood of motion occurrence. However, this is typically achieved at the cost of signal-to-noise ratio (SNR). Consequently, researchers have shifted their focus toward software-based solutions. Existing software-based methods can be categorized into two classes: image-domain-based restoration, and motion estimation-based compensation.

Image-domain methods learn a mapping that captures artifact statistics and suppresses them directly in the image [10, 21]. Nevertheless, image-domain restoration under severe motion in WB-CBCT is inherently ill-posed, as a single corrupted voxel may correspond to multiple plausible targets. Pixel-wise losses, such as L_1 or L_2, tend to drive the network toward a statistical average or median, resulting in an averaging effect [14, 11] that suppresses fine details and leads to over-smoothing. The second category focuses on estimating patient motion and compensating for it during image reconstruction. Periodic motion modeling methods [5, 22] in cardiopulmonary CT infer motion directly from projection data but are largely restricted to strongly periodic motion patterns, making them unsuitable for the random, involuntary tremors encountered in WB-CBCT. Iterative optimization-based methods jointly update motion parameters or deformation fields through image registration until convergence, but they suffer from high computational cost due to repeated forward projections [20, 28]. To reduce this burden, some studies use image metrics to evaluate quality of re-

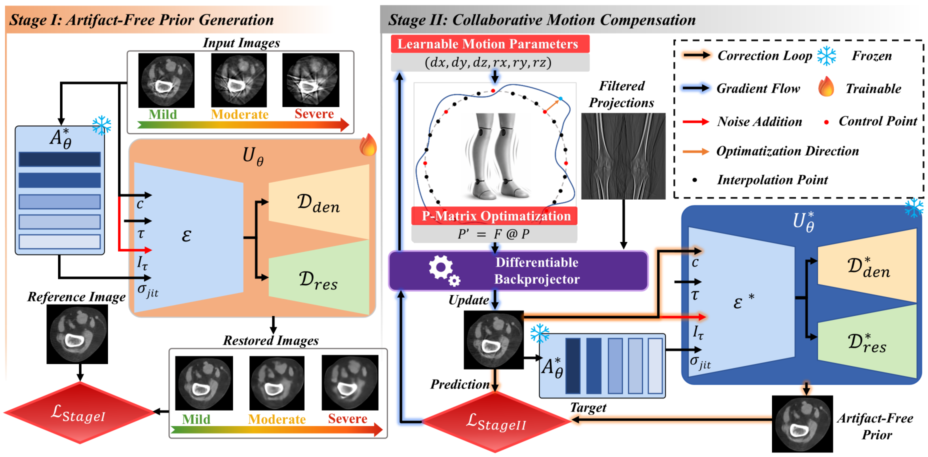

Fig. 2. Overview of the Co-MoCo framework. Stage I trains the MADN for artifact removal across varying motion severities; Stage II optimizes learnable motion parameters using the artifact-free prior through an end-to-end differentiable pipeline.

sults, such as TV or VIF [6, 25]. However, large inter-dataset variability in data characteristics hinders the reliable application of image quality metrics.

To overcome the limitations of isolated optimization of image-domain restoration or motion parameters estimation, which often yield suboptimal solutions, we propose Co-MoCo, a collaborative framework that allows image-domain restoration and parameter optimization tasks to mutually enhance each other.

2 Methodology

Co-MoCo collaboratively integrates an image-domain network U_θ and 3D motion parameters ϕ estimation. Specifically, a Motion-Aware Dual-Decoder Network (MADN) is designed to provide prior for motion estimation, while the motion-estimation model refines the 6-DoF motion parameters $\phi = (d_x, d_y, d_z, r_x, r_y, r_z)$ via an end-to-end differentiable pipeline. To build the relationship between estimated motion and the system geometry, the current projection matrix P_ϕ is formulated as the product of the initial projection matrix (P-Matrix) P_{init} and the motion-induced rigid transformation matrix F_ϕ:

$$P_\phi = P_{init} \cdot F_\phi, \tag{1}$$

where F_ϕ is the 4×4 homogeneous matrix constructed from the refined translation and rotation parameters (as illustrated in Fig. 1).

2.1 Collaborative Motion Compensation Paradigm

Inspired by the Regularization by Denoising (RED) [17], which posits that a high-performance denoiser can serve as an implicit regularizer, we formulate the optimization as a fixed-point consistency problem [18, 1]. Instead of constrainting by projections, we minimize the discrepancy between the motion-corrected reconstruction and its refined counterpart output from the prior network U_θ:

$$\phi^* = \arg\min_{\phi} \left\| \underbrace{\mathcal{R}(y; P_\phi)}_{\text{Physical Consistency}} - \underbrace{U_\theta(\mathcal{R}(y; P_\phi))}_{\text{Prior Regularizer}} \right\|_2^2 . \tag{2}$$

In this formulation, U_θ acts as a regularizer that defines the "artifact-free" manifold. By minimizing this objective, we iteratively update ϕ such that the resulting reconstruction $\mathcal{R}(y; P_\phi)$ aligns with the structural prior learned by the network.

2.2 Motion-Aware Dual-Decoder Network

To address varying severities of motion, we propose a MADN. Given that the reconstructed object is a 3D volume $V \in \mathbb{R}^{H \times W \times D}$, the network is designed to process data in a slice-wise manner. Specifically, we define a probability path following the OT-CFM framework [26]:

$$I_\tau = \tau I_{gt} + (1 - \tau)\epsilon, \tag{3}$$

where $\epsilon \sim \mathcal{N}(0, \mathbf{I})$ is Gaussian noise and $\tau \in [0, 1]$. The network inputs include the noisy image I_τ, the corrupted image c, and a motion intensity map σ_{jit}. Here, $\sigma_{jit} \in [0, 1]$ is derived from a pre-trained discriminator that quantifies the severity of motion. The resulting multi-scale features $\mathcal{F}$ are processed by two parallel branches: the Denoising Decoder $\mathcal{D}_{den}$ and the Restoration Decoder $\mathcal{D}_{res}$. Adopting a global residual learning strategy, the final outputs are formulated as:

$$\begin{cases} \hat{I}_{den} = I_\tau + \mathcal{D}_{den}(\mathcal{F}, \mathcal{F}_{skip}, \tau) \\ \hat{I}_{res} = c + \mathcal{D}_{res}(\mathcal{F}, \mathcal{F}_{skip}, \tau), \end{cases} \tag{4}$$

where $\mathcal{F}_{skip}$ denotes hierarchical features transferred via skip connections. To align with the OT-CFM framework, we optimize the network to predict I_{gt} directly, which is equivalent to matching the ideal vector field $v_t(I_\tau) = I_{gt} - \epsilon$:

$$\mathcal{L}_{stageI} = \mathbb{E}_{I_{gt}, \epsilon, c, \tau} \left[\|\hat{I}_{den} - I_{gt}\|_2^2 + \|\hat{I}_{res} - I_{gt}\|_2^2 \right] . \tag{5}$$

In the subsequent motion correction loop, $\hat{I}_{res}$ serve as the refined outputs of the MADN prior U_θ to guide the iterative optimization of motion parameters.

2.3 Differentiable Motion Matrix Estimation

In accordance with prior work [24, 10, 7, 8], the object's motion is parameterized as a time-varying rigid transformation, where motion parameters sets ϕ_i

are optimized at discrete intervals t_i (as illustrated in Fig. 2). To ensure temporal smoothness, a differentiable Gaussian-weighted interpolation is employed to derive a continuous trajectory.

For projection view k, we update the P-Matrix P_ϕ^k to $P_{init}^k F_\phi^k$. Equivalently, for any 3D point $\mathbf{X} \in \mathbb{R}^3$ in space, its geometric projection $\mathcal{S}(P_\phi^k, \mathbf{X})$ onto the detector plane is derived via homogeneous coordinates:

$$w \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = P_{init}^k F_\phi^k \begin{pmatrix} \mathbf{X} \\ 1 \end{pmatrix}, \quad \text{where } \mathcal{S}(P_\phi^k, \mathbf{X}) = (u, v). \tag{6}$$

Here, (u, v) denote the 2D pixel coordinates on the detector plane, and $w \in \mathbb{R}$ represents the homogeneous scaling factor required to map the coordinates back to the Euclidean space.

Based on the updated P-Matrix P_ϕ, volume V_ϕ can be reconstructed using a differentiable cone-beam backprojection operator. The projections y_k are preprocessed via cosine weighting and ramp filtering before backprojection. The reconstruction process is formalized as:

$$V_\phi = \mathcal{R}(y; P_\phi) = \sum_{k=1}^{N} \text{Interp}\left[\tilde{y}_k, \mathcal{S}(P_\phi^k, \mathbf{X})\right], \tag{7}$$

where $\tilde{y}_k$ represents the filtered projection data, $\text{Interp}[\cdot, \cdot]$ denotes the bilinear sampling operation in projections, which extracts the pixel value from $\tilde{y}_k$ at the projected coordinates determined by $\mathcal{S}$.

2.4 Optimization Pipeline

Algorithm 1: Optimization Pipeline of Co-MoCo

Input : Initial motion parameters ϕ_0, Raw projection data y, Frozen MADN U_θ
Output: Optimized motion-free volume V^*
Initialize $\phi \leftarrow \phi_0$;
while *not converged* **do**
    `// Motion compensation and reconstruction`
    $V_\phi \leftarrow \mathcal{R}(y; P_\phi)$;
    `// Generate high-quality prior using MADN` U_θ
    $\tilde{V} \leftarrow U_\theta(V_\phi)$;
    `// Compute StageII loss and gradient`
    $\mathcal{L}_{\text{StageII}} \leftarrow \|V_\phi - \tilde{V}\|^2$, $\mathbf{g} = \nabla_\phi \mathcal{L}_{StageII}$;
    `// Update motion parameters via gradient`
    $\phi \leftarrow \text{Adam}(\phi, \mathbf{g})$;
end
return $V^* = V_\phi$;

As illustrated in Fig. 2, proposed optimization pipeline follows an end-to-end differentiable correction loop: the current ϕ is first used to generate a motion-compensated volume V_ϕ, which is then passed through the frozen MADN to produce a artifacts-free prior $\tilde{V}$. By minimizing the $\mathcal{L}_{\text{StageII}}$ between the compensated volume and prior volume, gradients $\mathbf{g}$ are backpropagated through the end-to-end differentiable pipeline to update ϕ via the Adam optimizer.

3 Experiments

3.1 Datasets

We evaluated Co-MoCo on 140 clinical lower-limb CT scans, which were randomly divided into training, validation, and test sets with a ratio of 7:1:2. To enhance computational efficiency, motion-corrupted data were generated by applying rigid transformations directly to projection matrices (Eq. 6) rather than to 3D volumes. Simulations and reconstructions were implemented via the torch-mando toolkit [19] and a P-Matrix-based differentiable cone-beam backprojector [24]. Real-world validation was conducted using the First-Imaging Huashan Robotic-arm upright WB-CBCT system developed by First-Imaging Medical Equipment Co., Ltd., involving both human volunteers and porcine hind limbs from Beijing Jishuitan Hospital.

3.2 Experimental Setup

Experiments were conducted on an NVIDIA RTX 4090 GPU using PyTorch, with all methods sharing identical configurations. The MADN was trained for 500 epochs (batch size 8) using AdamW with an initial learning rate of 1×10^{-4} and an exponential decay ($\gamma = 0.95$ every 5 epochs) under $\mathcal{L}_{\text{StageI}}$. For joint optimization, 3D motion was modeled via 40 learnable control points over 100 iterations. This stage utilized independent SGD optimizers for rotation ($lr = 0.04$) and translation ($lr = 0.15$), regulated by a ReduceLROnPlateau scheduler (factor 0.5, patience 3) under $\mathcal{L}_{\text{StageII}}$.

3.3 Comparison Study

Co-MoCo was evaluated against five baseline methods, including image-domain network (IDN) and motion compensation (MC) methods, under mild and severe motion simulation datasets. Quantitative results in Table 1 show that Co-MoCo outperforms all methods. In mild motion cases, Co-MoCo achieves a PSNR of 40.95 dB, surpassing the Theies's by 1.21 dB. Under severe motion, where image-domain methods like Uformer fail to restore structural integrity, Co-MoCo maintains high fidelity by synergistically integrating image priors with MC. Visual results in Fig. 3 and absolute difference maps further confirm that while other methods leave substantial residual artifacts, Co-MoCo suppresses motion-induced artifacts while preserving high-frequency details.

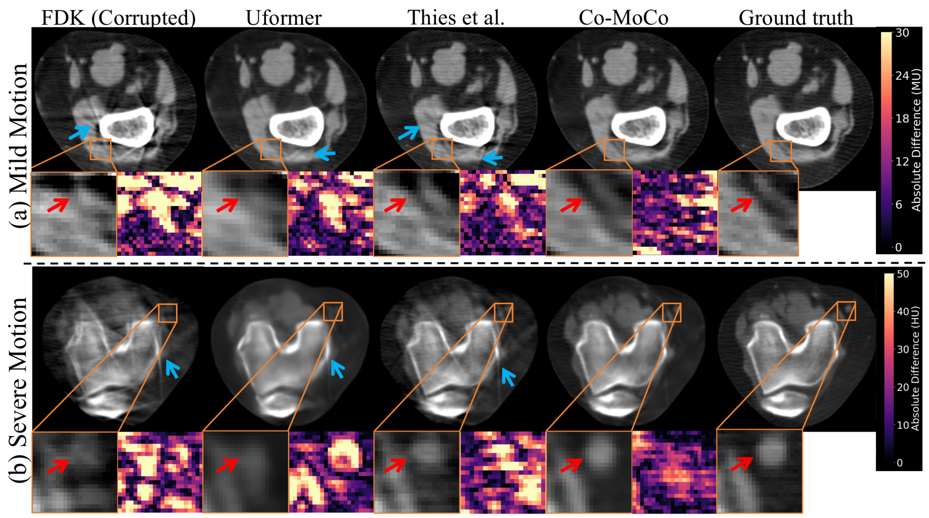

Fig. 3. Qualitative comparison of motion artifact removal methods across varying severity of simulated motion. (a) Mild motion (rotation: [0.3, 0.6]°, translation: [0.5, 1] mm); (b) Severe motion (rotation: [3, 5]°, translation: [4, 6] mm). Red arrows indicate residual blurring and blue arrows indicate residual motion artifacts

3.4 Ablation Study

An ablation study was conducted to analyze the contributions of the image domain network (IDN) and the motion compensation module (MC). As summarized in Table 2, introducing either IDN or MC independently yields significant performance gains in the mild motion dataset, boosting PSNR from 34.01 dB to over 40 dB. However, independent modules struggle under severe motion, yielding only marginal gains (e.g., +0.38 dB PSNR for IDN). In contrast, proposed Co-MoCo, which leverages the joint iterative optimization of IDN and MC,

Table 1. Quantitative performance comparison of various methods under mild and severe motion datasets. Pink and blue backgrounds highlight the best and second-best results, respectively.

Type	Methods	(a) Mild Motion			(b) Severe Motion		
		MSE ↓	SSIM ↑	PSNR ↑	MSE ↓	SSIM ↑	PSNR ↑
-	FDK (Moved)	1585.83	0.9490	34.01	9402.92	0.7892	26.28
IDN	ResUnet [29]	772.33	0.9681	37.15	8843.30	0.8599	26.55
	Uformer [27]	603.16	0.9738	38.22	9120.34	0.8624	26.42
	Ko et al. [10]	688.25	0.9618	37.65	8621.61	0.8391	26.65
	Su et al. [21]	621.46	0.9707	38.08	8402.99	0.8610	26.77
MC	Thies's [25]	424.51	0.9631	39.74	3591.95	0.8568	30.46
IDN&MC	**Co-MoCo**	**321.62**	**0.9849**	**40.95**	**2288.21**	**0.9152**	**32.45**

Table 2. Ablation study of different modules on two datasets.

IDN	MC	(a) Mild Motion			(b) Severe Motion		
		MSE ↓	SSIM ↑	PSNR ↑	MSE ↓	SSIM ↑	PSNR ↑
×	×	1585.83	0.9490	34.01	9402.92	0.7892	26.28
✓	×	363.27	0.9774	40.41	8434.33	0.8778	26.66
×	✓	381.99	0.9633	40.20	7157.66	0.8225	27.47
✓	✓	**321.62**	**0.9849**	**40.95**	**2288.21**	**0.9152**	**32.45**

achieves a dramatic performance leap. Specifically, it raises the PSNR from 26.28 dB to 32.45 dB in the severe motion dataset, alongside substantial improvements in MSE and SSIM. Above results confirm that coupling image priors and motion compensation is indispensable for handling complex motion degradations effectively.

3.5 Real-World Data Analysis

The performance of Co-MoCo in suppressing motion artifacts was further validated using the First-Imaging Huashan Robotic-arm WB-CBCT system deployed at Beijing Jishuitan Hospital. In a clinical case involving a patient with ankylosing spondylitis who was unable to remain stationary (as shown in Fig. 4 (f)), Co-MoCo successfully suppressed motion artifacts (highlighted by red arrows in Fig. 4 (a) and (c)) and restored soft-tissue details (highlighted by yellow arrows in Fig. 4 (b) and (d)). Furthermore, we conducted a study using a porcine limbs. By manually inducing motion during the scan via a rigid rod connected

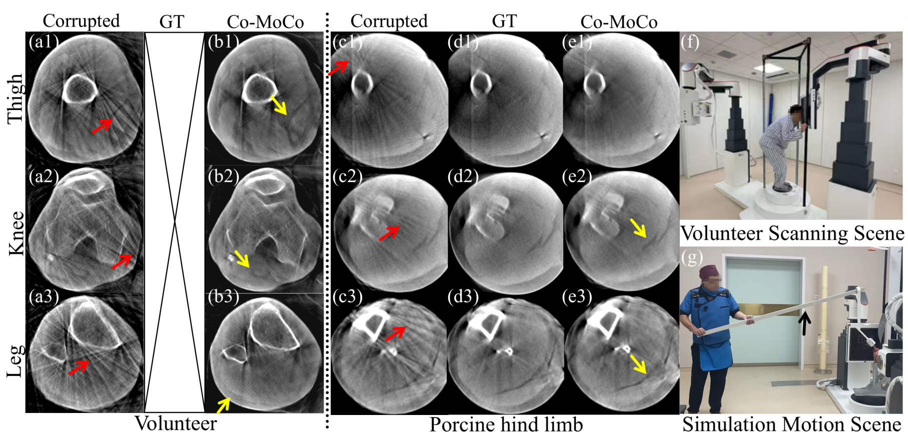

Fig. 4. Qualitative results on real-world datasets across thigh, knee, and leg. (a-b) Evaluation on a human volunteer. (c-e) Evaluation on a porcine hind limbs with reference. (f) Experimental setups for volunteer scanning and (g) Simulated motion datasets.

via bearings (highlighted by black arrow in Fig. 4 (g)), we demonstrated that Co-MoCo effectively suppressed the resulting artifacts, yielding image quality close to the fixed image (as shown in Fig. 4 (d) and (e)).

4 Conclusion

In this study, we proposed Co-MoCo, a collaborative motion compensation framework that synergistically integrates image-domain restoration with 6-DoF motion estimation through an end-to-end differentiable reconstruction pipeline. By establishing a mutually regularizing correction between a MADN and geometric optimization, Co-MoCo overcomes the limitations of existing methods, ensuring both physical consistency and motion artifact suppression. Extensive experiments further validated Co-MoCo effectively suppresses the motion artifacts. While the current 6-DoF rigid assumption performs robustly for bone-dominant motion, future work will explore non-rigid deformation modeling to address complex joint articulations. In general, Co-MoCo offers a promising solution for orthopedic and musculoskeletal diagnostics in WB-CBCT.

Acknowledgments. This work was supported by the International (Regional) Cooperation and Exchange Program of the National Natural Science Foundation of China (Grant 62561160154), the National Natural Science Foundation of China (Grants T2225-025, 62401134), the Zhishan Young Scholars' Program of Southeast University for Yikun Zhang and the Big Data Computing Center of Southeast University.

Disclosure of Interests. The authors have no competing interests to declare.

References

1. Bai, S., Kolter, J.Z., Koltun, V.: Deep equilibrium models. Advances in neural information processing systems **32** (2019)
2. Chen, Y., Yu, F., Rong, F., Lv, F., Lv, F., Li, J.: Analysis of spatial patellofemoral alignment using novel three-dimensional measurements based on weight-bearing cone-beam ct. Insights into Imaging **16**(1), 1 (2025)
3. Faict, S., Burssens, A., Van Oevelen, A., Maeckelbergh, L., Mertens, P., Buedts, K.: Correction of ankle varus deformity using patient-specific dome-shaped osteotomy guides designed on weight-bearing ct: a pilot study. Archives of Orthopaedic and Trauma Surgery **143**(2), 791–799 (2023)
4. Fritz, B., Fritz, J., Fucentese, S., Pfirrmann, C., Sutter, R.: Three-dimensional analysis for quantification of knee joint space width with weight-bearing ct: comparison with non-weight-bearing ct and weight-bearing radiography. Osteoarthritis and Cartilage **30**(5), 671–680 (2022)
5. Gong, H., Ahmed, Z., Chang, S., Koons, E.K., Thorne, J.E., Rajiah, P., Foley, T.A., Fletcher, J.G., McCollough, C.H., Leng, S.: Motion artifact correction in cardiac ct using cross-phase temporospatial information and synergistic attention gate and spatial transformer sub-networks. Physics in Medicine & Biology **69**(3), 035023 (2024)

6. Huang, H., Siewerdsen, J.H., Zbijewski, W., Weiss, C.R., Unberath, M., Ehtiati, T., Sisniega, A.: Reference-free learning-based similarity metric for motion compensation in cone-beam ct. Physics in medicine & biology **67**(12), 125020 (2022)
7. Jang, S., Kim, S., Kim, M., Ra, J.B.: Head motion correction based on filtered backprojection for x-ray ct imaging. Medical physics **45**(2), 589–604 (2018)
8. Jang, S., Kim, S., Kim, M., Son, K., Lee, K.Y., Ra, J.B.: Head motion correction based on filtered backprojection in helical ct scanning. IEEE transactions on medical imaging **39**(5), 1636–1645 (2019)
9. Kalra, M.K., Maher, M.M., D'Souza, R., Saini, S.: Multidetector computed tomography technology: current status and emerging developments. Journal of computer assisted tomography **28**, S2–S6 (2004)
10. Ko, Y., Moon, S., Baek, J., Shim, H.: Rigid and non-rigid motion artifact reduction in x-ray ct using attention module. Medical Image Analysis **67**, 101883 (2021)
11. Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al.: Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4681–4690 (2017)
12. Lin, T., Lyu, T., Wang, J., Wu, Z., Xi, Y., Jean-Louis, D., Zhu, W., Tang, H., Xie, S., Chen, Y.: Helical-like scan and upright cbct imaging algorithms based on robotic-arm system. Medical Physics **52**(7), e17997 (2025)
13. Lintz, F., de Cesar Netto, C., Barg, A., Burssens, A., Richter, M., Group, W.B.C.I.S., et al.: Weight-bearing cone beam ct scans in the foot and ankle. EFORT open reviews **3**(5), 278–286 (2018)
14. Lugmayr, A., Danelljan, M., Van Gool, L., Timofte, R.: Srflow: Learning the super-resolution space with normalizing flow. In: European conference on computer vision. pp. 715–732. Springer (2020)
15. Lullini, G., Belvedere, C., Busacca, M., Moio, A., Leardini, A., Caravelli, S., Maccaferri, B., Durante, S., Zaffagnini, S., Marcheggiani Muccioli, G.M.: Weight bearing versus conventional ct for the measurement of patellar alignment and stability in patients after surgical treatment for patellar recurrent dislocation. La radiologia medica **126**(6), 869–877 (2021)
16. Rahim, S., Korte, J., Hardcastle, N., Hegarty, S., Kron, T., Everitt, S.: Upright radiation therapy—a historical reflection and opportunities for future applications. Frontiers in Oncology **10**, 213 (2020)
17. Romano, Y., Elad, M., Milanfar, P.: The little engine that could: Regularization by denoising (red). SIAM journal on imaging sciences **10**(4), 1804–1844 (2017)
18. Ryu, E., Liu, J., Wang, S., Chen, X., Wang, Z., Yin, W.: Plug-and-play methods provably converge with properly trained denoisers. In: International Conference on Machine Learning. pp. 5546–5557. PMLR (2019)
19. SEU-CT-Recon: torch-mando: Differentiable domain transform in pytorch based on mandoct, https://github.com/SEU-CT-Recon/torch-mando
20. Shao, H.C., Mengke, T., Pan, T., Zhang, Y.: Dynamic cbct imaging using prior model-free spatiotemporal implicit neural representation (pmf-stinr). Physics in Medicine & Biology **69**(11), 115030 (2024)
21. Su, B., Wen, Y., Liu, Y., Liao, S., Fu, J., Quan, G., Li, Z.: A deep learning method for eliminating head motion artifacts in computed tomography. Medical Physics **49**(1), 411–419 (2022)
22. Tang, Q., Cammin, J., Srivastava, S., Taguchi, K.: A fully four-dimensional, iterative motion estimation and compensation method for cardiac ct. Medical physics **39**(7Part1), 4291–4305 (2012)

23. Thawait, G.K., Demehri, S., AlMuhit, A., Zbijweski, W., Yorkston, J., Del Grande, F., Zikria, B., Carrino, J.A., Siewerdsen, J.H.: Extremity cone-beam ct for evaluation of medial tibiofemoral osteoarthritis: initial experience in imaging of the weight-bearing and non-weight-bearing knee. European journal of radiology **84**(12), 2564–2570 (2015)
24. Thies, M., Wagner, F., Maul, N., Folle, L., Meier, M., Rohleder, M., Schneider, L.S., Pfaff, L., Gu, M., Utz, J., et al.: Gradient-based geometry learning for fan-beam ct reconstruction. Physics in Medicine & Biology **68**(20), 205004 (2023)
25. Thies, M., Wagner, F., Maul, N., Yu, H., Goldmann, M., Schneider, L.S., Gu, M., Mei, S., Folle, L., Preuhs, A., et al.: A gradient-based approach to fast and accurate head motion compensation in cone-beam ct. IEEE Transactions on Medical Imaging **44**(2), 1098–1109 (2024)
26. Tong, A., Malkin, N., Huguet, G., Zhang, Y., Rector-Brooks, J., Fatras, K., Wolf, G., Bengio, Y.: Conditional flow matching: Simulation-free dynamic optimal transport. arXiv preprint arXiv:2302.00482 **2**(3) (2023)
27. Wang, Z., Cun, X., Bao, J., Zhou, W., Liu, J., Li, H.: Uformer: A general u-shaped transformer for image restoration. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 17683–17693 (2022)
28. Xie, J., Shao, H.C., Zhang, Y.: Time-resolved dynamic cbct reconstruction using prior-model-free spatiotemporal gaussian representation (pmf-stgr). Physics in Medicine & Biology **70**(16), 165011 (2025)
29. Zhang, Z., Liu, Q., Wang, Y.: Road extraction by deep residual u-net. IEEE Geoscience and Remote Sensing Letters **15**(5), 749–753 (2018)