FlowReg: Reconstruction-Guided Flow Matching for Universal Spinal CT/X-ray Registration

Ao Shen[1], Junfeng Jiang[2] ✉, Xueming Fu[3], Qiang Zeng[4], Ye Tang[1], Zhengming Chen[1], Luming Nong[5], Feng Wang[6], and Shaohua Kevin Zhou[3] ✉

[1] College of Information Science and Engineering, Hohai University (HHU), Changzhou Jiangsu, 213200, China
[2] College of Artificial Intelligence and Automation, HHU, Changzhou Jiangsu, 213200, China
jiangjf.hhu@gmail.com
[3] School of Biomedical Engineering, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei Anhui, 230026, China
skevinzhou@ustc.edu.cn
[4] Wu Jieh Yee School of Interdisciplinary Studies, Lingnan University, Hong Kong SAR,China
[5] The Third Affiliated Hospital of Nanjing Medical University, Changzhou Jiangsu, 213164, China
[6] Tuodao Medical Technology Co., Ltd., Nanjing Jiangsu, 210012, China

Abstract. Precise and efficient spinal Computed Tomography (CT)/X-ray registration remains a long-standing challenge for image-guided navigation. Methodologies have evolved from intensity-based optimization to learning-based acceleration, transitioning from the traditional "render and compare" paradigm toward "reconstruction and compare" strategies. However, current learning-based methods still face difficult convergence and limited generalization, and a versatile framework capable of accommodating both single- and multi-vertebrae registration is yet to be established. To resolve these issues, we propose FlowReg, a versatile end-to-end framework that leverages CT volume reconstruction from intraoperative biplanar X-rays as a robust spatial prior to guide precise registration. FlowReg utilizes a $SE(3)$-aware flow matching branch to learn deterministic pose gradients in $\mathfrak{se}(3)$ while learning shared representations across vertebrae. To bridge the domain gap, we introduce a Difference-Aware Feature Fusion module to reconcile cross-modal disparities. Extensive experiments on synthetic and real clinical datasets demonstrate that FlowReg achieves state-of-the-art performance in both reconstruction quality and registration accuracy, successfully satisfying real-time intraoperative clinical demands. The code is available at: github.com/shenao1995/FlowReg.

Keywords: CT/X-ray Registration · 3D Reconstruction · Flow Matching · Multi-object Pose Estimation.

1 Introduction

Image-guided navigation is critical for minimally invasive spinal surgery, which is characterized by its stringent requirements for accuracy and real-time performance. [14]. Over the past two decades, numerous methods have been devised to address these critical clinical demands. Traditionally, intensity-based registration has achieved high precision but often depends on accurate anatomical segmentation or manual landmark identification [6,11,4]. Moreover, these optimization-driven approaches suffer from low efficiency, hindering real-time surgical workflows. To overcome these bottlenecks, learning-based methods have gained prominence by accelerating the registration process [10,23,8]. However, most existing learning-based methods employ direct parameter regression, which relies heavily on favorable initialization and is prone to local minima [3], necessitating frequent fine-tuning (Fig. 1(a)) [23]. In computer vision, generative diffusion models address these ambiguities but exhibit difficult convergence and high training costs (Fig. 1(b))[25,24]. As a cutting-edge learning strategy, flow matching fundamentally resolves these challenges by learning deterministic pose trajectories (Fig. 1(c)), providing the efficiency required for real-time clinical applications [12,19]. Despite its potential, flow matching remains unexplored in Computed Tomography (CT)/X-ray registration tasks. Furthermore, traditional registration methods generally follow the "render and compare" paradigm [23,4,3]. These methods optimize poses by comparing Digitally Reconstructed Radiographs (DRRs) with target X-rays, though 2D projection inherently entails 3D information loss and cross-modal domain gaps. While "reconstruction and compare" strategies have been explored to mitigate these issues, their application is currently restricted to single-vertebra or single-object registration [17,8]. This limitation is critical because simultaneous multi-vertebrae registration is vital for multi-level spinal interventions like short-segment fixation [1]. Nevertheless, we observe that current multi-vertebrae registration typically relies on global spinal-to-spinal alignment [8] or level-by-level processing [6], which either overlooks relative inter-vertebral motion or remains computationally sluggish.

To resolve above-mentioned challenges, we propose **FlowReg**, a versatile **Flow** Matching **Reg**istration method for spinal CT/X-ray registration capable of accommodating both single- and multi-vertebrae scenarios. We introduce a Registration Flow Matching (RFM) branch that learns the directional gradients of pose trajectories while implicitly modeling inter-vertebral pose correlations. Concurrently, a parallel reconstruction branch enables biplanar X-rays vertebral recovery to serve as a robust spatial prior for guiding precise registration. Furthermore, we propose a Difference-Aware Feature Fusion (DAFF) module to effectively reconcile the disparities between DRRs and X-rays. Ultimately, FlowReg establishes a generalized framework that achieves high-precision registration through synergistic volume reconstruction.

Our primary contributions are summarized as follows: (1) FlowReg, a versatile end-to-end framework for single- and multi-vertebrae registration that leverages 3D volume reconstruction as a robust spatial prior to guide precise registration; (2) an RFM branch that learns deterministic pose gradients in $\mathfrak{se}(3)$ with

SE(3)-aware geometric supervision while learning shared representations across vertebrae; (3) a DAFF module used for cross-modal feature alignment between DRRs and X-rays; and (4) a sim-to-real training strategy and extensive evaluation demonstrating that FlowReg achieves state-of-the-art performance in both reconstruction quality and registration accuracy.

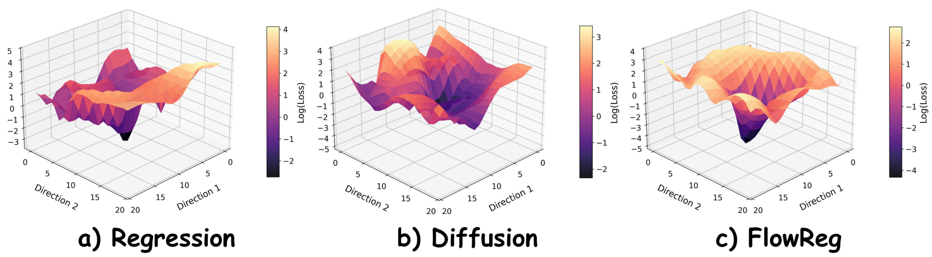

Fig. 1. Comparison of pose learning trajectories. (a) Direct regression: Unstable, jittery paths prone to local minima. (b) Diffusion-based methods: Slower convergence due to stochastic sampling. (c) FlowReg (Ours): Smooth, deterministic convergence in $se(3)$ with $SE(3)$-consistent outputs.

2 Method

We propose FlowReg, a versatile end-to-end framework leveraging 3D reconstruction as a spatial prior for precise single- and multi-vertebrae registration. As in Fig. 2(a), intraoperative Anterior–Posterior (AP)/LATeral (LAT) X-rays and preoperative CTs are cropped via vertebra-level detection [7] and semantic segmentation [13] of all vertebrae respectively. The initial DRRs are then generated via Siddon's ray-tracing using camera extrinsic parameters [18]. FlowReg integrates a pseudo-Siamese ResNet-50 for feature extraction, a DAFF module for cross-modal alignment, and a 3D reconstruction branch that recovers density volumes from X-rays to provide spatial priors guiding RFM pose estimation. It outputs semantically indexed pairs $(\boldsymbol{V}_i, \boldsymbol{T}_i)$, where reconstructed volumes $\{\boldsymbol{V}_i\}_{i=1}^{N}$ and 6-DoF poses $\{\boldsymbol{T}_i\}_{i=1}^{N}$ correspond to each input vertebra ($N \leq 3$), ensuring level-specific registration for short-segment fixation [1]. During inference (Fig. 2(b)), the target vertebrae are localized, enhanced with CLAHE [15], and processed for joint reconstruction and pose prediction.

Difference-Aware Feature Fusion Module To effectively bridge the domain gap between target X-rays and initial DRRs, we introduce a DAFF module designed to emphasize cross-modal disparities and commonalities. Specifically, as illustrated in Fig. 2(c), the module explicitly computes the difference and element-wise product of target (f_{tar}) and moving (f_{mov}) features. These maps, along with

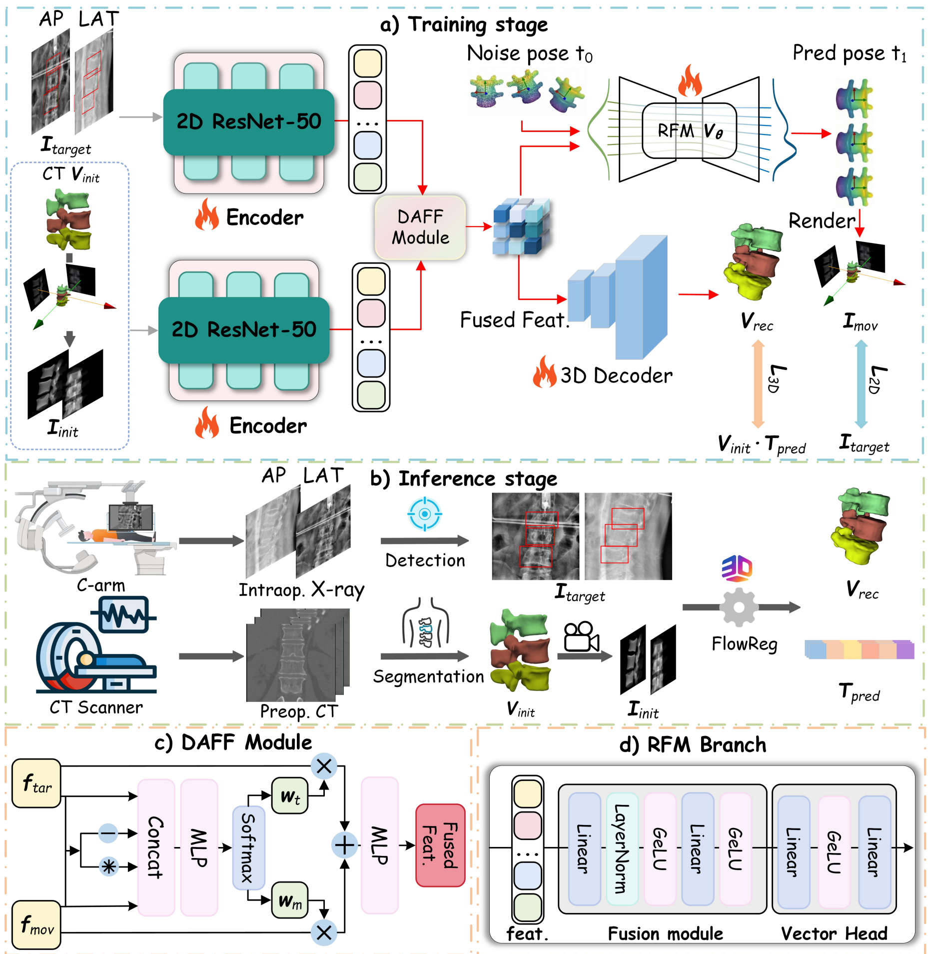

Fig. 2. The framework of FlowReg (a-b). (c) DAFF module and (d) RFM branch.

the original features, are concatenated and processed by a Multi-Layer Perceptron (MLP) to capture global misalignments. A weight generator then applies attention-based Softmax coefficients to produce a dynamically weighted fused feature. This mechanism enables the network to prioritize informative cross-modal features, enhancing robustness in reconstruction and registration under low-overlap scenarios.

Registration Flow Matching Branch We introduce an RFM branch to learn deterministic pose trajectories while implicitly capturing inter-vertebral pose relationships. Specifically, as shown in Fig. 2(d), we formulate the registration as a flow in the Lie algebra $\mathfrak{se}(3)$ with $SE(3)$-aware geometric constraints. We define the probability path $\boldsymbol{T}_t = (1-t)\boldsymbol{T}_0 + t\boldsymbol{T}_1$, where $t \sim \mathcal{U}[0,1]$, $\boldsymbol{T}_0$ is the initial noise, and $\boldsymbol{T}_1$ is the Ground-Truth (GT) pose in Lie algebra coordinates. Under this parameterization, the trajectory corresponds to a constant-velocity interpolation in the tangent space, with velocity given by $(\boldsymbol{T}_1 - \boldsymbol{T}_0)$. The network $\boldsymbol{v}_\theta$ encodes the time step t, the current pose $\boldsymbol{T}_t \in \mathbb{R}^{N\times 6}$ ($\boldsymbol{T}_{pred}$), and the DAFF condition features $\mathcal{C}$. These embeddings are concatenated and processed by an MLP with LayerNorm and GELU activations to output the 18-dimensional velocity vector (for $N = 3$) by minimizing the flow matching objective defined in Eq. (1):

$$\mathcal{L}_{RFM} = \mathbb{E}_{t,\boldsymbol{T}_1,\boldsymbol{T}_0} ||\boldsymbol{v}_\theta(\boldsymbol{T}_t, t, \mathcal{C}) - (\boldsymbol{T}_1 - \boldsymbol{T}_0)||^2 \tag{1}$$

By performing flow matching in $\mathfrak{se}(3)$ and mapping predictions to $SE(3)$ via the exponential map, the learned trajectories remain consistent with the geometric structure of rigid transformations. To further constrain the $SE(3)$ structure, we introduce a geodesic loss $\mathcal{L}_{geo}$ to penalize errors in the matrix space. The total registration loss is formulated as defined in Eq. (2):

$$\mathcal{L}_{reg} = \mathcal{L}_{RFM} + \mathcal{L}_{NCC}(\boldsymbol{I}_{mov}, \boldsymbol{I}_{target} \cdot \boldsymbol{M}) + \alpha \mathcal{L}_{geo} \tag{2}$$

where $\mathcal{L}_{NCC}$ evaluates projection similarity [6], α balances manifold constraints, and $\boldsymbol{M}$ is a vertebral bounding box mask for localized loss calculation.

3D Reconstruction Branch The 3D reconstruction branch recovers fine-grained vertebral volumes, providing a spatial prior that guides precise RFM pose estimation. We design a CNN-based 3D decoder that begins with a linear projection layer, mapping 1D latent features into a base $C \times 4^3$ ($C = 1027$) tensor. This base volume undergoes a series of upsampling stages, each comprising a trilinear interpolation layer followed by a 3D convolution, BatchNorm3D, and GELU activation. These stages progressively increase the resolution, yielding a final 64^3 single-channel density volume $\boldsymbol{V}_{rec}$. To ensure structural fidelity, we supervise the reconstruction using a hybrid loss between the reconstruction volume $\boldsymbol{V}_{rec}$ and the preoperative volume $\boldsymbol{V}_{init}$ transformed by the predicted pose $\boldsymbol{T}_{pred}$), denoted as $\mathcal{W}(\boldsymbol{V}_{rec} \cdot \boldsymbol{T}_{pred})$, as described in Eq. (3):

$$\mathcal{L}_{rec} = \mathcal{L}_1(\boldsymbol{V}_{rec}, \mathcal{W}(\boldsymbol{V}_{init} \cdot \boldsymbol{T}_{pred})) + \beta \mathcal{L}_{SSIM}(\boldsymbol{V}_{rec}, \mathcal{W}(\boldsymbol{V}_{init} \cdot \boldsymbol{T}_{pred})) \tag{3}$$

where $\mathcal{W}$ is a spatial warping operator [5], $\mathcal{L}_1$ measures voxel-wise density errors, and $\mathcal{L}_{SSIM}$ computes the Structural Similarity Index Measure (SSIM).

FlowReg Training Strategy FlowReg was trained end-to-end by optimizing a joint loss that synergizes registration and reconstruction as defined in Eq. (4):

$$\mathcal{L}_{total} = \mathcal{L}_{reg} + \mathcal{L}_{rec} \tag{4}$$

This joint optimization ensured co-adaptation between the 3D reconstruction and RFM branches. To guarantee convergence without separate pre-training, the reconstruction task served as a geometric anchor, providing spatial priors that constrained the RFM feature space. This synergy regularized pose estimation, preventing suboptimal local minima. Additionally, The specific sim-to-real training protocols and implementation details are provided in Sec. 3.2.

3 Experiments

3.1 Data Preparation

Datasets were derived from the public VERSE '20 benchmark [16] and a proprietary real-world dataset. To facilitate robust pre-training, we utilized 253 VERSE '20 CT scans to generate 37,950 synthetic pairs of biplanar DRRs and perturbed volumes. Samples cover 1–3 level configurations, augmented via two-stage perturbations: global rigid transformations (rotations $\in [-20°, 20°]$, translations $\in [-50mm, 50mm]$) followed by local inter-vertebral deviations (rotations: $x\pm12°, y\pm6°, z\pm2°$; translations: ±2 mm) [21]. For the real-world dataset, we collected 30 intraoperative cases with preoperative CTs and paired AP/LAT X-rays (976×976, 1,124 mm focal length), yielding 62 samples (30 single-, 21 two-, 11 three-level). GT poses were calibrated via C-arm cone-beam acquisitions to ensure geometric consistency.

3.2 Experimental Setup and Evaluation Metrics

To implement the data-aware strategy, FlowReg was first pre-trained on the large-scale synthetic dataset, comprising 37,950 pairs of DRRs and randomly perturbed volumes. This stage facilitated hyperparameter optimization; based on five-fold cross-validation results, the weighting coefficients were empirically set to $\alpha = 0.02$ in Eq. (2) and $\beta = 0.05$ in Eq. (3). Subsequently, the model underwent fine-tuning on the real-world dataset to bridge the domain gap between synthetic and clinical data. The final results were analyzed using five-fold cross-validation, reporting performance metrics on the aggregated validation sets of the real-world dataset.

The evaluation results are presented in two main aspects: first, the SSIM and Peak Signal-to-Noise Ratio (PSNR) were used to gauge the quality of the reconstructed vertebra volumes. Second, the mean Target Registration Error (mTRE), mean Projection Distance (mPD), time efficiency, and Success Rate (SR) were used to measure the performance of 2D/3D registration [23]. The success of SR was defined as the mTRE of less than 5 mm.

All experiments were conducted on an AMD Core EYPC 7R32 48-core processor with 128 GB RAM and an NVIDIA GeForce RTX 4090 GPU (24 GB memory). The Adam optimizer was used for training, which lasted 300 epochs with an initial learning rate of 0.1, decaying by 0.5 every 50 epochs. The batch size was set to 8 to fit within GPU memory, and the time step t in Eq. (1) was set to 30.

3.3 Results

Table 1. Quantitative comparison on reconstruction and registration tasks. Statistical significance ($p < 0.05$, Wilcoxon signed-rank test [20]) is denoted by $*$. Top results are first, second, and third. (mean±std)

Method	$*$	Reconstruction		Registration			
		SSIM(%) ↑	PSNR(dB) ↑	mTRE(mm)↓	mPD(mm)↓	SR(%)↑	RT(s)↓
SAX-NeRF	✓	30.77±7.08	15.81±1.44	-	-	-	-
R^2-GS	✓	45.29±6.33	18.65±2.23	-	-	-	-
GeoA	✓	32.71±4.53	18.09±1.60	-	-	-	-
RadGS	✓	55.09±5.29	16.67±1.55	5.83±2.95	4.91±2.55	33.33	1.52
PolyPose	✓	-	-	11.75±8.70	9.98±7.45	25.40	122.61
4D-foot	✓	-	-	15.32±10.33	13.08±9.28	19.05	70.41
Ours	-	90.01±1.72	30.29±0.95	2.75±1.23	2.35±1.05	98.41	0.42

Comparison with Existing Methods We conducted experiments on 3D reconstruction and single- and multi-vertebrae registration, as detailed in Table 1. For reconstruction, we compared NeRF-based (SAX-NeRF [2]), 3DGS-based (RadGS [17], R^2-GS [22]) and learning-based (GeoA [9]) methods using $\mathcal{L}_{rec}$ from Eq. (3). For registration, our baselines included multi-object frameworks (PolyPose [5], 4D-foot [11]) and learning-based models (RadGS [17]). FlowReg consistently outperformed all competitors in both tasks. As shown in Fig. 3, across both single- and multi-vertebrae tasks, FlowReg robustly reconstructs accurate geometric structures and voxel intensity distributions that closely resemble the GT volume, despite the severe noise in input X-rays. Furthermore, Fig. 3 demonstrates that our approach yields the most precise alignment for both single- and multi-vertebrae registration. This performance is primarily attributed to FlowReg elevating the ill-posed 2D/3D registration problem into 3D space, effectively bridging the inherent domain gap. By explicitly learning the flow trajectory on the pose manifold, our method fundamentally resolves geometric ambiguities and achieves highly accurate multi-vertebrae registration.

Ablation Study Ablation studies were conducted to evaluate the pre-training strategy, the DAFF module, pose paradigms, and joint training (Table 2). Comparing (1) and (2) validated the necessity of the sim-to-real pipeline (D_2). Specifically, D_1 involved training solely on synthetic data with direct testing on clinical

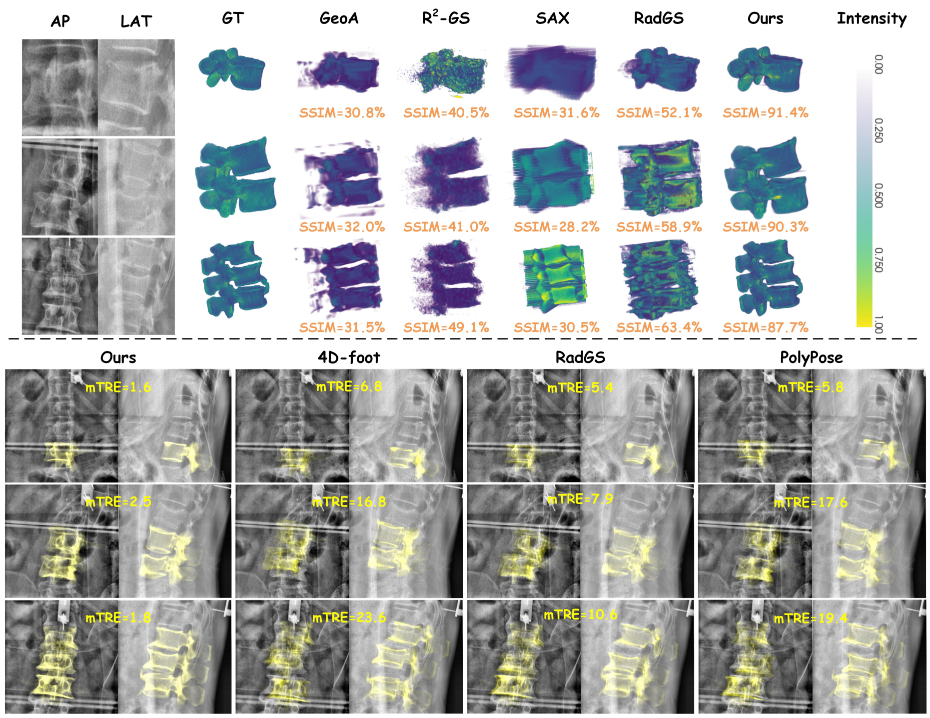

Fig. 3. Qualitative comparison with different reconstruction/registration methods. Each row represents an individual case.

datasets, while D_2 utilized synthetic pre-training followed by real-world fine-tuning via five-fold cross-validation, which reduced the mTRE to 8.92 mm. The DAFF module's effectiveness was evidenced in (2)-(3) and (6)-(7), where capturing cross-modal misalignments improved both pose estimation and reconstruction quality. Furthermore, (3)-(5) compared distinct paradigms: direct regression (Rgres.), diffusion-based prediction (Diffu.), and the proposed RFM. RFM outperformed both alternatives, reaching a 50.79% SR and proving the superiority of deterministic manifold trajectories. Finally, comparing isolated tasks (5, 7) with FlowReg (8) highlighted the impact of joint training. By co-optimizing both branches, FlowReg attained an optimal 98.41% SR and 90.01% SSIM, confirming that 3D geometric recovery provided essential spatial priors that fundamentally raised accuracy.

4 Conclusion

We present FlowReg, a versatile end-to-end framework designed for universal single- and multi-vertebrae CT/X-ray registration by synergizing 3D volume reconstruction with a deterministic RFM branch. Our approach effectively ad-

Table 2. Ablation study of our method. The top three results are highlighted as first, second, and third, respectively. (mean±std)

No.	Method	Reconstruction		Registration		
		SSIM(%) ↑	PSNR(dB) ↑	mTRE(mm)↓	mPD(mm)↓	SR(%)↑
(1)	Rgres.+D_1	-	-	17.59±11.11	14.86±9.57	14.29
(2)	Rgres.+D_2	-	-	8.92±6.15	7.56 ±5.43	22.22
(3)	Rgres.+DAFF+D_2	-	-	7.89±6.01	6.60±5.28	33.33
(4)	Diffu.+DAFF+D_2	-	-	6.64±5.64	5.47±4.86	38.10
(5)	RFM+DAFF+D_2	-	-	6.30±6.20	5.43±5.44	50.79
(6)	Recon+D_2	85.89±1.25	25.59±3.29	-	-	-
(7)	Recon+DAFF+D_2	87.50±1.16	27.24±3.51	-	-	-
(8)	FlowReg+D_2	90.01±1.72	30.29±0.95	2.75±1.23	2.35±1.05	98.41

dresses of traditional "render and compare" schemes [3,6,23,4] by utilizing reconstructed geometry as a robust spatial prior for registration. Key innovations include the DAFF module for reconciling cross-modal disparities, an RFM pose estimation branch that ensures SE(3)-aware geometric consistency in $\mathfrak{se}(3)$ while learning shared representations across vertebrae. FlowReg outperforms existing methods, achieving 90.01% SSIM in reconstruction and a 98.41% success rate with 2.75 mm mTRE in registration. By resolving the stochastic inefficiencies of generative diffusion models, its 0.42s runtime successfully meets the requirements for real-time intraoperative guidance. Future work will focus on incorporating non-rigid deformation models and validating the framework on broader multi-center clinical cohorts.

Acknowledgments. This work was supported partially by Jiangsu Province Key Research Program Grant BE2022718, Natural Science Foundation of China under Grant 62271465, National Key R&D Program of China under Grant 2025YFC3408300, and Suzhou Basic Research Program under Grant SYG202338.

Disclosure of Interests. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

References

1. Aly, T.A.: Short segment versus long segment pedicle screws fixation in management of thoracolumbar burst fractures: meta-analysis. Asian spine journal **11**(1), 150 (2017)
2. Cai, Y., Wang, J., Yuille, A., Zhou, Z., Wang, A.: Structure-aware sparse-view x-ray 3d reconstruction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11174–11183 (2024)

3. Gao, C., Liu, X., Gu, W., Killeen, B., Armand, M., Taylor, R.H., Unberath, M.: Generalizing spatial transformers to projective geometry with applications to 2d/3d registration. In: Martel, A.L., Abolmaesumi, P., Stoyanov, D., Mateus, D., Zuluaga, M.A., Zhou, S.K., Racoceanu, D., Joskowicz, L. (eds.) Medical Image Computing and Computer Assisted Intervention - MICCAI 2020 - 23rd International Conference, Lima, Peru, October 4-8, 2020, Proceedings, Part III. Lecture Notes in Computer Science, vol. 12263, pp. 329–339. Springer (2020)
4. Gopalakrishnan, V., Dey, N., Golland, P.: Intraoperative 2d/3d image registration via differentiable x-ray rendering. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11662–11672 (2024)
5. Gopalakrishnan, V., Dey, N., Golland, P.: Polypose: Deformable 2d/3d registration via polyrigid transformations. In: The Thirty-ninth Annual Conference on Neural Information Processing Systems (2025)
6. Grupp, R.B., Unberath, M., Gao, C., Hegeman, R.A., Murphy, R.J., Alexander, C.P., Otake, Y., McArthur, B.A., Armand, M., Taylor, R.H.: Automatic annotation of hip anatomy in fluoroscopy for robust and efficient 2d/3d registration. International journal of computer assisted radiology and surgery **15**, 759–769 (2020)
7. Jiang, J., Shuai, L., Zeng, Q.: Ablspinelevelcheck: Localization of vertebral levels on fluoroscopy via semi-supervised abductive learning. In: 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). pp. 3347–3350. IEEE (2024)
8. Li, T., Chen, M., Li, M., Li, C., Kong, Y.: Automatic x-ray to ct registration using embedding reconstruction and lite cross-attention. Medical Physics (2025)
9. Liu, Z., Fang, Y., Li, C., Wu, H., Liu, Y., Shen, D., Cui, Z.: Geometry-aware attenuation learning for sparse-view cbct reconstruction. IEEE Transactions on Medical Imaging **44**(2), 1083–1097 (2024)
10. Miao, S., Wang, Z.J., Liao, R.: A cnn regression approach for real-time 2d/3d registration. IEEE transactions on medical imaging **35**(5), 1352–1363 (2016)
11. Mizoe, S., Otake, Y., Miyamoto, T., Soufi, M., Nakao, S., Tanaka, Y., Sato, Y.: 4d-foot: a fully automated pipeline of four-dimensional analysis of the foot bones using bi-plane x-ray video and ct. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 182–192. Springer (2021)
12. Ouyang, W., Ye, Q., Wang, J., Xu, Z., Chen, J.: Rfmpose: Generative category-level object pose estimation via riemannian flow matching. In: The Thirty-ninth Annual Conference on Neural Information Processing Systems (2025)
13. Payer, C., Stern, D., Bischof, H., Urschler, M.: Coarse to fine vertebrae localization and segmentation with spatialconfiguration-net and u-net. In: VISIGRAPP (5: VISAPP). pp. 124–133 (2020)
14. Perdomo-Pantoja, A., Ishida, W., Zygourakis, C., Holmes, C., Iyer, R.R., Cottrill, E., Theodore, N., Witham, T.F., Lo, S.f.L.: Accuracy of current techniques for placement of pedicle screws in the spine: a comprehensive systematic review and meta-analysis of 51,161 screws. World neurosurgery **126**, 664–678 (2019)
15. Pizer, S.M., Amburn, E.P., Austin, J.D., Cromartie, R., Geselowitz, A., Greer, T., ter Haar Romeny, B., Zimmerman, J.B., Zuiderveld, K.: Adaptive histogram equalization and its variations. Computer vision, graphics, and image processing **39**(3), 355–368 (1987)
16. Sekuboyina, A., Husseini, M.E., Bayat, A., Löffler, M., Liebl, H., Li, H., Tetteh, G., Kukačka, J., Payer, C., Štern, D., et al.: Verse: a vertebrae labelling and segmentation benchmark for multi-detector ct images. Medical image analysis **73**, 102166 (2021)

17. Shen, A., Fu, X., Jiang, J., Zeng, Q., Tang, Y., Chen, Z., Nong, L., Wang, F., Zhou, S.K.: Radgs-reg: Registering spine ct with biplanar x-rays via joint 3d radiative gaussians reconstruction and 3d/3d registration. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 456–466. Springer (2025)
18. Siddon, R.L.: Fast calculation of the exact radiological path for a three-dimensional ct array. Medical physics **12**(2), 252–255 (1985)
19. Sun, T., Zhu, L., Huang, S., Song, S., Armeni, I.: Rectified point flow: Generic point cloud pose estimation. In: Advances in Neural Information Processing Systems (NeurIPS) (2025)
20. Wilcoxon, F.: Individual comparisons by ranking methods. In: Breakthroughs in statistics: Methodology and distribution, pp. 196–202. Springer (1992)
21. Yamamoto, I., Panjabi, M.M., Crisco, T., Oxland, T.: Three-dimensional movements of the whole lumbar spine and lumbosacral joint. Spine **14**(11), 1256–1260 (1989)
22. Zha, R., Lin, T.J., Cai, Y., Cao, J., Zhang, Y., Li, H.: R^2-gaussian: Rectifying radiative gaussian splatting for tomographic reconstruction. In: Advances in Neural Information Processing Systems (NeurIPS) (2024)
23. Zhang, B., Faghihroohi, S., Azampour, M.F., Liu, S., Ghotbi, R., Schunkert, H., Navab, N.: A patient-specific self-supervised model for automatic x-ray/ct registration. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 515–524. Springer (2023)
24. Zhang, J.Y., Lin, A., Kumar, M., Yang, T.H., Ramanan, D., Tulsiani, S.: Cameras as rays: Pose estimation via ray diffusion. In: International Conference on Learning Representations (ICLR) (2024)
25. Zhang, J., Wu, M., Dong, H.: Generative category-level object pose estimation via diffusion models. Advances in Neural Information Processing Systems **36**, 54627–54644 (2023)