SIIM CAIMI26 Conference on Artificial Intelligence in Medical Imaging
OCT 26-27 | University of Pennsylvania, Philadelphia, PA

Patient-Specific Volumetric CT Synthesis from Routine Biplanar Knee Radiographs

Amirali Khosravi, MD, Mayo Clinic; Linjun Yang, PhD; Kellen Mulford, PhD
Bardia Khosravi, MD, MPH, MHPE; Cody Wyles, MD; Bradley J. Erickson, MD, PhD, CIIP, FSIIM

Introduction/Background
Three-dimensional bone modeling depends on CT, which is costly, delivers higher radiation dose, and is not universally accessible. Radiographs are inexpensive, widely available, and may encode sufficient information to recover volumetric bone morphology. We developed a conditioned latent diffusion model reconstructing volumetric CT from AP and lateral radiographs, evaluated in robotic-assisted total knee arthroplasty (TKA), where every candidate receives a pre-operative CT for implant planning.

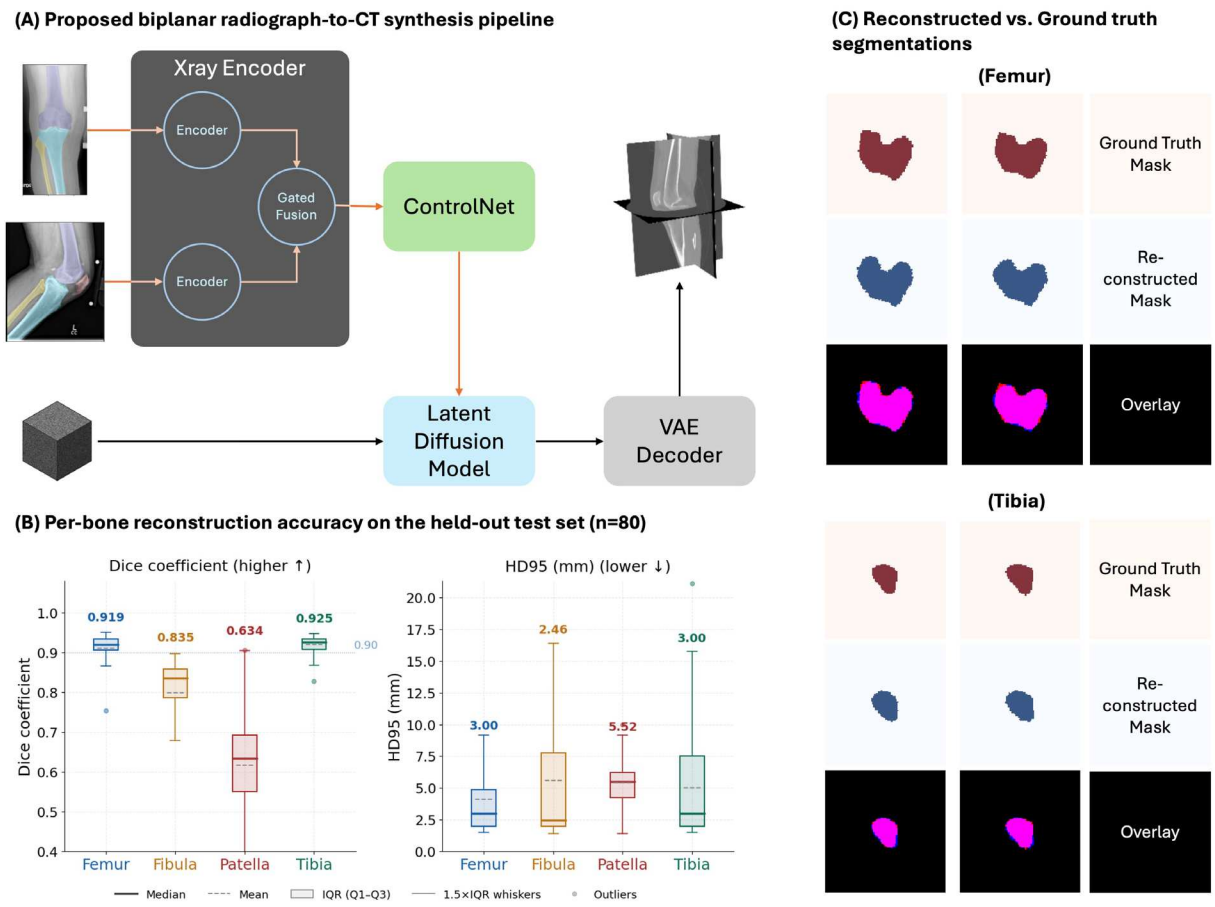
Methods/Intervention
A variational autoencoder and an unconditional latent diffusion model were trained on 6,300 knee CT volumes. Conditioning was implemented through a dual-arm X-ray encoder coupled to a ControlNet. These components were trained on 3,700 clinical radiograph–CT pairs acquired within a one-month interval, with 80 pairs held out for testing. Each arm receives one view, AP or lateral, concatenated with its bone segmentation mask; the resulting embeddings condition synthesis. All volumes were resampled to 1.5 × 1.5 × 1.5 mm isotropic spacing. Synthesized and reference volumes were segmented with TotalSegmentator, and Dice coefficient and 95th-percentile Hausdorff distance (HD95) were computed per bone following registration.

Results/Outcome
Femur reconstruction achieved median Dice 0.919 (IQR 0.906–0.933) and median HD95 3.00 mm (IQR 2.00–4.85). Tibia reconstruction achieved median Dice 0.925 (IQR 0.907–0.933) and median HD95 3.00 mm (IQR 2.00–7.50). The fibula was intermediate but variable, with median Dice 0.835 and median HD95 2.46 mm. The patella performed least well, with median Dice 0.634 and median HD95 5.52 mm. Narrow patellar IQR indicates consistent underperformance, attributable to AP occlusion by the distal femur and reliance on a single lateral view at variable flexion.

Conclusion
At 1.5 mm isotropic resolution, femoral and tibial median HD95 of 3.00 mm approaches the voxel diagonal of approximately 2.60 mm, placing surface error within roughly one voxel width. Patellar reconstruction and model behavior in severely deformed knees remain open questions. Work is ongoing on a deterministic, non-probabilistic formulation of the same reconstruction task, which eliminates sampling variability and the requirement for post-hoc registration.

Statement of Impact
Volumetric reconstruction from radiographs already present in the clinical record could reduce dependence on dedicated CT across musculoskeletal imaging, with the most immediate application to procedures requiring

3D osseous models, including robotic arthroplasty planning.



(A) Dual-arm X-ray encoder with gated fusion conditions a latent diffusion model via ControlNet; the VAE decoder reconstructs volumetric CT from paired AP and lateral radiographs and their bone masks. (B) Per-bone Dice and HD95 on the held-out test set (n = 80); boxes show IQR, solid lines medians (annotated), dashed lines means, whiskers 1.5 × IQR. (C) Representative femoral and tibial axial masks, reference versus reconstructed, with overlay (magenta agreement, red reference-only, blue reconstruction-only).

Keywords

Cross-modal image reconstruction; Radiograph-to-CT synthesis; Latent diffusion; Musculoskeletal imaging