Evaluation of AI-Based Synthetic CT Generation Methods in Pelvic MRI-Only Radiotherapy

Asieh Tavakol[1], Boyd McCurdy[1], Niranjan Venugopal[1,2]

[1] Department of Physics and Astronomy, University of Manitoba, Winnipeg, Manitoba, Canada

[2] Department of Radiology, University of Manitoba, Winnipeg, Manitoba, Canada

Introduction:

- MRI provides excellent soft-tissue contrast but lacks the electron-density information required for radiotherapy dose calculation (1–7).
- AI-generated synthetic CT (sCT) may enable MRI-only radiotherapy; however, pelvic anatomical variability can reduce reconstruction accuracy (7–10).
- Whole-image evaluation may conceal clinically relevant errors within individual organs, supporting the need for organ-specific assessment (7–10).
- This study compared three AI-based sCT algorithms using organ-specific evaluation of major pelvic structures.

Materials and Methods:

- Publicly available radiotherapy data from 10 pelvic cases, including CT and MRI images, were used for this study (12).
- Three AI-based algorithms for MR-to-sCT conversion were evaluated: Algorithm A (modified nnU-Net), Algorithm B (2.5D CNN), and Algorithm C (GAN-based Pix2Pix-SwinUNet) (13).
- Major pelvic structures, including the hip bones, femoral heads, bladder, rectum, and colon, were individually segmented.
- Algorithm performance was evaluated using both intensity-based metrics, including Mean Absolute Error (MAE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index Measure (SSIM), and geometric metrics, including Dice Similarity Coefficient (DSC), Sensitivity, Hausdorff Distance (HD), and Mean Average Surface Distance (MASD)(14).

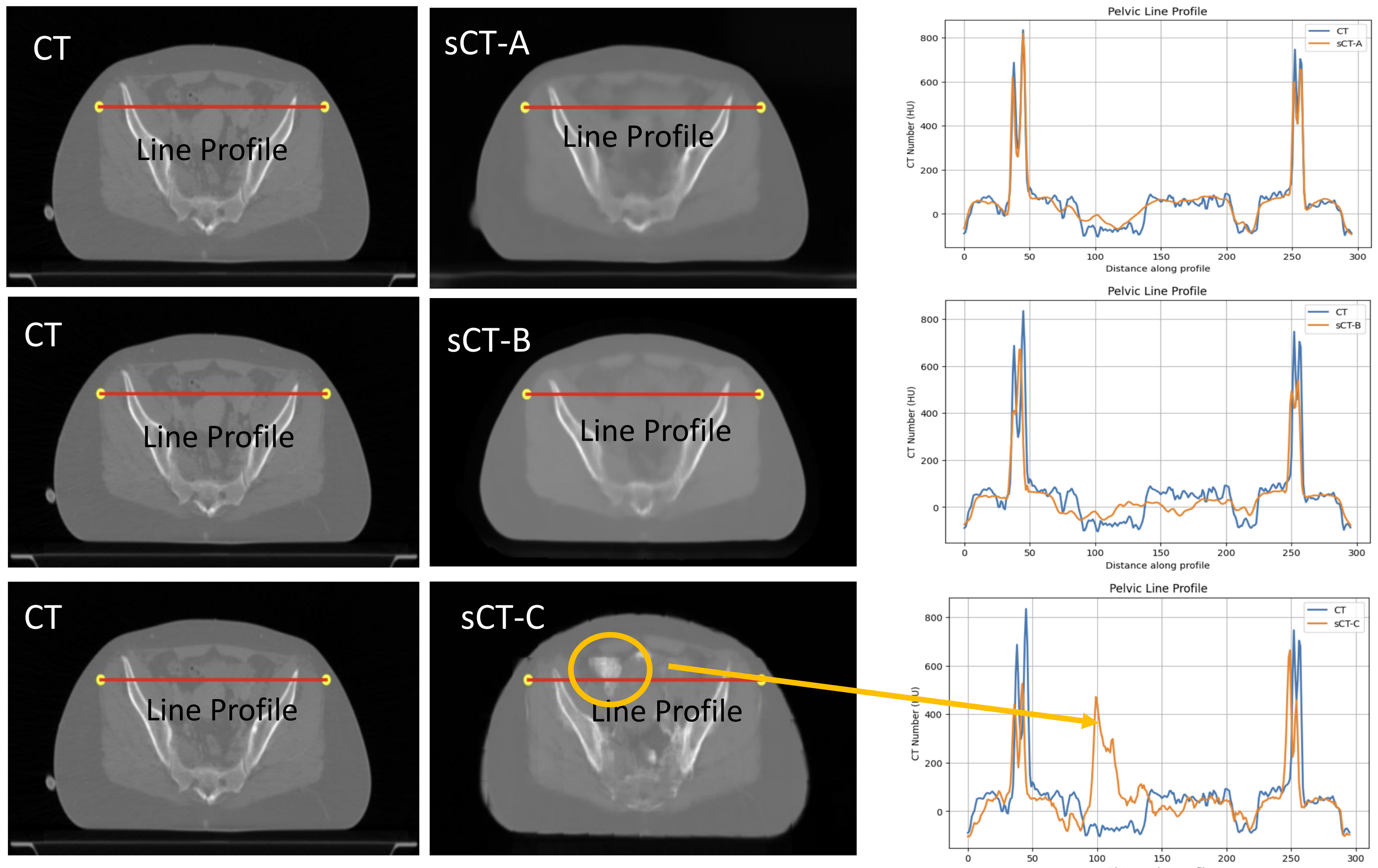
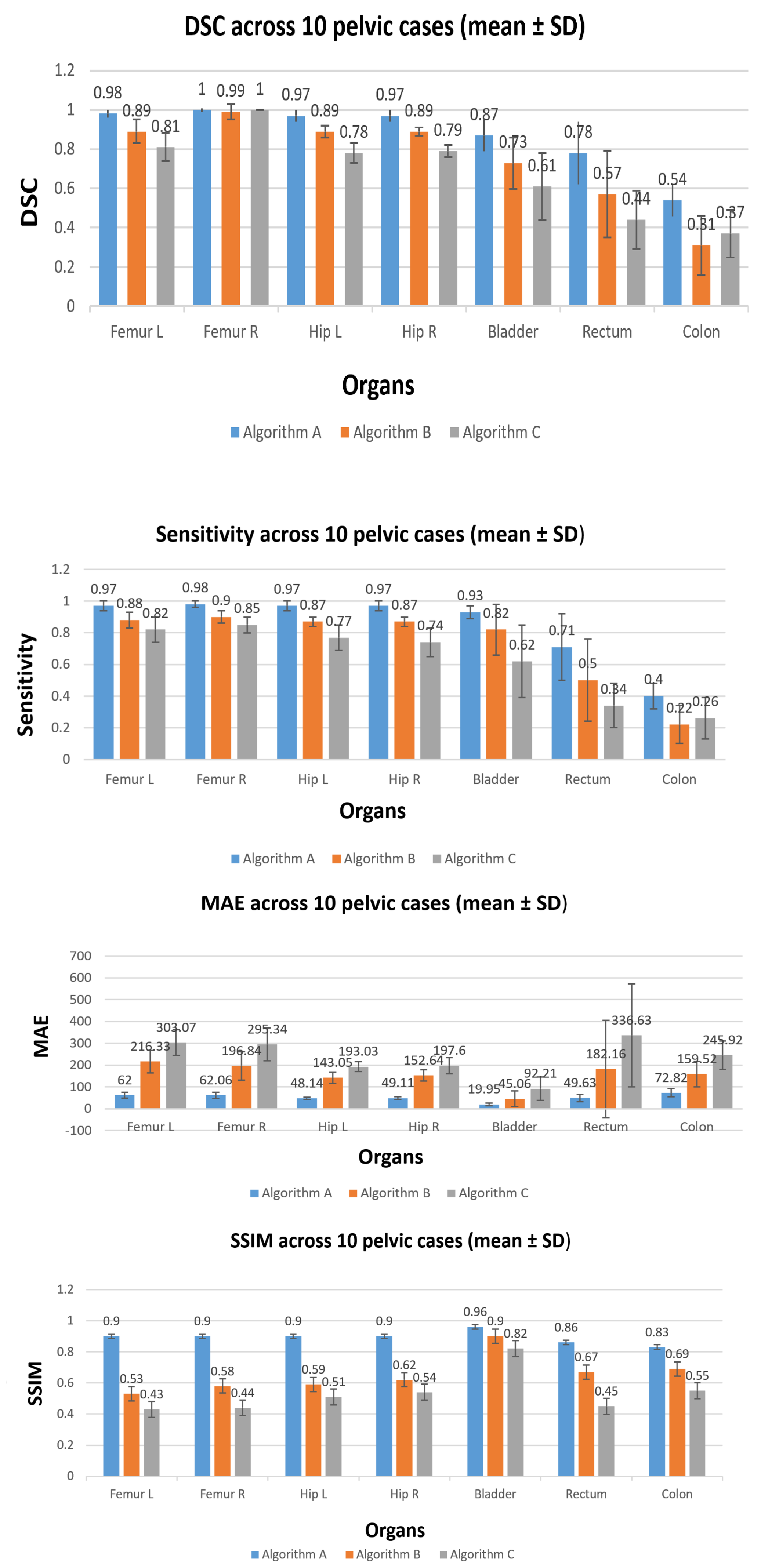


Figure 1. CT and sCT line profiles showing mild smoothing in A, greater detail loss in B, and soft-tissue misclassification as bone in C, highlighted by the orange circle and corresponding high-HU peak.



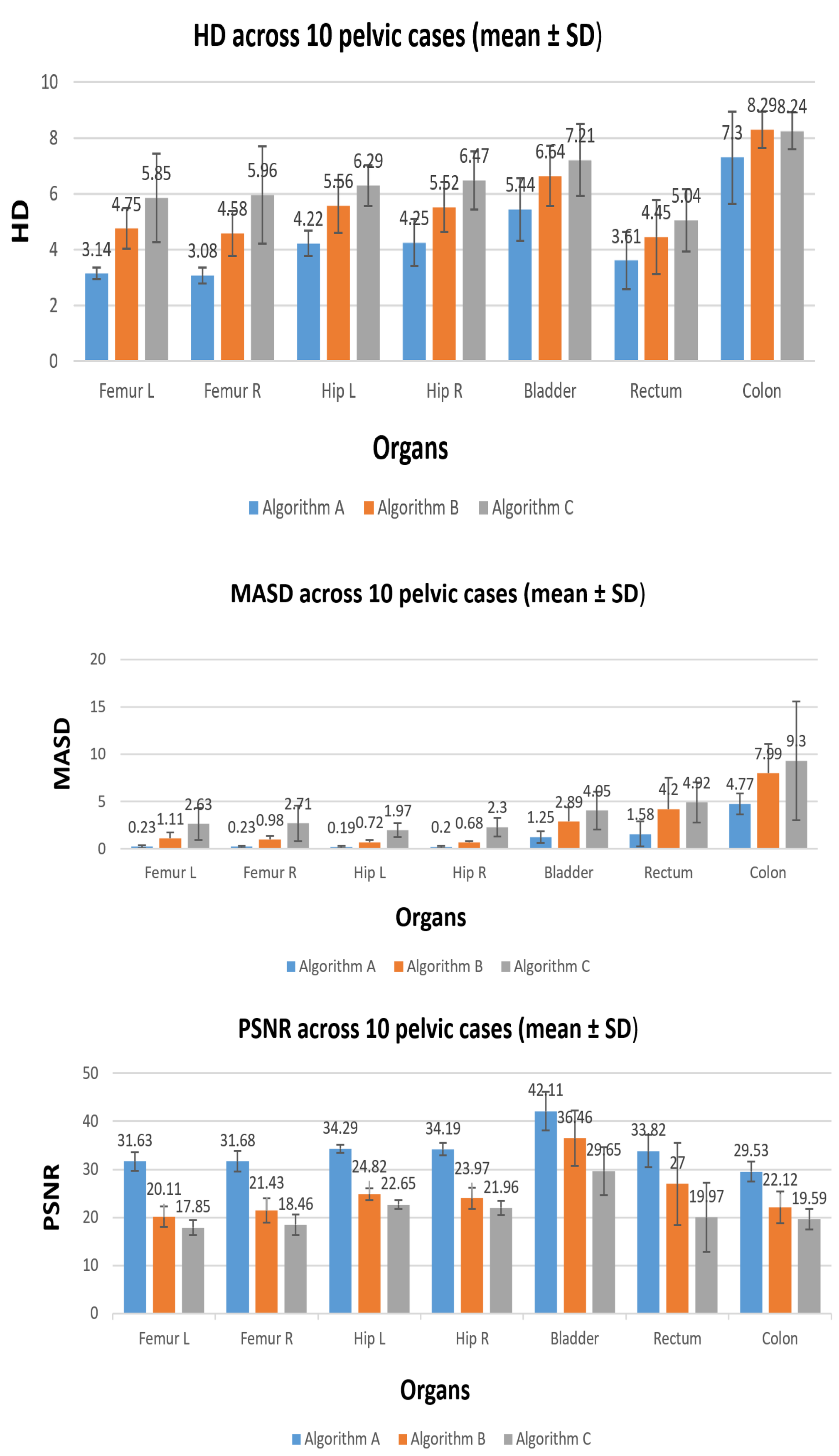


Figure 2. Quantitative organ-at-risk–specific performance of three AI-based sCT algorithms across geometric and intensity-based metrics, highlighting strong agreement in rigid structures and reduced accuracy in complex soft-tissue regions

Results:

- Algorithm A performed best overall, with near-perfect agreement in bone but lower accuracy in the bladder, rectum, and colon (Fig. 2).
- Line profiles showed slight smoothing for Algorithm A, greater oversmoothing and loss of detail for B, and bone-like soft-tissue misclassification for C (Fig. 1)
- The colon had the lowest agreement (DSC ≈ 0.54 for Algorithm A), highlighting organ-specific differences not evident in whole-image assessment (Fig. 2).

Discussions:

- Higher accuracy in rigid structures may reflect their stable anatomy and well-defined boundaries.
- Variable soft-tissue organs were more challenging, and organ-specific analysis revealed localized errors that whole-image metrics may conceal.

Conclusion:

- Algorithm A appears most promising for pelvic sCT generation, but improved reconstruction of variable soft-tissue organs is needed for robust clinical application.

References:

1. Boulanger et al. (2021); 2. Autret et al. (2023); 3. Pathmanathan et al. (2019); 4. Kerkmeijer et al. (2018); 5. Jonsson et al. (2019); 6. Owrangi et al. (2018); 7. Johnstone et al. (2018); 8. Stanescu et al. (2008); 9.Dowling et al. (2012); 10. Hsu et al. (2013); 11. Lui et al. (n.d.); 12. Thummerer et al. (2023); 13. Huijben et al. (2024); 14. Arabi et al. (2018)

Contact: tavako1@myumanitoba.ca