Beyond Hardware Limits: A Super-Resolution Convolutional Neural Network for Photon-Counting-Detector CT in Temporal Bone Imaging

Maryam Sadeghian, MS[1]; Shaojie Chang, PhD[1]; Du Zhang, PhD[1]; Cynthia H. McCollough, PhD[1]; Matthew Carlson, MD[2]; Shuai Leng, PhD[1]
[1]Department of Radiology, Mayo Clinic, Rochester, MN, USA; [2]Department of Otolaryngology-Head and Neck Surgery, Mayo Clinic, Rochester, MN, USA

INTRODUCTION

CT is commonly used to visualize temporal bone anatomy, pathology, prostheses, and implant in patients with hearing loss. Spatial resolution is critical in detecting and delineating fine structures in the temporal bone.[1−3] Photon-counting-detector (PCD) CT has shown great impact on high-resolution imaging, but the fine osseous and membranous structures of the temporal bone approach the limits of current hardware capabilities.[4,5] The present work develops a novel AI-based super-resolution neural network to extend effective resolution beyond current hardware constraints and open new opportunities for temporal bone assessment.

Combining PCD-CT with AI-based super-resolution extends spatial resolution beyond current hardware constraints for temporal bone imaging.

METHODS

Data Collection and Image Reconstruction

Following IRB approval, 10 patients underwent temporal bone imaging on a PCD-CT system (NAEOTOM Alpha, Siemens) using clinical protocol (120 kV, UHR: 120 × 0.2 mm). Images were reconstructed at 0.2/0.1 mm using quantum iterative reconstruction (QIR, strength 3). Clinical images served as network input, while sharper Hr96 reconstructions were used as training targets. Phantom and patients with prosthesis and implants were also included in the study.

Data Preparation

The Hr96 images were first processed using a dedicated CNN denoising network to reduce image noise associated with sharp kernel.[6] A total of 10,000 paired image patches (128 × 128 pixels) were extracted from the input and target images.

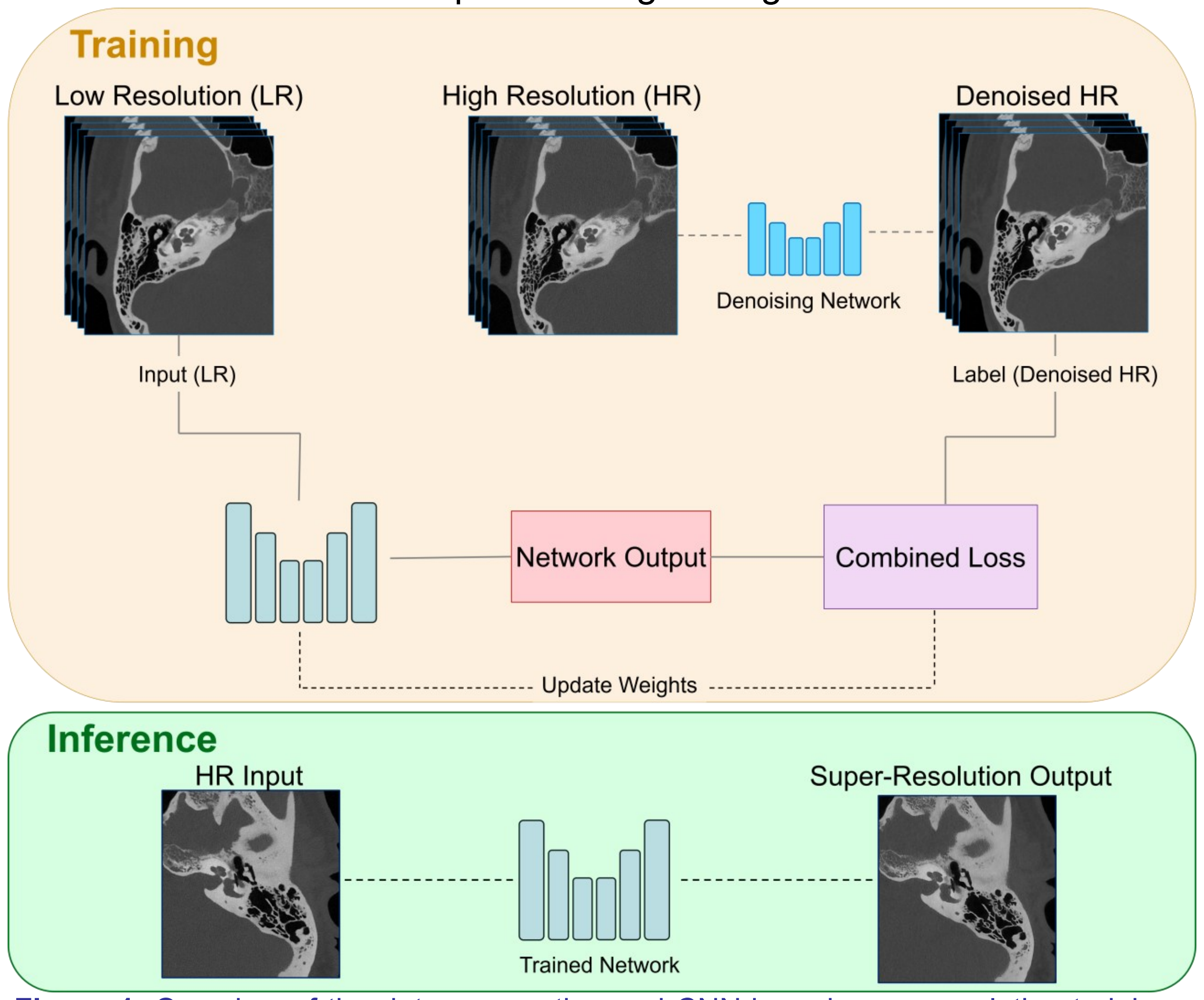

Figure 1. Overview of the data preparation and CNN-based super-resolution training framework. Low-resolution (LR) images were divided into patches for network input. High-resolution (HR; Hr96) images were first processed using a separate denoising CNN to reduce noise and then used as patched training labels. The trained network was subsequently applied to clinically acquired images.

Neural Network

A CNN-based super-resolution network employed a modified U-Net architecture, consisting of 4 max-pooling and 4 up-convolutional layers with exponential linear unit (ELU) activation. The network was trained using the Adam optimizer at a fixed learning rate of 0.001, with a combined loss function comprising MSE, SSIM, and gradient loss, weighted 0.3, 0.5, and 0.2, respectively. The trained model was applied to clinically reconstructed Hr84 images to generate super-resolution outputs.

Model Performance Evaluation (qualitatively and quantitatively)

Image quality improvement was assessed by visual inspection. Spatial resolution was quantified by measuring the 10–90% rise distance from line profiles drawn through corresponding structures in the input and super-resolution network output images for both phantom and patient datasets.

RESULTS

Network output showed substantial improvement in spatial resolution compared to the commercial (clinical) PCD-CT images, improving the visualization and delineation of fine osseous and membranous structures, including delicate trabecular and cellular patterns that were previously poorly resolved. Similar improvement was observed in patients with middle ear prostheses or cochlear implants in situ, as well as in the phantom prosthesis.

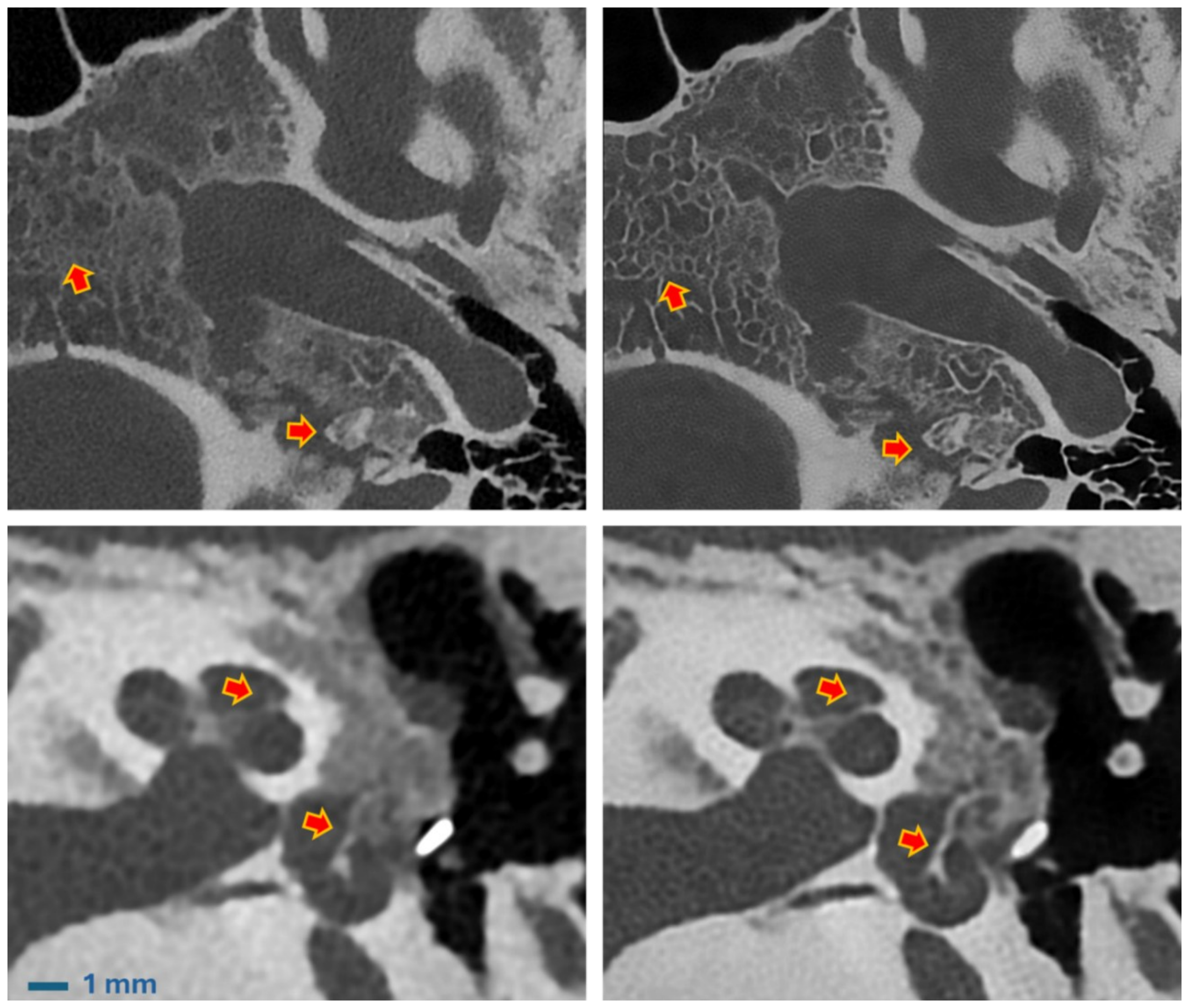

Figure 2. Representative patient temporal bone images reconstructed with commercial PCD-CT (left column) and super-resolution network output (right column). Arrows show regions of improved visualization with super-resolution.

Quantitative line-profile measurements through the phantom prosthesis showed a 10–90% rise distance of 0.234 mm for clinically reconstructed images versus 0.156 mm for network output, a 33.3% improvement. Figure 3 shows sample images of the middle ear prosthesis from Hr84 and network output, with the latter demonstrating sharper image appearance, better delineation of individual structures, and a steeper line-profile slope indicating improved spatial resolution. Figure 2 shows sample patient temporal bone images, demonstrating improved delineation of fine osseous structures, including trabecular and cellular bone patterns, with network output compared to clinically reconstructed images.

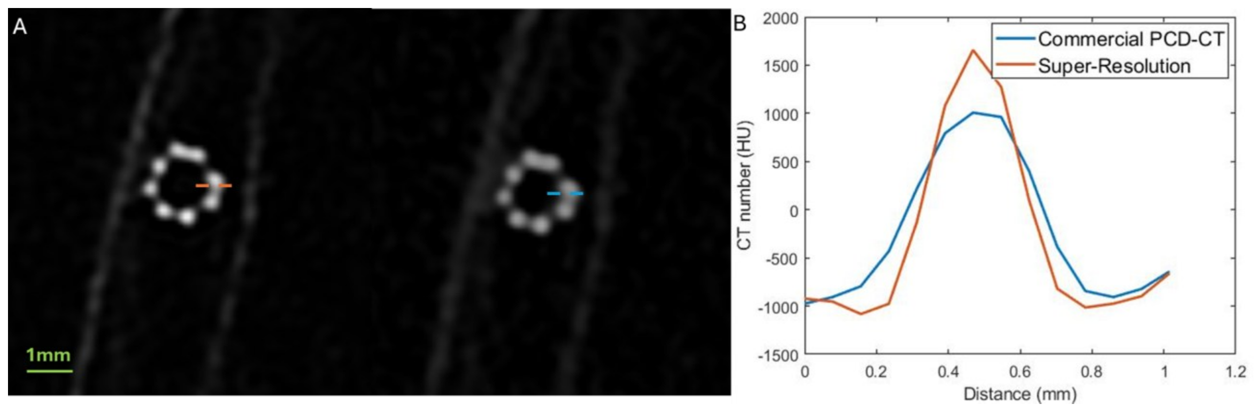

Figure 3. Comparison of super-resolution network output and commercial PCD-CT imaging of a representative prosthesis from phantom scan. A. Sample images showing super-resolution output (left) and Hr84 kernel input from commercial PCD-CT (right). B. Line profiles along the dashed line in panel A, showing improved resolution with super-resolution network.

DISCUSSION

- CNN-based super-resolution enhanced the effective spatial resolution of clinical PCD-CT images without requiring additional radiation dose, scan time, or hardware modifications.
- Training with AI-denoised high-resolution target images enabled improved sharpness while avoiding the increased noise typically associated with edge-enhancing reconstruction kernels.
- The improved depiction of fine temporal bone anatomy and implanted devices may increase diagnostic confidence in otologic imaging.
- These findings demonstrate the potential of AI-based super-resolution to extend the capabilities of PCD-CT.
- Further validation in larger patient cohorts and across a broader range of pathologies is warranted.

CONCLUSIONS

- CNN-based super-resolution substantially improved the effective spatial resolution of clinical PCD-CT temporal bone images.
- Fine anatomical structures and implanted devices were depicted with greater clarity than standard clinical reconstructions.
- This approach enhances existing PCD-CT acquisitions without additional radiation dose or scan time and shows promise for improving temporal bone imaging in clinical practice.

ACKNOWLEDGEMENTS

Research reported in this work was supported in part by the NIH under award number R01 DC022228 and the Mayo Clinic Graduate School of Biomedical Sciences.

REFERENCES

1. Pham N, Raslan O, Strong EB, et al. High-Resolution CT Imaging of the Temporal Bone: A Cadaveric Specimen Study. J Neurol Surg B Skull Base. 2022;83(5):470-475. Published 2022 Jan 31. doi:10.1055/s-0041-1741006
2. Sadeghian M, Benson JC, Farnsworth PJ, et al. Impact of Spatial Resolution on CT Imaging of Middle Ear Prostheses: Comparison of Photon-Counting and Energy-Integrating Detectors. AJNR Am J Neuroradiol. 2026;47(5):1341-1348. Published 2026 May 4. doi:10.3174/ajnr.A9103
3. Altmann S, Abello Mercado MA, Ucar FA, et al. Ultra-High-Resolution CT of the Head and Neck with Deep Learning Reconstruction-Assessment of Image Quality and Radiation Exposure and Intraindividual Comparison with Normal-Resolution CT. Diagnostics (Basel). 2023;13(9):1534. Published 2023 Apr 24. doi:10.3390/diagnostics13091534
4. Rajendran K, Voss BA, Zhou W, et al. Dose Reduction for Sinus and Temporal Bone Imaging Using Photon-Counting Detector CT With an Additional Tin Filter. Invest Radiol. 2020;55(2):91-100. doi:10.1097/RLI.0000000000000614
5. Benson JC, Rajendran K, Lane JI, et al. A New Frontier in Temporal Bone Imaging: Photon-Counting Detector CT Demonstrates Superior Visualization of Critical Anatomic Structures at Reduced Radiation Dose. AJNR Am J Neuroradiol. 2022;43(4):579-584. doi:10.3174/ajnr.A7452
6. Chang S, Benson JC, Lane JI, et al. Ultra-High-Resolution Photon-Counting-Detector CT with a Dedicated Denoising Convolutional Neural Network for Enhanced Temporal Bone Imaging. AJNR Am J Neuroradiol. 2025;46(6):1188-1195. Published 2025 Jun 3. doi:10.3174/ajnr.A8572

Contact information: sadeghiansharafi.maryam@mayo.edu