ORIGINAL RESEARCH

Ultra-High-Resolution Photon-Counting-Detector CT with a Dedicated Denoising Convolutional Neural Network for Enhanced Temporal Bone Imaging

Shaojie Chang, John C. Benson, John I. Lane, Michael R. Bruesewitz, Joseph R. Swicklik, Jamison E. Thorne, Emily K. Koons, Matthew L. Carlson, Cynthia H. McCollough, Shuai Leng

ABSTRACT

BACKGROUND AND PURPOSE: Ultra-high-resolution (UHR) photon-counting-detector (PCD) CT improves image resolution but increases noise, necessitating use of smoother reconstruction kernels that reduce resolution below the system's 0.110 mm maximum spatial resolution. To address this, a denoising convolutional neural network (CNN) was developed to reduce noise in images reconstructed with the available sharpest reconstruction kernel while preserving resolution for enhanced temporal bone visualization.

MATERIALS AND METHODS: With IRB approval, CNN was trained on 6 clinical temporal bone patient cases (1,885 images) and tested on 20 independent cases using a dual-source PCD-CT (NAEOTOM Alpha, Siemens). Images were reconstructed using iterative reconstruction at strength 3 (QIR3) with both clinical routine (Hr84) and the sharpest available head kernel (Hr96). The CNN was applied to images reconstructed with Hr96 and QIR1. Three image series (Hr84-QIR3, Hr96-QIR3, and Hr96-CNN) for each case were randomized for review by two neuroradiologists, assessing overall quality and delineation of the modiolus, stapes footplate, and incudomallear joint.

RESULTS: CNN reduced noise by 80% compared to Hr96-QIR3 and 50% relative to Hr84-QIR3, while maintaining high resolution. When compared to the conventional method at the same kernel (Hr96-QIR3), Hr96-CNN significantly decreased image noise (from 204.63 HU to 47.35 HU) and improved SSIM (from 0.72 to 0.99). Hr96-CNN images ranked higher than Hr84-QIR3 and Hr96-QIR3 in overall quality ($p<0.001$). Readers preferred Hr96-CNN for all three structures.

CONCLUSIONS: The proposed CNN significantly reduced image noise in UHR PCD-CT, enabling the use of sharpest kernel. This combination greatly enhanced diagnostic image quality and anatomical visualization.

ABBREVIATIONS: PCD = Photon-counting-detector; UHR = Ultra-high-resolution; IR = Iterative reconstruction; CNN = Convolutional neural network; SSIM: Structural similarity index.

Received month day, year; accepted after revision month day, year.
From the Department of Radiology (Shaojie Chang, John C. Benson, John I. Lane, Michael R Bruesewitz, Joseph R Swicklik, Jamison E. Thorne, Emily K. Koons, Cynthia H. McCollough, Shuai Leng) and Department of Otolaryngology–Head and Neck Surgery (Matthew L. Carlson), Mayo Clinic, Rochester, MN, 55905, USA.

The authors declare no conflicts of interest related to the content of this article.

Please address correspondence to Shuai Leng, PhD, Department of Radiology, Mayo Clinic, 200 1st St SW, Rochester, MN, 55905, USA; Leng.Shuai@mayo.edu

SUMMARY SECTION

PREVIOUS LITERATURE: Ultra-high-resolution (UHR) photon-counting detector (PCD) CT can enhance image resolution and improve visualization of temporal bone structures. However, the system's maximum spatial resolution has not been fully explored in previous studies due to the associated increase in noise.

KEY FINDINGS: A dedicated convolutional neural network enhanced high-resolution temporal bone imaging using the sharpest kernel in Photon-Counting-Detector CT, outperforming conventional methods, and significantly improving diagnostic quality and visualization of critical anatomical structures.

KNOWLEDGE ADVANCEMENT: The resulting high-resolution images feature acceptable noise levels that not only improve anatomical delineation but also more precisely define the interfaces between metal prostheses and surrounding structures, enhancing temporal bone visualization.

INTRODUCTION

Temporal bone structures, including the facial nerve and labyrinth, are submillimeter in scale and require high-spatial-resolution imaging.[1-4] Recently, photon counting detector (PCD) CT systems have demonstrated the ability to provide ultra-high-resolution (UHR) images, with in-plane resolutions reaching down to 0.125 mm.[5-9] However, using this level of spatial resolution is challenging clinically because images at the highest resolution exhibit excessive noise as demonstrated in Fig 1. Therefore, it is crucial to maintain acceptable noise levels while preserving detailed structures, particularly in high-resolution imaging at clinical dose levels.

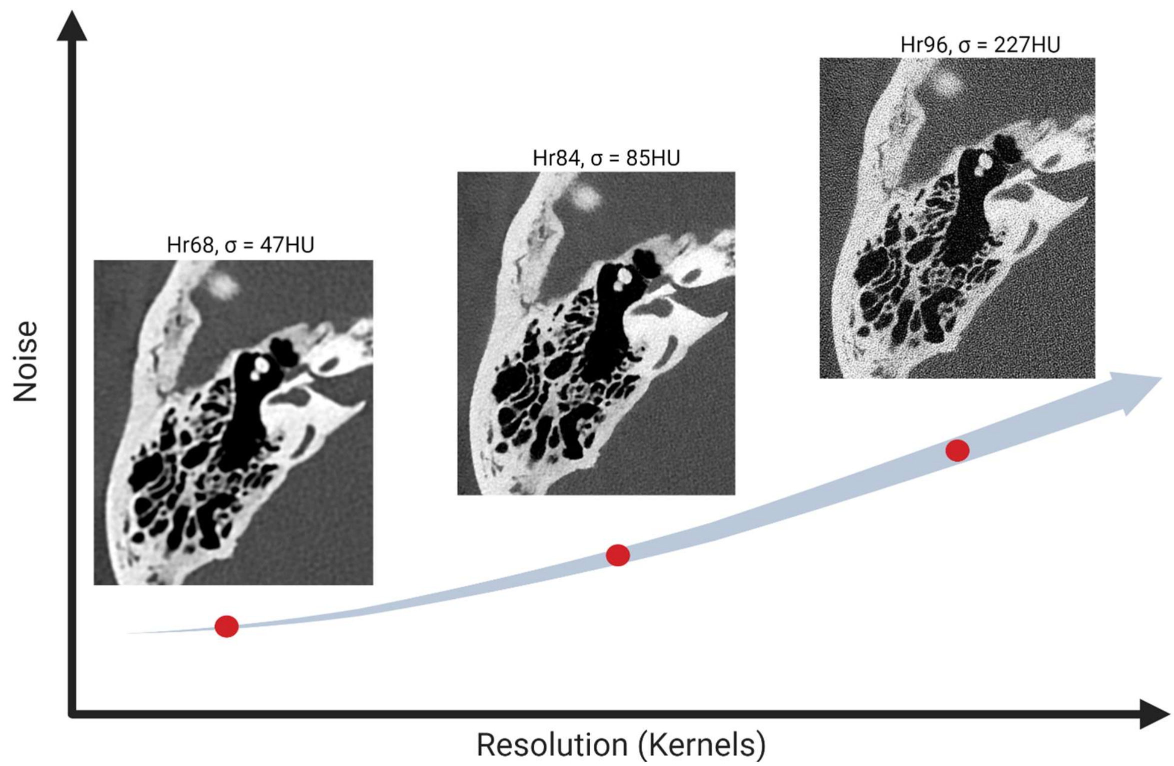

FIG 1. Image resolution vs. noise in PCD-CT: As image resolution increases, image noise also increases, which can limit the utility of high-resolution settings in PCD-CT.

In practice, PCD-CT temporal bone exams typically use smoother kernels (e.g., Hr84) instead of the sharpest available kernel (e.g., Hr96) to keep the noise at acceptable levels. The choice of kernel determines the final image resolution, with higher numbers indicating greater spatial resolution.[10,11] For instance, the Hr84 kernel results in an in-plane resolution of 0.154 mm, which is inferior to the system's maximum resolution (0.110 mm).[12] Noise reduction is an important topic in CT imaging and various methods have been investigated, such as traditional iterative reconstruction (IR)[13] and deep learning-based methods.[14,15] As highlighted by C. Niu et al.[16], deep learning approaches, which train a neural network to remove noise from a noisy image, demonstrate superior denoising performance compared to IR methods. However, these approaches often require extensive collections of spatially aligned low- and routine-dose patient images, which are challenging to obtain.

In this study, we developed a dedicated denoising convolutional neural network (CNN) to significantly reduce image noise in UHR PCD-CT, enabling the use of the sharpest kernel with acceptable noise levels for enhanced temporal bone visualization. Our approach utilizes a training dataset from the scanner, comprising both thin-slice and thick-slice IR images. All images are obtained from the routine clinical scan, without the need of images at different dose levels. This strategy not only ensures high-quality training data but also simplifies the replication of our methods by other researchers and facilitates adaptation to various clinical applications.

MATERIALS AND METHODS

Data Collection

This retrospective study was approved by our institutional review board and was Health Insurance Portability and Accountability Act compliant, with informed consent waived. The methodology proposed in the CLAIM checklist (supplementary material) was followed. Data from six adult clinical temporal bone PCD-CT scans were utilized for CNN training and validation, while 20 independent patient cases were used for testing. All exams were conducted in UHR mode (120 × 0.2 mm collimation, 120 kV) on a dual-source Photon Counting Detector CT (NAEOTOM Alpha, Siemens Healthineers, Forchheim, Germany), 1 second rotation time, 0.5 helical pitch, and automatic exposure control with 220 CARE keV IQ level, resulting in a volume CT dose index (CTDIvol) of 34 mGy for standard size patients. The training and validation dataset included 1,885 CT images, reconstructed using both thin-slice (0.2 mm slice thickness with 0.1 mm increment) and thick-slice (0.4 mm slice thickness with 0.2 mm increment) settings. Iterative reconstruction was applied at strength 1 (QIR1) for thin slices and strength 3 (QIR3) for both thin and thick slices. All images utilized the sharpest available kernel (Hr96) and were processed with a 1024 × 1024 matrix size within an 80 mm clinical standard field of view. The trained CNN was then applied to test cases, using images reconstructed at 0.2 mm thickness with the Hr96 kernel and QIR1. For reference, images reconstructed with the clinical routine kernel (Hr84) at a 0.2 mm slice thickness and QIR3 were also collected.

The Dedicated Denoising CNN Training Workflow

Figure 2 outlines our denoising CNN training workflow which begins by creating 'Noise_thin' images. These are generated by subtracting thin-slice iterative reconstructions with strengths 1 (QIR1_thin) and 3 (QIR3_thin), both set at a 0.2 mm slice thickness and increment, at the same anatomical location (higher strength settings perform more aggressive denoising). To prevent overfitting and introduce variability, we applied spatial decoupling techniques to "Noise_thin" images through random translations (ranging from 1 to 16 pixels) and inversions (using multipliers of +1 or -1) to create a set of randomized noise images.

Additionally, a set of thicker slice images (0.4 mm thickness, 0.2 mm increment), referred to as 'QIR3_thick,' was reconstructed as a low-noise reference. CNN inputs were formed by combining noise-only and reference patches (QIR3_thick + α * Noise_thin) from 7 adjacent slices, with α empirically set at 2.0 to balance noise reduction and detail preservation.[15] The central slice patch of QIR3_thick was the training target. As demonstrated in a previous study[17], training the CNN with thick reference images results in significantly improved noise reduction compared to using single-slice reference images. Moreover, the CNN's performance remains consistent when applied to single-slice images, even though it was trained on thick images. This is because the network focuses on learning noise patterns rather than the underlying tissue structures. Finally, the trained CNN's denoising efficacy was then tested on thin QIR1 images.

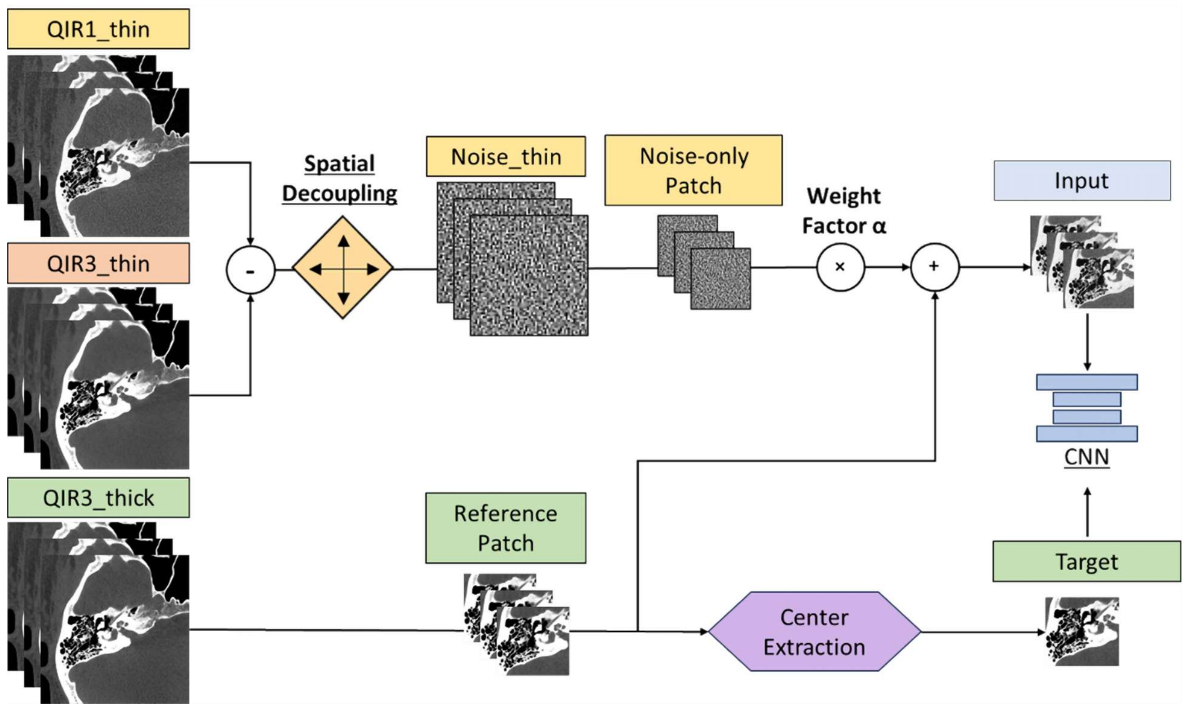

FIG 2. The overall workflow of the proposed deep CNN denoising method. All training data originated from patient image series reconstructed using two iterative reconstruction strengths, QIR1 and QIR3, with thin-slice (0.2 mm) and thick-slice (0.4 mm) thicknesses, respectively. A multiple-slice input strategy was implemented to enhance the CNN's performance.

Network Architecture and Training Details

We utilized a simplified U-Net architecture[18] with nine modules for our study. Each module sequentially performs convolution, batch normalization, and exponential linear unit activation operations. The architecture includes max pooling layers, convolution transpose operators, and concatenation to maintain input-output similarity. The mean-squared-error loss function was optimized during training. Our final training set comprised 18,864 patches of 128 × 128 pixels from the training data and 2,096 from validation data, at a 9:1 ratio. Training began with an initial learning rate of 0.001, progressively reduced to 0.00001, using the Adam optimizer[19] to minimize mean-squared-error loss function. We set the training for 100 epochs to ensure model convergence.

Phantom Experiments to Evaluate Noise and Spatial Resolution

A 20-cm-diameter American College of Radiology (ACR) CT accreditation phantom (Gammex) was scanned to evaluate the noise power spectrum (NPS) and the contrast-dependent modulation transfer function (MTFc). All acquisitions and reconstructions were performed according to the clinical protocol settings used in this study. MTFc and NPS were calculated using data from Hr84-QIR3, Hr96-QIR3, and Hr96-CNN, utilizing the online platform (https://www.ctpro.net) to illustrate the noise and resolution changes across different methods. In this study, NPS was computed by placing ten square regions of interest (ROIs) in the uniform section (module 3) of the ACR phantom. MTFc was calculated using bone cylindrical inserts (25 mm diameter, 4 cm depth) in module 1 of the phantom, based on 40 consecutive axial slices, to assess in-plane spatial resolution under high-contrast conditions. A circular ROI was placed around the insert, and a circular-edge technique was employed to measure the edge spread function (ESF) by plotting each pixel's HU value as a function of the distance from the center of the insert. The line spread function (LSF) was then derived from the ESF. After zero-padding, a fast Fourier transform (FFT) was applied to the LSF to compute the in-plane MTFc.

Objective Image Quality Assessment

For image quality evaluation, noise was measured in axial images as the standard deviation (SD) of CT numbers in a circular ROI drawn in a uniform soft-tissue area for each dataset. The size and location of ROIs were matched among 3 image sets (Hr84-QIR3, Hr96-QIR3, and Hr96-CNN).

Reader Evaluation

Two fellowship-trained neuroradiologists (>10 years experience each) assessed the overall image quality and delineation of three key anatomical structures— modiolus, stapes footplate, and incudomallear joint—for each of the 20 test cases. They assessed three image series per case (Hr84-QIR3, Hr96-QIR3, and Hr96-CNN), which were displayed side-by-side in a randomized and blinded manner. Images were ranked on a scale from 1 to 3, with 1 as the most preferred and 3 as the least preferred. Equal ranking was permitted.

Statistical Analyses

Statistical analyses were conducted using Python's statistical package *scikit-posthocs*. Pairwise comparisons were performed with Conover's post hoc test, applying a Bonferroni correction, to evaluate differences between two variables: the average rankings from two readers on overall image quality and diagnostic confidence for discerning three anatomical structures across Hr84-QIR3, Hr96-QIR3, and Hr96-CNN. A p-value < 0.05 was considered statistically significant.

Denoising Performance Comparison

The Residual Encoder-Decoder Convolutional Neural Network (RED-CNN) and U-Net are two of the most widely used models for CT image denoising[20]. Using our proposed dataset preparation workflow, we conducted a comparative analysis of the denoising performance between RED-CNN and the U-Net model we developed. Hr96-QIR1 was used as the input reference, while Hr96-QIR3 served as the conventional denoised reconstruction. We compared the difference images generated by subtracting the reference noisy input image from the conventional method, RED-CNN, and U-Net denoised images. Image quality was objectively assessed using the Structural Similarity Index (SSIM) and image noise measurements for each image set. This analysis was repeated at the patient level, with the mean and SD of these metrics reported.

RESULTS

NPS and MTFc performance on ACR phantom

Figure 3 illustrates the noise textures, NPS, and MTFc for the bone insert from axial slices of the ACR phantom across three configurations: Hr84-QIR3, Hr96-QIR3, and Hr96-CNN. The noise levels in Hr84-QIR3 ($\sigma = 90$ HU) and Hr96-QIR3 ($\sigma = 264$ HU) are higher than those in Hr96-CNN ($\sigma = 36$ HU), with the noise in these images exhibiting higher spatial frequency components. This is consistent with the NPS measurements shown in the middle panel of Figure 3. The NPS peak for Hr96 decreases from 30.4 cm^{-1} to 4.4 cm^{-1} after applying CNN denoising, whereas the NPS peak for Hr84-QIR3 is at 15.6 cm^{-1}. The MTFc results for the bone insert are displayed in the bottom panel of Figure 3. The spatial frequencies at 10% indicate that both Hr96-QIR3 and Hr96-CNN, utilizing the sharp kernel, offer better MTFc performance than the clinical routine Hr84-QIR3. The CNN denoising method preserved resolution in the Hr96 sharp kernel, with the 10% MTFc at 36.8 cm^{-1} and 38.4 cm^{-1} for the CNN and QIR3, respectively

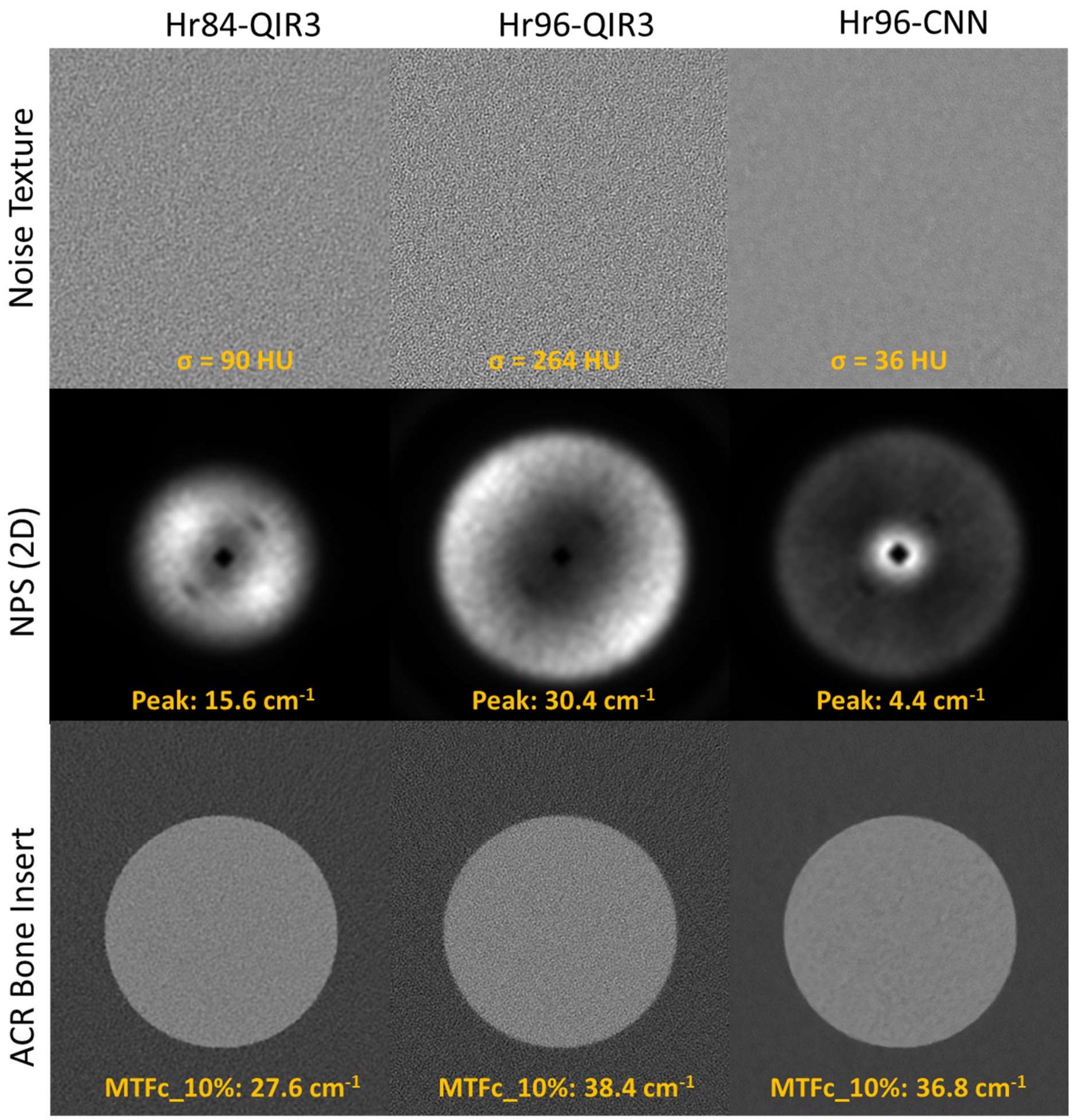

FIG 3. Noise textures, NPS, and MTFc for the bone insert from axial slices of the ACR phantom for Hr84-QIR3, Hr96-QIR3, and Hr96-CNN, displayed under a fixed window and level.

Example Images

Figure 4 displays representative images of modiolus, stapes footplate, and incudomallear joint using Hr84-QIR3, Hr96-QIR3, and Hr96-CNN and highlights the enhanced capability of CNN denoising at the highest resolution (Hr96) to clearly delineate each evaluated structure with acceptable noise levels. The CNN effectively reduced image noise—by approximately 80% compared to the highest resolution commercial images (Hr96-QIR3), and by 50% relative to clinical routine images (Hr84-QIR3)—while demonstrating ultra-high resolution (as assessed visually). Compared to the routine images (Hr84-QIR3), CNN denoised high-resolution images (Hr96-CNN) show substantially improved spatial resolution and better delineation of key anatomic structures.

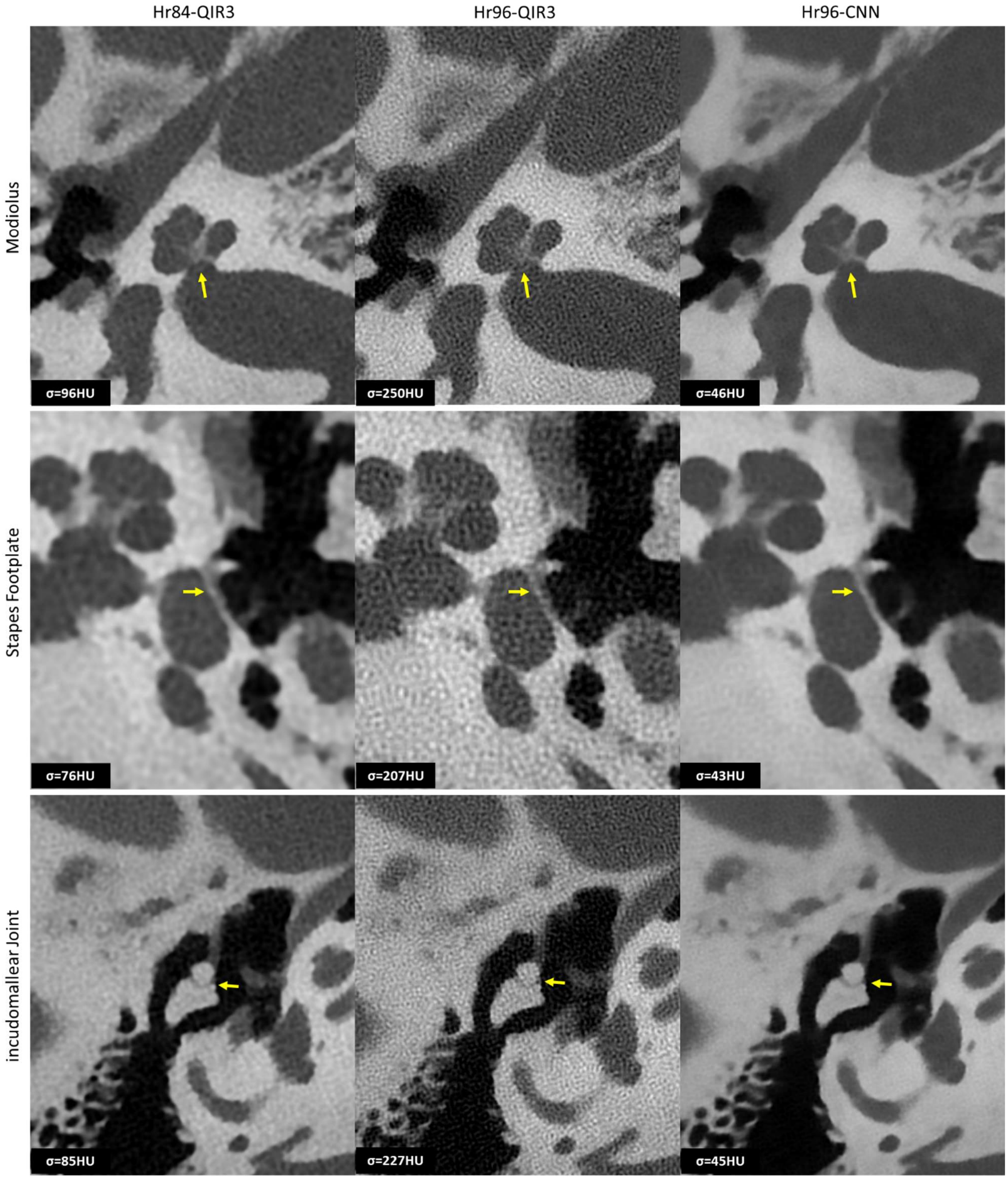

FIG 4. Representative images of the modiolus, stapes footplate, and incudomallear joint using three different reconstructions: Hr84-QIR3, Hr96-QIR3, and Hr96-CNN (W/L:4000/1000 HU). Enhanced visualization with improved resolution and reduced noise is demonstrated, with significant details indicated by yellow arrows. Image noise was quantified by measuring the standard deviation of CT numbers within a circular ROI placed in a uniform soft-tissue area, with values recorded in the lower left corner of each image.

Reader Evaluation

Figure 5 shows the results of the reader study. For overall image quality, Hr96-CNN images were ranked significantly higher than both Hr84-QIR3 ($p<0.001$) and Hr96-QIR3 ($p<0.001$). Both readers preferred CNN denoising images for visualization of all three anatomical structures: the modiolus (Hr96-CNN/Hr84-QIR3/HR96-QIR3: 1/1.8/2.8, $p<0.001$), the stapes footplates (Hr96-CNN/Hr84-QIR3/HR96-QIR3: 1/1.94/2.88, $p<0.001$), and the incudomallear joint (Hr96-CNN/Hr84-QIR3/HR96-QIR3: 1/1.64/2.53, $p<0.001$). It is noted that in both evaluations, equal ranking was allowed.

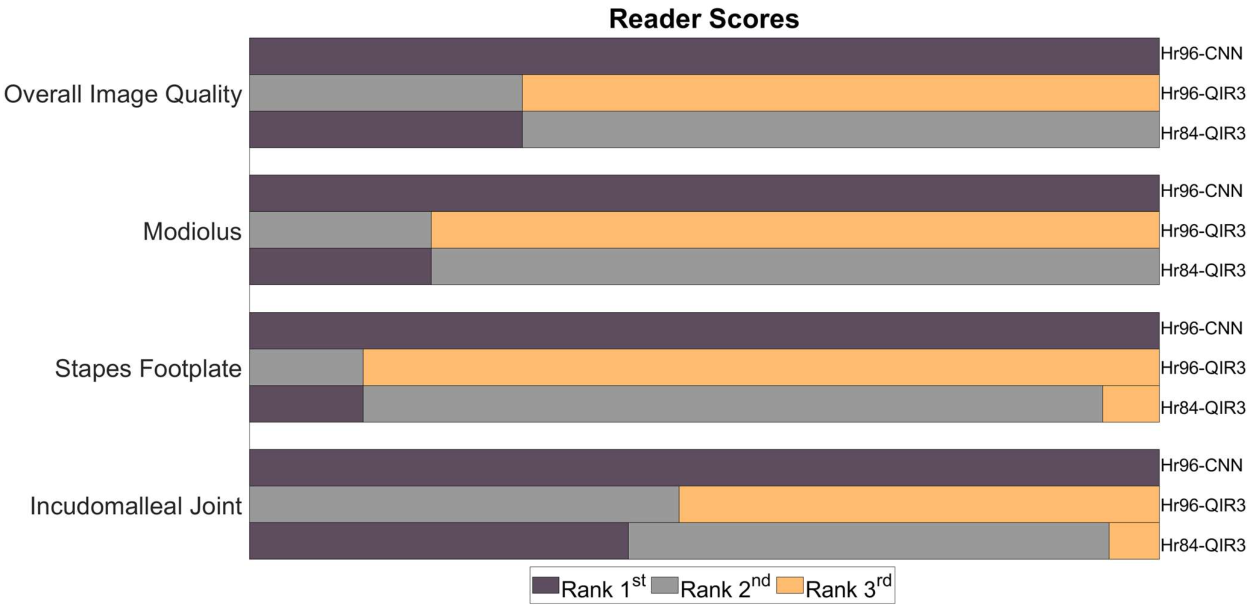

FIG 5. Rankings from 2 readers regarding overall image quality and delineation of 3 key anatomic structures. For all 3 structures and overall image quality, CNN-Hr96 images rank highest. Dull purple indicates the first rank; Med gray, the second rank; Gold, the third rank.

Denoising Performance Comparison

Figure 6 presents an example slice from one subject in the test dataset, processed using the conventional Hr96-QIR3, RED-CNN, and the proposed U-Net, along with their corresponding difference images when compared to the reference input. Both the conventional and CNN-based denoising methods successfully reduce noise relative to the reference. Specifically, noise levels were reduced from 572 HU to 235 HU, 53 HU, and 43 HU using Hr96-QIR3, RED-CNN, and the proposed U-Net, respectively. As demonstrated in the difference images, the CNN-based denoising methods primarily remove noise, whereas the conventional method (Hr96-QIR3) also removes subtle structures, as indicated by the yellow arrow. This observation is further supported by the SSIM values of 0.7045, 0.9865, and 0.9883 for Hr96-QIR3, RED-CNN, and U-Net, respectively. Moreover, the comparison between RED-CNN and U-Net indicates that the proposed data preparation workflow performs effectively across different network architectures, achieving satisfactory denoising performance while preserving fine structural details. At the patient-level comparison, shown in Table 1, the proposed U-Net notably improves image quality. Compared to the conventional method (Hr96-QIR3), the U-Net significantly reduces image noise (from 204.63 HU to 47.35 HU) and enhances the SSIM (from 0.72 to 0.99).

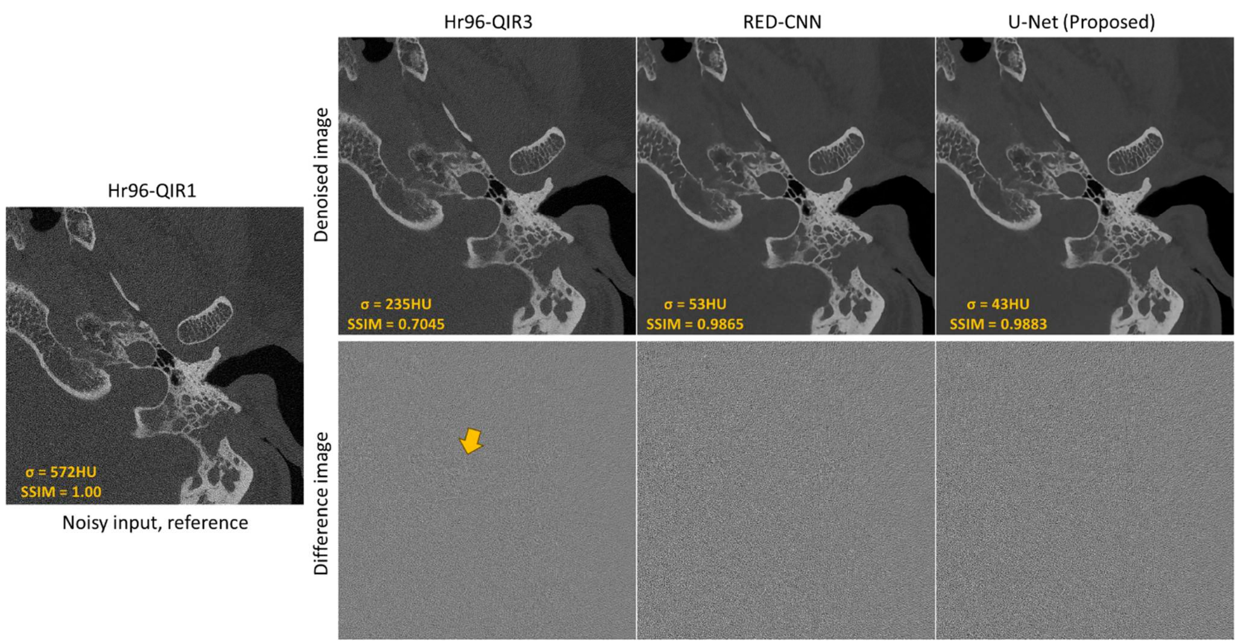

FIG 6. Example slice from the test dataset processed with Hr96-QIR3, RED-CNN, and the proposed U-Net, alongside their corresponding difference images compared to the reference input. The CNN-based methods (RED-CNN and U-Net) primarily reduce noise, while the conventional Hr96-QIR3 also removes subtle anatomical structures (highlighted by the yellow arrow). The display

window is (W/L:4000/1000 HU) for patient images, and (3000/1000 HU) for difference images.

Table 1: Quantitative comparison (mean ± standard deviation) at patient level across the conventional method (Hr96-QIR3), RED-CNN and U-Net.

Quality metrics	Hr96-QIR1	Hr96-QIR3	RED-CNN	U-Net (Proposed)
Image Noise	500.15 ± 52.38	204.63 ± 21.70	52.82 ± 1.17	47.35 ± 2.62
SSIM	1.00 ± 0.00	0.72 ± 0.05	0.98 ± 0.02	0.99 ± 0.01

DISCUSSION

In this work, the proposed CNN significantly reduced image noise in UHR PCD-CT, allowing the use of the sharpest kernel with acceptable noise levels, unlocking the full potential of the UHR PCD-CT system. This combination of CNN denoising and UHR PCD-CT substantially enhanced diagnostic image quality and improved visualization of critical anatomical structures.

Previous research[21,22] demonstrated that spatial resolution of PCD-CT was not fully utilized in routine practice. Graafen et al.[10] investigated the impact of kernel sharpness on image quality and concluded that soft reconstruction kernels yield the best overall quality for the evaluation of hepatocellular carcinoma in PCD-CT. The primary reason is the extensive noise associated with sharper kernels, which can compromise diagnostic clarity. Although some studies have employed noise reduction techniques, including deep learning methods, to denoise UHR PCD-CT images, these efforts have not maximized the use of the sharpest available kernels. Our approach utilizing CNN denoising allows for the use of the sharpest kernel in UHR PCD-CT, maintaining acceptable noise levels and broadening its clinical applicability. The proposed method is fully based on images from patients' routine clinical exams, without the need for additional low-dose images or any proprietary information. This flexibility allows the method to be adapted to any scanner. Furthermore, noise-only images in the workflow were generated from the same kernel reconstructions at different strengths, making it a kernel-based approach that can be applied to both smooth and sharp kernels for various imaging tasks, such as coronary and abdominal imaging. Additionally, the weight factor in the workflow allows for control over the level of noise reduction, accommodating the reader's preference for noise acceptability.

This proof-of-concept study has several limitations. Firstly, the sample size of twenty testing patient cases was relatively small. Future research will expand the sample size for more robust statistical analysis and validation of findings. Secondly, the study focused solely on demonstrating image quality improvement without targeting specific diagnostic tasks. This represents the initial step in demonstrating the feasibility and potential of the algorithm. Future studies on diagnostic performance and clinical significance are warranted, including assessing how denoising affects the visibility of pathologies. Furthermore, the proposed CNN denoising offers significant potential for radiation dose reduction while maintaining clinically acceptable image quality. As spatial resolution, image noise, and radiation dose are interrelated to each other, the CNN denoising can be used to either improve spatial resolution, reduce image noise, or reduce radiation dose, or a combination.

CONCLUSIONS

In summary, we developed and evaluated a dedicated deep learning-based denoising method for UHR PCD-CT. This method utilizes a training dataset sourced from commercially available images, requiring no additional data preparation. The application of this algorithm in temporal bone imaging shows high-resolution and low-noise images with improved anatomical delineation. This advancement significantly enhances temporal bone visualization by fully utilizing the spatial resolution capabilities of PCD-CT.

ACKNOWLEDGMENTS

Research reported in this work was supported by the National Institutes of Health under award number (anonymized for review only). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Health.

REFERENCES

1. Lane JI, Lindell EP, Witte RJ, DeLone DR, Driscoll CL. Middle and inner ear: improved depiction with multiplanar reconstruction of volumetric CT data. *Radiographics*. 2006;26(1):115-124.
2. Goldfeld M, Glaser B, Nassir E, Gomori JM, Hazani E, Bishara N. CT of the ear in Pendred syndrome. *Radiology*. 2005;235(2):537-540.
3. Zhou W, Lane JI, Carlson ML, et al. Comparison of a Photon-Counting-Detector CT with an Energy-Integrating-Detector CT for Temporal Bone Imaging: A Cadaveric Study. *AJNR Am J Neuroradiol*. 2018;39(9):1733-1738.
4. 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.
5. Leng S, Rajendran K, Gong H, et al. 150-μm Spatial Resolution Using Photon-Counting Detector Computed Tomography Technology: Technical Performance and First Patient Images. *Invest Radiol*. 2018;53(11):655-662.
6. Rajendran K, Petersilka M, Henning A, et al. First Clinical Photon-counting Detector CT System: Technical

Evaluation. *Radiology*. 2022;303(1):130-138.
7. Willemink MJ, Grist TM. Counting Photons: The Next Era for CT Imaging? *Radiology*. 2022;303(1):139-140.
8. Leng S, Yu Z, Halaweish A, et al. Dose-efficient ultrahigh-resolution scan mode using a photon counting detector computed tomography system. *J Med Imaging (Bellingham)*. 2016;3(4):043504.
9. Leng S, Bruesewitz M, Tao S, et al. Photon-counting Detector CT: System Design and Clinical Applications of an Emerging Technology. *Radiographics*. 2019;39(3):729-743.
10. Graafen D, Müller L, Halfmann MC, et al. Soft Reconstruction Kernels Improve HCC Imaging on a Photon-Counting Detector CT. *Acad Radiol*. 2023;30 Suppl 1:S143-s154.
11. Patzer TS, Kunz AS, Huflage H, et al. Ultrahigh-Resolution Photon-Counting CT in Cadaveric Fracture Models: Spatial Frequency Is Not Everything. *Diagnostics (Basel)*. 2023;13(10).
12. Mergen V, Sartoretti T, Baer-Beck M, et al. Ultra-High-Resolution Coronary CT Angiography With Photon-Counting Detector CT: Feasibility and Image Characterization. *Invest Radiol*. 2022;57(12):780-788.
13. Chen GH, Tang J, Leng S. Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. *Med Phys*. 2008;35(2):660-663.
14. Huber NR, Ferrero A, Rajendran K, et al. Dedicated convolutional neural network for noise reduction in ultra-high-resolution photon-counting detector computed tomography. *Phys Med Biol*. 2022;67(17).
15. Chang S, Huber NR, Marsh JF, et al. Pie-Net: Prior-information-enabled deep learning noise reduction for coronary CT angiography acquired with a photon counting detector CT. *Med Phys*. 2023;50(10):6283-6295.
16. Niu C, Li M, Fan F, et al. Noise Suppression With Similarity-Based Self-Supervised Deep Learning. *IEEE Trans Med Imaging*. 2023;42(6):1590-1602.
17. Huber NR, Ferrero A, Rajendran K, et al. Dedicated convolutional neural network for noise reduction in ultra-high-resolution photon-counting detector computed tomography. *Physics in Medicine & Biology*. 2022.
18. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Paper presented at: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015; 2015//, 2015; Cham.
19. Kingma D, Ba J. Adam: A Method for Stochastic Optimization. International Conference on Learning Representations; 12/22, 2014.
20. Sadia RT, Chen J, Zhang J. CT image denoising methods for image quality improvement and radiation dose reduction. *Journal of Applied Clinical Medical Physics*. 2024;25(2):e14270.
21. Mannil M, Hickethier T, von Spiczak J, et al. Photon-Counting CT: High-Resolution Imaging of Coronary Stents. *Invest Radiol*. 2018;53(3):143-149.
22. Baffour FI, Rajendran K, Glazebrook KN, et al. Ultra-high-resolution imaging of the shoulder and pelvis using photon-counting-detector CT: a feasibility study in patients. *Eur Radiol*. 2022;32(10):7079-7086.

SUPPLEMENTAL FILES

CLAIM: Checklist for Artificial Intelligence in Medical Imaging

Section / Topic	No.	Item	
TITLE / ABSTRACT			
	1	Identification as a study of AI methodology, specifying the category of technology used (e.g., deep learning)	✓
	2	Structured summary of study design, methods, results, and conclusions	✓
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the AI approach	✓
	4	Study objectives and hypotheses	✓
METHODS			
Study Design	**5**	Prospective or retrospective study	✓
	6	Study goal, such as model creation, exploratory study, feasibility study, non-inferiority trial	✓
Data	**7**	Data sources	✓

	8	Eligibility criteria: how, where, and when potentially eligible participants or studies were identified (e.g., symptoms, results from previous tests, inclusion in registry, patient-care setting, location, dates)	✓
	9	Data pre-processing steps	✓
	10	Selection of data subsets, if applicable	✓
	11	Definitions of data elements, with references to Common Data Elements	✓
	12	De-identification methods	✓
	13	How missing data were handled	**N/A**
Ground Truth	**14**	Definition of ground truth reference standard, in sufficient detail to allow replication	**N/A**
	15	Rationale for choosing the reference standard (if alternatives exist)	✓
	16	Source of ground-truth annotations; qualifications and preparation of annotators	**N/A**
	17	Annotation tools	**N/A**
	18	Measurement of inter- and intrarater variability; methods to mitigate variability and/or resolve discrepancies	✓
Data Partitions	**19**	Intended sample size and how it was determined	✓
	20	How data were assigned to partitions; specify proportions	✓
	21	Level at which partitions are disjoint (e.g., image, study, patient, institution)	✓
Model	**22**	Detailed description of model, including inputs, outputs, all intermediate layers and connections	✓
	23	Software libraries, frameworks, and packages	✓
	24	Initialization of model parameters (e.g., randomization, transfer learning)	✓
Training	**25**	Details of training approach, including data augmentation, hyperparameters, number of models trained	✓
	26	Method of selecting the final model	✓
	27	Ensembling techniques, if applicable	**N/A**
Evaluation	**28**	Metrics of model performance	✓
	29	Statistical measures of significance and uncertainty (e.g., confidence intervals)	✓
	30	Robustness or sensitivity analysis	**N/A**
	31	Methods for explainability or interpretability (e.g., saliency maps), and how they were validated	**N/A**
	32	Validation or testing on external data	**N/A**
RESULTS			
Data	**33**	Flow of participants or cases, using a diagram to indicate inclusion and exclusion	✓
	34	Demographic and clinical characteristics of cases in each	**N/A**

		partition	
Model performance	**35**	Performance metrics for optimal model(s) on all data partitions	✓
	36	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	**N/A**
	37	Failure analysis of incorrectly classified cases	**N/A**
DISCUSSION			
	38	Study limitations, including potential bias, statistical uncertainty, and generalizability	✓
	39	Implications for practice, including the intended use and/or clinical role	✓
OTHER INFORMATION			
	40	Registration number and name of registry	**N/A**
	41	Where the full study protocol can be accessed	✓
	42	Sources of funding and other support; role of funders	✓

Mongan J, Moy L, Kahn CE Jr. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): a guide for authors and reviewers. Radiol Artif Intell 2020; 2(2):e200029. https://doi.org/10.1148/ryai.2020200029