



Published in final edited form as:

Med Phys. 2022 February ; 49(2): 901–916. doi:10.1002/mp.15414.

Reconstruction of three-dimensional tomographic patient models for radiation dose modulation in CT from two scout views using deep learning

Juan C. Montoya*

Department of Medical Physics, University of Wisconsin School of Medicine and Public Health, Madison, WI 53705, USA

Chengzhu Zhang*

Department of Medical Physics, University of Wisconsin School of Medicine and Public Health, Madison, WI 53705, USA

Yinsheng Li,

Department of Medical Physics, University of Wisconsin School of Medicine and Public Health, Madison, WI 53705, USA

Ke Li,

Department of Medical Physics, University of Wisconsin School of Medicine and Public Health, Madison, WI 53705, USA

Department of Radiology, University of Wisconsin School of Medicine and Public Health, Madison, WI 53792, USA

Guang-Hong Chen

Department of Medical Physics, University of Wisconsin School of Medicine and Public Health, Madison, WI 53705, USA

Department of Radiology, University of Wisconsin School of Medicine and Public Health, Madison, WI 53792, USA

Abstract

Background: A tomographic patient model is essential for radiation dose modulation in x-ray computed tomography (CT). Currently, two-view scout images (also known as topograms) are used to estimate patient models with relatively uniform attenuation coefficients. These patient models do not account for the detailed anatomical variations of human subjects, and thus may limit the accuracy of intra-view or organ-specific dose modulations in emerging CT technologies.

Purpose: The purpose of this work was to show that 3D tomographic patient models can be generated from two-view scout images using deep learning strategies, and the reconstructed 3D

Corresponding Author: Guang-Hong Chen Department of Medical Physics and Department of Radiology, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI 53705, USA, gchen7@wisc.edu.

***Note:** J. C. Montoya and C. Zhang equally contributed to this work and are co-first authors

CONFLICTS OF INTEREST

The authors have no conflicts to disclose.

patient models indeed enable accurate prescriptions of fluence-field modulated or organ-specific dose delivery in the subsequent CT scans.

Methods: CT images and the corresponding two-view scout images were retrospectively collected from 4,214 individual CT exams. The collected data were curated for the training of a deep neural network architecture termed ScoutCT-NET to generate 3D tomographic attenuation models from two-view scout images. The trained network was validated using a cohort of 55,136 images from 212 individual patients. To evaluate the accuracy of the reconstructed 3D patient models, radiation delivery plans were generated using ScoutCT-NET 3D patient models and compared with plans prescribed based on true CT images (gold-standard) for both fluence-field modulated CT and organ-specific CT. Radiation dose distributions were estimated using Monte Carlo simulations and were quantitatively evaluated using the Gamma analysis method. Modulated dose profiles were compared against state-of-the-art tube current modulation schemes. Impacts of ScoutCT-NET patient model-based dose modulation schemes on universal-purpose CT acquisitions and organ-specific acquisitions were also compared in terms of overall image appearance, noise magnitude, and noise uniformity.

Results: The results demonstrate that (1) The end-to-end trained ScoutCT-NET can be used to generate 3D patient attenuation models and demonstrate empirical generalizability. 2) The 3D patient models can be used to accurately estimate the spatial distribution of radiation dose delivered by standard helical CTs prior to the actual CT acquisition; compared to the gold-standard dose distribution, 95.0% of the voxels in the ScoutCT-NET based dose maps have acceptable gamma values for 5 mm distance-to-agreement and 10% dose difference. 3) The 3D patient models also enabled accurate prescription of fluence-field modulated CT to generate a more uniform noise distribution across the patient body compared to tube current modulated CT. 4) ScoutCT-NET 3D patient models enabled accurate prescription of organ-specific CT to boost image quality for a given body region-of-interest under a given radiation dose constraint.

Conclusion: 3D tomographic attenuation models generated by ScoutCT-NET from two-view scout images can be used to prescribe fluence-field modulated or organ-specific CT scans with high accuracy for the overall objective of radiation dose reduction or image quality improvement for a given imaging task.

Keywords

Deep learning; image reconstruction; fluence-field modulated CT; low dose CT; sparse-view reconstruction

I. INTRODUCTION

The year 2021 marks the 50th anniversary of the first CT scan of a human patient. Over a half-century journey, CT imaging has demonstrated an unprecedented positive impact on the diagnosis and treatment of symptomatic patients. Continuous technological advances in CT have made it an integral component of modern medicine.¹ However, the issue of the small, but non-zero, theoretical cancer risk associated with the use of ionizing radiation in CT remains a public concern despite its invaluable contributions to modern patient care.^{2,3} To address the public concerns regarding ionizing radiation risks due to the use of CT in diagnosis and image-guided therapeutic procedures, low dose CT and radiation dose

optimization have become a key task in clinical medical physics practices and a central topic of research in the CT imaging community.^{4,5}

The research and development efforts from academic, industrial, and clinical societies in the past decade have resulted in the development of a variety of radiation dose reduction strategies for CT exams. Software-related techniques include noise reduction reconstruction or image processing algorithms that operate on either raw data, reconstructed image data, or both.^{6–13} The initial anticipation for these iterative reconstruction (IR) algorithms was to reduce radiation dose by 60–70%, with the potential for even further reduction.^{14,15} However, recent clinical results have concluded that, after a decade-long effort to reduce CT radiation dose with IR software techniques, the status quo is up to 25% dose reduction with IR techniques, far short of the originally expected 60–70% or even 90% claims.^{16,17} On the other hand, hardware improvements aim to effectively improve the dose efficiency of CT systems. Namely, given the necessary diagnostic image performance with the improved dose efficiency offered by the scanner hardware, the required radiation dose to a patient can be reduced with improvements in hardware. Current hardware dose reduction strategies include 1) newer generations of CT detectors (such as photon-counting detectors) with improved radiation dose efficiency,^{18–24} 2) beam shaping filters and dynamic z-axis collimators that limit dose at the periphery of the patient and edge of the scan range²⁵, 3) automatic exposure control systems with the flexibility to modulate tube potential and tube current along the projection view angle direction.^{26–29}, and 4) intra-view fluence-field modulation schemes and devices have been extensively studied in recent years to better account for the heterogeneous nature of human anatomy to further reduce dose.^{30–40}

Earlier researches on intra-view fluence-field modulated CT demonstrated a promising 30–50% radiation reduction compared to the current standard-of-care when the entrance fluence field is properly modulated based on the x-ray attenuation properties of the image objects.^{32,33,35} The success of fluence-field modulated CT-based dose reduction relies on the availability of an accurate 3D tomographic attenuation model for prescription of a personalized dose modulation scheme before the actual CT scan is executed. In the current clinical CT imaging workflow, the needed patient attenuation information for dose modulations is estimated from the planar scout images (also known as topograms) acquired at two orthogonal view angles (e.g., 90° and 180°) under a scanning-beam acquisition to overcome the narrow coverage of the CT detector along the superior-inferior (SI) direction. During a scout scan, the x-ray beam is narrowly collimated (e.g. to 5 mm); the patient table translates continuously while the detector is read out at a uniform frame rate. The recorded frames are stitched together based on the frame rate and table speed and then scaled to generate the overall scout image that covers the desired patient anatomy for the purpose of patient positioning and generation of the patient model for dose modulations. After the scout images are acquired, patient attenuation information for other view angles can be estimated by interpolating two-view scout images based on a cosine function or other models.⁴¹ However, these estimation strategies do not account for the heterogeneous patient attenuation distributions and thus are not able to provide sufficiently accurate patient information necessary for prescribing more detailed and accurate dose modulations. The concept of ultra-low-dose “3D CT scout” has been investigated^{42,43} to obtain the needed 3D patient models for prescribing intra-view or organ-specific CT dose modulations.

Despite its conceptual simplicity, a wide adoption of 3D CT scout strategies in practice encountered several challenges: 1) Additional radiation dose is introduced to the patient by this additional CT acquisition; 2) Noise streaks caused by strong photon starvation in these ultra-low-dose CT are often too severe to allow reliable estimations of patient- and organ-specific attenuation distributions; 3) The quantitative accuracy of CT numbers, which is important for the calculation of radiologic paths in intra-view dose modulation schemes, is significantly compromised at ultra-low-dose levels.^{44,45} Besides using 3D CT scout images and two-view 2D scouts, tube current modulations can also be prescribed using online feedback of the previously-measured 180° conjugate projection data at a different but nearby z-location (helical or step-and-shoot axial acquisitions) or at the same z-location but at a prior time (dynamic CT scans). These inter-view tube current modulation schemes have been successfully implemented by different CT manufacturers to enable their own vendor-specific tube current modulation schemes to reduce dose while maintaining the overall image quality.^{46,47} The limitation of the online feedback strategy lies in the fact that it is not applicable to most task- or organ-based dose deliveries which require the optimization of the modulation schemes ahead of the actual CT scan. Additionally, for scan applications with a single axial gantry rotation (e.g., coronary CT performed on 256- or 320-slice scanners with 16 cm z coverage), the online feedback strategy is not applicable even for inter-view tube current modulations.

In summary, there remains an unmet need for a robust method to generate 3D tomographic patient models to account for the heterogeneous attenuation distributions of individual patients. It is highly desirable to generate a 3D tomographic patient attenuation model from two 2D scout images to leverage its clinically favorable features: super low radiation dose (about 0.04mGy) and broad anatomical coverage along the SI direction. To accomplish this overall objective, the following two scientific challenges must be addressed: 1) Technically, what is the reconstruction method that enables tomographic reconstruction using two scout images stitched together from many small pieces? 2) If a technical solution is available, given the severe ill-posedness of the two-view reconstruction problem itself, it is anticipated that the reconstruction accuracy may be limited. Therefore, the practical value of the reconstructed 3D patient attenuation model in radiation dose modulations must be evaluated to carefully examine whether the 3D patient model generated from two-view scout images provides sufficient information for the purpose of prescribing intra-view or organ-specific dose-modulated CT.

Regarding the technical feasibility to reconstruct a 3D tomographic attenuation patient model from two projection view angles, it would be extremely challenging, if not impossible, to use traditional analytical or iterative image reconstruction strategies due to the difficulty of hand-crafting the appropriate regularizers to regularize the severe ill-posedness of this reconstruction problem. However, recent advances in data-driven deep learning approaches offer a new hope to regularize ill-posed inverse problems by learning the needed regularizers from well-curated training data sets. As a matter of fact, by leveraging the improved regression capabilities of deep neural networks, deep learning-based methods have been introduced into the CT image processing and reconstruction field in recent years and have demonstrated immense promise towards addressing difficult problems previously not solvable with traditional analytical and iterative methods. Earlier works focused primarily

on the use of image-to-image neural network models under a supervised learning paradigm to reduce image noise.^{48–53} Recent works in the context of image reconstruction have integrated model-based iterative reconstruction methods with neural networks via learned regularizers.^{54–58} Direct learning of CT reconstruction from projection data acquired under a wide variety of conditions has also been demonstrated: Li *et al.* showed promising results for low-dose CT and sparse-view CT reconstructions,⁵⁹ but not under the extreme condition of only one or two projection views.

Recently, several groups have pioneered tomographic reconstruction from one or two views projection images.^{60–63} It has been shown that under the condition that a prior CT or cone-beam CT volume of the same patient is available, digitally reconstructed radiographs (DRRs) can be generated from prior CT volumes by a numerical forward cone-beam projection operation. These DRRs and their corresponding image volumes can be used to generate synthetic training data pairs to train a variety of deep neural network architectures to accomplish 3D reconstruction from one or two synthetic views of projection data. For example, in the context of four-dimensional cone-beam CT (4D CBCT), using the DRRs and the corresponding image volumes from six out of ten respiratory phases, a deep learning model was trained to reconstruct 3D image volumes using DRRs from a single view angle.⁶⁴ The performance of the proposed reconstruction strategy was evaluated by applying the trained model to the DRRs of the remaining four respiratory phases of the same patient. However, it has not been shown that the proposed strategy can be used to reconstruct 4D CBCT images from the experimentally acquired cone-beam projection data yet. It is also unclear how the proposed strategy would generalize from one patient to another. In another interesting work, Ying *et al.*⁶⁵ trained a deep learning model termed the X2CT-GAN network to reconstruct a CT image volume from DRRs. Due to the differences between DRRs and the available chest x-ray radiography images in public databases that were acquired using x-ray cone-beam image acquisitions, the authors used a cycle-GAN to perform style transfer from available chest x-ray radiography images to DRRs and then used the output DRRs to reconstruct the CT-like images. Due to the differences between the slit-scanning acquisitions used in scout imaging⁶⁶ and the cone-beam acquisitions used in x-ray radiography, the authors in Ref. 65 were not able to generate CT-like images from actual scout images. Additionally, there is no reference CT image volume of the patient available to evaluate the accuracy of the reconstructed CT-like images. Similarly, Henzler *et al.* also trained a network using paired DRRs and CT images to generate volumetric images and rendering of high-contrast structures (e.g. skulls) inside the image objects from single-view DRR images.⁶⁷ However, there is no soft tissue reconstruction demonstrated yet in their work. In another work by Liu *et al.*,⁶⁸ the AUTOMAP neural network⁶⁹ was adapted for sparse-view reconstruction using two-view or four-view DRRs. Some good initial results have also been achieved despite the encountered challenges of severe organ distortions in reconstructed images as pointed out by the authors.

To summarize, despite the amazing progress made, the aforementioned two scientific questions remain to be answered in an attempt to reconstruct a 3D patient attenuation model for radiation dose prescription in CT. This motivated the purpose of this work: 1) To develop a deep learning-based method referred to as ScoutCT-NET to generate 3D patient attenuation models directly from two-view scout images acquired in clinical CT exams; 2)

To demonstrate the trained ScoutCT-NET can be used to directly reconstruct patient-specific 3D patient models from 2D scout images, and 3) To demonstrate both intra-view dose modulations and organ-specific dose modulations can be accurately prescribed based on ScoutCT-NET 3D patient models.

II. METHODS

A. ScoutCT-NET Image Formation – Problem Formulation

Classical statistical image tomographic reconstruction techniques aim to solve the following optimization problem:

$$\vec{x}^* = \underset{\vec{x}}{\operatorname{argmin}} [L(\vec{y} - A\vec{x}) + \lambda R(\vec{x})], \quad (1)$$

where \vec{x} denotes a vectorized image matrix to be reconstructed, \vec{y} denotes the vectorized projection data, and A denotes the CT system matrix. In general, $L(\vec{y} - A\vec{x})$ is considered a function that quantifies data consistency (e.g weighted least squares), and $\lambda R(\vec{x})$ represents a regularization function whose strength is controlled by the parameter λ . From a statistical inference perspective, the classical statistical image reconstruction techniques aimed to seek for a set of statistical parameters that maximize the *a posteriori* probability.

In contrast, the central objective of ScoutCT-NET aims to directly learn a transform G to generate a 3D attenuation model of each patient using two orthogonal-view scout images as inputs. To parameterize the desired target transform G , a deep neural network was used to leverage its powerful expression capacities. The proposed two-view ScoutCT-NET 3D patient model estimation problem is formulated as the following optimization problem:

$$G^* = \underset{G}{\operatorname{argmin}} \underset{D}{\operatorname{max}} \left\{ \underset{\vec{y}}{\mathbb{E}} \left[\|\vec{y} - AG(\vec{y})\|_2^2 \right] + \lambda \underset{\vec{y}, \vec{x}_p}{\mathbb{E}} \left[\|\vec{x}_p - G(\vec{y})\|_1 \right] + \underset{\vec{x}_p}{\mathbb{E}} \left[\ln D(\vec{x}_p) \right] + \underset{\vec{y}}{\mathbb{E}} [1 - \ln D(G(\vec{y}))] \right\} \quad (2)$$

In the above equation, \mathbb{E} denotes the expected value using the empirical statistical distribution, \vec{y} is the vector representation of the input two-view scout images, the 3D model of the image object is given by $\vec{x} = G(\vec{y})$, and \vec{x}_p denotes true CT images for the purpose of learning the neural network transform G . The factor λ is a scalar that controls the strength of the $L1$ regularization term and D represents a discriminator deep neural network that serves as a learned regularizer to enforce G to approximate the distribution of \vec{x}_p . The needed mapping $\vec{x} = G(\vec{y})$ can be learned in an end-to-end manner using the well-established adversarial neural network training methodology, in which G aims to minimize the data consistency loss and the $L1$ -norm between the resulting image and the prior image; at the same time, a cross-entropy loss is maximized to enforce G to

approximate the distribution of \vec{x}_p . This prevents the model from simply learning an identity function that only minimizes a data consistency loss.

B. Neural Network Architecture and Implementation

The overall architecture of ScoutCT-NET is shown in Fig. 1. In this architecture, the deep neural network G in Eq. (2) takes a 888×5 section from each of the two scout images as input to generate a sub-volume of tomographic images with $288 \times 288 \times 5$ voxels. This is based on the fact that in reality, scout images are acquired under a narrow-beam scanning mode on MDCT scanners. This is different from the cone-beam acquisition geometry used by most x-ray radiography and fluoroscopy systems. In order to map the narrow section of the scout images onto the 3D image space, the first layer of the network consists of a direct backprojection operation which is then followed by 24 convolutional layers that are arranged in a simplified U-Net architecture with four vertical levels⁷⁰. The network was trained using a conditional adversarial loss^{71,72} with a 24-layer convolutional discriminator using a ResNet architecture^{73,74}.

The exact neural network layer configuration is as follows: the U-Net encoder uses C64C64-sc1-MPs2-C128-C128-sc2-MPs2-C256-C256-sc3-MPs2-C512-C512-sc4-MPs2-C1024-C1024, where Ck stand for a convolutional layer with k filters, sc_i represents a skip connection at level i and $MPs2$ represents a max-pooling operation with stride 2. Similarly, the U-Net decoder layers are as follows: TC512-sc4-C512-C512-TC256-sc3-C256-C256-TC128-sc2-C128-C128-TC64-sc1-C64-C64-C64-C5, where TCk represents a transpose convolution operation with k filters. All convolutional layers use a 3×3 spatial filter configuration with ReLU activations and no batch normalization. Backprojection and forward projection layers are implemented as fully connected layers with weights predefined by the system geometry and no trainable parameters. The ResNet configuration for the discriminator network used only during training is the following: C32s2-C32-C64-sc-C64-C64-sc-C64-C64-sc-C64C64-sc-(C128s2,AP-C128)-C128-sc-C128-C128-sc-C128-C128-sc-(C256s2, AP-C256)-C256sc-C256-C256-sc-C256-C256-sc-FC1, where Ck stand for a convolutional layer with k filters and $Cks2$ represents a convolutional layer with k filters and stride 2. sc represents the characteristic skip connection of the ResNet architecture. Finally, the $AP-Ck$ notation represents an auxiliary layer consisting of a 3×3 average pooling operation with stride 2 followed by a convolutional layer of $1 \times 1 \times k$ filter to match the size and number of channels when spatial downsampling occurs in the network. The final FC1 layer represents a fully connected layer with one output. The first layer uses a 5×5 receptive field and all other convolutional layers use a 3×3 spatial filter configuration with ReLU activations and no batch normalization.

The model was implemented using TensorFlow⁷⁵ version 1.3. Network parameters were initialized using the variance scaling method⁷⁶ and trained from scratch using stochastic gradient descent with the Adam optimizer with a learning rate of 2.0×10^{-4} , momentum parameters of $\beta_1 = 0.9$ and $\beta_2 = 0.999$ and batch size of 32. To avoid exploding gradients, we adopted gradient clipping in the range $[-1.0, 1.0]$. The model was pretrained for 4 epochs to optimize the generator loss and then we used early stopping when the adversarial loss

reached equilibrium in the testing cohort after an additional 4 epochs. The model was trained using two GTX 1080 Ti (NVIDIA, CA) GPUs.

C. Training data collection and curation

More than 1.1 million CT images and their corresponding scout image data from 4,214 clinically indicated CT studies were retrospectively collected. The dataset was used to train the ScoutCT-NET in order to reconstruct 3D patient attenuation models from 2-view scout images. Once the ScoutCT-NET model was trained, it was used to generate 3D attenuation models for patients in the testing cohort. Based on the 3D attenuation model, intra-view or organ-based dose modulation plans were prescribed and virtually delivered. The overall study schema is shown in Fig. 2.

1. Data collection—All image data were deidentified and collected under an Institutionally Review Board approved protocol. The requirement for informed consent was waived for this retrospective study. Scout and CT images from 4,214 clinically indicated CT exams acquired over the period of nine months at our institution were retrospectively collected from PACS. Inclusion criteria were 1) CT exams acquired without contrast media, 2) Scanned with the anatomical region of the chest, and 3) Patients scanned in a supine position. To avoid potential data inconsistency, CT exams that presented significant image artifacts from metallic objects or data truncation caused by anatomical regions outside the scanning field of view were excluded.

2. Data curation—Each CT exam included in this work consisted of two orthogonal-view scout images and one volume of CT images. Scout and CT images are typically acquired under the same tube potential to assure consistent attenuation measurements in the scout images for the purpose of dose-modulated CT prescriptions. The two CT scanner models used to acquire the data are Discovery CT750 HD and Revolution CT (GE Healthcare, Waukesha, WI). Tube potentials used to acquire the data range from 70 to 140 kV and were determined based on the scan protocol and actual patient size. In order to account for possible differences in CT image reconstruction parameters (i.e. image thickness, voxel size and field of view position), each CT image volume was translated and scaled to the CT system global coordinates using an affine transformation. The matrix size of every image was 288×288 with $1.89 \text{ mm} \times 1.89 \text{ mm} \times 1.00 \text{ mm}$ voxel size. There is no special reason for using a matrix size of 288×288 ; instead, it was just a convention used in our work and one can also choose to use a matrix size of 256×256 if it is preferred. However, the physical size of the image voxel needs to accommodate the following practical considerations: 1) First is to reduce the computational burden of Monte Carlo simulations; 2) Second is to meet the condition of charged particle equilibrium (CPE) for dose calculations. Note that the range of the secondary electrons for the energies used in diagnostic CT is typically less than 0.5 mm in human tissue, thus the choice of a $1.89 \text{ mm} \times 1.89 \text{ mm} \times 1.00 \text{ mm}$ voxel size satisfies the CPE condition.

3. Data partition for training, validation, and testing—Data were split into three cohorts consisting of 3,790 (90%), 212 (5%) and 212 (5%) of the total number of CT studies for training, validation and testing of the ScoutCT-NET deep learning model respectively.

The validation and testing cohorts were randomly selected from the subset of patients who underwent a single CT exam with one study per exam in order to avoid potential overlap between the training and validation/testing data.

D. Quantitative evaluation of 3D patient attenuation model generated from ScoutCT-NET (I): Comparison of radiation dose distribution generated from the ScoutCT-NET 3D patient model in Monte Carlo dose simulations

After training, ScoutCT-NET was used to process scout images from the testing cohort to generate 3D patient models. As the first step of evaluating the ScoutCT-NET based 3D patient models, standard helical CT scans were virtually delivered to both the 3D patient models and the corresponding gold-standard patient models to evaluate whether the absorbed dose distributions in the ScoutCT-NET based attenuation model match that of the gold standard. The virtual CT scan was performed using the Monte Carlo platform (MC-GPU version 1.3).⁷⁷ A standard chest CT with a helical pitch of 1.375, a tube potential of 120 kV, 4 mm of Al filtration, and a scanner geometry matching that of the GE Discovery CT750 HD scanner was simulated. After the virtual CT scans were delivered and radiation dose distributions were calculated for both the ScoutCT-NET patient model and the CT-based gold-standard patient model, the gamma analysis method⁷⁸ was used to evaluate the quantitative agreement between their dose distribution maps.

To quantify the agreement between CT and ScoutCT-NET dose distributions and to test the empirical generalizability of the ScoutCT-NET, the percentage of voxels with a passing gamma value ($\gamma \geq 1$) was calculated for each of the 212 testing patients. The median and inter-quartile range (IQR) for the percentage were calculated. Three passing criteria were used: 1) $d_m = 5\text{mm}$, $D_m = 5\%$; 2) $d_m = 5\text{mm}$, $D_m = 10\%$; and 3) $d_m = 10\text{mm}$, $D_m = 10\%$, where D_m denotes the dose difference and d_m denotes the distance to agreement (DTA) as defined in Ref. 78.

E. Quantitative evaluation of 3D patient attenuation models generated from ScoutCT-NET (II): Accuracy of radiation dose prescription in fluence-field modulated CT applications

With the 3D patient model generated from ScoutCT-NET, the expected x-ray attenuation along an arbitrary ray direction was calculated by forward-projecting the 3D patient model using the Siddon method and the geometry of a given CT system.⁷⁹ Based on the attenuation information and the “minimal mean FBP noise variance method”,^{27,80,81} fluence-field modulation plans were calculated. The plans were determined by solving the following constrained optimization problem:

$$\min_{n_{j,\theta}} \overline{\sigma^2}(n_{j,\theta}), \text{ s . t . } n_{\text{total}} = \sum_{\theta} \sum_j n_{j,\theta}, \quad (3)$$

where $n_{j,\theta}$ represents the expected x-ray number for detector element j at projection angle θ , $\overline{\sigma^2}$ is the mean noise variance of a reconstructed CT image determined by $n_{j,\theta}$, and n_{total} is the total photon number budget. The well-known Lagrange multiplier method can be used to find a closed-form solution as shown by Harpen:⁸¹

$$n_{j,\theta} = \frac{e^{-\frac{l(j,\theta)}{2}}}{\sum_{j,\theta} e^{-\frac{l(j,\theta)}{2}}} n_{\text{total}}. \quad (4)$$

In Eq. (4), $l(j,\theta)$ denotes the line integral of the patient's linear attenuation coefficient along a ray corresponding to the j th detector element and a projection angle of θ . Once the fluence-field modulation plan was determined by solving Eq. (3), it was used to modulate the x-ray fluence for each ray, which will be virtually delivered to the patient using the geometry of the Discovery CT750 HD CT scanner (GE Healthcare, Waukesha, WI). The voxelized ground truth "patient" used for virtual delivery was established based on the true CT image volume. FBP was applied to the acquired projection dataset to reconstruct CT images resulted from the virtual fluence-field modulated CT scans. The radiation dose distribution map was also provided by the Monte Carlo method (MC-GPU version 1.3).

Using a similar workflow, we also generated images and dose distributions for 1) Interview modulated CT prescribed using the average of the central 100 detector channels from two-view scouts and the cosine interpolation method⁴¹ and 2) Inter-view modulated CT prescribed using online feedback from the 180° conjugate projection measurements^{46,47}. Specifically, the "online 180°" method used a bi-dimensional moving average of the projection data and stored the maximum patient attenuation of each angle over the previous half rotation for tube current modulation. All simulations were performed at matched overall patient dose levels. The results were assessed using noise-only image and noise standard deviation (STD) maps generated from 50 repeated virtual CT scans. Radiation dose distribution maps were also compared. The similarity of the dose distributions between the dose modulation plans calculated using ScoutCT-NET images and the actual CT images was assessed using the gamma analysis method previously described. Two passing criteria were used: 1) $d_m = 1\text{mm}$, $D_m = 1\%$; 2) $d_m = 3\text{mm}$, $D_m = 3\%$.

F. Quantitative evaluation of 3D patient attenuation model generated from ScoutCT-NET (III): Accuracy of radiation dose prescription in organ-specific CT dose delivery applications

Image data of a female patient were used to demonstrate the potential application of the ScoutCT-NET based 3D patient model in organ-specific CT dose delivery. The selected organs-of-interest are breast, spinal canal, and ascending aorta. Each virtual dose delivery focuses on one of the three organs. As a proof-of-concept, the objective of the dose prescription was to obtain the highest image quality within the organ-of-interest while keeping radiation dose delivered to other body regions as low as possible. No matter which organ was chosen, the overall dose delivered to the patient was fixed at 1.6 mGy. For each organ, two dose modulation plans were prescribed: one based on the 3D patient model generated from ScoutCT-NET; the other one based on the reference CT images. During the dose prescription, an ROI was placed on the organ identified based on the 3D patient model (or CT images). Next, the tube current for a given projection view was determined based on the standard that rays intersecting the ROI have a reference entrance photon fluence; all other non-intersecting rays were kept at only 6% of the reference fluence. The quality of the

organ-specific modulation plans was assessed using noise-only images and noise STD maps generated from 50 repeated virtual scans. Dose distribution maps were compared between plans prescribed using the ScoutCT-NET 3D patient model and those using true CT images.

G. Empirical studies of generalizability of ScoutCT-NET

One potential challenge with the existing deep learning studies lies in the fact that there is no theoretical guarantee in their generalizability at this time. Therefore, it is important to extensively test the generalizability of a trained deep learning model despite the caveat that there is no general principle to guide this empirical generalizability test. In this study, empirical generalizability was tested from two aspects: as shown in sub-section IID, Monte Carlo dose distributions have been computed using the output of the trained ScoutCT-NET over the entire test patient cohort (212 patients). To further test how well the network model can be generalized in practice, qualitative evaluations have also been performed to determine whether the trained model can be applied to extremal cases such as pediatric patients and anthropomorphic chest phantoms.

III. RESULTS

A. 3D tomographic patient model accuracy studies: Qualitative and quantitative assessment of ScoutCT-NET based dose distribution maps

Table I summarizes the results of the gamma analysis for the agreement between the radiation dose distributions virtually delivered to the ScoutCT-NET 3D patient model and those to the CT images for all 212 testing subjects. Median values for the percent of voxels with satisfactory gamma values ($\gamma(r_m) \geq 1$) across the 212 testing patients were 97.9%, 95.0% and 92.8% for the ($d_m = 10\text{mm}$, $D_m = 10\%$), ($d_m = 5\text{mm}$, $D_m = 10\%$), and ($d_m = 5\text{mm}$, $D_m = 5\%$) criteria, respectively. These results indicate that one could estimate the radiation dose from a helical chest CT using ScoutCT-NET images with 95.0% of the patient's voxels being within 5 mm and 10% dose difference of the actual CT scan. This is important because it opens up the possibility of having accurate patient-specific and organ-specific x-ray modulation prescriptions.

Figure 3 shows axial and coronal slices through the ScoutCT-NET 3D attenuation model of a patient from the testing cohort. A similar image impression between the ScoutCT-NET patient model and CT images was achieved with all major anatomical structures and organs clearly delineated in the ScoutCT-NET patient model. In addition, the radiation dose distribution from a virtual helical CT scan delivered to the 3D attenuation model closely matched the radiation dose distribution calculated based on the CT image volume. For this subject, the gamma analysis demonstrated that 96.0% of image voxels had a gamma value ≥ 1 based on the $d_m = 5\text{mm}$ distance to agreement and $D_m = 10\%$ criterion.

B. 3D tomographic patient model accuracy studies: Accuracy in radiation dose prescription in fluence-field modulated CT applications

Figure 4 compares x-ray fluence profiles used in the inter-view and intra-view dose modulated CT prescribed and virtually delivered to a patient from the testing cohort. For each projection angle, the x-ray fluence fields prescribed based on the ScoutCT-NET

3D attenuation model closely matched those prescribed based on CT images. The figure also demonstrates that, even with the use of the bowtie beam filter, the two-view scout based “cosine” inter-view modulation and the “online 180°” inter-view scheme generated less conformal fluence profiles: for example, for projection views with an overall high attenuation due to the presence of the spine, the breast tissue also receive more doses.

Figure 5 shows imaging outcomes from virtual deliveries of four different CT dose modulation schemes, which includes two-view scout image-based “cosine” tube current modulation, “online 180°” tube current modulation, reference CT image based intra-view fluence-field modulation, and ScoutCT-NET 3D patient model-based intra-view fluence-field modulation. The results were simulated at matched overall absorbed dose level (1.6 mGy). For each dose delivery scheme, the reconstructed CT image, noise-only image, noise STD map, and absorbed dose map are provided. The noise-only images and STD images show highly non-uniform noise distributions for both the “cosine” and the “online 180°” tube current modulation methods: while lower noise can be observed in the lungs, much higher noise can be found in the periphery region of the body where the bowtie shape mismatched the actual patient attenuation. In contrast, intra-view fluence-field modulation led to much more uniform noise patterns across the entire body cross-section with significantly fewer noise streaks. The spatially uniform image quality provided by fluence-field modulated CT can benefit the detection of lesions both within and outside the lungs as well as detection of other incidental findings. As shown by the third and fourth columns of Fig. 5, both the image quality and dose distributions of the ScoutCT-NET patient model-based fluence-field modulated CT matched those prescribed using the true CT volume: the percent of voxels within the patient with satisfactory gamma values are 98.4% and 100.0% for the 1) $d_m = 1\text{mm}$, $D_m = 1\%$ and 2) $d_m = 3\text{mm}$, $D_m = 3\%$ criteria, respectively. As a reminder, in reality, the true CT volume is generally unavailable before the fluence-field modulated CT is performed. In this case, only ScoutCT-NET enabled fluence-field modulated CT with uniform image quality.

C. 3D tomographic patient model accuracy studies: Accuracy in radiation dose prescription in organ-specific dose prescription CT applications

Figure 6 shows imaging outcomes of organ-specific CT dose deliveries. Plans of these deliveries were prescribed using the ScoutCT-NET 3D patient attenuation model, or using the CT volume as the reference standard. For each virtual dose delivery, one of the following three organs was chosen: breast, the ascending aorta, or the spinal canal. For the ScoutCT-NET-based plan, the ROIs of these organs were drawn based on the 3D patient attenuation model without using any information from the CT image volume. As demonstrated by Row (1) in Fig. 6, organ ROIs determined based on the ScoutCT-NET patient model matched those based on the CT image volume. Even for glandular tissues in the breast, its ROI can be easily drawn based on the ScoutCT-NET patient model. Fig. 6 Row (2) and Row (3) show the differing image quality between the ROIs and the rest of the body cross-section. When the CT dose modulation was prescribed for a given ROI, noise in that ROI is significantly lower. The noise-only images in Row (4) and the STD maps in Row (5) further confirm the much lower noise magnitude in the selected ROI of each plan. The radiation dose maps in Row (6) demonstrate that the dose distribution was substantially modulated: the

organ-of-interest received higher (but not overly high) dose while the rest of the patient body received significantly less radiation. The overall dose was maintained at a constant level of 1.6 mGy. These results demonstrate that ScoutCT-NET 3D patient model can be used to prescribe organ-specific dose modulated CT to enhance the image quality of specific regions in the patient body while meeting the radiation dose constraint. The results also confirm that the ScoutCT-NET patient model generated equivalent imaging outcomes as the gold-standard patient attenuation model.

D. Generalizability test

To demonstrate the generalizability of ScoutCT-NET, the network was applied not only to adult patients, but also to a pediatric patient and an anthropomorphic chest phantom. Fig. 7 shows multiplanar reformatted images and volume renderings of 3D attenuation models for the generalizability-testing image objects: although some anatomical structures in the lung were not perfectly reproduced in the ScoutCT-NET attenuation models due to the low attenuation contrast of these structures in the planar scout images, the overall image impression and organ locations of ScoutCT-NET patient models accurately resemble the corresponding CT images even for patients with abnormal anatomy in addition to a 2-year-old pediatric patient.

IV. DISCUSSION AND CONCLUSIONS

In this paper, the following key results have been obtained: 1) It has been shown that the proposed ScoutCT-NET can be trained in an end-to-end manner to generate 3D patient attenuation models directly from true clinical two-view scout (i.e. topogram) images. Once the ScoutCT-NET model is trained, it can be directly used to reconstruct 3D patient attenuation models without any prior knowledge of the patient anatomy. 2) One trained model using a large training data set with diverse data conditions can be applied to reconstruct 3D patient tomographic models for patients with a similar anatomical coverage. 3) The reconstructed 3D patient attenuation models provided the needed accuracy to prescribe radiation dose in both fluence-field modulated CT and organ-specific dose deliveries. 4) Since the 3D patient model is directly reconstructed from two-view scout images with a long z-coverage, the reconstructed patient tomographic model can be readily used to perform nearly optimal inter-view tube current modulations similar to that achieved in online modulation schemes. In addition, the ScoutCT-NET patient model does not depend on the CT scan modes employed in various applications.

Towards reconstructing CT images from one or two planar x-rays, prior efforts have been made to train deep neural networks using CT images and their digital forward-projections, namely DRRs.^{82–84} The majority of the validations of these networks used DRRs (over which the developers have precise control and knowledge during the forward-projection process) instead of real x-ray images, not to mention real scout images. To our knowledge, the only prior work that performed evaluations using real x-ray radiography images, but not actual scout images, is Ref. 65. However, no reference-standard CT images were provided to validate the “quite plausible”⁶⁵ images generated from real x-ray images and thus their values in medical applications remain unclear due to the potential challenges

between DRRs and scout images. It is worth mentioning that, in reality, there exist multiple differences between DRRs, x-ray radiography images, and scout images: for example, the signal intensity of scout images is scaled differently from radiography images; both the pixel size and material of CT detectors are different from the flat panel detectors used in radiography and the idealized detector used in generating DRRs; x-ray radiography images are usually acquired under the wide-cone geometry with much poorer scatter rejection capability and stronger x-ray heel effect compared to CTs. In contrast, scout images are acquired by operating the MDCT scanner under a scanning-beam acquisition mode with narrowly collimated beams. The scout images were stitched from a series of image segments along the scanning direction. Therefore, as shown in Fig. 2 in Subsection IIB, a proper preparation/curation of the scout images during the training phase of ScoutCT-NET is essential for its initial success. In the context of 4D CBCT acquisitions, it is also worth mentioning that earlier work⁸⁵ has demonstrated the possibility to estimate the motion field from a single view of DRR. If the estimated motion field is accurate enough, one can potentially generate a new image volume that is consistent with the single view DRR. However, there is no study to look into this possibility in a systematic way yet.

Our study also has some limitations. First, we only studied the application of ScoutCT-NET for anatomical regions in the chest. This deliberate choice is partially due to the fact that the chest anatomy is highly heterogeneous and contains important radiosensitive organs with major differences in density, e.g. lung tissue vs soft tissue. As a result, this anatomic region has a great potential for intra-view dose modulation CT to offer greater potential for radiation dose reduction. But our general methodology is not limited to chest scans and in principle, one could easily extend and re-train a ScoutCT-NET network for other anatomical regions. Second, we have only validated our methodology using images from a single CT vendor, meaning that our trained ScoutCT-NET model might not be directly generalizable to other vendors' scout image data without further re-training or network fine-tuning. However, it is important to notice that the general methodology of ScoutCT-NET does not make any assumptions about the equipment used to acquire the scout images and CT images or any vendor proprietary data processing steps other than the CT system geometry. Therefore, in principle, the ScoutCT-NET should be able to be implemented under multiple manufacturers' platforms without significant drawbacks. Finally, an inherent limitation of the presented methodology is the fact that both the AP and LAT scout images, as well as the CT scans used as prior images for network training and data analysis, come from separated scans acquired at different patient breath holds with CT table translation between scans. This means that the scout images used as input of the ScoutCT-NET network might not be perfectly registered to the exact same anatomy or the target images used during network training regularization. This limitation is somewhat alleviated by the use of a data consistency term in the objective function, but at the same time it prevents us from drawing even stronger conclusions regarding the accuracy of the reconstructed images and their potential diagnostic value. Also, at the current research stage, neural network architecture design is purely empirical, and we did not perform any ablation studies or neural architecture search optimization, as these problems will often be NP-hard in order to determine the minimal required amount of data and the neural network architecture needed to achieve the desired performance. However, it is important to notice that we did not

observe any case in our cohort where this limitation had a significant adverse effect on the practical use of ScoutCT-NET for radiation dose optimization purposes.

In conclusion, a deep learning-based method has been developed in this work to generate 3D patient attenuation models directly from real two-view scout images without requiring any prior knowledge of the patients. This method opens a new way for prospectively prescribing dose delivery schemes for fluence-field modulated and organ-specific dose modulated CT towards accomplishing personalized medicine.

ACKNOWLEDGMENT

This work is partially supported by the National Heart, Lung, and Blood Institute of the NIH under grant R01HL153594. The authors would like to thank Dalton Griner, Daniel Bushe, and Kevin Treb for their editorial assistance.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

1. van Randen A, Laméris W, van Es HW, et al. , “A comparison of the accuracy of ultrasound and computed tomography in common diagnoses causing acute abdominal pain.” *European radiology* 21, 1535–45 (2011). [PubMed: 21365197]
2. Schauer DA and Linton OW, “NCRP Report No. 160, Ionizing radiation exposure of the population of the united states, medical exposure—are we doing less with more, and is there a role for health physicists?” *Health Physics* 97, 1–5 (2009). [PubMed: 19509507]
3. Brenner DJ and Hall EJ, “Computed tomography — An increasing source of radiation exposure,” *New England Journal of Medicine* 357, 2277–2284 (2007).
4. McCollough CH, Chen GH, Kalender W, et al. , “Achieving routine submillisievert CT scanning: Report from the summit on management of radiation dose in CT,” *Radiology* 264, 567–580 (2012). [PubMed: 22692035]
5. McCollough CH, Bartley AC, Carter RE, et al. , “Low-dose CT for the detection and classification of metastatic liver lesions: Results of the 2016 Low Dose CT Grand Challenge,” *Med. Phys* 44, e339–e352 (2017). [PubMed: 29027235]
6. Hsieh J, “Adaptive streak artifact reduction in computed tomography resulting from excessive X-ray photon noise,” *Med. Phys* 25, 2139–2147 (1998). [PubMed: 9829238]
7. Manduca A, Yu L, Trzasko JD, et al. , “Projection space denoising with bilateral filtering and CT noise modeling for dose reduction in CT: Projection space denoising with bilateral filtering in CT,” *Med. Phys* 36, 4911–4919 (2009). [PubMed: 19994500]
8. Gomez-Cardona D, Hayes JW, Zhang R, Li K, Cruz-Bastida JP, and Chen G-H, “Low-dose cone-beam CT via raw counts domain low-signal correction schemes: Performance assessment and task-based parameter optimization (Part II. Task-based parameter optimization),” *Med. Phys* 45, 1957–1969 (2018). [PubMed: 29532480]
9. Hayes JW, Gomez-Cardona D, Zhang R, Li K, Cruz-Bastida JP, and Chen G-H, “Low-dose cone-beam CT via raw counts domain low-signal correction schemes: Performance assessment and task-based parameter optimization (Part I: Assessment of spatial resolution and noise performance),” *Med. Phys* 45, 1942–1956 (2018). [PubMed: 29532483]
10. Nuyts J, De Man B, Fessler JA, Zbijewski W, and Beekman FJ, “Modelling the physics in the iterative reconstruction for transmission computed tomography,” *Phys. Med. Biol* 58, R63–R96 (2013). [PubMed: 23739261]

11. Beister M, Kolditz D, and Kalender WA, "Iterative reconstruction methods in X-ray CT," *Physica Medica* 28, 94–108 (2012). [PubMed: 22316498]
12. Dang H, Siewerdsen JH, and Stayman JW, "Prospective regularization design in prior-image-based reconstruction," *Phys. Med. Biol* 60, 9515–9536 (2015). [PubMed: 26606653]
13. Zhang H, Gang GJ, Dang H, and Stayman JW, "Regularization analysis and design for prior-image-based X-ray CT reconstruction," *IEEE Trans. Med. Imaging* 37, 2675–2686 (2018). [PubMed: 29994249]
14. Pickhardt PJ, Lubner MG, Kim DH, et al. , "Abdominal CT with model-based iterative reconstruction (MBIR): initial results of a prospective trial comparing ultralowdose with standard-dose imaging." *AJR. American journal of roentgenology* 199, 1266–74 (2012). [PubMed: 23169718]
15. Pooler BD, Lubner MG, Kim DH, et al. , "Prospective evaluation of reduced dose computed tomography for the detection of low-contrast liver lesions: direct comparison with concurrent standard dose imaging," *European Radiology* 27, 2055–2066 (2017). [PubMed: 27595834]
16. Lubner MG, Pickhardt PJ, Kim DH, Tang J, del Rio AM, and Chen G-H, "Prospective evaluation of prior image constrained compressed sensing (PICCS) algorithm in abdominal CT: a comparison of reduced dose with standard dose imaging," *Abdominal Imaging* 40, 207–221 (2015). [PubMed: 24943136]
17. Mileto A, Guimaraes LS, McCollough CH, Fletcher JG, and Yu L, "State of the art in abdominal CT: the limits of iterative reconstruction algorithms," *Radiology* 293, 491–503 (2019). [PubMed: 31660806]
18. Shefer E, Altman A, Behling R, et al. , "State of the art of CT detectors and sources: a literature review," *Current Radiology Reports* 1, 76–91 (2013).
19. Duan X, Wang J, Leng S, et al. , "Electronic noise in CT detectors: impact on image noise and artifacts," *American Journal of Roentgenology* 201, W626–W632 (2013). [PubMed: 24059402]
20. Symons R, Pourmorteza A, Sandfort V, et al. , "Feasibility of dose-reduced chest CT with photon-counting detectors: initial results in humans," *Radiology* 285, 980–989 (2017). [PubMed: 28753389]
21. Yu Z, Leng S, Kappler S, et al. , "Noise performance of low-dose CT: comparison between an energy integrating detector and a photon counting detector using a whole-body research photon counting CT scanner," *Journal of Medical Imaging* 3, 043503 (2016).
22. Yu Z, Leng S, Jorgensen SM, et al. , "Evaluation of conventional imaging performance in a research whole-body CT system with a photon-counting detector array," *Phys. Med. Biol* 61, 1572–1595 (2016). [PubMed: 26835839]
23. Leng S, Yu Z, Halaweish A, et al. , "Dose-efficient ultrahigh-resolution scan mode using a photon counting detector computed tomography system," *Journal of Medical Imaging* 3, 043504 (2016).
24. Willeminck MJ, Persson M, Pourmorteza A, Pelc NJ, and Fleischmann D, "Photon-counting CT: technical principles and clinical prospects," *Radiology* 289, 293–312 (2018). [PubMed: 30179101]
25. Christner JA, Zavaletta VA, Eusemann CD, Walz-Flannigan AI, and McCollough CH, "Dose reduction in helical CT: dynamically adjustable z-axis X-ray beam collimation," *American Journal of Roentgenology* 194, W49–W55 (2010). [PubMed: 20028890]
26. Kalender WA, Wolf H, Suess C, Gies M, Greess H, and Bautz WA, "Dose reduction in CT by on-line tube current control: principles and validation on phantoms and cadavers," *European Radiology* 9, 323–328 (1999). [PubMed: 10101657]
27. Gies M, Kalender WA, Wolf H, Suess C, and Madsen MT, "Dose reduction in CT by anatomically adapted tube current modulation. I. Simulation studies," *Med. Phys* 26, 2235–2247 (1999). [PubMed: 10587204]
28. Kalender WA, Wolf H, and Suess C, "Dose reduction in CT by anatomically adapted tube current modulation. II. Phantom measurements," *Med. Phys* 26, 2248–2253 (1999). [PubMed: 10587205]
29. Mastora I, Remy-Jardin M, Delannoy V, et al. , "Multi-detector row spiral CT angiography of the thoracic outlet: dose reduction with anatomically adapted online tube current modulation and preset dose savings," *Radiology* 230, 116–124 (2004). [PubMed: 14645883]

30. Graham S, Siewerdsen J, and Jaffray D, "Intensity-modulated fluence patterns for task-specific imaging in cone-beam CT," in *Medical Imaging 2007: Physics of Medical Imaging*, Vol. 6510 (International Society for Optics and Photonics, 2007) p. 651003.
31. Bartolac S, Graham S, Siewerdsen J, and Jaffray D, "Fluence field optimization for noise and dose objectives in CT: Fluence field modulated computed tomography," *Med. Phys* 38, S2–S17 (2011). [PubMed: 21978114]
32. Szczykutowicz TP and Mistretta CA, "Design of a digital beam attenuation system for computed tomography: Part I. System design and simulation framework," *Med. Phys* 40, 021905 (2013).
33. Szczykutowicz TP and Mistretta CA, "Design of a digital beam attenuation system for computed tomography. Part II. Performance study and initial results: Digital beam attenuation system initial results," *Med. Phys* 40, 021906 (2013).
34. Hsieh SS and Pelc NJ, "The piecewise-linear dynamic attenuator reduces the impact of count rate loss with photon-counting detectors," *Phys. Med. Biol* 59, 2829 (2014). [PubMed: 24819415]
35. Hsieh SS, Fleischmann D, and Pelc NJ, "Dose reduction using a dynamic, piecewise-linear attenuator," *Med. Phys* 41, 021910 (2014).
36. Hsieh SS, Peng MV, May CA, Shunhavanich P, Fleischmann D, and Pelc NJ, "A prototype piecewise-linear dynamic attenuator," *Phys. Med. Biol* 61, 4974–4988 (2016). [PubMed: 27284705]
37. Stayman JW, Mathews A, Zbijewski W, et al. , "Fluence-field modulated X-ray CT using multiple aperture devices," *Proc.SPIE* 9783, 97830X (2016).
38. Gang GJ, Mao A, Wang W, Siewerdsen JH, Mathews A, Kawamoto S, Levinson R, and Stayman JW, "Dynamic uence _eld modulation in computed tomography using multiple aperture devices," *Phys. Med. Biol* 64, 105024 (2019).
39. Wang W, Gang GJ, Siewerdsen JH, Levinson R, Kawamoto S, and Stayman JW, "Volume-of-interest imaging with dynamic uence modulation using multiple aperture devices," *Journal of Medical Imaging* 6, 1 (2019).
40. Huck SM, Fung GSK, Parodi K, and Stierstorfer K, "Technical Note: Sheet-based dynamic beam attenuator - A novel concept for dynamic uence _eld modulation in X-ray CT," *Med. Phys* 46, 5528{5537 (2019).
41. Kalra MK, Maher MM, Toth TL, et al. , "Techniques and applications of automatic tube current modulation for CT," *Radiology* 233, 649–657 (2004). [PubMed: 15498896]
42. Yin Z, Yao Y, Montillo A, et al. , "Acquisition, preprocessing, and reconstruction of ultralow dose volumetric CT scout for organ-based CT scan planning," *Med. Phys* 42, 2730–2739 (2015). [PubMed: 25979071]
43. Gomes J, Gang GJ, Mathews A, and Stayman JW, "An investigation of low-dose 3D scout scans for computed tomography," *Proc.SPIE* 10132, 10132 – 10132 – 6 (2017).
44. Fessler J, "Hybrid Poisson/polynomial objective functions for tomographic image reconstruction from transmission scans," *IEEE Transactions on Image Processing* 4, 1439–1450 (1995). [PubMed: 18291975]
45. Zhang R, Cruz-Bastida JP, Gomez-Cardona D, Hayes JW, Li K, and Chen G-H, "Quantitative accuracy of CT numbers: Theoretical analyses and experimental studies," *Med. Phys* 45, 4519–4528 (2018). [PubMed: 30102414]
46. Popescu S, Hentschel D, Strauss K-E, and Wolf H, "Adaptive dose modulation during CT scanning," (1999), US Patent 5,867,555.
47. McMillan K, Bostani M, Cagnon CH, et al. , "Estimating patient dose from CT exams that use automatic exposure control: Development and validation of methods to accurately estimate tube current values," *Med. Phys* 44, 4262–4275 (2017). [PubMed: 28477342]
48. Kang E, Min J, and Ye JC, "A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction," *Med. Phys* 44, e360–e375 (2017). [PubMed: 29027238]
49. Wolterink JM, Leiner T, Viergever MA, and Isgum I, "Generative adversarial networks for noise reduction in low-dose CT," *IEEE Trans. Med. Imaging* 36, 2536–2545 (2017). [PubMed: 28574346]
50. Chen H, Zhang Y, Kalra MK, et al. , "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE Trans. Med. Imaging* 36, 2524–2535 (2017). [PubMed: 28622671]

51. “Artifact removal using improved GoogLeNet for sparse-view CT reconstruction.”
52. Yang Q, Yan P, Zhang Y, et al. , “Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss,” *IEEE Trans. Med. Imaging* 37, 1348–1357 (2018). [PubMed: 29870364]
53. Shan H, Zhang Y, Yang Q, et al. , “3-D convolutional encoder-decoder network for low-dose CT via transfer learning from a 2-D trained network,” *IEEE Trans. Med. Imaging* 37, 1522–1534 (2018). [PubMed: 29870379]
54. Chun IY, Huang Z, Lim H, and Fessler J, “Momentum-Net: Fast and convergent iterative neural network for inverse problems,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1–1 (2020).
55. Ye S, Ravishankar S, Long Y, and Fessler JA, “SPULTRA: Low-dose CT image reconstruction with joint statistical and learned image models,” *IEEE Trans. Med. Imaging* 39, 729–741 (2020). [PubMed: 31425021]
56. Chen H, Zhang Y, Chen Y, et al. , “LEARN: Learned experts’ assessment-based reconstruction network for sparse-data CT,” *IEEE Trans. Med. Imaging* 37, 1333–1347 (2018). [PubMed: 29870363]
57. Hayes JW, Montoya J, Budde A, et al. , “High pitch helical CT reconstruction,” *IEEE Trans. Med. Imaging* 40, 3077–3088 (2021). [PubMed: 34029189]
58. Zhang C, Li Y, and Chen G-H, “Accurate and robust sparse-view angle CT image reconstruction using deep learning and prior image constrained compressed sensing (DLPICCS),” *Med. Phys* 48, 5765–5781 (2021). [PubMed: 34458996]
59. Li Y, Li K, Zhang C, Montoya J, and Chen G-H, “Learning to reconstruct computed tomography images directly from sinogram data under a variety of data acquisition conditions,” *IEEE Trans. Med. Imaging* 38, 2469–2481 (2019). [PubMed: 30990179]
60. Montoya J, Li K, and Chen G-H, “Volumetric CT scouts from conventional two-view projection scout images using deep learning,” in 2018 AAPM Annual Meeting, Nashville, TN, USA (2018).
61. Montoya J, Zhang C, Garrett J, Li K, and Chen GH, “Three-dimensional CT scout from conventional two-view radiograph localizers using deep learning,” in Radiological Society of North America 2018 Scientific Assembly and Annual Meeting, Chicago, IL, USA (2018).
62. Montoya J, Zhang C, Li K, and Chen G-H, “Volumetric scout CT images reconstructed from conventional two-view radiograph localizers using deep learning,” in 2019 SPIE Medical Imaging Conference, San Diego, CA, USA, Vol. 10948 (2019) p. 1094825.
63. Shen L, Zhao W, and Xing L, “Harnessing the power of deep learning for volumetric CT imaging with single or limited number of projections,” in 2019 SPIE Medical Imaging Conference, San Diego, CA, USA, Vol. 10948 (2019) p. 1094826.
64. Shen L, Zhao W, and Xing L, “Patient-specific reconstruction of volumetric computed tomography images from a single projection view via deep learning,” *Nature Biomedical Engineering* 3, 880–888 (2019).
65. Ying X, Guo H, Ma K, Wu J, Weng Z, and Zheng Y, “X2CT-GAN: reconstructing CT from biplanar X-rays with generative adversarial networks,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019).
66. Kalender WA, *Computed Tomography: Fundamentals, System Technology, Image Quality, Applications* (3rd edition) (Publicis, 2011).
67. Henzler P, Rasche V, Ropinski T, and Ritschel T, “Single-image tomography: 3D volumes from 2D cranial X-Rays,” *Computer Graphics Forum* 37, 377–388 (2018).
68. Liu C, Huang Y, Maier J, Klein L, Kachelrieß M, and Maier A, “Robustness investigation on deep learning CT reconstruction for real-time dose optimization,” *arXiv preprint arXiv:2012.03579* (2020).
69. Zhu B, Liu JZ, Cauley SF, Rosen BR, and Rosen MS, “Image reconstruction by domain-transform manifold learning,” *Nature* 555, 487–492 (2018). [PubMed: 29565357]
70. Ronneberger O, Fischer P, and Brox T, “U-net: Convolutional networks for biomedical image segmentation,” in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, LNCS, Vol. 9351 (Springer, 2015) pp. 234–241.

71. Goodfellow IJ, Pouget-Abadie J, Mirza M, et al. , “Generative adversarial nets,” in Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2, NIPS’14 (MIT Press, Cambridge, MA, USA, 2014) p. 2672–2680.
72. Isola P, Zhu J-Y, Zhou T, and Efros AA, “Image-to-image translation with conditional adversarial networks,” in Proceedings of the IEEE conference on computer vision and pattern recognition (2017) pp. 1125–1134.
73. He K, Zhang X, Ren S, and Sun J, “Deep residual learning for image recognition,” in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016).
74. He K, Zhang X, Ren S, and Sun J, “Identity mappings in deep residual networks,” in European conference on computer vision (Springer International Publishing, 2016) pp. 630–645.
75. Abadi M, Agarwal A, Barham P, et al., “TensorFlow: Large-scale machine learning on heterogeneous systems,” (2015), software available from [tensorflow.org](https://www.tensorflow.org).
76. He K, Zhang X, Ren S, and Sun J, “Delving deep into rectifiers: Surpassing humanlevel performance on ImageNet classification,” in 2015 IEEE International Conference on Computer Vision, ICCV 2015, Santiago, Chile, December 7–13, 2015 (2015) pp. 1026–1034.
77. Badal A. and Badano A, “Accelerating monte carlo simulations of photon transport in a voxelized geometry using a massively parallel graphics processing unit,” *Med. Phys* 36, 4878–4880 (2009). [PubMed: 19994495]
78. Low DA, Harms WB, Mutic S, and Purdy JA, “A technique for the quantitative evaluation of dose distributions,” *Med. Phys* 25, 656–661 (1998). [PubMed: 9608475]
79. Siddon RL, “Fast calculation of the exact radiological path for a three-dimensional CT array,” *Med. Phys* 12, 252–255 (1985). [PubMed: 4000088]
80. Bartolac S. and Jaffray D, “Compensator models for fluence field modulated computed tomography: Compensator models for fluence field modulated CT,” *Med. Phys* 40, 121909 (2013).
81. Harpen MD, “A simple theorem relating noise and patient dose in computed tomography,” *Med. Phys* 26, 2231–2234 (1999). [PubMed: 10587203]
82. Zhao J, Chen Z, Zhang L, and Jin X, “Few-view CT reconstruction method based on deep learning,” in 2016 IEEE Nuclear Science Symposium, Medical Imaging Conference and Room-Temperature Semiconductor Detector Workshop (NSS/MIC/RTSD) (2016) pp. 1–4.
83. Kim H, Anirudh R, Mohan KA, and Champley K, “Extreme few-view CT reconstruction using deep inference,” arXiv:1910.05375 (2019).
84. Ye DH, Buzzard GT, Ruby M, and Bouman CA, “Deep back projection for sparseview CT reconstruction,” arXiv:1807.02370 (2018).
85. Xu Y, Yan H, Ouyang L, et al. , “A method for volumetric imaging in radiotherapy using single X-ray projection,” *Med. Phys* 42, 2498–2509 (2015). [PubMed: 25979043]

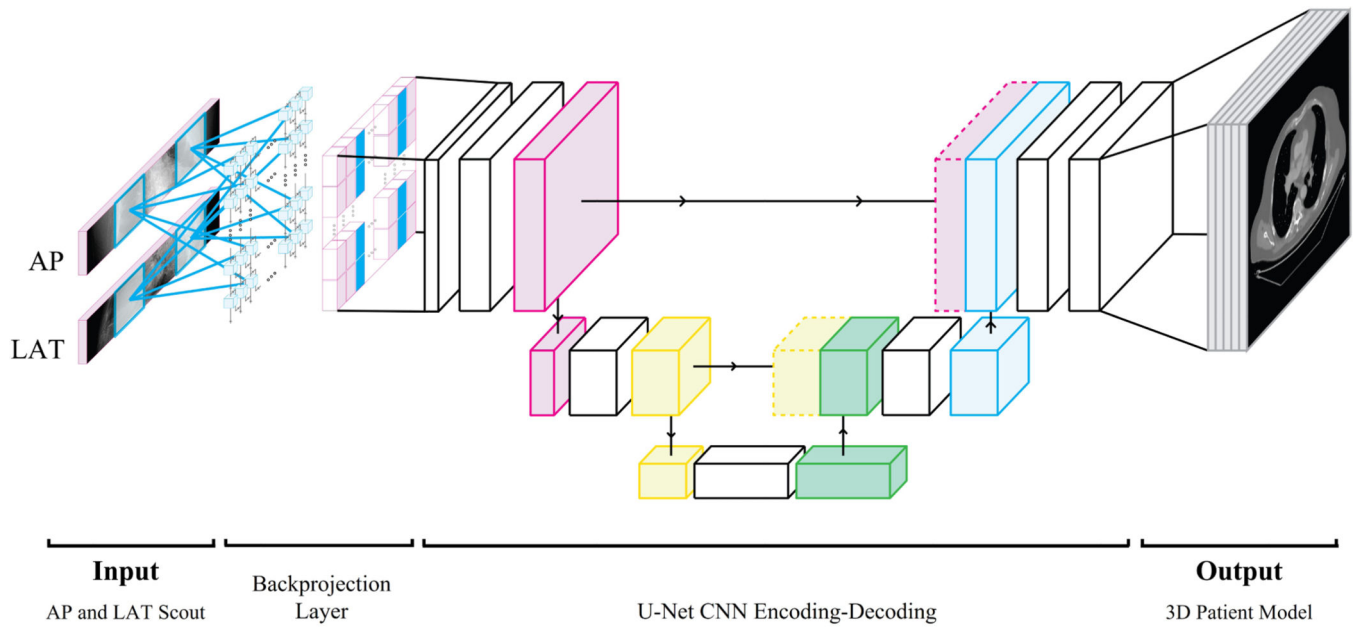


Figure 1: ScoutCT-NET deep neural network architecture. The network takes a 888×5 narrow section from each of the two scout images as input to generate a sub-volume of tomographic images with $288 \times 288 \times 5$ voxels. This process is repeated for each section of the scout images. In order to map the 2D projections into the 3D image space, the first layer of the network performs a direct backprojection operation which is then followed by 24 convolutional layers arranged into a simplified U-Net architecture with four vertical levels (only 3 vertical levels are shown for diagram simplicity).

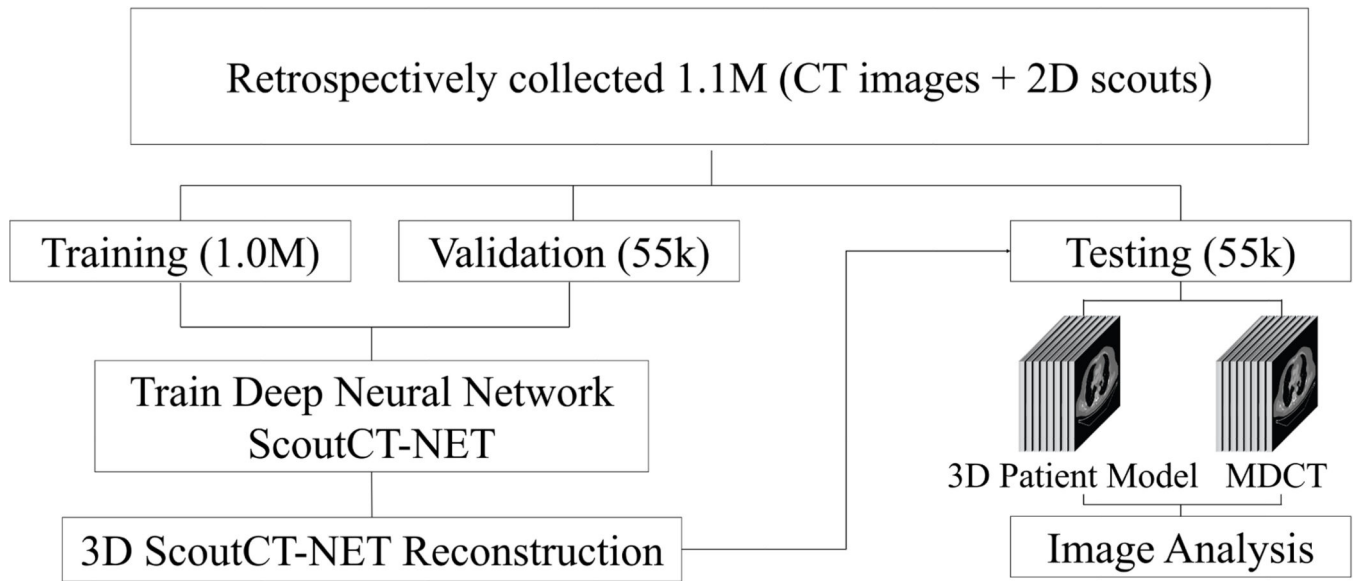


Figure 2: Overall study schema. More than 1.1 million CT images and their corresponding scout image data from 4,214 clinically indicated CT studies were retrospectively collected. Data were split into three cohorts consisting of 3,790 (90%), 212 (5%) and 212 (5%) CT studies for training, validation and testing of the deep learning model, respectively. The total number of unique data samples used was 1,001,496 for training, 55,136 for validation, and 55,136 for testing. After the ScoutCT-NET model was trained, 3D image localizers were calculated for the 212 testing patients and image analysis was performed.

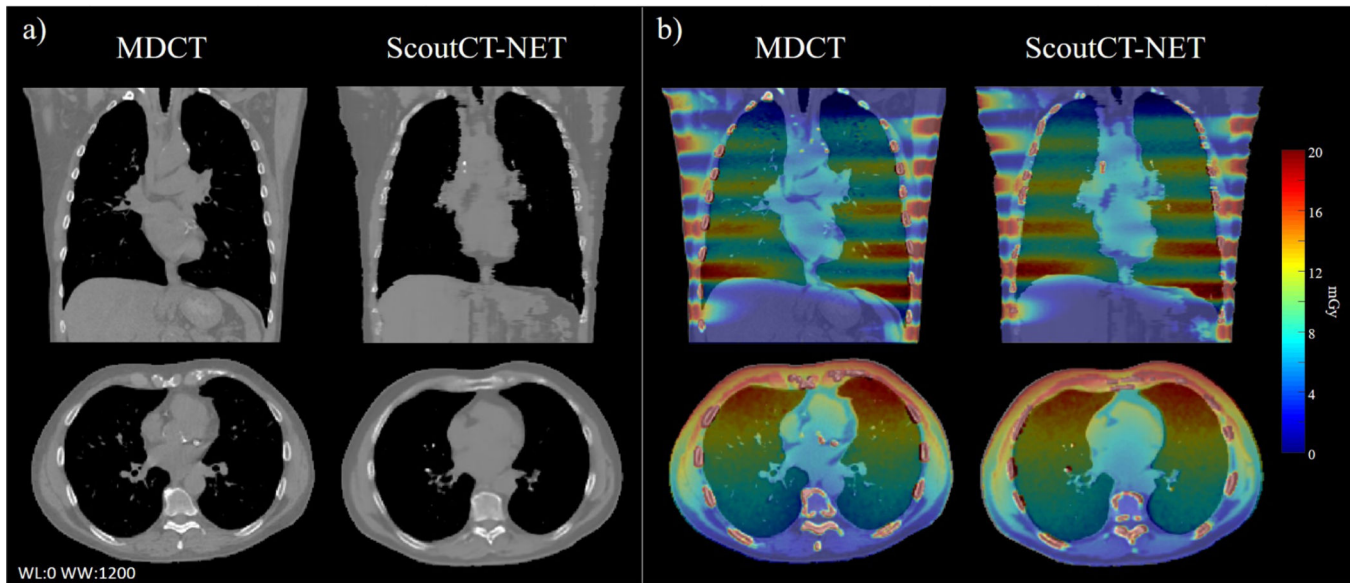


Figure 3:

(a) Coronal and axial slices through the ScoutCT-NET 3D patient attenuation model and the corresponding CT images. (b) Comparison of the spatial distribution of absorbed doses from a virtual helical CT scan delivered to the ScoutCET-NET attenuation model with those to the CT-based attenuation model. 96.0% of the voxels had a gamma value ≥ 1 based on the $d_m = 5\text{mm}$ distance to agreement and $D_m = 10\%$ criterion.

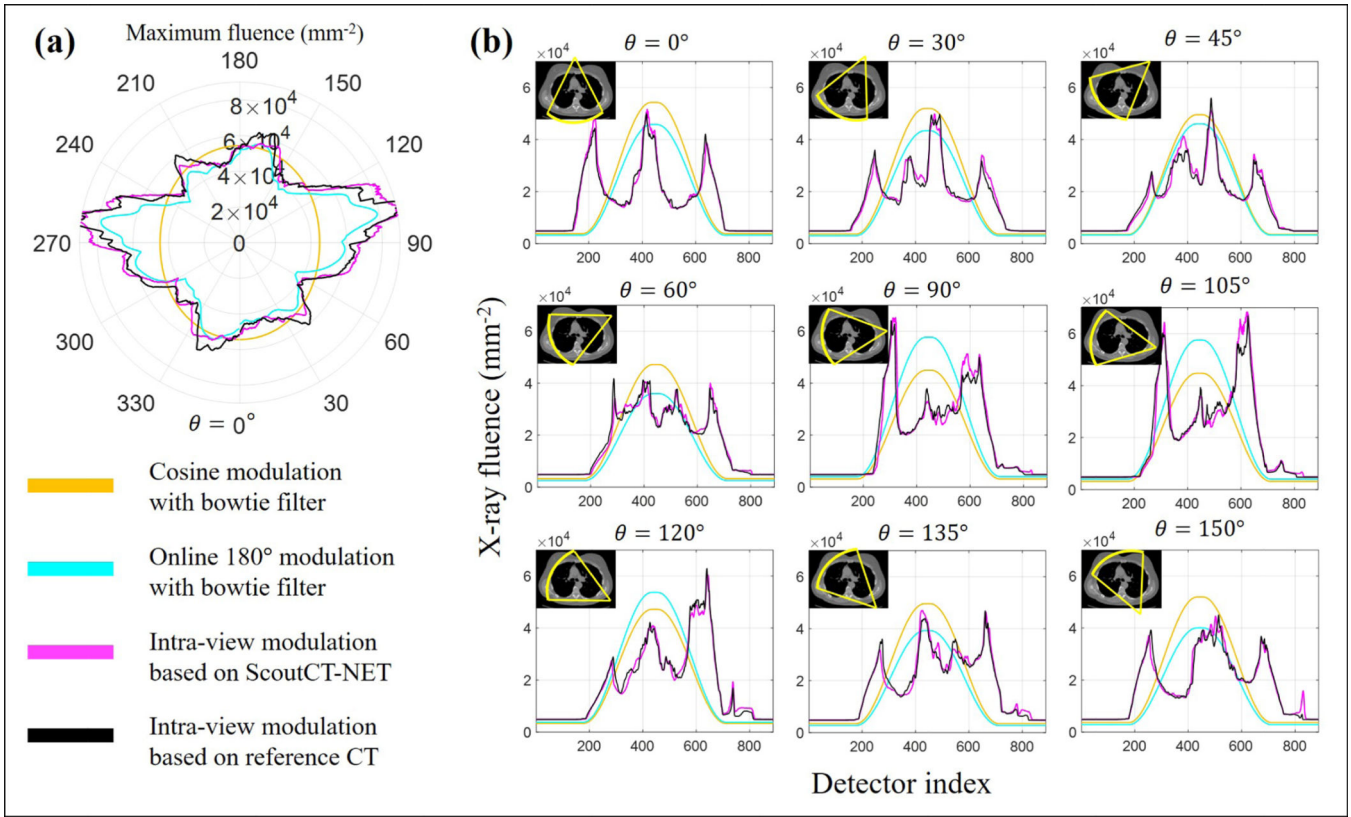


Figure 4: Comparison of x-ray fluence profiles for different CT dose delivery techniques applied to a test subject. These techniques include the “cosine” tube current modulation (TCM) combined with a bowtie filter, “online 180° ” TCM with a bowtie filter, intra-view fluence-field modulation prescribed based on the CT image volume, and intra-view fluence-field modulation prescribed based on the ScoutCT-NET 3D patient model. The sub-figure (a) on the upper left plots the maximal fluence of each view as a function of projection angle. Other sub-figures in (b) show x-ray fluence profiles for 9 representative projection angles. For the “cosine” and the “online 180° ” methods, the modulation within the fan beam for each projection view was caused by the bowtie filter while both methods used angular TCMs. Each detector element has a physical size of $1 \text{ mm} \times 1 \text{ mm}$.

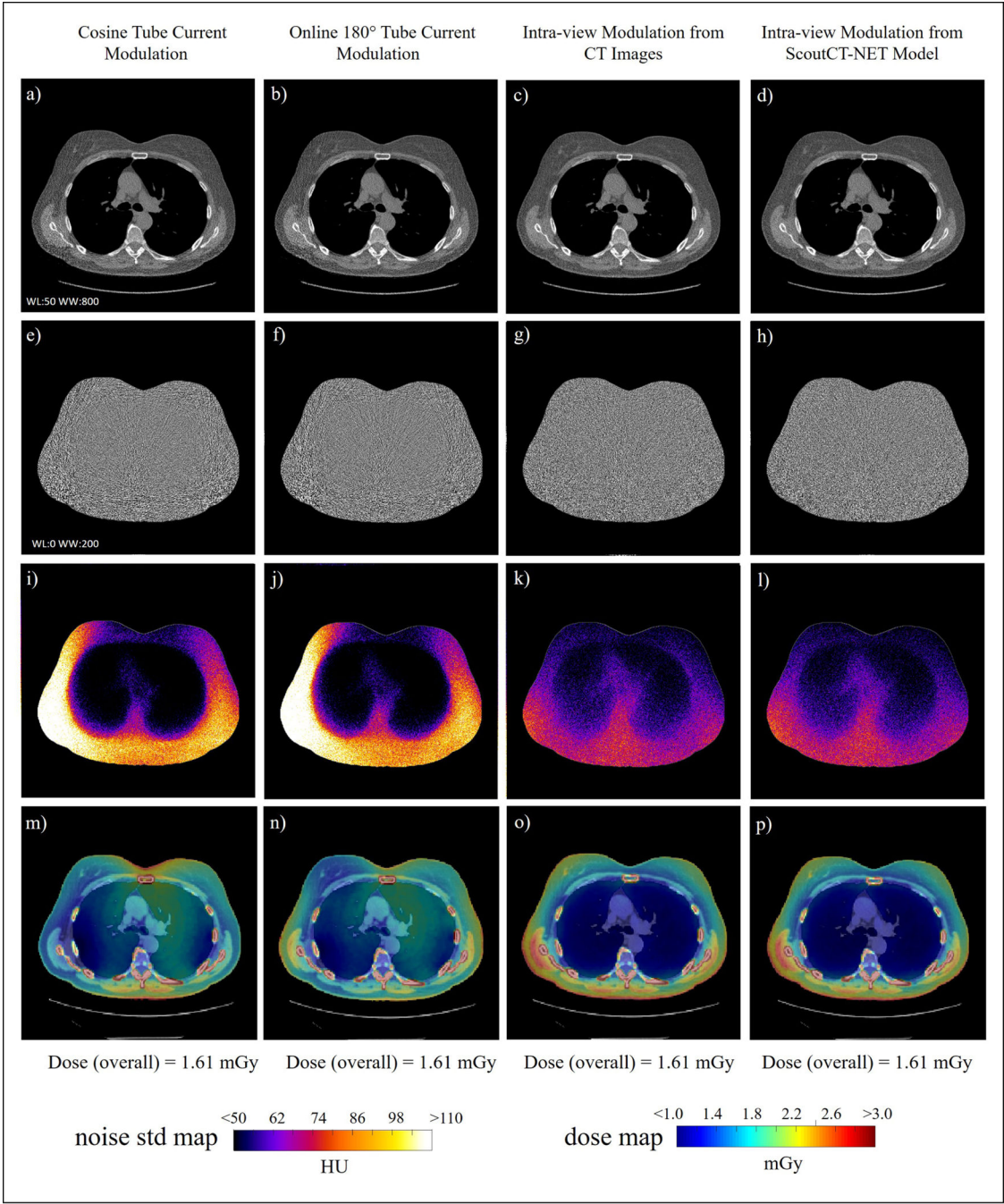


Figure 5: Comparison of four different dose modulation techniques that were virtually applied to a female patient from the test cohort. The four techniques are “cosine” tube current modulation prescribed based on two-view 2D scout images (column 1), “online 180°” tube current modulation prescribed based on conjugate projections (column 2), intra-view (fluence-field) modulation prescribed based on the reference CT images (column 3), and intra-view (fluence-field) modulation prescribed based on the ScoutCT-NET 3D patient model (column 4). The first row shows CT images from different dose modulation schemes;

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

the second row shows noise-only images; the third row displays noise standard deviation (STD) maps calculated from 50 repeated virtual experiments; the fourth row shows radiation dose distributions.

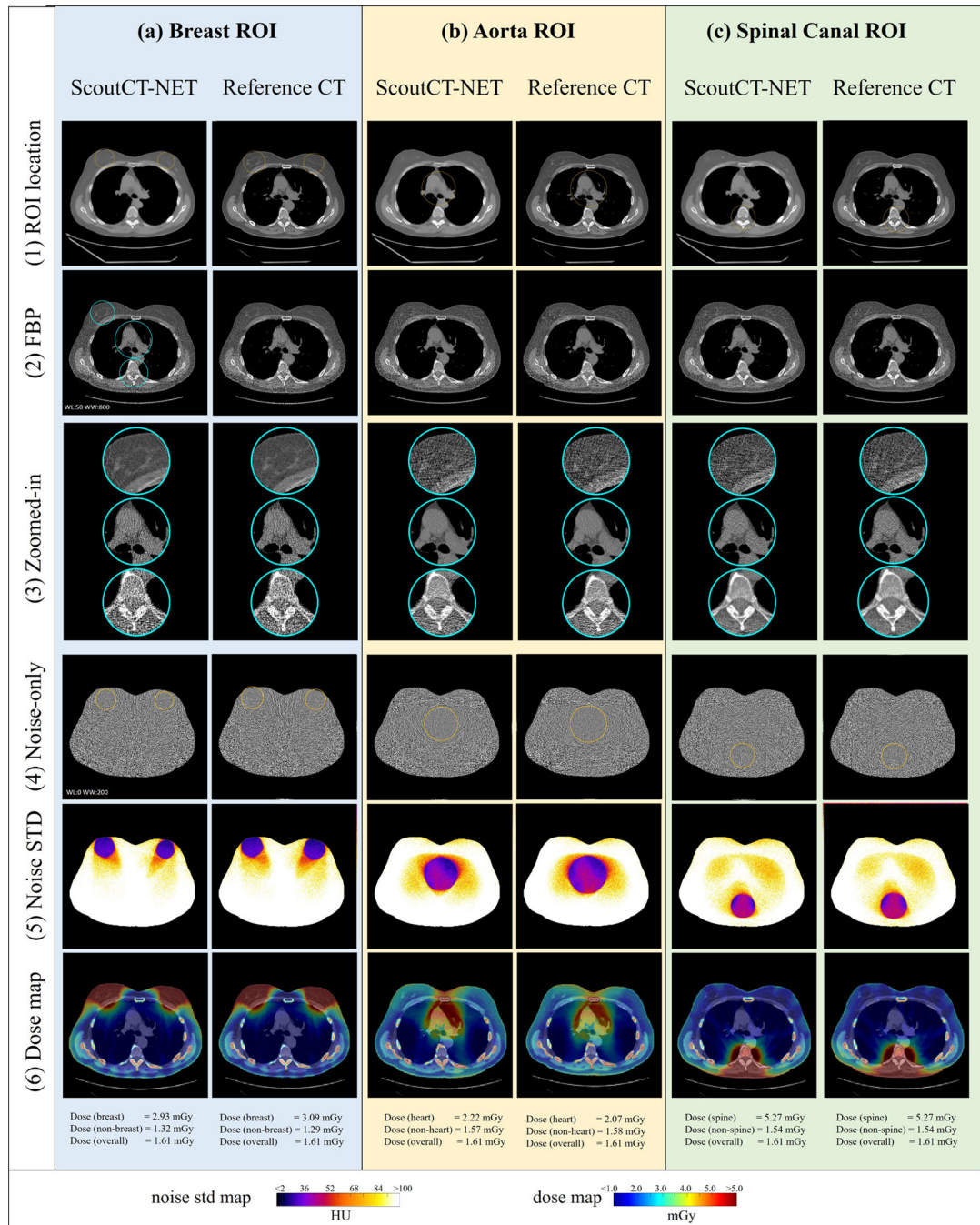


Figure 6: Results from organ-specific CT dose deliveries. The region-of-interest (ROI) is in the breast (a), the ascending aorta (b), or the spinal canal (c). Row (1) illustrates the locations of the ROI selected based on the ScoutCT-NET 3D patient model or the reference CT image volume. Row (2) shows the CT image virtually acquired using different organ-specific dose modulation plans and reconstructed using FBP. Row (3) shows zoomed-in images at three locations for each dose modulation plan. Row (4) shows noise-only images corresponding to the results in Row (2). Row (5) shows noise standard deviation (STD) maps calculated

from 50 repeated virtual deliveries of each dose modulation plan. Row (6) shows radiation dose distribution of each dose modulation plan. The dose levels are reported within the organ-of-interest and excluding the organ-of-interest. All modulation plans are conducted with the overall dose delivered to the patient being matched.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

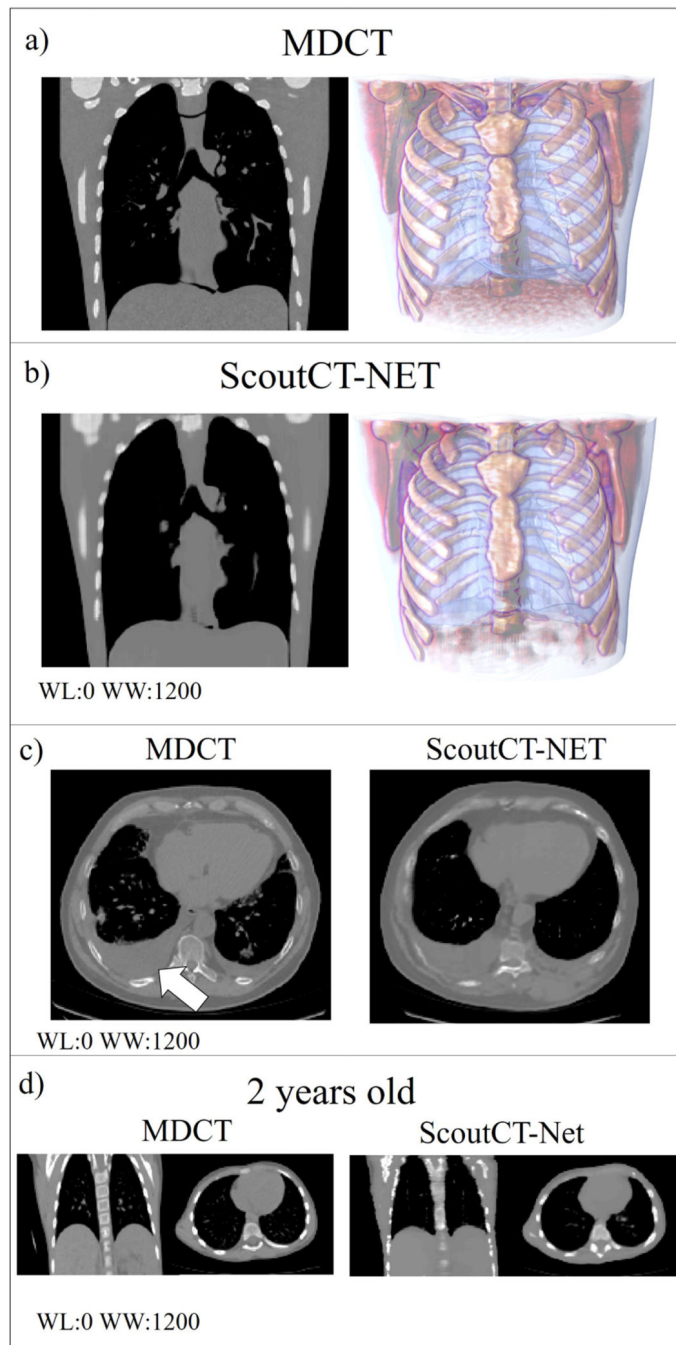


Figure 7: Coronal and volume rendering of (a) MDCT and (b) ScoutCT-NET images of an anthropomorphic phantom. The ScoutCT-NET images reconstructed in (b) served as the basis 3D patient attenuation model to calculate the optimal intra-view dose modulation plan. (c) Shows axial MDCT and ScoutCT-NET images of a patient with pleural effusion as pointed by the arrow and (d) shows axial and coronal images of a 2-years old pediatric patient. These images show the generalization of the ScoutCT-NET model to patients with

different medical conditions as well as pediatric patients, despite being trained using images from adult patients.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table I.

Percent of voxels with passing gamma criteria ($\gamma(r_m) \geq 1$) across the entire testing cohort with 212 patients.

DTA: distance to agreement (DTA).

DTA/Dose Difference	Median	IQR	Min	Max
5 mm/5%	92.8 %	89.7% – 94.2 %	78.1 %	96.8 %
5 mm / 10%	95.0 %	93.2% – 96.0 %	83.1 %	97.6 %
10 mm / 10%	97.9 %	97.4% – 98.3 %	93.3 %	99.1 %

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript