Constructing a tissue-specific texture prior by machine learning from previous full-dose scan for Bayesian reconstruction of current ultralow-dose CT images

Yongfeng Gao
Jiaxing Tan
Yongyi Shi
Siming Lu
Amit Gupta
Haifang Li
Zhengrong Liang

SPIE.

Constructing a tissue-specific texture prior by machine learning from previous full-dose scan for Bayesian reconstruction of current ultralow-dose CT images

Yongfeng Gao,[a] Jiaxing Tan,[a] Yongyi Shi,[a] Siming Lu,[a,b] Amit Gupta,[a] Haifang Li,[a] and Zhengrong Liang[a,b,*]

[a]State University of New York, Department of Radiology, Stony Brook, New York, United States
[b]State University of New York, Department of Biomedical Engineering, Stony Brook, New York, United States

Abstract

Purpose: Bayesian theory provides a sound framework for ultralow-dose computed tomography (ULdCT) image reconstruction with two terms for modeling the data statistical property and incorporating *a priori* knowledge for the image that is to be reconstructed. We investigate the feasibility of using a machine learning (ML) strategy, particularly the convolutional neural network (CNN), to construct a tissue-specific texture prior from previous full-dose computed tomography.

Approach: Our study constructs four tissue-specific texture priors, corresponding with lung, bone, fat, and muscle, and integrates the prior with the prelog shift Poisson (SP) data property for Bayesian reconstruction of ULdCT images. The Bayesian reconstruction was implemented by an algorithm called SP-CNN-T and compared with our previous Markov random field (MRF)-based tissue-specific texture prior algorithm called SP-MRF-T.

Results: In addition to conventional quantitative measures, mean squared error and peak signal-to-noise ratio, structure similarity index, feature similarity, and texture Haralick features were used to measure the performance difference between SP-CNN-T and SP-MRF-T algorithms in terms of the structure and tissue texture preservation, demonstrating the feasibility and the potential of the investigated ML approach.

Conclusions: Both training performance and image reconstruction results showed the feasibility of constructing CNN texture prior model and the potential of improving the structure preservation of the nodule comparing to our previous regional tissue-specific MRF texture prior model.

1 Introduction

Concerns about the radiation risk from the widely used computed tomography (CT) have motivated tremendous efforts toward producing clinically useful images at the lowest dose possible, i.e., ultralow-dose CT (ULdCT). However, the image quality will degrade seriously in the ULdCT case. To address ULdCT image reconstruction challenges, Bayesian theory establishes a sound framework with two terms for modeling the data statistics property (fidelity term) and incorporating *a priori* knowledge for the to-be-reconstructed image (prior term).[1] This study aimed to investigate the feasibility of using a machine learning (ML) strategy, particularly a convolutional neural network (CNN), to construct a tissue-specific texture prior from previous

full-dose CT (FdCT) and to integrate the prior with the prelog shift Poisson (SP) data property for Bayesian reconstruction of ULdCT images.

Tissue textures are recognized as useful biomarkers for clinical tasks and play an important role in computer-aided detection and diagnosis.[2,3] Previously, we proposed a Markov random field (MRF)-based tissue-specific texture (MRF-T) prior[4–6] to preserve the tissue textures in the ULdCT images, which demonstrated clinical benefit in the task-based evaluation.[7,8] However, the MRF-T model assumed that the center pixel within the patch has a linear relationship with its neighbor pixels. Therefore, this study aims to investigate the feasibility of "learning" the relationship without any assumption (of linear, nonlinear, or other forms) to benefit from the development of ML, which has been successfully applied in various areas in medical imaging.[9–15] Particularly, we propose to construct a CNN-based tissue-specific texture (CNN-T) prior taking advantage of its strength in pattern recognition, which may have the potential to preserve more detail, compared to our previous MRF-T model.

This study further integrates the proposed CNN-T model with the prelog SP data property for Bayesian reconstruction denoted as SP-MRF-T to address the low photon counts and high noise level in ULdCT. In a CT scanner, noise distribution is complicated by a cascade of random processes, such as photon generation, transmission, detection, electronic readout noise, and so on.[16,17] The prelog model[18–20] directly uses the transmission data from the detector, which can better utilize the statistical property of the x-ray quanta and the electronic background noise in its cost function. It also overcomes many difficulties associated with the log transformation of the postlog domain algorithm[21–23] in the ultralow-dose situation, such as undefined negative values,[24,25] weight estimation from the noisy data,[23,26,27] and so on. Therefore, it has shown theoretical advantages in ULdCT image reconstruction. In the prelog domain algorithms, the SP model is a good approximation to describe the signal property[20,28] and thus it is adapted for usage in this study.

In summary, the main contribution of this work is twofold. (1) We investigated the feasibility of constructing the CNN-T prior with no assumptions, comparing it to our previous MRF-T model. Compared to some existing CNN priors, which were proposed as a denoiser where the model parameters were trained with a pair of noise-free and noisy images,[5] the presented CNN prior in this study not only denoises but also preserves tissue image textures by training the model parameters to learn the tissue textures, thereby improving the model's clinical impact. (2) Comparing the presented SP-CNN-T method with our previous SP-MRF-T algorithm, we explored the potential benefit of CNN-T in texture preservation in ULdCT image reconstruction. Preliminary results were reported at the 15th international meeting on fully three-dimensional image reconstruction.[29]

The rest of this paper is organized as follows. Section 2 describes the proposed SP-CNN-T algorithm and the network design of the CNN priors. Section 3 describes the experiment design and presents the results. Section 4 describes discussion and conclusions.

2 Method

In this paper, we propose a machine-learning approach to construct a tissue-specific texture prior and integrate the prior with the prelog data statistical model under the Bayesian image reconstruction framework for the ULdCT imaging. This section first introduces the Bayesian image reconstruction framework. Then, we describe the "learnt" CNN-T prior within the Bayesian framework. The overall architecture of SP-CNN-T algorithm is presented in the following.

2.1 *Bayesian Image Reconstruction Framework*

Given a set of acquired transmission data, denoted by a vector $Y \in R^{I \times 1}$, where I is the number of data elements, we are interested in a solution, denoted by a vector $\mu \in R^{J \times 1}$ with J number of image voxels, which maximizes the posterior probability $p(\mu|Y)$. By the Bayesian theorem, we have

$$p(\mu|Y) = \frac{p(Y|\mu)p(\mu)}{p(Y)}, \tag{1}$$

where $p(Y)$ becomes a constant when maximizing the posterior probability; therefore, $p(Y)$ will be ignored.

For the acquired low-dose transmission data, following the description in the study of Xing et al.,[28] we introduce an artificial quantity $Y_i : Y_i = Y_i^M - m_i + \sigma_i^2$, which yields

$$[Y_i]_+ \approx \text{Poisson}\{b_i e^{-l_i} + \sigma_i^2\}, \tag{2}$$

where Y_i^M is the measured photon counts at the detector, and m_i and σ_i are the mean and standard deviation of the electronic noise, respectively. Notation b_i is the incident photon counts, and l_i is the i'th line integral along the x-ray path. Let the log data likelihood or fidelity term be $\ln[p(Y|\mu)]$ and the log prior term, $\ln[p(\mu)]$, be described by an MRF model, then the solution for Eq. (1) can be obtained by minimizing the objective function Φ, in which

$$\hat{\mu} = \arg \min_{\mu} - \sum_i \{Y_i \ln(b_i e^{-[A\mu]_i} + \sigma_i^2) - (b_i e^{-[A\mu]_i} + \sigma_i^2)\} + \lambda R(\mu), \tag{3}$$

where $\mathbf{A}$ is the projection matrix and $R(\mu)$ is the unknown prior, which will be learnt by CNN. Notation λ is the hyperparameter to balance the strength between the fidelity term and prior term. The introduced artificial quantity enables us to model the fidelity term with the SP distribution in the prelog domain for more accurate modeling. Here $R(\mu)$ enables us to incorporate our prior knowledge as constraint to improve the image reconstruction performance, which could be any assumption, such as piecewise smooth and low-rank property. There are many types of prior or regularization with the purpose of image denoising, deblurring, edge preserving, piecewise smoothing, etc.[30–32] In this study, we are more interested in the clinically important texture. In next section, we introduce how to construct the tissue-specific texture prior using the ML approach.

2.2 CNN-Based Texture Prior Learning

Tissue texture reflects the voxel gray-level distribution across the field of view or the contrast between one voxel and its neighboring voxels. Given a pixel and its surrounding neighbors, we define the tissue-specific texture prior to be the relationship between them, where the center pixel value could be inferred by its neighbors based on the pattern specified by its texture. Therefore, we proposed the tissue-specific texture prior in the following analytical form:[4–6]

$$w_r = \arg \min|_{w_r} \sum_r \sum_{m \in \text{Region}(r)} (\mu_m - w_r^T \boldsymbol{\mu}_{\Omega_m})^2, \tag{4}$$

where w_r represents the texture prior weights of tissue r, m represents voxels index, Ω_m denotes a small fixed neighborhood region of the voxel m, and $\boldsymbol{\mu}_{\Omega_m}$ are voxels within Ω_m. This tissue-specific texture prior has demonstrated certain advantages in clinical tasks.[7,8] According to Eq. (4), the texture prior assumes that the center voxel has a linear dependence on its neighboring voxels. Therefore, we aim to construct such a prior without the linear assumption benefiting from the ML technology. We denote the trained CNN model as $G(.)$ and the predication model can be expressed as

$$\mu_m = G(\boldsymbol{\mu}_{\Omega_m}). \tag{5}$$

CNNs $G(.)$ could learn a linear/nonlinear or other form model for the relationship between a pixel and its neighbors. Full-dose data were used to train this CNN network. After training, the CNN network can represent the relationship that all neighboring pixels follow.

Following the idea of tissue texture, constructing the CNN texture prior is straightforward. We can input the neighboring voxel values to the machine model and ask the machine to predict the center voxel value. By training so, the machine can learn the relationship or the texture. The idea and network are shown in Fig. 1. We design a three-layer CNN model to learn such relationship from data. As our texture prior is tissue-specific, we build four CNN models for lung, bone, fat, and muscle, respectively. This model is trained offline to learn the texture-constrained
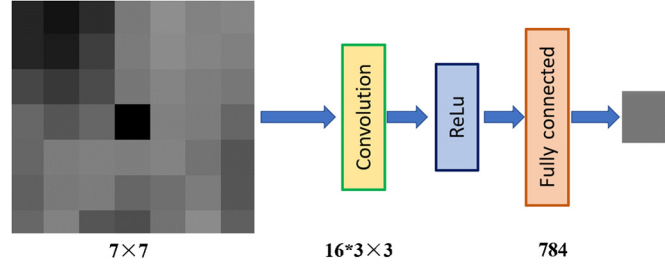
Fig. 1 The network design of the proposed CNN model.

predictions using the task of predicting a pixel value from its neighbors. Our CNN model has one convolution layer with 16 kernels of size 3×3. After the convolution, we use rectified linear unit as the activation function. In the last layer, we apply one fully connected layer with 784 and 1 neurons. The input of the model is a patch from $\boldsymbol{\mu}$ with center pixel removed and to-be-estimated. The output is the estimated value of the center pixel. Mean square error (MSE) is used as training loss.

2.3 Proposed SP-CNN-T Algorithm

We construct one CNN texture prior, which is a built network saved in the computer. Now, we need to integrate this prior term with the Bayesian image reconstruction framework.

Let an auxiliary variable be z, which satisfies $\mathbf{z} = \boldsymbol{\mu}$. Equation (3) can be expressed as

$$\hat{\boldsymbol{\mu}} = \arg \min_{\boldsymbol{\mu}} - \sum_i \{Y_i \ln(b_i e^{-[\mathbf{A}\boldsymbol{\mu}]_i} + \sigma_i^2) - (b_i e^{-[\mathbf{A}\boldsymbol{\mu}]_i} + \sigma_i^2)\} + \lambda R(z), \quad \text{s.t. } z = \boldsymbol{\mu}. \quad (6)$$

By half–quadratic splitting method, we have

$$\hat{\boldsymbol{\mu}} = \arg \min_{\boldsymbol{\mu}} - \sum_i \{Y_i \ln(b_i e^{-[\mathbf{A}\boldsymbol{\mu}]_i} + \sigma_i^2) - (b_i e^{-[\mathbf{A}\boldsymbol{\mu}]_i} + \sigma_i^2)\} + \lambda R(z) + \frac{\beta}{2} \|z - \boldsymbol{\mu}\|^2, \quad (7)$$

where the quadratic-type penalty was used. Here, β is the penalty parameter, which varies in a nondescending order. Then, Eq. (7) can be solved by the following splitting iterative scheme:[31]

$$\begin{cases} \boldsymbol{\mu}_{k+1} = \arg \min_{\boldsymbol{\mu}} - \sum_i \{Y_i \ln(b_i e^{-[\mathbf{A}\boldsymbol{\mu}]_i} + \sigma_i^2) - (b_i e^{-[\mathbf{A}\boldsymbol{\mu}]_i} + \sigma_i^2)\} + \frac{\beta}{2} \|\boldsymbol{\mu} - z_k\|^2 \\ z_{k+1} = \arg \min_z \frac{\beta}{2} \|z - \boldsymbol{\mu}_{k+1}\|^2 + \lambda R(z) \end{cases} \quad (8)$$

where k indicates the iteration number. Using Eq. (8), the fidelity term and regularization term are decoupled. The fidelity term is associated with a quadratic regularized minimization problem, i.e., the first part of Eq. (8). Surrogate function strategy was used to detangle pixels in the first part of Eq. (8) as done in the study of Gao et al.[33]. Replacing the prior term of Eq. (4) from the study of Gao et al.,[33] we obtained the updated formula for the first part of Eq. (8)

$$\boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k - \frac{\sum_i \frac{\mathbf{A}_{ij}}{\sum_j \mathbf{A}_{ij}} b_i e^{-l_i^k} \left(\frac{Y_i^{\mathbf{A}}}{b_i e^{-l_i^k} + \sigma_i^2} - 1 \right) + \beta(\boldsymbol{\mu}_k - z_k)}{\sum_i \mathbf{A}_{ij} c_i(l_i^k) \left(\sum_j \mathbf{A}_{ij} \right) + \beta}. \quad (9)$$

The regularization term is involved in the second part of Eq. (8) with two terms: $\beta \|z - \boldsymbol{\mu}_{k+1}\|^2$ constrains that the solution z should not be too far away from $\boldsymbol{\mu}$ updated by the fidelity. $\lambda R(z)$ constrains that z should satisfy some neighborhood relationship, which is produced by the CNN texture prior defined in Eq. (5), i.e., substituting R with $\|z - G\|^2$, we obtained

$$z_{k+1} = \arg \min_z \frac{\beta}{2} \|z - \boldsymbol{\mu}_{k+1}\|^2 + \lambda \|z - G(\boldsymbol{\mu}_\Omega^{k+1})\|^2 \quad (10)$$

$$\mu_{k+1} = \arg \min_{\mu} - \sum_i \left\{ Y_i \ln\left(b_i e^{-[A\mu]_i} + \sigma_i^2\right) - \left(b_i e^{-[A\mu]_i} + \sigma_i^2\right)\right\} + \frac{\beta}{2} \|\mu - z_k\|^2$$

μ_k: Generated CT in iteration k

z_k: Reconstructed CT

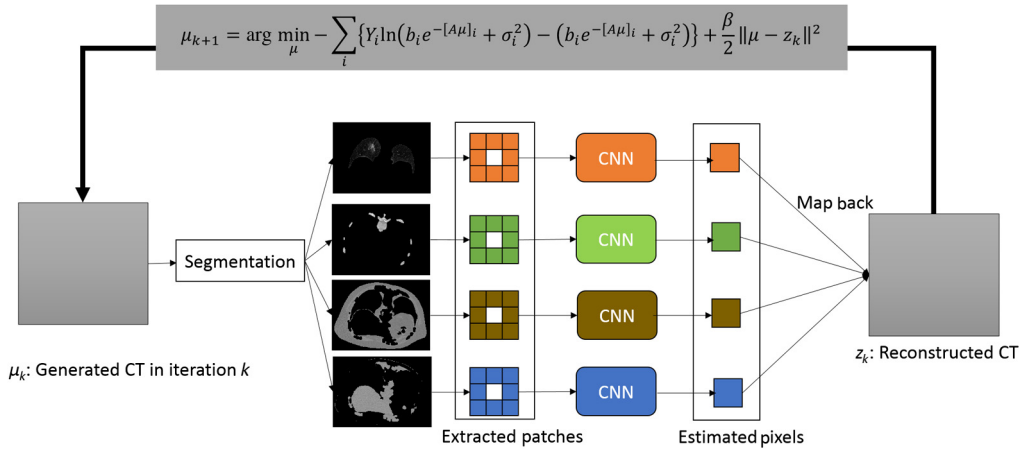Extracted patches   Estimated pixels

Fig. 2 Architecture of the proposed SP-CNN-T.

where the above equation builds the connection between the proposed CNN texture prior and the data fidelity constraint under Bayesian framework. Quadratic form penalty was used for both constraint and prior knowledge. As introduced above, the CNN texture prior $G(.)$ was trained based on the FdCT. During the iteration, it takes the intermediate of low-dose CT (LdCT) and is used to predict the center pixel based on its neighboring pixels of LdCT. This is the key point of incorporating prior knowledge from FdCT into LdCT reconstruction. Unlike universal model for the whole image, such as Huber MRF weights, it considers the tissue types. In addition, this learning method makes the tissue-specific texture more image-adaptive, which means it can recognize more substructure. This is further discussed in Sec. 3.

Equation (10) could be simplified to Eq. (11) with α representing the balance between fidelity and texture constraints

$$z_{k+1} = \arg \min_z \|z - \mu_{k+1}\|^2 + \alpha \|z - G(\mu_\Omega^{k+1})\|^2, \quad \alpha = 2\lambda/\beta. \tag{11}$$

The solution of Eq. (11) is

$$z_{k+1} = \frac{\mu_{k+1} + \alpha G(\mu_\Omega^{k+1})}{1 + \alpha}. \tag{12}$$

Initially both β and α are set to zero. The attenuation map with zero initial was updated only by the fidelity term. After about 500 iterations, we start to increase the β. Empirically, we varied β in an exponentially increasing way, meaning β changes slowly at the beginning, then much faster at later stages. We used $\beta = 1.23^k$, where k is the iteration number since we start to change β. The β changed from around 1 to 10^7 before the convergence criteria were met. The convergence criteria are MSE between z and $\mu < 1 \times 10^{-6}$. As discussed above, α represents the balance between fidelity and prior term. We choose $\alpha = 1$, which gives us current reasonably good results.

The flowchart of the proposed SP-CNN-T is shown in Fig. 2. The pseudocode of SP-CNN-T is summarized in Table 1.

3 Experiments and Results

Two patients, who were scheduled for CT-guided lung nodule needle biopsy at Stony Brook University Hospital, were recruited for this study under informed consent after approval by the institutional review board. The patients were scanned with full-dose settings, the voltage at 120 kVp, and the tube current at 100 mAs. From here, we obtained the raw transmission data of two sets of FdCT for the following study.

Based on two patients' scans, we first used one dataset to study the feasibility of constructing texture prior using the ML. Then, without changing any network design in the ML model

Table 1 Pseudocode of the proposed SP-MRF-T algorithm.

Algorithm: SP-CNN-T
Learn the tissue-specific MRF coefficients:
Segment the full-dose filtered back projection (FBP) reconstructed image into four tissue types.
Learn the MRF coefficients for each tissue type using the CNN model.
Prelog reconstruction for ULdCT:
Initialization μ by FBP.
Set parameters: α and β
While stop criterion is not met:
Update μ by Eq. (9).
Segment μ to lung, fat, muscle, and bone.
for pixel j inside lung, fat, muscle, and bone:
determine tissue type for pixel j;
update z by Eq. (12).
End
End until the stop criterion is satisfied.

and parameters in the image reconstruction, we performed the experiments on the other dataset. It not only validates the proposed method in a more diverse data but also helps us further understand whether the proposed CNN prior model is sensitive to the data variation or not. In addition, based on one patient's scans, we first performed experiments using the learnt CNN prior and simulated ULdCT transmission data from the same FdCT slice, which means that the machine has "seen" the exact image during the training. Then, we performed the experiments using the CNN prior from the neighboring adjacent slices for the target slice reconstruction. Because this proposed tissue-specific CNN prior does not require strict registration between previous FdCT and current ULdCT scan, the second experiment can investigate the performance with the mismatch of slice location and the tissue regions.

3.1 CNN-Based Tissue-Specific Texture Prior

Following the instruction in Sec. 2, we first segmented the FdCT image into four tissue types. For each tissue type, patches with 7×7 size were extracted and used as the training samples. For the experiment data, there are around 20,000 training samples for lung, fat, and muscles, and around 2500 training samples for bone region. To train CNN models, we select RMSprop optimizer,[34] with learning rate set to 5×10^{-5}. Early stop[35] is adapted to prevent overfitting in training. We initialize the weights of convolution kernels with the method introduced in the study of He et al.[36]

Figure 3 shows the training performance of the CNN model for lung region. Figure 3(R) refers to reference image, which is the ground truth. Figure 3(a) shows the CNN-predicted image by inputting patch of reference image and center pixels removed and Fig. 3(b) shows the linear MRF model-predicted image with the same input of CNN model. The structure predicted by CNN model keeps sharp at the edges and the blood vessels can also be seen clearly. There is a blurring in analytical MRF texture model, since we retreated the whole lung tissue as a region previously. Comparing the nodule region (yellow box in Fig. 3) of the two prediction models with ground truth, the CNN model can effectively learn the texture consistency constraint with better the local structures preserved.
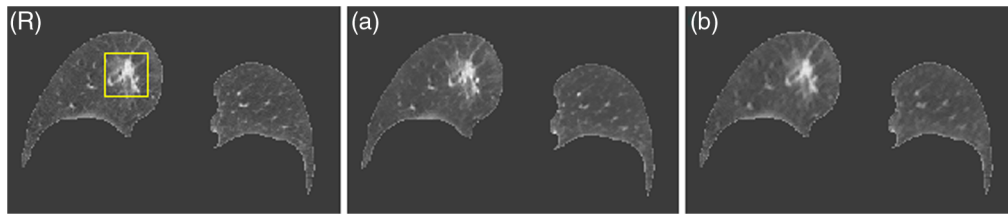
Fig. 3 The training performance of the CNN model for lung. (R) Reference image (ground truth). (a) Reference image predicted by CNN model. (b) Reference image by the linear analytical model. The display window is $(-0.007, 0.022)$ mm^{-1}.

3.2 ULdCT Image Reconstruction

From the transmission data, we simulated the transmission data of the ULdCT level based on the physics mechanism of the CT signal generation. The signal is approximated as a combination of Poisson statistics and Gaussian statistics. The mean of the Poisson distribution for one certain mAs level can be obtained from the semiexperimental model in Refs. [27] and [37], where the photon count per ray is linearly dependent with the x-ray current settings at the fixed voltage setting. We then added Poisson noise to the x-ray quanta and combined it with the Gaussian-distributed electronic noise. In this study, the mean and standard deviation of the electronic noise are 10 and 25, respectively, similar to the setup in the study of Li et al.[38]

For comparison purpose, we implemented our previous proposed SP-MRF-T method, which use the SP model for the fidelity term and linear analytical MRF texture model for the prior term in the cost function. Moreover, SP-No Prior (with no prior applied) and SP-Huber-MRF (Huber MRF weights[6]) were also employed for comparison. The element of well-established Huber MRF weights is inversely proportional to the Euclidean distance between two voxels, which consider the piecewise smoothing spatially invariantly inside one region and preserve the edge sharpness of the region. For the analytical prior algorithms, all of them are MRF-type weights and with the same window size. They share the same optimized parameters. For example, the reconstruction stopped after 1500 iterations. The hyperparameter to tune the prior strength is set as 800. For CNN prior, the parameters have been discussed above.

The reconstruction results are compared in Fig. 4. First row shows the reference images. Second to fifth rows are the results of SP-No Prior, SP-Huber-MRF, SP-MRF-T, and SP-CNN-T, respectively. First and second columns are the reconstructed images and their zoomed-in images with display window $(0, 0.035)$ mm^{-1}. Third and fourth columns are difference images comparing to the reference images with display window $(-0.001, 0.020)$ mm^{-1}. The nodule region, which is also our region of interest (ROI), is labeled in yellow box. In addition, one muscle region is also marked in a red circle for comparison. Without prior, high-level random noise is observed in images of SP-No Prior. It is clear that the prior can effectively suppress the noise. Comparing the ROI region, the SP-CNN-T outperforms other methods with least difference image.

Since the nodule is of the most clinical interest, we then focus our evaluation on the nodule. In this case, the nodule is glass-ground like, which is an important indicator for high chance to be cancer. We compared the nodules reconstructed by the aforementioned four methods. The comparison is presented in Fig. 5. According to Fig. 5(a), without any prior, the structure is buried in the noise. It is hard to identify the edges and blood vessels. The noise is suppressed significantly by applying Huber MRF prior in Fig. 5(b) and tissue-specific MRF prior in Fig. 5(c). However, the tiny structures were blurred to some extent. For example, the contrast of the "glass edge" of nodule (labeled by the yellow arrow in Fig. 5) is hard to identify. Our proposed CNN texture prior preserves the tiny structures well while suppressing the noise.

Quantified measures of the ROI are summarized in Table 2. We use the MSE, peak signal-to-noise ratio (PSNR), structure similarity (SSIM), feature similarity (FSIM),[39] and Haralick feature (HF)[40] metrics to quantify the image quality. MSE gives a quality score based on the pixel-based variation. PSNR evaluates the image quality in terms of noise suppression. SSIM evaluates the image similarity at the structure-based stage. FSIM mimics the human visual system to evaluate the image quality.[39] HF gives the quality score from the task-based evaluation view, since HFs
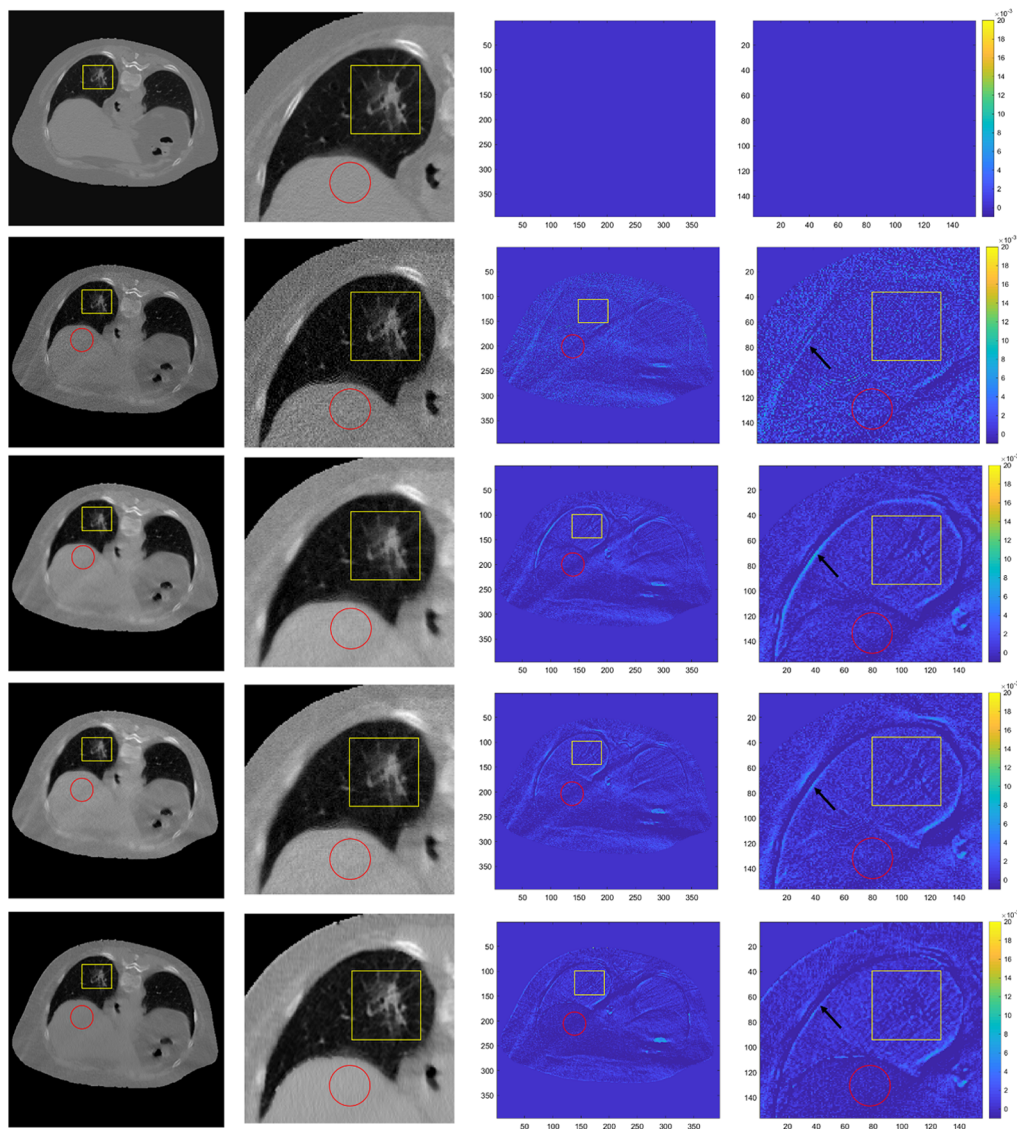
Fig. 4 Comparison of reconstructed images using different prior models. First row consists of reference images. Second to fifth rows contain results of SP-No Prior, SP-Huber-MRF, SP-MRF-T, and SP-CNN-T, respectively. First and second columns are reconstructed images and their zoomed-in images with display window $(0, 0.035)$ mm^{-1}. Third and fourth columns are difference images comparing to the reference images with display window $(-0.001, 0.020)$ mm^{-1}.

have been widely and successfully used in various clinical tasks.[40,41] Since the HF is a set of quantities or in a vector form, we use the Dice coefficient to quantify the quality.[42] For MSE, a smaller value means better image quality. For the other four metrics, higher values mean better image quality.

According to Table 2, quantified measures agree well with visual inspection. Without any prior, the image quality was worst for the five metrics. SP-Huber-MRF method improved the image quality. SP-MRF-T method achieved a further improvement over SP-Huber-MRF method, which agrees well with our previous study.[4–6] The proposed SP-CNN-T outperforms the other models for all five metrics, which demonstrates the potential of ML to further improve the analytical model. There may be two main reasons for the gain of the CNN-T over MRF-T. One is the linear assumption as mentioned in Sec. 1. The ML model removes this assumption and could be any form inside its model, which may bring in the gain by higher order approximation. Another reason could be the power of ML model in texture pattern or substructure recognition. Therefore, the CNN-T can adapt the weights in the model according to the texture pattern
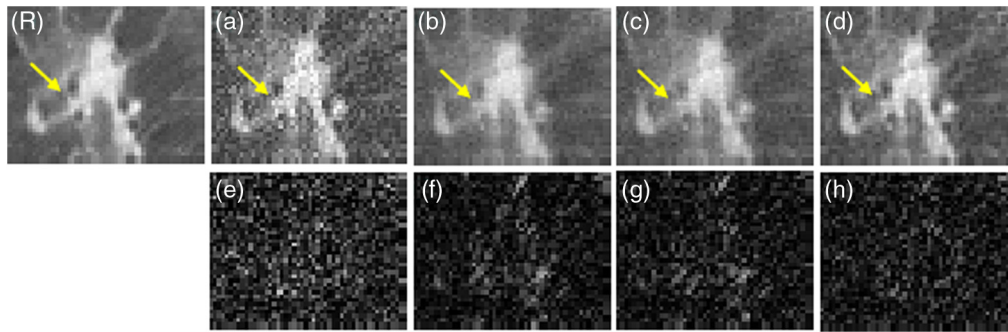
Fig. 5 Comparison of ROI in reconstructed images using different prior models. (R) Full-dose image reconstructed by the FBP, which is treated as the golden standard reference. (a)–(d) The used priors are no prior, Huber MRF prior, tissue-specific MRF prior, and CNN prior, respectively. The display window is $(-0.007, 0.022)$ mm^{-1}. (e)–(h) Difference between (a)–(d) and the reference (R).

Table 2 Quantified measures ROI in reconstructed images using different prior models.

	MSE	PSNR	SSIM	FSIM	HF
SP-No Prior	4.4437×10^{-6}	53.5226	0.9946	0.9914	0.1333
SP-Huber-MRF	1.9420×10^{-6}	57.1175	0.9971	0.9935	0.2500
SP-MRF-T	1.9098×10^{-6}	57.1902	0.9972	0.9946	0.4444
SP-CNN-T	1.3760×10^{-6}	58.6137	0.9980	0.9960	0.6667

described by the neighboring voxels. For example, there are some subsubstructures in the lung region, such as vessel, alveolus, etc. The CNN model can somehow recognize such substructures and adapt the corresponding weights to make the texture prior more accurate. Similar work was also explored in our previous analytical MRF texture model.[4] In the initial work, we derived the texture MRF weights for each individual patch of the FdCT and then used the weights as prior for the corresponding patch in its LdCT image reconstruction.[4] This patch-based MRF texture prior works perfectly, but it requires strict image registration between the FdCT and LdCT. Then, we developed the regional texture MRF model, which derives one set of MRF weights from the whole tissue and applies it into that tissue region, e.g., lung, bone, fat, and muscle.[6] Despite the small variation within the same tissue region, the regional MRF texture prior model works well.[7,8] Although the patched-based model works better than the regional model, the regional model makes it possible in practical use. Similarly, the ML model may have the ability to recognize substructures within the tissue region and thus works better than the analytical model, where one set weight is used for the whole tissue region.

3.3 Robustness Study

The proposed SP-CNN-T outperforms other comparison methods in one of our experimental dataset. Without changing any settings in the CNN training model and hyperparameter values in the image reconstruction, we evaluated this method on another dataset. Since experiments above have already shown the superiority of the texture prior, we now mainly compare the analytical texture prior model with the ML prior model.

The reconstruction results of another dataset are shown in Fig. 6. The result on top is from SP-MRF-T method, and of the bottom from SP-CNN-T. For each row, from left to right, the picture is reconstructed image, the image difference from FdCT (reference), the image difference of lung from FdCT, and the image difference of nodule from FdCT, respectively. The nodule region is marked in yellow box. The difference image shows larger errors at lung nodule (labeled by the red arrows) in SP-MRF-T than SP-CNN-T but comparable error in the rest nodule region.
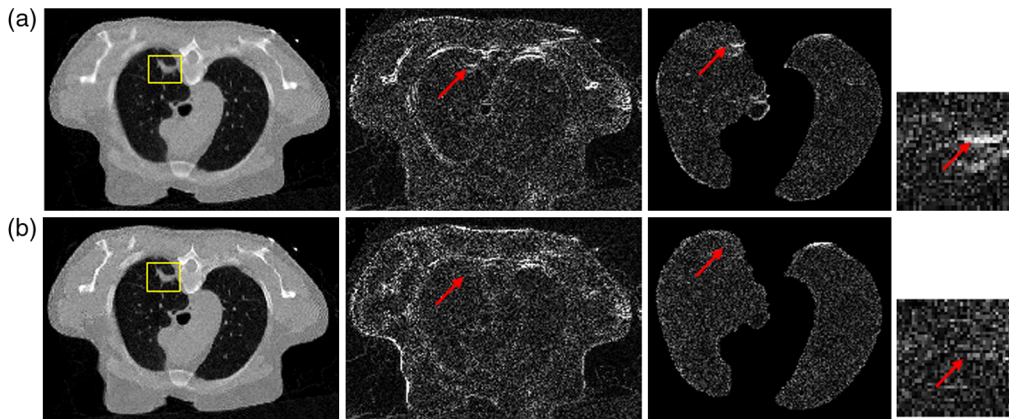
Fig. 6 Comparison of reconstructed images by (a) SP-MRF-T and (b) SP-CNN-T. For each row, from left to right, the picture is the reconstructed image [the display window is $(0, 0.035)$ mm^{-1}], the image difference from FdCT (reference), the image difference of lung from FdCT, and the image difference of nodule from FdCT. The nodule is labeled in yellow box.

Table 3 Quantified measures ROI in reconstructed images using different prior models.

Methods	MSE	PSNR	SSIM	FSIM	HF
SP-MRF-T	2.8991×10^{-6}	55.3773	0.9941	0.9958	0.3529
SP-CNN-T	1.6470×10^{-6}	57.8329	0.9977	0.9982	0.6000

This also indicated our analysis above. MRF-T and CNN-T perform similarly in most of tissue region, but differently at some substructure location. Quantified measures of the ROI are presented in Table 3. The metric values agree well with our visual judgment and also agree with our observation on the other dataset. Improved performance was observed in SP-CNN-T over SP-MRF-T through five quantified metrics in terms of noise suppression and texture preservation. Possible reasons for this gain have been discussed above in reference to the study by Zhang et al.[4] because of their consideration of the neighboring information.

Experimental results from both datasets agree well with each other. First, we demonstrate the superior performance of the proposed SP-CNN-T with different types of lung nodules. Second, the "learning" network designed based on the CNN is stable to the data variation. However, comparing both cases, we found that the training samples of each tissue type are comparable. For example, the lung region area of two patients is very similar. To further study the sensitivity of the network may require various data samples. For the hyperparameter, we chose it as 800 according to our previous experience. Usually, it will give us comparable and reasonable results when the parameter is within one range. There are also some studies[43,44] to tune the hyperparameters using the ML method. This will be one of our future interests.

3.4 CNN-Based Tissue-Specific Texture Prior from Adjacent Slices

In our previous study,[6] it is observed that the tissue textures from the nearest three neighboring slices are very similar to that of the target slice. This observation relieves the constraint of strict slice registration. Therefore, we further investigate the feasibility of extracting the CNN-T prior from the neighboring slices for the target slice. We first use the CNN prior, which was trained by adjacent slice (CNN-TA) around the target slice of the full-dose scan to predict the target slice of the FdCT. Then, we applied the CNN-TA prior to reconstruct the ULdCT images of the target slice.

We denote the target as slice 11, of which two adjacent slices 12 and 13 are used to train the CNN model. Then, we use the trained CNN-TA to predict target slice directly by feeding in the patches of slice 11. Figure 7 shows the zoomed ROI of predicted target slice from CNN models
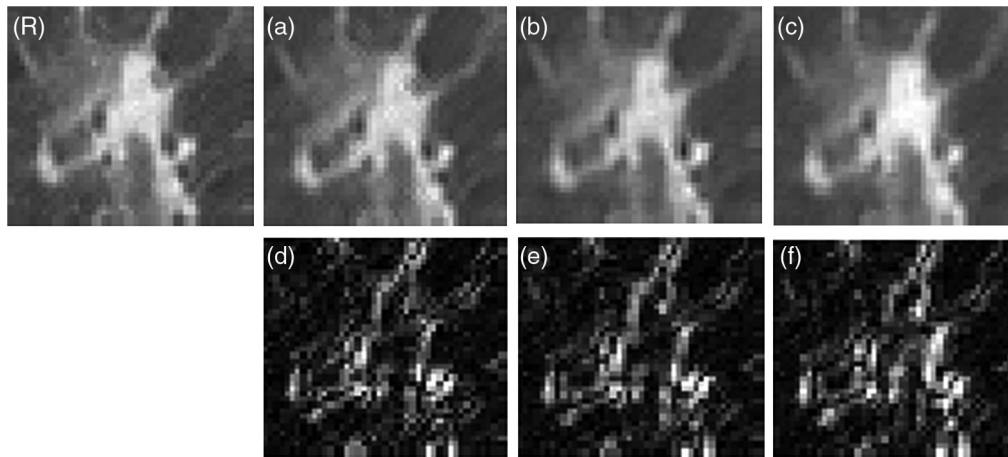
Fig. 7 Comparison of ROI in full-dose reconstructed images predicted by different CNN prior. (R) Full-dose image. (a)–(c) The used CNN priors are trained by slices 11 (target slice), 12 (adjacent slice), and 13 (adjacent slice) of the full-dose scans, respectively. The display window is $(-0.007, 0.022)$ mm^{-1}. (d)–(f) The difference between (a)–(c) and the reference (R).

trained by different slices as well as their difference to the reference. By visual assessment, the predicted images are very similar to each other, which agreed with our expectation that the adjacent two slices share the similar texture. From their difference, the farthest slice predicted images with relatively large difference comparing to the other two slices as expected.

We use the CNN-TA as the prior in the SP-CNN-T algorithm to reconstruct the ULdCT scan of slice 11. The zoomed ROIs of reconstructed images are shown in Fig. 8. The CNN-TA has a similar performance with the CNN-T, where no obvious difference is shown among Figs. 8(a)–8(c). However, the difference images of Figs. 8(d)–8(f) show dissimilar pattern compared with Figs. 7(d)–7(f). In Fig. 7, the difference images have larger values around structure. The difference images of Fig. 5 seem following random distribution. This observation indicates the effect of different noise strength of the FdCT and ULdCT. In both FdCT and ULdCT cases, the CNN prior trained by adjacent slices can be used for the reconstruction of the target slice. In other words, no strict registration is required for the proposed CNN texture prior (Fig. 9).

Quantitative measures on the results of different reconstruction methods mentioned above are summarized and shown in Fig. 7. The quantified measures agree well with our visual judgment. Overall, the proposed texture prior whether analytical or CNN type performs better than other comparison models in terms of noise suppression and texture preservation. The CNN texture
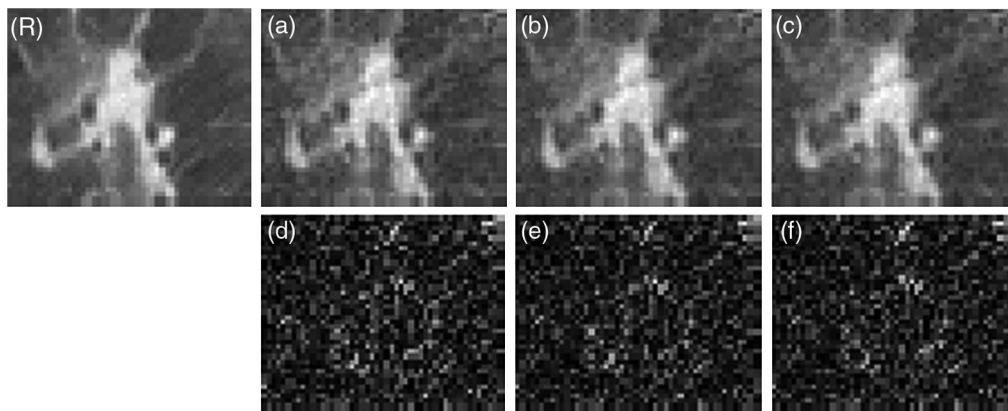

Fig. 8 Comparison of ROI in low-dose reconstructed images using different CNN prior. (R) Full-dose image reconstructed by the FBP, which is treated as the golden standard reference. (a)–(c) The used CNN priors are trained by slices 11 (target slice), 12 (adjacent slice), and 13 (adjacent slice) of the full-dose scans, respectively. The display window is $(-0.007, 0.022)$ mm^{-1}. (d)–(f) The difference between (a)–(c) and the reference (R).
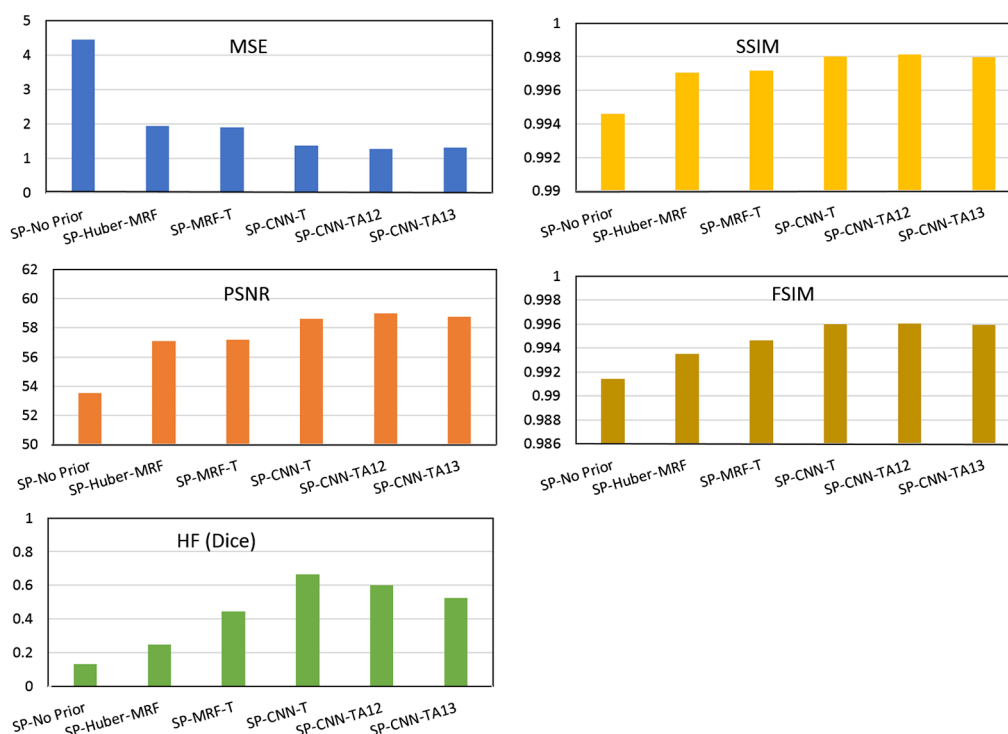
Fig. 9 Performance comparison of different methods with MSE, PSNR, SSIM, and Haralick measures.

prior performs better than the analytical texture model, which fits our expectation that the CNN prior removes the line assumption of the analytical prior, which may enable it a better approximation of texture pattern. Similar performance of CNN prior from adjacent slices is observed, which also makes the proposed CNN-T friendly and applicable in practice.

4 Conclusions and Discussion

In this paper, we proposed a way of constructing tissue-specific texture prior by ML from previous FdCT scan for current ULdCT image reconstruction. This work expanded our previously proposed analytically derived texture prior by removing the linear assumption of the previous model.[6] We further integrated the machine-learnt texture, CNN-T, with the prelog SP model (SP-CNN-T) under the Bayesian law for ULdCT imaging reconstruction. The proposed SP-CNN-T method can adaptively extract the MRF textures because of its benefit from the power of CNN in pattern or substructure recognition. By applying this adaptive tissue texture prior, the nodule texture can be preserved significantly. The feasibility and potential of this investigated ML approach is evaluated by clinical patient datasets.

This paper utilized CNN to learn information from FdCT as *a prior* for ULdCT image reconstruction. This method can be adapted for applications with available previous FdCT, e.g., interval evaluation, image-guided interventions, dynamic studies, biopsy, etc. Even though promising results have been demonstrated, a task-based evaluation is needed by more clinical datasets for each specific clinical task. This is one of our future research tasks. Another remaining issue that shall be addressed is the availability of previous FdCT. Although the proposed CNN-T does not require strict registration between the FdCT and ULdCT, it still extracts information from its previous FdCT, which may not always be available in all clinical cases, such as lung screening. We have performed a pilot study of extracting texture information from previous FdCT database to remove this constraint based on the analytical MRF texture model,[7] where encouraging results have been shown. In the future, we will also develop this CNN-T model to extract prior information from FdCT database and evaluate the developed model toward its application in clinical practice.

Disclosures

The authors of this work declare no conflicts of interest.

Acknowledgments

References

1. I. Elbakri and J. Fessler, "Efficient and accurate likelihood for iterative image reconstruction in x-ray computed tomography," *Proc. SPIE* **5032**, 1839–1850 (2003).
2. R. Gatenby, O. Grove, and R. Gillies, "Quantitative imaging in cancer evolution and ecology," *Radiology* **269**(1), 8–15 (2013).
3. F. Ng et al., "Assessment of primary colorectal cancer heterogeneity by using whole-tumor texture analysis," *Radiology* **266**(1), 177–184 (2013).
4. H. Zhang et al., "Deriving adaptive MRF coefficients from previous normal-dose CT scan for low-dose image reconstruction via penalized weighted least-squares minimization," *Med. Phys.* **41**(4), 041916 (2014).
5. H. Han et al., "Texture-preserved penalized weighted least-squares reconstruction of low-dose CT image via image segmentation and high-order MRF modeling," *Proc. SPIE* **9783**, 97834F (2016).
6. H. Zhang et al., "Extracting information from previous full-dose CT scan for knowledge-based Bayesian reconstruction of current low-dose CT images," *IEEE Trans. Med. Imaging* **35**(3), 860–870 (2016).
7. Z. Liang et al., "Different lung nodule detection tasks at different dose levels by different computed tomography image reconstruction strategies," in *IEEE Nucl. Sci. Symp. and Med. Imaging Conf.* (2018).
8. Y. Gao et al., "A feasibility study of extracting tissue textures from a previous full-dose CT database as prior knowledge for Bayesian reconstruction of current low-dose CT images," *IEEE Trans. Med. Imaging* **38**(8), 1981–1992 (2019).
9. H. Chen et al., "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE Trans. Med. Imaging* **36**(12), 2524–2535 (2017).
10. K. H. Jin et al., "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.* **26**(9), 4509–4522 (2017).
11. E. Kang, J. Min, and J. C. Ye, "A deep convolutional neural network using directional wavelets for low-dose x-ray CT reconstruction," *Med. Phys.* **44**(10), e360–e375 (2017).
12. C. Ledig et al., "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 4681–4690 (2017).
13. J. M. Wolterink et al., "Generative adversarial networks for noise reduction in low-dose CT," *IEEE Trans. Med. Imaging* **36**(12), 2536–2545 (2017).
14. H. Chen et al., "LEARN: learned experts' assessment-based reconstruction network for sparse-data CT," *IEEE Trans. Med. Imaging* **37**(6), 1333–1347 (2018).
15. Q. Yang et al., "Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss," *IEEE Trans. Med. Imaging* **37**(6), 1348–1357 (2018).
16. L. Fu et al., "Comparison between pre-log and post-log statistical models in ultra-low-dose CT reconstruction," *IEEE Trans. Med. Imaging* **36**(3), 707–720 (2017).
17. G. Wang et al., "Hybrid pre-log and post-log image reconstruction for computed tomography," *IEEE Trans. Med. Imaging* **36**(12), 2457–2465 (2017).
18. H. Erdogan and J. A. Fessler, "Monotonic algorithms for transmission tomography," *IEEE Trans. Med. Imaging* **18**(9), 801–814 (1999).
19. B. D. Man et al., "An iterative maximum-likelihood polychromatic algorithm for CT," *IEEE Trans. Med. Imaging* **20**(10), 999–1008 (2001).

20. I. Elbakri and J. Fessler, "Statistical image reconstruction for polyenergetic x-ray computed tomography," *IEEE Trans. Med. Imaging* **21**(2), 89–99 (2002).

21. J. A. Fessler et al., "Grouped coordinate ascent algorithms for penalized-likelihood transmission image reconstruction," *IEEE Trans. Med. Imaging* **16**(2), 166–175 (1997).

22. Q. Li and R. Leahy, "Statistical modeling and reconstruction of randoms precorrected PET data," *IEEE Trans. Med. Imaging* **25**(12), 1565–1572 (2006).

23. J. Wang et al., "Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose x-ray CT," *IEEE Trans. Med. Imaging* **25**(10), 1272–1283 (2006).

24. L. Kaufman, "Maximum likelihood, least squares and penalized least square for PET," *IEEE Trans. Med. Imaging* **12**(2), 200–214 (1993).

25. J. Fessler, "Penalized weighted least-squares image reconstruction for positron emission tomography," *IEEE Trans. Med. Imaging* **13**(2), 290–300 (1994).

26. B. De Man et al., "Reduction of metal streak artifacts in x-ray computed tomography using a transmission maximum a posteriori algorithm," *IEEE Trans. Nucl. Sci.* **47**(3), 977–981 (2000).

27. J. Ma et al., "Variance analysis of x-ray CT sinograms in the presence of electronic noise background," *Med. Phys.* **39**(7), 4051–4065 (2012).

28. Y. Xing et al., "Ultralow-dose CT image reconstruction with pre-log shifted-Poisson model and texture-based MRF prior," in *Int. Conf. Fully 3D Image Reconstr. Radiol. and Nucl. Med.*, pp. 254–258 (2017).

29. Y. Gao et al., "A machine learning approach to construct a tissue-specific texture prior from previous full-dose CT for Bayesian reconstruction of current ultralow-dose CT images," *Proc. SPIE* **11072**, 1107204 (2019).

30. Y. Romano, M. Elad, and P. Milanfar, "The little engine that could: regularization by denoising (RED)," *SLAM J. Imaging Sci.* **10**(4), 1804–1844 (2017).

31. K. Zhang et al., "Learning deep CNN denoiser prior for image restoration," in *IEEE Conf. Comput. Vison and Patten Recognit.*, pp. 3929–3938 (2017).

32. J. He et al., "Optimizing a parameterized plug-and-play ADMM for iterative low-dose CT reconstruction," *IEEE Trans. Med. Imaging* **38**(2), 371–382 (2019).

33. Y. Gao et al., "Characterizing CT reconstruction of pre-log transmission data toward ultralow dose imaging by texture measures," in *IEEE Nucl. Sci. Symp. and Med. Imaging Conf.*, pp. 1–4 (2018).

34. G. Hinton, N. Srivastava, and K. Swersky, "Neural networks for machine learning lecture, overview of mini-batch gradient descent," Vol. 14, p. 8 (2012).

35. F. Girosi, M. Jones, and T. Poggio, "Regularization theory and neural networks architectures," *Neural Comput.* **7**(2), 219–269 (1995).

36. K. He et al., "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Trans. Pattern Anal. Mach. Intell.* **37**(9), 1904–1916 (2015).

37. D. Zeng et al., "A simple low-dose x-ray CT simulation from high-dose scan," *IEEE Trans. Nucl. Sci.* **62**(5), 2226–2233 (2015).

38. Y. Li et al., "Radiation dose reduction in CT myocardial perfusion imaging using SMART-RECON," *IEEE Trans. Med. Imaging* **36**(12), 2557–2568 (2017).

39. Z. Lin et al., "FSIM: a feature similarity index for image quality assessment," *IEEE Trans. Image Process.* **20**(8), 2378–2386 (2011).

40. R. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst. Man Cybern.* **SMC-3**(6), 610–621 (1973).

41. Y. Hu et al., "Texture feature extraction and analysis for polyp differentiation via computed tomography colonography," *IEEE Trans. Med. Imaging* **35**(6), 1522–1531 (2016).

42. L. Dice, "Measures of the amount of ecologic association between species," *Ecology* **26**(3), 297–302 (1945).

43. T. Bai et al., "Z-index parameterization for volumetric CT image reconstruction via 3-D dictionary learning," *IEEE Trans. Med. Imaging* **36**(12), 2466–2478 (2017).

44. C. Shen et al., "Quality-guided deep reinforcement learning for parameter tuning in iterative CT reconstruction," *Proc. SPIE* **11072**, 1107203 (2019).

Yongfeng Gao received her PhD from the School of Physics at Peking University, Beijing, China, in 2018. She held a visiting scholar position in the Colliding and Accelerator Department, Brookhaven National Lab from January 2016 to October 2017. In 2017, she joined Stony Brook University as a specialist and now is a postdoctoral researcher in the Department of Radiology. Her research interests include low-dose CT reconstruction and machine learning (ML).

Jiaxing Tan received PhD in computer science from the City University of New York in 2019 with research area of computer vision and medical imaging. This work was done during his research internship at IRIS Lab, the Department of Radiology, State University of New York, Stony Brook, in 2018.

Yongyi Shi received his BSc and MSc degrees from Xi'an Jiaotong University, Xi'an, China, in 2014 and 2016, respectively. He is currently pursuing his PhD in the Institute of Image Processing and Pattern Recognition, Xi'an Jiaotong University. He was a visiting scholar in the Department of Radiology, Stony Brook University, New York, USA, in 2019. His research interests include CT reconstruction and ML.

Siming Lu received his MS degree in medical physics from Duke University in 2014, followed by two years of therapeutic medical physics residency training at Henry Ford Hospital. Then, in 2016, he became a therapeutic medical physicist at the Memorial Sloan Kettering Cancer Center. In 2019, he joined as a senior medical physicist in the Department of Radiation Oncology, Stony Brook University Hospital.

Amit Gupta completed medical school (2005) and a radiology residency (2009) in India. Then, he pursued four additional years of training at Cleveland Clinic, Cleveland, Ohio, USA, which included three clinical fellowships (one year each) in thoracic imaging, vascular and interventional radiology, and nuclear imaging and one year of clinical research fellowship in radiology. He is board certified by the American Board of Radiology and currently serving as an associate professor of clinical radiology at Stony Brook University Hospital, Stony Brook, New York, USA. The Society of Interventional Radiology has awarded him fellowship status. His research interests include postprocessing techniques to optimize quality of images obtained using low x-ray dose.

Haifang Li received his BS degree from the Department of Chemistry, University of Science and Technology, Heifei, China. He obtained his PhD in chemistry from the State University of New York, Stony Brook, USA. He is now a research faculty in the Department of Radiology, Stony Brook University.

Zhengrong Liang received his PhD in physics from the City University of New York in 1987, followed by a one-year research fellowship in nuclear medicine and radiation oncology at the Albert Einstein College of Medicine. He had been a research associate and then an assistant professor in radiology at Duke University Medical Center. He joined State University of New York, Stony Brook, in 1992 and currently holds a professorship in the Departments of Radiology and Biomedical Engineering.