Unsupervised/Supervised Hybrid Deep Learning Framework for Low Dose Phase Contrast Imaging

Guogang Zhu1,2,*, Jian Fu1,2,3 and Feng Zhao4
1School of Mechanical Engineering and Automation, Beijing University of Aeronautics and Astronautics, Beijing, 100191, China.
2Research Center of Digital Radiation Imaging, Beijing University of Aeronautics and Astronautics, Beijing, 100191, China.
3Jiangxi Research Institute, Beijing University of Aeronautics and Astronautics, Nanchang, 330000, China.
4Hongxia Chemical Co., Ltd., Hohhot, 010070, China

*Corresponding author email: buaa_zgg@buaa.edu.cn

Abstract. X-ray phase contrast computed tomography (PCCT) has better imaging quality than conventional attenuation X-ray CT and has demonstrated promising application prospects in medical diagnosis. However, reducing the radiation dose during PCCT imaging still remains a major challenge. Recently, deep learning (DL) techniques have been applied to low dose CT and obtain significant progress. Most of them require massive paired images to train the network in a supervised manner, which may hamper their practical applications because the ground-truth images are hard to be obtained in most cases. To address this issue, we report a hybrid deep learning framework for low dose PCCT which capsules unsupervised and supervised learning manners. It combines the advantages of convolutional neural network (CNN) and total variation (TV) and is suitable for both unlabelled datasets and labelled datasets. This framework has been validated and demonstrated with experimental data. It will be helpful to push the practical application of low dose PCCT.

1. Introduction
X-ray phase contrast computed tomography has been proposed over the last decades and yields better imaging contrast for soft-tissues compared to conventional CT. Several techniques have been proposed to form the phase contrast images such as crystal interferometer imaging [1], diffraction enhanced imaging [2,3,4], propagation-based imaging [5] and grating-based imaging [6,7,8,9]. Nevertheless, the aforementioned imaging techniques still exhibit significant problems with high radiation dose. Reducing tube current is an effective and universal method to reduce the radiation dose, while it will introduce quantum noise to projections and degrade the final image quality.

Many methods have been proposed to address the image degradation problem in low dose PCCT. Among them, total variation and its variants [10,11] based on sparse constraints have gained satisfactory effects. However, these methods are not capable of determining a uniform function and could only be implemented on each single CT image separately. Recently, deep learning techniques have attracted more and more attention in the field of the imaging processing, which suggests a potential future for low dose PCCT. Many researchers have applied DL techniques to low dose CT [12,13,14] and have achieved remarkable results both in image quality and processing speed. However, most of these methods are severely reliant on numerous pre-collected
paired CT images which consist of low dose CT images and their corresponding high dose ones to optimize the network. While these paired CT images may not be obtained due to limitations of human labor costs and collecting time consumptions and so on in practice. Thus it is vital to develop a hybrid framework to meet different data sets, both unlabelled data sets and labelled data sets.

In this paper, we report a hybrid deep learning framework for low dose PCCT which capsules unsupervised and supervised learning manners. It integrates the advantages of both convolutional neural network and total variation. This framework is capable of meeting different data sets and is suitable for both unsupervised and supervised learning. When there are no labelled images, the network takes sparse constrain as the prior of images to guide the training with TV regularization term. While there are labelled images, the network takes high dose as ground truth to guide the training. Data augmentation shall be implemented to enhance the data sets if necessary. Taking grating-based imaging as an example, this framework has been validated and demonstrated with experimental data obtained from the Applied Biophysics (E17), Technical University of Munich. It will promote the practical application of DL techniques in the field of low dose PCCT imaging.

The structure of remaining sections is as follows. In section 2, the overview of proposed framework as well as its implementation details including structure of network, design of loss function, reconstruction algorithm of PCCT and noise model of low dose imaging are presented. Experiments of both unsupervised and supervised learning on real experimental data are performed in section 3, which intend to demonstrate the validity of the proposed framework. Conclusion of the proposed framework is given in section 4.

2. Method

2.1. Framework Overview

As shown in figure 1, the proposed framework capsules unsupervised learning and supervised learning manners. For unlabelled images, the CNN could be trained through unsupervised learning based on sparse constraints, as shown in the upper portion of figure 1. As for labelled images, the CNN takes high dose PCCT images as ground truth to guide the training, as shown in the lower portion of figure 1. Thus, this framework can match different data sets in practice.

![Figure 1. Hybrid deep learning framework for low dose PCCT.](image)

2.1.1. Neural network. The structure of CNN is presented in figure 2. At first, the input image is fed into two successive convolutional layers indicated by blue solid rectangles to extract primary features. Then, these primary features are sent to 16 successive residual blocks surrounded by gray dashed rectangle for future feature extraction. These residual blocks are based on the idea of wide activation, which has better performance without additional computational overhead compared to common
residual blocks [15]. Finally, these high-level features are compressed to a single-channel image by convolutional layer indicated by blue solid rectangle and added with the original image to generate output image.

![Structure of the convolutional neural network.](image)

2.1.2. Loss function. The uniform loss function shown in equation (1) is the essence of the proposed framework. In this equation, $\|I_o - I_{ref}\|^2$ represents the mean square error (MSE) between the output image $I_o$ and the reference image $I_{ref}$, $\|I_o\|_{TV}$ the total variation of the output image and $\lambda_{TV}$ the weight of the TV regularization term. $\|I_o\|_{TV}$ is calculated as shown in equation (2), where $\nabla_h$ and $\nabla_v$ denote the horizontal and vertical difference operators.

For unsupervised learning manner, $I_{ref}$ is replaced by input image $I_{id}$ to constrain the difference between $I_o$ and $I_{id}$, as shown in equation (3). For supervised learning manner, $I_{ref}$ is replaced by high dose image $I_{hd}$ and $\lambda_{TV}$ is set to 0, as shown in equation (4).

$$L = \|I_o - I_{ref}\|^2 + \lambda_{TV} \|I_o\|_{TV}$$  \hspace{1cm} (1)

$$\|I_o\|_{TV} = \sqrt{(\nabla_h I_o)^2 + (\nabla_v I_o)^2}$$  \hspace{1cm} (2)

$$L_{unsup} = \|I_o - I_{id}\|^2 + \lambda_{TV} \|I_o\|_{TV}$$  \hspace{1cm} (3)

$$L_{sup} = \|I_o - I_{hd}\|^2$$  \hspace{1cm} (4)

2.1.3. Running modes. This framework has two modes: training and working. The training mode is as follows:

(1) A set of training data is fed into the aforementioned CNN. For supervised learning manner, this data set consists of low dose PCCT images and their corresponding high dose PCCT images. For unsupervised learning manner, this data set only consists of low dose PCCT images.

(2) The loss function value is calculated with equations (3) or (4) and used to optimize the network iteratively.

(3) Repeat above steps until the learning converges. In this paper, the number of iteration epochs is taken as the termination condition, which is set to 100.

In working mode, the trained CNN is fed with a low dose PCCT image and outputs improved PCCT image.

2.2. Grating-based Phase Contrast Imaging

Grating-based phase contrast imaging has shown promising application prospects since its implementation by conventional X-ray tube sources [16]. Grating-based phase contrast imaging is based on Talbot effect as shown in figure 3. From left to right, the X-ray tube, source grating G0, phase grating G1, absorption grating G2 and detector. During imaging, G2 will translate laterally...
several times and the detector will record the interferometric patterns of the sample to obtain the intensity curve. The differential phase information could be demodulated by applying the analytical or statistical analysis methods to the intensity curve.

Expressed in equation (5), two-dimensional filter back projection (FBP) algorithm is commonly adopted to reconstruct the PCCT image. In this equation, \( \delta(x, y) \) represents the phase contrast image, \( U \) the geometrical weight factor, \( \alpha \) the differential phase contrast projection, \( h \) the Hilbert filter shown in equation (6), \( v \) the frequency variant and \( \theta \) the rotation angle.

\[
\delta(x, y) = \frac{1}{2} \int_{0}^{2\pi} U \times \alpha \ast h(\nu)d\theta
\]

\[
h(\nu) = \frac{1}{2\pi} i \text{sgn}(\nu)
\]

Figure 3. The typical grating-based phase contrast imaging device. From left to right, the X-ray tube, source grating G0, phase grating G1, absorption grating G2 and detector.

2.3. Noise Model
Low dose PCCT imaging mainly suffers from the X-ray quantum noise, which could be approximated with Poisson process [17, 18]. In this paper, we adopted quantum noise model expressed in equation (7) to generate low dose PCCT projections from high dose PCCT projections. In this equation, \( P_{ld}(x, y) \) represents the value of low dose PCCT projection at point \((x, y)\), \( P_{hd}(x, y) \) the value of high dose projection at point \((x, y)\), \( \lambda \) a constant to control the noise level and \text{Poisson()} \) the Poisson sequence generator. Then low dose PCCT images can be obtained.

\[
P_{ld}(x, y) = \frac{1}{\lambda} \text{Poisson}(\lambda \cdot P_{hd}(x, y)) \tag{7}
\]

3. Experiments
3.1. Data Preparation
The experimental data is obtained from Applied Biophysics (E17), Technical University of Munich, which was carried out on a grating-based PCCT imaging device. The sample was immersed in a plastic cylinder and consisted of tissue that is common in the mammary such as fatty tissue, fibrous tissue and leafy tumor. There are 195 phase contrast images with the size 341×341 pixels in total. For unsupervised learning, we take 100 images for training and the other 95 images for testing. For supervised learning, we take 10 images and enrich them to 110 images by data augmentation for training, the other 185 images are taken for testing.

To obtain low dose PCCT images, quantum noise model described in equation (7) is first applied to projections and then FBP in equation (5) is applied to the projections to reconstruct PCCT images. For unsupervised learning, the value of the parameter \( \lambda \) in quantum noise model is set to 0.05 and 0.1, respectively. For supervised learning, the value of the parameter \( \lambda \) is set to 0.01, 0.05 and 0.1, respectively. They correspond to the typical cases of different levels of low dose noise. The \( \lambda \) value of 0.01 is used to prove the robustness of supervised learning.
3.2. Implementation
This method is implemented with Python 3.7.3 and Tensorflow 1.14.0. It runs in a workstation with a CPU Intel(R) Core(TM) i7-7700 HQ and a GPU Nvidia GTX 1060 with Max-Q Design 6GBytes. Adam optimizer algorithm [19] is applied to training. The initial learning rate is $10^{-3}$ and gradually reduced to $10^{-6}$. All the models are trained for 100 epochs and take about 2 hours to complete the training. For unsupervised learning, the value of the parameter $\lambda_{yy}$ is set to $2 \times 10^{-6}$.

3.3. Image Evaluation
According to the research of Zhang et al [20], the feature similarity (FSIM) [21] and information content weighted SSIM index (IW-SSIM) [22] are more accurate than other image evaluation methods. Therefore, we take FSIM and IW-SSIM as quantitative indexes for image quality. The FSIM is expressed in equation (8), where $I$ is the whole image domain, $S_k(x)$ the similarity at location $x$ and $PC_m(x)$ the index to weight the importance of $S_k(x)$.

$$FSIM = \frac{\sum_{x \in I} S_k(x) \cdot PC_m(x)}{\sum_{x \in I} PC_m(x)}$$

(8)

IW-SSIM can be calculated by equation (9), where $M$ is the number of scales, $\beta_j$ the scale weight.

$$IW-SSIM = \prod_{j=1}^{M} (IW-SSIM_j)^{\beta_j}$$

(9)

$IW-SSIM_j$ is the IW-SSIM measure defined as

$$IW-SSIM_j = \frac{\sum_{i,j}^w w_{j,i} c(x_{j,i}, y_{j,i}) s(x_{j,i}, y_{j,i})}{\sum_{i,j}^w w_{j,i}}$$

(10)

for $j = 1, \ldots, M-1$, and

$$IW-SSIM_M = \frac{\sum_{i,j}^I l(x_{j,i}, y_{j,i}) c(x_{j,i}, y_{j,i}) s(x_{j,i}, y_{j,i})}{N_j}$$

(11)

for $j = M$, where $N_j$ is the number of evaluation windows, $x_{j,i}$ and $y_{j,i}$ the $i$th image patches at $j$th scale, $w_{j,i}$ the information content weight, $l(x_{j,i}, y_{j,i}), c(x_{j,i}, y_{j,i}), s(x_{j,i}, y_{j,i})$ the luminance, contrast, structure similarities respectively.

3.4. Results
Figures 4 and 5 present the unsupervised learning results of one of the 95 images for testing. Figure 4 corresponds to the case with a $\lambda$ value 0.05 and figure 5 0.1. Some regions indicated with yellow squares are enlarged for better visualization. As shown in figures 3-5, the unsupervised method is capable of suppressing noise as well as reserving structure details.

Figures 6-8 present supervised learning results of one of the 185 images for testing. Figure 6 corresponds to the case with a $\lambda$ value 0.01, figure 7 0.05 and figure 8 0.1. Some regions indicated with yellow squares are enlarged for better visualization. As shown in figures 6-8, the supervised method is also capable of suppressing noise and reserving structure details. Moreover, the supervised method can deal with severer noise.
Figure 4. Results of unsupervised learning with a $\lambda$ value 0.05. From left to right, high dose PCCT image, low dose PCCT image and output image of the network. Some regions, indicated by yellow squares, are enlarged for better visualization.

Figure 5. Results of unsupervised learning with a $\lambda$ value 0.1. From left to right, high dose PCCT image, low dose PCCT image and output image of the network. Some regions, indicated by yellow squares, are enlarged for better visualization.

Tables 1 and 2 show the averaged FSIM and IW-SSIM values of unsupervised learning for testing data. Tables 3 and 4 show the averaged FSIM and IW-SSIM values of supervised learning for testing data. Both in unsupervised learning and supervised learning, these values are larger than the ones in original low dose PCCT images. It verifies the proposed framework quantitatively.

Figure 6. Results of supervised learning with a $\lambda$ value 0.01. From left to right, high dose PCCT image, low dose PCCT image and output image of the network. Some regions, indicated by yellow squares, are enlarged for better visualization.
Figure 7. Results of supervised learning with a $\lambda$ value 0.05. From left to right, high dose PCCT image, low dose PCCT image and output image of the network. Some regions, indicated by yellow squares, are enlarged for better visualization.

Figure 8. Results of supervised learning with a $\lambda$ value 0.1. From left to right, high dose PCCT image, low dose PCCT image and output image of the network. Some regions, indicated by yellow squares, are enlarged for better visualization.

Table 1. Averaged FSIM value of unsupervised learning.

| $\lambda$ | Low Dose Image | Output Image |
|-----------|----------------|--------------|
| 0.05      | 0.8971         | 0.9330       |
| 0.1       | 0.9372         | 0.9447       |

Table 2. Averaged IW-SSIM value of unsupervised learning.

| $\lambda$ | Low Dose Image | Output Image |
|-----------|----------------|--------------|
| 0.05      | 0.9253         | 0.9474       |
| 0.1       | 0.9583         | 0.9687       |

Table 3. Averaged FSIM value of supervised learning.

| $\lambda$ | Low Dose Image | Output Image |
|-----------|----------------|--------------|
| 0.01      | 0.7411         | 0.8831       |
| 0.05      | 0.8955         | 0.9384       |
| 0.1       | 0.9358         | 0.9555       |

4. Conclusion

In this paper, we report a hybrid deep learning framework for low dose phase contrast computed tomography. It is based on convolutional neural network and total variation and is suitable for both unsupervised and supervised learning. Taking grating-based imaging as an example, this framework has been validated and demonstrated with experimental data. It is helpful to push the application of low dose phase contrast computed tomography.
Table 4. Averaged IW-SSIM value of supervised learning.

| \( \lambda \) | Low Dose Image | Output Image |
|------------|----------------|--------------|
| 0.01       | 0.7685         | 0.8931       |
| 0.05       | 0.9253         | 0.9556       |
| 0.1        | 0.9574         | 0.9687       |

Acknowledgments
We acknowledge Applied Biophysics (E17), Technical University of Munich for providing the experimental data. We acknowledge support from the National Science and Technology Major Project of China (2018ZX04018001-006), the Joint Fund of Research Utilizing Large-scale Scientific Facilities by National Natural Science Foundation of China and Chinese Academy of Science (U1932111), National Natural Science Foundation of China (11574023, 51975026).

References
[1] Momose A 1995 Demonstration of phase-contrast X-ray computed tomography using an X-ray interferometer Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 352 622-8
[2] Zhu P, Wang J, Yuan Q, Huang W, Shu H, Gao B, Hu T and Wu Z 2005 Computed tomography algorithm based on diffraction-enhanced imaging setup Applied Physics Letters 87 264101
[3] Chapman D, Thomlinson W, Johnston R E, Washburn D, Pisano E, Gmür N, Zhong Z, Menk R, Arfelli F and Sayers D 1997 Diffraction enhanced x-ray imaging Phys.med.biol 42 2015-25
[4] Muehlemann C, Li J, Connor D, Parham C, Pisano E and Zhong Z 2009 Diffraction-Enhanced Imaging of Musculoskeletal Tissues Using a Conventional X-Ray Tube Academic Radiology 16 918-23
[5] Snigirev A, Snigireva I, Kohn V, Kuznetsov S and Schelokov I 1995 On the possibilities of x-ray phase contrast microimaging by coherent high-energy synchrotron radiation Review of scientific instruments 66 5486-92
[6] Zhu P, Zhang K, Wang Z, Liu Y, Liu X, Wu Z, McDonald S A, Marone F and Stampanoni M 2010 Low-dose, simple, and fast grating-based X-ray phase-contrast imaging Proceedings of the National Academy of Sciences 107 13576-81
[7] Sato G, Kondoh T, Itoh H, Handa S and Den T 2011 Two-dimensional gratings-based phase-contrast imaging using a conventional x-ray tube Optics Letters 36 3551-3
[8] Jian F, Xianhong S, Guo W and Peng P 2019 Fast X-ray Differential Phase Contrast Imaging with One Exposure and without Movements Scientific Reports (Nature Publisher Group) 9
[9] Marschner M, Willner M, Potdevin G, Fehringer A, Noël P, Pfeiffer F and Herzen J 2016 Helical X-ray phase-contrast computed tomography without phase stepping Scientific reports 6 23953
[10] Nolchian M, Vonesch C, Lefkimmiatis S, Modregger P, Stampanoni M and Unser M 2013 Constrained regularized reconstruction of X-ray-DPCI tomograms with weighted-norm Optics Express 21 32340
[11] Kostenko A, Batenburg K J, King A, Offerman S E and van Vliet L J 2013 Total variation minimization approach in in-line x-ray phase-contrast tomography Optics Express 21 12185
[12] Dong J, Fu J and He Z 2019 A deep learning reconstruction framework for X-ray computed tomography with incomplete data PloS one 14
[13] Chen H, Zhang Y, Kalra M K, Lin F, Chen Y, Liao P, Zhou J and Wang G 2017 Low-dose CT with a residual encoder-decoder convolutional neural network IEEE transactions on medical imaging 36 2524-35
[14] Fu J, Dong J and Zhao F 2019 A Deep Learning Reconstruction Framework for Differential Phase-Contrast Computed Tomography With Incomplete Data IEEE Transactions on Image Processing 29 2190-202
[15] Yu J, Fan Y, Yang J, Xu N, Wang Z, Wang X and Huang T 2018 Wide activation for efficient and accurate image super-resolution arXiv preprint arXiv:1808.08718
[16] David C, Nöhammer B, Solak H H and Ziegler E 2002 Differential x-ray phase contrast imaging using a shearing interferometer Applied physics letters 81 3287-9
[17] Chan C L, Sullivan B J, Sahakian A V, Katsaggelos A K, Swiryn S, Hueter D C and Frohlich T 1990 Simulation of quantum mottle in digital angiographic images. In: Biomedical Image Processing: International Society for Optics and Photonics) pp 104-10
[18] Schulz T J and Snyder D L 1991 Imaging a randomly moving object from quantum-limited data: applications to image recovery from second- and third-order autocorrelations Journal of the Optical Society of America A 8 801-7
[19] Kingma D P and Ba J 2014 Adam: A method for stochastic optimization arXiv preprint arXiv:1412.6980
[20] Zhang L, Zhang L, Mou X and Zhang D 2012 A comprehensive evaluation of full reference image quality assessment algorithms. In: 2012 19th IEEE International Conference on Image Processing: (IEEE) pp 1477-80
[21] Zhang L, Zhang L, Mou X and Zhang D 2011 FSIM: A feature similarity index for image quality assessment IEEE transactions on Image Processing 20 2378-86
[22] Wang Z and Li Q 2010 Information content weighting for perceptual image quality assessment IEEE Transactions on image processing 20 1185-98