Low Dose Mammography via Deep Learning

Guogang Zhu1,2,*, Jian Fu1,2,3 and Jianbing Dong1,2

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.

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

Abstract. X-ray mammography has been widely applied to breast cancer diagnosis due to its simplicity and reliability. However, X-ray will do harm to the health of patients or even cause cancer. Low dose mammography by reducing the tube current is an effective method to reduce radiation dose and has attracted more and more interests. In this paper, we implemented a method to improve the image quality of low dose mammography via deep learning. It is based on a convolutional neural network (CNN) and focuses on reducing the noise in low dose mammography. After training, it can obtain a high quality image from a low dose mammography image. This method is validated with experimental data sets obtained from The Cancer Imaging Archive (TCIA). It will promote the application of state-of-art deep learning technique in the field of low dose mammography.

1. Introduction

According to a report by American Cancer Society, breast cancer is the most common disease among women [1]. This report also illustrates that the incidence of breast cancer is much higher than other cancers for women, such as colon cancer and bronchial lung cancer. As one of the cancers with the highest morbidity and mortality [2], breast cancer has become a serious threat to women's health. Among various methods for breast examination, X-ray mammography is the most effective and reliable method to diagnose breast diseases which is based on penetrability of X-ray [3]. However, excessive doses of X-ray can harm to the body and even induce cancer [4]. The radiation dose can be reduced by lowering the tube current, but it will cause image degradation and introduce spot, reticulate grain, snowflake shape and other abnormal structures to mammography image. The reason for image degradation is X-ray quantum noise, which depends on pixel values acquired from detector and approximately obeys Poisson distribution [5]. In this paper, Poisson noise is used to simulate quantum noise in low dose mammography and deep learning technique is used to remove the quantum noise.

In recent years, deep learning (DL) technique has made a series of key breakthroughs in the field of image processing such as segmentation [6], image classification [7], image denoise [8] and so on. CNN has attracted much more interests due to its characteristics of sparse connection, local connection and parameter sharing. It has a promising future towards low dose mammography.

In this paper, we implemented a CNN method to improve the image quality of low dose mammography. It works at two stages: training and working. In training, the low dose images and the corresponding normal dose images are fed into the CNN to optimize the network parameters iteratively. In working, the trained CNN takes low dose images as input and outputs the high quality
image with a reduction on noise. This technique was validated with experimental data sets obtained from TCIA. It will promote the application of advanced DL technique in low dose mammography. In the following sections, the implemented CNN framework is first outlined. Then, the noise model is described. Next, the data sets, and implementation details are presented. Finally, the results are given out and discussed.

2. Method

2.1. Neural Network

The architecture of the used CNN is modified from our previous work [9]. It takes fully advantages of U-net [10] and DenseNet [11] to remove the noise caused by low dose. As depicted in figure 1, the CNN is fed high noise image and outputs low noise image. It consists of four function components: i) The first one is the two layers of filters indicated by the two yellow solid rectangles surrounded with a red dashed rectangle. They are used to extract primary features. ii) The next two parts surrounded by two grey dashed rectangles are DenseBlock function components. They act as encoders and obtain multi-scale features. iii) The two parts surrounded by two green dashed rectangles are decoder components. They restore high resolution feature images from the low resolution images generated by the corresponding encoders. iv) The last component is the channel compression layer with a filter size $1 \times 1$, indicated by the red solid rectangle.

![Figure 1. The architecture of the used CNN. It consists of four function components, indicated by the red, grey, green dashed rectangles and a red solid rectangle.](image)

2.1.1. Skip connection. In the CNN in figure 1, skip connection technique [12] is adopted to address the learning degradation in training. It could be formulated as equation (1). Here, $F(x)$ represents the underlying mapping with few stacked layers. $y$ is obtained by concatenating the identity mapping of input $x$ and $F(x)$.

$$y = F(x) + x$$

(1)

2.1.2. DenseBlock. The structure of DenseBlock is shown in figure 2. Each layer takes all preceding layers as inputs and sends its feature maps to all subsequent layers. It is helpful to condense models and could obtain high learning accuracy and efficiency due to the feature reuse. With a strong connection, the global features could reach any component within the network. It makes the network easier to train and enables state-of-art performance with fewer parameters.

2.1.3. Parameters optimization. All parameters in the CNN are initialized by Guassian distribution with mean value zero and standard deviation $\sqrt{\frac{2}{n_{in}}}$ in which $n_{in}$ is the number of input units.
2.1.4. Loss function. Multi-scale structure similarity (MS-SSIM) is adopted to estimate the structure similarity between outputs and inputs and build the loss function. The bigger the MS-SSIM value is, the better the similarity of the two images become. The loss function is expressed in equation (2). Here, \( Y \) denotes the output image, \( \hat{Y} \) the corresponding normal dose mammography image.

\[
\text{loss} = 1 - \text{MS-SSIM}(Y, \hat{Y})
\]  

(2)

![Figure 2. The structure of DenseBlock.](image)

2.2. Noise Model

The behaviour of receiving photons by the detector can be approximated with Poisson process. Low dose mammography mainly suffers from the quantum noise of X-ray. There exists several quantum noise models for low dose X-ray [13,14]. In this paper, we adopted quantum noise model expressed in equation (3) to generate low dose mammography images from normal dose mammography images. \( I_{ld}(x,y) \) represents the gray value of low dose image at point \( (x,y) \), \( I_{nd}(x,y) \) the gray value of normal dose image at point \( (x,y) \), \( \lambda \) a constant and \( \text{Poisson}() \) the Poisson sequence generator.

\[
I_{ld}(x,y) = \frac{1}{\lambda} \text{Poisson}(\lambda \cdot I_{nd}(x,y))
\]  

(3)

2.3. Running Modes

This method works at two modes: training and working. The training mode is as follows:

1) A set of training data consisting of low dose mammography images and its corresponding ground truth are fed into the CNN depicted in figure 1 successively.
2) For each pair of data, the MS-SSIM value is calculated between the output image and its corresponding ground truth and the network parameters are optimized by gradient descent.
3) One outer learning iteration is accomplished once all training data are used.
4) Repeat steps 2) and 3) until the training requirements are met.

In working mode, a low dose mammography image is fed into the CNN determined by training. The CNN will output high quality mammography image with a reduction on noise.

3. Experiments

3.1. Data Preparation

The experimental data sets are obtained from Curated Breast Imaging Subset of Digital Database for Screening Mammography (CBIS-DDSM) [15,16] in The Cancer Imaging Archive [17]. This data sets include 10239 images. We randomly chose 500 images for training and 50 images for testing. The image has a size 505 × 298 pixel.

Quantum noise is added into the mammography images by using the noise model discussed in subsection 2.2. To improve the training efficiency and accuracy, normalization is applied to low dose images and normal dose images. It is expressed in equation(4) in which \( I \) denotes the normalized image, \( I \) the original image, \( \text{min}() \) and \( \text{max}() \) the operator to obtain minimum and maximum value of the image respectively.
\[ I_n = \frac{I - \min(I)}{\max(I) - \min(I)} \]  

3.2. Implementation

This method is implemented with Python 3.5.2 and Tensorflow 1.8.0. It runs in a workstation with a CPU i7 9700 and a GPU Nvidia GTX 1080Ti 11GBytes. The parameter \( \lambda \) is set to be 0.001, 0.005, 0.01 to generate three levels of quantum noise. They correspond to three typical cases range from lower noise to higher noise that makes most of the details of the image disappear, which are used to prove the robustness of the method. These three generated data sets are used to train a model separately.

The popular denoising algorithm BM3D[18] is taken as the comparison of the proposed method. Anscombe transformation is adopted to transform the Poisson distribution into Gaussian distribution. Adam optimizer algorithm [19] is applied to training. The batch size is 4 and the learning rate is 0.001. All the models are trained for 100 epochs and take about 1 hour to complete the training.

3.3. Image Evaluation

Zhang et al [20] carried out research on several image evaluation methods. The results show that feature similarity (FSIM) [21] and information content weighted SSIM index (IW-SSIM) [22] are more accurate than others. Therefore, FSIM and IW-SSIM are used to execute image evaluation.

3.4. Results

Figures 3-5 present results of one of the 50 mammography image for testing. Figure 3 corresponds to the case with a \( \lambda \) value 0.001, figure 4 0.005 and figure 5 0.01. Some regions indicated with red boxes are enlarged for better visualization.

As shown in figures 3-5, the implemented method has the ability to suppress quantum noise. Meanwhile, the structure details are reserved. Comparison results with BM3D also demonstrate the validity of proposed method.

Tables 1 and 2 show the averaged FSIM and IW-SSIM values for testing. The values of the proposed method are significantly larger than that of the original low dose mammography images and BM3D. These results validate the proposed method quantitatively.

4. Conclusion

In this paper, we implemented a CNN method to suppress the noise and improve the image quality of low dose mammography. It is built based on U-net and DenseNet. After training, the CNN could obtain the high quality image from a low dose mammography image. This method is validated with experimental data sets of CBIS-DDSM obtained from TCIA. The result shows that this method is effective to remove the quantum noise. This work will promote the application of advance DL techniques in the field of low dose mammography.

![Figure 3](image-url)  

**Figure 3.** Results with a \( \lambda \) value 0.001. From left to right, normal dose, low dose, proposed method and BM3D. Some regions, indicated by red boxes, are enlarged for better visualization.
Figure 4. Results with a $\lambda$ value 0.005. From left to right, normal dose, low dose, proposed method and BM3D. Some regions, indicated by red boxes, are enlarged for better visualization.

Figure 5. Results with a $\lambda$ value 0.01. From left to right, normal dose, low dose, proposed method and BM3D. Some regions, indicated by red boxes, are enlarged for better visualization.

Table 1. Averaged FSIM value for testing.

| $\lambda$ | Low Dose Image | Proposed Method | BM3D   |
|-----------|----------------|-----------------|--------|
| 0.001     | 0.7847         | 0.9161          | 0.8993 |
| 0.005     | 0.8956         | 0.9445          | 0.9353 |
| 0.01      | 0.9311         | 0.9578          | 0.9459 |

Table 2. Averaged IW-SSIM value for testing.

| $\lambda$ | Low Dose Image | Proposed Method | BM3D   |
|-----------|----------------|-----------------|--------|
| 0.001     | 0.7498         | 0.8890          | 0.8383 |
| 0.005     | 0.9074         | 0.9394          | 0.9310 |
| 0.01      | 0.9459         | 0.9584          | 0.9536 |

Acknowledgments
We acknowledge support from the Cancer Imaging Archive for permitting us to use the data sets of Curated Breast Imaging Subset of DDSM. 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 (U1932111,51975026) and National Natural Science Foundation of China (11574023).

References
[1] Siegel R L, Miller K D and Jemal A 2019 Cancer statistics, 2019 CA: a cancer journal for clinicians 69 7-34
[2] Bray F, Ferlay J, Soerjomataram I, Siegel R L, Torre L A and Jemal A 2018 Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries CA: a cancer journal for clinicians 68 394-424
[3] Islam S and Aziz S 2012 Mammography is the most effective method of breast cancer screening Mymensingh medical journal: MMJ 21 366-71
[4] de Gonzalez A B and Darby S 2004 Risk of cancer from diagnostic X-rays: estimates for the UK and 14 other countries The lancet 363 345-51
[5] Sperl J, Bequé D, Claus B, De Man B, Senzig B and Brokate M 2010 Computer-assisted scan protocol and reconstruction (CASPAR)—Reduction of image noise and patient dose IEEE transactions on medical imaging 29 724-32
[6] Noh H, Hong S and Han B 2015 Learning deconvolution network for semantic segmentation. In: Proceedings of the IEEE international conference on computer vision, pp 1520-8
[7] Krizhevsky A, Sutskever I and Hinton G E 2012 Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp 1097-105
[8] Lefkimmiatis S 2017 Non-local color image denoising with convolutional neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 3587-96
[9] 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 2190-2202
[10] Ronneberger O, Fischer P and Brox T 2015 U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention: Springer) pp 234-41
[11] Huang G, Liu Z, Van Der Maaten L and Weinberger K Q 2017 Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4700-8
[12] He K, Zhang X, Ren S and Sun J 2016 Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770-8
[13] Chan C L, Sullivan B J, Sahakian A V, Katseggelos A K, Swiryn S, Huetter 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
[14] 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 JOSA A 8 801-7
[15] Rebecca Sawyer Lee, Francisco Gimenez, Assaf Hoogi, Kanae Kawai Miyake, Mia Gorovoy & Daniel L. Rubin. (2017) A curated mammography data set for use in computer-aided detection and diagnosis research. Scientific Data  volume 4, Article number: 170177 DOI: 10.1038/sdata.2017.177
[16] Rebecca Sawyer Lee, Francisco Gimenez, Assaf Hoogi, Kanae Kawai Miyake, Mia Gorovoy & Daniel L. Rubin. (2017) A curated mammography data set for use in computer-aided detection and diagnosis research. Scientific Data  volume 4, Article number: 170177 DOI: 10.1038/sdata.2017.177
[17] Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M, Tarbox L, Prior F. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository, Journal of Digital Imaging, Volume 26, Number 6, December, 2013, pp 1045-1057. DOI: 10.1007/s10278-013-9622-7
[18] Babov K, Foi A, Katkovnik V and Egiazarian K 2007 Image denoising by sparse 3-D transform-domain collaborative filtering IEEE Transactions on image processing 16 2080-95
[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