2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)
Venice, Italy, April 8-11, 2019

USING VIRTUAL DIGITAL BREAST TOMOSYNTHESIS FOR DE-NOISING OF LOW-DOSE PROJECTION IMAGES

Pranjal Sahu[1], *Hailiang Huang*[2], *Wei Zhao*[2], *Hong Qin*[1]

[1]Computer Science Department, Stony Brook University
[2]Department of Radiology, Stony Brook School of Medicine

ABSTRACT

Digital Breast Tomosynthesis (DBT) provides a quasi-3D impression of the breast volume resulting in a better visualization of mass. However, one serious drawback of Tomosynthesis is that compared to Mammography, each projection gets lower x-ray dose resulting into higher quantum noise which seriously hampers the visibility of calcifications. To solve this problem we propose a Convolutional Neural Network model based on Adversarial loss. We train the deep network using synthetic data obtained from Virtual Clinical Trials. Unlike earlier works which tested model on phantoms only, we performed experiments on real samples obtained in clinical settings as well. Our approach shows encouraging results in de-noising the projections. De-noised projections show higher perceptual similarity with mammograms and superior signal-to-noise ratio. The reconstructed volume also enhances calcification visibility. Our work shows the viability of utilizing synthetic data for training the deep network for de-noising purposes.

Index Terms— Digital Breast Tomosynthesis, Low dose projection de-noising, Generative Adversarial Network.

1. INTRODUCTION

Digital Breast Tomosynthesis (DBT) is a relatively new modality for capturing the breast projections for cancer diagnosis purpose. In DBT multiple low-dose projections of breast volume are taken over an angular range which can be later reconstructed as a quasi-3D volume in the form of parallel thin image slices, as shown in Figure 1. Numerous studies have shown the advantages of using DBT over Mammography [1, 2]. Having multiple projections over an angular range reduces the overlapping tissues artifact resulting in better mass detection in case of DBT. However, a severe drawback of taking multiple projections is that each projection needs to be low-dose compared to the dose used in conventional mammography. This is necessary to make the total x-ray dose within the permissible limit. Lower x-ray dose results into higher quantum noise which severely impacts the calcification visibility in DBT volume [3, 4]. And therefore, a method which could reduce the quantum noise in DBT projections will be of great importance.

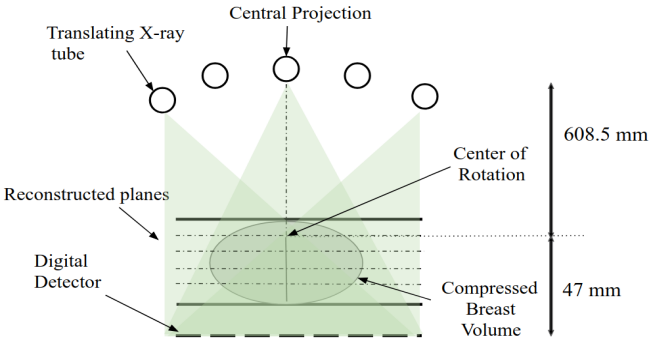

Fig. 1: Geometry of the Siemens Inspiration Digital Breast Tomosynthesis used in our experiments [14].

Classical de-noising methods such as BM3D [5] and KSVD [6] require considerable parameter tuning on case by case basis to work properly [7]. Moreover, in certain occasions their output result loses the details of important features such as calcifications [8]. Recently, multiple works using Convolutional Neural Network models have been proposed for de-noising CT projections [7, 8, 9, 10]. These models are typically trained by minimizing the mean squared error between the low and high dose projection pairs. Recently, the usage of Generative Adversarial Networks (GAN) [11] which use the adversarial loss have shown a huge improvement in de-noising tasks [12, 13]. In GAN based methods, the Generator learns to map the low-dose to high-dose images while the discriminator learns to distinguish between the de-noised image and real high-dose image.

In [13], authors compared three different approaches for training the de-noising GAN network. In their experiments, they found that a loss function which minimizes a combination of mean squared error and adversarial loss performs better in comparison to only minimizing the adversarial loss. However, one crucial requirement for using mean squared error is that the low-dose and high-dose projections should be perfectly aligned. Perfect alignment of projections for two different clinical trials is difficult because of patient motion

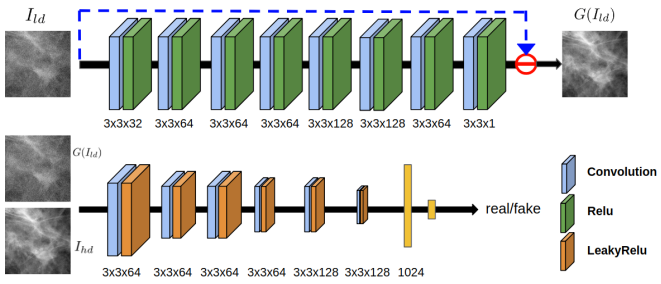

Fig. 2: Generator and Discriminator architectures used in our model. Generator minimizes both mean squared error and adversarial loss whereas discriminator minimizes the least squares objective between correct and predicted value. Input of Generator is 200x200 patch from DBT projection.

which is inevitable during the course of the trial which typically lasts for around 2 minutes. Moreover, obtaining a large amount of paired real data for training deep network is a problem in itself.

To circumvent such problems we propose a novel approach where the adversarial loss based network is trained entirely using the synthetic data. In the field of Breast Tomosynthesis, Virtual Clinical Trials (VCT) are utilized to study the impact of various parameters present in the acquisition systems for example system geometry, angular range, detector elements, x-ray dose etc. We generate perfectly aligned low and high dose projection pairs by varying the x-ray dose parameter in a Virtual Clinical Trial software called OpenVCT. The network is then trained using the synthetic projection pairs and tested on real projection pairs obtained from clinical settings. In our knowledge, this is first such attempt to study the feasibility of synthetic data for training the de-noising network in bio-medical domain. Our contributions in this paper are:-

1. Adversarial loss based de-noising of low-dose Digital Breast Tomosynthesis projections.
2. Evaluating training of neural network using synthetic data (phantoms) obtained from Virtual Clinical Trials.

2. NETWORK ARCHITECTURE

Earlier attempts in designing de-noising networks utilized only mean squared loss (mse). However, only using mean squared error results in smoothed output which lacks the detail present in high-dose images. Moreover, unrealistic looking smoothed images are not desirable for radiologists while reading. Therefore, to add realism and to avoid smoothing the output recent works utilized the concept of Generative Adversarial Networks popularly known as GANs. The main intuition behind GAN is to construct a generative model which learns the data distribution given some sample training data. It is trained by having a Discriminator (D) and a Generator network (G) being trained alternatively such that after the training is complete, the Generator (G) learns the distribution of the data. For our problem, the task of the Generator (G) is to take a low-dose image I_{ld} and to map it to its corresponding high-dose image I_{hd}. To achieve this objective the Generator minimizes loss l_G defined below which is a combination of mean squared loss and adversarial loss A:

$$l_G = \lambda \frac{\|G(I_{ld}) - I_{hd}\|^2}{S} + (1-\lambda)A(D(G(I_{ld})), 1), \quad (1)$$

where, $\lambda = 0.99$, S= pixels in an image I, $G(I_{ld})$ is the Generator's output for a low-dose image I_{ld} and $A(D(G(I_{ld})), 1)$ is the error between the Discriminator's (D) output for $G(I_{ld})$ and the label for high-dose images i.e. 1.

The task of Discriminator (D) is to accurately predict the label of the input image I. Therefore the formulation of the loss l_D for Discriminator (D) is:

$$l_D = A(D(I_{hd}), 1) + A(D(G(I_{ld})), 0), \quad (2)$$

where $A(I_{hd}, 1)$ is the error between Discriminator's (D) output for high-dose image I_{hd} and target label for high-dose images i.e. 1. Similarly, $A(D(G(I_{ld})), 0)$ is the error between Discriminator's output for $G(I_{ld})$ and the target label for low-dose images I_{ld} i.e. 0.

The main objective of this work is to demonstrate the feasibility of synthetic data for training de-noising networks and therefore we adopt very simple architectures for Generator and Discriminator networks as shown in Figure. 2. Filter size of 3x3 is used everywhere in both Generator and Discriminator. $Relu$ activation is used in all the layers in Generator (G), while $LeakyRelu$ activation with $\alpha = 0.2$ is used after all the layers in Discriminator (D). All the convolutions in Generator have a stride of 1 while in Discriminator stride of 2 is used in alternate convolution layers instead of max-pooling to reduce the feature dimension [15]. Fully connected layer with a dimension of 1024 is added after the last convolution layer in Discriminator followed by the final layer which outputs the prediction value. We do not use any activation for the final layer in Discriminator. The Generator uses skip connection, see Figure 2 and subtracts the output of the last convolution layer from the low-dose input image as shown below:

$$G(l_{ld}) = I_{ld} - \nu, \quad (3)$$

where ν is the output of the last convolution layer and can be interpreted as the quantum noise present in the projections [16].

3. EXPERIMENTS AND RESULTS

In this section we describe the experiment details in four sub-sections. First, we describe the synthetic data generation pipeline using OpenVCT [17] and the training of the

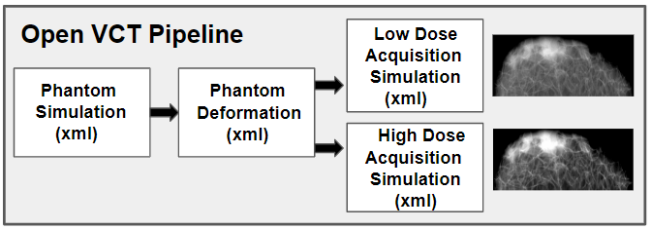

Fig. 3: Data generation pipeline for obtaining low and high dose projection pairs. Each stage takes parameters in the form of xml file.

de-noising GAN network in subsection 3.1. The de-noising of real projections using the GAN network is described in subsection 3.2 and finally the reconstruction of the volume using the de-noised projections is described in subsection 3.3.

3.1. Network Training

Training Data Generation Pipeline: We used OpenVCT for generating the synthetic dataset. OpenVCT provides a UI based simulation software to generate breast phantoms and takes the parameters of the system like system geometry, system parameters such as detector element, x-ray dose, phantom dimension etc. in the form of XML files, see Figure 3. Scripts were written to automate the phantom generation procedure and a total of 30 phantoms were generated with each phantom having 25 projections. We replicated the system geometry of the Siemens Inspiration DBT, see Figure 1, by providing the desired system geometry in XML file to the software. The detector element size is 0.085x0.085 mm^2. Random seed values are used to generate different distribution of tissues in the phantom and random breast density is selected for each phantom to obtain breast phantoms of both dense and non-dense types. Low and high dose projections for each phantom are simulated by using exposure equivalent to DBT and mammogram respectively. A sample of a synthetic low-dose and high-dose projection obtained from OpenVCT is shown in Figure 3. To only obtain the portion of projection containing the impression of breast we used otsu's thresholding. Mask of the non-breast region obtained by thresholding is then set to zero. While training we only consider image patches having breast occupancy greater than 90%.

GAN Training: In traditional GAN architectures adversarial loss A is taken as the binary cross-entropy, however, in [18] authors introduced Least Squares GAN where they showed that minimizing the mean squared error gives better results and hence we use the same. While sigmoid cross-entropy loss yields minimizing the KL divergence between the data distribution, the least squares loss minimizes Pearson χ^2 divergence [18]. Adam optimizer is used for both the Generator and Discriminator network with a learning rate of 0.0002 and $\beta_1 = 0.5$. Batch size of 32 samples is used. In each epoch crops of size 200x200 are sampled from random locations in a projection. The model is trained for a maximum of 20000 epochs as we found that discriminator's loss saturates after that. In each epoch the discriminator is trained 5 times while the generator is trained once.

3.2. De-noising of Real Projections

We test our de-noising network using images acquired in a pilot clinical study which investigates contrast-enhanced digital mammography and contrast-enhanced Digital Breast Tomosynthesis [19]. This study was approved by Institutional Review Board (IRB). Informed consents were obtained for research subjects prior to their participation in the study. Women who have suspicious breast lesions classified as BI-RADS grade 4 or 5 and scheduled for surgical biopsy are recruited in this study. A prototype system for dual energy contrast-enhanced breast imaging modified from Siemens Mammomat Inspiration DBT system is used for image acquisition. This system can generate high-energy (HE) and low-energy (LE) x-rays for imaging iodinated contrast agent. It acquires 25 projections in 25 seconds spanning an angular range of 50 degrees in one DBT scan, see Figure 1. Prior to imaging, the research subjects are injected with iodinated contrast agent. Approximately 2 minutes later, the breast with suspicious lesions is positioned and compressed. HE and LE mammograms are acquired first, followed by HE and LE DBT scans. The image acquisition is completed in one breast compression in approximately 3.5 minutes.

DBT projections have a resolution of the order of 2000 px x1000 px, therefore we de-noise the projections (LE) by feeding patches of size 200px x 200px into the generator network (G). To compare the model's performance we compare the de-noised projections (LE) with the mammograms (LE). While taking mammograms only one central projection is taken and hence higher-dose is used. MSE ignores the perceptual information present in the image and hence has limitations while quantifying the performance of the de-noising algorithm. Therefore, we use a recently introduced metric called HaarPSI (Haar Wavelet-Based Perceptual Similarity Index) [20] which takes the perceptual similarity also into account. HaarPSI metric determines the similarity of two images in the range [0,1], with higher value meaning greater similarity. The formulation of the metric can be found in [20]. Results of the de-noising network on real projections is shown in Figure 4 where the HaarPSI and PSNR (Peak Signal to Noise Ratio) metrics are shown below image.

3.3. Volume Reconstruction

The breast volume was reconstructed from projections using OS-SART (Ordered Subsets Simultaneous Algebraic Reconstruction Technique) method [21]. We used a subset size of 5 and reconstruction is done for 5 iterations. Comparison of

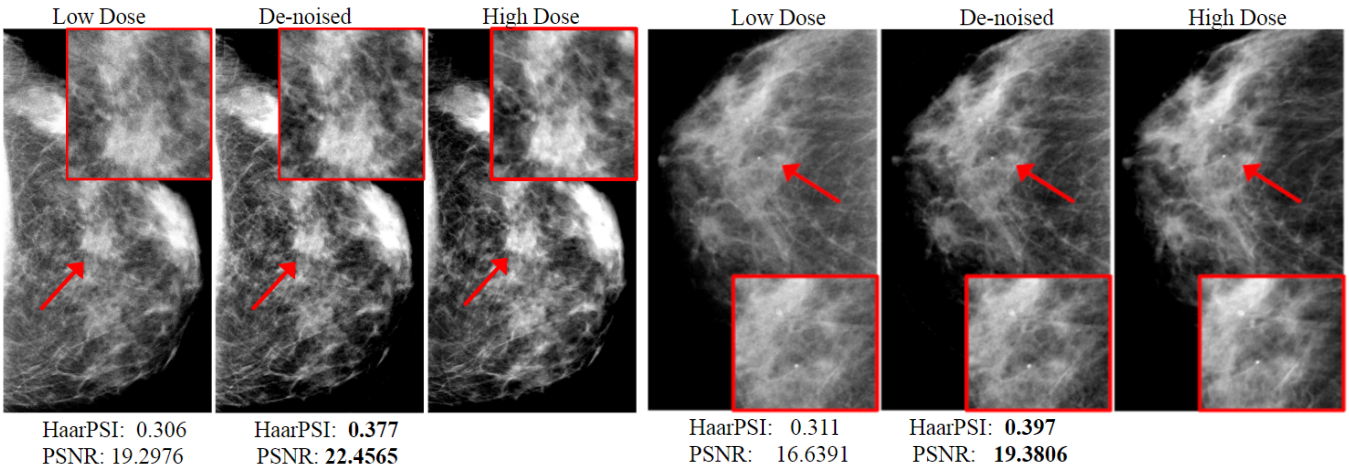

Fig. 4: Results of our GAN based de-noising model tested on real DBT central projections and their comparison with the Mammogram which acts as high-dose. HaarPSI and PSNR values is displayed below for each image patch.

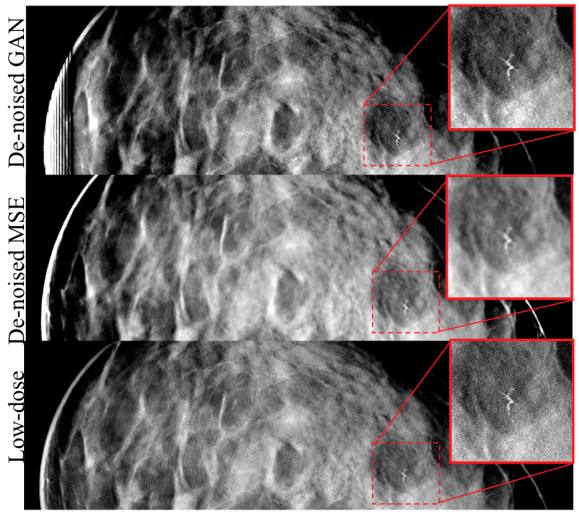

Fig. 5: Comparison between volume reconstructed using de-noised projections and original low-dose projections. Visibility of calcification is enhanced in case of adversarial loss while mean squared loss function gives blurry result.

volume reconstructed using de-noised projections and original low-dose projections is shown in Figure 5. De-noising done using mean squared error is also compared with the adversarial loss based de-noising in Figure 5. Blurring of calcifications is less in case of GAN based de-noising compared to mean squared error de-noising. Performance comparison of traditional algorithms such as K-SVD and BM3D is performed on six clinical cases with the metrics calculated on the crop extracted from central region. The summary of results is shown in Table 1. The GAN based de-noising method outperforms the K-SVD and BM3D methods based on HaarPSI metric. Comparison on a sample crop is shown in Figure 6.

Table 1: Comparison of KSVD, BM3D, MSE loss vs GAN based de-noising method for projection patches.

Metric	MSE	GAN	KSVD	BM3D
HaarPSI (mean)	0.367	**0.383**	0.295	0.2851
PSNR (mean)	19.90	18.91	**20.16**	18.44

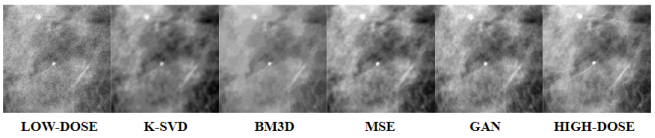

Fig. 6: Qualitative comparison of different de-noising algorithms on a real projection crop.

4. CONCLUSION AND FUTURE WORK

Our work demonstrates that synthetic data has good potential to be used for training the de-noising network specially when obtaining paired training data is difficult, for example in the case of Digital Breast Tomosynthesis. Our future work will explore the usage of GAN methodologies for decreasing out-of-plane artifacts in reconstructed volume caused due to limited angle acquisition which can improve lesion detection in dense breast. Another promising research direction is to explore the impact of angular range in reconstruction by varying the angular range in VCT. We also want to explore alternative digital breast phantoms which can further improve the performance of our approach [22].

5. ACKNOWLEDGMENT

We gratefully acknowledge the support of Bruno Barufaldi for helping us setup OpenVCT which is available at https://sourceforge.net/projects/openvct/.

References

[1] TM Svahn, Chakraborty, et al., "Breast tomosynthesis and digital mammography: a comparison of diagnostic accuracy," *The British journal of radiology*, vol. 85, no. 1019, pp. e1074–e1082, 2012.

[2] Julianne S Greenberg et al., "Clinical performance metrics of 3d digital breast tomosynthesis compared with 2d digital mammography for breast cancer screening in community practice," *American Journal of Roentgenology*, vol. 203, no. 3, pp. 687–693, 2014.

[3] M Lee Spangler, Margarita L Zuley, et al., "Detection and classification of calcifications on digital breast tomosynthesis and 2d digital mammography: a comparison," *American Journal of Roentgenology*, vol. 196, no. 2, pp. 320–324, 2011.

[4] Steven P Poplack et al., "Digital breast tomosynthesis: initial experience in 98 women with abnormal digital screening mammography," *American Journal of Roentgenology*, vol. 189, no. 3, pp. 616–623, 2007.

[5] Kostadin Dabov et al., "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Transactions on image processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

[6] Michal Aharon et al., "K-svd: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on signal processing*, vol. 54, no. 11, pp. 4311, 2006.

[7] Hu Chen, Yi Zhang, et al., "Low-dose ct with a residual encoder-decoder convolutional neural network," *IEEE transactions on medical imaging*, vol. 36, no. 12, pp. 2524–2535, 2017.

[8] Junchi Liu, Amin Zarshenas, Qadir, et al., "Radiation dose reduction in digital breast tomosynthesis (dbt) by means of deep-learning-based supervised image processing," in *Medical Imaging 2018: Image Processing*. International Society for Optics and Photonics, 2018, vol. 10574, p. 105740F.

[9] Hu Chen, Yi Zhang, et al., "Low-dose ct via convolutional neural network," *Biomedical optics express*, vol. 8, no. 2, pp. 679–694, 2017.

[10] Kenji Suzuki et al., "Neural network convolution (nnc) for converting ultra-low-dose to virtual high-dose ct images," in *International Workshop on Machine Learning in Medical Imaging*. Springer, 2017, pp. 334–343.

[11] Ian Goodfellow et al., "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[12] Qingsong Yang, Pingkun Yan, Zhang, et al., "Low dose ct image denoising using a generative adversarial network with wasserstein distance and perceptual loss," *IEEE transactions on medical imaging*, 2018.

[13] Jelmer M Wolterink, Tim Leiner, Max A Viergever, and Ivana Išgum, "Generative adversarial networks for noise reduction in low-dose ct," *IEEE transactions on medical imaging*, vol. 36, no. 12, pp. 2536–2545, 2017.

[14] Ioannis Sechopoulos, "A review of breast tomosynthesis. part i. the image acquisition process," *Medical physics*, vol. 40, no. 1, 2013.

[15] Alec Radford, Luke Metz, and Soumith Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv preprint arXiv:1511.06434*, 2015.

[16] Kai Zhang, Wangmeng Zuo, et al., "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, 2017.

[17] Bruno Barufaldi et al., "Openvct: a gpu-accelerated virtual clinical trial pipeline for mammography and digital breast tomosynthesis," in *Medical Imaging 2018: Physics of Medical Imaging*. International Society for Optics and Photonics, 2018, vol. 10573, p. 1057358.

[18] Xudong Mao, Li, et al., "Least squares generative adversarial networks," in *Computer Vision (ICCV), 2017 IEEE International Conference on*. IEEE, 2017, pp. 2813–2821.

[19] Hailiang Huang, David Scaduto, et al., "Comparison of contrast-enhanced digital mammography and contrast-enhanced digital breast tomosynthesis for lesion assessment," *Journal of Medical Imaging*, (in press).

[20] Rafael Reisenhofer et al., "A haar wavelet-based perceptual similarity index for image quality assessment," *Signal Processing: Image Communication*, vol. 61, pp. 33–43, 2018.

[21] Yihuan Lu, Boyu Peng, David A Scaduto, Wei Zhao, and Gene Gindi, "Application of the ordered-subsets transmission reconstruction algorithm to contrast-enhanced dual-energy digital breast tomosynthesis," in *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2014 IEEE*. IEEE, 2014, pp. 1–5.

[22] Stephen J Glick and Lynda C Ikejimba, "Advances in digital and physical anthropomorphic breast phantoms for x-ray imaging," *Medical physics*, vol. 45, no. 10, pp. e870–e885, 2018.