# Breast Tomosynthesis Reconstruction Using Artificial Neural Networks with Deep Learning

Davi Duarte de Paula*
São Paulo State University (Unesp)
Institute of Geosciences and Exact Sciences,
Av. 24A, 1515,
Rio Claro/SP, Brazil, 13506-900
Email: davi_duarte@unesp.br

Denis H. P. Salvadro
São Paulo State University (Unesp)
Institute of Geosciences and Exact Sciences,
Av. 24A, 1515,
Rio Claro/SP, Brazil, 13506-900
Email: denis.salvadeo@unesp.br

***Abstract*—The Filtered Backprojection (FBP) algorithm for Computed Tomography (CT) reconstruction can be mapped entire in an Artificial Neural Network (ANN), with the back-projection (BP) operation simulated analytically in a layer and the Ram-Lak filter simulated as a convolutional layer. Thus, this work adapt the BP layer for DBT reconstruction, making possible the use of FBP simulated as a ANN to reconstruct DBT images. For evaluation, Structural Similarity Index Measure (SSIM) and Peak Signal-to-Noise Ratio (PSNR) metrics were calculated to measure the improvement of the images made by the ANN, regarding a dataset containing 100 virtual breast phantoms to perform the experiments. We shown that making the Ram-Lak layer trainable, the reconstructed image can be improved in terms of noise reduction. And, considering an additional post-filtering step performed by Denoising Convolutional Neural Network (DnCNN), it shown comparable and superior results than a state-of-the-art DBT reconstruction method, averaging 37.644 dB and 0.869 values of PSNR and SSIM, respectively. Finally, this study enables additional proposals of ANN with Deep Learning models for DBT reconstruction and denoising.**

## I. Introduction

Digital Breast Tomosynthesis (DBT) is a radiographic imaging technique, with acquisition of projections at limited angles using a reduced dose of radiation. It aims to reconstruct tomographic slices from the inside of the breast, enabling the early diagnosis of possible lesions and, consequently, increasing the probability of a patient's cure. However, due to the fact that DBT uses low doses of radiation, the generated image contains more noise than digital mammography. Although the quality of the exam is directly related to the dose used, it is expected that the radiation dose used in the exam will be as low as possible, but still with sufficient quality so that the diagnosis can be made, according to the *As principle. Low As Reasonably Achievable* (ALARA). One of the important steps to pursue the ALARA principle is tomographic reconstruction, which consists of a software that generates the slices inside the breast from a set of acquired 2D DBT projections. On the other hand, Machine Learning techniques, especially neural networks with deep learning, which have recently considerably evolved the state of the art in several Computer Vision and Image Processing problems, have adequate characteristics to be applied also in the reconstruction stage. Thus, this work investigated a basic architecture of artificial neural network with deep learning that is capable of reconstructing DBT images, especially focused on noise reduction. Still, considering an additional filtering step using the *Denoising Convolutional Neural Networks* (DnCNN) method after the reconstruction, the proposed method was superior to a state-of-the-art DBT reconstruction approach in terms of SSIM (*Structural Similarity Index Measure*) and PSNR (*Peak Signal-to-Noise Ratio*) using data from virtual DBT specimens, reaching values of 37.644 db in PSNR and 0.869 in SSIM.

* This work is related to the author's M.Sc. dissertation.

## II. Related Works

In this section the methods for DBT Reconstruction will be introduced. Basically, there are three categories of methods based on machine learning for tomographic reconstruction [1]: a) Learned Iterative Scheme (LIS); b) Learned Regularizer (LR) and; c) Learned Denoiser (LD).

The LIS approach is the basic tomographic reconstruction scheme, they are methods simulating backprojection and forward projection processes. The LD approach comprise methods focused on denoising problem of tomographic data, applying some approach in some step of reconstruction process (projection or reconstructed image). Therefore, this approach uses a conventional reconstruction algorithm, like FPB, acting like a pre- or post- processing to improve the final image. Finally, the methods in LR approach aims to learn a regularization term from a set of tomographic data, using some reconstructing equation.

The work of Ref. 2 is in LIS approach, because this implements the FBP algorithm in an ANN architecture using fan-beam CT projections. The main idea is to increase the interpretability and facilitate combine others layers (like denoising layers) and/or approaches to improve the final image quality. Thus, this implementation has no trainable parameters. An extension for cone-beam was published. The authors proposed another layer to move their method from fan-beam CT to cone-beam CT 3. In this extension, the method uses 2D projections as input data, a layer for Parker weights [4] to correct 2D projections, and, finally, a layer for 2D cosine weighting.

The work made by Ref. 5 fall in the LIS approach, because the proposed method uses Convolutional Neural Network (CNN) and Reinforcement Learning (RL) to find the best parameters of reconstructions made by the Total Variation (TV) method, using fan-beam computed tomography. The TV uses several parameters that can result in non-optimal results if manually chosen. However, using this approach, this method overpass this manually limitation using a ANN to automatically choose the TV parameter.

In the LR approach, Ref. 6 uses a Convolutional Neural Network (CNN) with Residual Learning [7] and deconvolution [8] as a regularization function, avoiding the need to manually select one. Unlike most approaches that uses Machine Learning in tomographic image reconstruction, Ref. [6] do not use a dataset containing tomographic images to train its architecture, but the COCO [9] dataset, containing approximately 80,000 images of photographic objects, such as animals, people, cars, etc. However, the use of this dataset is justified because the architecture proposed by Ref. [6] aims to eliminate the noise present in the images, preserving the characteristic structures of tomographic images. Their experiments revealed that the resulting ANN reduces noise, blur and preserve the structures of tomographic images.

The authors of 10 propose a LR approach that simulates the Fields of Experts (FoE) [11] regularization function as an ANN. They replaced the main terms of the FoE equation with an ANN that uses convolutional layers. They used an iterative approach to reconstruction, such that, in each iteration, the reconstruction of the previous stage is added to the estimated reconstruction of the current stage, making the proposed ANN categorized by the authors as a recurrent CNN. In their tests, using CT images, the proposed approach obtained better results compared to traditional methods of tomographic reconstruction.

In turn, for the LD approach, it can be cited the work of Ref. [12] that proposes to use Adversarial Neural Networks (GAN) [13] to simulate our perception of good CT images, which was determined using Wasserstein distance and perceptual similarity, defining a loss function. The authors argue that other Deep Learning-based filtering approaches that use the Mean Squared Error (MSE) cost function improve the image in terms of PSNR, but can hide important structural details.

## III. Methodology

This section describes the operator written by Ref. 3 (Section III-A) and the modifications made by this work to propose the new operator to reconstruct DBT images (Section III-B).

### *A. Fundamental step: modeling backprojection as an ANN layer*

As described by Ref. [14], analytical reconstruction of CT images using parallel beam can be described as

$$f(x,y) = \int_0^{\pi} p(u,\theta) * h(u)|_{u=x*cos(\theta)+y*sin(\theta)} d\theta, \quad (1)$$

where $f(x,y)$ is a voxel in the reconstructed image, $p(u,\theta)$ is the projection value detected at the position $u$ in an acquisition angle $\theta$, $*$ is the operator of convolution and $h(u)$ is the discrete Ramp Filter [14]. In this way, a set of projections $p(.)$ can be reconstructed into a volumetric image $f(.)$.

However, it is possible to use an optimization method to perform the reconstruction of tomographic projections, resulting in an iterative method, in which the solution of the problem (i.e., the reconstructed image) is provided by finding the minimum point of a given function $L(f)$. This function can be defined by a subtraction of the reconstructed image projections from the original projections acquired from the imaged body, and may be written as [3]

$$L(f) = \frac{1}{2}||Af - p||_2^2, \quad (2)$$

where $A \in R^{M*P}$ is the forward projection transform (Radon Transform) [15], $f \in R^N$ denotes the expected reconstruction, $M$ is the number of pixels in the projection (size detector), $P$ is the number of projections and $N$ is the number of voxels in the reconstructed image.

Is possible to follow the inverse direction of gradient of Equation 2, in relation to the variable $f$, and find the desired reconstruction, mathematically defined by [3]

$$\frac{\partial L(f)}{\partial f} = 0, \quad (3)$$

whose tomographic reconstruction problem can be obtained by the solution of the system in Equation 2, leads to the following equation to obtain $f$ [3]

$$f = A^T(AA^T)^{-1}p. \quad (4)$$

Alternatively, the reconstruction of tomographic images can be written according to the following operators. Since the $A$ operator represents the Radon Transform, the reconstruction can be written by [3]

$$f = A_{bp}^T C p, \quad (5)$$

where $A_{bp}^T$ is the back projection operator, and $C$ is a convolution operator to filter the projections $p$ with the discrete ramp filter $h(u)$. So, it can be observed that $A_{bp}^T C$ corresponds to the term $(AA^T)^{-1}$ of Equation 4 [3].

Following the definition presented by [3], a completely connected layer of any ANN can be represented by the following equation [3]

$$x^{i+1} = Wx^i, \quad (6)$$

where $x^{i+1}$ is the layer output, $x^i$ the layer input and $W$ is an array containing the layer weights adjusted in the training step. In the case of CT image reconstruction, the $A_{bp}^T$ operator could be mapped to a fully connected layer. Thus, the $W$ array would have $N * M * P$ parameters, which could make its implementation unfeasible due to the high memory cost [3].

However, Ref. 3 implemented a specific layer for this operator as the analytical backprojection. Although this is a Fully Connected Layer, this has no trainable parameters, making its implementation feasible. This new layer aims to equate the value of $W$ with the computation of $A_{bp}^T$.

Furthermore, the Ref. 3 also proposed other layers to act joint with the backprojection to simulate the entire FBP method. This layers are before the backprojection layer and adapts the method for the cone-beam input data, in addition to perform filtering or other operations directly in the projection domain. So the error that will be passed back can be written as [3]

$$e_{l-1} = W^T e_l, \tag{7}$$

$$e_{l-1} = (A_{pb}^T)^T e_l = A_{pb} e_l, \tag{8}$$

where $e_{1-1}$ is the error fetched from the backprojection layer and $e_l$ is the error resulting from the underlying layer. This allows the feedfoward and backfoward step to be calculated efficiently.

### B. ANN-DL model to DBT from CT

Just as done by Ref. 3, is possible use the Equation 5 and add more operator to deal with the two-dimensional CBCT projections. Then, the reconstruction of cone-beam CT (CBCT) can be written as [3],

$$f = A_{cb}^T C_{2D} W_{red_{2D}} W_{cos_{2D}} p, \tag{9}$$

where $A_{cb}^T$ is the backprojection operator for CBCT projections, $C_{2D}$ is a convolution layer that simulates the discrete ramp filter for two-dimensional projections, $W_{red_{2D}}$ is a layer that implements Parker [4] weighting to handle redundant projections, $W_{cos_{2D}}$ is a two-dimensional cosine weighting of the projection data $p$, also implemented as a layer. The Figure 1 illustrates the ANN architecture given by Equation 9.

In this sense, it is possible to use the same concepts to enable DBT images to be reconstructed as well. Starting from Equation 5, it is possible to reach the operator equation that describes the reconstruction of DBT images as

$$f = A_{dbt}^T C_{2D} W_{red_{2D}} W_{cos_{2D}} p, \tag{10}$$

where $A_{dbt}^T$ is the backprojection operator for DBT projections. Figure 1 illustrates an ANN architecture given by Equation 10.

The difference in the $A_{cb}^T$ and $A_{dbt}^T$ operator is how projection pixels are summed to result the voxel of the final reconstructed image. While the former considers the projections acquired through a cone beam ray, the latter considers a truncated cone beam ray, as well as the specific geometry of the DBT system. To avoid negative values in the final reconstructed image, the ReLU activation function was added at the end of the architecture.

Implementation of FBP for DBT normally not use the Parker weights [16], [17], but it was decided to keep the operator $W_{red_{2D}}$ in Equation 10 as a trainable layer initialized as Parker weights to verify in the experiments if the training step can adapt the initial values for a layer more suitable to DBT reconstruction.

However, the basic proposed ANN architecture for DBT reconstruction do not consider the operator $W_{red_{2D}}$, being described by the following equation

$$f = A_{dbt}^T C_{2D} W_{cos_{2D}} p, \tag{11}$$

in which is visually presented in Figure 2.

It is important to emphasize that the ANN architectures implemented from Equations 10 and 11 have only one layer with adjustable parameters, the $W_{red_{2D}}$ layer, initialized with Parker weights, and $C_{2D}$ layer, initialized with ramp filter, respectively. These parameters will be modified by the training step to achieve better results in noise reduction of the final reconstruction.

Lastly, in the next Section, experiments will be carried out to verify the efficiency of these architectures in noise reduction.

## IV. Experiments and Results

Following the 2 ANN-DL architectures presented in last section (Parker and Ram-Lak Architecture), this section will evaluate the ability to reduce noise in DBT images. In subsection IV-A, the dataset is described. In Subsection IV-B, the experimental evaluation is detailed. In the Subsection IV-C, the results are presented.

### A. Dataset

To evaluate this work, we used a software named OpenVCT to generate 100 simulated phantoms. This software allows the generation of virtual phantoms very similar to real breast images, including internal injuries, structures, mechanical compression, and specifics details of the DBT scanners geometry. For the evaluation phase, this work used the Hologic geometry, that is, for each phantom, 15 projections were made by a hypothetical radiation source that moved within an angular range of $-7.5$ to $+7.5$ above the breast. The distance from detector to radiation source is 700mm. In order to compare the capability of model to remove noise from images, the OpenVCT can generate noisy and noise-free projections from a virtual phantom. In case of noisy DBT projections, a Poisson noise model was simulated.

Therefore, for each of these 100 phantoms, noisy and noisy free version of the projections are available. Each volume must contain 64 slices with dimensions $2048 \times 1792$ voxels and 1mm of thickness.

### B. Experimental Setup

The dataset composed of the 100 virtual phantons were divided into three subsets: (i) a training set, containing 79 samples; (ii) a validation set, containing 1 sample and; (iii) a test set, containing 20 samples. Then, the proposed ANN-DL architectures were trained using the training set and the validation set, the latter for early stop. Each sample in the dataset is composed of (a) a set of noisy projections of the phantom and (b) expected DBT volume, i. e., DBT slices

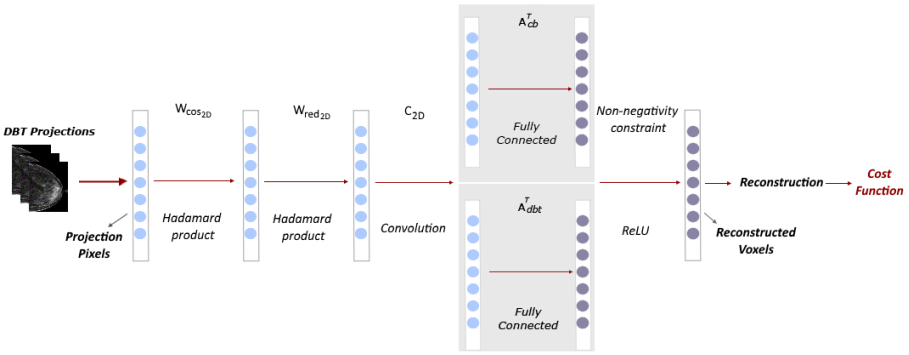



Fig. 1: This architecture can use the layer $A^T_{cb}$ or $A^T_{dbt}$ (gray blocks). If the former layer is used, the final architecture represents the Equation 9, proposed by Ref. 3 for CBCT. If the latter layer is used, the final architecture represents the Equation 10, proposed by this work.

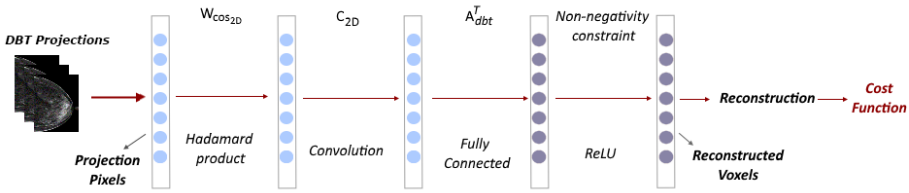



Fig. 2: ANN architecture for DBT reconstruction without using Ram-Lak weights. Adapted from Ref. 3.

reconstructed from noise-free projections, which is obtained from FBP method. The architectures proposed in Section III would be able to map noisy reconstructions to their noiseless versions.

For proposal evaluation, the trained networks received as input a set of noisy projections of a certain phantom in the test set, generating the volume. This volume was compared with the volume reconstructed by FBP from correspondent noise-free projections of that phantom in terms of quantitative and visual evaluation. For quantitative evaluate were calculated the Peak Signal-to-Noise Ratio (PSNR) [18] and Structural Similarity Index Measure (SSIM) [18]. It is important to inform that the figures in visual analysis were submitted to a linear adjustment for better viewing, but the SSIM and PSNR measures were calculated without this adjustment.

In addition, the results of the proposed methods were compared with: (a) FBP reconstruction; b) a combination of ANN-DL with DnCNN [19] for denoising slices; c) FBP reconstruction from filtered projections using Non Local Means (NLM) applied on Anscombe Domain [20]; d) FBP reconstruction from filtered projections using Block Matching and 3D Collaborative Filtering (BM3D) [20], state-of-the-art method for a complete DBT reconstruction process.

For running the experiments, a machine equipped with a 12GB Tesla K80 GPU, 6 GPUs and 54GB of memory was used. However, due to the high computational cost, only the center slice (32th slice) was used in the experiments.

Finally, the proposal was implemented by using the Tensorflow framework [21] and adapting to DBT the source code available for CT [3].

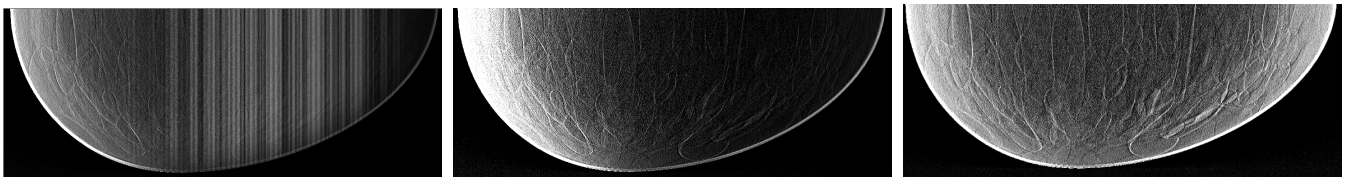
### C. Results

Following the experimental methodology detailed in Section IV-A and IV-B, results were generated to evaluate the Parker Architecture (Equation 10), detailed in Section IV-C1, and Ram-Lak Architecture (Equation 11), detailed in Section IV-C2. Section IV-C3 analyzes and compares the influence of the proposed architectures when applying a post-reconstruction and pre-reconstruction denoising method.

*1) Parker Architecture:* According to Equation 10, in this architecture, the only trainable layer is $W_{red_{2D}}$ (Figure 1). The training was carried out with the learning rate ($LA$) $LA = 1$, using the training set and the validation set (to early stop). The training time was approximately 12 hours.

After the training, the samples of the test set were submitted to the trained Parker Architecture and the values of SSIM and PSNR were calculated, whose mean and standard deviation are shown in Table I, in the row Parker Architecture. The row FBP shows the results for the FBP reconstruction method. The values were calculated on segments of $300 \times 550$px extracted from interior of breasts, to prevent the black background interfering with the results.

Evaluating the Table I, it is possible to notice the worsens in the PSNR and SSIM values, which can be confirmed by visually evaluating a sample of the test set as displayed in Figure 3a. In this Figure, it is observed that there are predominance of artifacts (vertical lines) in the right half of the sample. This shows that the $W_{red_{2D}}$ layer was not able to learn the parameters needed to properly reconstruct the DBT projections. That can makes sense, since the Parker weights were developed for image reconstructions with projections acquired over a wide angular range. In this way, the results show that changes in Parker weights in the training stage were not able to result in a suitable reconstruction.



(a) Noisy segment (b) Phantom filtered by Ram-Lak Architecture + $L_1$. (c) Phantom filtered by Ram-Lak Architecture + $L_2$.

Fig. 3: Reconstructions using Parker [4] weighting.

However, it is also possible to note that the use of the original Parker [4] weighting causes an incorrect reconstruction of the DBT images as displayed in Figure 3b. It can be noted a incorrect light distribution in the phantom, in comparison with the reconstruction made without the Parker Weighting as displayed in Figure 3c.

| | SSIM | PSNR |
|---|---|---|
| **FBP** | 0.396 $\pm$ 0.021 | 26.931 $\pm$ 0.18 |
| **Parker Architecture** | 0.202 $\pm$ 0.009 | 21.162 $\pm$ 0.273 |
| **Ram-Lak Architecture + L1** | **0.456 $\pm$ 0.0155** | **28.054 $\pm$ 0.131** |
| **Ram-Lak Architecture + L2** | 0.446 $\pm$ 0.0184 | 27.827 $\pm$ 0.186 |

TABLE I: Experimental results for reconstruction methods.

*2) Ram-Lak Architecture:* According to Equation 10, in this neural architecture, only the $C_{2D}$ layer is trainable (Figure 2). The training step was carried out with the learning rate $LA = 1 * 10^{-9}$. The training time was approximately 1 hour.

After the training step, the samples of the test set were submitted to the Ram-Lak Architecture and the values of SSIM and PSNR were calculated, which are displayed in Table I. The FBP row shows the results for the FBP reconstruction method, the Ram-Lak Architecture + $L_1$ row shows the results of the trained architecture using the $L_1$ cost function, and the Ram-Lak Architecture + $L_2$ row shows the results of the trained architecture using the $L_2$ cost function. The values were calculated on segments of $300 \times 550$px extracted from interior of breasts, to prevent the black background interfering with the results.

When analyzing the Table I, it can be noted that Ram-Lak Architecture + $L_1$ presented a better result than FBP and Ram-Lak Architecture + $L_2$. For a qualitative evaluation, Figure 4a shows a noisy segment from a test set phantom reconstructed by the FBP method, Figure 4b shows the same segment reconstructed by Ram-Lak Architecture + $L_1$, Figure 4c shows the segment reconstructed by Ram-Lak Architecture + $L_2$ and Figure 4 shows the ideal segment (no noise). It is possible to observe that Ram-Lak architecture achieved a improvement in term of noise, but still far from the ideal result.

*3) Denoising:* In the previous Subsection, it can be seen that Ram-Lak Architecture + $L_1$ presented better results than the traditional FBP reconstruction. However, the result was still inferior to the ideal image. In this way, this Section evaluate if the image reconstructed via Ram-Lak Architecture + $L_1$ shows improvement when filtered by a denoising approach in comparison to the same image reconstructed by the FBP and filtered by the same denoising method.

| | SSIM | PSNR |
|---|---|---|
| **BM3D+FBP** | 0.396 $\pm$ 0.021 | 37.102 $\pm$ 0.334 |
| **NLM+FBP** | 0.636 $\pm$ 0.009 | 31.91 $\pm$ 0.160 |
| **FBP+DnCNN** | 0.864 $\pm$ 0.010 | 37.412 $\pm$ 0.623 |
| **Ram-Lak Architecture + L1 + DnCNN** | **0.869 $\pm$ 0.010** | **37.644 $\pm$ 0.521** |

TABLE II: Experimental results for denoising methods.

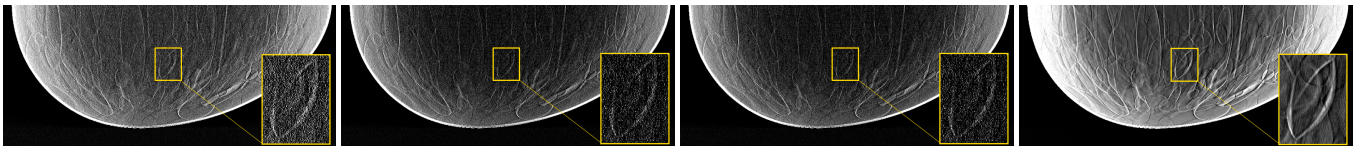
The chosen denoising method was Denoising Convolutional Neural Networks (DnCNN) [19], which uses convolutional layers to map an image on its initial noisy state to a final state without noise. The method was trained separately, after the image reconstruction, as post-denoising method, as seen in Figure 5. For DnCNN training, the simulated dataset was divided into a training set, containing 80 samples, and a test set, containing 20 samples. Due to computational costs, a region of interest (RoI) of $180 \times 180$px from the inside of the breast was used for training and testing.

The Block Matching and 3D Collaborative Filtering (BM3D) [22] and Nonlocal Means (NLM) [23] denoising methods were used for comparison to DnCNN. They were applied to the 15 projections before reconstruction via FBP, namely BM3D+FBP and NLM+FBP here, respectively. The Anscombe transform was performed before applying the method, and its inverse was applied after denoising, as done by Ref. [24] and illustrated in Figure 6. The use of Anscombe transform is necessary to stabilize the Poisson noise of the projections, such that projections in Anscombe Domain are approximately corrupted by a Gaussian noise, with mean 0 and unitary variance [25].

In Table II, we find the mean and standard deviation of the SSIM and FBP values for the $180 \times 180$px segments extracted from the samples in the test set. The rows BM3D+FBP and NLM+FBP show the result for pre-denoising using, respectively, the BM3D and NLM denoise algorithms. The rows FBP+DnCNN and Ram-Lak Architecture + $L_1$ + DnCNN show the results for post-denoising using, respectively, the DnCNN applied after reconstruction via FBP and the DnCNN applied after reconstruction via Ram-Lak Architecture + $L_1$.

In Table II is shown that the Ram-Lak Architecture + $L_1$ + DnCNN overpass the BM3D+FBP and NLM+FBP in terms of both SSIM and PSNR. In comparison with FBP+DnCNN, it achieved a little improvement.

For visual inspection, Figure 7 shows reconstruction results of an image segment of a phantom in the test set: (a) noisy (FBP result), (b) BM3D+FBP result, (c) NLM+FBP, (d) FBP+DnCNN, (e) Ram-Lak architecture + DnCNN, (f) ideal, noise-free segment. It can be observed that denois-



(a) Noisy segment (b) Phantom filtered by Ram-Lak Architecture + $L_1$. (c) Phantom filtered by Ram-Lak Architecture + $L_2$. (d) Ideal phantom

Fig. 4: Comparison of a phantom from test set with a highlighted segment.

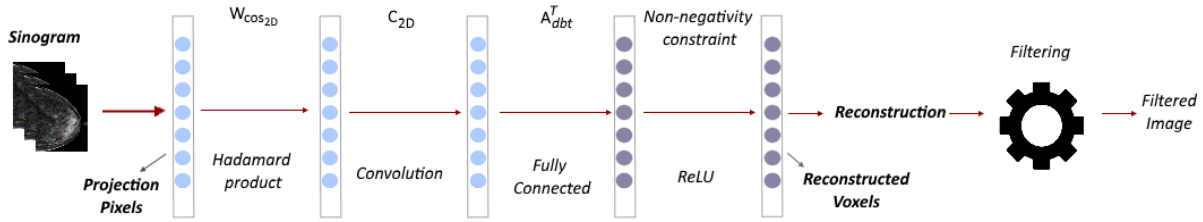



Fig. 5: DBT reconstruction followed by a denoiser.

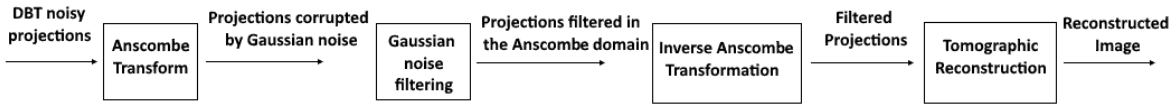



Fig. 6: Anscombe Transformation scheme for filtering in projections.

ing performed by BM3D+FBP reduced the noise and preserved the details, while the denoising by the NLM+FBP still achieved a high noise level. In turn, denoising performed by FBP+DnCNN reduced the significantly noise more than BM3D+FBP, but can be observed that some details were blurred in comparison with BM3D+FBP, even though the values of SSIM and PSNR were better (Table II). It is also observed that the denoising performed after the reconstruction by the Ram-Lak Architecture + $L_1$ + DnCNN presents an increase in noise reduction in comparison with FBP+DnCNN.

## V. Conclusions and Future Works

In this paper, two artificial neural network models were developed and evaluated for DBT reconstruction problem. The first model named, Parker Architecture (Section IV-C1), were not able to improve the image of DBT, unlike what happens with CT images [3]. The second model, named Ram-Lak Architecture (Section IV-C2), achieved the best results in the experiments in terms of PSNR and SSIM.

This enabled the additional analysis present in Section IV-C3, which compares several additional filtering approaches. In that Section, it was seen that the use of Ram-Lak Architecture achieve the best results when used together with a denoising approach (DnCNN applied on reconstructed images), being quantitatively superior to a state-of-the-art DBT reconstruction approach (FBP reconstruction from filtered projections by BM3D) [20], mainly in terms of PSNR. However, the Ram-Lak Architecture achieved subtle inferior results when qualitatively analyzed. It may have been caused by the ANN-DL hyperparameters and the limited dataset used. So, additional experiments for a further evaluation of the ANN hyperparameters can be explored in a future work.

Finally, possible extensions of this paper includes also: a) use DBT clinical data for evaluation; b) analysis of Parker and Ram-Lak trainable layers before and after training; c) join and evaluation of Ram-Lak architecture with DnCNN in a only one ANN-DL for DBT reconstruction; d) comparison with other traditional and machine learning DBT reconstruction methods; e) use of a bigger dataset; f) add a filtering step in a pre-reconstruction stage of Ram-Lak architecture + DnCNN, performing a double denoising [20] in DBT.

## Acknowledgments

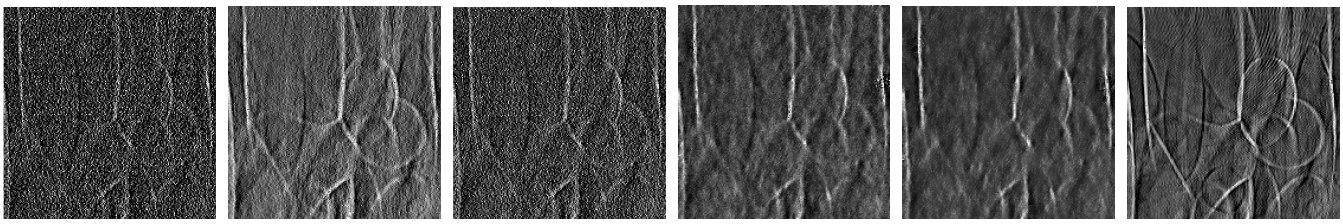
This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) - Finance Code 001 (provided a scholarship for Davi Duarte de Paula). We also would like to thank Andrew Douglas Arnold Maidment and Bruno Barufaldi from University of Pennsylvania for provide the DBT virtual phantoms used in these experiments.

## References

[1] J. Adler and O. Öktem, "Learned primal-dual reconstruction," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1322–1332, 2018.

[2] T. Würfl, F. C. Ghesu, V. Christlein, and A. Maier, "Deep learning computed tomography," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2016, pp. 432–440.

[3] T. Würfl, M. Hoffmann, V. Christlein, K. Breininger, Y. Huang, M. Unberath, and A. K. Maier, "Deep learning computed tomography: Learning projection-domain weights from image domain in limited angle problems," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1454–1463, 2018.



(a) Noisy segment (b) BM3D+FBP (c) NLM+FBP (d) FBP+DnCNN (e) Ram-Lak+DnCNN (f) Ideal (noise-free)

Fig. 7: Comparison results regarding an additional step for denoising pre- or post-reconstruction.


[4] D. L. Parker, "Optimal short scan convolution reconstruction for fan beam ct," *Medical physics*, vol. 9, no. 2, pp. 254–257, 1982.

[5] C. Shen, Y. Gonzalez, L. Chen, S. B. Jiang, and X. Jia, "Intelligent parameter tuning in optimization-based iterative ct reconstruction via deep reinforcement learning," *arXiv preprint arXiv:1711.00414*, 2017.

[6] B. Chen, K. Xiang, Z. Gong, J. Wang, and S. Tan, "Statistical iterative cbct reconstruction based on neural network," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1511–1521, 2018.

[7] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[8] M. D. Zeiler, G. W. Taylor, and R. Fergus, "Adaptive deconvolutional networks for mid and high level feature learning," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 2018–2025.

[9] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *European conference on computer vision*. Springer, 2014, pp. 740–755.

[10] H. Chen, Y. Zhang, Y. Chen, J. Zhang, W. Zhang, H. Sun, Y. Lv, P. Liao, J. Zhou, and G. Wang, "Learn: Learned experts' assessment-based reconstruction network for sparse-data ct," *IEEE transactions on medical imaging*, 2018.

[11] S. Roth and M. J. Black, "Fields of experts," *International Journal of Computer Vision*, vol. 82, no. 2, p. 205, 2009.

[12] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, Y. Zhang, L. Sun, and G. Wang, "Low-dose ct image denoising using a generative adversarial network with wasserstein distance and perceptual loss," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1348–1357, 2018.

[13] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[14] A. C. Kak and M. Slaney, *Principles of computerized tomographic imaging*. IEEE press New York, 1988.

[15] B. Z. Des Plantes, "Eine neue methode zur differenzierung in der roentgenographie (planigraphie)," *Acta Radiologica*, no. 2, pp. 182–192, 1932.

[16] R. B. Vimieiro, L. R. Borges, and M. A. Vieira, "Open-source reconstruction toolbox for digital breast tomosynthesis," in *XXVI Brazilian Congress on Biomedical Engineering*. Springer, 2019, pp. 349–354.

[17] B. Barufaldi, D. Higginbotham, P. R. Bakic, and A. D. Maidment, "Openvct: a gpu-accelerated virtual clinical trial pipeline for mammography and digital breast tomosynthesis," in *Medical Imaging 2018: Physics of Medical Imaging*, vol. 10573. International Society for Optics and Photonics, 2018, p. 1057358.

[18] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? a new look at signal fidelity measures," *IEEE signal processing magazine*, vol. 26, no. 1, pp. 98–117, 2009.

[19] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "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.

[20] D. H. Salvadeo, R. B. Vimieiro, M. A. Vieira, and A. D. Maidment, "Bayesian reconstruction for digital breast tomosynthesis using a non-local gaussian markov random field a priori model," in *Medical Imaging 2019: Physics of Medical Imaging*, vol. 10948. International Society for Optics and Photonics, 2019, p. 109485C.

[21] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard *et al.*, "Tensorflow: a system for large-scale machine learning." in *OSDI*, vol. 16, 2016, pp. 265–283.

[22] L. R. Borges, P. R. Bakic, A. Foi, A. D. Maidment, and M. A. Vieira, "Pipeline for effective denoising of digital mammography and digital breast tomosynthesis," in *Medical Imaging 2017: Physics of Medical Imaging*, vol. 10132. International Society for Optics and Photonics, 2017, p. 1013206.

[23] M. A. Vieira, H. C. de Oliveira, P. F. Nunes, L. R. Borges, P. R. Bakic, B. Barufaldi, R. J. Acciavatti, and A. D. Maidment, "Feasibility study of dose reduction in digital breast tomosynthesis using non-local denoising algorithms," in *Medical Imaging 2015: Physics of Medical Imaging*, vol. 9412. International Society for Optics and Photonics, 2015, p. 94122C.

[24] D. C. Scarparo, D. H. P. Salvadeo, D. C. G. Pedronette, B. Barufaldi, and A. D. Maidment, "Evaluation of denoising digital breast tomosynthesis data in both projection and image domains and a study of noise model on digital breast tomosynthesis image domain," *Journal of Medical Imaging*, vol. 6, no. 3, p. 031410, 2019.

[25] F. J. Anscombe, "The transformation of poisson, binomial and negative-binomial data," *Biometrika*, vol. 35, no. 3/4, pp. 246–254, 1948.