Deep learning reconstruction of digital breast tomosynthesis images for accurate breast density and patient-specific radiation dose estimation

Jonas Teuwen [a,b,1], Nikita Moriakov [a,b,1], Christian Fedon [a], Marco Caballo [a], Ingrid Reiser [e], Pedrag Bakic [f,g], Eloy García [c], Oliver Diaz [d], Koen Michielsen [a], Ioannis Sechopoulos [a,h,*]

[a] Department of Medical Imaging, Radboud University Medical Center, the Netherlands
[b] Department of Radiation Oncology, Netherlands Cancer Institute, the Netherlands
[c] Vall d'Hebron Institute of Oncology, VHIO, Spain
[d] Department of Mathematics and Computer Science, University of Barcelona, Spain
[e] Department of Radiology, The University of Chicago, USA
[f] Department of Radiology, University of Pennsylvania, USA
[g] Department of Translational Medicine, Lund University, Sweden
[h] Dutch Expert Centre for Screening (LRCB), the Netherlands

ARTICLE INFO

ABSTRACT

The two-dimensional nature of mammography makes estimation of the overall breast density challenging, and estimation of the true patient-specific radiation dose impossible. Digital breast tomosynthesis (DBT), a pseudo-3D technique, is now commonly used in breast cancer screening and diagnostics. Still, the severely limited 3rd dimension information in DBT has not been used, until now, to estimate the true breast density or the patient-specific dose. This study proposes a reconstruction algorithm for DBT based on deep learning specifically optimized for these tasks. The algorithm, which we name DBToR, is based on unrolling a proximal-dual optimization method. The proximal operators are replaced with convolutional neural networks and prior knowledge is included in the model. This extends previous work on a deep learning-based reconstruction model by providing both the primal and the dual blocks with breast thickness information, which is available in DBT. Training and testing of the model were performed using virtual patient phantoms from two different sources. Reconstruction performance, and accuracy in estimation of breast density and radiation dose, were estimated, showing high accuracy (density $< \pm 3\%$; dose $< \pm 20\%$) without bias, significantly improving on the current state-of-the-art. This work also lays the groundwork for developing a deep learning-based reconstruction algorithm for the task of image interpretation by radiologists.

1. Introduction

Breast cancer screening with mammography has proven effective in reducing breast cancer-related mortality (Plevritis et al., 2018). However, since mammography is a 2D imaging modality, it results in the projection of the internal tissue of the breast onto a single plane, yielding tissue superposition. This results in a lower sensitivity and specificity, especially with dense breasts.

1.1. Digital breast tomosynthesis

Digital breast tomosynthesis (DBT) has been introduced over the last two decades to decrease the impact of tissue superposition in mammography, by providing limited depth information, resulting in improved detection and diagnosis performance (Zackrisson et al., 2018; Zuley et al., 2013). Hence, DBT is rapidly replacing digital mammography as the primary X-ray technique for breast imaging (Niklason et al., 1997). DBT imaging consists of acquiring several low-dose planar X-ray projections over a limited angle. These projections are then used to reconstruct a pseudo-3D volume, albeit with very limited vertical spatial resolution. A secondary benefit of the introduction of DBT for widespread use for screening of-

---

* Corresponding author at: Department of Medical Imaging, Radboud University Medical Center, the Netherlands.
E-mail address: ioannis.sechopoulos@radboudumc.nl (I. Sechopoulos).
[1] Equal contribution.

fers the opportunity, for the first time, to obtain accurate estimates of the breast density and, subsequently, of the fibroglandular dose.

1.2. Breast density

In mammography, the lack of information on the true tissue distribution in the vertical (X-ray source to detector) direction limits the ability to accurately estimate the breast density (i.e., the proportion of the breast that consists of fibroglandular tissue). Up to now, estimates of breast density from mammography necessitated the use of models of the image acquisition process and assumptions and simplifications regarding tissue distribution. Although these models have been characterized for consistency and precision, their accuracy is unknown. Estimating breast density is of intense interest because it is an important factor for both masking and breast cancer risk (McCormack, 2006). As a result of its significance, breast density reporting is now mandated by law in several states in the USA. This makes it especially important to use methods that provide objective and quantitative breast density estimates. The feasibility of accurate breast density estimation from DBT images has been recognized before. However, given the very poor spatial resolution in the vertical direction obtained with all current DBT reconstruction algorithms, the localization of the fibroglandular tissue in DBT images has proven extremely challenging. Previous efforts to achieve this have not resulted in improvements over the model-based dose estimates obtained from 2D imaging (Geeraert, 2014), or have involved algorithms that require a lot of manual input, making them challenging to implement clinically (Förnvik et al., 2018).

1.3. Radiation dose

In breast imaging, the dose of interest is only that to the fibroglandular tissue, since this is the tissue most at risk of developing breast cancer (Dance and Sechopoulos, 2016), and to determine this dose it is necessary to know its vertical location within the breast. Therefore, currently, all breast dosimetry is based on dose estimates using a model breast, which does not reflect the dose deposited in the actual patient breast. It has been shown that the use of a model breast results in an average overestimation of the true patient breast of 30% and that this error can be as high as 120%, if not more (Dance et al., 2005; Sechopoulos et al., 2012; Hernandez et al., 2015). Even if a more accurate, unbiased model of the average breast is developed, an effort that is currently ongoing (Arana Peña et al., 2020), the over 100% error in patient-specific dose estimates using a population-wide model will not be ameliorated. To obtain accurate radiation dose estimates, the actual amount and position of the fibroglandular tissue in the individual patient's breast needs to be considered. However, due to its complete lack of information on the vertical position of tissues, this is impossible to achieve with mammography. Only with the introduction of digital breast tomosynthesis, is it now feasible to gather this knowledge for each imaged breast.

Therefore, in this work, we propose a new approach to DBT reconstruction, based on our earlier work (Moriakov et al., 2019), using deep learning methods, that results in the 3D representation of the imaged breast optimized for estimation of the true distribution of the fibroglandular tissue. This in turn allows for accurate estimation of both the breast density and the radiation dose imparted on it, improving significantly upon the state-of-the-art.

1.4. DBT reconstruction

Both FBP-based and iterative methods are in clinical use to reconstruct breast tomosynthesis images, and both methods result in severely limited resolution in the direction perpendicular to the detector plane (Sechopoulos et al., 2012; Vedantham et al., 2015). Despite this, most research is focused on improving image quality in the high-resolution planes parallel to the detector plane since only this direction is examined by radiologists. In particular, the choice of the reconstruction algorithm and regularization parameters can greatly influence the reconstruction quality, as was shown in previous work (Rodriguez-Ruiz et al., 2017; Michielsen et al., 2016). Approaches that try to improve image quality in the vertical direction typically require strong prior knowledge to sufficiently constrain the inverse problem. In industrial settings this is feasible if the scanned object contains a small number of known materials by applying a discrete tomography method such as poly-DART (Six et al., 2019). A different approach recently proposed by Zhang et al. (2021) obtained promising results under the assumption that the true total variation in each direction of the scanned object is known.

1.5. Deep learning for reconstruction

A recent and very promising development in medical imaging reconstruction uses methods that rely on deep learning (Arridge et al., 2019). The goal of this paper is to show the potential of such methods for the problem of DBT reconstruction, using the quantitative estimation of breast density and radiation dose as the target application. Our method combines a deep learning network with an inductive bias given by the forward and backward models (and therefore considering part of the physics processes involved in image acquisition). This is in contrast to other established methods that postprocess initial reconstructions with a deep learning network to improve image quality (Kang et al., 2017; Jin et al., 2017). To build this algorithm, we extended our previous results on DBT reconstruction (Moriakov et al., 2019) and investigate a data-driven reconstruction algorithm called Deep Breast Tomographic Reconstruction (DBToR), where the reconstruction is computed from projection data with a deep neural network. To train the model, and to test the performance of DBToR for the tasks of breast density and radiation dose estimation, we used dedicated breast CT images, where the tissue distribution is known, and use a finite-element model to simulate the change in the tissue distribution when under compression in DBT. In addition to these images, we also evaluate the model on simulated breast phantoms.

1.6. Our contribution

In this work, we show: (i) the feasibility of reconstruction of DBT using deep learning, having developed a novel deep learning-based model, to reconstruct DBT; (ii) that the resulting reconstructions have greatly improved vertical resolution compared to state-of-the-art analytical and iterative reconstruction methods; (iii) that the proposed reconstruction method is able to provide accurate breast density estimates; and (iv) that the dense tissue location information results in accurate patient-specific dosimetric estimates.

This paper is organized as follows: Section 2 sets the context, describing DBT reconstruction as an inverse problem, and presents the architecture and model. Section 3 describes the dataset and methods used for the dosimetric evaluation. We follow in Section 4 with the results obtained with the DBToR algorithm, and providing several comparisons against existing algorithms. Finally, we conclude in Section 5.

2. Deep learning-based reconstruction

Before we describe our model in more detail, we give a brief overview on how to formulate the reconstruction problem as an inverse problem.

2.1. Inverse problems and DBT reconstruction

DBT reconstruction can be formulated as an inverse problem. In a mathematically simplified setting this means that given an object to be imaged $x \in X$ and measured projection data $y \in Y$, we have

$$y = \mathcal{P}x + \eta \tag{1}$$

where $\mathcal{P} : X \to Y$ is the forward, or projection, operator, that models how the object x gives rise to the projection $\mathcal{P}x$ in the absence of noise, and η is a Y-valued random variable modeling the noise component of the measurements. The forward model used is given by

$$y_i(\mathbf{x}) = b_i \exp\left(-\sum_j l_{ij} x_j\right), \tag{2}$$

where y_i is the projection data, b_i the number of X-ray photons emitted towards detector pixel i and l_{ij} the length of the intersection between voxel j and the line between the source and detector pixel i. The linear attenuation in voxel j is denoted by x_j.

The noise vector η is typically assumed to be additive Gaussian, which is a good approximation for high photon counts, which are common in transmission imaging. In our work, we do not require an explicit noise model for inversion, and assume a more realistic noise model given by Poisson noise in Section 3 for the simulations of the projection data.

The goal of reconstruction is to retrieve the object x from measured (and noise-corrupted) projection data y. Inversion of the operator $\mathcal{P}$ is an ill-posed problem, and we will consider the estimation of x from a Bayesian perspective and consider X and Y as probability spaces, therefore identifying in this work the spaces X and Y with $L^2(\mathbb{R}^2)$. The goal for the Bayes estimator is to minimize the expected loss over all estimators $\hat{x} : Y \to X$ that give an estimate of x given measurements y. That is,

$$\hat{x}_{\text{Bayes}} := \operatorname*{argmin}_{\hat{x}:Y\to X} \mathbf{E}_{x\sim X}(\hat{x}(y) - x)^2. \tag{3}$$

In our approach, we will obtain a neural network approximation to the Bayes estimator $\hat{x}_{\text{Bayes}}$, where argmin in Eq. (3) is taken over the estimators given by a family of neural networks and where optimization is performed using minibatch stochastic gradient optimization, with expectation $\mathbf{E}_{x\sim X}$ being approximated by sampling $x \sim X$. For a complete overview on statistical inverse problems, we refer to the book by Kaipio and Somersalo (2005).

2.2. DBT neural network

The DBToR algorithm is a data-driven algorithm, which extends the Learned Primal-Dual (LPD) reconstruction algorithm (Adler and Öktem, 2018), by incorporating additional prior information about the geometry in the form of the thickness measurement of the breast under compression in DBT.[2] The LPD algorithm itself is an unrolled iterative scheme, based on the proximal primal-dual hybrid gradient method, where proximal operators are replaced by neural networks. The algorithm is trained to reconstruct the images directly from projection data. The DBToR neural network consists of several 'reconstruction blocks', which take in projection data, together with information on the thickness of the breast under compression as the initial input, perform a forward and a backward pass by taking projections and back-projections, and use a convolutional neural network to produce an intermediate reconstruction result, which is then improved further by each successive reconstruction block. The architecture and the training algorithm are illustrated in Fig. 1 and Algorithm respectively. In all our experiments we set the number of primal blocks N_{prim} and dual

blocks N_{dual} to 10. The blocks, denoted by $\Gamma_{\theta_i^d}$ and $\Lambda_{\theta_i^p}$ are all ResNet-type blocks consisting of three convolutional layers with kernel size 3×3 followed by a PreLU layer (with slope initialized as 0.25) and 64, 64 and 5 filters respectively. As the operators $\mathcal{P}$ and $\mathcal{P}^*$ we select the forward and backward operators respectively. The backward operator $\mathcal{P}^*$ with $\mathcal{P}$ as in (1) is the (matrix) adjoint of the forward operator $\mathcal{P}$. In our model, these are implemented using ASTRA (van Aarle et al., 2016).

At test time, the algorithm takes the input projection data y and the breast thickness information to compute the reconstruction using the function `compute_reconstruction`.

In what follows, we denote the training set of objects by $\mathcal{D}_{\text{train}}$. For an object x, we let $y = \text{sinogram}_x$ be the corresponding projection data for clarity. As is common, we assume that the training data $\mathcal{D}_{\text{train}}$ is a representative sample from the domain of DBT images that we want to reconstruct.

The neural network is trained in a supervised fashion as follows. We repeatedly sample an image $x \sim \mathcal{D}_{\text{train}}$ and the corresponding input projection data $y = \text{sinogram}_x$ from the training dataset. The corresponding thickness mask is denoted by $m = \text{thickness\_mask}_x$ and is represented by a rectangular mask with the same width as the detector and where the height is given by the measured breast thickness during compression. These measurements are provided by the DBT system, and are available both at training and at test time.

To find the parameters θ of the neural network `compute_reconstruction`, we train the network with the L^2-loss $l_\theta := \|x - z\|_2^2$. The parameters θ are updated using the Adam optimizer with a cosine annealing learning rate schedule (Loshchilov and Hutter, 2019), i.e. the learning rate in step t was

$$\eta_t = \frac{\eta_0}{2}\left(1 + \cos\left(\pi \frac{t}{N_{\text{iter}}}\right)\right)$$

starting at a learning rate η_0 of 10^{-4}. For the other Adam parameters, we choose the default parameters of $\beta = (0.9, 0.999)$, $\epsilon = 10^{-8}$ and weight decay 0. The total number of iterations N_{iter} and batch size differ per dataset, and are detailed in the Methods section. Before feeding the data into the network, the input projection data is log-transformed, and scaled such that the standard deviation and mean over the training set is 1.

Algorithm 1 Pseudocode of the DBToR reconstruction and training algorithm.

1: **procedure** COMPUTE_RECONSTRUCTION(m, g)
2:     $f_0 \leftarrow 0 \in X^{N^{prim}}$ ▷ Initialize primal vector
3:     $h_0 \leftarrow 0 \in U^{N^{dual}}$ ▷ Initialize dual vector
4:     **for** $i \leftarrow 0, N$ **do**
5:         $h_i \leftarrow \Gamma_{\theta_i^d}(h_{i-1}, \mathcal{P}(f_{i-1}^{(2)}), g, \mathcal{P}(m)))$
6:         $f_i \leftarrow \Lambda_{\theta_i^p}(f_{i-1}, \mathcal{P}^*(h_i^{(1)}), m)$
7:     **end for**
8:     **return** $f_I^{(1)}$
9: **end procedure**
10: **for** $j \leftarrow 0, N_{\text{iter}} - 1$ **do**
11:     $x \sim \mathcal{D}_{\text{train}}$ ▷ Sample train data
12:     $y \leftarrow \text{sinograms}_x$ ▷ Sample sinograms
13:     $m \leftarrow \text{thickness\_mask}_x$ ▷ Create masks
14:     $z \leftarrow \text{compute\_reconstruction}(m, y)$
15:     $\text{loss} \leftarrow \|z - x\|_2^2$
16:     change parameters $\theta_i^p, \theta_i^d, i = 1, \ldots, N$ to reduce loss
17: **end for**

---

[2] This information is measured by the DBT device and stored in the DICOM image header.
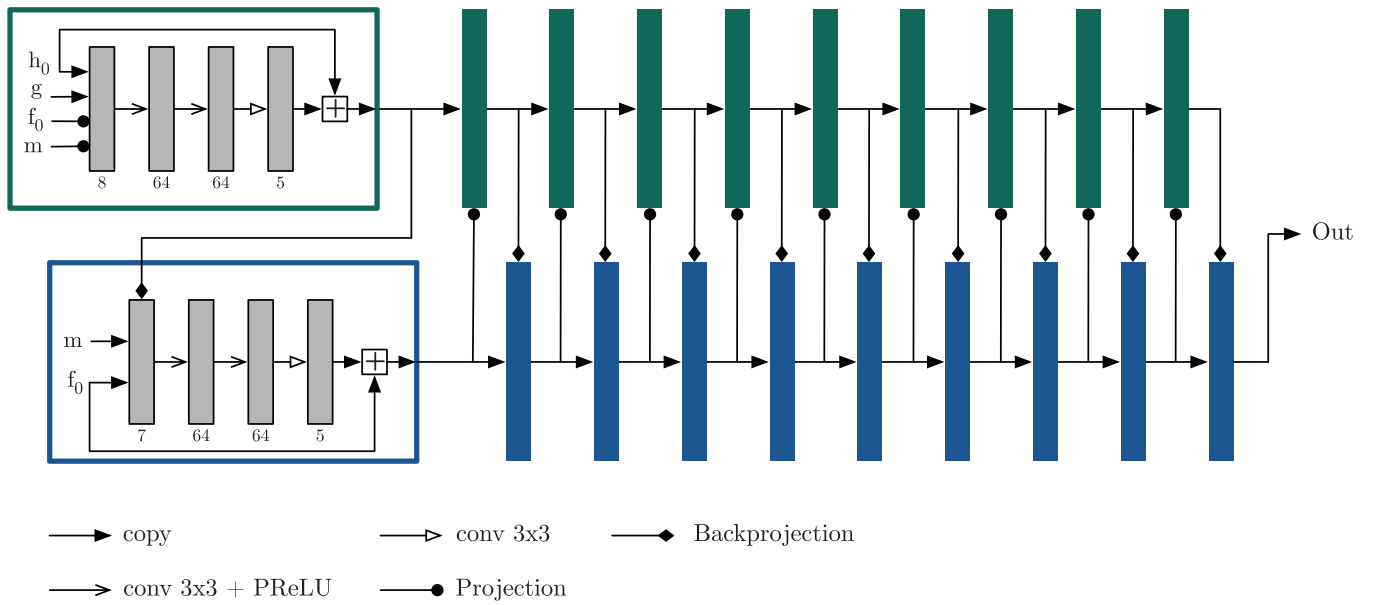
Fig. 1. Network architecture of DBToR. Dual blocks (green) are on the upper row and primal blocks (blue) are in the bottom row. The blocks have the same architecture, elaborated in the first blocks. m: the breast thickness mask, h_0: initial dual vector, f_0: initial primal vector, g: sinogram data, Out: final reconstruction. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.3. Reconstructed image classification

The final step of the reconstruction for estimation of breast density and radiation dose involved in image acquisition, is the classification of the reconstructed image into skin, adipose, and fibroglandular tissue voxels. This is required because rather than relying on voxel attenuation values, which can be quite similar for different types of tissue, we need correct tissue labels to compute the breast density and the dose absorbed by the fibroglandular tissue. Given its high contrast, the skin layer was segmented through a fast seeded region-growing algorithm (Adams and Bischof, 1994), which grows the segmented region starting from a subset of seeds (corresponding to the voxels located on the outer edge of the skin layer) by subsequently including voxels whose intensity was higher than or equal to the mean seed intensity value. For fibroglandular tissue classification, we extend our previous work on breast CT classification (Caballo et al., 2018), and in the first step remove the skin temporarily from the image, with the resulting representation undergoing a well-established, automatic thresholding method based on fuzzy c-means clustering (Bezdek, 1981), an algorithm generally adopted in the case of images with low noise content, accompanied by a non-negligible degree of blurring, as is the case for our images. Briefly, voxels are iteratively assigned to a given class (adipose or fibroglandular tissue) in an unsupervised fashion, with the iteration stopping criterion aiming at maximizing the distance between the average voxel values of the two classes. As opposed to traditional cluster analysis, this method allows for a degree of fuzzy overlap between the classes over each iteration, which helps classify the boundary voxels in each subsequent iteration. The fuzzy partition term (Bezdek, 1981) was experimentally tuned to a value of 2.0.

3. Materials and methods

To train and evaluate the algorithm, we created two datasets of 3D breast phantoms from which we extracted the coronal slices and their corresponding DBT projections. The first dataset consists of virtual 3D breast phantoms generated using a stochastic model, while the second is based on patient dedicated BCT (Breast Computed Tomography) images. DBT projections of these phantoms were simulated using deterministic simulation methods, with the posterior addition of Poisson noise. The use of virtual phantoms not only provided training data, but also allowed for assessing the accuracy of the density and dose estimates, since ground truth is known.

Each voxel of these phantoms was indexed with a label denoting the corresponding tissue type: skin, adipose tissue, fibroglandular tissue, and Cooper's ligaments. The elemental compositions of these materials were obtained from the work of Hammerstein et al. (1979), except for the composition of Cooper's ligaments, which was assumed to be identical to that of fibroglandular tissue. Linear attenuation coefficients at 20 keV, a typical average energy of the spectra in clinical DBT imaging, were calculated for each material using the software from Boone and Chavez (1996). The resulting linear attenuation coefficients were 0.512 cm^{-1} for adipose tissue, 0.798 cm^{-1} for fibroglandular tissue (and Cooper's ligaments), and 0.854 cm^{-1} for skin.

This process was done for the virtual phantom dataset and the patient BCT dataset.

3.1. Virtual phantom

We extracted 41,499 coronal slices from 50 breast phantoms generated using the method of Lau et al. (2012). This method generates breast phantoms in two steps: first, the breast structure is simulated on a coarse scale by generating large compartments of adipose tissue Zhang et al. (2008); Bakic et al. (2011). Second, finer detail for fibroglandular tissue is added subsequently in the form of power-law noise (Reiser and Nishikawa, 2010). The resulting images have dimensions 1000×300 and a resolution of 200 µm. An example of the simulated phantoms is shown in Fig. 2.

These 41,499 coronal slices were used for training and validating the algorithm. Each breast phantom was either included in a training or validation fold completely or not at all, in order to prevent data contamination and bias.

```
 1: procedure COMPUTE_RECONSTRUCTION(m, g)
 2:     f_0 ← 0 ∈ X^{N_prim}              ▷ Initialize primal vector
 3:     h_0 ← 0 ∈ U^{N_dual}             ▷ Initialize dual vector
 4:     for i ← 0, N do
 5:         h_i ← Γ_{θ_i^d}(h_{i-1}, 𝒫(f_{i-1}^{(2)}), g, 𝒫(m)))
 6:         f_i ← Λ_{θ_i^p}(f_{i-1}, 𝒫*(h_i^{(1)}), m)
 7:     end for
 8:     return f_I^{(1)}
 9: end procedure
10: for j ← 0, N_iter − 1 do
11:     if j is divisible by freq then
12:         x ∼ 𝒟_train                   ▷ Sample train data
13:         y ← sinograms_x               ▷ Sample sinograms
14:         m ← thickness_mask_x          ▷ Create masks
15:     end if
16:     z ← compute_reconstruction(m, y)
17:     loss ← ‖z − x‖_2^2
18:     change parameters θ_i^p, θ_i^d, i = 1, …, N to reduce loss
19: end for
```

Fig. 2. 2D coronal breast phantom containing skin (darkest gray), adipose tissue (dark gray), fibroglandular tissue (light gray), and Cooper's ligaments (black).
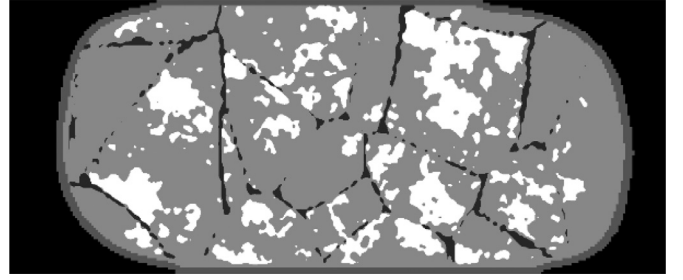

Fig. 3. (a) Coronal slice of a breast CT image, (b) the same image classified into skin (white), adipose (dark gray) and fibroglandular (light gray) tissue voxels, and (c) the classified deformed image with the technique described in Section 3.2.

3.2. Patient dedicated breast CT phantoms

Patient dedicated breast CT images were acquired for an unrelated, ethical-board approved patient study evaluating this imaging technology. The images were released for other research purposes after anonymization. In order to compute the density and the accumulated dose to the fibroglandular tissue, the patient breast CT images were automatically classified into four categories (air, skin, adipose and fibroglandular tissue) using a previously developed algorithm for BCT image classification (Caballo et al., 2018) (see Section 2.3). The classified breasts then underwent simulated mechanical deformation as previously described (Fedon et al., 2019; García et al., 2020). Briefly, the breasts were converted into a finite element (FE) biomechanical model using the package iso2mesh (v.1.8; Matlab v.13a). A large number of 4-node tetrahedral elements, between 100k and 500k, were used in order to minimize the numerical error during the FE analysis (del Palomar et al., 2008). Nearly incompressible (Poisson ratio equal to 0.495), homogeneous and isotropic Neo-Hookean material models for each tissue were used to describe their mechanical behavior. The Youngs modulus for fibroglandular, adipose and skin tissue were set to 4.46 kPa, 15.10 kPa, and 60.00 kPa, respectively (Wellman, 1999). The mechanical compression was then simulated using the open-source package NiftySim (v.2.3.1; University College London) (Johnsen et al., 2015), which uses a Total Explicity Dynamic Lagrangian approach to solve the mechanical FE problem (Miller et al., 2007).

Each breast model was compressed to the thickness recorded in the corresponding DICOM header of the cranio-caudal DBT view of that breast, which was acquired, for clinical purposes, during the same visit as the acquisition of the BCT image.

The total phantom population includes compressed breast thicknesses from 3.0 cm to 5.6 cm and chest wall-to-nipple distances from 5.8 cm to 18.0 cm with an isotropic voxel size of 0.2 mm × 0.2 mm × 0.2 mm, which is more than sufficient for dosimetric applications (Fedon et al., 2019).

Using this method, of which an example is given in Fig. 3, we obtained a total of 28891 deformed BCT slices extracted from 91 different patient breasts. Similar to the virtual phantoms the resulting images have dimensions of 1000×300 pixels and a resolution of 200 μm. Given that the number of deformed BCT slices was substantially lower than that of virtual phantom slices, we pretrained the model with the latter, and then fine-tuned the model using the BCT slices from 46 patient BCT images. The other 45 patient BCT image phantoms were used for testing the reconstruction performance of the model and the accuracy of density and dosimetry estimates. Each patient breast was either completely included or excluded when selecting slices for fine-tuning and testing the model in order to prevent data contamination and bias.

3.3. Projection data

Limited angle fan-beam projections were simulated for all coronal phantom slices using a geometry with the center of rotation placed at the center of the phantom as seen in Fig. 4. The X-ray source was placed 65cm above the center of rotation, and the source-detector distance was 70 cm. A total of 25 equally spaced projections between $-24°$ and $24°$ were generated, with the detector rotating with the X-ray source. The detector was a perfect photon counting system (100% efficiency) consisting of 1280 elements with a resolution of 0.2mm. The forward model (2) was used for the simulations.

For the virtual phantom data, we generated a series of data sets at 3 noise levels from the noiseless simulated projections. This was accomplished by setting photon count $b_i = 1000 \cdot \sqrt{2}^N$ with $N = \{4, 8, 12\}$. For each noise level, a single Poisson noise realization was generated. For the deformed BCT phantoms, only photon counts corresponding to $N = 8$ were used.

Baseline reconstructions were generated for both noiseless and noisy data using 100 iterations of the Maximum Likelihood for Transmission (MLTR) algorithm (Nuyts et al., 1998), using the compressed breast thickness to set the size of the reconstruction volume and with no additional regularization.

3.4. Density computation

Breast density by mass, also called glandularity (G), was computed as follows:

$$G = \frac{N_g \rho_g}{N_g \rho_g + N_a \rho_a} \tag{4}$$

where N_g and N_a are the number of voxels classified as fibroglandular and adipose tissue in the full image, respectively, and ρ_g and ρ_a are the corresponding density for fibroglandular and adipose tissue, respectively, according to Hammerstein et al. (1979). The true breast density of each BCT phantom was obtained by applying this equation to the phantom volumes themselves. The estimated breast density resulting from the proposed method was obtained by applying the equation to the classified reconstructed images.
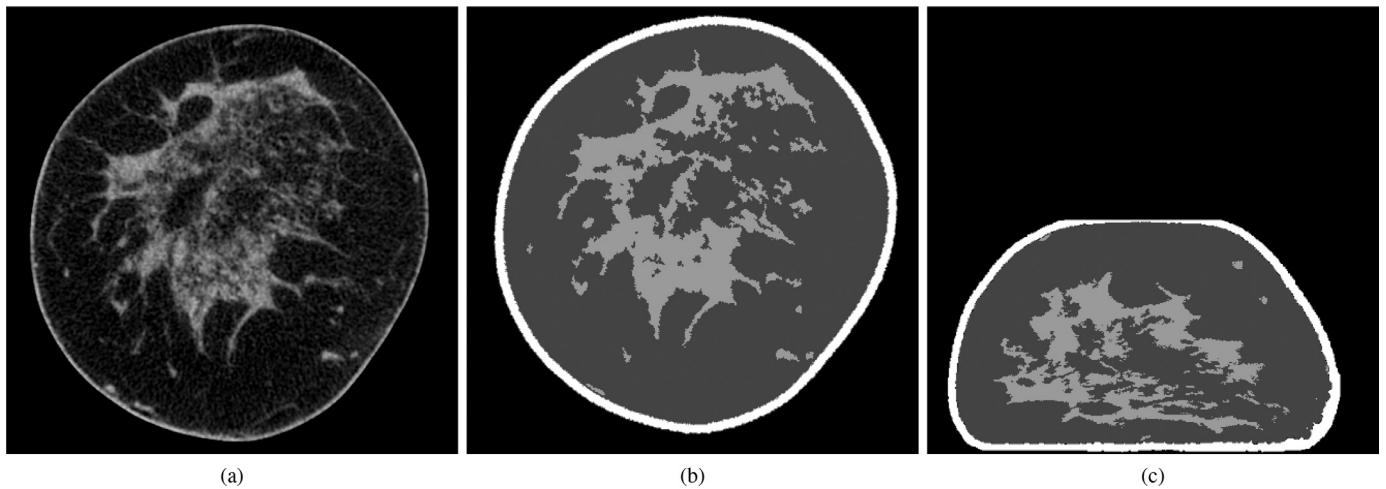
(a)   (b)   (c)

Fig. 4. Imaging geometry implemented in the Monte Carlo simulation: the X-ray source is placed at 70 cm from the detector, a 3 mm thick polyethylene terephthalate (PET) compression paddle was simulated and a large water cuboid was included to take into account the patient-body backscatter. The X-ray field irradiated the breast model at different angles (from 24° to 24°). The center of rotation is placed at 65 cm from the X-ray source. Drawing is not to scale and rotation of the detector is not shown.
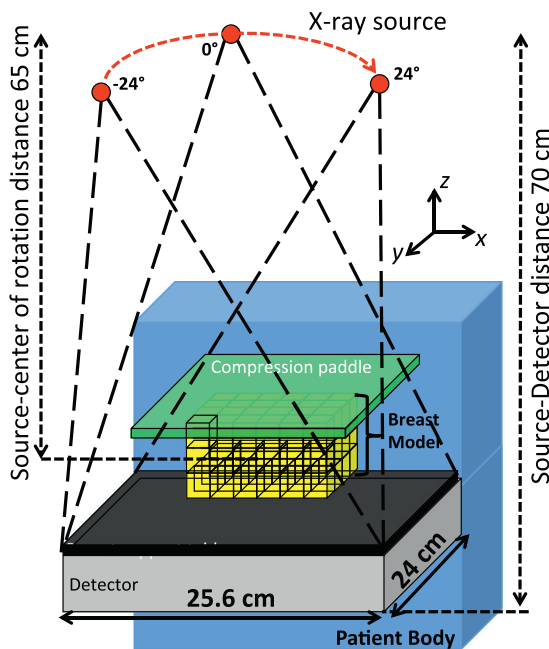

Fig. 5. Example of U-Net reconstruction artifact.

Table 1
X-ray spectra used in the Monte Carlo simulation. HVL: 1st half value layer.

Breast thickness (mm)	Spectrum	HVL (mm Al)	# Cases
30–39	W/Rh - 27 kV	0.519	2
40–49	W/Rh - 28 kV	0.530	7
50–59	W/Rh - 29 kV	0.538	6
60–69	W/Rh - 30 kV	0.547	18
70–79	W/Rh - 31 kV	0.557	12

where E_i is the energy deposited at the interaction event i, and M_g is the total fibroglandular breast mass.

10^7 primary X-rays were emitted by an isotropic point source placed at 70 cm from the detector and collimated to irradiate the entire detector. In order to replicate the tomosynthesis acquisition mode, a total of 25 projections were simulated from $-24°$ to $24°$, every 2°. The projection at 0° replicates the mammographic acquisition. The number of primary particles ensured a statistical uncertainty on the total dose of less than 0.7%, evaluated using the method proposed by Sempau et al. (2001). Photoelectric interactions, and coherent and incoherent scatter were included in the simulations without modifying the default cut range for photons (1 mm, corresponding to an energy of 2.45 keV and 2.88 keV for adipose and fibroglandular tissue, respectively).

The X-ray spectra were modeled using the TASMICS model (Hernandez and Boone, 2014) by adjusting the thickness of the modeled rhodium filter to match the first half-value layer measured with a solid state detector (RaySafe X2-MAM sensor, Billdal, Sweden) in the modelled system as shown in Table 1. The Monte Carlo simulations for estimating the MGD were performed twice for each breast; once for each of the 45 BCT phantoms, and once each for the corresponding labelled DBToR reconstructions. In this way, the accuracy of the resulting patient-specific dosimetry estimates could be assessed.

3.5. Dose calculation

The mean glandular dose (MGD) estimations were performed using a previously described and validated Monte Carlo code (Fedon et al., 2018a; 2018b), based on the Geant4 toolkit (release 10.05, December 2018). The Monte Carlo geometry replicates the one used to generate the projections and is shown in Fig. 5.

As during simulation of the DBT projections above, each voxel was labelled with an index related to its composition: air, adipose tissue, fibroglandular tissue, and skin, using the chemical compositions reported by Hammerstein et al. (1979). The energy deposited in the fibroglandular voxels was recorded and then converted into dose according to the formula

$$MGD = \frac{\sum_i E_i}{M_g} \tag{5}$$

3.6. Comparison to current reconstructions

The results of the DBToR reconstruction, prior to the voxel classification for estimation of breast density and dose, were compared to the baseline iterative MLTR reconstruction algorithm, the LPD

Table 2
Results on noise-free phantom projections, mean $\pm$ standard deviation (in bold best result).

Model	L^2-loss	SSIM	PSNR
MLTR	$0.007 \pm 4.3 \times 10^{-5}$	$0.83 \pm 4.2 \times 10^{-3}$	$20.2 \pm 2.9 \times 10^{-2}$
LPD	$0.014 \pm 6.9 \times 10^{-3}$	$0.87 \pm 2.9 \times 10^{-2}$	19.3 ± 1.8
U-Net	$\mathbf{0.002 \pm 7.0 \times 10^{-4}}$	$0.80 \pm 1.4 \times 10^{-1}$	$\mathbf{27.8 \pm 2.3}$
DBToR	$0.003 \pm 8.1 \times 10^{-4}$	$\mathbf{0.91 \pm 2.3 \times 10^{-2}}$	24.8 ± 1.3

algorithm and the U-Net trained on FBP reconstructions for both noiseless and noisy data.

3.7. Model training and evaluation

In total, we trained three versions of the DBToR algorithm: two versions were trained on virtual phantom projections, both without any noise and with varying levels of noise, and one version was pretrained on noisy virtual phantom projections and subsequently finetuned on noisy deformed BCT phantoms, on the data described in Sections 3.1 and 3.2. For comparison, we trained the basic LPD algorithm (achieved by removing the height mask from the input) in addition to DBToR on virtual phantom projections. Finally, we also trained U-Net baselines on FBP reconstructions of DBT slices as a classical deep learning reconstruction baseline. In our U-Net baselines, the standard U-Net architecture Ronneberger et al. (2015) with depth 4, 32 filters in the initial double convolution block and instance normalization layers was used.

For the DBToR and LPD models, we used a batch size of 8 for the pretraining on the virtual phantoms, and for the finetuning on the BCT phantoms for the dosimetry application, we chose a batch size of 1. For the U-net baseline, we use a batch size of 4. The number of iterations N_{iter} was set to 10^5 for all models trained on virtual phantoms, and to 4×10^5 for the final model finetuning on BCT phantoms. This leads to a training time to about 48 h and 24 h for the DBToR/LPD and the U-net models respectively. At inference stage, the DBToR reconstruction takes less than 1 s for a single coronal slice.

We use the L^2 loss, Structural Similarity Index (SSIM) (Wang et al., 2004) and Peak Signal-to-Noise Ratio (PSNR) as performance metrics.

4. Results

For DBToR, LPD and U-Net trained on noise-free virtual phantom data we report the corresponding L^2 loss, SSIM and PSNR on noise-free virtual phantom test data in Table 2. Lower L^2 loss and higher SSIM and PSNR values indicate better reconstruction performance. For DBToR, LPD and the U-Net trained on noisy virtual phantom projections, we report these metrics for noise levels $N = 4, 8, 12$ in Table 3. In both tables, we report the mean and standard deviation of the metrics obtained using 3 cross-validation folds with 60% of the data used for training and 40% used for testing in each fold.

The original LPD algorithm is significantly outperformed by DBToR for all noise levels. Visual inspection of the slices produced by LPD revealed that the LPD often reconstructs breast regions adjacent to the compression paddle very poorly for both test and training data. In particular, it frequently fails to reproduce the flatness of the part of the skin surface which is in contact with the compression paddle. We ruled out overfitting of LPD as the cause of the artifacts, since the validation metrics remained low throughout the training process. While it is possible that a much larger version of LPD with more reconstruction blocks would learn to correct these artifacts, we will see that DBToR resolves them without
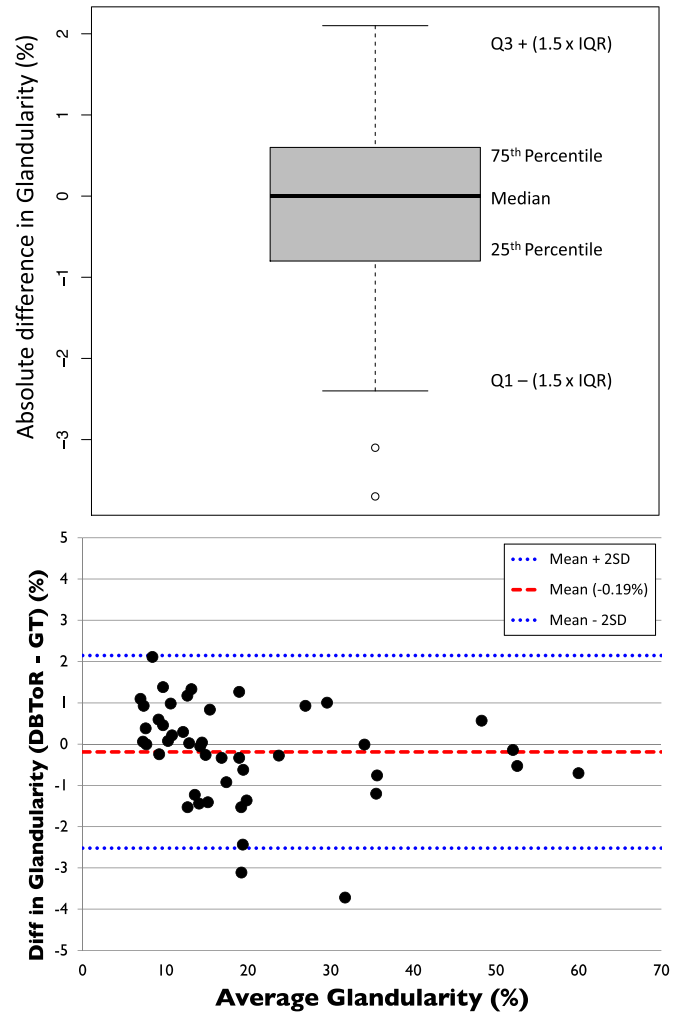

Fig. 6. Bland-Altman plot of the difference (DBToR - GT) of glandularity (in percentage points) and matching box plot with Tukey whiskers.

any further increase in the number of parameters. The U-Net architecture is outperformed by DBToR for noise levels $N = 4$ and $N = 8$, but it outperforms DBToR in terms of PSNR for noise level $N = 12$ and in the noise-free case. However, U-Net yields reconstructions that suffer from noticeably lower SSIM compared to DBToR, and a visual inspection of the volumes revealed that the U-Net also struggles with reconstructing the breast shape correctly, which can have a strong negative effect on the dosimetry accuracy. Higher PSNR for the U-Net can be, potentially, explained by the fact that U-Net reconstructs the background better, which is not important for dose computations. Additionally, the variance in the performance metrics is higher for the U-Net, indicating that it is more sensitive to the data split. Overfitting was more pronounced for the U-Net baseline due to the much higher parameter count (7.8M parameters for U-Net and 872k parameters for DBToR). These considerations make DBToR a preferred candidate for our application. The proposed DBToR algorithm also outperforms the iterative MLTR reconstruction algorithm at all noise levels and for all metrics being considered, while yielding visually more accurate reconstructions as well. It is also interesting to note from Table 3 that the performance of DBToR at noise level $N = 4$ is superior to the MLTR reconstruction algorithm at noise level $N = 12$, which corresponds to 8 times higher photon count. At noise level $N = 12$, the performance of MLTR is slightly below the performance of MLTR for the

Table 3

Results on noisy phantom projections for different noise levels *N*, mean ± standard deviation (in bold best result)

| Model	L^2-loss	SSIM	PSNR
MLTR ($N = 4$)	$0.0096 \pm 1.0 \times 10^{-4}$	$0.69 \pm 9.0 \times 10^{-3}$	$19.17 \pm 5.8 \times 10^{-2}$
LPD ($N = 4$)	$0.014 \pm 5.4 \times 10^{-3}$	$0.85 \pm 2.4 \times 10^{-2}$	18.79 ± 1.6
U-Net ($N = 4$)	$0.0085 \pm 1.4 \times 10^{-3}$	$0.64 \pm 1.7 \times 10^{-1}$	20.96 ± 1.0
DBToR ($N = 4$)	$\mathbf{0.0044 \pm 9.1 \times 10^{-4}}$	$\mathbf{0.90 \pm 1.8 \times 10^{-2}}$	$\mathbf{23.39 \pm 9.6 \times 10^{-1}}$
MLTR ($N = 8$)	$0.0082 \pm 5.2 \times 10^{-5}$	$0.74 \pm 7.2 \times 10^{-3}$	$19.87 \pm 3.6 \times 10^{-2}$
LPD ($N = 8$)	$0.013 \pm 5.3 \times 10^{-3}$	$0.84 \pm 3.0 \times 10^{-2}$	19.06 ± 1.6
U-Net ($N = 8$)	$0.0043 \pm 1.6 \times 10^{-3}$	$0.78 \pm 1.3 \times 10^{-1}$	24.13 ± 1.45
DBToR ($N = 8$)	$\mathbf{0.0033 \pm 6.3 \times 10^{-4}}$	$\mathbf{0.91 \pm 1.2 \times 10^{-2}}$	$\mathbf{24.71 \pm 8.0 \times 10^{-1}}$
MLTR ($N = 12$)	$0.0078 \pm 4.47 \times 10^{-5}$	$0.80 \pm 5.5 \times 10^{-3}$	$20.06 \pm 3.0 \times 10^{-2}$
LPD ($N = 12$)	$0.015 \pm 6.3 \times 10^{-3}$	$0.86 \pm 2.7 \times 10^{-2}$	19.04 ± 1.6
U-Net ($N = 12$)	$0.0034 \pm 1.2 \times 10^{-3}$	$0.72 \pm 1.6 \times 10^{-1}$	$\mathbf{25.58 \pm 2.3}$
DBToR ($N = 12$)	$\mathbf{0.0034 \pm 9.4 \times 10^{-4}}$	$\mathbf{0.91 \pm 1.6 \times 10^{-2}}$	24.47 ± 1.6
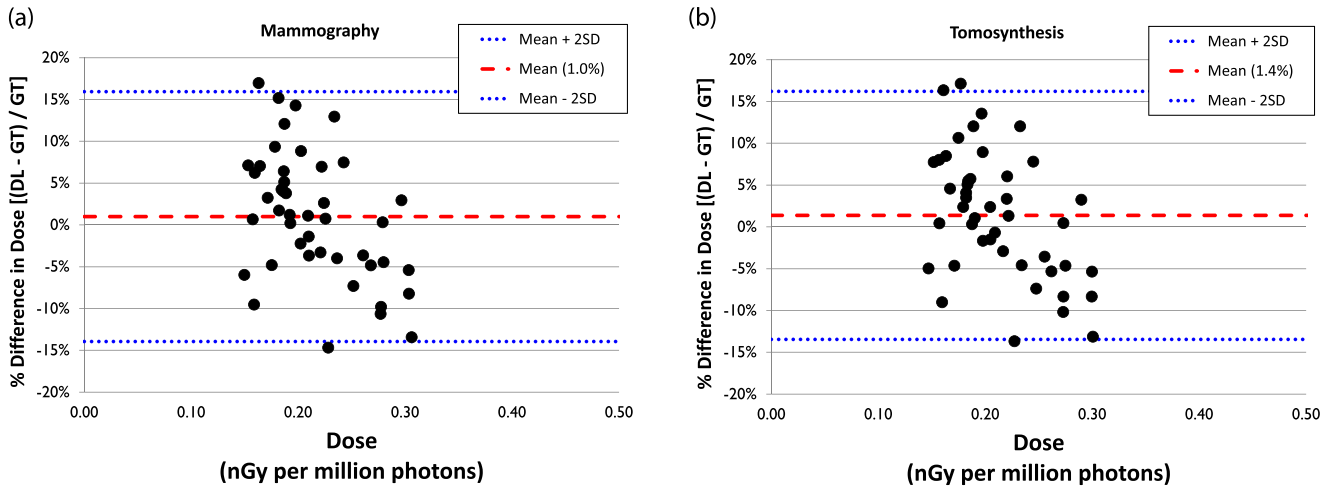
Fig. 7. Bland-Altman plot of the difference in dose estimates resulting from the DBToR reconstruction and the ground truth, in percentage, for both mammography (a) and DBT (b). The red line represents the mean, while the two blue-dashed lines represent the 95% limits of agreement. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
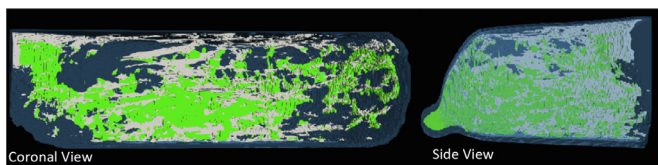

Fig. 8. Example of fibroglandular tissue distribution for an outlier case: the ground truth (in white) depicts a higher amount of fibroglandular tissue in the top breast layer (i.e., facing the X-ray tube), while the DBToR model (in green) predicts a fibroglandular distribution spread towards the anterior, and more inferior, part of the breast. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 9. (a) Coronal slice of a breast CT phantom, (b) MLTR reconstruction, (c) DBToR reconstruction, and (d) classification of DBToR reconstruction.

noise-free case. At the same time, DBToR at $N = 12$ reaches comparable level of performance to that of the DBToR on noise-free data.

For DBToR trained on noisy virtual phantom projections and subsequently finetuned on deformed breast CT slices, where we used noise level $N = 8$ only during training and finetuning, we summarize the reconstruction performance in Table 4, and in Figs. 9, 10, and 11 we give examples of coronal, axial, and sagittal slices of the virtual breast phantom and corresponding MLTR reconstruction, DBToR reconstruction, and DBToR classification (all on noisy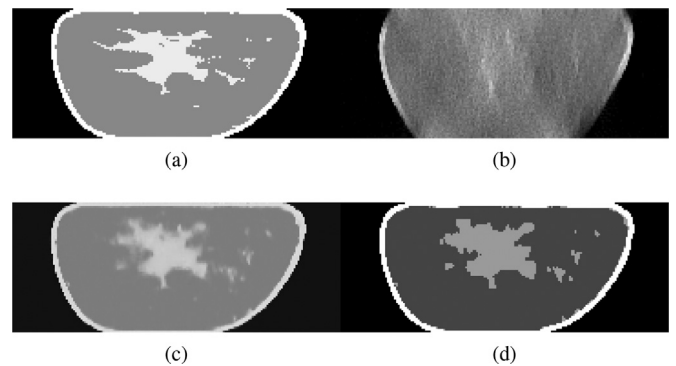 deformed breast CT slices). We observe that DBToR out-performs the baseline MLTR algorithm in terms of the reported metrics and visual quality of the slices, particularly noticeable for coronal and sagittal directions.

4.1. Breast density estimation

Fig. 6 shows a box-whisker plot of the absolute percentage difference in glandularity between the DBToR estimate and the
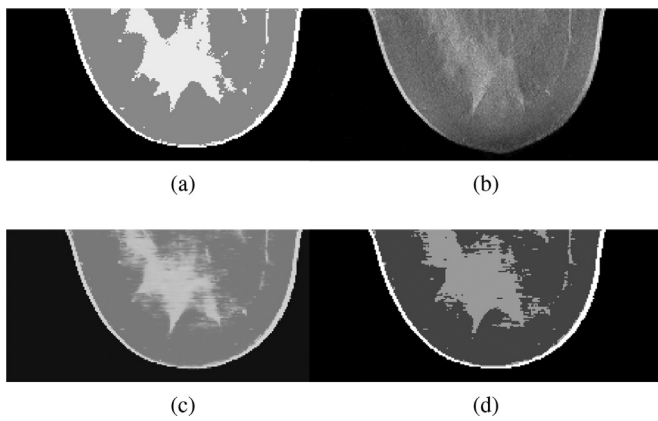
Fig. 10. (a) Axial slice of a breast CT phantom, (b) MLTR reconstruction, (c) DBToR reconstruction, and (d) classification of DBToR reconstruction. This view is created from the volume assembled by stacking all coronal slices from the same case.

Table 4
Results on noisy BCT phantom projections (in bold best result).

Model	L^2-loss	SSIM	PSNR
MLTR	0.006	0.82	21.45
DBToR	**0.0018**	**0.93**	**27.03**

ground truth (GT). It can be seen that, on average, no bias is observed (*p*-value 0.3), and that the breast density estimates are accurate to within 2%.

4.2. Mean glandular dose estimation

The comparison between the MGD evaluated from the DBToR reconstructions and the GT is shown in the Bland-Altman plots of Fig. 7, for both mammography and DBT geometries.

As can be seen, no bias is observed in the proposed dose estimation method (*p*-value 0.23), with the data points equally scattered around zero, and that the largest error in the dose estimation is less than 20%. Visual inspection of the cases that lie beyond the $\pm$2SD limits reveals that this is the consequence of differences in the reconstructed fibroglandular distribution obtained with the DBToR model compared to the GT, as shown in an example in Fig. 8.

For this case, the absolute difference on the glandularity is 1.5% (namely 12.0% for the DBToR and 13.5% for GT). Thus, due to a higher amount of fibroglandular tissue closest to the X-ray source in the case of the GT, the radiation dose estimated by DBToR is lower by 13.5%.
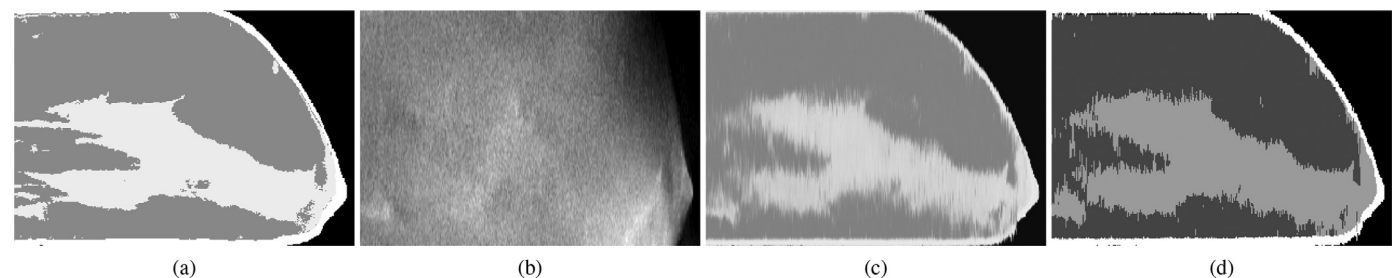
5. Discussion and conclusion

We presented a deep learning-based method for the reconstruction of DBT, which we call DBToR. The model is both data driven and model-based, since the forward and backprojection operators for a given DBT geometry are a part of our neural network architecture and at the same time the model is trained to reduce the tomographic artifacts of the reconstruction. As training data we used two sources of data, one based on random samples with statistical properties similar to real breast volumes, and one dataset of patient breast CT images that have been compressed with a finite element model to simulate the same breast under compression in a DBT system.

Compared to LPD, in DBToR we added the compressed breast thickness as prior information. Since the limited angle causes a severely ill-posed problem, it was expected that this information would be definitely required, and the experiments (Table 2) confirmed that the result dramatically degrades, compared to the 'full' DBToR, when this prior knowledge is not provided to the algorithm. Requiring this additional information does not limit the generalizability of the method, since it is readily available in all DBT systems.

The results indicate that the proposed algorithm outperforms the MLTR iterative reconstruction in terms of reconstruction quality for this application. Furthermore, the algorithm generalizes well even when trained on a small dataset and is robust to noise.

The simulated acquisitions in this work used a mono-energetic beam and did not include X-ray scattered radiation, so it remains to be seen how our new reconstruction method will handle these factors. In practice the effect of not modeling the spectrum will likely be minimal as regular filtered backprojection reconstructions also do not account for these physical effects and apply a series of precorrection steps to the projection data instead, such as the beam hardening correction described by Herman (1979). We foresee using the same approach to extend our method to work in clinical data.

We have shown that the method achieves robust and accurate predictions of breast density, which is an important metric relating to masking and cancer risk. As opposed to current density estimation methods based on mammography and DBT projections, which require assumptions and modeling of the image acquisition process, the use of the images produced by DBToR allows for a direct estimation of the amount of dense tissue present in the volume, resulting in estimates to within 3%. Accurate determination of breast density opens up further opportunities for personalized risk-based screening. As is crucial for an accurate dosimetric estimate, the location in the vertical direction of the dense tissue is also estimated with accuracy, resulting in a state-of-the-art dosimetric estimate. It is known that current model dose estimates introduce an average bias of 30%, and can misrepresent the actual patient-specific dose by up to 120% (Dance et al., 2005; Sechopoulos et al., 2012; Hernandez et al., 2015). In comparison, the results

Fig. 11. (a) Sagittal slice of a breast CT phantom, (b) MLTR reconstruction, (c) DBToR reconstruction, and (d) classification of DBToR reconstruction. This view is created from the volume assembled by stacking all coronal slices of this case.

obtained here achieve errors below 20% with no systematic bias. True patient-specific dosimetry could be used, for the first time, to gather dose registries, especially for screening, ensuring the optimal use of this imaging technology, and allowing for continuous monitoring of dose trends and providing valuable data for additional optimization and development of existing and new imaging technologies.

The main limitation of the current work is that it works on a slice by slice basis rather than on a full 3D volume. With this simplification we were able to concentrate on the network structure rather than on the logistics of handling the enormous datasets needed to train a 3D model.

The logical next step for our algorithm is an extension to fully 3D data instead of 2D slices. From there on, we could extend the model-based parts of the deep learning network instead of starting from precorrected projection data, as in current filtered back-projection methods, by including a polychromatic X-ray spectrum, X-ray scatter, and other relevant factors in the forward model. It could also be valuable to optimize the network output specifically for artificial intelligence reading by training the composite network in an end-to-end fashion. Finally, the reconstruction of a diagnostic-quality volume, for interpretation by radiologists, with the potential for the higher vertical resolution obtained here, would be a valuable improvement for clinical performance.

To conclude, we created a deep learning based reconstruction for DBT that was able to achieve accurate predictions of breast density and from there an accurate calculation of patient specific MGD.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Ioannis Sechopoulos has research agreements with Siemens Healthcare, Canon Medical Systems, Volpara Solutions, Sectra Benelux, and ScreenPoint Medical. None of these research agreements are related to the work described here.

CRediT authorship contribution statement

Jonas Teuwen: Conceptualization, Writing - original draft, Project administration. **Nikita Moriakov:** Methodology, Software, Writing - original draft. **Christian Fedon:** Writing - review & editing, Software, Validation. **Marco Caballo:** Writing - review & editing, Software. **Ingrid Reiser:** Writing - review & editing, Software. **Pedrag Bakic:** Writing - review & editing, Software. **Eloy García:** Writing - review & editing, Data curation. **Oliver Diaz:** Writing - review & editing, Data curation. **Koen Michielsen:** Methodology, Software, Data curation, Writing - review & editing. **Ioannis Sechopoulos:** Conceptualization, Methodology, Writing - original draft, Supervision, Project administration.

Acknowledgments

References

van Aarle, W., Palenstijn, W.J., Cant, J., Janssens, E., Bleichrodt, F., Dabrovolski, A., De Beenhouwer, J., Batenburg, K.J., Sijbers, J., 2016. Fast and flexible X-ray tomography using the ASTRA toolbox. Opt. Express 24 (22), 25129–25147. doi:10.1364/OE.24.025129.

Adams, R., Bischof, L., 1994. Seeded region growing. IEEE Trans. Pattern Anal. Mach. Intell. 16 (6).

Adler, J., Öktem, O., 2018. Learned primal-dual reconstruction. IEEE Trans. Med. Imaging 37 (6), 1322–1332. doi:10.1109/TMI.2018.2799231.

Arana Peña, L.M., Fedon, C., Garcia, E., Diaz, O., Longo, R., Dance, D.R., Sechopoulos, I., 2020. Monte Carlo dose evaluation of different fibroglandular tissue distribution in breast imaging. In: Van Ongeval, C., Marshall, N., Bosmans, H. (Eds.), 15th International Workshop on Breast Imaging (IWBI2020). SPIE, p. 76. doi:10.1117/12.2564278.

Arridge, S., Maass, P., Öktem, O., Schönlieb, C.-B., 2019. Solving inverse problems using data-driven models. Acta Numer. 28, 1174. doi:10.1017/S0962492919000059.

Bakic, P.R., Zhang, C., Maidment, A.D.A., 2011. Development and characterization of an anthropomorphic breast software phantom based upon region-growing algorithm. Med. Phys. 38 (6Part1), 3165–3176. doi:10.1118/1.3590357.

Bezdek, J.C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum, New York, NY.

Boone, J.M., Chavez, A.E., 1996. Comparison of X-ray cross sections for diagnostic and therapeutic medical physics. Med. Phys. 23 (12), 1997–2005. doi:10.1118/1.597899.

Caballo, M., Boone, J.M., Mann, R., Sechopoulos, I., 2018. An unsupervised automatic segmentation algorithm for breast tissue classification of dedicated breast computed tomography images. Med. Phys. 45, 2542–2559.

Dance, D.R., Sechopoulos, I., 2016. Dosimetry in X-ray-based breast imaging. Phys. Med. Biol. 61 (19), R271–R304. doi:10.1088/0031-9155/61/19/R271.

Dance, D.R., et al., 2005. Breast dosimetry using high-resolution voxel phantoms. Radiat. Prot. Dosim. 114 (1-3), 359–363. doi:10.1093/rpd/nch510.

Fedon, C., Caballo, M., Longo, R., Trianni, A., Sechopoulos, I., 2018. Internal breast dosimetry in mammography: experimental methods and Monte Carlo validation with a monoenergetic X-raybeam. Med. Phys. 45 (8), 1724–1737. doi:10.1002/mp.12792.

Fedon, C., Caballo, M., Sechopoulos, I., 2018. Internal breast dosimetry in mammography:Monte Carlo validation in homogeneous and anthropomorphic breast phantom with a clinical mammography system. Med. Phys. 45 (8), 3950–3961. doi:10.1002/mp.13069.

Fedon, C., et al., 2019. Monte Carlo study on optimal breast voxel resolution for dosimetry estimates in digital breast tomosynthesis. Phys. Med. Biol. 64 (1), 015003. doi:10.1088/1361-6560/aaf453.

Förnvik, D., Timberg, P., Zackrisson, S., Tingberg, A., Petersson, H., Dustler, M., 2018. Towards determination of individual glandular dose. In: Chen, G.-H., Lo, J.Y., Gilat Schmidt, T. (Eds.), Medical Imaging 2018: Physics of Medical Imaging. SPIE, p. 3. doi:10.1117/12.2293454.

García, E., Fedon, C., Caballo, M., Martí, R., Sechopoulos, I., Diaz, O., 2020. Realistic compressed breast phantoms for medical physics applications. In: Van Ongeval, C., Marshall, N., Bosmans, H. (Eds.), 15th International Workshop on Breast Imaging (IWBI2020). SPIE, p. 73. doi:10.1117/12.2564273.

Geeraert, N., 2014. Quantitative Evaluation of Fibroglandular Tissue for Estimation of Tissue-Differentiated Absorbed Energy in Breast Tomosynthesis. KU Leuven. Faculty of medicine Ph.D. thesis.

Hammerstein, R., Miller, D.W., White, D.R., Masterson, E., Woodard, H.Q., Laughlin, J.S., 1979. Absorbed radiation dose in mammography. Radiology 130 (2), 485–491. doi:10.1148/130.2.485.

Herman, G.T., 1979. Correction for beam hardening in computed tomography. Phys. Med. Biol. 24 (1), 008. doi:10.1088/0031-9155/24/1/008.

Hernandez, A.M., Boone, J.M., 2014. Tungsten anode spectral model using interpolating cubic splines: unfiltered X-ray spectra from 20 kV to 640 kV. Med. Phys. 41, 042101.

Hernandez, A.M., Seibert, J.A., Boone, J.M., 2015. Breast dose in mammography is about 30% lower when realistic heterogeneous glandular distributions are considered. Med. Phys. 42 (11), 6337–6348. doi:10.1118/1.4931966.

Jin, K.H., McCann, M.T., Froustey, E., Unser, M., 2017. Deep convolutional neural network for inverse problems in imaging. IEEE Trans. Image Process. 26 (9), 4509–4522. doi:10.1109/TIP.2017.2713099. http://ieeexplore.ieee.org/document/7949028/

Johnsen, S.F., et al., 2015. NiftySim: a GPU-based nonlinear finite element package for simulation of soft tissue biomechanics. Int. J. Comput. Assist. Radiol. Surg. 10 (7), 1077–1095. doi:10.1007/s11548-014-1118-5.

Kaipio, J., Somersalo, E., 2005. Statistical and Computational Inverse Problems. Applied Mathematical Sciences, 160. Springer-Verlag, New York doi:10.1007/b138659.

Kang, E., Min, J., Ye, J.C., 2017. A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction. Med. Phys. 44 (10), e360–e375. doi:10.1002/mp.12344.

Lau, B.A., Reiser, I., Nishikawa, R.M., Bakic, P.R., 2012. A statistically defined anthropomorphic software breast phantom. Med. Phys. 39 (6Part1), 3375–3385. doi:10.1118/1.4718576.

Loshchilov, I., Hutter, F., 2019. SGDR: Stochastic gradient descent with warm restarts. In: 5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings, pp. 1–16.

McCormack, V.A., 2006. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol. Biomark. Prev. 15 (6), 1159–1169. doi:10.1158/1055-9965.EPI-06-0034.

Michielsen, K., Nuyts, J., Cockmartin, L., Marshall, N., Bosmans, H., 2016. Design of a model observer to evaluate calcification detectability in breast tomosynthesis and application to smoothing prior optimization. Med. Phys. 43 (12), 6577–6587. doi:10.1118/1.4967268.

Miller, K., Joldes, G., Lance, D., Wittek, A., 2007. Total Lagrangian explicit dynamics finite element algorithm for computing soft tissue deformation. Commun. Numer. Methods Eng. 23 (2), 121–134. doi:10.1002/cnm.887.

Moriakov, N., Michielsen, K., Adler, J., Mann, R., Sechopoulos, I., Teuwen, J., 2019. Deep learning framework for digital breast tomosynthesis reconstruction. In:

Bosmans, H., Chen, G.-H., Gilat Schmidt, T. (Eds.), Medical Imaging 2019: Physics of Medical Imaging. SPIE, p. 220. doi:10.1117/12.2512912.

Niklason, L.T., et al., 1997. Digital tomosynthesis in breast imaging. Radiology 205 (2), 399–406. doi:10.1148/radiology.205.2.9356620.

Nuyts, J., De Man, B., Dupont, P., Defrise, M., Suetens, P., Mortelmans, L., 1998. Iterative reconstruction for helical CT: a simulation study. Phys. Med. Biol. 43 (4), 729–737. doi:10.1088/0031-9155/43/4/003.

del Palomar, A.P., Calvo, B., Herrero, J., Lopez, J., Doblare, M., 2008. A finite element model to accurately predict real deformations of the breast. Med Eng Phys 30 (9), 1089–1097. doi:10.1016/j.medengphy.2008.01.005.

Plevritis, S.K., et al., 2018. Association of screening and treatment with breast cancer mortality by molecular subtype in US women, 2000–2012. JAMA 319 (2), 154–164. doi:10.1001/jama.2017.19130.

Reiser, I., Nishikawa, R.M., 2010. Task-based assessment of breast tomosynthesis: Effect of acquisition parameters and quantum noisea). Med. Phys. 37 (4), 1591–1600. doi:10.1118/1.3357288.

Rodriguez-Ruiz, A., et al., 2017. New reconstruction algorithm for digital breast tomosynthesis: better image quality for humans and computers. Acta Radiol..

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Interventio MICCAI 2015, pp. 234–241.

Sechopoulos, I., Bliznakova, K., Qin, X., Fei, B., Feng, S.S.J., 2012. Characterization of the homogeneous tissue mixture approximation in breast imaging dosimetry. Med. Phys. 39 (8), 5050–5059. doi:10.1118/1.4737025.

Sempau, J., Sanchez-Reyes, A., Salvat, F., Tahar, H., Jiang, S.B., Fernandez-Varea, J.M., 2001. Monte Carlo simulation of electron beams from an accelerator head using PENELOPE. Phys. Med. Biol. 46, 1163–1186.

Six, N., De Beenhouwer, J., Sijbers, J., 2019. poly-DART: a discrete algebraic reconstruction technique for polychromatic X-ray CT. Opt. Express 27 (23), 33670. doi:10.1364/oe.27.033670.

Vedantham, S., Karellas, A., Vijayaraghavan, G.R., Kopans, D.B., 2015. Digital breast tomosynthesis: state of the art. Radiology 277 (3), 663–684. doi:10.1148/radiol.2015141303.

Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. 13 (4), 600–612. doi:10.1109/TIP.2003.819861. http://ieeexplore.ieee.org/document/1284395/

Wellman, P.S., 1999. Tactile Imaging. Harvard University Ph.D. thesis.

Zackrisson, S., et al., 2018. One-view breast tomosynthesis versus two-view mammography in the malm breast tomosynthesis screening trial (MBTST): a prospective, population-based, diagnostic accuracy study. Lancet Oncol. 19 (11), 1493–1503.

Zhang, C., Bakic, P.R., Maidment, A.D.A., 2008. Development of an anthropomorphic breast software phantom based on region growing algorithm. In: Miga, M.I., Cleary, K.R. (Eds.), Medical Imaging 2008: Visualization, Image-Guided Procedures, and Modeling, p. 69180V. doi:10.1117/12.773011.

Zhang, Z., Chen, B., Xia, D., Sidky, E.Y., Pan, X., 2021. Directional-TV algorithm for image reconstruction from limited-angular-range data. Med. Image Anal. 70, 102030. doi:10.1016/j.media.2021.102030.

Zuley, M.L., et al., 2013. Digital breast tomosynthesis versus supplemental diagnostic mammographic views for evaluation of noncalcified breast lesions. Radiology 266 (1), 89–95. doi:10.1148/radiol.12120552.