Published in final edited form as:

Phys Med Biol. ; 63(22): 225017. doi:10.1088/1361-6560/aae78d.

Development of 3D patient-based super-resolution digital breast phantoms using machine learning

Marco Caballo¹, Christian Fedon^{1,2}, Luca Brombal^{2,3}, Ritse Mann¹, Renata Longo^{2,3}, and Ioannis Sechopoulos^{1,4}

¹Department of Radiology and Nuclear Medicine, Radboud University Medical Center, PO Box 9101, 6500 HB Nijmegen, The Netherlands ²Istituto Nazionale di Fisica Nucleare (INFN), sezione di Trieste, 34127 Trieste, Italy ³Department of Physics, University of Trieste, 34127 Trieste, Italy ⁴Dutch Expert Center for Screening (LRCB), PO Box 6873, 6503 GJ Nijmegen, The Netherlands

Abstract

Digital phantoms are important tools for optimization and evaluation of x-ray imaging systems, and should ideally reflect the three-dimensional structure of human anatomy and its potential variability. In addition, they need to include a good level of detail at a high enough spatial resolution to accurately model the continuous nature of the human anatomy. A pipeline to increase the spatial resolution of patient-based digital breast phantoms that can be used for computer simulations of breast imaging is proposed. Given a tomographic breast image of finite resolution, the proposed methods can generate a phantom and increase its resolution at will, not only simply through super-sampling, but also by generating additional random glandular details to account for glandular edges and strands to compensate for those that may have not been detected in the original image due to the limited spatial resolution of the imaging system used. The proposed algorithms use supervised learning to predict the loss in glandularity due to limited resolution, and then to realistically recover this loss by learning the mapping between low and high resolution images. They were trained on high-resolution synchrotron images (detector pixel size 60 μ m) reconstructed at seven voxel dimensions (60 μ m–480 μ m), and applied to patient images acquired with a clinical breast CT system (detector pixel size 194 μ m) to generate super-resolution phantoms (voxel sizes 68 μ m). Several evaluations were made to assess the appropriateness of the developed methods, both with the synchrotron (relative prediction error 0.010 $\pm$ 0.004, recovering accuracy 0.95 $\pm$ 0.04), and with the clinical images (average glandularity error at 194 μ m: 0.15% $\pm$ 0.12%). Finally, a breast radiologist assessed the realism of the developed phantoms by blindly comparing original and phantom images, resulting in not being able to distinguish the real from the phantom images. In conclusion, the proposed method can generate super-resolution phantoms from tomographic breast patient images that can be used for future computer simulations for optimization of new breast imaging technologies.

Keywords

breast CT; breast glandularity; breast imaging; digital phantom; machine learning; super-resolution

1. Introduction

In the last decade, the introduction of dedicated breast computed tomography (breast CT) has been aimed, mainly, at improving detection and diagnosis of breast cancer by overcoming the issue of tissue superposition in 2D imaging. Breast CT can produce isotropic 3D images of the breast, being optimized, in terms of acquisition geometry and spatial and contrast resolution, for breast imaging requirements (Lindfors *et al* 2008, Kalender *et al* 2017). However, research and development in breast CT is ongoing, especially for the optimization of dynamic contrast enhanced (DCE) breast CT (Prionas *et al* 2010, He *et al* 2016, Caballo *et al* 2018a), which might have the potential to replace MRI for certain clinical applications. To aid in this optimization work, computer simulations, both deterministic and stochastic, of the image acquisition process are an invaluable tool to allow for the improvement and validation of an acquisition system before it is built or for further optimization of an existing system. For the former case, the methods described in this present work can be applied to a 4D phantom we previously developed (Caballo *et al* 2018a) for simulating perfusion contrast-enhanced breast CT imaging. For this, working with patient-based phantoms deriving from a clinical breast CT system is essential to evaluate the effect of the iodinated contrast medium within the whole breast volume, in order to optimize the acquisition settings based on a wide population. For the latter case, most of the work performed on radiation dose estimation and image quality in mammography and breast tomosynthesis has been conducted under the hypothesis of homogenous breasts, thus not accounting for the complex distribution of breast tissue and its variability among different patients. This issue can be solved by using breast CT-based phantoms which can display the real patterns of breast tissue in three dimensions (as opposed to homogeneous phantoms) and which can span over different complete breast patient volumes (as opposed to specimen images). These simulations require the inclusion of a representation of the imaged breast (i.e. a phantom image) that is realistic, flexible enough to represent various conditions, and, for some application, of known truth.

The development of voxelized breast phantoms for breast imaging simulations has already been pursued by a number of investigators. Some are based on mathematical models (Dance *et al* 2005, Bliznakova *et al* 2010, Pokrajac *et al* 2012, Baneva *et al* 2017), thus allowing a complete flexibility in the generation of different anatomical structures and a virtually unlimited resolution, but with limited realism. Others, based on empirical data, are based on the acquisition of images of breast specimens, therefore improving the realism but limiting the inter-patient variability (Hoeschen *et al* 2005). Finally, others are based on patient-images acquired with breast CT (Caballo *et al* 2018a, Li *et al* 2009, Hsu *et al* 2013). These phantoms have the advantage of including the realism and variability of actual patient data, but their accuracy is limited by the spatial resolution of the breast CT device.

The limited voxel dimensions cause, under the assumption of breast glandular fraction lower than 50 % (a fact that has been shown to be true in the majority of women, Yaffe *et al* 2009), a loss in the detected glandularity, since long edges and fine strands of glandular tissue may not be detected during the image acquisition process. This loss in glandular fraction, although small, limits the accuracy of all subsequent analysis. Furthermore, the use of digital phantoms with a spatial resolution similar to that of the system may not approximate a real breast well, leading to inaccuracies when using the phantoms in image-based simulations. In fact, to simulate an image correctly, phantoms with higher spatial resolution than that of the simulated projections are needed, so as to approximate the true condition of acquiring a discrete image of a continuous object (Saunders *et al* 2003). For this, interpolation and super-sampling could be used to reduce the nominal voxel size of the phantom, but this would not address the loss of fine glandular detail and overall loss in glandular fraction of the originating image acquisition process. However, to the best of our knowledge, no methodology to generate super-resolution breast phantoms from patient images has been proposed, with the term super-resolution referring to generating patient image-based voxelized phantoms whose resolution goes beyond the intrinsic spatial resolution capabilities of the imaging system used to acquire the underlying patient images. To address this issue, in this study we propose a machine learning-based pipeline to first estimate the glandular tissue loss due to the limited voxel size, and ultimately recover this loss and generate realistic super-resolution voxelized breast phantoms from patient images. The machine learning algorithms have been trained on high-resolution breast specimen images acquired with a synchrotron system, and then applied, for generating the phantoms, on patient images acquired with a clinical breast CT device. The need for the clinical images for breast phantoms development is due to the fact that directly using the breast specimen images acquired with the synchrotron system, therefore already having a high spatial resolution, is not feasible, since these images involve only a small portion of the breast, thus not covering the entire breast volume. In addition, they are not representative of in vivo patient breasts due to their shape and size.

The phantoms deriving from this work could be used for evaluation or optimization of breast imaging systems in terms of image acquisition, reconstruction and radiation dose, both for tomographic imaging (e.g. breast CT) and for digital breast tomosynthesis and mammography, with the important advantage of overcoming the limitations in resolution across different imaging techniques.

2. Methods

Our pipeline is composed by several steps devoted to super-sample and segment the input patient breast CT image reconstructed at a given voxel size, predict the glandularity at a smaller target voxel size, and recover the estimated glandular detail loss. Prediction and recovery steps are both handled using machine learning, through a multi-step regression model to relate glandularity and voxel size (prediction), and a convolutional neural network which is trained to map between low- and high-resolution binary images (recovery). The whole process outputs, for a given patient breast CT image, a phantom with the desired higher spatial resolution, accounting for the loss of glandular fraction. Since clinical breast CT systems have limited resolution, the data used to design and train the two models of our

pipeline (i.e. estimate the loss in glandularity due to limited voxel size, and perform the glandular detail recovery) are extracted from high-resolution images of human breast tissue samples acquired at a synchrotron light source equipped with a 60 μm -detector pixel size. The reason for this is that, although the synchrotron images may not be fully predictive of what clinical breast CT images would look like if acquired at high resolution, the purpose of the study is to train our algorithms on better images in terms of spatial resolution and glandular details detection, and then transfer these properties to the breast phantoms themselves, not to the subsequent simulated breast CT images.

Although the proposed pipeline can work for the whole continuous range of all possible voxel sizes, we describe the whole pipeline considering a 1:4 spatial resolution increase, due to the adequacy of this ratio in approximating the continuous condition in computer simulations of digital radiographic imaging systems (Saunders *et al* 2003).

2.1 Image Acquisition

To train the algorithms we propose, we used data extracted from five images of human breast tissue samples (two examples are shown in figure 1) acquired with a breast CT system based on synchrotron radiation (SR) (Longo *et al* 2016). The system is under development at the SYRMEP beamline the Elettra synchrotron light source in Trieste (Italy) in the framework of the SYRMA-3D project funded by the INFN (Italian Institute of Nuclear Physics) (Delogu *et al* 2017). The beamline is equipped with a CdTe photon counting detector (PIXIRAD-8) with a nominal pixel size of 60 μm (Delogu *et al* 2016). Acquisition protocol and data processing are discussed in Brombal *et al* 2018a and Brombal *et al* 2018b. The monochromatic SR beam coherence and the placement of the sample 1.8 meters away from the detector allows to exploit the phase-contrast effects in free-propagation modality (Sarno *et al* 2016). The scatter contribution is also reduced thanks to the parallel geometry of the beam (Johns and Yaffe 1982). To further improve image quality, a phase-retrieval algorithm was applied to the acquired projections prior to reconstruction, resulting in an increased contrast-to-noise ratio with a preserved spatial resolution (Gureyev *et al* 2017, Brombal *et al* 2018a). The acquired breast sample images were reconstructed at seven different isotropic voxel dimensions (60 μm , 120 μm , 180 μm , 240 μm , 300 μm , 360 μm and 480 μm). The images are used within this present work according to the European Parliament directive 2004/23/EC, and the directive of the Council (March 31st 2004) on setting standards of quality and safety for the donation, procurement, testing, processing, preservation, storage and distribution of human tissues. The synchrotron images reported in this study were acquired according to the standard procedures of the Pathology Unit of the Academic Hospital of Cattinara, Trieste University, accredited by the Joint Commission International (JCI). The samples were prepared from specimens of breast lumpectomy and were derived from surgical material sent to the Pathology Unit according to local guidelines for histological examination.

For the generation of high-resolution breast phantoms with recovered glandular fraction, images of patient breasts were acquired by trained radiographers with a dedicated breast CT clinical system (Koning Corp., West Henrietta, NY) installed at Radboud University Medical Center (Nijmegen, the Netherlands) (O'Connell *et al* 2010). These images were acquired as

part of an ethics board-approved patient trial on dedicated breast CT, and all patients provided written informed consent. The clinical breast CT device (source-to-imager distance of 92.3 cm, source-to-isocenter distance of 65 cm) is equipped with an x-ray tube with a tungsten target and aluminium filter set to a voltage of 49 kV for all acquisitions. The x-ray source (half-cone beam geometry) has a nominal focal spot size of 0.3 mm and a first half value layer of 1.39 mm Al (Boone *et al* 2001, Sechopoulos *et al* 2010). The detector is 397 mm × 298 mm (4030CB, Varian Medical Systems, Palo Alto, California, USA) with a nominal pixel size of 194 µm, and the tomographic reconstruction of the images was performed using a filtered backprojection (FBP) algorithm (reconstruction voxel size of 273 µm).

2.2 Glandular Tissue Segmentation and Data Extraction

The aim of this processing step is to automatically calculate, for each synchrotron image, the glandularity in different volumes of interest (VOIs) over the seven reconstructed voxel size images, and then fit these values with a polynomial function. For this, twelve different VOIs were positioned in different locations of the one of the breast specimen images reconstructed at 60 µm. The VOIs had different dimensions and were manually located within the breast specimen (i.e. the air was excluded) so that different amounts of glandular tissue were included within each VOI.

For each of these twelve VOIs, corresponding VOIs were selected at the same locations in each of the other six reconstructed images of the same tissue sample, with the number of voxels in each VOI automatically adjusted to maintain the overall VOI size according to the image voxel size. As a result of this process, twelve groups of seven VOIs each were identified, with volumes within each group containing the same image patch scaled at different voxel sizes.

For all groups of VOIs, glandular tissue voxels were classified using a previously described dedicated breast CT image segmentation algorithm (Caballo *et al* 2018b). This algorithm adopts an automatic energy minimizing framework based on image first and second order features for the separation of glandular tissue from adipose tissue. By way of example, three slices of the VOIs from the same group reconstructed at 60 µm, 240 µm and 480 µm and the respective segmented glandular tissue are displayed in figure 2. After segmentation, the glandular fraction was calculated for each of the 84 VOIs according to the following equation:

$$G = \frac{d_G \cdot (\# \text{ Glandular Voxels})}{d_G \cdot (\# \text{ Glandular Voxels}) + d_A \cdot (\# \text{ Adipose Voxels})} \quad (1)$$

where d_G and d_A are the density values for the glandular and adipose tissue (1.04 g/cm³ and 0.93 g/cm³, respectively Hammerstein *et al* 1979).

This same process was repeated for the other four synchrotron images, but with the VOIs automatically positioned and adjusted on the 60 µm image of each sample so as to match the same glandularity values of the corresponding VOIs on the first evaluated sample image.

This allowed assessing how the loss in glandularity varies according to the glandular tissue patterns, starting from equal glandular tissue densities. To train the learning algorithm described in the following subsection, the mean values for each VOI (averaged over the five breast images) were taken.

2.3 Glandular Detail Loss Estimation

A smooth function was estimated, per VOIs group, from the glandular variation with voxel size data. For each group, a machine learning-based regression algorithm previously proposed (Caballo *et al* 2018c) was used to fit the glandularity values over the different voxel dimensions. The algorithm aims to find the best polynomial regression by minimizing the following cost function:

$$J(\theta) = \frac{1}{7} \sum_{i=1}^7 [h_{\theta}(x^{(i)}) - y^{(i)}]^2 \quad (2)$$

where θ is a vector containing the regression parameters θ_j , $h_{\theta}(x)$ is the hypothesis function (i.e. the regression to be estimated), y is the true glandularity value calculated from each VOI, and x the different voxel dimensions. The minimization of J as a function of the parameters θ is achieved by applying stochastic gradient descent (Burgess *et al* 2005) for 150,000 training iterations, according to the following:

$$\theta_j = \theta_j - \alpha \frac{d}{d\theta_j} J(\theta) \quad (3)$$

where α is the algorithm learning rate (experimentally set to 10^{-3}).

After the training, the process resulted in a functional regression which minimizes the distance between the training data (i.e. the real glandularity values detected via segmentation at each voxel size) and the predicted regression function itself. For all the twelve groups of VOIs, the function with the lowest error had two polynomial features (linear and square root) and a bias term:

$$h_{\theta}(x) = \theta_0 + \theta_1 x + \theta_2 \sqrt{x} \quad (4)$$

This low number of polynomial features also allowed to prevent overfitting of the model.

As expected, different glandularity values resulted in different parameters θ_0 , θ_1 , θ_2 for the model, i.e. each of the 12 regressions was different. This is due to the loss of glandular tissue varying according to the amount of glandular tissue voxels found within each group of VOIs (in general, the more the glandular tissue, the more the absolute loss of detail due to a finite voxel size). Therefore, we developed a subsequent machine learning algorithm applied in cascade to fit the parameters θ_0 , θ_1 , θ_2 (over all possible values of glandularity). The algorithm performs a different polynomial fitting for each θ_j , resulting in regression

functions that can estimate the value of each θ_j , for any amount of glandular tissue within a breast tomographic image. The reference voxel size can be chosen at will, but since our aim is to ultimately apply this trained algorithm to the clinical breast CT patient images, we set it to 273 μm .

The whole process allows to correctly tune the regression parameters according to the amount of glandular tissue found on an input image, and then use these parameters to design a proper regression function to estimate the glandularity at the desired voxel size. Figure 3 shows a schematic diagram of the whole process described in this section.

2.4 Glandular Detail Recovery

To adequately represent a continuous object (the breast) with a discrete model (the phantom), simply reducing the voxel size through interpolation does not recover the fine glandular details that have been lost during the image acquisition process due to limited system spatial resolution. Therefore, we propose an additional processing step which, after image interpolation, can randomly add edges and strands of glandular tissue until reaching the glandularity value predicted by the first part of the proposed pipeline. As mentioned, in this work we consider as an adequate approximation of continuity a resolution increase of a factor of 4 (Saunders *et al* 2005).

Since we aim at applying this process to breast CT images acquired with the clinical breast CT system, the voxel size was first reduced from 273 μm to 68 μm . Before super-sampling, the breast CT image underwent segmentation where each voxel was classified as either skin, glandular or adipose tissue using the segmentation algorithm used in subsection 2.2. The glandularity at the original resolution was calculated for the whole breast from the segmented image, and then the expected breast glandularity at 68 μm was predicted for each patient image using the process described above.

The task of recovering glandular details was designed as iteratively thickening edges and extending strands of the breast glandular tissue, adding at realistic locations only a few voxels per iteration.

Therefore, we implemented a convolutional neural network (CNN) trained to map between low resolution and high resolution binary images. The network we implemented is inspired by the work of Dong *et al* (2014). The network, shown in figure 4, receives as input an upsampled low-resolution image, performs the feature extraction and the non-linear mapping to high-resolution patches, and finally reconstructs the final output image from the high-resolution patches.

The network contains only convolutional blocks, whose output is propagated using rectified linear units (ReLU) except for the last layer which, performing the reconstruction of the final image, is designed with a linear activation function.

We trained the network using pairs of image patches (voxel size of 240 μm as input, voxel size of 60 μm as labels) taken from the synchrotron breast CT images, after upscaling the input images so as to match the output patch dimensions. Although not strictly equal to the desired voxel size mapping (from 273 μm to 68 μm), the regression function relating

glandularity loss and voxel size resulted in an average deviation from linearity lower than 5% for the interval [240–60 μ m]. Therefore, we could assume the mapping 240–60 μ m approximately equal to the mapping 273–68 μ m.

The network was trained for 50,000 epochs on a dataset of 10,000 pairs of image patches (90% used for training, 10% for cross-validation) with dimensions of 60 $\times$ 60 voxels, and tested on another 500 image patches with the same dimensions. The mean-squared error loss is minimized using stochastic gradient descent with the standard backpropagation (Lecun *et al* 1998), using a batch size of 80 samples and a learning rate (which was set experimentally) of 10^{-4} .

After the training was complete, the network is applied at random locations along the edges of the segmented glandular tissue in the upsampled patient breast CT images, adding glandular details until reaching the predicted glandularity value given by the algorithm described in the previous section.

2.5 Validation

To validate our methods, five different tests (summarized in Table I) were performed.

To test the first part of the proposed pipeline, five additional groups of VOIs were obtained from different locations of the synchrotron images, segmented, and the glandularity within each volume calculated. The standard error between the true glandularity values (obtained after segmentation) and the ones predicted by the regression function was calculated (test 1).

To validate the second part of the pipeline, 500 different image patches (dimensions: 60 $\times$ 60 voxels) were used to test the performance of the network (test 2). Dice similarity coefficient was used to compare the results of the network (high-resolution predictions) with the corresponding real high-resolution image patches (high-resolution labels).

Finally, to evaluate the global performance of the entire system using the clinical breast CT images, first we segmented ten breast CT images acquired with the clinical system (reconstruction voxel size of 273 μ m). From these images, corresponding super-resolution phantoms were generated and then downsampled back from 68 μ m to 273 μ m. The resulting glandularity values were calculated and compared to the original ones at 273 μ m (test 3).

Second, another ten breast CT images acquired with the clinical system were reconstructed both at 273 μ m and at 194 μ m (i.e. the detector pixel size), and the glandularity for the latter was calculated after segmentation. We then applied our whole algorithm to the 273 μ m-images, targeting this time a voxel size of 194 μ m. For this, the same network presented in the previous subsection was re-trained using the synchrotron image pairs 240–180 μ m. The resulting glandularity between the images was again compared (test 4). Finally, to evaluate the realism of these generated phantoms, we created a group of 30 pairs of slices by extracting three slices from each pair of images (194 μ m-reconstructed images after segmentation, and generated super-resolution 194 μ m-phantoms). We showed the group of images to an experienced breast radiologist, who was requested to blindly recognize the estimated image from the real one within each pair (test 5). The radiologist was free to evaluate all images with no limit of time and with the possibility to zoom at will within any

image region, and was requested to always provide an answer for each image pair (even in case of uncertainty).

3. Results

Figure 5 (top panel) shows the regression parameters θ_0 , θ_1 , θ_2 found for twelve groups of ROIs of the training set, and the corresponding fitting used for the glandular detail loss estimation.

Some examples of high-resolution phantoms generated with the proposed pipeline are shown in figure 6.

For all five groups of VOIs of the test set the regression model estimated from the training VOIs could fit the real data well, with an average relative standard error of 0.010 ± 0.004 (test 1). Two examples of real and estimated (continuous) values for two groups of VOIs of the test set are shown in figure 6 (bottom panel).

As expected, the glandular detail loss over images with different voxel dimensions tends to increase with the amount of glandular tissue voxels. For example, the total absolute glandular detail loss from $60 \mu\text{m}$ to $480 \mu\text{m}$ for the first group of VOIs of the test set was 1.52 % (starting from a glandularity of 8.03 % detected at $60 \mu\text{m}$), while for the last group the total loss was 7.7 % (starting from a glandularity of 48.04 % detected at $60 \mu\text{m}$). This points out that a higher percentage of glandular tissue within the breast corresponds to a higher number of thin glandular tissue strands and longer edges which may be missed due to image spatial discretization, which leads to a higher global loss of detail due to a larger reconstructed voxel size. This also affects the absolute error between real and estimated percentage of glandularity, whose values increase (as expected) with the glandularity amount detected at the reference voxel size. All numerical values are reported in Table A1 and Table A2 in the online Appendix.

The validation for the second part of the pipeline resulted in an average Dice similarity between the true high-resolution labels and the predicted high-resolution image patches of 0.95 ± 0.04 (test 2).

The results of the evaluation of the global performance of our methods (test 3 and 4) are reported in Table 2 and Table 3. The comparison of glandularity values between the downsampled super-resolution phantoms (from $68 \mu\text{m}$ to $273 \mu\text{m}$) and the corresponding original breast CT images reconstructed at $273 \mu\text{m}$ resulted in an average error of $0.27 \% \pm 0.13 \%$. The comparison between the original breast CT images reconstructed at $194 \mu\text{m}$ and the phantoms generated (from the same patient cases) from $273 \mu\text{m}$ to $194 \mu\text{m}$ also resulted in a small average error ($0.15 \% \pm 0.12 \%$).

Finally, the radiologist assessment (test 5) confirmed the realism of the generated phantoms, with the test resulting in random outcomes (47 % accuracy). Using the normal distribution to approximate binomial probabilities, and assuming random outcomes as the null hypothesis, the test resulted in a p-value of 0.3 (random outcomes).

4. Discussion

The pipeline we propose can estimate the expected glandularity value over different voxel dimensions, and can then use these results in combination with deep learning to recover the loss in glandularity due to the finite reconstruction voxel size by generating further glandular details, ultimately resulting in the generation of realistic, high-resolution breast phantoms. The whole pipeline resulted in an accurate glandularity prediction over different voxel dimensions, and realistic glandular fraction recovery.

The main limitation of the present work is the limited availability of images used to train the algorithms, since these images need to be acquired with a sufficiently small detector pixel size which, to the best of our knowledge, has up to now only been used for limited research studies. Although, from our results, the loss in glandularity due to limited voxel size does not seem to be strongly dependent on the shape of the glandular tissue, with a larger dataset more accurate and reliable measurements might be achieved. The limited number of samples also limits the application of more advanced machine learning algorithms such as perceptron or deep neural networks to predict the glandularity loss over different voxel sizes. In fact, although the dimension of the feature-space is not large, applying these algorithms to a non-exhaustive dataset could drastically increase the risk of overfitting of the model, therefore increasing the prediction error. Another limitation is related to the physical detector pixel size of the synchrotron system used in this study, which does not allow to depict glandular details below 60 μm . Consequently, the resolution of the generated phantoms is still limited by the best resolution of the imaging system. This is a common issue for all patient image-based digital phantoms which are, by definition, constrained to the resolution of the acquired patient image. However, if higher spatial resolution systems become available, our methods could be easily trained on the new images, allowing the mapping between any resolution range.

The deep learning-based approach that we adopted could recover the loss of glandularity in a realistic manner, with the advantage of directly learning from real high-resolution images, a process which could lead to a high level of realism. Although the process of adding details and enlarging the boundaries of foreground objects in binary images has been widely exploited by mathematical models (e.g. mathematical morphology), the need of adapting these models to a wide-case scenario limits their realism (Banon *et al* 2007). For example, the traditional morphological closing operation (a well-established approach for binary edge thickening and detail recovery) performs the binary detail recovery operations based on structuring elements of simple shapes, limiting the level of realism for complex shapes such as the glandular tissue within the breast (an example is shown in figure 7).

The network adds, at each iteration, glandular details which respect the matching between the starting lower resolution and the desired higher resolution. However, due to the random positioning in different areas of the breast, it could happen that the network spans over the same location more than once. This could happen especially in dense breasts, which statistically show a higher glandularity loss due to partial volume effect. In real images the loss in glandularity could be not only related to edges, but also to isolated small areas (within a region of interest) which disappear in low-resolution images and affect the

glandularity loss. Although the network can recover the glandularity in a realistic manner by adding details to the already detected fibroglandular tissue, it is not able to create isolated, realistic spots from scratch. This leads, in case of large glandularity losses, to the glandularity not being fully recovered unless the network spans over some locations which had already been evaluated. This also explains the uncertainties in the results reported in Table 2 and Table 3. Perhaps more accurate results could be achieved if a much larger dataset of high-resolution images was available, from which the network could learn the task of increasing the phantom resolution on a global basis, as opposed to on regions of interest.

In this work, we used voxel-wise binary classification to isolate the breast glandular tissue and subsequently calculate the glandular fraction. A different approach could be to use an alternative segmentation method that results in a continuous (or semi-continuous) scale representing mixture of adipose and glandular tissue per voxel. Such an approach would require a different method for determining and compensating for the glandularity loss across different image resolutions. In addition, a continuous distribution of voxel values does not eliminate the need for glandular detail recovery, since predicting how the mixture adipose/glandular of a given voxel splits into smaller voxels remains to be investigated. By performing binary classification using a dedicated segmentation method whose objective function depends on global image features (and not just on a gray levels-based threshold), and by subsequently recovering the glandularity through machine learning, we could reduce the dynamics of the phantoms to their simplest representation, but accounting for the loss derived from the application of a discrete segmentation method across different image resolutions.

Future work includes the evaluation of our pipeline on a larger dataset, both for the synchrotron and for the clinical breast CT images. With a larger number of cases also different types of learning approaches may be taken into account; this would allow to reduce the complexity of our method and merge the entire pipeline into a single algorithm which, with a greater number of training examples, could provide comparable or better results with a low variance error.

5. Conclusion

We present a method to develop super-resolution breast phantoms from patient tomographic breast images, accounting for and recovering the loss in glandularity and glandular detail due to limited spatial resolution. Our method was trained on high-resolution synchrotron images reconstructed at different voxel sizes, therefore directly learning from real data, and was applied to patient breast CT images acquired with a clinical system for the final generation of super-resolution phantoms. Such generated digital phantoms can be used to develop, optimize, and evaluate x-ray breast imaging systems.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This research was supported in part by grant R01CA181171 from the National Cancer Institute, National Institutes of Health, and grant IIR13262248 from the Susan G. Komen Foundation for the Cure. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute, the National Institutes of Health, or the Komen Foundation. The authors thank the Italian SYRMA-3D collaboration, supported by INFN and Elettra-Sincrotrone Trieste, for the synchrotron-based data-set used in this study.

References

- Baneva Y, Bliznakova K, Cockmartin L, Marinov S, Buliev I, Mettievier G, Bosmans H, Russo P, Marshall N, Bliznakov Z 2017 Evaluation of a breast software model for 2D and 3D X-ray imaging studies of the breast *Phys Med.* 41 78–86 [PubMed: 28483356]
- Banon GJF, Barrera J, Braga-Neto UM (Eds.) 2007 *Mathematical Morphology and its Applications to Signal and Image Processing Proceedings of the 8th international symposium on mathematical morphology (ISMM'07)* (Fontainebleau, France, Springer)
- Bliznakova K, Suryanarayanan S, Karellas A, Pallikarakis N 2010 Evaluation of an improved algorithm for producing realistic 3D breast software phantoms: application for mammography *Med Phys.* 37(11) 5604–17 [PubMed: 21158272]
- Boone JM, Nelson TR, Lindfors KK, Seibert JA 2001 Dedicated breast CT: radiation dose and image quality evaluation *Radiology* 221(3) 657–67 [PubMed: 11719660]
- Brombal L, Golosio B, Arfelli F, Bonazza D, Contillo A, Delogu P, Donato S, Mettievier G, Oliva P, Rigon L, Taibi A, Tromba G, Zanconati F, Longo R. 2018a Monochromatic breast CT: absorption and phase-retrieved images *Proc. of SPIE Vol.* 10573 1057320–1
- Brombal L, Donato S, Brun F, Delogu P, Fanti V, Oliva P, Rigon L, Di Trapani V, Longo R, Golosio B 2018 Large-area single-photon-counting CdTe detector for synchrotron radiation computed tomography: a dedicated pre-processing procedure *J. Synchrotron Rad* 25
- Burges C, Shaked T, Renshaw E, Deeds M, Hamilton N, Hullender G 2005 Learning to rank using gradient descent *ICML* 89–96
- Caballo M, Boone JM, Mann R, Sechopoulos I 2018b An unsupervised automatic segmentation algorithm for breast tissue classification of dedicated breast computed tomography images *Med Phys* 45(6) 2542–2559 [PubMed: 29676025]
- Caballo M, Mann R, Sechopoulos I 2018a Patient-Based 4D Digital Breast Phantom for Perfusion Contrast-Enhanced Breast CT Imaging *Med Phys*, doi: 10.1002/mp.13156
- Caballo M, Fedon C, Brombal L, Michielsen K, Longo R, Sechopoulos I 2018c Automatic Estimation of Glandular Tissue Loss due to Limited Reconstruction Voxel Size in Tomographic Images of the Breast *Proc. SPIE* 10718, 14th International Workshop on Breast Imaging (IWBI 2018), 107181G (6 July 2018)
- Dance DR, Hunt RA, Bakic PR, Maidment AD, Sandborg M, Ullman G, Alm Carlsson G 2005 Breast dosimetry using high-resolution voxel phantoms *Radiat Prot Dosimetry.* 114(1–3) 359–63 [PubMed: 15933137]
- Delogu P, Oliva P, Bellazzini R, Brez A, de Ruvo PL, Minuti M, Pinchera M, Spandre G, Vincenzi A 2016 Characterization of Pixirad-1 photon counting detector for X-ray imaging *JINST* 11 P01015
- Delogu P, Golosio B, Fedon C, Arfelli F, Bellazzini R, Brez A, Brun F, Di Lillo F, Dreossi D, Mettievier G, Minuti M, Oliva P, Pichera M, Rigon L, Russo P, Sarno A, Spandre G, Tromba G, Longo R. 2017 Imaging study of a phase-sensitive breast-CT system in continuous acquisition mode *JINST* 12 C01016
- Dong C, Loy CC, He K and Tang X 2014 Learning a Deep Convolutional Network for Image Super-Resolution *ECCV Lecture Notes in Computer Science* 8692
- Gureyev TE, Nesterets YI, Kozlov A, Paganin DM, Quiney HM 2017 On the “unreasonable” effectiveness of transport of intensity imaging and optical deconvolution *Journal of the Optical Society of America A* 34(12) 2251–2260
- Hammerstein GR, Miller DW, White DR, Masterson ME, Woodard HQ, Laughlin JS 1979 Absorbed radiation dose in mammography *Radiology* 130(2) 485–91 [PubMed: 760167]

- He N, Wu YP, Kong Y, Lv N, Huang ZM, Li S, Wang Y, Geng ZJ, Wu PH, Wei WD 2016 The utility of breast cone-beam computed tomography, ultrasound, and digital mammography for detecting malignant breast tumors: A prospective study with 212 patients *Eur J Radiol* 85(2) 392–403 [PubMed: 26781145]
- Hoeschen C, Fill U, Zankl M, Panzer W, Regulla D, Döhring W 2005 A high-resolution voxel phantom of the breast for dose calculations in mammography *Radiat Prot Dosimetry*. 114(1–3) 406–9 [PubMed: 15933147]
- Hsu CM, Palmeri ML, Segars WP, Veress AI, Dobbins JT, 3rd 2013 Generation of a suite of 3D computer-generated breast phantoms from a limited set of human subject data *Med Phys*. 40(4) 043703 [PubMed: 23556929]
- Johns PC, Yaffe M 1982 Scattered radiation in fan beam imaging systems *Med Phys*. 9(2) 231–9 [PubMed: 7087909]
- Kalender WA, Kolditz D, Steiding C, Ruth V, Lück F, Rößler AC, Wenkel E 2017 Technical feasibility proof for high-resolution low-dose photon-counting CT of the breast *Eur Radiol*. 27(3) 1081–1086 [PubMed: 27306559]
- Lecun Y, Bottou L, Bengio Y, Haffner P 1998 Gradient-based learning applied to document recognition *Proceedings of the IEEE* 86(11) 2278–2324
- Li CM, Segars WP, Tourassi GD, Boone JM, Dobbins JT, 3rd 2009 Methodology for generating a 3D computerized breast phantom from empirical data *Med Phys*. 36(7) 3122–31 [PubMed: 19673211]
- Lindfors KK, Boone JM, Nelson TR, Yang K, Kwan AL, Miller DF 2008 Dedicated Breast CT: Initial Clinical Experience *Radiology* 246(3) 725–33 [PubMed: 18195383]
- Longo R, et al. 2016 Towards breast tomography with synchrotron radiation at Elettra: first images *Phys Med Biol*. 61(4) 1634–49 [PubMed: 26836274]
- O’Connell A, Conover DL, Zhang Y, Seifert P, Logan-Young W, Lin CF, Sahler L, Ning R 2010 Cone-beam CT for breast imaging: Radiation dose, breast coverage, and image quality *AJR Am J Roentgenol*. 195(2) 496–509 [PubMed: 20651210]
- Pokrajac DD, Maidment AD, Bakic PR 2012 Optimized generation of high resolution breast anthropomorphic software phantoms *Med Phys*. 39(4) 2290–302 [PubMed: 22482649]
- Prionas ND, Lindfors KK, Ray S, Huang SY, Beckett LA, Monsky WL, Boone JM 2010 Contrast-enhanced dedicated breast CT: initial clinical experience *Radiology* 256(3) 714–23 [PubMed: 20720067]
- Sarno A, Mettivier G, Golosio B, Oliva P, Spandre G, Di Lillo F, Fedon C, Longo R, Russo P 2016 Imaging performance of phase-contrast breast computed tomography with synchrotron radiation and a CdTe photon-counting detector *Phys. Med* (2016), 32(5) 681–690 [PubMed: 27160829]
- Saunders RS, Jr, Samei E 2003 A method for modifying the image quality parameters of digital radiographic images *Med Phys*. 30(11) 3006–17 [PubMed: 14655948]
- Sechopoulos I, Feng S, D’Orsi CJ 2010 Dosimetric characterization of a dedicated breast computed tomography clinical prototype *Med Phys*. 37(8) 4110–20 [PubMed: 20879571]
- Yaffe MJ, Boone JM, Packard N, Alonzo-Proulx O, Huang SY, Peressotti CL, Al-Mayah A, Brock K 2009 The myth of the 50–50 breast *Med Phys*. 36(12) 5437–43 [PubMed: 20095256]

Fig. 1.
Two example slices of the five breast tissue specimen images acquired with the synchrotron breast CT system.

Fig. 2. Three example VOIs from the same group reconstructed at (a) 60 μm , (b) 240 μm and (c) 480 μm , and (d), (e), (f) the respective segmented glandular tissue.

Fig. 3.

Schematic representation of the entire processing step described in subsection 2.3

Fig. 4. Schematic representation of the convolutional neural network implemented in this work to perform the glandular detail recovery. The network extracts features from low resolution images and maps them into the corresponding high resolution patches. The network uses rectified linear units (ReLUs) as activation functions between each layer

Fig. 5.

Top panel: estimation of parameters θ_0 , θ_1 , θ_2 over a full range of glandularity given the regression parameters found for the twelve groups of VOIs, with r-squared (coefficient of determination) reported. Bottom panel: examples of regression functions for two of the five groups of VOIs of the test set of the first part of the proposed pipeline (glandular tissue loss estimation). The green line was estimated with the proposed algorithms based on the training set; the red dots are real glandularity values obtained with segmentation. The average relative errors between the continuous green line and the red dots are (left panel) 0.003 and (right panel) 0.04.

Fig. 6. (a) Example of original, segmented breast CT slice from a patient case, and (b) the corresponding slice of the generated super-resolution phantom. Panels (c) and (d) show a magnification of the same region of both slices.

Fig. 7.

Comparison of glandular detail recovery through simple morphological closing and the proposed CNN. Panel (a) shows a segmented slice of a patient breast CT image reconstructed at 273 μm ; panel (b) and (c) show the corresponding slice in two super-resolution phantoms generated from the same patient case, using simple morphological closing with a circular structuring element (b) and the proposed CNN-based method (c).

Table 1.

Summary of the five tests to validate our methods.

Test #	Description
1	Standard error of predicted vs true glandularity values (double fitting)
2	Dice similarity between true and predicted high-resolution image patches (CNN)
3	Glandularity of true images reconstructed at 273 μm vs high-resolution phantoms downsampled back at 273 μm
4	Glandularity of true images reconstructed at 194 μm vs high-resolution phantoms generated with a target voxel size of 194 μm
5	Assessment of the realism of the phantoms by the radiologist

Table 2.

Comparison of glandularity values between ten downsampled super-resolution phantoms (from 68 μm to 273 μm) and the corresponding original breast CT images reconstructed at 273 μm and segmented.

Breast CT Image	Original image (273 μm) [%]	Downsampled phantom (273 μm) [%]	Percentage difference [%]	Super-resolution phantom (68 μm) [%]
1	36.42	36.55	0.13	40.27
2	38.24	38.04	0.20	41.91
3	28.52	27.96	0.56	32.17
4	22.69	22.84	0.15	26.16
5	21.63	21.32	0.31	23.49
6	43.91	43.72	0.19	47.43
7	24.28	23.91	0.37	27.95
8	39.58	39.81	0.23	43.24
9	46.38	45.97	0.41	49.85
10	13.32	13.18	0.14	15.33

Table 3.

Comparison of glandularity values between ten original breast CT images reconstructed at 194 μm and segmented, and phantoms generated (from the same patient cases) from 273 μm to 194 μm .

Breast CT Image	Original image (194 μm) [%]	Corresponding super-resolution phantom (from 273 μm to 194 μm) [%]	Percentage difference [%]
1	37.83	37.91	0.08
2	40.13	40.04	0.09
3	29.59	29.53	0.06
4	24.08	24.12	0.04
5	22.57	22.49	0.08
6	45.72	46.01	0.29
7	26.02	25.87	0.15
8	41.12	40.83	0.29
9	47.38	46.98	0.4
10	14.01	13.96	0.05