Denoising Diffusion Probabilistic Models for Generating Tissue Type Breast Image Dataset

V. Khoshdel *, N. Abharian, A. Attar, and J. LoVetri
Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB R3T 5V6, Canada

Abstract

Various supervised deep learning techniques have been used to enhance microwave imaging for breast cancer detection, but such techniques require large amounts of labelled training data which remains a significant challenge. In this research, we addressed the scarcity of realistic numerical permittivity breast phantoms by developing a generative model. This model uses high-level annotations, including breast types defined by BI-RADS categories, to produce a large, labeled tissue-type breast image dataset. Our prototype model integrates a U-Net with a denoising diffusion probabilistic model to generate tissue-specific 2D sagittal cross-section images of the segmented breast. Initial results demonstrate the model's proficiency in generating realistic images that closely resemble segmented MRI scans, with quantitative assessments through Fréchet Inception Distance (FID) scores confirming their authenticity. The prototype model will be extended to 3D in future work.

1. Introduction

Microwave Imaging (MWI) attempts to reconstruct the internal structure of an object or region of interest using electromagnetic (EM) waves at microwave frequencies (300 MHz–30 GHz). It performs the reconstruction using measurements of the EM fields scattered by the target when illuminated by MW energy [10, 6]. MWI has been used for a broad range of applications, from medical imaging [1,2] and security [3] to industrial engineering [4], geophysical surveying [5,6] and stored-grain monitoring [7]. Quantitative MWI is performed by solving an associated inverse problem that is fundamentally ill-posed. Traditionally, nonlinear optimization techniques are used for reconstructing 3D images wherein the voxel value is representative of a particular physical property of the target. In medical quantitative MWI, two images are typically reconstructed related to the complex-valued permittivity which is known to be specific to different biological tissues as well as an indicator for the existence of disease [9]. Although MWI has lower resolution compared to many other medical imaging modalities, for some medical applications it can have higher tissue specificity [8].

Recently, Supervised Deep Learning (SDL) techniques have been used to overcome traditional MWI challenges and enhance imaging results [19, 10]. Our recent research on the use of SDL for EM imaging applications has focused on MW breast imaging [21-19, 20] and lower frequency EM imaging for stored grain monitoring [10]. As is true in all SDL research, these techniques require large amounts of training data which remains a significant challenge [20, 19]. Using relatively small data sets to train SDL models limits their robustness [21].

Researchers have used generative machine learning (GML) models to create large databases in other imaging modalities. The Generative Adversarial Neural Network (GAN) was one of the first GML approaches used in several medical imaging tasks [11-15]. Most recently, Diffusion Models (DM) have gained significant traction as a novel generative method, known for creating highly realistic natural images [16-17]. Unlike GANs and variational autoencoders, which can be challenging to train, interpret, and often yield unsatisfactory image quality—DMs are easily trainable, are based on solid analytical principles, and deliver top-tier image quality. Research indicates that for various image generation tasks, DMs outperform GANs and variational auto-encoders [18].

In the present work we take the first steps in addressing the lack of realistic numerical permittivity phantoms, especially phantoms containing tumors. A large dataset of quantitative breast images that provides the complex-valued permittivity at each voxel of an image is essential. In the past, researchers have relied on converting small databases of available MRI breast scans [14-15, 12]. As such, the conversion of the qualitative intensity in an MRI scan to a representative permittivity is well understood [12]. However, with this technique, the resulting MWI dataset will still be the size of the original MRI dataset and the issues of inadequate size and imbalanced data (e.g., under- and over-representation of phantoms with different types of tumors) will persist. To generate a large-scale and balanced MWI dataset, for the initial phase and as a demonstration of feasibility, we are investigating the development of generative models that use high-level annotations as input to create labeled synthetic MW breast phantoms. Using high-level quantitative annotations that describe breast phantoms based on BI-RADS classification (as a pre-condition for the generative process) has allowed us to generate novel 2D images that faithfully embody the characteristics specified in the annotations, thus allowing us to create a large labeled dataset.

While the ultimate aim of our research is to generate quantitative 3D numerical breast phantoms in terms of the complex-valued permittivity at each voxel, the initial study presented in this paper focuses on generating 2D images that are segmented according to tissue type. The plan is to extend this approach to that of producing 3D phantoms with complex-valued permittivity voxels.

2. Proposed Conditional Diffusion

Our generative model is a relatively complex system that integrates the architecture of a scaled-down U-Net with a Denoising Diffusion Probabilistic Model (DDPM) framework to generate tissue type images in an unsupervised fashion. The model is tailored for conditioning on a class label, taking inspiration from the techniques used in Classifier-Free Diffusion Guidance [24], which involves the incorporation of time-step embeddings, context embeddings, and the utilization of U-Net activations at a certain layer.
For this work, the condition imposed during image generation corresponds to BI-RADS categories for the breast—effectively corresponding to the ratio of dense to non-dense breast tissues. Our model incorporates this classification by identifying ranges for the relative proportion of fibroglandular to adipose (fat) tissues as a condition for image synthesis. This conditioning enables the model to generate images that not only exhibit variations in tissue composition but also are directly labeled with these percentages, thereby providing a dataset that is, by design, annotated with useful information relevant to breast tissue characterization. In addition to providing a basis for creating complex-valued permittivity phantoms that can be used for MWI, such a labeled dataset would be invaluable for subsequent analytical tasks, including training diagnostic algorithms or conducting detailed statistical analyses.

At the core of the model's training regimen is the iterative process of denoising, where the model learns to iteratively map a sample from random noise back to the data distribution. This is mathematically represented by a series of reverse diffusion steps, each characterized by a noise schedule that is learned during training. The guiding principle of this process is encapsulated in the formula:

$$x_{t-1} = \sqrt{\alpha_t} x_t + \sqrt{1 - \alpha_t}\epsilon \qquad (1)$$

Here, x_{t-1} represents the denoised image at the previous time step, x_t is the current image, and ϵ is the noise term. The coefficient α_t is a predetermined constant that dictates the noise level at each diffusion step.

To fine-tune the specificity and diversity of the generated images, a guidance scale ω has been employed, which directly influences the classifier-free guidance. The guidance scale is a crucial hyperparameter that controls the trade-off between the fidelity and variety of the generated images. This is governed by the following equation:

$$x_{t-1} = (1 + \omega)\sqrt{\alpha_t} x_t + \sqrt{1 - \alpha_t}\epsilon \qquad (2).$$

The presence of ω in this equation allows us to modulate the contribution of the class-conditional information during the generation process. By adjusting ω, we can produce images that are either more typical of a given class (at higher values of ω) or more diverse but less class-specific (at lower values of ω).

Training of the model is efficient, with the entire process being completed in approximately 200 epochs. This efficiency, coupled with the high quality of the generated images, confirms the model's potential for extensive applicability in generative tasks across different domains.

3. Network Training and Dataset

For our diffusion model architecture, we implemented all components in Python 3.6, utilizing Keras 2.0.6 with a TensorFlow backend to manage the neural network layers. The computational work was carried out on a Windows 10 system equipped with a Tesla P100-PCIE-12GB graphics processor and an Intel(R) CPU running at 3.50 GHz. To optimize the initialization of our convolutional layers, we employed the widely recognized Xavier initialization method, ensuring an appropriate scale of weights [46].

The model was trained using a batch size of 10 across 200 epochs, employing the Adam optimizer for efficient learning. To robustly evaluate our diffusion model, we adopted a four-fold cross-validation strategy, ensuring comprehensive testing across different dataset segments.

The I-SPY TRIAL database featuring MRI scans and corresponding tissue type images (TTIs) from 222 participants, with a total of 386,528 images was used [22,23]. After executing preprocessing steps that involved removing the chest wall and excluding the initial and final slices, the training dataset contains a total of 2175 images, 723 images in Class A, 784 in Class B, and 668 in Class C. To reduce computational costs, the image dimensions were downscaled to 64x64 pixels from their original size. As a result, the images produced also maintain a dimension of 64x64 pixels. It should be highlighted that, owing to insufficient samples for the fourth BI-RADS category, only the initial three breast phantom models are incorporated, namely Almost Entirely Fat, Scattered Fibroglandular Density, and Heterogeneously Dense.

3. Results

Figure 1 provides a visual comparison between real MRI segmented scans and synthetic images generated by our model across three distinct classes: Class A (Mostly Fat), Class B (Scattered Fibroglandular and Gladular), and Class C (Heterogeneously Dense). Each column represents a different class, while the rows differentiate between real and synthetic images. The synthetic images were generated to exemplify the characteristics of the corresponding real MRI scans, and, as depicted, there is a striking resemblance in terms of structural features between the synthetic and real images.

The top three rows exhibit random examples of real MRI segmented scans from our dataset, showcasing the inherent diversity within each class. The bottom three rows display the synthetic images generated by the model. The color scheme in these images is as follows: dark blue for air, light blue for fat, light green for skin, dark green for fibroglandular tissue, and yellow for identifying tumor regions. These images underscore the model's capability to reproduce both tumor-present and tumor-absent scenarios, as it has been trained on a varied set of examples. A notable observation is the model's ability to localize tumors predominantly within the fibroglandular regions, aligning with the established medical understanding. This indicates a successful learning outcome, as the model has not been explicitly labeled for tumor presence.

Although the current model does not differentiate between tumor and non-tumor images, future work could incorporate such annotations to refine the classification further. The potential for expanding the model's labeling to include the tumor’s presence will enhance the utility and accuracy of synthetic image generation for diagnostic purposes.

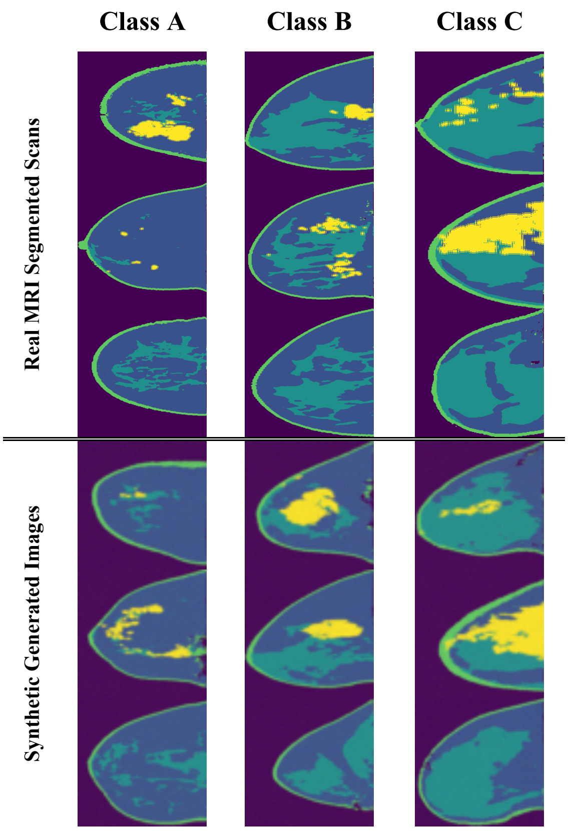

Figure 1. Comparative Analysis of MRI Segmented Scans and Machine-Generated Synthetic Images Across Three BI-RAD Classes.

In addition to the qualitative analysis, we performed a quantitative evaluation, employing the FID score to measure the authenticity of the generated images for each category. Table 1 presents this quantitative analysis for images from the three distinct classes. The diffusion-based model generated images with FID of 116.2, 120.7, and 120.6 for classes A, B, and C, respectively. For comparison, we created images using a region-growing algorithm to create fibroglandular and tumor regions within the breast. This involved selecting 40 images from each training class, preserving the fat areas, and generating fibro and tumor tissues starting from a central circular region. The initial regions are grown by adding pixels using a probabilistic procedure biased to the center-of-mass of the growing region. The result was a set of images with centrally-focused continuous tissues within the breast phantom. One sample image is shown in figure 2. The FID scores for these algorithmically synthesized images are 124.5, 141.5, and 146.4 FID for classes A, B, and C.

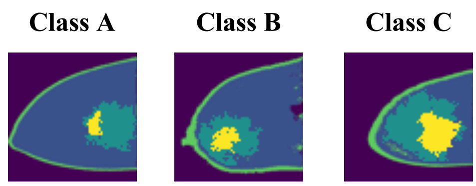

Figure 2. Representative Samples of Synthesized Images for Each Class, Generated Using the Algorithm-Based Technique.

These results demonstrate the effectiveness of our diffusion model in generating realistic images compared to the algorithm-based technique, as evidenced by the closer FID scores to the training data samples. It is noteworthy that while the algorithm-based technique retains the fat region from the original segmented MRI scans, potentially aiding in achieving lower FID values, the generative-based technique, which generates all four tissue types, results in even lower FID scores.

Table 1: Quantitative results of comparing different classes of datasets to the training dataset

	FID Score		
	Class A	Class B	Class C
Real MRI derived Tissue Type Scans	64.2	69.3	63.2
Diffusion-Based Synthesized Images	**116.2**	**120.7**	**120.6**
Algorithm-Based Synthesized Images	124.5	141.5	146.4

4. Conclusion

This study has illustrated the effectiveness of diffusion models in generating realistic tissue-type 2D breast images for addressing the problem of data scarcity. Our generative model out-performs an algorithm-based technique in terms of FID scores; the images are of a higher degree of similarity to actual MRI scans. This study highlights the potential of employing diffusion models to generate medical datasets and could be applicable to other scenarios where data is limited.

References

[1] P. M. Meaney, M. W. Fanning, Dun Li, S. P. Poplack, and K. D. Paulsen, “A clinical prototype for active microwave imaging of the breast," IEEE Transactions on Microwave Theory and Techniques, vol. 48, no. 11, pp. 1841-1853, 2000.

[2] A. H. Golnabi, P. M. Meaney, S. Geimer, and K. D. Paulsen, “Microwave imaging for breast cancer detection and therapy monitoring," in IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing Systems, 2011.

[3] D. M. Sheen, D. L. McMakin, and T. E. Hall, “Three-dimensional millimeter wave imaging for concealed weapon detection," IEEE Transactions on Microwave Theory and Techniques, vol. 49, no. 9, pp. 1581-1592, 2001.

[4] C. M. Bishop, Pattern Recognition and Machine Learning. Springer, 2006.

[5] Abubakar and P. M. Van den Berg, “Non-linear three-dimensional inversion of cross-well electrical measurements," Geophysical Prospecting, vol. 48, no. 1, pp. 109-134, 2000.

[6] I. Aliferis, C. Pichot, J.-Y. Dauvignac, and E. Guillanton, “Tomographic reconstructions of buried objects using nonlinear and regularized and inversion method," in Non-Linear Electromagnetic Systems, ser. Studies in Applied Electromagnetics and Mechanics. IOS Press, 2000,

[7] M. Asefi, I. Jeffrey, J. LoVetri, C. Gilmore, P. Card, and J. Paliwal, “Grain bin monitoring via electromagnetic imaging," Computers and Electronics in Agriculture, vol. 119, pp. 133-141, 2015.

[8] Kurrant, D., Omer, M., Abdollahi, N., Mojabi, P., Fear, E., & LoVetri, J. (2021). Evaluating Performance of Microwave Image Reconstruction Algorithms: Extracting Tissue Types with Segmentation Using Machine Learning. Journal of Imaging, 7(1), 5.

[9] Lazebnik M, McCartney L, Popovic D, Watkins CB, Lindstrom MJ, Harter J, Sewall S, Magliocco A, Booske JH, Okoniewski M, Hagness SC. A large-scale study of the ultrawideband microwave dielectric properties of normal breast tissue obtained from reduction surgeries. Phys Med Biol. 2007 May 21;52(10):2637-56.

[10] V. Khoshdel, “Using deep learning approaches for microwave imaging,” Ph.D. dissertation, Dept. Elect. Comp. Eng., Univ. of. Manitoba, MB, 2021.

[11] S. Kazeminia, C. Baur, A. Kuijper, B. van Ginneken, N. Navab, S. Albarqouni, A. Mukhopadhyay, “GANs for medical image analysis”. Artificial Intelligence in Medicine, 2020

[12] E. Omer and E. Fear, "Anthropomorphic breast model repository for research and development of microwave breast imaging technologies," Naure Scientific Data, vol. 5, no. 180257, 2018.

[13] E. Omer and E. Fear, "Automated 3D method for the construction of flexible and reconfigurable numerical breast models from MRI scans.," Med. Biol. Eng. Comput., no. 56, p. 1027–1040, 2018.

[14] Zastrow, E., Davis, S. K., Lazebnik, M., Kelcz, F., Van Veen, B. D., & Hagness, S. C. (2008). Development of anatomically realistic numerical breast phantoms with accurate dielectric properties for modeling microwave interactions with the human breast. IEEE Transactions on Biomedical Engineering, 55(12), 2792-2800.

[15] Zastrow, E., Davis, S. K., Lazebnik, M., Kelcz, F., Van Veen, B. D., & Hagness, S. C. (2007). Database of 3D grid-based numerical breast phantoms for use in computational electromagnetics simulations.

[16] O’Loughlin, D.; O’Halloran, M.J.; Moloney, B.M.; Glavin, M.; Jones, E.; Elahi, M.A. Microwave Breast Imaging: Clinical Advances and Remaining Challenges. IEEE Trans. Biomed. Eng. 2018, 65, 2580–2590. [CrossRef]

[17] Bolomey, J.C. Crossed Viewpoints on Microwave-Based Imaging for Medical Diagnosis: From Genesis to Earliest Clinical Outcomes. In The World of Applied Electromagnetics: In Appreciation of Magdy Fahmy Iskander;

[18] Pastorino, M. Microwave Imaging; John Wiley & Sons: Hoboken, NJ, USA, 2010.

[19] V. Khoshdel, A. Ashraf, M. Asefi, J. LoVetri. (2020). Full 3D Microwave Breast Imaging Using a Deep-Learning Technique. Journal of Imaging.

[20] K. Edwards, K. Krakalovich, R. Kruk, V. Khoshdel, J. LoVetri, C. Gilmore, I. Jeffrey. (2020). The implementation of neural networks for phase-less parametric inversion. Symposium of the International Union of Radio Science,

[21] V. Khoshdel, A. Ashraf, M. Asefi, J. LoVetri. (2021). A Multi-Branch Deep Convolutional Fusion Architecture for 3D Microwave Inverse Scattering: Stored Grain Application. Neural Computing and Applications.

[22] Newitt, D., Hylton, N., & the I-SPY 1 Network and ACRIN 6657 Trial Team. (2016). Multi-center breast DCE-MRI data and segmentations from patients in the I-SPY 1/ACRIN 6657 trials. The Cancer Imaging Archive.

[23] Hylton, N. M., Gatsonis, C. A., Rosen, M. A., Lehman, C. D., Newitt, D. C., Partridge, S. C., Bernreuter, W. K., Pisano, E. D., Morris, E. A., Weatherall, P. T., Polin, S. M., Newstead, G. M., Marques, H. S., Esserman, L. J., & Schnall, M. D. (2016). Neoadjuvant Chemotherapy for Breast Cancer: Functional Tumor Volume by MR Imaging Predicts Recurrence-free Survival—Results from the ACRIN 6657/CALGB 150007 I-SPY 1 TRIAL. Radiology, 279(1), 44-55.

[24] Ho, J., & Salimans, T. (2022). Classifier-Free Diffusion Guidance. arXiv. https://arxiv.org/abs/2207.12598