ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

Reconstruction of Mammography Projections using Image-to-Image Translation Techniques

Joana Cristo Santos[1] and Miriam Seoane Santos[2,3] and Pedro Henriques Abreu[1] *

1- University of Coimbra, CISUC, Department of Informatics Engineering, Coimbra 3030-290, Portugal

2- Laboratory of Artificial Intelligence and Decision Support (LIAAD – INESC TEC), Porto, Portugal

3- Department of Computer Sciences, Faculty of Sciences, University of Porto (FCUP), Porto, Portugal

Abstract. Mammography imaging is the gold standard for breast cancer detection and involves capturing two projections: mediolateral oblique and craniocaudal projections. The implementation of an approach that allows the acquisition of only one projection and reconstructs the other could mitigate patient burden, minimize radiation exposure, and reduce costs. Image-to-image translation has showcased the ability to generate realistic synthetic images in different medical imaging modalities which make these techniques a great candidate for the novel application in mammography. This study aims to compare five image-to-image translation approaches to assess the feasibility of reconstructing a mammography projection from its counterpart. The results indicate that ResViT shows the best overall performance in translating between both projections.

1 Introduction

Breast Cancer (BC) is one of the most prevalent cancers affecting women worldwide and remains the second leading cause of mortality among them [1, 2]. Mammography imaging, the gold standard for BC detection, requires capturing images in two projections: mediolateral oblique (MLO) and craniocaudal (CC) views per breast. To accurately diagnose BC, both MLO and CC projections are indispensable for clinical examination due to their complementary information. However, this procedure requires breast repositioning, a time-consuming process often associated with patient discomfort due to tissue recompression [3].

During the acquisition process, the integrity of information within one or both projections may be compromised [4, 5]. Image-to-image translation, proficient in transforming one image type into another, has presented the ability to generate realistic synthetic images in different medical imaging modalities which makes the hold promise in its application for mammography projection translation [6].

This study aims to investigate the feasibility of reconstructing the compromised projections, to potentially reduce the need for repeating the examination

*This work was funded by the FCT - Foundation for Science and Technology, I.P./MCTES through national funds (PIDDAC), within the scope of CISUC R&D Unit - UIDB/00326/2020 or project code UIDP/00326/2020. The work of Joana Cristo Santos was financially supported by the Portuguese Funding Institution Foundation for Science and Technology (FCT) under Grant 2020.05488.BD.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

and consequently, mitigate patient burden, minimize radiation exposure, and reduce costs. Five image-to-image translation approaches (pix2pix [7], CycleGAN [8], DiscoGAN [9], UNIT [10] and ResViT [11]) are compared to assess mammography reconstruction feasibility. Despite its potential, to the best of our knowledge no such approach has been applied to mammography for projection reconstruction. Two open-source datasets are included to demonstrate methodology versatility across contrasts and abnormalities. Performance evaluation metrics include mean absolute error (MAE), peak signal-to-noise ratio (PSNR), and structural similarity index (SSIM), revealing ResViT as the outperforming approach.

2 Related Work

Image-to-image translation approaches predominantly utilizing generative adversarial networks (GANs), have been extensively applied in medical imaging, including computed tomography (CT), positron emission tomography (PET), and magnetic resonance imaging (MRI) [12].

Armanious et al. [13] propose a novel conditional GAN architecture with non-adversarial losses for PET to CT translation, MRI motion correction, and PET denoising, achieving $SSIM = 0.92$ and $PSNR = 24.62$. Kong et al. [14] introduce a GAN capable of translating between T1 and T2 MRI for brain imaging, demonstrating slight performance improvements with $SSIM = 0.83$ and $PSNR = 24.00$. Liu et al. [15] enhance CycleGAN for MRI to CT translation, achieving impressive results of $MAE = 0.04$ and $PSNR = 39.10$. Moreover, Dalmaz et al. [11] present ResViT, a GAN architecture combining vision transformers with convolution operators for multimodal medical image synthesis, outperforming existing methodologies.

In mammography imaging, however, the focus has mainly been on generating synthetic images for dataset augmentation and contralateral mammographies [16, 17]. While some approaches have successfully generated realistic mammography images, none have explicitly addressed reconstructing a new mammography projection from another.

3 Methodology

In selecting approaches for this study, the aim is to encompass the main commonly employed approaches in other image modalities, namely pix2pix and CycleGAN, along with less commonly used such as DiscoGAN, UNIT, and ResViT. Pix2Pix [7] employs a conditional GAN architecture with a U-Net generator and PatchGAN discriminator, focusing on paired image translation. In contrast, CycleGAN [8] introduces cycle consistency loss for unpaired image translation, using ResNet generators and PatchGAN discriminators. Also designed for unpaired image-to-image translation, DiscoGAN [9] uses cycle consistency loss but with simpler generator and discriminator structures, incorporating an identity mapping loss and UNIT [10] employing cycle consistency and identity mapping losses alongside a multi-scale discriminator. ResViT [11], designed for medical image synthesis, combines residual connections and visual transformers, featuring aggregated residual transformer blocks for diverse information blending and a unified implementation for various source-target modalities.

The used mammography images originate from two distinct image datasets: the Cohort of Screen-Aged Women (CSAW) [18] and Curated Breast Imaging Subset of DDSM (CBIS-DDSM) [19]. The pairs of CC and MLO projections for each breast are selected from each dataset and in total, the CSAW dataset has 1,000 image pairs and the CBIS-DDSM dataset has 1,216 image pairs. Prior to analysis, pre-processing involved background removal, resizing to a standardized size of 256x256 pixels, and normalization.

The translation capability is evaluated in both MLO to CC and CC to MLO directions. The datasets are randomly split into training (80%) and testing (20%) set for each run, with the experimental setup executed 30 times to mitigate bias. The five image-to-image translation approaches are adapted to train and test on the mammography grayscale images and fine-tuned to the two image datasets [1]. The evaluation is performed considering visual comparison and the evaluation metrics MAE, PSNR, and SSIM. A paired t-test at an alpha value of 0.05 is used to test for a statistically significant difference.

4 Results

4.1 Visual Comparison

The visual results for both datasets are depicted in Figures 1 and 2.

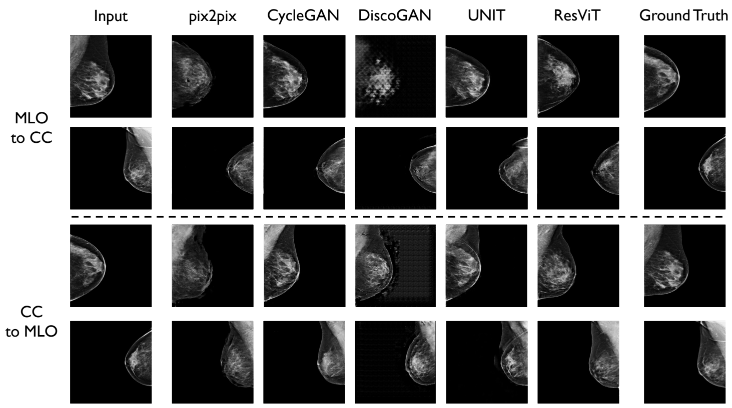

Fig. 1: Visual Comparison of the results for the CSAW dataset.

The approaches demonstrate the capability to produce realistic mammography projections. DiscoGAN performed poorly, producing distorted images

[1] Code available at `https://github.com/joanacsantos/MammographyReconstruction`

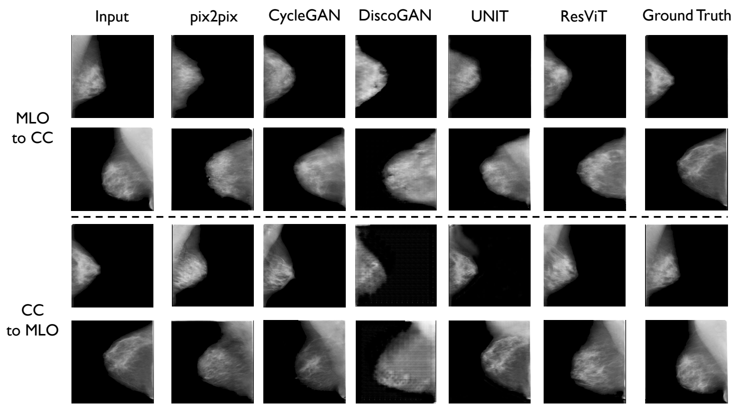

Fig. 2: Visual Comparison of the results for the CBIS-DDSM dataset.

with noticeable noise. Pix2pix and UNIT struggle with consistency, leading to misshapen breast contours. In contrast, both CycleGAN and ResViT excelled, accurately capturing the intricate shape and texture of breast tissue.

Additionally, significant disparities are observed between the CSAW and CBIS-DDSM datasets, with CSAW images exhibiting higher contrast and more intricate details compared to CBIS-DDSM. Despite these differences, all approaches demonstrated versatility in handling mammography images with diverse characteristics.

4.2 Evaluation Metric Results

The analysis of evaluation metrics for both datasets is summarized in Table 1.

Table 1: Evaluation Metrics Values for the CSAW and CBIS-DDSM Datasets. The best metric values for each dataset are highlighted in bold.

		MLO to CC			**CC to MLO**		
		MAE ↓	PSNR ↑	SSIM ↑	MAE ↓	PSNR ↑	SSIM ↑
CSAW Dataset	pix2pix	**0.035**	**21.506**	0.710	0.049	19.754	0.644
	CycleGAN	0.043	19.969	0.694	0.056	18.589	0.645
	DiscoGAN	0.042	20.133	0.677	0.078	17.695	0.378
	UNIT	0.045	19.684	0.685	0.058	18.520	0.622
	ResViT	0.038	20.654	**0.718**	**0.049**	**19.763**	**0.659**
CBIS-DDSM Dataset	pix2pix	0.060	18.541	0.739	**0.076**	**17.254**	**0.685**
	CycleGAN	0.068	17.579	0.724	0.083	16.545	0.679
	DiscoGAN	0.075	16.964	0.678	0.116	14.722	0.355
	UNIT	0.076	16.824	0.702	0.091	15.856	0.654
	ResViT	**0.058**	**18.577**	**0.741**	0.078	17.135	0.682

Results show similarity in performance for both translation directions:

- For translation from MLO to CC projection, pix2pix performs better on the CSAW dataset, while ResViT excels on the CBIS-DDSM dataset. However, statistical tests reveal no significant performance difference between pix2pix and ResViT in the CSAW dataset for MAE and SSIM metrics and in the CBIS-DDSM dataset.
- For the translation from CC to MLO projection, ResViT achieves the best results for the CSAW dataset, while pix2pix performs better for the CBIS-DDSM dataset. Again, statistical tests show no significant performance disparity between pix2pix and ResViT in the CSAW dataset for PSNR metrics and in the CBIS-DDSM dataset for the PSNR and SSIM metrics.

Consequently, due to the lack of statistical relevance in both translation directions, relying solely on evaluation metrics fails to definitively establish a superior approach between pix2pix and ResViT. Upon visual examination, pix2pix exhibits deficiencies in image structure, leading to the determination that ResViT emerges as the most effective methodology for the translation of both directions.

Overall, evaluation metrics indicate lower performance in reconstructing MLO from CC projection, possibly due to missing vital information in the CC projection, such as pectoral muscle location, crucial for accurate MLO reconstruction. These findings suggest a greater potential for further exploration of the opposite translation, offering promising avenues for future research.

5 Conclusions

This work presents a comprehensive study comparing five image-to-image translation approaches, aimed at assessing the feasibility of reconstructing a mammography projection from its counterpart. Identifying an effective approach holds the potential to alleviate patient burden, reduce costs, and minimize radiation exposure. The results show that ResViT demonstrates superior performance in translating both directions, with metrics indicating better performance in reconstructing MLO from CC projection. Future work will focus on further refining the reconstruction of MLO from CC projection, prioritizing enhancements in breast shape and texture to validate its clinical applicability.

References

[1] Hyuna Sung, Jacques Ferlay, Rebecca L Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal, and Freddie Bray. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a Cancer Journal for Clinicians*, 71(3):209–249, 2021.

[2] Joana Cristo Santos, Miguel Henriques Abreu, Miriam Seoane Santos, Hugo Duarte, Tiago Alpoim, Inês Próspero, Susana Sousa, and Pedro Henriques Abreu. Bone Metastases Detection in Patients with Breast Cancer: Does Bone Scintigraphy Add Information to PET/CT? *The Oncologist*, 28(8):e600–e605, 2023.

[3] Carlos Gutierrez, Alyssa Owens, Lori Medeiros, Donnette Dabydeen, Nithya Sritharan, Pradyumna Phatak, and Satish G Kandlikar. Breast cancer detection using enhanced IRI-numerical engine and inverse heat transfer modeling: model description and clinical validation. *Scientific Reports*, 14(1):1–17, 2024.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

[4] Ricardo Cardoso Pereira, Joana Cristo Santos, José Pereira Amorim, Pedro Pereira Rodrigues, and Pedro Henriques Abreu. Missing Image Data Imputation using Variational Autoencoders with Weighted Loss. In *ESANN*, pages 475–480, 2020.

[5] Joana Cristo Santos, Pedro Henriques Abreu, and Miriam Seoane Santos. The identification of cancer lesions in mammography images with missing pixels: analysis of morphology. In *2022 IEEE 9th International Conference on Data Science and Advanced Analytics (DSAA)*, pages 1–8. IEEE, 2022.

[6] Shizuo Kaji and Satoshi Kida. Overview of image-to-image translation by use of deep neural networks: denoising, super-resolution, modality conversion, and reconstruction in medical imaging. *Radiological Physics and Technology*, 12(3):235–248, 2019.

[7] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1125–1134, 2017.

[8] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *IEEE International Conference on Computer Vision*, pages 2223–2232, 2017.

[9] Taeksoo Kim, Moonsu Cha, Hyunsoo Kim, Jung Kwon Lee, and Jiwon Kim. Learning to discover cross-domain relations with generative adversarial networks. In *International Conference on Machine Learning*, pages 1857–1865. PMLR, 2017.

[10] Ming-Yu Liu, Thomas Breuel, and Jan Kautz. Unsupervised image-to-image translation networks. *Advances in Neural Information Processing Systems*, 30, 2017.

[11] Onat Dalmaz, Mahmut Yurt, and Tolga Çukur. ResViT: Residual vision transformers for multimodal medical image synthesis. *IEEE Transactions on Medical Imaging*, 41(10):2598–2614, 2022.

[12] Junhua Chen, Shenlun Chen, Leonard Wee, Andre Dekker, and Inigo Bermejo. Deep learning based unpaired image-to-image translation applications for medical physics: a systematic review. *Physics in Medicine & Biology*, 2023.

[13] Karim Armanious, Chenming Jiang, Marc Fischer, Thomas Küstner, Tobias Hepp, Konstantin Nikolaou, Sergios Gatidis, and Bin Yang. MedGAN: Medical image translation using GANs. *Computerized Medical Imaging and Graphics*, 79:101684, 2020.

[14] Lingke Kong, Chenyu Lian, Detian Huang, Yanle Hu, Qichao Zhou, et al. Breaking the dilemma of medical image-to-image translation. *Advances in Neural Information Processing Systems*, 34:1964–1978, 2021.

[15] Yanxia Liu, Anni Chen, Hongyu Shi, Sijuan Huang, Wanjia Zheng, Zhiqiang Liu, Qin Zhang, and Xin Yang. CT synthesis from MRI using multi-cycle GAN for head-and-neck radiation therapy. *Computerized Medical Imaging and Graphics*, 91:101953, 2021.

[16] Juhun Lee and Robert M Nishikawa. Simulating breast mammogram using conditional generative adversarial network: application towards finding mammographically-occult cancer. In *Medical Imaging 2020: Computer-Aided Diagnosis*, volume 11314, pages 291–297. SPIE, 2020.

[17] Annie Julie Joseph, Priyansh Dwivedi, Jiffy Joseph, Seenia Francis, PN Pournami, PB Jayaraj, Ashna V Shamsu, and Praveen Sankaran. Prior-guided generative adversarial network for mammogram synthesis. *Biomedical Signal Processing and Control*, 87:105456, 2024.

[18] Karin Dembrower, Peter Lindholm, and Fredrik Strand. A multi-million mammography image dataset and population-based screening cohort for the training and evaluation of deep neural networks-the cohort of screen-aged women (CSAW). *Journal of Digital Imaging*, 33(2):408–413, 2020.

[19] Rebecca Sawyer Lee, Francisco Gimenez, Assaf Hoogi, Kanae Kawai Miyake, Mia Gorovoy, and Daniel L. Rubin. A curated mammography data set for use in computer-aided detection and diagnosis research. *Scientific Data*, 4:170177, 2017. Available at https://doi.org/10.7937/K9/TCIA.2016.7O02S9CY.