JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Pix-2-Pix Neural Network Proposal for Automatic Breast Images Contrast Correction

[1]Anna Carolina Falanga Brenha, [1]Pedro Cunha Carneiro, [1]Ana Claudia Patrocínio
[1,2,3] Faculty of Electrical Engineering
[1,2,3] Federal University of Uberlândia, Uberlândia, Brazil

Abstract: In a scenario where breast cancer is still a worldwide challenge to healthcare professionals, with more than half a million deaths registered annually, the early detection of the disease faces a range of obstacles. The need of great image contrast quality, often associated with high radiation doses, is opposed by the constant efforts to reduce radiation absorbance for patients and technicians during breast imaging exams, as outlined in ALARA (As Low As Reasonably Achievable) guidelines. The present study proposed a Pix-2-Pix neural network implementation as an image contrast enhancement tool, projected to reproduce high-dose contrast quality in low-dose images, implementing the necessary conditions for proper diagnosis while preserving ALARA standards for patient and technician safety. The Carneiro Contrast Index (CCI) was employed for quantitative assessments, demonstrating significant contrast improvement without additional radiation. Also, the results successfully validate this technique's potential to support safer breast cancer screening clinical practices. Further works include different datasets validation and optimization of other learning parameters.

IndexTerms **– 2D Mammogram, Artificial Intelligence Breast Cancer, Image Contrast, Neural Networks**

I. INTRODUCTION

According to the World Cancer Research Fund (WCRF) [1], 2.296.840 new breast cases were globally registered in 2022, making it the most frequent cancer among women, and the second most frequent cancer worldwide. In terms of mortality rates, countries such as China, India, and the United States register the majority of the 666.103 deaths registered worldwide. In 2024, 310.720 women were diagnosed with invasive breast cancer only in the United States, with 16% of cases occurring in women under 50 years old [2]. This numbers reflect the still present diagnostic difficulties faced by healthcare professionals, that often lead to late disease detection, especially in younger woman [3].

Even after almost a century of the invention of the 2D mammogram, the current gold standard exam for breast cancer detection, breast anatomy is still a key challenge for radiographic imaging [4]. Two main tissues constitute the breast structure, the fat tissue, responsible for its overall form, the fibroglandular tissue, that is significant more present in younger individuals and forms the structures where cancerous lesions usually grow, such as the lobules, ducts, and Cooper's ligaments [5].

Concerning radiographic images production, the X-Ray dose is one of the principal factors to be considered, once the high exposure to radioactivity has significant long-term impact to the patients and technicians, such as cellular damage, DNA damage, weakened bones and other chronic conditions [6]. In the 2D mammogram acquisition process, the fat breast tissue tends to mitigate a significant amount of the received radiation, resulting in darker gray shades on the final image, while the fibroglandular tissues tend to mitigate very few of the radiation amount, resulting in a limited amount of lighter gray shades to represent the range of fibroglandular structures [7].

The mammogram spatial contrast refers to the exam's capacity to conscientiously represent subtle structure differences and it its directly related to the X-Ray dose amount received during the exam, resulting in lower spatial contrast quality when lower dose exposures are used. It plays a massive role on the fibroglandular structures differentiation in the image, once a cancerous lesion can be easily lost or unrecognized among the other breast fibroglandular structures that attenuate similar radiation amounts [8].

The ALARA (As Low As Reasonably Achievable) Principle [9] is a fundamental guideline for medical imaging, and establishes norms and protocols that aim to minimize radiation exposure to all individuals involved in the radiographic image acquisition procedure, taking social, economic, and public policies considerations into account. In this scenario, the safety of patients and technicians can be compromised by human errors during the ALARA protocols implementation, as in the unuse of appropriate shielding and lack of constant mammograph calibration.

As a solution to increase spatial contrast on 2D mammograms without increasing the X-Ray dose exposure, already subject to increases caused by the unfollowing of the ALARA Principle guideline, Artificial Intelligence systems such as Neural Networks are presented as a promising alternative, once its capacity to process a massive amount of data without being affected by the human error factor.

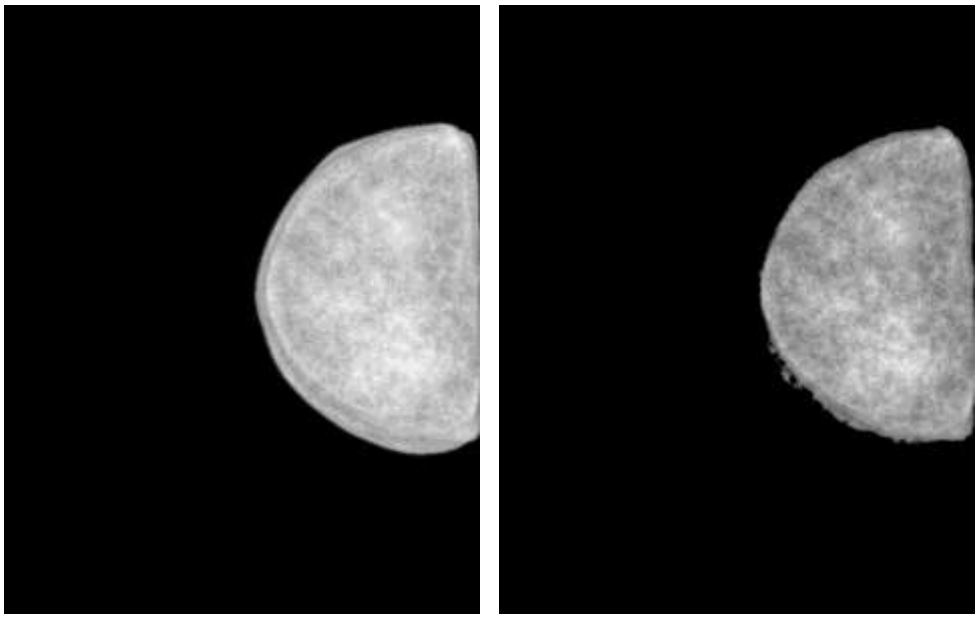
II. Materials and Methods

Dataset

The chosen dataset [9] included 35 DICOM images of a clinically validated breast phantom in 12 different fat-tissue arrangements, acquired from a *General Electric Senographe Essential* equipment. For 11 out of the 12 arrangements, the initial dataset presented 3 images in different acquisition settings: low-dose, with an average entrance dose of 3.359 mGy, automatic dose, with an average entrance dose of 8.407 mGy, and high dose, with an average entrance dose of 11.523 mGy. Once the remaining configuration only featured low-dose and automatic dose images it was removed from the final dataset.

From the filtered dataset, the 33 images were split into two X-Ray doses groups: Group A, containing the low-dose image of each configuration, and Group B, containing the high-dose version of the same phantom arrangement. All the automatic-dose images were discarded, resulting in a final dataset of 22 elements. Figure 1 exemplifies a low dose and high dose image pair.

Figure 1: (a) Phantom arrangement in low dose configuration; (b) Phantom arrangement in high dose configuration;



(a) (b)

The exclusion criteria aimed to align the final results with the ALARA principle, in order to produce a high-contrast image from the lowest dose possible.

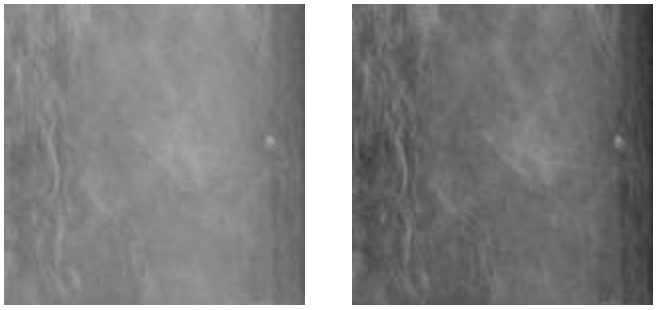
Data Pre-Processing Protocol

In order to better train network models, data augmentation techniques are often implemented [10]. For this study, the filtered dataset was submitted, through an authorial Python script, to a patch extraction procedure that cropped the phantom images into 150x150 patches and transformed them into PNG format. After the exclusion of the elements that display mostly X-Ray image background, the dataset resulted in 1504 patches, 752 from each dose group.

After the data augmentation, 188 images were randomly sorted from each group in order to separate the testing set from the training set that was further presented to the network. The training images were sorted with the support of a second authorial Python script that assured that both group A testing set and group B testing set received the exact same path, but in different dose levels.

Furthermore, the training datasets were established as: Training Dataset A, containing 564 patches in low dose mode, and Training Dataset B, containing the same 564 patches but in high dose mode, while the testing datasets were defined as: Testing Dataset A, containing 188 patches in low dose mode and Testing Dataset B, containing the same patches in high-dose mode. Figure 2 illustrates an example of low dose and high dose patch pair.

Figure 2: (a) Patch in low dose configuration; (b) Patch in high dose configuration;



(a) (b)

Pix-2-Pix Network Model

The network used in the present study is based on a pre-trained model [11][12], which suffered the required adaptations in order to better suit the implementation purpose.

The Pix2Pix approach was chosen due its high efficiency in generative systems such as image-to-image translation [12], since it is composed by a Generative Adversarial Network (GAN) that, in this case, aims to learn how to translate the low dose patch into its correspondent high dose pair, and a Discriminator Network, that tries to tell apart the real low dose images from the fake ones. Based on the analysis of the loss function behavior for both networks for different pre-selected learning hyperparameters, the network improvement was constructed. Table 1 shows the role of each loss function, pre-determined and adopted from the original pix-2-pix model, during the network training.

Table 1: Impact of the loss function on the network training

Loss Function	Target	Description
G_GAN	GAN	Measurement of the GAN's ability to produce fake images that will successfully be classified as real by the Discriminator.
G_L1	GAN	Numeric representation of the pixel map difference between a real and a fake patch.
D_real	Discriminator Network	Measurement of the Discriminator's capacity of labeling the high-dose patches from the dataset as real.
D_fake	Discriminator Network	Measurement of the Discriminator's capacity of labeling the high-dose patches created by the GAN as fake.

Table 2 describes the selected hyperparameters, that aimed to tune finely both networks in order to create the desirable prominence of the generator over the discriminative network.

Table 2: Evaluated Hyperparameters

Hyperparameter	Target	Description
Learning rate	GAN	Defines the
Epochs	GAN	Amount of times the network will look at the presented dataset.
Lambda L1	GAN	
GAN Mode	GAN	Adversarial loss function used in the GAN's training.
GAN Filters	GAN	Number of filters, or characteristics map, in the generator.
Discriminator Filters		Number of filters, or characteristics map, in the discriminator. Determines the discriminator's capacity to distinguish real and fake images.
Discriminator Layers	Discriminator Network	Number of layers in the discriminator. Tends to influence its capacity
Initial Type	GAN and Discriminator Network	Affects model's convergency time and quality.

Carneiro Contrast Index (CCI)

The Carneiro Contrast Index (CCI) [13] was implemented as a quantitative image evaluation metric, in order to describe and compare contrast behavior along the network training process.

CCI is a local contrast metric analysis, created and validated for Breast Tomosynthesis analysis and 2D Mammography images, that consists in a 3x3 spatial filter based on standard deviation calculation. After the normalization of the target image, every pixel value is replaced for its kernel's standard deviation value, and, once the standard deviation matrix is produced, its mean value is calculated in order to obtain the final CCI result. Since the metric was originally developed for 12-bit image analysis, the normalization formula was adapted to properly fit the 8-bit dataset. For group A full-images, the CCI mean was 1.995, while for group B the mean was 7.410, confirming the dose-contrast relation. Also, the metric was further implemented in the final network outputs, aiming to reinforce quantitative validation.

III. Results and Discussion

With the prepared dataset, the network was trained in different learning conditions. The starting point was established with the training of the model under its original parameters, described in Table 3. The obtained loss function results, shown in Table 4, reveal considerable balanced high values between D_real and D_fake, suggesting that the network is not being able to tell which image is fake, but neither which image is real. Also, the high values in G_GAN and G_L1 shows that the generator is having difficulties in tricking the discriminator, and the pixel maps from the real and fake images are not properly converging. Figure 3 graphically represents the obtained results, with a CCI value of 0.696 for image (a), 1.024 for image (b), and 0.899 for image (c).

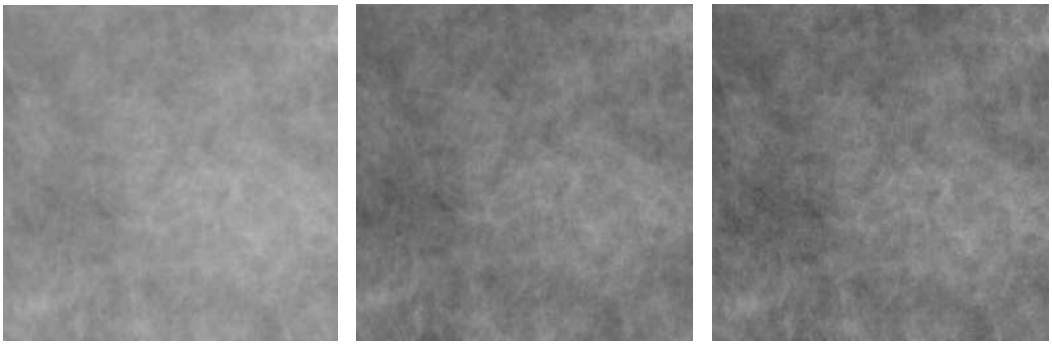
Table 3: Initial hyperparameters values.

Hyperparameter	Value
Learning rate	0.0002
Epochs	200
Lambda L1	100
GAN Mode	"vanilla"
GAN Filters	64
Discriminator Filters	64
Initial Type	"normal"

Table 4: Initial version loss function results.

Loss Function	Value
G_GAN	1.735
G_L1	3.371
D_real	0.379
D_fake	0.284

Figure 3: (a) Real input example; (b) Real output example; (c) First run's generated output example;



(a) (b) (c)

For the next runs, the initial settings were adjusted aiming to induce an unbalance between the discriminator's loss functions, once low values for D_real and higher values for D_fake suggests that the discriminator is labeling real images correctly, but is being successfully tricked by the generator, once it cannot tell fake images apart from the real ones. Furthermore, the generator's parameters were also tuned, pursuing better conversion values to the pixel maps.

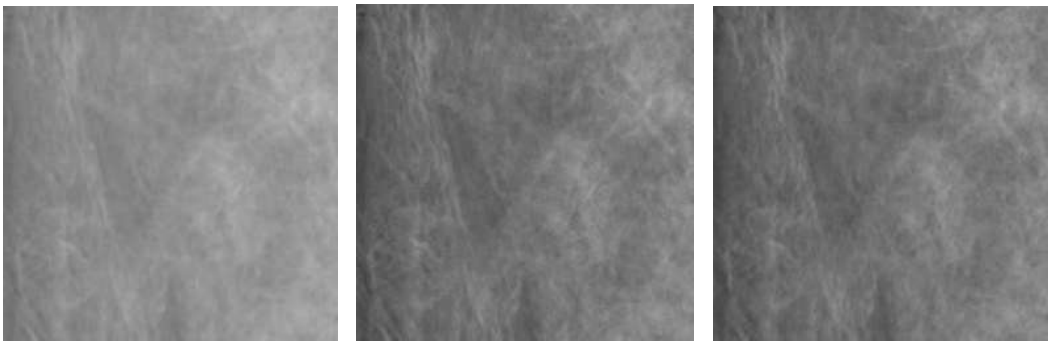
In the second version of training, the Lambda L1 value was reduced to 80 and the epochs number was increased to 300. The results in this version, shown in table 5, indicates a slight improvement of the pixel maps conversion, however, the increase in G_GAN and the non-significant changes in D_real and D_fake suggests an undesirable growth in the discriminator's capacity.

Table 5: Second version loss function results.

Loss Function	Value	Individual Performance (compared to last version)
G_GAN	1.842	Decreased
G_L1	3.248	Increased
D_real	0.182	Increased
D_fake	0.223	Decreased

Figure 4 describes the graphic results on the second run, reinforcing the previous insights obtained by the loss functions values, such as the subtle pixel maps convergence improvement, but highlighting a considerable visual image quality loss. The CCI values were 0.692 for image (a), 0.992 for image (b), and 1.016 for image (c).

Figure 4: (a) Real input example; (b) Real output example; (c) Second run's generated output example;



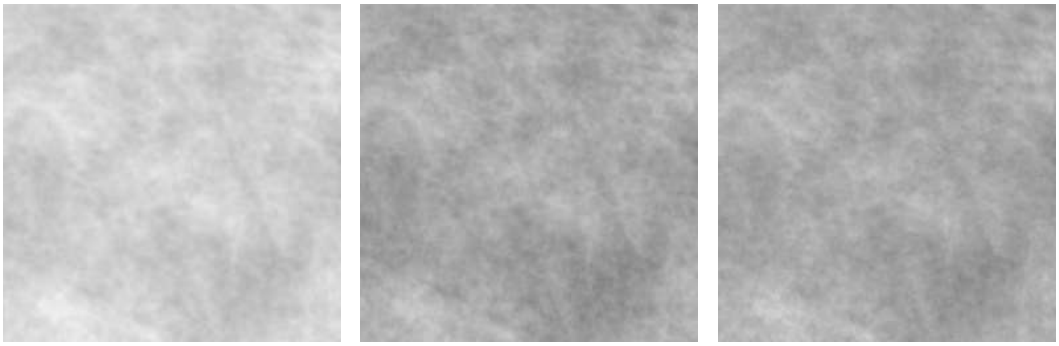
(a) (b) (c)

For the third run, the previous changed parameters were set to their original configuration, and, in addition, the initial mode was set as "xavier", that defines the layers weights based on the entrance and exit numbers in each layer, and the GAN mode was set as "lsmode", that substitutes the traditional binary loss function, as in "vanilla", with a least-squares based one. The results described in Table 6 show a significantly positive impact of the GAN mode change on G_GAN, but a subtle increase in G_L1, indicating that the network maintained a similar difficulty to make the pixel maps converge. Visually, Figure 5 reveals the improvements observed on the loss functions, with a more precise fine details positioning, but still lacking pixel mapping precision. The CCI values obtained on this run were 0.629 for image (a), 0.826 for image (b), and 0.861 for image (c).

Table 6: Third version loss function results.

Loss Function	Value	Individual Performance (compared to last version)
G_GAN	0.480	Increased
G_L1	3.357	Decreased
D_real	0.025	Increased
D_fake	0.108	Decreased

Figure 5: (a) Real input example; (b) Real output example; (c) Third run's generated output example;



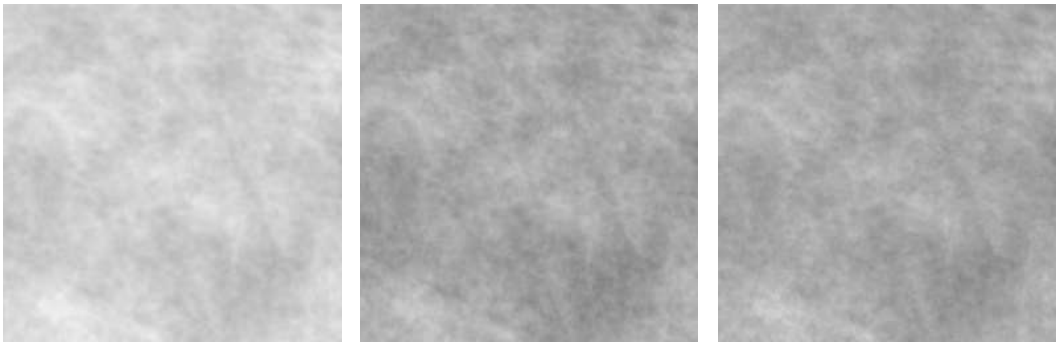
(a) (b) (c)

The model's peak performance was obtained on the fourth run, with the increasing of layers in the generator to 96, and the decreasing of layers in the discriminator to 32. The results presented in Table 7 reveal a stable balance between the loss functions, with the decreasing of G_GAN and G_L1, that implies in a considerable improvement of the pixel maps convergence and ability of the generator to trick. Also, the decreasing in D_real along with the increasing in D_fake shows that the main goal of the training process was successfully achieved. Finally, Figure 6 graphic represents the run's results, where the obtained CCI values were 0.964 for image (a), 1.181 for image (b), and 1.310 for image (c).

Table 7: Fourth version loss function results.

Loss Function	Value	Individual Performance (compared to last version)
G_GAN	0.425	Increased
G_L1	2.535	Increased
D_real	0.041	Increased
D_fake	0.152	Decreased

Figure 6: (a) Real input example; (b) Real output example; (c) Third run's generated output example;



(a) (b) (c)

IV. Conclusion

In fine, the obtained results were able to validate the Pix-2-Pix model as a promising technique to automatically improve 2d mammogram images, and a potential ally to reinforce the ALARA guideline, once it was able to implement a high dose contrast pattern in low dose the presented breast phantom patches dataset. The present study was able to prove the significant contribution of the number of layers on both generative and discriminative networks, that were satisfactorily balanced and promoted the generator's growth without compromising the pixel map similarity. Furthermore, the CCI demonstrated to be a key quantitative comparison metric to complement the qualitative analysis provided by both networks' loss functions.

In order to improve the system's quality and preparate it to further clinical trials, it is still necessary to validate the technique on real breasts dataset, as well as investigate other learning settings that may improve even more the goaled balance between the loss functions. Further studies on the subject will aim to improve the network tuning with the evaluation of other loss metrics and hyperparameters, as well as to validate the model on entire 2D breast images.

References

[1] *Breast Cancer Statistics*. (2022). World Cancer Research Fund (WCRF). https://www.wcrf.org/preventingcancer/cancerstatistics/breast-cancer-statistics/

[2] *Breast Cancer Facts and Statistics*. (2022). BreastCancer.Org. https://www.breastcancer.org/facts-statistics

[3] Partridge AH, Hughes ME, Ottesen RA, Wong YN, Edge SB, Theriault RL, Blayney DW, Niland JC, Winer EP, Weeks JC, Tamimi RM. The effect of age on delay in diagnosis and stage of breast cancer. Oncologist. 2012;17(6):775-82. doi: 10.1634/theoncologist.2011-0469. Epub 2012 May 3. PMID: 22554997; PMCID: PMC3380876.

[4] Vedantham S, Karellas A. Breast Cancer Screening: Opportunities and Challenges with Fully 3D Tomographic X-Ray Imaging. Bridge (Wash D C). 2022 Spring;52(1):33-42. Epub 2022 Mar 28. PMID: 35431425; PMCID: PMC9012464.

[5] Lee CI and Elmore JG. Chapter 10. Breast cancer screening, in Harris JR, Lippman ME, Morrow M, Osborne CK. Diseases of the Breast, 5th edition. Lippincott Williams and Wilkins, 2014.

[6] *AAPM Position Statements, Policies and Procedures - details*. (2023, November). AAPM Position Statement on Radiation Risks From Medical Imaging Procedures. https://www.aapm.org/org/policies/details.asp?id=3615

[7] Barnes G T. Contrast and scatter in x-ray imaging Radiographics. :307- 23.

[8] Fico N, Di Grezia G, Cuccurullo V, Salvia AAH, Iacomino A, Sciarra A, Gatta G. Breast Imaging Physics in Mammography (Part I). Diagnostics (Basel). 2023 Oct 17;13(20):3227. doi: 10.3390/diagnostics13203227. PMID: 37892053; PMCID: PMC10606465.

[9] Sousa, M. A. Z., Matheus, B. R. N., & Schiabel, H. (2018). *Development of a structured breast phantom for evaluating CADe/Dx schemes applied on 2D mammography. Biomedical Physics & Engineering Express, 4(4), 045018.*
[10] Hernández-García, A., & König, P. (2018). *Further Advantages of Data Augmentation on Convolutional Neural Networks. Lecture Notes in Computer Science, 95–103.* doi:10.1007/978-3-030-01418-6_10

[11] Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*.

[12] Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-Image Translation with Conditional Adversarial Networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[13] Carneiro P.C., et al. Proposta de metodologia para realce de contraste em imagens de mamas densas utilizando decomposição multi-escala com transformada discreta wavelet, 2019.