Applying a Conditional GAN for Bone Suppression in Chest Radiography Images

Hugo Eduardo Ziviani[2], **Guillermo Cámara Chávez**[2], **Mateus Coelho Silva**[2]

[1] Brazil - Minas Gerais - Ouro Preto

[2]Universidade Federal de Ouro Preto (UFOP) – Campus Morro do Cruzeiro

{hugo.ziviani, mateus.silva1}@aluno.ufop.edu.br, guillermo@ufop.edu.br

Keywords: Bone Suppression; Generative Adversarial Networks; cGAN's.

Abstract. *Bone suppression in radiography is a suitable technique to evaluate the health of soft tissues in exams. For instance, these techniques are essential in evaluating chest radiography images during the COVID-19 outbreak. The purpose of this work is to propose an alternative to solve the bone suppression task in chest radiography images using Generative Adversarial Networks (GANs). Specifically, we used a conditional GAN type (cGAN) to provide a bone-suppressed version of the initial image. To quantify the results, it was necessary to review the main metrics and some state-of-the-art papers related to ours. We compared our result to works from the literature that used the same dataset as the proposal or related techniques. The most used dataset was the Japanese Society of Radiological Technology (JSRT) in these works. With this set of images, we reached a PSNR index of 34.96, which was better than that reviewed in the literature, and a similarity coefficient, known as SSIM, of 0.94. As for the loss calculated by MS-SSIM, we obtained the lowest compared to the reviewed works.*

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001

1. Introduction

The computer's power and advanced methods are growing, and it is helping daily of various professionals on their tasks. The techniques such as optimized algorithms and processing distributed are advancing, and they can be used to auxiliary on difficult and tiring human functions. Sometimes a professional spend hours making analyses and diagnoses of medical exams. The technology and computer-aided systems can help these professionals spend some computer processing.

According to [Wang et al. 2019] a challenge for medicine and computer-aided diagnosis is to make the lung analysis. Detect any disease, for example, pneumonia, tumors, or the evolution of any illness, can be made better without bones shadow. In particular, for the soft tissue analysis, it would be ideal if all bones could be suppressed in an image. Carestream company [Matters 2014], a company on the medical image analysis, showed up the benefits to analysis an image without bones to detect and observe every possible lesion or disease on their White Paper.

With the growth of Graphic Process Unity power (GPU), techniques using Machine Learning and Deep Learning are being very frequently studied and researched to solve medical problems such as [Rajaraman et al. 2021a, Sujath et al. 2020, Oliveira et al. 2021]. The bone shadow elimination by computer processing is a way not to expend much money buying dual-energy hardware and is an alternative to update single energy equipment to do the task. A study as [Juhász et al. 2010] relates an experiment using image processing in a GPU and an object detection approach to eliminate the shadows. We can use artificial intelligence techniques, such as deep learning, as it is a specific task. In this way, a more recent work solved the task in [Gusarev et al. 2017], which uses deep learning models to clean the soft-tissue image.

The main objective of this study is to use deep learning techniques to generate chest images without bones using Adversarial Networks and some computer image processing. Specifically, we perform this task using Conditional Generative Adversarial Networks (CGANs). As we could observe on the related works, we can prove the necessity of systems to clean the chest radiography to support the final diagnosis towards a model that can be improved to be used in any computer-aided diagnosis (CAD) system. The final model will make bone suppression to other applications that use the lung image.

For this matter, in Section 2, we display the main works from the literature that relates to ours. Section 3 displays the whole materials and methods used in the process of performing this study. In Section 4, we display and discuss the results obtained from the experiments. Finally, we discuss our results and conclusions in Section 5.

2. Related Works

For this paper, we looked for research studies and articles that assess bone shadow elimination or bone suppression. A range of possibilities was found using dual-energy subtraction, deep learning, auto-encoder, convolutional networks, and adversarial networks. The research was made looking for works with a considerable result and accuracy and a scenario similar to ours that allowed us to verify the obtained results according to the dataset used. We considered the results that preserved the maximum of the features presented in the soft-tissue lung part.

The study made by [Gusarev et al. 2017] was made with a non-identified dataset. The dataset was composed of 35 images from different sources. From this 35, they generated an augmented dataset with 4000 images. Ten images were used to test, and the rest were used in the training process. A Contrast Limited Adaptive Histogram Equalization (CLAHE) was applied to improve the quality of training images and feature extraction. This study proposed a Convolutional Neural Network (CNN) with 6-layers filters. The input layer of the model has the dimension 440×440 pixels. They did not calculate the PSNR and the Loss of Multi-Scale Structural Similarity Index (MS-SSIM) obtained was 0.093. The other work related to our research was [Oh and Yun 2018]. They used a no-identified dataset composed of 348 paired images of bones and suppressed. The division of train and test was not specified and the model architecture used is composed of a CNN, GAN, and *Haar Wavelets*. The model input size was differential because of the resolution of 1024×1024 pixels. The MS-SSIM reached was 0.930 and the PSNR 24.08.

In another study, [Yang et al. 2017] employed the same dataset used in our experiments, reaching 0.976 of SSIM and 38.7 of PSNR. They propose a Deep Learning method for bone suppression in a single x-ray using cascade architecture of deep Convolutional Neural Networks (ConvNets) to map the bones gradient domain. The main idea was fusing with multi-scale bone gradients to improve prediction quality. More specifically, they used a Cascade of Multi-scale ConvNets (CamsNet). Their method does not require teaching from DES, but it requires segmentation and the border locations of bony structures. A positive point in their research is that their method works and performs considerably with different types of x-ray sources.

[Oh and Yun 2018] present two approaches; the first uses a conditional generative adversarial network, and the second a Haar 2D wavelet decomposition. They used the Euclidean distance between pairwise outputs to calculate the final result precision. They add on the experiments adversarial training to maintain the sharpness of specific lesions to avoid suppressing them. The main objective was minimizing the pixel-wise differences in bone suppression. The objective was to propose an image-to-image translation better than disposed on literature. Furthermore, they used a 2D wavelet decomposition as a perceptual guideline to minimize generic and ground truth differences. Finally, it is proposed a rigorously evaluated model to suppress bones from Dual Energy X-rays (DXRs).

In 2018, [Zhou et al. 2018], dividing the dataset into 170 images to train and 40 to test, they used a Multi-scale Conditional Adversarial Network (MCA-Net). Their process produces a high-resolution virtual chest soft-tissue image from a synthetics rib chain generated. The process is divided into two parts. The first is to generate the bone images using a multi-scale fully convolutional network. The second part is to generate the soft tissue chest image using bone suppression of the standard CR with the virtual bone image generated. Their model was tested with the JSRT dataset, and the images were divided into 170 for train and 40 to test. Their study reached a PSNR of 39.7 and an SSIM of 0.884.

[Zarshenas et al. 2019] propose a study that has significant results. They propose generating virtual dual-energy images and separating ribs and clavicles from soft-tissue chest radiographs. The propose is an Orientation-frequency-specific Deep Neural Network Convolution. They tested with their own dataset, composed of 118 chest images, reaching a PSNR index of 29.82 and an SSIM of 0.912.

[Chen et al. 2019] proposed a Cascaded Convolutional Network Model in Wavelet Domain Decomposition to do the bone shadow elimination using 504 images from a proprietary dataset, divided into 404 for training and 100 for testing. The trained network is used to predict the wavelet coefficients of the bone images. Thus, the predicted bone image is subtracted from the source, generating a bone-free image to train the model. Their trained model reached an SSIM of 0.977 and a PSNR of 39.7.

Other similar work, we can find on [Zhou et al. 2020]. They propose a neural network model for bone suppression based on image-to-image translation. The model consists of dilated convolutions to avoid contextual information loss. Furthermore, the proposed method enforces pixel intensity similarity to improve the suppression quality using a deep convolutional network between the generated chest X-ray and the ground truth. Basically, the model proposed by [Zhou et al. 2020] consists of a generator and a discriminator. The generator uses a U-net-like architecture with dilated convolutions/deconvolutions. The discriminator is based on PatchGAN to enforce the similarity of high-level feature representations. They tested the model using the JSRT, divided into 192 images to train and 42 to test. They reached 0.97 for the SSIM index and 33.5 for PSNR.

In 2019, [Matsubara et al. 2020], proposed a Convolutional Neural Filter (CNF) for a spatial filtering via CNN regression. This filter outputs a value for the bone component according to the neighborhood of the target pixel. In this process, a bone image is generated and subtracted from the original chest X-ray image. The images for their study were obtained from Computer Tomography (CT) data. These CT images were converted in isotropic voxels, projecting them in the ventral-dorsal direction and applying a nonlinear transformation for bone enhancement. After that, the filter is applied, isolating the bone-specific signal. Finally, the bone-extracted image is obtained by subtracting the bone isolated from the original chest X-ray. Using the JSRT to evaluate the trained model, they reached a PSNR of 36.23 and an SSIM of 0.96.

Another study is [Liang et al. 2020], the proposal is based on a Generative Adversarial Network (GAN) that learns bone suppression from dual-energy chest radiographs. A GAN is composed of two networks: a generator and a discriminator. The former creates images similar to the training set, while the latter discriminates them, classifying them as natural or artificial. The authors evaluate two variations of GANs, namely Pix2Pix [Isola et al. 2017] with paired radiographs and Cycle-GAN with unpaired radiographs. With a private dataset composed of 1,867 anonymized dual-energy images, the data was divided into 70% to train, 20% to test, and 10% to validation. The authors got an SSIM of 0.867 and a 36.078 on the PSNR index for the suppression task.

On [Sirazitdinov et al. 2020] was used the ChestX-ray-14, a public dataset provided by [Gusarev et al. 2017]. They used 24 images for train 7 to test and 4 to validation, with different models and architecture such as *autoencoder, U-Net, cGAN*. With all of them, the best precision was reached by the U-Net approach with 0.95 SSIM and 33.45 PSNR. Another study we can present is [Eslami et al. 2020] which used the augmented dataset of JSRT. Composed by 1.235 images, the study did not mention the division size of the train and test for the bone suppression part, just for the lung segmentation, but this is not our focus. For the bone suppression part, the architecture was pix2pix, the model input size was 512x512, and they calculated just the MS-SSIM, which was around 0.96 and 0.97.

[Gozes and Greenspan 2020] presented a different approach, building their own dataset from a Digital Reconstructed Radiographs (DRR) from a 664 Computer Tomography from a cancer dataset, the LIDC-IDRI. The division was 386 images to train, 129 to test, and 129 to validation. The model and technique used were based on segmentation of the bone structures in the CT domain to generate a bone suppressed image to train a Fourier Convolutional Neural Network (FCNN) model available on [Pratt et al. 2017]. The input of the trained network is 512x512, and they got 0.7 on the SSIM index and 22.6 for PSNR. The differential was to apply a dilated convolution and build an own dataset from a different source.

Another approach that makes bone segmentation for diagnoses uses a neural network to segment the chest region. [Eslami et al. 2020] propose a multitask model that does organ segmentation, and one of those processes on their pipeline is bone shadow elimination. As a model architecture, a CNN-based PatchGAN is used to do the bone suppression task. This architecture produces a matrix of size $k*k*1$ from an input tensor where k is the size of the image. They used the JSRT dataset with 247 CXRs, including lung nodules images. All the images have the 2048×2048 pixel dimension, and it was resized to 512×512 to adapt to the model entrance. The architecture used to translate the images was Pix2Pix [Isola et al. 2017]. The results were evaluated by the SSIM, looking for similarity estimation and the MSE to measure the difference between the predicted and ground truth values. Calculating the MS-SSIM, the authors got a state of the art results of 0.97.

[Rajaraman et al. 2021b] used the JSRT dataset ass well. They enhanced the contrast of the pixels values by 1%. The dataset composed of 4500 images was divided into 90% to train, 10% to test, and 10% to validation. The proposed architecture is a Residual Network model (ResNet-BS), where BS means Bone Suppression. The input size is 256×256, and even none of the proposed methods was the adversarial method, they experimented with four different architectures with the same dataset. For this last study, they got for SSIM 0.9492 and PSNR 34.0678.

3. Methodology

In this section, we display the methodology used to develop this work. Initially, we present brief concepts of Generative Adversarial Networks (GAN) and how a conditional GAN works. After this, we present the step-by-step used to build our model and illustrate the general architecture.

A GAN is a type of Machine Learning (ML) model that uses two neural networks as its core. These networks are called Generative Network (Generator) and Discriminative Network (Discriminator). Respectively, the Generator is a Convolutional Neural Network (CNN), and the Discriminator is a Deconvolutional Neural Network (DNN). The Generator's goal is to produce data as close as possible to the train data. Moreover, the Discriminator classifies the generated data. For instance, to illustrate in this work context, we have images without bones to generate artificial images like that. The Discriminator will work classifying these generated images as real or fake. To illustrate, we show the Figure 1.

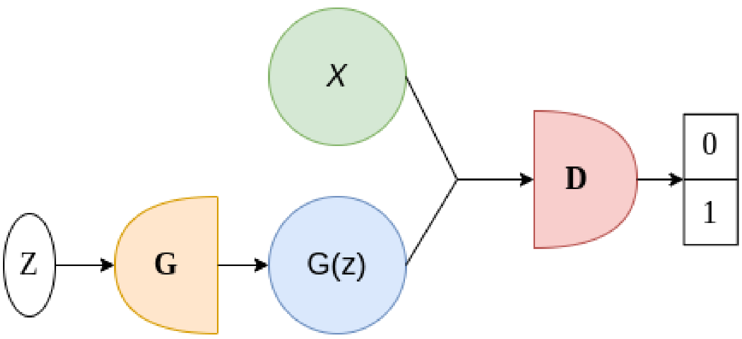

Figure 1. GAN schema based on [Oh and Yun 2018]

A Conditional GAN can be comprehended according to [Mirza and Osindero 2014]. To illustrate, for example, in a traditional GAN, we do not have control over the generated data. This method is called an unconditioned generative model. Nevertheless, we can direct the model predictions to the objective, establishing conditions and class labels or part of the target data. In the figure below based on [Oh and Yun 2018], we can illustrate, through the 2, the discriminator **x** and **y** presented as inputs to a discriminative function.

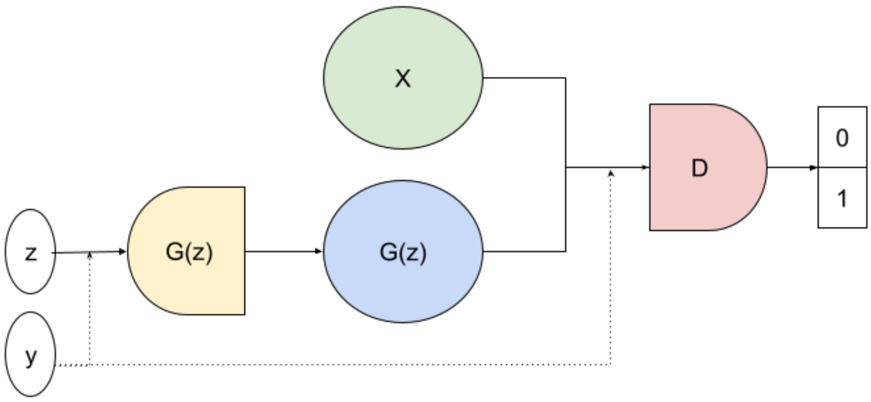

Figure 2. Condictional GANs schema

As described, our proposal is based on a Conditional GAN. The Generator and Discriminator are conditioned to an auxiliary function that helps our model reach the target image. The training process can be divided into two steps. The first moment, we created the folders to dispose of the data to be consumed by the training task. After that, we created a folder to receive the model checkpoints after every 5000 iterations. During the training process, a set of few images are getting to readjust the model weights called validation set.

3.1. Model Composition

Our proposal is composed of a generator based on a U-Net architecture and a discriminator represented by a PatchGAN, similar as proposed by [Isola et al. 2017]. In a few words, the PatchGAN is a type of Discriminator that only penalizes the scale of local image patches. Each patch of images is classified as to whether a sample is real or fake. Below we keep describing the details of the model composition.

3.1.1. Downsample - Encoder layer

This layer starts with a random normal initialize and is composed of a Sequential Keras model, which uses a standard 2D convolutional network. We deactivated the Batch normalization in the first convolutional block, but the other layers are activated by default. The last step is applied and the Leaky ReLU. The down_stack is composed of 8 layers, and it starts with the shape (256, 256, 3). In the future, we can adapt the model to a single-channel image. However, initially, just for tests and study, we decided to leave it.

3.1.2. Upsample - Decoder layer

This layer starts with random normal initialization and is composed of a Sequential Keras model that uses a standard 2D convolutional network. As on the first Sequential process mentioned, both were built to provide training and inference features on the model. Some of them were applied the *dropout* on intention to reduce the processing time.

3.1.3. Generator

As described in [Isola et al. 2017], GANs learn a loss and adapt it to the data. So the output which is distant from the target is penalized. A sigmoid cross-entropy represents the generator loss in this study. Another metric used was an L1 loss, calculated between the generated image and the target image based on mean absolute error (MAE). The formula to calculate it proposed by the authors was:

$$Generator\_loss = gan\_loss + \lambda \times l1\_loss$$

3.1.4. Discriminator

The Discriminator is a convolutional PatchGAN classifier. It tries to classify each image as artificial or real. The classifier receives the target and generated images, and the Discriminator classifies both images. The steps are composed of Convolution layers, Batch normalization, and a Leaky ReLU as the activation function.

To evaluate the Discriminator, we calculate its loss function. It means how the model is performing classifying real and artificial images. We input the real and the generated images to the discriminator loss function and do the data verification for each image classified correctly by the Discriminator. The final loss is the sum of the real and the generated losses. The real and the generated loss are calculated using the sigmoid cross-entropy.

3.2. Training the model

We fed the input and target images into the network in the training step. After that, the generator calculates the discriminator loss. The gradient loss is optimized over each interaction. To train our GAN, we use a loop interaction. This loop involves the processes of generating, discriminating, and validation during the train. The generated images are displayed every 1000 steps to show the progress, and the model checkpoints are saved every 5000 steps.

3.3. Dataset

Our dataset is composed of images from the Japanese Society of Radiological Technology (JSRT), which have the original image and the correspondent bone suppressed. It is composed of 240 pairs, disposed by [Hyunh 2021] and avaliable on `https://dx.doi.org/10.21227/xnb5-hg35`, already augmented the data, and is destined to research. To do our experiments, we divided the image pairs into three categories: 3.828 for the train, 226 for the test, and 26 for validation. The validation process was made during the training to measure and balance the network weights. Finally, the test set was used to measure how accurate the model was. Figures 3 and 4 presents some images from the JSRT dataset.

Figure 3. Complete rib cage

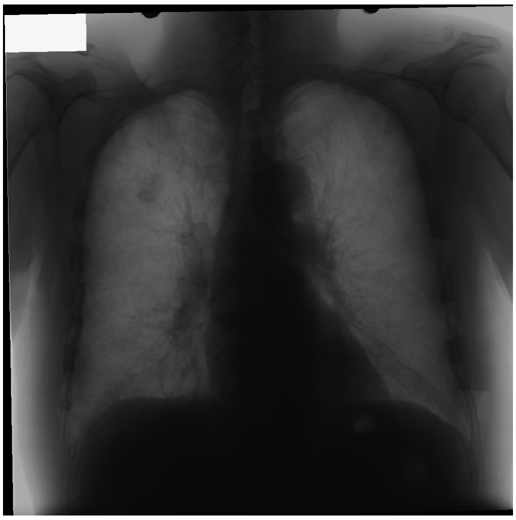

Figure 4. Bone suppressed

3.4. Metrics

Below, we presented a brief description and explanation of the most used metrics to measure the model quality on the bone suppression task in the literature review. In our proposal, we decided to use PSNR and SSIM.

3.4.1. Peak Signal-to-Noise Ratio (PSNR)

Given the images $\boldsymbol{f}$ and $\boldsymbol{g}$, both of $M \times N$ size [Hore and Ziou 2010] calculate the PSNR index with the Equation 1. Where MSE means the Mean Square Error.

$$PSNR(f,g) = 10 \times log_{10}\left(\frac{255^2}{MSE(f,g)}\right) \tag{1}$$

3.4.2. Structural Similarity Index Measure (SSIM)

the formula can be reduced to Eq. 2:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{2}$$

To better understand the variables, c_1, c_2 and c_3 are constants to stabilize the division and avoid a null denominator. Those μ_x and μ_y are the average of x and y signal values. The σ_x and σ_y are the variance of x and y signal values. We use in our work the SSIM according to [Rajaraman et al. 2021b], the index provides a measurement of the similarity between the ground truth and predicted images.

4. Results

Figure 5 display a sample generated from our GAN model after 45k training iterations. The output is of the same size as the input algorithm. It has a size of 256×256. The model is prepared to receive three channels, but our source is grayscale. It generates the same color output. Next, we present the Table 1 summarizing the main results.

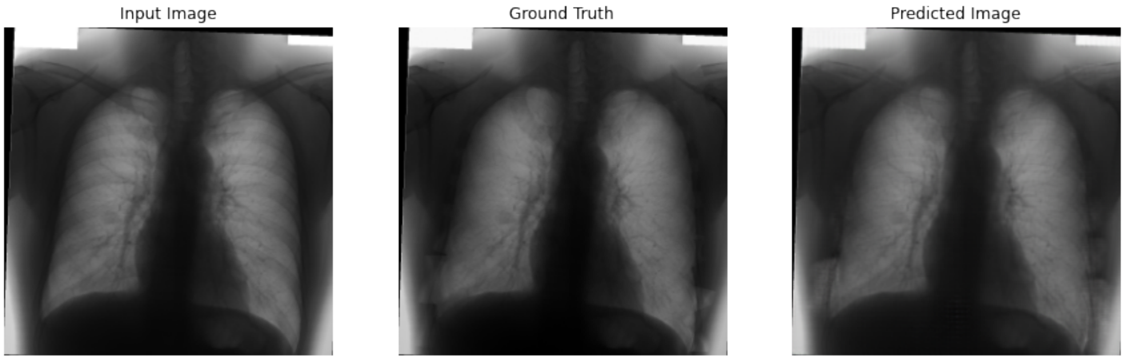

Figure 5. Input image (Source), Target image (Ground truth) and Predicted images

Table 1. Results from related works for bone suppression.

Name	Model and techniques	SSIM	PSNR
[Gusarev et al. 2017]	Auto-encoder and convolutional layers, Auto-encoder and without down/up sample operations	0.907	-
[Oh and Yun 2018]	Haar 2d Wavelet decomposition and vanilla pix2pix	0.930	24.08
[Eslami et al. 2020]	Condictional GAN and dilated convolutions	0.97	
[Zhou et al. 2020]	Dilated convolution to expand the receptive field	0.97	33.5
[Yang et al. 2017]	Cascade of multiscale CNN	0.976	38.7
[Zarshenas et al. 2019]	Anatomy-specific orientation-frequency-specific deep neural network convolution	0.912	29.82
[Chen et al. 2019]	Cascade of multiscale CNN & wavelet decomposition	0.977	39.40
[Zhou et al. 2018]	Multi-scale and conditional adversarial network	0.884	39.7
[Matsubara et al. 2020]	Bone suppression for chest X-ray image susing a convolutional neural	0.930	24.08
[Liang et al. 2020]	Cycle-GAN - Image-to-image translation	0.867	36.078
[Sirazitdinov et al. 2020]	autoencoder, U-net, FPN, cGAN	0.955	33.45
[Oh and Yun 2018]	CNN + GAN + Haar Wavelets	0.930	24.08
[Gozes and Greenspan 2020]	Hounsfield unit (HU) based segmentation and FCNN	0.70	22.6
[Rajaraman et al. 2021b]	Residual Network Model (ResNet-BS), where BS means Bone Suppression.	0.9492	34.0678
Our approach	Conditional GAN	0.943	34.967

In Table 1, we present the results of the first evaluations. Our research result at this first moment can be summarized in a few words. Our proposal is an Adversarial model using [Isola et al. 2017] framework architecture, and in the first train, we used all the augmented datasets of JSRT to train, validate and test. The results are for PSNR index 34.967 and 0.943 for SSIM.

We can observe, according to Table 1, that our result is close to the library and better than eight studies from 14 analyzed. Going into the analysis, in all of the reviewed papers, [Eslami et al. 2020, Zhou et al. 2020, Zhou et al. 2018, Rajaraman et al. 2021b] and ours used the JSRT dataset. When comparing our result with the study that used the same dataset and a similar technique, it is possible to see that our model achieved significant results. Studies like [Eslami et al. 2020, Zhou et al. 2020, Yang et al. 2017, Chen et al. 2019] reached an SSIM around 0.97, but we cannot reproduce what they did.

For an illustration of our model results, we present Figure 5. This image was obtained from our model after 45k training iterations. The output is of the same size as the input algorithm. It has a size of 256×256. The model is prepared to receive three channels, but our source is gray-scale. It generates the same color output.

5. Conclusions

Developing models that can be used in medical software to assist the clinical diagnosis is a big challenge. Although this work is a simple study, with low resolution, it shows the potential of the approach. In a compromise with the society, democracy, and inclusion, dispose of the code and the researched techniques open and free, allow others researches develop a low cost or free solution for public hospitals.

This work evaluated the application of Conditional Generative Adversarial Networks (CGANs) to perform the bone suppression task in chest radiography images. We used two traditional computer vision metrics to evaluate this method: PSNR and SSIM. Our results display that this approach produces results among the best found in the literature.

As we can see in the literature review, the bones on the chest image can sometimes be noisy when the soft tissue is diagnosed. There are plenty of ways to attenuate the shadows to improve the medical analysis. Some of those require a specific type of equipment that the costs are not accessible and exposes the patient to a double x-ray emission. Other approaches use classical computer image processing, looking for contours and edges. Other approaches use neural networks as filters or feature extractors. The choice of our work was to use deep learning with Adversarial Networks.

Our proposal is a CNN-based solution that learns from a dataset source. Our solution uses a DES dataset to show the cGAN the source and the target, and the model will learn how to generate an artificial image like the target dataset. In the literature review, we could observe some not covered points. Some studies did the experiments and did not mention how many iterations or code parameters were used. Our results were near from the review using the same metrics. For example, our approach is the third-best using PSNR, and for SSIM, we got results closer to the average performances.

To conclude the first part of our study, we could analyze the literature review and compare our results. As we can see, the model is promising and could even be improved if it is trained with real-size images such as 1024×1024. We figured that our approach got better results than classical methods and other deep learning strategies. Furthermore, the proposed model is a proof of concept and with comparable and better results than presented in the literature.

6. Future Works

Is known, that if a doctor analyzes an image generated by a system like that, the diagnosis and the whole clinical condition require more specific exams, it will be required. Our study is just a start for others more complex and complete. In the future, we aim to increase the input model for a high resolution, such as 1024×1024. We are planning, as well, to apply more image processing techniques to increase the image quality to reduce the noise and attenuate the shadows. Another future work is to use a dual-energy dataset, make the bone subtraction with classic techniques, and with the resulting train our cGAN based model.

References

Chen, Y., Gou, X., Feng, X., Liu, Y., Qin, G., Feng, Q., Yang, W., and Chen, W. (2019). Bone suppression of chest radiographs with cascaded convolutional networks in wavelet domain. *IEEE Access*, 7:8346–8357.

Eslami, M., Tabarestani, S., Albarqouni, S., Adeli, E., Navab, N., and Adjouadi, M. (2020). Image-to-images translation for multi-task organ segmentation and bone suppression in chest x-ray radiography. *IEEE transactions on medical imaging*, 39(7):2553–2565.

Gozes, O. and Greenspan, H. (2020). Bone structures extraction and enhancement in chest radiographs via cnn trained on synthetic data. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pages 858–861. IEEE.

Gusarev, M., Kuleev, R., Khan, A., Rivera, A. R., and Khattak, A. M. (2017). Deep learning models for bone suppression in chest radiographs. In *2017 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pages 1–7. IEEE.

Hore, A. and Ziou, D. (2010). Image quality metrics: Psnr vs. ssim. In *2010 20th international conference on pattern recognition*, pages 2366–2369. IEEE.

Hyunh, M.-C. (2021). X-ray bone shadow suppression dataset. IEEE Dataport.

Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134.

Juhász, S., Horváth, Á., Nikházy, L., and Horváth, G. (2010). Segmentation of anatomical structures on chest radiographs. In *XII Mediterranean Conference on Medical and Biological Engineering and Computing 2010*, pages 359–362. Springer.

Liang, J., Tang, Y.-X., Tang, Y.-B., Xiao, J., and Summers, R. M. (2020). Bone suppression on chest radiographs with adversarial learning. In *Medical Imaging 2020: Computer-Aided Diagnosis*, volume 11314, page 1131409. International Society for Optics and Photonics.

Matsubara, N., Teramoto, A., Saito, K., and Fujita, H. (2020). Bone suppression for chest x-ray image using a convolutional neural filter. *Physical and Engineering Sciences in Medicine*, 43(1):97–108.

Matters, I. (2014). Carestream's new bone suppression software receives fda clearance, now available worldwide.

Mirza, M. and Osindero, S. (2014). Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*.

Oh, D. Y. and Yun, I. D. (2018). Learning bone suppression from dual energy chest x-rays using adversarial networks. *arXiv preprint arXiv:1811.02628*.

Oliveira, B., Ziviani, H., Oliveira, J., Viegas, A., and Calvo, D. (2021). Suporte para diagnóstico de covid-19 por meio de classificação automática de imagens de raio-x e modelos explicáveis. In Filho, C. J. A. B., Siqueira, H. V., Ferreira, D. D., Bertol, D. W., and ao de Oliveira, R. C. L., editors, *Anais do 15 Congresso Brasileiro de Inteligência Computacional*, pages 1–8, Joinville, SC. SBIC.

Pratt, H., Williams, B., Coenen, F., and Zheng, Y. (2017). Fcnn: Fourier convolutional neural networks. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 786–798. Springer.

Rajaraman, S., Cohen, G., Antani, S., et al. (2021a). A bone suppression model ensemble to improve covid-19 detection in chest x-rays. *arXiv preprint arXiv:2111.03404*.

Rajaraman, S., Zamzmi, G., Folio, L., Alderson, P., and Antani, S. (2021b). Chest x-ray bone suppression for improving classification of tuberculosis-consistent findings. *Diagnostics*, 11(5):840.

Sirazitdinov, I., Kubrak, K., Kiselev, S., Tolkachev, A., Kholiavchenko, M., and Ibragimov, B. (2020). Evaluation of deep learning methods for bone suppression from dual energy chest radiography. In *International Conference on Artificial Neural Networks*, pages 247–257. Springer.

Sujath, R., Chatterjee, J. M., and Hassanien, A. E. (2020). A machine learning forecasting model for covid-19 pandemic in india. *Stochastic Environmental Research and Risk Assessment*, 34:959–972.

Wang, S., Yang, D. M., Rong, R., Zhan, X., Fujimoto, J., Liu, H., Minna, J., Wistuba, I. I., Xie, Y., and Xiao, G. (2019). Artificial intelligence in lung cancer pathology image analysis. *Cancers*, 11(11):1673.

Yang, W., Chen, Y., Liu, Y., Zhong, L., Qin, G., Lu, Z., Feng, Q., and Chen, W. (2017). Cascade of multi-scale convolutional neural networks for bone suppression of chest radiographs in gradient domain. *Medical image analysis*, 35:421–433.

Zarshenas, A., Liu, J., Forti, P., and Suzuki, K. (2019). Separation of bones from soft tissue in chest radiographs: Anatomy-specific orientation-frequency-specific deep neural network convolution. *Medical physics*, 46(5):2232–2242.

Zhou, B., Lin, X., Eck, B., Hou, J., and Wilson, D. (2018). Generation of virtual dual energy images from standard single-shot radiographs using multi-scale and conditional adversarial network. In *Asian Conference on Computer Vision*, pages 298–313. Springer.

Zhou, Z., Zhou, L., and Shen, K. (2020). Dilated conditional gan for bone suppression in chest radiographs with enforced semantic features. *Medical Physics*, 47(12):6207–6215.