Austin Publishing Group

## Research Article

# The Effectiveness of Data Augmentation for Bone Suppression in Chest Radiograph using Convolutional Neural Network

**Ren G[1], Lam S-K[1], Ni R[1], Yang D[1], Qin J[2] and Cai J[1]\***

[1]Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong SAR, China

[2]School of Nursing, The Hong Kong Polytechnic University, Hong Kong SAR, China

**\*Corresponding author:** Jing Cai, Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong SAR, China

## Abstract

**Objective:** Bone suppression of chest radiograph holds great promise to improve the localization accuracy in Image-Guided Radiation Therapy (IGRT). However, data scarcity has long been considered as the prime culprit of developing Convolutional Neural Networks (CNNs) models for the task of bone suppression. In this study, we explored the effectiveness of various data augmentation techniques for the task of bone suppression.

**Methods:** In this study, chest radiograph and bone-free chest radiograph are derived from 59 high-resolution CT scans. Two CNN models (U-Net and Generative Adversarial Network (GAN)) were adapted to explore the effectiveness of various data augmentation techniques for bone signal suppression in the chest radiograph. Lung radiograph and bone-free radiograph were used as the input and target label, respectively. Impacts of six typical data augmentation techniques (flip, cropping, noise injection, rotation, shift and zoom) on model performance were investigated. A series of statistical evaluating metrics, including Peak Signal-To-Noise Ratio (PSNR), Structural Similarity (SSIM) and Mean Absolute Error (MAR), were deployed to comprehensively assess the prediction performance of the two networks under the six data augmentation strategies. Quantitative comparative results showed that different data augmentation techniques exhibited a varying degree of influence on the performance of CNN models in the task of CR bone signal suppression.

**Results:** For the U-Net model, flips, rotation (10 to 20 degrees), all the shifts, and zoom (1/8) resulted in improved model prediction accuracy. By contrast, other studied augmentation techniques showed adverse impacts on the model performance. For the GAN model, it was found to be more sensitive to the studied augmentation techniques than the U-Net. Vertical flip was the only augmentation method that yielded enhanced model performance.

**Conclusion:** In this study, we found that different data augmentation techniques resulted in a varying degree of impacts on the prediction performance of U-Net and GAN models in the task of bone suppression in CR. However, it remains challenging to determine the optimal parameter settings for each augmentation technique. In the future, a more comprehensive evaluation is still warranted to evaluate the effectiveness of different augmentation techniques in task-specific image synthesis.

**Keywords:** Data augmentation; Bone suppression; Chest radiograph

## Introduction

Lung cancer is one of the second commonly occurring cancer worldwide, contributing about 11.4% of the new cancer cases [1]. One of the standard treatments for lung cancer is radiation therapy [2,3]. With the help of On-Board Imaging (OBI) systems, Image-Guided Radiation Therapy (IGRT) has been able to deliver a more accurate dose to the tumor region and reduce the radiation toxicity to the normal tissues [4,5]. The 2D Chest Radiograph (CR) generated by the OBI system is commonly used to determine the patient position and decrease the patient position variations during the IGRT course of lung cancer [6]. However, the bony structure in CR often obscures the localization of the target or landmarks, causing a maximum error of 22mm during the IGRT of lung cancer [7]. To improve the localization accuracy of IGRT, bone suppression in CR is regarded as a promising solution [8].

Various efforts have been made for bone suppression in CR. Dual-energy (DE) radiographic imaging attempts to leverage the difference in attenuation coefficients between bones and soft tissues for obtaining the separation of bone and soft tissue images using two levels of X-ray exposures [9]. Despite the increased diagnostic sensitivity, its clinical application in radiation oncology is still largely restricted. More recently, multiple deep learning techniques have been extensively studied, a variety of CNNs have achieved remarkable progress and have been successfully applied to the task of bone signal

suppression, including multiple massive-training artificial neural networks [10], filter learning [11], massive training artificial neural network [12], cascade of multi-scale convolutional neural networks [13], frequency-specific deep neural network convolution [14], to name a few. These methods suppress bone structures by regression prediction, in which the bone-free Digital Radiograph (DR) is used as training ground truth [10-16]. Although such bone suppression methods provide the radiologists with an unobstructed view of the lung tissue, streamlining diagnostic sensitivity of CR without incurring additional radiation dose to patients, the prediction accuracy of the deep learning models still heavily relies on the availability of large-scale high-quality data [17,18]. Undoubtedly, this poses a practical challenge in real-world scenario, since massive expenses and manual efforts are required to obtain enormous amount of datasets in demand, especially in the context of sparse availability of the desired label of interest [19].

Confronted with this roadblock in building deep learning models, data augmentation, which is the process of applying one or more geometric deformations for inflating the size of training dataset artificially [20,21], has been widely adopted. As deep learning models treat a geometrically transformed image as a meaningful image, CNN models can be trained using the deformed dataset to generate more "unseen" data. As such, data augmentation plays a vital role in enhancing the performance of classification and segmentation since it increases the data variability [22,23] and does not affect the semantic validity of the original dataset [24]. The effectiveness of data augmentation has been tested in many natural image datasets, including MNIST handwritten digit recognition, CIFAR-10/100, ImageNet, tiny-imagenet-200, SVHN (street view house numbers), Caltech-101/256, MIT places, MIT-Adobe 5K dataset, Pascal VOC, and Stanford Cars [19]. Hussain et al. compared training model performance utilizing different augmentation strategies, and their results suggested that both discriminative and generative performance were drastically affected [25]. Several data augmentation approaches, such as flips and rotations, have been widely studied in the literatures involving raw medical images. Nevertheless, the impact of various data augmentation techniques in the case of medical synthesis problems, particularly in the aspect of bone signal suppression in CR, remains to be investigated.

In this study, we investigated the impacts of six typical data augmentation techniques (flip, cropping, noise injection, rotation, shift, and zoom), each with varying intensities of augmentation, for the task of bone signal suppression in CR on two popular deep learning architectures, U-Net and GAN. A series of statistical evaluating metrics, including Peak Signal-To-Noise Ratio (PSNR), Structural Similarity (SSIM) and Mean Absolute Error (MAR), were deployed to comprehensively evaluate the prediction performance of the deep learning models under different data augmentation strategies. Our overarching purpose was to provide insights into the optimal adoption of data augmentation initiatives in synthesized bone-free CR using two typical deep learning architectures, U-Net and GAN.

## Materials and Methods

### Dataset and image processing

A publicly available dataset, RIDER Lung CT dataset, from The Cancer Imaging Archive (TCIA) [26,27] was used as the raw data to derive the lung Digital Radiograph (DR) images and bone-free lung DR image, which were used as the input and target, respectively. This dataset contains 59 high-resolution CT scans of the chest from non-small cell lung cancer patients. Each CT slice was constructed into a $512 \times 512$ matrix with $0.576 \times 0.576$ mm$^2$ pixel spacing and 1.25mm slice spacing.

The image processing workflow for bone-free DR images derivation is illustrated in Figure 1. The lung was first automatically segmented in CT using the U-net (R231) model [20,28], which was pretrained on a widely diverse lung CT scans. Bony structures were segmented by thresholding of +300 HU. The bone and lung segments were subsequently applied to the high-resolution CT to generate the bone CT and lung CT images. The digital radiographs were simulated from 3D volumetric CT images using the Insight Segmentation and Registration Toolkit (ITK) package in the Python environment. The ITK is an open-source SDK for image analysis and image processing and is widely used in medical image processing. To focus our analysis on the lung region, a lung mask was built on the lung region and applied on CRs for segmenting the bone DR and bone-free DR. The processed DRs were multiplied to generate the lung DR images. 47 cases of the generated pairs were used for downstream model
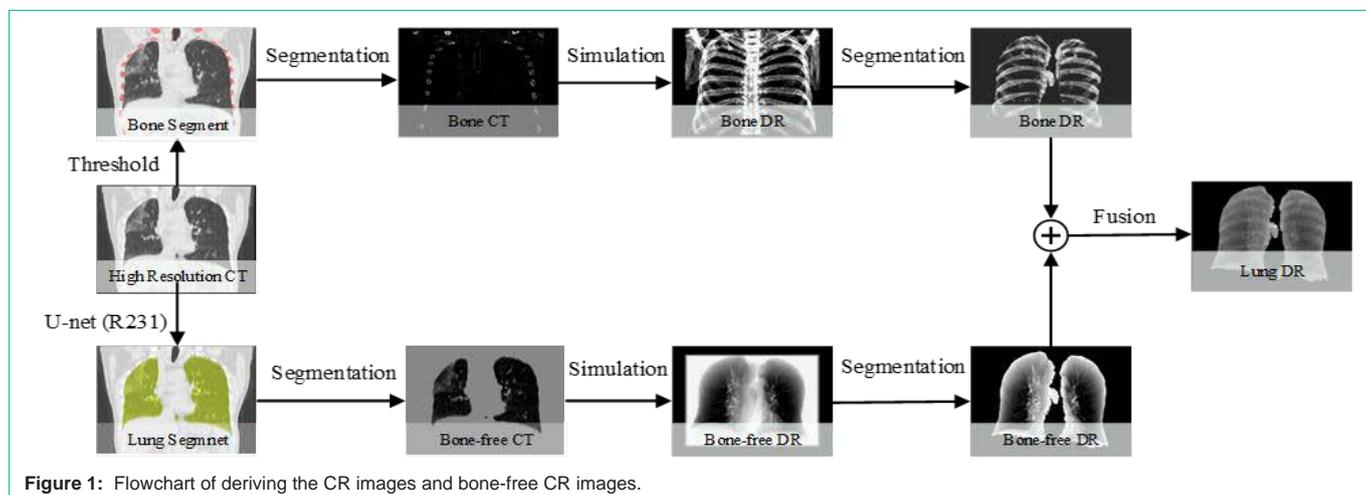


**Figure 1:** Flowchart of deriving the CR images and bone-free CR images.
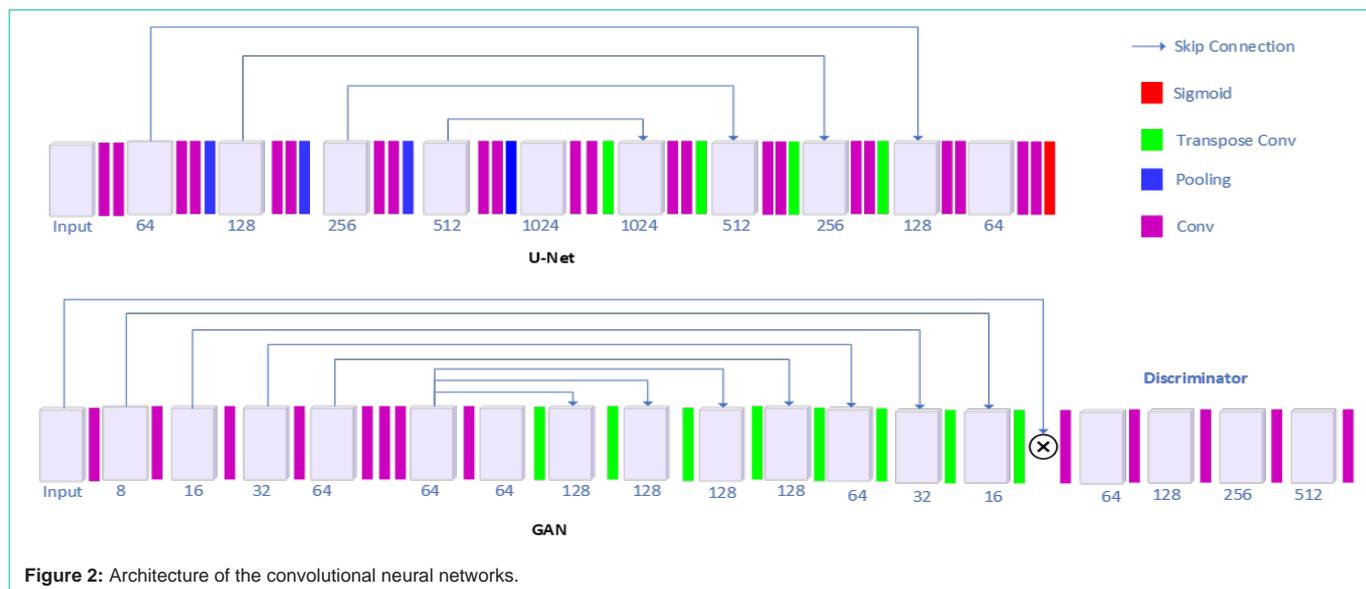
**Figure 2:** Architecture of the convolutional neural networks.

training, and 12 cases were used for model testing.

## Deep CNNs

Two commonly studied convolutional neural network architectures, U-Net [29], and Generative Adversarial Network (GAN) [30] were utilized for bone suppression of lung DR images in this study. The U-Net model used an encoding-decoding structure, as shown in Figure 2. Four skip attentions were used for original shape recovery. The convolution captures the hierarchical texture features of the input. To extract global texture features, four 2×2 pooling convolutions were used to increase the size of the receptive field. Accordingly, four transpose convolutions were used to recover to the original image size. Each convolution has a size of 3×3, and is followed by a batch normalization, and a Parametric Rectified Linear Unit (PReLU). At the last layer, a Sigmoid function sums up the results of the previous layers. The predicted values are in the range of (0,1). Binary Cross Entropy (BCE) loss was used as the loss function to minimize the difference between the ground truth and the output of the network.

The GAN model is composed of a generator and a discriminator. The two networks were trained simultaneously by an adversarial process. The generator used a fully convolutional architecture. 16 layers of convolutions were used to learn to create bone-free DR images. The discriminator has 5 layers and learns to distinguish between real and fake bone-free DR images. The maximum width of the generator and discriminator are 128 and 512, respectively. The output of the generator was concatenated with the original lung DR images to be discriminated by the discriminator. The generator used BCE as the loss function, while the discriminator used Mean Square Error (MSE) as the loss function. The overall loss function of CCNN was:

$$L = L_{MSE} + \gamma L_{BCE} \qquad (1)$$

γ was the weighing factor of the second network loss and was empirically set to 5.

The proposed method learns the optimal parameter values by

minimizing the loss function between the bone-free lung DR and the output of the network. Both models were updated using error backpropagation with an Adaptive Moment Estimation Optimizer (ADAM). The number of epochs was 400, and each epoch includes five iterations. We implemented our network using the Pytorch 1.1 framework. All the preprocessing steps were coded in python. All the experiments were performed using a workstation with CPU Intel Core i7-8700 @ 3.2GHz, GPU NVIDIA GTX 2080 TI with 11GB memory, and 32GB of RAM.

## Data augmentation

During model training, we applied six augmentation techniques to generate six categories of new training images. In each training epoch, either the generated new image or the original image was used for training. The augmentation techniques are described as follows:

**Flip:** Horizontal flip, vertical flip, or the combination of both was applied to each of the training images. Conventionally, a horizontal flip is more common used, compared to the vertical flip. This augmentation is one of the easiest approaches to implement and has been proven to be useful for classification tasks on datasets CIRAR-10 and ImageNet.

**Cropping:** Cropping of the image is regarded as a practical augmentation technique for data with changing height and width dimensions. In the process of cropping, γ percent of the original image was cropped and zoomed in relation to the original image size, where γ is 7/8, 6/8, and 5/8.

**Noise injection:** In the classification dataset, adding noise helps the CNN model to learn more robust features [31]. In this study, an array, N, was generated. The element of N was a sample from a Gaussian distribution with μ=0 and σ as an integer in the range of (0.1, 0.5). For each image I, the noise-injected image I' = I + N was obtained.

**Rotation:** Rotation augmentation was done by rotating the image right or left on an axis between certain degrees. The slight rotation was a desirable augmentation in the classification problem. We
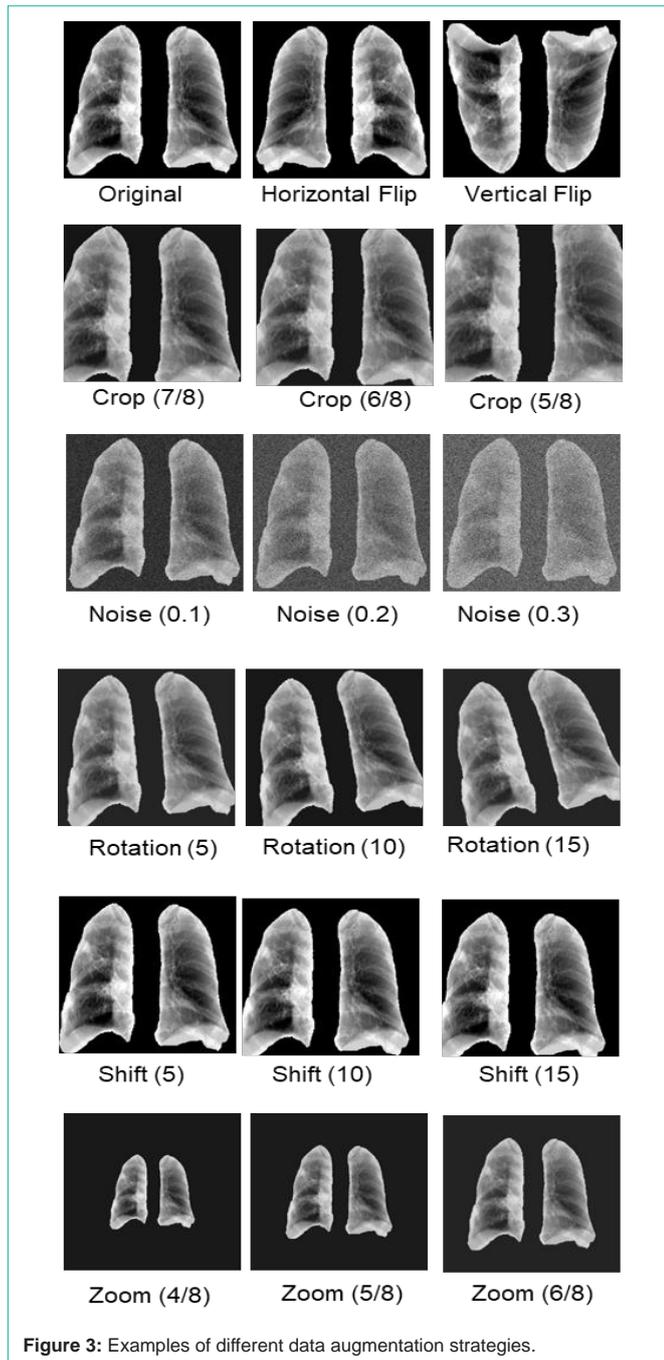
**Figure 3:** Examples of different data augmentation strategies.

randomly rotated the images between (-θ, θ), where θ is 5, 10, 15, 20, 25, 30 degrees.

**Shift:** Shifting the image can be a useful geometry augmentation to avoid positional bias in the dataset. In this study, the image was shifted randomly in the left, right, up, or down direction. The remaining pixels were filled with 0. Here, we shifted the image along both axes for β pixels, where β is 5, 10, 15, and 20.

**Zoom:** Zooming the image can be adopted as a processing step for the image object with different sizes. Here, we reduced the image size to α percent of the original image size, where α is 1/8, 2/8, 3/8,

4/8, 5/8, and 6/8 in different experiments. The background regions were filled with 0.

**Evaluation**

In addition to the commonly used image evaluation metrics, such as Peak Signal to Noise Ratio (PSNR) and Structural Similarity (SSIM), we adopted the Mean Absolute Error (MAE) to quantitatively assess the prediction performance of the proposed neural network in comparison to our ground truth, bone-free lung DR, in different augmentation settings (Figure 3).

PSNR is a quality metric assessing image quality between the original and compressed, or augmented images to describe how its fidelity is affected. The PSNR can be computed using the following equation:

$$PSNR = 20 \cdot \log_{10}\left(\frac{MAX}{\sqrt{MSE}}\right) \qquad (2)$$

where *MAX* is the maximum fluctuation in the input image data type. *MAX* is 1 in this study, as our input images are the double-precision floating-point data type. MSE is the mean squared error, defined as:

$$MSE = \frac{\sum_{M,N}[I_1(m,n) - I_2(m,n)]^2}{M*N} \qquad (3)$$

where $I_1$ and $I_2$ are the two images. *M* and *N* are the dimension sizes along the x-axis and the y-axis.

The SSIM contains three terms as the comparisons of three measurements: luminance term, contrast term, and the structural term [32]. The overall SSIM is a multiplicative combination of these terms:

$$SSIM = \frac{2\mu_y\mu_p + C_1}{\mu_y^2 + \mu_p^2 + C_1} \cdot \frac{2\sigma_{yp} + C_2}{\sigma_y^2 + \sigma_p^2 + C_2} \qquad (4)$$

where $\mu_P$, $\mu_I^*$, $\sigma_P$, $\sigma_I^*$ and $\sigma_{II}^*$ are the local means, standard deviations, and cross-variance for image *y* and *p*. $C_1=(k_1 L)^2$, $C_2=(k_2 L)^2$ are the two variables that stabilize the division with weak denomination.

MAE measures the matching error between the original and augmented images, calculating as:

$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n} \qquad (5)$$

To compute the arithmetic average of the absolute errors. In this study, $y_i$ is the augmented image and $x_i$ is the reference.

## Results

The accuracy of the proposed CNN models (U-Net and GAN) is presented in Figure 5. The baseline models of U-Net and GAN achieved MAE = 0.0340 ± 0.0095 and 0.0357 ± 0.0050; PSNR = 27.3681 ± 1.9674 and 27.0374 ± 1.0381; SSIM = 0.9408 ± 0.0121 and 0.01502, respectively. To further evaluate the performance, a representative case is visualized in Figure 4. The bone-free image synthesized by U-Net did not recover small vessels in the periphery region. In comparison, GAN failed to recover most of the detailed information in the periphery region and predicted most pixels as white. The U-Net model outperformed GAN in 24 out of 27 data augmentation strategies, while only in vertical flip, zoom (5/8), zoom (4/8) scenarios GAN has better performance.

For the U-Net model, the performance increased with different flips, rotation from 10 to 20 degrees, all the shift, and zoom (1/8) augmentations. With both vertical and horizontal flip, the
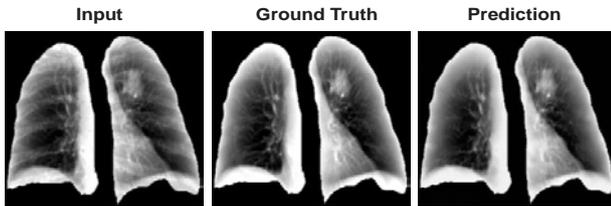
**Figure 4:** Illustration of the synthesized bone suppression image.

performance increased by -8.0069% (MAE), 3.3454% (PSNR), and 0.9878% (SSIM). Rotation with 10-degree lead to -10.8744% (MAE), 4.6000% (PSNR), and 1.2339% (SSIM) performance increase, which is the largest increment among all degrees of rotations tested.

Zoom augmentation to 1/8 of the original image also improved the performance by -13.9694% (MAE), 4.0462% (PSNR), and 0.8736% (SSIM).

The GAN models demonstrated a larger variation in performance (Variance = 0.0102, 28.3833, and 0.0420 for SAM, PSNR, and SSIM, respectively) with different data augmentation techniques in training, compared to U-Net (variance = 0, 1.9595 and 0.0003 for SAM, PSNR, and SSIM, respectively). With a vertical flip, the performance increased by -15.7407% (MAE), 4.7206% (PSNR), and 1.8480% (SSIM).

The correlation between different evaluations metrics was also evaluated using the Spearman correlation in this study (Figure 6). For the U-Net model, the evaluation metrics MAE, PSRN, and SSIM
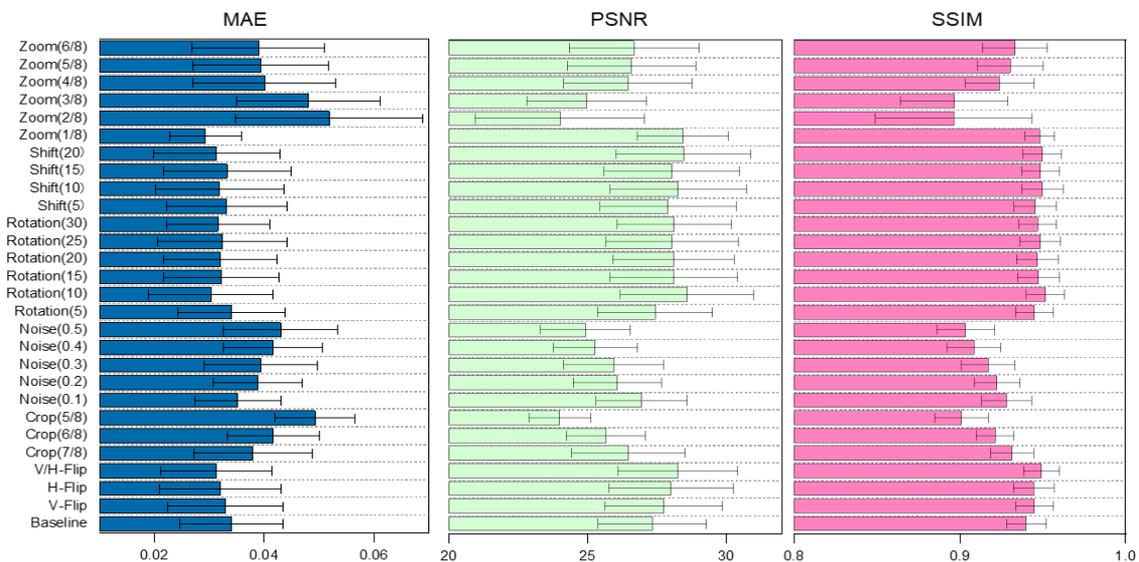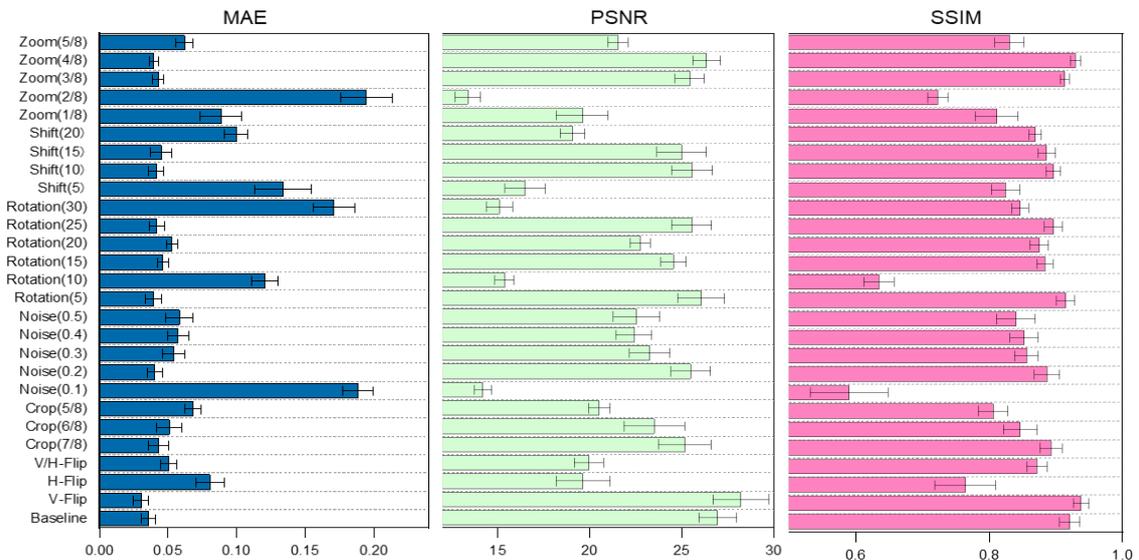


**Figure 5:** Performance of different data augmentation strategies using the U-Net model (a) and GAN model (b).

**Figure 6:** Correlation of different evaluation metrics.

had a high absolute correlation with each other (r=0.96-0.98). For the GAN model, MAE and SSIM are highly correlated (r= -0.95), followed by MAE and PSNR (r= -0.89) and SSIM and PSNR (r= 0.83).

## Discussion

Our quantitative results demonstrated that different data augmentation techniques resulted in a varying degree of impact on prediction performance of both U-Net and GAN deep learning models in the task of chest X-ray bone signal suppression. For the U-Net model, flips, rotation from 10 to 20 degrees, all the shift, and zoom (1/8) enhanced the accuracy of the synthesized image, while the rest of the studied augmentation techniques incurred adverse consequences. Compared to the U-Net model, the GAN model was found to be more sensitive to the studied augmentation techniques, among which only the vertical flip resulted in a positive effect on model predictability.

For the image synthesis task of bone signal suppression, some augmentation methods capture medical image statistics more effectively than others. For example, flips, rotation, and shift were the desirable augmentation techniques for the training of U-Net. We also revealed that a combination of all optimal augmentation techniques did not improve the model performance. Other augmentation techniques, like crop and noise injection, generally decreased the performance of the CNN model. According to a recent publication examining impact of noise injection on model performance, adding noise can help the neural network learn more robust features for the classification task [31]. Nevertheless, results from this study found that, in image synthesis task of bone signal suppression, adding noise deteriorated the prediction accuracy on the testing group. Such discrepancy could be intuitively elucidated, in part, that the Gaussian noises seldom appear on the real DR images. This augmentation compromises the prediction ability on the real lung DR images. This effect also showed the special of medical images is different from the natural images. Special attention should be paid to augment the medical image and understand types of the underlying variant.

Furthermore, we observed that the studied augmentation strategies led to a varying extent of impacts on the two assessed CNN models. For example, flips, rotation, and shift were found to be desirable augmentation techniques for training U-Net models, while vertical flip was the only augmentation approach that reinforced the performance of GAN models. This could partly be attributed to the fact that GAN model is often more complex, involving more

parameters and consisting of two networks, as in contrast to the U-Net model. These features of GAN models render difficulties for modelers to fine-tune the modeling parameters in order to achieve optimal model performance [33]. For the training of CNN models, it is important to determine the desirable augmentation techniques.

In spite of the wide application scope of data augmentations in image synthesis, a thorough evaluation of these techniques in relation to CNN model performance has not yet been discussed. In this study, we explored the effectiveness of different geometry augmentation techniques and attempted to identify desirable augmentation techniques. However, it remains challenging to determine the optimal parameter settings for each augmentation technique. In the future, a more comprehensive evaluation is still warranted to evaluate the effectiveness of different augmentation techniques in task-specific image synthesis.

## Conclusions

In this study, we found that different data augmentation techniques resulted in a varying degree of impacts on prediction performance of U-Net and GAN models in the task of CR bone signal suppression. For the U-Net model, flips, rotation (10 to 20 degrees), all the shift, and zoom (1/8) resulted in improved model prediction accuracy. By contrast, other studied augmentation techniques showed adverse impacts on the model performance. For the GAN model, it was found to be more sensitive to the studied augmentation techniques than the U-Net. Vertical flip was the only augmentation method that yielded enhanced model performance. However, it remains challenging to determine the optimal parameter settings for each augmentation technique. In the future, a more comprehensive evaluation is still warranted to evaluate the effectiveness of different augmentation techniques in task-specific image synthesis.

## Acknowledgement

## References

1. Sung H, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021; 71: 209-249.

2. Vansteenkiste J, et al. Randomized controlled trial of resection versus radiotherapy after induction chemotherapy in stage IIIA-N2 non-small cell lung cancer. Journal of Thoracic Oncology. 2007; 2: 684-685.

3. Albain KS, et al. Radiotherapy plus chemotherapy with or without surgical resection for stage III non-small-cell lung cancer: a phase III randomised controlled trial. Lancet. 2009; 374: 379-386.

4. Chen GT, GC Sharp and S Mori. A review of image-guided radiotherapy. Radiol Phys Technol. 2009; 2: 1-12.

5. Korreman S, et al. The European Society of Therapeutic Radiology and Oncology-European Institute of Radiotherapy (ESTRO-EIR) report on 3D CT-based in-room image guidance systems: A practical and technical review and guide. Radiotherapy and Oncology. 2010; 94: 129-144.

6. Xiao Y. Image-Guided Radiation Therapy (IGRT): kV Imaging. In Encyclopedia of Radiation Oncology. LW Brady and TE Yaeger, Editors. 2013, Springer Berlin Heidelberg: Berlin, Heidelberg. 343-351.

7. Stanley DN, et al. Comparison of Initial Patient Setup Accuracy between Surface Imaging and Three Point Localization: A Retrospective Analysis. Medical Physics. 2017; 44.

8. Tanaka R, et al. Improved accuracy of markerless motion tracking on bone suppression images: preliminary study for image-guided radiation therapy (IGRT). Phys Med Biol. 2015; 60: N209-218.

9. Lehmann L, et al. Generalized image combinations in dual KVP digital radiography. Medical physics. 1981; 8: 659-667.

10. Chen S and K Suzuki. Separation of bones from chest radiographs by means of anatomically specific multiple massive-training ANNs combined with total variation minimization smoothing. IEEE transactions on medical imaging. 2013; 33: 246-257.

11. Loog M, B van Ginneken and AM Schilham. Filter learning: application to suppression of bony structures from chest radiographs. Medical image analysis. 2006; 10: 826-840.

12. Suzuki K, et al. Image-processing technique for suppressing ribs in chest radiographs by means of massive training artificial neural network (MTANN). IEEE Transactions on medical imaging. 2006; 25: 406-416.

13. Yang W, et al. Cascade of multi-scale convolutional neural networks for bone suppression of chest radiographs in gradient domain. Medical Image Analysis. 2017; 35: 421-433.

14. Zarshenas A, et al. Separation of bones from soft tissue in chest radiographs: Anatomy-specific orientation-frequency-specific deep neural network convolution. Medical Physics. 2019; 46: 2232-2242.

15. Chen S, et al. Enhancement of chest radiographs obtained in the intensive care unit through bone suppression and consistent processing. Physics in Medicine & Biology. 2016; 61: 2283.

16. Suzuki K, et al. Suppression of the contrast of ribs in chest radiographs by means of massive training artificial neural network. In Medical Imaging 2004: Image Processing. International Society for Optics and Photonics. 2004.

17. Sun C, et al. Revisiting unreasonable effectiveness of data in deep learning era. In Proceedings of the IEEE international conference on computer vision. 2017.

18. Halevy A, P Norvig and F Pereira. The unreasonable effectiveness of data. 2009; 24: 8-12.

19. Shorten C and TM Khoshgoftaar. A survey on image data augmentation for deep learning. 2019; 6: 60.

20. Simard PY, D Steinkraus and JC Platt. Best practices for convolutional neural networks applied to visual document analysis. In Icdar. 2003.

21. Krizhevsky A, I Sutskever and GE Hinton. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems. 2012.

22. Shin HC, et al. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. IEEE Trans Med Imaging. 2016; 35: 1285-1298.

23. Shin HC, et al. Medical Image Synthesis for Data Augmentation and Anonymization Using Generative Adversarial Networks. Simulation and Synthesis in Medical Imaging. 2018; 11037: 1-11.

24. Salamon J and JP Bello. Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification. IEEE Signal Processing Letters. 2017; 24: 279-283.

25. Hussain Z, et al. Differential Data Augmentation Techniques for Medical Imaging Classification Tasks. AMIA Annu Symp Proc. 2017; 2017: 979-984.

26. Clark K, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. Journal of digital imaging. 2013; 26: 1045-1057.

27. Zhao B, et al. Evaluating variability in tumor measurements from same-day repeat CT scans of patients with non-small cell lung cancer. Radiology. 2009; 252: 263-272.

28. Hofmanninger J, et al. Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem. 2020. arXiv: 2001.11767.

29. Oktay O, et al. Attention U-Net: Learning Where to Look for the Pancreas. arXiv e-prints. 2018; arXiv: 1804.03999.

30. Goodfellow IJ, et al. Generative Adversarial Nets. Advances in Neural Information Processing Systems. 2014; 27: 27.

31. Moreno-Barea FJ, et al. Forward Noise Adjustment Scheme for Data Augmentation. 2018 IEEE Symposium Series on Computational Intelligence (IEEE Ssci). 2018: 728-734.

32. Wang Z, et al. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process. 2004; 13: 600-612.

33. Liu HD, XF Gu and D Samaras. Wasserstein GAN with Quadratic Transport Cost. 2019 IEEE/CVF International Conference on Computer Vision (ICCV 2019). 2019: 4831-4840.