# Synthetic nodule generation using one-shot generative learning for CXR nodule detection

Guillaume Sallé, Dimitris Visvikis and Vincent Jaouen

UMR 1101 Inserm LaTIM, UBO, IMT Atlantique, Brest, France
guillaume.salle@univ-brest.fr
https://github.com/GuigzoS/NODE21_NoduleGeneration

**Abstract.** Lung cancer is the most diagnosed and deadliest cancer worldwide. Early stage is characterized by pulmonary nodules that can be detected using computed tomography (CT) or at a cheaper technical cost using chest radiography (CXR). However, training a detection neural network requires large numbers of diseased patients that may not represent the complete diversity of possible cases during model deployment. In the context of the generation track of the NODE21 challenge, we propose a method for nodule synthesis based on a deep generative multi-scale cascaded network trained on a single image that learns how to compose projections of volumetric CT nodules with the style of a CXR image. The proposed approach produces images with realistic appearance that corrects for some unrealistic visual artefacts produced by the competitor baseline model.

**Keywords:** Nodule Generation · One-shot Learning · Domain adaptation

## 1 Introduction

Generative adversarial networks (GAN) have achieved outstanding performances in various fields of medical image analysis including image segmentation, cross-modality image synthesis, super-resolution or denoising [1]. Due to their generative nature, they are also often considered for data augmentation in machine learning-based models to address data scarcity, a recurring issue in medical imaging [2]. However, most GAN-based methods rely on large training sets in order to expose both the generator and the discriminator to sufficient number of examples in order to reach Nash equilibrium. Departing from these usual requirements, the one-shot GAN SinGAN model [3] has recently emerged as a new paradigm for deep generative learning using multiple adversarial generators in a cascaded multi-scale architecture. SinGAN enables to learn the rich complexity of natural images using as little as one training image. It has been applied to various computer vision tasks including super-resolution or style transfer, but has not yet been considered for realistic pathology insertion in medical imaging. In the context of the NODE21 lung nodule generation challenge, we propose to explore to what extent a SinGAN model trained on a single CXR image can be used for realistic nodule synthesis to facilitate the training of a downstream nodule detection model. We show that this approach partly addresses some visual limitations of the current baseline competitor approach based on Poisson image editing [4, 5]. In particular, we generally achieve smoother transitions between nodule and background.

## 2 Method

Our method is based on a 2D one-shot multi-stage generative SinGAN model [3]. Given a single image, SinGAN learns the image distribution at $N$ different scales using $N$ scale-specific generators trained successively in a coarse-to-fine fashion. The general principles of the SinGAN model are summarized in Fig. 1.

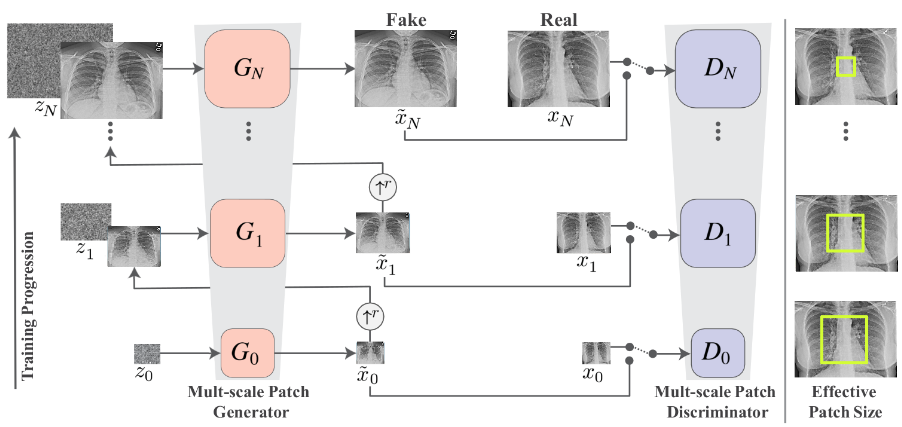


**Fig. 1.** SinGAN training pipeline. The image is initially downsampled to a very low scale, numbered 0. At each scale $n$, the generator $G_n$ learns to synthesize realistic image patches. It sharpens the output from the previous scale $n-1$ (or composes the image for $n=0$), while the discriminator $D_n$ learns to distinguish real and generated samples. After training scale $n$, the result is upsampled by factor $r$ for the next scale. Modified from [3]

In this work, we use the SinGAN architecture to achieve realistic style transfer between projected CT nodule patches and a 2D CXR image. First, we naively paste a 2D CT nodule on a CXR image at the desired location. We then learn a SinGAN model by fixing a scale $n_0$ and applying all generators linked to above scales $n \geq n_0$ successively on the pasted nodule to produce a globally harmonized CXR image containing a nodule.

Given the projective nature of chest radiography, this suppresses useful attenuation information from other regions outside the nodule along the X-ray beam. As a workaround, after CT nodule harmonization, we add back the original image to the SinGAN output using a weighted average depending on nodule location.

In our experiments we noticed that nodules in lower regions of the lungs and therefore on average appearing as hypersignals due to higher attenuation can be better harmonized using a separated SinGAN model trained on images cropped within this higher intensity background. The decision whether to apply one model or the other is based on Otsu thresholding the CXR image into four classes. The higher intensity SinGAN model is then selected when a majority of voxels in the bounding box belongs to the higher intensity class. The CT nodule mean intensity value is then adjusted using a contrast matching method provided in the generation baseline. A scalar is calculated as the maximum between the logarithm of intensity ratio $I_{nodule}/I_{cxr}$ and a threshold value (0.6 for hyposignal nodules, 0.7 for hypersignal ones). The nodule is finally multiplied by this scalar before being naively pasted on the CXR prior to SinGAN harmonization.

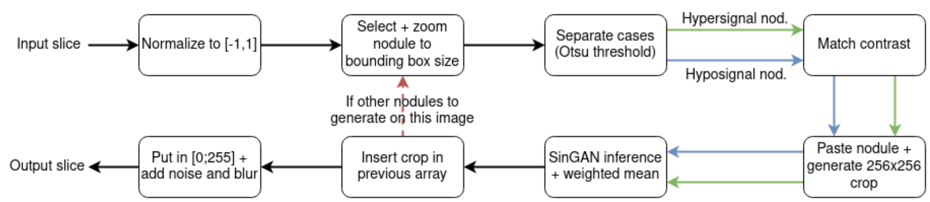


**Fig. 2.** General workflow of the proposed approach.

## 3 Experiments and results

The organizers of NODE21 challenge provided 4882 CXR images, respectively 1134 with nodules and 3748 healthy, all of them of size $1024 \times 1024$. A set of 1186 volumetric CT nodules with corresponding segmentation masks was also available. To evaluate performance regarding to state-of-the-art method, a generation baseline based on Poisson image editing [4] was also provided.

To build our training set, we first randomly selected two images corresponding to the aforementioned hypersignal and hyposignal cases. We then selected randomly 1200 healthy images to perform nodule generation : 1000 for training our future evaluation detection network, 200 for validation. We finally built a detection testing set composed of 200 healthy CXR (different from the first 1200) and 200 containing real nodules.

For time efficiency purpose, we processed all 3D CT nodules prior and store them for any further generation. All nodules are extracted according to their masks and the function *utils.generate_2d* provided in the generation baseline is applied on each of them. Each 2D array is finally stored as numpy array with its corresponding mask. A CSV file containing all nodule diameters, computed in the same way as the baseline, is also saved.

To avoid overfitting, we added Gaussian noise with $\mu = 0$ and $\sigma = 3$ and Gaussian blur of scale 3 pixels. Figure 2 summarizes nodule generation process. The nodule to be harmonized during model inference was selected exactly in the same way as the generation baseline by randomly picking among nodules having a diameter close to the the bounding box size. This nodule was then resampled to the desired bounding box size without any distortion. For GPU capacity issues, we could not train a SinGAN model on a $1024 \times 1024$ CXR. Given a training image containing a nodule, we cropped a $256 \times 256$ square near the nodule and use it for training.

Regarding SinGAN parameters, we varied the ratio factor between each GAN scale (from 0.75 to 0.85, creating 16 scales instead of 9) for smoother transition between scales. We also modified the kernel size (from 7 to 5) to avoid artefacts or strong blurry effects. All other parameters were kept unchanged with respect to the original SinGAN model.

We trained the SinGAN model on $256 \times 256$ CXR crops around the pasted nodule. The same processing was performed for the mask, as both masks and image are required by the model. After inference, the original image was reconstructed using the harmonized cropped image. If another nodule needed to be generated on the same CXR, we selected a new nodule and repeated the entire process.

A baseline faster-RCNN detection model was trained to adjust model parameters in-house. We tested them on our testing set composed of 400 real images and evaluated AUC on predictions. The maximum scores we obtained for the generation baseline and for our method are summarized in Table 1, which shows AUC results slighly reduced compared to baseline model.

Visually, our method appears nevertheless often more satisfying. Fig. 3 shows some cherry-picked generated samples for the baseline method and ours. In general, the generated nodules for both methods seem realistic. However, the baseline image often produces unnatural artefacts consisting of sharp transitions between object and background. The proposed approach generally allows for a smoother merging of the object to its background without such artefacts.

**Table 1.** Maximum AUC obtained on real data.

| Metric | Baseline | Our method |
|---|---|---|
| AUC | $\mathbf{0.71 \pm 0.03}$ | $0.68 \pm 0.02$ |

## 4 Conclusion

In this study, we proposed a new method for the fusion of synthetic nodules onto healthy CXR images based on a generative adversarial network trained on a single image. The proposed approach produces realistic images that allow for the training of a detection network applied on real images. Further experiments

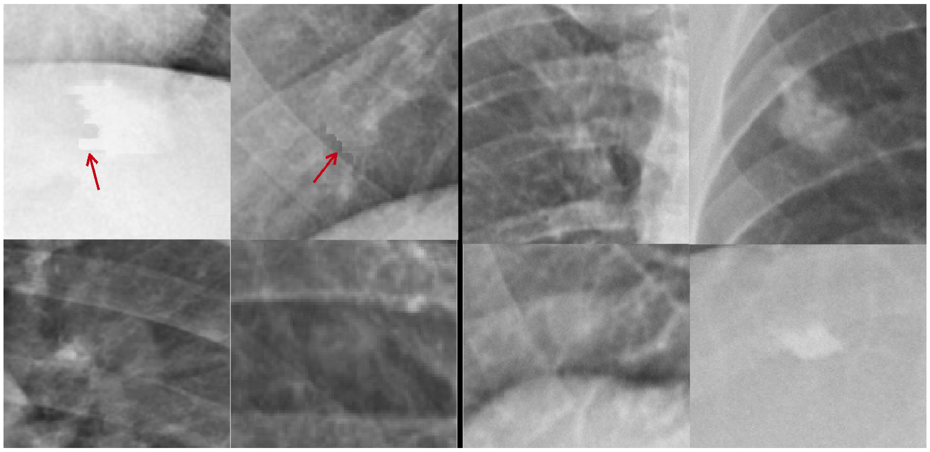

**Fig. 3.** Visual harmonization results for baseline (columns 1 and 2) and the proposed approach (columns 3 and 4). Baseline results often show unnatural artefacts (top row, indicated with red arrows). On the other hand, current parameters for our approach may make nodules appear too visible (bottom row)

will focus on adjusting model parameters, which are clearly sub-optimal at the time of submission. The realistic features in the generated images support the potential interest of the method for improving the generalization power of nodule detection models, especially in combination with other approaches based on other principles.

## References

1. Xin Yi, Ekta Walia, and Paul Babyn, "Generative adversarial network in medical imaging: A review," *Medical image analysis*, vol. 58, pp. 101552, 2019.
2. Junzhao Liang and Junying Chen, "Data augmentation of thyroid ultrasound images using generative adversarial network," in *2021 IEEE International Ultrasonics Symposium (IUS)*, 2021, pp. 1–4.
3. Tamar Rott Shaham, Tali Dekel, and Tomer Michaeli, "Singan: Learning a generative model from a single natural image," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 4570–4580.
4. Geert JS Litjens, Laurens Hogeweg, Arnold MR Schilham, Pim A de Jong, Max A Viergever, and Bram van Ginneken, "Simulation of nodules and diffuse infiltrates in chest radiographs using ct templates," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2010, pp. 396–403.
5. Patrick Pérez, Michel Gangnet, and Andrew Blake, "Poisson image editing," in *ACM SIGGRAPH 2003 Papers*, pp. 313–318. 2003.