Synthetic Lung Nodule 3D Image Generation Using Autoencoders

Steve Kommrusch, Louis-Noël Pouchet
Colorado State University, Fort Collins Colorado
{steveko,pouchet}@colostate.edu

Abstract

One of the challenges of using machine learning techniques with medical data is the frequent dearth of source image data for training. A representative example is automated lung cancer diagnosis, where nodule images need to be classified as suspicious or benign. In this work we propose an automatic synthetic lung nodule image generator. Our 3D shape generator is designed to augment the variety of 3D images. Our proposed system takes root in autoencoder techniques, and we provide extensive experimental characterization that demonstrates its ability to produce quality synthetic images.

1 The LuNG System

The LuNG (Lung Nodule Generator) system we developed is based on a neural network trained to produce realistic 3D lung nodule images from a small set of seed examples to help improve automated cancer screening.

Our work is aimed at creating synthetic images in cases where input images are difficult to get. For example, the Adaptive Lung Nodule Screening Benchmark (ALNSB) from the NSF Center for Domain-Specific Computing uses a flow that leverages compressive sensing to reconstruct images from low-dose CT scans. These images are slightly different than those built from filtered backprojection, a technique which has more samples readily available (such as LIDC/IDRI). To evaluate our results, we integrate our work with the ALNSB system ([Shen *et al.*, 2015]) that computes features on nodule images and classifies each nodule as benign or suspicious. We use original patient data to train LuNG, and then use LuNG to generate synthetic nodules that are processed by ALNSB. We create a network which optimizes 3 metrics: (1) increase the percentage of generated images accepted by the nodule analyzer; (2) increase the variation of the generated output images relative to the limited seed images; and (3) decrease the error of the seed images with themselves when input to the autoencoder.

To begin, guided training is used in which each nodule is modified to create 15 additional training samples. We call the initial nodule set, of which we were provided 51 samples, the 'seed' nodules, examples of which are shown in figure 2. The 'base' nodules include 15 modified samples per seed nodule for a total of 816 samples. The base nodules are used to train an autoencoder neural network with 3 latent feature neurons in the bottleneck layer. The output of the autoencoder goes through a cleanup algorithm to increase the likelihood that viable fully connected nodules are being generated. A nodule analyzer program then extracts relevant 3D features from the nodules and prunes away nodules outside the range of interesting feature values; these nodules are the final output of LuNG. We use the ALNSB ([Shen *et al.*, 2015]) nodule analyzer and classifier code for the LuNG project, but similar analyzers compute similar 3D features to aid in classification. To evaluate generated nodules, we develop a statistical distance metric similar to the Mahalanobis distance. Given the set of nodules output by LuNG (which can be used to augment image classifiers), we explore adding them to the autoencoder training set to improve the generality of the generator. In our full 17-page paper we evaluate various interface options and network hyperparameters.

While the LuNG use model relies on having both an encoder and decoder network as provided by autoencoder training, future work could merge our technique with a generative adversarial network to enhance the generator or test whether convolution/deconvolution layers can help improve our overall quality metrics ([Li *et al.*, 2017]).

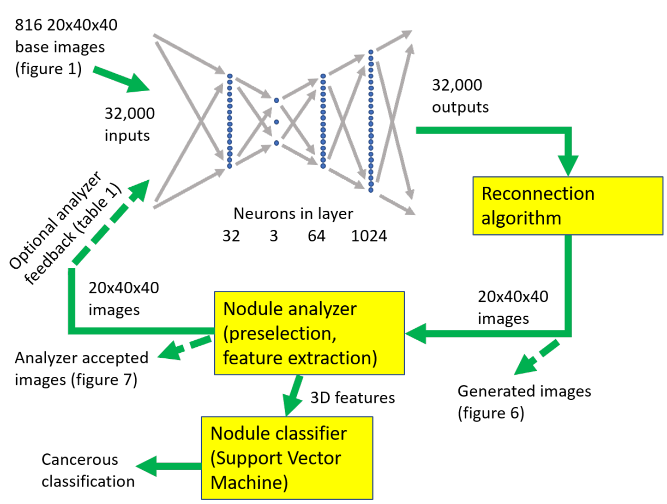

Figure 1: Interaction between autoencoder, nodule analyzer, and support vector machine. Figure numbers are in reference to figures in our full 17-page paper.

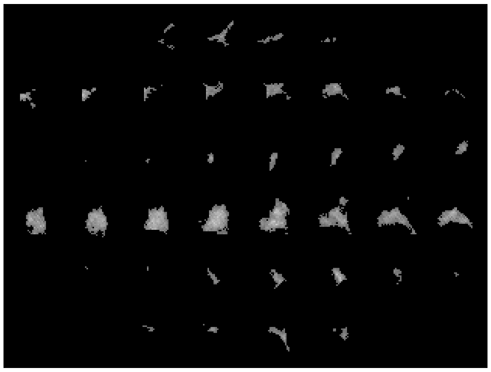

Figure 2: 6 of 51 original seed nodules showing middle 8 2D slices of 3D image from the CT scan

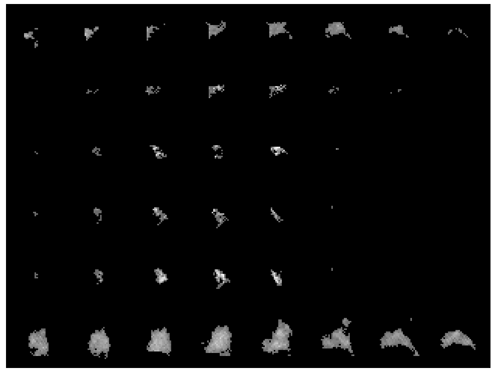

Figure 3: Generated images of 6 steps through latent feature space from seed nodules 2 to 4.

Analyzer feature	Ratio
2D area	1.1
2D max(xL,yL)	1.0
2D perimeter	1.2
2D area/peri^2	0.8
3D volume	1.3
3D rad/MeanSqDis	1.0
min(xL,yL)/max(xL,yL)	1.0
minl/maxl	1.0
surface area^3/volume^2	1.2
mean breadth	1.3
euler3D	1.1
maskTem area/peri^2	1.0

Table 1: Feature means ratios for 400 generated vs 51 seed nodules

2 Autoencoder network, analyzer, and optional image feedback

Figure 1 shows the general structure of the LuNG system. The autoencoder can be split into both an encoder and generator network for different use models. For example, given 2 seed nodules one can use the feature network to find their latent space coordinates and then step from one nodule to another with inputs to the generator network. The reconnection algorithm insures that all outputs of the network result in plausible nodule shapes for further analysis. The nodule analyzer checks that the nodules fit within the legal range used by the classifier to process actual CT scan candidates. The support vector machine is an example classifier to which LuNG can provide augmented data. Such augmented data is helpful in overcoming drawbacks in current lung nodule classification work ([Valente *et al.*, 2016]).

3 Results

By using the trained encoder network to find the latent feature coordinates for seed nodules 2 and 4, figure 3 shows 6 steps between these nodules. The top and bottom nodules in the image can be seen to accurately reproduce seed nodules 2 and 4 from figure 2. The 4 intermediate nodules are novel images from the generator which can be used to improve an automated classifier system.

As an example of the scoring metrics we use to evaluate networks architectures and interface options, we analyze the 12 3D features for nodules computed by the nodule analyzer (ALNSB). Table 1 shows that when 400 novel random images are generated by LuNG, the mean feature value for all 12 3D features stays within 30% of the seed nodules. Based on this alignment, we plot SVM distance values for 1,000 nodules and the 51 seed nodules in figure 4. After the support vectors are applied, nodules closer to the positive than the negative centroid are classified as suspicious. The results show that LuNG generated nodules augment the available nodules for analysis well, including providing many nodules which are near the existing boundary and can be useful for improving the sensitivity of the classifier.

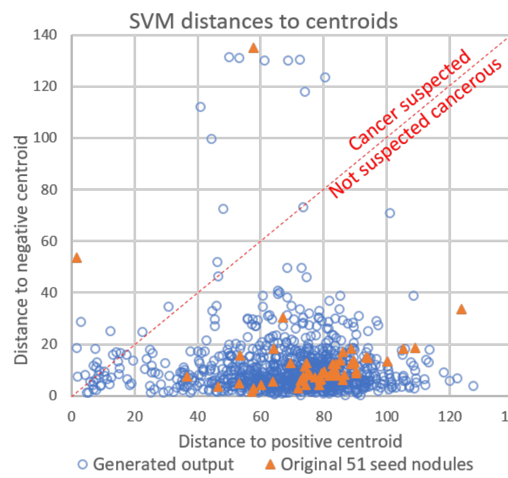

Figure 4: Support Vector Machine (SVM) coordinates for 1,000 generated and 51 seed nodules

Acknowledgments

This work was supported in part by the U.S. National Science Foundation award CCF-1750399.

References

[Li *et al.*, 2017] Jun Li, Kai Xu, Siddhartha Chaudhuri, Ersin Yumer, Hao Zhang, and Leonidas Guibas. GRASS: Generative Recursive Autoencoders for Shape Structures. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2017)*, 36(4), 2017.

[Shen *et al.*, 2015] Shiwen Shen, Prashant Rawat, Louis-Noel Pouchet, and William Hsu. Lung Nodule Detection C Benchmark, 2015.

[Valente *et al.*, 2016] Igor Rafael S. Valente, Paulo César Cortez, Edson Cavalcanti Neto, José Marques Soares, Victor Hugo C. de Albuquerque, and João Manuel R.S. Tavares. Automatic 3D Pulmonary Nodule Detection in CT Images. *Comput. Methods Prog. Biomed.*, 124(C):91–107, February 2016.