ILLINOIS
Bioengineering Department
GRAINGER COLLEGE OF ENGINEERING

Evaluating Generative Adversarial Networks as Stochastic Object Models

Dimitrios Gotsis

November 17th, 2021

Outline

- Basics of Medical Imaging and Imaging Science
  - Problem Set Up
  - Image Quality (IQ) Assessment
  - Stochastic Object Models (SOMs)

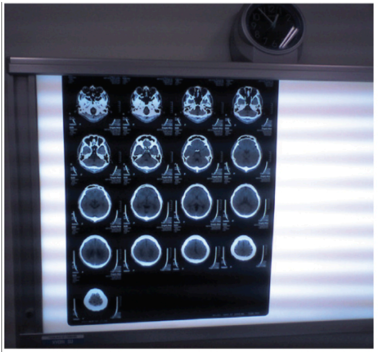
- Generative Adversarial Networks (GANs)
  - Basics
  - GAN Evaluation

- Virtual Imaging Trials (VITs)

- Numerical Studies and Results
  - Ultrasound Simulation
  - Results
  - Discussion
  - Future Work and Conclusion

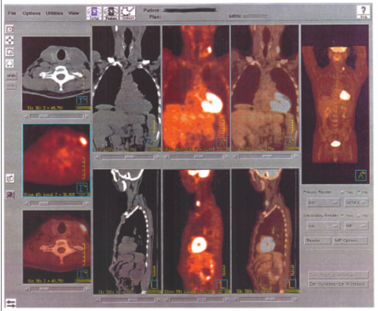
Medical Imaging Basics

- The purpose of medical imaging is to collect some information of internal organs or tissue in order to aid a diagnosis (e.g. lesion present or absent).

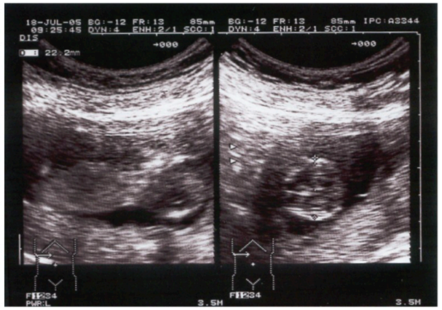


(a)

(b)



(c)



(d)

- However, the information we collect (measurements) are not a direct representation of the object (actual thing in space, e.g. brain) we are imaging and contain less information.

Medical Imaging Basics (cont.)

- The imaging process can be mathematically described as $y = Hx + n$ where y are our measurements, x is the object, n is noise and H is the forward operator.

- As mentioned before, information is lost in the imaging process and this is due to H having a null space.

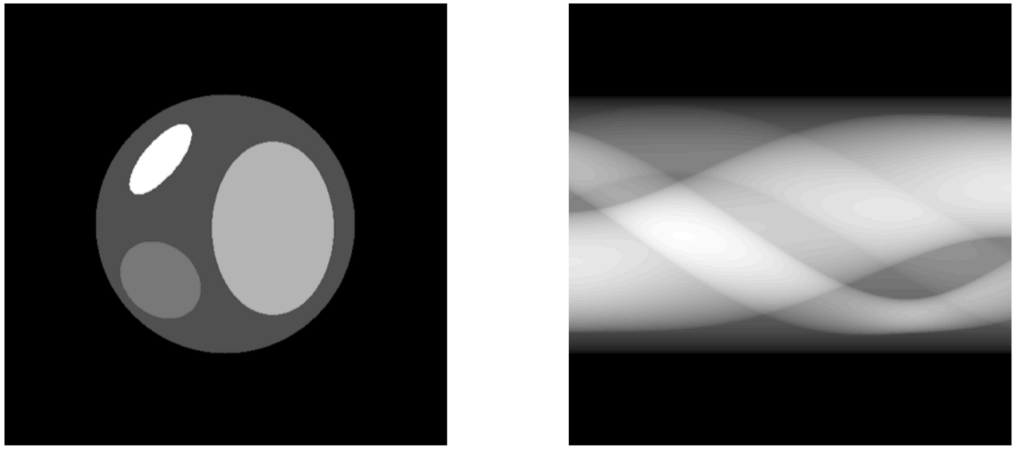


Figure: Object (left), Measurement (right)

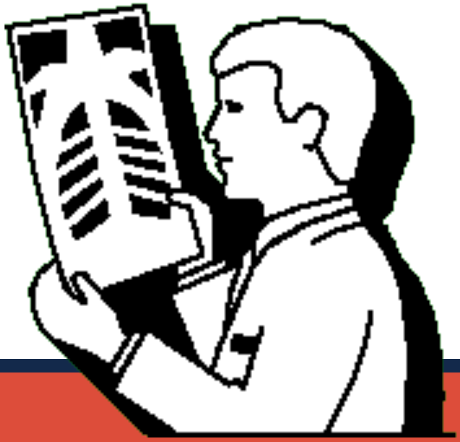
Image Quality (IQ) Assessment

- Missing information can lead to an incorrect diagnosis and since the purpose of medical imaging is to aid in a diagnosis, it is natural for us to consider task based image quality (IQ) for evaluating and optimizing imaging systems. Our observer for the task will either be a person and/or a machine.

- However, task based IQ requires information about sources of randomness in the measurement data.

- Randomness could come from measurement noise or object variation.

Optimize IQ using task-based measures



Generative Models and SOMs

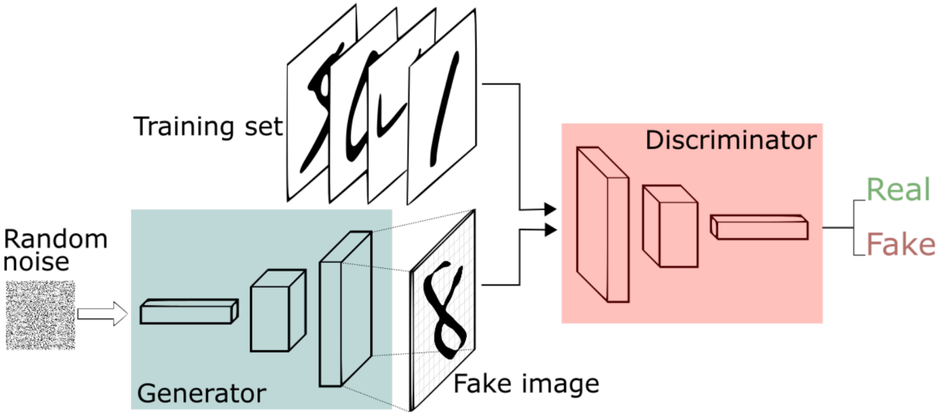
- Stochastic Object Model (SOM): defined as a mathematical or computational model which describes the randomness in a distribution of objects we are interested in imaging.

- Generative Model:
  - Estimates distribution of a dataset using either an explicit or implicit representation.
  - Once trained, the generative model can be sampled to produce realizations from the estimated distribution.

- A trained generative model on object data can thus be used to represent a SOM.

Generative Adversarial Networks (GANs)

- Introduced by Goodfellow et. al (2014), GANs are generative models which implicitly estimate the data distribution
  - Consist of a generator and discriminator which compete during training.
  - The generator tries to produce images that come from the data distribution, while the discriminator tries to decided whether an incoming image is from the training set (real) or from the generator (fake).
- More recent improvements to Goodfellow's GAN include StyleGAN2 and ProGAN



Credit: Thalles Silva (freecodecamp.org)

GAN Evaluation

- So do how do we know if GANs are ‘good’ SOMs?

- What statistics do GANs learn?
  - Important since our observers can be machines and thus different statistics between the real and fake images can lead to different task based results.
  - Dataset specific statistics specifically investigated as well as general first and second order statistics (e.g. histogram, autocorrelation).

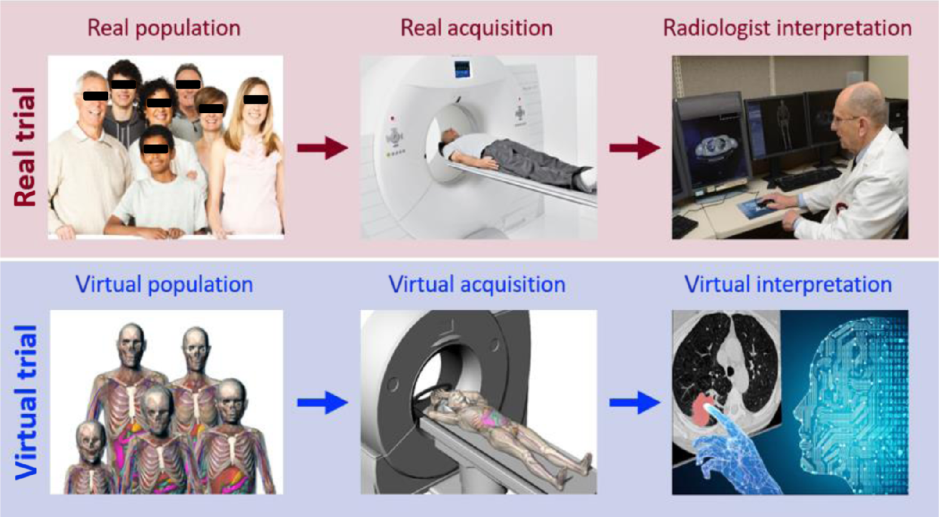
- How does it perform for task based evaluations?
  - Lesion detection
  - Benign vs. Malignant

- How do we put together our dataset for our GAN?

Virtual Imaging Trials (VITs)

- Many times we do not have access to large medical image datasets due to ethical limitations, expense, or time requirements.

- Thus, we can use virtual image trials (VITs) which simulate the whole imaging process virtually (i.e. patients, imaging systems, and interpreters) [1].



[1] Ehsan Abadi, William P. Segars, Benjamin M. W. Tsui, Paul E. Kinahan, Nick Bottenus, Alejandro F. Frangi, Andrew Maidment, Joseph Lo, Ehsan Samei, "Virtual clinical trials in medical imaging: a review," J. Med. Imag. 7(4) 042805 (11 April 2020) https://doi.org/10.1117/1.JMI.7.4.042805

Ultrasound (US) Simulation

- From Wear et al. (1997):
  - Choose Scatterers per Number Density (SND). Then choose the total area, carrier frequency and wave velocity.
  - Poisson distribute the scatterers over the image using a 2D uniform distribution.
  - With each scatterer apply $\exp\left(\frac{j2\pi x}{\lambda}\right)$.
  - Apply a 2D Gaussian blur over the image. The standard deviations are determined by the wavelength λ and the device dimensions [2]
- The resulting image is the envelope $|E|$ and the intensity is $I = |E|^2$.
- The signal to noise ratio of a US image is $SNR^2 = \left(\frac{\mu_I}{\sigma_I}\right)^2$ where μ_I is the intensity mean and σ_I is the standard deviation of the intensity.

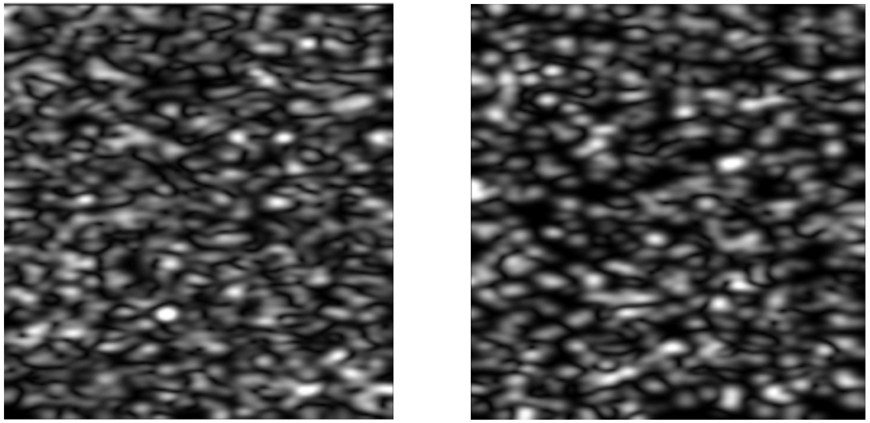


Figure: Example US simulated images

[2] K. A. Wear, R. F. Wagner, D. G. Brown, and M. F. Insana, "Statistical properties of estimates of signal-to-noise ratio and number of scatterers per resolution cell," The Journal of the Acoustical Society of America, vol. 102, no. 1, pp. 635–641, 1997. DOI: 10.1121/1.419738. eprint: https://doi.org/10.1121/1.419738. [Online]. Available: https://doi.org/10.1121/1.419738

Numerical Studies - Ultrasound

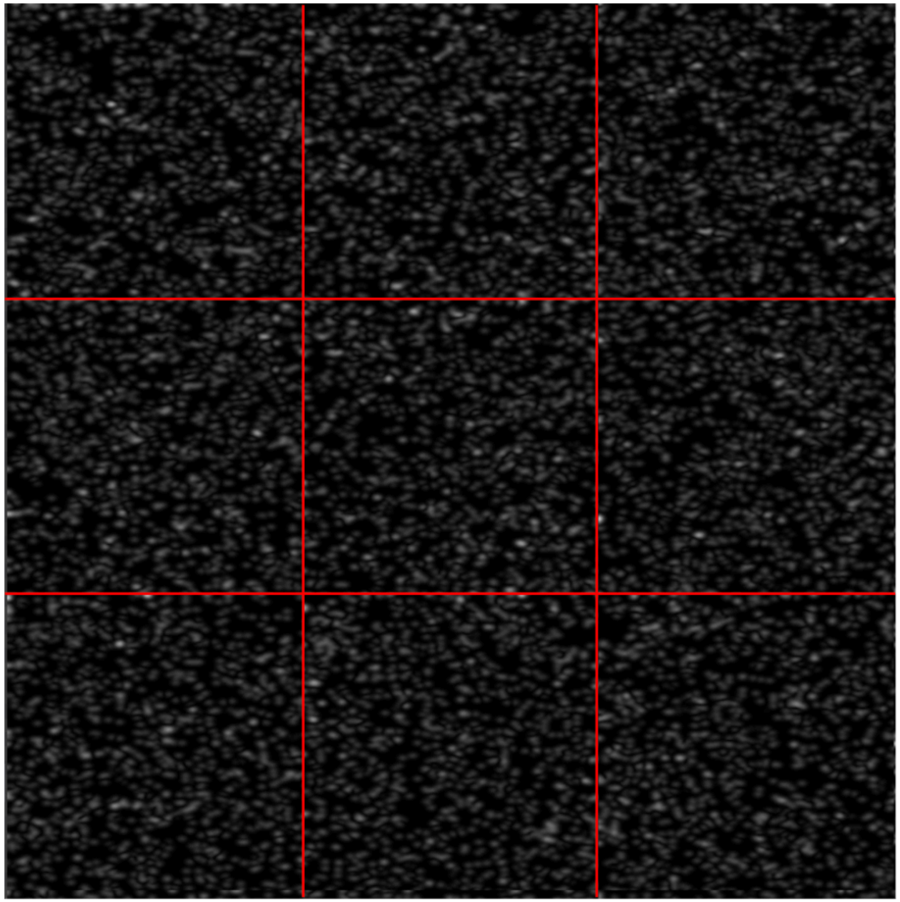
- Since the Ultrasound simulation requires convolution by a point spread function (PSF), this is an image and thus the following studies are a discussion on Stochastic Image Models (SIMs).

- Datasets
  - Fully Developed Speckle (SND = 30 dataset)
  - Scatterers per Number Density (SND = 1, 3 and 30)
  - Frequency set to 3.5 MHz for all three datasets

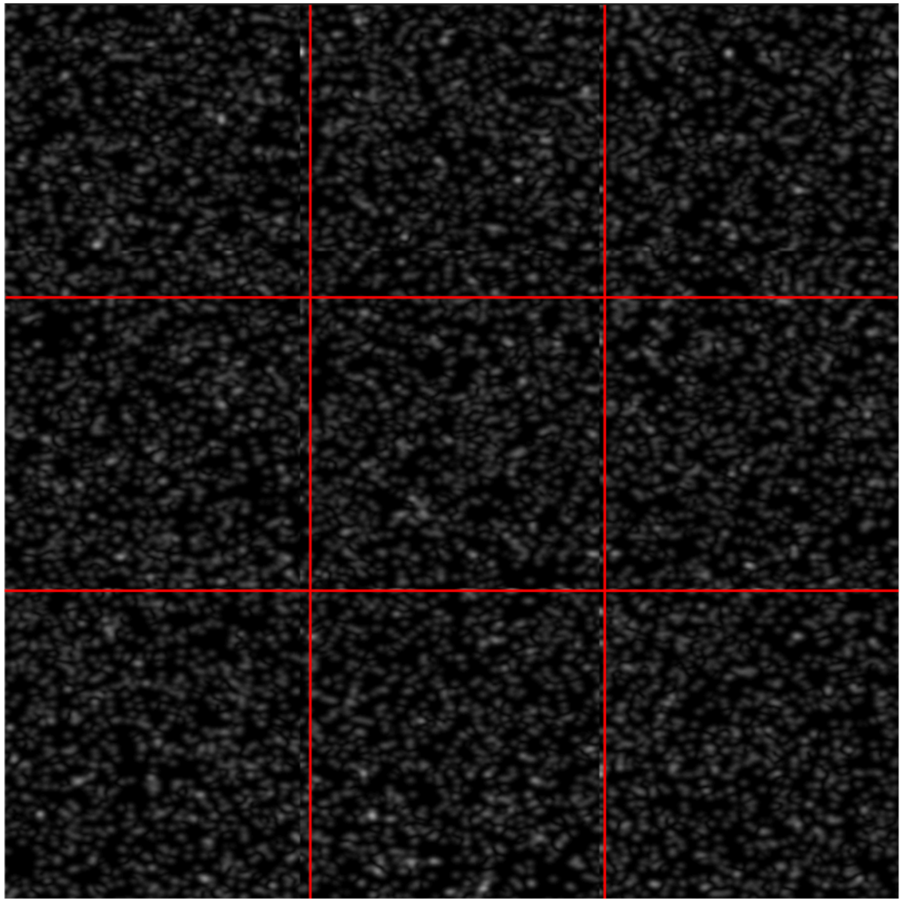
- Statistics Computed
  - Ensemble Envelope Histogram (Fully developed Speckle) (follows Rayleigh Distribution)
  - Ensemble Intensity Histogram (Fully developed Speckle) (follows Exponential Distribution)
  - Autocorrelation (Fully Developed Speckle)
  - Meaningful metric: SNR^2 (used to describe speckle type)

- The resulting distributions are compared between real and fake images. 20,000 real and fake images were used to calculate the above statistic distributions.

Results - SND-1

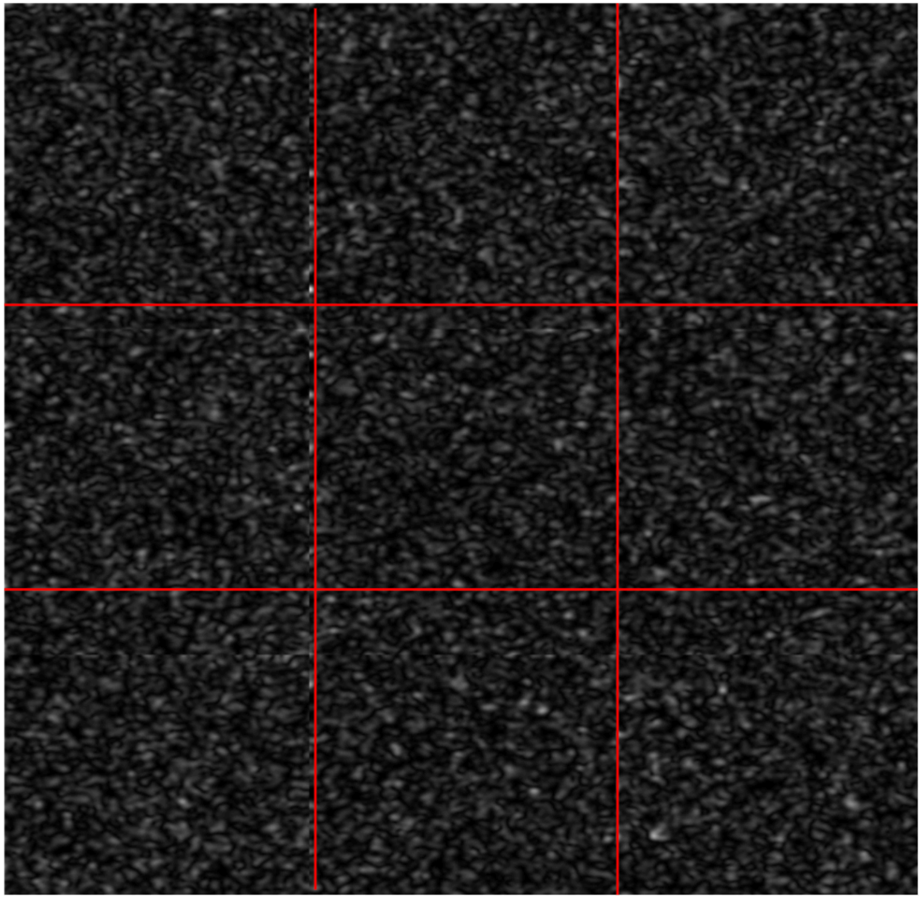


Reals

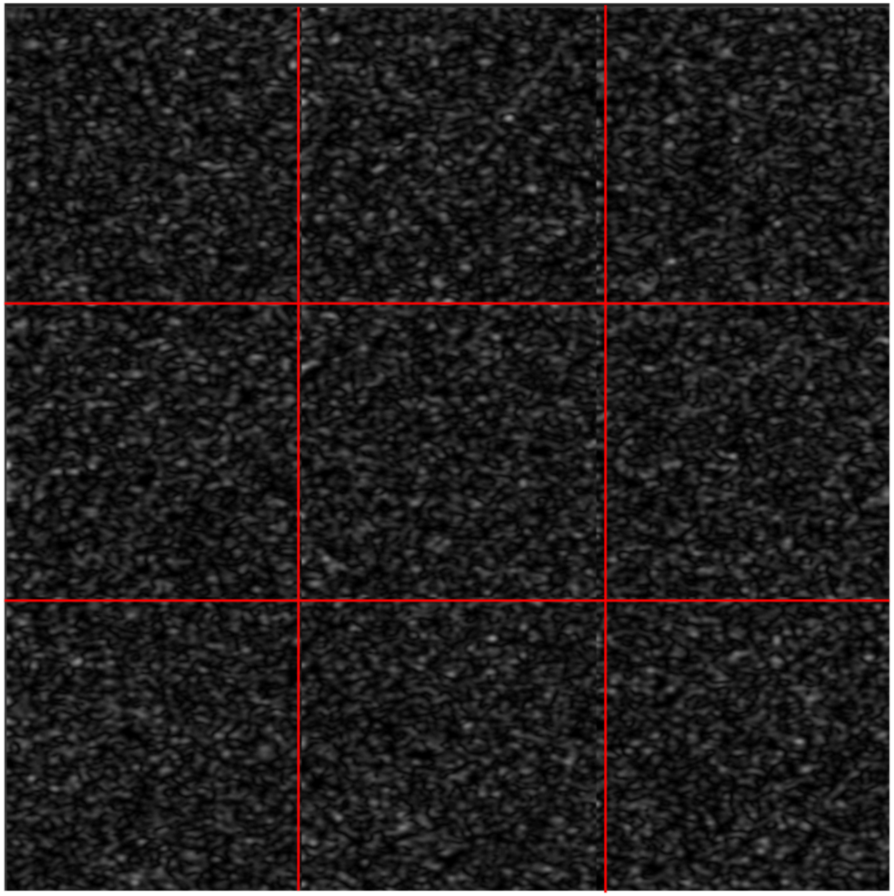


Fakes

Results - SND-3

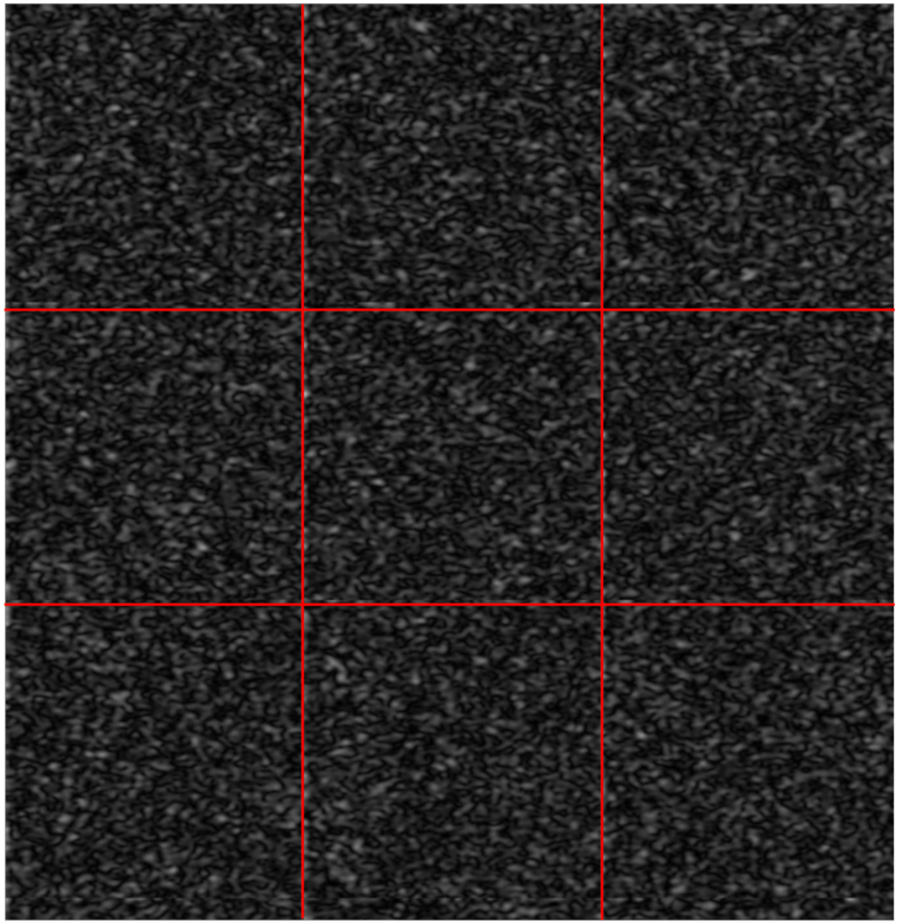


Reals

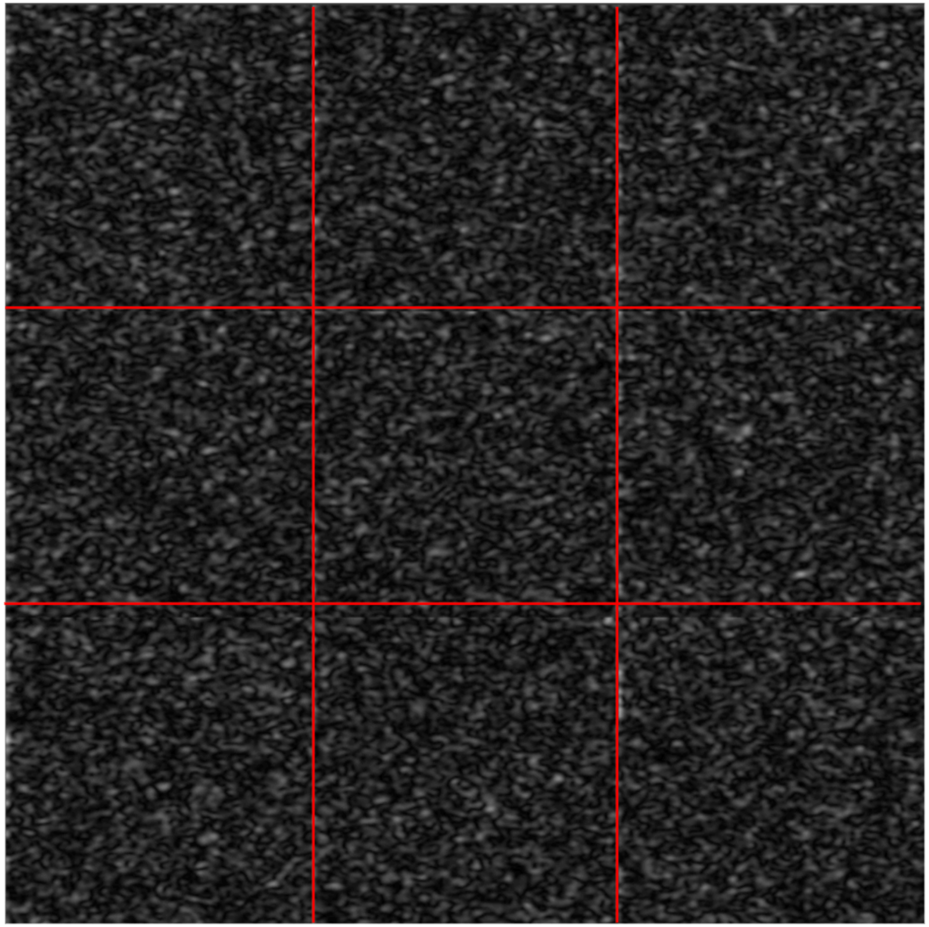


Fakes

GAN Results - SND-30

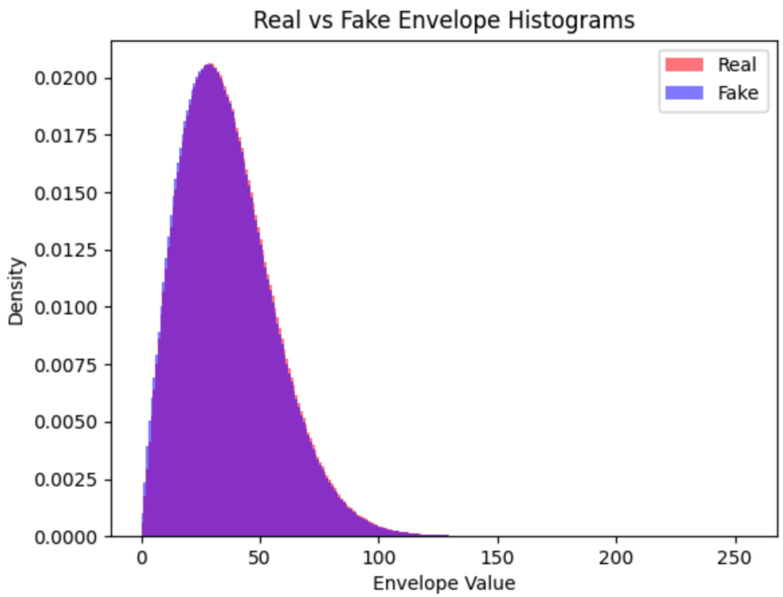


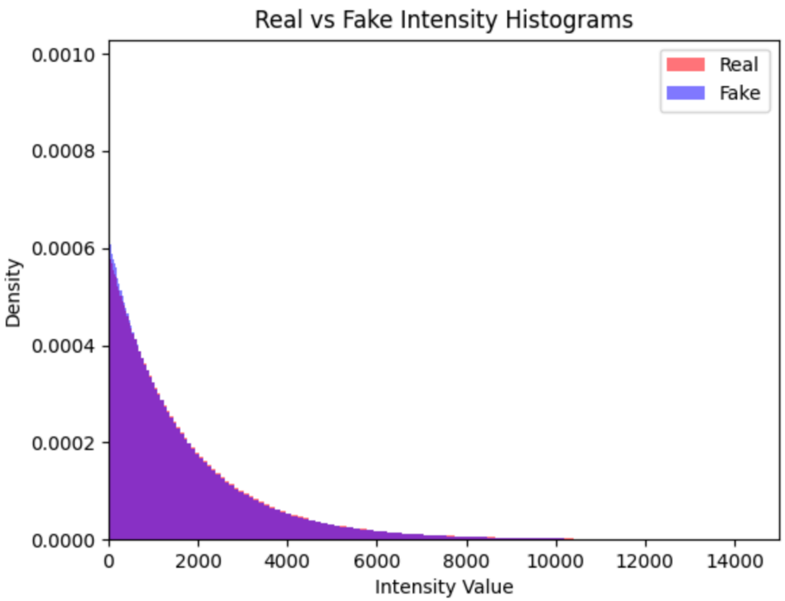
Reals

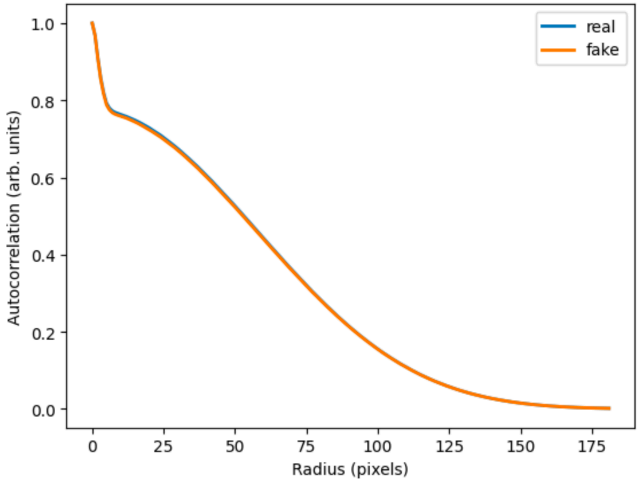


Fakes

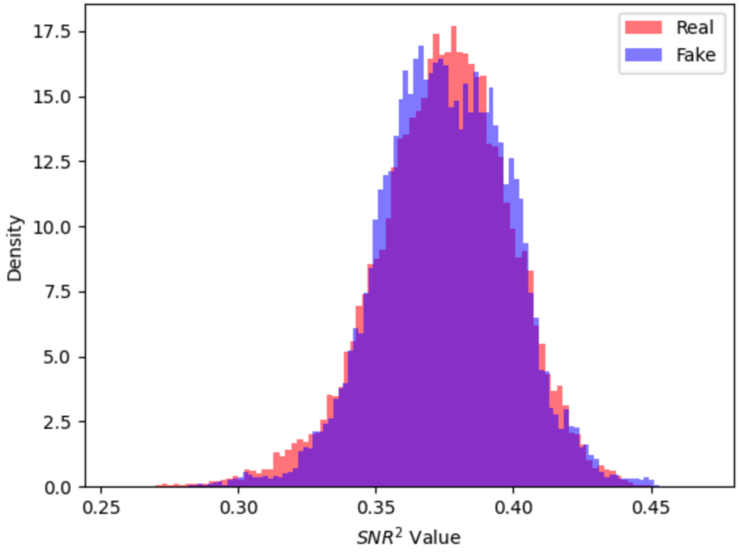
SND-30 Envelope, Intensity and Autocorrelation

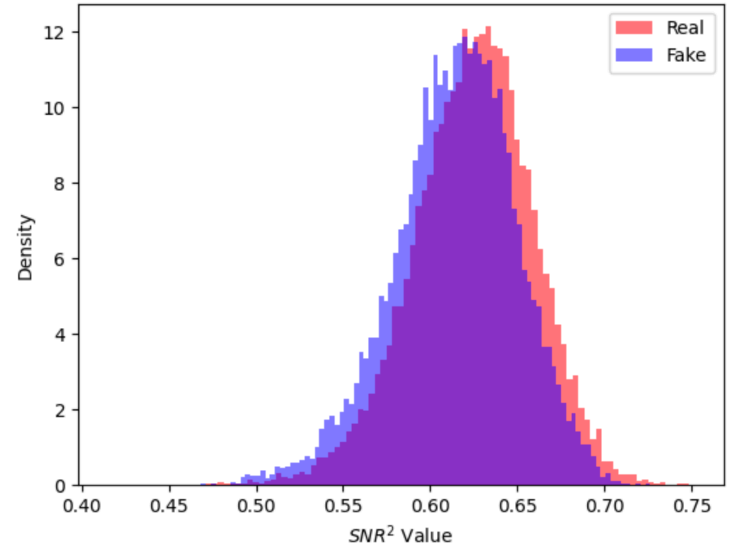




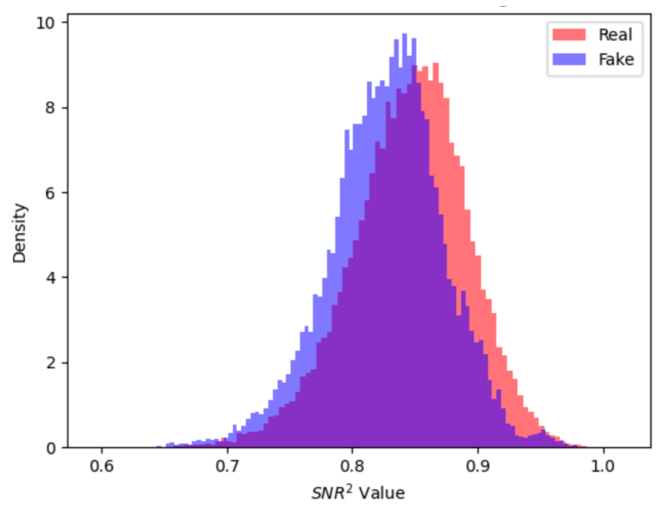


Signal to Noise Ratio





SND-1



SND-3

SND-30

Discussion

- The GAN succeeds in producing images which look visually similar to the training set.
- It also succeeds in reproducing the envelope and intensity histograms as well as the autocorrelation.
- However, the GAN consistently underestimates the SNR^2 or fails to properly reproduce the distribution which could lead to incorrect classification of the US.

Future Work

- Task based GAN evaluation (such as lesion detection) for our Clustered Lumpy Background (CLB), B-Mode Ultrasound Speckle and Angiogram datasets.

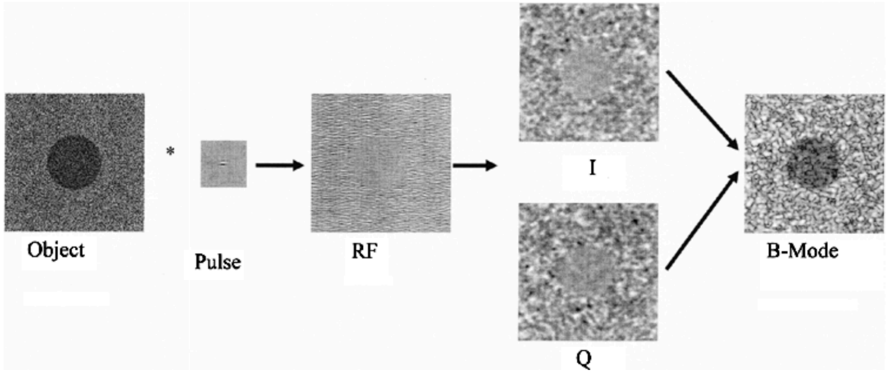


Fig. 1. Model of the object function magnitude of a circular hypo-echoic lesion, the pulse, RF and IQ data, and the B-mode image. [3]

- These studies will be then expanded to the AmbientGAN which trains using measurements instead of objects and has the forward operator as the last layer of the generator.

[3] R. J. Zemp, M. D. Parry, C. K. Abbey and M. F. Insana, "Detection performance theory for ultrasound imaging systems," in *IEEE Transactions on Medical Imaging*, vol. 24, no. 3, pp. 300-310, March 2005, doi: 10.1109/TMI.2004.841226.