CS230

Generative Adversarial Networks for Ultrasound Compressed Sensing

Soheil Hor, Ahmad Ghalayini, Khaled Saab
Department of Electrical Engineering
Stanford University
Email:{soheilh,ahmad2, ksaab}@stanford.edu

Abstract

Due to power, space, and bandwidth limitations in novel miniature ultrasound implants, there has been a growing interest in decreasing the computational requirements on these devices. One method towards achieving this goal is by reducing the amount of sampling that is done by such implants. In this project, we are proposing using generative adversarial networks (GANs) for the reconstruction of ultrasound B-mode images from 8% of raw time series transducer measurements. Using peak signal to noise ratio (PSNR) as our metric, we compare the GAN's performance to an FFT-based compressed sensing (CS) reconstruction method. Our GAN produces a reconstruction quality that is around 5 times better than the CS method, with a 5 times speed-up.

1 Introduction

1.1 Motivation

Due to Food and Drug Administration (FDA) safety regulations, ultrasound implants generally cannot have strong batteries, are usually powered wirelessly, and cannot heat up beyond a certain threshold. Therefore, they have strict power constraints. Similar regulations impose limits on their dimensions and thus the transducer array size. Power and dimension constraints then translate into bandwidth constraints, limiting the amount of data acquired in the body that the implants can send to an external receiver [1,2]. An example of this is the capsule ultrasound device (Fig. 1) currently in development at Stanford's research labs [3,4,5].

Therefore, to meet the aforementioned constraints, it would be desirable to reduce the amount of data sampling done by the implants. The challenge in undersampled ultrasound measurements is reconstructing the image with high diagnostic quality while dealing with aliasing artifacts. There exist some techniques in the compressed sensing (CS) literature to reconstruct undersampled data (e.g. using wavelet transforms) [6], but they are slow due to their iterative nature and would not perform well with measurements from a small transducer array size. In this project, we are trying to use GANs to learn the end-to-end mapping from highly compressed (a compression ratio of 12.5 in our tests) raw time-series ultrasound data to their corresponding high-quality B-mode images.

1.2 Our Approach

Summarizing our approach in one sentence, and motivated by the above, we are proposing an 1) end-to-end 2) low-latency compression and reconstruction solution for ultrasound imaging based on

CS230: Deep Learning, Winter 2018, Stanford University, CA. (LateX template borrowed from NIPS 2017.)

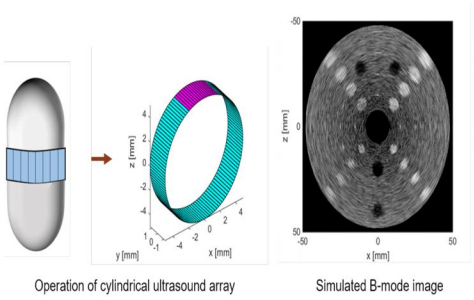

Figure 1: Pill-sized ultrasound device (credit: Arbabian Lab)

generative adversarial networks (GANs), with the aim of reconstructing high quality B-mode images from highly compressed raw time-series ultrasound measurements.

1.3 Why GANs?

Even though Generative Adversarial Networks (GANs) have been recently successful in image restoration in similar biomedical applications (e.g. MRI reconstruction [7]), they have not been investigated in the context of ultrasound image reconstruction.

2 Dataset and Features

After investigating the "Plane-wave Imaging Challenge in Medical Ultra Sound" (PICMUS) dataset [18], we found that this publicly available ultrasound dataset is not big enough for training complex networks like GANs. Moreover, our investigation about data simulation concluded that unlike in the usual ultrasound benchmarks, trying to reconstruct b-mode ultrasound images for simple but unrealistic topologies (circles of different sizes and boxes of different sizes) will not make training models for the reconstruction of natural b-mode images any easier. As a result, we decided to simulate raw-sensor data from a dataset of real ultrasound images publicly available at [15] using a Matlab toolbox used for time-domain simulation of acoustic wave fields known as "k-wave" [19].

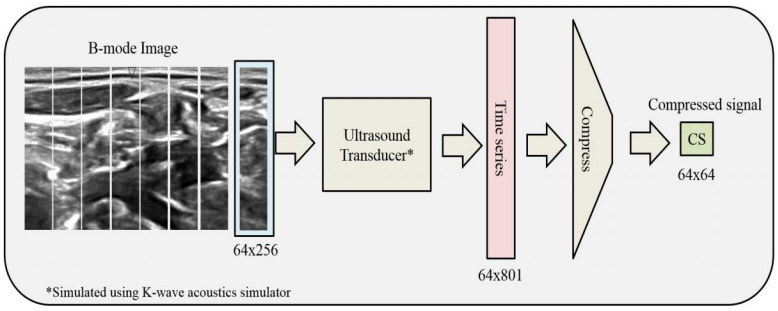

Figure 2: Data preprocessing overview

Moreover, we decided that a deep but narrow b-mode image is the most probable shape of a real-life b- mode image for our application because ultrasound implants usually have small transducer size (resulting in small width of b-mode image) but can provide a high resolution in time (which translates to higher number of pixels in depth direction of the image). Each image of the dataset provided in [15] has the size of 580x420. Considering our available computation resources we decided that b-mode image size of 64x256 would be proper for our experiment. Cropping each original b-mode image into smaller non-overlapping images resulted in 9 images per slice per patient. Considering that we have 47 patients and more than 100 b-mode slices per patient, the total available training set is now over 42000 images that should be enough for training a GAN. The boundary conditions of the k-wave simulator yielded that each b-mode images should have a padding 10 pixels resulting in a b-mode image of size 44*236. This image is translated to a raw sensor data of the size 44x801. 3 shows an example of the raw sensor data for a few of the sensors.

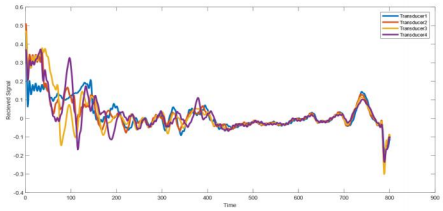

Figure 3: Time series raw data extracted from b-mode images

To model the sampling procedure in a real scenario, this raw signal data is multiplied by a randomly generated matrix of size 801xM (M representing the compressed dimension of the sampled signal) resulting in a matrix of size 44xM. The GAN will use this sampled data to reconstruct the original b-mode image.

3 Methods

As a starting point for our GAN, we used an open source baseline code from [17]. The input to the generative model is the under sampled time-series signal plus noise and the output is the reconstructed ultrasound image. The discriminator's job is to distinguish between the generated images and the original images. As a first step we have used a fully-connected network for both the generator and discriminator, similar to the approach in [13]. Unfortunately, in the current implementation of our GAN, the losses of the D and G networks and the accuracy of the network tend to oscillate and occasionally the accuracy of the D network gets fixed on 100. As a result, the sample output from the generator does not resemble the ultrasound images that the GAN has been trained on. Moreover, the generated images seem highly correlated (independent of the input).

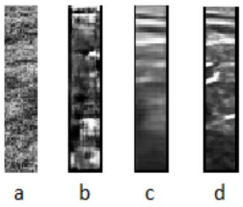

Figure 4: an example of performance of different models on a test image.
a. Mlp-GAN, b. Convolutional GAN c. Conditinal GAN with U-net, d. Target image

Considering that the dimensions of the inputs are very large for a dense network, we planed to use an adversarial auto-encoder using a convolutional network, along with a convolutional network for the discriminator. Unlike the previous model that resulted in noisy images (salt and pepper noise) in the image, images generated by this method where smooth to the point that the general outline of the constructed image resembled a low-quality b-mode image(see 4. However, the resulting GAN was very unstable. Meaning that the generator network would generate similar image for all of the test samples. At each epoch the output of the generator network would update but still would output the same image independent of the input. This behavior is known in literature as mode collapse which seemed to be the case independent of the number of training epochs.

Inspired by the pix2pix paper [11] we decided to use a conditional GAN for the end to end construction based on the implementation publicly available at [16]. The generator network of this model consists of a U-net network (5 and the discriminator is simply a multi layer fully connected network with the condition (compressed signal) and image (the generated or real b-mode image) concatenated together as the input to the network. After iteration over the number of layers, shape of layers and the learning rate, we came up with the current version of the conditional GAN which generates the image shown in 4. The general structure of the used model is shown in 6.

Although our current implementation is able to generate relatively realistic images, it still seems to have stability issues because the accuracy of the discriminator is very high (in range of 95 to 100

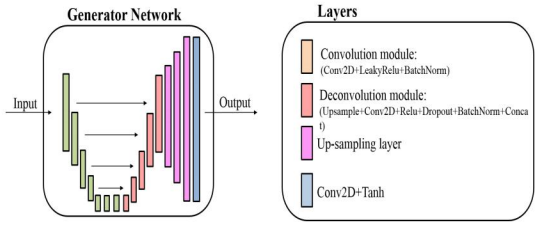

Figure 5: U-net network used as the Generator

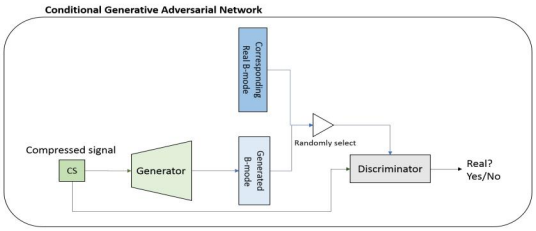

Figure 6: Structure of Conditional GAN

percent). We have tried some of the methods used in literature like training the discriminator and generator with different rates, replacing the model weights with models of previous iterations and using soft labels but none of these methods showed a considerable improvement in stability. The only alternative that seemed to be more stable than our current model was our implementation of wGAN [12] which yielded in a discriminator with accuracy of around 50 percent but for an unknown reason did not result in visually acceptable results.

4 Experiments/Results/Discussion

To understand how well the GAN is doing, we compare the reconstructed B-mode images to the target B-mode images. As a reference, we also compare to a traditional CS approach.

4.1 Traditional CS Approach

CS is a popular method that reconstructs a signal, $x \in \mathbb{R}^n$, from a linear combination of observed measurements,$y \in \mathbb{R}^m$, with $m < n$ [6]. The measurements are acquired from the transformation $y = \Phi x$, where Φ is an mxn matrix that consists in the delta functions basis, which models random under-sampling. CS then assumes that the signal to be reconstructed, x, can be expressed in an orthonormal basis, (e.g. Fourier or wavelets), which is chosen such that v is sparse, where $x = \Theta v$. Thus, $y = \Phi\Theta v = Av$, where A is an mxn full rank matrix. Solving for v results in the convex optimization shown below [7,8].

$$\hat{v} = arg \quad min\|v\|_{l1}, \quad subject \quad to \quad y = Ax \quad (1)$$

Once we solve for $\hat{v}$, x can be computed from $x = \Phi v$. In ultrasound reconstruction, three of the most common orthonormal bases used are wavelets, Fourier and wave-atoms. However, in this comparison we use the Fourier basis.

In our application, the signal, x, refers to the fully sampled RF data, while the observed signal, y, refers to the under-sampled RF data. The procedure using CS will consist of estimating the RF data, $\hat{x}$, then reconstructing the B-mode image.

4.2 Image Quality Comparison

In the CS literature, a common method to compare how close two images are is by computing the Peak Signal-to-Noise Ratio (PSNR) between the images. The equations to calculate the PSNR are shown below.

	Target	FFT
GAN	**15 dB**	7.3 dB
CS1	7.5 dB	22 dB
CS2	7.6 dB	X

Table 1: Average PSNR between images generated from CS with compression, CS without compression, GAN, and target over 374 test images

	Latency	Compression Percentage
GAN	**4 ms**	8%
CS1	20 ms	8%
CS2	22 ms	100%

Table 2: Reconstruction time and compression percentage for the three approaches

$$MSE = \frac{\Sigma_{M,N}[I_1(m,n) - I_2(m,n)]^2}{M \times N}$$

$$PSNR = 10\log(\frac{R^2}{MSE})$$

The results in Table 1 show the average PSNR over 374 test images. The GAN generates images closer to the target image than the traditional CS approach. As shown in Table 2, CS1 corresponds to CS with compression, while CS2 is CS without compression. We also took the average reconstruction times over the test images and found that GAN generates images around 5 times faster than CS.

Fig. 7 displays an example of the reconstructed images of the three approaches along with the target B-mode image on the far left.

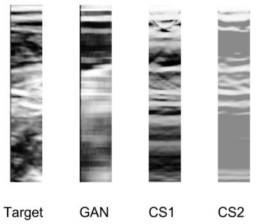

Figure 7: Example of reconstructed B-mode images from GAN, CS1, and CS2

5 Conclusion/Future Work

Due to the constraints of state-of-the-art ultrasound implants, reconstruction of B-mode images from under-sampled RF data has gained interest. The main limitations of traditional CS methods is the low lateral resolution when dealing with a small transducer array size. In this project, we experimented using a conditional GAN to generate the B-mode images, which was conditioned by the under-sampled RF data.

We have shown the potential of the superiority of using a GAN over these traditional methods when faced with such hardware constraints. Using a GAN revealed two critical benefits over the traditional CS method: (1) it was able to generate images closer to the target images in the PSNR sense, and (2) it reconstructs images five times faster.

An important next step would be to have a physician verify that the generated images from the GAN are indeed legitimate reconstructions. We also want to train our GAN on a larger and more diverse dataset. Additionally, comparison with other reconstruction methods would be beneficial.

6 Contributions

All authors had equal contributions, with continuous cooperation and planning between them, sharing and dividing tasks as needed. Specifically, for presentation and preparation of posters and reports, all authors were equally involved. For data collection, choosing a good GAN architecture with suitable inputs and outputs, designing the model and choosing its parameters, Soheil Hor was the main contributor; Ahmad Ghalayini and Khaled Saab were the main contributors towards experimenting to train the model on Amazon Web Services (AWS) as well as finding CS methods to compare our model's performance to.

References

[1] J. Charthad, M. J. Weber, T. C. Chang, and A. Arbabian, "A mm-sized implantable medical device (imd) with ultrasonic power transfer and a hybrid bi-directional data link," IEEE Journal of Solid-State Circuits, vol. 50, no. 8, pp. 1741–1753, Aug 2015.

[2] J. Charthad, M. J. Weber, T. C. Chang, M. Saadat, and A. Arbabian, "A mm-sized implantable device with ultrasonic energy transfer and rf data uplink for high-power applications," in Proceedings of the IEEE 2014 Custom Integrated Circuits Conference, Sept 2014, pp. 1–4.

[3] F. Memon, G. Touma, J. Wang, S. Baltsavias, A. Moini, C. Chang, M. F. Rasmussen, A. Nikoozadeh, J. W. Choe, A. Arbabian, R. B. Jeffrey, E. Olcott, and B. T. Khuri-Yakub, "Capsule ultrasound device," in 2015 IEEE International Ultrasonics Symposium (IUS), Oct 2015, pp. 1–4.

[4] F. Memon, G. Touma, J. Wang, S. Baltsavias, A. Moini, C. Chang, M. F. Rasmussen, A. Nikoozadeh, J. W. Choe, E. Olcott, R. B. Jeffrey, A. Arbabian, and B. T. Khuri-Yakub, "Capsule ultrasound device: Further developments," in 2016 IEEE International Ultrasonics Symposium (IUS), Sept 2016, pp. 1–4.

[5] J. Wang, F. Memon, G. Touma, S. Baltsavias, J. H. Jang, C. Chang, M. F. Rasmussen, E. Olcott, R. B. Jeffrey, A. Arbabian, and B. T. Khuri- Yakub, "Capsule ultrasound device: Characterization and testing results," in 2017 IEEE International Ultrasonics Symposium (IUS), Sept 2017, pp. 1–4.

[6] D. L. Donoho, "Compressed sensing," IEEE Transactions on Information Theory, vol. 52, no. 4, pp. 1289–1306, April 2006.

[7] M. Mardani, E. Gong, J. Y. Cheng, S. Vasanawala, G. Zaharchuk, M. T. Alley, N. Thakur, S. Han, W. J. Dally, J. M. Pauly, and L. Xing, "Deep generative adversarial networks for compressed sensing automates MRI," CoRR, vol. abs/1706.00051, 2017. [Online]. Available: http://arxiv.org/abs/1706.00051

[8] M. W. E.J. Candes, "An introduction to compressive sampling," IEEE Signal Processing Magazine, vol. 25, pp. 21–30, 2008.

[9] E. Candes, "The restricted isometry property and its implications for compressed sensing," Compte Rendus de lAcademie des Sciences, no. 346, pp. 589–592, 2008.

[10] T. T. E.J. Candes, "Decoding by linear programming," IEEE Transactions on Information Theory, no. 51, pp. 4203–4215, 2005.

[11] Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." arXiv preprint (2017).

[12] Arjovsky, Martin, Soumith Chintala, and Léon Bottou. "Wasserstein gan." arXiv preprint arXiv:1701.07875 (2017).

[13] Radford, Alec, Luke Metz, and Soumith Chintala. "Unsupervised representation learning with deep convolutional generative adversarial networks." arXiv preprint arXiv:1511.06434 (2015).

[15] URL: https://www.kaggle.com/c/ultrasound-nerve-segmentation/rules

[16] URL: https://github.com/eriklindernoren/Keras-GAN/tree/master/pix2pix

[17] URL: https://github.com/gwaygenomics/keras-gan