
This is the Accepted version of the article

Ultrasound speckle reduction using generative adversarial networks

Fabian Dietrichson, Erik Smistad, Andreas Ostvik, Lasse Løvstakken

Citation:

Fabian Dietrichson, Erik Smistad, Andreas Ostvik, Lasse Løvstakken(2018) Ultrasound speckle reduction using generative adversarial networks. In: 2018 IEEE International Ultrasonics Symposium (IUS) Kobe, Japan, 22-25 Oct. 2018

DOI: <https://doi.org/10.1109/ULTSYM.2018.8579764>

This is the Accepted version.
It may contain differences from the journal's pdf version

This file was downloaded from SINTEFs Open Archive, the institutional repository at SINTEF
<http://brage.bibsys.no/sintef>

Ultrasound speckle reduction using generative adversarial networks

Fabian Dietrichson

Norwegian University of Science and Technology (NTNU)
Trondheim, Norway
fabiansd1402@gmail.com

Andreas Østvik

Centre for Innovative Ultrasound Solutions (CIUS)
Norwegian University of Science and Technology (NTNU)
SINTEF Medical Technology
Trondheim, Norway
andreas.ostvik@ntnu.no

Erik Smistad

Centre for Innovative Ultrasound Solutions (CIUS)
Norwegian University of Science and Technology (NTNU)
SINTEF Medical Technology
Trondheim, Norway
erik.smistad@ntnu.no

Lasse Lovstakken

Centre for Innovative Ultrasound Solutions (CIUS)
Norwegian University of Science and Technology (NTNU)
Trondheim, Norway
lasse.lovstakken@ntnu.no

Abstract—Generative adversarial networks (GANs) have shown its ability to create realistic and accurate image-to-image transformation. The goal of this work was to investigate whether deep convolutional GANs can learn to perform advanced ultrasound speckle reduction in real-time. The GAN was trained using a dataset of cardiac images from 200 patients and tested on a separate dataset from 55 patients. A U-net type of generator was used together with a patch-wise discriminator. Three different generator sizes were tested in order to see the tradeoff between speckle reduction accuracy and runtime. The results show that GANs can learn ultrasound speckle reduction. Even though the training set consisted only of cardiac ultrasound images, results from other parts of the body and scanners indicate that the method learns speckle reduction in general, and not just for cardiac images. By reducing the number of filters in the generator, real-time performance was achieved with an average of 11 ms per frame.

I. INTRODUCTION

Raw ultrasound B-mode images go through a significant amount of post-processing before being displayed. State-of-the-art image processing algorithms may be complex, difficult to implement and not necessarily efficient enough to satisfy real-time requirements of ultrasound devices. One example is speckle reduction for improving image contrast. Speckle is a type of granular image degrading noise pattern that arise due to interference of reflected ultrasound waves.

A generative adversarial network (GAN) is an unsupervised deep learning approach introduced in 2014 by Goodfellow et al. [1]. GANs have shown to be able to create realistic and

accurate image-to-image transformations, including denoising low-dose CT images as shown by Wolterink et al. [2] in 2017.

GANs consist of two networks: a generative network (G), and a discriminative network (D). The goal is to train a generator which can create new samples which lay in the same distribution as the training data. The discriminator are given samples from the training set and samples from the generator and have to determine if the samples are real or fake. Thus, the training procedure for G is to maximize the probability of D making a mistake, while D is trained to minimize the misjudging error on whether the images it receives are real or fake.

The aim of this work was to investigate if GANs can learn an advanced ultrasound image processing algorithm, and to see if this approach can provide a practical path to real-time performance by utilizing the optimized neural network GPU backends. Simultaneously with this work, Mishra et al. [3] published a GAN based ultrasound speckle reduction method, training with liver data from a high-end scanner as target images, and simulation corrupted images from the high-end scanner and images from a low-end scanner as input images. Instead of using simulated speckle and low-end scanner images to learn speckle reduction as done by Mishra et al. [3], the proposed method learns from a known complex filtering method, and is thus trained and validated on real speckle noise and filtered output of the images. We also show how this method can be optimized for real-time runtime and that it works on ultrasound images from different parts of the body.

II. METHODS

A. Dataset

A dataset of 600 samples from 200 patients was used to train the networks. This dataset consisted of cardiac B-mode

ultrasound images from the apical four chamber view acquired with a Vivid E9 scanner and a M5S phased-array transducer (GE Vingmed Ultrasound, Horten, Norway). Images from 10% of the patients were used for parameter optimization during training. A second dataset of 170 samples from 55 patients was created to calculate the accuracy of the network. The scan converted images were resized to 512×512 . All images were processed using the optimized non-local low-rank (NLLR) speckle reduction algorithm by Zhu et al. [4]. These speckle reduced images were the target for the training and used for evaluation on the test set.

B. Generative adversarial network setup

In this work, the goal is to have the generator learn how to transform a raw B-mode ultrasound image into a speckled reduced image similar to that of the NLLR algorithm. The discriminator, on the other hand, is trained to separate the speckle reduced NLLR images from the output of the generator as shown in Fig. 1.

The generator model was created using a fully convolutional U-net type architecture [5]. Instead of max pooling, strided convolutions were used to downsample the feature maps in the encoder part. For the decoder part, simple upsampling was used and concatenated with cross-over connections from the encoder. Batch normalization was used for all convolutional layers except the last. Leaky ReLU was used as the activation function for all layers in the encoder, ReLU for layers in the decoder and tanh for the last layer. The non-saturating cross-entropy loss function was used on the generator which allows better learning for both the generator and discriminator because it ensures that both networks have access to a strong error gradient when one of the networks is losing. Since this work not only focus on good accuracy, but also real-time performance, three different sizes of the generator were tested: large, medium and small. All three networks had the same depth of 7, but with varying amount of filters for each level in the encoder and decoder stage:

- Large: 64, 128, 256, 512, 512, 512, 512, 512, 512, 1024, 1024, 1024, 512, 256, 1
- Medium: 32, 32, 64, 64, 128, 128, 256, 256, 256, 256, 256, 128, 128, 64, 64, 1
- Small: 2, 2, 4, 4, 8, 8, 16, 16, 16, 16, 16, 8, 8, 4, 4, 1

A convolutional patch classifier, also known as a *PatchGAN*, was used as the discriminator. Compared to a normal convolutional image classifier which considers the whole image, the patch classifier divides the image in $N \times N$ patches and classifies each patch as fake or real, then averages the responses of all patches to provide the final output of the discriminator [6]. The idea is that the discriminator doesn't need global context of the entire image, and thus instead focus the discriminator on small local structures in the image to determine its trueness. This results in a much smaller and faster discriminator network. In this work, $N = 128$ was used, totalling in 16 patches.

The standard GAN training approach from [1] was used by alternating the backpropagation steps on the discriminator and

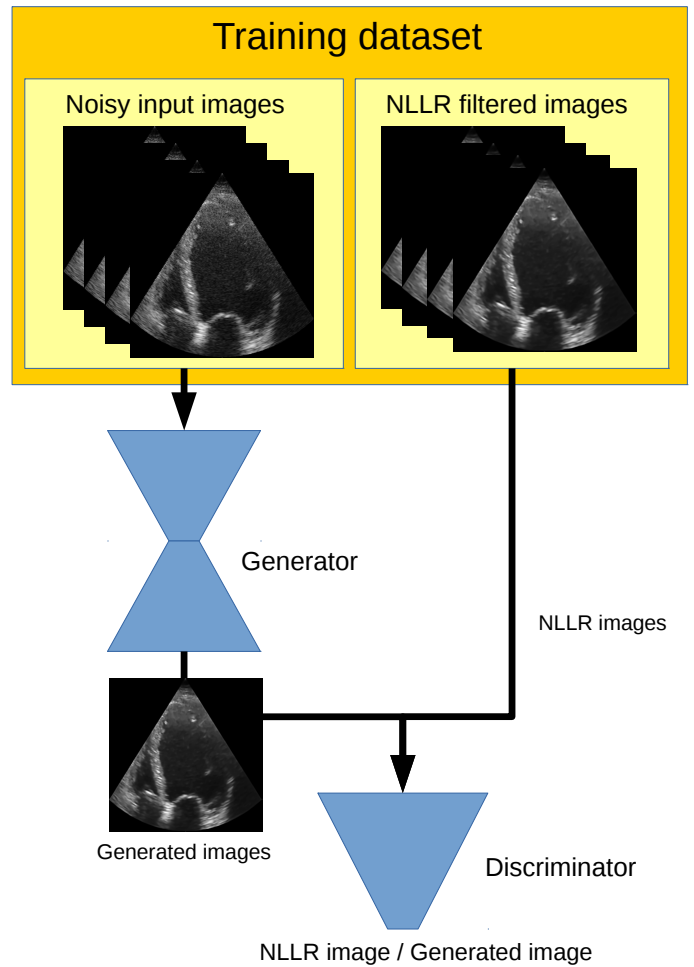


Fig. 1. GAN speckle reduction training setup. The training dataset consists of a set of raw B-mode ultrasound images and their NLLR speckle reduced counterparts. The generator is trained to modify the raw B-mode input image to resemble the NLLR processed images. The discriminator is trained to differentiate the NLLR images and the ones coming from the generator, this drives the generator to learn how to reproduce the NLLR algorithm.

the generator on a batch of 50/50 NLLR images and generated images. These two steps were repeated until the networks converge toward a local Nash equilibrium, meaning that the discriminator is as good as guessing at random and thus not able to distinguish between the NLLR and the despeckled images from the generator.

C. Deployment

Training was done in Python with the Keras and TensorFlow frameworks [7]. After training, the neural network and learned weights were exported as a TensorFlow graph and stored in a single protobuf file. Only the generator network is exported, as the discriminator is only needed for the training process. This file can then be loaded directly in C++ using the TensorFlow C++ API. Thus, the process of going from a research prototype to testing it on a live stream of ultrasound images is short. Assuming a modern GPUs is available, the runtime is low as TensorFlow is highly optimized with the

NVIDIA cuDNN backend. The FAST framework [8] provides a wrapper for TensorFlow, GPU optimized processing and visualization and real-time streaming of ultrasound data, thus making deployment and prototype testing even easier. Using FAST, the exported generator was tested in real-time by streaming ultrasound data from a GE Vivid E95 scanner. A video of this experiment is accessible online¹.

III. RESULTS

For comparison, the results of using a simple Gaussian blur filtering method was also included. Fig. 2 shows a comparison of the Gaussian blur, GAN medium and NLLR methods on two images from the test set.

Table I shows the mean runtime, absolute error (MAE) and standard deviation between the different GAN network sizes and the target image. The accuracy was calculated for each pixel and example in the test set using image intensities in the scale of 0 to 255. The runtime of the neural networks was measured using an NVIDIA Geforce Titan X GPU, with TensorFlow and the cuDNN backend from NVIDIA.

In addition to the large cardiac test set, the GAN method was tested on two images of the liver as shown in Fig. 3. This was done to see if the trained network is applicable to other organs.

TABLE I
RUNTIME AND MEAN ABSOLUTE ERROR (MAE) FOR THE TEST DATASET.

Filtering method	Runtime (seconds)	MAE \pm std. dev.
Input image	-	4.0 \pm 0.4
Gaussian filter	0.0063	4.9 \pm 0.7
GAN Large	0.286	2.2 \pm 0.3
GAN Medium	0.011	2.6 \pm 0.3
GAN Small	0.004	3.7 \pm 0.5
NLLR (target)	160	-

IV. DISCUSSION

The results in Fig. 2 and Table I show that DNNs are able to learn principal features of NLLR, which is to smooth out the speckle while preserving tissue boundaries. While a simple Gaussian blurring method can also smooth out the speckle, it is not able to preserve edges.

Comparison of the three different generator sizes in Table I show the tradeoff between accuracy and runtime. The small network does not give much improvement from the input image, while the medium network achieve much better accuracy with a real-time runtime. The large generator on the other hand, has a high runtime of 286 ms, and only slightly better accuracy than the medium GAN. In comparison, the method of Mishra et al. [3] used 300 ms on the same GPU. More efficient runtime was achieved by reducing the number of convolutions, other optimizations such as quantization may increase performance even further. This runtime is a huge improvement from the 160 seconds used by the original NLLR implementation

by Zhu et al. [4]. However, this implementation was written in Matlab with parts written in C using MEX files².

The test on liver images in Fig. 3 show that although the method was trained on cardiac images, it is able to reduce speckle on images of the liver as well. However, the output is not as identical as in the cardiac examples in Fig. 2.

The proposed method can only perform one level of speckle reduction, however NLLR supports adjusting the aggressiveness of the speckle reduction through a parameter. Future work includes using a conditional GAN to train a neural network which can adjust the level of speckle reduction through an extra input parameter.

V. CONCLUSION

The results indicate that generative adversarial networks (GANs) can learn advanced ultrasound image processing algorithms by only having input and target images, and perform these image transformations efficiently in real-time using modern GPUs.

REFERENCES

- [1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversial Nets," in *Neural Information Processing Systems (NIPS)*, mar 2014. [Online]. Available: <http://arxiv.org/abs/1406.2661>
- [2] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Išgum, "Generative Adversarial Networks for Noise Reduction in Low-Dose CT," *IEEE transactions on medical imaging*, vol. 36, no. 12, pp. 2536–2545, 2017.
- [3] D. Mishra, S. Chaudhury, M. Sarkar, and A. S. Soin, "Ultrasound image enhancement using structure oriented adversarial network," *IEEE Signal Processing Letters*, vol. 25, no. 9, pp. 1349–1353, 2018.
- [4] L. Zhu, C.-W. Fu, M. S. Brown, and P.-A. Heng, "A Non-local Low-Rank Framework for Ultrasound Speckle Reduction," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, jul 2017, pp. 493–501. [Online]. Available: <http://ieeexplore.ieee.org/document/8099543/>
- [5] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *MICCAI*, pp. 234–241, 2015.
- [6] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, jul 2017, pp. 5967–5976. [Online]. Available: <http://ieeexplore.ieee.org/document/8100115/>
- [7] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," mar 2016. [Online]. Available: <http://arxiv.org/abs/1603.04467>
- [8] E. Smistad, M. Bozorgi, and F. Lindseth, "FAST: framework for heterogeneous medical image computing and visualization," *International Journal of Computer Assisted Radiology and Surgery*, vol. 10, no. 11, pp. 1811–1822, 2015. [Online]. Available: <http://link.springer.com/10.1007/s11548-015-1158-5>

¹https://www.youtube.com/watch?v=phY2_-Y-r6A

²https://sites.google.com/site/indexlzh/webpage_despeckling_cvpr2017/index

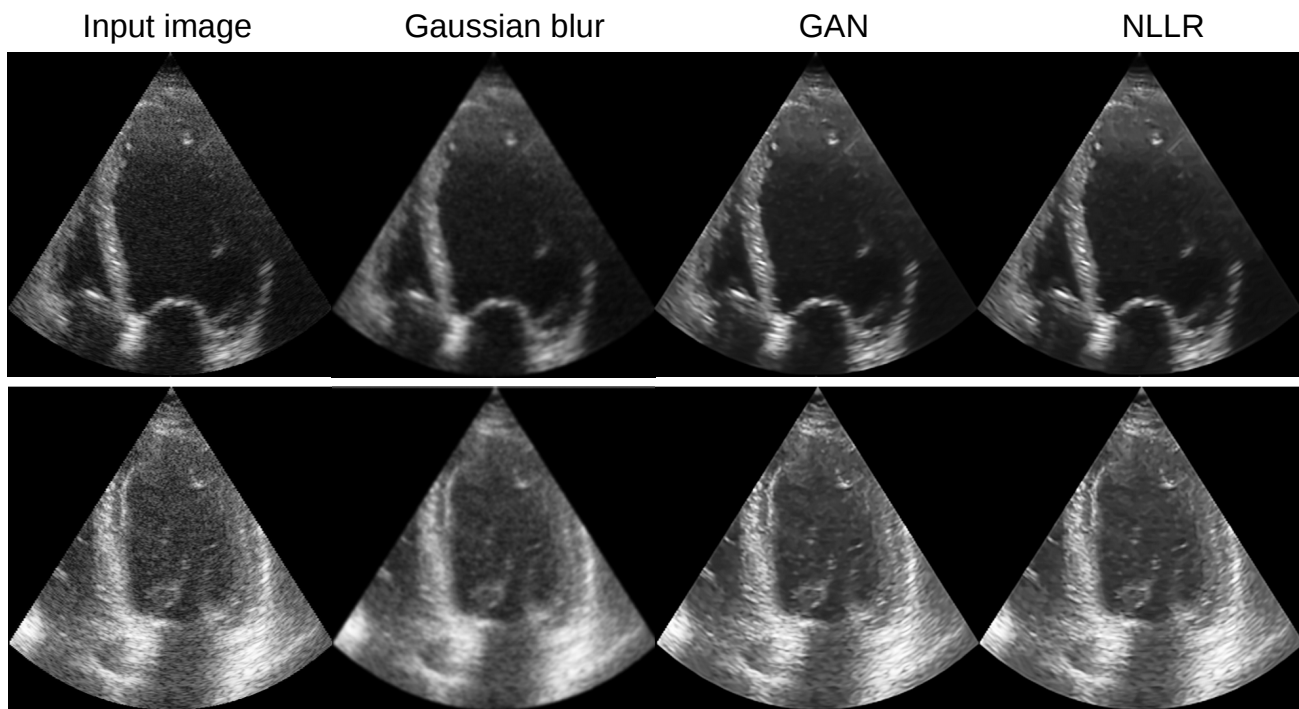


Fig. 2. Cardiac ultrasound despeckling examples.

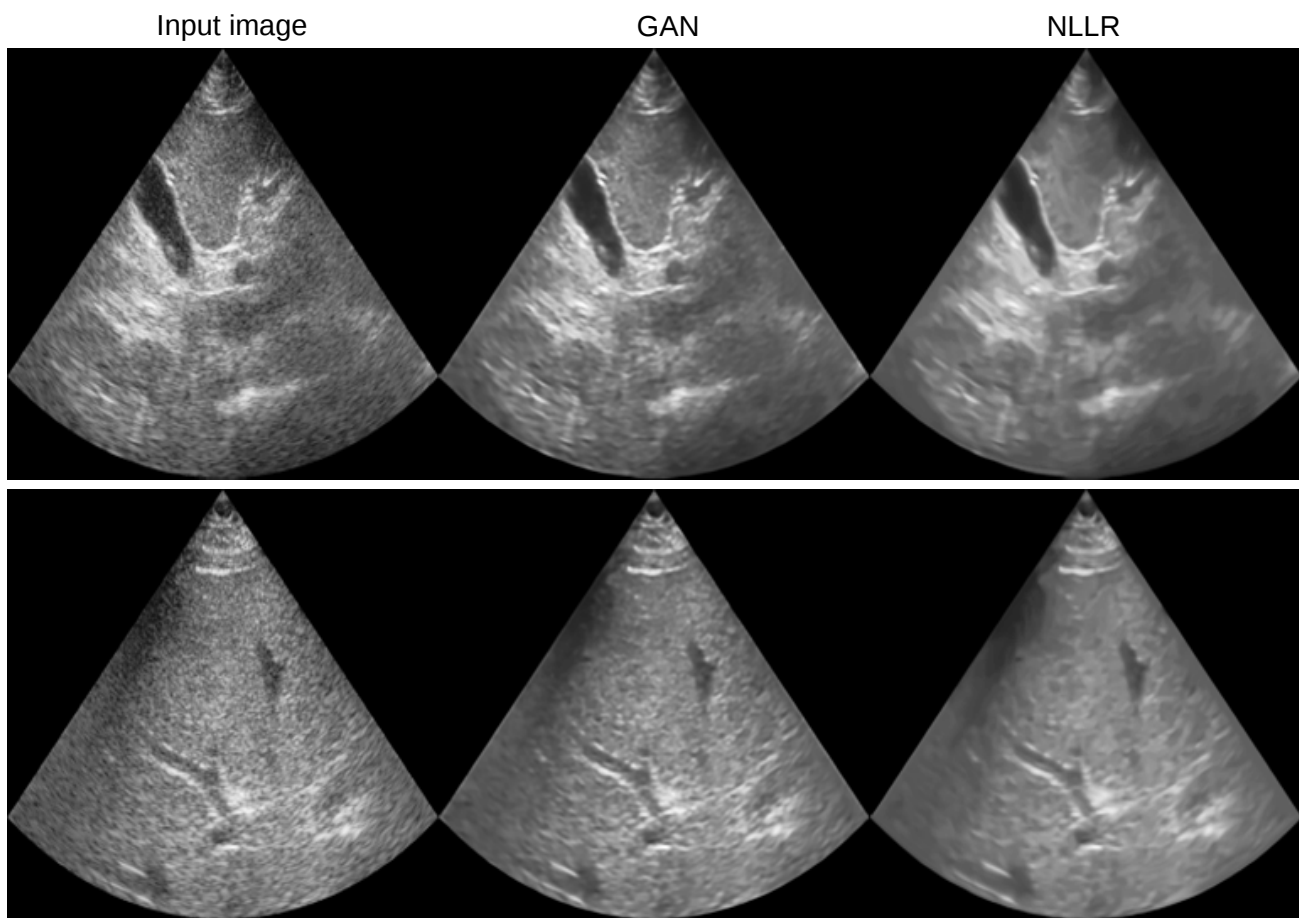


Fig. 3. Liver ultrasound despeckling examples.