Proceedings of the 2020 Design of Medical Devices Conference
DMD2020
April 6, 7-9, 2020, Minneapolis, MN, USA

DMD2020-9109

USDL: Inexpensive Medical Imaging Using Deep Learning Techniques and Ultrasound Technology

Manish Balamurugan[1], Kathryn Chung[1], Venkat Kuppoor[1], Smruti Mahapatra[2]

[1]Fairfax High School, Fairfax, VA, United States
[2]Dept of Biomedical Engineering-Johns Hopkins University, Baltimore, MD, United States

Aliaksei Pustavoitau, MD
Dept. of Anesthesiology-Johns Hopkins University
Baltimore, MD, United States

Amir Manbachi, PhD
Dept. of Neurosurgery-Johns Hopkins University
Baltimore, MD, United States

Abstract
In this study, we present USDL, a novel model that employs deep learning algorithms in order to reconstruct and enhance corrupted ultrasound images. We utilize an unsupervised neural network called an autoencoder which works by compressing its input into a latent-space representation and then reconstructing the output from this representation. We trained our model on a dataset that compromises of 15,700 *in vivo* images of the neck, wrist, elbow, and knee vasculature and compared the quality of the images generated using the structural similarity index (SSIM) and peak to noise ratio (PSNR). In closely simulated conditions, the architecture exhibited an average reconstruction accuracy of 90% as indicated by our SSIM. Our study demonstrates that USDL outperforms state of the art image enhancement and reconstruction techniques in both image quality and computational complexity, while maintaining the architecture efficiency.

Keywords: Ultrasound imaging, Autoencoders, Deep Learning, Speckle Noise, Denoising, *In Vivo* Ultrasounds, PSNR, SSIM, MSE

INTRODUCTION
Ultrasound imaging is the most commonly performed cross-sectional diagnostic imaging method in the medical field today, ranging from musculoskeletal to abdominal scans. It is broadly used to examine the thoracic pelvic and abdominal regions to guide procedures. Ultrasound imaging utilizes the interaction of sound waves with living tissue in order to produce images of the tissue in real time. [1] These images can be processed in order to obtain quantitative structural and functional information from the region of interest. [1] Among imaging modalities, ultrasound images are the most widely used for initial diagnostic purposes due to its non- invasive, and non-ionizing nature, along with it being adaptable, portable, and economical for medical imaging. However, despite the advantages of ultrasound technology, it has its disadvantages. When consulting clinicians, the key issue emphasized is that while ultrasound technology is convenient and cost-effective, inherent flaws such as image quality and lack of consistency has rendered ultrasound imaging an initial diagnostic tool at best.

Ultrasound images are often corrupted due to various noises such as additive noise, system noise, along with speckle (multiplicative) noise. The impact of speckle noise is the most significant in terms of the visual quality of medical ultrasound images. Speckle noise is a noise that is generated mainly caused by the interference of the returning wave at the transducer aperture. The wavelength of the highly sensitive signals are scattered, resulting in random black and white dots of speckle noise.

Speckle is a multiplicative, granular noise that inherently exists in and degrades the quality of medical ultrasound. Speckle is mainly caused by the interference of the returning wave at the transducer aperture. Speckle noise blur features which are essential for diagnosis and assessment. Often, these variables lead to misdiagnosis where an ultrasound image might identify a potential concern that is not malignant. This leads to further testing that might not be required. These issues make speckle reduction an important step in the processing and analysis of medical ultrasound images.

In order to correct the issues with ultrasound image formation, we propose the utilization of a denoising autoencoder algorithm (DAE) in order to reconstruct and enhance corrupted ultrasound images. A DAE consists of three layers: an input layer, a hidden (encoding) layer, and a decoding layer. A DAE is an unsupervised learning algorithm that applies backpropagation, which is the practice of fine-tuning the weights of a neural network input based on the error rate obtained from the previous epoch (iteration), thereby improving the reliability of the model.

METHODS

This study presents: Ultrasound Deep Learning (USDL), a novel tool that employs deep learning algorithms to reconstruct corrupted ultrasound scans. In closely simulated conditions, the architecture exhibited an average reconstruction accuracy of 90% as measured by structural similarity index (SSIM). We show that USDL is able to outperform state of the art filtration techniques for the recovery of ultrasound images in both quality and reconstruction time.

In order to train USDL, we utilized developed a dataset consisting of 15,700 *in vivo* ultrasound images that we obtained through the Johns Hopkins Hospital. This dataset is comprised of images from various regions in the body as a form of cross-validation that the efficiency of our model is applicable to a wide range of medical regions and purposes.

The images were specifically targeted towards the neck, knee, wrist, and elbow region. These regions have a substantial amount of vasculature, allowing the blood velocity to be easily detected through the B-mode ultrasound. In the neck region, the carotid artery was the region of interest (ROI). The popliteal vein in the knee , the ulnar and radial arteries/veins in the wrist, and the median cubital vein in the antecubital fossa (elbow) were additional ROI's respectively.

In closely simulated conditions, the architecture exhibited an average reconstruction accuracy of 90% as measured by structural similarity index (SSIM). This platform can facilitate inexpensive and timely diagnosis by improving upon the efficiency of ultrasound technology as an initial diagnostic tool.

Framework

The system architecture was built around a denoising autoencoder (DAE). As stated earlier, a standard autoencoder has the risk of learning the "Identity Function" where the output equals the input, thereby marking the autoencoder output essentially useless. We use a DAE in order to resolve this issue by randomly corrupting data that the autoencoder must reconstruct. Thus, we trained the DAE to reconstruct a "clean" version of the input from a corrupted version of said input - and this process is known as denoising. The training process of the DAE involves the following processes [13]:

1. The initial input x is corrupted into $\tilde{x}$ by means of stochastic mapping $\tilde{x} \sim q_D(\tilde{x} \mid x)$
2. The corrupted image is then mapped to the hidden representation $y = f_\theta(\tilde{x}) = s(W\tilde{x} + b)$
3. From the hidden representation, the DAE then reconstructs $z = g_{\theta'}(y) = s(W'y + b')$

We utilize backpropagation in order to minimize the average reconstruction error over the training data, specifically the loss between the output values with the original input. It is important to make the distinction that input corruption only occurs during the training phase of the DAE, and once the model has been optimally trained to recognize specific parameters, no corruption is required to the original data in order to extract key representations.

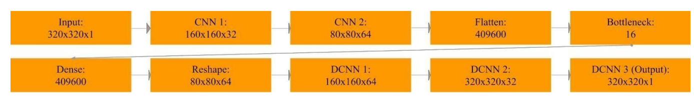

Fig. 1: System Architecture

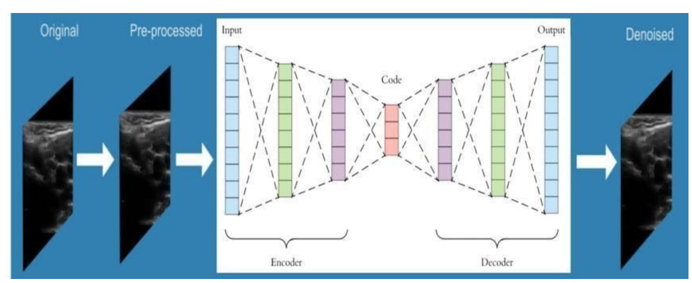

Fig. 2: Autoencoder Architecture

Preprocessing

In order to raise the efficiency of the DAE, we utilized various techniques in order to optimize our dataset. The data was converted to grayscale using the OpenCV Image Processing API, because RGB data is 3 times the size of grayscale data, and the color is not essential for this specific task. Images were also resized using OpenCV in order to normalize the dataset as each image in the data set varied in dimensions. Images were downscaled to 320 by 320 pixels for testing purposes to train the model efficiently. By downscaling these images, the data set is normalized, maximizing efficiency.

Overview

After the images were preprocessed in OpenCV, we utilized convolution. The spatial dimensions of the results were then condensed via pooling. The outputs from each convolutional hidden layers were fed into a rectified linear unit (ReLU) activation function. This process occurs twice then enters a bottleneck of only a few neurons to discard unnecessary information. The entire process is fully reversed to form the output, except that the final layer is determined through a sigmoid activation function. The autoencoder will generate a reconstructed image using the determined predictions.

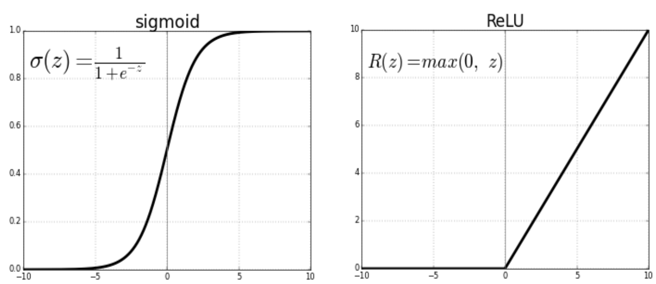

Fig.3 Activation function

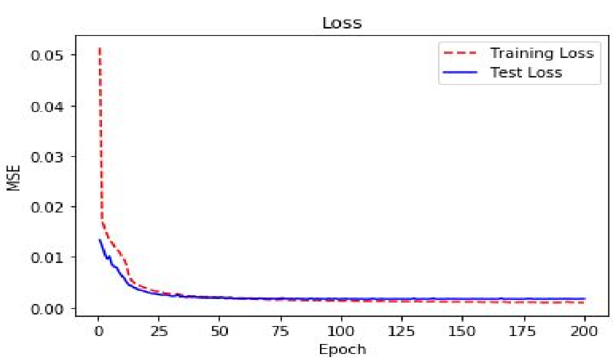

Fig.4: Loss Graph of Model

A dataset consisting of 15700 images from the right external carotid artery of a single healthy human using a Clarius Ultrasound Probe at 4.8 cm penetration level was used to train and test the neural network. 15500 images of the dataset were used to train the algorithm, and a corruption algorithm was used on top of the remaining 200 images to test the algorithm. The images used were highly varied in angle and position to promote accuracy in the algorithm's ability to reconstruct.

The *in vivo* images were acquired utilizing a linear array transducer- the optimal method for capturing the intricacies of the carotid artery. Linear array transducers contain a group of piezoelectric crystals, which convert electrical signals into mechanical energy or mechanical energy into electrical signals, arranged in a linear fashion. The transducer produced real time 2D images with a frequency ranging between 2.5-12 Mhz.

RESULTS

In order to assess the performance of our algorithm, a structural similarity index (SSIM) and peak noise-to-signal ratio (PSNR) were used to measure the similarity between the images that were extracted from our data set and their corresponding reconstructions.

SSIM is a mathematical index that measures the similarity between images by using an initial uncompressed image as a reference. It bases its foundation on the concept of comparing change in terms of structure. It is different from PSNR and MSE in that it takes into account luminance and contrast masking. By attaching a value to the amount of compression or loss there is in data transmission, the SSIM calculates the structural differences. The higher the SSIM value, the more structurally similar the initial and reconstructed image is. PSNR is an alternative technique that represents the ratio between the maximum power between a signal and the power of corrupting noise in grayscale images. They are used to compare two images, a referenced and processed one; they can be a good indicator of image degradation or clarification. The higher the PSNR value, the better the quality of the reconstructed image. In order to compute the extent of error, a numerical value is calculated using MSE (mean-square error). MSE assesses the performance of algorithms consisting of regressions. As the algorithm was trained using an increased amount of data to introduce more variability, the MSE value progressively decreased. The fluctuation in the MSE value indicates that the algorithm improved and reduced the amount of randomized error, which can be represented by the loss graph (Fig. 4).

The algorithm was tested at 200 epochs (reiterations), generating 200 SSIMs. A confidence interval was constructed to determine how the data can be related to all possible tests. We are 99 percent confident that the population mean of all possible test cases falls between 0.798 and 0.804 for SSIM and 39.758 and 40.368 for PSNR.

To compare the efficiency of our algorithm in comparison to current noise reduction methods, the configurations are shown in Fig. 5 and Fig.6. It clearly indicates that our SSIM was high-reaching 0.904, while the other filters were in the 0.20-0.40 range. Furthermore, the PSNR comparison chart represents that our algorithm had a PSNR of 40.0628, while the others merely had 25.000-26.000.

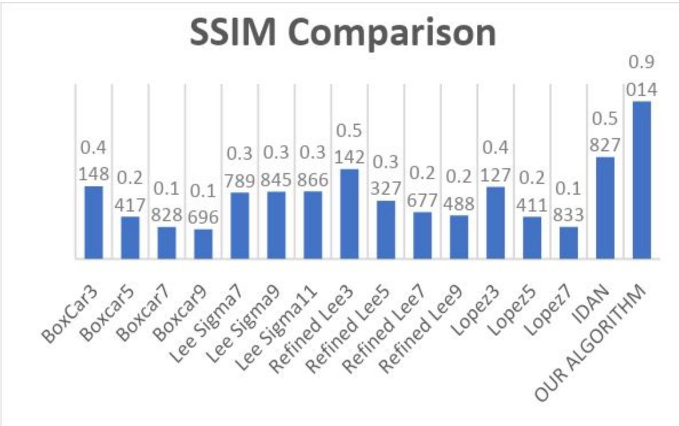

Figure 5: SSIM Comparison with current noise reduction models

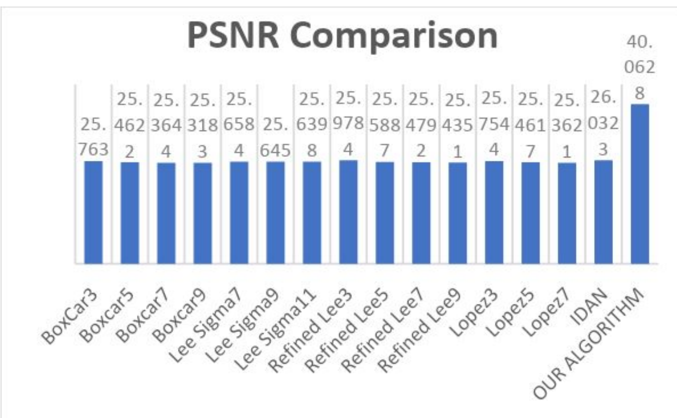

Figure 6: PSNR Comparison with current noise reduction models

DISCUSSION

Medical ultrasound images are corrupted with a specific type of noise called *speckle noise*, which causes inaccuracies on image-based interpretation and diagnostic procedure. This is why Speckle Reduction is an important step prior to the processing and analysis of medical ultrasound.

As previously mentioned in the introduction, speckle noise blurs features which are essential for medical diagnosis and assessment. The wavelength of the highly sensitive constructive and destructive signals are scattered, resulting in randomized black and white dots of speckle noise. Noise reduction is an important tool that can be used to increase clarity and remove all undesired confounding discrepancies.

Currently, there are noise reduction methods using filters, such as the Lee filter and Frost filter. These filters suppress noise and statistical values such as MSE and PSNR are used to compare the effectiveness. However, Lee filters are unable to remove noise at edges and while filtering methods utilizing scalar filters are good for removing high frequency noise, they cannot preserve the details of the edges, whereas Adaptive filter methods are able to preserve the details but require a significant amount of more computation time.

We were able to train our model to reconstruct corrupted ultrasound scans (Fig.7). In closely simulated conditions our deep learning architecture exhibited results that indicated an average reconstruction accuracy of 90 percent measured by the Structural Similarity Index (SSIM) which is vastly superior to existing state of the art methods of speckle denoising via filtration techniques such as Refined Lee3 or IDAN.

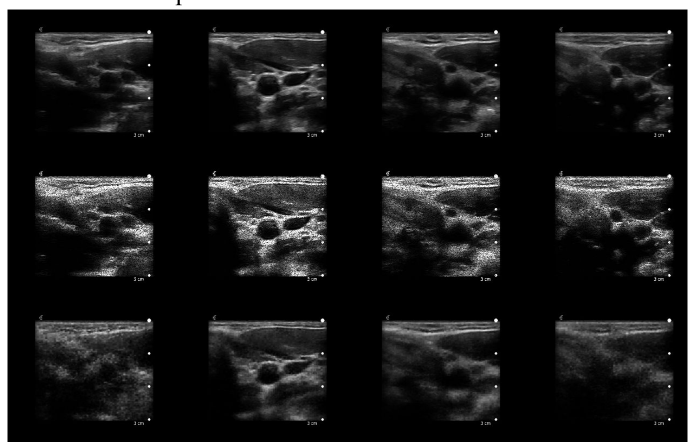

Figure 7: Top row - original images; Middle row - corrupted imagesBottom row - de-corrupted images

We used a relatively small dataset in order to train our model. This was because patient ultrasound data is not easily accessible on open-source platforms. The algorithm is also limited to only one type of vasculature structure: the carotid artery.

Our goal currently is to present a proof of concept and provide a foundation for AI driven medical imaging based on ultrasound technology. Our algorithm model is efficient and can successfully enhance and reconstruct ultrasound distortions. It is able to corroborate its benefits and efficiencies, in comparison to the current state of the art denoising algorithm.

CONCLUSION

We demonstrate the use of USDL, a low-cost solution to reduce noise levels in ultrasound diagnostic imaging and obtaining high resolution images with high accuracy. We envision the conjunctional use of ultrasound technology and this algorithm will enable clinicians to improve accuracy for detection of a wide variety of abnormalities, and will hopefully encourage an increased rate of early detection of diseases.

FURTHER RESEARCH AND APPLICATION

Development of a more accurate and robust model
Currently, our model is a proof of concept for a method to image the brain. Our model simulates the images that would be found when taking ultrasound scans of the brain by artificially adding 72% speckle noise to images of the carotid artery at a 4.8cm penetration level. To improve the algorithm, we plan to acquire a larger dataset of images and expand the dataset to cover more vascular structures at varying levels of penetration.

Integration of USDL with ultrasound in real-time.
Our platform is currently intended to be used on static images after an ultrasound scan has been captured. Our next goal is to implement our model on ultrasound for real-time imaging in order to capture movement.

Augmenting USDL with classification techniques.
For further research, we plan to expand our Deep Learning Model by cascading two modules for Image Denoising and Image Classification. This will allow USDL to classify neurological conditions from its denoised scans. Further extension of this will be a great tool for patients to get preliminary diagnosis before visiting a doctor as ultrasound technology is inexpensive and portable.

REFERENCES

[1. Litjens, Geert, et al. ”A survey on deep learning in medical image analysis.” Medical image analysis 42 (2017): 60-88.

[2. Pizurica, Aleksandra, et al. ”A review of wavelet denoising in MRI and ultrasound brain imaging.” Current medical imaging reviews 2.2 (2006): 247-260.

[3. Luchies, Adam C., and Brett C. Byram. ”Deep neural networks for ultrasound beamforming.” IEEE transactions on medical imaging 37.9 (2018): 2010-2021.

[4. Errico, Claudia, et al.” Transcranial functional ultrasound imaging of the brain using microbubble-enhanced ultrasensitive Doppler.” NeuroImage 124 (2016): 752-761.

[5. Nair, Arun Asokan, et al.” A DEEP LEARNING BASED ALTERNATIVE TO BEAMFORMING ULTRASOUND IMAGES.”

[6. Hough-CNN: deep learning for segmentation of deep brain regions in MRI and ultrasound.” Computer Vision and Image Understanding 164 (2017): 92-102.

[7. iek, zgn, et al. ”3D U-Net: learning dense volumetric segmentation from sparse annotation.” International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2016.

[8. Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., Manzagol, P. A. (2010). Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. Journal of machine learning research, 11(Dec), 3371-3408.

[9. Perdios, D., Besson, A., Arditi, M., Thiran, J. P. (2017, September). A deep learning approach to ultrasound image recovery. In Ultrasonics Symposium (IUS), 2017 IEEE International (pp. 1-4). Ieee.

[10. Medasani, S., & Reddy, G. U. (2018, February). Analysis and Evaluation of Speckle Filters for Polarimetric Synthetic Aperture Radar (PolSAR).

[11. Shen, C., Xu, J., Fang X. N., Jing, Y. (2014, November). Anisotropic Complementary Acoustic Metamaterial for Canceling out Aberrating Layers.

[12. Jung, H., Lee, S., Park, S., Kim, B., Kim, J., Lee, I., & Ahn, C. (2015, January). Development of deep learning-based facial expression recognition system. In 2015 21st Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV) (pp. 1-4). IEEE.

[13. Arora, S., & Bhatia, M. P. S. (2018, July). A Robust Approach for Gender Recognition Using Deep Learning. In 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1-6). IEEE.

[14. Nkemelu, D. K., Omeiza, D., & Lubalo, N. (2018). Deep Convolutional Neural Network for Plant Seedlings Classification. arXiv preprint arXiv:1811.08404.