Enhancing Ultrasound Image Quality: A Comprehensive Framework Using Apodization, Augmentation, and Modified U-Net

Eric Dong, Reshma Shaji, Anthony Bertnyk, Christopher Kallweit

University of Waterloo

ABSTRACT

Ultrasound image quality for single angle plane wave images remains a barrier for clinical adoption of high frame rate ultrasound imaging (HiFRUS). Deep learning techniques could be used for enhancing image quality, but data availability and patient privacy make it difficult to acquire images for training. To address this limitation, we employ data augmentation techniques such as apodization and image transformations to increase the amount of data available for training. Using the PICMUS dataset, we train a U-net and a Modified U-net architecture with PixelShuffle layers, evaluating the performance of the Modified U-net models before and after adding the augmented data to the base U-net model. The proposed U-net, trained using the augmented images, showed improvements in PSNR and SSIM: PSNR increased from 28.73 (U-net) to 30.29 (Modified U-net trained with augmentations) for *in vivo* data, 29.10 to 32.68 for *in vitro* data, and *29.34* to 33.10 for *in silico* data. SSIM increased from 0.741 to 0.871 *in vivo*, 0.767 to 0.858 *in vitro* and 0.537 to 0.794 *in silico*. Our framework demonstrates that augmenting data with different apodizations and image transformations strategies can enhance the training of deep learning models for single plane-wave ultrasound imaging.

Index Terms— *Ultrasound Imaging, deep learning, image enhancement, data augmentation, apodization*

1. INTRODUCTION

Clinical ultrasound imaging has become well established as a diagnostic tool for healthcare [1]. By providing real-time images of tissues and organs without the need of ionizing radiation, ultrasound has become the modality of choice for clinicians [2]. The conventional ultrasound imaging approach is to transmit acoustic pulses into the tissue and reconstruct images based on received echoes [3]. Despite its widespread adoption, this traditional approach to ultrasound imaging has inherent trade-offs between the spatial resolution, penetration depth, and frame rate [3]. These limitations stem from constraints of mechanical wave propagation and the scanline image formation paradigm. Ultrasound pulses with higher frequencies can achieve sharper ultrasound images, at a consequence of lowering the penetration of the ultrasound pulses, since acoustic attenuation increases with frequency [4]. Additionally, resolution can be enhanced by using more scanlines to form an image, at the cost of needing more time to form an image [4].

While the wave propagation limitations are currently unavoidable, a new approach known as high frame rate ultrasound (HiFRUS) imaging has allowed for significantly higher frame rates to be achieved [5]. HiFRUS makes use of unfocused transmissions, either plane or diverging, to form images using a single transmission. This technology has allowed for advancements to different clinical applications such as shear wave elastography [6], functional cardiac imaging [7], and complex blood flow visualizations [8]. While HiFRUS has enhanced framerates, it suffers from a limited spatial resolution due to the lack of spatial focusing [4]. One technique to improve the spatial resolution is to perform image compounding with multiple unfocused transmissions [9], but this necessarily reduces the effective frame rate that could be achieved. As such, there is a need to develop image enhancement techniques which can improve the resolution from single plane wave acquisitions to preserve the high frame potential.

Deep learning approaches, such as U-net [10] and other such convolutional neural network architectures [10-11], have been successful in image enhancement tasks. Using complex mappings, learned using low-quality and high-quality image pairs, these architectures have demonstrated success in effectively enhancing images. While a promising solution to the HiFRUS image enhancement problem, the application of these techniques face challenges related to data availability and image diversity [13]. Unlike natural image domains where many datasets are available for training, ultrasound data faces limitations due to patient privacy, acquisition quality, and the requirement for specialized expertise for annotation and data collection [13]. Furthermore, ultrasound scanners often contain proprietary algorithms and hidden features which can limit the generalizability of the images for training. Until large-scale datasets become publicly available, researchers need to make do with the currently available datasets.

In this work, we propose a framework to address the data limitation challenges in super resolution training by investigating different data augmentation strategies to modify the performance of a Modified U-net model. First, we incorporate different apodization techniques applied during the beamforming stage, using Hanning [14], Tukey [15], and Boxcar [14] windowing functions to create different but

physically realistic ultrasound images from the same radiofrequency data. Next, we apply a series of image augmentation transformations such as: horizontal flip, zoom, rotation, crop, translation, and contrast changes Using these data augmentation strategies, we then train a Modified U-net model on the Plane-wave Imaging Challenge in Medical UltrasSound (PICMUS) dataset [16] and evaluate the model's performance before and after including the augmented data. For performance evaluation, we will investigate Peak Signal-To-Noise Ratio (PSNR) [17] and Structural Similarity Index Measure (SSIM) [18], which are a mathematical metric comparing pixel-level differences and a perceptual metric for assessing structural information between images, respectively.

2. METHODS

2.1 PICMUS Dataset

Our study leveraged PICMUS dataset [16]. The goal of this dataset was to challenge researchers to find an optimal strategy to create image enhancement tools for ultrasound plane-wave imaging. The dataset includes four *in silico* (simulation data), four *in vitro* (experimental data, measured on something which mimics the human body), and four *in vivo* (measured on a living organism) acquisitions. Each dataset includes the raw radiofrequency (RF) data for a series of 75-single angle plane wave transmissions, ranging from [-16°:16°]. For training, the input low-resolution/high-resolution image pair consisted of an image generated from a single plane wave transmission (low-resolution) and a compounded image generated by compounding all 75 single angle plane wave transmissions. As such, each of the 12 acquisition sets contain the RF data which can generate 900 total images for training or validation.

2.2 Delay and Sum Beamforming

To convert the RF data to ultrasound images for training, delay and sum beamforming was used [19]. To briefly summarize, for a certain pixel of interest, a delay for each transducer element channel is calculated based on the two-way propagation time of flight of the ultrasound pulse between the element and the pixel. These delays are then applied to the RF data to extract the data sample which corresponds to that pixel, and the resulting samples from each channel are summed together to determine the reflected ultrasound energy from that pixel. By repeating this for a range of pixels, and mapping the intensity to a colormap, an ultrasound image is formed. The summing expression can be expressed by Equation (1) as:

$$I(x,y) = \frac{1}{K} \sum_{i=1}^{K} RF(sample, i) \qquad (1)$$

where, *I(x,y)* is the image intensity for a pixel at position *(x,y)*, RF is the radio-frequency data matrix, *sample* is the sample estimated from the two-way time of flight to the pixel, and K is the number of transducer elements.

2.3 Apodization

When evaluating the spatial resolution of an ultrasound image, the aperture size (the width of the transducer) used in the beamforming process plays an important role [3]. As the aperture size increases, so does the lateral resolution for the far field of the ultrasound image. However, using all the ultrasound elements can cause side-lobe artifacts [3], deteriorating overall image quality. These side-lobes can be suppressed using apodization windows, such as Boxcar, Hanning, and Tukey [3]. The formulation for beamforming with an apodization can be seen in Equation (2):

$$I(x,y) = \frac{1}{K} \sum_{i=1}^{K} w(i) * RF(sample, i) \qquad (2)$$

Here, *w(i)* represents the weighting function used for a specific transducer element. The other variables are the same as Equation (1). Equation (1), in turn, represents a Boxcar weighting function using all transducer elements. The trade-off of adding apodization to remove side-lobes is that resolution is sacrificed [14]. For our analysis, we generated beamformed images using four different apodization schemes: No Apodization, Boxcar(K/2), Hanning(K), and Tukey 25%. Using these various apodization strategies, the number of images available for training is quadrupled.

2.4 Image Augmentation

A common method to address a low amount of training images for CNNs is image augmentation [20]. This technique increases the amount and diversity of training data for CNNs without physically collecting extra samples [21].

Within medical imaging, geometrical and instrument-specific transformations are used. Geometrical transformations include rotation, flipping, scaling, cropping, translation [20][21]. Ultrasound-specific transformations include augmentations that modify the noise level, occlusions of regions of an image, and FOV changing [20].

In ultrasound imaging, geometrical augmentations are more commonly used and are often more effective with identification and segmentation tasks versus ultrasound-specific augmentations [20]. Ultrasound-specific transformations are typically limited to certain organs or rely on accurate segmentations of an image [20]. Further, implementations of these ultrasound specific augmentations are often not made public and only evaluated with a few conditions/tasks [20].

Among 85 studies that used image augmentation to train CNNs for ultrasound images, geometrical transformations including zoom, translation, rotation, horizontal flipping, and cropping, were often evaluated as one of the top augmentations across six ultrasound environments evaluated using a DICE score (which represents the similarity between two sets of data) [20]. Across the same six environments, a combination of five to six augmentations was consistently evaluated with the highest DICE score in comparison with up to fourteen augmentations [20].

2.5 Super Resolution Model

Image enhancement in this study was approached using convolutional neural networks (CNNs), specifically leveraging the U-Net architecture, originally developed for image segmentation [22]. The standard U-Net features an encoding pathway that captures contextual information and a symmetric decoding pathway that enables precise localization, connected via skip connections to preserve high-resolution details. Traditionally, the decoding pathway utilizes deconvolution layers for upsampling.

While U-Net has been previously applied to plane wave image enhancement [23], reviews comparing it to other CNNs for super-resolution showed that earlier U-Net implementations were outperformed by models like Efficient Sub-Pixel Convolutional Neural Network (ESPCN) [24], particularly noted for its innovative PixelShuffle operation which replaces deconvolutions for more effective upsampling [25]. However, subsequent modifications to the U-Net architecture demonstrated improved performance on specific tasks like ex-vivo tissue imaging [26], suggesting potential for further optimization.

Motivated by these findings, we investigated modifications to the U-Net architecture. As a baseline for comparison, a standard U-Net model was trained using the Tukey25 window subset of images. Our proposed methodology involved adapting the U-Net architecture by incorporating elements from ESPCN. Specifically, inspired by ESPCN's success [27], we replaced the deconvolution layers in the U-Net's decoding pathway with PixelShuffle operations. The resulting proposed architecture, depicted in Figure 1, maintains the core U-Net structure, but substitutes deconvolutions with four sequential PixelShuffle layers in the decoder; each designed for a 2x upscaling factor. Crucially, when evaluated, this proposed U-Net architecture modified with PixelShuffle layers demonstrated superior performance compared to the baseline U-Net architecture trained on the same Tukey25 window subset of images.
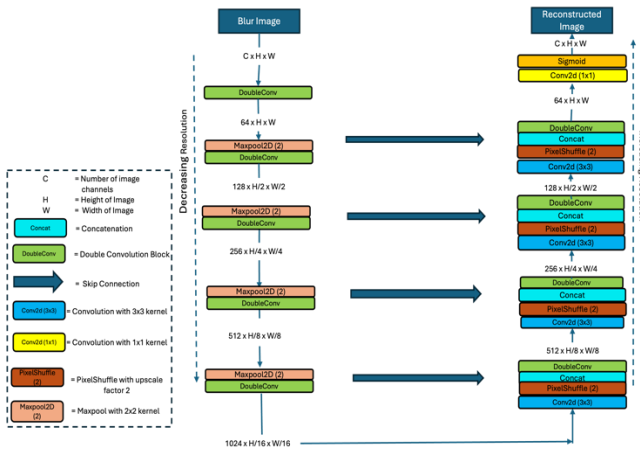

Figure 1: Proposed U-Net Architecture

2.6 Model Training

The model training utilized a dataset comprising a total of 27000 augmented low-resolution and high-resolution image pairs generated with various apodization functions, specifically including Hanning, Tukey25, and Boxcar windows. This diverse dataset, containing augmented examples from all three apodization types, was partitioned for model development: 80% (21600 images) were allocated to the training set, and the remaining 20% (5400 images) were reserved for the validation set, maintaining the mix of apodization types in both sets.

To analyze the impact of the loss function on performance, separate training iterations were conducted on this diverse training set. The training cycle optimized the model using Mean Squared Error (MSE). The training cycle was initially performed using the Tukey25 window subset of images within the training set. Subsequently, data augmentation techniques were applied to the training data. The training cycle was then repeated using the augmented dataset, comprising of all three windows, allowing for comparison of performance outcomes.

2.6.1. Methods to prevent overfitting

To ensure that the trained model generalizes well to new, unseen data, as well as mitigates the risk of overfitting to the training set, preventative methods were employed during the training process. Firstly, the continuous monitoring of the model's performance on the validation set, serving as a proxy for generalization ability. The validation loss was tracked for each training epoch. The model weights with the lowest validation loss throughout the training run were then saved, ensuring that the final model represents peak performance on data not used for gradient updates.

Furthermore, employing early stopping was a key regularization technique; a method which also relied on the monitoring of the validation loss. If the validation loss did not show improvement for a patience threshold of 5 consecutive epochs, the training process was automatically halted. This prevented the model from continuing to learn noise or specific patterns that are unique to the training data after it has stopped improving on the validation set. This effectively determined the optimal training duration and prevented overfitting.

Lastly, an adaptive learning rate scheduler was utilized for training dynamics refinement. The scheduler adjusted the learning rate based on the validation loss trend. More specifically, if the validation loss had slowed or stalled with the current learning rate, the scheduler would kick in and reduce the learning rate. This allowed for finer adjustments to the model weights during later training stages, helping to fine-tune convergence towards a more ideal minimum in the loss landscape, and potentially improving the final model performance (beyond what may be achievable using a fixed learning rate).

2.7 Quality Metric Comparisons

For evaluating the performance of our model, both PSNR and SSIM were used. These metrics compare the high-resolution reference image (I_r) to the output model (I_o).

2.7.1. PSNR

Peak signal-to-noise ratio is a simple and computationally efficient metric often used for image quality assessment, with a higher PSNR indicating better image quality. PSNR is defined by Equation (3) as:

$$PSNR = 20 \log_{10} \frac{I_{max}}{\sqrt{MSE}} \qquad (3)$$

where I_{max} is the maximum possible pixel value, and MSE is the mean squared error. The MSE is derived as seen in Equation (4):

$$MSE = \frac{1}{NM} \sum_{n=1}^{N} \sum_{m=1}^{M} \left(I_0(n,m) - I_r(n,m) \right)^2 \qquad (4)$$

where N and M represent the size of the matrix, and n and m represent the individual pixel locations in the image.

2.7.2. SSIM

Structural Similarity Index Measure is a tool which estimates how similar two images are as though perceived by human eyes. This is done by evaluating the luminance (L), contrast (C), and structure (S). The final value ranges from -1 to 1, with -1 indicating completely different images, and 1 indicating perfect similarity. The general form of SSIM [18] is often defined as in Equation (5):

$$SSIM = \frac{(2\mu_{I_0}\mu_{I_r} + C_1)(2\sigma_{I_0 I_r} + C_2)}{(\mu_{I_0}^2 + \mu_{I_0}^2 + C_1)(\sigma_{I_0}^2 + \sigma_{I_r}^2 + C_2)} \qquad (5)$$

where μ_{I_0} and μ_{I_r} refer to the mean of the mean of I_0 and I_r respectively, σ_{I_0} and σ_{I_r} refer to the variance of I_0 and I_r, $\sigma_{I_0 I_r}$ refers to the variance between I_0 and I_r, and C_1 and C_2 are constants to maintain numerical stability.

3. RESULTING IMAGES AND COMPARISONS

3.1 Apodization Adjustment

A subset of the differently apodized images can be seen in Figure 2. From the *in vitro* point target images, Figure 2(a)-(d), slight changes in the resolution of the point targets can be identified, as is expected from the different apodization schemes. Here, the Boxcar and Tukey25 apodization schemes provide the sharpest targets.

The *in vivo* images, with different apodizations, showing the long-axis view of the carotid artery can be seen Figure 2(e)-(h). These *in vivo* images highlight the effect of the different apodizations more clearly. While the speckle pattern in Figure 2(e) appears to have a finer resolution, there are ringing artifacts near the top of the image. These artifacts are suppressed with different apodizations, as shown in Figure 2(f)-(h).
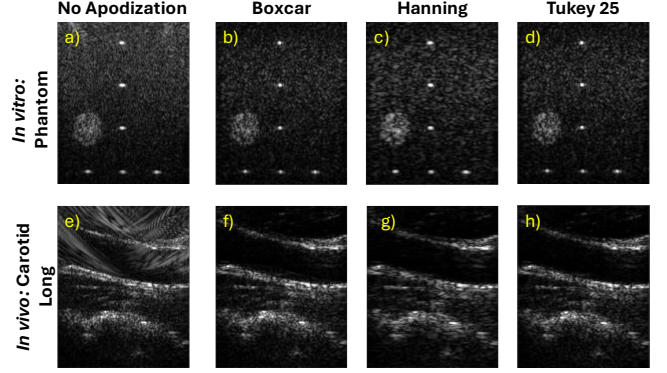

Figure 2: Example images with different apodizations. Images (a)-(d) show the point targets in the CIRS phantom with different apodization types: No Apodization, Boxcar, Hanning, and Tukey25, respectively. Images (e)-(h) show the in vivo images with the same respective apodizations

3.2 Image Augmentation

Six image augmentations were used, including: horizontal flip, zoom, rotation, crop, translation, and contrast changes. An example is shown in Figure 3(a) and 3(b), showing the original low-resolution image and the resulting image after augmenting, respectively. The augmentations include zooming, rotation, cropping, and contrast changes.
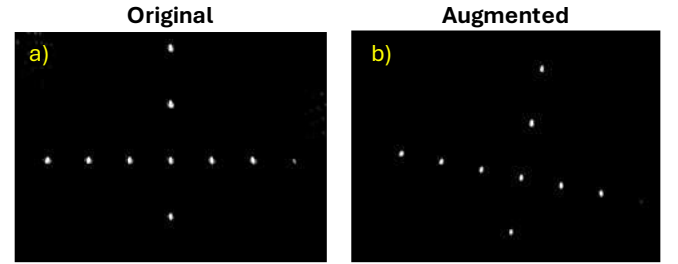

Figure 3: Original low-resolution image (a) and augmented image (b)

3.3 Quantitative Assessment of Model Enhancements

The evaluation was conducted on the original images consisting of 900 low-resolution and high-resolution image pairs.

Table 1 and Table 2 reveal clear performance improvements when comparing the Original U-Net and the Modified U-Net across three different data domains, *in vivo*, *in vitro*, and *in silico*. The metrics under scrutiny, PSNR (Table 1) and SSIM (Table 2), consistently favor the Modified U-Net, indicating a better ability to enhance image quality.

The *in vivo* results show that even under real biological conditions, where imaging noise and variability are high, the Modified U-Net achieves higher PSNR and SSIM than the Original U-Net. Although the magnitude of these gains is more modest compared to *in silico* results, they nevertheless

confirm that the proposed architecture refinements improve image fidelity and perceptual quality in live-tissue scenarios.

For *in vitro* data acquired under more controlled settings, the Modified U-Net delivers a noticeable increase in PSNR (from around 29.1 to 32.7) and SSIM (from about 0.77 to 0.86). These improvements illustrate that, even when images are less affected by real-world noise, the adjusted network architecture can still extract and leverage finer structural details for enhanced reconstruction quality.

The largest quantitative gains appear in *in silico*. PSNR rises substantially from approximately 29.3 with the Original U-Net to over 33.1 with the Modified U-Net, while SSIM sees a similarly significant jump. This outcome emphasizes how the enhanced architecture excels in datasets that are more uniform and structured, allowing the network to capture nuanced spatial features with high consistency.

PSNR	*In vivo*	*In vitro*	*In silico*
Original U-net	28.73	29.10	29.34
Proposed U-net	28.78	26.19	32.46
Proposed U-net with augmented images	30.29	32.68	33.10

Table 1: PSNR comparison: Original, Proposed U-Net and Proposed U-Net with augmented images (*in vivo*/*in vitro*/*in silico*))

SSIM	*In vivo*	*In vitro*	*In silico*
Original U-net	0.741	0.767	0.537
Proposed U-net	0.817	0.767	0.704
Proposed U-net with augmented images	0.871	0.858	0.794

Table 2: SSIM comparison: Original, Proposed U-Net and Proposed U-Net with augmented images (*in vivo*/*in vitro*/*in silico*)
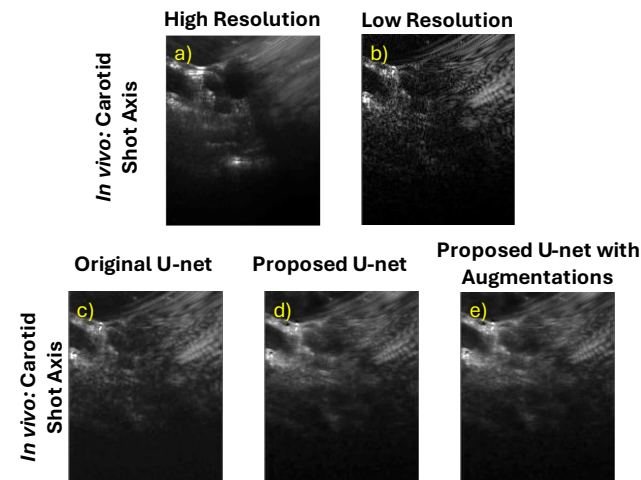

Figure 4: Example of reconstructed and the original images. Images (a) and (b) show a high and low resolution image pair from the test set. Images (c)-(e) show the reconstructed image by the original U-Net model trained with

Tukey25 window images, the proposed U-Net model trained with Tukey25 window images, and the proposed U-Net model trained with augmented images respectively

4. DISCUSSION

Our study demonstrates that the training of deep learning models for plane-wave ultrasound image quality enhancement can be improved using different image augmentation strategies. By quadrupling the effective dataset through the various apodization schemes and applying additional image transformations, we addressed one of the limitations in medical deep learning – insufficient data.

Additionally, our findings highlight the robustness of the Modified U-Net design. Whether dealing with naturally noisy *in vivo* images, semi-controlled *in vitro* conditions, or highly structured *in silico* datasets, the revised architecture consistently yields superior quantitative (PSNR) and perceptual (SSIM) outcomes. The results indicate initial improvements in image quality due to the PixelShuffle operation added to U-Net.

4.1 Additional Considerations for Future Training

While our study demonstrates the usefulness of the image augmentation strategies, there is still a limit to the generalizability of the framework. To test and improve the model further, additional studies should be performed to investigate its performance across different anatomical structures beyond the carotid artery.

The evaluation of different apodization schemes could also be investigated further. We have only looked at a subset of the many available apodization schemes, and so there is the potential for further improvement by adding more apodization schemes. Future work could benefit from training additional models with different apodizations to identify the best apodization windows for ultrasound data augmentation without adding unnecessary data to the dataset.

Additional image quality metrics could also be of benefit in both the training of the model along with the analysis of the model's performance. Here, we have only looked at two common metrics, but other metrics such as contrast to noise ratio [28] and full width at half maximum [29] could also be analyzed to further demonstrate the capabilities of the model.

5. CONCLUSION

U-Net is a valuable tool for medical image processing. In this work, we have demonstrated how the use of apodization, augmentation, and the PixelShuffle operation improve U-Net's ability to enhance plane-wave ultrasound images. The proposed architecture and training methodology robustly improves quality across a variety of testing conditions. Further experimentation is required to determine the ideal set of apodization schemes for training this architecture.

6. REFERENCES

[1] M. A. Schellpfeffer, "Ultrasound imaging in research and clinical medicine," *Birth Defects Res. Part C Embryo Today Rev.*, vol. 99, no. 2, pp.

[2] M. G. Andreassi, "The biological effects of diagnostic cardiac imaging on chronically exposed physicians: the importance of being non-ionizing," *Cardiovasc. Ultrasound*, vol. 2, no. 1, p. 25, Nov.

[3] R. S. C. Cobbold, *Foundations of Biomedical Ultrasound*. Oxford University Press, 2006.

[4] A. Ng and J. Swanevelder, "Resolution in ultrasound imaging," *Contin. Educ. Anaesth. Crit. Care Pain*, vol. 11, no. 5, pp. 186–192, Oct.

[5] J. S. Au, R. L. Hughson, and A. C. H. Yu, "Riding the Plane Wave: Considerations for In Vivo Study Designs Employing High Frame Rate Ultrasound," *Appl. Sci.*, vol. 8, no. 2, Art. no. 2, Feb.

[6] M. Strachinaru *et al.*, "Local myocardial stiffness variations identified by high frame rate shear wave echocardiography," *Cardiovasc. Ultrasound*, vol. 18, no. 1, p. 40, Sep.

[7] A. Azarine, F. Scalbert, and P. Garçon, "Cardiac functional imaging," *Presse Médicale*, vol. 51, no. 2, p. 104119, Jun.

[8] J. A. Jensen, S. I. Nikolov, A. C. H. Yu, and D. Garcia, "Ultrasound Vector Flow Imaging—Part I: Sequential Systems," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 63, no. 11, pp. 1704–1721, Nov.

[9] "Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography | IEEE Journals & Magazine | IEEE Xplore." Accessed: Apr. 14, 2025.

[10] P. Tang, L. Li, Y. Xue, M. Lv, Z. Jia, and H. Ma, "Real-World Underwater Image Enhancement Based on Attention U-Net," *J. Mar. Sci. Eng.*, vol. 11, no. 3, Art. no. 3, Mar.

[11] Z. Zhou, Y. Wang, J. Yu, Y. Guo, W. Guo, and Y. Qi, "High Spatial–Temporal Resolution Reconstruction of Plane-Wave Ultrasound Images With a Multichannel Multiscale Convolutional Neural Network," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 65, no. 11, pp. 1983–1996, Nov.

[12] Z. Zhou, Y. Wang, Y. Guo, Y. Qi, and J. Yu, "Image Quality Improvement of Hand-Held Ultrasound Devices With a Two-Stage Generative Adversarial Network," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 1, pp. 298–311, Jan.

[13] M. Mendez, S. Sundararaman, L. Probyn, and P. N. Tyrrell, "Approaches and Limitations of Machine Learning for Synthetic Ultrasound Generation," *J. Ultrasound Med.*, vol. 42, no. 12, pp. 2695–2706, 2023.

[14] X. Zhao and B. Zhao, "The Study of Apodization of Imaging Fourier Transform Spectrometer in the Spectrum Reconstruction," in *Advances in Computer Science and Information Engineering*, D. Jin and S. Lin, Eds., Berlin, Heidelberg: Springer, 2012, pp. 677–682.

[15] P. S. Sheeran, T. J. Czernuszewicz, K. H. Martin, and B. C. Byram, "Binary apodization schemes for plane wave transmits," in *2012 IEEE International Ultrasonics Symposium*, Oct. 2012, pp. 2145–2148.

[16] H. Liebgott, A. Rodriguez-Molares, F. Cervenansky, J. A. Jensen, and O. Bernard, "Plane-Wave Imaging Challenge in Medical Ultrasound," in *2016 IEEE International Ultrasonics Symposium (IUS)*, Sep. 2016, pp. 1–4.

[17] A. Horé and D. Ziou, "Image Quality Metrics: PSNR vs. SSIM," in *2010 20th International Conference on Pattern Recognition*, Aug. 2010, pp. 2366–2369.

[18] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr.

[19] V. Perrot, M. Polichetti, F. Varray, and D. Garcia, "So you think you can DAS? A viewpoint on delay-and-sum beamforming," *Ultrasonics*, vol. 111, p. 106309, Mar.

[20] A. Tupper, Gagné Christian, "Revisiting Data Augmentation for Ultrasound Images," *arXiv*, 2025

[21] F. Garcea, Serra, Lamberti, Morra Alessio, Fabrizio, Lia, "Data augmentation for medical imaging: A systematic literature review," *Comput. Biol. Med.*, vol.

[22] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2015*, 2015, pp. 234–241.

[23] D. Perdios, M. Vonlanthen, A. Besson, F. Martinez, M. Arditi, and J. P. Thiran, "Deep Convolutional Neural Network for Ultrasound Image Enhancement," in *2018 IEEE International Ultrasonics Symposium (IUS)*, 2018, pp. 1–4.

[24] E. Agustsson and R. Timofte, "NTIRE 2017 Challenge on Single Image Super-Resolution: Dataset and Study,"

[25] W. Shi et al, "Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 1874–1883.

[26] Sobhan Goudarzi and Hassan Rivaz, "Deep Ultrasound deposing without clean data," 2022

[27] F. Anjidani and D. Rakhmatov, "Coupling Fast Superresolution CNN with Fast Plane-Wave Fourier-Domain Beamforming," in *2022 IEEE International Ultrasonics Symposium (IUS)*, Venice, Italy: IEEE, Oct. 2022, pp. 1–5.

[28] P. F. Stetson, F. G. Sommer, and A. Macovski, "Lesion contrast enhancement in medical ultrasound imaging," *IEEE Trans. Med. Imaging*, vol. 16, no. 4, pp. 416–425, Aug.

[29] R. A. Harris, D. H. Follett, M. Halliwell, and P. N. Wells, "Ultimate limits in ultrasonic imaging resolution," *Ultrasound Med. Biol.*, vol. 17, no. 6, pp. 547–558, Jan.