Dynamic Range-Invariant GAN Reconstruction via Optimized Target Training in Medical Ultrasound Imaging

Dynamic Range-Invariant GAN Reconstruction via Optimized Target Training in Medical Ultrasound Imaging

Silvia Seoni[1], Giulia Matrone[2], Massimo Salvi[1], Kristen M. Meiburger[1]

[1]Biolab, Polito[BIO]Med Lab, Department of Electronics and Telecomunications, Politecnico di Torino, Torino, Italy
[2]Department of Electrical, Computer and Biomedical Engineering, University of Pavia, via Ferrata 5, 27100 Pavia, Italy

***Abstract*—Ultrasound imaging is sensitive to operator-dependent parameters such as dynamic range (DR), when considering 8-bit reconstructed images, which can compromise both clinical interpretation and the reliability of artificial intelligence (AI)-based reconstruction pipelines. This work validates an automatic dynamic range optimization (autoDR) method as a standardized training reference for Generative Adversarial Network (GAN) models. By evaluating GAN robustness to DR variability, we demonstrate that autoDR enables consistent, high-quality reconstructions across diverse acquisition settings, outperforming classical enhancement techniques. These findings highlight autoDR as a practical solution for reducing operator dependence, improving reproducibility, and a robust reference standard for GAN-based ultrasound image reconstruction when there is no access to raw radiofrequency data.**

Keywords—ultrasound, beamforming, generative adversarial networks, dynamic range

I. Introduction

Ultrasound (US) imaging is widely employed in diagnostic and interventional medicine due to its portability, real-time capability, and absence of ionizing radiation. The visual quality and interpretability of 8-bit US images strongly depend on post-processing parameters, among which the beamforming algorithm and dynamic range (DR) play a critical role [1]. Beamforming algorithms take in input the raw signals and reconstruct the image based on a specific method, such as the traditional delay and sum (DAS) or other advanced algorithms, like e.g. the filtered delay multiply and sum (FDMAS) [2]. After beamforming, to create a final 8-bit image, a DR value must be selected, which can alter contrast, visibility of fine anatomical structures and quantitative image analysis outcomes [3].

Recent developments in deep learning and generative models, in particular generative adversarial networks (GANs) have shown promise in reconstructing higher-quality US images from lower-quality or alternative beamforming methods. The majority of methods found in literature depend on ready access to the raw data for image reconstruction [4, 5] or are focused on image denoising [6, 7]. In one of our recent studies, we presented a GAN-based method for beamforming domain transfer [8] where plane wave (PW) 8-bit DAS images were provided as input to the GAN, and the target image was the focused wave 8-bit FDMAS reconstructed image. This study demonstrated how the generative adversarial approach can be a viable option when there is no access to raw radiofrequency (RF) data and it furthermore underlined how radiomic texture features can be employed to validate effective beamforming domain transfer [8].

Moreover, it is well known that data harmonization is crucial for many artificial intelligence (AI) image-based tasks [9], and GANs have effectively been employed for color constancy and stain normalization applications [10]. In this study, we investigate the performance of three GAN architectures (Pix2Pix, CycleGAN, and Pyramidal Pix2Pix) trained to reconstruct high-quality focused FDMAS images from PW DAS inputs with varying DRs. We specifically evaluate whether training against an automatic DR optimization (autoDR) [3] reference image yields robust reconstructions across diverse input conditions.

II. Materials and Methods

A. Dataset description

The dataset consisted of 980 images collected from 14 healthy volunteers, after the study protocol received approval from the Local Ethics Committee (n. 24780/2022, approved on 19/07/2022). Informed consent was obtained from all participants prior to data collection. Imaging covered five musculoskeletal regions: gastrocnemius lateralis, gastrocnemius medialis, vastus lateralis, vastus medialis, and biceps. Additionally, carotid artery images were acquired bilaterally in both transverse and longitudinal orientations.

The RF data were collected using both PW and focused transmitting sequences using the Verasonics Vantage™ Research Ultrasound system equipped with an L11-5v linear array. Five different dynamic range values were employed to reconstruct the PW DAS images: 50 dB, 55 dB, 60 dB, 65 dB, and 70dB. An optimized automatic DR was employed instead for reconstructing the target focused FDMAS images. Further details on image acquisition parameters and beamforming methods are reported in our previous study [8]. In total, 840 images were employed for training and 140 images were employed for testing, and the dataset was partitioned on a patient-level basis to avoid data leakage. Fig. 1 displays an example image of the input PW DAS images with two different DR values at 50 dB and 70 dB (first row), highlighting strong intensity variation. The second row displays the target focused FDMAS images, which provide a consistent representation across DRs due to autoDR preprocessing.

B. GAN architectures

We implemented a domain transfer framework where input images consisted of 8-bit PW DAS reconstructions generated at the various DR levels and the target images were consistently defined as focused FDMAS reconstructions processed through an automatic dynamic range (autoDR)

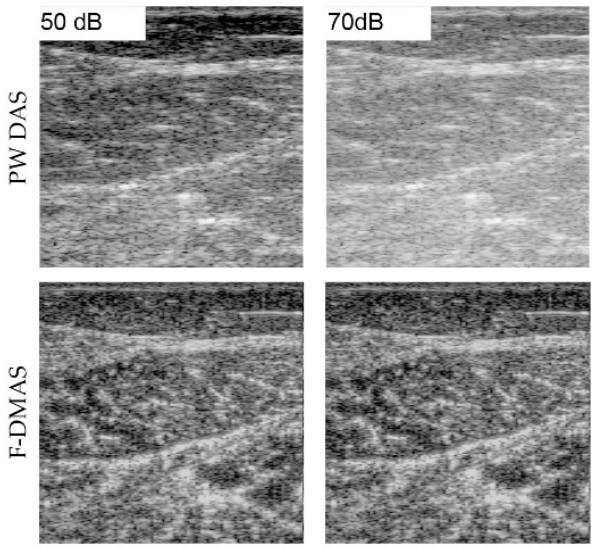

Fig. 1. Example images. Top row: PW DAS images with 2 different DR values (GAN input). Second row: focused FDMAS reconstructed images using the autoDR method (GAN target).

pipeline, ensuring that the target intensity distribution was standardized.

Three GAN models were evaluated:

- Pix2Pix, a paired supervised image-to-image translation network [11];
- CycleGAN [12], capable of learning mappings between domains without paired training;
- Pyramidal Pix2Pix [13], which integrates multi-scale feature extraction to enhance structural preservation.

For benchmarking, two widely used non-AI enhancement methods were included:

- Contrast-Limited Adaptive Histogram Equalization (CLAHE) [14];
- Histogram Matching (HM).

C. Validation metrics

Performance was assessed using the Root Mean Square Error (RMSE) and the Peak Signal to Noise Ratio (PSNR) between the GAN outputs or benchmark-enhanced images and the autoDR-optimized FDMAS targets.

III. Results

Fig. 2 depicts example input and target images along with the output images from the three trained GANs, which effectively normalized input discrepancies and reproduced a target-like representation, preserving both contrast and fine structural detail.

Table I reports the overall average and standard deviation of the RMSE and PSNR values for all analyzed methods. Fig. 3 instead reports the RMSE values for all tested approaches across the different DR conditions on the test set. The GAN-based models consistently achieved low and stable RMSE scores irrespective of the input DR, with the Pyramidal Pix2Pix variant showed slightly superior stability compared to Pix2Pix and CycleGAN. In contrast, classical enhancement methods, especially CLAHE, displayed noticeably higher and more variable RMSE values, indicating reduced robustness to DR variability.

TABLE I. Validation results comparing the methods against the target focused FDMAS reconstructed image

Method	RMSE ↓	PSNR (dB)↑
DAS	0.161 ± 0.031	16.0 ± 1.6
CLAHE	0.171 ± 0.029	15.5 ± 1.4
HM	0.131 ± 0.010	17.7 ± 0.7
Pix2Pix GAN	0.131 ± 0.008	17.6 ± 0.5
Cycle GAN	0.131 ± 0.007	17.7 ± 0.5
Pyramidal GAN	**0.126 ± 0.008**	**18.0 ± 0.6**

The best values are highlighted in bold

IV. Discussion and Conclusions

These findings support two major conclusions:

- they confirm that GANs learn robust mappings independent of input DR in a medical US setting, as also shown in previous studies focusing on other imaging modalities: all three architectures generalized well to varying DR inputs, demonstrating their ability to distinguish acquisition variability from underlying tissue features;
- the autoDR method can serve as an effective training reference: by providing a standardized target, autoDR mitigated operator dependence and enabled the networks to converge toward consistent reconstructions.

When compared to traditional methods, the GAN-based techniques provided more robust results with lower RMSE and higher PSNR values, although the HM method also demonstrated similar results. It should be emphasized that lower RMSE values are not necessarily expected for the original DAS, CLAHE, and HM approaches, since the

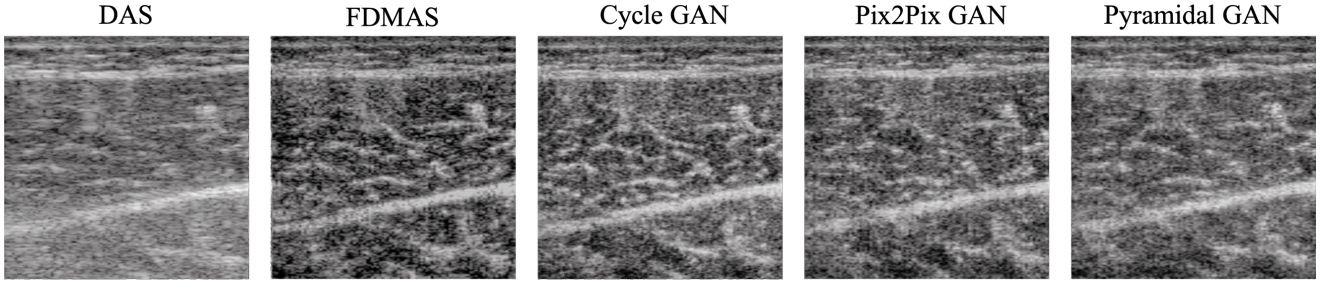

Fig. 2. Example images. DAS: plane-wave delay-and-sum input image for GAN training; FDMAS: focused filtered-delay-multiply-and-sum target image for GAN training; Cycle GAN: unpaired Cycle GAN output; Pix2Pix GAN: classic Pix2Pix paired GAN output; Pyramidal GAN: pyramidal paired GAN output.

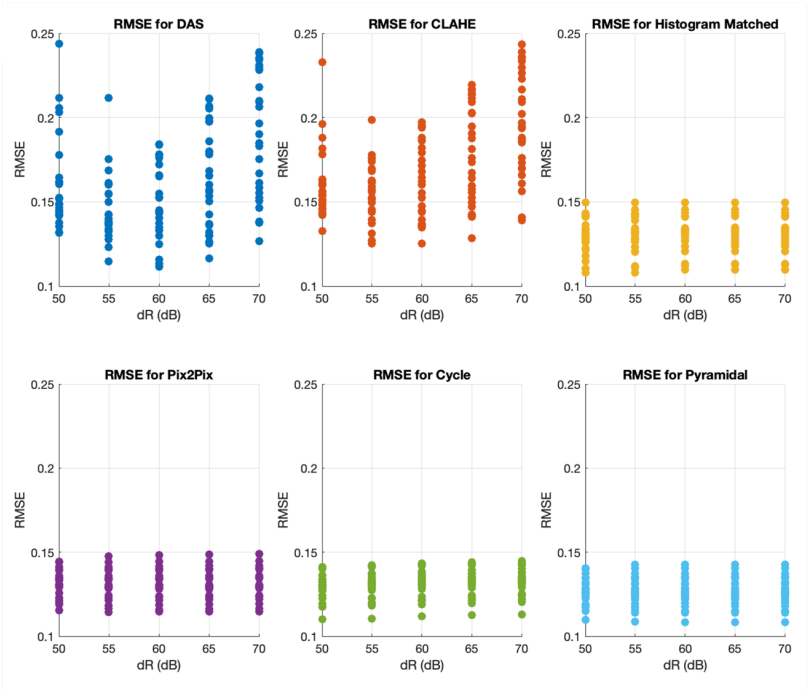

Fig. 3. RMSE plots for the investigated methods when compared against the target focused FDMAS reconstructed image. DAS: plane-wave delay and sum; CLAHE: contrast limited adaptive histogram equalization; Pix2Pix: paired Pix2Pix; Cycle: unpaired CycleGAN; Pyramidal: paired pyramidal Pix2Pix.

reconstructions are obtained through different techniques. Nevertheless, the results were comparable, which also highlights an important limitation of RMSE when used in isolation for medical imaging assessment. As highlighted in recent literature, conventional full-reference metrics often fail to reflect perceptual or clinically meaningful differences [15]. For this reason, it is crucial to assess the RMSE together with other qualitative and quantitative parameters, such as visual assessment, texture-oriented measures, and task-specific evaluation criteria, as we did in our recent study [8].

This study validates automatic dynamic range optimization (autoDR) as a robust reference standard for GAN-based ultrasound image reconstruction. GANs trained with autoDR targets were able to effectively normalize intensity inconsistencies caused by user-defined DR variations, yielding stable, robust outputs. Our future work will extend this framework toward exploring interpretability methods to better understand how GANs achieve robustness to DR variability.

REFERENCES

[1] S. Seoni, G. Matrone, and K. M. Meiburger, “Texture analysis of ultrasound images obtained with different beamforming techniques and dynamic ranges–A robustness study”, Ultrasonics, vol. 131, p. 106940, 2023, doi: 10.1016/j.ultras.2023.106940

[2] G. Matrone, A. S. Savoia, G. Caliano, and G. Magenes, “The delay multiply and sum beamforming algorithm in ultrasound B-mode medical imaging”, IEEE Trans Med Imaging, vol. 34, no. 4, pp. 940–949, 2015, doi: 10.1109/TMI.2014.2371235.

[3] K. M. Meiburger, S. Seoni, and G. Matrone, “Automatic Dynamic Range Estimation for Ultrasound Image Visualization and Processing”, IEEE Int. Ultrason. Symp. IUS, no. 2, pp. 2–5, 2020.

[4] E. Bosco, E. Spairani, E. Toffali, V. Meacci, A. Ramalli, and G. Matrone, “A Deep Learning Approach for Beamforming and Contrast Enhancement of Ultrasound Images in Monostatic Synthetic Aperture Imaging: A Proof-of-Concept”, IEEE Open J Eng Med Biol, vol. 5, pp. 376–382, 2024, doi: 10.1109/OJEMB.2024.3401098.

[5] D. Hyun et al., “Deep Learning for Ultrasound Image Formation: CUBDL Evaluation Framework and Open Datasets”, IEEE Trans Ultrason Ferroelectr Freq Control, vol. 68, no. 12, pp. 3466–3483, Dec. 2021, doi: 10.1109/TUFFC.2021.3094849.

[6] 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. 2020, doi: 10.1109/TBME.2019.2912986.

[7] H. G. Khor, G. Ning, X. Zhang, and H. Liao, “Ultrasound Speckle Reduction Using Wavelet-Based Generative Adversarial Network”, IEEE J Biomed Health Inform, vol. 26, no. 7, pp. 3080–3091, Jul. 2022, doi: 10.1109/JBHI.2022.3144628.

[8] S. Seoni, M. Salvi, G. Matrone, F. Lapia, C. Busso, M. A. Minetto, and K. M. Meiburger, “Adversarial learning for beamforming domain transfer in ultrasound medical imaging”, Ultrasonics, p. 107749, 2025, doi: 10.1016/j.ultras.2025.107749.

[9] S. Seoni, A. Shahini, K. M. Meiburger, F. Marzola, G. Rotunno, U. R. Acharya, and M. Salvi, “All you need is data preparation: A systematic review of image harmonization techniques in multi-center/device studies for medical support systems”, Comput. Methods Programs Biomed., vol. 250, p. 108200, 2024, doi: 10.1016/j.cmpb.2024.108200.

[10] M. Salvi, F. Branciforti, F. Molinari, and K. M. Meiburger, “Generative models for color normalization in digital pathology and dermatology: Advancing the learning paradigm”, Expert Syst. Appl., vol. 245, p. 123105, 2024, doi: 10.1016/j.eswa.2023.123105.

[11] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-image translation with conditional adversarial networks”, Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, vol. 2017-January, pp. 5967–5976, Nov. 2017, doi: 10.1109/CVPR.2017.632.

[12] J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks”, Proceedings of the IEEE International Conference on Computer Vision, vol. 2017-October, pp. 2242–2251, Dec. 2017, doi: 10.1109/ICCV.2017.244.

[13] H. Yin, J. Xiao, and H. Chen, “CSPA-GAN: A Cross-Scale Pyramid Attention GAN for Infrared and Visible Image Fusion”, IEEE Trans Instrum Meas, vol. 72, 2023, doi: 10.1109/TIM.2023.3317932.

[14] H. Yin, J. Xiao, and H. Chen, “CSPA-GAN: A Cross-Scale Pyramid Attention GAN for Infrared and Visible Image Fusion”, IEEE Trans Instrum Meas, vol. 72, 2023, doi: 10.1109/TIM.2023.3317932

[15] A. Breger et al., “A Study of Why We Need to Reassess Full Reference Image Quality Assessment with Medical Images”, Journal of Imaging Informatics in Medicine 2025, pp. 1–26, Mar. 2025, doi: 10.1007/S10278-025-01462-1.