Ultrasound Image Beamforming Optimization Using a Generative Adversarial Network

(Article begins on next page)

Ultrasound Image Beamforming Optimization Using a Generative Adversarial Network

Silvia Seoni
Polito[BIO]MedLab, Biolab, Dept. of Electronics and Telecommunications, Politecnico di Torino
Torino, Italy
silvia.seoni@polito.it

Massimo Salvi
Polito[BIO]MedLab, Biolab, Dept. of Electronics and Telecommunications, Politecnico di Torino
Torino, Italy
massimo.salvi@polito.it

Giulia Matrone
Dept. of Electrical, Computer and Biomedical Engineering
University of Pavia
Pavia, Italy
giulia.matrone@unipv.it

Kristen M. Meiburger
Polito[BIO]MedLab, Biolab, Dept. of Electronics and Telecommunications, Politecnico di Torino
Torino, Italy
kristen.meiburger@polito.it

***Abstract*—Recently, research has been focusing on the development of artificial intelligence ultrasound beamforming methods to improve the contrast and resolution of B-mode images. In this work, we propose an innovative beamforming domain transfer method using a generative adversarial network (GAN). The GAN takes as input a plane-wave (PW) delay and sum (DAS) image and generates an image as if it had been acquired using the focused modality and reconstructed with the filtered Delay Multiply and Sum (F-DMAS) beamforming technique. A Verasonics Vantage 256 system (L11-5v linear array) was used to acquire 560 (480 and 80 for train and test set, respectively) *in-vivo* musculoskeletal US images. Images were acquired on five muscles (gastrocnemius lateralis, gastrocnemius medialis, vastus lateralis, vastus medialis, and biceps) on both sides of 14 healthy volunteers (50% female). RF data were acquired both in plane-wave (PW) and focused mode and beamformed using the UltraSound ToolBox (USTB). The DAS beamforming method was employed for PW data, whereas the focused data were reconstructed using F-DMAS. Various dynamic ranges (dR) were employed to create the final 8-bit PW DAS images (dR = 55, 65, 75, 85 dB) while an automatic dR was employed to optimize focused F-DMAS images. A Pix2Pix GAN architecture was designed to formulate the task of beamforming as the translation from one domain (PW DAS image) to another (focused F-DMAS image). Our GAN employed a UNet as the generator and a 3-layer fully convolutional PatchGAN as the discriminator. The proposed GAN architecture shows promising results, generating a GAN image comparable to the F-DMAS image, i.e., in terms of SSIM (0.5183 ± 0.0437 and 0.5152± 0.0519 for GAN images vs DAS images and F-DMAS images vs DAS images). Overall, our GAN enhances image quality and simulates focused F-DMAS beamforming starting from a PW DAS image without needing to access the raw RF data, which is typically unavailable with clinical ultrasound devices.**

Keywords*— ultrasound image enhancement, ultrasound beamforming, deep learning, generative adversarial networks, filtered delay multiply and sum beamforming

I. Introduction

Ultrasound (US) imaging is widely used in clinical settings, as it is a non-invasive and safe method [1]. B-mode images are reconstructed from the raw radiofrequency (RF) data acquired by the transducer through a process called *beamforming*. In addition to the traditional beamforming Delay and Sum (DAS) method [2], there are several alternative methods that aim to improve image quality in terms of resolution and contrast, such as filtered Delay Multiply and Sum (F-DMAS), minimum variance (MV), and coherence factor (CF) methods [2]–[6].

In recent years, several deep learning models have been proposed for beamforming and enhancing ultrasound images [7]–[13]. The first works aimed to develop deep learning models that optimize specific steps of the beamforming technique, such as the estimation of apodization weights in the minimum-variance methods [8]. Other studies showed the feasibility of using deep learning models for the formation of high-quality images from RF data acquired in plane-wave mode and synthetic aperture with a subsampled array [7], [11]. Several deep learning models have also been developed for simultaneously producing the beamformed image and segmentation masks of circular structures present in the image [12], [13]. Nair et al. [14]developed the first model based on a Generative Adversarial Network (GAN) for producing a DAS-like image and segmentation mask from RF data. Other studies designed a GAN model for the beamforming task [15]–[17].

However, all these methods show some limitations. All models were trained with raw data, which are often very difficult to obtain from conventional ultrasound scanners. This entails using mainly simulated data to train the models, which limits the model's generalizability and performance on experimental *in-vivo* images. Furthermore, models were often trained with images generated with standard visualization parameters, such as a dynamic range (dR) value of 60 dB for all image types. Using a standard value of dR, some images may not be optimally displayed. In addition, most of the models focus on generating images in PW mode, although the focused mode is more widely used in clinical settings.

The aim here is to propose a novel beamforming domain transfer method to improve PW DAS image quality by replicating focused imaging with F-DMAS using a GAN, without the need to access raw RF data. This is obtainable thanks to the paradigm shift proposed here in using a Pix2Pix GAN architecture for direct domain transfer from the pixels of the reconstructed PW DAS image to those of the focused F-DMAS image. The rest of this paper is organized as follows: Section II describes the dataset, the network architecture, and the training modality; Section III reports the results obtained using our model; Section IV includes a discussion about the results and future works.

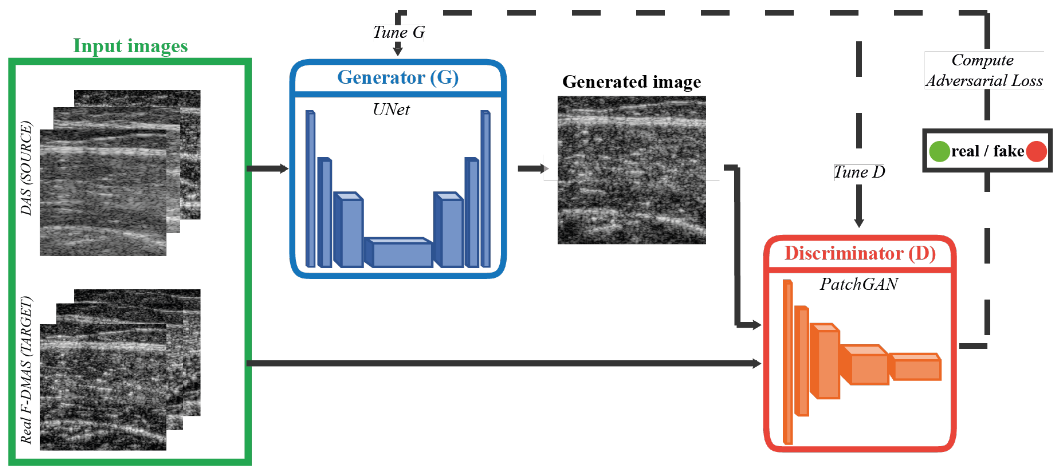

Fig. 1. Overview of our GAN-based beamforming domain transfer algorithm. The PW DAS image is used as the source domain, while the focused F-DMAS image is employed as the target domain. The output of the Generator (generated image), together with the actual F-DMAS image, is passed to the Discriminator (D).

II. Materials and Methods

A. Dataset

Fourteen healthy volunteers were enrolled for dataset creation. The study protocol was approved by the Local Ethics Committee and all enrolled subjects signed an informed consent before the acquisition. The Verasonics Vantage™ Research Ultrasound System was used to acquire RF data from 5 different muscles (gastrocnemius lateralis, gastrocnemius medialis, vastus lateralis, vastus medialis, and biceps) on both sides (right and left). Musculoskeletal ultrasound images were considered based on previous studies that have shown how a quantitative texture analysis can discriminate between healthy and pathological images of the muscles and tendons [18], [19]. An ad-hoc Verasonics acquisition protocol was written to acquire PW and focused images simultaneously, which were then beamformed using the USTB Toolbox [20]. The DAS method was used to reconstruct the PW images, while F-DMAS was employed for RF data acquired in the focused mode. Four different values of dR were employed to display the PW DAS images (55, 65, 75, and 85 dB), while an automatically adjusted dynamic range value was used for the focused F-DMAS images [21]. The final dataset included 560 *in-vivo* musculoskeletal US images, 480 for the training set, and 80 for the test set. The training set was subject-specific, meaning that all images acquired from subjects of the test set were present only in the test set. In this way, the model was tested on all muscle images (gastrocnemius lateralis, gastrocnemius medialis, vastus lateralis, vastus medialis, and biceps) of unseen subjects.

B. Generative Adversarial Network (GAN)

We present here a GAN for domain transfer between US images, from PW DAS to focused F-DMAS images. GANs are unsupervised generative networks that take advantage of adversarial training [22]. Adversarial training consists of a generative and a discriminative model trained through an objective function using a two-player min-max game. The model trains so that the generator learns to fool the discriminator by generating images as similar as possible to the target, while the discriminator classifies the image as real or false, trying not to be fooled. In this study, a Pix2Pix GAN was developed for the beamforming domain transfer task. Our GAN employed a UNet architecture [23] as a generator network and a three-layer fully convolutional PatchGAN [24] as a discriminator. This framework allows the architecture to learn a direct pixel to pixel image transformation, conditioned on the input image. The input size of both generator and discriminator was set to 2048x2048 pixels (original size of the US images employed in this study). Fig. 1 displays the overall architecture of the GAN.

Both generator and discriminator models were trained with the Adam optimizer with an initial learning rate of 10-4 for 150 epochs, and the entire models were updated every 16 images (batch size). The training was performed on a NVIDIA RTX 3090 24 GB using Pytorch framework. After training, the best-stored model of the generator was selected according to loss values.

C. Validation parameters

Two rectangular regions of interest were selected in the muscle images to estimate the contrast, generalized contrast to noise ratio (gCNR) [25], signal to noise ratio (SNR) [26] and peak SNR (PSNR) [26]. We placed one ROI inside the muscle (ROI_1) and the second one on the aponeurosis (ROI_2). Fig. 2 shows an example of the ROIs in an example vastus lateralis image. Structural Similarity index (SSIM) and root-square-mean error (RMSE) were computed on the entire images. These validation parameters were chosen as image quality indices (contrast, gCNR, and SNR) and similarity indicators between the F-DMAS and GAN image (PSNR, SSIM, and RMSE). To estimate the similarity between the real F-DMAS and GAN

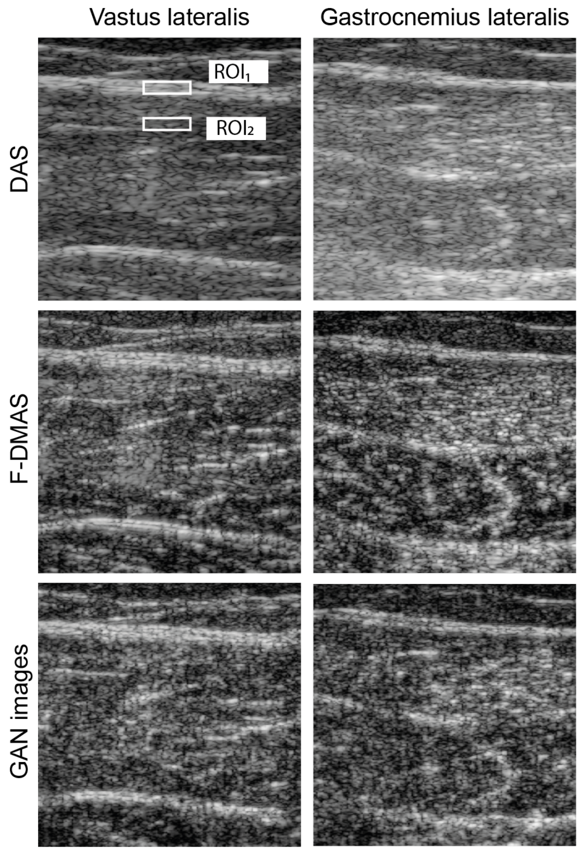

Fig. 2. Comparison between the real DAS, real F-DMAS and GAN images for two sample muscles of the test set.

image, PSNR, SSIM, and RMSE were computed by comparing the F-DMAS image with the DAS image and the GAN image with the DAS image.

III. Results

The results of the validation metrics are listed in Table I and II. Table I reports the quality metrics estimated in DAS, F-DMAS and GAN images. Table II reports the similarity metrics between F-DMAS and GAN images.

Contrast, gCNR, and SNR values show that the F-DMAS and the GAN images are comparable. Indeed, the metrics values estimated in GAN images are similar to the values estimated in F-DMAS images (gCNR=0.44± 0.11 vs 0.43± 0.11 for the F-DMAS and GAN images), while differing from those of DAS images (gCNR=0.37± 0.15 vs 0.43± 0.11 for DAS and GAN images). The obtained PSNR, SSIM, and RSME values demonstrate that the focused F-DMAS images and GAN images are comparable, showing similar values when computed in comparison to the PW DAS images (Table II). Fig. 2 shows the comparison between the real F-DMAS and GAN images of two example images of the vastus lateralis and gastrocnemius lateralis muscles, which were generated from the PW DAS images of the test set.

IV. Discussion

In this work, we propose a Pix2Pix GAN that is designed to formulate the task of beamforming as the translation from one domain (DAS PW image) to another (F-DMAS focused image), enhancing image quality and simulating the F-DMAS beamforming method without the need to access RF data, which is typically not available with clinical ultrasound devices.

The developed model makes it possible to obtain enhanced ultrasound images with increased contrast and resolution that is observable with the F-DMAS beamforming method, starting from an already reconstructed PW DAS image. This can be useful for both clinical and quantitative analyses, for example, texture parameter analysis, which can be influenced by the beamforming method [27]. In addition, enhanced image quality can help the quantitative analyses to determine tendon abnormalities [18] or for the assessment of changes in muscle mass and structure in patients with steroid myopathy [19].

The proposed approach overcomes some of the major limitations of previous works in literature [7]–[13], [28]. The first advantage of the proposed model is the possibility of improving image quality by generating an image with a different acquisition mode from the original one. In fact, the model generates an image as if it were acquired in focused mode starting from an image acquired in PW mode. Our model is therefore useful for improving PW DAS image quality, and unlike other studies in the literature, [8], [12], [13], it does not require raw data which are often not accessible from clinical ultrasound devices. Furthermore, our dataset consists of experimental *in-vivo* muscle images, which were acquired using the Verasonics Vantage™ research ultrasound system and not simulated data, as is usually the case [8], [12], [13].

Moreover, the proposed model also takes into consideration the optimization of visualization parameters, such as the dynamic range. Indeed, by training the model with images obtained with various dR values as source images coupled with one image obtained with an optimized automatic dR as the target, the model learns not only the pixel-to-pixel translation between beamforming methods, but it also learns to generate images with an optimized dynamic range value [21], which improves visualization.

TABLE I. Validation metrics in the test images using manually selected ROIs. Contrast values are expressed in dB

Metrics	PW DAS	Focused F-DMAS	GAN images
Contrast (dB)	3.02± 5.27	2.05± 0.32	2.10± 0.37
gCNR	0.37± 0.15	0.44± 0.11	0.43± 0.11
SNR	4.45± 2.24	2.48± 0.52	2.44± 0.46

TABLE II. PSNR, SSIM, and RSME estimated in test images. PSNR values are expressed in dB

Metrics	PW DAS vs Focused F-DMAS	PW DAS vs GAN images
$PSNR_{ROI1}$ (dB)	18.25±3.39	18.38±3.31
$PSNR_{ROI2}$ (dB)	16.40±3.04	16.02±3.02
SSIM	0.5152 ± 0.0519	0.5183 ± 0.0437
RMSE	47.92 ± 13.17	48.62 ± 13.51

This work is not free from limitations. First, the proposed model was trained and tested using only musculoskeletal ultrasound images. In future works, we will investigate the use of the proposed model with B-mode images acquired from different locations, such as the carotid artery or the thyroid. Moreover, the proposed GAN model can only be used to generate focused F-DMAS images. However, in the literature, there are several other beamforming methods that can improve image quality and can therefore be employed for image quality improvement [3]–[6].

V. Conclusion

In this work, we present a novel method of beamforming domain transfer that does not require access to raw RF data. This is obtained through a GAN that generates a focused image as if it were reconstructed by the F-DMAS method starting from an image acquired in PW mode and reconstructed by the DAS method. Overall, promising results have been obtained, which encourages further analyses aimed at generalizing the use of GANs for the pixel-to-pixel enhancement of ultrasound images.

References

[1] A. Carovac, F. Smajlovic, and Dzelaludin Junuzovic, "Application of Ultrasound in Medicine," *Medicine Review | AIM*, vol. 19, no. 3, pp. 168–171, 2011, doi: 10.5455/aim.2011.19.168-171.

[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] G. Matrone and A. Ramalli, "Spatial coherence of backscattered signals in multi-line transmit ultrasound imaging and its effect on short-lag Filtered-Delay Multiply and Sum beamforming," *Applied Sciences (Switzerland)*, vol. 8, no. 4, 2018, doi: 10.3390/app8040486.

[4] K. W. Hollman, K. W. Rigby, and M. O. Donnell, "Coherence Factor," pp. 1257–1260, IEEE Ultrasonics Symposium. Proceedings. International Symposium1999.

[5] J. Camacho, M. Parrilla, and C. Fritsch, "Phase coherence imaging," *IEEE Trans Ultrason Ferroelectr Freq Control*, vol. 56, no. 5, pp. 958–974, 2009, doi: 10.1109/TUFFC.2009.1128.

[6] J. F. Synnevåg, A. Austeng, and S. Holm, "Benefits of minimum-variance beamforming in medical ultrasound imaging," *IEEE Trans Ultrason Ferroelectr Freq Control*, vol. 56, no. 9, pp. 1868–1879, Sep. 2009, doi: 10.1109/TUFFC.2009.1263.

[7] S. Khan, J. Huh, and J. C. Ye, "Adaptive and Compressive Beamforming Using Deep Learning for Medical Ultrasound," *IEEE Trans Ultrason Ferroelectr Freq Control*, vol. 67, no. 8, pp. 1558–1572, 2020, doi: 10.1109/TUFFC.2020.2977202.

[8] B. Luijten *et al.*, "Deep Learning for Fast Adaptive Beamforming," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2019-May, pp. 1333–1337, 2019, doi: 10.1109/ICASSP.2019.8683478.

[9] A. C. Luchies and B. C. Byram, "Deep Neural Networks for Ultrasound Beamforming," *IEEE Trans Med Imaging*, vol. 37, no. 9, pp. 2010–2021, 2018, doi: 10.1109/TMI.2018.2809641.

[10] Y. Qi, Y. Guo, and Y. Wang, "Image Quality Enhancement Using a Deep Neural Network for Plane Wave Medical Ultrasound Imaging," *IEEE Trans Ultrason Ferroelectr Freq Control*, vol. 68, no. 4, pp. 926–934, 2021, doi: 10.1109/TUFFC.2020.3023154.

[11] B. Luijten *et al.*, "Adaptive Ultrasound Beamforming Using Deep Learning," *IEEE Trans Med Imaging*, vol. 39, no. 12, pp. 3967–3978, 2020, doi: 10.1109/TMI.2020.3008537.

[12] A. A. Nair, K. N. Washington, T. D. Tran, A. Reiter, and M. A. Lediju Bell, "Deep Learning to Obtain Simultaneous Image and Segmentation Outputs from a Single Input of Raw Ultrasound Channel Data," *IEEE Trans Ultrason Ferroelectr Freq Control*, vol. 67, no. 12, pp. 2493–2509, 2020, doi: 10.1109/TUFFC.2020.2993779.

[13] Seoni, G. Matrone, N. Casali, E. Spairani and K. M. Meiburger, "Generalization of a deep learning network for beamforming and segmentation of ultrasound images," *IEEE International Ultrasonics Symposium(IUS),2021*,pp.1-4,*doi:10.1109/IUS52206.2021.9593883*

[14] A. A. Nair, T. D. Tran, A. Reiter, and M. A. L. Bell, "A Generative Adversarial Neural Network for Beamforming Ultrasound Images: Invited Presentation," *2019 53rd Annual Conference on Information Sciences and Systems, CISS 2019*, pp. 1–6, 2019, doi: 10.1109/CISS.2019.8692835.

[15] Y. Wang, K. Kempski, J. U. Kang, and M. A. L. Bell, "A Conditional Adversarial Network for Single Plane Wave Beamforming".

[16] Z. Zhou, Y. Guo, and Y. Wang, "Ultrasound deep beamforming using a multiconstrained hybrid generative adversarial network," *Med Image Anal*, vol. 71, p. 102086, Jul. 2021, doi: 10.1016/J.MEDIA.2021.102086.

[17] J. Tierney, A. Luchies, C. Khan, B. Byram, and M. Berger, "Domain Adaptation for Ultrasound Beamforming," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12262 LNCS, pp. 410–420, 2020, doi: 10.1007/978-3-030-59713-9_40/FIGURES/4.

[18] K. M. Meiburger *et al.*, "Quantitative Analysis of Patellar Tendon Abnormality in Asymptomatic Professional 'Pallapugno' Players: A Texture-Based Ultrasound Approach," *Applied Sciences 2018, Vol. 8, Page 660*, vol. 8, no. 5, p. 660, Apr. 2018, doi: 10.3390/APP8050660.

[19] M. A. Minetto *et al.*, "Ultrasound-based detection of glucocorticoid-induced impairments of muscle mass and structure in Cushing's disease," vol. 42, pp. 757–768, 2019, doi: 10.1007/s40618-018-0979-9.

[20] A. Rodriguez-Molares, O. M. H. Rindal, O. Bernard, H. Liebgott, A. Austeng, and L. Lovstakken, "The ultrasound toolbox," *IEEE International Ultrasonics Symposium, IUS*, Oct. 2017, doi: 10.1109/ULTSYM.2017.8092026.

[21] K. M. Meiburger, "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, doi: 10.1109/IUS46767.2020.9251470.

[22] I. J. Goodfellow *et al.*, "Generative Adversarial Nets", Accessed: Sep. 21, 2022. [Online]. Available: http://www.github.com/goodfeli/adversarial

[23] W. Weng and X. Zhu, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *IEEE Access*, vol. 9, pp. 16591–16603, May 2015, doi: 10.48550/arxiv.1505.04597.

[24] 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. 2016, doi: 10.48550/arxiv.1611.07004.

[25] A. Rodriguez-Molares *et al.*, "The Generalized Contrast-to-Noise Ratio: A Formal Definition for Lesion Detectability," *IEEE Trans Ultrason Ferroelectr Freq Control*, vol. 67, no. 4, pp. 745–759, 2020, doi: 10.1109/TUFFC.2019.2956855.

[26] A. J. Dinu, R. Ganesan, A. A. Kebede, and B. Veerasamy, "Performance analysis and comparison of medical image compression techniques," *2016 International Conference on Control Instrumentation Communication and Computational Technologies, ICCICCT 2016*, pp. 738–745, 2017, doi: 10.1109/ICCICCT.2016.7988050.

[27] S. Seoni, G. Matrone, and K. M. Meiburger, "Robustness Analysis of Texture Features with Different Beamforming Techniques," *IEEE Int. Ultrason. Symp. IUS*, no. 2, 2020.

[28] B. Luijten *et al.*, "Adaptive Ultrasound Beamforming Using Deep Learning," *IEEE Trans Med Imaging*, vol. 39, no. 12, pp. 3967–3978, 2020, doi: 10.1109/TMI.2020.3008537.