UltraUP: A Deep Learning Framework for Real-Time Processing of Ultrasound Images

Simone Cammarasana and **Giuseppe Patanè**
CNR-IMATI, Genova - Italy {simone.cammarasana,giuseppe.patane}@cnr.it

Abstract

Ultrasound (US) is widespread across medical specialities due to its non-invasive nature, portability, cost-effectiveness, and real-time imaging capability. Nevertheless, portable and handheld US systems are limited by speckle noise, low spatial resolution, and computational constraints, which reduce diagnostic accuracy and broader clinical adoption. To address these challenges, we introduce *UltraUP*, a hardware-software prototype for real-time denoising and super-resolution of ultrasound data. UltraUP integrates adaptive edge-preserving denoising with deep learning-based super-resolution to suppress noise, reconstruct missing scan lines, and reveal fine anatomical structures. Optimised for both CPU and GPU platforms, the system delivers low-latency performance suitable for clinical workflows, while balancing resolution and penetration depth, particularly in deep tissue imaging. UltraUP applies to both 2D images, videos, and volumetric acquisitions increasing the field of view and frame rate, and supports device interoperability, reducing operator dependence and standardising output quality. UltraUP is designed for mobile and resource-limited contexts, including bedside diagnostics, field hospitals, and telemedicine. Finally, UltraUP enhances diagnostic confidence, improves reproducibility, and enables integration into different healthcare settings.

Keywords: Ultrasound imaging, denoising, super-resolution, artificial intelligence, real-time processing, software prototype

1 Introduction

Ultrasound (US) is widespread across medical specialities for its unique advantages over other modalities: it is non-invasive, cost-effective, portable, and capable of real-time acquisition. The recent diffusion of portable US devices has further expanded accessibility, supporting applications in remote patient monitoring and personalised medicine. In *point-of-care ultrasound* (POCUS), compact scanners allow physicians to perform bedside diagnostics, even in rural clinics or domestic environments, shifting US from a predominantly secondary-care tool to an essential component of primary care. POCUS is also integrated with *telemedicine*, extending high-quality imaging to underserved regions, including developing countries where trained specialists are often unavailable.

AI in Science
Conference 2025
EU2025DK

Despite its advantages, US imaging still faces important technical and clinical limitations. US images are affected by *speckle noise*, a granular artefact that reduces contrast, obscures anatomical boundaries, and interferes both with human interpretation and automated analysis (e.g., segmentation). Another major challenge is the limited *spatial and temporal resolution* of images, particularly with low-cost handheld probes or high-frame-rate acquisitions. Achieving real-time imaging while simultaneously applying advanced processing tasks such as denoising or super-resolution requires highly optimised algorithms and efficient hardware. Moreover, strict memory constraints are necessary to keep portable systems affordable and compact. These trade-offs between image quality, penetration depth, acquisition speed, and device cost constrain the performance of portable US imaging.

AI has greatly advanced US imaging in denoising and super-resolution. Self-supervised approaches (e.g., Noise2Noise [Lehtinen et al., 2018], Noise2Void [Krull et al., 2019], Noise2Self [Batson and Royer, 2019]) enable a good denoising quality, complemented by CNN-based methods integrating regularisation [Fang and Zeng, 2020], low-rank representations [Fu et al., 2021], wavelet CNNs [Wu et al., 2020], and hybrid models like BM-CNN [Ahn et al., 2018, Zhang et al., 2017]. For super-resolution, the field evolved from dictionary-based models [Peleg and Elad, 2014] to CNNs (e.g., SRCNN [Dong et al., 2014, 2015], EDSR [Lim et al., 2017]) and GAN-based architectures such as SRGAN [Ledig et al., 2017], ESR-GAN [Wang et al., 2019], and ESRGAN+ [Rakotonirina and Rasoanaivo, 2020], which improved perceptual fidelity and efficiency [Yu et al., 2020]. A broader overview of supervised, unsupervised, and domain-specific methods is provided in [Wang et al., 2020, Cammarana and Patanè, 2025].

To address the limitations of portable US systems, we introduce UltraUP (Sect. 2), a hardware-software prototype for real-time super-resolution and denoising of US images and videos. Designed for mobile and resource-limited scenarios (e.g., bedside diagnostics, field hospitals, telemedicine), UltraUP improves image quality, resolution, and acquisition frequency while preserving anatomical fidelity. Its modular and AI-based framework combines adaptive denoising, edge-preserving enhancement, and super-resolution to remove noise, reconstruct non-acquired scan lines, and preserve fine anatomical details. UltraUP is general with respect to the device type (e.g., ultra-handheld probes), anatomical district (e.g., cardiac, abdominal), and image properties (e.g., penetration depth, field of view, image resolution). Furthermore, UltraUP is compliant with real-time industrial and clinical requirements (Sect. 3).

2 UltraUP: real-time US super-res & denoising

UltraUP (Fig. 1) is a tool related to the field of biomedical image processing, specifically to a hardware/software prototype for the real-time super-resolution and denoising of US images and videos acquired by portable US systems. It addresses a critical limitation in existing imaging solutions: the low quality of US data produced by compact and mobile US devices due to physical, technical, and computational constraints. UltraUP has been developed to support diagnostic imaging in scenarios with limited infrastructure or mobile use cases, such as bedside diagnosis, field hospitals, and telemedicine. UltraUP addresses real-time denoising and super-resolution of US 2D/3D images and videos, which (i) take into account the high variability of US signals, (ii) are general for the underlying anatomical (e.g., muscle-skeletal, abdominal, cardiac) districts, and (iii) comply with industrial requirements in terms of computational cost and memory requirements (e.g., on commercial US machines). UltraUP can operate as a standalone or integrated tool, improving image quality, resolution, and acquisition frequency, while preserving anatomical features and assisting physicians in both diagnosis and treatment stages. It addresses the following fundamental challenges in US imaging.

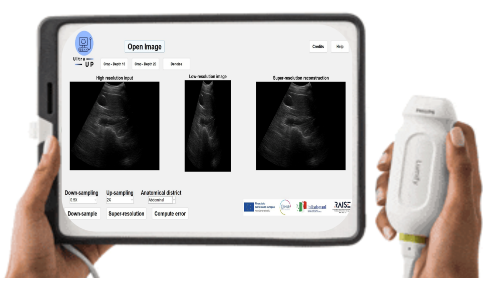

Figure 1: Our UltraUP prototype.

Concerning *denoising and enhancement*, effective suppression of noise while preserving fine anatomical structures and tissue texture is critical for both clinical interpretation and automated analysis. The Ultra-UP prototype employs adaptive algorithms to remove noise without introducing artefacts, ensuring edge preservation and the retention of diagnostically relevant details. This feature not only improves the visual quality perceived by the physician but also enhances the performance of downstream post-processing methods, such as segmentation and quantitative analysis. Concerning *super-resolution*, current US probes face a trade-off between acquisition frequency and spatial resolution, with higher frequencies often resulting in reduced resolution. The Ultra-UP prototype tackles this limitation by reconstructing non-acquired scan lines and increasing image resolution through super-resolution techniques. Applied to low-resolution, high-frequency US videos (e.g., cardiac imaging), this approach reconstructs high-resolution 2D sequences where each frame exhibits enhanced spatial detail. It allows the visualisation of fine anatomical and functional features that are otherwise inaccessible, thus overcoming one of the primary hardware constraints of conventional probes.

Through the combination of hardware and a modular, AI-based software library for denoising, enhancement, and super-resolution, UltraUP solves the following challenges. In terms of *deep tissue imaging and resolution*, US suffers from an inherent trade-off between penetration depth and resolution: high frequencies yield sharper images but are quickly attenuated, while low frequencies penetrate deeper but lose detail. These limitations are particularly critical in obese patients or when imaging deep organs. UltraUP addresses this challenge by combining advanced signal processing with super-resolution techniques to improve penetration without sacrificing image quality. In addition, integration with innovative transducer designs can further optimise the balance between depth and resolution. In terms of *real-time performance and extended field of view*, delivering high-quality images at clinical frame rates is essential for real-time US. However, the simultaneous demands of denoising, super-resolution, and image reconstruction often introduce latency. UltraUP employs efficient algorithms optimised for CPU and GPU execution, ensuring low-latency processing and real-time rendering. Furthermore, it extends beyond traditional 2D imaging to process 3D and 4D acquisitions, broadening the field of view and enabling more comprehensive diagnostic assessments. In terms of *standardisation, interoperability, and operator support*, US examinations are highly dependent on operator expertise, and variations across devices further reduce consistency. UltraUP is designed to be device-agnostic, supporting interoperability across systems and healthcare platforms. By guiding acquisition, assisting interpretation, and standardising output quality, it reduces operator dependency and improves reproducibility. Moreover, its modular framework can be adapted to specialised modalities, such as Doppler and contrast-enhanced US, to enhance sensitivity in blood flow and tissue characterisation.

3 Validation and further developments

UltraUP results are validated on the international US datasets *trackless 3D freehand ultrasound reconstruction* (TUS-REC) [Li et al., 2023] and *breast ultrasound images* [Al-Dhabyani et al., 2020] through pixel-wise, structural, and functional quantitative metrics (e.g., PSNR, SSIM, FSIM). UltraUP applies a real-time denoising and enhancement with an average accuracy of 97% with respect to state-of-the-art non-real-time methods (e.g., [Ahn et al., 2018]), and applies real-time super-resolution with an accuracy of 99% (2X up-sampling) and 95% (4X up-sampling) with respect to high-resolution target images. For further details on the experimental validation and comparison with state-of-the-art methods, we refer the reader to [Cammarasana and Patane, 2022] and [Cammarasana and Patanè, 2025]. UltraUP is available for proof-of-concept and can be applied through a user interface with US 2D and 3D images and 2D videos acquired by any US probe and from any anatomical district. As further developments, we are validating UltraUP in clinical and industrial contexts with different types of data (e.g., PET, MRI) and applications (e.g., focused US, spine segmentation). We refer the reader to the supplementary material for additional examples and demonstrations `https://drive.google.com/file/d/1XwqkDM_n_1JdpYr6xuRNyeexeFhspQIU/view?usp=sharing`.

Funding This work has been partially supported by the European Commission, NextGenerationEU, Missione 4 Componente 2, "*Dalla ricerca all'impresa*", Innovation Ecosystem RAISE "*Robotics and AI for Socio-economic Empowerment*", ECS00000035.

References

B. Ahn, Y. Kim, G. Y. Park, and N. I. Cho. Block-matching convolutional neural network (BMCNN): Improving cnn-based denoising by block-matched inputs. pages 516–525, 2018.

W. Al-Dhabyani, M. Gomaa, H. Khaled, and A. Fahmy. Dataset of breast ultrasound images. *Data in brief*, 28:104863, 2020.

J. Batson and L. Royer. Noise2self: Blind denoising by self-supervision. *arXiv:1901.11365*, 2019.

S. Cammarana and G. Patanè. *Real-Time Processing of Ultrasound Images: State-of-the-Art and Future Perspectives*, volume Synthesis Lectures on Visual Computing: Computer Graphics, Animation, Computational Photography and Imaging (SLVC-CGACPI). Springer Nature, 2025.

S. Cammarasana and G. Patane. Learning-based low-rank denoising. *Signal, Image and Video Processing*, pages 1–7, 2022.

S. Cammarasana and G. Patanè. Learning-based and quality preserving super-resolution of noisy images. *Multimedia Tools and Applications*, 84(9):6007–6023, 2025.

C. Dong, C. C. Loy, K. He, and X. Tang. Learning a deep convolutional network for image super-resolution. In *European Conference on Computer Vision*, pages 184–199. Springer, 2014.

C. Dong, C. C. Loy, K. He, and X. Tang. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2):295–307, 2015.

Y. Fang and T. Zeng. Learning deep edge prior for image denoising. *Computer Vision and Image Understanding*, 200:103044, 2020. ISSN 1077-3142.

X. Fu, S. Jia, L. Zhuang, M. Xu, J. Zhou, and Q. Li. Hyperspectral anomaly detection via deep plug-and-play denoising cnn regularization. *Transactions on Geoscience and Remote Sensing*, 59(11):9553–9568, 2021.

A. Krull, T.-O. Buchholz, and F. Jug. Noise2void-learning denoising from single noisy images. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pages 2129–2137, 2019.

C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4681–4690, 2017.

J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila. Noise2noise: Learning image restoration without clean data. *arXiv:1803.04189*, 2018.

Q. Li, Z. Shen, Q. Li, D. C. Barratt, T. Dowrick, M. J. Clarkson, T. Vercauteren, and Y. Hu. Long-term dependency for 3d reconstruction of freehand ultrasound without external tracker. *Transactions on Biomedical Engineering*, 71(3):1033–1042, 2023.

B. Lim, S. Son, H. Kim, S. Nah, and K. Mu Lee. Enhanced deep residual networks for single image super-resolution. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition Workshops*, pages 136–144, 2017.

T. Peleg and M. Elad. A statistical prediction model based on sparse representations for single image super-resolution. *Transactions on Image Processing*, 23(6):2569–2582, 2014.

N. C. Rakotonirina and A. Rasoanaivo. Esrgan+: Further improving enhanced super-resolution generative adversarial network. In *ICASSP 2020-2020 International Conference on Acoustics, Speech and Signal Processing*, pages 3637–3641. IEEE, 2020.

X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. C. Loy. Esrgan: Enhanced super-resolution generative adversarial networks. In L. Leal-Taixé and S. Roth, editors, *Computer Vision – ECCV 2018 Workshops*, pages 63–79, Cham, 2019. Springer International Publishing. ISBN 978-3-030-11021-5.

Z. Wang, J. Chen, and S. C. Hoi. Deep learning for image super-resolution: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(10): 3365–3387, 2020.

T. Wu, W. Li, S. Jia, Y. Dong, and T. Zeng. Deep multi-level wavelet-cnn denoiser prior for restoring blurred image with cauchy noise. *Signal Processing Letters*, 27: 1635–1639, 2020.

J. Yu, Y. Fan, and T. Huang. Wide activation for efficient image and video super-resolution. In *30th British Machine Vision Conference, BMVC 2019*, 2020.

K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *Trans. on Image Processing*, 26 (7):3142–3155, 2017.