ORIGINAL ARTICLE

Real-time denoising of ultrasound images based on deep learning

Simone Cammarasana[1] · Paolo Nicolardi[2] · Giuseppe Patanè[1]

Abstract

Ultrasound images are widespread in medical diagnosis for muscle-skeletal, cardiac, and obstetrical diseases, due to the efficiency and non-invasiveness of the acquisition methodology. However, ultrasound acquisition introduces noise in the signal, which corrupts the resulting image and affects further processing steps, e.g. segmentation and quantitative analysis. We define a novel deep learning framework for the real-time denoising of ultrasound images. Firstly, we compare state-of-the-art methods for denoising (e.g. spectral, low-rank methods) and select WNNM (*Weighted Nuclear Norm Minimisation*) as the best denoising in terms of accuracy, preservation of anatomical features, and edge enhancement. Then, we propose a tuned version of WNNM (*tuned-WNNM*) that improves the quality of the denoised images and extends its applicability to ultrasound images. Through a deep learning framework, the tuned-WNNM qualitatively and quantitatively replicates WNNM results in real-time. Finally, our approach is general in terms of its building blocks and parameters of the deep learning and high-performance computing framework; in fact, we can select different denoising algorithms and deep learning architectures.

Keywords Image denoising · Deep learning · Real-time denoising · Biomedical data · Ultrasound images

1 Introduction

Ultrasound imaging uses high-frequency sound waves to visualise soft tissues, such as internal organs, and support medical diagnosis for muscle-skeletal, cardiac, and obstetrical diseases, due to the efficiency and non-invasiveness of the acquisition methodology. Ultrasonic sound waves are reflected off from different layers of body tissues. The main issues of the ultrasound techniques are a significant loss of information during the reconstruction of the signal, the dependency of the signal from the direction of acquisition, an underlying noise that corrupts the image and significantly affects the evaluation of the morphology of the anatomical district. In this context, the denoising of ultrasound images is relevant both for post-processing and visual evaluation by the physician.
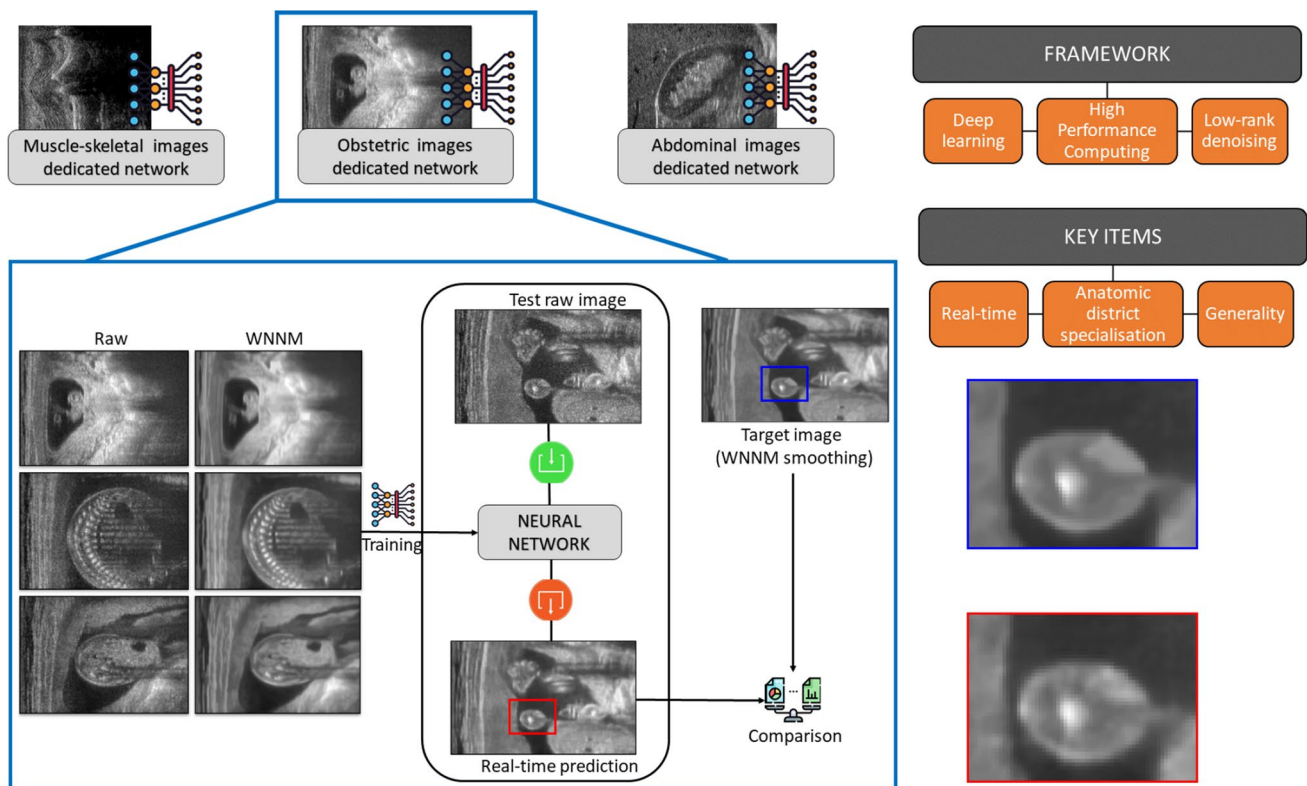
✉ Simone Cammarasana
simone.cammarasana@ge.imati.cnr.it

Paolo Nicolardi
paolo.nicolardi@esaote.com

Giuseppe Patanè
patane@ge.imati.cnr.it

1   CNR-IMATI, Via De Marini 6, Genoa, Italy

2   Esaote S.p.A., Via E. Melen 7, Genoa, Italy

Our main goal is the definition of a novel deep learning framework for the real-time denoising of ultrasound images (Fig. 1). After the design of a training data set, composed of raw images and the corresponding denoised images, we train a neural network that replicates the denoising results. Then, the real-time denoising is achieved through the prediction of the trained network. The proposed framework combines three elements: *low-rank denoising*, *deep learning*, and *high-performance computing* (HPC, for short).

We select WNNM-*Weighted Nuclear Norm Minimisation* [27] as the best denoising method, which is then specialised to ultrasound images as a "new" *tuned-WNNM* denoising, by tuning its parameters. The choice of WNNM is based on a qualitative and quantitative comparison of five denoising methods, i.e. WNNM, SAR-BM3D - *SAR Block-Matching 3D* [48], BM-CNN - *Block Matching Convolutional Neural Network* [5], NCSR - *Non-Locally Centralised Sparse Representation* method [19], PCA-BM3D *Principal Component Analysis Block Matching 3D* [16]) belonging to the spectral, low-rank, and deep learning classes (Section 2).

To achieve a *real-time denoising of ultrasound images*, we propose a deep learning framework that is based on the learning of the tuned-WNNM and HPC tools (Section 3). The training is performed offline and can be further improved with new data, a priori information on the

Fig. 1 Pipeline of the proposed framework, with a magnification of prediction (red) and target (blue) images (right side). Our framework applies deep learning and HPC to learn and replicate the denoising results of state-of-the-art low-rank method (i.e. tuned-WNNM), in real-time and with a specialisation to anatomic districts

input images or the anatomical district, and denoised images selected after experts' validation. Through our framework, the execution time of the denoising only depends on the network prediction, which is achieved in real-time on standard medical hardware.

As the main contribution, the proposed denoising of ultrasound images runs in real-time and is general in terms of the input data, in terms of resolution of the input images (e.g. isotropic, anisotropic images), acquisition methodology, anatomical district, noise (e.g. speckle or Gaussian noise), and the dimension of the images (i.e. 2D, 3D images). Our approach is also general in terms of the building blocks and parameters of the deep learning framework; in fact, we can select different denoising algorithms (e.g. WNNM, SARBM3D) and deep learning architectures (e.g. Pix2Pix, VGG19).

As experimental validation (Section 4), we perform a quantitative (e.g. PSNR, SSIM) and qualitative evaluation of the selected denoising methods on ultrasound images acquired from different anatomical (e.g. muscle-skeletal, obstetric, and abdominal) districts. Then, the results of the deep learning and HPC frameworks are quantitatively and qualitatively analysed on a large collection of ultrasound images. The industrial requirement of real-time denoising

is verified in terms of the execution time of the network prediction on GPU-based hardware installed in commercial ultrasound machines. Finally, we discuss main results (Section 5), conclusions, and future work (Section 6).

2 Related work

We review previous work on image denoising (Section 2.1) and deep learning methods for denoising and regression (Section 2.2).

2.1 Image denoising

Non-local methods The *Non-Local Means* (NLM) denoising [11] uses the patterns' redundancy of the input image, and each patch is restored with a weighted average of all the other patches, where each weight is proportional to the similarity among the patches. The *Bayesian non-local mean filter* [30] improves the NLM with the introduction of a Bayesian estimator as distance measure among the patches, which allows the user to better determine the amount of denoising by the noise variance of the patch. The anisotropic neighbourhood in NLM [44] uses image gradient to estimate the

edge orientation and then adapts the patches to match the local edges. The characterisation of the patches through a redundancy index [45] improves the self-similarity computation among patches. The improvement on the structure of the search window is achieved through the computation of an optimal search window for each pixel [63], according to the denoising degree of the related patch.

Anisotropic methods　The denoised image is computed as the solution to an anisotropic diffusion equation [49, 50], where the gradient of the image guides the diffusion process. The variant [72] exploits the Lee [35] and Forst [24] filters, which are edge-sensitive to speckle noise. An improvement of the previous results [6] is achieved by applying the Kuan filter [34] in the diffusion equation and revising the selection of the neighbourhood used for the estimation of the statistical parameters. The anisotropic method introduces a class of fractional-order anisotropic diffusion equations [7], using the Fourier transform to compute the fractional derivatives, and the discrete Fourier transform to compute the fractional-order differences.

Spectral denoising　Denoising based on *spectral decomposition* transforms a signal into its spectral domain and exploits the sparsity of the transformed signal to remove noise, through a threshold operation. Several transformations have been applied to image denoising, such as Wavelets [13, 39, 46, 51], Curvelets [61], Contourlets [14], and Shearlets [71]. To reconstruct the denoised image, the *3D block-matching* [15] computes and stacks similar patches through NLM; each stack is transformed into its spectral domain with wavelet decomposition, denoised through a hard/soft threshold, and reconstructed in the space domain. The denoised patches are aggregated by a collaborative filter. The *synthetic aperture radar block matching 3D* (SAR-BM3D) [48] introduces a speckle-based variant of 3D block matching; the similarity among the patches is computed by considering the probability distribution of the speckle noise as a distance metric. Furthermore, the hard/soft threshold of the wavelet transformed signal is substituted by a Local Linear Minimum Mean Square Error (LLMMSE) filter. The *principal component analysis block matching 3D* (PCA-BM3D) [16] improves the stacking operation of 3D block-matching by using shape-adaptive neighbourhoods, which enable its local adaptability to image features. The 3D transformation of each stack to the spectral domain is performed through the PCA [66] and an orthogonal 1D transformation in the third dimension.

Low-rank methods　Low-rank approximation computes the denoised image as the solution to a weighted minimisation problem, whose cost function is the Frobenius norm [12, 49, 60] or the ℓ_1 norm [20], between the input and the target

images. The relation between local and non-local information [18] allows us to estimate signal variances, by interpreting the *Singular Value Decomposition* (SVD, for short) through a bilateral variance estimation. In [54], a high-order SVD is applied to 3D blocks, and the denoised image is achieved with hard thresholding of the decomposed signal. The *Weighted Nuclear Norm Minimisation* (WNNM) [27] computes the stacks as in the 3D block-matching method, performs the SVD on the stacks and applies a weighted threshold to the singular values, where higher weights correspond to lower singular values, which capture the noisy component of the image. The *collaborative filtering* of WNNM for the aggregation of the denoised patches is performed as in the 3D block-matching method. The weighted nuclear norm and the histogram preservation [74] are combined in a single constrained optimisation problem, which is solved through the alternating direction method of multipliers [10]. The WNNM is extended to image deblurring with several types of noise [40].

External learning　The K-SVD algorithm [4] represents the signal as a linear combination of an over-complete dictionary of atoms, which are iteratively updated through the SVD of the representation error to better fit the data. A *learned simultaneous sparse coding* method [43] integrates sparse dictionary learning with non-local self-similarities of natural images. The *non-locally centralised sparse representation* (NCSR) [19] exploits the non-local redundancies, combined with local sparsity properties, to estimate the coefficients of the sparse representation of the input image. The dictionary is learned by clustering the patches of the image into K clusters through the K-means [42] method and then learning a PCA sub-dictionary for each cluster. This method has been further improved in [69] with a fast version based on a pre-learned dictionary and achieving an improvement of the computational efficiency. The structured sparse model selection over a family of learned orthogonal bases [41] is applied to the deblurring of images with Gaussian noise.

2.2 Deep learning for denoising and regression

Deep learning methods for denoising　In the *Noise2Noise* algorithm [36], the network learns to denoise images only considering the noisy data, without any knowledge of the ground truth. The *Noise2Void* algorithm [33] further expands this idea, and it does not require couples of noisy images for the training. This approach is relevant in biomedical fields, where there are no ground truth images. The *Noise2Self* method [8] proposes a self-supervised algorithm that does not require any prior information on the input image, estimation on the noise, or ground truth data. The denoising of images [21] is achieved through the extraction of features from the noisy image through a convolutional

neural network (CNN), and combining the edge regularisation with the total variation regularisation. The combination of CNN and low-rank representation [25] is applied to detect anomalous pixels in hyperspectral images. The multilevel wavelet convolutional neural network [67] is applied for restoring blurred images affected by Cauchy noise. The *block matching Convolutional Neural Network* (BM-CNN) [5] integrates a deep learning approach with the 3D block-matching method; the denoising of the stacks is predicted through a DnCNN [73], which is trained with a data set of 400 images corresponding to more than 250K training samples. A feed-forward Convolutional Neural Network smooths images, independently from the noise level, by exploiting residual learning and batch normalisation. Then, the blocks are aggregated and the image is reconstructed, as in the 3D block-matching algorithm.

Deep learning methods for image-to-image regression The *VGG19* [56] introduces *Convolutional Neural Networks* (CNN) pushing the depth to 16–19 weight layers, using small convolution filters of 3×3 size, with an application to large scale images classification. The *Pix2Pix* [29] method is a *Generative Adversarial Network* (GAN), where the generator is a U-net [55], the discriminator is an encoding network [32], and the loss function is based on the binary cross-entropy. The *deep convolutional generative adversarial network* [53] applies unsupervised learning for image classification and the generation of natural images, exploiting batch normalisation, rectified linear unit activations, and removing fully connected hidden layers.

3 Method

We propose a novel deep learning framework for the real-time denoising of ultrasound images. Firstly, we introduce the data sets and metrics for the selection of the denoising method that best fits our requirements for ultrasound images (Section 3.1). Then, we optimise the parameters of the selected method to the denoising of ultrasound images (Section 3.2). Finally, we introduce a deep learning (Section 3.3) and HPC (Section 3.4) framework, which achieves real-time denoising. For a detailed discussion on the results, we refer the reader to Section 4.

3.1 Data sets and metrics for denoising evaluation

We compare five denoising state-of-the-art methods, which are either specific for speckle noise (e.g. SAR-BM3D) or independent of the type of noise (e.g. WNNM). We consider the *Esaote private data set*, which contains more than 10K ultrasound images at different resolutions, and is acquired from different (e.g. muscle-skeletal, obstetric, abdominal)

anatomical districts. On this data set, we compare the performance of different denoising methods applied to ultrasound images, and analyse the performance of the proposed framework, through the training and the prediction of the learning-based network, with a specialisation to anatomic districts.

As *quantitative metrics*, we consider the *peak-signal-to-noise ratio* (PSNR) and the *structural similarity index measure* (SSIM) for the comparison of the raw input with the target denoised image provided by the proposed framework. Furthermore, we integrate the quantitative metrics with a qualitative discussion on the quality of the denoised images, in terms of blurring and edge preservation.

3.2 Tuned-WNNM

According to the results in Section 4, WNNM has been selected as the best denoising method among five state-of-the-art methods belonging to the spectral, low-rank, and deep learning classes. To improve the quality of the denoised image, we propose a novel approach to the tuning of WNNM parameters, and we refer to this method as *tuned-WNNM*.

Given a pixel x, the *patch P_x* is the set of pixels in the neighbourhood of x; each pixel of the image has a related patch. The *search window* is the set of patches selected for searching the closest ones to a reference patch, under a certain metric. The *stack*, or *3D block*, is the set of patches that are similar to a reference patch; these patches are stored in a 3D structure, and the redundancy of the stack is exploited to remove the noise. Within this setting, the tuning of these parameters (i.e. search window, stack, patch size) improves the results of tuned-WNNM with respect to WNNM. Our framework maximises the performance of the denoising method; in fact, the real-time requirement is achieved by the network prediction, while the denoising is applied offline for the generation of the training data set.

3.3 Real-time denoising with deep learning

The main requirements of a denoising algorithm for ultrasound images are the magnitude of the removed noise, edge preservation, and real-time computation. The tuned-WNNM satisfies these requirements, except for the execution time, which does not satisfy the real-time need of ultrasound applications. To achieve a real-time computation and to maintain the good results of the tuned-WNNM method in terms of denoising and edges preservation, we identify two strategies. We develop (i) a computationally optimised version of the tuned-WNNM method, exploiting HPC tools, CPUs and GPUs, and low-level programming languages. We design and implement (ii) a deep learning framework that uses the tuned-WNNM as an instance of denoising methods.

The implementation of a computationally optimised, and potentially real-time, version of the tuned-WNNM

is a very tough requirement; the main iterative cycle of the algorithm is not parallelisable, due to the dependency of the data among the iterations. Furthermore, the cubic computational cost for the evaluation of the SVD of each block is no further optimisable. The real-time requirement needs a specific hardware-based implementation, and any modification to the method requires a new implementation of the parallel algorithm. This approach needs dedicated and more expensive hardware, which is in contrast with the cheapness of the ultrasound acquisition.

Proposed approach and contributions The proposed real-time denoising is based on the training of a neural network to learn and replicate the tuned-WNNM behaviour. In the first phase, the network is trained on a data set of ultrasound images of the same district. The data set for the training of the learning method is composed of a set of couples of ultrasound images: the input (i.e. the raw image) and the target (i.e. the image denoised with the tuned-WNNM filter). The ground truth is not available in ultrasound applications; for this reason, the target of the learning method is the output of the tuned-WNNM filter. Then, the trained network provides the denoised output through a real-time prediction of the test images. As the main contribution, the proposed deep learning framework is general in terms of the input data, i.e. type of noise (e.g. speckle, Gaussian noise), resolution (e.g. isotropic, anisotropic) of the input images, acquisition methodology, and the anatomical district. Our deep learning framework is also general in terms of building blocks and parameters: since different methods (e.g. NCSR, SAR-BM3D, custom methods) have good performances, the generality of our framework allows us to use a different denoising algorithm and to exploit its different characteristics in terms of denoising and edge preservation. Alternative denoising methods can be used for different types of noise (e.g. speckle, Gaussian noise) and signals (e.g. 3D images, time-dependent ultrasound videos).

In our approach, we specialise the training phase to specific anatomical districts or types of noise. For instance, we train a specific network for each district, thus obtaining a more precise result when predicting the denoised image, as each network is specialised to the input anatomical features. We also improve the WNNM denoising in terms of real-time computation based on offline training. The real-time computation depends only on the execution time of the network prediction. Furthermore, the offline training can be improved with new data, a priori and/or additional information on the data (e.g. input anatomical district, noisy type/intensity, image resolution, acquisition methodology/protocol). The update of the existing training data set is always performed offline, through the addition of new images after expert validation of the denoising results.

Deep learning architecture To evaluate the proposed framework, we analyse several networks and perform an image-to-image regression; among them (Section 2), we select Pix-2Pix, which guarantees good results in terms of learning. We specialise Pix2Pix to ultrasound images, with two updates: (i) the introduction of a validation data set of the same district of the training data set, which forces the exit condition when the validation error increases, and (ii) the introduction of padding and masking pre-processing operations, which allow us to deal with images of different resolution, without an image resize that would imply a distortion artefact. A comparison between Pix2Pix and *Matlab* CNN is discussed in Section 4.5.

Training data sets We generate and test different data sets, by varying the number of images for the training phase, and the anatomical district for the prediction phase. The custom Pix2Pix network is trained on four data sets of obstetric images, with respectively: (a) 500, (b) 1500, (c) 3500, and (d) 5000 images. Each data set is composed of the input images (i.e. the raw Esaote data set) and the target images (i.e. the corresponding images denoised with the tuned-WNNM). The validation data set is composed of an additional set of different images (i.e. about 10% of the training data set) of the same district. Then, we evaluate each of the four networks (i.e. the networks trained with a different number of images) with two different test data sets of 50 images each, respectively from the (i) obstetric and (ii) muscular anatomical districts. For each test data set, we compute the quantitative PSNR and SSIM metrics between the prediction of the network and the expected target; furthermore, the experts visually evaluate the prediction results.

3.4 HPC framework for learning

We define an HPC implementation of the proposed deep learning framework, taking advantage of a large ultrasound data set with 5K ultrasound images, and of the CINECA-Marconi100 cluster, exploiting both CPUs (IBM POWER9 AC922) and GPUs (NVIDIA Volta V100). Given a training data set, composed of raw images and the corresponding denoised images, we implement a parallel and distributed deep learning framework in TensorFlow2. Then, we define a batch file for the execution of the deep learning framework on the cluster, which specifies the number of nodes, CPUs, GPUs, and memory of the cluster. Through the proposed HPC framework, we train multiple networks with large data sets in a reasonable time for the target medical application, thus increasing the specialisation to anatomical districts, and consequently the accuracy of the deep learning framework. The HPC framework generates a network model that is stored and used for predicting the output results. Furthermore, we can improve the offline training with new data, a

priori and/or additional information on the input data (e.g. input anatomical district, noisy type/intensity, image resolution, acquisition methodology/protocol). The training data set can be periodically updated with the denoised images after the expert validation of the denoising results.

4 Results

We present the results of denoising methods with a specialisation to ultrasound images (Section 4.1), a comparison between baseline and tuned-WNNM (Section 4.2), deep learning (Section 4.3) and HPC (Section 4.4) framework. Finally, we compare the real-time denoising with different neural network architectures (Section 4.5).

4.1 Comparison of denoising methods

We evaluate the denoising results on different anatomical districts of the Esaote data set (Figs., 2, 3, 4): WNNM, NCSR, and PCA-BM3D have been judged as the best methods in terms of denoising, and WNNM outperforms all the other methods in terms of edge preservation and enhancement. In particular, WNNM well preserves the edges of the muscular fibres (Fig. 2) and the internal organs (Figs. 3, 4). The output of SAR-BM3D shows a granular effect, which negatively affects the preservation of the anatomical features, and BM-CNN generates artefacts, which are typical of a deep learning approach. According to these results, we select WNNM as the best method for the denoising of ultrasound images. However, we underline that the other methods have their characteristics in terms of denoising and edge preservation, and they could be included in the framework as alternative denoising algorithms.

Execution time Our tests (Table 1) are executed with Matlab R2020a, on a workstation with 2 Intel i9-9900KF CPUs (3.60GHz), 32 GB RAM, and none of these methods achieves real-time computation. In particular, WNNM takes more than three minutes to process a 600×485 image, and the fastest method (i.e. SAR-BM3D) takes about one minute; however, real-time computation in an ultrasound environment requires a processing time in the order of a few milliseconds. This result motivates the proposed development of a deep learning framework for the real-time denoising of ultrasound images, further optimised with a HPC framework (Section 3.4).

4.2 Tuned-WNNM for US images

We implement the tuned-WNNM through the optimisation of the following parameters. The number of patches is no

**Table 1** Execution time computed as an average value of a set of Esaote images at 600×485 resolution

Method	WNNM	SAR-BM3D	BM-CNN	NCSR	PCA-BM3D
Execution time [s]	215	**55**	356	380	95
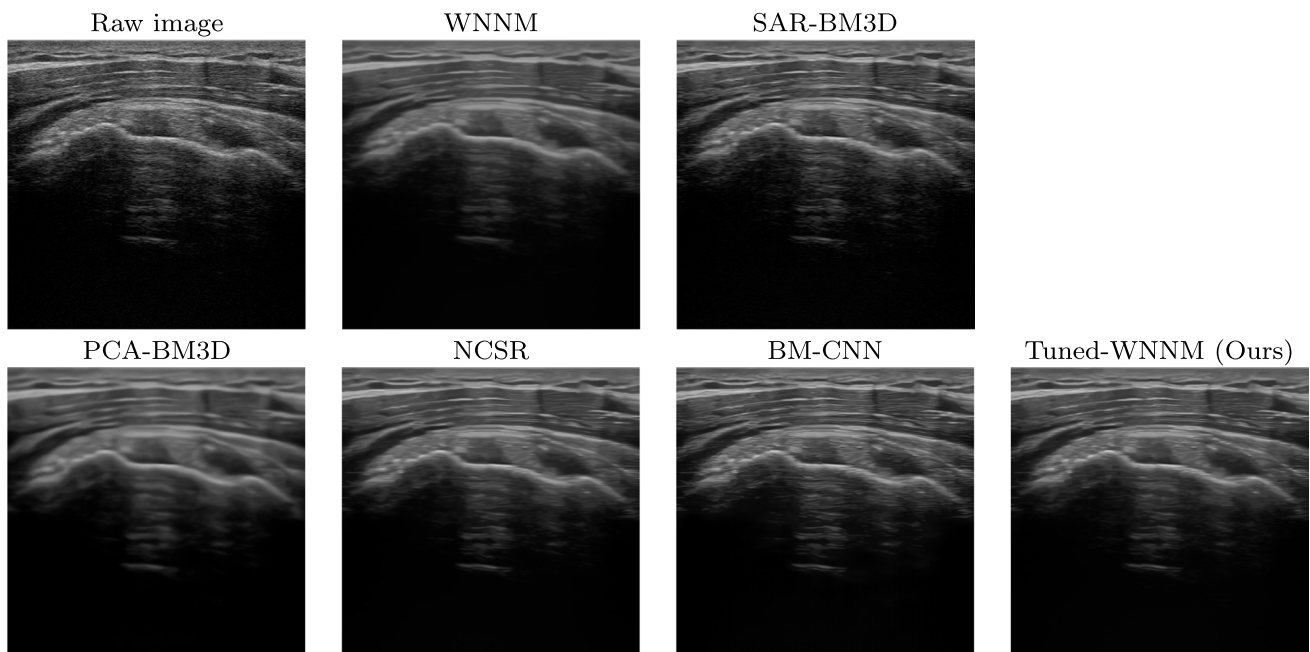
Fig. 2 Raw images of a muscle-skeletal district and denoised images, visualised in a scan-converted format
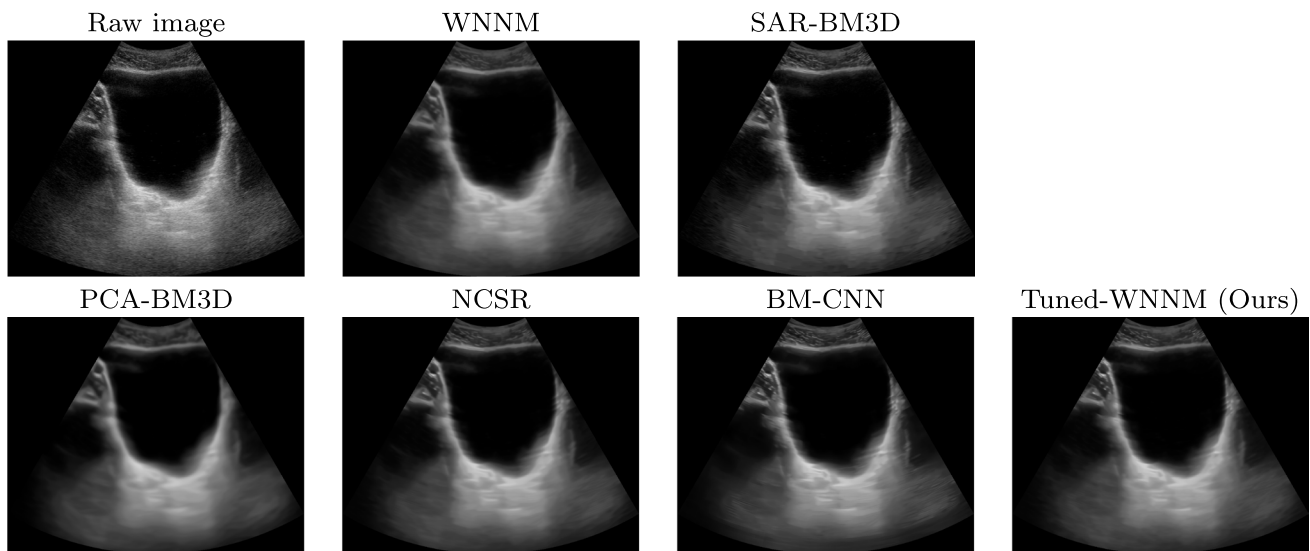
Fig. 3 Raw data set of an abdominal district and denoised images, visualised in a scan-converted format
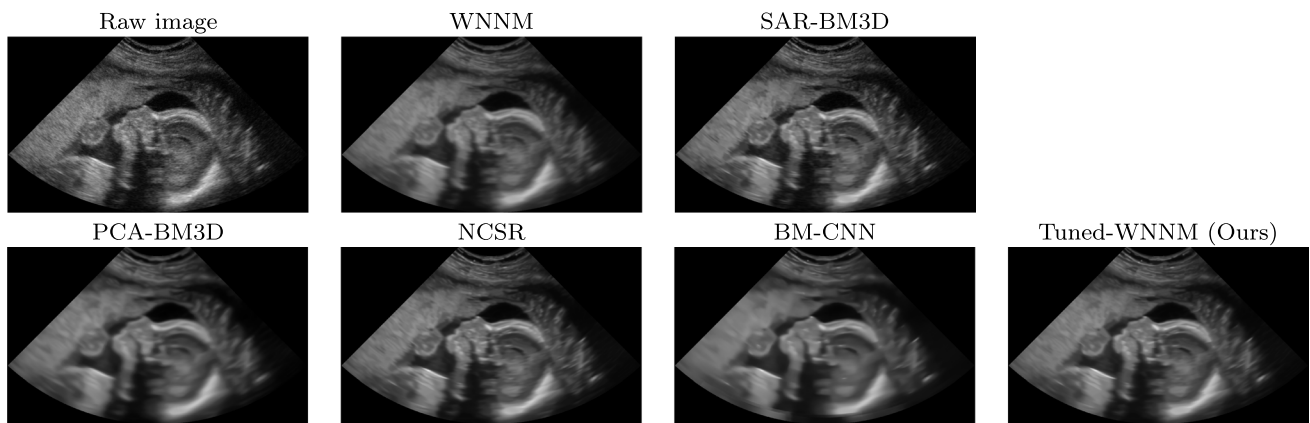
Fig. 4 Raw data set of an obstetric district and denoised images, visualised in a scan-converted format

more limited by the step value (e.g. 1 patch every 2 or 3 pixels) and we assign a patch for each pixel; this parameter allows us to increase the number of processed patches, thus improving the data redundancy. The block-matching algorithm is now performed every iteration, instead of one every two iterations; the selection of the searching window and the size of the stack are now larger than previous work. These parameters allow us to improve the measure of the similarity among 3D blocks and the accuracy of the denoising method.

Furthermore, we specialise the tuned-WNNM method to ultrasound images, by varying the denoising intensity through a parameter that affects the threshold of the singular values of the SVD. Increasing this parameter, the method improves in terms of removed noise, though introducing a low blurring effect. To select the best tuning for denoising intensity, we select the output image that best fits the medical requirements, among three different levels of denoising

intensity (Fig. 5). In particular, Fig. 5b shows the best result as a compromise between noise removal, edges preservation, and blurring effect; in fact, it preserves the geometry of the internal tissues, while enhancing the edges of the anatomical structures.

4.3 Learning-based denoising

Qualitative results Regarding the deep learning framework, and the large ultrasound data set (Section 3.3), Fig. 6 shows the prediction results of the four networks, when tested with obstetric images (i). The predicted images are very close to the target image in all four cases; the edges and the greyscale values are well reproduced by the network. Furthermore, the predictions do not generate artefacts or spurious patterns. Varying the number of images of the training data set from 500 to 5K (Fig. 6a–d), the predicted images are

Input image   Tuned-WNNM: (a)   Tuned-WNNM: (b)   Tuned-WNNM: (c)
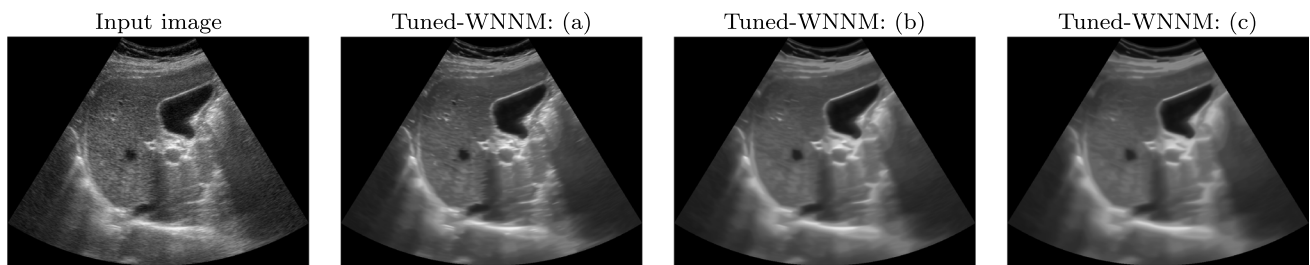
Fig. 5 Ultrasound image of an abdominal district and denoised images achieved by applying the tuned-WNNM and varying the denoising intensity from low (**a**) to high (**c**)

Raw image   Prediction:(a)   Prediction: (b)

Target   Prediction: (c)   Prediction: (d)
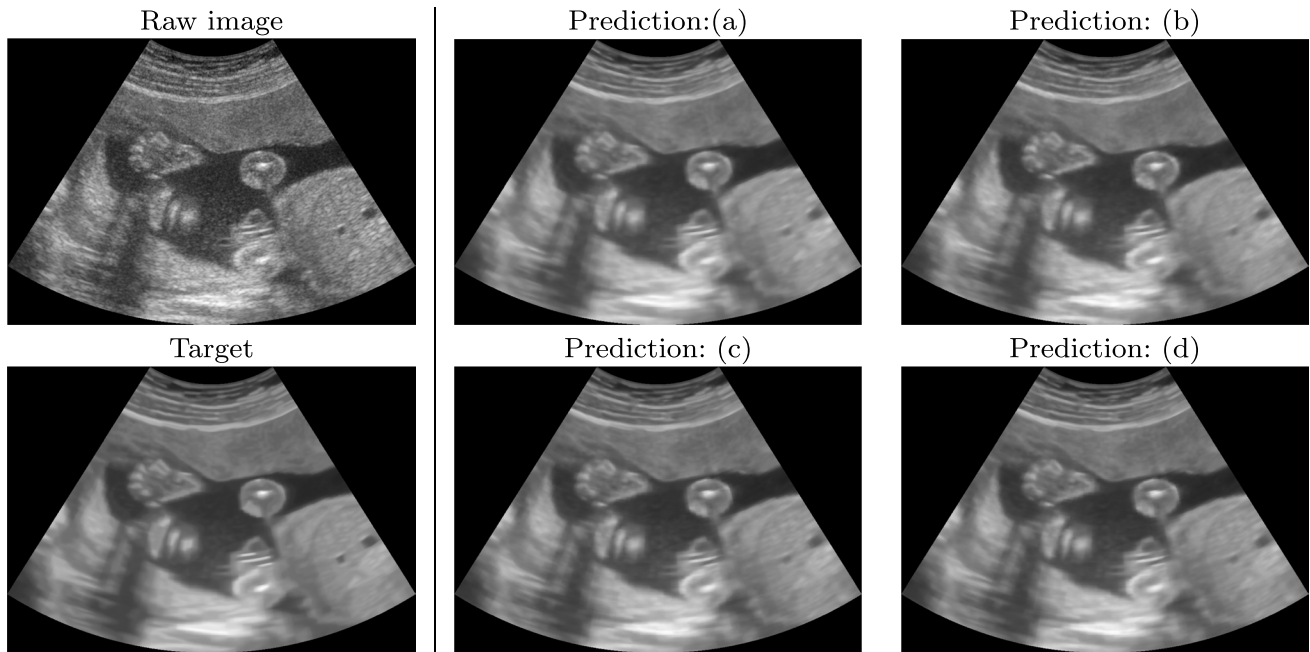
Fig. 6 Raw, target, and prediction images, related to the obstetric data set (i). Training set: (**a**) 500 images, (**b**) 1500 images, (**c**) 3500 images, (**d**) 5000 images (Section 3.3)

slightly better than the target denoised images. Nevertheless, the results are good even with a small training data set of 500 images. Figure 7 shows the prediction results of the four networks when tested with muscle-skeletal images (ii). Predicting the output images with the networks trained with obstetric images (Fig. 7a–d), the results are slightly worse than the corresponding case in Fig. 6, even if the predicted images do not show any artefact of pattern repetition. These networks are trained with images from a different (i.e. obstetric) district, with different anatomical features. This result confirms that each district requires a specific network and that a single network for all the districts gives lower quality results.

Quantitative results Table 2 reports the quantitative metrics (Section 3.2) computed between the target and the predicted images. The network trained with 5K images (d) tested with

obstetric images (i) has a median PSNR and SSIM value of 36.13 and 0.964, respectively, while the same network tested with muscle-skeletal images (ii) has a median PSNR and SSIM value of 26.58 and 0.881. Both the metrics have a very slight improvement when passing from a training set of 500 to a training set of 3500 images, confirming the results of the qualitative analysis. An additional increase of the size of the training data set to 5K images further improves the quantitative results for both the test data sets. Figure 8 shows the box plot of the PSNR and SSIM metrics for four training data sets and two test data sets. Increasing the number of images of the training data set, the range of the metrics tends to decrease; this behaviour has a lower variability on the prediction of the output image. These results confirm that a network specialised in a single anatomical district reaches the best denoising quality. The prediction of muscle-skeletal images from a network trained with obstetric images highly

Fig. 7 Raw, target, and prediction images, related to the muscle-skeletal data set (ii). Training set (Section 3.3) with (**a**) 500, (**b**) 1500, (**c**) 3500, and (**d**) 5000 images
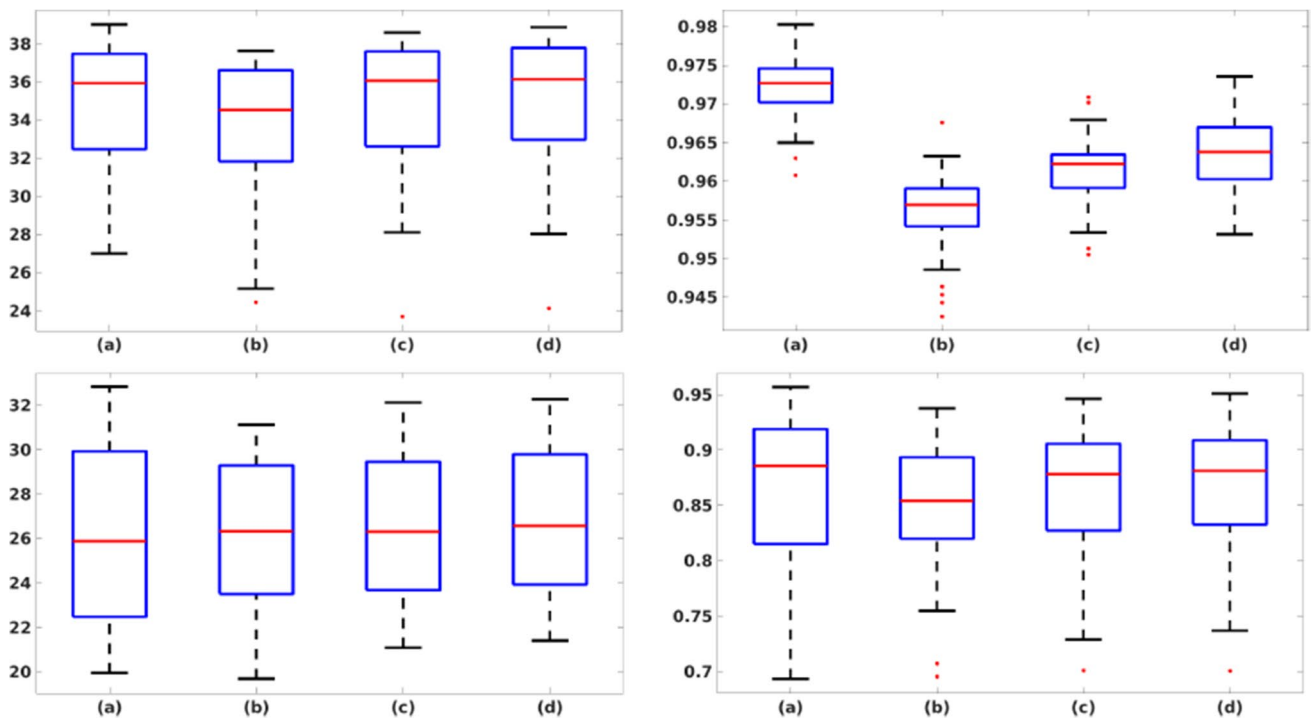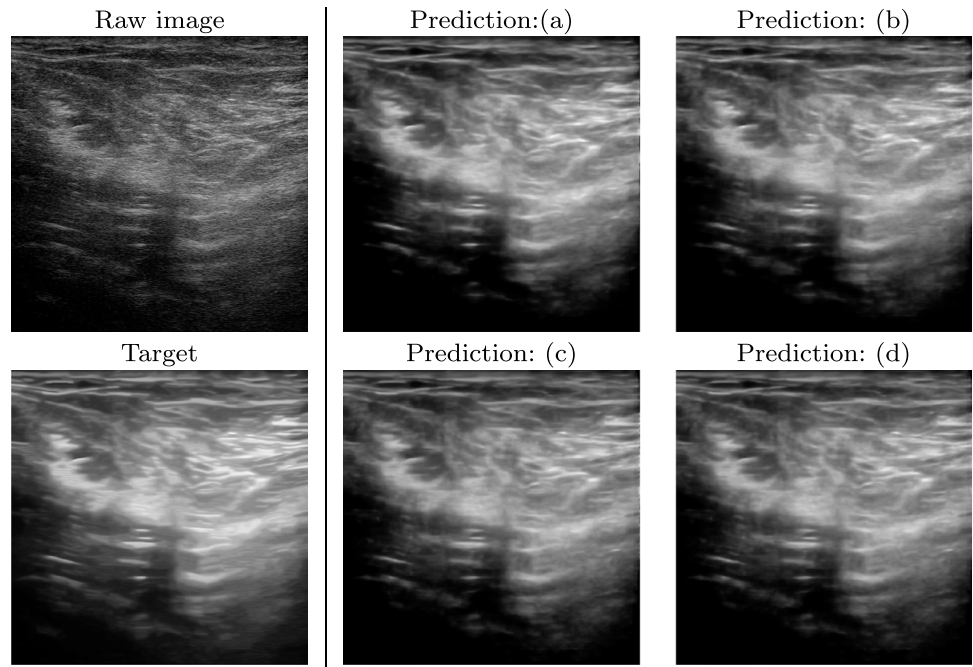
Fig. 8 PSNR and SSIM boxplot for each of the four training data set (i.e. (**a**–**d**)) and the two test data sets (Section 3.3): (top-left) PSNR, obstetric test data set; (top-right) SSIM, obstetric test data set; (bottom-left) PSNR, muscle-skeletal test data set; (bottom-right) SSIM, muscle-skeletal test data set

reduces the performance of our framework; in fact, the network learns that replicates not only the denoising algorithm itself but also its adaptation to the anatomic structures and features of each district.

Single versus multiple districts We compare the prediction results of our framework with three different training data sets. The first two data sets have 500 and 1500 ultrasound images of the same district (i.e. the obstetric one),

Table 2 With reference to the four training data sets and the two test data sets (i.e. obstetric: *Ob.*, and muscle-skeletal: *Msk.*) described in Section 3.3, we report the PSNR and SSIM metrics computed between the target and the prediction images, as median value among the 50 test images

Metrics	PSNR		SSIM	
Test data set	Ob.	Msk.	Ob.	Msk.
Training with				
(a) 500 images	35.93	25.88	0.973	0.886
(b) 1500 images	34.52	26.33	0.957	0.854
(c) 3500 images	36.07	26.31	0.962	0.878
(d) 5000 images	36.13	26.58	0.964	0.881

respectively. The third one is composed of 1500 images of different districts; in particular, we select 500 images from the cardiac, obstetric, and muscle-skeletal districts. Due to the different resolutions of the images, the padding has been applied to obtain the same input resolution for each network. We evaluate the prediction results on four test data sets: the first three are composed of 50 images from the cardiac, obstetric, and muscle-skeletal districts, respectively. The fourth is composed of 50 images randomly selected from the aforementioned three districts. The prediction results (Fig. 9 and Table 3) show that the networks trained with obstetric images (i.e. the single district networks) give the best results with the obstetric test data set: the predicted image of the single district network shows fewer scattering artefacts than the multiple districts network. Also, the single district networks have better results in terms of quantitative metrics: adding further images from different districts to the training data set worsens the results; in fact, the single district network with 500 obstetric images has a PSNR value

of 35.93, while the multiple districts network has a PSNR value of 33.70.

Comparing the results on the other test data sets (e.g. cardiac Fig. 10 and Table 3), the network trained with images of multiple districts has better results than the networks trained with obstetric images only. The multiple district network better generalises on the denoising algorithm, more than on the anatomic features, thus generating fewer artefacts on the prediction. Furthermore, the multiple district network includes 500 cardiac images in its training data set, thus improving the prediction results on this district. As the main conclusion, a dedicated network for each anatomic district is the best solution for the prediction of the denoised ultrasound images of each specific district, if a sufficiently large data set is available for the training.

## 4.4 Execution time and computational cost

To test the training phase of the deep learning framework (Section 3.3) on the HPC framework (Section 3.4), we exploit 8 nodes, each one composed of 32 cores and 4 accelerators, for a theoretical computational performance of 260 TFLOPS, and 220 GB of memory per node. The parallel implementation of the deep learning framework and the high hardware performance reduce the computation time of the training phase by at least 100 orders less than a serial implementation on a standard workstation.

The execution time of the prediction is crucial for the real-time implementation of our framework. We test the denoising prediction on GPU-based hardware, which replicates the hardware of an ultrasound scanner currently in use. Given a set of ultrasound input images from different districts, the average execution time is 25 milliseconds; this

Fig. 9 Prediction results of the obstetric district, with the networks trained with 500 (**a**) and 1500 (**b**) images from the obstetric district, and 1500 images from multiple districts (500 obstetric, 500 cardiac, 500 muscle-skeletal images)
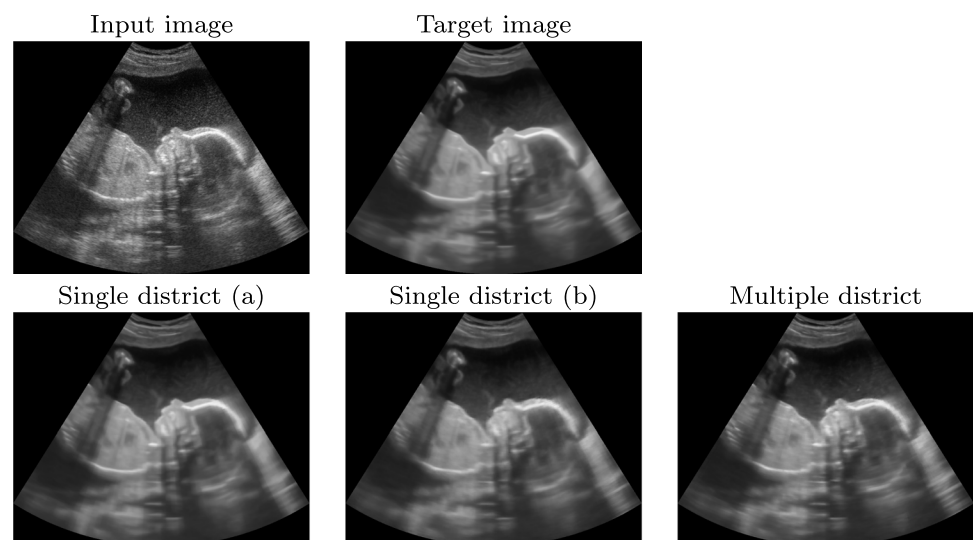
Fig. 10 Prediction results of the cardiac district, with the networks trained with 500 (**a**) and 1500 (**b**) images from the obstetric district, and 1500 images from multiple districts
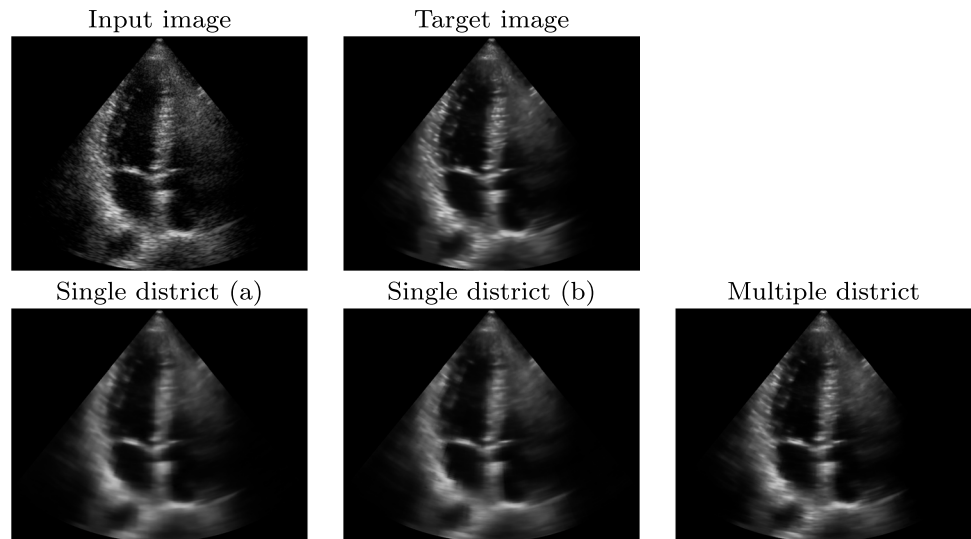
Table 3 With reference to the results in Figs. 9 and 10, we report the PSNR metric computed between the target and the prediction images, as average value among the 50 images of each test data set: obstetric (Ob.), muscle-skeletal (Msk.), cardiac (Card.), and multiple districts (Multi.). The network are trained with: single district (a, 500 obstetric images), single district (b, 1500 obstetric images) and multiple district images

Test data set	Ob.	Msk.	Card.	Multi.
Training with				
Single district (a)	35.93	25.88	28.46	29.47
Single district (b)	34.52	26.33	28.63	29.51
Multiple district	33.70	28.41	33.82	30.33

result confirms that we achieve the real-time computation target, required by the industrial constraint.

We underline that the input resolution of the network is 600×600, which is reached through the zero-padding of each input image. The computational cost of the prediction depends on the resolution of the input image and on the architecture of the network: in particular, the computational cost of a convolution operation is $\mathcal{O}(r/s_r \cdot c/s_c) \cdot (f_r \cdot f_c) \cdot f$; in our application, the input image has a resolution of $r = c = 600$, the kernel-filter size on rows and columns is $f_r = f_c = 4$, the stride on rows and columns is $s_r = s_c = 2$, we use 10 convolution and 10 deconvolution operators, and a number of kernel-filters from 32 to 512.

4.5 Denoising on different learning architectures

We compare the prediction results of two different networks: Pix2Pix and the (*Matlab*) CNN, as part of our deep learning framework. Figure 11 shows that Pix2Pix has better results than the CNN, in terms of blurring reduction, noise removal, and edge preservation. We also compare the quantitative

metrics between the target images and the predicted images, on the test data set of 50 ultrasound images of the obstetric anatomic district (Section 3.3). Pix2Pix has a PSNR average value of 36.07, and an SSIM average value of 0.878, while the CNN has a PSNR average value of 25.69, and an SSIM average value of 0.651. This result underlines that Pix2Pix outperforms the CNN as network architecture for our deep learning framework.

5 Discussion

Several ultrasound machines manufactured by main competitors (e.g. Esaote, Philips) are equipped with GPU cards [1, 2] Furthermore, some recent denoising methods for ultrasound images are developed on GPUs [9, 22], which are also used for denoising. Furthermore, the application of GPUs to image processing for future medical ultrasound imaging systems [59] presents the advantages of GPUs over CPUs in terms of performance, power consumption, and cost.

Denoising of ultrasound images is relevant both for postprocessing and visual evaluation by medical experts. Despite some relevant works consider raw ultrasound images and videos for cardiac segmentation [38, 47], several works show the benefits of denoising for *segmentation* [62, 70, 75], *feature extraction* [28], *classification* [57, 65], *super-resolution* [31], *registration* [17], and *texture analysis* [52]. Furthermore, main ultrasound machine manufacturers include a denoising filter in their scanners [1, 3]. In ultrasound denoising, the main goal is to achieve the best compromise between noise removal, features preservation, and real-time execution. The use of deep learning allows us to reach real-time performance, which is required by the clinical practice while preserving the denoising results of state-of-the-art methods, which are time-consuming. In our framework, deep learning

Fig. 11 (**a**) Input, (**b**) target, (**c**) our prediction based on Pix-2Pix, and (**d**) CNN prediction for the obstetric district

(a) Input

(b) Target

Prediction: (c) Ours

(d) CNN
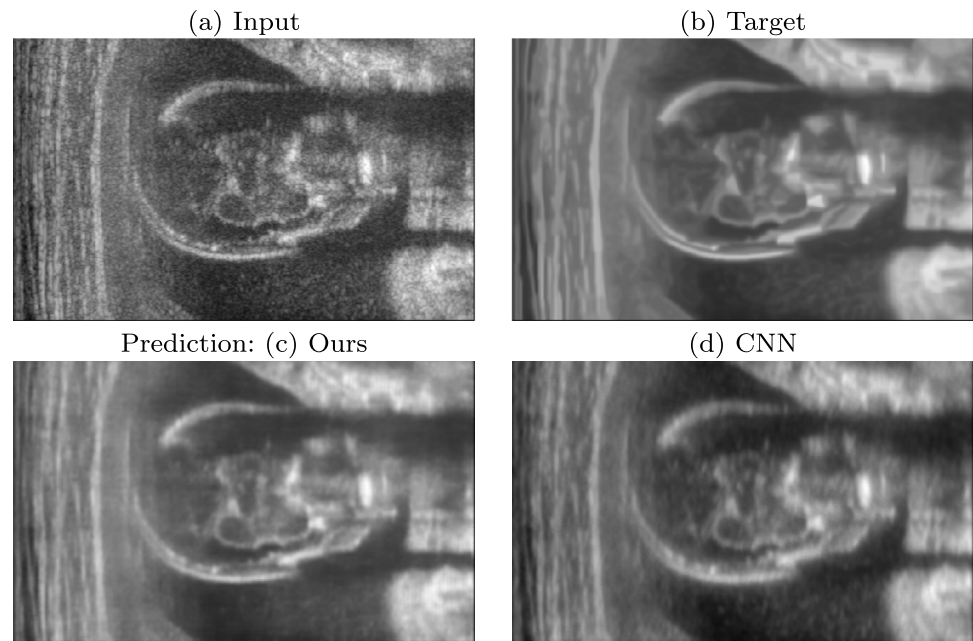
preserves the quality of WNNM and reaches its results in real-time. Furthermore, the deep learning for the real-time processing of ultrasound images has been applied in several works [37, 58].

Our method reaches real-time performance and high-quality denoising results, through a learning-based approach. In contrast, fast handcrafted methods [26] have lower results in terms of noise removal and edges enhancement; GPU-based methods [23] have higher hardware requirements than our method; other denoising methods [68] have good results in terms of noise removal, but they cannot reach a real-time implementation, due to high computational cost. Our framework also allows us to tune the denoising algorithm to obtain the best denoising results, as this tuning only affects the training phase, while the real-time computation of the denoised image is performed through the prediction of the network.

6 Conclusions and future work

We have presented a novel deep learning framework for real-time denoising of ultrasound images, which is general enough to be applied to different anatomical districts and noise levels. As the main contribution, the proposed real-time denoising of ultrasound images is general in terms of the input data, i.e. type of noise (e.g. speckle, Gaussian noise), the resolution and the dimensionality of the input images (e.g. isotropic/anisotropic, 2D/3D images), the acquisition methodology, and the anatomical district. We also mention its generality in terms of building blocks and parameters of the deep learning framework, i.e. the

denoising algorithms (e.g. WNNM, SAR-BM3D) and the deep learning architecture (e.g. Pix2Pix, VGG19).

As future work, we plan to apply our framework to data acquired with different methodologies (e.g. 3D ultrasound, MRI), also taking into account time-dependent data (e.g. ultrasound videos). Finally, the industrial and clinical validations of the proposed framework are under development, by comparing our results with tools currently used in medical clinics.

Appendix: . Additional material

A.1 Quantitative comparison

We compare the five selected denoising methods (Section 4) on synthetic images, by adding speckle noise with different levels of noise intensity: given a noisy image $Y = X + NX$, where X is the normalised ground truth image, we define the artificial multiplicative noise $N(x) = \sqrt{12}\sigma u$, where $u \sim \mathcal{U}(-0.5, 0.5)$, $\mathcal{U}$ is a uniform distribution, σ is the noise intensity, and x is a pixel of the image.

The *SIPI data set* [64] is composed of 44 ground truth images of different sizes, organised in different classes (e.g. humans, landscapes). We evaluate the efficiency of the denoising methods: WNNM has very good results in terms of noise removal, edge preservation (e.g. vehicles shape (Fig. 12), and hat feathers (Fig. 13). SAR-BM3D has the best results in terms of noise removal; however, it does not correctly preserve the grey-scale values (e.g. boy's sleeve in Fig. 13) and it generates a blurred effect (e.g. grass and bushes in Fig. 12). PCA-BM3D and NCSR show minor

preservation of edges and details than WNNM (e.g. boy's face in Fig. 13). Finally, BM-CNN is not able to correctly remove the noise; this result underlines the importance of the training data set (e.g. the type and the intensity of the applied noise) when using a deep learning approach, and the necessity of using datsa-specific networks, instead of a generic-purpose one.

Concerning the metrics introduced in Section 3.1, Table 4 summarises the results of the five denoising methods on the SIPI data set; we compute the average value of each metric (i.e. PSNR and SSIM) among 44 images of the data set, and report the average values when varying the intensity of the speckle noise. SAR-BM3D has the best results under these metrics, outperforming all the other methods. The NCSR, WNNM, and PCA-BM3D methods have good and similar results in terms of PSNR and SSIM indices. These four methods show a small degradation of the metrics values when increasing the noise intensity; this result is significant

Fig. 12 Input (SIPI data set, van image), noisy (speckle noise intensity $\sigma = 0.10$), and denoised images. For error metrics, we refer the reader to Table 4
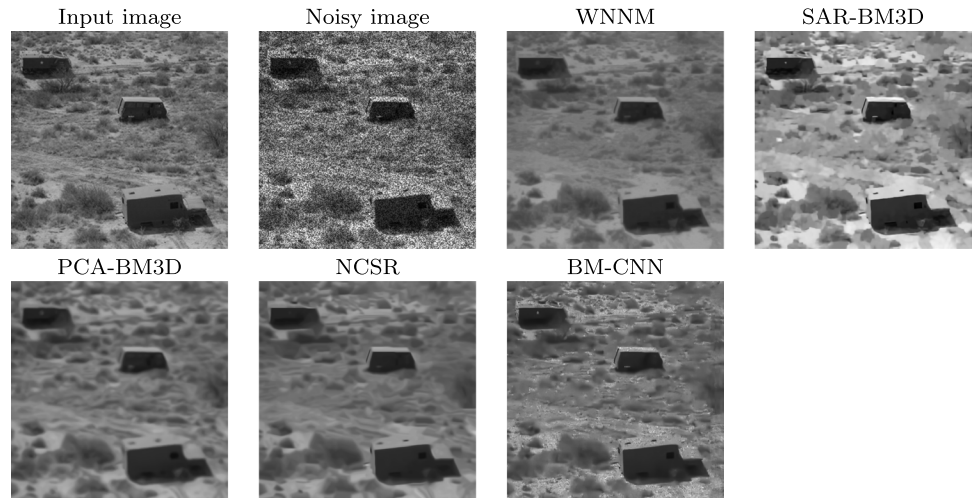
Fig. 13 Input (SIPI data set, man image), noisy (speckle noise intensity $\sigma = 0.20$), and denoised images. For error metrics, we refer the reader to Table 4
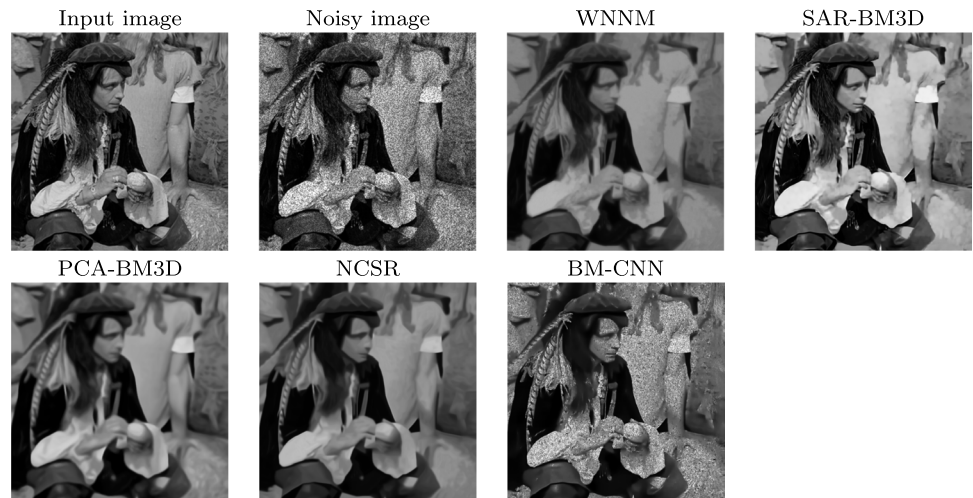
Table 4 PSNR and SSIM metrics of the denoising methods tested on the SIPI data set. For each σ value (i.e. the intensity of the speckle noise), we report the average metric computed on the 44 images of the data set

Metric	PSNR				SSIM			
Method | σ	0.05	0.1	0.2	0.3	0.05	0.1	0.2	0.3
WNNM	25.57	24.68	23.35	22.32	0.681	0.659	0.630	0.602
SAR-BM3D	**27.36**	**26.01**	**24.71**	**23.65**	0.730	**0.699**	**0.673**	**0.651**
PCA-BM3D	25.09	24.36	23.05	22.10	0.652	0.640	0.614	0.585
NCSR	26.60	25.36	23.61	22.36	0.665	0.669	0.619	0.588
BM3D-CNN	26.85	24.20	20.21	17.26	**0.733**	0.569	0.31	0.215

Table 5 PSNR and SSIM metrics of WNNM and tuned-WNNM, tested on the SIPI data set. For each σ value (i.e. the intensity of the speckle noise), we report the average metric computed on the 44 images of the data set

Metric	PSNR				SSIM			
Method | σ	0.05	0.1	0.2	0.3	0.05	0.1	0.2	0.3
WNNM	25.57	24.68	23.35	22.32	0.681	0.659	0.630	0.602
tuned-WNNM	**25.60**	**24.76**	**23.49**	**22.61**	**0.685**	**0.663**	**0.642**	**0.627**

Fig. 14 Input (256 × 256), noisy (speckle noise intensity $\sigma = 0.05$), and denoised images with WNNM and the tuned-WNNM

Input image   Noisy image   WNNM   Tuned-WNNM
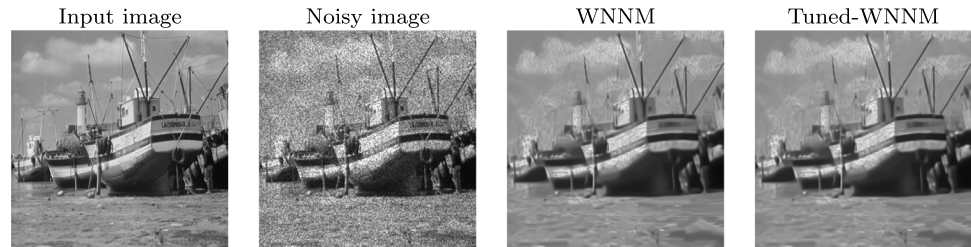
for ultrasound images, which generally have a different noise intensity, according to the anatomical district, the type of probe, and the data acquisition modality. Finally, BM-CNN shows a higher degradation of the PSNR and SSIM values, when increasing the noise intensity. The quantitative analysis is useful to compare methods with numerical measures, instead of performing only a visual evaluation. However, the main comparison among methods is the qualitative evaluation performed by the medical experts on ultrasound images, through the evaluation of the speckle noise removal and the preservation of anatomical features. We underline that, even if SAR-BM3D has better results than WNNM on synthetic images, WNNM has better performance on ultrasound images. Furthermore, our framework is general enough to use different denoising methods; two different learning networks can be trained, with WNNM and SAR-BM3D, to offer the physician the comparison between two denoising results.

A.2 Tuned-WNNM

Comparing the baseline WNNM with the tuned-WNNM on synthetic images, we improve the denoising quality (Fig. 14) in terms of quantitative metrics; in fact, the output of WNNM has a PSNR value of 26.67, while the output of tuned-WNNM has a PSNR value of 26.74. Nevertheless, the execution time of WNNM is 94 seconds, while the execution time of tuned-WNNM is 260 seconds. We also compare tuned-WNNM and WNNM on the SIPI data set (Table 5). The aggregated results show that tuned-WNNM has slightly better performance with low noise intensity, while the results improve when the speckle noise is higher. For example, WNNM has an average PSNR value of 22.32 with a speckle

noise of intensity $\sigma = 0.3$, while tuned-WNNM has a PSNR value of 22.61.

References

1. Esaote mylab-9. https://www.esaote.com/ultrasound/ultrasound-systems/p/mylab-9/. Accessed: 2022-02-11
2. Philips epiq elite. https://www.philips.it/c-dam/b2bhc/it/events/siumb/Epiq_Elite_GISS_6.0_452299156181_LR200-vb.pdf. Accessed: 2022-02-11
3. Philips xres pro. https://www.usa.philips.com/healthcare/education-resources/technologies/ultrasound/xres. Accessed: 2022-02-11

4. Aharon M, Elad M, Bruckstein A (2006) K-svd: An algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans on Signal Processing 54(11):4311–4322

5. Ahn B, Cho NI (2017) Block-matching convolutional neural network for image denoising. arXiv:1704.00524

6. Aja-Fernández S, Alberola-López C (2006) On the estimation of the coefficient of variation for anisotropic diffusion speckle filtering. Trans Image Process 15(9):2694–2701

7. Bai J, Feng X (2007) Fractional-order anisotropic diffusion for image denoising. Trans Image Process 16(10):2492–2502

8. Batson J, Royer L (2019) Noise2self: Blind denoising by self-supervision. arXiv:1901.11365

9. Biswas B, Sen BK, Dey KN (2018) Ultrasound medical image deblurring and denoising method using variational model on cuda. In: Advanced computing and systems for security, pp 95–108. Springer

10. Boyd S, Parikh N, Chu E, Peleato B, Eckstein J (2011) Distributed optimization and statistical learning via the alternating direction method of multipliers. Found Trends Mach Learn 3(1):1–122

11. Buades A, Coll B, Morel JM (2005) A non-local algorithm for image denoising. In: Conf. on computer vision and pattern recognition, vol 2, pp 60–65

12. Cammarasana S, Nicolardi P, Patanè G (2021) A universal deep learning framework for real-time denoising of ultrasound images. arXiv:2101.09122

13. Chang SG, Yu B, Vetterli M (2000) Adaptive wavelet thresholding for image denoising and compression. Trans Image Process 9(9):1532–1546

14. Da Cunha AL, Zhou J, Do MN (2006) The nonsubsampled contourlet transform: theory, design, and applications. Trans Image Process 15(10):3089–3101

15. Dabov K, Foi A, Katkovnik V, Egiazarian K (2006) Image denoising with block-matching and 3D filtering. In: Image processing: algorithms and systems, neural networks, and machine learning, vol 6064, p 606414

16. Dabov K, Foi A, Katkovnik V, Egiazarian K (2009) BM3D image denoising with shape-adaptive principal component analysis. In: Gribonval R (ed) Signal processing with adaptive sparse structured representations

17. De Silva T, Fenster A, Cool DW, Gardi L, Romagnoli C, Samarabandu J, Ward AD (2013) 2d-3d rigid registration to compensate for prostate motion during 3d trus-guided biopsy. Med Phys 40 (2):022904

18. Dong W, Shi G, Li X (2012) Nonlocal image restoration with bilateral variance estimation: a low-rank approach. Trans Image Process 22(2):700–711

19. Dong W, Zhang L, Shi G, Li X (2012) Nonlocally centralized sparse representation for image restoration. Trans Image Process 22(4):1620–1630

20. Eriksson A, Van Den Hengel A (2010) Efficient computation of robust low-rank matrix approximations in the presence of missing data using the L1 norm. In: 2010 IEEE Conf. on computer vision and pattern recognition, pp 771–778

21. Fang Y, Zeng T (2020) Learning deep edge prior for image denoising. Comput Vis Image Underst 200:103044

22. Palhano Xavier de Fontes F, Andrade Barroso G, Coupé P, Hellier P (2011) Real time ultrasound image denoising. J Real-time Image Process 6(1):15–22

23. Fredj AH, Malek J (2016) Real time ultrasound image denoising using nvidia cuda. In: 2016 2nd International conference on advanced technologies for signal and image processing (ATSIP), pp 136–140. IEEE

24. Frost VS, Stiles JA, Shanmugan KS, Holtzman JC (1982) A model for radar images and its application to adaptive digital filtering of multiplicative noise. Trans Pattern Anal Mach Intell 4(2):157–166

25. Fu X, Jia S, Zhuang L, Xu M, Zhou J, Li Q (2021) Hyperspectral anomaly detection via deep plug-and-play denoising cnn regularization. Trans Geosci Rem Sens 59(11):9553–9568

26. Garg A, Khandelwal V (2019) Despeckling of medical ultrasound images using fast bilateral filter and neighshrinksure filter in wavelet domain. In: Advances in signal processing and communication, pp 271–280. Springer

27. Gu S, Zhang L, Zuo W, Feng X (2014) Weighted nuclear norm minimization with application to image denoising. In: Proc. of the IEEE conf. on computer vision and pattern recognition, pp 2862–2869

28. Iakovidis DK, Keramidas EG, Maroulis D (2008) Fuzzy local binary patterns for ultrasound texture characterization. In: Inter. Conf. image analysis and recognition, pp 750–759. Springer

29. Isola P, Zhu J, Zhou T, Efros AA (2017) Image-to-image translation with conditional adversarial networks. In: Conf. on Computer vision and pattern recognition, pp 1125–1134

30. Kervrann C, Boulanger J, Coupé P. (2007) Bayesian non-local means filter, image redundancy and adaptive dictionaries for noise removal. In: Inter. Conf. on scale space and variational methods in computer vision, pp 520–532. Springer

31. Khavari P, Asif A, Rivaz H (2018) Non-local super resolution in ultrasound imaging. In: 20th International workshop on multimedia signal processing (MMSP), pp 1–6. IEEE

32. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp 1097–1105

33. Krull A, Buchholz TO, Jug F (2019) Noise2void-learning denoising from single noisy images. In: Proc. of the IEEE conf. on computer vision and pattern recognition, pp 2129–2137

34. Kuan DT, Sawchuk AA, Strand TC, Chavel P (1985) Adaptive noise smoothing filter for images with signal-dependent noise. Trans Pattern Anal Mach Intell 7(2):165–177

35. Lee J (1980) Digital image enhancement and noise filtering by use of local statistics. IEEE Trans Pattern Anal Mach Intell 2(2):165–168

36. Lehtinen J, Munkberg J, Hasselgren J, Laine S, Karras T, Aittala M, Aila T (2018) Noise2noise: learning image restoration without clean data. arXiv:1803.04189

37. Liu S, Wang Y, Yang X, Lei B, Liu L, Li S, Ni D, Wang T (2019) Deep learning in medical ultrasound analysis: a review. Engineering 5(2):261–275

38. Liu X, Fan Y, Li S, Chen M, Li M, Hau WK, Zhang H, Xu L, Lee APW (2021) Deep learning-based automated left ventricular ejection fraction assessment using 2-d echocardiography. Am J Physiol-Heart Circul Physiol 321(2):H390–H399

39. Liu X, Zhang H, Cheung Ym, You X, Tang YY (2017) Efficient single image dehazing and denoising: an efficient multi-scale correlated wavelet approach. Comput Vis Image Understand 162:23–33

40. Ma L, Xu L, Zeng T (2017) Low rank prior and total variation regularization for image deblurring. J Sci Comput 70(3):1336–1357

41. Ma L, Zeng T (2016) Image deblurring via total variation based structured sparse model selection. J Sci Comput 67(1):1–19

42. MacQueen J, et al. (1967) Some methods for classification and analysis of multivariate observations. In: Proc. of the symposium on mathematical statistics and probability, vol 1, pp 281–297

43. Mairal J, Bach F, Ponce J, Sapiro G, Zisserman A (2009) Non-local sparse models for image restoration. In: IEEE Inter. conf. on computer vision, pp 2272–2279

44. Maleki A, Narayan M, Baraniuk RG (2013) Anisotropic nonlocal means denoising. Appl Comput Harmon Anal 35(3):452–482

45. Mei F, Zhang D, Yang Y (2020) Improved non-local self-similarity measures for effective speckle noise reduction in ultrasound images. Comput Methods Programs Biomed 196:105670

46. Mihcak MK, Kozintsev I, Ramchandran K, Moulin P (1999) Low-complexity image denoising based on statistical modeling of wavelet coefficients. Signal Process Lett 6(12):300–303

47. Ouyang D, He B, Ghorbani A, Yuan N, Ebinger J, Langlotz CP, Heidenreich PA, Harrington RA, Liang DH, Ashley EA et al (2020) Video-based ai for beat-to-beat assessment of cardiac function. Nature 580(7802):252–256

48. Parrilli S, Poderico M, Angelino CV, Verdoliva L (2011) A non-local sar image denoising algorithm based on llmmse wavelet shrinkage. Trans Geosci Rem Sens 50(2):606–616

49. Patané G (2015) Diffusive smoothing of 3d segmented medical data. J Adv Res 6(3):425–431

50. Perona P, Malik J (1990) Scale-space and edge detection using anisotropic diffusion. Trans Pattern Anal Mach Intell 12(7):629–639

51. Portilla J, Strela V, Wainwright MJ, Simoncelli EP (2003) Image denoising using scale mixtures of gaussians in the wavelet domain. Trans Image Process 12(11):1338–1351

52. Puri M, Patil K, Balasubramanian V, Narayanamurthy V (2005) Texture analysis of foot sole soft tissue images in diabetic neuropathy using wavelet transform. Med Biol Eng Comput 43(6):756–763

53. Radford A, Metz L, Chintala S (2015) Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv:1511.0643

54. Rajwade A, Rangarajan A, Banerjee A (2012) Image denoising using the higher order singular value decomposition. Trans Pattern Anal Mach Intell 35(4):849–862

55. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In: Inter. Conf. on medical image computing and computer-assisted intervention, pp 234–241. Springer

56. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556

57. Sivanandan R, Jayakumari J (2021) A new cnn architecture for efficient classification of ultrasound breast tumor images with activation map clustering based prediction validation. Med Biol Eng Comput 59 (4):957–968

58. van Sloun RJG, Cohen R, Eldar YC (2020) Deep learning in ultrasound imaging. Proc IEEE 108(1):11–29. https://doi.org/10.1109/JPROC.2019.2932116

59. So H, Chen J, Yiu B, Yu A (2011) Medical ultrasound imaging: To gpu or not to gpu? Micro 31(5):54–65

60. Srebro N, Jaakkola T (2003) Weighted low-rank approximations. In: Proc. of the inter. conf. on machine learning, pp 720–727

61. Starck JL, Candès E. J, Donoho DL (2002) The curvelet transform for image denoising. Trans Image Process 11(6):670–684

62. Tang J, Guo S, Sun Q, Deng Y, Zhou D (2010) Speckle reducing bilateral filter for cattle follicle segmentation. BMC genomics 11(2):1–9

63. Verma R, Pandey R (2017) Adaptive selection of search region for nlm based image denoising. Optik 147:151–162

64. Weber AG (1997) The usc-sipi image database version 5 USC-SIPI Report 315(1)

65. Wei M, Du Y, Wu X, Su Q, Zhu J, Zheng L, Lv G, Zhuang J (2020) A benign and malignant breast tumor classification method via efficiently combining texture and morphological features on ultrasound images. Comput Math Methods Med, 2020

66. Wold S, Esbensen K, Geladi P (1987) Principal component analysis. Chemom Intell Lab Syst 2(1–3):37–52

67. Wu T, Li W, Jia S, Dong Y, Zeng T (2020) Deep multi-level wavelet-cnn denoiser prior for restoring blurred image with cauchy noise. Signal Process Lett 27:1635–1639

68. Xu J, Zhang L, Zhang D (2018) A trilateral weighted sparse coding scheme for real-world image denoising. In: Proceedings of the European conference on computer vision (ECCV), pp 20–36

69. Xu S, Yang X, Jiang S (2017) A fast nonlocally centralized sparse representation algorithm for image denoising. Signal Process 131:99–112

70. Yang F, Qin W, Xie Y, Wen T, Gu J (2012) A shape-optimized framework for kidney segmentation in ultrasound images using nltv denoising and drlse. Biomed Eng Online 11(1):1–13

71. Yang HY, Wang XY, Niu PP, Liu YC (2014) Image denoising using nonsubsampled shearlet transform and twin support vector machines. Neural Netw 57:152–165

72. Yu Y, Acton ST (2002) Speckle reducing anisotropic diffusion. Trans Image Process 11 (11):1260–1270

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

74. Zhang M, Desrosiers C (2018) Structure preserving image denoising based on low-rank reconstruction and gradient histograms. Comput Vis Image Underst 171:48–60

75. Zhuang Z, Lei N, Raj ANJ, Qiu S (2019) Application of fractal theory and fuzzy enhancement in ultrasound image segmentation. Med Biol Eng Comput 57(3):623–632

Simone Cammarasana is a Ph.D. student in Computer Science at the University of Genova-DIBRIS. He obtained a post-lauream Master in Scientific Computing at the University of Sapienza-Roma (2018), and a Master degree in Engineering at the University of Pisa. His research interests include signals analysis, optimisation problems, and linear systems.

Paolo Nicolardi is image processing and algorithms technical leader at Esaote. He obtained a Master degree in Engineering at Politecnico di Milano, in 2005. His research interests include image processing, computer vision, pattern recognition, machine learning, and medical images.