

Deep Learning in Ultrasound Imaging

This article provides an overview of use of deep, data-driven learning strategies in ultrasound systems, from the front-end to advanced applications. The authors discuss the use of these new computational approaches in all aspects of ultrasound imaging, ranging from ideas that are at the interface of raw signal acquisition (including adaptive beam forming) and image formation, to learning compressive codes for color Doppler acquisition to learning strategies for performing clutter suppression.

By RUUD J. G. VAN SLOUN¹, Member IEEE, REGEV COHEN, Graduate Student Member IEEE,
AND YONINA C. ELДАР², Fellow IEEE

ABSTRACT | In this article, we consider deep learning strategies in ultrasound systems, from the front end to advanced applications. Our goal is to provide the reader with a broad understanding of the possible impact of deep learning methodologies on many aspects of ultrasound imaging. In particular, we discuss methods that lie at the interface of signal acquisition and machine learning, exploiting both data structure (e.g., sparsity in some domain) and data dimensionality (big data) already at the raw radio-frequency channel stage. As some examples, we outline efficient and effective deep learning solutions for adaptive beamforming and adaptive spectral Doppler through artificial agents, learn compressive encodings for the color Doppler, and provide a framework for structured signal recovery by learning fast approximations of iterative minimization problems, with applications to clutter suppression and super-resolution ultrasound. These emerging technologies may have a considerable impact on ultrasound imaging, showing promise across key components in the receive processing chain.

KEYWORDS | Beamforming; compression; deep learning; deep unfolding; Doppler; image reconstruction; super resolution; ultrasound imaging.

Manuscript received April 11, 2019; revised June 24, 2019 and July 25, 2019; accepted July 27, 2019. Date of publication August 21, 2019; date of current version December 26, 2019. (Corresponding author: Ruud J. G. van Sloun.)

R. J. G. van Sloun is with the Department of Electrical Engineering, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands (email: r.j.g.v.sloun@tue.nl).

R. Cohen is with the Department of Electrical Engineering, Technion, Haifa, Israel.

Y. C. Eldar is with the Faculty of Mathematics and Computer Science, Weizmann Institute of Science, Rehovot 7610001, Israel (e-mail: yonina.eldar@weizmann.ac.il).

Digital Object Identifier 10.1109/JPROC.2019.2932116

I. INTRODUCTION

Diagnostic imaging plays a critical role in healthcare, serving as a fundamental asset for timely diagnosis, disease staging, and management, as well as for treatment choice, planning, guidance, and follow-up. Among the diagnostic imaging options, ultrasound imaging [1] is uniquely positioned, being a highly cost-effective modality that offers the clinician an unmatched and invaluable level of interaction, enabled by its real-time nature. Its portability and cost effectiveness permit point-of-care imaging at the bedside, in emergency settings, rural clinics, and developing countries. Ultrasonography is increasingly used across many medical specialties, spanning from obstetrics to cardiology and oncology, and its market share is globally growing.

On the technological side, ultrasound probes are becoming increasingly compact and portable, with the market demand for low-cost “pocket-sized” devices expanding [2], [3]. Transducers are miniaturized, allowing, e.g., in-body imaging for interventional applications. At the same time, there is a strong trend toward 3-D imaging [4] and the use of high-frame-rate imaging schemes [5]; both accompanied by dramatically increasing data rates that pose a heavy burden on the probe-system communication and subsequent image reconstruction algorithms. Systems today offer a wealth of advanced applications and methods, including shear wave elasticity imaging (SWEI) [6], ultrasensitive Doppler [7], and ultrasound localization microscopy (ULM) for super-resolution microvascular imaging [8].

With the demand for high-quality image reconstruction and signal extraction from unfocused planar wave transmissions that facilitate fast imaging, and a push toward

miniaturization, modern ultrasound imaging leans heavily on innovations in powerful receive channel processing. In this article, we discuss how artificial intelligence and deep learning methods can play a compelling role in this process and demonstrate how these data-driven systems can be leveraged across the ultrasound imaging chain. We aim to provide the reader with a broad understanding of the possible impact of deep learning on a variety of ultrasound imaging aspects, placing particular emphasis on methods that exploit both the power of data and signal structure (for instance sparsity in some domain) to yield robust and data-efficient solutions. We believe that methods that exploit models and structure together with learning from data can pave the way to interpretable and powerful processing methods from limited training sets. As such, throughout this article, we will typically first discuss an appropriate model-based solution for the problems considered and then follow by a data-driven deep learning solution derived from it.

We start by briefly describing a standard ultrasound imaging chain in Section II. We then elaborate on several dedicated deep learning solutions that aim at improving key components in this processing pipeline, covering adaptive beamforming (see Section III-A), adaptive spectral Doppler (see Section III-B), compressive tissue Doppler (see Section III-C), and clutter suppression (see Section III-D). In Section IV, we show how the synergistic exploitation of deep learning and signal structure enables robust super-resolution microvascular ultrasound imaging. Finally, we discuss future perspectives, opportunities, and challenges for the holistic integration of artificial intelligence and deep learning methods in ultrasound systems.

II. ULTRASOUND IMAGING CHAIN AT A GLANCE

A. Transmit Schemes

The resolution, contrast, and overall fidelity of ultrasound pulse–echo imaging relies on careful optimization across its entire imaging chain. At the front end, imaging starts with the design of appropriate transmit schemes.

At this stage, crucial tradeoffs are made, in which the frame rate, imaging depth, and attainable axial and lateral resolution are weighted carefully against each other: improved resolution can be achieved through the use of higher pulse modulation frequencies; yet, these shorter wavelengths suffer from increased absorption and thus lead to reduced penetration depth. Likewise, a high frame rate can be reached by exploiting parallel transmission schemes based on, e.g., planar or diverging waves. However, use of such unfocused transmissions comes at the cost of loss in lateral resolution compared to line-based scanning with tightly focused beams. As such, optimal transmit schemes depend on the application.

Today, an increasing amount of ultrasound applications rely on high-frame-rate (dubbed ultrafast) imaging.

Among these are, e.g., ULM (see Section IV), highly sensitive Doppler, and shear wave elastography, where the former two mostly exploit the incredible vastness of data to obtain accurate signal statistics, and the latter leverages high-speed imaging to track ultrasound-induced shear waves propagating at several meters per second.

With the expanding use of ultrafast transmit sequences in modern ultrasound imaging, a strong burden is placed on the subsequent receive channel processing. High data rates not only raise substantial hardware complications related to power consumption, data storage, and data transfer; the corresponding unfocused transmissions require much more advanced receive beamforming and clutter suppression to reach satisfactory image quality.

B. Receive Processing, Sampling, and Beamforming

Modern receive channel processing is shifting toward the digital domain, relying on computational power and very-high-bandwidth communication channels to enable advanced digital parallel (pixel-based) beamforming and coherent compounding across multiple transmit/receive events. For large channel counts, e.g., in dense matrix probes that facilitate high-resolution 3-D imaging, the number of coaxial cables required to connect all probe elements to the back-end system quickly becomes infeasible. To address this, dedicated switching and processing already take place in the probe head, e.g., in the form of multiplexing or microbeamforming. Slow-time¹ multiplexing distributes the received channel data across multiple transmits, by only communicating a subset of the number of channels to the back end for each such transmit. This consequently reduces the achieved frame rate. In microbeamforming, an analog prebeamforming step is performed to compress channel data from multiple (adjacent) elements into a single focused line. This, however, impairs flexibility in subsequent digital beamforming, limiting the achievable image quality. Other approaches aim at mixing multiple channels through analog modulation with chipping sequences [9]. Additional analog processing includes signal amplification by a low-noise amplifier (LNA) as well as depth (i.e., fast time) dependent gain compensation (TGC) for attenuation correction.

Digital receive beamforming in ultrasound imaging is dynamic, i.e., receive focusing is dynamically optimized based on the scan depth. The industry standard is delay-and-sum beamforming, where depth-dependent channel tapering (or apodization) is optimized and fine-tuned based on the system and application. Delay-and-sum beamforming is commonplace due to its low complexity, providing real-time image reconstruction, albeit at a high sampling rate and nonoptimal image quality.

¹In ultrasound imaging, we make a distinction between slow time and fast time: slow time refers to a sequence of snapshots (i.e., across multiple transmit/receive events) at the pulse repetition rate, whereas fast time refers to samples along depth.

Performing beamforming in the digital domain requires sampling the signals received at the transducer elements and transmitting the samples to a back-end processing unit. To achieve sufficient delay resolution for focusing, the received signals are typically sampled at four to ten times their bandwidth, i.e., the sampling rate may severely exceed the Nyquist rate. A possible approach for sampling rate reduction is to consider the received signals within the framework of the finite rate of innovation (FRI) [10], [11]. Tur *et al.* [12] modeled the received signal at each element as a finite sum of replicas of the transmitted pulse backscattered from reflectors. The replicas are fully described by their unknown amplitudes and delays, which can be recovered from the signals' Fourier series coefficients. The latter can be computed from low-rate samples of the signal using compressed sensing (CS) techniques [10], [13]. Wagner *et al.* [14] and Wagner *et al.* [15] extended this approach and introduced compressed beamforming. It was shown that the beamformed signal follows an FRI model, and thus, it can be reconstructed from a linear combination of the Fourier coefficients of the received signals. Moreover, these coefficients can be obtained from low-rate samples of the received signals taken according to the Sampling framework [16]–[18]. Chernyakova and Eldar [3] showed that this Fourier domain relationship between the beam and the received signals holds irrespective of the FRI model. This leads to a general concept of the frequency-domain beamforming (FDBF) [3] that is equivalent to beamforming in time. FDBF allows sampling the received signals at their effective Nyquist rate without assuming a structured model; thus, it avoids the oversampling dictated by digital implementation of beamforming in time. Furthermore, when assuming that the beam obeys an FRI model, the received signals can be sampled at sub-Nyquist rates, leading to up to 28-fold reduction in the sampling rate [19]–[21].

C. B-Mode, M-Mode, and Doppler

Ultrasound imaging provides anatomical information through the so-called brightness mode (B-mode). B-mode imaging is performed by envelope-detecting the beamformed signals, e.g., through the calculation of the magnitude of the complex in-phase and quadrature (IQ) data. For visualization purposes, the dynamic range of these envelope-detected signals is subsequently compressed via a logarithmic transformation or specifically designed compression curves based on a lookup table. Scan conversion then maps these intensities to the desired (Cartesian) pixel coordinate system. The visualization of a single B-mode scan line (i.e., brightness over fast time) across multiple transmit–receive events (i.e., slow time), is called motion-mode (M-mode) imaging.

Beyond anatomical information, ultrasound imaging also permits the measurement of functional parameters related to blood flow and tissue displacement. The

extraction of such velocity signals is called Doppler processing. We distinguish between two types of velocity estimators: color Doppler and spectral Doppler. Color Doppler provides an estimate of the mean velocity through the evaluation of the first lag of the autocorrelation function for a series of snapshots across slow time [22]. Spectral Doppler provides the entire velocity distribution in a specified image region through estimation of the full power spectral density and visualizes its evolution over time in a spectrogram [23]. Spectral Doppler methods are relevant for, e.g., detecting turbulent flow in stenotic arteries or across heart valves. Besides assessing blood flow, Doppler processing also finds applications in measurement of tissue velocities (tissue Doppler), e.g., for assessment of the myocardial strain.

D. Advanced Applications

In addition to B-mode, M-mode, and Doppler scanning, ultrasound data are used in a number of advanced applications. For instance, elastography methods aim at measuring mechanical parameters related to tissue elasticity and rely on the analysis of displacements following some form of imposed stress. Stress may be delivered manually (through gentle pushing), naturally (e.g., in the myocardium of a beating heart) or acoustically, as done in acoustic radiation force impulse imaging (ARFI) [24]. Alternatively, the speed of laterally traveling shear waves induced by an acoustic push-pulse can be measured, with this speed being directly related to the shear modulus [6]. SWEI also permits measurement of tissue viscosity in addition to stiffness through the assessment of wave dispersion [25]. All the above-mentioned methods rely on the adequate measurement of local tissue velocity or displacement through some form of tissue Doppler processing.

While Doppler methods enable estimation of blood flow, detection of low-velocity microvascular flow is challenging since its Doppler spectrum overlaps with that of the strong tissue clutter. Contrast-enhanced ultrasound (CEUS) permits visualization and characterization of microvascular perfusion through the use of gas-filled microbubbles [26], [27]. These intravascular bubbles are sized similar to red blood cells, reaching the smallest capillaries in the vascular net, and exhibit a particular nonlinear response when insonified. The latter is specifically exploited in contrast-enhanced imaging schemes that aim at isolating this nonlinear response through the dedicated pulse sequences. Unfortunately, this does not lead to complete tissue suppression since the tissue itself also generates harmonics [28]. Thus, clutter rejection algorithms are becoming increasingly popular, in particular when used in conjunction with ultrafast imaging [29].

Recent developments also leverage the microbubbles used in CEUS to yield super-resolution imaging [30]–[33]. ULM is a particularly popular approach to achieve this [8]. ULM methods rely on adequate detection, isolation,

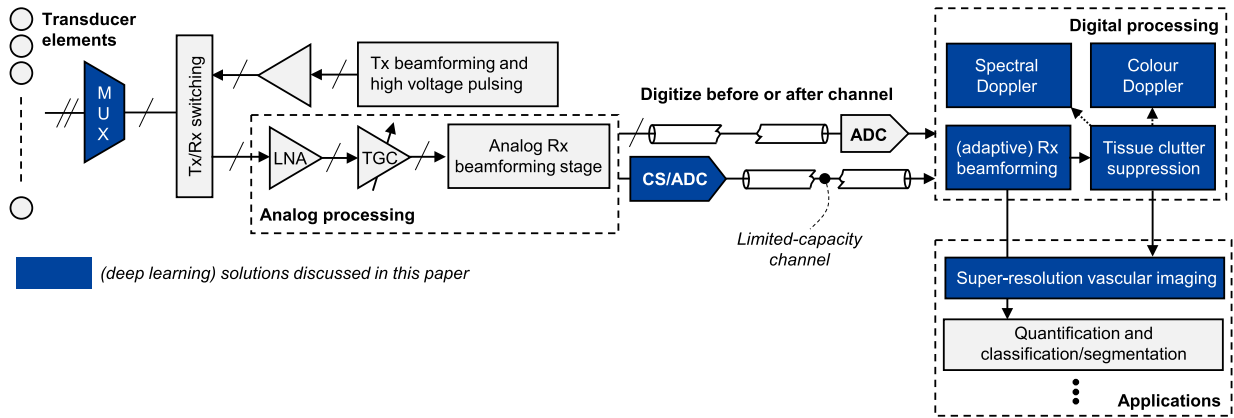


Fig. 1. Overview of the ultrasound imaging chain, along with the deep learning solutions discussed in this article. Note that, today, analog processing at the front end typically comprises some form of lossy (micro-)beamforming to reduce data rates, in contrast to the here advocated paradigm based on compressive sub-Nyquist sampling, intelligent ASICs with neural edge computing, and subsequent remote deep-learning-based processing of the low-rate channel data.

and localization of the microbubbles, typically achieved through precisely tuned tissue clutter suppression algorithms and by posing strong constraints on the allowable concentrations. We will further elaborate on this approach and its limitations in Section IV, where we discuss a dedicated deep learning solution for super-resolution ultrasound that aims at addressing some of these disadvantages.

III. DEEP LEARNING FOR (FRONT-END) ULTRASOUND PROCESSING

The effectiveness of ultrasound imaging and its applications are dictated by adequate front-end beamforming, compression, signal extraction (e.g., clutter suppression), and velocity estimation. In this section, we demonstrate how neural networks, being universal function approximators [34], can learn to act as powerful artificial agents and signal processors across the imaging chain to improve resolution and contrast, adequately suppress clutter, and enhance spectral estimation. We here refer to artificial agents [35] whenever these learned networks impact the processing chain by actively and adaptively changing the settings or parameters of a particular processor depending on the context.

Deep learning is the process of learning a hierarchy of parameterized nonlinear transformations (or layers), such that it performs a desired function. These elementary nonlinear transformations in a deep network can take many forms and may embed structural priors. A popular example of the latter is the translational invariance in images that is exploited by convolutional neural networks, but we will see that in fact many other structural priors can be exploited.

The methods proposed throughout this article are both model based and learn from data. We complement this approach with *a priori* knowledge on the signal structure to develop deep learning models that are both effective and data efficient, i.e., “fast learners.” An overview is

given in Fig. 1. We assume that the reader is familiar with the basics of (deep) neural networks. For a general introduction to deep learning, we refer the reader to [36].

A. Beamforming

1) *Deep Neural Networks as Beamformers*: The low complexity of delay-and-sum beamforming has made it the industry standard and commonplace for real-time ultrasound beamforming. There are, however, a number of factors that cause deteriorated reconstruction quality of this naive spatial filtering strategy. First, the channel delays for time-of-flight correction are based on the geometry of the scene and assume a constant speed of sound across the medium. As a consequence, variations in the speed of sound and the resulting aberrations impair proper alignment of echoes stemming from the same scatterer [37]. Second, the *a priori* determined channel weighting (apodization) of pseudo-aligned echoes before summation requires a tradeoff between mainlobe width (resolution) and sidelobe level (leakage) [38].

Delay-and-sum beamformers are typically hand-tailored based on knowledge of the array geometry and medium properties, often including specifically designed array apodization schemes that may vary across imaging depth. Interestingly, it is possible to learn the delays and apodizations from paired channel-image data through gradient-descent by dedicated “delay layers” [39]. To show this, unfocused channel data were obtained from echocardiography of six patients for both single- and multiple-line acquisitions. While the latter allows for increased frame rates, it leads to deteriorated image quality when applying standard delay-and-sum beamforming. The authors, therefore, propose to train a more appropriate delay-and-sum beamforming chain that takes multi-line channel data as an input and produces beamformed images that are as close as possible to those obtained from single-line acquisitions, minimizing their ℓ_1 distance. Since the introduced delay and apodization layers are differentiable,

efficient learning is enabled through backpropagation. Although such an approach potentially enables discovery of a more optimal set of parameters dedicated to each application, the fundamental problem of having *a priori*-determined static delays and weights remains.

Several other data-driven beamforming methods have recently been proposed. In contrast to [39], these are mostly based on “general-purpose” deep neural networks, such as stacked autoencoders [40], encoder–decoder architectures [41], and fully convolutional networks, that map predelayed channel data to beamformed outputs [42]. In the latter, a 29-layer convolutional network was applied to a 3-D stack of array response vectors for all lateral positions and a set of depths to yield a beamformed IQ output for those lateral positions and depths. Others exploit neural networks to process channel data in the Fourier domain [43]. To that end, axially gated sections of the predelayed channel data first undergo discrete Fourier-transformation. For each frequency bin, the array responses are then processed by a separate fully connected network. The frequency spectra are subsequently inverse Fourier transformed and summed across the array to yield a beamformed radio-frequency (RF) signal associated with that particular axial location. The networks were specifically trained to suppress off-axis responses (outside the first nulls of the beam) from the simulations of ultrasound channel data for point targets.

Beyond beamforming for suppression of off-axis scattering, Hyun *et al.* [44] propose deep convolutional neural networks for joint beamforming and speckle reduction. Rather than applying the latter as a post-processing technique, it is embedded in the beamforming process itself, permitting exploitation of both channel and phase information that is otherwise irreversibly lost. The network was designed to accept 16 beamformed subaperture RF signals as an input, and outputs speckle-reduced B-mode images. The final beamformed images exhibit comparable speckle-reduction as postprocessed delay-and-sum images using the optimized Bayesian nonlocal means algorithm [45], yet at an improved resolution. Additional applications of deep learning in this context include removal of artifacts in time-delayed and phase-rotated element-wise IQ data in multi-line acquisitions for high-frame-rate imaging [46] and synthesizing multiple-focus images from single-focus images through generative adversarial networks [47]. In [48], such generative adversarial networks were used for joint beamforming and segmentation of cyst phantoms from unfocused RF channel data acquired after a single plane-wave transmission.

While the flexibility and capacity of very deep neural networks in principle allow for learning context-adaptive beamforming schemes, such highly overparameterized networks notoriously rely on vast RF channel data to yield robust inference under a wide range of conditions. Moreover, large networks have a large memory footprint, complicating resource-limited implementations.

2) *Leveraging Model-Based Algorithms*: One approach for constraining the solution space while explicitly embedding adaptivity is to borrow concepts from model-based adaptive beamforming methods. These techniques steer away from the fixed-weight presumption and calculate an array apodization depending on the measured signal statistics. In the case of pixel-based reconstruction, apodization weights can be adaptively optimized per pixel. A popular adaptive beamforming method is the minimum variance distortionless response (MVDR), or Capon, beamformer, where optimal weights are defined as those that minimize signal variance/power, while maintaining distortionless response of the beamformer in the desired source direction. This amounts to solving

$$\begin{aligned} \hat{\mathbf{w}} &= \arg \min_{\mathbf{w}} \mathbf{w}^H \mathbf{R}_x \mathbf{w} \\ \text{s.t. } \mathbf{w}^H \mathbf{a} &= 1 \end{aligned} \quad (1)$$

where \mathbf{R}_x denotes the covariance matrix calculated over the receiving array elements and \mathbf{a} is a steering vector. When the receive signals are already time-of-flight corrected, \mathbf{a} is a unity vector.

Solving (1) involves the inversion of \mathbf{R}_x , whose computational complexity grows cubically with the number of array elements [50]. To improve stability, it is often combined with subspace selection through eigendecomposition, further increasing the computational burden. Another problem is the accurate estimation of \mathbf{R}_x , typically requiring some form of averaging across subarrays and the fast- and slow-time scales. While this implementation of MVDR beamforming is impractical for typical ultrasound arrays (e.g., 256 elements) or matrix transducers (e.g., 64×64 elements), it does provide a framework in which deep neural networks can be leveraged efficiently and effectively.

Instead of attempting to replace the beamforming process entirely, a neural network can be used specifically to act as an artificial agent that calculates the optimal apodization weights \mathbf{w} for each pixel, given the received predelayed channel signals at the array. By only replacing this bottleneck component in the MVDR beamformer, and constraining the problem further by promoting close-to-distortionless response during training (i.e., $\sum_i w_i \approx 1$), this solution is highly data efficient, interpretable, and has the ability to learn powerful models from only a few images [49].

The neural network proposed in [49] is compact, consisting of four fully connected layers comprising 128 nodes for the input and output layers and 32 nodes for the hidden layers. This dimensionality reduction enforces a compact representation of the data, mitigating the impact of noise. Between every fully connected layer, dropout is applied with a probability of 0.2. The input of the network is the predelayed (focused) array response for a particular pixel (i.e., a vector of length N , with N being the number of array elements), and its outputs are the corresponding

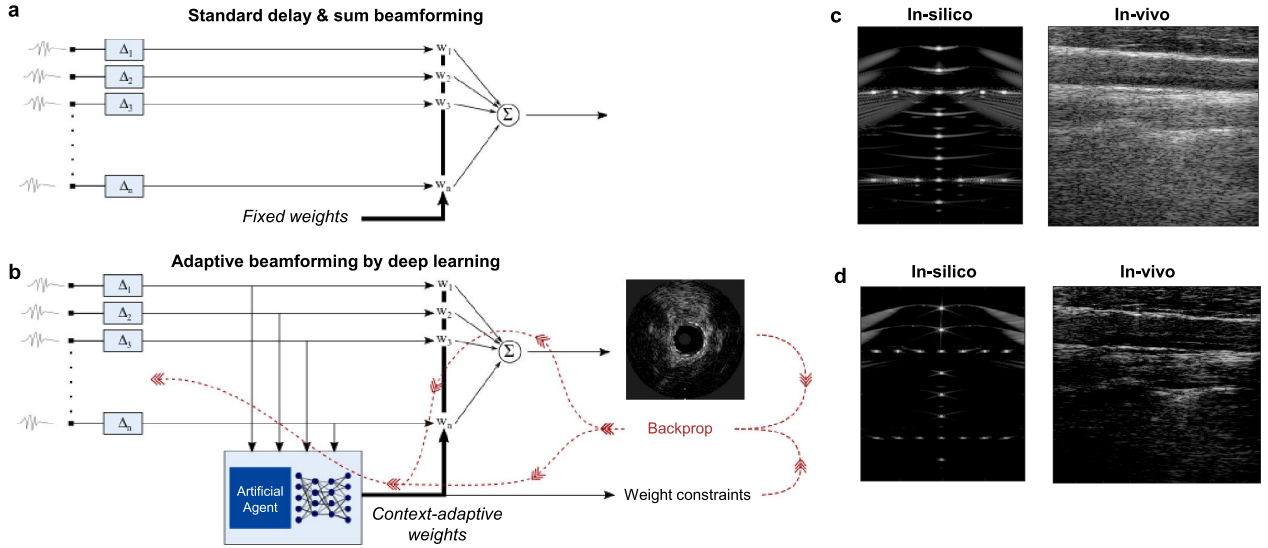


Fig. 2. (a) Flow charts of standard delay-and-sum beamforming using fixed apodization weights and (b) adaptive beamforming by deep learning [49], along with (c) and (d) illustrative reconstructed images (in silico and in vivo) for both methods, respectively. Adaptive beamforming by deep learning achieves notably better contrast and resolution and generalizes very well to unseen data sets.

array apodizations w . This apodization is subsequently applied to the network inputs to yield a beamformed pixel. Since pixels are processed independently by the network, a large amount of training data is available per acquisition. Inference is fast, and real-time rates are achievable on a GPU-accelerated system. For an array of 128 elements, adaptive calculation of a set of apodization weights through MVDR requires $> N^3 (= 2097152)$ floating point operations (FLOPS), while the deep-learning architecture only requires 74656 FLOPS [49], leading to a more than $400\times$ speedup in the reconstruction time. Additional details regarding the adopted network and training strategy are given in Section III-A3.

Fig. 2 exemplifies the effectiveness of this approach on plane-wave ultrasound acquisitions obtained using a linear array transducer. Compared to standard delay-and-sum, adaptive beamforming with a deep network serving as an artificial agent visually provides reduced clutter and enhanced tissue contrast. Quantitatively, it yields a slightly elevated contrast-to-noise ratio (CNR) (10.96 versus 11.48 dB), along with significantly improved resolution (0.43 versus 0.34 mm, and 0.85 versus 0.70 mm in the axial and lateral directions, respectively).

Interestingly, the neural network exhibits increased stability and robustness compared to the MVDR weight estimator. This can be attributed to its small bottleneck latent space, enforcing apodization weight realizations that are represented in a compact basis.

3) *Design and Training Considerations:* The large dynamic range and modulated nature of RF ultrasound channel data motivate the use of specific nonlinear activation functions. While rectified linear units (ReLU) are typically used in image processing, popular for their sparsifying nature and ability to avoid vanishing gradients due to their positive unbounded output, it inherently causes many

“dying nodes” (neurons that do no longer update since their gradient is zero) for ultrasound channel data, as a ReLU does not preserve (the abundant) negative values. To circumvent this, a hyperbolic tangent function could be used. Unfortunately, the large dynamic range of ultrasound signals makes it difficult to be in the “sweet spot,” where gradients are sufficiently large, thereby avoiding vanishing gradients during backpropagation across multiple layers.

A powerful alternative that is, by nature, unbounded and preserves both positive and negative values is the class of concatenated ReLUs [51]. A particular case is the anti-rectifier function

$$f(\mathbf{x}) = \begin{pmatrix} \left[\frac{\mathbf{x} - \bar{\mathbf{x}}}{\|\mathbf{x} - \bar{\mathbf{x}}\|_2} \right]_+ \\ \left[-\frac{\mathbf{x} - \bar{\mathbf{x}}}{\|\mathbf{x} - \bar{\mathbf{x}}\|_2} \right]_+ \end{pmatrix} \quad (2)$$

where $[\cdot]_+ = \max(\cdot, 0)$ is the positive part operator, \mathbf{x} is a vector containing the linear responses of all neurons (before activation) at a particular layer, and $\bar{\mathbf{x}}$ is its mean value across all those neurons. The anti-rectifier does not suffer from vanishing gradients, nor does it lead to dying nodes for negative values, yet provides the nonlinearity that facilitates learning complex models and representations. This dynamic-range preserving activation scheme is, therefore, well suited for processing RF or IQ-demodulated ultrasound channel data and is also used for the results presented in Fig. 2. These advantages come at the cost of a higher computational complexity compared to a standard ReLU activation.

When training a neural-network-based ultrasound beamforming algorithm, it is important to consider the impact of subsequent signal transformations in the processing chain. In particular, envelope-detected beamformed signals typically undergo significant dynamic range compression (e.g., through a logarithmic transformation) to

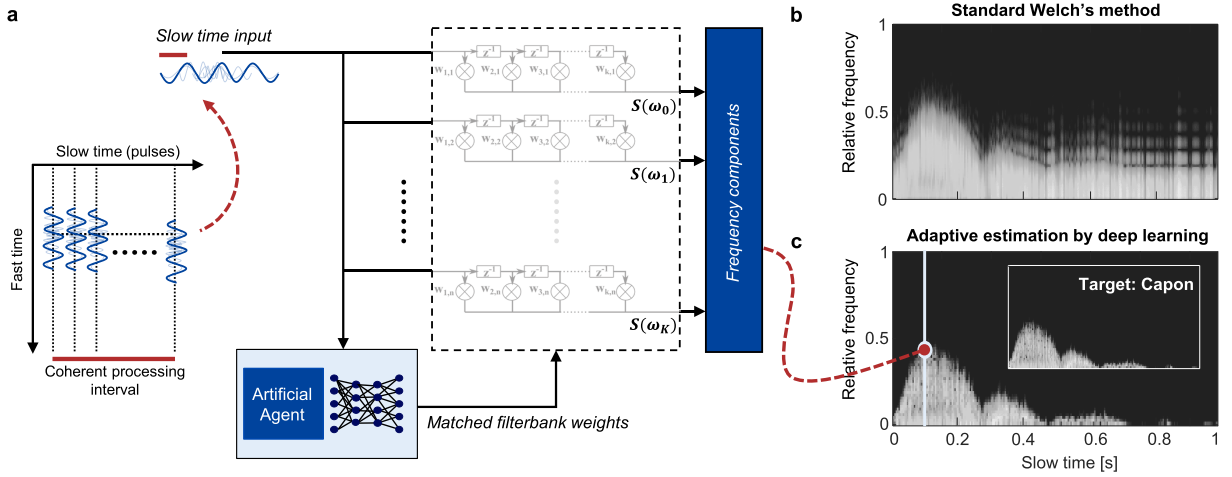


Fig. 3. Adaptive spectral Doppler processing using deep learning, displaying (a) illustrative overview of the method, comprising an artificial agent that adaptively sets the optimal matched filterbank weights according to the input data, and (b) and (c) spectral estimates using Welch’s method and the deep learning approach, respectively. The input was phantom data for the arteria femoralis, with spectra estimated from a coherent processing interval of 64 slow-time samples.

project the high dynamic range of the backscattered ultrasound signals onto the limited dynamic range of a display and allow for improved interpretation and diagnostics. To incorporate this aspect in the neural network’s training loss, beamforming errors can be transformed to attain a mean squared logarithmic error

$$\mathcal{L} = \|\log_{10}([\hat{\mathbf{y}}]_+) - \log_{10}([\mathbf{y}]_+)\|_2^2 + \|\log_{10}([-\hat{\mathbf{y}}]_+) - \log_{10}([-\mathbf{y}]_+)\|_2^2 \quad (3)$$

where $\hat{\mathbf{y}}$ is a vector containing the neural-network-based prediction of the beamformed responses for all pixels and \mathbf{y} contains the target beamformed signals. For our model-based adaptive beamforming solution [49], \mathbf{y} contains the MVDR beamformer outputs for each pixel, and $\hat{\mathbf{y}}$ is the corresponding set of pixel responses after application of the apodization weights calculated by the neural network.

B. Adaptive Spectral Estimation for Spectral Doppler

As mentioned in Section II, beamformed ultrasound signals are not only used to visualize anatomical information in B-mode but also permit the extraction of velocities by processing subsequent frames across slow time.

Spectral Doppler ultrasound enables measurement of blood (and tissue) velocity distributions through the generation of a Doppler spectrogram from slow-time data sequences, i.e., a series of subsequent pulse–echo snapshots. In commercial systems, spectra are estimated using the Fourier-transform-based periodogram methods, e.g., the standard Welch approach. Such techniques, however, require long observation windows (denoted as “coherent

processing intervals”) to achieve high spectral resolution and mitigate spectral leakage. This deteriorates the temporal resolution.

Data-adaptive spectral estimators alleviate the strong time–frequency resolution tradeoff, providing superior spectral estimates and resolution for a given temporal resolution [52]. The latter is determined by the coherent processing interval, which is, in turn, defined by the pulse repetition frequency and the number of slow-time snapshots required for a spectral estimate. Adaptive approaches steer away from the standard periodogram methods and rely on the content-matched filterbanks. The filter coefficients for each frequency of interest ω are adaptively tuned to, e.g., minimize signal energy while being constrained to unity frequency response. This Capon spectral estimator is given by solving [52]

$$\begin{aligned} \hat{\mathbf{w}}_\omega &= \arg \min_{\mathbf{w}_\omega} \mathbf{w}_\omega^H \mathbf{R}_y \mathbf{w}_\omega \\ \text{s.t. } &\mathbf{w}_\omega^H \mathbf{e}_\omega = 1 \end{aligned} \quad (4)$$

where \mathbf{R}_y is the covariance matrix of the (slow-time) input signal vector \mathbf{y} , and \mathbf{e}_ω is the corresponding Fourier vector. While this adaptive spectral estimator indeed improves upon standard approaches and significantly lowers the required observation window while gaining spectral fidelity, it, unfortunately, suffers from high computational complexity stemming from the need for inversion of the signal covariance matrix.

As for the MVDR beamformer (see Section III-A), we here demonstrate that neural networks can also be exploited to provide fast estimators for the optimal matched filter coefficients, acting as an artificial agent. An overview of this approach is given in Fig. 3, for

pulsed-wave phantom data for the arteria femoralis [53]. The neural network takes a beamformed slow-time RF signal as input and outputs a set of filter coefficients for each filter in the filterbank. The slow-time input signal is then passed through this filterbank to attain a spectral estimate. The neural network is trained by minimizing the mean squared logarithmic error (3) between the resulting spectrum and the output spectrum of the high-quality adaptive Capon spectral estimator. It comprised 128 four-layer fully connected subnetworks, each of those predicting the coefficients for one of the 128 filters in the filterbank. The optimization problem is then regularized by penalizing deviations from unity frequency response (4). The length of the slow-time observation window was only 64 samples, taken from a single depth sample. Compared to Welch’s periodogram-based method, adaptive spectral estimation by deep learning achieves far less spectral leakage and higher spectral resolution [see Fig. 3(b) and (c)].

Training the artificial agent is subject to similar considerations outlined in Section III-A3. First, slow-time input samples have a large dynamic range, such that a nonsaturating activation scheme is preferred, as in (2). Second, Doppler spectra are typically presented in decibels, advocating for the use of a log-transformed training loss, as in (3). Third, training is regularized by adding an additional loss to penalize predicted filterbanks that deviate from unity frequency response.

The above-mentioned approach is designed to processes uniformly sampled slow-time signals. In practice, there is a desire to expand these techniques to estimators that have the ability to cope with “gaps,” or even sparsely sampled signals, since the spectral Doppler processing is typically interleaved with B-mode imaging for navigation purposes (Duplex mode). To that end, extensions of data-adaptive estimators for periodically gapped data [54] and recovery for nested slow-time sampling [55] can be used.

C. Compressive Encodings for Tissue Doppler

From a hardware perspective, a significant challenge for the design of ultrasound devices and transducers is coping with the limited cable bandwidth and related connectivity constraints [56]. This is particularly troublesome for catheter transducers used in the interventional applications (e.g., intravascular ultrasound or intracardiac echography), where data need to pass through a highly restricted number of cables. While this is less of a concern for transducers with only few elements, the number of transducer elements have expanded greatly in recent devices to facilitate high-resolution 2-D or 3-D imaging [57]. Beyond the limited capacity of miniature devices, (future) wireless transducers will pose similar constraints on data rates [58]. Today, front-end connectivity and bandwidth challenges are addressed through, e.g., application-specific integrated circuits that perform microbeamforming [59] or simple summation of the receive signals across neighboring

elements [60] to compress the full channel data into a manageable amount, and multiplexing of the receive signals. This inherently entails information loss and typically leads to reduced image quality.

Instead of Nyquist-rate sampling of pre-beamformed and multiplexed channel data, compressive sub-Nyquist sampling methods permit reduced-rate imaging without sacrificing quality [3], [19]. After (reduced-rate) digitization, additional compression may be achieved through neural networks that serve as application-specific encoders. Advances in low-power neural edge computing may permit placing such a trained encoder at the probe, further alleviating probe-scanner communication, and a subsequent high-end decoder at the remote processor [61].

Instead of aiming at decoding the full input signals from the encoded representation, one can also envisage decoding only a specific signal or source that is to be extracted from the input. This may enable stronger compression during encoding whenever this component has more restricted entropy than the full signal. In ultrasound imaging, such signal-extracting compressive deep encoder–decoders can, e.g., be used for velocity estimation in the color Doppler [62]. Fig. 4 shows how these networks enable decoding of the tissue Doppler signals from the encoded IQ-demodulated input data acquired in an *in vivo* open-chest experiment of a porcine model, using intra-cardiac diverging-wave imaging in the right atrium at a frame rate of 474 Hz.

Here, the encoding neural network comprised a series of three identical blocks, each composed of two subsequent convolutional layers across fast- and slow time, followed by an aggregation of this processing through spatial downsampling (max pooling). The decoder had a similar, mirrored, architecture. The degree of IQ data compression achieved by the encoder can be changed by varying the number of channels (in the context of image processing often referred to as feature maps) at the latent layer. The encoder and decoder network parameters can then be learned by mimicking the phase (and therewith, velocity) estimates obtained using the well-known Kasai autocorrelator on the full input data [see Fig. 4(b)]. Interestingly, IQ compression rates as high as 32 can be achieved [see Fig. 4(c)] while retaining reasonable Doppler signal quality, yielding a relative phase root-mean-squared error (RMSE) of approximately 0.02. These errors drop when requiring lower compression rates. Higher compression rates lead to an increased degree of spatial consistency, displaying fewer spurious variations that could not be represented in the compact latent encoding.

The design of the traditional Doppler estimators involves careful optimization of the slow- and fast-time range gates, across which the estimation is performed, amounting to a tradeoff between the estimation quality and spatiotemporal resolution [22]. For many practical applications, the optimal settings not only vary across measurements and desired clinical objectives but also within a single mea-

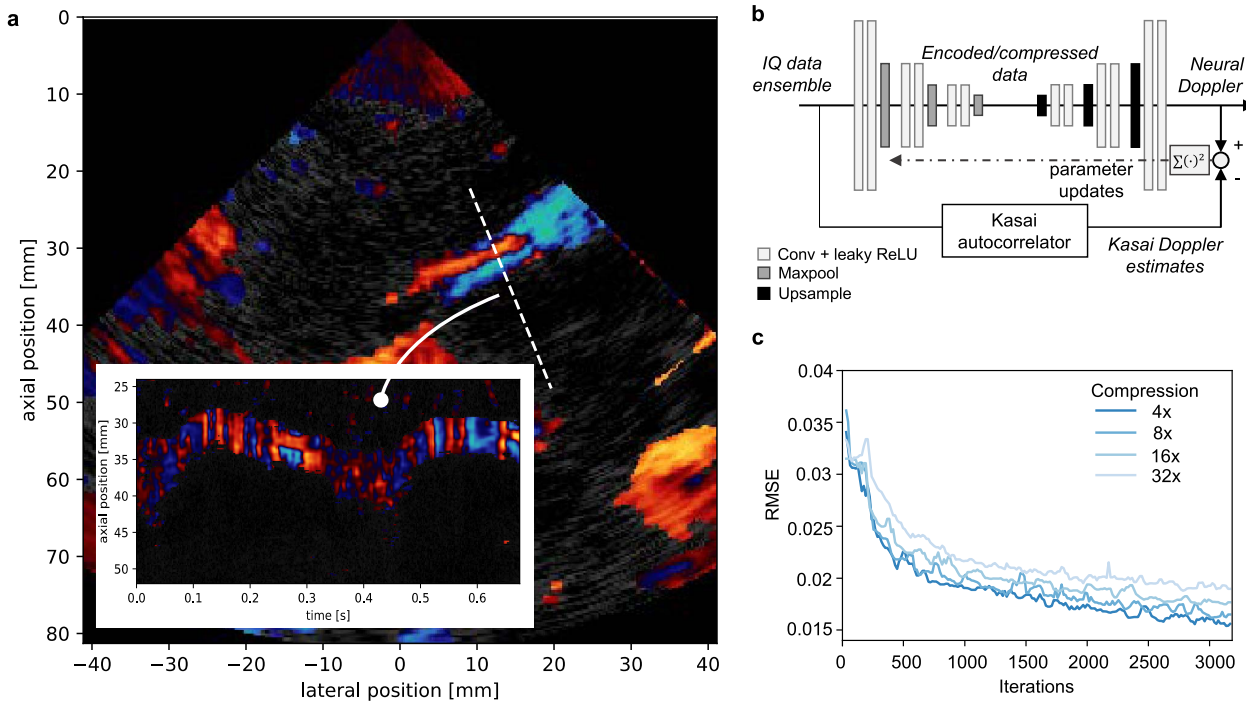


Fig. 4. (a) Tissue Doppler processing using a deep encoder–decoder network for an illustrative intracardiac ultrasound application [62], displaying the wall between the right atrium and the aorta. (b) Deep network architecture is designed to encode input IQ data into a compressed latent space via a series of convolutional layers and spatial (max) pooling operations while maintaining the functionality and performance of a typical Doppler processor (the Kasai autocorrelator [22]) using full uncompressed IQ data. (c) Convergence of the network parameters during training, showing the relative RMSEs on a test data set for four data compression factors.

surement. In contrast, a convolutional encoder–decoder network can learn to determine the effective spatiotemporal support of the given input data required for adequate Doppler encoding and prediction.

D. Unfolding Robust PCA for Clutter Suppression

An important ultrasound-based modality is CEUS [63], which allows the detection and visualization of small blood vessels. In particular, CEUS is used for imaging perfusion at the capillary level [64], [65] and for estimating different vascular properties such as relative volume, velocity, shape, and density. These physical parameters are related to different clinical conditions, including cancer [66].

The main idea behind CEUS is the use of encapsulated gas microbubbles, serving as ultrasound contrast agents (UCAs), that are injected intravenously and can flow throughout the vascular system due to their small size [67]. To visualize them, strong clutter signals originating from stationary or slowly moving tissues must be removed, as they introduce significant artifacts in the resulting images [68]. The latter poses a major challenge in ultrasonic vascular imaging, and various methods have been proposed to address it. In [69], a high-pass filtering approach was presented to remove tissue signals using finite impulse response (FIR) or infinite impulse response (IIR) filters. However, this approach is prone to failure in the presence of fast tissue motion. An alternative

strategy is the second-harmonic imaging [70] that exploits the nonlinear response of the UCAs to separate them from the tissue. This technique, however, does not remove the tissue completely, as it also exhibits a nonlinear response.

One of the most popular approaches for clutter suppression is the spatio–temporal filtering based on the singular value decomposition (SVD). This strategy has led to various techniques for clutter removal [68], [71]–[80]. SVD filtering includes collecting a series of consecutive frames, stacking them as vectors in a matrix, performing SVD of the matrix, and removing the largest singular values, assumed to be related to the tissue. Hence, a crucial step in SVD filtering is determining an appropriate threshold that discriminates between tissue- and blood-related singular values. However, the exact setting of this threshold is difficult to determine and may vary dramatically between different scans and subjects, leading to significant defects in the constructed images.

To overcome these limitations, in [81]–[83], the task of clutter removal was formulated as a convex optimization problem by leveraging a low-rank-and-sparse decomposition. Solomon *et al.* [81] then proposed an efficient deep learning solution to this convex optimization problem through an algorithm-unfolding strategy [84]. To enable explicit embedding of signal structure in the resulting network architecture, the following model for the signal after beamforming was proposed.

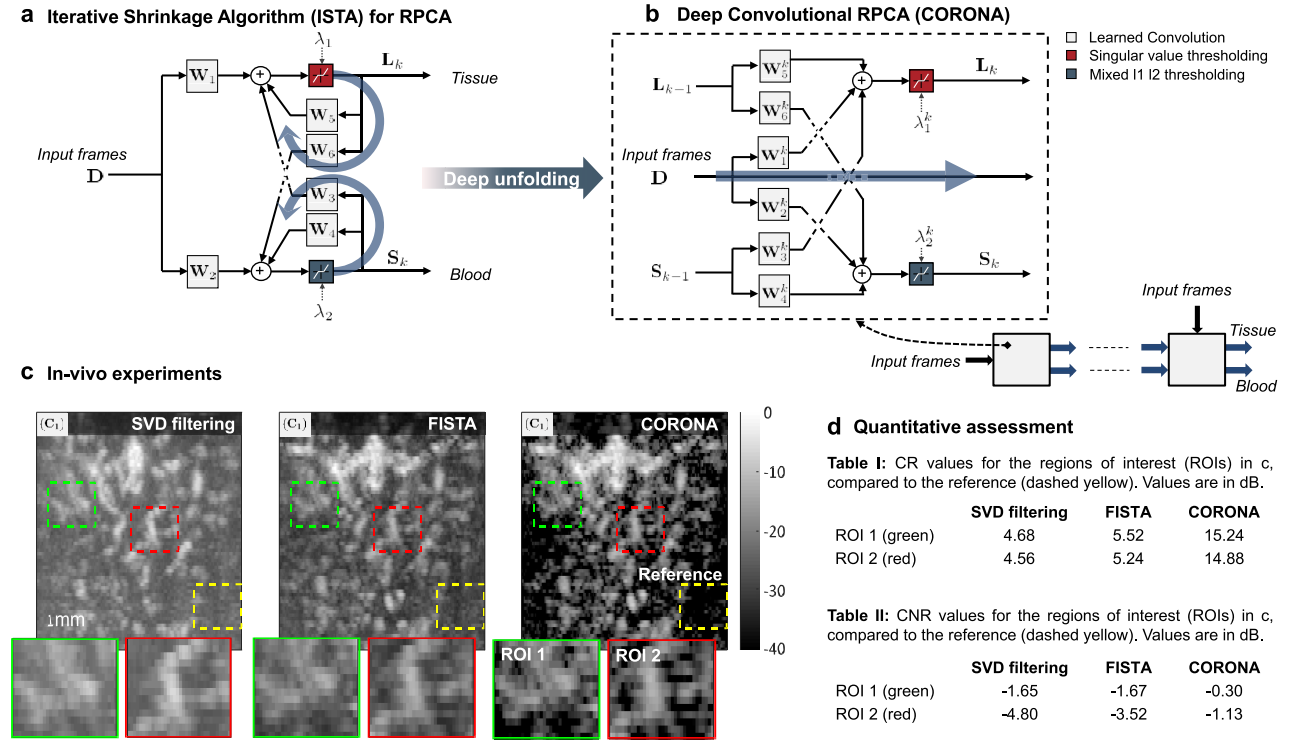


Fig. 5. (a) ISTA diagram for solving RPCA. (b) Diagram of a single layer of CORONA [81]. (c) Qualitative assessment of clutter removal performed by SVD filtering, FISTA, and CORONA, shown in panels c_1 – c_3 , respectively. Below each panel, we present enlarged views of selected areas, indicated by the green and red rectangles. (d) Quantitative assessment of clutter removal performed by the mentioned methods.

Denote the received beamformed signal at snapshot time t by $D(x, z, t)$, where (x, z) are image coordinates. Then, we may write

$$D(x, z, t) = L(x, z, t) + S(x, z, t) \quad (5)$$

where the term $L(x, z, t)$ represents the tissue and $S(x, z, t)$ is the signal stemming from the blood. Similar to SVD filtering, a series of consecutive snapshots ($t = 1, \dots, T$) is acquired and stacked as vectors into a matrix, leading to the matrix model

$$\mathbf{D} = \mathbf{L} + \mathbf{S}. \quad (6)$$

The tissue exhibits high spatio-temporal coherence; hence, the matrix \mathbf{L} is assumed to be low rank. The matrix \mathbf{S} is considered to be sparse since small blood vessels sparsely populate the image plane.

These assumptions on the rank of \mathbf{L} and the sparsity in \mathbf{S} enable formulation of the task of clutter suppression as a robust principle component analysis (RPCA) problem [85]

$$\min_{\mathbf{L}, \mathbf{S}} \frac{1}{2} \|\mathbf{D} - (\mathbf{L} + \mathbf{S})\|_F^2 + \lambda_1 \|\mathbf{L}\|_* + \lambda_2 \|\mathbf{S}\|_{1,2} \quad (7)$$

where λ_1 and λ_2 are threshold parameters. The symbol $\|\cdot\|_*$ stands for the nuclear norm that sums the

singular values of \mathbf{L} . The term $\|\cdot\|_{1,2}$ is the mixed $l_{1,2}$ norm [33], [86] that promotes the sparsity of the blood vessels along with consistency of their locations over consecutive frames. RPCA is widely used in the area of computer vision and can be solved iteratively using the fast iterative shrinkage/soft-thresholding algorithm (FISTA) [87], leading to the following update rules:

$$\begin{aligned} \mathbf{L}^{k+1} &= \mathcal{ST}_{\lambda_1/2} \left(\frac{1}{2} \mathbf{L}^k - \mathbf{S}^k + \mathbf{D} \right) \\ \mathbf{S}^{k+1} &= \mathcal{MT}_{\lambda_2/2} \left(\frac{1}{2} \mathbf{S}^k - \mathbf{L}^k + \mathbf{D} \right) \end{aligned} \quad (8)$$

where $\mathcal{MT}_\alpha(\mathbf{X})$ is the mixed $l_{1,2}$ soft-thresholding operator that applies the function $\max(0, 1 - (\alpha/\|\mathbf{x}\|))\mathbf{x}$ on each row \mathbf{x} of the input matrix \mathbf{X} . Assuming the input matrix is given by its SVD $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$, the singular value thresholding (SVT) is defined as $\mathcal{ST}_\alpha(\mathbf{X}) = \mathbf{U}\mathcal{S}_\alpha(\mathbf{\Sigma})\mathbf{V}^H$, where $\mathcal{S}_\alpha(x) = \max(0, x - \alpha)$ is applied point-wise on $\mathbf{\Sigma}$. A diagram of this iterative solution is given in Fig. 5(a).

As shown in Fig. 5(c), the iterative solution (8) outperforms SVD filtering and leads to improved clutter suppression. However, it suffers from two major drawbacks. The threshold parameters λ_1 and λ_2 need to be properly tuned, as they have a significant impact on the final result. Moreover, depending on the dynamic range between the tissue and the blood, FISTA may require many iterations to converge, thus making it impractical for real-time

imaging. This motivates the pursuit of a solution with fixed complexity, in which the threshold parameters are adjusted automatically.

Such a fixed-complexity solution can be attained through unfolding [88], [89], in which a known iterative solution is unrolled as a feedforward neural network. In this case, the iterative solution is the FISTA algorithm (8) that can be rewritten as

$$\begin{aligned}\mathbf{L}^{k+1} &= \mathcal{ST}_{\lambda_1/2}(\mathbf{W}_1\mathbf{D} + \mathbf{W}_3\mathbf{S}^k + \mathbf{W}_5\mathbf{L}^k) \\ \mathbf{S}^{k+1} &= \mathcal{MT}_{\lambda_2/2}(\mathbf{W}_2\mathbf{D} + \mathbf{W}_4\mathbf{S}^k + \mathbf{W}_6\mathbf{L}^k)\end{aligned}\quad (9)$$

where $\mathbf{W}_1 = \mathbf{W}_2 = \mathbf{I}$, $\mathbf{W}_3 = \mathbf{W}_6 = -\mathbf{I}$, and $\mathbf{W}_4 = \mathbf{W}_5 = (1/2)\mathbf{I}$. From this, the deep multilayer network takes a form, in which the k th layer is given by

$$\begin{aligned}\mathbf{L}^{k+1} &= \mathcal{ST}_{\lambda_1^k}(\mathbf{W}_1^k * \mathbf{D} + \mathbf{W}_3^k * \mathbf{S}^k + \mathbf{W}_5^k * \mathbf{L}^k) \\ \mathbf{S}^{k+1} &= \mathcal{MT}_{\lambda_2^k}(\mathbf{W}_2^k * \mathbf{D} + \mathbf{W}_4^k * \mathbf{S}^k + \mathbf{W}_6^k * \mathbf{L}^k).\end{aligned}\quad (10)$$

In (10), the matrices $(\mathbf{W}_1^k, \dots, \mathbf{W}_6^k)$ and the regularization parameters λ_1^k and λ_2^k differ from one layer to another and are learned during training. Moreover, $(\mathbf{W}_1^k, \dots, \mathbf{W}_6^k)$ were chosen to be convolution kernels, where $*$ denotes the convolution operator. The latter facilitates spatial invariance along with a notable reduction in the number of learned parameters. This results in a CNN that is specifically tailored for solving RPCA, whose nonlinearities are the soft-thresholding and SVT operators, and is termed Convolutional rObust pRincipal cOmpoNent Analysis (CORONA). A diagram of a single layer from CORONA is given in Fig. 5(b).

The training process of CORONA is performed by back-propagation in a supervised manner, leveraging both simulations, for which the true decomposition is known, and *in vivo* data, for which the decomposition of FISTA (8) is considered as the ground truth. Moreover, data augmentation is performed, and the training is done on 3-D patches extracted from the input measurements. The loss function was chosen as the sum of mean squared errors (mse)

$$\mathbf{E}(\theta) = \frac{1}{2N} \left(\sum_{i=1}^N \|\mathbf{S}_i - \hat{\mathbf{S}}_i(\theta)\|_F^2 + \|\mathbf{L}_i - \hat{\mathbf{L}}_i(\theta)\|_F^2 \right) \quad (11)$$

where $\{\mathbf{S}_i, \mathbf{L}_i\}_{i=1}^N$ are the ground truth and $\{\hat{\mathbf{S}}_i, \hat{\mathbf{L}}_i\}_{i=1}^N$ are the network's outputs. The learned parameters are denoted by $\theta = \{\mathbf{W}_1^k, \dots, \mathbf{W}_6^k, \lambda_1^k, \lambda_2^k\}_{k=1}^K$, where K is the number of layers. Backpropagation through the SVD was done using PyTorch's Autograd function [90].

Fig. 5 shows how CORONA effectively suppresses clutter on CEUS scans of two rat's brains, outperforming SVD filtering and RPCA through FISTA (8). The recovered CEUS (blood) signals are given in Fig. 5(c), including the enlarged views of regions of interest. Visually judging, FISTA achieves moderately better

contrast than SVD filtering, while CORONA outperforms both approaches by a large margin. For a quantitative comparison, the CNR and the contrast ratio (CR) were assessed, defined as

$$\text{CNR} = \frac{|\mu_s - \mu_b|}{\sqrt{\sigma_s^2 + \sigma_b^2}} \quad \text{CR} = \frac{\mu_s}{\mu_b} \quad (12)$$

where μ_s and σ_s^2 are the mean and variance of the regions of interest in Fig. 5(c), and μ_b and σ_b^2 are the mean and variance of the noisy reference area indicated by the yellow box. In both metrics, higher values imply higher CRs, which suggests better noise suppression. FISTA obtained slightly better performance than SVD filtering (CR \approx 4.6 dB and \approx 5.4 dB, respectively), and CORONA outperformed both (CR \approx 15 dB). In most cases, the performance of CORONA was about an order of magnitude better than that of SVD. Thus, combining a model for the separation problem with a data-driven approach leads to improved separation of UCA and tissue signals, together with noise reduction compared to the popular SVD approach.

The complexity of all three methods is governed by the singular-value decomposition that requires $O(MN^2)$ FLOPS for an $M \times N$ matrix, where $M \geq N$. However, FISTA may require thousands of iterations, i.e., thousands of such SVD operations. Hence, FISTA for RPCA is computationally significantly heavier than regular SVD-filtering. On the other hand, for CORONA, up to ten layers were shown to be sufficient (i.e., up to ten SVD operations), therewith offering a dramatic increase in performance at the expense of only a moderate increase in complexity. All three methods can benefit from using inexact decompositions that exhibit reduced computational load, such as the truncated SVD and randomized SVD.

IV. DEEP LEARNING FOR SUPER-RESOLUTION

A. Ultrasound Localization Microscopy

While the above-described advances in front-end ultrasound processing can boost resolution, suppress clutter, and drastically improve tissue contrast, the attainable resolution of ultrasonography remains fundamentally limited by wave diffraction, i.e., the minimum distance between separable scatters is half a wavelength. Simply increasing the transmit frequency to shorten the wavelength, unfortunately, comes at the cost of reduced penetration depth since higher frequencies suffer from stronger absorption compared to waves with a higher wavelength. This tradeoff between resolution and penetration depth particularly hampers deep high-resolution microvascular imaging, being a cornerstone for many diagnostic applications.

Recently, this tradeoff was circumvented by the introduction of ULM [91], [92]. ULM leverages principles that formed the basis for the Nobel-prize-winning concept from optics of super-resolution fluorescence microscopy

and adapts these to ultrasound imaging: if individual point sources are well isolated from diffraction-limited scans, and their centers subsequently precisely pinpointed on a subdiffraction grid, then the accumulation of many such localizations over time yields a super-resolved image. In optics, stochastic “blinking” of subsets of fluorophores is exploited to provide such sparse point sources. In ULM, intravascular lipid-shelled gas microbubbles fulfill this role [93]. This approach permits achieving a resolution that is up to ten times smaller than the wavelength [8].

Since the fidelity of ULM depends on the number of localized microbubbles and the localization accuracy, it gives rise to a new tradeoff that balances the required microbubble sparsity for accurate localization and acquisition time. To achieve the desired signal sparsity for straightforward isolation of the backscattered echoes, ULM is typically performed using a very diluted solution of microbubbles. On regular ultrasound systems, this constraint leads to tediously long acquisition times (on the order of hours) to cover the full vascular bed. Using an ultrafast plane-wave ultrasound system rather than regular scanning, Errico *et al.* [8] performed ultrafast ULM (uULM) in a rat’s brain. Empowered by high frame rates (500 frames/s), the acquisition time was lowered to minutes instead of hours. Ultrafast imaging indeed enables taking many snapshots of individual microbubbles, as they transport through the vasculature, thereby facilitating very high-fidelity reconstruction of the larger vessels. Nevertheless, mapping the full capillary bed remains dictated by the requirement of microbubbles to pass through each of the capillaries. As such, long acquisitions of tens of minutes are required, even with uULM [94]. To boost the achieved coverage in a given time-span, methods that enable the use of higher concentrations can be leveraged [32], [33], [95]–[97].

B. Exploiting Signal Structure

To strongly relax the constraints on microbubble concentration and therewith cover more vessels in a shorter time, standard ULM can be extended by incorporating knowledge of the measured signal structure, in particular, its sparsity in a transform domain. To that end, a received contrast-enhanced image frame can be modeled as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w} \quad (13)$$

where \mathbf{x} is a vector that describes the sparse microbubble distribution on a high-resolution image grid, \mathbf{y} is the vectorized image frame of the ultrasound sequence, \mathbf{A} is the measurement matrix, where each column of \mathbf{A} is the point-spread-function shifted by a single pixel on the high-resolution grid, and \mathbf{w} is a noise vector.

Leveraging this signal prior, i.e., assuming that the microbubble distribution is sparse on a sufficiently high-resolution grid (or, the number of nonzero entries in \mathbf{x} is low), we can formulate the following ℓ_1 -regularized

inverse problem:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1 \quad (14)$$

where λ is a regularization parameter that weighs the influence of $\|\mathbf{x}\|_1$.

Equation (14) may be solved using a numerical proximal gradient scheme, such as FISTA [87]. We will discuss this FISTA-based solution in Section IV-C2. After estimating \mathbf{x} for each frame, the estimates are summed across all frames to yield the final super-resolution image.

Beyond sparsity on a frame-by-frame basis, signal structure may also be leveraged across multiple frames. To that end, a multiple-measurement vector model [98] and its structure in a transformed domain can be considered, e.g., by assuming that a temporal stack of frames \mathbf{x} is sparse in the temporal correlation domain [32], [33]. Considering the temporal dimension, sparse recovery may be improved by exploiting the motion of microbubbles, allowing the application of a prior on the spatial microbubble distribution through the Kalman tracking [99].

Exploiting signal structure through sparse recovery indeed enables improved localization precision and recall for high microbubble concentrations [95], [97]. Unfortunately, proximal gradient schemes, such as FISTA, typically require numerous iterations to converge (yielding a very time-consuming reconstruction process), and their effectiveness is strongly dependent on careful tuning of the optimization parameters (e.g., λ and the step size). In addition, the linear model in (13) is an approximation of what is actually a nonlinear relation between the microbubble distribution and the resulting beamformed and envelope-detected image frame. While this approximation is valid for microbubbles that are sufficiently far apart, the significant image-domain implications of the RF interference patterns of very closely spaced microbubbles cannot be neglected.

C. Deep Learning for Fast High-Fidelity Sparse Recovery

1) *Encoder–Decoder Architectures*: In pursuit of fast and robust sparse recovery for the nonlinear measurement model, we leveraged deep learning to solve the complex inverse problem based on adequate simulations of the forward problem [95], [96]. This data-driven approach, named deep-ULM, harnesses a fully convolutional neural network to map a low-resolution input image containing many overlapping microbubble signals, to a high-resolution sparse output image, in which the pixel intensities reflect recovered backscatter levels. This process is illustrated in Fig. 6(a). The network comprises an encoder and a decoder, with the former expressing input frames in a latent representation and the latter decoding such representation into a high-resolution output. The encoder is composed of a contracting path of three blocks, each block consisting of two successive 3×3 convolution

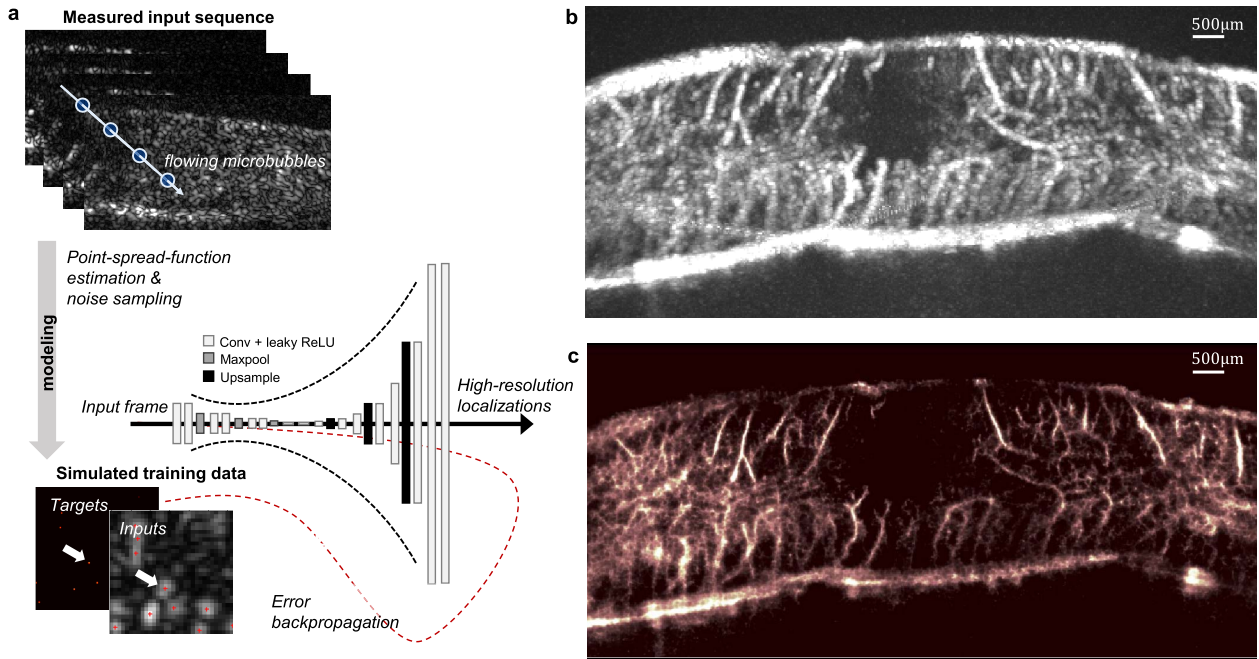


Fig. 6. (a) Fast ULM through deep learning (deep-ULM) [95], [96], using a convolutional neural network to map low-resolution CEUS frames to highly resolved sparse localizations on an eight times finer grid. The network is trained using realistic simulations of the corresponding ultrasound acquisitions, incorporating a point-spread-function estimate, the modulation frequency, pixel spacing, and background noise as sampled from real data sets. (b) Standard maximum intensity projection across a sequence of frames for a rat's spinal cord. (c) Corresponding deep-ULM reconstruction.

layers and one 2×2 max-pooling operation. This is followed by two 3×3 convolutional layers and a dropout layer that randomly disables nodes with a probability of 0.5 to mitigate overfitting. The subsequent decoder also consists of three blocks; the first two blocks encompassing two 5×5 convolution layers, of which the second has an output stride of 2, followed by 2×2 nearest-neighbor up-sampling. The last block consists of two convolution layers, of which the second again has an output stride of 2, preceding another 5×5 convolution that maps the feature space to a single-channel image through a linear activation function. All other activation functions in the network were leaky ReLUs [100]. The full deep encoder–decoder network [see Fig. 6(a)] effectively scales the input image dimensions up by a factor 8 and provides a powerful model that has the capacity to learn the sparse decoding problem while yielding simultaneous denoising through the compact latent space.

The network is trained on simulations of CEUS acquisitions, using an estimate of the real system point-spread-function, the RF modulation frequency, and pixel spacing. Noise, clutter, and artifacts were included by randomly sampling from real measurements across frames, in which no microbubbles are present. Similar to [101], we adopt a specific loss function that acts as a surrogate for the real localization error

$$\mathcal{L}(\mathbf{Y}, \mathbf{X}_t | \theta) = \|f(\mathbf{Y}|\theta) - \mathbf{G}(\sigma) * \mathbf{X}_t\|_2^2 + \gamma \|f(\mathbf{Y}|\theta)\|_1 \quad (15)$$

where \mathbf{Y} and \mathbf{X}_t are the low-resolution input and sparse super-resolution target frames, respectively, $f(\mathbf{Y}|\theta)$ is the nonlinear neural network function, and $\mathbf{G}(\sigma)$ is an isotropic Gaussian convolution kernel. Jointly, the ℓ_1 penalty that acts on the reconstructions and the kernel $\mathbf{G}(\sigma)$ that operates on the targets yield a loss function that increases when the reconstructed images exhibit less sparsity and when the Euclidean distances between the localizations and the targets become larger. We note that the selection of the relative weighting of this sparsity penalty by γ is less critical than the thresholding parameter λ adopted in the sparse recovery problem (14) since the measurement model \mathbf{A} (characterized by the point-spread-function) exhibits a much smaller bandwidth than $\mathbf{G}(\sigma)$ for low values of σ as adopted here. Consequently, the degree of bandwidth extension necessary to yield sparse outputs is less in the latter case.

Fig. 6(c) displays the super-resolution ultrasound reconstruction of a rat's spinal cord [102], qualitatively showing how deep-ULM achieves a significantly higher resolution and contrast than the diffraction-limited maximum intensity projection image [see Fig. 6(b)]. Deep-ULM achieves a resolution of about 20–30 μm , being a 4–5-fold improvement compared to standard imaging with the adopted linear 15-MHz transducer [95]. In terms of speed, recovery on a 4096×1328 grid takes roughly 100 ms/frame using GPU acceleration, making it about four orders of magnitude faster than a Fourier-domain

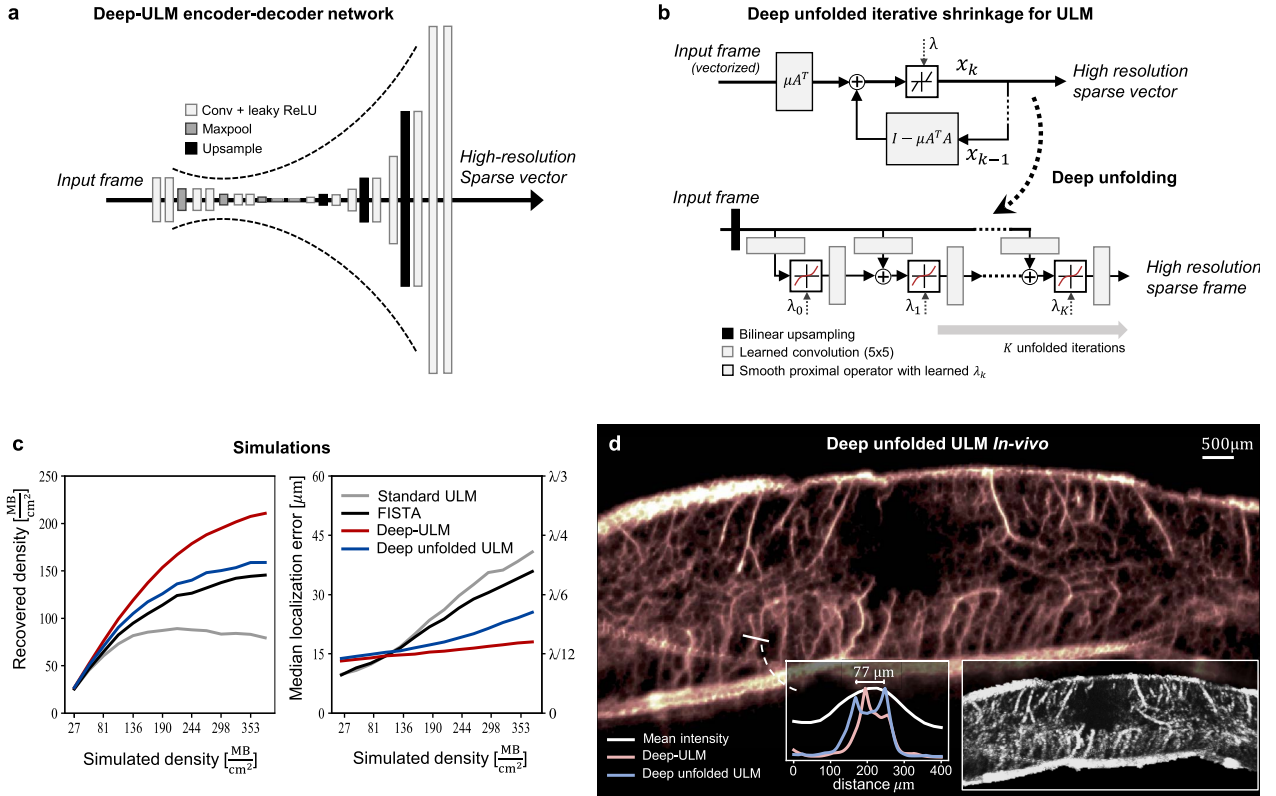


Fig. 7. (a) Deep encoder–decoder architecture used in Deep-ULM [95], [96]. (b) Deep unfolded ULM architecture obtained by unfolding the ISTA scheme, as shown in Section IV-C2. (c) Performance comparison of standard ULM, sparse-recovery, deep-ULM, and deep unfolded ULM on simulations. (d) Deep unfolded ULM for super-resolution vascular imaging of a rat’s spinal cord. Both deep learning approaches outperform the other methods. While Deep-ULM shows a higher recall and slightly lower localization error as compared to deep unfolded ULM on simulation data, the latter seems to generalize better toward in vivo acquisitions, qualitatively yielding images with higher fidelity [see Fig. 6(c) for comparison].

implementation of sparse recovery through the FISTA proximal gradient scheme [33].

2) *Deep Unfolding for Robust and Fast Sparse Decoding:* While deep encoder–decoder architectures (as used in deep-ULM) serve as a general model for many regression problems and are widely used in computer vision, their large flexibility and capacity also likely make them over-parameterized for the sparse decoding problem at hand. To promote robustness by exploiting knowledge of the underlying signal structure (i.e., microbubble sparsity), we propose using a dedicated and more compact network architecture that borrows inspiration from the proximal gradient methods introduced in Section IV-B [87].

To do so, we first briefly describe the ISTA scheme for the sparse decoding problem in (14)

$$\mathbf{x}^{k+1} = \mathcal{T}_\lambda(\mathbf{x}^k - \mu\mathbf{A}^T(\mathbf{A}\mathbf{x}^k - \mathbf{y})) \quad (16)$$

where μ determines the step size, and $\mathcal{T}_\lambda(\mathbf{x})_i = (|x_i| - \lambda)_+ \text{sgn}(x_i)$ is the proximal operator of the ℓ_1 norm. Equation (16) is compactly written as

$$\mathbf{x}^{k+1} = \mathcal{T}_\lambda(\mathbf{W}_1\mathbf{y} + \mathbf{W}_2\mathbf{x}^k) \quad (17)$$

with $\mathbf{W}_1 = \mu\mathbf{A}^T$ and $\mathbf{W}_2 = \mathbf{I} - \mu\mathbf{A}^T\mathbf{A}$. Similar to our approach to robust PCA in Section III-D, we can unfold this recurrent structure into a K -layer feedforward neural network, as in LISTA (“learning ISTA”) [88], with each layer consisting of trainable convolutions \mathbf{W}_1^k and \mathbf{W}_2^k , along with a trainable shrinkage parameter λ^k . This enables learning a highly efficient fixed-length iterative scheme for fast and robust ULM, with an optimal set of kernels and parameters per iteration, which we term deep unfolded ULM. Different than LISTA, we avoid vanishing gradients in the “dead zone” of the proximal soft-thresholding operator \mathcal{T}_λ , by replacing it by a smooth sigmoid-based soft-thresholding operation [103]. An overview of this approach is given in Fig. 7(b), contrasting this dedicated sparse-decoding-inspired solution with a general deep encoder–decoder network architecture in Fig. 7(a). Both networks are trained on the same, synthetically generated, data.

Tests on synthetic data show that both deep learning methods significantly outperform standard ULM and sparse decoding through FISTA for high microbubble concentrations [see Fig. 7(c)]. On such simulations, the deep encoder–decoder used in deep-ULM yields higher recall and lower localization errors compared to the deep

unfolded ULM. Interestingly, when applying the trained networks to *in vivo* ultrasound data, we instead observe that deep unfolded ULM yields super-resolution images with higher fidelity. Thus, it is capable of translating much better toward real acquisitions than the large deep encoder–decoder network [see Figs. 6(c) and 7(d) for comparison].

Our ten-layer deep unfolded ULM comprising 5×5 convolutional kernels has much fewer parameters (merely 506, compared to almost 700 000 for the encoder–decoder scheme), therefore exhibiting a drastically lower memory footprint and reduced power consumption, in addition to achieving higher inference rates. The encoder–decoder approach requires over four million FLOPS for mapping a low-resolution patch of 16×16 pixels into a super-resolution patch of 128×128 pixels. The unfolded ISTA architecture is much more efficient, requiring just over 1000 FLOPS.

The lower number of trainable parameters may also explain the improved robustness and better generalization toward real data compared to its over-parameterized counterpart. On the other hand, complex image artifacts, such as the strong bone reflections visible in the bottom left of Fig. 7(d), remain more prominent using the compact unfolding scheme.

V. OTHER APPLICATIONS OF DEEP LEARNING IN ULTRASOUND

While this article predominantly focuses on deep learning strategies for ultrasound-specific receive processing methods along the imaging chain, the initially most thriving application of deep learning in ultrasound was spurred by computer vision: automated analysis of the images obtained with traditional systems [104]. Such image analysis methods aim at dramatically accelerating (and potentially improving) current clinical diagnostics.

A classic application of ultrasonography lies in prenatal screening, where fetal growth and development are monitored to identify possible problems and aid diagnosis. These routine examinations can be complex and cumbersome, requiring years of training to swiftly identify the scan planes and structures of interest. Baumgartner *et al.* [105] effectively leverage deep learning to drastically simplify this procedure, enabling real-time detection and localization of standard fetal scan planes in freehand ultrasound. Similarly, in [106] and [107], deep learning was used to accelerate echocardiographic exams by automatically recognizing the relevant standard views for further analysis, even permitting automated myocardial strain imaging [108]. In [109], CNN was trained to perform thyroid nodule detection and recognition. Similar applications of deep learning include automated identification and segmentation of tumors in breast ultrasound [110]–[112], localization of clinically relevant B-line artifacts in lung ultrasonography [113], and real-time segmentation of anatomical zones on

transrectal ultrasound (TRUS) scans [114]. Hu *et al.* [115] show how such anatomical landmarks and boundaries can be exploited by a deep neural network to attain accurate voxel-level registration of TRUS and MRI.

Beyond these computer-vision applications, other learning-based techniques aim at extracting relevant medium parameters for tissue characterization. Among such approaches is data-driven elasticity imaging [116], [117]. In these works, the authors propose neural-network-based models that produce spatially varying linear elastic material properties from force–displacement measurements, free from prior assumptions on the underlying constitutive models or material properties. In [118], a deep convolutional neural network is used for speed-of-sound estimation from the (single-sided) B-mode channel data. Vishnevskiy *et al.* [119] address the problem by introducing an unfolding strategy to yield a dedicated network based on the iterative wave reflection tracking algorithm. The ability to measure the speed of sound not only permits tissue characterization but also adequate refraction-correction in beamforming.

VI. DISCUSSION AND FUTURE PERSPECTIVES

Over the past years, deep learning has revolutionized a number of domains, spurring breakthroughs in computer vision, natural language processing, and beyond. In this article, we aimed to signify the potential that this powerful approach carries when leveraged in ultrasound image and signal reconstruction. We argue and show that deep learning methods profit considerably when integrating signal priors and structure, embodied by the proposed deep unfolding schemes for clutter suppression and super-resolution imaging, and the learned beamforming approaches. In addition, several ultrasound-specific considerations regarding suitable activation and loss functions were given.

We designed and showcased a number of independent building blocks, with trained artificial agents and neural signal processors dedicated to distinct applications. Some of the presented methods operate on images (see Sections III-D and IV) or IQ data (see Section III-C), while others process channel data directly (see Sections III-A and III-B). A full processing chain may easily comprise a number of such components, which can be optimized holistically. This proposition enables imaging chains that are dedicated to the application and fully adaptive.

Designing neural networks that can efficiently process channel data in real time comes with a number of challenges. First, in contrast to images, channel data has a very large dynamic range and is RF modulated. This makes typical activation functions as used in image analysis (often ReLUs or hyperbolic tangents) less suited. In Section III-A3, we argue that the class of concatenated

ReLUs provides a possible alternative. Second, channel data is extremely large, in particular, for large arrays or matrix transducers and when sampled at the Nyquist rate. This may be alleviated significantly by leveraging sub-Nyquist sampling schemes [3], [14], [15], [17], [55], permitting high-end processing of low-rate channel data after (wireless) transfer to a remote (or cloud) processor. Such a new scheme, with a wireless probe that streams low-rate channel data for subsequent deep learning in the cloud, would open up many new possibilities for intelligent image formation and advanced processing in ultrasonography.

Deep learning typically relies on vast amounts of training data. Although several approaches to make learning more data-efficient and robust have been discussed throughout this article, a significant amount of data is still required. In the framework of supervised learning, training data typically consist of input data and desired targets. What these targets are and how they should be obtained depends on the application and goal. Sometimes, it is, for instance, desirable to mimic an existing high-performance algorithm that is too complex and costly to implement in real time. Examples of this are the adaptive beamforming and spectral Doppler applications described in Sections III-A and III-B, respectively. At other times, training data may only be obtainable through simulations or measurements on well-characterized *in vitro* phantoms. In such cases, the performance of a deep learning algorithm on *in vivo* data stands or falls with the realism of these training data and its coverage of the real-world data distribution. As shown in Section IV-C2, leveraging structural signal priors in the network architecture strongly aids generalization beyond simulations.

Once trained, inference can be fast through the exploitation of high-performance GPUs. While advanced high-end imaging systems may be equipped with GPUs to facilitate the deployment of deep neural networks at the remote processor, FPGAs or ASICs may be more appropriate for resource-limited low-power settings [120]. In the consumer market, small neural- and tensor-processing units (NPU and TPU, respectively) are enabling neural network inference at the edge [121]—one can envisage a similar paradigm for the front-end ultrasound processing. As such, the relevance of designing compact and efficient neural networks for memory-constrained (edge) settings is considerable and becomes particularly relevant

for miniature and highly portable ultrasound systems, where memory size, inference speed, and network bandwidth are all strictly constrained. This may be achieved by favoring (multiple) artificial agents that have very specific and well-defined tasks (see Sections III-A and III-B), as opposed to a single highly complex end-to-end deep neural network. We also showed that embedding signal priors in neural architectures permits drastically reduced memory footprints. In that context, the difference between a deep convolutional encoder–decoder network (no prior) and a deep unfolded ISTA network (structural sparsity prior) is illustrative, where the former consists of almost 700 000 parameters and the latter can perform super-resolution recovery with just over 500. Additional strategies to condense large models include knowledge distillation [122] and parameter pruning, as well as weight quantization [123].

Once deployed in the field, artificial agents in next-generation ultrasound systems ultimately should be able to embrace the vastness of data at their disposal to continuously learn throughout their “lifetime.” To that end, unsupervised or self-supervised learning become increasingly relevant [124]. This holds true for many artificial intelligence applications and extends beyond ultrasound imaging.

The promise that deep learning holds for ultrasound imaging is significant; it may spur a paradigm shift in the design of ultrasound systems, where smart wireless probes facilitated by sub-Nyquist and neural edge computing are connected to the cloud and with AI-driven imaging modes and algorithms that are dedicated to specific applications. Empowered by deep learning, next-generation ultrasound imaging may become a much stronger modality with devices that continuously learn to provide better images and clinical insight, leading to improved and more widely accessible diagnostics through cost-effective, highly portable, and intelligent imaging. ■

Acknowledgments

The authors would like to thank B. Luijten, F. de Bruijn and H. Schmeitz for their contribution to the adaptive beamforming and spectral Doppler applications. They would also like to thank M. Bruce and Z. Khaing for acquiring the spinal cord data used to evaluate the super-resolution algorithms.

REFERENCES

- [1] T. L. Szabo, *Diagnostic Ultrasound Imaging: Inside Out*. New York, NY, USA: Academic, 2004.
- [2] J. M. Baran and J. G. Webster, “Design of low-cost portable ultrasound systems: Review,” in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Sep. 2009, pp. 792–795.
- [3] T. Chernyakova and Y. Eldar, “Fourier-domain beamforming: The path to compressed ultrasound imaging,” *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 61, no. 8, pp. 1252–1267, Aug. 2014.
- [4] J. Provost et al., “3D ultrafast ultrasound imaging *in vivo*,” *Phys. Med. Biol.*, vol. 59, no. 19, p. L1, 2014.
- [5] M. Tanter and M. Fink, “Ultrafast imaging in biomedical ultrasound,” *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 61, no. 1, pp. 102–119, Jan. 2014.
- [6] J. Bercoff, M. Tanter, and M. Fink, “Supersonic shear imaging: A new technique for soft tissue elasticity mapping,” *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 51, no. 4, pp. 396–409, Apr. 2004.
- [7] C. Deme ne et al., “Spatiotemporal clutter filtering of ultrafast ultrasound data highly increases Doppler and fullrange sensitivity,” *IEEE Trans. Med. Imag.*, vol. 34, no. 11, pp. 2271–2285, Nov. 2015.
- [8] C. Errico et al., “Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging,” *Nature*, vol. 527, no. 7579, p. 499, 2015.
- [9] A. Besson, D. Perdios, M. Arditi, Y. Wiaux, and J.-P. Thiran, “Compressive multiplexing of ultrasound signals,” in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Oct. 2018, pp. 1–4.
- [10] Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2015.
- [11] K. Gedalyahu, R. Tur, and Y. C. Eldar,

- “Multichannel sampling of pulse streams at the rate of innovation,” *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1491–1504, Apr. 2011.
- [12] R. Tur, Y. C. Eldar, and Z. Friedman, “Innovation rate sampling of pulse streams with application to ultrasound imaging,” *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1827–1842, Apr. 2011.
- [13] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
- [14] N. Wagner, Y. C. Eldar, A. Feuer, G. Danin, and Z. Friedman, “Xampling in ultrasound imaging,” *Proc. SPIE*, vol. 7968, p. 796818, Mar. 2011.
- [15] N. Wagner, Y. C. Eldar, and Z. Friedman, “Compressed beamforming in ultrasound imaging,” *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4643–4657, Sep. 2012.
- [16] M. Mishali, Y. C. Eldar, and A. E. Jeron, “Xampling: Signal acquisition and processing in union of subspaces,” *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4719–4734, Oct. 2011.
- [17] M. Mishali, Y. C. Eldar, O. Dounaevsky, and E. Shoshan, “Xampling: Analog to digital at sub-Nyquist rates,” *IET Circuits, Devices Syst.*, vol. 5, no. 1, pp. 8–20, Jan. 2011.
- [18] T. Michaeli and Y. C. Eldar, “Xampling at the rate of innovation,” *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1121–1133, 2012.
- [19] T. Chernyakova et al., “Fourier-domain beamforming and structure-based reconstruction for plane-wave imaging,” *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 65, no. 10, pp. 1810–1821, Oct. 2018.
- [20] A. Burshtein, M. Birk, T. Chernyakova, A. Eilam, A. Kempinski, and Y. C. Eldar, “Sub-Nyquist sampling and Fourier domain beamforming in volumetric ultrasound imaging,” *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 63, no. 5, pp. 703–716, May 2016.
- [21] A. Lahav, T. Chernyakova, and Y. C. Eldar, “FoCUS: Fourier-based coded ultrasound,” *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 64, no. 12, pp. 1828–1839, Dec. 2017.
- [22] T. Loupas, J. T. Powers, and R. W. Gill, “An axial velocity estimator for ultrasound blood flow imaging, based on a full evaluation of the Doppler equation by means of a two-dimensional autocorrelation approach,” *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 42, no. 4, pp. 672–688, Jul. 1995.
- [23] P. D. Welch, “The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms,” *IEEE Trans. Audio Electroacoust.*, vol. AE-15, no. 2, pp. 70–73, Jun. 1967.
- [24] K. Nightingale, “Acoustic radiation force impulse (ARFI) imaging: A review,” *Current Med. Imag. Rev.*, vol. 7, no. 4, pp. 328–339, Nov. 2011.
- [25] R. J. G. van Sloun, R. R. Wildeboer, H. Wijkstra, and M. Mischi, “Viscoelasticity mapping by identification of local shear wave dynamics,” *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 64, no. 11, pp. 1666–1673, Nov. 2017.
- [26] B. B. Goldberg, J.-B. Liu, and F. Forsberg, “Ultrasound contrast agents: A review,” *Ultrasound Med. Biol.*, vol. 20, no. 4, pp. 319–333, 1994.
- [27] R. J. G. van Sloun, L. Demi, A. W. Postema, J. J. de la Rosette, H. Wijkstra, and M. Mischi, “Ultrasound-contrast-agent dispersion and velocity imaging for prostate cancer localization,” *Med. Image Anal.*, vol. 35, pp. 610–619, Jan. 2017.
- [28] M. F. Hamilton et al. *Nonlinear Acoustics*, vol. 237. San Diego, CA, USA: Academic, 1998.
- [29] Y. Desailly, A.-M. Tissier, J.-M. Correias, F. Wintzenrieth, M. Tanter, and O. Couture, “Contrast enhanced ultrasound by real-time spatiotemporal filtering of ultrafast images,” *Phys. Med. Biol.*, vol. 62, no. 1, p. 31, 2016.
- [30] O. M. Viessmann, R. J. Eckersley, K. Christensen-Jeffries, M. X. Tang, and C. Dunsby, “Acoustic super-resolution with ultrasound and microbubbles,” *Phys. Med. Biol.*, vol. 58, no. 18, p. 6447, Sep. 2013.
- [31] M. A. O’Reilly and K. Hynynen, “A super-resolution ultrasound method for brain vascular mapping,” *Med. Phys.*, vol. 40, no. 11, 2013, Art. no. 110701.
- [32] A. Bar-Zion, C. Tremblay-Darveau, O. Solomon, D. Adam, and Y. C. Eldar, “Fast vascular ultrasound imaging with enhanced spatial resolution and background rejection,” *IEEE Trans. Med. Imag.*, vol. 36, no. 1, pp. 169–180, Jan. 2017.
- [33] A. Bar-Zion, O. Solomon, C. Tremblay-Darveau, D. Adam, and Y. C. Eldar, “SUSHI: Sparsity-based ultrasound super-resolution hemodynamic imaging,” *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 65, no. 12, pp. 2365–2380, Dec. 2018.
- [34] K. Hornik, M. Stinchcombe, and H. White, “Multilayer feedforward networks are universal approximators,” *Neural Netw.*, vol. 2, no. 5, pp. 359–366, 1989.
- [35] V. Mnih et al., “Human-level control through deep reinforcement learning,” *Nature*, vol. 518, no. 7540, p. 529, 2015.
- [36] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. Cambridge, MA, USA: MIT Press, 2016.
- [37] R. Mallart and M. Fink, “Sound speed fluctuations in medical ultrasound imaging comparison between different correction algorithms,” in *Acoustical Imaging (Acoustical Imaging Book Series)*, vol. 19. Springer, 1992, pp. 213–218.
- [38] H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. Hoboken, NJ, USA: Wiley, 2004.
- [39] S. Vedula, O. Senouf, G. Zurakhov, A. Bronstein, O. Michailovich, and M. Zibulevsky, “Learning beamforming in ultrasound imaging,” 2018, arXiv:1812.08043. [Online]. Available: <https://arxiv.org/abs/1812.08043>
- [40] D. Perdios, A. Besson, M. Arditi, and J.-P. Thiran, “A deep learning approach to ultrasound image recovery,” in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Sep. 2017, pp. 1–4.
- [41] W. Simson et al., “End-to-end learning-based ultrasound reconstruction,” 2019, arXiv:1904.04696. [Online]. Available: <https://arxiv.org/abs/1904.04696>
- [42] S. Khan, J. Huh, and J. C. Ye, “Universal deep beamformer for variable rate ultrasound imaging,” 2019, arXiv:1901.01706. [Online]. Available: <https://arxiv.org/abs/1901.01706>
- [43] A. C. Luchies and B. C. Byram, “Deep neural networks for ultrasound beamforming,” *IEEE Trans. Med. Imag.*, vol. 37, no. 9, pp. 2010–2021, Sep. 2018.
- [44] D. Hyun, L. L. Brickson, K. T. Looby, and J. J. Dahl, “Beamforming and speckle reduction using neural networks,” *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 66, no. 5, pp. 898–910, May 2019.
- [45] P. Coupé, P. Hellier, C. Kervran, and C. Barillot, “Nonlocal means-based speckle filtering for ultrasound images,” *IEEE Trans. Image Process.*, vol. 18, no. 10, pp. 2221–2229, Oct. 2009.
- [46] O. Senouf et al., “High frame-rate cardiac ultrasound imaging with deep learning,” in *Proc. Int. Conf. Med. Image Comput. Assist. Intervent (MICCAI)*. Granada, Spain: Springer, 2018, pp. 126–134.
- [47] S. Goudarzi, A. Asif, and H. Rivaz, “Multi-focus ultrasound imaging using generative adversarial networks,” in *Proc. IEEE Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 1118–1121.
- [48] A. A. Nair, T. D. Tran, A. Reiter, and M. A. L. Bell, “A generative adversarial neural network for beamforming ultrasound images: Invited presentation,” in *Proc. 53rd Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2019, pp. 1–6.
- [49] B. Luijten et al., “Deep learning for fast adaptive beamforming,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 1333–1337.
- [50] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [51] W. Shang, K. Sohn, D. Almeida, and H. Lee, “Understanding and improving convolutional neural networks via concatenated rectified linear units,” in *Proc. Int. Conf. Mach. Learn.*, 2016, pp. 2217–2225.
- [52] F. Gran, A. Jakobsson, and J. A. Jensen, “Adaptive spectral Doppler estimation,” *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 56, no. 4, pp. 700–714, Apr. 2009.
- [53] J. A. Jensen, *Estimation of Blood Velocities Using Ultrasound: A Signal Processing Approach*. Cambridge, U.K.: Cambridge Univ. Press, 1996.
- [54] P. Liu and D. Liu, “Periodically gapped data spectral velocity estimation in medical ultrasound using spatial and temporal dimensions,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2009, pp. 437–440.
- [55] R. Cohen and Y. C. Eldar, “Sparse convolutional beamforming for ultrasound imaging,” *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 65, no. 12, pp. 2390–2406, Dec. 2018.
- [56] M. W. Rashid et al., “Front-end electronics for cable reduction in intracardiac echocardiography (ICE) catheters,” in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Sep. 2016, pp. 1–4.
- [57] D. E. Dausch, K. H. Gilchrist, J. B. Carlson, S. D. Hall, J. B. Castellucci, and O. T. V. Ramm, “In vivo real-time 3-D intracardiac echo using PMUT arrays,” *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 61, no. 10, pp. 1754–1764, Oct. 2014.
- [58] A. Bar-Zion, D. Adam, M. Alessandrini, J. D’hooge, and Y. C. Eldar, “Towards sub-Nyquist tissue Doppler imaging using non-uniformly spaced stream of pulses,” in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Oct. 2015, pp. 1–4.
- [59] D. Wildes et al., “4-D ICE: A 2-D array transducer with integrated ASIC in a 10-Fr catheter for real-time 3-D intracardiac echocardiography,” *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 63, no. 12, pp. 2159–2173, Dec. 2016.
- [60] D. Bera, J. G. Bosch, M. D. Verweij, N. D. Jong, and H. J. Vos, “Dual stage beamforming in the absence of front-end receive focusing,” *Phys. Med. Biol.*, vol. 62, no. 16, p. 6631, 2017.
- [61] S. Teerapittayanon, B. McDanel, and H. T. Kung, “Distributed deep neural networks over the cloud, the edge and end devices,” in *Proc. IEEE 37th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Jun. 2017, pp. 328–339.
- [62] R. J. V. Sloun, H. Belt, K. Janse, and M. Mischi, “Learning Doppler with deep neural networks and its application to intra-cardiac echocardiography,” in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Oct. 2018, pp. 1–4.
- [63] B. Furlow, “Contrast-enhanced ultrasound,” *Radiol. Technol.*, vol. 80, no. 6, pp. 547S–561S, 2009.
- [64] N. Lassau, L. Chami, B. Benatsou, P. Peronneau, and A. Roche, “Dynamic contrast-enhanced ultrasonography (DCE-US) with quantification of tumor perfusion: A new diagnostic tool to evaluate the early effects of antiangiogenic treatment,” *Eur. Radiol. Supplements*, vol. 17, no. 6, pp. 89–98, 2007.
- [65] J. M. Hudson et al., “Dynamic contrast enhanced ultrasound for therapy monitoring,” *Eur. J. Radiol.*, vol. 84, no. 9, pp. 1650–1657, Sep. 2015.
- [66] T. Opacic et al., “Motion model ultrasound localization microscopy for preclinical and clinical multiparametric tumor characterization,” *Nature Commun.*, vol. 9, no. 1, p. 1527, 2018.
- [67] N. D. Jong, F. T. Cate, C. Lancee, J. Roelandt, and N. Bom, “Principles and recent developments in ultrasound contrast agents,” *Ultrasonics*, vol. 29, no. 4, pp. 324–330, 1991.
- [68] S. Bjærum, H. Torp, and K. Kristoffersen, “Clutter filter design for ultrasound color flow imaging,” *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 49, no. 2, pp. 204–216, Feb. 2002.
- [69] L. Thomas and A. Hall, “An improved wall filter for flow imaging of low velocity flow,” in *Proc. IEEE Ultrason. Symp.*, vol. 3, Oct. 1994, pp. 1701–1704.
- [70] P. J. A. Frinking, A. Bouakaz, J. Kirkhorn, F. J. Ten Cate, and N. de Jong, “Ultrasound contrast imaging: Current and new potential methods,” *Ultrasound Med. Biol.*, vol. 26, no. 6, pp. 965–975, Jul. 2000.
- [71] A. C. Yu and L. Lovstakken, “Eigen-based clutter filter design for ultrasound color flow imaging: A review,” *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 57, no. 5, pp. 1096–1111, May 2010.

- [72] F. W. Mauldin, F. Viola, and W. F. Walker, "Complex principal components for robust motion estimation," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 57, no. 11, pp. 2437–2449, Nov. 2010.
- [73] F. W. Mauldin, Jr., D. Lin, and J. A. Hossack, "The singular value filter: A general filter design strategy for PCA-based signal separation in medical ultrasound imaging," *IEEE Trans. Med. Imag.*, vol. 30, no. 11, pp. 1951–1964, Nov. 2011.
- [74] C. M. Gallippi, K. R. Nightingale, and G. E. Trahey, "BSS-based filtering of physiological and ARFI-induced tissue and blood motion," *Ultrasound Med. Biol.*, vol. 29, no. 11, pp. 1583–1592, 2003.
- [75] L. Lovstakken, S. Bjaerum, K. Kristoffersen, R. Haaverstad, and H. Torp, "Real-time adaptive clutter rejection filtering in color flow imaging using power method iterations," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 53, no. 9, pp. 1597–1608, Sep. 2006.
- [76] D. E. Kruse and K. W. Ferrara, "A new high resolution color flow system using an eigendecomposition-based adaptive filter for clutter rejection," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 49, no. 10, pp. 1384–1399, Oct. 2002.
- [77] C. Errico et al., "Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging," *Nature*, vol. 527, no. 7579, pp. 499–502, Nov. 2015.
- [78] P. Song, A. Manduca, J. D. Trzasko, and S. Chen, "Ultrasound small vessel imaging with block-wise adaptive local clutter filtering," *IEEE Trans. Med. Imag.*, vol. 36, no. 1, pp. 251–262, Jan. 2017.
- [79] A. J. Chee and C. Alfred, "Receiver-operating characteristic analysis of eigen-based clutter filters for ultrasound color flow imaging," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 65, no. 3, pp. 390–399, Mar. 2018.
- [80] M. Kim, Y. Zhu, J. Hedhli, L. W. Dobrucki, and M. F. Insana, "Multidimensional clutter filter optimization for ultrasonic perfusion imaging," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 65, no. 11, pp. 2020–2029, Nov. 2018.
- [81] O. Solomon et al., "Deep unfolded robust PCA with application to clutter suppression in ultrasound," 2018, *arXiv:1811.08252*. [Online]. Available: <https://arxiv.org/abs/1811.08252>
- [82] M. Ashikuzzaman, C. Belasso, C. Gauthier, and H. Rivaz, "Suppressing clutter components in ultrasound color flow imaging using robust matrix completion algorithm: Simulation and phantom study," in *Proc. IEEE Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 1–4.
- [83] M. Bayat and M. Fatemi, "Concurrent clutter and noise suppression via low rank plus sparse optimization for non-contrast ultrasound flow Doppler processing in microvasculature," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 1080–1084.
- [84] Y. Li, M. Tofghi, J. Geng, V. Monga, and Y. C. Eldar, "Deep algorithm unrolling for blind image deblurring," 2019, *arXiv:1902.03493*. [Online]. Available: <https://arxiv.org/abs/1902.03493>
- [85] R. Otazo, E. J. Candès, and D. K. Sodickson, "Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components," *Magn. Reson. Med.*, vol. 73, no. 3, pp. 1125–1136, 2014.
- [86] O. Solomon, Y. C. Eldar, M. Mutzafi, and M. Segev, "SPARCOM: Sparsity based super-resolution correlation microscopy," *SIAM J. Imag. Sci.*, vol. 12, no. 1, pp. 392–419, 2019.
- [87] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, 2009.
- [88] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proc. 27th Int. Conf. Int. Conf. Mach. Learn.*, 2010, pp. 399–406.
- [89] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.
- [90] A. Paszke et al., "Automatic differentiation in PyTorch," in *Proc. 31st Conf. Neural Inf. Process. Syst. (NIPS)*, Long Beach, CA, USA, 2017.
- [91] M. Siepmann, G. Schmitz, J. Bzyl, M. Palmowski, and F. Kiessling, "Imaging tumor vascularity by tracing single microbubbles," in *Proc. IEEE Int. Ultrason. Symp.*, Oct. 2011, pp. 1906–1909.
- [92] O. Couture, B. Besson, G. Montaldo, M. Fink, and M. Tanter, "Microbubble ultrasound super-localization imaging (MUSLI)," in *Proc. IEEE Int. Ultrason. Symp.*, Oct. 2011, pp. 1285–1287.
- [93] O. Couture, V. Hingot, B. Heiles, P. Muleki-Seya, and M. Tanter, "Ultrasound localization microscopy and super-resolution: A state of the art," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 65, no. 8, pp. 1304–1320, Aug. 2018.
- [94] V. Hingot, C. Errico, B. Heiles, L. Rahal, M. Tanter, and O. Couture, "Microvascular flow dictates the compromise between spatial resolution and acquisition time in ultrasound localization microscopy," *Sci. Rep.*, vol. 9, no. 1, p. 2456, 2019.
- [95] R. J. V. Sloun et al., "Super-resolution ultrasound localization microscopy through deep learning," 2018, *arXiv:1804.07661*. [Online]. Available: <https://arxiv.org/abs/1804.07661>
- [96] R. J. V. Sloun, O. Solomon, M. Bruce, Z. Z. Khaing, Y. C. Eldar, and M. Mischi, "Deep learning for super-resolution vascular ultrasound imaging," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 1055–1059.
- [97] R. J. V. Sloun, O. Solomon, Y. C. Eldar, H. Wijkstra, and M. Mischi, "Sparsity-driven super-resolution in clinical contrast-enhanced ultrasound," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Sep. 2017, pp. 1–4.
- [98] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–2488, Jul. 2005.
- [99] O. Solomon, R. J. V. Sloun, H. Wijkstra, M. Mischi, and Y. C. Eldar, "Exploiting flow dynamics for super-resolution in contrast-enhanced ultrasound," 2018, *arXiv:1804.03134*. [Online]. Available: <https://arxiv.org/abs/1804.03134>
- [100] B. Xu, N. Wang, T. Chen, and M. Li, "Empirical evaluation of rectified activations in convolutional network," 2015, *arXiv:1505.00853*. [Online]. Available: <https://arxiv.org/abs/1505.00853>
- [101] E. Nehme, L. E. Weiss, T. Michaeli, and Y. Shechtman, "Deep-STORM: Super-resolution single-molecule microscopy by deep learning," *Optica*, vol. 5, no. 4, pp. 458–464, 2018.
- [102] Z. Z. Khaing et al., "Contrast-enhanced ultrasound to visualize hemodynamic changes after rodent spinal cord injury," *J. Neurosurgery Spine*, vol. 29, no. 3, pp. 306–313, 2018.
- [103] X.-P. Zhang, "Thresholding neural network for adaptive noise reduction," *IEEE Trans. Neural Netw.*, vol. 12, no. 3, pp. 567–584, May 2001.
- [104] S. Liu et al., "Deep learning in medical ultrasound analysis: A review," *Engineering*, vol. 5, no. 2, pp. 261–275, Apr. 2019.
- [105] C. Baumgartner et al., "SonoNet: Real-time detection and localisation of fetal standard scan planes in freehand ultrasound," *IEEE Trans. Med. Imag.*, vol. 36, no. 11, pp. 2204–2215, Nov. 2017.
- [106] A. Madani, R. Arnaout, M. Mofrad, and R. Arnaout, "Fast and accurate view classification of echocardiograms using deep learning," *NPJ Digit. Med.*, vol. 1, no. 1, p. 6, 2018.
- [107] A. Østvik, E. Smistad, S. A. Aase, B. O. Haugen, and L. Lovstakken, "Real-time standard view classification in transthoracic echocardiography using convolutional neural networks," *Ultrasound Med. Biol.*, vol. 45, no. 2, pp. 374–384, 2019.
- [108] A. Østvik, E. Smistad, T. Espeland, E. A. R. Berg, and L. Lovstakken, "Automatic myocardial strain imaging in echocardiography using deep learning," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. DLMIA, ML-CDS* (Lecture Notes in Computer Science), vol. 11045, D. Stoyanov et al., Eds. Cham, Switzerland: Springer, 2018, pp. 309–316.
- [109] W. Song et al., "Multitask cascade convolutional neural networks for automatic thyroid nodule detection and recognition," *IEEE J. Biomed. Health Inform.*, vol. 23, no. 3, pp. 1215–1224, May 2019.
- [110] T.-C. Chiang, Y.-S. Huang, R.-T. Chen, C.-S. Huang, and R.-F. Chang, "Tumor detection in automated breast ultrasound using 3-D CNN and prioritized candidate aggregation," *IEEE Trans. Med. Imag.*, vol. 38, no. 1, pp. 240–249, Jan. 2019.
- [111] S. Y. Shin, S. Lee, I. D. Yun, S. M. Kim, and K. M. Lee, "Joint weakly and semi-supervised deep learning for localization and classification of masses in breast ultrasound images," *IEEE Trans. Med. Imag.*, vol. 38, no. 3, pp. 762–774, Mar. 2018.
- [112] M. Xian, Y. Zhang, H. D. Cheng, F. Xu, B. Zhang, and J. Ding, "Automatic breast ultrasound image segmentation: A survey," *Pattern Recognit.*, vol. 79, pp. 340–355, Jul. 2018.
- [113] R. J. G. van Sloun and L. Demi, "Deep learning for automated detection of b-lines in lung ultrasonography," *J. Acoust. Soc. Amer.*, vol. 145, no. 3, p. 1674, 2019.
- [114] R. J. V. Sloun et al., "Deep learning for real-time, automatic, and scanner-adapted prostate (Zone) segmentation of transrectal ultrasound, for example, magnetic resonance imaging–transrectal ultrasound fusion prostate biopsy," *Eur. Urol. Focus*, to be published. doi: [10.1016/j.euf.2019.04.009](https://doi.org/10.1016/j.euf.2019.04.009).
- [115] Y. Hu et al., "Weakly-supervised convolutional neural networks for multimodal image registration," *Med. Image Anal.*, vol. 49, pp. 1–13, Oct. 2018.
- [116] C. Hoerig, J. Ghaboussi, and M. F. Insana, "An information-based machine learning approach to elasticity imaging," *Biomech. Model. Mechanobiol.*, vol. 16, no. 3, pp. 805–822, 2017.
- [117] C. Hoerig, J. Ghaboussi, and M. F. Insana, "Data-driven elasticity imaging using cartesian neural network constitutive models and the autoprogressive method," *IEEE Trans. Med. Imag.*, vol. 38, no. 5, pp. 1150–1160, May 2018.
- [118] M. Feigin, D. Freedman, and B. W. Anthony, "A deep learning framework for single-sided sound speed inversion in medical ultrasound," 2018, *arXiv:1810.00322*. [Online]. Available: <https://arxiv.org/abs/1810.00322>
- [119] V. Vishnevskiy, S. J. Sanabria, and O. Goksel, "Image reconstruction via variational network for real-time hand-held sound-speed imaging," in *Machine Learning for Medical Image Reconstruction. MLMI* (Lecture Notes in Computer Science), vol. 11074, F. Knoll, A. Maier, and D. Rueckert, Eds. Cham, Switzerland: Springer, 2018, pp. 120–128.
- [120] J. Johansson, M. Gustafsson, and J. Delsing, "Ultra-low power transmit/receive ASIC for battery operated ultrasound measurement systems," *Sens. Actuators A, Phys.*, vol. 125, no. 2, pp. 317–328, 2006.
- [121] N. Jouppi, C. Young, N. Patil, and D. Patterson, "Motivation for and evaluation of the first tensor processing unit," *IEEE Micro*, vol. 38, no. 3, pp. 10–19, May/Jun. 2018.
- [122] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," 2015, *arXiv:1503.02531*. [Online]. Available: <https://arxiv.org/abs/1503.02531>
- [123] I. Hubara, M. Courbariaux, D. Soudry, R. El-Yaniv, and Y. Bengio, "Quantized neural networks: Training neural networks with low precision weights and activations," *J. Mach. Learn. Res.*, vol. 18, no. 1, pp. 6869–6898, 2017.
- [124] P. Sermanet et al., "Time-contrastive networks: Self-supervised learning from video," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 1134–1141.

ABOUT THE AUTHORS

Ruud J. G. van Sloun (Member, IEEE) received the B.Sc. and M.Sc. degrees (*cum laude*) in electrical engineering and the Ph.D. degree (*cum laude*) from the Eindhoven University of Technology, Eindhoven, The Netherlands, in 2012, 2014, and 2018, respectively.

Since then, he has been an Assistant Professor with the Department of Electrical Engineering, Eindhoven University of Technology. He is also a Visiting Professor with the Department of Mathematics and Computer Science, Weizmann Institute of Science, Rehovot, Israel. His current research interests include artificial intelligence and deep learning for front-end signal processing, model-aware deep learning, compressed sensing, ultrasound imaging, and signal and image analysis.



Regev Cohen (Graduate Student Member, IEEE) received the B.Sc. degree (*summa cum laude*) in electrical engineering from the Technion—Israel Institute of Technology, Haifa, Israel, in 2015, where he is currently working toward the Ph.D. degree.

His current research interests include theoretical aspects of signal processing, sampling theory, compressed sensing, optimization methods, sparse array design, deep learning, and advanced signal processing methods for ultrasonic imaging.

Mr. Cohen received the Meyer Foundation Excellence Award and the Elias-Perlmutter Award in 2015. In 2017, he was awarded the Israel and Debora Cederbaum Scholarship.



Yonina C. Eldar (Fellow, IEEE) received the B.Sc. degree in physics and the B.Sc. degree in electrical engineering from Tel Aviv University (TAU), Tel Aviv, Israel, in 1995 and 1996, respectively, and the Ph.D. degree in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 2002.

She was a Professor with the Department of Electrical Engineering, Technion, Haifa, Israel, where she held the Edwards Chair in Engineering. She was a Visiting Professor with Stanford University, Stanford, CA, USA. She is currently a Professor with the Department of Mathematics and Computer Science, Weizmann Institute of Science, Rehovot, Israel. She is also a Visiting Professor with MIT, a Visiting Scientist with the Broad Institute, Cambridge, MA, USA, and an Adjunct Professor with Duke University, Durham, NC, USA. She is the author of the book *Sampling Theory: Beyond Bandlimited Systems* and a coauthor of the books *Compressed Sensing* and *Convex Optimization Methods in Signal Processing and Communications* (Cambridge University Press). Her current research interests include the broad areas of statistical signal processing, sampling theory and compressed sensing, learning and optimization methods, and their applications to medical imaging, biology, and optics.

Dr. Eldar was a member of the Young Israel Academy of Science and Humanities and the Israel Committee for Higher Education. She was a member of the IEEE Signal Processing Theory and Methods and Bio Imaging Signal Processing Technical Committees. She was a Horev Fellow of the Leaders in Science and Technology Program at the Technion and an Alon Fellow. She is also a member of the IEEE Sensor Array and Multichannel Technical Committee and serves on several other IEEE committees. She is also a member of the Israel Academy of Sciences and Humanities (elected in 2017) and a EURASIP Fellow. She received many awards for excellence in research and teaching, including the IEEE Signal Processing Society Technical Achievement Award in 2013, the IEEE/AESS Fred Nathanson Memorial Radar Award in 2014, and the IEEE Kiyo Tomiyasu Award in 2016. She also received the Michael Bruno Memorial Award from the Rothschild Foundation, the Weizmann Prize for Exact Sciences, the Wolf Foundation Krill Prize for Excellence in Scientific Research, the Henry Taub Prize for Excellence in Research (twice), the Hershel Rich Innovation Award (three times), the Award for Women with Distinguished Contributions, the Andre and Bella Meyer Lectureship, the Career Development Chair at the Technion, the Muriel & David Jacknow Award for Excellence in Teaching, and the Technion's Award for Excellence in Teaching (two times). She also received several best paper awards and best demo awards together with her research students and colleagues, including the SIAM Outstanding Paper Prize, the UFFC Outstanding Paper Award, the Signal Processing Society Best Paper Award, and the IET Circuits, Devices and Systems Premium Award. She was selected as one of the 50 most influential women in Israel. She was the co-chair and the technical co-chair of several international conferences and workshops. She was a Signal Processing Society Distinguished Lecturer. She has served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, *EURASIP Journal on Advances in Signal Processing*, *SIAM Journal on Matrix Analysis and Applications*, and *SIAM Journal on Imaging Sciences*. She is also the Editor-in-Chief of *Foundations and Trends in Signal Processing*.

