Guest Editorial: Deep Learning in Medical Ultrasound—From Image Formation to Image Analysis

I. Introduction

OVER the past years, deep learning has established itself as a powerful tool across a broad spectrum of domains. While deep neural networks initially found nurture in the computer vision community, they have quickly spread over medical imaging applications, ranging from image analysis and interpretation to—more recently—image formation and reconstruction. Deep learning is now rapidly gaining attention in the ultrasound community, with many groups around the world exploring a wealth of opportunities to improve ultrasound imaging in several key aspects, ranging from beamforming and compressive sampling to speckle suppression, segmentation, and super-resolution imaging.

The goal of this Special Issue is to share these recent advances and methods with the UFFC community, highlight open challenges in this area, and motivate further research in this exciting direction. This Special Issue explores the use of deep learning for image formation, segmentation, and inference of specific ultrasound parameters. It covers applications in cardiovascular imaging, elastography, and photoacoustic imaging. Overall, this Special Issue contains 17 articles.

II. Image Formation

The first two articles present deep-learning approaches for image formation. Lu *et al.* propose a convolutional neural network (CNN) architecture for image reconstruction from diverging wave (DW) transmissions, referred to as Inception for DW Network (IDWNet). Multiscale convolutional kernels are adopted throughout an inception module to cope with the sectorial geometry of the acquisition strategy. The results using only three DW transmissions show image quality comparable to those obtained by standard compounding of 31 DWs. This allows boosting the frame rate while preserving image quality.

Also in the area of ultrafast ultrasound imaging, Nair *et al.* show for the first time the possibility of performing simultaneous B-mode image reconstruction and segmentation of anechoic areas using a single plane wave of raw IQ data (rather than RF data) as the input to a CNN. The CNN architecture consists of one encoder and two decoders, which simultaneously output a reconstructed B-mode image along with the segmentation results. This approach reformulates the problem of image formation into image transformation directly from channel data, circumventing the standard image formation processing chain, with future potential to reduce computational complexity.

III. Image Segmentation

The next four articles are focused on image segmentation. Amiri *et al.* optimize a U-Net architecture for ultrasound image segmentation. To cope with the availability of limited datasets, domain adaptation is usually employed by transfer learning and fine-tuning a network that is already trained on a large dataset from a different domain. The authors investigate the effect of fine-tuning specific sets of layers from a pretrained U-Net. The results provide indications for the selection of the layers to be fine-tuned. As shallow layers learn low-level features that are critical for segmentation, fine-tuning these layers may be preferred over fine-tuning the full network. The remaining segmentation articles focus on specific tissue structures.

Leclerc *et al.* introduce the use of a localization U-Net (LU-Net) pipeline to improve the segmentation accuracy of left ventricular structures (endocardial and epicardial borders) while enhancing the estimation of relevant clinical parameters such as volumes (end-systolic and end-diastolic) and ejection fraction. Zhou *et al.* address the segmentation of the myotendinous junction in muscular ultrasound. A region adaptive network is proposed to localize and segment the myotendinous junction in a single shot. A composite architecture is adopted that incorporates a region-based multitask learning network for identification of the junction and a U-shaped segmentation network to extract the structures of the myotendinous junction. Antico *et al.* propose a deep-learning classifier to assess the image quality in a relationship with the detection of the femoral cartilage. This is a relevant task for supporting navigation and proper identification of the surgical site during arthroscopic knee surgery. For classification, a Gaussian process deep neural network, consisting of a densely connected CNN, is combined with a Gaussian regressor.

IV. Inference of Specific Ultrasound Parameters

The following four articles present the use of deep learning for inference of specific ultrasound parameters, generating maps or images from the estimated parameters. Zhang *et al.* propose the use of a CNN for probabilistic scatterer estimation from ultrasound data, with the aim of overcoming the limitations of deconvolution approaches. A known statistical distribution is imposed on the scatterers, and the mapping

Digital Object Identifier 10.1109/TUFFC.2020.3026598

between ultrasound image and distribution parameter map is learned by training a CNN on synthetic images. Learning the scatterer distributions of tissues creates a new opportunity for the implementation of realistic simulation frameworks. The synthesized images from the estimated scatterer representations closely match the observations with varying acquisition parameters, such as compression and rotation of the imaged domain.

Evain *et al.* investigate Flow-Net2-based networks for motion estimation. In particular, the authors focus on the estimation of rotation, which is known to be a challenging task by standard tracking methods. To this end, a dedicated experimental phantom setup, made of a rotating disc, was employed for the evaluation of different Flow-Net2 solutions. Transfer learning of pre-trained networks was adopted for domain adaptation. The results provide clear indications of the best choices for improved motion estimation.

Wiacek *et al.* propose the use of the CohereNet deep neural network for learning spatial coherence functions. This technique is successfully proposed to support short-lag spatial coherence beamforming by processing aligned RF channel data. Moreover, based on the proven generalizability over different imaging systems, data types, and probe geometries, the method can be suitable for several clinical applications requiring the calculations of cross-correlations, such as elastography, strain imaging, and sound speed correction. The latter is specifically addressed by Bernhardt *et al.* in their article proposing the use of variational networks for speed-of-sound image reconstruction. The speed of sound is a relevant biomarker for breast cancer detection. Although variational networks have shown promise for inverse-problem optimization, their generalizability remains limited. To address this problem, the authors incorporate simulations of varying complexity in the training set. The training is regularized by loop unrolling of gradient descent with momentum, with an exponentially weighted loss of outputs at each unrolled iteration. Validation with a breast mimicking phantom shows promising results.

V. Cardiovascular Applications

Three articles are dedicated to cardiovascular applications. Smistad *et al.* present the use of several deep-learning components for real-time automatic assessment of ejection fraction, volumes, and apical foreshortening in the left ventricle using 2-D echocardiography. The proposed real-time pipeline can considerably reduce the time required for echocardiographic analysis in a standard clinical workflow.

Jahren *et al.* present a deep-learning solution for automatic detection of the end-diastole timing in cardiac spectral Doppler. This is particularly relevant when no electrocardiographic reference is present. For the detection, the authors propose the combination of a CNN for feature extraction with a recurrent neural network for modeling the temporal relations. Multiple spectral-Doppler modalities, including continuous-wave, pulsed wave, and tissue Doppler, can be processed. Spectrograms with end-diastole predictions of low confidence are automatically discarded, resulting in a practical method that has the potential to shorten the time required for echocardiographic examinations.

Nahas *et al.* address the problem of aliasing in ultrasound color Doppler, which may hamper the interpretation of the flow patterns. Combinations of several Doppler features are used for training a U-Net for identification of the aliased regions, which are then corrected by adaptive phase unwrapping. Validation with data from the femoral bifurcation shows good agreement of the method with manual dealiasing.

VI. Elastography

Two articles present the use of deep learning in elastography. Deep learning has been widely employed for the estimation of displacement in computer vision, but the different data characteristics of ultrasound RF data limit their direct application. To overcome these limitations, Tehrani and Rivaz propose two networks, namely, a modified pyramid warping and cost volume network (MPWC-Net) and the RFMPWC-Net. The proposed networks show a tenfold reduction in the number of parameters compared to the FlowNet2 architecture. A new loss function is also proposed for fine-tuning the network and improving the estimation performance. The proposed method shows similar performance as state-of-the-art optical-flow solutions, and it outperforms other deep-learning solutions proposed in the literature.

Considering that shear-wave elastography is not available in all ultrasound systems, Wildeboer *et al.* propose the use of a U-Net architecture to generate synthetic shear-wave elastographic maps using B-mode ultrasound data only. The method is validated for the ability to reproduce elastographic maps of the prostate. The generalizability of the method to different scanners is also proven, while the generalizability to other organs, such as the thyroid, is more critical, and requires dedicated retraining of the network. The promising results obtained by the authors suggest that B-mode features also incorporate information about tissue elasticity.

VII. Photoacoustic Imaging

The final two articles are dedicated to applications of deep learning in photoacoustics. The strength of a photoacoustic signal is fundamentally proportional to the local optical fluence, which decreases with increasing depth because of optical attenuation. This limits the ability to investigate deep structures in tissue. To address this limitation, Johnstonbaugh *et al.* present an encoder–decoder CNN architecture that is trained to identify the origin of the photoacoustic wavefronts in a deep, optically scattering medium. More specifically, the proposed deep-learning architecture makes use of an atrous convolution layer with Nyquist-rate dilation and stride in the temporal dimension, and a differentiable spatial-to-numerical transform layer to directly regress to coordinates, while simultaneously combining the design elements of U-Net and ResNet. Training and test data consist of simulated noisy measurements using a 256-element ultrasound transducer array. Simulation and experimental validation is presented, showing the ability to localize photoacoustic wavefronts located up to 4.6 cm in depth, in conditions where conventional beamforming approaches fail.

Awasthi *et al.* address the problems related to tomographic reconstruction in photoacoustics. A U-Net architecture with exponential linear unit activations is proposed for super-resolution, denoising, as well as bandwidth enhancement of raw tomographic photoacoustic data (i.e., sinograms). Down-sampled, noisy sinograms were input to the network, targeting fully sampled clean sinograms for training. Validation was performed with numerical and experimental phantom data, and results demonstrated image quality improvements when compared to more standard reconstruction methods and other U-Net architectures.

VIII. Conclusion

We conclude this overview by thanking all who helped to make this Special Issue possible, including all of the authors who chose to contribute their excellent work to this collection. In addition, we are grateful to the editorial board members involved with commissioning this Special Issue, as well as all editors and reviewers who made valuable comments and contributions during the peer review process. We hope that you will find the articles in this Special Issue interesting, stimulating, and useful for your own research, and that this Special Issue will continue to inspire future work on the use of deep learning in ultrasound imaging.

Massimo Mischi, *Guest Editor*
Department of Electrical Engineering
Eindhoven University of Technology
De Rondom 70
5612 AP Eindhoven, The Netherlands

Muyinatu A. Lediju Bell, *Guest Editor*
Department of Electrical and Computer Engineering
Johns Hopkins University
Baltimore, MD 21218 USA

Ruud J. G. van Sloun, *Guest Editor*
Department of Electrical Engineering
Eindhoven University of Technology
De Rondom 70
5612 AP Eindhoven, The Netherlands

Yonina C. Eldar, *Guest Editor*
Department of Mathematics and Computer Science
Weizmann Institute of Science
Rehovot 7610001, Israel

Massimo Mischi (Senior Member, IEEE) received the M.Sc. degree in electrical engineering from La Sapienza University, Rome, Italy, in 1999, and the Ph.D. degree from the Eindhoven University of Technology (TU/e), Eindhoven, The Netherlands, in 2004.

In 2007, he was an Assistant Professor with Electrical Engineering Department, TU/e. In 2011, he was an Associate Professor with TU/e and founded the Biomedical Diagnostics Research Laboratory, focusing on model-based quantitative analysis of biomedical signals with applications ranging from electrophysiology to diagnostic imaging, with a focus on ultrasound technology. He is currently a Full Professor with the Electrical Engineering Department, TU/e. He has coauthored over 300 peer-reviewed publications, 13 patents, one book, and several book chapters.

Prof. Mischi is a Board Member of the Imaging Section of the European Association of Urology, the Secretary of the Dutch Society of Medical Ultrasound, and the Chairman of the IEEE EMBS Benelux Chapter. He was a recipient of the STW VIDI Grant in 2009, the ERC Starting Grant in 2011, and the ERC Proof of Concept in 2019 for his research on angiogenesis imaging.

Muyinatu A. Lediju Bell (Senior Member, IEEE) received the B.S. degree in mechanical engineering (biomedical engineering minor) from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2006, and the Ph.D. degree in biomedical engineering from Duke University, Durham, NC, USA, in 2012.

From 2009 to 2010, she was a Whitaker International Fellow with The Institute of Cancer Research and Royal Marsden Hospital (Surrey), Sutton, U.K. From 2012 to 2016, she was a Post-Doctoral Fellow with the Engineering Research Center for Computer-Integrated Surgical Systems and Technology, Johns Hopkins University, Baltimore, MD, USA. She is currently an Assistant Professor with the Department Electrical and Computer Engineering with a joint appointment with the Department of Biomedical Engineering, Johns Hopkins University, where she founded and directs the Photoacoustic and Ultrasonic Systems Engineering (PULSE) Lab. She holds patents for short-lag spatial coherence imaging and photoacoustic-guided surgery. Her research interests include ultrasound and photoacoustic imaging, coherence-based beamforming, deep learning for ultrasound and photoacoustic image formation, image-guided surgery, robotics, and medical device design.

Dr. Bell's awards and honors include the NSF CAREER Award, the NIH Trailblazer Award, and MIT Technology Review's Innovator Under 35 Award. She also serves as Associate Editor-in-Chief of IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control and as Associate Editor of the IEEE Transactions on Medical Imaging.

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 2018, he has been an Assistant Professor with the Department of Electrical Engineering, Eindhoven University of Technology. Since January 2020, he has been a Kickstart-AI Fellow with Philips Research, Eindhoven. From 2019 to 2020, he was 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 (ultrasound) signal processing, model-based deep learning, compressed sensing, ultrasound imaging, and probabilistic signal and image analysis.

Dr. van Sloun is an NWO Rubicon Laureate and received the Google Faculty Research Award in 2020.

Yonina C. Eldar (Fellow, IEEE) is currently a Professor with the Department of Math and Computer Science, Weizmann Institute of Science, Rehovot, Israel, where she heads the Center for Biomedical Engineering and Signal Processing. She is also a Visiting Professor with the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, and the Broad Institute, Cambridge, an Adjunct Professor with Duke University, Durham, NC, USA, and a Visiting Professor with Stanford University, Stanford, CA, USA.

Prof. Eldar is a member of the Israel Academy of Sciences and Humanities and a fellow of EURASIP. She has received many awards for excellence in research and teaching, including the IEEE Signal Processing Society Technical Achievement Award, the IEEE/AESS Fred Nathanson Memorial Radar Award, the IEEE Kiyo Tomiyasu Award, the Michael Bruno Memorial Award from the Rothschild Foundation, the Weizmann Prize for Exact Sciences, and the Wolf Foundation Krill Prize for Excellence in Scientific Research. She is the Editor-in-Chief of *Foundations and Trends in Signal Processing*. She serves the IEEE on several technical and award committees.