Ultrasound in Med. & Biol., Vol. 46, No. 3, pp. 598–609, 2020
Copyright © 2019 World Federation for Ultrasound in Medicine & Biology. All rights reserved.
Printed in the USA. All rights reserved.
0301-5629/$ - see front matter

https://doi.org/10.1016/j.ultrasmedbio.2019.12.002

● *Original Contribution*

DEEP LEARNING-BASED SUPER-RESOLUTION ULTRASOUND SPECKLE TRACKING VELOCIMETRY

Jun Hong Park, Woorak Choi, Gun Young Yoon, and Sang Joon Lee
Department of Mechanical Engineering, Pohang University of Science and Technology (POSTECH), Nam-gu, Pohang, Republic of Korea

(*Received* 13 *March* 2019; *revised* 25 *November* 2019; *in final from* 1 *December* 2019)

Abstract—Deep ultrasound localization microscopy (deep-ULM) allows sub-wavelength resolution imaging with deep learning. However, the injection of contrast agents (CAs) in deep-ULM is debatable because of their potential risk. In this study, we propose a deep learning-based super-resolution ultrasound (DL-SRU), which employs the concept of deep-ULM and a convolutional neural network. The network is trained with synthetic tracer images to localize positions of red blood cells (RBCs) and reconstruct vessel geometry at high resolution, even for CA-free ultrasound (US) images. The proposed algorithm is validated by comparing the full width at half-maximum values of the vascular profiles reconstructed by other techniques, such as the standard ULM and the US average intensity under *in silico* and *in vitro* conditions. RBC localization by DL-SRU is also compared with that by other localization approaches to validate its performance under *in vivo* condition, especially for veins in the human lower extremity. Furthermore, a two-frame particle tracking velocimetry (PTV) algorithm is applied to DL-SRU localization for accurate flow velocity measurement. The velocity profile obtained by applying the PTV is compared with a theoretical value under *in vitro* condition to verify its compatibility with the flow measurement modality. The velocity vectors of individual RBCs are obtained to determine the applicability to *in vivo* conditions. DL-SRU can achieve high-resolution vessel morphology and flow dynamics in vasculature, mapping 110 super-resolved images per second on a standard PC, regardless of various imaging conditions. As a result, the DL-SRU technique is much more robust in localization compared with previous deep-ULM. In addition, the performance of DL-SRU is nearly the same as that of deep-ULM in rapid computational processing and high measurement accuracy. Thus, DL-SRU might become an effective and useful instrument in clinical practice. (E-mail: sjlee@postech.ac.kr) © 2019 World Federation for Ultrasound in Medicine & Biology. All rights reserved.

Key Words: **Convolutional neural network, Ultrasound imaging, Particle tracking velocimetry, Great saphenous vein.**

INTRODUCTION

For the clinical diagnosis of circulatory vascular diseases, the *in vivo* imaging of vascular morphology and blood flow in vasculature is essential (Feingold et al. 2010; Gessner et al. 2013). Ultrasound (US) imaging has been widely employed as one of the main angiographic modalities because of its inherent advantages, such as non-invasiveness, economy and rapid imaging process. Nonetheless, the diffraction limit and the trade-off between spatial resolution and penetration depth preclude its application in deep vascular imaging with high resolution.

Address correspondence to: Sang Joon Lee, Department of Mechanical Engineering, Pohang University of Science and Technology (POSTECH), 77, Cheongam-ro, Nam-gu, Pohang, 37679, Republic of Korea. E-mail: sjlee@postech.ac.kr

The advent of ultrasound localization microscopy (ULM) has recently enabled the super-resolution US imaging of deep vasculature by harnessing the concept of super-resolution optical microscopies, such as photoactivation localization microscopy. ULM detects microbubbles streaming in blood vessels by conducting signal decorrelation within hundreds of thousands of US images. The centroids of the detected bubbles are extracted by deconvolving their echoes from point-spread functions (PSFs), finally constructing super-resolved images by accumulating their local positions. ULM can be used for sub-wavelength structural imaging and hemodynamic analysis of rodent cerebral microvasculature (Desailly et al. 2015; Errico et al. 2015). In addition, recently, ULM has allowed super-resolved imaging of well-perfused breast cancer in a clinical setup (Dencks et al. 2019).

Although ULM can attain vascular images of sub-wavelength resolution, the number of localized microbubbles determines the accuracy of ULM. Mainly two factors govern the number of localized microbubbles, namely, the image acquisition rate (frame rate) and the concentration of microbubbles. The concentration of microbubbles is more important for accurate imaging than the image acquisition rate, the maximum of which is set systematically. However, the high concentration of microbubbles increases the difficulty of single-PSF deconvolution because of the overlapping of microbubble signals. Recently, sparse recovery algorithms were introduced to overcome this problem. However, other drawbacks, such as the time-consuming process and the frequent tuning for an optimal setting depending on the microbubble concentration, were concurrently observed (Bar-Zion et al. 2017, 2018).

Deep learning-based ULM (deep-ULM) was recently developed to achieve robust super-resolved US images under diverse imaging conditions such as high contrast agent (CA) concentration. A convolutional neural network (CNN) is employed in deep-ULM and trained with data sets of high-concentration microbubbles to localize their positions. This fast and robust method was reported to reconstruct super-resolved images of rodent and human microvasculature without manual modification (van Sloun et al. 2018).

Although deep-ULM is a cost-effective and robust localization technique under challenging *in vivo* conditions, it requires the injection of CAs. Several US CAs have been clinically approved and widely used as intravascular acoustic sources because of their cost effectiveness. Their prices range from \$50−\$100 and account for roughly 10%−15% of the ultrasonic examination fee. However, the use of CAs in the human body remains controversial because of their adverse effects. According to previous studies on adverse reactions to US CAs, 0.002% fatal adverse effects and 0.012% serious non-fatal adverse effects have been reported (Dijkmans et al. 2005). Patients are reluctant to receive CA injections, especially pregnant women because of anxiety about the health of their babies. In addition, some clinical conditions may prevent the seeding of microbubbles to real blood flows as tracer particles. Complex and fast blood flows in cardiovascular and carotid arteries may make the use of CA difficult because of the fast dissolution and rapid passage of microbubbles in the regions of interest (ROIs).

In several previous studies, US speckle patterns of red blood cells (RBCs) were used as tracer particles to obtain flow dynamic information (Nam et al. 2012; Lee et al. 2017). Although the speckle signal of RBCs has a lower intensity and more random shape than that of CA, it would be used as the source of RBC localization and vessel reconstruction similar to the speckle signal of CAs in deep-ULM.

In this study, we propose a deep learning-based super-resolution ultrasound (DL-SRU) technique that can handle RBC speckle localization by harnessing the deep-ULM approach and using a U-net, which is an architecture of a CNN (Ronneberger et al. 2015; van Sloun et al. 2018). We checked the feasibility of this concept through simulations and *in vitro* experiments. Then, we compared RBC localization under the CA-free US acquisition condition with that of other previous localization methods to verify the performance of the neural network trained with a RBC speckle pattern. Furthermore, a particle tracking velocimetry (PTV) algorithm was applied to the captured super-resolved US images to obtain accurate flow information. We checked the feasibility and medical applicability of the DL-SRU technique through US vascular imaging of the human lower extremity. As a result, US imaging of vascular morphologies and flow diagnosis using DL-SRU are sufficiently fast, accurate and robust for CA-free US imaging, compared with deep-ULM and a conventional clinical US imaging system.

METHODS

The synthetic training data were generated to train the proposed deep neural network utilized in DL-SRU. Thereafter, the performance of DL-SRU was validated through *in silico, in vitro* and *in vivo* experiments. Vascular bifurcation phantom models were simulated to verify the performance under simple conditions. Two vascular phantom models with lumens of different sizes were employed to test DL-SRU localization and the two-frame PTV algorithm under *in vitro* conditions at different flow rates. The tests for different lumen sizes and flow rates supported the feasibility of DL-SRU regardless of Reynolds number (Re). The applied US B-mode images were acquired with a clinical US imaging system. Spatiotemporal filtering and tissue motion compensation methods were implemented to reduce unintended noises in the raw B-mode images.

The proposed deep learning architecture and two-frame PTV algorithm were applied to speckle signals. Then, this approach and the application of the PTV algorithm achieved the localized signals and their velocity information. All proposed techniques and algorithms including the deep learning architecture were implemented in MATLAB R2018 b (The MathWorks, Inc., Natick, MA, USA). Image pre-processing for the imported images was conducted in ImageJ (National Institutes of Health, Bethesda, MD, USA).

Generation of synthetic training data

We prepared 64 × 64 image patches (2.59 × 2.59 mm in physical dimensions) containing many microbubbles and

RBCs with various US speckle features such as concentration, intensity and shape. These various speckle features were obtained from the real US B-mode images of microbubbles and RBCs, captured with a clinical US imaging system. The concentration was evaluated by counting the speckle localizations in the images using PSF deconvolution or manual detection per unit area. Their elliptical shape information was acquired by measuring the lengths of the elliptical axes and the rotating angles of the speckle signal. The backscattering intensity of target tracers was obtained by measuring the intensity profiles along the minor and major axes of the speckle signal.

These features were then reflected for the formation of synthetic US images. First, the target tracers randomly dispersed in the range 0−260 tracers/cm^2 were simulated to verify the performance of the proposed algorithm by comparing it with that of the previous study under the same experimental conditions (van Sloun et al. 2018). Next, their elliptical shapes and intensities were adjusted to the randomly dispersed positional information. Finally, Gaussian noise, with a mean value of 0.01 and a variance value of 0.001, was applied to formulate synthetic images similar to real images. The randomly distributed tracer image without adding backscatter intensity, elliptical shape and noise was used as the training output. The final simulated patches downsized to 16×16 pixels were turned into contrast-enhanced ultrasound (CEUS) images with the aid of the contrast-limited adaptive histogram equalization (CLAHE) method (Reza 2004). By adoption of the CLAHE method, the image contrast was amplified in the near-constant regions of the image without noise amplification by setting an amplification limit. The simulated CEUS images were used as input data for training the automatic tracer localization algorithm. The synthetic images were utilized to make various training data sets, because their parameters were quantitatively varied by changing the US speckle features. The synthetic training data sets generated by reflecting the features of the original US images were employed to train the devised neural network of DL-SRU even under challenging imaging conditions, such as poor signal-to-noise ratio (SNR).

Simulation of bifurcation network

We designed synthetic tracers that move along a vascular bifurcation network by distributing tracers uniformly in the entire region of the branches. The vascular width in the simulated bifurcation network was 200 μm, and the whole length of the vascular bifurcation was 31.9 mm. The synthetic tracers were randomly distributed in the range 0−260 tracers/cm^2 to compare with the simulated results of the previous study. A total of 5000 frame images were generated to sufficiently reconstruct the simulated vascular bifurcation and tested for training the localization algorithm. DL-SRU localized over 50,000 tracers in the generated 5000 frames.

In vitro *microbubbles and* in vivo *human US imaging*

We designed two bifurcation phantoms made of hydrogel to validate the algorithm through *in vitro* microbubble experiments. One of the bifurcation models had a lumen of 700 μm in diameter and 11.5 mm in whole length. The other had a lumen of 200 μm in diameter and 10 mm in whole length, respectively. The hydrogel phantoms were fabricated according to the procedure described in our previous study (Lee et al. 2017). The upper parts of the phantoms were filled with distilled water to reduce the acoustic attenuation at their interfaces, and 0.5% Sonazoid (GE Healthcare, Oslo, Norway) diluted with phosphate buffered saline solution at pH 7.4 (Bio Solution, Korea) was supplied into the bifurcation phantoms through a syringe pump (neMESYS, Cetoni, Germany). The flow rates of the working fluid in the bifurcation models with lumens of 700 and 200 μm were set to 0.5 and 0.1 mL/min, respectively. The tracks of the microbubbles were observed on the axial sectional B-mode images.

In vivo feasibility tests were conducted for healthy volunteers as experimental patients. The human experiments were performed according to the Declaration of Helsinki with a permission from the institutional review board (IRB) of Pohang University of Science and Technology (POSTECH; PIRB-2017-E060). All volunteers were briefed about the test purpose, procedures and possible risks before signing the written consent. The RBC speckle patterns were obtained from the human posterior accessory great saphenous vein (GSV) in the region near the ankle perforator sitting in a comfortable position. Digital image processing techniques were applied to the captured US images to enhance RBC signal and remove tissue signal. The processed US images exhibit a blood-to-tissue signal ratio of 7.7 dB, while the ratio of the unprocessed US images is −2.62 dB. The localization algorithms were applied to the processed US images.

US B-mode images were acquired at a high frame rate using a SmartUS EXT-1 M system and a L18-10 L30 H-4 linear probe (TELEMED, Milan, Italy; http://www.pcultrasound.com, accessed 23 July 2019). The transducer frequency, focal length and dynamic range of the device were set to 18 MHz, 12 mm and 90 dB, respectively. The image acquisition rate was adjusted in the range 78−79 Hz. US speckle patterns of microbubble tracers and RBCs were recorded into 798 frames over 10 s. The acquired US data were saved as Digital Imaging and Communications in Medicine (DICOM)-based files. Saved data were exported by RadiAnt, which is a Windows DICOM viewer (Medixant, Poznan, Poland), as Joint Photographic Experts Group (JPEG)-based image sequences. Then, the JPEG images were imported on ImageJ (National Institutes of Health, Bethesda, MD, USA). They were converted to 8-bit gray-scale images, and the (ROIs) were cropped by the image processing

software. Finally, post-image processing techniques such as CLAHE and singular value decomposition (SVD) were applied to the pre-processed JPEG image bundles by ImageJ to generate the training data sets for the proposed neural network.

Spatiotemporal filtering and tissue motion compensation

The following spatiotemporal clutter filtering technique based on SVD, was applied to differentiate the US signals of tracers from the background signal (Demené et al. 2015):

$$s_{\text{blood or CA}}(x,z,t) = s(x,z,t) - \sum_{i=1}^{n} \lambda_i I_i(x,z) V_i(t) \quad (1)$$

where $s(x, z, t)$ denotes the spatiotemporal cine loop corresponding to the ultrasonic raw data. The ROI was defined at the beam-focused microvasculature region. x and z represent the lateral dimension along the transducer array and the depth of the medium in front of the US probe, respectively, and t stands for time. The matrix S was made by rearranging two spatial dimensions and one temporal dimension simultaneously. Whole 2-D US images at each frame were included in the matrix completely. λ_i denotes the ordered singular value. $I_i(x, z)$ denotes the 2-D spatial image with the numbers of samples along the x- and z-directions. $V_i(t)$ denotes the temporal signal with the number of time samples. n denotes the number of singular vectors removed from the raw signal. The filtered images (without background signal) were utilized with the trained algorithm.

Image acquisition during a certain time span was hampered by unnecessary tissue motions, which were caused mainly by respiration. Hence, such inadequate images were compensated to obtain motion-free image sequences. Motion-free images could generally be attained from the normalized cross-correlation (CC) between frames in the ROI.

$$\mathrm{CC}(m) = \frac{\sum_u \sum_v \mathrm{IQ}(u,v,m) \cdot \mathrm{IQ}^*(u,v,m+1)}{\sqrt{\sum_u \sum_v \mathrm{IQ}(u,v,m) \cdot \mathrm{IQ}^*(u,v,m) \cdot \sum_u \sum_v \mathrm{IQ}(u,v,m+1) \cdot \mathrm{IQ}^*(u,v,m+1)}} \quad (2)$$

where IQ represents the in-phase/quadrature data. m, u and v represent the time, lateral and axial indexes within the ROI. The asterisk (*) denotes the complex conjugation form of the data. The frames affected by severe motion were discarded when their correlation values were locally low. The ROI was selectively determined in specific vasculatures to calculate the frame-to-frame correlation values (Supplementary Fig. S1a, online only). In the images captured at cardiac pulsation and respiration stages, the frame-to-frame correlation values sharply decreased (Supplementary Fig. S1b, online only). Thus, these frames exhibiting large decreases in the correlation values were neglected to minimize the effects of tissue motion. For this, the threshold of CC was set to 0.965 and the rejection rate was 11.16 % (28/251 images) for 3.21 s. With the repetition of this compensation process, motion-free images could be attained appropriately (Foiret et al. 2017).

Deep learning architecture

We implemented a U-net architecture by integration of the CNN-related MATLAB toolboxes (MATLAB R2018 b) including the Deep Learning Toolbox and Computer Vision Toolbox. The Deep Learning Toolbox is utilized to design and implement deep neural networks with algorithms, pre-trained models and apps. The Computer Vision Toolbox is utilized for designing and testing computer vision and video processing systems with several algorithms, functions and apps. They are indispensable in the semantic segmentation of multispectral images using deep learning architecture. In general, U-net integrates a contracting path and an expanding path. In this case, the contracting path consisted of three-layer blocks. Each block included two 5×5 convolution layers with a leaky rectified linear unit (ReLU) and a 2×2 max pooling operator with stride 1. The leaky ReLU was used to prevent the inactive neural nodes from decreasing the capacity of the model (Xu et al. 2015). The next layer comprised two 3×3 convolution layers and a dropout layer with 50% probability. Although we set the dropout rate of 50% probability on the hidden layers, it randomly ignored some number of the layer outputs with 50% probability. Therefore, the dropout layer involved nearly 50% of the layer features during the training process. Then, the space of this layer enabled the conversion of images into high-resolved images by using the expanding path. The expanding path included three blocks. The first two groups had two 5×5 deconvolution layers followed by a 2×2 upsampling layer (Long et al. 2015). The latter deconvolution layer had an output stride of 2. The last expanding block consisted of two deconvolution blocks; the second also had an output stride of 2. Another 7×7 convolution layer followed the last block. Then, the processed information was morphed into a single-channel image using a linear activation method. The completed neural network converted low-resolved 64×64 input images, which were upsized from the original 16×16 input images, to high-resolved 64×64 output images. The additional

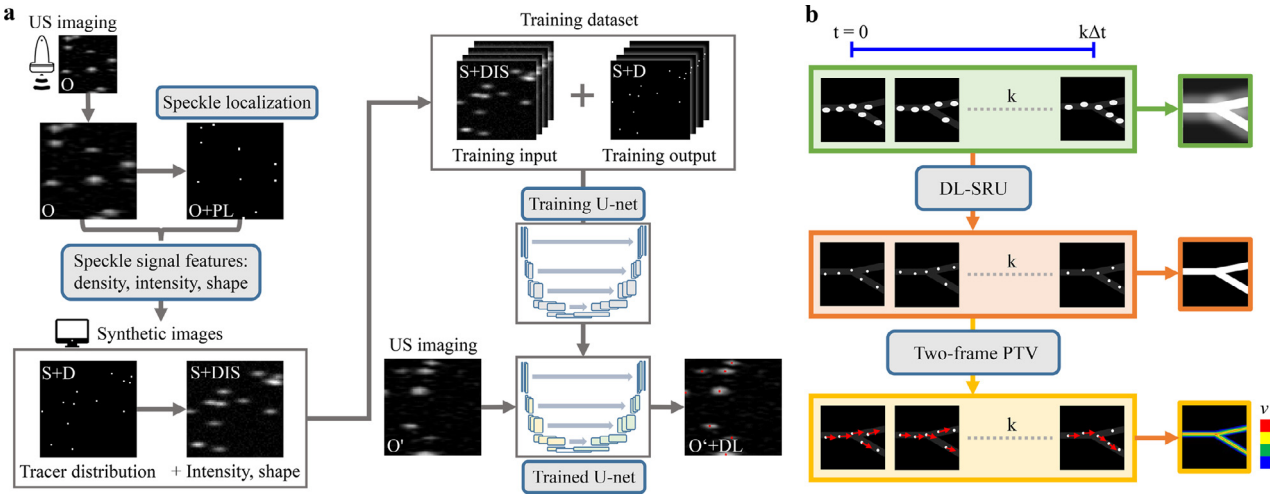

Fig. 1. Overview of DL-SRU and two-frame PTV applications. (a) Training images were synthesized reflecting the target tracer's features extracted from the corresponding original US images (O) and their speckle localized images (O + PL) by using point spread function deconvolution. The U-net architecture was trained by using realistic synthetic images (S + DIS) as the training inputs in which parameters, such as concentration, intensity and shape, were varied. The corresponding position-marked images (S + D) were used as the desired outputs. The tracer localization by DL-SRU (O' + DL) upscaled the low-resolved original US images (O'). (b) DL-SRU localization was reconstructed from a certain number of frames (*k*) to obtain a highly resolved geometry (*orange box*), compared with the original US accumulation (*green box*). A two-frame PTV algorithm based on match probability was applied to DL-SRU localization to obtain the flow dynamic information for the time span (*k*Δ*t*) (*yellow box*). The colored contours and color bar represent the velocity magnitude information measured with the PTV technique. DL-SRU = deep learning-based super-resolution ultrasound; PTV = particle tracking velocimetry; US = ultrasound.

information of each layer of the U-net is described in Supplementary Figure S2 (online only).

We employed the adaptive moment estimation (Adam) method with a learning rate of 0.0001. The network was trained with 18,000 iterations to optimize the cost function as follows (Nehme et al. 2018)

$$c(x,y|\theta) = \|f(x|\theta) - G * y\|_2^{\,2} + \lambda \|f(x|\theta)\|_1 \qquad (3)$$

where x and y represent the input and high-resolution output, respectively; $f(x|\theta)$ denotes the non-linear neural network function with θ, weight and bias; λ denotes the regularization parameter; G and * represent the 2-D Gaussian filter and convolution, respectively. Training and test runs were conducted using an NVidia GeForce GTX 1070 Ti with a video memory of 8 GB. The performance of DL-SRU was evaluated using the ThunderSTORM plugin of ImageJ for a standard ULM (Ovesný et al. 2014).

Speckle tracking velocimetry

Speckle tracking velocimetry (STV) is composed of DL-SRU, which can localize the exact positions of the speckle signal, and two-frame PTV, which provides flow information by tracking the localized speckle signals from consecutive US images. The details of the two-frame PTV technique are described in a previous study (Baek and Lee 1996).

Deep learning approach and PTV application

Figure 1a illustrates the procedure for the proposed DL-SRU technique. Real US B-mode images of the target tracers (CAs or RBCs) were captured (O) and then several features, such as backscatter intensity, concentration and shape, were extracted from the images. Next, their speckles were localized by PSF deconvolution (O + PL). Synthetic tracers were randomly dispersed based on the concentration information obtained from the real speckle signal (S + D). Then, the extracted backscattered intensities and shapes of the real target tracers were applied to the positional information of individual synthetic tracers to make the images look similar to the real ones (S + DIS). U-net exploited both the synthetic tracer input (S + DIS) and the ground truth images (S + D). Thereafter, DL-SRU, which was trained with the synthetic data, localized the positions of the target tracers (O' + DL) from the real US test images (O') captured under *in vitro* and *in vivo* conditions. Figure 1b illustrates a super-resolved vessel geometry (*orange box*) which was reconstructed from the localized positions within a certain number of frames (*k*), while the real US data exhibit a poor-resolved vessel geometry (*green box*). The two-frame PTV algorithm was applied to the localized target tracers to obtain their flow dynamic information for a given time span (kΔt) (*yellow box*).

Statistical analysis

Throughout this article, all data are expressed as the mean ± standard deviation. The significance of differences between velocity magnitudes along the radial direction were estimated using Student's *t*-test to compare near-wall velocities and centerline velocities. Each paired comparison has its own *p* value according to different thresholds depending on measurement positions. A *p* value <0.05 can be considered to indicate statistical significance at the corresponding comparisons. Statistical analyses were carried out using OriginPro 8.5 (OriginLab Corp., Northampton, MA, USA).

RESULTS

DL-SRU for in vitro *microbubble and RBC images*

Figure 2 illustrates the typical results of DL-SRU application on real microbubble images, trained with synthetic tracer data sets. The locations predicted by the proposed algorithm are well-correlated with the true locations obtained by manual pinpointing. DL-SRU especially distinguishes the overlapped speckle signals properly (Fig. 2b). Its localization accuracy and recall were estimated as a function of microbubble concentration, and the results are illustrated in Figure 2c and 2d. The localization error caused by the tracer concentration of 14.77 MB/cm^2 has the lowest value. The errors increase with the increase in tracer concentration, but the increasing rate of error becomes smaller. The mean value of localization errors does not exceed 0.1 mm. The highest recall rate occurs at the lowest tracer concentration, and the recalls decrease with the increase in tracer concentration. However, their decreasing rate becomes smaller and all the recalls in the graph exceed 60%. These results indicate a high localization accuracy and a high recall rate even at high bubble concentrations, although the best performances were observed at the lowest tracer concentration. Van Sloun et al. (2018) reported that the U-net-based localization algorithm provides a median localization error that is smaller than the original pixel size. The recall rate was high for high microbubble concentrations, as reported in the previous study. Supplementary Figure S3 (online only) compares the localization results for additional microbubble and RBC images, obtained with the standard ULM (Ovesný et al. 2014) and DL-SRU. DL-SRU outperformed the standard ULM in localizing microbubbles and RBCs. The standard ULM did not pinpoint the various single or overlapped microbubble signals adequately, while DL-SRU localized microbubbles within a 1-pixel radius of the true locations (Supplementary Fig. S3a, 3b, online only). Likewise, RBC localizations by DL-SRU were achieved within a 1-pixel radius of their true locations. However, the standard ULM could not detect the exact locations of RBC signals because of their weak and overlapping signals (Supplementary Fig. S3c, online only). The localization performance of the DL-SRU for both microbubbles and RBCs was better than that of the standard ULM, except at the lowest tracer concentration. DL-SRU was found to maintain high localization accuracy even at high tracer concentrations, while the localization accuracy of the standard ULM decreased with increasing tracer concentration, as illustrated in Supplementary Figure S3(d, e) (online only). The time required to obtain super-resolved US imaging is less than 10 ms with the proposed DL-SRU, compared with approximately 24 ms with the standard ULM. These results indicate that DL-SRU, which was trained with synthetic tracer data sets adequately, could discern real tracer signals better and faster than the standard ULM.

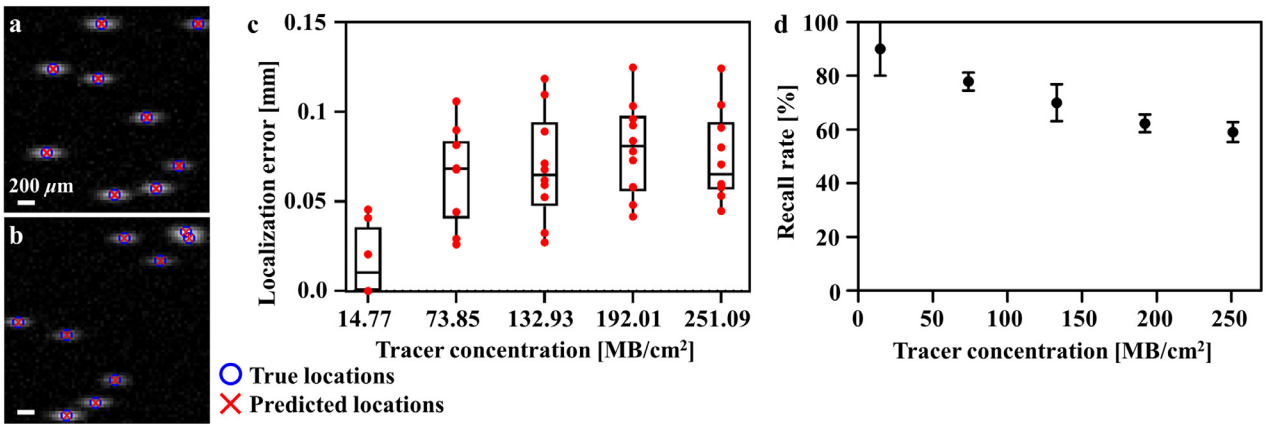

Fig. 2. Typical examples of *in vitro* DL-SRU localization and variations of localization error and recall rate according to tracer concentration. (a, b) Typical *in vitro* test results of DL-SRU trained with synthetic image data sets. The corresponding true and the predicted locations are marked as *blue circles* and *red crosses*, respectively. (c) Variations of localization error and (d) the corresponding recall rate according to tracer concentration. The error bar indicates the standard deviation of localization errors and recall rate values. DL-SRU = deep learning-based super-resolution ultrasound; MB = microbubbles.

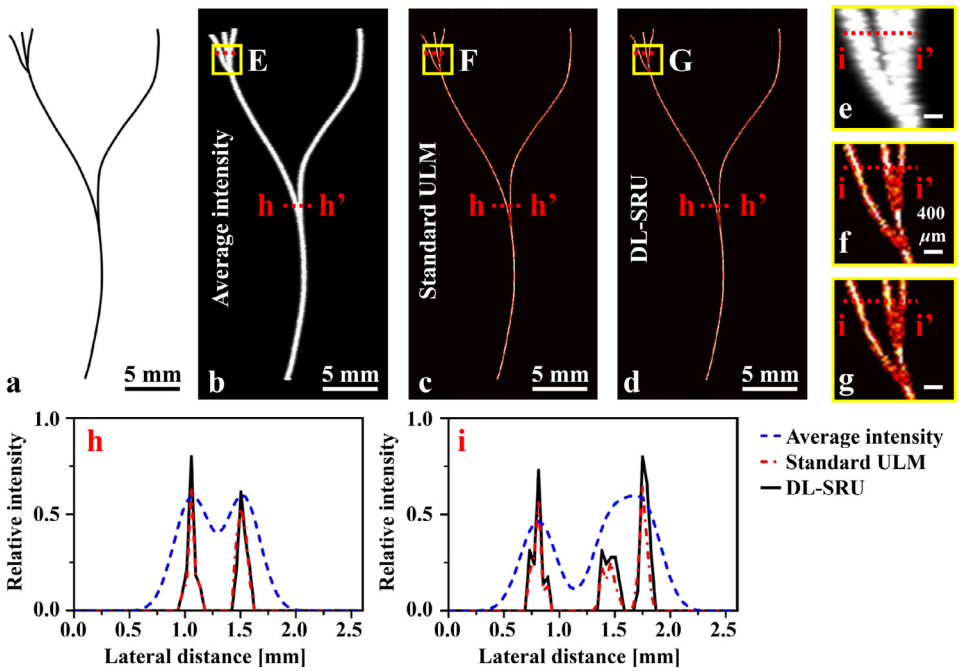

Fig. 3. Comparison of localization performance of DL-SRU and the standard ULM for *in silico* bifurcation. (a) Simulated vascular bifurcation model. (b) Diffraction-limited average intensity image. (c, d) Vasculatures reconstructed by the (b) standard ULM and (d) DL-SRU. (e–g) Zoomed images of (b)–(d). (h, i) Comparison of intensity profiles along dashed lines (h) h–h' and (i) i–i' in the reconstructed bifurcation images and their zoomed images. DL-SRU provided good image quality beyond the diffraction limit like the standard ULM. DL-SRU = deep learning-based super-resolution ultrasound; ULM = ultrasound localization microscopy.

DL-SRU for in silico *vascular bifurcation*

We investigated the effect of the DL-SRU algorithm on resolution recovery for the *in silico* vascular bifurcation model with synthetic tracer signals. Figure 3 compares the spatial resolution of DL-SRU with that of the average intensity method and the standard ULM. The standard ULM and DL-SRU delineated the bifurcation and trifurcation with high resolution, while the average intensity method failed to recover the vasculature geometry appropriately in the magnified ROI (Fig. 3e–g). Figure 3(h, i) illustrates that the DL-SRU method and the standard ULM could resolve the vascular models better than the diffraction limit of the clinical US imaging system. The relative intensity of DL-SRU was slightly higher than that of the standard ULM because of its superior signal recovery and reconstruction for the bifurcation and trifurcation models. We evaluated their full width at half-maximum (FWHM) values to compare their spatial resolutions quantitatively. In Figure 3h, the FWHM evaluated by average intensity was 829.1 μm because of the overlapping of two peak profiles. However, the FWHMs by the standard ULM were 61.2 and 191.0 μm at the peaks from left to right. On the other hand, the FWHMs by DL-SRU were 55.1 and 118.1 μm from the left-hand peaks, which were smaller than any other FWHMs. In Figure 3i, the FWHMs by average intensity were 382.3 and 619.2 μm at the peaks from left to right. However, the FWHMs by standard ULM were 76.4, 152.2 and 105.0 μm at the peaks from left to right. The FWHMs by DL-SRU were smaller than any other FWHMs, with values of 65.6, 142.2 and 100.0 μm from the left-hand peaks. Thus, the proposed algorithm trained with multiple overlaps in high-concentration input images could provide high-resolved images.

DL-SRU for in vitro *vascular bifurcation phantom models*

We continued to test the DL-SRU algorithm for two *in vitro* bifurcation phantom models by using a clinical US system (Fig. 4a). Approximately 85,000 localizations were obtained from the acquired 798 frames. The average intensity of the US images and the reconstruction of the standard ULM exhibited boundary noise at the bottom of the vessel (Fig. 4b, 4c). However, the boundary geometry reconstructed by DL-SRU was clearer with less noise compared with those reconstructed by the two other methods. The rough boundaries reconstructed by the average intensity and standard ULM methods were regarded as image noise, because the original bifurcation phantom models had smooth lumen surfaces. In addition, the width of the reconstruction intensity profiles by DL-SRU was narrower than those by the two other methods. This implies that the spatial resolution of DL-SRU is better than that of the other methods and is beyond the

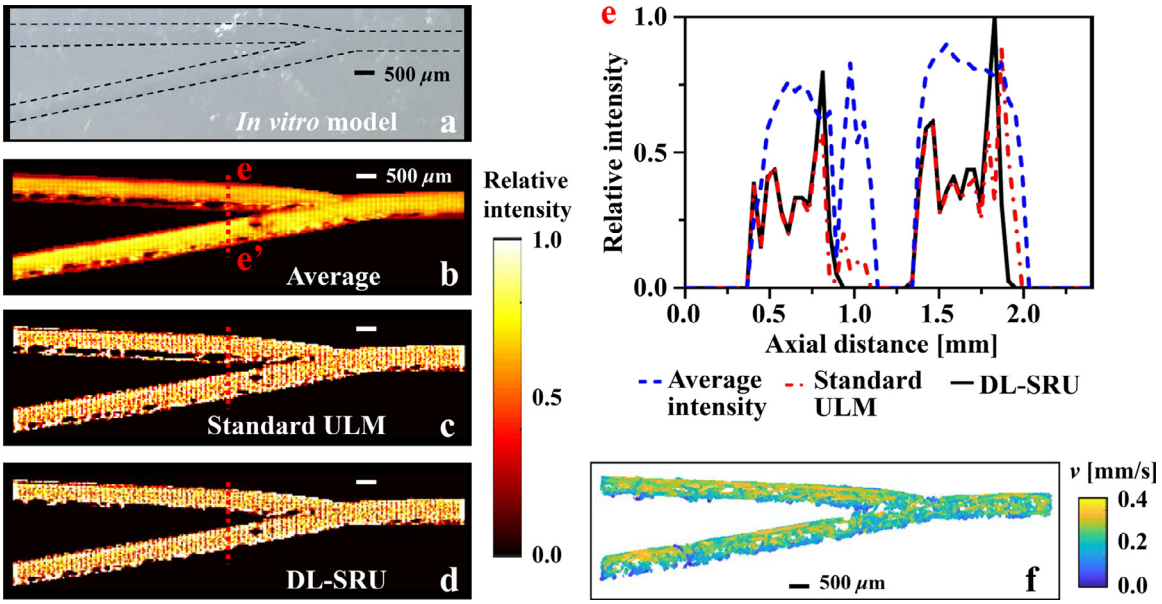

Fig. 4. Comparison of the performance of DL-SRU and the standard ULM for *in vitro* bifurcation, and the application of the PTV algorithm to DL-SRU localization. (a) *In vitro* model of vascular bifurcation made of hydrogel, and (b) its diffraction-limited average intensity image. (c, d) Vasculature reconstructed by the (c) standard ULM and (d) DL-SRU. (e) Comparison of intensity profiles along dashed line e-e' in the reconstruction images. DL-SRU resolved *in vitro* US images better than the original US and the standard ULM. (f) A typical velocity field of microbubbles measured by two-frame PTV algorithm. DL-SRU = deep learning-based super-resolution ultrasound; PTV = particle tracking velocimetry; ULM = ultrasound localization microscopy; US = ultrasound.

diffraction limit of the commercial clinical US imaging system (Fig. 4d). Figure 4e compares the spatial resolution of DL-SRU with those of the average intensity and the standard ULM along line e−e'. The average intensity and the standard ULM had unexpected boundary errors at an axial distance of approximately 400 μm. The widths in the relative intensity for both methods were wider than that of the proposed DL-SRU.

Figure 4f illustrates the velocity and trajectory of each tracer localized by DL-SRU, measured by applying the two-frame PTV algorithm. Most of the trajectories of the individual tracers were parallel to the phantom wall when they passed through the vessel. However, the trajectories of a few tracers roamed in certain regions because of the rough and uneven phantom surface resulting from the failure in the manufacturing process. The tracers propagating along the centerlines had the highest speed. Their propagating speeds had lower values in the region near the bottom wall than those near the upper wall because of the sedimentation of microbubbles.

Supplementary Figure S4(a) (online only) illustrates a vascular phantom model with a lumen of 200 μm, reconstructed by DL-SRU. The spatial resolution of DL-SRU is compared with that of other techniques along the *blue solid line* in Supplementary Figure S4(b) (online only). The average intensity method fails to recover the vascular bifurcation, while the standard ULM and DL-SRU are observed to reconstruct high-resolution phantom images beyond the diffraction limit. Their FWHMs were evaluated to compare their spatial resolutions quantitatively. The spatial resolutions of DL-SRU and the standard ULM are 126 and 130 μm, respectively, while the average intensity method does not provide a resolved result because of the overlapping intensity signals.

The two-frame PTV algorithm was employed to obtain the velocity vector of each tracer localized by DL-SRU (Supplementary Fig. S4c, online only). Most individual trajectories move well along the phantom wall. Supplementary Figure S3(d) (online only) compares the measured average velocity profile of microbubbles with the theoretical velocity profile in a circular pipe flow. Statistical differences between near-wall velocities and centerline velocities are distinguishable ($p < 0.05$). In addition, the experimental velocity profile matches well with the theoretical one. This implies that DL-SRU successfully localized the tracers with high resolution, and the two-frame PTV algorithm can be applied to localized tracers to measure their velocity vectors, regardless of flow rate or Re.

In vivo *DL-SRU for human posterior accessory GSV*

The applicability of DL-SRU in clinical practice was illustrated by applying the trained algorithm to

clinical US images obtained without CAs. In Figure 5a are typical clinical US B-mode images acquired from the human posterior accessory GSV near the ankle perforator. The RBCs localized by the standard ULM and DL-SRU methods are illustrated in Figure 5(b, c). Unlike the standard ULM, DL-SRU could pinpoint the RBC speckle patterns that critically overlapped with low intensity and various shapes (Fig. 5b, 5c). The excellent performance of DL-SRU in RBC localization was clearly illustrated in the magnified ROI, and this finding is compared with that of the standard ULM in Figure 5d. DL-SRU recovered RBC localizations from the bottom, center and upper wall regions (Fig. 5d1, d2, d3), respectively, whereas the standard ULM failed.

Figure 5e illustrates velocity vectors and trajectories of the RBCs localized by DL-SRU, measured by applying the two-frame PTV algorithm. The high velocity in the center region gave rise to a tracking error caused by missing tracers and the excessive maximum distance threshold employed in the application of the PTV algorithm. We especially focused on the tracers whose velocities were smaller than 1 mm/s to track the motions of near-wall RBCs accurately. Most of the RBCs in the near-wall regions moved along the near-wall geometry substantially (Fig. 5e). The moving velocities of the RBCs were low when they were located nearer the wall. Few trajectories and the high velocities of some RBCs contained errors because of the rapid flow change in the presence of the ankle perforator (Fig. 5f). These results indicate that DL-SRU could localize RBC speckle signals better than the standard ULM. In addition, the RBC localizations from DL-SRU can be utilized for tracking RBC speckles to obtain velocity information with the aid of the two-frame PTV algorithm.

DISCUSSION

Deep-ULM technique allowed researchers to overcome the diffraction limit of US imaging and achieve the high-resolution imaging of microvasculature (van Sloun

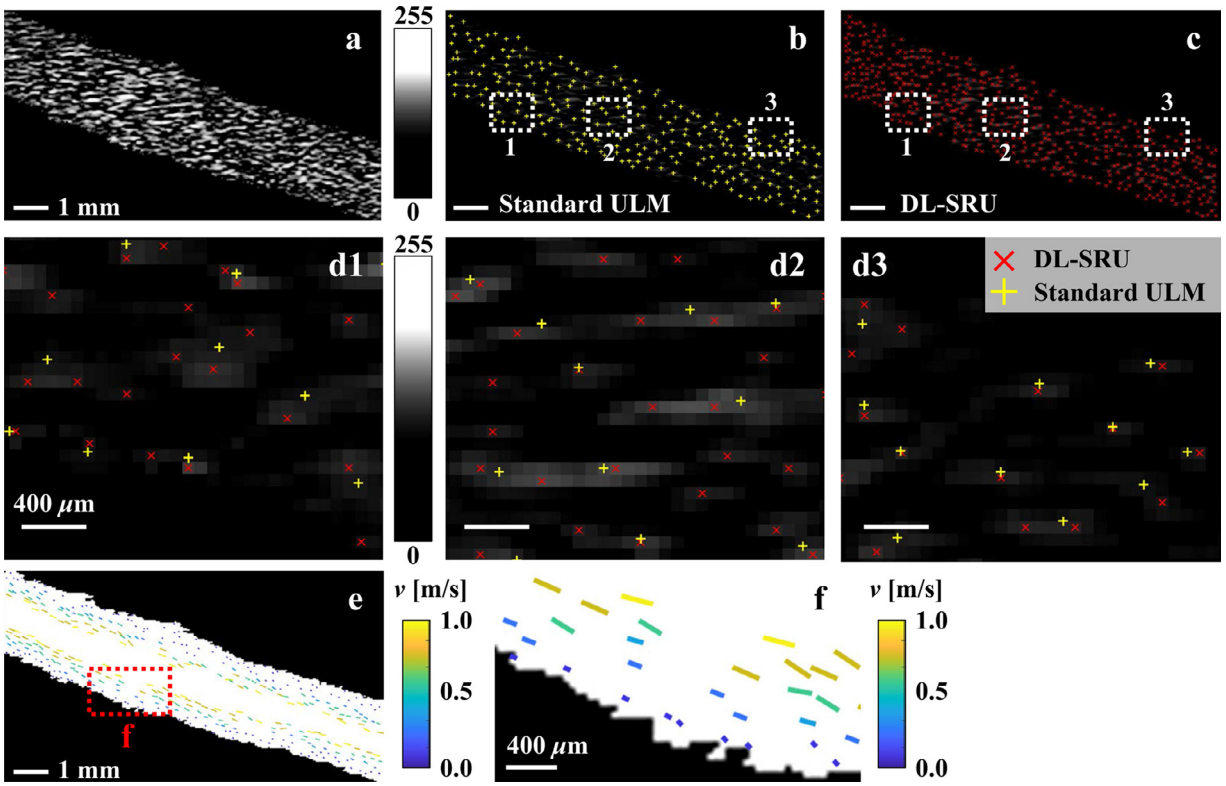

Fig. 5. Comparison of the *in vivo* locations of DL-SRU and the standard ULM for the human posterior accessory great saphenous vein, and the application of two-frame PTV algorithm to *in vivo* DL-SRU localization. (a) Original US images captured with the contrast-limited adaptive histogram equalization method at a frame rate of 78 Hz. (b, c) Examples of localization with (b) DL-SRU and (c) the standard ULM. (d1–d3) Zoomed images of the standard ULM and DL-SRU in the selected areas (1–3) in (b) and (c), respectively. DL-SRU could recover the overlapped RBC signals, while the standard ULM failed. (e) Velocity vectors of RBCs across three consecutive frames using the two-frame PTV algorithm, and (f) the corresponding zoomed image of (e). The RBC velocities were up to 1 mm/s. DL-SRU = deep learning-based super-resolution ultrasound; PTV = particle tracking velocimetry; RBC = red blood cell; ULM = ultrasound localization microscopy.

et al. 2018). However, the clinical side effects on the human body of the CAs that the system employs as tracers remain controversial. In this study, we trained DL-SRU with RBC synthetic images for clinical practice by performing localization on real US images of RBCs without CAs. The proposed neural network can be applied to various biomedical fields, depending on the quality and type of training data sets (Ronneberger et al. 2015). RBCs could be localized by the present neural network by training well-synthesized RBC speckle images under challenging *in vivo* conditions, such as low frame rate, high concentration and unintended organ motion.

Synthetic data set of in vitro *microbubble image*

DL-SRU employed the CNN that was trained with synthetic data sets of microbubbles or RBCs. The performance of the algorithm was determined by checking how the input data were generated similarly to the real tracer speckle patterns. With varying ellipsoidal shapes, tracer densities, backscattering intensities and SNRs, various synthesized data sets were incorporated to improve the robustness of the trained network. In particular, the concentration of tracers in the synthetic images was a dominant factor because it caused frequent overlapping. In this study, DL-SRU was found to discern the positions of tracers accurately even at high concentrations. Nevertheless, relatively lower tracer concentrations were ideal for DL-SRU localization because the overlapping of speckle signals may cause overestimation in the localization of tracers at high concentration.

U-net consisted of a contracting encoder for expanding the decoder structures. This technique has been widely applied in multiple segmentation in biomedical fields (Ronneberger et al. 2015; Badrinarayanan et al. 2017; Chen et al. 2017; van Sloun et al. 2018). The network had 15 layers to secure the capability of tracer localization without overfitting. The dropout layer with 50% probability was favorable for the algorithm's performance and computational cost. DL-SRU allowed the recovery of super-resolved images within 10 ms.

In silico *vascular bifurcation and* in vitro *bifurcation phantom*

The results of the *in silico* and *in vitro* experiments indicated that DL-SRU can pinpoint the locations of microbubbles and reconstruct the vascular geometry satisfactorily, because it was trained to handle highly concentrated tracers and many noises. In particular, DL-SRU performed better than the other approaches despite the imperfect *in vitro* phantom for the localization test. This result was consistent with the performance of deep-ULM that was validated by using images of microbubbles (van Sloun et al. 2018). The *in vitro* trajectories and velocity vectors, extracted from the microbubble localizations, illustrated that the PTV algorithm could be applied to DL-SRU adequately. The velocity fields derived by the PTV algorithm were well correlated with the previous results calculated by computational fluid dynamics (CFD) (Gharahi et al. 2016). As the flow passed through the bifurcation region, the speed decreased, and the flow direction deviated in the near-wall region because of flow recirculation. Furthermore, the flow was slower in the region near the bottom outside wall because of the sedimentation effect of the microbubbles.

In vivo *GSV study with a clinical system*

DL-SRU was developed to improve the localization and tracking of RBCs compared with those of deep-ULM method, regardless of clinically challenging imaging situations. In addition, it improved spatial resolution compared with the other localization methods. The trajectories and velocity fields of RBCs were derived well from their localizations by DL-SRU. The derived velocity fields agreed with previous results obtained by using nuclear magnetic resonance and US speckle image velocimetry (Kerr et al. 1991; Nam et al. 2012). In the near-wall region, the measured velocity was less than 1 mm/s, and the RBCs moved along the surface morphology of the wall. However, the trajectories and velocity fields of a few RBCs were distorted by the perforators (Fig. 5f). The combination of DL-SRU with the two-frame PTV algorithm facilitates the collection of flow information of high accuracy in clinical practice. We examined the feasibility of DL-SRU for CA-free US imaging and the effect of modality combination in comparison with deep-ULM.

Nevertheless, a few limitations remain. First, although DL-SRU successfully achieved the RBC localization and PTV combination, the motions of vessels and the surrounding organs may complicate the recovery of super-resolved images. The large fluctuations in tissue and organ signals were removed by adopting digital image processing techniques. However, the removal of many motion-affected frames could lead to reduction in the number of available frames for vessel reconstruction and velocity field measurement. This condition also affects the accuracy of image reconstruction.

Second, we could not assess the best computation performance of DL-SRU because of the absence of the latest GPU computation. Therefore, DL-SRU could still be improved, for example, by upgrading the hardware or accelerating the neural network using model compression (Cheng et al. 2017).

Third, the applied two-frame PTV algorithm is limited in terms of the selection of the maximum displacement threshold for a short time interval. When the PTV algorithm is applied to discern localization for a very

fast flow, the method may produce some errors because of the excessive maximum displacement threshold.

Deep-ULM technique was introduced to overcome the diffraction limit of US and acquired high-resolution images of microvasculature (van Sloun et al. 2018). However, the clinical side effects of the CA, which was employed as the tracer on the human body, remain controversial. In this study, we trained DL-SRU with RBC synthetic images for clinical practice by conducting localization on real US images of RBCs without CAs. The main merit of DL-SRU is that the deep learning architecture can be trained with a synthetic training data set that fully reflects the original US speckle features. Thus, many training data sets were generated for various imaging conditions, and they were exploited to achieve more robust and accurate localization, compared with other ULM techniques. In other words, RBCs can be localized by the present neural network by training well-synthesized RBC speckle images under challenging *in vivo* conditions, such as low frame rate, high concentration and unintended organ motion. In addition, it takes a very short time to localize the speckle signal with low computational cost with the aid of pre-training for the neural network. Finally, the localized results are used for analysis of flow dynamics.

CONCLUSIONS

The proposed DL-SRU successfully recovered RBC localization from clinical US images without the injection of CAs. We mapped the super-resolved vasculature and measured its flow dynamics. DL-SRU was combined with the two-frame PTV algorithm to develop new STV. STV could effectively obtain the flow dynamic information of vasculature with high spatial resolution.

Given that DL-SRU was trained with the synthetic training data sets obtained under various imaging conditions and achieved accurate and fast localization of real tracers, it could be used in various biomedical applications including clinically difficult-to-diagnose cases. In addition, it could be combined with a flow measurement modality for blood flow analysis. This study serves as a stepping-stone in the field of clinical US imaging because of its robust, fast and accurate RBC localization, compared with other ULM techniques. Moreover, it might broaden the deep learning application to the field of biomedical imaging, because the deep learning-based algorithm illustrated that deep learning training with synthetic data sets can be used for real image segmentation.

Acknowledgments—This research was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIP) (NRF-2017 R1 A2 B3005415).

Conflict of interest disclosure—The authors declare no competing interests.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.ultrasmedbio.2019.12.002.

REFERENCES

Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans Pattern Anal Mach Intell 2017;39:2481–2495.

Baek S, Lee S. A new two-frame particle tracking algorithm using match probability. Exp Fluids 1996;22:23–32.

Bar-Zion A, Tremblay-Darveau C, Solomon O, Adam D, Eldar YC. Fast vascular ultrasound imaging with enhanced spatial resolution and background rejection. IEEE Trans Med Imaging 2017;36:169–180.

Bar-Zion A, Solomon O, Tremblay-Darveau C, Adam D, Eldar YC. SUSHI: Sparsity-based ultrasound super-resolution hemodynamic imaging. IEEE Trans Ultrason Ferroelectr Freq Control 2018;65:2365–2380.

Chen H, Zhang Y, Kalra MK, Lin F, Chen Y, Liao P, Zhou J, Wang G. Low-dose CT with a residual encoder-decoder convolutional neural network. IEEE Trans Med Imaging 2017;36:2524–2535.

Cheng Y, Wang D, Zhou P, Zhang TA, survey of model compression and acceleration for deep neural networks. arXiv preprint arXiv 1710.09282. 2017.

Demené C, Deffieux T, Pernot M, Osmanski BF, Biran V, Gennisson JL, Sieu LA, Bergel A, Franqui S, Correas JM. Spatiotemporal clutter filtering of ultrafast ultrasound data highly increases Doppler and fultrasound sensitivity. IEEE Trans Med Imaging 2015;34:2271–2285.

Dencks S, Piepenbrock M, Opacic T, Krauspe B, Stickeler E, Kiessling F, Schmitz G. Clinical pilot application of super-resolution US imaging in breast cancer. IEEE Trans Ultrason Ferroelectr Freq Control 2019;66:517–526.

Desailly Y, Pierre J, Couture O, Tanter M. Resolution limits of ultrafast ultrasound localization microscopy. Phys Med Biol 2015;60:8723.

Dijkmans P, Visser C, Kamp O. Adverse reactions to ultrasound contrast agents: Is the risk worth the benefit? Eur J Echocardiogr 2005;6:363–366.

Errico C, Pierre J, Pezet S, Desailly Y, Lenkei Z, Couture O, Tanter M. Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging. Nature 2015;527:499–502.

Feingold S, Gessner R, Guracar IM, Dayton PA. Quantitative volumetric perfusion mapping of the microvasculature using contrast ultrasound. Invest Radiol 2010;45:669–674.

Foiret J, Zhang H, Ilovitsh T, Mahakian L, Tam S, Ferrara KW. Ultrasound localization microscopy to image and assess microvasculature in a rat kidney. Sci Rep 2017;7:13662.

Gessner RC, Frederick CB, Foster FS, Dayton PA. Acoustic angiography: A new imaging modality for assessing microvasculature architecture. J Biomed Imaging 2013;2013:14.

Gharahi H, Zambrano BA, Zhu DC, DeMarco JK, Baek S. Computational fluid dynamic simulation of human carotid artery bifurcation based on anatomy and volumetric blood flow rate measured with magnetic resonance imaging. Int J Adv Eng Sci Appl Math 2016;8:46–60.

Kerr TM, Cranley JJ, Johnson JR, Lutter KS, Duldner JE, Sampson MG. Measurement of blood flow rates in the lower extremities with use of a nuclear magnetic resonance based instrument. J Vasc Surg 1991;14:649–657.

Lee SJ, Park JH, Kim JJ, Yeom E. Quantitative analysis of helical flow with accuracy using ultrasound speckle image velocimetry: In vitro and in vivo feasibility studies. Ultrasound Med Biol 2017;44:657–669.

Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. Proc IEEE Conf Comput Vis Pattern Recognit 2015;3431–3440.

Nam KH, Yeom E, Ha H, Lee SJ. Velocity field measurements of valvular blood flow in a human superficial vein using high-

frequency ultrasound speckle image velocimetry. Int J Card Imaging 2012;28:69–77.

Nehme E, Weiss LE, Michaeli T, Shechtman Y. Deep-STORM: Super-resolution single-molecule microscopy by deep learning. Optica 2018;5:458–464.

Ovesný M, Křížek P, Borkovec J, Švindrych Z, Hagen GM. ThunderSTORM: A comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. Bioinformatics 2014;30:2389–2390.

Reza AM. Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. J VLSI Signal Process Syst Signal Image Video Technol 2004;38:35–44.

Ronneberger O, Fischer P, Brox T. 2015 U-net: Convolutional networks for biomedical image segmentation. In: Proceedings, International Conference on Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015. Lecture Notes Comput Sci 9351, 234–241.

van Sloun RJ, Solomon O, Bruce M, Khaing ZZ, Wijkstra H, Eldar YC, Mischi M. Super-resolution ultrasound localization microscopy through deep learning. arXiv preprint arXiv 1804.07661. 2018.

Xu B, Wang N, Chen T, Li M. Empirical evaluation of rectified activations in convolutional network. arXiv preprint arXiv 1505.00853. 2015.