Localization free super-resolution microbubble velocimetry using a long short-term memory neural network

Xi Chen[1,2,§], Matthew R. Lowerison[1,2,§], Zhijie Dong[1,2], Nathiya Chandra Sekaran[1,3], Chengwu Huang[4], Shigao Chen[4], Timothy M. Fan[5,6], Daniel A. Llano[1,3], and Pengfei Song[1,2,6*]

[1] Beckman Institute for Advanced Science and Technology, University of Illinois Urbana-Champaign, Urbana, IL
[2] Department of Electrical and Computer Engineering, University of Illinois Urbana-Champaign, Urbana, IL
[3] Department of Molecular and Integrative Physiology, University of Illinois Urbana-Champaign, Urbana, IL
[4] Department of Radiology, Mayo Clinic College of Medicine and Science, Mayo Clinic, Rochester, MN
[5] Department of Veterinary Clinical Medicine, University of Illinois Urbana-Champaign, Urbana, IL
[6] Cancer Center at Illinois, University of Illinois Urbana-Champaign, Urbana, IL
[§] These authors contributed equally to this work

*Corresponding Author:
Dr. Pengfei Song
Department of Electrical and Computer Engineering
Beckman Institute for Advanced Science and Technology
University of Illinois at Urbana-Champaign
405 N. Mathews Ave.
Urbana, IL 61801
Email: songp@illinois.edu

Abstract

Ultrasound localization microscopy is a super-resolution imaging technique that exploits the unique characteristics of contrast microbubbles to side-step the fundamental trade-off between imaging resolution and penetration depth. However, the conventional reconstruction technique is confined to low microbubble concentrations to avoid localization and tracking errors. Several research groups have introduced sparsity- and deep learning-based approaches to overcome this constraint to extract useful vascular structural information from overlapping microbubble signals, but these solutions have not been demonstrated to produce blood flow velocity maps of the microcirculation. Here, we introduce Deep-SMV, a localization free super-resolution microbubble velocimetry technique, based on a long short-term memory neural network, that provides high imaging speed and robustness to high microbubble concentrations, and directly outputs blood velocity measurements at a super-resolution. Deep-SMV is trained efficiently using microbubble flow simulation on real *in vivo* vascular data and demonstrates real-time velocity map reconstruction suitable for functional vascular imaging at super-resolution. The technique is successfully applied to a wide variety of imaging scenarios, include flow channel phantoms, chicken embryo chorioallantoic membranes, and mouse kidney, tumor, and brain imaging.

Introduction

Ultrasound localization microscopy (ULM) is a super-resolution imaging technique that relies on the localization and tracking of intravascular microbubble (MB) contrast agents to reconstruct microvasculature. It has been demonstrated by multiple research groups that MBs can act as acoustic point sources that can be localized at a sub-diffraction precision, achieving micrometer-scale vascular fidelity at clinical frequencies without sacrificing imaging depth (1–5). Furthermore, a unique feature of ULM is the ability to measure the blood flow velocity of the microvasculature, which serves as a sensitive biomarker for numerous physiological and pathological states. The ULM MB localization protocol side steps a fundamental trade-off between imaging resolution, which is dictated by wavelength, and imaging penetration depth, which is limited by attenuation. However, ULM introduces a new consideration between MB localization precision and acquisition time. High fidelity MB localization requires non-overlapping MB signals; but the dilution of MBs used to achieve spatially sparse signals necessitates a long imaging duration to ensure the full perfusion of all patent vasculature (6, 7). This exasperates the pragmatic challenges associated with super-resolution vascular reconstruction, including sources of tissue motion and time-varying changes in vascular flow, which severely limit the clinical application of the technology. High concentration MB injections will perfuse through more of the vasculature over the same duration, permitting shorter imaging acquisition windows but at the expense of MB localization precision with increased spurious events and erroneously reconstructed features. Rectifying these two contradictory requirements for ULM is an area of ongoing and active research for super-resolution ultrasound.

Several ULM approaches have been specifically introduced to handle high-concentration contrast injections overlapping MB signals. Sparse recovery strategies (8, 9), such as sparsity-based ultrasound super-resolution hemodynamic imaging (SUSHI), utilize a structurally sparse prior-representing vasculature to resolve overlapping MB point-spread functions (PSF). This

localization-free approach improves imaging speed but cannot measure blood flow velocity. Compressive sensing algorithms have demonstrated the ability to resolve high density MB localizations (10, 11) and deconvolution-based approaches can iteratively shrink the imaging PSF, resulting in better localization accuracy for overlapping MB signals (12, 13). These approaches improve localization accuracy and can generate high-fidelity microvascular density maps but sacrifice the ability to track individual flowing MB trajectories. Deep-learning (DL) solutions are also gaining popularity for ULM imaging, with numerous classes of neural network architecture and training data generation showing promise in improving localization accuracy, particularly for high MB concentrations, and for accelerating the time-consuming ULM processing pipeline. Specifically, van Sloun *et al.* (14) and Liu *et al.* (15) used neural network architectures to extract point target locations representing MB centroids from B-mode images in the spatial domain. Lok *et al.* (16) and Milecki *et al.* (17) used spatiotemporal data and a fully convolutional network to improve localization confidence. These proposed DL solutions directly generate super-resolved microvessel maps but, as with the above approaches, have not been demonstrated to produce blood flow velocity maps. Thus, despite these promising results, a true high-speed super-resolution technology that can handle high concentration MB injections and can offer all the features of conventional ULM (that is, microvascular structure and velocity) has not been previously reported.

Here we introduce DL-based super-resolution microvessel velocimetry (Deep-SMV), a localization free approach that provides high imaging speed and robustness to high MB concentrations, and directly outputs blood velocity measurements at a super-resolution. Deep-SMV leverages the rich spatiotemporal information embedded in ultrasound MB imaging data with a convolutional neural network including long short-term memory (LSTM) blocks to provide both structural super-resolution imaging and time-resolved velocimetry of tissue vasculature. LSTM is a type of recurrent neural network (RNN) architecture (18, 19), which is typically characterized as

neural networks with a "feedback loop" mechanism that uses previous output as input and maintains internal hidden states. Compared to using feedforward networks for temporal inference by treating the temporal dimension as an extra spatial dimension, RNNs are more suited in learning temporal features because they ensure that temporal information flows in the direction of the progression of time. By keeping internal hidden states that are constantly updated with new input, RNNs can better retain temporal information. LSTM addresses the vanishing gradient problem in classical RNNs by introducing gates to modulate information flow within the network. RNNs and LSTM have been successfully applied to various tasks such as machine translation, object tracking and motion prediction (20–22). In this implementation, Deep-SMV was trained to recognize diffraction-limited MB signals, ranging from spatially sparse to heavily overlapping conditions, to generate super-resolved microvascular structural and velocity maps. By eliminating the need for the highly inefficient process of MB localization and tracking, Deep-SMV provides effective and fast super-resolution imaging by maximizing the spatiotemporal information from the input data. We first evaluate the performance of Deep-SMV using *in vitro* flow channel phantoms and then on *in vivo* ultrasound MB data taken from the chorioallantoic membrane (CAM) of chicken embryos, which provides an optical reference standard. We then demonstrate Deep-SMV's ability to handle more complex microvascular dynamics by applying the technique to super-resolution velocimetry of the mouse kidney, tumor, and brain.

Results

Deep-SMV architecture efficiently uses spatiotemporal MB data

The LSTM-UNet structure of the Deep-SMV model (**Figure 1**, **Supplementary Figure 1**) utilizes LTSM blocks in the UNet bottleneck layers to take advantage of the spatiotemporal information present in the input ultrasound data. The feature extraction path of the conventional UNet architecture functions to generate spatial feature maps from each frame of the input sequence without awareness of their temporal relations. The LSTM units (**Supplementary Figure 1, top**) perform spatiotemporal inference on the sequence of temporally contiguous features, generating feature maps consisting of both spatial and temporal information. Finally, the reconstruction block of the UNet functions to perform predictions based on the spatiotemporal feature output by the LSTM units. This approach allows for the Deep-SMV network to directly train for the output of velocity magnitude and flow direction from the raw spatiotemporal features, without explicitly relying on inefficient MB localization and tracking processes. This permits the network to use high concentration and overlapping MB signals more efficiently for large vascular lumens while retaining high fidelity for microvascular reconstruction.

An example of *in vivo* informed data simulation for Deep-SMV training is demonstrated in **Figure 2**. CAM optical images (**Figure 2A**) provide a physiologically relevant *in vivo* vascular

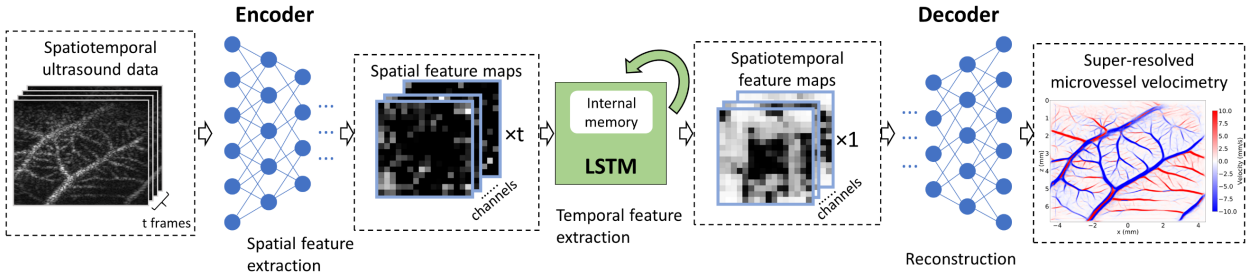


Figure 1. Deep-SMV data workflow. The Deep-SMV network takes a spatiotemporal (2D spatial + time) input of ultrasound data, passes this through encoder blocks to extract spatial features, uses LSTM blocks to extract temporal features, and then reconstructs a super-resolved velocimetry map from the spatiotemporal features via decoder blocks. Further details are presented in Supplementary Figure 1.

structure that can be readily segmented (**Figure 2B**) to generate binarized representations of the vessel space. The process of medial axis skeletonization can then yield structural connectivity maps (**Figure 2C**) as well as distance transform maps (**Figure 2D**) of this vessel bed, which is informative of the local vascular diameter. This serves as the basis for an undirected vessel graph for MB flow simulation (**Figure 2E**), which is assigned direction using a breadth-first search (BFS) traversal algorithm (**Figure 2F**). The resulting MB locations were fed into a Field-II (23, 24) simulation (see **Figure 2G**, detailed in **Methods**) to generate training samples of 16 frames in the temporal dimension, and 256 * 256 pixels in the spatial dimension. The loss of the NN on the training set and the validation set gradually decreases as a function of time, showing no sign of overfitting until around 75 epochs, when the validation loss plateaus, and training loss continues to show minor improvement. The trained network was tested on 32 simulation samples of 16 *

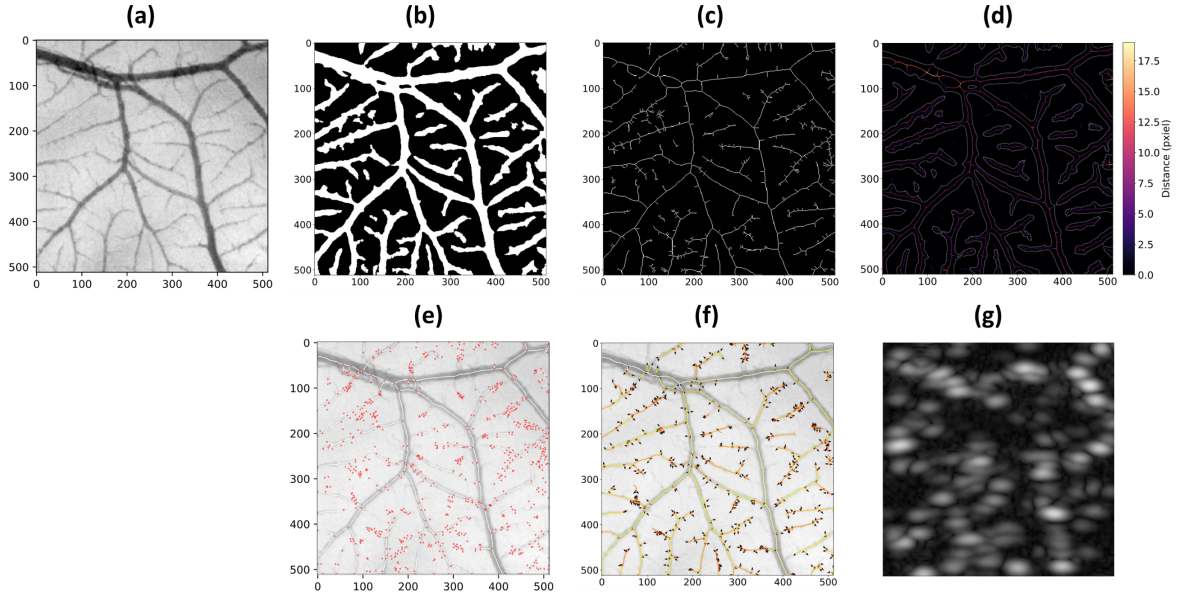


Figure 2. MB flow simulation procedure. (A) Grayscale image of the green channel pixel intensity of a 512-by-512 pixels region in a CAM optical image. **(B)** Binary segmentation result of (A) by applying adaptive thresholding. (**C**) Skeleton of (B) obtained by applying medial axis skeletonization on (B). **(D)** Distance transform of (B) obtained by applying medial axis skeletonization on (B). **(E)** Undirected graph constructed from the skeleton (B), displayed on top of the original CAM image. Lines represent edges and dots represent vertices. **(F)** Final directed vessel graph with arrows indicating flow direction on each vessel segment. **(G)** Example frame of Field-II ultrasound simulation from MB flow simulation.

256 * 256 simulation data. It was able to achieve a mean absolute prediction error of 23.24 degrees for flow direction prediction, and 1.59 mm/s for flow velocity prediction.

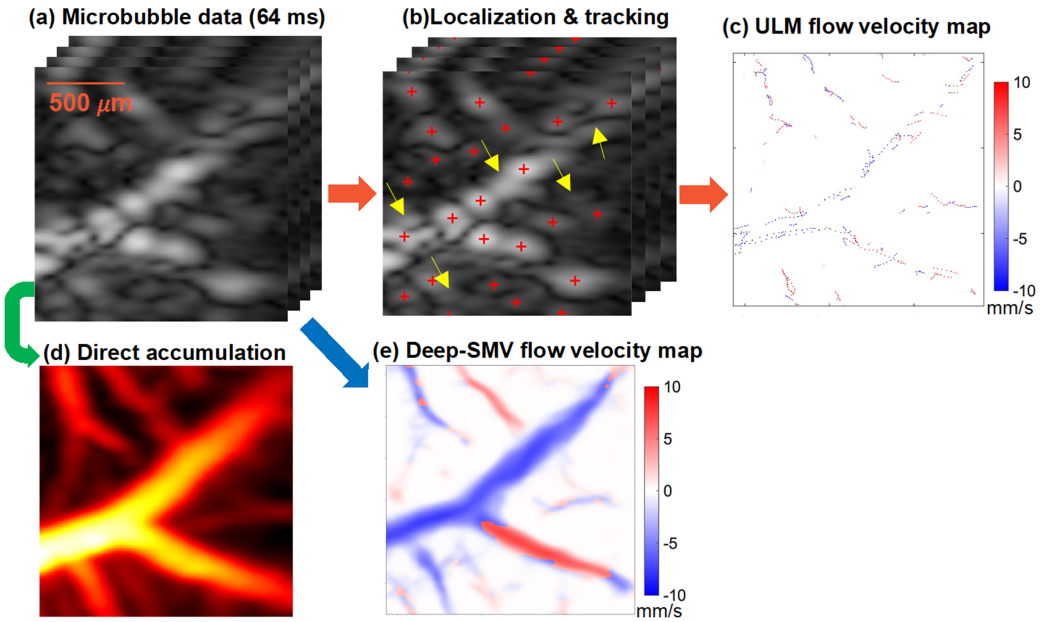


Figure 3. Conventional ULM and Deep-SMV comparison on short data segments. (A) An example short data segment of MB data (64 ms) was processed using **(B)** conventional ULM localization and tracking, resulting in **(C)** a very sparse flow velocity map where it is difficult to visualize the microvessel anatomy. **(D)** Direct accumulation of MB signal data reveals diffraction-limited vessel structure. **(E)** Processing the same data segment with Deep-SMV allowed for super-resolution velocity estimation with more structural connectivity than conventional ULM. The red color in (C) and (E) indicates flow toward the transducer, and blue is flow away from the transducer.

Deep-SMV outperforms conventional ULM in short data segments

The use of Deep-SMV for the generation of super-resolved velocity maps takes advantage of the fast forward-processing speed of NNs and GPU-based parallelized computing to achieve real-time super-resolution imaging capability. The example 300 x 300-pixel image patch with 64 temporal frames in **Figure 3A** took approximately 2,500 ms to undergo the conventional ULM localization and tracking steps (**Figure 3B**) to produce the sparsely populated super-resolution flow velocity map shown in **Figure 3C**. The direct accumulation of the MB signals, in **Figure 3D**, demonstrates that most of the vascular luminal space is missed in conventional ULM for a short data segment due to the inefficient MB localization process. Applying Deep-SMV to the same

data segment yielded a super-resolution flow velocity map (**Figure 3E**) which only took 28 ms to process on a NVIDIA® GeForce® RTX 2080 Ti GPU, a near 100X acceleration compared to conventional ULM. The magnitude and direction of flow estimated via Deep-SMV is comparable to the point estimates derived from ULM; however, it is notable that the proportion of vascular lumen space that was reconstructed is much higher than conventional processing. The convention used in all results presented is that the red color indicates flow toward the transducer, and blue indicates flow away from the transducer.

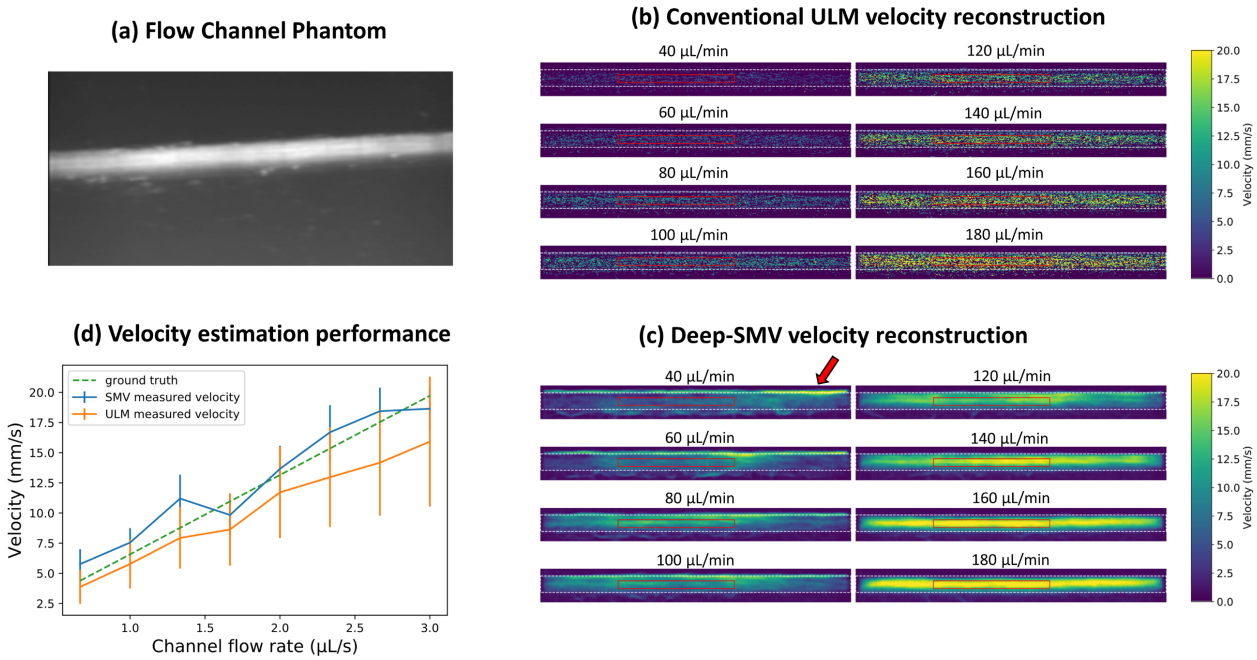


Figure 4. Flow channel phantom validation. (A) A contrast power Doppler image of the experimental flow phantom. **(B)** Conventional ULM localization and tracking was able to sparsely populate the flow channel with velocity estimates. Due to the inefficient use of MB signals, a much longer acquisition time would be required to completely scout out the channel lumen. **(C)** Deep-SMV was also able to provide velocity estimates for this flow channel and was much more efficient in using the MB signal, resulting in a higher proportion of reconstructed luminal space. **(D)** The performance of velocity estimates for different flow volume rates between ULM and Deep-SMV, in comparison to the theoretical gold-standard. ULM demonstrated an underestimation bias for velocity estimates, particularly for high flow volume rates. The analysis ROIs were placed to avoid the upper edge enhancement artifact from buoyant MBs at low flow volume rates (arrow).

Flow channel phantom validates Deep-SMV reconstructions

We then tested the performance of Deep-SMV on an *in vitro* flow channel phantom (**Figure 4**) where the ground-truth velocity is known from the flow volume rate of the connected syringe pump. A contrast power image of the flow channel (**Figure 4A**) demonstrates that the flow channel was well perfused with MB contrast agent, yielding a high MB concentration scenario where conventional ULM does not perform efficiently. This is apparent in the conventional ULM velocity reconstructions (**Figure 4B**) where MB localization and tracking results in relatively sparse tracing of the flow channel lumen. Deep-SMV was also used to process the same dataset (**Figure 4C**). Deep-SMV demonstrated a better realization of the parabolic flow profile within the flow channel phantom with a much higher perfusion efficiency. In comparison to the mass-conservation calculated ground-truth (**Figure 4D**), conventional ULM showed an underestimation bias for velocity estimates, particularly for high flow volume rates, whereas Deep-SMV was more consistent across all tested flow velocities. The analysis ROIs were placed to avoid the upper edge enhancement artifact from buoyant MBs at low flow volume rates (arrow). Each 1,600-frame flow volume dataset took approximately 442 seconds to process using conventional ULM without MB separation (25), and only 190 seconds to process using Deep-SMV. These processing time benchmarks include SVD filtering and data transfer times.

Chicken embryo CAM velocimetry reveals super-resolved vascular pulsatility

After demonstrating good performance in the flow phantom experiment, we then sought to test Deep-SMV on real *in vivo* contrast ultrasound data. **Figure 5** demonstrates Deep-SMV and conventional ULM results taken from the surface vasculature of the CAM of chicken embryos. We have previously shown that the CAM is an excellent model for testing ultrasound super-resolution processing methods (25–27), providing low attenuation and minimal tissue motion that is accessible to optical imaging to serve as a reference standard. Conventional ULM processing

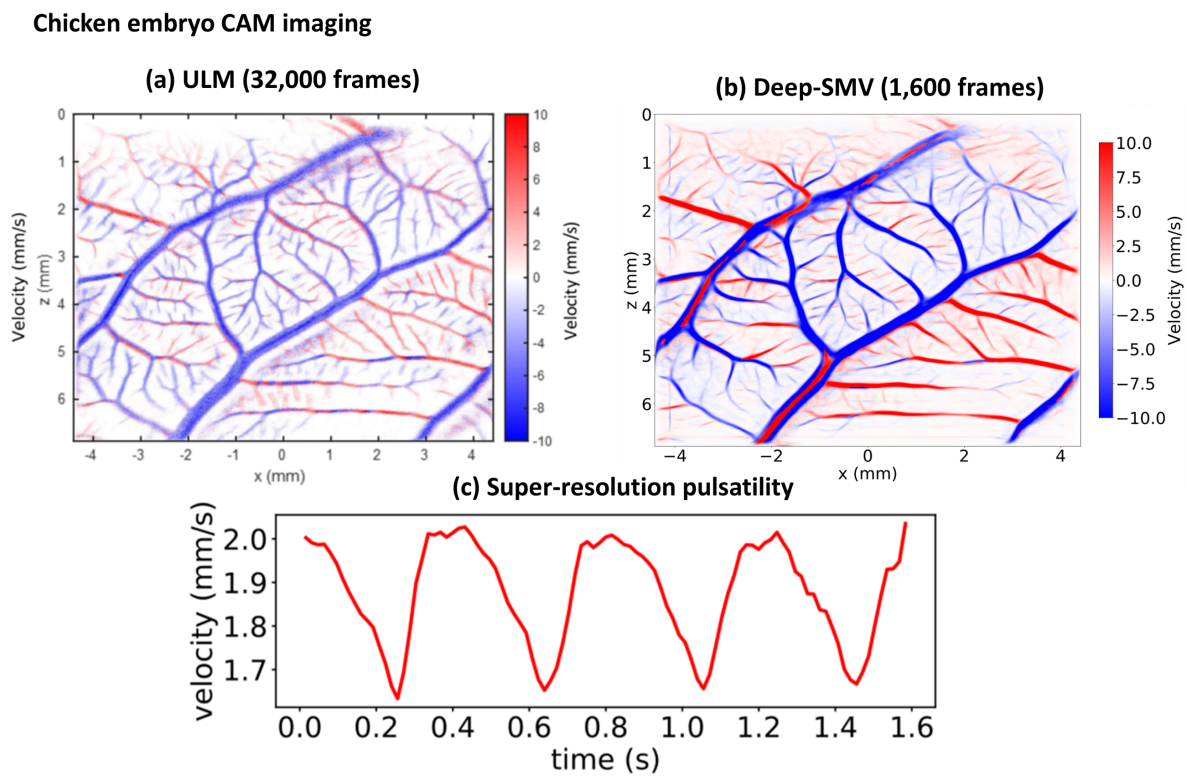


Figure 5. Chicken embryo CAM vessel velocimetry. (A) ULM reconstruction using 32,000 frames of total data, in comparison to **(B)** Deep-SMV applied to the CAM surface vasculature using 16 frames of data (16 ms), accumulated over a total duration of 1.6 seconds. **(C)** Deep-SMV achieved a 16ms temporal resolution, revealing a cyclic pulsatility in the blood velocity which is not visible in conventional ULM processing.

(**Figure 5A**) reveals an interdigitated vessel structure that is characteristic of the CAM vascular topology. The small branching arterioles on the order of 20-30 μm in diameter, which are adjacent to the main site of gas exchange, are visible in the ULM reconstruction of this planar vascular membrane. A total of 32,000 frames (32 seconds of acquisition data) were used to generate this super-resolution image. The ULM processing time was approximately 5,400 seconds for each 1,600-frame acquisition, including the application of MB separation filters, SVD filters, MB pairing and tracking, and data accumulation. The final accumulated image of 20 acquisitions took several hours to process. In comparison, the Deep-SMV image of this CAM network (**Figure 5B**), generated by accumulating multiple 16ms subsets of data for a total duration 1.6 seconds, used only a twentieth of the data (1,600 frames) and took 294 seconds to process. This Deep-SMV result shows comparable performance to the much more expensive ULM algorithm with the added

advantage of providing a higher temporal resolution. As demonstrated in **Figure 5C**, Deep-SMV revealed a cyclic pulsatility in the blood flow velocity that was not visible in conventional ULM processing. The frequency of the change in blood velocity was around 2-3 Hz, which matches with the power Doppler fluctuation, and is close to the reported values for the cardiac cycle of chicken embryos at this stage of development (28).

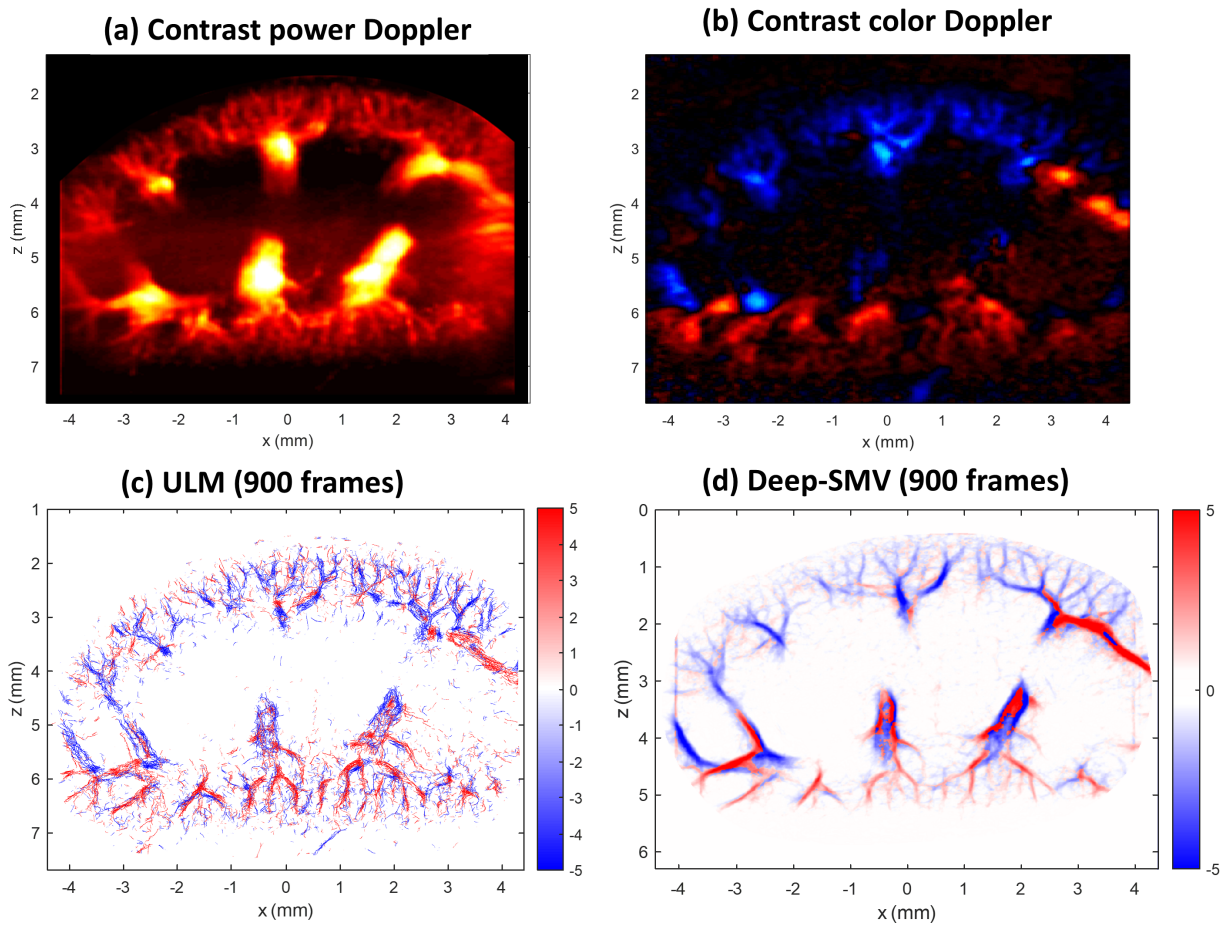


Figure 6. Mouse kidney results. (A) Contrast power Doppler imaging of the mouse kidney. **(B)** Color flow Doppler demonstrates the directional flow of the larger renal arterioles and venules. **(C)** Conventional ULM imaging was able to detect some of the smaller vessels in this imaging cross section, with relatively poor performance in the larger vessels. **(D)** Deep-SMV demonstrates directional velocity of both the large and small vessels, with better visualization of the luminal space of the hepatic arteries.

Deep-SMV demonstrates hierarchical in vivo vasculature

Contrast power Doppler of the mouse kidney (**Figure 6A**) demonstrates a typical renal anatomical vascular feature of large, deep feeding vessels that branch outward into the medulla as a fine, densely packed microvascular bed approaching the renal cortex. Applying color Doppler processing to this dataset (**Figure 6B**) reveals the directionality of the blood flow but lacks fidelity for the smallest vessels evident on the power Doppler image. Conventional ULM processing was able to capitalize on the relatively low MB concentration in this dataset to reconstruct microvessel velocity flow information (**Figure 6C**) but did not perform as well for the larger vessels in the FOV where the local MB concentration was higher. It took approximately 7,200 seconds of processing time to apply the MB separation filters, localize and track MBs, this 900-frame dataset (respiratory gating was applied to remove high tissue motion from three 720-frame acquisitions). In comparison, Deep-SMV only took 135 seconds to generate the image presented in **Figure 6D**. The Deep-SMV image also demonstrated super-resolved velocity estimation with a higher fidelity than color Doppler imaging, while outperforming ULM in the larger vessels of the mouse kidney. Deep-SMV was able to better handle the overlapping MB signals in the large renal vessels to fill in the vascular luminal space more appropriately. Similar findings are presented for a mouse Hepa1-6 tumor in **Figure 7**. Contrast power Doppler demonstrated high peripheral vasculature (**Figure 7B**) and color Doppler processing was only able to detect the larger vessels (**Figure 7C**). Conventional ULM reconstructed a very sparse vascular bed (**Figure 7D**) using a dataset of 1,600 frames and took over 9100 seconds to process. In comparison, Deep-SMV produced a superior visualization of larger vessels and could better depict the smaller tortuous microvessels in this tumor (**Figure 7E**). Deep-SMV processing took approximately 580 seconds to accumulate this image.

Finally, we applied Deep-SMV to the extremely complex vascular structure that is present in the brain. This complexity is evident in the spatiotemporal MB data presented in **Figure 8A**.

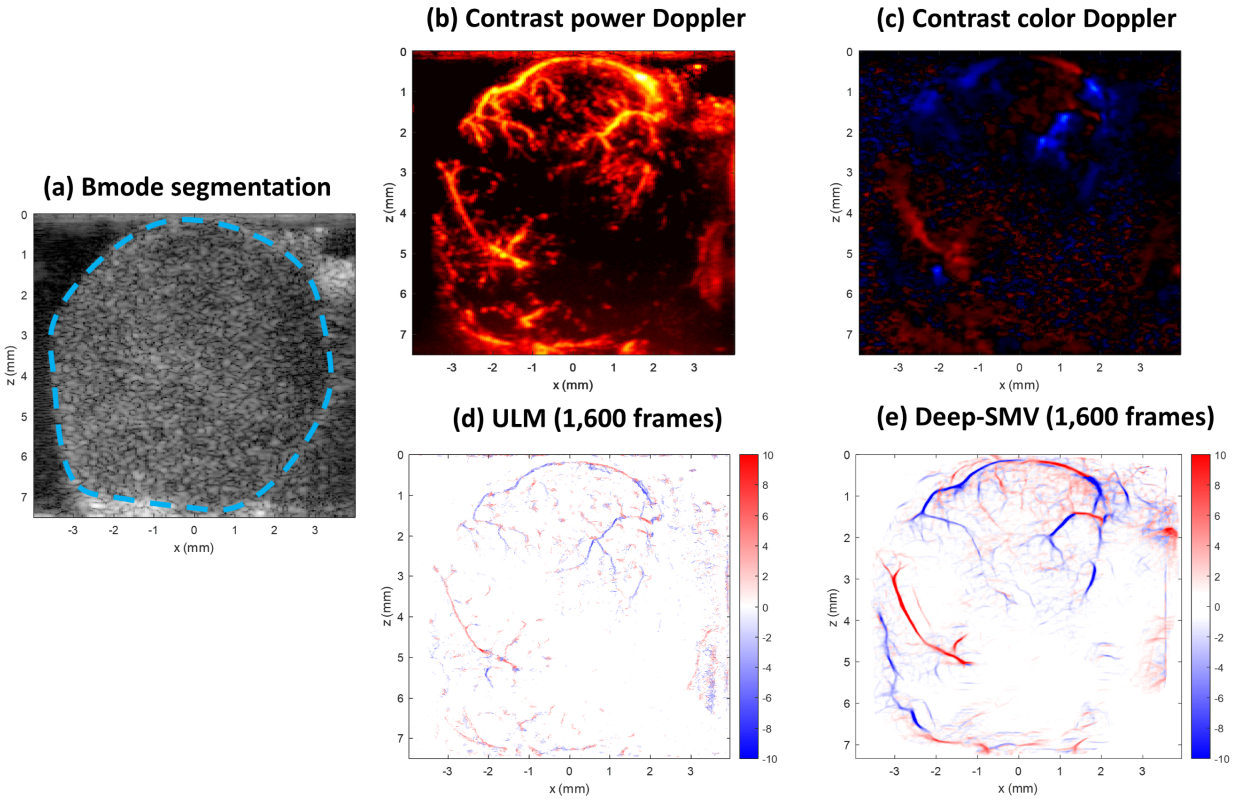


Figure 7. Mouse Hepa1-6 tumor results. (A) Manual segmentation of tumor boundary. **(B)** Hepa1-6 tumors demonstrated relatively high vessel density in the periphery with reduced intratumoral vascularization. **(C)** Color Doppler could only detect the larger vessels, performing poorly in the deeper regions of the tumor. **(D)** ULM imaging of this acquisition demonstrates relatively sparse vascular reconstruction in comparison to **(E)** Deep-SMV processing which had superior visualization of larger vessels and could demonstrate better connectedness in the smaller tortuous microvessels.

The densely packed and hierarchical vascular organization of the mouse brain presents a substantial challenge for super-resolution vascular imaging, typically requiring very low concentration MB injections and prohibitively long data accumulation times to accurately reconstruct cerebral features. A case point for this difficulty is evident in the example shown in **Figure 8B**, a short duration (1,600 frame) ULM reconstruction of the brain vasculature. This example exemplifies the graduality of ULM accumulation, with the sparsity of the reconstruction evident in both the columnar cortical vessels and the subcortical supply vasculature. Sporadic MB localization errors are also evident, particularly in the largest vessels, leading to spurious velocity estimates which bias the final velocity map. In comparison, the Deep-SMV velocity map (**Figure**

8C) demonstrates multiple scales of vasculature, with obvious branching connectivity ranging from the interior feeding arterioles to the cortical capillary layers. The processing time for this 1,600-frame acquisition was approximately 3,540 seconds for the ULM pipeline, and 167 seconds for Deep-SMV.

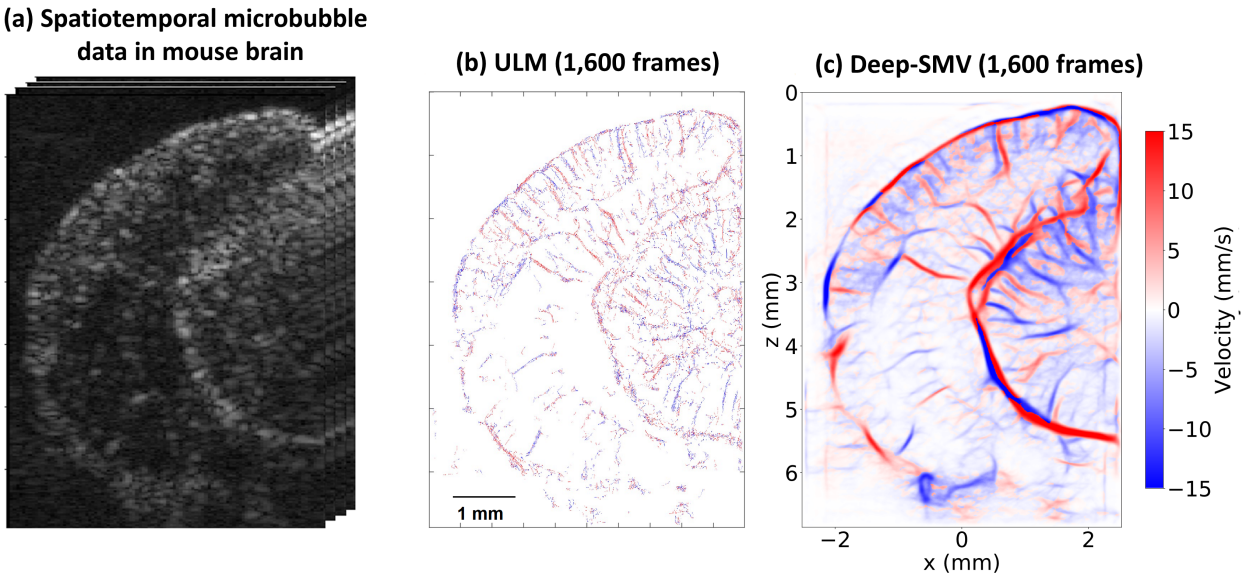


Figure 8. Mouse brain results. (A) The SVD-filtered spatiotemporal MB data from the mouse brain demonstrates a complex flow pattern with dense hierarchical vascularization. (B) Applying conventional ULM to 1,600 frames of mouse brain data yielded sparsely reconstructed vascular trajectories. (C) Deep-SMV applied to the same dataset was better able to reconstruct the vascular lumen of brain vasculature, particularly in the deeper regions of the brain.

Discussion

In this study, we introduced a DL-based super-resolution microvessel velocimetry approach for contrast-enhanced ultrasound imaging, named Deep-SMV, by employing LSTM blocks in the bottleneck layers of a UNet architecture to take advantage of the rich spatiotemporal information that is present in ultrafast ultrasound MB data. Using this architecture, we can bypass the extremely expensive MB localization and tracking processing in conventional ULM to provide super-resolved blood velocity measurements that are more robust to high MB concentrations and at a high imaging speed. The training data for the Deep-SMV network was generated using ultrasound MB flow simulation on *in vivo* optical images taken from the CAM of chicken embryos. The technique was validated using a flow channel phantom, where less velocity estimation bias in comparison to ULM was noted, and we demonstrated that Deep-SMV can produce super-resolved velocity maps for several challenging *in vivo* imaging scenarios, including chicken embryo CAM imaging, mouse kidney imaging, mouse tumor imaging, and mouse brain imaging.

Previous research groups have applied CNNs to the problem of extracting useable ULM localization data from overlapping MB signals (14–17); however, none of these approaches yield velocity information. While it is possible to perform inference on spatiotemporal data with a conventional CNN by treating the temporal dimension as multiple channels or as an additional spatial dimension, the LSTM-based spatiotemporal model has several advantages over conventional CNN approaches. The scale of temporal perception of a multi-channel CNN or 3D CNN is limited by the number of frames in a single input sequence, which is often subject to memory constraints. Longer sequences need to be partitioned into shorter segments to accommodate the CNN input size. The CNN will not form any connection between different input segments, which means that it can only detect local temporal features. LSTM, on the other hand, can handle arbitrary length inputs. With the help of internal hidden states, LSTM can recall

information obtained from previous inputs for a long time, making it more suited for the task of velocity estimation.

We also demonstrated that the Deep-SMV approach substantially improves super-resolution reconstruction speed, particularly for regions with high vessel densities, allowing for real-time visualization of the microvascular flow dynamics. This is a particularly enticing accomplishment for super-resolution ultrasound imaging, potentially allowing the translation of the technology into functional ultrasound applications where high temporal resolution is critical for monitoring vascular response(s). Although other DL-based approaches have demonstrated rapid localization times, these solutions still rely on conventional ULM pairing and tracking algorithms to estimate velocity. The processing time burden of these algorithms scale with the MB count, making these solutions ineligible for dynamic real-time velocimetry applications. The Deep-SMV network processing time does not depend on the MB concentration, performing consistently across multiple imaging scenarios. This is evident in the performance gain of Deep-SMV in comparison to ULM for the imaging experiments presented in this manuscript: Deep-SMV modestly reduced processing time for the flow channel phantom (a low vessel density application) and yielded the best improvement for the *in vivo* cases, where the processing time requirement was reduced by an order of magnitude.

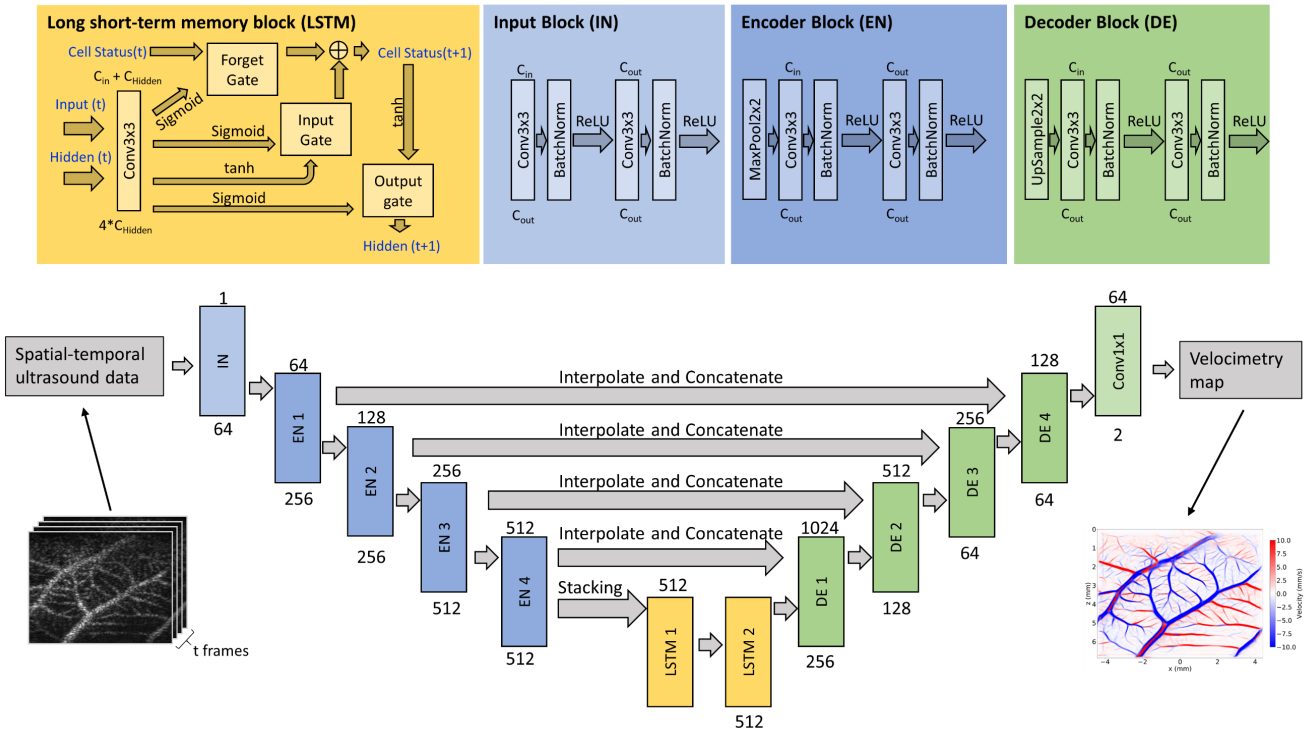
Although the Deep-SMV approach shows substantial promise over conventional ULM and competing DL solutions, the results presented in this manuscript must be understood within the context of the limitations of the study. The most critical is that there is no established gold-standard technique for measuring microvascular flow velocity to validate the output of the Deep-SMV network. As a compromise, we have used the velocity results of the much more computationally expensive ULM reconstructions as a reference standard; however, this approach comes with some caveats. As discussed, ULM processing is susceptible to spurious MB localizations which may bias the velocity point estimates. We have attempted to account for this by averaging the

ULM velocity references across a long acquisition time, which should reduce the impact of this source of error. The network was trained using Field II simulations of MB flow, where the velocity distribution was taken from experimental ULM reconstructions of vasculature. This implicitly tunes some aspects of the Deep-SMV network to the parameters used in the initial ULM reconstruction and is also dependent on the assumptions applied during vascular modeling (e.g., laminar flow). Field II also cannot simulate nonlinear MB responses. Furthermore, although Deep-SMV demonstrates better handling of high MB concentrations, it is still susceptible to signal interference of densely packed MBs leading to a retrograde flow aliasing artifact in some larger vessels. The incorrect flow direction is also apparent in ULM reconstructions, but the long acquisition duration can reduce this effect by accumulating data from the wash-out phase of contrast enhancement. We are currently exploring adjustments to the Deep-SMV output to better encode for the circular flow angle topology to help address this limitation.

In the current training setup, we used SSIM loss, which is designed to provide a similarity measurement between two images that reflects how humans perceive them. Although SSIM has been successfully applied to various image processing tasks, it is not necessarily the most efficient in capturing errors in our desired output data. We believe that a loss function specifically tailored for velocimetry can potentially improve the performance of Deep-SMV.

Deep-SMV provides efficient, effective, and fast super-resolution microvascular structural and velocity reconstruction without requiring MB localization, pairing, and tracking. By eliminating this inefficient ULM processing pipeline, Deep-SMV can capitalize on the rich spatiotemporal information present in the input IQ data, enabling real-time super-resolution imaging that is robust to a wide range of imaging conditions. This technique demonstrated a high temporal resolution, permitting the application of super-resolution velocimetry to functional imaging scenarios. By relaxing the strict imaging paradigm that ULM typically requires, Deep-SMV can facilitate the

translation of super-resolution ultrasound to a wider range of experimental and clinical applications.



Supplementary Figure 1. Deep neural network architecture for Deep-SMV. The main network design is a classic UNet structure with long short-term memory (LSTM) blocks in the bottleneck layers to provide flow velocity measurements, with input consisting of spatial-temporal (2D spatial + time) ultrasound data and output of super-resolution velocity and structure maps. The LSTM, input, encoder, and decoder block designs are provided in detail at the top of the figure. The text above each block indicates the input channel size to the unit, and the text under each block indicates the output channel size from the unit.

Author contributions

XC, MRL, and PS designed and wrote the paper. NCS and DL prepared the mouse model and performed craniotomies. XC, MRL, ZD, and PS performed the ultrasound imaging on mice. CH and SC designed the mouse kidney study and protocol, prepared the mouse kidney model, and performed the ultrasound imaging. CH and SC evaluated and preprocessed the mouse kidney ultrasound data. TMF designed and produced the mouse tumor model used in this study. ZD designed the ultrasound transducer holder and programmed the motorized imaging stage. XC, MRL, and PS developed and applied the super-resolution ULM algorithm.

Data Availability: The data that support the findings of this study are available from the corresponding authors on request.

Competing Interests: The authors declare no competing interests.

Funding: The study was partially supported by the National Cancer Institute, the National Institute of Biomedical Imaging and Bioengineering, the National Institute on Deafness and Other Communication Disorders, and the National Institute on Aging of the National Institutes of Health under grant numbers R00CA214523, R21EB030072, R21DC019473 and R03AG059103, as well as a grant to DL from the Kiwanis Neuroscience Research Foundation, and by the Seed Grant from the Cancer Center at Illinois. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. NCS is supported by a Beckman Institute Postdoctoral Fellowship.

Methods

Deep-SMV architecture and design

Our NN model adopts the convolutional LSTM-UNet structure, where the bottleneck layers of the classical UNet were replaced by LSTM layers (**Figure 1**, **Supplementary Figure 1**). The feature extraction path of the network starts with an input block that contains two convolution layers with 3*3 kernel size, each followed by a batch normalization layer and a Rectified Linear Unit (ReLU) activation function. The encoder block contains a 2*2 max pooling layer, two 3*3 convolution layers each followed by a batch normalization layer and ReLU activation function. Each encoder block reduces all spatial dimensions of its input by a factor of two and expands the feature dimension by a factor of two. Each frame of 2-D spatial input will be fed into the feature extraction blocks separately. The resulting feature maps were concatenated along the temporal dimension before being fed to the convolutional-LSTM layers as a feature map sequence.

The convolutional-LSTM block takes a new input and the previous hidden state output as it inputs, while maintaining an internal cell status variable. The inputs will first go through a 3*3 convolution layer. The convolution layer output will be split along the channel dimension into four parts, where three will go through a sigmoid activation function and operate as the forget gate, the input gate, and the output gate, respectively. A hyperbolic tangent (tanh) function will be applied to the remaining part. The forget gate determines which part of the previous cell status will be discarded. The input gate determines whether parts of the new input contribute to updating the cell status. A tanh activation function will be applied to the new cell status. Finally, the output gate determines how the cell status propagates to the output. The convolutional LSTM block will reduce the input sequence along the temporal dimension.

The decoder block is constructed similarly to the encoder block, with the max pooling layer replaced by spatial upsampling layer. Each decoder block will expand the spatial dimension of its input by 2 and reduce the channel dimension by 2. The output block is a 3*3 convolution layer

with output channel size of 2. The final output of the network will have the same spatial dimension as the input data with two channels: magnitude of the flow velocity and direction of the flow velocity with respect to the lateral direction.

We used the structural similarity (SSIM) as the loss function for training the NN, defined as:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y+c_1)(2\sigma_{xy}+c_2)}{(\mu_x^2+\mu_y^2+c_1)(\sigma_x^2+\sigma_y^2+c_2)},$$

where x, y are the two images being compared, μ_x, μ_y are the mean pixel values of the images, σ_x , σ_y are the standard deviations, and σ_{xy} is the covariance of x and y. $c_1$and c_2 are constants to enforce numerical stability.

CAM optical image-based training data simulation

<u>CAM optical image acquisition:</u> Optical imaging was performed using a Nikon SMZ800N stereomicroscope (Nikon, Tokyo, Japan) with a Plan Apo 1x objective at 1x zoom. Images were captured at 16 bits using an attached DS-Fi3 microscope camera, digitized using NIS Elements D software, and saved as .nd2 format. These were later exported as RGB .tiff files for external analysis (**Figure 2A**). The optical imaging field of view was 11.01 mm x 15.48 mm yielding 2048-by-2880 pixel images.

<u>Vessel flow dynamics model</u>: A binary vessel map of each CAM vasculature (**Figure 2B**) was segmented by applying the threshold_adaptive function available in the scikit-image python package (29) to the green channel of each acquired optical image (**Figure 2A**) as it provides the best contrast for blood vessels. Non-overlapping regions of 256-by-256 pixels were sampled from 35 2048-by-2880 pixel binarized CAM optical images, resulting in 3080 binarized ROIs, from

which 1356 regions with acceptable segmentation performance were manually selected for further processing.

Each binary vessel map was skeletonized using the medial axis skeletonization function, also available in the scikit-image package. The medial axis skeletonization function returns the skeletonized image (**Figure 2C**), as well as the distance transform map (**Figure 2D**) which can be used as an estimation of local width of the vasculature. The skeleton image is converted to an undirected graph G = {V, E}, where E is the set of edges that represents vessel segments, and V is the set of vertices that represents junctions of the vessel segments (**Figure 2E**). The graph is stored as a NetwokX (30) undirected graph. Conversion from skeleton image to NetworkX graph is achieved using the sknw function available in the ImagePy image processing framework (31).

The undirected vessel graph was then separated into a set of subgraphs, where each subgraph is a connected component of G. A spanning tree (a connected acyclic subgraph that covers all vertices in the original graph) was computed for each subgraph using the breadth-first search (BFS) algorithm starting from a source point defined as the degree-1 vertex with the largest local width in the corresponding distance map from the skeletonization step. Directions of the edges in the tree were assigned to start from the vertex that appeared earlier in the BFS traversal and point to the vertex that appeared later in the BFS traversal. The final directional vessel graph (**Figure 2F**) is a forest that contains all the directional spanning trees obtained from each connected component of the undirected vessel graph.

<u>Calculating reference velocity</u>: We used the CAM ULM result as a reference to calculate a distribution of blood flow velocity on the central line of vessels for a given radius. We apply adaptive thresholding on the ULM velocity map to generate a binary vessel map, then perform medial axis skeletonization to obtain a local vessel radius estimation for each pixel on the medial axis of the vessel structure. All local radius-velocity pairs of pixels on the medial axis were sorted

in ascending order of the local distances. The estimated reference velocity on the center line of a vessel with radius d was drawn from a normal distribution N(μ, σ^2), where μ is the average velocity of all radius-velocity pairs with radius within d±0.5 pixels range in the list of radius-velocity pairs, and σ^2 is the variance of the selected pairs.

<u>Simulating MB motion</u>**:** Once the vessel flow dynamics model was constructed, we can use it to generate a sequence of MB locations that flow in the vessel network. Each MB was stored as a tuple (e_orig, e_curr, d_ax, d_lat, amp), where e_orig is the vessel segment (edge) that it originated from, e_curr is the edge that it is currently traveling along, d_ax is the axial distance with respect to the current edge, d_lat is the lateral distance with respect to the current edge, and amp is the amplitude of the MB. The spatial location of the MB with respect to the entire field-of-view (FOV) can be easily inferred from e, d_ax and d_lat. The MB location was updated using the following rules:

- Initialization: place $n = \frac{a_e}{A} N$ MBs on each edge, where a_e is the volume of the current vessel segment, estimated using the local radius at points on the edge e, $a_e = \sum_{i \in \text{points on } e} \pi r_i^2$. Set both e_orig and e_curr to be e. d_ax and d_lat were both randomly chosen so that the MB stays within the current vessel segment. Amp was randomly selected from a uniform distribution on (0, 1]. E_orig and amp will remain unchanged until the MB is removed from the vessel network.
- For each timestep, perform the following for each MBs:
    - Calculate the axial velocity along the direction of the current vessel segment of each MB according to the Poiseuille flow condition, $v_{ax} = v_{ref} \cdot \left(1 - \frac{d_{lat}}{r}\right)$, where v_{ref} is the reference flow velocity on the centerline calculated using the method

described in the previous section, and r is the local radius of the vessel segment. Calculate the displacement $v_{ax} \cdot \Delta t$ where Δt is the temporal resolution.

- If the displacement is smaller than the remaining length of the current edge:
  - Update d_ax with the displacement.
  - Update d_lat with a number randomly generated small distance so that the MB still stays within the vessel segment.
  - E_curr remains unchanged.
- If the displacement is larger than the remaining length of the current edge:
  - If the current edge ends at a leaf vertex (i.e., no outgoing edge from the vertex), add a new MB to vessel segment e_orig of the current MB, remove the old MB from the vessel network.
  - If the current edge has outgoing adjacent edges, the MB will be moved to one of the adjacent edge e_{aj} with probability $P(e_{aj}) = \frac{\pi r_{aj}^2}{\sum_{i \in outgoing\ edges} \pi r_i^2}$, where r is the local radius of the starting vertex of each edge. Subtract the remaining length on the current edge from the total displacement. Update e_curr to e_aj. Repeat the steps above until the remaining displacement is less than the remaining length of the current edge. Then, update d_ax and d_lat with the same method in the previous step using the final remaining displacement.
- Calculate the spatial coordinates of each updated MB using e, d_ax and d_lat. Store the list of locations and amplitudes of each MB. After repeating the update steps for a desired duration of time, save the list of MB locations and amplitudes of each timestep as input to Field-II simulation.

Field-II simulation: Ultrasound simulation of MB motion was performed in Field-II (23, 24) using the sequence of MB locations and amplitudes obtained in the previous simulated MB motion step. For each timestep, we generate one frame of ultrasound simulation of point scatterers specified by the corresponding list of locations and amplitudes in the sequence. The imaging sequence was configured according to the CAM ultrasound imaging experiment setting, with a high frequency linear array transducer, 20MHz center frequency, and 9-angle plane wave compounding. Detailed simulation parameters were summarized in **Table I**.

Table I
Field-II Simulation Parameters

Variable name	Value
Transducer type	Linear array transducer
Number of transmit elements	128
Number of receive elements	128
Element height	80μm
Element width	62μm
Pitch	70μm
Transmit apodization	Kaiser
Receive apodization	Rectangle
Center frequency	20MHz
Sampling frequency	125MHz
Speed of sound	1540m/s
Number of compounding angles	9

Deep-SMV training: Training samples of 16 frames in the temporal dimension, and 256 * 256 pixels in the spatial dimension were fed to the model in mini-batches of 4 samples per batch. The model was trained with an Adam optimizer to minimize the SSIM loss between the model output and the ground truth using back-propagation with a learning rate of 0.001. The ground truth contains two channels of the magnitudes in pixels/ms and angles of the flow velocity map. To avoid complications of training with negative values in the ground truth, the angle map which

originally contains values in the range $[-\pi, \pi]$ was scaled and shifted to the scale [0, 1]. The model was trained for a total of 100 epochs.

Ethics approval

All procedures performed on mice at the University of Illinois Urbana-Champaign in this manuscript were approved by the Institutional Animal Care and Use Committee (IACUC). All experiments were performed in accordance with the IACUC guidelines. Mice were housed in an animal care facility approved by the American Association for Assessment and Accreditation of Laboratory Animal Care.

All experimental procedures involving animals at the Mayo Clinic were performed according to the Guide for the Care and Use of Laboratory Animals (NIH publication Nos. 80–23, revised 1996) and followed the institutional ethical guidelines for animal experiments. All experiments were approved by local Ethics Committee of the Mayo Clinic.

No IACUC approval was necessary to conduct the chicken embryo imaging presented in this manuscript, as avian embryos are not considered to be “live vertebrate animals” under NIH PHS policy.

Ultrasound imaging

All flow phantom, chicken embryo, and mouse brain ultrasound data were acquired with a Vantage 256 system (Verasonics Inc., Kirkland, WA) with a L35-16vX high-frequency linear array transducer. Imaging was performed with a center frequency of 20 MHz, using 9-angle plane wave compounding (1-degree increments) with a post-compounding frame rate of 1,000 Hz. Ultrasound data was saved as in-phase quadrature (IQ) datasets of 1,600 frames each for external

processing. Mouse kidney imaging was performed at a center frequency of 25 MHz, and a post-compounding frame rate of 500 Hz.

Microbubble flow phantom

A pair of holes were drilled on the opposite walls of a clear rectangular plastic container. Stainless steel dispensing needles were attached to the outside wall of the container on top of the holes using epoxy adhesive. A stainless-steel rod with a diameter of 500μm was inserted horizontally into the container through the needle openings. A 20% gelatin mixture was poured into the container until the surface was approximately 5mm above the metal tube. The phantom was placed in a 4ºC refrigerator until the gelatin completely solidified. The metal tube was removed prior to the experiment to create a flow channel, the diameter of which constricted to approximately 450μm during the gelatin solidification process. The inlet to the flow channel was connected to a programmable syringe pump (NE-300, New Era Pump Systems Inc., Farmingdale, NY) with a soft plastic tube to provide constant flow volume rate through the channel. A clinically available ultrasound contrast agent (DEFINITY®, Lantheus Medical Imaging, Inc.) was diluted 1000-fold with 0.9% saline (0.9 sodium chloride, BD, Franklin Lakes, NJ) and was perfused through the flow channel. Volume rates of 40-180μL/min with 20μL increment was used in this experiment. The L35-16vX transducer was placed at the top surface of the phantom and positioned to provide a longitudinal view of the flow channel (**Figure 4A**). For each volume rate, we waited until a stable flow was formed before starting the acquisition of 1.6s duration (1,600 frames).

Ex ovo CAM preparation

Fertilized leghorn chicken eggs (*Gallus gallus domesticus*) were provided locally by the University of Illinois Poultry Research Farm. Eggs were housed in a tilting incubator (Digital Sportsman Cabinet Incubator 1502, GQF manufacturing Inc.) until embryonic development day 4 (EDD-04). The eggshells were removed using a rotary Dremel tool and the egg contents (yolk, albumin, and embryos) were transferred into plastic weigh boats. The now *ex ovo* chicken embryos were then placed into a humidified incubator (Darwin Chambers HH09-DA) until imaging on EDD-17.

A glass capillary needle was produced using a PC-100 glass puller (Narishige, Setagaya City, Japan) and borosilicate tube (B120-69-10, Sutter Intruments, Novato, CA, USA) to aid contrast MB injection. The glass needle was attached to a 1mL syringe using Tygon R-3603 laboratory tubing, and the surface vasculature of the CAM was injected with 70 µL of DEFINITY® at clinical bolus concentration immediately prior to ultrasound imaging. The L35-16vX transducer was placed on the side of the plastic weigh boat to image the surface vasculature of the CAM (**Figure 5**). A total acquisition length of 32 seconds (32,000 frames) was acquired from the CAM.

Mouse kidney imaging

Kidney imaging was performed on an anesthetized mouse (2% isoflurane mixed with oxygen) following a 20 µL bolus injection of Bracco Lumason (Bracco, Milan, Italy) microbubble solution via tail vein. The L35-16vX was positioned to produce a long axis view (**Figure 6**). A total of 12 IQ datasets each with 720 frames of data (8,640 frames in total) corresponding to a data length of 17.28 seconds, were stored for post-processing. Contrast power images were generated by accumulating the MB signal along the temporal dimensions, and color Doppler images were generated using the technique by Loupas *et al*. (32).

Mouse tumor preparation and imaging

Athymic nude mice were inoculated with 5 x 10^6 Hepa1-6 cells subcutaneously into their right flank. After 10 days of growth, the mice were anesthetized with isoflurane and were then secured to a heated imaging platform, and a tail vein catheter was inserted to aid with MB delivery. The L35-16vx transducer was placed to produce the largest cross-section of the tumor (**Figure 7**), and then a 50uL bolus of DEFINITY® was injected before imaging. IQ data was stored as 1,600 frame acquisitions for post-processing. Contrast power and color Doppler images were produced as above.

Mouse craniotomy preparation

Anesthesia was induced by placing each CBA/CAJ (stock # 000654) mouse in a gas induction chamber supplied with 4% isoflurane mixed with medical oxygen. Then the mice were placed in a stereotaxic frame with nose cone supplying 1-2% isoflurane with oxygen for maintenance. Homoeothermic heating pads were used to maintain body temperature at 36.5°C. The surgical area was shaved and cleaned with iodine before being injected with Lidocaine (1%) for local anesthesia. Ear bars were used to secure the mouse head to the stereotaxic imaging stage. The scalp of the mouse was removed, and a cranial window was opened on the left side of the skull using a rotary Dremel tool, starting at the sagittal suture and moving laterally to expose the cerebral cortex. The tail vein of the mouse was cannulated with a 30-gauge catheter and vessel patency was confirmed with a 0.1 mL injection of sterile saline. The L35-16vX transducer was positioned to acquire a coronel imaging plane of the brain through the cranial window (**Figure 8**). A fresh 50uL bolus of DEFINITY® was injected before every 10 imaging acquisitions (1,600 frames per acquisition). A total of 80 acquisitions (128,000 frames, or 128 seconds of data) was acquired for this imaging plane.

Conventional ULM image processing and analysis

MBs were extracted for each ultrasound IQ dataset by applying a spatiotemporal singular value decomposition (SVD)-based clutter filter (25, 33–35), where the low-order singular value threshold representing tissue signal was determined adaptively (36). A noise-equalization profile was applied to account for depth-dependent effects (37). An isolated MB signal was manually identified and fit to a multivariate Gaussian function to represent the point-spread function (PSF) of the system. A MB separation filter (25) was applied to each SVD-filtered dataset, and then the imaging data was spatially interpolated to an isotropic 4.9 μm axial/lateral resolution using spline interpolation (38). MBs were then localized by applying a 2D normalized cross-correlation between the empirical PSF and each frame of interpolated IQ data. MB centroid pairing and frame-to-frame trajectory estimation was performed using the uTrack algorithm (39), or a Kalman filtering algorithm (27), where velocity was calculated from centroid displacement. Each processed data acquisition was accumulated into a final reconstruction.

References

1. C. Errico, *et al.*, Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging. *Nature* **527**, 499–502 (2015).

2. K. Christensen-Jeffries, R. J. Browning, M.-X. Tang, C. Dunsby, R. J. Eckersley, In vivo acoustic super-resolution and super-resolved velocity mapping using microbubbles. *IEEE Trans Med Imaging* **34**, 433–440 (2015).

3. T. Opacic, *et al.*, Motion model ultrasound localization microscopy for preclinical and clinical multiparametric tumor characterization. *Nat Commun* **9**, 1527 (2018).

4. O. Couture, V. Hingot, B. Heiles, P. Muleki-Seya, M. Tanter, Ultrasound Localization Microscopy and Super-Resolution: A State of the Art. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **65**, 1304–1320 (2018).

5. P. Song, *et al.*, Improved Super-Resolution Ultrasound Microvessel Imaging with Spatiotemporal Nonlocal Means Filtering and Bipartite Graph-Based Microbubble Tracking. *IEEE Trans Ultrason Ferroelectr Freq Control* **65**, 149–167 (2018).

6. V. Hingot, *et al.*, Microvascular flow dictates the compromise between spatial resolution and acquisition time in Ultrasound Localization Microscopy. *Scientific Reports* **9**, 2456 (2019).

7. K. Christensen-Jeffries, *et al.*, Poisson Statistical Model of Ultrasound Super-Resolution Imaging Acquisition Time. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **66**, 1246–1254 (2019).

8. A. Bar-Zion, C. Tremblay-Darveau, O. Solomon, D. Adam, Y. C. Eldar, Fast Vascular Ultrasound Imaging With Enhanced Spatial Resolution and Background Rejection. *IEEE Trans Med Imaging* **36**, 169–180 (2017).

9. A. Bar-Zion, O. Solomon, C. Tremblay-Darveau, D. Adam, Y. C. Eldar, SUSHI: Sparsity-based Ultrasound Super-resolution Hemodynamic Imaging. *arXiv:1712.00648 [physics]* (2017) (August 20, 2019).

10. Y. Shu, C. Han, M. Lv, X. Liu, Fast Super-Resolution Ultrasound Imaging With Compressed Sensing Reconstruction Method and Single Plane Wave Transmission. *IEEE Access* **6**, 39298–39306 (2018).

11. J. Kim, Q. Wang, S. Zhang, S. Yoon, Compressed sensing-based super-resolution ultrasound imaging for faster acquisition and high quality images. *IEEE Trans Biomed Eng* **PP** (2021).

12. Q. Chen, *et al.*, Ultrasound super-resolution imaging provides a noninvasive assessment of renal microvasculature changes during mouse acute kidney injury. *Kidney International* **98**, 355–365 (2020).

13. J. Yu, L. Lavery, K. Kim, Super-resolution ultrasound imaging method for microvasculature in vivo with a high temporal accuracy. *Sci Rep* **8**, 13918 (2018).

14. R. J. G. van Sloun, *et al.*, Super-Resolution Ultrasound Localization Microscopy Through Deep Learning. *IEEE Trans Med Imaging* **40**, 829–839 (2021).

15. X. Liu, *et al.*, Deep Learning for Ultrasound Localization Microscopy. *IEEE Transactions on Medical Imaging* **39**, 3064–3078 (2020).

16. U.-W. Lok, *et al.*, Fast super-resolution ultrasound microvessel imaging using spatiotemporal data with deep fully convolutional neural network. *Phys. Med. Biol.* **66**, 075005 (2021).

17. L. Milecki, *et al.*, A Deep Learning Framework for Spatiotemporal Ultrasound Localization Microscopy. *IEEE Trans Med Imaging* **40**, 1428–1437 (2021).

18. S. Hochreiter, J. Schmidhuber, Long short-term memory. *Neural Comput* **9**, 1735–1780 (1997).

19. J. L. Elman, Finding structure in time. *Cognitive Science* **14**, 179–211 (1990).

20. B. McCann, J. Bradbury, C. Xiong, R. Socher, Learned in Translation: Contextualized Word Vectors. *arXiv:1708.00107 [cs]* (2018) (September 22, 2021).

21. G. Ciaparrone, *et al.*, Deep learning in video multi-object tracking: A survey. *Neurocomputing* **381**, 61–88 (2020).

22. J. Martinez, M. J. Black, J. Romero, On human motion prediction using recurrent neural networks. *arXiv:1705.02445 [cs]* (2017) (September 22, 2021).

23. J. A. Jensen, Field: A Program for Simulating Ultrasound Systems: 10th Nordic-Baltic Conference on Biomedical Imaging. *Medical & Biological Engineering & Computing* **34**, 351–353 (1997).

24. J. A. Jensen, N. B. Svendsen, Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **39**, 262–267 (1992).

25. C. Huang, *et al.*, Short Acquisition Time Super-Resolution Ultrasound Microvessel Imaging via Microbubble Separation. *Scientific Reports* **10**, 1–13 (2020).

26. M. R. Lowerison, *et al.*, In Vivo Confocal Imaging of Fluorescently Labeled Microbubbles: Implications for Ultrasound Localization Microscopy. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **67**, 1811–1819 (2020).

27. S. Tang, *et al.*, Kalman Filter–Based Microbubble Tracking for Robust Super-Resolution Ultrasound Microvessel Imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 1–1 (2020).

28. A. E. Aubert, *et al.*, Heart rate and heart rate variability in chicken embryos at the end of incubation. *Exp Physiol* **89**, 199–208 (2004).

29. S. Van der Walt, *et al.*, scikit-image: image processing in Python. *PeerJ* **2**, e453 (2014).

30. A. A. Hagberg, D. A. Schult, P. J. Swart, Exploring Network Structure, Dynamics, and Function using NetworkX in *Proceedings of the 7th Python in Science Conference*, G. Varoquaux, T. Vaught, J. Millman, Eds. (2008), pp. 11–15.

31. A. Wang, X. Yan, Z. Wei, ImagePy: an open-source, Python-based and platform-independent software package for bioimage analysis. *Bioinformatics* **34**, 3238–3240 (2018).

32. T. Loupas, J. T. Powers, 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 Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **42**, 672–688 (1995).

33. Y. Desailly, *et al.*, Contrast enhanced ultrasound by real-time spatiotemporal filtering of ultrafast images. *Phys Med Biol* **62**, 31–42 (2017).

34. M. R. Lowerison, C. Huang, F. Lucien, S. Chen, P. Song, Ultrasound localization microscopy of renal tumor xenografts in chicken embryo is correlated to hypoxia. *Scientific Reports* **10**, 1–13 (2020).

35. P. Song, *et al.*, Improved Super-Resolution Ultrasound Microvessel Imaging With Spatiotemporal Nonlocal Means Filtering and Bipartite Graph-Based Microbubble Tracking. *IEEE Trans Ultrason Ferroelectr Freq Control* **65**, 149–167 (2018).

36. P. Song, A. Manduca, J. D. Trzasko, S. Chen, Ultrasound Small Vessel Imaging With Block-Wise Adaptive Local Clutter Filtering. *IEEE Transactions on Medical Imaging* **36**, 251–262 (2017).

37. P. Song, A. Manduca, J. D. Trzasko, S. Chen, Noise Equalization for Ultrafast Plane Wave Microvessel Imaging. *IEEE Trans Ultrason Ferroelectr Freq Control* **64**, 1776–1781 (2017).

38. P. Song, A. Manduca, J. D. Trzasko, R. E. Daigle, S. Chen, On the Effects of Spatial Sampling Quantization in Super-Resolution Ultrasound Microvessel Imaging. *IEEE Trans Ultrason Ferroelectr Freq Control* **65**, 2264–2276 (2018).

39. K. Jaqaman, *et al.*, Robust single-particle tracking in live-cell time-lapse sequences. *Nature Methods* **5**, 695–702 (2008).