End-to-end deep learning for tuning-free non-contrast ultrasound microvessel imaging

Ahmed Tahseen Minhaz
Department of Biomedical Engineering
Case Western Reserve University
Cleveland, USA
axm1287@case.edu

Michaela Cooley
Department of Biomedical Engineering
Case Western Reserve University
Cleveland, USA
mbc84@case.edu

Ananya Subramaniam
Departments of Biomedical Engineering
Case Western Reserve University
Cleveland, USA
axs1189@case.edu

Agata Exner, PhD
Department of Radiology and Biomedical Engineering
Case Western Reserve University
Cleveland, USA
agata.exner@case.edu

Faruk Orge, MD
Pediatric Ophthalmology and Adult Strabismus
Rainbow Babies and Children's Hospital
Cleveland, USA
Faruk.Orge@uhhospitals.org

David Wilson, PhD
Departments of Biomedical Engineering
Case Western Reserve University
Cleveland, USA
dlw@case.edu

Mahdi Bayat, PhD
Department of Electrical Engineering & Computer Science
Case Western Reserve University
Cleveland, USA
mxb871@case.edu

***Abstract*— Microvessel (MV) imaging using non-contrast US is challenging due to low blood echo SNR, the presence of spectrally overlapping tissue clutter, and noise. Spatiotemporal clutter removal, e.g., via singular value decomposition (SVD), requires parameter tuning and are computationally expensive. In this work, we proposed an end-to-end deep learning approach to remove tissue clutter and extract blood signals from US ensemble in a purely data-driven fashion. Our deep learning approach accelerates the decluttering process and provides a generalizable framework to adapt to realistic scenarios via diversly augmented data. In this preliminary study, we scanned a day 8 developing chicken embryo to collect US ensembles of different regions of the embryo. Training data was created using a fine-tuned SVD for tissue clutter removal and to train a 2D U-net. The proposed deep learning approach can reconstruct MV from US ensembles (structural similarity index of 0.83 ± 0.02 compared to processing via SVD) implying good reconstruction. Our end-to-end deep learning approach is also 2x faster than the traditional SVD approach. Deep learning based MV imaging using non-contrast US shows promise in reconstructing MV in presence of different tissue motions and can be extended to perform under various SNR, noise, and motion conditions.**

Keywords— deep learning, microvessel, ultrasound

I. Introduction

Microvasculature imaging has a diverse range of applications such as microvascular disease, neuroimaging, tumor microenvironment monitoring, etc. Microvascular imaging with US without a contrast agent is difficult due to the presence of tissue clutter. Singular value decomposition (SVD) has been used to remove tissue clutter echo from blood flow signals [1,2]. Low-rank approximation via SVD requires choosing an optimum threshold that can remove the tissue clutter. This threshold can vary greatly and requires fine-tuning. Moreover, depending on the ensemble length (number of frames) and field-of-view, the computational cost of SVD thresholding (SVT) based tissue removal can be very high. Previously, we presented a model-based deep residual network that demonstrated the ability of a simplified network to compensate for the not carefully adjusted rank value when using SVT [1].

A powerful Deep learning (DL) based vascular imaging approach can approximate or extend SVT for variable ensemble lengths with reduced computational cost during test time, making vascular imaging possible with a fast acquisition, especially on resource-constrained portable devices. A properly trained DL approach will be independent of tissue rank and allow integration of tissue motion, low SNR, etc. in form of data augmentation. Therefore, we proposed an end-to-end deep convolutional neural network-based approach for removing tissue clutter from US data to image microvessel (MV).

II. Methods

A. Data Acquisition

We scanned a day 8 chicken embryo using a Vevo 3100 (VisualSonics Inc.) system and an MX250 transducer (center transmit frequency: 21 MHz). The focus was set halfway down the depth of the embryo at 7 mm. Embryos were grown according to Meijlink et al.[2] and acquired locally from Meyer Hatchery (Polk, OH). Briefly, embryos were stored at 37°C and 50-60% humidity for 8 days in an incubator with automatic egg turning. On day 8, embryos were carefully extracted from the shell and placed in a weigh boat that was kept at 37°C until imaging began. The embryo chorioallantoic membrane (CAM) surface was directly coupled with the ultrasound transducer using ultrasound gel. 3D imaging data was acquired by imaging the entire embryo with a step size of 50µm. Steps were generated precisely with a customized motorized stage (MTS50/M-Z8, Thorlabs Inc.). The experiment took approximately 3 hours and the embryo's heart continued

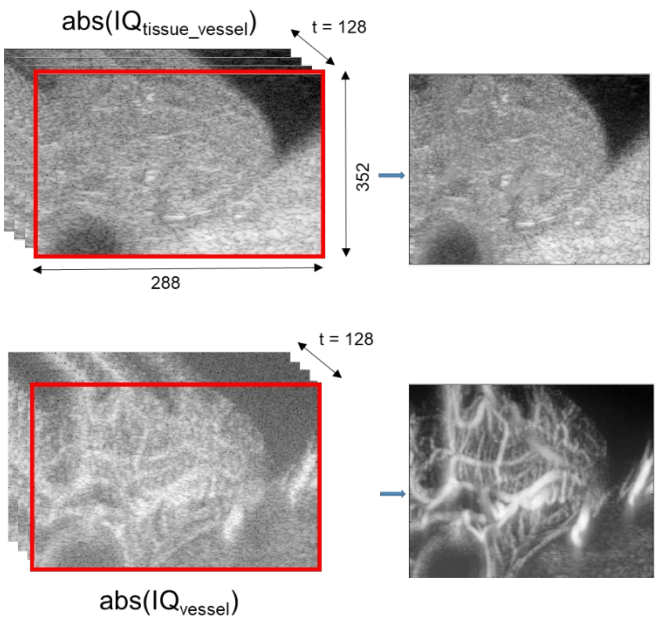

Figure 1. Magnitude of the complex IQ data, $|s|$ before (top) and after (bottom) SVT. Top-left and bottom-left show the US ensemble. Top-right and bottom right show log-normalized intensity projection of the ensemble.

beating for the duration. Spatiotemporal (IQ) data were extracted from the experiment.

B. SVT processing of RF data

Spatiotemporal IQ data s (x,y,t) was processed via SVT reported before [3], [4]. Spatiotemporal data s contains IQ data containing amplitude and phase of the US echo which are both used for removing tissue clutter. Blood consists of flowing heterogeneous scattering elements (e.g., blood cells, platelets, etc.) in a fluid. This produces a complicated, high-rank spatiotemporal echo matrix. We computed a low-rank approximation of the complex IQ data and removed the first $k = 30$ singular vectors spanning the signal to remove tissue clutter.

C. Deep neural network training

The magnitude of IQ data, $|s|$ before and after thresholding was used as input and ground truth to train a deep neural network respectively, shown in Figure 2. The input and the output dimensions were $x \times y \times t$, where x×y was the size of the area imaged and t represents the number of frames. In our experiment, x = 352, y = 288, t = 128. Input volumes first were reshaped into a 2D input array of size (352×288) ×128. Patches of size 352×128 were fed to a 2D U-net [5] network for its prior use in medical image regression tasks [6]. Processing the input this way led to each A-line getting processed individually over time (Figure 2). We trained a U-net (encoder depth 3) for 100 epochs with an initial learning rate of 0.01 and a learning rate drop factor of 0.2. Early stopping was used to avoid overfitting. Mean squared error (MSE) was used as the loss function to train the network. We collected IQ data from 131 locations, 17 of which were discarded due to poor data quality. Data from three separate regions of the embryo were divided into training (69), validation (29), and testing (16) to ensure no overlap. The training and validation data were collected from embryo heart, while the testing data were collected from the brain region of the embryo. During test time, patches of test ensemble were processed individually and assembled.

III. Results

Figure 3 shows that deep neural network can successfully remove tissue clutter. We found structural similarity index of 0.83 ± 0.02 in log-normalized vascular images between SVT images and DL predictions, implying good reconstruction. DL based approach was also 2x faster than SVT approach. However, while SVT works on both amplitude and phase data,

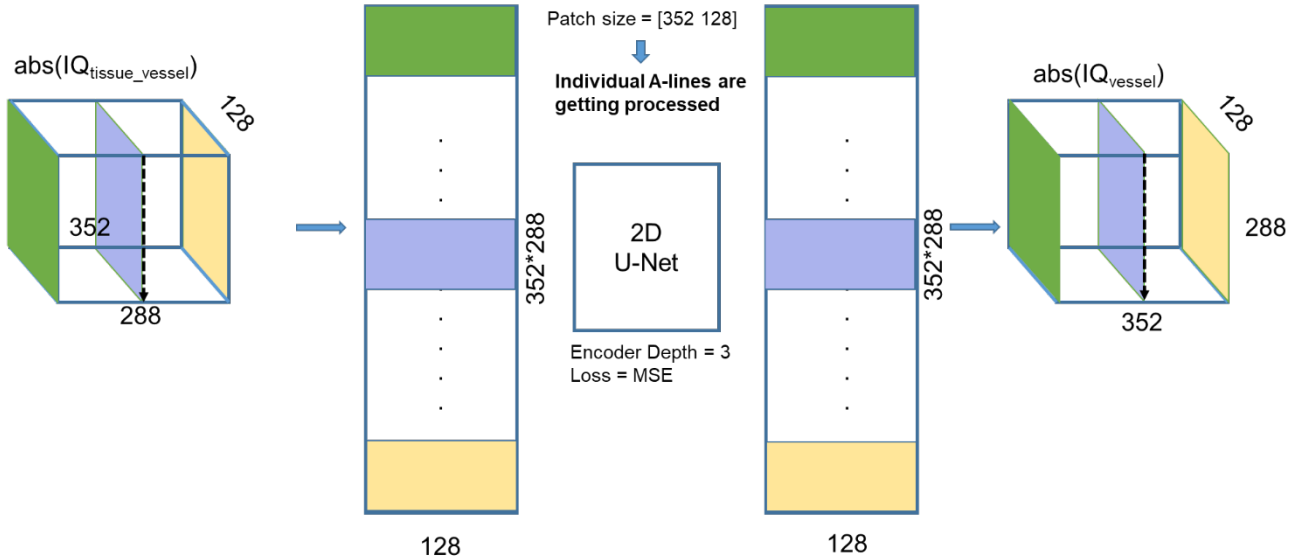

Figure 2. Proposed end-to-end DL approach to extract MV from US ensemble. A 2D U-net was used to train the input patches. During testing, patches were evaluated and stitched together to create the MV image.

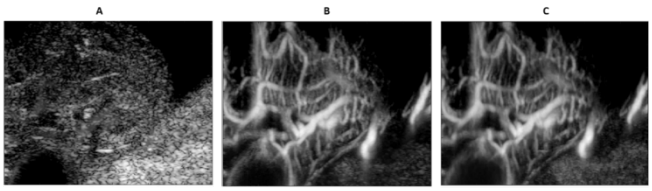

Figure 3. A. Intensity image acquired from raw IQ data acquired near brain and spine of day 8 chicken embryo, B. Log-normalized intensity projection of MV using SVT, C. Log-normalized intensity projection of MV proposed DL approach.

proposed deep learning predictions are from amplitude of IQ data alone so it can be applied to any imaging system without requiring raw data. Addition of phase information, via an appropriate network architecture, can potentially improve performance.

IV. Conclusion

Our proposed deep neural network can successfully remove tissue clutter from US ensemble and makes the process faster than traditional SVT approach. While SVT relies on careful selection of rank, DL can be extended to adapt to different scenarios to admit realistic tissue motions and under noisy conditions. While SVT requires complex IQ data of high ensemble, DL can also extract vessel from the amplitude signal for a shorter ensemble length.

Acknowledgment

This project is funded by an internal source at CWRU, the Case-Coulter Translational Research Partnership, and research funds from the state of Ohio i.e., JobsOhio and Ohio Technology (ODSA) (TECG20200278). This is also funded by the Department of Defense–CDMRP Vision Research Program (W81XWH2110659). The veracity guarantor, Juhwan Lee, affirms to the best of his knowledge that all aspects of this paper are accurate. This work made use of the High Performance Computing Resource in the Core Facility for Advanced Research Computing at Case Western Reserve University. This research was conducted in a space renovated using funds from an NIH construction grant (C06 RR12463) awarded to Case Western Reserve University. The presentation of this work was made possible, in part, through financial support from the School of Graduate Studies at Case Western Reserve University.

References

[1] S. Ehrenstein, E. Abenojar, R. Perera, A. Exner, and M. Bayat, "Rank-Assisted Deep Residual Reconstruction Network for Non-Contrast Ultrasound Imaging of Blood Microvessels," in *2021 IEEE International Ultrasonics Symposium (IUS)*, Xi'an, China, Sep. 2021, pp. 1–4. doi: 10.1109/IUS52206.2021.9593817.

[2] United States. *The Preparation of Chicken Ex Ovo Embryos and Chorioallantoic Membrane Vessels as In Vivo Model for Contrast-Enhanced Ultrasound Imaging and Microbubble-Mediated Drug Delivery Studies.*, (Feb. 09, 2021). doi: 10.3791/62076.

[3] C. Demene *et al.*, "Spatiotemporal Clutter Filtering of Ultrafast Ultrasound Data Highly Increases Doppler and fUltrasound Sensitivity," *IEEE Trans. Med. Imaging*, vol. 34, no. 11, pp. 2271–2285, Nov. 2015, doi: 10.1109/TMI.2015.2428634.

[4] M. Bayat, M. Fatemi, and A. Alizad, "Background Removal and Vessel Filtering of Noncontrast Ultrasound Images of Microvasculature," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 3, pp. 831–842, Mar. 2019, doi: 10.1109/TBME.2018.2858205.

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, vol. 9351, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham: Springer International Publishing, 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.

[6] W. Yao, Z. Zeng, C. Lian, and H. Tang, "Pixel-wise regression using U-Net and its application on pansharpening," *Neurocomputing*, vol. 312, pp. 364–371, Oct. 2018, doi: 10.1016/j.neucom.2018.05.103.