AI-enabled high-speed photoacoustic endomicroscopy through a multimode fibre

Tianrui Zhao[†,*], Mengjie Shi[†], Sebastien Ourselin, Tom Vercauteren, and Wenfeng Xia

School of Biomedical Engineering & Imaging Sciences, King's College London, 4th Floor, Lambeth Wing St Thomas' Hospital, London SE1 7EH, United Kingdom

[†] These authors contributed equally.
[*]Correspondence should be addressed to: Tianrui Zhao, E-mail: tianrui.zhao@kcl.ac.uk

ABSTRACT

Photoacoustic (PA) endoscopy promises to be useful in a variety of clinical contexts including intravascular imaging, gastrointestinal tracts imaging and surgical guidance. Recent advancements of optical wavefront shaping allow the development of ultrathin endoscopy probes based on multimode optical fibres, which can provide higher spatial resolution than previously reported fibre bundle-based endoscopes. In this work, we developed a forward-viewing PA endomicroscopy imaging system and further improved its performance with a deep image prior (DIP) neural network. Laser was focused and scanned through a multimode fibre via wavefront shaping, in which a real-valued intensity transmission matrix approach was used for fibre characterisation, and a digital micromirror device (DMD) was used for light modulation. The excited ultrasound waves at the distal fibre tip were detected by an ultrasound transducer. High fidelity images of *ex vivo* mouse red blood cells were acquired. A DIP neural network was then used to improve the spatial resolution with unsupervised learning. Convolutional filters were used to learn features of low-level images as priors and reconstruct high-resolution images accordingly. The performance of the DIP approach was evaluated using a structural similarity index measure (SSIM) at a level of 0.85 with 25% effective pixels, which outperformed the bicubic method. The use of DIP allows reducing scanning positions by several times, and thus improves the speed of pixel-wise PA microscopy imaging. With further miniaturisation of the ultrasound detector, we anticipate that this system could be used for real-time guidance of minimally invasive surgeries by providing micro-structural, molecular, and functional information of tissue.

Keywords: Photoacoustic endoscopy, wavefront shaping, neural network

1. INTRODUCTION

Photoacoustic (PA) imaging is an emerging modality with demonstrated potential for a wide range of biomedical and clinical applications [1,2]. Photoacoustic endoscopy (PAE) greatly extends the clinical applicability landscape of PA imaging by inserting a miniature probe for imaging from inside human bodies, promising to be useful in various clinical fields such as intravascular imaging, gastrointestinal tracts imaging, and surgical guidance [3,4].

PAEs can be generally categorised into two groups: side-viewing and forward-viewing PAE. With side-viewing PAEs, excitation light is delivered through a side-firing (angled) optical fibre, or deflected by a mirror in front of the tip of a flat-cleaved optical fibre, to illuminate biological tissues surrounding the probe [3,4]. An US transducer integrated with the fibre detects the excited US waves. Circumferential scans of surrounding tissue are enabled by rotating the side-viewing probe with a frame rate up to ~20 frame per second (fps). In contrast, forward-viewing PAEs usually rely on a micro-electromechanical system (MEMS) or a galvo mirror to raster-scan a focused laser beam through a coherent fibre bundle, and a miniature US sensor integrated at the probe tip for US detection. The imaging speed of 2 fps was achieved in such a setup [5]. Fibre bundles have also been coated with Fabry-Perot sensors for the development of PAE in PA tomography mode, which were demonstrated for high-fidelity imaging. However, this method required a lower image acquisition speed [6]. In recent years, multimode fibres (MMF) were studied as an alternative to coherent fibre bundles for forward-viewing PAEs, due to its higher resolution, thinner size, and lower cost. Wavefront shaping was employed to calibrate the scrambled laser transmission through a MMF for light focusing and scanning at the distal tip of the MMF for microscopy imaging.

Early works used liquid-crystal spatial light modulators (LC-SLMs) to modulate the incident light field, and thus the imaging speed was limited by the slow rate of LC-SLMs (<100 Hz). In a previous study, it took 30 s for the acquisition of a single image [7]. In our previous work, a high-speed modulator called digital micromirror device (DMD) was used for wavefront shaping and improved the scanning rate to 23 kHz. However, since signal averaging was required to achieve high image fidelity, the acquisition of a single PA image of mouse red blood cells (RBCs) cost ~7 s [8], which needed to be improved for clinical translation.

Recently, deep learning has shown promise for improving the image acquisition rate of PA microscopy. Convolutional neural networks (CNNs) were employed to improve the under sampled PA microscopy images and showed a good inference performance on *in vivo* vascular data [9,10], with a structural similarity index (SSIM) of up to 0.819 using 2% of the fully sampled pixels [10]. However, these supervised models required a large training dataset with full-sampled images as the ground truths. Besides, the previous models explicitly trained with vascular images were highly likely to fail on unseen patterns [11]. The unsupervised deep image prior (DIP) model has been proven efficient for improving the sparse PA microscopy images neither with pre-training nor fully sampled ground truths [12]. The vascular images with different sparsity were restored from the noise input via iterative optimization, showing comparable performances with the state-of-art pre-trained model.

In this work, we developed a DIP neural network with unsupervised learning to improve the imaging speed of the MMF-based PAE. Convolutional filters were used to learn features of sub-sampled PA images as priors and reconstruct high-resolution images accordingly. The performance of the DIP approach was evaluated using the SSIM and peak signal-to-noise ratio (PSNR), which outperformed the bicubic method with upsampling PAM images of mouse red blood cells (RBCs).

2. METHODS

2.1 Imaging system

The configuration of the PAE system was detailed in our previous work [8] and a schematic diagram is shown in Figure 1. Briefly, a 532 nm laser (2 ns, 50 kHz, SPOT-10-200-532, Elforlight, Daventry, United Kingdom) was used as the excitation light source. A graded-index MMF (Ø100 μm, 0.29 NA, Newport) with a length of 20 cm was used for laser delivery. A DMD (768×1080 elements, DLP7000, Texas Instruments, Texas, USA) was placed in front of the proximal tip of the MMF to project binary patterns onto the fibre tip via a tube lens (AC254-050-A-ML, Thorlabs, New Jersey, USA), a circular polariser (CP1L532, Thorlabs, New Jersey, USA), and a 20× objective (0.4 NA, RMS20X, Thorlabs, New Jersey, USA). A sub-region of the DMD covering 128×128 micromirrors was controlled for light modulation.

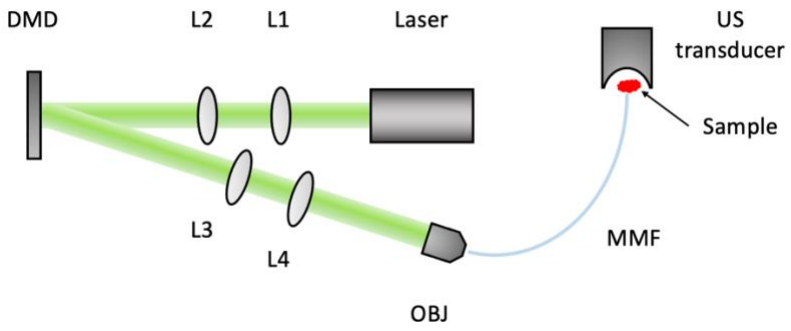

Figure 1. Configuration of the photoacoustic endoscopy system for imaging implementation. DMD, digital micromirrors; L1-4, lenses; OBJ, objective; MMF, multimode fibre; US, ultrasound. Characterisation module with a CMOS camera, an objective, and a lense was not shown.

For fibre characterisation, light transmitted through the MMF was captured with a CMOS camera (C11440-22CU01, Hamamatsu Photonics, Shizuoka, Japan) after magnification by an objective (20×, 0.4 NA, RMS20X, Thorlabs, New

Jersey, USA) and a tube lens (AC254-0100-A-ML, Thorlabs, New Jersey, USA). A series of binary Hadamard patterns were displayed on the DMD whilst the corresponding speckles were recorded. Then, the light transmission characteristics through the MMF was determined using a real-valued intensity transmission matrix (RVITM) algorithm developed in our group [15,16]. Subsequently, optimal DMD patterns for light focusing at all output positions across the distal fibre tip were calculated [16]. For imaging, the camera was replaced with an imaging module comprising imaging samples, and a piezoelectric US transducer (V358, central frequency: 50 MHz; diameter: 0.25 inches, Olympus, Japan,) for US detection. An acoustic lens (LC4210, f = -25 mm, Thorlabs, New Jersey, USA) was attached to the active surface of the US transducer for focusing at the optical focal plane of the MMF. The ultrasound signals were digitised and processed for the reconstruction of PA images.

2.2 Network architecture and implementation

The deep image prior (DIP) model proposed in Ref. [12] employed the classical U-Net [13] with some modifications. The DIP model implemented in this paper followed the primal features of the modified U-Net in Ref. [12] as shown in Figure 2. It has an encoder-decoder architecture where the encoder path learns the low-resolution features and the decoder projects the learnt features on the pixel space and forms the output. Compared to the architecture in Ref. [12], fewer scales were used due to the smaller input size by our system. 64 convolutional filters were used at all the scales and a medium kernel size (7×7) was incorporated for efficiently extracting nearby features from low-resolution images. As indicated in Ref. [12], bilinear interpolation was employed at the upsampling path to mitigate checkboard artefacts related to transposed convolution. The DIP model was trained for 4,000 iterations using the Mean Square Error (MSE) loss and AMSGrad optimizer with a clip value of 10. Training was performed with NVIDIA DGX-1 using Python v3.6.1 and Keras 2.2.5 with Tensorflow 1.15.0 backend. Noise-based regularization [12][14] was used during the training. The noise input was perturbed by adding a regularization term z' with a Gaussian distribution after each iteration. The standard deviation of z' was experimentally determined for each image.

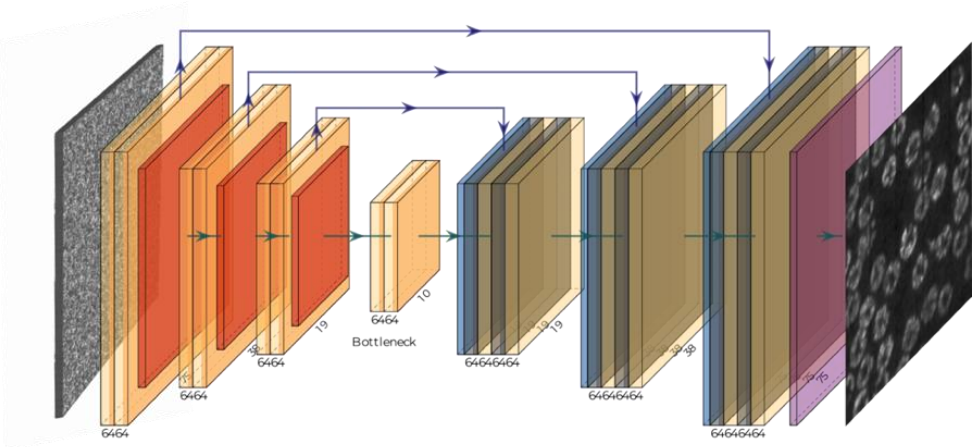

Figure 2. The U-Net based deep image prior (DIP) model. (a) The input is a 32-channel 100×100 random noise, and the output is a recovered photoacoustic microscopy image in grayscale. (b) all layers employed 64 convolutional filters for parameter reduction. Bilinear upsampling was used in the expansion path.

PAE images of mouse blood cells were used for training, and evaluation. Mouse blood was obtained from culled mice: these procedures involving mice were ethically reviewed and carried out in accordance with the Animals (Scientific Procedures) Act 1986 (ASPA) UK Home Office regulations governing animal experimentation. The undersampled images were generated by pixel-wisely multiplying binary undersampling masks with their corresponding full sampled images. Three different sampling ratios, namely 1/2, 1/3, 1/4 along both x and y axes were considered. To evaluate the model, the SSIM and PSNR were measured between the reconstructed images by the bicubic interpolation, DIP and the corresponding fully sampled image for quantifying their performances. The training time for processing an input of 100×100 pixels with different undersampling conditions by the DIP model was recorded.

3. RESULTS

Representative results using the DIP and bicubic interpolation with mouse red blood cells are shown in Figure 3. For the small sampling ratio at 1/2, DIP presents better reconstructed details compared to the bicubic. Most of the recovered blood cells by DIP remain smooth and clear at the boundaries while the bicubic shows ragged edges. In terms of the quantitative metrics SSIM and PSNR, DIP slightly outperforms the bicubic with the 1/2 sampling ratio (SSIM: 0.8458 vs 0.7368; PSNR: 20.6352 vs 22.1455). The reconstruction results by DIP show less improvements for higher undersampling ratios: while the biconcave structure of the RBCs can be easily resolved with the ratio at 1/3, it was challenging to visualise the cellular structures with the under-sampling ratio of 1/4. Nevertheless, DIP reconstructs the cells with flat and continuous surfaces, which indicates its primal attribute for recovering graphical/physiological features and reducing artefacts.

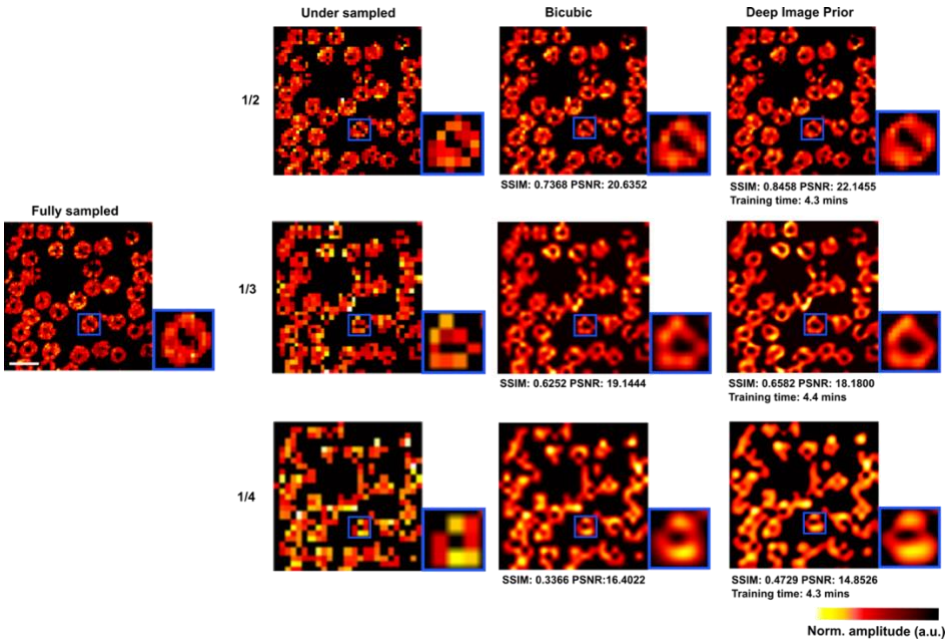

Figure 3. Representative results of deep image prior (DIP) reconstruction of photoacoustic images of mouse red blood cells. Bicubic interpolation is employed for comparison. Scale bar: 10 μm. SSIM: structural similarity index measure; PSNR: peak signal-to-noise ratio.

4. DISCUSSION

In this work, we demonstrated the use of a DIP method for upsampling reconstruction of PA microscopy images achieved with am MMF-based endomicroscopy probe. The bi-concave structure of mouse red blood cells was recovered from 1/3 undersampling, as such, the speed of the PAE system can be significantly enhanced by up to 9 times with slight degradation in image fidelity.

Experimental results indicated that the U-Net based DIP model could learn low-level statistics prior from the sparse images at the contracting path and present the prior on the reconstructed image via the expansive path. Therefore, compared to the interpolation algorithms, DIP performs well for keeping major visual features of the imaging targets and maintaining high image fidelity. However, based on our experimental results, DIP or the bicubic interpolation could hardly recover completely missing features due to the down sampling and their performance will be extensively limited by the original image quality. Images in large size or high-resolution could fit in deeper neural networks that DIP would benefit from [10]. Supervised models have also been studied for reconstructing under sampling data, however, they require large-scale datasets with fully-sampled ground truth which is expensive and time-intensive. In contrast, DIP enables upsampling reconstruction without pre-training on dataset or fully sampled ground truth. The DIP method was implemented using the NVIDIA DGX-1 cluster, and it took around 4 min for reconstructing an image of 100×100 pixels. The inference time could be further optimised for real-time applications by employing lightweight models or involving pre-trained modules.

5. CONCLUSION

In conclusion, we developed a PA endomicroscopy imaging system and employed a DIP-based upsampling method to improve the imaging speed. With high-performance upsampling from 1/2 undersampled data, the imaging speed was improved by 4 times with slight degradation in image quality. The DIP was demonstrated for upsampling reconstruction of PAM images of red blood cells, which allowed an enhancement of imaging speed by 9 time. With further optimisation and the use of more powerful computer, this method is promising to be used for the development of a real-time PAE imaging system.

ACKNOWLEDGEMENT

This project was supported by the Academy of Medical Sciences/the Wellcome Trust/ the Government Department of Business, Energy and Industrial Strategy/the British Heart Foundation/Diabetes UK Spring board Award [SBF006/1136], Wellcome Trust, United Kingdom (203148/Z/16/Z, WT101957, 203145Z/16/Z), Engineering and Physical Sciences Research Council, United Kingdom (NS/A000027/1, NS/A000049/1, NS/A000049/1). Mengjie Shi acknowledges the support of King's – China Scholarship Council PhD Scholarship Program (K-CSC) (202008060071).

REFERENCES

[1] Beard, Paul. "Biomedical photoacoustic imaging." Interface focus 1, no. 4 (2011): 602-631.
[2] Wang, Lihong V. "Multiscale photoacoustic microscopy and computed tomography." Nature photonics 3, no. 9 (2009): 503-509.
[3] Zhao, Tianrui, Adrien E. Desjardins, Sebastien Ourselin, Tom Vercauteren, and Wenfeng Xia. "Minimally invasive photoacoustic imaging: Current status and future perspectives." Photoacoustics 16 (2019): 100146.
[4] Zhou, Jingcheng, and Jesse V. Jokerst. "Photoacoustic imaging with fiber optic technology: A review." Photoacoustics (2020): 100211.
[5] Hajireza, Parsin, Tyler J. Harrison, Alexander Forbrich, and Roger J. Zemp. "Optical resolution photoacoustic microendoscopy with ultrasound-guided insertion and array system detection." Journal of Biomedical Optics 18, no. 9 (2013): 090502.
[6] Ansari, Rehman, Edward Z. Zhang, Adrien E. Desjardins, and Paul C. Beard. "All-optical forward-viewing photoacoustic probe for high-resolution 3D endoscopy." Light: Science & Applications 7, no. 1 (2018): 1-9.
[7] Mezil, Sylvain, Antonio M. Caravaca-Aguirre, Edward Z. Zhang, Philippe Moreau, Irène Wang, Paul C. Beard, and Emmanuel Bossy. "Single-shot hybrid photoacoustic-fluorescent microendoscopy through a multimode fiber with wavefront shaping." Biomedical Optics Express 11, no. 10 (2020): 5717-5727.
[8] Zhao, Tianrui, Michelle T. Ma, Sebastien Ourselin, Tom Vercauteren, and Wenfeng Xia. "Video-rate dual-modal photoacoustic and fluorescence imaging through a multimode fibre towards forward-viewing endomicroscopy." Photoacoustics (2021): 100323.
[9] Zhou, Jiasheng, Da He, Xiaoyu Shang, Zhendong Guo, Sung-liang Chen, and Jiajia Luo. "Photoacoustic Microscopy with Sparse Data Enabled by Convolutional Neural Networks for Fast Imaging." arXiv preprint arXiv:2006.04368 (2020).
[10] DiSpirito, Anthony, Daiwei Li, Tri Vu, Maomao Chen, Dong Zhang, Jianwen Luo, Roarke Horstmeyer, and Junjie Yao. "Reconstructing undersampled photoacoustic microscopy images using deep learning." IEEE transactions on medical imaging 40, no. 2 (2020): 562-570.
[11] Tommasi, Tatiana, Novi Patricia, Barbara Caputo, and Tinne Tuytelaars. "A deeper look at dataset bias." In Domain adaptation in computer vision applications, pp. 37-55. Springer, Cham, 2017.
[12] Vu, Tri, Anthony DiSpirito III, Daiwei Li, Zixuan Wang, Xiaoyi Zhu, Maomao Chen, Laiming Jiang et al. "Deep image prior for undersampling high-speed photoacoustic microscopy." Photoacoustics 22 (2021): 100266.
[13] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.
[14] Ulyanov, Dmitry, Andrea Vedaldi, and Victor Lempitsky. "Deep image prior." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 9446-9454. 2018.

[15] Zhao, Tianrui, Sebastien Ourselin, Tom Vercauteren, and Wenfeng Xia. "Seeing through multimode fibers with real-valued intensity transmission matrices." Optics Express 28, no. 14 (2020): 20978-20991.
[16] Zhao, Tianrui, Sebastien Ourselin, Tom Vercauteren, and Wenfeng Xia. "Focusing light through multimode fibres using a digital micromirror device: a comparison study of non-holographic approaches." Optics Express 29, no. 10 (2021): 14269-14281.