UC Irvine
UC Irvine Previously Published Works

Title
Deep Learning–Assisted Multiphoton Microscopy to Reduce Light Exposure and Expedite Imaging in Tissues With High and Low Light Sensitivity

Permalink
https://escholarship.org/uc/item/1kq1x4s7

Journal
Translational Vision Science & Technology, 10(12)

ISSN
2164-2591

Authors
McAleer, Stephen
Fast, Alexander
Xue, Yuntian
et al.

Publication Date
2021-10-20

DOI
10.1167/tvst.10.12.30

Peer reviewed

Deep Learning–Assisted Multiphoton Microscopy to Reduce Light Exposure and Expedite Imaging in Tissues With High and Low Light Sensitivity

Stephen McAleer[1,2], Alexander Fast[3,4], Yuntian Xue[5], Magdalene J. Seiler[6–8], William C. Tang[5], Mihaela Balu[3], Pierre Baldi[1,2], and Andrew W. Browne[5,8,9]

[1] Department of Computer Science, University of California, Irvine, Irvine, CA, USA
[2] Institute for Genomics and Bioinformatics, University of California, Irvine, Irvine, CA, USA
[3] Beckman Laser Institute and Medical Clinic, University of California, Irvine, Irvine, CA, USA
[4] InfraDerm, LLC, Irvine, CA
[5] Department of Biomedical Engineering, University of California, Irvine, Irvine, CA, USA
[6] Department of Physical Medicine & Rehabilitation, University of California, Irvine, Irvine, CA, USA
[7] Sue and Bill Gross Stem Cell Research Center, University of California, Irvine, Irvine, CA, USA
[8] Gavin Herbert Eye Institute, Department of Ophthalmology, University of California, Irvine, Irvine, CA, USA
[9] Institute for Clinical and Translational Science, University of California, Irvine, Irvine, CA, USA

Correspondence: Andrew W. Browne, Department of Biomedical Engineering, University of California, Irvine, 402 E Peltason Drive, Irvine, CA 92617, USA.
e-mail: abrowne1@uci.edu
Pierre Baldi, Department of Computer Science, University of California, Irvine, 6210 Donald Bren Hall, Irvine, CA 92697, USA.
e-mail: pfbaldi@uci.edu

Purpose: Two-photon excitation fluorescence (2PEF) reveals information about tissue function. Concerns for phototoxicity demand lower light exposure during imaging. Reducing excitation light reduces the quality of the image by limiting fluorescence emission. We applied deep learning (DL) super-resolution techniques to images acquired from low light exposure to yield high-resolution images of retinal and skin tissues.

Methods: We analyzed two methods: a method based on U-Net and a patch-based regression method using paired images of skin (550) and retina (1200), each with low- and high-resolution paired images. The retina dataset was acquired at low and high laser powers from retinal organoids, and the skin dataset was obtained from averaging 7 to 15 frames or 70 frames. Mean squared error (MSE) and the structural similarity index measure (SSIM) were outcome measures for DL algorithm performance.

Results: For the skin dataset, the patches method achieved a lower MSE (3.768) compared with U-Net (4.032) and a high SSIM (0.824) compared with U-Net (0.783). For the retinal dataset, the patches method achieved an average MSE of 27,611 compared with 146,855 for the U-Net method and an average SSIM of 0.636 compared with 0.607 for the U-Net method. The patches method was slower (303 seconds) than the U-Net method (<1 second).

Conclusions: DL can reduce excitation light exposure in 2PEF imaging while preserving image quality metrics.

Translational Relevance: DL methods will aid in translating 2PEF imaging from bench-top systems to in vivo imaging of light-sensitive tissues such as the retina.

Introduction

Traditional fluorescence microscopy, also known as single-photon microscopy, illuminates a sample using short-wavelength light to excite fluorescent molecules, which then release the energy by fluorescing at a longer wavelength.[1] A computer-based imaging system is used to collect the fluorescent light and reconstruct the fluorescence image. Multiphoton microscopy, in contrast, splits the energy required for fluorescence excitation into two or three lower energy photons of light (typically within the infrared spectrum). Two-photon excitation fluorescence (2PEF) occurs with simultaneous absorption of two photons by a molecular fluorophore and is a nonlinear process proportional to the square of the instantaneous excitation light intensity.[2,3] The major benefits of 2PEF over single-photon microscopy are twofold. First, 2PEF induces fluorescence excitation in a small focal volume where two photons interact simultaneously with the tissue, as opposed to single photon techniques where the fluorescence occurs along the light path length, thus achieving better spatial resolution. Second, 2PEF uses infrared light to excite fluorescence, which allows deeper tissue penetration and imaging than visible light.

Microscopic image resolution is determined by how much focused light is detected by the imaging system. Therefore, increasing light originating from a sample, refining optical focus, and optimizing light detector sensitivity can each improve image resolution. Scan averaging and higher power fluorescence excitation allow imaging systems to collect more light from a sample to produce an image with higher resolution. Acquiring many low-energy excitation images of a sample and averaging the images into a single image are equivalent to increased light originating from the sample (Fig. 1i). Alternatively, sample fluorescence is amplified by increasing excitation power so that fewer scans are required to produce an image (Fig. 1ii).
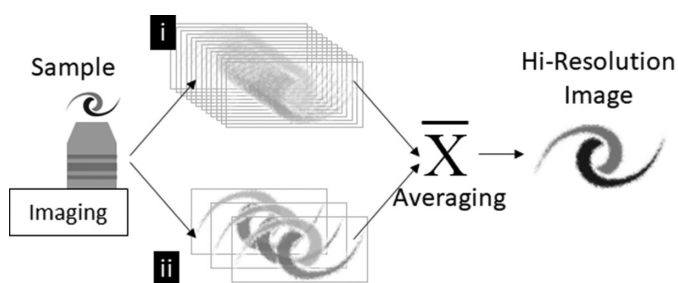

Figure 1. Methods to increase image resolution by increasing light originating from a sample. (i) Many low-power excitation scans are acquired or (ii) few high-power scans are acquired and averaged to produce a high-resolution image.

Safety is a primary concern for all optical technologies used to image human tissues, especially the retina. Fortunately, all preclinical results in murine eyes demonstrated no damage from 2PEF, as evaluated using visual function testing, in vivo structural imaging, ex vivo histology, and ex vivo biochemical analysis.[4]

2PEF has been comprehensively evaluated in rodents in vitro and in vivo and was shown to be safe.[4] Briefly, two-photon infrared light (1.2-mW laser power, imaging duration of 130 seconds, 100 frames, 790 nm, 20-fs pulse duration, 80-MHz pulse frequency, and total dose of 15.6 J/cm^2) was compared with no light and white light exposure. Structural imaging (optical coherence tomography, confocal scanning laser ophthalmoscopy, and autofluorescence), cellular electrophysiology (electroretinogram), and cellular biochemistry (quantity of rhodopsin and 11-*cis* retinal) showed no structural nor functional difference between eyes exposed to no light and those exposed to the two-photon infrared laser. However, all endpoints were significantly diminished in eyes exposed to visible white light. Schwarz et al.[5] demonstrated in macaques that infrared two-photon light exposure (0.5-mW laser power, imaging duration of 40 seconds, 900 frames, 730 nm, 55-fs pulse duration, 80-MHz pulse frequency, and total dose of 20.4 J/cm^2) resulted in changes in infrared autofluorescence immediately following light exposure but no detectable functional changes. These changes were noted to revert to normal over a period of 22 weeks of observation. No other structural or functional alterations were detected by other imaging techniques for any of the lower dose exposures. Further evaluation in macaques demonstrated that 856-J/cm^2, pulsed infrared two-photon light exposure (730 nm, 55-fs pulse, and 80-MHz repetition frequency) resulted in changes in infrared adaptive optics findings in one class of retinal photoreceptors (S cones, blue).[6] However, they also demonstrated that this effect was not seen with lower power laser energy doses ranging from 214 to 489 J/cm^2. Despite the safety profile of 2PEF in preclinical models, it is essential to reduce risks of light toxicity to human tissues in every way possible. Both novel biophotonic principles and optimal data processing methods can reduce light exposure, expedite image acquisition, and optimize fluorescence image quality. This work explores improving human tissue imaging by applying deep learning (DL) to maximize image quality while reducing fluorescence excitation exposure.

Deep learning is a branch of machine learning based on artificial neural networks. Taking advantage of computing power with graphics processing units, these DL approaches are currently the method of
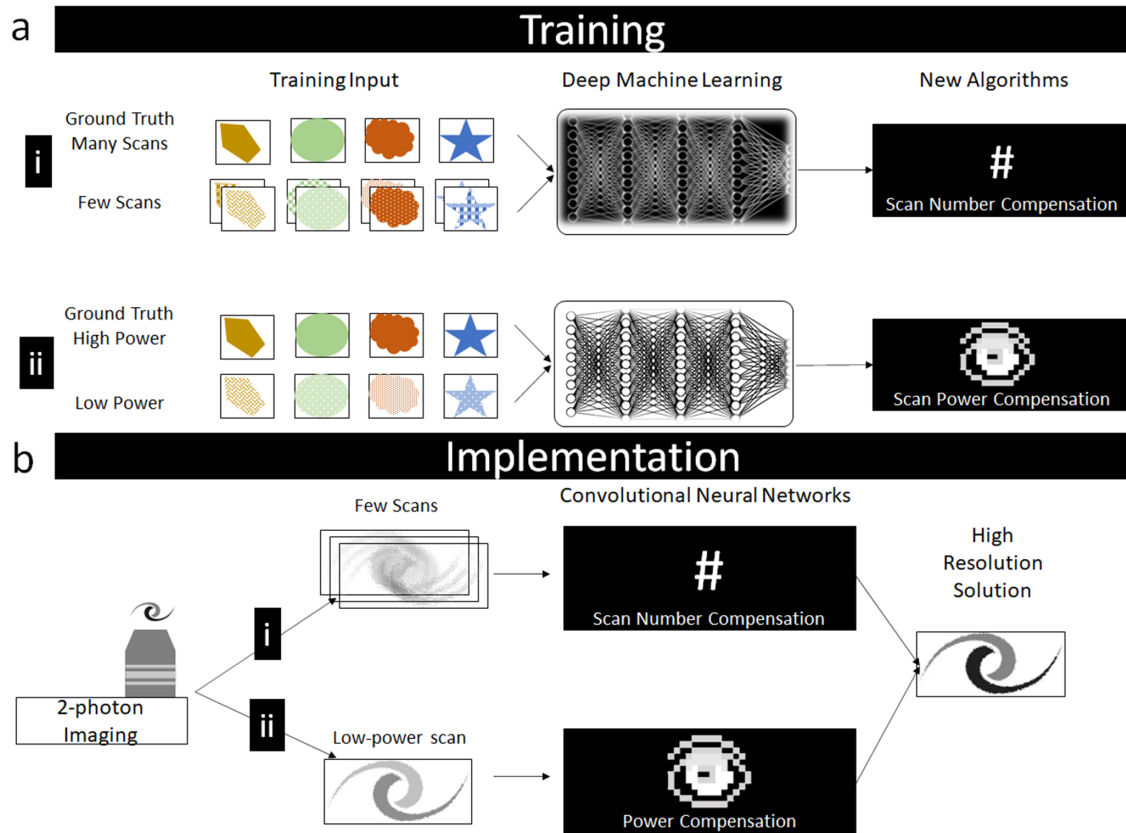
Figure 2.　Training and implementing deep machine learning algorithms. (a) Training convolutional neural networks using paired datasets composed of (i) images produced from multiple or few scans or (ii) high-resolution scans acquired at high or low laser excitation power. (b) Imaging data were acquired as follows: (i) a few scans at low resolution were processed by a scan number compensation algorithm, or (ii) a low-power scan was processed by a power compensation algorithm to produce a high-resolution solution.

choice in artificial intelligence and machine learning for computer vision applied to biomedical imaging.[7–9] In image analysis, DL algorithms distinguish themselves from other approaches because they do not require manually entering a list of features. DL algorithms, in contrast, learn relevant features directly from training data and use them for classification, regression, and other tasks. In ophthalmology, for example, DL has been applied to diabetic retinopathy, age-related macular degeneration,[10,11] and glaucoma screening.[12] Recently, DL has enabled superresolution microscopy applied to fluorescence microscopy techniques[13] and mobile phone microscopes.[14] Within our group, DL has been used in many biomedical imaging tasks, such as identifying gastrointestinal polyps,[15] classifying genetic mutations in gliomas,[16] detecting cardiovascular diseases,[17] detecting spinal metastasis,[18] counting hair follicles,[19] identifying fingerprints,[20] and analyzing vascularization images.[21]

With the goal of reducing excitation light exposure and expediting imaging time, we explored two strategies to enhance 2PEF imaging in human stem cell–derived retinal organoids[22] and human skin.[3] First, we trained a convolutional neural network using paired datasets from multiple image samples acquired by averaging many scans or few scans (Fig. 2ai). Second, we trained a convolutional neural network using paired images acquired at low and higher scan powers (Fig. 2aii). We then finally evaluated the performance of two DL algorithms to produce high-resolution images based on few input scans (Fig. 2bi) or a single low-power scan (Fig. 2aii).

Methods

Ex Vivo Imaging of Human Skin

De-identified institutional review board–exempt human skin was obtained from the University of California, Irvine, Dermatology Clinic and consisted of excess normal tissue discarded during wound closure procedures. The specimens were imaged fresh, immediately upon collection.

Images were taken with a home-built fast, large area multiphoton exoscope (FLAME), a device based on laser scanning multiphoton microscopy and optimized for rapid, depth-resolved imaging of large skin tissue areas with submicron resolution.[23–25] Briefly, a frequency-doubled, ytterbium-doped, 780-nm, 90-fs, 80-MHz amplified fiber laser (Carmel X-series; Calmar Laser, Palo Alto, CA) was used to excite endogenous skin components such as keratin, melanin, NAD(P)H, collagen, and elastin. Laser power was set to 45 mW at the focus of a 1.05-NA, 25× objective lens (XLPL25XWMP; Olympus, Tokyo, Japan) for all measurements. Emission was separated from excitation with a 705-nm shortpass dichroic mirror (FF705-Di01; Semrock, Rochester, NY) and further filtered with a 620-nm shortpass filter and a 535-nm centered bandpass filter (FF01-620/SP and FF01-535/150, respectively; Semrock). To detect the signal, we employed a sensitive photomultiplier tube (R9880-20; Hamamatsu Photonics, Hamamatsu, Japan) in photon counting mode. A resonant scanner and galvonometric mirror pair were used to raster scan the excitation beam over the sample. Image frames were taken at 1024 × 1024 pixels (900 × 900 μm^2) with an 88-ns dwell time per pixel and consecutively averaged. High signal-to-noise ratio ground-truth (GT) images were acquired by averaging 70 frames (~10 seconds), and lower resolution input images for training the neural network were acquired by averaging the first seven and 15 frames (~1 and 2 seconds, respectively) of the same video stack. Images were acquired to train DL algorithms from eight excisions including pigmented and nonpigmented skin at different depths within viable epidermis. Among the 713 images acquired at both frame counts, 641 images were used for neural net training, and 72 images were held out for testing the neural net. Images used for neural net testing were selected randomly.

In Vitro Imaging of Human Retinal Organoids

A genetically modified registered embryonic line (WA01 line expressing CRX-GFP[26]) was used for retinal organoid generation in this study; the retinal organoids were produced by the force aggregation method[27] and differentiated as described previously.[28] Images were taken with a Zeiss LSM 510 microscope (Carl Zeiss Meditec, Jena, Germany) with a Zeiss EC Plan Neofluar 20×/0.5 Ph2 objective. The imaging sample included eight retinal organoids on day 61 of differentiation; each organoid was imaged at different depths and sections to obtain unique images of different cross-sections. The excitation wavelength was 740 nm, produced by a Chameleon laser source (COHER-

ENT MRU X1, coupled with a cooling system, Neslab Refrigerated Recirculator; Thermo Fisher Scientific, Waltham, MA). The laser power values used were 300 mW (high power) and 50 mW (low power); laser pulse frequency was about 518.84 Hz as measured using a light power meter (PM100D; Thorlabs, Newton, NJ). Except for laser power, other imaging parameters were held constant. The scan mode was planar, 12 bits, with frame size of 512 × 512 (450 μm × 450 μm); pixel dwell time was 1.61 μs. The laser passed through a Zeiss HFT KP 650 dichroic beam splitter, mirror, and Zeiss NFT 490 beam splitters. Each imaging region was first imaged under low power and then switched to high power. Paired images ($n = 300$ low power, $n = 300$ high power) were acquired with two separate imaging channels for nicotinamide adenine dinucleotide (390–465 nm) and flavin adenine dinucleotide (500–550 nm). A total of 1200 images were acquired to train the DL algorithms. Among the 1200 images acquired at each power, 1169 images were used for neural net training, and 31 images were randomly designated for testing.

DL Methods

We applied two different DL methods, patch-based and content-aware image restoration (CARE), to the two datasets. The patch-based method partitioned the input into small tiles and then trained a neural network using supervised learning with the objective to predict the high-resolution of the middle pixel. The CARE method used U-Net, a popular model that used contractive convolutional neural networks and expansive up-convolutional neural networks for fine-grained prediction. CARE was selected for its broad success and adoption in the literature.[29–32]

Patch-Based Regression

In this method, we followed the patch-based method of Ciresan et al.[33] We first created patches of the input image. Each patch was a 40 × 40-pixel square around each pixel in the input image. For pixels near the edge, portions of the 40 × 40-pixel square fell outside the image, so we padded the extra values with 0. These patches were then compiled as a set of inputs. The target corresponding to a given patch was the center pixel of the patch in the high-resolution image. The input image and target image both had two color channels representing two intrinsic fluorophores (nicotinamide adenine dinucleotide and flavin adenine dinucleotide) that provide structural contrast in 2PEF imaging. We then trained a neural network to predict the target pixel from the input of the patch using mean squared error (MSE) as the loss function. We used a neural network with two convolutional layers followed

by three fully connected layers. The first convolutional layer had 64 filters with a kernel size of 4×4. The second convolutional layer had 32 layers with a kernel size of 3×3. A batch normalization layer was between the two convolutional layers. The fully connected layers had sizes of 1024, 512, 32, and 2, respectively. All layers had a rectified linear unit (ReLU) activation. We trained the neural network with Adam optimization.[34] We used a neural network with two convolutional layers followed by three fully connected layers. The first convolutional layer had 64 filters with a kernel size of 4×4. The second convolutional layer had 32 layers with a kernel size of 3×3. A batch normalization layer was between the two convolutional layers. The fully connected layers had sizes of 256, 128, 32, and 2, respectively. We trained the neural network with Adam optimization with a learning rate of 0.001.[34] We used a batch size of 256 and 64 steps per epoch. A pretrained model will be used in future work, as a pretrained model would likely yield better performance than patches from scratch.

U-Net

We used CSBDeep CARE,[35] a model that is based on a residual version of U-Net.[36] Our dataset included source images and high-resolution GT images. We split the images into patches and then ran the CARE U-Net model on the patches. To create the training set, we extracted 128×128 patches from each image, and each patch was centered on a pixel in the original image with padding. The depth of the U-Net was 2, the kernel size was 5, the last activation was linear, and the training loss was Laplace. We trained the CARE U-Net on 100 training epochs with 30 training steps per epoch, with a training learning rate of 0.0004 and a batch size of 16.

Cross-Validation

Five-fold cross validation was performed for each dataset. Each image in the datasets was numbered. We used the last digit to assign each image to a fold for cross-validation. For the first fold, we trained on images with digits ending in 2 to 9 and tested on images with digits ending in 0 and 1. For the second fold, we trained on images ending in 0, 1, and 4 to 9 and tested on images ending in 2 and 3. The same process was repeated across five cross-validation steps.

Quantitative Evaluation of DL Output

Quantitative comparison of images produced by the two DL strategies for agreement with GT images was performed by two conventional unitless measures: MSE and structural similarity index measure (SSIM). MSE assessed the cumulative error between two datasets (images in this context). Pixels or groups of pixels between a target and source image were compared, and the average squared difference between the estimated pixel values and the actual pixels was calculated. Image degradation might occur with data compression, data lost during transmission or acquisition, or data prediction. SSIM quantified similarity between two images to produce an image degradation metric.[37]

Results

The patch-based and CARE methods applied to two different datasets were compared using MSE and SSIM metrics applied to images restored from low-resolution images. MSE and SSIM were determined for the held-out test set of images for each dataset (Table 1). Table 2 presents performance metrics for Patches and CARE models.

Figure 3 shows representative images restored using the two DL approaches and by averaging the restored images to produce an average image produced by the two DL approaches. Associated MSE and SSIM images are shown for each of the three approaches. As shown in Table 1, the patches method achieved much lower MSE and slightly higher SSIM.

Cross-validation was performed for both skin and retinal organoid datasets as detailed in the Supplementary Materials. For the skin dataset, the mean

Table 1. Quantitative Output for All Test Images From the Retinal Organoid and Skin Datasets[a]

Dataset		Patches	Care	Average (Patches + Care)
Retinal organoid	Average MSE	27,611 (2819)	146,855 (26633)	55,546 (5899)
	Average SSIM	0.636 (0.01)	0.607 (0.04)	0.620 (0.01)
Skin samples	Average MSE	3.768 (0.4)	4.032 (2.1)	3.803 (0.3)
	Average SSIM	0.824 (0.01)	0.783 (0.4)	0.816 (0.02)

[a] Average values were determined for 31 values in the retinal organoid set and 72 values in the skin set. Values in parentheses indicate standard deviations across five-fold cross-validation.
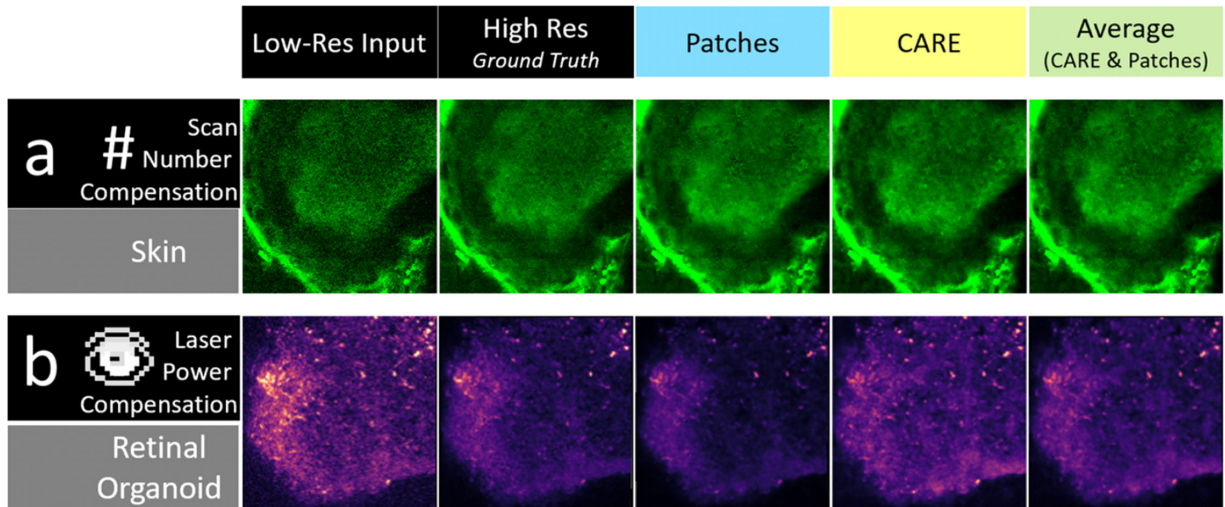
Figure 3. Representative images demonstrating performance of DL methods and the average of the output of each method. (a) Human skin sample imaging (epidermal keratinocytes) using DL to predict higher resolution images using a low scan number for image acquisition. Ground-truth images were acquired using a high scan number. (b) Retinal organoid imaging using DL to predict higher resolution images using a low-power laser for image acquisition. Ground-truth images were acquired using a high laser power.

Table 2. Comparison of the Two Methods by Number of Parameters and Time to Predict

	Patches	Care
Number of parameters	20,260,162	923,876
Time to predict one image	303.68 s	0.89 s

MSE values across five trials for CARE, patches, and average of CARE and patches were 10.41, 4.43, and 5.92, respectively. For the skin dataset, the mean SSIM values across five trials for CARE, patches, and average of CARE and patches were 0.69, 0.82, and 0.77, respectively. For the retinal organoid dataset, the mean MSE values across five trials for CARE, patches, and average of CARE and patches were 164,154, 26,424, and 59,154, respectively. For the retinal organoid dataset, the mean SSIM values across five trials for CARE, patches, and average of CARE and patches were 0.59, 0.64, and 0.61, respectively.

Discussion

One of the primary concerns with the application of advanced laser imaging to humans is safety. The retina is the tissue most sensitive to light. Therefore, minimizing the light exposure required to image the retina can safeguard imaging all human tissues in vivo. However, image quality is generally improved by increasing the light captured, which presents a contradictory requirement for tissue safety. In multipho-ton microscopy, fluorescence is excited by an ultrafast femtosecond laser, and the emitted light is detected. We sought to determine if multiphoton excitation of intrinsic fluorophores to image human skin and retinal organoid tissue can be enhanced using DL methodology without increasing light exposure.

In this paper, we presented two DL methods for improving the image quality of microscopy images acquired using two-photon excited fluorescence. The first method was based on U-Net, and the second was a patch-based regression model. We evaluated these methods on two datasets. The first dataset was acquired from human skin, and the second was from human retinal organoids. Both methods achieved good performance on both datasets. Although the patch-based method outperformed CARE in image-quality metrics, the former was trained for longer than the CARE U-Net method. Because both methods used different architectures, training parameters, and training time, these results should not be used to infer that the patch-based method would always outperform the CARE method. Rather, these results demonstrate that both methods could achieve good performance on both datasets. Therefore, we have demonstrated that DL could be used to reconstruct high-resolution images from lower resolution images acquired using lower excitation laser power or fewer excitation light scans. Although patch-based regression achieved higher image quality than the U-Net, it was slower at predicting new images. Qualitative evaluation of images produced by averaging CARE U-Net and patches suggested that a combination

of DL approaches might produce images with high resolution and faster computational times.

In conclusion, we have demonstrated that DL is a valuable approach to reconstructing high-resolution images from multiphoton microscopy on light-sensitive tissues with minimal light exposure needed to acquire an image when phototoxicity was a demonstrable concern.

Acknowledgments

Disclosure: **S. McAleer**, None; **A. Fast**, None; **Y. Xue**, None; **M.J. Seiler**, None; **W.C. Tang**, None; **M. Balu**, None; **P. Baldi**, None; **A.W. Browne**, None

References

1. Sanderson MJ, Smith I, Parker I, Bootman MD. Fluorescence microscopy. *Cold Spring Harb Protoc*. 2014;2014(10):pdb.top071795.
2. Masters BR, So PT. Antecedents of two-photon excitation laser scanning microscopy. *Microsc Res Tech*. 2004;63(1):3–11.
3. Masters BR, So PTC, Buehler C, et al. Mitigating thermal mechanical damage potential during two-photon dermal imaging. *J Biomed Opt*. 2004;9(6):1265–1270.
4. Abdelgawad M, Watson MWL, Young EWK, Mudrick JM, Ungrin MD, Wheeler AR. Soft lithography: masters on demand. *Lab Chip*. 2008;8(8):1379–1385.
5. Schwarz C, Sharma R, Fischer WS, et al. Safety assessment in macaques of light exposures for functional two-photon ophthalmoscopy in humans. *Biomed Opt Express*. 2016;7(12):5148–5169.
6. Schwarz C, Sharma R, Cheong SK, Keller M, Williams DR, Hunter JJ. Selective S cone damage and retinal remodeling following intense ultrashort pulse laser exposures in the near-infrared. *Invest Ophthalmol Vis Sci*. 2018;59(15):5973–5984.
7. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. Cambridge, MA: MIT Press, 2016.
8. Schmidhuber J. Deep learning in neural networks: an overview. *Neural Netw*. 2015;61:85–117.
9. Baldi P. Deep learning in biomedical data science. *Annu Rev Biomed Data Sci*. 2018;1:181–205.
10. Parmar C, Barry JD, Hosny A, Quackenbush J, Aerts HJWL. Data analysis strategies in medical imaging. *Clin Cancer Res*. 2018;24(15):3492–3499.
11. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316(22):2402–2410.
12. Asaoka R, Tanito M, Shibata N, et al. Validation of a deep learning model to screen for glaucoma using images from different fundus cameras and data augmentation. *Ophthalmol Glaucoma*. 2019;2(4):224–231.
13. Wang H, Rivenson Y, Jin Y, et al. Deep learning enables cross-modality super-resolution in fluorescence microscopy. *Nat Methods*. 2019;16(1):103–110.
14. Rivenson Y, Koydemir HC, Wang H, et al. Deep learning enhanced mobile-phone microscopy. *ACS Photonics*. 2018;5(6):2354–2364.
15. Urban G, Tripathi P, Alkayali T, et al. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. *Gastroenterology*. 2018;155(4):1069–1078.e8.
16. Chang P, Grinband J, Weinberg BD, et al. Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas. *AJNR Am J Neuroradiol*. 2018;39(7):1201–1207.
17. Wang J, Ding H, Bidgoli FA, et al. Detecting cardiovascular disease from mammograms with deep learning. *IEEE Trans Med Imaging*. 2017;36(5):1172–1181.
18. Wang J, Fang Z, Lang N, Yuan H, Su M-Y, Baldi P. A multi-resolution approach for spinal metastasis detection using deep Siamese neural networks. *Comput Biol Med*. 2017;84:137–146.

19. Urban G, Feil N, Csuka E, et al. Combining deep learning with optical coherence tomography imaging to determine scalp hair and follicle counts. *Lasers Surg Med*. 2021;53(1):171–178.

20. Baldi P, Chauvin Y. Neural networks for fingerprint recognition. *Neural Comput*. 1993;5(3):402–418.

21. Urban G, Bache KM, Phan D, et al. Deep learning for drug discovery and cancer research: automated analysis of vascularization images. *IEEE/ACM Trans Comput Biol Bioinform*. 2019;16(3):1029–1035.

22. Browne AW, Arnesano C, Harutyunyan N, et al. Structural and functional characterization of human stem-cell-derived retinal organoids by live imaging. *Invest Ophthalmol Vis Sci*. 2017;58(9):3311–3318.

23. Fast A, Lal A, Durkin AF, et al. Fast, large area multiphoton exoscope (FLAME) for macroscopic imaging with microscopic resolution of human skin. *Sci Rep*. 2020;10(1):18093.

24. Balu M, Potma E, Tromberg B, Mikami H, inventors. Imaging platform based on nonlinear optical microscopy for rapid scanning large areas of tissue [patent application WO2018075562]. Oakland, CA: University of California; 2018.

25. Balu M, Mikami H, Hou J, Potma EO, Tromberg BJ, et al. Rapid mesoscale multiphoton microscopy of human skin. *Biomed Opt Express*. 2016;7(11):4375–4387.

26. Collin J, Mellough CB, Dorgau B, Przyborski S, Moreno-Gimeno I, Lako M. Using zinc finger nuclease technology to generate CRX-reporter human embryonic stem cells as a tool to identify and study the emergence of photoreceptors precursors during pluripotent stem cell differentiation. *Stem Cells*. 2016;34(2):311–321.

27. Wahlin KJ, Maruotti JA, Sripathi SR, et al. Photoreceptor outer segment-like structures in long-term 3D retinas from human pluripotent stem cells. *Sci Rep*. 2017;7(1):766.

28. Zhong X, Gutierrez C, Xue T, et al. Generation of three-dimensional retinal tissue with functional photoreceptors from human iPSCs. *Nat Commun*. 2014;5:4047.

29. Zhang Z, Liu Q, Wang Y. Road extraction by deep residual U-Net. *IEEE Geosci Remote Sens Lett*. 2018;15:749–753.

30. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60–88.

31. Falk T, Mai D, Bensch R, et al. U-Net: deep learning for cell counting, detection, and morphometry. *Nat Methods*. 2019;16(1):67–70.

32. Dong H, Yang G, Liu F, Mo Y, Guo Y. Automatic brain tumor detection and segmentation using u-net based fully convolutional networks. In: Valdés Hernández M, González-Castro V, eds. *Medical Image Understanding and Analysis. MIUA 2017*. Cham: Springer; 2017:506–517.

33. Ciresan DC, Giusti A, Gambardella L, Schmidhuber J, et al. Deep neural networks segment neuronal membranes in electron microscopy images. In: Pereira F, Burges CJC, Bottou L, Weinberger KQ, eds. Advances in Neural Information Processing Systems 25 (NIPS 2012). La Jolla, CA: Neural Information Processing Systems Foundation; 2012:2852–2860.

34. Kingma DP, Ba J. Adam: a method for stochastic optimization. ArXiv. 2017, arXiv:1412.6980v9.

35. Weigert M, Schmidt U, Boothe T, et al. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat Methods*. 2018;15(12):1090–1097.

36. Ronneberger O, Fischer P, Brox T., Navab N, Hornegger J, Wells WM, Frangi AF, eds. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*. Cham: Springer International Publishing; 2015:234–241.

37. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process*. 2004;13(4):600–612.