

Problem: DEEP-TFM still needs hundreds of patterned excitations to de-scatter at depth, limiting throughput for *in vivo* imaging.

ARTICLE

Open Access

DEEP-squared: deep learning powered De-scattering with Excitation Patterning

Navodini Wijethilake^{1,2}, Mithunjha Anandakumar¹, Cheng Zheng^{3,4}, Peter T. C. So^{1,3,4,5}, Murat Yildirim^{1,4,6,7} and Dushan N. Wadduwage¹✉

Abstract

Limited throughput is a key challenge in *in vivo* deep tissue imaging using nonlinear optical microscopy. Point scanning multiphoton microscopy, the current gold standard, is slow especially compared to the widefield imaging modalities used for optically cleared or thin specimens. We recently introduced “De-scattering with Excitation Patterning” or “DEEP” as a widefield alternative to point-scanning geometries. Using patterned multiphoton excitation, DEEP encodes spatial information inside tissue before scattering. However, to de-scatter at typical depths, hundreds of such patterned excitations were needed. In this work, we present DEEP², a deep learning-based model that can de-scatter images from just tens of patterned excitations instead of hundreds. Consequently, we improve DEEP’s throughput by almost an order of magnitude. We demonstrate our method in multiple numerical and experimental imaging studies, including *in vivo* cortical vasculature imaging up to 4 scattering lengths deep in live mice.

Introduction

Imaging biological structures such as neurons or blood vessels deep inside scattering tissue is an important yet challenging microscopy problem. For such *in vivo* deep tissue experiments, today, the gold standard is the point-scanning two- (or three-) photon microscopy (PSTPM)^{1–3}. Especially, PSTPM occupies a unique position in neuroscience due to its high spatial resolution, low phototoxicity, and ability to penetrate through deep tissue. In operation, PSTPM focuses femtosecond laser light at long wavelengths through scattering tissue to excite fluorescent molecules. Due to long wavelengths, the excitation light sees only slight scattering. The microscope then collects emission fluorescence through the same scattering tissue onto a detector. This emission light, however, is at a much shorter wavelength and hence

encounters significant scattering on its way to the detector. For this reason, multiple points cannot be excited and resolved simultaneously. So imaging is performed sequentially, one point at a time. Thus, PSTPM is inherently a slow imaging technique. When speed is needed, either resolution or the imaging field of view should be compromised⁴.

Despite the limitations due to emission light scattering, widefield two- and three-photon microscopes (WFTPM) have been proposed. For instance, temporal focusing microscopy (TFM) focuses amplified femtosecond laser pulses temporally for depth-selective WFTPM. TFM provides excellent optical sectioning and also penetrates well through scattering tissue due to long multiphoton excitation wavelengths^{5–8}. Yet, the emission light scatters. Therefore, in widefield deep tissue imaging, some photons are mapped onto incorrect detector pixels, degrading the signal-to-noise ratio and spatial resolution. A few groups, including us, recently proposed to overcome this limitation by combining TFM with structured light illumination. Since excitation light penetrates through the scattering medium, structured illuminations can modulate the imaging field of view before scattering. Thus, the

Correspondence: Dushan N. Wadduwage (wadduwage@fas.harvard.edu)

¹Center for Advanced Imaging, Faculty of Arts and Sciences, Harvard University, Cambridge, MA, USA

²Department of Electronic and Telecommunication Engineering, University of Moratuwa, Moratuwa, Sri Lanka

Full list of author information is available at the end of the article

These authors contributed equally: Navodini Wijethilake, Mithunjha Anandakumar

information about the structured illuminations can be used to de-scatter TFM images despite degraded detection. Using this principle, Escobet et al.⁹ proposed “TempoRAL Focusing microscopy with single-pIXel detection (TRAFIX).” They used coded excitations and single-pixel detection. Subsequently, we proposed “De-scattering with Excitation Patterning (DEEP)”¹⁰. Instead of single-pixel detection, we used widefield detection. Single-pixel detection completely relies on coded illuminations for de-scattering. Widefield detection, however, retains some spatial information and only partially depends on coded illumination. Thus, while TRAFIX needs tens of thousands of illumination patterns to reconstruct a typical WFTPM image, DEEP requires only hundreds. Despite such speed improvement, DEEP is still slower than other depth-resolved widefield imaging modalities used for optically cleared or thin specimens. For instance, one-photon fluorescence modalities, that optically section with planar or structured illumination, are either single-shot or few-shot acquisition¹¹. Therefore, in this work, we attempt to further increase the throughput of DEEP by utilizing prior image information within a deep-learning-based computational imaging framework.

From a computational imaging point of view, the imaging process that describes the translation of the ideal image (x) to the observed image (y) is called the forward model (f). During the forward imaging process, the observed image (y) is degraded by image noise; low-pass filtering; pixel-value quantization; sub-sampling; and, in our case, scattering¹². An inverse model is needed to map the observed (y) back to the ideal expected image (x). In our original work¹⁰, we used an analytical inverse model with wavelet sparsity priors. However, recently, for such inverse problems, deep learning has proven impressively capable^{12–15}. For example, deep learning has shown outstanding success in image classification^{16,17}, biomedical imaging^{18–20}, segmentation²¹, prediction²², denoising²³ and other linear and nonlinear inverse problems^{14,24–26}. In our previous work, we demonstrated that deep learning can reconstruct fine biological structures such as dendritic spines from scattered TFM images²⁷. Similarly, in this work, we present a deep learning-based inverse reconstruction method for DEEP-TFM.

Our proposed method is termed DEEP². DEEP² consists of a learning-powered inverse model that can reconstruct a de-scattered image from only 32 DEEP measurements (instead of 256 in our original work, DEEP). The model architecture is inspired by the UNet but modified to include a self-attention mechanism on the expanding path and a terminating reconstruction block. Generating paired DEEP-TFM and ground-truth images to train the DEEP² inverse model is impractical. So we trained DEEP² using simulated data. To simulate training data, we first

carefully modeled the forward imaging process of the DEEP microscope. As a part of the forward model, we also modeled the optical behavior of scattering tissue using Monte Carlo simulations. We then used the forward model to simulate DEEP images from PSTPM-like image stacks (synthetically generated or experimentally acquired). The DEEP², trained on the simulated data, was finally tested on the experimentally acquired DEEP images of fluorescent beads and mouse cortical vasculature. Our results suggest that the proposed method can reconstruct deep tissue images up to 4 scattering lengths (at the emission wavelengths)—accounting for ~200 μm—deep in mouse cortical vasculature, using only tens of patterned excitations instead of hundreds.

Results

Deep learning powered DEEP microscopy

In this work, we propose a deep-learning-based inverse model to reconstruct images from DEEP measurements. Figure 1a shows the optical schematic of the DEEP-TFM microscope. Amplified femtosecond laser pulses are temporally focused for depth-resolved multiphoton excitation. In the excitation path, a digital micro-mirror device (DMD) is placed at a conjugate image plane before the excitation tube lens. The DMD mirrors project a binary-patterned excitation onto the sample at the focal plane. The emission photons are collected by the objective lens and imaged onto a camera detector. During the experiment, multiple patterns generate multiple encoded image measurements. Our goal is to learn an inverse model to reconstruct de-scattered images using the knowledge about the excitation patterns and measured images (see Fig. 1b). To train such a model, we first generated simulated training data using a physics-based forward image model of the DEEP-TFM microscope and the scattering process. In the Methods section, we discuss the forward model, the electron multiplication CCD (EMCCD) detector model, and the scattering model in detail. The excitation patterns are parameters of the forward model. We first measured these patterns in a calibration experiment. Next, PSTPM-like ground-truth images were input to the forward model to simulate DEEP images (see Fig. 1b). This allowed us to generate paired DEEP and PSTPM data to train the inverse model. The inverse model was based on the seminal UNet architecture. We further modified the UNet by adding an attention mechanism to the expanding path. Details about the inverse model are presented in the Method section.

To test our method, we experimented on three datasets: (1) a mixture of fluorescent beads, (2) mouse pyramidal neurons with dendritic arbors, and (3) mouse cortical vasculature. We synthetically generated PSTPM-like beads images for training (see methods). Mouse pyramidal neurons and cortical vasculature were imaged

Contributions: Cuts required patterns from $\sim 256 \rightarrow 32$ ($\approx 10\times$ faster), validates on simulations + experiments (beads, neurons, cortical vasculature), and demonstrates *in vivo* mouse vasculature de-scattering up to ~ 4 scattering lengths ($\sim 200 \mu\text{m}$) using tens of patterns.

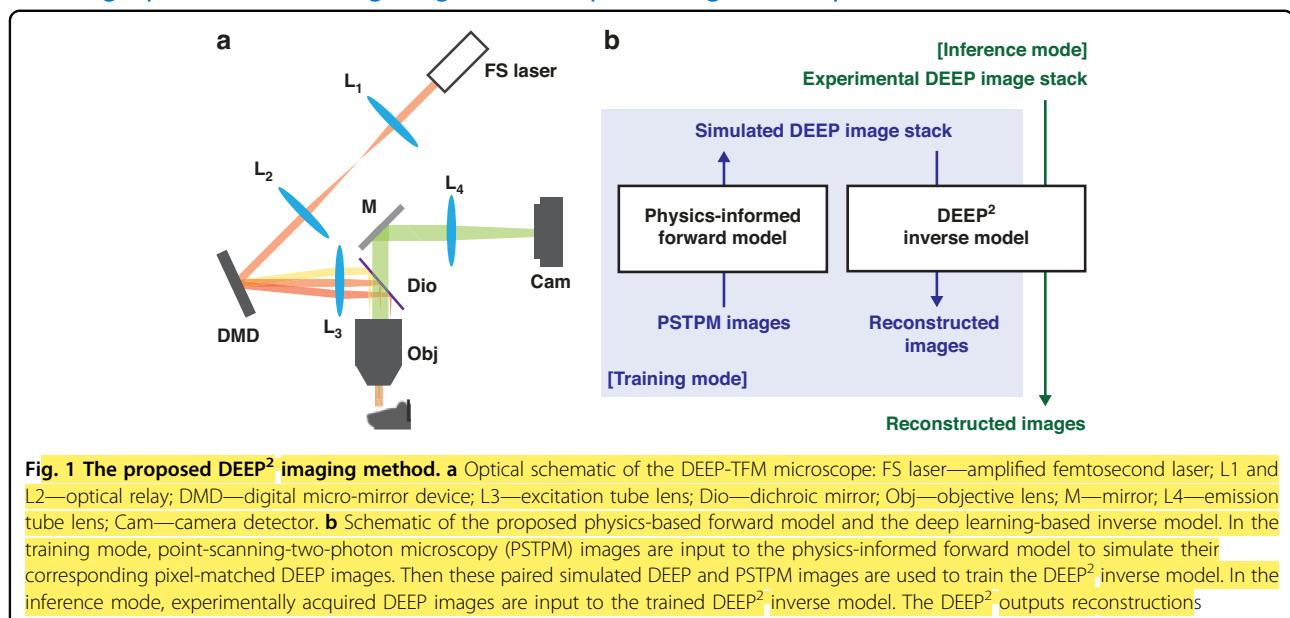


Fig. 1 The proposed DEEP² imaging method. **a** Optical schematic of the DEEP-TFM microscope: FS laser—amplified femtosecond laser; L1 and L2—optical relay; DMD—digital micro-mirror device; L3—excitation tube lens; Dio—dichroic mirror; Obj—objective lens; M—mirror; L4—emission tube lens; Cam—camera detector. **b** Schematic of the proposed physics-based forward model and the deep learning-based inverse model. In the training mode, point-scanning-two-photon microscopy (PSTM) images are input to the physics-informed forward model to simulate their corresponding pixel-matched DEEP images. Then these paired simulated DEEP and PSTM images are used to train the DEEP² inverse model. In the inference mode, experimentally acquired DEEP images are input to the trained DEEP² inverse model. The DEEP² outputs reconstructions

from multiple animals using PSTM to create the training datasets. We then performed physical DEEP experiments on a mixture of fluorescent beads and cortical vasculature of anesthetized mice (see Methods). We finally used our inverse model (trained on simulated data) to successfully de-scatter experimental DEEP measurements.

In the next two sections, we present numerical reconstruction results on unseen test datasets for each case; and de-scattering results for experimental DEEP measurements. We present all results in terms of “scattering lengths” or “scattering length-depths.” A scattering length is defined as the mean-free-path, i.e., the inverse of the scattering coefficient of the tissue at the emission wavelength. In experimental results, we assumed a scattering length of $50 \mu\text{m}$ ^{10,28}. Therefore 2, 4, and 6 scattering lengths, respectively, refer to 100, 200, and 300 μm .

Numerical results on simulated test data

This section presents numerical results on simulated test datasets to validate our trained inverse model. All the experiments used only 32 patterned illuminations, which is nearly an order of magnitude less than the current state of the art¹⁰.

Reconstruction results on synthetic beads data

Synthetic 3D bead stacks were used as the input to the forward model to obtain the simulated DEEP-TFM-like instances at 4 scattering lengths. Then, the DEEP²-based inverse model was trained on the synthetic beads instances and was validated on similarly generated unseen synthetic beads instances. Figure 2 presents two validation instances, where the average input is shown in the first column, the ground truth in the second column, and the reconstruction

of the DEEP² in the third column. In addition, the intensities along the lateral directions (G, H, I, J, K and L in Fig. 2) were plotted for further clarification. The DEEP² successfully reconstructed the small 1 μm beads (i.e., beads with 0.5 μm radius).

Mouse pyramidal neuron reconstructions

Figure 3 shows the validation cohort results on the mouse pyramidal neurons with a dendritic arbor. We present two instances corresponding to 2 and 6 scattering length depths. The respective intensities along the lateral directions on each instance are shown in the last column. Figure 3a, d is two separate instances, each representing average of 32 patterned DEEP-TFM images generated using the forward model with the PSTM mouse pyramidal neuron images at 2 scattering lengths. Figure 3b, e represents the PSTM images, which were used as the ground truth for the corresponding Fig. 3a, d instances. Figure 3c, f is the reconstruction of DEEP². The intensity along the yellow line on Fig. 3a–c and Fig. 3d–f are visualized in Fig. 3m, n graphs, respectively. Similarly, Fig. 3g, j shows an average of 32 patterned DEEP-TFM images synthesized using the forward model at six scattering lengths. Figure 3h, k shows corresponding ground truth, and Fig. 3i, l shows the reconstruction of DEEP² at six scattering lengths. The intensity along the yellow line on Fig. 3g–i and Fig. 3j–l are represented on the Fig. 3o, p graphs, respectively.

Mouse cortical vasculature reconstructions

Next, we trained DEEP² to reconstruct mouse cortical vasculature images from numerical DEEP data. Figure 4

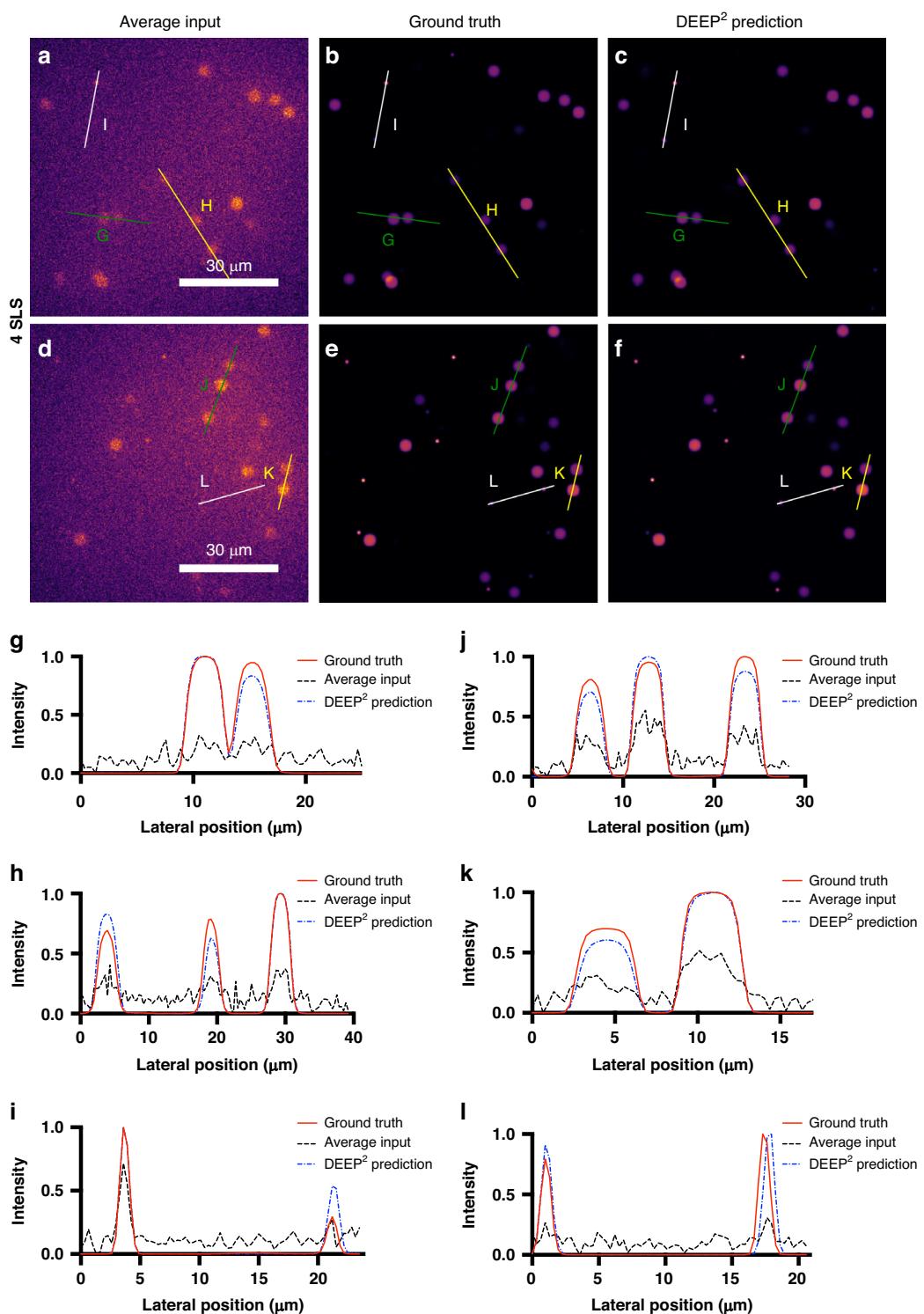
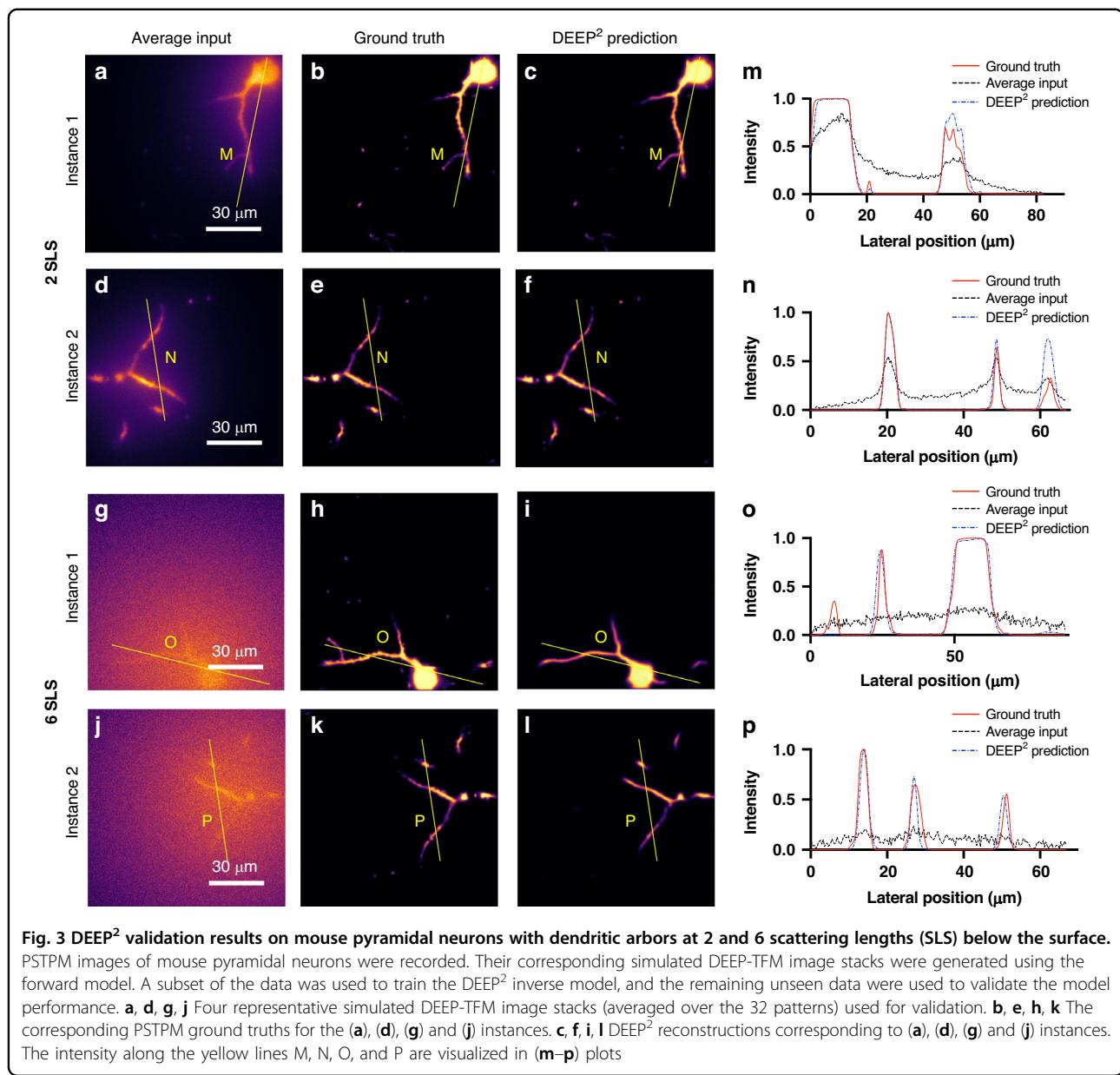


Fig. 2 DEEP² validation results on Synthetic Fluorescent Beads at 4 scattering lengths (SLS) below the surface. Synthetic beads-objects of the size of 1 and 4 μm , with the intensity of the 1 μm beads 5x higher than the 4 μm beads, were simulated. Their corresponding simulated DEEP-TFM image stacks were generated using the forward model. A subset of the data was used to train the DEEP² inverse model, and the remaining unseen data were used to validate the model performance. **a, d** Two representative simulated DEEP-TFM image stacks (averaged over the 32 patterns). **b, e** Corresponding synthetic ground truths for the **(a)** and **(d)** instances. **c, f** DEEP² reconstructions corresponding to **(a)** and **(d)** instances. The intensity along the lines G, H, and I shown in green, yellow and white on **(a)** is visualized in **(g–i)**. Similarly, the intensity along the lines J, K, and L are shown in **(j–l)** plots

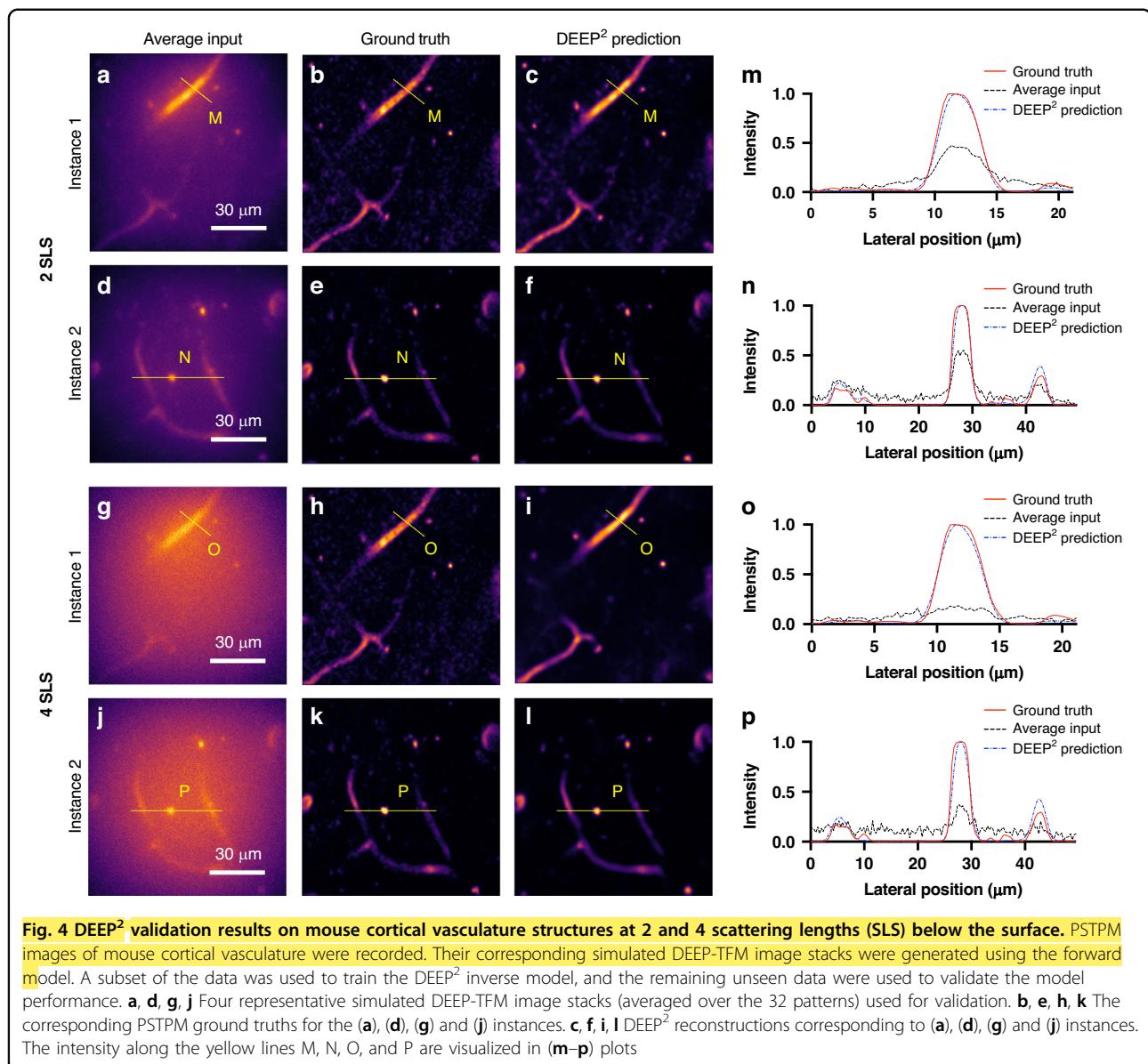


shows two validation instances at 2 and 4 scattering lengths below the surface. The intensity profiles along the highlighted lateral directions are presented in the last column. Figure 4a, d is the average of 32 patterned DEEP-TFM images, simulated using the forward model, for two separate mouse cortical vasculature instances. Figure 4b, e shows the corresponding ground truths. The DEEP² reconstruction for the same two instances at 2 scattering lengths are presented in Fig. 4c, f. The intensity profiles along the marked yellow lines on Fig. 4a–c and Fig. 4d–f are indicated in Fig. 4m, n graphs, respectively. Similarly, for the same vasculature instances, the average over the 32 patterned DEEP-TFM instances synthesized using the forward model at

4 scattering lengths are shown in Fig. 4g, j. The corresponding ground truths are shown in Fig. 4h, k, and the DEEP² reconstructions at 4 scattering lengths are shown in Fig. 4i, l. The intensity along the lateral directions marked in yellow on Fig. 4g–i and Fig. 4j–l are represented in Fig. 4o, p.

Results on experimental data

In this section, we present experimental validation of DEEP² by successfully de-scattering experimentally acquired DEEP measurements using the inverse model trained on simulated data. Experiments used only 32 patterned illuminations. Thus DEEP² is nearly an order of magnitude faster than the current state of the art¹⁰.



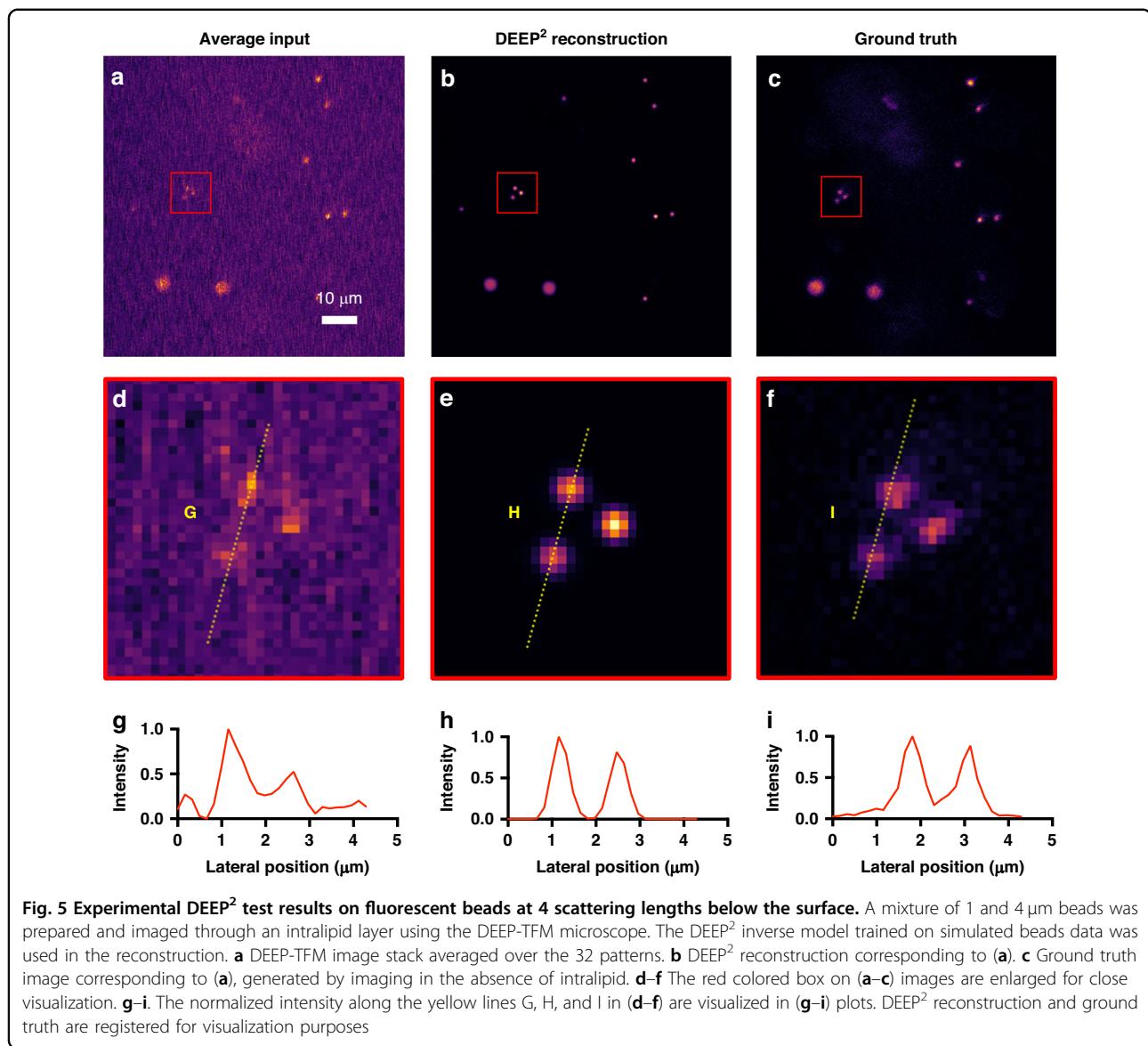
DEEP² imaging and reconstruction of a mixture of fluorescent beads

The DEEP² inverse model trained on the artificial beads data at 4 scattering lengths was tested on experimentally acquired DEEP-TFM images. In accordance with the training data, similar-sized beads at a similar composition were imaged through a scattering intralipid layer (resulting in 4–5 scattering lengths) using the DEEP-TFM microscope with 32 patterned excitations. The images were then input to the trained inverse model to output the reconstructions. The results are shown in Fig. 5. Figure 5a shows the average of the 32 patterned DEEP-TFM images, and Fig. 5b shows the DEEP² reconstruction. Figure 5c shows the ground truth acquired without the scattering intralipid layer. A region that contains 1 μm beads,

marked as a red box on Fig. 5a, was closely observed to evaluate the performance. As shown in Fig. 5d, the average of 32 patterned DEEP-TFM images contained faint signs of those structures. But the proposed DEEP² could reconstruct those fine beads, eliminating the background noise as shown in Fig. 5e. Finally, we also show the intensity cross-sections across the same beads, along the yellow lines drawn in Fig. 5d–f in Fig. 5g–i.

DEEP² imaging and reconstruction of *in vivo* mouse cortical vasculature

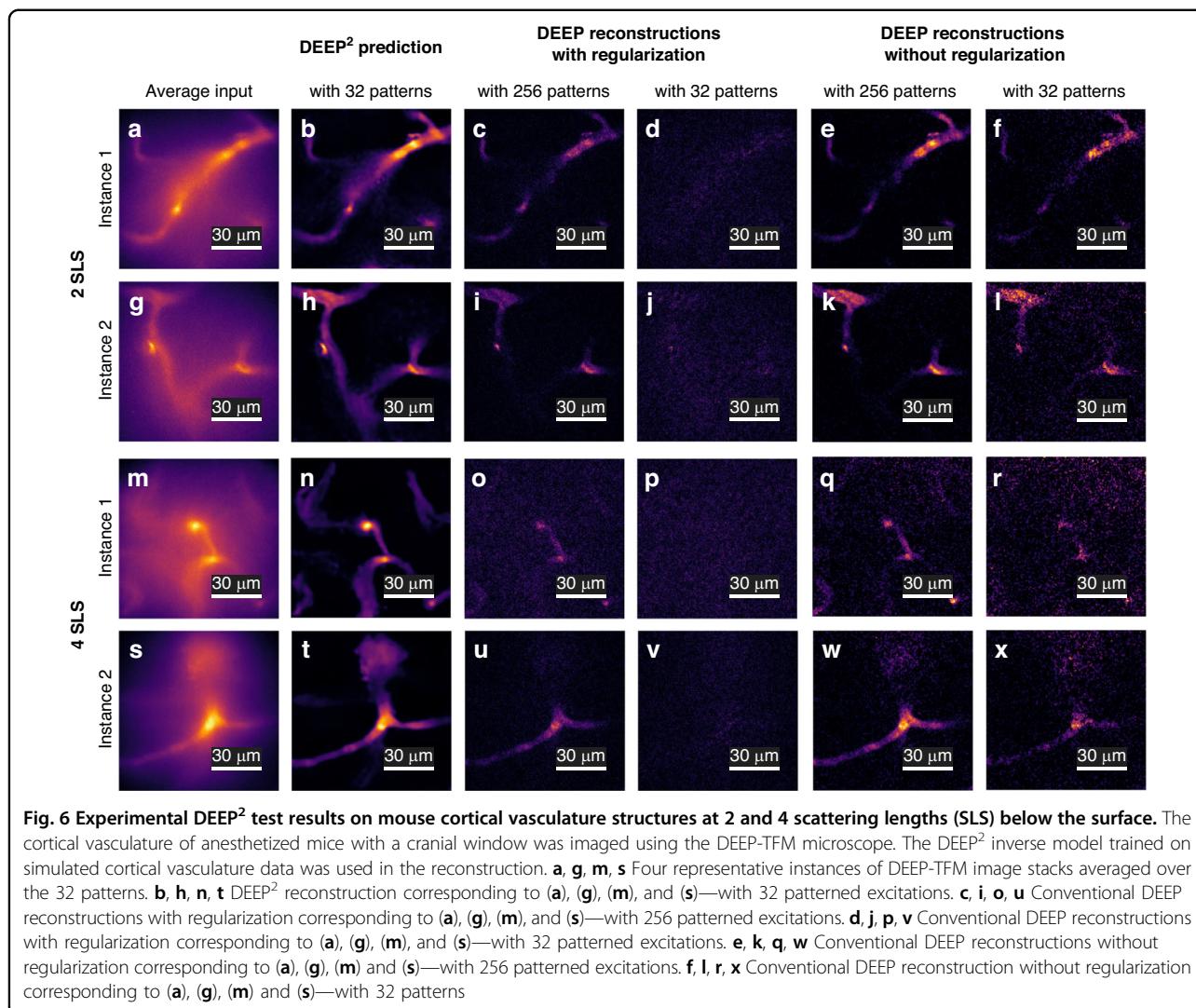
Next, the DEEP² model trained on simulated cortical vasculature data was tested on experimentally acquired DEEP-TFM images of mouse cortical vasculature (see Methods section for imaging details). Figure 6 shows several



mouse cortical vascular instances at 2 and 4 scattering length depths. Figure 6a, g shows the average of the 32 patterned DEEP-TFM images for two separate FOVs at 2 scattering lengths. Figure 6b, h is the corresponding DEEP² reconstructions with 32 DEEP-TFM images. Figure 6c, i are the DEEP-TFM reconstructions with 256 DEEP-TFM images using the original DEEP algorithm with a regularizer¹⁰. Figure 6d, j is the DEEP-TFM reconstruction with 32 DEEP-TFM images using the original DEEP inverse algorithm with the same regularizer¹⁰. We noticed that using a regularizer to remove noise may also remove dim image features (we also investigated this issue for DEEP² as shown in Supplementary Fig. S14). Therefore, a similar DEEP reconstruction was also performed using a pseudo-inverse solver with no regularization (see corresponding results in Fig. 6e, f, k, l¹⁰).

Similarly, Fig. 6m, s shows the average of the 32 patterned DEEP-TFM images for two separate FOVs at 4 scattering lengths. Figure 6n, t is the DEEP² reconstruction with 32 DEEP-TFM images corresponding to Fig. 6m, s, respectively. Figure 6o, u are the DEEP inverse algorithm reconstructions with regularization using 256 DEEP-TFM images at 4 scattering lengths. Figure 6p, v is the DEEP inverse algorithm reconstruction with regularization using 32 DEEP-TFM images at 4 scattering lengths. Figure 6q, r, w, x shows the same corresponding results without using the regularizer.

The DEEP² reconstruction (with 32 DEEP-TFM images) showed additional details missing in the inverse algorithm reconstructions even with 256 DEEP-TFM images. This was more evident for DEEP reconstructions with a regularized inverse solver. However,



the average of raw data (see Fig. 6a, g, m, s) suggests that these structures are real and not false artifacts generated by the network. The inverse algorithm with regularization failed to reconstruct with 32 DEEP-TFM images (see Fig. 6d, j, p, v), possibly due to a sub-optimal regularization loss term. But the inverse solver without a regularizer managed a noisy reconstruction for 2 scattering lengths.

Discussion

In this paper, we developed a physics-informed-forward model and a deep-learning-based inverse model called DEEP² to de-scatter DEEP-TFM images up to 4 scattering lengths below the surface. Our model uses only 32 DEEP-TFM raw images in the reconstruction, while the original DEEP used 256 such images. Therefore DEEP² is almost an order of magnitude faster than DEEP. We validated our approach numerically and tested it experimentally to

image *in vitro* samples through scattering media, as well as to image *in vivo* mouse cortical vasculature.

One major technical contribution of our work is the physics-based forward model that can simulate realistic DEEP-TFM images using PSTPM image data. This approach eliminates the need to collect paired PSTPM and DEEP data—which is nearly impossible for *in vivo* experiments—to train the neural network inverse model. The inverse model, trained on simulated data, was first validated on unseen simulated data (validation experiments). The same model was then tested on new experimental data (test experiments). In the validation experiments, the model worked up to six scattering lengths on multiple datasets (see Supplementary Figs. S2–S4). In *in vivo* test experiments, however, the model worked only up to 4 scattering lengths; the model failed at six scattering lengths (see Supplementary Fig. S6). This suggests a model mismatch between the forward model

and the experimental system at higher scattering lengths. Therefore, an interesting future direction is to investigate reasons for the model mismatch in controlled test experiments using *in vitro* samples.

We tested the sensitivity of the inverse model to the changes in the scattering length. The models each trained on 2, 4, and 6 scattering lengths were tested on all 3 scattering lengths (see Supplementary Fig. S7). The models were indeed sensitive to the training conditions. We also further experimented with the model trained on 2 scattering lengths on validation data from a range of 1.6–2.4 scattering lengths. We observed a gradual drop in performance on off-training-data conditions (see Supplementary Fig. S10). To overcome this issue, we then trained a single model with a mix of scattering lengths (see results in Supplementary Fig. S7d, e). This model generalized to all scattering lengths in the training data mix. This observation suggests that a single model may be trained to generalize to a range of scattering lengths by mixing these conditions in the simulated training data. We also tested its sensitivity to off-training data conditions at 1.6–2.4 scattering lengths. The performance loss was lower than the model trained only on 2 scattering lengths training data (compare the red plot to the black—see Supplementary Fig. S10). The model trained on a mix also seemed to perform better at higher scattering lengths (see the last column in Supplementary Fig. S7e). Furthermore, we performed an ablation study to investigate the importance of different parts of the forward model (see results in see Supplementary Fig. S11). We observed that the EMCCD noise model and scattering model were critical parts of the simulated data generation process.

We also investigated the minimum number of patterns needed for reconstruction by training an ensemble of models that reconstruct from 1, 2, 4, 8, 16, and 32 patterns at 2, 4, and 6 scattering lengths. As expected, the minimum number of patterns needed for satisfactory performances depended on the scattering length. As shown in Supplementary Fig. S9e for the 2-scattering-length mouse cortical vasculature data, a stable and satisfactory performance was achieved even with four patterns. However, 6-scattering-length mouse cortical vasculature data required a minimum of 32 patterns to perform satisfactorily. We believe that a higher number of patterns could give better performance for deeper scattering lengths. But there is a tradeoff between performance and computational requirements, and we were not able to successfully experiment with a higher number of patterns. As a conclusion, we chose 32 as the number of patterns considering the computational limitations and the model's performance with the different number of patterns. We also extensively compared the original DEEP¹⁰ under the same conditions (see Supplementary Figs. S12 and S13). In numerical validations, DEEP could reconstruct

marginally satisfactorily at 2 scattering lengths with 32 patterns (see structural similarity index (SSIM), peak signal-to-noise ratio (PSNR), and qualitative results in Supplementary Fig. S12b, d). But at 4 and 6 scattering lengths, DEEP reconstructions were not satisfactory (see Supplementary Fig. S12). We observed similar trends in the DEEP reconstruction on experimental cortical vasculature data (Supplementary Fig. S13). This benchmarking experiment clearly demonstrates the power of DEEP² compared to DEEP.

We also tested the effect of architecture and cost function of the inverse model with respect to the forward model mismatch. To this end, we tested five variations of the inverse model on three datasets. The validation experiments did not show significant differences in qualitative and quantitative performance among the five variations (see Supplementary Figs. S1–S4). The test experiments on real data, however, showed that the combination of KL divergence loss with the channel and spatial attention mechanism (scSE-UNet) helps with the forward model mismatch (see Supplementary Figs. S5 and S6). Further numerical experiments and interpretation are needed to definitively interpret the reasons for this behavior with different forward model variations. Our current DEEP² approach requires training a separate inverse model for every forward model variation. Therefore, testing a wide range of forward model variations is not realistic. A future direction to overcome this limitation is to develop forward-model-agnostic inverse approaches using deep generative priors. Such an approach will not require re-training a separate inverse model for each forward model variation; hence, one can run many numerical experiments to investigate the model mismatch problem. This will also help investigate other behaviors, such as the effect of the number of DEEP-TFM raw images in the reconstruction and the effect of the excitation pattern selection. In terms of inverse model selection, another interesting direction is exploring efficient architectures that can be trained for more patterns than 32.

For experimental results, some DEEP² reconstructions looked visually better than the DEEP reconstructions using even full 256 patterns. We think this is mostly because the DEEP² inverse network acting as a denoizer (in addition to a de-scatterer). For DEEP, we attempted some denoising using regularized inverse solvers with limited success. Moreover, DEEP² (as well as DEEP) demonstrated enhanced depth sectioning that could make the reconstructed results look better than the pseudo ground truth results (for example, in the case of the beads). This is because DEEP² (as well as DEEP) uses structured (i.e., patterned) illumination that may lead to better depth sectioning compared to widefield illumination used to generate the pseudo ground truth. Nevertheless, further controlled experiments are needed to definitively rule out the possibility of false reconstructions due to model overfitting. Finally, while we have improved

the speed of DEEP² by almost an order of magnitude compared to DEEP, both DEEP² and DEEP still require further optimization in instrument design to outperform traditional point-scanning and line-scanning systems. DEEP uses an EMCCD detector that is orders of magnitude slower than photomultiplier tubes (PMTs) used in point-scanning systems. To improve the practical speed in DEEP², the next-generation instrument needs to be built with fast sensors like multi-anode PMTs or avalanche photodiode arrays. Nevertheless, these instrumentation advances are orthogonal to the algorithmic advanced presented in this work. Hence, DEEP² can improve the speed of any potential DEEP system with advanced optical designs.

In conclusion, in this work, we demonstrate DEEP², a physics-informed forward model, and a deep-learning-based inverse model for DEEP-TFM. DEEP² is an order of magnitude faster than the original DEEP¹⁰. We validated our approach on unseen numerical data, as well as experimentally acquired test data. We demonstrate that DEEP² can de-scatter *in vivo* cortical vasculature images up to 4 scattering lengths below the surface. Our work set the foundation to develop data-driven inverse solvers for DEEP-TFM using deep learning, and hence an important milestone for *in vivo* widefield multiphoton microscopy.

Materials and methods

Forward model

Model architecture

To be used as inputs to our deep learning model for training, we generate DEEP-TFM image stacks synthetically from the PSTPM image stacks using a forward imaging model (see Fig. 7). Let $exPSF(x, y, z)$ be the 3D excitation point spread function (PSF) and $H_t(x, y)$ be the pattern on the DMD placed at an image plane at the excitation side. Now, $exPSF(x, y) *_{3D} H_t$ is the 3D patterned excitation on the sample. Here $*_{3D}$ is the 3D convolution operation. The result is multiplied element-wise with the object, $X0(x, y, z)$ to get the excited object cube. Then the resulting 3D excited object is convolved plane-by-plane with the stack of 2D scattering PSFs at each depth ($sPSF(x, y, z)$). This operation is represented by $*_{2D}$. Each $sPSF$ at a certain depth, represents the effective scattered light distribution from a point source at that depth. We introduce and discuss the sPSFs model in the next section in detail. The resulting 3D cube is then convolved in 3D with the emission PSF ($emPSF(x, y, z)$). These operations constitute the following equation:

$$Y_t(x, y, z) = \{ \{ (exPSF(x, y, z) *_{3D} H_t(x, y)) \\ \circ X0(x, y, z) \} *_{2D} sPSF(x, y, z) \} *_{3D} emPSF(x, y, z) \quad (1)$$

Finally, we get the image on the detector, by selecting the z -plane that corresponds to the focal plane (z_{focal}) from

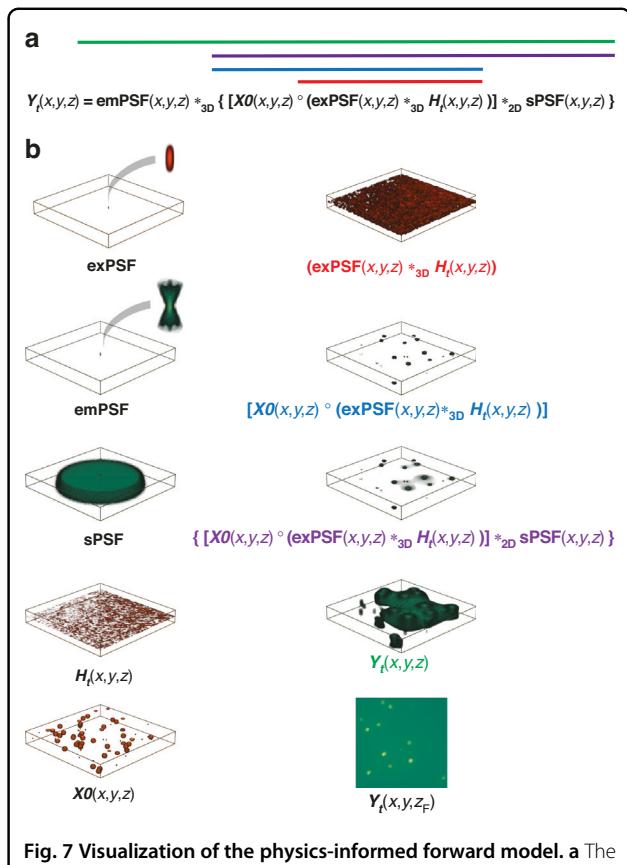


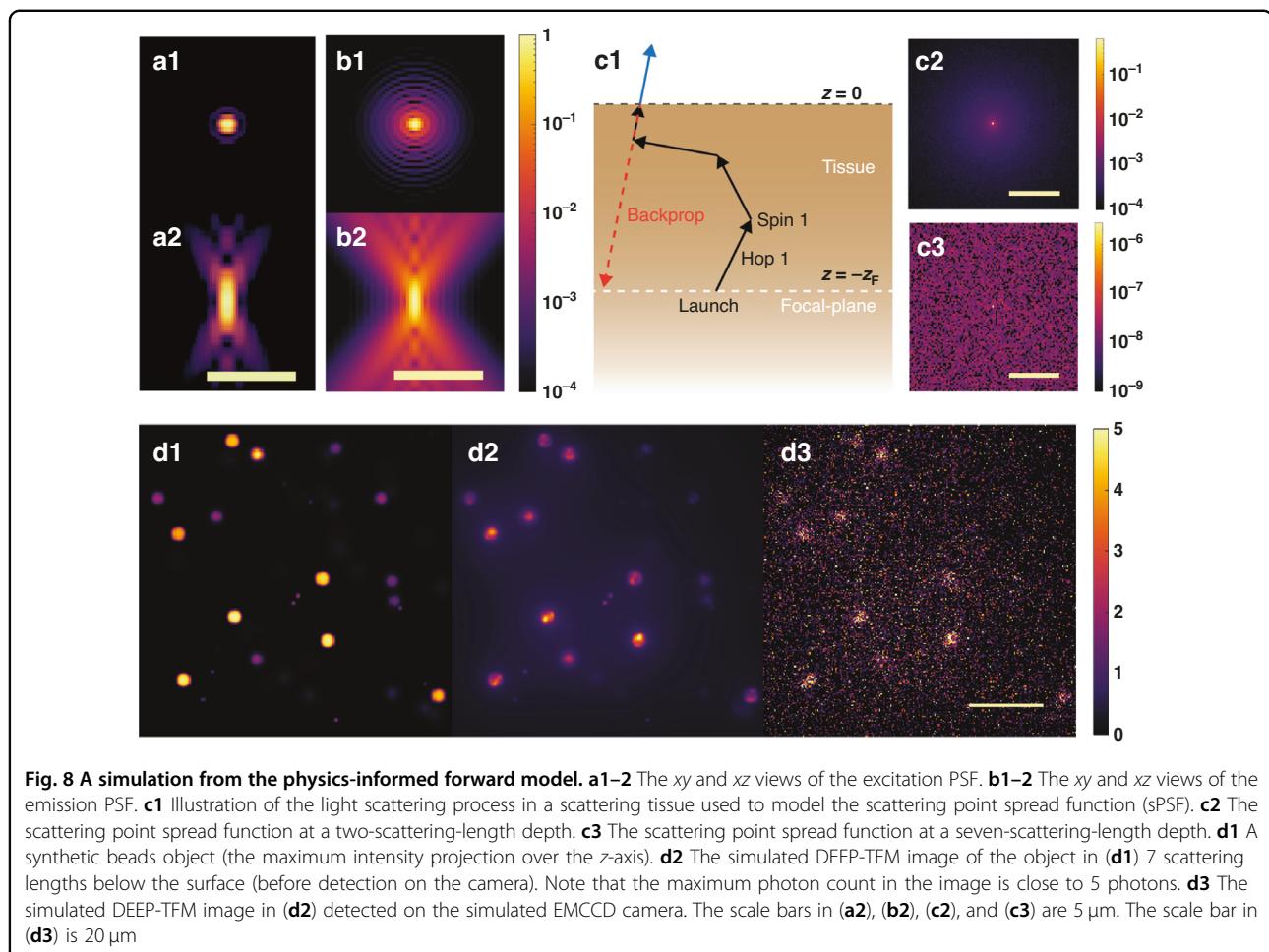
Fig. 7 Visualization of the physics-informed forward model. a The forward model equation. **b** A representative simulation of the forward model for a 3D beads specimen. Note that the different sections of the equation in (a) are shown in color-coded bars on the equation in (b)

the last output cube as shown in Fig. 7. Then the noise is added to the resulting image using the following noise model discussed in the “EMCCD noise model” section.

$$\hat{Y}_t(x, y) = f_{EM}(\sim Poiss(Y_t(x, y, z_{focal})) \\ + \sim Poiss(\bar{Y}_{Dark})) + \sim Norm(0, \sigma_{Read}) \quad (2)$$

Here, $\sim Poiss(\mu)$ denotes the observations drawn from a Poisson distribution of mean μ ; $\sim Norm(\mu, \sigma)$ denotes the observations drawn from a Normal distribution of mean μ and standard deviation σ ; \bar{Y}_{Dark} is the expected value of dark current of the camera and $f_{EM}(.)$ models the electron multiplying (EM) process of the EMCCD camera. Figure 8 shows the cross-sections of the simulated PSFs and representative simulated DEEP-TFM images.

For the beads dataset, fully simulated objects were used as $X0$ s. For mouse pyramidal neuron and mouse cortical vasculature datasets, experimentally acquired PSTPM images were used as $X0$ s after pre-processing them to match the voxel size of the DEEP-TFM experiment. During training, the PSTPM images were simulated and used as the output (or the ground truth)



of the network. The object X_0 was convolved with the excitation PSF to generate the PSTPM image.

Scattering point spread function (sPSF)

In this section, we discuss the mathematical model of the scattering PSF. We define the scattering PSF as the effective photon distribution of a certain z -plane below the tissue surface due to a point source placed at that z -plane. Using Monte Carlo simulations of light transport through scattering tissue from a point source, we model how the downstream optics see the light distribution from the plane where the point source is located. Figure 8c1 illustrates how the light scattering process is modeled using Monte Carlo methods. The surface is treated as the $z = 0$ axial plane. A fluorescent light source is placed at $z = -z_0$ plane. The lateral location of this point source is considered as $x = y = 0$. A photon is then launched from the source, placed at $(0, 0, -z_0)$, in the random direction (u_x, u_y, u_z) . This photon is propagated for s distance (call this a hop). Then a scattering event occurs, and the direction of the photon is changed by a deflection angle, θ , and an azimuth angle, ψ (call this a spin). Then the photon is hopped and spun

repeatedly until it reaches the surface. Once it reaches the surface, the photon's final direction is recorded. If the direction is within the range captured by the numerical aperture of the objective, the photon is then back-propagated until it reaches the original launch plane (i.e., $z = -z_0$), and its location on the launch plane is recorded. When the imaging system is focused on this launch plane, the photon appears to have originated from this location. This process is repeated a sufficiently large number of times to generate the sPSF at the launch plane.

In the simulation, we need to randomly generate the hop distance (s) and spin angles (θ and ψ) according to the scattering properties of the tissue. First, the probability distribution of s is related to the scattering coefficient of the tissue, μ_s . By the definition of μ_s ²⁹, the probability of transmission of a photon without encountering a scattering event after path-length s is given by,

$$P(S > s) = e^{-\mu_s s} \quad (3)$$

Therefore, we can derive the cumulative distribution function (c.d.f.) of s , $F(s)$, and probability density function

(p.d.f) of $s, f(s)$ as follows:

$$F(s) = P(S \leq s) = 1 - e^{-\mu_s s} \quad (4)$$

$$f(s) = \frac{d(F(s))}{dt} = \mu_s e^{-\mu_s s} \quad (5)$$

In Monte Carlo methods, $f(s)$ is sampled using a computer-generated, uniformly distributed random number rnd_1 such that,

$$rnd_1 = F(s) \quad (6)$$

Therefore, we can derive s as a function of rnd_1 as follows:

$$s = -\ln(1 - rnd_1)/\mu_s \quad (7)$$

Second, the deflection angle, θ , is related to the anisotropy, g , of the tissue according to the Henyey–Greenstein scattering function (HG function) that mimics the angular dependence of light scattering by small particles²⁹. According to the HG function, the p.d.f. of $\cos(\theta)$ can be written as follows:

$$f(\cos(\theta)) = \frac{1}{2} \frac{(1 - g^2)}{(1 + g^2 - 2g \cos(\theta))^{(3/2)}} \quad (8)$$

From the same principles of sampling $f(\cos(\theta))$ using a second random number rnd_2 it can be shown that,

$$\cos(\theta) = \frac{1}{2g} \left(1 + g^2 - \left(\frac{1 - g^2}{1 - g + 2grnd_2} \right)^2 \right) \quad (9)$$

Last, the azimuth angle, ψ , can be picked at random, i.e., $\psi = rnd_3$, where rnd_3 is a third computer-generated normally distributed random number. For more details, we refer interested readers to Jacques and Wang²⁹.

EMCCD noise model

Next, we model the noise added to the image, as shown in Eq. (2). We use an EMCCD camera for detection. Let Y_{Shot} be the number of electrons \bar{e} from the signal detected before the EM process. Note that the expected value of Y_{Shot} (i.e., \bar{Y}_{Shot}) is equal to $Y(x, y, z_{focal})$ in Eq. (2). Let Y_{Dark} be the number of electrons generated by the dark current of the camera. Dark current is usually listed in the camera specifications in “ $\bar{e}/pixel/s$ ” and can be used to calculate the expected value of Y_{Dark} (i.e., \bar{Y}_{Dark}) when the exposure time is known. Both Y_{Shot} and Y_{Dark} are Poisson distributed. Let N_{Read} be the read noise of the camera. The standard deviation of the read noise (σ_{Read}) is listed in the camera specifications in \bar{e} . Read noise can be described by a normal distribution with zero mean and (σ_{Read}) standard deviation. Let us denote the EM process by a statistical function $f_{EM}(.)$. Then we can write a formulation for the output signal, Y_{Out}

(equals to $\hat{Y}_t(x, y)$ in Eq. (2)), in \bar{e} as follows:

$$Y_{Out} = f_{EM}(Y_{Shot} + Y_{Dark}) + N_{Read} \quad (10)$$

We can then write the mean and variation equations for the signals from Eq. (10) as follows:

$$\bar{Y}_{Out} = g_{EM}(\bar{Y}_{Shot} + \bar{Y}_{Dark}) \quad (11)$$

$$\sigma_{Out}^2 = g_{EM}^2 F^2 (\sigma_{Shot}^2 + \sigma_{Dark}^2) + \sigma_{Read}^2 \quad (12)$$

Here \bar{Y} denotes the mean of the random variable Y . g_{EM} is the EM gain of the camera, and it is the average gain added by the EM process. The range of g_{EM} is listed in camera specs and the exact value is set during imaging. F quantifies the noise added by the EM process, $f_{EM}(.)$, and is discussed in the next subsection.

Electron multiplication process, $f_{EM}(.)$

The EM process in practice is noisy and g_{EM} is only the average value of the EM gain. This noise is quantified by the excess noise factor, F , defined in Eq. (12). We can derive an expression for F^2 from Eq. (12) as follows:

$$F^2 = \frac{\sigma_{Out}^2 - \sigma_{Read}^2}{g_{EM}^2 (\sigma_{Shot}^2 + \sigma_{Dark}^2)} \quad (13)$$

Here $(\sigma_{Shot}^2 + \sigma_{Dark}^2)$ is the variation of the signal that goes into the EM process, and $(\sigma_{Out}^2 - \sigma_{Read}^2)$ is the variation of the signal that comes out of the EM process (if F is 1 there is no noise added to by the EM process). Robins and Hadwen³⁰ derived an expression for F as follows:

$$F^2 = \frac{1}{g_{EM}} \left(\frac{2g_{EM} + \alpha - 1}{\alpha + 1} \right) = 2(g_{EM} - 1)g_{EM}^{(-N+1)/N} + \frac{1}{g_{EM}} \quad (14)$$

Here, N is the number of EM gain stages. EM process is treated as a Bernoulli process at each EM stage, and α is the probability of an electron multiplication event happening, i.e., the probability of a success event. The value of α is usually small and is in the order of 1–2%³¹. Considering all input electrons to a stage, the gain for each EM stage can be described by a Binomial distribution with a probability mass function,

$$pmf(x_{Ad}; x_{in}, \alpha) = Pr(X_{Ad} = x_{Ad}) = \binom{x_{in}}{x_{Ad}} \alpha^{x_{Ad}} (1 - \alpha)^{(x_{in} - x_{Ad})} \quad (15)$$

Here x_{in} is the number of input electrons to the gain stage, and X_{Ad} is the number of added electrons through the EM gain stage. The output electrons from the gain, therefore, equal to $X_{Ad} + x_{in}$. N such gains stages are cascaded to get the final EM output signal $f_{EM}(.)$.

However, simulating a cascade of $pmf(.)$ functions of the form Eq. (15) to add noise to every pixel in every simulated image is prohibitively slow. Therefore, for each potential input value to $f_{EM}(.)$, we generated a distribution of $f_{EM}(.)$ output values. Note that this is doable as the inputs to $f_{EM}(.)$ are integers that come from Poisson distributed variables (i.e., $Y_{Shot} + Y_{Dark}$). Then during the simulations, we randomly sampled the output distribution corresponding to the input value to get the output of $f_{EM}(.)$ in Eq. (10).

Inverse model

Network architecture

In this work, we use a deep learning network, called DEEP², as the inverse model. Our proposed DEEP² is inspired by the UNet³² architecture with some modifications, as illustrated in Fig. 9. DEEP² consists of 9 convolution blocks, each block containing two convolutional layers followed by batch normalization and a ReLU activation. In the encoder block, the output of each block is sent through a max pooling layer for downsampling. An additional block, which we call scSE, is integrated into the decoder blocks. The importance of this block is discussed in the next section. The feature maps are upsampled in the decoder path before feeding into the next decoder block. The skip connections between the downsample and upsample parts enable the model to learn the fine-grained details in the biological structures more efficiently. We exploit a separate block at the end, which consists of a convolutional layer and a ReLU activation layer, as the final reconstruction step.

UNet with spatial and channel squeeze and excitation

In any deep learning model, the main purpose of the convolution layer is to learn to capture all local spatial information within each channel and to create feature maps jointly encoding the spatial and channel dependencies. Improving the joint encoding of spatial and channel information through convolution requires much effort. Thus, the Concurrent Spatial and Channel “Squeeze and Excitation” (scSE) block is introduced³³. This improves the feature map by suppressing the weak features and enhancing the important features. Despite this, recalibrating non-significant features towards zero with scSE may increase the sparsity in deeper layers as well as reduce the parameter learning³⁴. We denote the performance enhancement of the UNet with scSE over the vanilla UNet in our experiments.

Loss function

Our learning algorithms focus on solving the optimization problem given by structural risk minimization, $\min_f \frac{1}{N} \sum_{i=1}^N L_\theta(I_f \hat{Y}_t, X_0)$, where \hat{Y}_t is the set of DEEP-TFM images given to the Inverse model I_f and the X_0 is the corresponding ground truth (unscattered) image. In the past few years, the attention of the machine learning community has been drawn to the selection of the loss function, which provides a significant theoretical and practical value to the optimization task. Therefore, we have exploited four loss functions with the above-described architecture as the

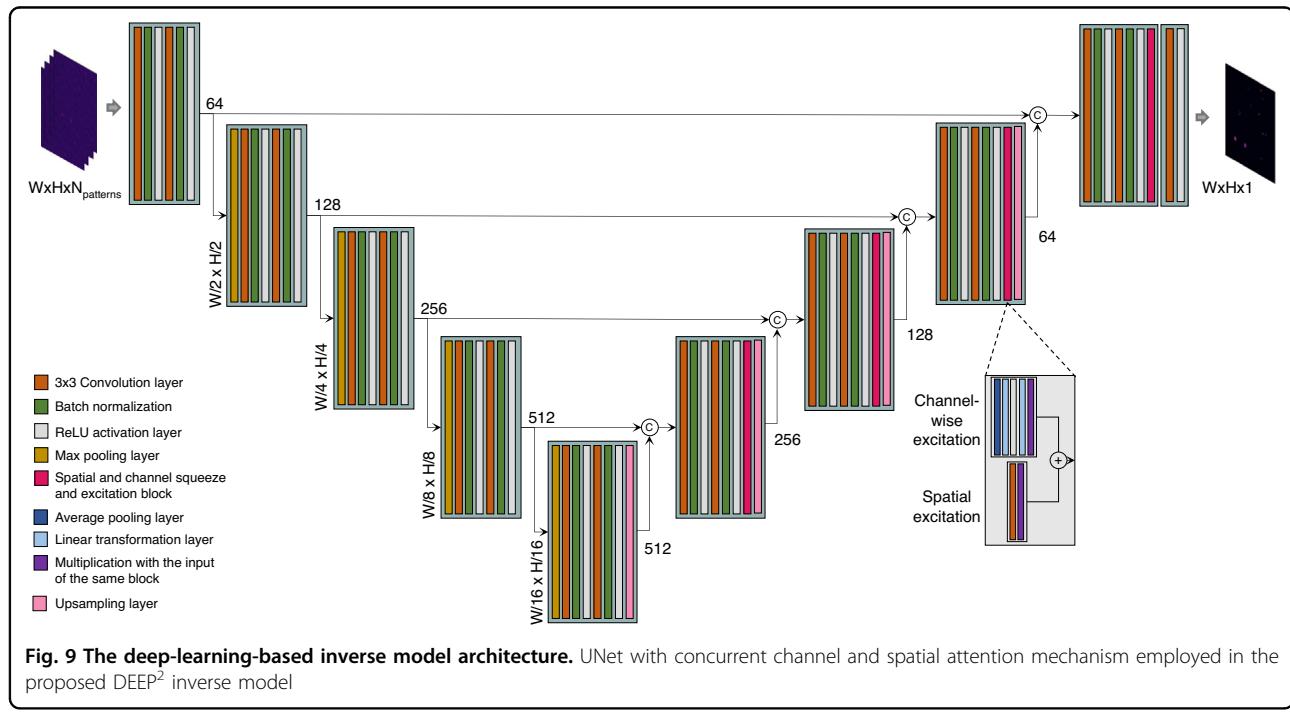


Fig. 9 The deep-learning-based inverse model architecture. UNet with concurrent channel and spatial attention mechanism employed in the proposed DEEP² inverse model

inverse model to obtain the optimal learning algorithm. The least squared error given by Eq. (16) is the most frequently used loss function in regression tasks. Let us assume the output of the inverse model is \hat{X}_0 ,

$$\text{Loss}_{\text{MSE}} = \frac{1}{n} \sum_{t=1}^n (X_0 - \hat{X}_0)^2 \quad (16)$$

This loss function can converge fast as the gradient of the loss decreases when the loss approach 0. However, a drawback of this loss function is the high sensitivity for outliers. Root Mean Squared Logarithmic Error (RMSLE) is more robust to the outliers as it nullifies the effect by considering the log value.

$$\text{Loss}_{\text{RMSLE}} = \sqrt{\frac{1}{n} \sum_{t=1}^n (\log(X_0) - \log(\hat{X}_0))^2} \quad (17)$$

Therefore, the above loss function is also utilized in our experiments. Next, we also explored the Smooth L1 loss, which is also more robust for the outliers compared to the mean squared error (MSE).

$$\text{Loss}_{\text{SmoothL1}} = \begin{cases} \frac{1}{n} \sum_{t=1}^n \frac{1}{2} (X_0 - \hat{X}_0)^2 & \text{if } \frac{1}{n} \sum_{t=1}^n |X_0 - \hat{X}_0| < 1 \\ \frac{1}{n} \sum_{t=1}^n |X_0 - \hat{X}_0| - \frac{1}{2} & \text{otherwise} \end{cases} \quad (18)$$

Smooth L1 loss conditionally chooses either the mean squared loss or the absolute loss, while combining the advantages of both loss functions. Finally, we also explored the usability of KL divergence as a loss function, which can handle distributions more efficiently despite the normality of the approximation error.

$$\text{Loss}_{\text{KLDivergence}} = \frac{1}{n} \sum_{t=1}^n X_0 \cdot \log\left(\frac{X_0}{\hat{X}_0}\right) \quad (19)$$

Training and evaluating the inverse model

The inverse model was trained using the Adam optimizer with learning rate (*lr*), *beta1*, *beta2* and batch size set to, 10^{-3} , 0.9, 0.999 and 10, respectively, for 100 epochs. All the experiments are conducted with the PyTorch³⁵ learning framework. All the experiments are done with the Titan RTX available on the FASRC Cannon cluster at Harvard University.

Evaluation metrics

We mainly used three metrics to evaluate the performance of our model on the validation data (see supplementary): (1) SSIM³⁶, (2) MSE, (3) PSNR. The quantitative evaluation of the proposed model on the validation dataset is presented in Table 1.

$$\text{SSIM} = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (20)$$

Table 1 Quantitative comparison (MSE/PSNR/SSIM) of the proposed model with KL divergence loss

Dataset	SLS	Evaluation metrics		
		MSE (10^{-4})	PSNR	SSIM (10^2)
Mouse cortical vasculature	2 SLS	4.33 ± 2.46	34.11 ± 1.87	92.45 ± 1.97
	4 SLS	6.77 ± 3.01	32.03 ± 1.62	88.79 ± 2.08
	6 SLS	29.07 ± 12.71	25.72 ± 1.72	80.76 ± 3.86
Mouse pyramidal neuron	2 SLS	2.76 ± 1.49	36.18 ± 2.28	97.10 ± 1.29
	6 SLS	15.46 ± 9.51	28.99 ± 2.89	85.81 ± 4.74
Synthetic beads	4 SLS	1.33 ± 0.58	39.16 ± 1.82	78.75 ± 12.95

$$\text{MSE} = \frac{1}{n} \sum (x_i - y_i)^2 \quad (21)$$

$$\text{PSNR} = 10 \log_{10} \left(\frac{R^2}{\text{MSE}} \right) \quad (22)$$

where $\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \sigma_{xy}$ are mean, standard deviation, and covariance of ground truth image x and output of the inverse model y . $C1 = (k_1 \times L)^2$ and $C2 = (k_2 \times L)^2$ are regularization parameters with $L = 1.0$, $k_1 = 0.01$ and $k_2 = 0.03$. n denotes the number of pixels in the image, and R denotes the maximum intensity of the pixels in input images.

Training and validation datasets

Artificial fluorescent beads mixture dataset

We experimentally acquired fluorescent beads images using our DEEP-TFM system to test DEEP². To train DEEP² inverse model under similar conditions, we simulated a synthetic dataset. Artificial beads data, in the size of the experimentally acquired DEEP-TFM instance, were generated to use as the input to the forward model. Here 3D volumes with beads in radius 0.5 μm (1.5 pixels) and 2 μm (6 pixels) were generated along with randomly selected intensities from a normal distribution with mean 1 and standard deviation 0.1. The 0.5 μm radius beads intensities were enhanced by 5 times compared to the 2 μm radius beads. These artificially generated 3D volumes with beads were given as the input to the forward model to generate the simulated DEEP-TFM instances. Since the experimental beads data acquired using DEEP-TFM was obtained at 4 scattering lengths, the same depth was used for the synthetic bead dataset creation. Thus, for 4 scattering lengths, 3455 instances were obtained, which we split into a 4:1 ratio for training and validation. In addition, 128 artificial beads instances were generated for

independent testing. Each instance consists of 32 patterned DEEP-TFM-like images of size 256 pixels by 256 pixels.

Mouse pyramidal neuron dataset

Experimentally obtained images of pyramidal neurons with the dendritic arbor, using a PSTPM, were also used for simulated DEEP-TFM image generation using the forward model. Each neuronal image volume was 607 pixels \times 607 pixels \times 263 pixels with a spatial resolution of 0.25 μm . We used 11 such neuronal image volumes to create simulated DEEP-TFM neuronal instances. To enhance the intensity of the fine structures such as dendritic spines, the intensity was thresholded at 20, by assigning 20 for all the pixels with intensity over 20. The successive images along the z -plane of the image volume contained similar structures. Thus, to avoid repeating similar training instances, the sub-image volumes that match the third dimension of the expSF, were extracted from the neuronal volume, to use as the input to the forward model, with 5 image planes apart from each other. Furthermore, to avoid the image volumes with no significant structures, the sub-image volumes with the mean exceeding the mean of the whole neuronal image volume were considered as the input of the forward model. The 3895 simulated DEEP-TFM neuronal instances were generated for 2 (~100 μm) and 6 (~300 μm) scattering lengths. Each instance consists of 326 pixels by 326 pixels ground truth ($\hat{X}_0(x, y)$) and 32 patterned images ($t = 32$) of size 326 pixels by 326 pixels ($\hat{Y}_t(x, y)$) as the input of the inverse model.

Mouse cortical vasculature dataset

In addition to the experimentally acquired beads, we also imaged mouse cortical vasculature at increasing depths using the DEEP-TFM. Hence, to train the network, we used an equivalent vascular structural image volume acquired using a three-photon Fluorescent microscope³⁷. The vascular image stack consists of 800 z -planes, each with an increment of 1.5 μm and each sized 512 pixels \times 512 pixels with a spatial resolution of 0.6 μm corresponding to a field of view of ~300 μm .

Artificial vasculature dataset

In addition, artificial blood vessel volumes were employed^{38,39} to demonstrate the applicability of DEEP² on vasculature structures furthermore. The synthetic volumes originally consisted of blood vessels with radii ranging from 1 to 7 pixels. The finest blood vessels with low intensities were removed by applying a threshold at 190. The blood vessel volumes were originally sized 325 pixels by 304 pixels by 600 pixels, and 20 such volumes were downloaded. Each volume was rescaled by factor 1.09, normalized by dividing by the maximum intensity of each volume and extracted volumes of 326 pixels by 326

pixels by 200 pixels volumes to redirect as the input to the Forward model. Similarly, we generated datasets at 2 and 6 scattering lengths for blood vessels.

Mouse surgery and imaging

Experiments were carried out under protocols approved by MIT's Animal Care and Use Committee and conformed to NIH guidelines. All data in this study were collected from adult (>8 weeks old) mice of either sex. The wild-type mice were acquired from the Jackson laboratory (#000664). Mice were initially anesthetized with 4% isoflurane in oxygen and maintained on 1.5–2% isoflurane throughout the surgery. Buprenorphine (1 mg kg^{-1} , subcutaneous) and/or meloxicam (1 mg kg^{-1} , subcutaneous) was administered preoperatively and every 24 h for 3 days to reduce inflammation. Ophthalmic ointment was used to protect the animal's eyes during the surgery. Body temperature was maintained at 37.5 °C with a heating pad. The scalp overlying the dorsal skull was sanitized and removed. The periosteum was removed with a scalpel, and a craniotomy (5 mm) was made over the primary visual cortex (V1, 4.2 mm posterior, 3.0 mm lateral to Bregma) on either the left or right hemisphere, leaving the dura intact. Last, a custom-designed stainless steel head plate (eMachineShop.com) was affixed to the skull using dental acrylic. Experiments were performed at least 5 days after head plate implantation to allow animals to recover. For labeling blood vessels with a fluorescent dye, a rhodamine + dextran dye (70 kDa; D1841, Thermo Fisher Scientific) mixed with saline solution at 5% (wv⁻¹) concentration was applied retroorbitally with 100 μl volume⁴⁰. During the retroorbital injection, the animal was anesthetized with 2% isoflurane in oxygen. During both 2-photon and 3-photon imaging, the animal was anesthetized with ketamine + xylazine mixture with 0.1 ml volume, and this mixture was applied as needed after checking the reflexes. The imaging sessions lasted for a maximum of 2 h.

Acknowledgements

This work was supported by the Center for Advanced Imaging at Harvard University (D.N.W., N.W., M.A), 5-P41EB015871-32 (D.N.W., P.T.C.S.), R21 MH130067 (P.T.C.S., D.N.W.), R21 NS105070 (P.T.C.S.), and R00EB027706 (M.Y.). D.N.W. is also supported by the John Harvard Distinguished Science Fellowship Program within the FAS Division of Science of Harvard University. We also thank Dr Josiah R. Boivin for providing point-scanning two-photon datasets of pyramidal neurons. We also thank Dr Ranga Rodrigo for his feedback during the early stages of the project.

Author details

¹Center for Advanced Imaging, Faculty of Arts and Sciences, Harvard University, Cambridge, MA, USA. ²Department of Electronic and Telecommunication Engineering, University of Moratuwa, Moratuwa, Sri Lanka.

³Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, USA. ⁴Laser Biomedical Research Center, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, USA. ⁵Department of Biological Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, USA. ⁶Picower Institute for Learning and Memory,

Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, USA. ⁷Department of Neuroscience, Cleveland Clinic Lerner Research Institute, Cleveland, OH 44195, USA

Author contributions

D.N.W. conceived the idea. N.W. designed and implemented the neural network algorithms. D.N.W. formulated and implemented the forward model. N.W. and M.A. simulated the training data and trained the neural network-based inverse models. M.A. designed and performed all ablation studies. C.Z. and M.Y. performed the DEEP-TFM imaging experiments under P.T.C.S.'s guidance. M.Y. performed animal surgeries and point-scanning three-photon experiments. N.W., D.N.W., and M.A. wrote the manuscript. P.T.C.S. and M.Y. contributed with valuable suggestions during discussions and revised the manuscript. D.N.W. supervised the project.

Data availability

The dataset and excitation patterns used in the forward model of this study are available on <https://zenodo.org/record/8161051>. The datasets simulated by the forward model are available from the corresponding author upon request.

Code availability

The codes used in this study are available through a public GitHub repository: <https://github.com/wadduwagelab/DEEP-squared>.

Conflict of interest

The authors declare no competing interests.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41377-023-01248-6>.

Received: 1 March 2023 Revised: 21 July 2023 Accepted: 29 July 2023

Published online: 13 September 2023

References

1. Rocheleau, J. V. & Piston, D. W. Two-photon excitation microscopy for the study of living cells and tissues. *Curr. Protoc. Cell Biol.* **20**, 4.11.1–4.11.15 (2003).
2. Yildirim, M. et al. Functional imaging of visual cortical layers and subplate in awake mice with optimized three-photon microscopy. *Nat. Commun.* **10**, 77 (2019).
3. Yildirim, M. et al. Label-free three-photon imaging of intact human cerebral organoids for tracking early events in brain development and deficits in Rett syndrome. *eLife* **11**, e78079 (2022).
4. Benninger, R. K. P. & Piston, D. W. Two-photon excitation microscopy for the study of living cells and tissues. *Curr. Protoc. Cell Biol.* **59**, 4.11.1–4.11.24 (2013).
5. Oron, D., Tal, E. & Silberberg, Y. Scanningless depth-resolved microscopy. *Opt. Express* **13**, 1468–1476 (2005).
6. Zhu, G. H. et al. Simultaneous spatial and temporal focusing of femtosecond pulses. *Opt. Express* **13**, 2153–2159 (2005).
7. Vaziri, A. & Shank, C. V. Ultrafast widefield optical sectioning microscopy by multifocal temporal focusing. *Opt. Express* **18**, 19645–19655 (2010).
8. Rowlands, C. J. et al. Wide-field three-photon excitation in biological samples. *Light Sci. Appl.* **6**, e16255 (2017).
9. Escobet-Montalbán, A. et al. Wide-field multiphoton imaging through scattering media without correction. *Sci. Adv.* **4**, eaau1338 (2018).
10. Zheng, C. et al. De-scattering with excitation patterning enables rapid wide-field imaging through scattering media. *Sci. Adv.* **7**, eaay5496 (2021).
11. Mertz, J. Optical sectioning microscopy with planar or structured illumination. *Nat. Methods* **8**, 811–819 (2011).
12. Belthangady, C. & Royer, L. A. Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction. *Nat. Methods* **16**, 1215–1225 (2019).
13. Weigert, M. et al. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat. Methods* **15**, 1090–1097 (2018).
14. Jin, K. H. et al. Deep convolutional neural network for inverse problems in imaging. *IEEE Trans. Image Process.* **26**, 4509–4522 (2017).
15. Yannai, K. et al. Deep learning for fast spatially varying deconvolution. *Optica* **9**, 96–99 (2022).
16. Krizhevsky, A., Sutskever, I. & Hinton, G. E. Imagenet classification with deep convolutional neural networks. In: *Proceedings of the 25th International Conference on Neural Information Processing Systems*, 1097–1105 (Curran Associates Inc., Lake Tahoe, USA, 2012).
17. Ziletti, A. et al. Insightful classification of crystal structures using deep learning. *Nat. Commun.* **9**, 2775 (2018).
18. Wei, Z., Liu, D. & Chen, X. D. Dominant-current deep learning scheme for electrical impedance tomography. *IEEE Trans. Biomed. Eng.* **66**, 2546–2555 (2019).
19. Liu, S. Q. et al. Early diagnosis of Alzheimer's disease with deep learning. In: *Proceedings of 2014 IEEE 11th International Symposium on Biomedical Imaging*, 1015–1018 (IEEE, Beijing, China, 2014).
20. Wang, S. S. et al. Accelerating magnetic resonance imaging via deep learning. In: *Proceedings of 2016 IEEE 13th International Symposium on Biomedical Imaging*, 514–517 (IEEE, Prague, Czech Republic, 2016).
21. Girshick, R., Donahue, J., Darrell, T. & Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: *Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition*, 580–587 (IEEE, Columbus, OH, USA, 2014).
22. Nielsen, A. A. K. & Voigt, C. A. Deep learning to predict the lab-of-origin of engineered DNA. *Nat. Commun.* **9**, 3135 (2018).
23. Eraslan, G. et al. Single-cell RNA-seq denoising using a deep count auto-encoder. *Nat. Commun.* **10**, 390 (2019).
24. Eulenberg, P. et al. Reconstructing cell cycle and disease progression using deep learning. *Nat. Commun.* **8**, 463 (2017).
25. Wei, Z. & Chen, X. D. Physics-inspired convolutional neural network for solving full-wave inverse scattering problems. *IEEE Trans. Antennas Propag.* **67**, 6138–6148 (2019).
26. Zhu, B. et al. Image reconstruction by domain-transform manifold learning. *Nature* **555**, 487–492 (2018).
27. Wei, Z. et al. 3D deep learning enables fast imaging of spines through scattering media by temporal focusing microscopy. Preprint at <https://arxiv.org/abs/2001.00520> (2019).
28. Ma, Y. et al. Wide-field optical mapping of neural activity and brain haemodynamics: considerations and novel approaches. *Philos. Trans. R. Soc. B Biol. Sci.* **371**, 20150360 (2016).
29. Jacques, S. L. & Wang, L. Monte Carlo modeling of light transport in tissues. In: *Optical-Thermal Response of Laser-irradiated Tissue* (eds. Welch, A. J. & Gemert, M. J. C.) 73–100 (Springer, Boston, 1995).
30. Robbins, M. S. & Hadwen, B. J. The noise performance of electron multiplying charge-coupled devices. *IEEE Trans. Electron Devices* **50**, 1227–1232 (2003).
31. EMCCD Tutorial. <https://www.nuvucameras.com/emccd-tutorial/>. Accessed February 24, 2012.
32. Ronneberger, O., Fischer, P. & Brox, T. U-net: convolutional networks for biomedical image segmentation. In: *Proceedings of the 18th International Conference on Medical Image Computing and Computer-assisted Intervention*, 234–241 (Springer, Munich, Germany, 2015).
33. Roy, A. G., Navab, N. & Wachinger, C. Concurrent spatial and channel 'squeeze & excitation' in fully convolutional networks. In: *Proceedings of the 21st International Conference on Medical Image Computing and Computer-Assisted Intervention*, 421–429 (Springer, Granada, Spain, 2018).
34. Uhrig, J. et al. Sparsity invariant CNNs. In: *Proceedings of 2017 International Conference on 3D Vision*, 11–20 (IEE, Qingdao, China, 2017).
35. Paszke, A. et al. Automatic differentiation in pytorch. In: *Neural Information Processing Systems 2017 Autodiff Workshop* (Long Beach, CA, USA, 2017).
36. Wang, Z. et al. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**, 600–612 (2004).
37. Yildirim, M. et al. Quantitative third-harmonic generation imaging of mouse visual cortex areas reveals correlations between functional maps and structural substrates. *Biomed. Opt. Express* **11**, 5650–5673 (2020).
38. Schneider, M. et al. Tissue metabolism driven arterial tree generation. *Med. Image Anal.* **16**, 1397–1414 (2012).
39. Todorov, M. I. et al. Machine learning analysis of whole mouse brain vasculature. *Nat. Methods* **17**, 442–449 (2020).
40. Yardeni, Y., Eckhaus, M., Morris, H.D., Huizing, M. & Hoogstraten-Miller, S. Retro-orbital injections in mice. *Lab. Anim. (NY)* **40**, 155–160 (2011).