Reinforcing neuron extraction and spike inference in calcium imaging using deep self-supervised denoising

Xinyang Li[1,2,3,4,8], Guoxun Zhang[1,3,4,5,8], Jiamin Wu [1,3,4,5,6], Yuanlong Zhang[1,3,4,5], Zhifeng Zhao[1,3,4,5], Xing Lin [1,3,4,6], Hui Qiao [1,3,4,5,6], Hao Xie[1,3,4,5,6], Haoqian Wang [2,3 ✉], Lu Fang [3,7 ✉] and Qionghai Dai [1,3,4,5,6 ✉]

Calcium imaging has transformed neuroscience research by providing a methodology for monitoring the activity of neural circuits with single-cell resolution. However, calcium imaging is inherently susceptible to detection noise, especially when imaging with high frame rate or under low excitation dosage. Here we developed DeepCAD, a self-supervised deep-learning method for spatiotemporal enhancement of calcium imaging data that does not require any high signal-to-noise ratio (SNR) observations. DeepCAD suppresses detection noise and improves the SNR more than tenfold, which reinforces the accuracy of neuron extraction and spike inference and facilitates the functional analysis of neural circuits.

Calcium imaging enables parallel recordings of large neuronal ensembles in living animals[1–4] and offers the possibility of deciphering information propagation, integration, and computation in neural circuits[5]. For accurate neuron extraction and spike inference for further analysis, high-SNR calcium imaging is desired. However, owing to the paucity of fluorescence photons caused by low-peak accumulations and fast dynamics of in vivo calcium transients[6,7], calcium imaging is vulnerable to noise contamination (that is, photon shot noise and electronic noise), especially in cases where high temporal resolution is important for analyzing neural activity[8].

The most direct way to capture sufficient fluorescence photons for high-SNR calcium imaging is to use high excitation dosage, but concurrent photobleaching, phototoxicity[9,10], and tissue heating[11,12] are detrimental for sample health and photosensitive biological processes. More effective strategies include using brighter calcium indicators[7,13] and more advanced photoelectric detection techniques[14], but their performances are still largely restricted in photon-limited conditions, such as dendritic imaging and deep-tissue imaging. Apart from these physical or biological approaches, data-driven methods can offer an alternative solution to recover faithful signals from degraded recordings and reduce the photon budget of calcium imaging. As an advanced signal-processing technique, deep learning has been adopted by microscopists and has achieved good performance in fluorescence imaging[15–18]. However, calcium transients constitute highly dynamic, nonrepetitive activities, and a firing pattern cannot be captured twice. Previous schemes for obtaining training ground truth (that is, images without noise contamination or high-SNR images with the same underlying scene as the low-SNR image) by extending the integration time or averaging multiple noisy frames are no longer feasible, posing an obstacle for conventional supervised learning methods.

Here we present DeepCAD, a self-supervised learning method for denoising calcium imaging data that achieves over tenfold SNR improvement without requiring any high-SNR observations for training. DeepCAD is based on the insight that a deep neural network can converge to a mean estimator even though the target image used for training is another corrupted sampling of the same scene, because the optimal network parameters are approximate to those trained with ground-truth images[19]. In the context of calcium imaging, we explored the temporal redundancy of video-rate (~30 Hz) imaging and found that any two consecutive frames can be regarded as two independent samplings of the same underlying firing pattern. Thus, image pairs composed of two consecutive frames can be used for the training of denoising models. Furthermore, the input and output data of DeepCAD are designed to be three-dimensional (3D) volumes rather than two-dimensional (2D) frames to fully exploit spatiotemporal information in the time-lapse stack. We quantitatively evaluated our method on both simulated and experimental data. We show that such a 3D self-supervised method is effective for calcium imaging denoising, and that even calcium fluctuations induced by a single action potential (AP) can be restored from severely corrupted images. Finally, we released a Fiji-based plugin to make our method easy to access and convenient to use.

Results

Self-supervised principle and performance validation. The general principle of DeepCAD is schematized in Fig. 1a. The underlying mechanism of our self-supervised strategy is that using noisy target images for training produces noisy gradients, but the mean gradient on the entire training set is approximately equal to the true gradient. Therefore, the training result will not be affected. In our network architecture, 3D U-Net[20] utilizes the spatiotemporal

[1]Department of Automation, Tsinghua University, Beijing, China. [2]Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen, China. [3]Institute for Brain and Cognitive Sciences, Tsinghua University, Beijing, China. [4]Beijing Key Laboratory of Multi-dimension & Multi-scale Computational Photography (MMCP), Tsinghua University, Beijing, China. [5]IDG/McGovern Institute for Brain Research, Tsinghua University, Beijing, China. [6]Beijing Innovation Center for Future Chips, Tsinghua University, Beijing, China. [7]Department of Electronic Engineering, Tsinghua University, Beijing, China. [8]These authors contributed equally to this work: Xinyang Li, Guoxun Zhang. ✉e-mail: wanghaoqian@tsinghua.edu.cn; fanglu@tsinghua.edu.cn; qhdai@mail.tsinghua.edu.cn
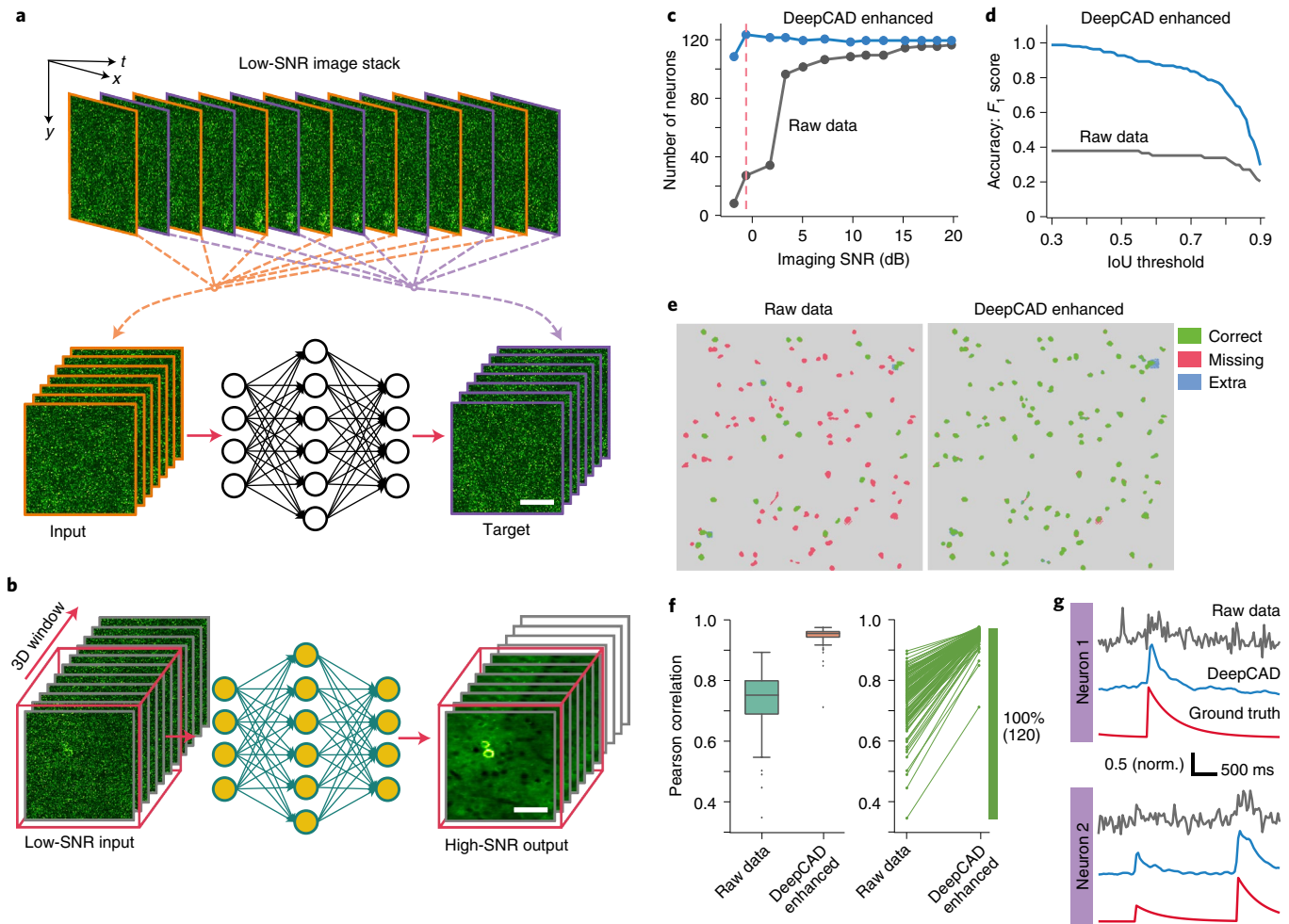
Fig. 1 | General principle and validation of DeepCAD. a, Self-supervised training strategy of DeepCAD. Consecutive frames in the original low-SNR stack are divided into two sub-stacks, which are then used as the input volume and corresponding target volume to train the deep neural network (3D U-Net[20]). After training, a denoising model can be established and memorized in network parameters. Scale bar, 50 μm. **b**, Deployment of the DeepCAD model. For subsequent acquisitions, a 3D (x–y–t) window traverses the entire stack, and 3D tiles are sequentially fed into the pretrained model. Denoised recordings will be obtained after processing by the model. Scale bar, 50 μm. **c**, The number of neurons extracted at different imaging SNRs before and after the enhancement of DeepCAD. n = 120 active neurons were simulated in the field of view (FOV). **d**, Accuracy of neuron segmentation quantified with F_1 score at different IoU thresholds (imaging SNR = −0.7 dB, indicated by the red dashed line in **c**). **e**, Spatial profiles of extracted neurons (imaging SNR = −0.7 dB). Correctly segmented regions (true positive) are colored green. Missing (false negative) and extra regions (false positive) are colored red and blue, respectively. Neuron extraction was implemented with the constrained non-negative matrix factorization (CNMF) algorithm[24]. **f**, Left, Tukey box-and-whisker plot showing the distribution of Pearson correlation coefficients with ground-truth traces before and after denoising (n = 120 independent traces). Right, increases of the trace correlation. Each line represents 1 of 120 calcium traces. **g**, Restoration of calcium transients indiscernible from noise (gray) by DeepCAD (blue). Traces without noise contamination (red) serve as ground truth for comparison.

correlations of calcium imaging in multiple frames using 3D convolutional layers (Extended Data Fig. 1), which endows DeepCAD with a better denoising capability than 2D architectures or other methods[21–23] (Supplementary Figs. 1 and 2). Benefiting from the self-supervised strategy, a single low-SNR stack composed of about 3,500 frames is sufficient as a training set. To generate the training set, we split two sub-stacks consisting of interlaced frames from the original low-SNR stack and extracted 3D tiles from these sub-stacks for training (Extended Data Fig. 2). Adjacent image frames contain approximately identical calcium transients if the original stack was imaged at ~30 Hz, which is widely accessible for most commercial or customized microscopes. After training, interpretable features are learned (Extended Data Fig. 3) and the model can be applied to subsequent acquisitions without additional training (Fig. 1b). Although we trained the DeepCAD network on a specific

combination of spatial and temporal resolutions, we found that it had acceptable performance on various spatial and temporal resolutions (Supplementary Figs. 3–5), illustrating the scalability and generalization for various applications of DeepCAD.

To quantitatively evaluate the performance of DeepCAD, we first validated it on simulated calcium imaging data of different SNRs and synchronous simulated noise-free recordings as ground truth for comparison (Supplementary Figs. 6 and 7 and Supplementary Notes 1 and 2). We used the constrained non-negative matrix factorization (CNMF) algorithm[24] for neuron extraction. After denoising with DeepCAD, more active neurons can be detected, especially when SNR is low (Fig. 1c). We also quantified the accuracy of neuron extraction using F_1 scores and observed substantial improvements across a wide range of intersection-over-union (IoU) thresholds (Fig. 1d,e). For a typical IoU threshold of 0.7, the segmentation accuracy was improved
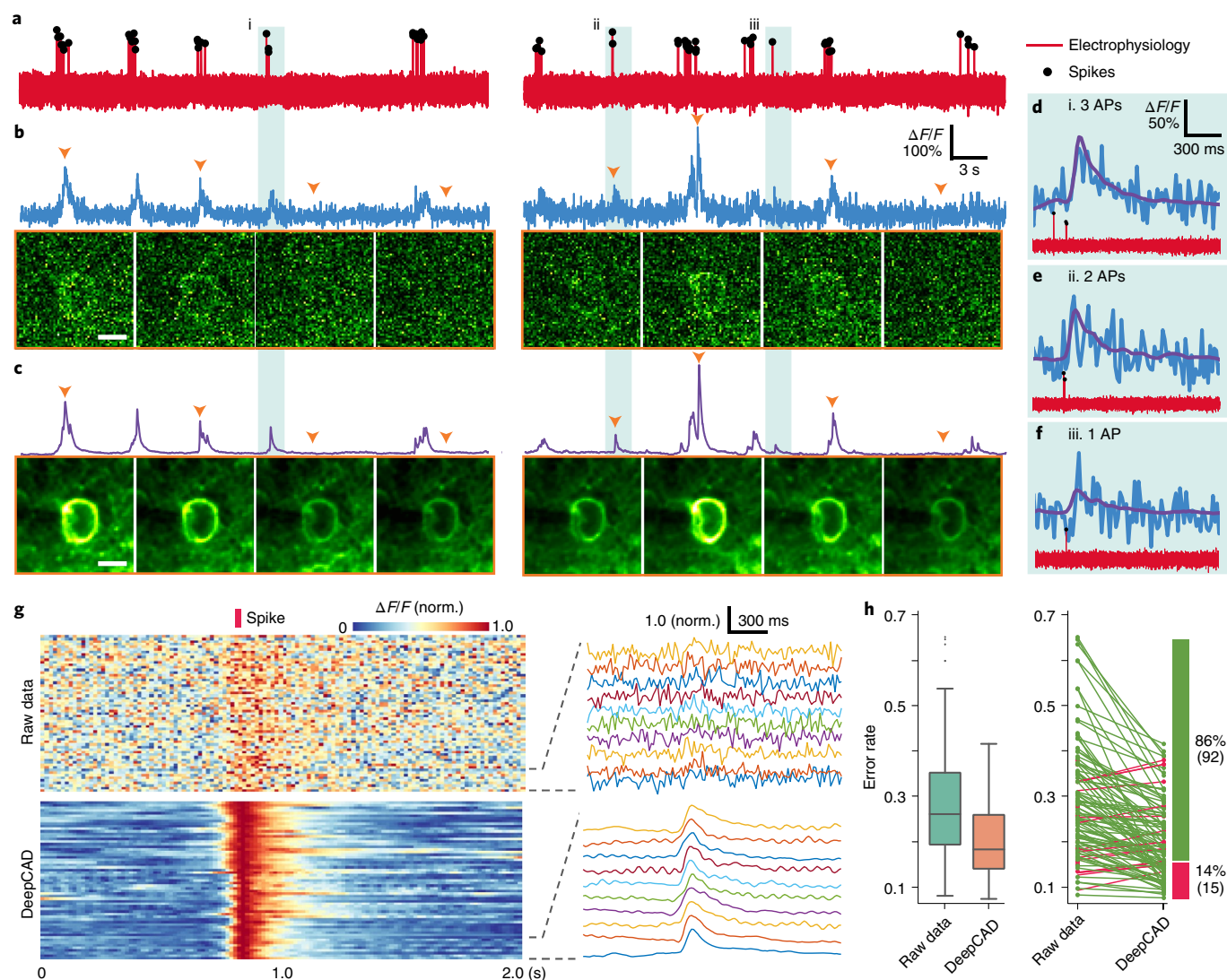
Fig. 2 | Spatiotemporal enhancement with DeepCAD. a, Single-neuron electrophysiology. Detected spikes are marked with black dots. **b**, Two-photon calcium imaging data of the same neuron synchronized with cell-attached electrophysiology. Representative frames indicated with orange triangles are presented below the trace. Scale bar, 10 μm. **c**, Fluorescence traces and representative frames after the enhancement with DeepCAD. Scale bar, 10 μm. **d–f**, Examples of calcium transients evoked by three APs (**d**), two APs (**e**), and one AP (**f**) before (blue) and after (purple) noise removal. **g**, Calcium fluctuations evoked by 61 isolated action potentials. All spikes were normalized and temporally aligned with the red bar. Zoom-in traces are shown in the right panel. **h**, Left, Tukey box-and-whisker plot showing the distribution of error rates of spike inference (lower is better) for calcium traces extracted from enhanced data compared with those extracted from raw noisy data ($n = 107$). Real spike timings were revealed by simultaneous cell-attached recordings. Right, decreases of the error rate. Each line represents one recording. Green, decreased error rates; red, increased error rates.

by 2.4-fold (0.84 for DeepCAD compared with 0.35 for the raw data). Benefiting from the improved data quality, calcium traces extracted from the denoised data possess higher fidelity. To investigate the temporal enhancement of DeepCAD, we extracted calcium traces of all neurons in our simulated dataset from both the raw noisy data and the enhanced counterpart. The Pearson correlation with the ground-truth traces was substantially improved after denoising (Fig. 1f). Even small calcium transients can be restored from the original noisy data (Fig. 1g and Supplementary Fig. 8). These findings suggest that the spatiotemporal enhancement of DeepCAD can improve the accuracy of neuronal localization and trace extraction and facilitate the analysis of neural circuits.

Spatiotemporal enhancement of single-neuron recording with DeepCAD. To verify the effectiveness and reliability of DeepCAD

on experimentally obtained data, we then demonstrated its performance on previously released two-photon calcium imaging data[7]. In this dataset, simultaneous cell-attached electrophysiological recordings (Fig. 2a) are synchronized with two-photon imaging and serve as reference for calcium transients and as ground truth for spike inference. Contaminated by detection noise, both the spatial footprint and temporal traces of neurons were severely corrupted in the original data (Fig. 2b). After we applied DeepCAD to enhance these data, the annular cytoplasms became recognizable and calcium traces were liberated from noise (Fig. 2c and Supplementary Video 1). Even barely perceptible calcium transients evoked by one AP, two APs, or three APs were distinguished and maintained their original dynamics (Fig. 2d–g), which otherwise would be overwhelmed by noise. For further comparison, we extracted single-pixel fluorescence from cytoplasmic pixels and found that calcium transients

can be unveiled at a single-pixel scale (Supplementary Fig. 9). Moreover, we performed spike inference on traces extracted from the original data as well as the corresponding denoised data. Due

to the improvement of SNR, the error rate of spike inference was decreased (Fig. 2h, Extended Data Figs. 4 and 5 and Supplementary Fig. 10) without aggravating the timing jitters of inferred spikes or
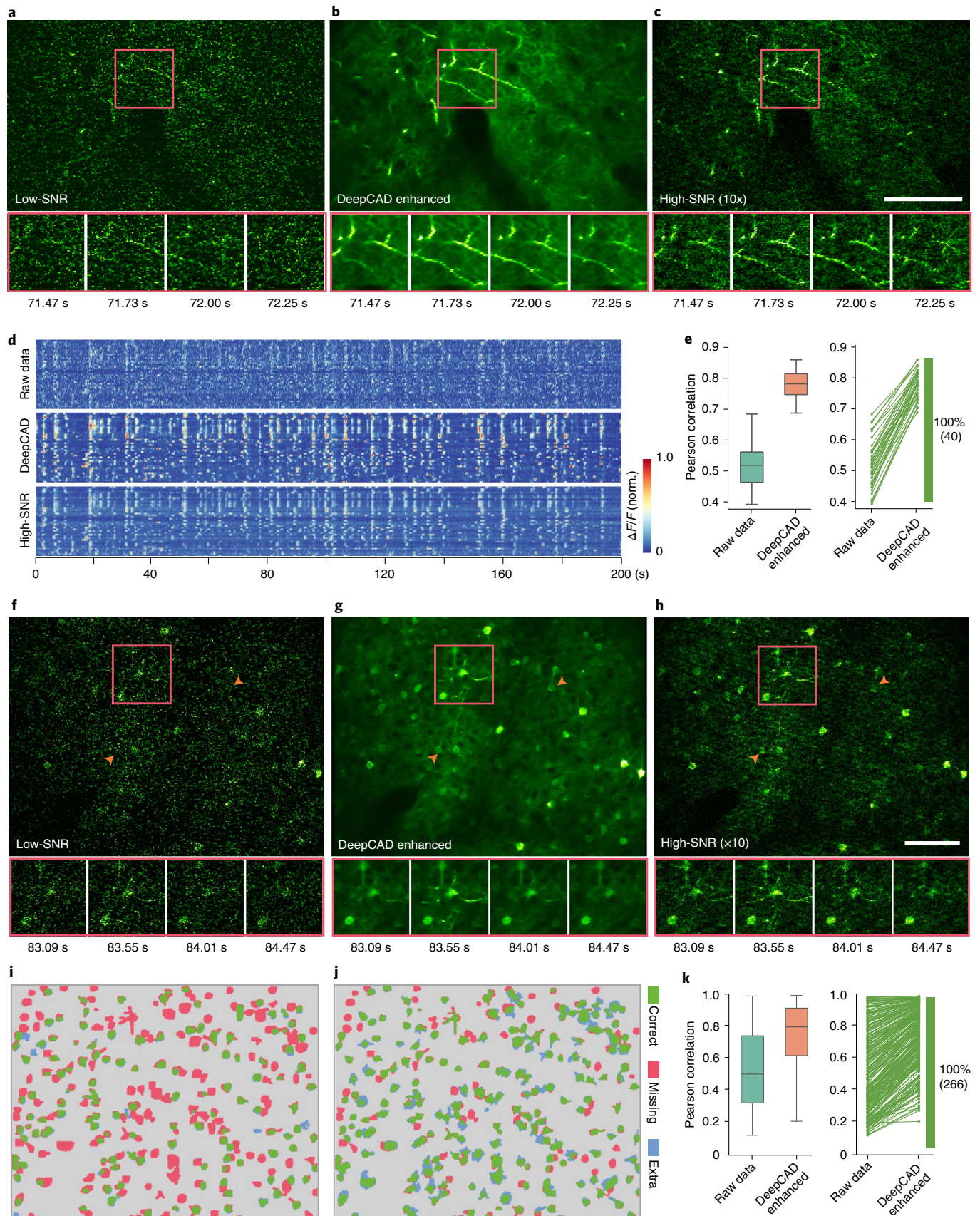
Fig. 3 | DeepCAD denoises calcium imaging data of large neuronal populations. a, Spontaneous neurite activity in layer 1 of the mouse cortex captured by the low-SNR detection path. **b**, Images restored from the low-SNR recording using DeepCAD. **c**, The synchronized recording acquired by the high-SNR detection path (tenfold imaging SNR). Magnified views of the boxed regions show calcium transients in a time window of ~0.8 s. Scale bar, 100 μm. **d**, Fluorescence traces extracted from 40 dendritic pixels. Top, low-SNR recording. Middle, DeepCAD enhanced recording. Bottom, high-SNR recording. **e**, Left, Tukey box-and-whisker plot showing the Pearson correlations of dendritic calcium traces before and after denoising ($n = 40$). High-SNR traces were used as the reference for correlation calculation. Right, increases of the trace correlation. Each line represents 1 of 40 dendritic pixels. Green, increased correlations; red, decreased correlations. **f**, Low-SNR recording of somatic signals in cortical layer 2/3. **g**, DeepCAD enhanced data. **h**, Synchronized high-SNR recording (tenfold imaging SNR). Orange arrows point to two neurons. Magnified views of the boxed regions show calcium transients in a ~1.4-second time window. Scale bar, 100 μm. **i**, Neurons extracted from the original low-SNR recording ($N = 150$). **j**, Neurons extracted from the DeepCAD enhanced data ($n = 229$). Manual annotations served as the ground truth. Correctly segmented regions (true positive) are colored green. Missing (false negative) and extra regions (false positive) are colored red and blue, respectively. **k**, Left, Tukey box-and-whisker plot showing the Pearson correlations of somatic calcium traces before and after denoising ($n = 266$). High-SNR traces were used as the reference for correlation calculation, and manually annotated masks were used for neuron segmentation. Right, increases of the trace correlation. Each line represents 1 of 266 somatic traces. Green, increased correlations; red, decreased correlations.

deteriorating the spike-rate estimation (Extended Data Fig. 6 and Supplementary Fig. 11). Eighty-six percent of 107 independent calcium traces had lower error rates.

DeepCAD facilitates analysis of large neuronal populations. Next, we employed DeepCAD for noise removal in calcium imaging data of large neuronal populations in awake mice. To obtain high-SNR recordings for validation of our method, we designed and built a two-photon imaging system with the capability of simultaneous low-SNR and high-SNR recording (Extended Data Fig. 7). The high-SNR detection path was strictly synchronized with the low-SNR detection path but with about tenfold higher imaging SNR (Extended Data Fig. 8), which can be used as the reference for our denoising results. We first imaged spontaneous neurite activity in cortical layer 1 of a transgenic mouse expressing GCaMP6f and found that calcium fluctuations indiscernible in the original low-SNR recordings can be effectively recovered by DeepCAD (Fig. 3a–c and Supplementary Video 2). The SNR was improved by more than tenfold with DeepCAD and surpassed the corresponding high-SNR reference. Fluorescence traces of dendritic pixels can be accurately resolved and keep high consistency with the high-SNR reference (Fig. 3d,e and Supplementary Fig. 12). We also applied DeepCAD to enhance calcium imaging data of somatic signals. After denoising, neuronal distribution and circuit state can be extracted from a single frame (Fig. 3f–h and Supplementary Video 3). Using CNMF as the source extraction method, we could extract 52.6% (229 compared with 150) more active neurons (Fig. 3i,j and Extended Data Fig. 9). Pearson correlation of somatic signals was improved from 0.50 to 0.79 (Fig. 3k, median value), and the trace peak SNR was also improved by more than twofold (Supplementary Fig. 13), indicating that the functional analysis of large neuronal populations can be strengthened owing to improved SNR. Apart from our own imaging system, we also extended DeepCAD to data obtained with other two-photon imaging systems. DeepCAD has good generalization and scalability and can achieve comparable performance on several different imaging systems, regardless of objectives and detectors (Extended Data Fig. 10 and Supplementary Video 4).

Discussion

In summary, DeepCAD is a method based on deep self-supervised learning for spatiotemporal enhancement of calcium imaging data. DeepCAD overcomes the reliance on training ground truth by directly training a denoising deep neural network with the original low-SNR time-lapse calcium imaging data. We have demonstrated its denoising performance on single-neuron recordings and neuronal-population recordings. To maximize its accessibility, we released a Fiji plugin (Supplementary Fig. 14 and Supplementary Note 3) and a pretrained DeepCAD model for denoising two-photon imaging data of neuronal populations. Our method can be effi-

ciently configured on a common desktop and generalized to other imaging systems, making it a compelling tool for calcium imaging denoising. Although we validated DeepCAD only on two-photon calcium imaging data, we believe it can be extended to other imaging modalities, such as wide-field microscopy, confocal microscopy, and light-sheet microscopy, or other functional imaging applications, such as cell migration observation and voltage imaging. We anticipate that this method could serve as a general processing step for time-lapse imaging data obtained in photon-limited conditions, and could promote long-term and high-fidelity recordings of biological dynamics.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41592-021-01225-0.

References

1. Grienberger, C. & Konnerth, A. Imaging calcium in neurons. *Neuron* **73**, 862–885 (2012).
2. Lu, R. et al. Video-rate volumetric functional imaging of the brain at synaptic resolution. *Nat. Neurosci.* **20**, 620–628 (2017).
3. Weisenburger, S. et al. Volumetric Ca²⁺ imaging in the mouse brain using hybrid multiplexed sculpted light microscopy. *Cell* **177**, 1050–1066 e1014 (2019).
4. Chow, D. M. et al. Deep three-photon imaging of the brain in intact adult zebrafish. *Nat. Methods* **17**, 605–608 (2020).
5. Calarco, J. A. & Samuel, A. D. Imaging whole nervous systems: insights into behavior from worms to fish. *Nat. Methods* **16**, 14–15 (2019).
6. Sabatini, B. L., Oertner, T. G. & Svoboda, K. The life cycle of Ca²⁺ ions in dendritic spines. *Neuron* **33**, 439–452 (2002).
7. Chen, T. W. et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).
8. Ji, N., Freeman, J. & Smith, S. L. Technologies for imaging neural activity in large volumes. *Nat. Neurosci.* **19**, 1154–1164 (2016).
9. Svoboda, K. & Yasuda, R. Principles of two-photon excitation microscopy and its applications to neuroscience. *Neuron* **50**, 823–839 (2006).
10. Skylaki, S., Hilsenbeck, O. & Schroeder, T. Challenges in long-term imaging and quantification of single-cell dynamics. *Nat. Biotechnol.* **34**, 1137–1144 (2016).
11. Podgorski, K. & Ranganathan, G. Brain heating induced by near-infrared lasers during multiphoton microscopy. *J. Neurophysiol.* **116**, 1012–1023 (2016).
12. Wang, T. et al. Quantitative analysis of 1300-nm three-photon calcium imaging in the mouse brain. *eLife* **9**, e53205 (2020).
13. Dana, H. et al. High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nat. Methods* **16**, 649–657 (2019).

14. Samantaray, N., Ruo-Berchera, I., Meda, A. & Genovese, M. Realization of the first sub-shot-noise wide field microscope. *Light Sci. Appl.* **6**, e17005 (2017).

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

16. Belthangady, C. & Royer, L. A. Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction. *Nat. Methods* **16**, 1215–1225 (2019).

17. Wang, H. et al. Deep learning enables cross-modality super-resolution in fluorescence microscopy. *Nat. Methods* **16**, 103–110 (2019).

18. Ouyang, W. et al. Deep learning massively accelerates super-resolution localization microscopy. *Nat. Biotechnol.* **36**, 460–468 (2018).

19. Lehtinen, J. et al. Noise2Noise: learning image restoration without clean data. in *Proc. 35th International Conference on Machine Learning* (eds Dy, J. & Krause, A.) 2965–2974 (PMLR, 2018).

20. Çiçek, Ö. et al. 3D U-Net: learning dense volumetric segmentation from sparse annotation. in *Medical Image Computing and Computer-Assisted Intervention* 424–432 (2016).

21. Maggioni, M., Katkovnik, V., Egiazarian, K. & Foi, A. Nonlocal transform-domain filter for volumetric data denoising and reconstruction. *IEEE Trans. Image Process.* **22**, 119–133 (2013).

22. Batson, J. & Royer, L. Noise2Self: blind denoising by self-supervision. in *Proc. 36th International Conference on Machine Learning* (eds Chaudhuri, K. & Salakhutdinov, R.) 524–533 (PMLR, 2019).

23. Krull, A., Buchholz, T.-O. & Jug, F. Noise2Void—learning denoising from single noisy images. in *Proc. IEEE Conference on Computer Vision and Pattern Recognition* (eds Davis, L., Torr, P. & Zhu, S. C.) 2129–2137 (2019).

24. Pnevmatikakis, E. A. et al. Simultaneous denoising, deconvolution, and demixing of calcium imaging data. *Neuron* **89**, 285–299 (2016).

Methods

Optical setup. A two-photon imaging system was designed to capture strictly synchronized low-SNR and high-SNR calcium recordings for validation of our method. Our system was based on a standard two-photon laser scanning microscope (2PLSM), and the detection path was specially designed to split the fluorescence in a ratio of 1:10. All components of our imaging system are commercially available or easy to fabricate. The schematic of the custom-built two-photon microscope is shown in Extended Data Fig. 7. At the forefront of the optical path, a titanium-sapphire laser system with tunable wavelength (Mai Tai HP, Spectra-Physics) was used as the illumination source to emit a linearly polarized, femtosecond-pulsed Gaussian excitation beam (920 nm central wavelength, pulse width <100 fs, 80 MHz repetition rate). A half-wave plate (AQWP10M-980, Thorlabs) was placed to adjust the polarization of the laser beam. Then the laser beam went through an electro-optic modulator (350-80LA-02, Conoptics) to modulate the excitation power and the half-wave plate was rotated to maximize the maximal extinction ratio. A 4f system composed of two achromatic lenses (AC508-200-B, Thorlabs) with the same focal length was followed to collimate the laser beam. Another 4f system (AC508-100-B and AC508-400-B, Thorlabs) with a fourfold magnification was used to expand the laser beam and guide the beam into a galvo-resonant scanner (8315K/CRS8K, Cambridge Technology) for fast optical scanning. The scanner mount was optimally designed for reliable and distortion-free scanning. Then, the beam went through a scan lens (SL50-2P2, Thorlabs) and a tube lens (TTL200MP, Thorlabs) and converged into a tight focus through a high numerical aperture (NA) water-dipping objective (×25/1.05 NA, XLPLN25XWMP2, Olympus). A high-precision piezo actuator (P-725, Physik Instrumente) was additionally used to drive the objective for fast axial scanning. The beam size at the back aperture of the objective was further restricted with an iris set behind the beam expander (L4) to keep the back aperture underfilled. The effective excitation NA was about 0.5 in our imaging experiments.

For the detection path, fluorescence excited by the Gaussian focus was first collected by the objective. High-NA detection is helpful to detect more fluorescence photons and improve the signal intensity. A long-pass dichroic mirror (DMLP650L, Thorlabs) was used to separate fluorescence by reflecting the fluorescence signals and transmitting the excitation light. A 1:9 (reflectance: transmission) non-polarizing plate beam splitter (BSN10, Thorlabs) was then placed in the detection path. All fluorescence going through the beam splitter will be split into a 10% component (low-SNR path) and a 90% component (high-SNR path), propagating in two orthogonal directions and detected by two photomultiplier tubes (PMT1001, Thorlabs). A pair of fluorescence filters (MF525-39, Thorlabs; ET510/80M, Chroma) was configured in front of each PMT to fully block wavelengths outside the emission passband of green fluorescent protein (GFP). To improve detection efficiency, we conjugated the back aperture of the objective to the sensor planes of the two PMTs using two 4f systems (TTL200-A and AC254-050-A, Thorlabs). The two detection paths can record synchronized fluorescence signals but with quite different imaging SNR. Although the high-SNR recording still suffers from noise, it can be used as the reference to identify underlying structures and calcium fluctuations. The maximal FOV of our two-photon imaging system is about 720 μm and the typical frame rate is 30 Hz for 512 × 512 pixels.

System calibration. To confirm the fluorescence intensity ratio between the high-SNR detection path and the low-SNR detection path, we imaged 1-μm green-fluorescent beads (G0100, ThermoFisher) for system calibration. The bead suspension was first diluted and embedded in 1.0% agarose and then mounted on a microscope slide to form a single bead layer composed of sparse beads. A specified region was continuously scanned to acquire 500 consecutive frames. These frames can be regarded as independent samplings of the same underlying scene. To reduce the impact of detection noise, we averaged these frames to obtain the noise-suppressed image of each path (Extended Data Fig. 8). All beads in the FOV were manually segmented and the intensity of each bead was calculated by averaging all pixels inside its segmentation mask. According to our statistical analysis, the fluorescence intensity of the high-SNR detection path was approximately tenfold higher than that of the low-SNR detection path.

Mouse preparation and calcium imaging. All experiments involving mice were performed in accordance with the institutional and ethical guidelines for animal welfare and have been approved by the Institutional Animal Care and Use Committee (IACUC) of Tsinghua University. Mice used in this study were housed in cages (24 °C, 50% humidity) in groups of 1–5 under a reverse light cycle. Both male and female mice were used without randomization or blinding.

Adult transgenic mice (Ai148D/Rasgrf2-dCre) at 8–12 postnatal weeks were anesthetized with 1.5% isoflurane, and craniotomy surgeries were conducted using a stereotaxic instrument (68018, RWD Life Science) under a bright-field binocular microscope (77001S, RWD Life Science). A custom-made coverslip fitting the shape of the cranial window (~6 mm in diameter) was embedded and cemented to the skull. A biocompatible titanium headpost was then cemented to the skull for stable head fixation. The edge of the cranial window was enclosed with dental cement to hold the immersion water of the objective. After the surgery, 0.25 mg per g (body weight) of trimethoprim (TMP) was intraperitoneally injected to induce

the expression of GCaMP6f genetically encoded calcium indicator (GECI) in layer 2/3 neurons across the whole brain. To reduce potential inflammation, 5 mg per kg (body weight) of ketoprofen was injected subcutaneously. Each mouse was housed in a separate cage for 1–2 weeks of postoperative recovery.

Imaging experiments were carried out when the cranial window became clear and no inflammation occurred. Mice were first rapidly anesthetized with 3.0% isoflurane and then fixed onto a custom-made holder with the headpost. The mouse holder was mounted on a precision translation stage with three motorized axes (M-VP-25XA-XYZL, Newport) to find the region of interest (ROI) for imaging. The correction ring of the objective was adjusted to compensate for the coverslip thickness and eliminate spherical aberrations. In all experiments, the highest excitation power after the objective was ~140 mW to avoid potential laser-induced tissue damage. Gaseous anesthesia was turned off and the mice kept awake during the whole imaging process. We performed single-plane imaging on 14 different planes at approximately 80–210 μm below the pia mater. Each plane was continuously imaged for 200 seconds. The frame rate was 30 Hz and the FOV was 550 × 575 μm (Supplementary Table 1).

Network architecture and training details. The network architecture of DeepCAD employs 3D U-Net, which is reported to have superior performance on the segmentation of volumetric data[20]. In general, the network is composed of a 3D encoder module (the contracting path), a 3D decoder module (the expanding path), and three skip connections from the encoder module to the decoder module (Extended Data Fig. 1). In the 3D encoder module, there are three encoder blocks. Each block consists of two $3 \times 3 \times 3$ convolutional layers, followed by a leaky rectified linear unit (LeakyReLU) and a $2 \times 2 \times 2$ max pooling with strides of two in three dimensions. In the decoder module, there are three decoder blocks, each of which contains two $3 \times 3 \times 3$ convolutional layers followed by a LeakyReLU and a 3D nearest interpolation. A group normalization[25] layer is configured after each convolutional layer. The skip connections link low-level features and high-level features by concatenating their feature maps. All operations (convolutions, max-poolings, and interpolations) in the network are in 3D to aggregate spatial information and temporal information because the spatiotemporal correlation of calcium imaging data is a key factor of our method (Supplementary Fig. 15). For the loss function, we used the arithmetic average of a L1-norm loss term and a L2-norm loss term. Each input stack was normalized by subtracting the minimum value and then dividing by the maximum value. The model was trained on 3D tiles with a spatial size of 64×64 pixels and a temporal size of 300 frames. Small spatial size can lower memory requirements and reduce the training time, and large temporal size is helpful to make full use of temporal information.

Adam optimizer[26] was used for network training, with a learning rate of 0.00005 and exponential decay rates of 0.5 for the first moment and 0.9 for the second moment. We used graphics processing units (GPU) to accelerate the training and testing process. It took about 12 hours to train our model for 20 epochs on a typical training set (about 1,200 3D tiles) with a single GPU (Nvidia TITAN RTX, 24 GB memory). On an already trained network, it took 502 seconds to process 3,500 frames (512 × 512 pixels). Training and testing time can be further shortened by using a more powerful GPU or parallelizing computation on multiple GPUs.

Generally speaking, a higher imaging frame rate, higher spatial sampling rate and more training frames can lead to better denoising performance (Supplementary Figure 16). The full 3D architecture of DeepCAD makes it easy to overfit because 3D convolutions usually involve more parameters than does the 2D counterpart. The best denoising performance is only achieved at the point where there is neither underfitting nor overfitting (Supplementary Figure 17). To screen out the model with the best generalization ability, we saved the network snapshot after each training epoch and evaluated its performance on a holdout validation set. We fed the validation data into each model and calculated the s.d. projection of the output stack of each model. Then, the average pixel intensity was calculated on unlabeled vascular regions for all standard deviation projections. Empirically, the best model tends to be the one with the smallest dark s.d. To avoid ignoring the best model, manual inspection is recommended as the final step of model screening. A good model often needs several rounds of training and screening.

Data simulation. Our simulation program includes the first step for synthesizing the noise-free video (ground truth) and the second step for adding the Mixed Poisson-Gaussian (MPG) noise. Firstly, based on the fact that a calcium imaging video can be modeled as the product of a spatial matrix and a temporal matrix[24], the ground-truth video was synthesized using a non-negative matrix factorization (NMF) model (Supplementary Note 2). To generate realistic simulated calcium imaging data, we constructed a neuron library containing the spatial profiles of 517 neurons. These neurons were extracted with constrained non-negative matrix factorization[24] (CNMF) from an experimentally obtained two-photon calcium imaging data of a virus-transfected wild-type mouse expressing GCaMP6f (layer 2/3 at the primary somatosensory cortex). For the spatial component that defines the location of each neuron, 120 neurons were randomly selected from the library to keep the sparsity of neurons. For the temporal component that defines the fluorescence fluctuations of each neuron, MLspike[27] was employed to generate calcium traces with GCaMP6f kinetics. Then, these two components were reshaped

into 2D matrices, and the simulated noise-free data (1 μm/pixel spatial sampling rate, 30 Hz frame rate) was synthesized as the product of the spatial matrix and the temporal matrix. The noise-contaminated counterpart was ultimately generated by adding the content-related MPG noise (Supplementary Note 1). Data with different imaging SNRs were simulated with different relative photon numbers and their relationship was investigated in Supplementary Fig. 6. All images were saved as uncompressed TIF files with the format of unsigned 16-bit integer (uint16).

Single-neuron recordings. The data of simultaneous two-photon imaging and electrophysiological recordings of single-neuron activity were released by the Svoboda lab[28] and were downloaded from the Collaborative Research in Computational Neuroscience (CRCNS) platform. Only recordings of GCaMP6f neurons were used in this study. The image stacks were fourfold downsampled to reduce the sampling rate, and some outlier recordings with very sparse spikes and low electrophysiological SNR were excluded. Fluorescence traces were extracted from the temporal stacks using manually annotated cytoplasmic and neuropil masks. In single-neuron recordings, the neuropil contamination was corrected using the formula $F_{corrected} = F_{cytoplasmic} - 0.7 \times F_{neuropil}$, since DeepCAD can restore both the cytoplasmic fluorescence and neuropil fluorescence indiscriminately (Supplementary Figs. 18 and 19). For spike inference, we used the MLspike algorithm[27], which was reported to rank first in the Spikefinder challenge[29]. All traces were divided by their mean values for normalization before being fed into the spike inference pipeline. Recommended model parameters for GCaMP6f indicator were used to ensure optimal performance of spike inference.

Data analysis of neuronal populations. Calcium imaging data of large neuronal populations were first registered with a non-rigid motion correction method[30], and the black edges of registered images were clipped. Then, CNMF[24] was employed as the source extraction method for neuron segmentation and trace extraction. A spatial matrix and a temporal matrix can be obtained from each video, storing the spatial footprints and corresponding calcium traces of all active neurons, respectively. The same set of parameters was used for the original low-SNR recording and corresponding DeepCAD enhanced counterpart, as well as the high-SNR recording. Simulated data were analyzed following the same pipeline except for the motion correction step. Along with automatic neuron extraction, we also performed manual annotations to inspect our results. High-SNR recordings were tenfold downsampled along the time axis by averaging each consecutive ten frames, which reduced the disturbance of detection noise and was helpful to improve annotation accuracy. Boundaries of all active components were annotated using the ROI Manager toolbox of Fiji. The final segmentation masks were generated through subsequent morphological operations of images and connected domain extraction implemented with customized MATLAB scripts.

Performance metrics. Two types of metrics were used for quantitative evaluation of the spatial and temporal performance of DeepCAD. For synthetic calcium imaging data, corresponding ground-truth images and calcium traces were available. SNR and PSNR were used as the spatial metric to evaluate pixel-level similarity between DeepCAD enhanced images and ground-truth images. Pearson's correlation coefficient (R) was used as the temporal metric to reflect the similarity between enhanced traces and ground-truth traces. The Pearson correlation between signal x and the reference signal y is defined as:

$$R = \frac{E[(x - \mu_x)(y - \mu_y)]}{\sigma_x \sigma_y}$$

Here μ_x and μ_y are the mean values of signal x and y, respectively; σ_x and σ_y are the standard deviations of signal x and y, respectively; and E represents arithmetic mean.

Furthermore, we also evaluated the performance of DeepCAD on the basis of more complex downstream tasks such as neuron extraction and spike inference, which are the most crucial prerequisites in functional analysis of neural circuits from calcium imaging data. We considered neuron extraction as an instance segmentation problem and adopted an object-level metric to evaluate the segmentation performance of CNMF before and after denoising[31]. Different intersection-over-union (IoU, defined as the intersection area divided by the union area of two objects) thresholds were selected to determine correctly segmented objects. For a specified IoU threshold, the segmentation accuracy (F_1 score) was defined as the harmonic mean of sensitivity and precision, which can be formulated as:

$$F_1 = \frac{2TP}{2TP + FP + FN}$$

Here TP, FP, and FN are the number of true positives, false positives, and false negatives, respectively. When CNMF was applied as the source extraction method, the SNR of calcium traces was quantified with the peak SNR automatically calculated by the CaImAn toolbox[32] with infinite outliers eliminated. For spike inference, we used the error rate (ER) to quantify the performance of spike inference, which is defined as $ER = 1 - F_1$. Spikes detected from simultaneous electrophysiological recordings were used as the ground truth for ER calculation. The evaluation process was implemented with customized MATLAB scripts. SNR, PSNR, Pearson correlation coefficient, and IoU were computed using built-in functions.

Statistics and reproducibility. All boxplots were plotted in standard Tukey box-and-whisker plot format. The box indicates the lower and upper quartiles while the line in the box shows the median. The lower whisker extends to the first data point greater than the lower quartile minus 1.5× the interquartile range (IQR). Similarly, the upper whisker extends to the last data point less than the upper quartile plus 1.5× the IQR. Data points outside of the whiskers were plotted as outliers with small black dots. Experiments in Fig. 3a–c,f–h were repeated on 5,000 test frames with similar results. Experiments in Extended Data Fig. 10 were repeated on 600 test frames, all achieving similar results.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The dataset of synchronized low-SNR and high-SNR two-photon calcium imaging (covering various imaging depths, excitation power, and cell structures) has been made publicly available at https://github.com/cabooster/DeepCAD/tree/master/dataset. The dataset of simultaneous two-photon imaging and electrophysiological recording can be downloaded from the Collaborative Research in Computational Neuroscience (CRCNS) platform at http://crcns.org/data-sets/methods/cai-1. Source data are provided with this paper.

Code availability

Our PyTorch implementation of DeepCAD is publicly available at https://github.com/cabooster/DeepCAD. The Fiji plugin and the pretrained model for denoising of large neuronal populations are readily accessible at https://github.com/cabooster/DeepCAD/tree/master/DeepCAD_Fiji. Because the plugin is only compatible with TensorFlow, a companion TensorFlow implementation of DeepCAD is also made publicly available at the same GitHub repository.

References

25. Wu, Y. & He, K. Group normalization. in *European Conference on Computer Vision* (ECCV) 3–19 (2018).
26. Kingma, D. P. & Ba, J. L. Adam: a method for stochastic optimization. in *International Conference on Learning Representations* 1–15 (2015).
27. Deneux, T. et al. Accurate spike estimation from noisy calcium signals for ultrafast three-dimensional imaging of large neuronal populations in vivo. *Nat. Commun.* **7**, 12190 (2016).
28. GENIE project. Simultaneous imaging and loose-seal cell-attached electrical recordings from neurons expressing a variety of genetically encoded calcium indicators. CRCNS.org (2015); https://doi.org/10.6080/K02R3PMN
29. Berens, P. et al. Community-based benchmarking improves spike rate inference from two-photon calcium imaging data. *PLoS Comput. Biol.* **14**, e1006157 (2018).
30. Pnevmatikakis, E. A. & Giovannucci, A. NoRMCorre: an online algorithm for piecewise rigid motion correction of calcium imaging data. *J. Neurosci. Methods* **291**, 83–94 (2017).
31. Caicedo, J. C. et al. Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl. *Nat. Methods* **16**, 1247–1253 (2019).
32. Giovannucci, A. et al. CaImAn an open source tool for scalable calcium imaging data analysis. *eLife* **8**, e38173 (2019).
33. Zeiler, M. D. & Fergus, R. Visualizing and understanding convolutional networks. in *European Conference on Computer Vision* (ECCV) 818–833 (2014).

Author contributions

Q. D., H. W., L. F., and X. Li conceived this project. Q. D., H. W., and L. F. supervised this research. X. Li and G. Z. designed detailed implementations and processed the data. X. Li designed and set up the imaging system. X. Li and G. Z. conducted the experiments. G. Z. developed the python code and the Fiji plugin. J. W., Y. Z., and X. Lin directed the experiments and data analysis. L. F., Y. Z., Z. Z, H. Q., and H. X. provided critical support on system setup and imaging procedure. J. W., L. F., Y. Z., X. Lin, H. Q., H. X., H. W., and Q. D. participated in critical discussions about the results. All authors participated in the writing of the paper.

Competing interests

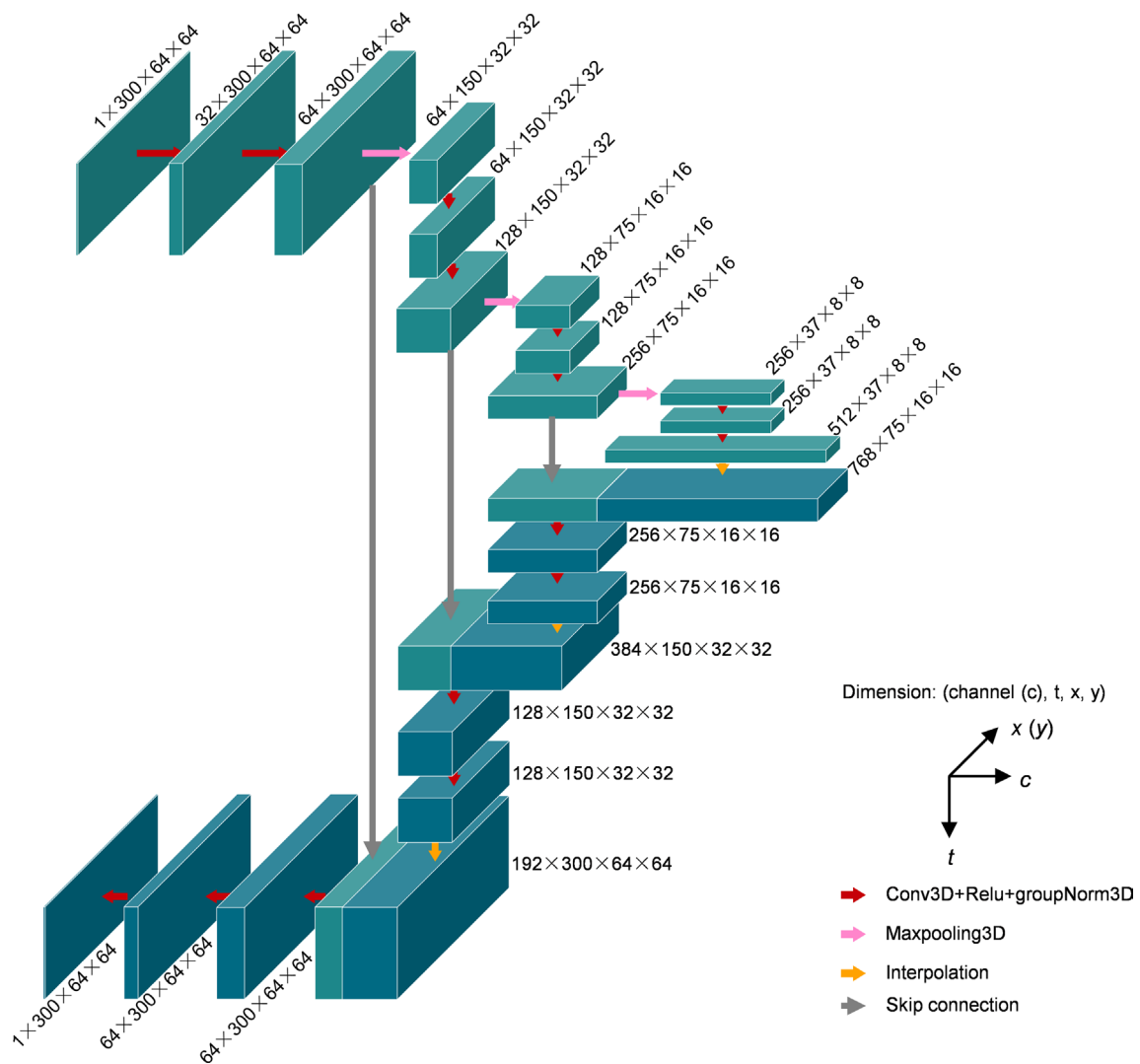The authors declare no competing financial interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41592-021-01225-0.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41592-021-01225-0.

**Extended Data Fig. 1 | Network architecture.** Our model adopted 3D U-net[20], which is composed of a 3D encoder module, a 3D decoder module, and three skip connections from the encoder module to the decoder module. In the encoder module, there are three encoder blocks. Each block consists of two $3 \times 3 \times 3$ convolutional layers followed by a leaky rectified linear unit (LeakyReLU), a group normalization layer, a $2 \times 2 \times 2$ max pooling with strides of 2 in three dimensions. In the decoder module, there are three decoder blocks, each of which contains two $3 \times 3 \times 3$ convolutional layers followed by a LeakyReLU, a group normalization layer, and a 3D nearest interpolation. The skip connections can pass feature maps from the encoder module to the decoder module to integrate low-level features and high-level features. Feature maps of the encoder module and the decoder module are represented in different colors. All operations are in 3D and feature maps are all 4D tensors. 3D (c, t, x) feature maps were used here to simplify representation.
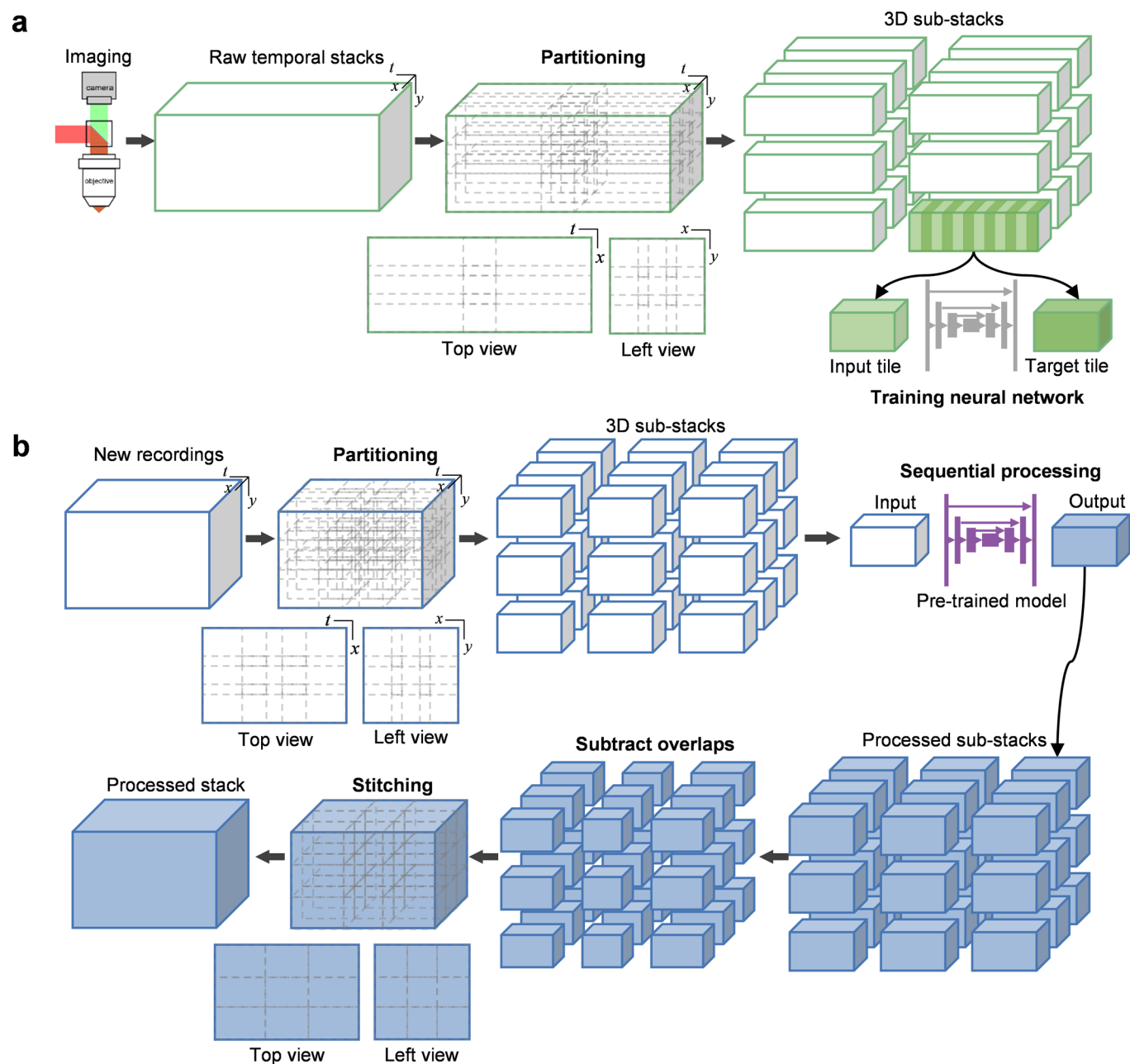
Extended Data Fig. 2 | Data processing pipeline. a, The training process. Raw data captured by the imaging system are organized in 3D (x, y, t) and saved as a temporal stack. The original noisy stack is partitioned into thousands of 3D sub-stacks (64×64×600 pixels) with about 25% overlap in each dimension. For temporal stacks with a small lateral size or short recording period, sub-stacks can be randomly cropped from the original stack to augment the training set. Then, interlaced frames of each sub-stack are extracted to form two 3D tiles (64 × 64 × 300 pixels). One of them serves as the input and the other serves as the target for network training. **b**, Deployment of the pre-trained model. New recordings obtained with the imaging system are partitioned into 3D sub-stacks (64 × 64 × 300 pixels) with 25% overlap in each dimension. Then, pre-trained models are loaded into memory and the sub-stacks are directly fed into the model. Enhanced sub-stacks are sequentially output from the network and overlapping regions (both the lateral and temporal overlaps) are subtracted from the output sub-stacks. The final enhanced stack can be obtained by stitching all sub-stacks.
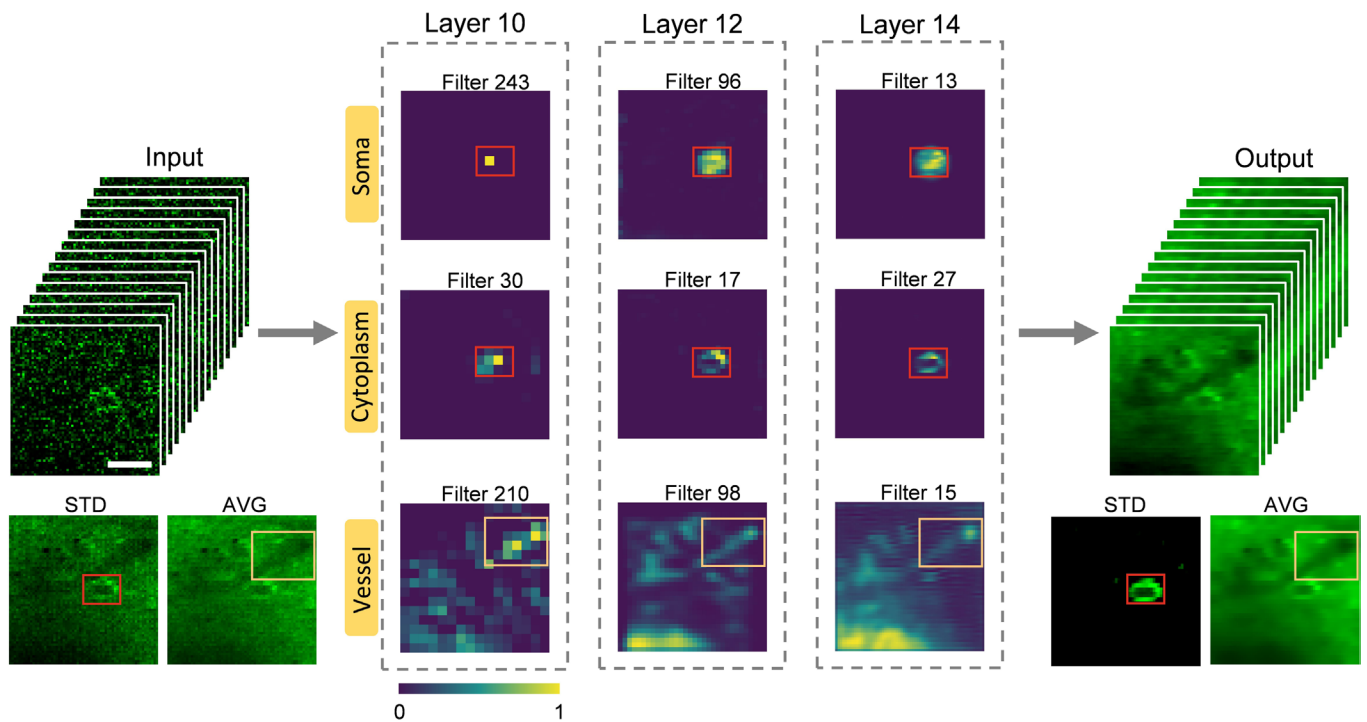
Extended Data Fig. 3 | Interpretability of DeepCAD model. To demonstrate the interpretability and reliability of our pre-trained DeepCAD model, a small 3D patch (64 × 64 × 300 pixels) was fed into the model and feature maps of the convolutional layers were visualized[33]. Scale bar, 20 μm. Example feature maps of three intermediate convolutional layers in the decoder module (Layer 10, Layer 12, and Layer 14) are shown here, displayed as the average intensity projection (AVG) of the original 3D feature maps. The feature representations learned by DeepCAD have substantial semantic meaning, such as soma-like structures, cytoplasm-like structures, and vessel-like structures (or shadows). These interpretable semantic representations would contribute to locating neurons, restoring cytoplasmic fluorescence, and avoiding unwanted intensity fluctuations in vascular regions.
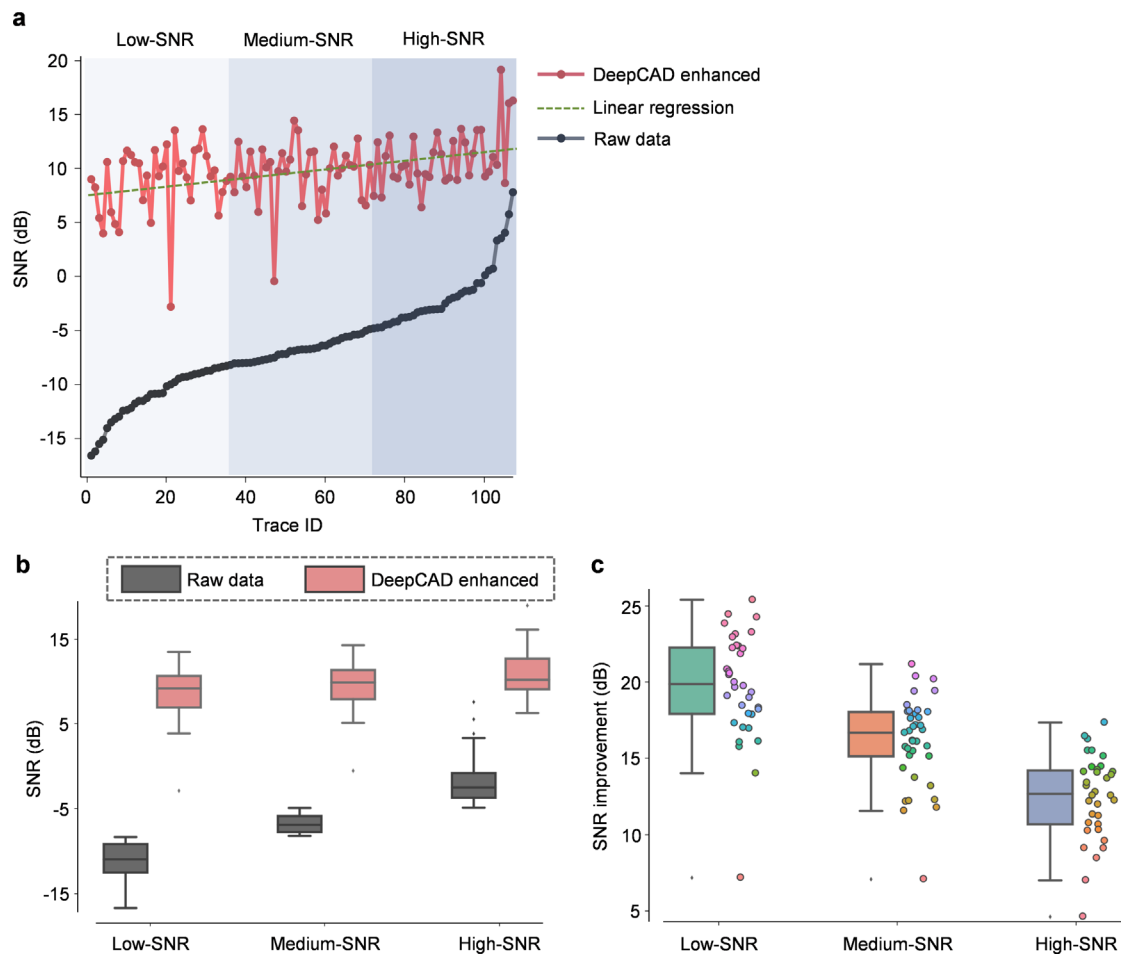
Extended Data Fig. 4 | SNR improvement of calcium traces after denoising. a, Trace SNR before and after denoising. Calcium traces (N = 107) were divided into three groups according to input SNR (36 low-SNR traces, 37 medium-SNR traces, 34 high-SNR traces). Quantitatively, low-SNR traces are those with SNR < −8.10 dB, medium-SNR traces are those with −8.10 dB ≤ SNR < −4.71 dB, high-SNR traces are those with SNR ≥ −4.71 dB. **b**, The distribution of trace SNR before and after denoising (N = 36 for low-SNR, N = 37 for medium-SNR, N = 34 for high-SNR). **c**, SNR improvements at different input SNR levels (N = 36 for low-SNR, N = 37 for medium-SNR, N = 34 for high-SNR). The trace SNR was calculated by $10\log(||x||/||y-x||_2)$, where x is the normalized calcium trace and y is corresponding normalized noise-free trace estimated by MLspike[27]. Boxplots were plotted in standard Tukey box-and-whisker plot format with outliers indicated with small black dots.
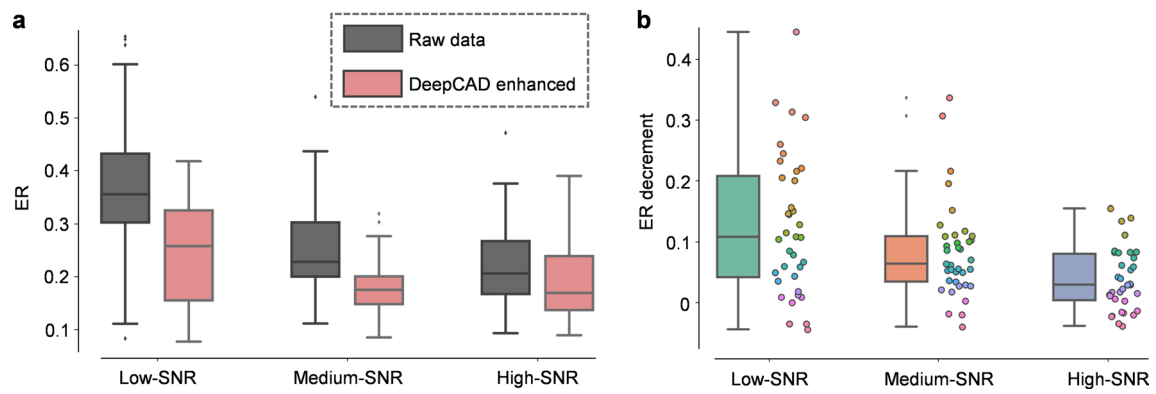
Extended Data Fig. 5 | DeepCAD reduces the error rate of spike inference at different input SNRs. a, The error rate (ER) of raw data and DeepCAD enhanced data at different input SNRs. **b**, The decrements of ER at different input SNRs. Sample size: N = 36 for low-SNR, N = 37 for medium-SNR, N = 34 for high-SNR. Boxplots were plotted in standard Tukey box-and-whisker plot format with outliers indicated with small black dots.
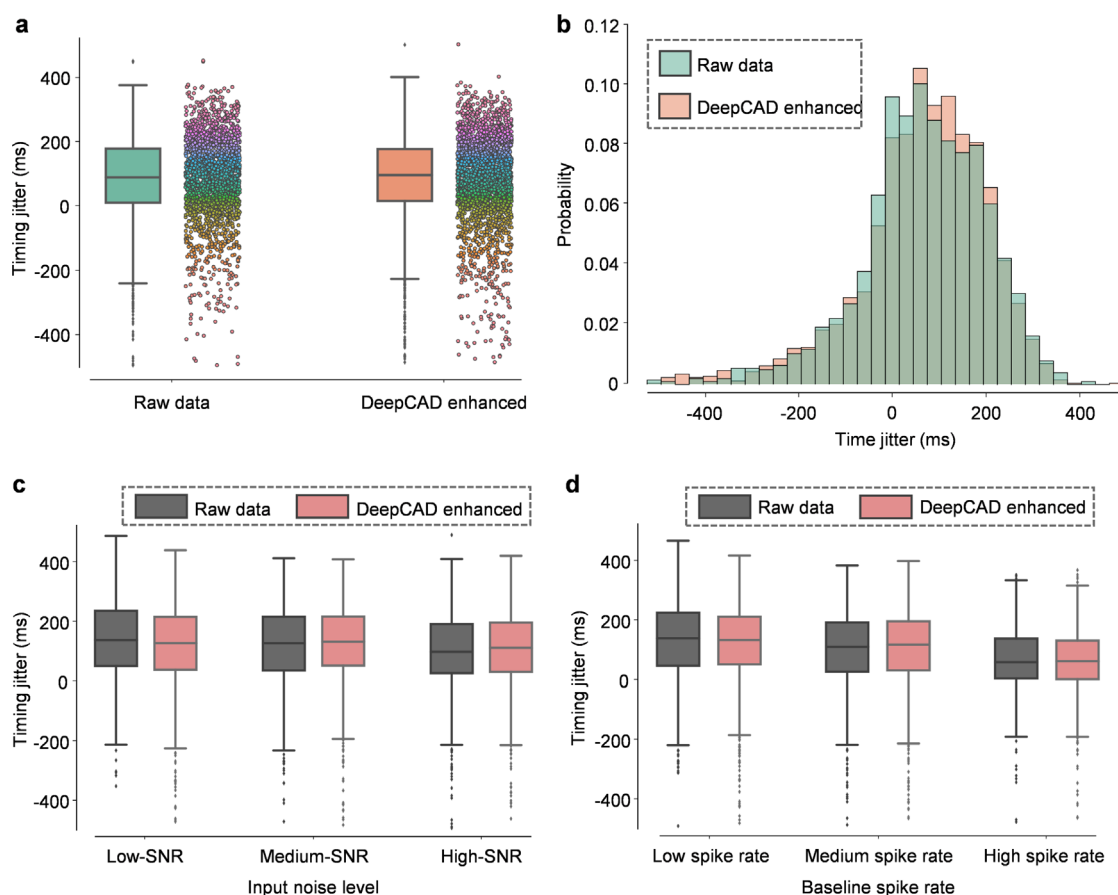
Extended Data Fig. 6 | Timing jitters of inferred spikes relative to real spikes before and after denoising. a, Boxplots showing the distribution of timing jitters relative to real spikes (electrophysiology) of all inferred spike pairs before (N = 2031) and after (N = 2574) denoising. **b**, Histograms showing the probability distributions of timing jitters before and after denoising. The two probability distributions were verified to be equivalent by Kolmogorov–Smirnov test (one-side, P ≤ 0.01, N = 2031 for raw data, N = 2574 for DeepCAD enhanced). **c**, Distributions of timing jitters at different input noise levels (Raw data, N = 326 for low-SNR, N = 689 for medium-SNR, N = 1016 for high-SNR; DeepCAD enhanced, N = 545 for low-SNR, N = 880 for medium-SNR, N = 1149 for high-SNR). **d**, Distributions of timing jitters at different baseline spike rates (Raw data, N = 663 for low spike rate, N = 766 for medium spike rate, N = 602 for high spike rate; DeepCAD enhanced, N = 1095 for low spike rate, N = 837 for medium spike rate, N = 642 for high spike rate). Baseline spike rates were calculated with 2 s binning time. All timing jitters were divided into three groups, *that is* low spike rate (baseline spike rate ≤ 2.0 spk/s), medium spike rate (2.0 spk/s < baseline spike rate ≤ 3.5 spk/s), and high baseline spike rate (baseline spike rate å 3.5 spk/s). These timing jitters were caused by the spike inference algorithm. Boxplots were plotted in standard Tukey box-and-whisker plot format with outliers indicated with small black dots.
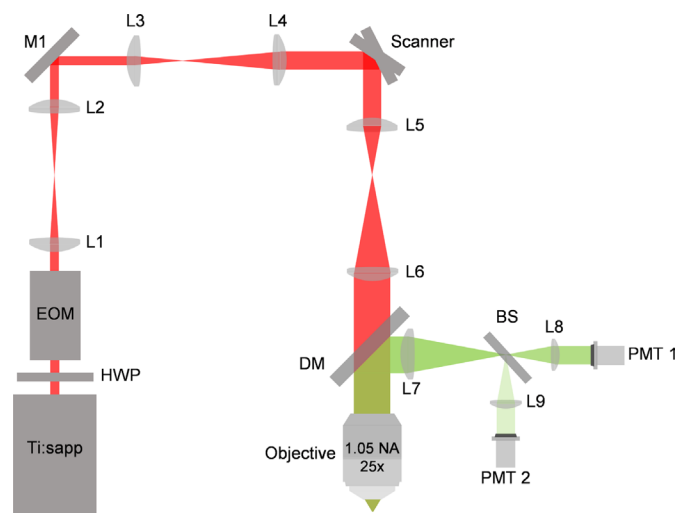
Extended Data Fig. 7 | Simultaneous low-SNR and high-SNR two-photon imaging system. Microscope set-up for simultaneous acquisition of high-SNR and low-SNR calcium imaging data. Ti:sapp: titanium-sapphire laser with tunable wavelength; HWP: half-wave plate; EOM: Electro-Optic Modulator; M1: mirror; L1, L2, L3, L4, L5, L6, L7, L8, L9: lens; Scanner: galvo-resonant scanners; DM: long-pass dichroic mirror to separate fluorescence signals (green path) from excitation light (red path); BS: 1:9 (reflectance: transmission) non-polarizing plate beam splitter; PMT1, PMT2: photomultiplier tubes. Fluorescence signals were split into a low-SNR (~10%) component and a high-SNR (~90%) component and were synchronously detected by two PMTs.
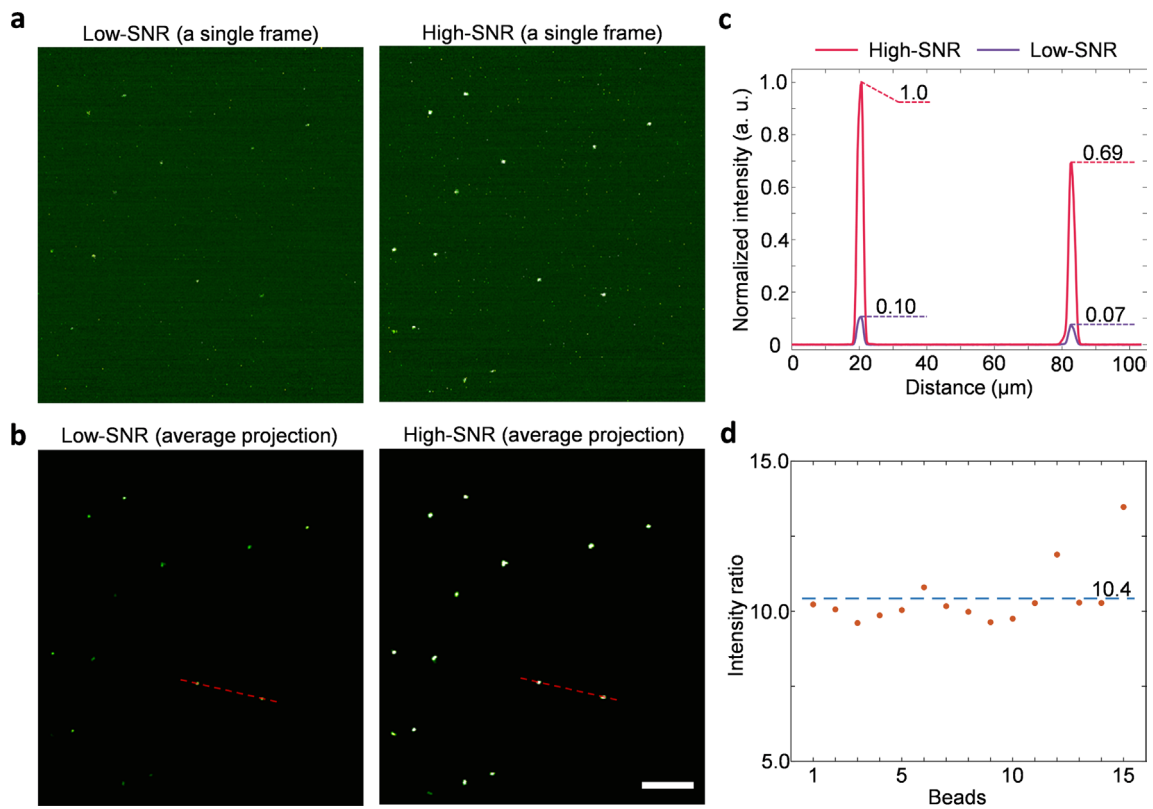
Extended Data Fig. 8 | System calibration. a, Example frames captured by the low-SNR detection path (left) and the high-SNR detection path (right). There were 15 isolated fluorescent beads (1 μm diameter) in the field of view (FOV). **b**, Average projection of 500 continuously acquired frames. Scale bar, 50 μm. **c**, Intensity profiles (normalized to the maximum of high-SNR recording) along the red dashed lines in **b**. **d**, The intensity ratios (high-SNR relative to low-SNR) of all 15 fluorescent beads. Each point represents one bead. The average intensity ratio is 10.4 (blue dashed line).
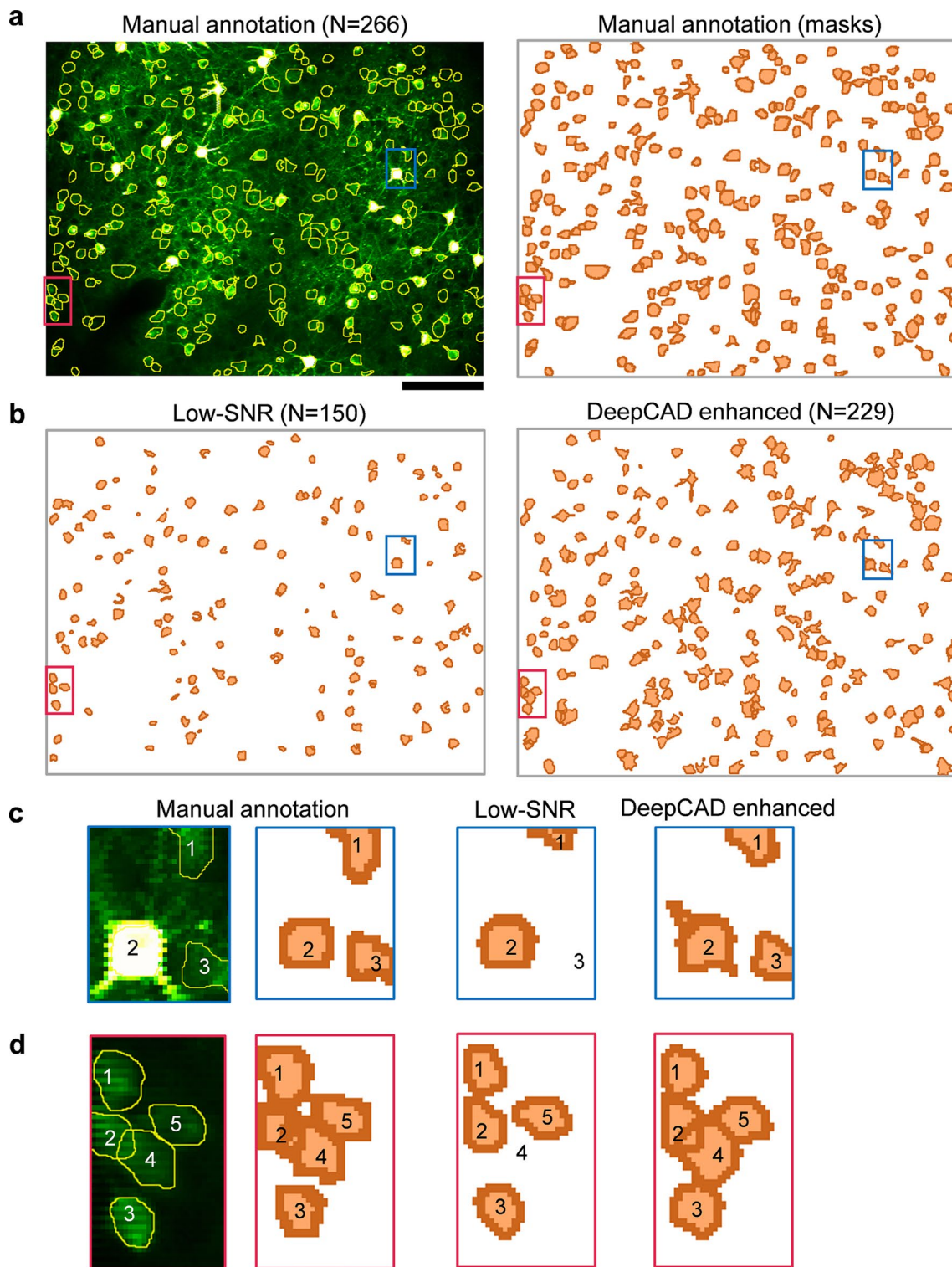
Extended Data Fig. 9 | Human inspection of segmentation results. a, Left: manually annotated neuron borders. The standard deviation projection served as the background image. Right: manually annotated segmentation masks. Scale bar, 100 μm. **b**, Left: segmentation masks of the Low-SNR recording. Right: segmentation masks of the DeepCAD enhanced recording. The constrained nonnegative matrix factorization (CNMF) algorithm[24,32] was used as the segmentation method. **c**, Magnified view of the blue boxed region showing the segmentation of three neurons. **d**, Magnified view of the red boxed region showing the segmentation results of five neurons.
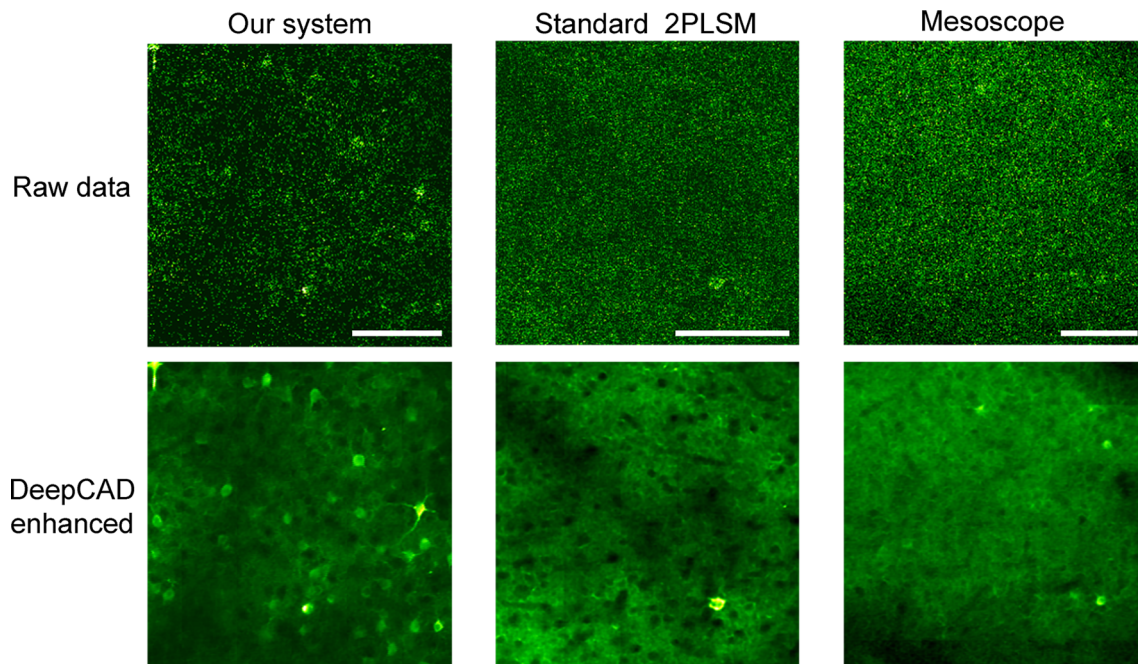
Extended Data Fig. 10 | Cross-system validation. Denoising performance of DeepCAD on three two-photon laser-scanning microscopes (2PLSMs) with different system setups. Our system was equipped with alkali PMTs (PMT1001, Thorlabs) and a 25×/1.05 NA commercial objective (XLPLN25XWMP2, Olympus). The standard 2PLSM was equipped with a GaAsP PMT (H10770PA-40, Hamamatsu) and a 25×/1.05 NA commercial objective (XLPLN25XWMP2, Olympus). The two-photon mesoscope was equipped with a GaAsP PMT (H11706-40, Hamamatsu) and a 2.3×/0.6 NA custom objective. The same pre-trained model was used for processing these data. All scale bars represent 100 μm.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The image acquisition process of our custom two-photon microscopy was controlled by ScanImage 5.5R1 (Free release, Vidrio). We used the GUI Software Package (version 4.0, Thorlabs) customized for Thorlabs PMTs to control our PMTs.
Data analysis	Data simulation and result analysis were performed using custom Matlab (R2017b, MathWorks) scripts. Deep learning models reported in this work were implemented with standard libraries of Python (3.6.0), Pytorch (1.3.1, Facebook), and TensorFlow (1.4.0, Google). Source extraction of calcium imaging data was implemented with the open source CaImAn toolbox (https://github.com/flatironinstitute/CaImAn). Our Fiji plugin was based on the CSBDeep (0.6.0) framework of Fiji (1.53c, https://imagej.net/CSBDeep). Spike inference was implemented with open-source MLspike (https://github.com/MLspike). Manual annotation of neural masks was implemented with ROI Manager toolbox of Fiji (1.53c). The complete code of DeepCAD has been made publicly available at https://github.com/cabooster/DeepCAD.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The dataset of synchronized low-SNR and high-SNR two-photon calcium imaging (covering various imaging depths, excitation power, and cell structures) used in Fig. 3 has been made publicly available at https://github.com/cabooster/DeepCAD/tree/master/dataset. The dataset of simultaneous two-photon imaging and

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For two-photon calcium imaging of mouse cortex, 14 independent field-of-views (FOVs) at different depths, excitation power and cell structures were recorded. For demonstrating our method, this sample size is sufficient.
Data exclusions	We performed imaging experiments on 3 transgenic mice (GCaMP6f) with different injection conditions, and only the one with the healthiest expressing condition was used in this study. For simultaneous two-photon imaging and electrophysiological recordings, outlier recordings with very sparse spikes and low electrophysiological SNR were excluded with manual screening.
Replication	Imaging and data analysis results were reliably reproduced and similar results have been obtained over 14 groups of recordings with various imaging depths, excitation power, and cell structures.
Randomization	Not relevant, as there were no such experimental groups in this study.
Blinding	Not relevant, as there were no such experimental groups in this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Male or female transgenic mice (Ai148D/Rasgrf2-dCre) expressing GCaMP6f after Trimethoprim (TMP) injection were used in this study. All mice were aged 8-12 weeks and were housed in cages (24°C, 50% humidity) in groups of 1-5 under a reverse light cycle before surgeries. Each mouse was housed solely for 1-2 weeks of postoperative recovery after surgeries.
Wild animals	Not involved in this study
Field-collected samples	Not involved in this study
Ethics oversight	All experiments involving mice were performed in accordance with institutional guidelines for animal welfare and have been approved by the Institutional Animal Care and Use Committee (IACUC) of Tsinghua University.

Note that full information on the approval of the study protocol must also be provided in the manuscript.