Supplementary Materials of MIN1PIPE

Contents

A Supplemental Figures

A.1 Supplemental Figure S1, related to Figure 1

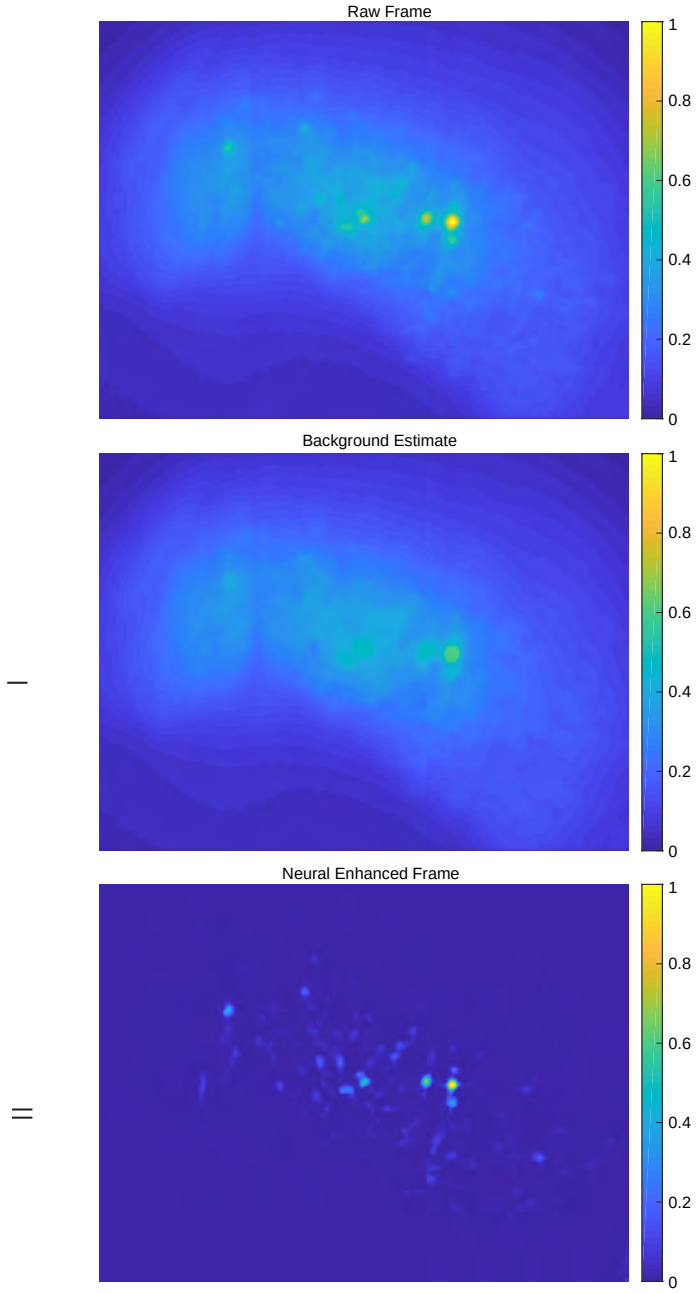

Figure S1: Background estimation and removal. Top panel: a raw frame before neural enhancing. Middle panel: the background estimation with morphological opening operation, with a disk shape structure element of size 9 (the default size for Inscopix miniscopes with a downsampling factor of 2). The raw frame is first denoised with anisotropic diffusion before the morphological opening. As shown here, the background structure and details are well approximated. Bottom panel: the neural enhanced version of the same frame. It is computed by subtracting the background from the raw frame. All the panels are in parula color map and appropriate color range for best visualization of the performance.

A.2 Supplemental Figure S2, related to Figure 1

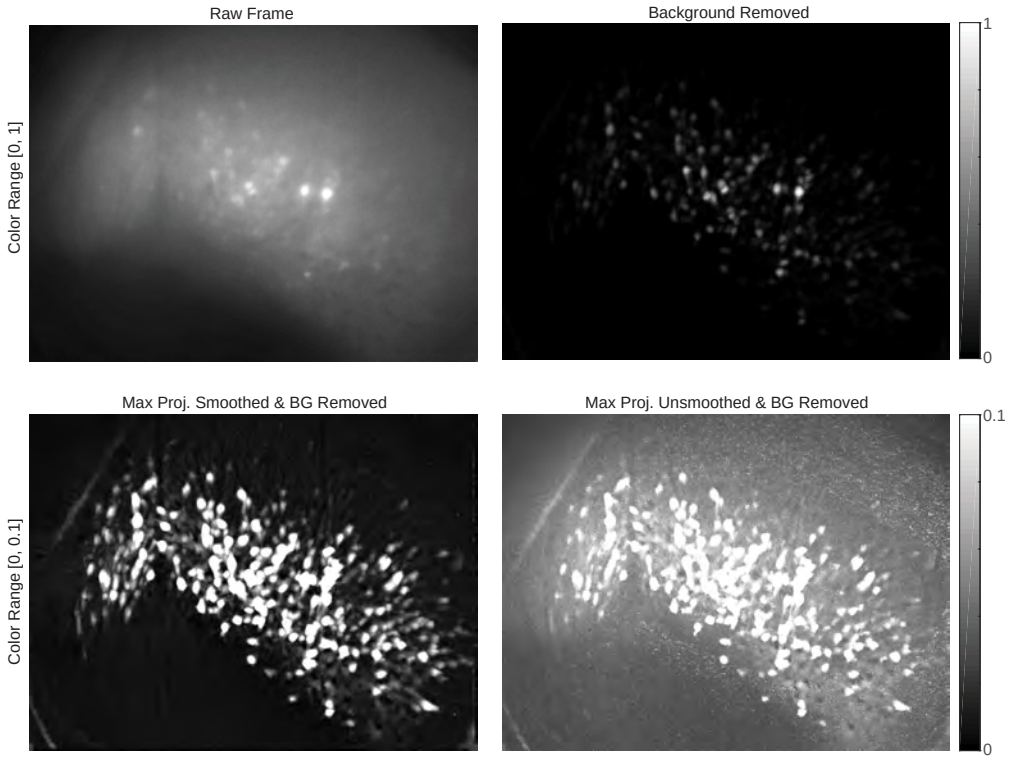

Figure S2: Necessity of the frame-wise denoising. Upper left panel: a raw frame. Upper right panel: the neural enhanced version of the same frame. Color range is [0, 1] for the both upper panels. Lower left panel: the max projection of neural enhanced video, including both denoising and background removal. Lower right panel: the max projection of background removed video without first denoising. The background noise is mostly removed with the complete neural enhancing module but remains large without anisotropic diffusion. Color range is [0, 0.1] for the both bottom panels. The color map is grayscale for best visualization of the background noise

A.3 Supplemental Figure S3, related to Figure 1

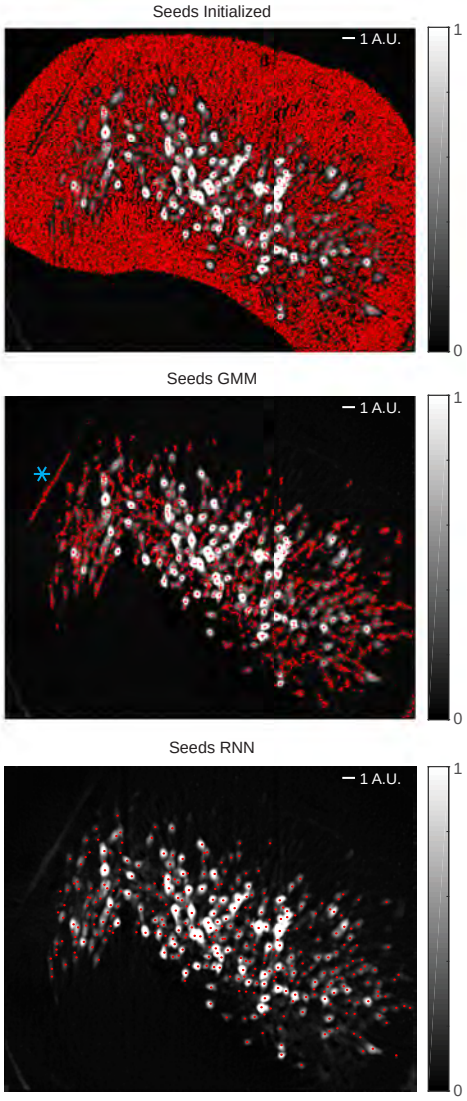

Figure S3: Visualization of the seeds cleansing process. Top panel: the over-complete set of seeds initially generated using the randomized max pooling operation, superimposed on the max projection of neural enhanced video. Middle panel: the set of seeds after GMM classifier. The seeds with low peak-valley difference are removed by the GMM. Bottom panel: the set of seeds after RNN classifier. Many seeds with relatively large peak-valley difference but without calcium spike dynamics are removed, such as the upper left line of noise seeds (blue * in the middle panel) indicating the edge of the prism probe.

A.4 Supplemental Figure S4, related to Figure 1

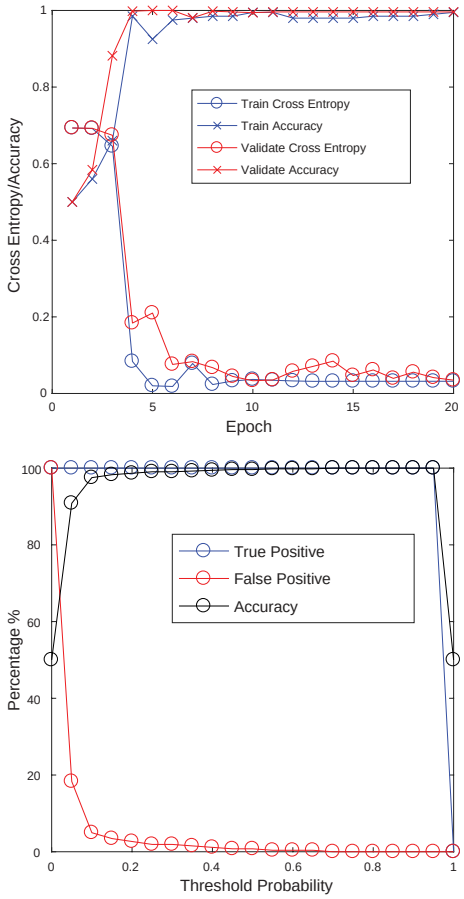

Figure S4: Training and performance of the RNN classifier. Upper panel: Training results of the RNN classifier. With the cross entropy decreasing in both training dataset and the validating dataset during training, which is a separate dataset from the training set, the accuracy of classification in both the training and validating sets increases and reaches plateau of near optima. Lower panel: the performance of the RNN classifier on the testing set.Though we present the performance as a function of the threshold probability, in practice we use 0.5 which naturally sets the correct label. The RNN classifier both selects as many real neurons as possible, and as few false positives as possible.

A.5 Supplemental Figure S5, related to Figure 2

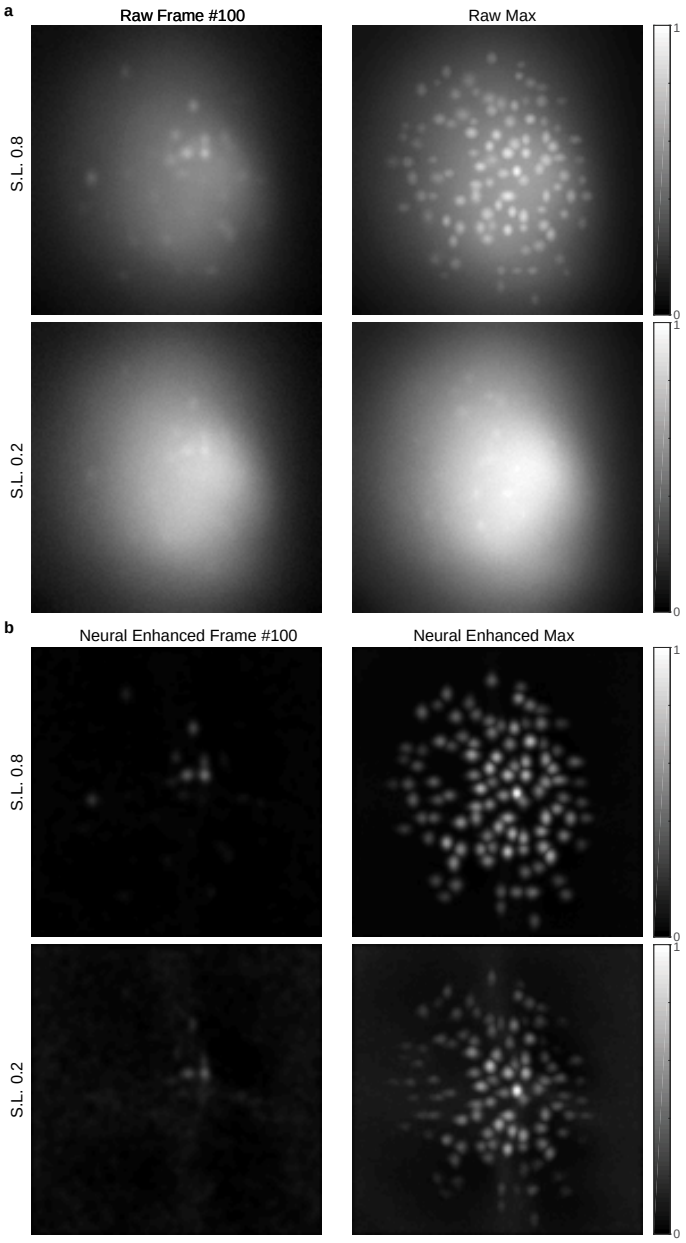

Figure S5: Demonstration using simulation datasets of different signal levels (S.L.). Upper panels: 0.8 signal level. Lower panels: 0.2 signal level. We show an example frame of each simulated video, the max projection of the raw video, an example frame of the neural enhanced version of the same frame, and the max projection of the neural enhanced video. As shown here, the 0.8 signal level shows a very clean field of view that can represent the highest quality of the miniscope imaging, while the 0.2 signal level can represent the lowest quality of miniscope imaging. On the other hand, the neural enhancing module successfully removes the background even in the very low signal level conditions.

A.6 Supplemental Figure S6, related to Figure 2

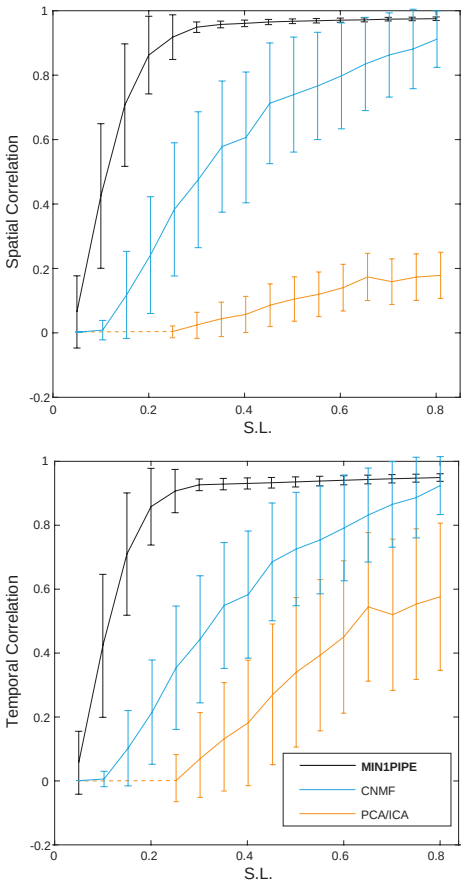

Figure S6: Comparison of detailed performance between MIN1PIPE and the other currently used methods in simulated datasets. The upper panel shows the spatial correlation and the lower panel shows the temporal correlation between the identified ROIs and the ground truth of the three methods. In general, MIN1PIPE has the most robust and sustainable performance over the range of signal levels that are most similar to the real conditions. CNMF has reasonably good performance when the signal level is high but degrades quickly when the signal level decreases. PCA/ICA does poorly in finding the real ROI spatial footprints due to the lack of localization constraints. Note that even when the signal level is extremely low, MIN1PIPE can still find a reasonable number of neurons. On the other hand, PCA/ICA fails in finding a single neuron when the signal level is lower than 0.2.

A.7 Supplemental Figure S7, related to Figure 2

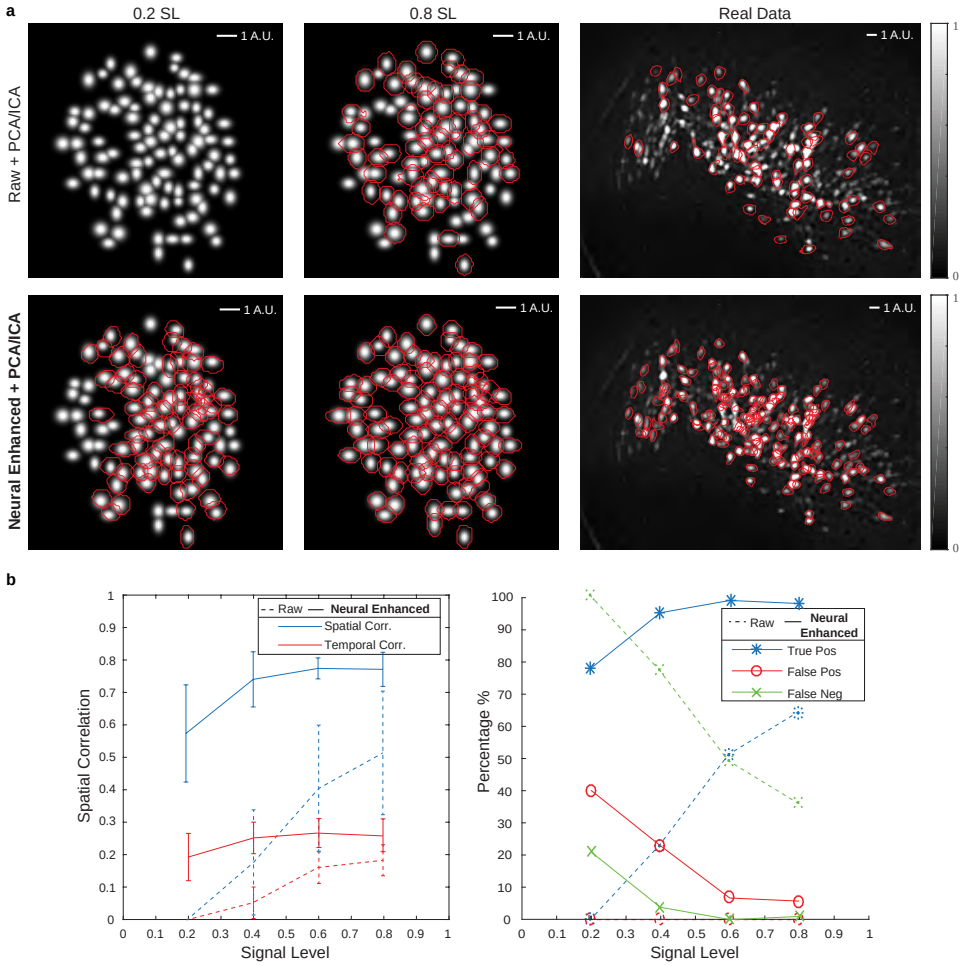

Figure S7: The neural enhancing module improves the accuracy of ROI identification when combined with PCA/ICA. **a** Visualization of the identified ROIs with contours using the raw or the neural enhanced video as an input to PCA/ICA superimposed on the max projection of ground truth (for simulation) and neural enhanced data (for real data). **b** A comparison of the average spatiotemporal correlation between the identified ROIs and the ground truth using the raw or the neural enhanced video, and the accuracy and precision of the two inputs. Our neural enhancing module significantly improves the performance of PCA/ICA. The number of identified ROIs and overall spatiotemporal correlation significantly increase at all signal levels. Similarly, in the case of real data, the number of identified ROI increases compared to that using raw data. There are still potential false positive and duplicated ROIs, and this can only be removed by manual intervention due to the limitations of PCA/ICA method itself.

A.8 Supplemental Figure S8, related to Figure 2

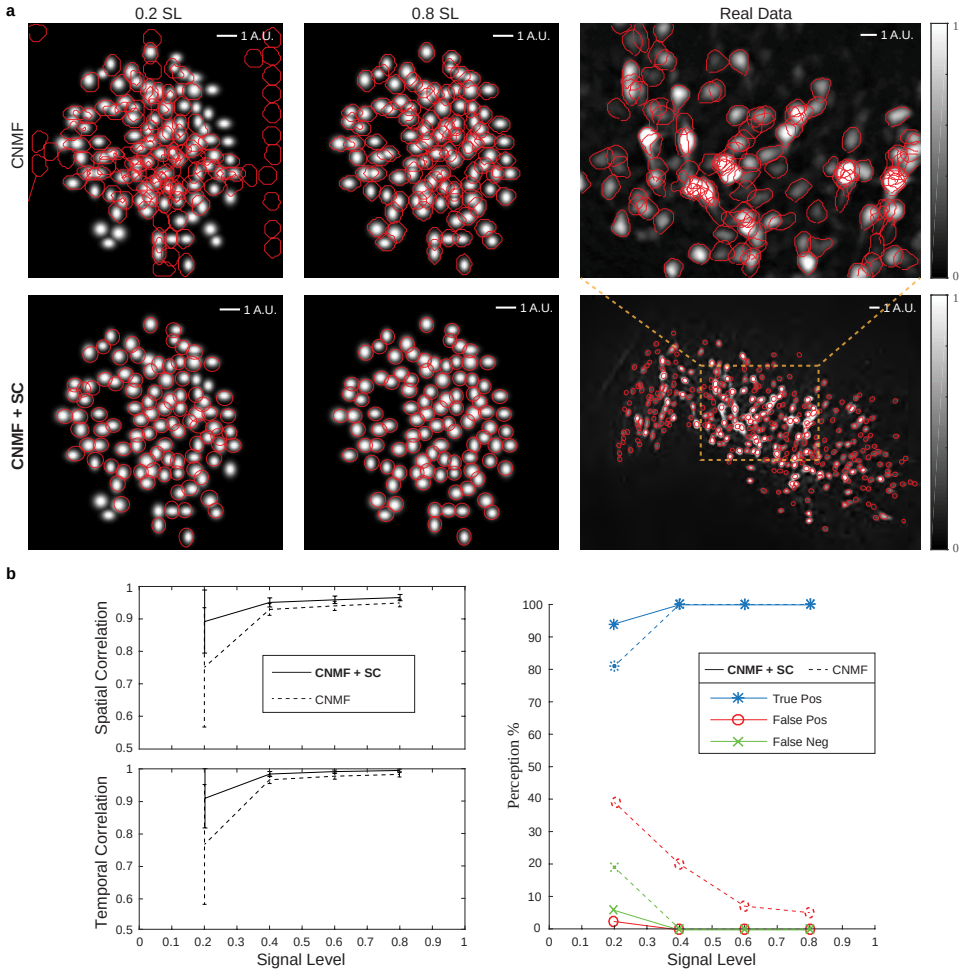

Figure S8: The seeds cleansing module improves the accuracy of ROI identification using CNMF. **a** Visualization of the identified ROIs with contours using the pure CNMF or in combination with the seeds cleansing module superimposed on the max projection of ground truth (for simulation) and neural enhanced data (for real data). **b** A comparison of the average spatiotemporal correlation between the identified ROIs and the ground truth with the pure CNMF or the seeds cleansing module + CNMF, and accuracy and precision of the two processes. The pure CNMF can (1) miss true ROIs, (2) generate duplicate seeds within a single ROI, and (3) include false positives when signal-to-noise ratio is not high enough. With the seeds cleansing module, a maximally correct set of ROIs is returned, without tuning the number of ROIs by trial-and-error.

A.9 Supplemental Figure S9, related to Figure 2

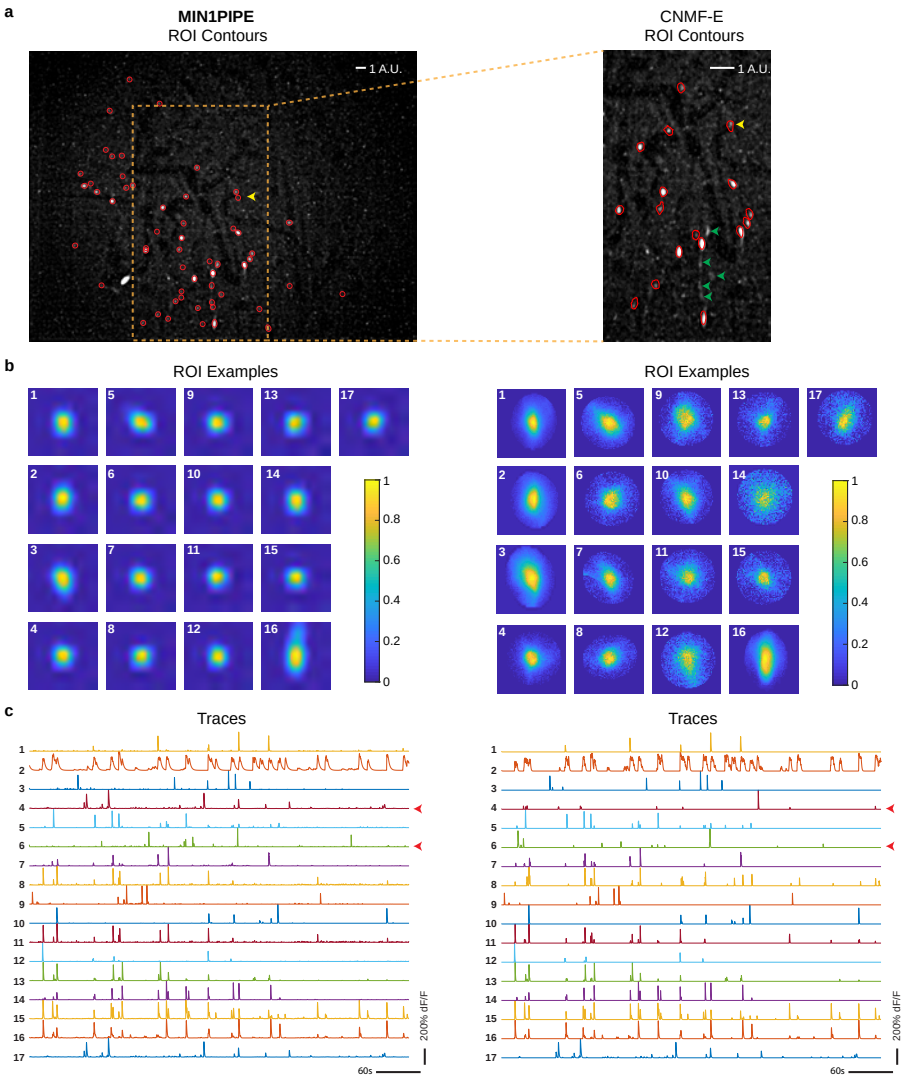

Figure S9: Comparison between MIN1PIPE and CNMF-E on analyses of miniscope imaging data from Area X in zebra finch. **a** Visualization of the identified ROIs with contours using the MIN1PIPE or CNMF-E superimposed on the max projection of neural enhanced data. The CNMF-E can only handle the cropped video due to memory issue. The CNMF-E detects 17 ROI components, which are all detected by MIN1PIPE (ROIs in red), and there are also several ROI components undetected with CNMF-E but detected with MIN1PIPE (arrows in green). **b** Visualization of the individual ROI components that are detected by the both methods. The MIN1PIPE returns smoothed, localized ROI components, whereas the CNMF-E returns ones contaminated by noise. **c** Visualization of the calcium traces of the individual ROIs. In general, the two methods returns comparable calcium traces. The only differences lies in the traces No.4 and No.6, where the No.4 detected in CNMF-E is actually composed of two ROI components (indicated in yellow arrow in **a**), and the No.6 contains some different calcium events.

B Notes on the Supplemental Videos

B.1 Supplemental Video S1, related to Figure 2

Demo of Simulated Datasets We show raw simulated datasets of two different colormaps, parula and grayscale. The three panels represent simulated datasets at signal level 0.05, 0.2, and 0.8.

B.2 Supplemental Video S2, related to Figure 2

Demo of the Results Obtained with MIN1PIPE at Signal Level 0.8 We show the *Raw*, *Neural Enhanced* and *Processed* simulated data at signal level 0.8. The *Raw* video shows the raw imaging data, while the *Neural Enhanced* and the *Processed* video represents the video after neural enhancing and the full signal extraction in MIN1PIPE separately. Specifically, after neural enhancing we can get the preprocessed video with every frame enhanced, but no separated spatiotemporal information of individual ROIs. However, after the whole process we do get every ROI and its corresponding temporal signal separated. We then take the product of the separated spatial and temporal matrix to reconstruct the processed video. The panels have two different colormaps, parula and grayscale.

B.3 Supplemental Video S3, related to Figure 2

Demo of the Results Obtained with MIN1PIPE at Signal Level 0.2 We show the *Raw*, *Neural Enhanced* and *Processed* simulated data at signal level 0.2. The panels have two different colormaps, parula and grayscale.

B.4 Supplemental Video S4, related to Figure 2

Demo of Movement Correction: Ferret Posterior Parietal Cortex The *Raw* video shows the raw two-photon imaging, the *Neural Enhanced* shows the processed video after applying the neural enhancing module, and the *Movement Corrected* shows the video after applying the hierarchical movement correction module. The video contains 1000 frames.

B.5 Supplemental Video S5, related to Figure 3

Demo of the Results Obtained with MIN1PIPE on the Mouse Barrel Cortex Imaging Data We show the *Raw*, *Neural Enhanced* and *Processed* barrel cortex data with full field of view (cropped field for CNMF). The panels have two different colormaps, parula and grayscale.

B.6 Supplemental Video S6, related to Figure 1 and 3

Demo of the Comparison between the Three Methods on the Mouse Barrel Cortex Imaging Data We show the *Raw*, and *Processed* (using the three methods, MIN1PIPE, CNMF and PCA/ICA) barrel cortex data with full field of view. The panels have two different colormaps, parula and grayscale.

B.7 Supplemental Video S7, related to Figure 4 and 5

Demo of the Results Obtained with MIN1PIPE on the Zebra Finch Area X Imaging Data We show the *Raw*, *Neural Enhanced* and *Processed* Area X data with full field of view. We concatenated demo singing epochs with 2 seconds before and 5 seconds after each singing event. The panels have two different colormaps, parula and grayscale.

C Supplementary Notes S1: Experimental Data

C.1 Simulated Data

We synthesized a series of 5-minute long videos with 10 frames per second (fps), containing 100 neurons with average neuron diameter γ_0 comparable to real data in a 128×128 field of view. Each neuron was simulated by a 2D Gaussian function with a variance in its two dimensions ($\gamma = \gamma_0 + \xi,\ \xi \sim \mathcal{N}(0, 1)$), and the calcium dynamics was generated by the first-order autoregressive process (decay time constant 0.95) with randomized spike events (spiking probability equals 1%), modified from [26]. The background was extracted from the real single-photon miniscope imaging videos separately, and temporal fluctuations of the background and spatial noises were added. The raw ground truth neural signals were then synthesized and added to the noisy background with a varying ratio, named here signal level, ranging from 0.05 to 0.8 with a step size 0.05 (*e.g.* $0.8\times$ signal $+$ background). Since the fluorescence intensity is positively correlated to the light stimulation, the spatial positions of simulated neurons were densely generated in the high background illuminated area and with a higher intensity of brightness.

C.2 Data Obtained from Mouse Barrel Cortex

The miniscope calcium imaging data were collected from the barrel cortex of an awake and freely behaving mouse using *nVista* system and prism probes (*Inscopix Inc.*). The LED power was maintained between 10 and 20%, corresponding to 0.12 and $0.24\ \mathrm{mW/mm^2}$. The size of the field of view was 1080×1440 pixels, or $900 \times 650\ \mu\mathrm{m}$. Pixel size was $0.625 \times 0.625\ \mu\mathrm{m}$. Imaging data were acquired at 20 fps, with 5933 frames. To test different methods, we spatially downsampled the videos with a factor of 2 (540×720 pixels), and the temporal downsampling rate was 2. For CNMF, the data field of view was cropped to 200×250 pixels after the downsampling.

C.3 Data Obtained from Area X in Zebra Finch

The LED power was maintained between 10 and 20%, corresponding to 0.12 and $0.24\ \mathrm{mW/mm^2}$. The size of the field of view was 1080×1440 pixels, or $900 \times 650\ \mu\mathrm{m}$. Pixel size was $0.625 \times 0.625\ \mu\mathrm{m}$. Imaging data were acquired at 10 fps, with 5701 frames, composed of concatenated 15 or $30\mathrm{sec}$ trials in which LED power was on and data were recorded. To test different methods, we spatially downsampled the videos with a factor of 2 (540×720 pixels) after cropping the edge of the field of view that contained some artifacts for MIN1PIPE and PCA/ICA, and further cropped a center patch (286×251 pixels after downsampling) for CNMF and CNMF-E. To analyze the neural correlation with singing, we epoched the neural traces of identified ROIs with 0.5 second before and 4.5 seconds after each song onset moment.

D Supplementary Notes S2: Experimental Details

D.1 Evaluation of Different Methods using Simulations or Real Data

We applied the three methods to the 16 simulated datasets and two real datasets (from mouse barrel cortex and Area X in zebra finch). For PCA/ICA, we used the commercially available *Mosaic* software (Inscopix Inc.) as PCA/ICA implementation, and processed the data by following the standard workflow in the software manual. In particular, we chose the number of principal components (PC) and independent components (IC) based on the sugguested rate (*e.g.* 20% more ICs and 50% more PCs than the estimated number of ROIs). After the neural signal extraction, an experienced neuroscientist manually rejected the unreal neurons based on their appearance and calcium traces. For CNMF, we used the default initialization strategy and the package as described in [26]. Specifically, we chose the number of neurons, K, to be 20% more than the expected neurons (for example $K = 120$ for simulated datasets).

To quantify the results, we employed a scoring criterion that calculates the spatial and temporal similarity between the ground truth and the identified ROIs. The spatial ROI footprints and the calcium traces were first rescaled to [0, 1] and then the cosine similarity were computed, as the correlation between the identified components and the ground truth. This measurement can reflect the level of spatial locality objectively. To measure the precision of the ROI identification results, we calculated the true positive, false positive and false negative of the identified population of ROIs. For real data, we used the contours to annotate the identified ROIs, and superimpose them to max projection of preprocessed videos to show the performance of ROI footprints identification. In addition, we showed some examples of the calcium traces and compared across the three methods. All methods and analyses and *in silico* experiments were implemented in Matlab.

D.2 Miniscope Calcium Imaging of Mouse Barrel Cortex

Adult C57BL/6 mice (2 months) were anesthetized with and placed in a custom stereotaxic apparatus. We made ~ 1 mm craniotomies in the skull above the sensorimotor cortex. Using a glass pipette attached to a pressure injection system (*Drummond Nanoject II*), AAV2/1-hsyn-GCaMP6f (*Penn Vector*) was injected (coordinates: right hemisphere 1.5 mm posterior to the bregma, 3.5 mm lateral to the midline, and three sites (0.2, 0.4 and 0.6 mm) in depth)), and a 1-mm diameter GRIN prism (*Inscopix*) was implanted at the injection site. After allowing 6 weeks for virus expression, a baseplate was placed on the skull to hold a miniscope nVista for imaging. After another period of recovery ($\sim$3 d), the mice were placed in a recording chamber and the miniature microscope was attached to the baseplate for imaging. The activity of barrel cortex neurons was imaged, and data (fluorescence) was collected using nVista software and a custom written acquisition program in Matlab.

All experiments were conducted according to protocols approved by the Duke University Institutional Animal Care and Use Committee.

D.3 Miniscope Calcium Imaging of Area X in Zebra Finch during Singing

Adult zebra finches (90-110 d) were anesthetized and placed in a custom stereotaxic apparatus. After applying a topical anesthetic (0.25% bupivacaine) and making a vertical incision in the skin over the skull, we made ~ 0.5 mm craniotomies in the skull at a predetermined distance from the bifurcation of a major blood vessel ('Y sinus'; Area X: head angle set to 43 degrees from horizontal plane angle - 5.1 mm anterior, 1.6 mm lateral, 3.0 mm deep). Using a glass pipette attached to a pressure injection system (*Drummond Nanoject II*), AAV2/9-CAG-GCaMP6s (*Penn Vector*) was injected. The skull was then placed over the craniotomy, sealed with bone wax, and the incision site in the skin closed with a tissue adhesive (*VetBond*). 20-25 days after injection, a 1-mm diameter prism gradient-index (GRIN) lens was implanted at the anterior border of Area X. After allowing 7-15 days for tissue clearing and recovery, a baseplate was placed on the bird's skull to hold a miniature microscope nVista for imaging. After another period of recovery ($\sim$3 d), birds were placed in a recording chamber and the miniature microscope was attached to the baseplate for imaging. The activity of Area X neurons was then imaged, and data (fluorescence and song activity) was collected using nVista software and a custom written acquisition program in LabVIEW (sound and frame times).

All experiments were conducted according to protocols approved by the Duke University Institutional Animal Care and Use Committee.

D.4 Two-photon Imaging of Posterior Parietal Cortex in Ferret

A spayed adult female ferret (Mustela putoris furo; housed in pairs in a 12 hr light/ 12 hr dark cycle) was initially anesthetized with intramuscular (IM) injection of ketamine/xylazine (30 $\mathrm{mg/kg}$ of ketamine, 1-2 mg/kg of xylazine). They were then intubated and deep anesthesia was maintained with isoflurane (0.5-2.5% in 100% oxygen) throughout the surgery. Partial oxygen saturation, end-title CO_2, electrocardiogram, and rectal temperature were monitored throughout the procedure. Body temperature was maintained at 38-39°C, and end-title CO2 at 30 to 50 mmHg. Tissue and muscle were first resected to expose the skull surface. A custom-machined stainless steel headpost was secured to the skull with bone screws and C&B metabond (*Parkell*). A small craniotomy was then made above posterior parietal cortex on the lateral gyrus (14 mm anterior to the caudal crest, 2.5 mm lateral to the midline). Through a small incision in the dura, 1 µL of AAV2/9-Syn-GCaMP6s-WPRE-SV40 (*Penn Vector*) was delivered to brain tissue using a pulled glass pipette and a nanoject microinjector (*Drummond Scientific*). Virus was allowed to express for 3 weeks before a second surgery was performed to install a glass cranial window. In this surgery, animals underwent the same pre-surgery and surgery setup as detailed above with the following differences. A custom-machined coverslip holding ring was secured using C&B metabond above the site where the virus injection was made. An eight-millimeter craniotomy was then opened up and dura was resected carefully. A composite glass coverslip (a 1.4 mm thick, 5 mm diameter glass coverslip was glued to a 0.2 mm thick, 8 mm diameter glass coverslip with NOA61 adhesive) (*Warner Instruments*; *Swiftglass*; *Edmond Optics*) was lowered into the coverslip holding ring and secured using flexible retaining ring (*McMaster-Carr*) and super glue. Animals were allowed to recover for a day before imaging sessions. Animals received pain and infection management medication for 7 days post-surgery. All surgical procedures were carried out under aseptic conditions. Calcium imaging data were acquired using a Mai-tai tai sapphire laser (920 nm; *Spectra Physics*) and a resonant-scanning 2-photon microscope (*Neurolabware*) with a 16x water immersion objective (*Nikon*, 0.8 NA, 3 mm working distance).

All animal procedures were performed in compliance with the National Institutes of Health guide for the care and use of laboratory animals (NIH Publications No. 8023, revised 1978) and approved by the Institute of Animal Use and Care of the University of North Carolina at Chapel Hill.

D.5 Computing Setup

All the analyses were run in a workstation with two Intel Xeon E5-2630 v3 CPUs, 2.4GHz, 128GB RAM. The implementation contains options of parallel or serial operations at each step.