Supplementary Note 1. Assembly and alignment of the Bessel module

Detailed documentation on assembling and aligning the mesoscope with conventional Gaussian focus scanning can be found at https://wiki.janelia.org/wiki/display/mesoscopy/Documentation. Below we provide a detailed assembly and alignment procedure for the Bessel module (laser safety guidelines should be followed at all time):

1. After installing the breadboard containing the Bessel module onto the mesoscope, set up all mechanical mounts and cage system (**Supplementary Fig. 1b**) and install all optical components except the axicon, mask, or lenses. Adjust the mirrors so that the excitation light passes through the center of the cage system. Install the axicon, mask, and lenses L1-L3.

2. Adjust the *xy* position of the axicon, so that the excitation light hits its center. The transmitted light should be circularly symmetric when viewed immediately before L1 with a near-infrared detector card (VRC4, Thorlabs).

3. Adjust the *xyz* position of the mask to maximize its transmission power throughput.

4. Place the near-infrared detector card below the focal plane of the mesoscope objective to visualize the transmitted annular illumination. Adjust mirror M2 till the annular illumination is circularly symmetric and centers at the same lateral position as the transmitted Gaussian beam.

5. Translate lens L2 along its optical axis (made easy by the cage system) to control the length of the Bessel focus (upwards for shorter, and downwards for longer foci).

Supplementary Note 2. **Summary of automatic detection of ROIs and extraction of calcium transients**

1. Compute the mean image of the video over time and filter the image with a mean-centered truncated 2D Gaussian template (gaussian width $\sigma = 2.5$ pixels and the truncated size is 10×10 pixels).
2. In the resulting image following step 1, find all pixels that are local maxima within an area of 5×5pixels. All pixels with values above a manually selected threshold will be selected as neuron centers.
3. Create ROI masks for each neuron by creating a circular region (diameter=10 pixels, 20 □m) surrounding the identified neuron centers.
4. Mask all ROIs out from the raw video data $Y \in \mathbb{R}^{d \times T}$ (d is the number of pixels and T is the number of frames) and fit the model $Y = Y_{bg} = UV + b_0 \cdot \mathbf{1}^T$ (note that Y has no neural signals after we mask all ROIs out) by running a truncated-k (k=3 here) singular value decomposition on the masked video data ($U \in \mathbb{R}^{d \times k}, V \in \mathbb{R}^{k \times T}, \boldsymbol{b_0} = \boldsymbol{mean}(\boldsymbol{Y}) \in \mathbb{R}^d$). Both $\hat{U}$ and $\hat{b}_0$ have missing values for pixels inside the ROIs.
5. Run MATLAB built-in function *regionfill* to fill the missing values in the estimated $\hat{U}$ and $\hat{b}_0$, and then reconstruct the background signal $\hat{Y}_{bg} = \hat{U}\hat{V} + \hat{b}_0 \cdot \mathbf{1}^T$, as shown in the "background" panel in **Supplementary Video 8**..
6. Let $Y_{signal} = Y - \hat{Y}_{bg}$ ("Raw-BG" panel in **Supplementary Video 8**). For every detected neuron in step 2, fit a model $Y_i = \boldsymbol{a_i} \cdot \boldsymbol{c_i}^T + a_{i,0} \cdot \mathbf{1}^T$ by requiring $\boldsymbol{a_i}, \boldsymbol{a_{i,0}} \in \mathbb{R}_{\geq \mathbf{0}}^{\boldsymbol{d}}, c_i \in \mathbb{R}^{\boldsymbol{T}}$ and $\boldsymbol{mean}(c_i) = 0$, where Y_i is constructed from Y_{signal} by only selecting pixels in the i-th neuron's ROI. After we process one neuron, we update Y_{signal} by subtracting $\hat{Y}_i = \hat{a}_i \cdot \hat{c}_i^T + \hat{a}_{i,0} \cdot \mathbf{1}^T$. We also fit the estimated $\hat{a}_{i,0}$ with a 2D Gaussian function $f(x,y) = \beta + \alpha \cdot e^{-(x-x_0)^2/(2{\sigma_x}^2)-(y-y_0)^2/(2{\sigma_y}^2)}$ and delete the neurons with R-squared values (variance explained by the model / total variance) smaller than a user-specified value (~0.55); these deleted neurons typically have non-neural shapes or can sometimes correspond to merges of multiple cell shapes.
7. Compute $\Delta F/F$ for all neurons. Calcium dynamics for the i-th neuron without neuropil contamination ($\Delta F + F$) is first reconstructed as $\sum \hat{Y}_i$ plus the sum of $\hat{b}_0$ within the i-th neuron's ROI. We then determine ΔF and F from the reconstructed traces, from which "Denoised" and "Demixed" panels in **Supplementary Video 8** are calculated.