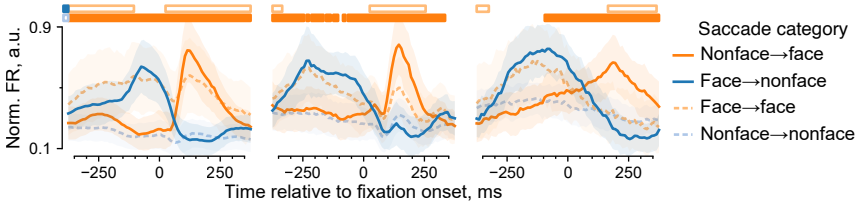
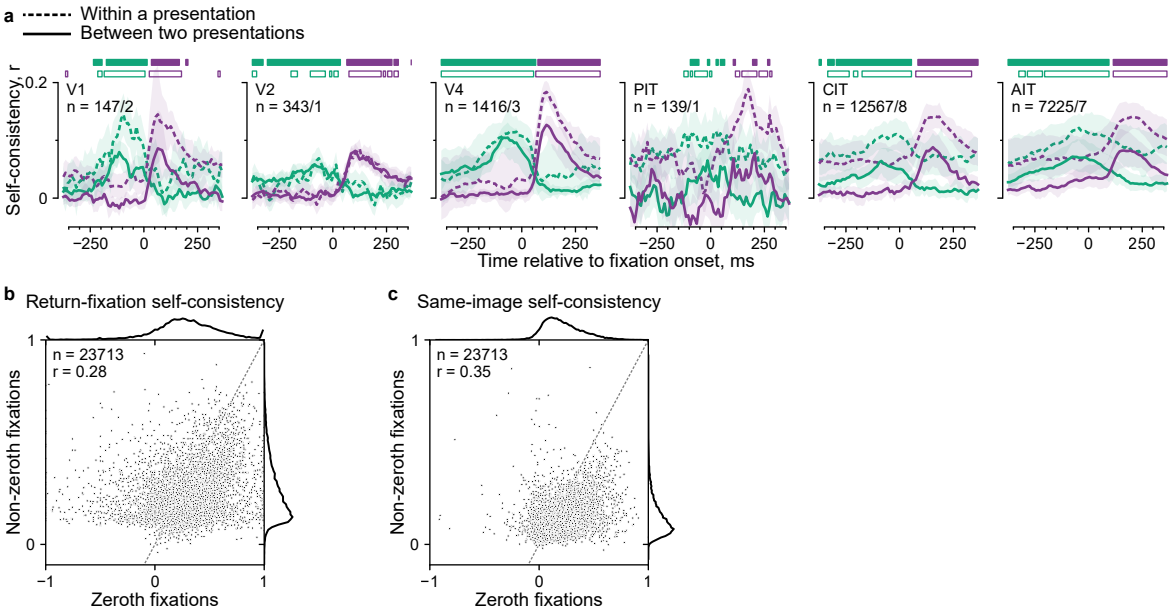
https://doi.org/10.1038/s41593-024-01631-5



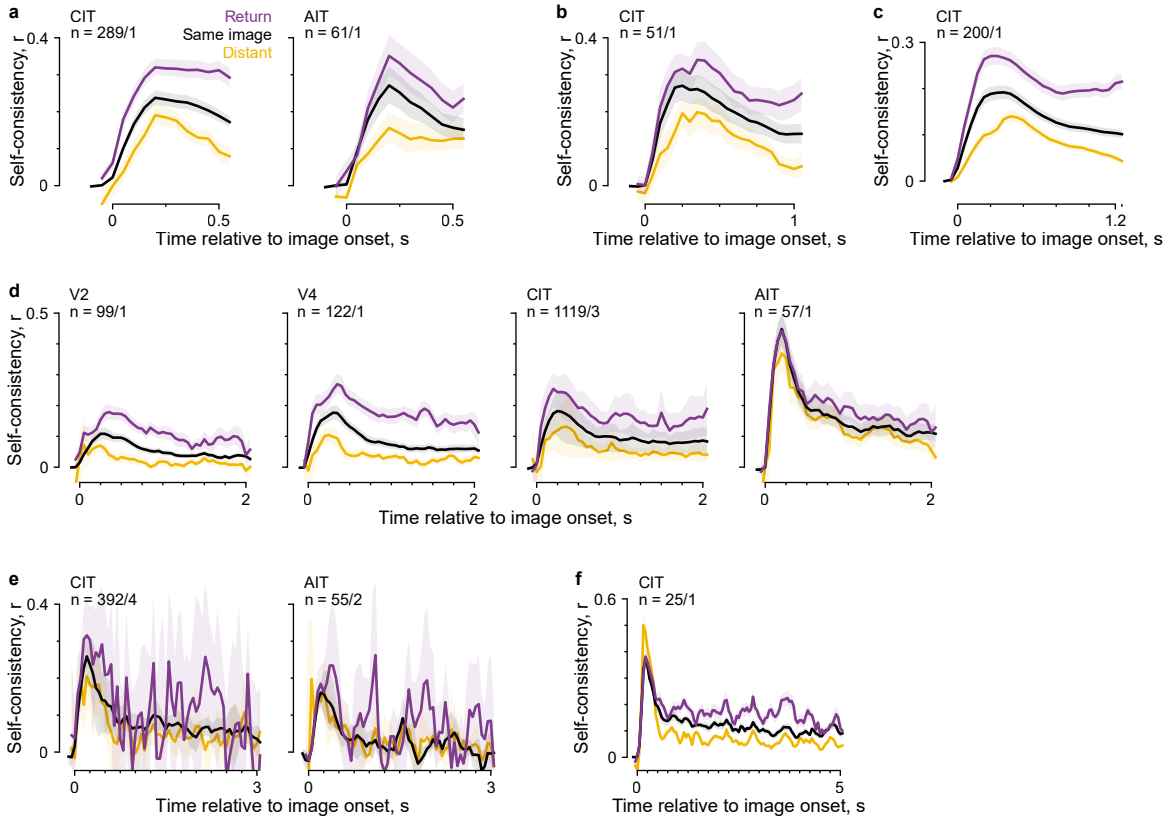
Extended Data Fig. 1 | Responses per category for saccades larger than 4 dva (related to Fig. 2e). Each subplot corresponds to the same monkey and face neurons as in Fig. 2c, e. Lines and shading indicate the median ± median absolute deviation (m.a.d.) across neurons.

https://doi.org/10.1038/s41593-024-01631-5



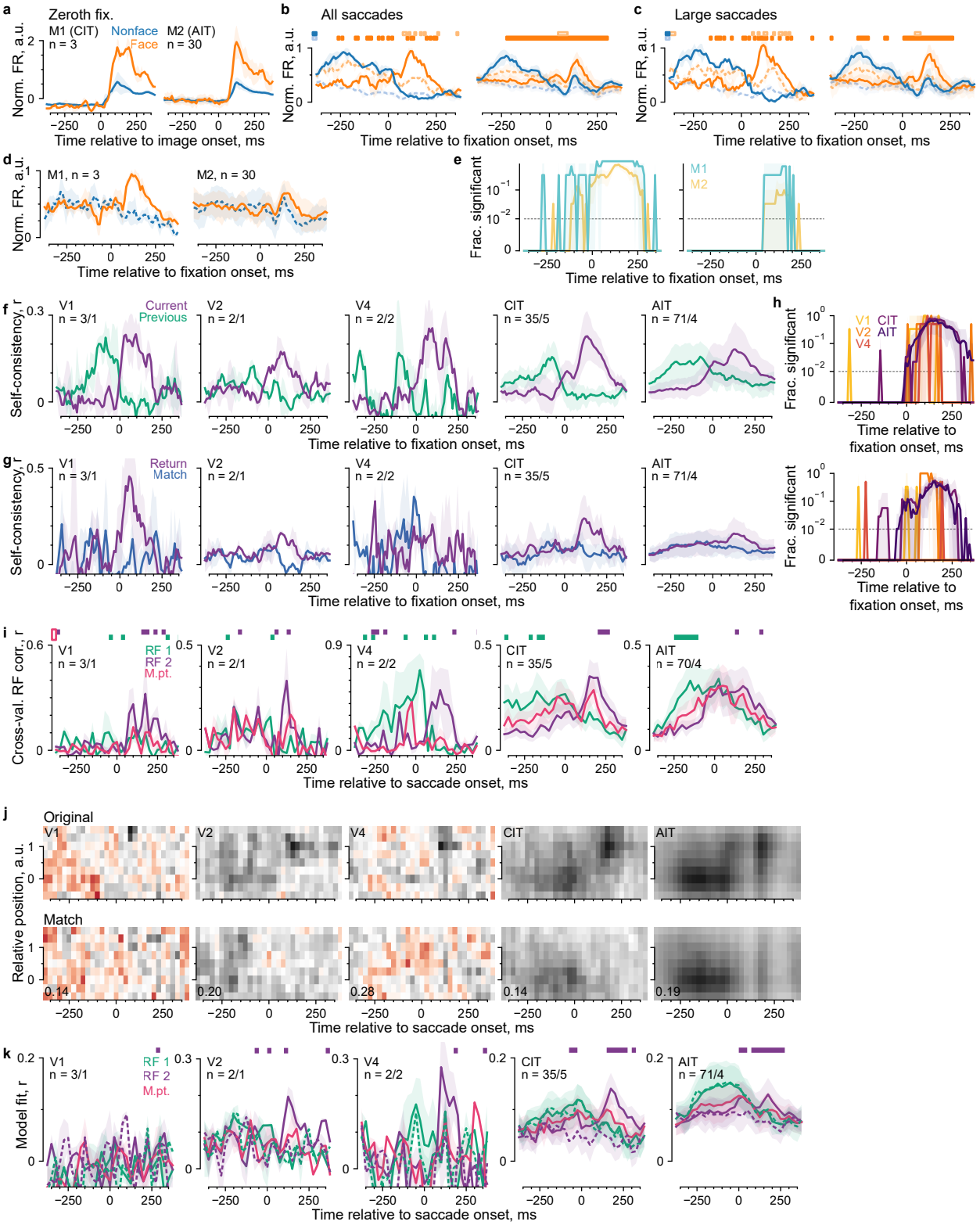
Extended Data Fig. 2 | Self-consistency comparison between within- and across-presentation fixations and between zeroth and non-zeroth fixations (related to Fig. 3). a, Return-fixation self-consistency was calculated separately for fixation pairs within a presentation or between two presentations. Lines and shading indicate the mean and its bootstrap 95%-CI; horizontal bars, time bins with significantly higher self-consistency for current than previous return fixations (purple) or vice versa (green; $p < 0.01$, one-tailed permutation test, FDR-corrected), separately for within-presentation and between-presentation types (open and filled bars, respectively). **b**, **c**, Return-fixation self-consistency, b, and same-image self-consistency, c, separately for zeroth fixations (x-axis) or non-zeroth fixations (y-axis). Each dot indicates a neuron, showing 5,000 examples; dashed line, identity; top and right subplots, marginal distributions across neurons (numbers indicated as n values). The self-consistency was correlated between zeroth and non-zeroth fixations (r = 0.28. 0.35) and slightly lower for non-zeroth fixations (difference = 0.03, 0.08; $p < 10\text{-}156$, ≈ 0, Wilcoxon signed-rank test, uncorrected), consistent with Fig. 4c.

https://doi.org/10.1038/s41593-024-01631-5



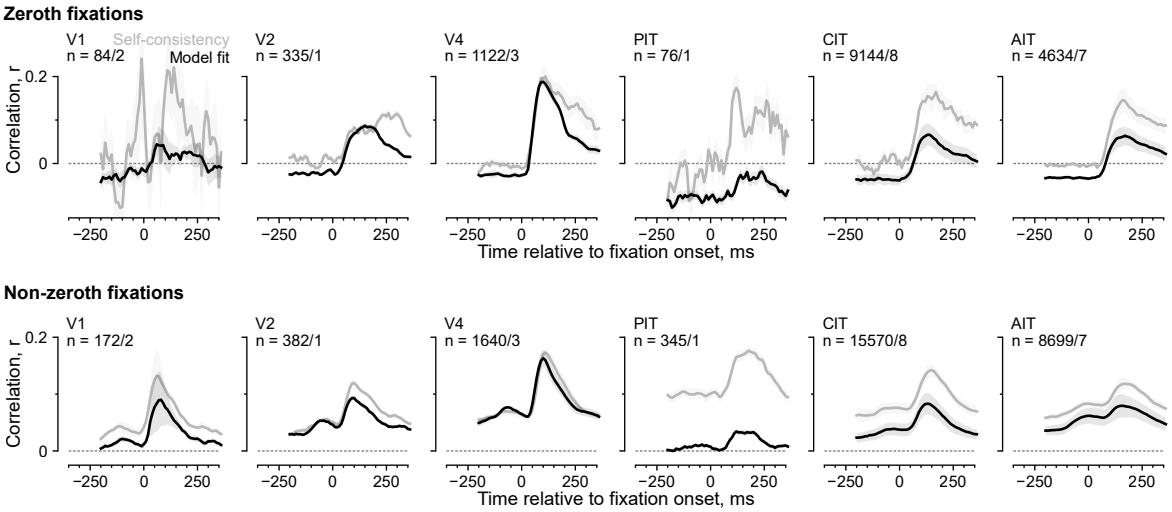
Extended Data Fig. 3 | Mean self-consistency time courses throughout an image presentation for various presentation durations (related to Fig. 4c). Each panel corresponds to a different presentation duration, showing trial durations and regions with at least 25 neurons. Lines and shading indicate the mean and its bootstrap 95%-CI.

https://doi.org/10.1038/s41593-024-01631-5



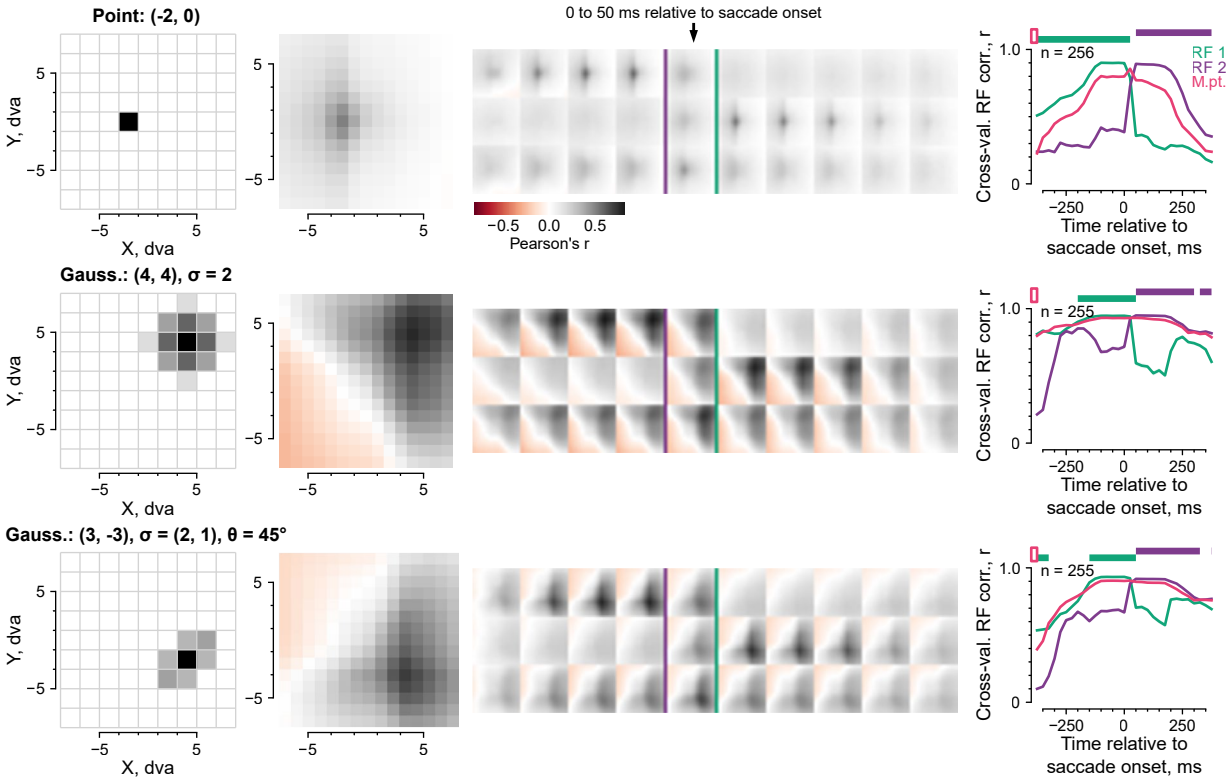
Extended Data Fig. 4 | Results separately for neurons showing negative latencies relative to fixation onsets (related to Figs. 2, 3, 5–7). The neurons are the same as in Fig. 5b, third subplot. Panels a–k respectively correspond to Fig. 2c (top row); 2e; Extended Data Fig. 1; Fig. 5e; f; 3h; 5i; h: 6f; 7c; and 7d. Lines and shading indicate the same center and spread estimates as in the corresponding figure panels.

https://doi.org/10.1038/s41593-024-01631-5



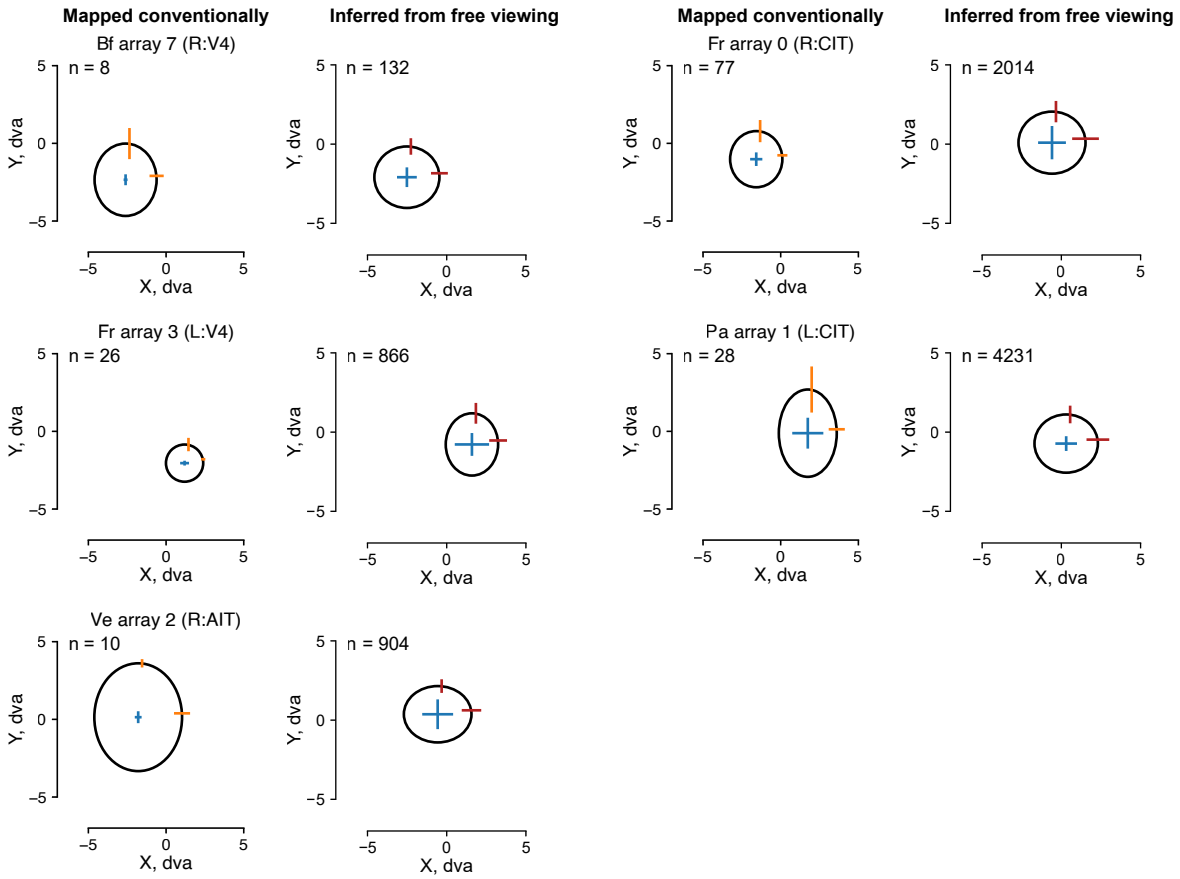
Extended Data Fig. 5 | Time-resolved model fits (related to Fig. 6b). Mean return-fixation self-consistency (gray) and unnormalized model fit (black) time courses for zeroth fixations (first row) and non-zeroth fixations (second row), separately per visual area. Lines and shading indicate the mean ± bootstrap s.e.m.

https://doi.org/10.1038/s41593-024-01631-5



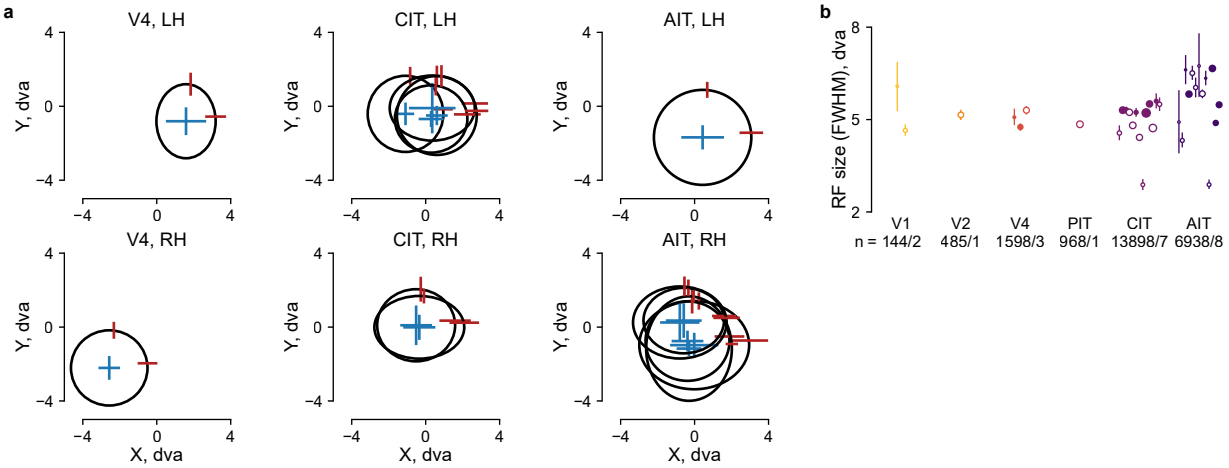
Extended Data Fig. 6 | Validation of model-based RF mapping using simulated responses (related to Fig. 6d–f). We simulated retinotopic, stimulus-selective responses using behavior data from an example session. The simulated responses were a weighted spatial combination of ResNet-50 features (layer3.4.conv1, 256 channels) over 2 × 2 dva image patches centered on the eye position (with no time lag). Each channel was treated as a different neuron. Here, each row corresponds to a different simulated RF. The first column illustrates the image patches and weights used to approximate the RF. The second column shows maps for fixational periods and corresponds to the third column averaged over -150 to 0 ms (relative to saccade onsets) for RF 1 and 50 to 200 ms for RF 2. In the third column, the first two rows in each set of maps correspond to Fig. 6e but are averaged over the 256 channels. The third row corresponds to the midpoint control. The arrow indicates the 0–50 ms time bin relative to saccade onsets; the purple and green lines indicate saccade onsets and typical saccade offsets. The fourth column corresponds to Fig. 6f. Lines and shading indicate the mean ± bootstrap s.e.m.

https://doi.org/10.1038/s41593-024-01631-5



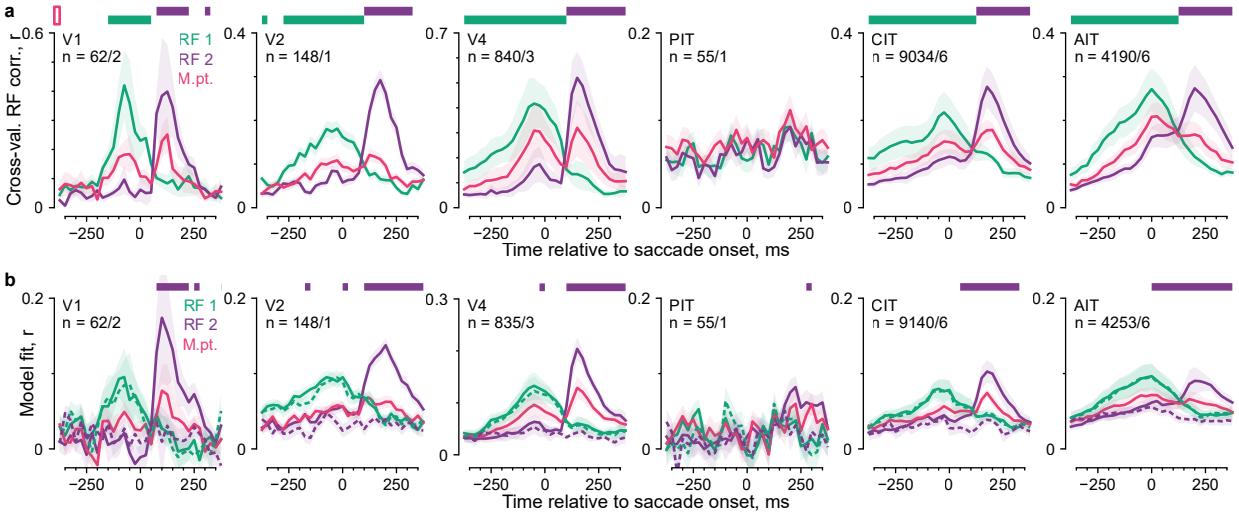
Extended Data Fig. 7 | Model-inferred RFs during free viewing matched conventionally mapped RFs (related to Fig. 6d). RFs recovered from traditional mapping data (firing rates as a function of location in response to small stimuli flashed at various locations) were compared to model-inferred RFs (Fig. 6c, d). In each subplot, the ellipse and error bars indicate the median and m.a.d. across neurons within an array of the center and size of RF Gaussian fits. The five pairs of plots show five example arrays; titles indicate the monkey, hemisphere, and visual area of the arrays. Columns 1 and 3 correspond to conventionally mapped RFs; columns 2 and 4 correspond to model-inferred RFs from free viewing data.

https://doi.org/10.1038/s41593-024-01631-5



Extended Data Fig. 8 | Summary of model-inferred RFs across arrays (related to Fig. 6d). a, Each subplot corresponds to a visual area and hemisphere. Each ellipse indicates the center and size of RF Gaussian fits for a chronic recording array. Error bars show the m.a.d. across neurons within an array. For clarity, only arrays with RF-center m.a.d. ≤ 1.5 dva are shown. **b**, Mean RF size per array. The RF size was quantified analogously to the full-width half-maximum (FWHM) of a circular 2D Gaussian. Specifically, FWHM = 2.355 × sqrt(ab), where a and b are the stdev. along the major and minor axes of the elliptical Gaussian fit. Dots and error bars indicate the mean and its bootstrap 95%-CI per array; filled dots, arrays in panel a; dot size, the number of neurons per array.

https://doi.org/10.1038/s41593-024-01631-5



Extended Data Fig. 9 | Modeling results separately for well-fit neurons (related to Figs. 6, 7). Neurons with normalized model fit of at least 0.5 in Fig. 6b are included. Panels a and b respectively correspond to Figs. 6f and 7d. Lines and shading indicate the mean ± bootstrap s.e.m.