Cell, Volume *177*

Supplemental Information

Evolving Images for Visual Neurons Using a Deep Generative Network Reveals Coding Principles and Neuronal Preferences

Carlos R. Ponce, Will Xiao, Peter F. Schade, Till S. Hartmann, Gabriel Kreiman, and Margaret S. Livingstone

Supplementary Tables for

Evolving images for visual neurons using a deep generative network reveals coding principles and neuronal preferences

Authors: Carlos R. Ponce, Will Xiao, Peter F. Schade, Till S. Hartmann, Gabriel Kreiman, and Margaret S. Livingstone

Table S1. Response rate change of neurons during evolution of synthetic images, averaged across all experiments for each subject, based on fit to exponential function. Related to Figure 5A.

Synthetic images				
Area	Median response change per evolution (spikes/s/gen)	25th, 75th percentile	# experiments with amplitude change different from zero (bootstrap test, 95% CI not including zero)	No. experiments using multiunits vs. single units SU | MU
PIT (monkey Ri)	81.4	67.4, 110.5	15/15	7 | 8
PIT (monkey Gu)	24.6	17.8, 34.2	4/4	0 | 4
P/CIT (monkey Ge)	38.2	32.8, 88.8	8/9	6 | 3
CIT (monkey B3)	47.0	24.8, 77.0	4/4	4 | 0
CIT (monkey Y1)	25.5	16.0, 39.5	6/6	0 | 6
CIT (monkey Jo)	53.7	39.5, 61.2	8/8	6 | 2
V1 (monkey Vi)	84.0	77.4, 91.2	6/6	1 | 5
Natural images				
PIT (monkey Ri)	−3.3	−15.1, 4.8	5/15	
PIT (monkey Gu)	8.7	2.3, 18.2	2/4	
P/CIT (monkey Ge)	-11.4	−18.8, −1.3	5/9	
CIT (monkey B3)	−10.4	−14.7, 4.4	1/4	
CIT (monkey Y1)	−8.8	−12.2, −2.4	0/6	
CIT (monkey Jo)	−1.6	−10.6, −13.4	2/8	
V1 (monkey Vi, gratings)	-32.45	-107.1, 43.1	4/6	

Table S2. Frequency that the closest ImageNet images to the evolved images had the following labels (mean frequency ± se, per bootstrap). Related to Figures 7, S5 and section Predicting neuronal responses to a novel image from its similarity to the evolved stimuli.

	ImageNet labels			
	"macaque"	"monkey"	"face" (human only)	"appliance"
frequency of label in sampled image set	9.97×10^{-4}	1.30×10^{-2}	5.99×10^{-3}	1.10×10^{-2}
Monkey Ri	0.021±0.014 (mean±SE)	0.092±0.030	0.001±0.002	0.010±0.009
Monkey Ge	0.007±0.008	0.033±0.017	0.002±0.005	0.013±0.010
Monkey B3	0.008±0.009	0.048±0.022	0.002±0.005	0.015±0.012
Monkey Gu	0.010±0.010	0.068±0.025	0.000±0.000	0.029±0.016
Monkey Y1	0.002±0.005	0.041±0.017	0.001±0.003	0.041±0.019
		Probability that the values in Ri and Y1 were the same under the null hypothesis: 0.070		Probability that values in Ri and Y1 were the same under the null hypothesis: 0.076

Table S3. Response statistics for fc6-prediction experiments, comparing evolved images and top predictions. Related to Figure 7 and section Predicting neuronal responses to a novel image from its similarity to the evolved stimuli.

				Linear regression between ordinal prediction distance and mean neuronal response	
Subject	Evolved images (mean response in spikes/s, across all experiments)	Top predictions (response to closest fc6 neighbors)	*P* value range across experiments (Wilcoxon rank sum test for equal medians, synthetic vs. natural)	Slope values (spikes/s per prediction group, t-test *P* value)	Range of slope values per experiment per animal, t-test *P* values
Ri	59.4±1.4, N=4	30.8±1.3	4.5×10^{-144} to 3.5×10^{-8}	-21.1 < 1x10-6	-25.3 to -4.9 1×10^{-6} to 7.9×10^{-7}
Gu	38.5±0.8 N = 3	26.3±1.4	5.8×10^{-309} to 5.1×10^{-2}	-5.9 4.3×10^{-114}	-15.1 to 2.2 1×10^{-6} to 8.2×10^{-2}
Y1	38.3±1.1 N = 3	21.4±2.2	8.7×10^{-23} to 1.0×10^{-2}	-5.7 1.4×10^{-16}	-12.0 to -1.8 1.4×10^{-16} to 5.0×10^{-2}
Relationship between distance in fc6 space and mean response *per image* For every site, we computed the fc6 distance between each site's evolved image and a sample of natural images, and compared those distance values with the same sites' mean response to the images. We also measured the trial-by-trial variability of the sites to the images (variability estimated by correlation across a random bipartition)					
	Distance-response correlation (Pearson); Each value corresponds to one experiment	*P*-values (under null hypothesis of zero correlation, Students' T-test)	Trial-by-trial correlation (Pearson)	*P*-value (under null hypothesis of zero correlation, Students' T-test)	
Ri	0.66,0.71, 0.51,0.55	$< 1.2 \times 10^{-3}$	0.83, 0.66, 0.57, 0.56	$\leq 1.4 \times 10^{-3}$	
Gu	-0.27,0.40,0.65	0.16, 1×10^{-4}, 0.03	0.84, 0.72, 0.90	$\leq 8 \times 10^{-6}$	
Y1	0.26, 0.77, 0.06	0.17,8×10^{-7}, 0.75	0.68, 0.66,0.83	$\leq 3 \times 10^{-5}$	

Table S4. (a) Response statistics for synthetic and natural images during evolution experiments (non-parametric), comparing mean and maximum responses reached during the experiment. Related to Figure 5B and section Testing XDREAM using the ground truth of primary visual cortex.

Subject (area)	Mean (spikes/ s, ±sem)			Max (spikes/s, ±se)		
	Synthetic	Reference (natural)	*P* < 0.03; Wilcoxon rank sum test, FDR correction	Synthetic	Reference (natural)	*P* < 0.03; randomization test, with FDR correction
Ri (PIT)	90.5±0.6	45.1±0.6	15 of 15	279.0±8.6	236.6±8.6	9 of 15 Synthetic larger than reference in 9/9 cases
Gu (PIT)	26.6±0.4	21.3±0.4	3 of 4	122.4±4.1	121.4±4.6	0 of 4
Ge (P/CIT)	66.9±0.5	15.1±0.5	8 of 9	220.3±7.1	209.3±8.5	5 of 9 Synthetic > reference in 4/5 cases
B3 (CIT)	45.0±0.4	5.9±0.3	4 of 4	213.1±4.9	169.9±18.2	3 of 4 Synthetic > reference in 3/3 cases
Y1 (CIT)	34.0±0.4	14.5±0.4	6 of 6	156.4±8.9	146.3±6.7	1 of 6 , Synthetic > reference
Jo (CIT)	57.6±0.5	11.0±0.5	8 of 8	180.6±4.9	117.2±7.2	7 of 8, Synthetic > reference in 7/7 cases
Total number of IT experiments: 46						
Vi (V1)	184.5±1.8	114.5±1.8	6 of 6	416.1±14.5	(gratings) 390.3±13.0	P values: 0.003, 0.003, 0.012, 0.050, 0.347 and 0.398

S4 (b). Response statistics for experiments testing previously-evolved synthetic images and ≥ 2,550 natural images.

Mean and maximum rates			
Subject	Natural (mean±sem, max±se, per bootstrap)	Synthetic	*P* value Wilcoxon rank sum test + permutation test (max)
Ri	24.7±0.5 104.2±1.4	72.3±1.9 130.3±5.8	$<1 \times 10^{-6}$ 1.0×10^{-3}
Ge	-8.4, 87.0±3.8	28.0, 83.5±4.4	$<1 \times 10^{-6}$ 1.0×10^{-3}

Table S5: Quantification of the goodness of fit by the substitute network. Related to Figure 7 and last paragraph of Discussion.

Unit Ri-10	Train	Test	Train	Val	Test
Fit to CaffeNet fc6	all natural images	evolved images	50% of natural images	50% of natural images	evolved images
	n=2458	n=244	n=1229	n=1229	n=244
Corr. coef.	0.79	0.68	0.82 ± 0.01	0.74 ± 0.01	0.68 ± 0.02
Mean abs. residual (spks/s)	21.4	55.9	23.3 ± 1.6	23.0 ± 1.6	56.3 ± 2.1
Slope (orthogonal distance regression)	1.34	8.00	1.30 ± 0.03	1.31 ± 0.03	7.52 ± 0.80

Unit Ge-7	Train	Test	Train	Val	Test
Fit to CaffeNet fc6	all natural images	evolved images	50% of natural images	50% of natural images	evolved images
	n=2551	n=179	n=1276	n=1275	n=179
Corr. coef.	0.80	0.18	0.85 ± 0.01	0.68 ± 0.01	0.20 ± 0.15
Mean abs. residual (spks/s)	10.8	24.1	12.6 ± 1.4	12.6 ± 1.3	25.6 ± 2.3
Slope (orthogonal distance regression)	1.31	34.64	1.29 ± 0.04	1.28 ± 0.05	37.70 ± 49.44

Table S6. Comparison of approaches. Related to Figure 2 and last paragraph of Discussion. Firing rate responses of two PIT units to images generated by three alternative methods: 1) real-time genetic algorithm with neurons combined with a deep generative network ('XDREAM'), 2) data-fitted substitute ConvNet combined with backpropagation directly to pixel space, and 3) substitute ConvNet combined with backpropagation to input space of the generative network (Nguyen et al., 2016).

	XDREAM (genetic algorithm + generative network)	**Substitute network optimization** (directly in pixel space, **with** jitter-robustness technique)	**Substitute network optimization** (in input space of generative network)	**Probability of medians coming from the same distribution** (Kruskal-Wallis test, Chi^2, DF)
Multi-unit	Median firing rate ± SE (per bootstrap, N_{boot} = 1000) measured within a 50-200 ms window minus 1-40 ms baseline (N = image presentations)			
Ri-10	225.8±2.3 (N = 1747)	152.2±2.1 (N = 1741)	178.2±1.9 (N = 3504)	< 0.001, 1550, 2
Ri-12	107.7±1.2 (N = 1763)	81.4±1.0 (N = 1741)	93.1±1.3 (N = 3516)	< 0.001, 753, 2
	Pairwise comparisons (*P*-value per Wilcoxon rank sum test)			
	XDREAM vs substitute network optimization in pixel space		**Substitute network optimization in pixel space, without vs with jitter**	**XDREAM vs substitute network optimization** in input space of generative network
Ri-10	225.8±2.3 vs. 152.2±2.1 *P* < 0.001		106.6±2.1 (N = 1774) vs. 152.2±2.1 *P* < 0.001	225.8±2.3 vs. 178.2±1.9 *P* < 0.001
Ri-12	107.7±1.2 vs. 81.4±1.0 *P* < 0.001		60.4±1.4 (N = 1770) vs. 81.4±1.0 *P* < 0.001	107.7±1.2 vs. 93.1±1.3 *P* < 0.001