Neural population control via deep image synthesis

Bashivan et al. 2019

General idea

Align ANN with neuron sites

- Similarity in structure (theoretically)

Generate images

- Synthesized controlled images

Specific behavior modification

- Stretch

- One-hot population control

Current Limitations

Comprehension

- Deep NNs are hard to comprehend

Generalization

- Similarity between train and test data
- Inability to be used on novel inputs

Where this idea comes from?

Similarity between the visually evoked ANN internal layer structures (artificial ‘neural’ representations) and the visually evoked neural representations of V4 (mid-level) and inferior temporal regions (high-level) of of the ventral stream

Controller images

Types of control

Stretch

- Stretch the maximal rate of firing beyond its natural occurring maximal rate
- For 79% of neural sites, the synthesis algorithm successfully found at least one image that it predicted to be at least 10% above the site’s naturally observed maximal firing rate

One-hot control

- High firing rate in one site while the rest are suppressed
- By single synthesized image
- Population of 5-40

Experiment Setup

Recorded sites across V4 cortex in left, right, and left hemisphere of 3 awake macaque monkeys

Implanted microelectrode array immediately anterior to the lunate sulcus and posterior to the inferior occipital sulcus, (targeting central visual representation)

Neural sites: spike responses maintain stability in its image-wise ‘fingerprint’ between the building days (mapping images to build ANN) and testing days (controller images) using image-wise Pearson correlation (>= .8)

Experiment Setup

Raw spike rates of each site were normalized from each recording session

Mean response - normalizer images / SD of responses of normalizer images

Test set: 3d-rendered object at random view + naturalistic image (overlay)

Model Setup

ANN model of ventral stream must contains a set of artificial features that span the same visual encoding space as the brain's population of neurons in the area

Alexnet trained on Imagenet

Images were transformed to be proportional to the retinae image (fisheye transformation)

Mapping: ANN-to-V4 using convolutional function which significantly reduces the neural prediction error

Model Setup cont.

$$\hat{y}_n = \left[\sum \left(W_{\mathrm{s}}^{(n)} \cdot X\right)\right] * W_{\mathrm{d}}^{(n)} + W_{\mathrm{b}}^{(n)} \qquad (1)$$

Learning spatial mask used to estimate receptive fields of each neuron

Regularization: L2 and Laplacian smoothing losses

$$\mathcal{L} = \mathcal{L}_{\mathrm{e}} + \mathcal{L}_{\mathrm{Laplace}} + \mathcal{L}_2$$

Two-fold cross validation 89% variance could be explained

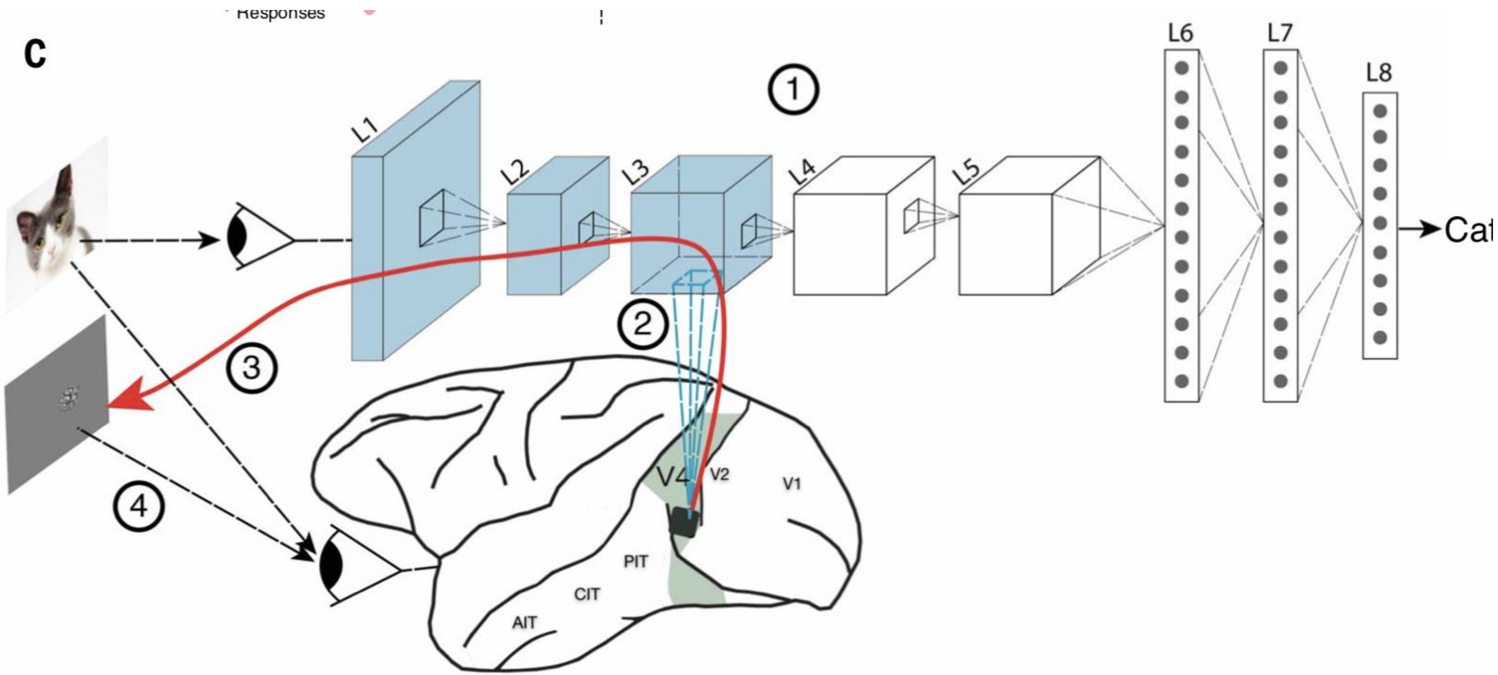
Retinae transformation

$$g(r') = \frac{b}{\sqrt{\pi}} \sum_{k=0}^{r'-1} d_k = \frac{b}{\sqrt{\pi}} \sum_{k=0}^{r'-1} \exp\left(\frac{ar}{2}\right)$$

Synthesized images

(takes account noise in gaze locations)

model-driven image synthesis algorithm

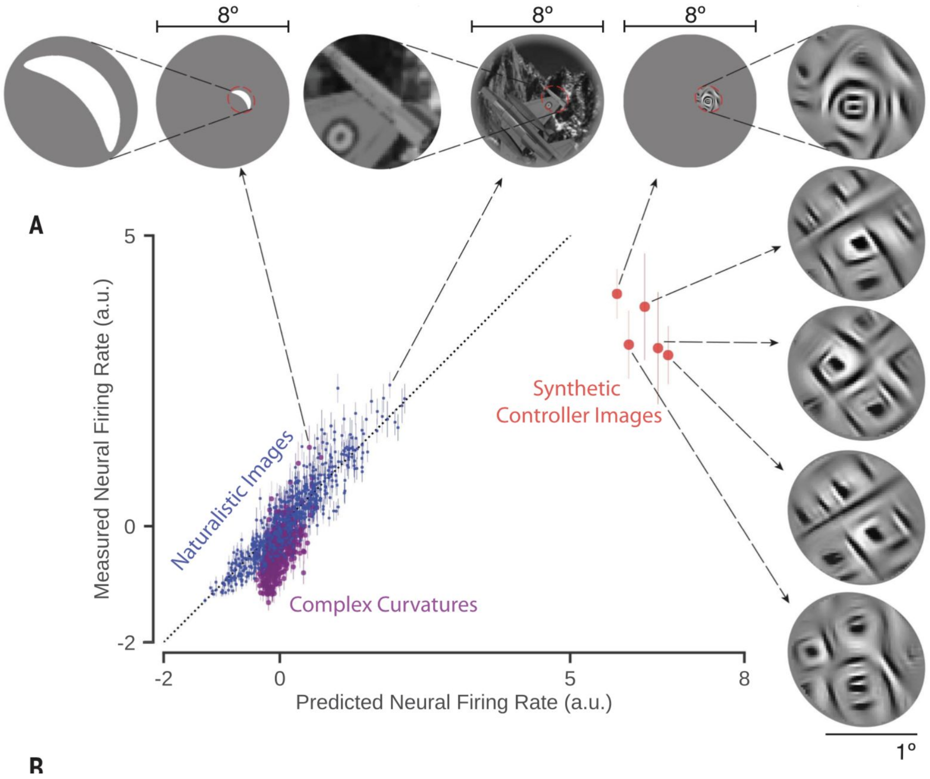


model-driven image synthesis algorithm

Restricted the algortihm to only operate on parts of the image within the cRFs (classical receptive fields) of the neural sites.

Particularly sensitive to angled convex curvature (middle) and concentric circles (lower left side)

Goodness of stretch



Stretch control is nontrivial

Reducing surrounding suppression effect (~14% lower)

Optimized size and location (cropped) (~0.1% lower)

contrast blends

One-hot

Randomly choose a subset of neural sites, generate synthetic contorller images for each site

Predicted softmax of > 0.5 for 77% population (not perfect but not bad) for one-hot population control

Still achieved one-hot (enhance activity in one without interfering the rest)

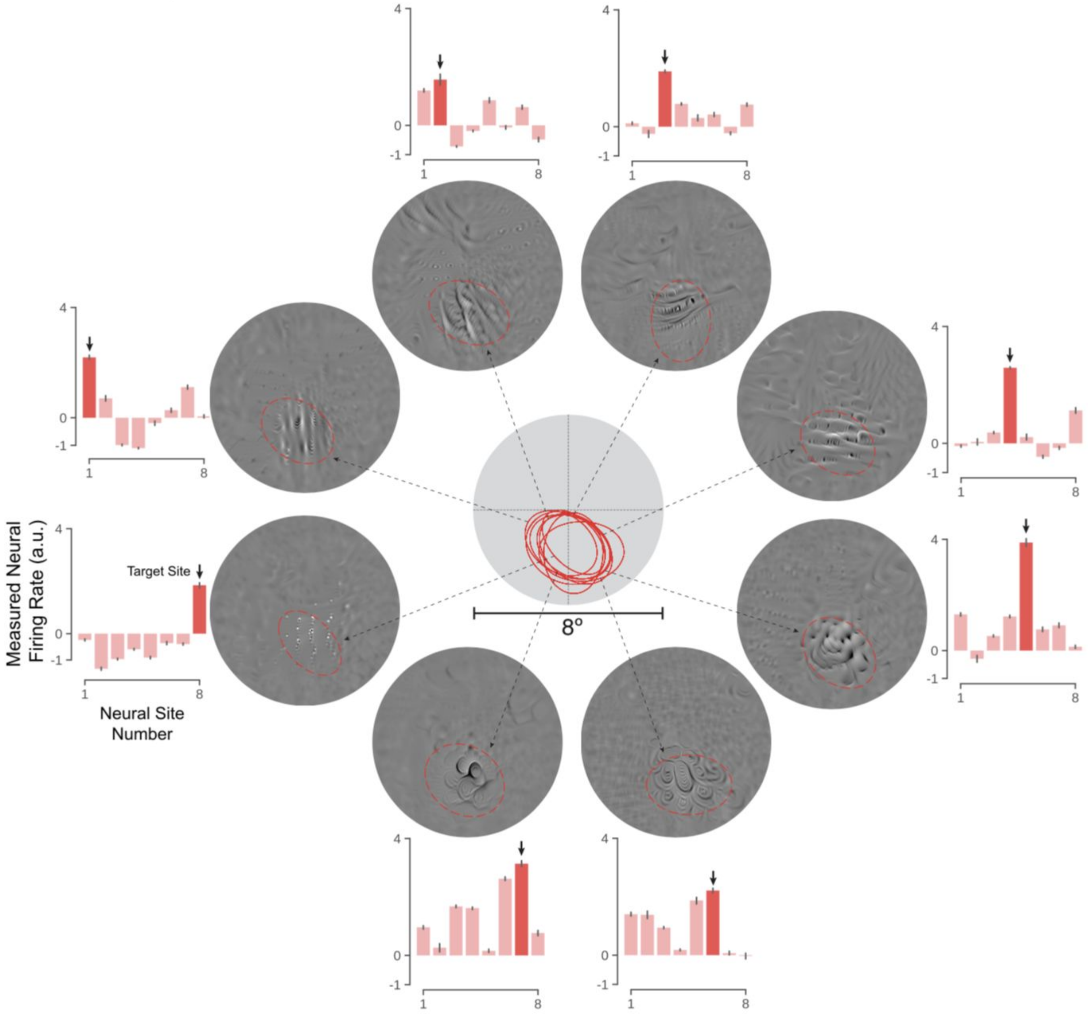
Improved control was statistically significant (comparing to the one-hot control score of the naturalistic images)

One-hot cont.

Possibly due to the nature of non-overlapping cRFs in neural sites

Conducted experiment on shared cRF of all neural sites

Produced very large gain (112%)



Measured Neural Firing Rate (a.u.)
Target Site
Neural Site Number
8°
4
0
-1
1
8

Generalization problem

Synthetic images were statistically less similar to naturalistic images (distances in pixel space, recorded V4 neural population space, and model-predicted V4 population space)

No parameter tuning, model generalizes to predict 54% V4 responses (different from training)

Interesting result

Combining best driver images actually weaker?

From Hubel and Wiesel to ANN

V1, V2, V4, Inferior temporal, …

With ANN, they could be studied in more detail

The ‘how’ question remains unsolved (although solved a bit)