Journal Club

eNeuro

Figure-ground organization in natural scenes: Performance of a recurrent neural model compared with neurons of area V2

Giulia D'Angelo, PhD student

ISTITUTO ITALIANO
DI TECNOLOGIA
EVENT-DRIVEN PERCEPTION
FOR ROBOTICS

Journal Club

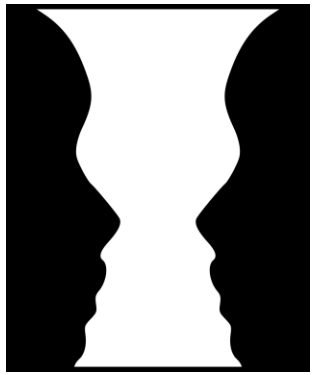
eNeuro

Figure-ground organization in natural scenes: Performance of a recurrent neural model compared with neurons of area V2

Publisher: Society for Neuroscience
Cite Score year 2019: 4.9

Figure-ground organization in natural scenes: Performance of a recurrent neural model compared with neurons of area V2

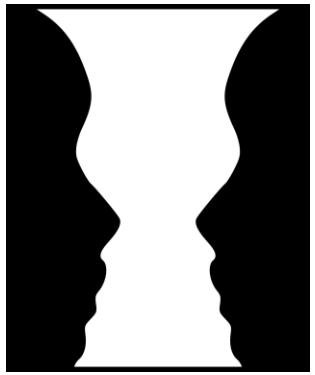
What do they do?



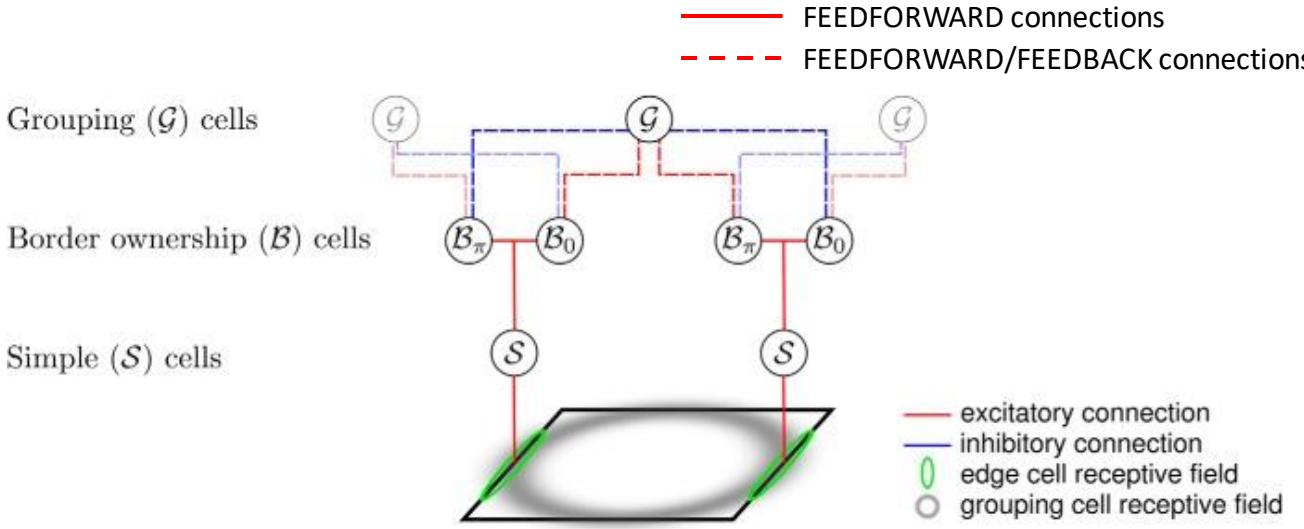
Journal Club

Figure-ground organization in natural scenes: Performance of a recurrent neural model compared with neurons of area V2

What do they do?



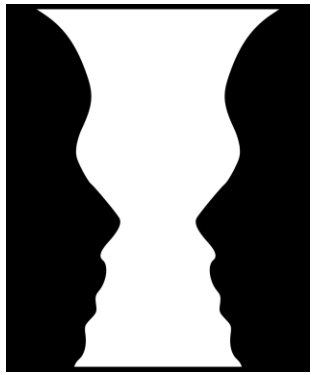
How?



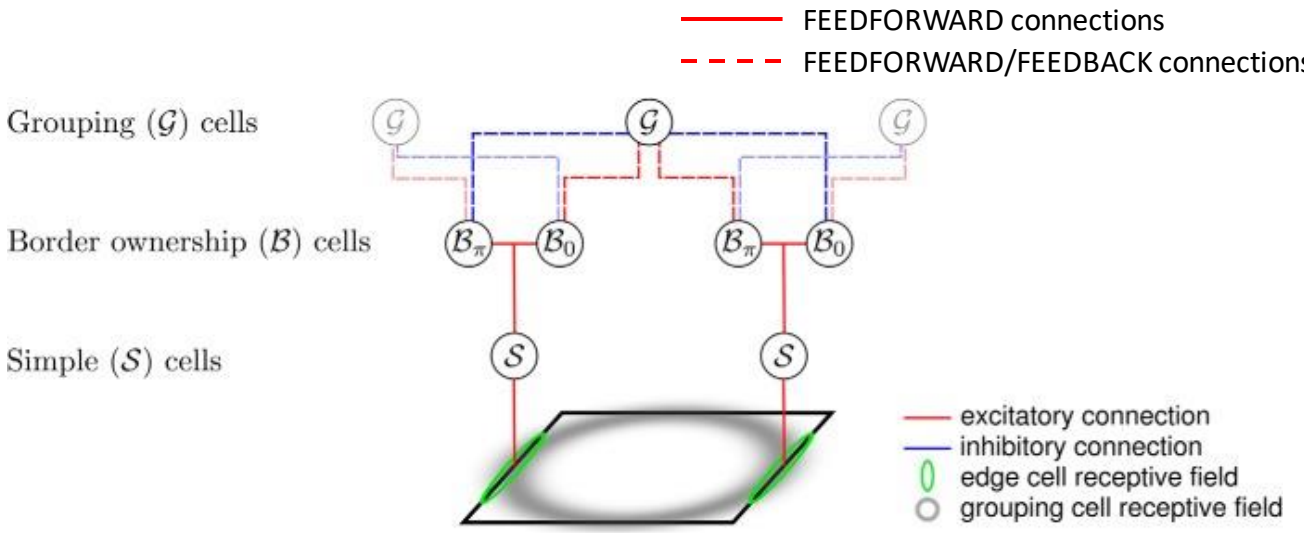
Journal Club

Figure-ground organization in natural scenes: Performance of a recurrent neural model compared with neurons of area V2

What do they do?



How?

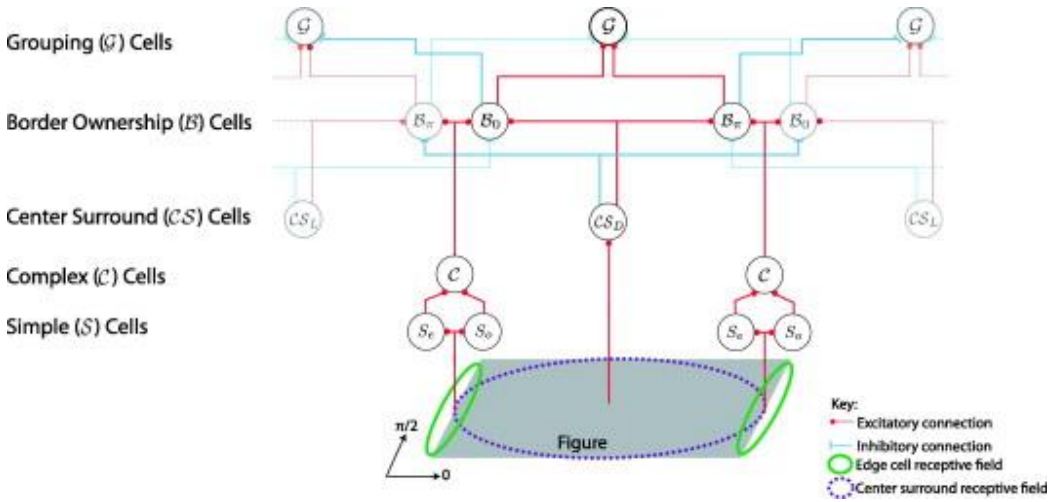


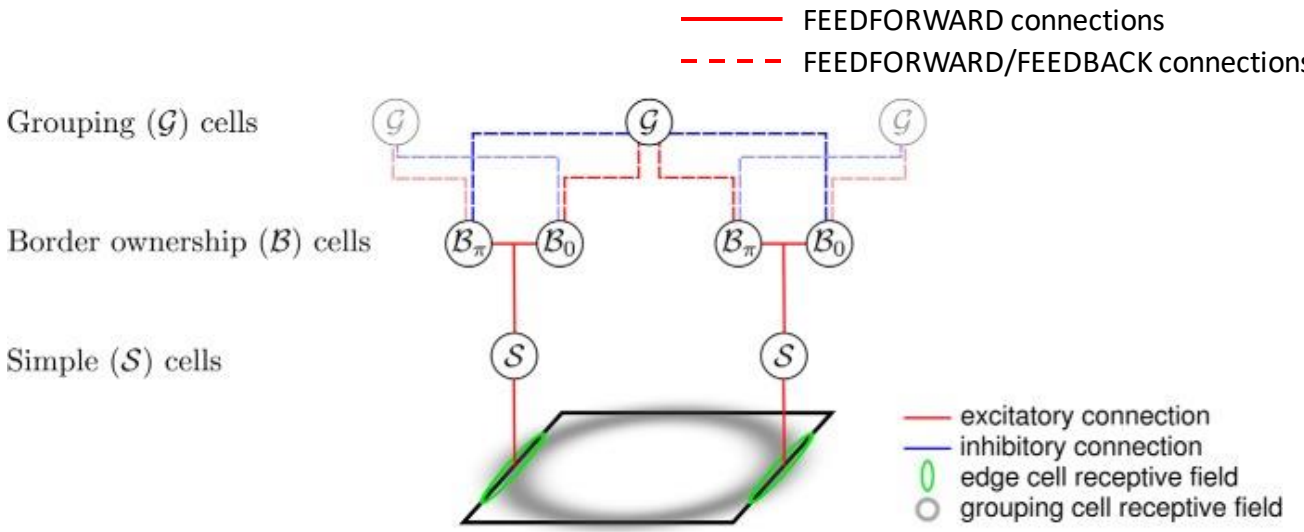
Why?

"some current computer vision approaches are able to achieve better performance than our model based on the evaluation metrics described above, but they require extensive training, i.e., tuning of a large number of parameters using large sets of training data. In contrast, our model is built based on first principles and does not require any specific form of training."

Journal Club

Figure-ground organization in natural scenes: Performance of a recurrent neural model compared with neurons of area V2





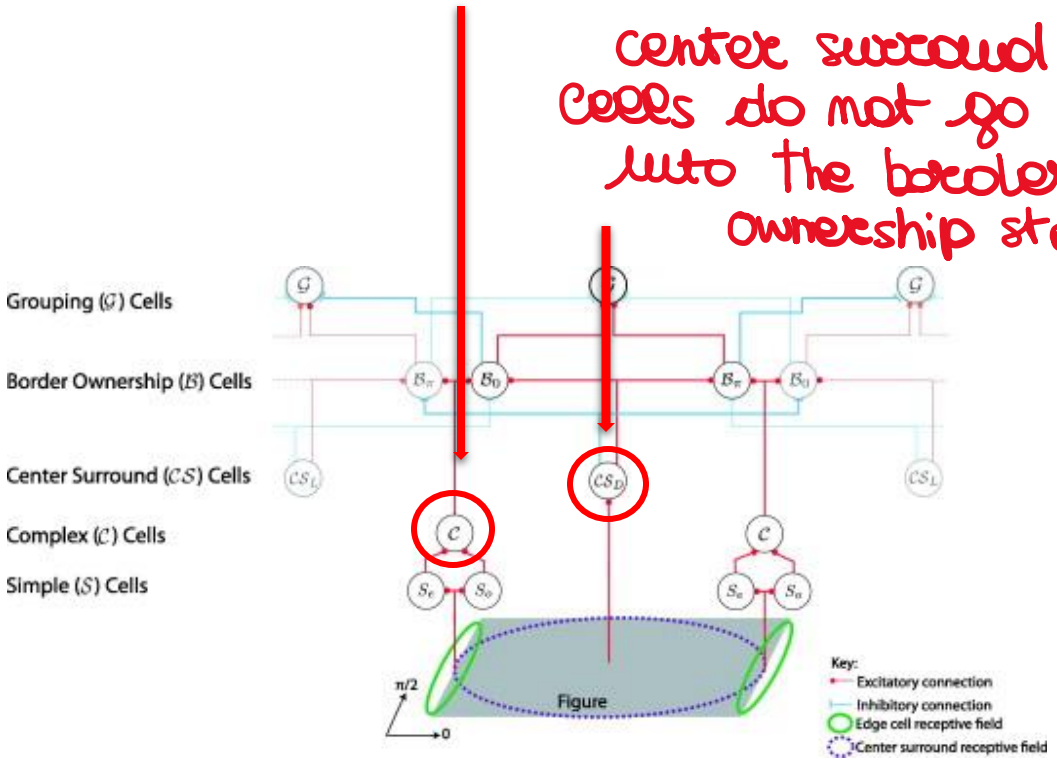
Russell, Alexander F., et al. "A model of proto-object based saliency." *Vision research* 94 (2014): 1-15.

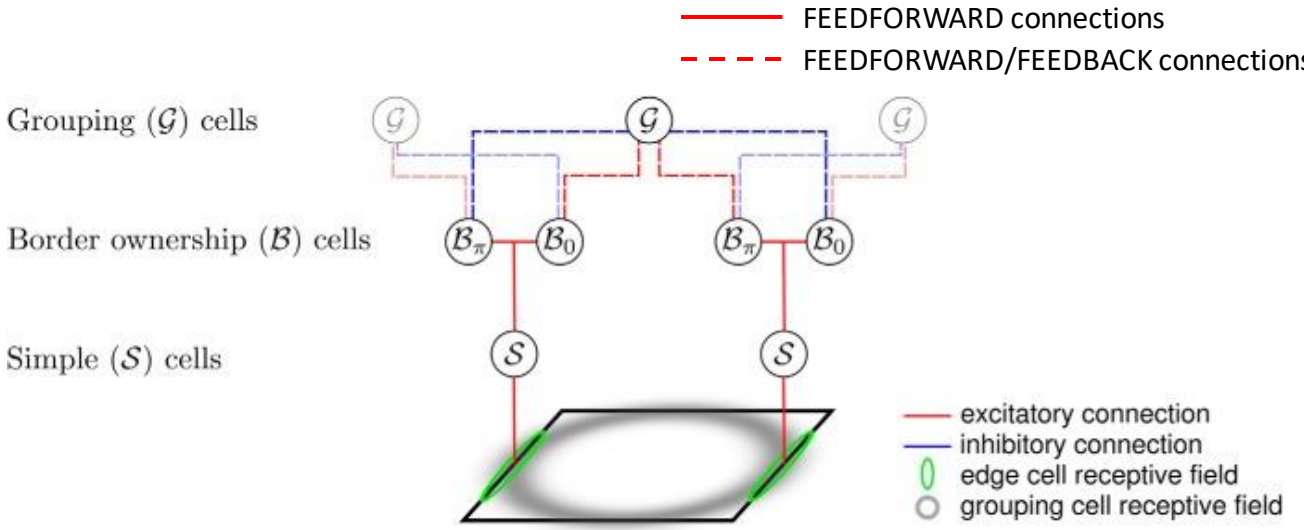
Journal Club

Figure-ground organization in natural scenes: Performance of a recurrent neural model compared with neurons of area V2

NO COMPLEX CELL!
they keep contrast sensitive cells!

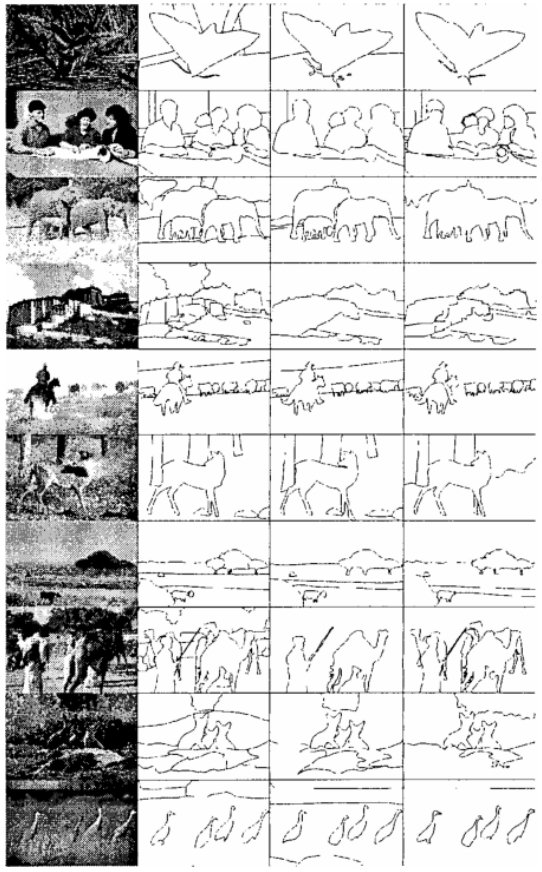
center surround cells do not go into the border ownership stage

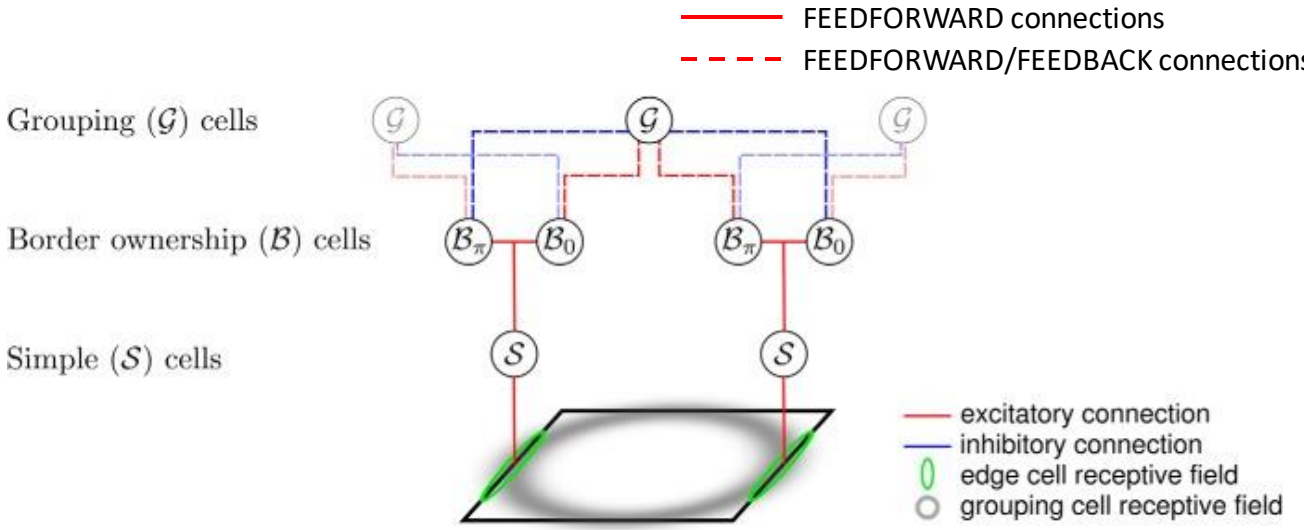




Journal Club

Figure-ground organization in natural scenes: Performance of a recurrent neural model compared with neurons of area V2

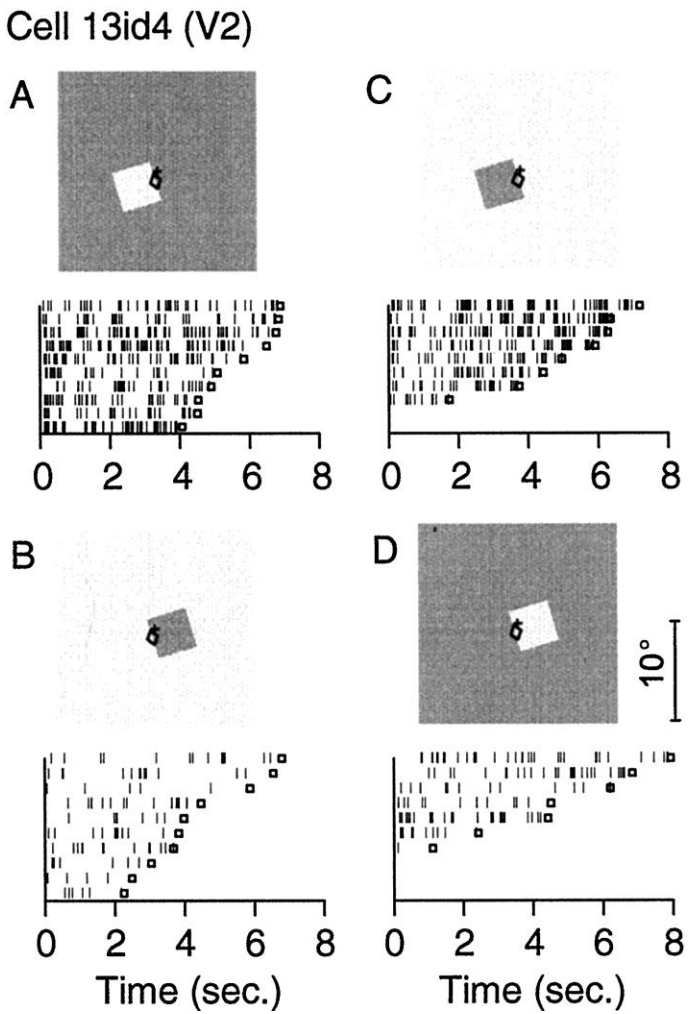


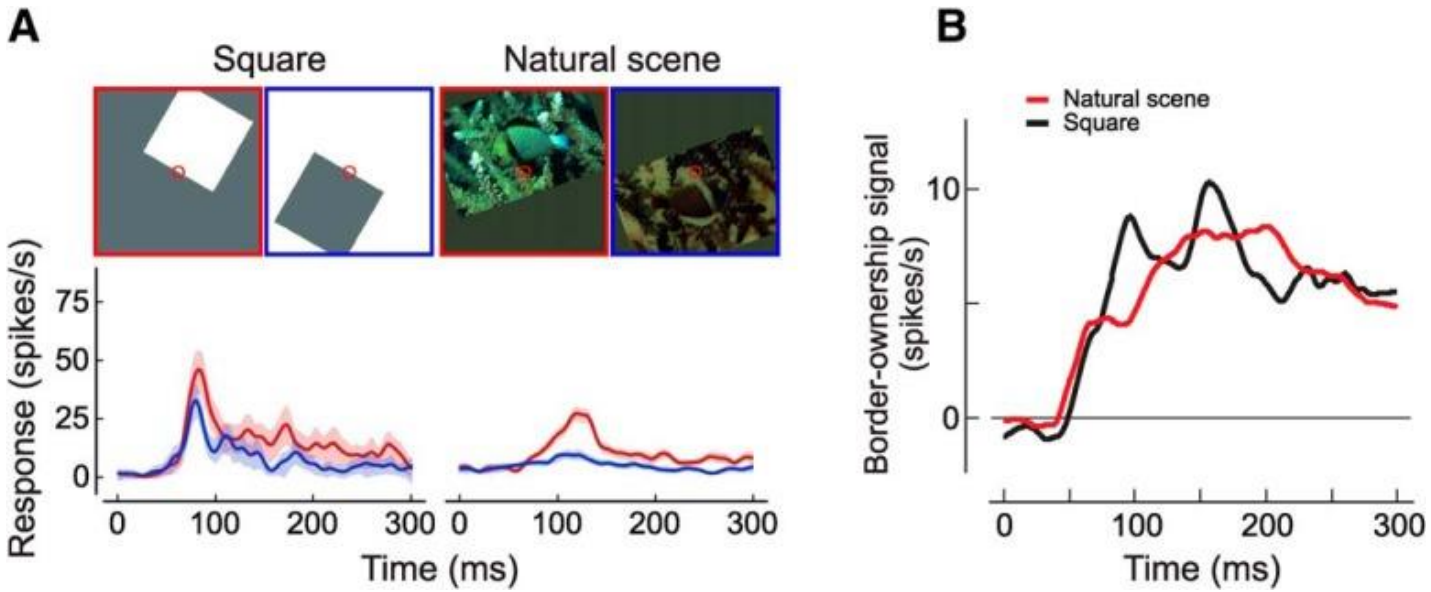


Martin, David, et al. "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics." *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001.* Vol. 2. IEEE, 2001.

Journal Club

Figure-ground organization in natural scenes: Performance of a recurrent neural model compared with neurons of area V2

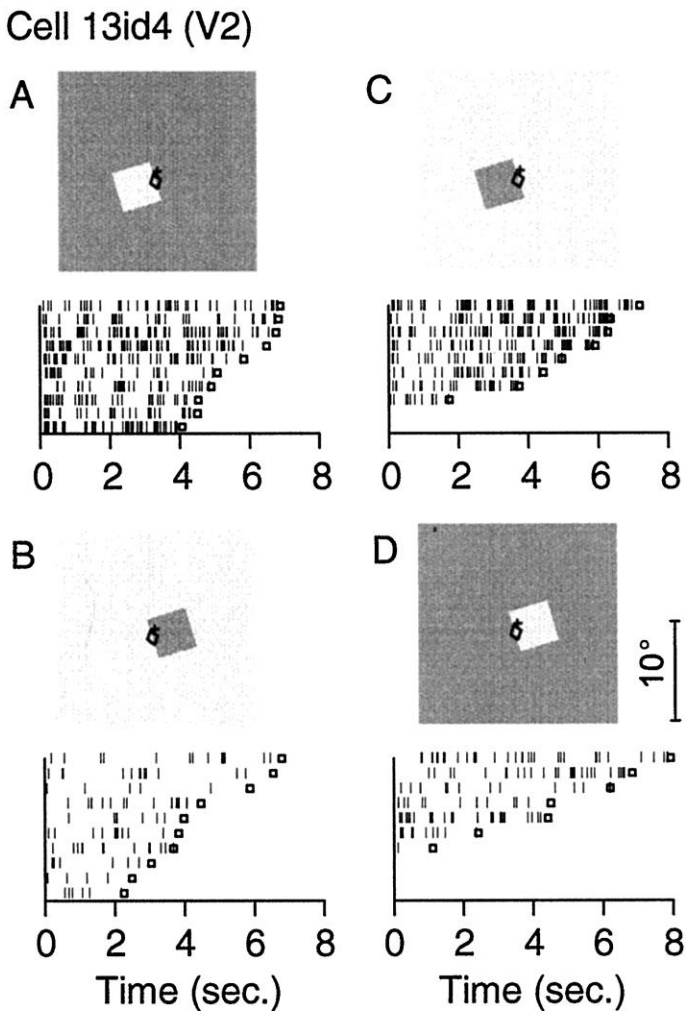


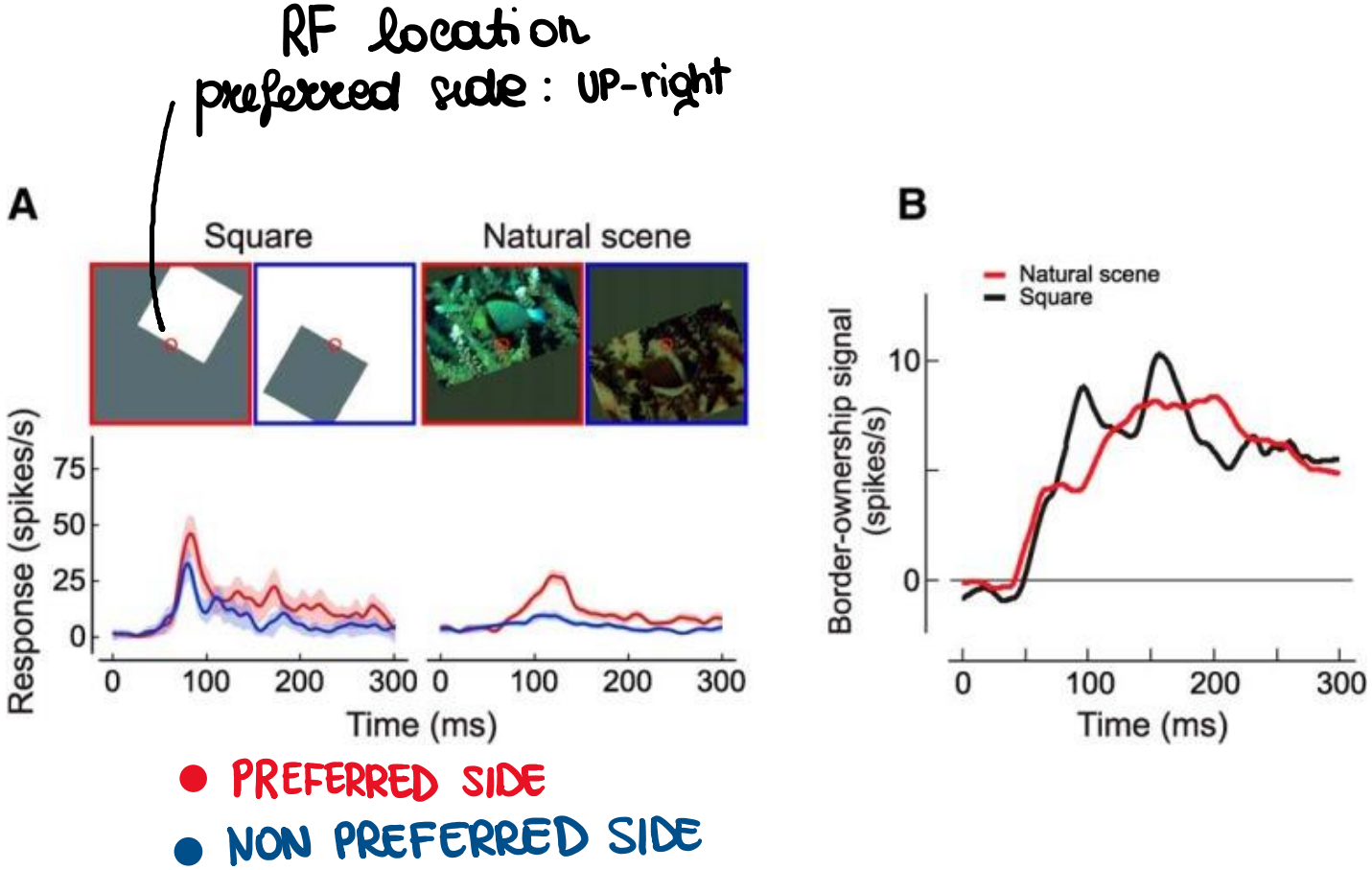


Zhou, Hong, Howard S. Friedman, and Rüdiger Von Der Heydt. "Coding of border ownership in monkey visual cortex." *Journal of Neuroscience* 20.17 (2000): 6594-6611.

Journal Club

Figure-ground organization in natural scenes: Performance of a recurrent neural model compared with neurons of area V2

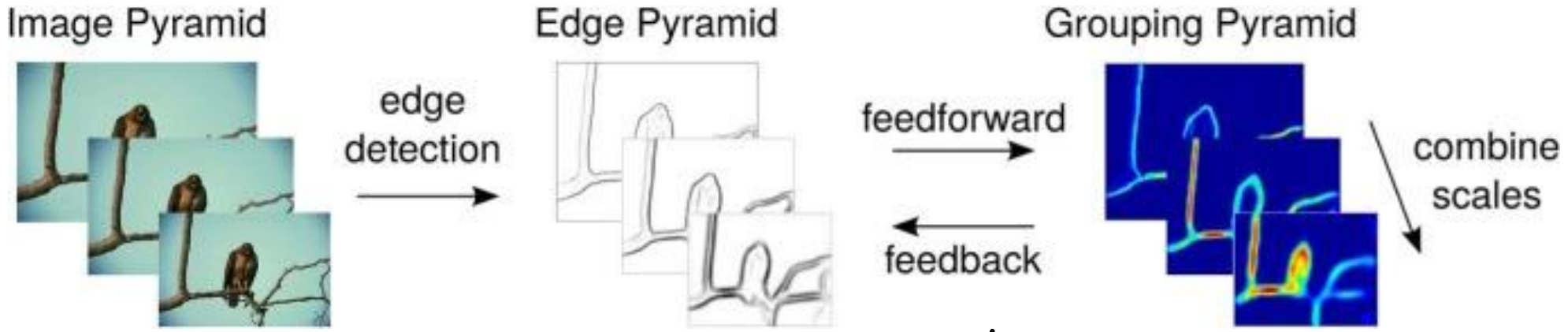




Zhou, Hong, Howard S. Friedman, and Rüdiger Von Der Heydt. "Coding of border ownership in monkey visual cortex." *Journal of Neuroscience* 20.17 (2000): 6594-6611.

Journal Club

Figure-ground organization in natural scenes:
Performance of a recurrent neural model compared with neurons of area V2



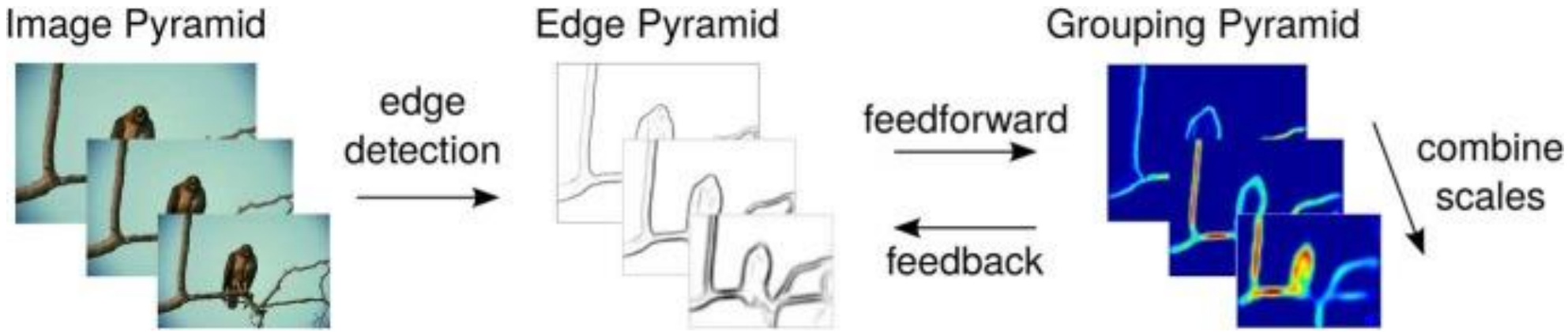
edge extraction CORF (contour operator) DoG

no complex cells & combination in grouping to get contrast invariance

this because obj maintain same contrast polarity along the edges

Journal Club

Figure-ground organization in natural scenes:
Performance of a recurrent neural model compared with neurons of area V2



- different scales of the same computation
- grouping cells using feedback & feedforward to group information coming from the border ownership

Journal Club

Figure-ground organization in natural scenes: Performance of a recurrent neural model compared with neurons of area V2

$$\mathcal{B}^k_{\theta,L}(x,y) = 2\mathcal{S}^k_{\theta,L}(x,y) \times \frac{1}{1+\exp(-(\sum_{j\geq k}\frac{1}{2^{j-k}}v_{\theta+\pi}(x,y)*\mathcal{G}^j_L(x,y) - \sum_{j\geq k}\frac{1}{2^{j-k}}v_\theta(x,y)*\mathcal{G}^j_D(x,y)))}$$

previous model was a complex cell

competition

FEEDBACK

logistic function is to impose the competition between B cells

and $\mathcal{B}^k_{\theta,D}$, the border ownership activity for a dark object on a light background is given by

$$\mathcal{B}^k_{\theta,D}(x,y) = 2\mathcal{S}^k_{\theta,D}(x,y) \times \frac{1}{1+\exp(-(\sum_{j\geq k}\frac{1}{2^{j-k}}v_{\theta+\pi}(x,y)*\mathcal{G}^j_D(x,y) - \sum_{j\geq k}\frac{1}{2^{j-k}}v_\theta(x,y)*\mathcal{G}^j_L(x,y)))}$$

$$\mathcal{G}^k_L(x,y) = \left\lfloor \sum_\theta [\mathcal{B}^k_{\theta,L}(x,y) - \mathcal{B}^k_{\theta+\pi,L}(x,y)] * v_\theta(x,y) \right\rfloor$$

$$\mathcal{G}^k_D(x,y) = \left\lfloor \sum_\theta [\mathcal{B}^k_{\theta,D}(x,y) - \mathcal{B}^k_{\theta+\pi,D}(x,y)] * v_\theta(x,y) \right\rfloor$$

$$\mathcal{G}^k_L(x,y) \leftarrow \begin{cases} \mathcal{G}^k_L(x,y) & \text{if } \mathcal{G}^k_L(x,y) > \mathcal{G}^k_D(x,y) \\ 0 & \text{otherwise} \end{cases}$$

$$\mathcal{G}^k_D(x,y) \leftarrow \begin{cases} \mathcal{G}^k_D(x,y) & \text{if } \mathcal{G}^k_D(x,y) > \mathcal{G}^k_L(x,y) \\ 0 & \text{otherwise} \end{cases}$$

ONLY ONE CELL IS ACTIVE FOR EACH LOCATION

Journal Club

Figure-ground organization in natural scenes: Performance of a recurrent neural model compared with neurons of area V2

$$\mathscr{B}^k_\theta(x,y) = \mathscr{B}^k_{\theta,L}(x,y) + \mathscr{B}^k_{\theta,D}(x,y)$$

dark/light response combined to obtain BOS

$$\mathscr{B}^k_\theta(x,y) - \mathscr{B}^k_{\theta+\pi}(x,y)$$

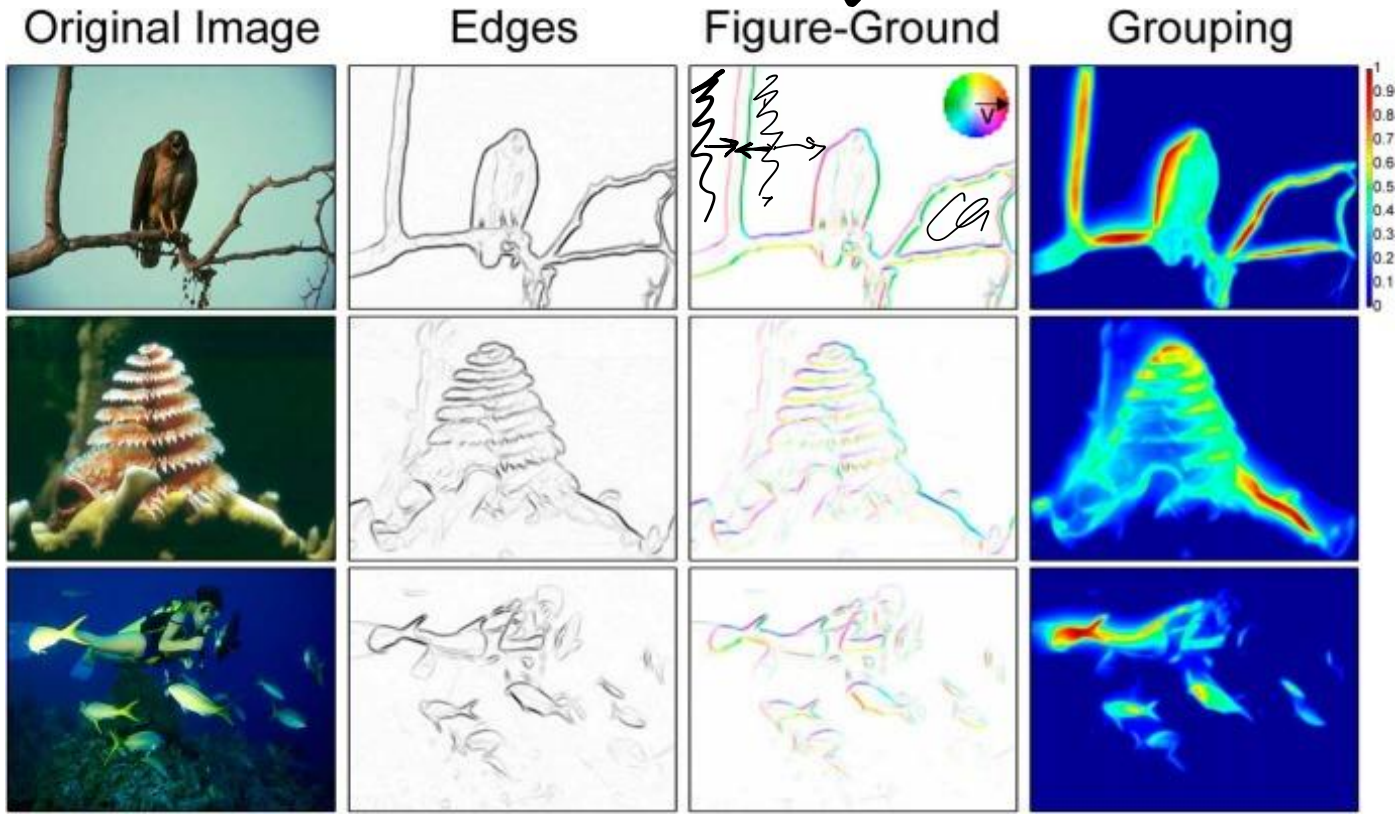
$$\mathscr{G}^k(x,y) = \mathscr{G}^k_L(x,y) + \mathscr{G}^k_D(x,y)$$

" " pooling from grouping cells

Journal Club

Figure-ground organization in natural scenes:
Performance of a recurrent neural model compared with neurons of area V2

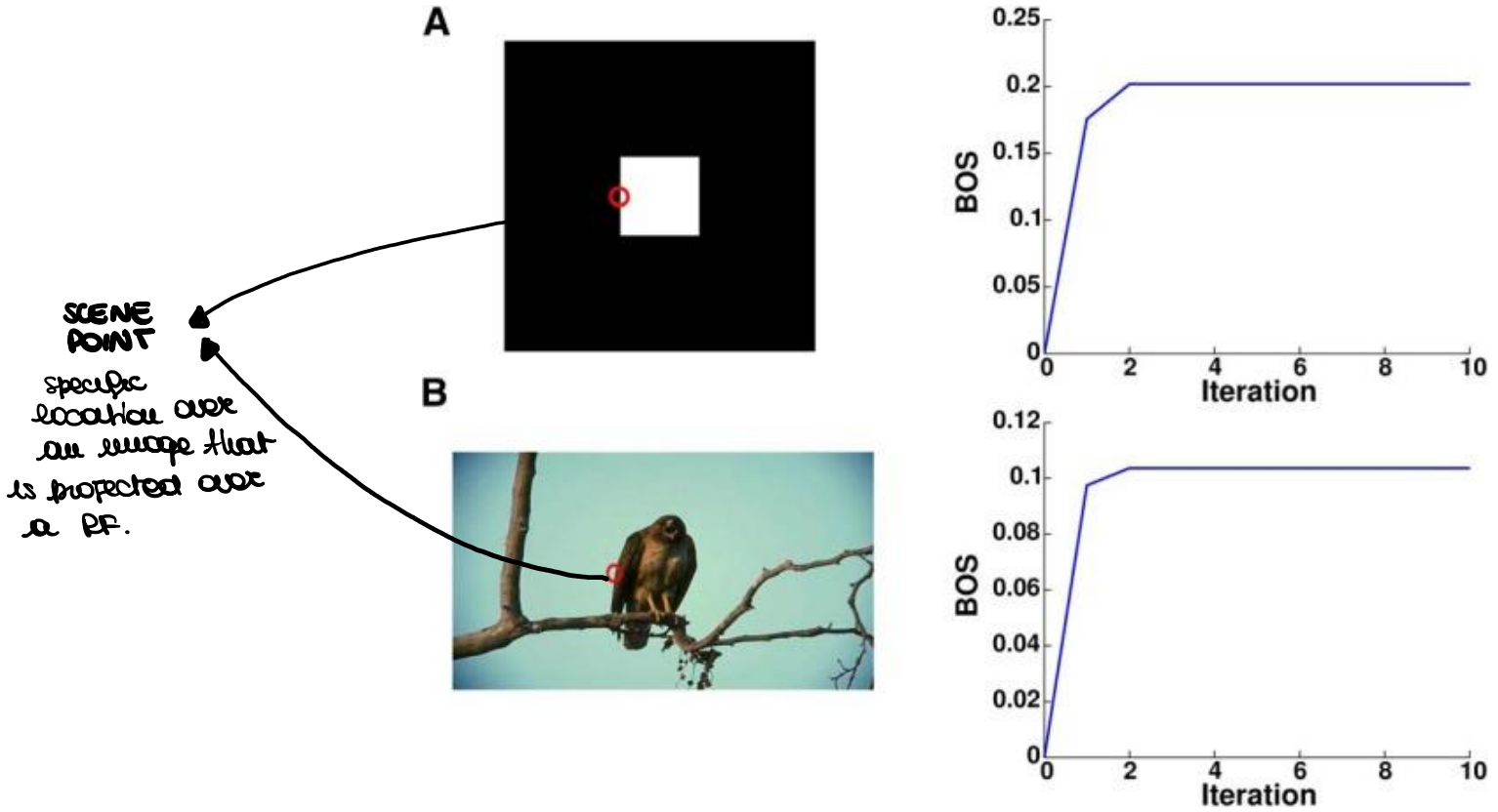
BOS activity summed over ≠ orientations!



THEY RAN THE MODEL FOR ONLY 10 ITERATIONS!
THEY DO NOT NEED TRAINING!

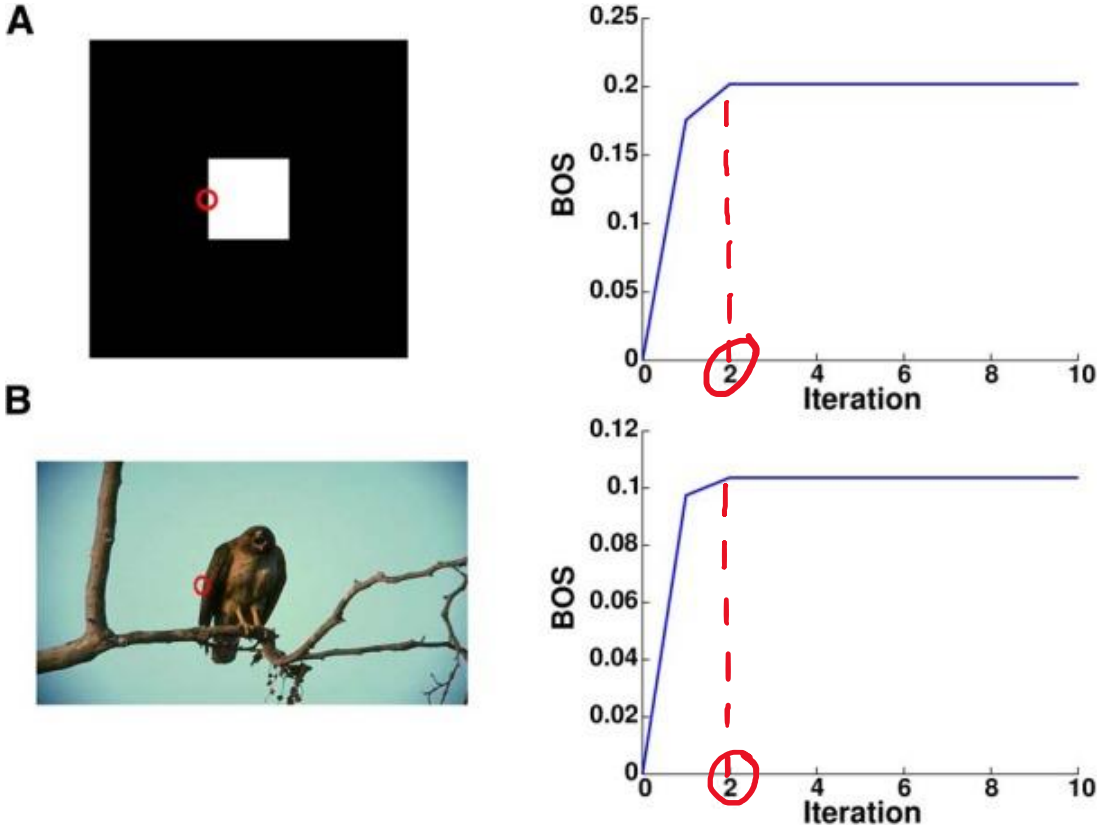
Journal Club

Figure-ground organization in natural scenes: Performance of a recurrent neural model compared with neurons of area V2



Journal Club

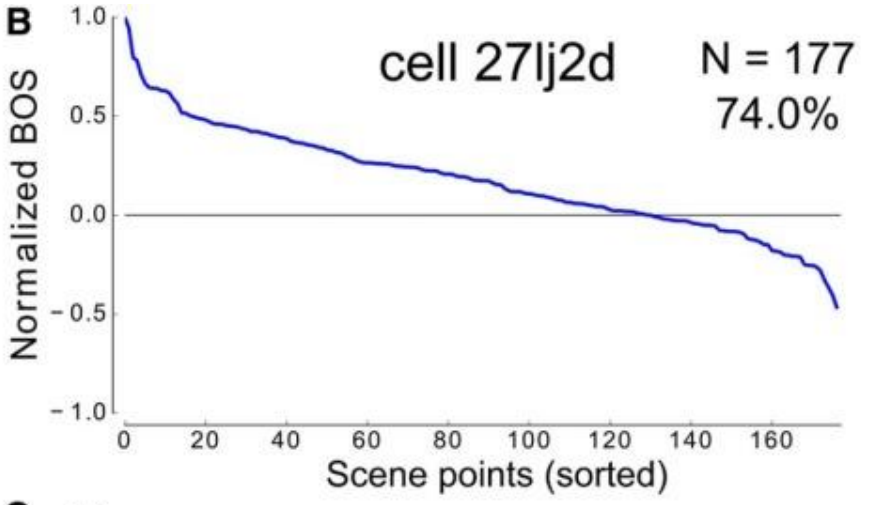
Figure-ground organization in natural scenes: Performance of a recurrent neural model compared with neurons of area V2

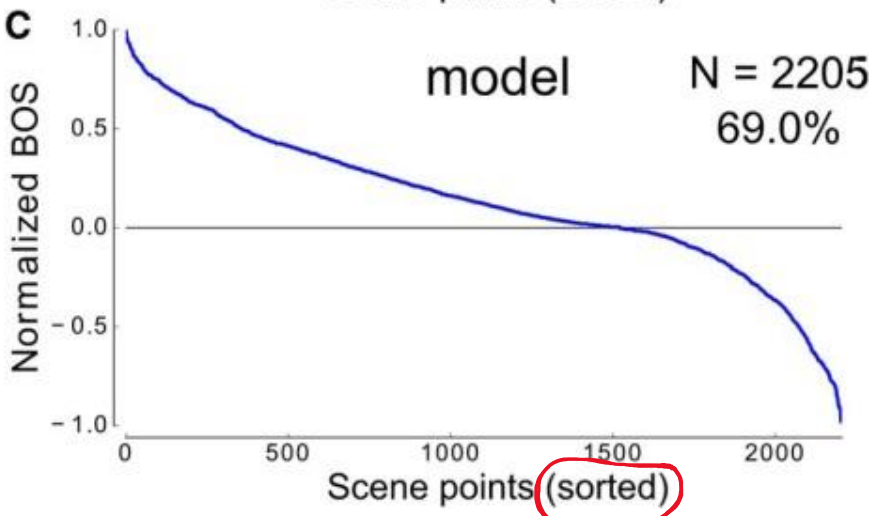


away with 2 iterations!

Journal Club

Figure-ground organization in natural scenes: Performance of a recurrent neural model compared with neurons of area V2

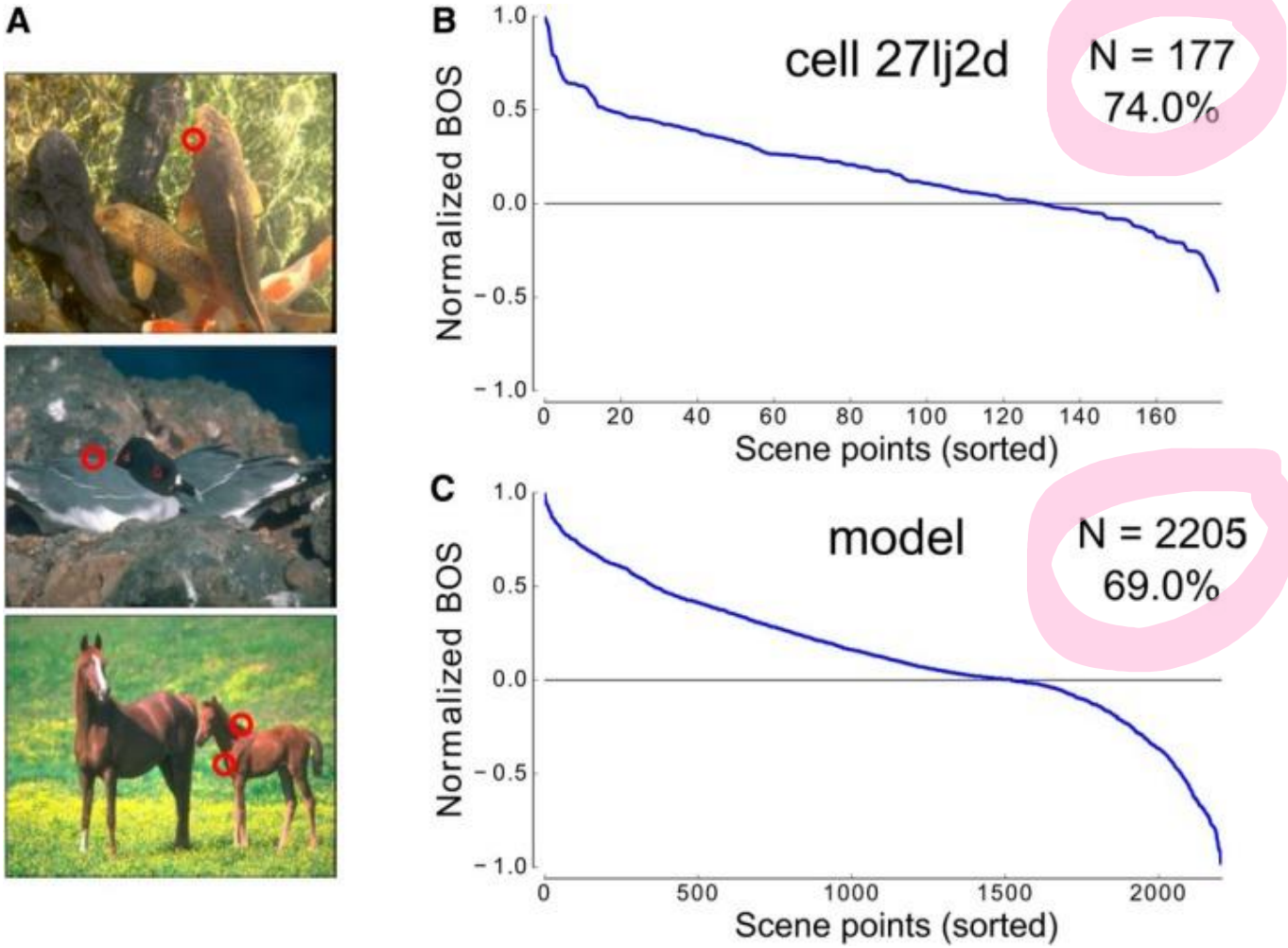




sorted by BOS strength in descending order

Journal Club

Figure-ground organization in natural scenes:
Performance of a recurrent neural model compared with neurons of area V2



POWER OF THE BRAIN!

average values over the population

Table 1.

Statistical analysis

Line	Data structure	Type of test	Power
a	Approximately normal	Bootstrap	$p = 0.11$
b	Approximately normal	Equivalence test	$p = 0.03$
c	Normal	Significance of correlation coefficient	$p < 0.5$

Table 2.

F-scores

Contour-detection results on the BSDS-500 dataset

	Contour		
	ODS	OIS	AP
Human	0.80	0.80	-
Our approach	0.64	0.65	0.51
gPb-owt-ucm	0.73	**0.76**	0.73
SE	0.73	0.75	**0.77**
SRF	0.73	0.74	0.76

entire dataset
per image
average precision

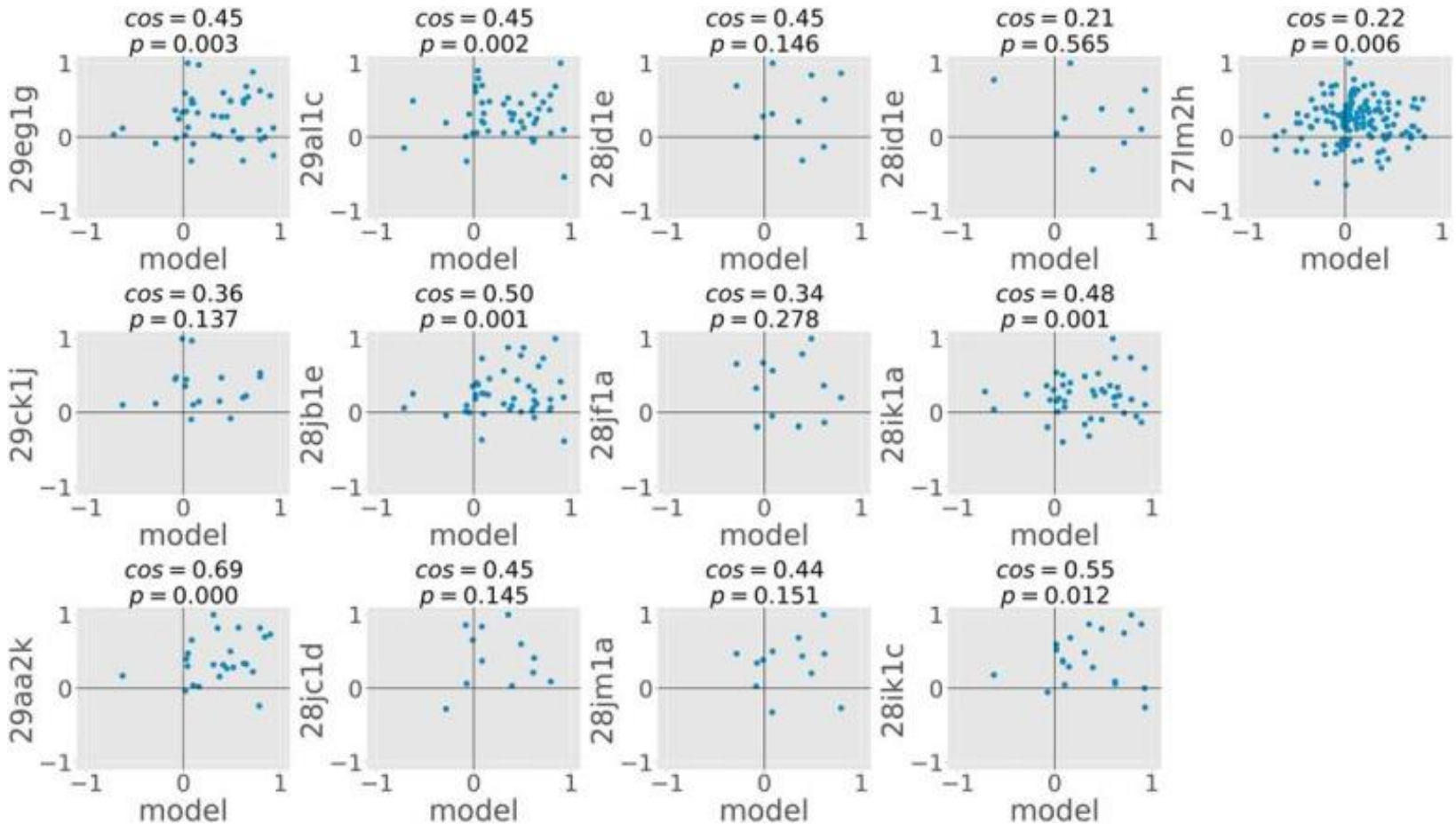
Table 3.

Figure-ground assignment results

	Figure-ground
	Mean accuracy
Human	83.9%
Our approach	71.5%
SRF	**74.7%**
Global-CRF	68.9%
2.1D-CRF	69.1%

Journal Club

Figure-ground organization in natural scenes: Performance of a recurrent neural model compared with neurons of area V2

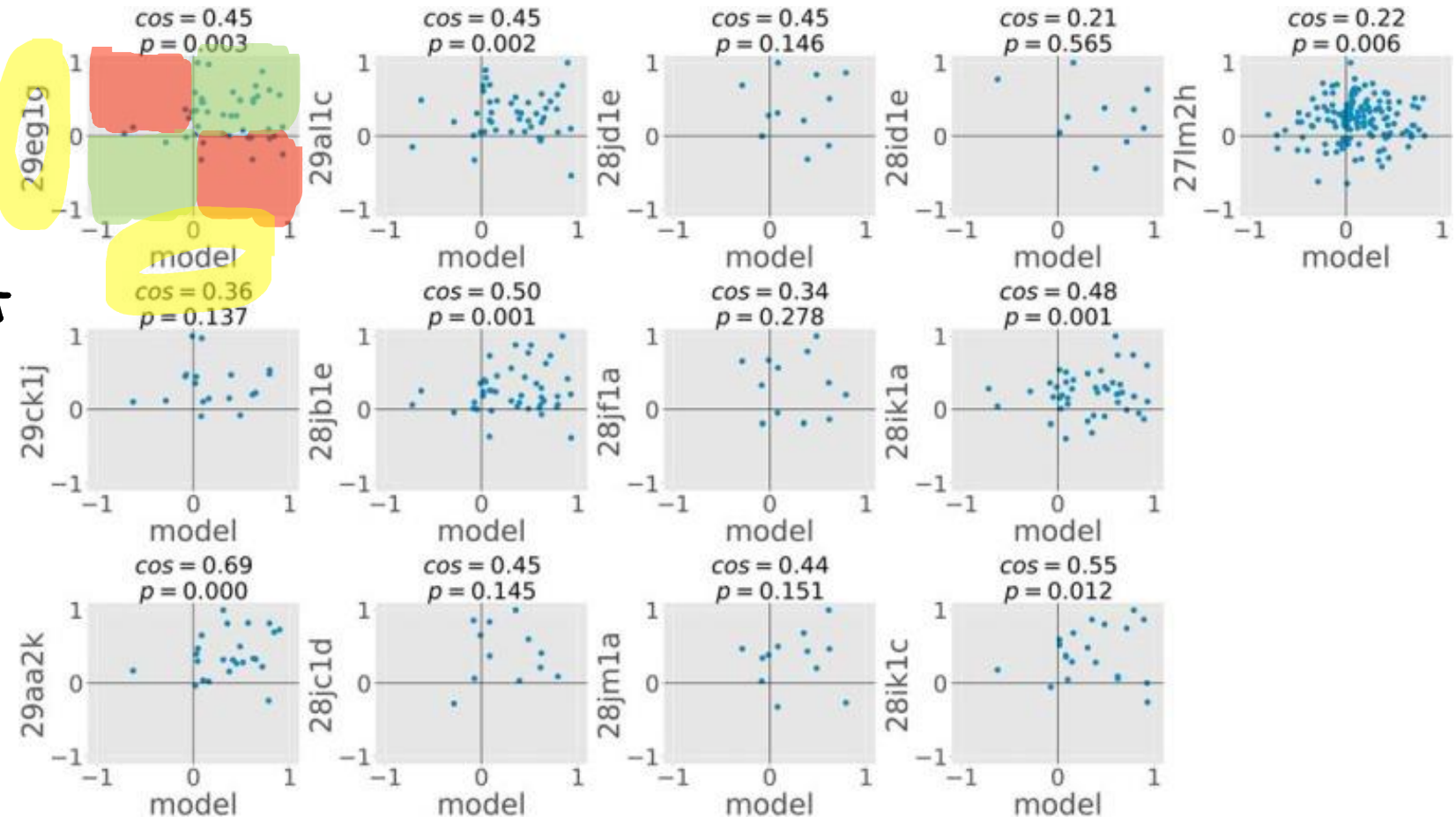


Journal Club

Figure-ground organization in natural scenes: Performance of a recurrent neural model compared with neurons of area V2

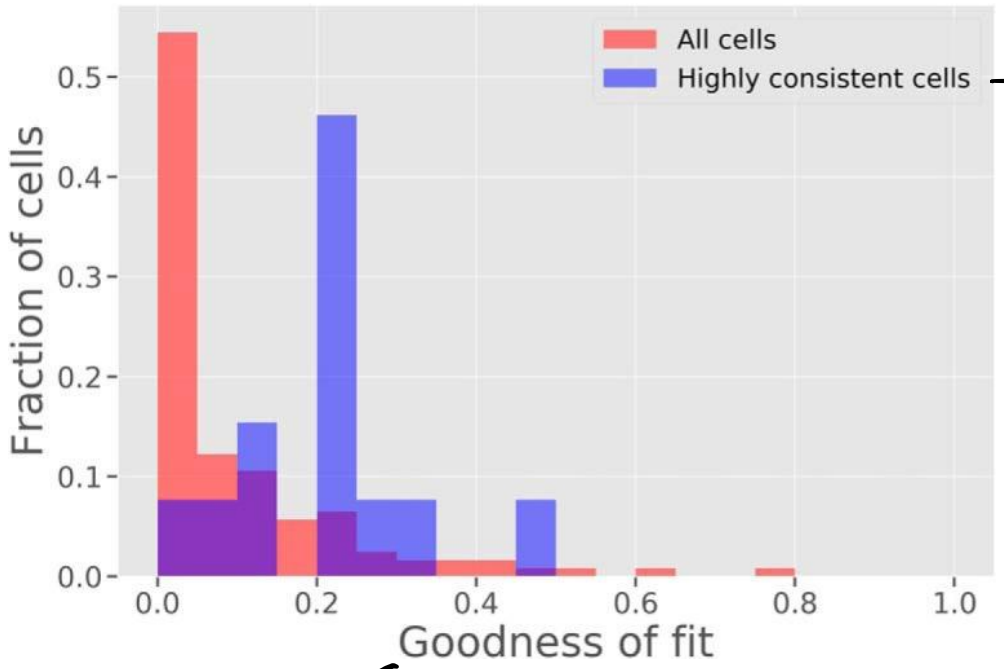
AGREEMENT

DISAGREEMENT



Journal Club

Figure-ground organization in natural scenes: Performance of a recurrent neural model compared with neurons of area V2



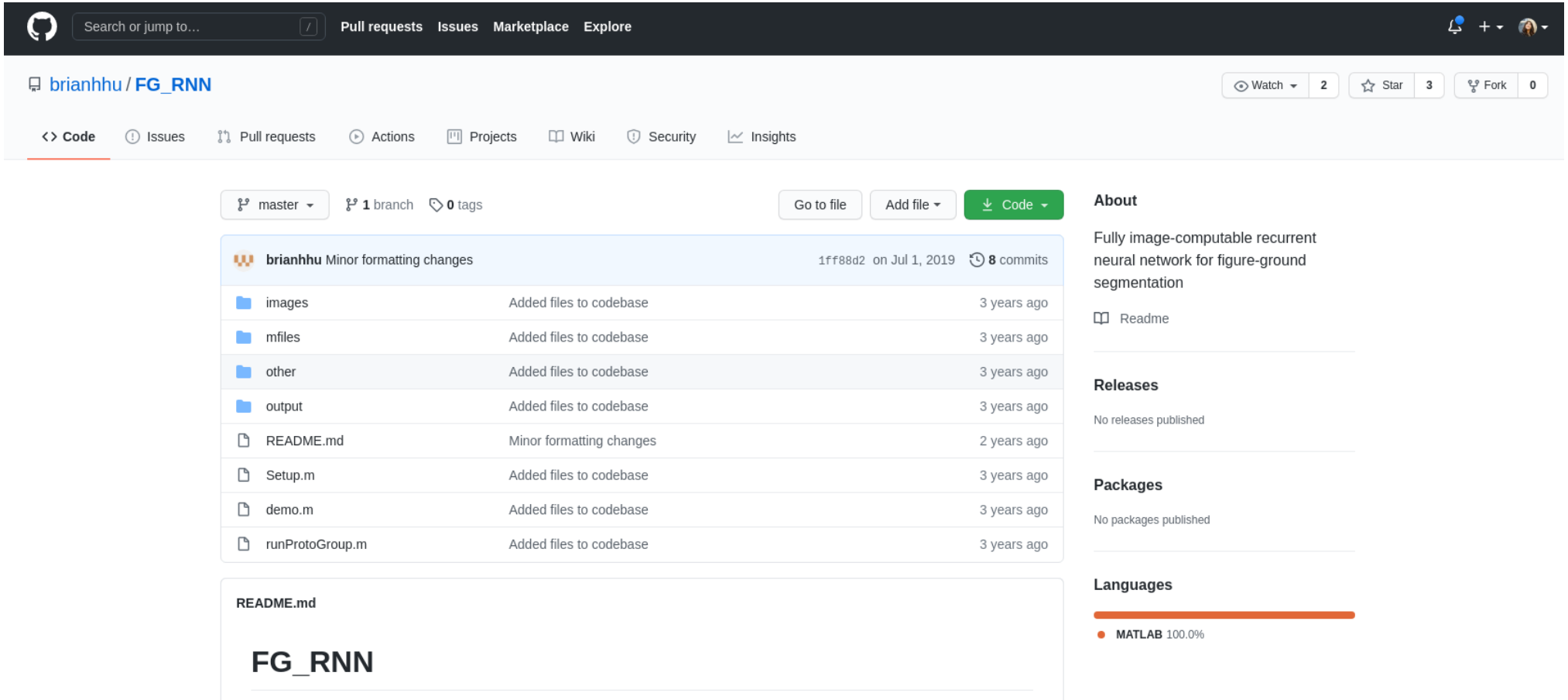
highly BOS response, the same region of border ownership (>80%)

how well the statistical model fits a set of observation

Journal Club

Code accessibility

https://github.com/brianhhu/FG_RNN



Journal Club

PROS

- good paper, it could be interesting to implement ED vision
- it seems to work well enough
- it does not reach amazing results but it does not need training
- good usage of the proto object model

CONS

- expressed badly, too long & complex for someone that knows nothing of BOS
- no good explanation of the model with diagrams how input to the output
- they only show 1 figure where they show the segmentation

Journal Club


Journal Club


Journal Club


Journal Club


Journal Club


Journal Club


Journal Club