Case Study:
Color encoding in CNNs and some parallelisms with human vision

Vision Research
Volume 151, Oct-2018, 7-17
https://doi.org/10.1016/j.visres.2018.03.010

IEEE Workshop on
Conference on Internation Conference on
Computer Vision (WICCV 2017)
https://doi.org/10.1109/ICCVW.2017.318

What?

We dissected a trained Convolutional Neural Network (CNN)
to understand how color is represented.

How?

We defined a methodology to visualize and understand neuron functionalities
based on a color selectivity index

Questions to answer:

- How many color selective neurons there are?
- Are they selective to one or more colors?
- Which colors are they selective to?
- Is color independent of shape?
- Is there any parallelism between our findings and known evidences in the human visual system?

Which network?

VGG-M Network proposed by Chatefield-et al. [1] and which is similar to Zeiler & Fergus CNN [2]

Why this one?

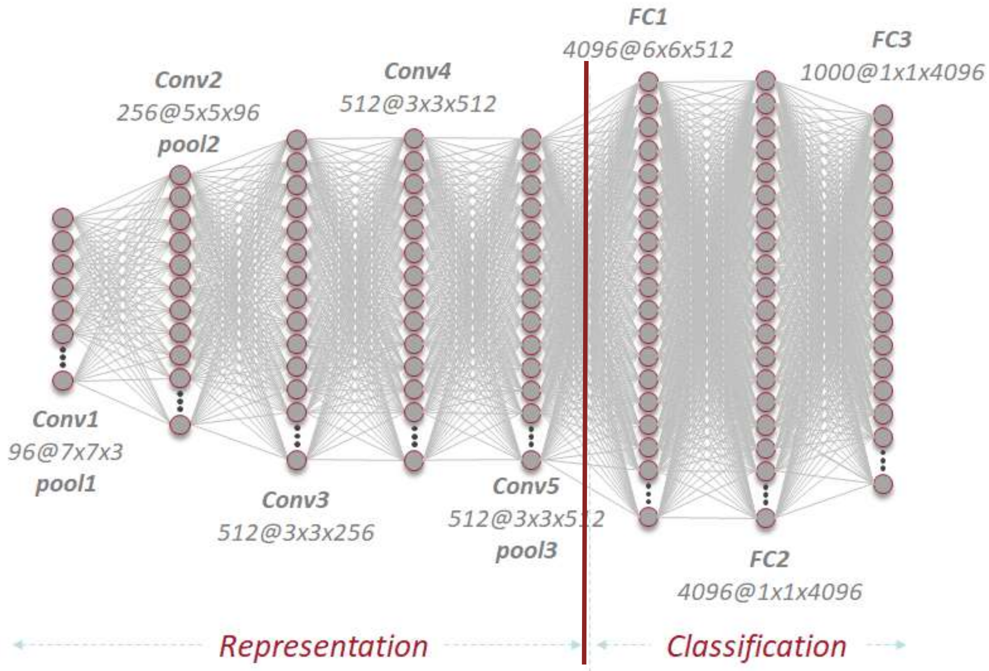
Cadieu *et al.* [3] proved that this sort of architecture shows representational capabilities that **rival primate performance** in visual recognition

[1] K. Chatefield, K.- Simonyan, A. Vedaldi, A. Zisserman. "Return of the devil in the details: Delving deep into convolutional nets". *British Machine Vision Conference*. 2014.
[2] M. Zeiler, R. Fergus. "Visualizing and Understanding Convolutional Networks". *13th European Conference (ECCV 2014)*. 2014
[3] C.F. Cadieu, H. Hong, D.L.K. Yamins, N. Pinto, D. Ardila, E. A. Solomon, N. J . Majaj, J. J. DiCarlo. "Deep Neural Networks Rival the Representation of Primate IT Cortex for Core Visual Object Recognition". *PLoS Comput. Biol. 10(12).* 2014.

Case study: VGG-M Network trained by Chatefield *et al.* [2]

Network architecture:

5 convolutional layers + 3 fully connected layers



Trained on ImageNet for an object recognition task [1]

- Large dataset with 1.2M images
- Labeled in 1000 different object categories
- JPEG compressed
- Uncalibrated RGB

Imagenet Dataset

[1] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... Fei-Fei, L. (2015). ImageNet large scale visual recognition challenge. International Journal of Computer Vision (IJCV), 115, 211–252.
[2] K. Chatefield, K.- Simonyan, A. Vedaldi, A. Zisserman. “Return of the devil in the details: Delving deep into convolutional nets”. *British Machine Vision Conference*. 2014.

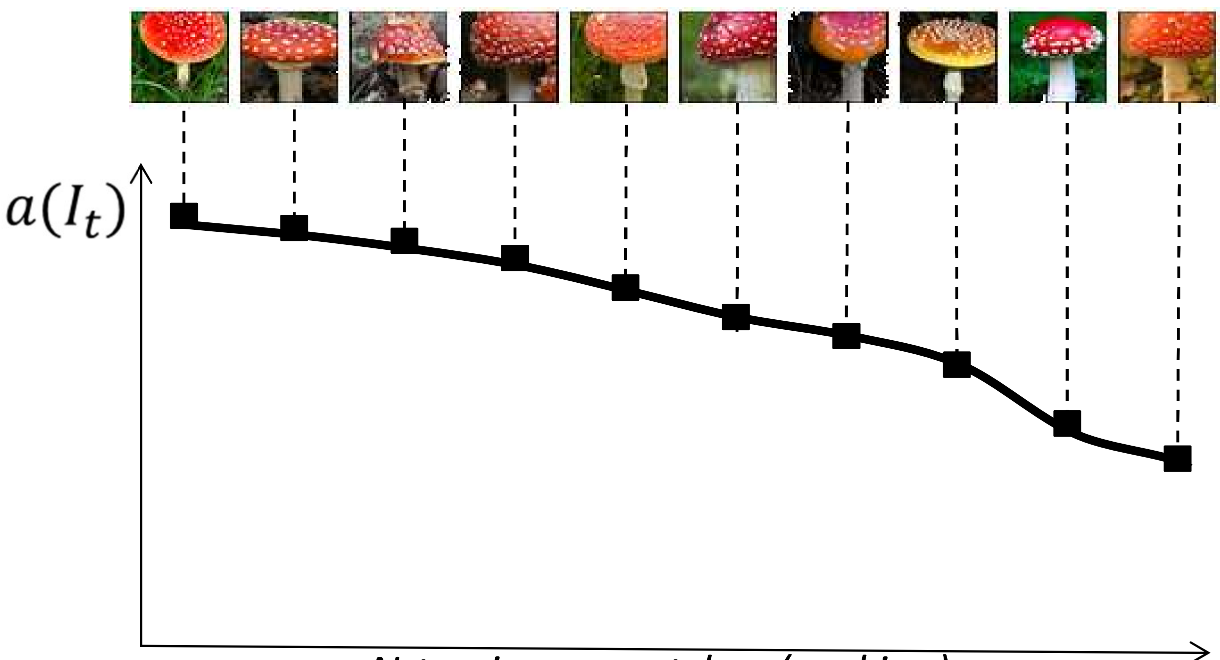
Before to start, we needed to **<u>visualize neuron activity</u>**

Neuron Feature (NF)

Definition: NF **visualizes** the features provoking a high activation of a specific neuron as the <u>weighted average of the first N top scoring images</u>.

Example: for a given neuron we get

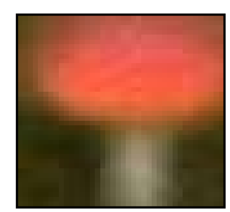
N-top imatge patches



N-top image patches (ranking)
(We used N=100 in all the experiments)

Construction:

$$\boxed{NF = \frac{1}{N_{max}} \sum_{t}^{N_{max}} a(I_t) \cdot I_t}$$



NF

For understanding, we proposed **Neuron Classification**

Brian A. Wandell in [1] states:

> *" [...] The most fundamental method of distinguishing categories of neurons is simply to study their **morphology**. A second type of data we can use is the neuron's electrical responsiveness to different signals, that is its **electrophysiology**. A third type of data we can use is to study the chemical substances used to build the neuron, that is the neuron's **biochemistry**. A fourth type of data is the **anatomical** pattern of interconnections a neuron makes with other neurons."*

So, we decided to classify neurons using a **Color selectivity Index**

[**1**] Brian A. Wandell. Foundations of Vision. Stanford University, 1995

Color selectivity Index (α)

Definition: measures the activity of a neuron to an input stimulus presenting color bias. It presents a high value when a neuron is sensitive to a color, and low value when is not.

Estimation:

$$\alpha = 1 - \frac{\sum_t^N a_t(I_{gray})}{\sum_t^N a_t(I_{color})}$$

N-top RGB patches

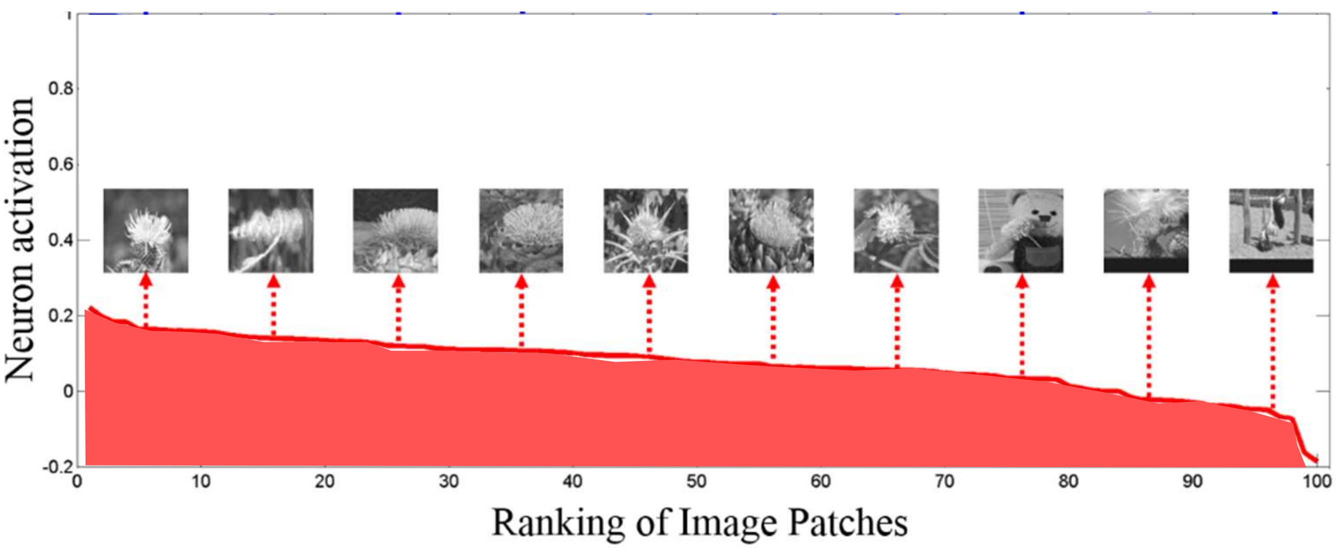
Gray-level versions of the RGBpatches

Color selectivity Index (α)

Definition: measures the activity of a neuron to an input stimulus presenting color bias. It presents a high value when a neuron is sensitive to a color, and low value when is not.

Estimation:

$$\alpha = 1 - \frac{\boxed{\sum_t^N a_t(I_{gray})}}{\sum_t^N a_t(I_{color})}$$



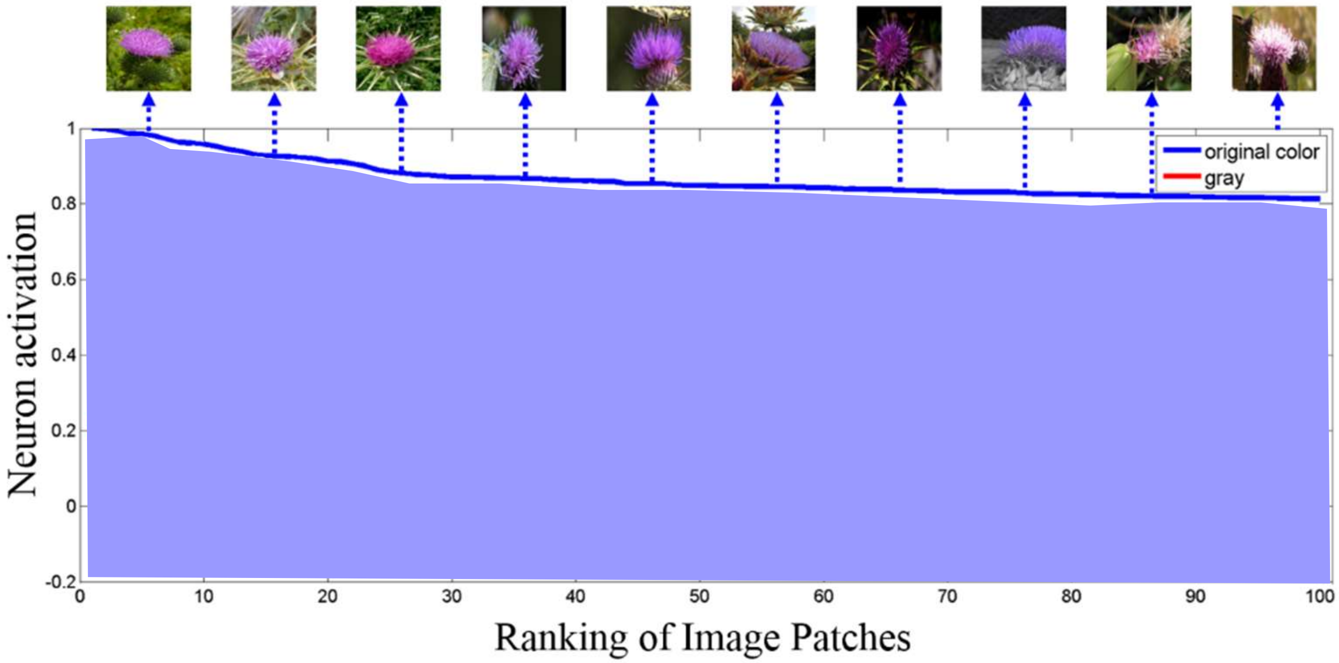
Example of a High color selective neuron

Color selectivity Index (α)

Definition: measures the activity of a neuron to an input stimulus presenting color bias. It presents a high value when a neuron is sensitive to a color, and low value when is not.

Estimation:

$$\alpha = 1 - \frac{\sum_t^N a_t(I_{gray})}{\boxed{\sum_t^N a_t(I_{color})}}$$



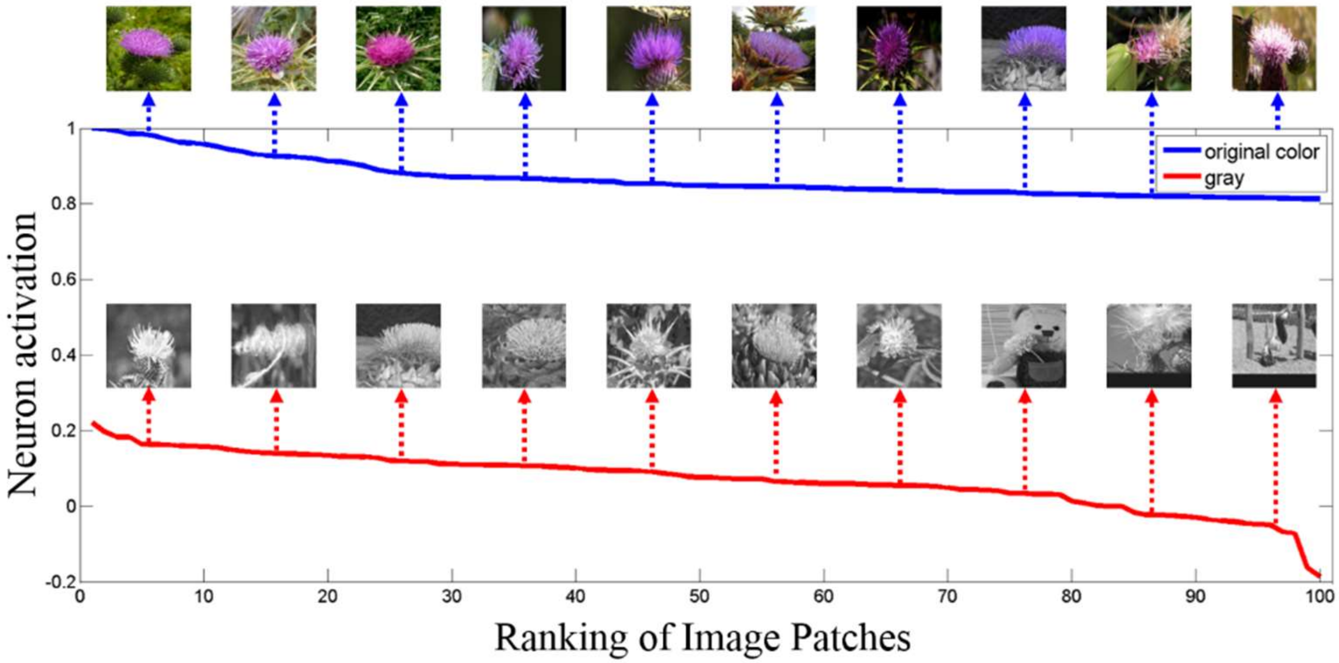
Example of a High color selective neuron

Color selectivity Index (α)

Definition: measures the activity of a neuron to an input stimulus presenting color bias. It presents a high value when a neuron is sensitive to a color, and low value when is not.

Estimation:

$$\alpha = 1 - \frac{\boxed{\sum_t^N a_t(I_{gray})}}{\boxed{\sum_t^N a_t(I_{color})}}$$



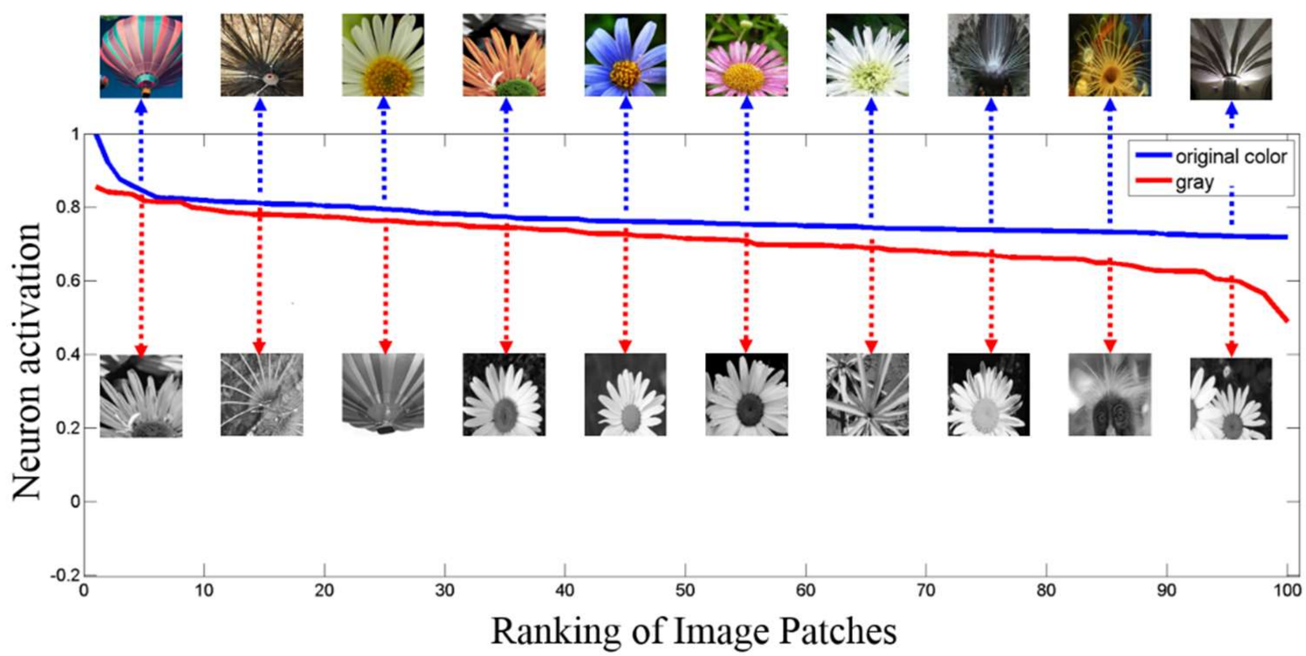
Example of a High color selective neuron
(selective to a purple cardoon)

Color selectivity Index (α)

Definition: measures the activity of a neuron to an input stimulus presenting color bias. It presents a high value when a neuron is sensitive to a color, and low value when is not.

Estimation:

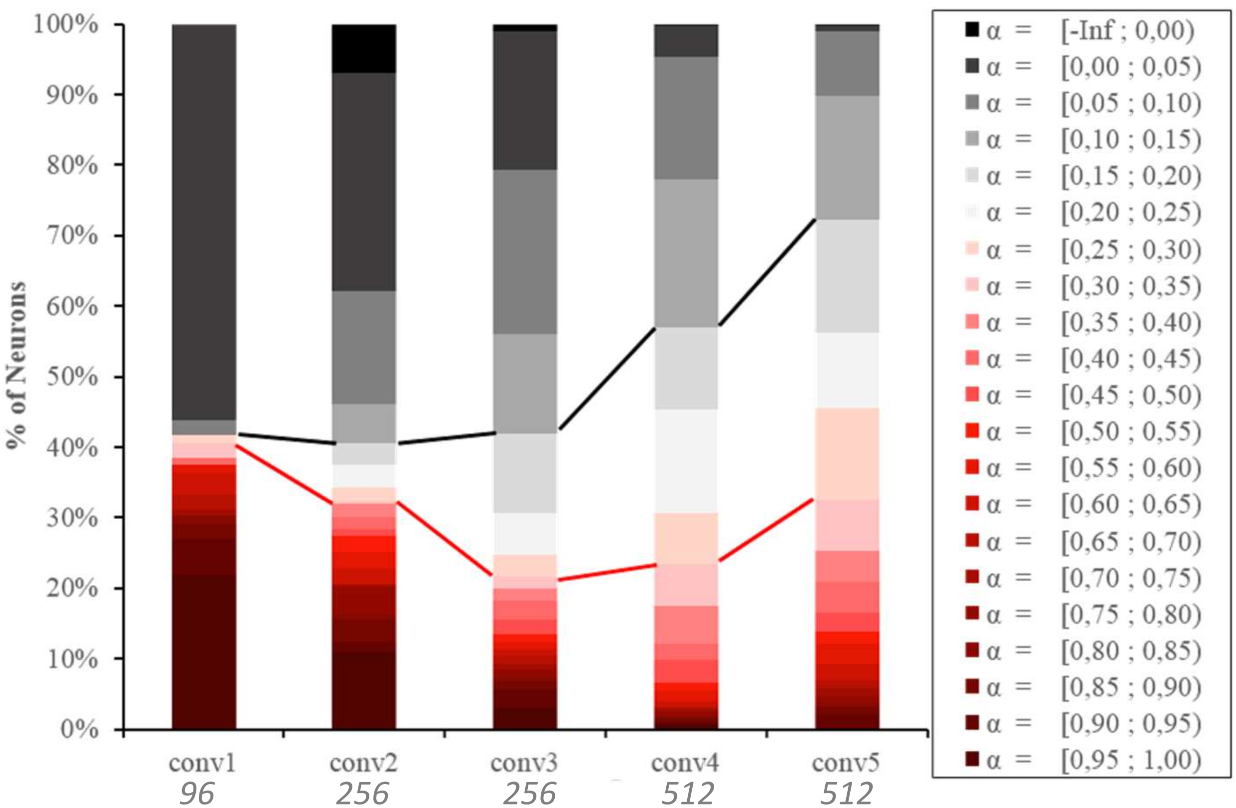
$$\boxed{\alpha = 1 - \frac{\boxed{\sum_t^N a_t(I_{gray})}}{\boxed{\sum_t^N a_t(I_{color})}}}$$



Example of a Non color selective neuron

(selective to a any color daisy)

Color selectivity through layers



1st parallelism

Color selective neurons can be found through **all the layers** like in HVS [**1**]

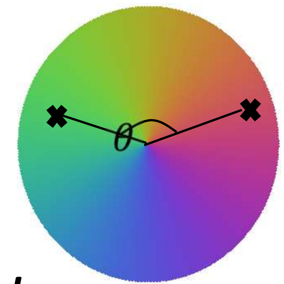
Conclusions:

- A high degree of color selectivity is found in all layers, even in deeper layers
- Low color-selectivity increases with depth

[**1**] B. R. Conway, S. Chatterjee, G. D. Field, G. D. Horwitz, E. N. Johnson, K. Koida, , and K. Mancuso. Advances in color science: from retina to behavior. *The Journal of Neuroscience*, 30(45),2010

Classification of neurons based on color selectivity

- *High Color selective*
  $\alpha > thr_1$
  - *Single*
    1 color
  - *Double*
    2 color
    - *Opponent*
      Close to 180º
    - *Non opponent*
      Far from 180º
- *Low Color selective*
  $\alpha \in [thr_1, thr_2]$
- *Non Color Selective*
  $\alpha > thr_2$



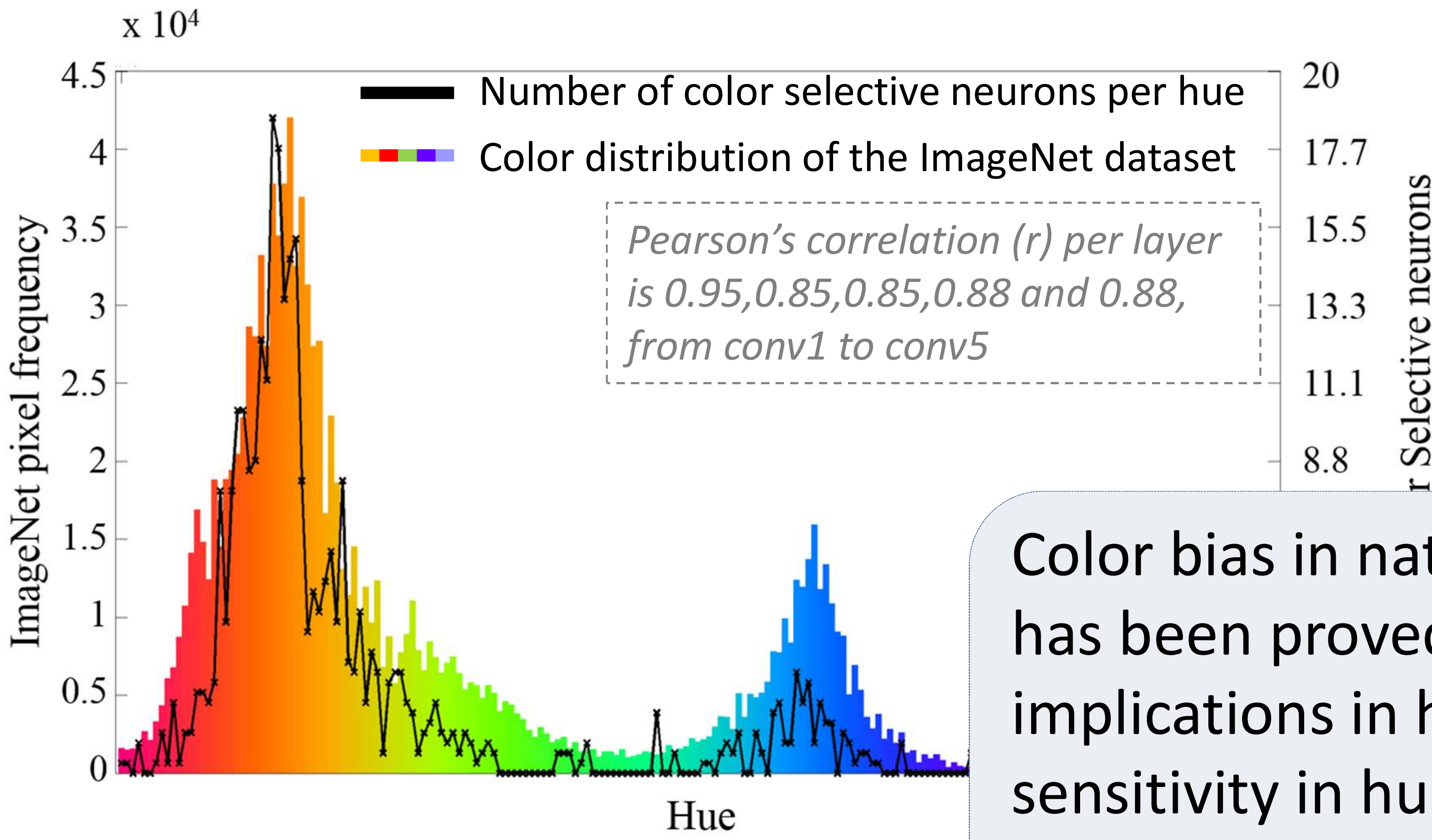
Gaussian Mixture Model was fitted with EM algorithm

K-means was used to find emergent color axes

Selectivity #Neurons	Conv1 96	Conv2 219	Conv3 512	Conv4 512	Conv5 512
Non Color	58 (60.42%)	83 (37.90%)	153 (29.88%)	107 (20.90%)	148 (28.91%)
Low Color Sel	0 (0%)	61 (27.85%)	268 (52.34%)	332 (64.84%)	285 (55.66%)
Color Sel	38 (39.58%)	75 (34.25%)	91 (17.77%)	73 (14.26%)	79 (15.43%)

Correlation dataset statistics

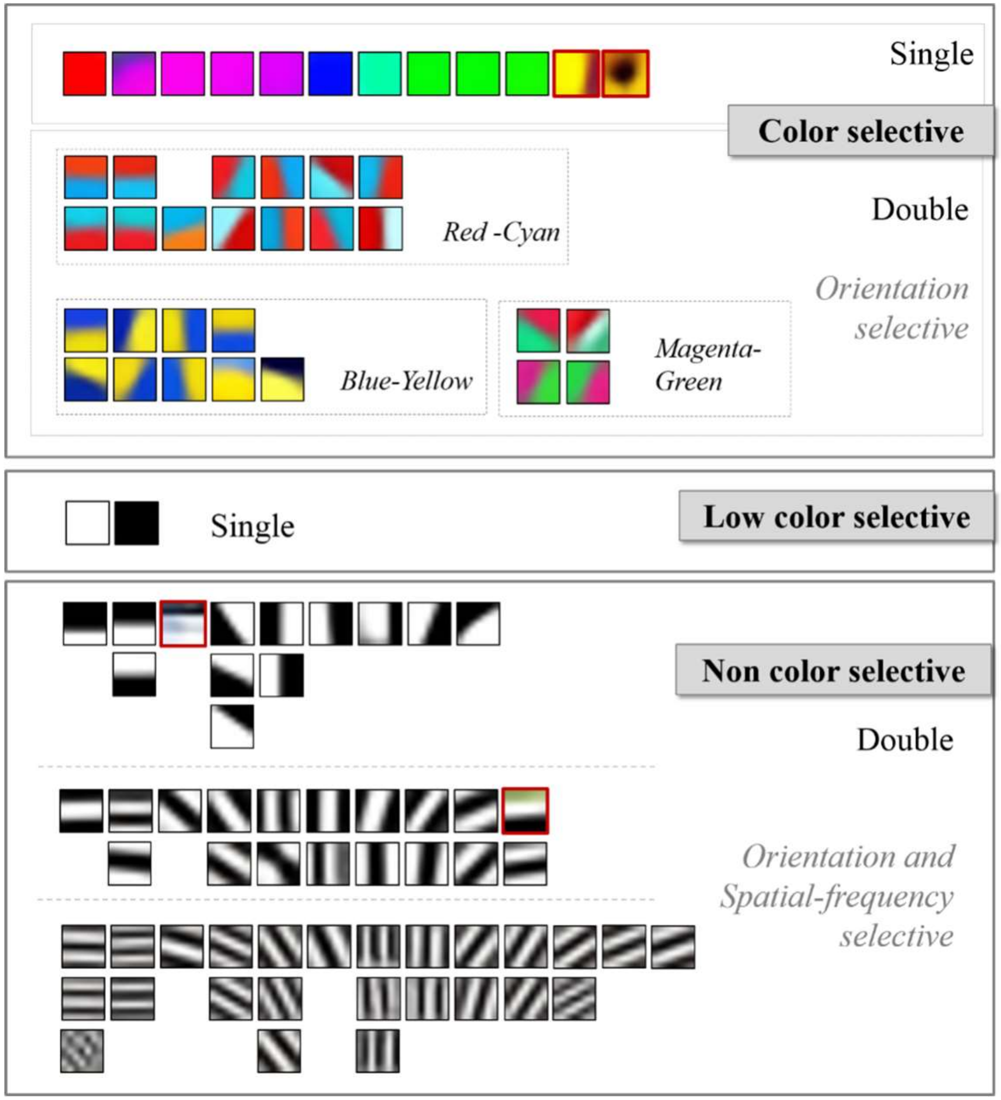
High correlation between the number of color selective neurons and the dataset color distribution



Color bias in natural scenes has been proved to have implications in higher color sensitivity in human visual system [1]

[1] K. C. McDermott and M. A. Webster. **Uniform color spaces and natural image statistics**. JOSA, 29(2):A18/2—A187. 2012

Neurons in Conv1

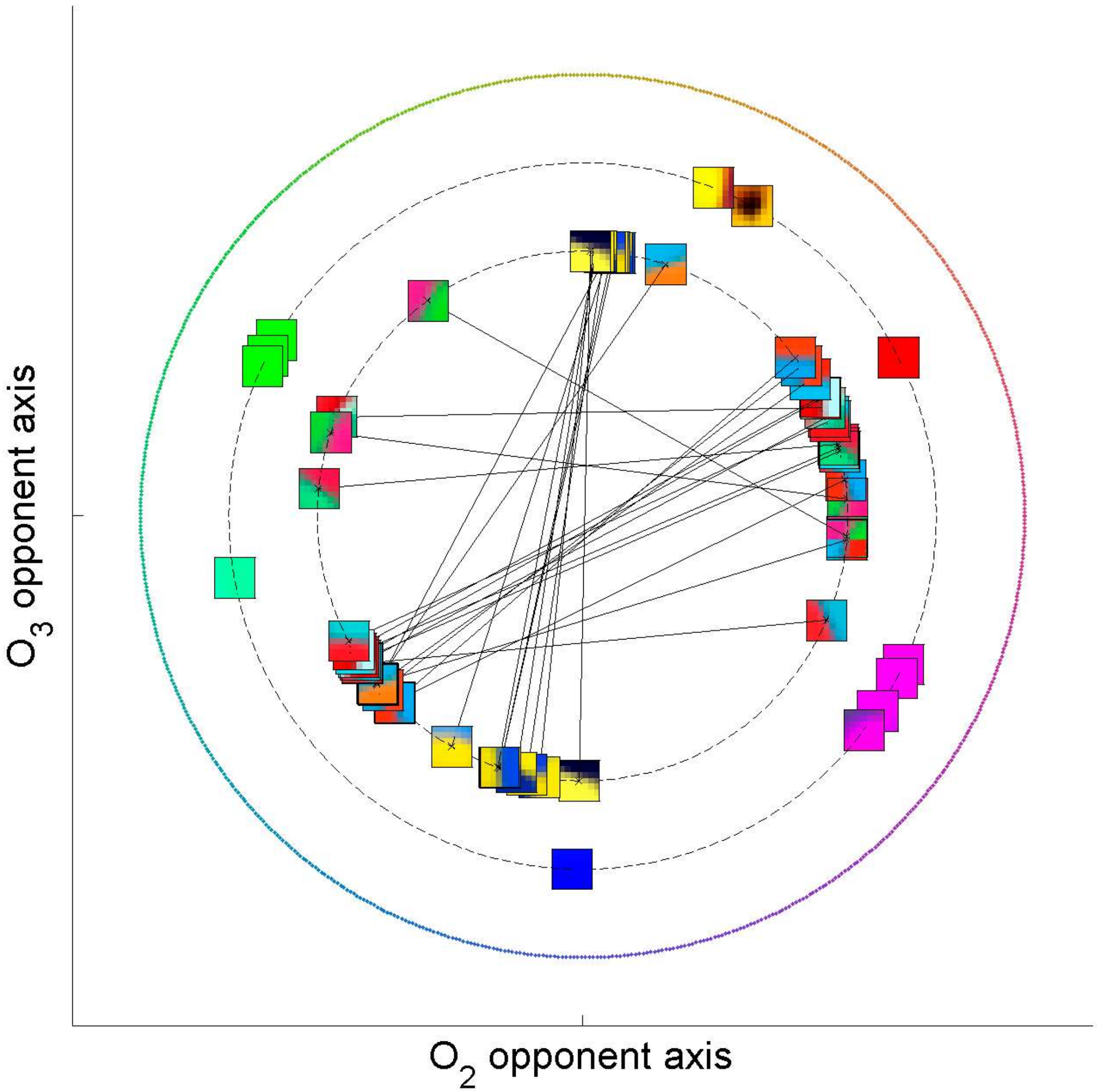


Selectivity #Neurons	Conv1 96
Non Color	58 (60.42%)
Low Color Sel	0 (0%)
Color Sel	38 (39.58%)
Single Color	12 (12.50%)
Double Color	26 (27.08%)
Opponent	21 (21.88%)
Non opponent	5 (5.20%)

A **higher spatial frequency selectivity in black-white neurons** compared to **low spatial frequency of color selective** neurons have also been reported in [1]

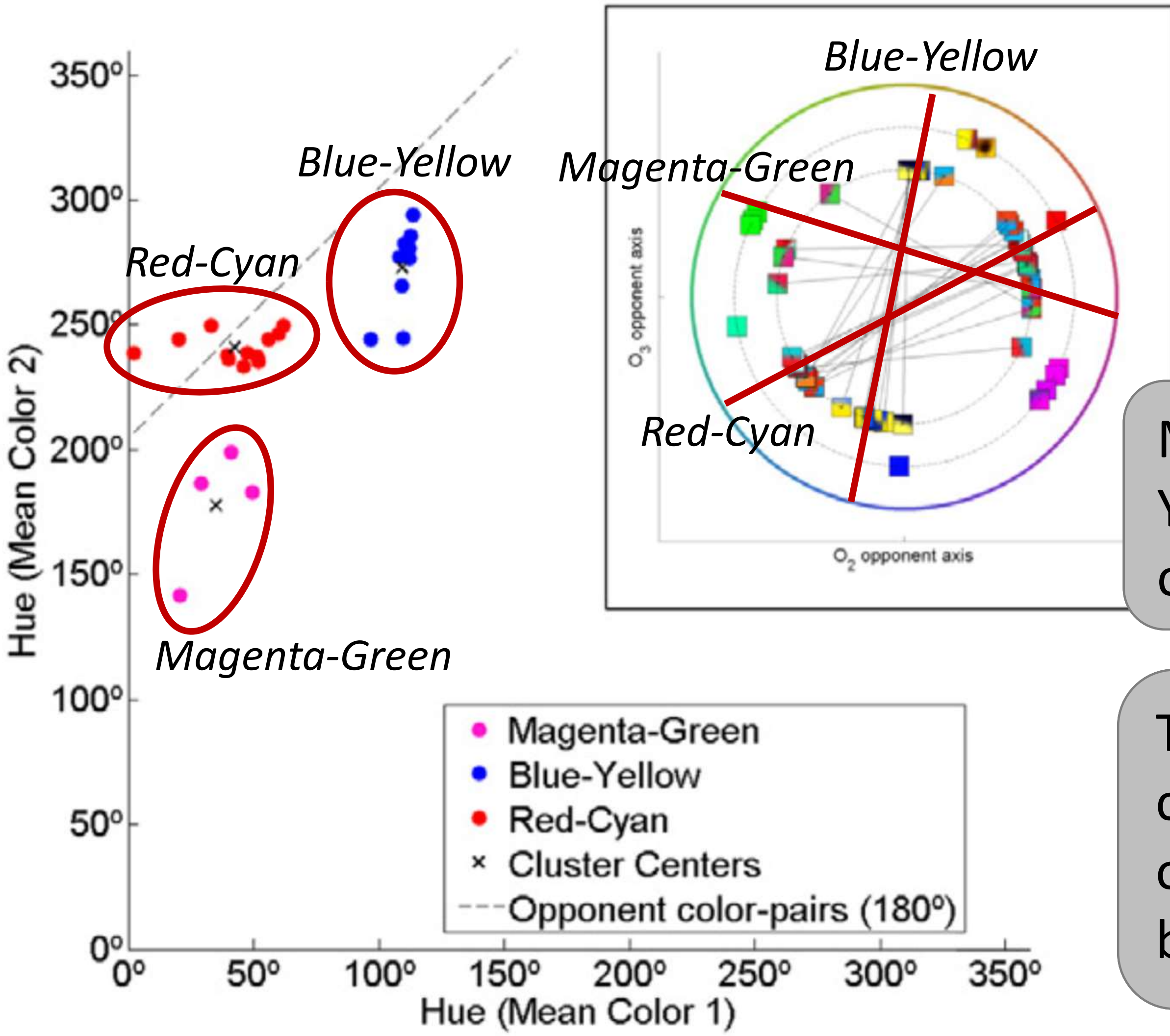
[1] R. Shapley, M. Hawken. "Color in the Cortex : single- and double-opponent cells". *VR, 51 (7).* 2011.

Single and Double neurons in Conv1



Selectivity #Neurons	Conv1 96
Non Color	58 (60.42%)
Low Color Sel	0 (0%)
Color Sel	38 (39.58%)
Single Color	12 (12.50%)
Double Color	26 (27.08%)
Opponent	21 (21.88%)
Non opponent	5 (5.20%)

Opponency in Conv1



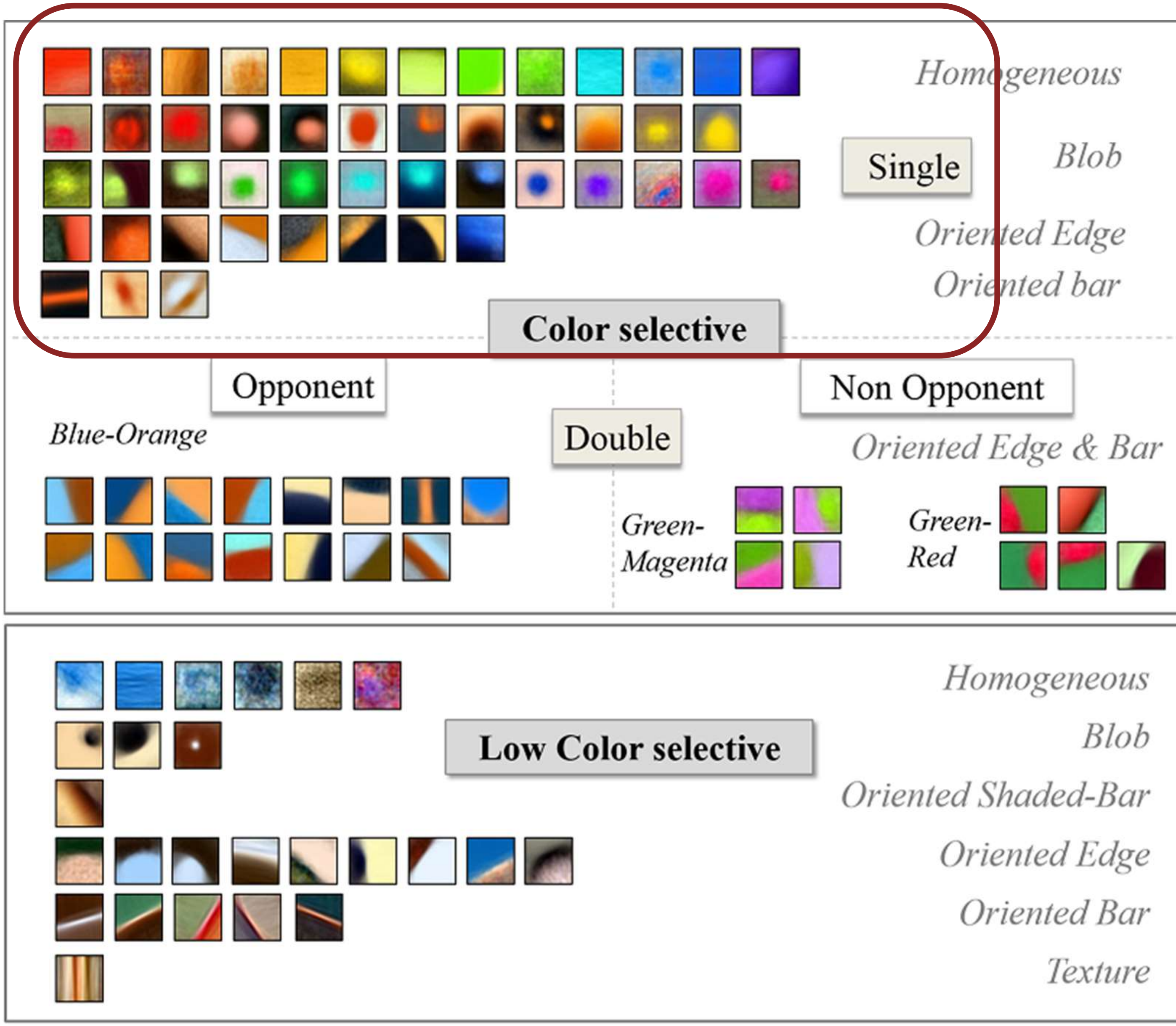
Selectivity #Neurons	Conv1 96
Non Color	58 (60.42%)
Low Color Sel	0 (0%)
Color Sel	38 (39.58%)
Single Color	12 (12.50%)
Double Color	26 (27.08%)

Magenta-Green and Blue-Yellow opponent channels correlate with findings in [1,2]

The Red-Cyan channel could correlate with a fourth opponent channel reported by Conway in [3]

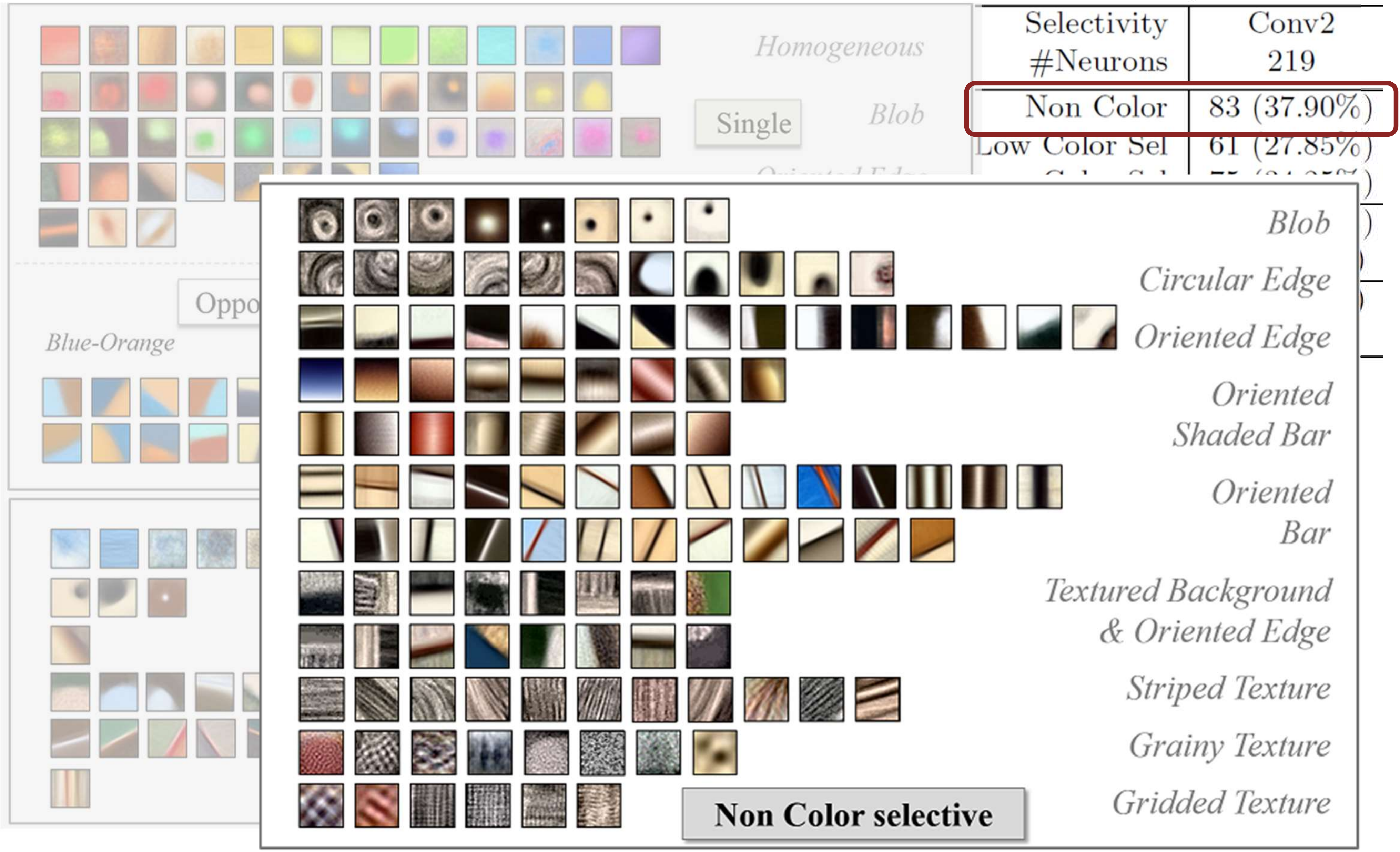
[1] Derrington, et-al (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. Journal of Physiology, 241–265.
[2] Lennie, P., & D'Zmura, M. (1988). Mechanisms of color vision. CRC Critical Reviews in Clinical Neurobiology, 333–400
[3] Conway, B. R. (2001). Spatial structure of cone inputs to color cells in alert macaque primary visual cortex (v-1). Journal of Neuroscience, 21, 2768–2783

Neurons in Conv2



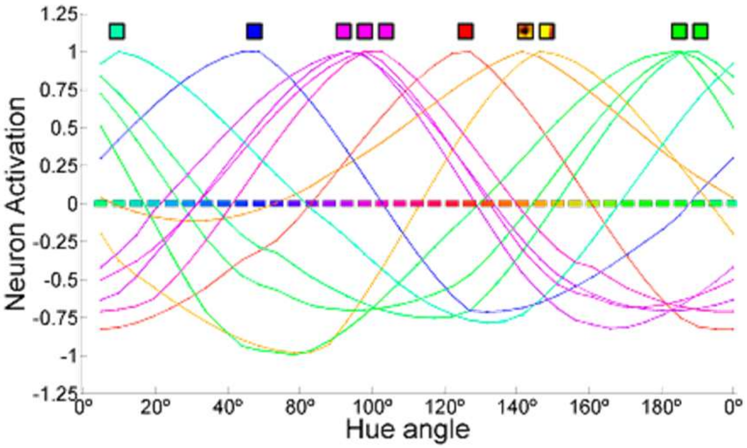
Selectivity #Neurons	Conv2 219
Non Color	83 (37.90%)
Low Color Sel	61 (27.85%)
Color Sel	75 (34.25%)
Single Color	54 (24.66%)
Double Color	21 (9.59%)
Opponent	14 (6.39%)
Non opponent	7 (3.19%)

Neurons in Conv2

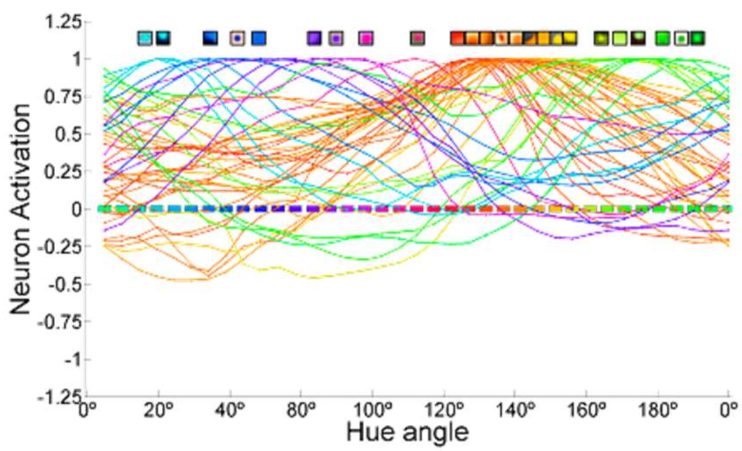


Single color neurons:

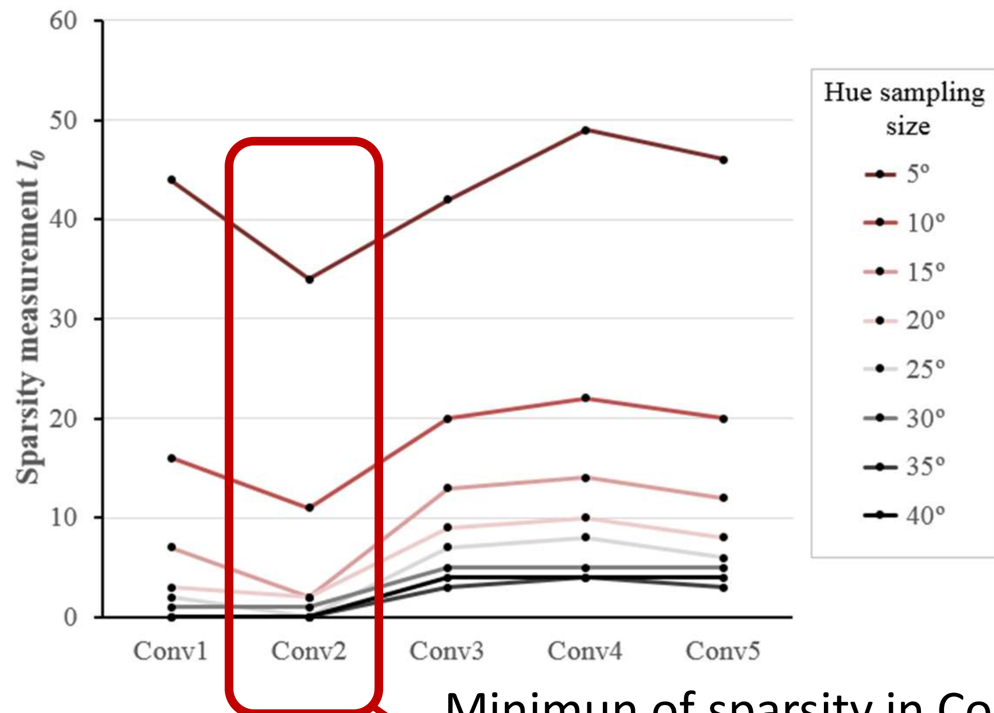
Conv 1



Conv 2



Hue Sampling in layer Conv2 is **more dense** (less sparse) than in the rest of layers

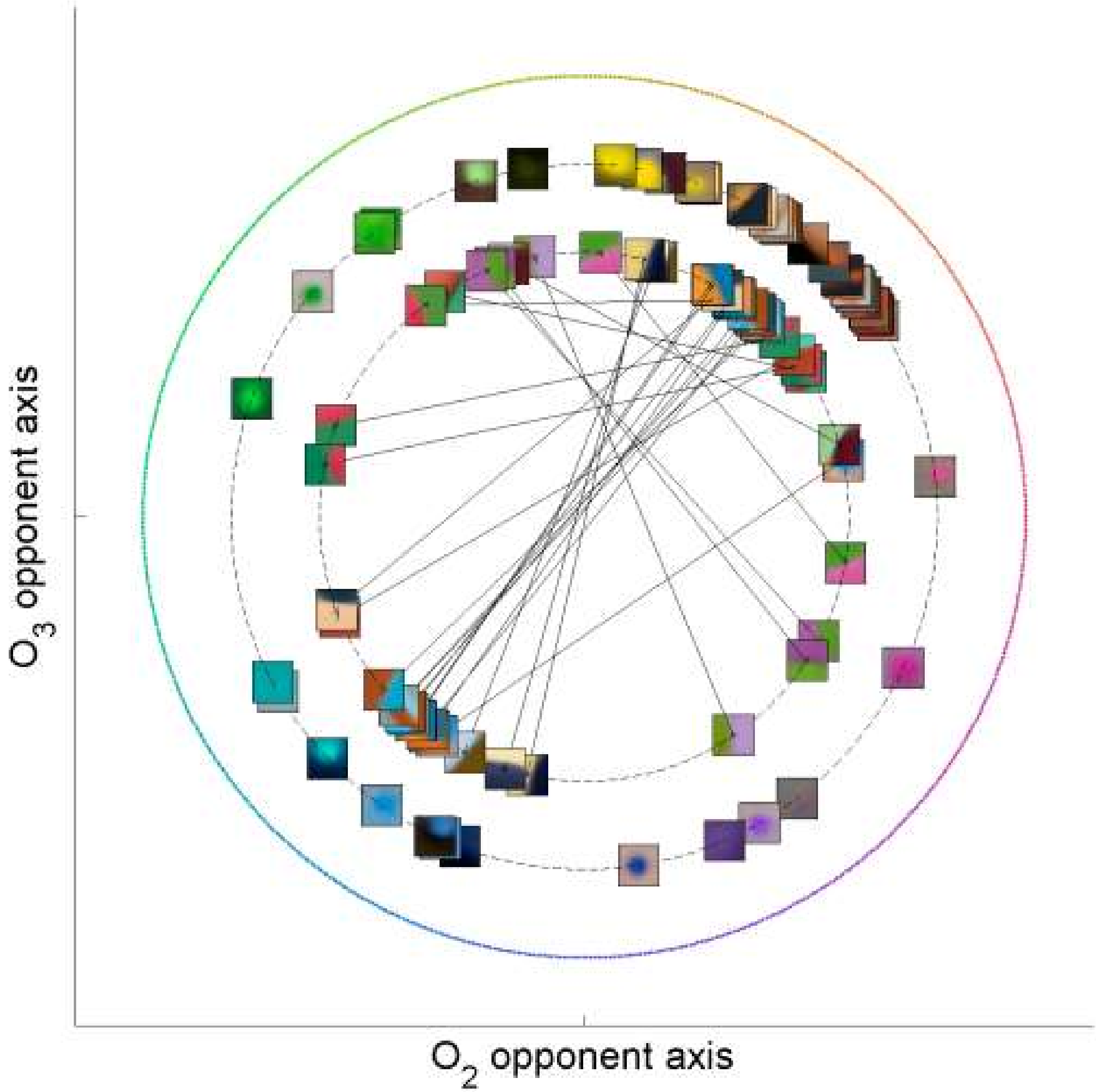


Minimun of sparsity in Conv2 (máximum sampling)

Hue maps in V2 [1,2,3].

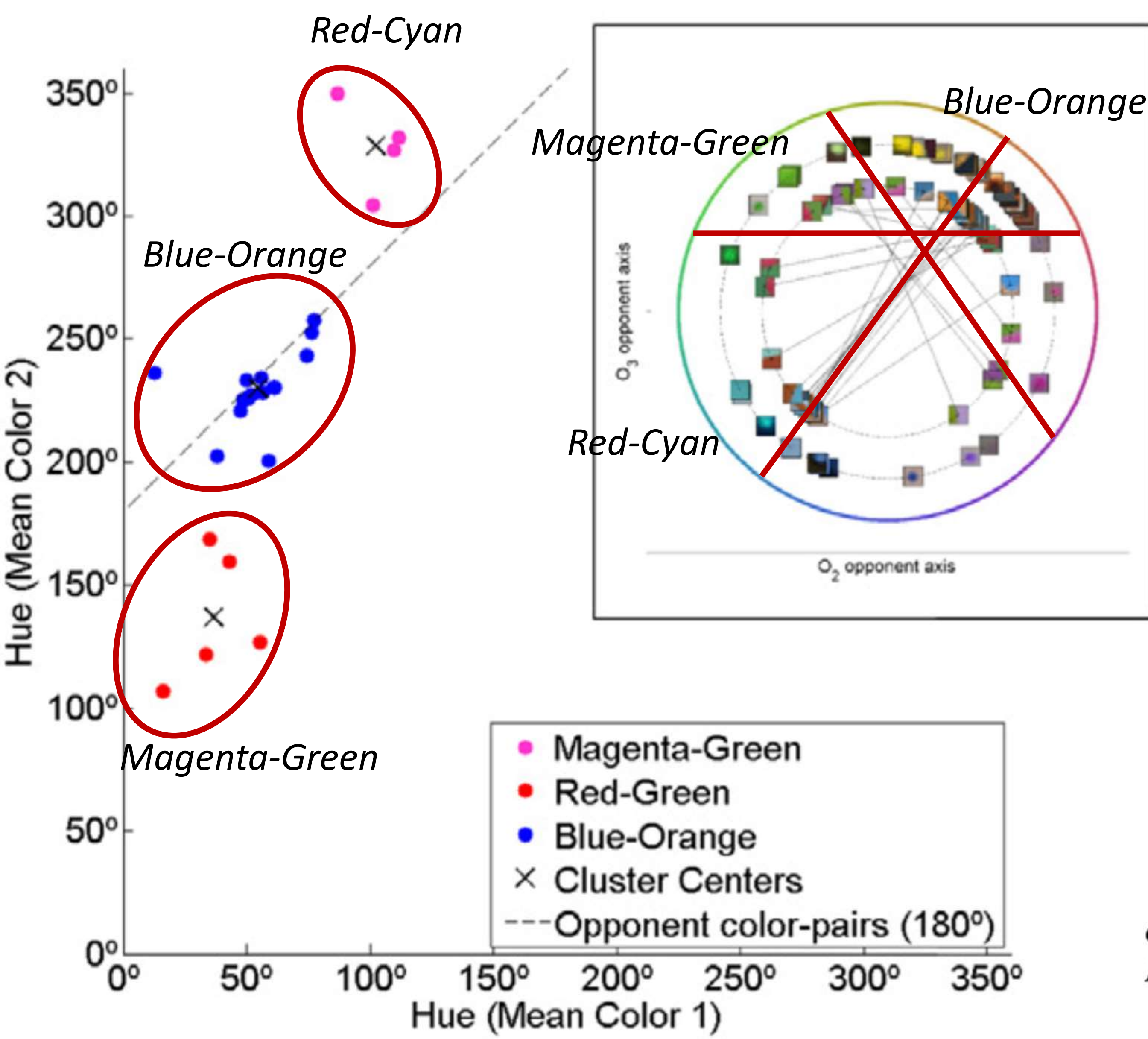
[**1**] L. H., W. Y., X. Y., H. M., and F. DJ. Organization of hue selectivity in macaque v2 thin stripes. *Journal of Neurophysiology,* 102(5).
[**2**]M. A. Webster, Y. Mizokami, and S. M. Webster. Hue maps in primate striate cortex. *NeuroImage*, 35(2).
[**3**] Y. Xiao and Y. W. D. Felleman. A spatially organized representation of colour in macaque cortical area v2. *Nature,*421.

Single and Double neurons in Conv2

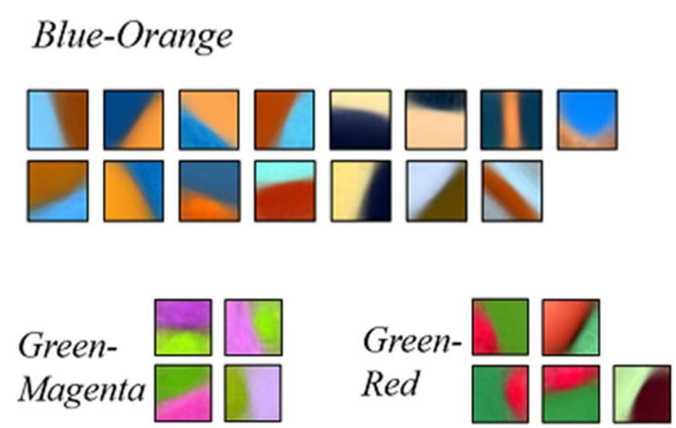


Selectivity #Neurons	Conv2 219
Non Color	83 (37.90%)
Low Color Sel	61 (27.85%)
Color Sel	75 (34.25%)
Single Color	54 (24.66%)
Double Color	21 (9.59%)
Opponent	14 (6.39%)
Non opponent	7 (3.19%)

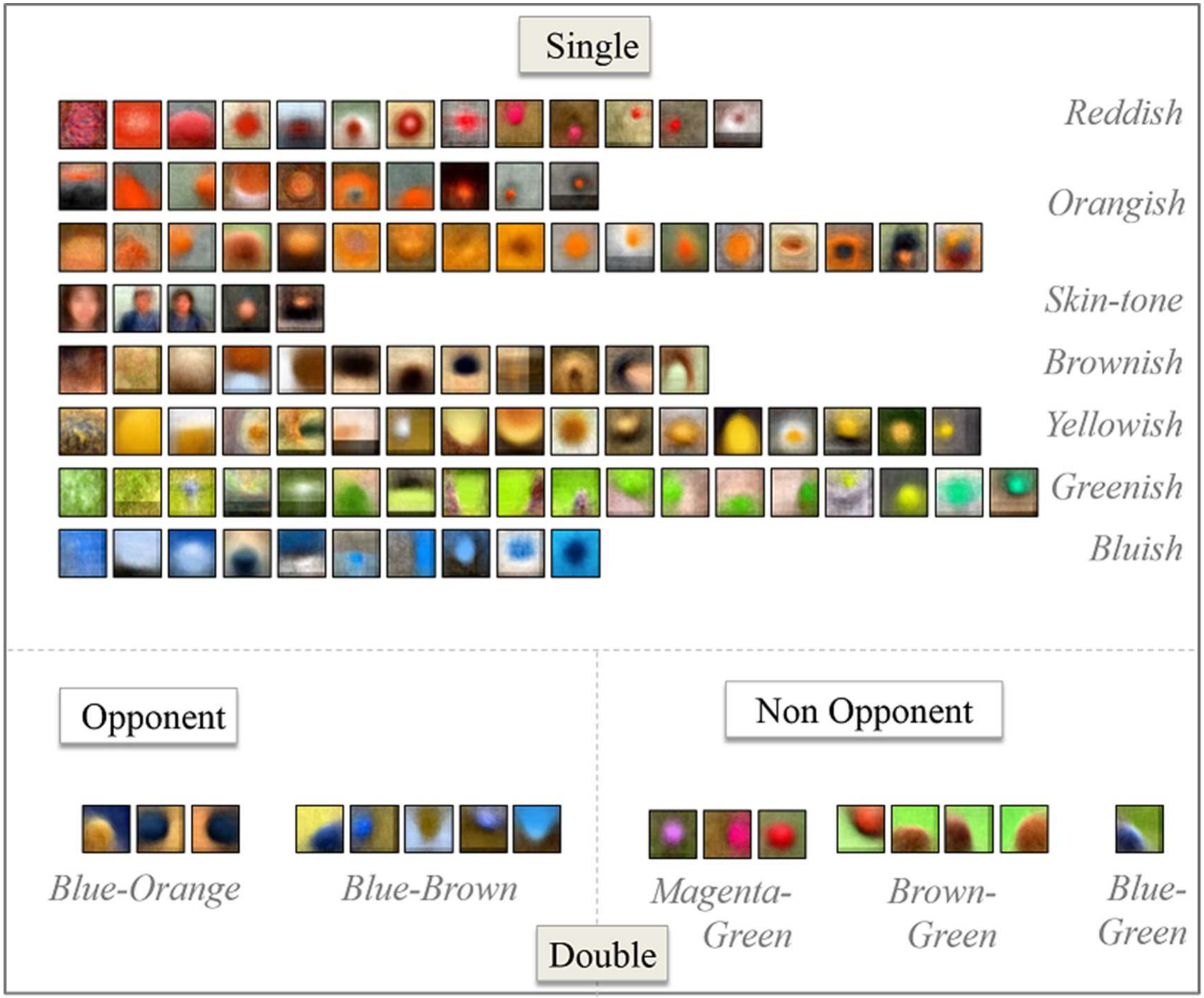
Opponency in Conv2



Selectivity #Neurons	Conv2 219
Non Color	83 (37.90%)
Low Color Sel	61 (27.85%)
Color Sel	75 (34.25%)
Single Color	54 (24.66%)
Double Color	21 (9.59%)
Opponent	14 (6.39%)
Non opponent	7 (3.19%)



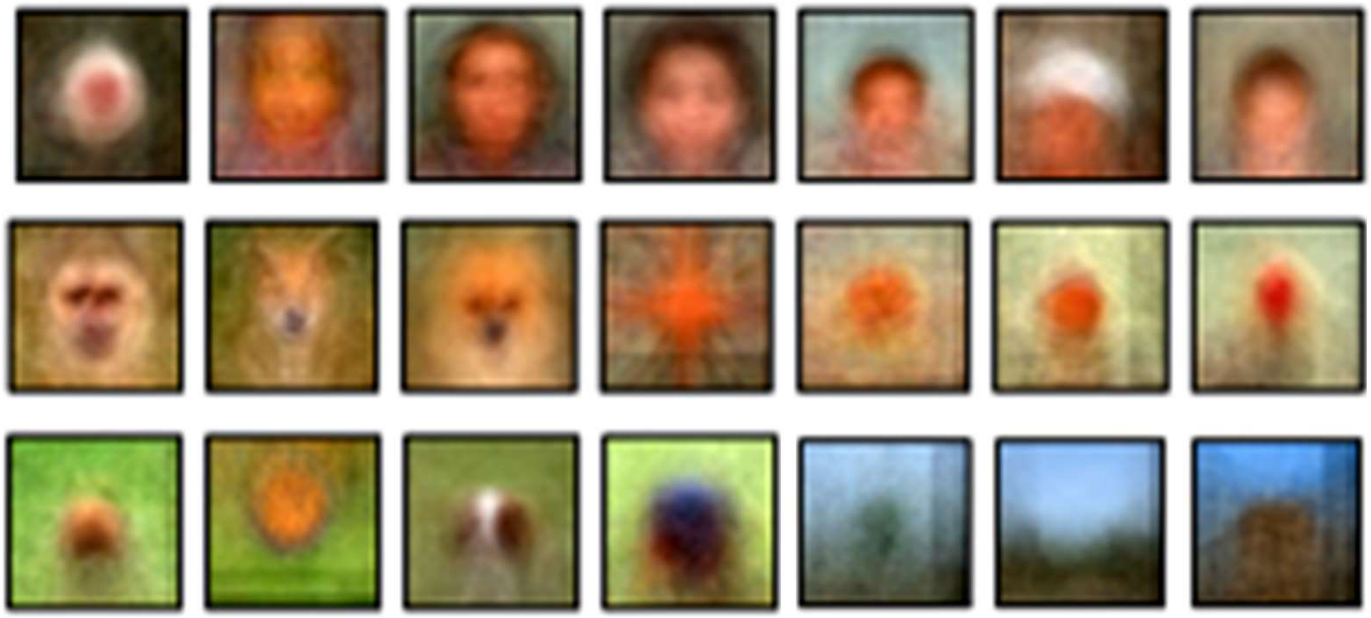
Neurons in Deeper layers (Conv3)



Neurons in Deeper layers (Conv 4)

Neurons in Deeper layers (Conv 5)

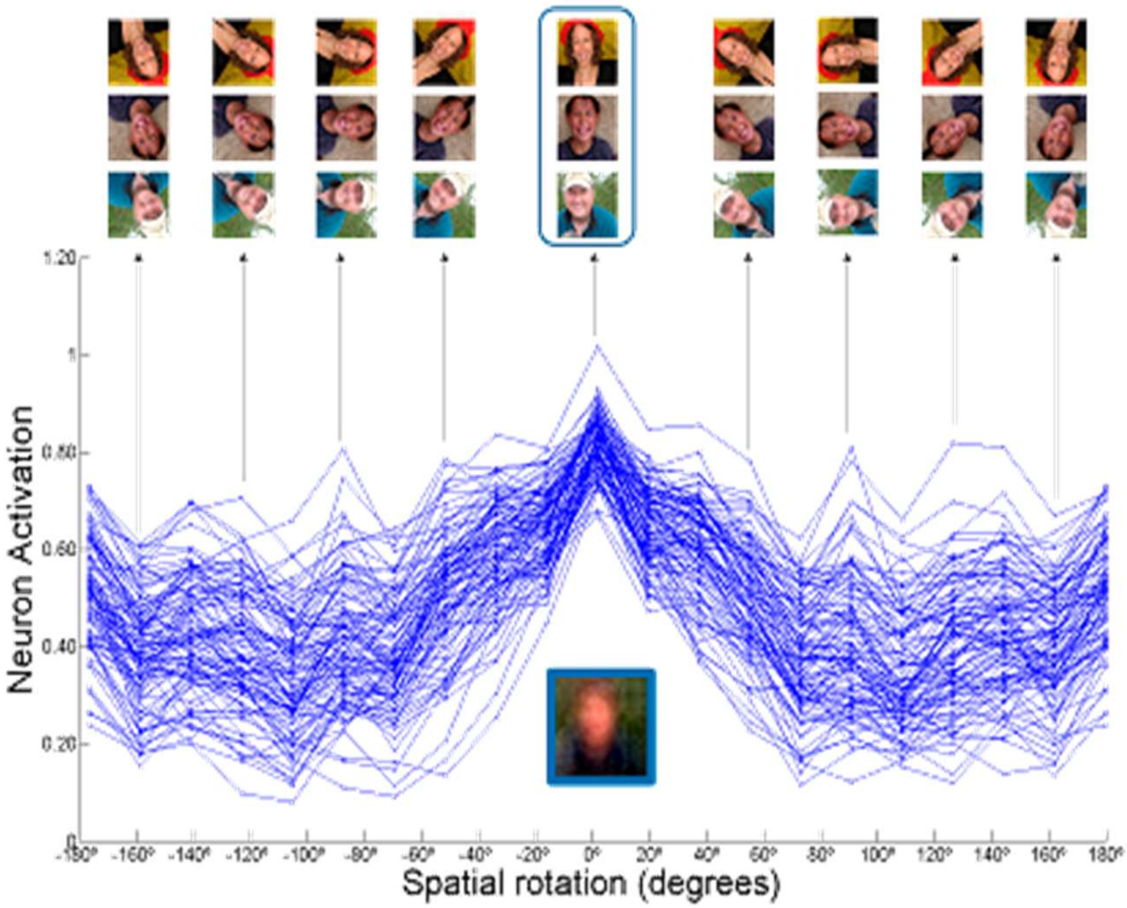
Neurons in V4 are selective to large range of colors and to white surfaces and also sensitive to surrounds that **may participate in the separation between object and background** [1].



[**1**] S. J. Schein, R. Desimone.Spectral properties of v4 neurons in the macaque.*VR* 51(7):701–717,2011.

Strong color-shape entanglement

Example: neuron in Conv4



Complex shape and color selectivity have been found **in V4 and PIT** [1]

[**1**] B. R. Conway, S. Chatterjee, G. D. Field, G. D. Horwitz, E. N. Johnson, K. Koida, , and K. Mancuso. Advances in color science: from retina to behavior. *The Journal of Neuroscience*, 30(45),2010.

Summary about initial questions:

- How many color selective neurons there are?
  A high number of color selective neurons are found in all layers

- Are they selective to one or more colors?
  We found double colors neurons, specially in shallower layers

- Which colors are they selective to?
  A high correlation of color selectivity with the dataset bias

- Is color independent of shape?
  Strong entanglement between color and shape

- Is there any parallelism between our findings and known evidences in the human visual system?
  - Similar color-opponent channels in shallow layers
  - Higher spatial frequency selectivity of non-color selective neurons
  - A more dense hue sampling in Conv2,
  - A strong entanglement between color and shape in deep layers too

Related Publications

Ivet Rafegas
Color in Visual Recognition: from flat to deep representations and some biological parallelisms.
Phd Thesis. Universitat Autònoma de Barcelona, 2017

Ivet Rafegas, Maria Vanrell
Color encoding in biologically-inspired convolutional neural networks
Vision Research, Volume 151, Oct-2018, 7-17

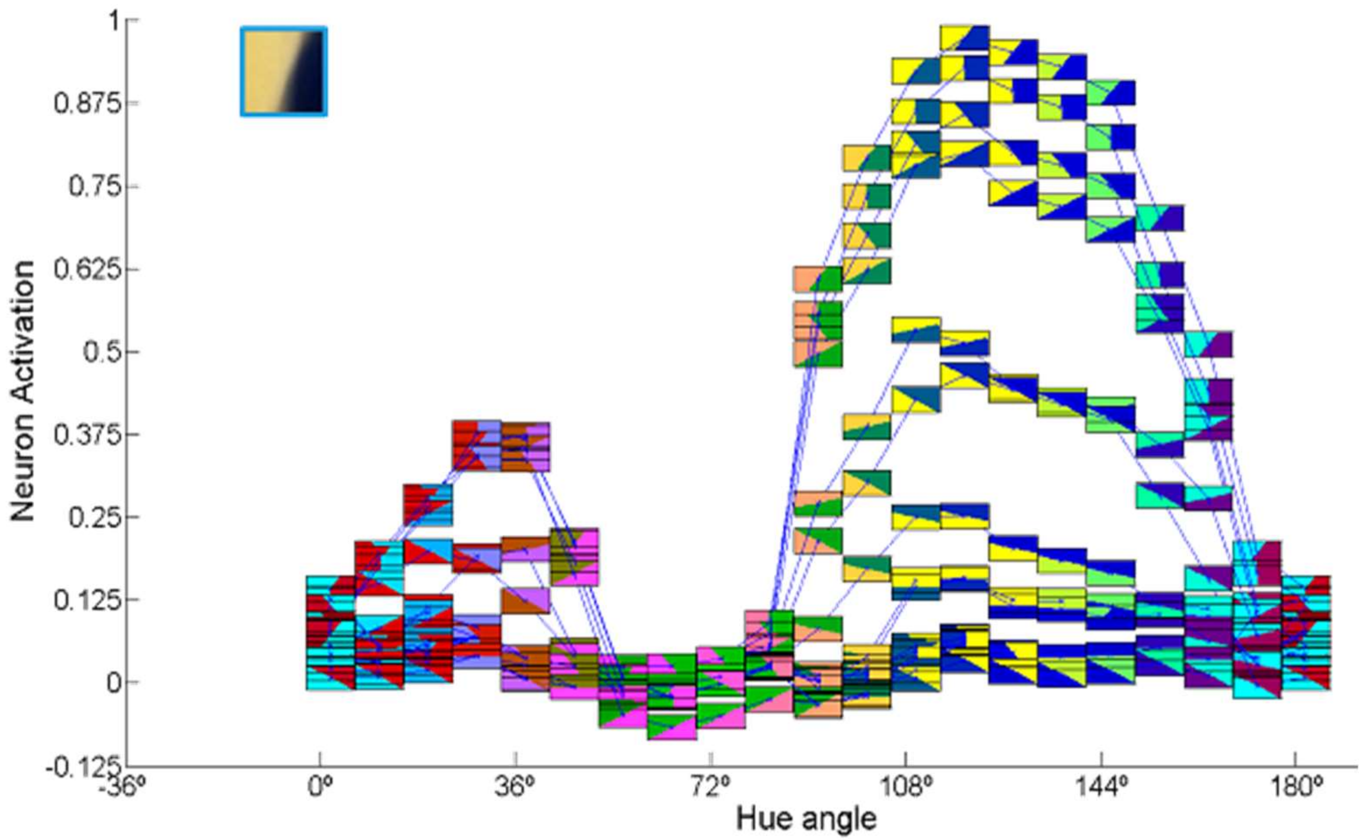
Ivet Rafegas, Maria Vanrell
Color representation in CNNs: parallelisms with biological vision
IEEE Workshop on International Conference on Computer Vision (WICCV) 2017

Ivet Rafegas, Maria Vanrell, Luís A Alexandre, Guillem Arias
Understanding trained CNNs by indesing neuron selectvity
Pattern Recognition Letters, 2019

Strong Color-Shape entanglement

Color always appears strongly linked with the intrinsic

Example 1: neuron in Conv2



[1] R. Shapley, M. Hawken. "Color in the Cortex : single- and double-opponent cells". *VR, 51 (7).* 2011.