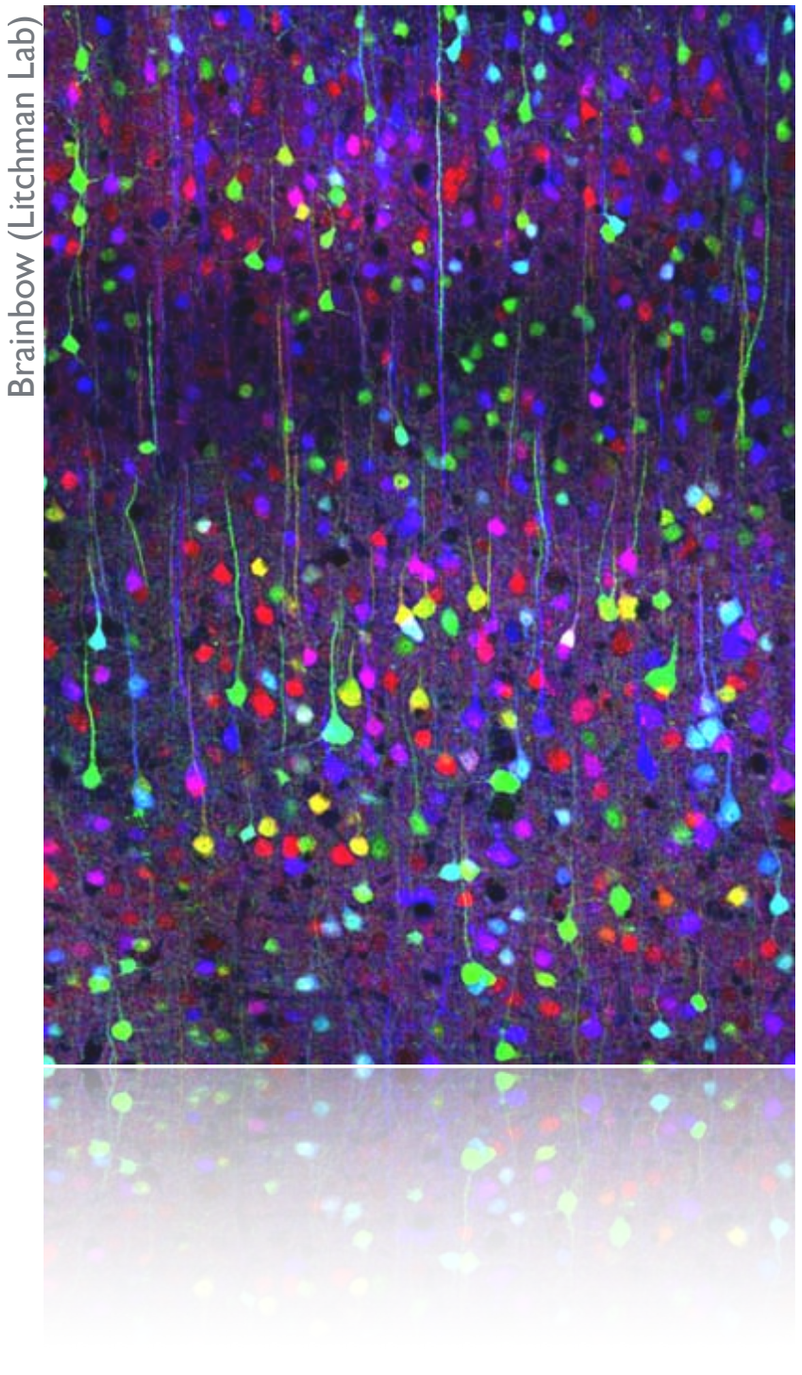
Information Processing & the Brain 2020/2021

University of BRISTOL

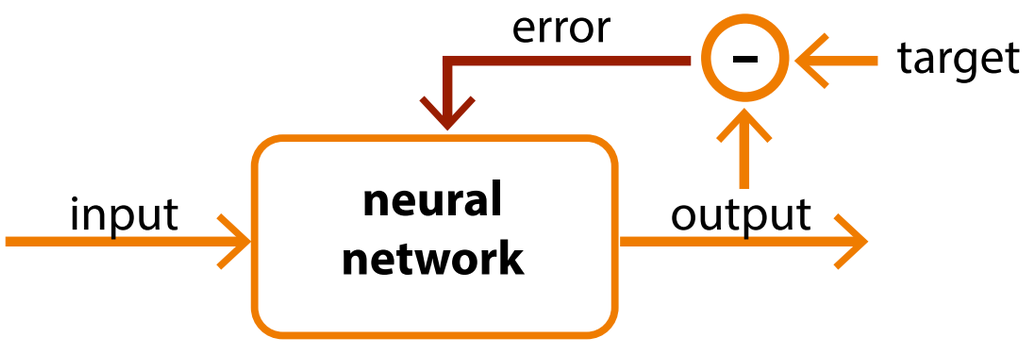


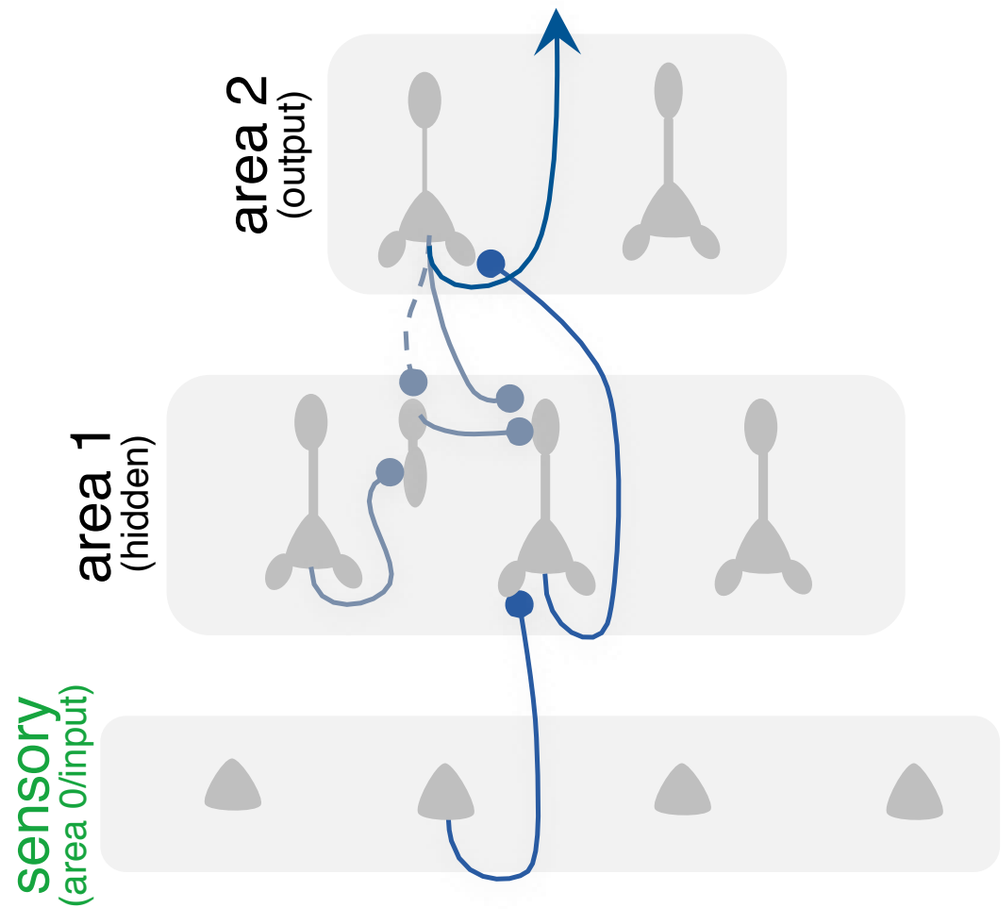
Lecture 3
Visual system: deep learning?

Previously on *IPB*…

Supervised learning:

Relies on a teaching signal

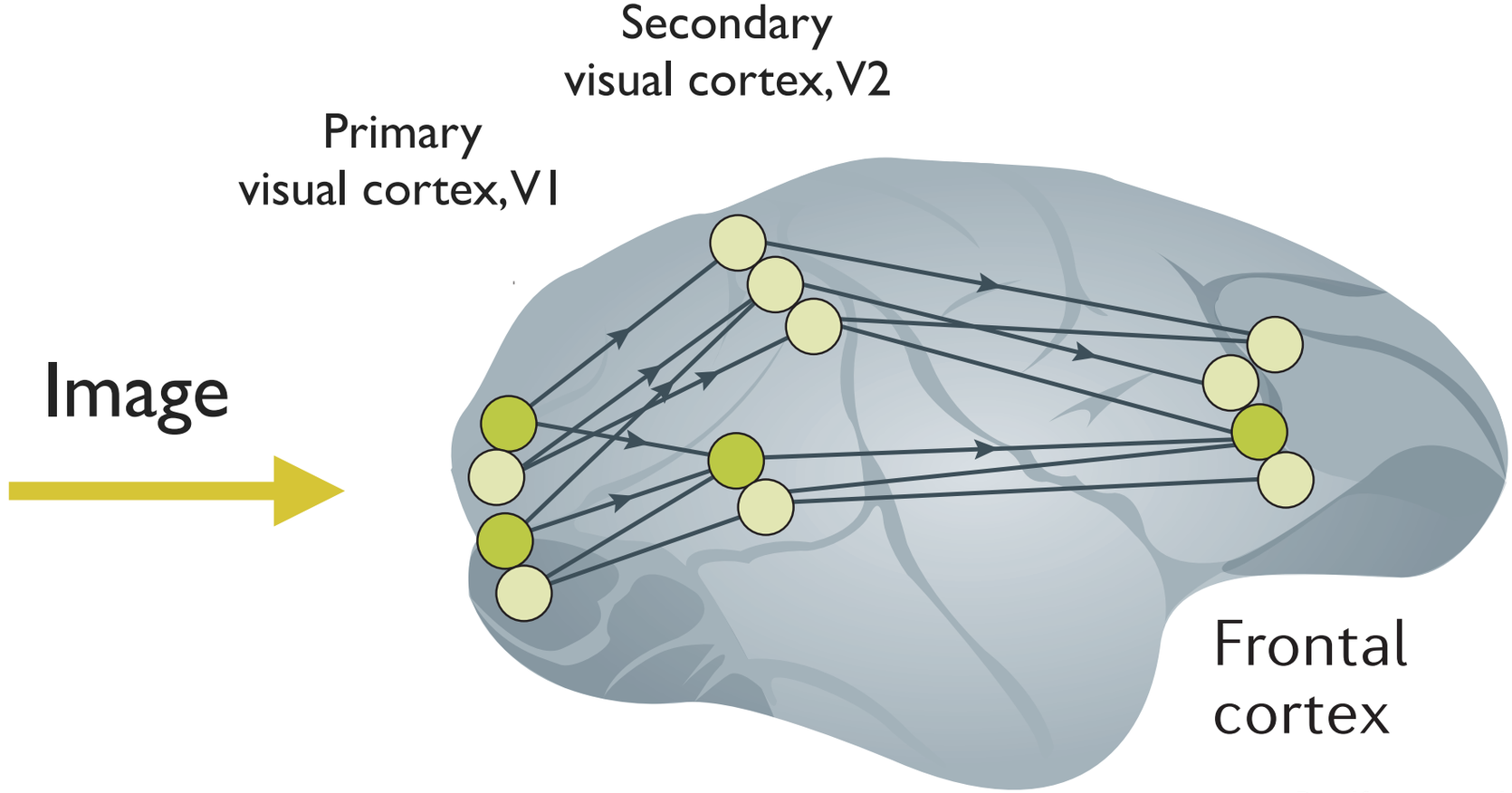




But, what to learn?

How to connect the different elements in the brain!

for example: the visual cortex



Roelfsema et al., Nature Neuroscience Rev (2018)

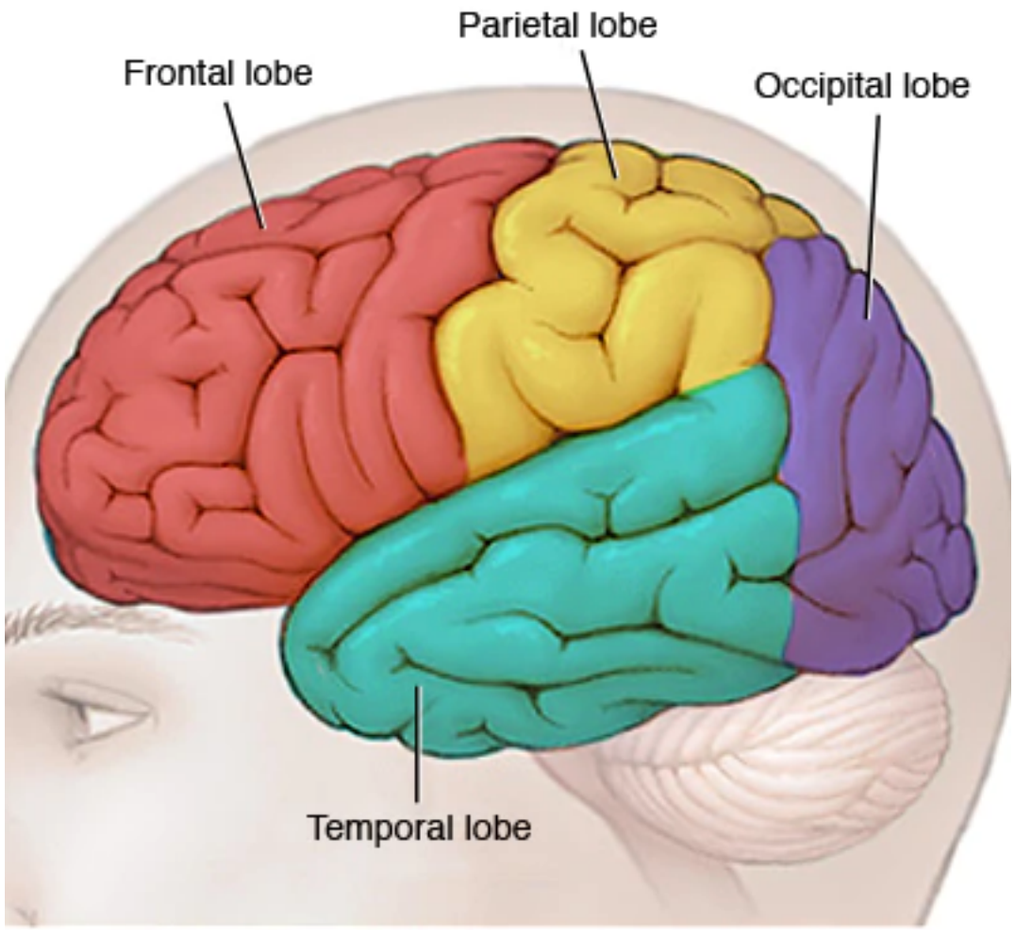
Outline

1. **Visual cortex: structure**

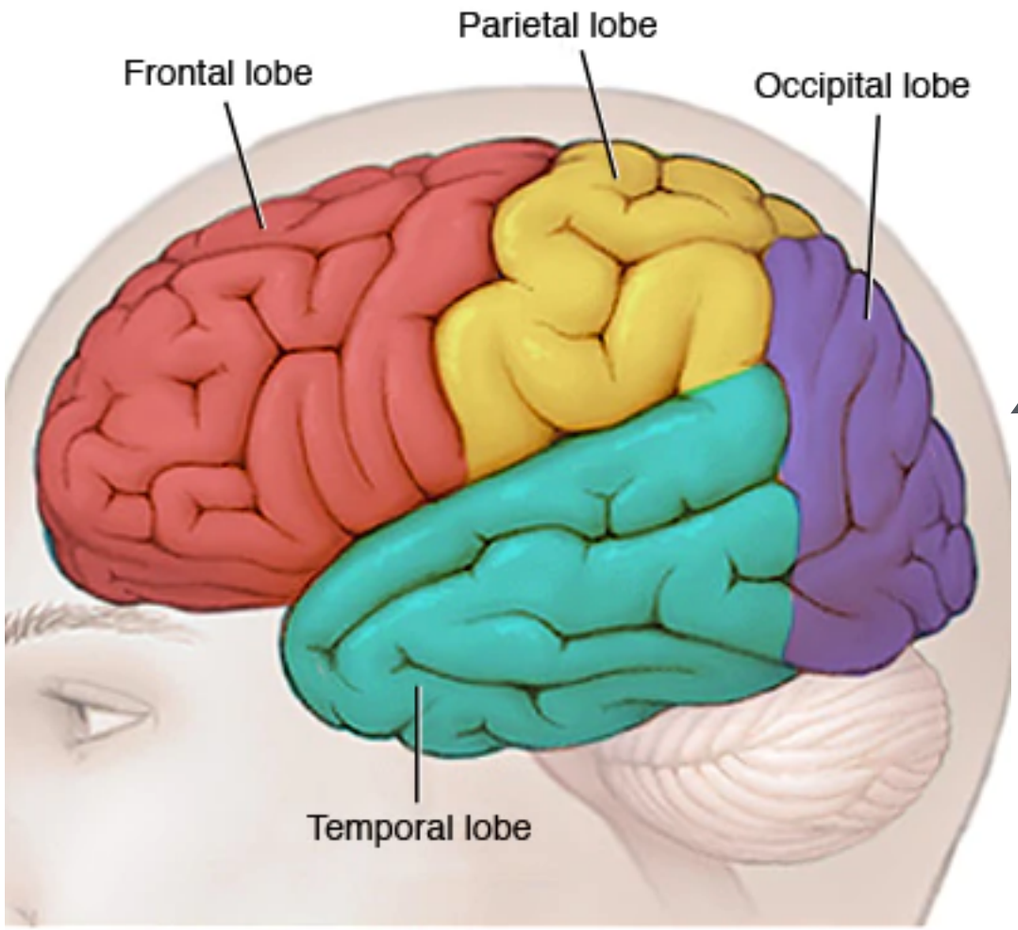
2. **Visual cortex: receptive fields**

3. **Convolutional neural networks as models of visual system**

Brain structure

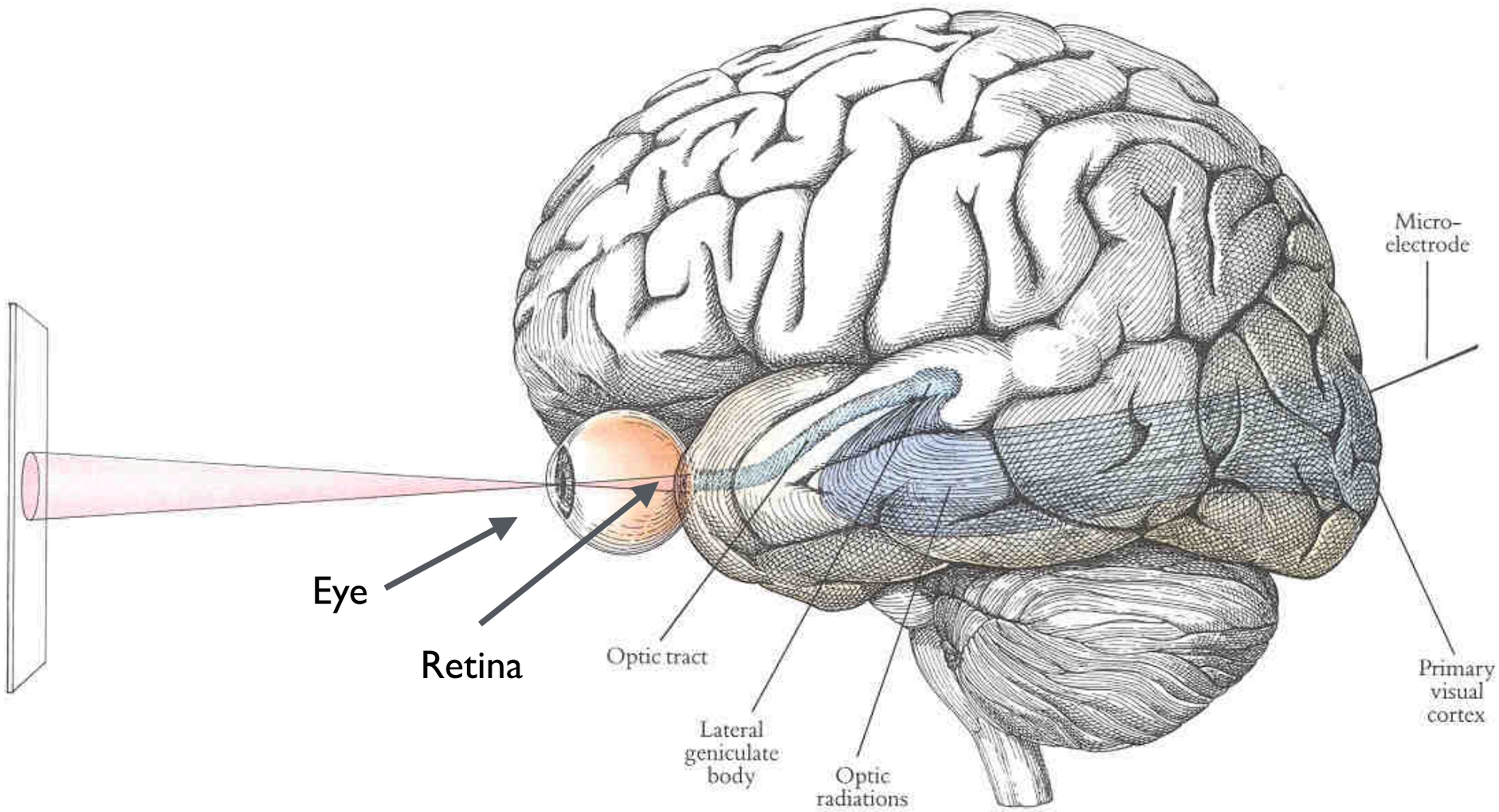


The visual cortex



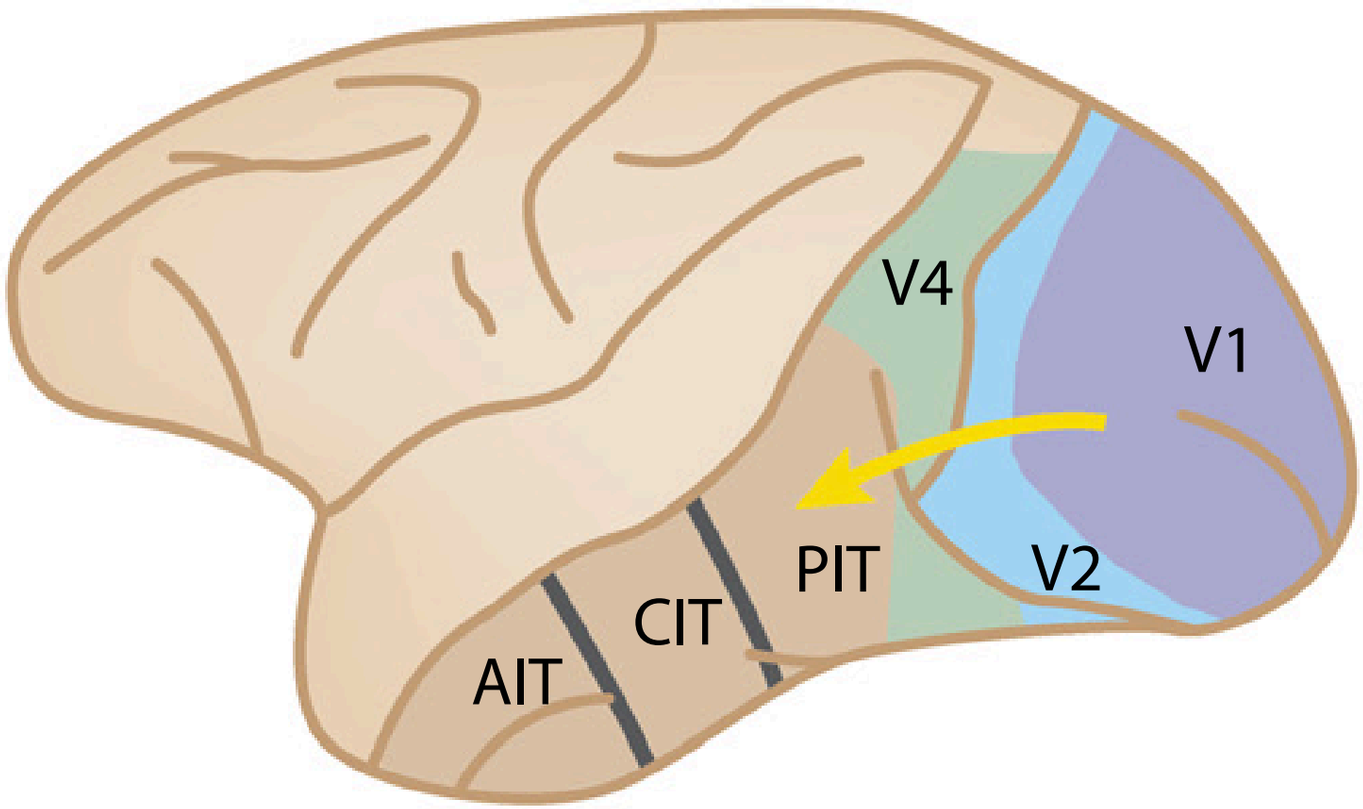
Visual cortex

The visual system



[http://www.cog.brown.edu/courses/cg00001/lectures/visualpaths.html]

Visual cortex: multiple sub-areas

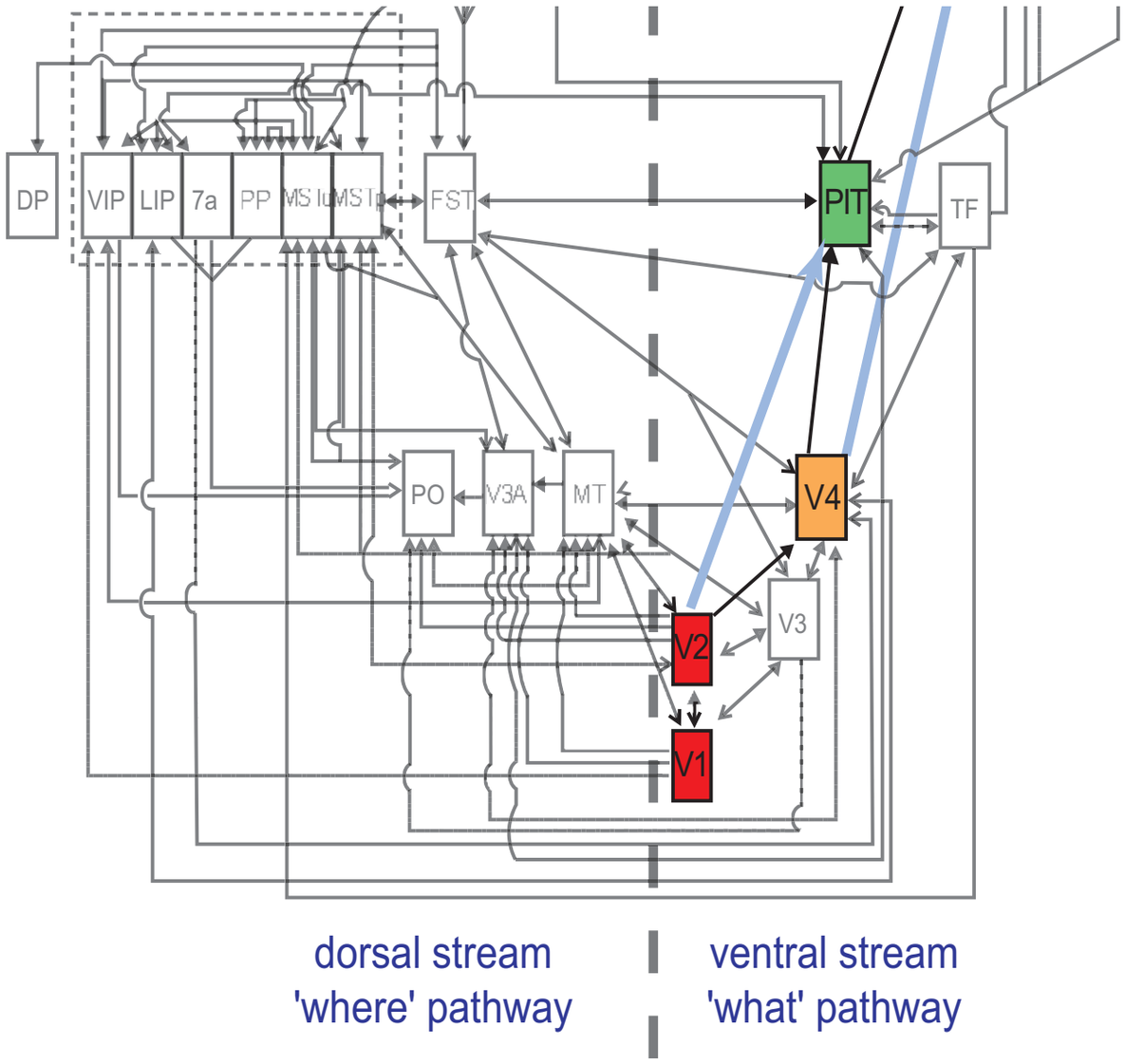


- V1: Primary visual cortex
- V2: Secondary visual cortex

- IT: Inferior temporal cortex

Yamins and DiCarlo, Nature Neuroscience (2016)

Visual cortex: what and where pathways

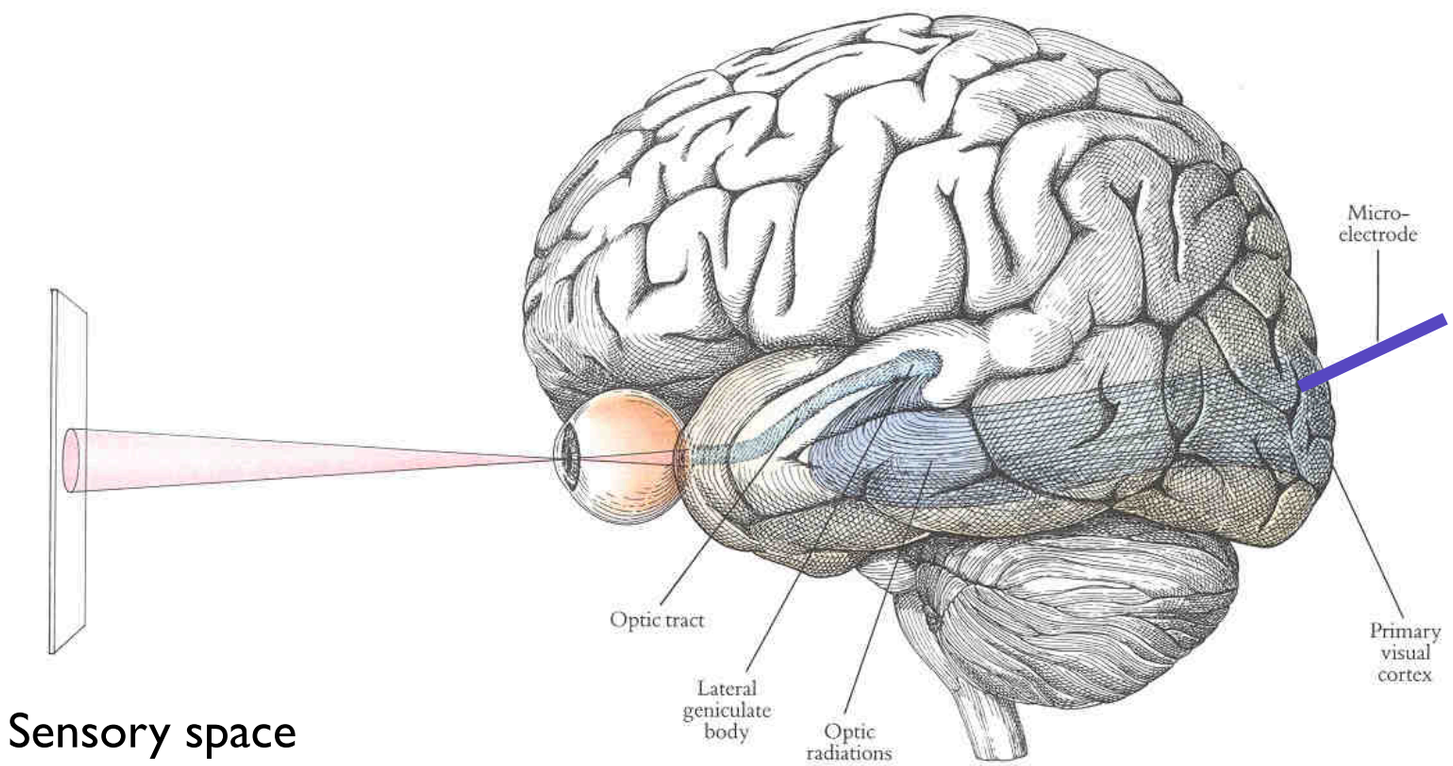


- **'What' pathway:** Relevant to detect what is in the image (e.g. lion vs eagle?)

- **'Where' pathway:** Processes the spatial location of objects (e.g. top vs bottom)

Serre et al., PNAS (2007)

Visual cortex function
receptive fields

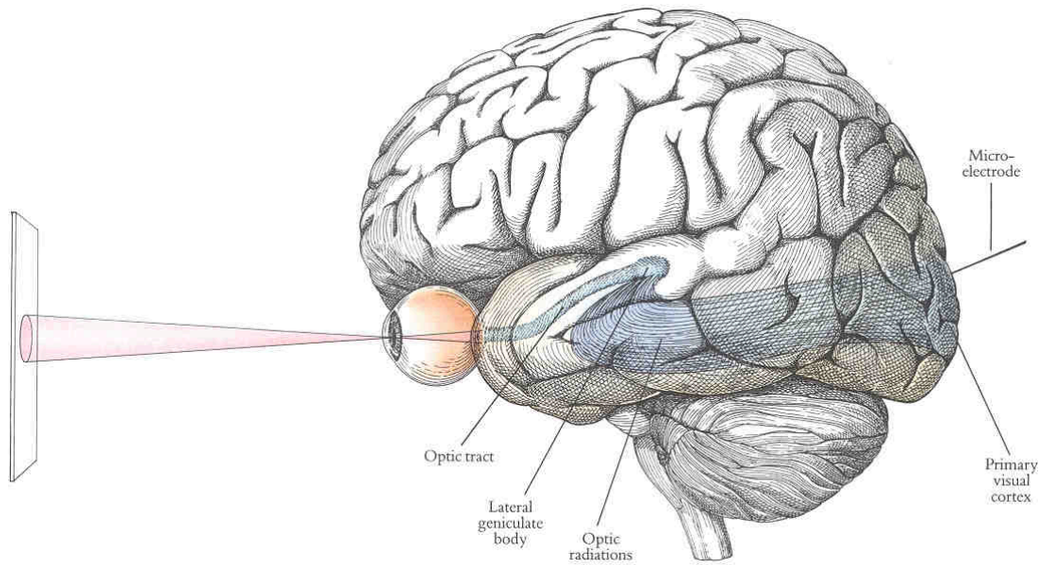


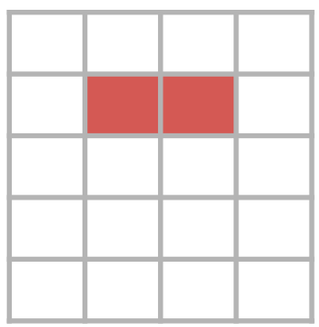
Sensory space
(e.g. visual input)

Electrode used to measure
receptive field

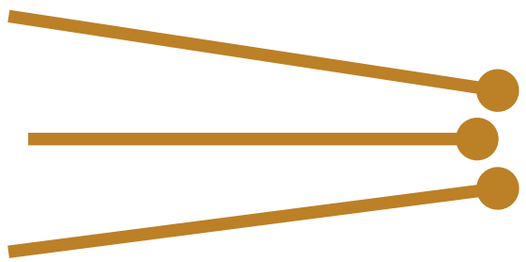
Receptive field: Region of sensory space in which a stimulus modifies (i.e. increases or decreases) the firing of the neuron.

Visual cortex function
a simple model of receptive field





Visual input, ***u***



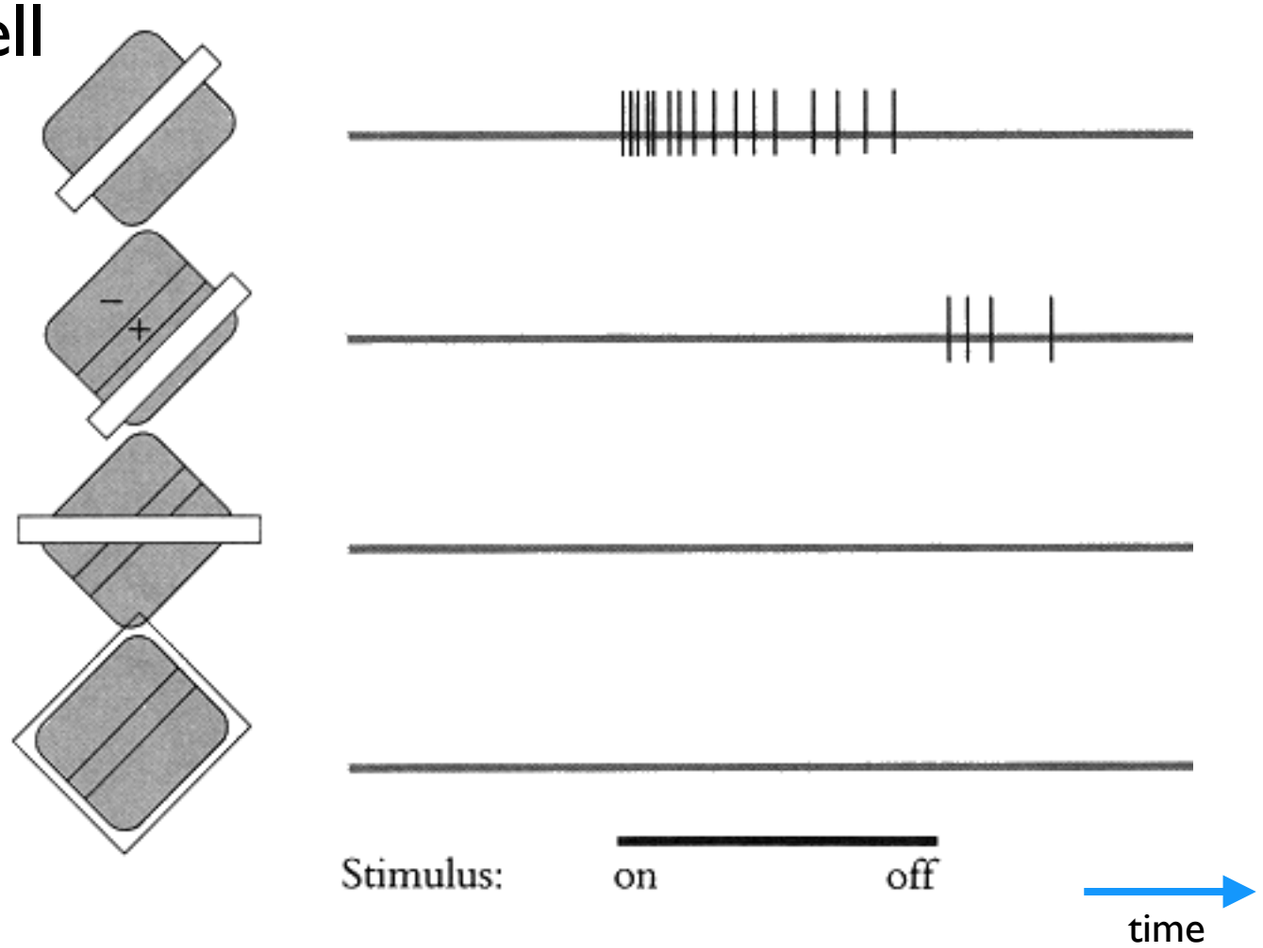
weights, ***w***

Visual cortex
neuron, ***v***

$$v = f(wu)$$

Classical receptive field:
simple cells in primary visual cortex

Simple cell



Video of Hubel & Wiesel experiments:
https://youtu.be/8VdFf3egwfg

Hubel and Wiesel, J Physiol (1962)

Classical receptive field:
Gabor filter as a model of simple cells

where **w** can be approximated by a Gabor filter:

$$w \sim g(x, y; \lambda, \theta, \phi, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2} \right) \cos\left(2\pi \frac{x'}{\lambda} + \phi \right),$$

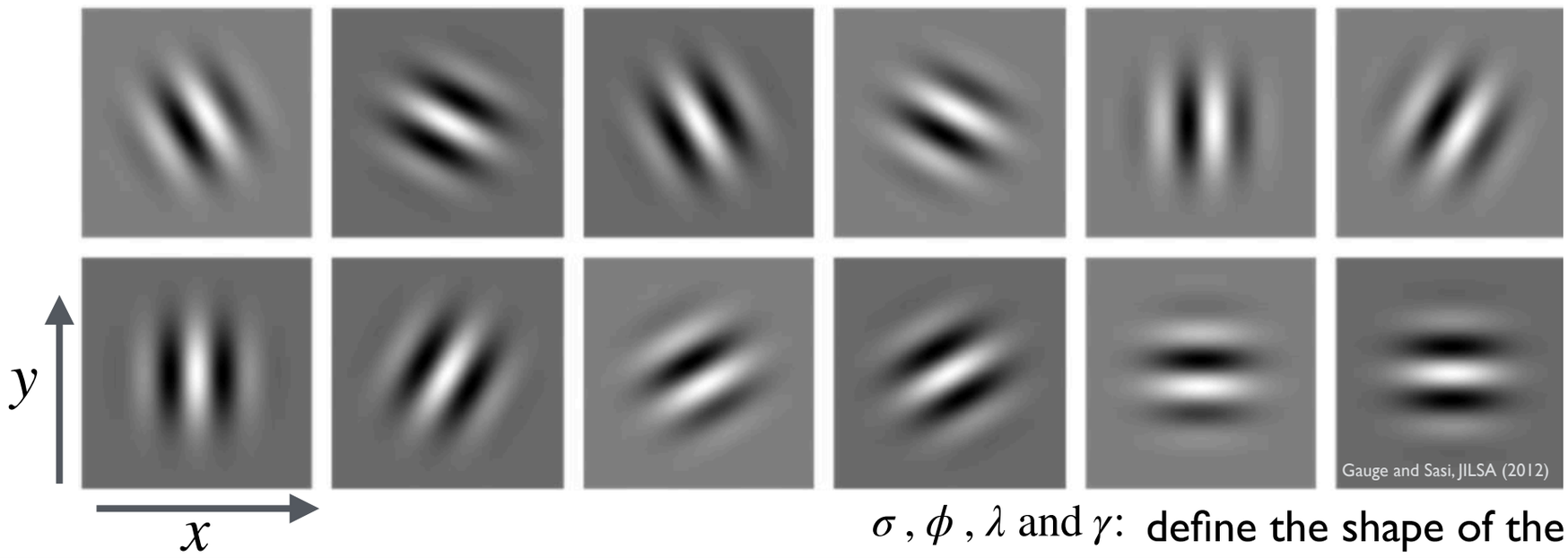
Classical receptive field:
Gabor filter as a model of simple cells

A Gabor filter is a product of a Gaussian and a sinusoid that can model a periodic pattern similar to simple cells.

$$w \sim g(x, y; \lambda, \theta, \phi, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2} \right) \cos\left(2\pi \frac{x'}{\lambda} + \phi \right),$$

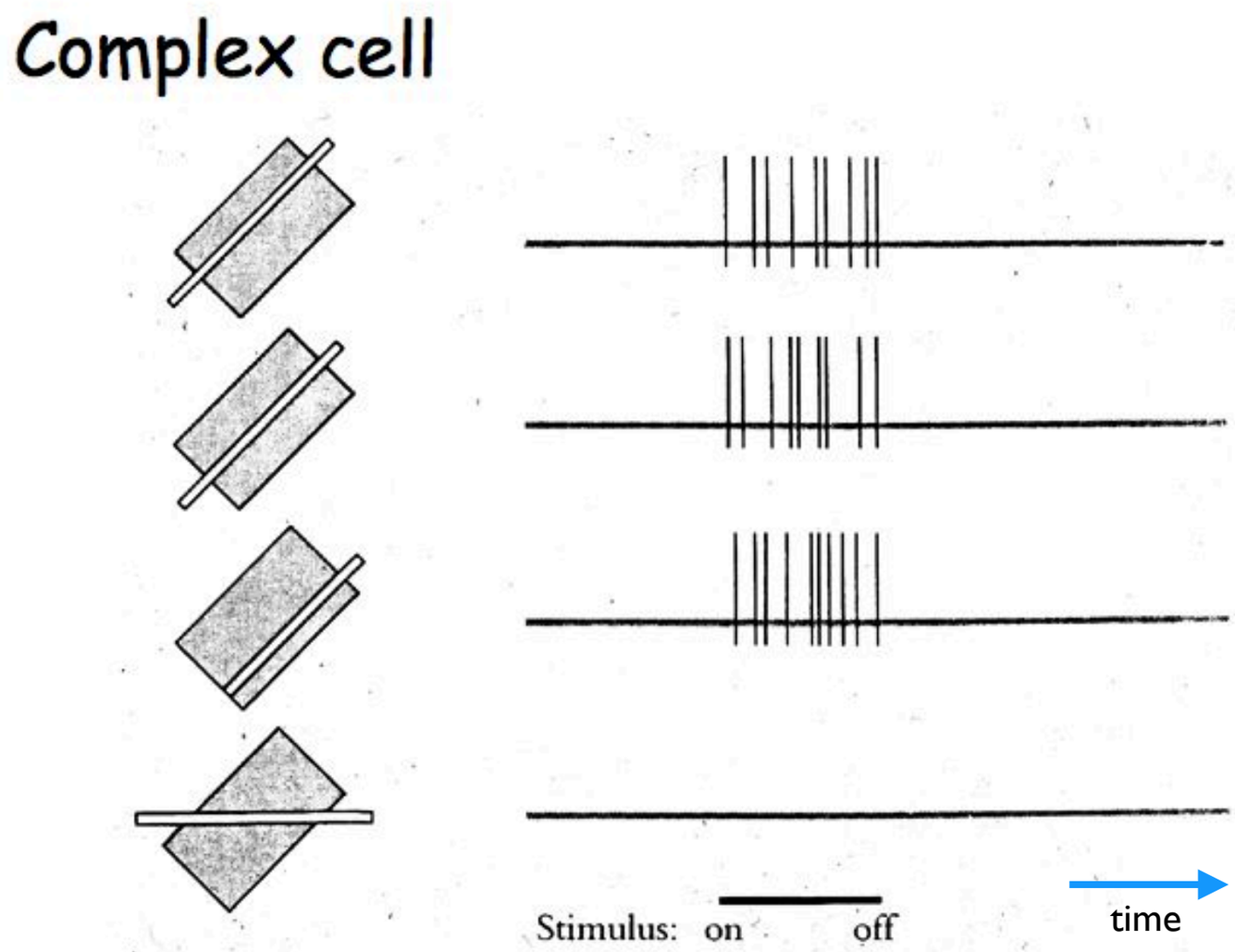
$$y' = -x \sin\theta + y \cos\theta .$$
$$x' = x \cos\theta + y \sin\theta,$$

Gabor filters for different angles θ :



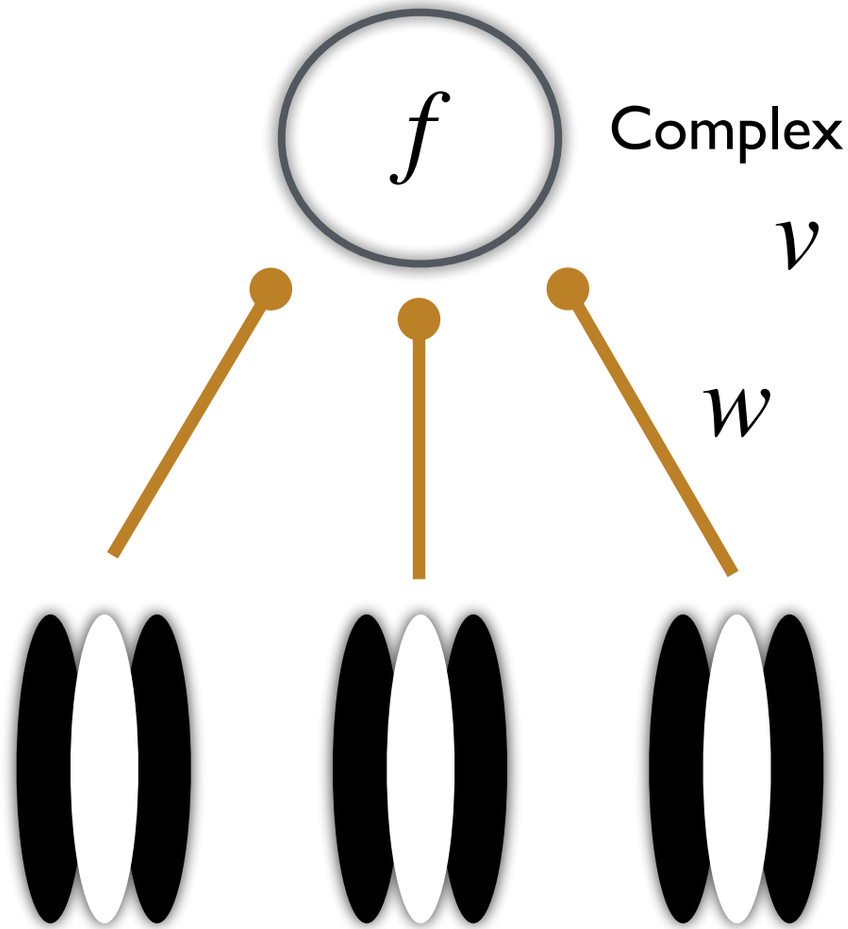
σ , ϕ , λ and γ: define the shape of the Gabor filter.

Visual cortex: complex cells



Hubel and Wiesel, J Physiol (1962)

Visual cortex: complex cells, a model



Set of **simple cells** with the same orientation but different locations

$$g(x, y; \lambda, \theta, \phi, \gamma)$$

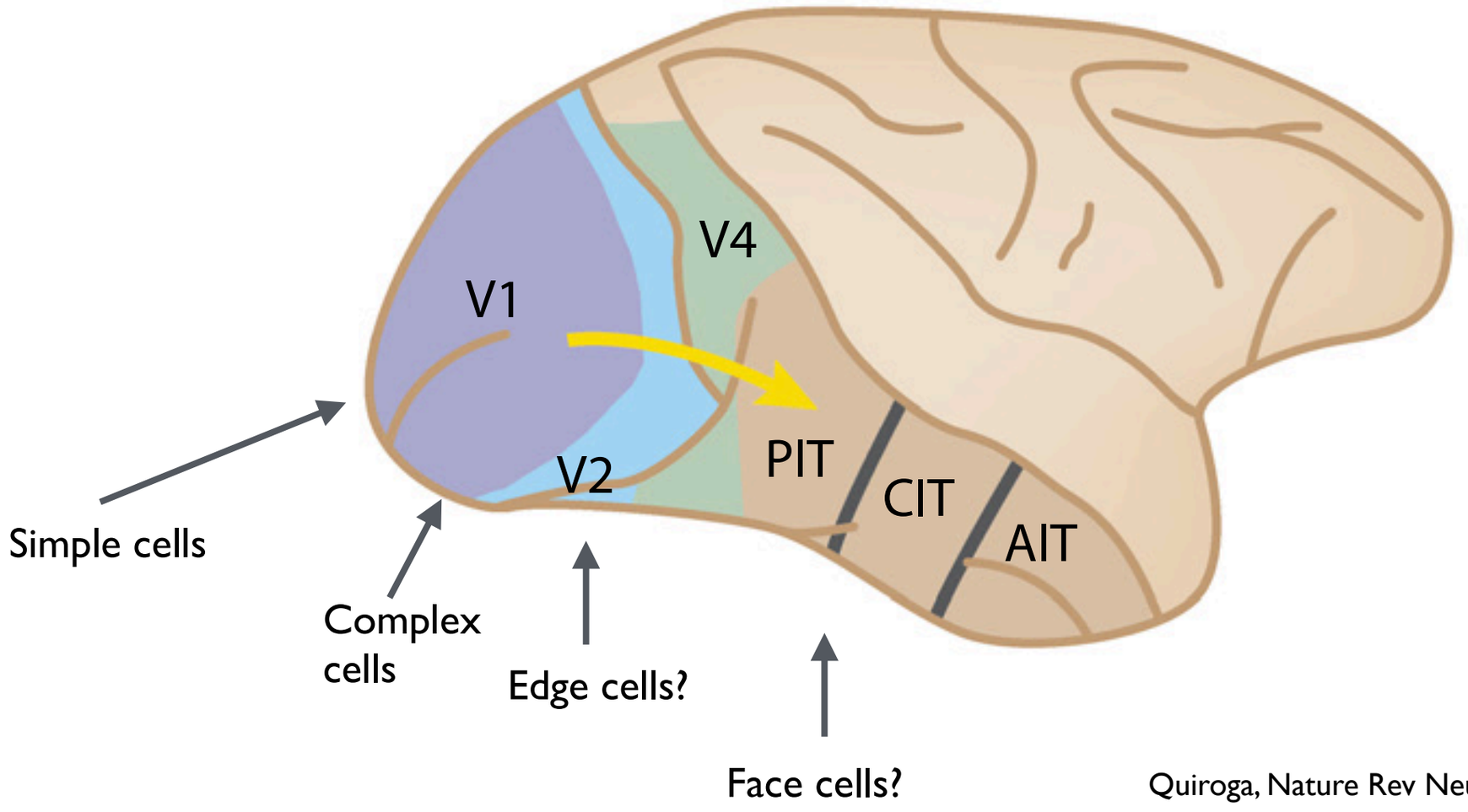
E.g.:
A set of Gabor filters

To think about..

- **What other receptive fields/features would be useful for the visual cortex to encode?**

Hint: look around you for common visual features

Visual cortex:
Higher order representations



Quiroga, Nature Rev Neurosci (2012)

Yamins and DiCarlo, Nature Neuroscience (2016)

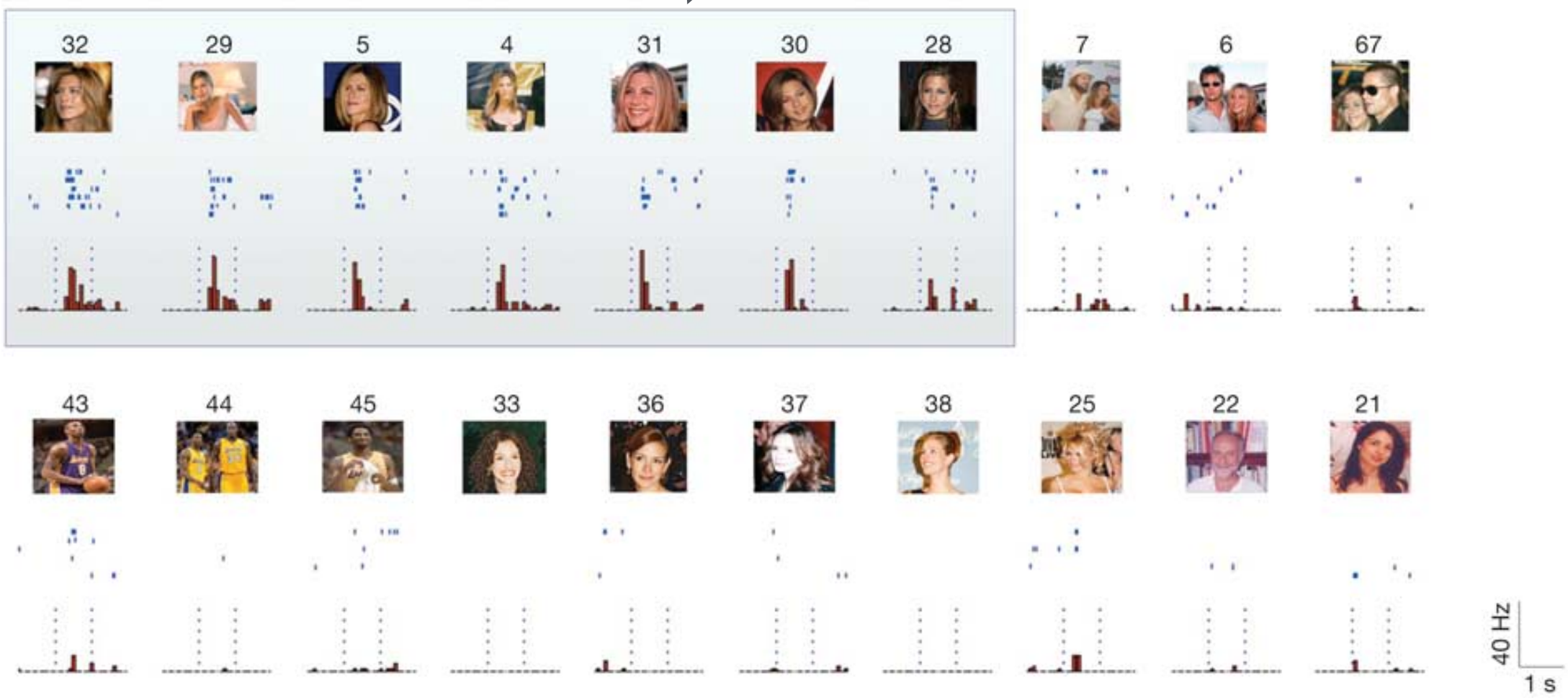
Visual cortex:

Grandmother cell (e.g. the Jennifer Aniston cell)

Grandmother cell: The idea that here are neurons responding specifically when you see people that you are familiar with.

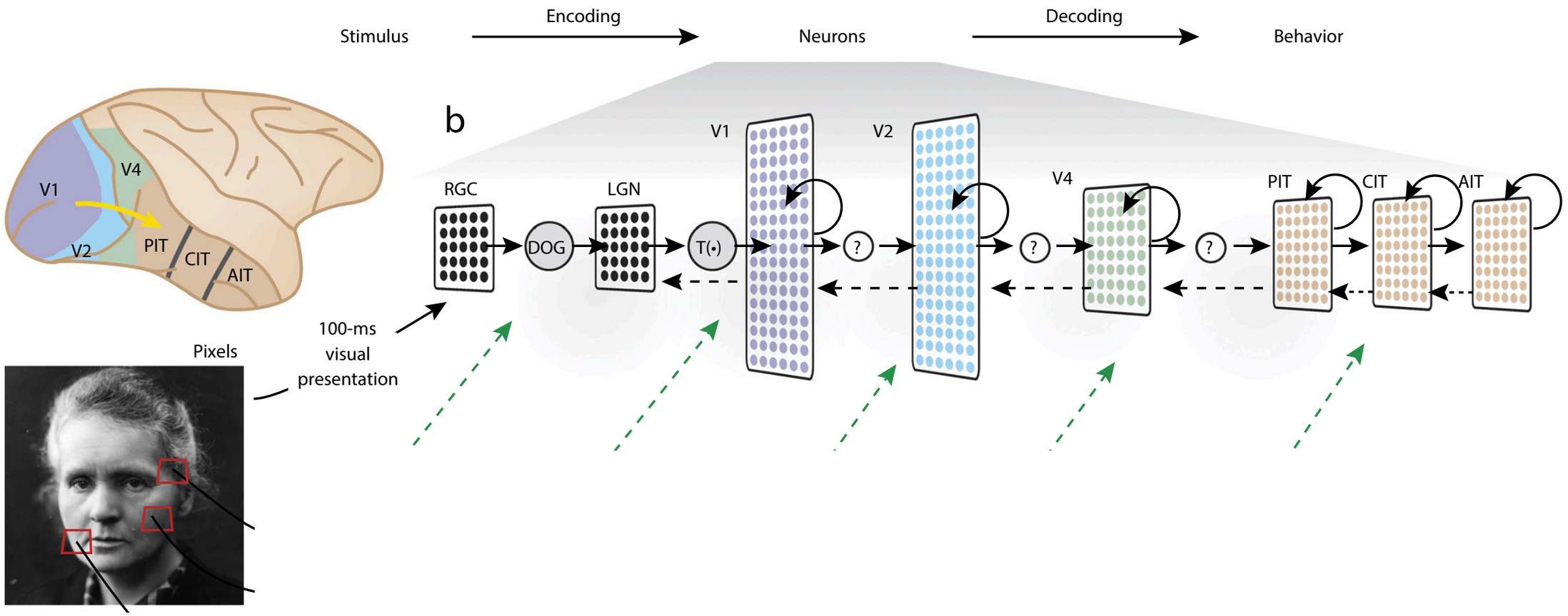
A neuron responding to images of Jennifer Aniston recorded in Medial Temporal Lobe in humans (epilepsy patients):

Quiroga et al., Nature (2005)



Convolutional neural networks (CNNs)

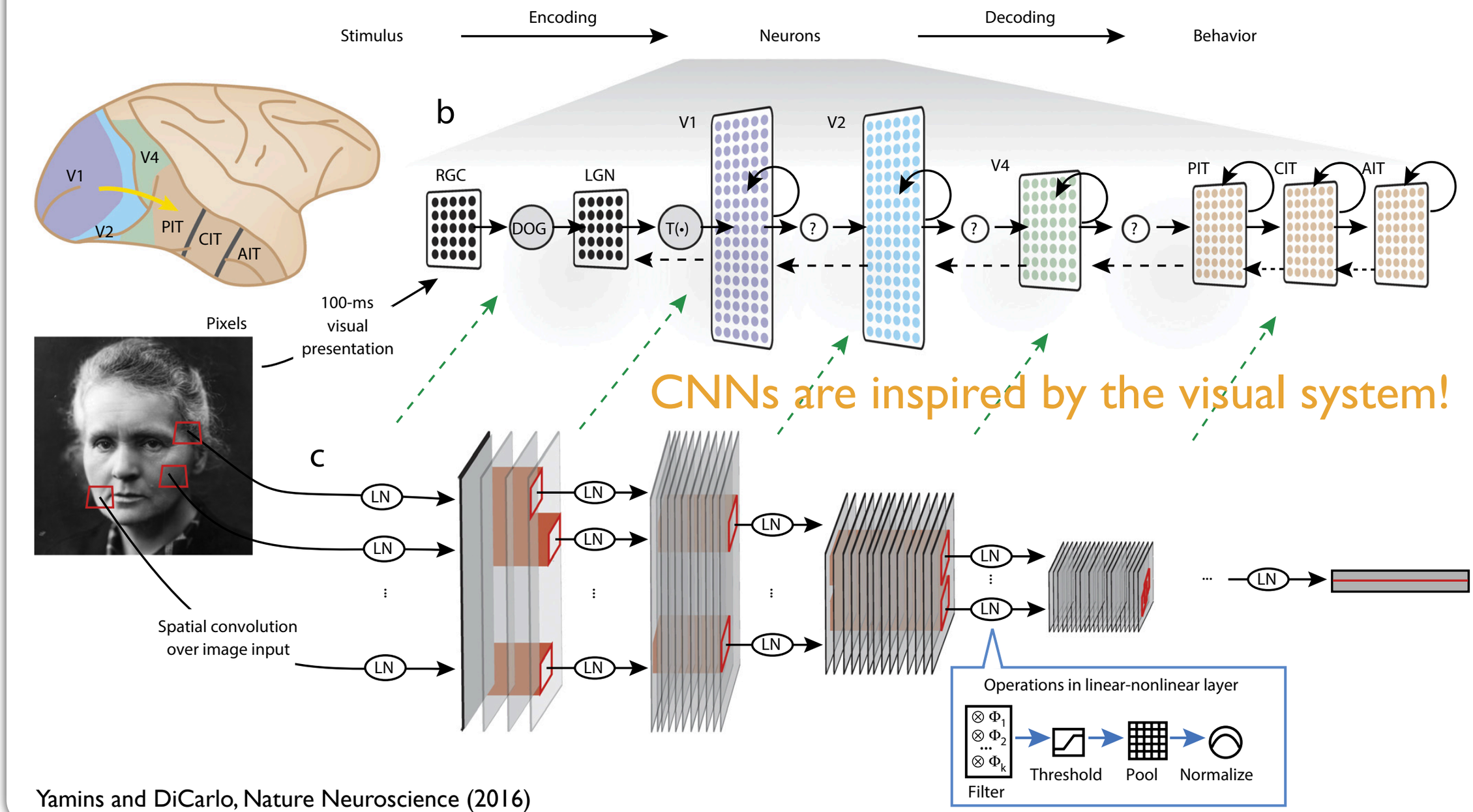
as a model of the visual system



Yamins and DiCarlo, Nature Neuroscience (2016)

Convolutional neural networks (CNNs)

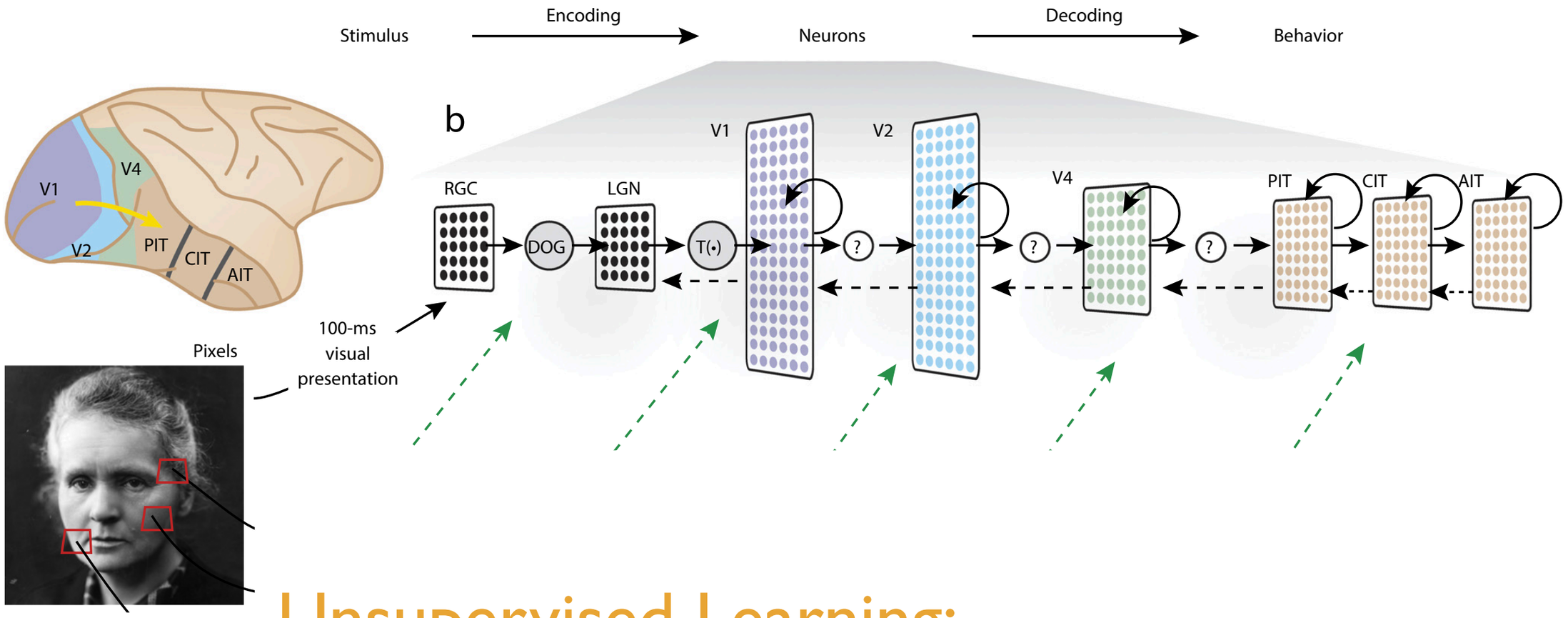
as a model of the visual system



Yamins and DiCarlo, Nature Neuroscience (2016)

Convolutional neural networks (CNNs)

as a model of the visual system



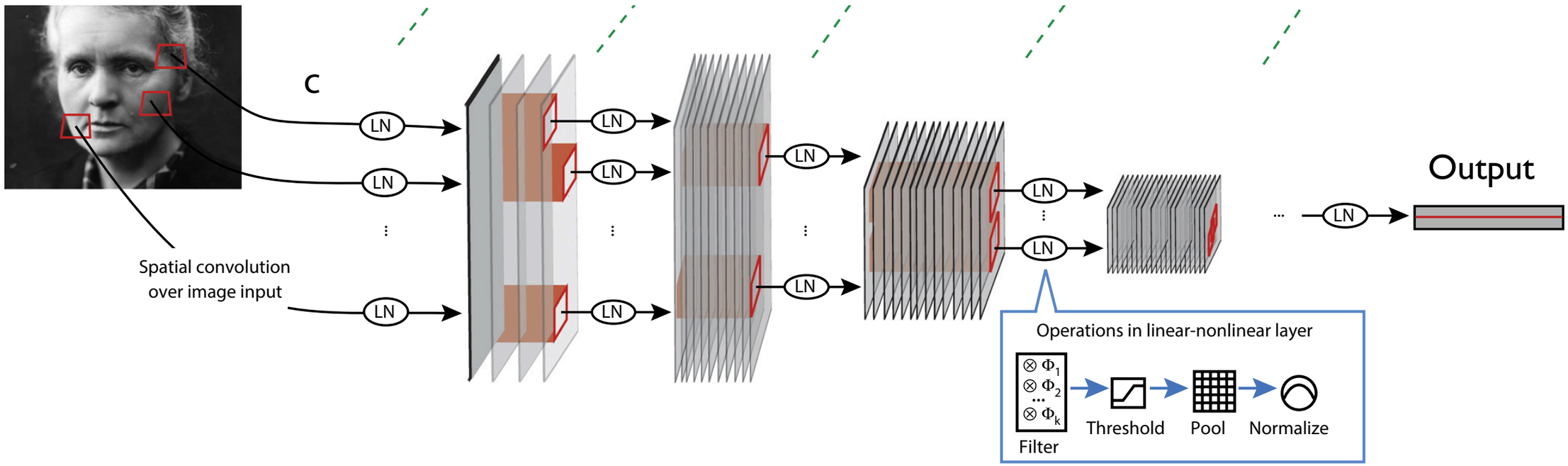
Unsupervised Learning:

In neuroscience the visual cortex is traditionally thought of learning useful representations of the environment.

Yamins and DiCarlo, Nature Neuroscience (2016)

Convolutional neural networks (CNNs)

as a model of the visual system

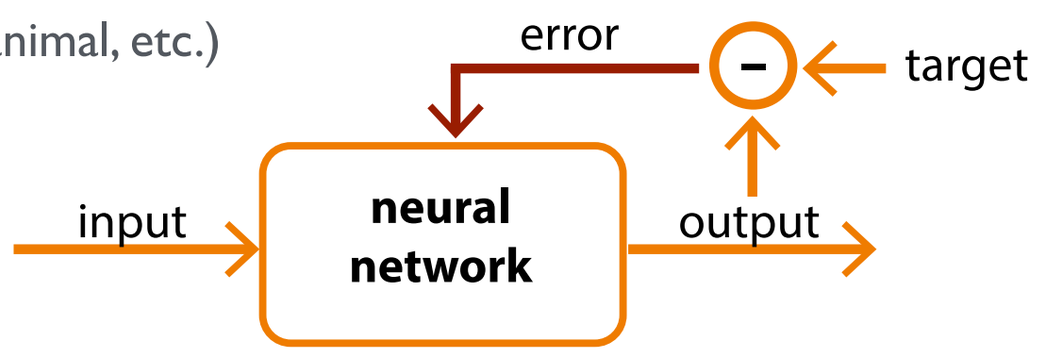


But CNNs rely on Supervised Learning:

CNNs are trained with the backprop algorithm (see prev. lecture)

Target: Image label/category (e.g. person, animal, etc.)

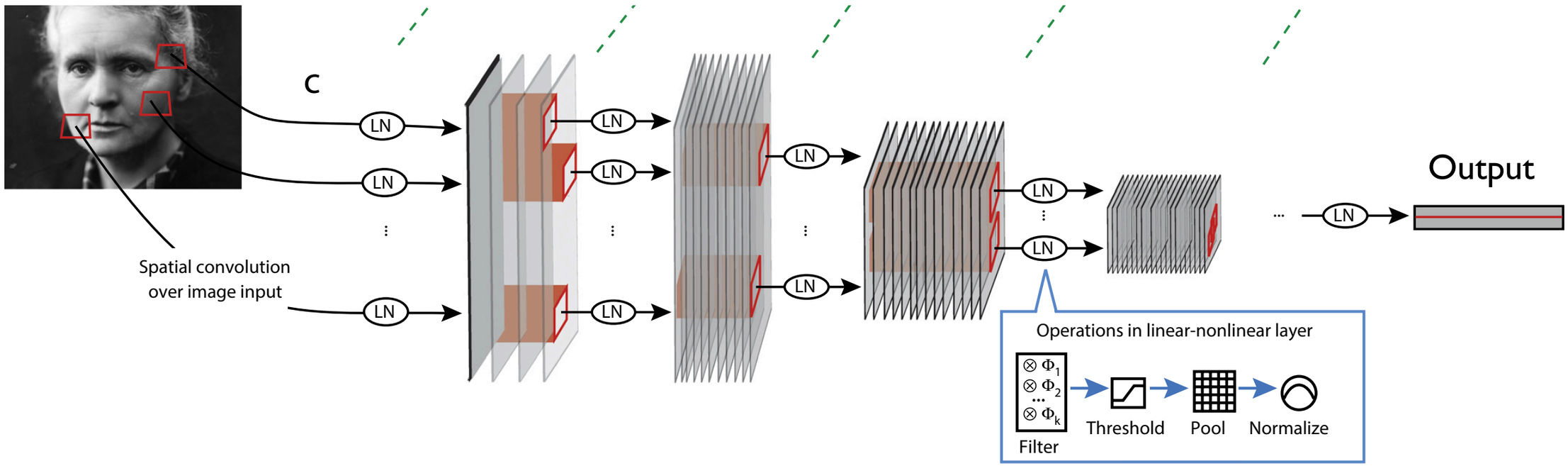
CNNs have revolutionised computer vision!



Yamins and DiCarlo, Nature Neuroscience (2016)

Convolutional neural networks (CNNs)

as a model of the visual system



Two key operations:

Spatial convolution: Scanning the image with a filter (red squares)

Pooling: Subsampling input (e.g. max operation: selecting most active unit)

Others: Thresholding and normalisation

Yamins and DiCarlo, Nature Neuroscience (2016)

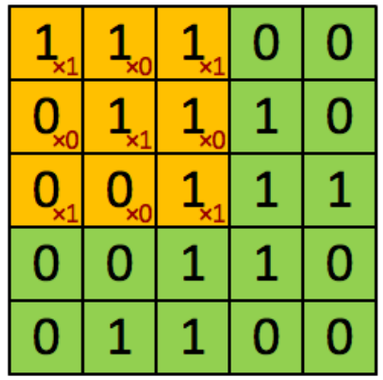
Convolutional neural networks (CNNs)

Convolution operator

Example of filter:

1	0	1
0	1	0
1	0	1

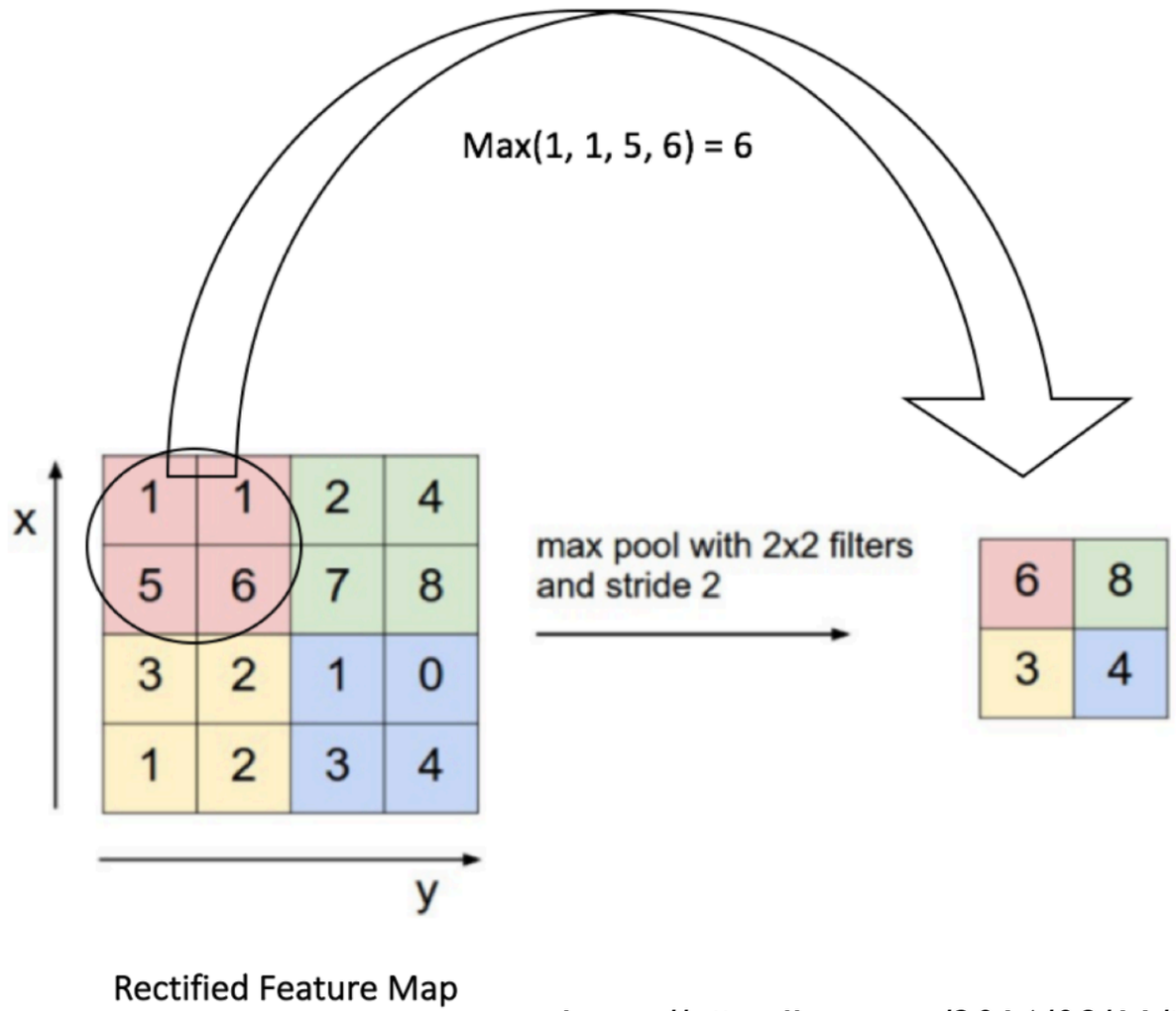
Convolutional operation:



https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets

Convolutional neural networks (CNNs)
Pooling

(Max) Pooling:



https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets

Convolutional neural networks (CNNs)
as a model of the visual system

A comparison between CNNs and the visual system:

CNNs (machine learning)	**Visual system** (neuroscience)	**Example model**
Filter	Simple cells	w
Convolution	Complex cells	$w * u$
Thresholding	Neuron Threshold	ReLU: $\boldsymbol{v}=\max(w * u, 0)$
Pooling/Subsampling	Connectivity between and within brain areas	$\max(\boldsymbol{v})$ average($\boldsymbol{v}$)
Layers	Brain areas (e.g. V1, V2, etc.)	$f(w_2 * f(w_1 * u))$

$*$: convolution operator

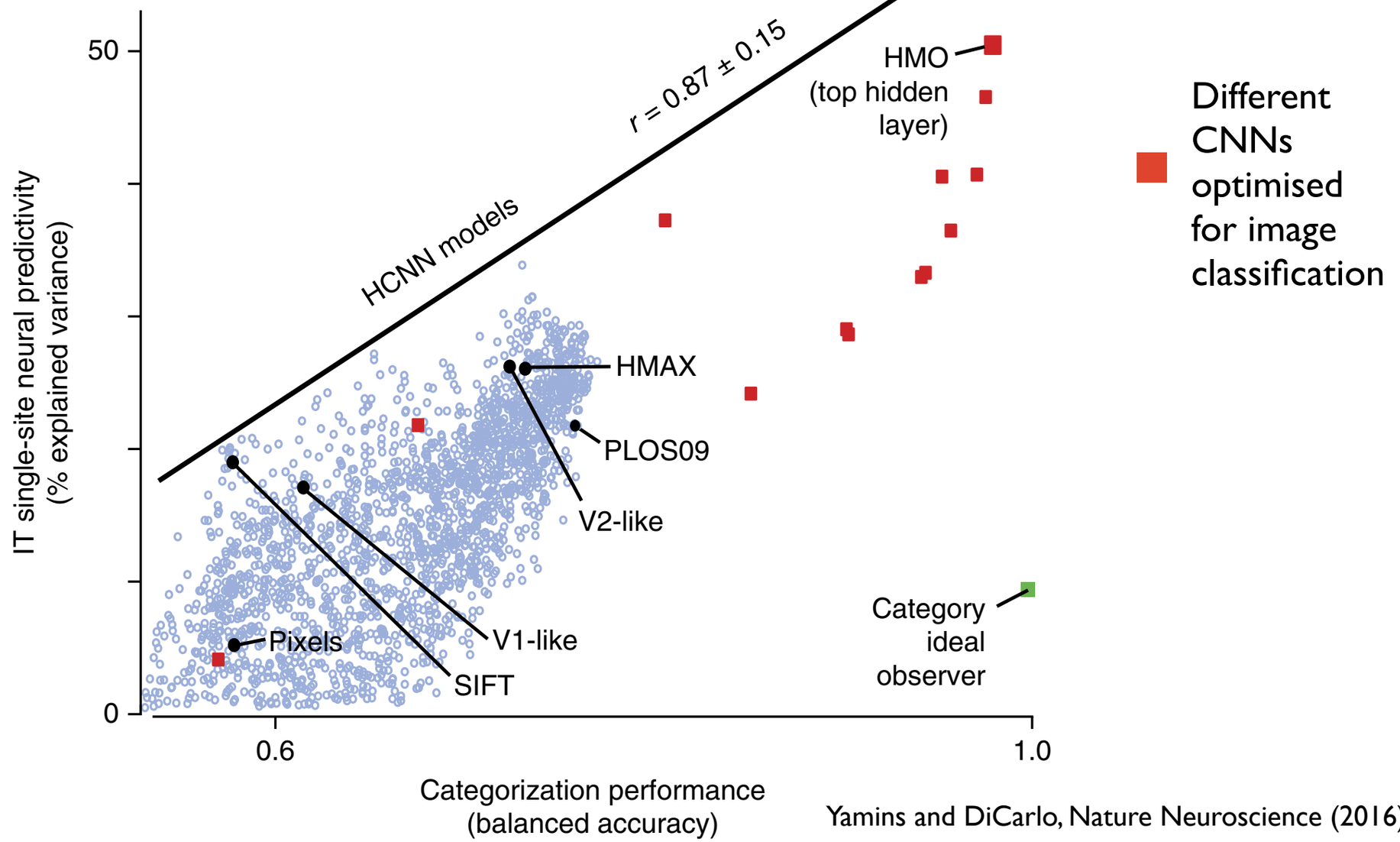
Quiz time!

Please go to BB
and solve quiz 3.

It should take you just a couple of minutes.

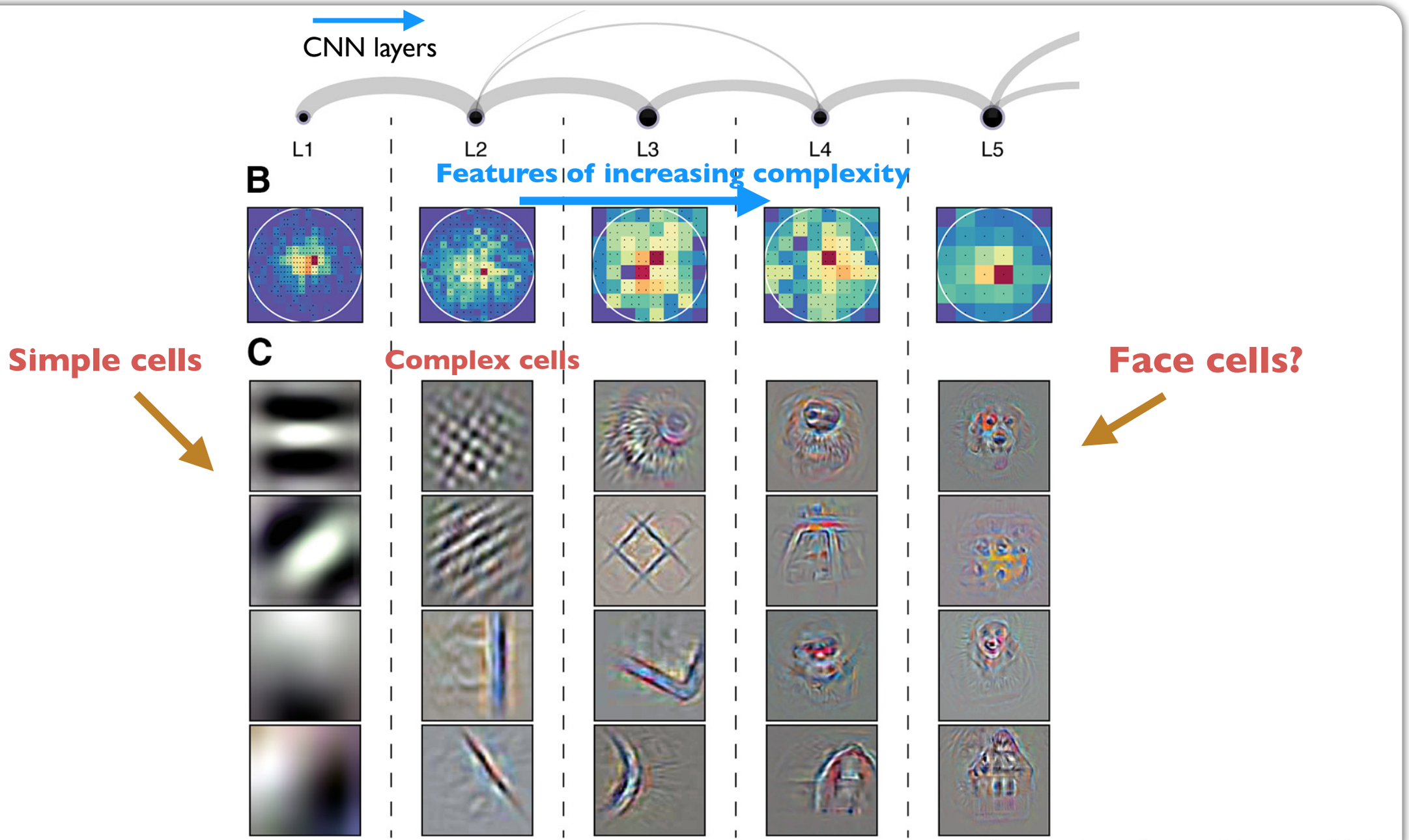
CNNs performance predicts visual cortex responses

CNNs were trained to classify natural images.
Their performance predicts their ability to explain real neural responses.



Yamins and DiCarlo, Nature Neuroscience (2016)

CNNs learn hierarchical features similar to the ones found in the brain



Guclu and van Gerven, JNeurosci (2015)

Summary

1. **Visual cortex: spread across multiple areas (V1, V2..)**
2. **Receptive fields of increasing complexity: simple, complex, face and grandmother cells**
3. **Mathematical models of simple and complex cells (e.g. gabor filters)**
4. **Convolutional neural networks as hierarchical models of the visual system**

Questions?

reddit.com/r/comsm0075/

References

Text books:

Theoretical neuroscience, Dayan and Abbott (Chapter 2: receptive fields in the visual system)

Deep Learning, Courville, Goodfellow and Bengio

Recommended: https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets

Relevant papers:

- Yamins and DiCarlo, Nature Neuroscience (2016) (review on how well CNNs capture neural data)
- Hubel and Wiesel, J Physiol (1962) (seminal work measuring receptive fields in visual cortex)
- Quiroga, Nature Rev Neurosci (2012) (review on concept/grandmother cells in the cortex)

Upcoming lectures

- L1[4]: Neural circuits and learning: introduction
- L2[4]: Supervised learning & backprop
- L3[5]: Visual system: deep learning?
- **L4[5]: Reinforcement learning**
- L5[6]: Unsupervised learning
- L6[6]: Temporal processing
- L7[7]: Recurrent neural networks
- L8[7]: Guest lecture