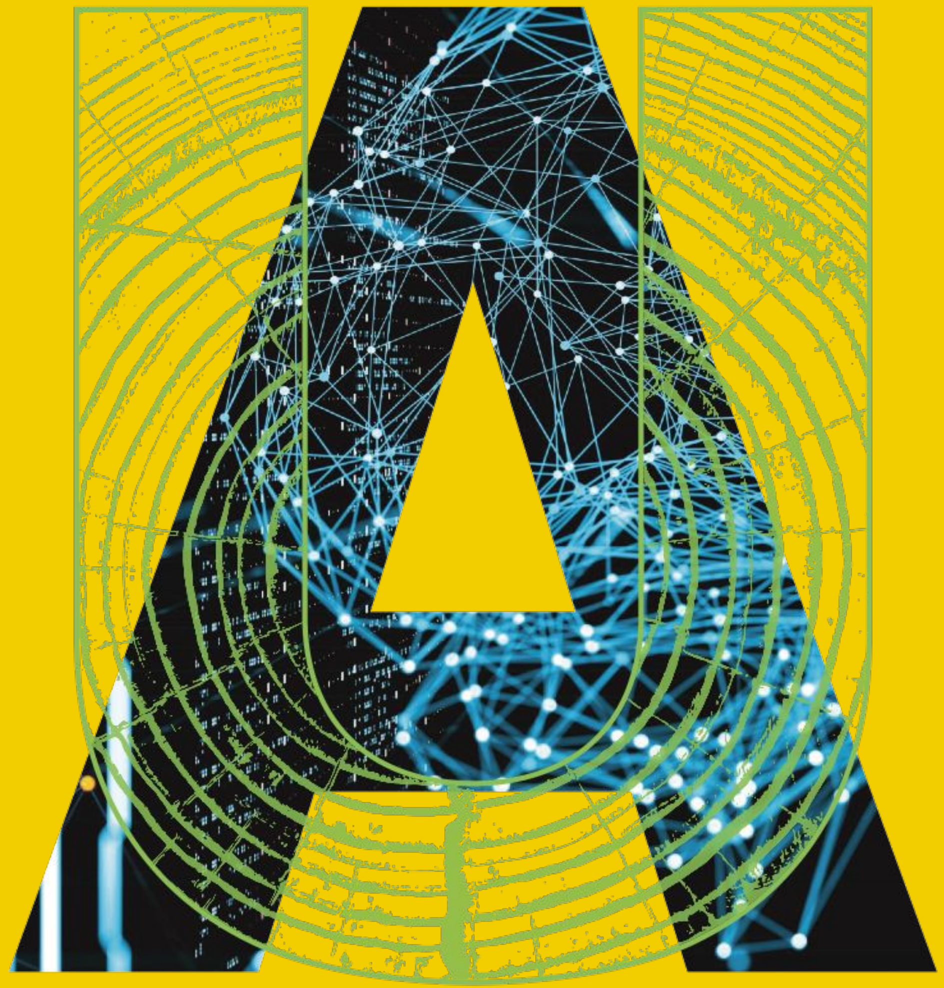
MACHINE LEARNING & THE BRAIN

Vision Neuroscience & Vision Models

Thursday 12 September 2023
Alex Murphy

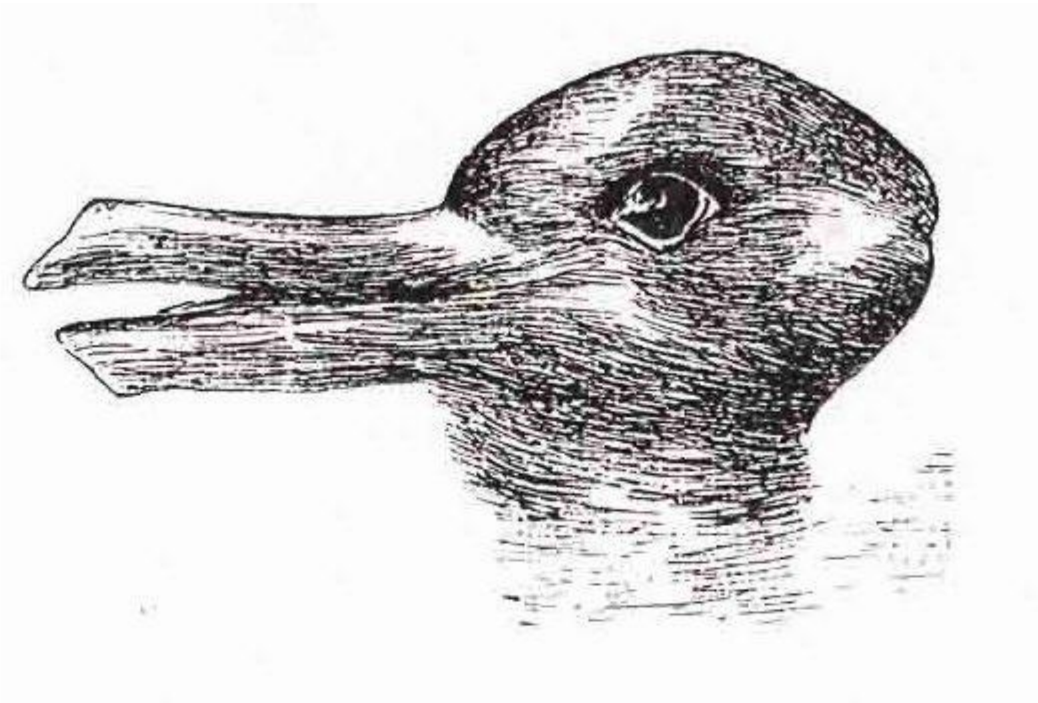
UNIVERSITY OF ALBERTA

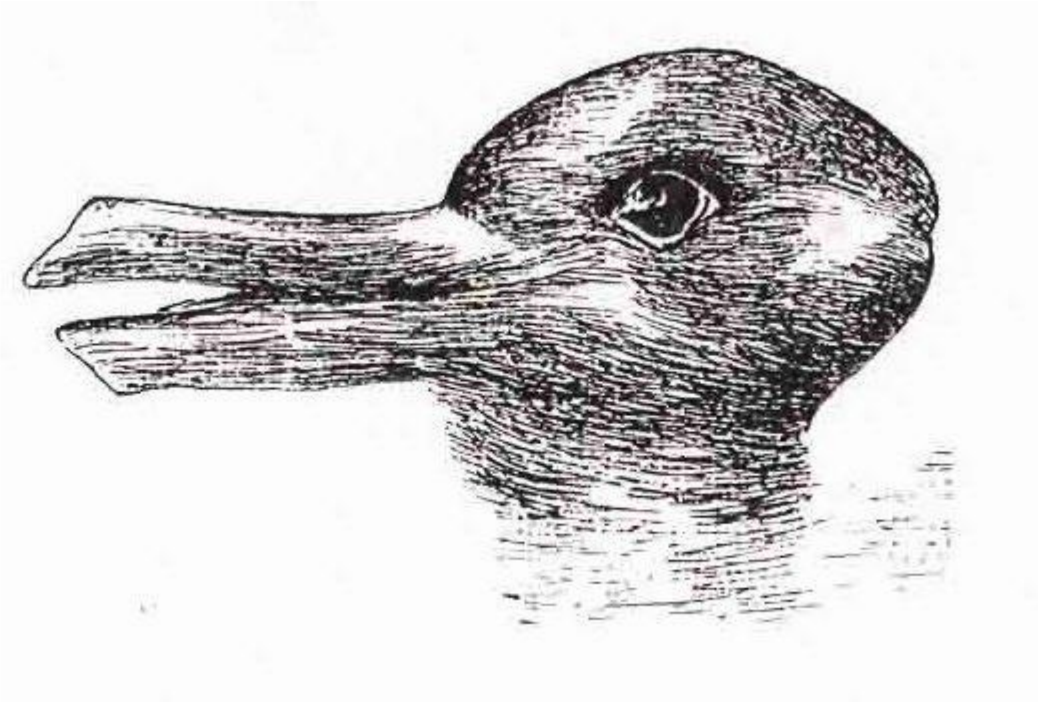


Credit

Dr Kyle Mathewson -> Dr Alona Fyshe -> Dr Alex Murphy

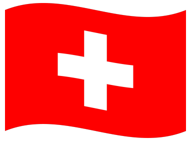
Visual Illusions

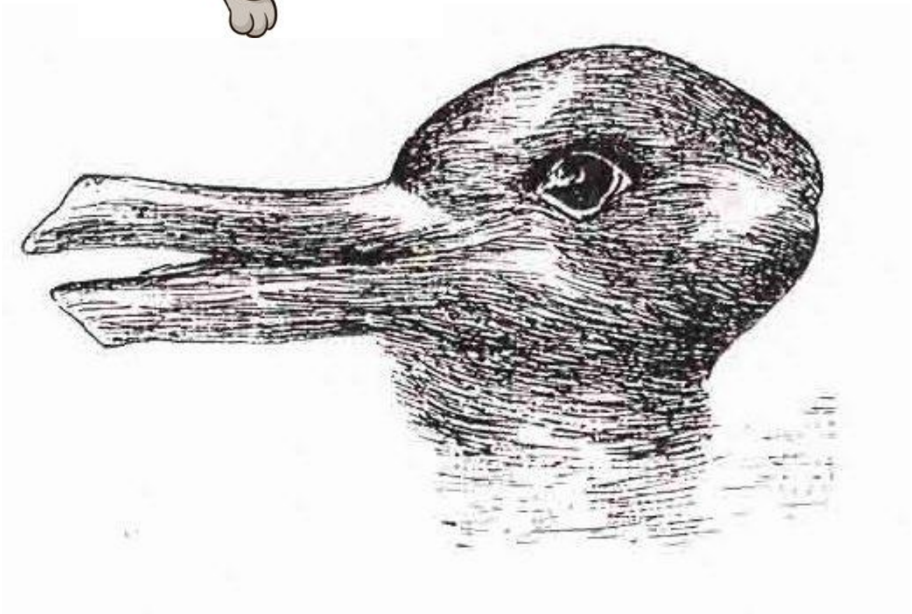


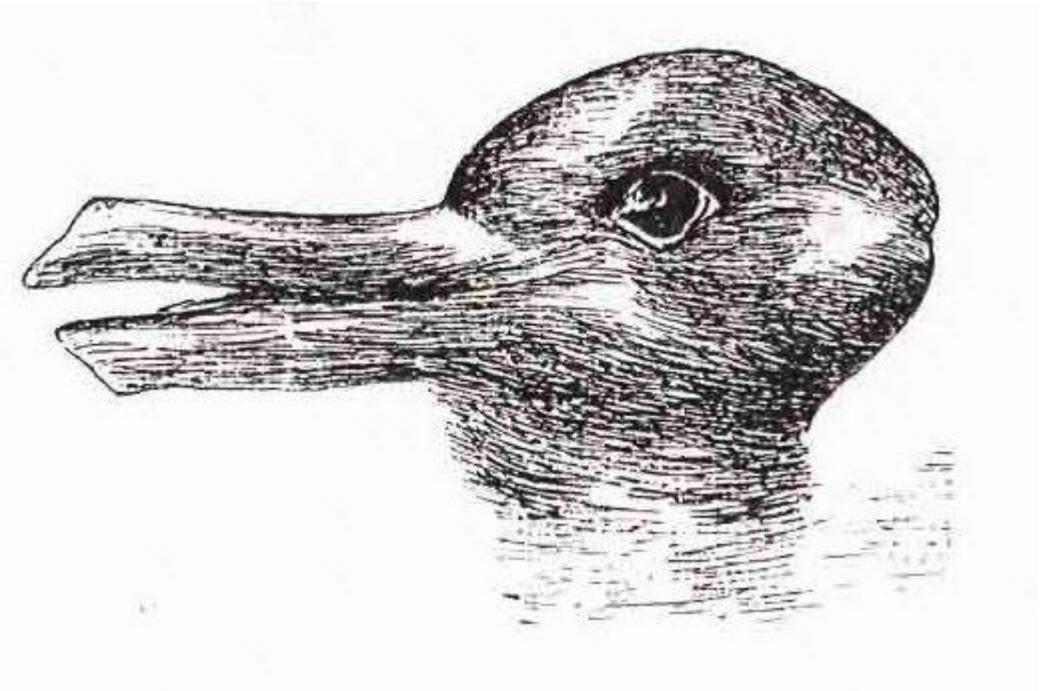


Jensen & Mathewson, 2011, Perception
Mathewson, 2017, Perception

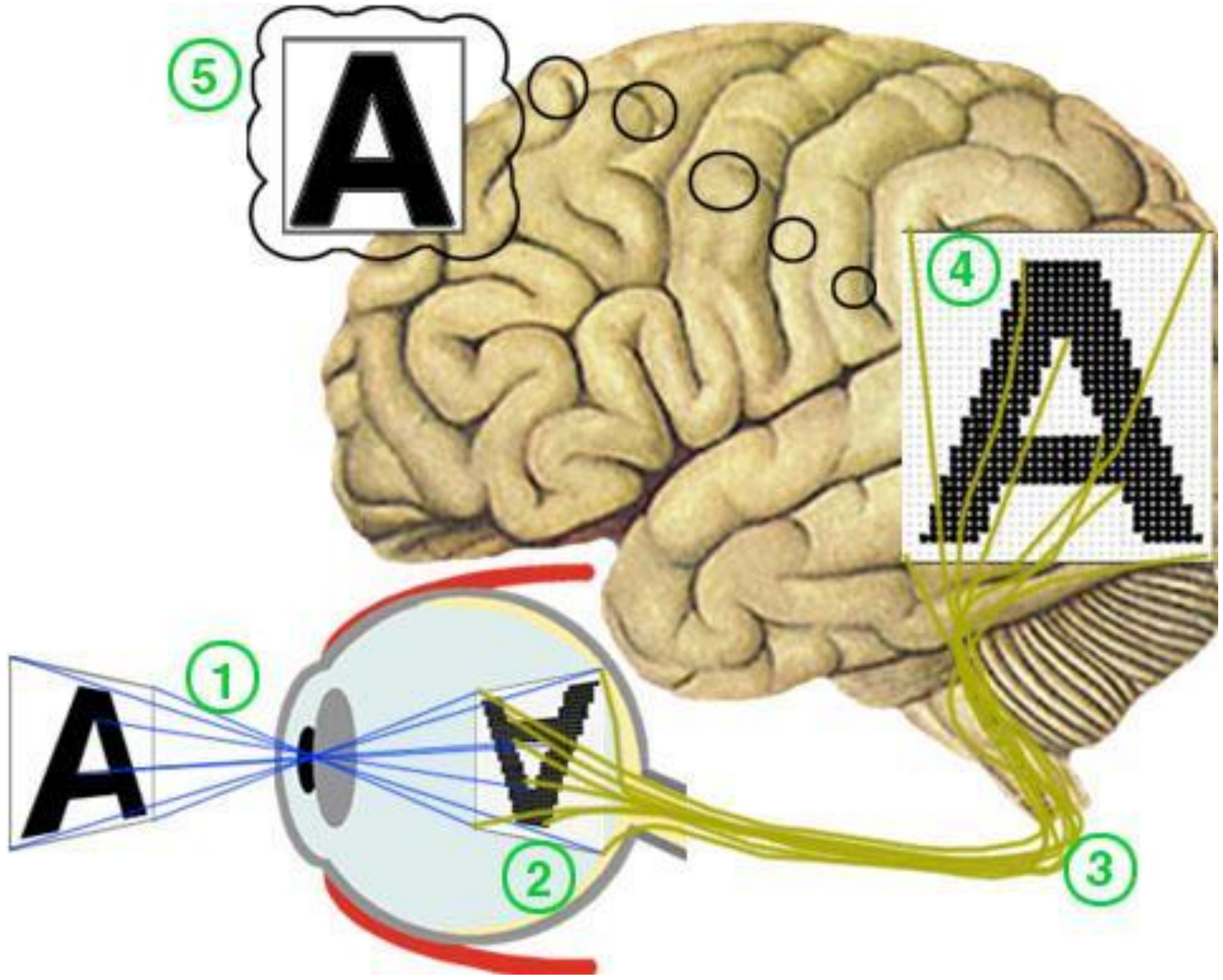
Visual Illusions



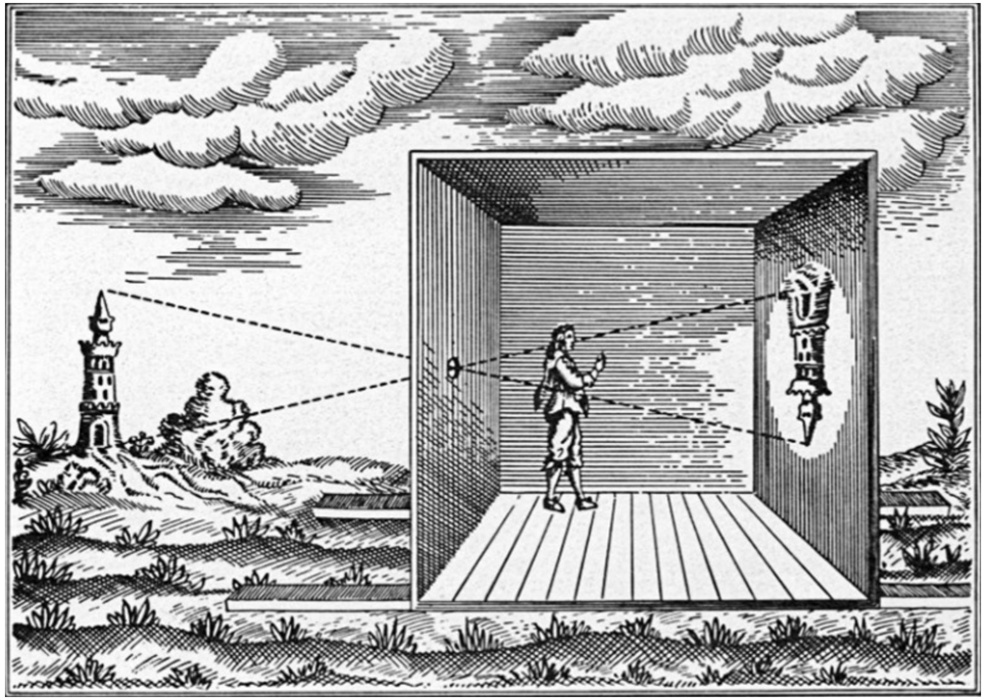


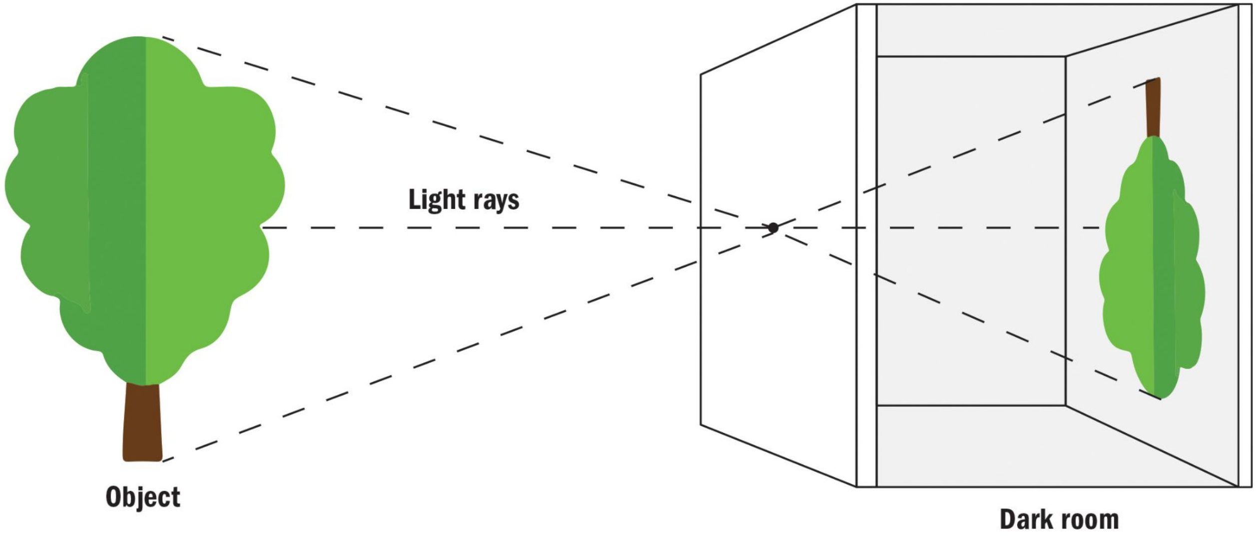


Transduction

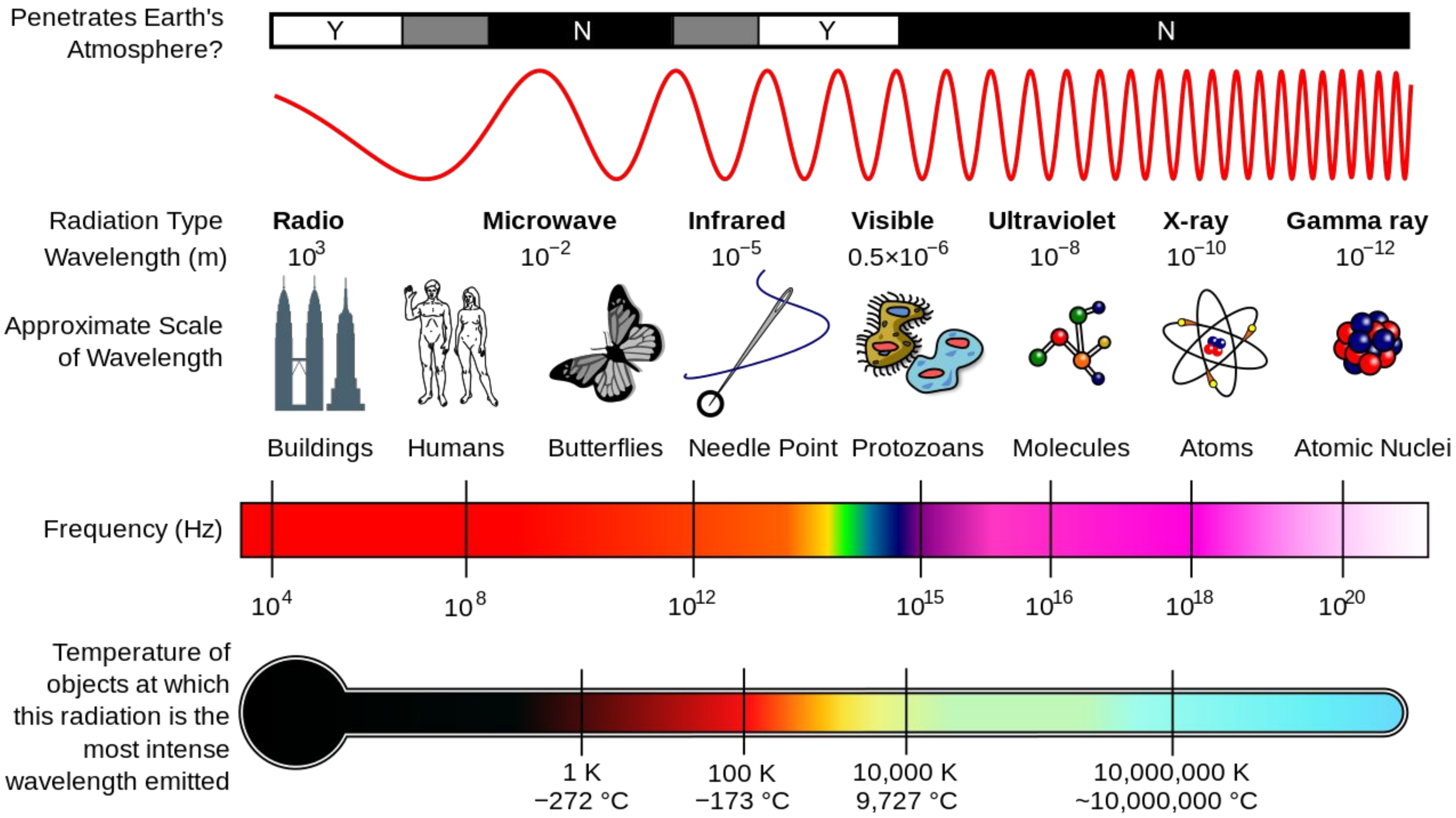


Camera Obscura





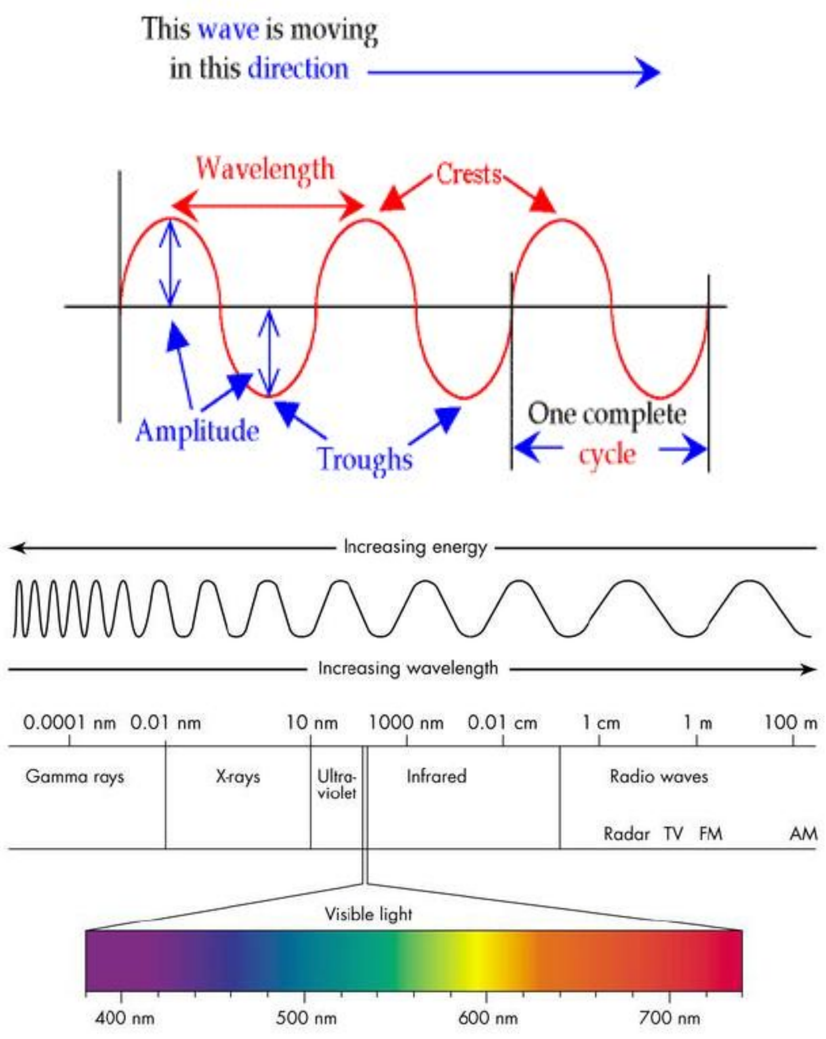
How We See



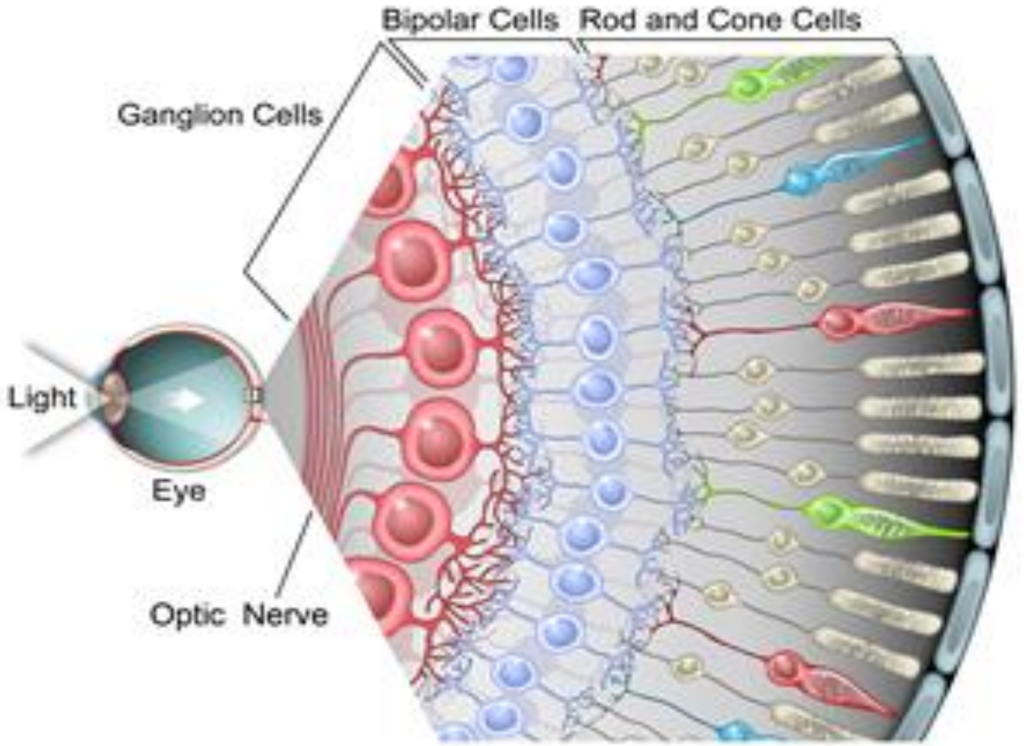
Penetrates Earth's Atmosphere?
Y
N
Y
N
Radiation Type
Wavelength (m)
Radio
10^3
Microwave
10^−2
Infrared
10^−5
Visible
0.5×10^−6
Ultraviolet
10^−8
X-ray
10^−10
Gamma ray
10^−12
Approximate Scale of Wavelength
Buildings
Humans
Butterflies
Needle Point
Protozoans
Molecules
Atoms
Atomic Nuclei
Frequency (Hz)
10^4
10^8
10^12
10^15
10^16
10^18
10^20
Temperature of objects at which this radiation is the most intense wavelength emitted
1 K
−272 °C
100 K
−173 °C
10,000 K
9,727 °C
10,000,000 K
~10,000,000 °C

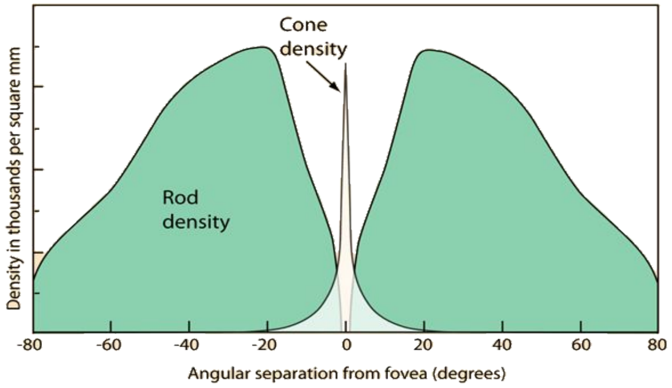
Properties of Light

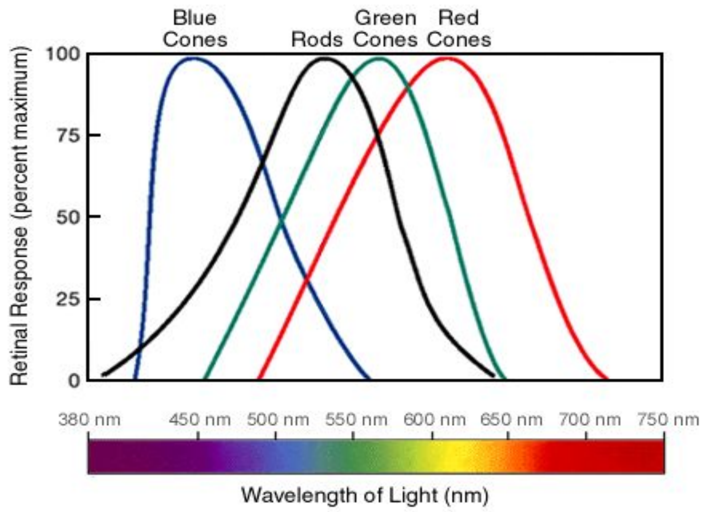
- **Wavelength**
  - Perception of colour
- **Intensity** (derived from amplitude of waveform)
  - Perception of brightness



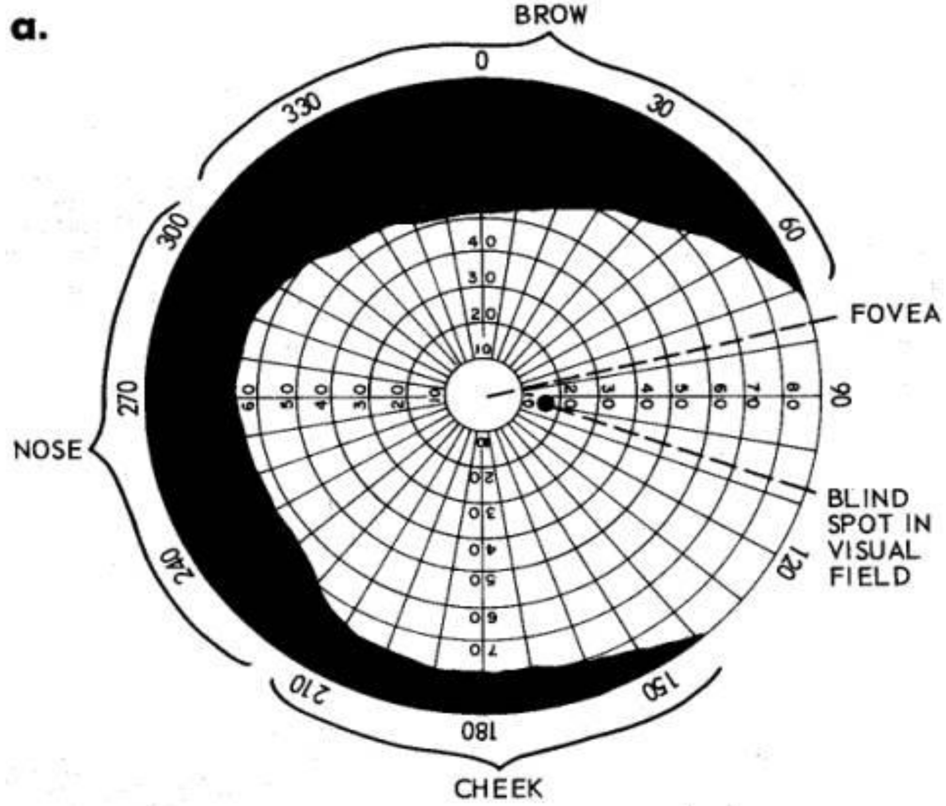
Retina

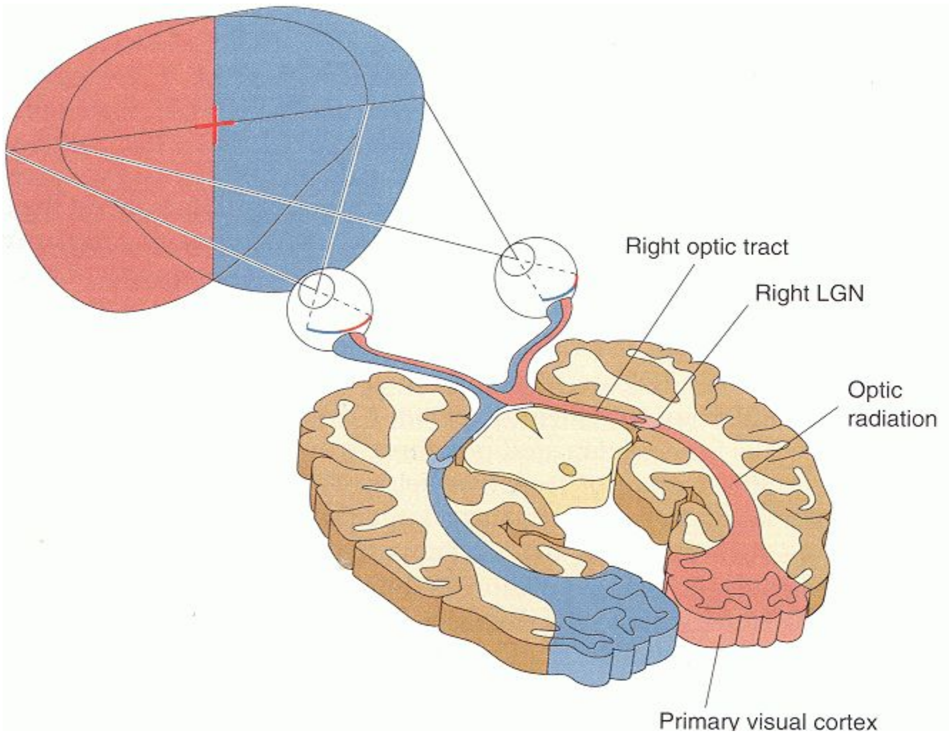

Bipolar Cells
Rod and Cone Cells
Ganglion Cells
Light
Eye
Optic Nerve


Cone
density
Rod
density
Density in thousands per square mm
-80
-60
-40
-20
0
20
40
60
80
Angular separation from fovea (degrees)


Blue
Cones
Rods
Green
Cones
Red
Cones
100
75
50
25
0
Retinal Response (percent maximum)
380 nm
450 nm
500 nm
550 nm
600 nm
650 nm
700 nm
750 nm
Wavelength of Light (nm)

Visual Field

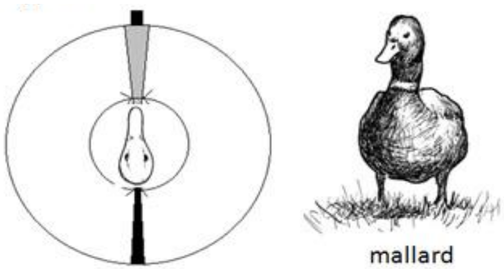


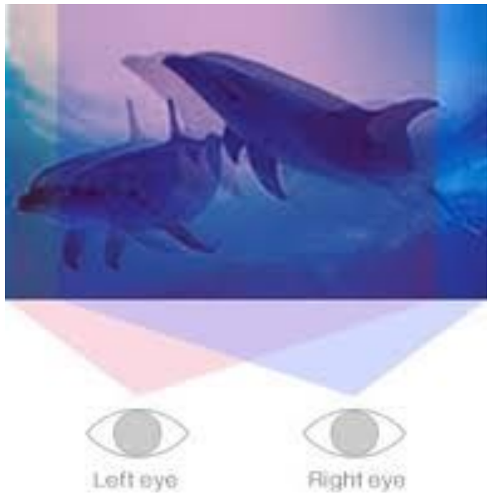


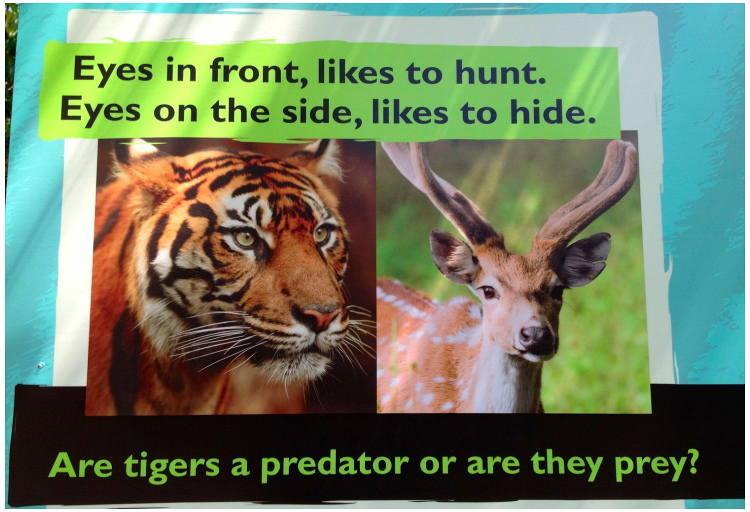
Eye Position

Eyes come in pairs

- Ducks - Eyes on both sides
  - Can see in every direction
- Humans - Eyes in front
  - Sacrifice backwards view
  - Can look forward through both eyes simultaneously
  - Create 3d perceptions form 2d retinal images
    - Depth



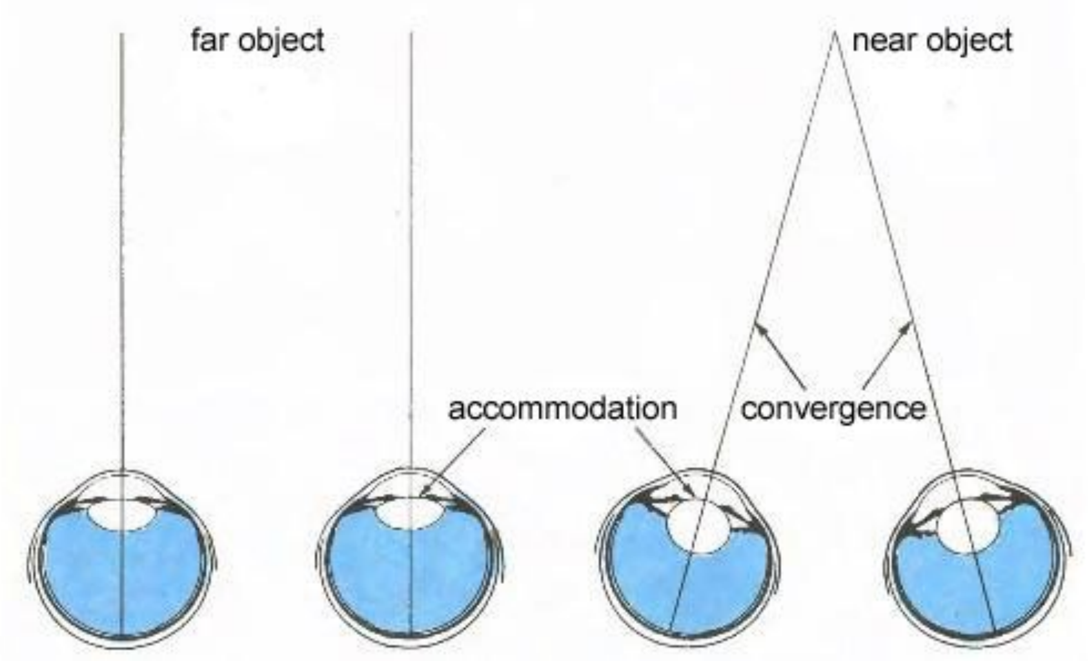




Eye Position

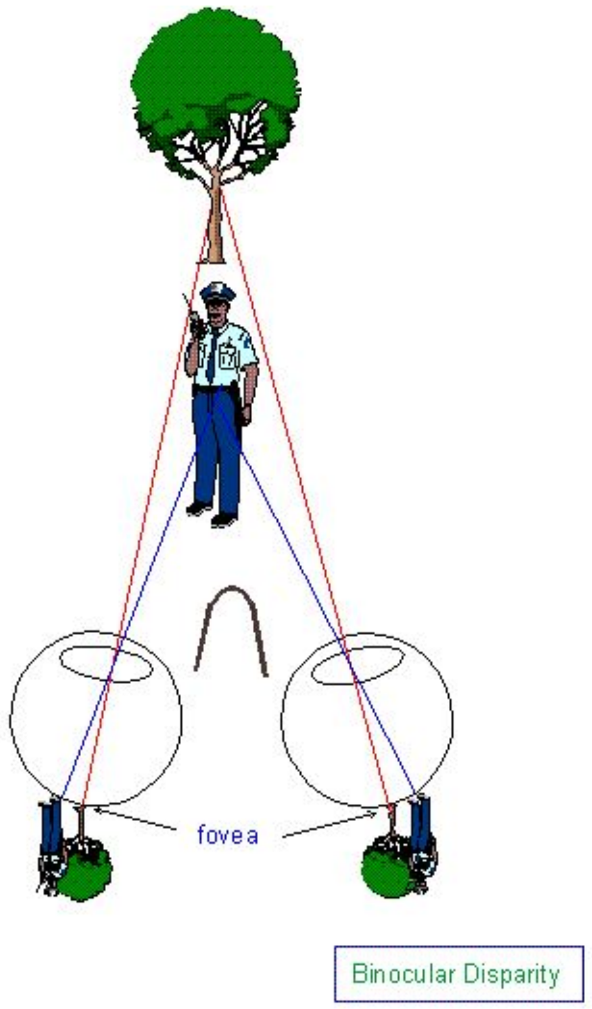
Binocular Coordination

- Each point of visual field project to corresponding points on 2 retinae
- Eyes converge to see close objects
- Correspondence never perfect; each eye views the world from slightly different position, gets different image

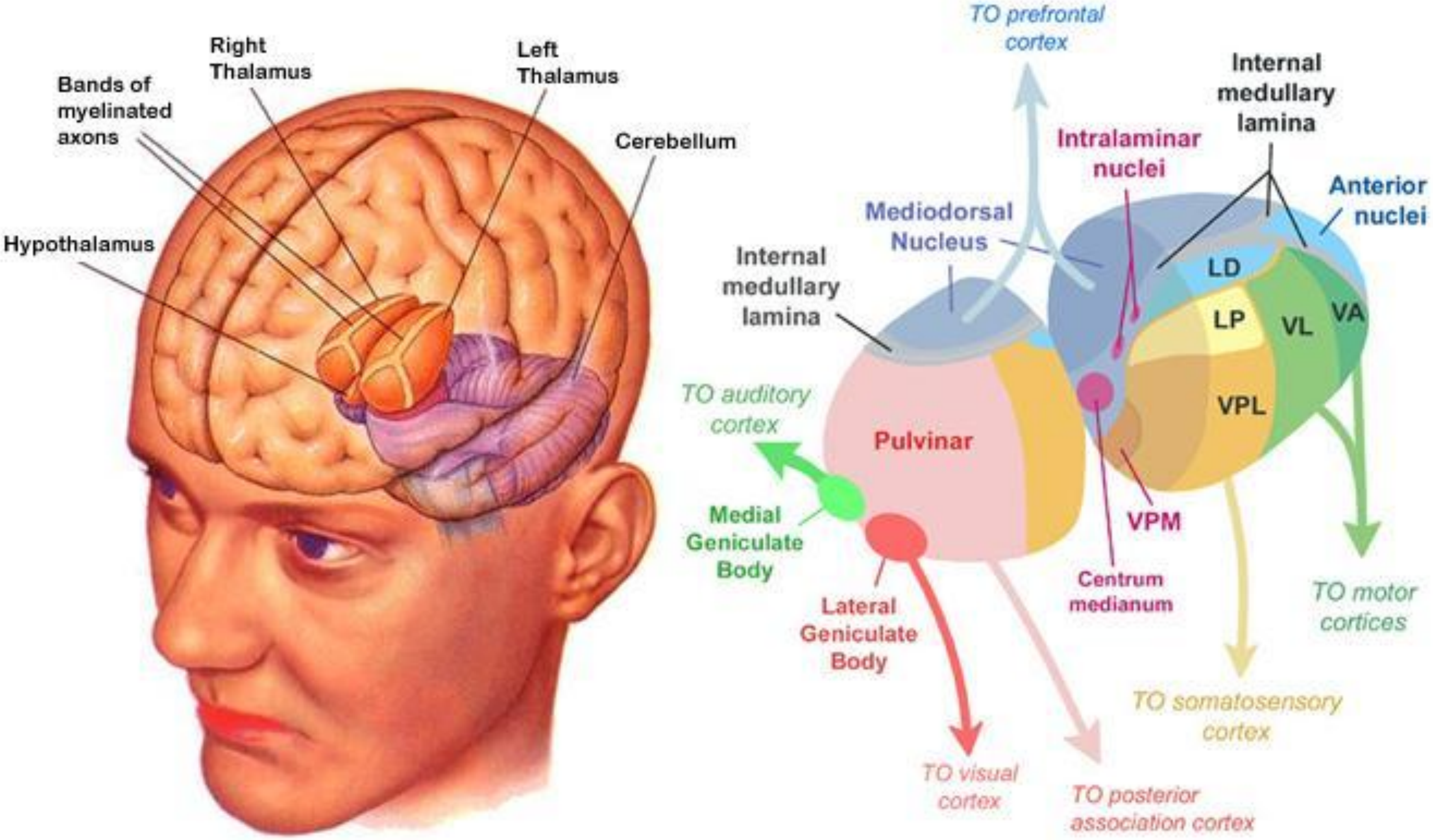


Binocular Disparity

- The difference in the same image on the 2 retinas
  - Greater for close than far objects
- Visual system constructs a 3d perception from 2 slightly different 2d retinal images
- Helps with perception of distance

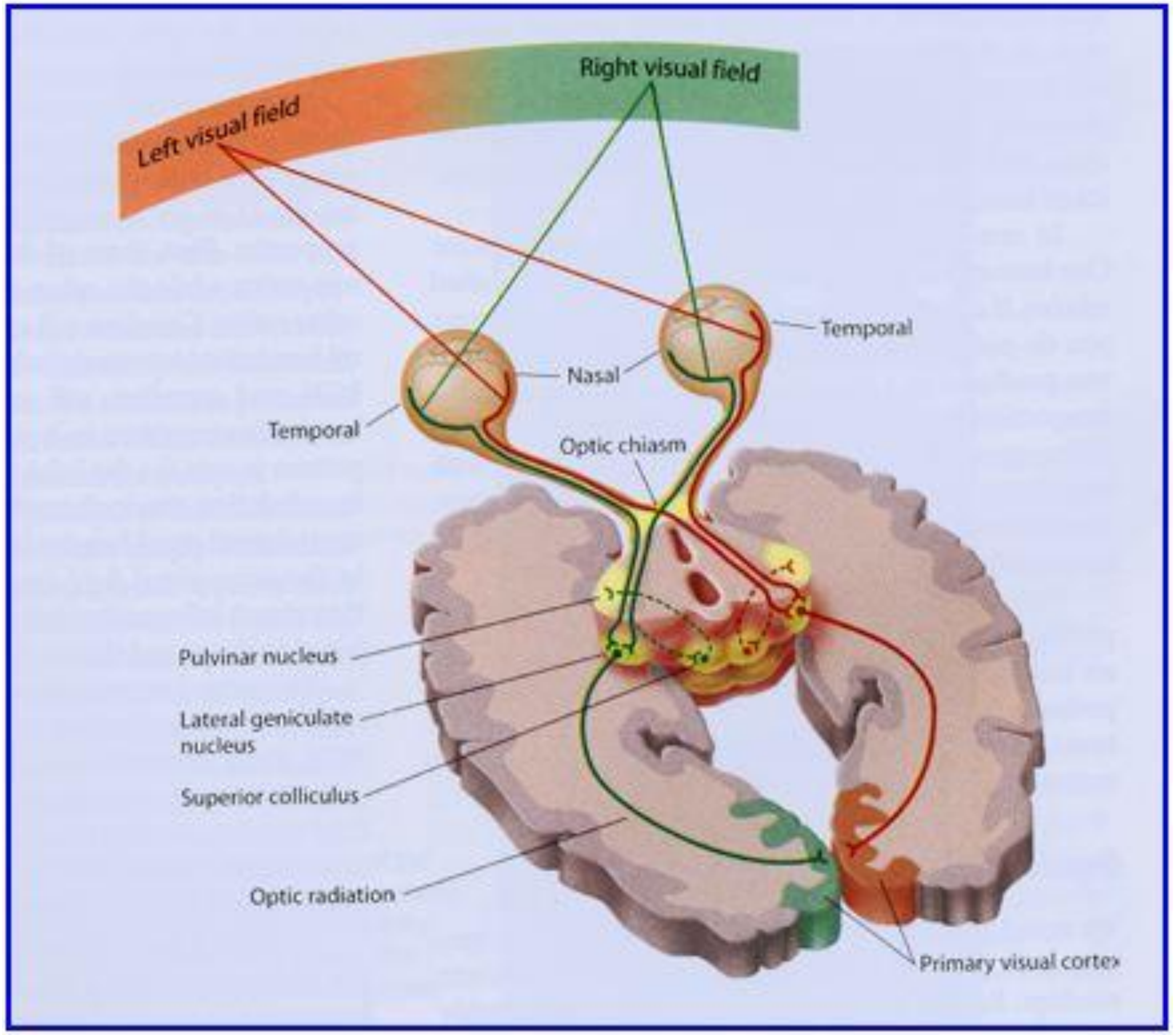


Thalamus



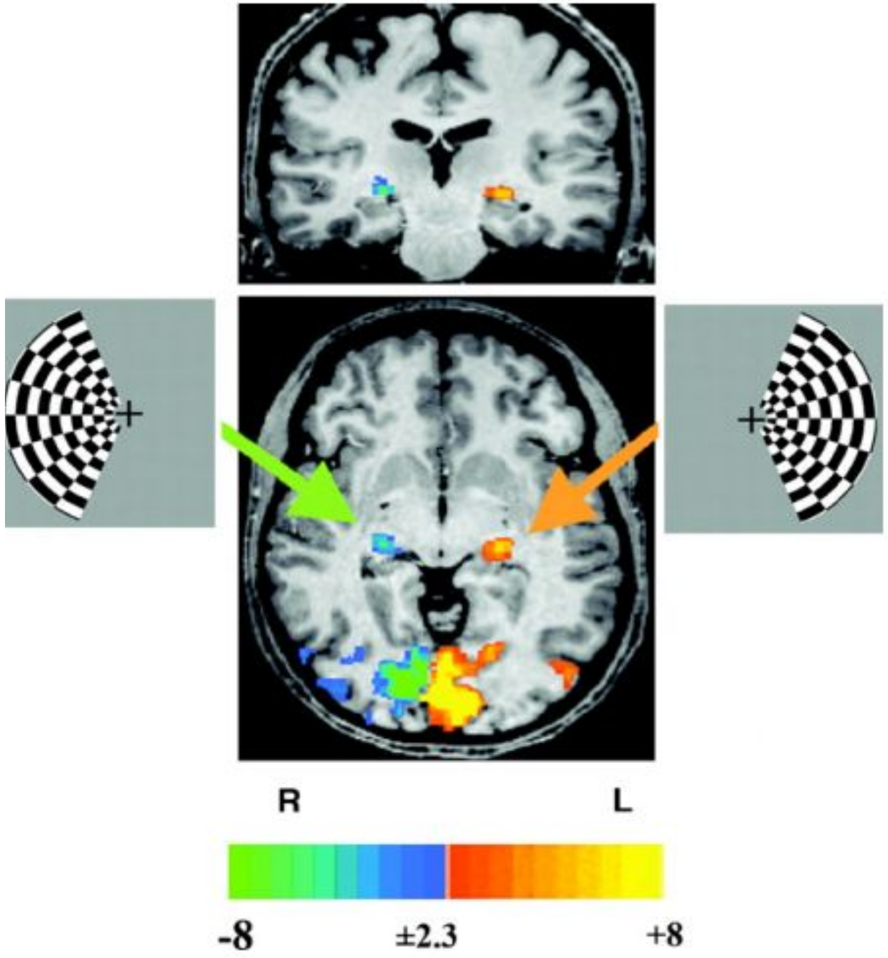
Right Thalamus
Left Thalamus
Bands of myelinated axons
Cerebellum
Hypothalamus
TO prefrontal cortex
Internal medullary lamina
Intralaminar nuclei
Anterior nuclei
Mediodorsal Nucleus
Internal medullary lamina
LD
LP
VL
VA
VPL
Pulvinar
TO auditory cortex
Medial Geniculate Body
Lateral Geniculate Body
VPM
Centrum medianum
TO motor cortices
TO somatosensory cortex
TO visual cortex
TO posterior association cortex

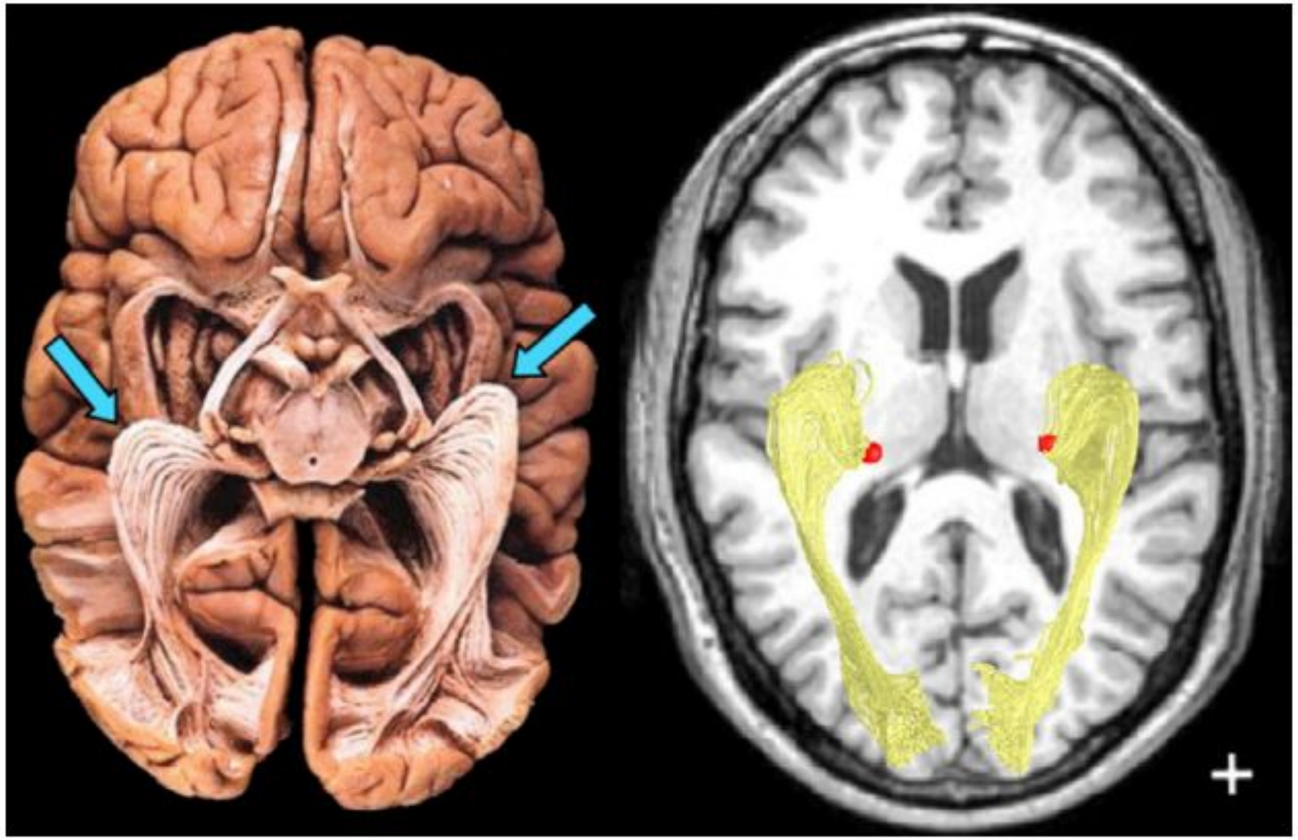
Lateralisation



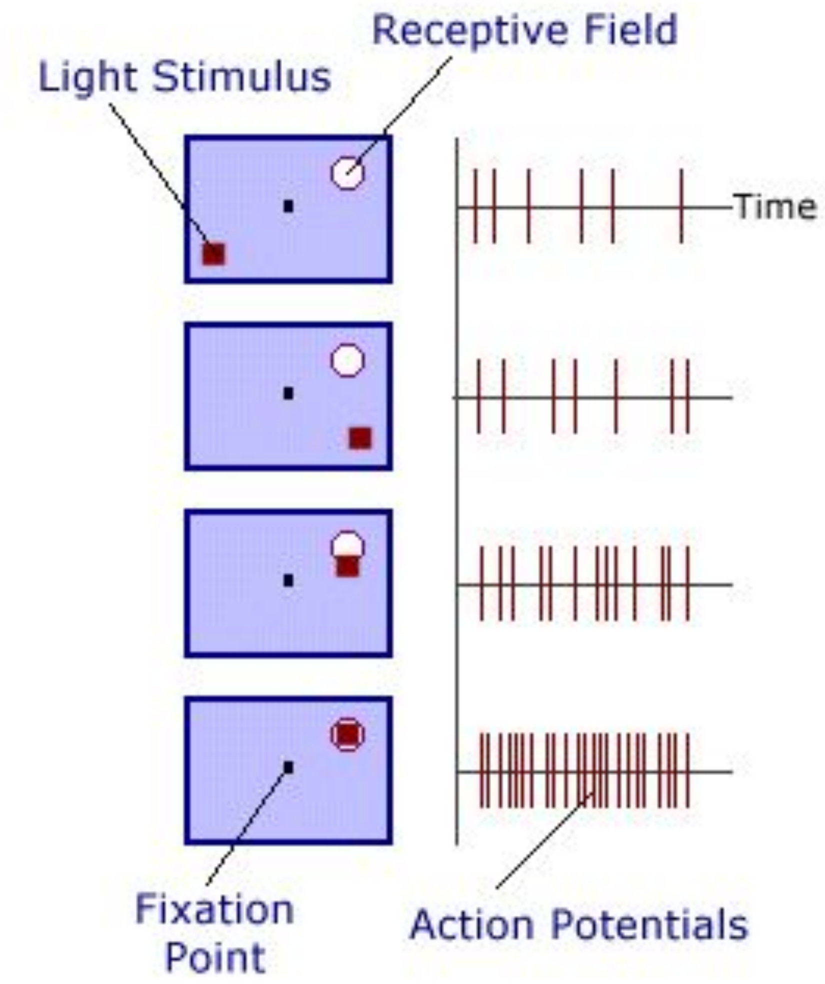
https://www.cns.nyu.edu/~david/courses/perception/lecturenotes/V1/lgn-V1.html

LGN

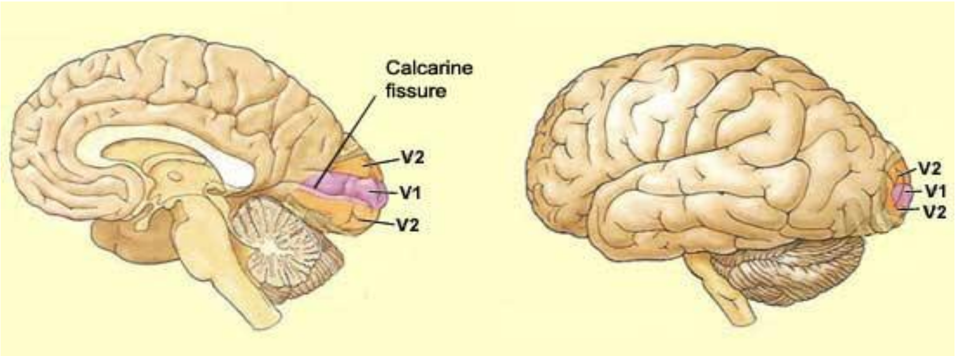


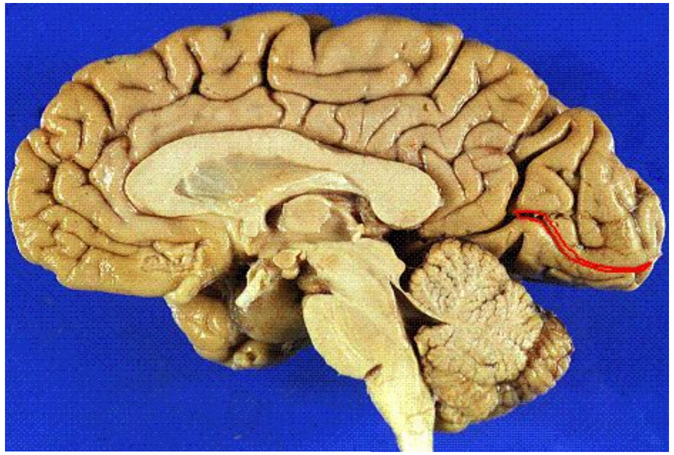


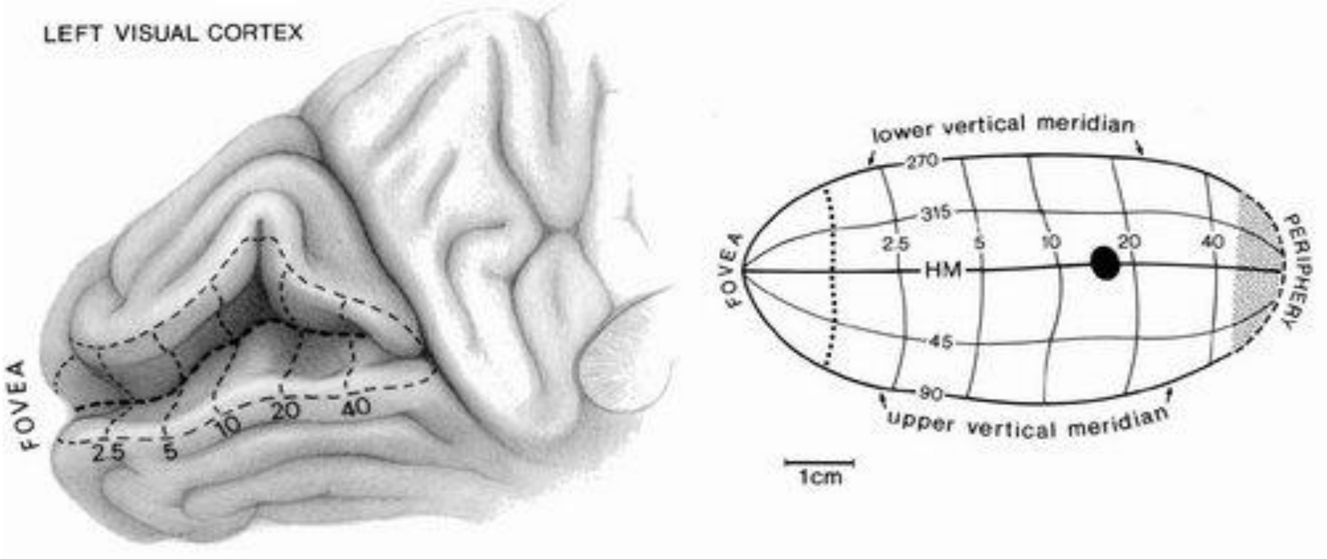
Receptive Field



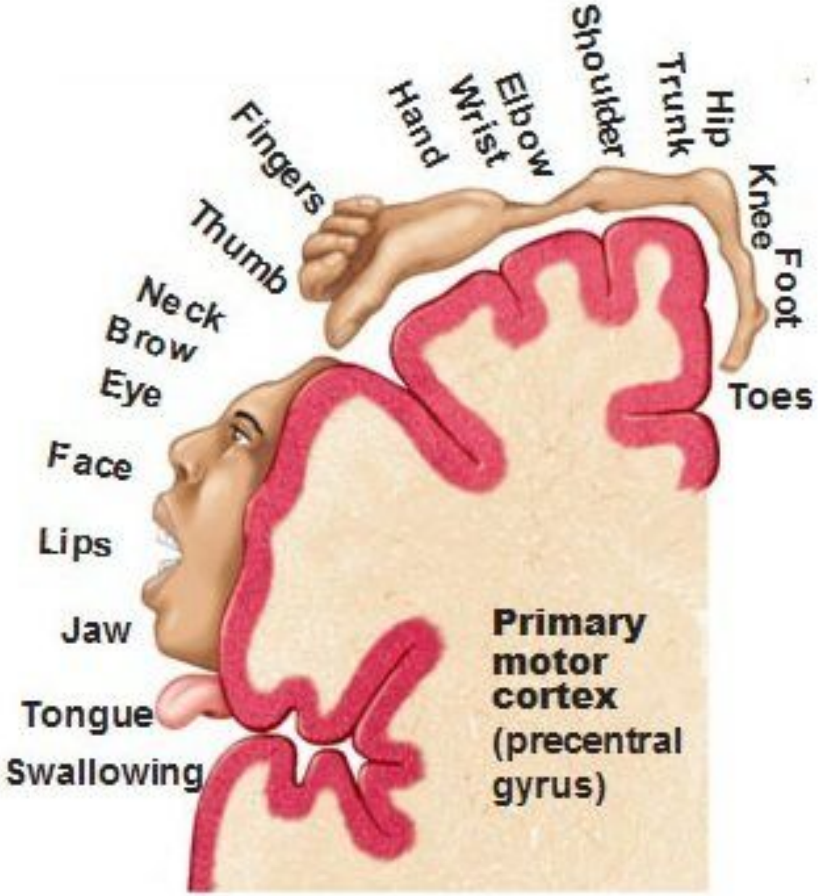
Primary Visual Cortex (V1)



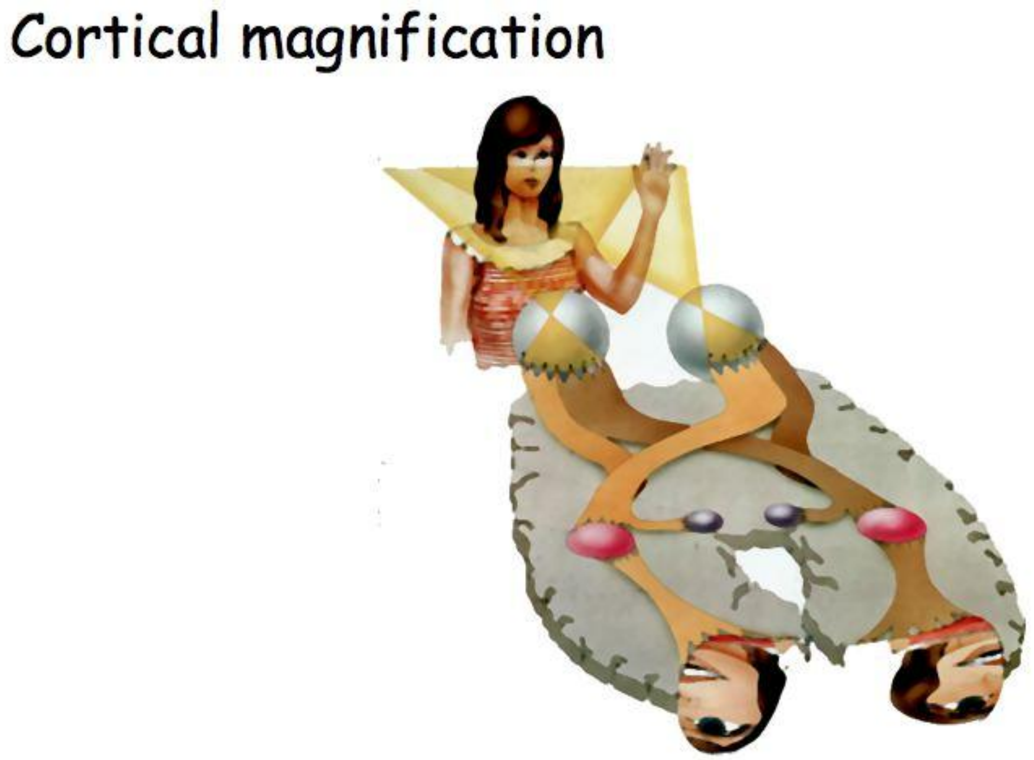


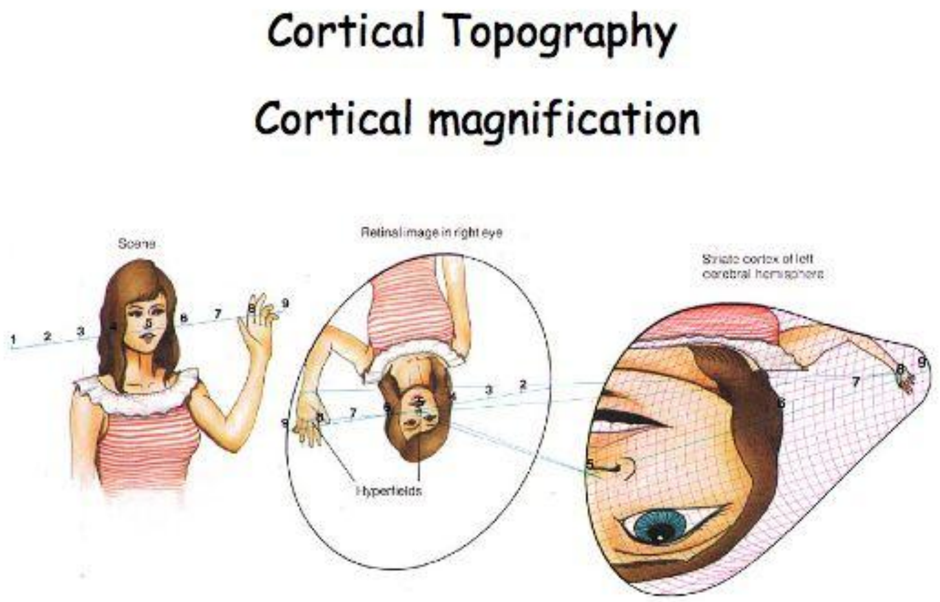


Sensory Homunculus



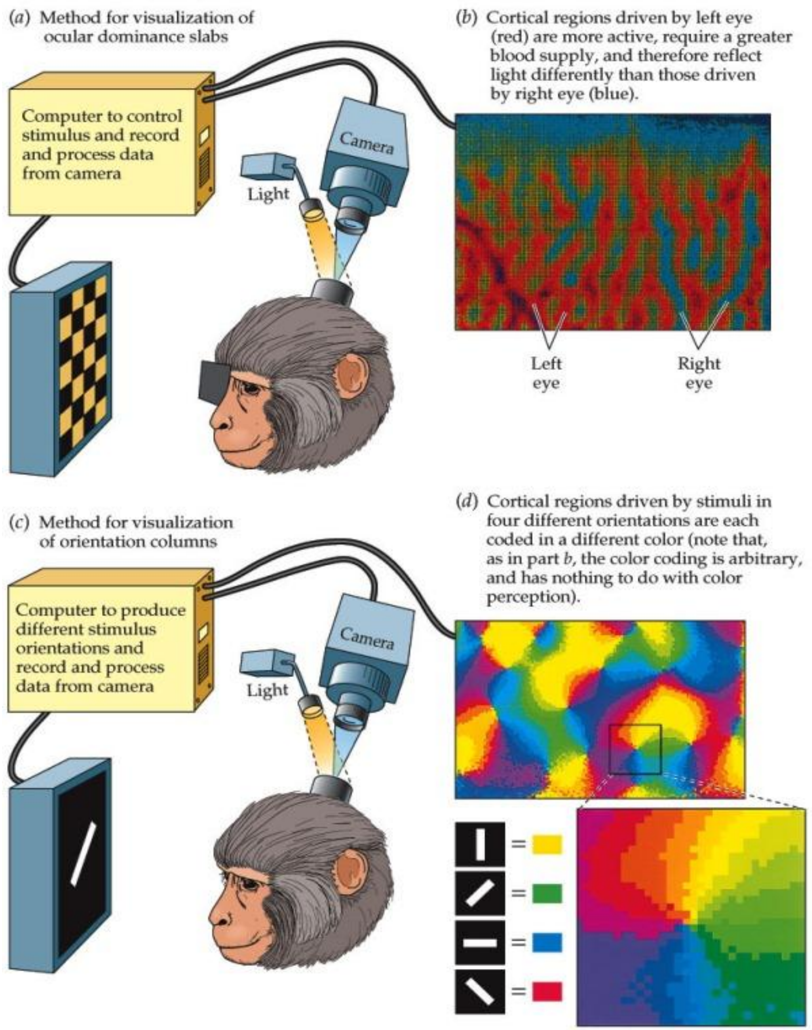
Cortical Magnification

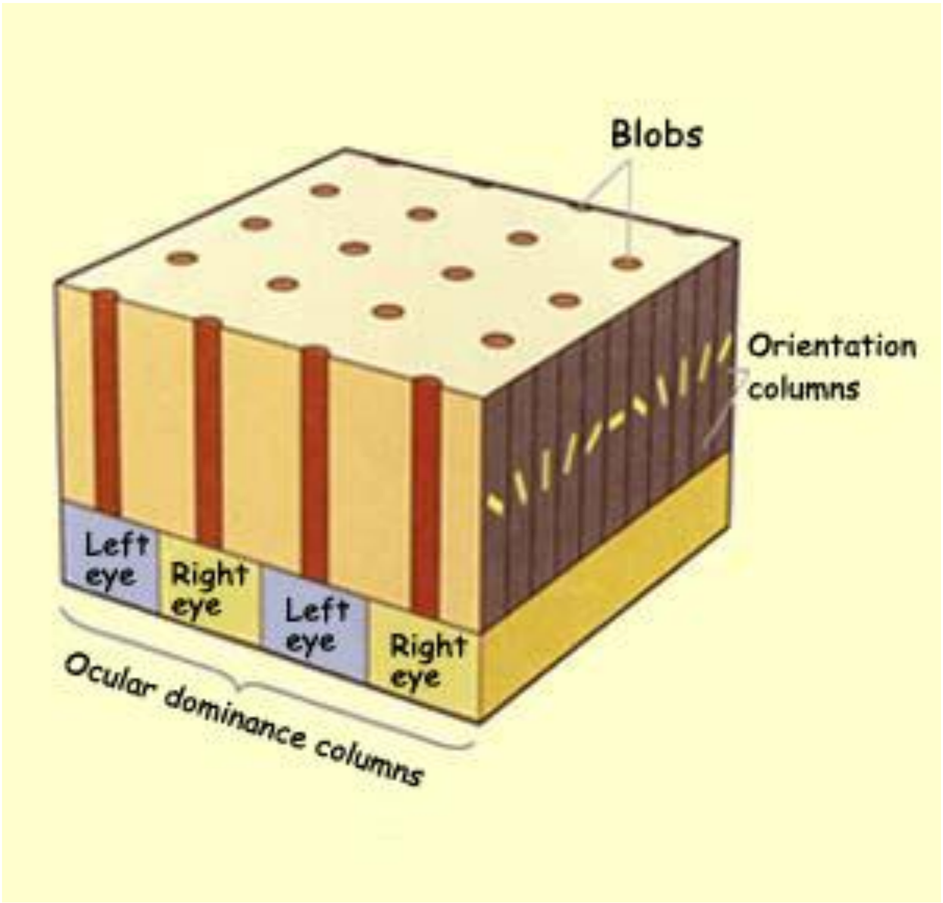




https://www.cns.nyu.edu/~david/courses/perception/lecturenotes/V1/lgn-V1.html

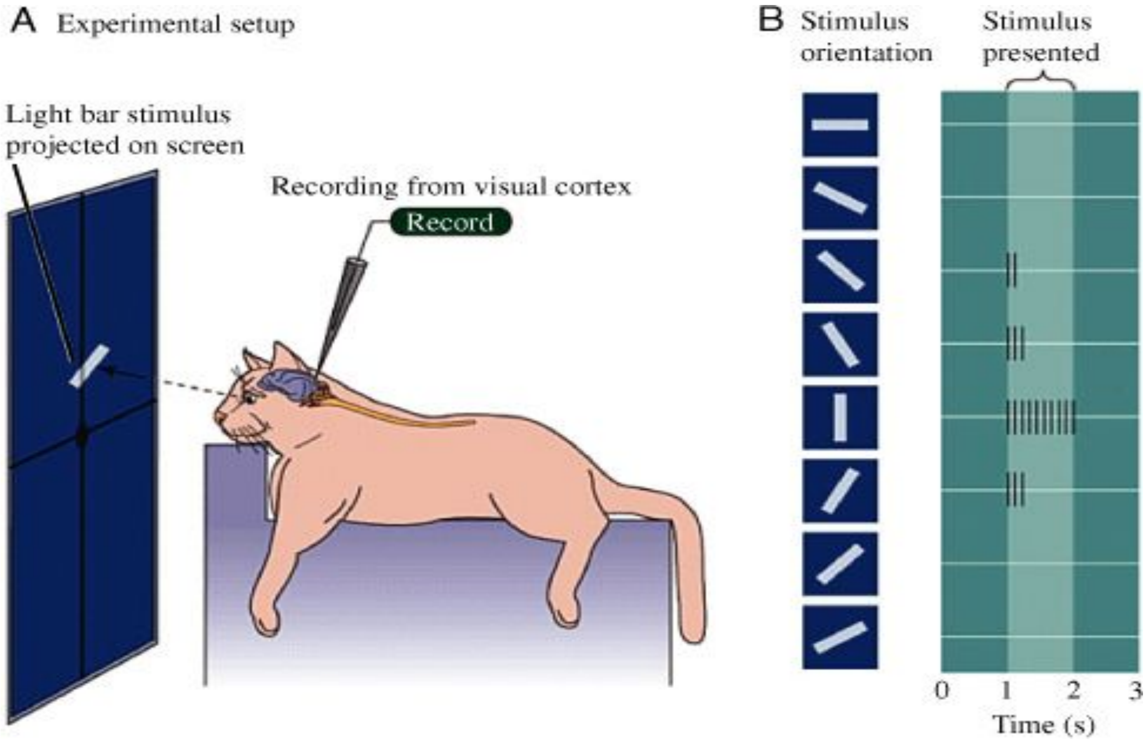
Hypercolumns & Blobs





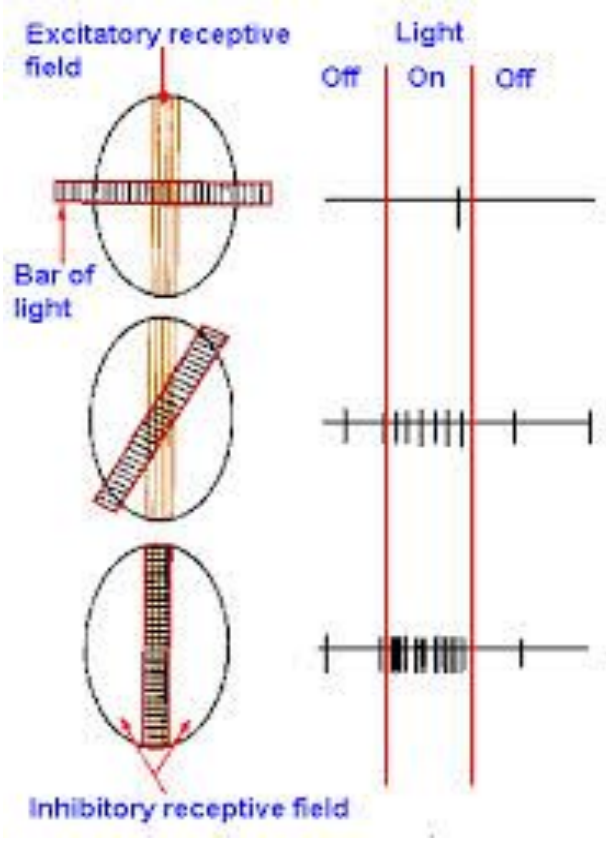
Receptive Fields

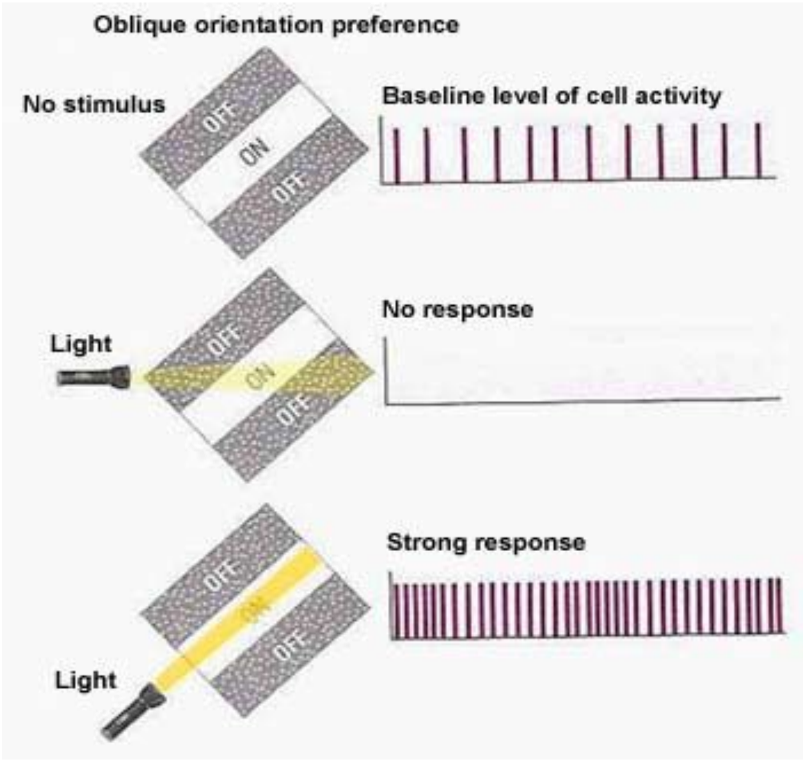
- Hubel and Wiesel (1970s)
  - Implant electrode near a single neuron
  - Map neuron's **receptive field**
  - Once field mapped different stimuli presented in receptive field to influence firing



Receptive Fields (Simple)

- V1 neurons (layer 4) tend to respond to straight edges instead of circular edges

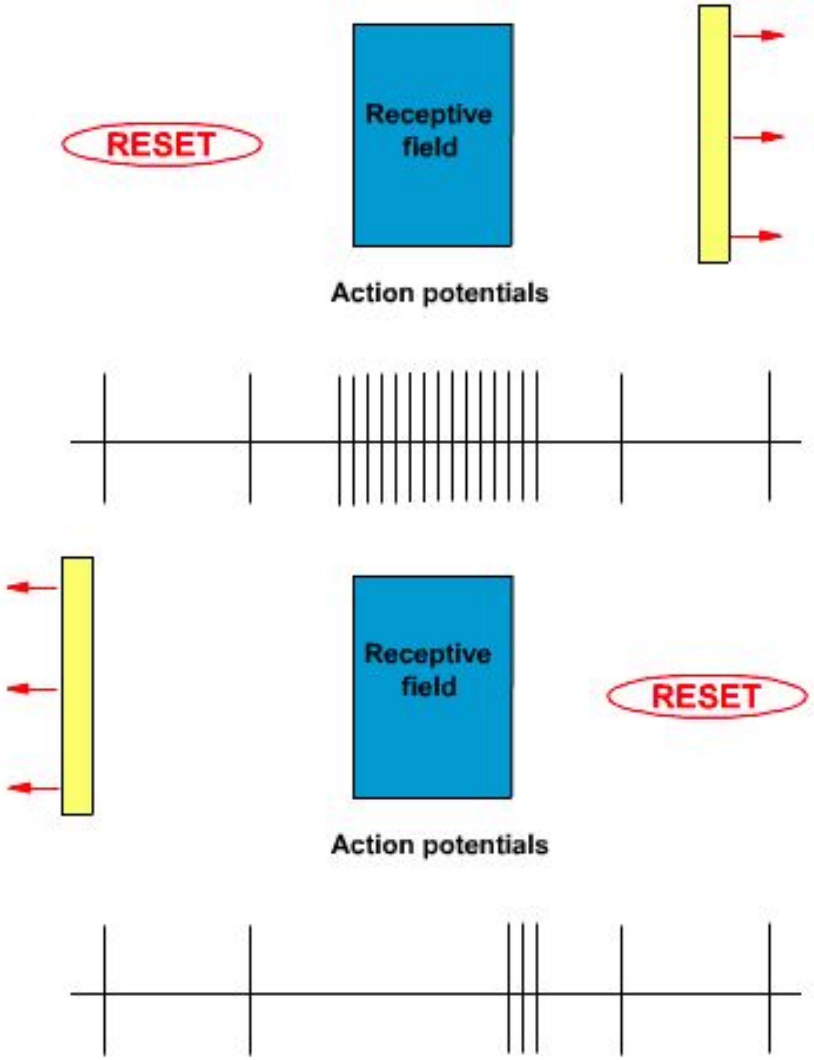




Receptive Fields (Complex)

Differences to simple cells

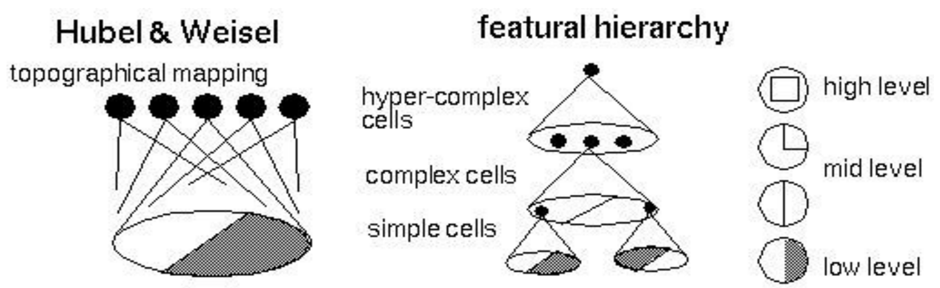
1. Larger receptive fields
2. They do not have static on and off regions
3. Many are *binocular*

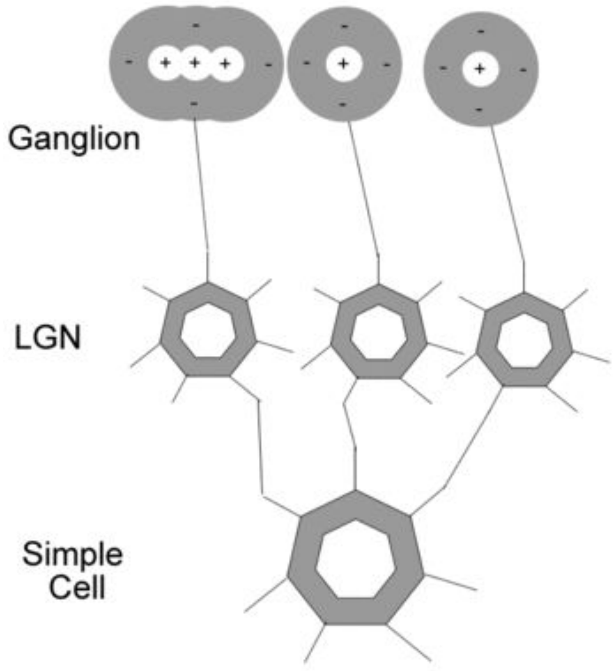


Hierarchical Model

- Hubel and Wiesel proposed that complexity of receptive fields at higher levels of processing is from convergence of preceding levels
  - LGN → simple cells → complex cells

- Hubel and Wiesel proposed that V1 divided into functionally independent **columns** that are responsible for analyzing input from different areas of visual field…

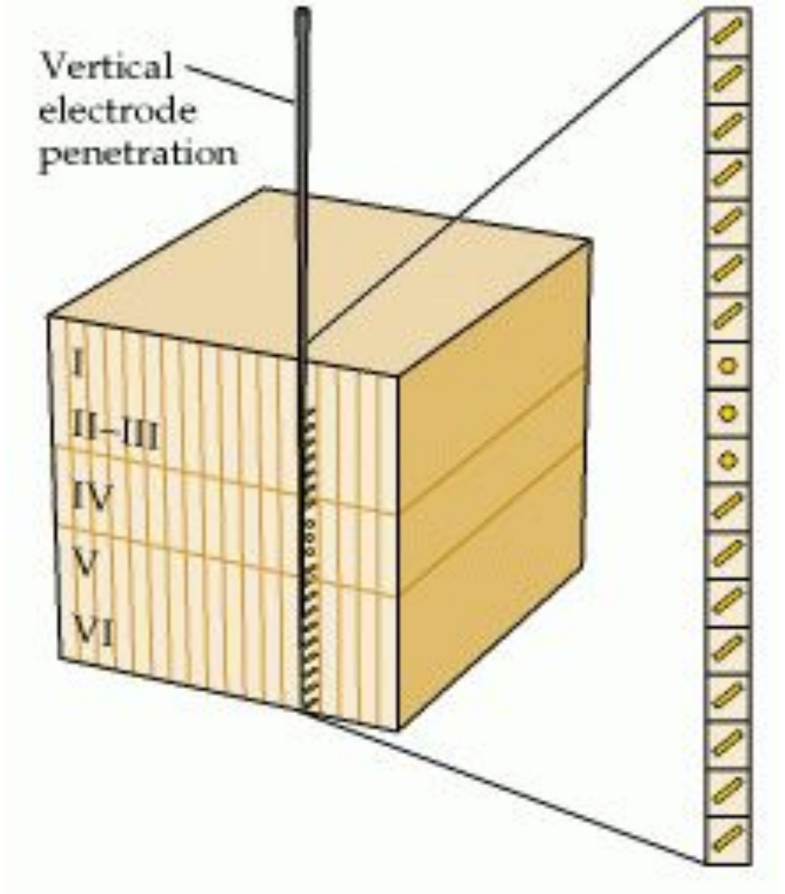




Columnar Organisation

Electrode lowered vertically into V1 and map receptive field of neurons along the way

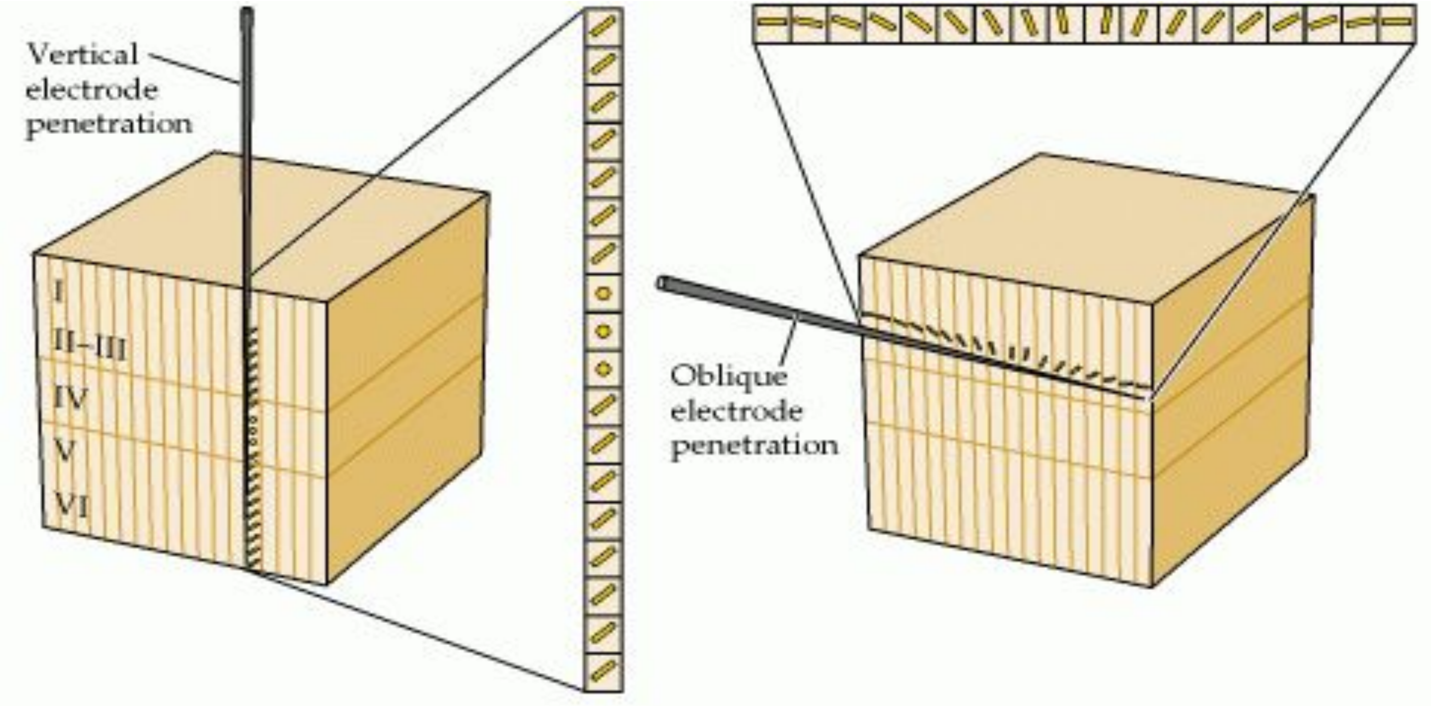
Common characteristics of cells in a column.



Columnar Organisation

Electrode passed horizontally through striate.

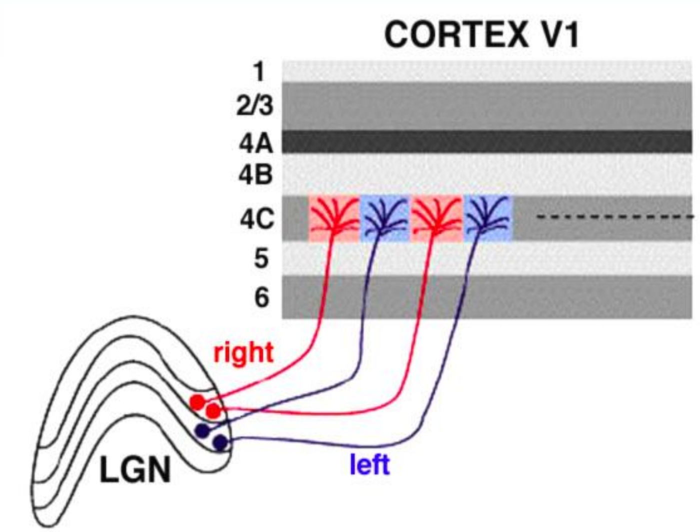
Each successive cell receptive field is in slightly different location in visual field. **Each cell responds to lines of slightly different orientations**

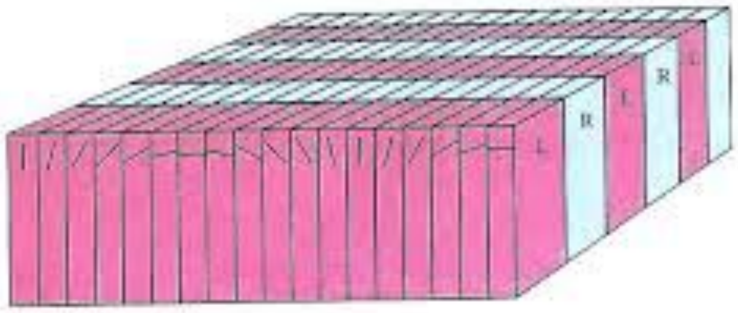


Ocular Dominance

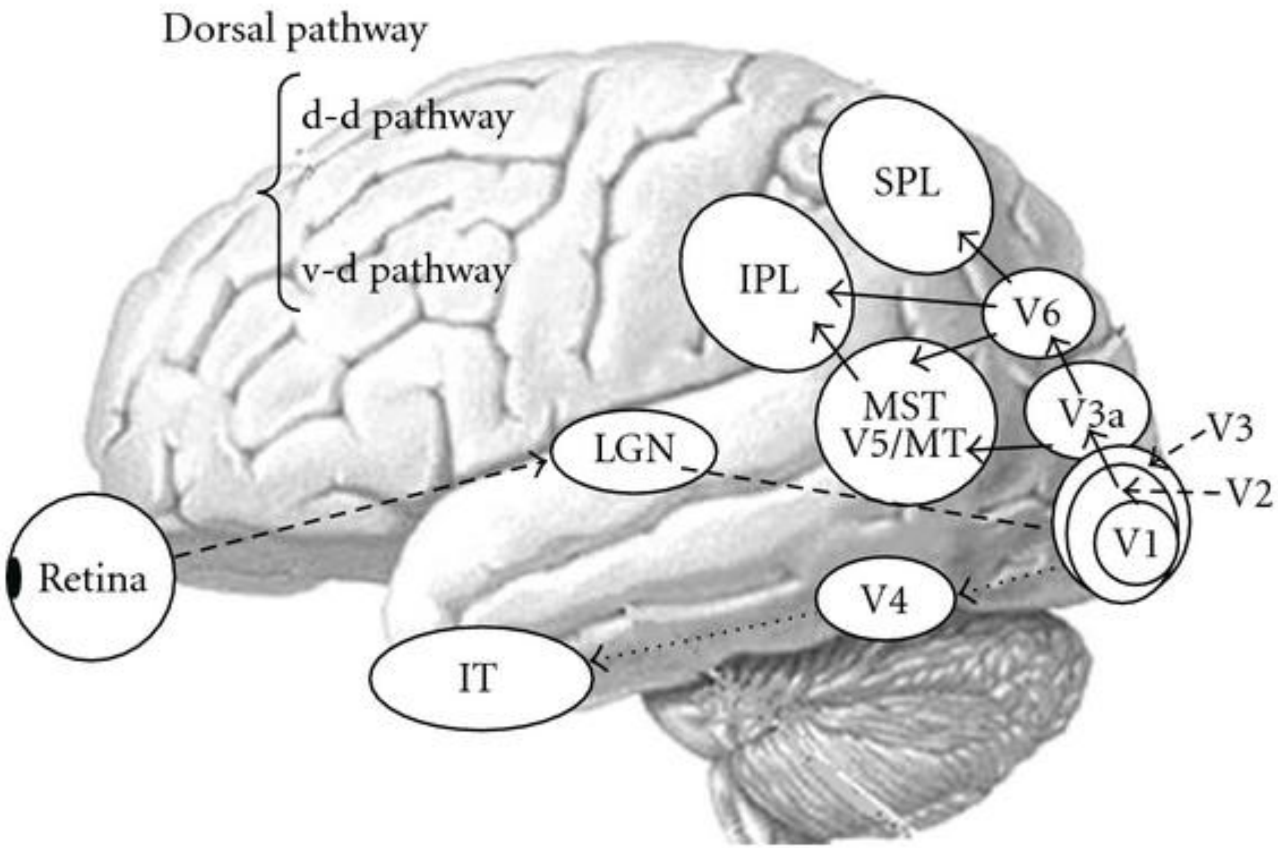
Horizontally, successive cells alternate with respect to left or right eye dominance.

Cortical column can be subdivided further into ocular dominance.



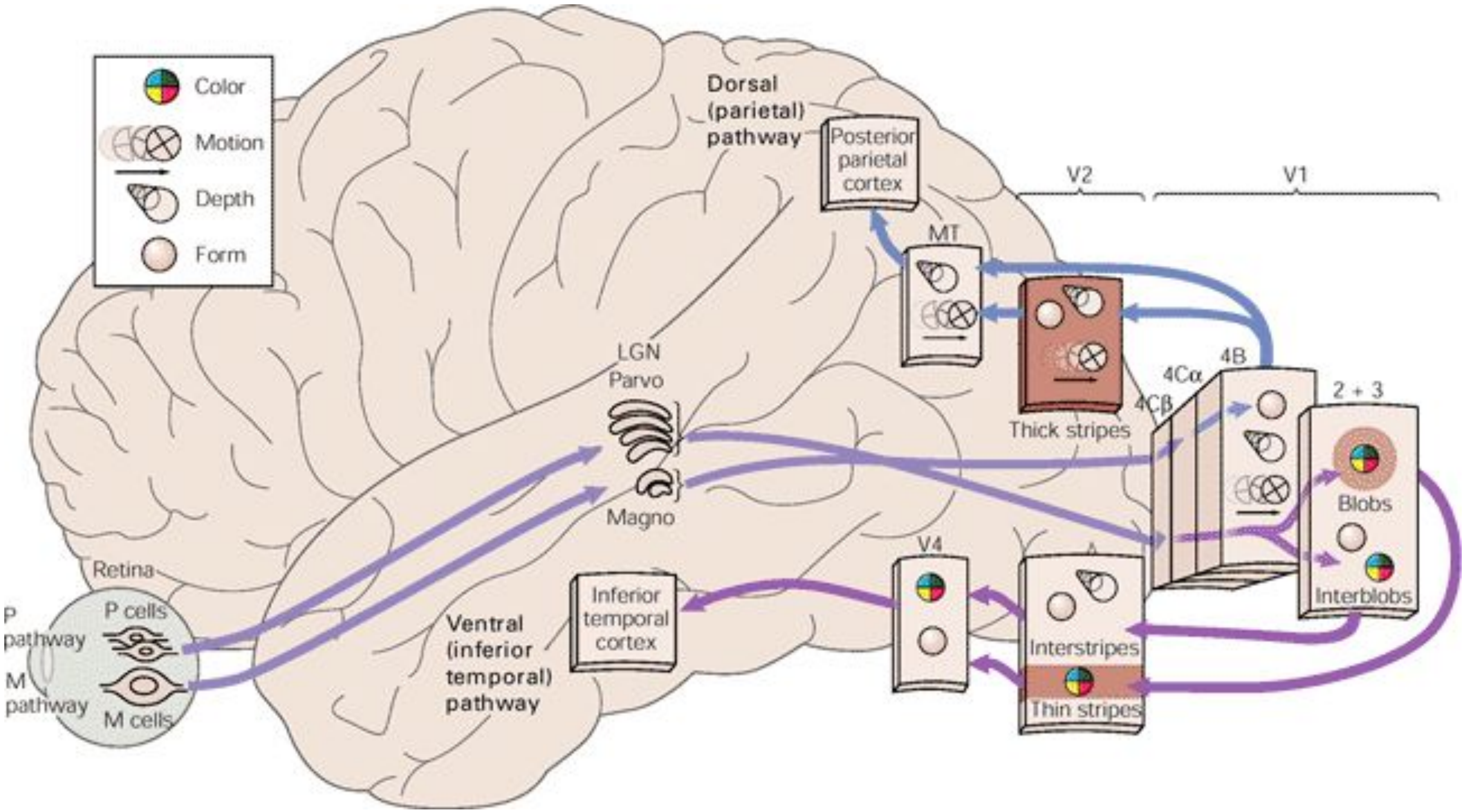


Visual System



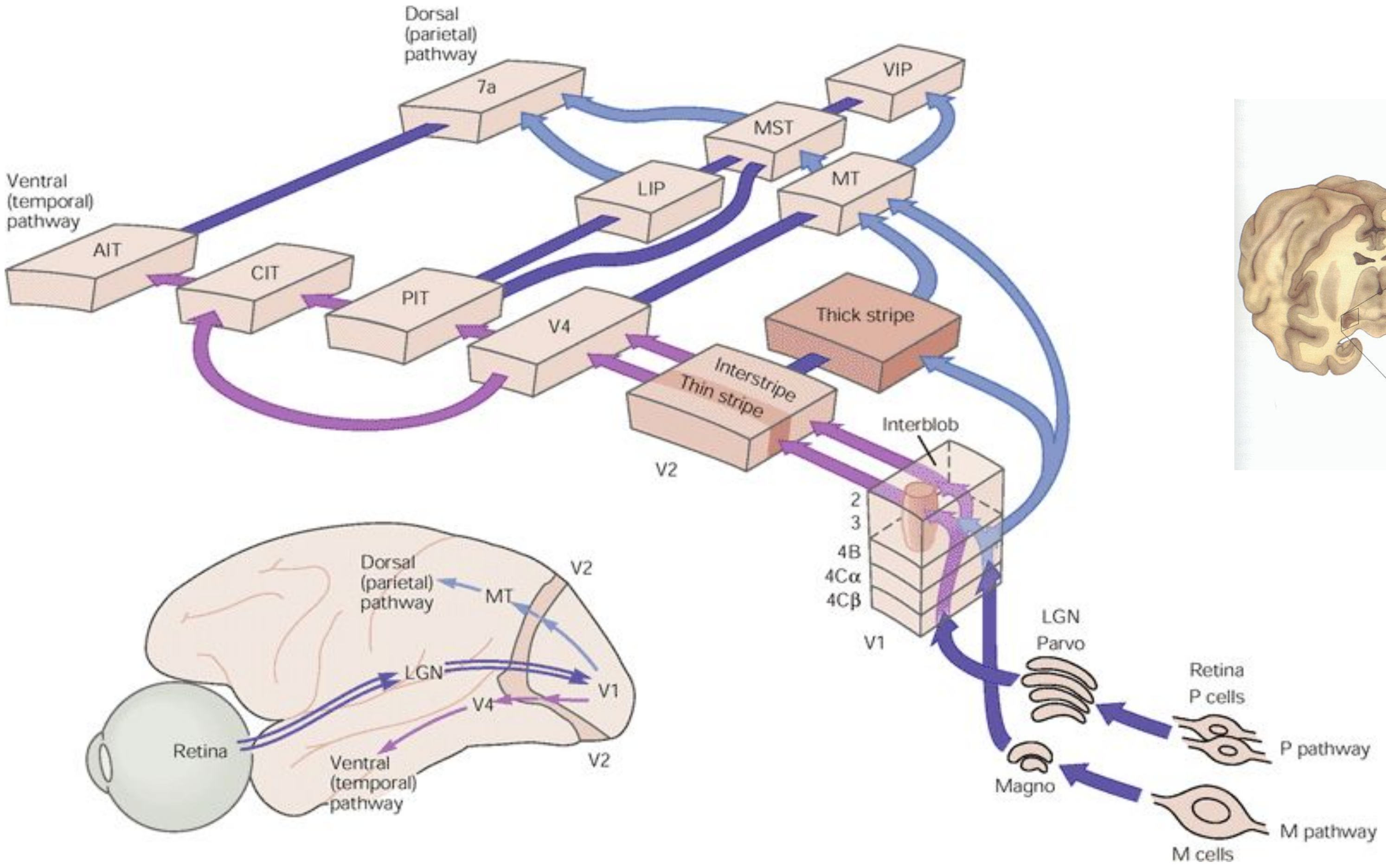
Yamasaki et al. (2012). Neurology research international.

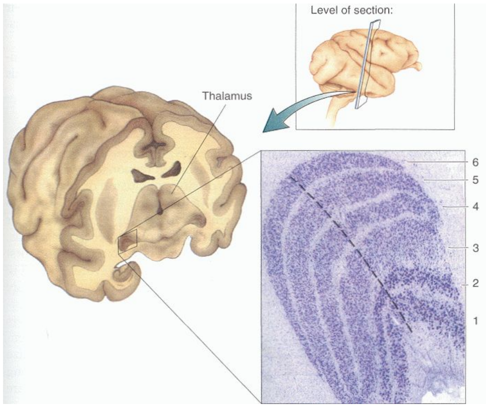
Dorsal / Ventral Stream



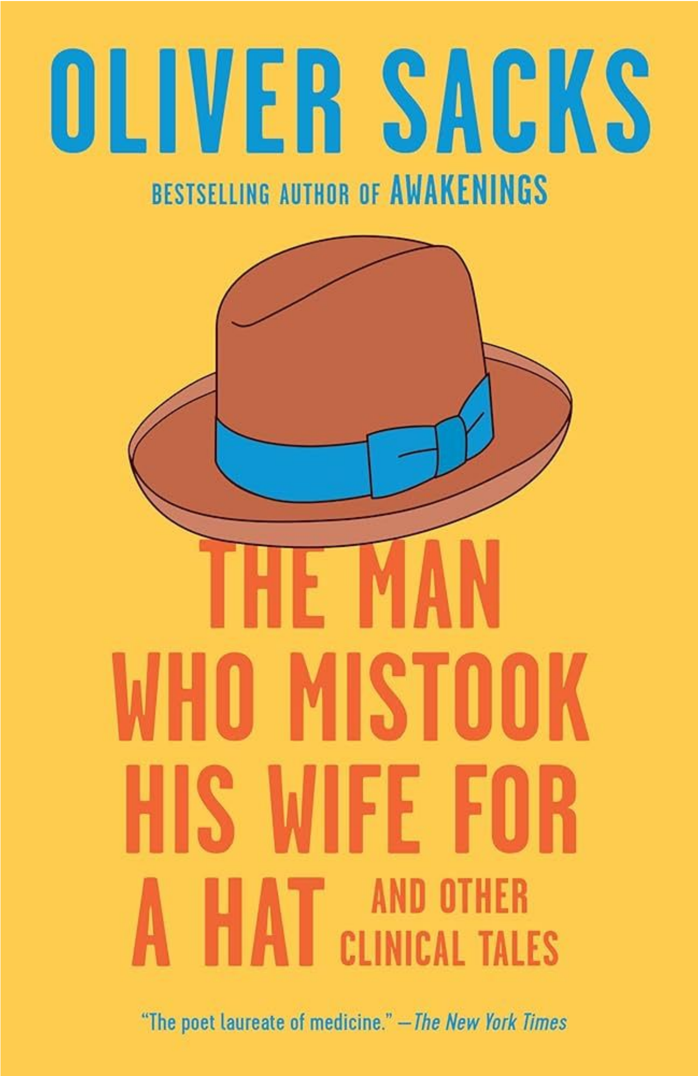
Magno/Parvo Pathways

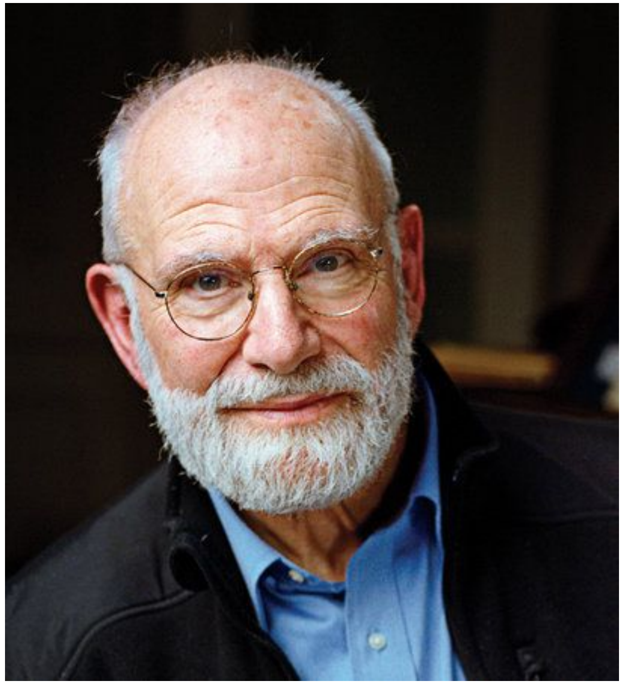
Magnocellular
Parvocellular



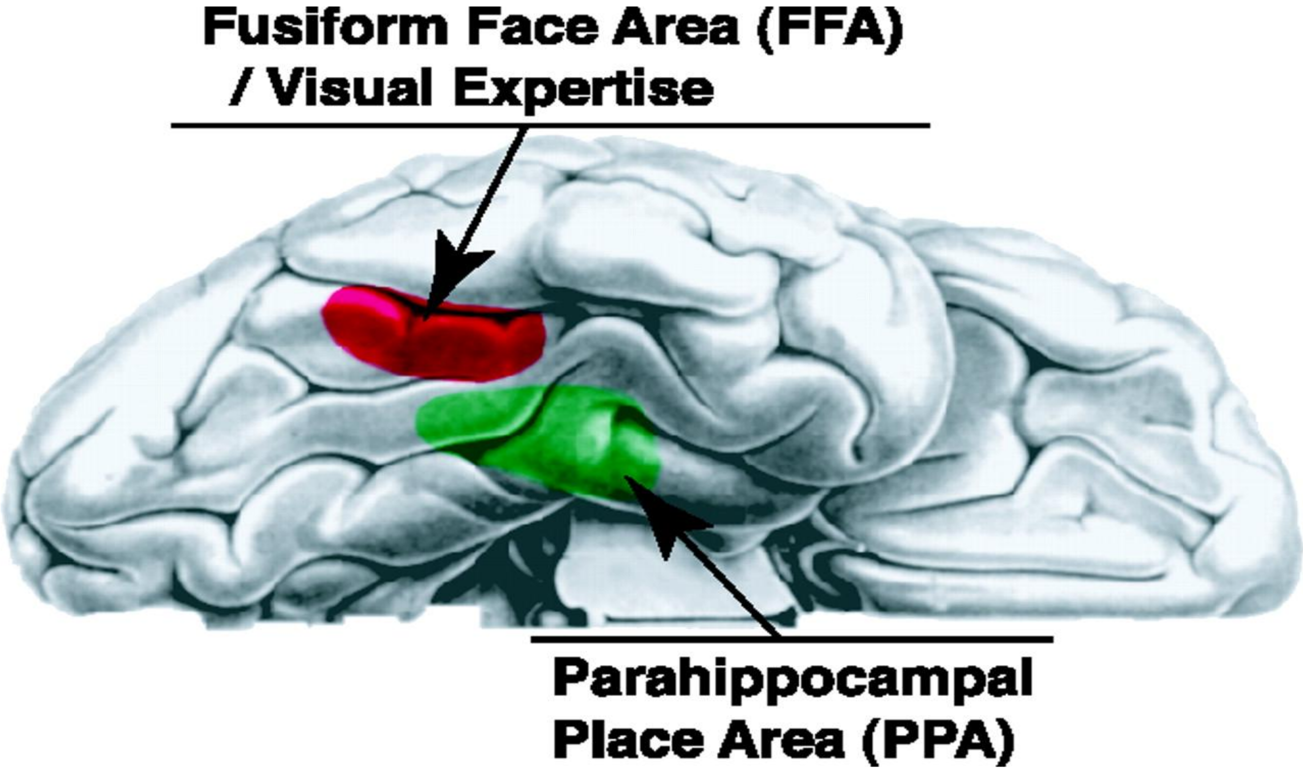


Prosopagnosia

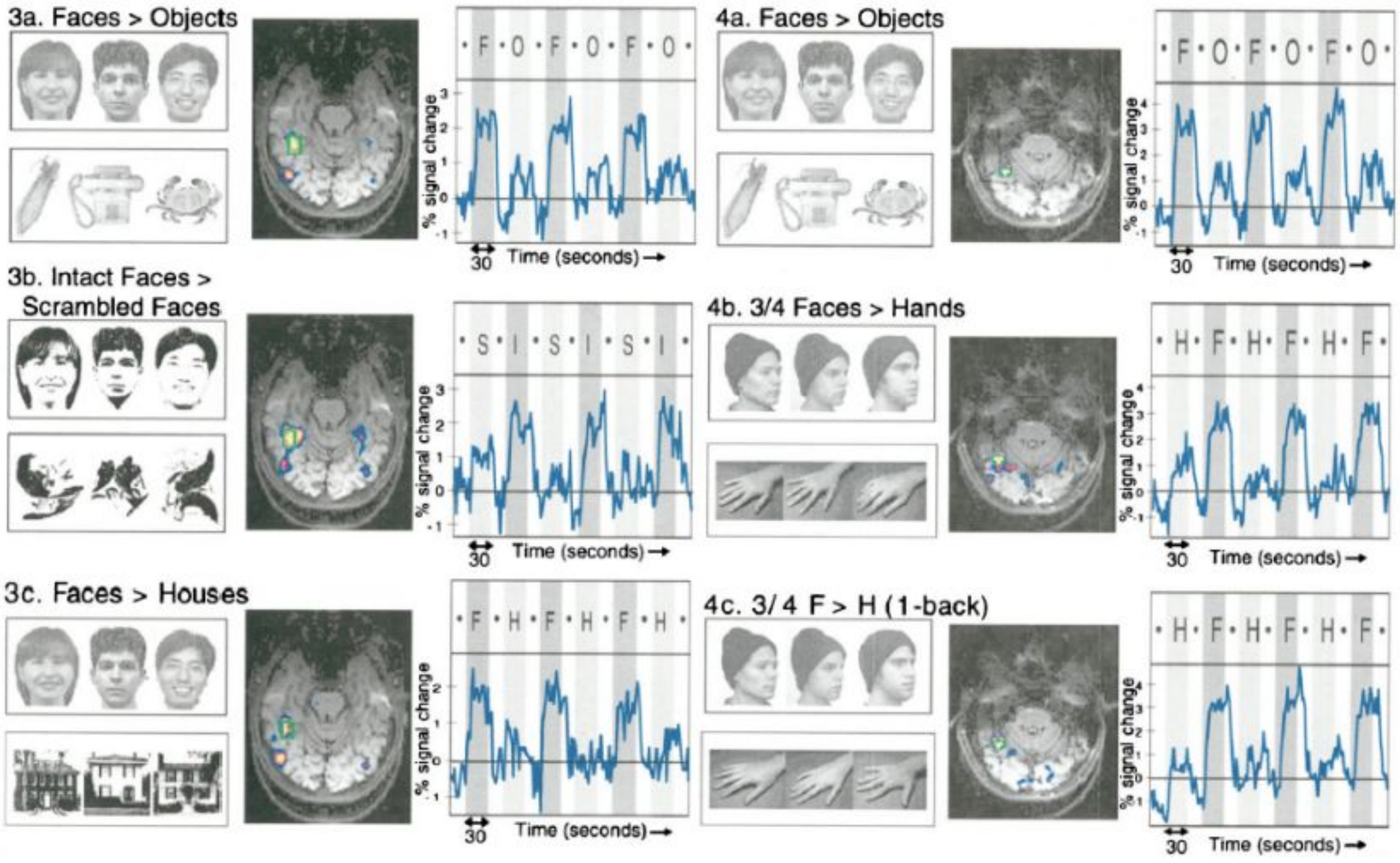




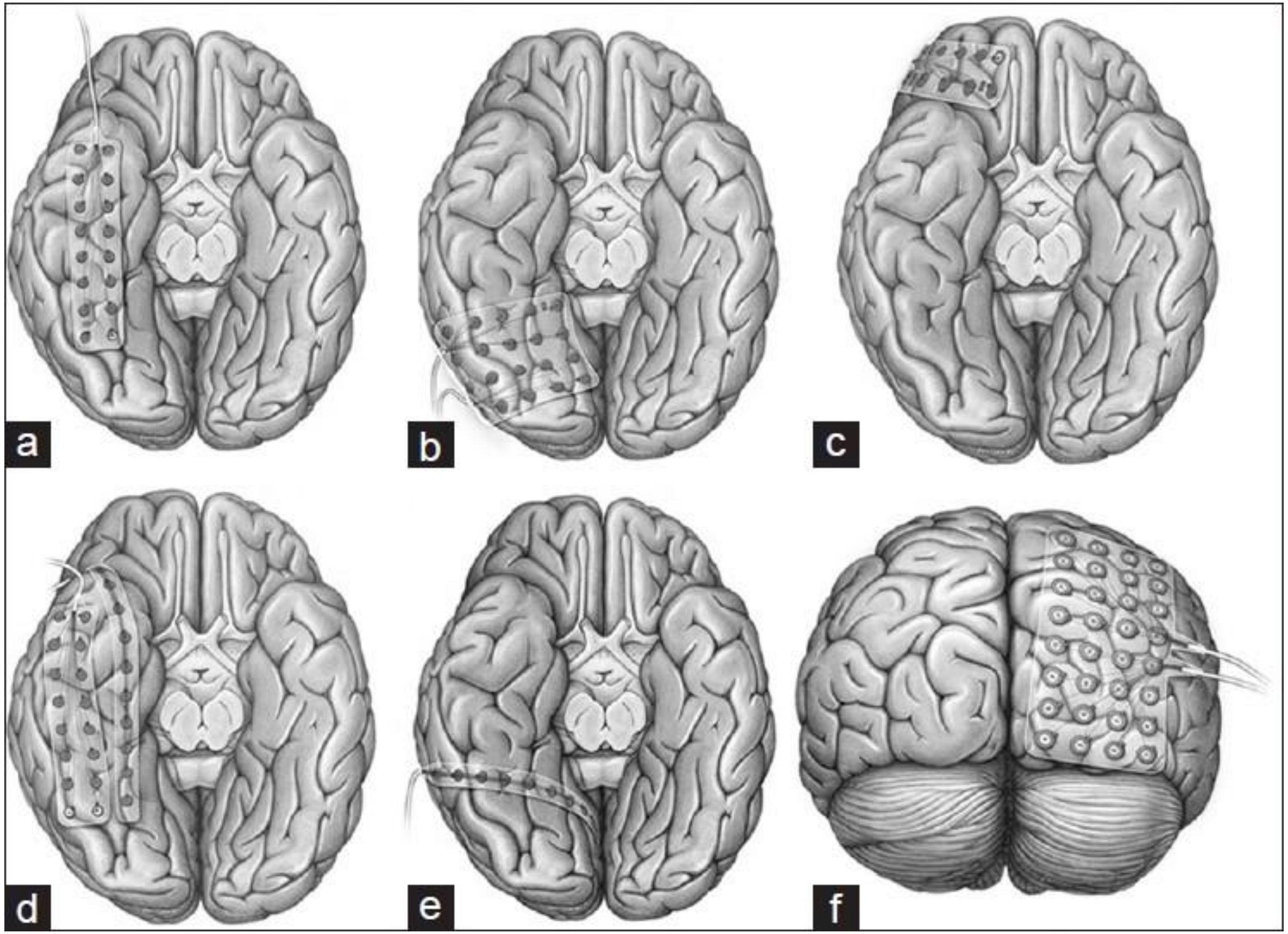
Fusiform Face Area



Fusiform Face Area



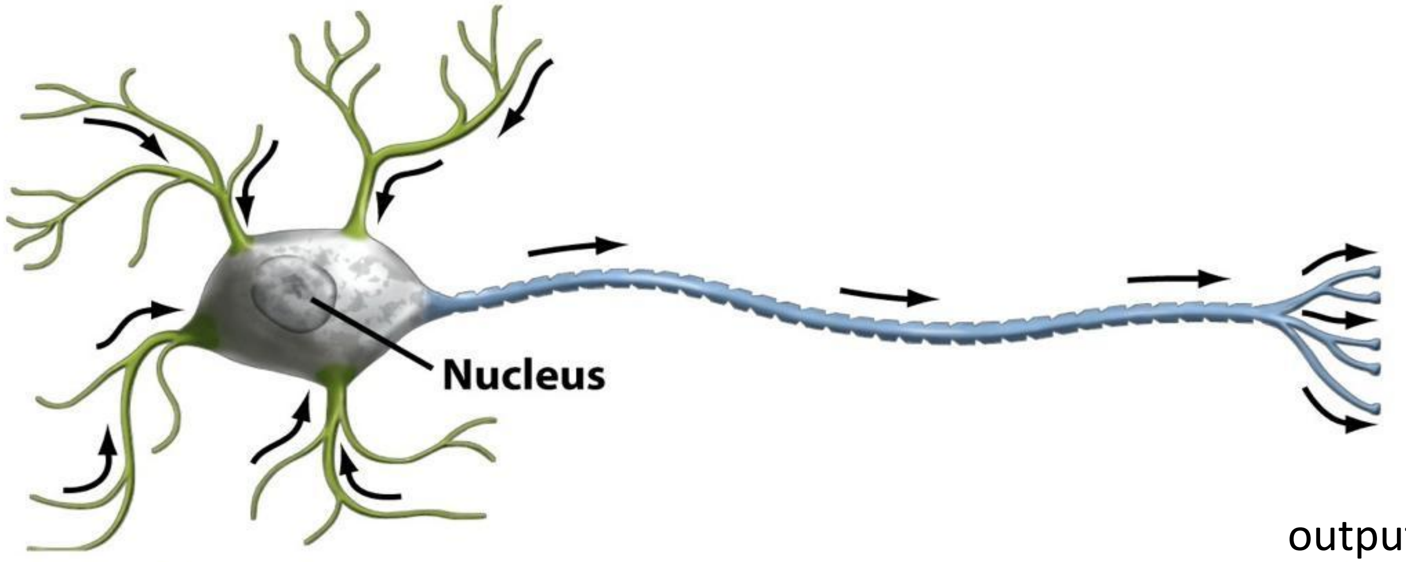
Electrocorticography



Electrocorticography

Vision Models & CNNs

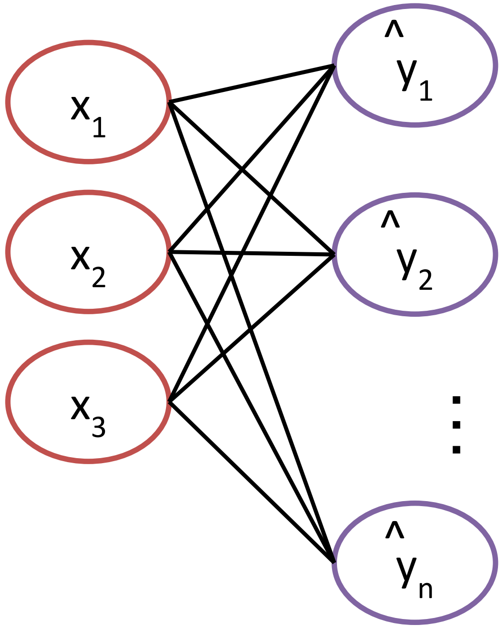
Inspiration



output
([1, 0]
fire, don't fire)

input
(many inputs from many
cells, weighted with
different importance)

Neural Nets



Can have multiple outputs

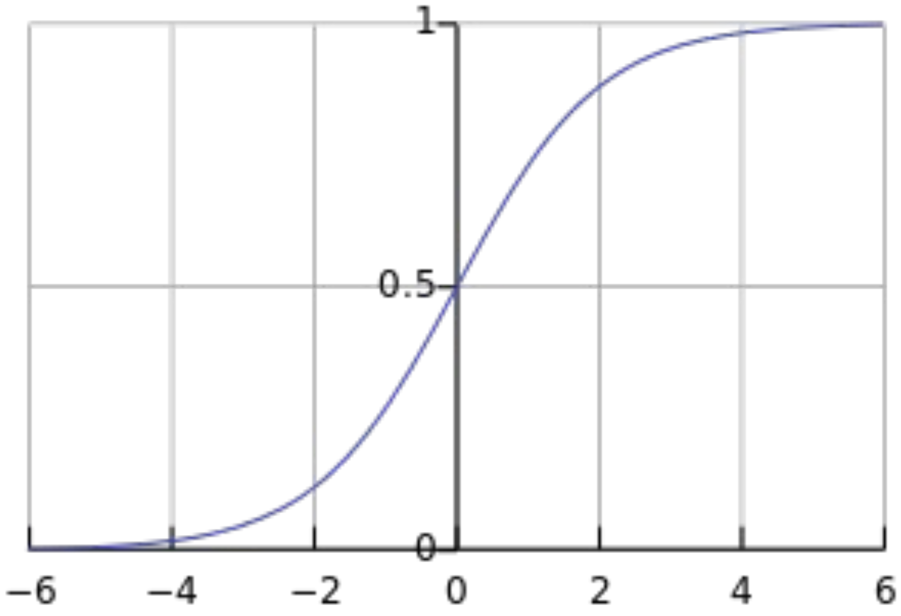
Activation Functions

The “original” nonlinearity was the sigmoid function, along with hyperbolic tangent (tanh).

$$S(z) = \frac{1}{1+e^{-z}}$$

More recently:

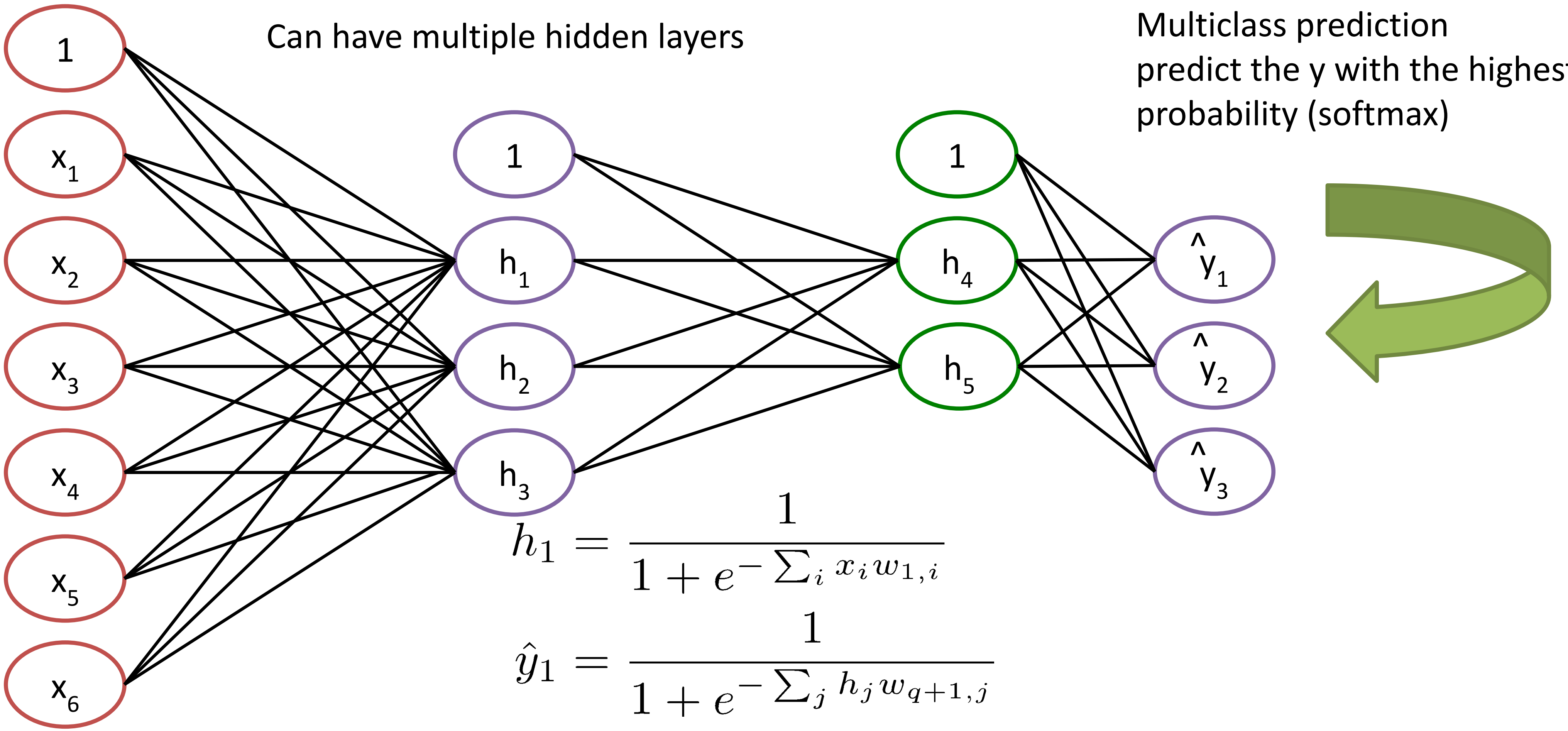
- RELU - Rectified Linear Unit
- GELU - Gaussian Error Linear Unit [Transformers, GPT3]
- Swish -
- Softplus - approximation to ReLU, smoother
- ELU (Exponential Linear Unit)
- SELU (scaled ELU)



Neural Nets

Can have multiple hidden layers

Multiclass prediction
predict the y with the highest probability (softmax)



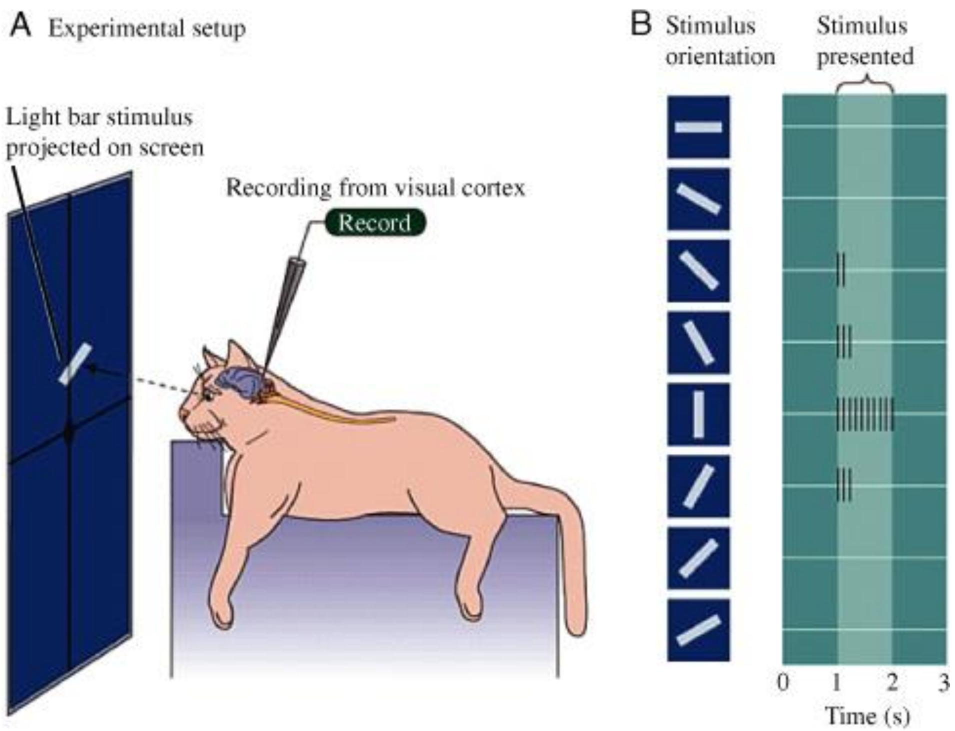
$$h_1 = \frac{1}{1 + e^{-\sum_i x_i w_{1,i}}}$$

$$\hat{y}_1 = \frac{1}{1 + e^{-\sum_j h_j w_{q+1,j}}}$$

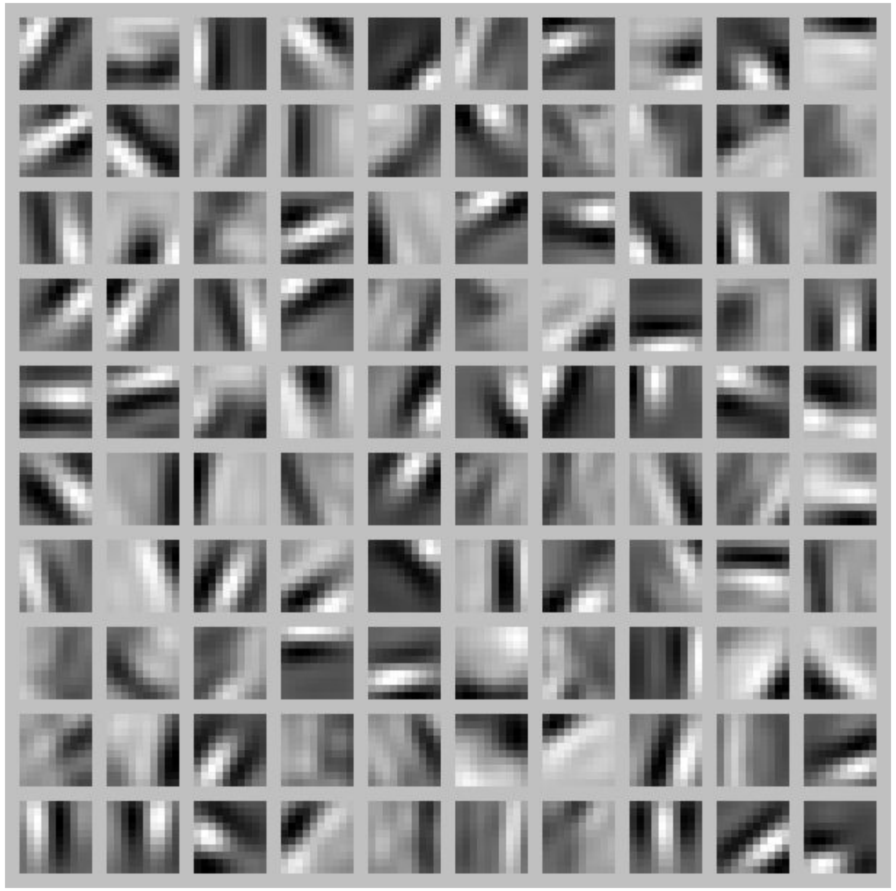
q is total # of hidden nodes across all layers

An Aside

- In this video, the static noise you hear is a representation of the neurons firing in response to the visual stimulus
  - https://www.youtube.com/watch?v=jw6nBWo21Zk



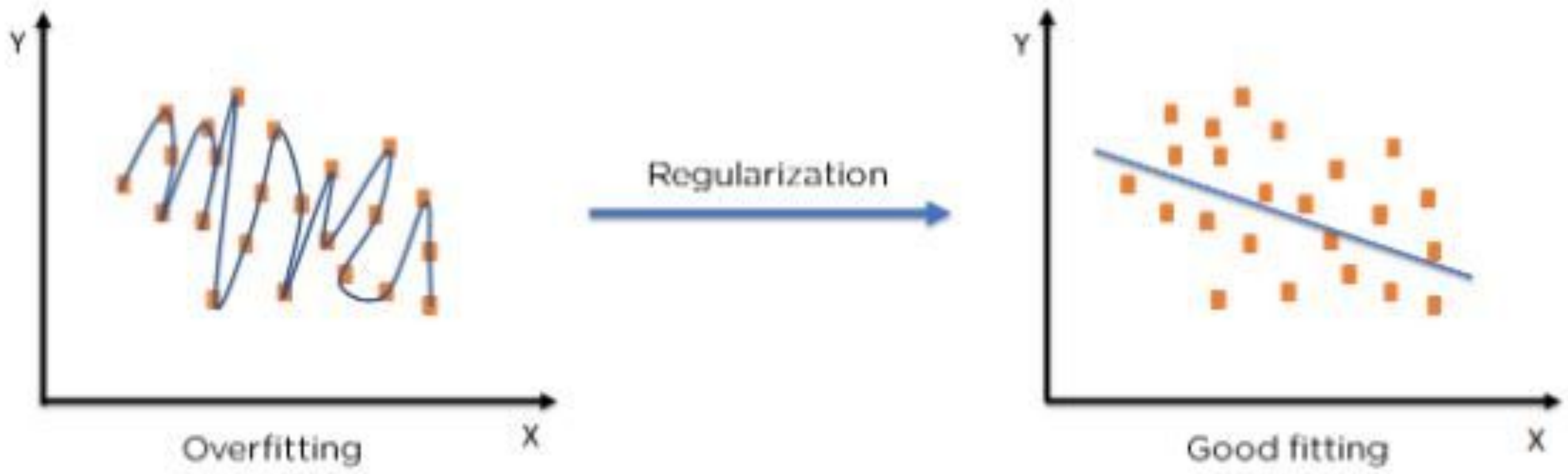
NNs learn something similar



http://deeplearning.stanford.edu/wiki/index.php/Visualizing_a_Trained_Autoencoder

Regularisation

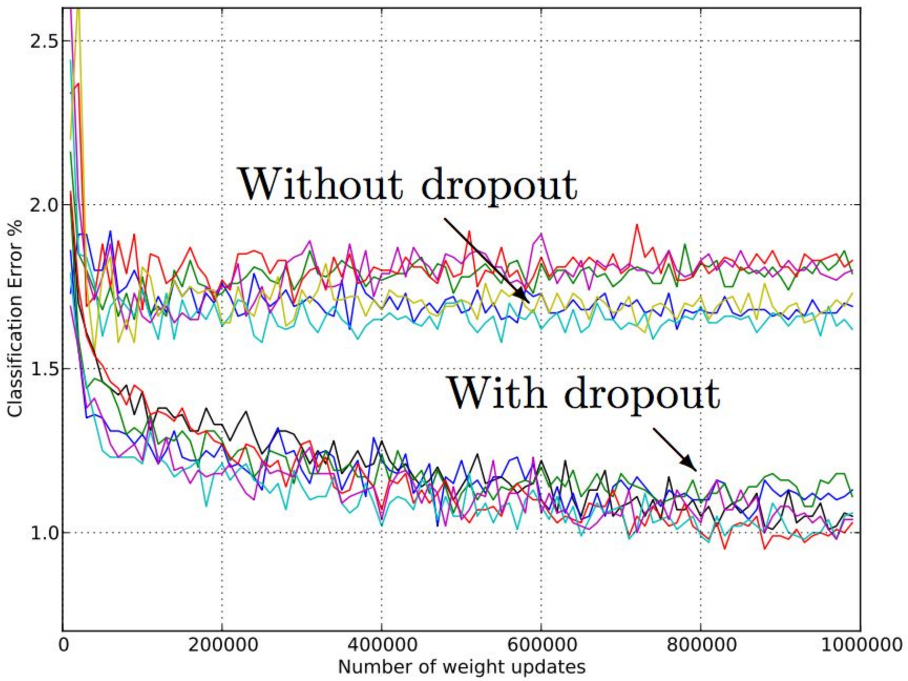
- We are learning a much larger number of parameters than, e.g. SVM
- This makes the model prone to overfitting
- One way to control: regularization



https://www.simplilearn.com/tutorials/machine-learning-tutorial/regularization-in-machine-learning

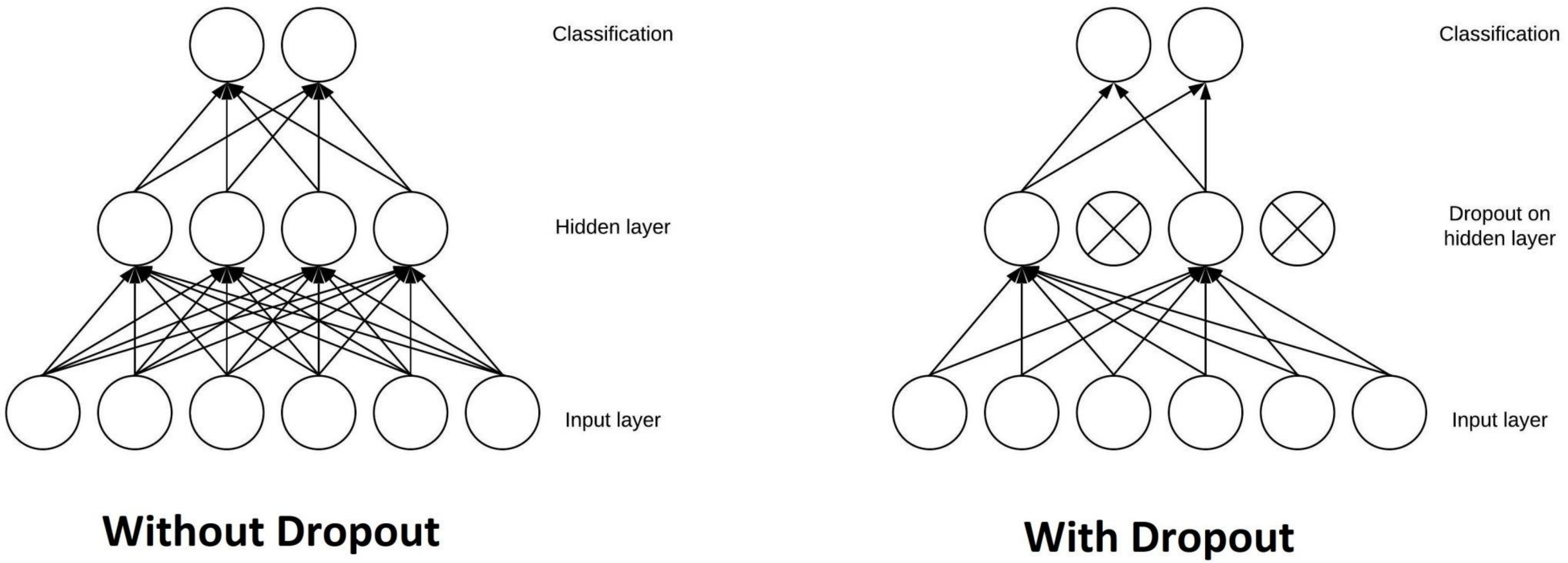
Dropout

- randomly remove nodes from the network
  - hidden and visible
- encourages redundant connections
- reduces overfitting
- kind of like an ensemble



https://www.cs.toronto.edu/~hinton/absps/JMLRdropout.pdf

Dropout



https://www.baeldung.com/cs/ml-relu-dropout-layers

Convolutional neural nets (CNNs)

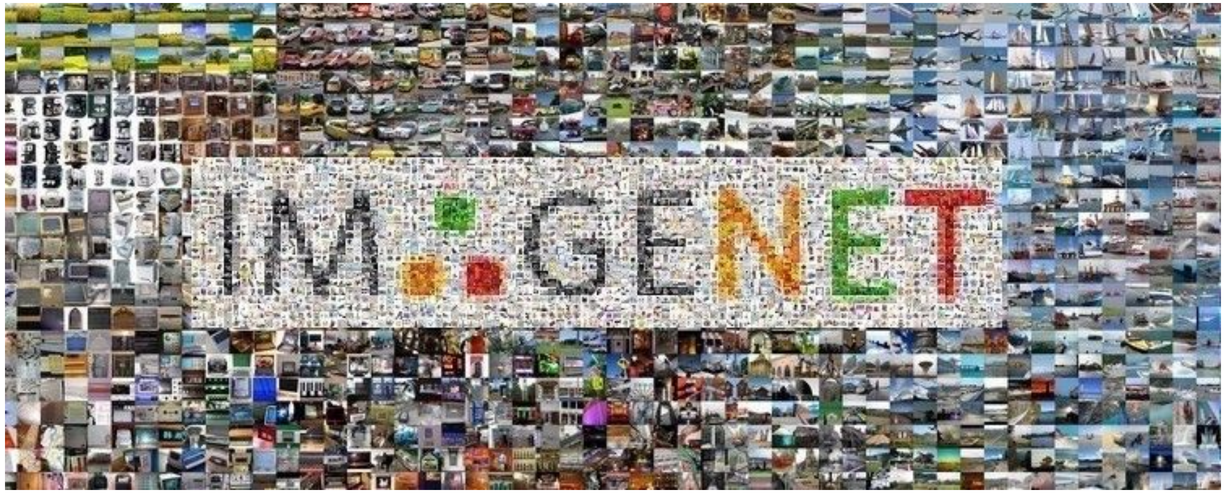
- Built to capture the invariances we see in images
  - objects can appear at any place in the image

Convolutional neural nets (CNNs)

- Learn filters that respond to particular visual features
  - e.g. edges, curves, etc
- Look for those visual features anywhere in the image
  - i.e. convolve the *same* filter with many patches of the image
- Hidden layers combine these filters to create more complex shapes
  - e.g. straight edges combined to form curves, combined to form the handle on a coffee mug

CNNs

- Amazing advances in computer vision
  - ImageNet
    - First released 2009
    - 1.2 million images
    - more than 1000 concepts
      - e.g. cup, oil filter, ptarmigan
  - Deep learning
    - CNNs



http://image-net.org/

ImageNet

http://cs.stanford.edu/people/karpathy/cnnembed/

ImageNet

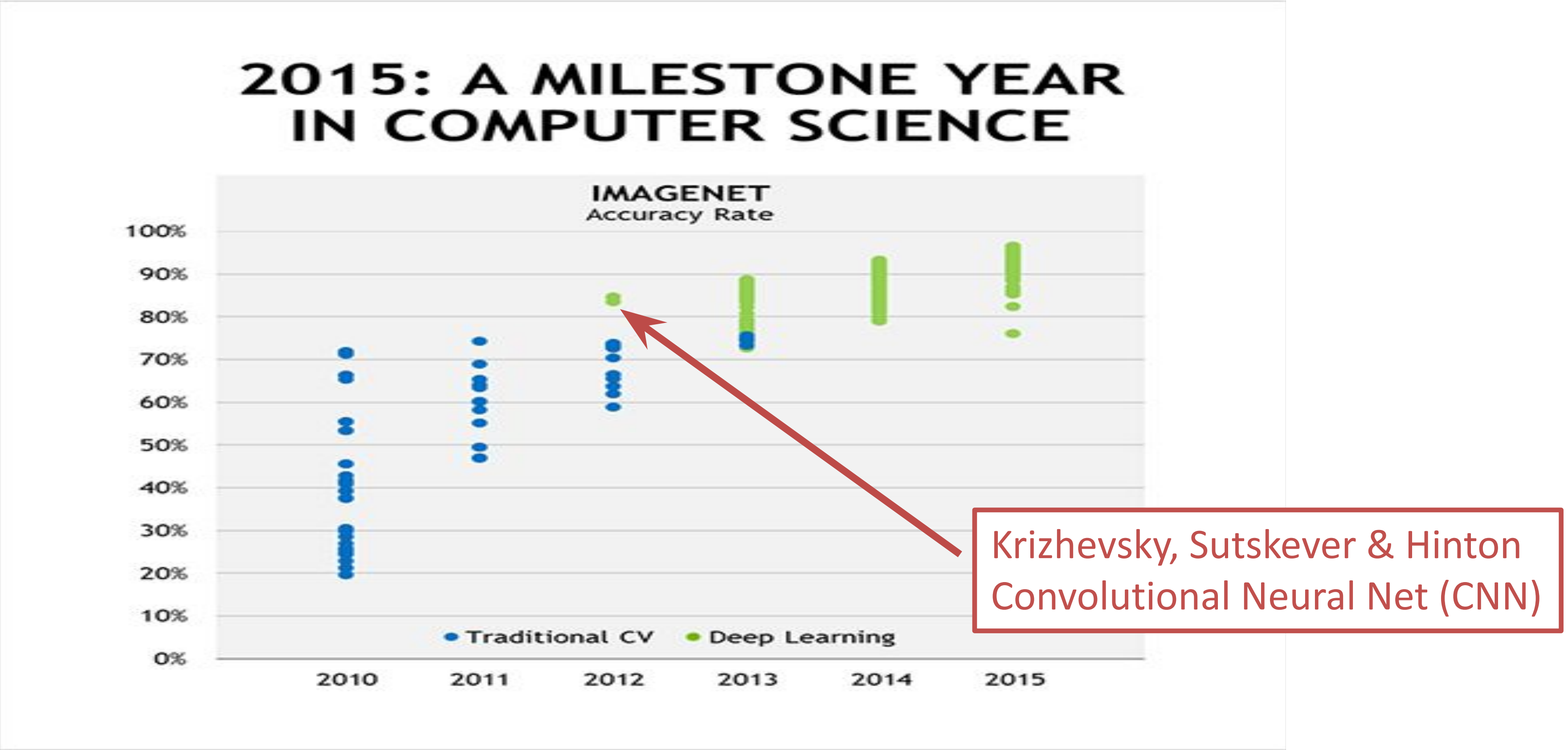
french fries mashed potato black olive face powder crab apple Granny Smith strawberry blueberry cranberry currant blackberry raspberry persimmon mulberry orange kumquat lemon grapefruit plum fig pineapple banana jackfruit cherry grape custard apple durian mango elderberry guava litchi pomegranate quince kidney bean soy green pea chickpea chard lettuce cress spinach bell pepper pimento jalapeno cherry tomato parsnip turnip mustard bok choy head cabbage broccoli cauliflower brussels sprouts zucchini spaghetti squash acorn squash butternut squash cucumber artichoke asparagus green onion shallot leek cardoon celery mushroom pumpkin cliff lunar crater valley alp volcano promontory sandbar dune coral reef lakeside seashore geyser bakery juniper berry gourd acorn olive hip ear pumpkin seed sunflower seed coffee bean rapeseed corn buckeye bean peanut walnut cashew chestnut hazelnut coconut pecan pistachio lentil pea peanut okra sunflower lesser celandine wood anemone blue columbine delphinium nigella calla lily sandwort pink baby's breath ice plant globe amaranth four o'clock Virginia spring beauty wallflower damask violet candytuft Iceland poppy prickly poppy oriental poppy celandine blue poppy Welsh poppy celandine poppy corydalis pearly everlasting strawflower yellow chamomile dusty miller tansy daisy common marigold China aster cornflower chrysanthemum mistflower cosmos dahlia coneflower blue daisy gazania African daisy male orchis butterfly orchid aerides brassavola spider orchid grass pink calypso cattleya red helleborine coelogyne cymbid lady's slipper marsh orchid dendrobium disa helleborine fragrant orchid fringed orchis lizard orchid laelia masdevallia odontoglossum oncidium bee orchid fly orchid spider orchid phaius moth orchid ladies' tresses stanhopea stelis vanda cyclamen centaury gentian begonia commelina scabious achimenes African violet streptocarpus scorpionweed calceolaria toadflax veronica bonsai star anise wattle huisache silk tree rain tree dita pandanus linden American beech New Zealand beech live oak shingle oak pin oak cork oak yellow birch American white birch downy birch alder fringe tree European ash fig witch elm Dutch elm cabbage tree golden shower tree honey locust Kentucky coffee tree Brazilian rosewood logwood coral tree Japanese pagoda tree kowhai palm Arabian coffee cork tree weeping willow pussy willow goat willow China tree pepper tree balata teak ginkgo pine ilang-ilang laurel magnolia tulip tree baobab kapok red beech cacao sorrel tree iron tree mangrove paper mulberry Judas tree redbud mountain ash ailanthus silver maple Oregon maple sycamore box elder Japanese maple holly dogwood truffle shiitake lichen hen-of-the-woods jelly fungus dead-man's-fingers earthstar coral fungus stinkhorn puffball gyromitra bolete polypore gill fungus morel agaric trilobite harvestman scorpion black and gold garden spider barn spider garden spider black widow tarantula wolf spider tick mite centipede millipede horseshoe crab

http://image-net.org/challenges/LSVRC/2010/browse-synsets

ImageNet

http://image-net.org/challenges/LSVRC/2012/analysis/

>> AlexNet Enters the Chat



https://blogs.nvidia.com/blog/2016/01/12/accelerating-ai-artificial-intelligence-gpus/

Self Driving Cars

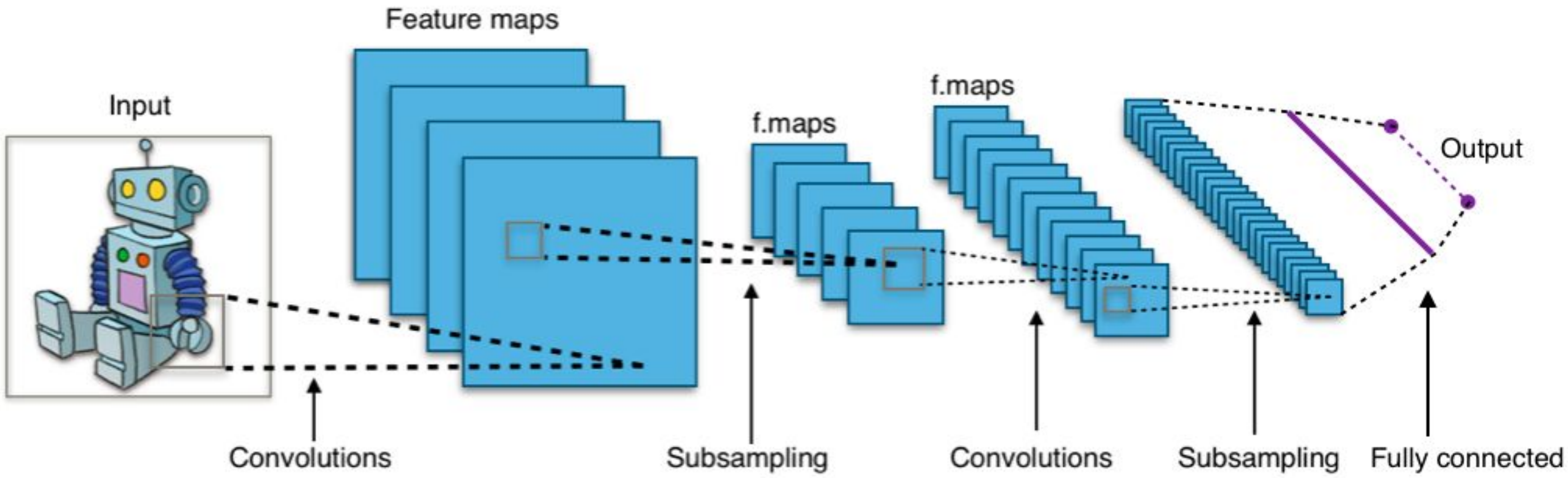
Self Driving Cars

CNNs

Two additional operations:

- Convolution
- Pooling/Subsampling

CNNs: Convolution



https://upload.wikimedia.org/wikipedia/commons/6/63/Typical_cnn.png

CNNs: Convolution

1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

Input image (here binary, RGB typical)

1	0	1
0	1	0
1	0	1

Filter (here binary, but typically continuous values)

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

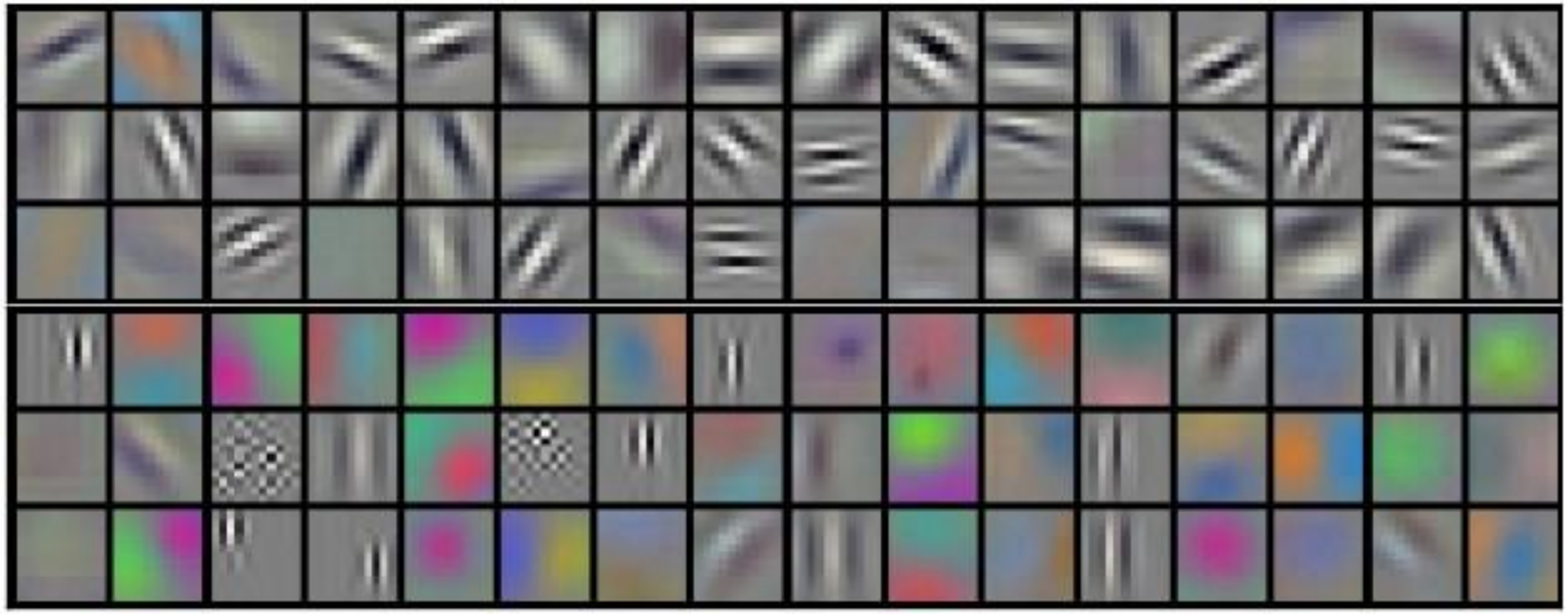
CNNs: Convolution

- Typical Filters (

1	0	1
0	1	0
1	0	1

)

Filters are learned

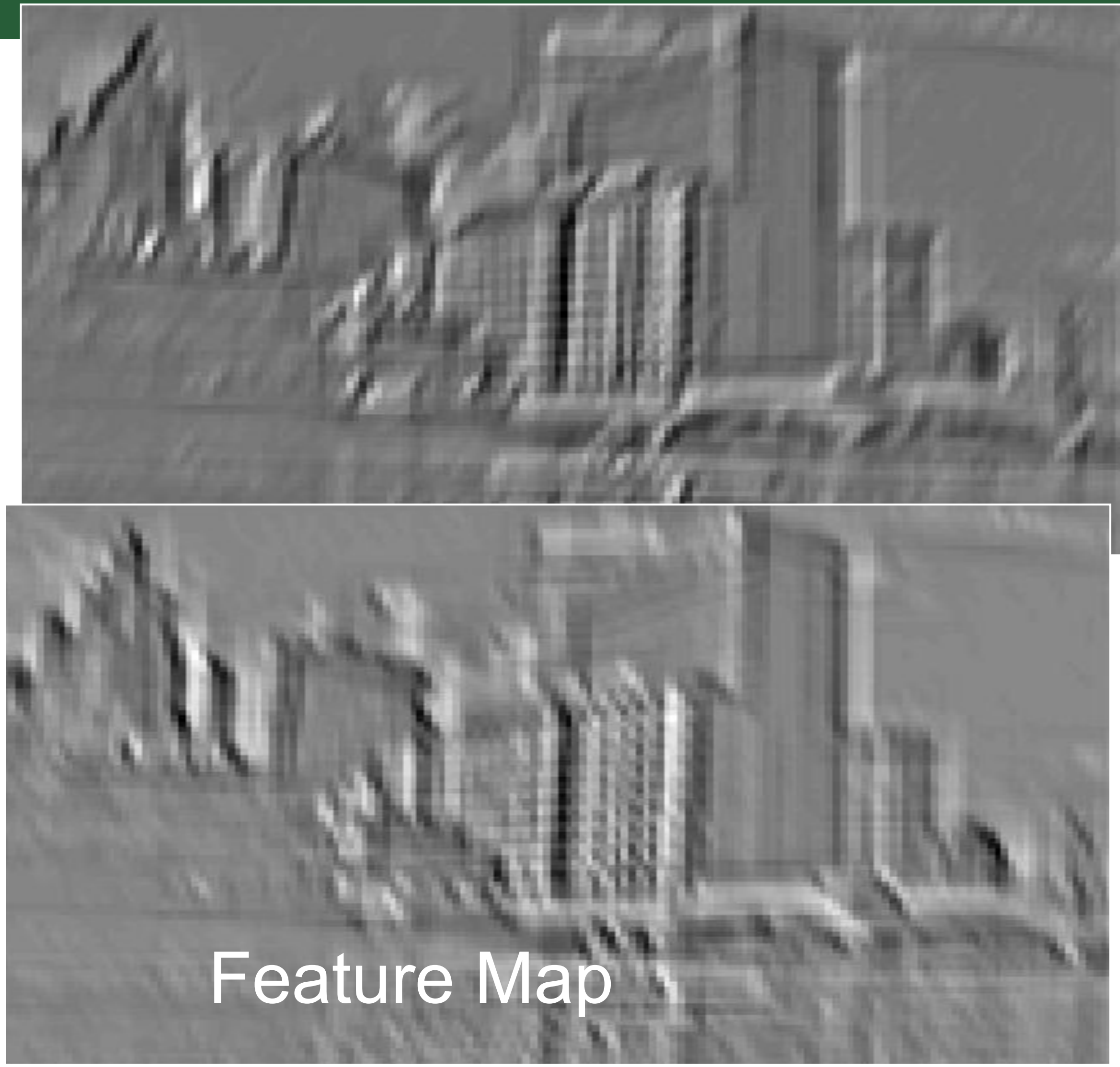


Krizhevsky et al.

CNNs: Convolution

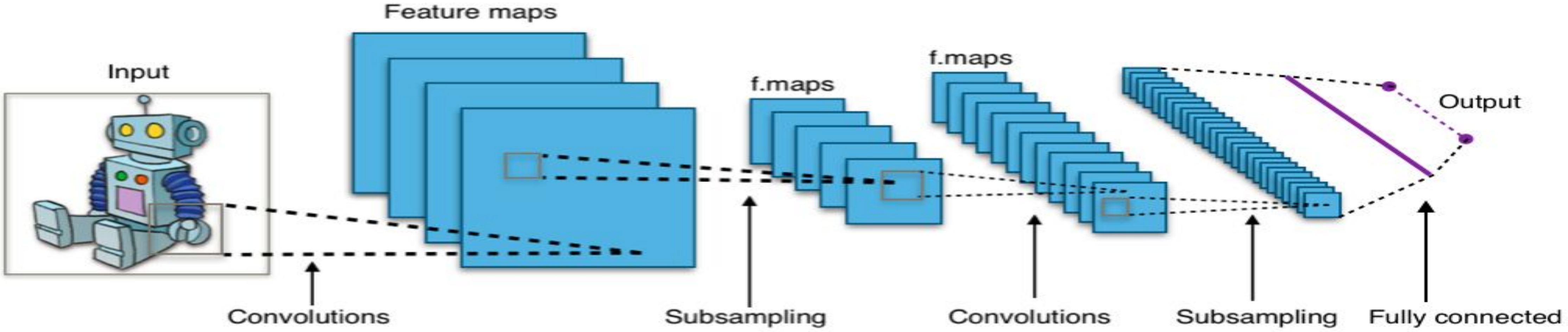
- Output of convolution layer is a feature map

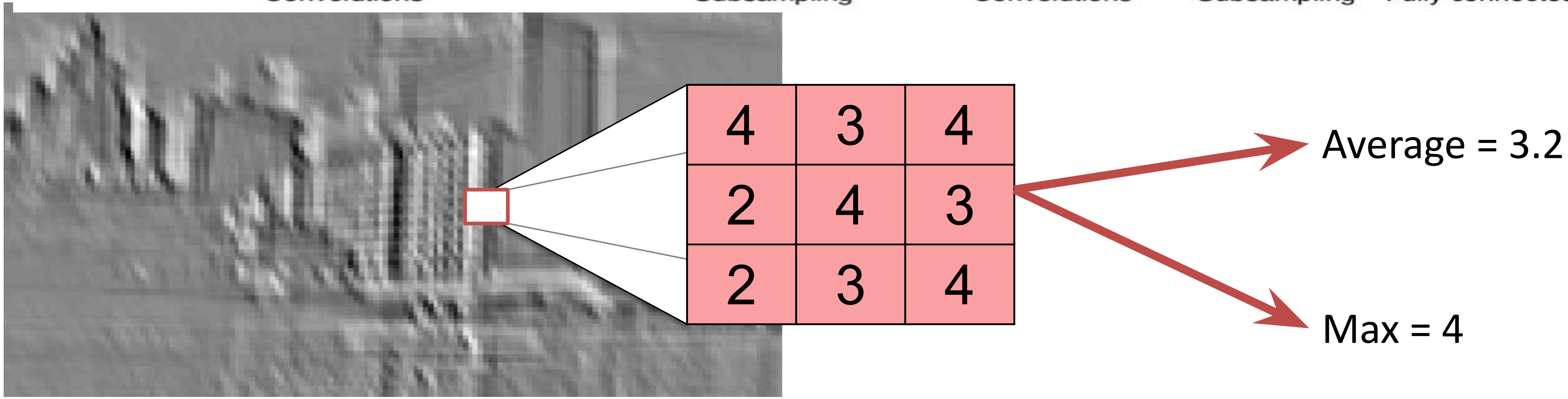
$$\left(\begin{array}{|c|c|c|} \hline 4 & 3 & 4 \\ \hline 2 & 4 & 3 \\ \hline 2 & 3 & 4 \\ \hline \end{array} \right)$$



http://cs.nyu.edu/~fergus/tutorials/deep_learning_cvpr12/ (Fergus intro)

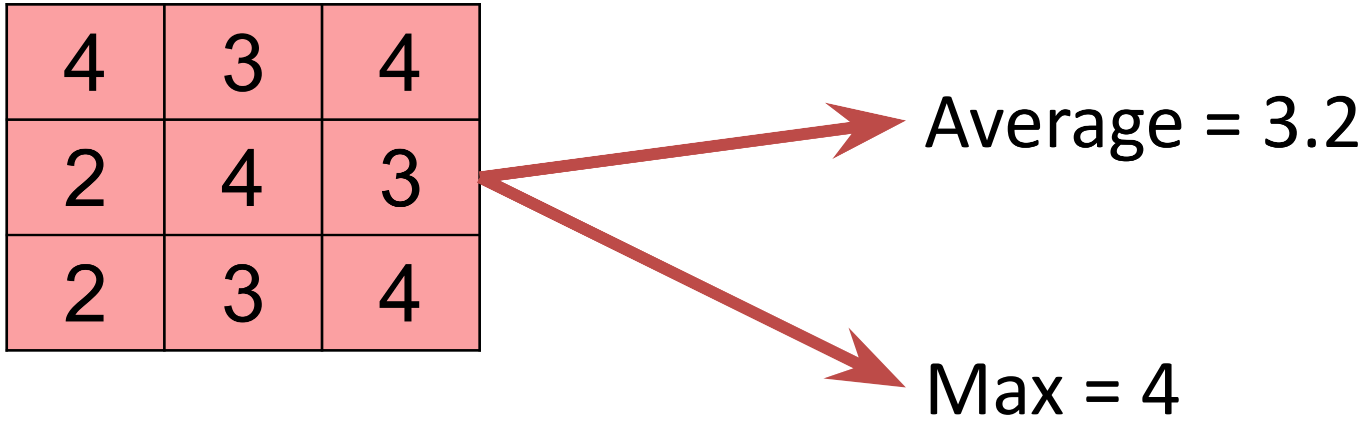
CNNs: Pool/Subsample





Krizhevsky et al.

CNNs: Pool/Subsample

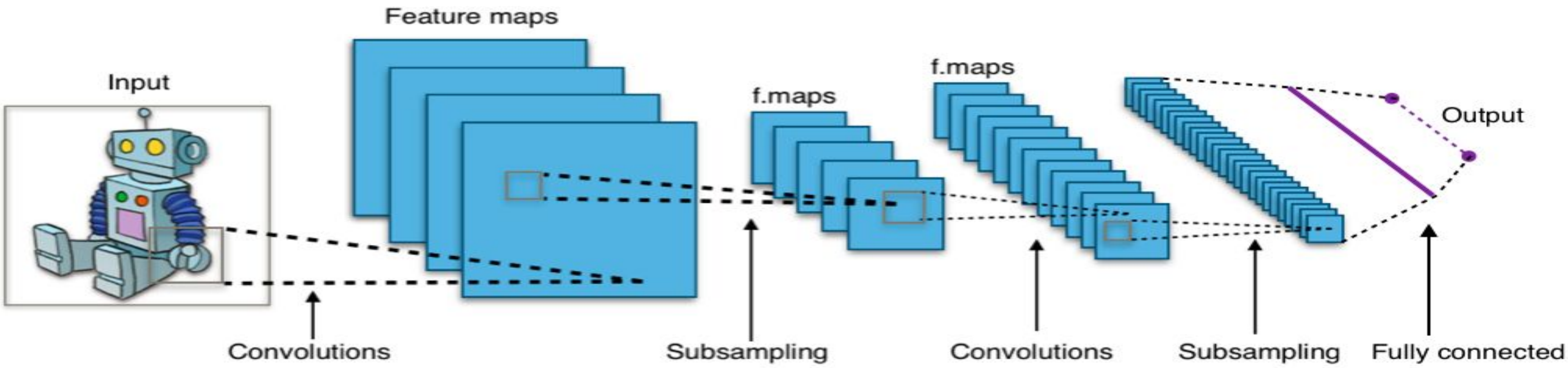


- Supplies some **invariance** to local translations
- Reduces the dimension of subsequent layers

Krizhevsky et al.

CNNs: Hidden Representations

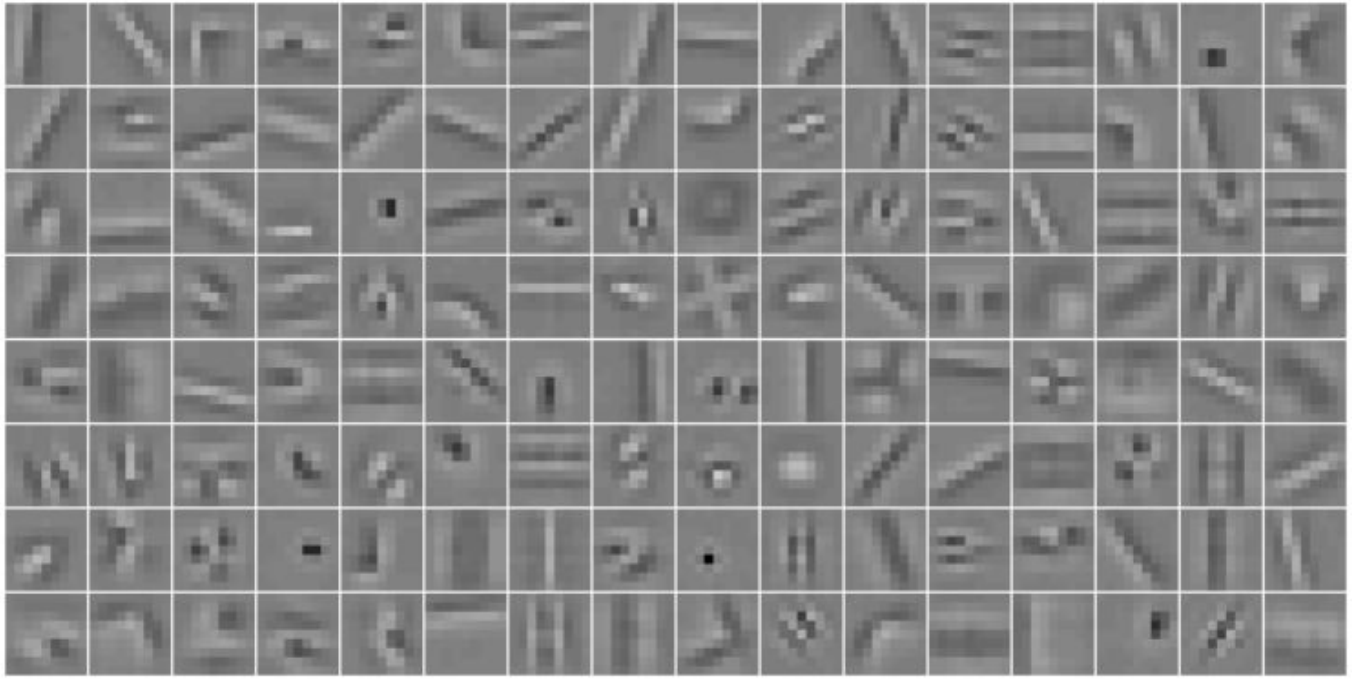
- Activations (feature maps) at each layer are a “hidden representation” of the image
- A compression of the information in image
  - not unlike PCA/SVD



What do CNNs learn?

1	0	1
0	1	0
1	0	1

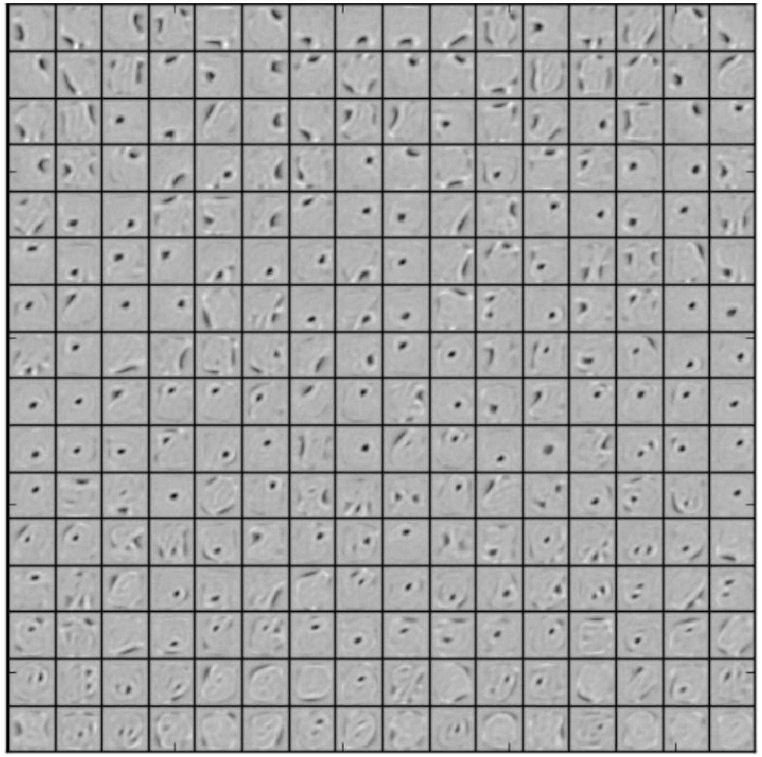
- When trained on images



http://yann.lecun.com/exdb/publis/pdf/lecun-eccv-12.pdf

What do CNNs Learn?

1	0	1
0	1	0
1	0	1

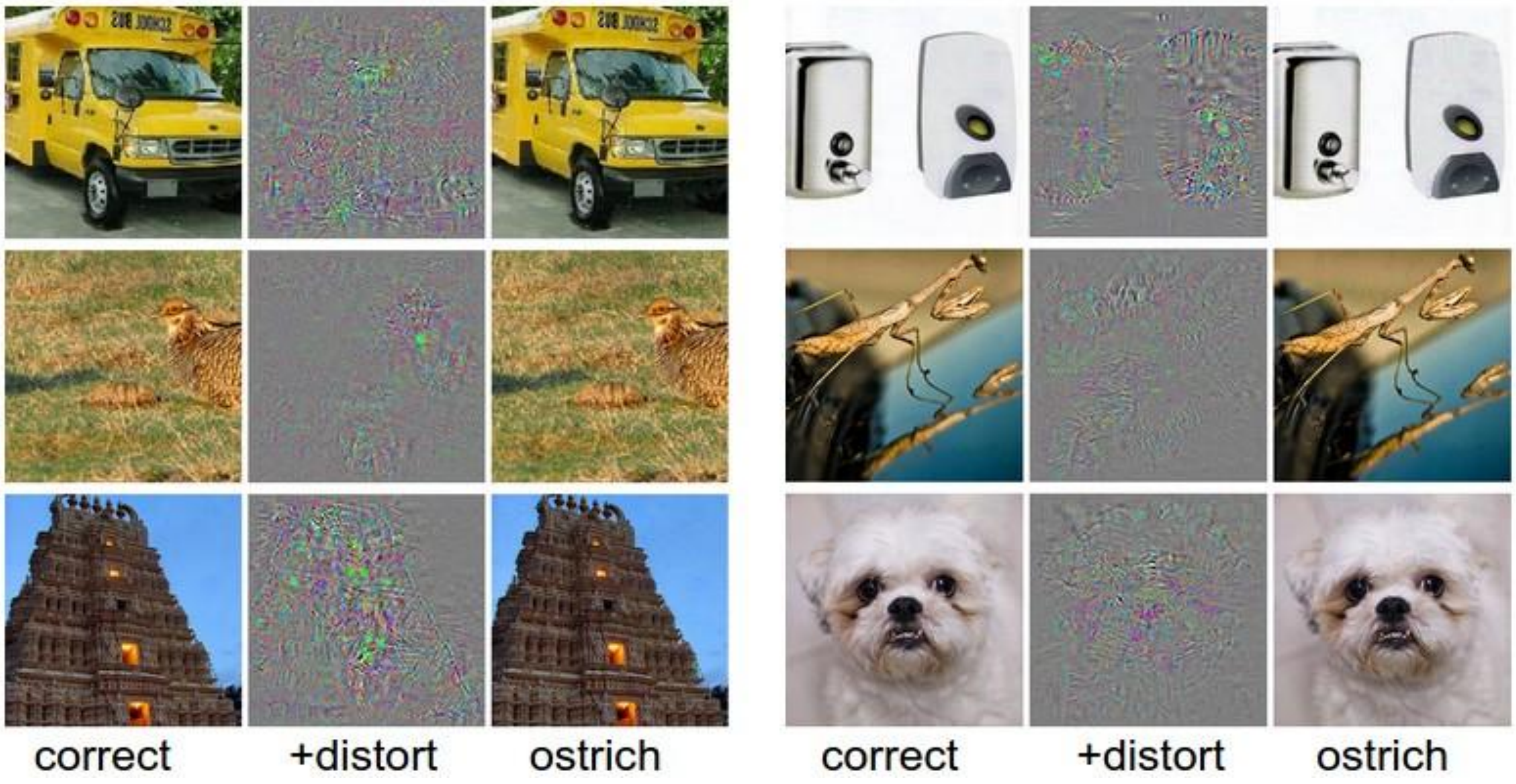


https://www.cs.toronto.edu/~hinton/absps/JMLRdropout.pdf

Adversarial Examples

- We can make small changes to an image (nearly imperceptible!) and cause a network to misclassify
- Extreme implications for e.g. self driving cars

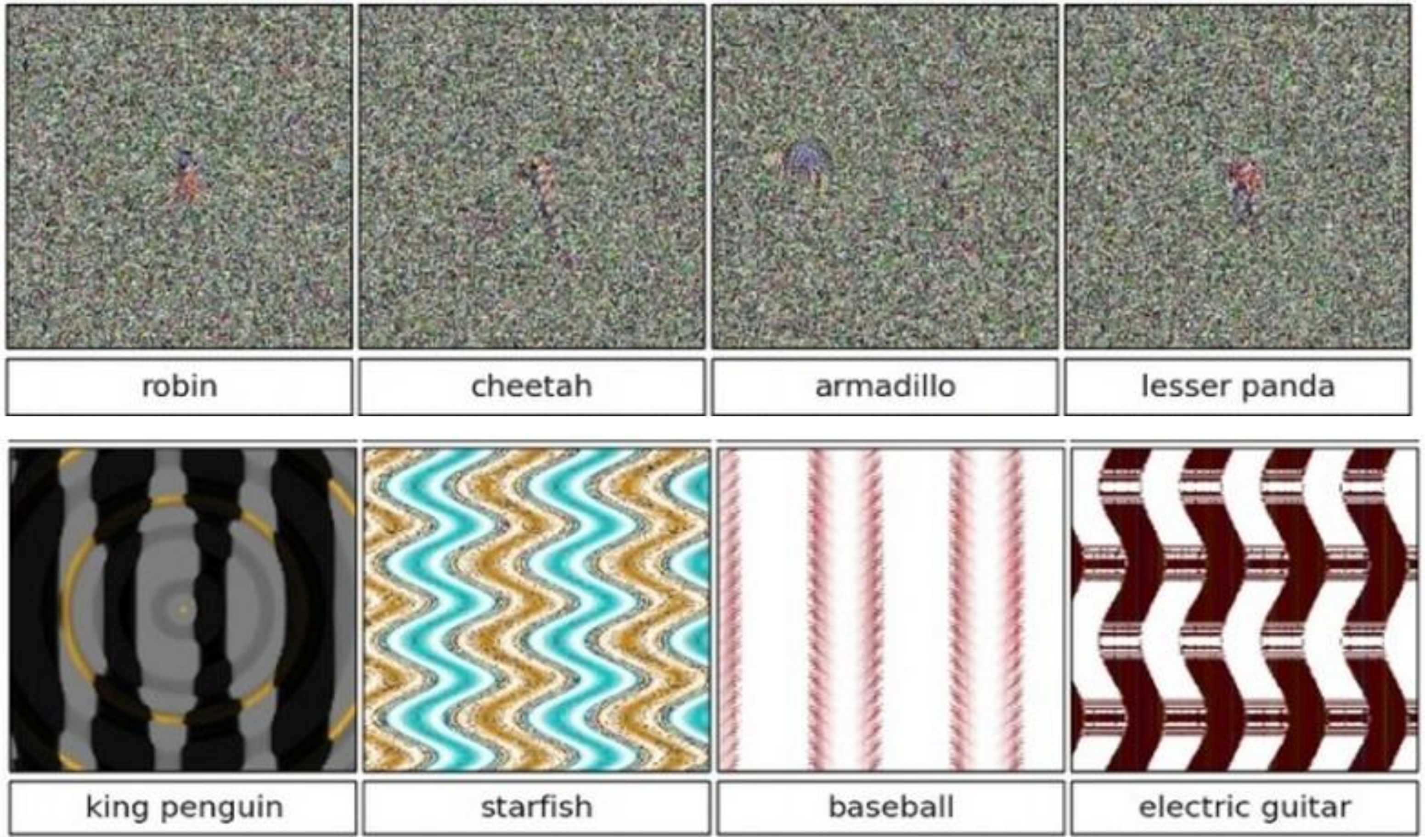
Adversarial Examples



http://karpathy.github.io/2015/03/30/breaking-convnets/

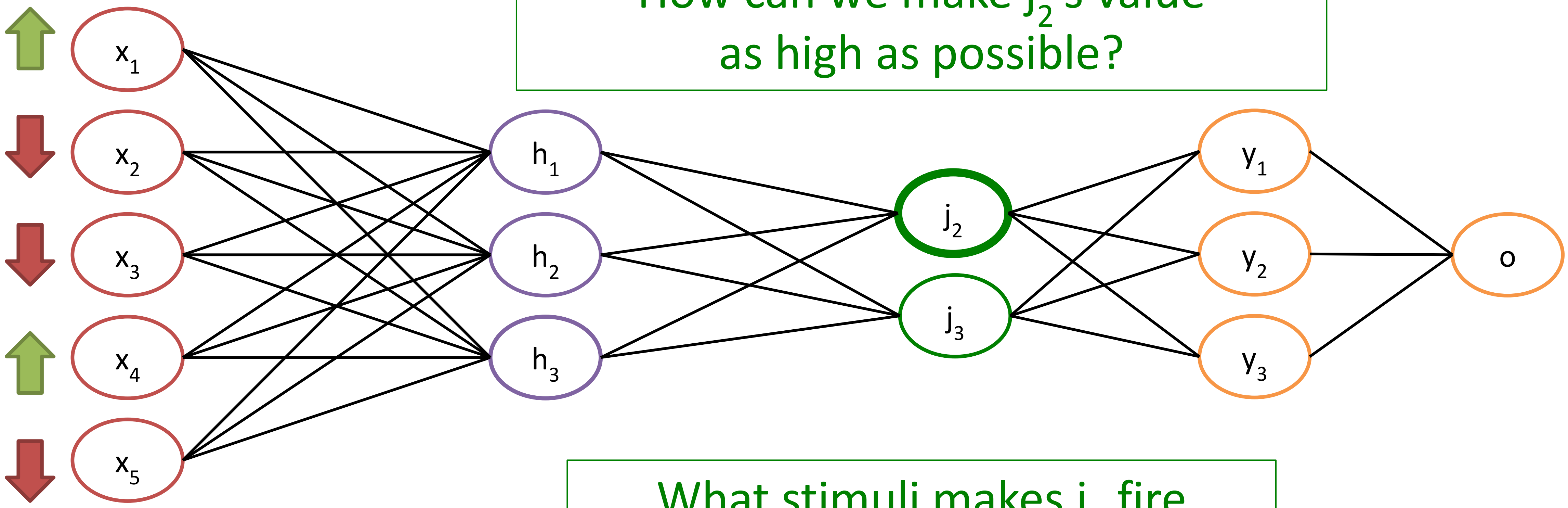
Adversarial Examples

Nonsense images that the classifier is very sure about (>96% prediction prob)



What do the neurons represent?

How can we make j_2's value as high as possible?



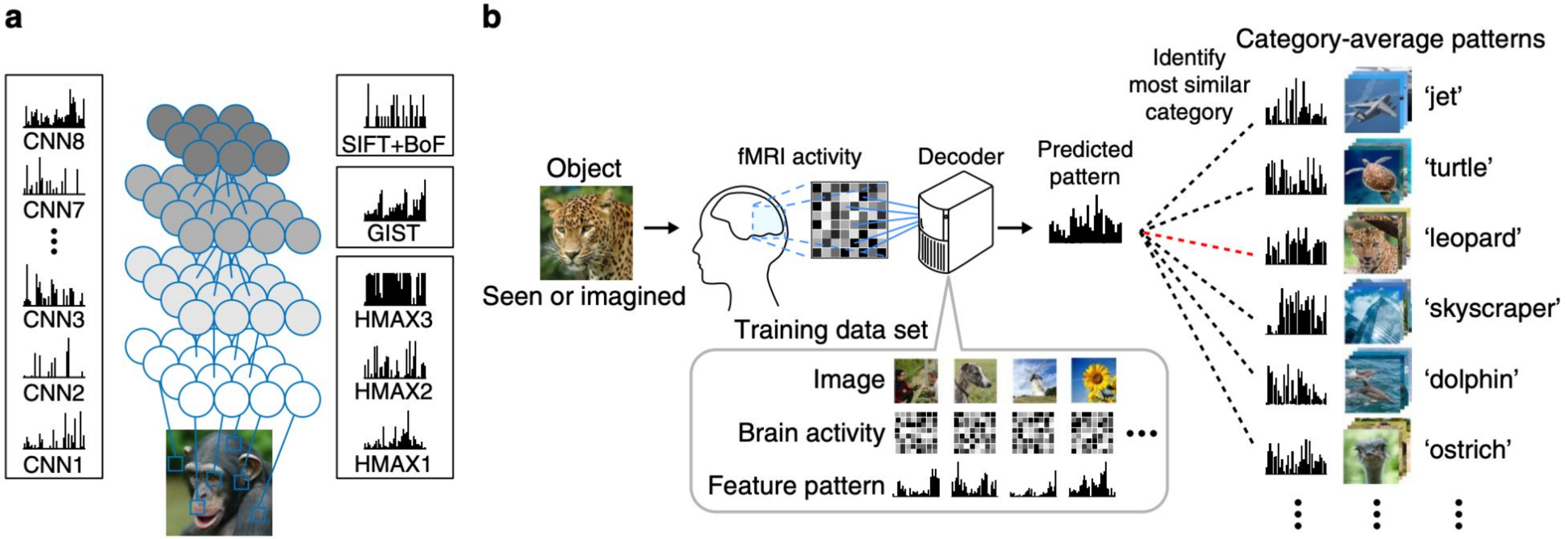
What stimuli makes j_2 fire maximally?

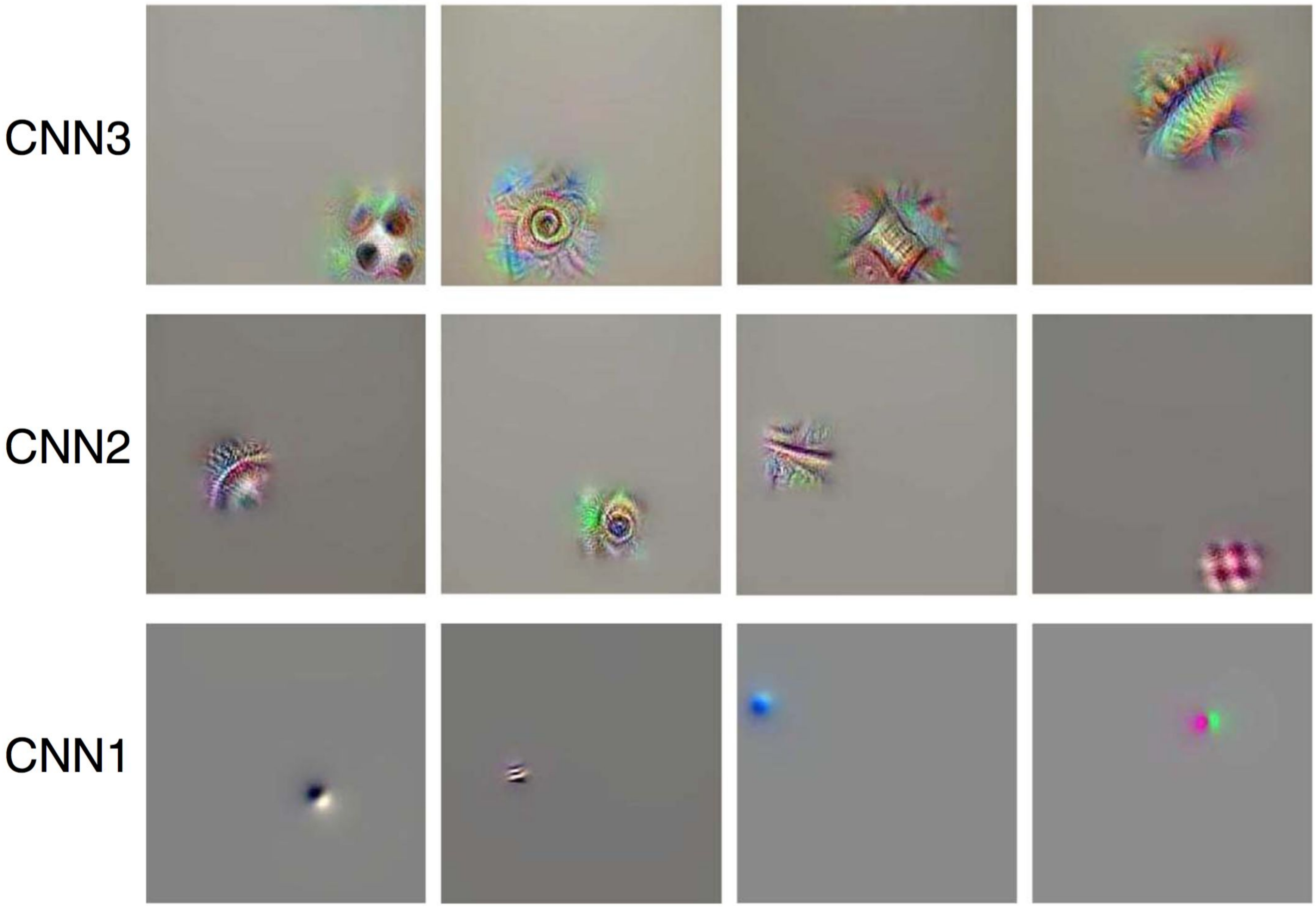
ARTICLE

Received 15 Oct 2015 | Accepted 21 Feb 2017 | Published 22 May 2017 DOI: 10.1038/ncomms15037 OPEN

Generic decoding of seen and imagined objects using hierarchical visual features

Tomoyasu Horikawa[1] & Yukiyasu Kamitani[1,2]





Horikawa & Kamitani (2017)

Horikawa & Kamitani (2017)

Questions?

UNIVERSITY OF ALBERTA