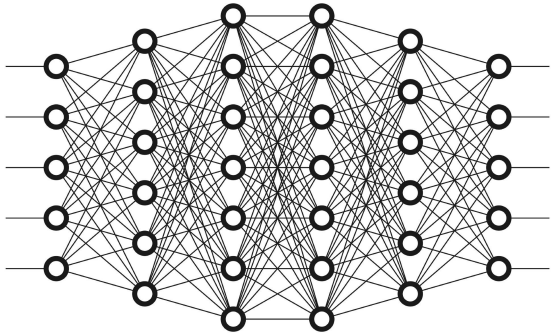
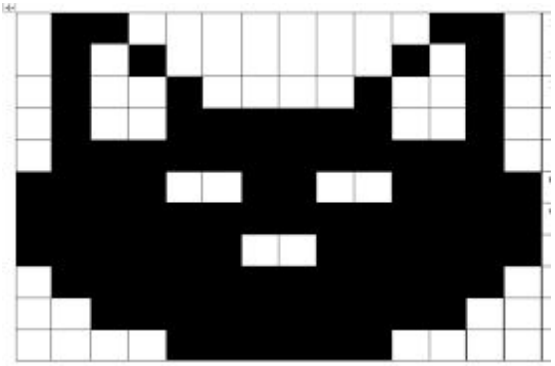


Introduction to Deep Learning

Fall 2018

Images and the Visual System

Elan Barenholtz



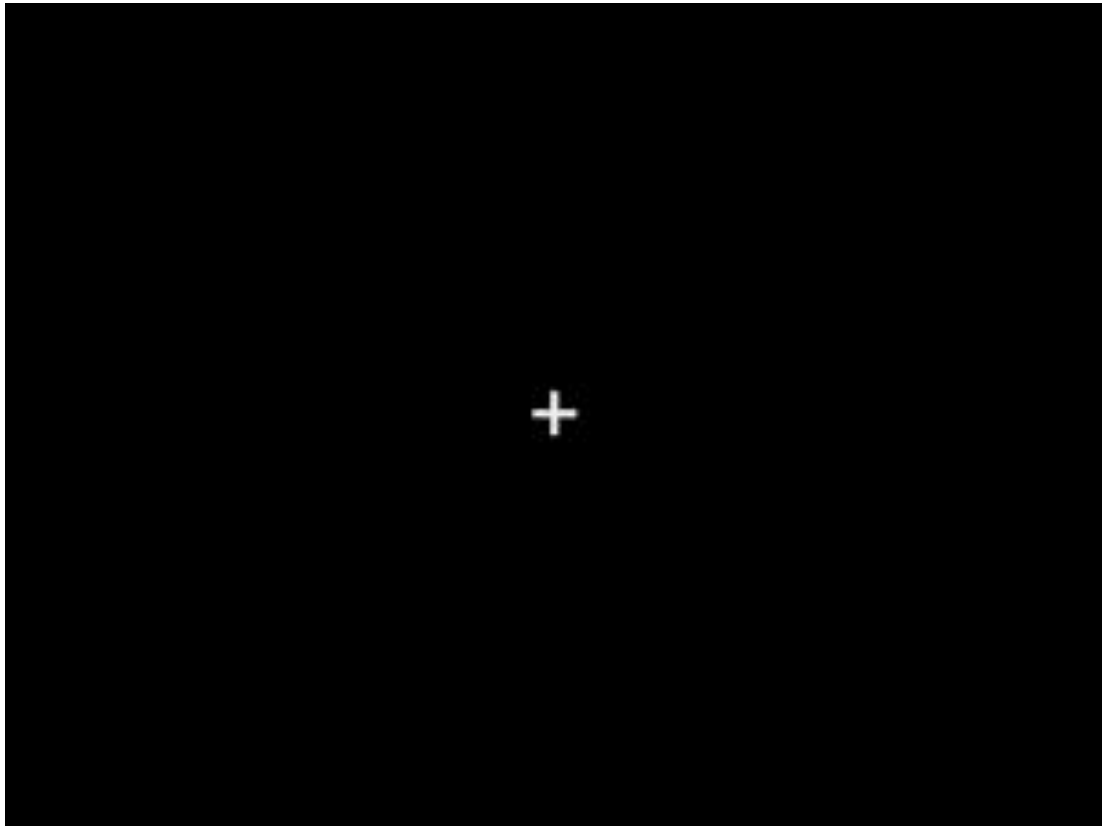
Binary Image: $2^{number\_of\_pixels}$

Grayscale: $256^{number\_of\_pixels}$

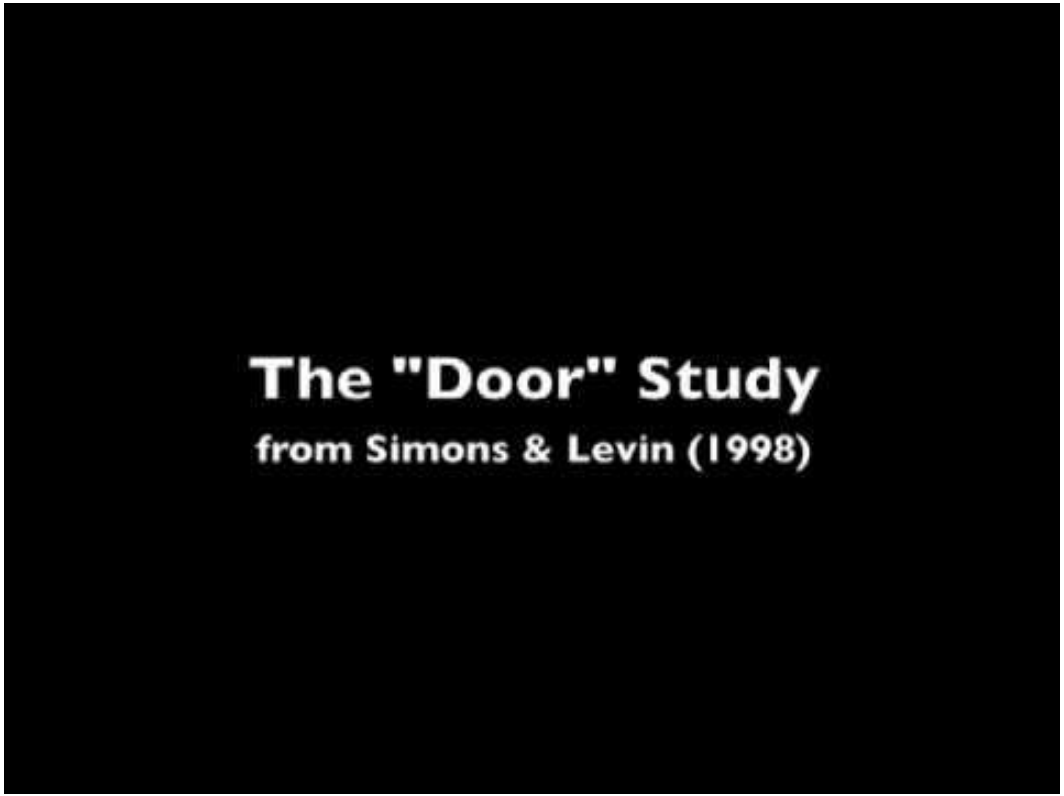
HD image: $(256)^{(3)^{3,145,728}}$

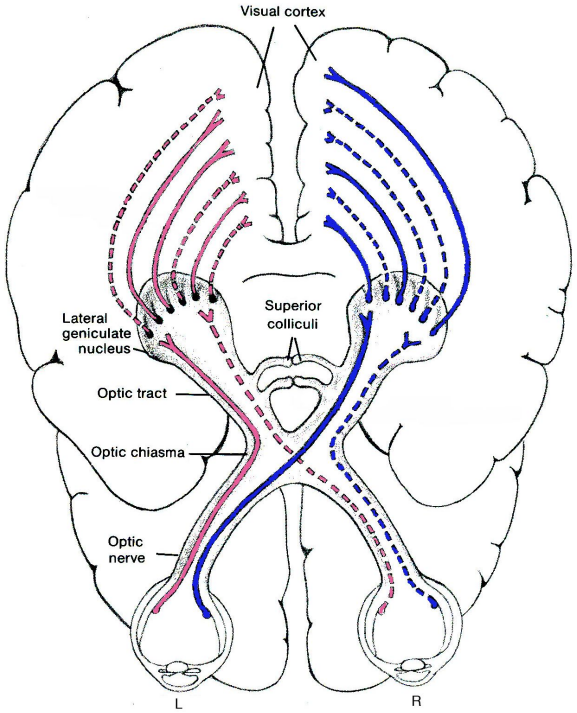
That's a lot of parameters!

Seeing feels easy...



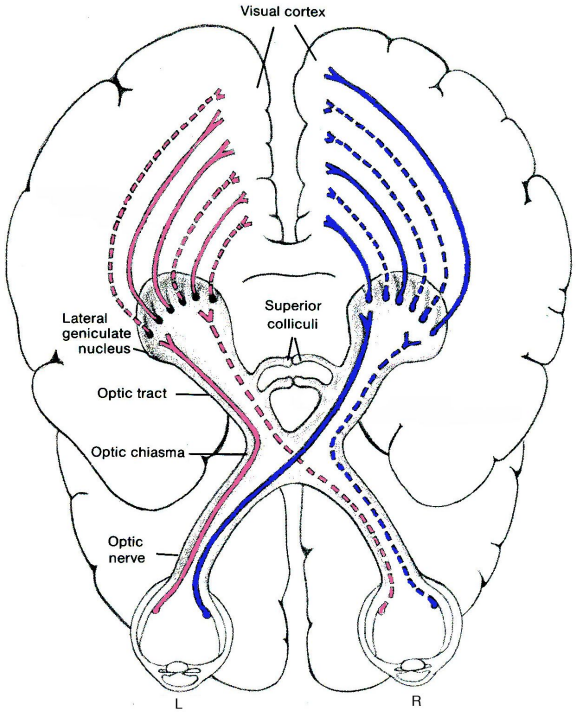
But it is hard!





3 Stages of processing

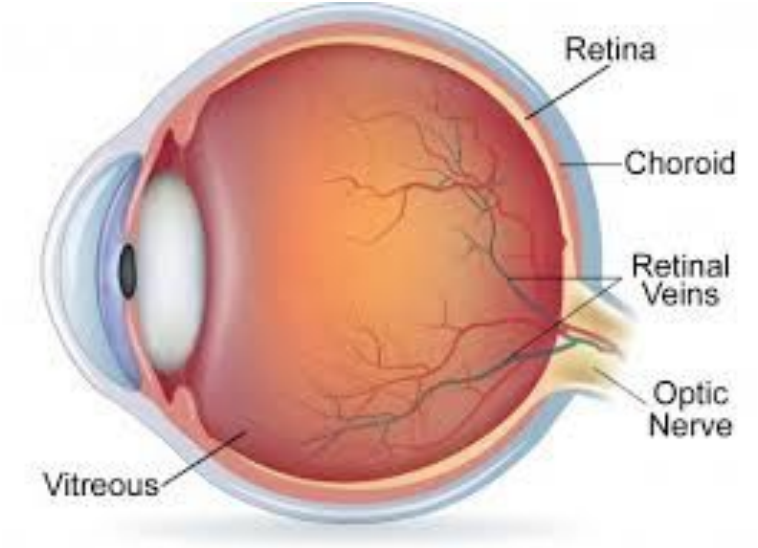
1. Retina (in the eyes)
2. Lateral geniculate nuclues (in the thalamus)
3. Visual Cortex (in the cerebrum)

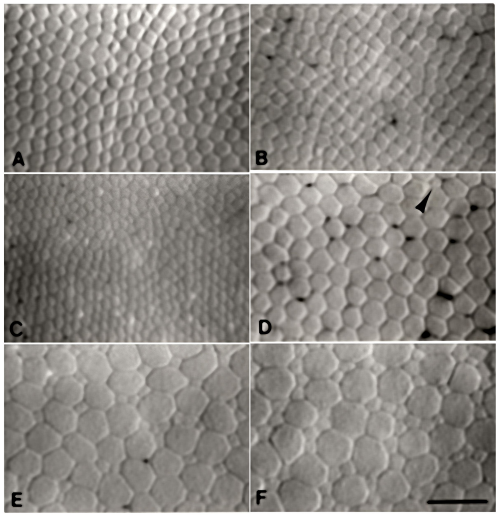


3 Stages of processing

1. Retina (in the eyes) COMPRESSION
2. Lateral geniculate nuclues (in the thalamus) PREPROCESSING
3. Visual Cortex (in the cerebrum) FEATURE LEARNING

Retina

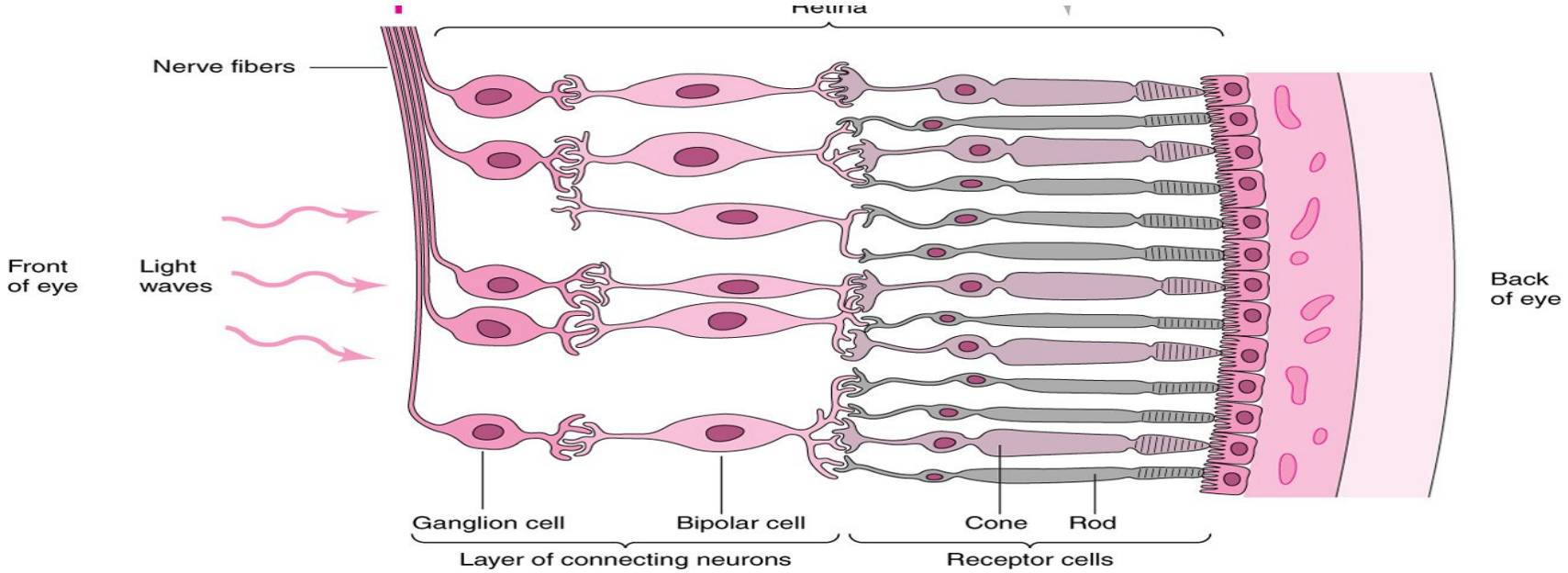




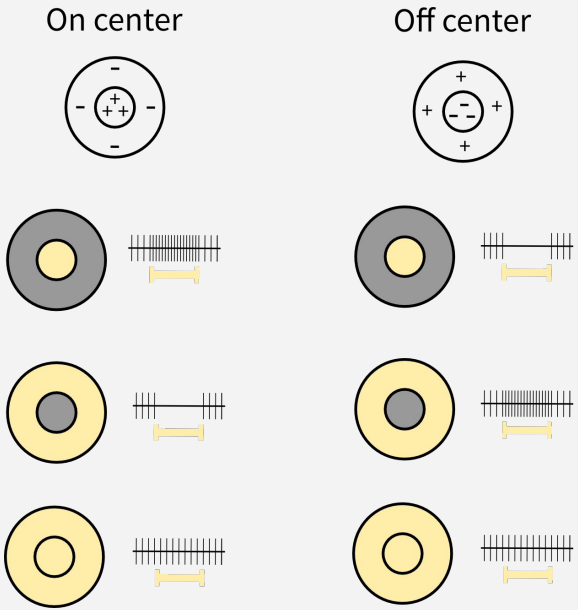
Photoreceptor Mosaic

- Photoreceptors to transduce light into neural signal
- 100 Milliion receptors per eye
- BIG Data!

Retinal Layers

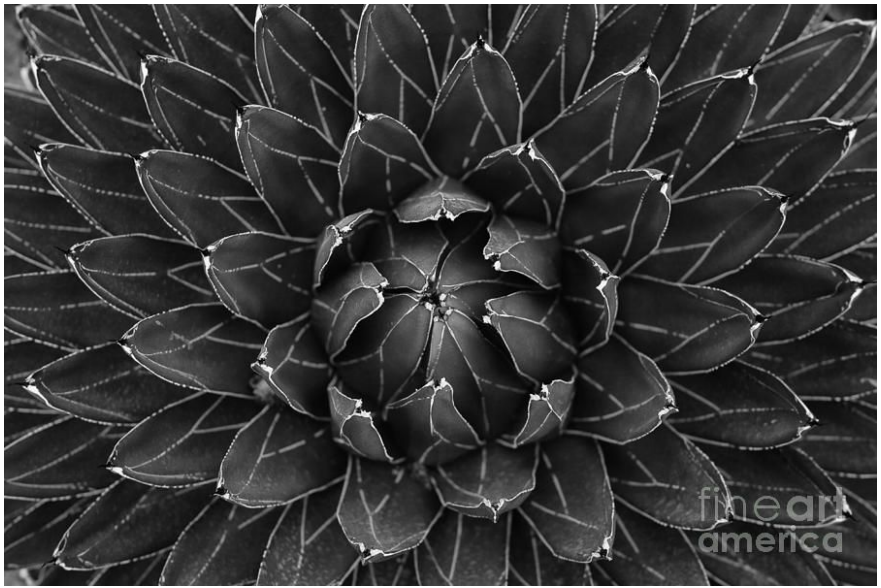


Retinal Ganglion Cells



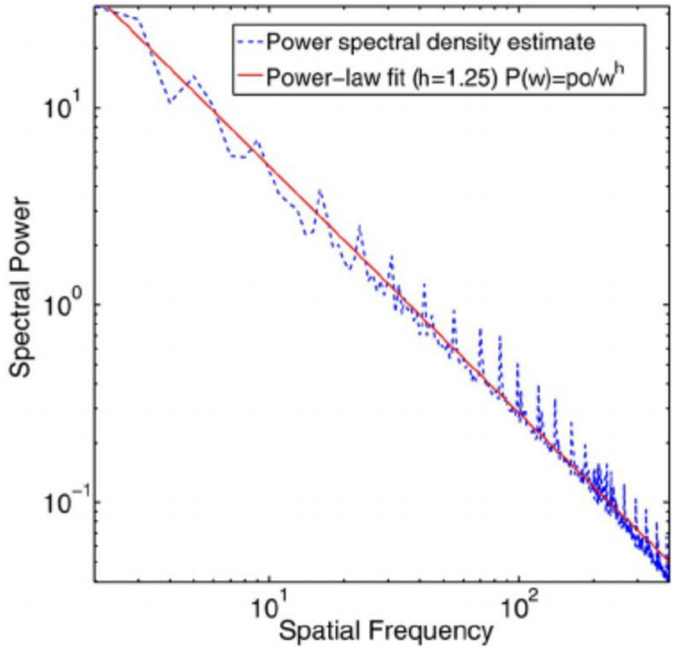
Center-Surround Response

Structure of Natural Images

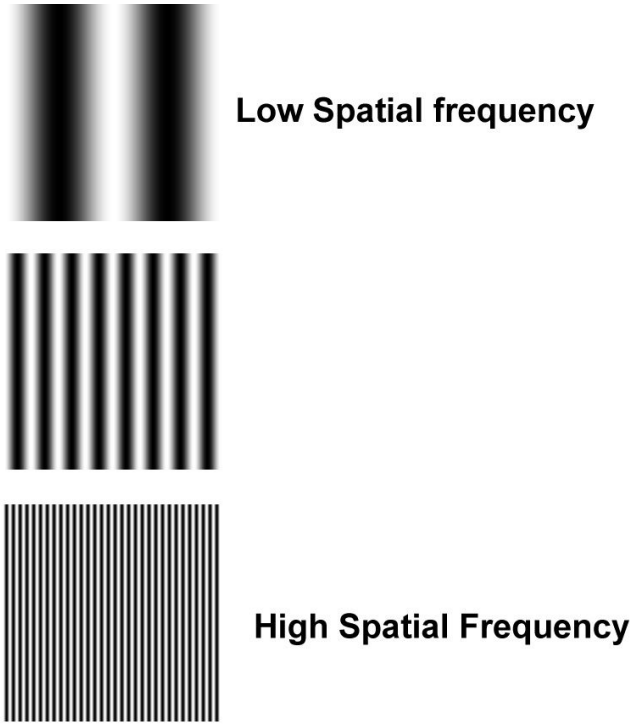


This is *not* random

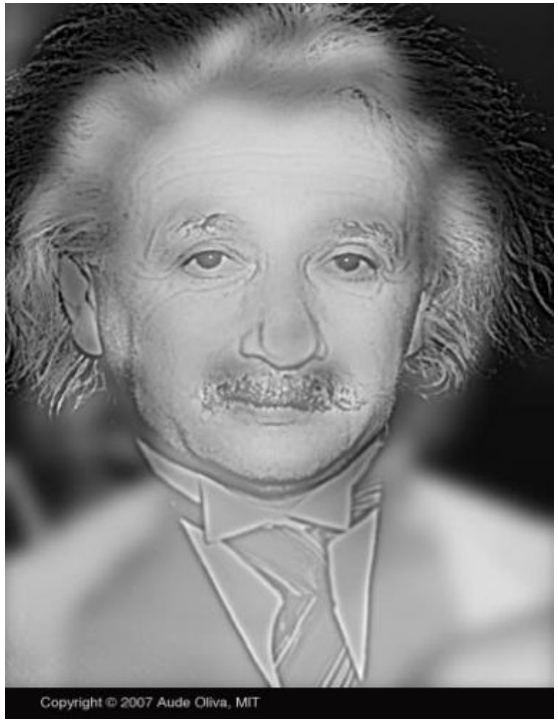
$1/f^2$ Distribution



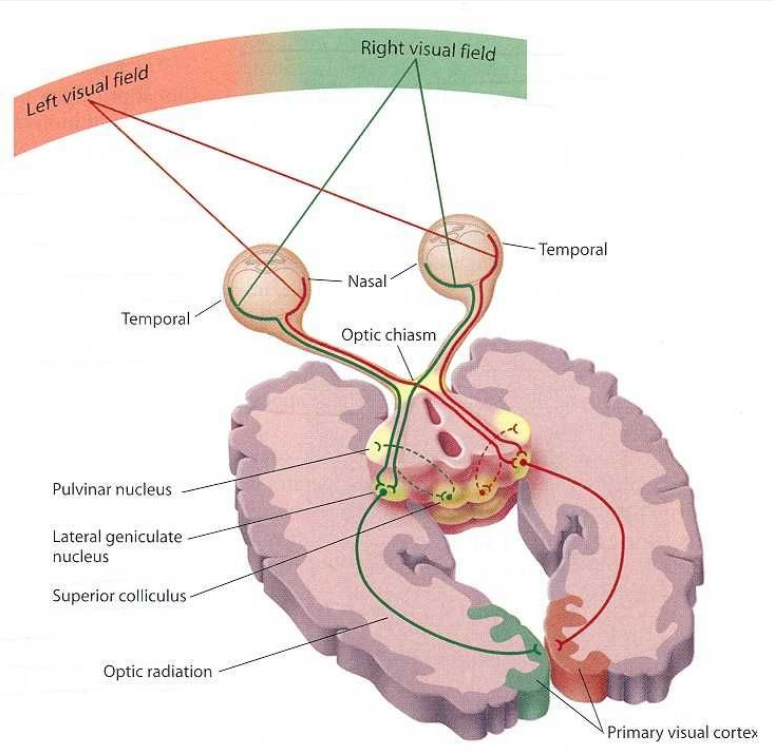
Spatial Frequency



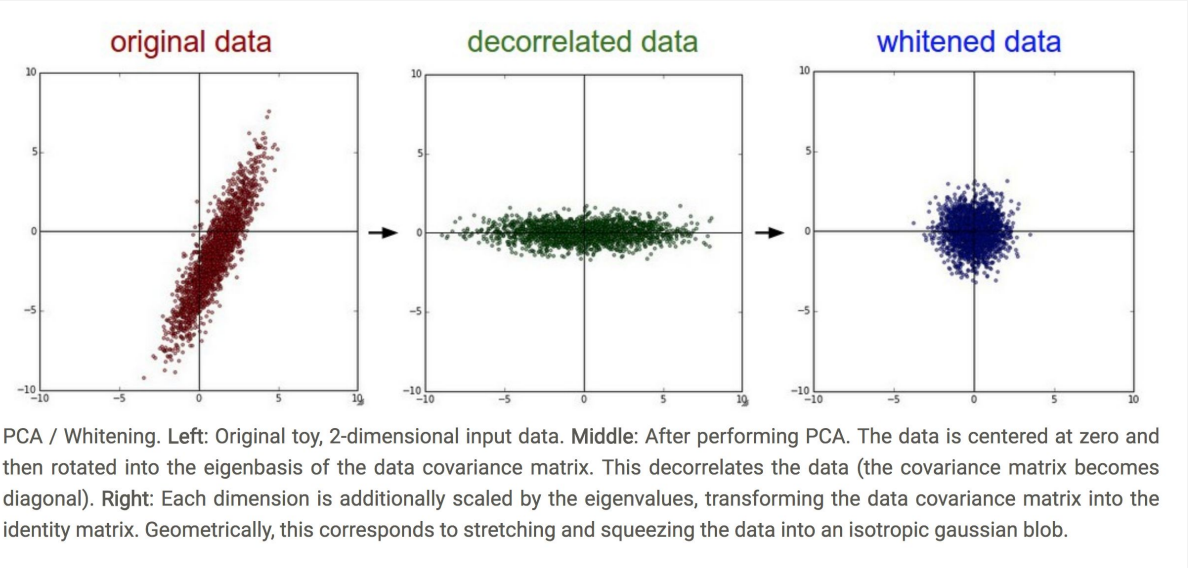
Spatial Frequency



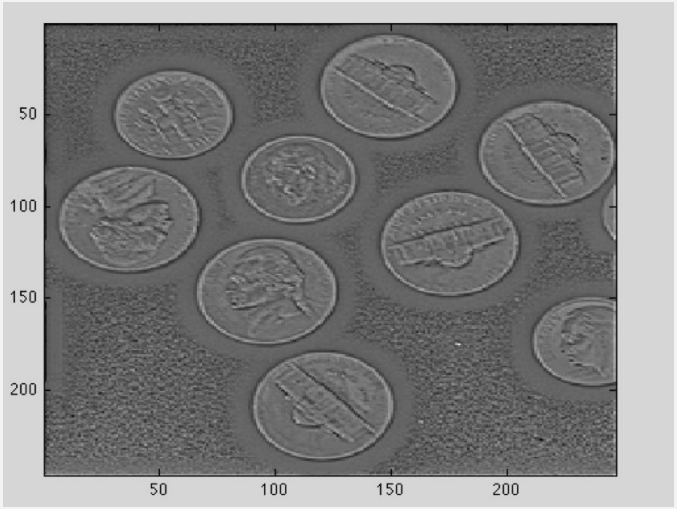
The Visual System

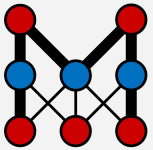


Decorrelation and Whitening

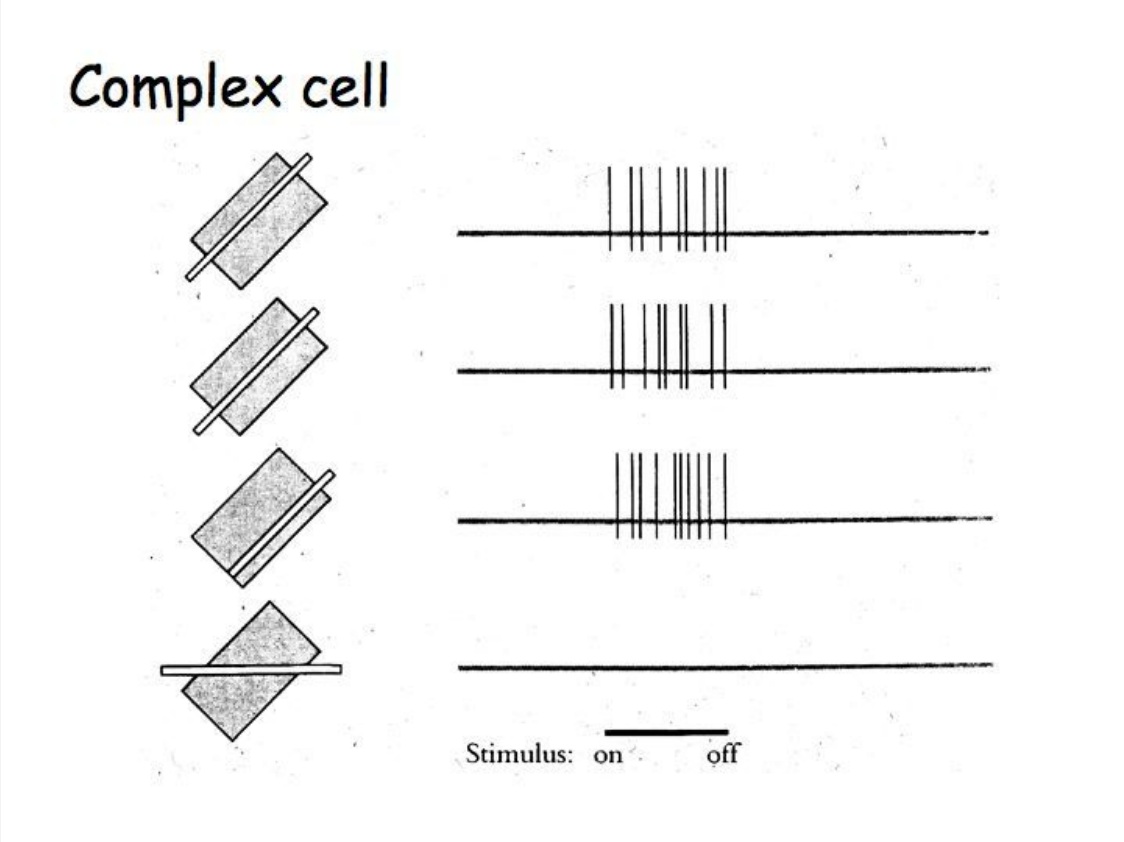


PCA / Whitening. **Left**: Original toy, 2-dimensional input data. **Middle**: After performing PCA. The data is centered at zero and then rotated into the eigenbasis of the data covariance matrix. This decorrelates the data (the covariance matrix becomes diagonal). **Right**: Each dimension is additionally scaled by the eigenvalues, transforming the data covariance matrix into the identity matrix. Geometrically, this corresponds to stretching and squeezing the data into an isotropic gaussian blob.





Location Invariance in cortical neurons



Visual Cortex

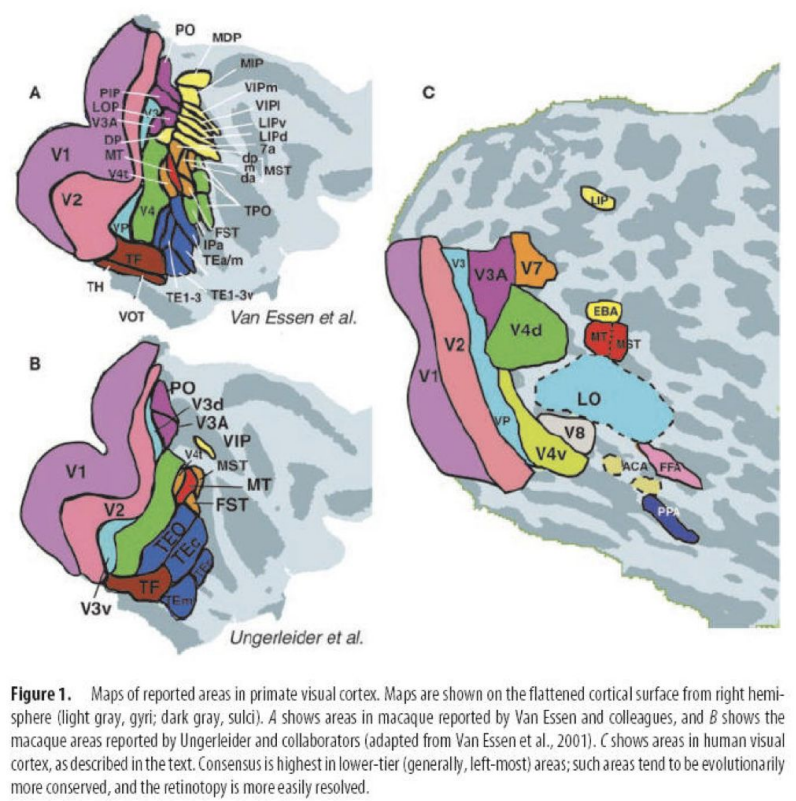
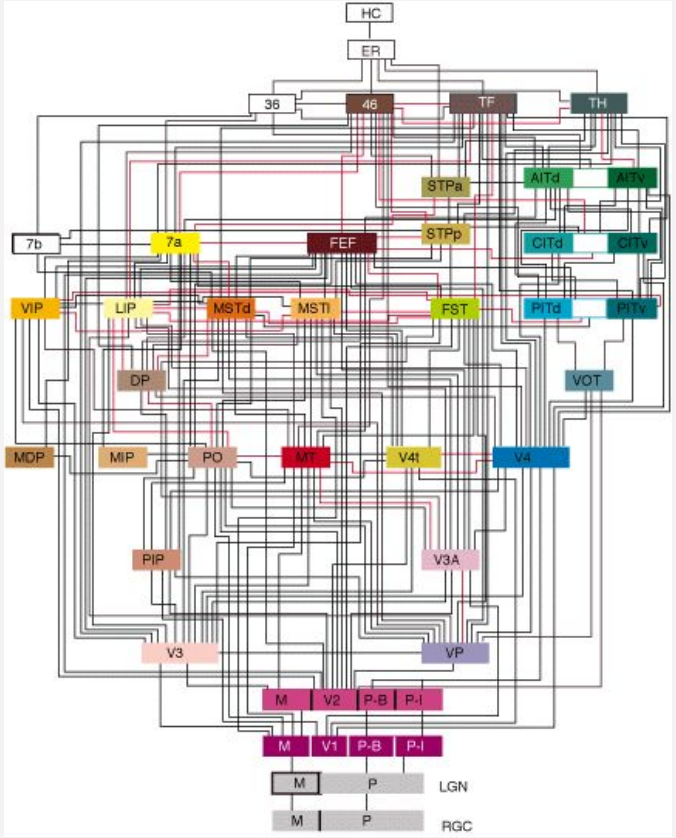
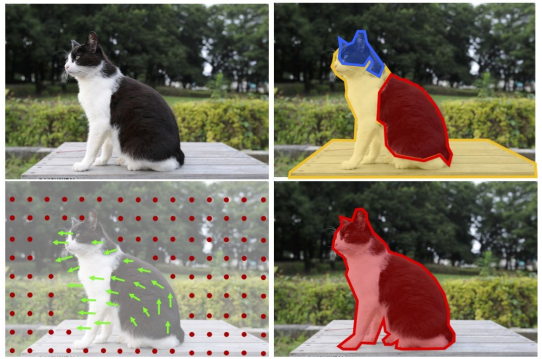
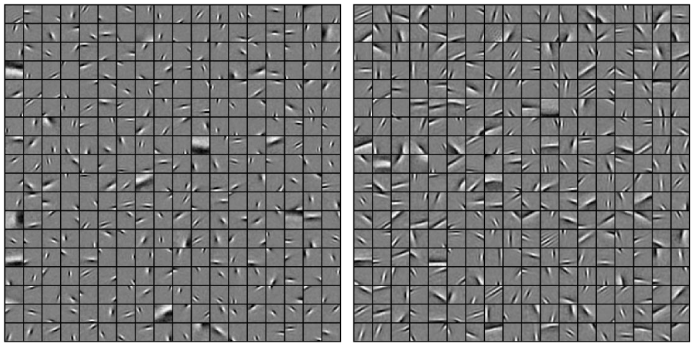


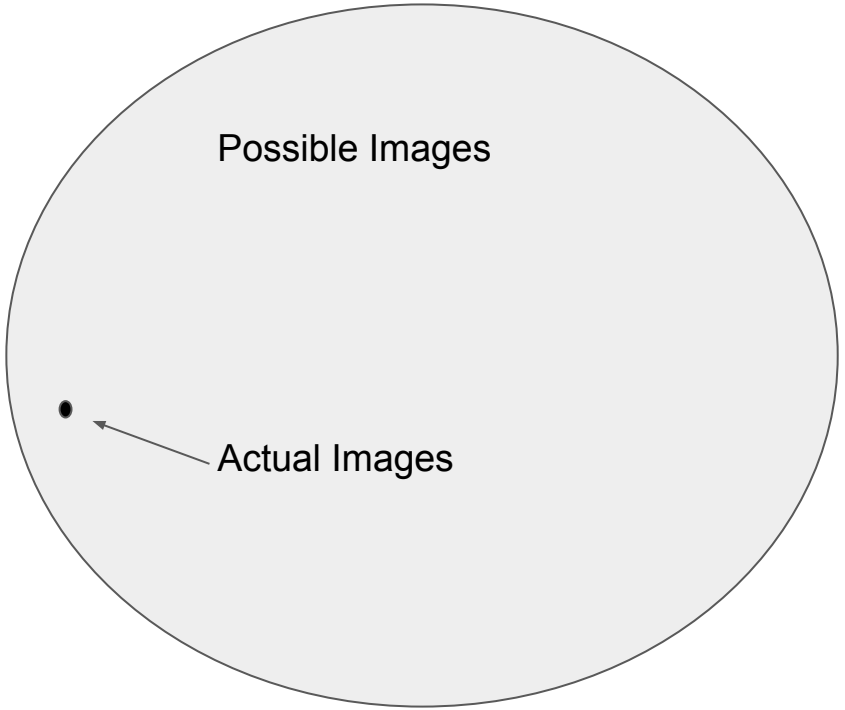
Figure 1. Maps of reported areas in primate visual cortex. Maps are shown on the flattened cortical surface from right hemisphere (light gray, gyri; dark gray, sulci). *A* shows areas in macaque reported by Van Essen and colleagues, and *B* shows the macaque areas reported by Ungerleider and collaborators (adapted from Van Essen et al., 2001). *C* shows areas in human visual cortex, as described in the text. Consensus is highest in lower-tier (generally, left-most) areas; such areas tend to be evolutionarily more conserved, and the retinotopy is more easily resolved.



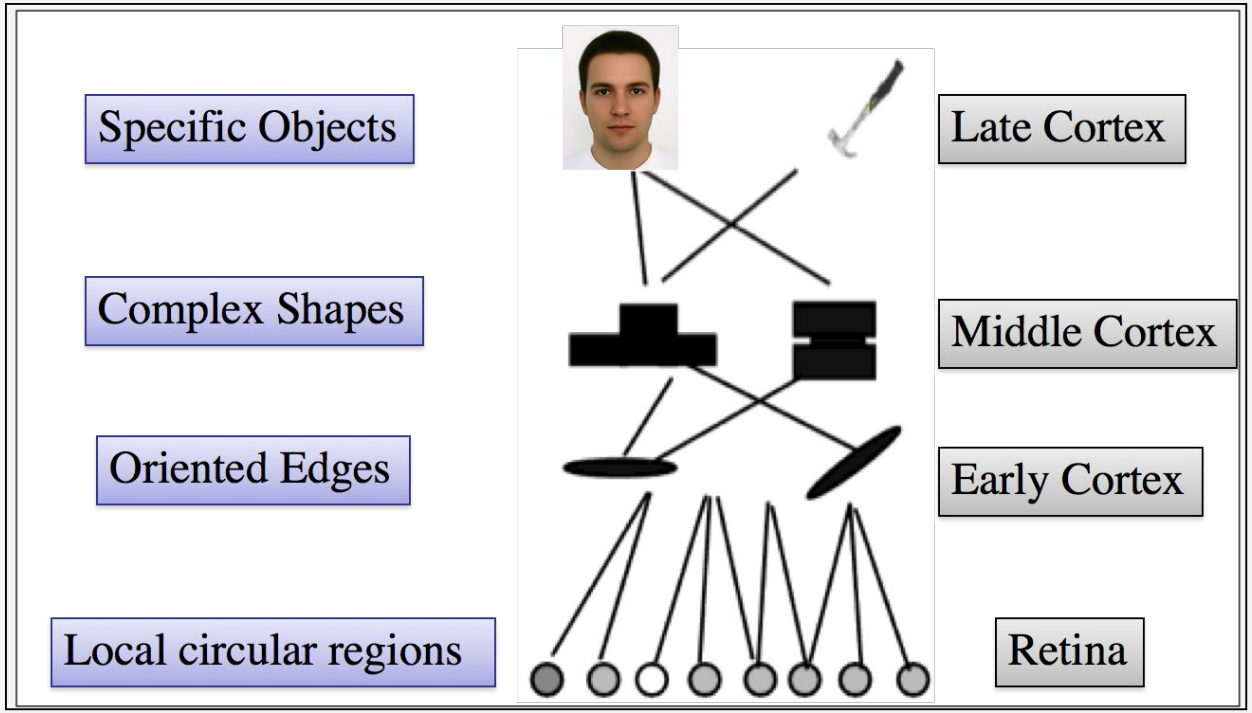
Learning Features to reduce dimensionality



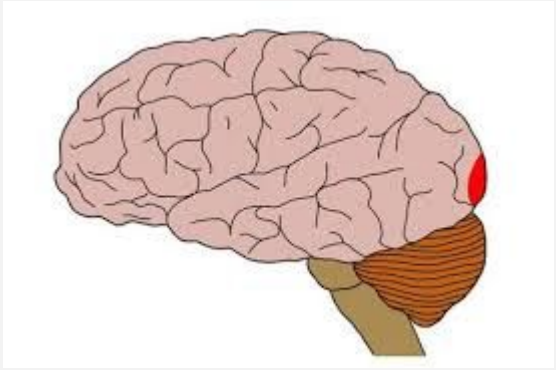




Visual Hierarchy

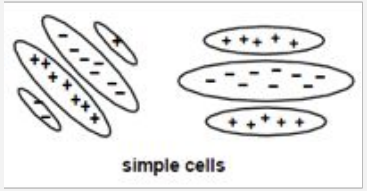


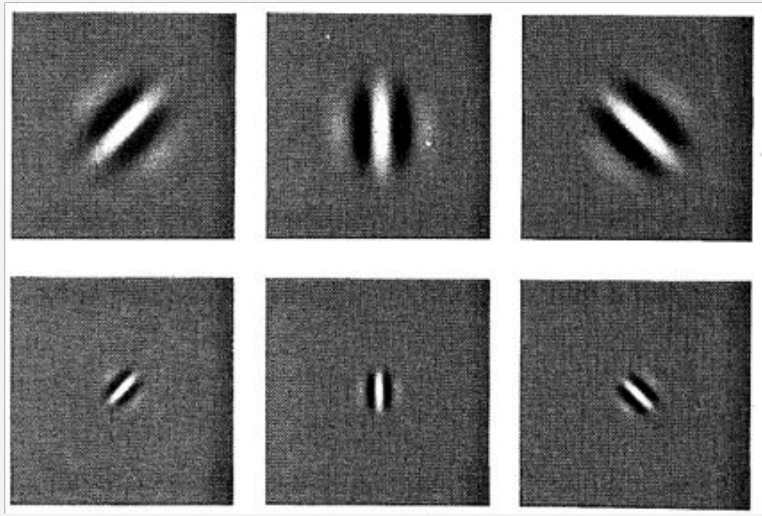
Primary Visual Cortex



Feature detectors in V1 have the properties of:

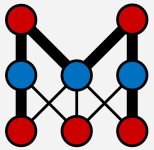
- Localized
- Oriented
- Bandpass



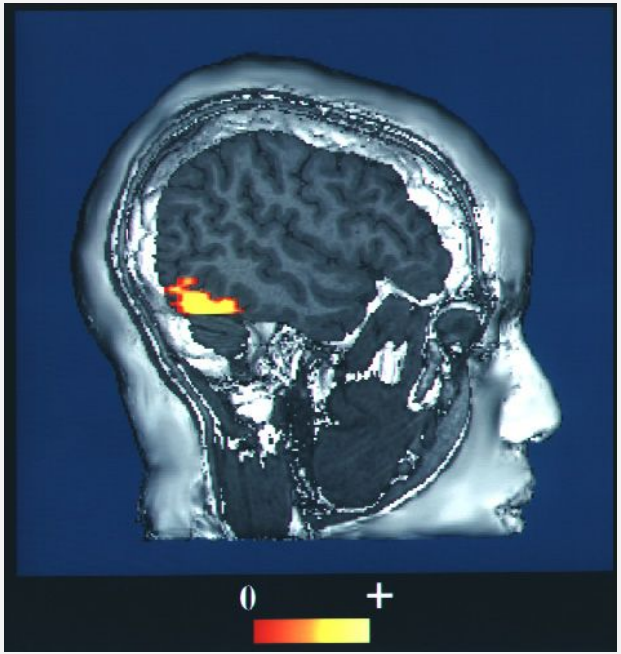


Gabors

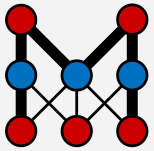
Sensitive to spatially and frequency-localized features



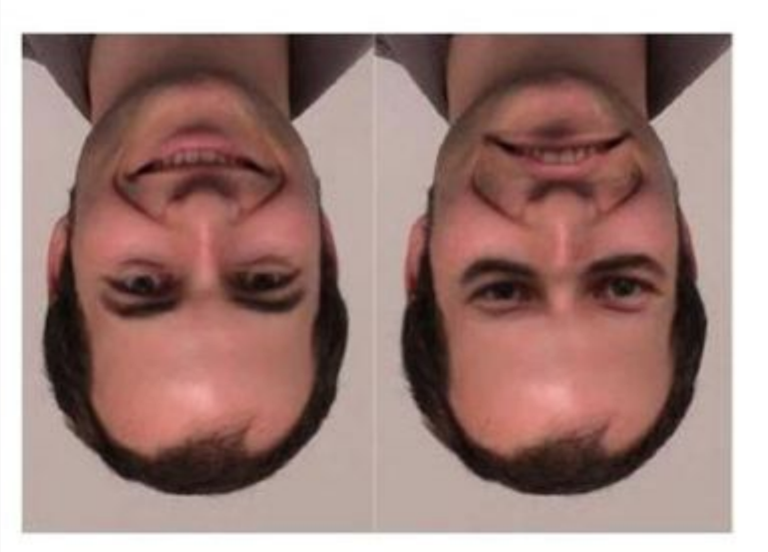
High-level features



Fusiform Face Area

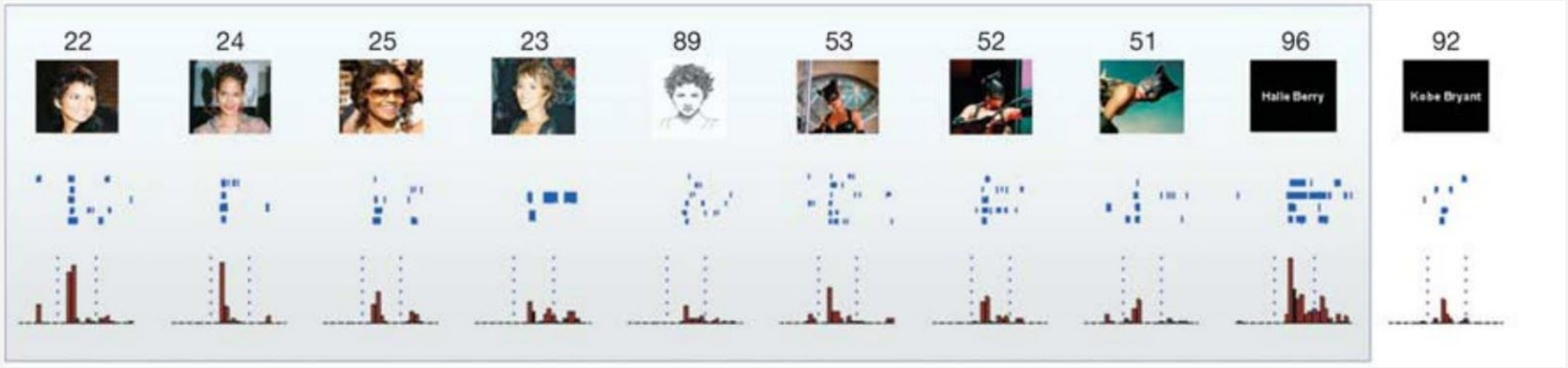


Neural Specialization



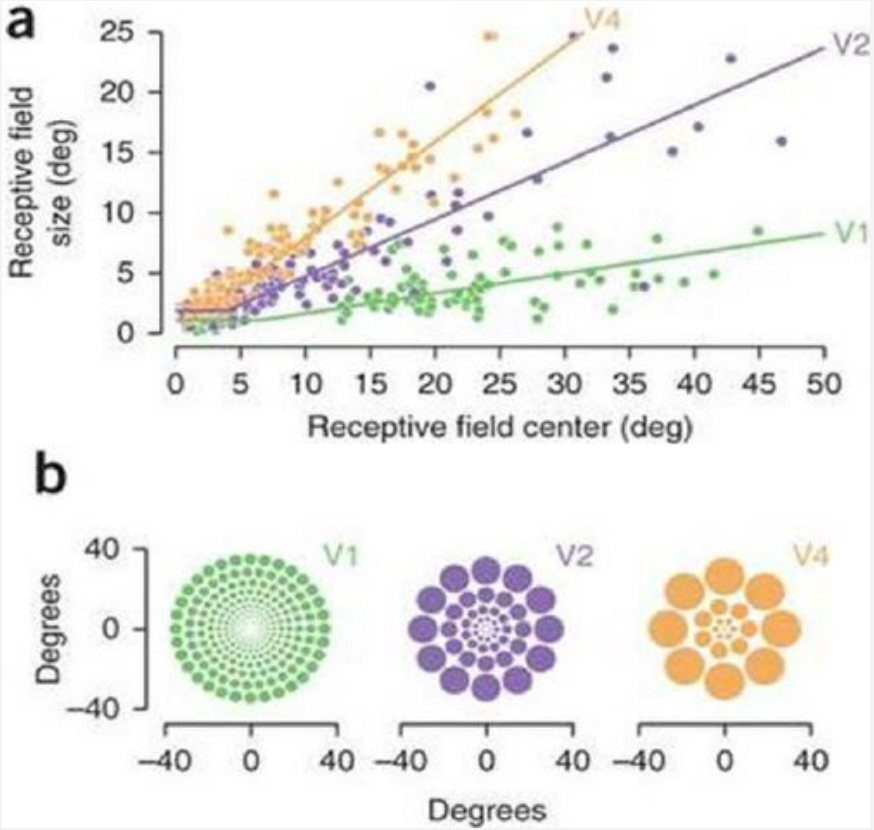
Neural Specialization

Neural Specialization

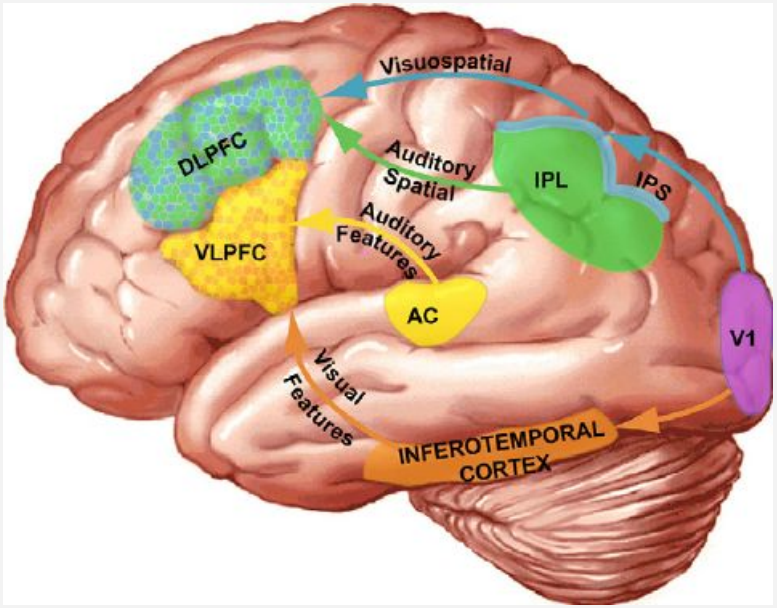


Quiroga et al (2005)

Receptive field size increases leading to spatial invariance



Multisensory Cortex



Multisensory Cortex

Convolutional Neural Networks

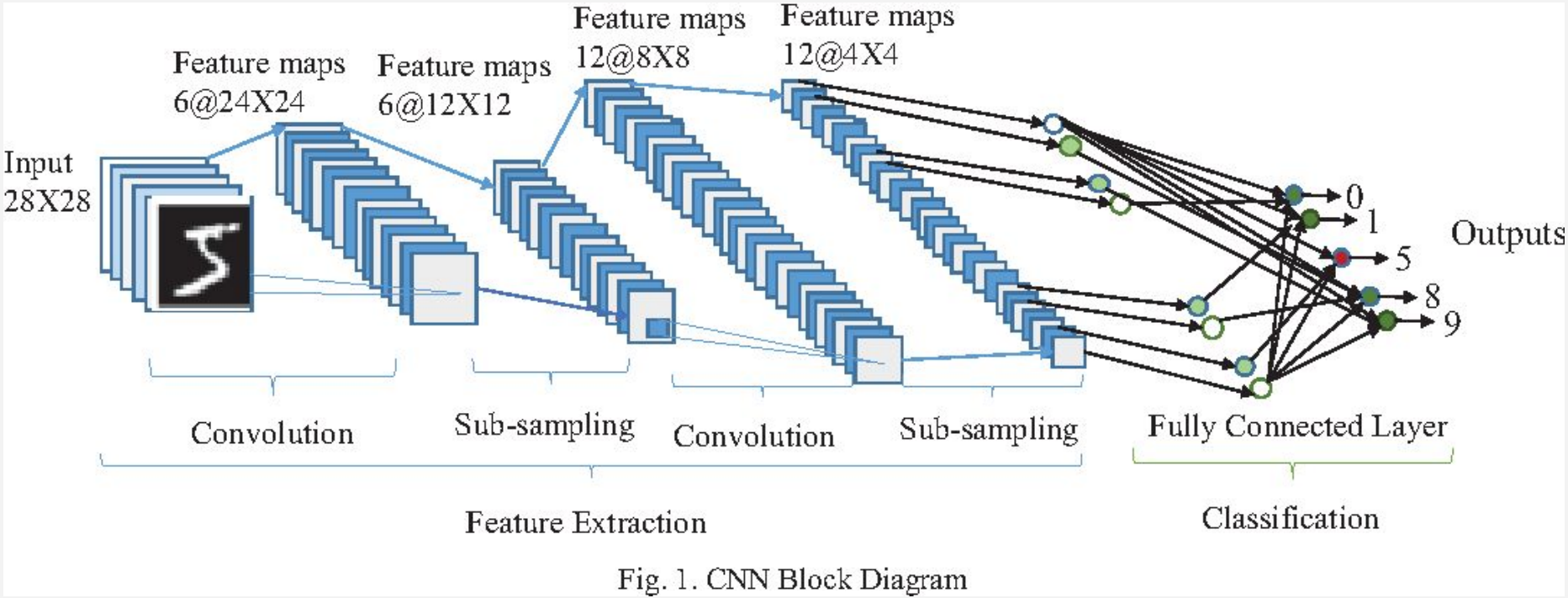


Fig. 1. CNN Block Diagram