

Big Data Technology

Paul Rad, Ph.D.

Associate Professor
Information Systems and Cyber Security, College of Business School
Electrical and Computer Engineering, College of Engineering

Neural Network in Your Browser

<https://playground.tensorflow.org>

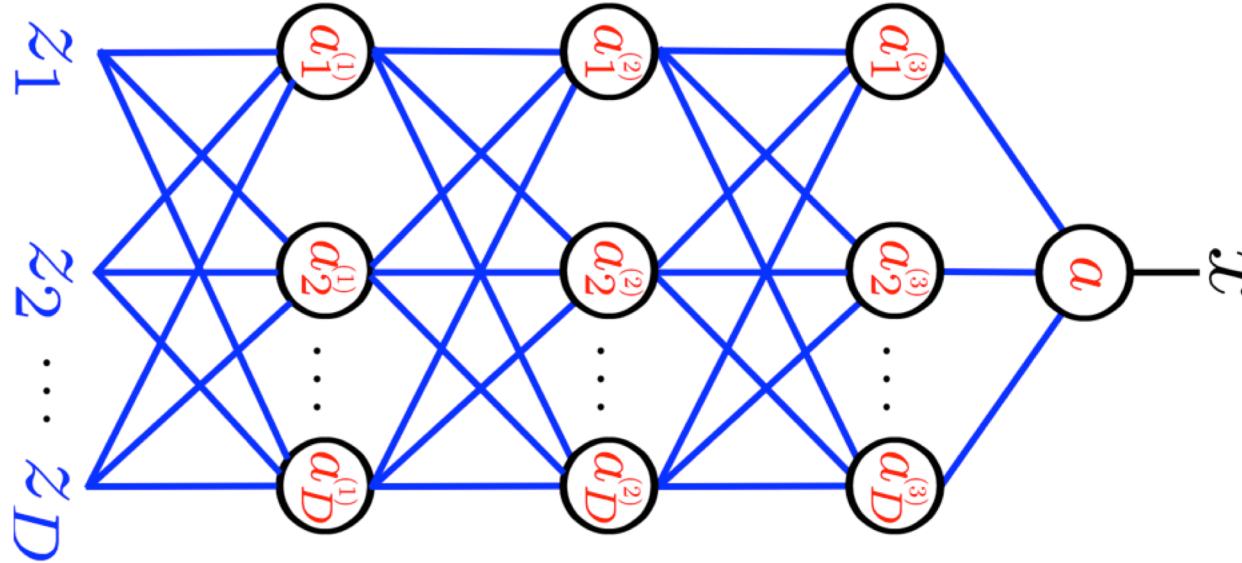
Outline

- **Logistics**
- **Summary and Review Previous Sections**
- **TensorFlow Playground**
- **Convolution Neural Network**

Stochastic Gradient Descent (SGD)

Loop:

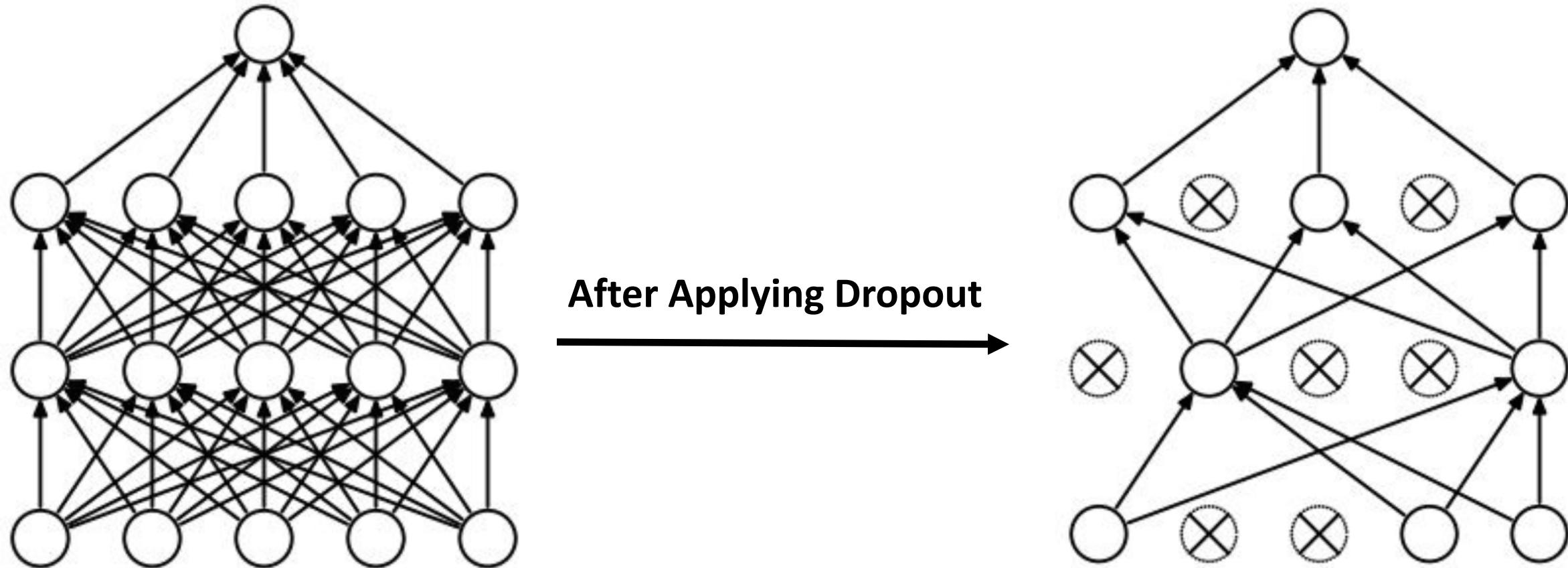
1. Sample a batch of data with their labels
2. Forward Propagate it through the graph, calculate the error
3. Backpropagate to calculate the gradients
4. Update the parameters (weights) using the gradient



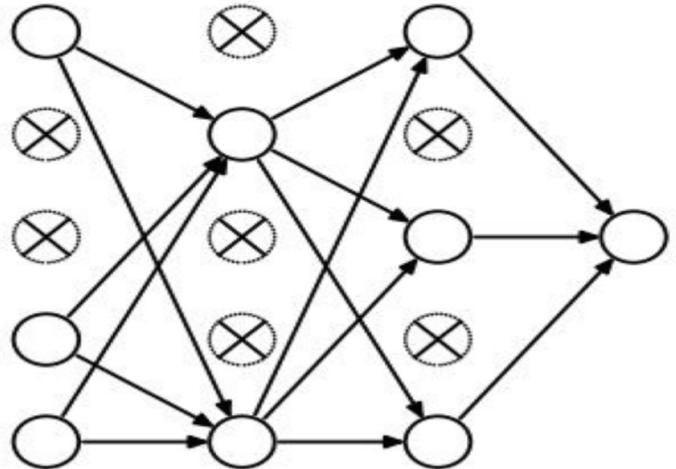
$$W_k = W_{k-1} - \epsilon \frac{\partial E(W)}{\partial W}$$

Dropout

Randomly set some neurons to zero in the forward pass



Dropout



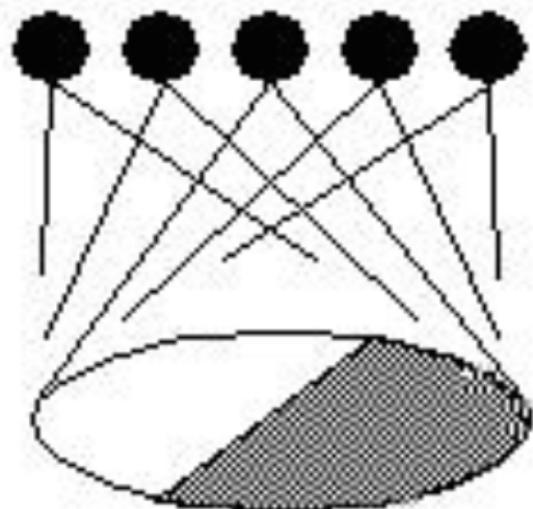
**Forces the network to have
a redundant representation**



Vision: Hierarchical Organization (Year: 1962)

Hubel & Weisel

topographical mapping

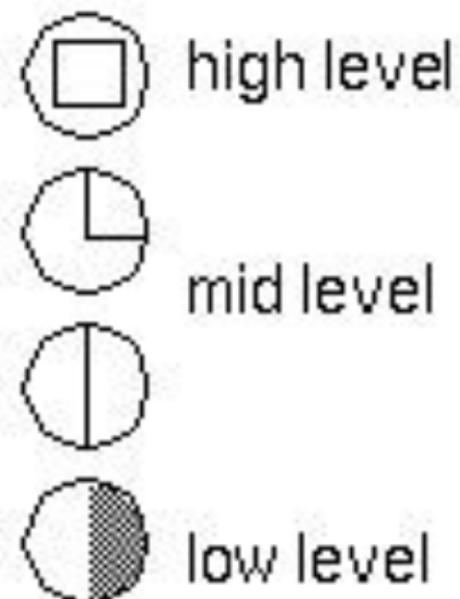
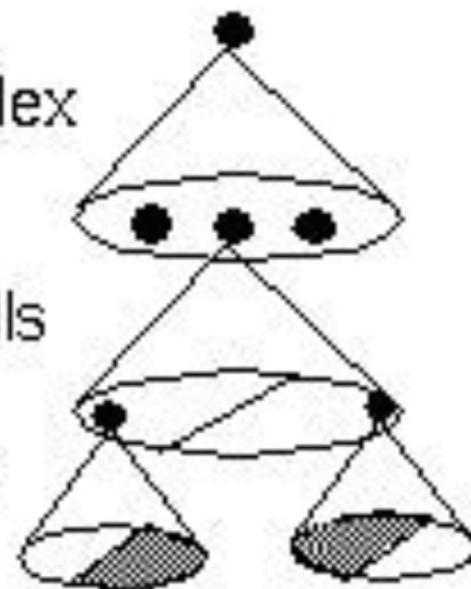


featural hierarchy

hyper-complex
cells

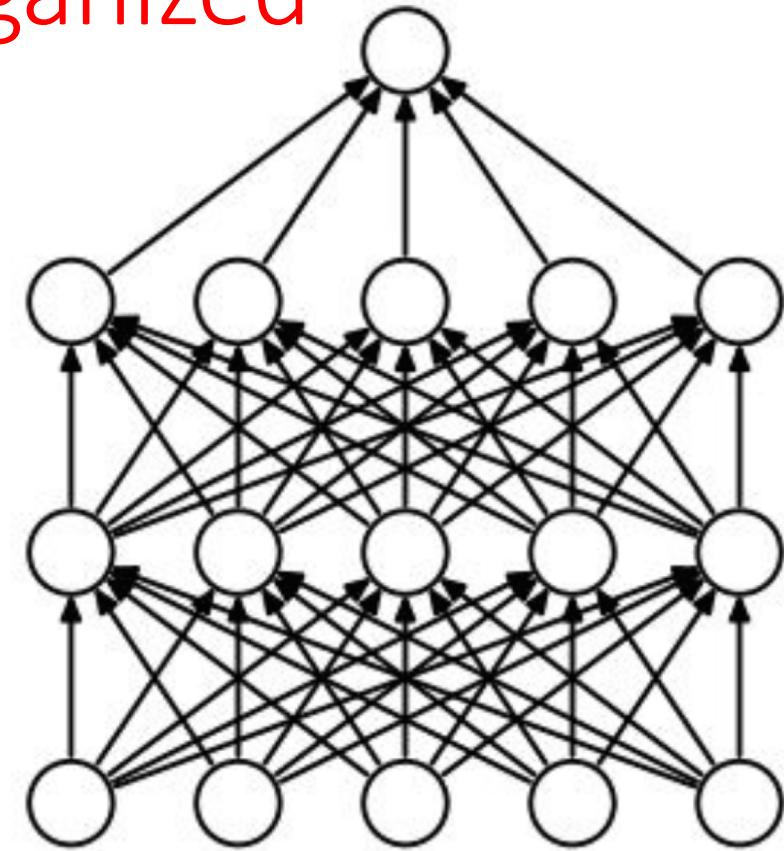
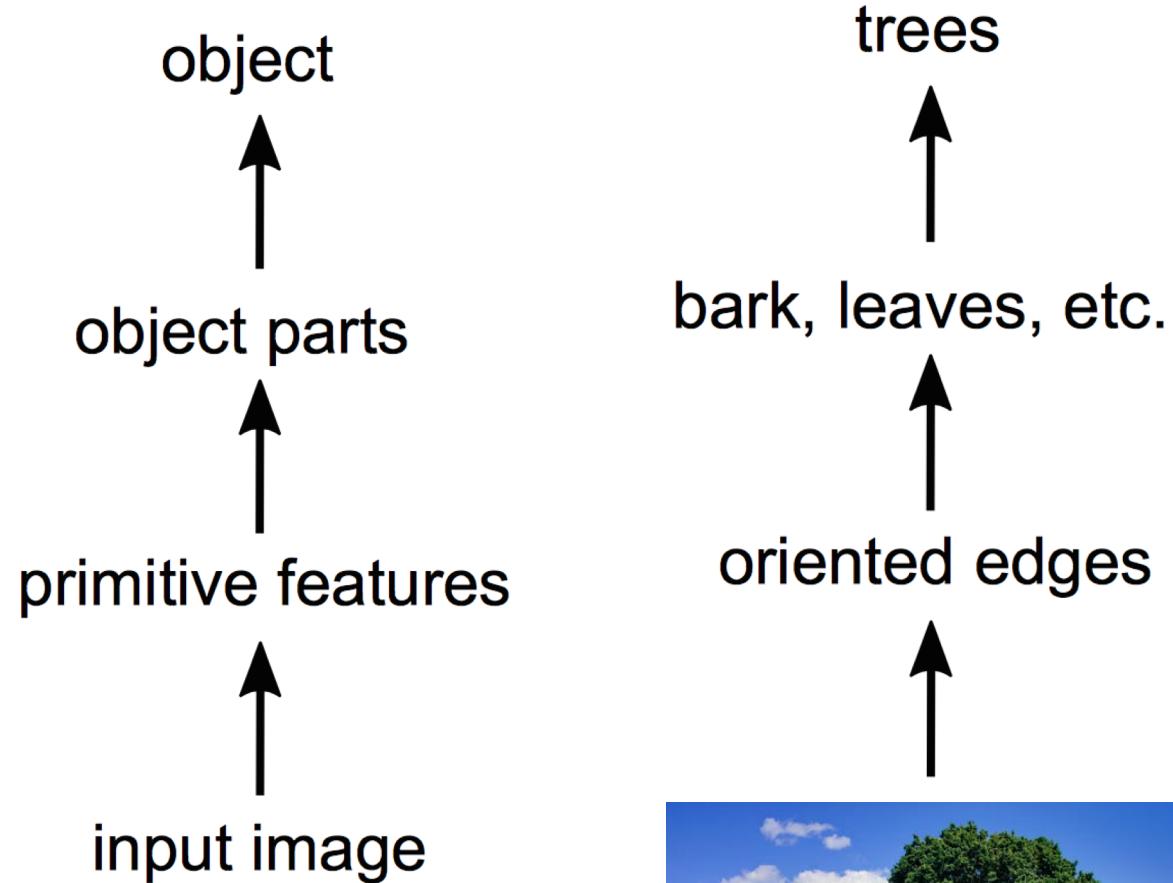
complex cells

simple cells



Why use hierarchical multi-layered models?

Biological vision is hierarchically organized



What's wrong with standard neural networks?

Hard to Train

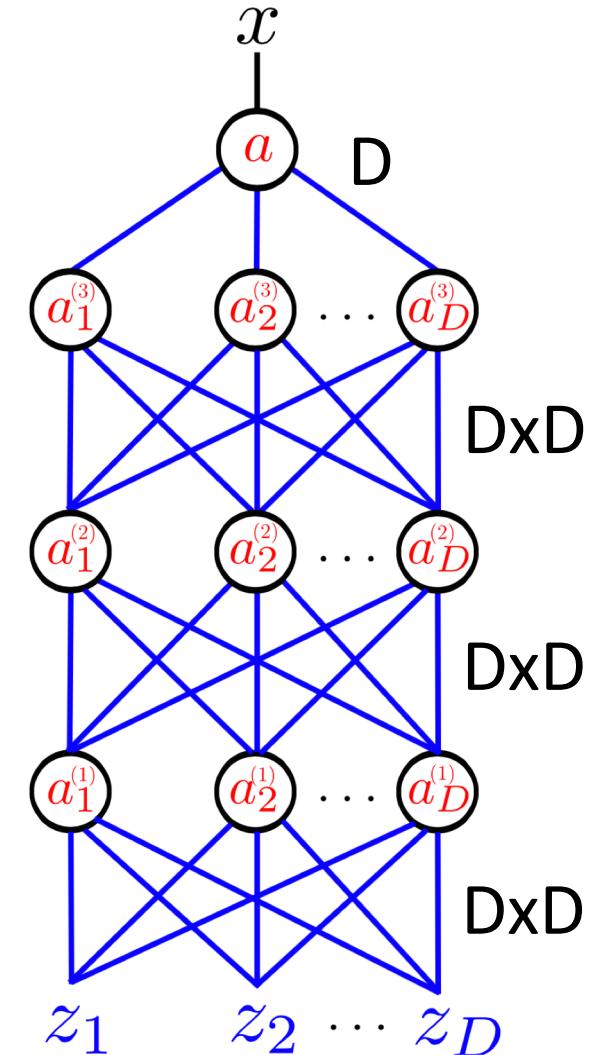
How many parameters does this network have?

Number of Parameters = $3 \times (D \times D) + D$

For a small $D = 32 \times 32 = 1024$ MNIST image:

Number of Parameters = $3 \times (1024 \times 1024) + 1024$

$\sim 3 \times 10^6$



Architecture of LeNet-5, Convolution Neural Network

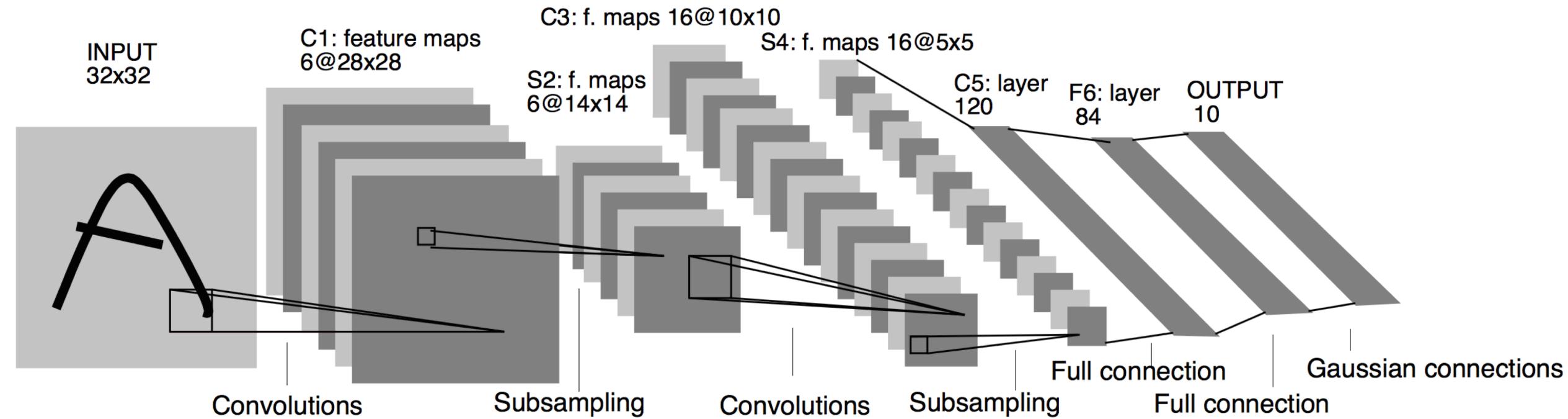
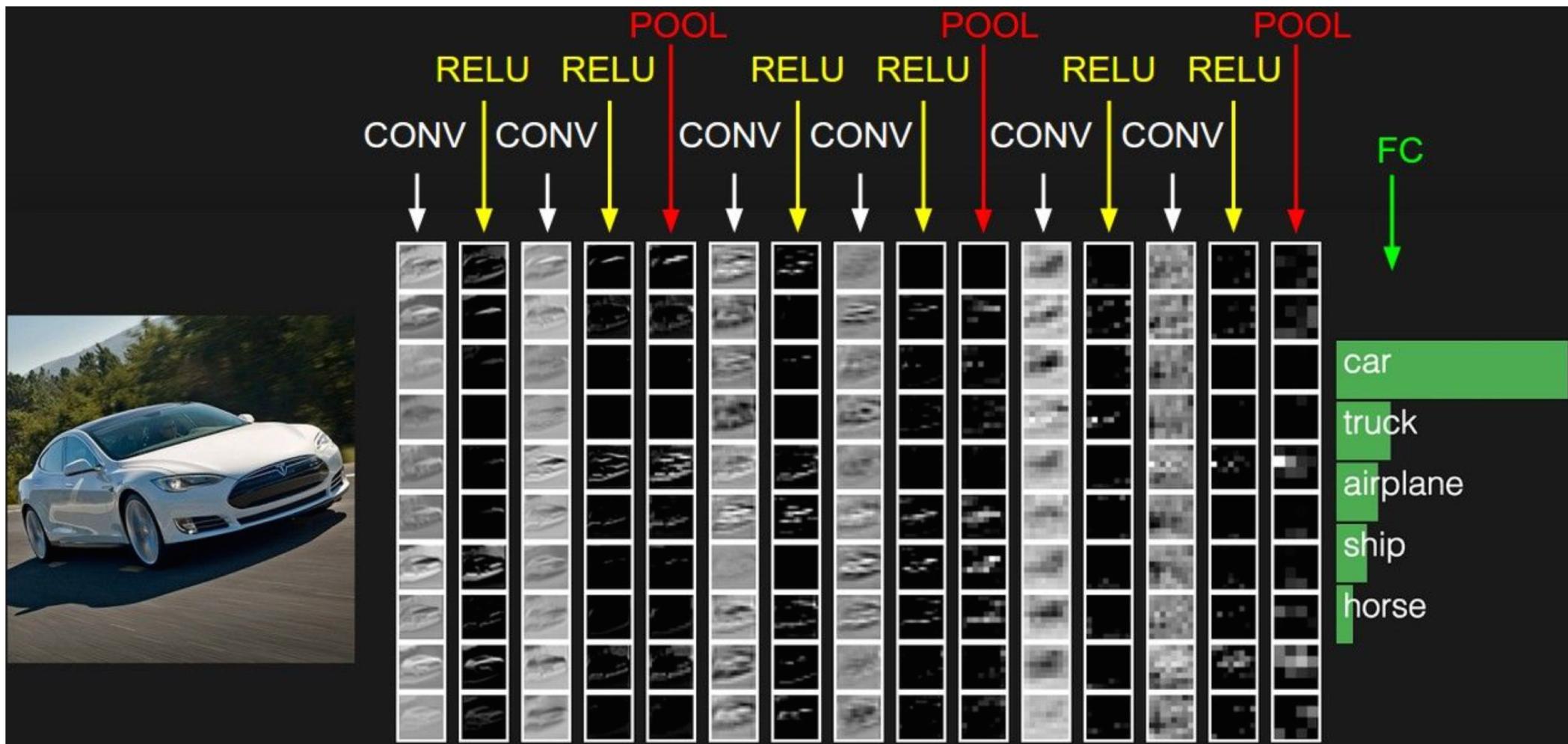


Fig. 2. Architecture of LeNet-5, a Convolutional Neural Network, here for digits recognition. Each plane is a feature map, i.e. a set of units whose weights are constrained to be identical.

Proc. Of the IEEE, November 1998, "Gradient-Based Learning Applied to Document Recognition"



Review: What is convolution?

- Convolution is an important operation from signal processing
- A convolution is an integral that expresses the amount of overlap of one function as it is shifted over another function .

$$f * g = \int_{-\infty}^{\infty} f(\tau) g(t - \tau) d\tau = \int_{-\infty}^{\infty} g(\tau) f(t - \tau) d\tau$$

- 2 Dimensional Discrete Function (Image)

$$f[x, y] * g[x, y] = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} f[n_1, n_2] \cdot g[x - n_1, y - n_2]$$