CSE578: Computer Vision

Spring 2019:

Feature Learning with Deep Learning

Anoop M. Namboodiri

Center for Visual Information Technology

IIIT Hyderabad, INDIA

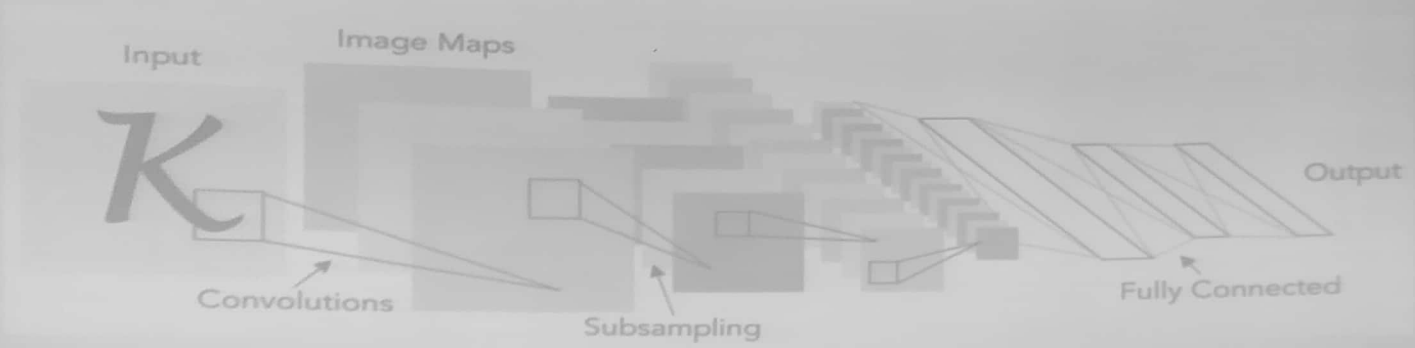
[Content Generously Borrowed from CS231n]

Convolutional Neural Networks

- Course on CNN in Computer Vision at Stanford
  - Fei-Fei Li, Justin Johnson and Serena Yeung
  - 2017 edition on YouTube: https://www.youtube.com/playlist?list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv

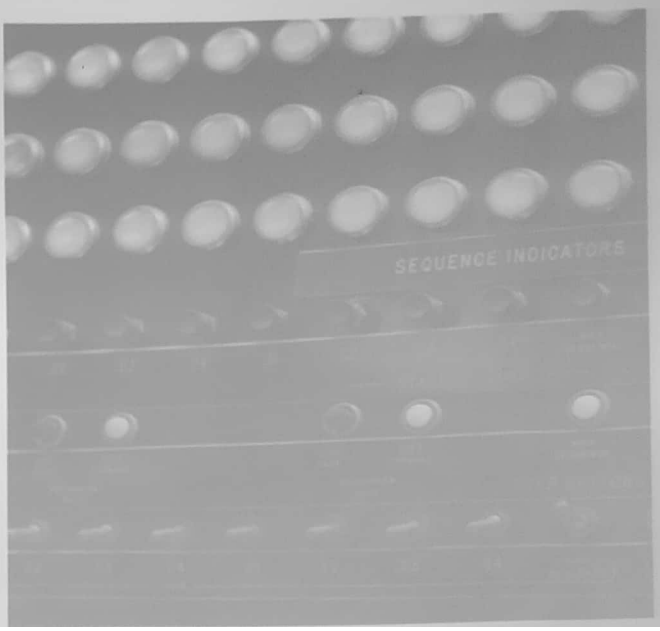
Convolutional Neural Networks: History

- LeNet: Digit / Character Recognition, LeCun, Bottou, Bengio, Haffner 1998.



Convolutional Neural Networks: History

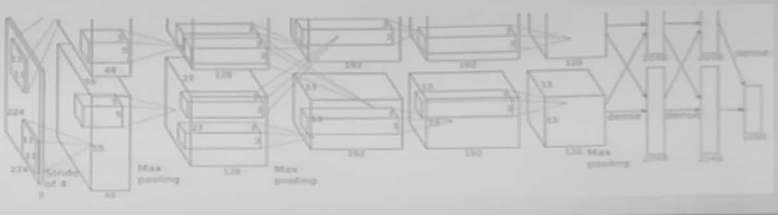
- The Mark 1 Perceptron machine (Frank Rosenblatt, ~1957)
- Connected to a camera that used 20×20 cadmium sulfide photocells to produce a 400-pixel image.
- Recognized letters of the alphabet
- Used gradient descent update rule for learning



Convolutional Neural Networks: History

Several other efforts

- Adaline/Madaline: Widrow and Hoff, 1960
- Backpropagation: Rumelhart et al. 1986
- RBMs: Pretraining: Hinton and Salakhutdinov 2006
- The watershed moment: "Imagenet Classification with Deep Convolutional Neural Networks": Alex Krizhevsky, Ilya Sutskever, Geoffrey E Hinton, 2012



ConvNets tops most vision tasks

Classification

Retrieval

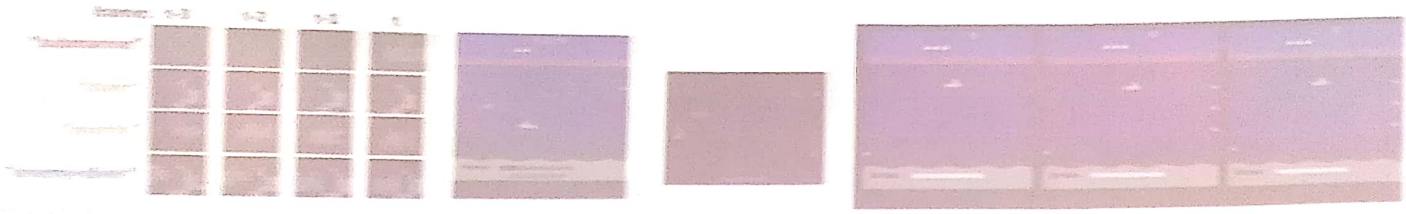
ConvNets tops most vision tasks

Self Driving

Street Sign Recognition

ConvNets tops most vision tasks

- Human Pose Estimation; Video game play



ConvNets tops most vision tasks

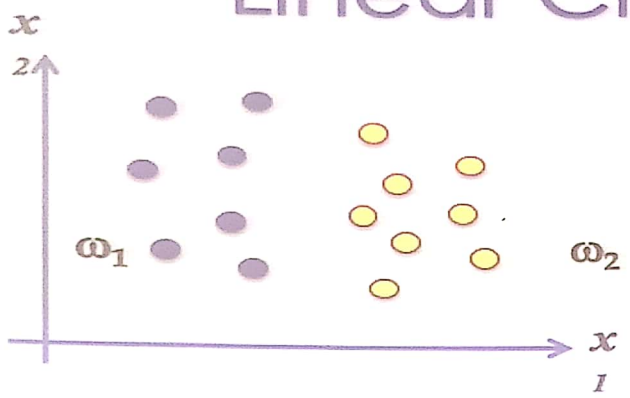
- Medical Diagnosis
- Street Detection

Whale Recognition (Kaggle)

ConvNets tops most vision tasks

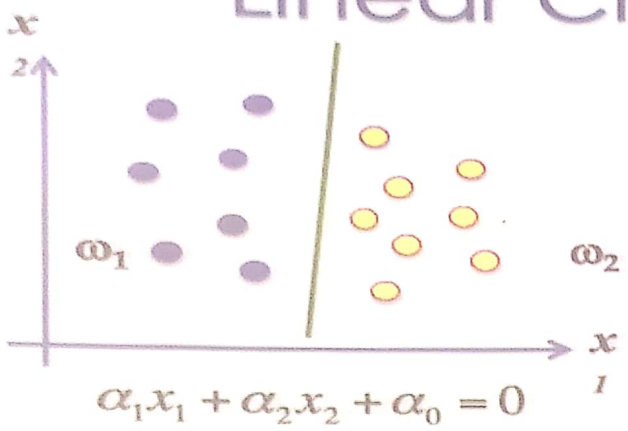
- Person Recognition
- Spoof Detection
- Video Activity Recognition
- Image Captioning
- Image Generation
- Style Transfer
- Image Super-resolution
- Image Coloring
- Lip Reading
- Visual QA
- Video Captioning
- Video Highlight Detection
- Single/few Image 3D Reconstruction
- And many others
  - Just see Kaggle

Linear Classifier



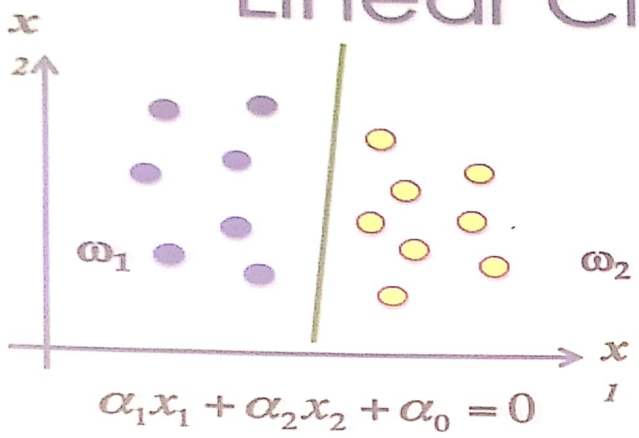
- A linear boundary separates two classes.

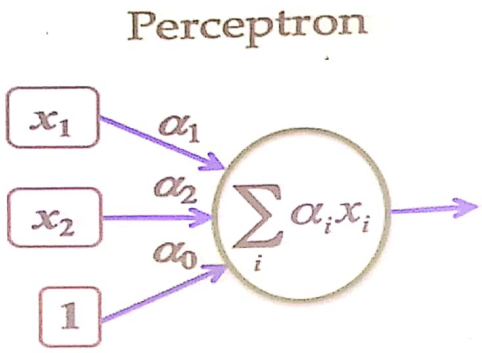
Linear Classifier



- A linear boundary separates two classes.

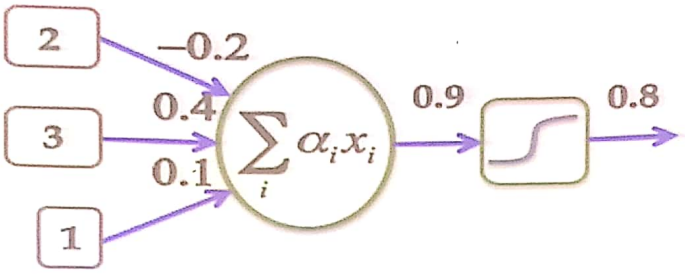
Linear Classifier





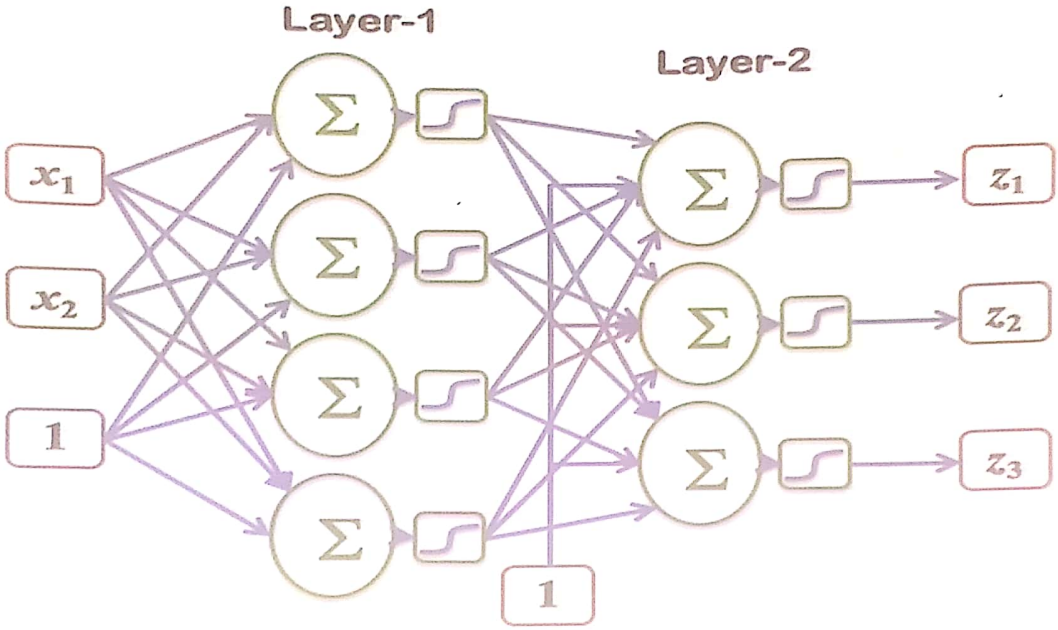
- A linear boundary separates two classes.

Perceptron Learning



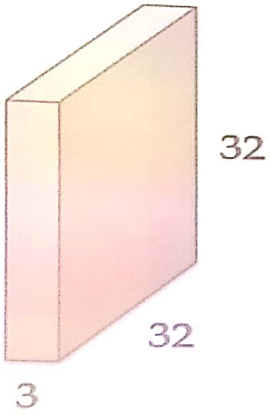
- Randomly Initialize the weights
- For each sample:
  - Feed a sample and find the output (forward pass)

Multi-layer Perceptron

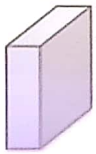


Convolution Layer

- 32 x 32 x 3 image

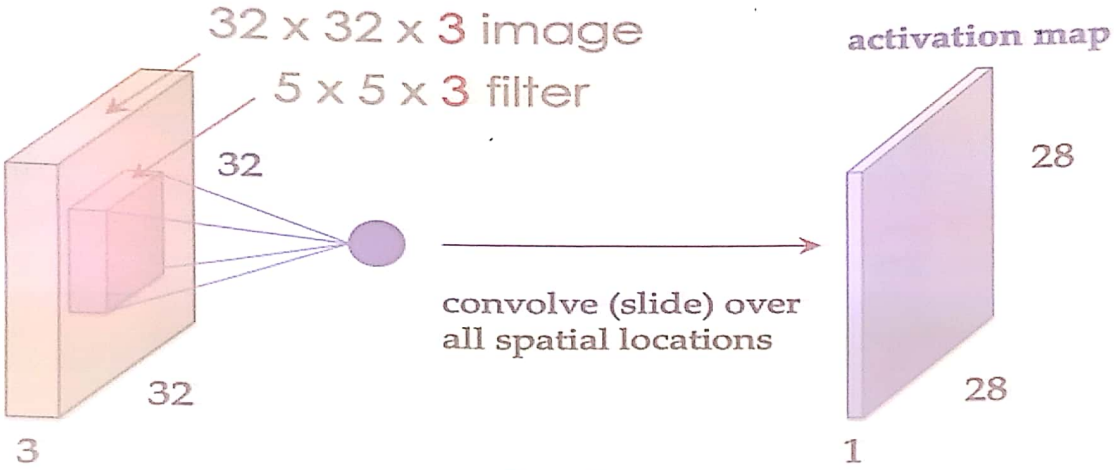


5 x 5 x 3
filter



Convolve the filter with the image. i.e. "Slide over the image spatially, computing dot products"

Convolution Layer



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