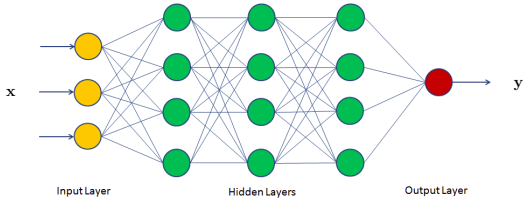
Biological Vision and Applications
Module 05-07: Introduction to neural networks

Hiranmay Ghosh

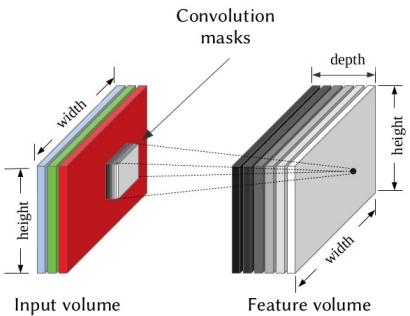
Neural Networks



- Feed-forward network – back-propagation algorithm for training
- Transfer function: $y = W.x$
- W is a constant: deterministic output
  - no learning from "experience" in deployment stage
- For a 640×480 color image
  - Number of input nodes = 927,360
  - Large number of parameters to be learned

Convolutional Neural Networks

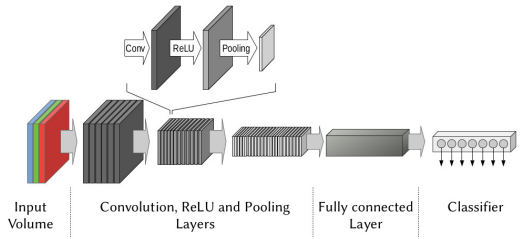
Why convolution?



- 2D organization exploits
  - Spatial context of a location in 2D
  - Identical operations repeated over the different spatial regions
- Drastic reduction in model parameters
  - For a 3×3 convolution filter, only 27 parameters to learn
  - Independent of image size

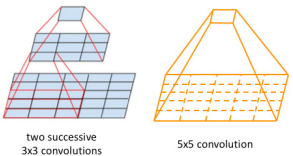
Convolutional Neural Networks

Structure

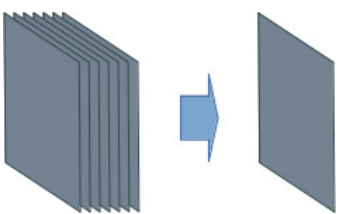


- Same operation to be repeated over different receptive fields
  - Do not need so many parameters
- Architecture motivated by early vision
  - Convolution: Aggregates information from receptive field
  - Filtering (ReLU): Non-linear transformation
  - Pooling (max): Reduces information volume

On filter sizes

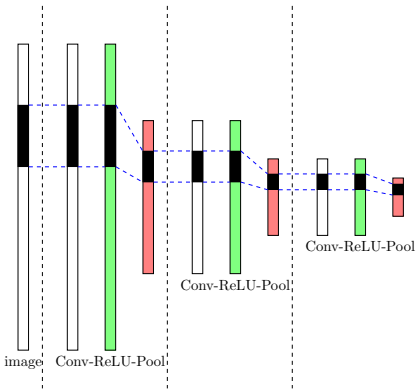


- A bank of two 3×3 filters in succession has a receptive field of 5×5
  - Can implement identical transfer function
- Which one would you prefer?



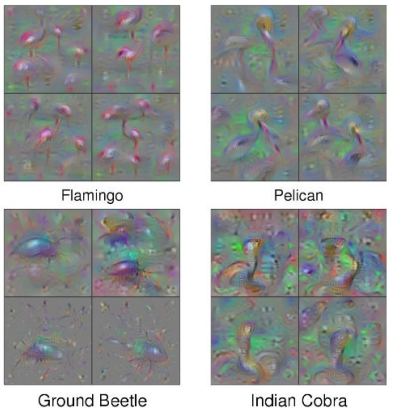
- Filter size $= 1$
- For "flattening" the layers
  - $y(i,j) = \sum_k w_k.x_k(i,j)$

Progressive abstraction



- Each location at any layer of a CNN holds information about some locality of the image
- A location in a deeper layer covers more visual field of the image than a shallower layer
  - A deeper layer incorporates more context than a shallower layer
- Visual information is progressively abstracted
- Depth of layer increases with the depth of the network

Does CNN really do progressive abstraction ?



Yosinski, et al. Understanding neural networks through deep visualization (2015)

Some notable CNN implementations (2012 – 2015)

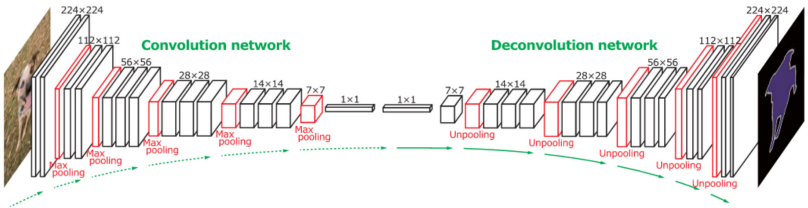
These implementations are reused in different contexts

- AlexNet
- VGG
- ResNet
- GoogleNet

Architecture comparisons

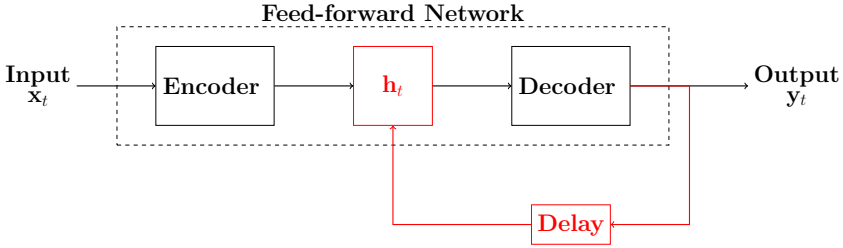
Fully Convolutional Network (FCNN)

Used for Image Segmentation



Deconvnet

Recurrent Neural Network (RNN)



- RNN incorporates a feedback loop (with delay)
- Transfer function
  - $h_t = f(W_1.x_t + W_2.y_{t-1})$
  - $y_t = g(W_3.h_t)$
  - h accumulates experience
- Tool for sequence processing tasks

Quiz

Quiz 05-07

End of Module 05-07