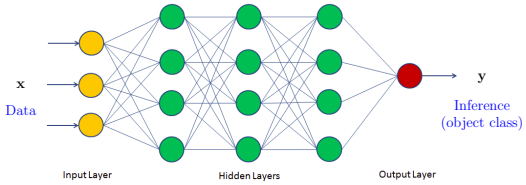
Biological Vision and Applications
Module 06-01: Introduction to Neural Networks

Hiranmay Ghosh

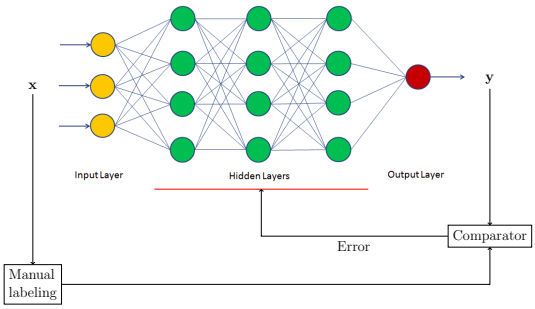
Neural Networks



- Transfer function: $\mathbf{y} = \mathbf{W}.\mathbf{x}$
- Feed-forward network: back-propagation algorithm for training
- $\mathbf{W}$ is a constant: deterministic output

Neural Networks

Training



- Back-propagation algorithm
  - Adjust network parameters based on error
  - Minimize error over many observations

Challenges for using neural network for image processing

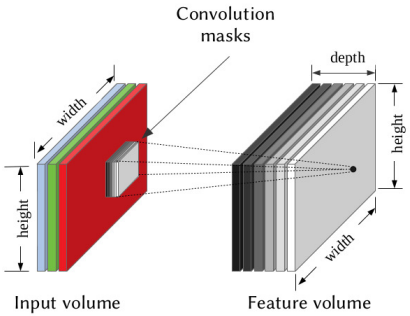
Training

- For a 640×480 color image
  - Number of input nodes $= 927{,}360$
  - Large number of parameters to be learned

- Early vision:
  - Image is organized in 2D
  - All image locations are to be similarly processed (contrast detection)

Convolutional Neural Networks (CNN)

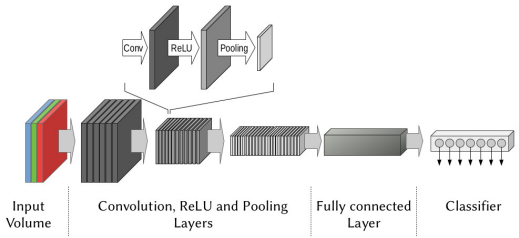
Exploits properties of early vision



- 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

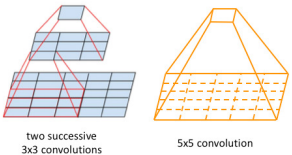
Convolutional Neural Networks

Structure

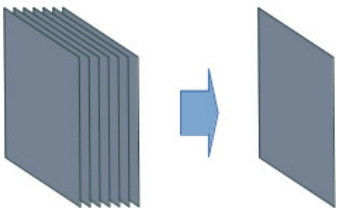


- Architecture motivated by early vision
  - Convolution: Aggregates information from receptive field
  - Filtering (ReLU): Non-linear transformation
  - Pooling (avg / max): Reduces information volume

On filter sizes



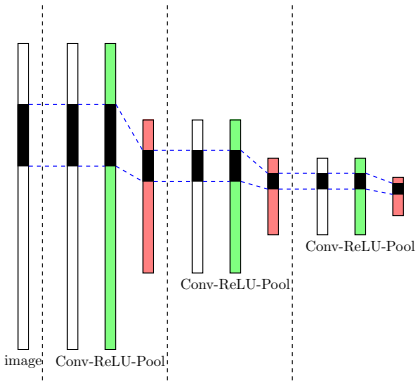
- A bank of two 3×3 filters in succession has a receptive field of 5×5
  - Can implement identical transfer function
- Which one would you prefer?



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

Progressive abstraction

Use of context



- 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

Some notable CNN implementations (2012 – 2015)

These implementations are reused in different contexts

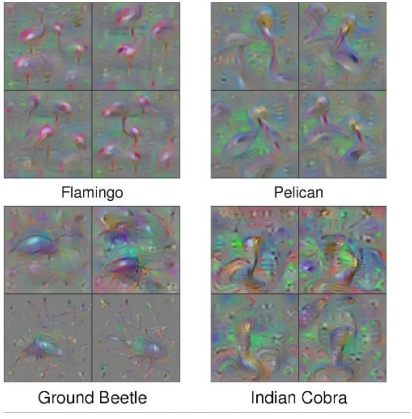
- AlexNet
- VGG
- ResNet
- GoogleNet

A feed-forward network does not learn from (runtime) experience

Architecture comparisons (blog)

Does CNN really do progressive abstraction ?

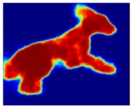
Visualization at the last layer – just before classification

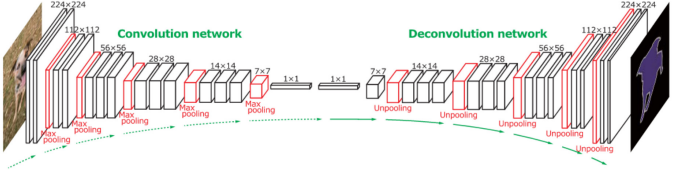


Yosinski, et al. Understanding neural networks through deep visualization

Fully Convolutional Neural Network (FCNN)

Used for Image Segmentation



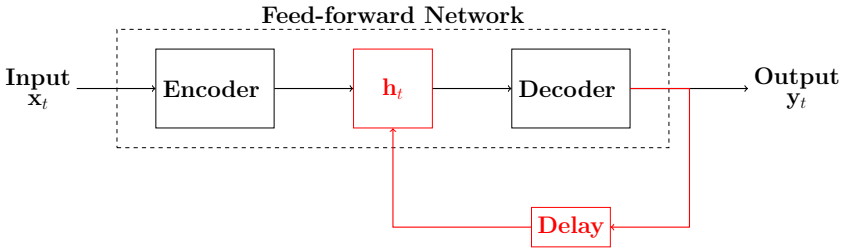


DeconvNet

Understanding DeconvNet

Recurrent Neural Network (RNN)

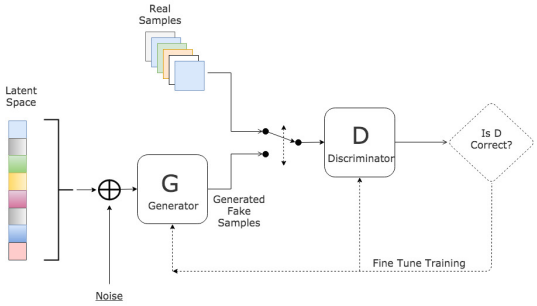
Tool for sequence processing tasks (natural language, video, ...)



- RNN incorporates a feedback loop (with delay)
- Transfer function
  - $\mathbf{h}_t = f(\mathbf{W_1}.\mathbf{x}_t + \mathbf{W_2}.\mathbf{y}_{t-1})$
  - $\mathbf{y}_t = g(\mathbf{W_3}.\mathbf{h}_t)$
  - $\mathbf{h}_t$ accumulates experience

Generative Adversarial Network (GAN)

Technology behind DeepFake, etc.



- Generator attempts to create realistic samples
- Discriminator tries to differentiate between real samples and the fakes (generated)
- Both networks are trained together – both improves with training

Quiz

Quiz 06-01

End of Module 06-01