Convolutional Neural Networks

INSTRUCTOR: HONGJIE CHEN

JUNE 21ST 2022

Types of NNs

- A normal DNN
  - Time expensive
  - Parameter explosion
  - Universal but unnecessary

- Data format
  - Images (2D), Convolutional Neural Networks
  - Time-series (variable length), Recurrent Neural Networks
  - Networks (graph structure), Graph Neural Networks
  - And more?

Convolution

- A mathematical operation on two functions $x()$ and $w()$ that produces a third function $y()$ that can be viewed as a modified version of one of the original functions $x()$

$$y(i) = \int_t x(t)w(i-t)dt, \text{ written as } y(i) = (x * w)(i)$$

- Demo: https://www.youtube.com/watch?v=das6mpj13XQ
- Smoothing effect (reading: link)

Discrete Convolution

- Discrete convolution

$$y(i) = \sum_{t=-\infty}^{\infty} x(t)w(i-t)$$

- Two dimensional convolution or even higher dimensions

$$y(i,j) = \sum_{t_1=-\infty}^{\infty} \sum_{t_2=-\infty}^{\infty} x(t_1,t_2)w(i-t_1, j-t_2)$$

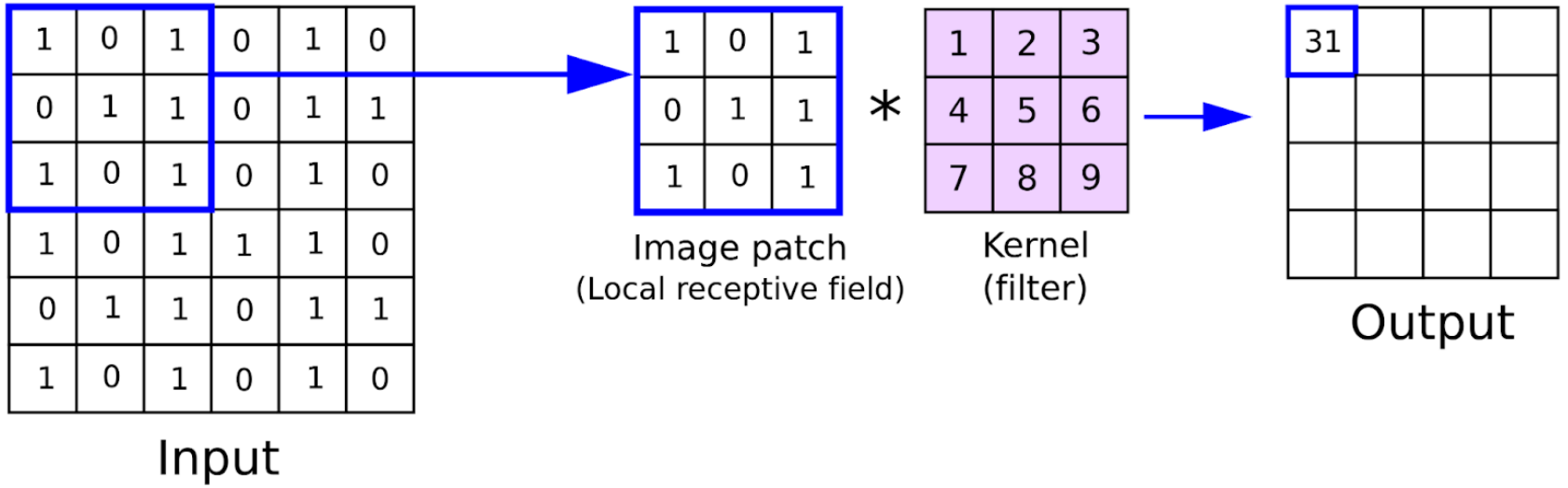
Convolution in Neural Networks

- A convolution denotes the linear combination of **a subset of units of a specific pattern** of weights

- Instead of all weights as in dense layers

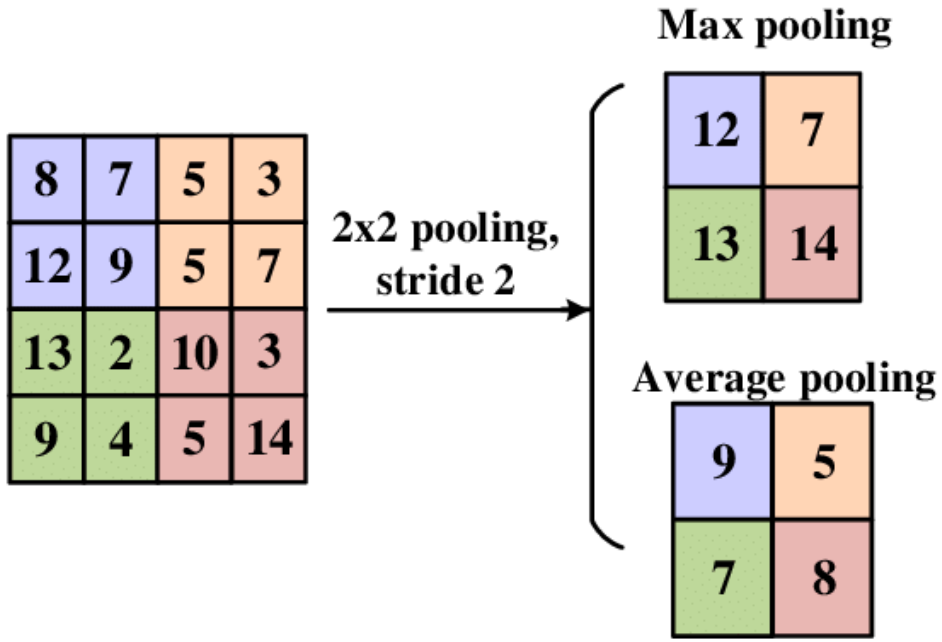
CNN

- A CNN is a NN that consists of convolutional layers (and optionally pooling layers).



Pooling

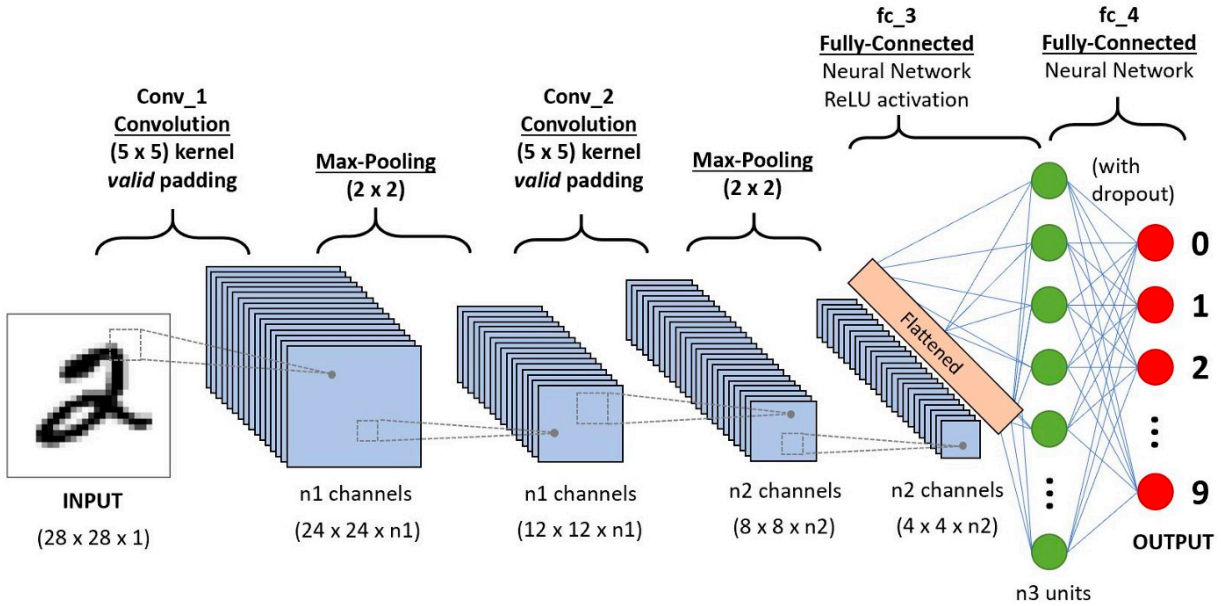
- Downsample features by aggregation



- Popular pooling functions
  - Max, mean, sum, product, etc.

Classical CNN Example

- Digit recognition



- Back in history: https://www.youtube.com/watch?v=FwFduRA_L6Q

Advantages of CNN

- Time: fewer weights to calculate, faster
- Sapce: fewer links, or equivalently weight parameters

- Locally equivariant representation
  - Capture the pattern
  - Better Generalization

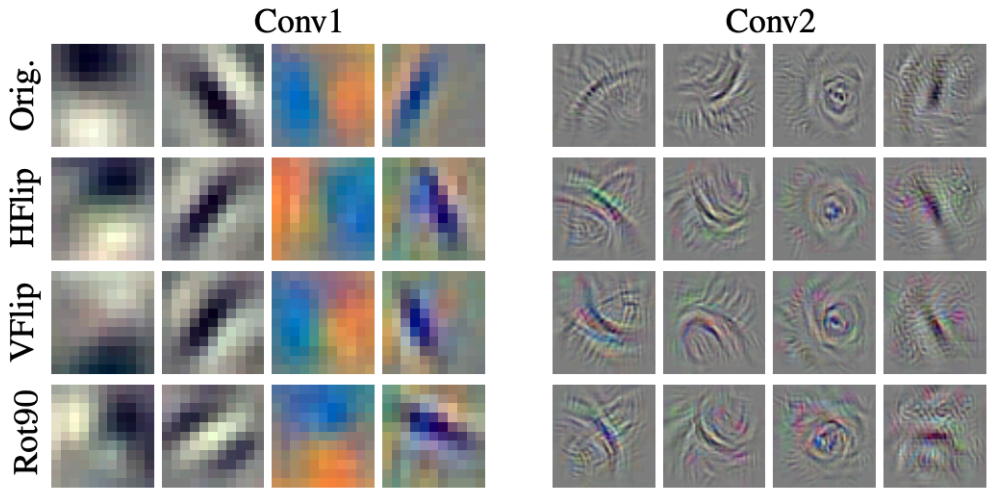


Figure credits: paper

Parameters of Convolutional Layers

- **Number of filters**: an integer indicating how many filters / channels for each window where one filter yield one scalar value
- **Kernel size**: a tuple (width, height) indicating the size of the window
- **Stride**: a tuple (horizontal, vertical) indicating shifting step between window
- **Padding**: “valid” or “same” where the former indicates no padding, thus the resulting values will be in a shrinking size, while the latter indicates padding zeros to let the resulting values remain the same size.

*Don’t get confused with the Kernel in kernel perceptron

CNN Parameters Example

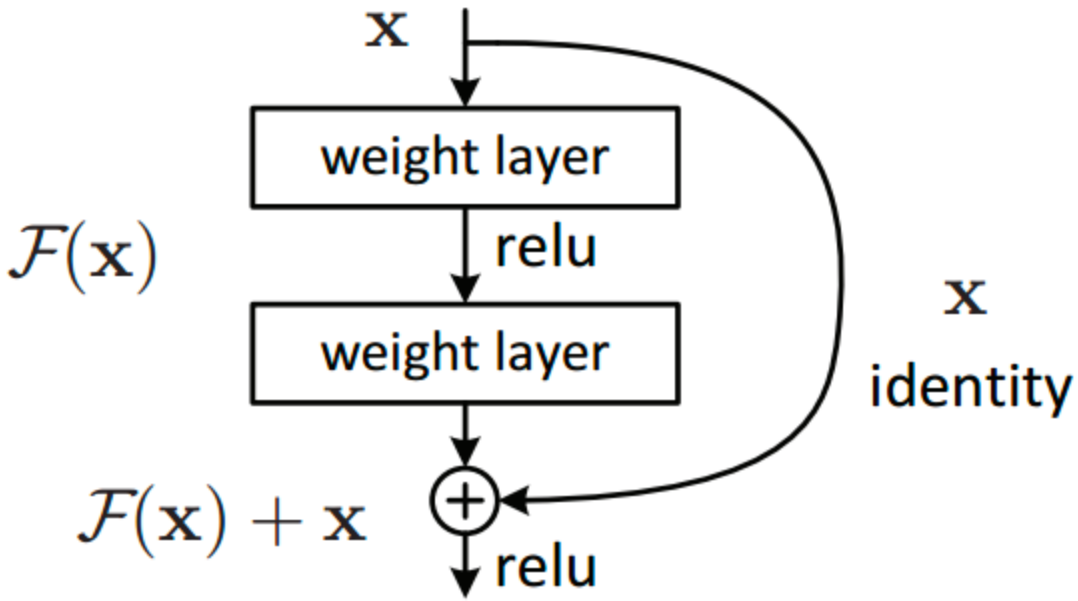
- **Number of filters, Kernel size, Stride, Padding**
- Let's draw a diagram

Fewer Large Filters or More Small Filters

- More small filters
  - Fewer parameters
  - Allow deeper structure

Residual

- Also called skip connection
  - Address vanishing gradients



ResNet for Image Recognition (Even deeper)