CONVOLUTIONAL NEURAL NETWORKS

Shingchern D. You
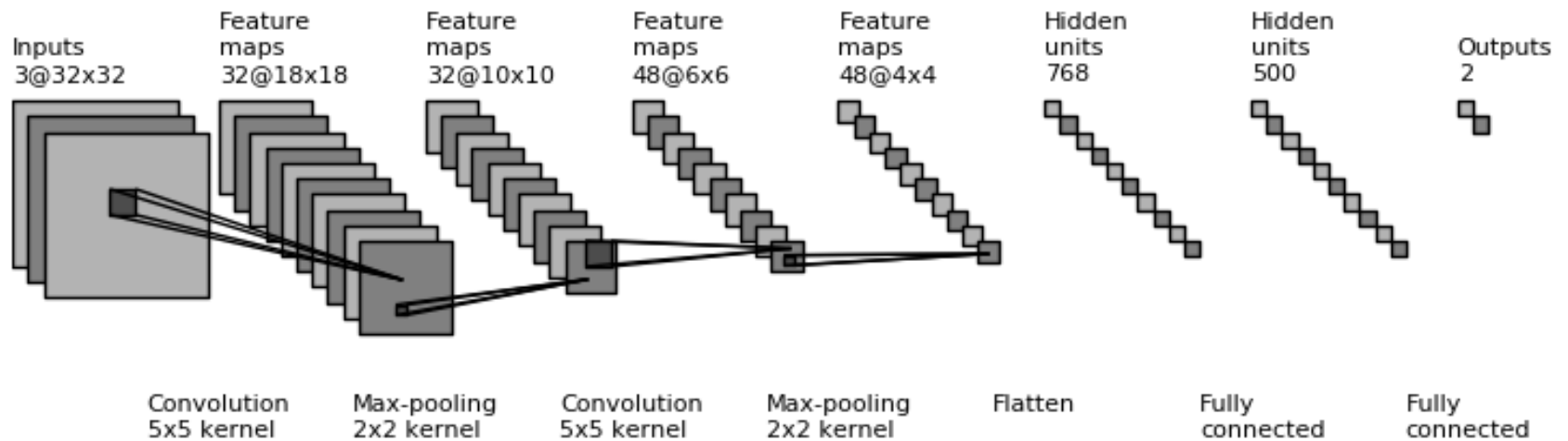
References

- A good source of reference is Stanford's CS231n lecture notes

Overview

□ A typical picture (picture source:
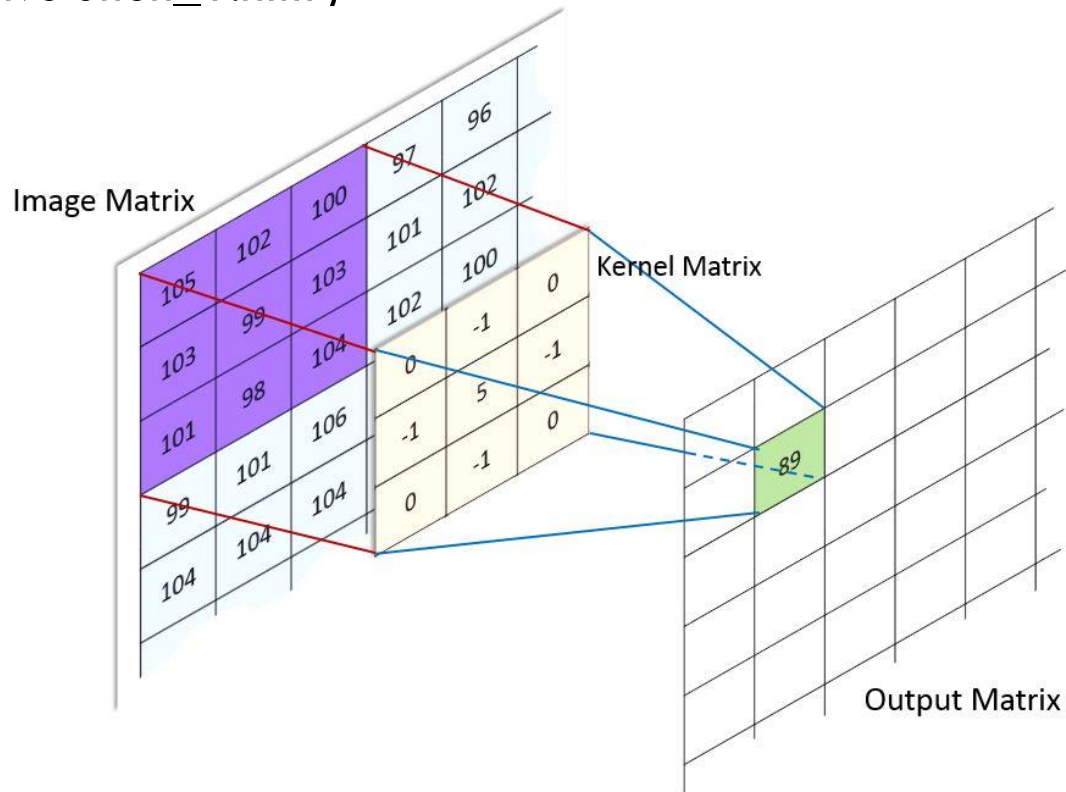https://datascience.stackexchange.com/questions/28339/convolutional-neural-networks-layer-sizes)

Layer partitioning

- Convolutional layer
- Max-pooling layer
- Fully-connected layer
- Output layer

Convolutional layer

□ **What does it mean by convolution** (Figure source: http://machinelearninguru.com/computer_vision/basics/convolution/image_convolution_1.html)

Convolutional layer

- Convolution means multiplication and add
- Some parameters in a filter (kernel)
  - Size: 5x5x3 (width, height, depth = color)
  - 5x5 is also called receptive field
  - Stride: (hop size)
  - Number of kernels (filters)
  - Zero-padding (for boundary pixels)
  - Meaning of 1x1 convolution (actually 1x1x?)

Convolutional layer

- Why is this layer called convolutional layer
    - Weight (parameter) sharing
    - Make features translation-invariant (not always good though)
- With weight sharing, fewer "free variables" are present in this layer

Convolutional layer

□ What is learned in the convolutional layer (picture source: https://www.slideshare.net/nmhkahn/case-study-of-convolutional-neural-network-61556303)
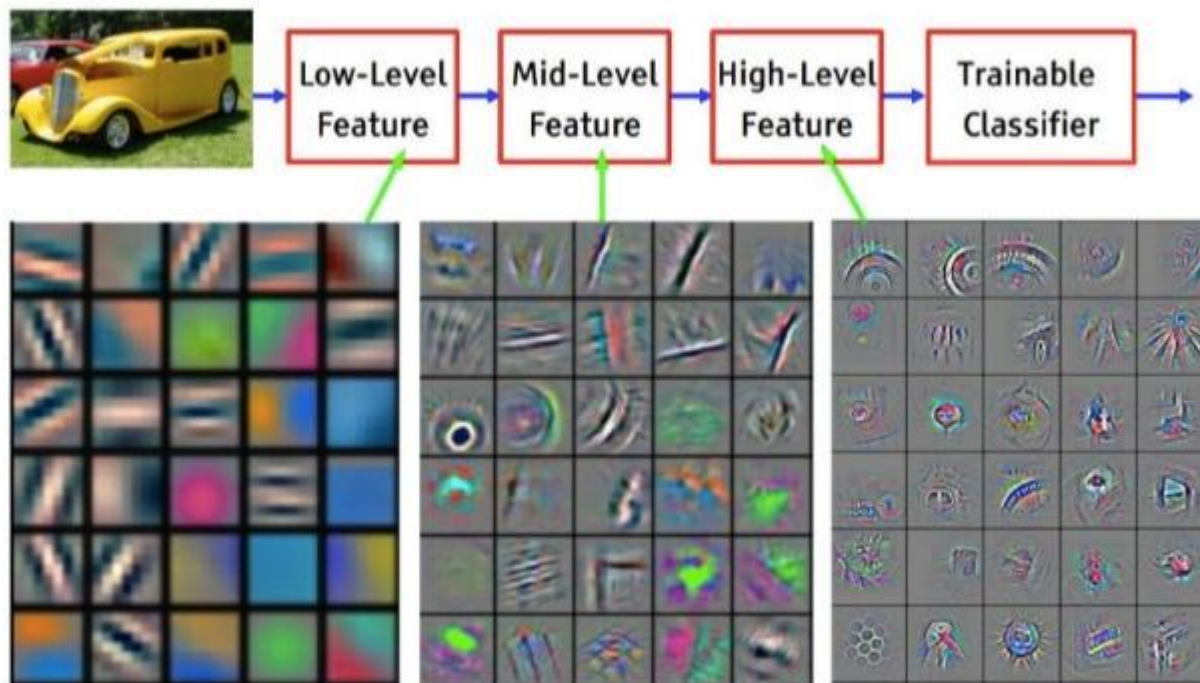
Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]

Pooling layer

- Pooling layer is used as a "down sampling" to reduce the size of internal representations (size of internal layers)

- Size (spatial extent) 2x2 (actually means 2x2**x1**) or 3x3

- Usually stride of pooling is set to 2

- Max pooling is most widely used

- Other types of pooling methods existing

Fully-connected layer

- Sometimes we use a term "flatten" to mean the conversion from 3-D structure to 1-D
- Fully-connected layer is similar to traditional neural networks
- May contain lots of weights

Training convolutional neural networks

- Depending on the problems, training may take a long time

- You may want to use existing tools (such as tensorflow and keras) to invoke GPU's computing power

- Our experience is that speed up is up to 10x with one GPU card