Feature Extraction and Image Processing for Object Detection

Shrey Gupta

Applied Machine Learning (HOUSECS 59-01), Duke University

November 14, 2018

Recall

- Recall the intuition behind convolutional neural networks: create *templates*, and use some similarity measure (e.g. dot product) to match to the image.
- This is important in *pattern recognition*, where, for example, we may want to detect edges in an image.

Edge Detection

- A common problem in computer vision is detecting the edges surrounding the objects in an image.

Image source: MathWorks

Fixed Edge Detection

$$G_x = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \quad G_y = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

- Consider G_x and G_y, simple convolutional templates (filters) for horizontal and vertical edge detection, respectively.
- Intuitively, think about the *gradient* in the horizontal and vertical directions.

Image source: Edge Detection (Image Processing)

Thresholds

- Performing a convolution between the template (filter) G and an image I will result in another image C.
- Larger values $|C_{(x,y)}|$, in magnitude, signify a higher probability of the presence of an edge.
- Apply thresholds across the image such that if $|C_{(x,y)}| \geq \tau$, (x, y) is part of an edge in the original image.

CNN Edge Detection

- There are several other more complicated edge filters, but many times they are chosen with prior knowledge of the image or data set.
- Instead, use a neural network that treats the elements of the template as parameters to "learn" the optimal filter for edge detection.

Deep Neural Networks

- The architectures covered in the previous lecture are *deep* convolutional neural networks, since they require stitching together several layers–making the network deep.
- However, without transfer learning and/or a large amount of data, it can be difficult to achieve the desired results.

Feature Extraction

- Feature extraction can be a useful tool for making your network more shallow, and hence require less (training) data.
- By choosing handpicked features, we might replace what a network may do in the first several layers.
- While the mechanics behind state-of-the-art feature extractors are beyond the scope of this class, two examples are HOG (histogram of oriented gradients) and SIFT (scale-invariant feature transform).
  - Or perhaps use a modified version of the image (e.g. with an edge filter applied) as the input.

Notebook

- Today's notebook will briefly explore the HOG feature extractor for pedestrian detection.