Homework 4

Fall 2019
(Due: Nov. 18, 2019)

Introduction

This assignment is on convolutional networks, which is covered in Chapter 9 of the recommended text.

Exercises

1. (The Convolution Operation)

   a) What is the difference between *convolution* and *cross-correlation* (state the formulations for the discrete case and explain the differences)? Which operation is commutative? Justify your answer.

   b) Which operation (between *convolution* and *cross-correlation*) is used in DNNs? Explain why choosing to use this particular operation in a DNN does not matter.

2. (Variants of the Basic Convolution Function)

   a) What is the difference between covolutional layers and locally connected layers? When should we use each of them?

   b) What kind of transformational equivariance does parameter sharing in convolutional layers, with pooling over spatial regions, cause in the output? Explain.

3. (Pooling)

   a) Explain the term *pooling*.

   b) Discuss the advantage of *pooling* in terms of:

      i. Computation

      ii. Generalization

   c) Define the following two terms:

      i. Pooling width

      ii. Stride

   d) What kind of output invariance does pooling over spatial regions introduce?

   e) How to make the network learn the type of output invariance that pooling introduces?

4. (Zero Padding)

   a) What is *downsampled convolution*? What is the stride in *downsampled convolution*?

   Consider the input to a Convolutional Neural Network to be a 16 pixel image (i.e., we have a 1×16 vector input). Suppose our convolutional kernel width is 6 and we are not using zero padding.

   b) How many convolutional layers can we have (given a kernel width of 6 in each layer and that we do not use zero padding in any layer)? Justify your answer.

   c) Is the kernel being moved in the last layer? Justify your answer.

d) What is the problem with using a smaller kernel width to address the problem raised by your answers above?

e) As convolutional layers are added to a neural network, the output size shrinks until we are eventually left with a 1×1 sample at the output at which point additional convolutional layers cannot be meaningfully considered. Explain how zero-padding is used to avoid this problem.

f) In MATLAB terminology, what is the name of the convolutional operation that employs the solution stated in (e)?

5. (Tiled Convolution)

a) What is *tiled convolution*? What is its associated memory requirement?

b) When using *tiled convolution*, our output locations can cycle through a set of t different choices in kernel stack in each direction. What values of t makes *tiled convolution* equivalent to:

i. Normal Convolution?

ii. Locally Connected Layers?

Bonus

6. (Data Types)

a) Give examples of each of the following data types for the single channel and multichannel case:

i. 1-D data types

ii. 2-D data types

iii. 3-D data types

b) When is using convolution for processing variably sized inputs suitable? Justify your answer.