Chapter 6

Simon Ward-Jones
simonwardjones16@gmail.com

October 18, 2020

Link to chapter

1 Convolutional Neural Networks

- Convolutional Neural Networks
- Translation invariance in images implies that all patches of an image will be treated in the same manner.
- Locality means that only a small neighbourhood of pixels will be used to compute the corresponding hidden representations.
- In image processing, convolutional layers typically require many fewer parameters than fully-connected layers.
- The core computation of a two-dimensional convolutional layer is a two-dimensional cross-correlation operation. In its simplest form, this performs a cross-correlation operation on the two-dimensional input data and the kernel, and then adds a bias.

2 Padding and Stride

- Padding can increase the height and width of the output. This is often used to give the output the same height and width as the input.
- The stride can reduce the resolution of the output, for example reducing the height and width of the output to only 1/n of the height and width of the input (n is an integer greater than 1).

3 Multiple Input and Multiple Output Channels

- Multiple channels can be used to extend the model parameters of the convolutional layer.
- The 1x1 convolutional layer is equivalent to the fully-connected layer, when applied on a per pixel basis.
- The 1x1 convolutional layer is typically used to adjust the number of channels between network layers and to control model complexity.

4 Pooling

- Taking the input elements in the pooling window, the maximum pooling operation assigns the maximum value as the output and the average pooling operation assigns the average value as the output.
- Padding and Stride - Similar to convolutions
- The pooling layer's number of output channels is the same as the number of input channels.

References

Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. (2020). *Dive into deep learning.* (`https://d2l.ai`)