

COMP9444

Neural Networks and Deep Learning

Term 2, 2024



Week 5 Tutorial: Convolutional Neural Networks

1. Convolutional Network Architecture

One of the early papers on Deep Q-Learning for Atari games (Mnih et al., 2013) contains this description of its Convolutional Neural Network:

“The input to the neural network consists of an $84 \times 84 \times 4$ image. The first hidden layer convolves 16 8×8 filters with stride 4 with the input image and applies a rectifier nonlinearity. The second hidden layer convolves 32 4×4 filters with stride 2, again followed by a rectifier nonlinearity. The final hidden layer is fully-connected and consists of 256 rectifier units. The output layer is full-connected linear layer with a single output for each valid action. The number of valid actions varied between 4 and 18 on the games we considered. ”

For each layer in this network, compute the number of:

- (a) weight per filter in this layer (including bias)
- (b) width and height of layer (convolutional layers)
- (c) neurons in this layer
- (d) connections into the neurons in this layer
- (e) independent parameters in this layer

You should assume the input images are gray-scale, there is no padding, and there are 18 valid actions (outputs).

2. Weight Initialization

Briefly describe the problem of vanishing or exploding gradients, and how Weight Initialization can be used to prevent it.

3. Batch Normalization

Briefly describe the Batch Normalization algorithm.

4. Residual and Dense Networks

Explain the difference between a Residual Network and a Dense Network.

5. Other Questions

Any questions or discussion about image processing, societal impact, or any other aspect of the course.