Advanced Robotic Perception: Lecture 6
Lecture Notes

Victor Risager

September 28, 2023

1 Recap

2 CNN

2.1 Individual neuron

2.2 Convolutional vs fully connected

Convolutional neural network, has multidimensional data. You can have 1d convolutions, but often we do 3×3 where the last dimension can e.g. be the color channels of the image. Thus x, y and colors.

A standard fully connected nerual network has only 2 dimensions.

Filters are run through each channel at the image, and values are summed across the filter, AND across the channels. This gives the new activation function.

$$a = \sum_{i=0}^{n} w_i \cdot x_i + b \qquad (1)$$

The filter values are learned and not handcrafted.

Equivariance is if you do the same in the input as you do in the output.

By design, the network is scaling invariant, but in practice, the size of the network can actually result in scaling equivariance.

ImageNet has 14 million images an d 20.000 classes. Use transfer learning.

3 AlexNet

It has 8 layers.

Convulutional $F \times F \times D \times K + K$

- F is the filter size

- D is input depth

- K number of filters

The output size parameters are padding and stride.

4 Activation function

4.1 ReLU

$$f(x) = max(0, x) \tag{2}$$

Easier to train.

5 Pooling

Sampling an image from a kernel.

- Max-pooling

- Average pooling

- Sum pooling

6 Batch normalization

Learn what transformation is best to do from every layer to the next layer.
The hyper parameter is you can choose which layers have batch normalization.
It is done during training, and it reduces the training time.

7 Optimization algorithms

Mini-batch gradient descent, is doing gradient descent based on smaller batches
of the pixels.

8 VGG-Net

Smaller filters than AlexNet.

9 GoogLeNet

Have different sizes and combinations of convolutions and then learn what combination and size is best.

10 ResNet

Allow gradient flow to bypass some parts of the network, to avoid degration.

11 General network training tips

There is more than one parameter to choose a network from.

12   Data augmentation

Use different techniques to alter the training images, to augment the training process. But be carefull, you could e.g. crop out the desired part of the image.