**Home Work Assignment 1**

Or Kozlovsky 305447476

Noa

Task 1: Deep Learning Introduction

Dry section

1. Working with a convolution network:
   1. the output dimensions of the following layers are with an Input RGB image of size 128X128X3:

|  |  |  |
| --- | --- | --- |
| Layer # | Layer description | output dimensions |
| 1 | convolution with 64 kernels of size 1X1X3 | 128X128X64 |
| 2 | max pooling of size 2x2 | 64X64X64 |
| 3 | convolution with 32 kernels of size 5X5X64 (no zero padding). | 60X60X64 |

* 1. We will explain the calculation of a 2D convolution with a kernel of size 1X1X3.

The output of the convolution is the sum of the same pixel in all 3 channels of the image. i.e. let’s assume we have the following kernel:



|  |  |  |
| --- | --- | --- |
| **1** | **1** | **1** |

And the following image channels:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **1** | **2** | **3** |  | **1** | **2** | **3** |  | **1** | **2** | **3** |
| **4** | **5** | **6** |  | **4** | **5** | **6** |  | **4** | **5** | **6** |
| **7** | **8** | **9** |  | **7** | **8** | **9** |  | **7** | **8** | **9** |

Then the output, as described above, will be:

|  |  |  |
| --- | --- | --- |
| **3** | **6** | **9** |
| **12** | **15** | **18** |
| **21** | **24** | **27** |

* 1. We will present the output of the convolution of the given sub-image with the following average filter:
     1. With stride = (1,1) and padding = (1,1):
     2. With stride = (2,1) and padding = (0,0):

1. We selcted to present the VGG16 classification architecture:

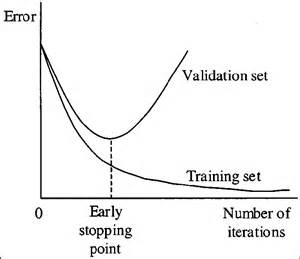
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Layer # | Layer description | output dimensions | Kernel size | Stride  (dx, dy) | Padding\dilation  (Dx, Dy) |
| 0 | input | 224x224x3 | - | - | - |
| 1 | Convolution | 224x224X64 | 3X3 | (1, 1) | (1, 1) |
| 2 | Convolution | 224x224X64 | 3X3 | (1, 1) | (1, 1) |
| 3 | max pooling | 112x112x64 | 2X2 | (2, 2) | (1, 1) |
| 4 | Convolution | 112x112x128 | 3X3 | (1, 1) | (1, 1) |
| 5 | Convolution | 112x112x128 | 3X3 | (1, 1) | (1, 1) |
| 6 | max pooling | 56X56x128 | 2X2 | (2, 2) | (1, 1) |
| 7 | Convolution | 56X56x256 | 3X3 | (1, 1) | (1, 1) |
| 8 | Convolution | 56X56x256 | 3X3 | (1, 1) | (1, 1) |
| 9 | Convolution | 56X56x256 | 3X3 | (1, 1) | (1, 1) |
| 10 | max pooling | 28X28x256 | 2X2 | (2, 2) | (1, 1) |
| 11 | Convolution | 28X28x512 | 3X3 | (1, 1) | (1, 1) |
| 12 | Convolution | 28X28x512 | 3X3 | (1, 1) | (1, 1) |
| 13 | Convolution | 28X28x512 | 3X3 | (1, 1) | (1, 1) |
| 14 | max pooling | 14X14x512 | 2X2 | (2, 2) | (1, 1) |
| 15 | Convolution | 14X14x512 | 3X3 | (1, 1) | (1, 1) |
| 16 | Convolution | 14X14x512 | 3X3 | (1, 1) | (1, 1) |
| 17 | Convolution | 14X14x512 | 3X3 | (1, 1) | (1, 1) |
| 18 | max pooling | 7X7x512 | 2X2 | (2, 2) | (1, 1) |
| 19 | Flatten | 1x25088 | - | - | - |
| 20 | Fully connected | 1x4096 | - | - | - |
| 21 | Fully connected | 1x4096 | - | - | - |
| 22 | Prediction (out) | 1x1000 | - | - | - |

* The output of one layer is the input of the sequential layer.
* The input size, kernel size, stride and padding dictamens the output size of each layer as describes in proviso sections

1. overfitting is when our trained model (net) doesn’t generalize well from our training data to unseen data. When a model learns the noise (training set) instead of the signal (observed (true) values) is considered “overfit” because it fits the training dataset but has poor fit with new datasets. We can recognize it for example, if our model saw 99% accuracy on the training set but only 55% accuracy on the validation set.

We can measure how well each iteration of the model performs When training a learning algorithm iteratively. Up until a certain number of iterations, new iterations improve the model. After that point, however, the model’s ability to generalize can weaken as it begins to overfit the training data.

Early stopping refers stopping the training process before the learner passes that point.



In the above graph we can recognize two cases:

1. The error (accuracy) of the Training set which decrease as monotonously as the number of iterations.
2. The error of the Validation set which from a certain point begins to increase as the iterations continue.

Our goal when building our learning algorithm is to recognize this point and make sure we not passing it.