Home Assignment 4

Tudor Andrei Dumitrascu 2682776

July 08, 2021

25

Overfitting occurs when the model performs really well on the train set and poorly on the validation set.

Methods:

- Regularization
- Normalization: Batch Normalization or Layer Normalization
- Dropout
- Data Augmentation

26

The result of the first conv is: For the kernel

$$\begin{bmatrix}1 & -2 & 0\\ 0 & 4 & -1\\ 0 & 2 & 0\end{bmatrix} \star \begin{bmatrix}0 & 1 & 2 & 0 & 1\\ 0 & 0 & 1 & 2 & 0\\ 0 & 0 & 0 & 0 & 0\\ 5 & 0 & 1 & 0 & 5\\ 1 & 1 & 1 & 1 & 1\end{bmatrix} = \begin{bmatrix}-2-1 & 1-4+4-2 & 2+8\\ 0 & 2-2 & 1-3\\ -1+2 & 4+2 & -5+2\end{bmatrix} = \begin{bmatrix}-3 & -1 & 10\\ 0 & 0 & -3\\ 1 & 6 & -3\end{bmatrix}$$

$$\begin{bmatrix}1 & -2 & 0\\ 0 & 4 & -1\\ 0 & 2 & 0\end{bmatrix} \star' \begin{bmatrix}0 & 1 & 0\\ 2 & 0 & 2\\ 1 & 1 & 1\end{bmatrix} =$$

First row

$$\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix} + \begin{bmatrix}1 & -2 & 0\\ 0 & 4 & -1\\ 0 & 2 & 0\end{bmatrix} + \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix} +$$

Second row

$$+\begin{bmatrix} 2 & -4 & 0 \\ 0 & 8 & -2 \\ 0 & 4 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 2 & -4 & 0 \\ 0 & 8 & -2 \\ 0 & 4 & 0 \end{bmatrix}$$

third row

$$+\begin{bmatrix} 1 & -2 & 0 \\ 0 & 4 & -1 \\ 0 & 2 & 0 \end{bmatrix} + \begin{bmatrix} 1 & -2 & 0 \\ 0 & 4 & -1 \\ 0 & 2 & 0 \end{bmatrix} + \begin{bmatrix} 1 & -2 & 0 \\ 0 & 4 & -1 \\ 0 & 2 & 0 \end{bmatrix}$$

Final matrix

$$+\begin{bmatrix} 0 & 1 & -2 & 0 & 0 \\ 0+2 & 0-4 & 4+0+2 & -1-4 & 0 \\ 0+0+1 & 0+8+0-2+1 & 2-2+0+0-2+1 & 0+8-2 & -2 \\ 0 & 4+4 & 0-1+4 & 4-1+4 & -1 \\ 0 & 2 & 0+2 & 0+2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & -2 & 0 & 0 \\ 2 & -4 & 6 & -5 & 0 \\ 1 & 7 & -1 & 6 & 0 \\ 0 & 8 & 3 & -4 & -1 \\ 0 & 2 & 0 & 2 & 0 \end{bmatrix}$$

27

Input: 32x32x3

- Kernel: 5x5, stride 1, channels 16
- params: 5*5*16*3=1200

Out1: 28x28x16

- AvgPool: 2x2, stride 2 -> params: 0

Out2: 14x14x16

- Kernel: 3x3, stride 1, channels 8
- params: 3*3*16*8 = 1152

Out3: 12x12x8

- AvgPool: 2x2, stride 2 -> params: 0

Out4: 6x6x8

Flatten: 288> params: 0

Dense: 32 -> params: 288*32 = 9216

Dense: 16 -> params: 32*16 = 512

- Total parameters: 12,080
- Number of classes: 16 classes, due to the fact that the last layer has 16 outputs, which will be used trough a Softmax to determine the prob for each class

28

Input: 32x32x1

- Conv1: 3x3, stride 1, channels 64
- params: 3*3*1*64 = 576

Out1: 30x30x64

- MaxPool: 2x2, stride 1

Out2: 29x29x64

- Conv2: 5x5, stride 1, channels 32
- params: 5*5*64*32 = 51200

Out3: 25x25x32

- MaxPool: 2x2, stride 2

Out4: 12x12x32

- Conv3: 3x3, stride 1, channels 16
- params: 3*3*32*16 = 4608

Out5: 10x10x16

- MaxPool: 2x2, stride 2

Out6: 5x5x16

Flatten: 400

Dense1: 16 -> params: 400*16 = 6400

Dense2: 10 -> params: 16*10 = 160

- Total params: 62,944

29

30

Input: 64x64x1

- Conv1: 3x3, stride 1, channels 32
- params: 3*3*1*32 = 288

Out1: 62x62x32

- MaxPool: 2x2, stride 1

Out2: 61x61x32

- Conv2: 5x5, stride 1, channels 32
- params: 5*5*32*32 = 25600

Out3: 57x57x32

- MaxPool: 2x2, stride 2

Out4: 28x28x32

- Conv3: 3x3, stride 1, channels 32
- params: 3*32*32 = 9216

Out5: 26x26x32

- MaxPool: 2x2, stride 2

Out6: 13x13x32

Flatten: 5408

Dense1: 128 -> params: 5408*128 = 692224

Dense2: 10 -> params: 10*128 = 1280

- Total params:728,608