Weiyue Leng, wl4817, 01351038

1. CNN_Introduction: Task 1

Design Choices Figure 1 demonstrates a bar graph which portrays how testing and training accuracy vary with respect to different design choices. A simple default architecture with one 2D convolution layer of 16 filters and 10 neurons dense layer, followed by a softmax activation function was used to start with.

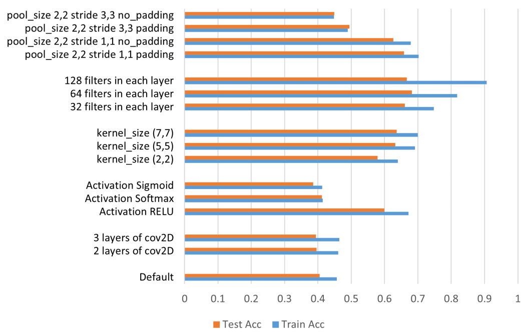

Figure 1. Training-Test accuracy against various design choices

Discussion As demonstrated in the Figure 1, the number of convolution layers does not affect the performance significantly. However, enabling the RELU activation layer massively improves both accuracies. It can be seen that as the number of filter increases, the testing accuracy roughly stays the same; whereas the training accuracy is boosted, due to the overfitting problem. Having a smaller stride size also contributes to a higher performance. This is because two pixels are less correlated when they are further apart from each other. Therefore, a large stride in the pooling layer leads to a high information loss. Another positive effect is observed when the padding is enabled in the Maxpooling layer. Padding is added to edge pixels of an image to allow more frame covered by the kernel; therefore, the image can be analysed more accurately. Hence, based on the above results, an optimal structure is proposed, as shown in Figure 2.

2. CNN Introduction: Task 2

Proposed architecture The proposed architecture in this task is very similar to the previous one. However, the final dense layer is replaced by one neuron, and the output activation function is changed to a linear function so that linear regression is performed. The approximate estimation error of this architecture is 54.54%. Figure 3 illustrates how the training and validation loss vary with respect to the number of epochs for different architectural design choices. As the number of epochs increases, the training loss tends to decrease; nonetheless, the validation loss of the green curve and yellow curve gradually increases after 80 epochs, meaning that overfitting is introduced. The validation loss of the orange curve is relatively stable than others; therefore, it is chosen as the optimal architecture.

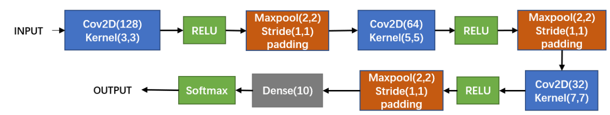

Figure 2. Optimal architecture with test accuracy of 74.1% and train accuracy of 85.9%

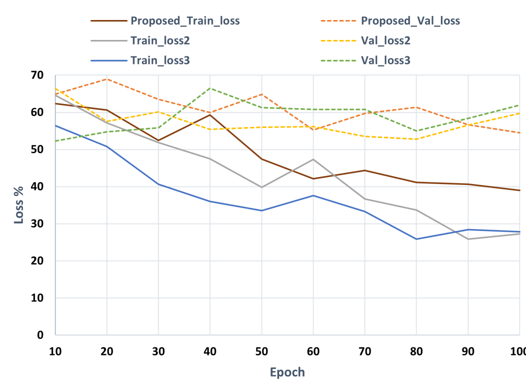

Figure 3. Training and Validation loss versus Epochs

Categories It can be seen from Table 1 that the proposed architecture performs the best for bedroom and bathroom images among two other back-up choices. Overall, the bathroom images give the worst validation loss.

Categories	Proposed	Back-up 1	Back-up 2
Kitchen	57.75%	64.96%	48.38%
Bedroom	46.76%	71.47%	49.96%
Bathroom	62.01%	74.43%	71.08%

Table 1. Validation loss for different categories of images with 3 architecture designs