jj1220, 01860908

1. CNN Introduction

Task 1.1. Shown in Figure 1, increasing number of filters per layer has significantly increased both the training and the validation accuracy as more weights are involved in learning and thus more features are extracted. Increasing filter size also increases accuracy as filter size determines the size of the receptive field and larger receptive field takes more values from neighbor pixels and hence increases accuracy. Nonetheless, if filter size gets too big, it risks of

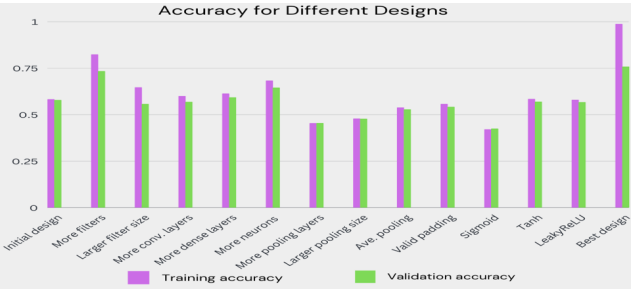

Figure 1. Training and validation accuracy for different designs

loosing local information/details. More convolution layers also increases accuracy as more complex features can be learned through hierarchical structure. More neurons in the dense layer increases number of connections and features and hence, increases the accuracy significantly. Number of pooling layers and pooling size has a noticeable effect on accuracy as it downsamples the activation map and hence decreases accuracy with less parameters. Finally, sigmoid in convolution layers decreases accuracy dramatically because the output of the sigmoid function can easily saturates with large positive and negative input values due to its small valid interval between -3 and 3, which kills neurons and causes vanishing gradient. Among all activation functions, relu has demonstrated the best performance. It can be observed from Figure 1, with increasing complexity and number of parameters especially in the best model with 7,480,202 parameters, the discrepancy between training and validation accuracy becomes more significant. This can be explained as high complexity has led to overfitting and bad generalisation.

Task 1.2. Figure 2 represents my best structure with 2 convolution layers with ReLu(128 and 256 filters) and 2 max pooling layers and 2 dense with ReLu (128 and 10 neurons) layers.

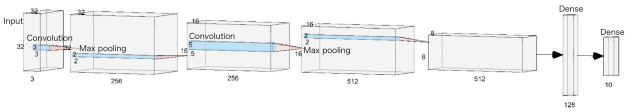

Figure 2. Best architecture

Task 2.1. Figure 3 4 convolution layers (16,32,64,64), two max pooling layers and two dense (4,1) layers.

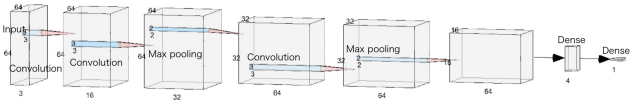

Figure 3. CNN architecture with 69% estimation error

Task 2.2. As displayed in Figure 4, the initial training loss is high and the gap between the initial training and validation loss is small as the initial model is relatively simple and has less parameters. When increasing layers and number of filters, the training loss decreases and the validation loss shows a rising trend as the model becomes more complex with more parameters/features which increases the generalisation error as a result of overfitting.

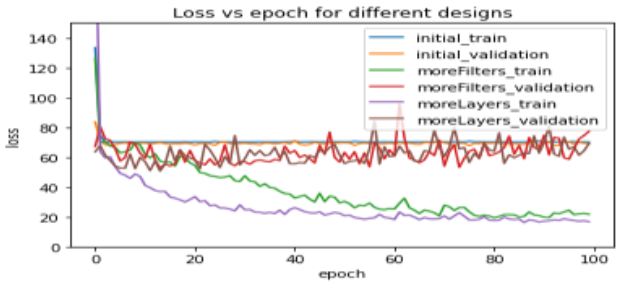

Figure 4. Loss vs epoch for different design choices

Task 2.3. All errors are lower than 75% in Table 1.

Image	Training loss	Validation loss	Estimation error
frontal	70.2301	69.0010	69.00%
kitchen	70.2317	69.0745	69.07%
bedroom	15.5139	57.6733	57.67%
bathroom	70.2312	69.0549	69.05%

Table 1. Estimation error for different images

References

2. Appendix