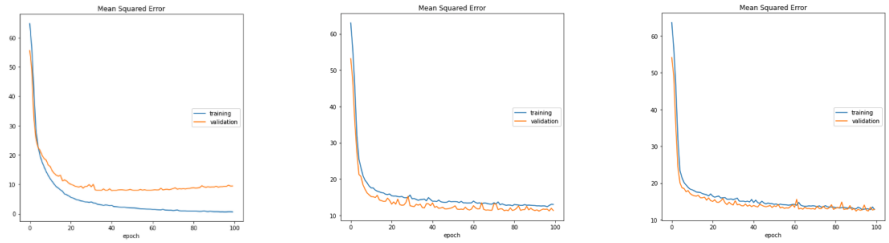
1 CNN For Image-To-Vector Regression

We use 'nn.MSELoss' as the loss function. The model took 17mins 4s to train on a NVidia GeForce GTX 1050 GPU. The final mean squared error (MSE) after training on training set was 0.62 and on validation set 9.43. The target ranges between -20 and 20, therefore we can say that the model is severely overfitting. We find an increasing MSE on the validation set which is a sign of overfitting.



(a) Baseline. (b) Global Pooling. (c) Data Augmentation.

Figure 1: Training and validation set MSE of 3 models.

2 CNN With Global Pooling

With global pooling, training time reduced to 9mins 51s (approx. 7mins reduction). The final MSE on training set was 13.0 and on validation set, it was 11.34. Although the MSE is higher than Baseline CNN, it generalises well.

3 CNN With Data Augmentation

The modification made in this model is a form of data augmentation, which is adding a layer of noise on top of our data before it is passed into the first convolutional layer. This ensures that the model cannot memorise the training data and prevents it from overfitting. Training time was significantly faster at 6mins 36s (approx. 10 mins reduction) and MSE on training set was 12.3 and on validation set was 11.8. The validation set performance is very close to the CNN trained with global pooling (MSE difference 0.46).