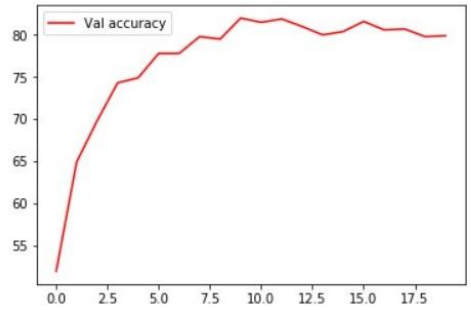
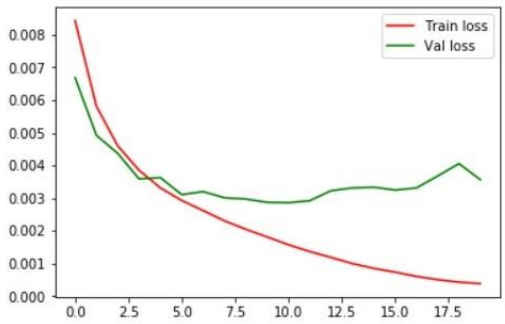
Question 2
Improve training of Convolutional Networks

Batch Normalization

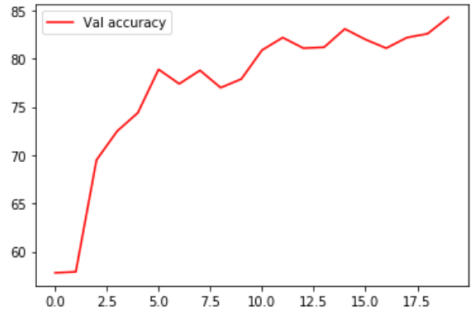
Results Q1a



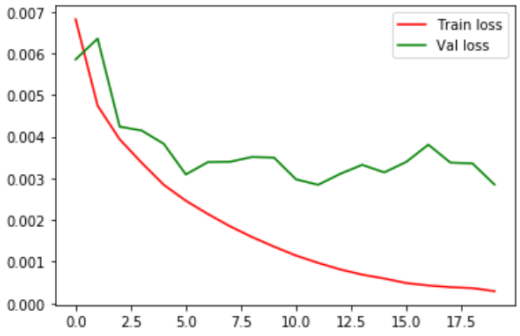
Accuracy of the network on the 1000 test images: 79.4 %



Results with batch normalization



Accuracy of the network on the 1000 test images: 81.8 %



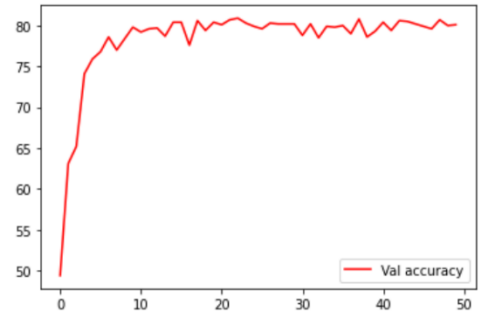
Comments:

As expected, the overall Accuracy with batch normalization is higher than without. The model with batch normalization was also faster to train. This method allows each layer to learn on a more stable distribution of inputs, and that accelerates de training of the network, also avoiding giving unfair higher weights to nodes that are dealing with features that are in a higher scale than other, event that can explain the improvement on the performance of the algorithm.

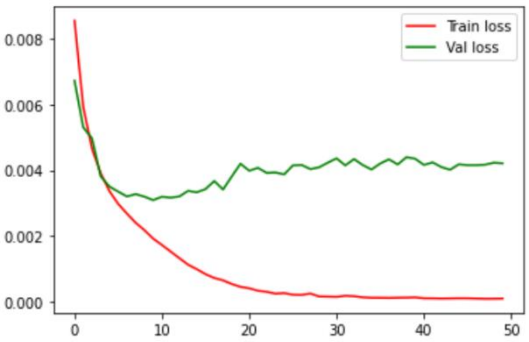
Early stopping

Training epochs: 50

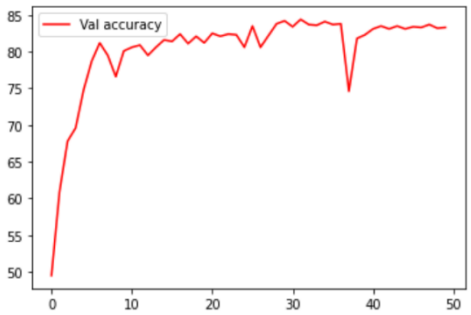
Results Q1a



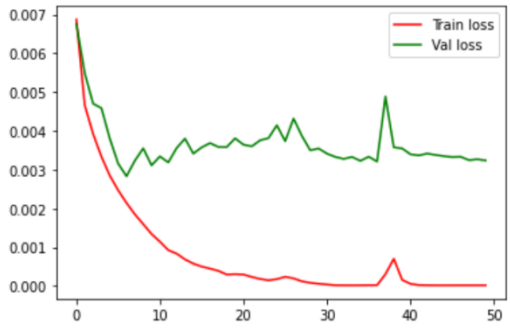
Accuracy of the network on the 1000 test images: 78.7 %



Results with batch normalization



Accuracy of the network on the 1000 test images: 82.0 %



Comments:

Early stopping is a way of avoiding overfitting in neural networks. It is not true that more epochs mean, better performance, there is a point where is more training, only means a better tracking of the error of the training set. While the training loss is getting smaller with each epoch, the validation loss, starts to stabilize, and eventually should start to rise. Considering that more epochs makes the algorithm more expensive, we think that 30 epochs for the model without batch normalization is a good estimation, and around 20 with batch normalization. This also matches the theory, that explains that batch normalization decreases the convergence time.