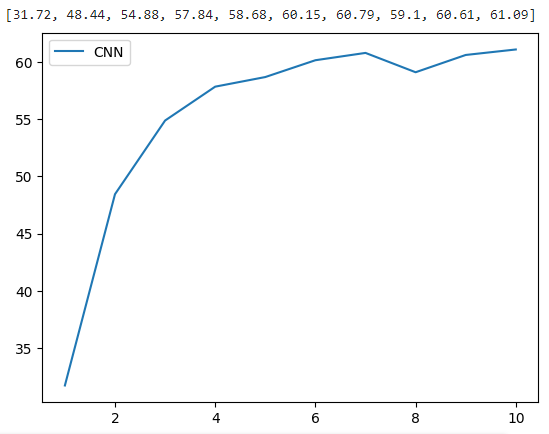
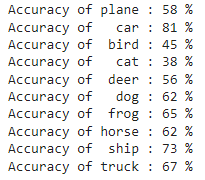
4102

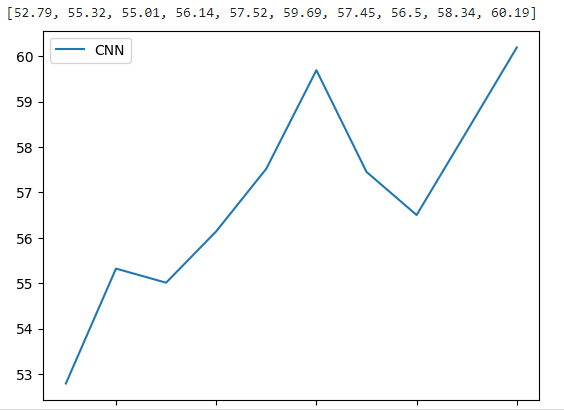
1.

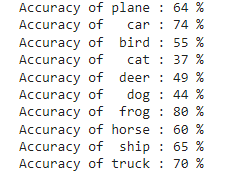
CNN:



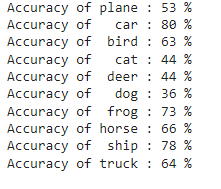
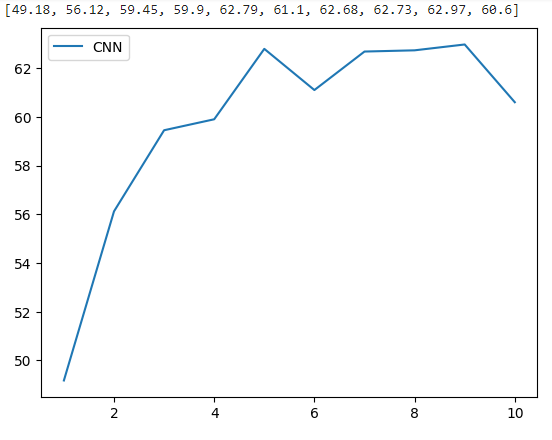


0 hidden layers:

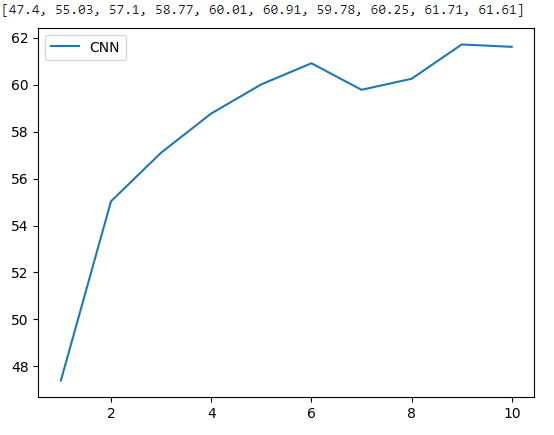


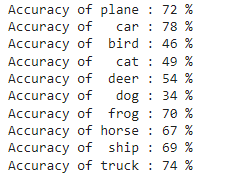


1 hidden layer:

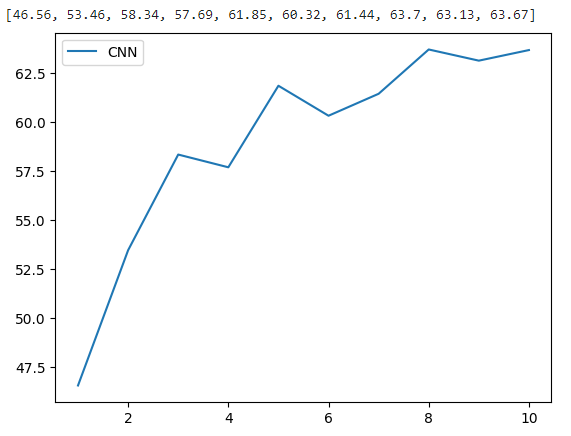


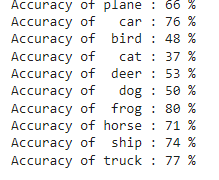
2 hidden layers:



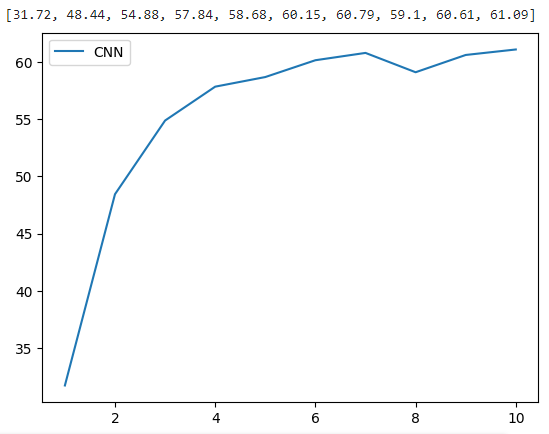


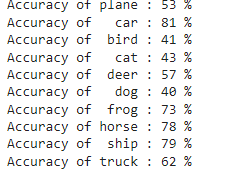
3 hidden layers:





4 hidden layers:

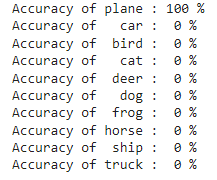


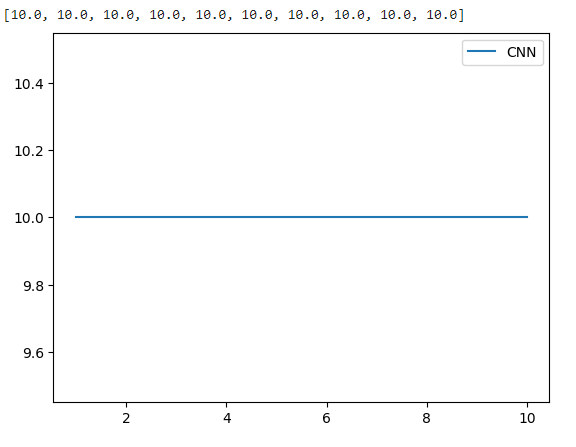


Comparing the accuracy of the CNN with the simple neural networks with 0, 1, 2, 3 and 4 hidden layers, it seems like there is an upward trend that correlates with the more layers there are, the more accurate the images become. This must be because more layers equals more applications, thus increasing the accuracy. However, the 4th layer seems less accurate which may be due to a slower rendering speed, and out of scale, which the program cannot detect, compared to 3 hidden layers.

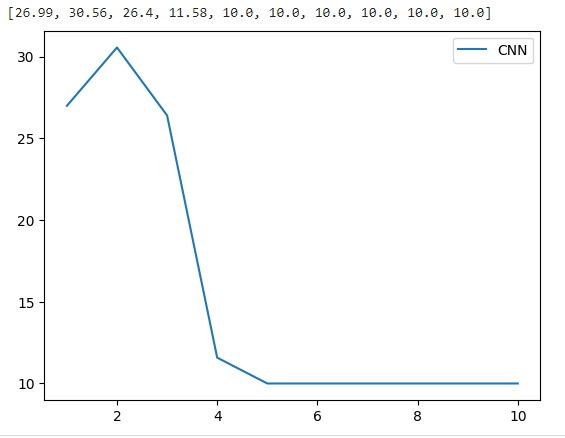
2.

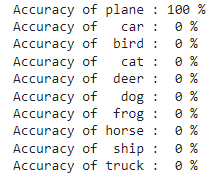
0.1:



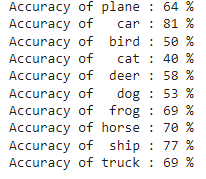


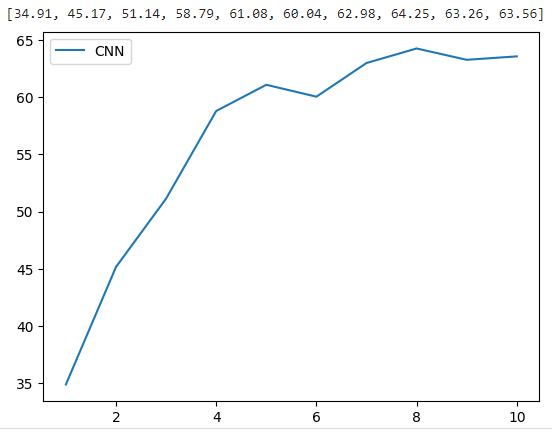
0.01:





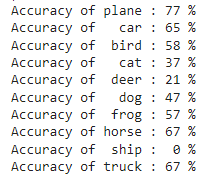
0.001:

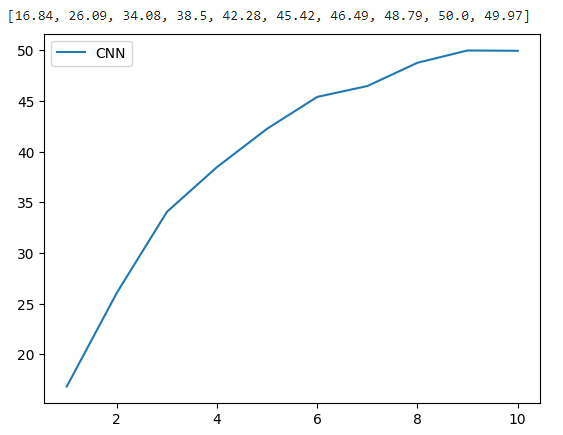






0.0001:

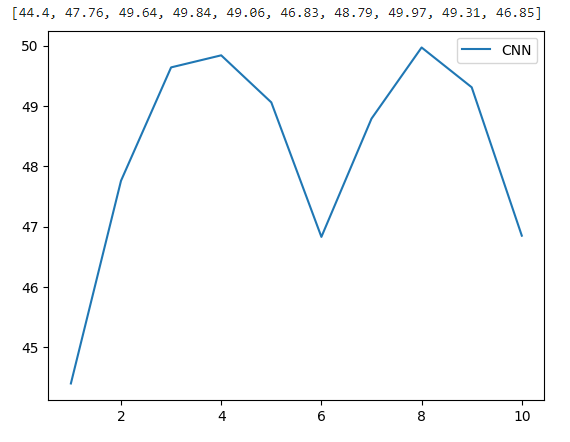




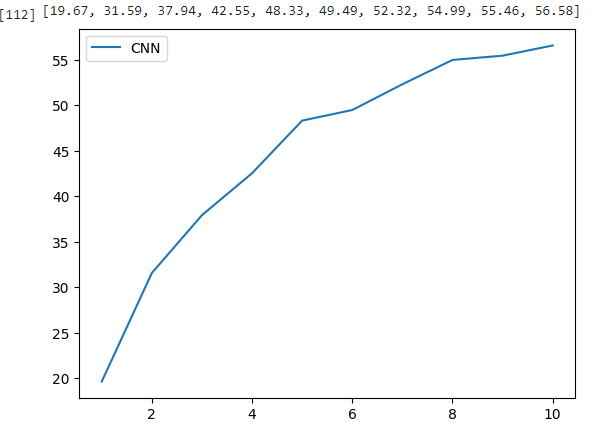
Comparing each graph, it seems the impact of learning rate on model convergence seems to increase the smaller the convergence meaning 0.0001 increasing the accuracy while a higher rate decreases the accuracy. This must be because

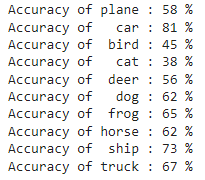
3.

1:

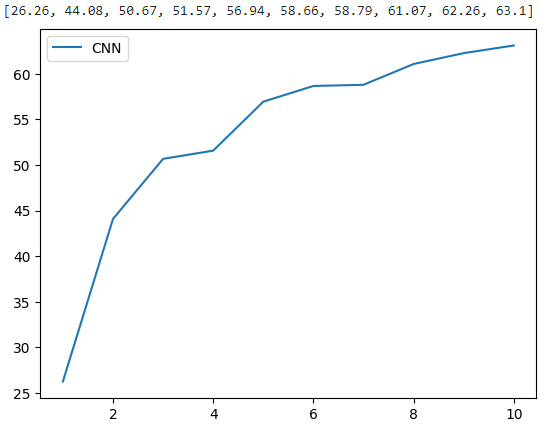


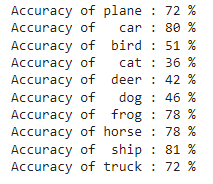
4:



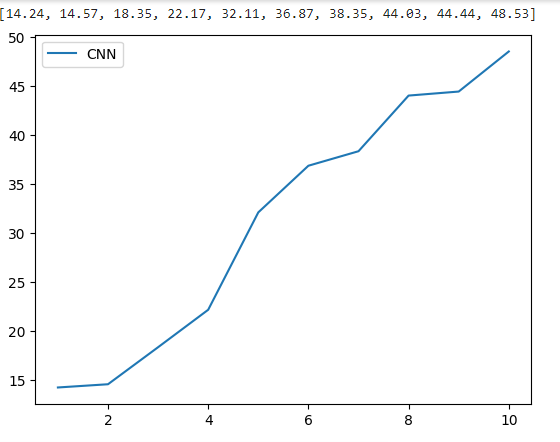


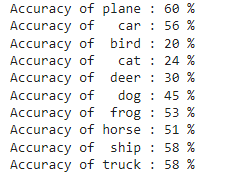
16:



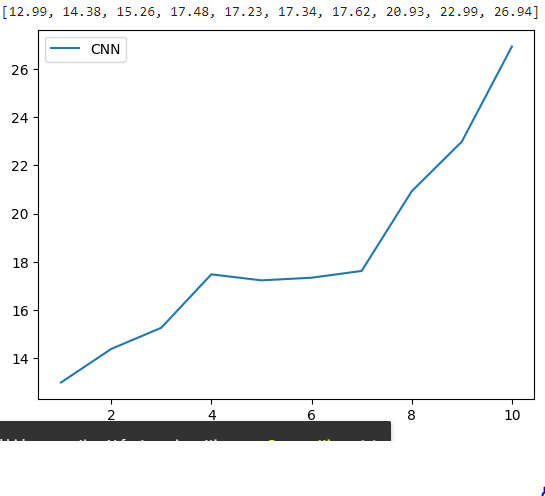


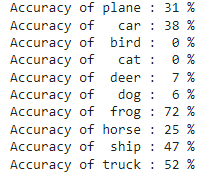
64:





256:

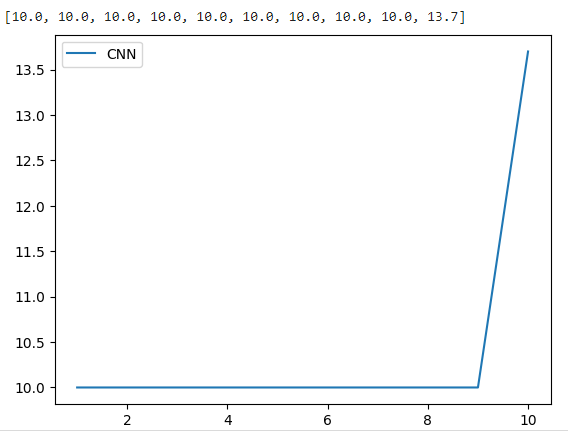


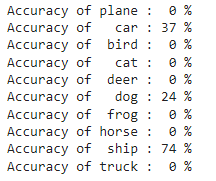


What I noticed for these graphs was that the higher the batch size, the more accurate the photos were, and a stronger linear progression. These means that batch size of 256 shows linear progression with less sharp linearities compared to a batch size of 64, while 16 and 4 both shows sharp linear progression straight from the beginning. 256 was a lot easier on my gpu as well and was more consistent with the accuracies compared to the rest.

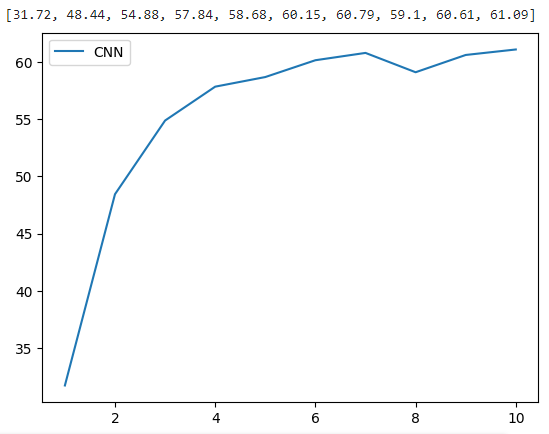
4.

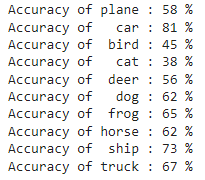
Sigmoid





Relu





It seems like sigmoid produces fundamentally different results than relu. That being, one shows linear progression while the other shows a sharp linear progression after constant equal progression. This makes sense because sigmoid is meant to show outputs in the range of [0, 1], while relu shows a subset of neurons, and shows more linear progression which is why the graph produces higher accuracies while sigmoid does not.