(Fall) 2020 Computer Vision CAP 5415

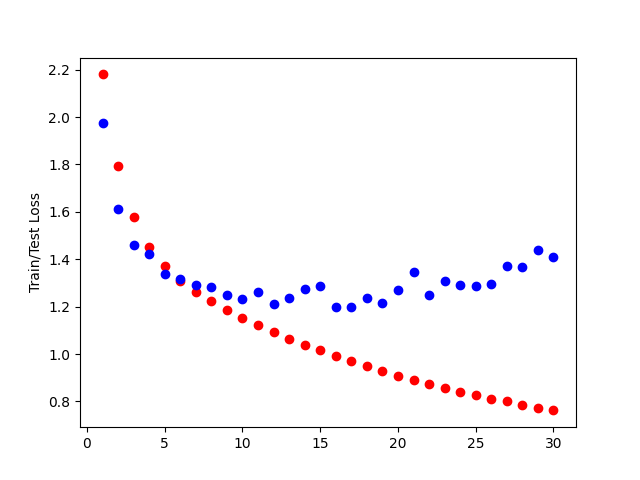
Programming Assignment-III – Report

Chakib Cerny

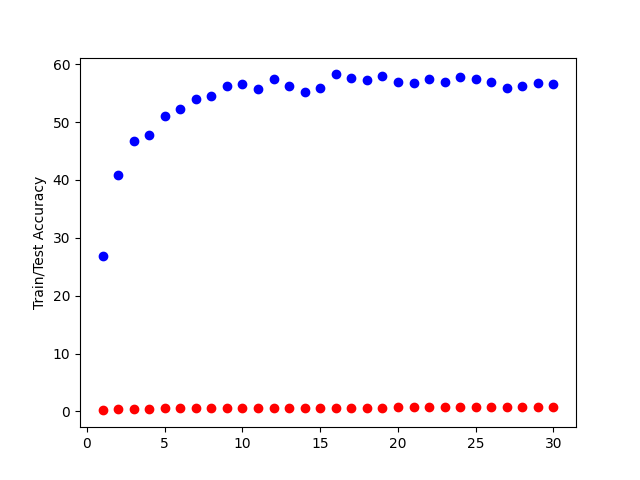
Computer Vision

# Question 1: Image Classification

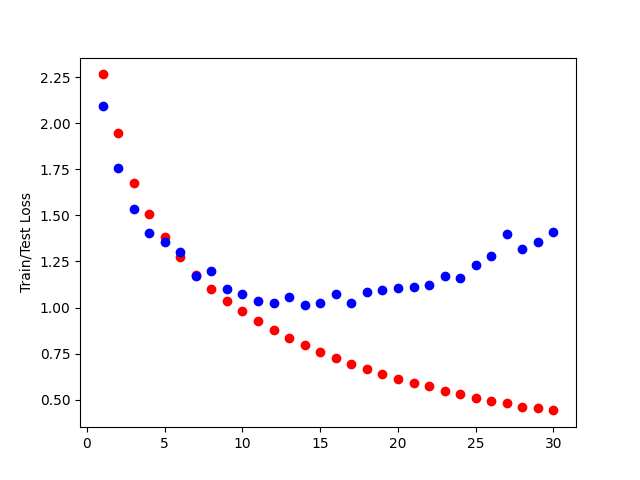
The first implementation consisted of 3 convolution layers and 3 fully connected layers which resulted in the following graphs:



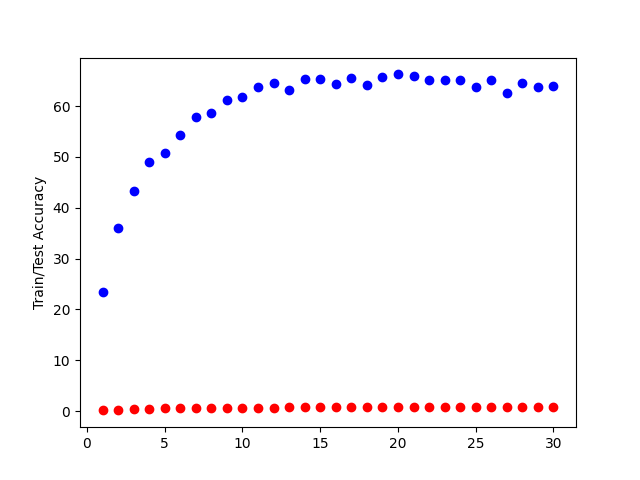
In this graph we can notice that the train lost (red) keeps degreasing and it seems that it did not reach its full potential by the 30th epoch. The assumption is that if we have more epochs or higher the learning rate, we could get even better result.



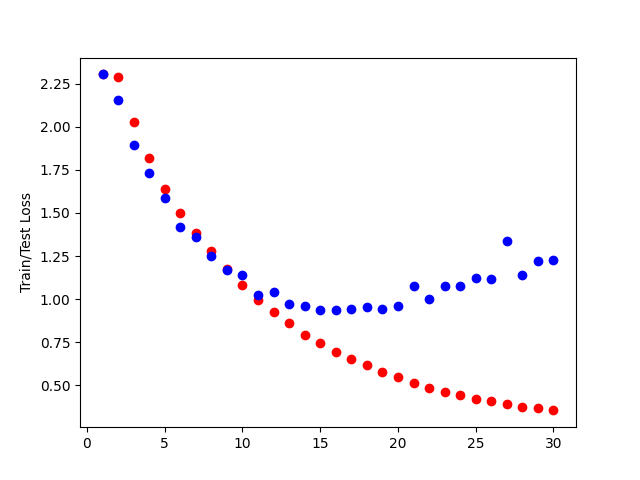
The test accuracy (blue) shows a result averaging 57% which, looking at the first graph could be higher if we are able to get lower train loss by having a higher learning rate.



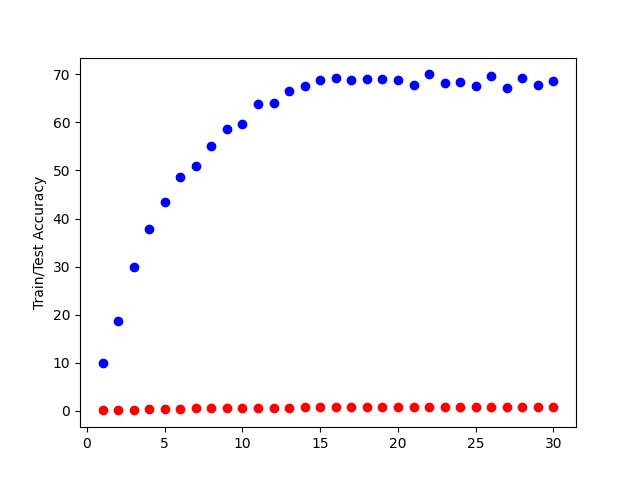
In this case, we added two more convention layers and we can notice that we have a better train loss (red) as it is dropping below 0.5 but as the first graph we can notice that the curb didn’t flatten yet so a higher learning could eventually give us an even better loss.



Similarly to the loss, the test accuracy is better as it passes the 60% mark. We can conclude that adding convolutional layers help with our CNN and on the next run we will add two more layers.



We have now added two more convolutional layers and again the train loss in red looks better but the improvement is not as obvious as the first time we added layers.



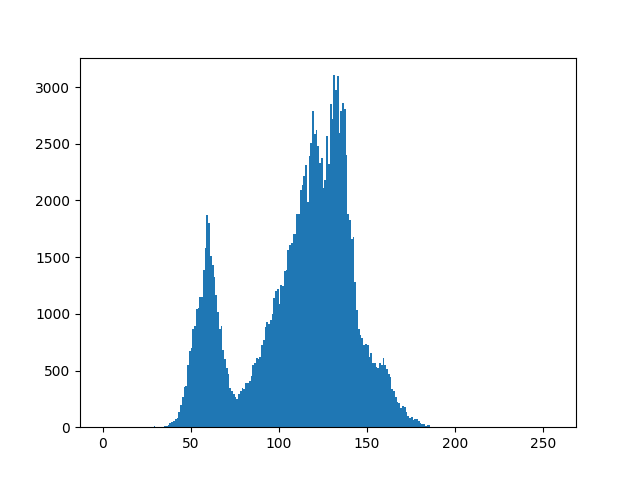
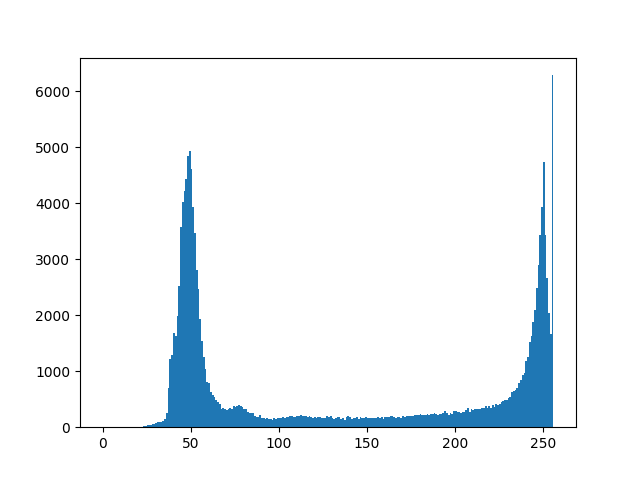
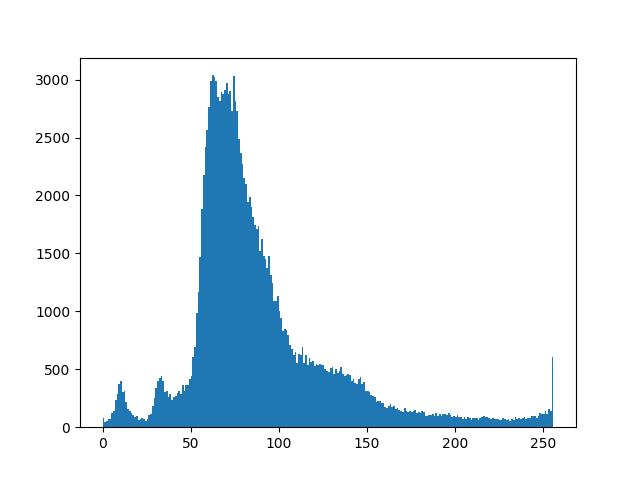
The accuracy is also getting better when adding even more convolutional layers as we are now getting closer to 70%. We can then conclude that adding layers is beneficial to our network which makes as it is adding more features. However, I notices that the GPU utilization was getting higher as we added convolutional layers which mean that the computation requires more power and we are dealing with more features.

# Question 2: Image segmentation

The 3 pictures I chose for this question are the following:

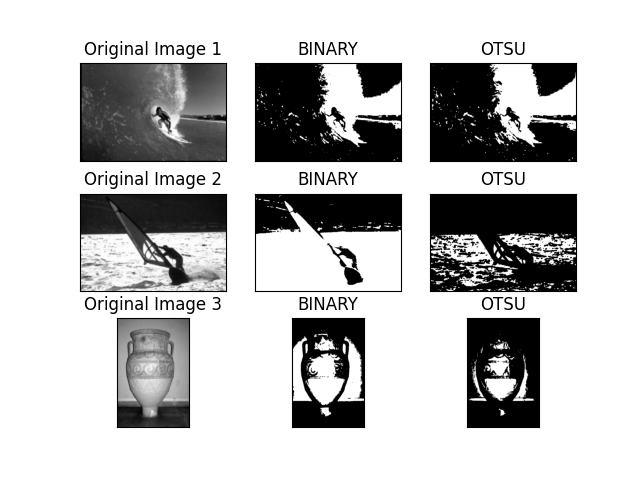


Which gives us the following graphs:



The first Threshold chosen for the binary threshold is right after the first high values, at 100, for the second given the big gap between the two max I decided to go for the end of the first peak at 60 and for the last one, I first picked a value around 75 but realized that the first peak represents the low luminosity of the floor and not the vase so I picked a value between the two middle peak at 126.

I then used the Otsu formula for the second part of the question and obtained the following results:



Looking at the result we can tell that the binary gives a better result for the 2nd and 3rd picture which can be explained by the fact that I also tried 2-3 different thresholds to get this result, I first got tricked by the floor on the 3rd picture, which would almost make the vase disappear and for the second picture we can notice that the water and the variation in it which cover the middle part on the histogram is what causes the issue.