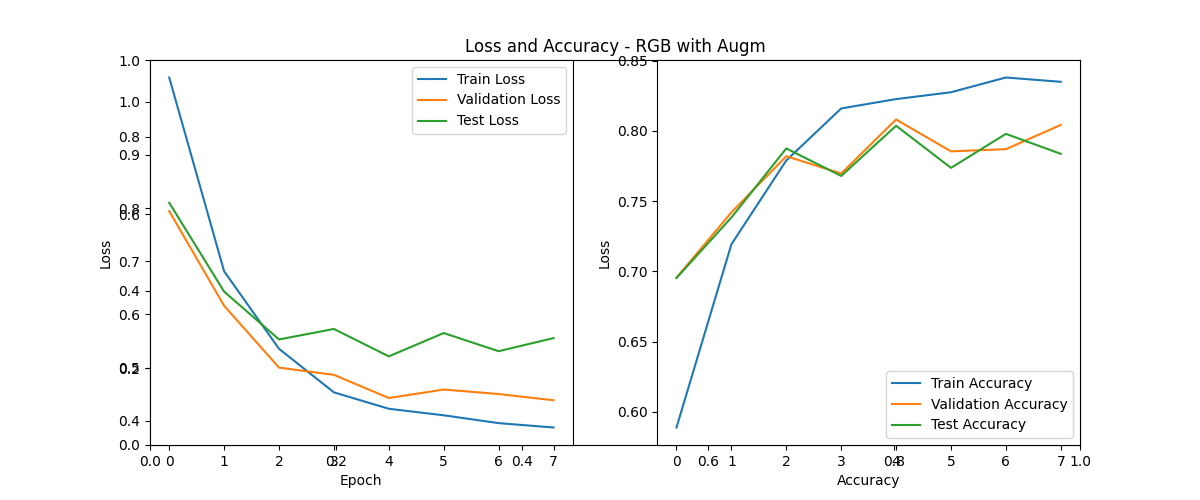
Task 4 results:

I implemented video generation for training dataset. Then I tested CNN models in the gymnasium environment by modifying the car\_racing script. However, my models never managed to do turning in the environment.

My dataset consist of 13000 images. For models I tried different combinations:

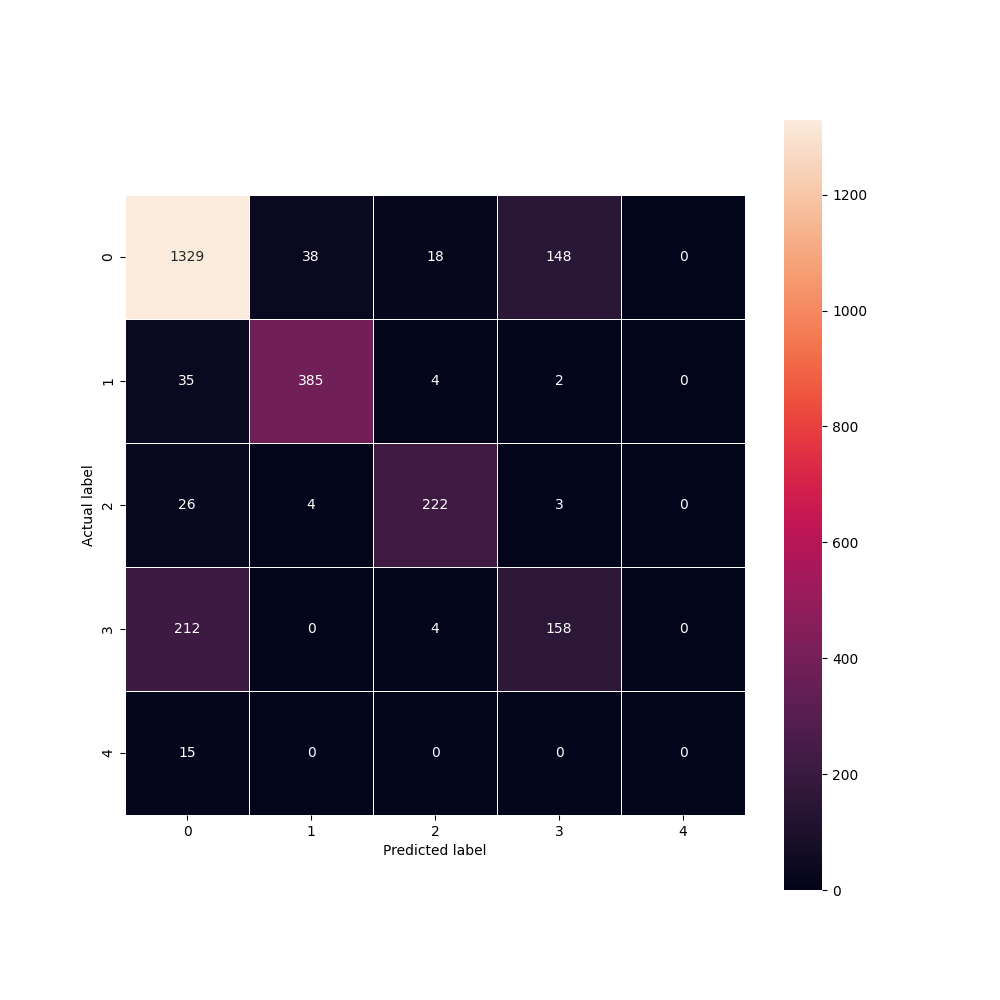
* With or without augmentations
* With or without weighted cross entropy loss (weights proportional to the number of samples or custom weights)
* With smaller part of the dataset
* With different handling of classes where 2 actions are pressed at the same time (do nothing or just turn and ignore gas/breaking)

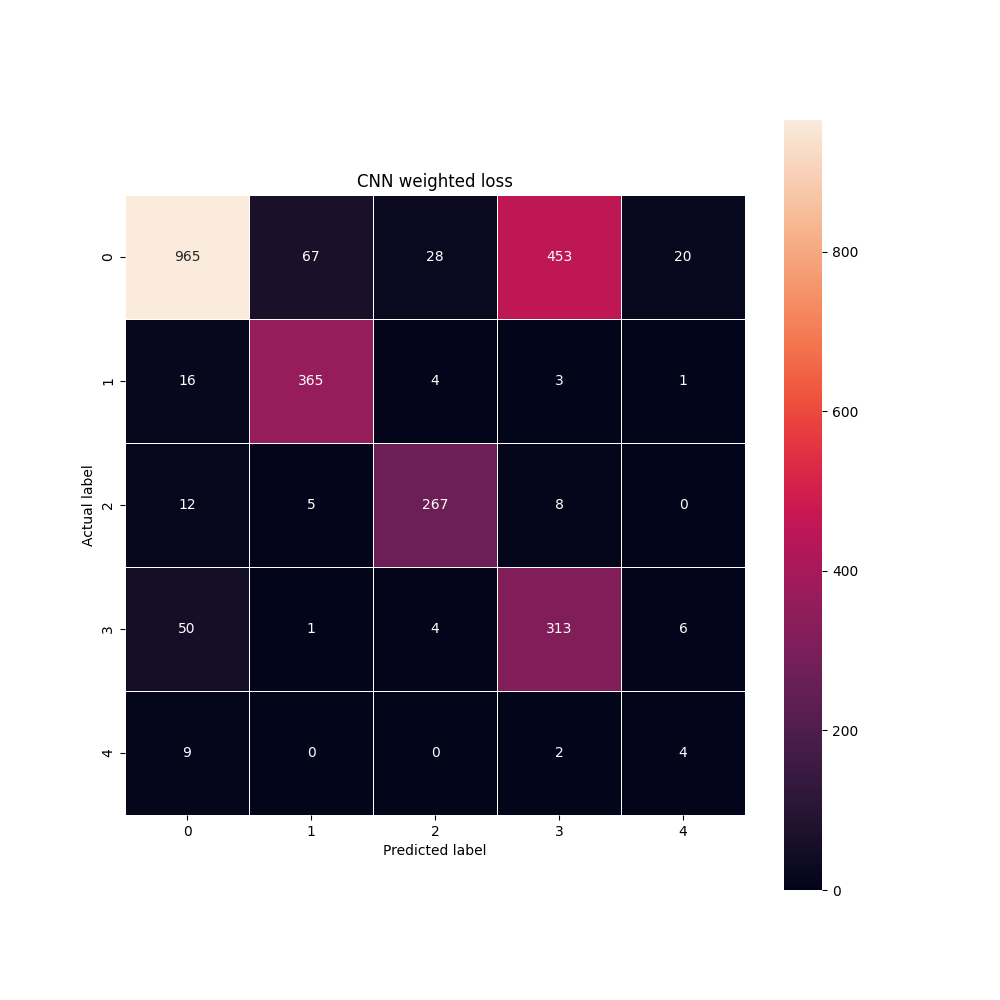
Same situation happened when using RNN. You can see the results on following graphs:

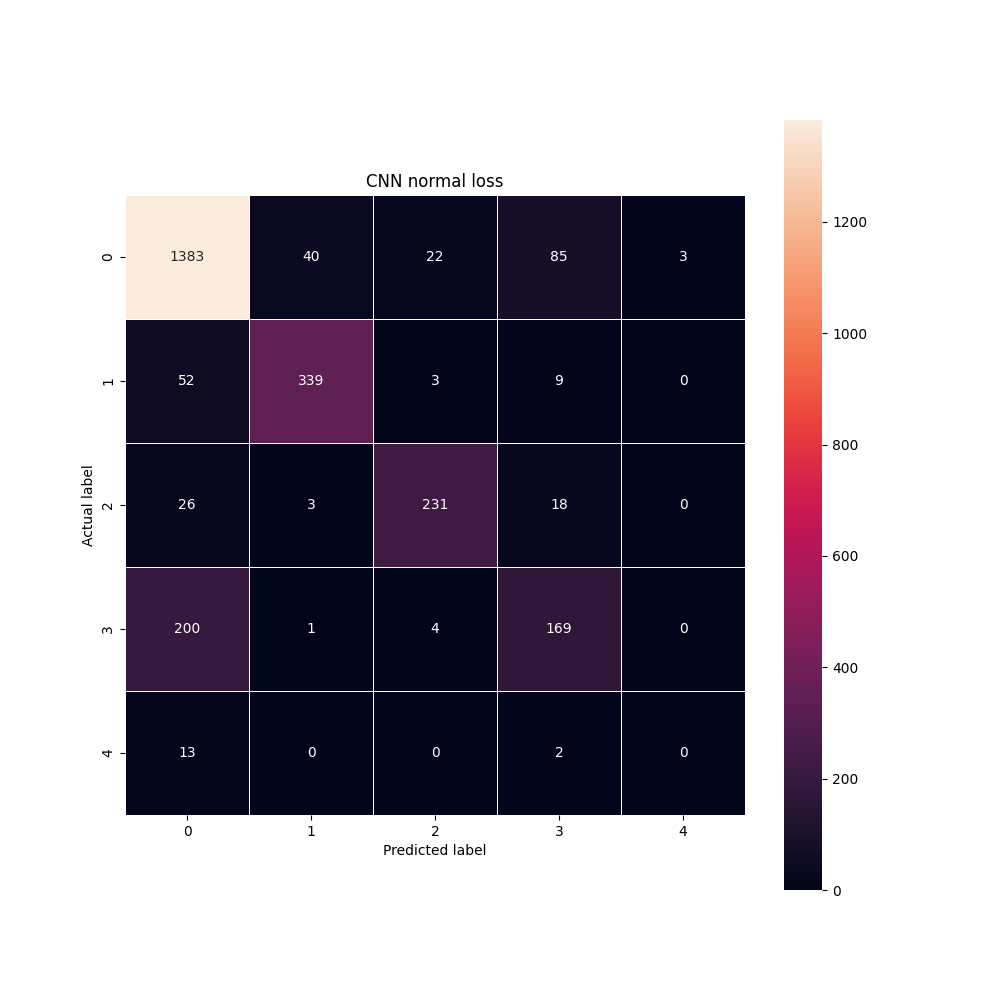


Training results of RNN

RNN confusion matrix:







From confusion matrices, I hypothesize that the model doesn’t learn how to do few of the turns on the track. We can see that all of the models predict 0 (do nothing) instead of 1 (left) or 2 (right) sometimes. Unfortunately, model probably doesn’t learn how to do the first turning so it never ends up showing what it actually learns.

It just goes straight and occasionally uses gas (command 3).

In the attachment are all of the models and scripts I used for training.