Image processing:

The image processing part includes crops, flips (horizontally) and more, but we are still unsure about the appropriate processing needing to be done and how helpful it can be. Also, we reshape our images to size 64(so far), this is still a work in progress and might be changed eventually.

Data information:

Our dataset contains over one thousand vehicle and non-vehicle images, but we only trained using 150 of our images (total). We plan to start using the whole dataset after we are confident about the parametrization of our neural network.

For the dataset we use an equal number of images of vehicles and non-vehicle. We use 80% of our dataset for training and 20% for testing, maintaining the 50-50 ratio of vehicle and non-vehicle images for both training and testing.

Neural network:

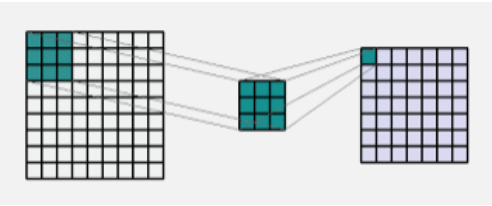
We created 14 Unit layers with max pooling between some layers and an average pool at the end:

3 Unit -> MaxPool -> 4 Unit -> MaxPool -> 4 Unit -> MaxPool -> 3 Unit -> AvgPool ->output

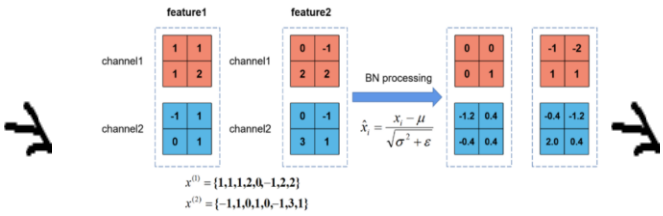
Each Unit layer applies a convolution to its input, a batch normalization and then a ReLU function to compute the output.

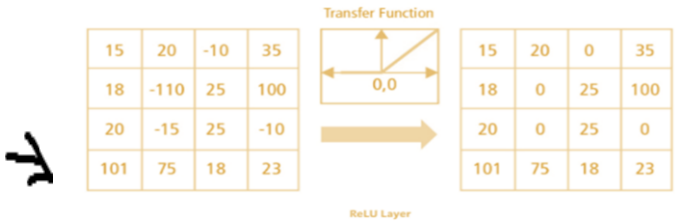
Each Unit layer basically does:

CONVOLUTION



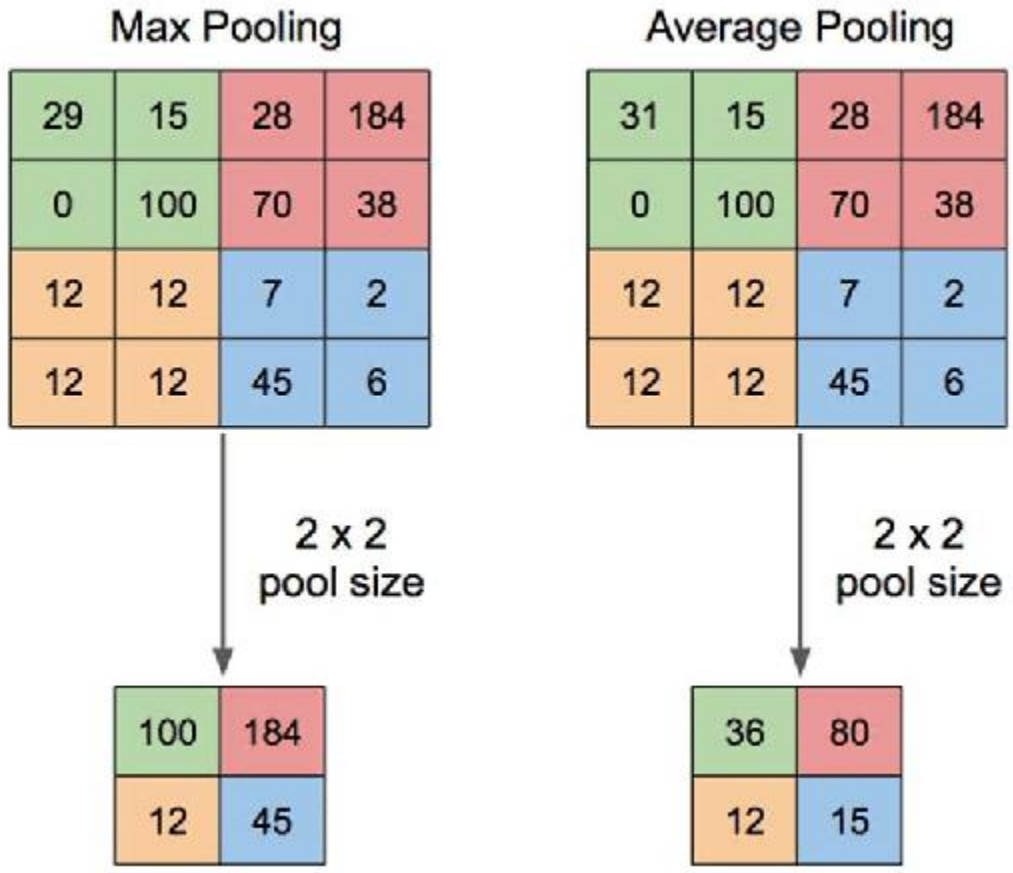
Normalization





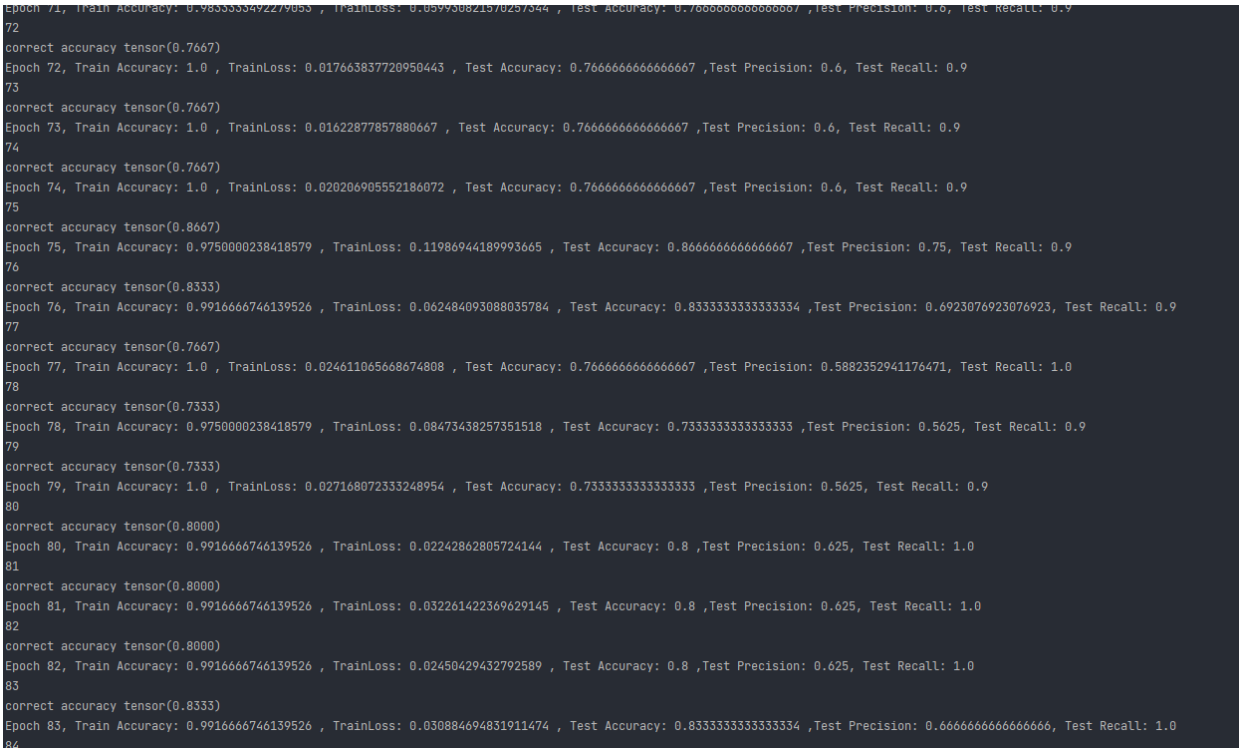
ReLU

A Max Pooling takes the maximum value for each window selected from the input, and the Avg Pooling computes the average from the window selected:

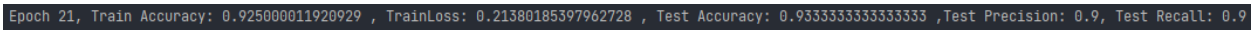


Performance:

We computed the accuracy, precision and recall, here are some results:



One of the best results:



NOTE: This has been done one a dataset of 150 images (total), using an image size of 32x32.