Aprendizagem TPC 5

João Marques 58722 Daniel Oliveira 87848

—Link to code—

Dividing the Image Set

We chose to keep the train set originally given. It seems to have a good variety of the various fruits in different angles and we saw no reason to alter it.

We obtained the validation and test sets by randomly splitting the original test set. The original test set has several sequences that are very similar with slight rotations, some very similar to the train sequences, others with fruits that are slightly different. Since there is some repetition it was expected (and observed) that randomly splitting it roughly puts half of each sequence on each side, making the sets similar but with a certain variation added by the random split.

CNN Architecture

Our final convolution neural network is composed by 1 layer with 20 filters with dimensions 12×12 then a 2×2 max-pool layer, followed by another layer with 50 filters with dimensions 12×12 and then another 2×2 max-pooling layer. After the convolution part of our CNN we end up with 50 features each one with 25×25 pixels. These features are then flattened into a 31250×1 tensor that is fed into our ANN. The Ann is composed by 31250 neurons in the input layer, 256 neurons in the hidden layer and 4 neurons in the output layer, as it is shown in figure 1.

Note that we did not use any filters downloaded from the internet. We trained our own filters using CUDA in a NVIDIA GeForce GTX 745 GPU.

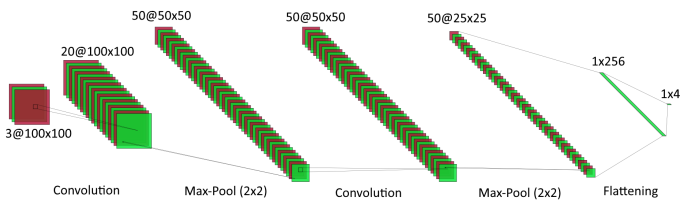

Figure 1: CNN Architecture

Comparing filter size

Three variations of the neural network were trained using xavier_glorot initialization and dropout regularization: The first one uses 3 by 3 filters, the second one 6 by 6 filters and the last one 12 by 12 filters.

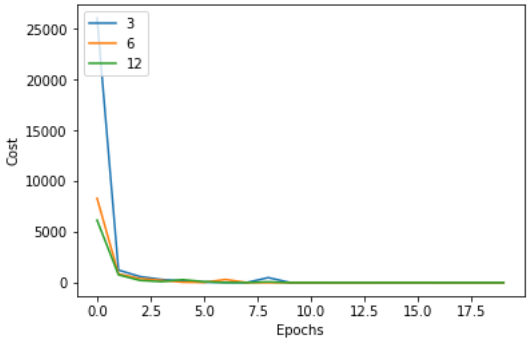

Figure 2: Cost progression for different filter sizes

As can be seen in figure 2, the cost starts out higher the smaller the filter number, but they all converge to near 0.

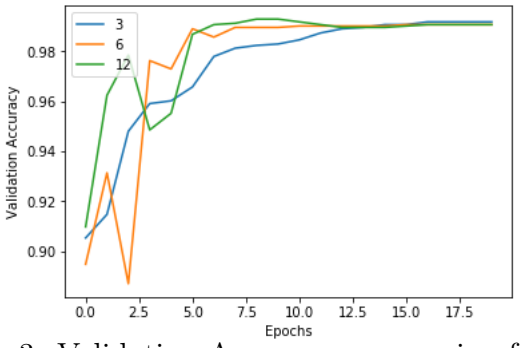

Figure 3: Validation Accuracy progression for different filter sizes

In figure 3 we see that, after some initial variations during the early epochs, the tendency for the higher number of filters to achieve a better validation accuracy asserts itself. The peak for 12 by 12 seems to be around 9 epochs, with the decrease after that presumably reflecting the network starting to overfit.

The best filter size is 12 by 12, however 3 by 3 did not perform much worse and is a lot easier to train, therefore the following comparisons will use 3 by 3. The final model will use 12 by 12.

Comparing initialization

Three variations of the neural network were trained using dropout regularization and 3 by 3 filters - as mentioned previously: The first one uses xavier_glorot[1] initialization, the second initializes with zeros, and the third with values from a random normal distribution.

[1] glorot10a.pdf

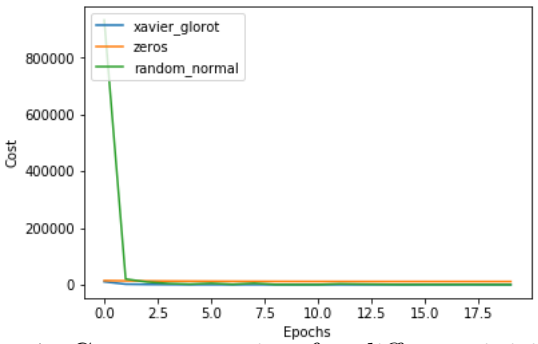

Figure 4: Cost progression for different initialization methods

The cost starts much higher for random normal but quickly converge to near 0 alongside xavier glorot, as shown in figure 4. Zeros on the other hand never goes below 10 000.

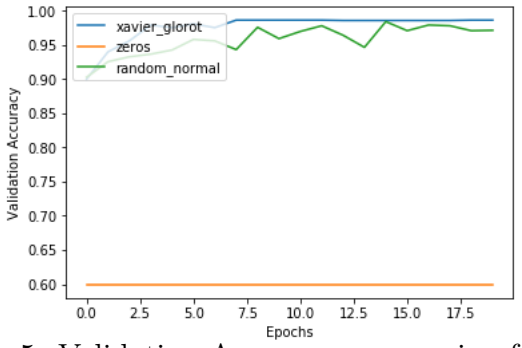

Figure 5: Validation Accuracy progression for different initialization methods

The validation accuracy is always higher for xavier glorot, as per figure 5, it is also more stable. The zeros initialization cannot learn since we are using ReLU as the activation function.

The superior initialization method is xavier glorot for this neural network.

Comparing regularization

Four variations of the neural network were trained using xavier_glorot initialization and 3 by 3 filters - as mentioned previously.The first one uses no regularization, and the other three use dropout, l1 and l2 regularization respectively.

In figure 6 we see that costs all have similar behaviour, tending to near 0, except for l1 that never goes below 600.

Figure 7 shows that after 7 epochs the validation accuracy is fairly close for all methods, with l2 slightly edging about the rest and dropout right behind it. l1 seems to be worse than having no regularization.

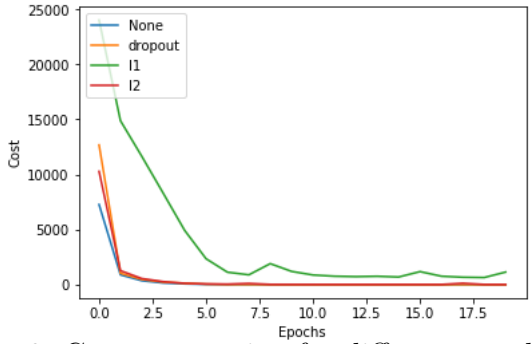

Figure 6: Cost progression for different regularizations

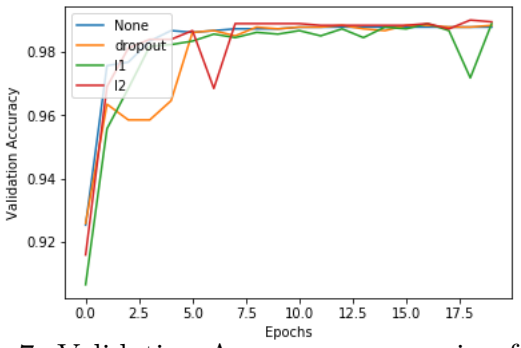

Figure 7: Validation Accuracy progression for different regularizations

Final parameters

For the final parameters we picked 12 by 12 filters, xavier glorot initialization and l2 regularization. The Neural Network was trained for 9 epochs, since that seems to be the beginning of the overfit for 12 by 12 filters.

The final validation accuracy in the final epoch was 0.992, with an epoch error of 18. The accuracy with the test set is slightly lower at 0.978.

Figure 8: Incorrectly classified examples

As was expected the failures to classify were between pears and apples, which in some angles look remarkably similar. In the specific network trained for these results, 38 pears were classified as apples and 1 apple was classified as a pear.

Some examples can be seen in figure 8. The lack of pixels in the apple is probably what lead to the error. Pear 1 is the rarer case (2 of them) where we could still tell that it is a pear, pear 2 is the most common type or error, a pear from a bottom angle. We would classify pear 2 as an apple so it is hardly surprising that the neural network does so as well.