Network Architecture

- three 2D convolutional layers with 32 filters, a 3 × 3 kernel, and ReLU activation
- 2 × 2 max pooling layer
- four 2D convolutional layers with 64 filters, a 3 × 3 kernel, and ReLU activation
- 4 × 4 average pooling layer
- linear layer

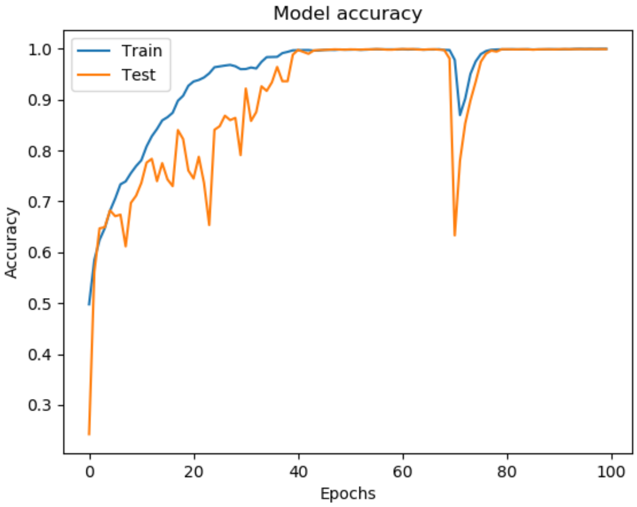

Figure 1: Train and test accuracy over number of epochs.

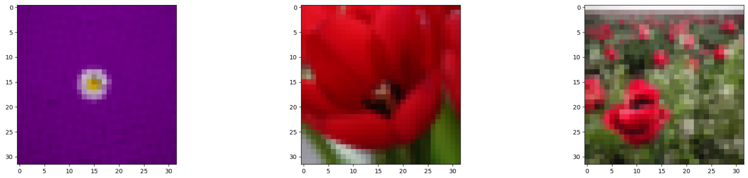

Figure 2: Misclassified examples.

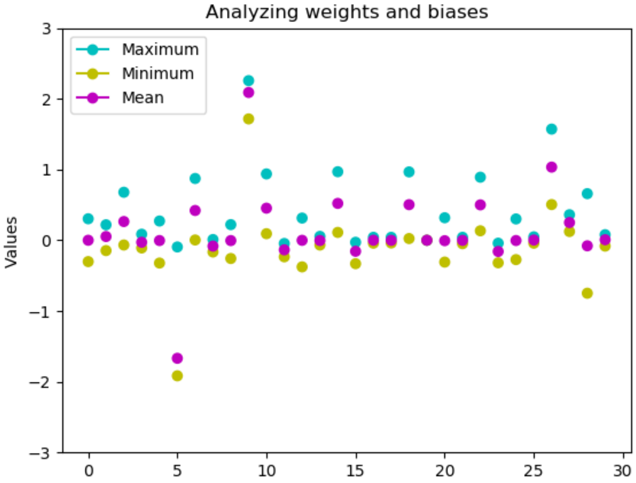

Figure 3: Final weights and biases of trained model. There are three colored points along each x-coordinate because these are the values associated with a parameter: maximum, minimum, and mean. These values indicate the largest, smallest, and average of values in the weight matrix/ bias vector.

Performance The model is trained for 100 epochs. The train accuracy reached 100% while the test accuracy reached 99.9% (Figure 1). The mean for both bias vectors and weight matrices are around 0, but the values themselves are usually between -1 and 1 (Figure 3). This network took about 10 minutes to train on a GPU.

Figure 2 shows misclassified examples in the test set. The left image is supposed to be labeled “daisy”, but our model predicts a “tulip” instead. The center image is labeled “tulip”, but our model predicts a “rose” instead. The right image is labeled “tulip”, but our model predicts a “rose” instead. These images are blurry and some tulips really look like roses.

Explanation The motivation for using a CNN is that we are classifying an image dataset. It turns out that the CNN works really well for the Flowers dataset. The convolutional layer can extract features from images through convolving with the filter. Then, the pooling layer can reduce dimensionality of the feature map and dropout can prevent overfitting. The ReLU activation function can introduce nonlinearity to the model. Average pooling takes the mean of all values in a block. I found it to increase accuracy because classifying flowers needs to take all information about image pixels into account.

The loss function is cross-entropy. The advantage of this loss is that the network can learn faster if error is higher. For a large dataset like Flowers, this is important because learning can take very long until the network converges.