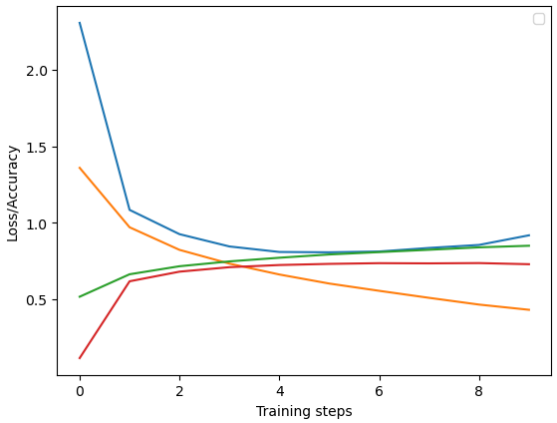
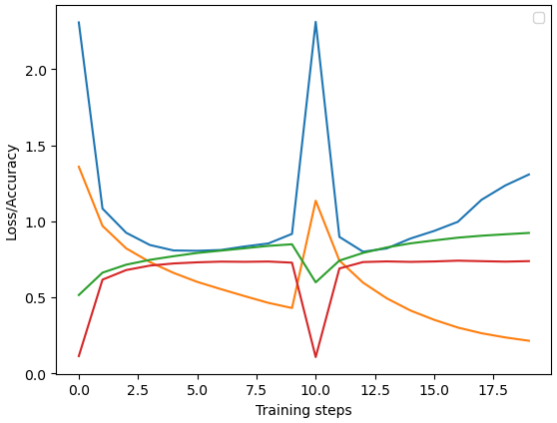
Initial settings:
Final test_accuracy = 0.728
Final train_accuracy = 0.849
Final test_loss = 0.917
Final train_loss = 0.430



Change 1:
Augment the data with flipping and rotation
Why? More data points should help the model generalize.

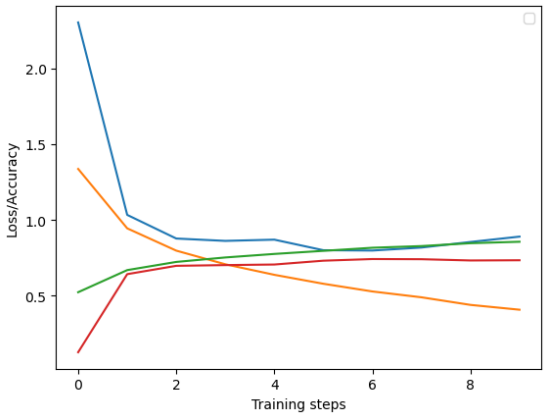
Final test_accuracy = 0.738
Final train_accuracy = 0.923
Final test_loss = 1.307
Final train_loss = 0.215



Dit it work? I helped a bit with the accuracy, although the loss is higher. Also, there seems to have been a problem while training. The exact reason remains unclear to us.

Change 2:
Add L2 regularization penalty
Why? Reduces the fluctuations in the coefficients.

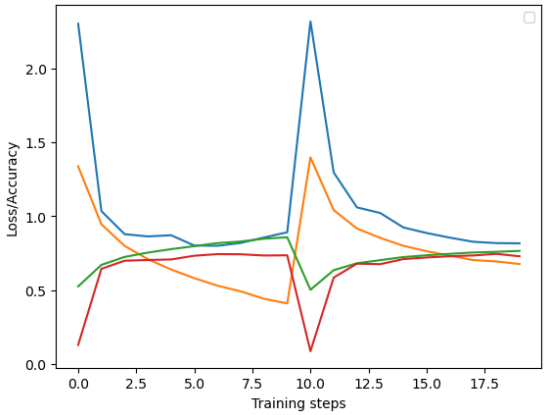
Final test_accuracy = 0.735
Final train_accuracy = 0.856
Final test_loss = 0.890
Final train_loss = 0.409



Did it work? It helped with both test loss and test accuracy. Still the test and train loss diverged in the later epochs.

Change 3:
Adding dropout layers before every pool
Why? Deletion of some weights should help with generalization as less information can be stored in the weights themselves.

Final test_accuracy = 0.728
Final train_accuracy = 0.764
Final test_loss = 0.815
Final train_loss = 0.676



Did it work? Yes, it did! The test loss and train loss are actually quite similar. In combination with other techniques, this might be very useful.