

[Q1]

The correct kernel size in the average pooling layer is 16×16 , since the input is $1 \times 32 \times 32$, and after the first two convolution layers, it will become $32 \times 32 \times 32$, then after the max pooling layer, it will become $32 \times 16 \times 16$, then after the four convolution layers, it will end up $512 \times 16 \times 16$, hence, in order to achieve 1×512 at the last, the kernel size should be $16/1 = 16$.

[Q2]

Dropout is a regularization technique for neural network models, where randomly selected neurons are ignored (dropped-out) while training. This means that their contribution to the activation of downstream neurons is temporarily removed on the forward pass and any weight updates are not applied to the neuron on the backward pass. The effect is that the network becomes less sensitive to the specific weights of neurons. This in turn results in a network that is capable of better generalization and is less likely to overfit the training data.

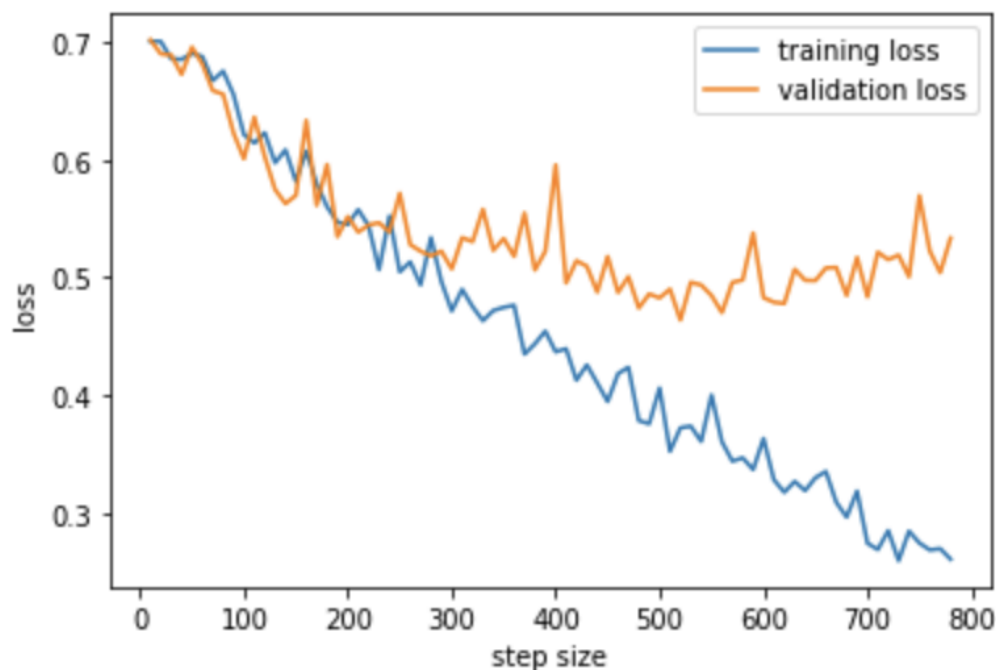
[Q3]

1.842465M (million)

[Q4]

The total number of trainable parameters will increase compared to the original settings if concatenation is used.

[Q5]



According to the plot, I should stop training at around step 520, since after that the training loss has kept decreasing but the validation loss has an increasing trend.

[Q6]

Optimal threshold : 0.477

Validation accuracy : 0.7957

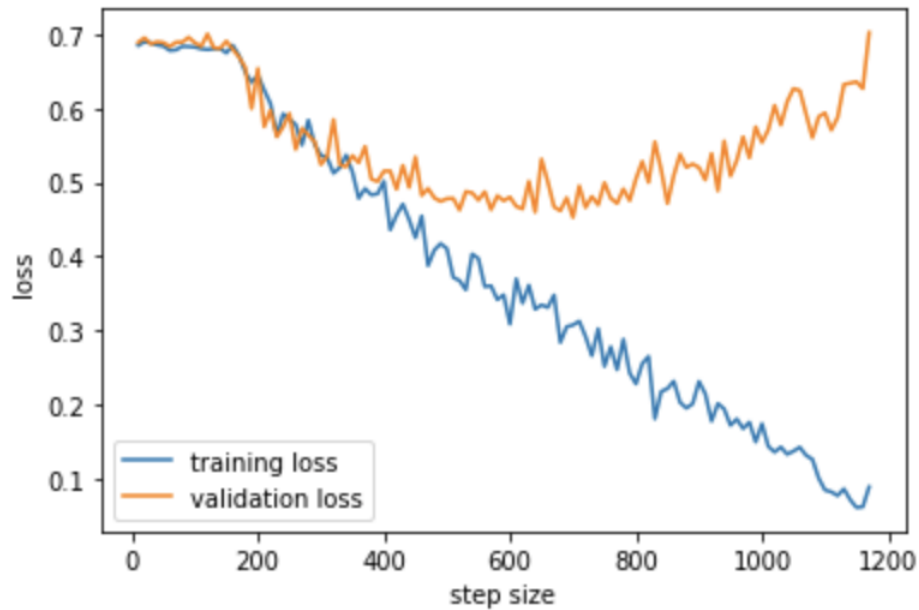
[Q7]

You have to set a relatively high threshold for each of the applications, since their faces might look similar, so we have to set a higher threshold to make sure that it can recognize the right person. If the threshold is low, it may occur that two people with similar faces will be recognized as the same person.

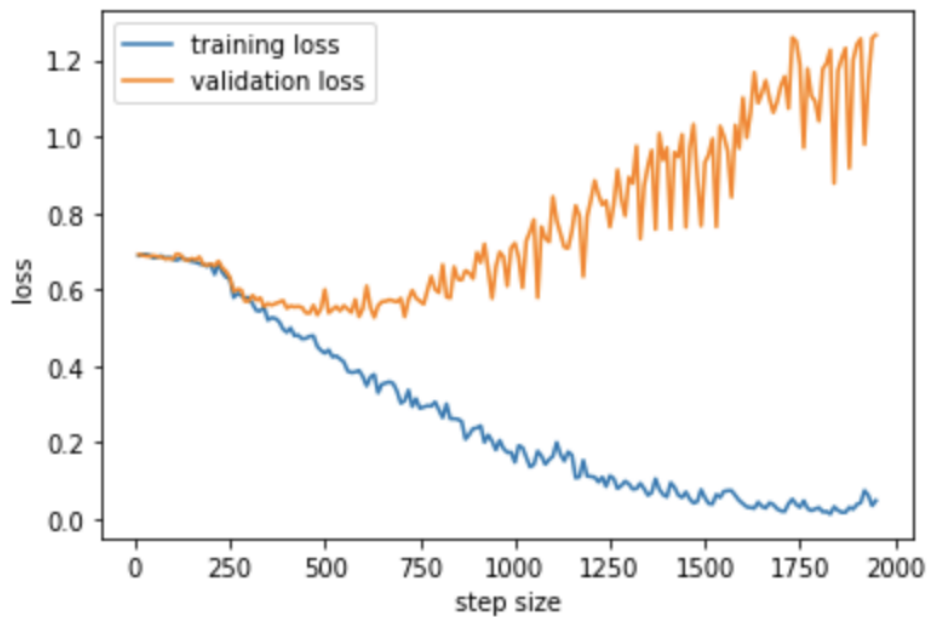
[Q8]

A.

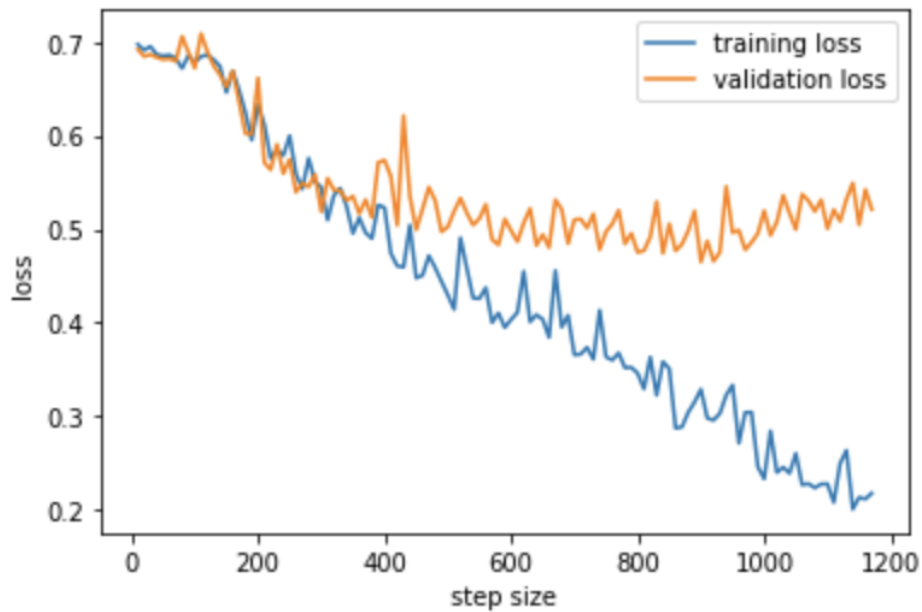
1. Change the number of epochs to 30, and change the learning rate to 0.0005
Validation accuracy : 0.7986 (better than the original settings)



2. Change the number of epochs to 50, and set the learning rate to 0.001
Validation accuracy : 0.74 (lower than the original settings)



3. Change the number of epochs to 30, change the learning rate to 0.001, and change the weight decay to 0.00001
Validation accuracy : 0.8028 (higher than the original settings)



C.

1. The two implementation methods that I added are RandomHorizontalFlip() and RandomRotation(10). The reason I think these two implementation methods are going to improve the comparison task is that by horizontal flipping and random rotating 10 degrees, the model would learn more about the face, which means the model would get more circumstances of the face rather than just a regular right angled picture, and the next time the model meets a face that is not at the right position, say crooked or upside down, it may still be able to differentiate who it is correctly.

Validation accuracy : 0.8 (better than the original settings)

