Summary of the improved architecture:

Model: "sequential_126"

_____

Layer (type)   Output Shape   Param #

==

conv2d_603 (Conv2D)   (None, 26, 26, 32)   320

_____

conv2d_604 (Conv2D)   (None, 24, 24, 128)   36992

_____

max_pooling2d_297 (MaxPoolin (None, 12, 12, 128)   0

_____

conv2d_605 (Conv2D)   (None, 8, 8, 128)   409728

_____

conv2d_606 (Conv2D)   (None, 6, 6, 64)   73792

_____

max_pooling2d_298 (MaxPoolin (None, 3, 3, 64)   0

_____

conv2d_607 (Conv2D)   (None, 1, 1, 64)   36928

_____

flatten_73 (Flatten)   (None, 64)   0

_____

dense_158 (Dense)   (None, 32)   2080

_____

dense_159 (Dense)   (None, 10)   330

==

Total params: 560,170

Trainable params: 560,170

Non-trainable params: 0

_____

Model performance:

     Accuracy   Loss

Train 0.943733 0.154034

Test   0.914300 0.256925