# ECS6P9U/P: NEURAL NETWORKS & DEEP LEARNING

**Task 4**

**Write the training script to train the model. Provide in the report:**

- The curves for the evolution of loss.
- The curves for the evolution of training and validation (test) accuracies.
- All training details including hyper-parameters used.

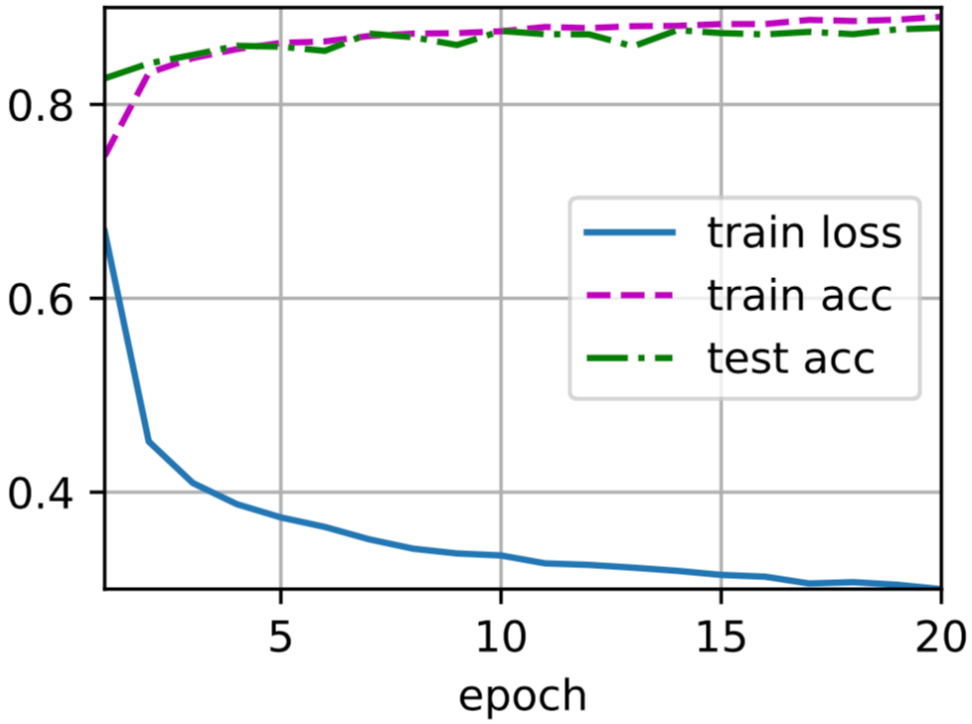
**Training details including hyper-parameters used**

**Try 1 – patch_size(4,4 ), learning_rate = 0.01, epochs = 50, batch_size=256**

```
Epoch   1 | Train: loss=0.666, acc= 75.0% | Test: loss=0.467, acc= 83.4%
Epoch   2 | Train: loss=0.448, acc= 83.4% | Test: loss=0.444, acc= 83.8%
Epoch   3 | Train: loss=0.414, acc= 84.7% | Test: loss=0.403, acc= 85.2%
Epoch   4 | Train: loss=0.379, acc= 86.0% | Test: loss=0.391, acc= 85.6%
Epoch   5 | Train: loss=0.370, acc= 86.5% | Test: loss=0.393, acc= 85.7%
Epoch   6 | Train: loss=0.358, acc= 86.8% | Test: loss=0.399, acc= 85.3%
Epoch   7 | Train: loss=0.344, acc= 87.4% | Test: loss=0.367, acc= 86.7%
Epoch   8 | Train: loss=0.345, acc= 87.2% | Test: loss=0.363, acc= 87.1%
Epoch   9 | Train: loss=0.333, acc= 87.7% | Test: loss=0.365, acc= 87.5%
Epoch  10 | Train: loss=0.329, acc= 87.9% | Test: loss=0.354, acc= 87.3%
Epoch  11 | Train: loss=0.325, acc= 88.0% | Test: loss=0.360, acc= 86.9%
Epoch  12 | Train: loss=0.321, acc= 88.0% | Test: loss=0.349, acc= 87.4%
Epoch  13 | Train: loss=0.315, acc= 88.2% | Test: loss=0.334, acc= 88.0%
Epoch  14 | Train: loss=0.313, acc= 88.4% | Test: loss=0.353, acc= 88.1%
Epoch  15 | Train: loss=0.310, acc= 88.6% | Test: loss=0.362, acc= 87.4%
Epoch  16 | Train: loss=0.303, acc= 88.8% | Test: loss=0.331, acc= 87.9%
Epoch  17 | Train: loss=0.302, acc= 88.8% | Test: loss=0.347, acc= 87.6%
Epoch  18 | Train: loss=0.304, acc= 88.7% | Test: loss=0.341, acc= 87.9%
Epoch  19 | Train: loss=0.300, acc= 88.8% | Test: loss=0.330, acc= 88.4%
Epoch  20 | Train: loss=0.297, acc= 89.0% | Test: loss=0.335, acc= 88.2%
Epoch  21 | Train: loss=0.293, acc= 89.1% | Test: loss=0.336, acc= 87.8%
Epoch  22 | Train: loss=0.296, acc= 88.9% | Test: loss=0.351, acc= 87.6%
Epoch  23 | Train: loss=0.292, acc= 89.1% | Test: loss=0.351, acc= 88.3%
Epoch  24 | Train: loss=0.288, acc= 89.2% | Test: loss=0.327, acc= 88.6%
Epoch  25 | Train: loss=0.290, acc= 89.2% | Test: loss=0.340, acc= 88.2%
Epoch  26 | Train: loss=0.290, acc= 89.3% | Test: loss=0.318, acc= 88.5%
Epoch  27 | Train: loss=0.285, acc= 89.3% | Test: loss=0.359, acc= 87.7%
Epoch  28 | Train: loss=0.280, acc= 89.6% | Test: loss=0.323, acc= 88.6%
Epoch  29 | Train: loss=0.281, acc= 89.4% | Test: loss=0.330, acc= 88.3%
Epoch  30 | Train: loss=0.283, acc= 89.5% | Test: loss=0.347, acc= 87.6%
Epoch  31 | Train: loss=0.278, acc= 89.5% | Test: loss=0.334, acc= 88.1%
Epoch  32 | Train: loss=0.277, acc= 89.7% | Test: loss=0.336, acc= 88.5%
Epoch  33 | Train: loss=0.276, acc= 89.7% | Test: loss=0.323, acc= 88.8%
Epoch  34 | Train: loss=0.277, acc= 89.6% | Test: loss=0.334, acc= 88.6%
Epoch  35 | Train: loss=0.278, acc= 89.7% | Test: loss=0.333, acc= 88.7%
Epoch  36 | Train: loss=0.271, acc= 90.0% | Test: loss=0.334, acc= 88.3%
Epoch  37 | Train: loss=0.277, acc= 89.9% | Test: loss=0.332, acc= 88.7%
Epoch  38 | Train: loss=0.274, acc= 89.8% | Test: loss=0.346, acc= 88.2%
Epoch  39 | Train: loss=0.273, acc= 89.9% | Test: loss=0.328, acc= 88.4%
Epoch  40 | Train: loss=0.271, acc= 89.8% | Test: loss=0.320, acc= 88.8%
Epoch  41 | Train: loss=0.268, acc= 90.0% | Test: loss=0.310, acc= 89.0%
Epoch  42 | Train: loss=0.267, acc= 90.0% | Test: loss=0.313, acc= 89.0%
Epoch  43 | Train: loss=0.268, acc= 90.0% | Test: loss=0.335, acc= 88.8%
Epoch  44 | Train: loss=0.265, acc= 90.1% | Test: loss=0.329, acc= 88.3%
Epoch  45 | Train: loss=0.266, acc= 90.1% | Test: loss=0.311, acc= 88.8%
Epoch  46 | Train: loss=0.266, acc= 90.0% | Test: loss=0.319, acc= 88.8%
Epoch  47 | Train: loss=0.262, acc= 90.2% | Test: loss=0.315, acc= 89.0%
Epoch  48 | Train: loss=0.264, acc= 90.2% | Test: loss=0.335, acc= 88.4%
Epoch  49 | Train: loss=0.263, acc= 90.2% | Test: loss=0.345, acc= 88.3%
Epoch  50 | Train: loss=0.260, acc= 90.1% | Test: loss=0.322, acc= 88.8%
```

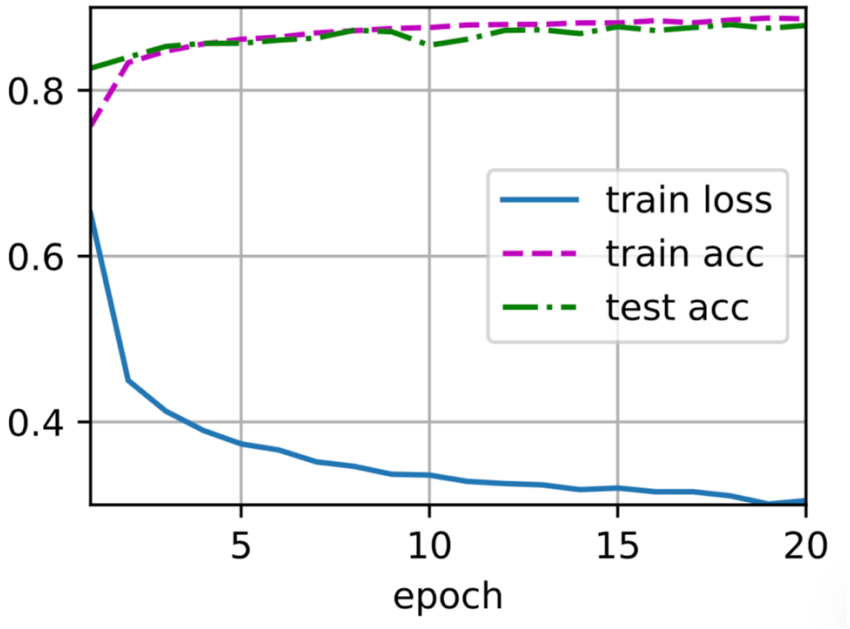
**Try 2 – patch_size(4, 4), learning_rate = 0.001, epochs=20, batch_size=256**

```
Epoch   1 | Train: loss=0.942, acc= 64.6% | Test: loss=0.534, acc= 80.8%
Epoch   2 | Train: loss=0.516, acc= 80.9% | Test: loss=0.471, acc= 83.0%
Epoch   3 | Train: loss=0.459, acc= 82.9% | Test: loss=0.449, acc= 84.0%
Epoch   4 | Train: loss=0.430, acc= 84.0% | Test: loss=0.420, acc= 84.8%
Epoch   5 | Train: loss=0.405, acc= 84.9% | Test: loss=0.403, acc= 85.4%
Epoch   6 | Train: loss=0.391, acc= 85.5% | Test: loss=0.403, acc= 85.2%
Epoch   7 | Train: loss=0.376, acc= 86.0% | Test: loss=0.378, acc= 86.2%
Epoch   8 | Train: loss=0.366, acc= 86.3% | Test: loss=0.373, acc= 86.8%
Epoch   9 | Train: loss=0.357, acc= 86.7% | Test: loss=0.380, acc= 86.0%
Epoch  10 | Train: loss=0.349, acc= 86.9% | Test: loss=0.357, acc= 87.0%
Epoch  11 | Train: loss=0.343, acc= 87.2% | Test: loss=0.354, acc= 87.1%
Epoch  12 | Train: loss=0.337, acc= 87.4% | Test: loss=0.345, acc= 87.8%
Epoch  13 | Train: loss=0.332, acc= 87.6% | Test: loss=0.341, acc= 87.6%
Epoch  14 | Train: loss=0.327, acc= 87.8% | Test: loss=0.344, acc= 87.5%
Epoch  15 | Train: loss=0.322, acc= 88.1% | Test: loss=0.336, acc= 87.9%
Epoch  16 | Train: loss=0.318, acc= 88.2% | Test: loss=0.339, acc= 87.9%
Epoch  17 | Train: loss=0.314, acc= 88.1% | Test: loss=0.332, acc= 88.1%
Epoch  18 | Train: loss=0.311, acc= 88.2% | Test: loss=0.352, acc= 87.3%
Epoch  19 | Train: loss=0.308, acc= 88.4% | Test: loss=0.327, acc= 88.2%
Epoch  20 | Train: loss=0.303, acc= 88.6% | Test: loss=0.335, acc= 87.8%
```




**Try 2 – patch_size(4, 4), learning_rate = 0.01, epochs=20, batch_size=256**

```
Epoch   1 | Train: loss=0.660, acc= 75.1% | Test: loss=0.469, acc= 82.9%
Epoch   2 | Train: loss=0.442, acc= 83.8% | Test: loss=0.436, acc= 84.1%
Epoch   3 | Train: loss=0.408, acc= 85.0% | Test: loss=0.385, acc= 86.1%
Epoch   4 | Train: loss=0.386, acc= 85.8% | Test: loss=0.377, acc= 86.2%
Epoch   5 | Train: loss=0.367, acc= 86.4% | Test: loss=0.382, acc= 86.6%
Epoch   6 | Train: loss=0.357, acc= 86.7% | Test: loss=0.380, acc= 86.2%
Epoch   7 | Train: loss=0.350, acc= 86.9% | Test: loss=0.402, acc= 85.9%
Epoch   8 | Train: loss=0.343, acc= 87.2% | Test: loss=0.377, acc= 86.4%
Epoch   9 | Train: loss=0.333, acc= 87.7% | Test: loss=0.364, acc= 86.5%
Epoch  10 | Train: loss=0.325, acc= 88.0% | Test: loss=0.355, acc= 86.9%
Epoch  11 | Train: loss=0.328, acc= 87.8% | Test: loss=0.405, acc= 86.1%
Epoch  12 | Train: loss=0.321, acc= 88.1% | Test: loss=0.359, acc= 87.0%
Epoch  13 | Train: loss=0.319, acc= 88.1% | Test: loss=0.357, acc= 87.2%
Epoch  14 | Train: loss=0.313, acc= 88.4% | Test: loss=0.358, acc= 87.5%
Epoch  15 | Train: loss=0.317, acc= 88.3% | Test: loss=0.346, acc= 87.4%
Epoch  16 | Train: loss=0.308, acc= 88.6% | Test: loss=0.338, acc= 87.7%
Epoch  17 | Train: loss=0.306, acc= 88.6% | Test: loss=0.355, acc= 88.0%
Epoch  18 | Train: loss=0.304, acc= 88.7% | Test: loss=0.365, acc= 86.5%
Epoch  19 | Train: loss=0.302, acc= 88.8% | Test: loss=0.335, acc= 87.6%
Epoch  20 | Train: loss=0.296, acc= 88.9% | Test: loss=0.346, acc= 87.8%
```




I have been experimenting with different hyper-parameters of the model, including dropout rate, learning rate, different weight initialisation methods, batch and patch sizes, etc. However, I have only included two of these attempts, both with batch size of 256, patch size of (4, 4), and learning rate of 0.01 and 0.001. I run the model with a patch size of (2, 2), but it took a very long time to carry out the training and testing, so I gave up on that direction.

**Task 5**

**Final model accuracy on Fashion-MNIST Validation Set**

The highest model accuracy on a Test set was 89.0% (see 'Try 1').

I would like to emphasise it here too, that my model architecture and understanding was highly influenced by others' work I found on the Internet, whether it was a YouTube video, articles, or any other sources. I tried to point to these sources both at the beginning through the referencing, and throughout the code too.