

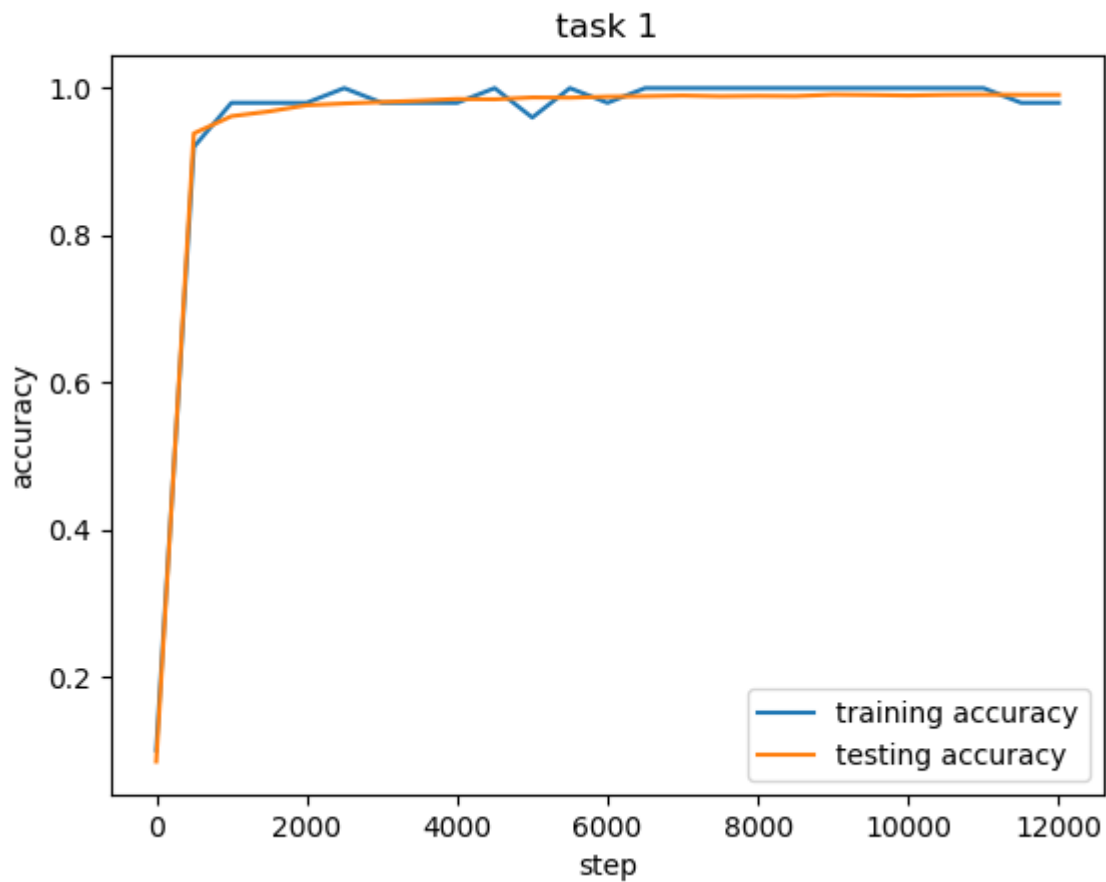
ML 2018 Ex1

By: Li Zhu(Julie)

A

There are $(5 \times 5 \times 1 + 1) \times 32 = 832$ free parameters in conv1 layer, $(5 \times 5 \times 32 + 1) \times 64 = 51264$ free parameters in conv2 layer, $(7 \times 7 \times 64 + 1) \times 1024 = 3212288$ free parameters in fc1 layer and $(1024 + 1) \times 10 = 10250$ free parameters in fc2 layer.

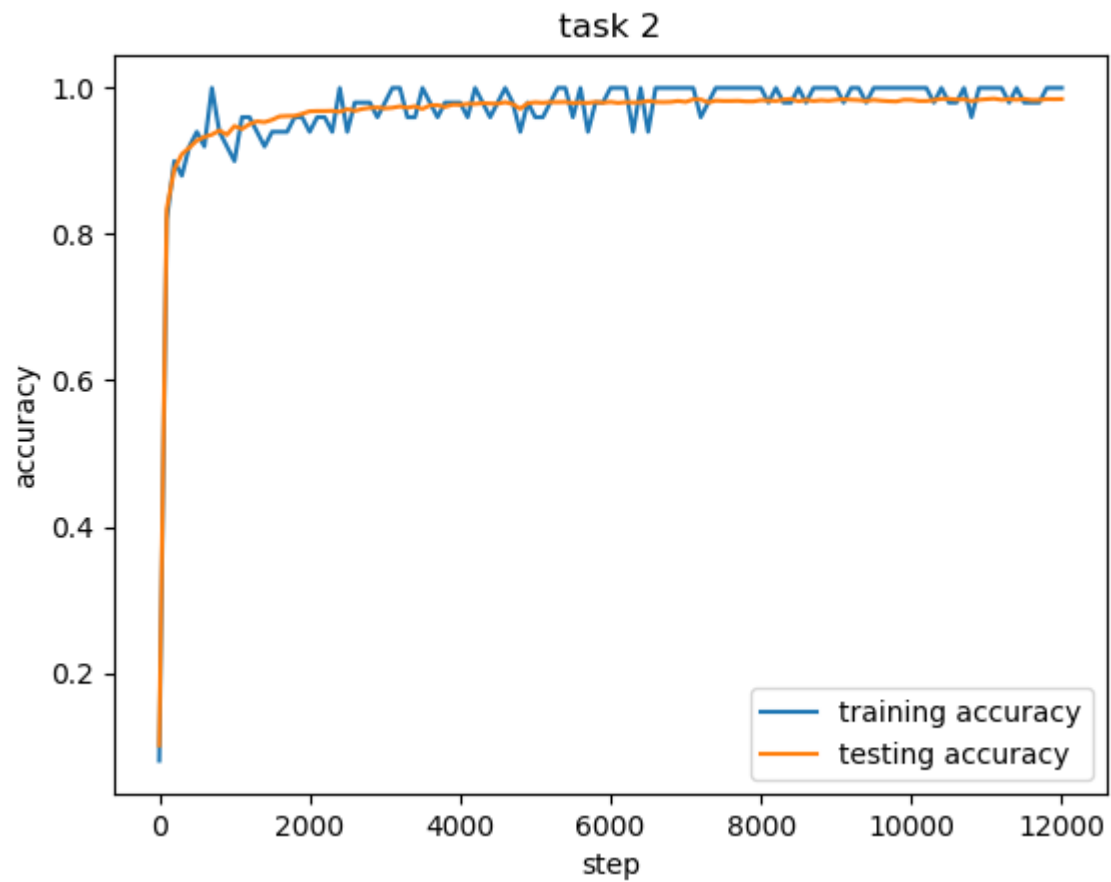
So this network has a total of 3274634 free parameters.



B

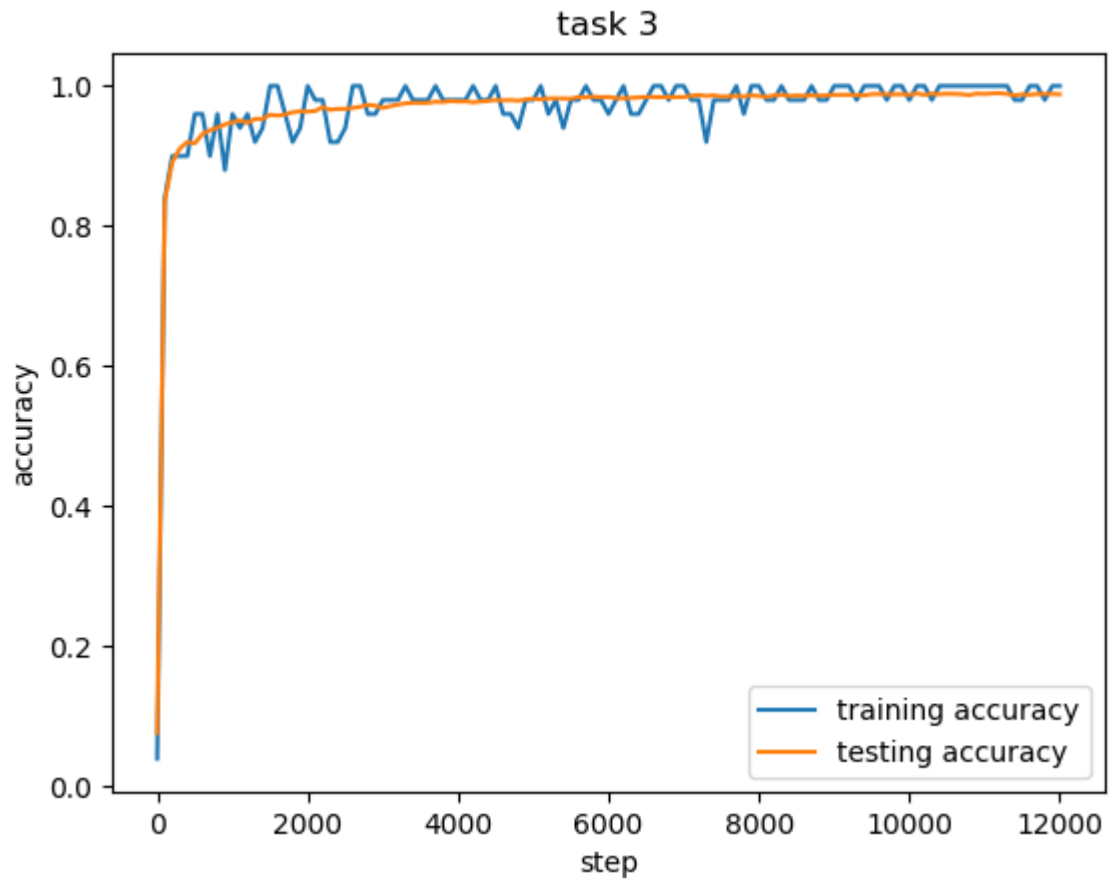
The ReLU function(the activation function) is non-linear. So I remove the ReLU functions. Later the network works slower, but the accuracy has nearly no change.

Maybe the figure looks like a little different, it's just because I set new steps.

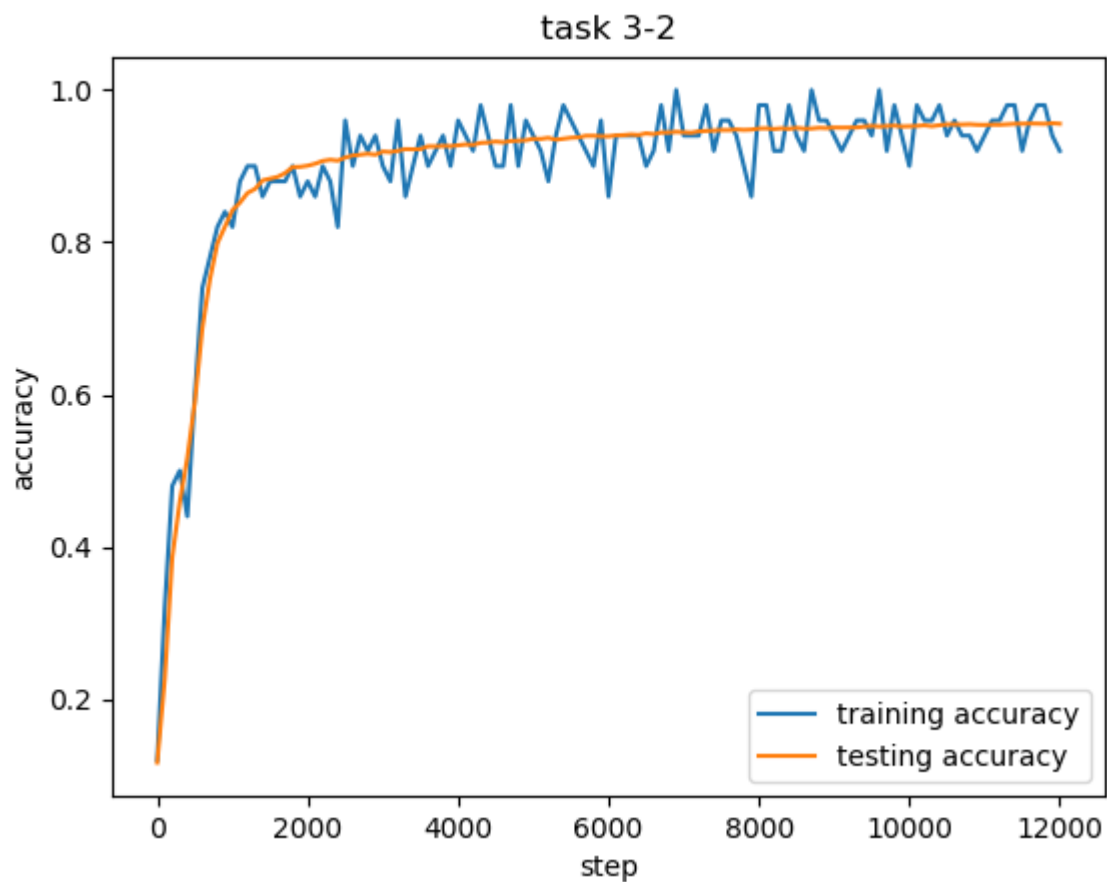


C

1. After removing the second conv layer and the second max pooling, the network still work well. The accuracy is as follow.



2. By decrease the number of the free parameters with the accuracy no less than 95%, I find that the number of the free parameters in conv1 layer can be reduced to $(5 * 5 + 1) * 3 = 78$, and the conv2 layer: $(5 * 5 + 1) * 6 = 156$, the fc1 layer: $(7 * 7 * 6 + 1) * 45 = 13275$, the fc2 layer: $(45 + 1) * 10 = 460$. so there are at least 13969 free parameters.



D

The network structure is similar to the original network, only a little difference.

The size of the input image becomes 28×56 (or we can also make it as 56×28) and the number of the class changed to 19(0~18).

The network is much slower. So by the way, I change the training batch size to 30(just have a try!).

