
Model-Agnostic Meta-Learning (MAML) - No 2(b)

David Ishak Kosasih (20195033)

June 28, 2020

For this experiment, I decreased the number of task for the task distribution to 50000 in order to reduce memory consumption. Moreover, I also fixed the maximum number of epoch to 30000; however, for some experiments I stopped the training process when the network started to converge.

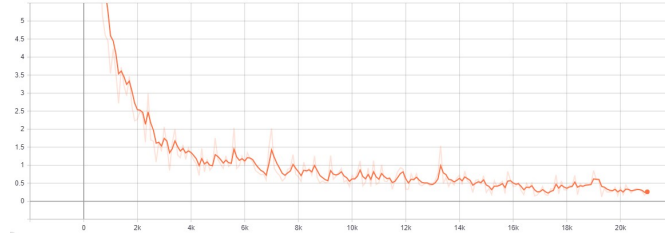


Figure 1: Meta-validation loss for omniglot 5 way 1 shot, $\alpha = 0.4$

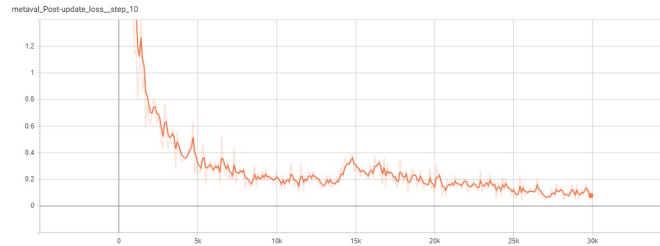


Figure 2: Meta-validation loss for omniglot 5 way 5 shot, $\alpha = 0.4$

Figures 1 to 4 showed that the model converge more quickly when it's evaluated on omniglot dataset compared to when it's evaluated on

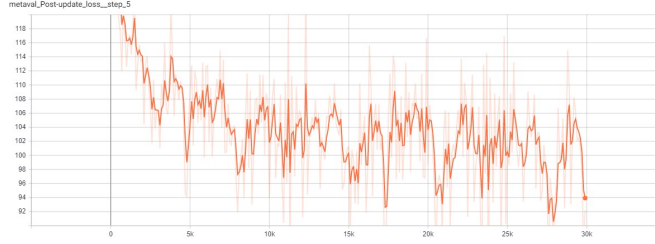


Figure 3: Meta-validation loss for *miniImagenet* 5 way 1 shot $\alpha = 0.01$

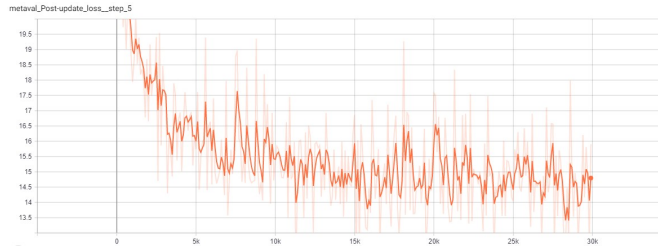


Figure 4: Meta-validation loss for *miniImagenet* 5 way 5 shot $\alpha = 0.01$

miniImagenet dataset. For omniglot dataset, it started to converge only after approximately 5000 epoch. On the other hand, when *miniImagenet* was used to evaluate the model, it didn't start to converge even after 30000 epoch. In my opinion, there are at least two things that contributed to this result :

1. Different learning rate.

In this experiment, I fixed the learning rate to 0.4 and 0.01 for omniglot and *miniImagenet* dataset respectively.

2. Dataset

As we already know, *miniImagenet* dataset is more challenging or complicated than omniglot dataset.

According to the theory, when we increase the learning rate, our model should learn more quickly. The prove of this hypothesis can be seen in figs. 5 and 6 as the model took less time to converge when the learning rate was fixed to 1. As for the accuracy, it also yielded better result in this experiment. The drawback of this approach is that our model may never converge if we set the learning rate too big since it can end up jumping

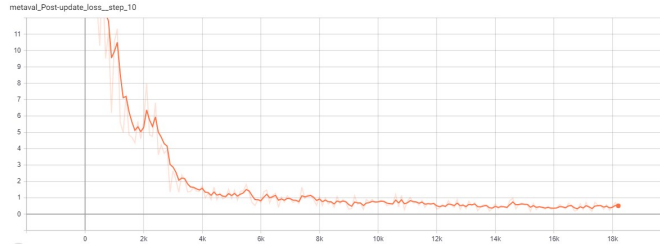


Figure 5: Meta-validation loss for omniglot 5 way 1 shot, $\alpha = 1$

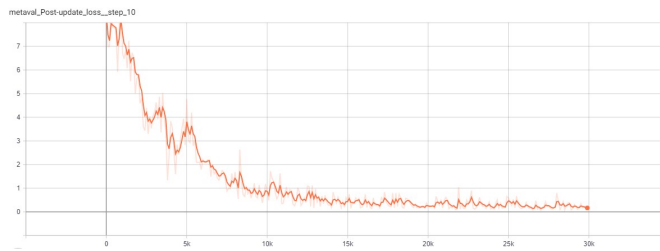


Figure 6: Meta-validation loss for omniglot 5 way 5 shot, $\alpha = 1$

randomly; however, since I fixed the learning epoch to a relatively small value for this kind of task, this problem didn't happen.