1. Simple definition of transfer learning.

Suppose we have a source data Ds(x) and source task Ts(x), another target data Dt(x') and its related task Tt(x'). We hope a model trained in the Ds for Ts can be used to solve Tt based on Dt.

2. Traditional method of transfer learning on image task.

In general, we think a model trained based on a huge image dataset like ImageNet with labels has strong ability to extract texture feature from images. So we hope we can use it to solve a similar problem. The general way is to modify the last layers' neuron size to adjust the number of classes in the new task. If the new data's distribution is very similar to ImageNet's, we will fix the previous layers' weights and only optimize the last few layers' weights. If the problem is much different from ImageNet, we will optimize whole the weights.

3. Benifit from transfer learning.

In some field, the data with labels is hard to get. For deep learning, a large number of data is necessary for a complicate model, and a complicate model is important if you want to get good performance. So a few number of training data can hardly help to train feature extraction ability for a big model. For different kind of images, the feature extraction direction we need is different because important texture and pixel distribution are different. So we hope a model has been trained on a large dataset which has basic feature extraction ability is easier to converge in a small target dataset. In another words, a pretrained model's weights is a good initialization for new task.

4. Simple transfer learning from cifar10 to cifar100.

I use two experiments to show its advantage.
Dataset: Cifar10 , Cifar 100
Model: ResNet-18

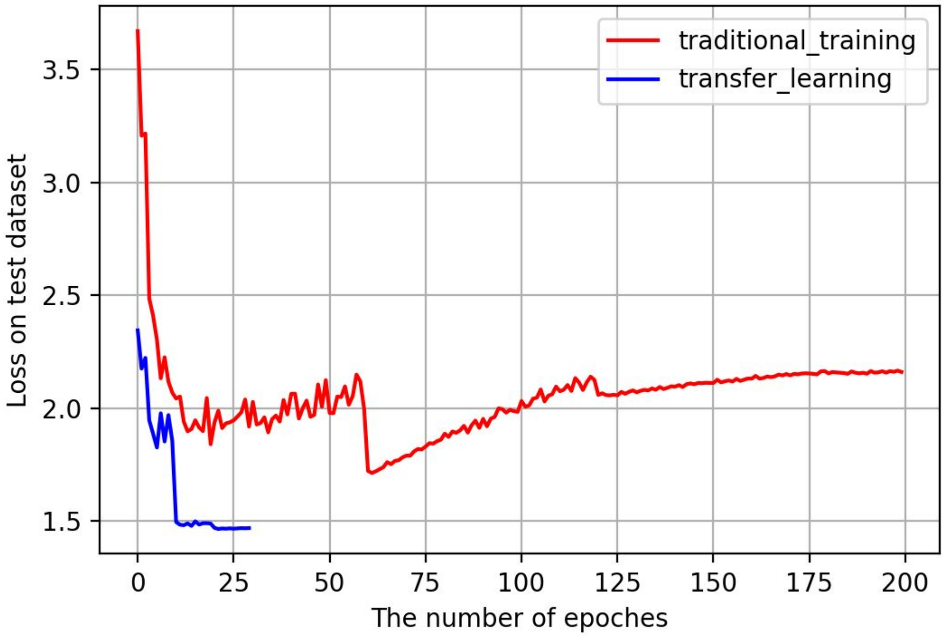
Experiment A:

Train model on Cifar100 from 0, which means I use random weights to initialize.

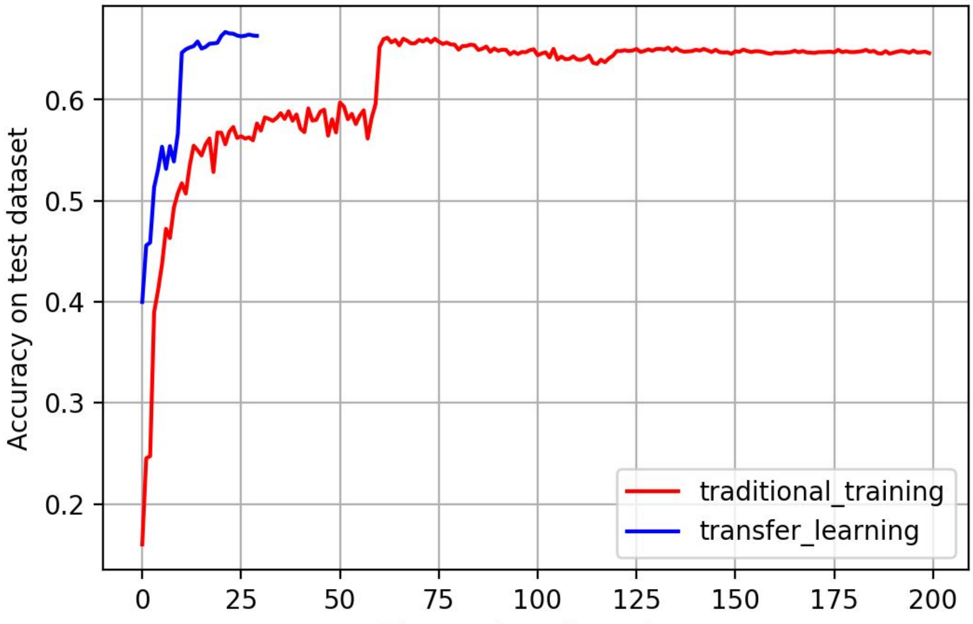
Experiment B:

Retrain a model which has been trained on Cifar10, modify the last layer's size to suit for 100 classes.

I use two pictures to show the result, both picture is the performance on test dataset.



Blue line is transfer learning and red line is training from begining. We can see that model trained in cifar10 can converge to much lower loss in less than 25 epochs. Training from beginning seems overfit after 60 epochs.



In this picture, we can see transfer learning can get better accuracy which training from beginning can hardly get in 30 epochs.

5. Conclusion

Cifar100 has 100 classes and 50000 images in train dataset, Cifar10 has 10 classes and 50000 images in train dataset. Because their images are similar, both them are object in real world, so transfer learning perform good in this experiment. Pre-trained on cifar10 makes the model’s weights to be a good initialization, which helps the model to converge in cifar100. Transfer learning shorter our training time.