

Weekly Report

Haozhe Feng

December 10, 2017

Weekly Report

The tasks I have finished this week

The tasks I plan to do next week

Continue the training of Faster R-CNN

Do researches based on the experiment

The sparse distribution of weight

Some other points

The tasks I have finished this week

This week, I complete the project of *Faster R-CNN* and begin the **training** of it. During the training process, I have fixed a large quantity of bugs myself, and gained much training experience. I can analysis the training results and try some **state-of-art** achievement of the papers I have read such as [DenseNet](#), [Mask R-CNN](#) which I should finished a few weeks ago. I have added the focal loss part in the *Faster R-CNN* and focal loss quiet shows a specific improvement in the decrease process of loss.

During this experiment, I mainly meet 3 classes of problems I haven't thought about when reading papers, and they are :

1. The numerical stability

I find that during the whole network framework, the problem of *numerical stability* is a big problem. And I mainly meet the problem of the initial choice of *learning rate* and the presentation of the instability is the explosion of gradient. Finally, the operation of adding BN layers solves this problem, which shows me the efficient of BN operation. I also find an interesting phenomenon that if we view the optimization progress as a function of the initial value of *learning rate*, **this system is a non-linear system of learning rate**. In my experience, the 0.001's change of *learning rate* will lead to the different result of convergence and divergence. I will take notice of this phenomenon in the next experiments.

2. The advance of vectorization

Matrix operation is largely used in *CNN related networks*, and how to write the operation of the net vectorially is an important problem. I find the efficient of network is highly improved when I write some operation into vectorization, and I think it is a good way to use our knowledge of *linear algebra* into the network.

3. The focal loss part

Paper [Focal Loss](#) proposes a strategy to process the imbalance problem of positive and negative samples. I add it into my *Faster R-CNN* framework, and drop the NMS operation during training. It really makes the loss function converge smoothly, and releases the imbalance problem.

I also meet and solve some technique problems like storage leaky and so on, but they aren't the main point.

The tasks I plan to do next week

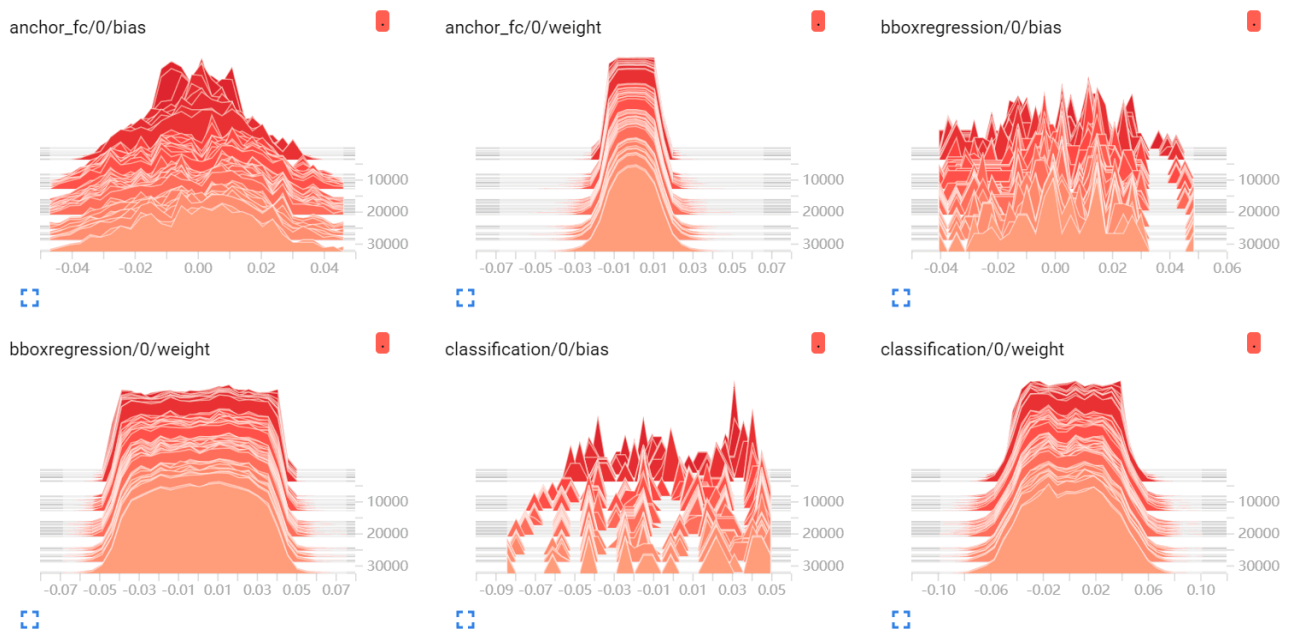
Continue the training of Faster R-CNN

Faster R-CNN is a **two-stage** network and I just begin the training of the first part. I will continue the training experiment of the network and learn how to merge and use 2 networks together. I hope I can get a result that can **achieve the result** mentioned in the *Faster R-CNN* paper.

Do researches based on the experiment

The sparse distribution of weight

I also plan to do some research based on this experiment. I use the *vgg16_bn* net as the basic, which is an old net. During the training, I find a statistical prove that the connection weight is **sparse** in the deep network framework, just like this :



The distribution of weight is clustered around 0, and it makes me think of the article [Identity Mappings in Deep Residual Networks](#), which is the theory basic of *ResNet*. Comparing with the experiment, I can learn a lot which I didn't notice when reading the article.

Some other points

I also plan to try many other structures such as *DenseNet* , *Deep Supervision* and *Mask R-CNN* using this **object detection task**, and fill the vacancy that I can't directly learn from reading articles.