Exeriese 3

Lan Zhang
954517477

1 Problem 1

Solved on 11/02/2019. The seed I used in bellowing experiments is 2612.

1.1 Batch Normalization

I implement batch normalization after the convontional layer. Figure 1 is the performance of the model with and without batch normalization. With batch normalization, the model can converge faster.

Then, I tried to compare my implements of batch normalization with tensorflow's implements. At first, I didn't add trainable variables and calculate their gredients. The results of tensorflow' batch normalization layer is highly different from mine. After applying gredients, Figure 2 is the performace of my batch normalization and tensorflow's batch normalization. The difference of two implement is quite small(less than 0.01). I extracted comparable parameters of two batch normalization layer – gamma and beta, and summerized their differences shown as Figure 3. The differences of parameters are gradually increase. I think the implements of tensorflow is more complex and may have other mechanism that lend the differences.

1.2 Layer Normalization

Layer normalization is implemented after the convontional layer. Figure 4 is the performance of the model with and without layer normalization.

Then, I compared my implements of layer normalization with tensorflow's implements. After applying gredients, Figure 5 is the performace of my layer normalization and tensorflow's layer normalization. I guess the reason why my layer normalization is better than tensorflow's is **. I extracted comparable parameters of two layer normalization – gamma and beta, and summerized their differences shown as Figure 6 .

1.3 Weight Normalization

1.4 The performance of different normalization

33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.

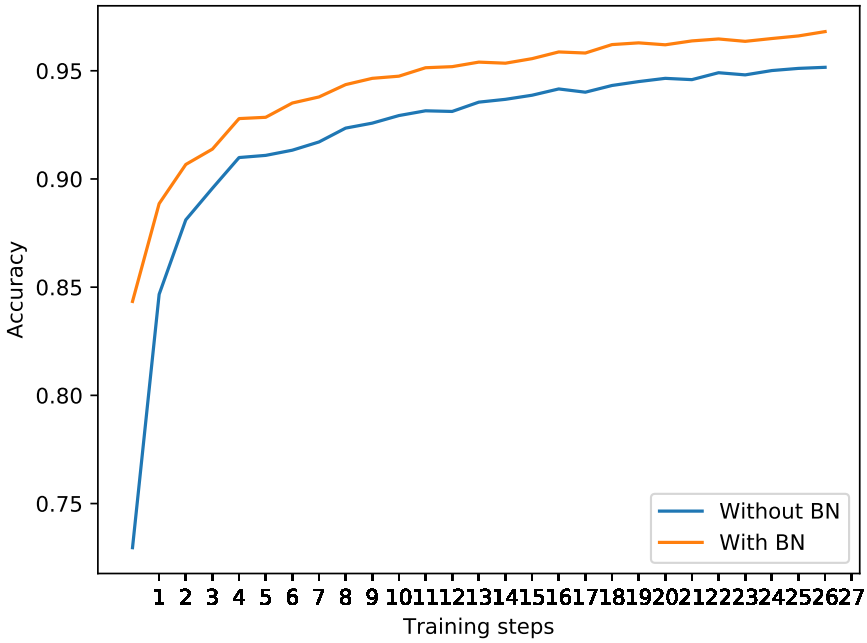

Figure 1: Performance of batch normalization on test dataset.

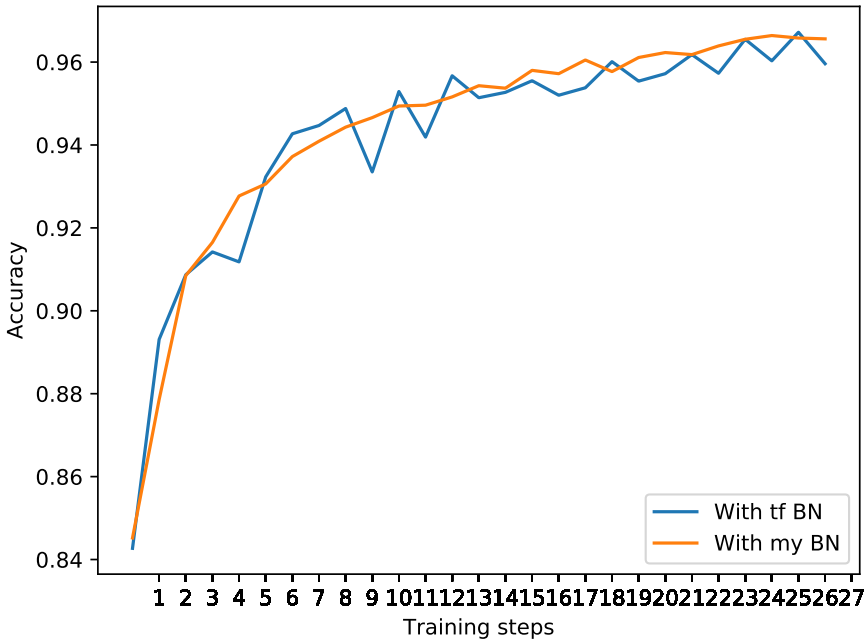

Figure 2: Compare my implements of batch normalization with tensorflow's implements

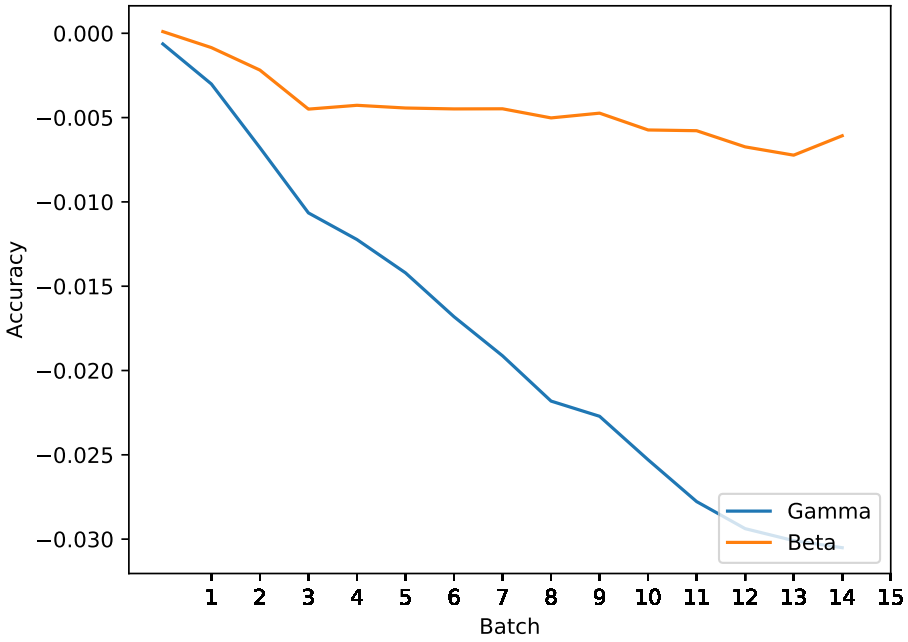

Figure 3: Compare parameters difference of batch normalization

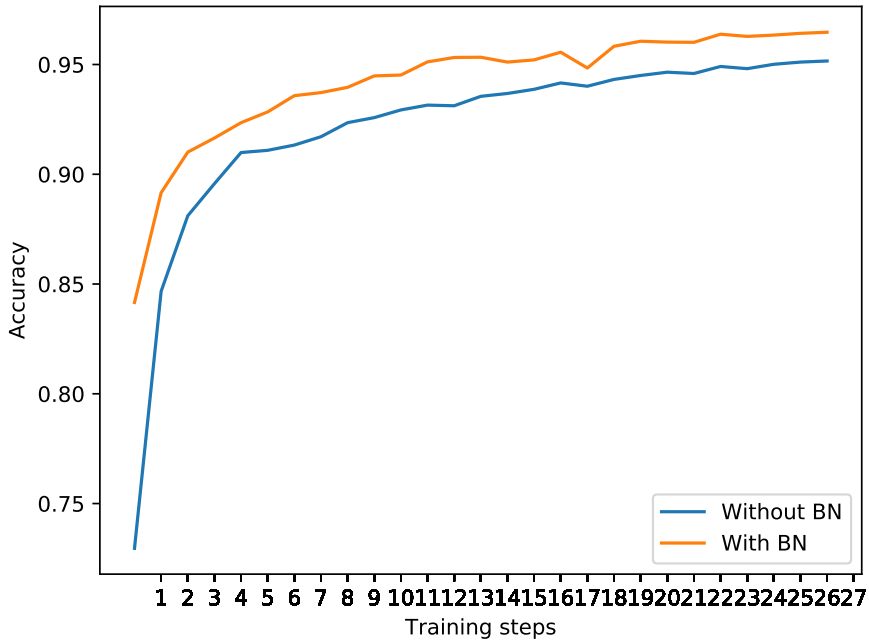

Figure 4: Performance of layer normalization on test dataset.

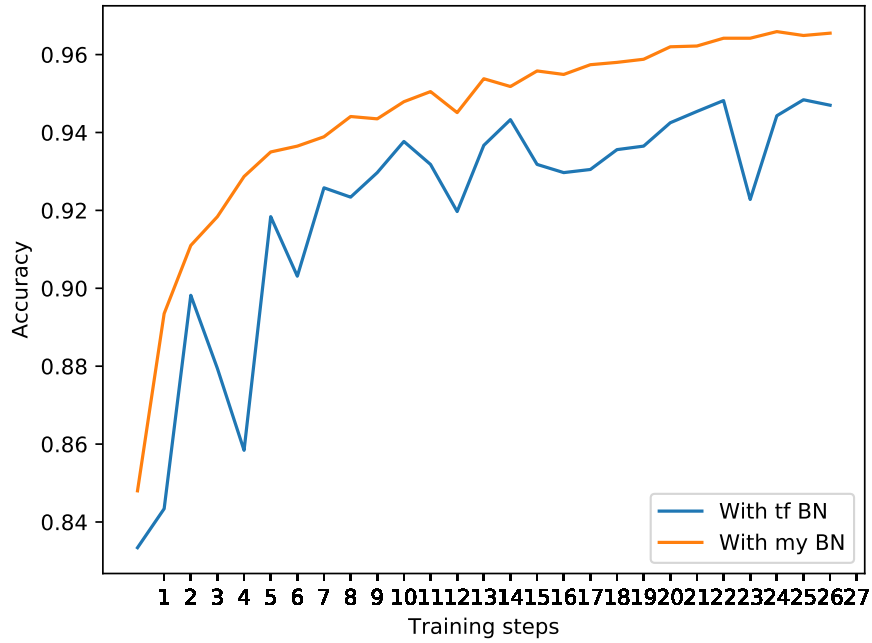

Figure 5: Compare my implements of layer normalization with tensorflow's implements

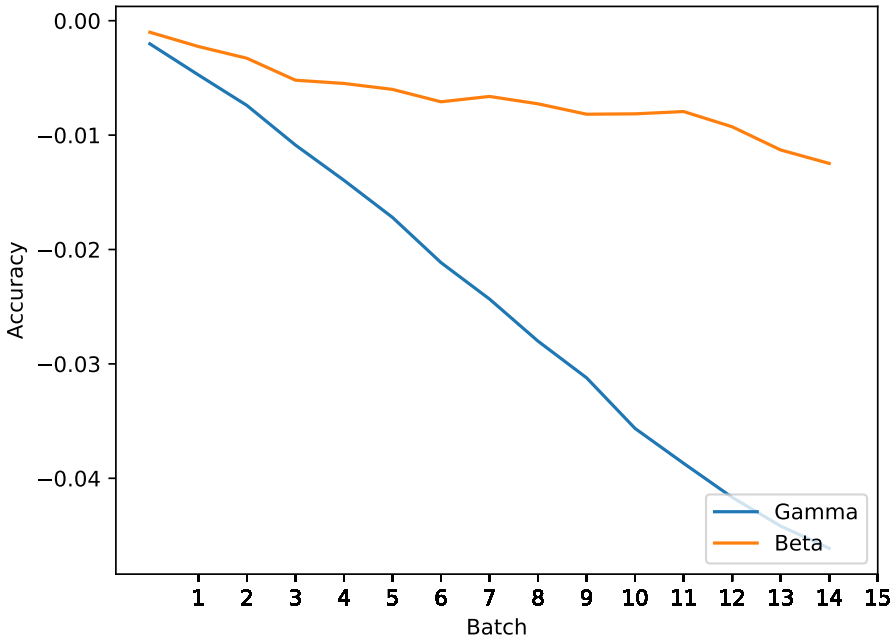

Figure 6: Compare my implements of layer normalization with tensorflow's implements