Deep learning course project

Jianing Dong
djn1996@tamu.edu

1 Network Architectures

The method used here is mainly an extension and modification on ResNet [1], that is, ResNEXT [2]. ResNet is a widely-used convolutional neural network and has achieved great success in image classification task, which demonstrates its representational power when dealing with images. ResNEXT repeats a building block that aggregates a set of transformations with the same topology. Shortly, it is a homogeneous, multi-branch architecture. The following figure shows the key idea of ResNEXT.

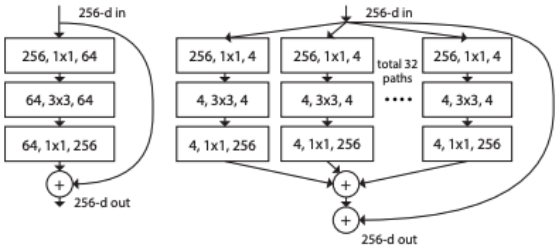


Figure 1. **Left**: A block of ResNet [14]. **Right**: A block of ResNeXt with cardinality = 32, with roughly the same complexity. A layer is shown as (# in channels, filter size, # out channels).

This strategy exposes a new dimension, which we call "cardinality" (the size of the set of transformations). The authors show that increasing cardinality is able to improve classification accuracy. I have seen it from the following results.

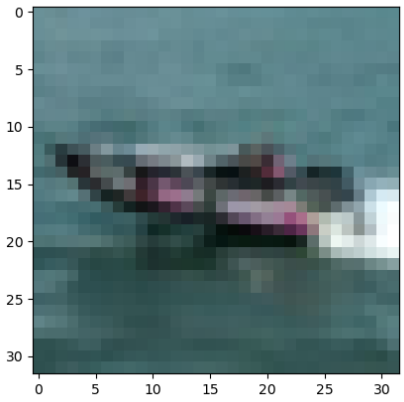
2 Experimental Results

I choose PyTorch 1.6 to implement the whole architecture. After tuning based on validation dataset, the final model is ResNEXT-29 (9n+2, n=3). I tried multiple cardinalities and other model structure parameters. In the end, I choose cardinality, base width, widen factor and number of first filters as 8, 64, 4 and 64 respectively. The detail of training hyperparameters is as follows. I start with a learning rate of 0.05, divide it by 10 at 50, 100 and 150 epochs, and terminate training at 200 iterations. I choose the best epoch number based on the accuracy of validation set. In addition, I set batch size, and weight decay as 128, and 5e-4 respectively.

After fixing hyperparameters, I retrain the model by combining training and validation dataset. The test accuracy for the public testing dataset is 90.82%. The prediction of private testing dataset is provided in the zip file.

The first picture of the private testing set is

32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, Canada.



3 Discussion and Future Direction

The authors claims that increasing cardinality is able to improve classification accuracy. However, they didn't propose a way to choose the cardinality for a specific task. On the one hand, we can think that larger cardinality means more blocks capturing the features of data. On the other hand, more cardinality requires more computation resources. It should be interesting to study on a strategy to choose a reasonable cardinality of the model.

References

[1] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[2] Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. Aggregated residual transformations for deep neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1492–1500, 2017.