All Convolutional Net for Recognition and Transfer learning

Xiangyi Zhang
ShanghaiTech University
zhangxy9@shanghaitech.edu.cn

Abstract

All convolutional neural network is proposed a new structure for image recognition without maxpooling layer, and replaced by a convolutional layer, which brings better performance. We follow the baseline in [1] and implement three models and other three additional models which represent on the paper. Apart from this work, we design a new transfer learning model, indicating that all convolutional neural network perform well on transfer learning program.

1. Introduction

The task of object recognition focus on the same convolutional neural net principle, which concludes alternative convolution and max-pooling layers followed by fully connected layers. However, Jost proposed that pooling layers can be replaced by a convolution neural layer. Followed by his work, we verify the role of pooling layer and model a new transfer learning structure.

2. Model review

We inplement three base network: Model A, Model B, Model C, and three additional network: Strided-CNN-C, ConvPool-CNN-C and All-CNN-C in separately. Each of them have same high level layers and different lower level layers. And three additional net are built on the All-CNN-C based net. All models are used for evaluating the importance of pooling in case of classification on CIFAR-10. Apart from this work, we did another experiment of transfer learning on two subset of CIFAR-100. On each subset, we firstly take part of ConvNet of Model All-CNN-C as fixed feature extractor and fine-tuning the Convnet later.

3. Experiment

3.1. Dataset

In our first experiment we compared six models from section 2 on the CIFAR-10 dataset. All networks were

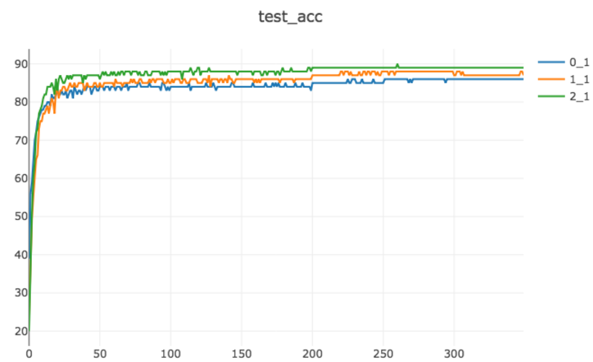

Figure 1. Test accuracy curves of three base networks on CIFAR-10 dataset. The highest green one is model C, outperform others due to stacked layers with 3x3 kernel size. Because of shallow layers, Model A show worst accuracy in blue color. The curve of model B is in the middle of the others.

Model	Accuracy (%)	SOTA (%)
ANet	86.4	87.53
BNet	88.7	89.80
CNet	90.1	90.26
SCC	89.6	89.81
CCC	90.2	90.69
ACC	90.9	90.92

Table 1. Comparison of test accuracy between our results and results by Jost on the CIFAR-10 dataset.

trained using stochastic gradient decent with fixed momentum of 0.9. The learning rate lr was adapted using from 1e-1, 1e-2, 1e-3, 1e-4 , and we finally choose lr = 1e-2 with best performance. And we discarded the dropout layer with no better performance. As for data augmentation, we use random crop, horizontal flip and normalized opreations. Finally, we find it helpful to use weight initialization to avoid entering into local optimal.

3.2. Result

Our result consists of two parts: Firstly we experiment on six models from section 2. secondly we do transfer learning on two subsets of CIFAR-100. On each subset, we alse

Figure 2. Test accuracy curves of three variation networks on CIFAR-10 dataset. Roughly speaking, three variation models perform better than base models. All-conv-net in blue color perform best, ConvPool-cnn-c in orange color is in the middle, and Strided-cnn-c in green color is in the bottom, but without large gaps.

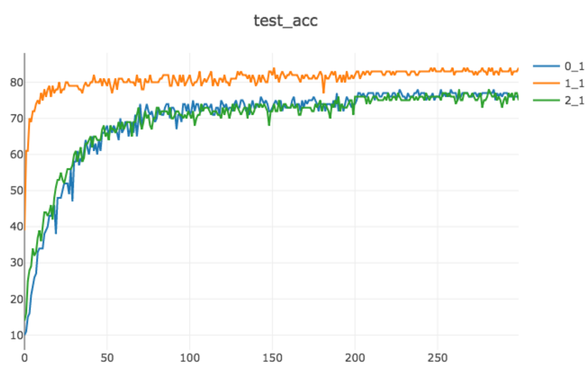

Figure 3. Test accuracy curves of transfer learning and from scratch training. The highest one is the curve of transfer learning on class 2 dataset, the blue and green one correspond to class 1 and class 2 dataset with training from scratch. From the curves, transfer learning converge faster than from scratch, and show better performance than from scratch.

Dataset	Training Strategy	Accuracy (%)
class 1	transfer learning	87.0
class 2	transfer learning	84.5
class 1	from scratch	78.4
class 2	from scratch	78.2

Table 2. Comparison of test accuracy on different training strategy — transfer learning and from scratch.

train an entire Convolutional Network from scratch (with random initialization) as a comparative experiment.

The result of all models that we considered are given in Table 1 and Table 2. For the first six models on Table 1, the model A from previous literature(Srivastava et.al.,2014) performs well of 86.4% accuracy. For second model B removing maxpooling layer and increasing the stride achieving higher performance. While Model C performs best among three based net of 90.1%. Among three based models, Model C performs well because of the largest receptive field. Basenet: Test accuracy curves of three base networks on CIFAR-10 dataset is in Fig 1.

About three additional models, All-CNN-C performs better than others on 90.9% which indicates that replacing maxpooling layer by a conv layer helps to improve the accuracy. While maxpooling layer seems to perform well than without maxpooling, each of which is 89.6% and 90.2%. Test accuracy curves of three variation networks on CIFAR-10 dataset in Fig 2.

Finally, for the results of transfer learning. From the table, the networks with transfer learning strategy outperform ones with from-scratch strategy. Transfer Learning Net: Test accuracy curves of transfer learning and from scratch training in Fig 3.

References

[1] J. T. Springenberg, A. Dosovitskiy, T. Brox, and M. Riedmiller. Striving for simplicity: The all convolutional net. *arXiv preprint arXiv:1412.6806*, 2014.