CS231n Assignment Summary

Chengzong Cai
toadjiang@163.com

March 15, 2019

Abstract

After months of learning,I finally completed CS231n lessons and assignments.The lessons brings me to the inner world of deep learning and computer vision.In this report,I'd like to share my acquisitions and confusions with teachers and buddies.

1 Assignment 1:Naive classifiers

In assignment 1,I learned the basic and simple classifiers including kNN,SVM and a small two-layer network.

1.1 kNN

The principle of a kNN classifier is quite simple.a kNN classifier doesn't have a training process, it just compare the input images and the whole data set,picking the most common label of the k nearest neighbors.The performance of this algorithm based on the given dataset is so poor that the highest accuracy is only nearly 0.3

1.2 Support Vector Machine

Compared with kNN,support vector machine classifier is more complex and it's performance is also better. A multitask support vector machine is formed by the linear mapping layer and hinge loss function.The classifier firstly multiplies the input data with a weight matrix W and then add a bias b to the result.The result can be considered as the predicted score of each class.And the hinge loss function computes the loss based on the following formula.

$$HingeLoss = \sum_{\hat{y}} max(0, 1 + S_{\hat{y}} - S_y)$$

The performance is a bit better than kNN,which achieves the accuracy of 0.4.

1.3 Softmax classifier

This assignment is analogous to the SVM one.The only difference is the replacement of hinge loss by softmax loss function.

1.4 Two Layer Network

In this section,a simle two layer network is implemented for classification.Different from the former classifier, this one added a activation function to the output of each linear mapping layer.The activation function bring non-linearity,which makes the fitting ability more powerful.And hence the performance the accuracy is nearly 0.5 is the best among those classifiers.

2 Assignment 2:CNN and Some Tricks

In assignment 2,I implemented a naive CNN and some useful tricks in training neural networks like BN and Dropout.

Batch Normalization Batch normalization is a commonly used trick applied in data among each layer.It changes the bias and variance of data to address the internal coviarance shift problem. In this assignment,I got confused by the backward process.I tried to brutally compute the analytic derivative using my calculation skills.But it failed.Dealing with tensors is quite different from dealing with scalars.So I changed my mind.I drew a data flow graph and compute derivative node by node.Then finally a correct answer emerges.By accomplishing this assignment,I get a deeper understanding of data flow graph and the backward pass of tensor derivative.

2.1 Dropout

Dropout is another normalization trick.Commonly speaking,it blocks a certain part of input image to force neural network to learn from the entirety not only some certain features.It does prevent overfitting and improves the test-time accuracy.

2.2 CNN

Convolutional neural network is the basic of those advanced computer vision algorithms.In this section,I accomplished a naive convolutional neural network using NumPy.The efficient of ”Fast Layers” written in Cython attracts my attention,so in the following study I want to learn more about the architecture of those deep learning frameworks.

3 Assignment 3:Application of CNN and RNN

Among these assignments,the third one is most interesting and challenging.Using popular framework like PyTorch and Tensorflow to address some practical task like style transfer brings me quite sense of achievement.

3.1 GAN

As Yann Lecun says,GAN is the most creative idea in recent years.Unlike other computer vision algorithms,GAN teaches computer to paint a image.In this assignment,I wrote a vanilla GAN,which generates images like MNIST.The writing part is not hard following the given instructions.I think the funnest part is the training process.In training process,the discriminator tries to tell whether this image is true or false and the generator tries to cheat the discriminator,just like buyers and sellers in a antique market. However,the vanilla GAN is hard to training and the performance is not satisfying.So there comes WGAN.Replacing the original loss function with wasserstein distance brings a faster and more stable training process and a better performance.

3.2 RNN and LSTM Captioning

Like a task of a preschool child,captioning is to learn to describe an image.Technically,we firstly extract features from raw images using CNN and then feed these features as hidden states to a vanilla RNN or LSTM which generates description.The difficult point of this assignment is the implementing of LSTM backward.Something weird comes out that when I alter a part of code without changing it's function,the result turns totally different and I can't still figure out why.At the end of the assignment,they referred character-wise RNN that draws my attention.Compared to the word-embedding RNN,I think the idea of character-wise RNN is much more smart.Without remembering every single word,it just processes words like a human brain.

3.3 Network Visualization

In this section,I was instructed to exploring the inner world of neural networks.First I try to find out where the neural network focus at when it is working.In next step I fooled the neural network.And finally I used gradient ascent to generate a image that the network think the object is.In completing this assignment,I realized that the current neural network is not as robust as I thought,it can be fooled by simple added noises.And I think this area about improving robustness of neural network will be interesting.

3.4 Style Transfer

In this assignment,I completed a style transferring neural network.The core idea of style transferring in the mentioned paper is cool.It says that to transferring a style of a image is to transfer low-level features of this image to the target one without changing the high-level feature of the target.So they designed a loss function to handle this.However even I uses many different paints in different artistic style as the style image,the generated image still looks like oil paintings.After completing the assignment,an idea comes to my mind.What if we can apply a text style transferring using the same idea?

4 Some Confusions

Q1:After completing those assignment,though I learned so much,but I still feel powerless when I try to write code as beautiful as those paper authors and my seniors do.I attempted to read Pytorch,Tensorflow documents,however soon I found it too boring to keep reading.What should I do to truly master Pytorch and Tensorflow that I can write code in their level? Q2:Some research areas draw my attention.They are

GAN,medical image processing and even meta learning(I will talk about it in next weekly report).Which one should I choose?