AI3607 Homework 2 Report

Weijiude 521030910418

May 14, 2023

1 Introduction

We imitate AlexNet to construct a CNN to solve the classification problem on CIFAR-10. We reload training set where only 10% data with first 5 labels are reserved. With the same network architecture, the classification accuracy over the first 5 labels significantly drops. We try data augmentation to restore the accuracy over first 5 labels and accuracy bounces back with restriction.

2 Network Architecture

Our network imitates AlexNet and contains five layers: three convolutional and subsequent two fully-connected.

The first convolutional layer receives input image of size $3 \times 32 \times 32$ and filters it with kernels of size $3 \times 3 \times 3$. The second convolutional layer receives former output and filters it with kernels of size $32 \times 3 \times 3$. Then the third convolutional layer filters it by kernels of size $64 \times 3 \times 3$. A max-pooling layer lies between each two convolutional layers to reduce overfitting.

Two fully-connected layers reduces the output features of convolutional layers from 2048 to 512 to 10. Dropout layers are used to reduce overfitting.

3 Primal Experiments

3.1 CIFAR-10 Baseline

We use a learning rate of 0.001 and weight decay of 0.0001 to train our model on the CIFAR-10 dataset. After 100 epochs, the training loss is steady around 0.624. The classification accuracy of our model on training set is 0.700.

The test result of out model on CIFAR-10 shows the overall classification accuracy is 0.565. The confusion matrix of classification over all labels is shown in Figure 1.

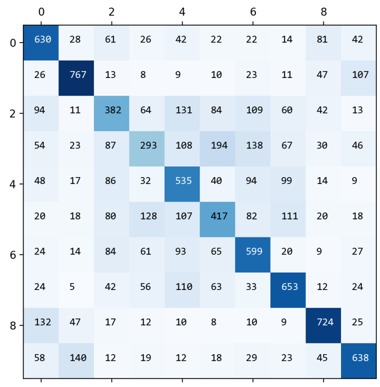

Figure 1: The confusion matrix over the 10 labels of the test on original CIFAR-10 dataset

3.2 Unbalanced CIFAR-10

We reload CIFAR-10 training set where only 10% data with the first 5 labels are reserved. We keep data with other 5 labels the same in training set.

We retrain our model on the reloaded dataset with same hyperparameters. The overall classification accuracy on test set drops to 0.448 on condition that the classification accuracy over the first 5 labels drops significantly while the accuracy over rest 5 labels slightly rises. The confusion matrix of classification over all labels is shown in Figure 2.

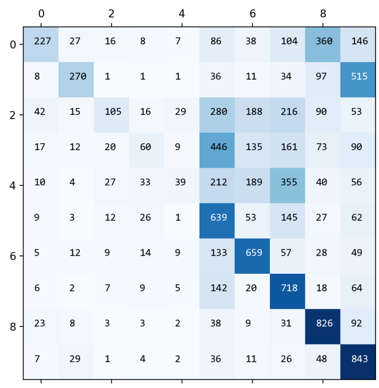

Figure 2: The confusion matrix over the 10 labels of the test on reloaded CIFAR-10 dataset

3.3 Analysis

Suppose we are solving a linear classification problem with SVM. It shows clear classification on full dataset. But it gives a poor classification on dataset where most blue pixels are thrown away (See Figure 4).

The classification depends on the sampling of data.

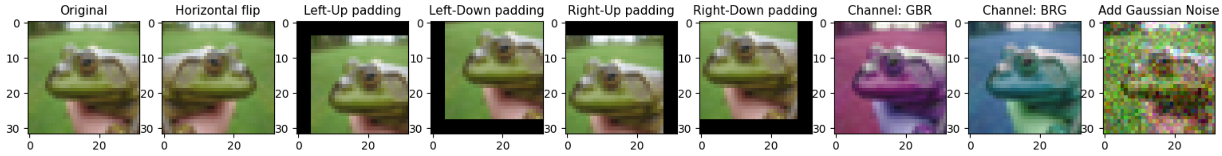

Figure 3: An implementation of data augmentation cast on a frog image.

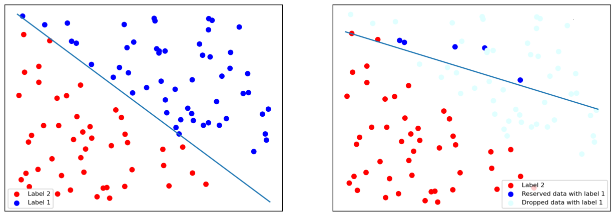

Figure 4: Due to the poor classification, most blue pixels are wrongly classfied.

We might have a better classification if we retreive the ratio of those labels among training set.

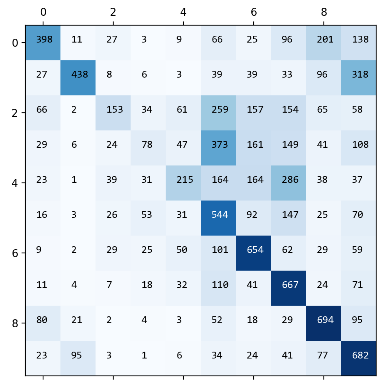

Figure 5: The confusion matrix over the 10 labels of the test on data-augmentated dataset.

Label / Scale	0	1	2	3	4
Original	0.63	0.78	0.39	0.29	0.54
UB 10%	0.23	0.27	0.10	0.06	0.04
DA 100%[1]	**0.40**	0.44	0.15	0.08	**0.22**
DA 150%[2]	0.35	**0.47**	0.20	0.10	0.13
DA 200%[3]	0.32	0.40	**0.27**	**0.10**	0.15

Table 1: Classification accuracy over first 5 labels based on various training methods

4 Restoration of classification

4.1 Data augmentation

Data augmentation is an effective technique to increase both the amount and diversity of data. Common augmentation methods include changing a few pixels or horizontally flipping the image. Tentatively, we cast one or a combination of following methods on original images to create new images. (See Figure 3)

1. horizontally flip
2. cycle RGB channels
3. randomly change a few pixels
4. move the object away from the center of image
5. add Gausian noises

4.2 Results

New images are yield by those 10% reserved images in training set. We extend new images into the training set, to enlarge the amount of images with first 5 labels in training set by ten times to match the original amount. Trained with same hyperparameters, our network shows bounce in classification accuracy over first 5 labels (See Figure 5). The median of bounce in classification accuracy over first 5 labels is 57%.

As we enlarge the amount of images with first 5 labels in training set, the classification accuracy is rising but constrained and will not keep up with the accuracy given by network trained on original dataset (See Table 1).

5 Discussion

Throwing away most data with certain labels may severely influence the classification result of neural network, resulting more wrong predictions in those labels and more right predictions in other labels. Simple data augmentation based on the reserved data can raise the classification to some extent, but cannot raise it to the accuracy given by network trained on original dataset.

Simple data augmentation based on the reserved data can increase the amount of data, but unable to enhance the diversity. Those thrown data cannot be simply derived by the reserved 10% data. Therefore, if we find a classification is poor, besides increasing the amount of parameters, we should focus on enhancing the diversity of all kinds of data to clearify the classification bound among different labels.

[1] Augmented by combination of method 1 and 4.
[2] Augmented by combination of method 1, 3 and 4.
[3] Augmented by combination of all five methods.