Classification of CIFAR-10 image data using Convolutional Neural Network

Jaeyong Lim
Seoul National University
jaeyonglim@snu.ac.kr

Abstract

This study investigates the effects of label and input perturbations on CIFAR-10 image classification using a simple convolutional neural network (CNN). We find that different types of perturbation have measurable impacts on both accuracy and class-wise performance. All source code and experimental results are publicly available at https://github.com/jaylim61/cifar10-classification.

1. Introduction

Image classification is one of the most fundamental tasks in computer vision. Among various bencmark datasets, CIFAR-10 serves as a canonical benchmark for evaluating model performance in supervised learning settings, containing 60,000 32 by 32 RGB images across 10 object categories.

While many architectures have achieved remarkable performance on clean datasets, real-world data often contains imperfections such as mislabeled samples or distribution shifts. This project aims to empirically study how various types of data and label perturbations affect the performance of CNN-based classifiers on CIFAR-10.

Specifically, we explore four experimental conditions: a clean baseline, random label shuffling, label noise affecting 20% of training data, and input perturbations applied through image transformations. We implement a simple CNN model using PyTorch and evaluate its performance under each condition.

2. Related works

Related works

3. Methods

3.1. Model Architectue

We use a simple convolutional neural network (CNN) composed of three convolution layers with increasing channel depth, followed by two fully connected layers. Each convolutional block consists of a 3x3 convolution, batch normalization, and max pooling. Dropout layers are added after each block to prevent overfitting. The last fully connected output layer uses a 10-way linear classifier suitable for CIFAR-10.

3.2. Dataset and Preprocessing

All experiments are conducted using the CIFAR-10 dataset, which consists of 60,000 color images (32x32 pixels) in 10 classes, with 50,000 training and 10,000 test samples. Images are normalized to the [-1, 1] range. For the input perturbation experiment, images transformations are introduced.

4. Experiments

4.1. Experiment Conditions

The experiment is conducted in the following four conditions:

- **Baseline:** Plain CIFAR-10 inputs and labels with no modifications.

- **Label Shuffle:** Training labels are randomly shuffled.

- **Label Noise (20%):** One fifth of training labels are replaced with incorrect labels.

- **Input Perturbation:** The following sequence of manipulations are applied to input images: random horizontal flip with probability 0.5, random vertical flip with probability 0.5, Gaussian blur, Color jitter, and Random erasing.

4.2. Training Settings

All models are trained using the Adam optimizer with a learning rate of 0.001 and a batch size of 32. Each experiments is run for 30 epochs. Cross-entropy loss is chosen as the training objective. Model performance is evaluated on the unmodified test set, and accuracy is reported as the primary metric.

Condition	Accuracy (%)
Baseline	78.61
Label shuffle	10.00
Label noise	73.41
Image perturbation	59.85

Table 1. Performance of models under four different conditions.

5. Results

5.1. Analyzing accuracy

Table 1 shows percentages of correctly labeled images by the model, under the four experiment conditions. The performance of the model was the highest when the input data was not modified with a precision of 78. 61%. With a portion of the labels noised, the model was still able to correctly identify 73. 41% of the train data, which is a decrease in accuracy but bearable. When train images undergone perturbation, accuracy dropped to 59.85%. Remarkably, the model identified one tenth of the images when trained with randomly shuffled labels.

5.2. Analyzing confusion matrix

Figure (a) through (d) shows the confusion matrices of each experiment. In most cases, the elements on the main diagonal are among the highest, since those entries indicate correct guesses of the trained model. In confusion matrix of label shuffle condition, every entry is zero but the last column. This indicates that the model identified all the test data as a truck.

6. Conclusion

In this study, we trained a CNN model with CIFAR-10 image dataset with four different conditions. We were able to observe that the performance of the model decreases as the noise in the train data increases. This suggests the importance of building a good dataset.

References