TP3 - Small Data and Deep Learning

Deep Learning in Practice

Tong ZHAO (tong.zhao@eleves.enpc.fr)

1   Introduction

In this homework, we focus on training deep neural networks on a dataset with few available labels. Several methods are implemented and then tested on CIFAR10. We use 100 images to train all models and then evaluate the models using the official test dataset of CIFAR10. Here we show a table comparing the results of all methods after 100 epochs.

Model	Pretained	Data Augmentation	Train Accuracy	Test Accuracy	Whole Set Accuracy
ResNet18	No	No	100.0	24	24.35
VGG11	No	No	100.0	27.67	27.63
Finetuned ResNet18	Yes	Yes	77.0	50.4	50.61
ResNet18	No	Yes	43.0	25.52	25.54
VGG11	No	Yes	79.0	27.49	27.26
DCGAN + SVM	Yes	No	100.0	27.7	-
Wavelet	No	Yes	89.0	29.4	28.96
Weak supervised GAN	No	Yes	100.0	16.2	15.4

In the following sections, we discuss the advantages and disadvantages of each method and their application scenes.

2   VGG Model

Advantages

- The model has a good performance and is used as baseline in many applications.

- The model contains a huge number of small 3x3 convolution filters, which helps to get a better feature representation.

Disadvantages

- The model is heavy. It takes longer to train the model and to do the inference than ResNet.

- The model suffers from vanishing gradient problem.

Application It is usually considered as a baseline model. It can be used if we have enough computational energy.

3 ResNet

Advantages

- The model can go deeper without optimization issue.

- The net structure can be easily modified by stacking residual blocks.

- The training and the inference are fast.

Disadvantages

- The feature maps have the same size thus we need to pad the data after each convolutional operations.

Application It is a very popular choice and is thus widely used in many domains.

4 Finetuned Model

Finetuning a model is a great choice for the cases where we don't have enough labeled training examples. It introduces extra information and improves a lot the prediction score.

Advantages

- The model contains information from a large dataset, which makes the model robust and powerful.

- The result is much better than the one without pretraining.

Disadvantages

- All the examples should be resized to the input size of the pretrained network, which may introduces some artifacts. (e.g. 32 to 224)

- Sometimes there is not a proper pretrained model for our task.

Application This method is prefered once we have an appropriate pretrained model.

5 Data Augmentation

Advantages

- Data augmentation helps us to reduce data overfitting.

- It can make the model more robust to noises.

Disadvantages

- The transformed image may be unreal then it behaves like an outlier.

- Small images are more sensitive to transformations.

Application It is used in most cases. We need to choose appropriate transformations according to the dataset.

6 DCGAN Features

Advantages

- Once we have a pretrained model, the features can be extracted directly from the model.

- The features have a good quality. We can achieve a good score even with a linear SVM.

Disadvantages

- It is complicated to train a GAN. Many hyper-parameters should be chosen carefully.

- Few pretrained models are available on the internet.

Application The features from DCGAN have a nice quality. This method can be used if we have a pretrained discriminator of DCGAN.

7 Wavelet Net

Advantages

- Wavelet preprocessing helps us to reduce noises from data thus the model will be more robust.

- The calculation is fast and it can be easily integrated into a deep learning framework.

Disadvantages

- It increases the channels of the input thus it consumes more energies during the training process and inference.

Application It is a good choice if we train a model from scratch.

8 Weakly Supervised Model

Advantages

- We can befinit from the large unlabeled dataset thus more information are provided to the model.

- Unlabeled data can help us to reduce overfitting issue.

Disadvantages

- Normally this kind of method needs more computational energies.

- The choice of models and the definition of loss functions are not obvious.

Application It can be considered if we have powerful machines.