



Report

Deep Learning

Author:
Manh Tu VU

Supervisor:
MarieBEURTON-AIMAR

July 31, 2017

Contents

1	Introduction	2
2	Preparing dataset	2
2.1	Fetch all images	2
2.2	Remove duplicate images	2
2.3	Remove broken images	3
3	Unsupervised Deep Learning	3
3.1	Trying to run the implement code	3
3.2	Preparing we own image dataset	3
4	Supervised Deep Learning	3
4.1	Train a CNN model with a small dataset	4
4.1.1	Preparing small dataset based on original dataset	4
4.1.2	The caffe model for this CNN	4
4.1.3	The slover definition	4
5	Result and Analysis	4

Abstract

In this project, we use Deep Learning method to automatic classify images from <https://heobs.org> into 4 classes, include:

Heritage

A place of cultural, historical, or natural significance for a group or society.

Beings

Any form of life, such as a plant or a living creature, whether human or other animal.

Scenery

Any form of landscapes which show little or no human activity and are created in the pursuit of a pure, unsullied depiction of nature, also known as scenery.

Other

Any other type of image that doesn't represent a photograph, such as painting, illustration, any object.

1 Introduction

2 Preparing dataset

2.1 Fetch all images

The entire image dataset described on the text file "photos.txt" line by line. Each line includes the image id and image description.

```
5a36f382-dbdf-11e6-95fd-d746d863c3eb | Những người ăn xin | vie
5a36f382-dbdf-11e6-95fd-d746d863c3eb | Mendiants | fra
17be8122-dbe0-11e6-860c-5fea02802d0a | Chợ Cũ (3) | vie
17be8122-dbe0-11e6-860c-5fea02802d0a | Vieux marché (3) | fra
400286c8-dbe1-11e6-bb4d-ff975c68de04 | Ngân hàng Đông Dương | vie
400286c8-dbe1-11e6-bb4d-ff975c68de04 | La Banque de l'Indochine | fra
```

In order to get the image dataset, we have to fetch each image one by one by join the image id with heobs cdn url <https://cdn.heobs.org/photo/>. For example, with the first line in the record above, we have the following URL:

```
https://cdn.heobs.org/photo/5a36f382-dbdf-11e6-95fd-d746d863c3eb
```

We wrote a small python script to automatic read this text file & download images one by one. Totally, we have 144564 images in our dataset.

2.2 Remove duplicate images

In the "photos.txt", some image has two languages and then, it consumes two lines. As the record above, we have six lines, but actually, a half of them was duplicated.

Because of the duplicate images doesn't help deep learning anything otherwise consume more

time to train. So, to reject those images, before fetching each image, we check if this image id already exists or not.

After removed duplicate, we have 142459 images left.

2.3 Remove broken images

After downloaded & look around all images, we found that it has a lot of broken images, which can't be displayable. So, we write a small script to filter all of those broken images automatically. Finally, we have 89850 images left in our dataset.

3 Unsupervised Deep Learning

Because of when classifying images by hand, it has some special case when one image may refer to more than one class. Thus, we need machine help us to make decisions by comparing two images are the same class or not. We also want to separate images of one class into multi unknown sub-classes. So, by using Unsupervision Deep Learning, we want to let the machine to classify a set of images unlabeled into some unknown classes.

After some research, we found the Unsupervised Deep Embedding for Clustering Analysis[1] paper, which propose a new method that simultaneously learns feature representations and cluster assignments using deep neural networks to classify unlabeled images. They also provide the implement code at <https://github.com/piiswrong/dec>

3.1 Trying to run the implement code

We already tried to run this implement code with the latest Caffe version (Jul 2017), but it doesn't work because this Caffe model required some deprecated parameters. However, because of the code delivery with Caffe and Docker, so, we continue to try with Docker and this Caffe version.

3.2 Preparing we own image dataset

4 Supervised Deep Learning

Supervised deep learning requires us to provide a training dataset, which includes a list of images labeled. However, our image dataset is not, and we can't classify the entire dataset by hand. So, we propose the following method, which includes three steps:

- Train a Convolution Neural Network (CNN) model with a small dataset based on the original dataset, which labeled by hand.
- Use the CNN model trained above to classify the entire original dataset
- Review & Train a CNN model with the original dataset.

4.1 Train a CNN model with a small dataset

4.1.1 Preparing small dataset based on original dataset

We created this dataset by random 1000 images from the original dataset and classify them by hand into four classes:

- Heritage
- Beings
- Scenery
- Other

4.1.2 The caffe model for this CNN

Our model is reuse from [bvlc_reference_caffenet](#) model, which is a replication of AlexNet with a few modifications. The original bvlc reference caffe net was designed for a classification problem with 1000 classes. However, we just need to classify into 4 classes. So, we change num output of the last InnerProduct layer from 1000 to 4.

4.1.3 The slover definition

Our model using Stochastic Gradient Descent solver method. We run our model with 4000 iterators, drop learning rate every 500 iterators and take a snapshot every 500 iterators.

```
net: "caffe_model/caffenet_train.prototxt"
test_iter: 500
test_interval: 500
base_lr: 0.001
lr_policy: "step"
gamma: 0.1
stepsize: 500
display: 50
max_iter: 4000
momentum: 0.9
weight_decay: 0.0005
snapshot: 1000
snapshot_prefix: "caffe_model/snapshot"
solver_mode: GPU
```

5 Result and Analysis

References

- [1] A. F. Junyuan Xie, Ross Girshick, "Unsupervised deep embedding for clustering analysis," *ICML'16 Proceedings of the 33rd International Conference on International Conference on Machine Learning*, vol. 48, 2016.