CIFAR10 ML Tutorial:

training with Caffe and quantizing/pruning with DeePhi

Daniele Bagni
DSP / ML Specialist

XILINX.

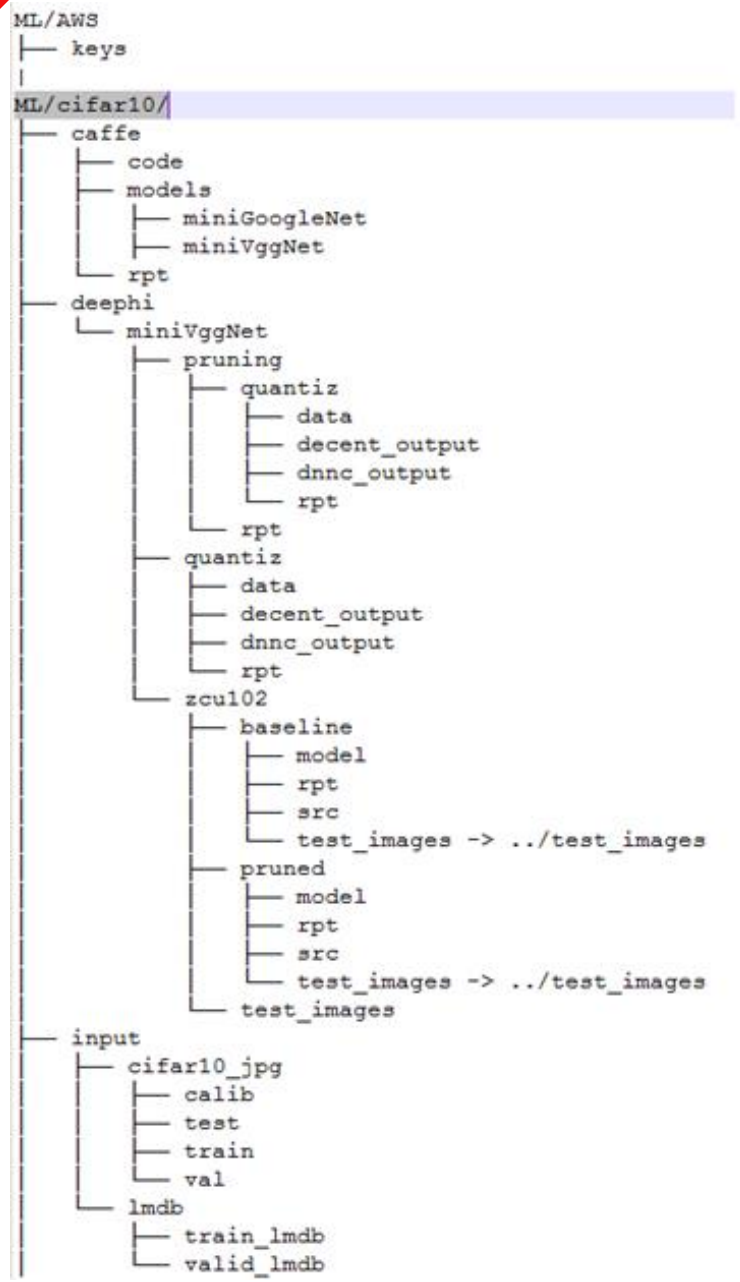
Why CIFAR10?

- CIFAR10 dataset is small enough to carry experiments about ML with different CNNs on any PC, even with a small GPU support
  - of course, having a powerful GPU would let you save a lot of time during training
- 10 classes of objects to be classified, so general enough to be meaningful
  - with the **same template** I have built another tutorial based on **Cats-vs-Dogs dataset** and AlexNet CNN
- 60000 RGB images of 32x32 size
- 50000 images in the training set and 10000 in the testing set to be organized in 3 databases:
  - **`train_lmdb`**: 50000 images in LMDB database for the forward/backward training process
  - **`valid_lmdb`**: 9000 images in LMDB database for the validation step during training process
  - **`test:`** 1000 images in JPEG plain format for the prediction measurements after training
  - Note that the last 2 datasets are created by the 10000 images of the original testing dataset
  - All the images are randomly shuffled before forming the database

Rationale

- 2 custom CNNs: **miniVggNet** and **miniGoogleNet**
  - custom CNNs with same layers of VGG and GoogleNet, just less deep
- ML tutorial on training with **Caffe** (or a fork of it):
  - Building the training, validation and testing databases based on CIFAR10 dataset
  - Training from scratch different models and plotting learning curves
  - Computing predictions on the etesting database
- ML tutorial on Quantization and Pruning with **DeePhi DNNDK** tools
- A lot of useful hyperlinks to Reference material
- WARNINGS:
  - 1) The larger the GPU RAM size and the shorter the training time
  - 2) miniVggNet can be trained also without a GPU, but the CPU time could be from 10x to 70x the GPU time, depending on your system
  - 3) miniGoogleNet requires a GPU with more than 2GB otherwise the training time explodes (if you do not have one, you can use a **p2.xlarge EC2 on AWS**, as shown in the following)

Overall project directories structure



I have organized the project in the following subdirectories:

1) there are 3 different Caffe forks in **`/caffe`**, but you need only one

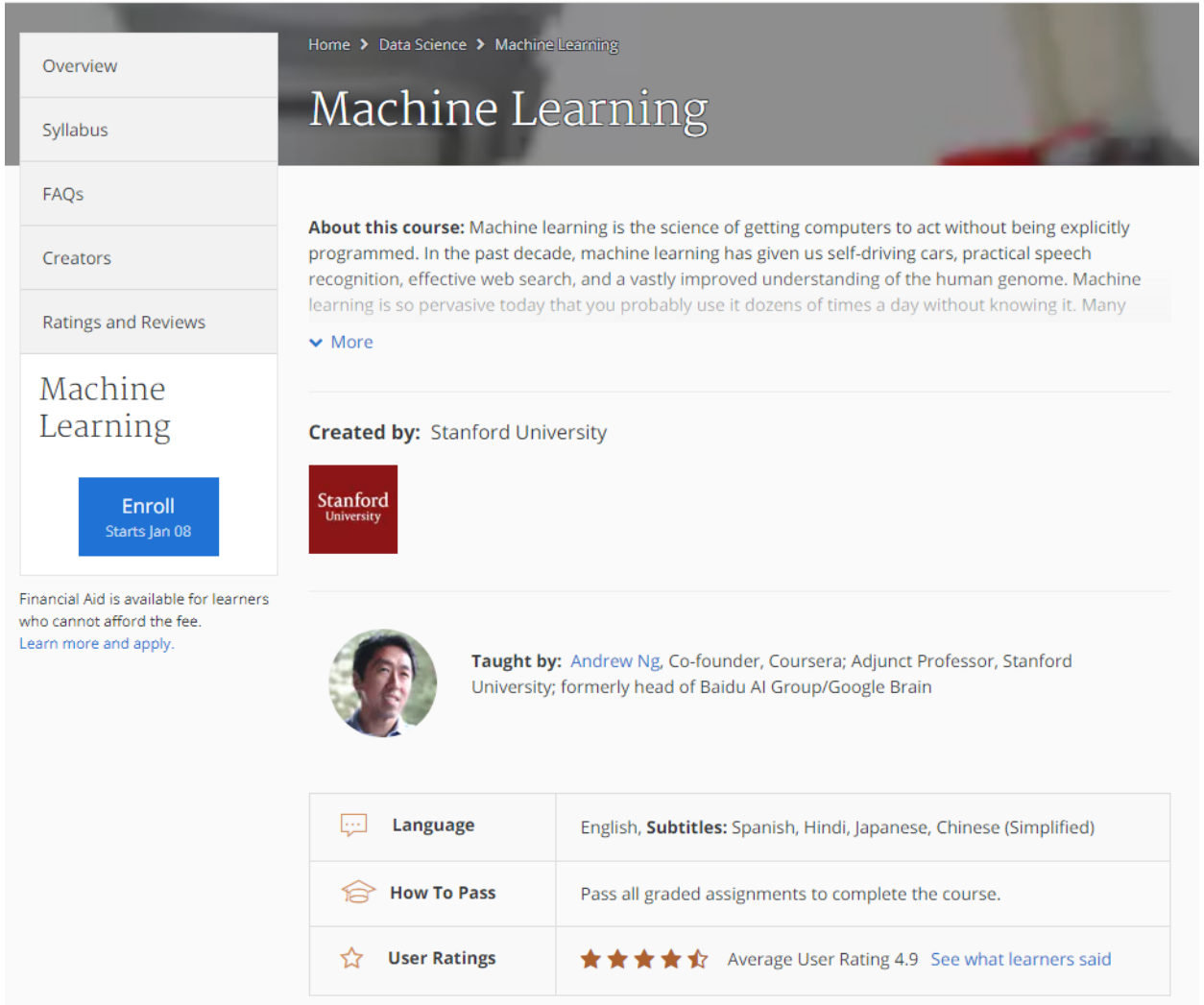
2) the CIFAR10 project in placed in **`<home>/ML/cifar10`**

3) the DeePhi tools are placed in **`<home>/ML/DNNDK`**

4) my AWS keys are placed in **`<home>/ML/AWS/keys`**

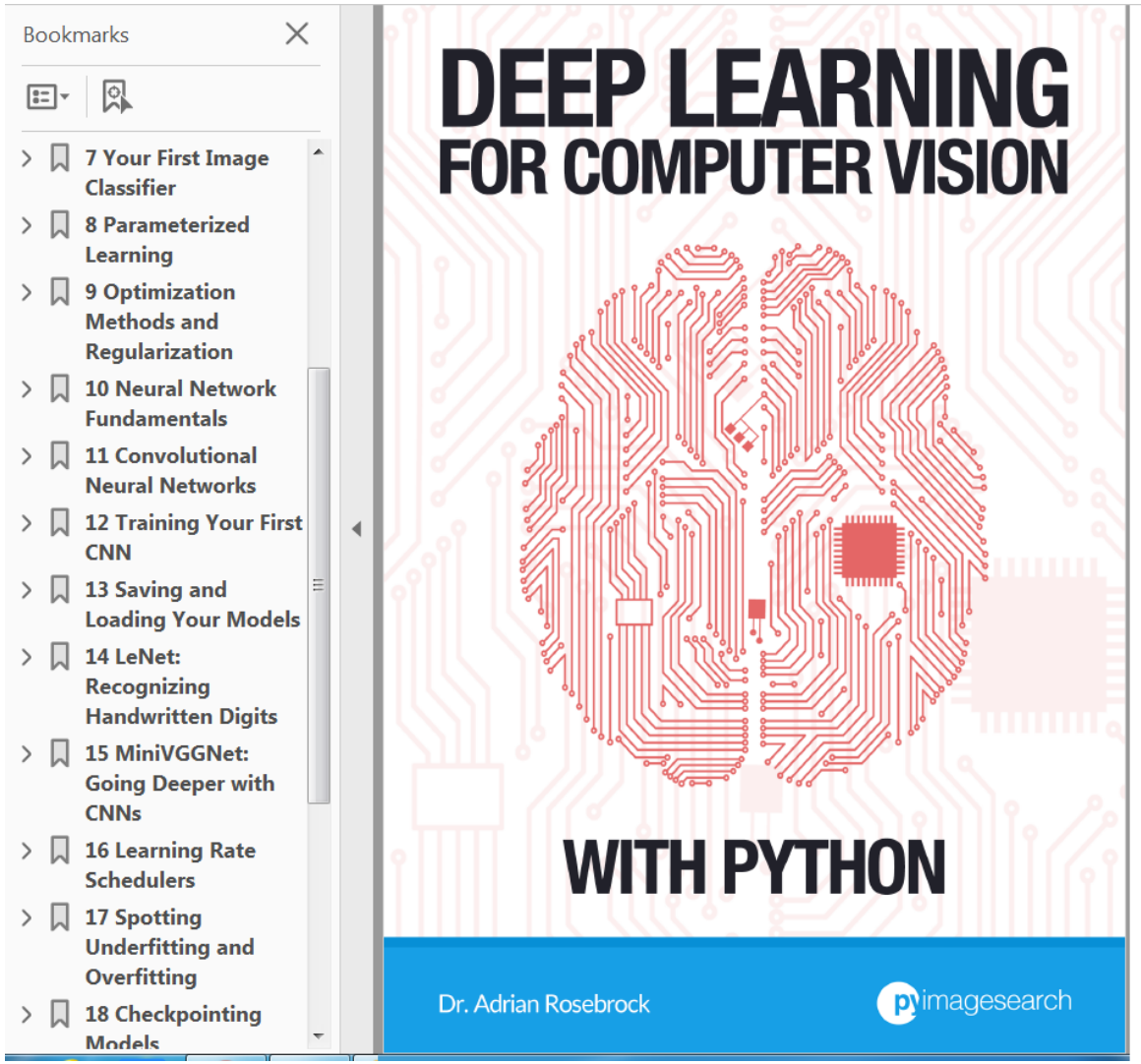
References

https://www.coursera.org/learn/machine-learning



- On-line training (costs <100$)
  - with online MATLAB tests to be passed
- 3 months MATLAB license
- 11 weeks course
  - I needed 8 working days full time
- Andrew NG is a widely cited name in the ML academic world

Deep Learning for Computer Vision with Python



- 2 eBooks:
  - Starter Bundle and Practitioner Bundle
  - About 300$ cost
- 3rd eBook
  - ImageNet Bundle
  - 300$ cost standalone
- Fully based on Python 2.7
- It comes with different VMs (including an AMI on AWS)
- Supplementary Material website (you need to register) with updates
- **GREAT MATERIAL!**

Basics on ML / CNNs

- https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/
- https://medium.com/technologymadeeasy/the-best-explanation-of-convolutional-neural-networks-on-the-internet-fbb8b1ad5df8
- https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/
- https://jalammar.github.io/visual-interactive-guide-basics-neural-networks/
- http://cs231n.stanford.edu/
- http://cs231n.github.io/convolutional-networks/
- http://www.wisdom.weizmann.ac.il/~vision/courses/2016_1/DNN/files/TA_lecture.pptx
- https://github.com/BVLC/caffe/wiki/Using-a-Trained-Network:-Deploy
- https://rdipietro.github.io/friendly-intro-to-cross-entropy-loss/

Good ML/CNNs material based on Python: 1

- http://paddlepaddle.org/docs/develop/book/02.recognize_digits/index.html
- http://karpathy.github.io/neuralnets/
- http://cs231n.github.io/
- http://shengshuyang.github.io/A-step-by-step-guide-to-Caffe.html
- http://rodriguezandres.github.io/2016/04/28/caffe
- https://github.com/humphd/have-fun-with-machine-learning
- http://alexthompson.ai/Setting-up-Caffe-on-AWS/
- http://deepdish.io/2015/04/28/creating-lmdb-in-python/
- https://jihongju.github.io/2017/05/10/caffe-filler/
- https://stackoverflow.com/questions/35529078/how-to-predict-in-pycaffe
- https://github.com/davidstutz/caffe-tools
- https://github.com/BVLC/caffe/wiki/Using-a-Trained-Network:-Deploy

Good ML/CNNs material based on Python: 2

- https://github.com/Franck-Dernoncourt/caffe_demos
- http://christopher5106.github.io/deep/learning/2015/09/04/Deep-learning-tutorial-on-Caffe-Technology.html
- https://github.com/nitnelave/pycaffe_tutorial/blob/master/03%20Custom%20solver.ipynb
- https://mohanadkaleia.com/caffe-tutorial-build-train-and-predict/
- https://github.com/adilmoujahid/deeplearning-cats-dogs-tutorial

About Keras/TensorFlow

- **https://www.pyimagesearch.com/2016/11/14/installing-keras-with-tensorflow-backend/**
- https://keras.io/
- https://www.tensorflow.org/tutorials/layers
- http://www.pythonexample.com/search?q=transfer+learning+keras
- https://nicolovaligi.com/converting-deep-learning-model-caffe-keras.html
- https://hackernoon.com/learning-keras-by-implementing-vgg16-from-scratch-d036733f2d5
- http://www.deepvisionconsulting.com/it/from-keras-to-caffe-2/
- https://github.com/balancap/SSD-Tensorflow

About HDF5 with Python

- http://cs231n.github.io/python-numpy-tutorial/
- https://docs.python.org/3/library/stdtypes.html#dict-views

- https://support.hdfgroup.org/HDF5/Tutor/HDF5Intro.pdf
- http://www.h5py.org/
- https://confluence.slac.stanford.edu/display/PSDM/How+to+access+HDF5+data+from+Python

Agenda

- Part 1. Caffe minimum background (CIFAR10_1_Caffe-Background.pdf)
- Part 2. Caffe Flow with CIFAR10 (CIFAR10_2_Caffe-Flow.pdf)
- Part 3. Training with Caffe
  - a) miniVggNet (CIFAR10_3a_Caffe-Training_miniVggNet.pdf)
- Part 4. DeePhi ML design flow (Installing / Quantizing / Pruning)
  - a) miniVggNet (CIFAR10_4a1_Deephi-Quantiz_miniVggNet.pdf, CIFAR10_4a2_Deephi-Pruning_miniVggNet.pdf)