CAFFE MODEL TRAINING - II

NRUPESH PATEL

nrupesh.patel@sjsu.edu
SJSU ID: 011425271

SAN JOSE STATE UNIVERSITY

REQUIREMENT

Training a Caffe model was necessary for following reasons:

- No specific model trained for classifying or detecting garbage.

- Model used in current solution is not open-source.

FINE-TUNING

Guidelines for Fine-tuning a model:

- Learn the last layer first (earlier layer weights won't change very much in fine-tuning)
    - Caffe layers have local learning rates: lr_mult
    - Freeze all but the last layer for fast optimization, i.e., lr_mult=0
    - Stop if good enough or keep fine-tuning other layers
    - This will speed up training times

- Alternatively, we can leave all learning rates as they are and increase the last two layers
    - Last layer by 10x
    - Second to last by 5x

- Reduce the learning rate
    - Drop the initial learning rate (in the solver_file.prototxt) by 10x or 100x

FINE-TUNING (CONT..)

Training requires:

- `train_val.prototxt`: defines the network architecture, initialization parameters, and local learning rates
- `solver.prototxt`: defines optimization/training parameters and serves as the actual file that's called to train a deep network

Fine-tuning Command:

```
./build/tools/caffe train -solver solver.prototxt –weights
trained_binary_model.caffemodel
```

Guide & Code: https://github.com/Nrupesh29/Clean-City-Research

Note: Require Lab access for executing training process on GPU enabled systems.

References

[1] Mittal, Gaurav, et al. "SpotGarbage: smartphone app to detect garbage using deep learning." *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2016.

[2] https://medium.com/@alexrachnog/using-caffe-with-your-own-dataset-b0ade5d71233#.m03nhhn5b

[3] https://github.com/christopher5106/FastAnnotationTool

[4] http://www.pyimagesearch.com/2016/11/28/macos-install-opencv-3-and-python-2-7/

[5] https://www.quora.com/How-do-I-fine-tune-a-Caffe-pre-trained-model-to-do-image-classification-on-my-own-dataset

[6] https://frankzliu.com/experimenting-with-different-penultimate-layers-in-caffe/

[7] http://rodriguezandres.github.io/2016/04/28/caffe/#dataset-preparation