How to start doing research in Deep Learning

An Tran

http://antran89.github.io/

whoami

- An Tran, PhD student in NUS, homepage: http://antran89.github.io/
- Research areas: Computer Vision, Machine Learning
- Research interests: Video classification, deep learning, action detection.

Contents

- Building infrastructure
- Software deployment
- Updating current researches

Building infrastructure

Building your own infrastructure

- GPUs are your workhorse.
- NVIDIA Titan X is good for DL.

Having a free GPU

- NVIDIA hardware donation program: https://developer.nvidia.com/academic_ hw_seeding
- Should apply through your supervisor.
- Having multiple GPUs is good for your research.

# Leverage your old system, if possible

- Planning out power of your system:
  - Mainboard ~ 300W
  - CPU ~ 300W
  - 2 GPUs ~ 300W + 250W
  - Total power: 1150W
  - PSU: 1300W with 90% efficient => 1170W
  - So still have chance to be in short of power, but it is rare.

Considering your GPUs

- Dimensions fit into your workstation.
- Power requirements.
- Power cables: 6-pin + 8-pin

Note: Must connect correct pin in order to

have enough power for Titan X GPUs.

- GPU's memory: Titan X has 12 GB.
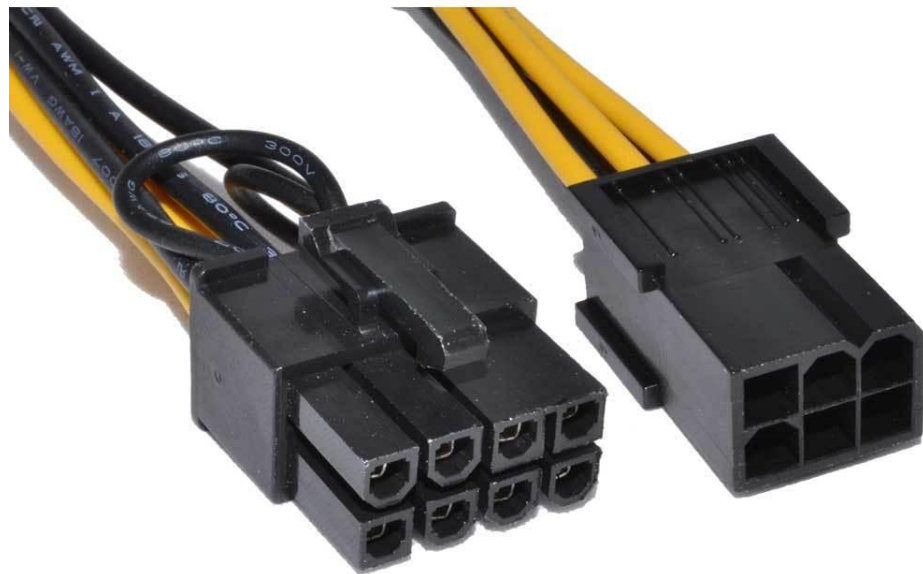
Cooling your GPUs

- System will restart if GPU's temperature reaches 80°C.
- If your workstation does not have a good cooling system, you can hack GPU's fan speed.
   - http://antechcvml.blogspot.sg/2015/12/how-to-control-fan-speed-of-multiple.html

Controlling clocks of GPUs

- To decrease power consumptions, heats, we can underclock GPU, turn off GPU auto-boost.
- To increase performances, we can overclock GPU.
- Easy to manage through *nvidia-smi*.

http://devblogs.nvidia.com/parallelforall/increase-performance-gpu-boost-k80-auto boost/

GPU in my research work

- Computing optical flow for videos: Faster than running on CPUs.
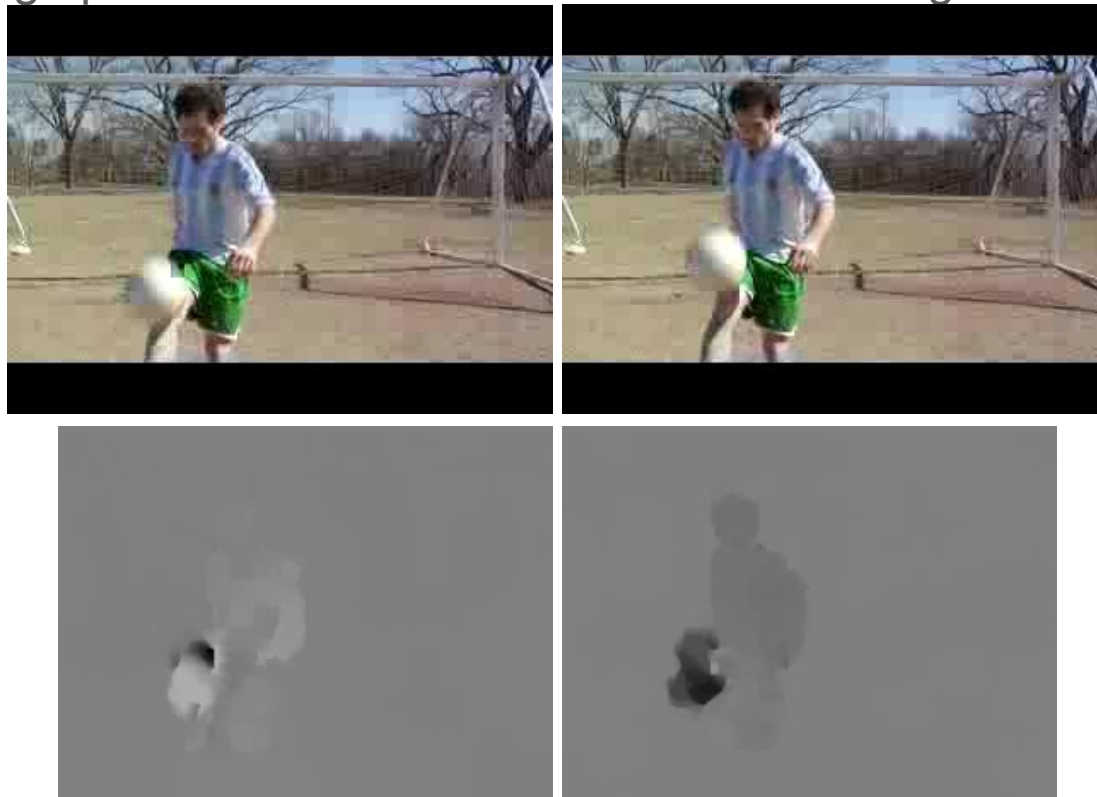
OpenCV

GPU in my research work

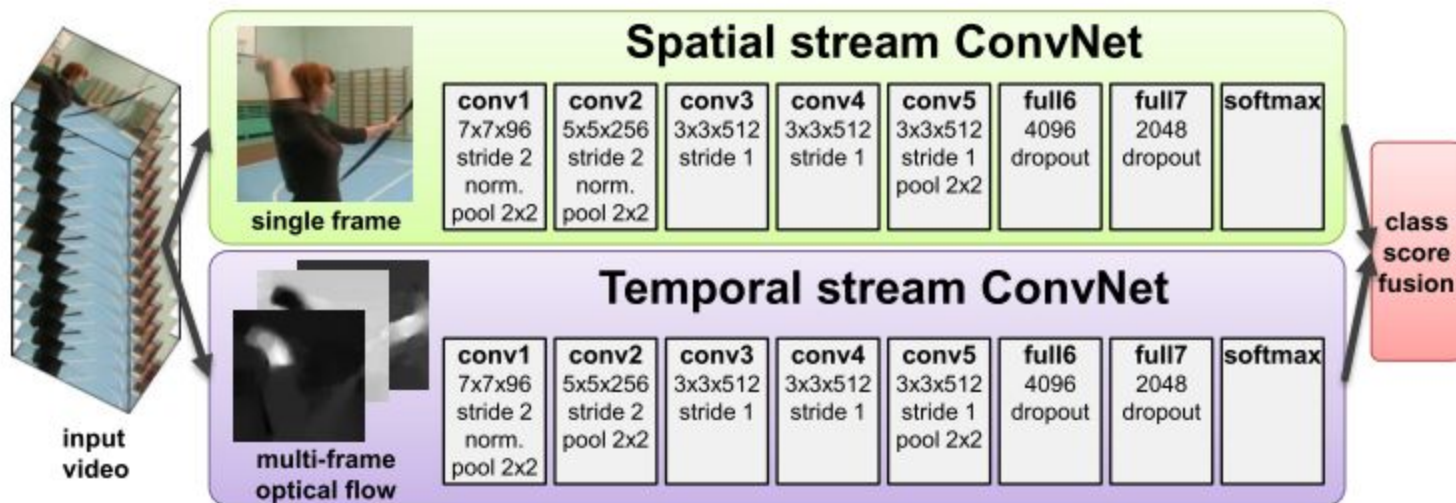- Current work is trying to improve performance of two-stream conv network.

Figure 1: **Two-stream architecture for video classification.**

Simonyan, K., & Zisserman, A. (2014). Two-Stream Convolutional Networks for Action Recognition in Videos. *arXiv Preprint arXiv:1406.2199*, 1–11. Retrieved from http://arxiv.org/abs/1406.2199

Software deployments

Deep Learning frameworks

- BLVC Caffe
  - Vision problems
  - Active community and developers
- Theano
  - Symbolic maths, more than just Deep Learning
  - Auto-differentiation
  - Wrappers: Keras, Lasage
- Torch
  - Twitter Cortex autograd (auto-diff)
- TensorFlow
  - Auto-differentiation
  - Faster time to production scale, native to Google Cloud Platform
- NVIDIA DIGITS:
  - A DL web app built on top of Caffe and Torch.

Updating current researches

Research updates

● Most challenging part of research.

Helpful tools

- http://gitxiv.com/, sharing new papers and codes.
- Building basics of DL: http://goodfeli.github.io/dlbook/
- Awesome lists of Deep Learning, Deep Vision
- Reading related DL papers from good conferences: CVPR, ICCV, ECCV, ACCV, NIPS, ICML, ICLR.
- Subscribe to http://arxiv.org/
- Joining an online community: G+ Deep Learning group.
- Reading materials from a DL course.
- Following some experts in the field on social networks, blogs, Quora.
- Balancing between reading and doing experiments. Applying techniques to solve my problems.

Thank you

Questions.