ENSC 424 - Multimedia Communications Engineering

Project

Jie Liang
Engineering Science
Simon Fraser University
jiel@sfu.ca

Project Topic

- The topic must be related to multimedia communications or deep learning

- "Default" project:
  - Based on the simple video codec that we will develop through assignments

  - Introduce some advanced coding tools from latest standards, such as H.264 and H.265
    - Intra prediction, Sub-pixel motion estimation, SVC, MVC, 3D video coding …

Past Project Topics

- Video coding:
  - motion estimation, scene change detection, multiple reference frames, intra prediction, Multiview video coding, Scalable video coding
- WebRTC (a web-based video communication protocol)
- Kinect depth image coding
- High dynamic range (HDR) image
- Deep learning projects
- ……

Deep Learning

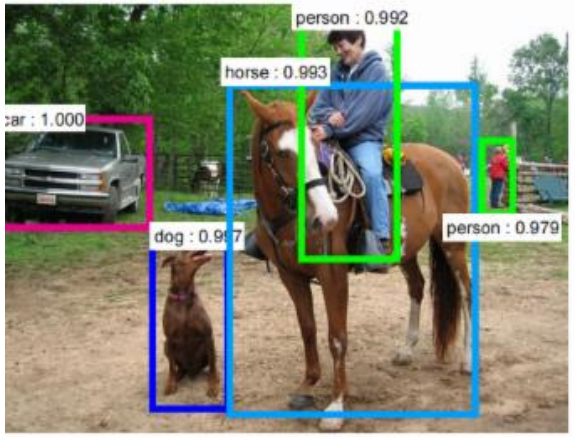
- Pioneered by Canadians:
  - Geoffrey E. Hinton: UoT (Google now)
  - Yoshua Bengio: U Montreal
  - Yann LeCun: NYU. Director of Facebook AI Research
  - 2019 Turing Award

- Record-breaking results in many fields
- Need powerful GPU to do serious research
- Possible to work with toy examples



Deep Learning

- Some deep learning topics:
  - Object detection
  - Object classification
  - Semantic segmentation
  - Generative adversarial network (GAN)
  - Face recognition
  - Object tracking
  - Multi-object tracking
  - Compression of DL network
  - Image/video captioning
  - Image/video compression
  - ……
- Many open-source codes



Deep Learning

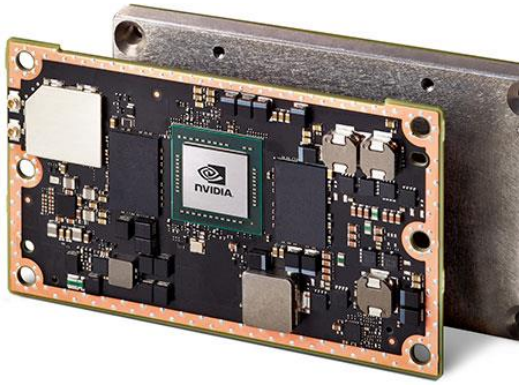
- Types of projects:
  - Design a simple neural network and train it
  - Train an existing network from scratch
    - Compare with pre-trained models
  - Implement DL on different platforms
  - Apply DL to some applications

Deep Learning

- Deep learning algorithms have millions of parameters
- The training needs millions of data and could take days
- The training and testing need powerful GPU cards
- Examples:
- GeForce GTX 1050 Ti: $169
  - 768 CUDA cores 4GB memory
- GeForce GTX 1070 Ti: $399
  - 1920 CUDA cores 8GB memory
- GeForce GTX 1080: $699
  - 2560 CUDA cores 8GB memory
- NVIDIA Titan X Pascal Architecture: US$1200
  - 3584 CUDA cores 12GB memory
- Caution: Some cards need special computer case and power.

Deep Learning

- Embedded Systems: NVIDIA TX2 board
  - US$299 after education discount
  - Has built-in camera
  - A Linux computer!
  - 256 NVIDIA CUDA Cores
  - HMP Dual Denver 2/2 MB L2 + Quad ARM® A57/2 MB L2
  - 8 GB Memory 32 GB Flash St
- Can run deep learning algorithms
- But not powerful enough to train



- ENSC has 25 sets.

Deep Learning

- Other Embedded Systems:
- Raspberry Pi 3:
  - https://medium.com/dt42/run-object-detection-using-deep-learning-on-raspberry-pi-3-1-55027eac26c3
  - https://medium.com/dt42/run-object-detection-using-deep-learning-on-raspberry-pi-3-2-66f43609bc85
  - https://github.com/DT42/BerryNet
- https://developer.movidius.com/
  - Movidius Neural Compute Stick

- Google Coral TPU:

Deep Learning

- Amazon DeepLens for developers:
  - A fully programmable video camera

- Amazon DeepRacer: (pre-order)
  - Use reinforcement learning

Deep Learning

- Many mobile phones have powerful GPUs:

- ARM Compute Library:
- https://community.arm.com/graphics/b/blog/posts/arm-compute-library-for-computer-vision-and-machine-learning-now-publicly-available

- Qualcomm AI:
- https://www.qualcomm.com/invention/cognitive-technologies/artificial-intelligence

- Apple Core ML:
- https://developer.apple.com/machine-learning/

Deep Learning

- Amazon EC2 Elastic GPUs
  - https://aws.amazon.com/ec2/elastic-gpus/

Amazon EC2 Elastic GPUs Pricing

With Amazon EC2 Elastic GPUs, you only pay for what you use. The pricing for Elastic GPUs is listed below.

US East (Ohio):

Elastic GPU Size	GPU Memory	Price
eg1.medium	1 GiB	$0.050/hour
eg1.large	2 GiB	$0.100/hour
eg1.xlarge	4 GiB	$0.200/hour
eg1.2xlarge	8 GiB	$0.400/hour

Where to find references?

- IEEE Xplore:
  http://ieeexplore.ieee.org/search/advsearch.jsp
  Search keywords, or browse journals and conferences.
  Accessible from SFU computers

  If you need to access from home, use the link at SFU library: http://www.lib.sfu.ca/
  - Click 'Journal articles & databases'
  - Enter 'ieee explore' into the text box below 'Find databases by database title and description:'
  - Click 'connect'

Where to find references?

- IEEE Xplore:
  http://ieeexplore.ieee.org/search/advsearch.jsp
  Search keywords, or browse journals and conferences.

- Some relevant journal publications:
  - IEEE Transactions on Circuits and Systems for Video Technology
  - IEEE Transactions on Image Processing
  - IEEE Transactions on Multimedia
  - ACM Transactions on Multimedia Computing, Communications and Applications

Where to find references?

- https://arxiv.org/
- Many latest papers
- Free access anywhere

- https://scholar.google.ca/
- Provide useful citation information

Where to find references?

- The Calls for Papers of major conferences are good places to know the latest topics in the field:

- Example: ICIP (International Conf on Image Processing)
  - Image/video coding and transmission
  - Image/video processing
  - Image formation
  - Image scanning, display, and printing
  - Image/video storage, retrieval, and authentication
  - Applications

Where to find references?

- ICME: International Conf Multimedia and Expo
  - Multimedia content analysis
  - Multimedia activity and event understanding
  - Multimedia search and retrieval
  - Mobile, location-based and other context-based multimedia
  - Social, user-generated, and cloud-based multimedia
  - 3D immersion and virtual reality
  - Multimedia security and forensics
  - Human computer interaction based on multimedia
  - Multimedia networking and communication
  - Multimedia coding and compression
  - Multimedia signal processing and enhancement
  - Multimedia systems, applications, services and implementations

Where to find references?

- ACM Multimedia:
  - Mobile Multimedia
  - Multimedia Systems and Middleware
  - Multimedia Telepresence and Virtual Reality
  - Multimedia Transport and Delivery
  - Multimedia for Education & Distributed Environments
  - Multimedia HCI and Quality of Experience
  - Music, Speech and Audio Processing in Multimedia
  - Multimedia Authoring and Enrichment
  - Multimedia Art, Entertainment and Culture
  - Multimedia Search and Recommendation
  - Social Multimedia
  - Emotional and Social Signals in Multimedia
  - Deep Learning for Multimedia
  - Multimodal Analysis and Description
  - Multimedia and Vision

Where to find references?

- CVPR: Computer Vision and Pattern Recognition
  - Image captioning
  - Object recognition and detection
  - Image restoration
  - Computational Photography
  - Actions and human pose estimation
  - Video understanding
  - Learning and CNN architectures
  - 3D reconstruction
  - Face recognition
  - 3D, stereo, matching and saliency estimation

- Other computer vision conferences:
  - ECCV
  - ICCV
  - NeurIPS

Source Codes

- JPEG 2000: http://www.openjpeg.org/
- H.264: http://iphome.hhi.de/suehring/tml/
- HEVC: http://hevc.hhi.fraunhofer.de/
- H.266/VVC: https://jvet.hhi.fraunhofer.de/

- Prof. Xin Li's List of Reproducible Research:
  - http://www.csee.wvu.edu/~xinl/source.html
  - Has links to many other source codes.

- Video Sequences:
  - http://trace.eas.asu.edu/yuv/
  - ftp://ftp.tnt.uni-hannover.de/testsequences/

Report Format & Grading

- Format: similar to a scientific publication
  - Title
  - Abstract
  - Introduction
  - Project description (can have multiple sections), including experimental results
  - Conclusion
    - Explanation of each team member' contribution.
    - Clearly state if you use any source codes from other places, and what changes you have made.
  - References
  - Software
- Length: No more than 20 pages, including references:
  - single column, single space, 12-point font.
  - Penalty for overlength report: 20% / page

Report Format & Grading

- Grading: the project is 60% of the course
  - Understanding of the topic
  - Difficulty and amount of work
  - Writing: clarity, correctness (no grammar/spelling errors)

- Always cite the references if you use something from other places
- Cheating caught in the project will fail the course!

Tips for the Projects

- Work on the project progressively:
  - Agile workflow
  - Build a simple but demo-able prototype first
  - Improve the feature of each component gradually
  - This way you can stop at any time and still have something to report
- Start early, don't wait till it is too late
- Define some milestones

Tips for the Projects

- Version control:
  - punch.cs.sfu.ca
  - Google Code
  - https://github.com/
    https://bitbucket.org/
- At least back up your codes frequently