ECE 599/692 – Deep Learning

Lecture 19 – Beyond BP and CNN

Hairong Qi, Gonzalez Family Professor
Electrical Engineering and Computer Science
University of Tennessee, Knoxville
http://www.eecs.utk.edu/faculty/qi
Email: hqi@utk.edu
Schedule

- Nov. 9: Beyond BP and CNN
- Nov. 14: Depth vs. Breadth
  - What’s beyond ReLU?
  - What’s beyond cross entropy?
- Nov. 16: Projects discussion
  - What’s beyond leaderboard?
- Nov. 21: How to design a new network structure?
- Nov. 28, 30, Dec. 5
  - Final project presentation (26 presentations)
From Big Data to Artificial Intelligence

• What is big data?
• The more the better?
• Can’t live without data?
• AlphaGo Zero vs. AlphaGo
  – No human input vs. Human knowledge
• Hinton’s two new papers on capsule networks (2017)
  – “I think the way we’re doing computer vision is just wrong… It works better than anything else at present but that doesn’t mean it’s right.”
  – Capsule (small groups of crude virtual neurons) networks for tracking different parts of an object
  – Abandon BP
  – “…humans should encode as little knowledge as possible into AI software, and instead make them figure things out for themselves from scratch…”
  – “The future of AI is determined by those graduate students who seriously doubt all what I have said.”

https://deepmind.com/blog/alphago-zero-learning-scratch/
A bit history

- 1943 (McCulloch and Pitts):
- 1957 - 1962 (Rosenblatt):
  - From Mark I Perceptron to the Tobermory Perceptron to Perceptron Computer Simulations
  - Multilayer perceptron with fixed threshold
- 1969 (Minsky and Papert):
- The dark age: 70’s ~25 years
- 1986 (Rumelhart, Hinton, McClelland): BP
- 1989 (LeCun et al.): CNN (LeNet)
- Another ~20 years
- 2006 (Hinton et al.): DL
- 2012 (Krizhevsky, Sutskever, Hinton): AlexNet
- 2014 (Goodfellow, Benjo, et al.): GAN

A bit of history - revisited

• 1956-1976
  – 1956, The Dartmouth Summer Research Project on Artificial Intelligence, organized by John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon
    We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College ... The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.
  – The rise of symbolic methods, systems focused on limited domains, deductive vs. inductive systems
    – 1976, the AI Winter

• 1976-2006
  – 1986, BP algorithm

• 2006-???
  – 2006, Hinton (U. of Toronto), Bingio (U. of Montreal, LeCun (NYU)
  – 2012, ImageNet by Fei-Fei Li (2010-2017) and AlexNet

https://en.wikipedia.org/wiki/Lighthill_report
ML vs. AI

- ML + inference $\rightarrow$ AI
Turing awardees in AI

- 1969, Marvin Minsky
- 1971, John McCarthy
- 1975, Allen Newell and Herbert A. Simon
- 1994, Edward Feigenbaum and Raj Reddy
- 2010, Leslie G. Valiant
- 2011, Judea Pearl

https://en.wikipedia.org/wiki/Turing_Award
Artificial Intelligence

• The three branches
  – Logic-based (Symbolic method)
  – Network-based
  – Behavior-based (Self adaptive and evolution)

• Artificial intelligence vs. Human intelligence
  – Logical
  – Linguistic
  – Spatial
  – Musical
  – Kinesthetic
  – Intra-personal
  – Inter-personal
  – Naturalist
  – Graphics
Government involvement

• China’s new generation artificial intelligence
  – Big data intelligence
  – Swarm intelligence
  – Multimedia (Cross-domain) intelligence (speech, image, text, natural language)
  – Human-machine coordination
  – Autonomous vehicle

• Applications
  – Computer vision
  – Speech recognition
  – Natural language processing
  – Human-machine interface
  – Robotics
Andrew Parker’s Light Switch Theory

- The Cambrian explosion: the explosion of life forms (550 million years ago)
- The theory:
  - It was the development of vision in primitive animals that caused the explosion.

Visual intelligence: Beyond ImageNet

- Fei-Fei Li’s talk at CNCC 2017
- Object recognition (ImageNet – Single Object Recognition)
  - \(0.28\) (2010) \(\rightarrow\) \(0.26\) (2011) \(\rightarrow\) \(0.16\) (2012) \(\rightarrow\) \(0.12\) (2013) \(\rightarrow\) \(0.07\) (2014) \(\rightarrow\) \(0.036\) (2015) \(\rightarrow\) \(0.03\) (2016) \(\rightarrow\) \(0.023\) (2017)
- Beyond object recognition \(\rightarrow\) Rich scene gist
  - The Visual Genome Dataset
  - Visual relationship \(\rightarrow\) Semantic scene retrieval \(\rightarrow\) Scene graph generation
- Beyond scene gist \(\rightarrow\) Vision + Language & Reasoning
  - The CLEVR Dataset
  - Image captioning \(\rightarrow\) Dense Captioning \(\rightarrow\) Paragraph generation
- https://www.leiphone.com/news/201710/CvdrhzTO0cndEArJ.html
What’s after ImageNet?

• **MS COCO (Microsoft Common Objects in Context)**
  – 4 Tasks: Detection challenge, Instance segmentation, Human keypoint challenge, Stuff segmentation
  – [http://cocodataset.org/#home](http://cocodataset.org/#home)

• **Places (MIT and CMU)**
  – 3 Tasks: Scene parsing, Instance segmentation, and Boundary detection
  – Winners (2017):
    – Scene parsing: CAS
    – Instance segmentation: Face++
Reference

- Xiangyang Shen, Microsoft, CNCC 2017
- Wen Gao, From Big Data to Artificial Intelligence, JDDiscovery, 2017, https://www.toutiao.com/a6485167794265522701/?tt_from=weixin&utm_campaign=client_share&app=news_article&utm_source=weixin&iid=9368489672&utm_medium=toutiao_android&wxshare_count=1