CS 6501: Deep Learning for Computer Graphics

Introduction / Overview

Connelly Barnes

Outline

- Course info:
 - Website
 - Goals / topics
 - Grading
 - Sign up sheet
- Background: machine learning, neural networks, computer vision
- Recent impact on computer graphics
- Examples of papers we will cover

Course Website

http://www.connellybarnes.com/work/class/2016/deep_learning_graphics/

- Other ways to find the website:
 - Linked to from UVa Collab
 - Linked to from Lou's list
 - Linked to from my website <u>connellybarnes.com</u>

Course Goals

- Learn math and computer science behind deep (artificial) neural networks
- Understand in a general way how deep learning has been applied to areas such as natural language processing, robotics, games, finance
- Understand in moderate depth how deep learning is used in computer vision
- Understand in greater depth how deep learning is used in graphics
- Gain experience building your own deep networks for visual computing problems (computer graphics, computer vision)

Course Topics (More Detail)

- Neural networks
 - Fully connected
 - Convolutional
 - Recurrent
 - Autoencoders
 - How to initialize
 - How to train
 - How to prevent problems: e.g. vanishing/exploding gradients

Grading

- Quizzes in class (15%)
- Student paper presentations (20%)
- Programming assignments (35%)
- Final course project (30%)
- No mid-term

Sign Up Sheet

- Everyone sign up for student presentation of a paper
- Also enter your username if you do not have CS department account
- Obtain sign up sheet URL from Collab.

What is Machine Learning?

 "The field of study that gives computers the ability to learn without being explicitly programmed." – Arthur Samuel

What is Machine Learning?

- 3 broad categories:
 - Supervised learning: computer presented with example inputs and desired outputs by a "teacher", goal is to learn general rule that maps inputs to outputs.
 - Unsupervised learning: No output labels are given to the algorithm, leaving it on its own to find structure in the inputs.
 - Reinforcement learning: An agent determines what actions to best take in an environment to maximize some notion of cumulative reward.

Background: Artificial Neural Networks

- "An artificial neural network is a network inspired by biological neural networks, which are used to estimate or approximate functions that can depend on a large number of inputs that are generally unknown."
 Wikipedia
- In this class, we can call these "neural networks" for short.

Historical Background: Neural Networks

- 1943: McCulloch and Pitts
 - Warren McCulloch (a psychiatrist and neuroanatomist)
 - Walter Pitts (a mathematician)
- Examined computational capabilities of networks made of very simple neurons.
- McCulloch-Pitts network fires if sum of excitatory inputs exceeds threshold, as long as it does not receive an inhibitory input.
- Can construct AND and OR gates, so one can construct any logical function.



Historical Background: Neural Networks

• 1949: Psychologist Donald Hebb created learning hypothesis for neurons in the brain.



Neuron diagram from Wikipedia

Hebbian Learning (1949)

"Let us assume that the persistence or repetition of a reverberatory activity (or "trace") tends to induce lasting cellular changes that add to its stability.... When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased." - Hebb

"Simultaneous activation of cells leads to pronounced increases in synaptic strength between those cells." – Wikipedia

- 400 pixel camera
- Designed for image recognition
- Weights encoded in potentiometers (variable resistors)
- Weight updates during learning performed by electric motors



- In machine learning, a classifier identifies which out of a set of categories a new observation belongs to, on the basis of a training set of data containing observations whose category is known.
- Perceptron is a binary classifier (classes are 0 and 1):

$$f(x) = egin{cases} 1 & ext{if} \, w \cdot x + b > 0 \ 0 & ext{otherwise} \end{cases}$$

• Weights w and bias b are learned from the training data.



• Can learn linearly separable patterns.



Diagram from Wikipedia

Multilayer Neural Networks (1965)

- Alexey Ivakhnenko trained multi-layer networks (as deep as 8 layers) that had polynomial activation functions.
- First deep learning?

Multilayer Perceptron (1960s)



Backpropagation (1960-1980s)

- Popular method used in training neural networks that have differentiable units.
- Early forms: Henry Kelley (1960), Arthur Bryson (1961), Stuart Dreyfus (1962).
- Modern variant: Seppo Linnainmaa (1970)
- If there is a loss function or objective that is to be minimized (reflecting the accuracy of the classifier), backpropagation finds the gradient of objective with respect to each parameter in the network.
- This allows optimization methods to improve the weights: for example by going "downhill" (negative gradient direction).

Backpropagation and Gradient Descent



Diagram from Wikipedia

Neocognitron (1980s)

- Kunihiko Fukushima
- Biologically inspired:
 - Local features extracted by S-cells
 - Deformation of those features is tolerated by C-cells
- Inspired convolutional neural networks

Convolutional Neural Networks (1990s)

Yann LeCun, Leon Bottou, Yoshua Bengio, Patrick Haffner



Diagram from Wikipedia

Fast GPUs + Big Data => Deep Learning (2000s)

- More data means we can train deeper networks
- Fast GPUs mean we can run the training process more efficiently
- Dramatic improvement in performance in computer vision and natural language processing.



What is Computer Vision?

• "An interdisciplinary field that deals with how computers can be made to gain high-level understanding from digital images or video.

From the perspective of engineering, it seeks to automate tasks that the human visual system can do." - Wikipedia

Computer Vision Example: Face Detection



Face Detection, Viola & Jones, 2001

What is Computer Graphics?

- "Pictures and movies created using computers --- usually referring to image data created by a computer specifically with help from specialized graphical hardware and software." – Wikipedia
- Computer graphics and computer vision are inverses of each other
 - Graphics: input: knowledge/models/descriptions, output: images
 - Vision: input: images, output: knowledge/models/descriptions

Computer Graphics Example: Special Effects

<u>Creating Gollum</u>



History of Computer Vision: Hubel & Wiesel (1959)

Hubel & Wiesel, 1959

History of Computer Vision: Block World (1963)

Block world

Larry Roberts, 1963

History of Computer Vision: David Marr (1970s)

Stages of Visual Representation, David Marr,

History of Computer Vision: 1980s

David Lowe 1987: 3D Object Recognition from Single 2D Images

History of Computer Vision: 1980s

Figure 13: The three-dimensional wire-frame model of the razor shown from a single viewpoint.

Figure 14: Successful matches between sets of image segments and particular viewpoints of the model.

David Lowe 1987: 3D Object Recognition from Single 2D Images

Scale-Invariant Feature Transform (SIFT, 1999)

"SIFT" & Object Recognition, David Lowe, 1999

Face Detection (2001)

Face Detection, Viola & Jones, 2001

HoG and Deformable Part Models (2000s)

Histogram of Gradients (HoG) Dalal & Triggs, 2005 Deformable Part Model Felzenswalb, McAllester, Ramanan, 2009

ImageNet (2009)

- Jia Deng, Wei Dong, Richard Socher, Lia-Jia Li, Kai Li, Li Fei-Fei
- Currently over 10 million images in tens of thousands of categories
- <u>http://image-net.org/explore</u>
- ImageNet Large Scale Visual Recognition Competition (ILSVRC)

Computer Vision: Deep Learning (2012-)

- In 2012, the AlexNet classifier gained a significant performance improvement on the ILSVRC by using deep neural networks.
- Deep learning on ImageNet <u>surpassed human performance</u> for object recognition in 2015
- Recently, computer vision predominantly uses deep learning.

Question: What is the Inverse of Artificial Intelligence?

Question: What is the Inverse of Artificial Intelligence?

<u>Artificial stupidity</u>

Recent Impact on Computer Graphics

- As computer vision works better, we can now use techniques from computer vision and deep learning for the generation of synthetic pictures in computer graphics.
- Increasing number of papers after 2014.
- Predict this will be a trend:
 - It is expensive and challenging to manually author computer graphics programs, especially those related to content in virtual worlds, e.g. shaders, geometric and texture variations, etc.
 - Cameras are cheap and there is an enormous amount of visual data to learn from to construct plausible virtual replicas.

Examples of Papers We Will Cover: Text

<u>Random generation of Shakespeare text</u> using recurrent neural networks (RNNs):

PANDARUS:

Alas, I think he shall be come approached and the day When little srain would be attain'd into being never fed, And who is but a chain and subjects of his death, I should not sleep.

 How Google DeepMind used reinforcement learning to train a neural network to <u>play Atari games</u>.

Examples of Papers: Computer Vision

• How convolutional neural networks can be trained to <u>segment video</u> into semantically meaningful regions.

Examples of Papers: Computer Vision

• How to use deep learning to <u>find body correspondences</u>.

• How to use deep learning to <u>replace faces</u>.

• How to use deep learning to synthesize character motion.

• How to use deep learning to <u>semantically label objects in 3D</u>.

- Shaders are (typically short) programs that run on each pixel during 3D rendering, to determine the pixel's color based on lighting, material, and camera parameters.
- Recent research demonstrates that we can use deep learning to replace the process of manually writing shaders.

Discussion Question: Suggested Topics / Papers?

- Are there any additional topics or papers people suggest?
- You can also email them to me.