CS590-DVC
Deep Visual Computing

Daniel G. Aliaga
Fall 2020
Tentative Main Topics

• Reprise of fundamentals: machine/deep learning, vision, graphics
• Tricks of the trade: synthetic vs real, supervised vs unsupervised, latent space transformations
• Deep Image Segmentation
• Deep Shape Segmentation
• Neural Rendering, Deep Light Transport
• Deep Single Image Reconstruction
• Deep 3D Reconstruction
• Deep Urban Modeling

• You tell me? Other topics?
Workload

• (10%) Attend and participate in lectures
• (40%) Assignments:
  • A1: Deep Test
  • A2: Deep Segmentation
  • A3: Deep Modeling
• (20%) Midterm
• (10%) Papers Presentation
• (20%) Final Project
• [no final exam]
Deep Visual Computing

• Since the beginning, it turns out visual computing and machine learning have been deeply connected

• Do you know why?

• Lets see... (get it: lets “see”)

A long time ago in a computer far, far inferior to your phone, it all began...

-Daniel Aliaga, August 25, 2020
ENIAC

• Completed in 1945
• Was called a “Giant Brain” by the press
• Cost $6.3M of today’s dollars

• However, computers then lacked a key prerequisite for intelligence: they could barely remember...they only executed a few commands
Logic Theorist (1956)

• A program designed to mimic the problem solving skills of a human

• From 1957-1974, AI flourished and failed and flourished...

• In 1968, A. Clarke and S. Kubrik said “by the year 2001 we will have machines with intelligence that matches or exceeded humans’s”

• In 1970, Marvin Minsky (MIT) said that in 3-8 years “we will have a machine with the general intelligence of an average human being”
AI Timeline

1938-1946: Golden Age of Science Fiction
1950: Can Machines Think? - Alan Turing
1955: Logic Theorists, the first AI program, is invented
1956: Dartmouth Summer Research Project on Artificial Intelligence
1965: Moore's Law
1968: "By the year 2001 we will have machines with intelligence that matched or exceeded human's" - Arthur Clarke and Steve Kubrik
1969: DARPA funds AI at MIT
1970: "From 9-18 years we will have a machine with the general intelligence of a human being" - M. Minsky
1980: Edward Feigenbaum introduces expert systems
1982: Japan's Fifth Generation Computer Project
1986: Navlab, the first autonomous car, is built by Carnegie Mellon
1997: Deep Blue defeats Gary Kasparov in chess
1997: First publicly available speech recognition software developed by Dragon Systems
1980s

• Expert systems became popular: dedicated systems

• “Deep learning techniques” was a coined phrase but with diverse meanings...

• I was around then, and even a paid undergraduate researcher in a major AI lab
  - our job was to create a robot that could be programmed remotely and could execute algorithms for navigating and deciding how to avoid obstacles (e.g., walls and boxes)
(Single Layer) Perceptron


- Model based on the human visual system
Fig. 2B. Venn diagram of the same perceptron (shading shows active sets for $R_1$ response).
Algorithm 1: Perceptron Learning Algorithm

**Input:** Training examples \( \{x_i, y_i\}_{i=1}^m \).

Initialize \( \mathbf{w} \) and \( b \) randomly.

**while** not converged **do**

### Loop through the examples.

**for** \( j = 1, m \) **do**

#### Compare the true label and the prediction.

\[
\text{error} = y_j - \sigma(\mathbf{w}^T \mathbf{x}_j + b)
\]

### If the model wrongly predicts the class, we update the weights and bias.

**if** \( \text{error} \neq 0 \) **then**

#### Update the weights.

\[
\mathbf{w} = \mathbf{w} + \text{error} \times \mathbf{x}_j
\]

#### Update the bias.

\[
b = b + \text{error}
\]

Test for convergence

**Output:** Set of weights \( \mathbf{w} \) and bias \( b \) for the perceptron.
Perceptrons

• Book by M. Minsky and S. Papert (1969)

• Was actually “An Introduction to Computational Geometry” – thus visual as well

• Commented on the limited ability of perceptrons and on the difficulty in training multi-layer perceptrons

• (Back propagation appeared in 1986 and helped a lot!)
Try this...

https://playground.tensorflow.org/

- First try something linear
- Then try something more complex...
Deep Learning Timeline

- 1940: Dark Era Until 1940
- 1943: Neural Nets McCulloch & Pitt
- 1948: Perceptron Rosenblatt
- 1950: Computing Machinery and Intelligence Alan Turing
- 1958: XOR problem Minsky & Papert
- 1960: ADALINE Widrow & Hoff
- 1969: Self Organizing Map Kohonen
- 1974: Backpropagation Werbos (and more)
- 1980: Neocogitron Fukushima
- 1982: Hopfield Network John Hopfield
- 1985: Boltzmann Machine Hinton & Sejnowski
- 1986: Restricted Boltzmann Machine Smolensky
- 1986: Multilayer Perceptron Rumelhart, Hinton & Williams
- 1986: RNNs Jordan
- 1990: LeNet Lecun
- 1997: LSTM Hochreiter & Schmidhuber
- 1997: Bidirectional RNN Schuster & Paliwal
- 2006: Deep Belief Networks - pretraining Hinton
- 2006: Dropout Hinton
- 2014: GANs Goodfellow

Made by Favio Vázquez
Reprise: Computer Vision

• In 1959, Russell Kirsch and colleagues developed an image scanner: transform an image into a grid of numbers so that a machine can understand it!

• One of the first scanned images:
  
  (176x176 pixels)
1982

- David Marr, British neuroscientists, published influential paper
  “Vision: A computational investigation into the human representation and
  processing of visual information”

  Among many things, he gave the insight that vision is hierarchical (i.e., primal
  sketch, 2.5D, and then 3D recognition)

  (now at CVPR, the Marr Prize exists)
1999

• David Lowe’s work “Object Recognition from Local Scale-Invariant Features” indicated a shift to feature-based visual object-recognition (instead of full 3D models as Marr proposed)

  • Scale-Invariant Feature Transform (SIFT)

  • and many subsequent derivatives
2010

- ImageNet Large Scale Visual Recognition Competition (ILSVRC) runs annually
  - 2010/2011: error rates were around 26% (using Lowe-style approaches)
  - 2012: the beginning of a new beginning – AlexNet – reduced errors to 16%!
AlexNet

- University of Toronto created a CNN model (AlexNet) that changed everything (Krizhevsky et al. 2012)
Just a note: 1980s

• Kunihiko Fukushima developed Neocognitron for visual pattern recognition which included several convolutional layers whose (typically rectangular) receptive fields had weight vectors (known as filters)

• This was perhaps the earliest deep and convolutional network
Just a note: 1989

• Yann LeCun applied backpropagation to Fukushima’s network and with other improvements released LeNet-5 – quite similar to today’s CNNs
ILSVRC (2011-2017)
Talk Slides

- Fei Fei and Deng – CVPR 2017
Reprise: Graphics

• First graphics **visual** image:
  • Ben Laposky used an oscilloscope in 1950s

(note: one of my undergrad senior projects was an oscilloscope based graphics engine)
Whirlwind Computer @ MIT

- Video display of real-time data:
1960s

- Ivan Sutherland used vector displays (=oscilloscope), light pens, and interaction
1965: The Ultimate Display...

• Fred Brooks using one of Ivan’s displays....the birth of VR/AR

• NOTE: Fred Brooks was on my PhD committee, I worked in his research group and my MS and PhD revolved around VR/AR and graphics.
1960s-today

• Fred Brooks had a huge role in CS and in Graphics
  • He defined the 8-bit byte
  • Revolutionized computing
  • Founded CS department at UNC Chapel Hill
  • Pioneer in AR/VR, architectural modeling, Design of Design, Mythical Man Month, and more!
  • Many awards: National Medal of Technology, Turing Award, etc...

• He is still with us, though almost 90 years old!
Computer Graphics History

- 1950: MIT Whirlwind (CRT)
- 1955: Sage, Radar with CRT and light pen
- 1958: Willy Higinbotham “Tennis”
- 1960: MIT “Spacewar” on DEC-PDP-1
- 1963: Ivan Sutherland’s “Sketchpad” (CAD)
- 1968: Tektronix storage tube
- 1968: Evans & Sutherland’s flight simulators
- 1968: Douglas Engelbart: computer mouse
- 1969: ACM SIGGRAPH
- 1970: Xerox GUI
- 1971: Gouraud shading
- 1974: Z-buffer
- 1975: Phong Model
- 1979: Eurographics
- 1981: Apollo Workstation, PC
- 1982: Whitted: Ray tracing
- 1982: SGI
- 1984: X Window System
- 1984: 1st SGI Workstation
- ->1995: SGI dominance
- ->2003: PC dominance
- Today: programmable graphics hardware (again)
Deep Learning in Computer Graphics

• Like in computer vision, since 2010’ish deep learning has revolutionized computational imaging and computational photography

• However, hand-crafted methods have significantly improved other domains such as geometry processing, rendering and animation, video processing, and physical simulations