UMassAmherst | Manning College of Information & Computer Sciences

682: Neural Networks: A Modern Introduction

Lecture 1: Introduction

Welcome to 682!

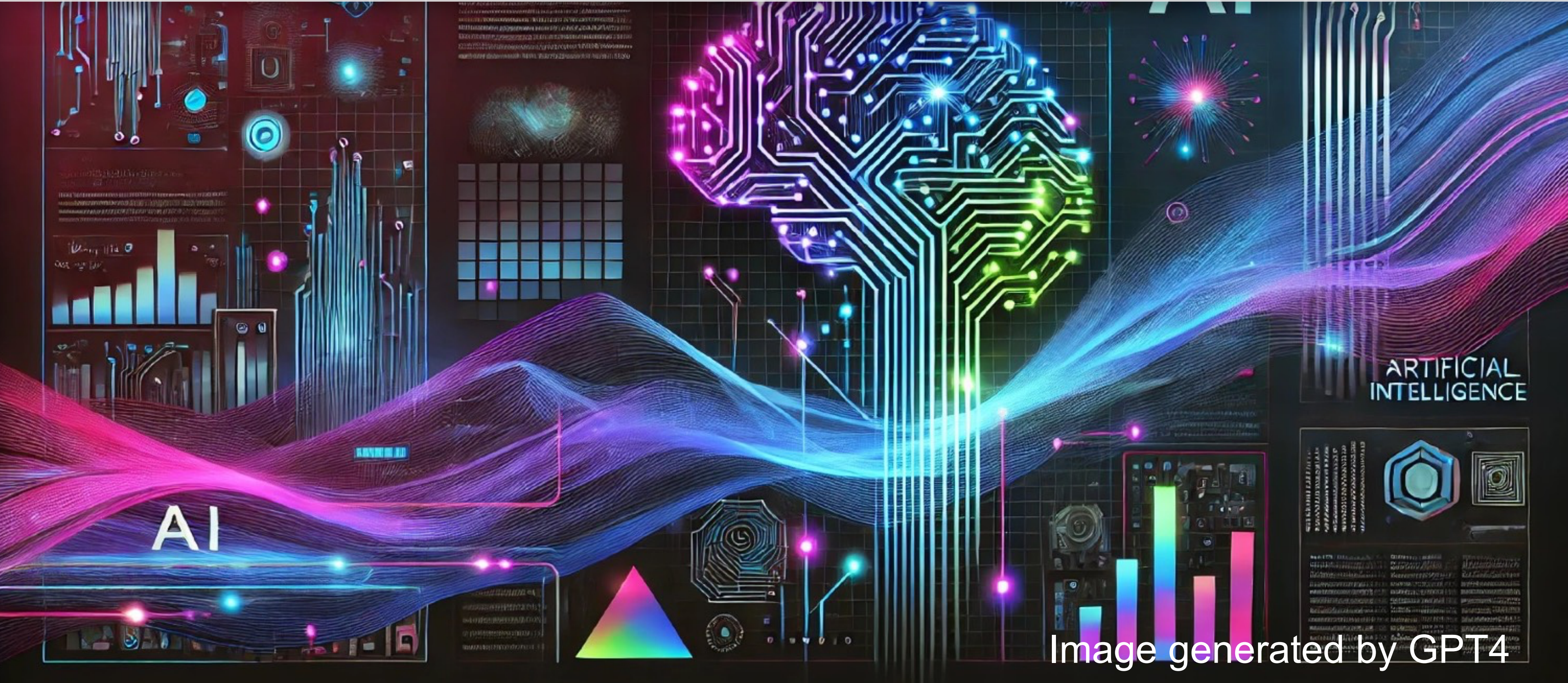


Image generated by GPT4

Who are we? Instructors

Subhransu Maji
Computer vision, AI,
AI4Science (ecology & remote sensing)

Chuang Gan
Computer vision, AI,
cognitive science, robotics

https://cvl-umass.github.io/compsci682-fall-2024

Who are we? TAs

Oindrila Saha

Chunru Lin

Junyan Li

Blossom Metevier

Ashish Singh

https://cvl-umass.github.io/compsci682-fall-2024

Who are you? Raise your hand if you are …

- First year student?
- Taking the first AI course?
- Interested in
  - Computer vision
  - Natural language processing
  - Robotics
  - Applications (remote sensing, medical imaging, climate change, etc.)
  - System building
  - Theory
  - Philosophy
  - …

Today's agenda

- 682 overview
- A brief history of computer vision and deep learning

Course page

https://cvl-umass.github.io/compsci682-fall-2024

UMassAmherst | Home | Lectures | Notes | Assignments | Policies | Project | Office Hours

COMPSCI 682 Neural Networks: A Modern Introduction

Note

- This is a tentative class outline and is subject to change throughout the semester.
- Regular lectures will be Tue & Th 1:00PM - 2:15PM, Thompson Hall, Room 106.
- There will be a few optional discussion sections organized by the TAs on Fridays (shown in green)
- Slides will be finalized after the lecture and Echo360 recordings accessible via Canvas.

Event Type	Date	Description	Course Materials
Lecture	Tuesday, Sep 3	Intro to deep learning, Historical context	[slides] [python/numpy tutorial] [software setup for assignments]
Lecture	Thursday, Sep 5	Image classification K-nearest neighbor Linear classification	
Lecture	Tuesday, Sep 10	Loss functions Optimization	

Optional Discussion Sections

- Fridays (date & location will be listed on the lecture page)
  - First one — 11-12am, CS 142 Friday, September 13
- Will cover background topics such as:
  - Python setup, and some basics such as
    - Slicing and broadcasting
    - Other parallelization techniques
  - Math techniques
    - Derivatives of vectors, matrices, etc.
    - Complex chain rule examples

Topics

- Intro to supervised learning
  - k-nearest neighbors, Support vector machines, Logistic regression for classification
- Feedforward neural nets
  - Network architecture, backpropagation, optimization, regularization, speed, etc
- Convolutional neural nets
- Beyond classification
  - Detection, segmentation, 3D understanding
- Visualization and understanding neural nets
- Other topics: generative AI, RNNs/Transformers, graphics, robotics, …
  - Will have some guest lectures

682 Neural Networks: A Modern Introduction

- Balance of theory vs. practice
  - Heavily tilted toward practice.
  - Examples:
    - Regularization will be used, but not much theory of it.
    - No proofs of convergence
  - Instead:
    - Develop applications “from scratch”
    - Build “layered” architectures from scratch so new models can be easily assembled
    - Implement popular add-ons such as batch normalization
    - Learn techniques for training and setting hyperparameters.

Topics

- Applications
  - Mostly **Computer Vision**: Object recognition in particular.
  - However, can easily be applied to other domains.
    - You will learn what you need to know to apply neural nets broadly.
  - Will cover some **Natural Language Processing** (or **Large Language Models**) this semester.

Topics

- What this course is *not*:
  - General course on machine learning
  - General course on graphical models
  - Not even a general class on deep learning!!!
    - No Bayes Nets
    - No restricted Boltzmann machines or deep Boltzmann machines
  - Not a computer vision survey class
    - No tracking, stereo, depth estimation, etc., etc.

Grading

- 3 Assignments: 15%, 15%, 20% = 50%
- Course project: 50% (teams of 2-3 members)
    - Proposal: 5%
    - Milestone: 15%
    - Final write-up: 25%
    - Presentation: 5% (in class)
    - We will have a lecture on project ideas and expectations later in the semester.
- Late Policy:
    - 7 free late days in total: use them as you see fit (no permission necessary)
    - Max 3 late days per homework.
    - Afterwards: 25% off per day late
    - Does not apply to the course project requirements (must be on time)

- Check course website for details and dates.

Assignment #1

- Will be posted soon on course website
- Due in 3 weeks (Thursday, Sept. 26, 11:55pm) (in GradeScope).
- It includes:
- Write/train/evaluate a kNN classifier
- Write/train/evaluate a Linear Classifier (SVM and Softmax)
- Write/train/evaluate a 2-layer Neural Network (backpropagation!)
- Requires writing numpy/Python code

Compute: Use your own laptops. Talk to TA if you don't have your own computer.

Communication

- **Piazza** for questions, announcements, etc.
  - Do not email us except for personal reasons
- **Course website** for syllabus, links to assignment downloads
  - https://cvl-umass.github.io/compsci682-fall-2024/
- **Gradescope** for homework and project submission
  - We will automatically enroll you
  - Automatically tracks late days, deadlines, etc.
  - Do not email us the submissions
- **Echo360 / Canvas**
  - For watching recorded lectures
  - For watching recorded discussion sections

Collaboration policy

Every year we deal with a number of cheating cases! Don't do it !!! Read and follow academic integrity policy at UMass. Some simple guidelines:

1. Let's start with an easy one. Don't copy any piece of the solution of any problem.
2. Never **look at** solutions to any of the homework problems. Most people who were caught cheating last semester claimed that they only "looked at" on-line solutions. This is NOT ALLOWED.
3. Do not look at discussions of the homework problems. These are likely to include methods for solving parts of the problem, which is cheating.
4. Don't look up pieces of the problem on Google. For example:
   a. "Computing the derivative of softmax"
   b. "Gradient updates for the multi-class SVM loss".
   Once you've done the search, you cheated. You are likely to see something you cannot forget. You can't "unsee" the answer once you've seen it.
5. Common sense. If you look at something on the web and it made the problem easier, then you're probably cheating. To be safe, stick to class materials, TAs, and Professors.

Advice

- Everyone knows you’re not supposed to cheat.
- What people don’t know is what you’re supposed to do when you’re desperate. Here’s some advice:

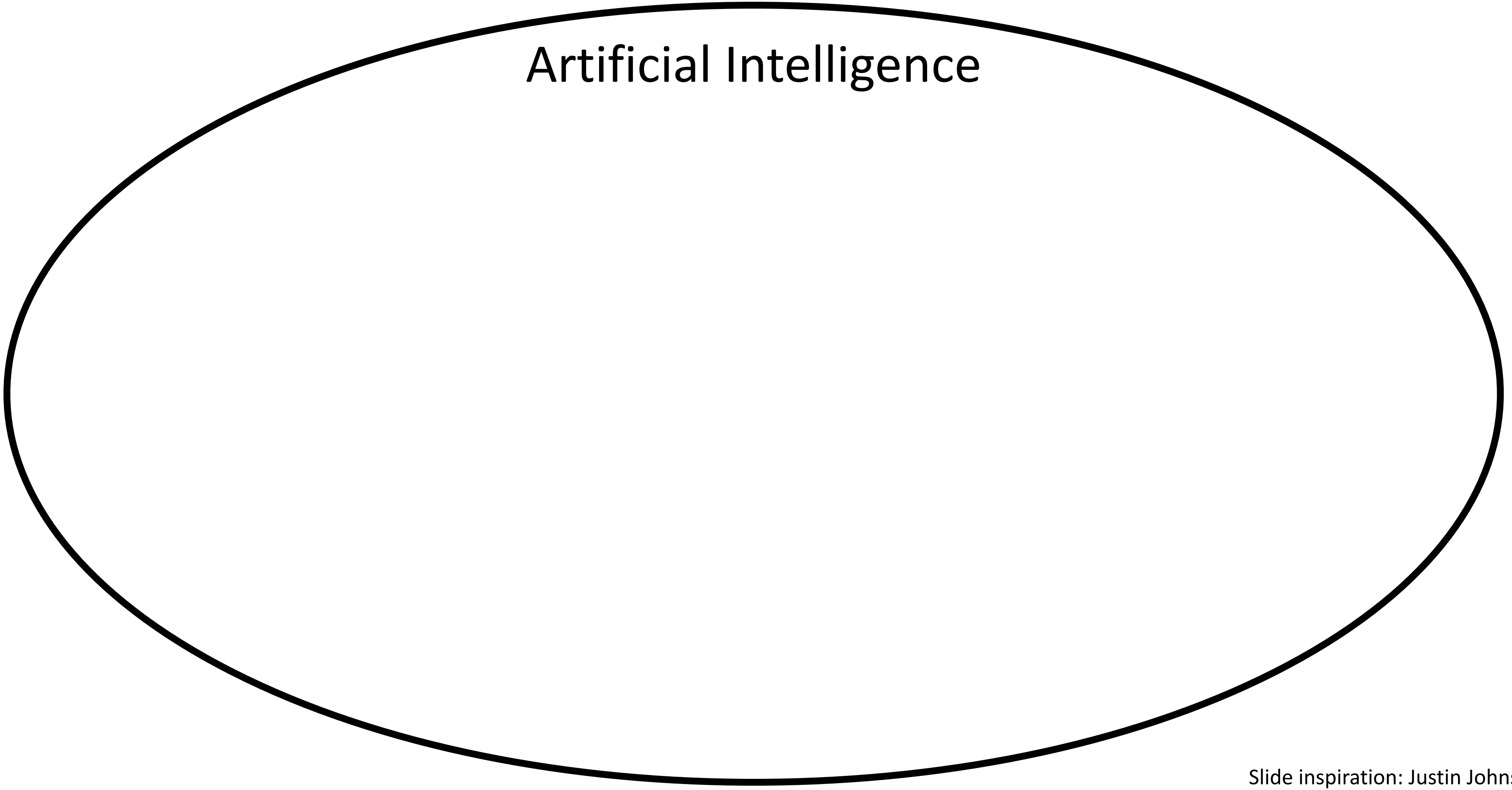
1) If you’re overloaded in the middle of the semester, consider dropping a class. Hopefully you can drop it without a “W”, but even a “W” is a lot better than an “F” and a record of cheating. A “W” will not influence your grade point average.
(I dropped the same class 4 times in grad school!)
2) Take a “0” on part of the problem set. Many people who did not do part of one problem set got an A-. Some people missed a whole problem set and still got a B for the course.

Questions about what is allowed

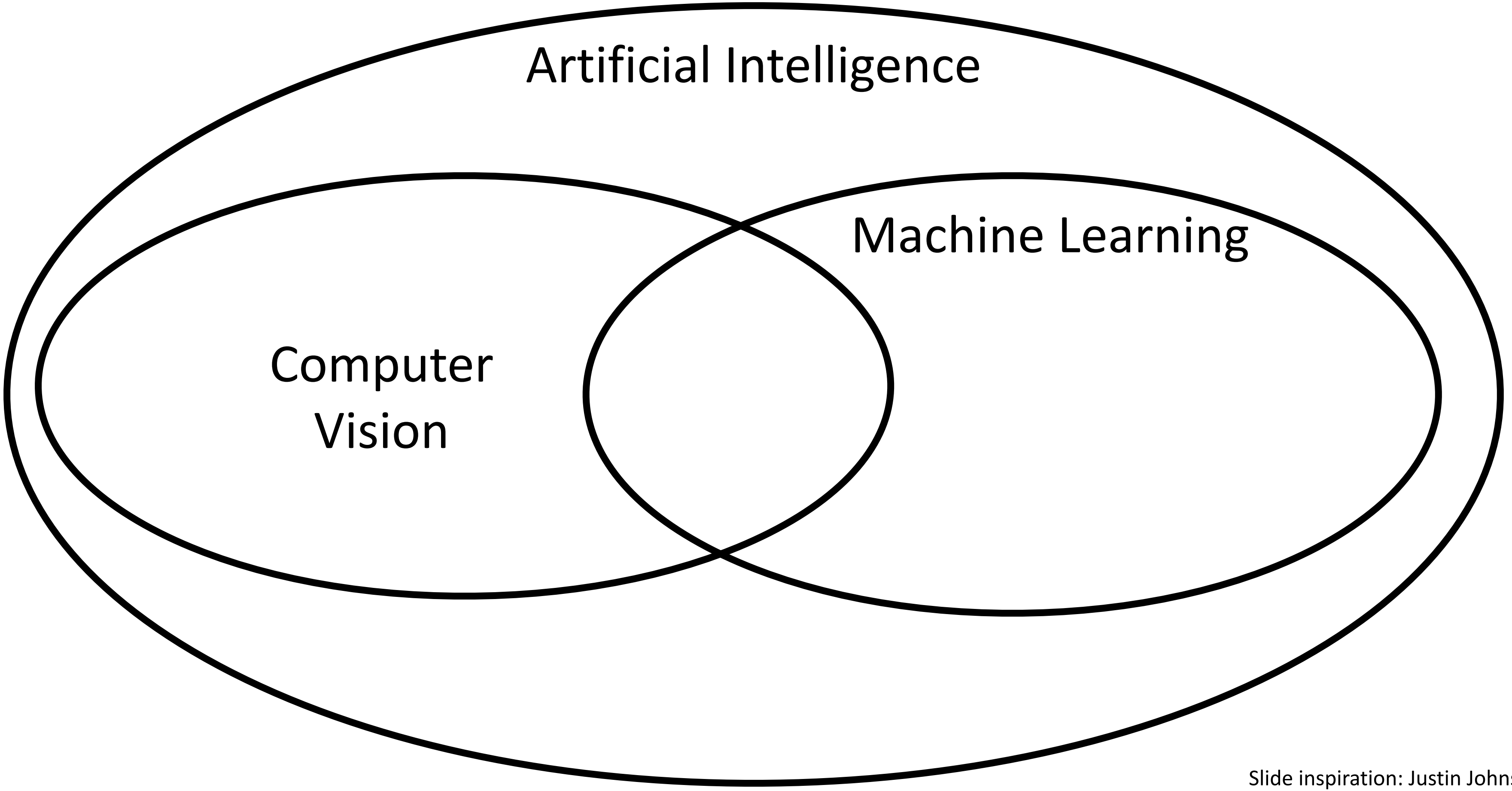
1. Question: Can I work with other students on the homeworks?
   Answer: You are allowed and encouraged to discuss high-level strategies. Include the list of students you discussed your solution with in the submission. However, code and answers must be written individually.
2. Question: Where can I get help?
   a. Look at the course notes
   b. Go to optional Friday sections
   c. Talk to the TAs
   d. Talk to the professor
3. Can I look at on-line materials that are not part of the course?
   a. Basically no. If you look at something and it's part of the solution, then you have cheated. So it's dangerous to go surfing around. Stick to the materials on the course web site. If there is something you want to look up, ask the TAs a question and we'll try to put materials on the course web site if it's appropriate.

Today's agenda

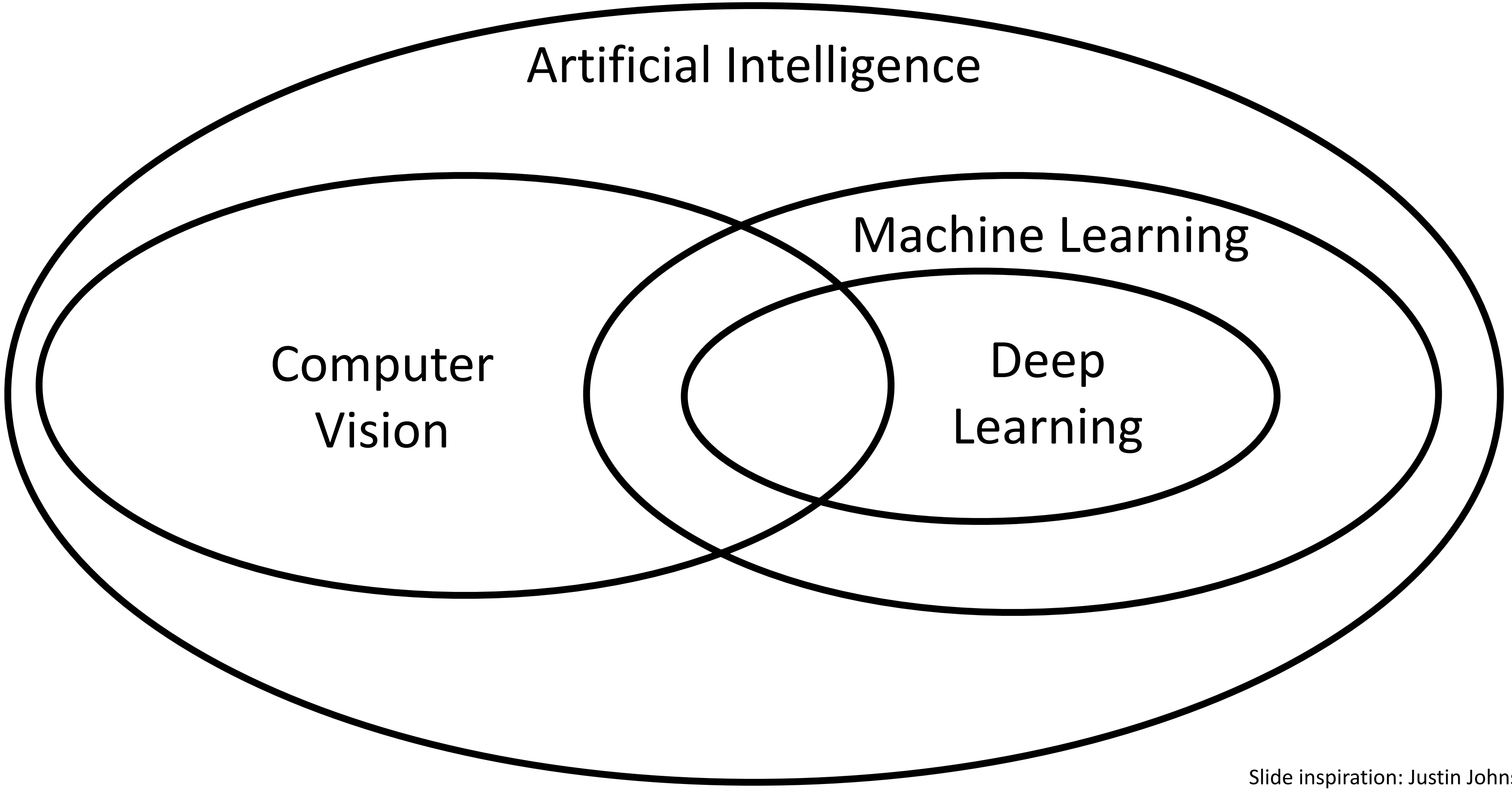
- 682 overview
- A brief history of computer vision and deep learning



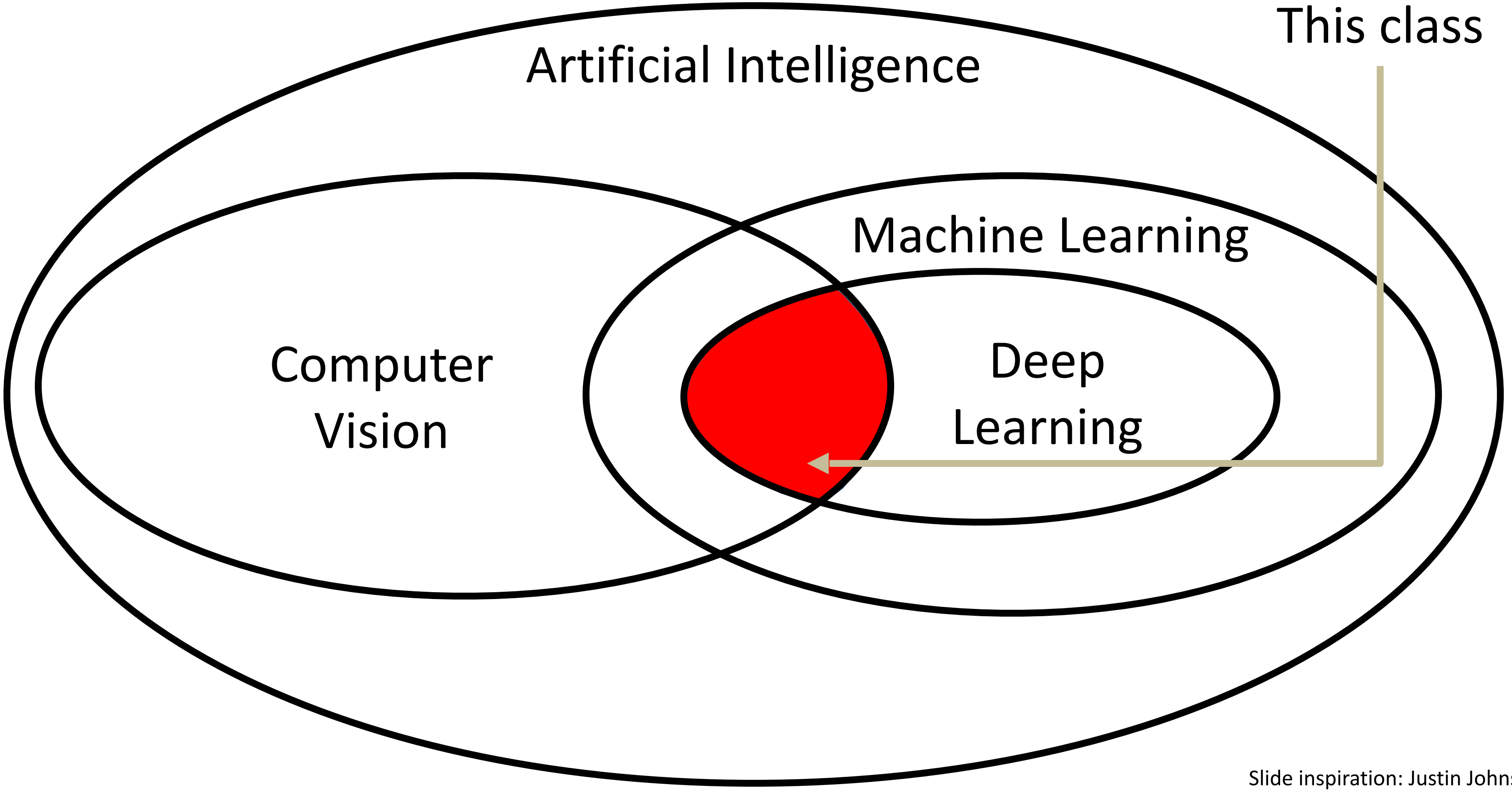
Slide inspiration: Justin Johnson



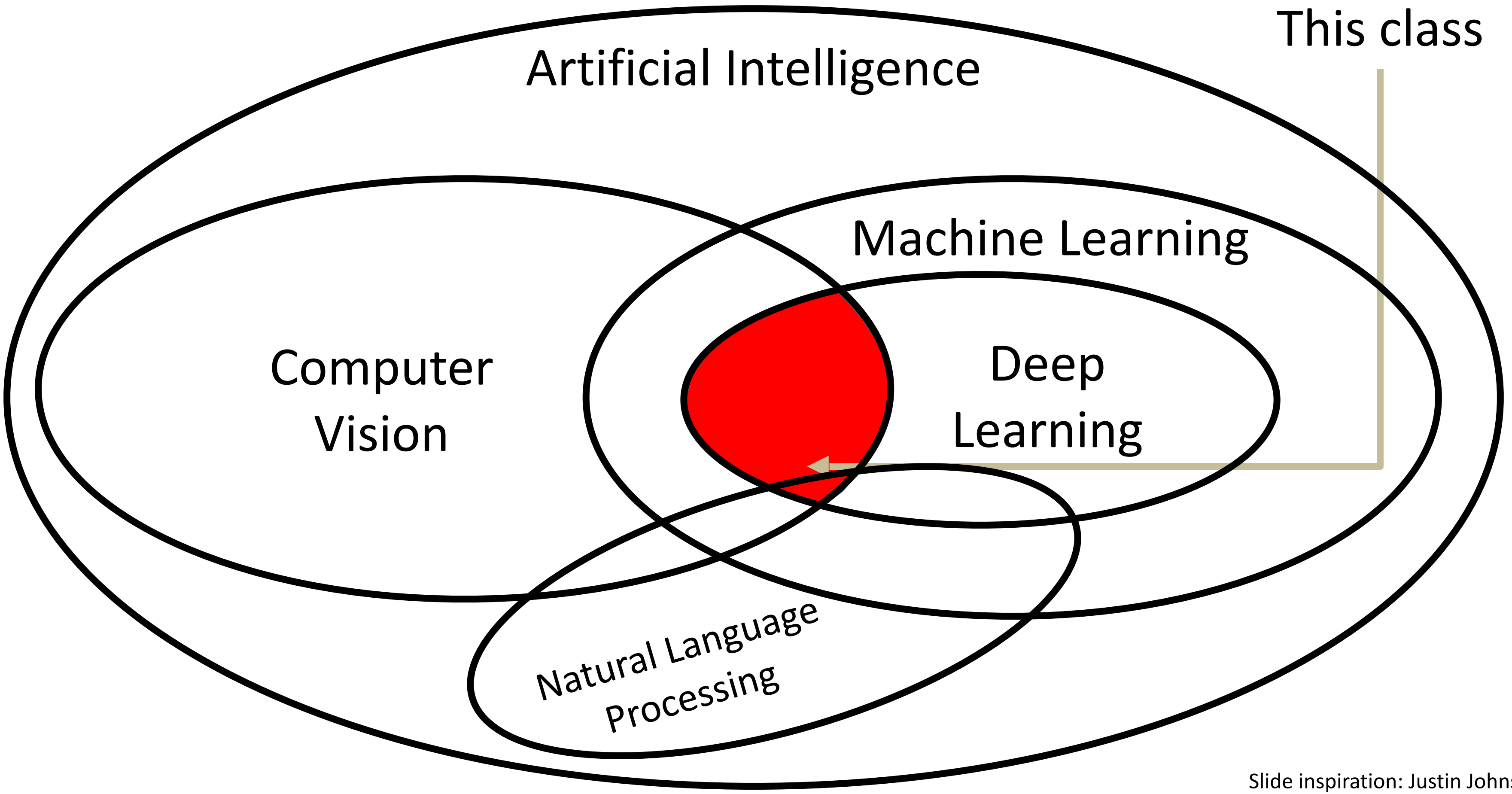
Slide inspiration: Justin Johnson



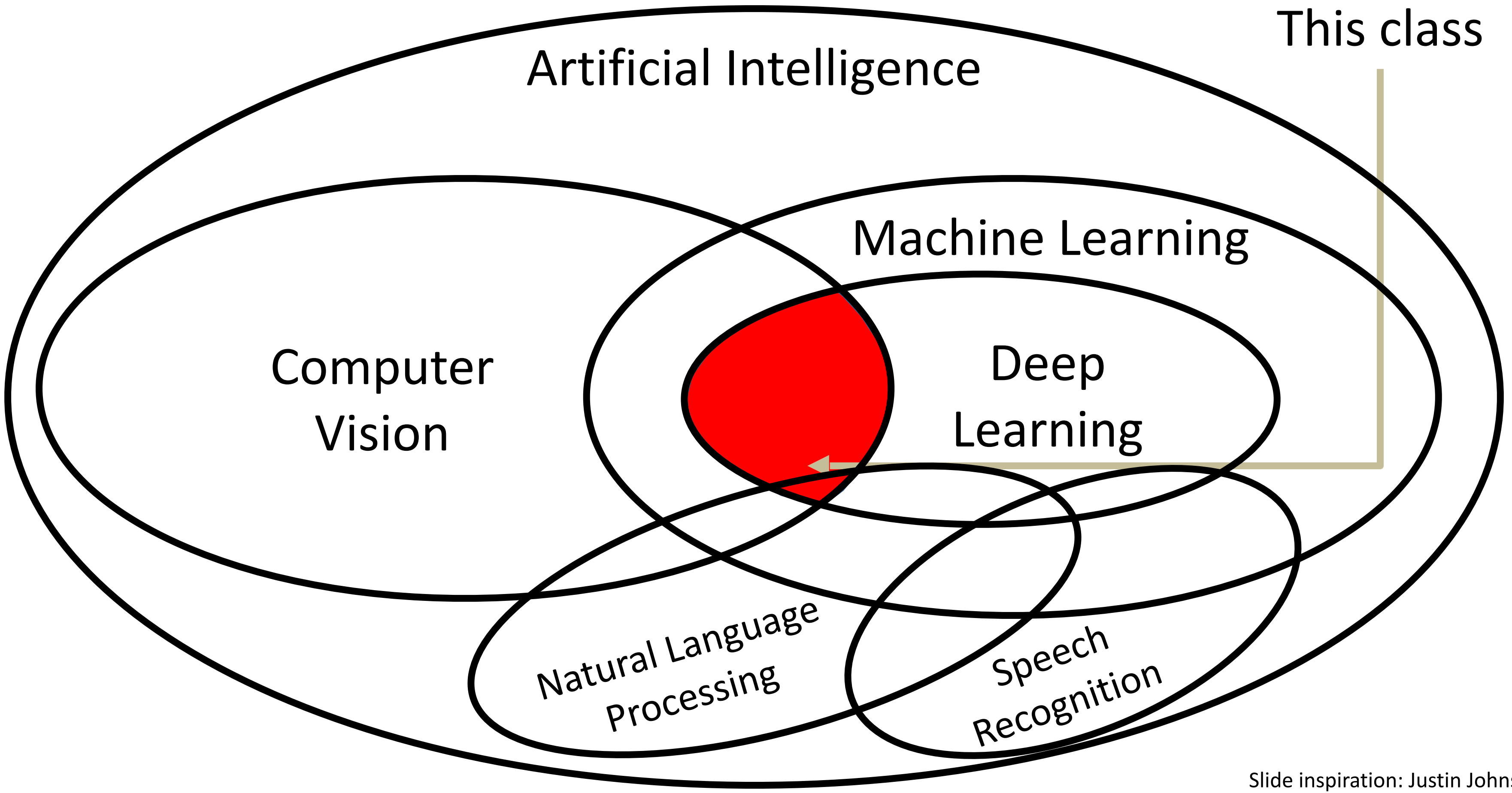
Slide inspiration: Justin Johnson



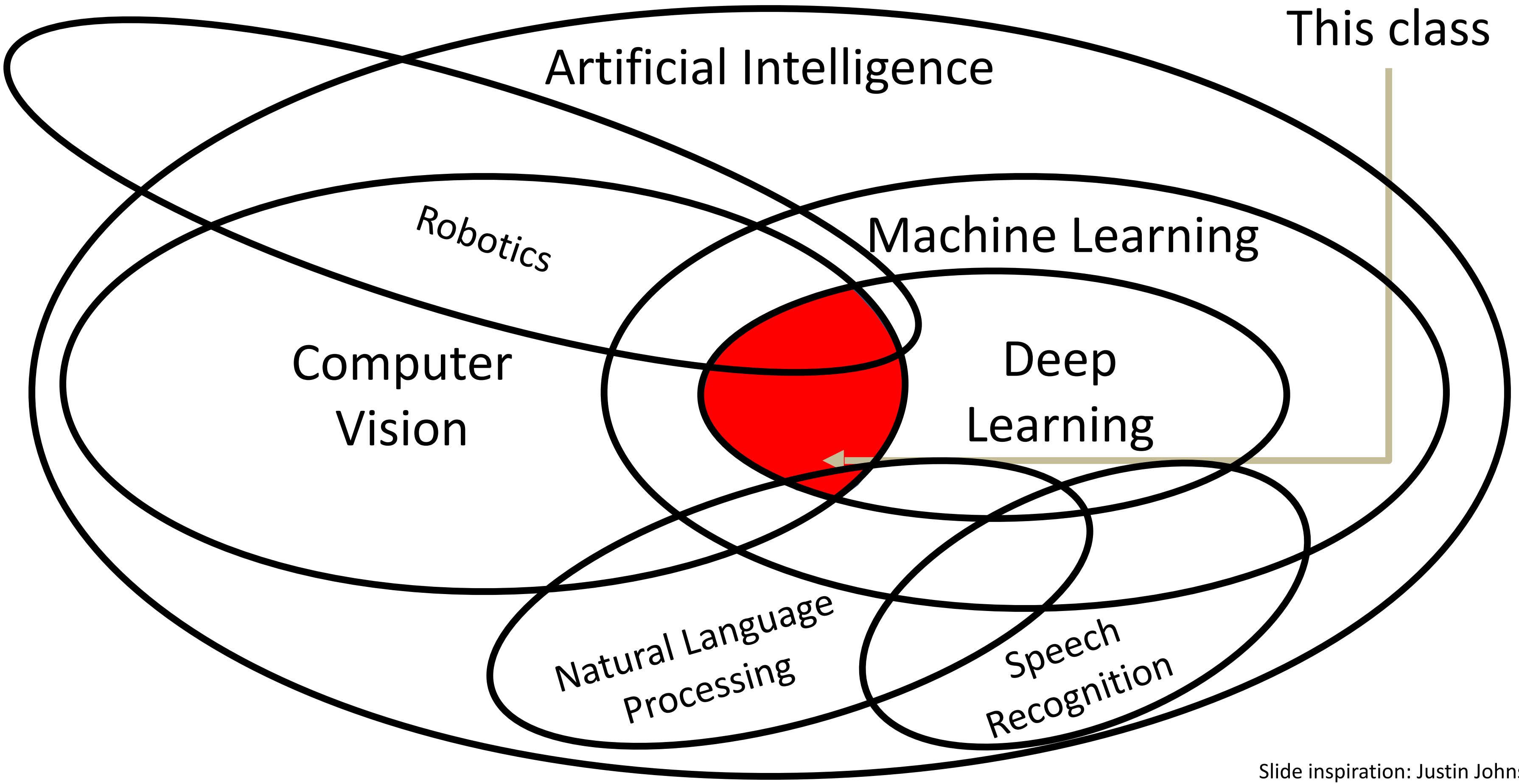
Slide inspiration: Justin Johnson



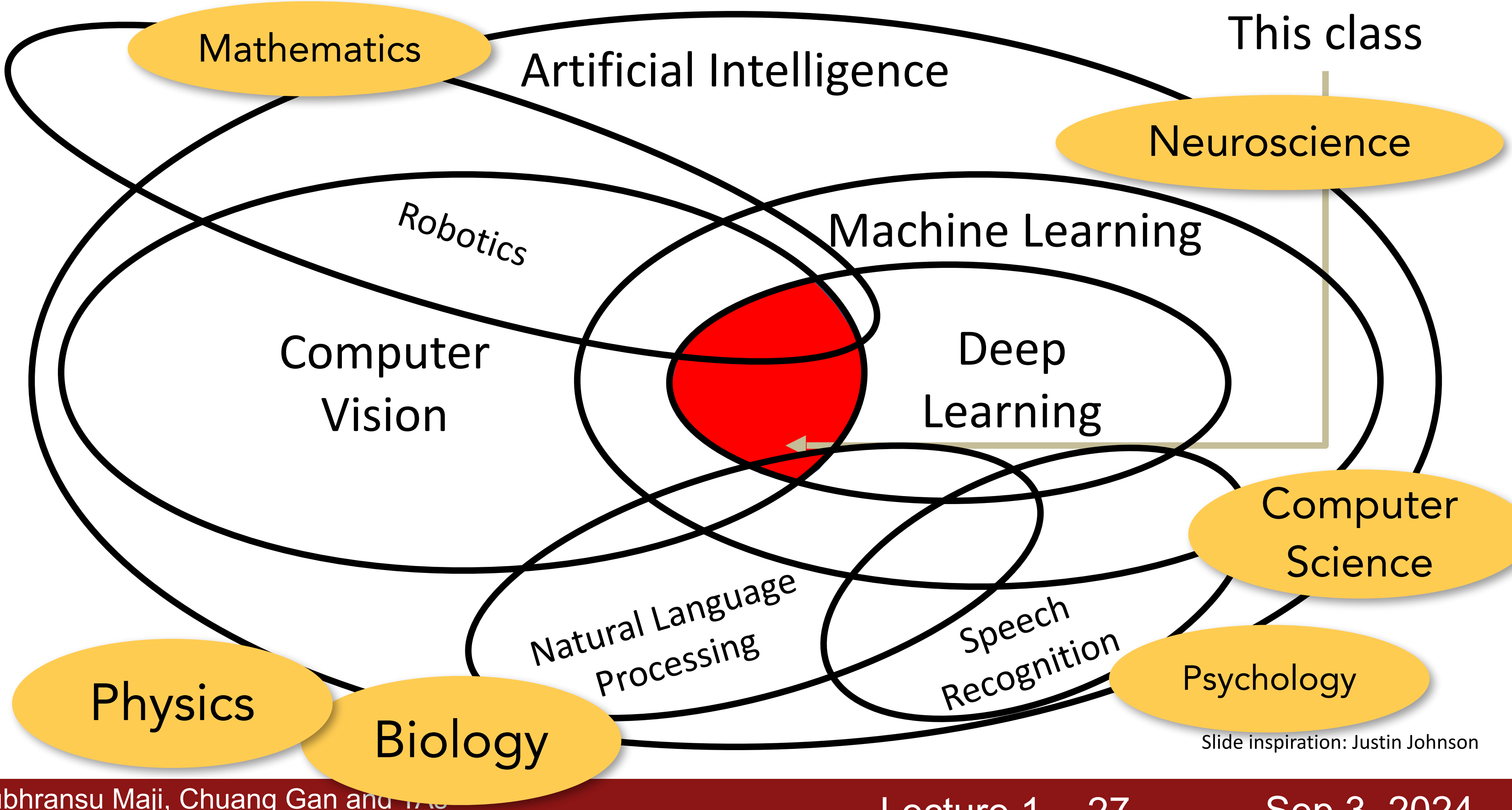
Slide inspiration: Justin Johnson



Slide inspiration: Justin Johnson



Slide inspiration: Justin Johnson

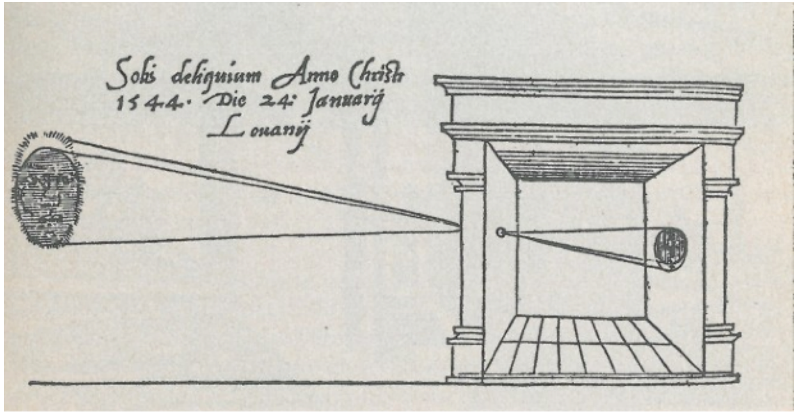


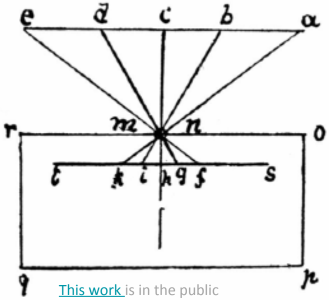
Slide inspiration: Justin Johnson

Evolution's Big Bang:
Cambrian Explosion, 530-540million years, B.C.

Camera Obscura

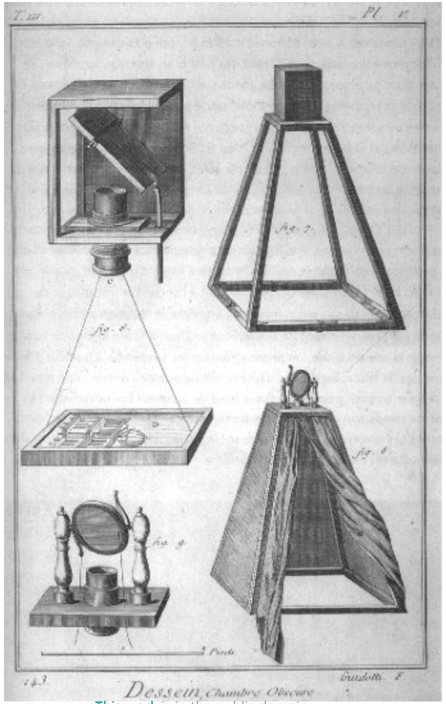
Gemma Frisius, 1545





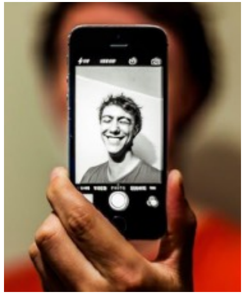
Leonardo da Vinci, 16th Century AD

Encyclopedia, 18th Century



Computer Vision is everywhere!



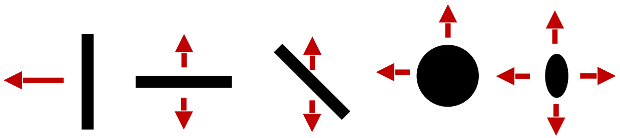




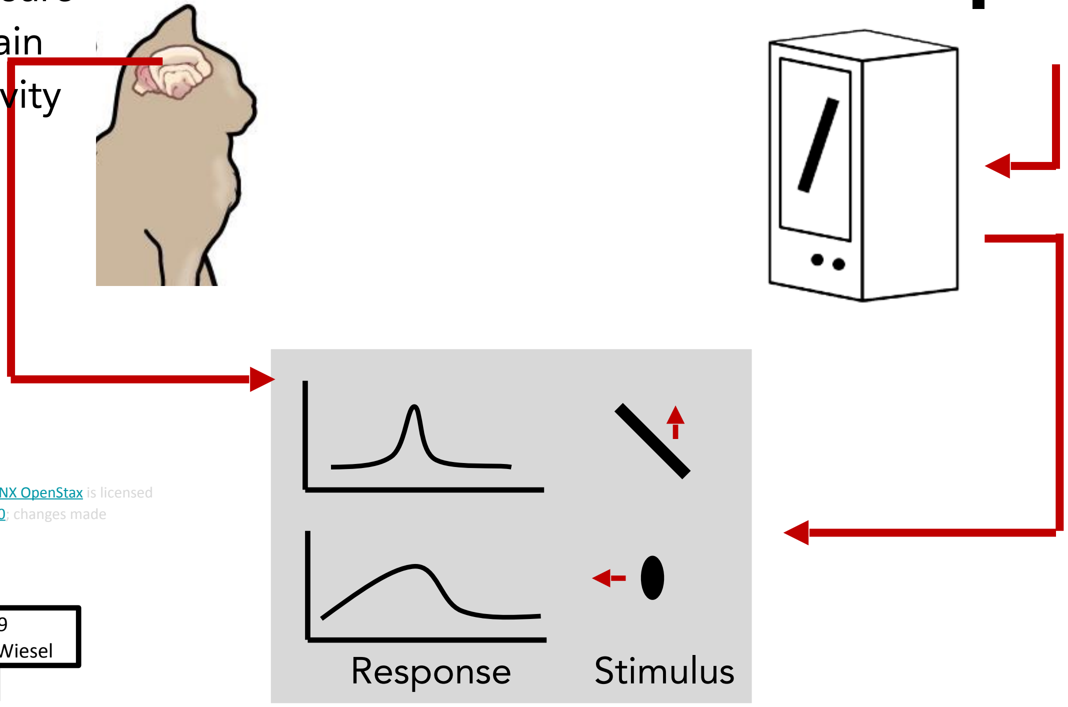


Where did we come from?

Hubel and Wiesel, 1959



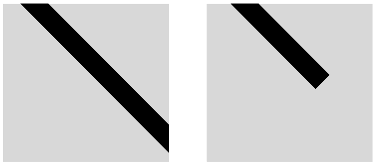
Measure brain activity



1959
Hubel & Wiesel

Simple cells:
Response to specific rotation and orientation

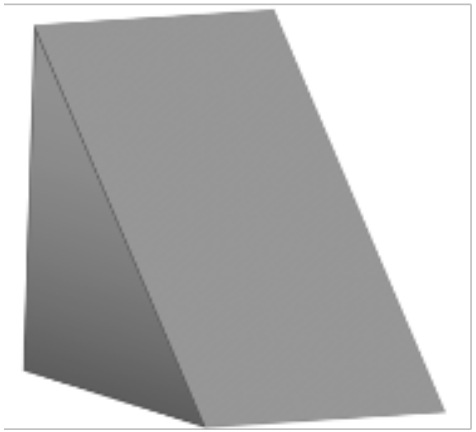
Complex cells:
Response to light orientation and movement, some translation invariance



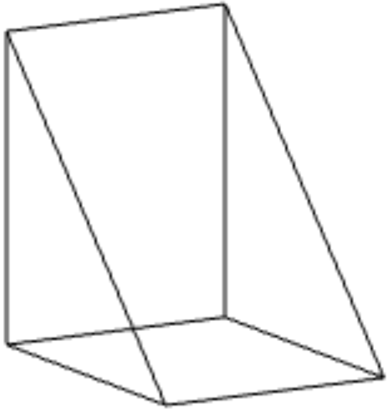
No response

Slide inspiration: Justin Johnson

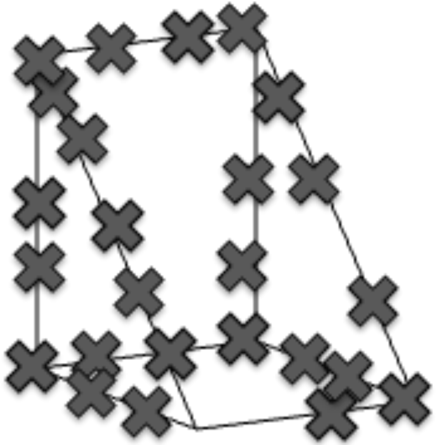
Larry Roberts, 1963



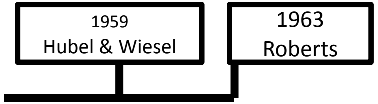
(a) Original picture



(b) Differentiated picture



(c) Feature points selected



Lawrence Gilman Roberts, “Machine Perception of Three-Dimensional Solids”, 1963

Slide inspiration: Justin Johnson

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

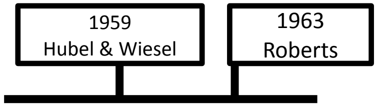
Artificial Intelligence Group July 7, 1966

Vision Memo. No. 100.

THE SUMMER VISION PROJECT

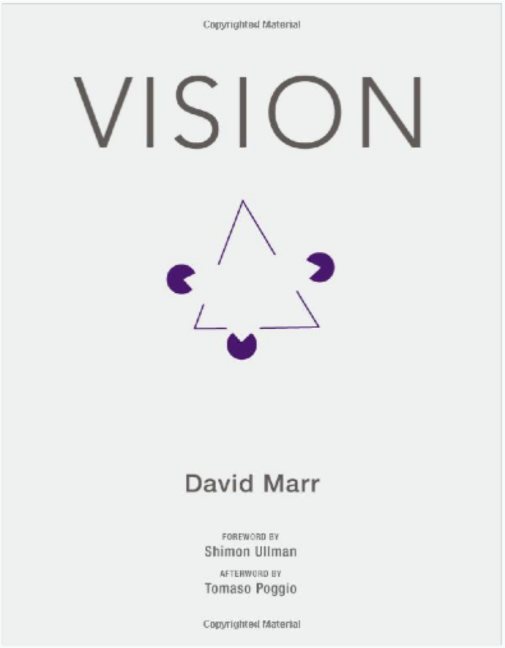
Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".



https://dspace.mit.edu/handle/1721.1/6125

Slide inspiration: Justin Johnson

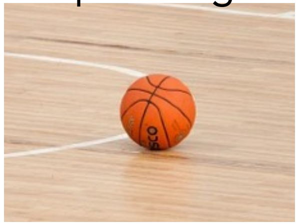


1959
Hubel & Wiesel

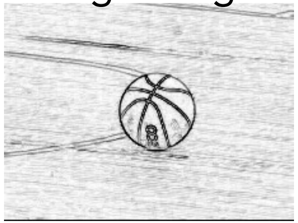
1963
Roberts

1970s
David Marr

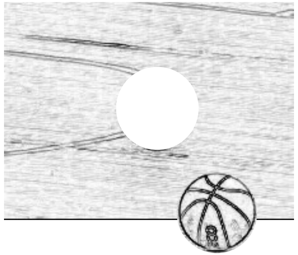
Input image



Edge image



2 ½-D sketch



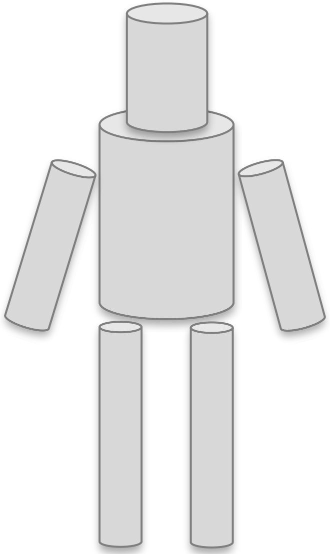
3-D model

Input Image	→	Primal Sketch	→	2 ½-D Sketch	→	3-D Model Representation
Perceived intensities		Zero crossings, blobs, edges, bars, ends, virtual lines, groups, curves boundaries		Local surface orientation and discontinuities in depth and in surface orientation		3-D models hierarchically organized in terms of surface and volumetric primitives

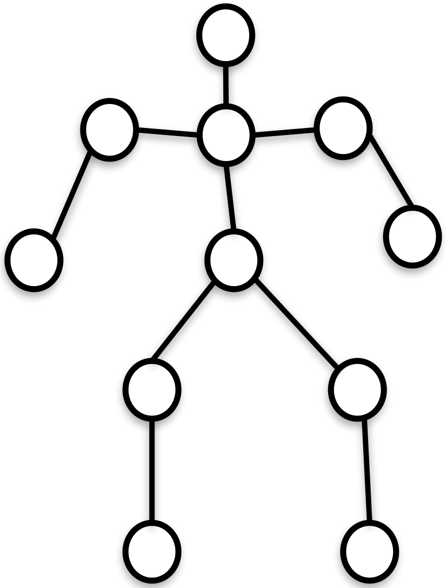
Stages of Visual Representation, David Marr, 1970s

Slide inspiration: Justin Johnson

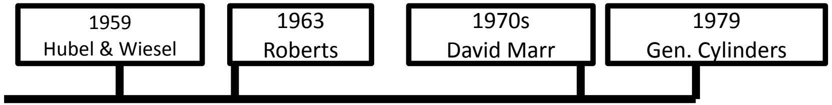
Recognition via Parts (1970s)



Generalized Cylinders,
Brooks and Binford,
1979

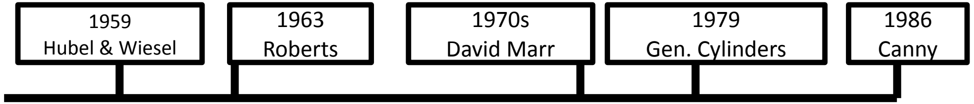


Pictorial Structures,
Fischler and Elshlager, 1973



Slide inspiration: Justin Johnson

Recognition via Edge Detection (1980s)

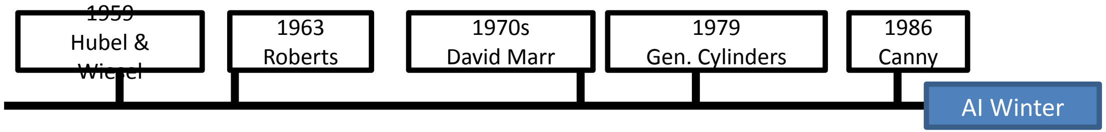


John Canny, 1986
David Lowe, 1987

Slide inspiration: Justin Johnson

Arriving at an "AI winter"

- Enthusiasm (and funding!) for AI research dwindled
- "Expert Systems" failed to deliver on their promises
- But subfields of AI continues to grow
  - Computer vision, NLP, robotics, compbio, etc.

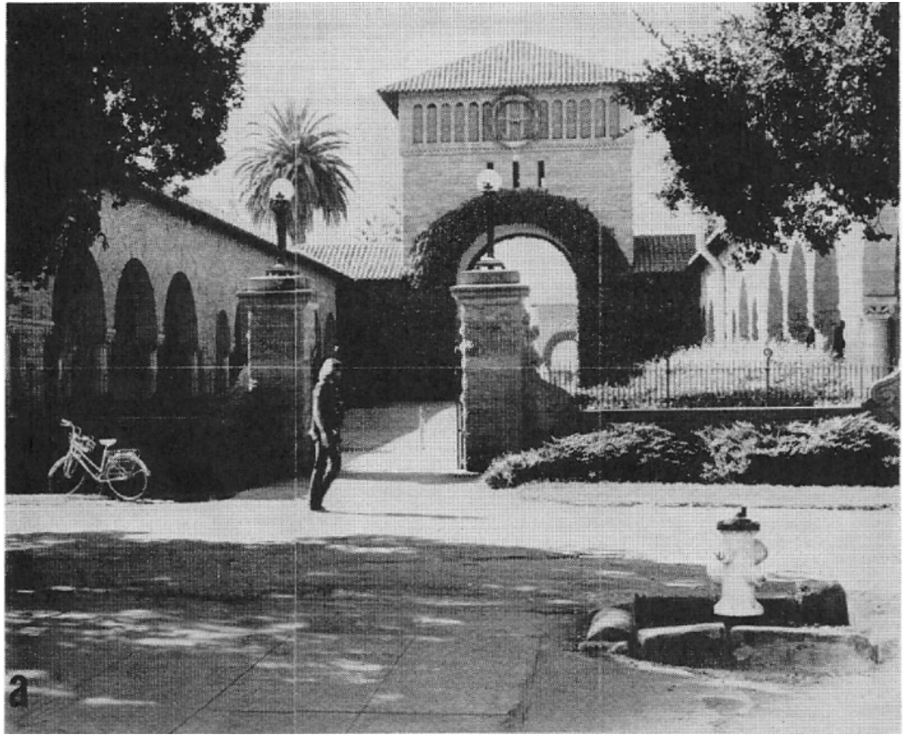


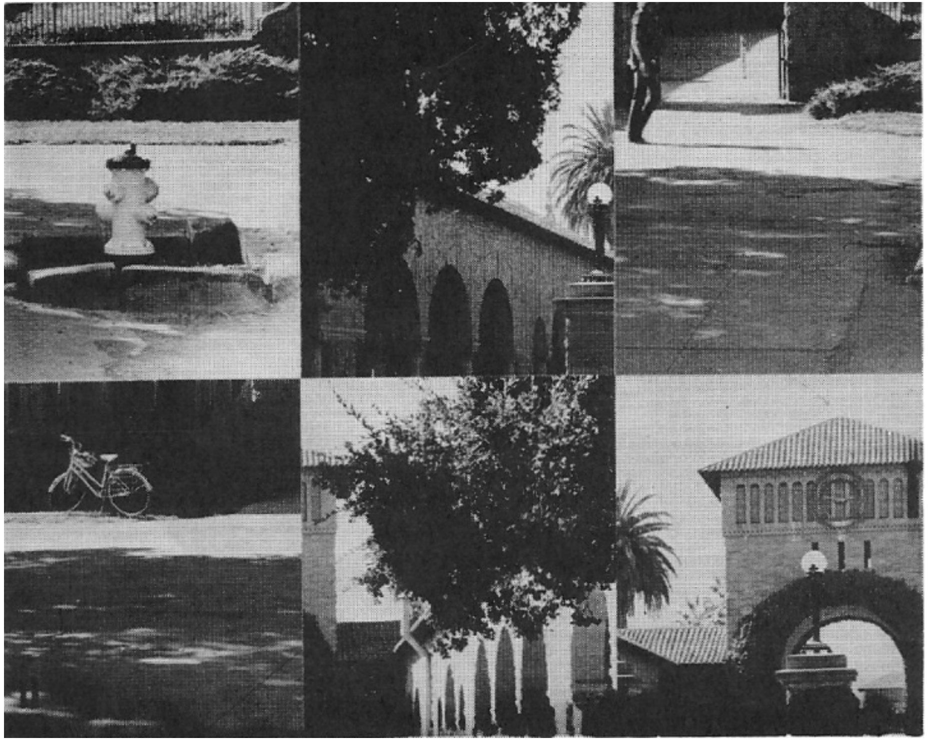
Slide inspiration: Justin Johnson

In the meantime…seminal work in cognitive and neuroscience

Perceiving Real-World Scenes

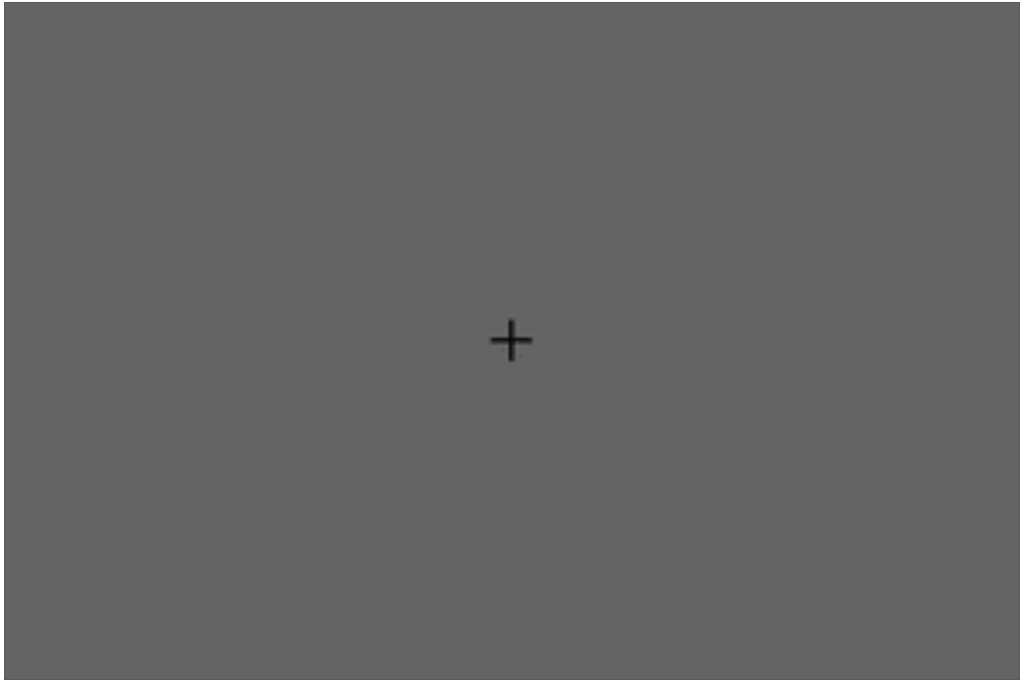
Irving Biederman





I. Biederman, *Science*, 1972

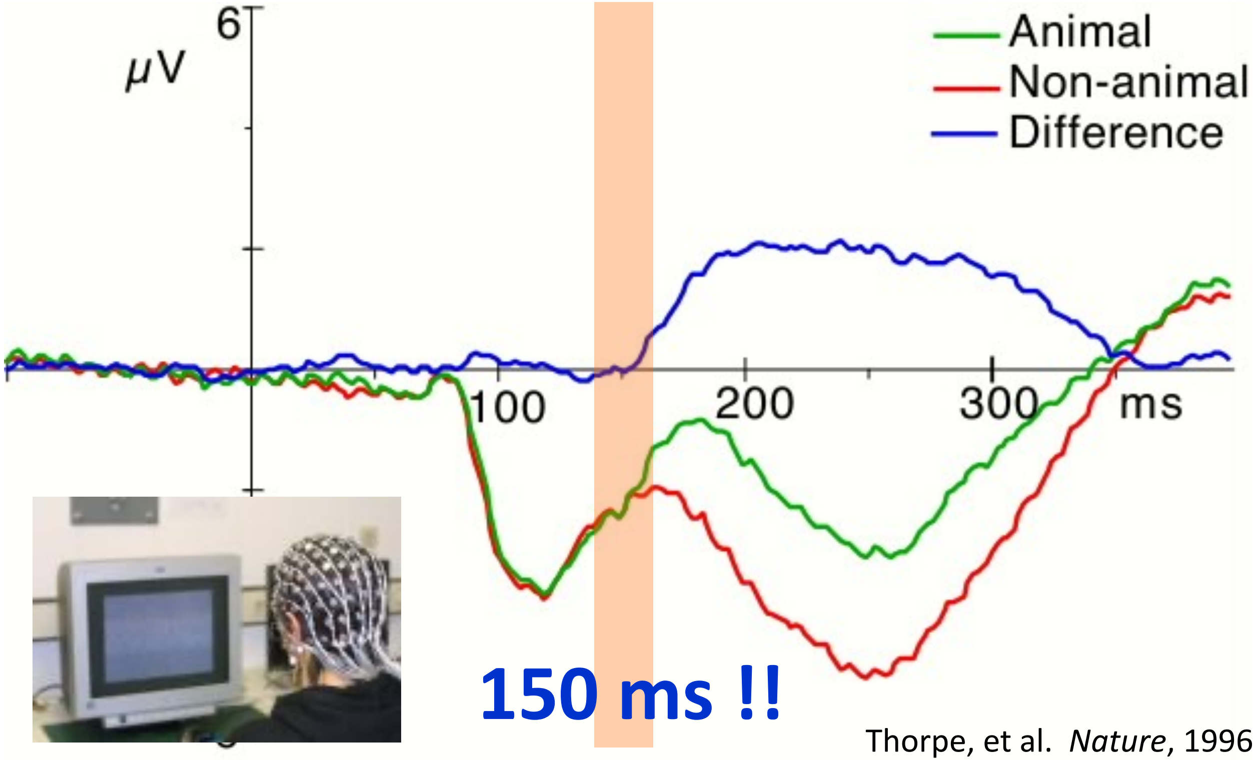
Rapid Serial Visual Perception (RSVP)



Potter, etc. 1970s

Speed of processing in the human visual system

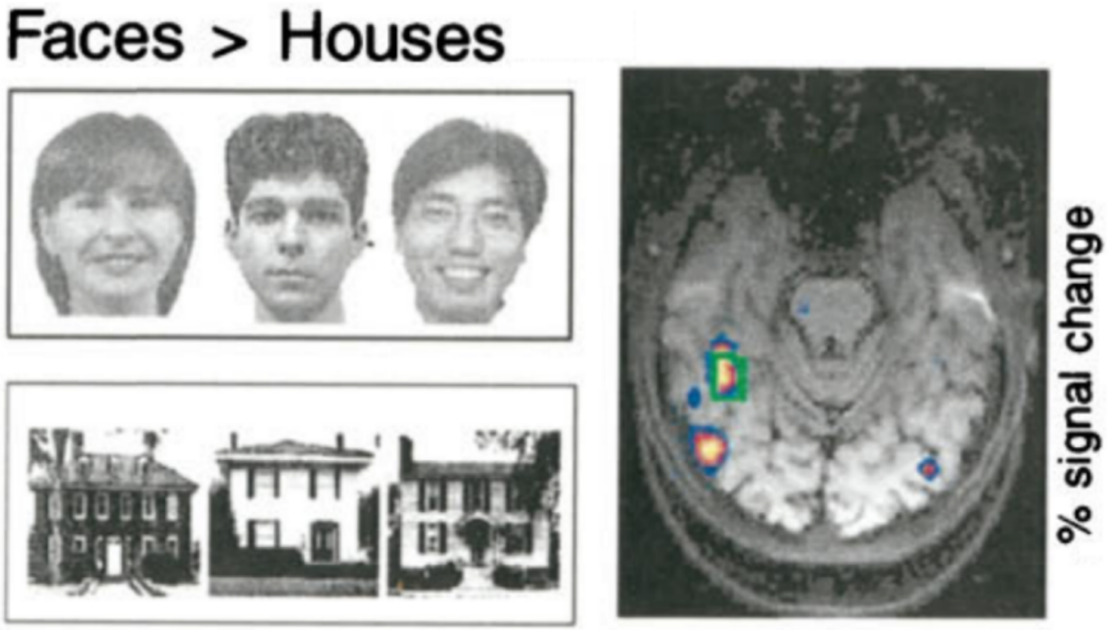
Simon Thorpe, Denis Fize & Catherine Marlot



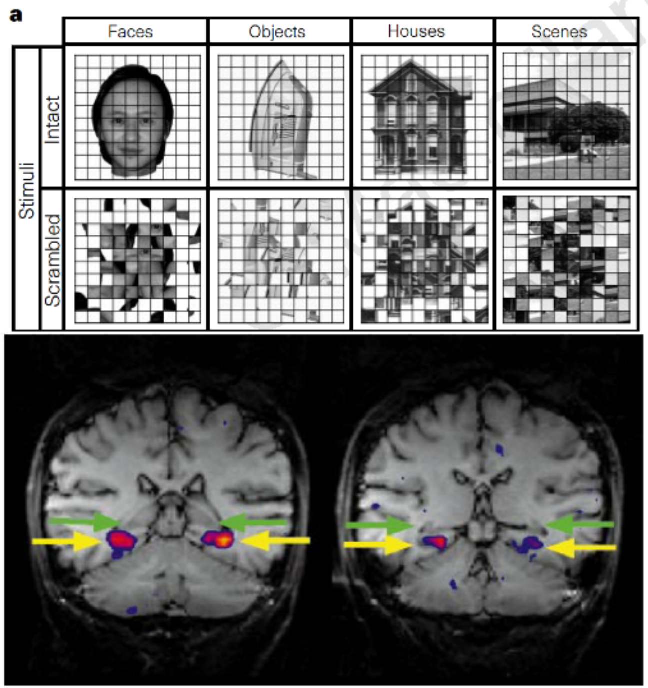
150 ms !!

Thorpe, et al. *Nature*, 1996

Neural correlates of object & scene recognition

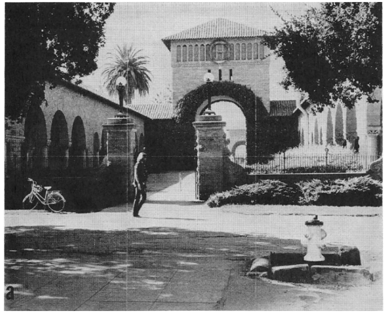


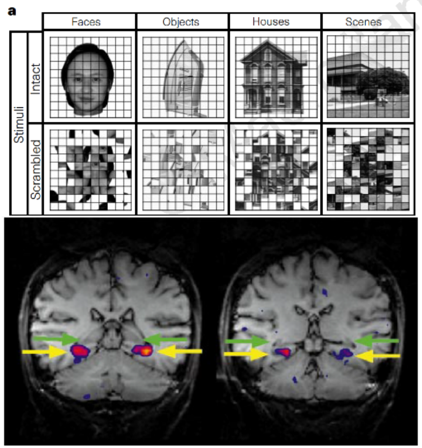
Kanwisher et al. J. Neuro. 1997



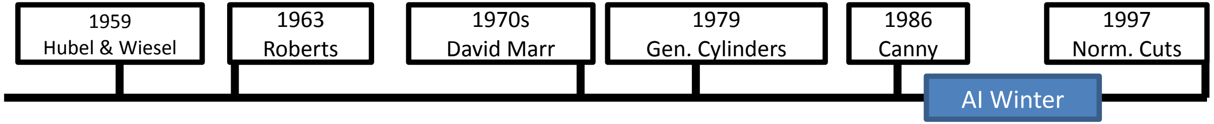
Epstein & Kanwisher, Nature, 1998

Visual recognition is a fundamental task for visual intelligence





Recognition via Grouping (1990s)



Normalized Cuts, Shi and Malik, 1997

Slide inspiration: Justin Johnson

Recognition via Matching (2000s)

[Image](#) is public domain

[Image](#) is public domain

1959 Hubel & Wiesel | 1963 Roberts | 1970s David Marr | 1979 Gen. Cylinders | 1986 Canny | AI Winter | 1997 Norm. Cuts | 1999 SIFT

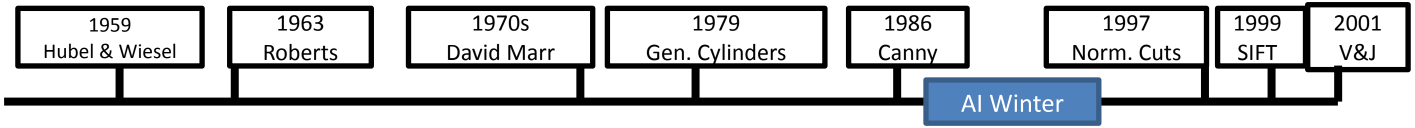
SIFT, David Lowe, 1999

Slide inspiration: Justin Johnson

Face Detection

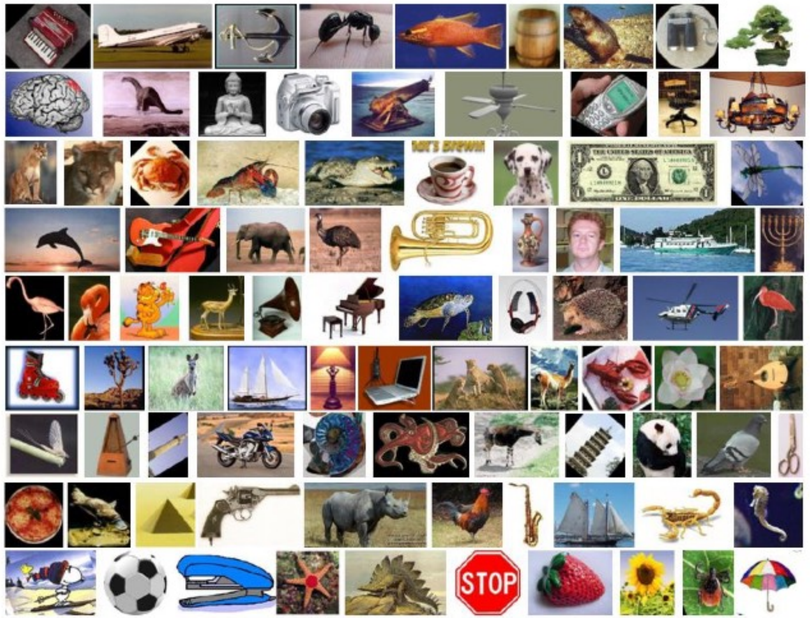
Viola and Jones, 2001

One of the first successful applications of machine learning to vision



Slide inspiration: Justin Johnson

Caltech 101 images



PASCAL Visual Object Challenge

Image is CC0 1.0 public domain

Image is CC0 1.0 public domain

1959 Hubel & Wiesel | 1963 Roberts | 1970s David Marr | 1979 Gen. Cylinders | 1986 Canny | AI Winter | 1997 Norm. Cuts | 1999 SIFT | 2001 V&J | 2004, 2007 Caltech101; PASCAL

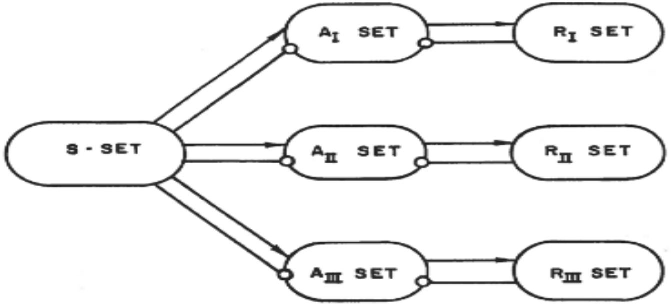
Slide inspiration: Justin Johnson

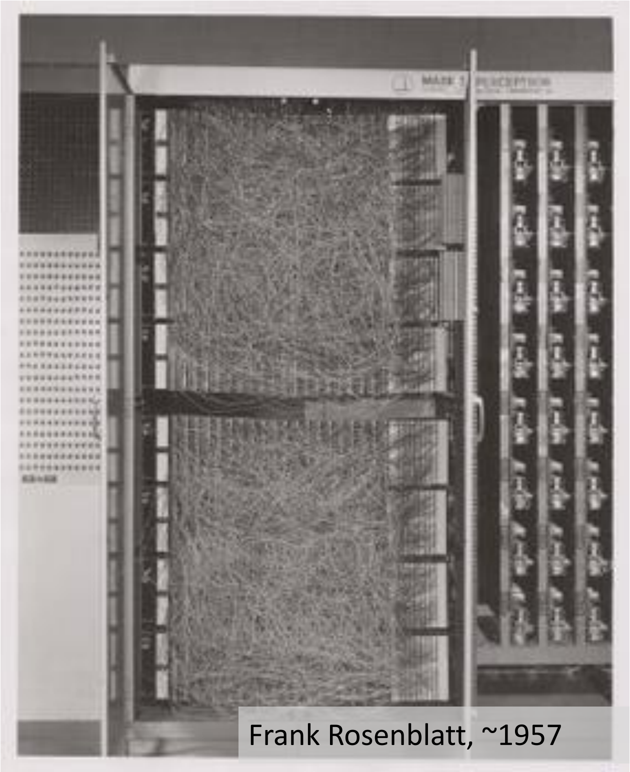
Perceptron

Learning representations by back-propagating errors

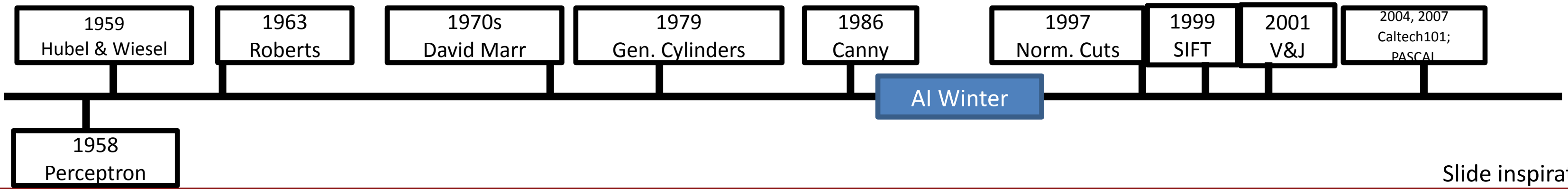
David E. Rumelhart*, Geoffrey E. Hinton† & Ronald J. Williams*

* Institute for Cognitive Science, C-015, University of California, San Diego, La Jolla, California 92093, USA
† Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Philadelphia 15213, USA





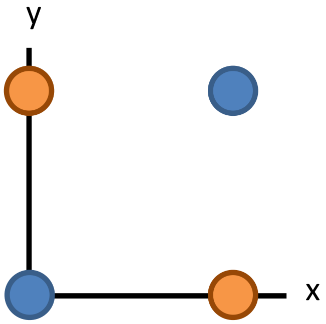
Frank Rosenblatt, ~1957

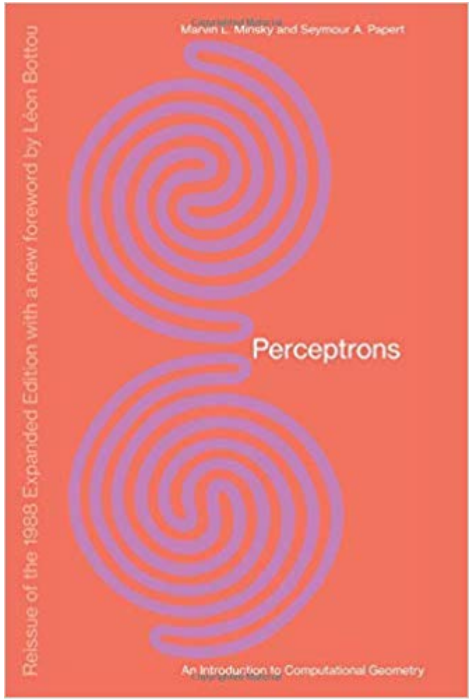


Slide inspiration: Justin Johnson

Minsky and Papert, 1969

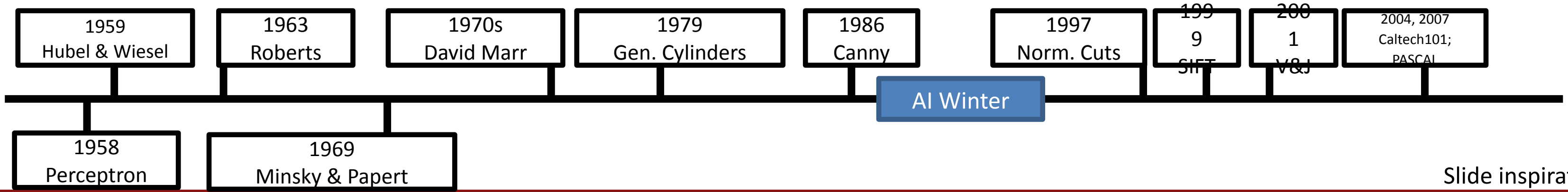
X	Y	F(x,y)
0	0	0
0	1	1
1	0	1
1	1	0





Showed that Perceptrons could not learn the XOR function

Caused a lot of disillusionment in the field



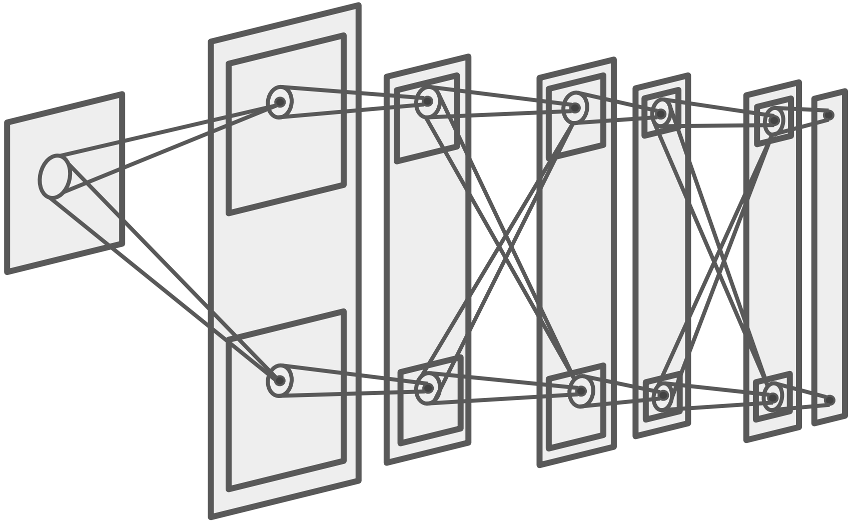
Slide inspiration: Justin Johnson

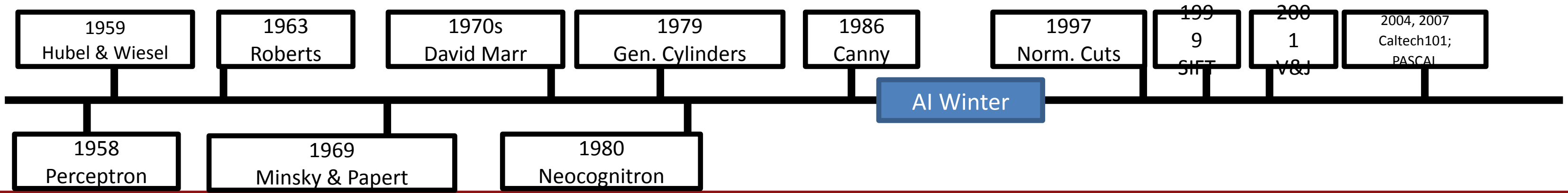
Neocognitron: Fukushima, 1980

Computational model the visual system, directly inspired by Hubel and Wiesel's hierarchy of complex and simple cells

Interleaved simple cells (convolution) and complex cells (pooling)

No practical training algorithm





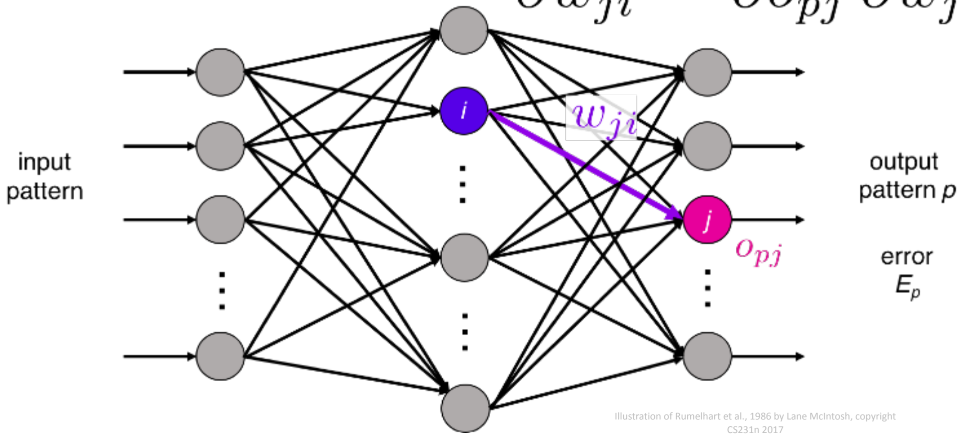
Backprop: Rumelhart, Hinton, and Williams, 1986

Introduced backpropagation for computing gradients in neural networks

Successfully trained perceptrons with multiple layers

$$\frac{\partial E_p}{\partial w_{ji}} = \frac{\partial E_p}{\partial o_{pj}} \frac{\partial o_{pj}}{\partial w_{ji}}$$

recognizable math



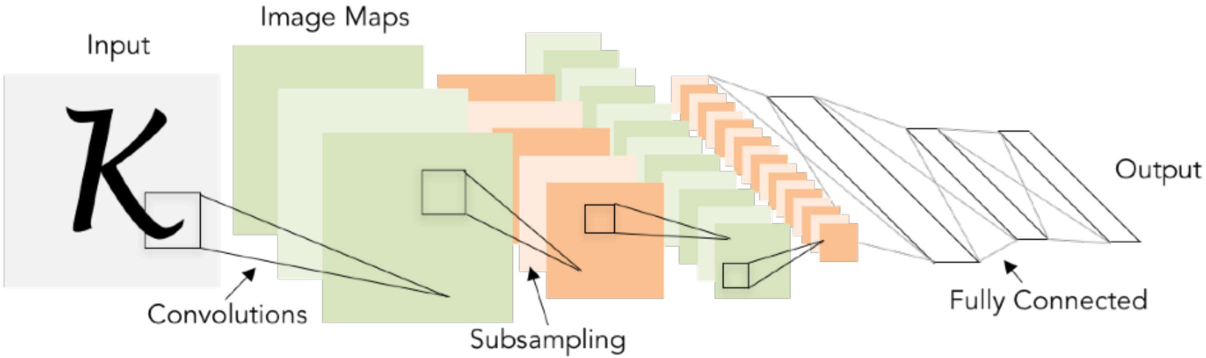
1959 Hubel & Wiesel | 1963 Roberts | 1970s David Marr | 1979 Gen. Cylinders | 1986 Canny | 1997 Norm. Cuts | 1999 SIFT | 2001 V&J | 2004, 2007 Caltech101; PASCAL

AI Winter

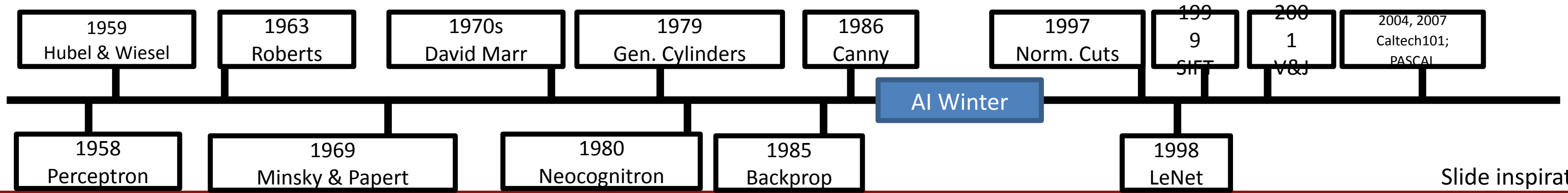
1958 Perceptron | 1969 Minsky & Papert | 1980 Neocognitron | 1985 Backprop

Slide inspiration: Justin Johnson

Convolutional Networks: LeCun et al, 1998



Applied backprop algorithm to a Neocognitron-like architecture
Learned to recognize handwritten digits
Was deployed in a commercial system by NEC, processed handwritten checks
Very similar to our modern convolutional networks!



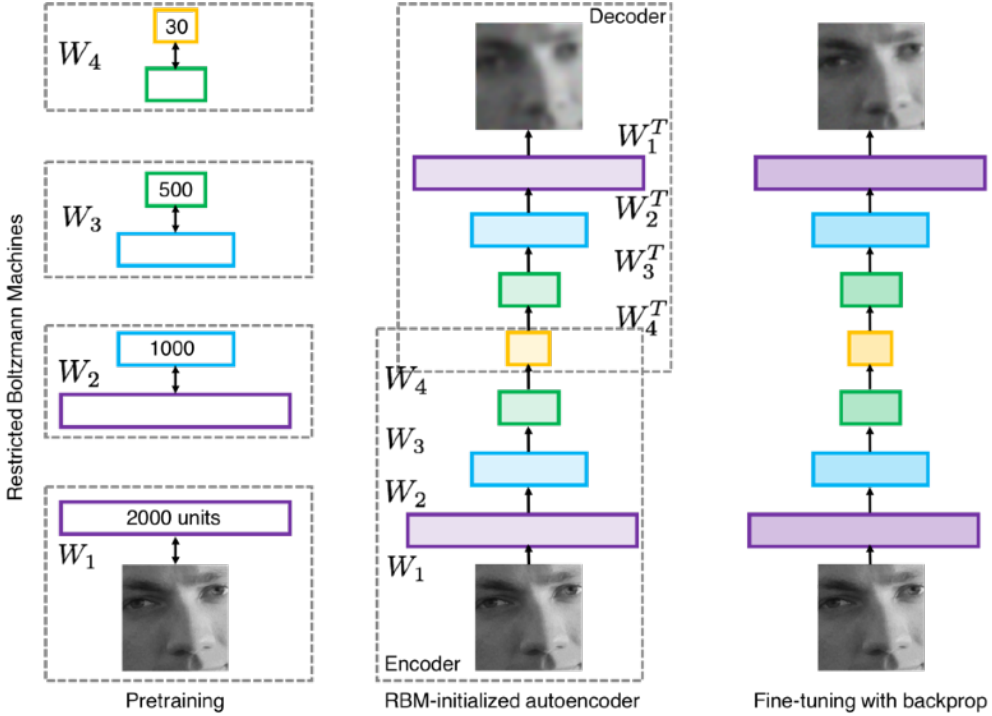
Slide inspiration: Justin Johnson

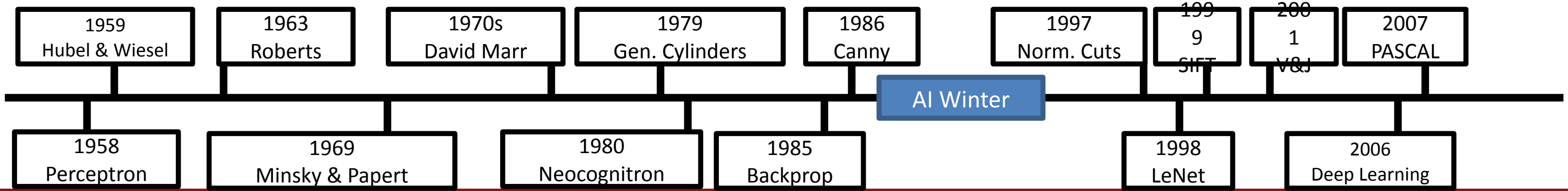
2000s: "Deep Learning"

People tried to train neural networks that were deeper and deeper

Not a mainstream research topic at this time

Hinton and Salakhutdinov, 2006
Bengio et al, 2007
Lee et al, 2009
Glorot and Bengio, 2010





Slide inspiration: Justin Johnson

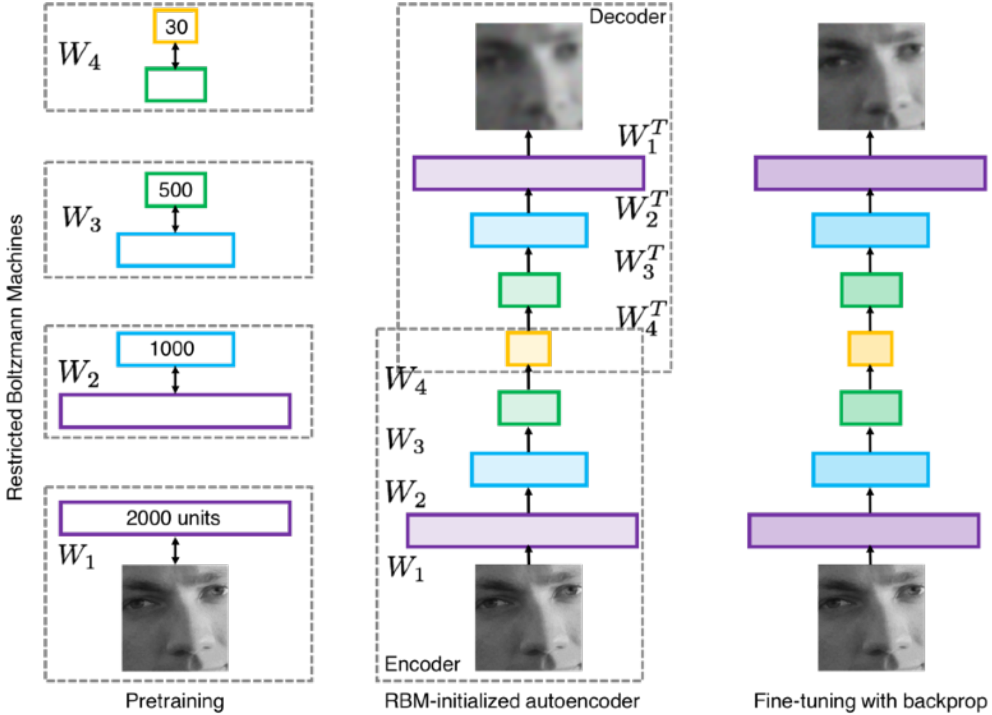
2000s: "Deep Learning"

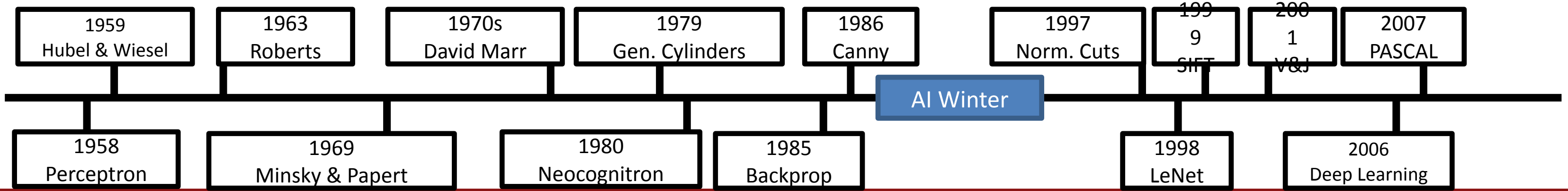
People tried to train neural networks that were deeper and deeper

Not a mainstream research topic at this time

No good dataset to work on

Hinton and Salakhutdinov, 2006
Bengio et al, 2007
Lee et al, 2009
Glorot and Bengio, 2010





Slide inspiration: Justin Johnson

IMAGENET Large Scale Visual Recognition Challenge

The Image Classification Challenge:

1,000 object classes

1,431,167 images

Output:
Scale
T-shirt
Steel drum
Drumstick
Mud turtle

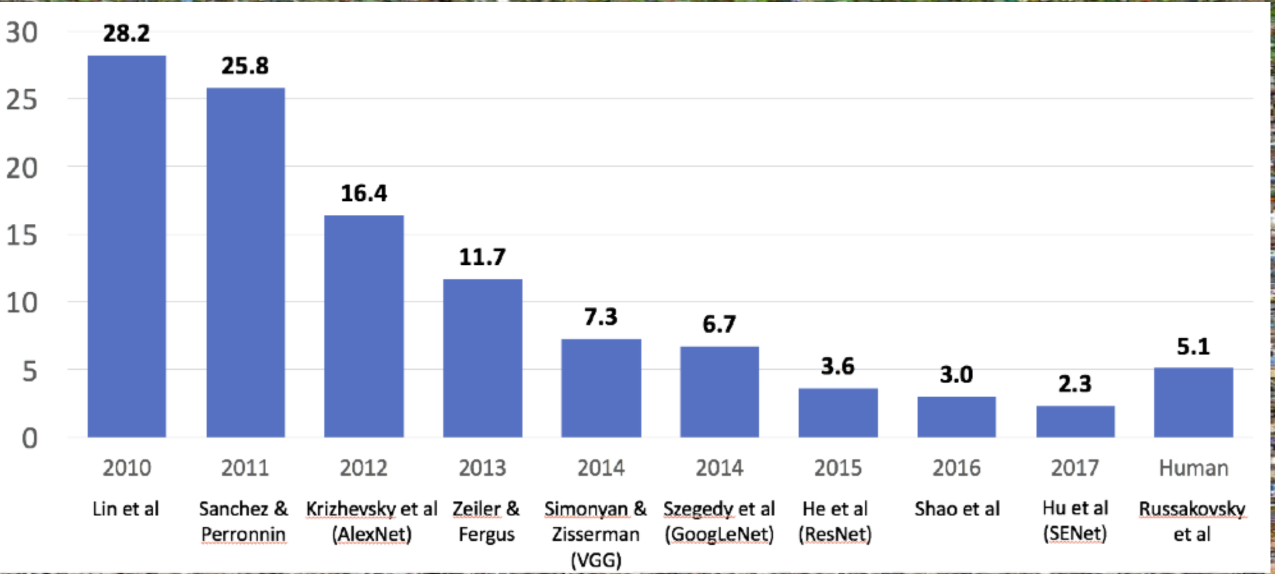
Deng et al, 2009
Russakovsky et al. IJCV 2015

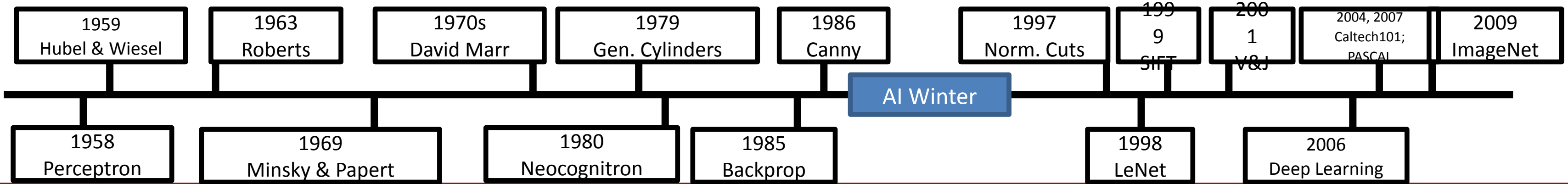
1959 Hubel & Wiesel | 1963 Roberts | 1970s David Marr | 1979 Gen. Cylinders | 1986 Canny | 1997 Norm. Cuts | 1999 SIFT | 2001 V&J | 2004, 2007 Caltech101; PASCAL | 2009 ImageNet

AI Winter

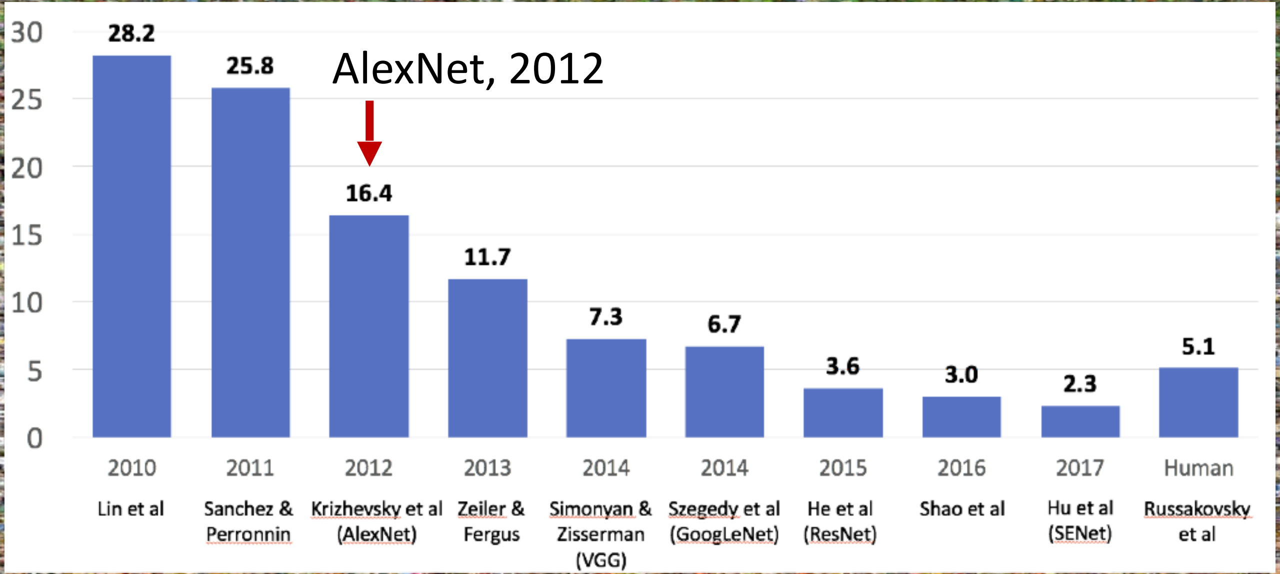
1958 Perceptron | 1969 Minsky & Papert | 1980 Neocognitron | 1985 Backprop | 1998 LeNet | 2006 Deep Learning

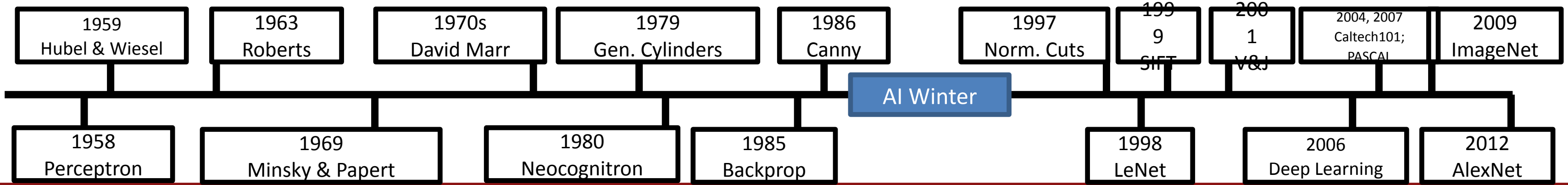
IMAGENET Large Scale Visual Recognition Challenge



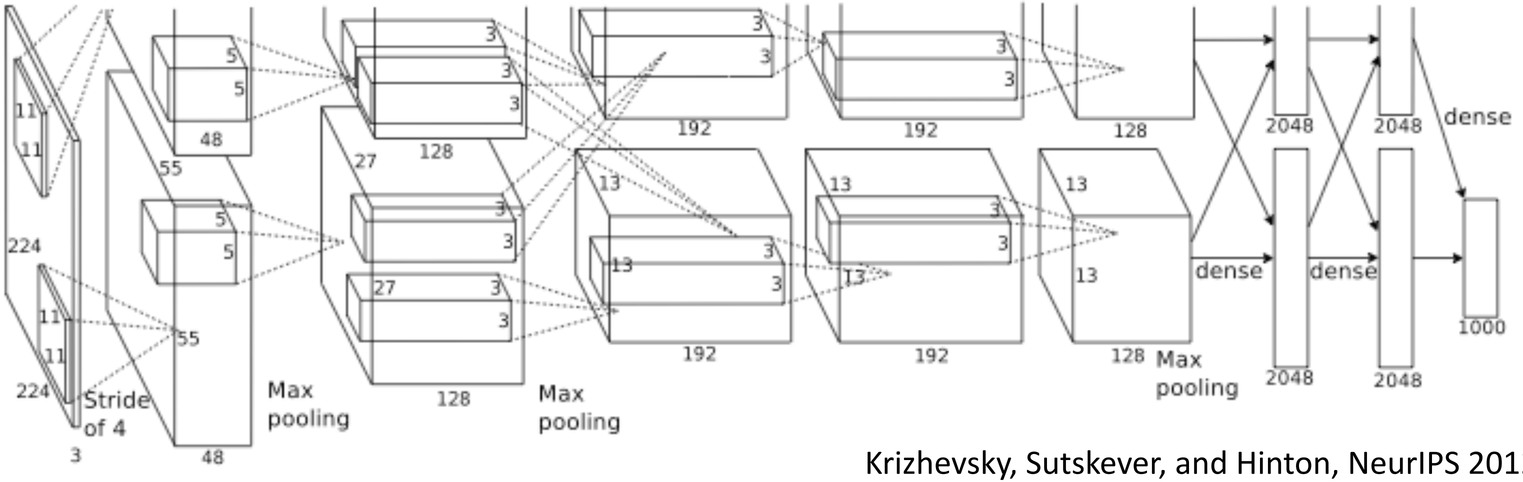


IMAGENET Large Scale Visual Recognition Challenge

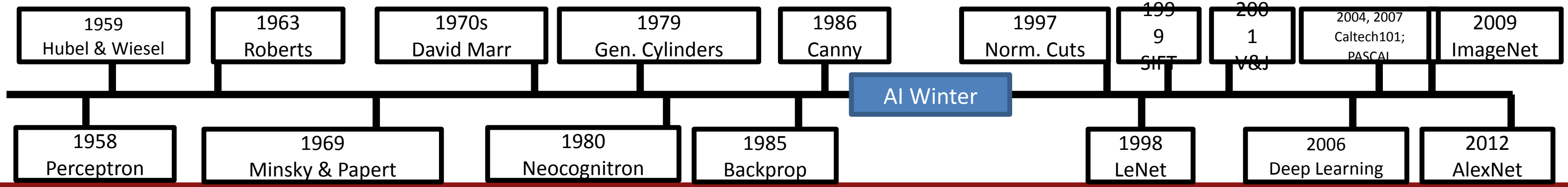




AlexNet: Deep Learning Goes Mainstream

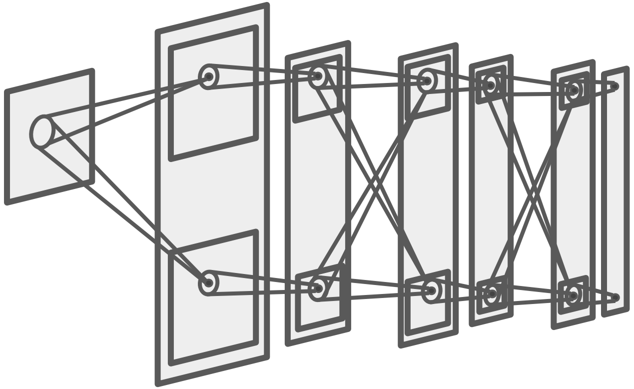


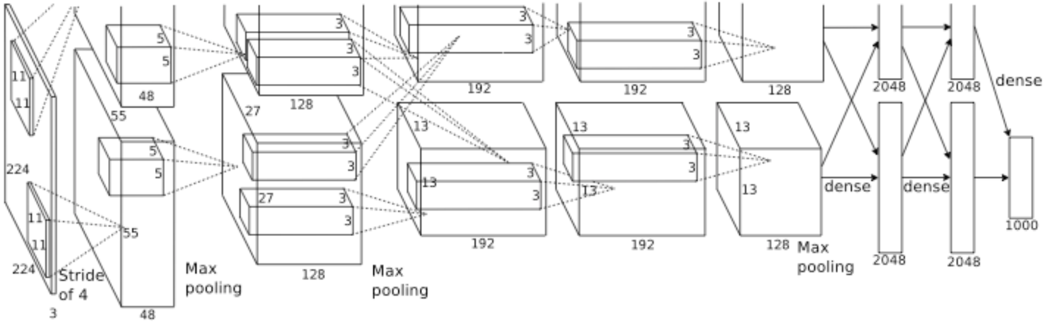
Krizhevsky, Sutskever, and Hinton, NeurIPS 2012



Slide inspiration: Justin Johnson

AlexNet vs. Neocognitron: 32 years apart





1959 Hubel & Wiesel | 1963 Roberts | 1970s David Marr | 1979 Gen. Cylinders | 1986 Canny | 1997 Norm. Cuts | 1999 SIFT | 2001 V&J | 2004, 2007 Caltech101; PASCAL | 2009 ImageNet

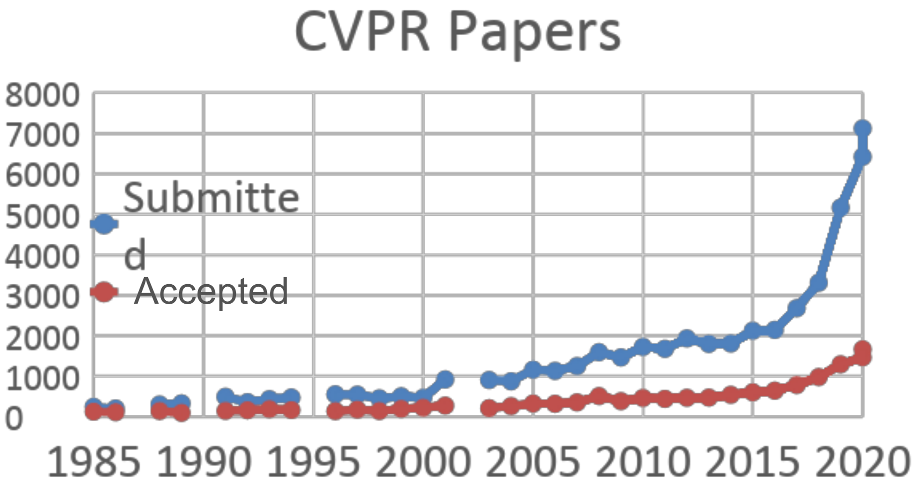
AI Winter

1958 Perceptron | 1969 Minsky & Papert | 1980 Neocognitron | 1985 Backprop | 1998 LeNet | 2006 Deep Learning | 2012 AlexNet

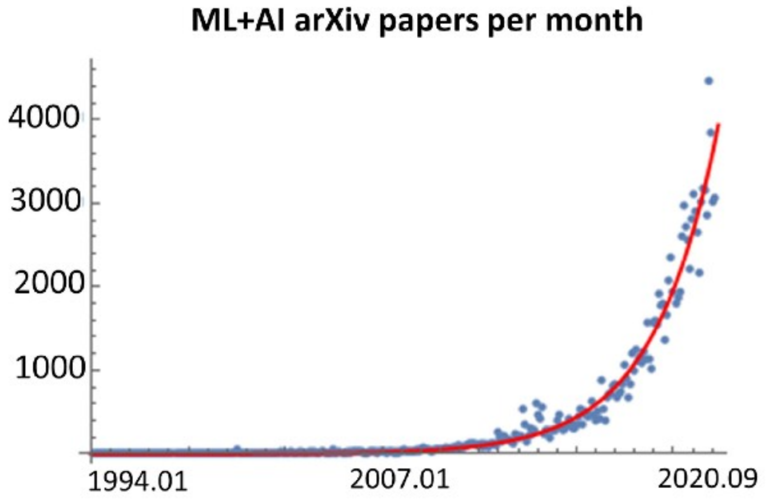
Slide inspiration: Justin Johnson

The AI winter thawed,
deep learning revolution arrived

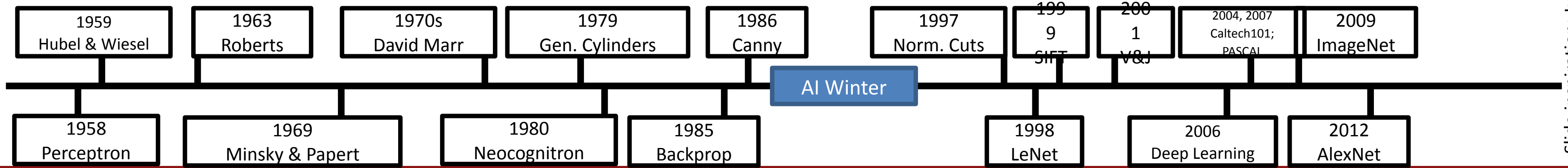
2012 to Present: Deep Learning Explosion



Publications at top Computer Vision conference



arXiv papers per month (source)

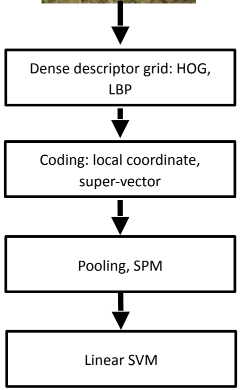


Slide inspiration: Justin Johnson

2012 to Present: Deep Learning is Everywhere

Year 2010

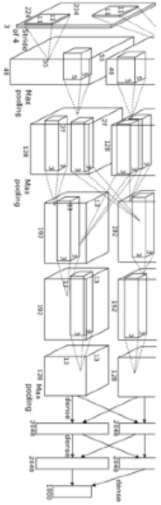
NEC-UIUC



[Lin CVPR 2011]

Year 2012

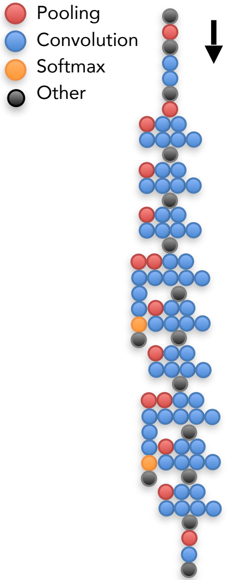
SuperVision



[Krizhevsky NIPS 2012]

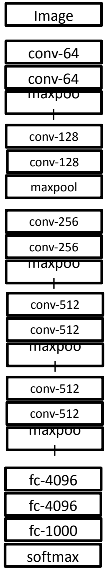
Year 2014

GoogLeNet



[Szegedy arxiv 2014]

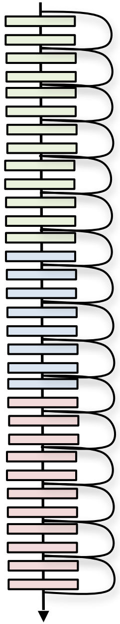
VGG



[Simonyan arxiv 2014]

Year 2015

MSRA



[He ICCV 2015]

2012 to Present: Deep Learning is Everywhere

Image Classification

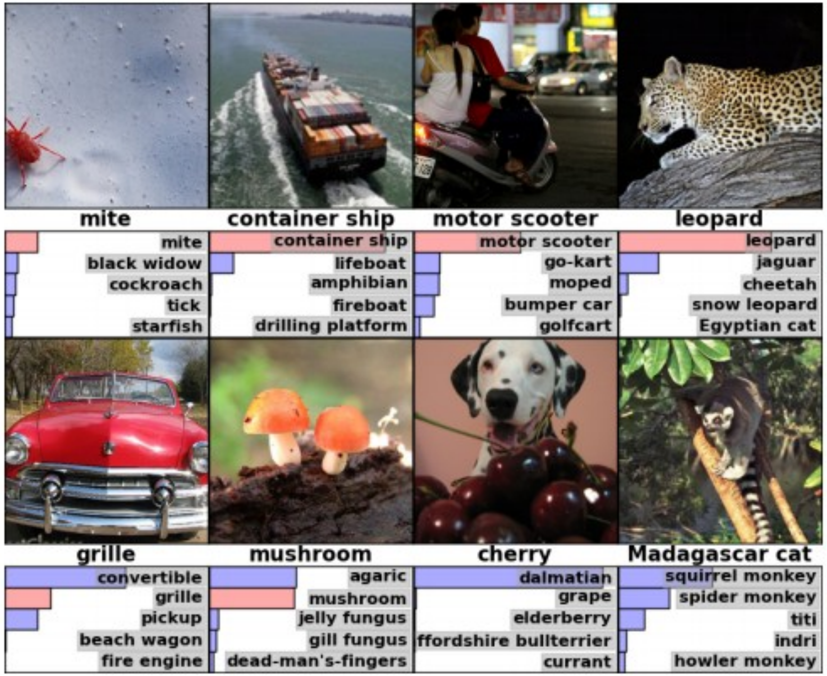
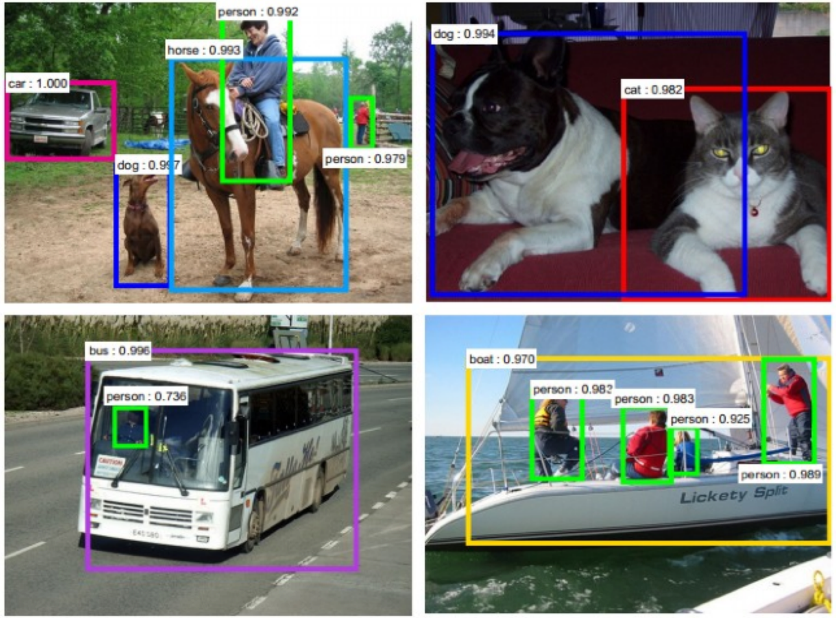


Image Retrieval

Slide inspiration: Justin Johnson

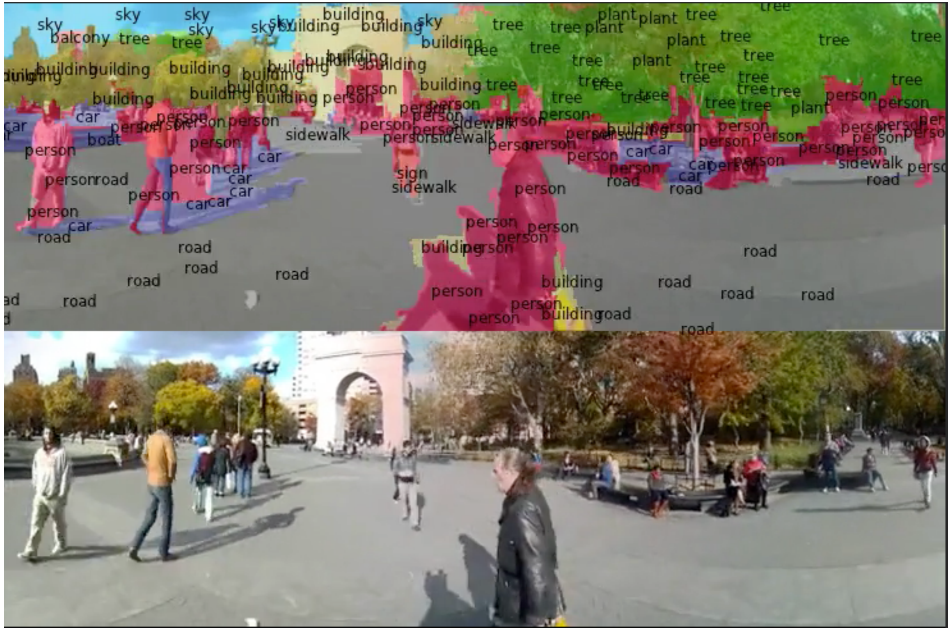
2012 to Present: Deep Learning is Everywhere

Object Detection



Ren, He, Girshick, and Sun, 2015

Image Segmentation

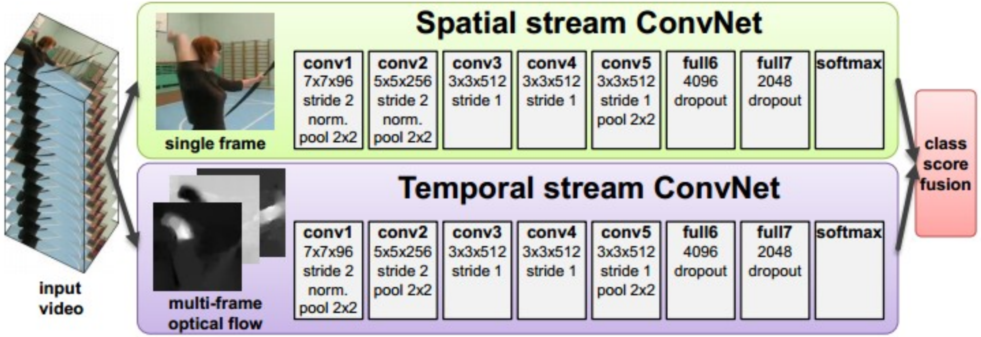


Fabaret et al, 2012

Slide inspiration: Justin Johnson

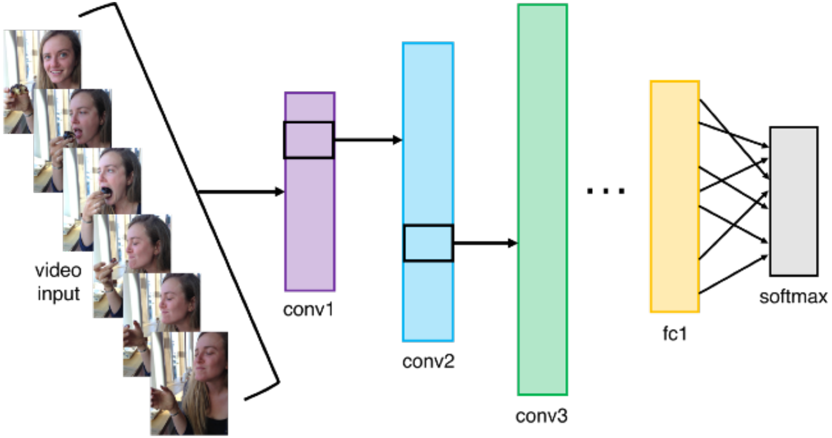
2012 to Present: Deep Learning is Everywhere

Video Classification



Simonyan et al, 2014

Activity Recognition

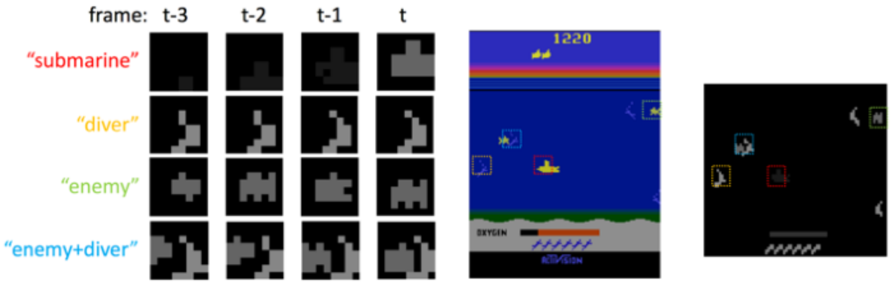


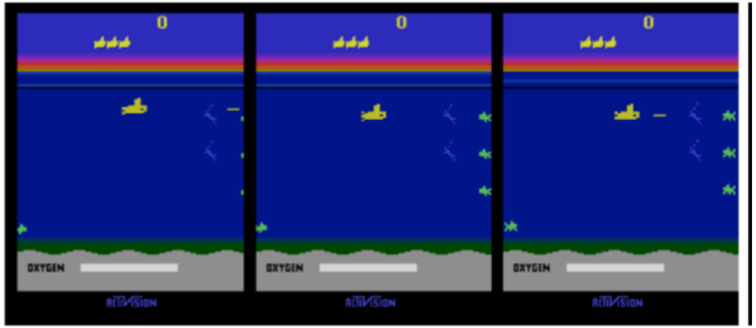
Slide inspiration: Justin Johnson

2012 to Present: Deep Learning is Everywhere

Pose Recognition (Toshev and Szegedy, 2014)

Playing Atari games (Guo et al, 2014)

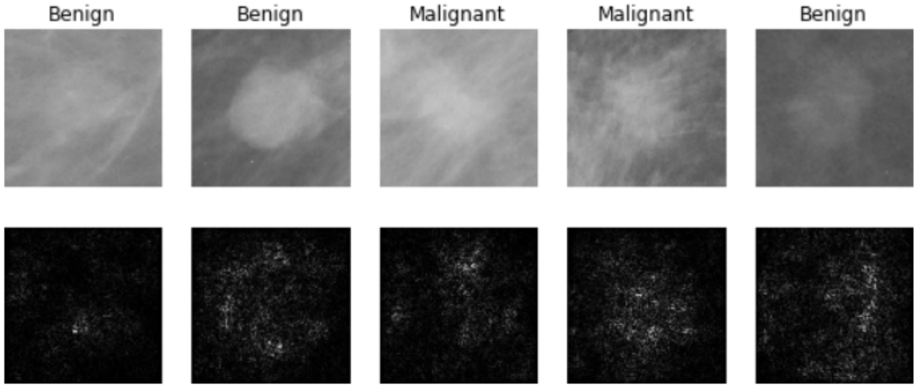




Slide inspiration: Justin Johnson

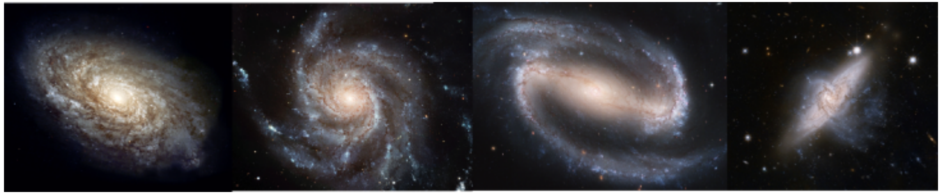
2012 to Present: Deep Learning is Everywhere

Medical Imaging



Levy et al, 2016 Figure reproduced with permission

Galaxy Classification



Dieleman et al, 2014

From left to right: public domain by NASA, usage permitted by ESA/Hubble, public domain by NASA, and public domain.

Whale recognition

Kaggle Challenge

This image by Christin Khan is in the public domain and originally came from the U.S. NOAA.

Slide inspiration: Justin Johnson

2012 to Present: Deep Learning is Everywhere

Image Captioning

Vinyals et al, 2015
Karpathy and Fei-Fei, 2015

A white teddy bear sitting in the grass

A man in a baseball uniform throwing a ball

A woman is holding a cat in her hand

A man riding a wave on top of a surfboard

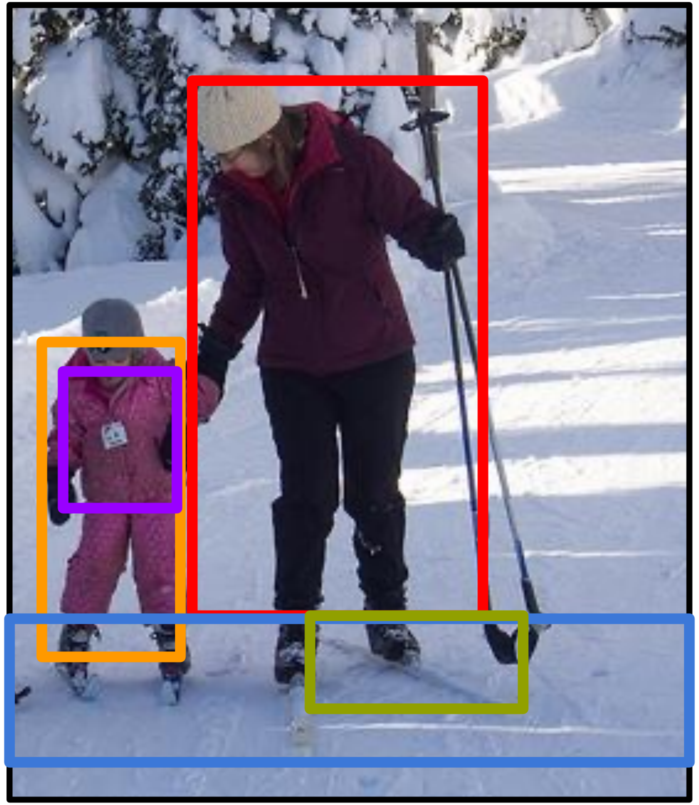
A cat sitting on a suitcase on the floor

A woman standing on a beach holding a surfboard

Captions generated by Justin Johnson using Neuraltalk2

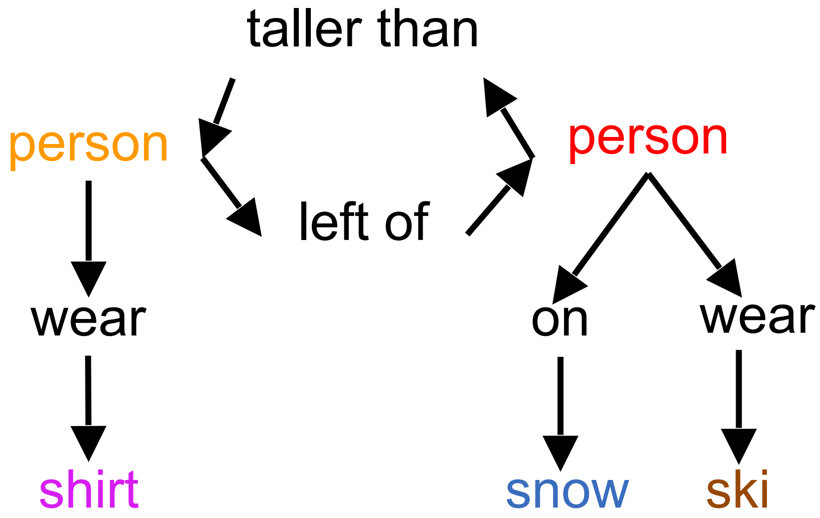
Slide inspiration: Justin Johnson

2012 to Present: Deep Learning is Everywhere

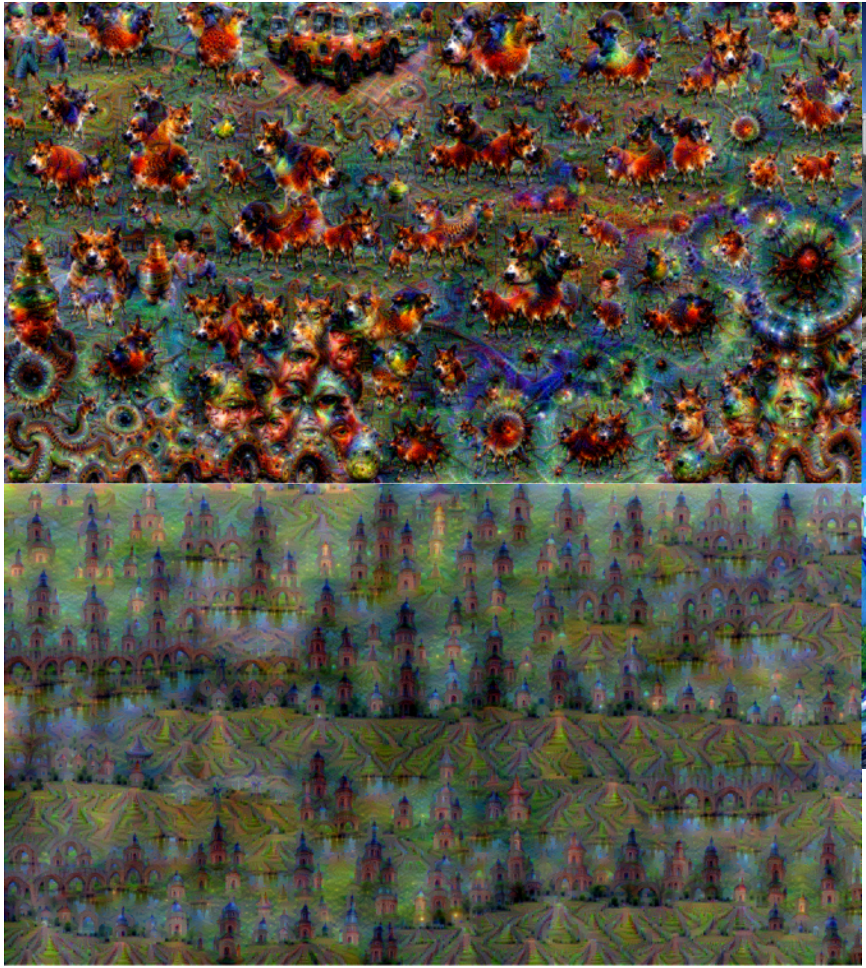


Results:

spatial, comparative, asymmetrical, verb, prepositional



Krishna*, Lu*, Bernstein, Fei-Fei, *ECCV* 2016



Figures copyright Justin Johnson, 2015. Reproduced with permission. Generated using the Inceptionism approach from a blog post by Google Research.

Original image is CC0 public domain
Starry Night and Tree Roots by Van Gogh are in the public domain
Bokeh image is in the public domain
Stylized images copyright Justin Johnson, 2017;
reproduced with permission

Mordvinsev et al, 2015

Gatys et al, 2016

Slide inspiration: Justin Johnson

2012 to Present: Deep Learning is Everywhere

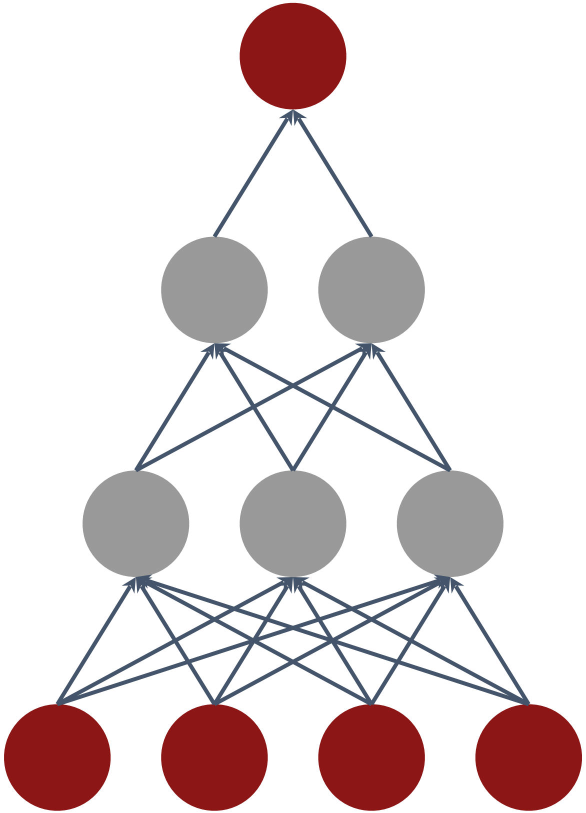
A Dutch still life of an arrangement of tulips in a fluted vase. The lighting is subtle, casting gentle highlights on the flowers and emphasizing their delicate details and natural beauty.

Slide inspiration: Justin Johnson

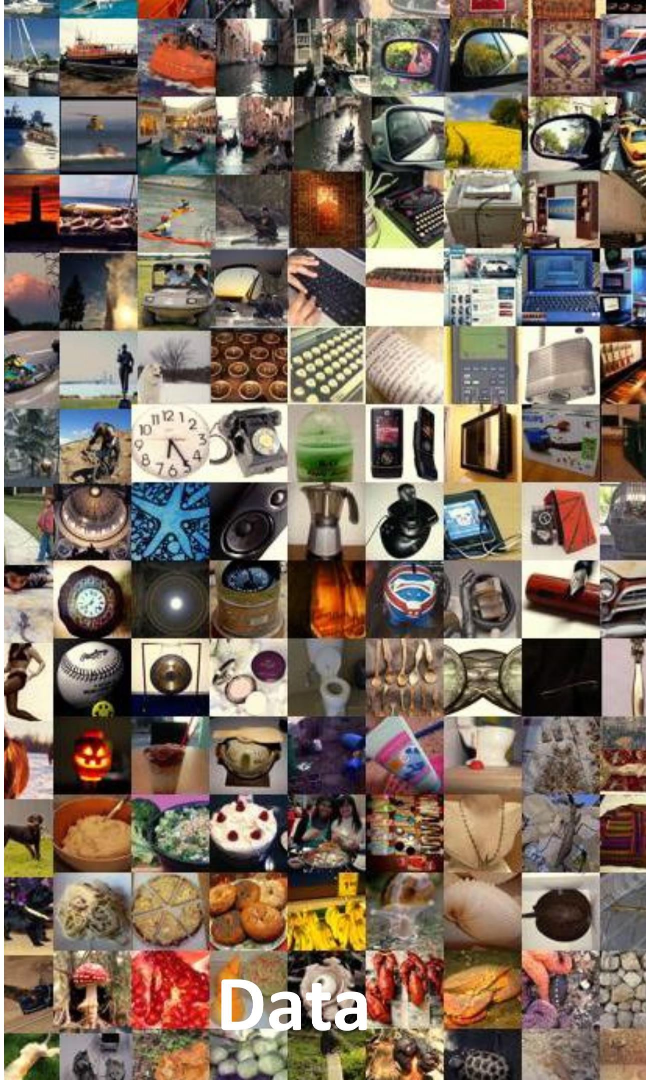
2012 to Present: Deep Learning is Everywhere

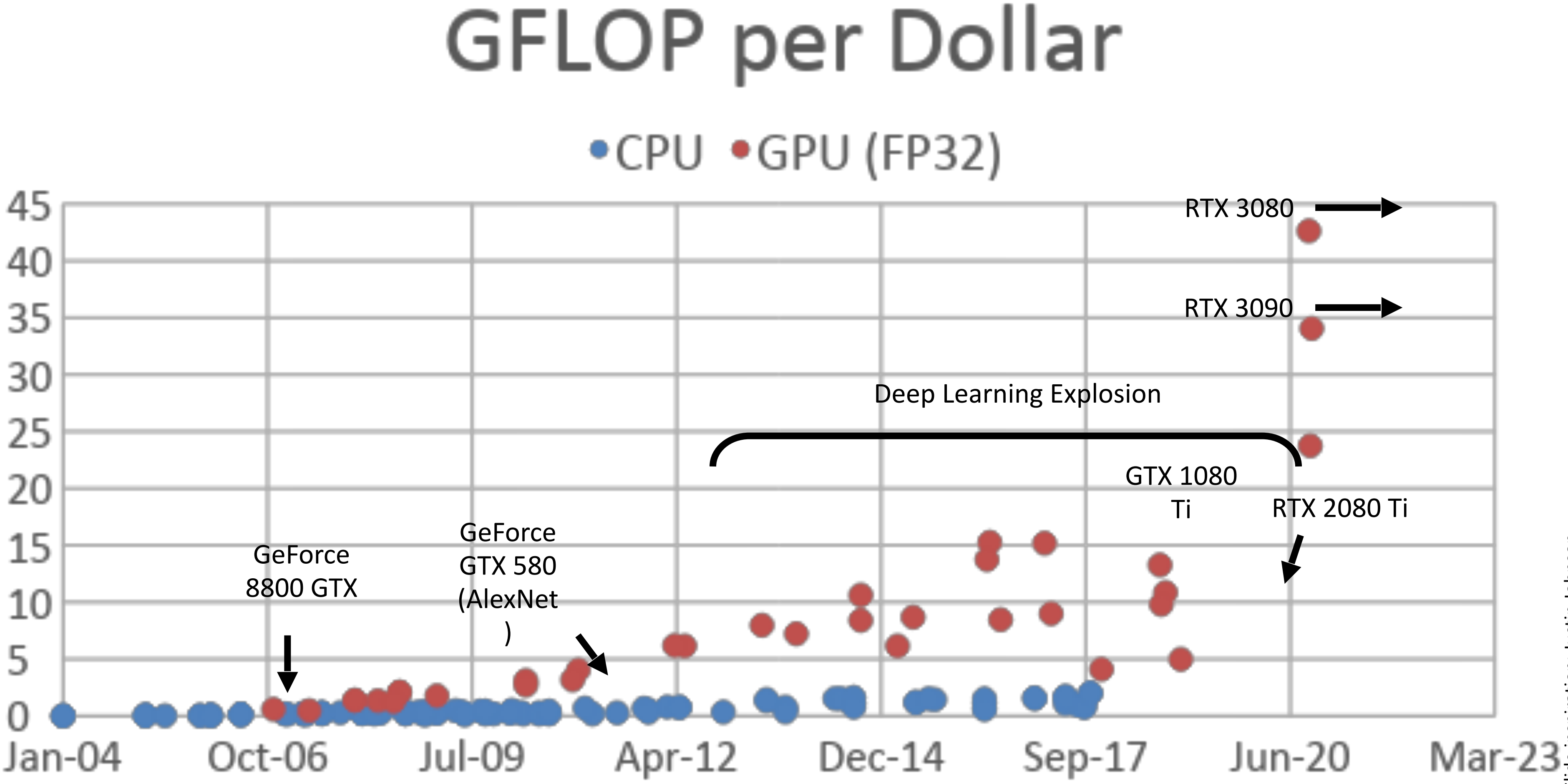
A 2D animation of a folk music band composed of anthropomorphic autumn leaves, each playing traditional bluegrass instruments, amidst a rustic forest setting dappled with the soft light of a harvest moon.

Slide inspiration: Justin Johnson



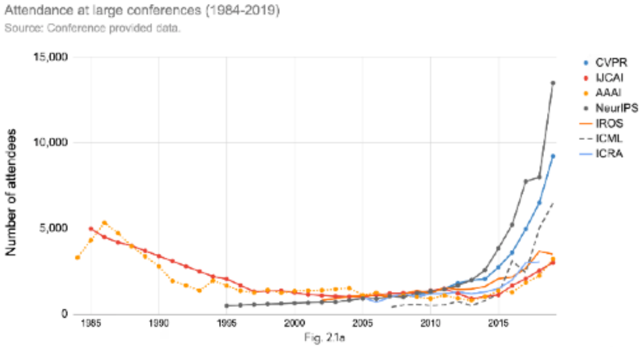
Algorithms



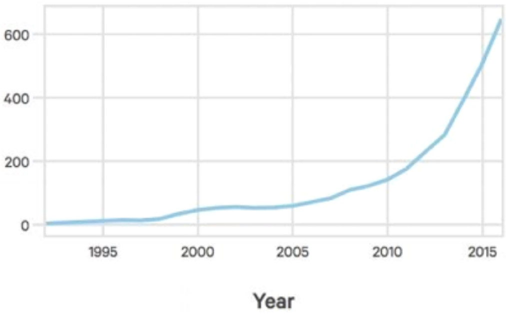


Slide inspiration: Justin Johnson

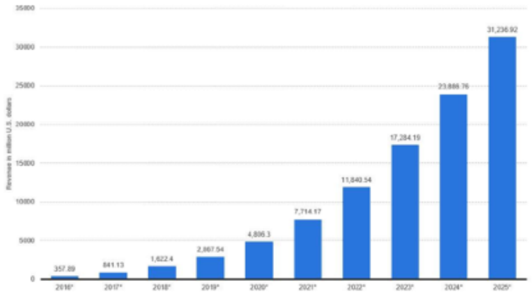
AI’s Explosive Growth & Impact



Number of attendance At AI conferences



Startups Developing AI Systems



Enterprise Application AI Revenue

Source: The Gradient

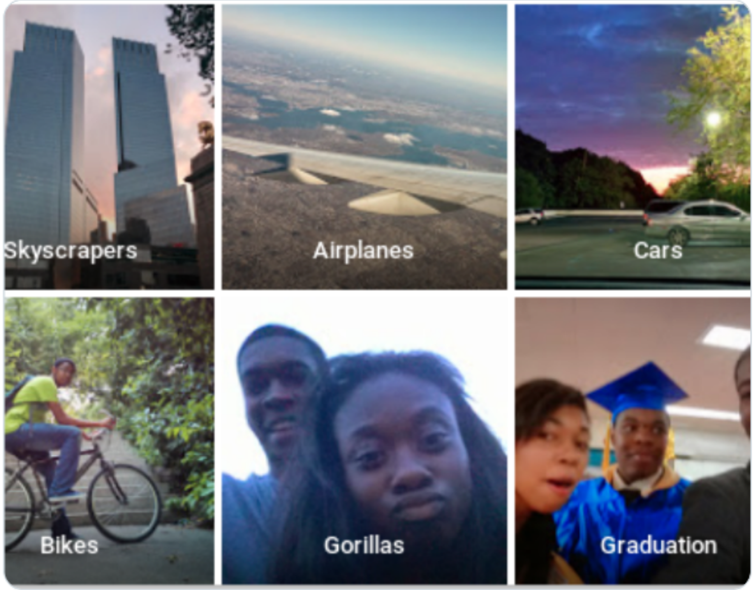
Source: Crunchbase, VentureSource, Sand Hill Econometrics

Source: Statista

Despite the successes, computer vision still has a long way to go

Computer Vision Can Cause Harm

Harmful Stereotypes



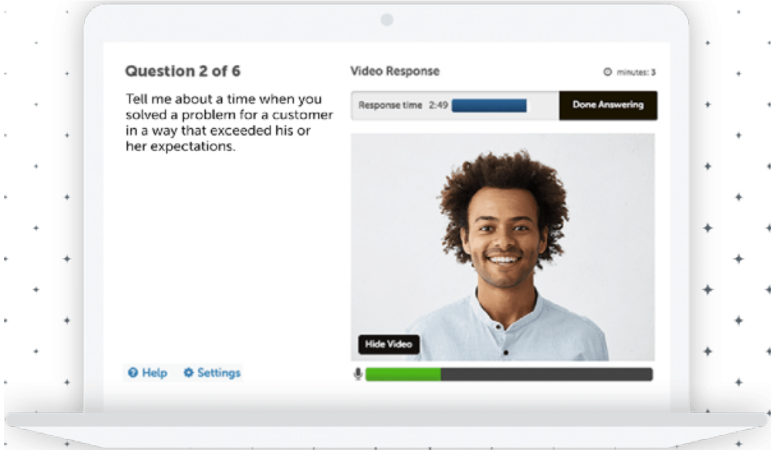
Barocas et al, "The Problem With Bias: Allocative Versus Representational Harms in Machine Learning", SIGCIS 2017
Kate Crawford, "The Trouble with Bias", NeurIPS 2017 Keynote
Source: https://twitter.com/jackyalcine/status/615329515909156865 (2015)

Affect people's lives

Technology

A face-scanning algorithm increasingly decides whether you deserve the job

HireVue claims it uses artificial intelligence to decide who's best for a job. Outside experts call it 'profoundly disturbing.'



Source: https://www.washingtonpost.com/technology/2019/10/22/ai-hiring-face-scanning-algorithm-increasingly-decides-whether-you-deserve-job/
https://www.hirevue.com/platform/online-video-interviewing-software

Example Credit: Timnit Gebru

And there is a lot we don't know how to do

https://fedandfit.com/wp-content/uploads/2020/06/summer-activities-for-kids_optimized-scaled.jpeg

Slide inspiration: Andrej Karpathy

Next lecture: Image Classification: A core task in Computer Vision

(assume given set of discrete labels)
{dog, cat, truck, plane, ...}

→ cat