DATA ANALYTICS USING DEEP LEARNING

GT 8803 // FALL 2019 // JOY ARULRAJ

LECTURE #01: COURSE INTRODUCTION

CREATING THE NEXT®

WELCOME TO 8803-DDL

- This is a cross-cutting course!
- Gain holistic understanding of three areas
  - Data Analytics
  - Machine Learning
  - Computer Vision
- Bridge the gap between systems and machine learning

CREDITS

- This course is derived from two courses
- Convolutional Neural Networks for Visual Recognition
  - Fei Fei Li, Andrej Karpathy, and Justin Johnson
  - http://cs231n.stanford.edu/
- Advanced database systems
  - Andy Pavlo
  - https://15721.courses.cs.cmu.edu/

TODAY'S AGENDA

- Course Overview
- Course Objectives
- Course Logistics
- History of Computer Vision
- Visual Recognition Overview

COURSE OVERVIEW

BIG DATA & DATA SCIENCE ERA

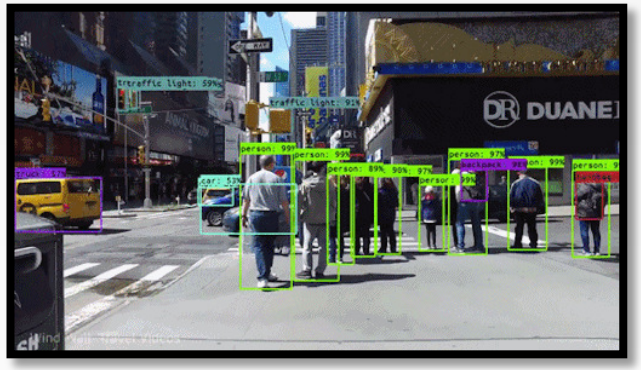
- Visual data is the biggest Big Data out there

MILLIONS OF **IMAGES**
UPLOADED EVERY DAY

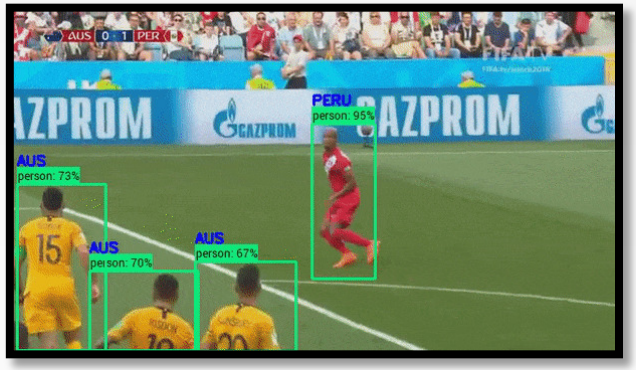
HOURS OF **VIDEOS**
UPLOADED EVERY MINUTE

NEXT-GENERATION APPS

- Apps will focus on visual data



SELF-DRIVING CARS



SPORTS ANALYTICS

CHALLENGES: TRADITIONAL DATABASE SYSTEMS

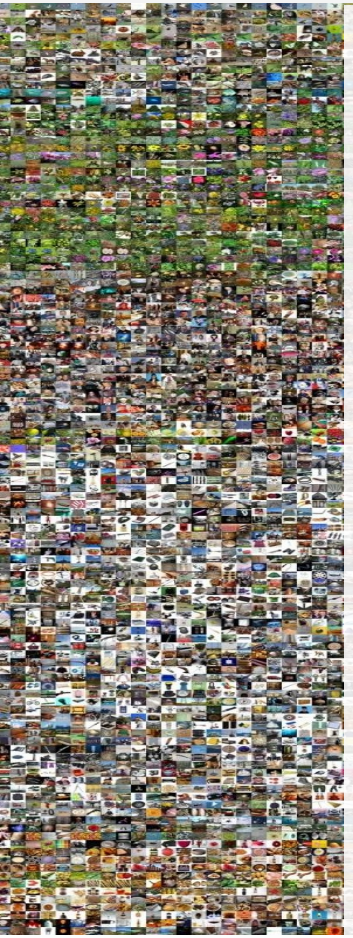
- Traditional database systems only support structured data

EMPLOYEE ID	NAME	AGE	SALARY
101	PETER	25	100K
102	JOHN	20	80K
103	MARK	30	120K

WHY IS THIS IMPORTANT NOW?

- Modern computer vision techniques have made great strides
  - Near human-levels of accuracy for several visual data analytics tasks

EXAMPLE: IMAGE CLASSIFICATION



IMAGENET

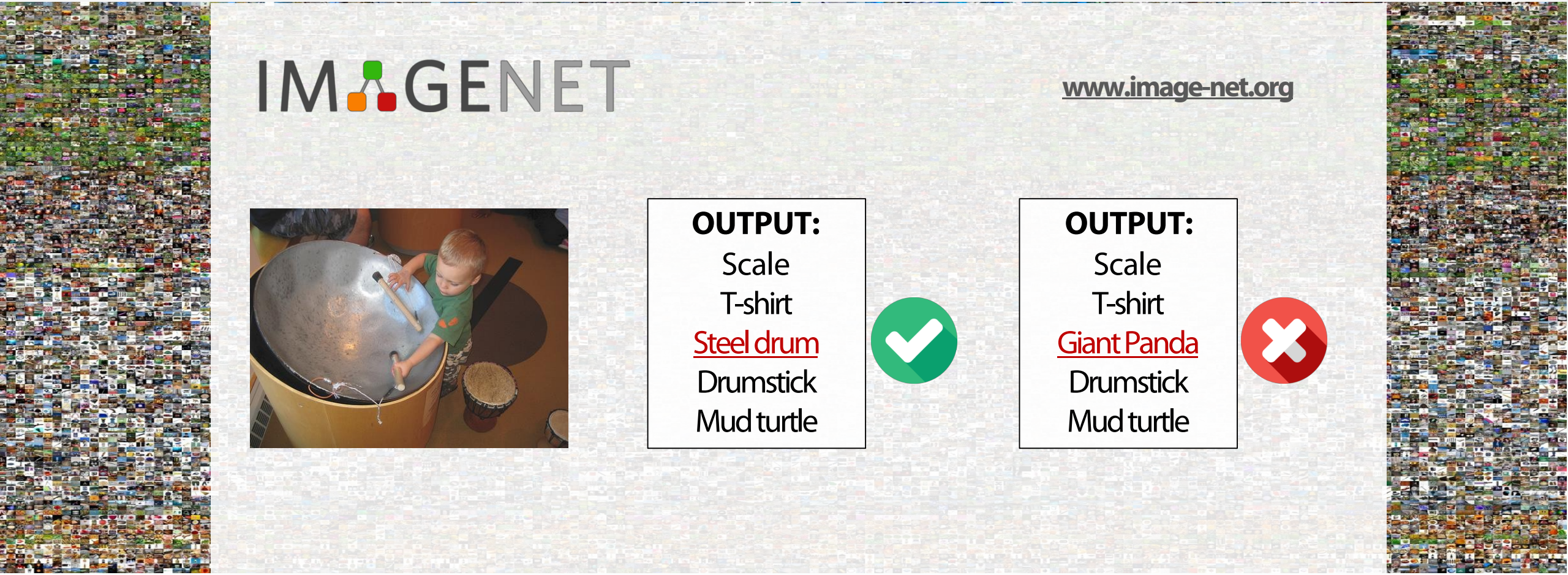
www.image-net.org

22K categories and 15 M images

- Animals
  - Bird
  - Fish
  - Mammal
  - Invertebrate
- Plants
  - Tree
  - Flower
- Food
- Materials
- Structures
- Artifact
  - Tools
  - Appliances
  - Structures
- Person
- Scenes
  - Indoor
  - Geological Formations
- Sport Activities

Deng, Dong, Socher, Li, Li, & Fei-Fei, 2009

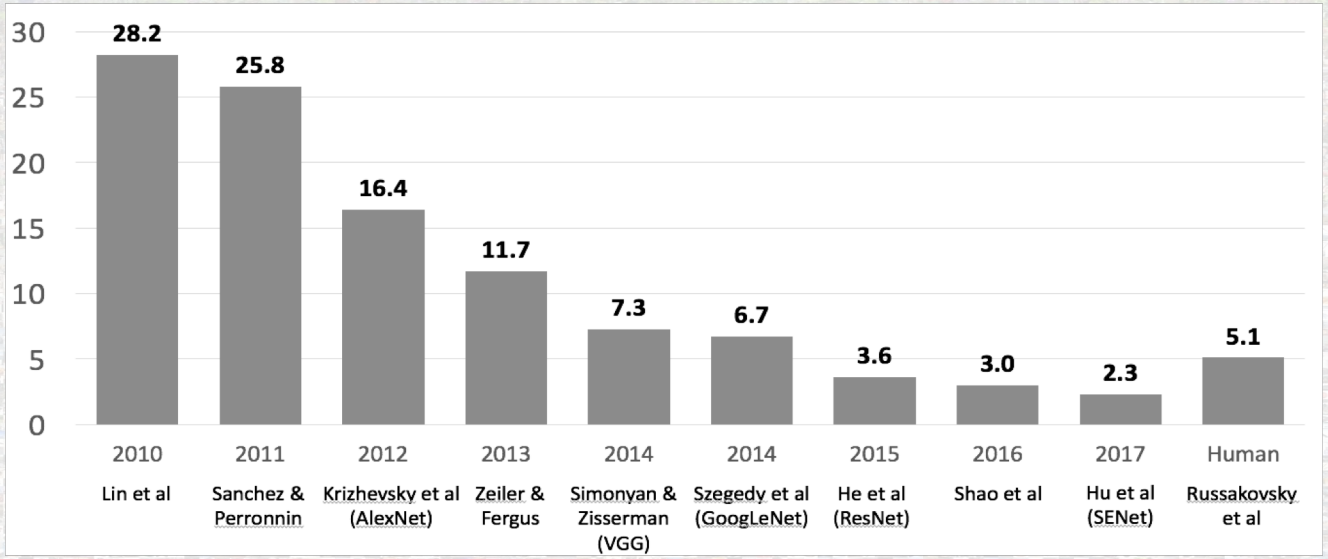
EXAMPLE: IMAGE CLASSIFICATION



EXAMPLE: IMAGE CLASSIFICATION

IMAGENET

www.image-net.org



Russakovsky et al., IJCV 2015

CHALLENGES: DEEP LEARNING MODELS

- Computational Efficiency
- Usability

CHALLENGES: COMPUTER VISION PIPELINES

- Computational Efficiency
  - These pipelines are computationally infeasible at scale
  - Example: State-of-the-art object detection models run at 3 frames per second (fps) (e.g., Mask R-CNN)
  - It will take 8 decades of GPU time to process 100 cameras over a month of video.

CHALLENGES: COMPUTER VISION PIPELINES

- Usability
  - These techniques require complex, imperative programming across many low-level libraries (e.g., Pytorch and OpenCV)
  - This is an ad-hoc, tedious process that ignores opportunity for cross-operator optimization
  - Traditional database systems were successful due to their ease of use (i.e., SQL is declarative)

GOAL: DECLARATIVE VIDEO ANALYTICS SYSTEM

COURSE OBJECTIVES

WHY SHOULD YOU TAKE THIS COURSE?

- There are many challenging problems in database systems & machine learning
- Systems + ML developers are in demand
- If you are good enough to write code for a ML-driven data analytics system, then you can write code on almost anything else

COURSE OBJECTIVES

- Learn about cutting-edge research topics in data analytics and deep learning
- Learn about modern practices in systems programming and deep learning
- We will cover state-of-the-art topics
- This is **not** a course on classical database systems

PRE-REQUISITES

- Proficiency in Python and some high-level familiarity with C++
  - All assignments will be in Python; but some of the deep learning libraries we may look at later in the class will be written in C++
  - A Python tutorial is available on course website
- Calculus, Linear Algebra
- Basic Probability and Statistics

PRE-REQUISITES

- Fundamentals of Machine Learning
  - We will be formulating cost functions, taking derivatives and performing optimization with gradient descent
- I am happy to have people from different backgrounds
  - But talk to me if you're not sure

COURSE LOGISTICS

COURSE LOGISTICS

- Office: Klaus 3324
- On-line Discussion through Piazza:
  - https://piazza.com/gatech/fall2019/cs8803ddl/home
- For all technical questions, please use Piazza
  - Don't email me directly
  - All non-technical questions should be sent to me

COURSE LOGISTICS

- Course Schedule
  - https://www.cc.gatech.edu/~jarulraj/courses/8803-f19/pages/schedule.html
  - We will post lecture slides and course materials on this page
- Course Policies
  - Students are expected to abide by the Georgia Tech Honor Code
  - If you are not sure, ask me

COURSE LOGISTICS

- Grading Tool: Gradescope
  - https://www.gradescope.com/courses/54455
  - You will get immediate feedback on your programming assignments
  - You can iteratively improve your score over time

GRADE BREAKDOWN

- The final grade for the course will be tentatively based on the following weights:
  - 30% Assignments
  - 30% Midterm Exam
  - 40% Group Project
- Emphasis on learning rather than testing you
  - If your project is truly amazing, you get an automatic A!

TEACHING ASSISTANTS

- TA #1: Jaeho Bang
  - Ph.D. student in Computer Science
  - B.S. from Carnegie Mellon

- TA #2: TBD

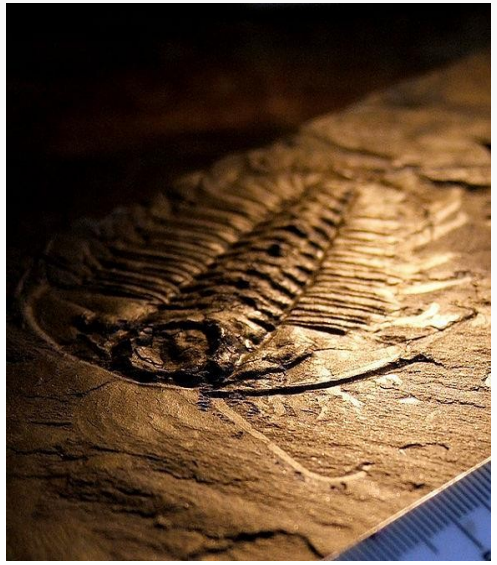
OFFICE HOURS

- Immediately before class
  - Me: Mon/Wed 3:30 – 4:30 PM
  - Jaeho: Tue/Thu 3:30 – 4:30 PM
  - Near my office (Klaus 3324)
- Things we can talk about
  - Questions related to lectures and assignments
  - Project ideas
  - Can't give relationship advice

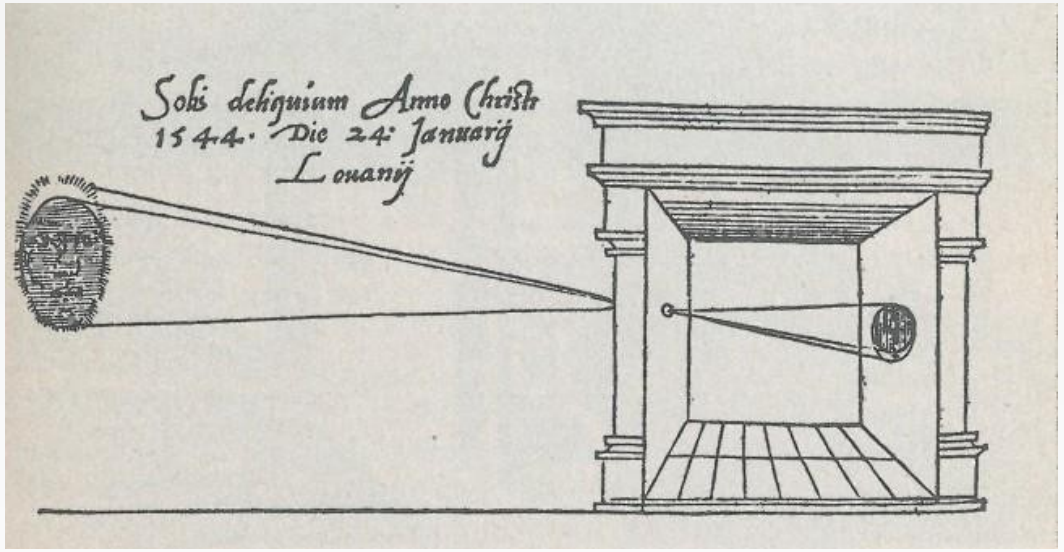
HISTORY OF COMPUTER VISION

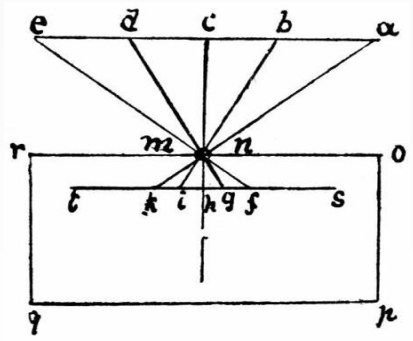
EVOLUTION'S BIG BANG

- ~543 million years
  - What was life like back then?
  - Onset of vision triggered evolution's Big Bang
  - Now biggest sensory system in most animals



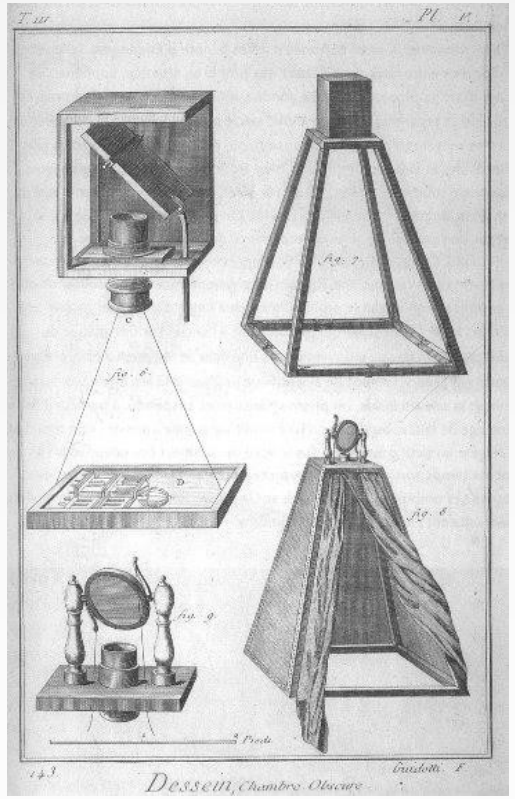
CAMERA OBSCURA





GEMMA FRISIUS (1545)

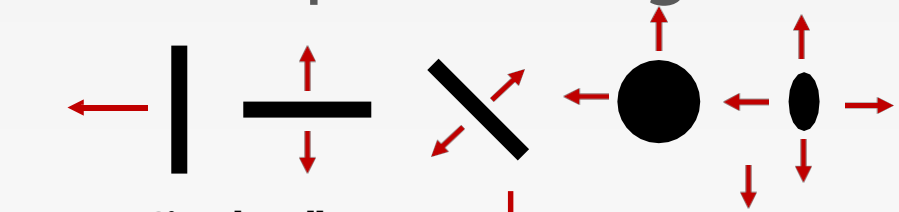
DA VINCI (~1500)



ENCYCLOPEDIE (~1800)

ELECTROPHYSIOLOGY (1959)

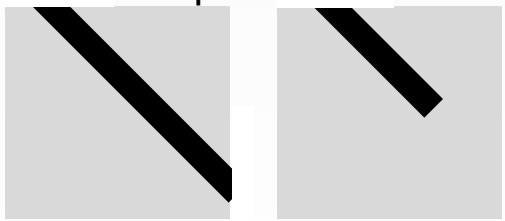
- Visual processing mechanism in mammals



Simple cells:
Response to light orientation

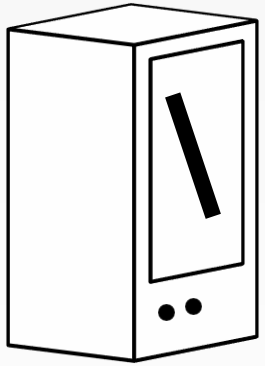
Complex cells:
Response to light orientation and movement

Hypercomplex cells:
Response to movement with an end point

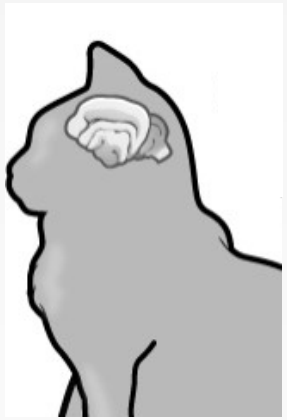


No response

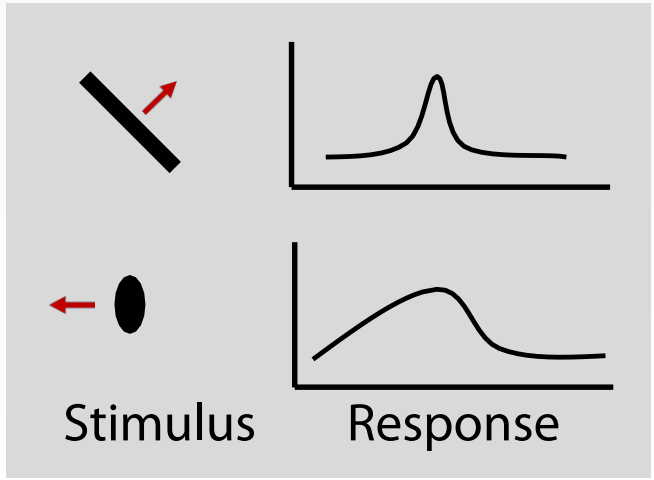
Response (end point)



Stimulus



Electrical signal from brain

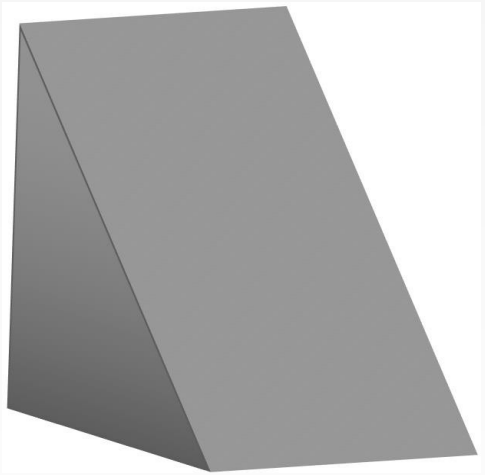


Stimulus

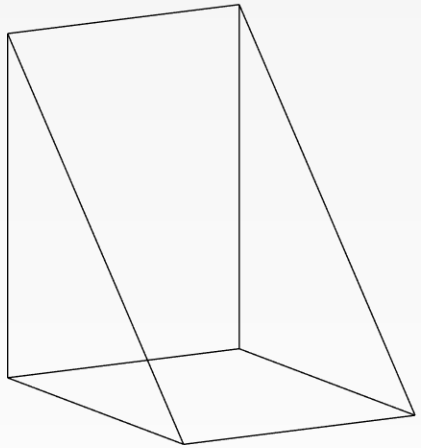
Response

BLOCK WORLD (1961)

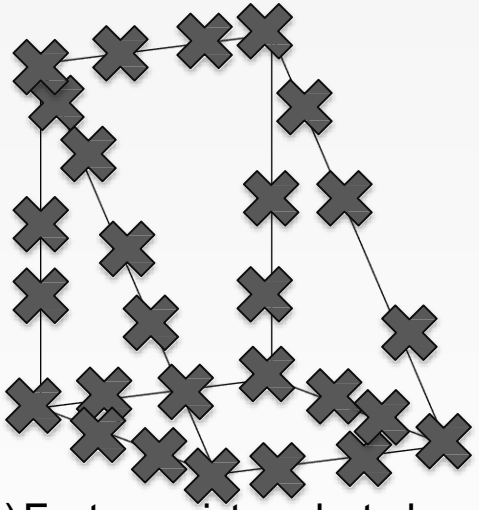
- Visual world simplified into geometric shapes



(a) Original picture

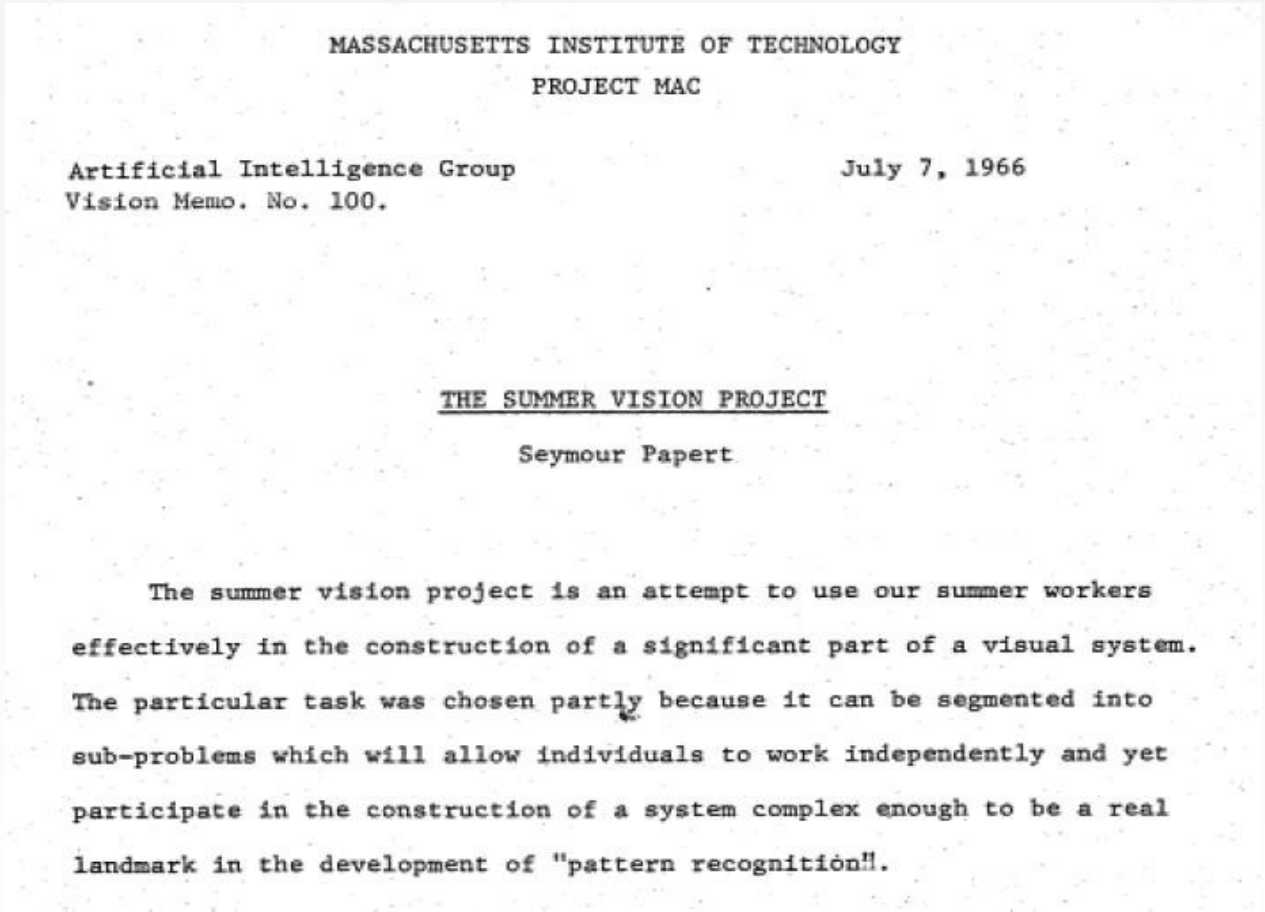


(b) Differentiated picture



(c) Feature points selected

PROJECT MAC (1966)



MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

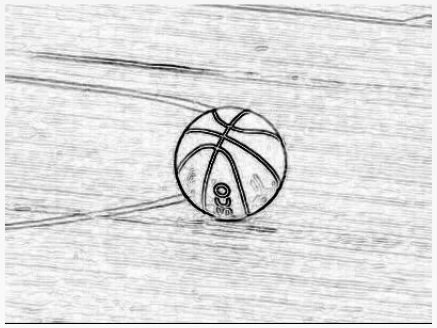
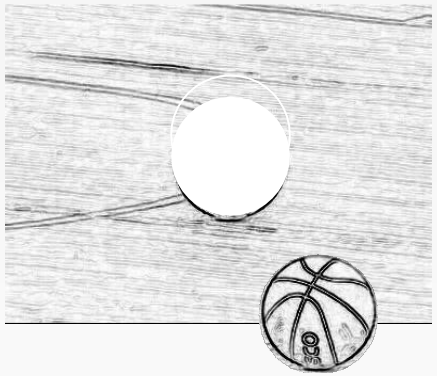
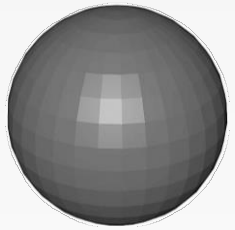
Artificial Intelligence Group — July 7, 1966
Vision Memo. No. 100.

THE SUMMER VISION PROJECT

Seymour Papert

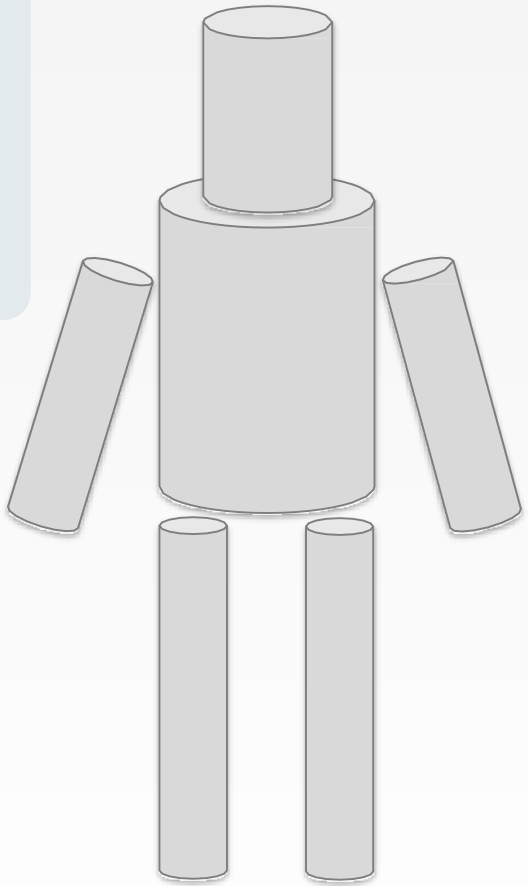
The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the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

STAGES OF VISUAL REPRESENTATION (1970s)

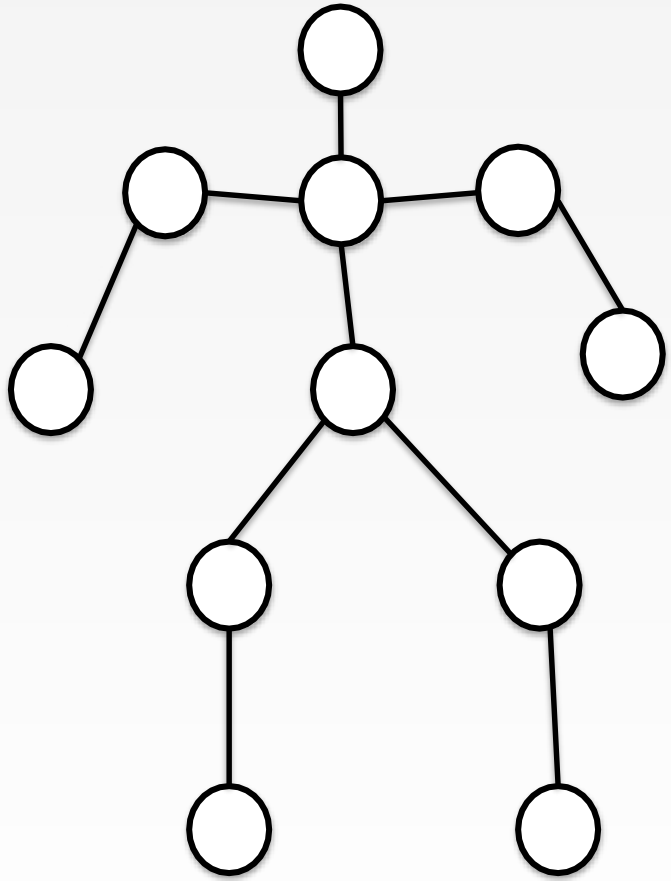
INPUT IMAGE	EDGE IMAGE	2.5-D MODEL	3-D MODEL
			
Perceived Intensities	Zero crossings, edges, bars, ends, virtual lines, groups, curves boundaries	Local surface orientation and discontinuities in depth and surface orientation	3-D models hierarchically organized in terms of surface and volumetric primitives

BETTER REPRESENTATIONS (1970s)

GENERALIZED CYLINDER (1979)



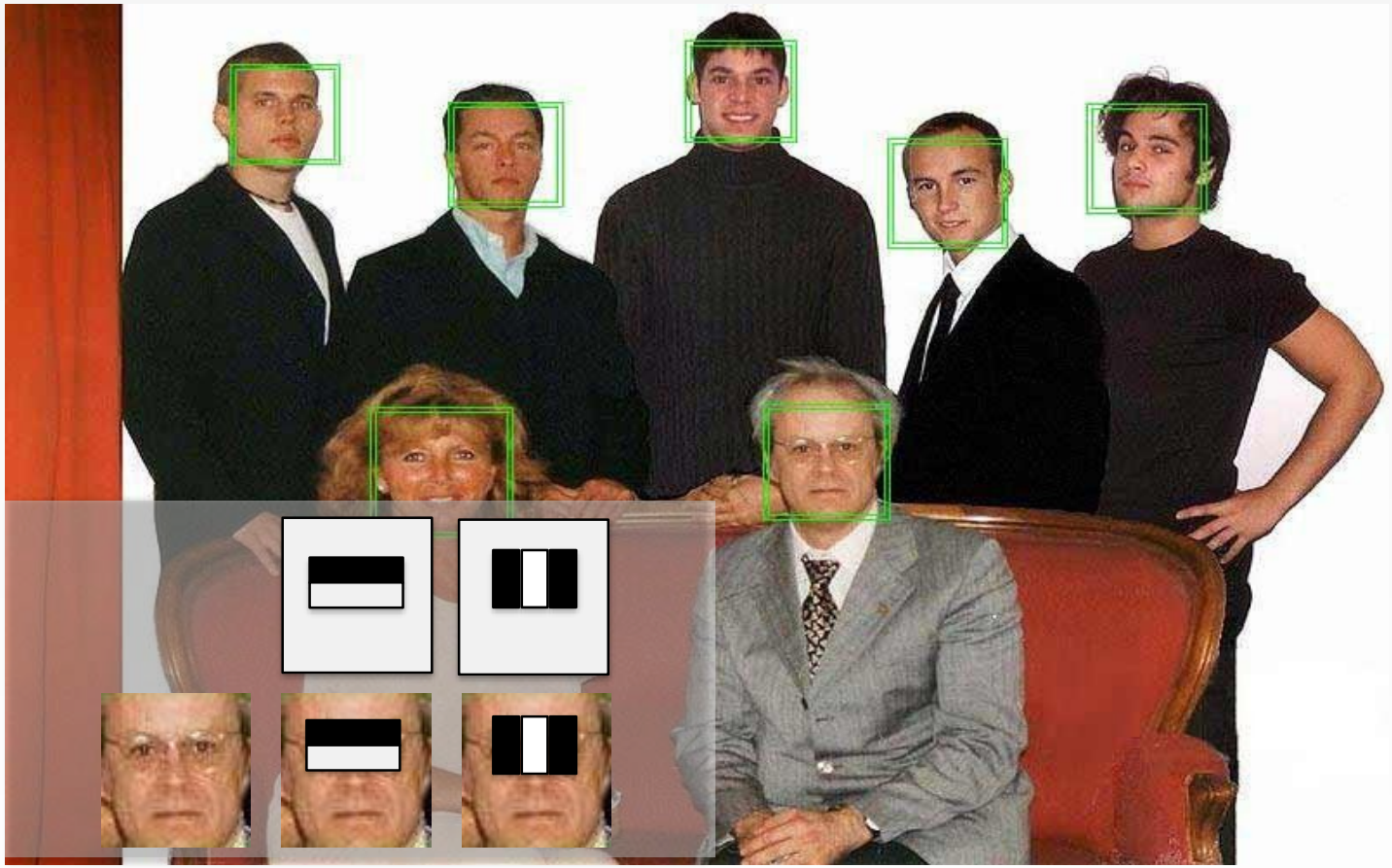
PICTORIAL STRUCTURE (1973)



OBJECT RECOGNITION (1987)

IMAGE SEGMENTATION (1987)

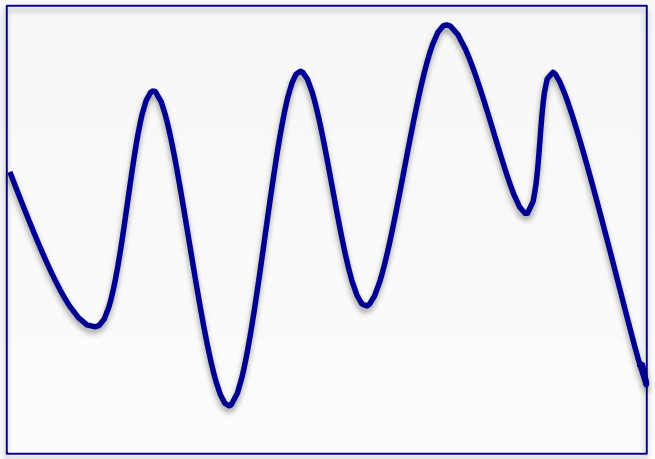
FACE DETECTION (2001)



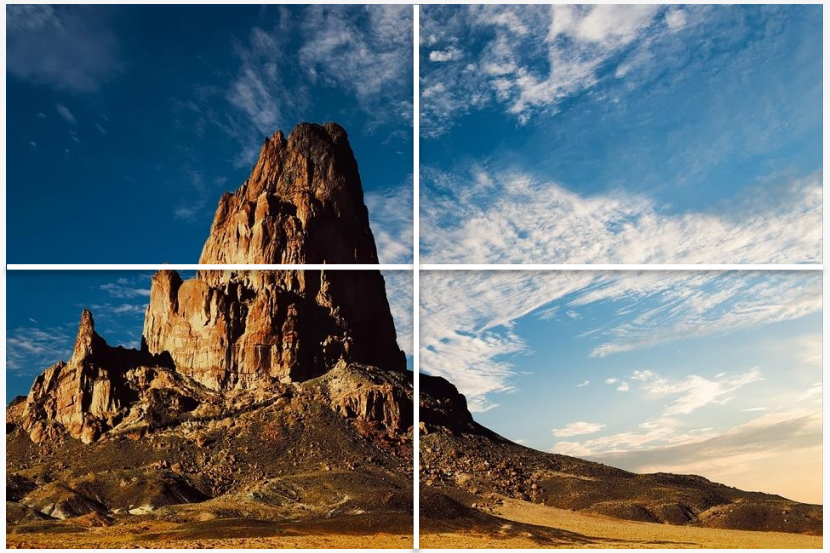
FEATURE-BASED OBJECT RECOGNITION (1999)

- Certain features are invariant to perspective

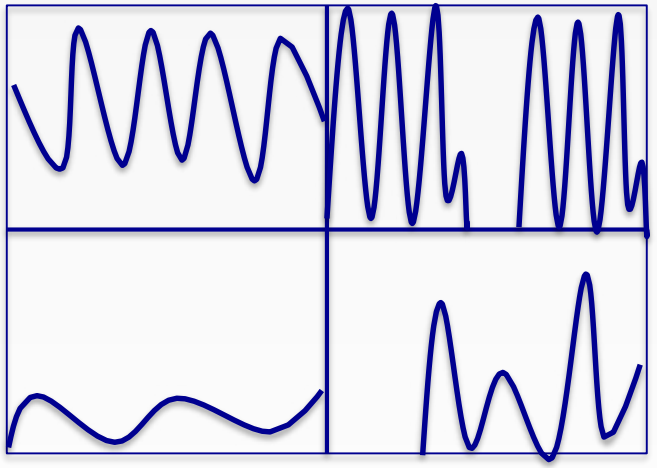
FEATURE MATCHING (2006)



LEVEL 0

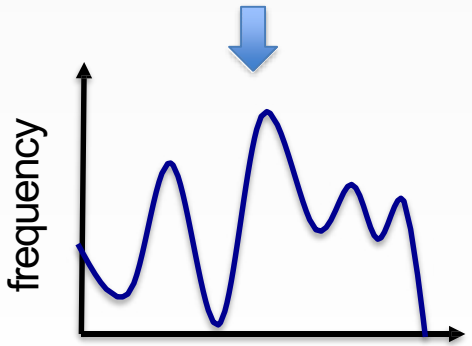


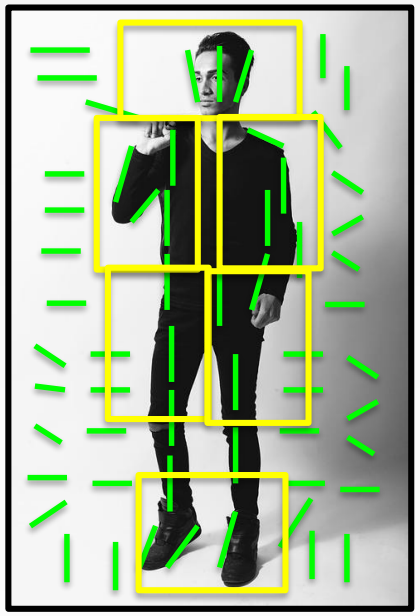
SPATIAL PYRAMID



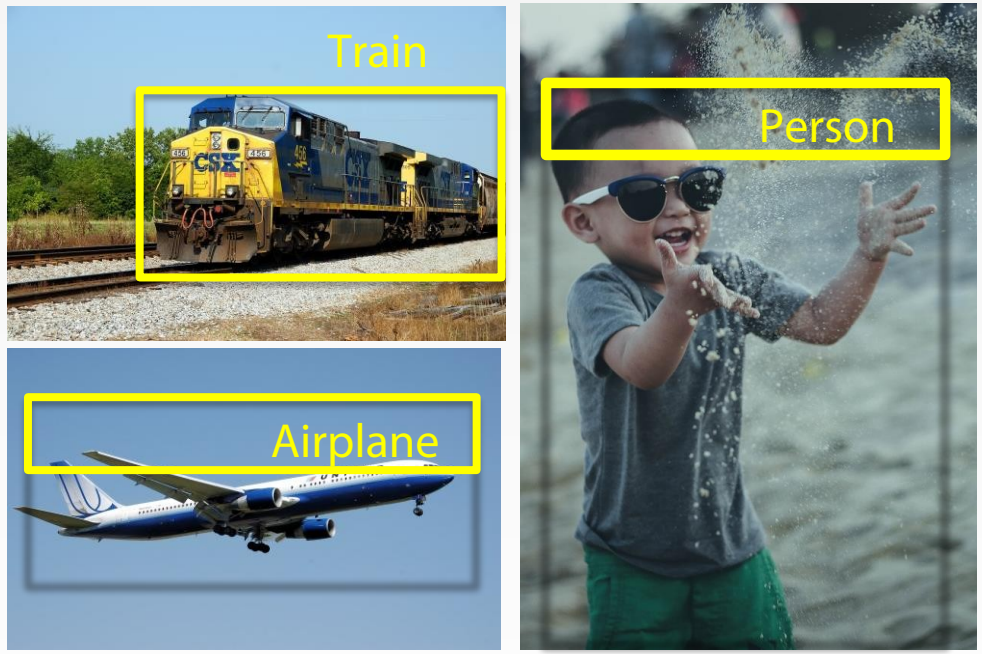
LEVEL 1

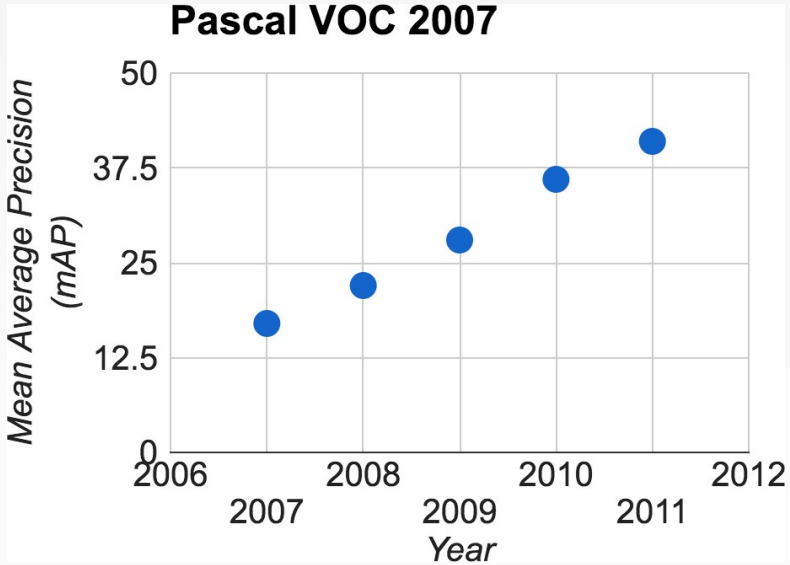
HUMAN POSE DETECTION (2005)





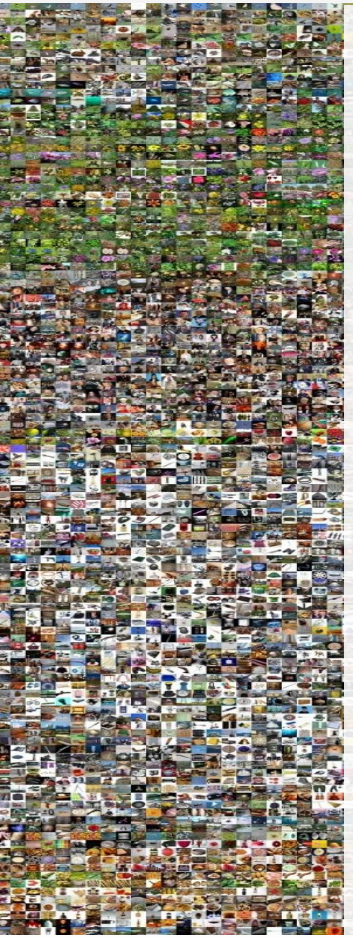
PASCAL VISUAL OBJECT CHALLENGE (2006~12)





20 OBJECT CATEGORIES

IMAGENET CHALLENGE (2009~17)



IMAGENET

www.image-net.org

22K categories and 15M images

- Animals
  - Bird
  - Fish
  - Mammal
  - Invertebrate
- Plants
  - Tree
  - Flower
- Food
- Materials
- Structures
- Artifact
  - Tools
  - Appliances
  - Structures
- Person
- Scenes
  - Indoor
  - Geological Formations
- Sport Activities

Deng, Dong, Socher, Li, Li, & Fei-Fei, 2009

IMAGENET CHALLENGE (2009~17)

IMAGENET

www.image-net.org

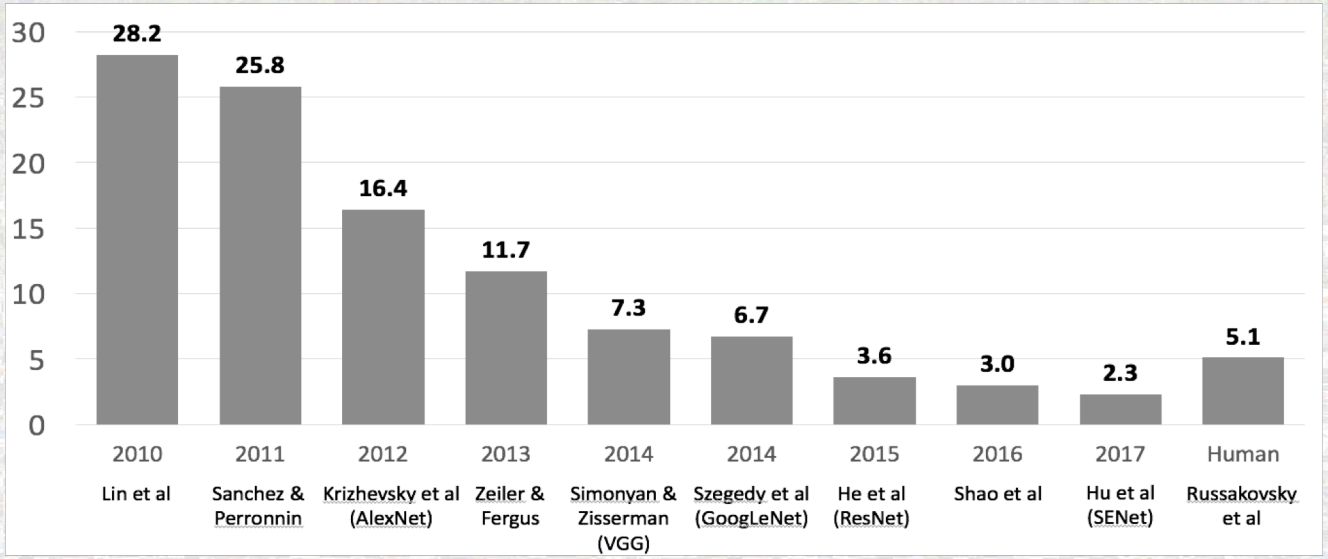
OUTPUT:
Scale
T-shirt
Steel drum
Drumstick
Mud turtle

OUTPUT:
Scale
T-shirt
Giant Panda
Drumstick
Mud turtle

IMAGENET CHALLENGE (2009~17)

IMAGENET

www.image-net.org



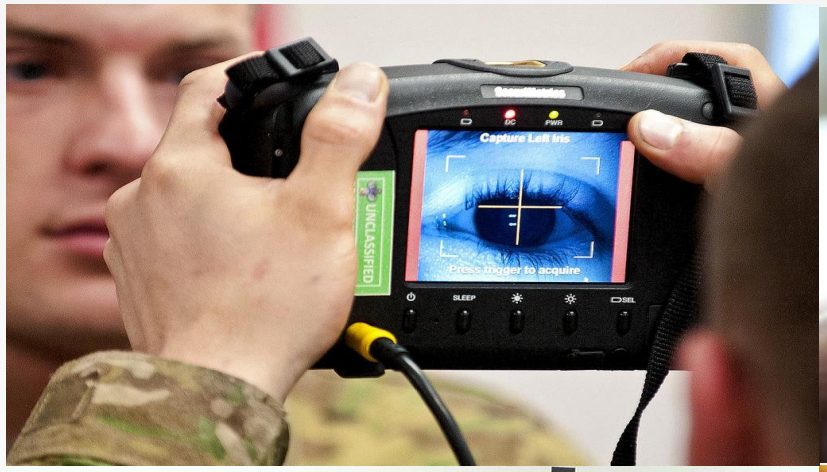
Russakovsky et al., IJCV 2015

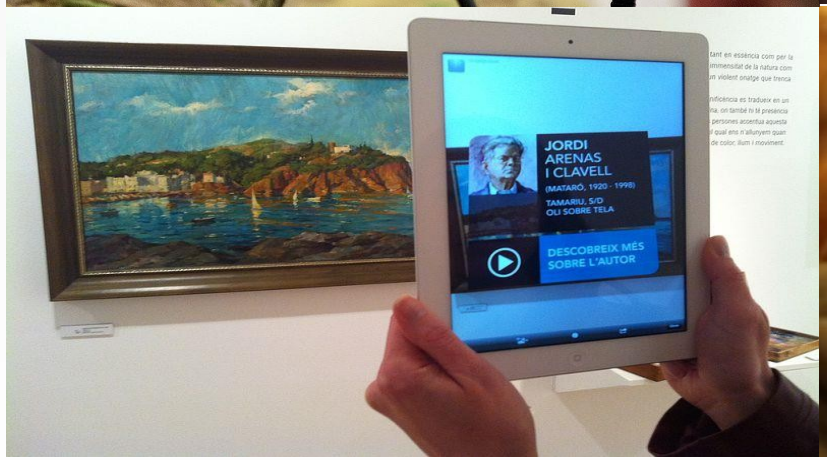
VISUAL RECOGNITION OVERVIEW

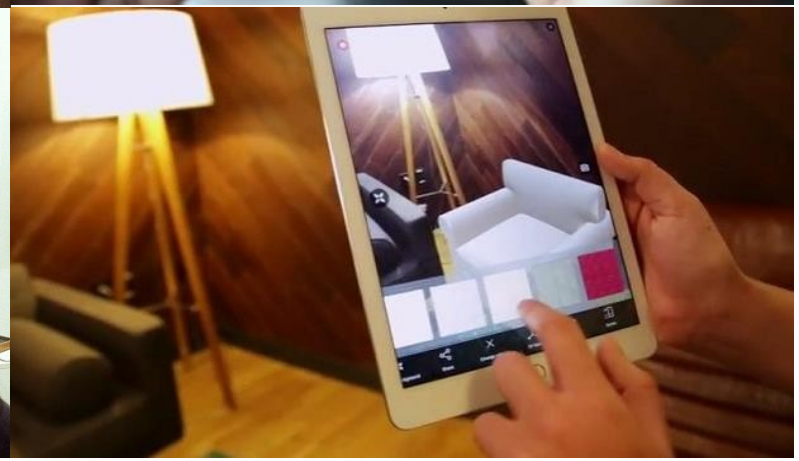
IMAGE CLASSIFICATION

- This course will focus on one of the most fundamental problems of visual recognition
  - Image classification
- This technique can be applied in many ways

IMAGE CLASSIFICATION



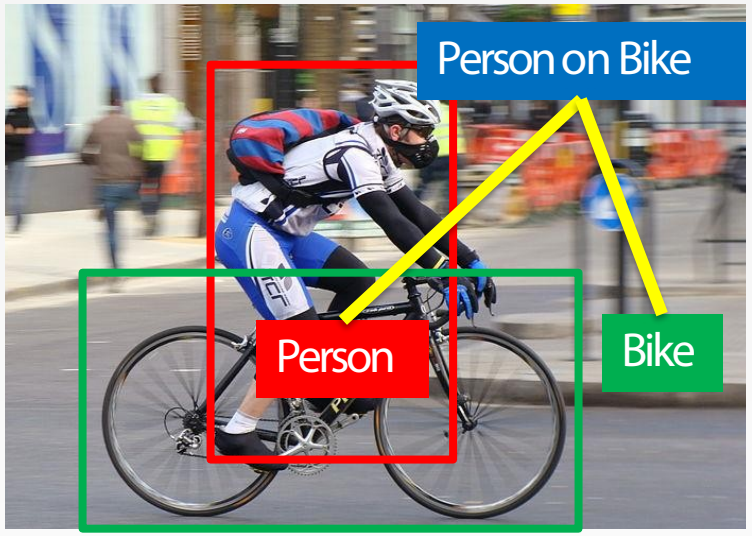




OTHER VISUAL RECOGNITION PROBLEMSS

- There are many visual recognition problems related to image classification
  - *Action classification*
  - *Image captioning*
  - *Object detection*
- Tools developed for image classification can be reused for these other problems as well

OTHER VISUAL RECOGNITION PROBLEMSS

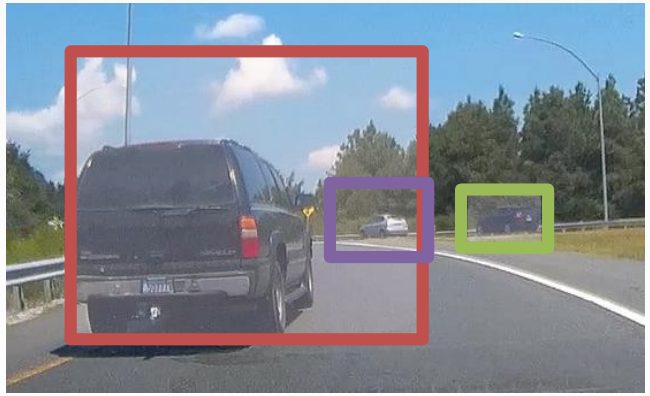


ACTION CLASSIFICATION

IMAGE CAPTIONING

Person

Hammer



OBJECT DETECTION

CONVOLUTIONAL NEURAL NETWORKS

- *Convolutional Neural Networks (CNNs)* have become an important tool for object recognition

IMAGENET CHALLENGE

Year 2010

NEC-UIUC

Dense descriptor grid: HOG, LBP

Coding: local coordinate, super-vector

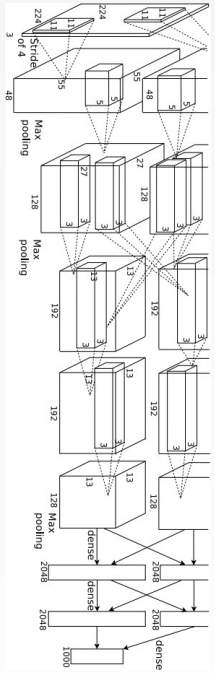
Pooling, SPM

Linear SVM

[Lin CVPR 2011]

Year 2012

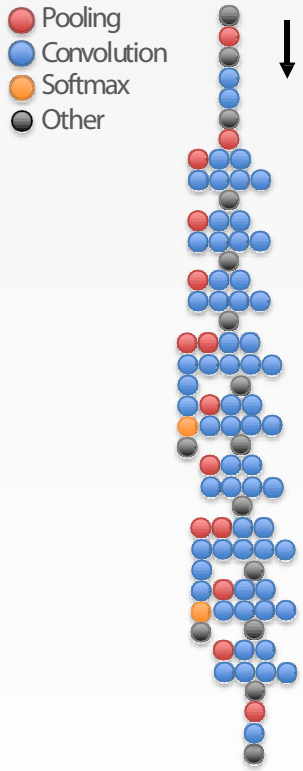
AlexNet



[Krizhevsky NIPS 2012]

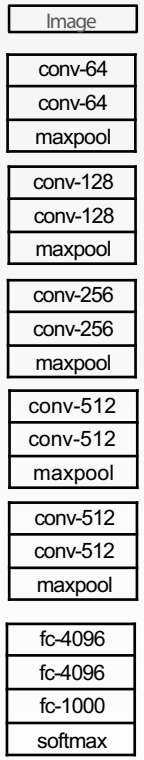
Year 2014

Google LeNet



[Szegedy arxiv 2014]

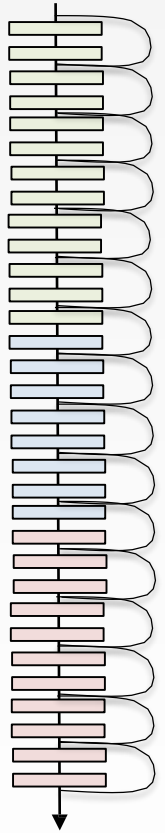
VGG



[Simonyan arxiv 2014]

Year 2015

MSR Asia



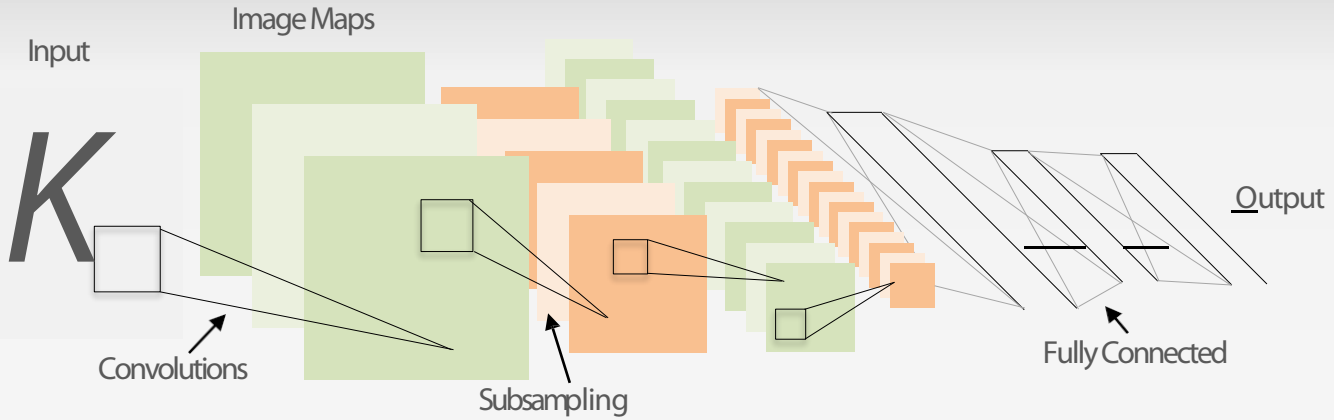
[He ICCV 2015]

CONVOLUTIONAL NEURAL NETWORKS

- They were not invented overnight

1998

LeCun et al.



OF TRANSISTORS

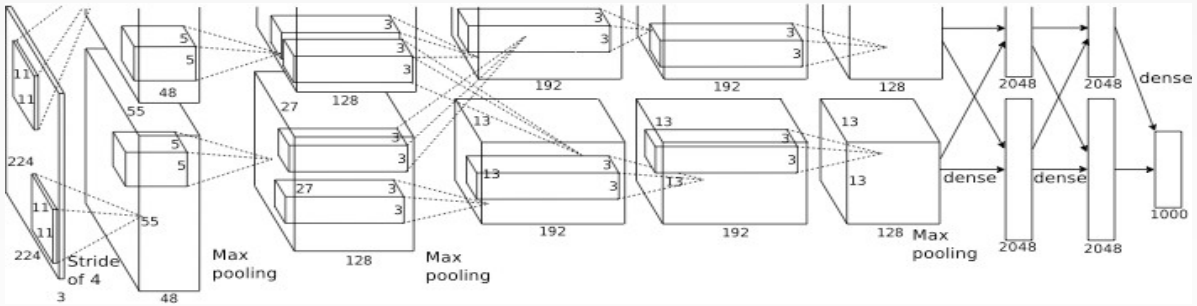
10^6

OF PIXELS USED IN TRAINING

10^7 NIST

2012

Krizhevsky et al.



OF TRANSISTORS

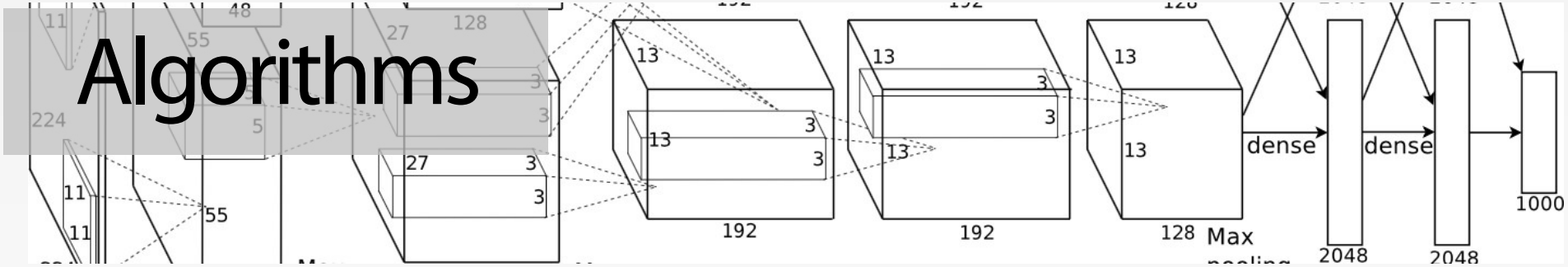
10^9

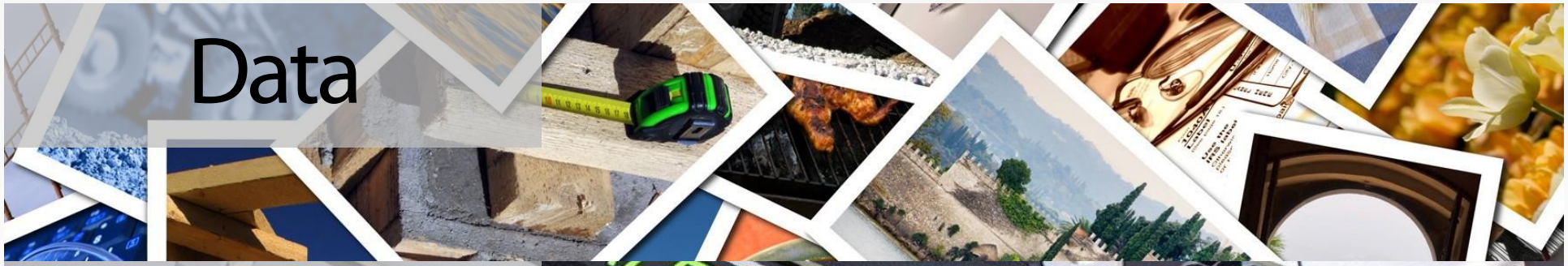
GPUs

OF PIXELS USED IN TRAINING

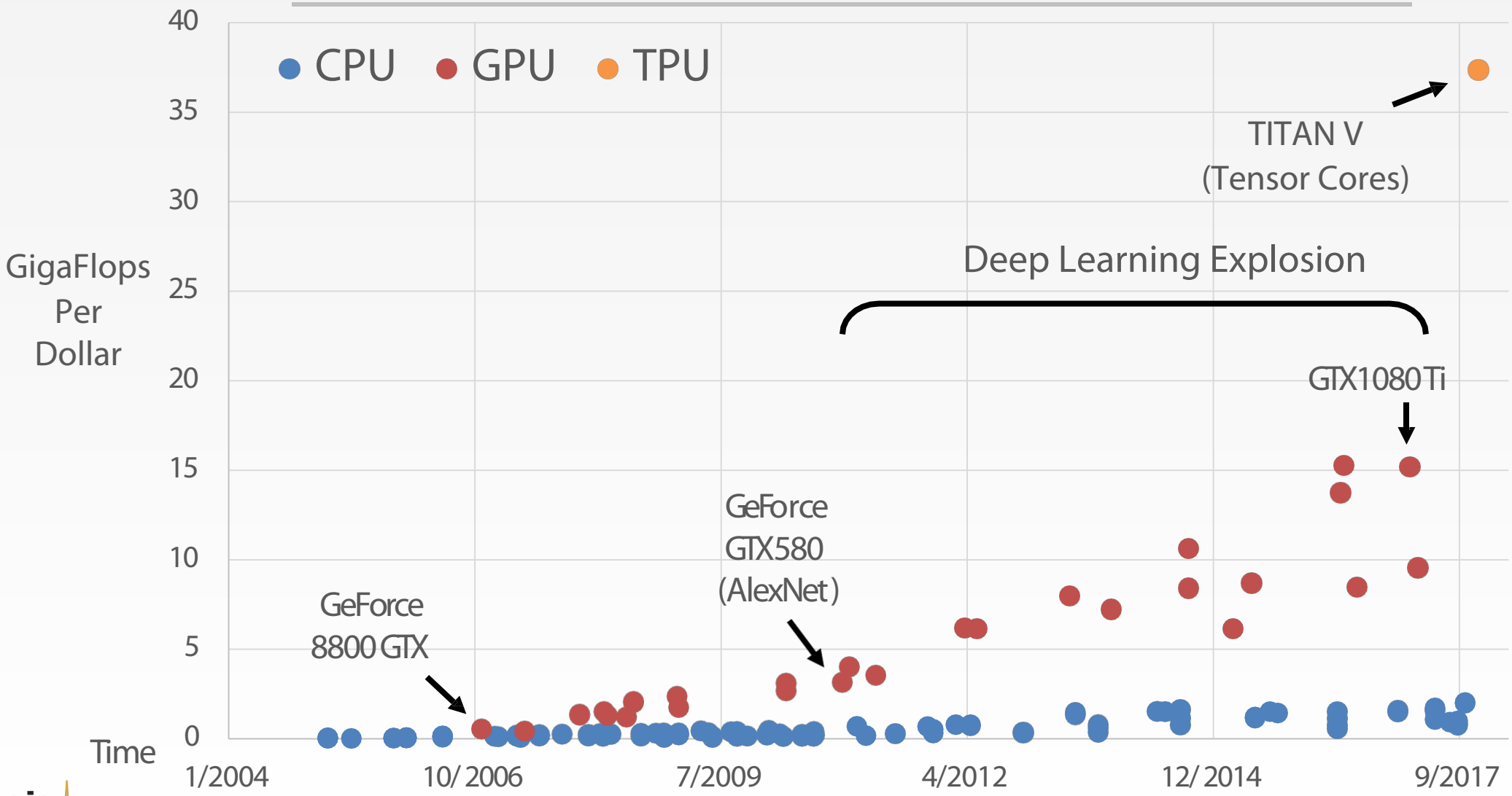
10^{14} IMAGENET

INGREDIENTS FOR DEEP LEARNING





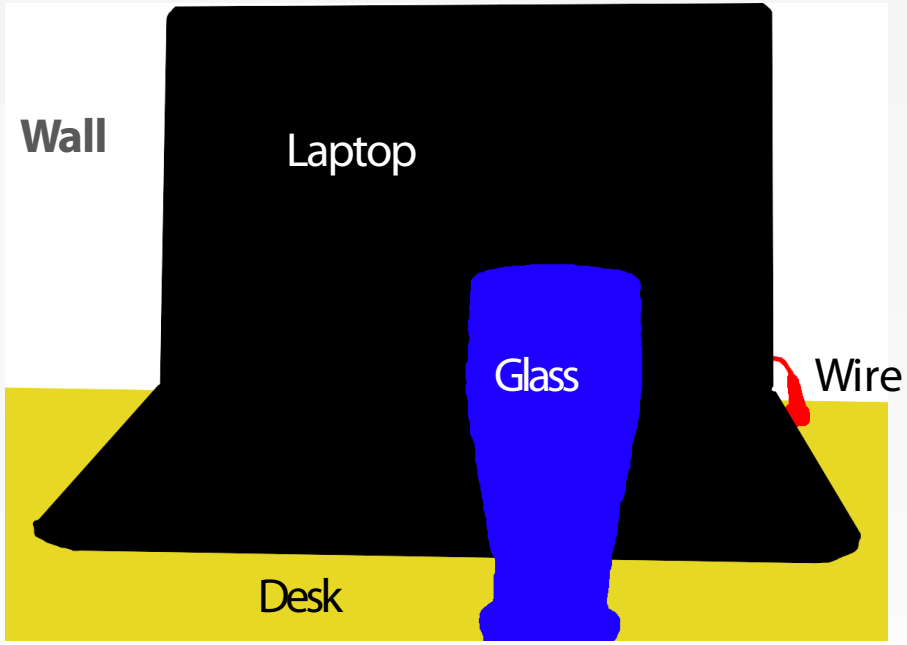
GIGAFLOPS PER DOLLAR



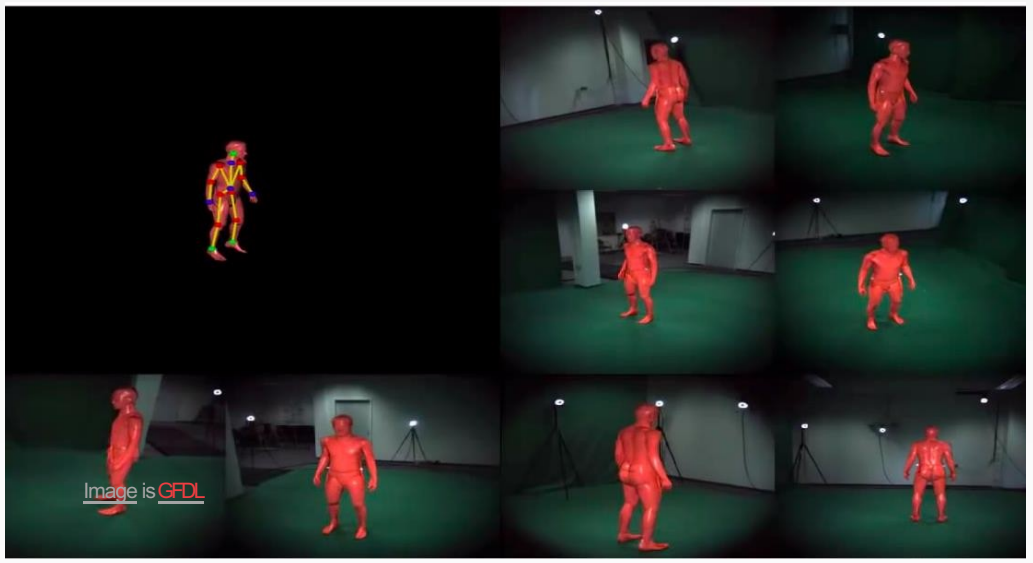
QUEST FOR VISUAL INTELLIGENCE

- The quest for visual intelligence goes far beyond object recognition

QUEST FOR VISUAL INTELLIGENCE

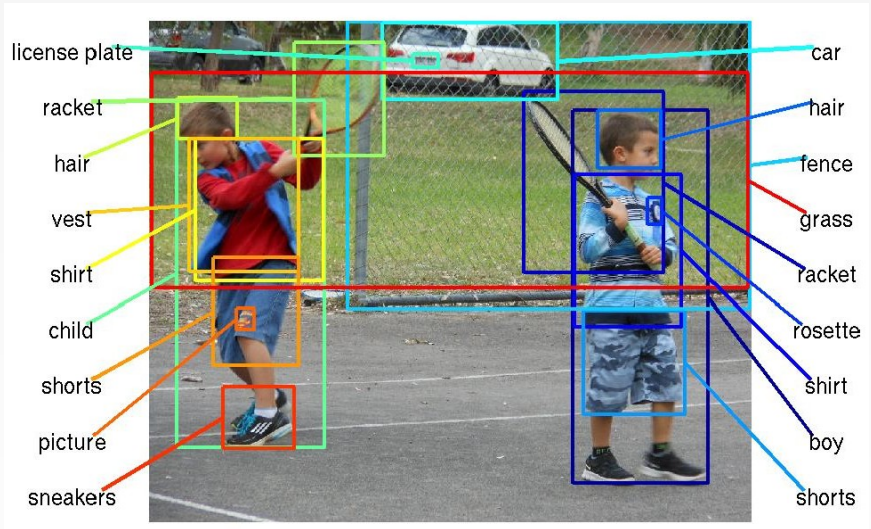


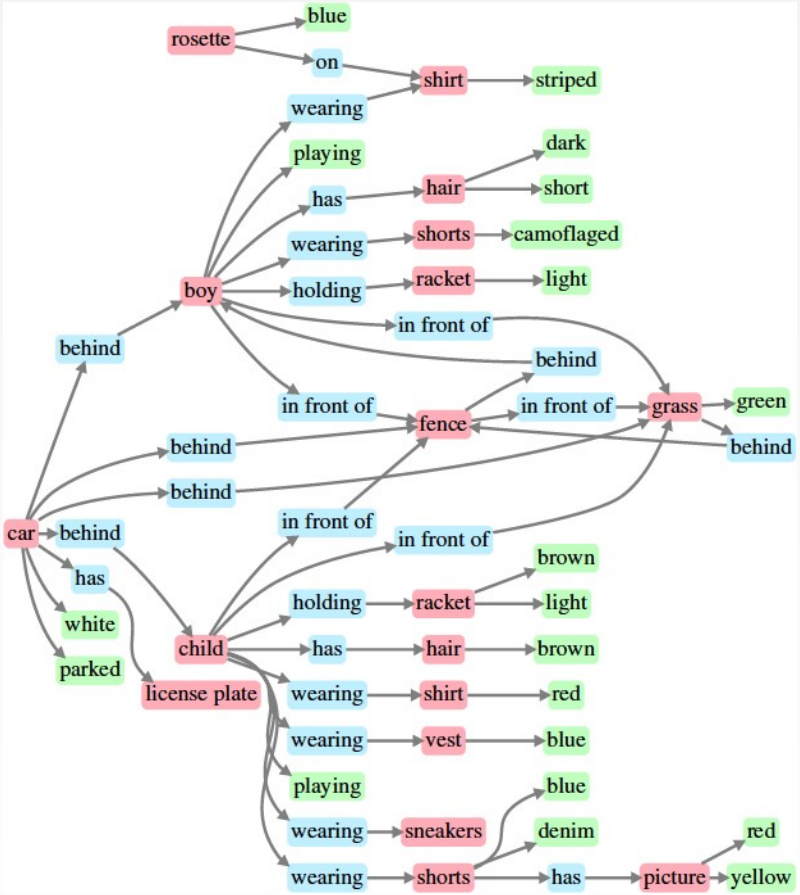
SEMANTIC SEGMENTATION



VIRTUAL REALITY

QUEST FOR VISUAL INTELLIGENCE





SCENE GRAPHS

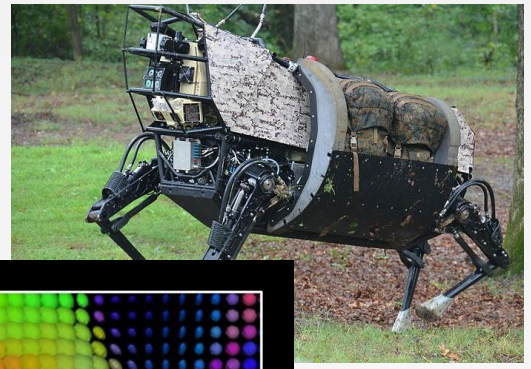
QUEST FOR VISUAL INTELLIGENCE

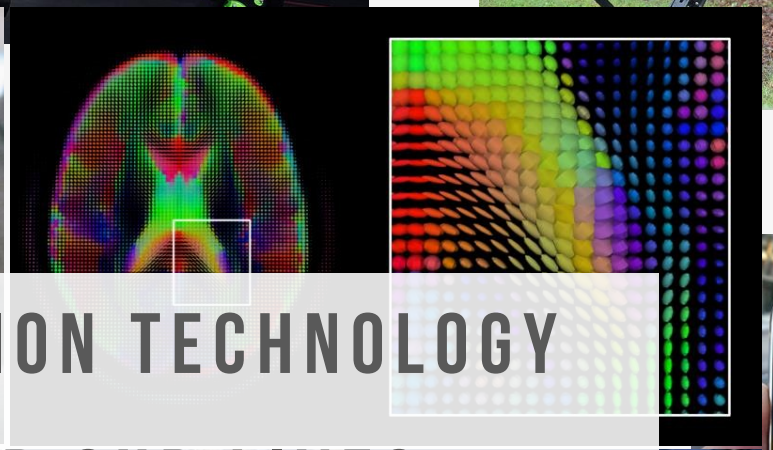
PT= 500ms

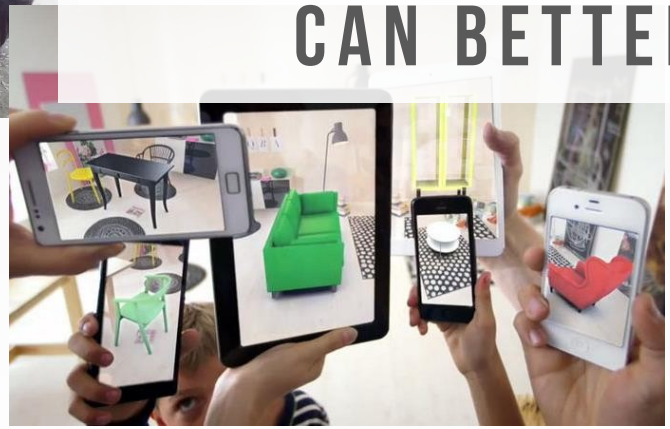
Some kind of game or fight. Twogroups of two men?Theman on the left is throwing something. Outdoors seemed like because i have an impression of grass and maybe lines on the grass?That would be why I think perhaps a game, rough game though, more like rugby than football because they pairs weren't in pads and helmets, though I did get the impression of similar clothing. maybe sometrees? in the background. (Subject: SM)

QUEST FOR VISUAL INTELLIGENCE





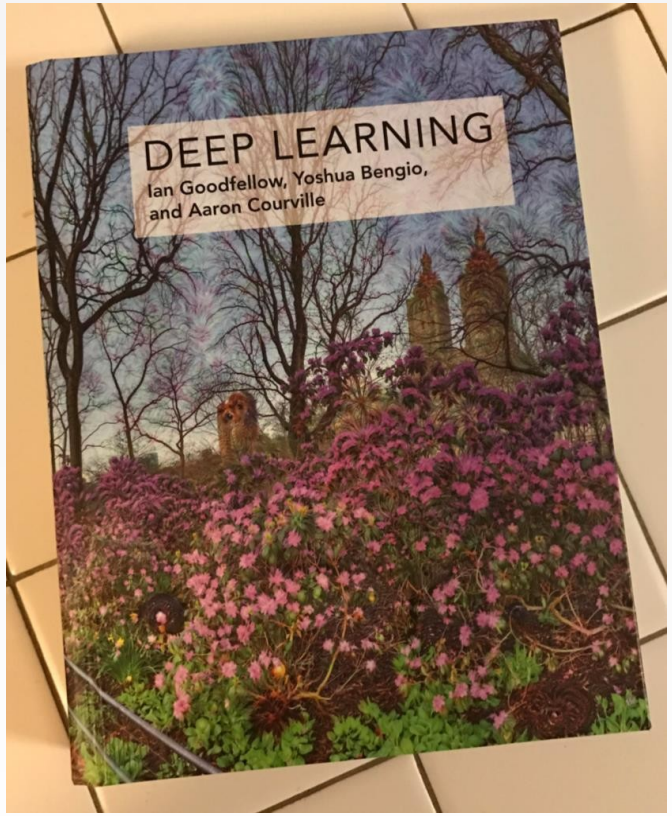



COMPUTER VISION TECHNOLOGY CAN BETTER OUR LIVES

OPTIONAL TEXTBOOK

- Deep Learning
  - Ian Goodfellow et. al.
  - Free online



COURSE PHILOSOPHY

- Thorough and detailed
  - Understand how to develop, train, and debug convolutional neural networks from scratch.
- Practical
  - Focus on practical techniques for training these networks at scale, and on GPUs. Cover deep learning frameworks.
- State of the art
  - Most materials are new from research world.

COURSE PHILOSOPHY

- Fun
  - We will cover some fun topics
  - Image Captioning (using RNN), NeuralStyle, etc.

NEXT CLASS: IMAGE CLASSIFICATION

K-NEAREST NEIGHBOURS

LINEAR CLASSIFIER