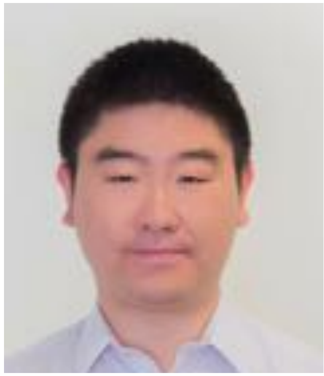
Lecture 1:
Introduction to Machine Learning

PEOPLE

Who Are We?



Prof. Liang Zheng
Instructor
Weeks 1-7

Prof. Wei Lu
Instructor
Weeks 8-14

zheng_liang@sutd.edu.sg

wei_lu@sutd.edu.sg

Teaching Assistants

Ngo Van Mao
Fri AM

Thilini Cooray
Thu PM

WHO ARE YOU?

class

Evaluation

- Homework (30%)
  - Programming and theory
  - Honor Code
    - Form study groups to work on the homework
    - You can discuss with other classmates as well
    - Write-up solutions on your own
    - List names of anyone you talked to
- Project (20%)
- Midterm Exam (25%)
- Final Exam (25%)

Course Goals

1. Curious to discover more
2. Confident of doing it yourself
3. Contemplative of the theory
4. Cautious of the dangers

Acknowledgement

- MIT 6.036 Introduction to Machine Learning
- SUTD 50.007 Machine Learning (Alex Binder)
- Stanford CS229 Machine Learning
- Stanford CS231n Convolutional Neural Network

Machine Learning

What is machine learning?

Hard-Coded

Trained

Giving computers the ability to learn
without being explicitly programmed
– Arthur Samuel (1959)

What is machine learning?

Task

Performance

Experience

Algorithms that improve their performance
at some task with experience
– Tom Mitchell (1998)

What is machine learning?

- A branch of **artificial intelligence**, concerned with the design and development of algorithms that allow computers to evolve behaviors based on empirical data.

- As intelligence requires knowledge, it is necessary for the computers to acquire knowledge.

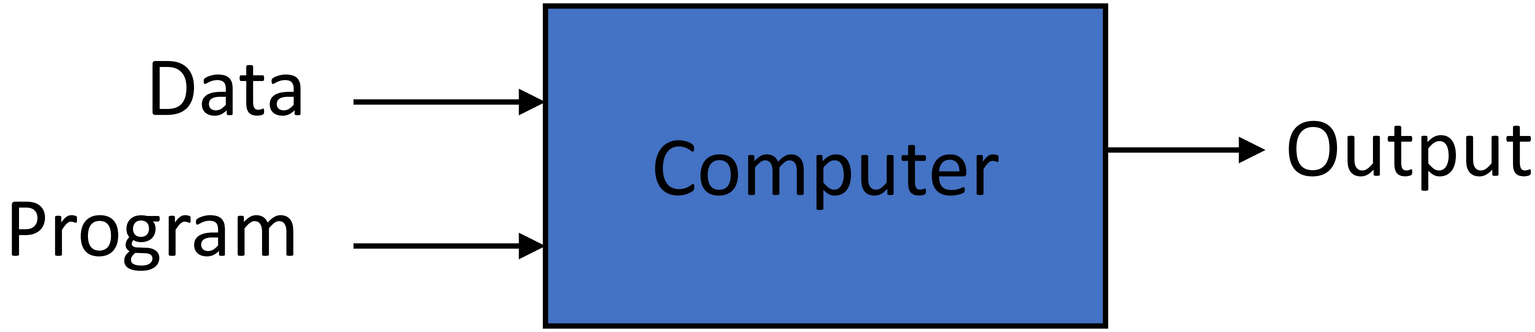
What is machine learning?

- Machine learning is programming computers to optimize a performance criterion using example data or past experience.
- There is no need to “learn” to calculate payroll
- Learning is used when:
  - Human expertise does not exist (navigating on Mars),
  - Human are unable to find the underlying insight from large volumes of data (image classification)
  - Solution needs to be adapted to particular cases (unsupervised domain adaptation)

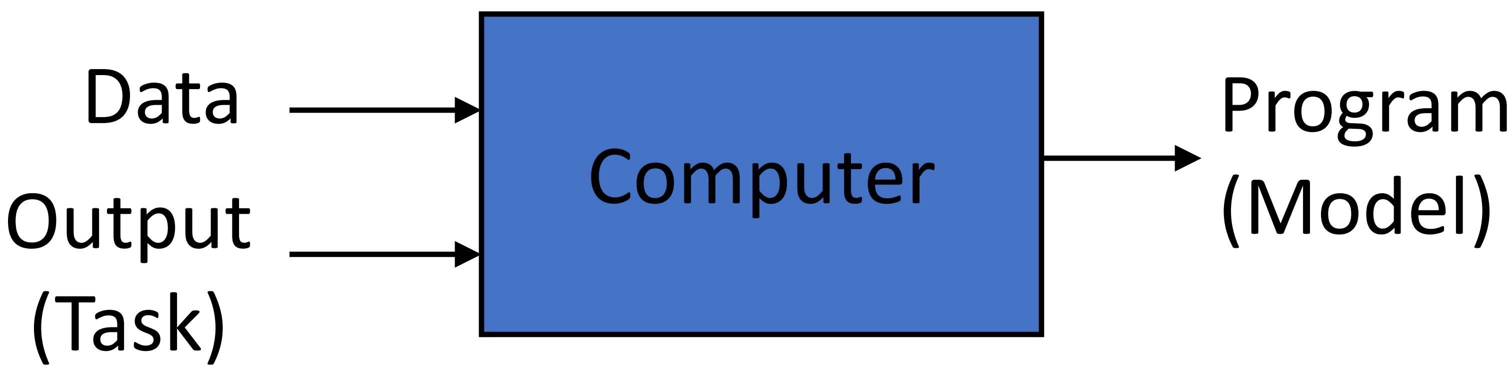
What is machine learning?

- Machine Learning
    - Study of algorithms that
    - improve their performance
    - at some task
    - with experience
- Optimize a performance criterion using example data or past experience.
- Role of Statistics: Inference from a sample
- Role of Computer science: Efficient algorithms to
    - Solve the optimization problem
    - Representing and evaluating the model for inference

Traditional Programming



Machine Learning



What is machine learning?

- We we have a model
- We predict
  - Given input
- Image classification
- Face recognition

input

model →

Dog
Building
Cat ✓
Human
Car

What we talk about when we talk about "learning"

- Learning general models from a data of particular examples
- Data is cheap (?) and abundant (data warehouses, data mars); knowledge is expensive and scarce.
- Example in retail: Customer transactions to consumer behavior:

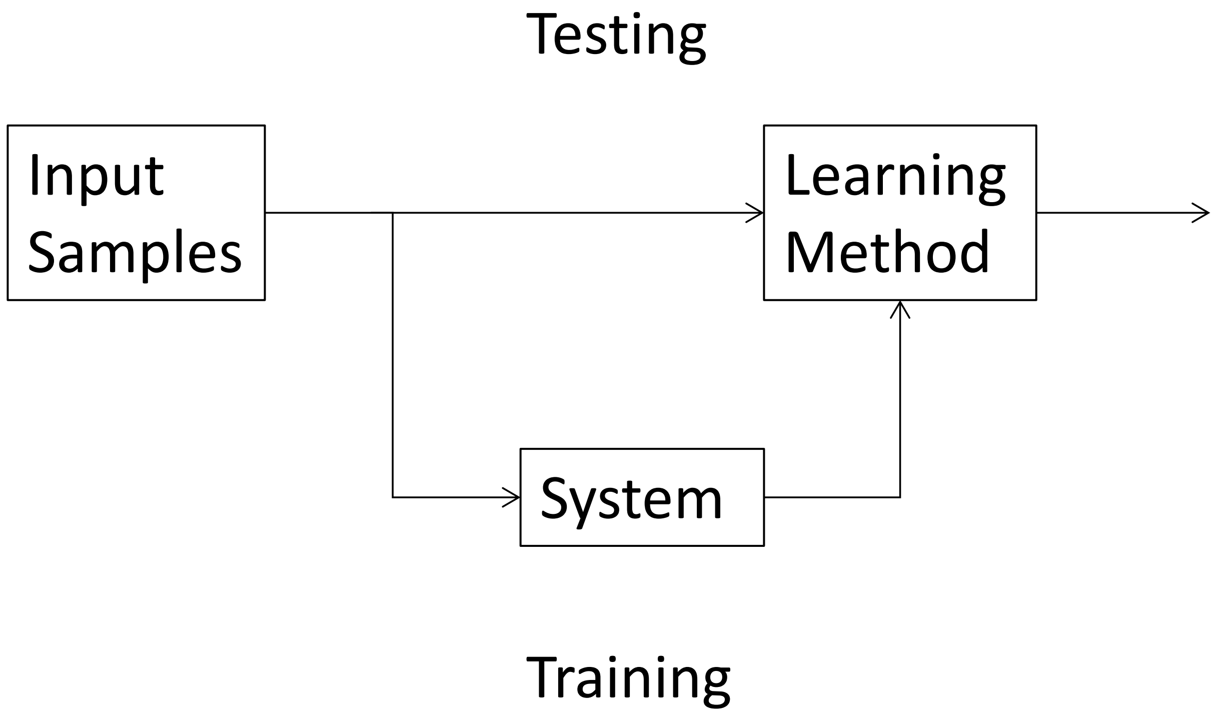
  People who bought "Da Vinci Code" also bought "The Five People You Meet in Heaven" (www.amazon.com)

- Build a model that is *a good and useful approximation* to the data.

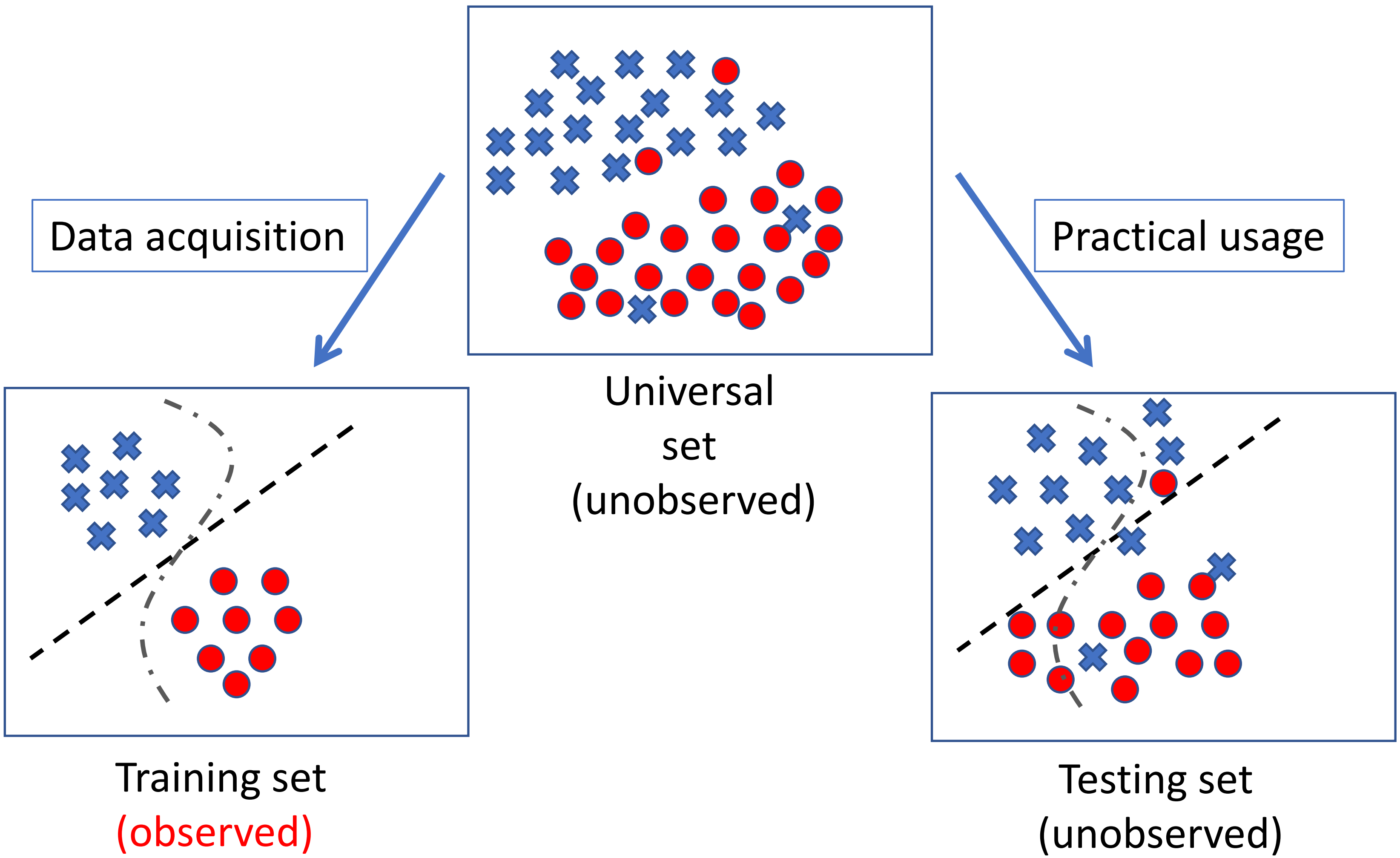
Growth of Machine Learning

- Machine learning is preferred approach to
  - Speech recognition, Natural language processing
  - Computer vision
  - Medical outcomes analysis
  - Robot control
  - Computational biology
- This trend is accelerating
  - Improved machine learning algorithms
  - Improved data capture, networking, faster computers
  - Software too complex to write by hand
  - New sensors / IO devices
  - Demand for self-customization to user, environment
  - It turns out to be difficult to extract knowledge from human experts → *failure of expert systems in the 1980's.*

Learning system model



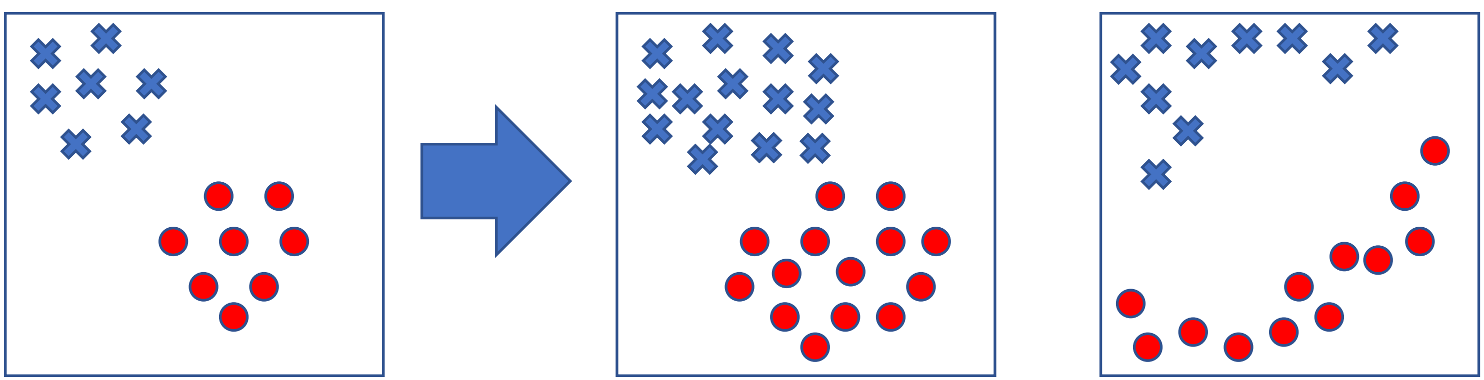
Training and testing



Training and testing

- Training is the process of making the system able to learn.

- No free lunch rule:
  - Training set and testing set come from the same distribution
  - Need to make some assumptions or bias



Training data

Testing data

Performance

- There are several factors affecting the performance:
  - **Quality of training data** provided
  - The form and extent of any initial **background knowledge**
  - The **type of feedback** provided
  - The **learning algorithms** used

- Two important factors:
  - Modeling
  - Optimization

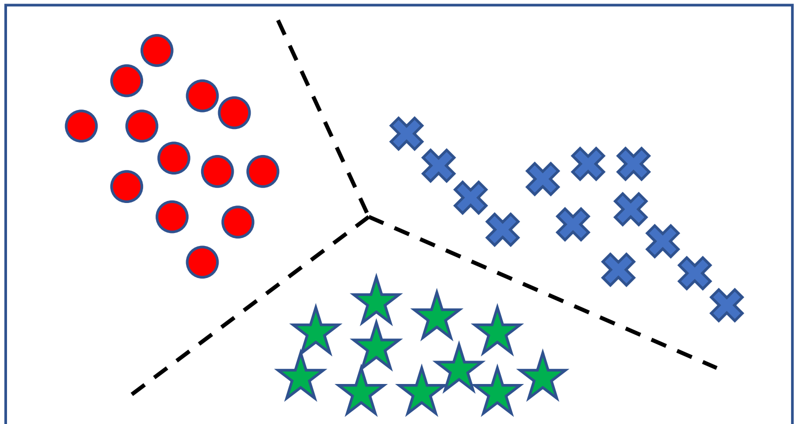
Algorithms

- The success of machine learning system also depends on the algorithms.
- The algorithms control the search to find and build the knowledge structures.
- The learning algorithms should extract useful information from training examples.

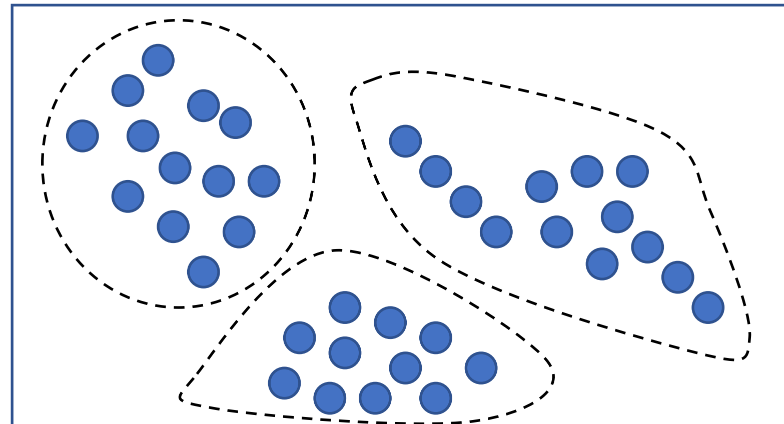
Types of machine learning

- **Supervised learning** ($\{x_n \in R^d, y_n \in R\}_{n=1}^N$)
  - Prediction
  - Classification (discrete labels), Regression (real values)
- **Unsupervised learning** ($\{x_n \in R^d\}_{n=1}^N$)
  - Clustering
  - Probability distribution estimation
  - Finding association (in features)
  - Dimension reduction
- **Semi-supervised learning**
- **Reinforcement learning**
  - Decision making (robot, chess machine)

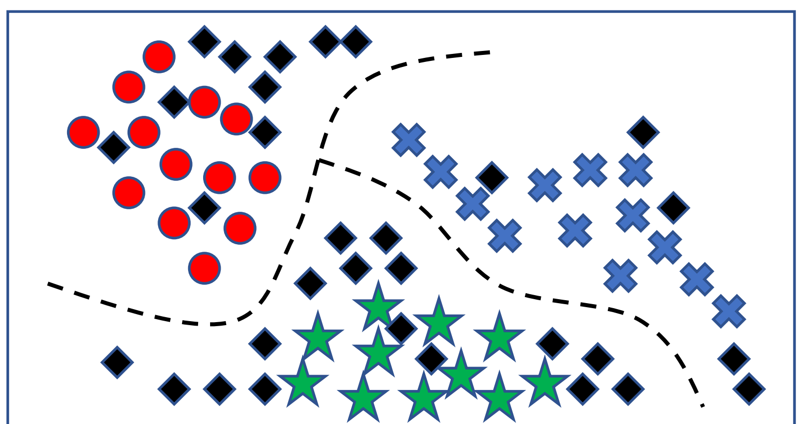
Types of machine learning



Supervised learning



Unsupervised learning



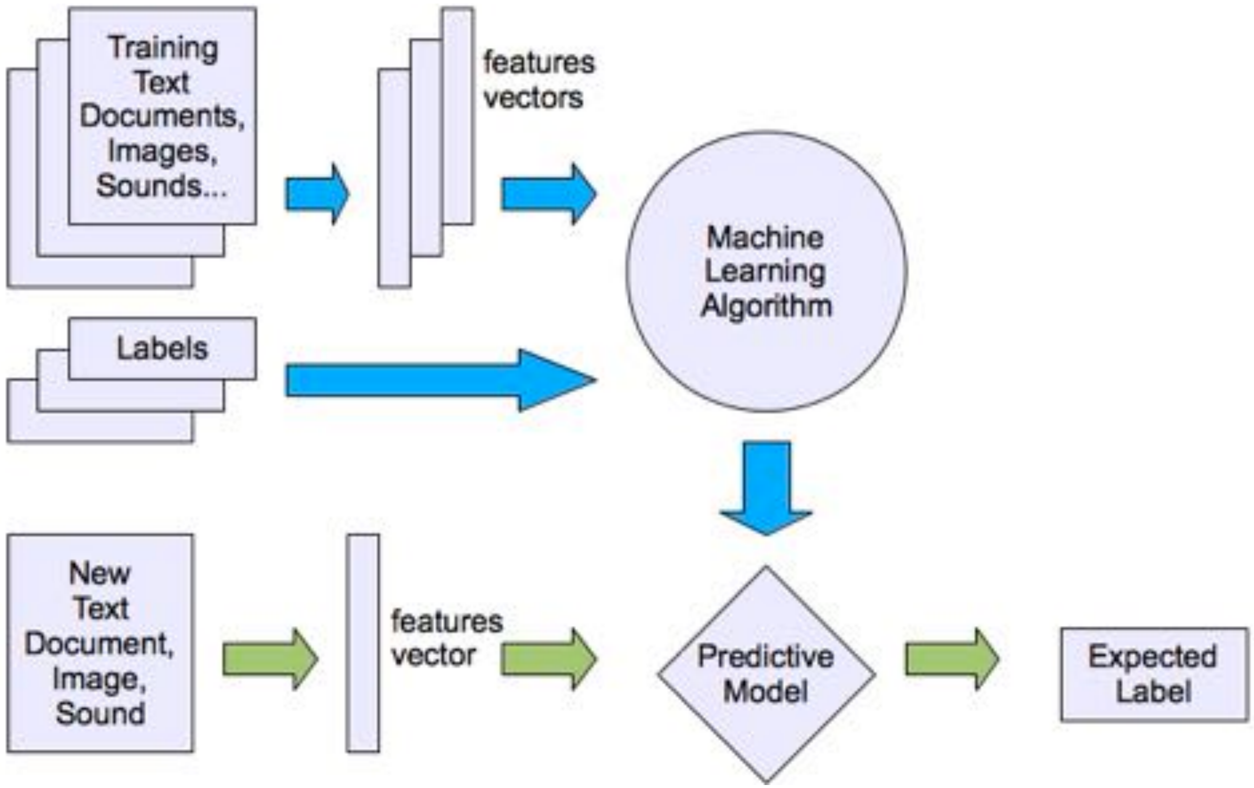
Semi-supervised learning

Types of machine learning

Supervised Learning

Types of machine learning

- Supervised learning

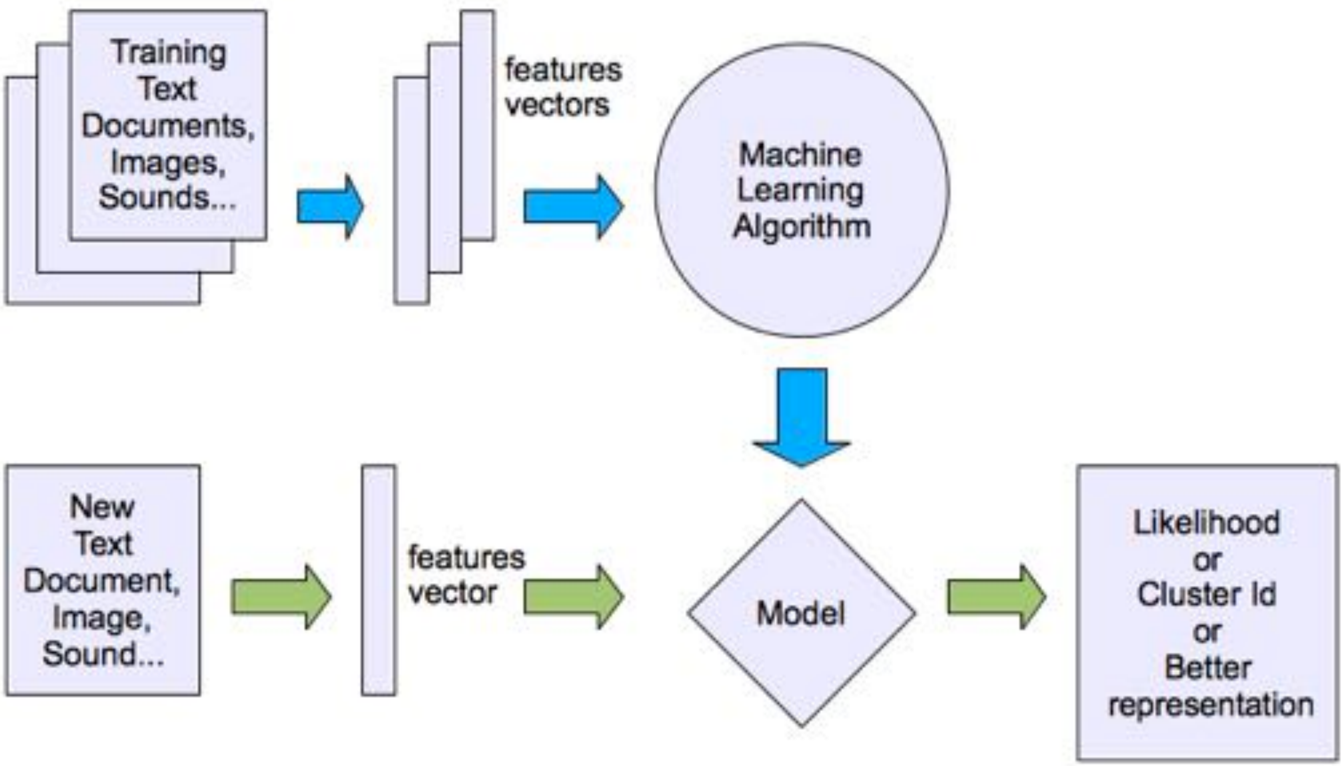


Types of machine learning

Unsupervised Learning

Types of machine learning

- Unsupervised learning

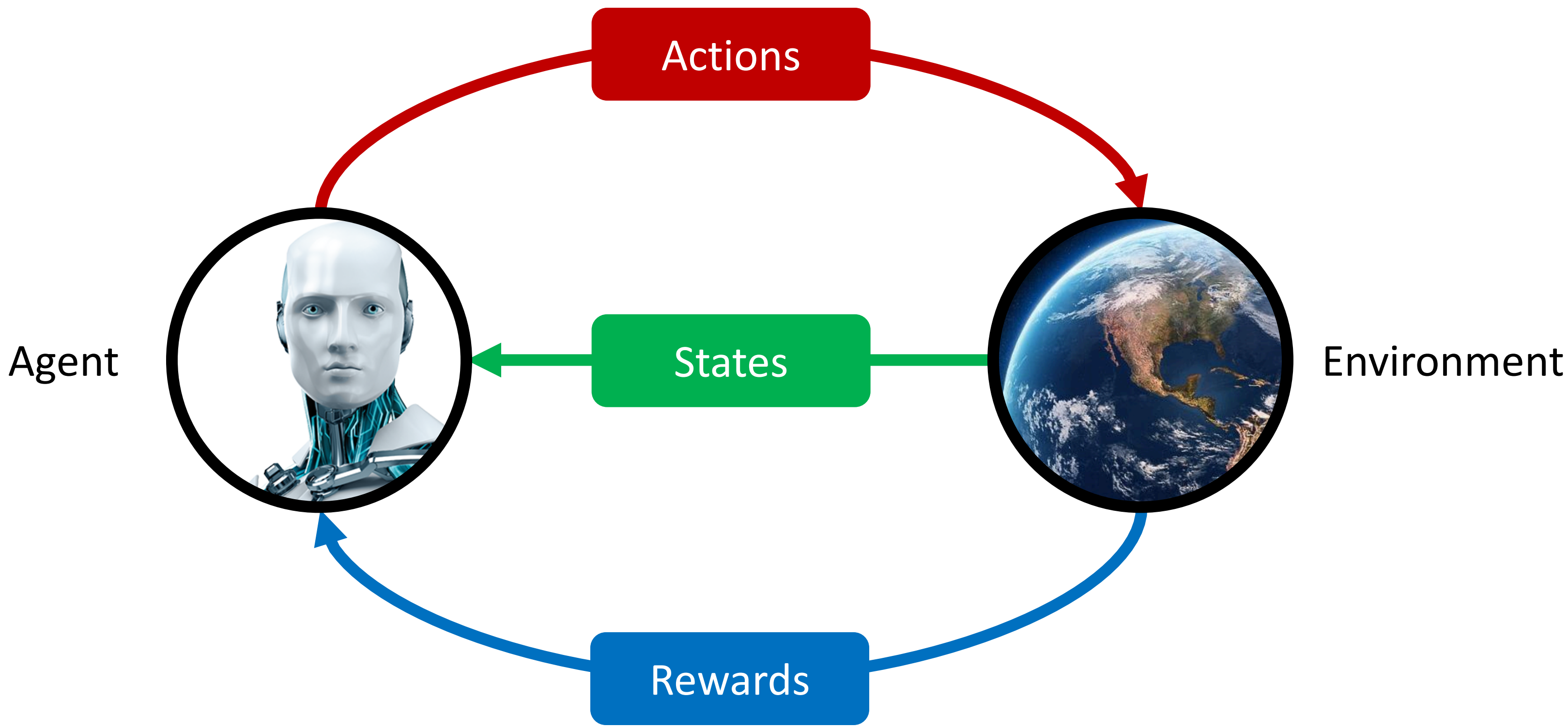


Types of machine learning

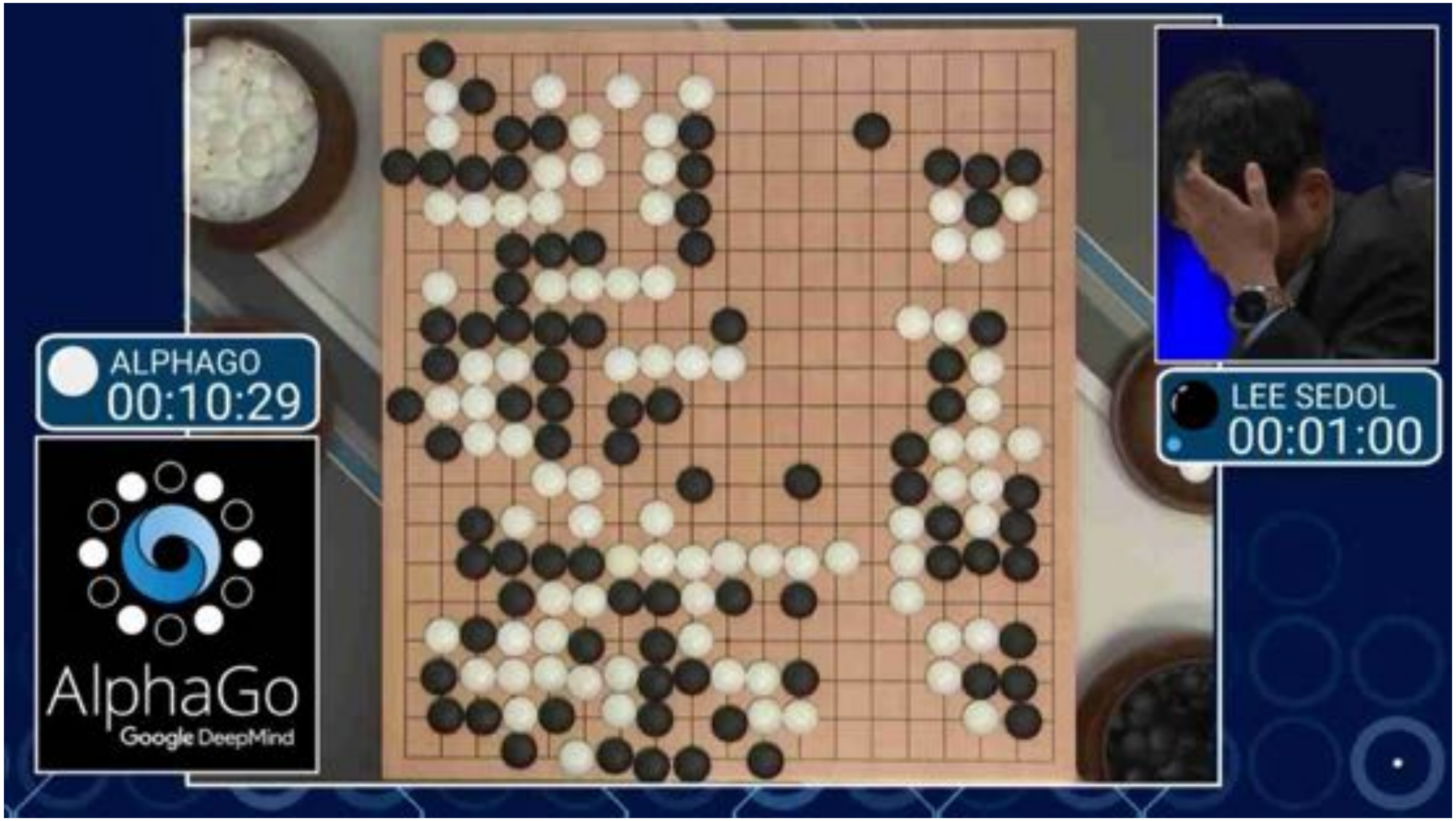
Reinforcement Learning

Rewards from a sequence of actions

Reinforcement Learning



alphago



https://www.youtube.com/watch?v=SUbqykXVx0A

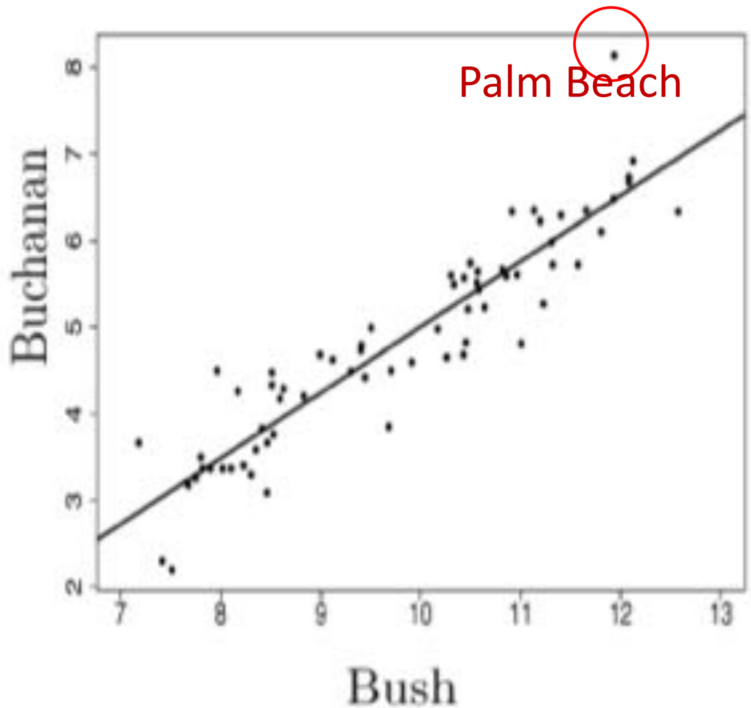
Types of machine learning

Transfer Learning

Supervised Learning

Learning a function
$y = f(x)$
$x \in \mathbb{R}$
$y \in \mathbb{R}$

Regression (Linear)



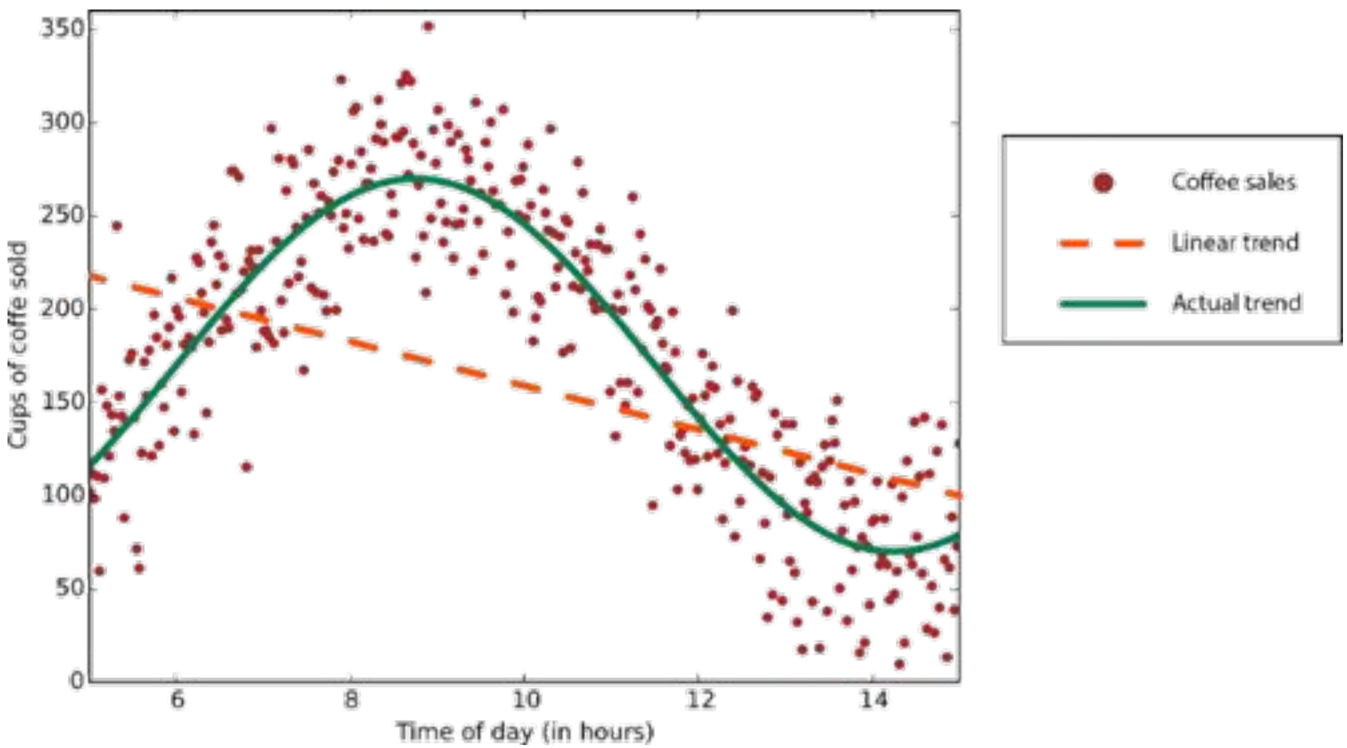
2000 USA Presidential Elections.

Votes for Buchanan and Bush in counties of Florida on a log scale.

https://en.wikipedia.org/wiki/United_States_presidential_election_in_Florida,_2000
http://andrys.com/flballot.html

Supervised Learning

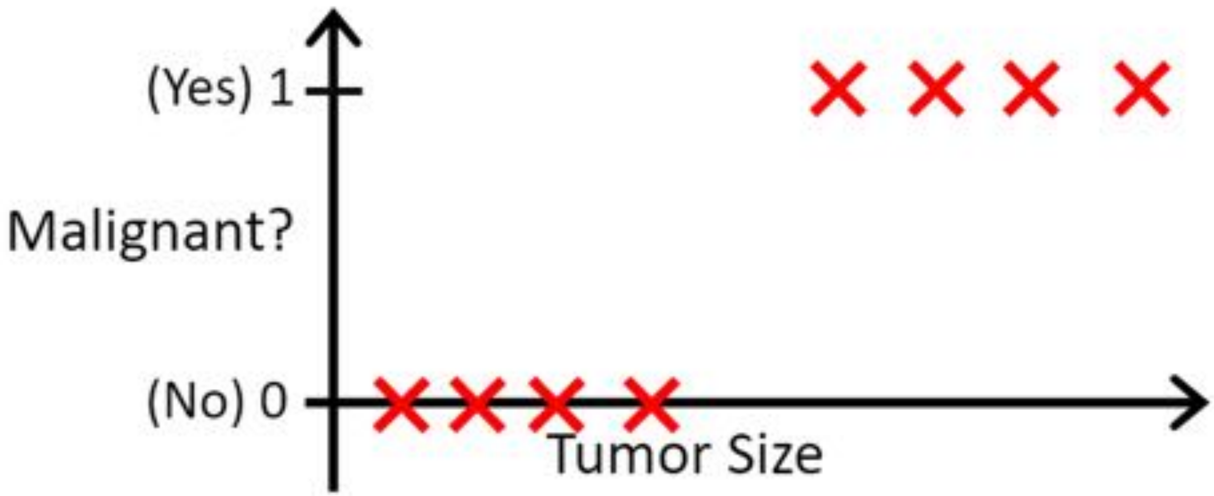
Regression (Non-linear)



Supervised Learning

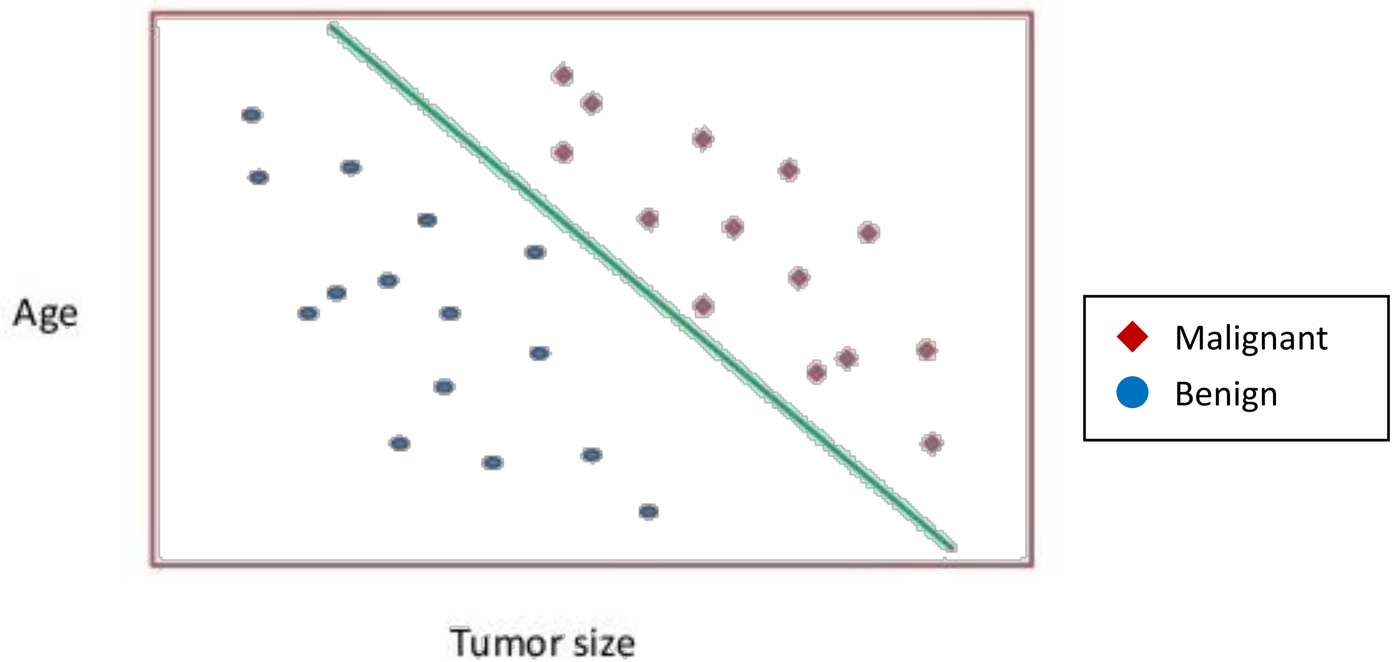
Learning a function
$y = f(x)$
$x \in \mathbb{R}$
$y \in \{1, 2, \dots, k\}$

Classification



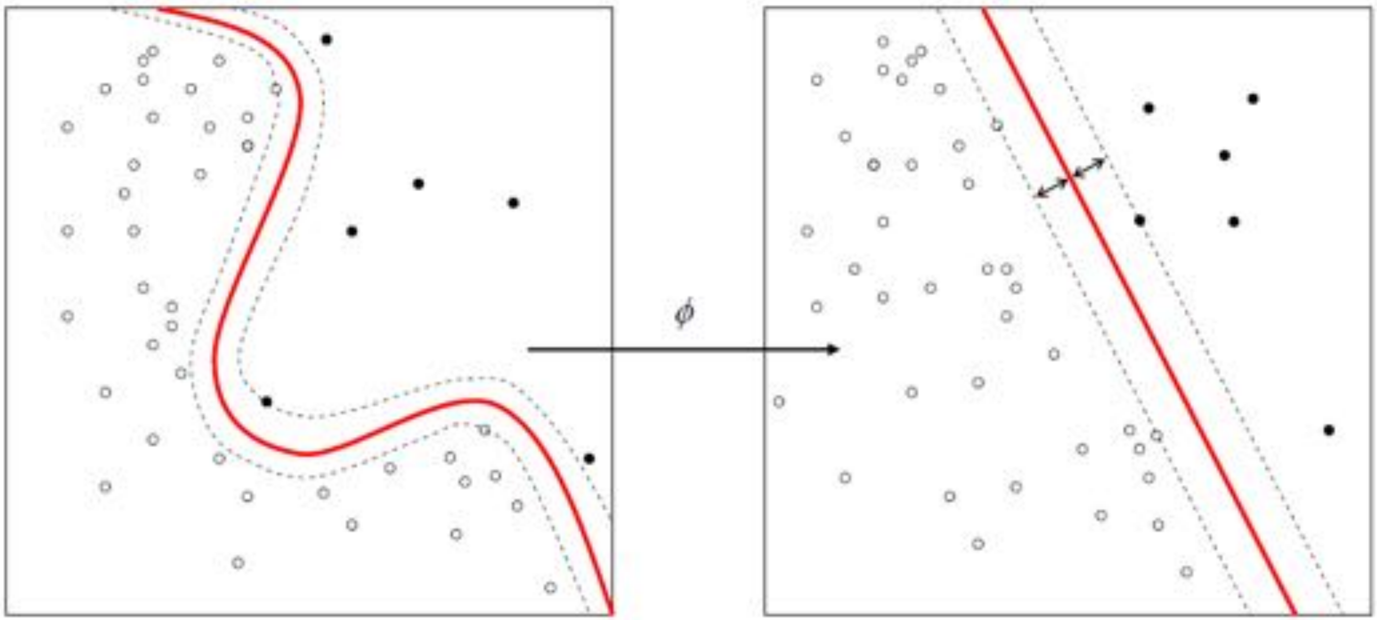
Supervised Learning

Classification (Linear)



Supervised Learning

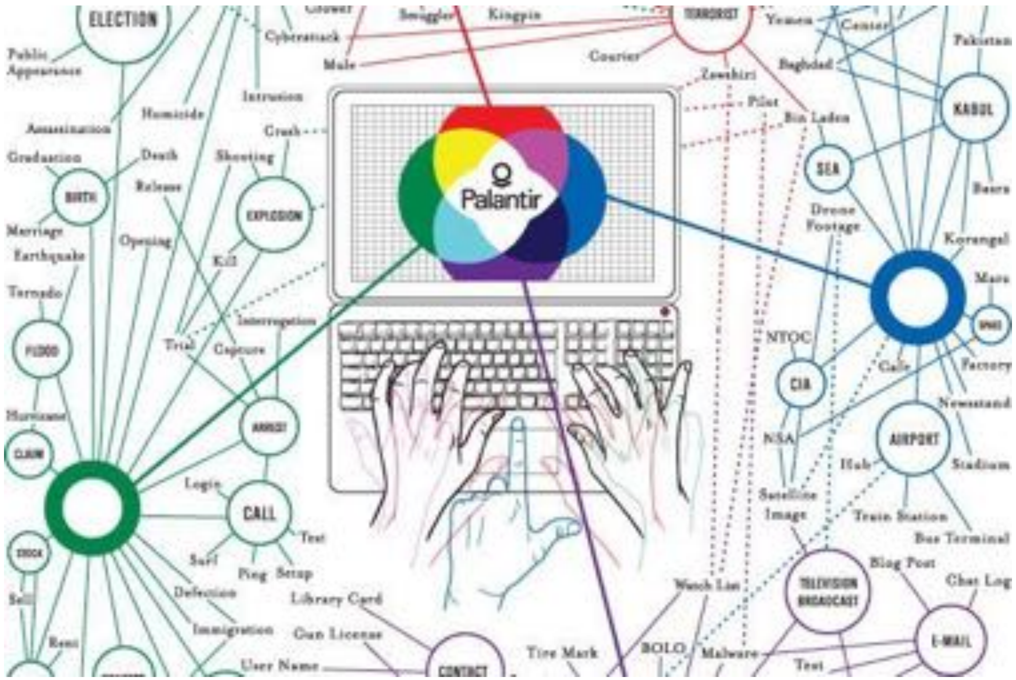
Classification (Non-linear)



Spam Filters

Bayesian Networks

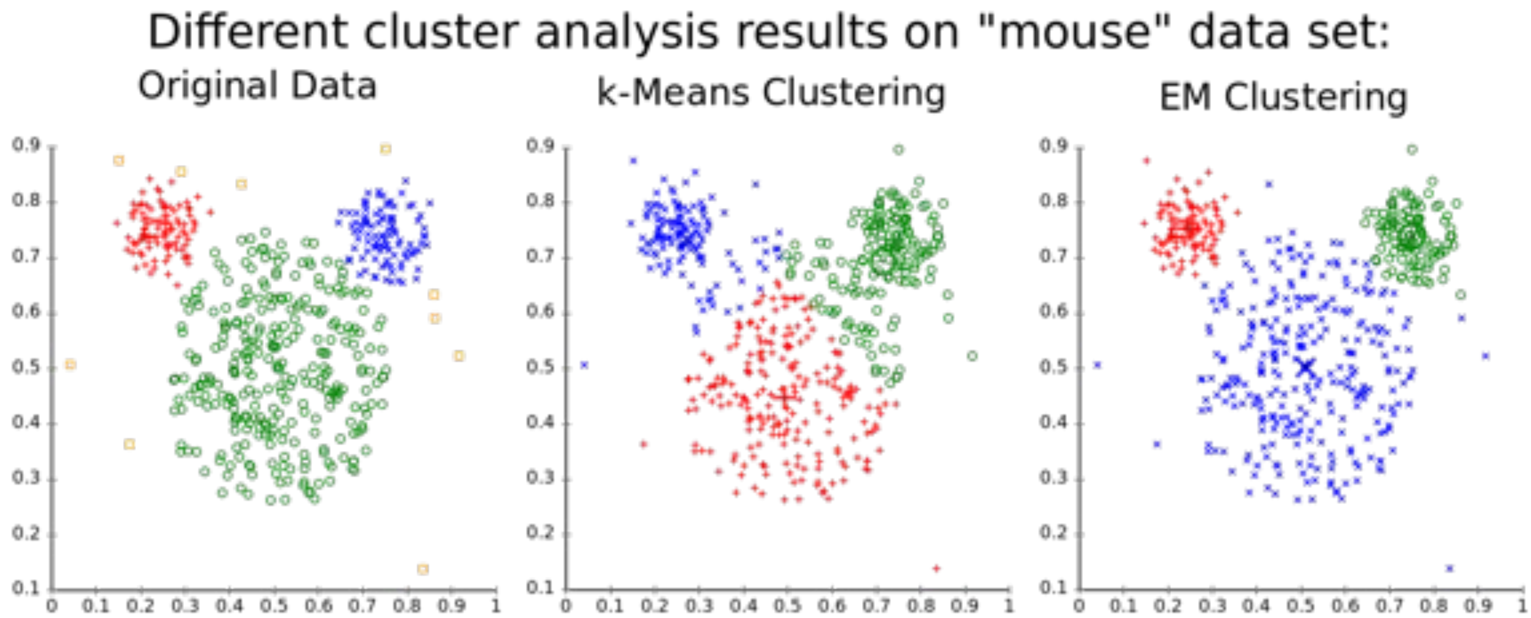
Fraud detection



https://www.washingtonian.com/2012/01/31/killer-app/

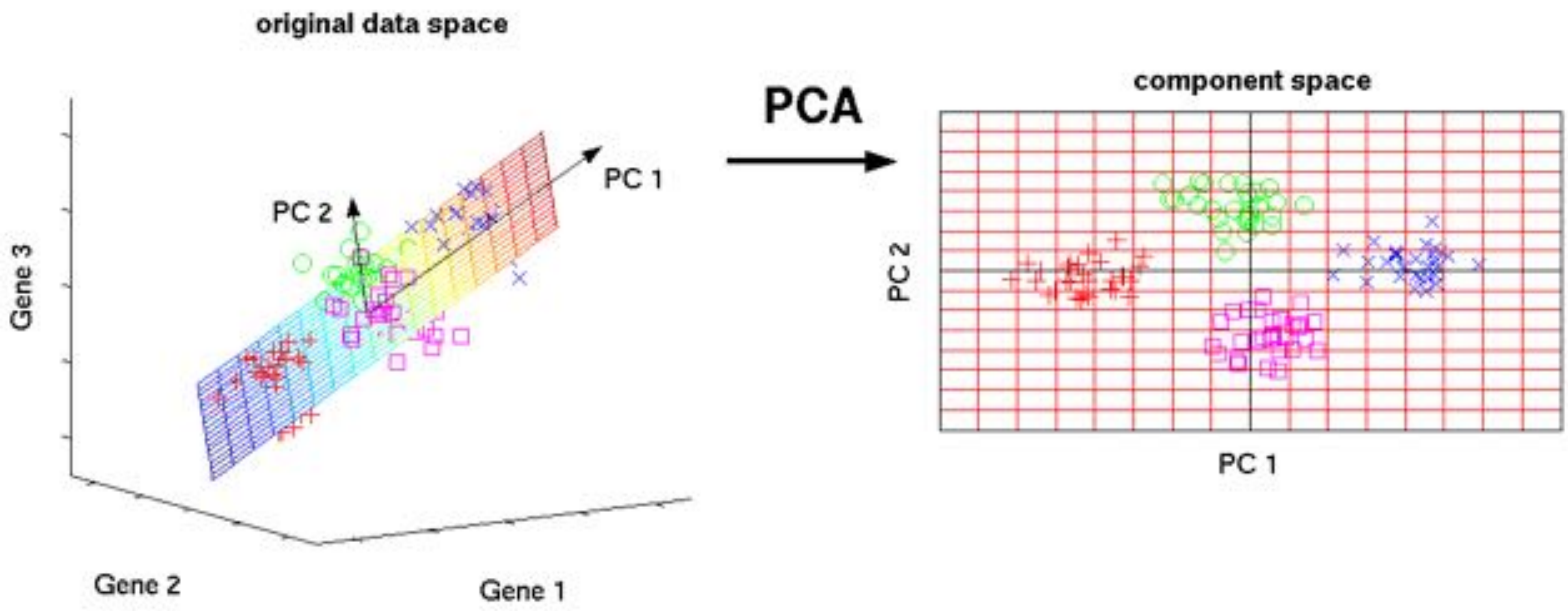
Unsupervised Learning

Clustering



Unsupervised Learning

Dimensionality Reduction: Subspace Learning



Deep Learning

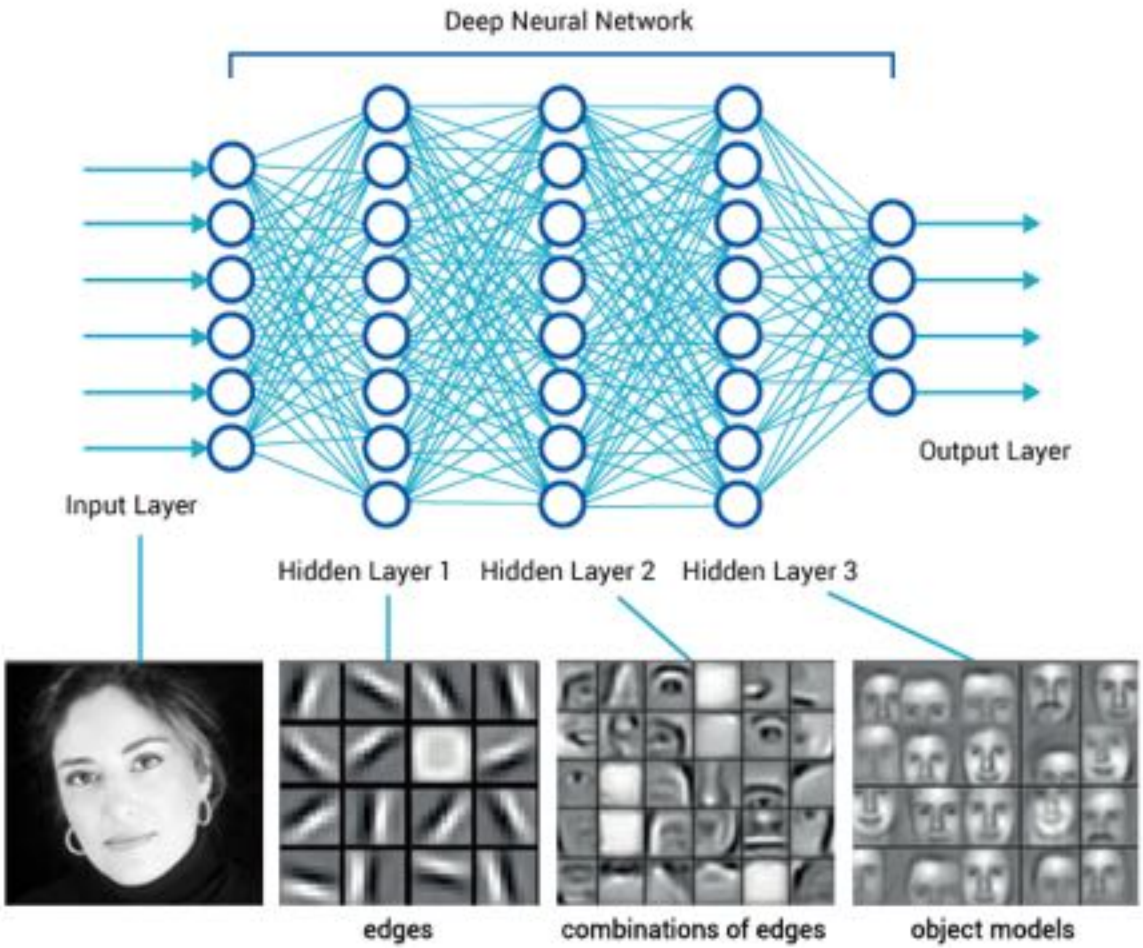


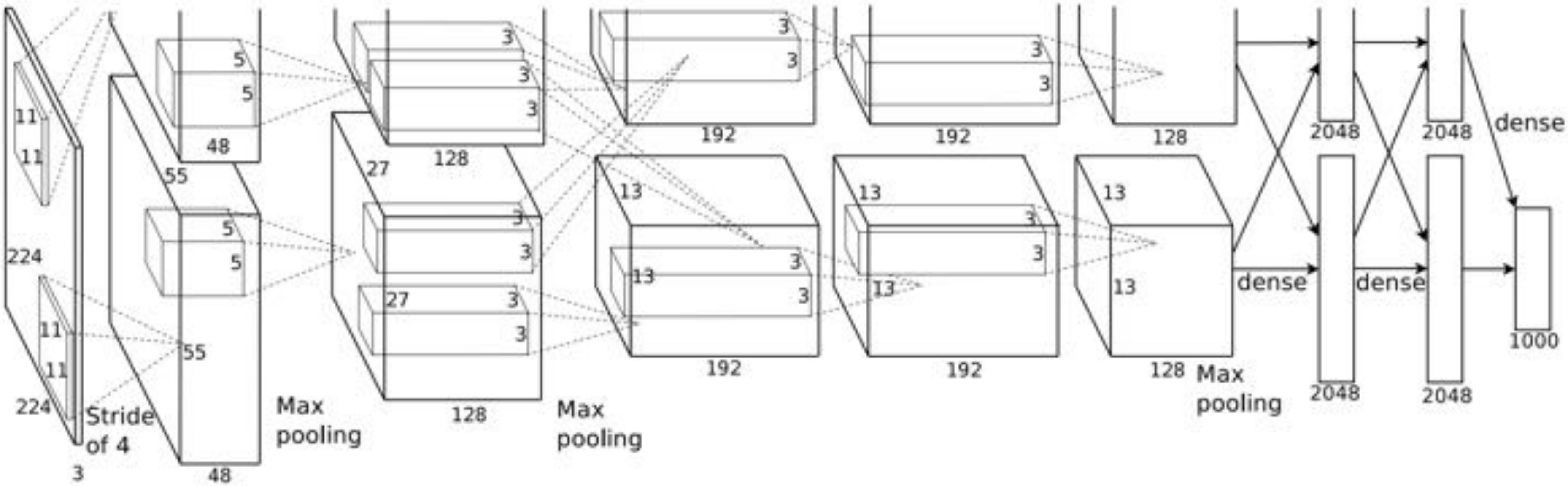
Image Classification

ImageNet dataset: 1,000 classes, 1.2 million images for training, 50k images for testing

Method	**Year**	**Top-1 error (%)**	**Top-5 error (%)**
Sparse coding	2010	47.1	28.2
SIFT + FV	2011	45.7	25.7
AlexNet	2012	37.5	17.0
VGGNet	2014	23.7	6.8
GoogleNet	2014	21.99	4.82
ResNet	2016	19.38	3.57

Human: 5.1%

Alexnet



Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." In Advances in neural information processing systems, pp. 1097-1105. 2012.

Image Classification

ImageNet dataset: 1,000 classes, 1.2 million images for training, 50k images for testing

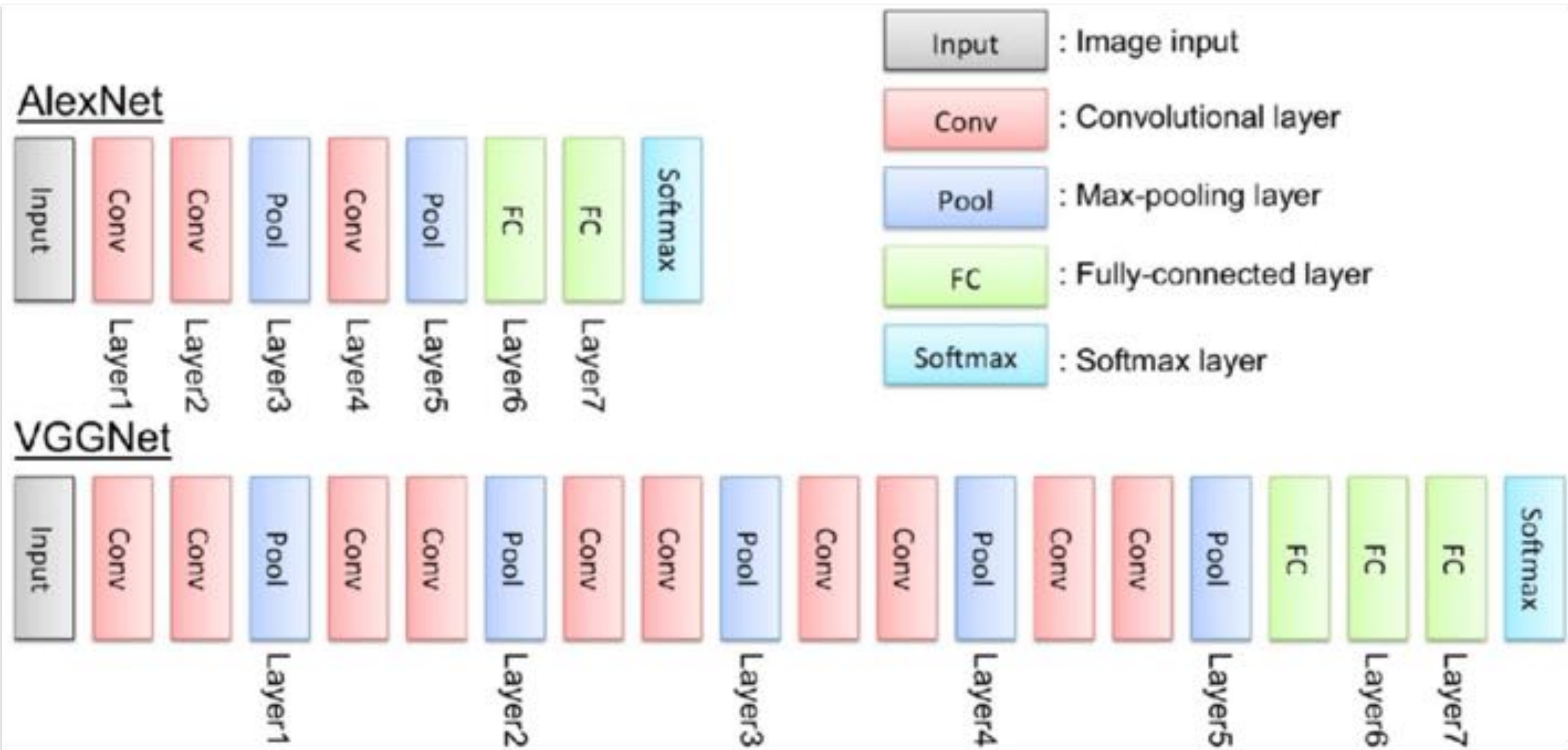
Method	**Year**	**Top-1 error (%)**	**Top-5 error (%)**
Sparse coding	2010	47.1	28.2
SIFT + FV	2011	45.7	25.7
AlexNet	2012	37.5	17.0
VGGNet	2014	23.7	6.8
GoogleNet	2014	21.99	4.82
ResNet	2016	19.38	3.57

Human: 5.1%

VGGNet

Karen Simonyan, and Andrew Zisserman. “Very deep convolutional networks for large-scale image recognition.” *ICLR 2015*.

VGGNet



Input
: Image input
Conv
: Convolutional layer
Pool
: Max-pooling layer
FC
: Fully-connected layer
Softmax
: Softmax layer
AlexNet
Input
Conv
Conv
Pool
Conv
Pool
FC
FC
Softmax
Layer1
Layer2
Layer3
Layer4
Layer5
Layer6
Layer7
VGGNet
Input
Conv
Conv
Pool
Conv
Conv
Pool
Conv
Conv
Pool
Conv
Conv
Pool
Conv
Conv
Pool
FC
FC
FC
Softmax
Layer1
Layer2
Layer3
Layer4
Layer5
Layer6
Layer7

Image Classification

ImageNet dataset: 1,000 classes, 1.2 million images for training, 50k images for testing

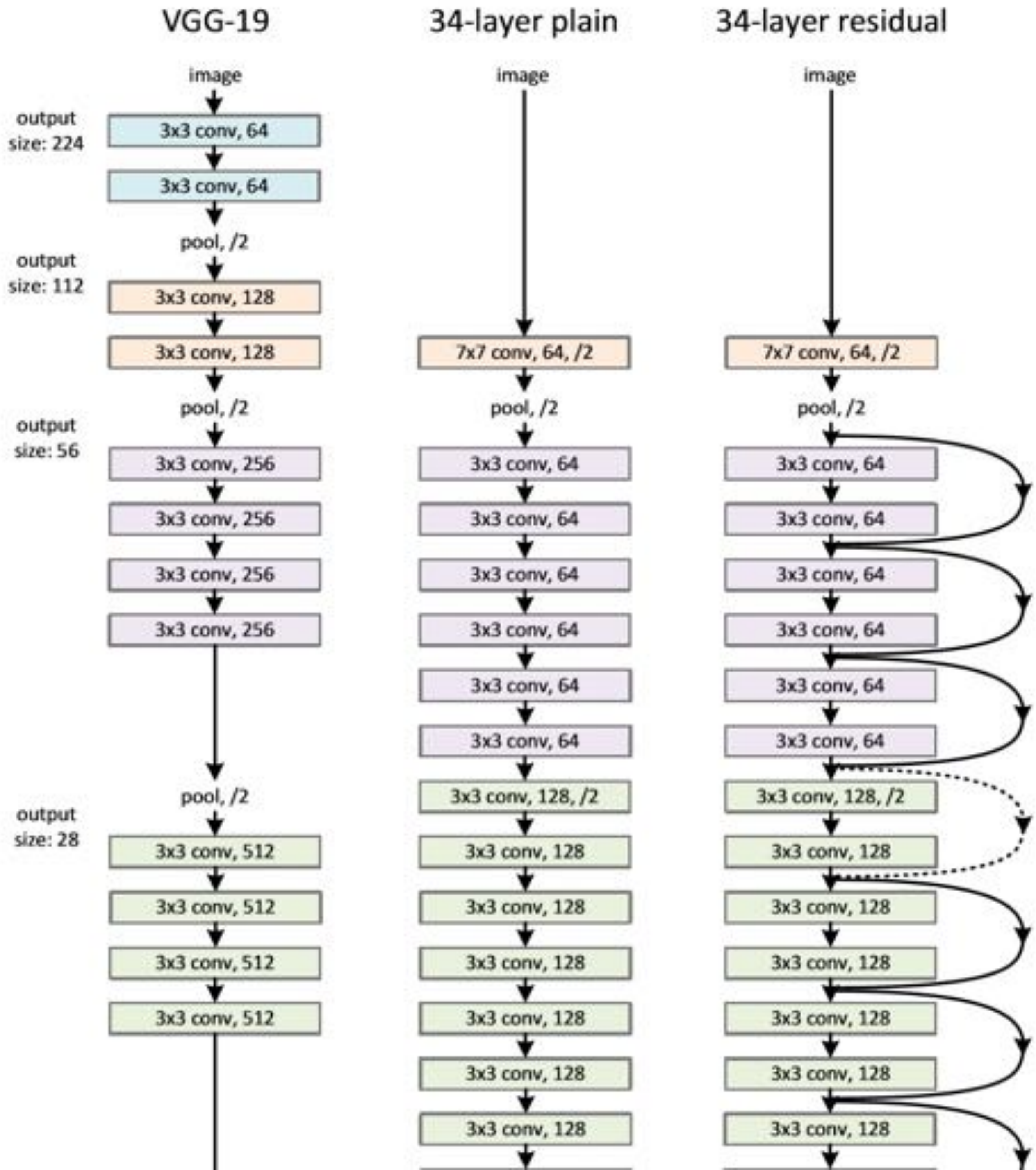
Method	**Year**	**Top-1 error (%)**	**Top-5 error (%)**
Sparse coding	2010	47.1	28.2
SIFT + FV	2011	45.7	25.7
AlexNet	2012	37.5	17.0
VGGNet	2014	23.7	6.8
GoogleNet	2014	21.99	4.82
ResNet	2016	19.38	3.57

Human: 5.1%

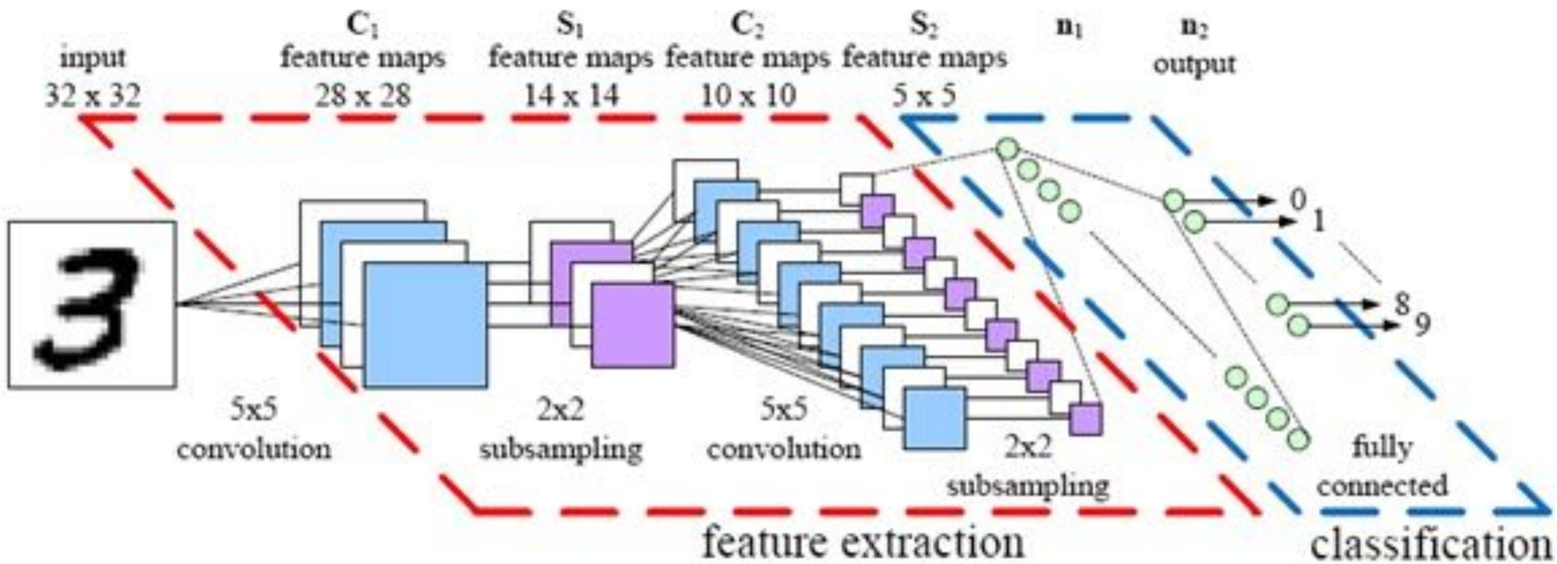
ResNet

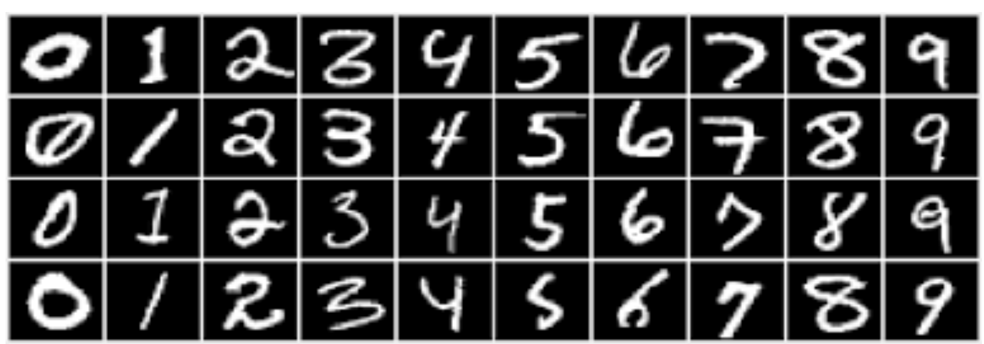
Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778. 2016.

Res

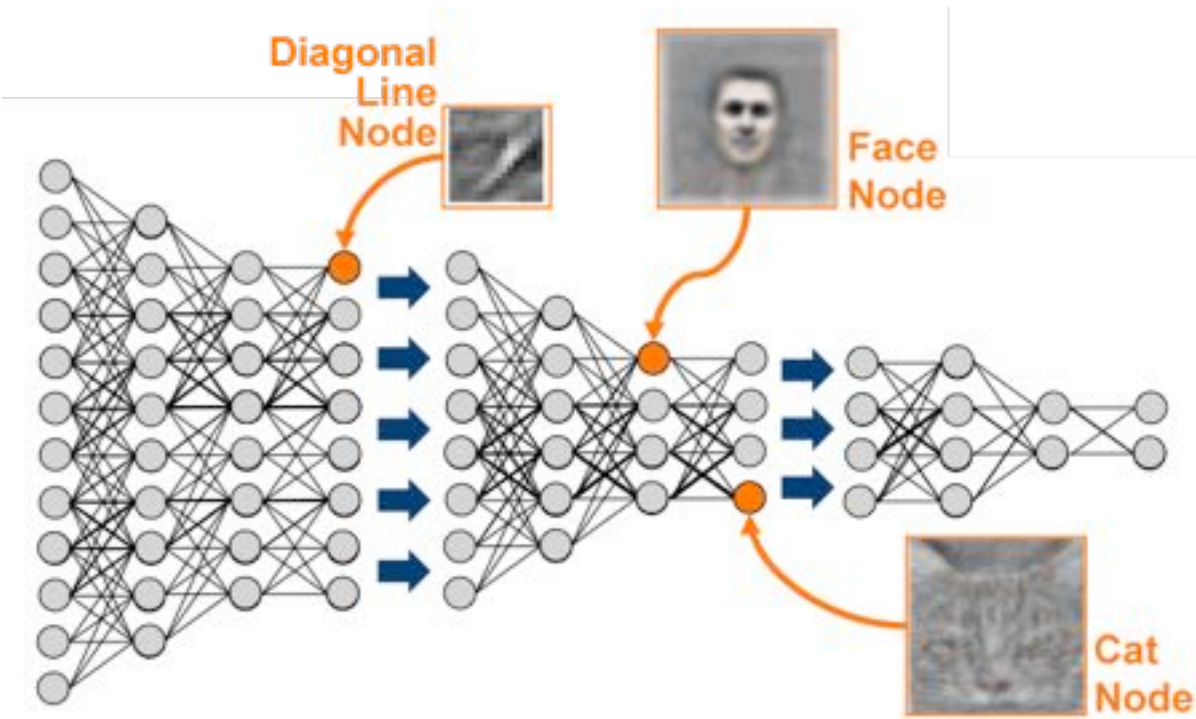

VGG-19
34-layer plain
34-layer residual
image
image
image
output
size: 224
3x3 conv, 64
3x3 conv, 64
pool, /2
output
size: 112
3x3 conv, 128
3x3 conv, 128
7x7 conv, 64, /2
7x7 conv, 64, /2
pool, /2
pool, /2
pool, /2
output
size: 56
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
pool, /2
output
size: 28
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
3x3 conv, 128, /2
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128, /2
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128

Handwriting Recognition



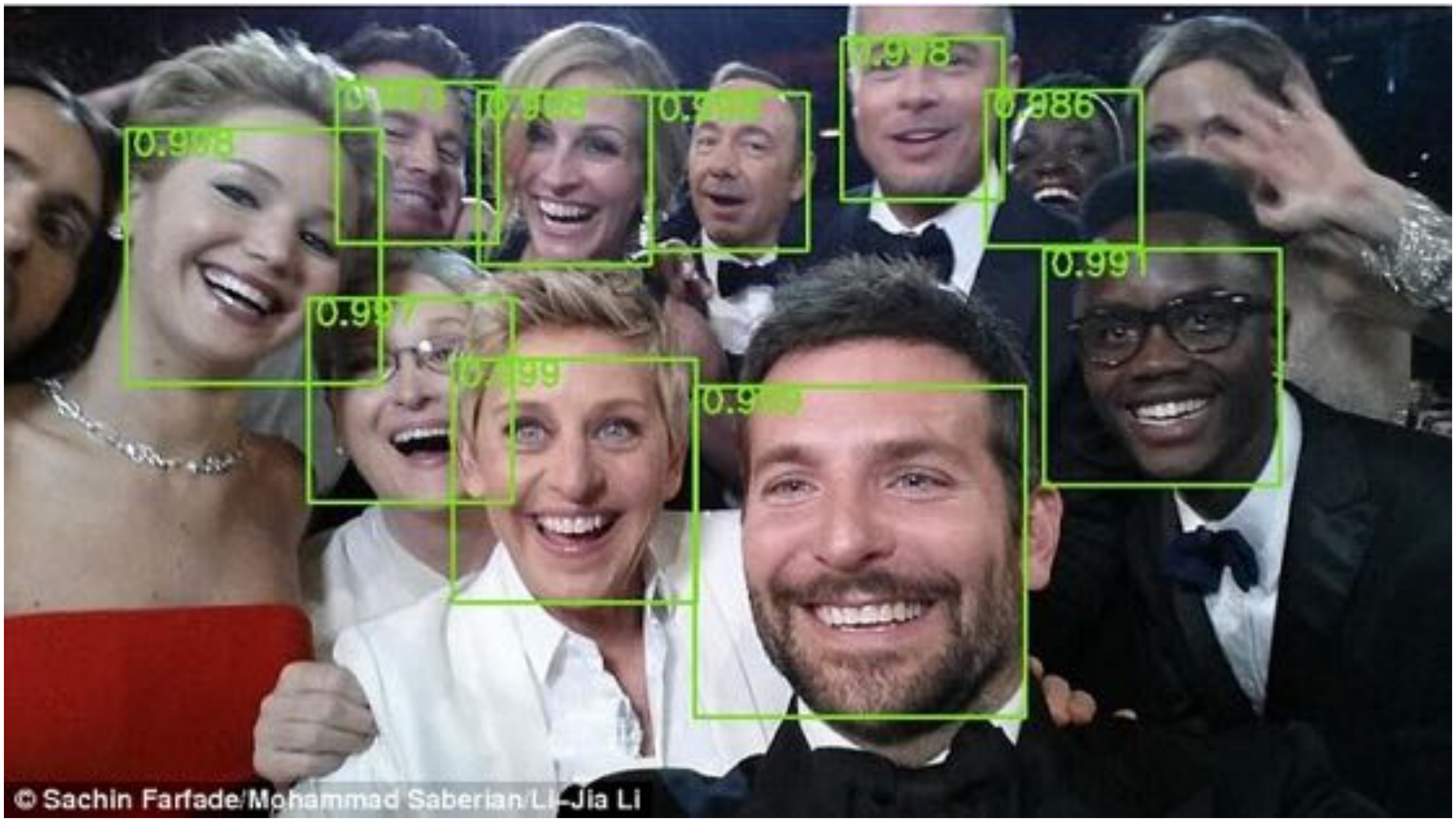


Google cat videos



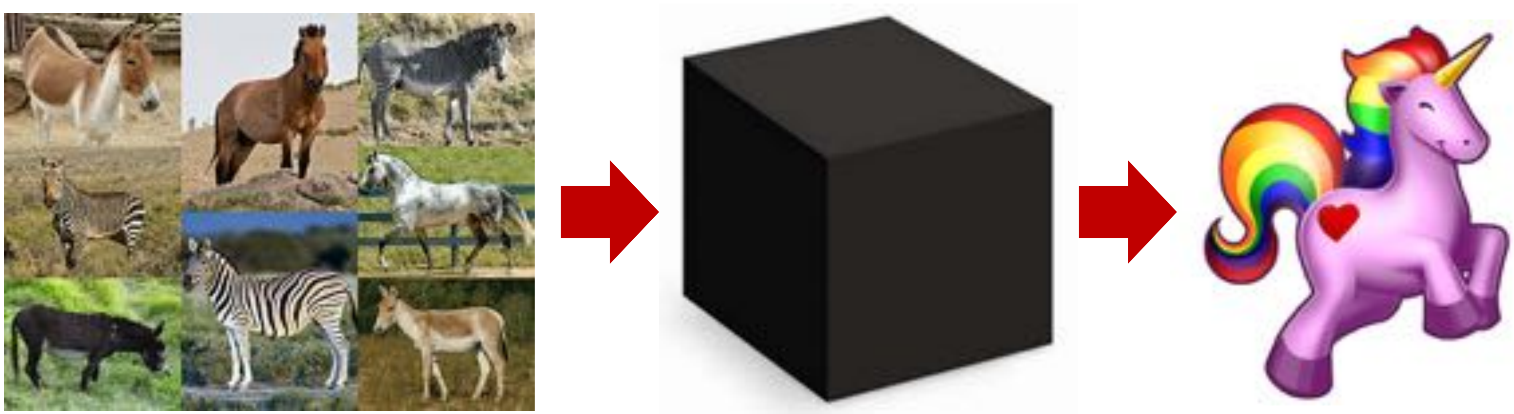
Deep Learning
with GPUs

Face Detection



Self-Driving Cars

Not a black box!



Importance

More Structure

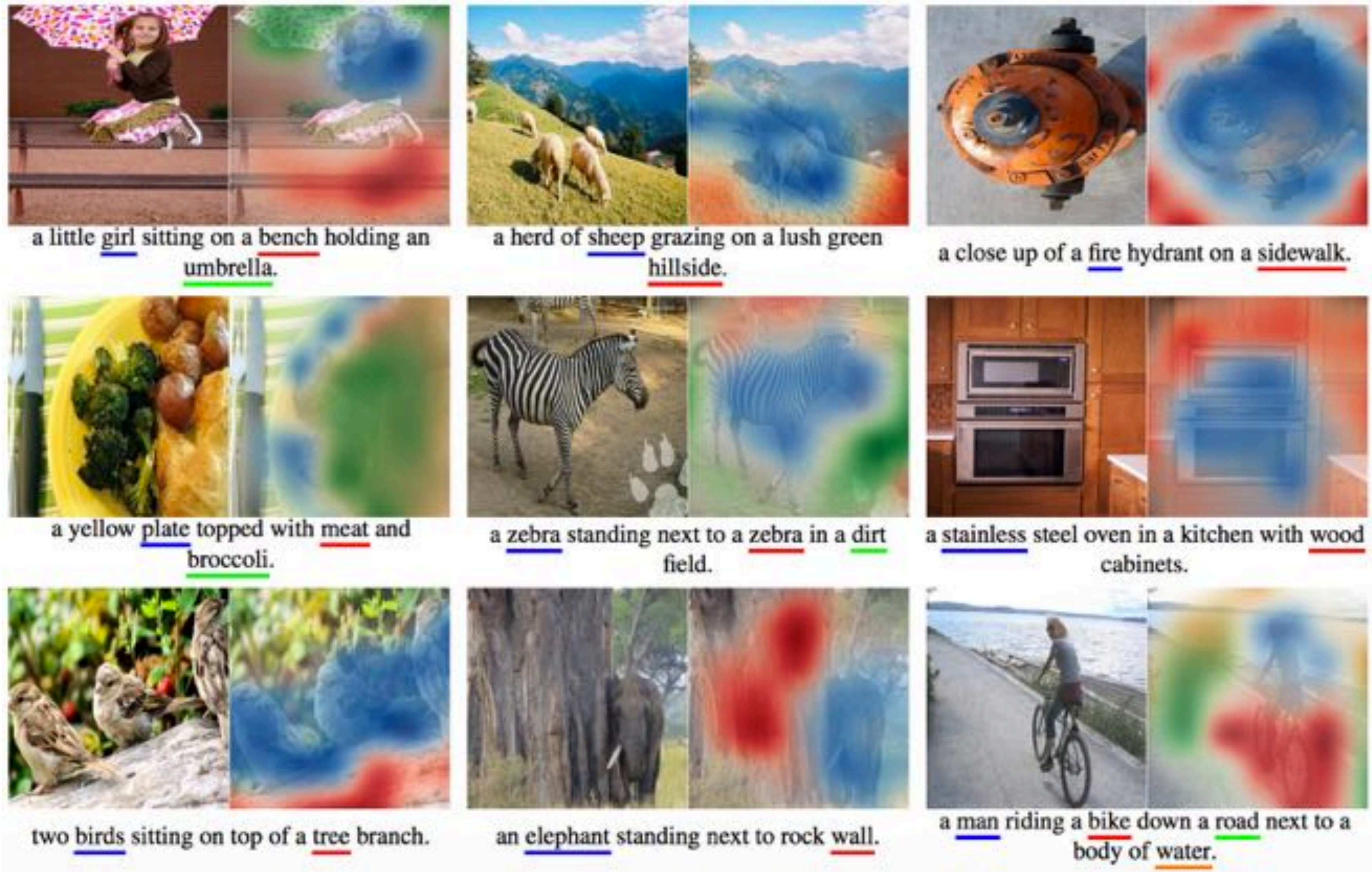
More Data

Better Machines

Better Algorithms

Object Detection

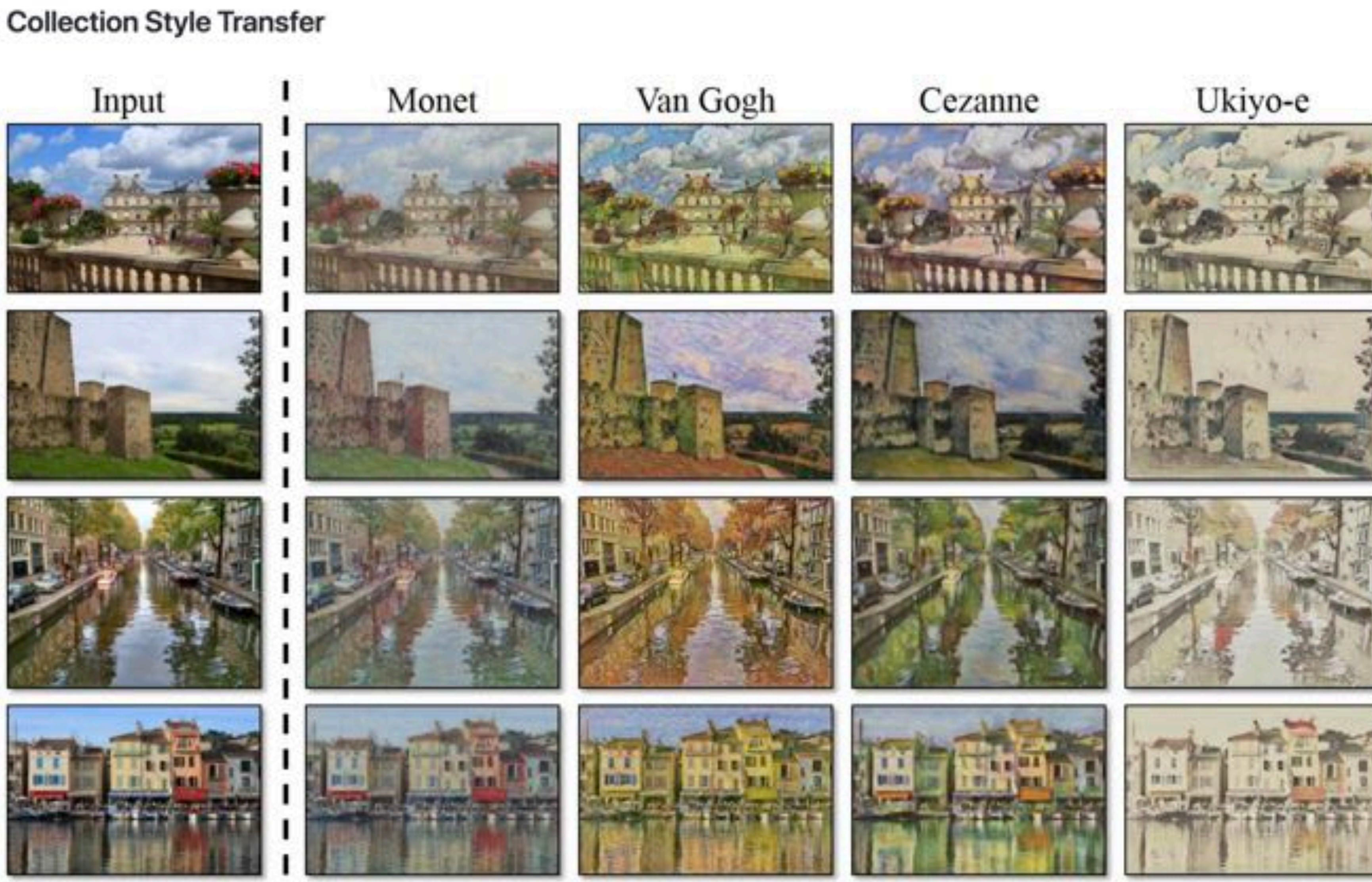
Image Captioning



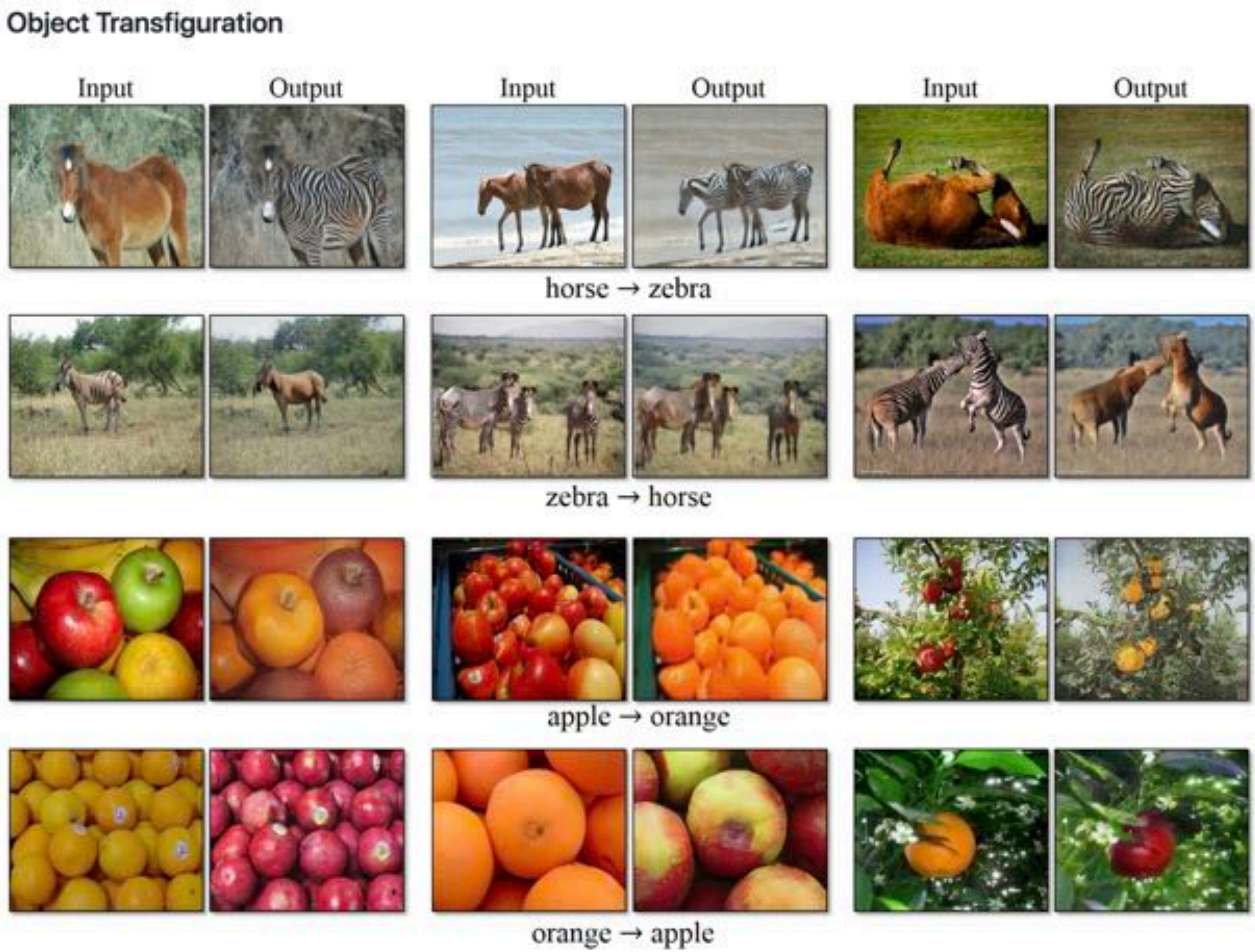
Video Prediction

Qi et al, Predicting Human Activities Using Stochastic Grammar , ICCV 2017

Generative Models



Generative Models



Generative Models

Generative Models

Season Transfer

winter Yosemite → summer Yosemite

summer Yosemite → winter Yosemite

Go

ARTICLE

doi:10.1038/nature16961

Mastering the game of Go with deep neural networks and tree search

David Silver[1]*, Aja Huang[1]*, Chris J. Maddison[1], Arthur Guez[1], Laurent Sifre[1], George van den Driessche[1], Julian Schrittwieser[1], Ioannis Antonoglou[1], Veda Panneershelvam[1], Marc Lanctot[1], Sander Dieleman[1], Dominik Grewe[1], John Nham[2], Nal Kalchbrenner[1], Ilya Sutskever[2], Timothy Lillicrap[1], Madeleine Leach[1], Koray Kavukcuoglu[1], Thore Graepel[1] & Demis Hassabis[1]

AlphaGo

- Policy Network
- Fast Rollout
- Value Network
- Monte Carlo Tree Search

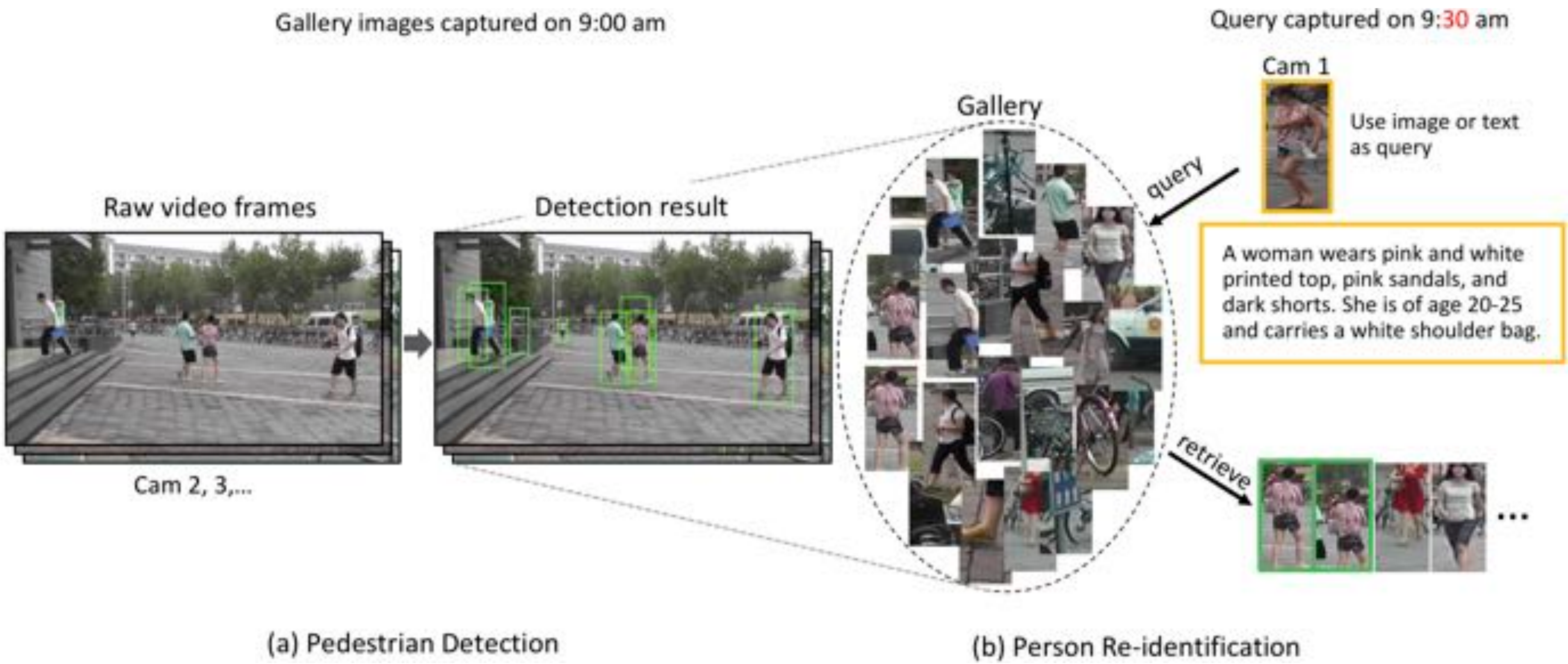
trained by a novel combination of supervised learning from human expert games, and reinforcement learning from games of self-play

Dota2

OpenAI created a bot which beats the world's top professionals at 1v1 and 5v5 matches of Dota 2 under standard tournament rules.

Person re-identification in video surveillance

- Given an image of a person-of-interest, we aim to tell whether this person has appeared in certain surveillance cameras.

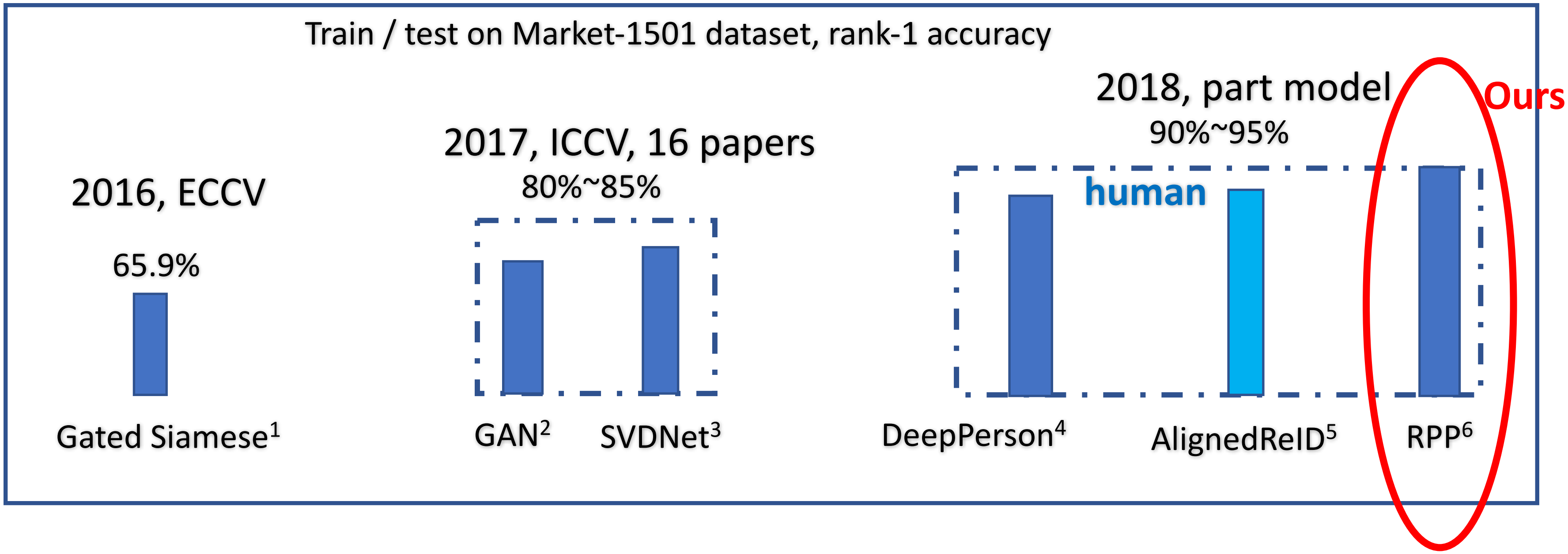


Person re-identification in video surveillance

- Given an image of a person-of-interest, we aim to tell whether this person has appeared in certain surveillance cameras.

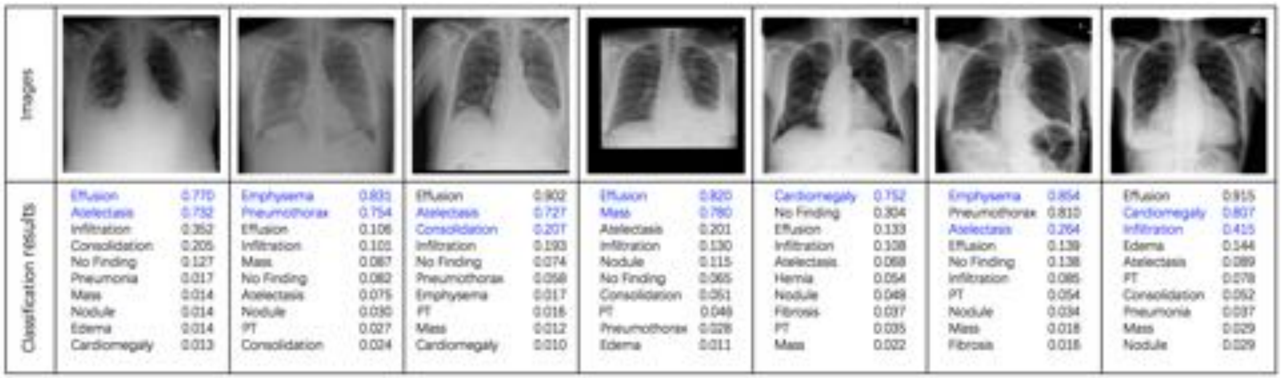
- Applications
  - Finding a missing child / elderly man in a shopping mall or a neighborhood.
  - Locating a suspect in an airport or even in much larger camera networks.

Comparison with the state of the art



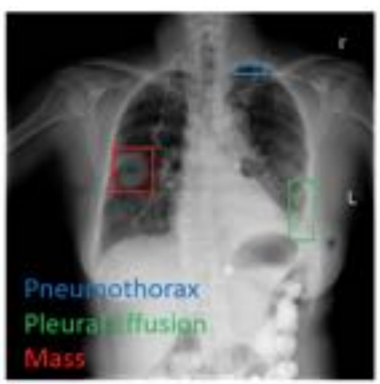
1. R. R. Varior, M. Haloi, and G. Wang. Gated Siamese convolutional neural network architecture for human reidentification. In ECCV, 2016.
2. Z. Zheng, L. Zheng, and Y. Yang. Unlabeled samples generated by gan improve the person re-identification baseline in vitro. In ICCV, 2017.
3. Y. Sun, L. Zheng,W. Deng, and S.Wang. SVDNet for pedestrian retrieval. In ICCV, 2017.
4. X. Bai, M. Yang, T. Huang, Z. Dou, R. Yu, and Y. Xu. Deepperson: Learning discriminative deep features for person reidentification. In arXiv: 1711.10658, 2017.
5. X. Zhang, H. Luo, X. Fan, W. Xiang, Y. Sun, et.al,. Alignedreid: Surpassing human-level performance in person re-identification. In arXiv:1711.0818, 2017
6. Y. Sun, L. Zheng, Y. Yang, Q. Tian, and S. Wang. Beyond part models: Person retrieval with refined part pooling. In arXiv:1711.09349, 2017.

Understanding Medical Images



Disease classification

Lesion area detection



Disease reporting

Findings: left apical small pneumothorax and small left pleural effusion remains. unchanged nodular opacity right mid lung field.

Impression: removal of left chest tube with tiny left apical pneumothorax and small left pleural fluid.

Intended Learning Outcomes

Introduction

- Define machine learning in terms of algorithms, tasks, performance and experience.
- List four main types of machine learning, e.g. supervised, unsupervised, reinforcement, and transfer learning.
- Describe some potential dangers in machine learning, e.g. applying an algorithm without understanding its assumptions, forgetting that the training data could be biased.