POLITECNICO
MILANO 1863

Cognitive Robotics

2018/2019

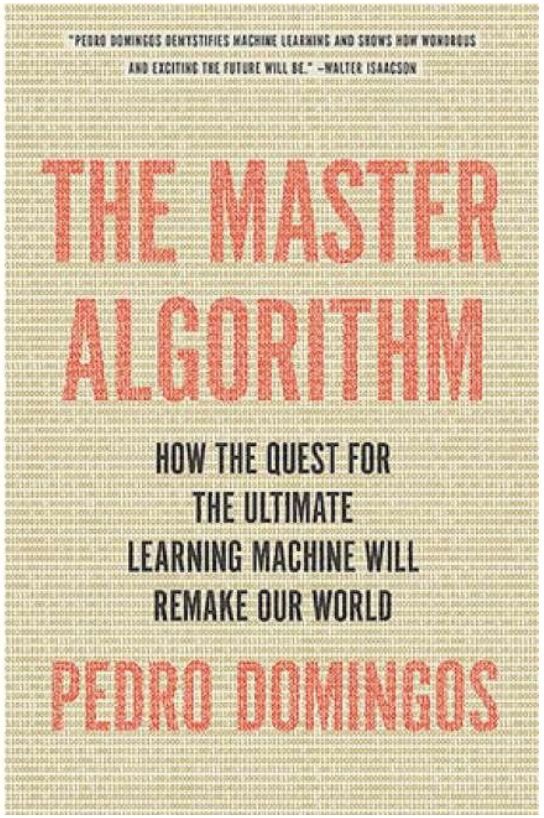
Introduction to Deep Learning

Matteo Matteucci
matteo.matteucci@polimi.it

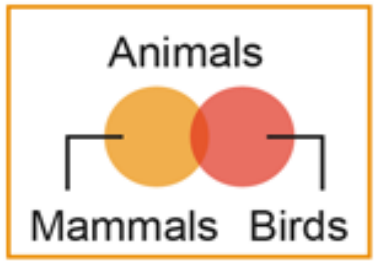
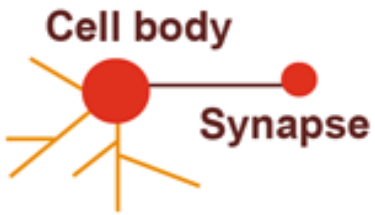
Artificial Intelligence and Robotics Lab - Politecnico di Milano

The Master Algorithm (Pedro Domingos, 2015)

"The master algorithm is the ultimate learning algorithm. It's an algorithm that can learn anything from data and it's the holy grail of machine learning …"

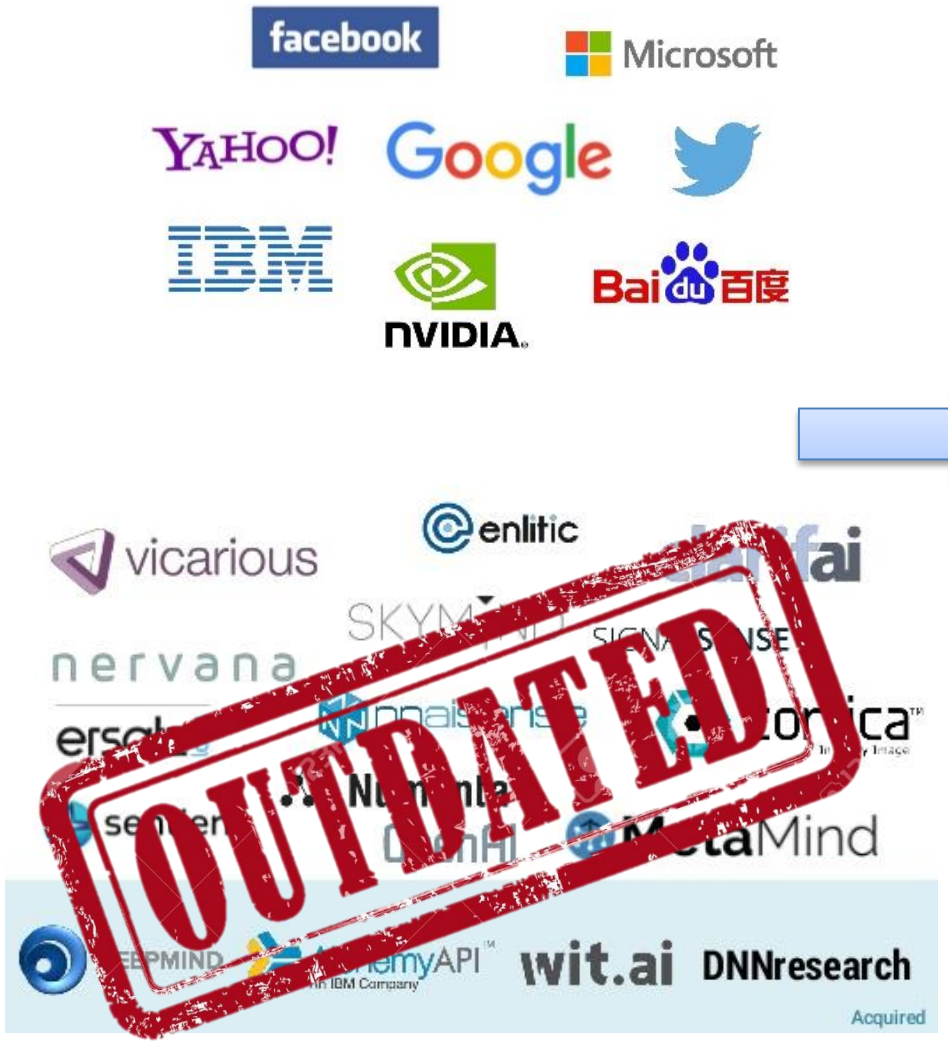


The Master Algorithm (Pedro Domingos, 2015)

Symbolists	Bayesians	Connectionists	Evolutionaries	Analogizers
				
Use symbols, rules, and logic to represent knowledge and draw logical inference	Assess the likelihood of occurrence for probabilistic inference	Recognize and generalize patterns dynamically with matrices of probabilistic, weighted neurons	Generate variations and then assess the fitness of each for a given purpose	Optimize a function in light of constraints ("going as high as you can while staying on the road")
Favored algorithm Rules and decision trees	**Favored algorithm** Naive Bayes or Markov	**Favored algorithm** Neural networks	**Favored algorithm** Genetic programs	**Favored algorithm** Support vectors

Source: Pedro Domingos, *The Master Algorithm*, 2015

Is Deep Learning the Master Algorithm?



MIT Technology Review

10 BREAKTHROUGH TECHNOLOGIES 2013

Introduction | The 10 Technologies | Past Years

Deep Learning
With massive amounts of computational power, machines can now recognize objects and translate speech in real time. Artificial intelligence is finally getting smart.

Temporary Social Media
Messages that quickly self-destruct could enhance the privacy of online communications and make people freer to be spontaneous.

Prenatal DNA Sequencing
Reading the DNA of fetuses will be the next frontier of the genomic revolution. But do you really want to know about the genetic problems or musical aptitude of your unborn child?

Additive Manufacturing
Skeptical about 3-D printing? GE, the world's largest manufacturer, is on the verge of using the technology to make jet parts.

Baxter: The Blue-Collar Robot
Rodney Brooks's newest creation is easy to interact with, but the complex innovations behind the robot show just how hard it is to get along with people.

Memory Implants
A maverick neuroscientist believes he has deciphered the code by which the brain forms long-term memories. Next: testing a prosthetic implant for people suffering from long-term memory loss.

Smart Watches
The designers of the Pebble watch realized that a mobile phone is more useful if you don't have to take it out of your pocket.

Ultra-Efficient Solar Power
Doubling the efficiency of a solar cell would completely change the economics of renewable energy. Nanotechnology just might make it possible.

Big Data from Cheap Phones
Collecting and analyzing information from simple cell phones can provide surprising insights into how people move about and behave – and even help us understand the spread of diseases.

Supergrids
A new high-power circuit breaker could finally make highly efficient DC power grids practical.

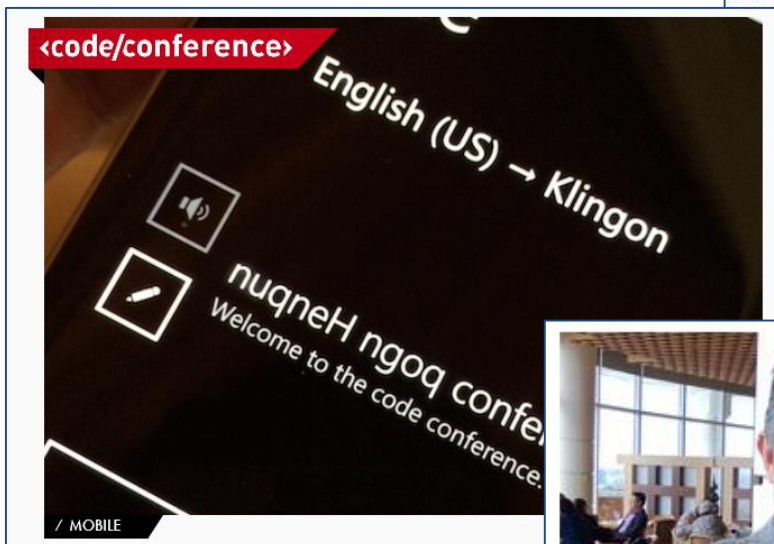
Enabling Cross-Lingual Conversations in Real Time

Microsoft Research
May 27, 2014 5:58 PM PT

The success of the team's progress to date was on display May 27, in a talk by Microsoft CEO Satya Nadella in Rancho Palos Verdes, Calif., during the Code Conference. During Nadella's conversation
Kara Swisher and Walt Mossberg of
Re/code tech website relating to a new
of personal computing, he asked
deep Pall to join him on stage. Pall, the
rosoft corporate vice president of
pe, demonstrated for the first time
licly the Skype Translator app, with Pall
versing in English with German-

View milestones on the path to Skype Translator

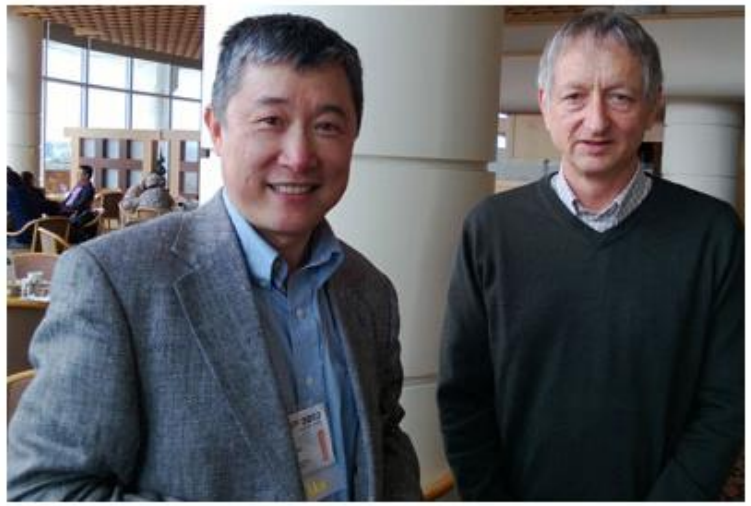
#speech2speech



Microsoft's Skype "Star Trek" Language Translator Takes on Tower of Babel

May 27, 2014, 5:48 PM PDT

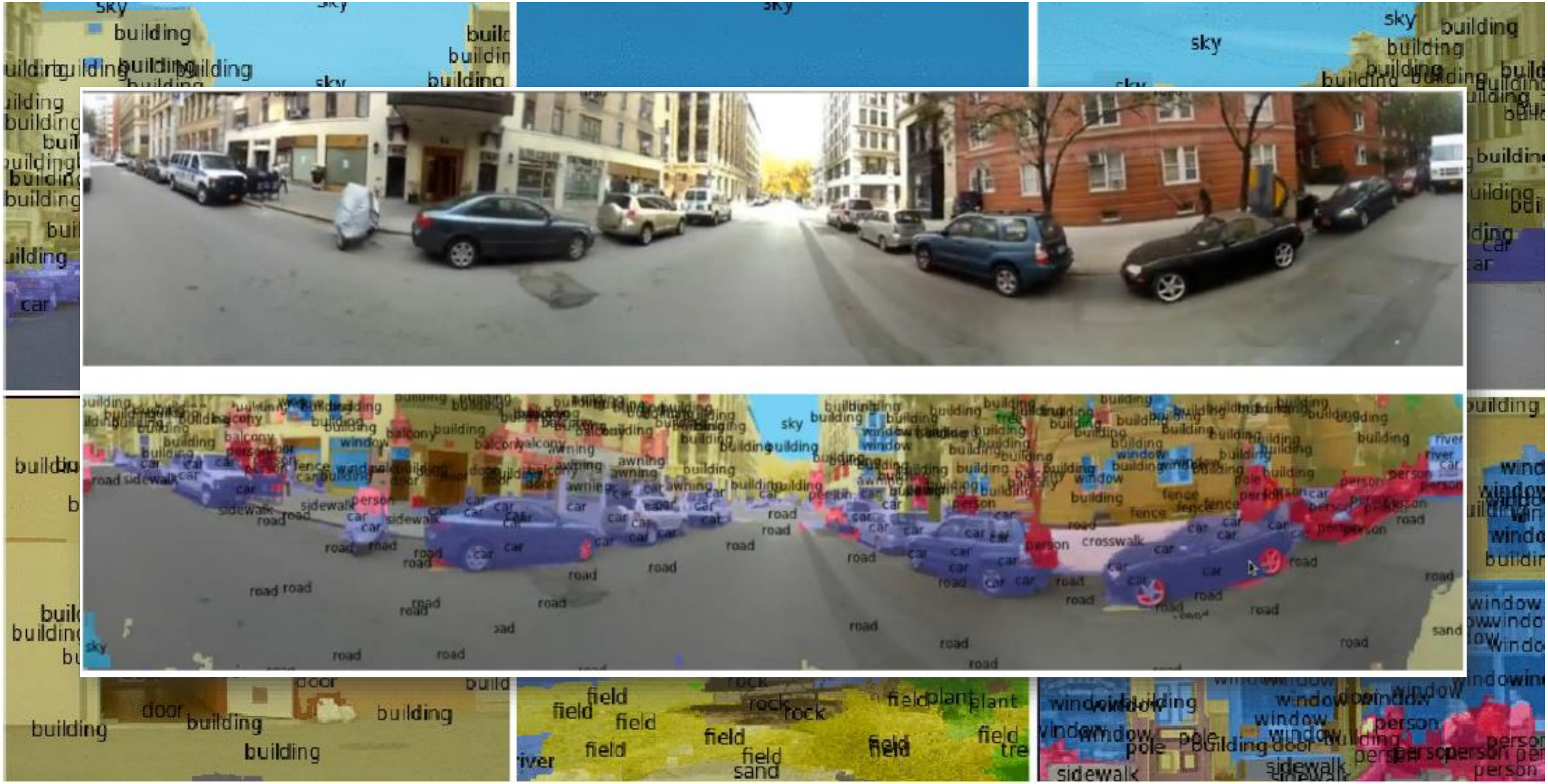
Remember the universal translator on Star Trek? The gadget th
to aliens?



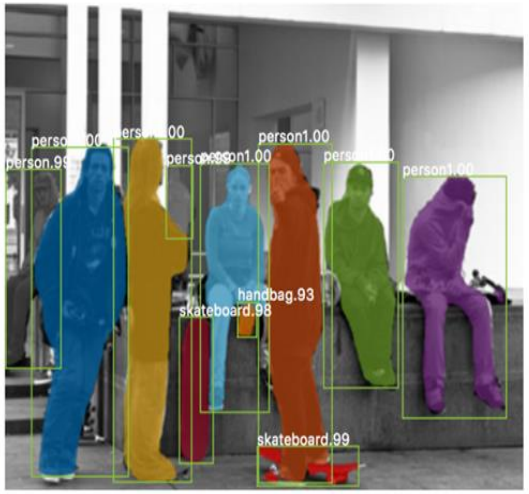
Li Deng (*left*) and Geoff Hinton.

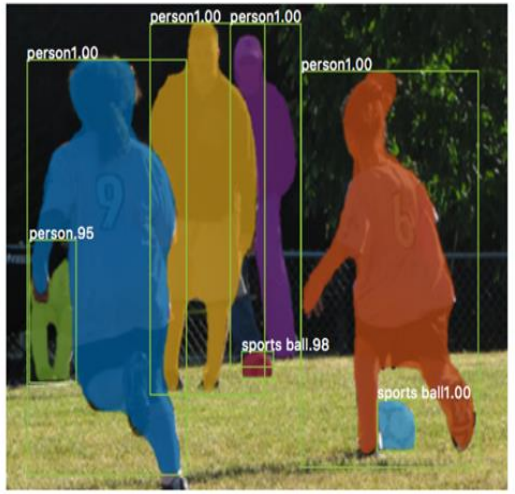
A core development that enables Skype translation came from Redmond researcher Li Deng. He invited Geoff Hinton, a professor at the University of Toronto, to visit Redmond in 2009 to work on new neural-network learning methods, based on a couple of seminal papers from Hinton and his collaborators in 2006 that had brought new

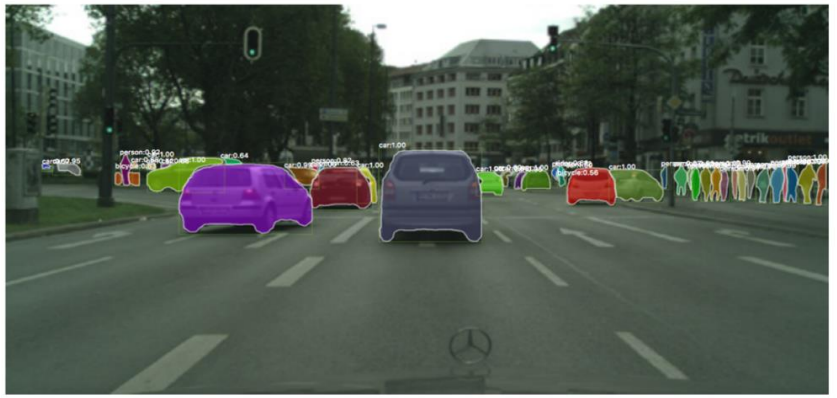
The path to the Skype Translator gained momentum with an encounter in the autumn of 2010. Seide and colleague Kit Thambiratnam had developed a system they called The Translating! Telephone for live speech-to-text and speech-to-speech translation of phone calls.



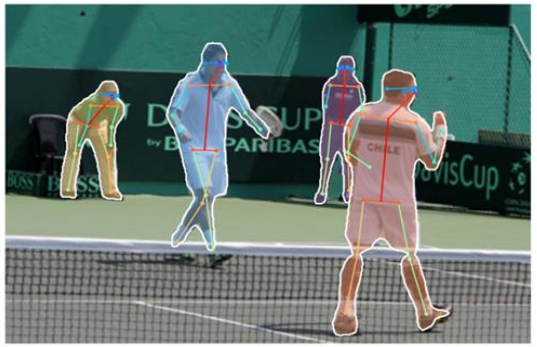

umbrella.99
person1.00
person1.00
person1.00
person1.00
car.99
car.93

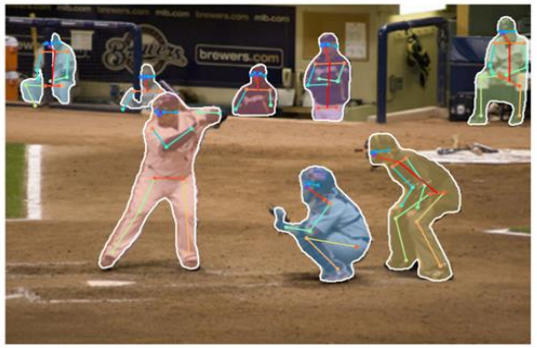

person1.00
handbag.93
skateboard.98
skateboard.99

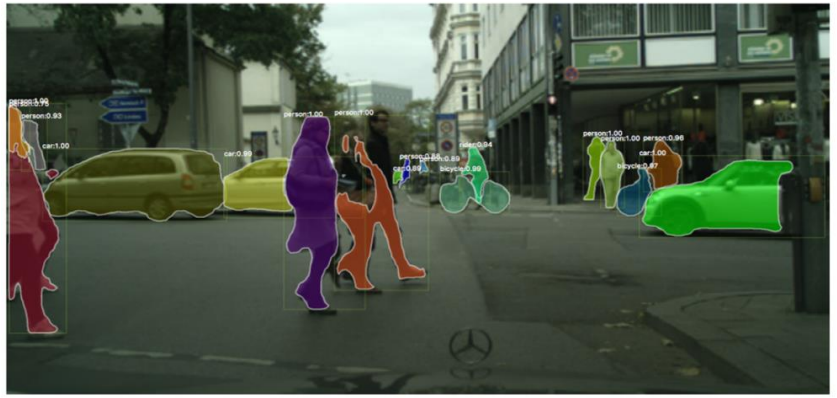

person1.00
person1.00
person1.00
person1.00
person.95
sports ball.98
sports ball1.00




imax PHOTOGRAPHY




brewers.com

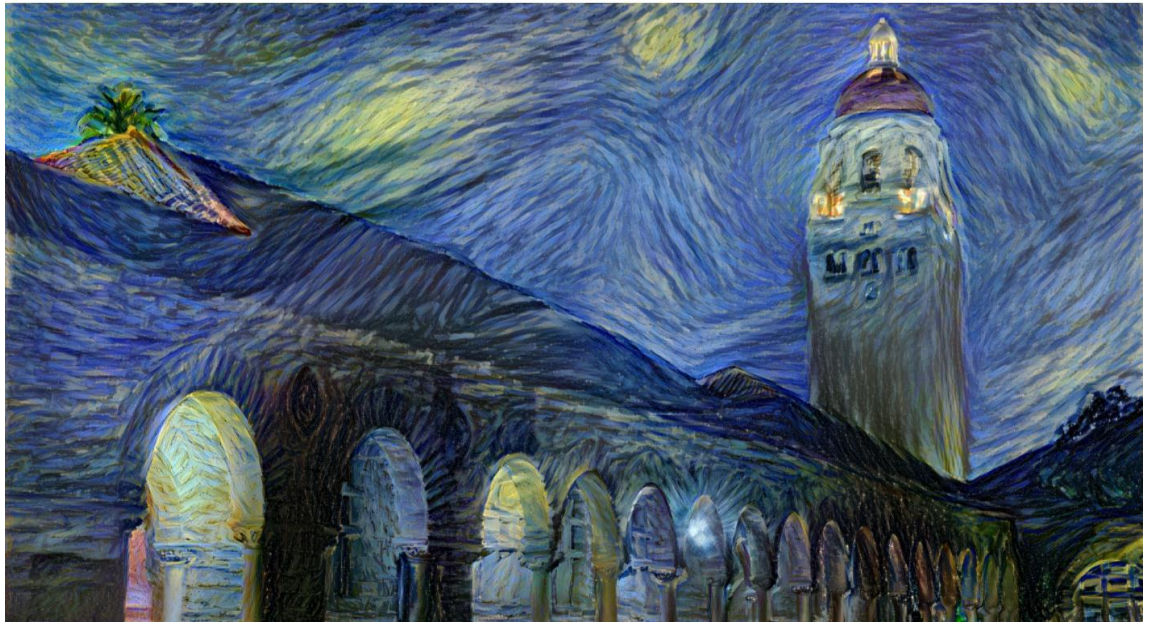


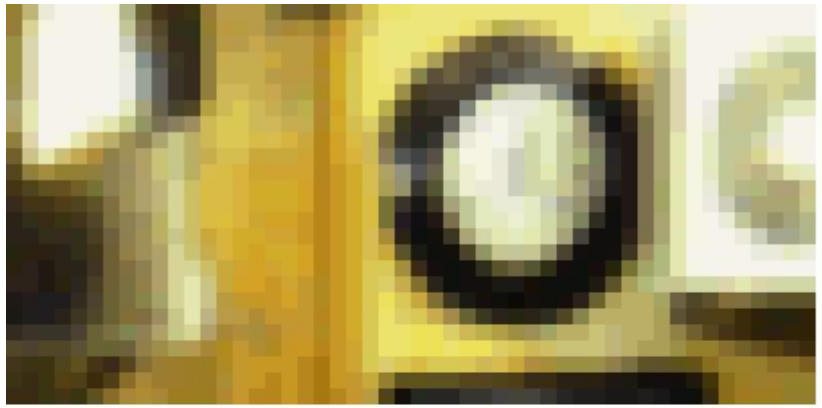
https://github.com/luanfujun/deep-photo-styletransfer

https://github.com/jcjohnson/neural-style

https://github.com/jcjohnson/fast-neural-style

https://ml4a.github.io/ml4a/style_transfer/

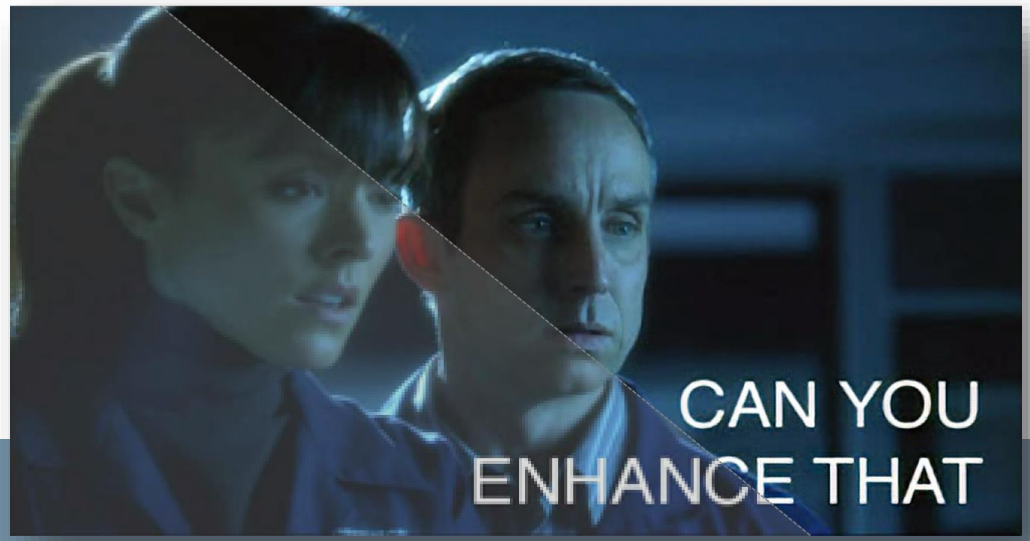









https://github.com/alexjc/neural-enhance



Text description	This flower has petals that are white and has pink shading	This flower has a lot of small purple petals in a dome-like configuration	This flower has long thin yellow petals and a lot of yellow anthers in the center	This flower is pink, white, and yellow in color, and has petals that are striped	This flower is white and yellow in color, with petals that are wavy and smooth	This flower has upturned petals which are thin and orange with rounded edges	This flower has petals that are dark pink with white edges and pink stamen
256x256 StackGAN							

Text description	This bird is red and brown in color, with a stubby beak	The bird is short and stubby with yellow on its body	A bird with a medium orange bill white body gray wings and webbed feet	This small black bird has a short, slightly curved bill and long legs	A small bird with varying shades of brown with white under the eyes	A small yellow bird with a black crown and a short black pointed beak	This small bird has a white breast, light grey head, and black wings and tail
256x256 StackGAN							

'Go is implicit. It's all pattern matching. But that's what deep learning does very well.'

—DEMIS HASSABIS, DEEPMIND

The win is more than a novelty. Online services like Google, Facebook, and Microsoft, already use deep learning to identify images, recognize spoken words, and understand natural

with a technology called reinfor
methods, point the way to a futu
can learn to perform physical ta
environment. "It's a natural fit f

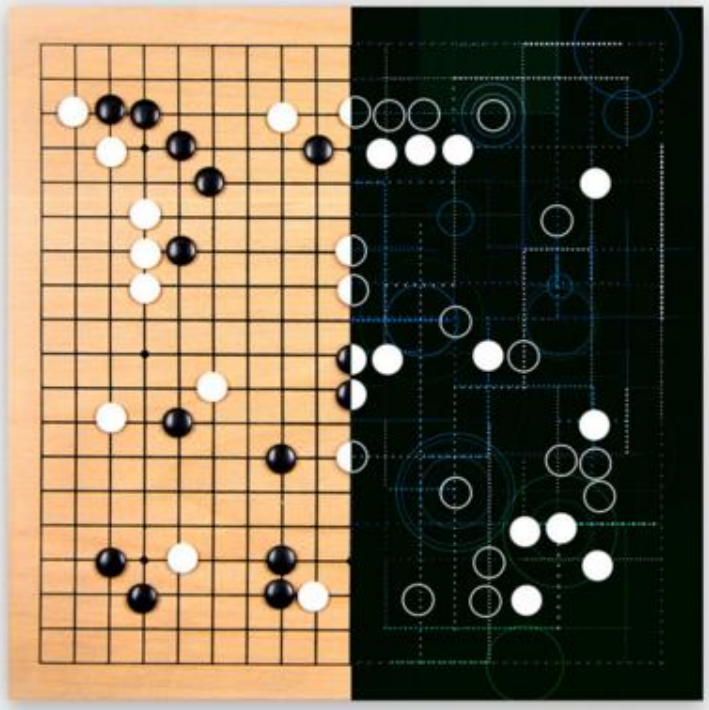
It's incredibly difficult to build a machine that duplicates the kind of intuition that makes the top human players so good at

In the mid-'90s, a computer program called Chinook beat the world's top player at the game of checkers. A few years later, IBM's Deep Blue supercomputer shocked the chess world when it wiped the proverbial floor with world champion Gary Kasparov. And more

BM machine, Watson, topped the best
dy!, the venerable TV trivia game.
o mastered Othello, Scrabble,
poker. But in the wake of Crazy Stone's
Coulom predicted that another ten years
a machine could beat a grandmaster
rt.

IN A HUGE BREAKTHROUGH, GOOGLE'S AI BEATS A TOP PLAYER AT THE GAME OF GO

WIRED



Deep Learning in a nutshell

Can you guess what it is in the picture?

What if I enlarge it?

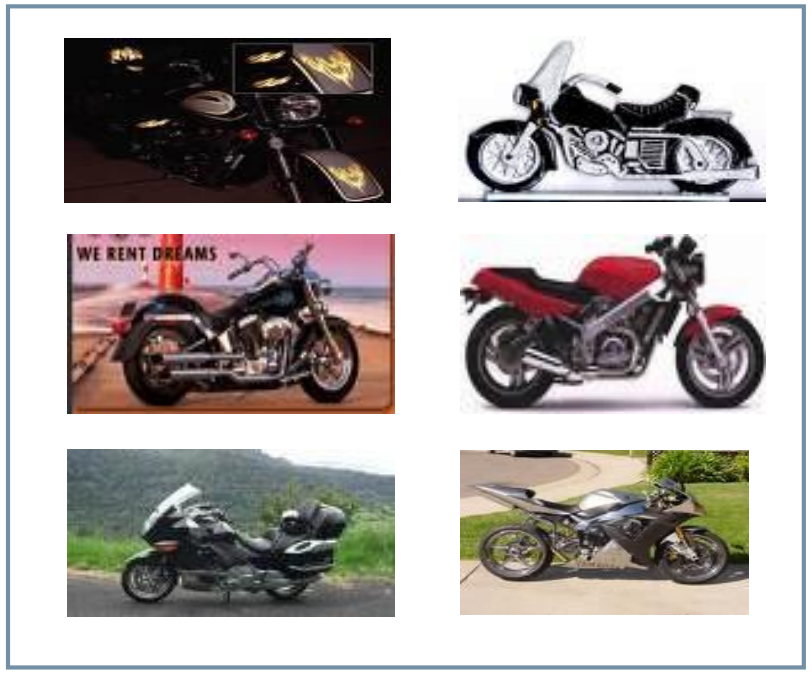
What if I enlarge it even more?

Let's use context + common knowledge

Deep Learning uses Big Data to learn common sense and context representation …

Recall about Supervised Learning

Cars



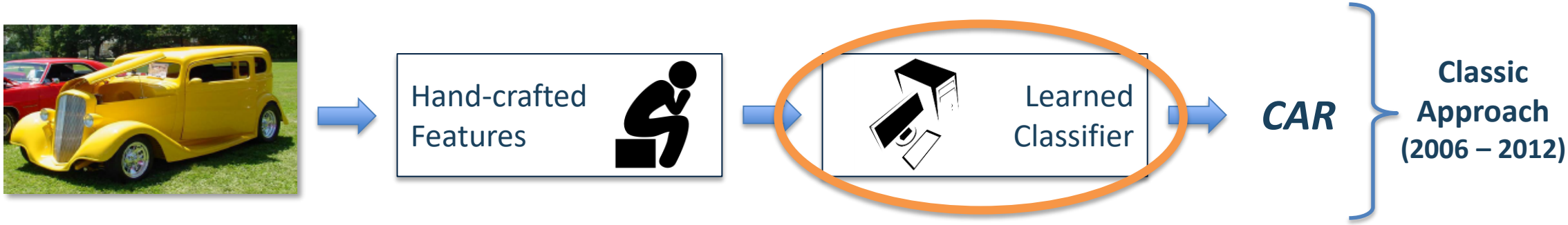
Motorcycles

Hand-crafted Features

Learned Classifier

CAR

Recall about Supervised Learning

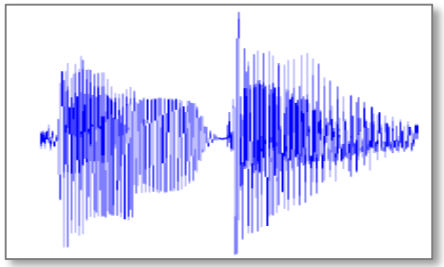


Features are based on domain knowledge or heuristics:

- Words in a Dictionary for text classification
- MFCC for Speech Recognition
- SIFT, HoG, BRIEF in Visual Tasks

Modern Pattern Recogniton

Speech recognition (early 90's – 2011)



Low level features

Fixed Features

MFCC

Mid level features

Mixture of Gaussian

Unsupervised Learning

Learned Classifier

Supervised Learning

Horn

Object rec 06 – 2012)

Low level features

Fixed Features

SIFT/HoG

Mid level features

K-means Sparse Coding

Unsupervised Learning

Pooling

Learned Classifier

Supervised Learning

Car

It’s all about features ...

What if we do not get these right?

Hand-crafted Features → Learned Feature Projection → Learned Classifier → ***CAR***

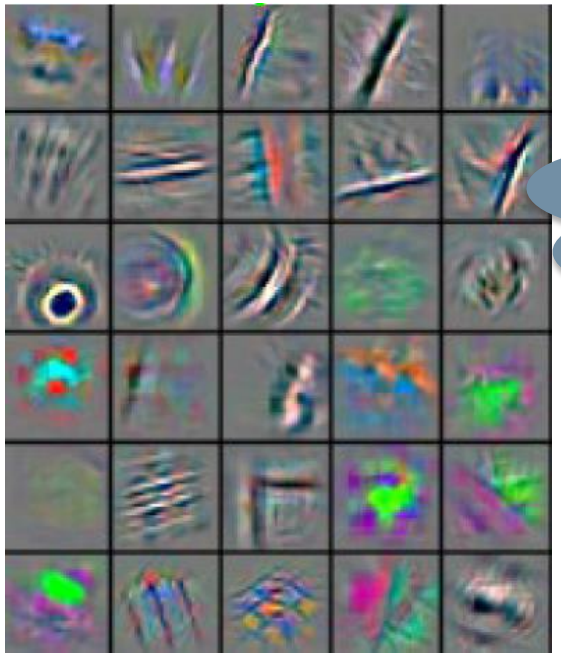
Classic approach (2006 – 2012)

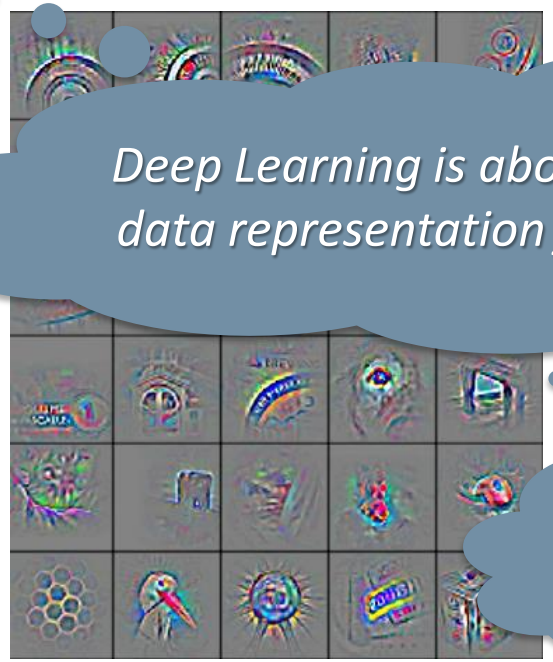
It's all about features ...

Learned Features → Learned Features → Learned Features → Learned Classifier → *CAR*

Deep Learning (2012 – ...)

Learn from data!





Deep Learning is about learning data representation from data!

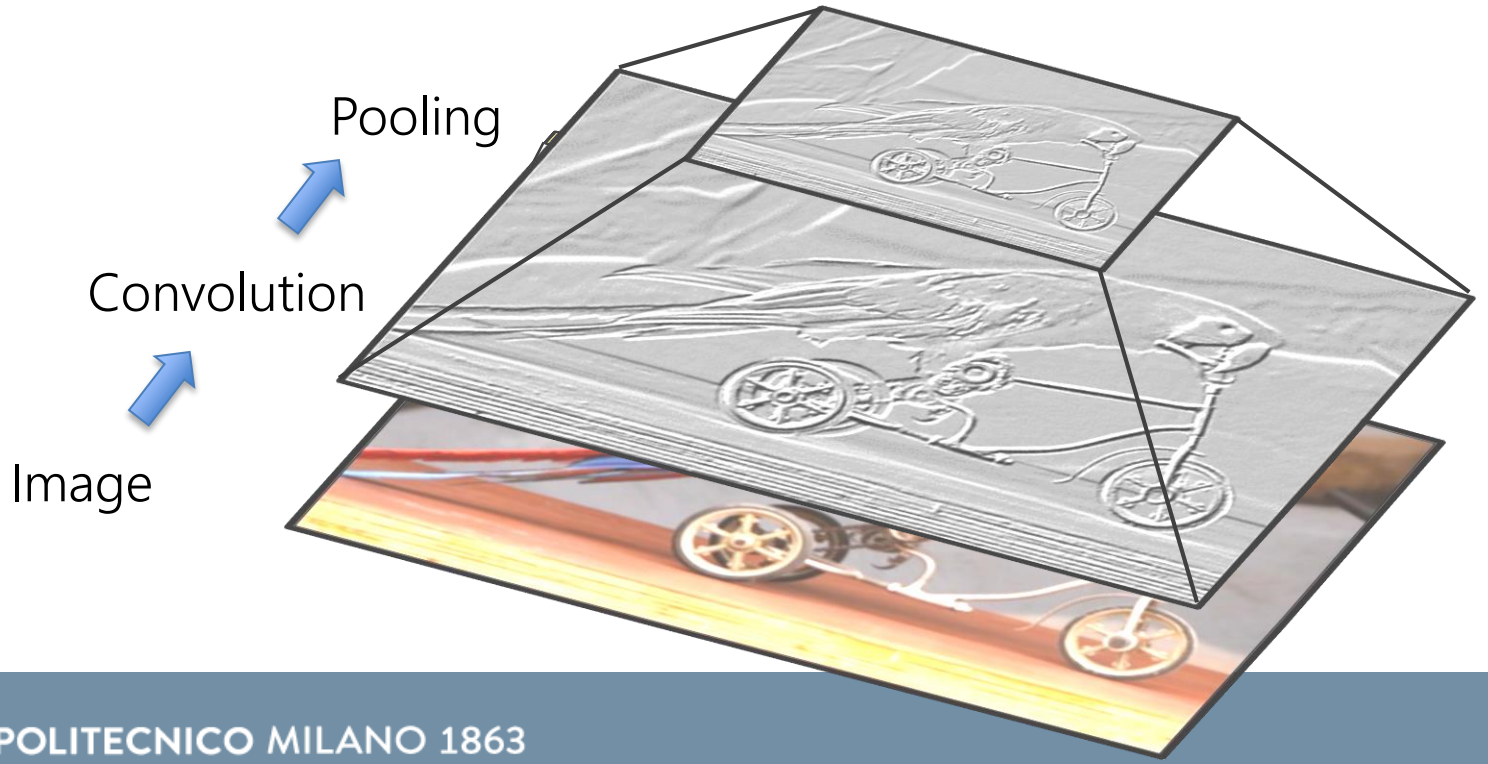
But which data?

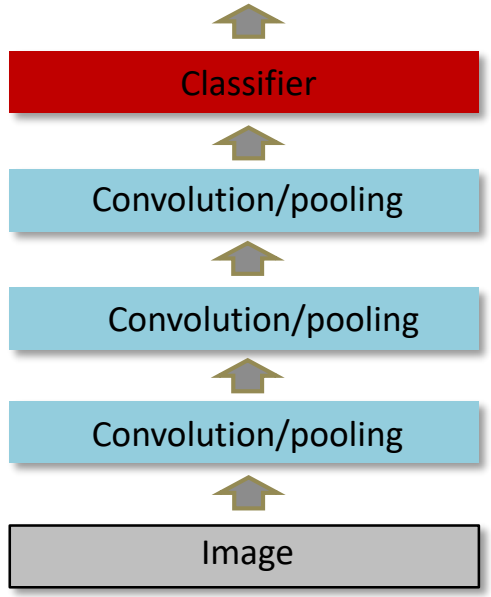

IN 60 SECONDS...
1 NEW DEFINITION IS ADDED ON urban
1,600+ READS ON Scribd.
13,000+ HOURS MUSIC STREAMING ON PANDORA
12,000+ NEW ADS POSTED ON craigslist
370,000+ MINUTES VOICE CALLS ON skype
98,000+ TWEETS
20,000+ NEW POSTS ON tumblr.
THE LARGEST SOCIAL READING PUBLISHING COMPANY!!
New Craigslist Ads
320+ NEW twitter ACCOUNTS
13,000+ iPhone APPLICATIONS DOWNLOADED
100+ NEW Linked in ACCOUNTS
1 associatedcontent NEW ARTICLE IS PUBLISHED
THE WORLD'S LARGEST COMMUNITY CREATED CONTENT!!
QUESTIONS ASKED ON THE INTERNET...
100+ Answers.com
40+ YAHOO! ANSWERS
600+ NEW VIDEOS
6,600+ NEW PICTURES ARE UPLOADED ON flickr
50+ WordPress DOWNLOADS
25+ HOURS TOTAL DURATION
70+ DOMAINS REGISTERED
695,000+ facebook STATUS UPDATES
125+ PLUGIN DOWNLOADS
60+ NEW BLOGS
1,500+ BLOG POSTS
168 MILLION EMAILS ARE SENT
694,445 SEARCH QUERIES
1,700+ Firefox DOWNLOADS
79,364 WALL POSTS
510,040 COMMENTS
Google
Google Search

... now we have Deep Learning!

Learned features → Learned features → Learned features → Learned Classifier → ***CAR***

Deep Learning (2012 – ...)

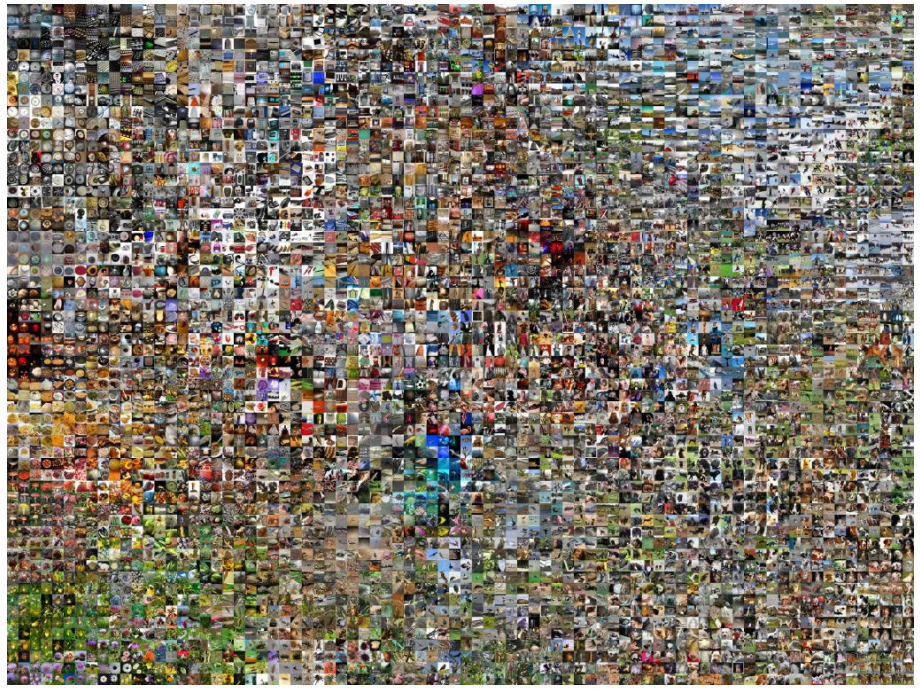




IMAGENET

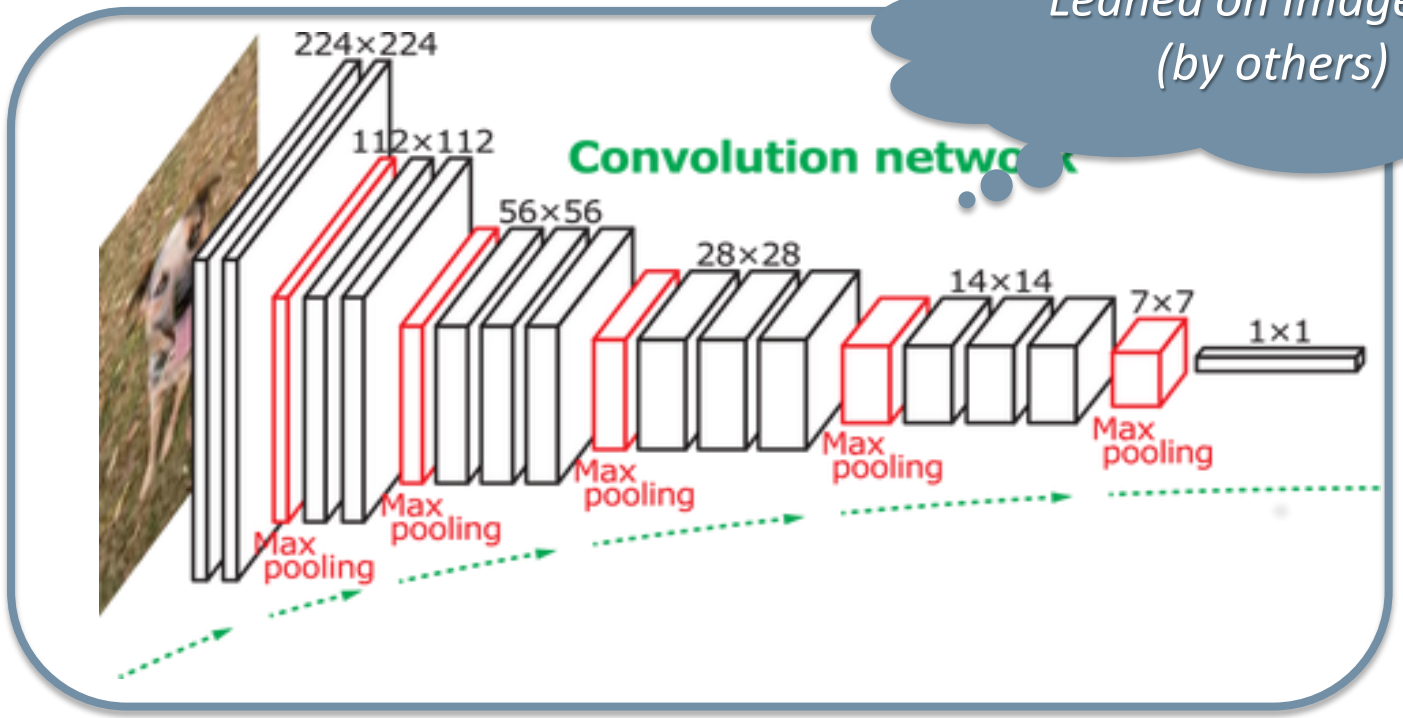
1000 categories, 1.5 M labeled images (2012)

Supervision (ImageNet - 2012)



What if our labels are not among these 1K?

Transfer learning!!!



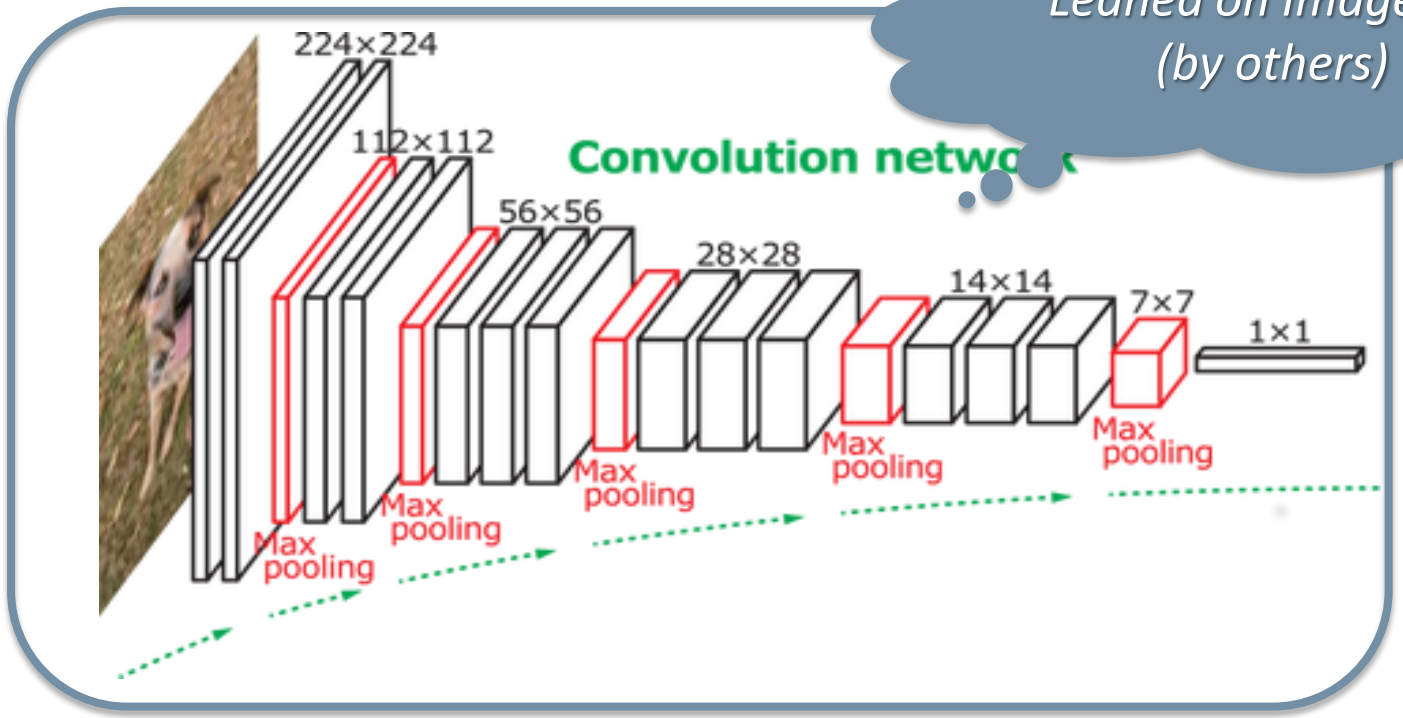
Leaned on ImageNet (by others)

Classifier

Car

Leaned on your data

Transfer learning!!!

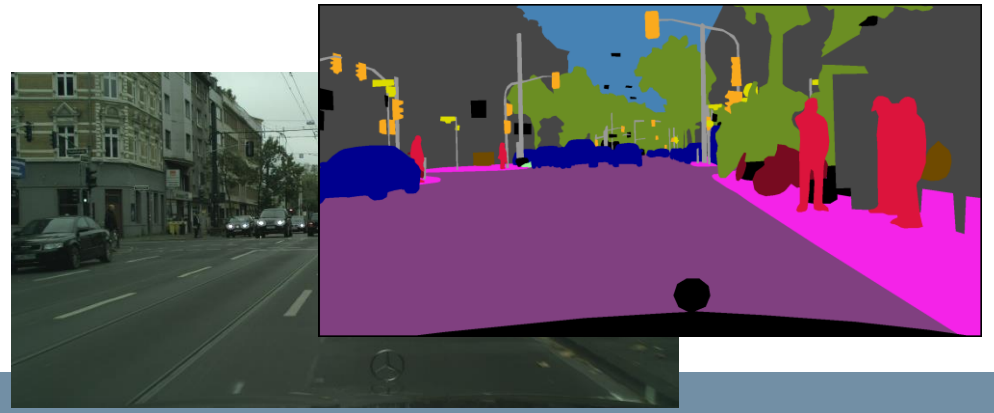


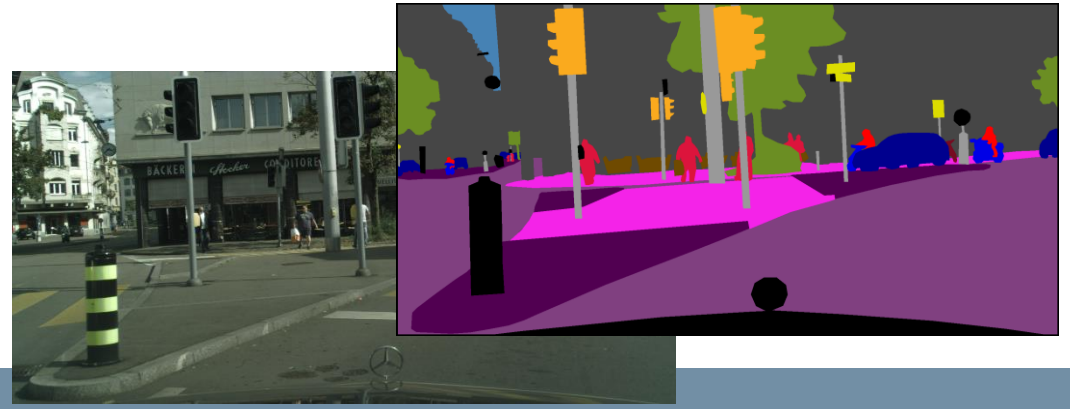
Leaned on ImageNet (by others)

Classifier

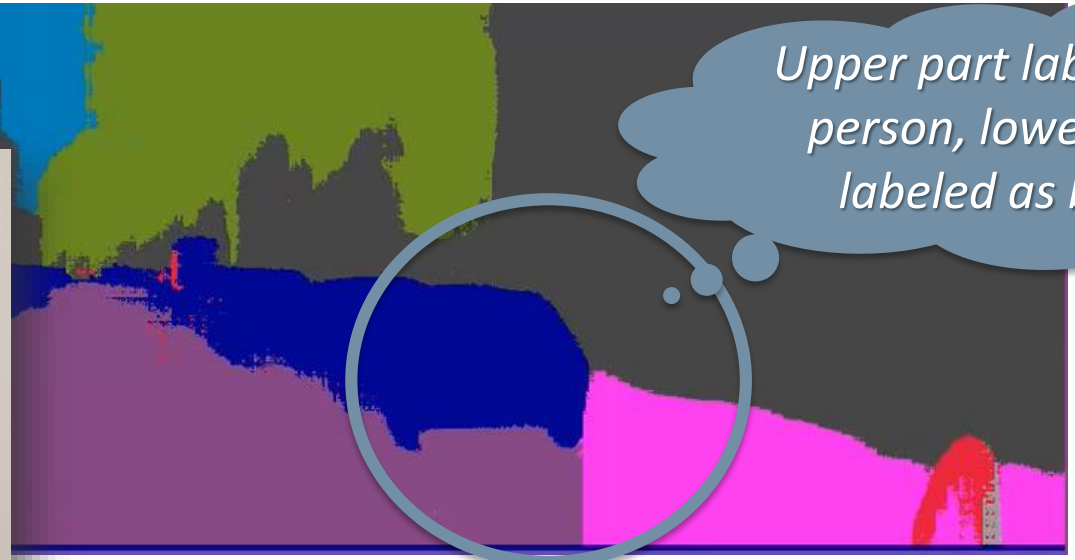
Car

Leaned on your data





«On every street»



Upper part labeled as person, lower part labeled as bike!

Not too different from this, isn't it? ;-)