Advanced Topics in Machine Learning

Paolo Favaro

Contents

Course Information

Focus

- This year the course will be focused on **deep learning (artificial intelligence)**
- Artificial Intelligence is about developing intelligent software to automate routine labor, understand speech or images, make diagnoses in medicine and support basic scientific research
- Modern AI follows the machine learning paradigm where rules are **learned from examples**, rather than being hard-coded

Motivation

- **Impact in industry**: Applied by top hi-tech companies including Google, Microsoft, Facebook, IBM, Baidu, Apple, Adobe, Netflix, NVIDIA and NEC.
- **Impact in academia**: e.g., Neuroscience (modern convolutional networks for object recognition are a model of visual processing), drug design (predict molecules interaction), Physics (search for subatomic particles), Microscopy (automatic parsing of images to construct a 3-D map of the human brain)

Course Information

- 31088/61088, Lectures and Exercises, 5.0 ECTS
- **Organized in 14 weeks**
- Lectures
  - Location: Online https://us02web.zoom.us/j/88157161793?pwd=Qjhad2thU2NWMzh0a1RPdHZtYlNnQT09
  - Time: Tuesdays, 09.15-12.00
- Course webpage
  - ILIAS
  - http://www.cvg.unibe.ch
- Sources
  - deeplearningbook.com, Stanford CS231n course
  - http://uvadlc.github.io/ UAV deep learning course, University of Amsterdam

Course Information

- Lecturer
  - Prof Dr Paolo Favaro, paolo.favaro@inf.unibe.ch
  - Office: Room 213, Neubrückstrasse 10
- Teaching assistants
  - Mr Adam Bielski, adam.bielski@inf.unibe.ch
  - Mr Llukman Cerkezi, llukman.cerkezi@inf.unibe.ch
  - Mr Aram Davtyan, aram.davtyan@inf.unibe.ch
  - Mr Alp Eren Sari, alp.sari@inf.unibe.ch
  - Office: Room 212, Neubrückstrasse 10

Course Registration

- At the beginning of each semester, students must register for the teaching units that they want to take during that semester
- Students who have registered for a teaching unit, but later change their mind about attending it, are requested to cancel their registration
- (De-)Registration in Academia through SWITCHaai with your account and password of your affiliated university
  http://mcs.unibnf.ch/admin

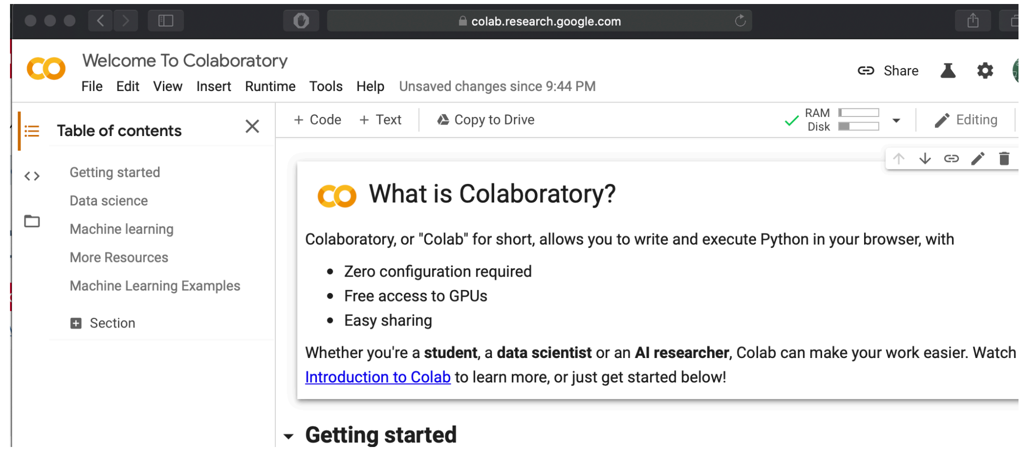
Exam Registration

- Students have to register for the exam on **15.06.21** before the registration deadline on **15.05.21**
- Students who do not register to take the exam at the end of the teaching unit are not allowed to take the repetition exam
- Registration for an exam can be cancelled or withdrawn (this procedure depends on when it takes place)
- (De-)Registration in Academia through SWITCHaai with your account and password of your affiliated university
  http://mcs.unibnf.ch/admin

Course Information

- **Resources**
  - Books
    - **Deep Learning** by Ian Goodfellow, Yoshua Bengio, Aaron Courville (available online at www.deeplearningbook.org)
    - **Neural Networks and Deep Learning** by Michael Nielsen (free online book at www.neuralnetworksanddeeplearning.com)
    - **Pattern Recognition and Machine Learning** by Christopher M. Bishop
  - Computing
    - **Google COLABORATORY**
    - **As a backup: CVG GPU cluster**

Computing:
Google Colab



Code + report

Colab notebooks allow you to combine **executable code** and **rich text** in a single document, along with **images**, **HTML**, **LaTeX** and more. When you create your own Colab notebooks, they are stored in your Google Drive account. You can easily share your Colab notebooks with co-workers or friends, allowing them to comment on your notebooks or even edit them. To learn more, see Overview of Colab. To create a new Colab notebook you can use the File menu above, or use the following link: create a new Colab notebook.

Colab notebooks are Jupyter notebooks that are hosted by Colab. To learn more about the Jupyter project, see jupyter.org.

Data science

With Colab you can harness the full power of popular Python libraries to analyze and visualize data. The code cell below uses **numpy** to generate some random data, and uses **matplotlib** to visualize it. To edit the code, just click the cell and start editing.

```
[ ]  import numpy as np
     from matplotlib import pyplot as plt

     ys = 200 + np.random.randn(100)
```

Course Information

- **Assessment**
  - Coursework will consist of
    - 3 personal assignments (every 3 weeks)
  - Final grade will be split as follows
    - [30%] personal assignments
    - [70%] final exam

Assignments

- Personal work (no group work)
- Aimed at practical understanding of deep learning methods
- Reports must be submitted every three weeks (when the new assignment is provided)
- Mark is used towards the final grade
- Output: **report (jupyter format — code + text)**

Timeline

week \	assign start	assign due	topics	tutorials
1				python & jupyter basics
2				pytorch
3	Assign.#1		machine learning foundations	pytorch
4				pytorch practice
5				regularization
6	Assign.#2	Assign.#1	convolutional neural networks	optimization
7			Easter Holiday	
8				mock exam part 1/2 + feedback on assign. #1
9				convnets
10	Assign.#3	Assign.#2	denoising autoencoders	autoencoders
11				generative adv. nets
12				feedback on assign. #2
13				unsupervised learning
14		Assign.#3		LSTMs
15				mock exam part 2/2 + feedback on assign. #3

Course Information

- **Review session**
  - last week of the course
  - review of the whole course material
- **Prerequisites**: applied math fundamentals like linear algebra, probability and numerical optimization (see **Chapters 2, 3, and 4** of Deep Learning by Goodfellow, Bengio, Courville)

Learning objectives

- At the end of this course you will

1. Learn how to cast and solve problems via supervised and unsupervised learning

2. Learn how modern deep learning tools work (how neural networks work, are designed, optimized and implemented)

3. Learn how to use modern deep learning libraries (PyTorch)

4. Learn soft skills: how to coordinate and plan work, scientific writing

Related Courses

- **Machine Learning (BSc) — UniBe**
    - Fall semester
    - Foundations of machine learning (supervised, unsupervised and reinforcement learning)
    - Highly recommended for this course

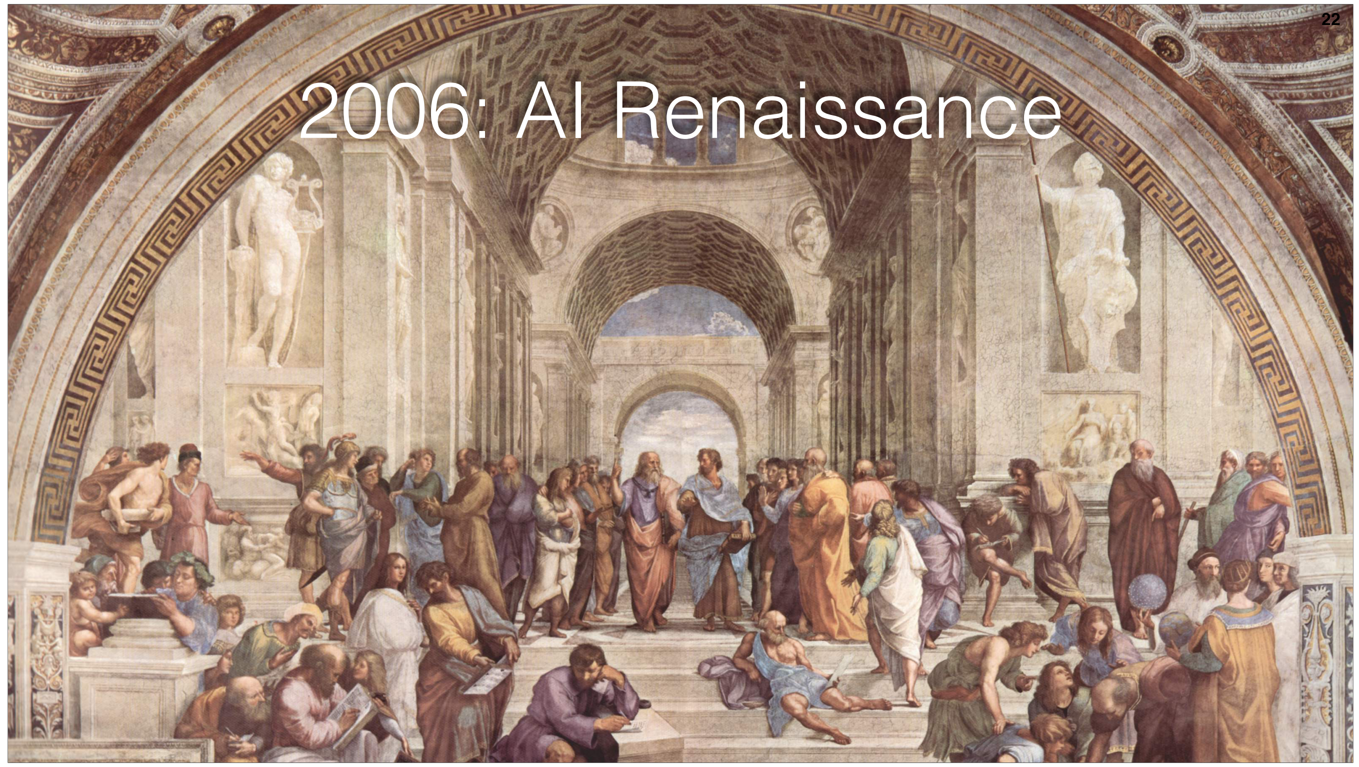
Related Courses

- **Computer Vision (MSc) — UniBe**
  - Fall semester
  - From image processing to 3D reconstruction from images; from energy optimization to bayesian methods
  - A complementary course

Introduction to Deep Learning

Contents

- Overview of applications and concepts in deep learning
- **Chapter 1** of Deep Learning by Goodfellow, Bengio, Courville



Artificial Intelligence and Neural Networks in particular are not new. They have a long history. They gained quite a lot of success in the 90s with back-propagation. Then, other techniques in ML performed better and NN became less popular. Since 2006 they have become again mainstream. This success is due to two main factors: 1) the availability of very large annotated (and not annotated) datasets (of images and videos); 2) the availability of better computational hardware (GPUs) and software architectures. Besides these two main factors, few improvements in the performance of NNs are due to novel algorithmic/theoretical ideas.

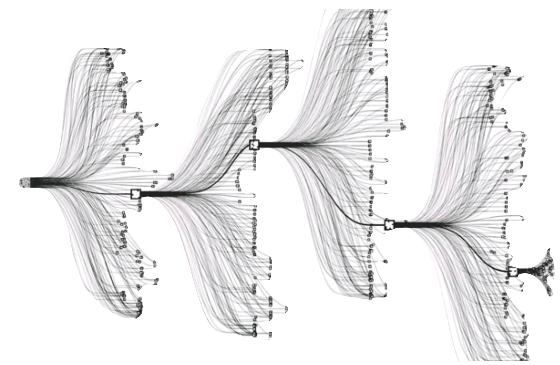
Deep Learning

- **Objective**: Build a machine that can learn from experience and understand the world as a hierarchy of concepts

Our ultimate goal is to build machines that learn from experience. This course will show the ultimate techniques (Deep Learning) to achieve this goal.

Knowledge-Based Approach

- List of all the knowledge and formal rules
  - works for games and simple systems
  - leads to a combinatorial problem
  - **not general (often we do not know the rules)**



One first approach is to encode a set of predefined rules.
It relies on the programmer having done a good analysis of
the problem and on having found the best set of rules to describe it (or its solution).
applies well to games (e.g. alpha go)
machines are good at combinatorial problems

Learning from Examples

- The machine automatically learns from examples
  - machine learning
  - no need to identify and explain rules
  - **general and flexible**

training set (data,label)

label?

new data

A second approach is not to describe the rules or the properties of the solution, but to just let data do that for us.
We only specify the objective (what we want to achieve) and let
the machine find out how to achieve that objective.
This methodology leads to Machine Learning.

Features

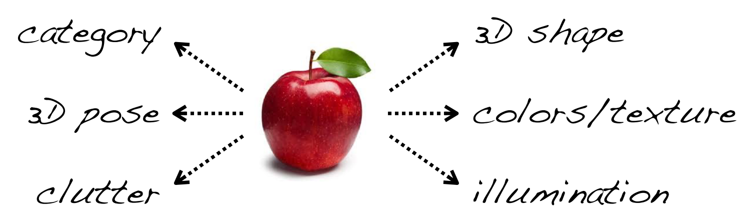
- Machines solve tasks/decisions by using the provided information (data)
- Data is often encoded into more focused relevant information (features) to simplify the decision

$$\underbrace{x}_{\text{data}} \rightarrow \underbrace{\phi(x)}_{\text{feature}}$$

- Features can be hand-made/encoded
  - <u>Operators often do not know the optimal features</u>

Representation Learning

- Features or, more in general, an **internal representation** or a hierarchy of concepts should be learned automatically
- The internal representation should separate all **factors of variation** (i.e., concepts that summarize important variation of the data)



examples of factors of variation
When analyzing a speech recording, the factors of variation include the speaker's age, their sex, their accent and the words that they are speaking. When analyzing an image of a car, the factors of variation include the position of the car, its color, and the angle and brightness of the sun.
Factors are not directly observed

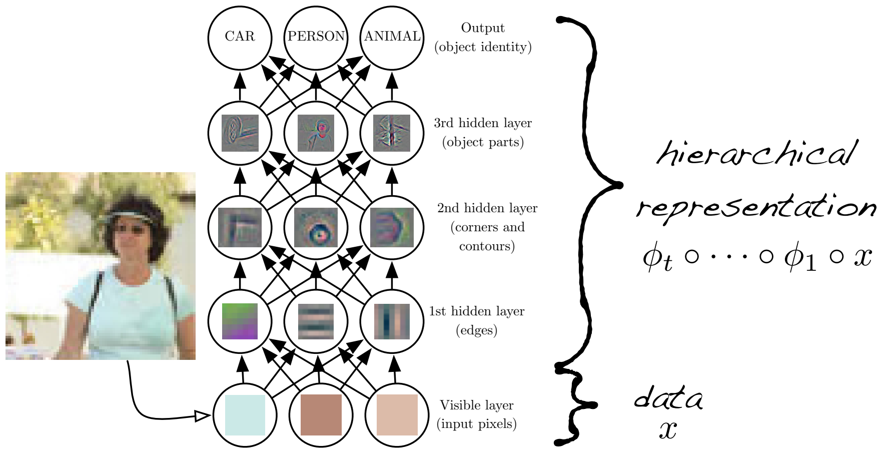
Distributed Representation

- Use many features to represent data and each feature should handle multiple data samples
- Example: Recognition of cars, trucks and birds and each can be red, green or blue
- Case #1: 1 feature for each case
  (3x3 = 9 features)
- Case #2: 3 features for identity and 3 for color
  (3+3 = 6 features)

in case1 each feature specializes in an identity and color
in case2 features generalize across identities or colors

Deep Learning

Introduces hierarchical representations
(from simple to complex, from low-level features to high-level features)



History of Vision and Machine Learning

Why do we focus on vision? There are theories that argue that vision was the fundamental trigger to evolution.

A report of the article in TIME

WHEN LIFE EXPLODED

For billions of years, simple creatures like plankton, bacteria and algae ruled the earth. Then, suddenly, life got very complicated

BY J. MADELEINE NASH

An hour later and he might not have noticed the rock, much less stooped to pick it up. But the early morning sunlight slanting across the Namibian desert in southwestern Africa happened to illuminate momentarily some strange squiggles on a chunk of sandstone. At first Douglas Erwin, a paleobiologist at the Smithsonian Institution in Washington, wondered if the meandering markings might be dried-up curls of prehistoric sea mud. But no, he decided after studying the patterns for a while, these were burrows carved by a small, wormlike creature that arose in long-vanished subtropical seas--an archaic organism that, as Erwin later confirmed, lived about 550 million years ago, just before the geological period known as the Cambrian.

As such, the innocuous-seeming creature and its curvy spoor mark the threshold of a critical interlude in the history of life. For the Cambrian is a period distinguished by the abrupt appearance of an astonishing array of multicelled animals--animals that are the ancestors of virtually all the creatures that now swim, fly and crawl through the visible world.

Indeed, while most people cling to the notion that evolution works its magic over millions of years, scientists are realizing that biological change often occurs in sudden fits and starts. And none of those fitful starts was more dramatic, more productive or more mysterious than the one that occurred shortly after Erwin's wormlike creature slithered through the primordial seas. All around the world, in layers of rock just slightly younger than that Erwin discovered, scientists have found the mineralized remains of organisms that represent the emergence of nearly every major branch in the zoological tree. Among them: bristle worms and roundworms, lamp shells and mollusks, sea cucumbers and jellyfish, not to mention an endless parade of arthropods, those spindly legged, hard-shelled ancient cousins of crabs and lobsters, spiders and flies. There are even occasional glimpses--in rock laid down not long after Erwin's Namibian sandstone--of small, ribbony swimmers with a rodlike spine that are unprepossessing progenitors of the chordate line, which leads to fish, to amphibians and eventually to humans.

Where did this extraordinary bestiary come from, and why did it emerge so quickly? In recent years, no question has stirred the imagination of more evolutionary experts, spawned more novel theories or spurred more far-flung expeditions. Life has occupied the planet for nearly 4 billion of its 4.5 billion years. But until about 600 million years ago, there were no organisms more complex than bacteria, multicelled algae and single-celled plankton. The first hint of biological ferment was a plethora of mysterious palm-shape, frondlike creatures that vanished as inexplicably as they appeared. Then, 543 million years ago, in the early Cambrian, within the span of no more than 10 million years, creatures with teeth and tentacles and claws and jaws materialized with the suddenness of apparitions. In a burst of creativity like nothing before or since, nature appears to have sketched out the blueprints for virtually the whole of the animal kingdom. This explosion of biological diversity is described by scientists as biology's Big Bang.

Over the decades, evolutionary theorists beginning with Charles Darwin have tried to argue that the appearance of multicelled animals during the Cambrian merely seemed sudden, and in fact had been preceded by a lengthy period of evolution for which the geological record was missing. But this explanation, while it patched over a hole in an otherwise masterly theory, now seems increasingly unsatisfactory. Since 1987, discoveries of major fossil beds in Greenland, in China, in Siberia, and now in Namibia have shown that the period of biological innovation occurred at virtually the same instant in geologic time all around the world.

What could possibly have powered such a radical advance? Was it something in the organisms themselves or the environment in which they lived? Today an unprecedented effort to answer these questions is under way. Geologists and geochemists are reconstructing the Precambrian planet, looking for changes in the atmosphere and ocean that might have put evolution into sudden overdrive. Developmental biologists are teasing apart the genetic toolbox needed to assemble animals as disparate as worms and flies, mice and fish. And paleontologists are exploring deeper reaches of the fossil record, searching for organisms that might have primed the evolutionary pump. "We're getting data," says Harvard University paleontologist Andrew Knoll, "almost faster than we can digest it."

Every few weeks, it seems, a new piece of the puzzle falls into place. Just last month, in an article published by the journal Nature, an international team of scientists reported finding the exquisitely preserved remains of a 1-in.- to 2-in.-long animal that flourished in the Cambrian oceans 525 million years ago. From its flexible but sturdy spinal rod, the scientists deduced that this animal--dubbed Yunnanozoon lividum, after the Chinese province in which it was found--was a primitive chordate, the oldest ancestor yet discovered of the vertebrate branch of the animal kingdom, which includes Homo sapiens.

Even more tantalizing, paleontologists are gleaning insights into the enigmatic years that immediately preceded the Cambrian explosion. Until last spring, when John Grotzinger, a sedimentologist from M.I.T., led Erwin and two dozen other scientists on an expedition to the Namibian desert, this fateful period was obscured by a 20 million--year gap in the fossil record. But with the find in Namibia, as Grotzinger and three colleagues reported in the Oct. 27 issue of Science, the gap suddenly filled with complex life. In layer after layer of late Precambrian rock, heaved up in the rugged outcroppings the Namibians call kopfs (after the German word for "head"), Grotzinger's team has documented the existence of a flourishing biological community on the cusp of a startling transformation, a community in which small wormlike somethings, small shelly somethings--perhaps even large frondlike somethings--were in the process of crossing over a shadow line into uninhabited ecospace.

Here, then, are highlights from the tale that scientists are piecing together of a unique and dynamic time in the history of the earth, when continents were rifting apart, genetic programs were in flux, and tiny organisms in vast oceans dreamed of growing large.

THE WEIRD WONDERS

Inside locked cabinets at the Smithsonian Institution nestle snapshots in stone as vivid as any photograph. There, engraved on slices of ink-black shale, are the myriad inhabitants of a vanished world, from plump Aysheaia prancing on caterpillar-like legs to crafty Ottoia, lurking in a burrow and extending its predatory proboscis. Excavated in the early 1900s from a geological formation in the Canadian Rockies known as the Burgess Shale, these relics of the earliest animals to appear on earth are now revered as priceless treasures. Yet for half a century after their discovery, the Burgess Shale fossils attracted little scientific attention as researchers concentrated on creatures that were larger and easier to understand--like the dinosaurs that roamed the earth nearly 300 million years later.

Then, starting in the late 1960s, three paleontologists--Harry Whittington of the University of Cambridge in England and his two students, Derek Briggs and Simon Conway Morris--embarked on a methodical re-examination of the Burgess Shale fossils. Under bright lights and powerful microscopes, they coaxed fine-grain anatomical detail from the shale's stony secrets: the remains of small but substantial animals that were overtaken by a roaring underwater mudslide 515 million years ago and swept into water so deep and oxygen-free that the bacteria that should have decayed their tissues couldn't survive. Preserved were not just the hard-shelled creatures familiar to Darwin and his contemporaries but also the fossilized remains of soft-bodied beasts like Aysheaia and Ottoia. More astonishing still were remnants of delicate interior structures, like Ottoia's gut with its last, partly digested meal.

Soon, inspired reconstructions of the Cambrian bestiary began to create a stir at paleontological gatherings. Startled laughter greeted the unveiling of oddball Opabinia, with its five eyes and fire-hose-like proboscis. Credibility was strained by Hallucigenia, when Conway Morris depicted it as dancing along on needle-sharp legs, and also by Wiwaxia, a whimsical armored slug with two rows of upright scales. And then there was Anomalocaris, a fearsome predator that caught its victims with spiny appendages and crushed them between jaws that closed like the shutter of a camera. "Weird wonders," Harvard University paleontologist Stephen Jay Gould called them in his 1989 book, Wonderful Life, which celebrated the strangeness of the Burgess Shale animals.

But even as Wonderful Life was being published, the discovery of new Cambrian-era fossil beds in Sirius Passet, Greenland, and Yunnan, China, was stripping some of the weirdness from the wonders. Hallucigenia's impossibly pointed legs, for example, were unmasked as the upside-down spines of a prehistoric velvet worm. In similar fashion, Wiwaxia, some scientists think, is probably allied with living bristle worms. And the anomalocaridids--whose variety is rapidly expanding with further research--appear to be cousins, if not sisters, of the amazingly diverse arthropods.

The real marvel, says Conway Morris, is how familiar so many of these animals seem. For it was during the Cambrian (and perhaps only during the Cambrian) that nature invented the animal body plans that define the broad biological groupings known as phyla, which encompass everything from classes and orders to families, genera and species. For example, the chordate phylum includes mammals, birds and fish. The class Mammalia, in turn, covers the primate order, the hominid family, the genus Homo and our own species, Homo sapiens.

EVOLVING AT SUPERSONIC SPEED

Scientists used to think that the evolution of phyla took place over a period of 75 million years, and even that seemed impossibly short. Then two years ago, a group of researchers led by Grotzinger, Samuel Bowring from M.I.T. and Harvard's Knoll took this long-standing problem and escalated it into a crisis. First they recalibrated the geological clock, chopping the Cambrian period to about half its former length. Then they announced that the interval of major evolutionary innovation did not span the entire 30 million years, but rather was concentrated in the first third. "Fast," Harvard's Gould observes, "is now a lot faster than we thought, and that's extraordinarily interesting."

What Knoll, Grotzinger and colleagues had done was travel to a remote region of northeastern Siberia where millenniums of relentless erosion had uncovered a dramatic ledger of rock more than half a mile thick. In ancient seabeds near the mouth of the Lena River, they spotted numerous small, shelly fossils characteristic of the early Cambrian. Even better, they found cobbles of volcanic ash containing minuscule crystals of a mineral known as zircon, possibly the most sensitive timepiece nature has yet invented.

Zircon dating, which calculates a fossil's age by measuring the relative amounts of uranium and lead within the crystals, had been whittling away at the Cambrian for some time. By 1990, for example, new dates obtained from early Cambrian sites around the world were telescoping the start of biology's Big Bang from 600 million years ago to less than 560 million years ago. Now, with information based on the lead content of zircons from Siberia, virtually everyone agrees that the Cambrian started almost exactly 543 million years ago and, even more startling, that all but one of the phyla in the fossil record appeared within the first 5 million to 10 million years. "We now know how fast fast is," grins Bowring. "And what I like to ask my biologist friends is, How fast can evolution get before they start feeling uncomfortable?"

FREAKS OR ANCESTORS?

The key to the Cambrian explosion, researchers are now convinced, lies in the Vendian, the geological period that immediately preceded it. But because of the frustrating gap in the fossil record, efforts to explore this critical time interval have been hampered. For this reason, no one knows quite what to make of the singular frond-shape organisms that appeared tens of millions of years before the beginning of the Cambrian, then seemingly died out. Are these puzzling life-forms--which Yale University paleobiologist Adolf Seilacher dubbed the "vendobionts"--linked somehow to the creatures that appeared later on, or do they represent a totally separate chapter in the history of life?

Seilacher has energetically championed the latter explanation, speculating that the vendobionts represent a radically different architectural solution to the problem of growing large. These "creatures"--which reached an adult size of 3 ft. or more across--did not divide their bodies into cells, believes Seilacher, but into compartments so plumped with protoplasm that they resembled air mattresses. They appear to have had no predators, says Seilacher, and led a placid existence on the ocean floor, absorbing nutrients from seawater or manufacturing them with the help of symbiotic bacteria.

UCLA paleontologist Bruce Runnegar, however, disagrees with Seilacher. Runnegar argues that the fossil known as Ernietta, which resembles a pouch made of wide-wale corduroy, may be some sort of seaweed that generated food through photosynthesis. Charniodiscus, a frond with a disklike base, he classifies as a colonial cnidarian, the phylum that includes jellyfish, sea anemones and sea pens. And Dickinsonia, which appears to have a clearly segmented body, Runnegar tentatively places in an ancestral group that later gave rise to roundworms and arthropods. The Cambrian explosion did not erupt out of the blue, argues Runnegar. "It's the continuation of a process that began long before."

The debate between Runnegar and Seilacher is about to get even more heated. For, as pictures that accompany the Science article reveal, researchers have returned from Namibia with hard evidence that a diverse community of organisms flourished in the oceans at the end of the Vendian, just before nature was gripped by creative frenzy. Runnegar, for instance, is currently studying the fossil of a puzzling conical creature that appears to be an early sponge. M.I.T.'s Beverly Saylor is sorting through sandstones that contain a menagerie of small, shelly things, some shaped like wine goblets, others like miniature curtain rods. And Guy Narbonne of Queen's University in Ontario, Canada, is trying to make sense of Dickinsonia-like creatures found just beneath the layer of rock where the Cambrian officially begins.

What used to be a gap in the fossil record has turned out to be teeming with life, and this single, stunning insight into late-Precambrian ecology, believes Grotzinger, is bound to reframe the old argument over the vendobionts. For whether they are animal ancestors or evolutionary dead ends, says Grotzinger, Dickinsonia and its cousins can no longer be thought of as sideshow freaks. Along with the multitudes of small, shelly organisms and enigmatic burrowers that riddled the sea floor with tunnels and trails, the vendobionts have emerged as important clues to the Cambrian explosion. "We now know," says Grotzinger, "that evolution did not proceed in two unrelated pulses but in two pulses that beat together as one."

BREAKING THROUGH THE ALGAE

To human eyes, the world on the eve of the Cambrian explosion would have seemed an exceedingly hostile place. Tectonic forces unleashed huge earthquakes that broke continental land masses apart, then slammed them back together. Mountains the size of the Himalayas shot skyward, hurling avalanches of rock, sand and mud down their flanks. The climate was in turmoil. Great ice ages came and went as the chemistry of the atmosphere and oceans endured some of the most spectacular shifts in the planet's history. And in one way or another, says Knoll, these dramatic upheavals helped midwife complex animal life by infusing the primordial oceans with oxygen.

Without oxygen to aerate tissues and make vital structural components like collagen, notes Knoll, animals simply cannot grow large. But for most of earth's history, the production of oxygen through photosynthesis--the metabolic alchemy that allowed primordial algae to turn carbon dioxide, water and sunlight into energy-- was almost perfectly balanced by oxygen-depleting processes, especially organic decay. Indeed, the vast populations of algae that smothered the Precambrian oceans generated tons of vegetative debris, and as bacteria decomposed this slimy detritus, they performed photosynthesis in reverse, consuming oxygen and releasing carbon dioxide, the greenhouse gas that traps heat and helps warm the planet.

For oxygen to rise, then, the planet's burden of decaying organic matter had to decline. And around 600 million years ago, that appears to be what happened. The change is reflected in the chemical composition of rocks like limestone, which incorporate two isotopes of carbon in proportion to their abundance in seawater--carbon 12, which is preferentially taken up by algae during photosynthesis, and carbon 13, its slightly heavier cousin. By sampling ancient limestones, Knoll and his colleagues have determined that the ratio of carbon 12 to carbon 13 remained stable for most of the Proterozoic Eon, a boggling expanse of time that stretched from 2.5 billion years ago to the end of the Vendian. But at the close of the Proterozoic, just prior to the Cambrian explosion, they pick up a dramatic rise in carbon 13 levels, suggesting that carbon 12 in the form of organic material was being removed from the oceans.

One mechanism, speculates Knoll, could have been erosion from steep mountain slopes. Over time, he notes, tons of sediment and rock that poured into the sea could have buried algal remains that fell to the sea floor. In addition, he says, rifting continents very likely changed the geometry of ocean basins so that water could not circulate as vigorously as before. The organic carbon that fell to the sea floor, then, would have stayed there, never cycling back to the ocean surface and into the atmosphere. As levels of atmospheric carbon dioxide dropped, the earth would have cooled. Sure enough, says Knoll, a major ice age ensued around 600 million years ago--yet another link in a complex chain that connects geological and geochemical events to a momentous advance in biology.

Biology also influenced geochemistry, says Indiana University biochemist John Hayes. In fact, in a paper published in Nature earlier this year, Hayes and his colleagues argue that guts, those simple conduits that take food in at one end and expel wastes at the other, may be the key to the Cambrian explosion. Their reasoning goes something like this: animals grazed on the algae, packaging the leftover organic material into fecal pellets. These pellets dropped to the ocean depths, depriving oxygen-depleting bacteria of their principal food source. The evidence? Organic lipids in ancient rocks, notes Hayes, underwent a striking change in carbon-isotope ratios around 550 million years ago. Again, the change suggests that food sources rich in carbon 12, like algae, were being "express mailed" to the ocean floor.

THE GENETIC TOOL KIT

The animals that aerated the precambrian oceans could have resembled the wormlike something that left its meandering marks on the rock Erwin lugged back from Namibia. More advanced than a flatworm, which was not rigid enough to burrow through sand, this creature would have had a sturdy, fluid-filled body cavity. It would have had musculature capable of strong contractions. It probably had a heart, a well-defined head with an eye for sensing light and, last but not least, a gastrointestinal tract with an opening at each end. What kind of genetic machinery, Erwin wondered, did nature need in order to patch together such a creature?

Over the summer, Erwin pondered this problem with two paleontologist friends, David Jablonski of the University of Chicago and James Valentine of the University of California, Berkeley. Primitive multicelled organisms like jellyfish, they reasoned, have three so-called homeotic homeobox genes, or Hox genes, which serve as the master controllers of embryonic development. Flatworms have four, arthropods like fruit flies have eight, and the primitive chordate Branchiostoma (formerly known as Amphioxus) has 10. So around 550 million years ago, Erwin and the others believe, some wormlike creature expanded its Hox cluster, bringing the number of genes up to six. Then, "Boom!" shouts Jablonski. "At that point, perhaps, life crossed some sort of critical threshold." Result: the Cambrian explosion.

The proliferation of wildly varying body plans during the Cambrian, scientists reason, therefore must have something to do with Hox genes. But what? To find out, developmental biologist Sean Carroll's lab on the University of Wisconsin's Madison campus has begun importing tiny velvet worms that inhabit rotting logs in the dry forests of Australia. Blowing bubbles of spittle and waving their fat legs in the air, they look, he marvels, virtually identical to their Cambrian cousin Aysheaia, whose evocative portrait appears in the pages of the Burgess Shale. Soon Carroll hopes to answer a pivotal question: Is the genetic tool kit needed to construct a velvet worm smaller than the one the arthropods use? Already Carroll suspects that the Cambrian explosion was powered by more than a simple expansion in the number of Hox genes. Far more important, he believes, were changes in the vast regulatory networks that link each Hox gene to hundreds of other genes. Think of these genes, suggests Carroll, as the chips that run a computer. The Cambrian explosion, then, may mark not the invention of new hardware, but rather the elaboration of new software that allowed existing genes to perform new tricks. Unusual-looking arthropods, for example, might be cobbled together through variations of the genetic software that codes for legs. "Arthropods," observes paleoentomologist Jarmila Kukalova-Peck of Canada's Carleton University, "are all legs"--including the "legs" that evolved into jaws, claws and even sex organs.

BEYOND DARWINISM

Of course, understanding what made the Cambrian explosion possible doesn't address the larger question of what made it happen so fast. Here scientists delicately slide across data-thin ice, suggesting scenarios that are based on intuition rather than solid evidence. One favorite is the so-called empty barrel, or open spaces, hypothesis, which compares the Cambrian organisms to homesteaders on the prairies. The biosphere in which the Cambrian explosion occurred, in other words, was like the American West, a huge tract of vacant property that suddenly opened up for settlement. After the initial land rush subsided, it became more and more difficult for naive newcomers to establish footholds.

Predation is another popular explanation. Once multicelled grazers appeared, say paleontologists, it was only a matter of time before multicelled predators evolved to eat them. And, right on cue, the first signs of predation appear in the fossil record exactly at the transition between the Vendian and the Cambrian, in the form of bore holes drilled through shelly organisms that resemble stacks of miniature ice-cream cones. Seilacher, among others, speculates that the appearance of protective shells and hard, sharp parts in the late Precambrian signaled the start of a biological arms race that did in the poor, defenseless vendobionts.

Even more speculative are scientists' attempts to address the flip side of the Cambrian mystery: why this evolutionary burst, so stunning in speed and scope, has never been equaled. With just one possible exception--the Bryozoa, whose first traces turn up shortly after the Cambrian--there is no record of new phyla emerging later on, not even in the wake of the mass extinction that occurred 250 million years ago, at the end of the Permian period.

Why no new phyla? Some scientists suggest that the evolutionary barrel still contained plenty of organisms that could quickly diversify and fill all available ecological niches. Others, however, believe that in the surviving organisms, the genetic software that controls early development had become too inflexible to create new life-forms after the Permian extinction. The intricate networks of developmental genes were not so rigid as to forbid elaborate tinkering with details; otherwise, marvels like winged flight and the human brain could never have arisen. But very early on, some developmental biologists believe, the linkages between multiple genes made it difficult to change important features without lethal effect. "There must be limits to change," says Indiana University developmental biologist Rudolf Raff. "After all, we've had these same old body plans for half a billion years."

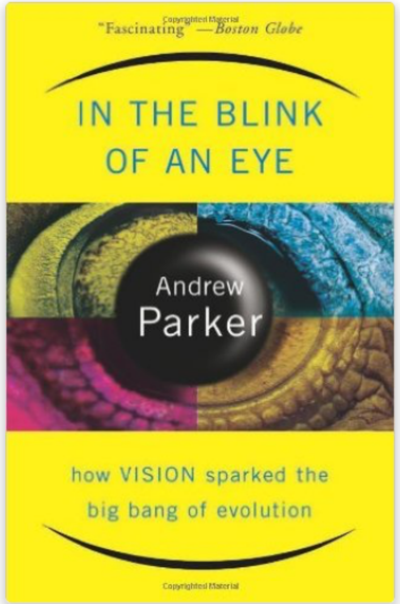
The more scientists struggle to explain the Cambrian explosion, the more singular it seems. And just as the peculiar behavior of light forced physicists to conclude that Newton's laws were incomplete, so the Cambrian explosion has caused experts to wonder if the twin Darwinian imperatives of genetic variation and natural selection provide an adequate framework for understanding evolution. "What Darwin described in the Origin of Species," observes Queen's University paleontologist Narbonne, "was the steady background kind of evolution. But there also seems to be a non-Darwinian kind of evolution that functions over extremely short time periods--and that's where all the action is."

In a new book, At Home in the Universe (Oxford University Press; $25), theoretical biologist Stuart Kauffman of the Santa Fe Institute argues that underlying the creative commotion during the Cambrian are laws that we have only dimly glimpsed--laws that govern not just biological evolution but also the evolution of physical, chemical and technological systems. The fanciful animals that first appeared on nature's sketchpad remind Kauffman of early bicycles, with their odd-size wheels and strangely angled handlebars. "Soon after a major innovation," he writes, "discovery of profoundly different variations is easy. Later innovation is limited to modest improvements on increasingly optimized designs."

Biological evolution, says Kauffman, is just one example of a self-organizing system that teeter-totters on the knife edge between order and chaos, "a grand compromise between structure and surprise." Too much order makes change impossible; too much chaos and there can be no continuity. But since balancing acts are necessarily precarious, even the most adroit tightrope walkers sometimes make one move too many. Mass extinctions, chaos theory suggests, do not require comets or volcanoes to trigger them. They arise naturally from the intrinsic instability of the evolving system, and superior fitness provides no safety net.

In fact, some of prehistory's worst mass extinctions took place during the Cambrian itself, and they probably occurred for no obvious reason. Rather, just as the tiniest touch can cause a steeply angled sand pile to slide, so may a small evolutionary advance that gives one species a temporary advantage over another be enough to bring down an entire ecosystem. "These patterns of speciations and extinctions, avalanching across ecosystems and time," warns Kauffman, are to be found in every chaotic system--human and biological. "We are all part of the same pageant," as he puts it. Thus, even in this technological age, we may have more in common than we care to believe with the weird--and ultimately doomed--wonders that radiated so hopefully out of the Cambrian explosion.

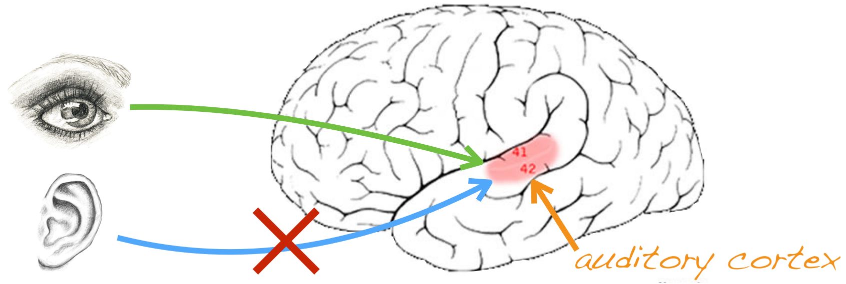
- **600M years ago**: bacteria, multicelled algae and single-celled plankton
- **543M years ago** (early Cambrian): within 10 million years, creatures with teeth and tentacles and claws and jaws suddenly materialized - **Biology's Big Bang**
- Simultaneously all around the world
- Andrew Parker argues that **vision** caused this big bang



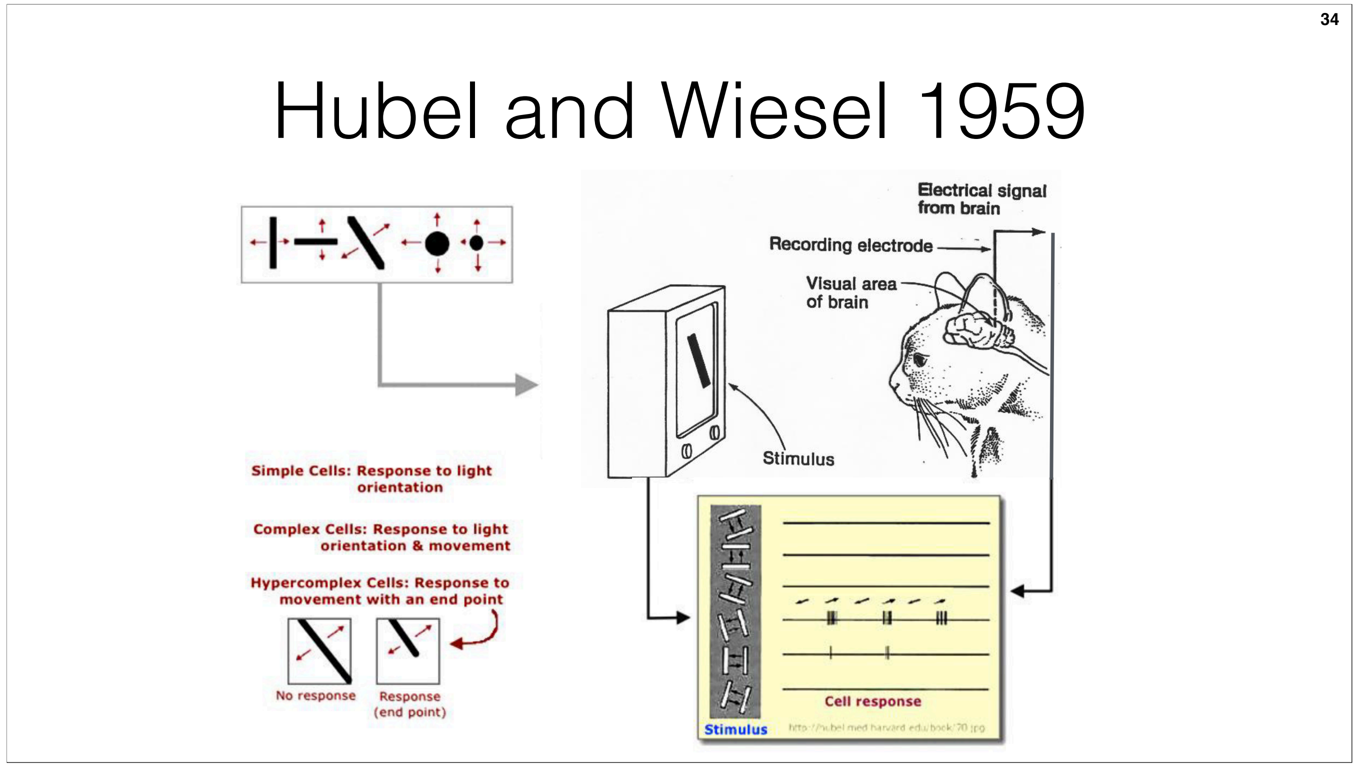
Oxford zoologist Andrew Parker reveals his theory of the great flourishing of life during the Cambrian. Parker's controversial but increasingly accepted "Light Switch Theory" holds that it was the development of vision in primitive animals that caused the explosion.

The Brain

- Neural Networks are inspired by the brain
- The “One Learning Algorithm” hypothesis
- Experiments on animals have shown that rewiring the optical nerves to either the auditory or the somatosensory cortex the animal learns to “see”



Besides vision, another important component we focus on is the brain.
The brain is an example of a high-performance machine learning algorithm. It is used as a guiding example of what can be achieved. Indeed neural networks take a lot of inspiration from the brain.
The brain can handle sound, vision, touch etc
Experiments have shown that if we remove the connection between the auditory system and the corresponding cortex and instead rewire the optic nerves to such part of the cortex, the brain readapts and learns to see. The same applies for the somatosensory cortex (related to touch). It shows that the brain is a sort of universal learning algorithm capable of dealing with any input patterns and eventually learn to distinguish and recognize them.
So is it the case that there is no need for parts of the brain that are built to serve a specific sense? If we design machines by imitating the brain, is it the case that we only need to design a single learning algorithm and then the senses just provide different patterns, whose nature is not relevant to learning?



How did we understand how the brain makes sense of vision?
The Nobel prizes Hubel and Wiesel determined the primary visual area of the cortex and that it was involved in the processing of low-level features line edges.

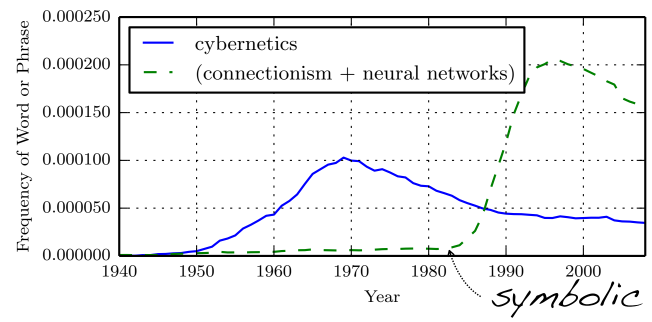
Their discovery was almost by chance as they recorded cell
responses only when they were changing the slides.
The edge of the slide (rather than the picture in the slide)
was responsible for the triggering signal.

They further introduced the idea of a hierarchical neuronal structure (deconstructing images and then reassembling details to create scene)
Dr. Hubel’s and Dr. Wiesel’s work further showed that sensory deprivation early in life can permanently alter the brain’s ability to process images. Their findings led to a better understanding of how to treat certain visual birth defects. Just exposure to light is not sufficient. Light patterns are necessary.

A Bit of History

- 1940s-1960s: Cybernetics
- 1980s-1990s: Connectionism
- 2006-now: Deep learning

artificial neural networks



symbolic reasoning

During the early 1980s, most cognitive scientists studied models of symbolic reasoning. Despite their popularity, symbolic models were difficult to explain in terms of how the brain could actually implement them using neurons. The connectionists began to study models of cognition that could actually be grounded in neural implementations

A Bit of History

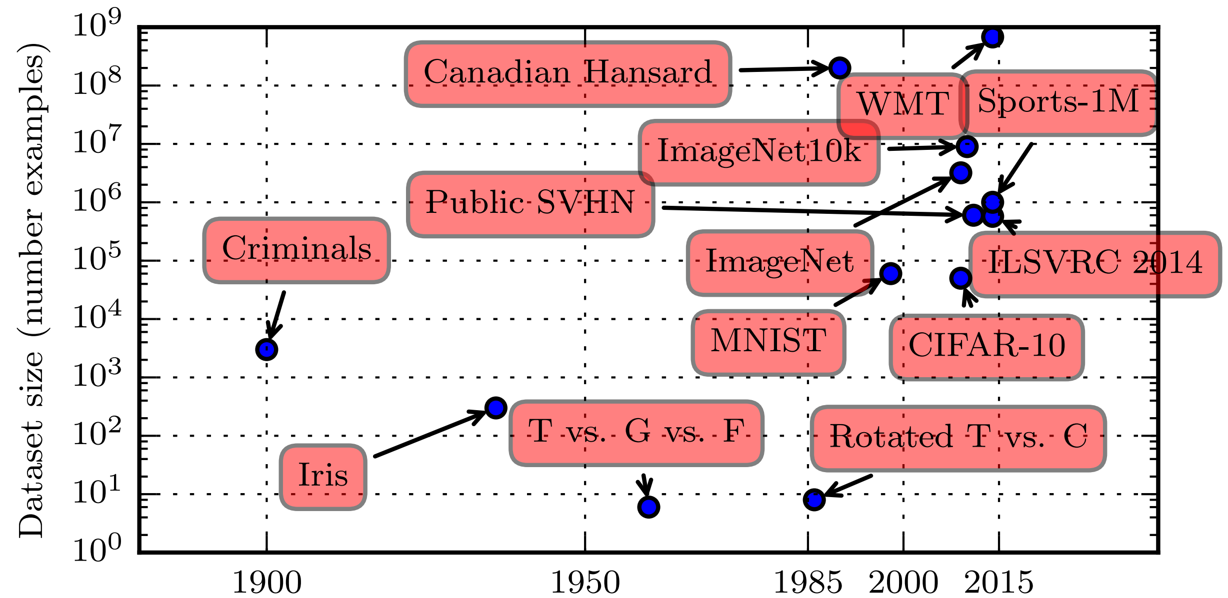
- AI Winter (mid-1990s)
  - Unrealistic claims were made and unfulfilled
  - Other Machine Learning areas advanced (e.g., kernel machines and graphical models)
  - Lack of understanding of how/why neural networks work
  - Funding and popularity decline

A Bit of History

- Concurrence of three events led to a revival

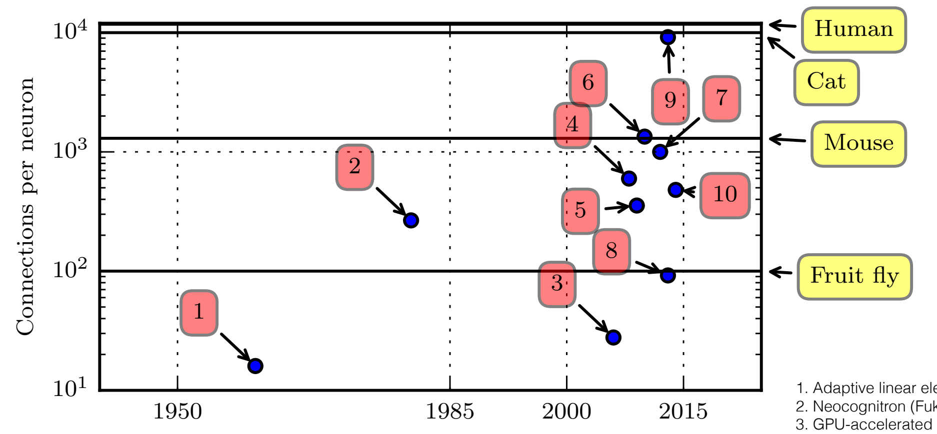
1. More data (millions of examples - e.g., ILSVRC)
2. More computing power (GPUs)
3. More efficient and deeper networks (e.g., convolutional NN) and better training algorithms (although still based on backpropagation)

More Data



Dataset sizes have increased greatly over time. In the early 1900s, statisticians studied datasets using hundreds or thousands of manually compiled measurements (Garson, 1900; Gosset, 1908; Anderson, 1935; Fisher, 1936). In the 1950s through 1980s, the pioneers of biologically inspired machine learning often worked with small, synthetic datasets, such as low-resolution bitmaps of letters, that were designed to incur low computational cost and demonstrate that neural networks were able to learn specific kinds of functions (Widrow and Hoff, 1960; Rumelhart et al., 1986b). In the 1980s and 1990s, machine learning became more statistical in nature and began to leverage larger datasets containing tens of thousands of examples such as the MNIST dataset (shown in figure 1.9) of scans of handwritten numbers (LeCun et al., 1998b). In the first decade of the 2000s, more sophisticated datasets of this same size, such as the CIFAR-10 dataset (Krizhevsky and Hinton, 2009) continued to be produced. Toward the end of that decade and throughout the first half of the 2010s, significantly larger datasets, containing hundreds of thousands to tens of millions of examples, completely changed what was possible with deep learning. These datasets included the public Street View House Numbers dataset (Netzer et al., 2011), various versions of the ImageNet dataset (Deng et al., 2009, 2010a; Russakovsky et al., 2014a), and the Sports-1M dataset (Karpathy et al., 2014). At the top of the graph, we see that datasets of translated sentences, such as IBM's dataset constructed from the Canadian Hansard (Brown et al., 1990) and the WMT 2014 English to French dataset (Schwenk, 2014) are typically far ahead of other dataset sizes.

Better and Deeper Models

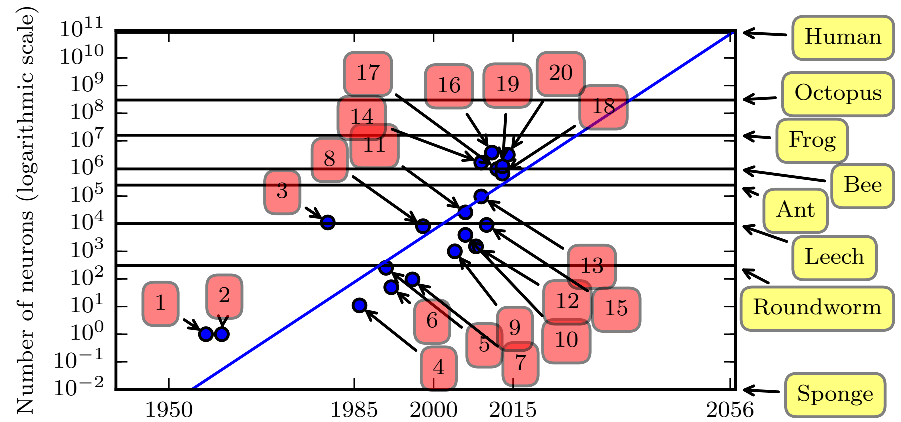


1. Adaptive linear element (Widrow and Hoff, 1960)
2. Neocognitron (Fukushima, 1980)
3. GPU-accelerated convolutional network (Chellapilla et al., 2006)
4. Deep Boltzmann machine (Salakhutdinov and Hinton, 2009a)
5. Unsupervised convolutional network (Jarrett et al., 2009)
6. GPU-accelerated multilayer perceptron (Ciresan et al., 2010)
7. Distributed autoencoder (Le et al., 2012)
8. Multi-GPU convolutional network (Krizhevsky et al., 2012)
9. COTS HPC unsupervised convolutional network (Coates et al., 2013)
10. GoogLeNet (Szegedy et al., 2014a)

COTS-HPC: Commodity Off-The-Shelf High Performance Computing (COTS HPC) technology: a cluster of GPU servers with Infiniband interconnects and MPI. This system is able to train 1 billion parameter networks on just 3 machines in a couple of days

Better and Deeper Models

Complexity of biological and artificial neurons considered the same;
biological neural networks may be even larger than this plot portrays



number of neurons matches that of humans in 2050

1. Perceptron (Rosenblatt, 1958, 1962)
2. Adaptive linear element (Widrow and Hoff, 1960)
3. Neocognitron (Fukushima, 1980)
4. Early back-propagation network (Rumelhart et al., 1986b)
5. Recurrent neural network for speech recognition (Robinson and Fallside, 1991)
6. Multilayer perceptron for speech recognition (Bengio et al., 1991)
7. Mean field sigmoid belief network (Saul et al., 1996)
8. LeNet-5 (LeCun et al., 1998b)
9. Echo state network (Jaeger and Haas, 2004)
10. Deep belief network (Hinton et al., 2006)
11. GPU-accelerated convolutional network (Chellapilla et al., 2006)
12. Deep Boltzmann machine (Salakhutdinov and Hinton, 2009a)
13. GPU-accelerated deep belief network (Raina et al., 2009)
14. Unsupervised convolutional network (Jarrett et al., 2009)
15. GPU-accelerated multilayer perceptron (Ciresan et al., 2010)
16. OMP-1 network (Coates and Ng, 2011)
17. Distributed autoencoder (Le et al., 2012)
18. Multi-GPU convolutional network (Krizhevsky et al., 2012)
19. COTS HPC unsupervised convolutional network (Coates et al., 2013)

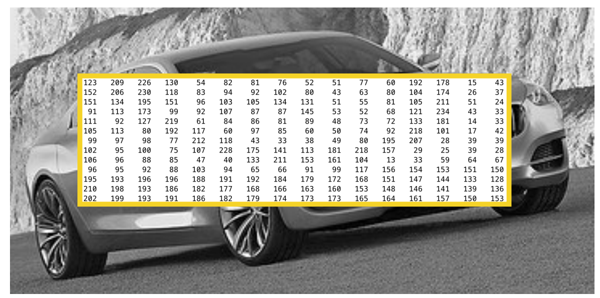
20. GoogLeNet (Szegedy et al., 2014a)

Recent Trends

- **Code-sharing**
  - Deep Learning libraries: Caffe, Torch/Lua, TensorFlow, MatConvNet, Theano, PyLearn2, DistBelief, MXNet, etc
  - Deep Learning models in GitHub
- **Fast dissemination of work**
  - arxiv (not peer-reviewed, low-quality writing, but cuts dissemination time by 6 months)

Focus: Computer Vision

- Computer Vision is about extracting high-level information from visual data (e.g., images, videos)
- Why is it so difficult for a machine?



this is what
the computer sees

The state of Computer Vision and AI: we are really, really far away.

Oct 22, 2012

The picture above is funny.

Andrej Karpathy blog

But for me it is also one of those examples that make me sad about the outlook for AI and for Computer Vision. What would it take for a computer to understand this image as you or I do? I challenge you to think explicitly of all the pieces of knowledge that have to fall in place for it to make sense. Here is my short attempt:

You recognize it is an image of a bunch of people and you understand they are in a hallway
You recognize that there are 3 mirrors in the scene so some of those people are “fake” replicas from different viewpoints.
You recognize Obama from the few pixels that make up his face. It helps that he is in his suit and that he is surrounded by other people with suits.
You recognize that there’s a person standing on a scale, even though the scale occupies only very few white pixels that blend with the background. But, you’ve used the person’s pose and knowledge of how people interact with objects to figure it out.
You recognize that Obama has his foot positioned just slightly on top of the scale. Notice the language I’m using: It is in terms of the 3D structure of the scene, not the position of the leg in the 2D coordinate system of the image.
You know how physics works: Obama is leaning in on the scale, which applies a force on it. Scale measures force that is applied on it, that’s how it works => it will over-estimate the weight of the person standing on it.
The person measuring his weight is not aware of Obama doing this. You derive this because you know his pose, you understand that the field of view of a person is finite, and you understand that he is not very likely to sense the slight push of Obama’s foot.
You understand that people are self-conscious about their weight. You also understand that he is reading off the scale measurement, and that shortly the over-estimated weight will confuse him because it will probably be much higher than what he expects. In other words, you reason about implications of the events that are about to unfold seconds after this photo was taken, and especially about the thoughts and how they will develop inside people’s heads. You also reason about what pieces of information are available to people.
There are people in the back who find the person’s imminent confusion funny. In other words you are reasoning about state of mind of people, and their view of the state of mind of another person. That’s getting frighteningly meta.

Finally, the fact that the perpetrator here is the president makes it maybe even a little funnier. You understand what actions are more or less likely to be undertaken by different people based on their status and identity.

The state of Computer Vision and AI: we are really, really far away.

- We need just a brief glance at a 2D array of RGB values
- Pixel values are just a tip of a huge iceberg
- Deriving the entire shape and size of the iceberg is the most difficult task ahead of us
- How about an algorithm that can reason about the scene like I did?
- How do we even begin to gather data that can support these inferences (for example how a scale works)?
- How do we go about even giving the computer a chance?

Andrej Karpathy blog

It is mind-boggling that all of the above inferences unfold from a brief glance at a 2D array of R,G,B values. The core issue is that the pixel values are just a tip of a huge iceberg and deriving the entire shape and size of the icerberg from prior knowledge is the most difficult task ahead of us. How can we even begin to go about writing an algorithm that can reason about the scene like I did? Forget for a moment the inference algorithm that is capable of putting all of this together; How do we even begin to gather data that can support these inferences (for example how a scale works)? How do we go about even giving the computer a chance?

Data in Vision-Related Tasks

Datasets and Challenges

- Circa 2001: five categories, hundreds of images per category
- Circa 2004: 101 categories, 40 to 800 images per category
- Circa 2006: 256 categories, 30607 images
- Today: up to thousands of categories, millions of images

*exponential growth

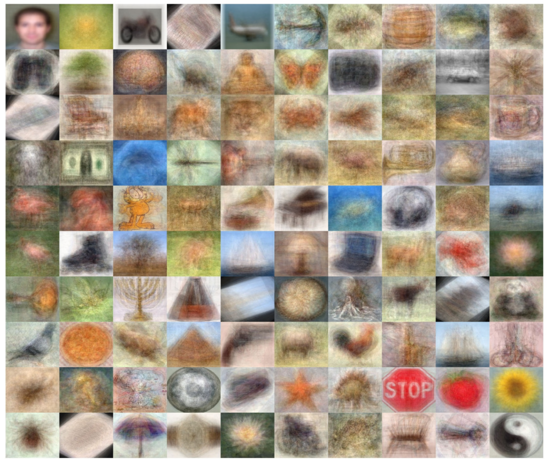
Caltech 101 & 256

Fei-Fei, Fergus, Perona, 2004

Griffin, Holub, Perona, 2007

http://www.vision.caltech.edu/Image_Datasets/Caltech101/
http://www.vision.caltech.edu/Image_Datasets/Caltech256/

Caltech 101
intra-class variability



per class averages

too aligned/object centric. classification is easy and unrealistic when applied to real images.

The PASCAL Visual Object Challenge 2005-2012

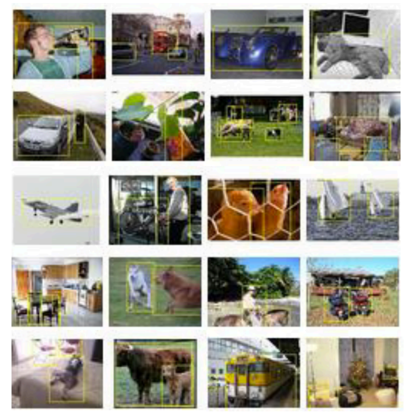
- 20 classes:
  Person: person
  Animal: bird, cat, cow, dog, horse, sheep
  Vehicle: aeroplane, bicycle, boat, bus, car, motorbike, train
  Indoor: bottle, chair, dining table, potted plant, sofa, tv/monitor

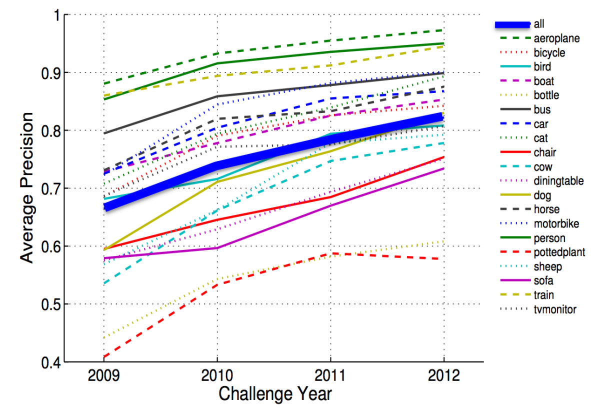
- Dataset size (by 2012):
  11.5K training/validation images
  27K bounding boxes
  7K segmentations

http://pascallin.ecs.soton.ac.uk/challenges/VOC/

The PASCAL Visual Object Challenge 2005-2012

Classification: For each of the twenty classes, predict presence/absence of an example of that class in the test image
Detection: Predict the bounding box and label of each object from the twenty target classes in the test image

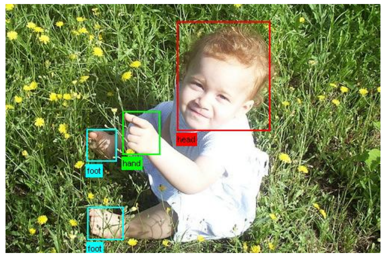




PASCAL competitions

- **Segmentation**: Generating pixel-wise segmentations giving the class of the object visible at each pixel, or "background" otherwise

- **Person layout**: Predicting the bounding box and label of each part of a person (head, hands, feet)



PASCAL competitions

- **Action/Pose classification** 10 classes

ImageNet - ILSVRC

IMAGENET

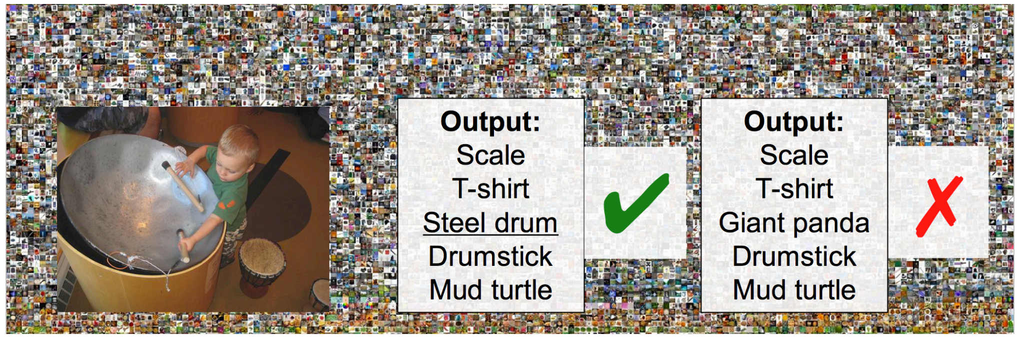
www.image-net.org

22K categories and **14M** images

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

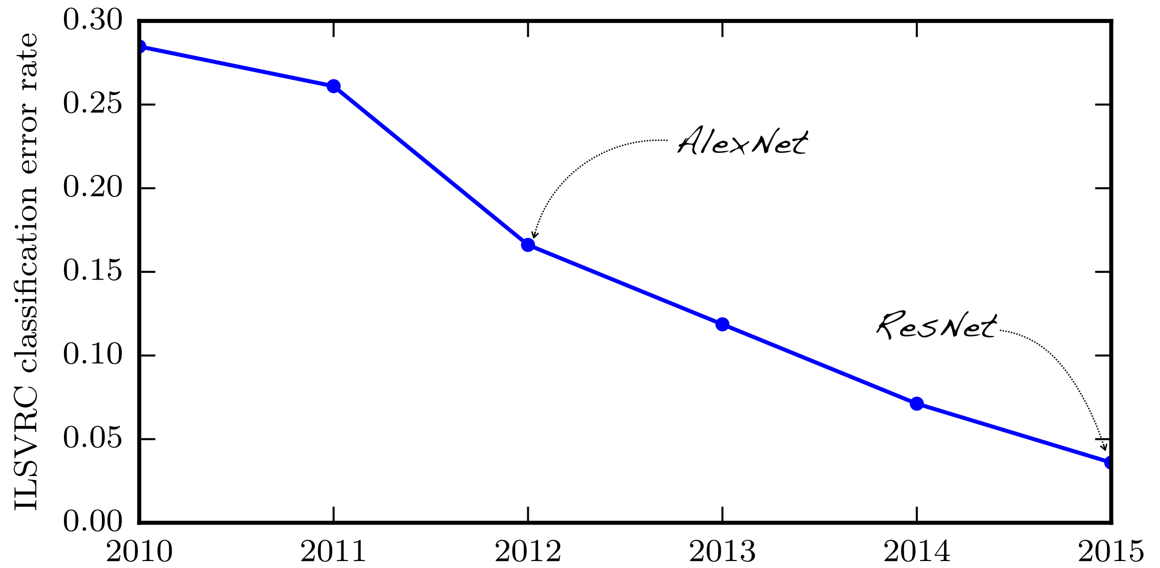
Datasets and Challenges

- The ILSVRC object classification challenge:
  1K categories 1.4 million images



top-5 score

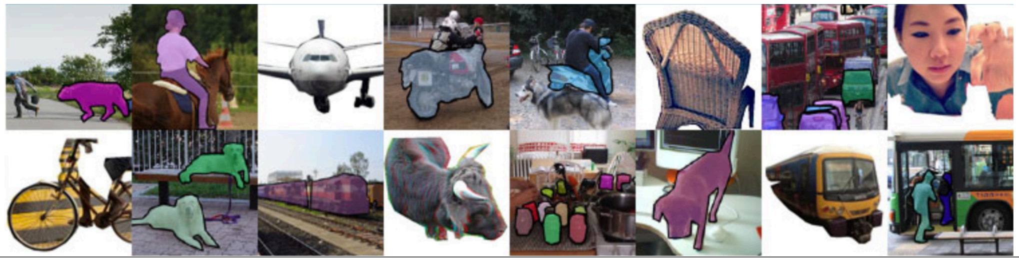
Datasets and Challenges



top 5 performance

Microsoft COCO
Common Objects in Context

- Object segmentation, recognition in context, multiple objects per image
  - 300K images
  - 2 Million instances
  - 80 Object categories
  - 5 captions per image
  - Keypoints on 100K people



Microsoft COCO
Common Objects in Context

- Object segmentation, recognition in context, multiple objects per image

Microsoft COCO
Common Objects in Context

- Challenges
  - Detection: Bounding box or segmentation
  - Captions: Generate text descriptions of images (rated by humans)
  - Keypoints: detect people and then localize human body key points

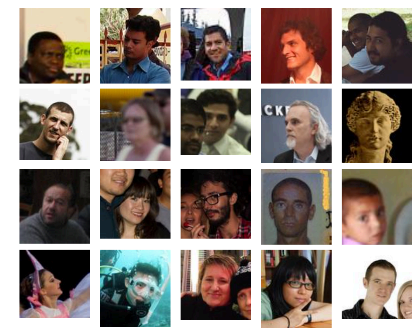
- Scene categorization
- +10 Million images
- +400 scene categories

http://visualgenome.org

MegaFace

http://megaface.cs.washington.edu

- Face recognition
- Training: 4.7M photos, 672K unique identities
- Distractors: 1 Million photos, 690K unique users
- Test: FaceScrub Celebrities, FGNet Age-Invariant non Celebrities



Distractors are faces of unknown people. FaceScrub is a smaller dataset (100K images of 530 celebs) that can be used for test purposes. FGNet has 975 images of 82 people (also used for test purposes).