Artificial Intelligence

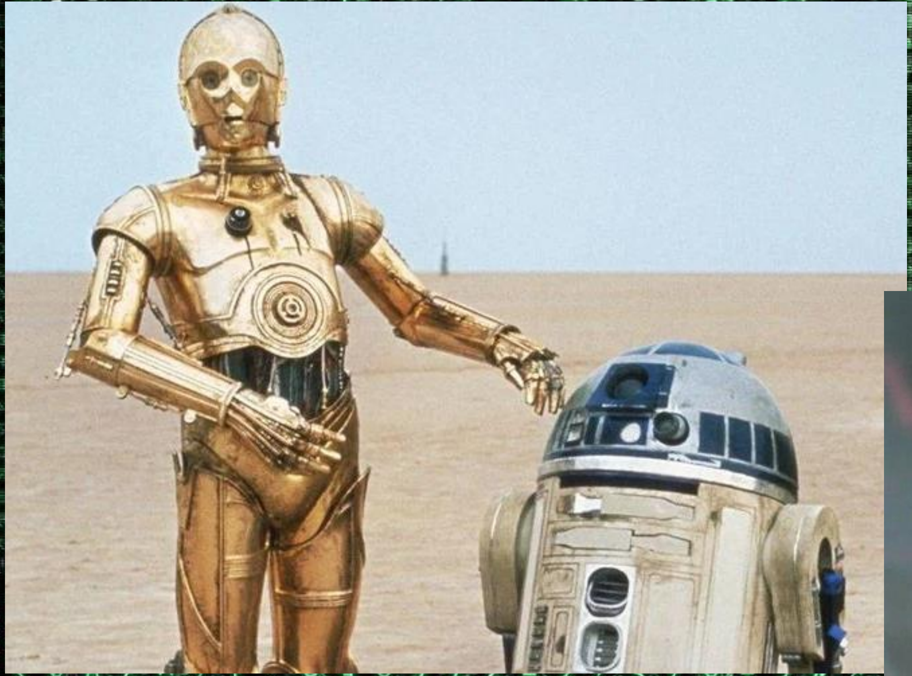
Lecture 11-1

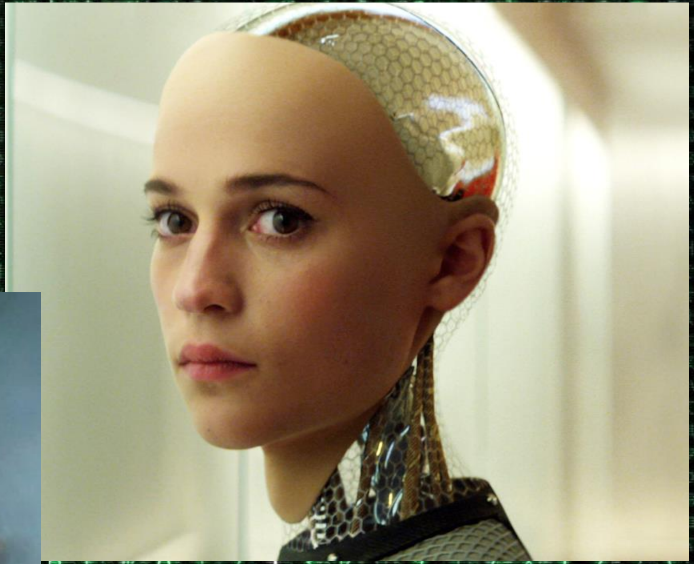
Computers & Society (CPSC 430)

Kevin Leyton-Brown

https://www.cs.ubc.ca/~kevinlb/teaching/cs430

Movies Help Us Think About AI

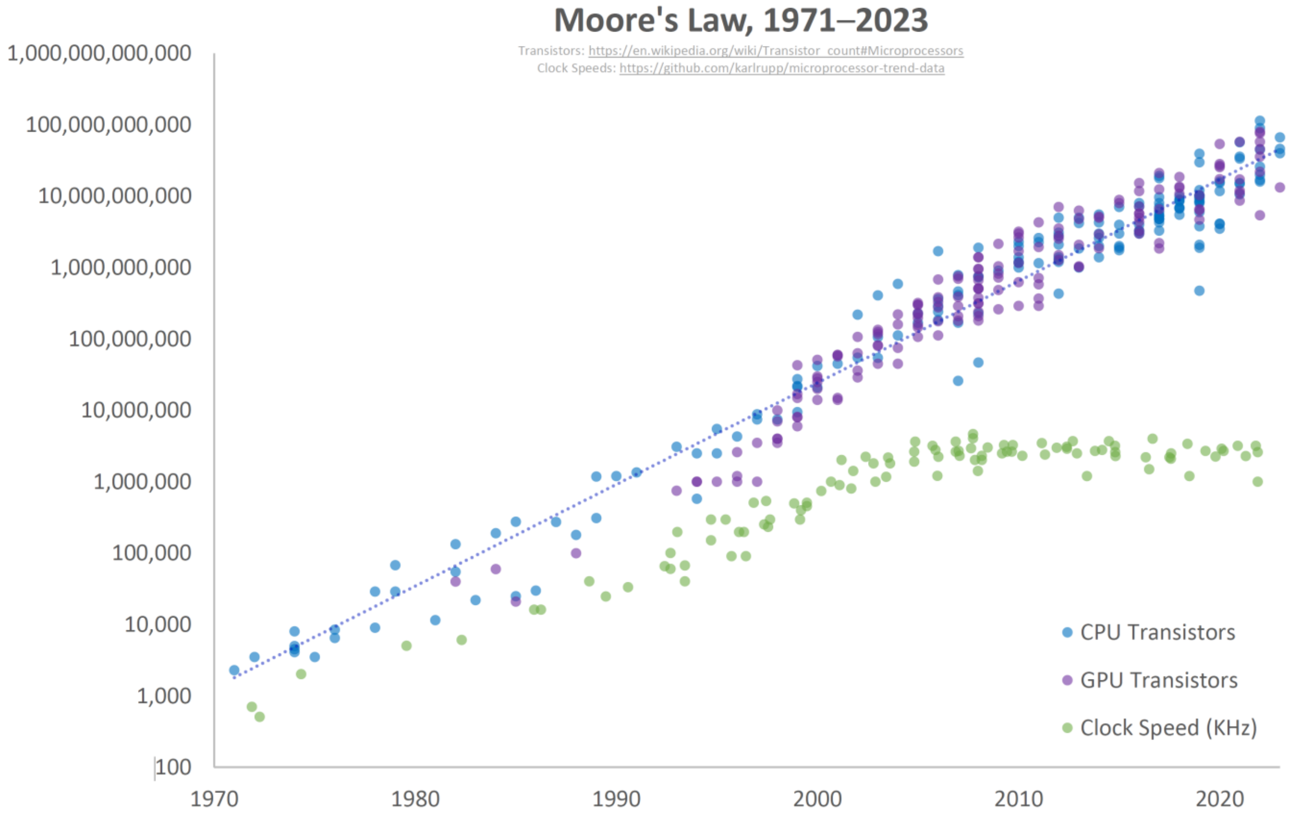




But most modern AI systems look like…

Why is it happening?

1. Scientific/mathematical **breakthroughs**, esp in machine learning
2. Growth in raw **computing power**



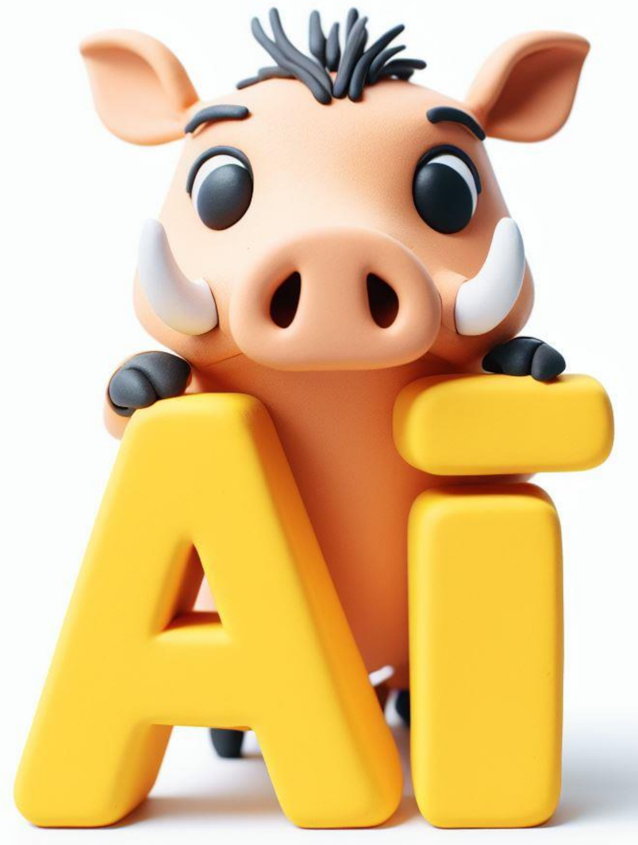
For comparison:

Species	# Neurons in Brain
Fruit Fly	100,000
Cat	1,000,000,000
Chimpanzee	10,000,000,000
Human	100,000,000,000

What is Artificial Intelligence?

AI is the study, design, and development of computational processes to solve problems that **previously required human intelligence**

The "**AI Paradox**": once we become familiar with a technology, we stop considering it AI



“Good Old-Fashioned AI”

- Early AI systems were **explicitly programmed**
  - reasoning systems were based on logic
  - rule-based “expert systems”
  - language systems explicitly modeled grammar
  - vision systems reasoned about optics, geometry

- Many important **conceptual foundations**

- Few **practical successes**
  - systems were brittle in practice
  - dealt poorly with noise, imperfect world models

A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence

August 31, 1955

John McCarthy, Marvin L. Minsky, Nathaniel Rochester, and Claude E. Shannon

■ The 1956 Dartmouth summer research project on artificial intelligence was initiated by this August 31, 1955 proposal, authored by John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon. The original typescript consisted of 17 pages plus a title page. Copies of the typescript are housed in the archives at Dartmouth College and Stanford University. The first 5 papers state the proposal, and the remaining pages give qualifications and interests of the four who proposed the study. In the interest of brevity, this article reproduces only the proposal itself, along with the short autobiographical statements of the proposers.

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

The following are some aspects of the artificial intelligence problem:

1. Automatic Computers

If a machine can do a job, then an automatic calculator can be programmed to simulate the machine. The speeds and memory capacities of present computers may be insufficient to simulate many of the higher functions of the human brain, but the major obstacle is not lack of machine capacity, but our inability to write programs taking full advantage of what we have.

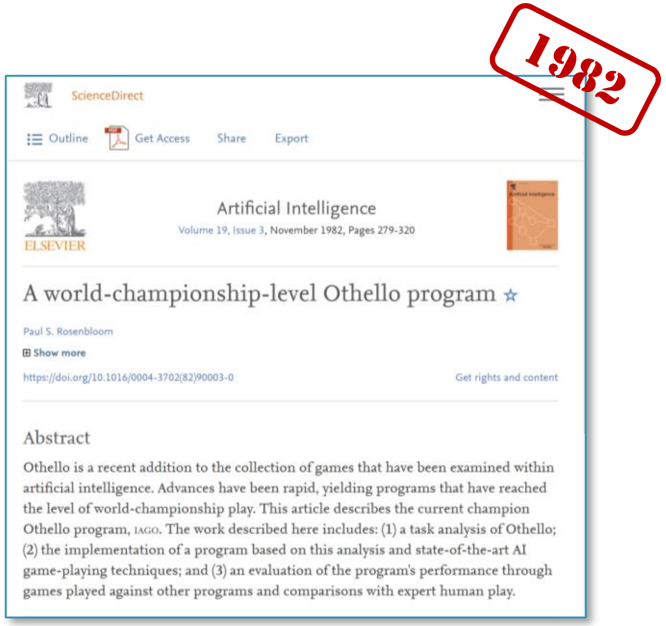
2. How Can a Computer be Programmed to Use a Language

It may be speculated that a large part of human thought consists of manipulating words according to rules of reasoning and rules of conjecture. From this point of view, forming a generalization consists of admitting a new

Search

- Instead of telling a computer how to solve a problem, tell it how to **recognize a solution & let it experiment**

- Drove many of AI's **early successes**:



ScienceDirect

Outline Get Access Share Export

Artificial Intelligence

Volume 19, Issue 3, November 1982, Pages 279-320

A world-championship-level Othello program

Paul S. Rosenbloom

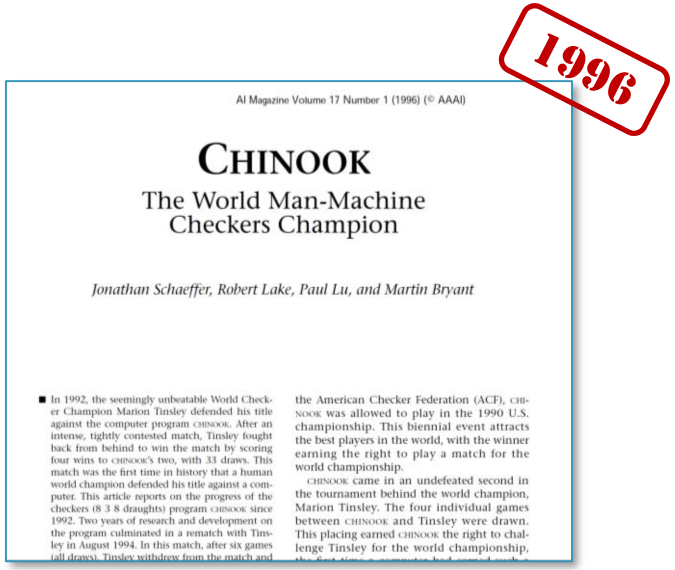
Show more

https://doi.org/10.1016/0004-3702(82)90003-0

Get rights and content

Abstract

Othello is a recent addition to the collection of games that have been examined within artificial intelligence. Advances have been rapid, yielding programs that have reached the level of world-championship play. This article describes the current champion Othello program, IAGO. The work described here includes: (1) a task analysis of Othello; (2) the implementation of a program based on this analysis and state-of-the-art AI game-playing techniques; and (3) an evaluation of the program's performance through games played against other programs and comparisons with expert human play.



AI Magazine Volume 17 Number 1 (1996) (© AAAI)

CHINOOK

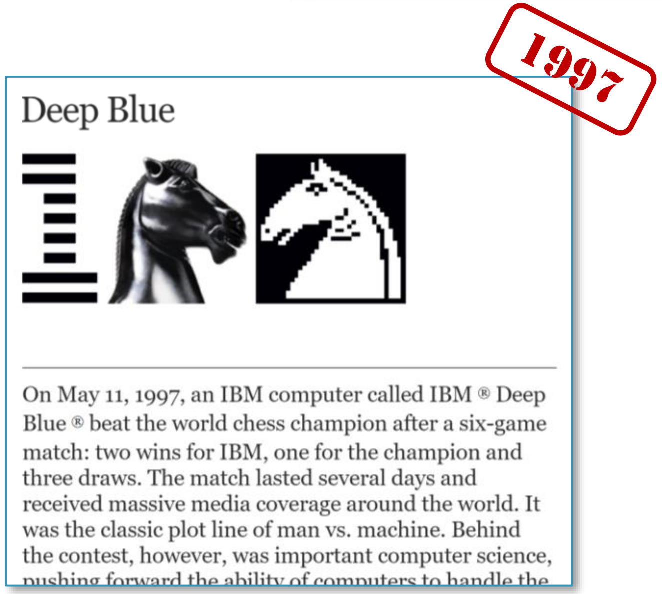
The World Man-Machine Checkers Champion

Jonathan Schaeffer, Robert Lake, Paul Lu, and Martin Bryant

In 1992, the seemingly unbeatable World Checker Champion Marion Tinsley defended his title against the computer program CHINOOK. After an intense, tightly contested match, Tinsley fought back from behind to win the match by scoring four wins to CHINOOK's two, with 33 draws. This match was the first time in history that a human world champion defended his title against a computer. This article reports on the progress of the checkers (8 3 8 draughts) program CHINOOK since 1992. Two years of research and development on the program culminated in a rematch with Tinsley in August 1994. In this match, after six games (all draws), Tinsley withdrew from the match and

the American Checker Federation (ACF), CHINOOK was allowed to play in the 1990 U.S. championship. This biennial event attracts the best players in the world, with the winner earning the right to play a match for the world championship.

CHINOOK came in an undefeated second in the tournament behind the world champion, Marion Tinsley. The four individual games between CHINOOK and Tinsley were drawn. This placing earned CHINOOK the right to challenge Tinsley for the world championship,

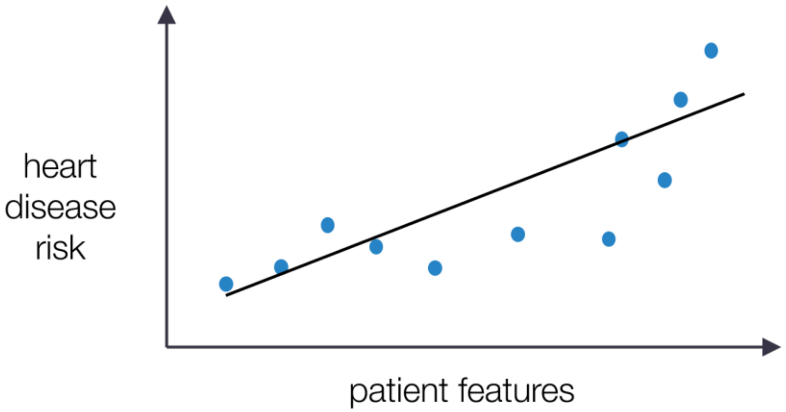


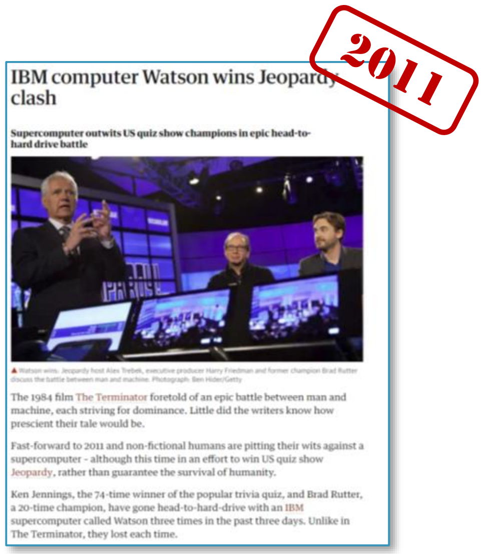
Deep Blue

On May 11, 1997, an IBM computer called IBM ® Deep Blue ® beat the world chess champion after a six-game match: two wins for IBM, one for the champion and three draws. The match lasted several days and received massive media coverage around the world. It was the classic plot line of man vs. machine. Behind the contest, however, was important computer science, pushing forward the ability of computers to handle the

Machine Learning

- Give a computer **examples of a pattern** and ask it to find a **rule**
- x: **features**; y: **labels**
- Example:
  - x = blood pressure, diet, exercise, age, gender
  - y = risk of heart disease





IBM computer Watson wins Jeopardy clash

Supercomputer outwits US quiz show champions in epic head-to-hard drive battle

Watson wins: Jeopardy host Alex Trebek, executive producer Harry Friedman and former champion Brad Rutter discuss the battle between man and machine. Photograph: Ben Hider/Getty

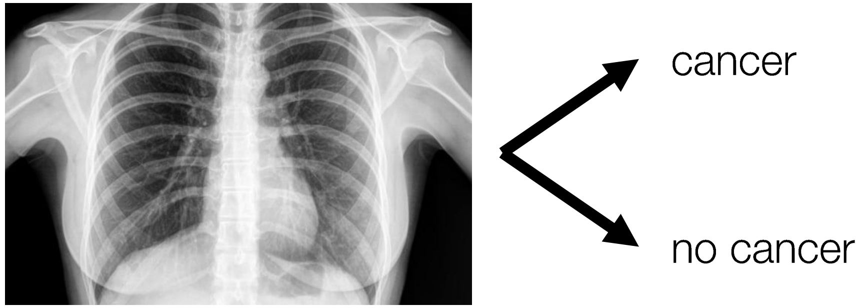
The 1984 film The Terminator foretold of an epic battle between man and machine, each striving for dominance. Little did the writers know how prescient their tale would be.

Fast-forward to 2011 and non-fictional humans are pitting their wits against a supercomputer - although this time in an effort to win US quiz show Jeopardy, rather than guarantee the survival of humanity.

Ken Jennings, the 74-time winner of the popular trivia quiz, and Brad Rutter, a 20-time champion, have gone head-to-hard-drive with an IBM supercomputer called Watson three times in the past three days. Unlike in The Terminator, they lost each time.

Deep Learning with Neural Networks

- Get **rid of features**!
  - build machine learning models that take raw inputs like pictures, sound recordings, text, …
- Architecture is loosely analogous to **brains**
- An **old idea** (60s; 80s)
  - Fundamental benefit: scalable model complexity
  - Breakthrough idea (2014): accelerate training with GPUs
- Example:
  - x = lung X-ray image
  - y = lung cancer diagnosis



Image, Face Recognition

- Understanding **images and faces** had long been seen as a fundamentally hard AI problem
- Deep learning was a **game changer**

DeepFace: Closing the Gap to Human-Level Performance in Face Verification

Yaniv Taigman   Ming Yang   Marc'Aurelio Ranzato   Lior Wolf

Facebook AI Research
Menlo Park, CA, USA
{yaniv, mingyang, ranzato}@fb.com

Tel Aviv University
Tel Aviv, Israel
wolf@cs.tau.ac.il

Abstract

In modern face recognition, the conventional pipeline consists of four stages: detect ⇒ align ⇒ represent ⇒ classify. We revisit both the alignment step and the representation step by employing explicit 3D face modeling in order to apply a piecewise affine transformation, and derive a face representation from a nine-layer deep neural network. This deep network involves more than 120 million parameters using several locally connected layers without weight sharing, rather than the standard convolutional layers. Thus we trained it on the largest facial dataset to-date, an identity labeled dataset of four million facial images belonging to more than 4,000 identities. The learned representations coupling the accurate model-based alignment with the large facial database generalize remarkably well to faces in unconstrained environments, even with a simple classifier. Our method reaches an accuracy of 97.35% on the Labeled Faces in the Wild (LFW) dataset, reducing the error of the current state of the art by more than 27%, closely approaching human-level performance.

1. Introduction

Face recognition in unconstrained images is at the forefront of the algorithmic perception revolution. The social and cultural implications of face recognition technologies are far reaching, yet the current performance gap in this domain between machines and the human visual system serves as a buffer from having to deal with these implications.

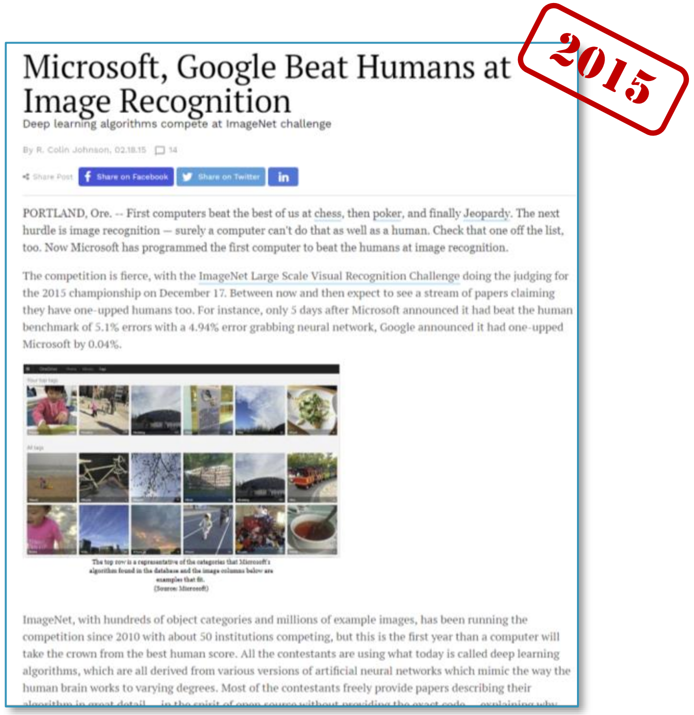
toward tens of thousands of appearance features in other recent systems [5, 7, 2].

The proposed system differs from the majority of contributions in the field in that it uses the deep learning (DL) framework [3, 21] in lieu of well engineered features. DL is especially suitable for dealing with large training sets, with many recent successes in diverse domains such as vision, speech and language modeling. Specifically with faces, the success of the learned net in capturing facial appearance in a robust manner is highly dependent on a very rapid 3D alignment step. The network architecture is based on the assumption that once the alignment is completed, the location of each facial region is fixed at the pixel level. It is therefore possible to learn from the raw pixel RGB values, without any need to apply several layers of convolutions as is done in many other networks [19, 21].

In summary, we make the following contributions : (i) The development of an effective deep neural net (DNN) architecture and learning method that leverage a very large labeled dataset of faces in order to obtain a face representation that generalizes well to other datasets; (ii) An effective facial alignment system based on explicit 3D modeling of faces; and (iii) Advance the state of the art significantly in (1) the Labeled Faces in the Wild benchmark (*LFW*) [18], reaching near human-performance; and (2) the YouTube Faces dataset (*YTF*) [30], decreasing the error rate there by more than 50%.

1.1. Related Work

Big data and deep learning In recent years, a large number of photos have been crawled by search engines, and up-



Microsoft, Google Beat Humans at Image Recognition

Deep learning algorithms compete at ImageNet challenge

By R. Colin Johnson, 02.18.15 14

Share Post Share on Facebook Share on Twitter in

PORTLAND, Ore. -- First computers beat the best of us at chess, then poker, and finally Jeopardy. The next hurdle is image recognition -- surely a computer can't do that as well as a human. Check that one off the list, too. Now Microsoft has programmed the first computer to beat the humans at image recognition.

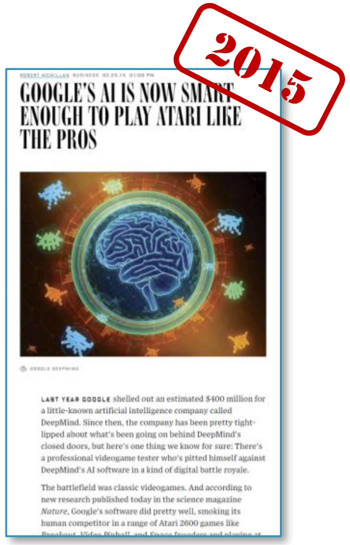
The competition is fierce, with the ImageNet Large Scale Visual Recognition Challenge doing the judging for the 2015 championship on December 17. Between now and then expect to see a stream of papers claiming they have one-upped humans too. For instance, only 5 days after Microsoft announced it had beat the human benchmark of 5.1% errors with a 4.94% error grabbing neural network, Google announced it had one-upped Microsoft by 0.04%.

The top row is a representative of the categories that Microsoft's algorithm found in the database and the image columns below are examples that fit.
(Source: Microsoft)

ImageNet, with hundreds of object categories and millions of example images, has been running the competition since 2010 with about 50 institutions competing, but this is the first year than a computer will take the crown from the best human score. All the contestants are using what today is called deep learning algorithms, which are all derived from various versions of artificial neural networks which mimic the way the human brain works to varying degrees. Most of the contestants freely provide papers describing their

Reinforcement Learning

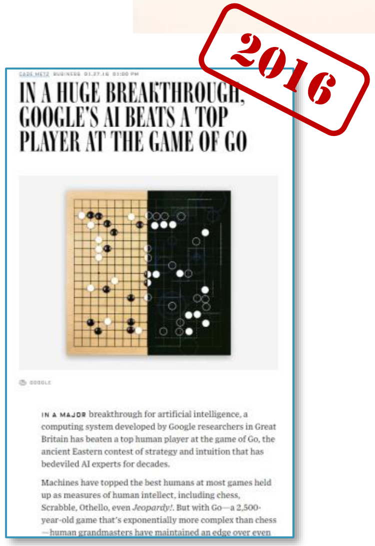
- Often a complex **sequence of actions** must be taken before reaching a **reward or punishment**
  - RL: an ML approach for such settings
- Example:
  - navigate a maze to reach a goal
  - you need a key to unlock the door
  - quicksand slows you down
- **Foundations** of RL laid in the 80s
- **Breakthroughs** in mid 2010s:
  - state representation using deep learning
  - Monte Carlo Tree Search
  - new policy search algorithms



GOOGLE'S AI IS NOW SMART ENOUGH TO PLAY ATARI LIKE THE PROS

LAST YEAR GOOGLE shelled out an estimated $400 million for a little-known artificial intelligence company called DeepMind. Since then, the company has been pretty tight-lipped about what's been going on behind DeepMind's closed doors, but here's one thing we know for sure: There's a professional videogame tester who's pitted himself against DeepMind's AI software in a kind of digital battle royale.

The battlefield was classic videogames. And according to new research published today in the science magazine *Nature*, Google's software did pretty well, smoking its human competitor in a range of Atari 2600 games like



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

IN A MAJOR breakthrough for artificial intelligence, a computing system developed by Google researchers in Great Britain has beaten a top human player at the game of Go, the ancient Eastern contest of strategy and intuition that has bedeviled AI experts for decades.

Machines have topped the best humans at most games held up as measures of human intellect, including chess, Scrabble, Othello, even *Jeopardy!*. But with Go—a 2,500-year-old game that's exponentially more complex than chess—human grandmasters have maintained an edge over even

AI shoves all in: DeepStack, Libratus poker bots battle Texas Hold 'em pros heads up

Both robo-players ace humans – in one-on-one matches

By Katyanna Quach 14 Jan 2017 at 08:53 15 SHARE

DeepStack is the first AI computer programme to beat professional poker players in a game of hands-on no-limit Texas hold'em, a team of researchers claim in a research paper out this week.

Self-Supervision and Generative AI

Claymation zebra standing in front of a chalkboard holding a yellow book

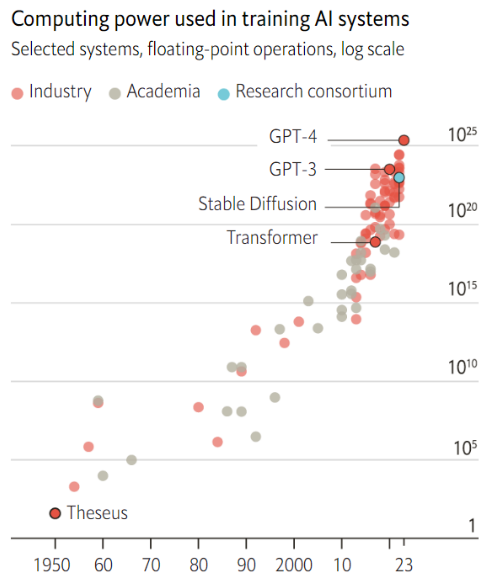
- How can we learn from **huge, unlabeled datasets**?
  - traditional ML needs class labels
  - RL needs rewards
- A really clever idea: turn raw data into **puzzles**

  > Stanley Park has a long history. The land was originally used by Indigenous peoples for thousands of years before British Columbia was colonized by the British during the 1858 Fraser Canyon Gold Rush and was one of the first areas to be

- The same idea works with **images**
  - add noise and ask the model to remove it
  - use existing image captions as “clues” to teach association between text and images
- Eventually, models can **generate** inputs like those they saw during training
  - give the first few lines of an essay and the model will complete it
  - start with random noise and a text description and ask the model to denoise it

Large Language Models (GPT-3, ...)

- Start with a huge corpus of text
  - the entire Internet, filtered for link spam
  - books and news
  - code repositories
  - transcripts of conversations
- Take an absolutely enormous neural net
  - GPT-2: 1.5B parameters
  - GPT-3: 175B parameters
  - GPT-4: 8 * 220B = 1.76T parameters
- Spend literally millions of dollars conducting self-supervised training
- The resulting "foundation model" can be "fine-tuned" to specific tasks
  - e.g., sentiment analysis on Twitter: X = tweet; Y = positive/negative



Chat Models (ChatGPT, ...)

- Supervised fine tuning
  - align to the **chat task** via examples of input and output
- Human alignment
  - generate **multiple texts** for each prompt
  - get humans to **rank them**
    - it matters who gets chosen to do this work!
    - in newer work, get an aligned LLM to rank them (RLAIF)
  - train a **reward model** to predict these human preferences
    - or possibly multiple models: helpfulness; honesty; harmlessness
- **Reinforcement learning**
  - initial state: context
  - actions: words
  - reward: human-aligned reward model