August 4, 1997: Skynet goes online
August 29, 1997, 2:14am ET: Skynet gains ‘consciousness’
August 29, 1997: Judgement Day
CURRENT DATE: 09.10.1999

TARGET DATE: 2007

ENABLED
What ‘features’ does T-800 have?
What ‘features’ does T-800 have?

• Vision/Perception
• Speech
• Motion/Planning
• Object Manipulation
• Self-Repair
• Goal-directed Exploration
• Social Skills
• Ethics/Morality
Back to our timeline..
1960’s (1st Wave)

- Single Layer networks

- XOR problem killed research for two decades
Mid-1980s (2nd Wave)

- Multi-layer networks
- Backpropagation algorithm
2010s (3rd Wave)

- **Big Data**
  - O(M) labeled images
- **Big Compute**
- ‘Deep’ Learning
ImageNet Architecture
Deep Learning for Computer Vision

- Bird
- Frog
- Person
- Dog
- Chair
- Person
- Hammer
- Flower pot
- Power drill
- Person
- Car
- Helmet
- Motorcycle
Imagenet ILSVRC Challenge

Error rate

1: ImageNet top 5 error rate
Source: ImageNet
Deep Learning for Segmentation
Deep Learning for Caption Generation

A woman is throwing a **frisbee** in a park.

A dog is standing on a hardwood floor.

A **stop** sign is on a road with a mountain in the background.

A little girl sitting on a bed with a teddy bear.

A group of **people** sitting on a boat in the water.

A giraffe standing in a forest with trees in the background.
Some things you can ask me:

“Did the Giants win?”
“Game schedules”
“My mom is Susan Conway”
“Read it again”
“Delete my 7:30 alarm”
Google Assistant

Hi, how can I help?

Make me a haircut appointment on Tuesday morning anytime between 10 and 12.
At last—a computer program that can beat a champion Go player

**ALL SYSTEMS GO**

CONSERVATION

SONGBIRDS À LA CARTE
Illegal harvest of millions of Mediterranean birds

RESEARCH ETHICS

SAFEGUARD TRANSPARENCY
Don’t let openness backfire on individuals

POPULAR SCIENCE

WHEN GENES GOT ‘SELFISH’
Dawkins’s calling card 40 years on
40 days – AlphaGo Zero surpasses all previous versions, becomes the best Go player in the world.

36 hours – AlphaGo Zero reaches level of Alpha Go Lee, which beat world champion Lee Sedol in 2016.

72 hours – AlphaGo Zero beats Alpha Go Lee, 100:0.
IBM Watson
IBM Project Debater

Project Debater

Dan Zafrir
DARPA Robotics Challenge
Object Manipulation
Deep Learning for Self-Driving Cars
Some Failures...
Can Deep Learning Work for Science?

• **Similarities**
  – Tasks:
    • Pattern Classification
    • Regression
    • Clustering
    • Feature Learning
    • Anomaly Detection

• **Differences**
  – Unique attributes of Scientific Data
    • Multi-channel / Multi-variate
    • Double precision floating point
    • Noise and Artefacts
    • Statistics are likely different
CAM5 0.25-degree simulation data
Challenge: Multi-Variate Data
Climate Science Tasks

Classification    Classification + Localization    Object Detection    Instance Segmentation

Images of cats and dogs with boxes and labels indicating different tasks.
## Supervised Convolutional Architecture

<table>
<thead>
<tr>
<th></th>
<th>Logistic Regression</th>
<th>K-Nearest Neighbor</th>
<th>Support Vector Machine</th>
<th>Random Forest</th>
<th>ConvNet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tropical Cyclone</td>
<td>Test</td>
<td>Test</td>
<td>Test</td>
<td>Test</td>
<td>Test</td>
</tr>
<tr>
<td></td>
<td>95.85</td>
<td>97.85</td>
<td>95.85</td>
<td><strong>99.4</strong></td>
<td>99.1</td>
</tr>
<tr>
<td>Atmospheric Rivers</td>
<td>82.65</td>
<td>81.7</td>
<td>83.0</td>
<td>88.4</td>
<td><strong>90.0</strong></td>
</tr>
<tr>
<td>Weather Fronts</td>
<td><strong>89.8</strong></td>
<td>76.45</td>
<td>90.2</td>
<td>87.5</td>
<td><strong>89.4</strong></td>
</tr>
</tbody>
</table>

![Supervised Convolutional Architecture Diagram]

- **Input**: Convolution
- **Pooling**: Convolution
- **Pooling**: Fully connect
- **Class score**
Semi-Supervised Convolutional Architecture (NIPS’17)

Encoder

Decoder

Classification + Bounding Box Regression

Contributors: Evan Racah, Chris Pal, Chris Beckham, Samira Kahou, Tegan Maharaj. MILA
Classification + Regression Results

Ground Truth
Prediction

Contributors: Thorsten Kurth, Jian Yang, Ioannis Mitliagkas, Chris Pal, Nadathur Satish, Narayanan Sundaram, Amir Khosrowshahi, Michael Wehner, Bill Collins, Intel, Stanford, LBL, MILA.
Deep Learning at 15PF (SC’17)
Segmentation

Contributors: Mayur Mudigonda, Thorsten Kurth, Sean Treichler, Josh Romero, Massimiliano Fatica, Mike Houston. UC Berkeley, LBL, NVIDIA
Segmentation Results
Deep Learning for Science

- Modeling galaxy shapes
- Generating cosmology mass maps
- Decoding speech from ECoG
- Clustering Daya Bay events
- LHC Signal/Background classification
- Oxford Nanopore sequencing
Deep Learning Hardware

**Edison: Cray XC-30**
- 5,576 nodes, 133K, 2.4GHz Intel "IvyBridge" Cores, 357TB RAM

**Cori: Cray XC-40**
- Ph1: 1630 nodes, 2.3GHz Intel "Haswell" Cores, 203TB RAM
- Ph2: >9300 nodes, >60cores, 16GB HBM, 96GB DDR per node

**Data-Intensive Systems**
- PDSF, JGI, KBASE, HEP
- 14x QDR

**Ethernet & IB Fabric**
- Science Friendly Security
- Production Monitoring
- Power Efficiency
- WAN

**HPSS**
- 250 TB
- NetApp 5460
- 50 PB stored, 240 PB capacity

**Global Scratch**
- 3.6 PB
- 5 x SFA12KE

**/project**
- 5 PB
- DDN9900 & NexSAN

**/home**
- 250 TB
- NetApp 5460

**2 x 10 Gb**
- Software Defined Networking

**1 x 100 Gb**
- Science Gateways

**Vis & Analytics**
- Data Transfer Nodes
- Adv. Arch. Testbeds
## Deep Learning Software

<table>
<thead>
<tr>
<th>Deep Learning Frameworks</th>
<th>Technologies</th>
</tr>
</thead>
<tbody>
<tr>
<td>TensorFlow, Caffe, PyTorch, Theano</td>
<td>Neon, CNTK, MXNet, ...</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Multi Node libraries</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MLSL, MPI, Horovod, GRPC</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Single Node libraries</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MKL-DNN, CuDNN</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hardware</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CPUs (KNL), GPUs, FPGAs, Accelerators</td>
<td></td>
</tr>
</tbody>
</table>
Open Challenges

1. Performance and Scaling
2. Complex Data
3. Hyper-Parameter Optimization
4. Scarcity of Labeled Data
5. Interpretability and Visualization
6. Theory
Assumptions...

- Communities will self-organize and conduct labeling campaigns
  - Active Learning systems can determine optimal strategies for seeking labels
- Incorporation of domain science principles into learning algorithms
  - Solution spaces that satisfy physical constraints
- Pattern Classification, Clustering, Anomaly Detection are solved problems
What is the role of humans?

- **Labels**
- **Mechanisms, Hypothesis**
- **Patterns, Clusters, Anomalies**

/Project

HPSS
How close are we to creating a T-800?

- Vision/Perception ✓ 😊🎉
- Speech 😊🚀💥
- Motion/Planning ✓ 🔥🚀
- Object Manipulation 😞
- Self-Repair 😞
- Goal-directed Exploration 😞
- Social Skills 😞😊😔
- Ethics/Morality 😭😱
How about the T-1000 and T-X?
Time Travel?
Conclusions

• AI appears to be working
  – Genuine breakthroughs in vision, speech, control
  – Wide range of commercial applications

• Tremendous potential for scientific applications
  – Low-hanging fruit, but hard questions are coming next

• NERSC is at the forefront of Deep Learning for Science
  – Applications
  – Hardware and Software
  – R&D on optimizations + scaling, methods
Thanks!

- Contact: prabhat@lbl.gov
- Connect on LinkedIn
- Internships, full-time opportunities