A Brief Introduction of Berkeley Deep Drive (BDD)

Presented by Ching-Yao Chan
PATH, UC Berkeley

on behalf of
Prof. Trevor Darrell
BDD Director
EECS, UC Berkeley & PATH Co-Director

Deep Learning: A Buzzword

- **Alpha Go**
- **Already broadly adopted at many high-tech companies**
- **A flurry of investments**
- **A cluster of start-ups**

Deep Learning

- **A Type of Machine Learning**
- **Machine Learning (Artificial Intelligence) systems acquire their knowledge, by extracting patterns from raw data.**
- **Data representation in deep learning**
  - **Mapping representation to output**
  - **Deep Learning introduces representation that are expressed in other simpler representations in multiple layers**
  - **Thus a Multi-Layer Deep Structure.**

Illustration of a Deep Learning Model

Output:
object identity

3rd hidden layer:
object parts

2nd hidden layer:
corners and contours

1st hidden layer:
edges

Visible layer:
input pixels

Source:

- **http://www.deeplearningbook.org/contents/intro.html**

Deep Learning History

- Deep Learning dated back to 1940s, known as Cybernetics in 1940s-1960s

- Connectionism in 1980s-1990s

- Current resurgence started in 2006

- In recent years, Deep Learning has advanced significantly due to several contributing factors
  - Greater computing power (CPU, GPU, Network)
  - Higher availability an affordability of GPUs
  - Large collection of data samples available to train and test algorithms (such as ImageNet, with 14,197,122 images, 21841 synsets indexed)

End-to-End Learning
for Vision, Text, Speech

What is Deep Learning?

Compositional Models
Learned End-to-End

Hierarchy of Representations
- vision: pixel, motif, part, object
- text: character, word, clause, sentence
- speech: audio, band, phone, word

concrete → abstract
learning

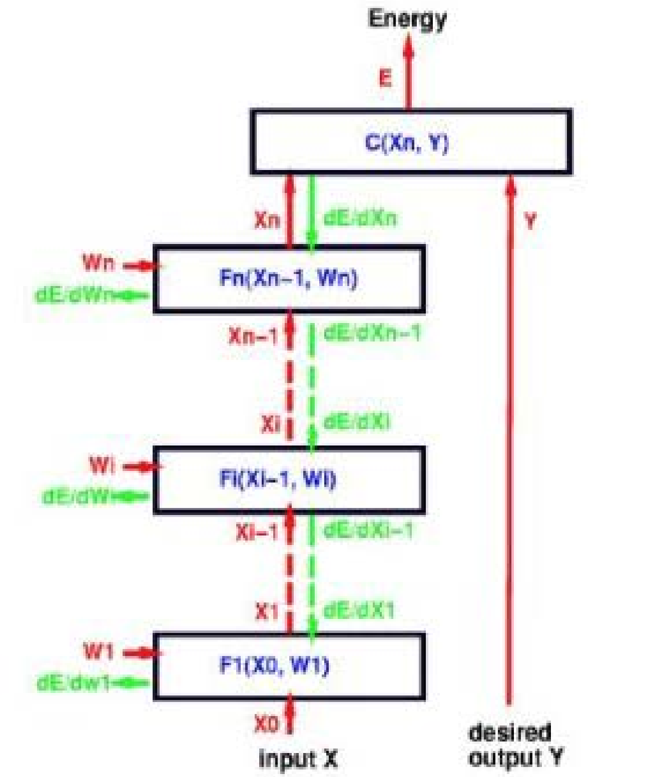
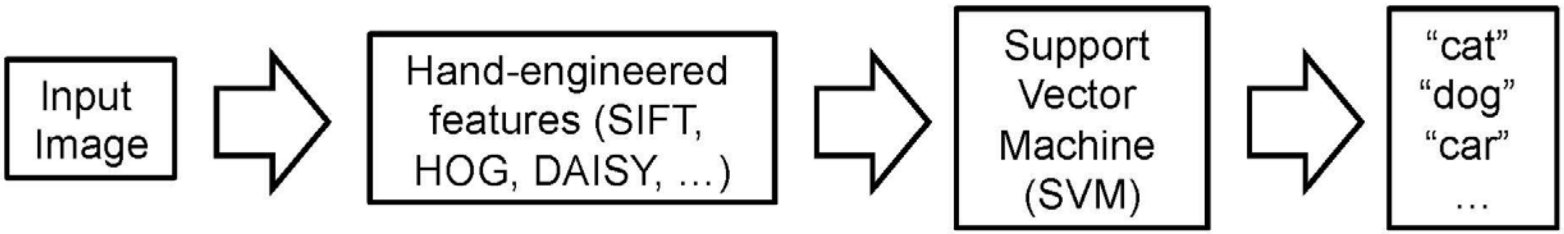


figure credit Yann LeCun, ICML '13 tutorial

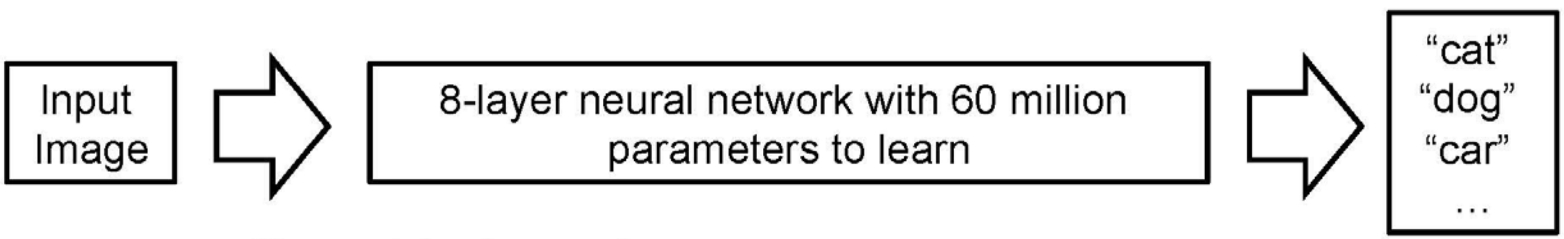
Source: Berkeley Vision and Learning Center
http://caffe.berkeleyvision.org/

Object Detection in Computer Vision

- State-of-the-art object detection until 2012:



- Deep Supervised Learning (Krizhevsky, Sutskever, Hinton 2012; also LeCun, Bengio, Ng, Darrell, …):

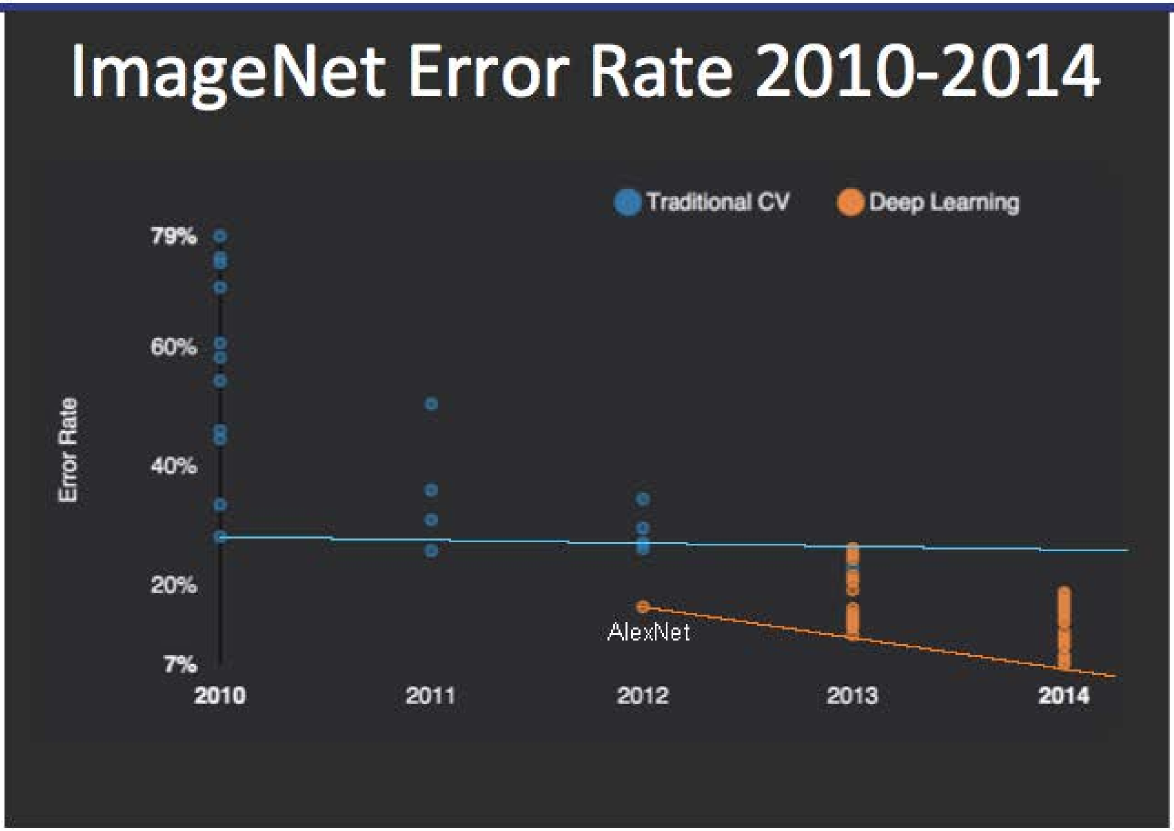


  - ~1.2 million training images from ImageNet [Deng, Dong, Socher, Li, Li, Fei-Fei, 2009]

Source: Pieter Abbeel presentation, 04/2016

Doing Better and Better

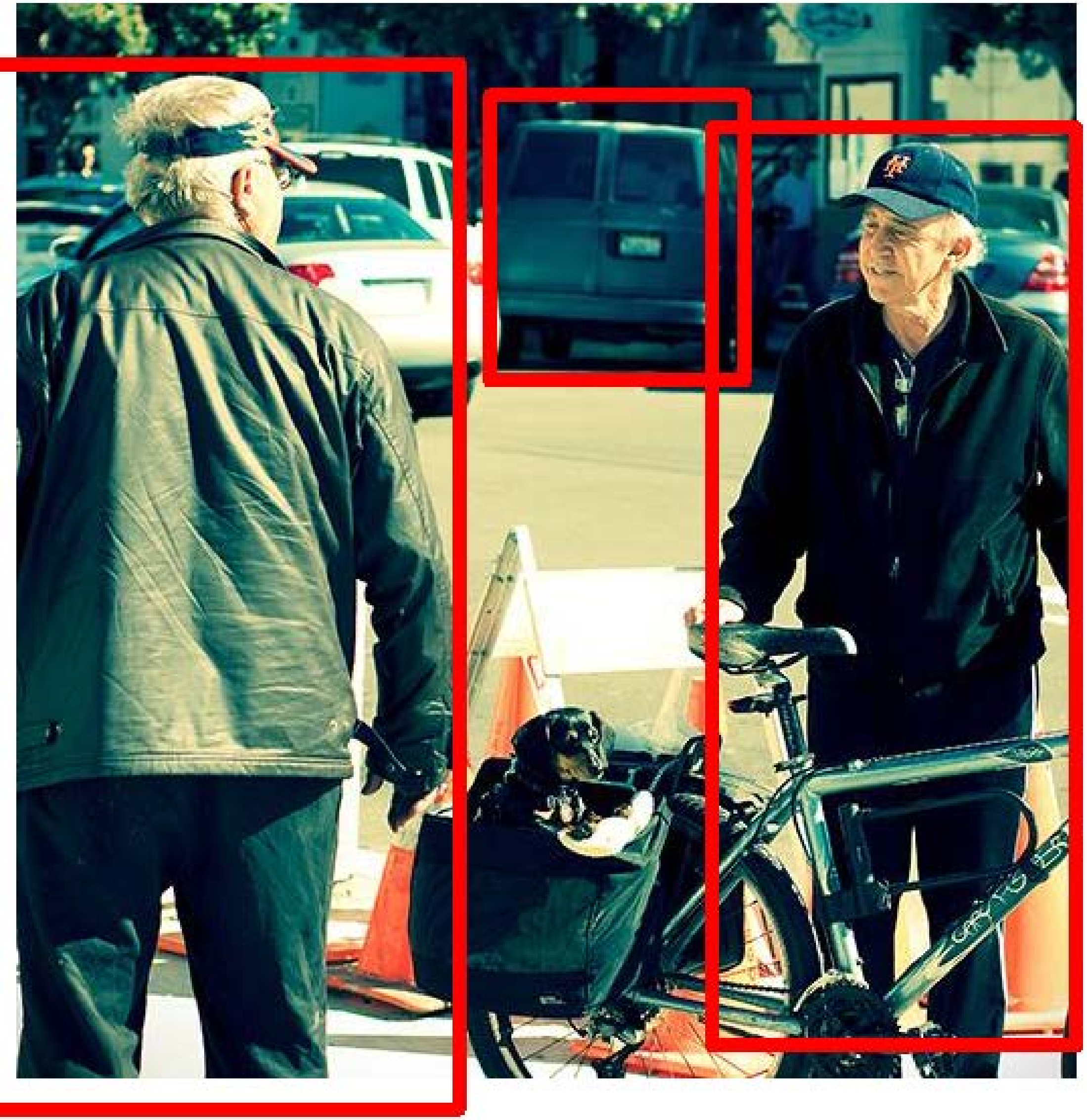
Performance



graph credit Matt Zeiler, Clarifai

Source: Pieter Abbeel presentation, 04/2016

Large-scale ***Semantic Description***



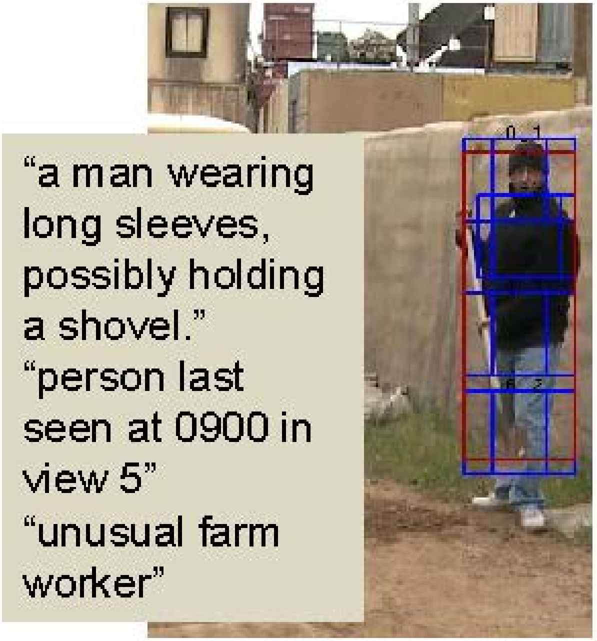
Object Detection

...

**Source:
Trevor Darrell
presentation**

*Large-scale **Semantic Description***

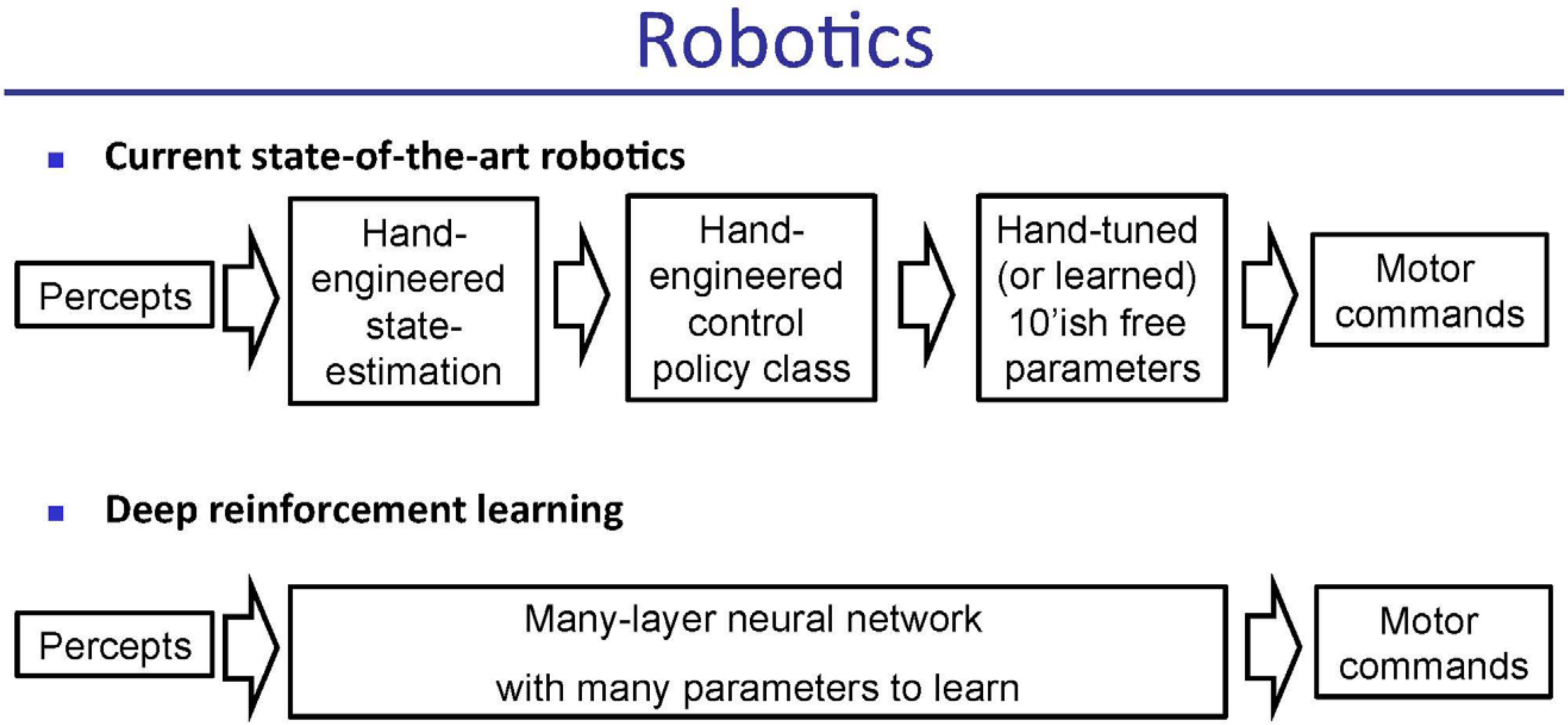
Object Detection
Semantic Segmentation
Pose Estimation
Attribute Classification
Fine-Grained Recognition
Action Recognition
…



Hypothetically,

- **Recent Tesla Incident (May 2016, Florida)**
    - **Supposedly, the Tesla (camera + radar) sensor did not recognize the "side of truck" versus the background sky;**
- **Can a "deep learning" system recognize an object that is "not the same" as a typical target?**

End-to-End Visuomotor



Source: Pieter Abbeel presentation, 04/2016

State of the Art – Nvidia Example

Nvidia demo of Visuomotor Control (April 2016)

- End-to-End learning, implementation on Drive-PX2 platform
- trained a convolutional neural network (CNN) to map raw pixels from a single front-facing camera directly to steering commands
- With only human steering inputs as training signals; does not explicitly train the system to detect the outline of roads
- Avoided the needs to recognize human-designated features, such as lane markings, other cars, etc., nor did it include a number of “if-then” rules
- As of 03/2016, 72 hours of training data collected
- Use simulation to enhance training
- **98% of autonomy** (see Nvidia paper) in field testing

Criticism and Challenges

- **Formal and complete safety design verification**
  - **Training data**
  - **Stability**
  - **Credit assignment**

- **Compliance with functional safety (such as ISO-26262)**
  - **safety assurance**

- **Disruptive proposition of end-to-end solution**
  - **Departure from conventional automotive model**

Deep Learning at Berkeley

- **Berkeley Vision Learning Center**
  - **A consortium that started in 2012**
  - **Tremendous advances in computer and deep learning.**
  - **Open-source CAFFE, widely used globally**

- **Berkeley Deep Drive (BDD) Center**
  - **A consortium that started in Spring 2016**
  - **Seeking to merge deep learning with automotive perception and bring computer vision technology to the forefront.**

Berkeley Deep Drive

A Research Alliance

to Investigate State-of-the-Art Technologies

in Computer Vision and Machine Learning

for Automotive Applications

Berkeley Deep Drive

- Current members include:
  - Audi/VW, Bosch, Ford, Honda, Hyundai, Nvidia, Panasonic, Qualcomm, Samsung, Toyota
  - GM, NXP, Sony recently joined
  - Nexar and Mapillary are contributing partners
    - Nexar provides 100,000 hours of driving videos yearly
    - Mapillary provides millions of images
- Several more are in agreement reviews

Berkeley Deep Drive

- Sponsors membership
  - Access to faculty and students
  - Shared use of research outcome, codes and data, in repository
  - Commercial use of BDD software repository, without further licensing agreement with UC Berkeley

- BDD project and scope of study
  - Proposals submitted by campus PIs
  - Sponsors review and rate proposals
  - Panel consolidates and decides on final selection of projects to sponsors

Berkeley Deep Drive

Categories of Exemplar Projects, 1 of 2

- **Deep Learning Methodologies**
  - Clockwork FCNs for Fast Video Processing
  - Cross-modal Transfer Learning
  - Deep Learning for Tracking
  - Fast Object Detection and Segmentation
  - Improving the Scaling of Deep Learning Networks by Characterizing and Exploiting Soft Convexity
  - Learning Deep Models Securely on Sensitive Imaging Data with Cryptographic Gaurantees
  - Unsupervised Representation Learning for Autonomous Driving
- **Deep Learning Implementation**
  - Benchmarks and Leaderboard for Deep Reinforcement Learning
  - Design Space Exploration for Deep Neural Nets for Advanced Driver Assistance Systems
  - FPGA PRET Accelerators of Deep Learning Classifiers for Autonomous Vehicles
  - Learning to Drive Under Unstructured Conditions
  - Low Latency Deep Inference for Self-Driving Vehicles
  - Secure and Privacy-Preserving Deep Learning

Berkeley Deep Drive

Categories of Exemplar Projects, 2 of 2

- **Detection and Perception**
  - **Pedestrian Models in Urban Environment for Autonomous Driving and Database of Video Sequences for Model Training, Testing, and Implementation**
  - **Real-Time Perception/Prediction of Traffic Scene with Deep Learning for Autonomous Driving**
  - **Motion Prediction for Urban Autonomous Driving Based on Stochastic Policy Learned via Deep Neural Network**
  - **Inference of Drivers' Intent at Intersections**
  - **Outdoor Semantic Scene Segmentation via Multi-modal Sensor Fusion**

- **Driver-Vehicle Interaction**
  - **Verifiable Control for (Semi)Autonomous Cars that Learn from Human (Re)Actions**
  - **Implicit Communication Through a Car's Motion**
  - **Understanding Driver Awareness for Smart Vehicles**
  - **Human-Machine Arbitration in Hybrid Driving Systems**

Future of AI / Deep Learning

- A progressive emergence

- A worldwide community

- Deep Learning for image and speech recognition widely deployed
  - Prospects for automated driving, medical imaging, robots, etc.

- Hardware for embedded applications needed
  - Nvidia, Qualcomm, etc.

- Probably a way to go for truly intelligent machine

Questions?

Check out BDD website
bdd.Berkeley.edu

Thank You!

Ching-Yao Chan
California PATH Program
University of California, Berkeley

cychan@berkeley.edu
TEL: 510-665-3621