

CSCI 599: Deep Learning and its Applications

Lecture 1

Fall 2017
Joseph J. Lim

Welcome to CSCI 599!

For today, we need your attendance checked.

Please check in with TAs after the class.

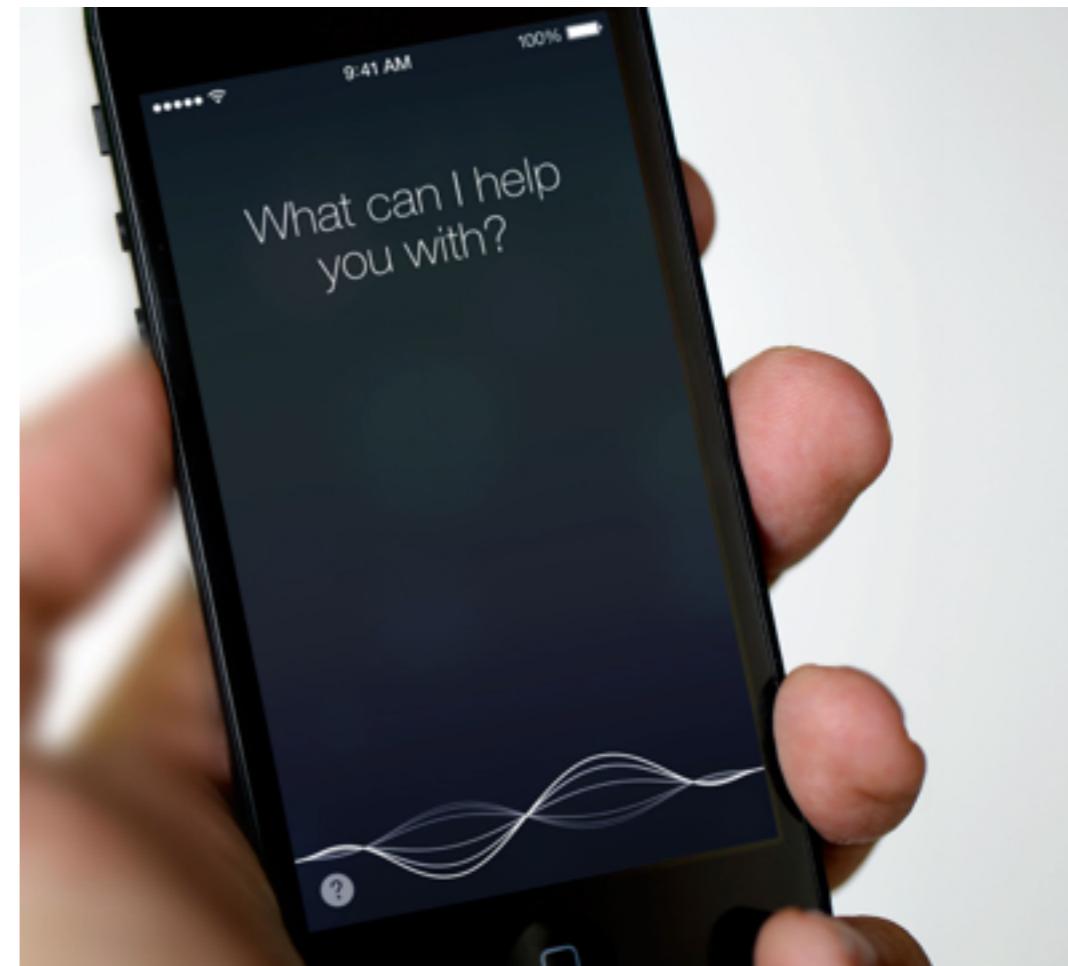
Welcome to CSCI 599!

This class will teach you

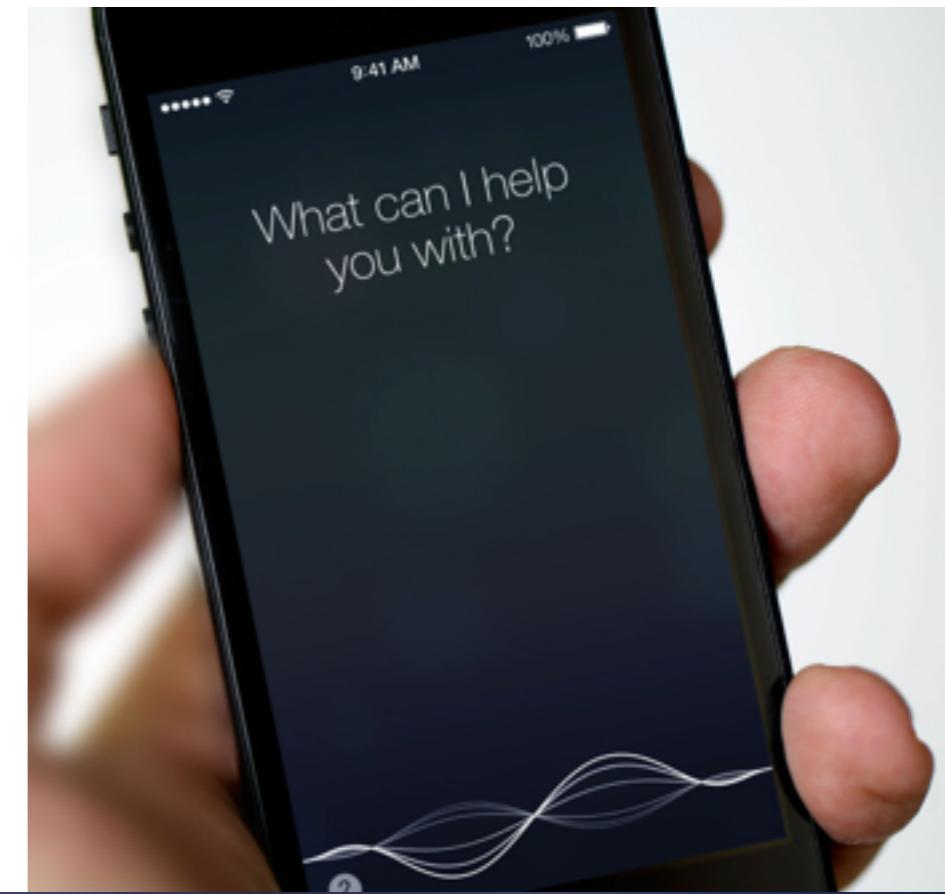
one of the most exciting developments in

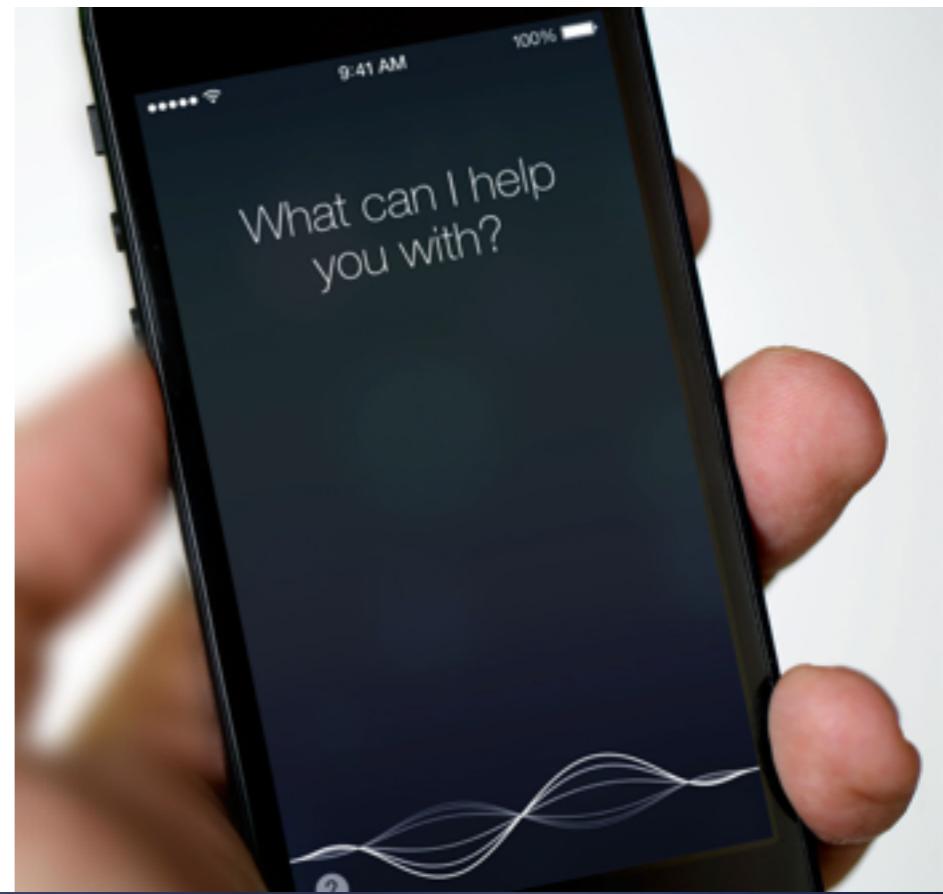
Machine Learning, Computer Vision, NLP, Robotics,
Other AI related fields

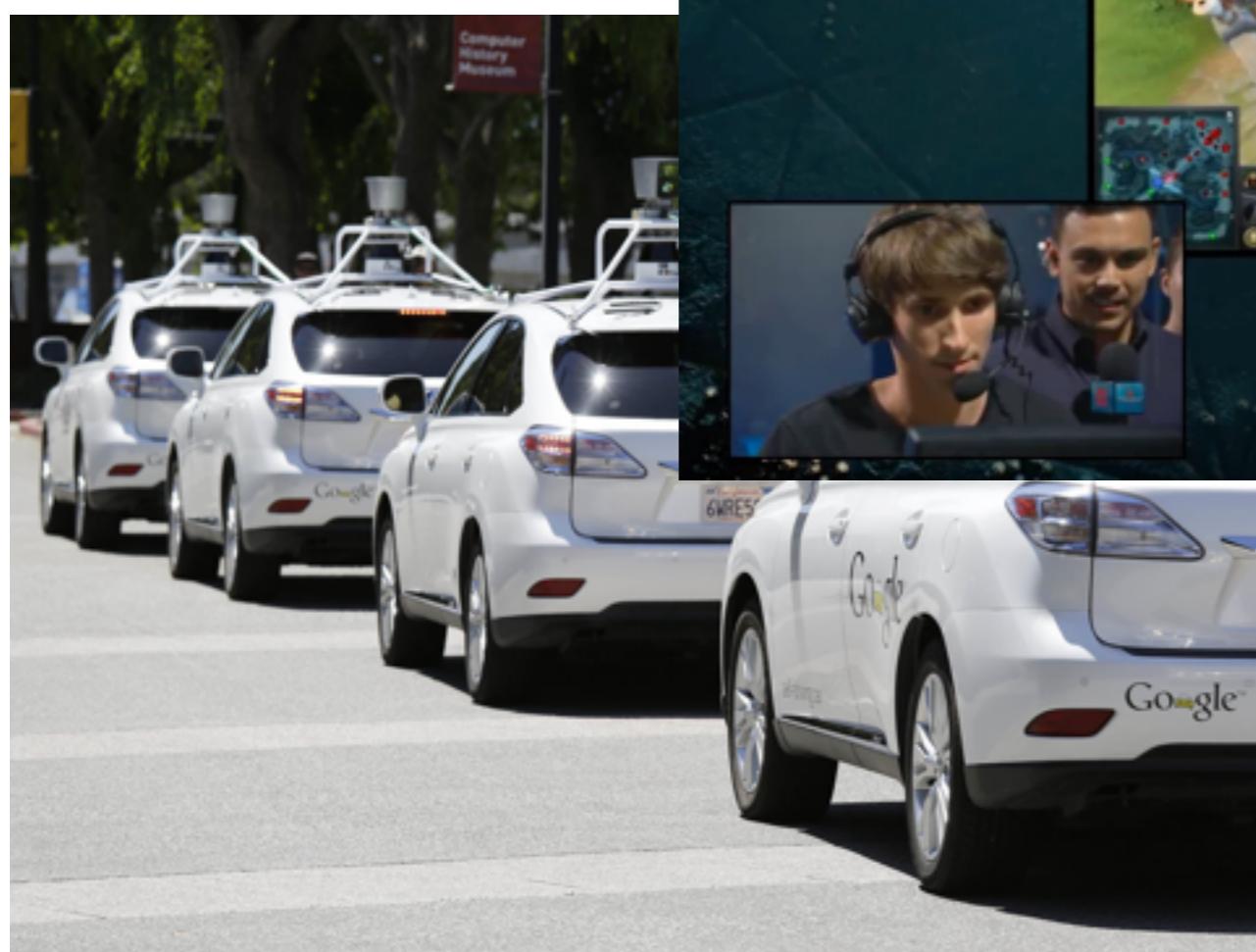
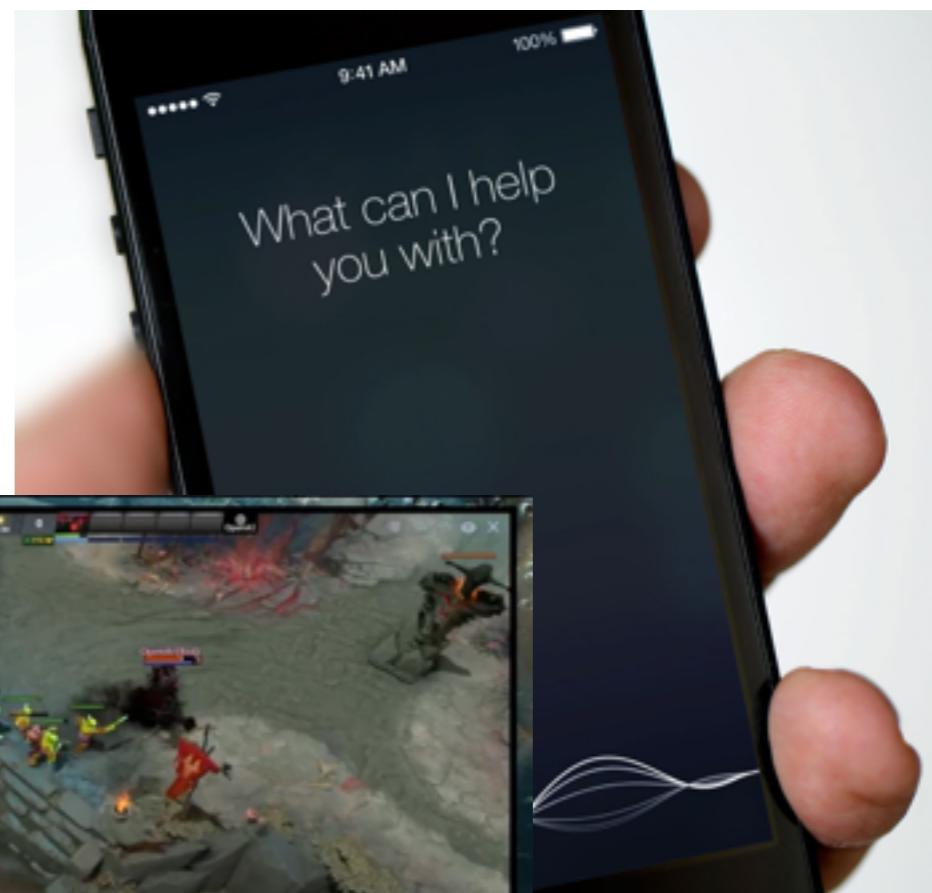
in the last decade!











Disclaimer

- This course is taught for the 1st time @ USC. This course is 599, and thus an **experimental** course.
- The syllabus, course policy, and grading details **may change** over the semester (**check website!**)
- If you prefer a well-structured course, this is **NOT** a course for you, and I encourage you to take the course next year. We really mean this.
- It will be fun but **demanding** and challenging!

Should be **already** familiar with Machine Learning

Do you know the following..?

- Probability and Statistical Learning
 - Density function, loss function, cross-validation

Should be **already** familiar with Machine Learning

Do you know the following..?

- Probability and Statistical Learning
 - Density function, loss function, cross-validation
- Supervised Learning
 - Nearest Neighbor, Kernels, Random Forest

Should be **already** familiar with Machine Learning

Do you know the following..?

- Probability and Statistical Learning
 - Density function, loss function, cross-validation
- Supervised Learning
 - Nearest Neighbor, Kernels, Random Forest
- Unsupervised Learning
 - Clustering, PCA, SVD

Should be **already** familiar with Machine Learning

- If not, please take other ML courses first!
 - For example, CSCI 567: Machine Learning

Today's agenda

- A brief introduction to Deep Learning
- Survey
- CSCI 599 overview

Today's agenda

- A brief introduction to Deep Learning
- Survey
- CSCI 599 overview

Deep Learning is impacting everywhere

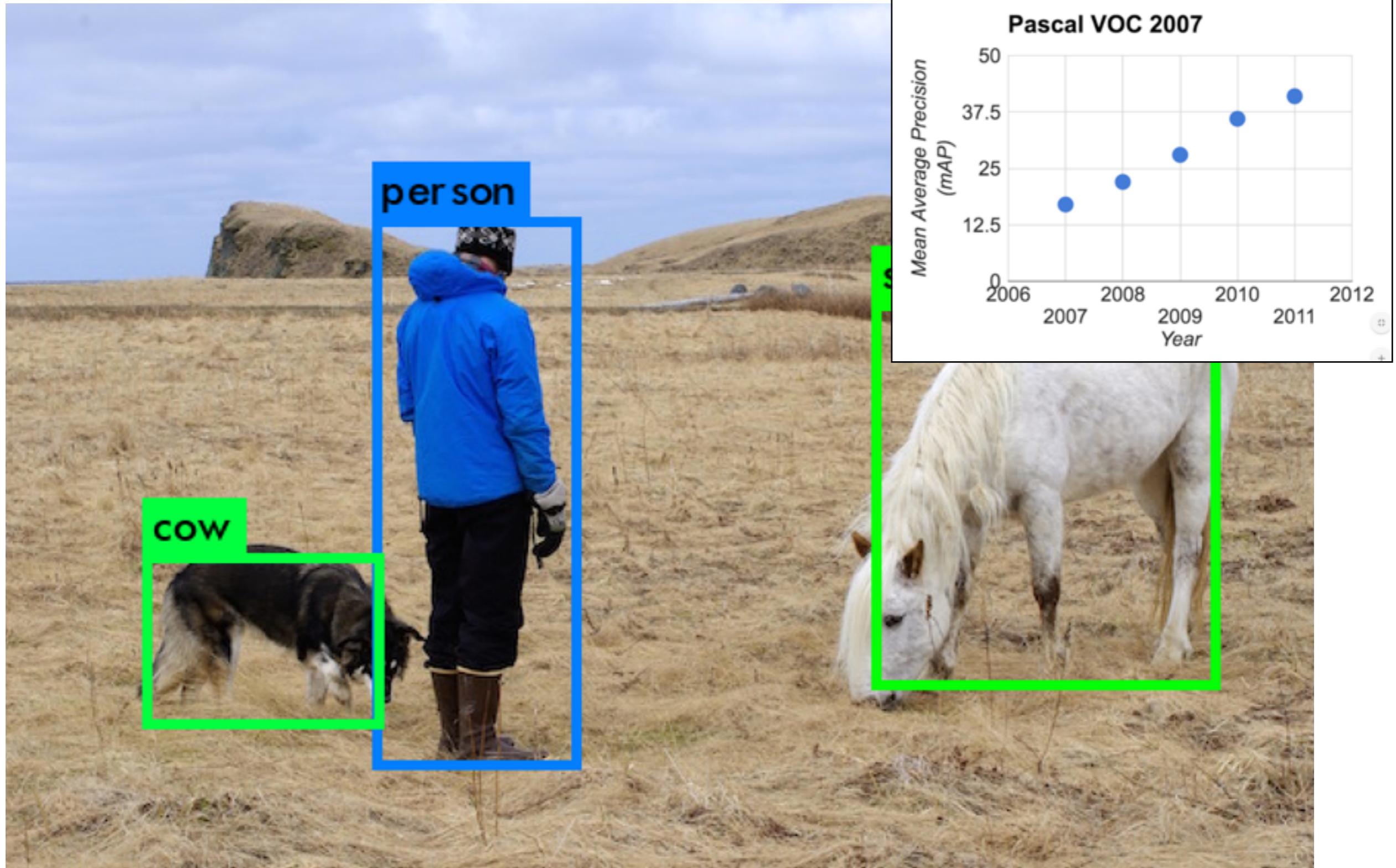
- Machine Learning
- Computer Vision
- Natural Language Processing
- Robotics
- Medical Application
- Graphics
- Finance
- and many more

Art Creation



Elgammal, Ahmed, et. al. CAN: Creative Adversarial Networks, Generating "Art" by Learning About Styles and Deviating from Style Norms. arXiv 2016.

Object Detection



Object Detection in Video



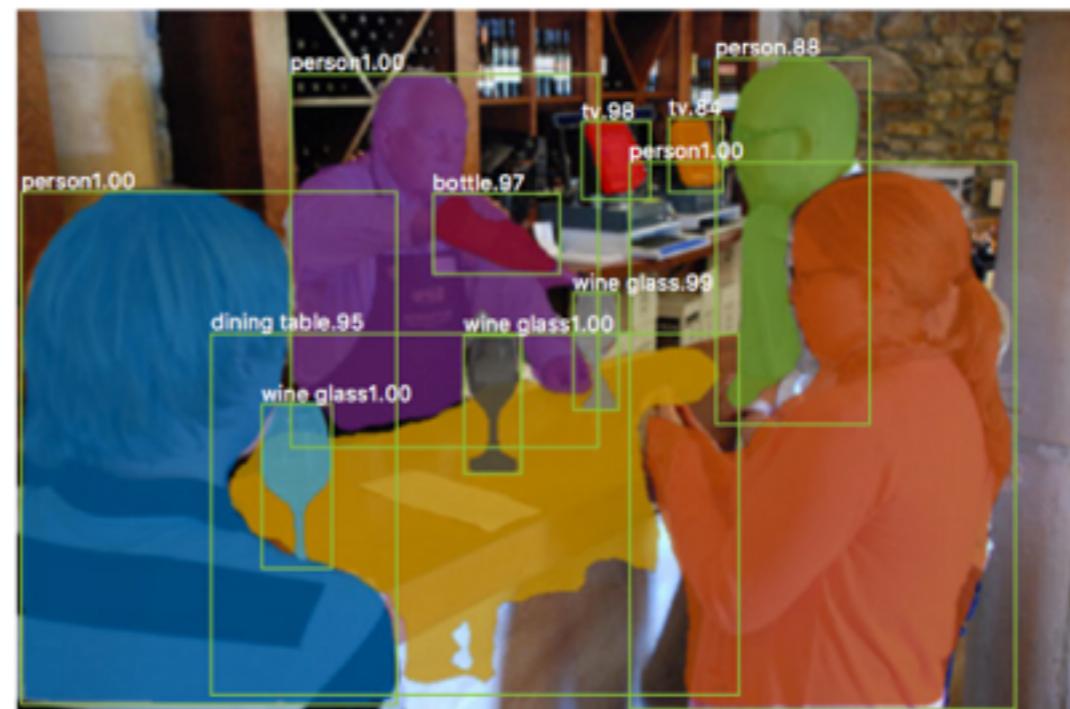
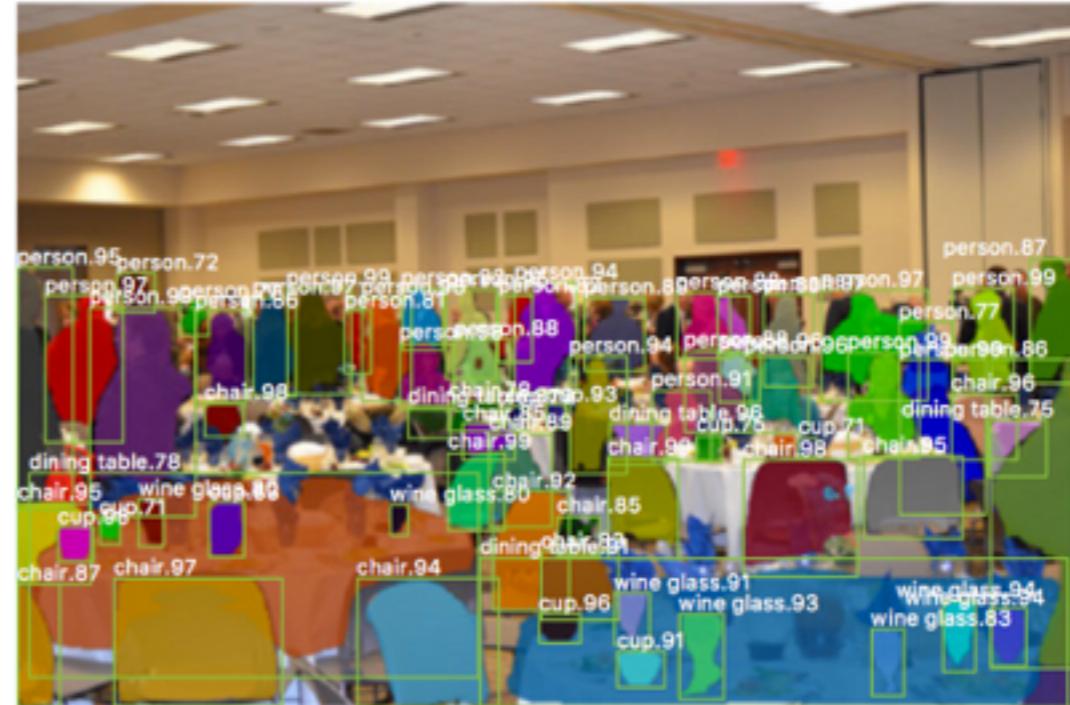
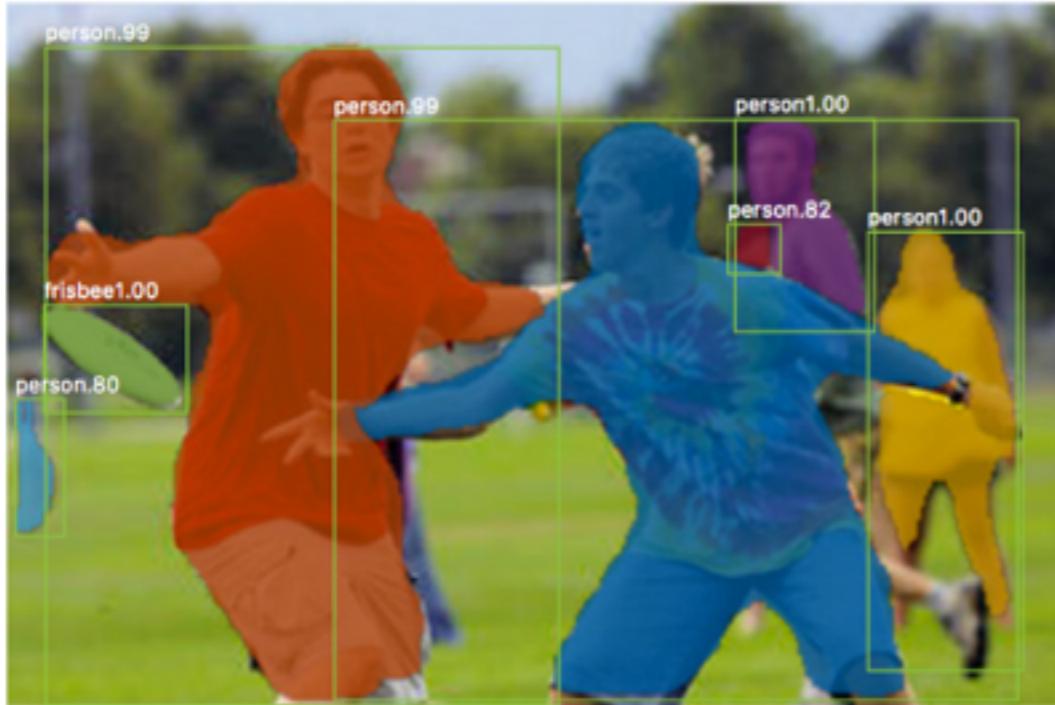
J. Redmon and A. Farhadi. YOLO9000: Better, Faster, Stronger. CVPR 2017.

Object Detection in Video



J. Redmon and A. Farhadi. YOLO9000: Better, Faster, Stronger. CVPR 2017.

Semantic Segmentation



3D Pose for Furniture



J. Lim, et. al. Parsing IKEA Objects: Fine Pose Estimation. ICCV 2013.

3D Pose for Furniture



J. Lim, et. al. Parsing IKEA Objects: Fine Pose Estimation. ICCV 2013.

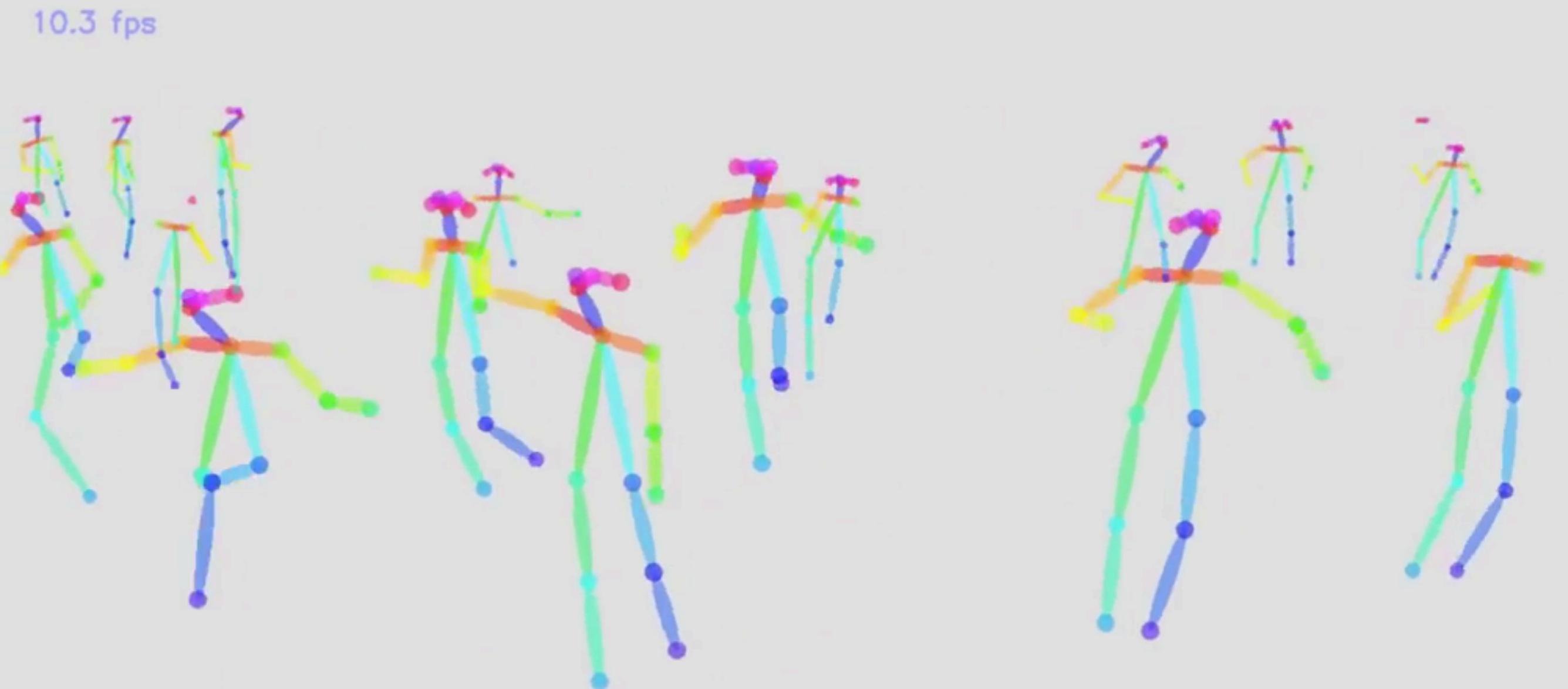
Human Pose

10.3 fps



Z. Cao, et. al. Realtime Multi-person 2D Pose Estimation using Part Affinity Fields. CVPR 2017.

Human Pose



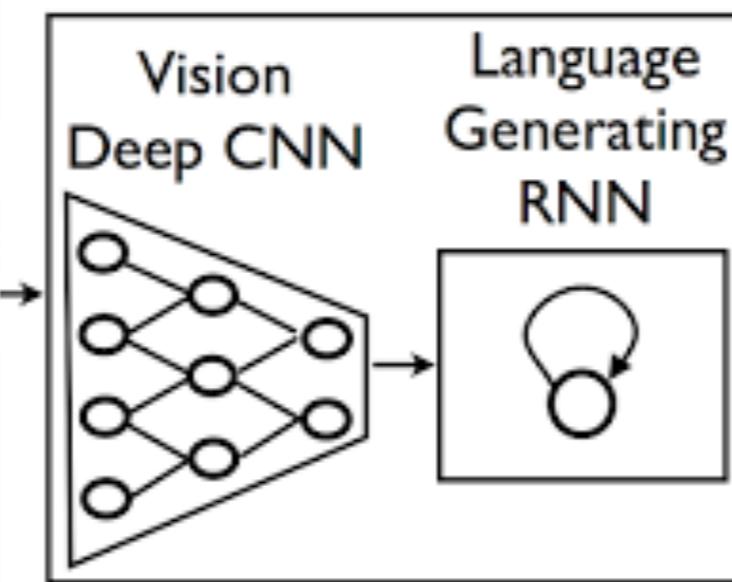
Z. Cao, et. al. Realtime Multi-person 2D Pose Estimation using Part Affinity Fields. CVPR 2017.

Image to Caption



From <https://research.googleblog.com/2014/11/a-picture-is-worth-thousand-coherent.html>

Image to Caption



A group of people shopping at an outdoor market.

There are many vegetables at the fruit stand.

From <https://research.googleblog.com/2014/11/a-picture-is-worth-thousand-coherent.html>

Caption to image

Caption

the flower shown has yellow anther red pistil and bright red petals

this flower has petals that are yellow, white and purple and has dark lines

the petals on this flower are white with a yellow center

Caption to image

Caption	Generated Images
the flower shown has yellow anther red pistil and bright red petals	
this flower has petals that are yellow, white and purple and has dark lines	
the petals on this flower are white with a yellow center	

Reed, Scott, et. al. Generative Adversarial Text to Image Synthesis. ICML 2016.

Visual Question Answering (VQA)



Who is holding the kite?

Submit

From <http://www.visualqa.org>

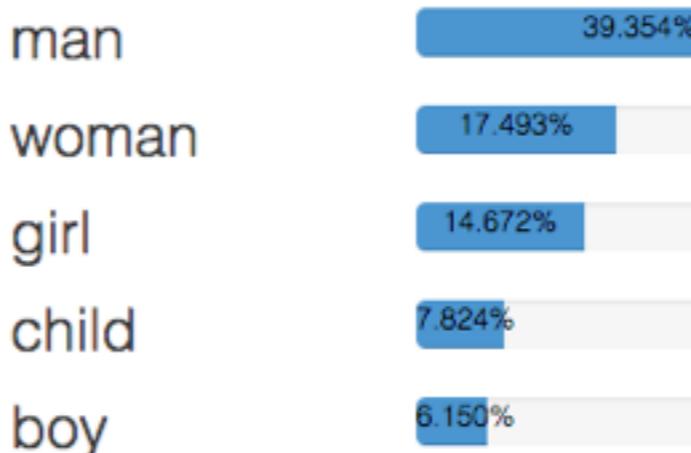
Visual Question Answering (VQA)



Who is holding the kite?

Submit

Predicted top-5 answers with confidence:



From <http://www.visualqa.org/>

Machine Translation

Korean English Chinese (Simplified) ▾ Translate

CS599 will be a fun class! CS599는 재미있는 수업이 될 것입니다!

26/5000 Suggest an edit

Korean English Chinese (Simplified) ▾ Translate

CS599 será una clase divertida! CS599将是一个有趣的课!

Suggest an edit

Korean English Chinese (Simplified) ▾ Translate

CS599 sera une classe amusante! CS599 wird eine lustige Klasse!

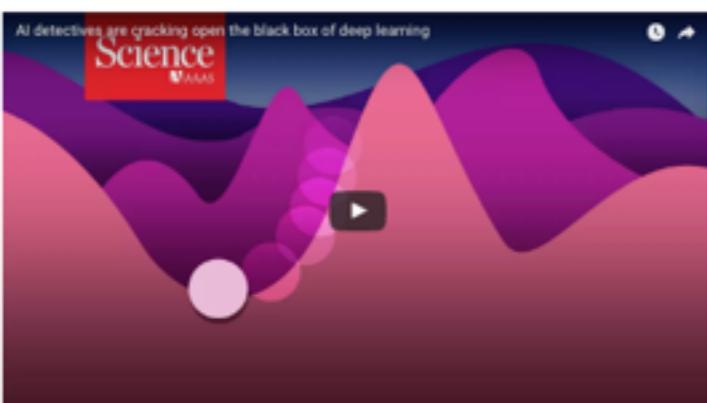
Suggest an edit

Google Translate

Machine Translation



The screenshot shows the homepage of the Science AAAS website. At the top, there are links for 'Log in', 'My account', and 'Contact us'. To the right is a 'Science' logo with a lemur image and options to 'Become a member', 'Renew my subscription', and 'Sign up for newsletters'. Below this is a navigation bar with 'Authors', 'Members', 'Librarians', and 'Advertisers'. The main menu includes 'Home', 'News', 'Journals', 'Topics', and 'Careers'. A search bar is on the right. Below the menu, categories like 'Latest News', 'ScienceInsider', 'ScienceShots', 'Sitter', 'From the Magazine', 'About News', and 'Quizzes' are listed. A 'SHARE' button on the left has icons for Facebook (3K), Twitter (90), and LinkedIn (1K). A large video thumbnail on the left features a purple and pink abstract design with a play button.



How AI detectives are cracking open the black box of deep learning

By Paul Voosen | Jul. 6, 2017, 2:00 PM

Jason Yosinski sits in a small glass box at Uber's San Francisco, California, headquarters, pondering the mind of an artificial intelligence. An Uber research scientist, Yosinski is performing a kind of brain surgery on the AI running on his laptop. Like many of the AIs that will soon be powering so much of modern life, including self-driving Uber cars, Yosinski's program is a deep neural network, with an architecture loosely inspired by the brain. And like the brain, the program is hard to understand from the outside: It's a black box.

This particular AI has been trained, using a vast sum of labeled images, to recognize objects as random as zebras, fire trucks, and seat belts. Could it recognize Yosinski and the reporter hovering in front of the webcam? Yosinski zooms in on one of the AI's individual computational nodes—the neurons, so to speak—to see what is prompting its response. Two ghostly white ovals pop up and float on the screen. This neuron, it seems, has learned to detect the outlines of faces. "This responds to your face and my face," he says. "It responds to different size faces, different color faces."

No one trained this network to identify faces. Humans weren't labeled in its training images. Yet learn faces it did, perhaps as a way to recognize the things that tend to accompany them, such as ties and cowboy hats. The network is too complex for humans to comprehend its exact decisions. Yosinski's probe had illuminated one small part of it, but overall, it remained opaque. "We build amazing models," he says. "But we don't quite understand them. And every year, this gap is going to get a bit larger."

SIGN UP FOR OUR DAILY NEWSLETTER

Get more great content like this delivered right to you!

Email Address

Sign Up

By providing your email address, you agree to send your email address to the publication.
Information provided here is subject to Science's [Privacy Policy](#).

CRACKING THE MYSTERY OF
FOOD ALLERGY

CSCI 599 @ USC

English

Google Translate

Joseph J. Lim

Lecture 1

Machine Translation

The screenshot shows the homepage of Science AAAS. At the top right, there are links for 'Log in', 'My account', and 'Contact us'. Below that is a 'Become a member' section with links for 'Renew my subscription', 'Sign up for newsletters', and social media icons. The main navigation bar includes 'Home', 'News', 'Journals', 'Topics', 'Careers', 'Authors', 'Members', 'Librarians', and 'Advertisers'. A search bar is at the top right. Below the navigation is a red banner with the text 'Latest News', 'ScienceInsider', 'ScienceShots', 'Sifter', 'From the Magazine', 'About News', and 'Quizzes'. The main content area features a large image of a brain with the text 'AI detectives are cracking open the black box of deep learning'. Below the image is the title 'How AI detectives are cracking open the black box of deep learning' by Paul Voosen on July 6, 2017, 2:00 PM. The article discusses how AI is being used to understand the inner workings of deep neural networks.

The screenshot shows a Google Translate interface. The URL in the address bar is 'http://www.sciencemag.org/news/2017/07/how-ai-detectives-are-cracking-open-b'. A message says 'This page was not retrieved from its original location over a secure connection.' There are buttons for 'Translate', 'From: English', 'To: French', and 'Dismiss'. Below the interface is a 'View' button with options for 'Translation' and 'Original'.

A large black arrow points from the Science AAAS article to the Google Translate interface, indicating the process of translating the article.

How AI detectives are cracking open the black box of deep learning
By Paul Voosen | Jul. 6, 2017, 2:00 PM

Jason Yosinski sits in a small glass box at Uber's San Francisco, California, headquarters, pondering the mind of an artificial intelligence. An Uber research scientist, Yosinski is performing a kind of brain surgery on the AI running on his laptop. Like many of the AIs that will soon be powering so much of modern life, including self-driving Uber cars, Yosinski's program is a deep neural network, with an architecture loosely inspired by the brain. And like the brain, the program is hard to understand from the outside: It's a black box.

This particular AI has been trained, using a vast sum of labeled images, to recognize objects as random as zebras, fire trucks, and seat belts. Could it recognize Yosinski and the reporter hovering in front of the webcam? Yosinski zooms in on one of the AI's individual computational nodes—the neurons, so to speak—to see what is prompting its response. Two ghostly white ovals pop up and float on the screen. This neuron, it seems, has learned to detect the outlines of faces. "This responds to your face and my face," he says. "It responds to different size faces, different color faces."

No one trained this network to identify faces. Humans weren't labeled in its training images. Yet learn faces it did, perhaps as a way to recognize the things that tend to accompany them, such as ties and cowboy hats. The network is too complex for humans to comprehend its exact decisions. Yosinski's probe had illuminated one small part of it, but overall, it remained opaque. "We build amazing models," he says. "But we don't quite understand them. And every year, this gap is going to get a bit larger."

SIGN UP FOR OUR DAILY NEWSLETTER
Get more great content like this delivered right to you!
Email Address Sign Up
By providing your email address, you agree to send your email address to the publication.
Information provided here is subject to Science's [Privacy Policy](#).

English

Google Translate

Machine Translation

The screenshot shows the main navigation bar with links for Home, News, Journals, Topics, Careers, and a search bar. Below the navigation is a large article thumbnail titled "AI detectives are cracking open the black box of deep learning". To the right of the thumbnail is a sidebar with a "SHARE" section and a note about metered access for AAAS members.

The screenshot shows the same page as above, but the content has been translated into French by Google Translate. The main title is now "Comment les détectives d'IA créent-ils la boîte noire d'apprentissage approfondi?". The sidebar and navigation links are also in French.

How AI detectives are cracking open the black box of deep learning
By Paul Voosen | Jul. 6, 2017, 2:00 PM

Jason Yosinski sits in a small glass box at Uber's San Francisco, California, headquarters, pondering the mind of an artificial intelligence. An Uber research scientist, Yosinski is performing a kind of brain surgery on the AI running on his laptop. Like many of the AIs that will soon be powering so much of modern life, including self-driving Uber cars, Yosinski's program is a deep neural network, with an architecture loosely inspired by the brain. And like the brain, the program is hard to understand from the outside: It's a black box.

This particular AI has been trained, using a vast sum of labeled images, to recognize objects as random as zebras, fire trucks, and seat belts. Could it recognize Yosinski and the reporter hovering in front of the webcam? Yosinski zooms in on one of the AI's individual computational nodes—the neurons, so to speak—to see what is prompting its response. Two ghostly white ovals pop up and float on the screen. This neuron, it seems, has learned to detect the outlines of faces. "This responds to your face and my face," he says. "It responds to different size faces, different color faces."

No one trained this network to identify faces. Humans weren't labeled in its training images. Yet learn faces it did, perhaps as a way to recognize the things that tend to accompany them, such as ties and cowboy hats. The network is too complex for humans to comprehend its exact decisions. Yosinski's probe had illuminated one small part of it, but overall, it remained opaque. "We build amazing models," he says. "But we don't quite understand them. And every year, this gap is going to get a bit larger."

SIGN UP FOR OUR DAILY NEWSLETTER

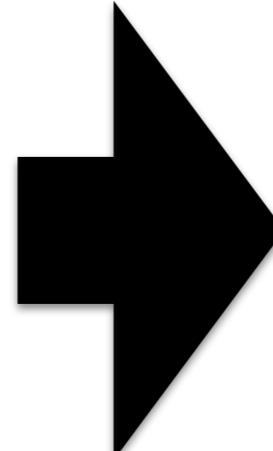
Get more great content like this delivered right to you!

Email Address

Sign Up

By providing your email address, you agree to send your email address to the publication.
Information provided here is subject to Science's [Privacy Policy](#).

English



Comment les détectives d'IA créent-ils la boîte noire d'apprentissage approfondi?

Par Paul Voosen 6 juillet 2017, 14h00

Jason Yosinski est assis dans une petite boîte en verre dans le quartier général d'Uber à San Francisco, en Californie, en considérant l'esprit d'une intelligence artificielle. Un scientifique de recherche d'Uber, Yosinski effectue une sorte de chirurgie du cerveau sur l'IA qui court sur son ordinateur portable. Comme beaucoup d'AI qui vont bientôt alimenter tant de la vie moderne, y compris les voitures Uber auto-conduisant, le programme de Yosinski est un réseau neuronal profond, avec une architecture peu inspirée par le cerveau. Et comme le cerveau, le programme est difficile à comprendre de l'extérieur: c'est une boîte noire.

Cette IA particulière a été formée, en utilisant une vaste somme d'images

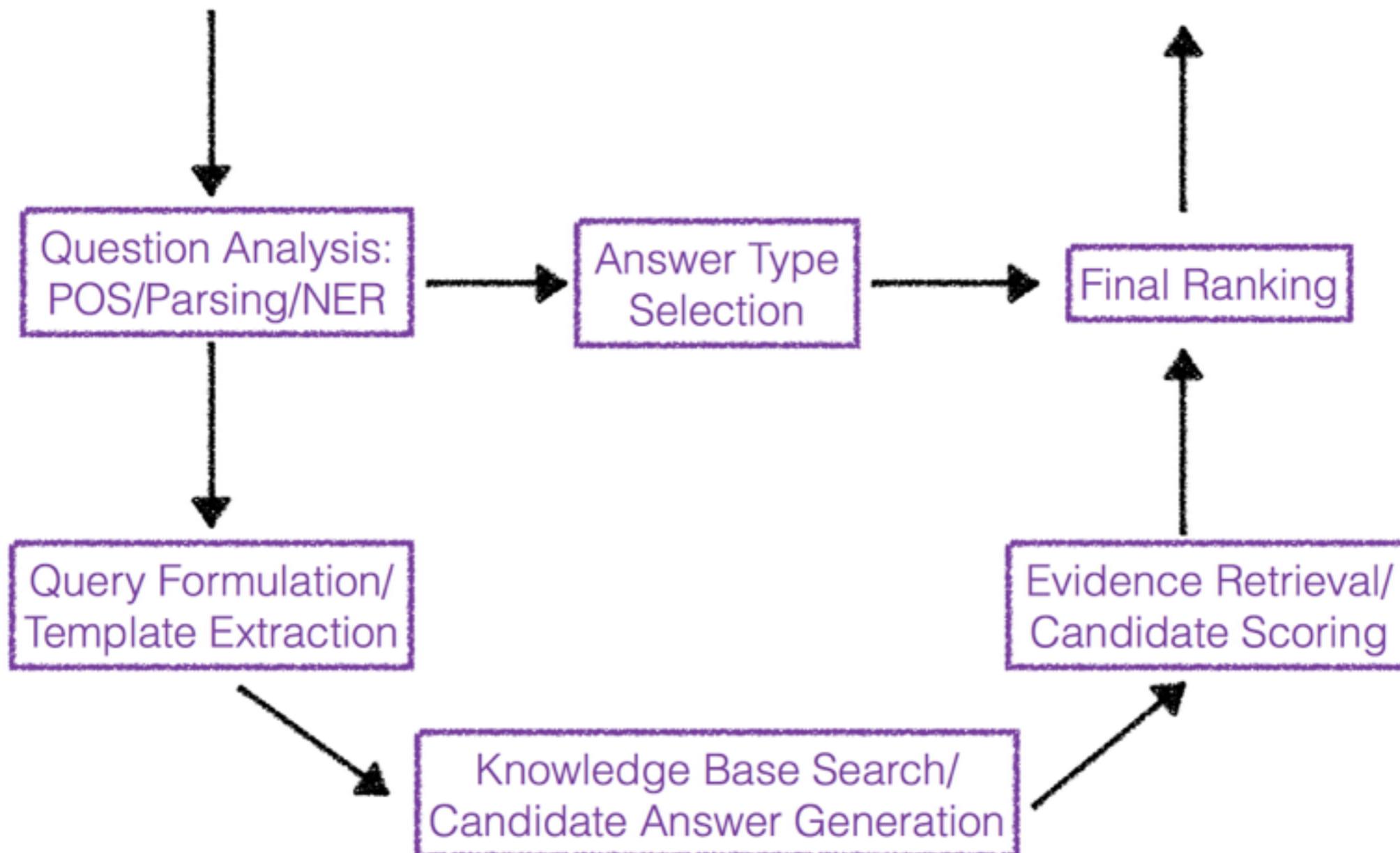
French

Google Translate

Question Answering

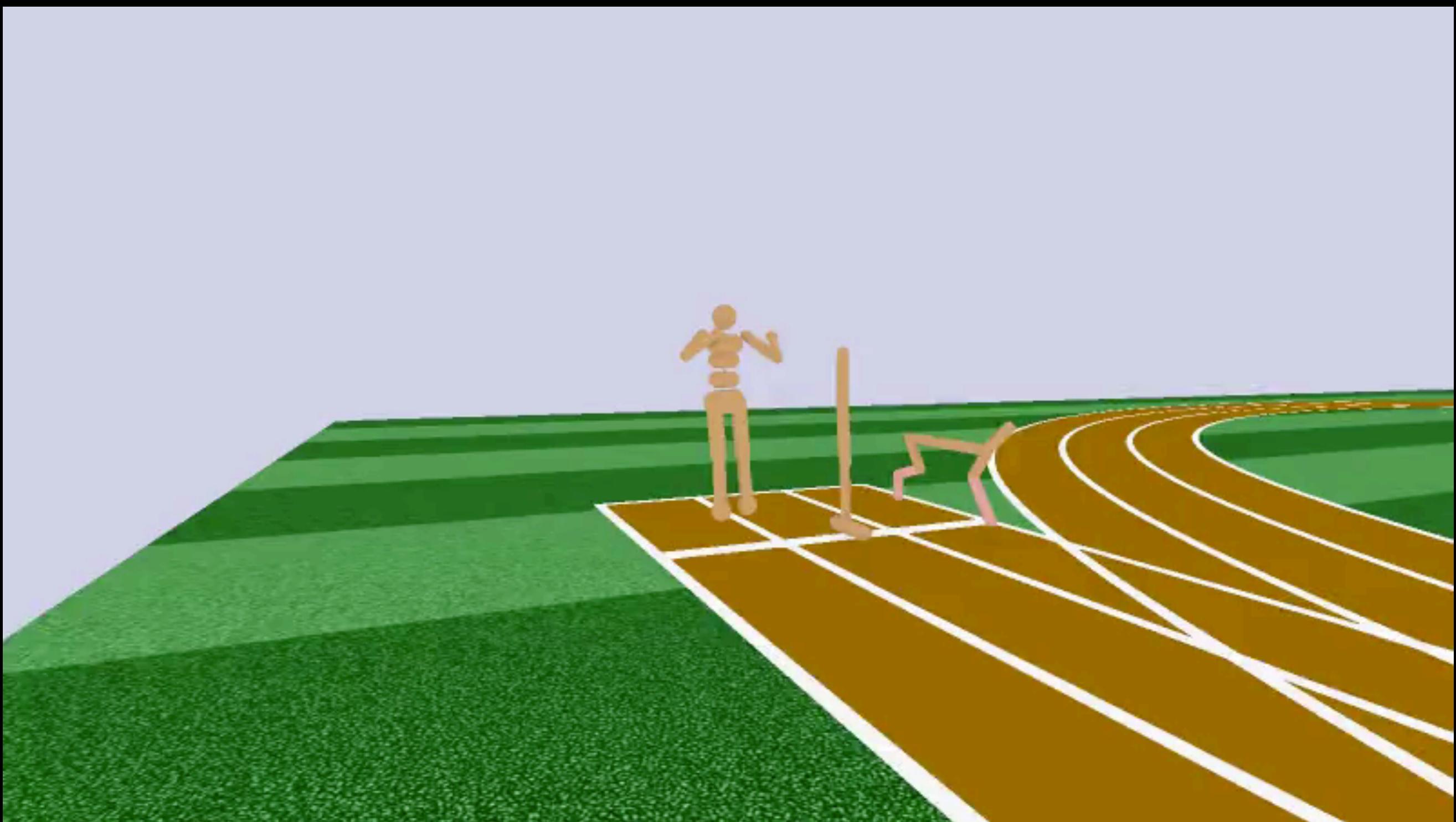
Who wrote the song
“Kiss from a Rose”?

Seal



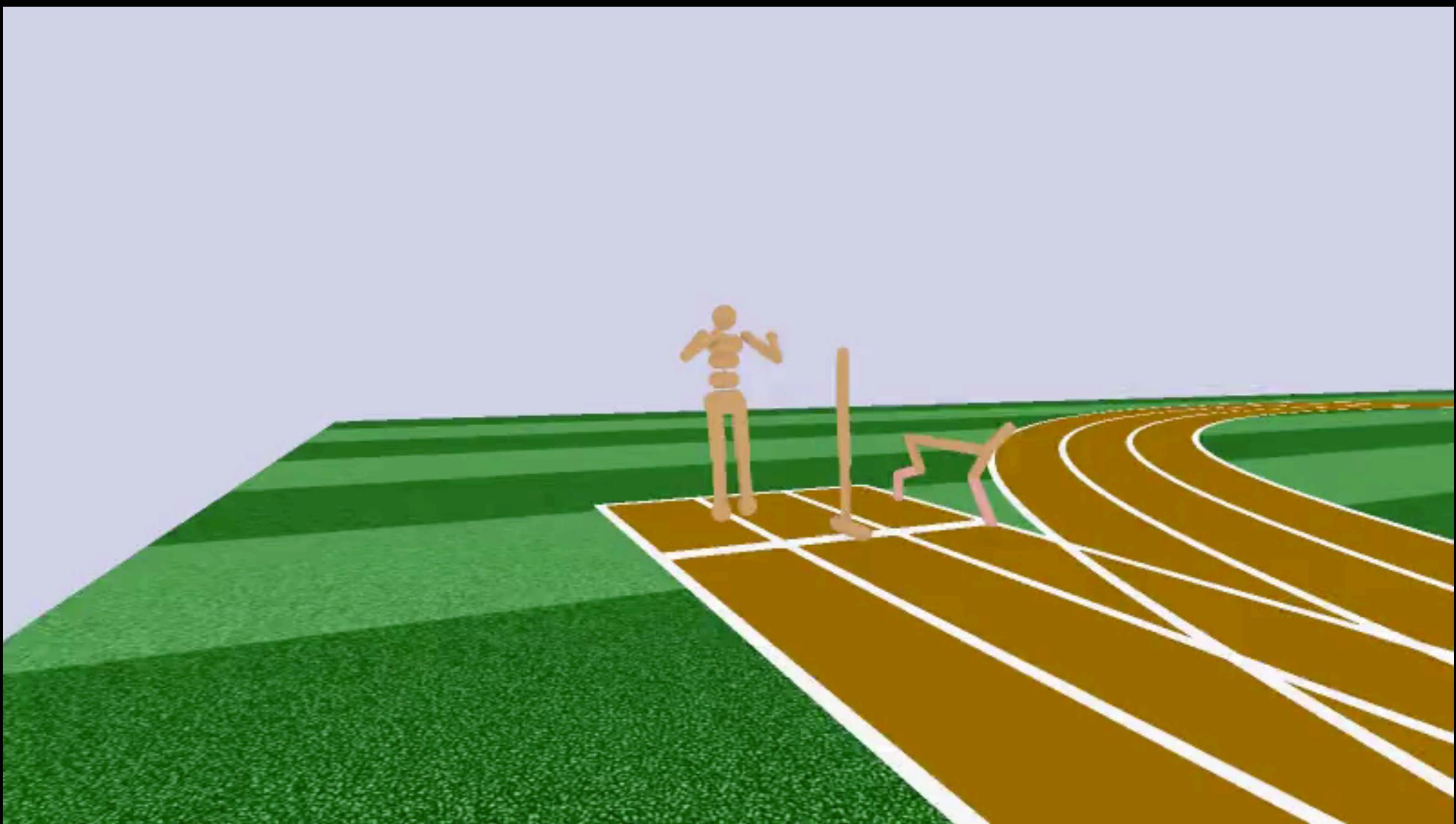
Kumar, Ankit, et. al. “Ask me anything: Dynamic memory networks for natural language processing.” ICML 2016.

Learning to Walk



From <https://blog.openai.com/roboschool/>

Learning to Walk



From <https://blog.openai.com/roboschool/>

Object Picking Robot



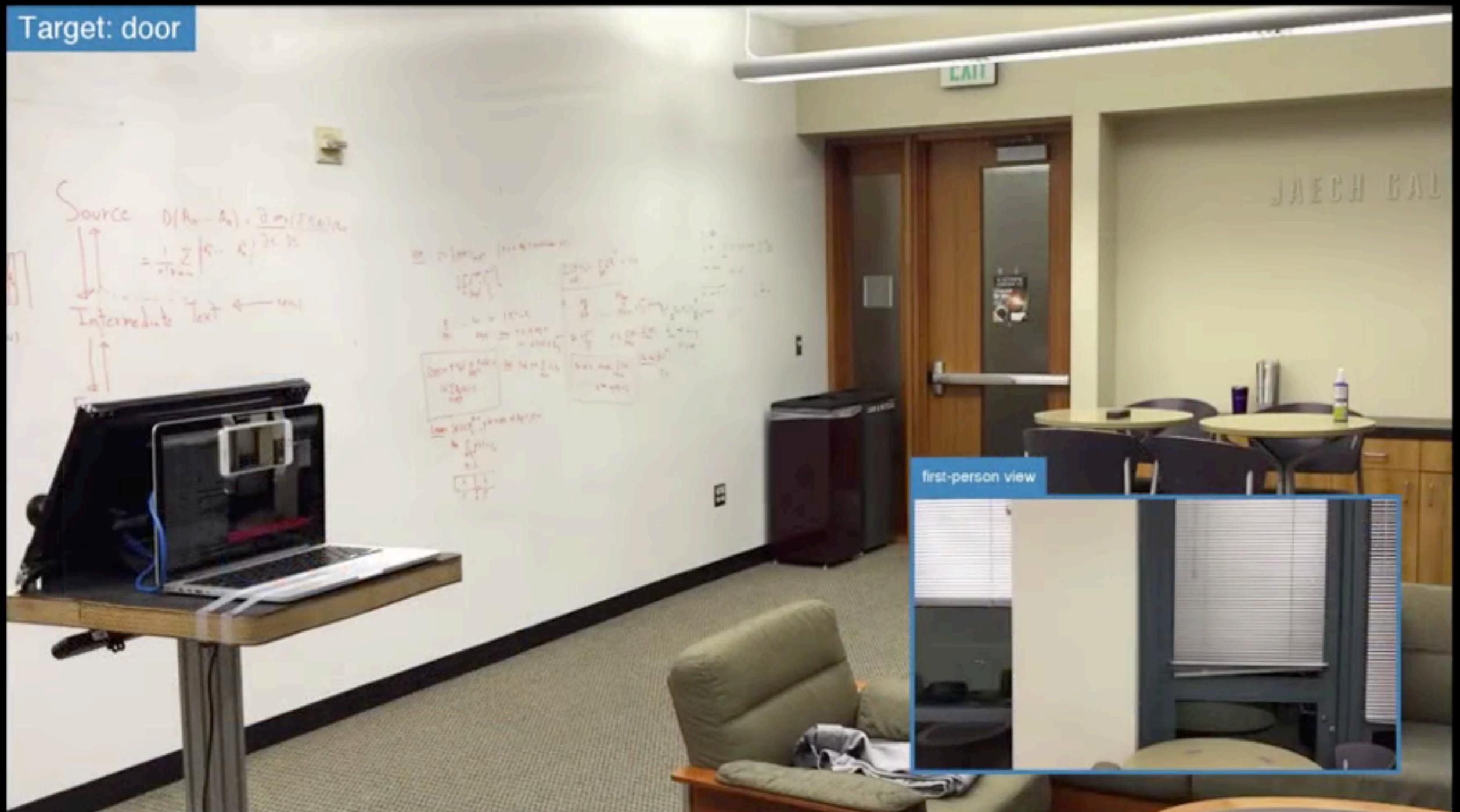
S. Levine, et. al. Learning Hand-Eye Coordination for Robotic Grasping
with Deep Learning and Large-Scale Data Collection.. IJRR 2017.

Object Picking Robot



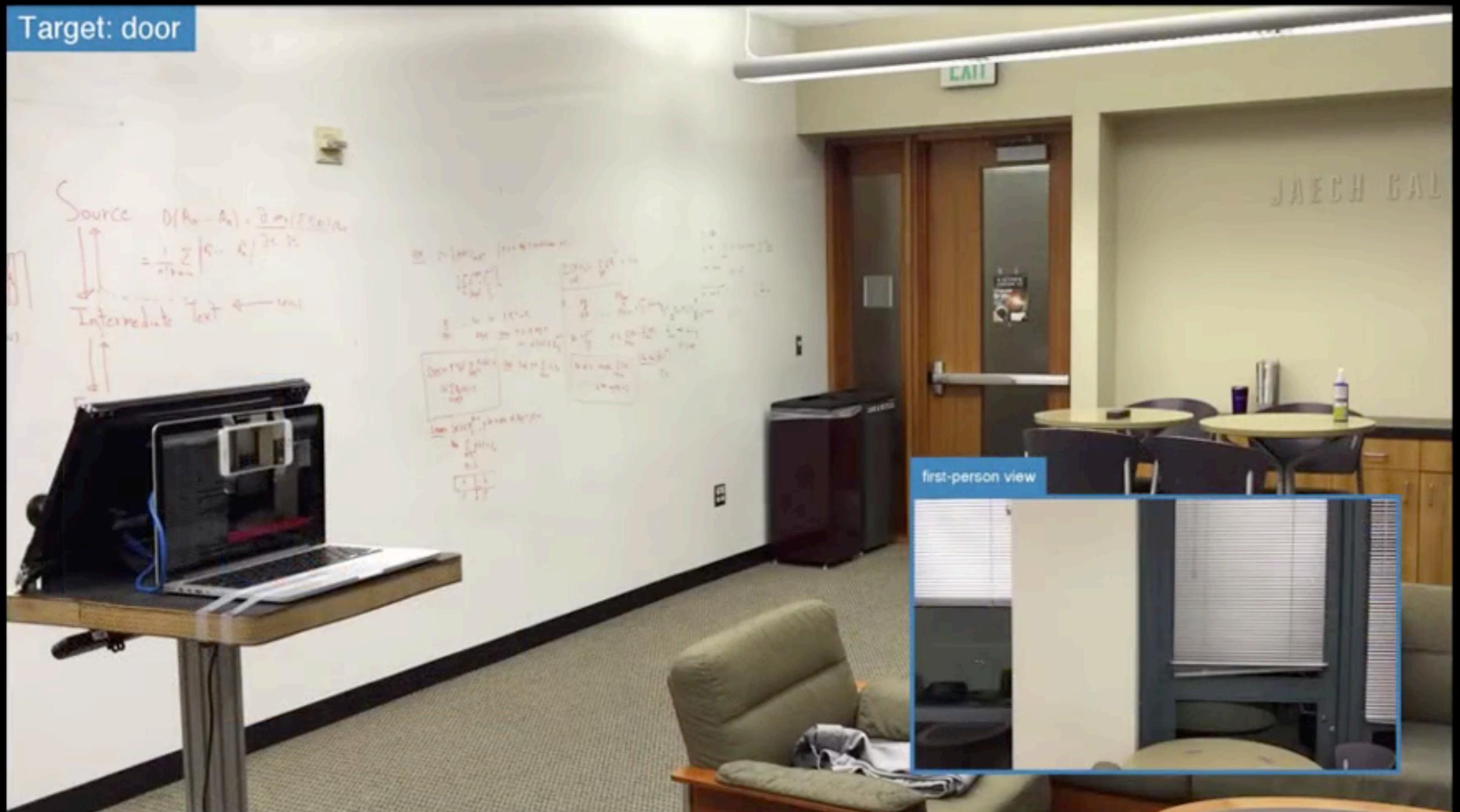
S. Levine, et. al. Learning Hand-Eye Coordination for Robotic Grasping
with Deep Learning and Large-Scale Data Collection.. IJRR 2017.

Navigation Robot



Y. Zhu, et. al. Target-driven Visual Navigation in Indoor Scenes using Deep Reinforcement Learning. ICRA 2017.

Navigation Robot



Y. Zhu, et. al. Target-driven Visual Navigation in Indoor Scenes using Deep Reinforcement Learning. ICRA 2017.

Autonomous Driving

AutoX
Democratizing autonomy



From AutoX

Autonomous Driving

AutoX
Democratizing autonomy



From AutoX



input performance



mouth camera



real-time
facial animation

K. Olszewski, **J. Lim**, S. Saito, H. Li. High-Fidelity Facial and Speech Animation for VR HMDs. SIGGRAPH Asia 2016.



input performance



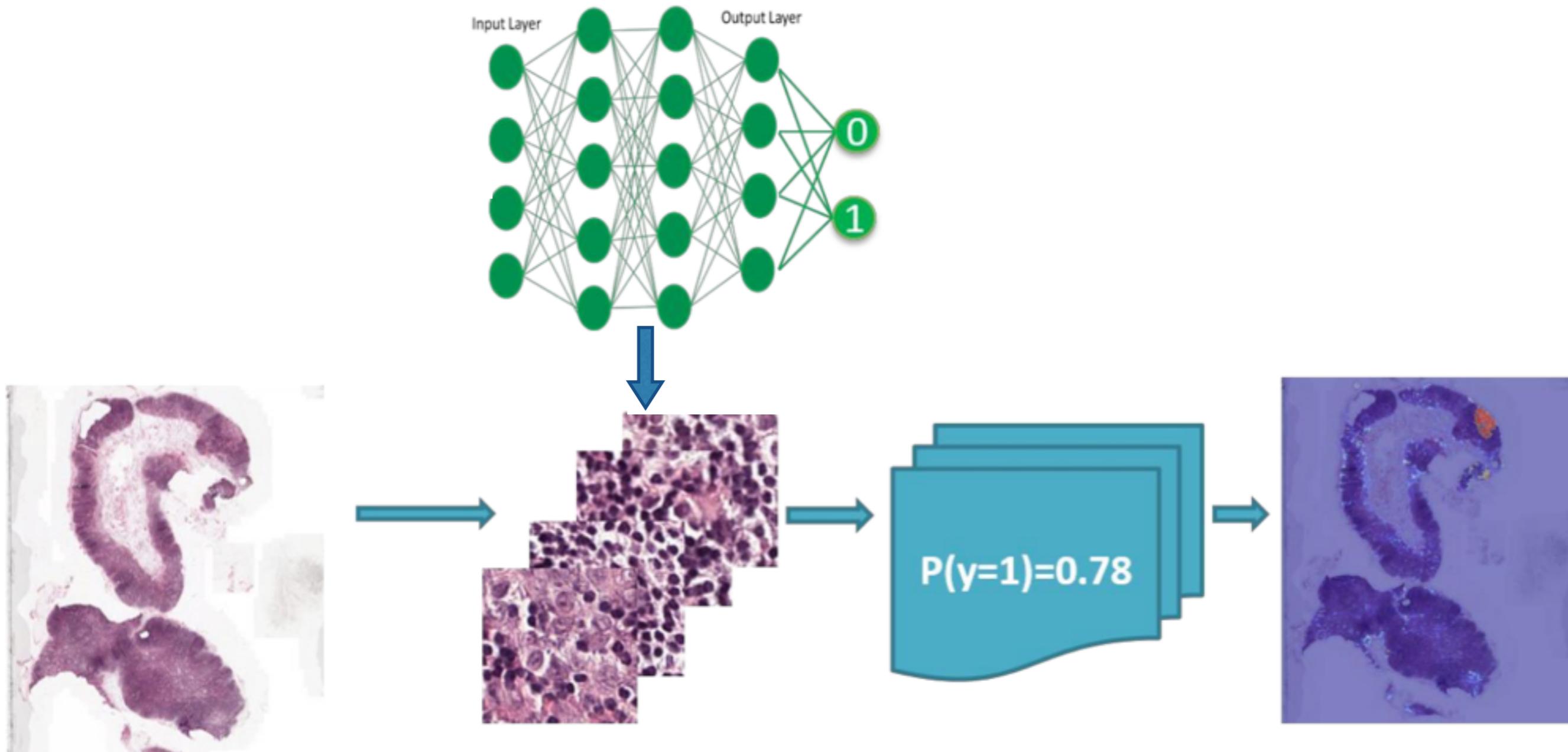
mouth camera



real-time
facial animation

K. Olszewski, **J. Lim**, S. Saito, H. Li. High-Fidelity Facial and Speech Animation for VR HMDs. SIGGRAPH Asia 2016.

Cancer Metastases Detection



From PathAI's submission to CAMELYON16.

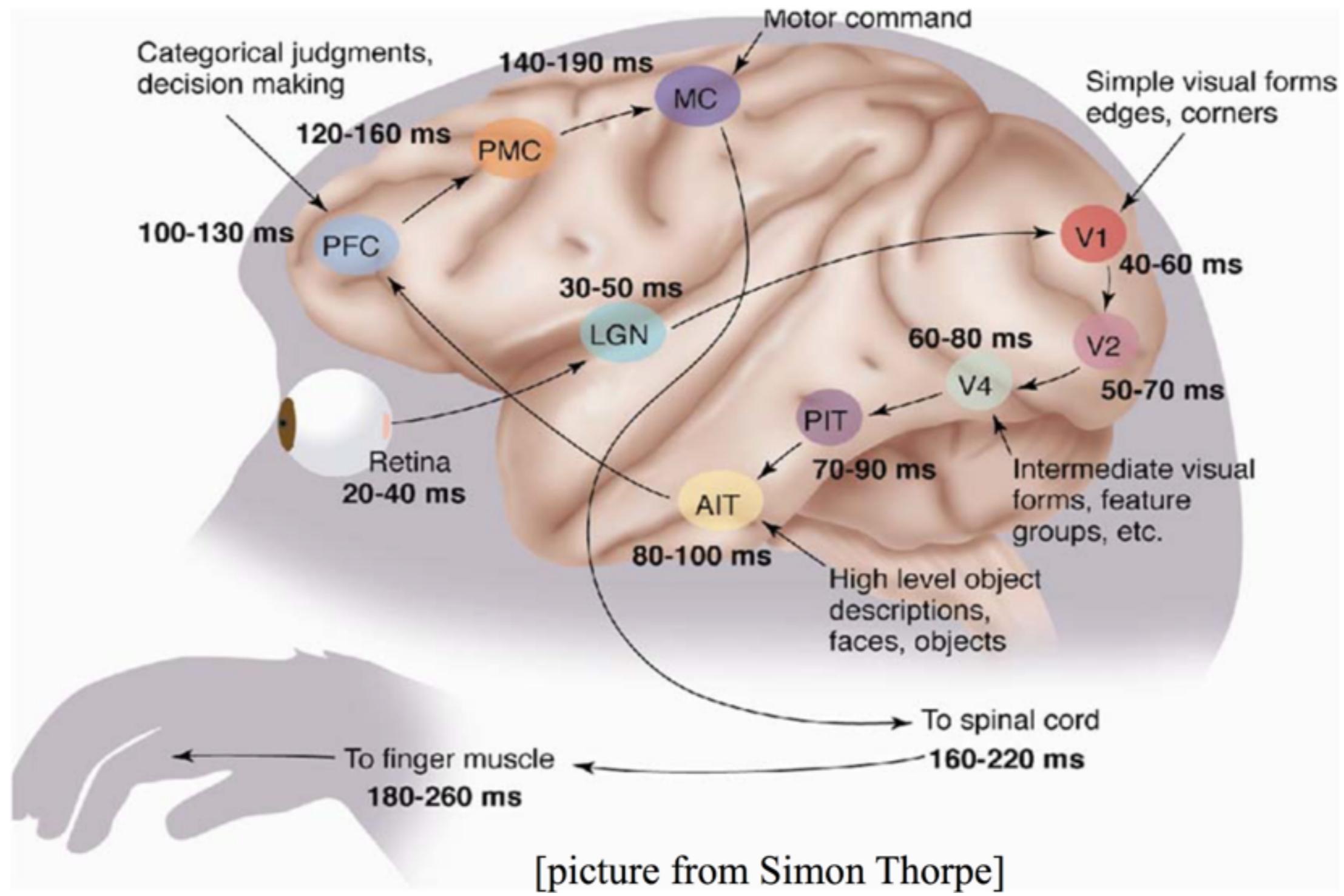
Deep Learning is impacting everywhere

- Machine Learning
- Computer Vision
- Natural Language Processing
- Robotics
- Medical Application
- Graphics
- Finance
- and many more

Why is DL so powerful?

Deep Learning is motivated by ?

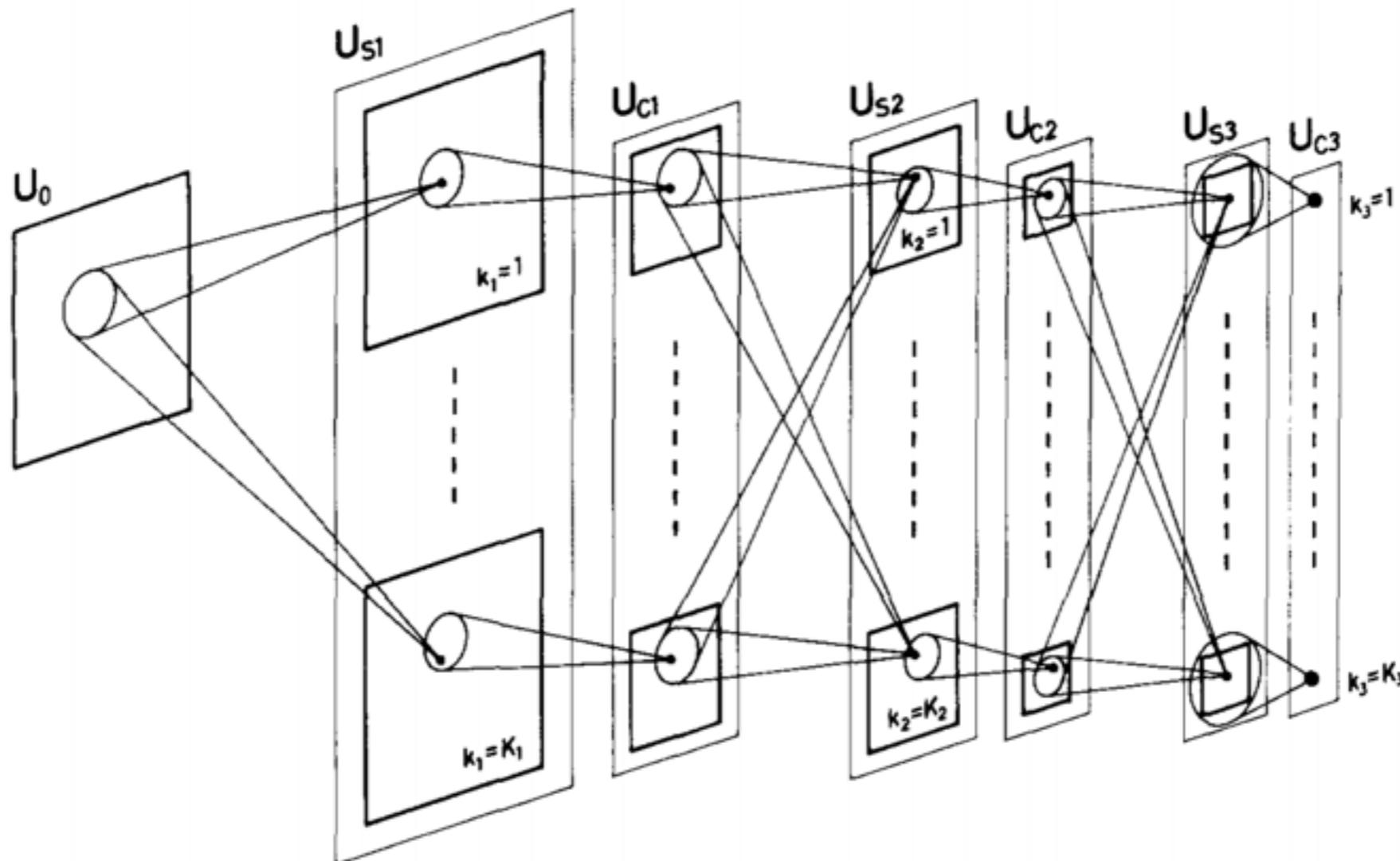
Deep Learning is motivated by human brain



Slide credit: Marc'Aurelio Ranzato, Yann LeCun

History of DL in 3 slides

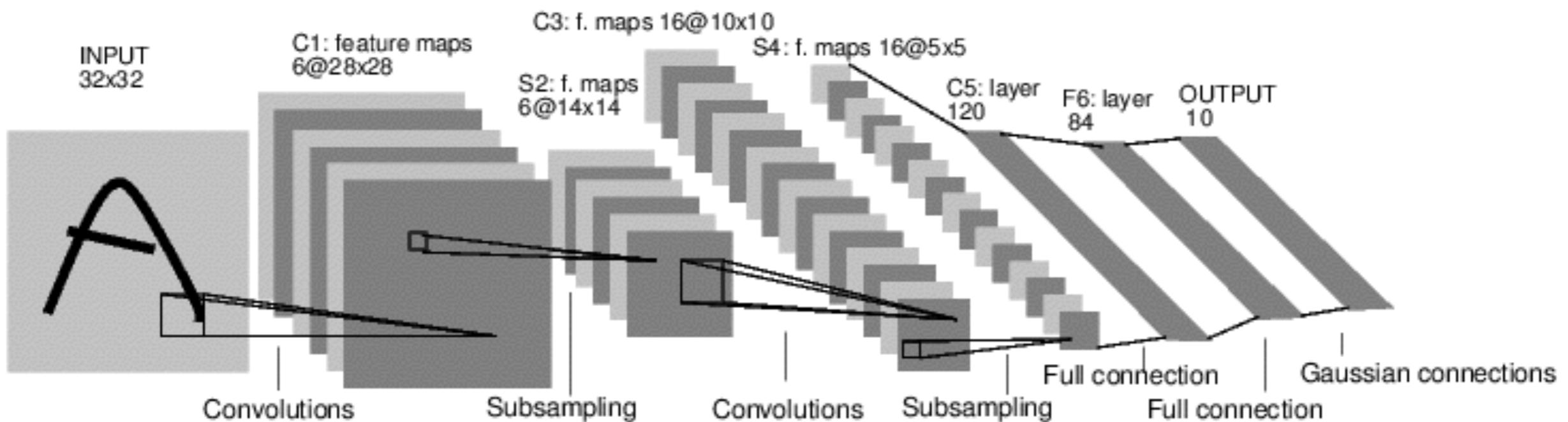
First idea



Fukushima, Kunihiko. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. Competition and cooperation in neural nets 1982.

History of DL in 3 slides

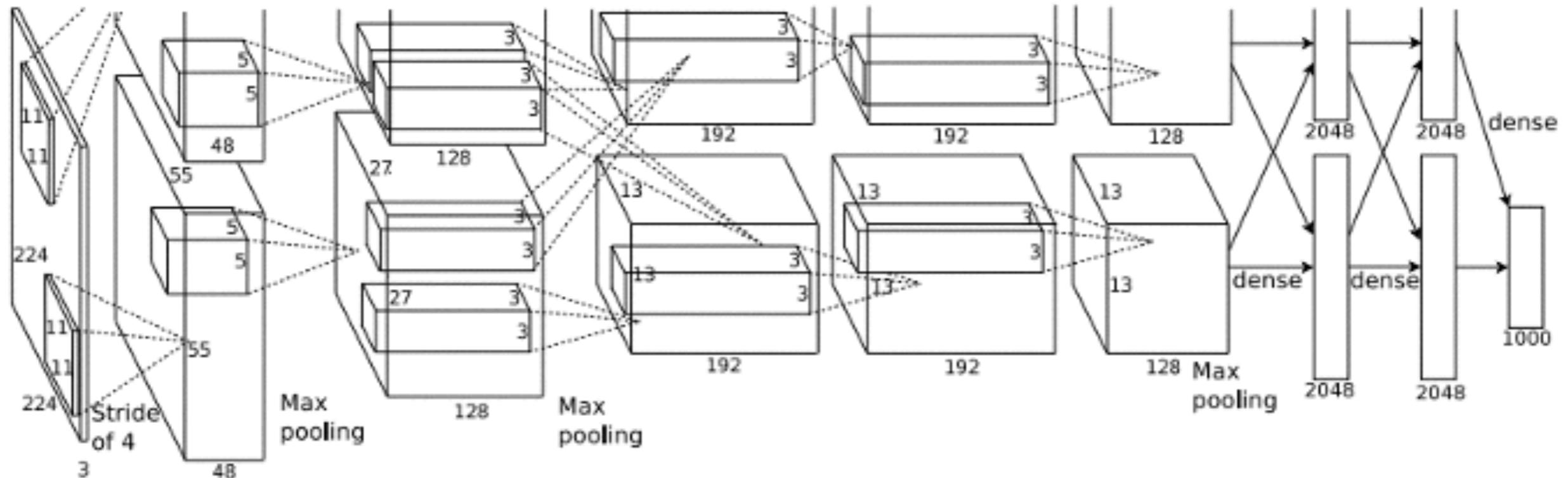
Backpropagation for training



Y. LeCun, et. al. Handwritten digit recognition with a back-propagation network.
NIPS 1989.

History of DL in 3 slides

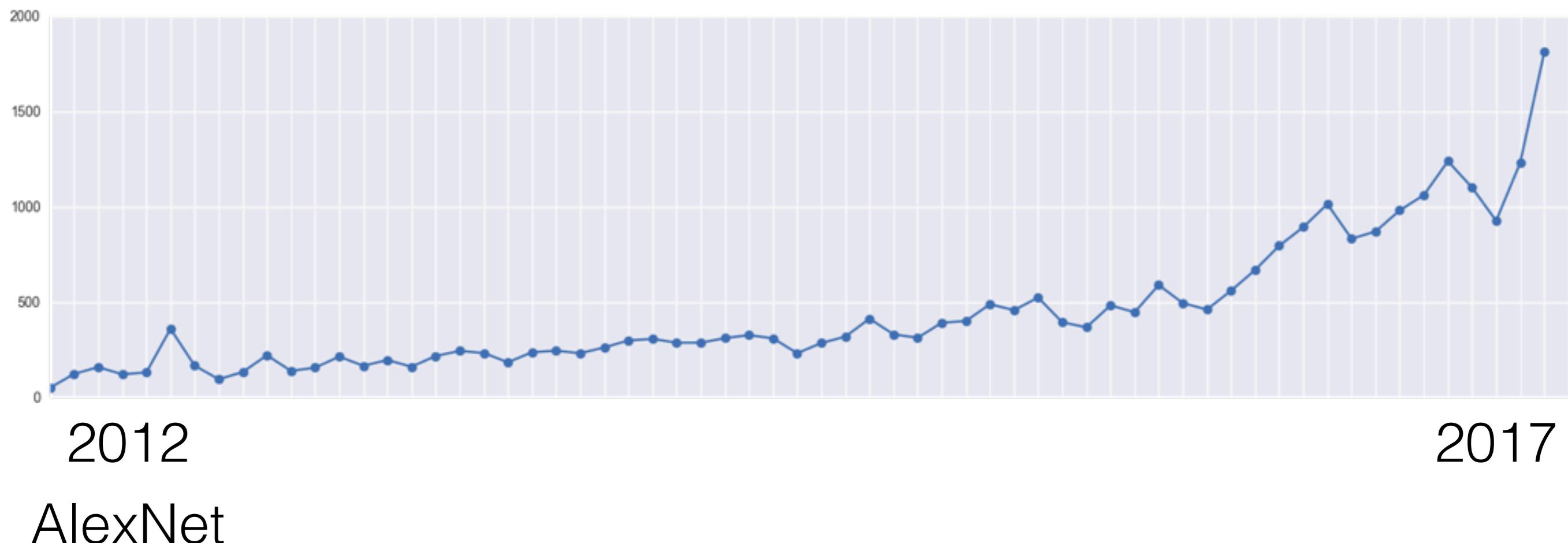
Large-scale data (& GPU)



A Krizhevsky, et. al. ImageNet Classification with Deep Convolutional Neural Networks. NIPS 2012.

Also known as, **AlexNet**

DL trend in arXiv



By Andrej Karpathy

DL trend in arXiv



By Andrej Karpathy

Deep Learning is impacting everywhere

- Machine Learning
- Computer Vision
- Natural Language Processing
- Robotics
- Medical Application
- Graphics
- Finance
- and many more

Is Artificial Intelligence solved?

What is Artificial Intelligence?

What is Artificial Intelligence?

- Artificial intelligence (AI, also machine intelligence, MI) is **intelligence exhibited by machines**, rather than humans or other animals (natural intelligence, NI).

From wikipedia

AI is **intelligence exhibited by machines**

- Object Detection
- Visual Question Answering
- Question Answering
- Autonomous Driving
- ...

Is AI solved (by deep learning)?!

AI is **intelligence exhibited by machines**

- Object Detection
- Visual Question Answering
- Question Answering
- Autonomous Driving
- ...

Is AI solved (by deep learning)?!

- **No!** There are many domains of problems that we don't even know how to tackle.

Is AI solved (by deep learning)?!

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning

Is AI solved (by deep learning)?!

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning

We are quite good

Is AI solved (by deep learning)?!

- Supervised Learning

We are quite good

- Unsupervised Learning
- Reinforcement Learning

Yet to explore more

Is AI solved (by deep learning)?!

- Supervised Learning

We are quite good

- Unsupervised Learning
- Reinforcement Learning

Yet to explore more

Specific AI

vs

General AI

Is AI solved (by deep learning)?!

- Supervised Learning

We are quite good

- Unsupervised Learning
- Reinforcement Learning

Yet to explore more

Specific AI

vs

General AI

ok..

Is AI solved (by deep learning)?!

- Supervised Learning

We are quite good

- Unsupervised Learning
- Reinforcement Learning

Yet to explore more

Specific AI

vs

General AI

ok..

Yet to explore more

Today's agenda

- A brief introduction to Deep Learning
- Survey
- CSCI 599 overview

Degree Level

1. Undergraduate
2. Master's
3. PhD

Department

1. Computer Science
2. Electrical Engineering
3. Science / Engineering
4. Medical
5. Business
6. Others

Background in Programming

1. Python Guru
2. Python Proficient
3. Python Intermediate
4. Python Beginner
5. Proficient in other languages
6. Others

Background in Deep Learning

1. Know all state-of-the-art topics (e.g. memory networks, relation networks, and A3C)
2. Worked on 1-2 projects
3. Heard about it
4. None of the above

Background in Machine Learning

1. Research experience in ML
2. Took some relevant courses
3. None of the above

Research experience?

1. Multiple projects in AI
2. Worked on at least 1 project in AI
3. Research experience in CS/EE
4. Research experience in Engineering
5. Research experience in others
6. No experience

When to eat dinner?

1. Before
2. After
3. During

Today's agenda

- A brief introduction to Deep Learning
- Survey
- CSCI 599 overview

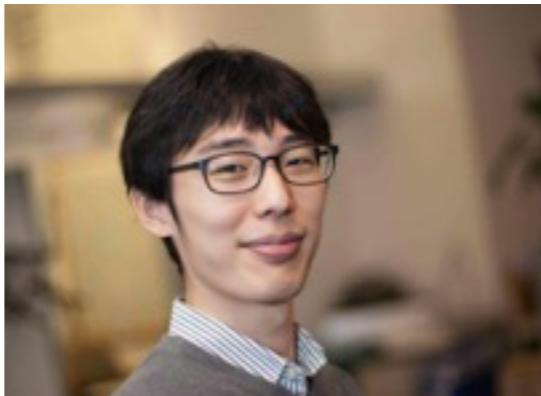
Our Goal

- Creating and tackling challenging yet interesting
deep learning projects
- Teaching **practical deep learning** knowledge

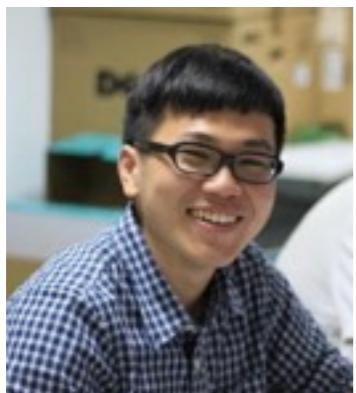
Topics

- Basic ML
- Convolutional Neural Networks (CNNs)
- Recurrent Neural Networks (RNNs)
- Generative Models
- Deep Reinforcement Learning
- Advanced Deep Learning (e.g. memory networks)

Course Staff



Prof. Joseph Lim



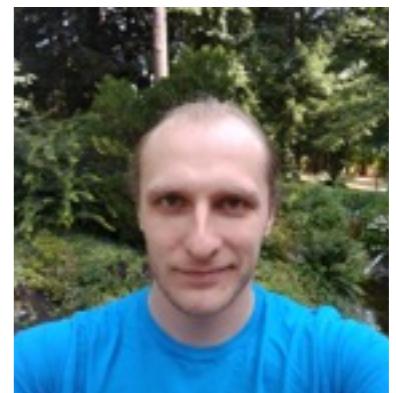
Hexiang Hu



Zeng Huang



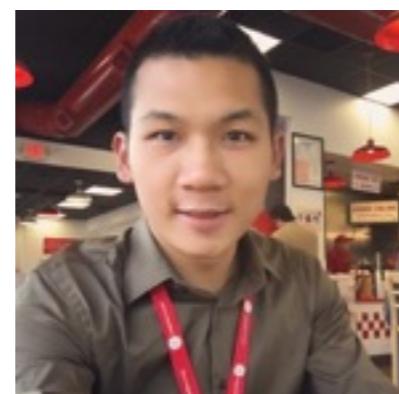
Youngwoon Lee



Artem Molchanov



Shunsuke Saito



Shao-Hua Sun



Te-Lin Wu



Sitao Xiang

Office Hours

- Instructor OH @ SAL 214
 - Wednesday 2-3pm
 - This is NOT for homework related questions.
- TA OH @ SAL 125
 - Tuesday 1-5pm

Pre-requisite

- Proficiency in Python
- College calculus, linear algebra
- Probability and statistics
- Equivalent knowledge of CSCI 567 (Machine Learning)
 - Online course may or may not be sufficient
- (Ideally) experience with cloud services

Entrance Exam

- You need to pass both
 - Take-home coding assignment (due next Wednesday 5pm)
 - In-class exam (next Wednesday)
Last name starts with
 - A-L: 5-6:30pm
 - M-Z: 6:45-8:15pm

Entrance Exam: In-class exam

- Background Knowledge
 - Calculus
 - Linear algebra
 - Probability and statistics
 - Machine Learning
- Open-ended questions (1 paragraph each):
 - Why are you taking this course?
 - Which project(s) are you excited about?

Entrance Exam: Take-home

- Individual assignment - **Do NOT collaborate**
- Python programming skills
 - Data manipulation
 - Basic ML algorithms
- Due by 5pm on Aug 30th
 - Submit your code online

Grading

Entry	% of grade
Assignment #1	15
Assignment #2	15
Midterm	25
Course project	45
TOTAL	100

Important Dates

- Assignment 1: week 6
- Assignment 2: week 11
- Project meeting with Instructor #1: week 8 (M-W)
- Project meeting with Instructor #2: week 11 (M-W)
- Project meeting with TA: 3 times (arranged later)
- Final presentation:
 - week 15 (5-9:30pm) **4.5 hours**

Subject to change!

Course Project

- Team-based project (3-4 students per team)
- Each team will have at least 1 dedicated TA
 - Mandatory meeting with TA at least once every 3 weeks
- Create your own problems (extra points)
 - **Talk and discuss** with your TAs and me!
 - In the worst case, we will give a project idea
 - Less fun, Less points!

Course Project

- Computational resource (**be conservative!**):
\$150 Google Cloud credit per student
\$125 Amazon AWS credit per student
- Tentative Schedule for Project
 - Week 4: Course Project Team
 - Week 8: Course Project Proposal
 - Week 12: Mid-report
 - Week 15: **Project Presentation** (5-9pm) + Report

Subject to change!

Collaboration Policy

Entry	Ok to discuss?
Entrance Take-home Exam	NO
Assignment #1	Yes
Assignment #2	Yes
Course project	Yes

Please list all names of people whom you had discussions with!

Other Policy

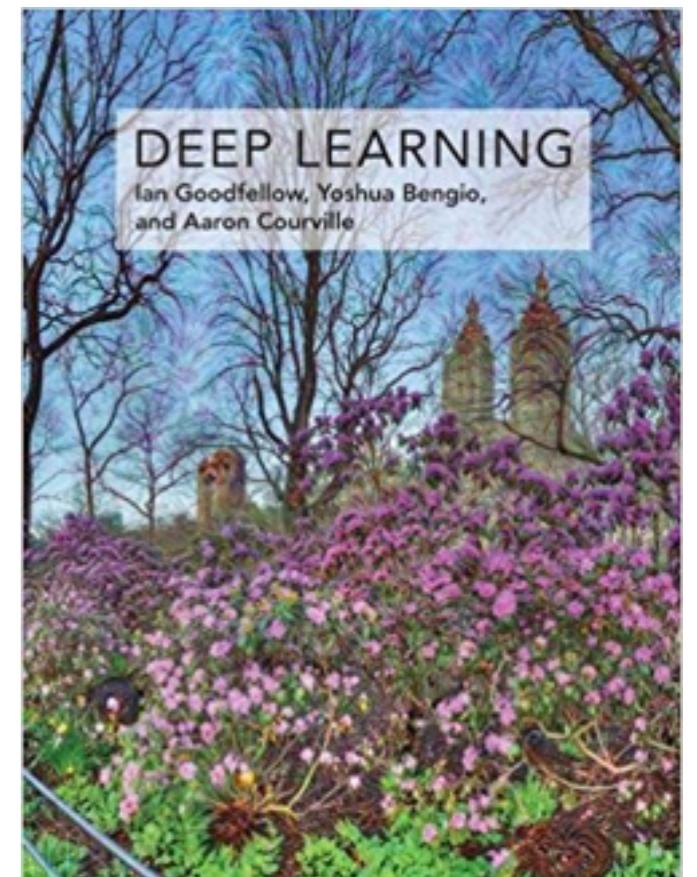
- Regrading: **within 1 week** after we hand out grades
- Late days: 5 days for the entire semester (no exception).
 - **No late day** for the final project presentation and report!

Communication

- Please use **Piazza** for any general communication including questions
<https://piazza.com/usc/fall2017/csci599/home>
- Use e-mail ONLY when it is necessary. Seriously I don't know when...
But, the staff e-mail address is: deeplearning-staff-l@usc.edu
- Any non-necessary e-mail will be ignored. Period.

Course material

- Deep Learning (MIT Press)
Ian Goodfellow, Yoshua Bengio, and
Aaron Courville
 - Free online version is available at
<http://www.deeplearningbook.org/>



Course website

- <https://csci599-dl.github.io/>

Syllabus

- Syllabus is on the course website.
- This will keep changing.

Disclaimer

- This course is taught for the 1st time @ USC. This course is 599, and thus an **experimental** course.
- The syllabus, course policy, and grading details **may change** over the semester (**check website!**)
- If you prefer a well-structured course, this is **NOT** a course for you, and I encourage you to take the course next year. We really mean this.
- It will be fun but **demanding** and challenging!

Thanks for generous support!



Google Cloud Platform



Questions?

Todo

- Check in before you leave with TAs
- Take-home assignment
- Next week: In-class exam

Welcome to CSCI 599!