Introduction to Deep Learning and Applications

ECE 176

Xiaolong Wang

xiw012@ucsd.edu

Logistics

- Website: https://xiaolonw.github.io/ece176/
- Assignments:
  - 6 Homeworks, 60% total
- Final Project:
  - Project proposal, 10%
  - Project report, 30%

Logistics

- TAs:
  - Yinbo Chen: yic026@ucsd.edu
  - Jiarui Xu: jix026@ucsd.edu
  - Yang Fu: yafu@ucsd.edu
  - Anjie Cheng: a8cheng@ucsd.edu

Logistics

Office Hour:

- Monday, 10:00 am - 11:00 am, Franklin Antonio Hall 3301.
- Friday, 4:00 pm - 5:00 pm, Franklin Antonio Hall 3301.

Logistics

The lectures are hosted in person.

We will also use zoom: https://ucsd.zoom.us/j/9086454206

Logistics

- Canvas (https://canvas.ucsd.edu/courses/51588):
  - Announcements
  - Zoom recordings
  - Slides and assignments

- Piazza:
  - https://piazza.com/class/lqvksn9zox065f
  - Discussions

- GradeScope:
  - https://www.gradescope.com/courses/690893
  - Entry Code:**PWNBJE**
  - Submit assignments

Date	Lecture	Materials	Assignments
Jan 9	Introduction		
Jan 11	Nearest Neighbor and Linear Classifiers		Assignment 1: KNN in Numpy
Jan 12	(Zoom Only) Jupyter Notebook Tutorial		
Jan 16	Linear Classifier and Optimization		
Jan 18	MLP and Back-Propagation		Assignment 2: Linear Classifiers in Numpy
Jan 19	Assignment 1 Due		
Jan 23	Intro to CNN and Back-Propagation with CNN		
Jan 25	Different Elements in Training CNNs 1		Assignment 3: Training MLP in Numpy (Toy Dataset)
Jan 26	Assignment 2 Due		
Jan 30	Different Elements in Training CNNs 2		
Feb 1	Tutorial on Pytorch		Assignment 4: Training MLP in Numpy (CIFAR10)
Feb 2	Assignment 3 Due		
Feb 6	Deep Network Architectures		
Feb 8	Image Segmentation		
Feb 9	Assignment 4 Due, Final Project Proposal Due		
Feb 13	Visualizing Deep Networks		
Feb 15	Object Detection 1: Box		Assignment 5: Pytorch CIFAR100 Classification
Feb 20	Object Detection 2: Mask and Pose		
Feb 22	Recurrent Neural Networks		
Feb 27	Temporal and 3D Convolution		
Feb 29	Self-Attention and Transformer		Assignment 6: Pytorch Segmentation
Mar 1	Assignment 5 Due		
Mar 5	Vision Transformer		
Mar 07	Generative Adversarial Networks		
Mar 12	Conditional Generative Adversarial Networks		
Mar 14	Self-supervised Learning		
Mar 15	Assignment 6 Due, Final Project Due		

Final Project

https://docs.google.com/document/d/1hB64kWtOLWlbX9yw1Ng_xJ2HH21UZLXYstPxWArVPhQ/edit?usp=sharing

Zoom / Podcast

- https://ucsd.zoom.us/j/9086454206

- Podcast

Deep Learning

- Computer Vision (Main focus in this course)
- Natural Language Processing

What is learning?

- The power of learning lies in generalization

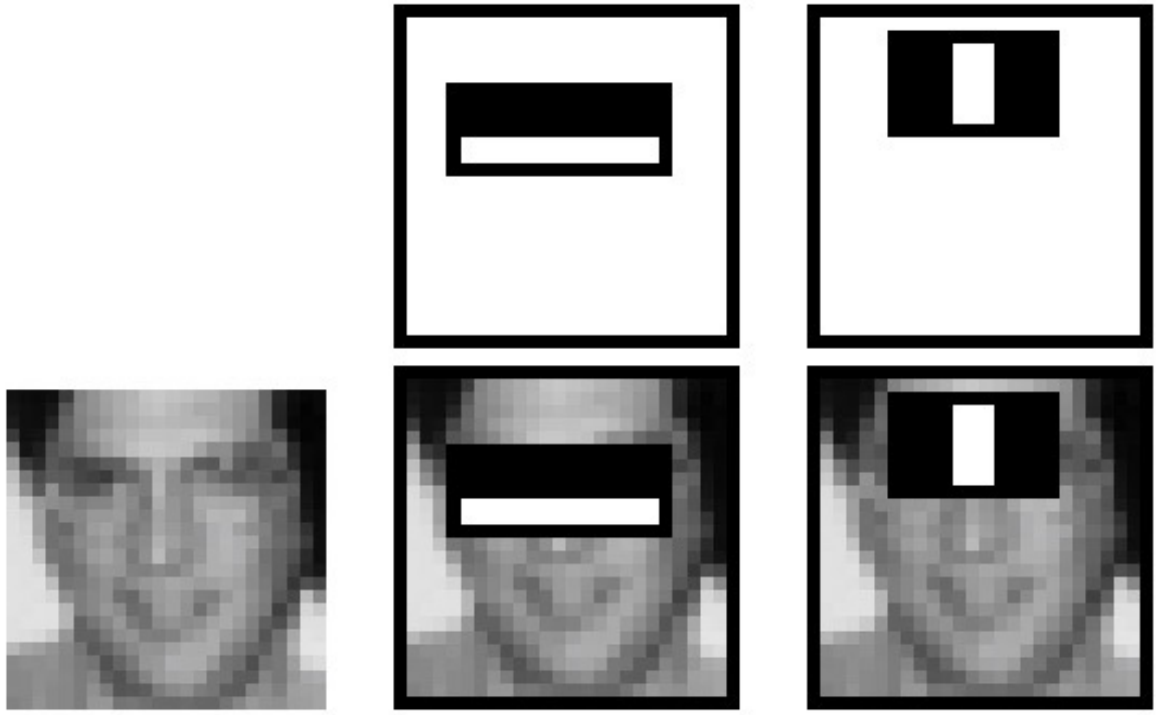


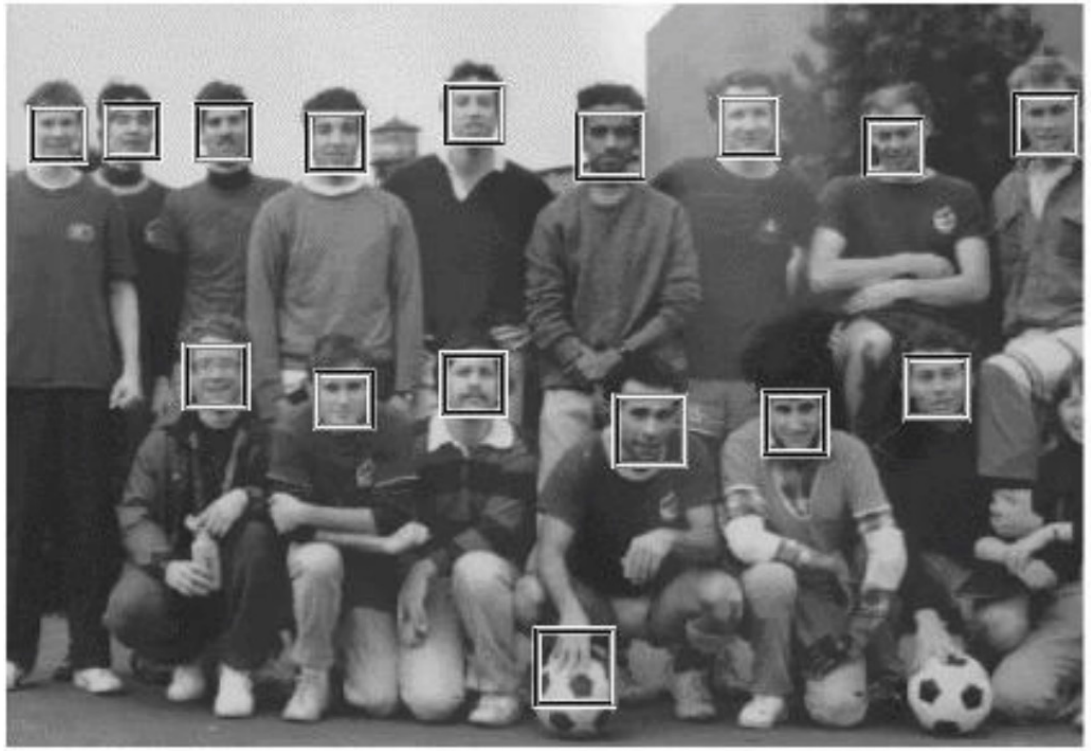

Training Data




Test Data

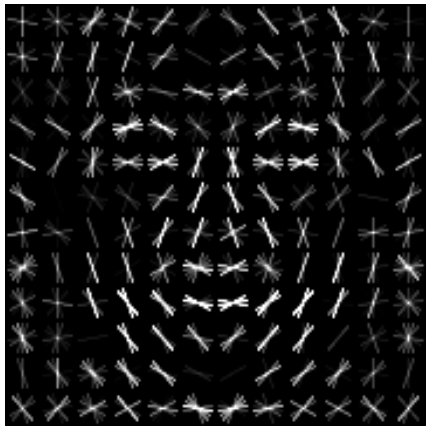
What is Deep about Deep Learning?

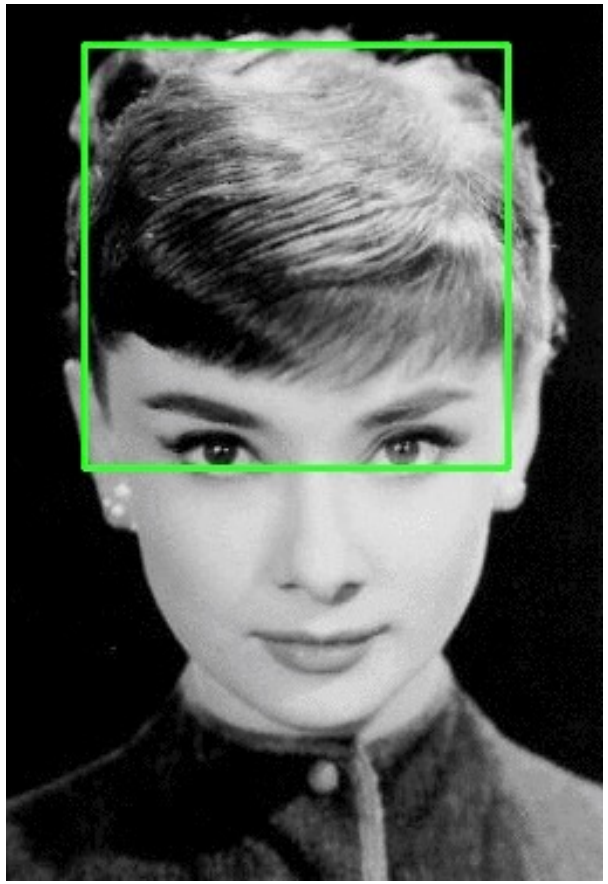


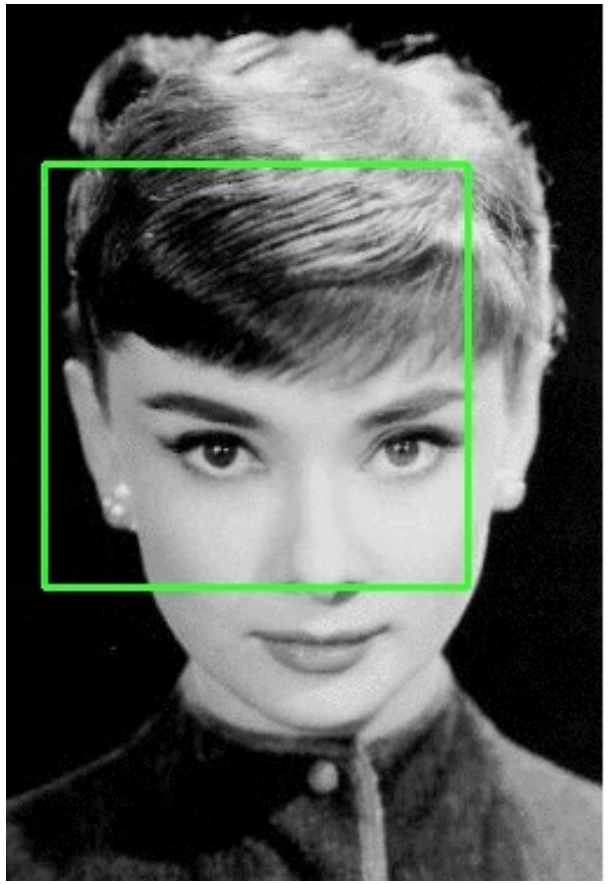


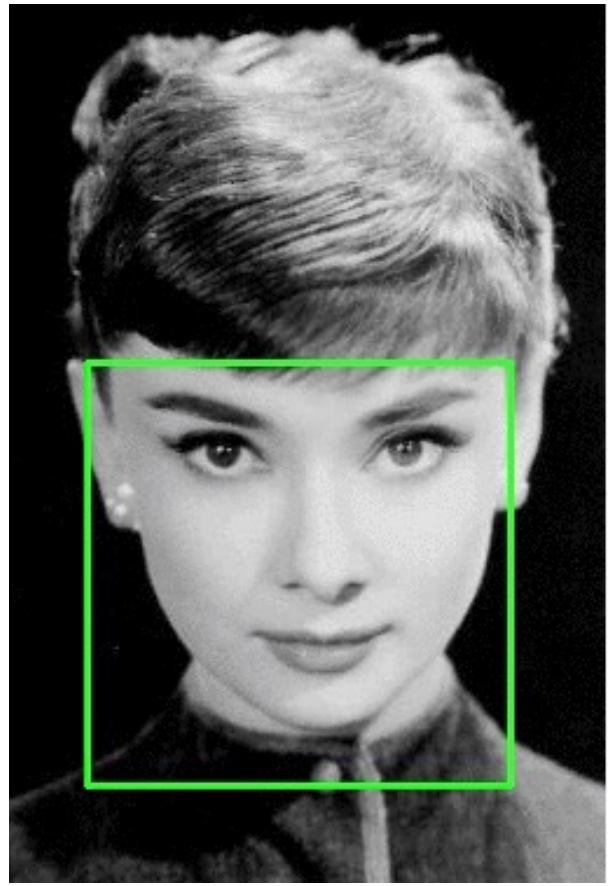
Viola et al. 2001

What is Deep about Deep Learning?







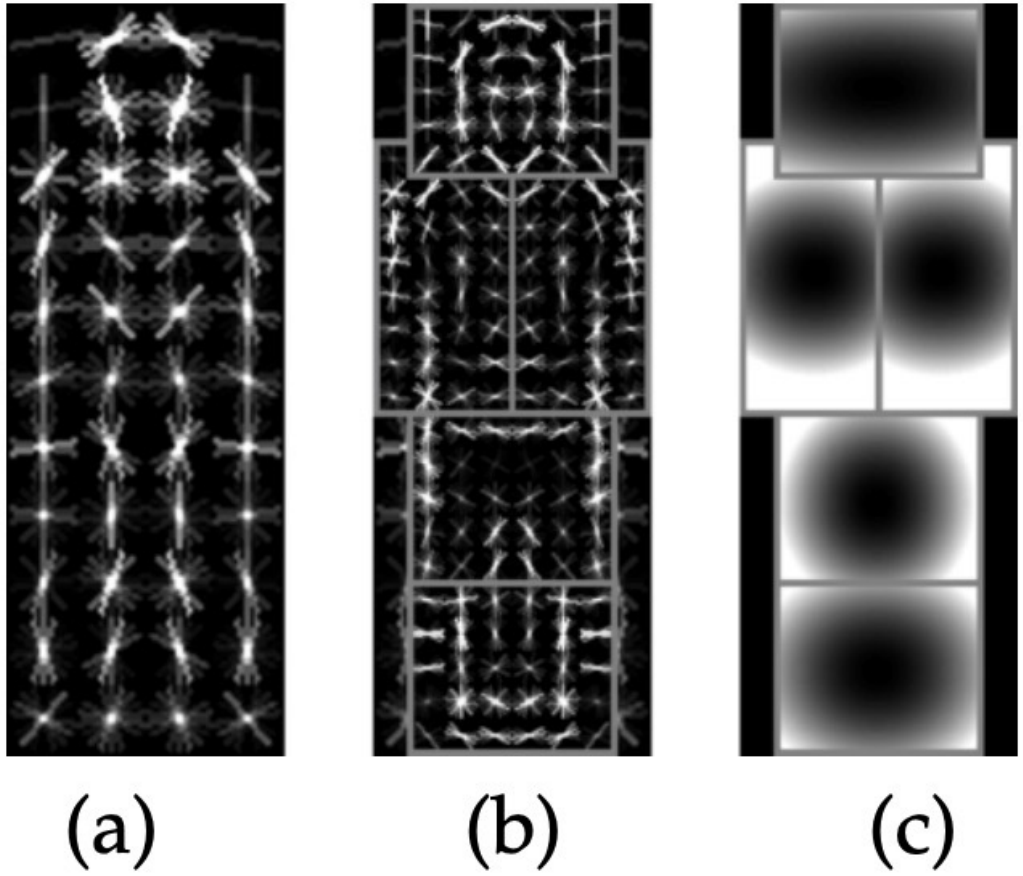


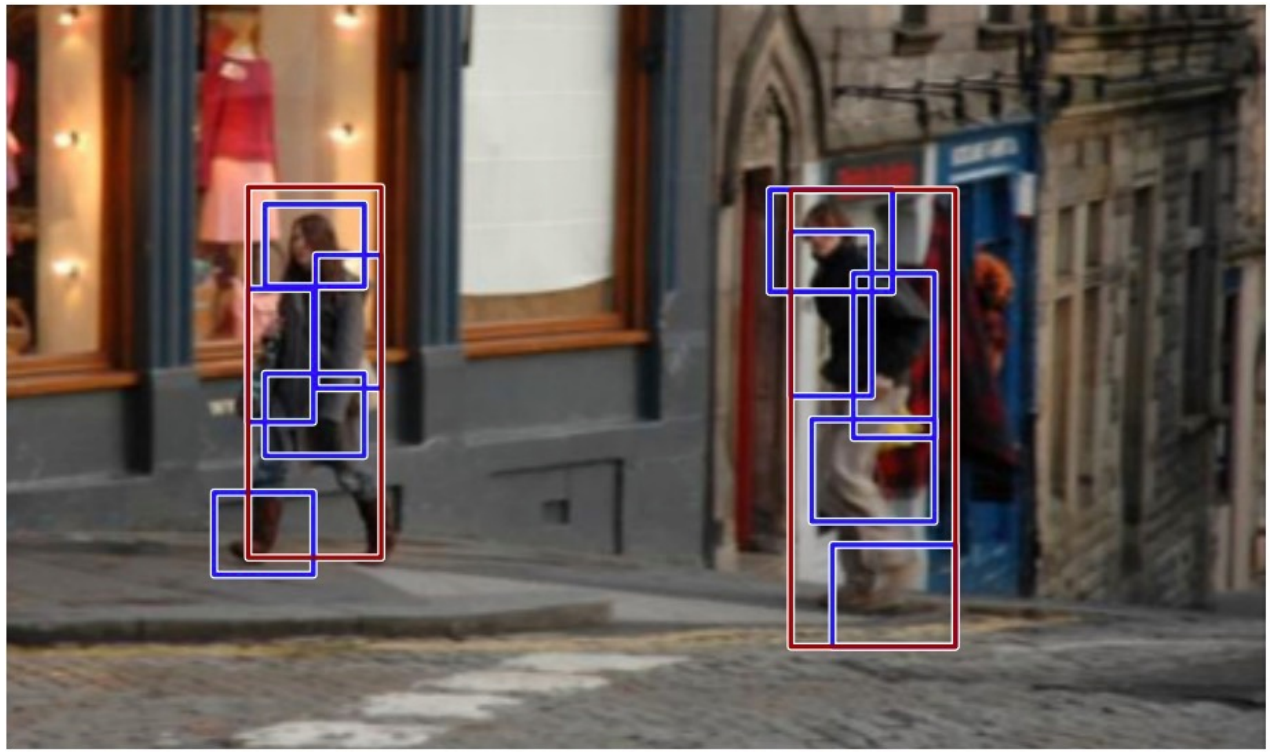
Histograms of Oriented Gradients. Dalal et al. 2005

What is Deep about Deep Learning?

Histograms of Oriented Gradients. Dalal et al. 2005

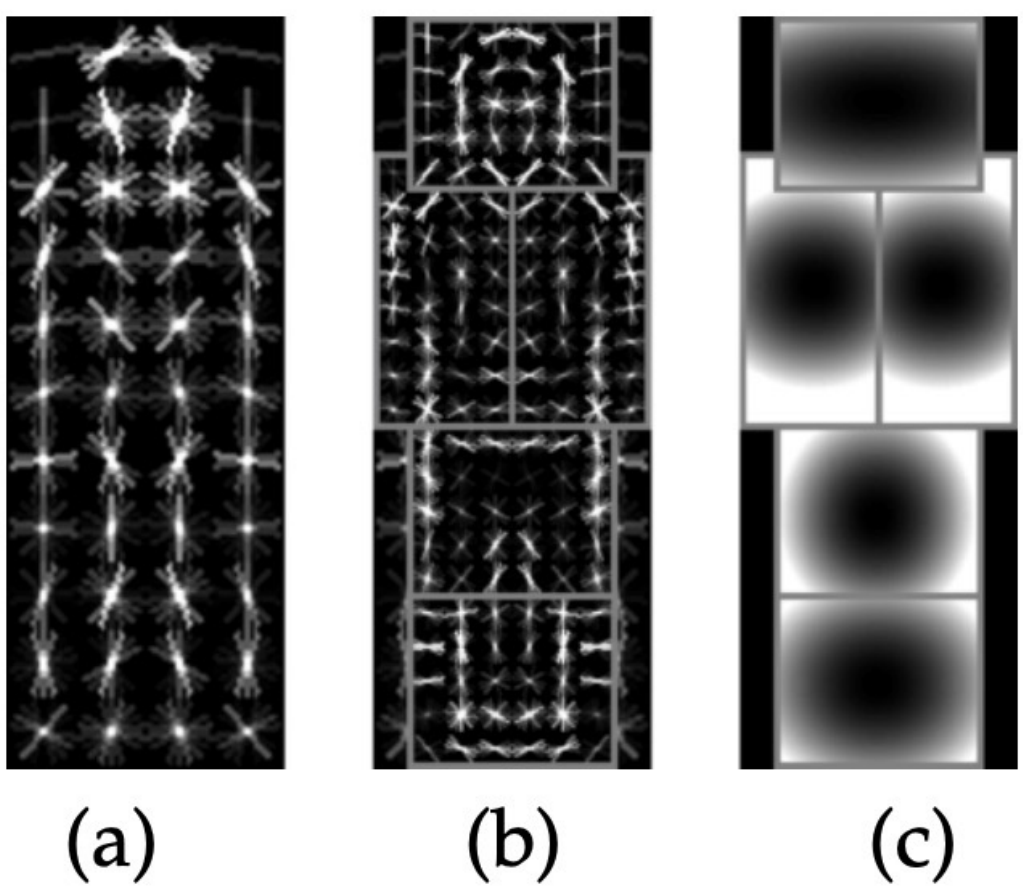
What is Deep about Deep Learning?

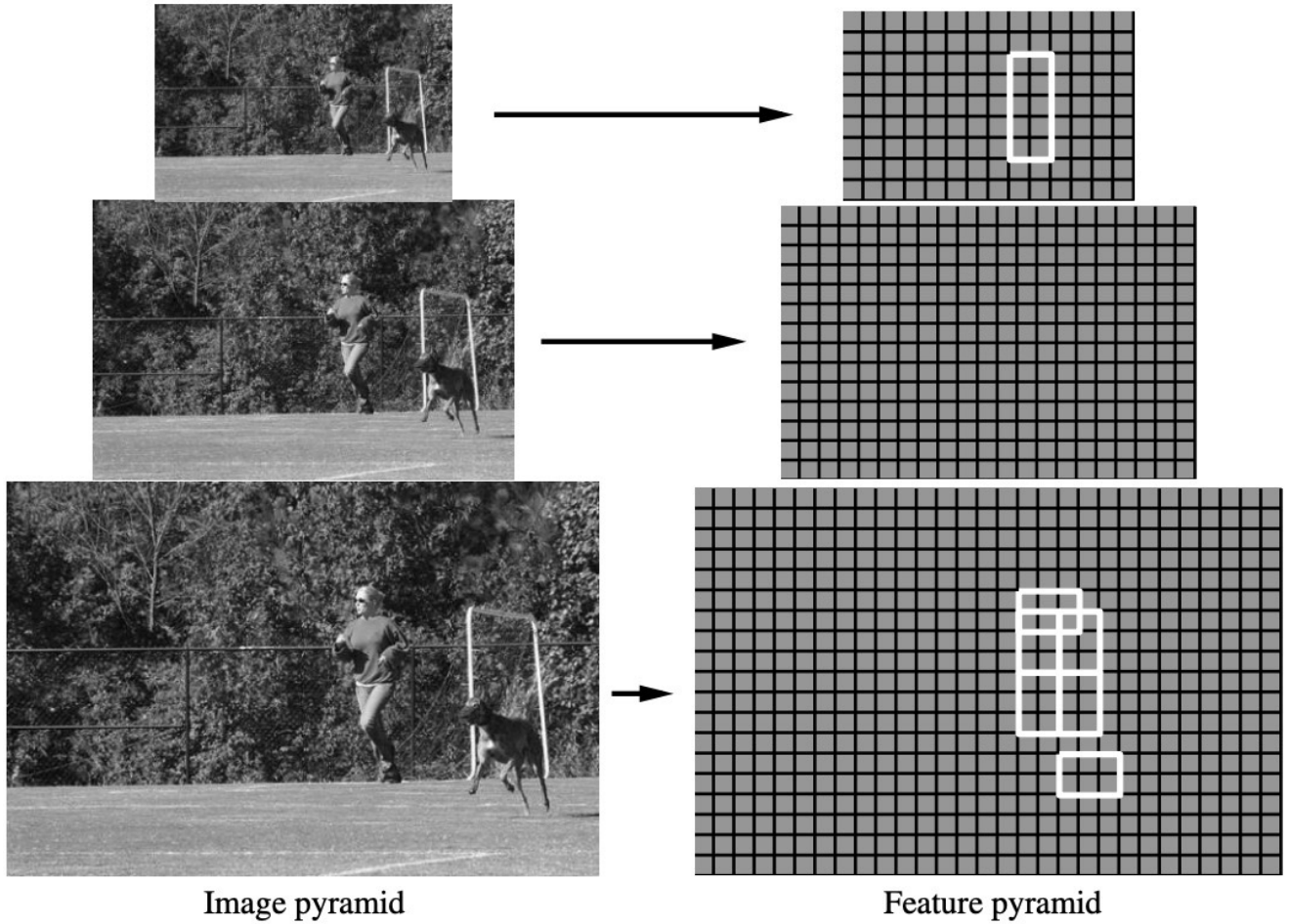




Discriminatively trained Part-based Models. Felzenszwalb et al. 2009

What is Deep about Deep Learning?





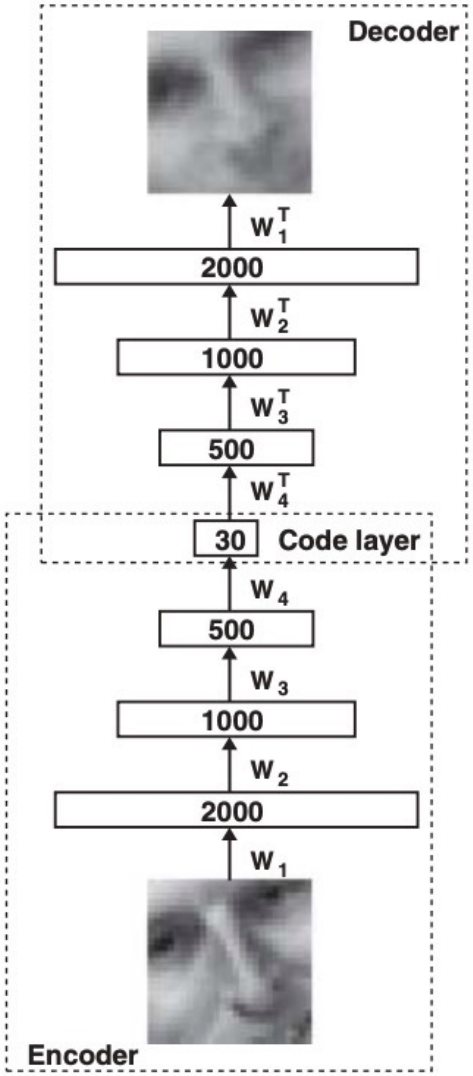
Discriminatively trained Part-based Models. Felzenszwalb et al. 2009

What is Deep about Deep Learning?

- More Layers
  - Previous method: 1-layer classifier (HoG), 2-layer classifier (DPM)
  - Deep Networks: 100, 1000 layers.

- End-to-End Training
  - Previous method: Training each layer of classifier individually.
  - Deep Networks: Training with back-propagation.

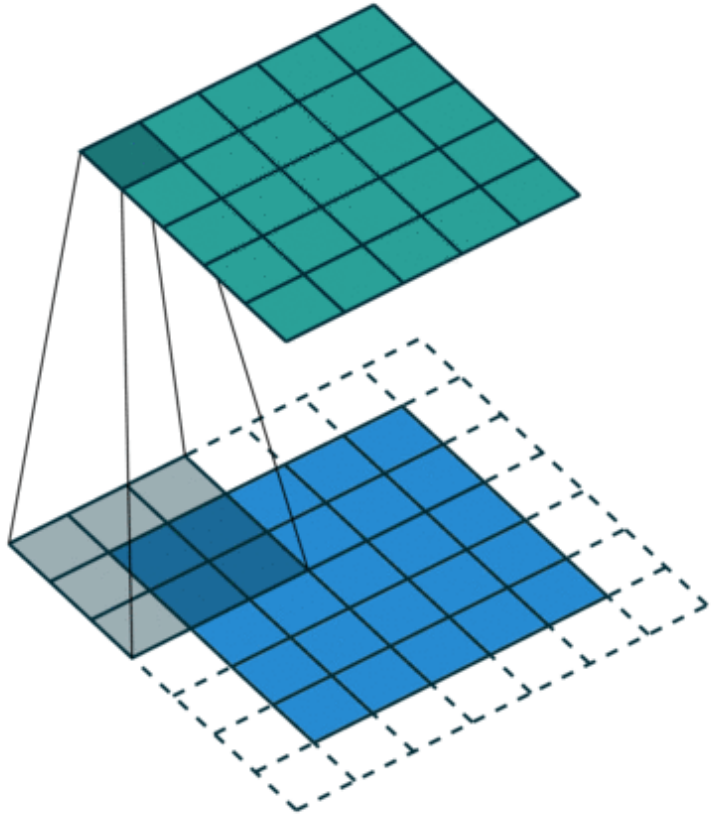
Different Types of Deep Networks

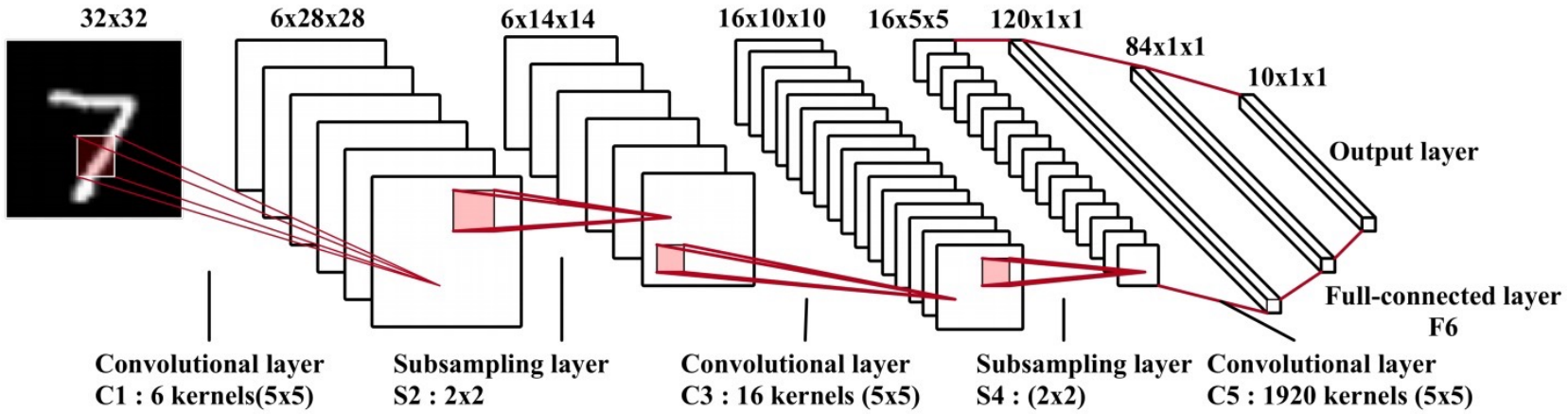


- Multilayer perceptron (MLP)
  - Input image I size : 32 x 32 = 1024
  - First hidden layer h_1 output size: 2000
  - First layer parameters W_1 size: 1024 x 2000
  - $h_1 = I\, W_1$

Hinton et al. 2006

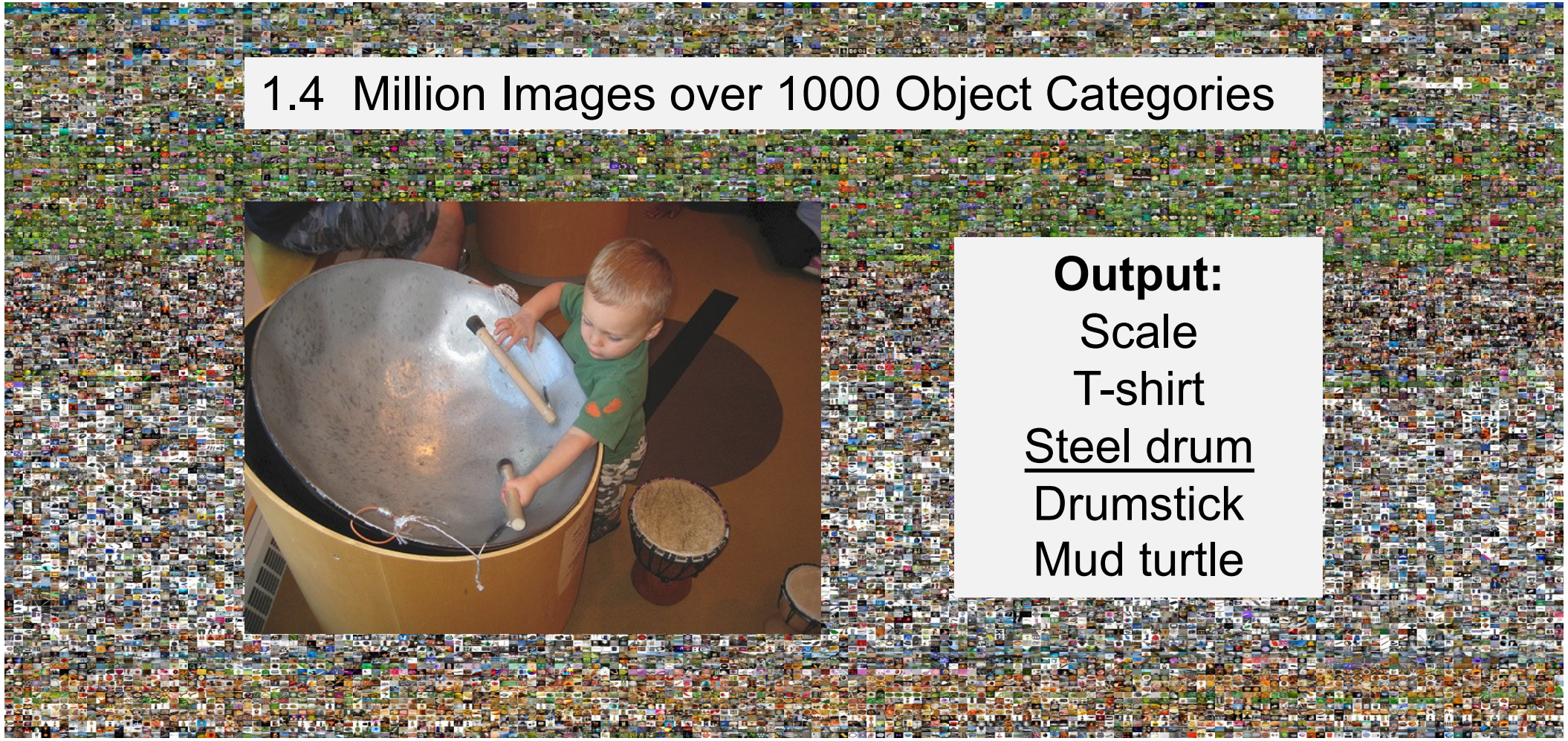
Different Types of Deep Networks





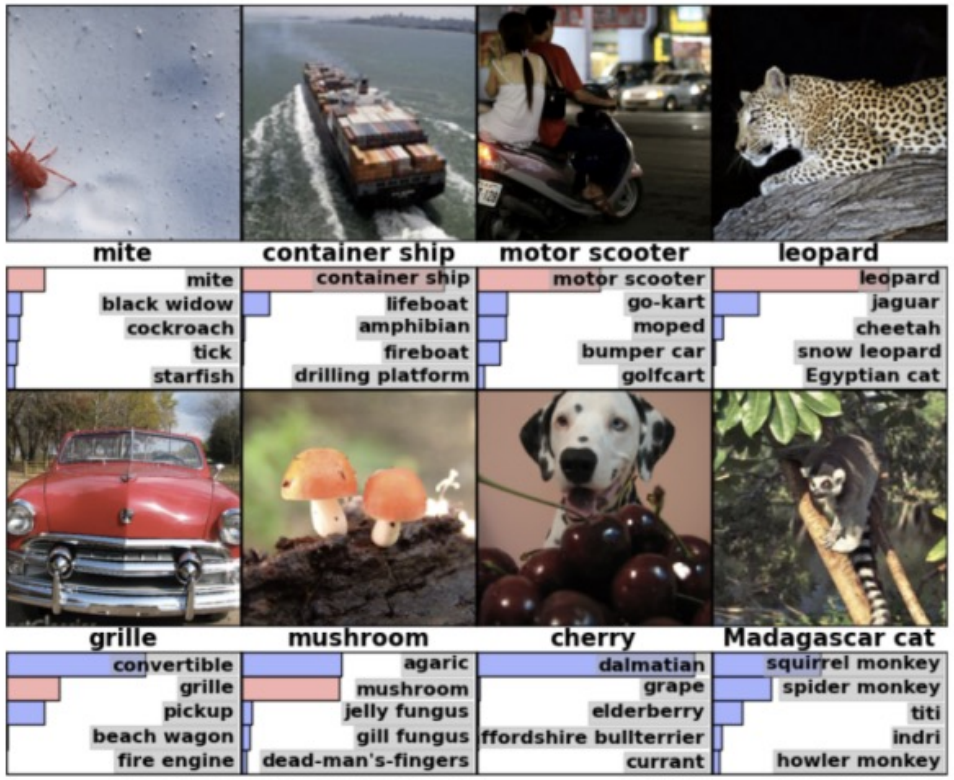
LeCun et al. 1990

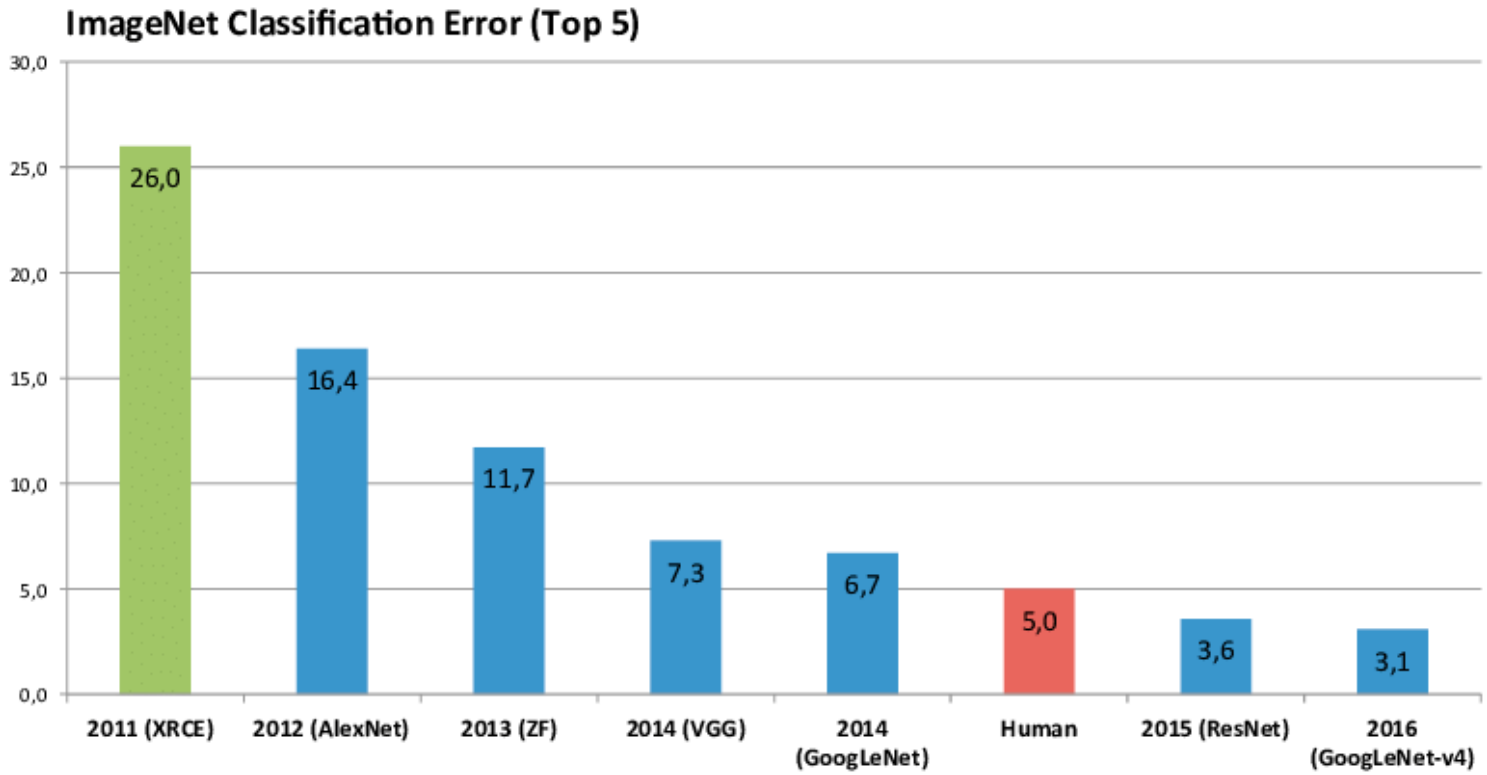
The ImageNet Challenge



Russakovsky et al. 2015

The ImageNet Challenge

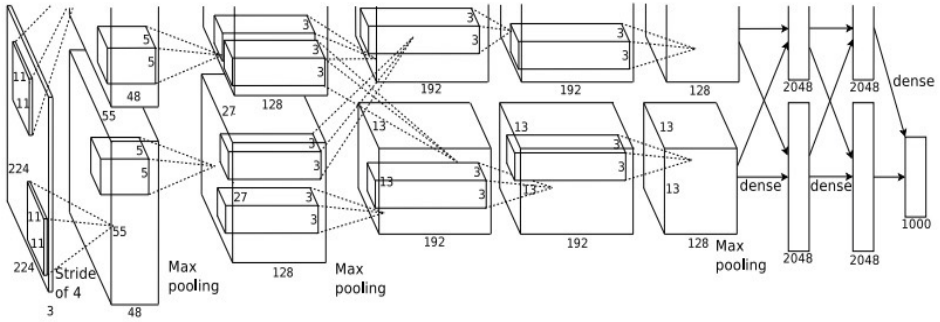


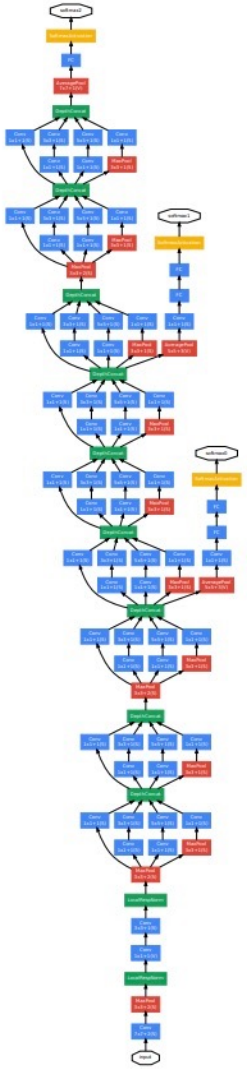


Where does the 5% error human performance come from?

http://karpathy.github.io/2014/09/02/what-i-learned-from-competing-against-a-convnet-on-imagenet/

Many-Layer Networks



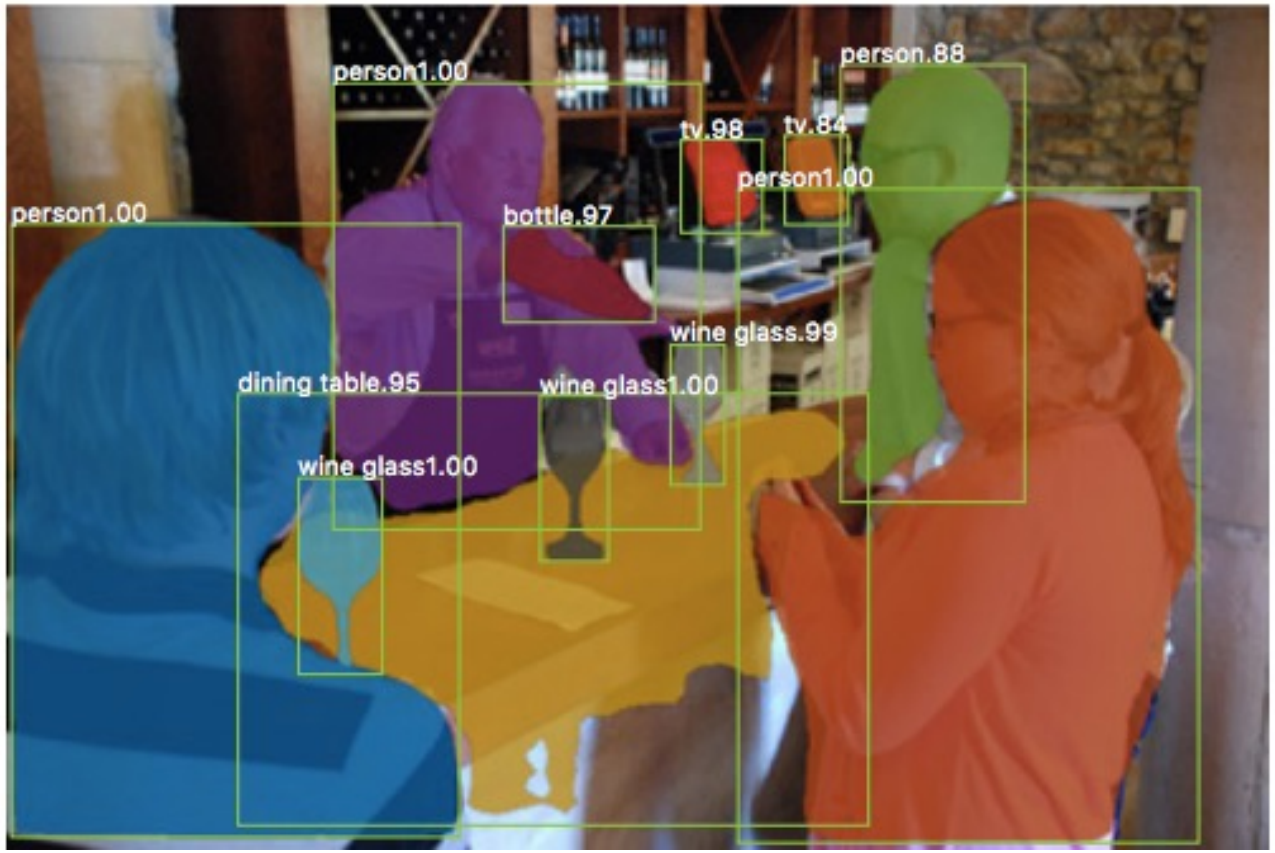


ResNets

Vision Transformers

Network	#param.	image size	throughput (image/s)	ImNet top-1
Convnets				
ResNet-18 [21]	12M	224^2	4458.4	69.8
ResNet-50 [21]	25M	224^2	1226.1	76.2
ResNet-101 [21]	45M	224^2	753.6	77.4
ResNet-152 [21]	60M	224^2	526.4	78.3
RegNetY-4GF [40]⋆	21M	224^2	1156.7	80.0
RegNetY-8GF [40]⋆	39M	224^2	591.6	81.7
RegNetY-16GF [40]⋆	84M	224^2	334.7	82.9
ViT-B/16 [15]	86M	384^2	85.9	77.9
ViT-L/16 [15]	307M	384^2	27.3	76.5
DeiT-Ti	5M	224^2	2536.5	72.2
DeiT-S	22M	224^2	940.4	79.8
DeiT-B	86M	224^2	292.3	81.8
DeiT-B↑384	86M	384^2	85.9	83.1

Object Detection/Segmentation



Mask R-CNN. He et al. 2017.

Human Pose Estimation

Mask R-CNN. He et al. 2017.

Action Recognition

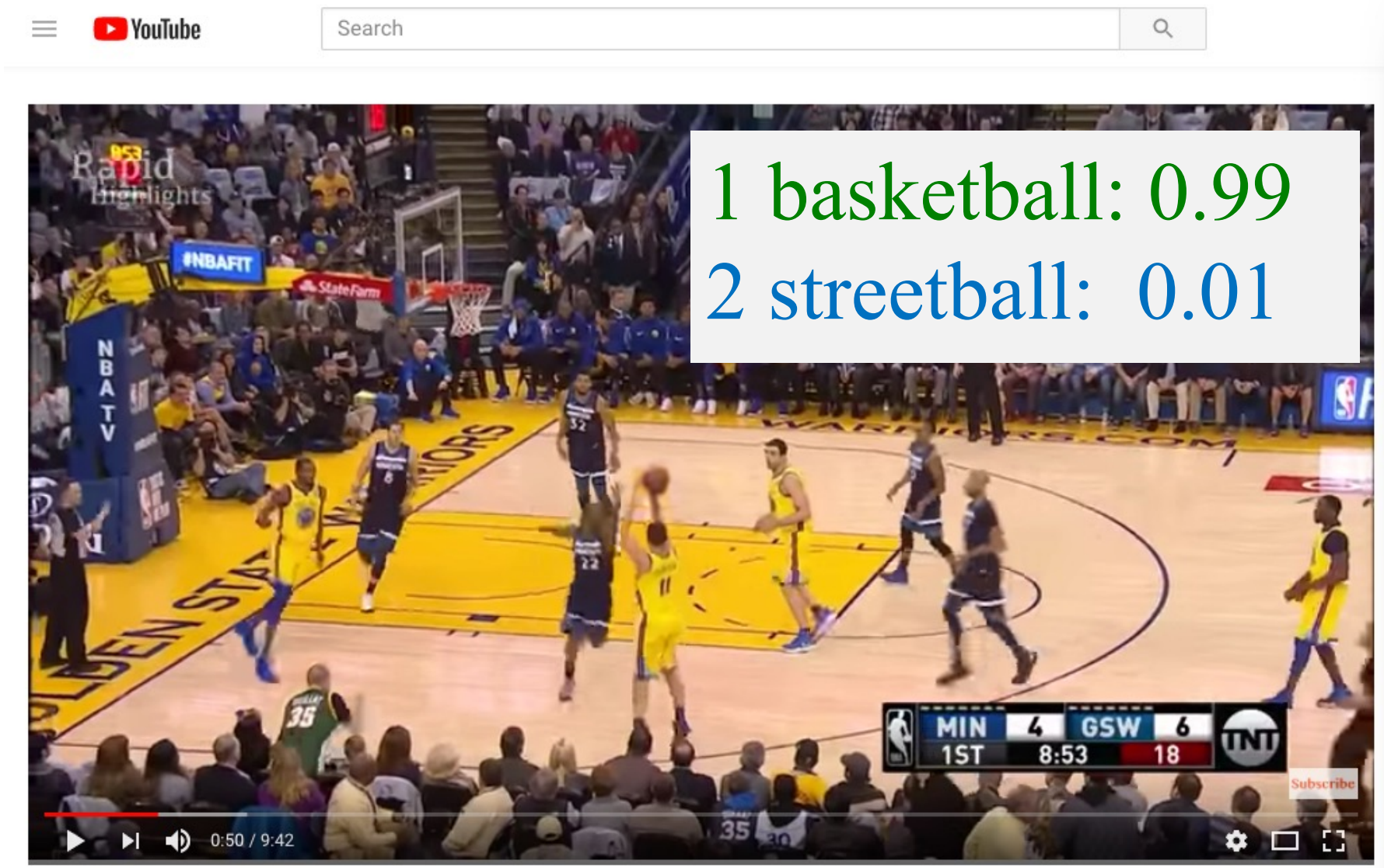
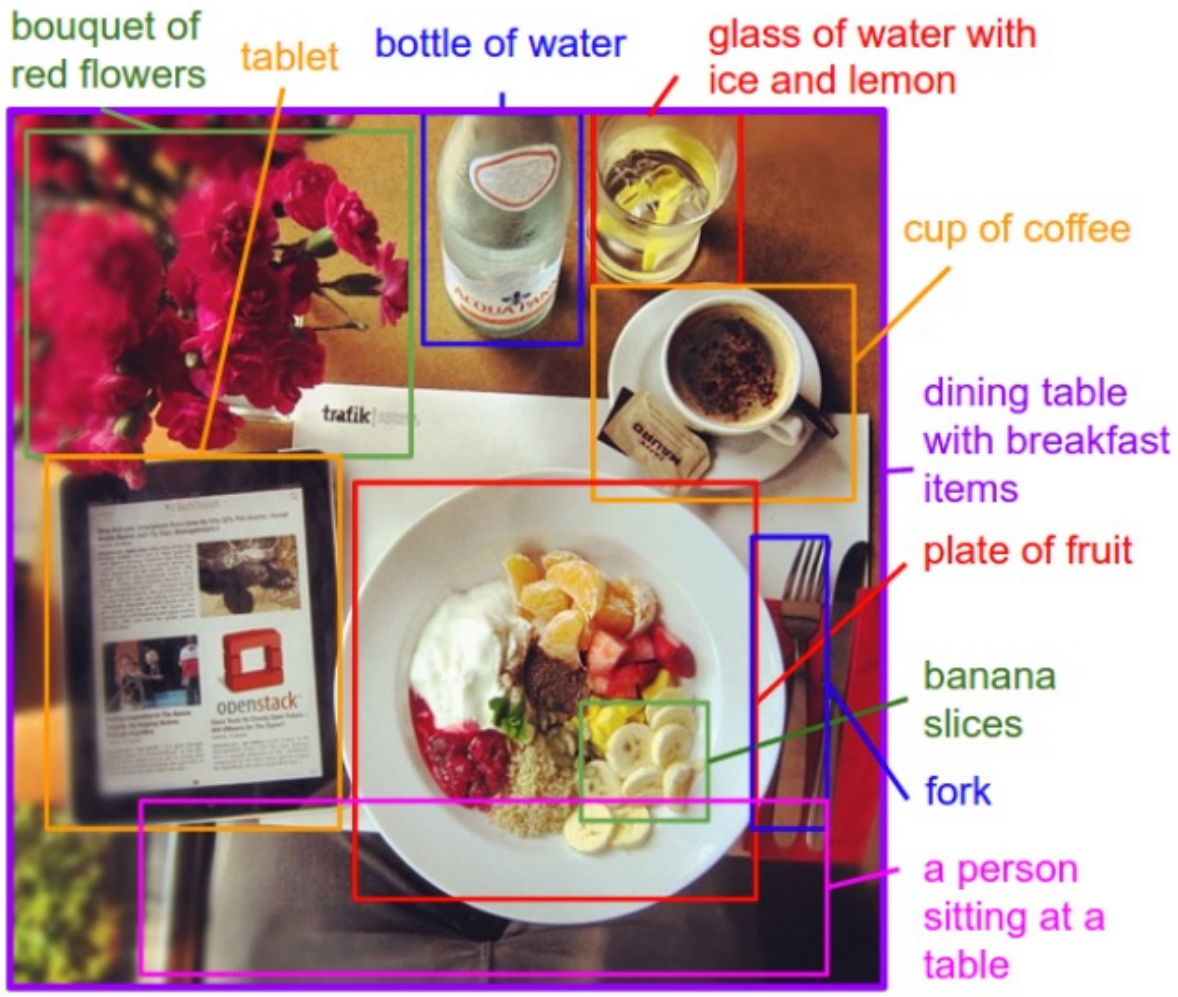


Image Captioning



Karpathy et al. 2015.

Open Vocabulary Segmentation

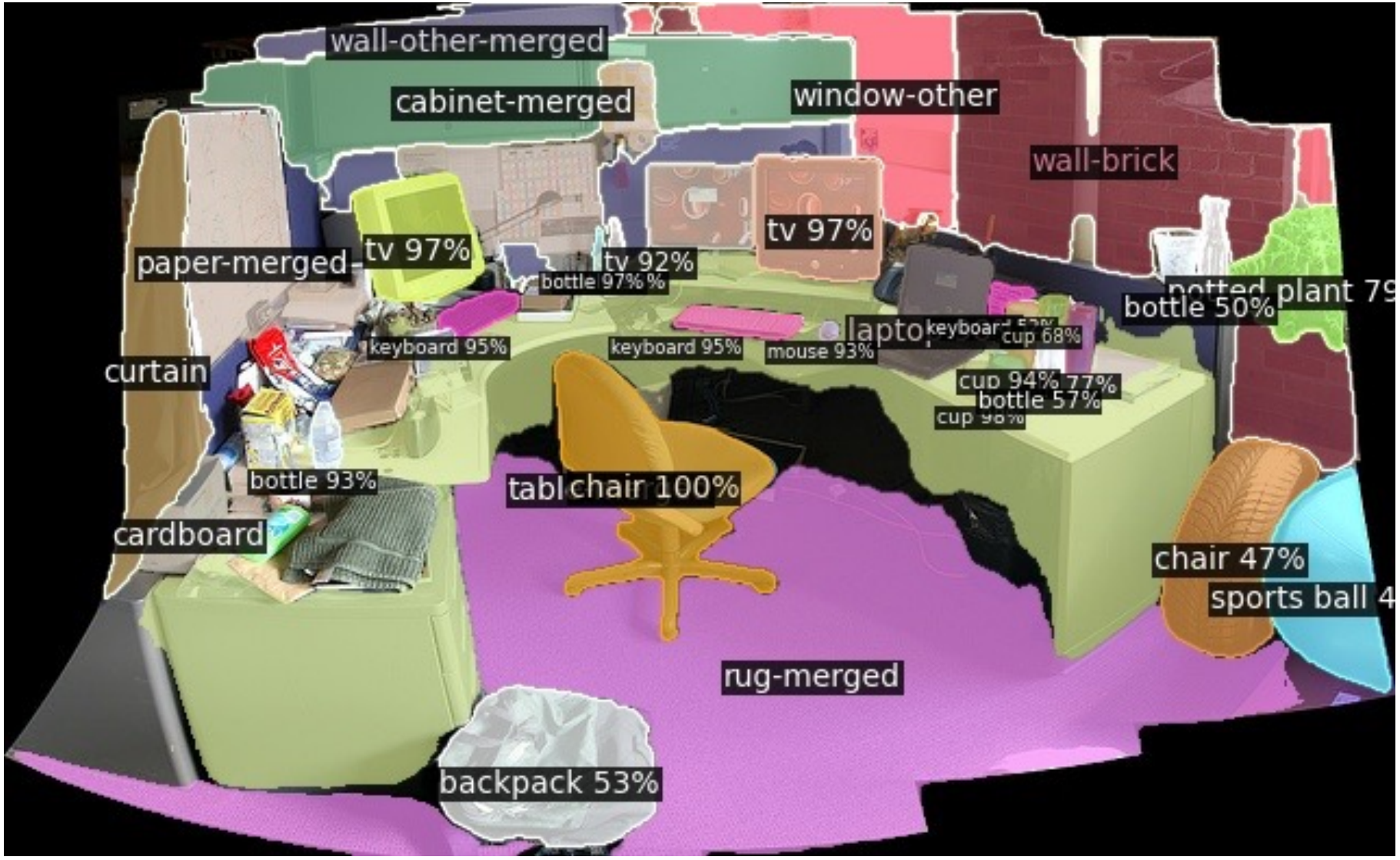
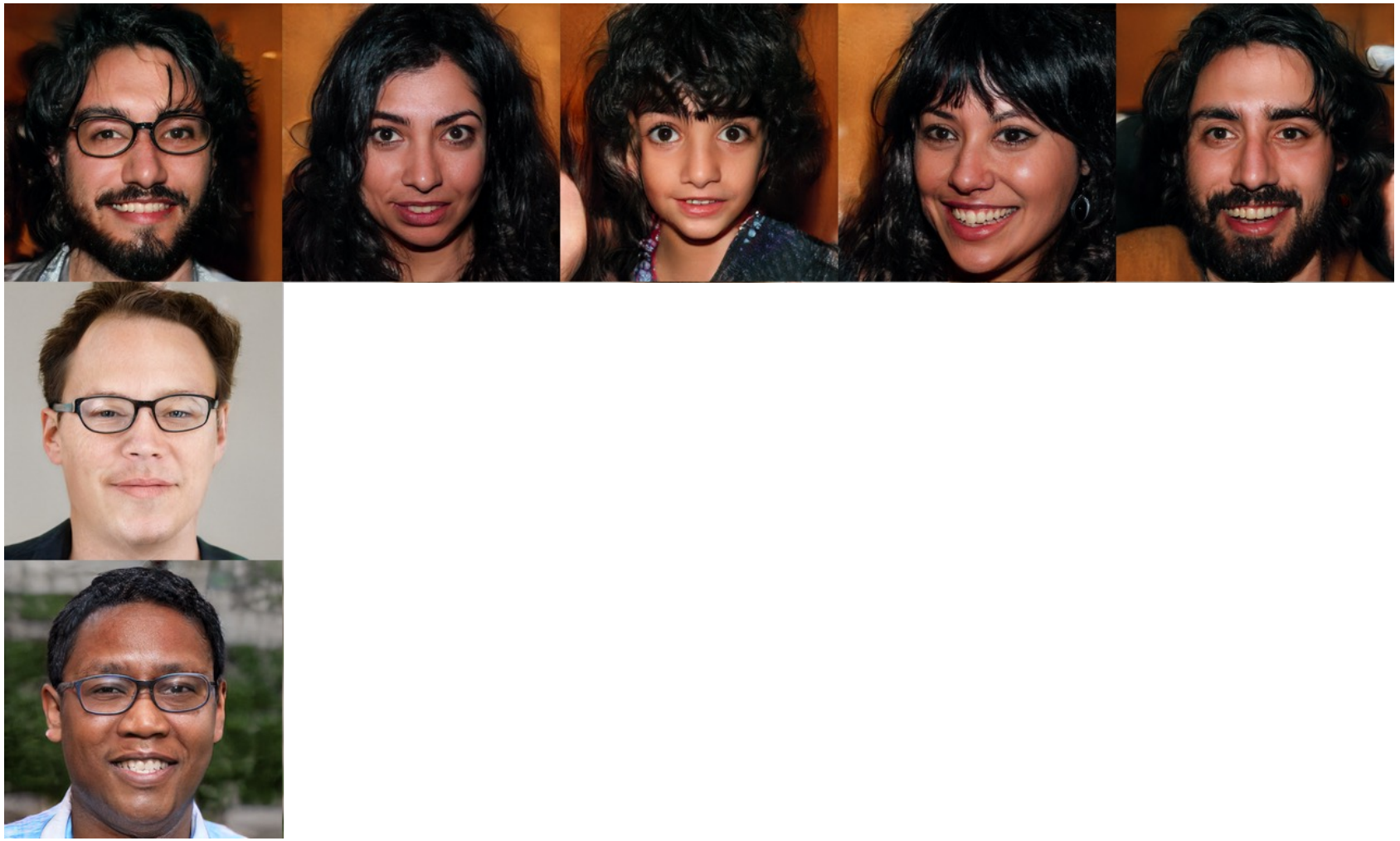


Image generation

BigGAN. Brock et al. 2019.

Image generation



StyleGAN. Karras et al. 2018.

Image generation

vibrant portrait painting of Salvador Dalí with a robotic half face

a shiba inu wearing a beret and black turtleneck

a close up of a handpalm with leaves growing from it

an espresso machine that makes coffee from human souls, artstation

panda mad scientist mixing sparkling chemicals, artstation

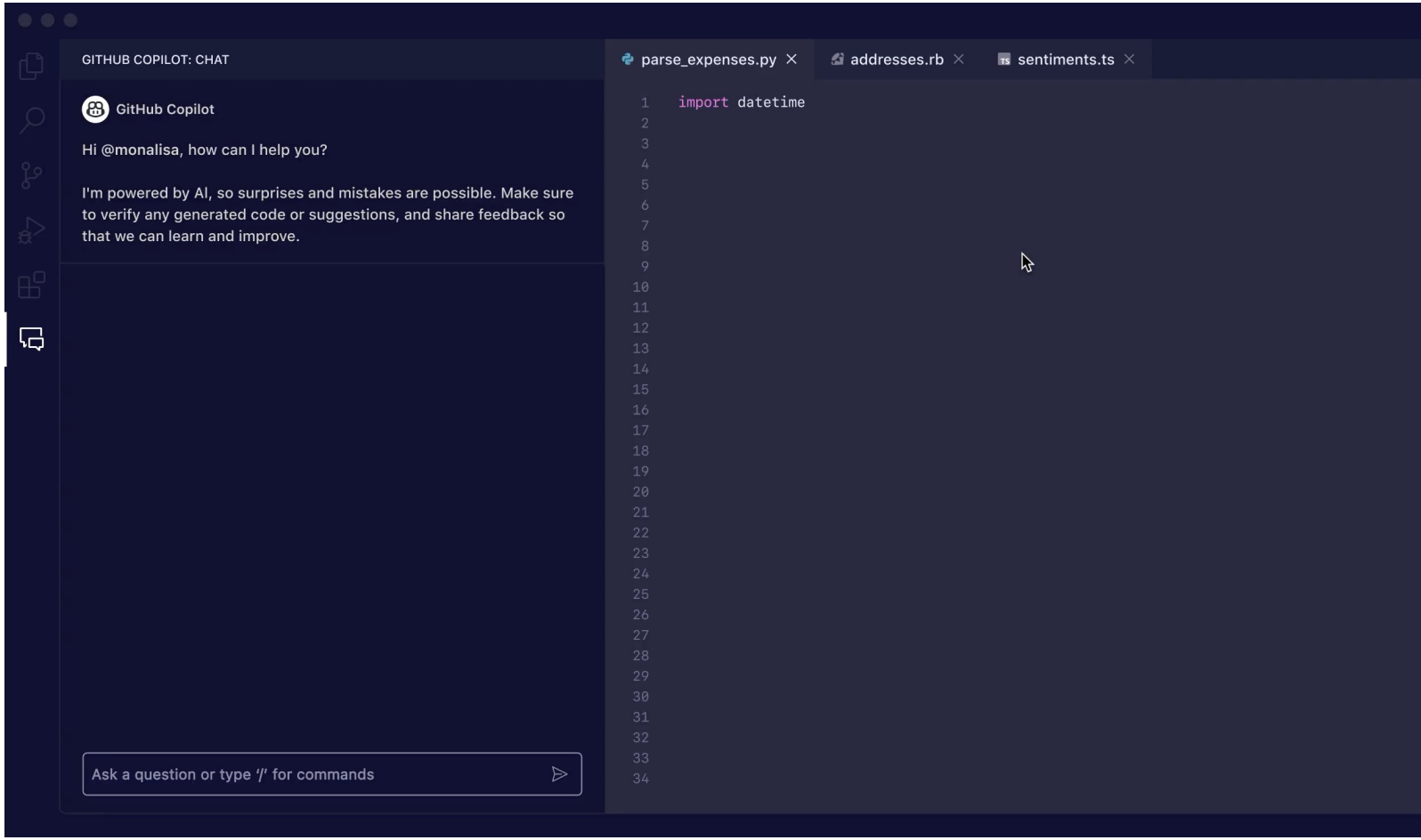
a corgi's head depicted as an explosion of a nebula

Language: Recurrent Neural Networks

```
PANDARUS:
Alas, I think he shall be come approached and the day
When little srain would be attain'd into being never fed,
And who is but a chain and subjects of his death,
I should not sleep.

Second Senator:
They are away this miseries, produced upon my soul,
Breaking and strongly should be buried, when I perish
The earth and thoughts of many states.

DUKE VINCENTIO:
Well, your wit is in the care of side and that.

Second Lord:
They would be ruled after this chamber, and
my fair nues begun out of the fact, to be conveyed,
Whose noble souls I'll have the heart of the wars.

Clown:
Come, sir, I will make did behold your worship.

VIOLA:
I'll drink it.
```

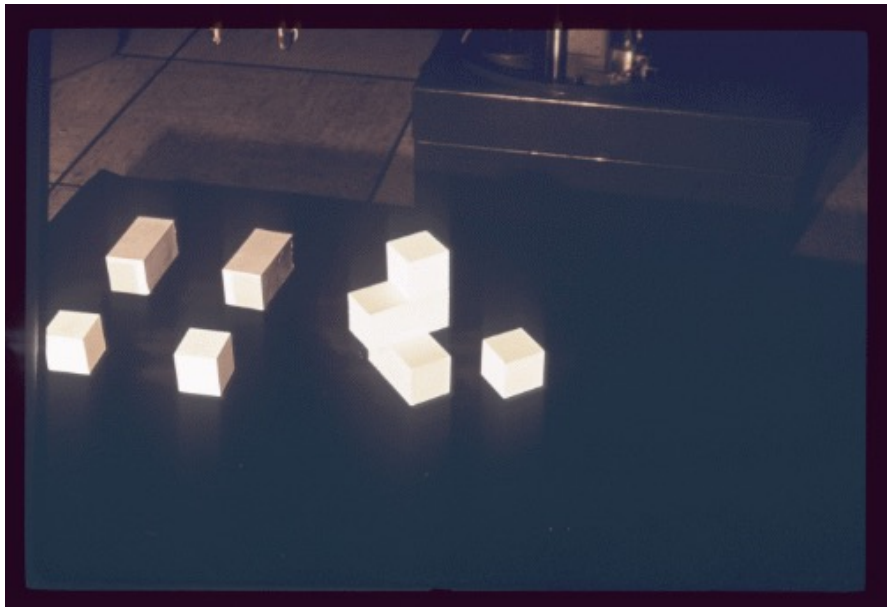
http://karpathy.github.io/2015/05/21/rnn-effectiveness/

Language: Transformer, GPT-4



https://github.com/features/copilot

Robotics



Blocks World
MIT, 1960s – 1970s
Copy demo (1970)

Gupta et al. 2018.

Robotics

https://leggedrobotics.github.io/rl-blindloco/

Lee et al. 2020.

Statistical learning, Training and Testing

- Training: Learning from the past experience:
  - training dataset
  - demonstrations

- Testing: Generalize to unseen inputs
  - Data that does not exist in training set

Image Classification

input	desired output
	apple
	pear
	tomato
	cow
	dog
	horse

Credit: Svetlana Lazebnik

Image Classification

input **desired output**

apple

pear

tomato

cow

dog

horse

training data

apple

pear

tomato

cow

dog

horse

Credit: Svetlana Lazebnik

Training time

Training Samples

Training Labels

Features

Training

Learned model

Testing time

Test Sample

Features

Learned model

Prediction

Credit: Svetlana Lazebnik

Supervised Learning

$$y = f(x)$$

output label

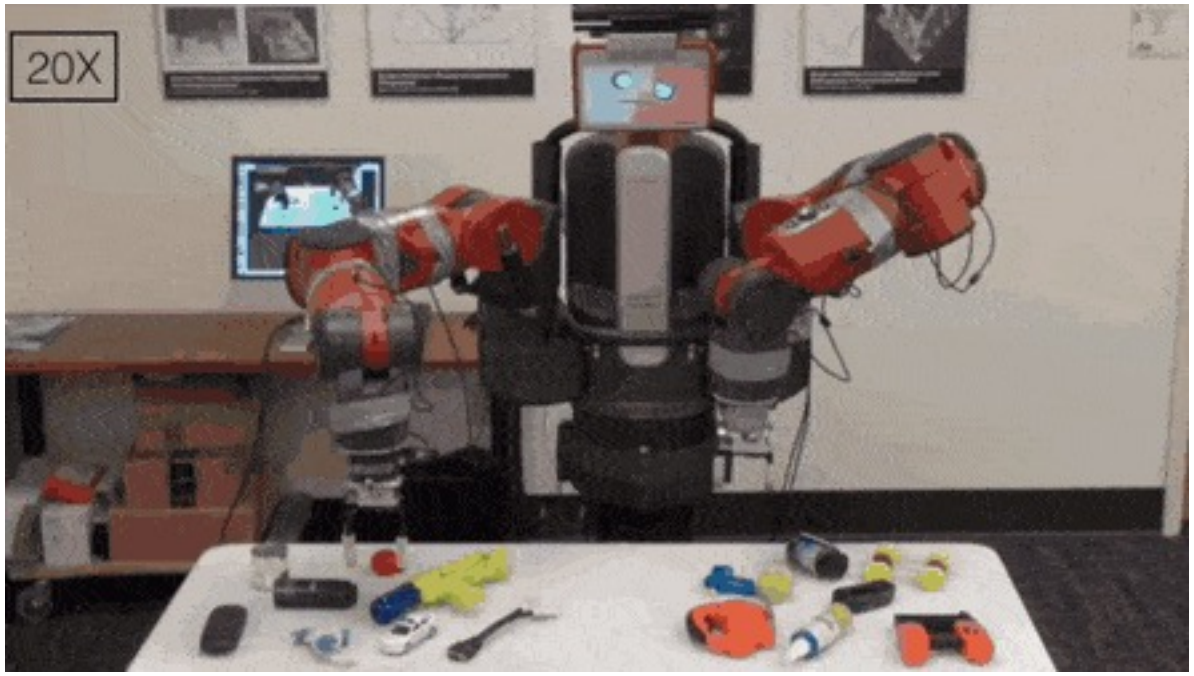
neural network

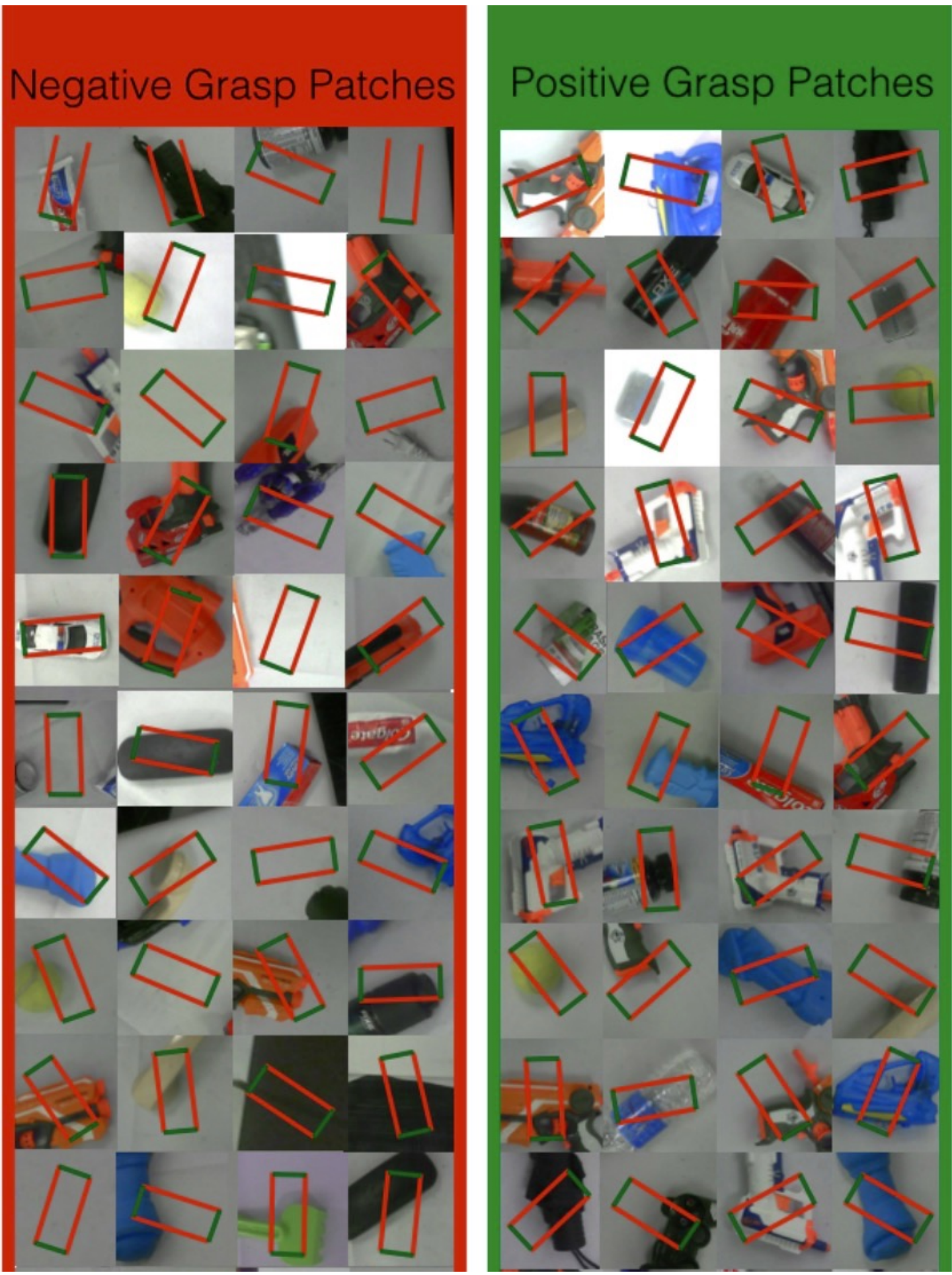
input image

- **Training** (or **learning**): given a *training set* of labeled examples $\{(x_1, y_1), \ldots, (x_N, y_N)\}$, train a neural network predictor f

- **Testing** (or **inference**): apply neural network f to a new *test example* x and output the predicted value $y = f(x)$

Supervised Learning and Self-Supervised Learning




Negative Grasp Patches
Positive Grasp Patches

Test Set

This Class

- Deep Learning backgrounds and applications
- Training and Testing

Next Class

- Nearest Neighbor Classifier
- Linear Classifier

Coming Assignments

- The first assignment will be announced in This Thursday after the class
- There will be a tutorial on how to do/submit assignments This Friday, 4:00 - 5:00 pm
- We will use the compute resources in https://datahub.ucsd.edu/