Introduction to Visual Learning

ECE 285

Xiaolong Wang
xiw012@ucsd.edu

Logistics

- Zoom: https://ucsd.zoom.us/j/9086454206
- Website: https://xiaolonw.github.io/ece285/
- Assignments:
  - 4 Homeworks, each 15%
- Final Project:
  - Project proposal, 10%
  - Project report, 30%

Logistics

- Main TAs:
  - Rishabh Jangir, rjangir@ucsd.edu ; Mohit Jain, m4jain@ucsd.edu

- Emergency TA:
  - Hanwen Jiang, h7jiang@eng.ucsd.edu ; Shaowei Liu, shl044@eng.ucsd.edu

Logistics

Office Hour, starting next week:

- Monday, 9:30 - 10:30 am
- Friday, 5:00 - 6:00 pm

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

Logistics

- Canvas (https://canvas.ucsd.edu/courses/25168):
  - Announcements
  - Zoom recordings
  - Slides and assignments

- Piazza:
  - https://piazza.com/ucsd/spring2021/ece285
  - Discussions

- GradeScope:
  - https://www.gradescope.com/courses/256233
  - Entry Code:**V84YGX**
  - Submit assignments

Date	Lecture	Materials	Assignments
Mar 30	Introduction to the Class		
Apr 1	Nearest Neighbor and linear classifiers		
Apr 6	MLP and back-propagation		Assignment 1
Apr 8	Convolutional Neural Networks 1		
Apr 13	Convolutional Neural Networks 2		
Apr 15	Tutorial on Pytorch		
Apr 20	Image Segmentation and Visualization		Assignment 2
Apr 22	Object Detection		
Apr 27	Recurrent Neural Networks		
Apr 29	Temporal and 3D convolution		Project Proposal Due
May 4	Video Prediction		Assignment 3
May 6	Self-Attention, GNN and Transformer 1		
May 11	Self-Attention, GNN and Transformer 2		
May 13	Generative Model 1		
May 18	Generative Model 2		Assignment 4
May 20	3D Vision		
May 25	Multi-task, Zero-shot, Few-shot Learning		
May 27	Trainsfer Learning and Domain Adaptation		
June 1	Self-Supervised Learning 1		
June 3	Self-Supervised Learning 2		Final project due

Final Project

https://docs.google.com/document/d/1aByplfb_VHFHTaFdZe2TBZXJQx1AR5Zhj3s4S7IGJ_A/edit?usp=sharing

Cannot re-use existing project that is online.

Class Interaction

- Use Chat to ask questions if you want to type the question
- You can directly ask the question by speaking after raising your hand

Computer Vision with Deep Learning

What is learning?

- The power of learning lies in generalization

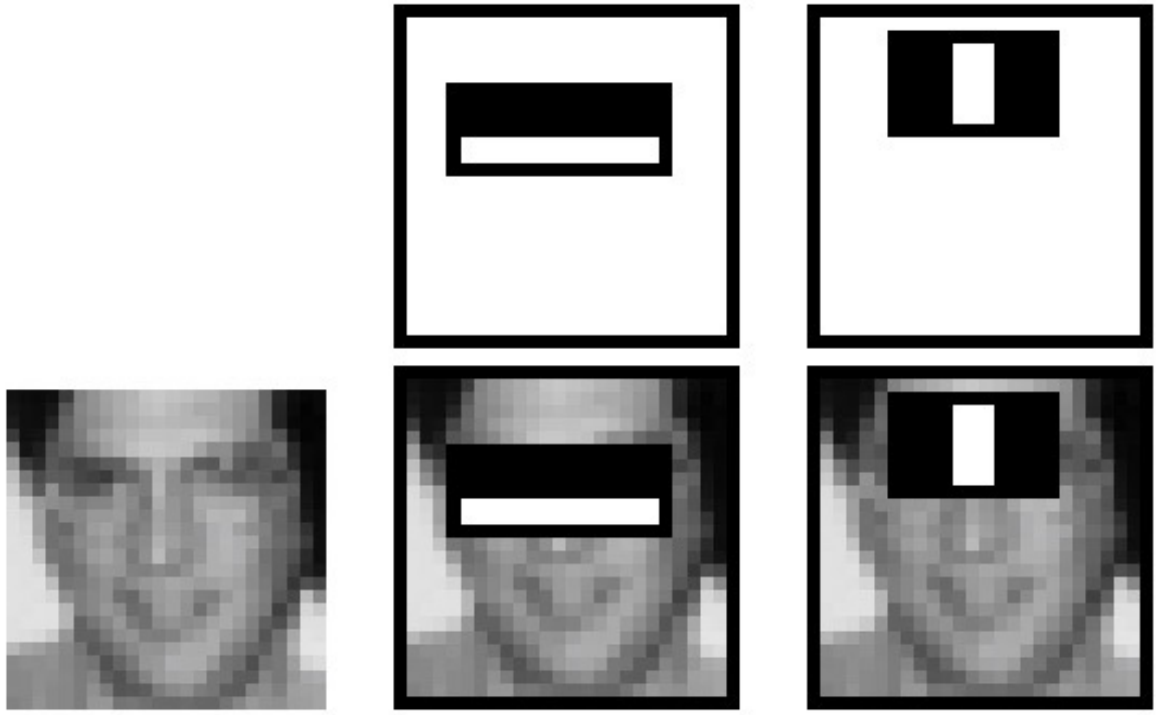


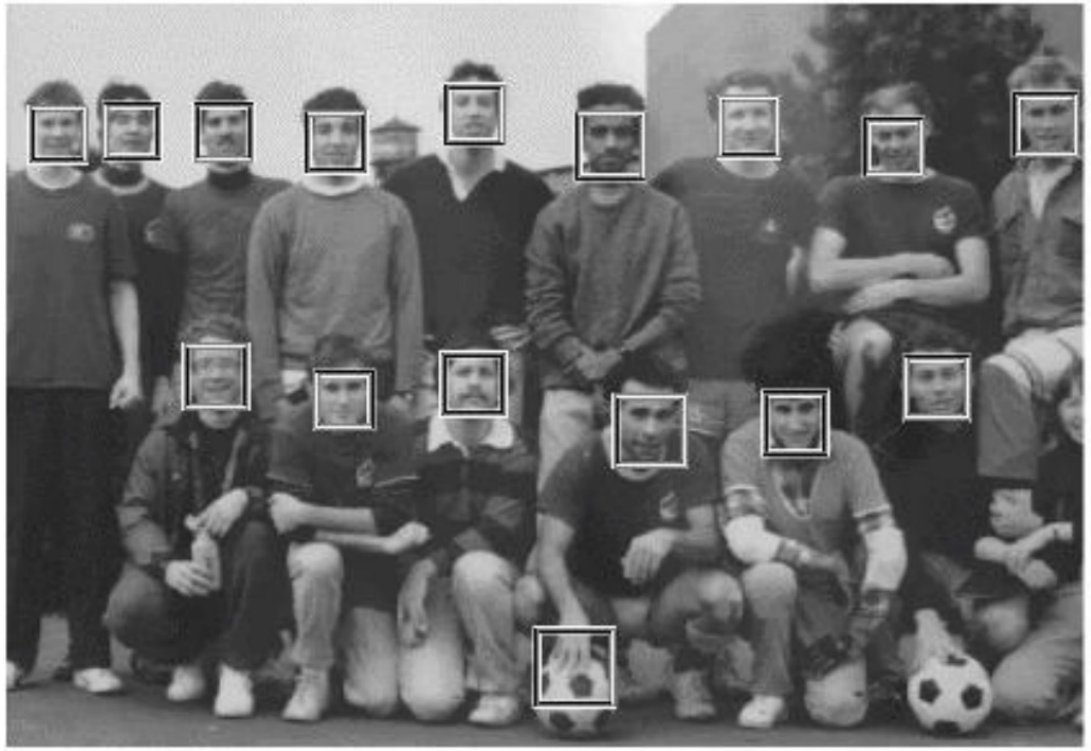

Training Data




Test Data

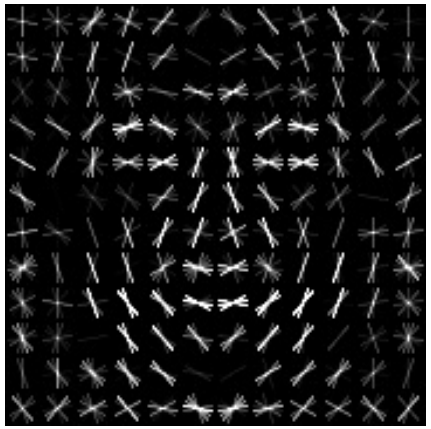
What is Deep about Deep Learning?

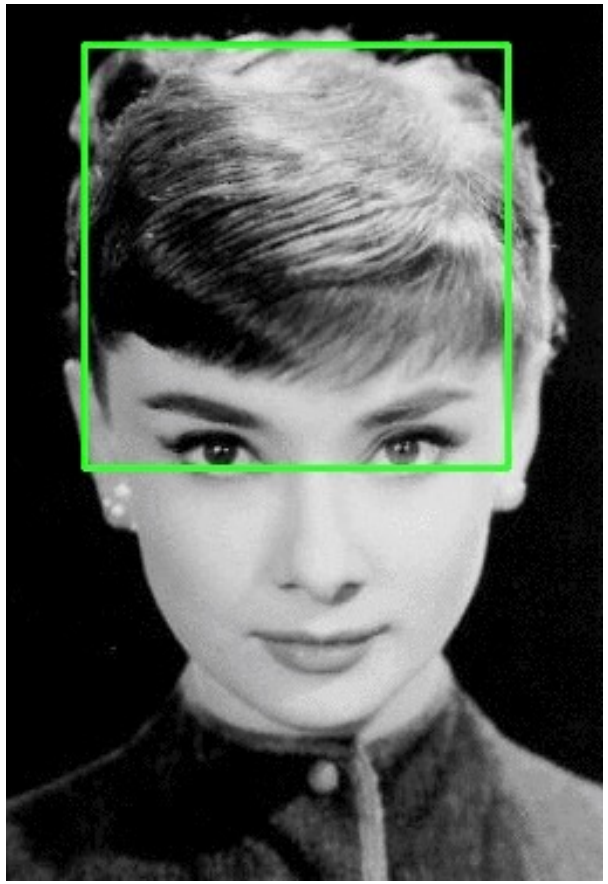


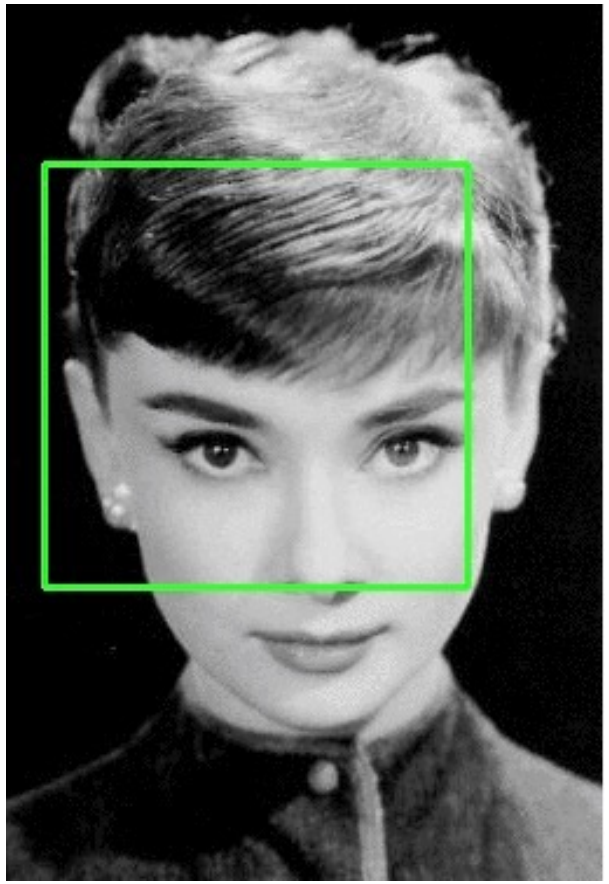


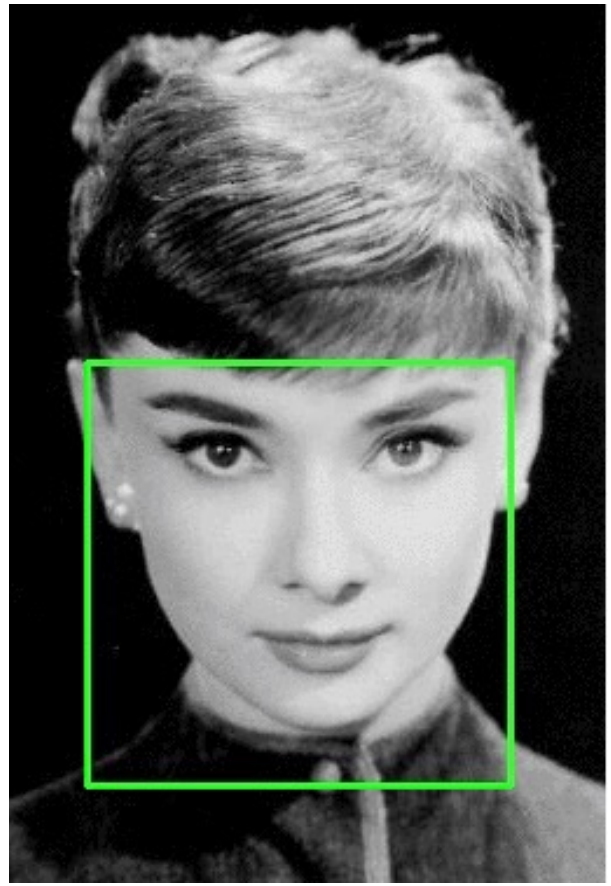
Viola et al. 2001

What is Deep about Deep Learning?







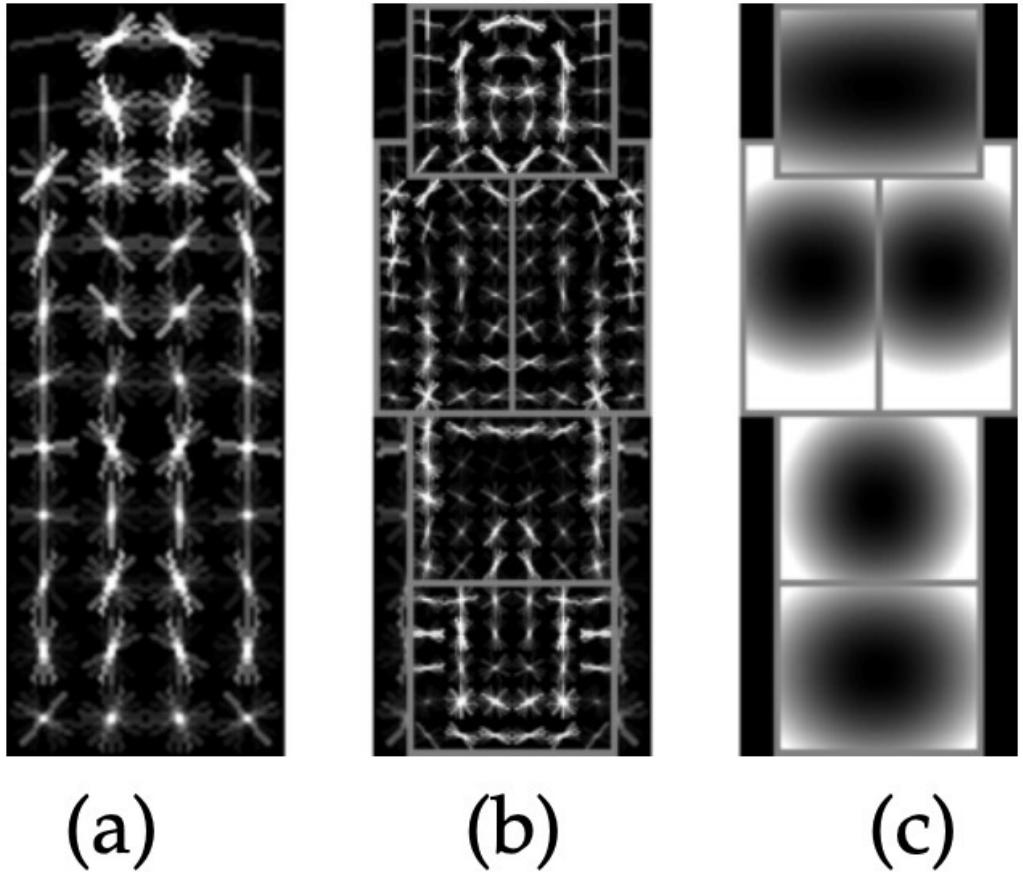


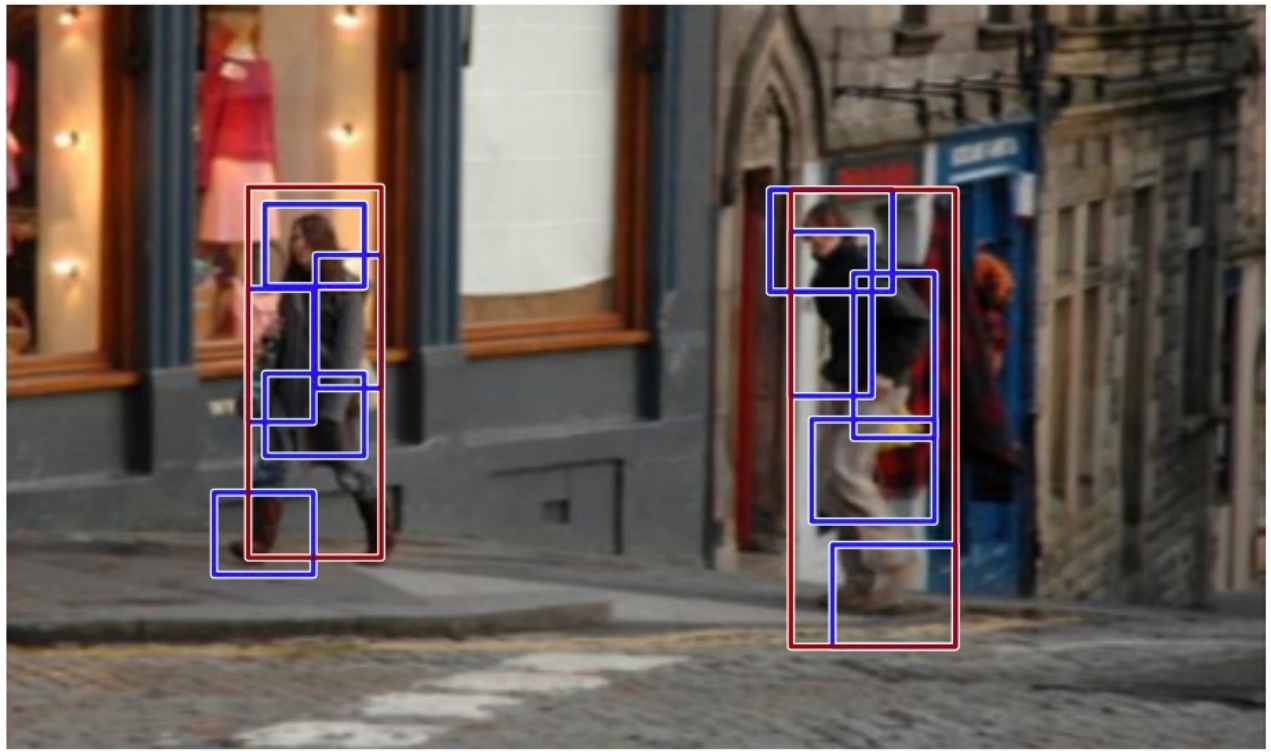
Histograms of Oriented Gradients. Dalal et al. 2005

What is Deep about Deep Learning?

Histograms of Oriented Gradients. Dalal et al. 2005

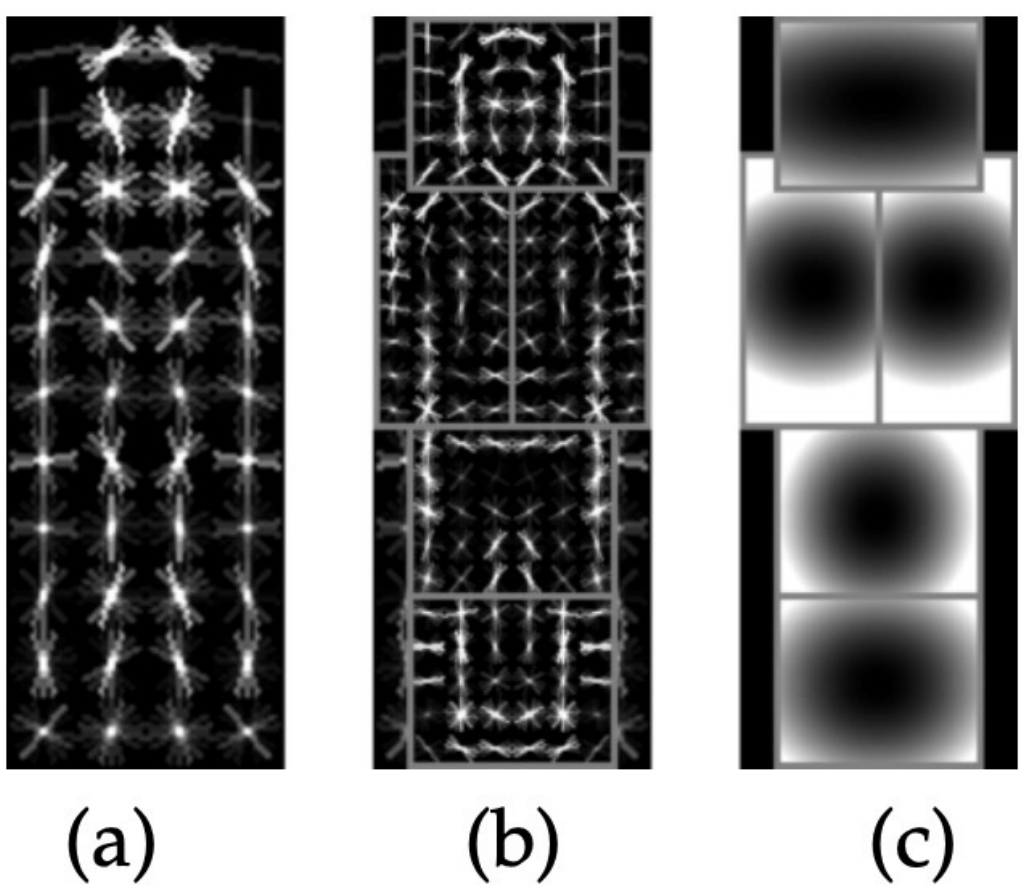
What is Deep about Deep Learning?

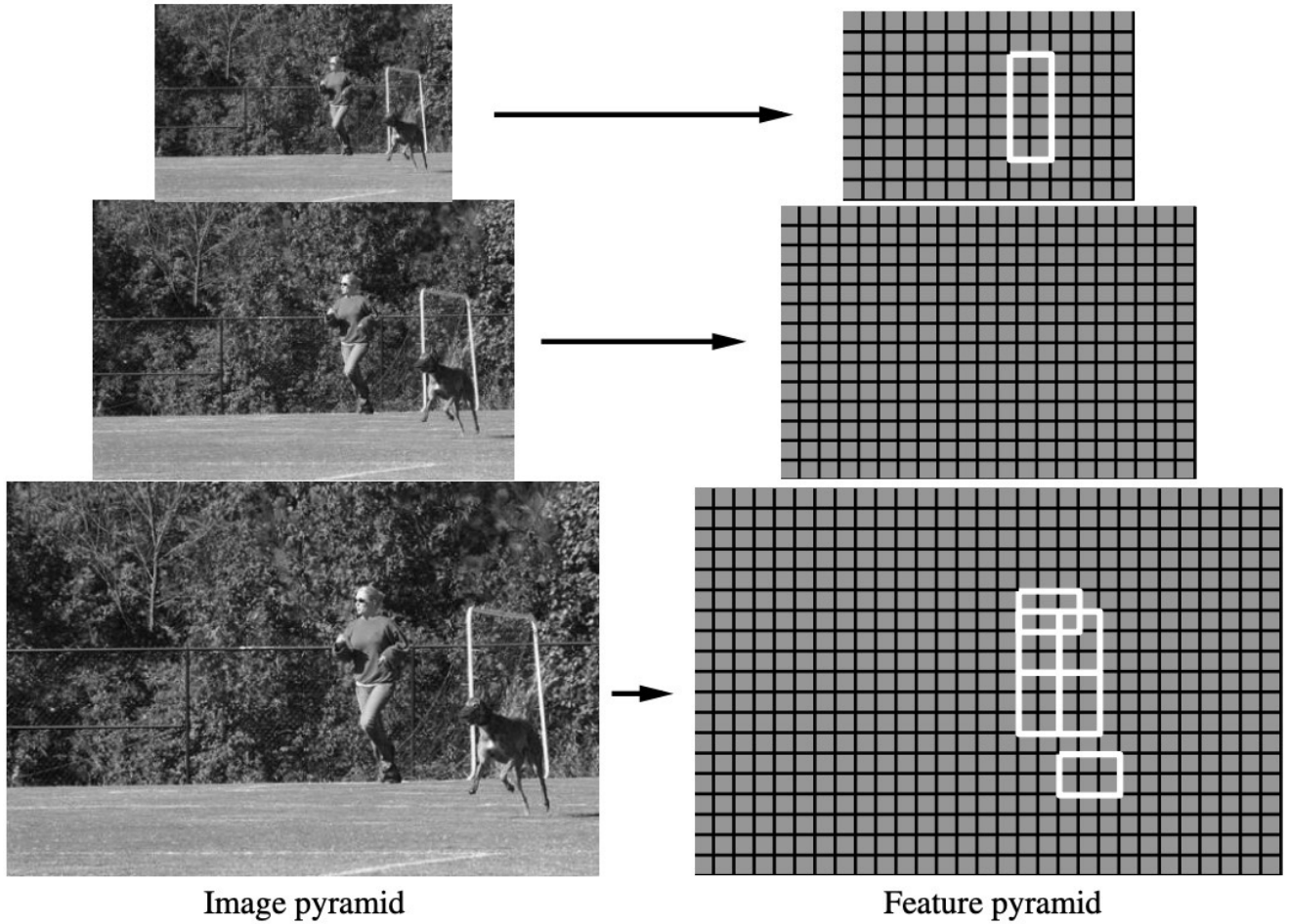




Discriminatively trained Part-based Models. Felzenszwalb et al. 2009

What is Deep about Deep Learning?





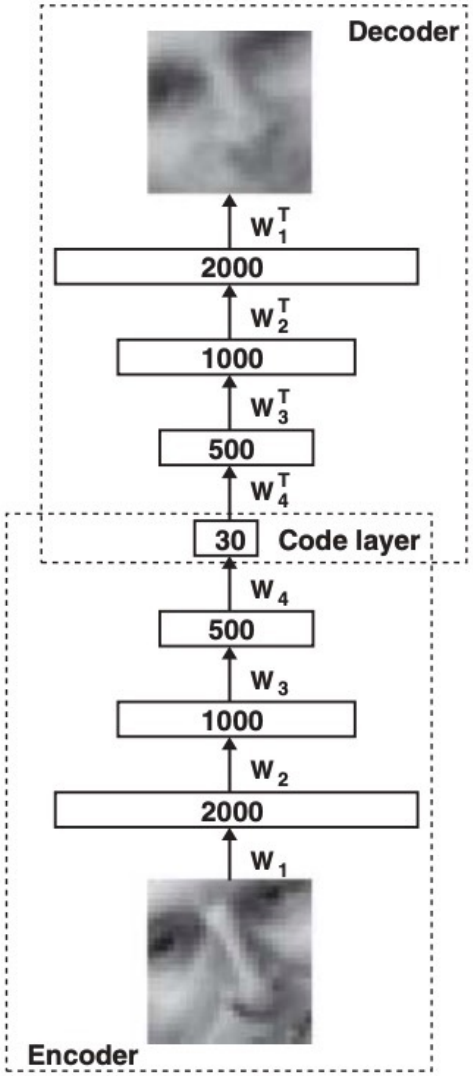
Discriminatively trained Part-based Models. Felzenszwalb et al. 2009

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.

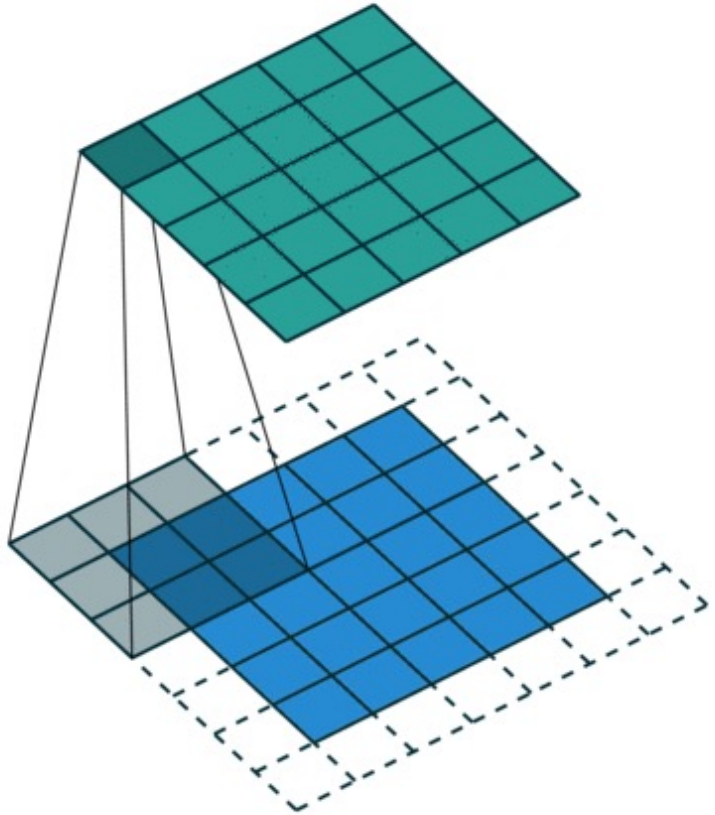
Different Types of Deep Networks

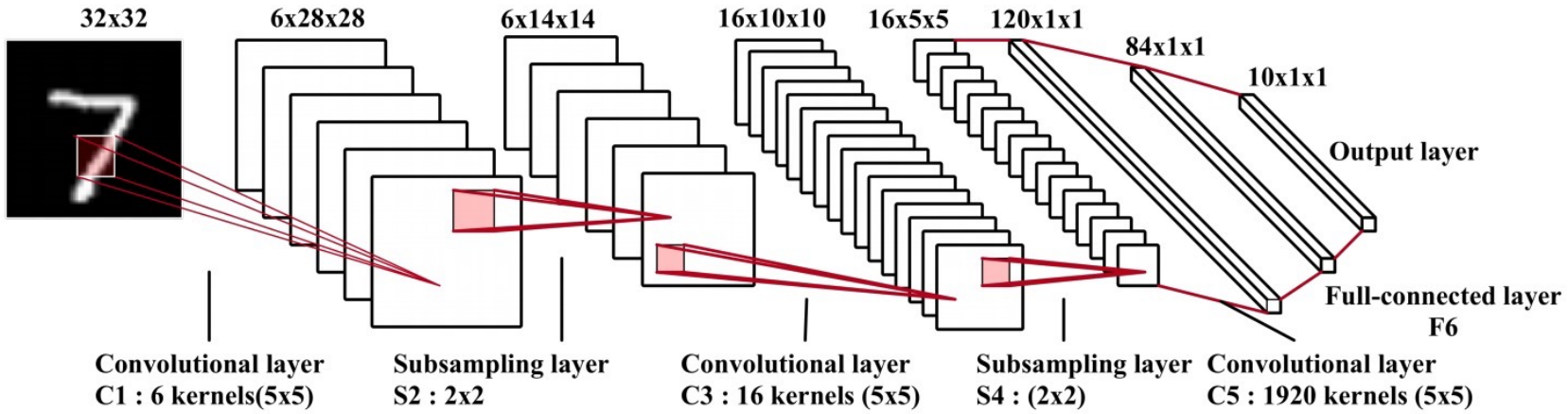


- Multilayer perceptron (MLP)
  - Input image I size : 32 x 32 = 1024
  - First hidden layer h_1 output size: 2000
  - First layer parameters W_1 size: 1024 x 2000
  - $h_1 = I\, W_1$

Hinton et al. 2006

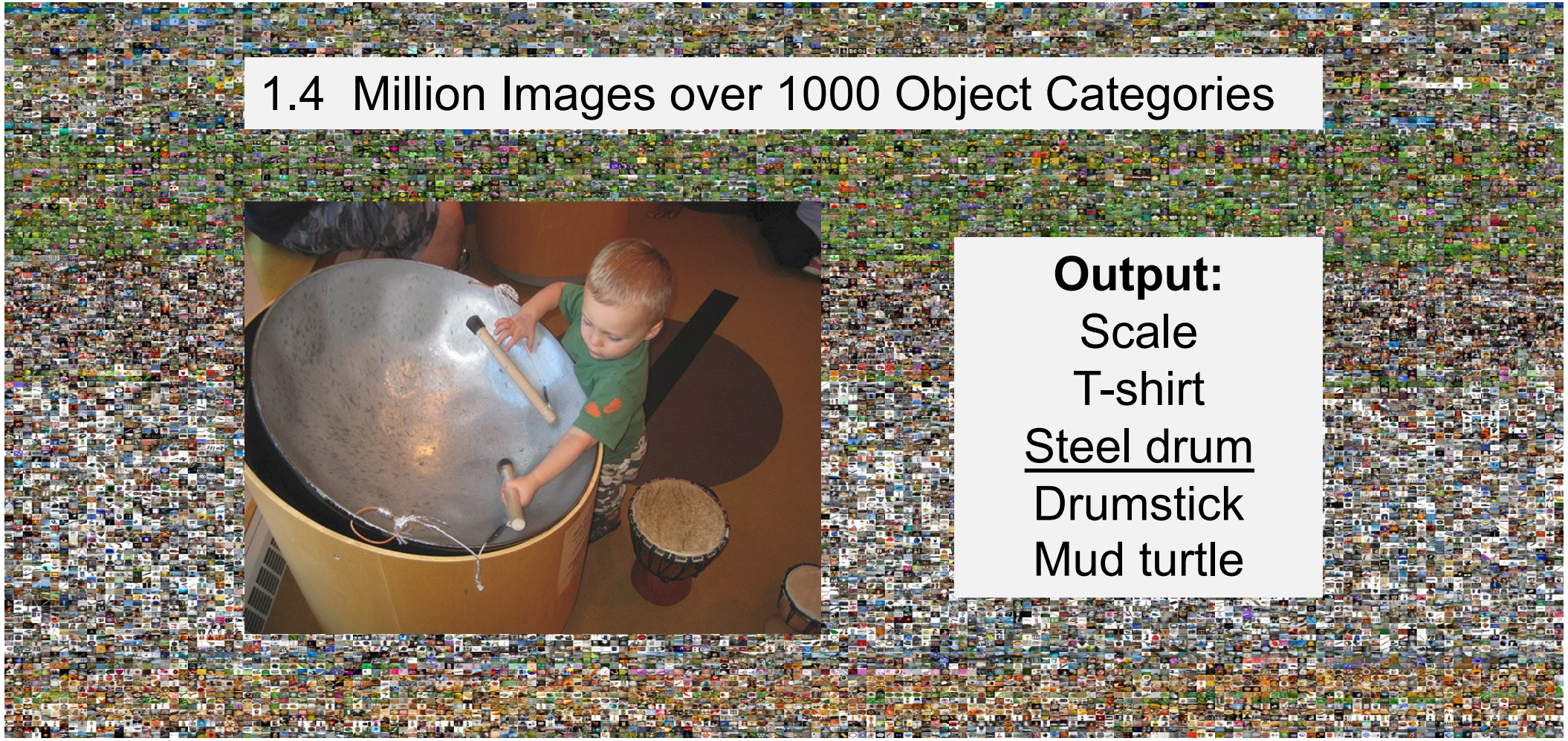
Different Types of Deep Networks





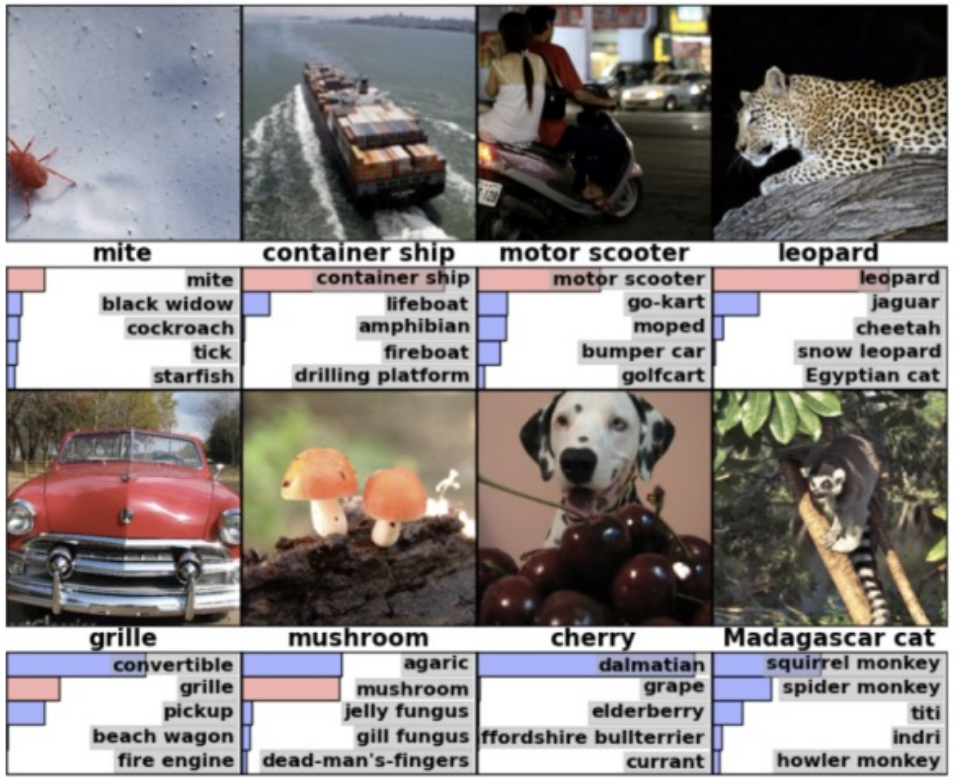
LeCun et al. 1990

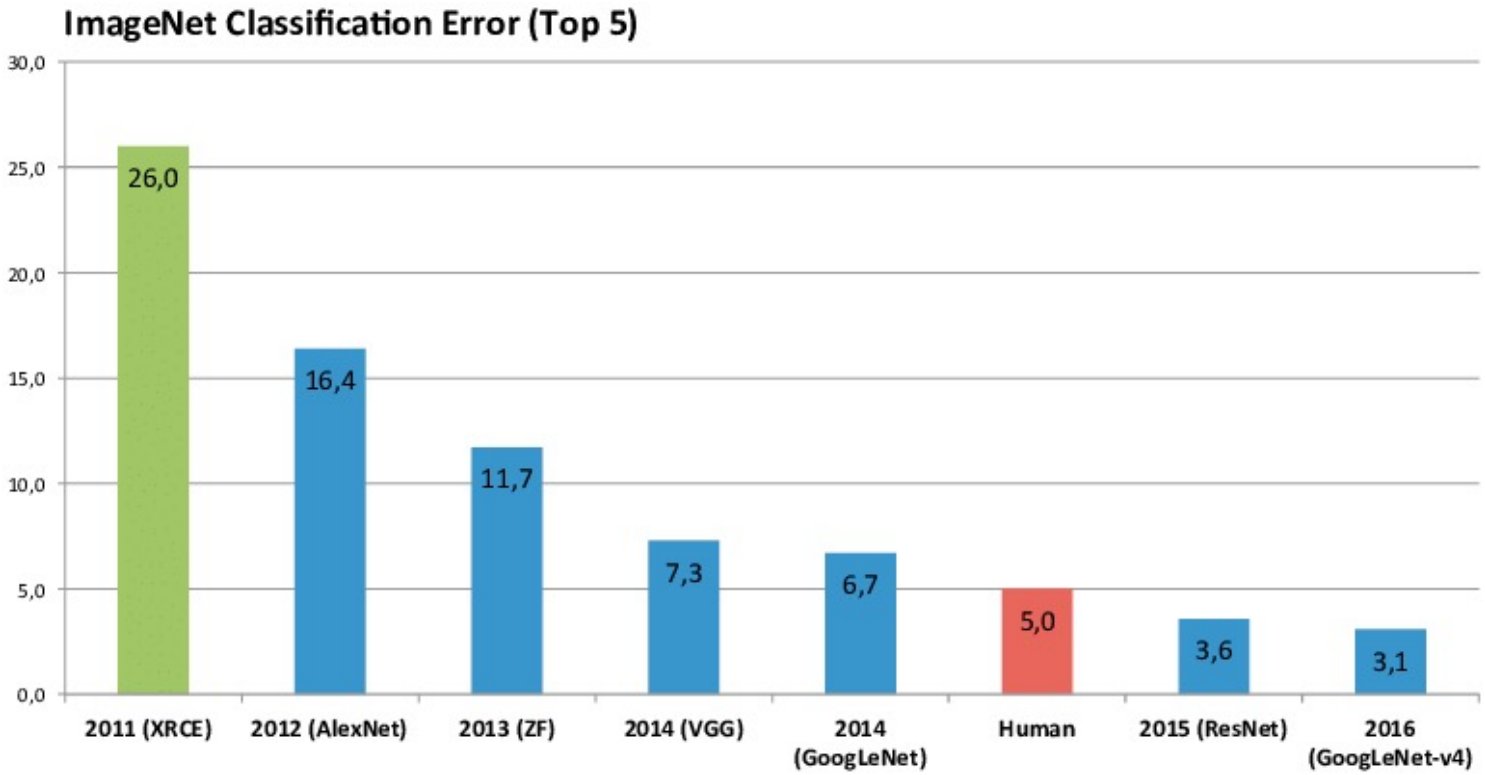
The ImageNet Challenge



Russakovsky et al. 2015

The ImageNet Challenge

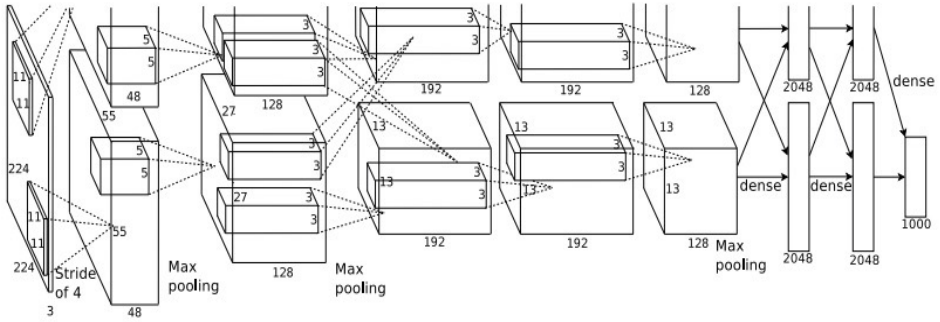


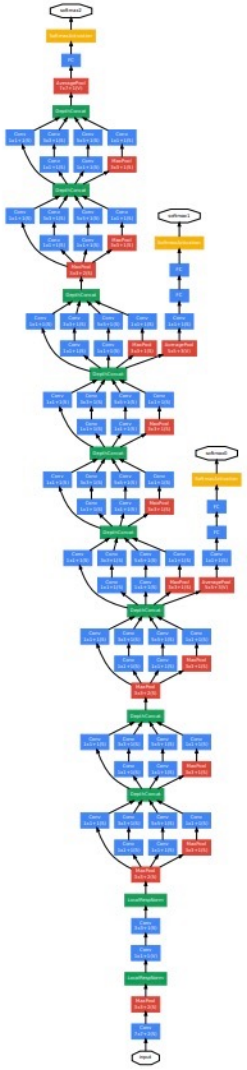


Where does the 5% error human performance come from?

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

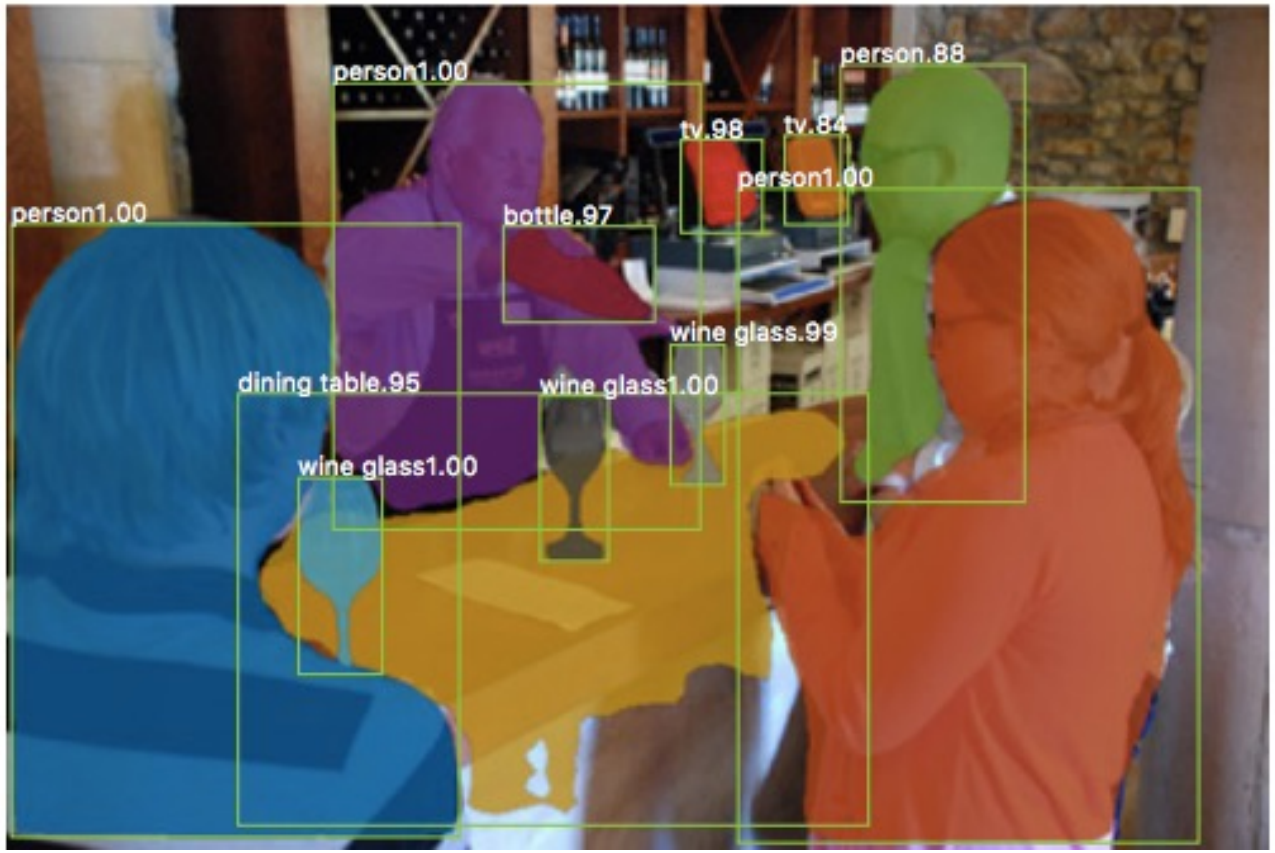
Many-Layer Networks





ResNets

Object Detection/Segmentation

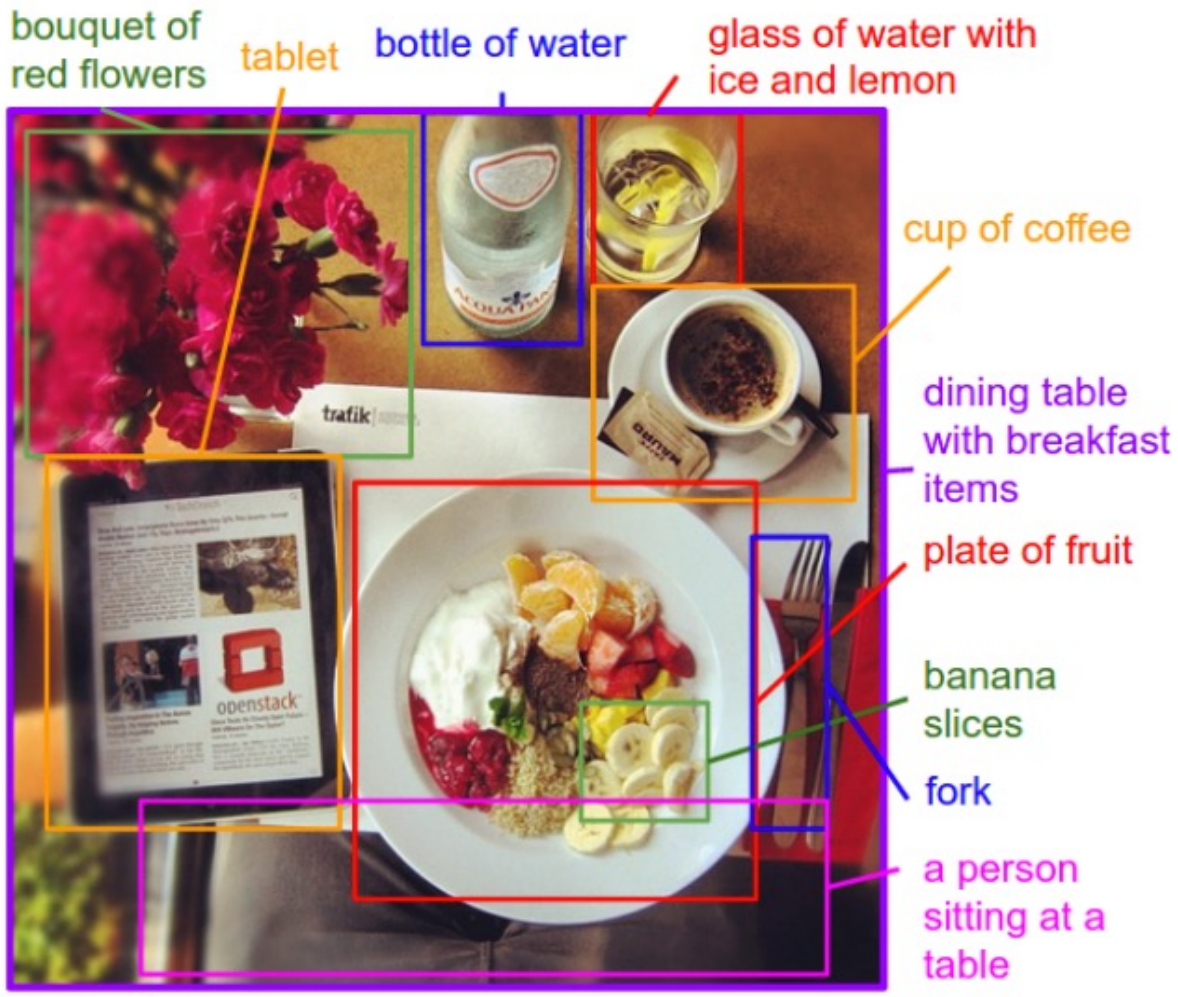


Mask R-CNN. He et al. 2017.

Human Pose Estimation

Mask R-CNN. He et al. 2017.

Image Captioning

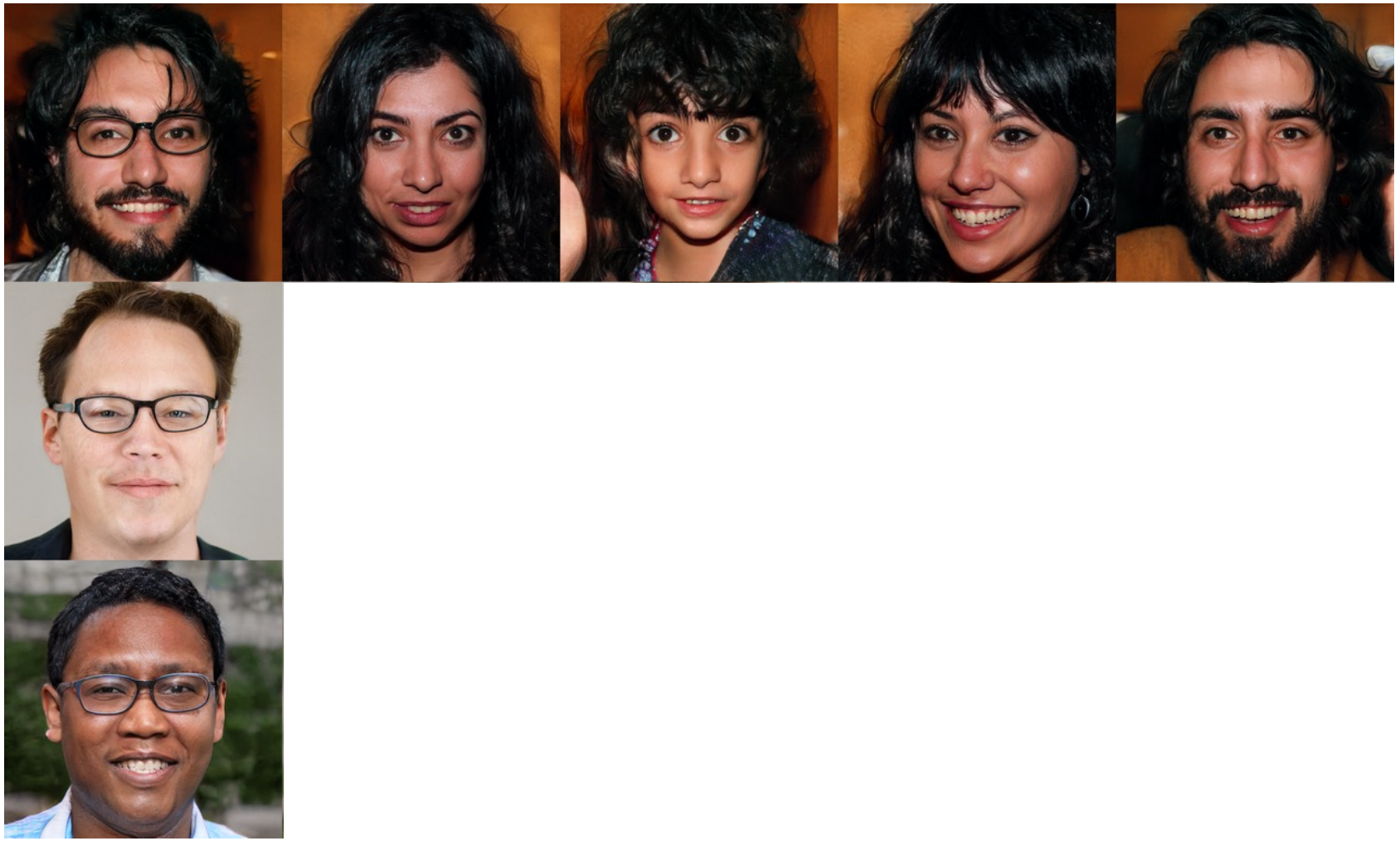


Karpathy et al. 2015.

Image generation

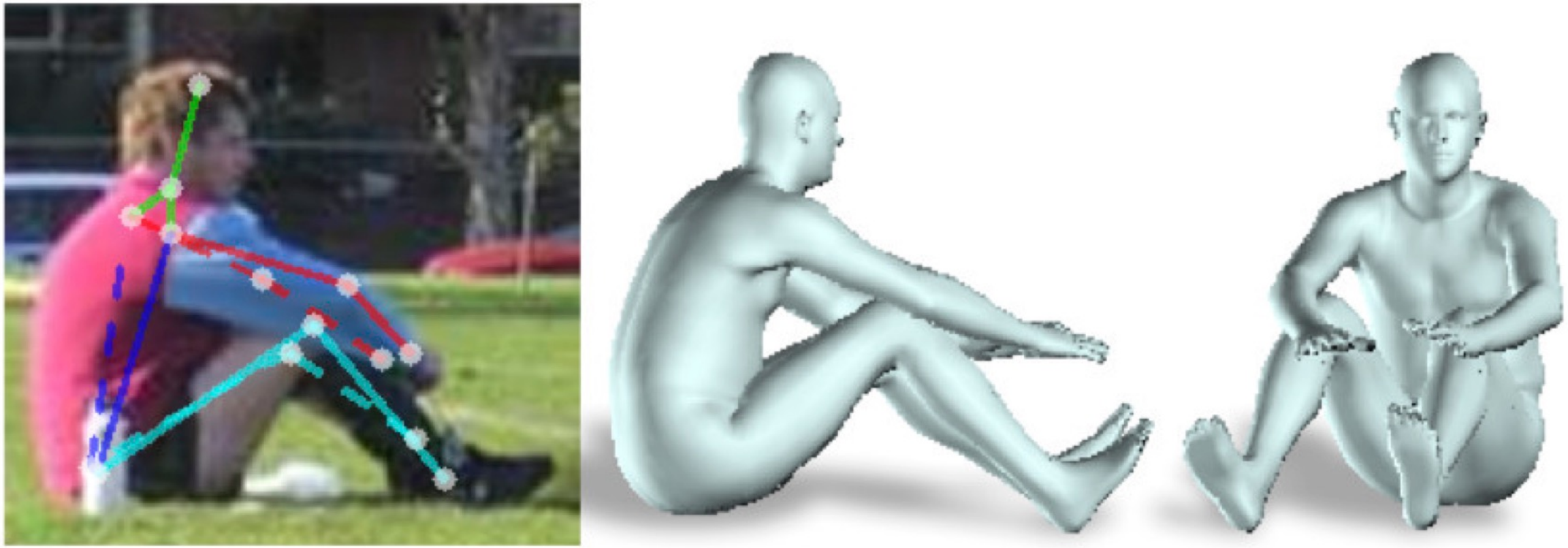
BigGAN. Brock et al. 2019.

Image generation



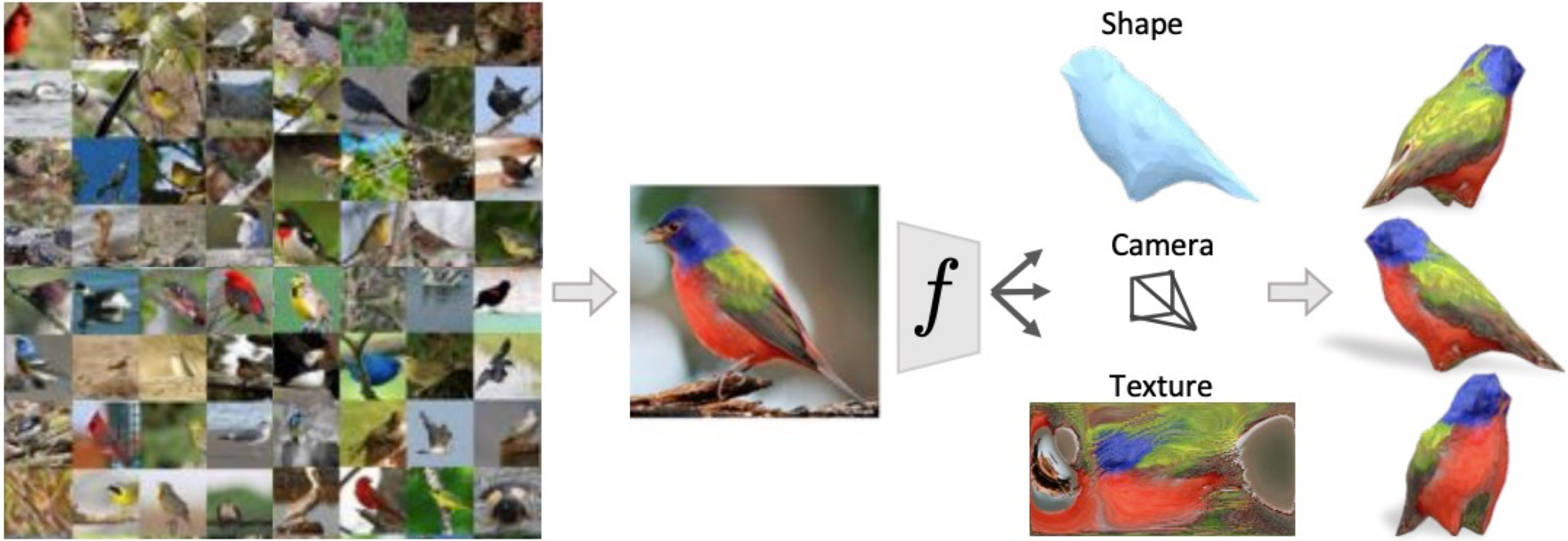
StyleGAN. Karras et al. 2018.

3D Reconstruction from a single image



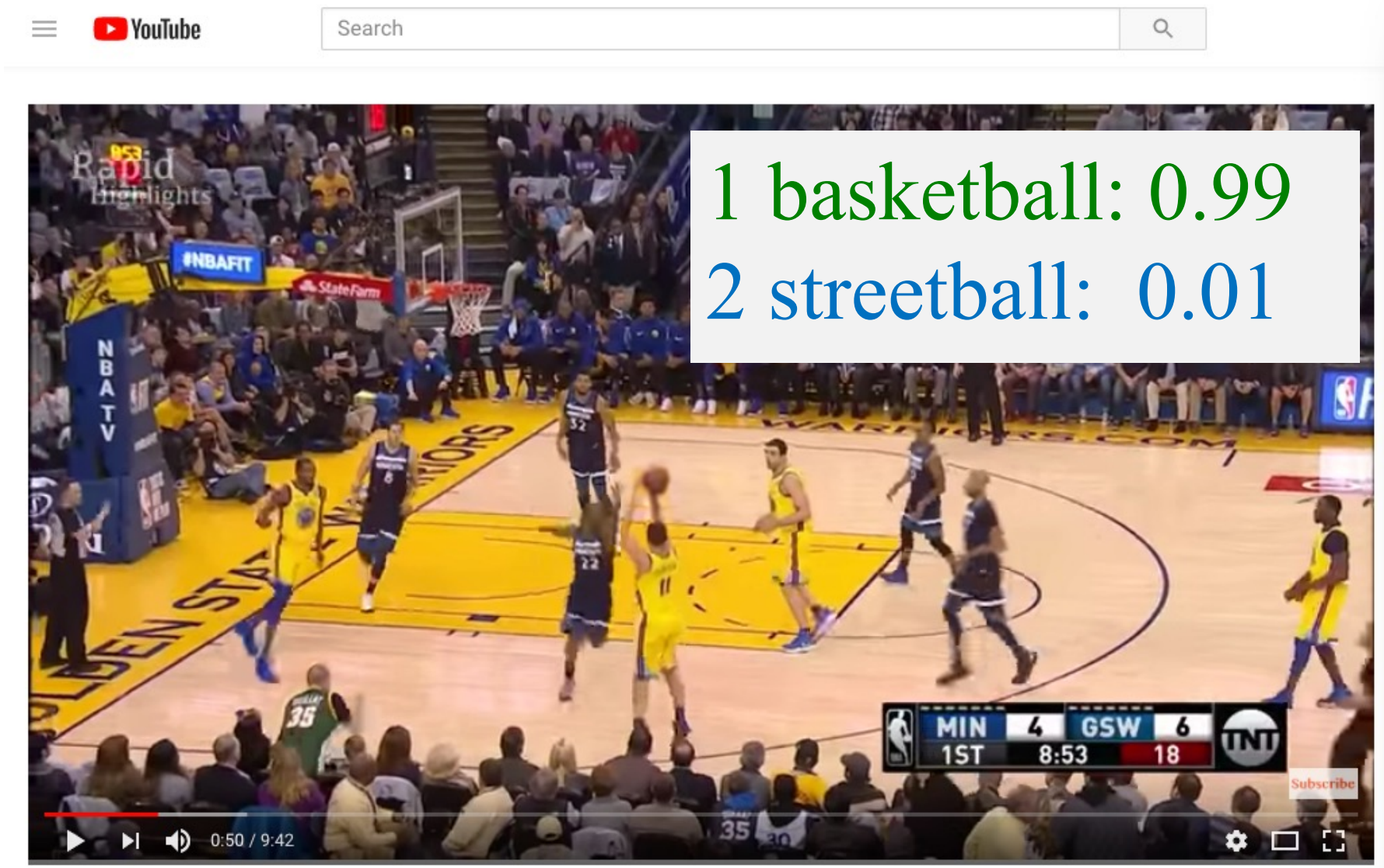
Bogo et al. 2016.

3D Reconstruction from a single image



Kanazawa et al. 2018.

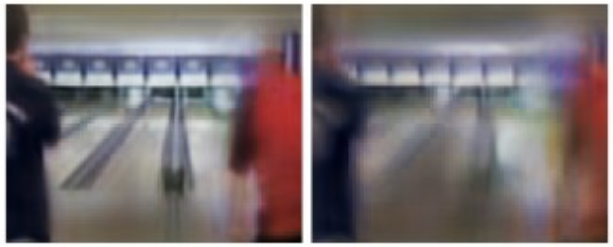
Action Recognition



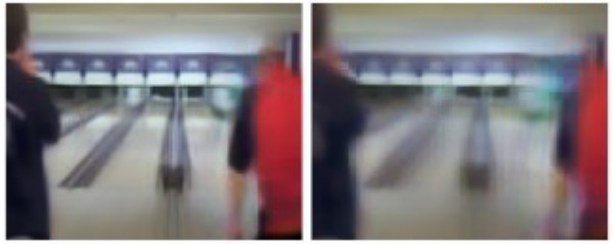
Video Prediction

Input frames

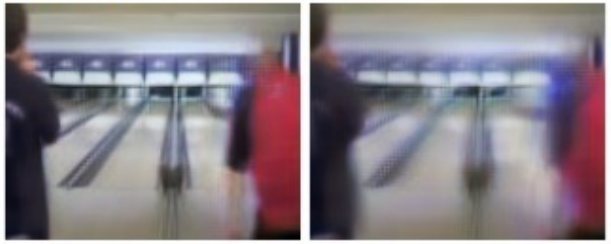
Ground truth



ℓ_2 result



ℓ_1 result



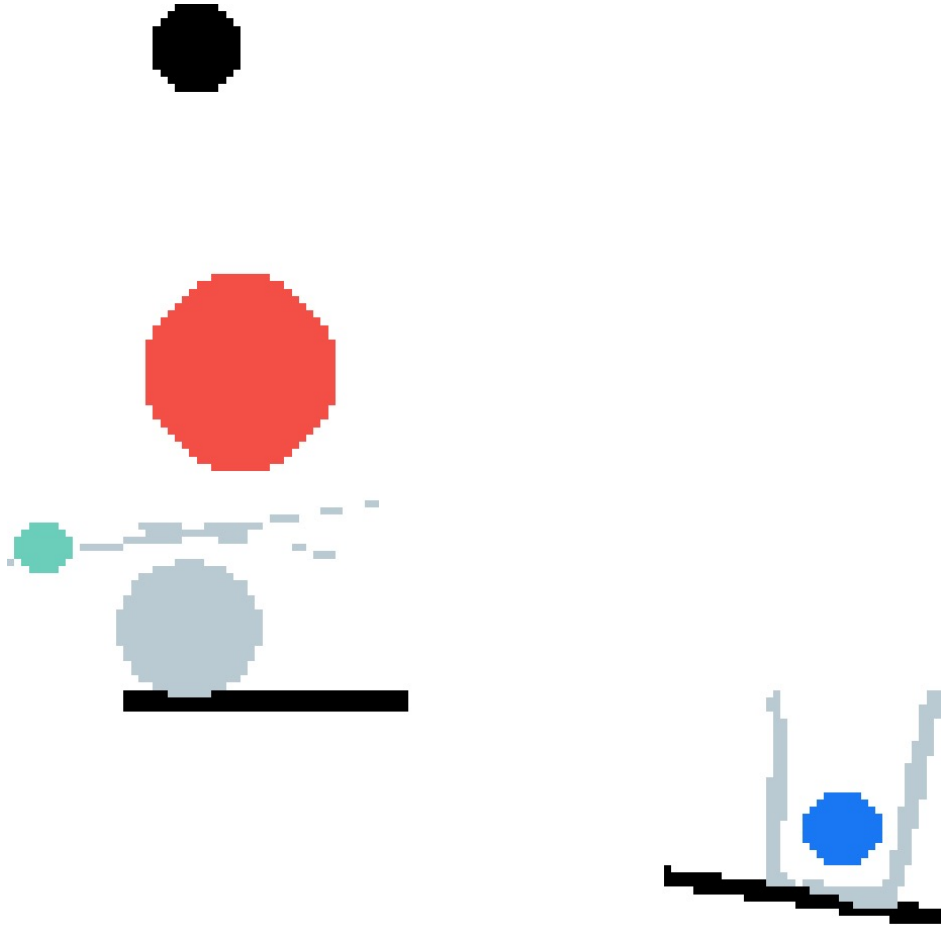
GDL ℓ_1 result

Adversarial result

Adversarial+GDL result

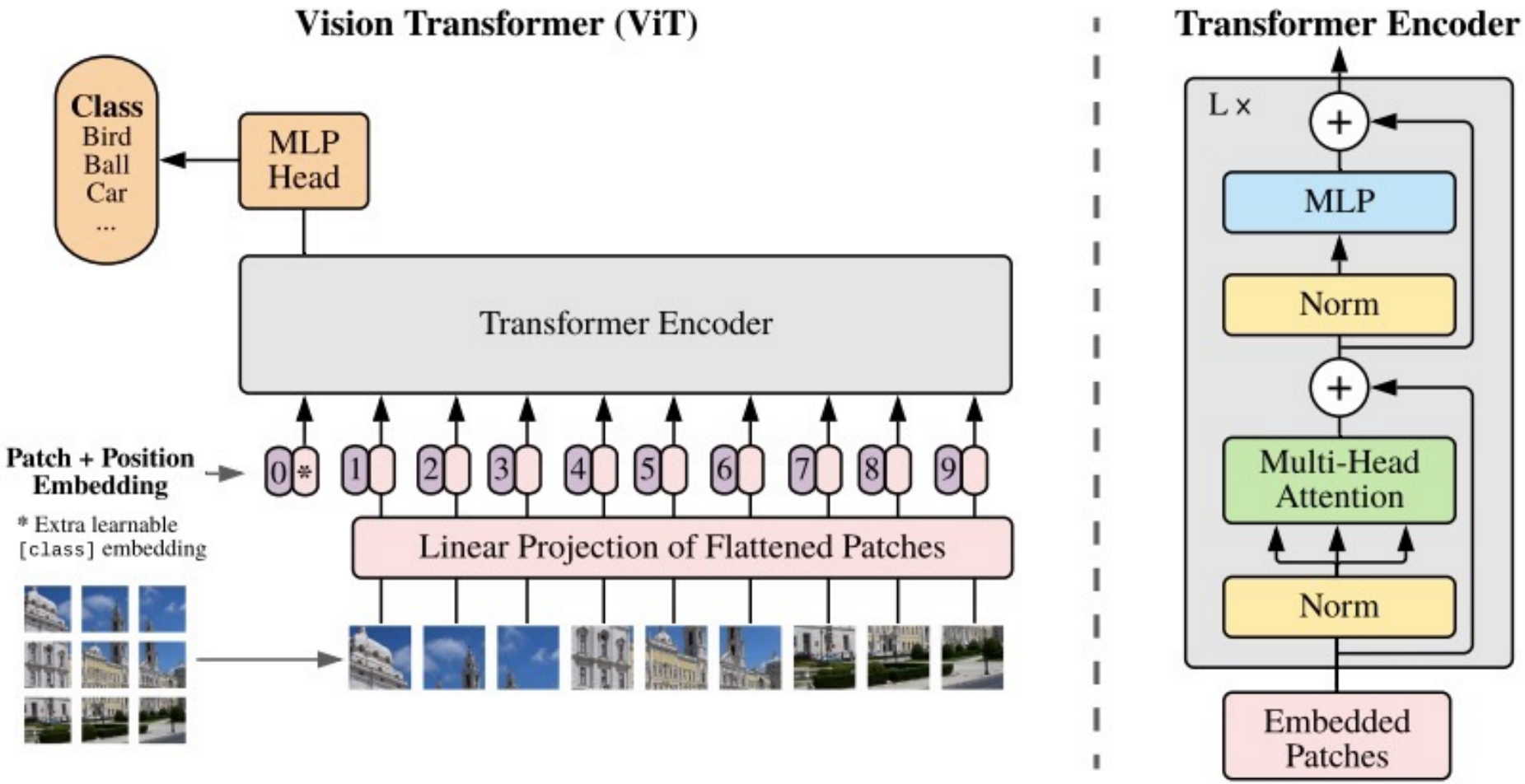
Mathieu et al., 2015

Physical Interaction Prediction



Qi et al., 2021

Transformer



Dosovitskiy et al., 2021

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

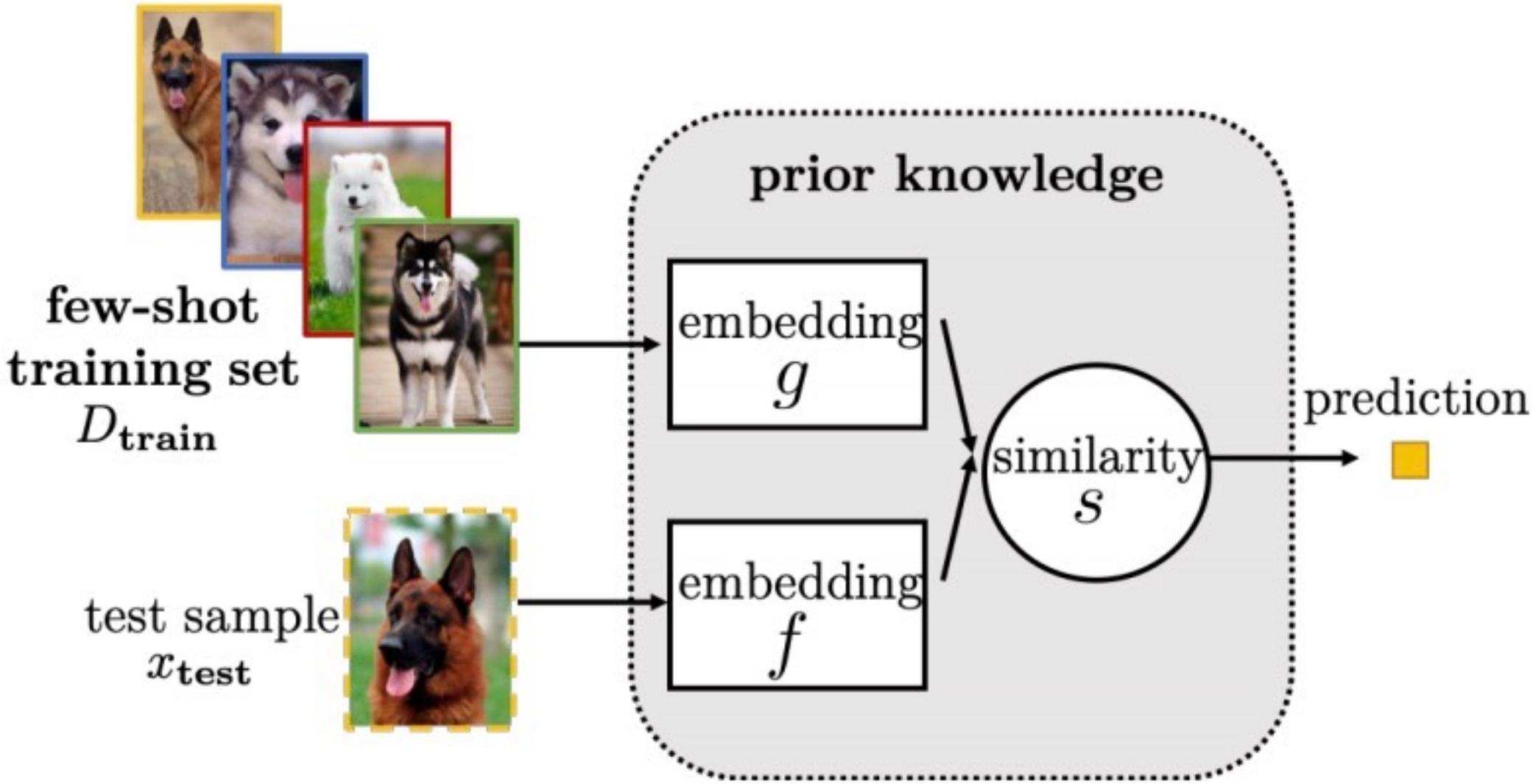
input image

- **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

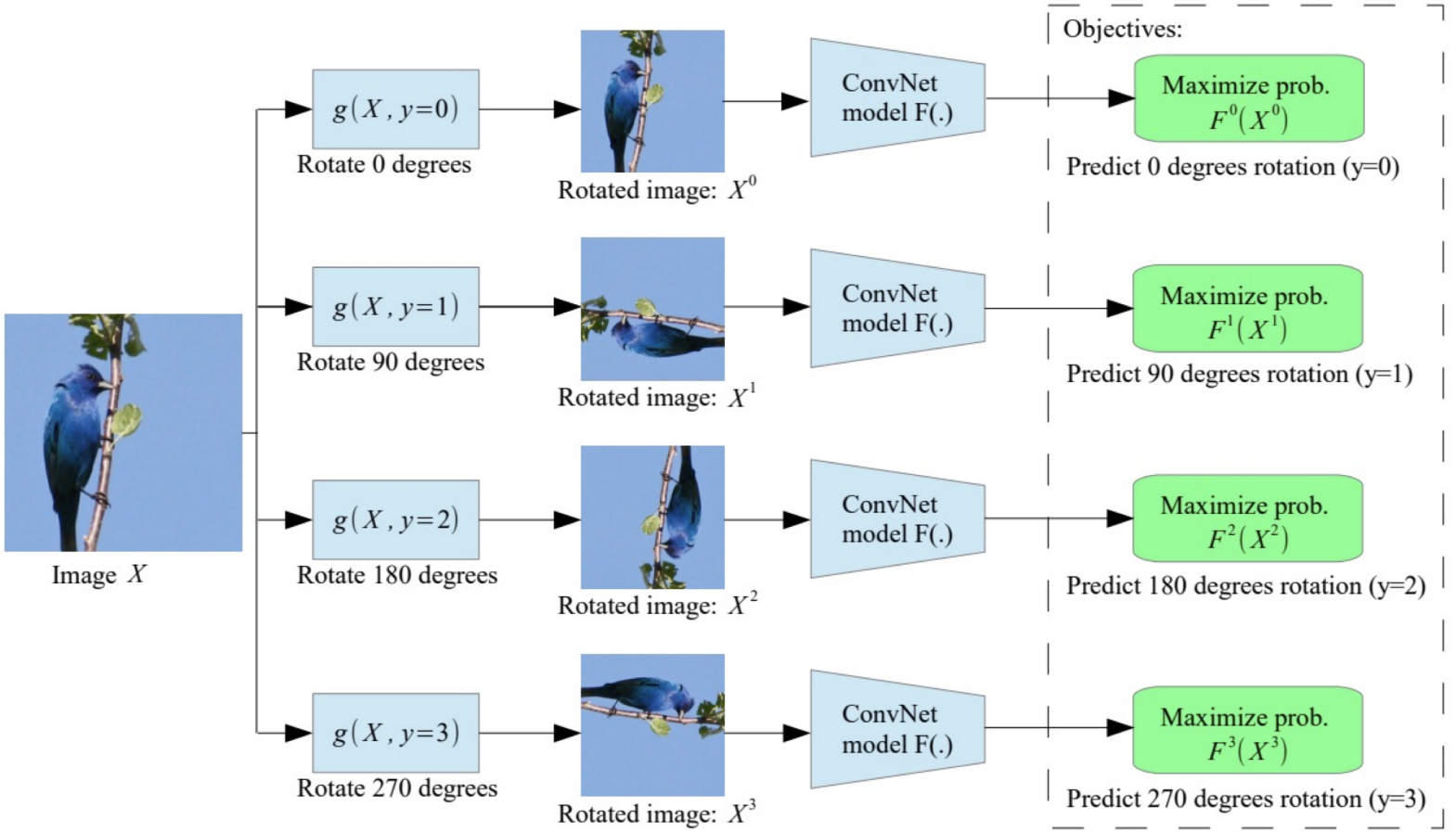
- **Testing** (or **inference**): apply neural network f to a new *test example* x and output the predicted value $y = f(x)$

Transfer Learning

Transfer Learning



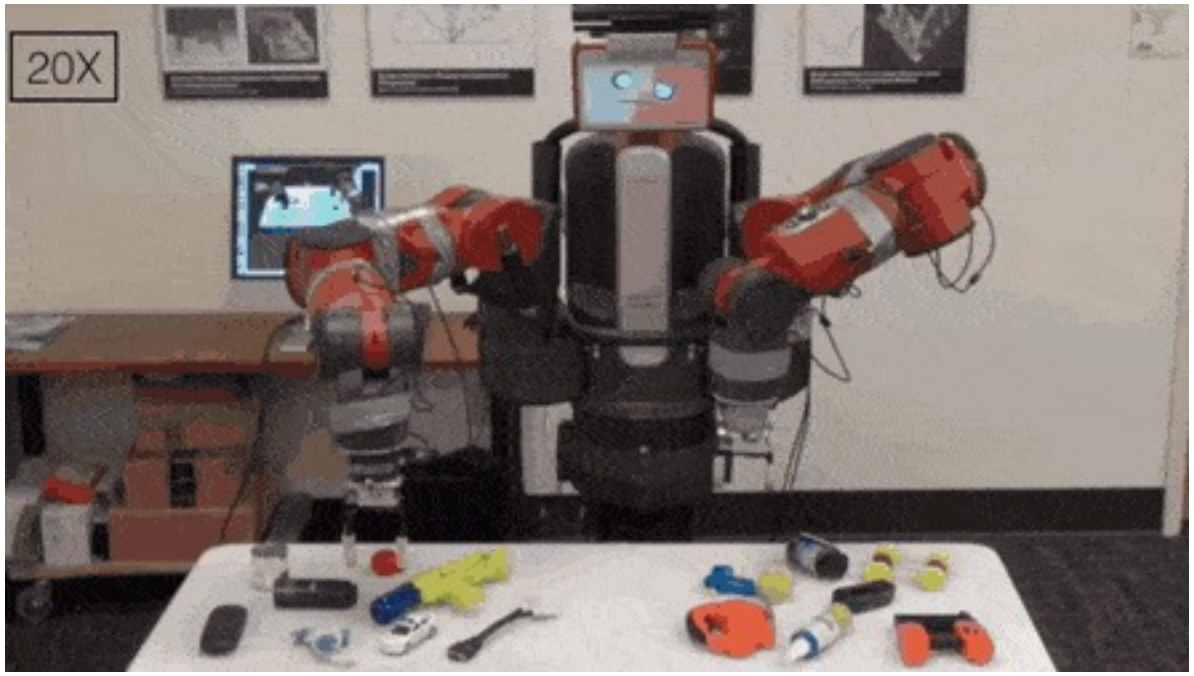
Self-Supervised Learning

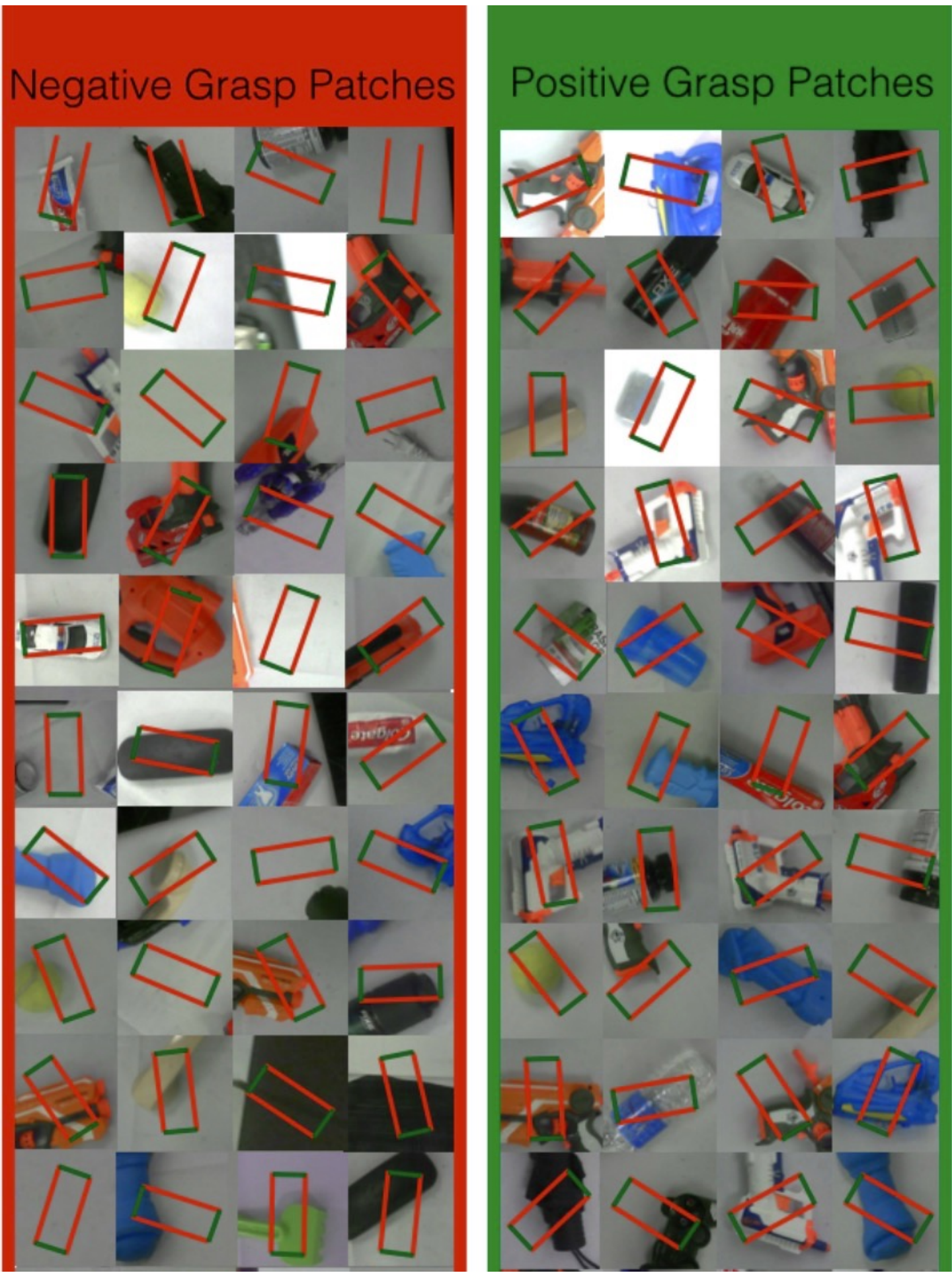


[Gidaris et al. 2018]

Self-Supervised Learning with Rotation Prediction

Self-Supervised Learning




Negative Grasp Patches
Positive Grasp Patches

Test Set

This Class

- Computer Vision Research with Deep Learning
- Training and Testing

Coming Tutorial

- There will be a tutorial on how to do/submit assignments NEXT Monday, 12:00 - 1:00 pm PST

- We will use the compute resources in https://datahub.ucsd.edu/

Last few things

- Piazza:
  - https://piazza.com/ucsd/spring2021/ece285
  - Discussions

- GradeScope:
  - https://www.gradescope.com/courses/256233
  - Entry Code:**V84YGX**
  - Submit assignments