Which one of these images is fake?
Type your answer in the chat
Class 1 [09/02/2020]. Introduction to Vision
Class 2 [09/14/2020]. Natural image statistics and the retina
Class 3 [09/21/2020]. The Phenomenology of Vision
Class 4 [09/28/2020]. Learning from Lesions
Class 5 [10/05/2020]. Primary Visual Cortex
October 12th: University Holiday
Class 6 [10/19/2020]. Adventures into terra incognita
Class 7 [10/26/2020]. From the Highest Echelons of Visual Processing to Cognition
Class 8 [11/02/2020]. First Steps into in silico vision
Class 9 [11/09/2020]. Teaching Computers how to see
Class 10 [11/16/2020]. Computer Vision
FINAL EXAM, PAPER DUE 12/14/2020. No extensions.
Some computer vision problems

Classification

Classification+Localization

Object detection

Instance segmentation

Face identification

Action recognition

Many more…
- Face detection
- Distance estimation
- Video prediction
- Image captioning
Many more architectures

VGG (2014) – 6.8%
Baidu (2015) – 5.33%

MSRA (2015) – 4.94%
- Spatial pyramid pooling
- Optimized PReLU
- Improved (random) initialization

GoogleNet (2014) – 6.67%
- Inception module
- Multi-scale convolutions (including 1x1 filters)
- Minimal dense layers
- Auxiliary classifiers

AlexNet (2012) – 15.3%
Clarifai (2013) – 11.7%

Inception Module
Note: lots of parameters!!!

- Image of 256 x 256 x 3 pixels = 196,608 inputs
- 1000 output categories (imagenet)
- Simplest scenario: go from pixels to outputs
- ~ 200 x $10^6$ parameters
- ~ $10^6$ training images in ImageNet
Data, data, data

- Galaxies
- Plants
- Clinical images
- Cell types

Many more …
- MNIST
- IMAGENET
- MSCOCO
- QUICKDRAW
ImageNet

~1,000 categories
~1,000 images/category
Computational models can approximate neuronal responses along the ventral visual cortex.
The better the biological approximation the better performance in computer vision tasks.
Predicting eye movements during visual search

Zhang et al, 2018
Predicting eye movements during visual search

Zhang et al 2018
Machines surpass humans in pattern recognition tasks

Face recognition better than forensic experts and human “superrecognizers” (Phillips et al 2018)

Plant and animal classification (iNaturalist, Van Horn et al. 2018) ~ 1M photos from 5,089 taxa and 13 “super-classes”: expert human levels and better than naïve observers

Pose tracking in animal biomedical research (Matthis 2018)
Computer vision can help segment biological images
Computer vision for action recognition

A
- PlayingCello 65
- BreastStroke 62
- BenchPress 56
- BrushingTeeth 91
- BodyWeightSquats 101
- BlowDryHair 85
- Bowling 13
- SoccerJuggling 62

B
- No
- Yes
- Drinking

C
- No
- Yes
- Reading
Automatic pose estimation for ethology research

Mathis et al 2018
Face recognition by computer vision

Same or different?

Phillips et al 2018
Face recognition by computer vision

Phillips et al 2018
Species classification and detection

Van Horn et al 2018
Figure 7. Sample detection results for the 2,854-class model that was evaluated across all validation images. Green boxes represent correct species level detections, while reds are mistakes. The bottom row depicts some failure cases. We see that small objects pose a challenge for classification, even when localized well.

Van Horn et al 2018
Applications of computer vision to clinical diagnosis

- Excellent performance in many clinical diagnosis tasks
  - E.g. breast tumor detection
  - E.g. diabetic retinopathy

- Reliability, consistency, accuracy

- Machines can discover properties in the data that humans never even thought of before
  - E.g. cardiovascular disease risk from fundus photographs

- Beware of incidental findings

- Beware of biases in training data
Generative adversarial networks (GANs)
Deep Dreaming

Simonyan et al. 2014
Kreiman 2019
Xdream: Discovering neuronal tuning preferences

Ponce, Xiao, et al 2019
Style transfer

Gatys 2015
The portrait of Edmond de Belamy

Sold at Christie’s auction: $432,500
Predicting the next video frames

William Lotter, David Cox
PredNet captures neurophysiological properties!

William Lotter, David Cox
Adversarial examples

- schoolbus
- add this "noise"
- ostrich

Szegedy 2013
Models of ventral visual cortex provide a first order approximation to visual behavior (e.g., recognition, eye movements)

Models of ventral visual cortex provide a first-order approximation to neural responses

Computer vision has shown major strides in the last decade in many applications
  - Face recognition
  - Clinical diagnosis
  - Object segmentation
  - Tracking behavior
  - Action recognition

Inverting recognition models yields powerful image generators

A model that predicts what will happen next can learn in a self-supervised manner and captures fundamental responses in visual cortex
Visual Object Recognition
Computational Models and Neurophysiological Mechanisms
Neuro 130/230. Harvard College/GSAS 78454

Web site: http://tinyurl.com/visionclass
Class notes, Class slides, Readings Assignments

Location: Biolabs 2062
Time: Mondays 03:00 – 05:00

Lectures: Faculty: Gabriel Kreiman (and invited guests)
TA: Will Xiao

Contact information:
Gabriel Kreiman
gabriel.kreiman@tch.harvard.edu
Will Xiao
xiaow@fas.harvard.edu

617-919-2530
Office Hours: Before class (Mondays 2pm), after class (Mondays 5pm). By appointment