

Visual Object Recognition

Computational Models and Neurophysiological Mechanisms

Neurobiology 230. Harvard College/GSAS 78454

Class 1 [09/01/2021]. Introduction to Vision

Note: no class on 09/06/2021

Class 2 [09/13/2021]. Natural image statistics and the retina

Class 3 [09/20/2021]. The Phenomenology of Vision

Class 4 [09/27/2021]. Learning from Lesions

Class 5 [10/04/2021]. Primary Visual Cortex

Note: no class on 10/11/2021

Class 6 [10/18/2021]. Adventures into *terra incognita*

Class 7 [10/25/2021]. From the Highest Echelons of Visual Processing to Cognition

Class 8 [11/01/2021]. First Steps into in silico vision [Will Xiao]

Class 9 [11/08/2021]. Teaching Computers how to see

Class 10 [11/15/2021]. Computer Vision

Class 11 [11/22/2021]. Connecting Vision to the rest of Cognition

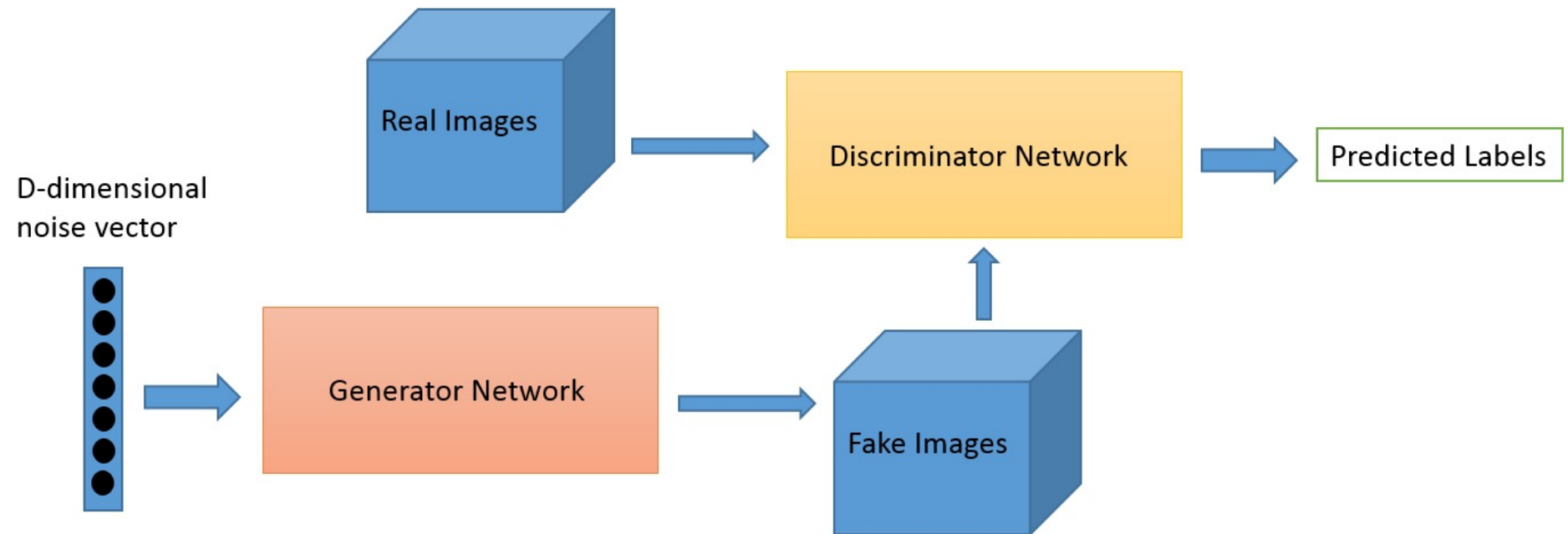
Class 12 [11/29/2021]. Visual Consciousness

FINAL EXAM, PAPER DUE 12/14/2021. No extensions.

What is common to all these faces?



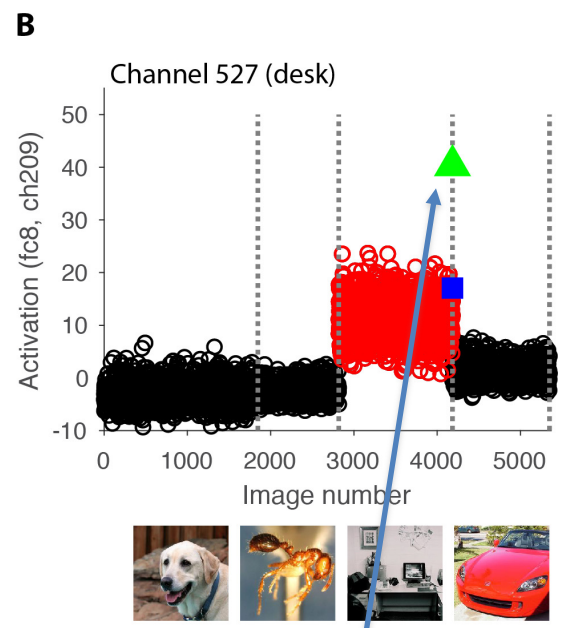
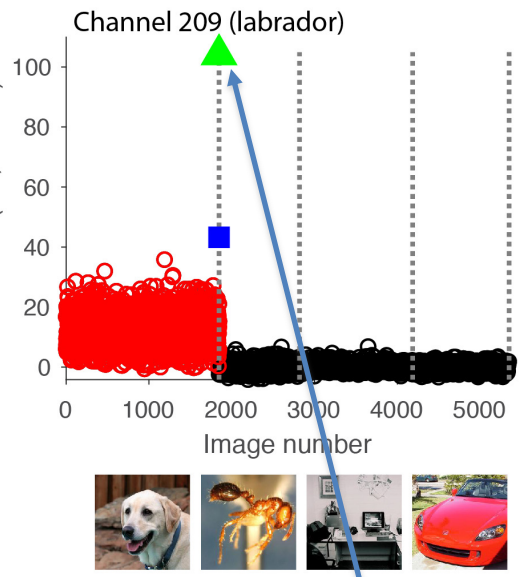
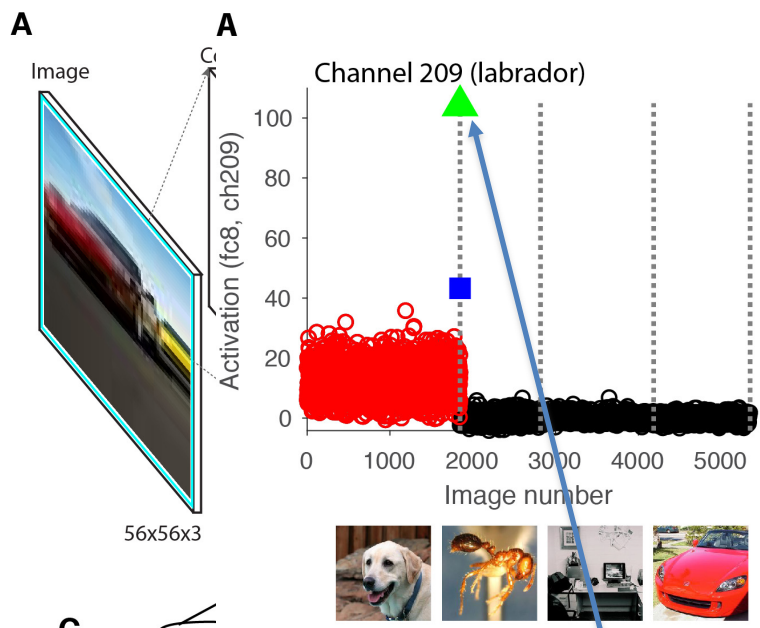
Generative adversarial networks (GANs)



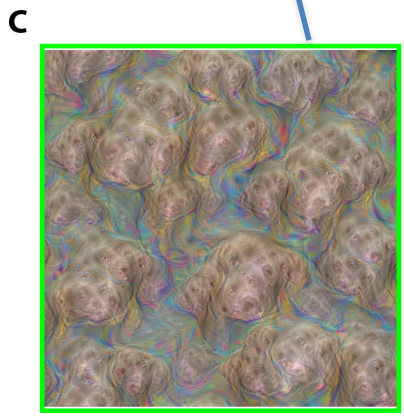
<https://www.nytimes.com/interactive/2020/11/21/science/artificial-intelligence-fake-people-faces.html?searchResultPosition=1>

Goodfellow 2014

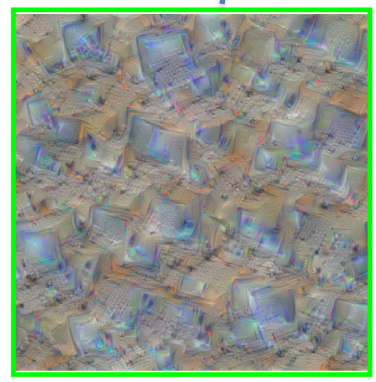
Deep Dreaming



C

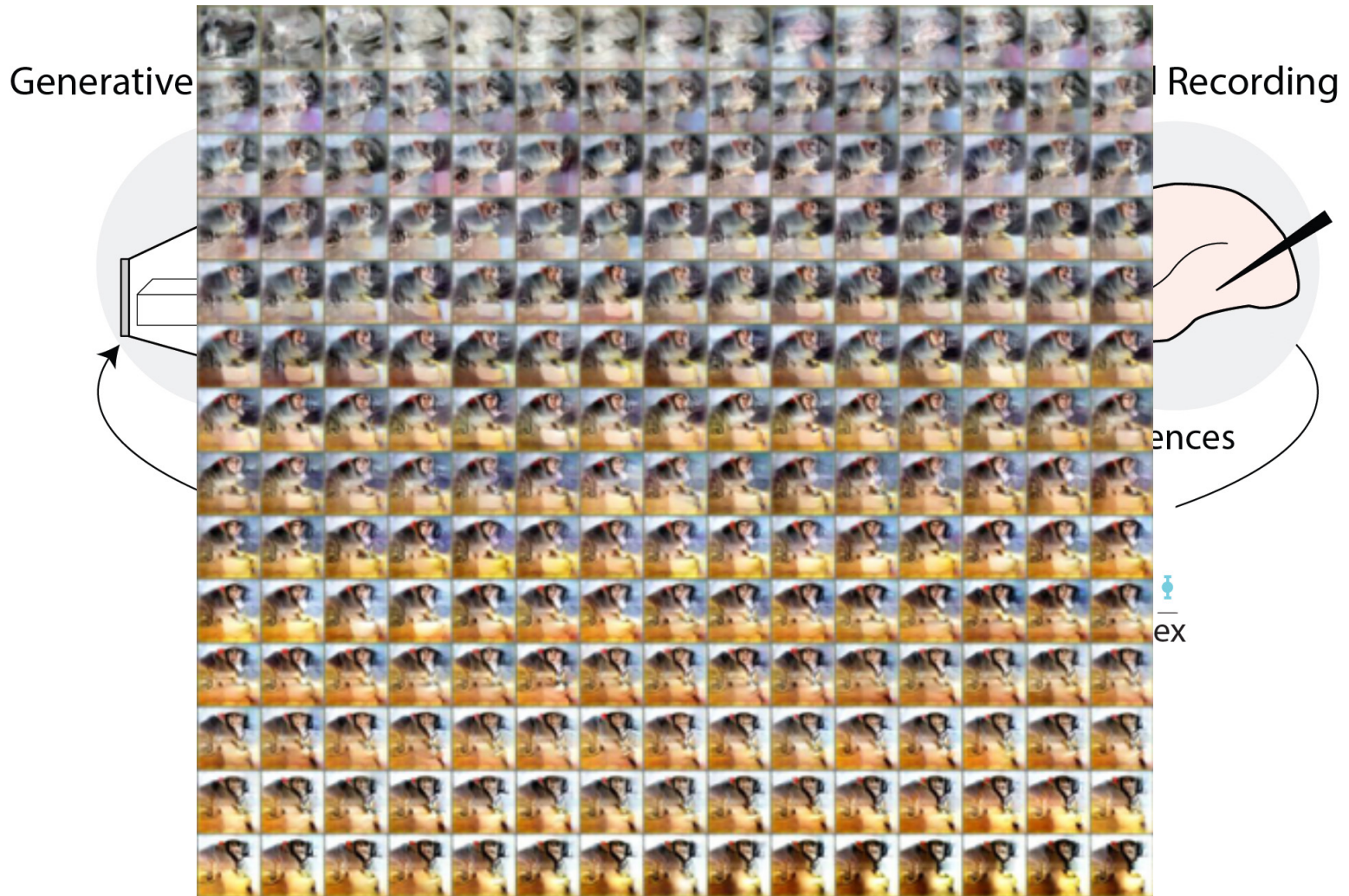


D

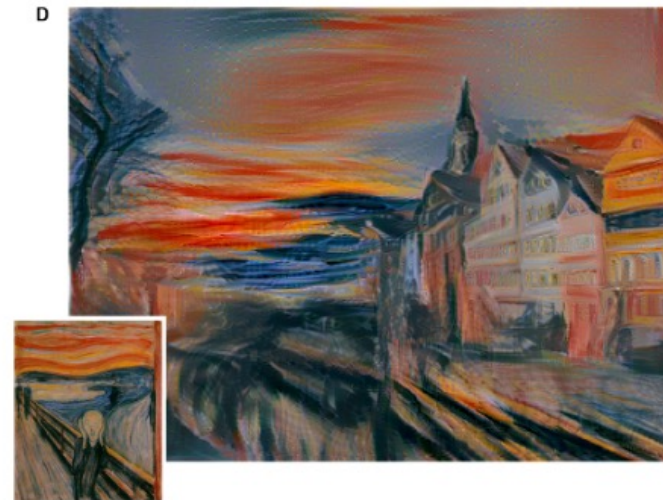


r)

Xdream: Discovering neuronal tuning preferences



Style transfer



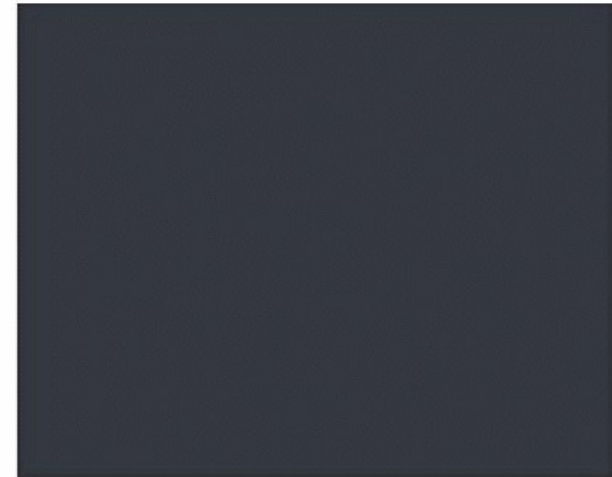
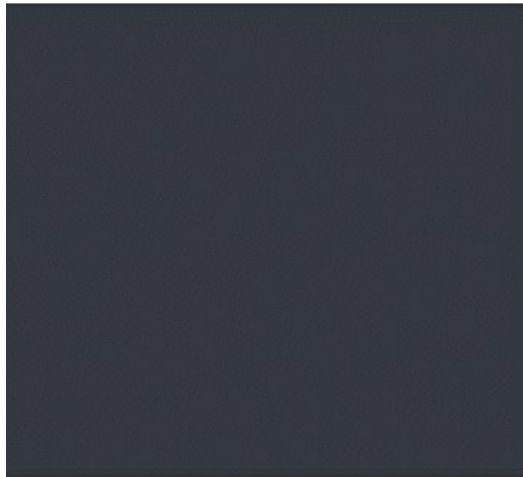
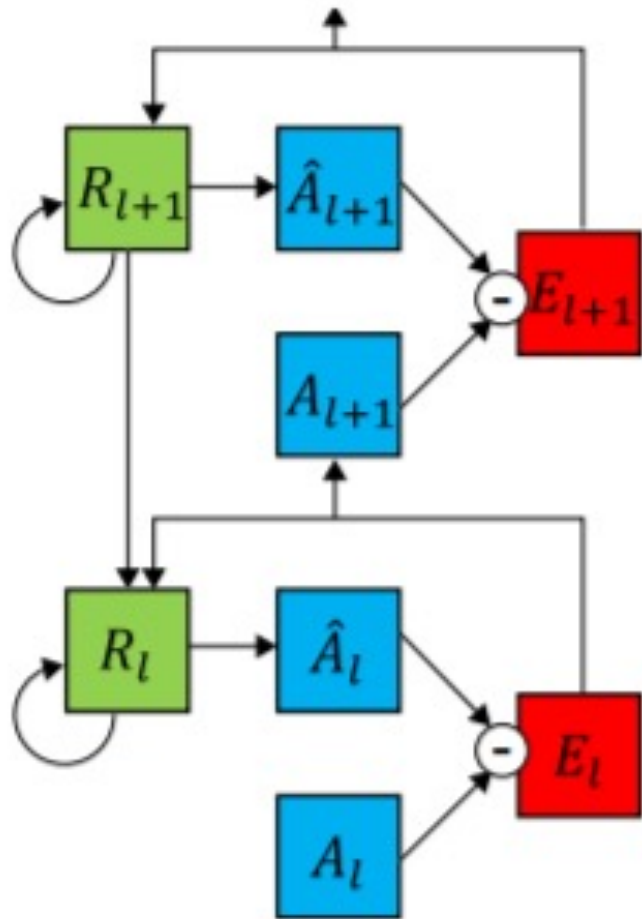
Gatys
2015

The portrait of Edmond de Belamy

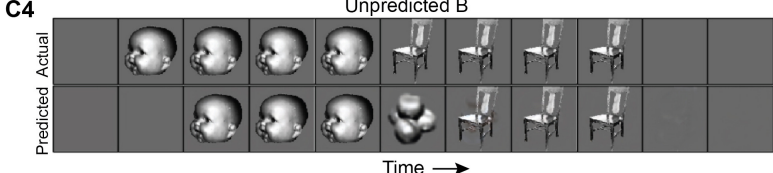
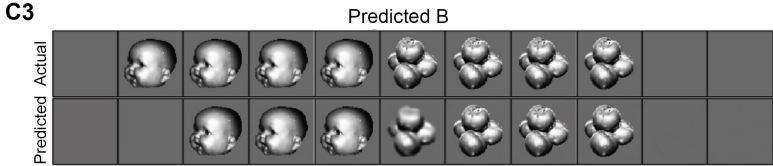
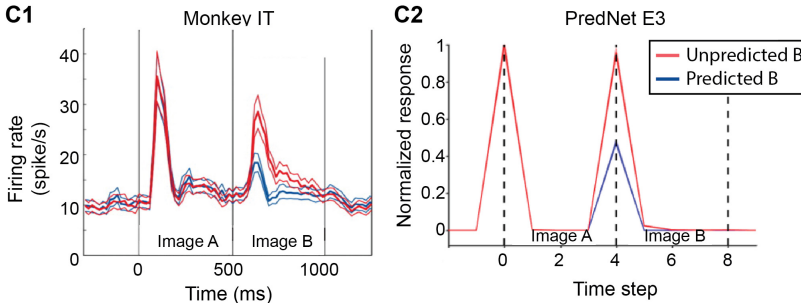
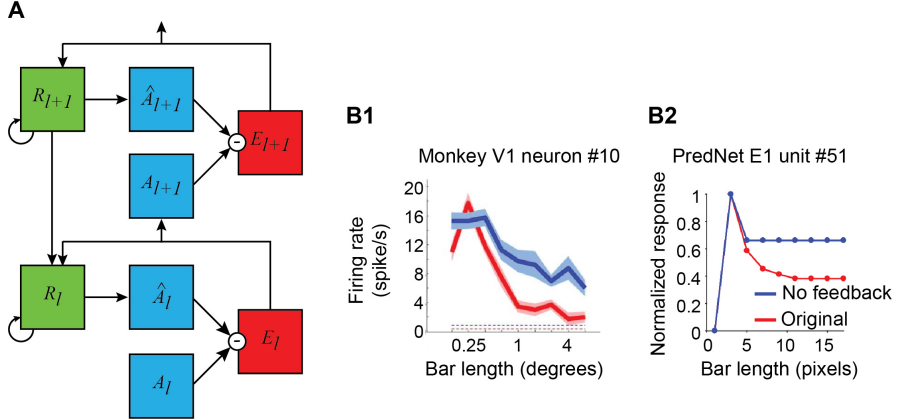


Sold at Christie's auction: \$432,500

Predicting the next video frames



PredNet captures neurophysiological properties!



The Turing test for vision



Adversarial examples

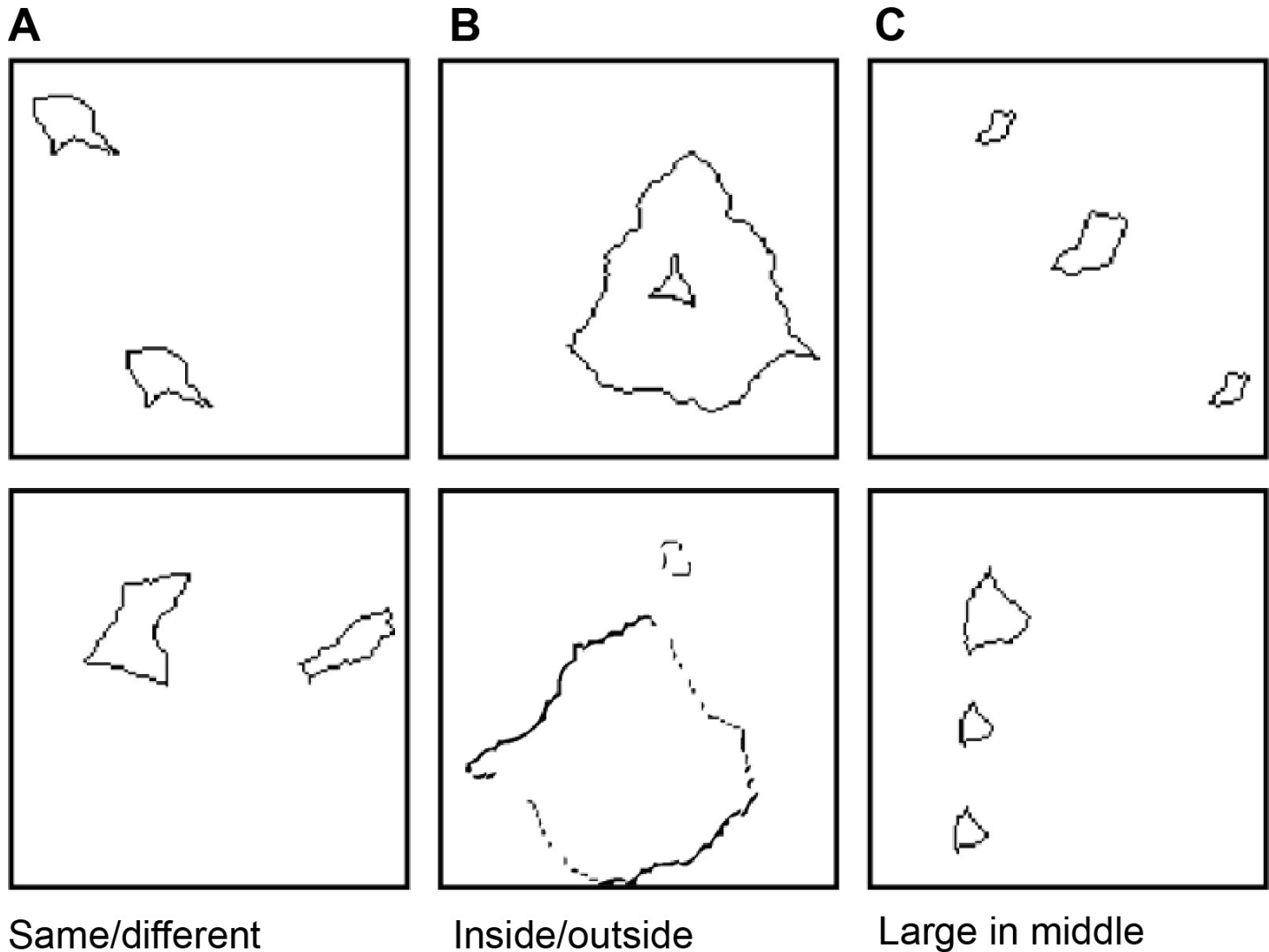
schoolbus

add this "noise"

ostrich



Example visual reasoning tasks



Answering questions on an image



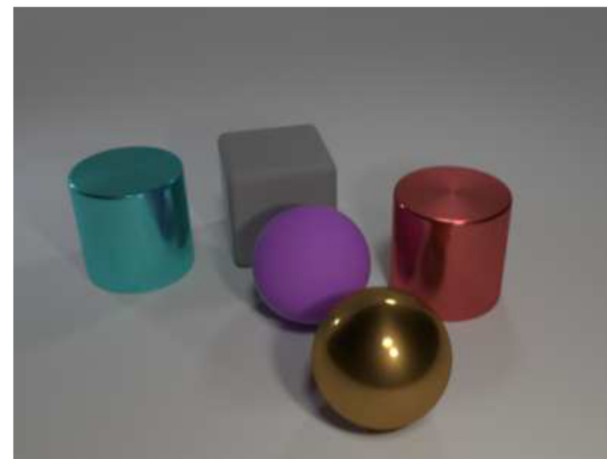
How many chairs are at the table?



Is there a pedestrian in my lane?

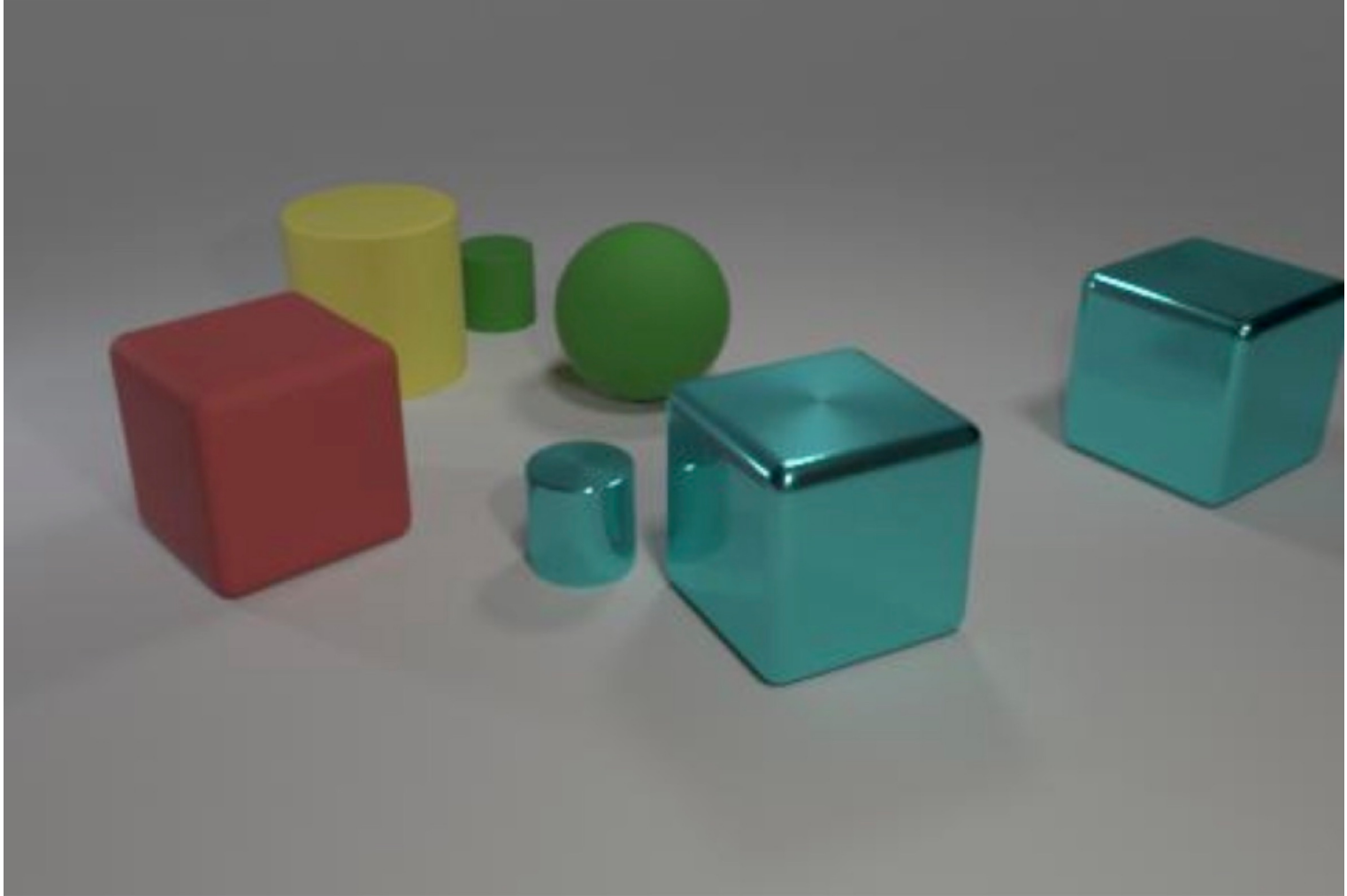


Is the person with the blue hat touching the bike in the back?



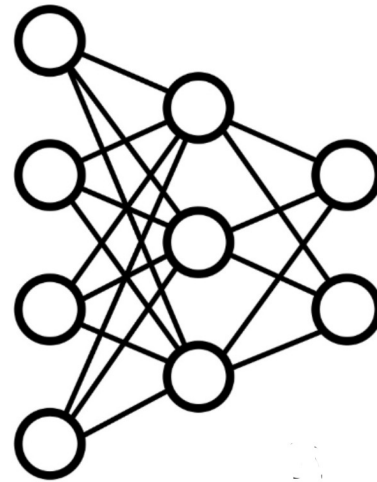
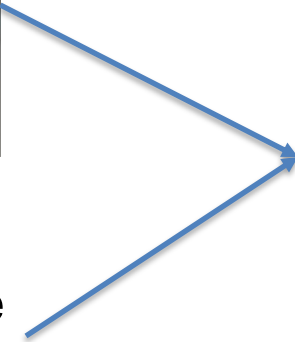
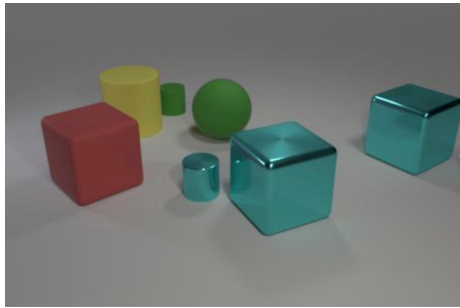
Is there a matte cube that has the same size as the red metal object?

Answering questions on an image



What color is the object with the same size as the blue metallic cylinder?

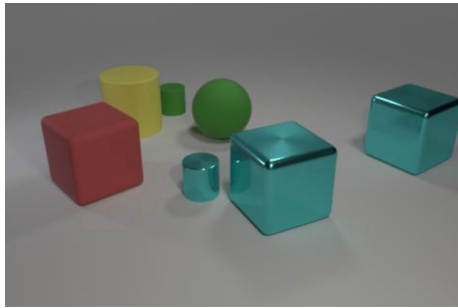
Answering questions on an image



Green

What color is the object with the same size as the blue metallic cylinder?

Sequential tasks

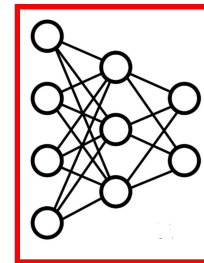
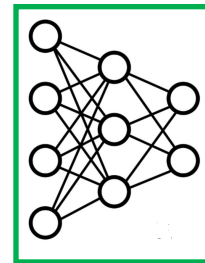
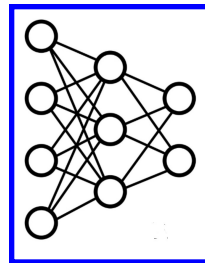


Find blue
metallic
cylinder

Compare
sizes

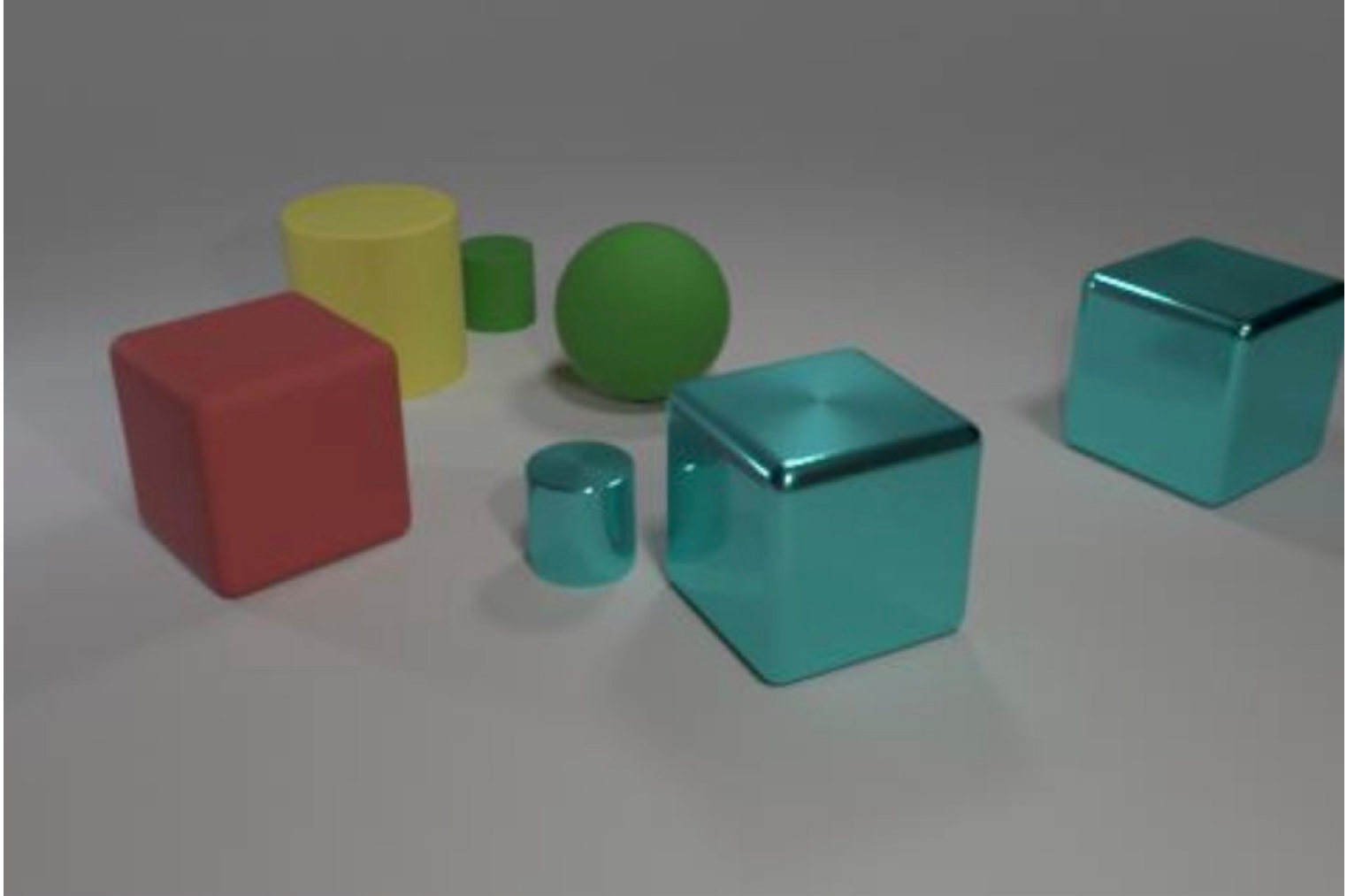
Describe
color

Green



What color is the
object with the same
size as the blue
metallic cylinder?

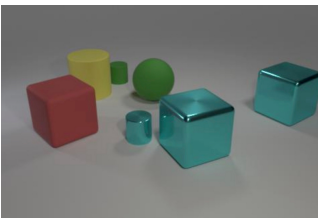
Reusable visual routines



How many objects are the same size as the ball?

Sequential tasks

What color is the object with the same size as the blue metallic cylinder?

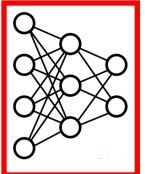
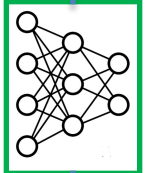
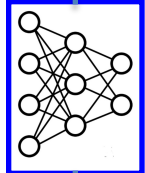


Find blue metallic cylinder

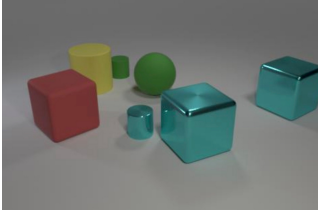
Compare sizes

Describe color

Green



How many objects are the same size as the ball?

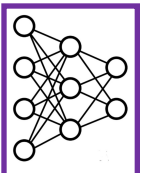


Find ball

Compare sizes

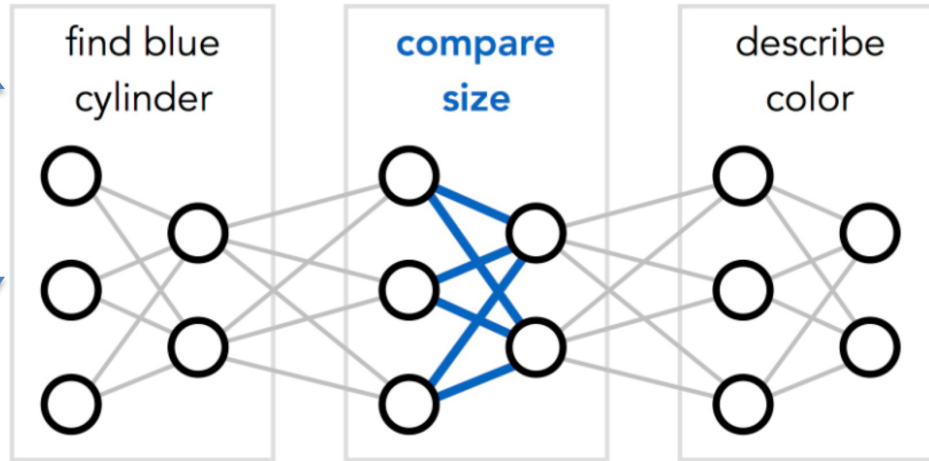
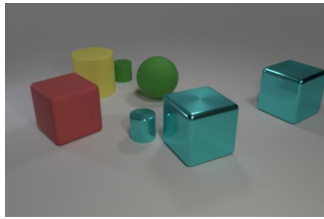
Count

Four



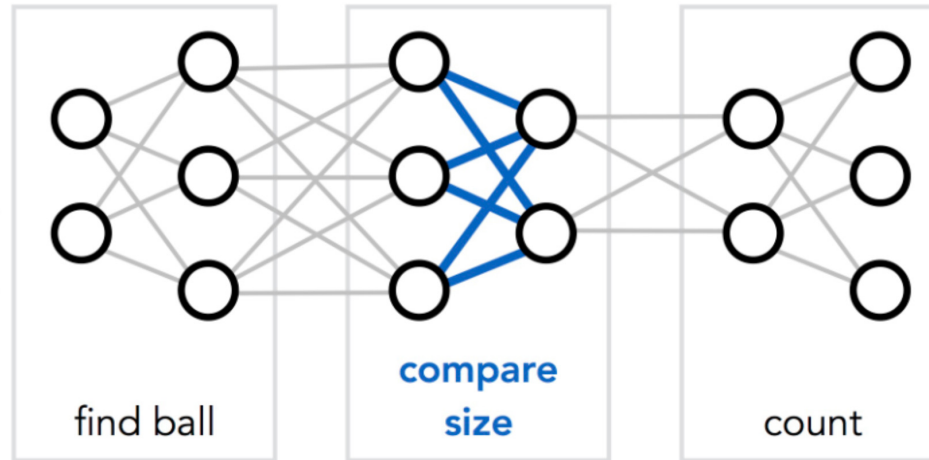
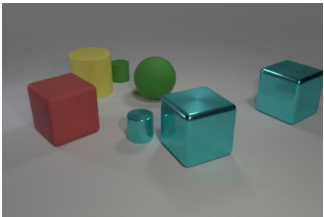
Sequential tasks

What color is the object with the same size as the blue metallic cylinder?

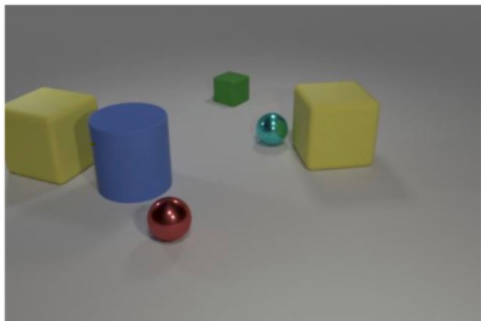
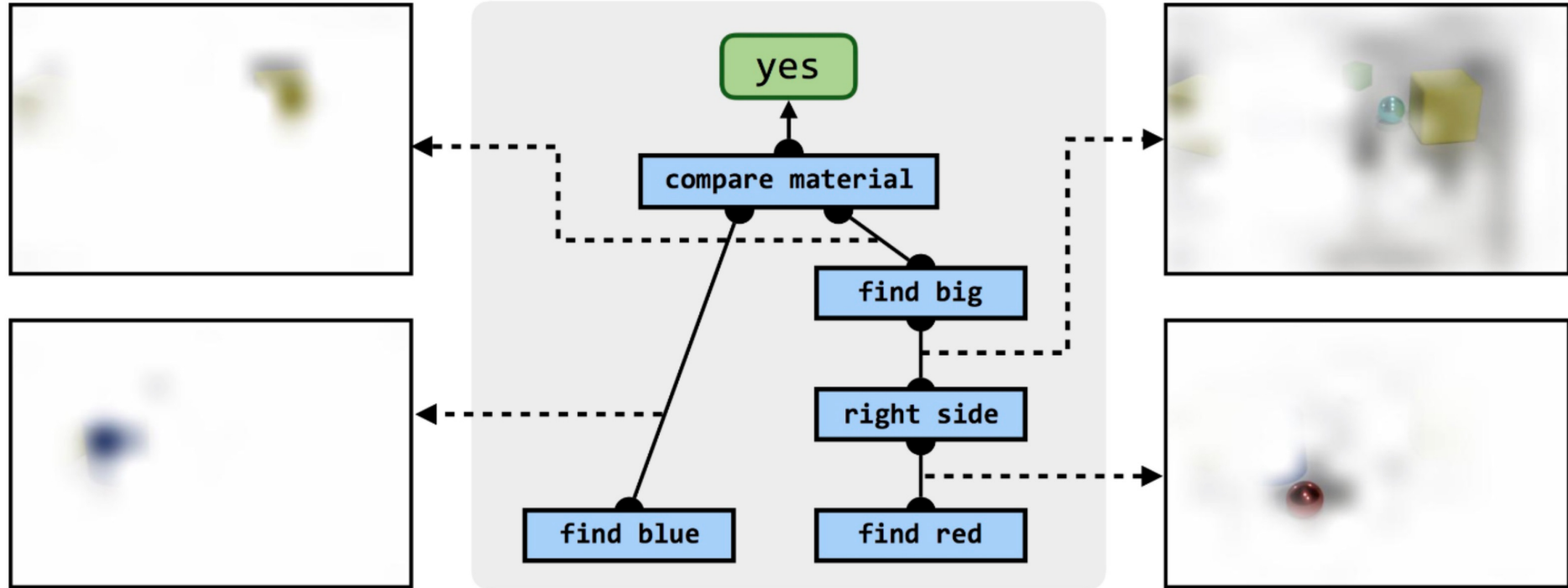


||

How many objects are the same size as the ball?



Compositional flexibility of visual routines



Does the blue cylinder have the same material as the big block on the right side of the red metallic thing?

Computer vision to help the blind

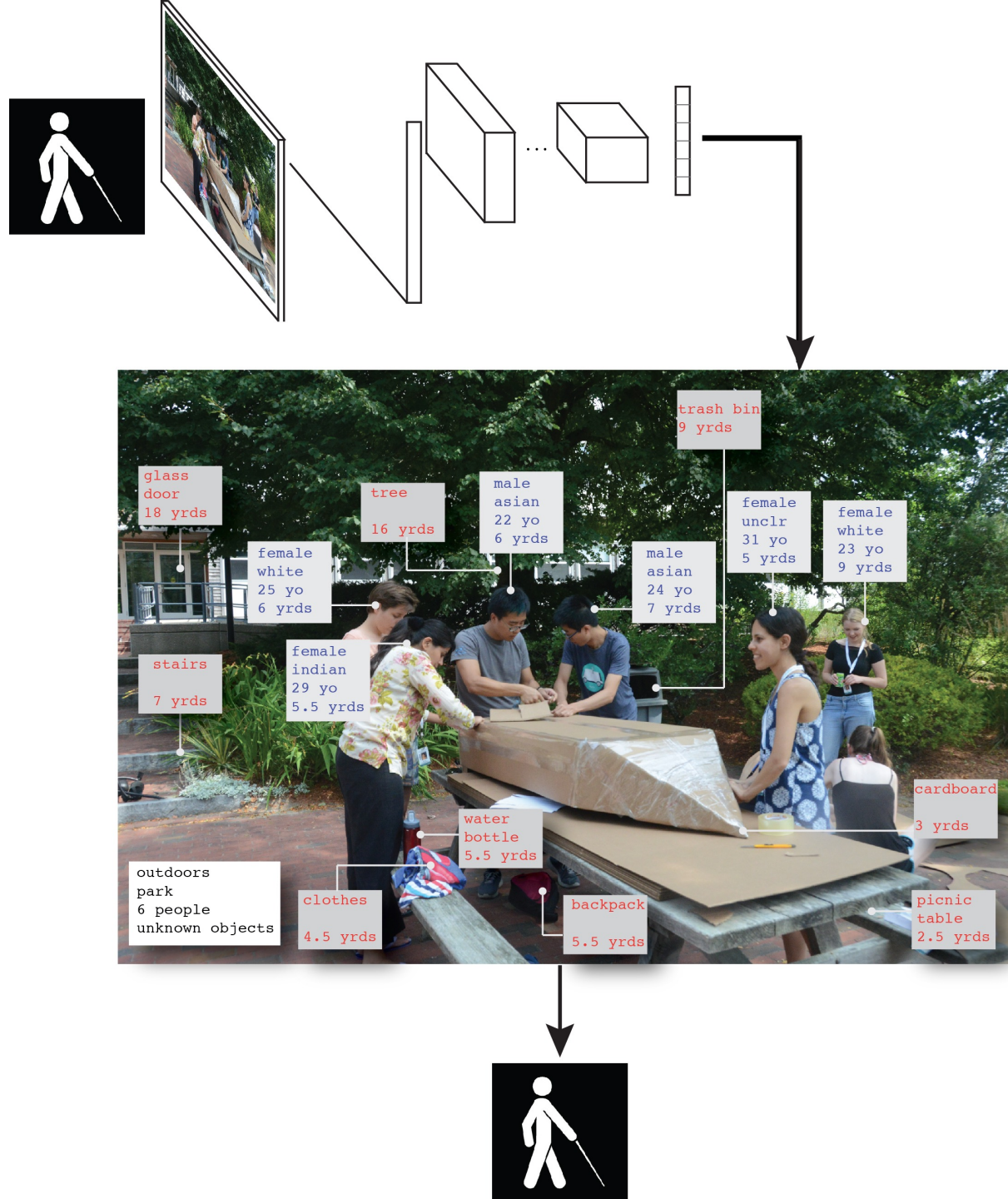


Image captioning

A

I think it's a person sitting at a table and she seems : |



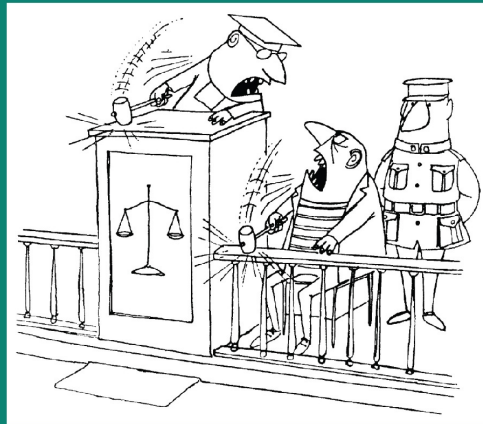
B

I think it's a group of people standing next to a body of water



C

I can't really describe line drawings : (



D

I think it's a group of people standing in front of a building and they seem :)



Summary

1. We have powerful image generators (→entertainment + science!)
2. Self-supervised predictive models mimic computational properties in neuroscience
3. We aspire to build a machine that can pass the Turing test for vision
4. State-of-the-art computer vision still has a long way to go (e.g., adversarial images, simple tasks that machines cannot solve, image captioning, etc.)
5. Compositional visual routines can connect perception to cognition
6. A computer vision system can help patients with severe visual deficits

References

Serre, T. Deep learning: the good, the bad and the ugly. *Annual Review of Vision* **5**, 399-426 (2019).

Kreiman, G. & Serre, T. Beyond the feedforward sweep: feedback computations in the visual cortex. *This Year in Cognitive Neuroscience* (2020).

Ponce, C. R. *et al.* Evolving images for visual neurons using a deep generative network reveals coding principles and neuronal preferences. *Cell* (2019).

Visual Object Recognition

Computational Models and Neurophysiological Mechanisms

Neuro 130/230. Harvard College/GSAS 78454

