



Object Recognition

in images



Image: two-dimensional matrix of pixels
Contains information, objects represented in it



Image: two-dimensional matrix of pixels
Contains information, objects represented in it

Extract this information?
Automate the process?



Machine Learning

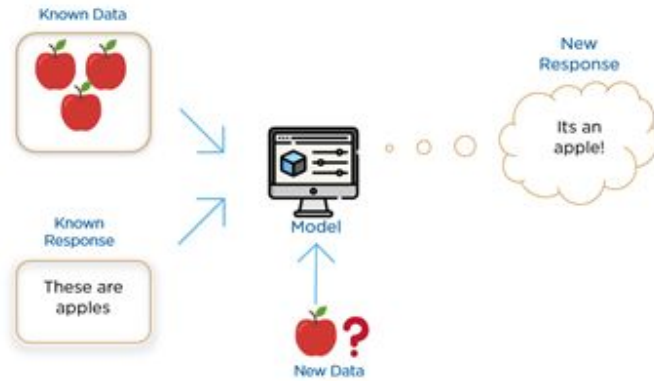
Machine Learning

In particular, Neural Networks

Neural network?

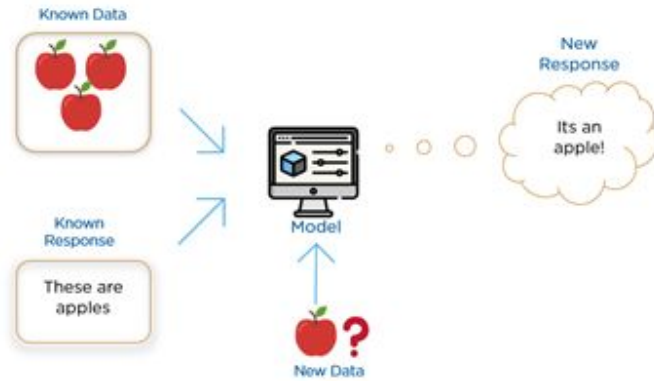


Neural network



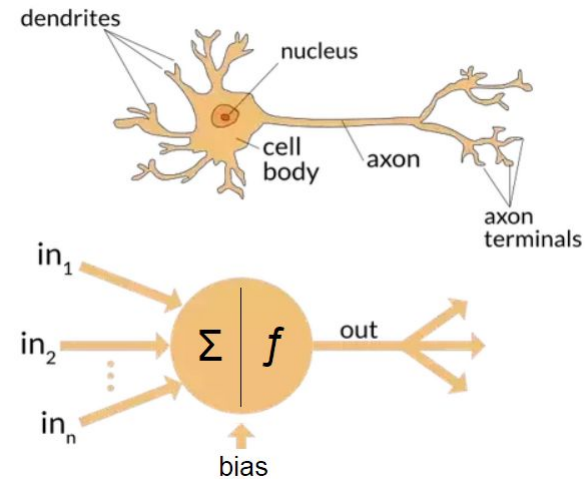
Supervised learning

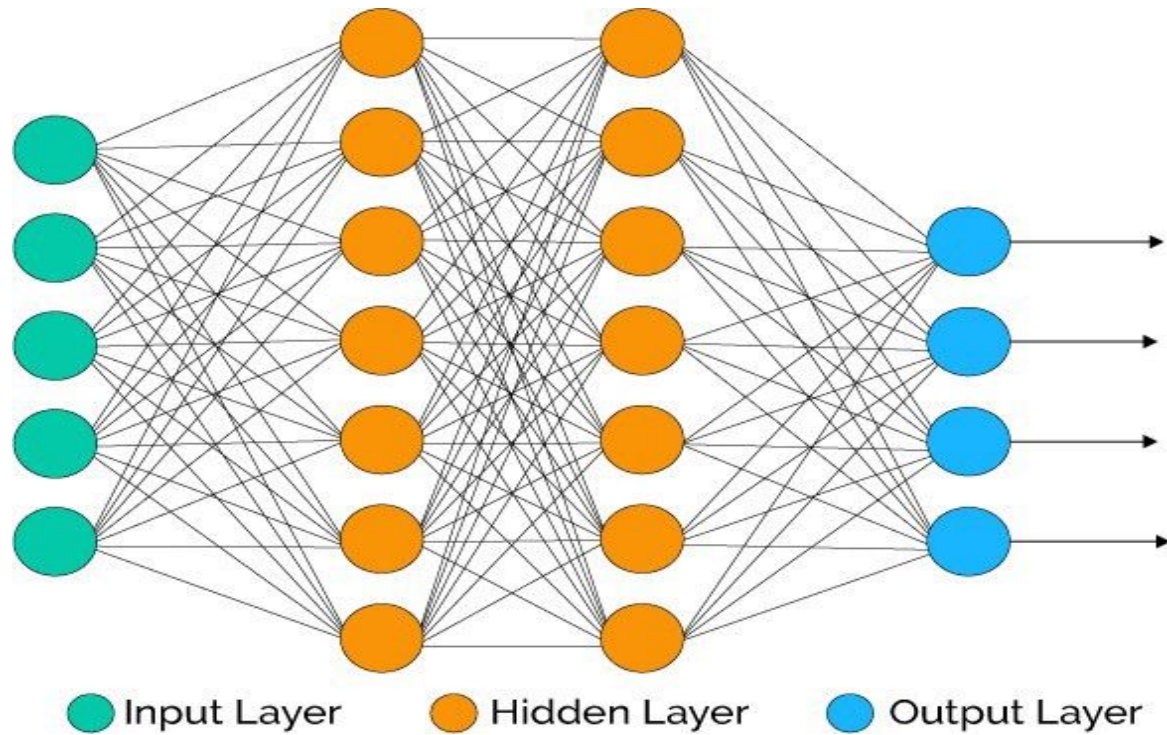
Neural network



Mimics the brain

Supervised learning





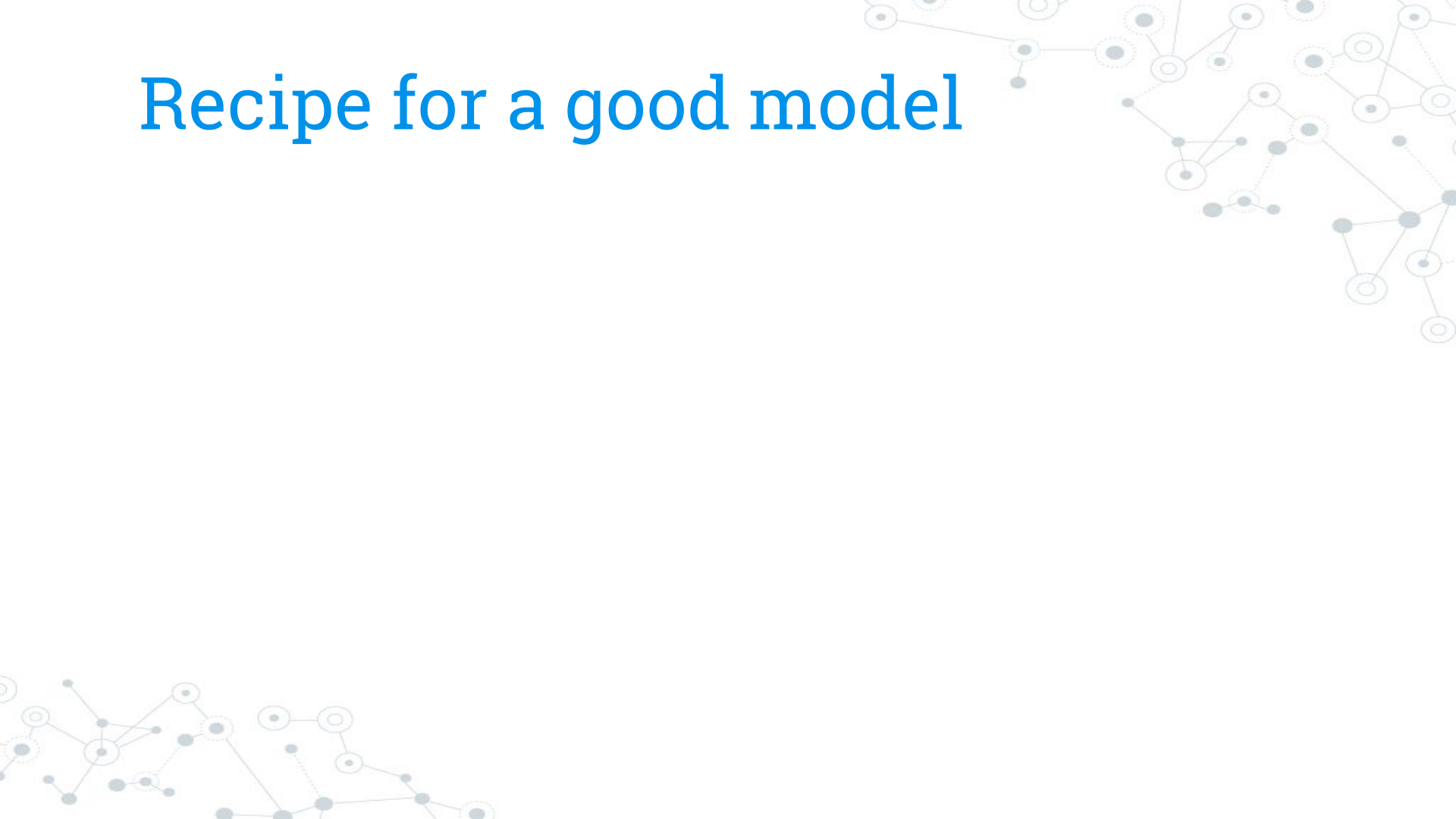
⇒ Interconnected layers of neurons

To learn more:

- <https://www.forbes.com/sites/bernardmarr/2018/09/24/what-are-artificial-neural-networks-a-simple-explanation-for-absolutely-anyone/#3c290e7a1245>
- <https://www.techradar.com/news/what-is-a-neural-network>
- <https://www.datascience.com/blog/convolutional-neural-network>



Recipe for a good model



Recipe for a good model

- A lot of training material (>1000 per class)

Recipe for a good model

- A lot of training material (>1000 per class)
- A good neural network

Recipe for a good model

- A lot of training material (>1000 per class)
- A good neural network
 - Many parameters

Recipe for a good model

- A lot of training material (>1000 per class)
- A good neural network
 - Many parameters
 - lots of tweaking

Recipe for a good model

- A lot of training material (>1000 per class)
- Computational power
- A good neural network
 - Many parameters
 - lots of tweaking

Recipe for a good model

- A lot of training material (>1000 per class)
- A good neural network
 - Many parameters
 - lots of tweaking
- Computational power
 - GPUs

Recipe for a good model

- A lot of training material (>1000 per class)
- A good neural network
 - Many parameters
 - lots of tweaking
- Computational power
 - GPUs
- A lot of time

Recipe for a good model

- A lot of training material (>1000 per class)
- A good neural network
 - Many parameters
 - lots of tweaking
- Computational power
 - GPUs
- A lot of time
 - Retraining if the model isn't optimal

A decorative network diagram in the top-left corner, featuring a complex web of interconnected nodes and lines. The nodes are represented by circles of varying sizes, some with concentric rings, and are connected by thin, light gray lines. The overall structure is dense and organic, resembling a molecular or biological network.

How do we do that?

A decorative network diagram in the bottom-right corner, similar to the one in the top-left. It consists of a cluster of interconnected nodes and lines, with nodes represented by circles of different sizes and some having concentric rings. The connections are thin, light gray lines, creating a complex, web-like structure.

A decorative network diagram in the top-left corner, consisting of various sized circles (nodes) connected by thin lines (edges). Some nodes are solid grey, while others are hollow with a grey outline. The connections form a complex, branching structure.

How do we do that?

We don't

A decorative network diagram in the bottom-right corner, similar to the one in the top-left, featuring a cluster of interconnected nodes and edges.

Lots of training datasets online



Lots of training datasets online

Pretrained models available for them





Image classification

Image classification



persian_cat



bat

Image classification



persian_cat



bat

Image classification



persian_cat

toothbrush







14m+ labeled images





14m+ labeled images
20k+ classes



ILSVRC



ILSVRC

- ImageNet Large Scale Visual Recognition Competition

ILSVRC

- ImageNet Large Scale Visual Recognition Competition
- Yearly competition

ILSVRC

- ImageNet Large Scale Visual Recognition Competition
- Yearly competition
- 2015 winner: ResNet50 -- available online



Label:	Score:
platypus	0.31578568
puffer	0.25567934
scuba_diver	0.1726397
tree_frog	0.049659904
electric_ray	0.036546305





Label:	Score:
platypus	0.31578568
puffer	0.25567934
scuba_diver	0.1726397
tree_frog	0.049659904
electric_ray	0.036546305



Label:	Score:
swimming_trunks	0.855407178401947
maillot	0.03095098026096821
miniskirt	0.008445939049124718
bathing_cap	0.0081565510481596
maillot (again)	0.007052644621580839

But...



But...

cat

???



But...

cat

???



cat

???



towel/
doormat





Huge dataset, many classes



Huge dataset, many classes

BUT





Huge dataset, many classes

BUT

Made for single-label classification



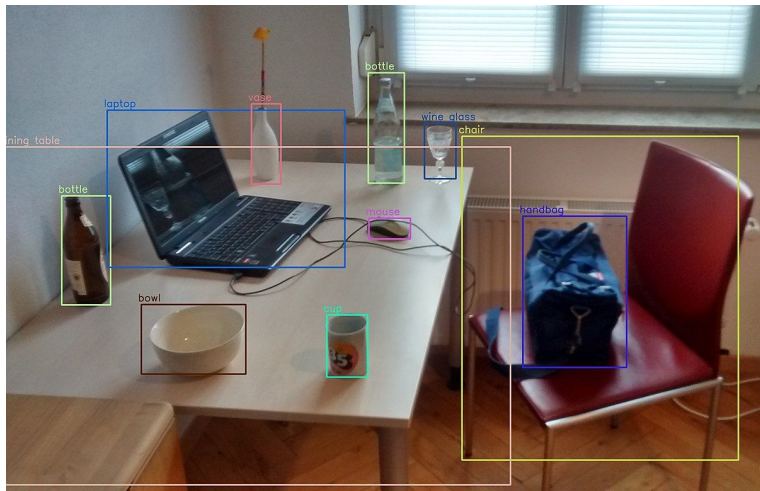
A decorative background featuring a network diagram. It consists of numerous nodes, represented by small circles, some of which are solid blue, some are grey with a blue outline, and others are grey with a blue outline and a smaller blue dot inside. These nodes are interconnected by thin, light grey lines, forming a complex web-like structure that is more dense on the left and right sides of the slide.

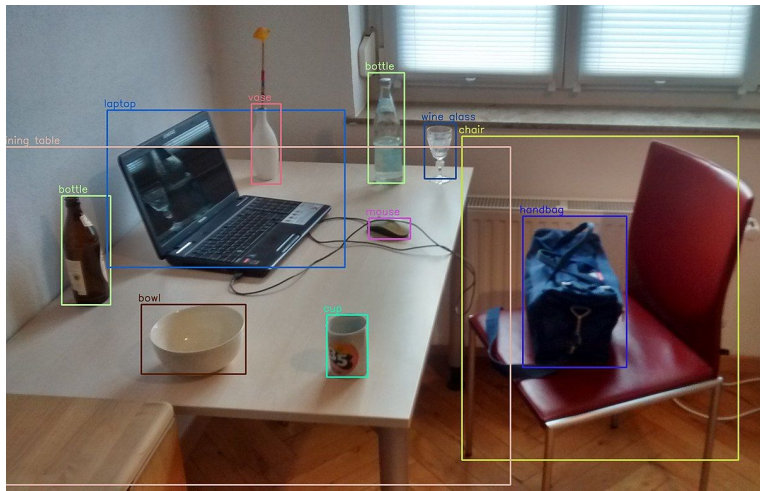
Object detection



Object detection

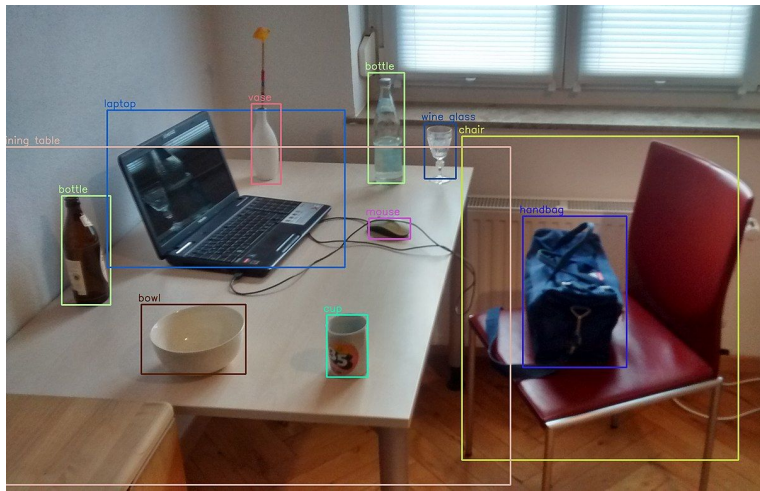
Image segmentation





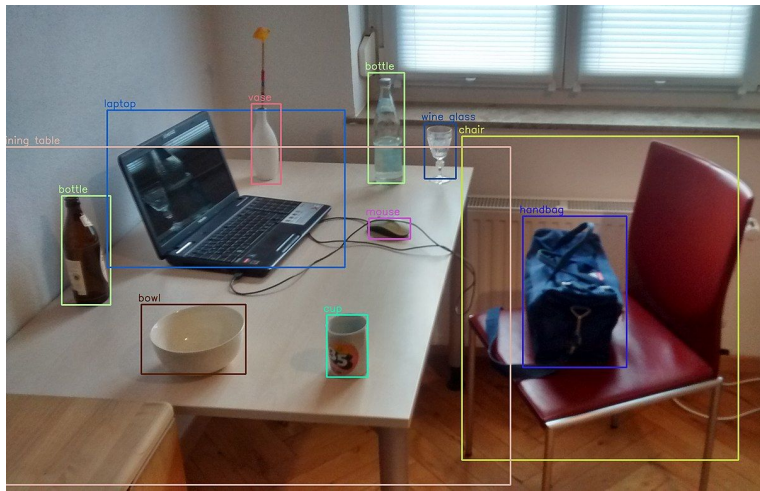
VS





VS

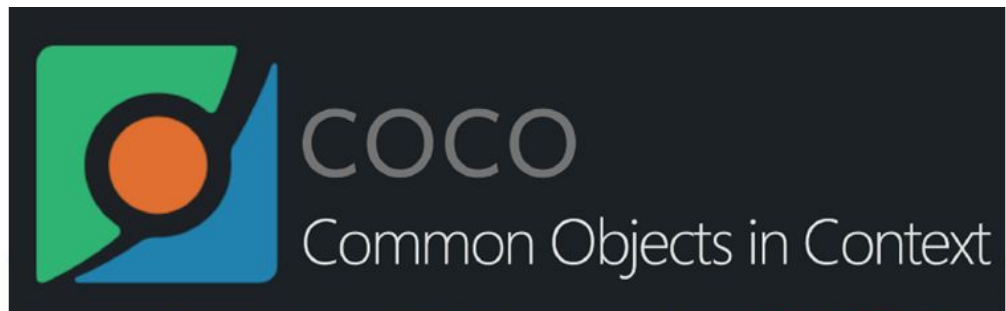


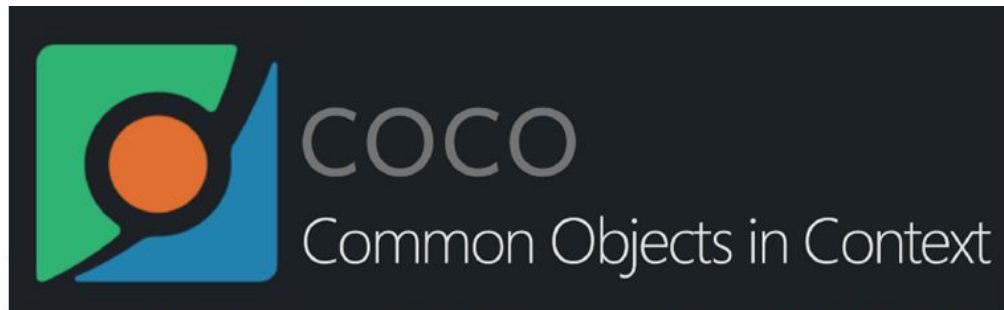


VS

⇒ multi-label
classification







200k+ labeled images

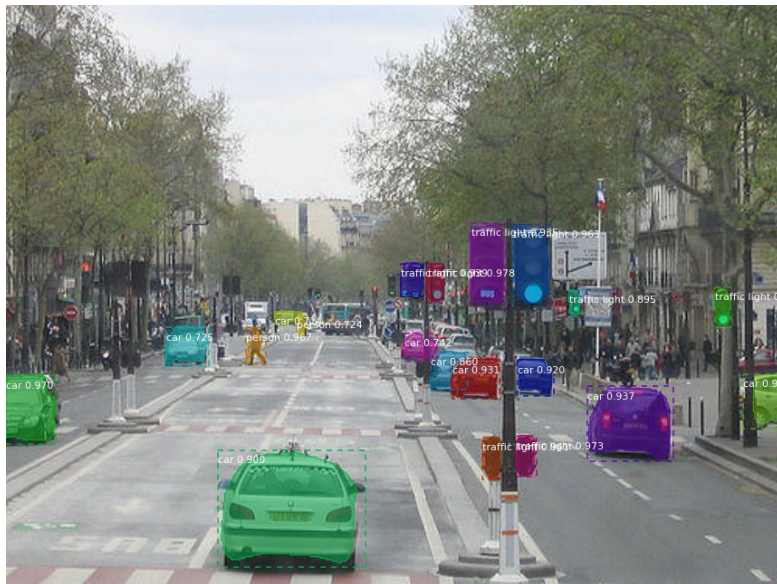


Mask-RCNN pretrained model

https://github.com/matterport/Mask_RCNN

Mask-RCNN pretrained model

https://github.com/matterport/Mask_RCNN



Good results...



Good results...



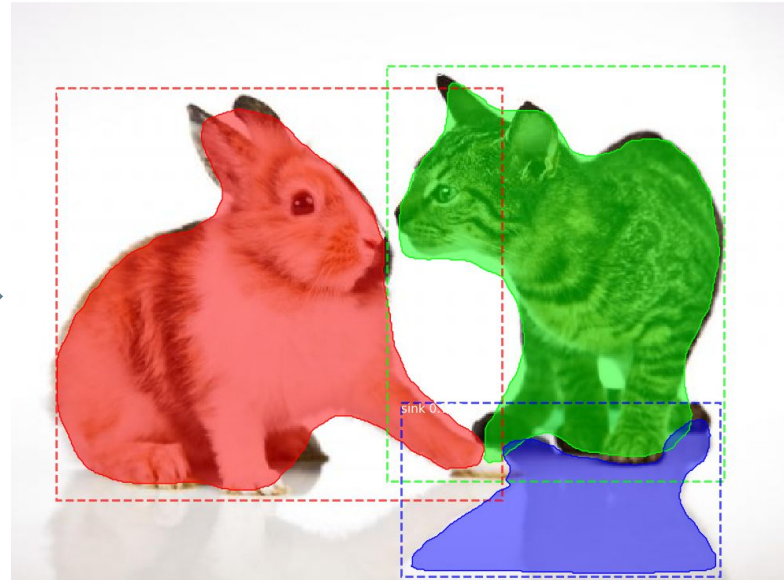
Good results...



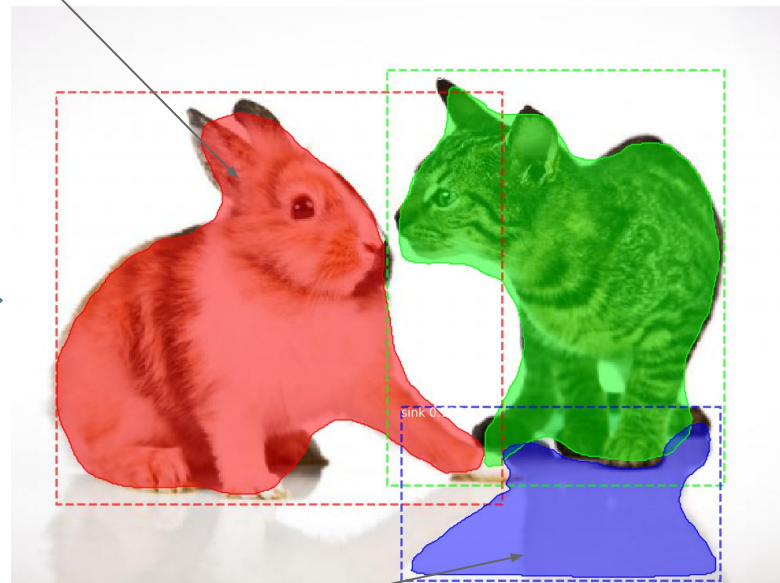
... mostly



... mostly



... mostly



sink (false positive)

Multi-label classification



Multi-label classification

BUT



Multi-label classification

BUT

Harder to create a good dataset



A decorative background featuring a network diagram. It consists of numerous nodes, represented by small circles, some of which are solid blue, some are grey with a blue outline, and others are grey with a blue outline and a smaller blue dot inside. These nodes are interconnected by thin, light grey lines, forming a complex web-like structure that is more dense in the corners and sparser in the center.

What's left to do

What's left to do



- Look into other datasets



- Look into other datasets
- Find additional models



- Look into other datasets
- Find additional models
- Investigate model retraining



A decorative network diagram in the top-left corner, featuring a complex web of interconnected nodes and lines. The nodes are represented by circles of varying sizes, some with concentric rings, and are connected by thin, light gray lines. The overall structure is organic and sprawling, extending from the top-left towards the center of the slide.

A promising axis of research

A decorative network diagram in the bottom-right corner, similar in style to the one in the top-left. It consists of a cluster of interconnected nodes and lines, with nodes represented by circles of different sizes and some having concentric rings. The lines are thin and light gray, creating a web-like pattern that extends from the bottom-right towards the center.

A decorative network diagram in the top-left corner, consisting of a complex web of interconnected nodes and lines. The nodes are represented by small circles, some of which are larger and have concentric rings, while others are smaller and solid. The lines are thin and grey, connecting the nodes in a non-linear fashion. The overall style is minimalist and technical.

A promising axis of research

Google's Open Images

A decorative network diagram in the bottom-right corner, similar to the one in the top-left. It features a cluster of nodes connected by lines. The nodes vary in size and some have concentric circles, suggesting a hierarchical or complex network structure. The lines are thin and grey, blending into the white background.

The background of the slide is a collage of various images, including a person, a bicycle, and a person's arm, each with colored bounding boxes (yellow, red, blue) indicating object detection. The text is overlaid on this collage.

Open Images Dataset V5 + Extensions

15,851,536 boxes on 600 categories

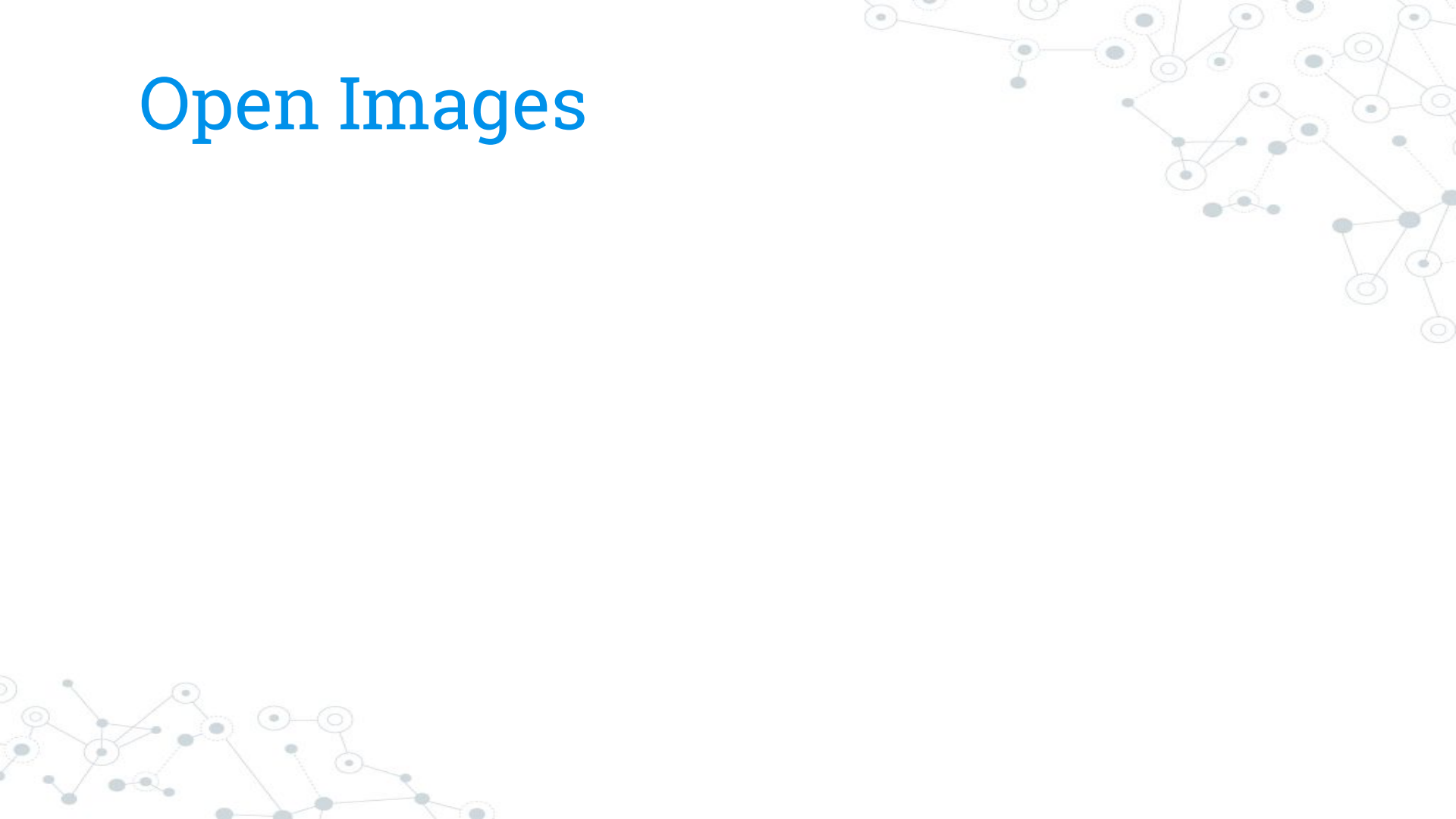
2,785,498 instance segmentations on 350 categories

36,464,560 image-level labels on 19,959 categories

391,073 relationship annotations of 329 relationships

Extension - 478,000 crowdsourced images with 6,000+ categories

Open Images



Open Images

- A recent dataset (V1 in 2016!)

Open Images

- A recent dataset (V1 in 2016!)
- Models harder to find

Open Images

- A recent dataset (V1 in 2016!)
- Models harder to find
- Looks like our best option so far

A decorative network diagram at the top of the slide, featuring a complex web of interconnected nodes and lines. The nodes are represented by small circles, some of which are highlighted with a dashed border. The lines are thin and gray, creating a mesh-like structure. A central node is highlighted with a solid blue border and contains a large blue quotation mark.

“

Thank you for your attention.



That's all Folks!

Th...tion.