

LEMMIE DO SOME

OBJECT DETECTION



AMAN AGARWAL

WHAT IS OBJECT DETECTION?

Detecting/localizing instances of semantic objects of a certain category in digital images or videos.

Classification



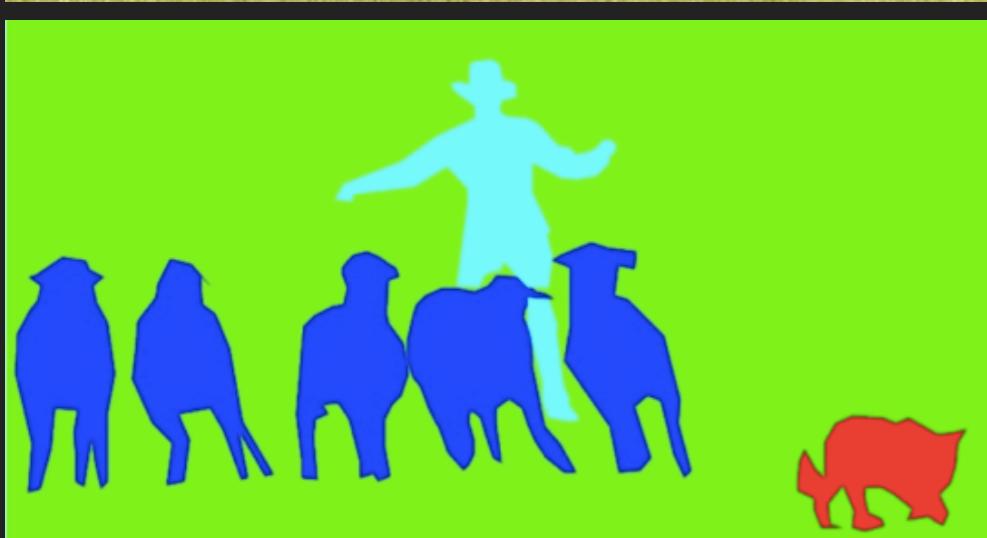
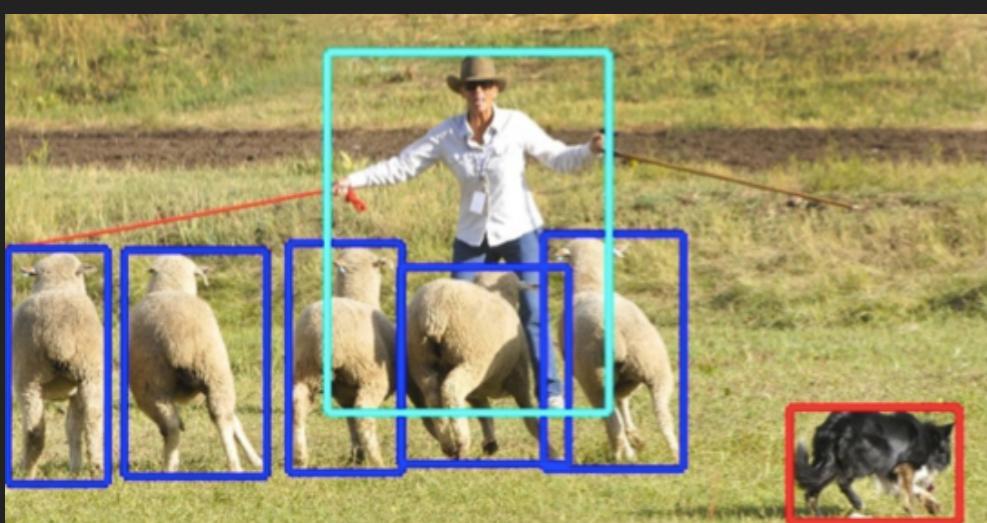
Car?

Localization

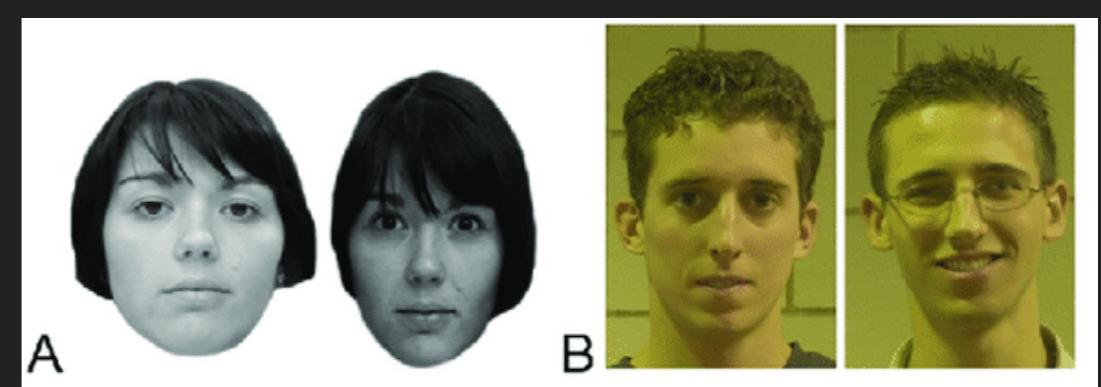
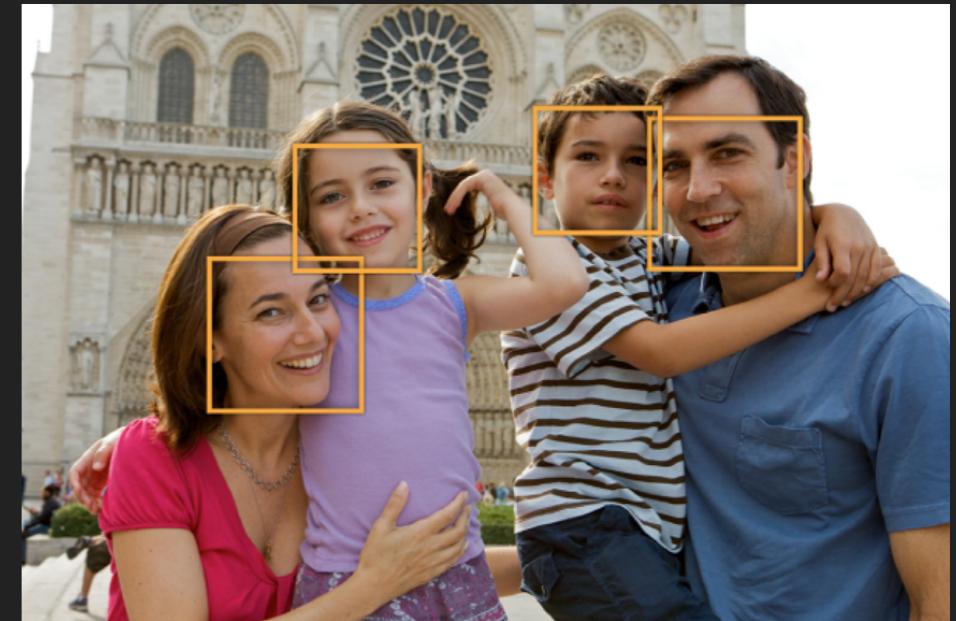


Where is the Car?

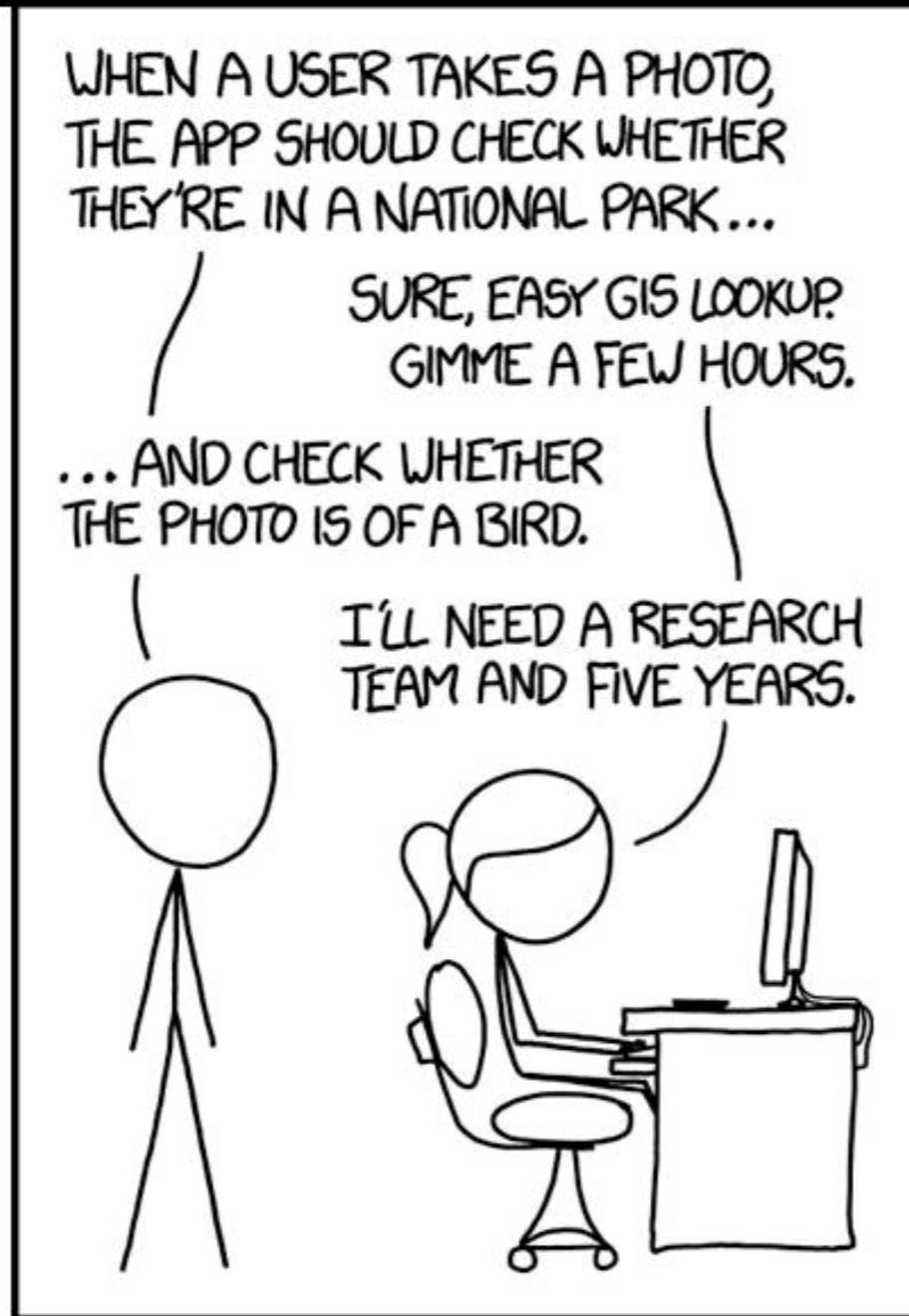
B'COZ, SOMETIMES YOU DON'T GET IT



Classification



2012:



IN CS, IT CAN BE HARD TO EXPLAIN
THE DIFFERENCE BETWEEN THE EASY
AND THE VIRTUALLY IMPOSSIBLE.

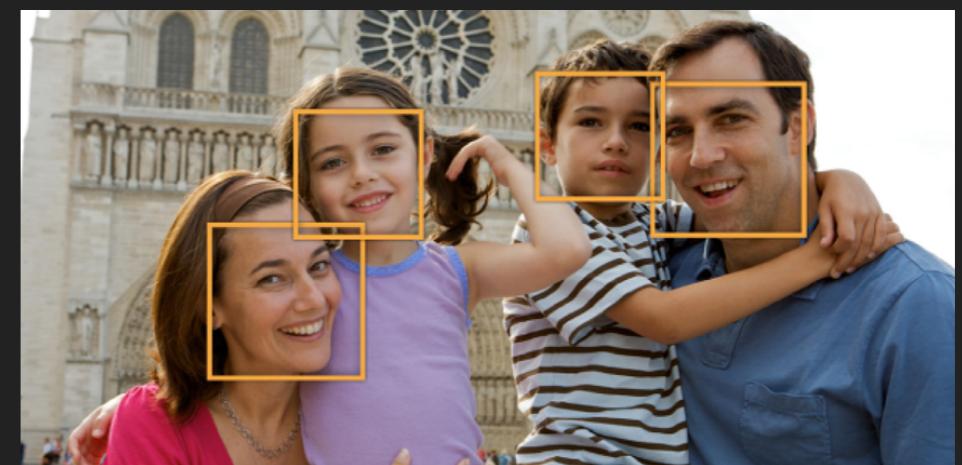
2019:

```
from tensorflow.keras.applications  
import PretrainedBirdDetector
```

OBJECT DETECTION

APPLICATION

- ▶ Self-driving car
- ▶ Google search
- ▶ Camera
- ▶ Conveyer belt
- ▶ Object tracking/counting
- ▶ OCR



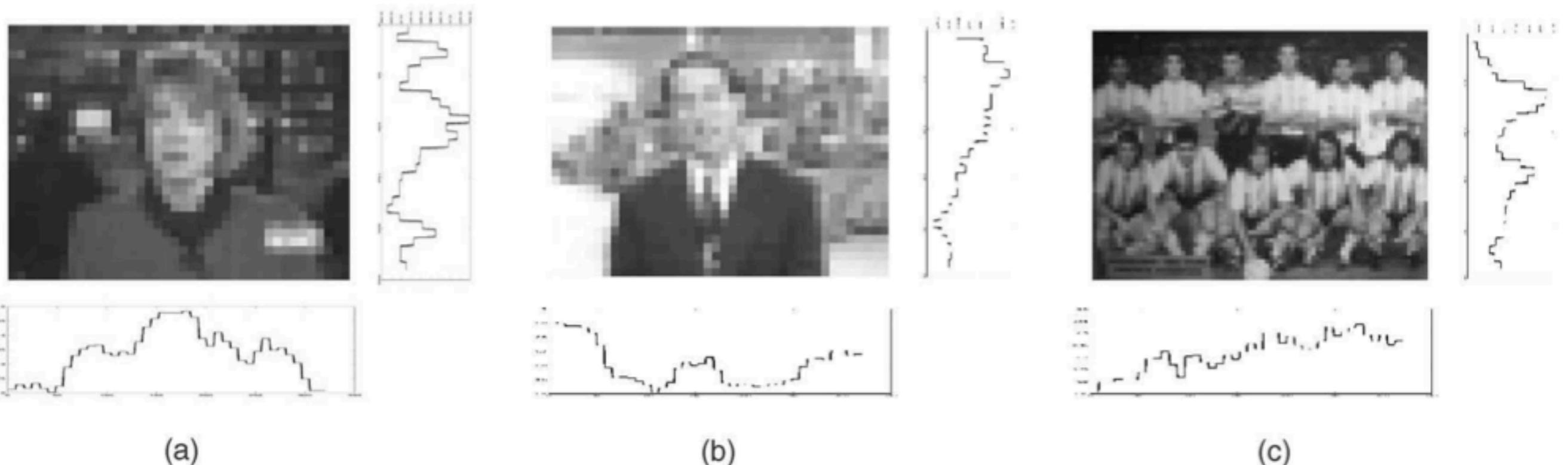
**A LONG
TIME AGO
IN A GALAXY
FAR, FAR AWAY...**

OBJECT DETECTION EARLIER

YANG ET AL. [1]



KOTROPOULOS ET AL. [2]

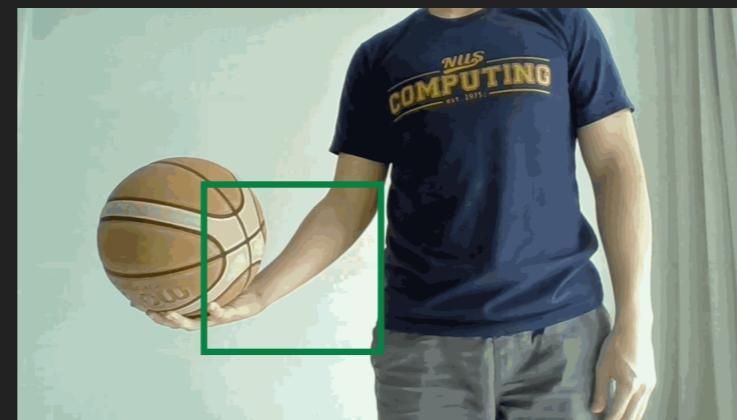


LIMITATIONS

- ▶ Works only on front profile.
- ▶ Requires hand-crafted features.
- ▶ Only works for one specific purpose.

NAÏVE APPROACH

USING SLIDING WINDOW (CLASSIFICATION)



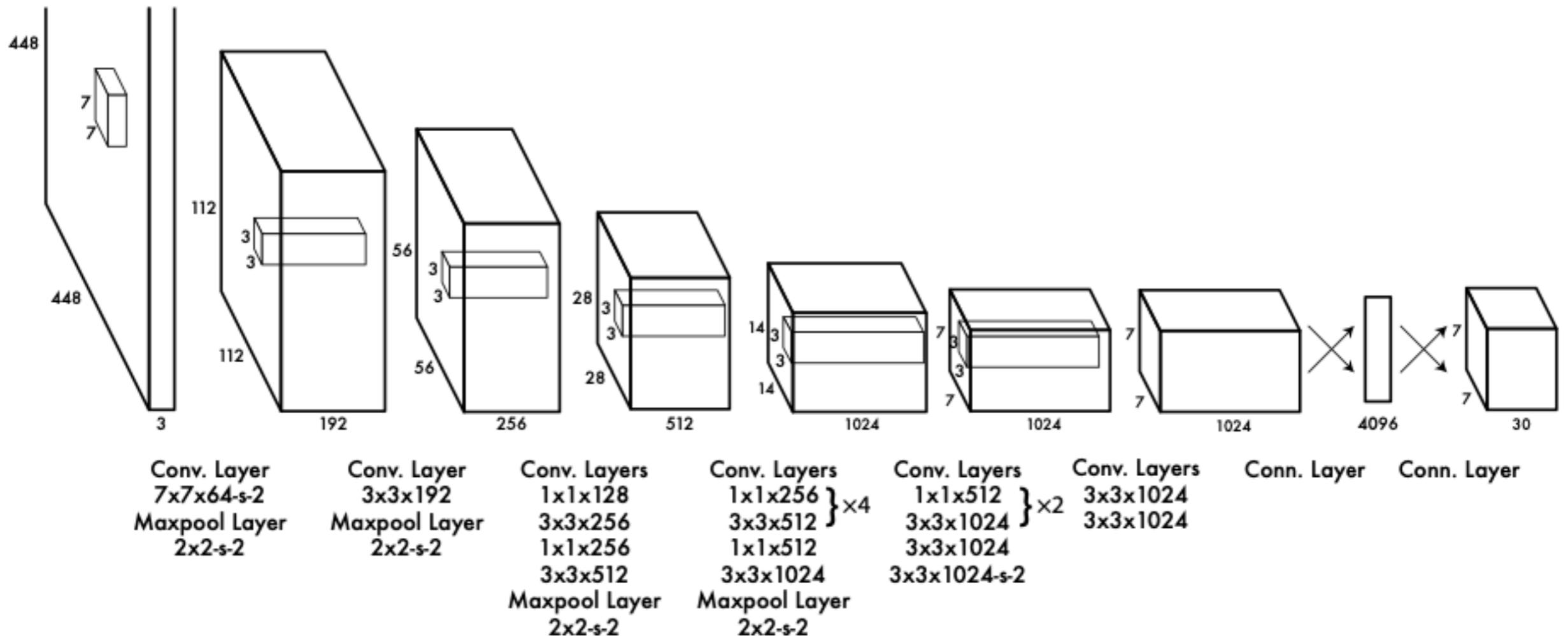
LIMITATIONS

- ▶ Time consuming.
- ▶ Inaccurate boundaries.
- ▶ Needs a well-trained classification algorithm.
- ▶ Unnecessary computation on background regions.



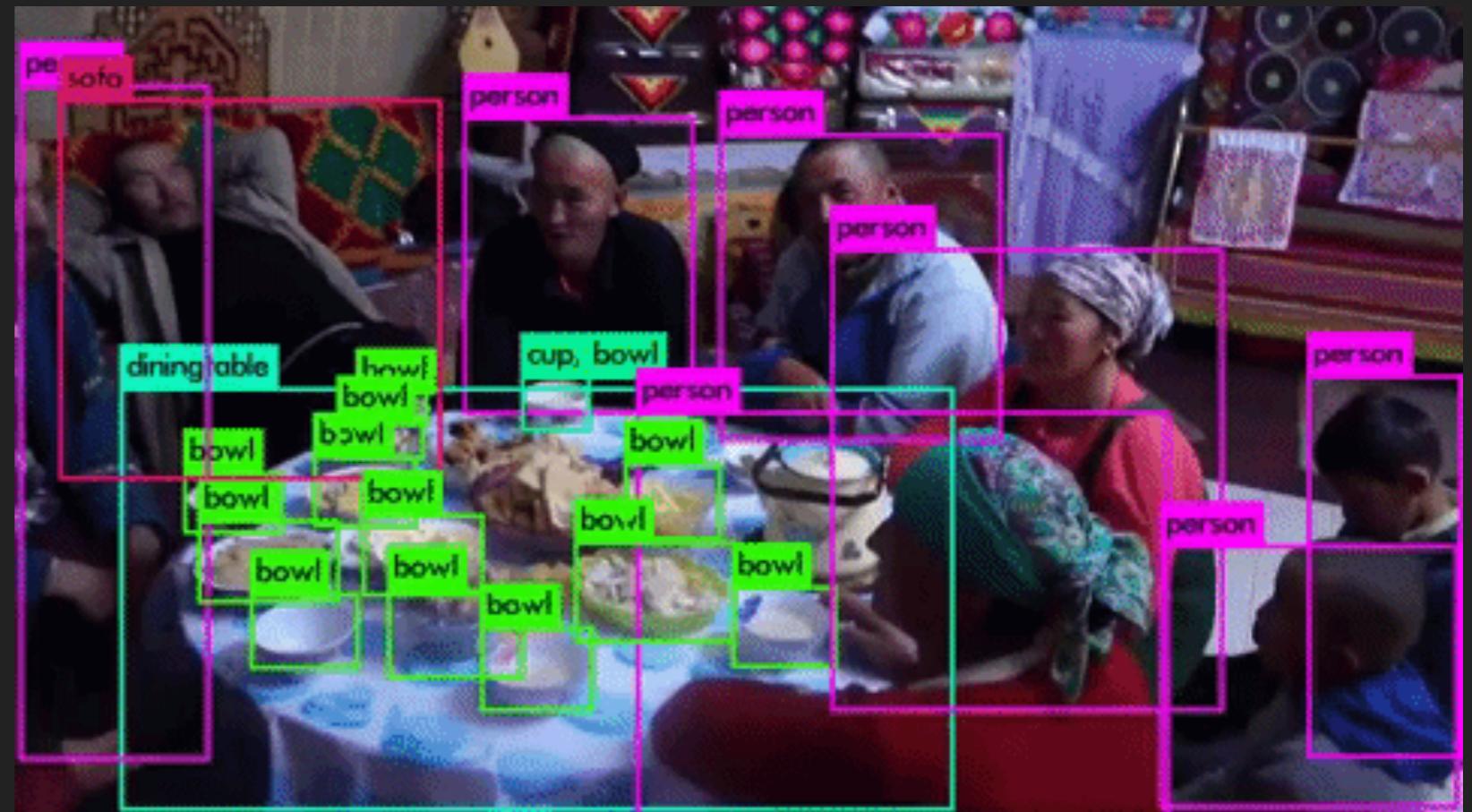
ADVANCED APPROACH

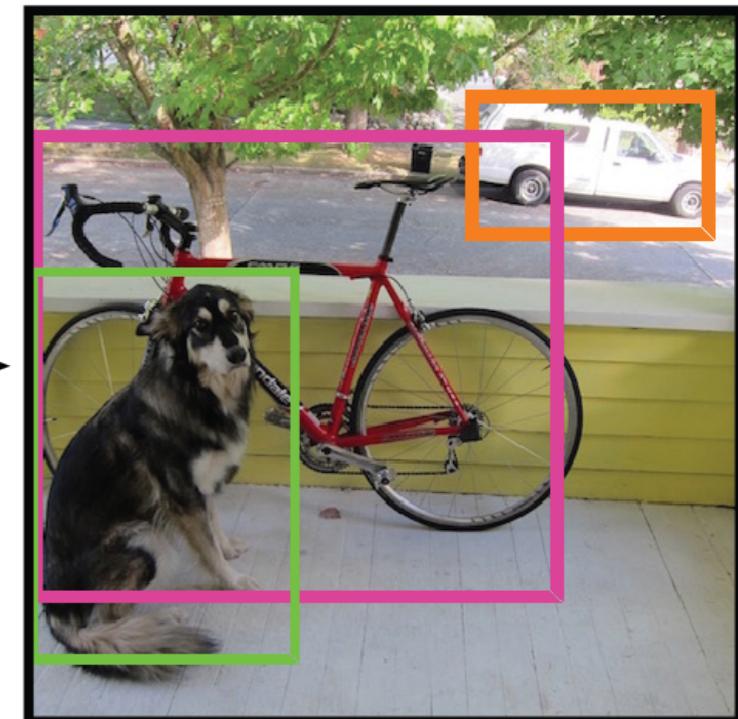
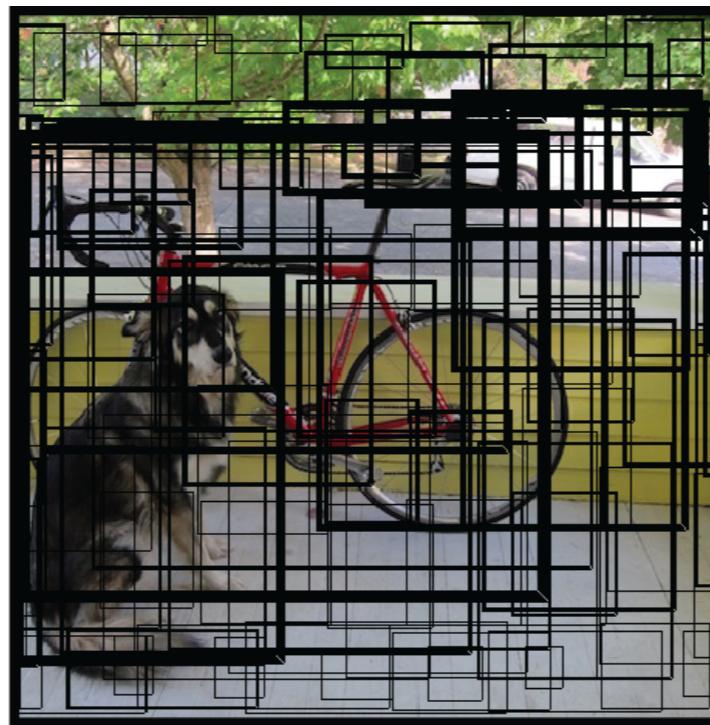
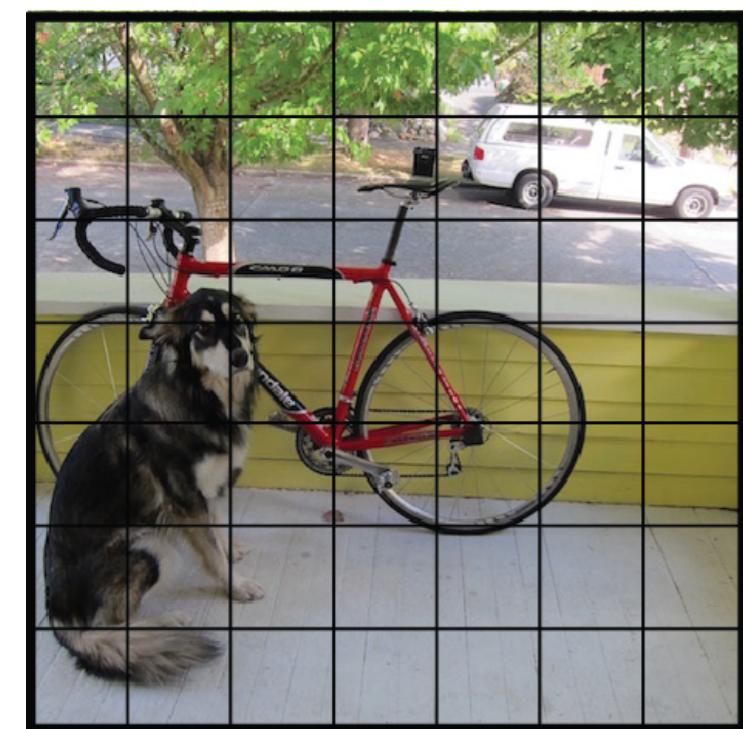
You Only Look Once



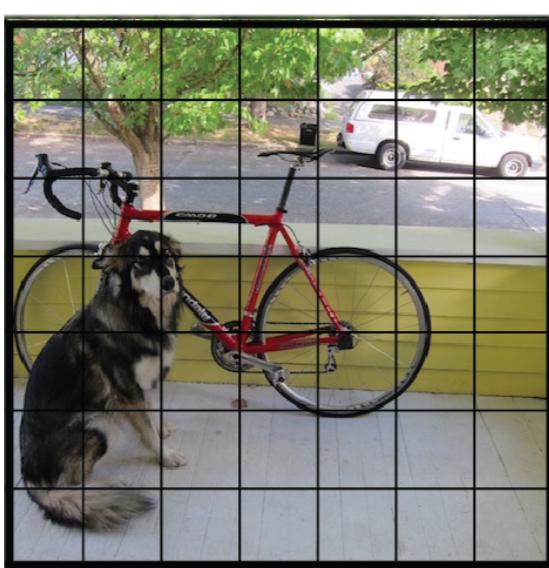
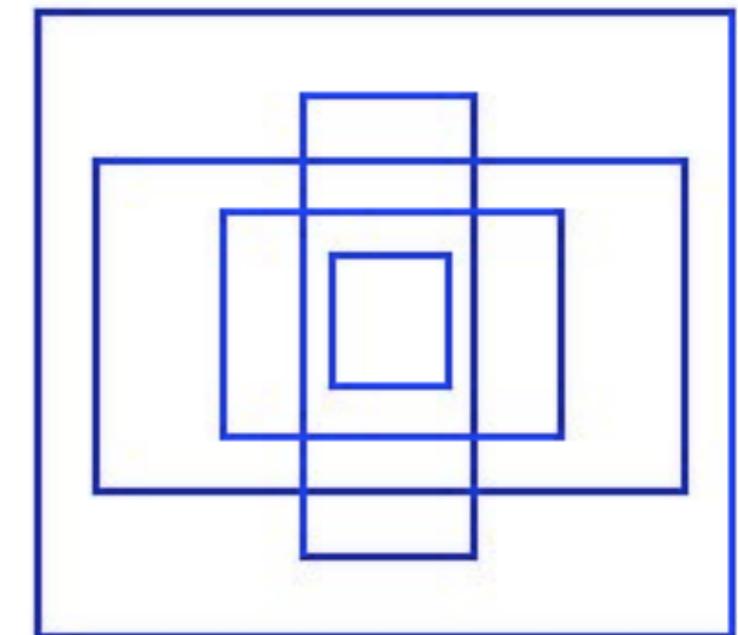
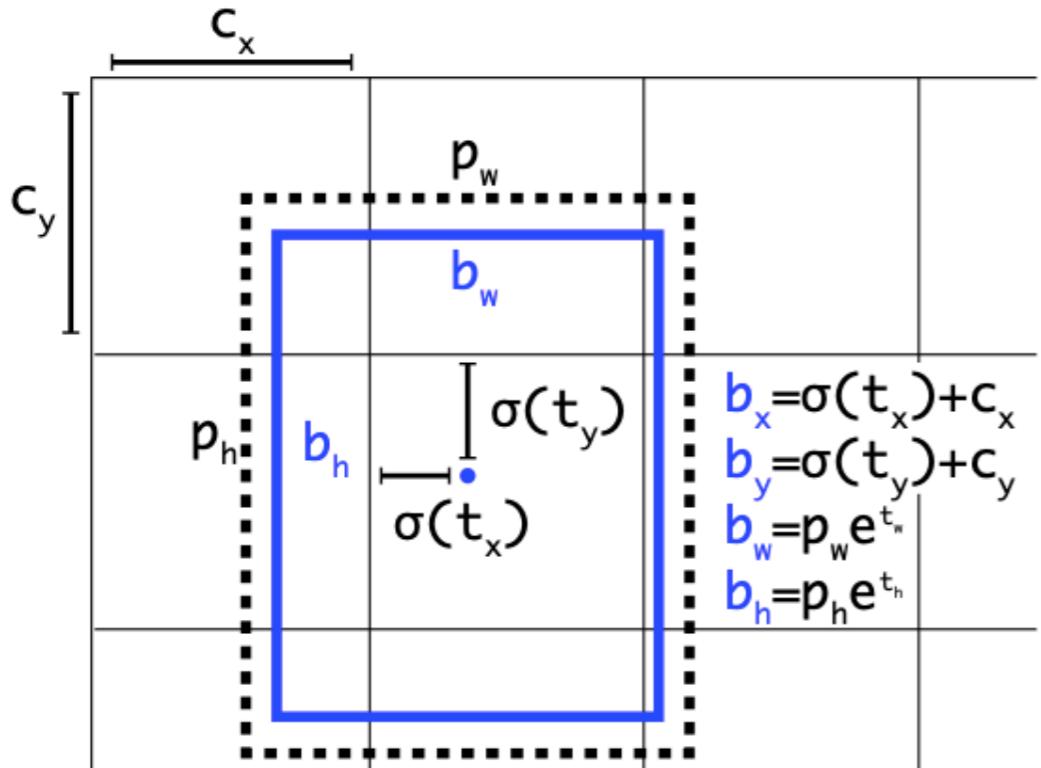
YOLO [3]

- ▶ Speed: 45-150 fps.
- ▶ Trained to work on 80 classes, but can even detect 9000.
- ▶ Can even work on smartphones.
- ▶ Very accurate.
- ▶ Sometimes struggles with small objects though.





Aman Agarwal

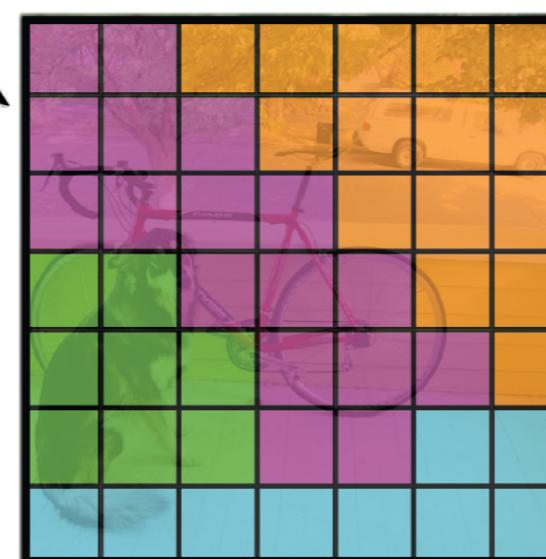


$S \times S$ grid on input

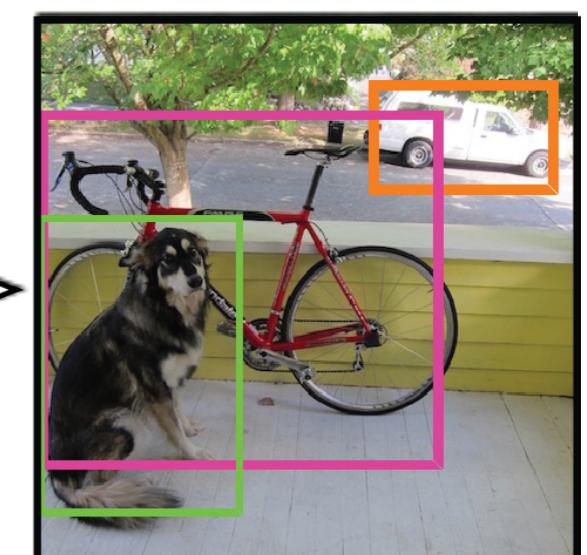
$y =$	pc
	bx
	by
	bh
	bw
	c1
	c2
	c3



Bounding boxes + confidence



Class probability map

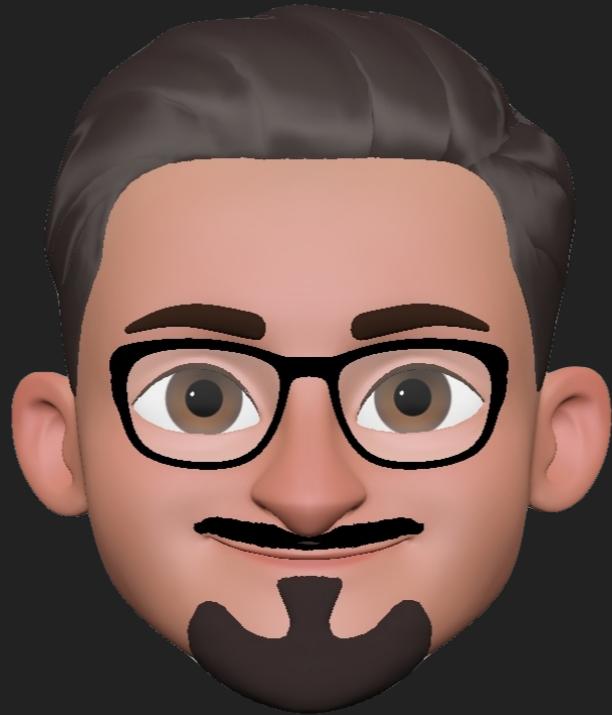


Final detections

HANDS-ON

<https://app.nanonets.com/>

<https://nanonets.com/blog/object-detection-tensorflow-js/>



AMAN AGARWAL

Software Developer @ HSBC | Intel Software Innovator of AI

+91 75740 14476 | amanag.11@gmail.com | amanbasu.github.io/portfolio/



theabecedarian



amanbasu



@TheAbecedarian_

REFERENCES

1. G. Yang and T. S. Huang, "Human Face Detection in Complex Background," *Pattern Recognition*, vol. 27, no. 1, pp. 53-63, 1994.
2. C. Kotropoulos and I. Pitas, "Rule-Based Face Detection in Frontal Views," *Proc. Int'l Conf. Acoustics, Speech and Signal Processing*, vol. 4, pp. 2537-2540, 1997.
3. J. Redmon and A. Farhadi. Yolov3: An incremental improvement. *arXiv*, 2018.