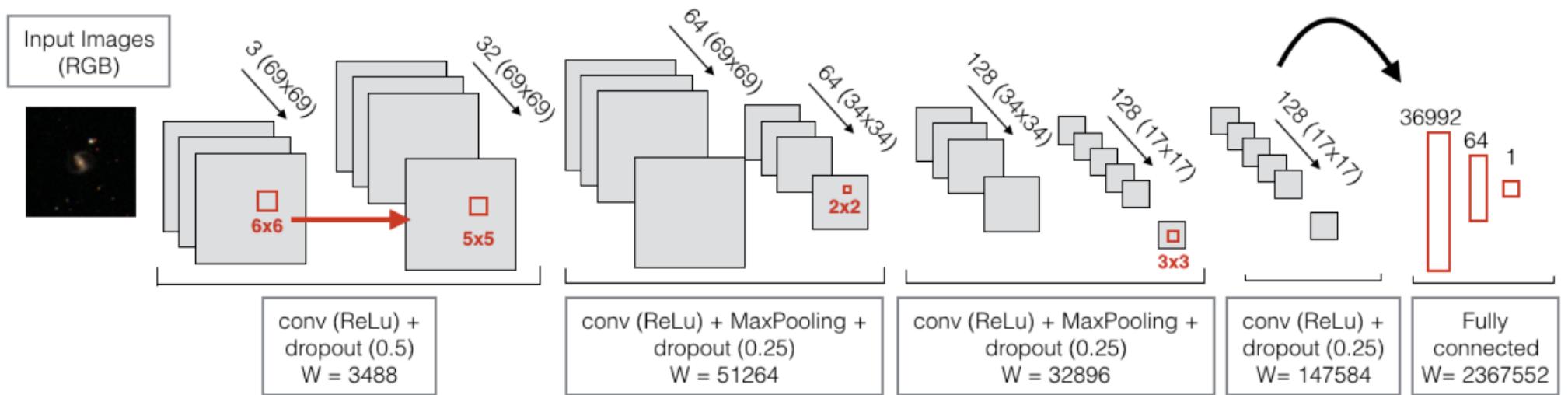


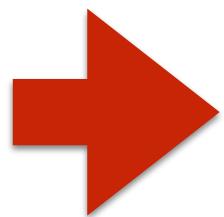
## PART IV: BEYOND CLASSIFICATION: IMAGE2IMAGE NETWORKS

# UP TO NOW CNNs MAP IMAGES (SIGNALS) INTO FLOATS



Dominguez-Sanchez+18

Classification has its limits ....



**HOW DO I CLASSIFY THIS IMAGE?**

Classification has its limits ....



classification

person, sheep, dog



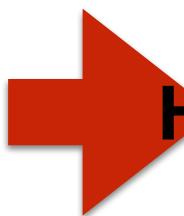
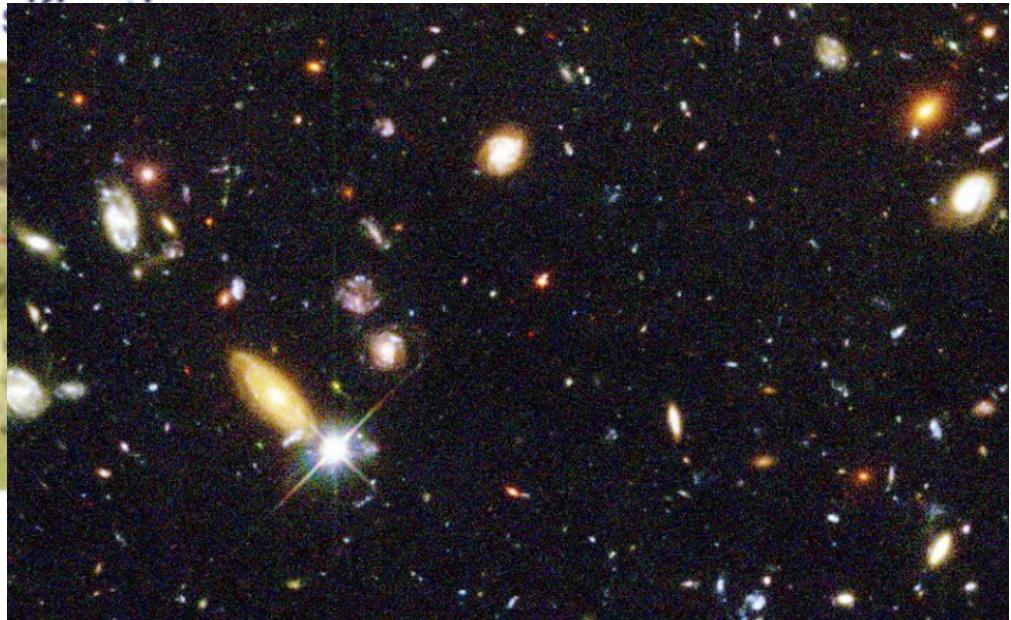
**→ HOW DO I CLASSIFY THIS IMAGE?**

Classification has its limits ....



classifi

per



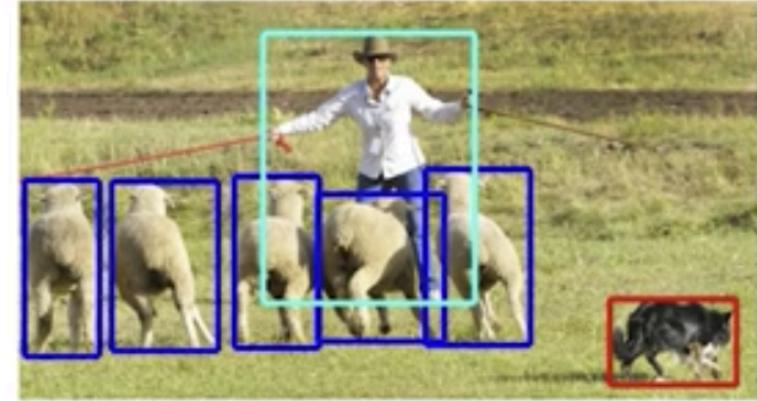
**HOW DO I CLASSIFY THIS IMAGE?**

# Going beyond classification: increasing complexity

classification



object detection



semantic segmentation

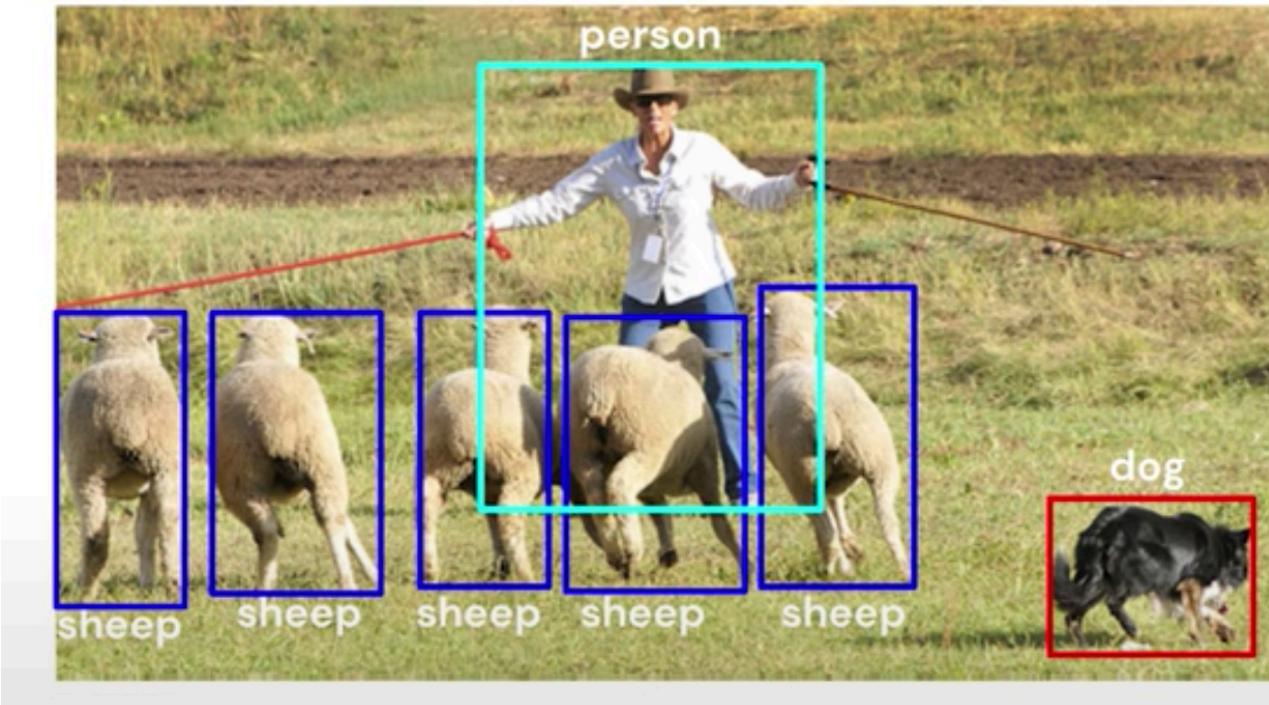


instance segmentation



# Object detection

First task is to find a bounding box for every object. How we do that?



## Inputs

- RGB image  $H \times W \times 3$

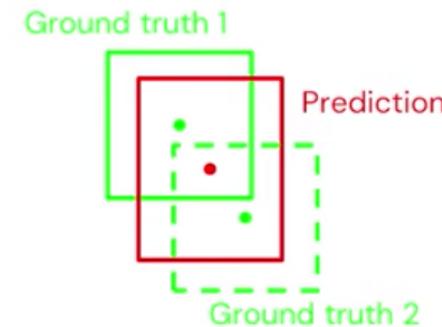
## Targets

- Class label one\_hot  $0\ 0\ 0\ 1\ 0\ ...$
- Object bounding box  
 $(x_c, y_c, h, w)$

for all the objects present in the scene

# WHAT WOULD BE THE LOSS FUNCTION OF SUCH A PROBLEM?

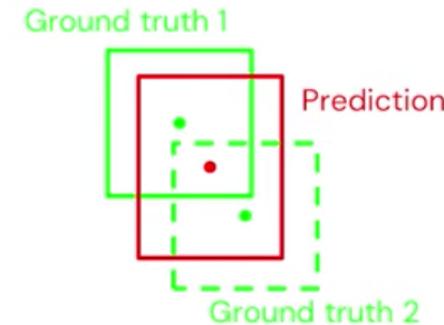
$$\frac{1}{N} \sum_{i=1}^N (y_i - p_i)^2$$



IT IS A REGRESSION WITH A SIMPLE QUADRATIC LOSS.  
WE TRY TO FIND THE BEST COORDINATES OF THE  
BOUNDING BOX.

# WHAT WOULD BE THE LOSS FUNCTION OF SUCH A PROBLEM?

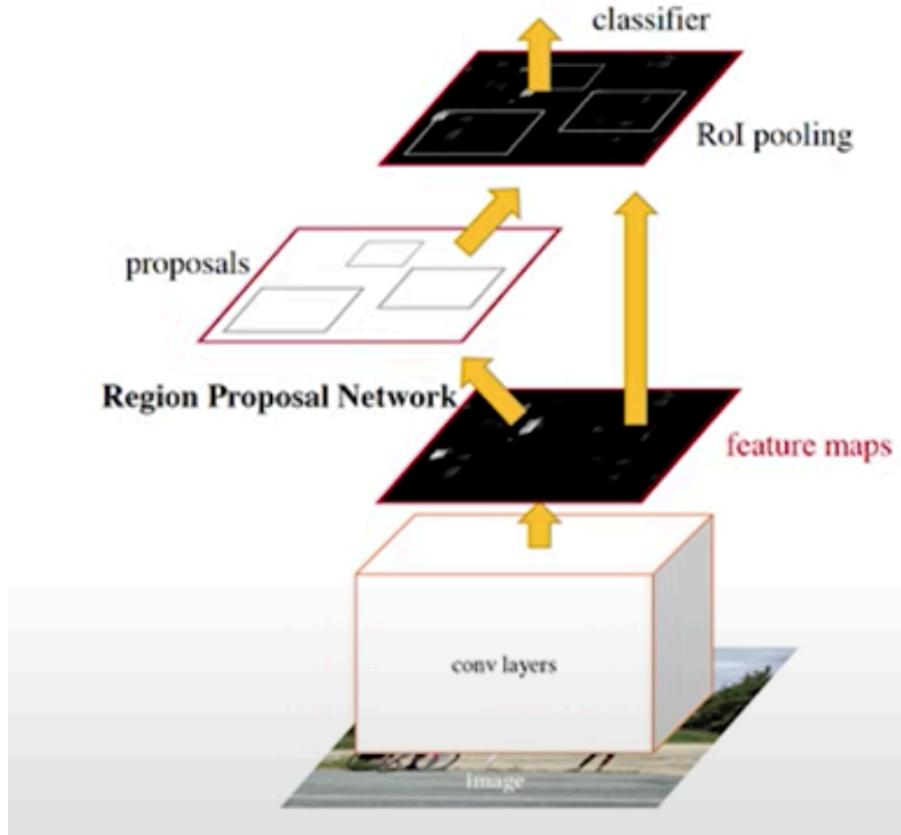
$$\frac{1}{N} \sum_{i=1}^N (y_i - p_i)^2$$



IT IS A REGRESSION WITH A SIMPLE QUADRATIC LOSS.  
WE TRY TO FIND THE BEST COORDINATES OF THE  
BOUNDING BOX.

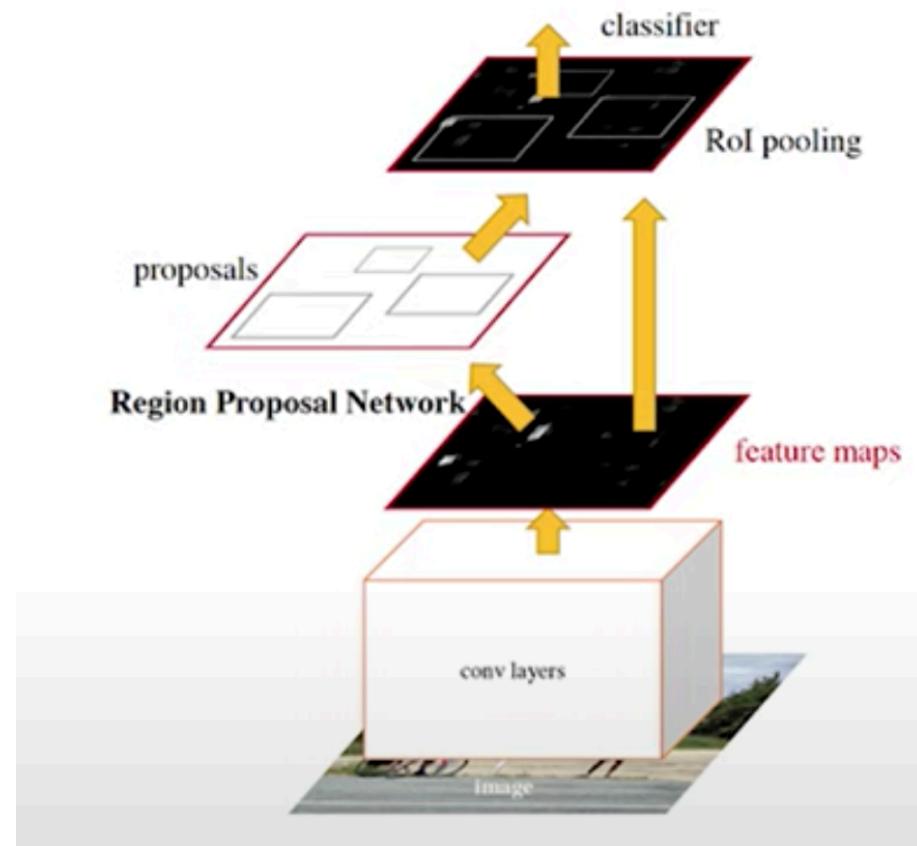
IT BECOMES MESSY QUITE RAPIDLY...

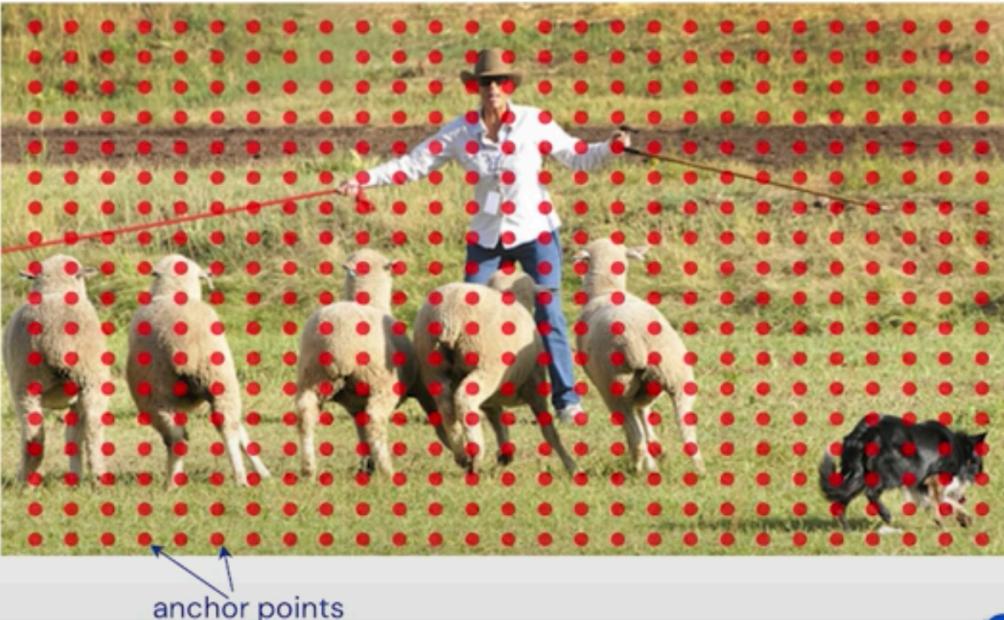
# EXAMPLE: FASTER R-CNN



We divide the task in 2 steps:

1. Identify bounding box candidates (CLASSIFICATION)
2. Classify and Refine (REGRESSION)





Discretize Bounding  
Box Space



Choose n candidates  
per position

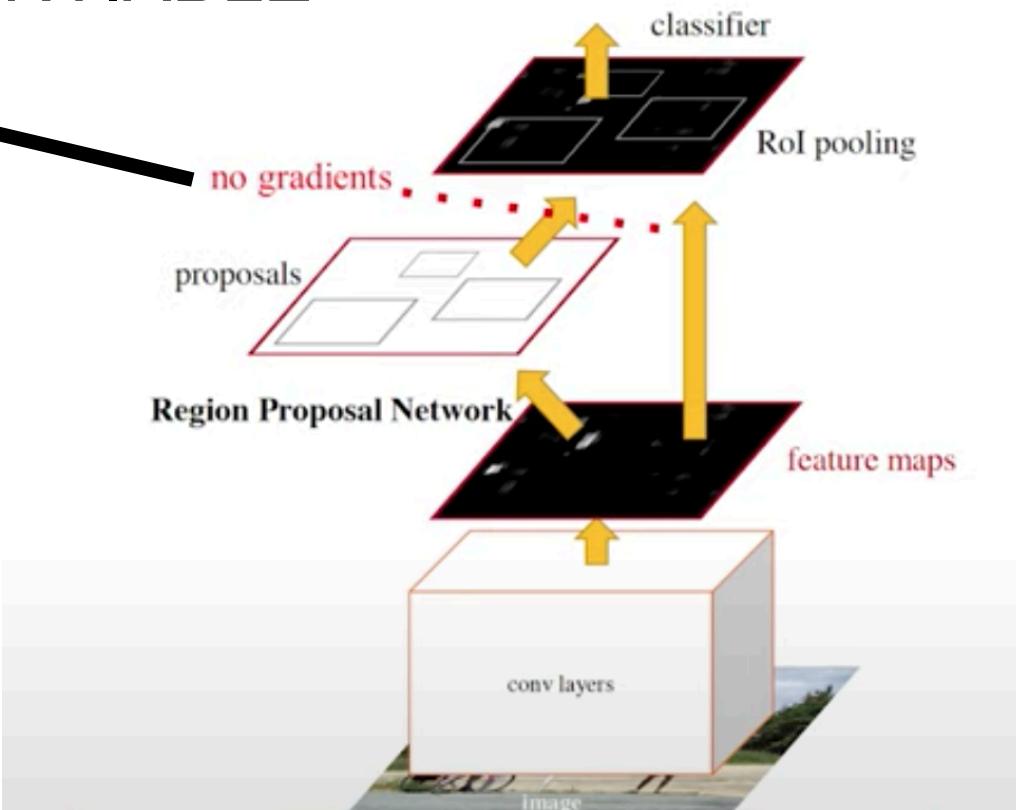


Predict objectness  
score (classification)

Sort and keep top K

# THIS IS NOT DIFFERENTIABLE

(FIXED)



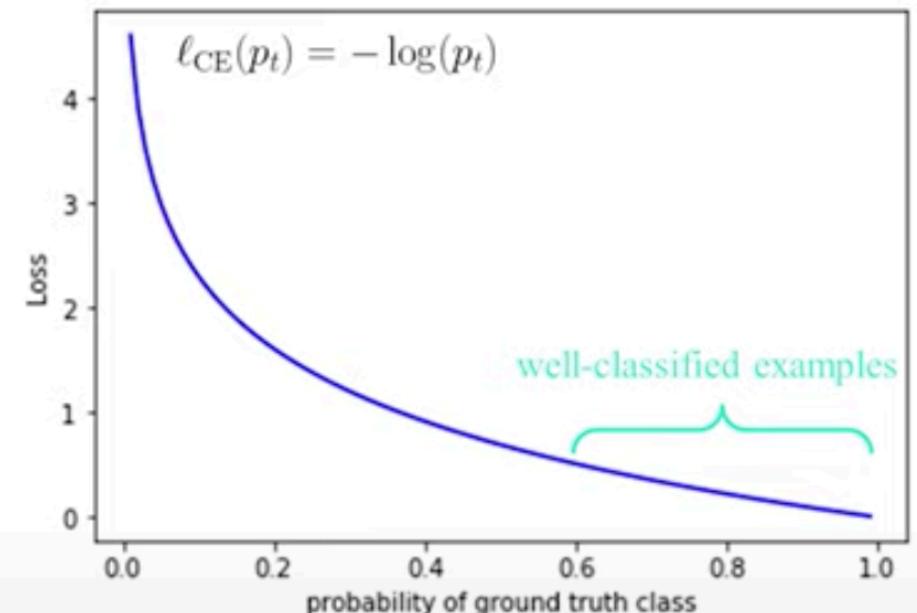
## WAYS OF MAKING THIS DIFFERENTIABLE:

Spatial Transformer Networks (Jaderberg+18)

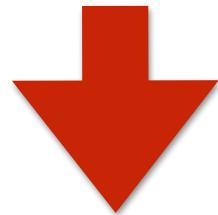
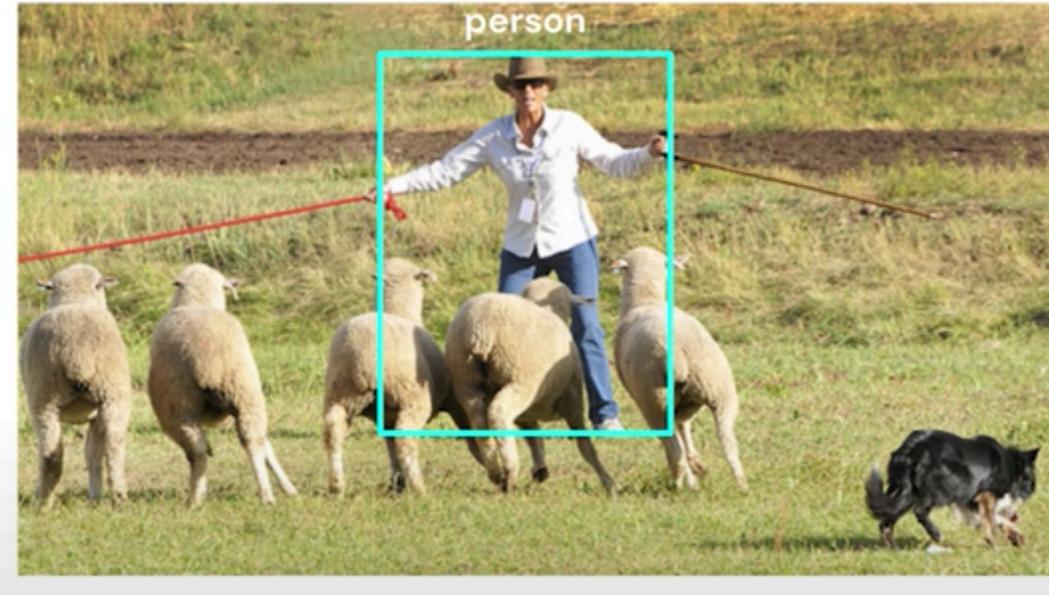
# WHY NOT DOING IT ONE STAGE?

MOST OF THE CANDIDATES  
ARE BACKGROUND,  
EASY TO IDENTIFY

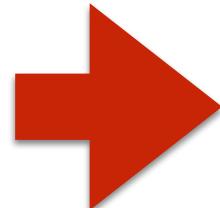
THE LOSS OF THE MANY  
EASY EXAMPLES  
DOMINATES OVER THE  
RARE USEFUL ONES



# LET'S GO A STEP FURTHER INTO SEMANTIC SEGMENTATION



BOUNDING BOXES  
ARE NOT ALWAYS  
GOOD  
REPRESENTATIONS



semantic segmentation

