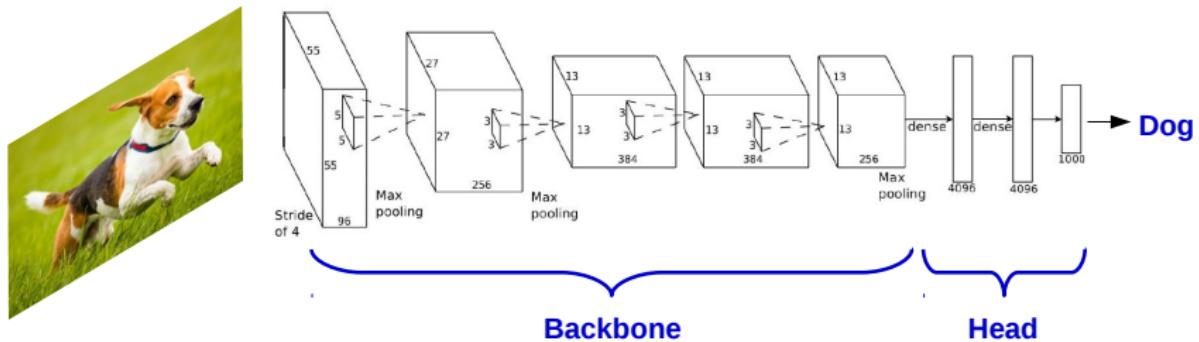


# Deep Learning for Visual Recognition: Application to Pose Recognition

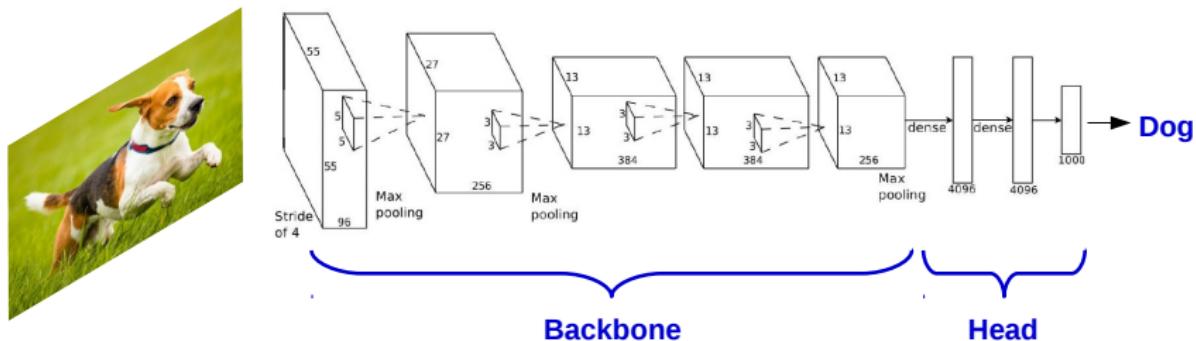
Alvaro Soto

Computer Science Department (DCC), PUC

# Visual Recognition with CNNs

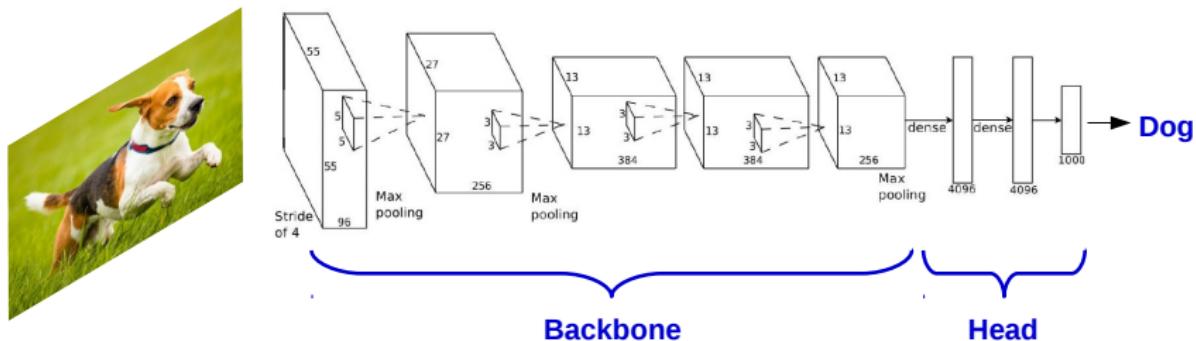


# Visual Recognition with CNNs



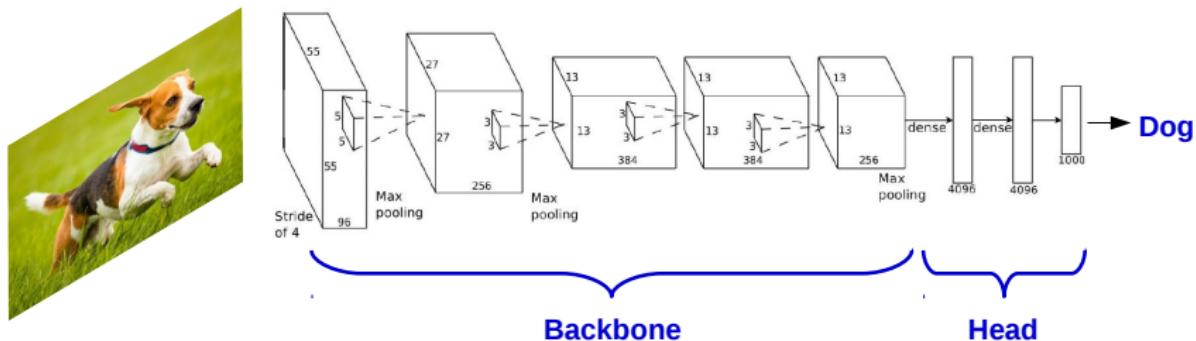
- Last class we discuss about the two main components of a CNN for visual recognition: Backbone and Head.

# Visual Recognition with CNNs



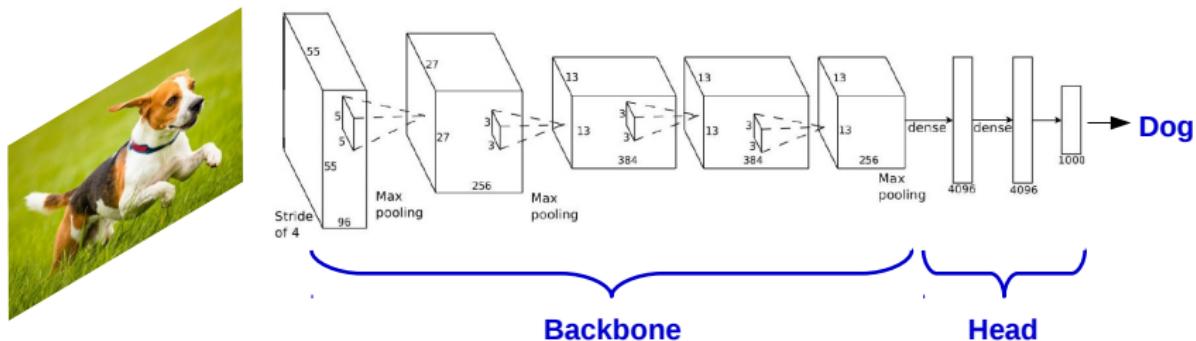
- Last class we discuss about the two main components of a CNN for visual recognition: Backbone and Head.
- Main job of backbone is **feature learning**.

# Visual Recognition with CNNs



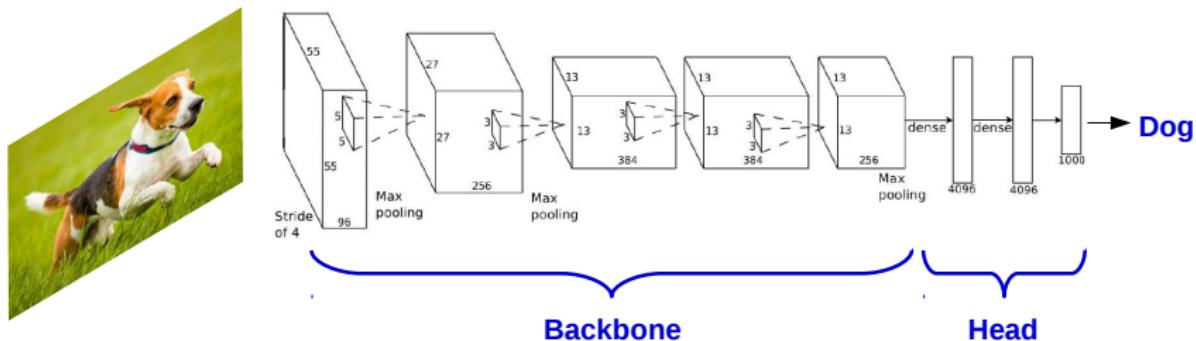
- Last class we discuss about the two main components of a CNN for visual recognition: Backbone and Head.
- Main job of backbone is **feature learning**.
- Main job of head is **classification**.

# Visual Recognition with CNNs



- Last class we discuss about the two main components of a CNN for visual recognition: Backbone and Head.
- Main job of backbone is **feature learning**.
- Main job of head is **classification**.

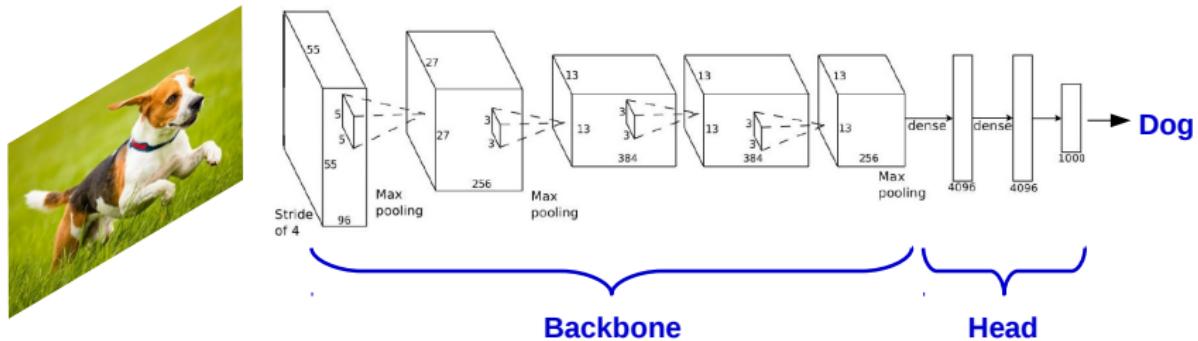
# Visual Recognition with CNNs



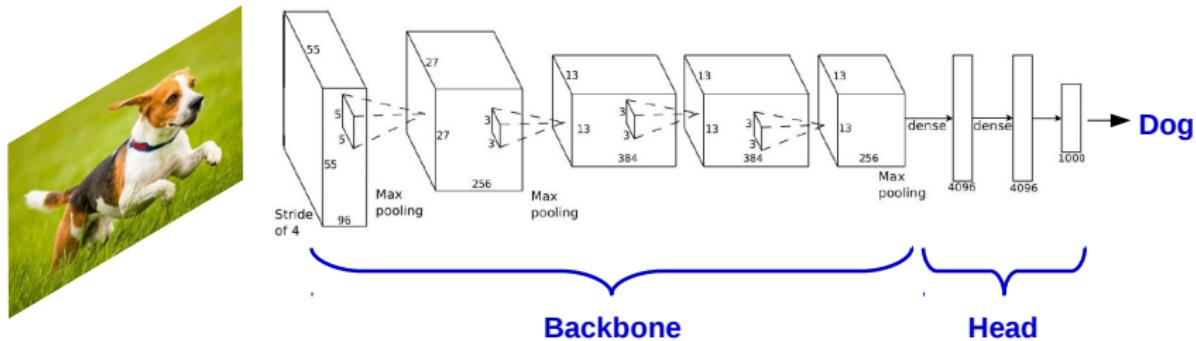
- Last class we discuss about the two main components of a CNN for visual recognition: Backbone and Head.
- Main job of backbone is **feature learning**.
- Main job of head is **classification**.

Note: This is not the same than the encoder/decoder process but it is related.

# Visual Recognition with CNNs

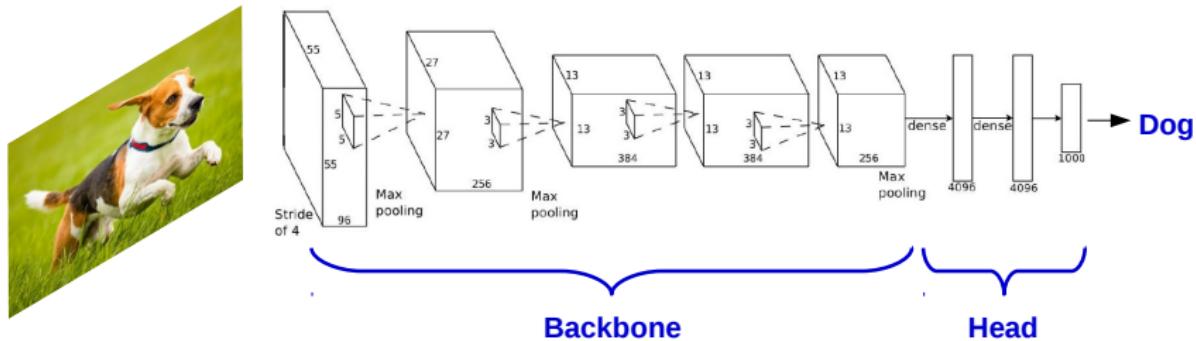


# Visual Recognition with CNNs



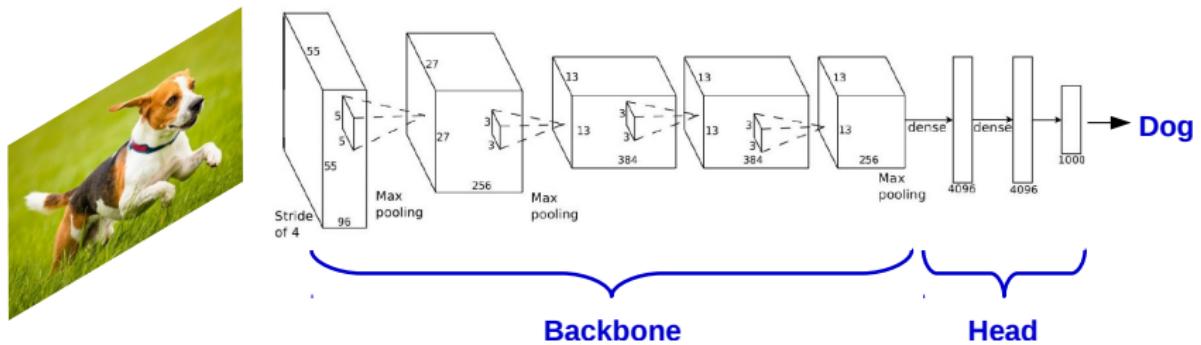
- Main secrets of backbone:
  - Hierarchical and compositional structure.

# Visual Recognition with CNNs

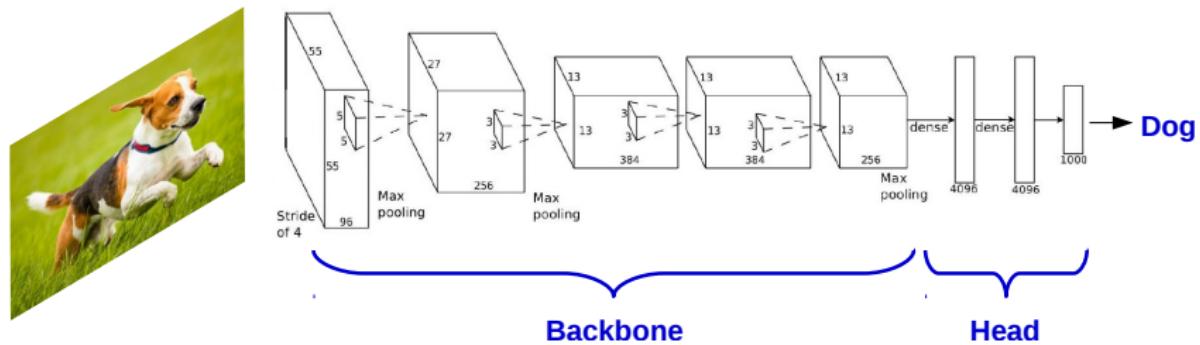


- Main secrets of backbone:
  - Hierarchical and compositional structure.
  - Translation invariance (convolution).

# Visual Recognition with CNNs

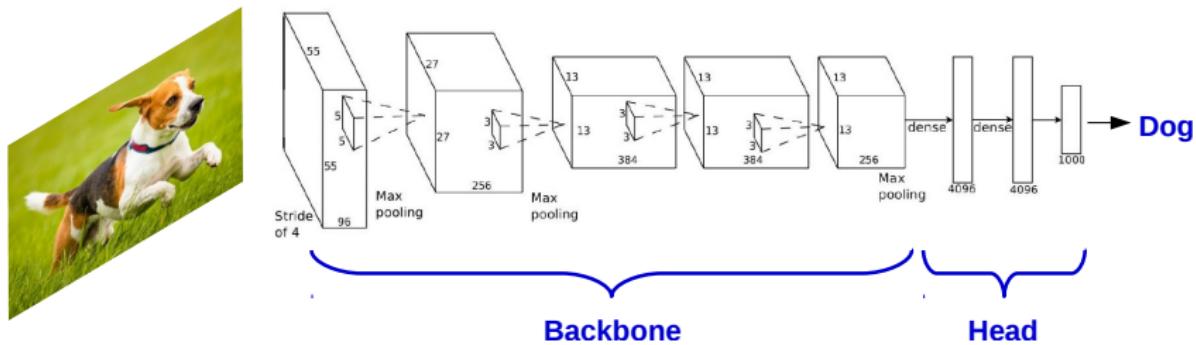


# Visual Recognition with CNNs



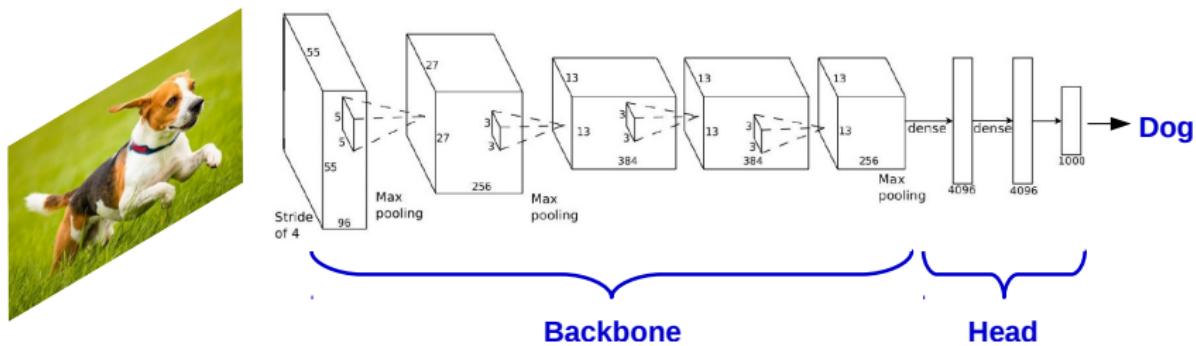
- Main secrets of head:
  - It is on top of the feature learning steps.

# Visual Recognition with CNNs



- Main secrets of head:
  - It is on top of the feature learning steps.
  - High representational capacity.

# Visual Recognition with CNNs

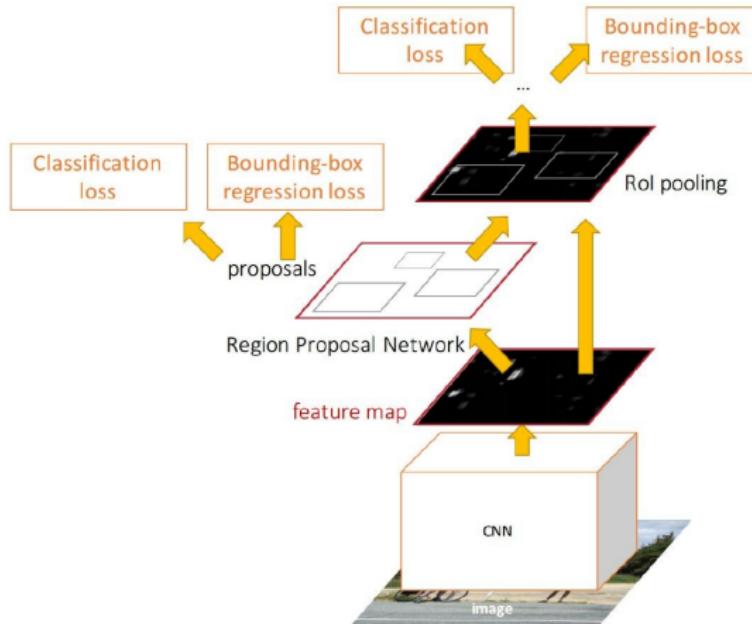


- Main secrets of head:
  - It is on top of the feature learning steps.
  - High representational capacity.
  - We can simultaneously train several heads: **cross-learning**

# Cross Learning



# Faster-RCNN: Multiple heads



Any cross-learning?

# This class

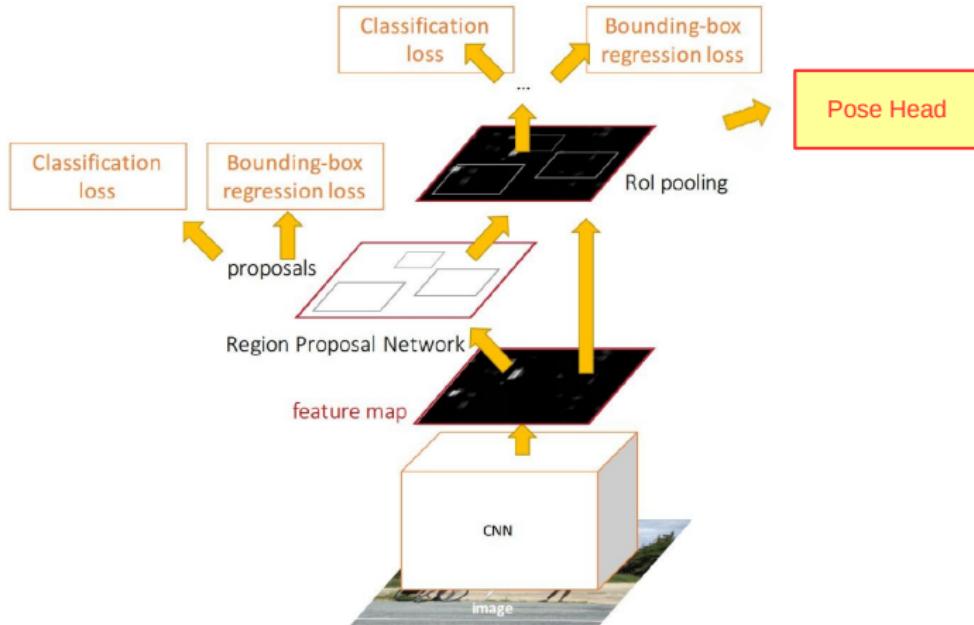
## Learning Human Poses

# Learning Human Poses

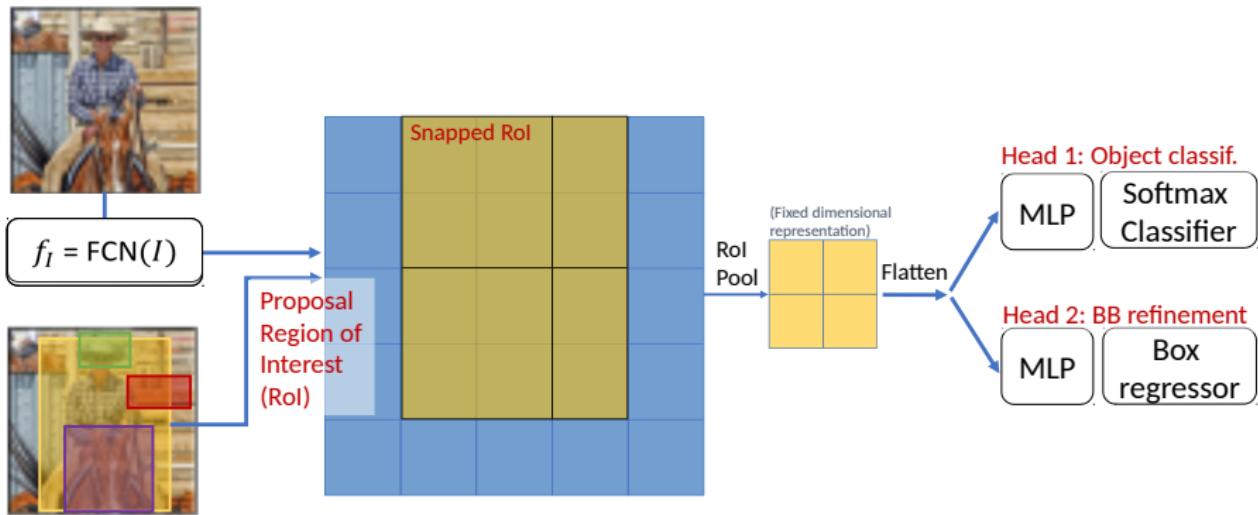


Any ideas?

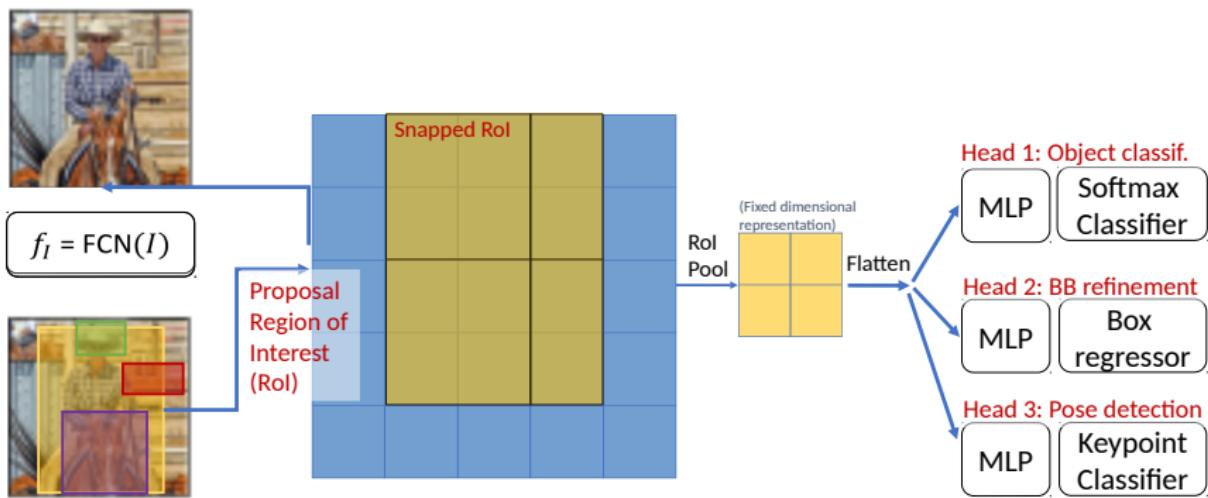
# Faster-RCNN: Multiple heads



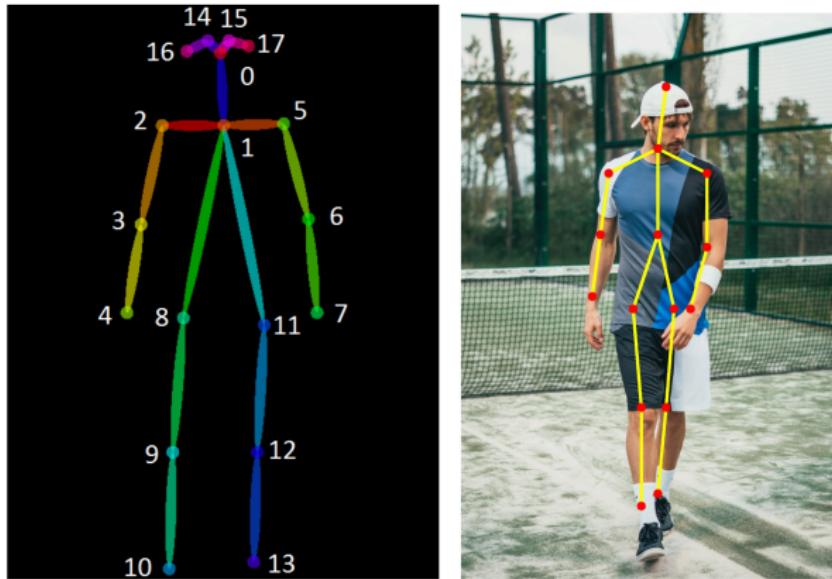
# Faster-RCNN: Multiple heads



# Faster-RCNN: Adding a head for pose detection

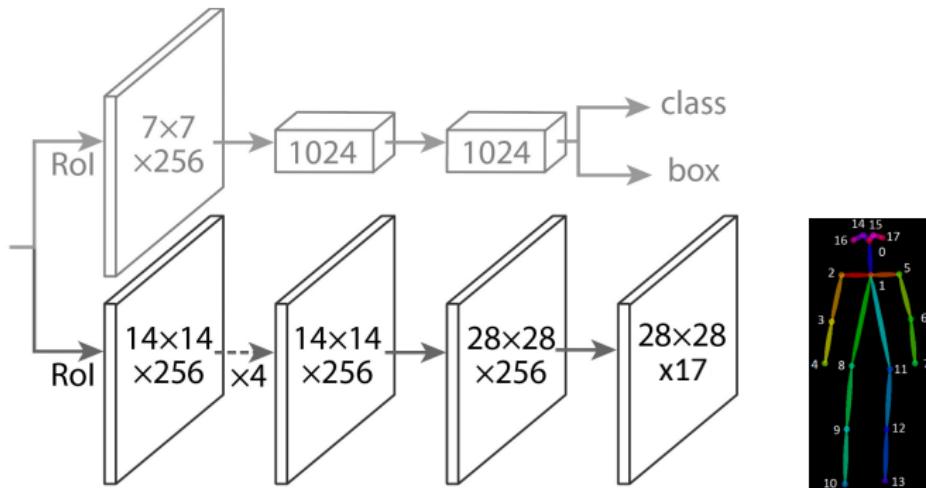


# Pose Representation

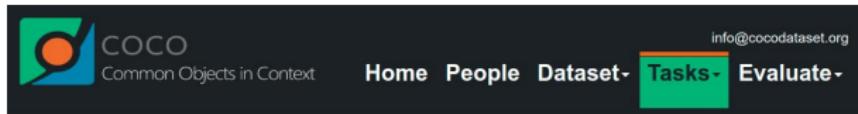


17 Keypoints, representing mostly body joints

# Adding New Head to Faster-RCNN



# Need of a Suitable Dataset



The image shows the COCO dataset homepage. At the top left is the COCO logo (a stylized 'C' inside a square). Next to it is the text "COCO" and "Common Objects in Context". On the right side of the header are links: "Home", "People", "Dataset", "Tasks", and "Evaluate". The "Tasks" link is highlighted with a green background and white text. The "Evaluate" link is also partially visible. The email "info@cocodataset.org" is located at the top right.

## COCO 2020 Keypoint Detection Task



<https://cocodataset.org/#keypoints-2020>

# Results



# Results

