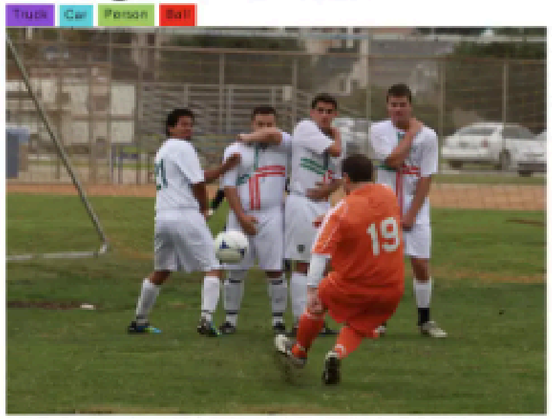
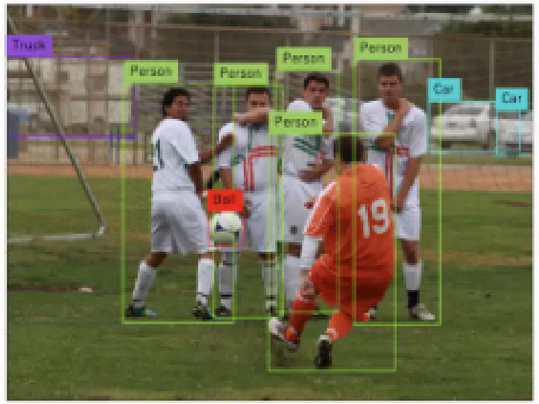
# Mask R-CNN

# What is Mask R-CNN

1. Image Classification

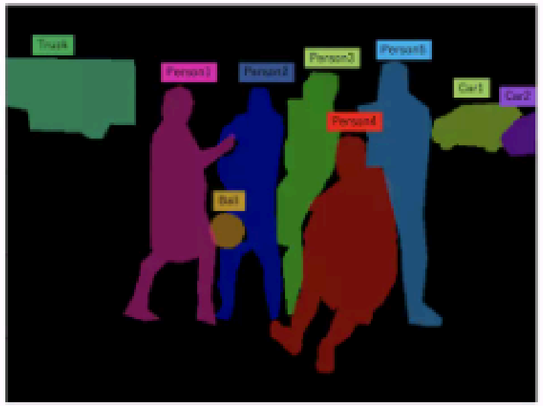



2. Object Detection



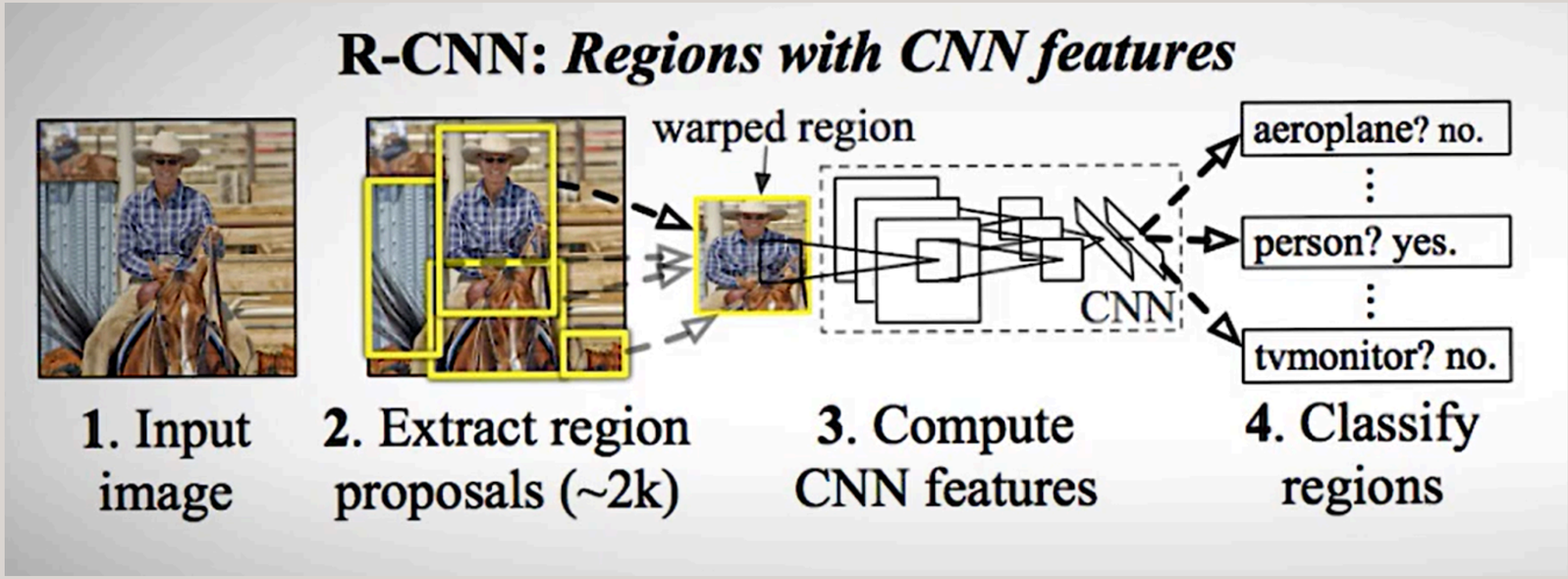

3. Semantic Segmentation


4. Instance Segmentation



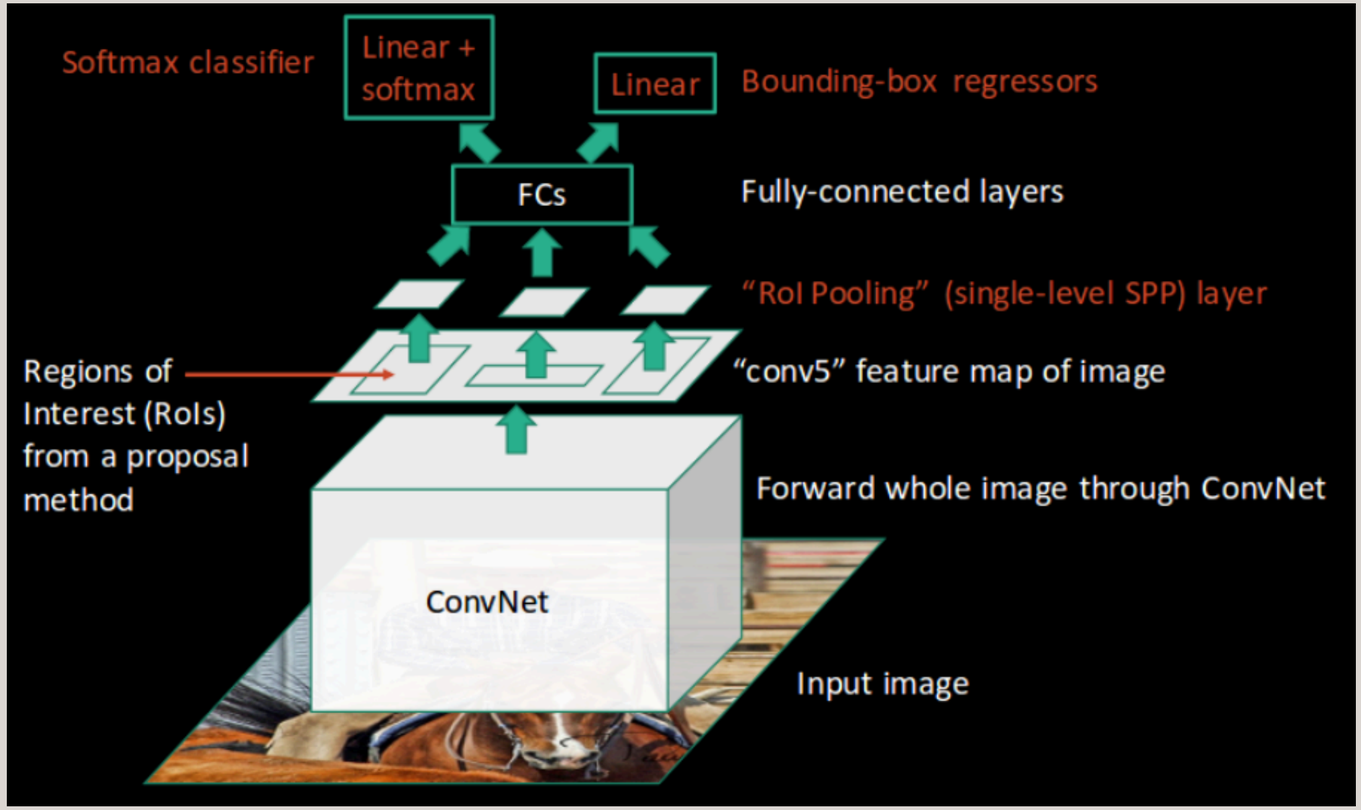

Mask RCNN

# R-CNN



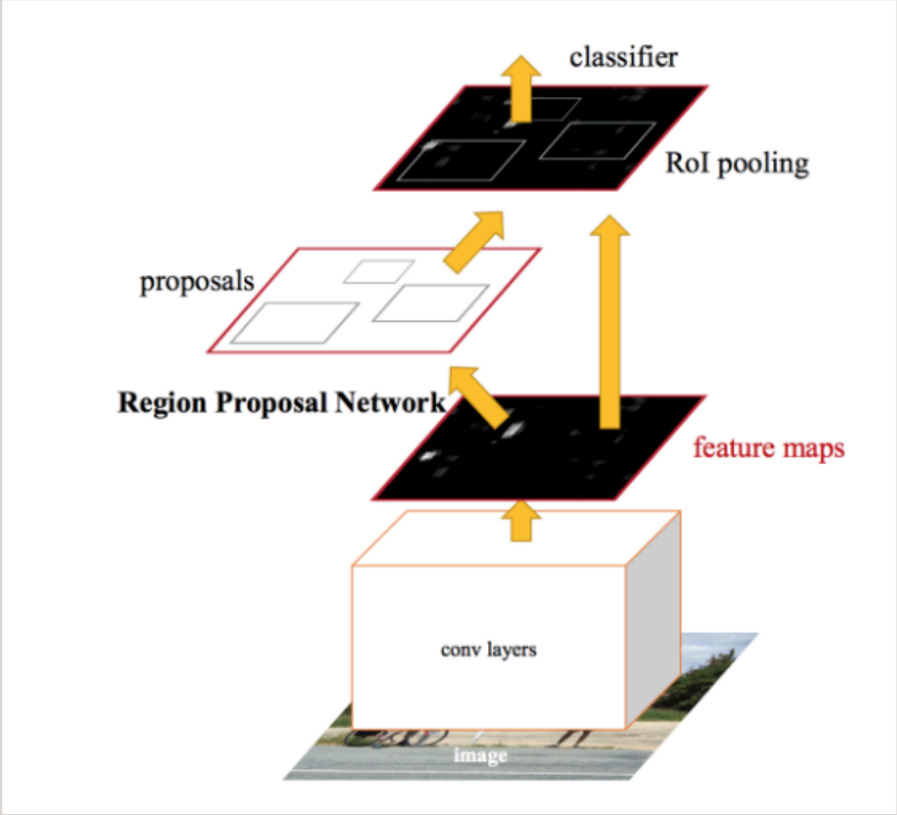

=> Extremely slow because has to process ~ 2000 region

# Fast R-CNN




=> Not fast enough, especially with large dataset

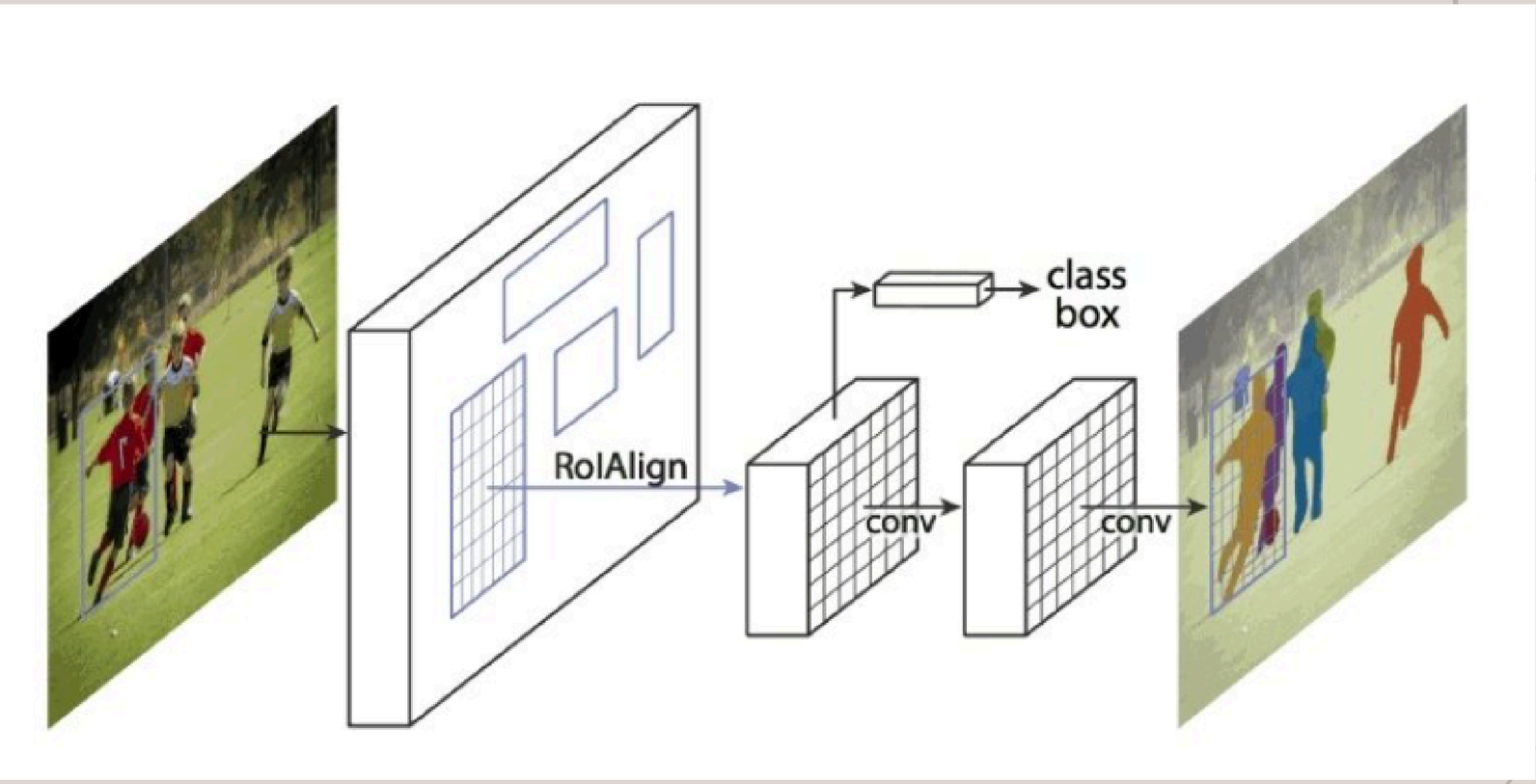
# Faster R-CNN


classifier
RoI pooling
proposals
Region Proposal Network
feature maps
conv layers
image

# Mask R-CNN

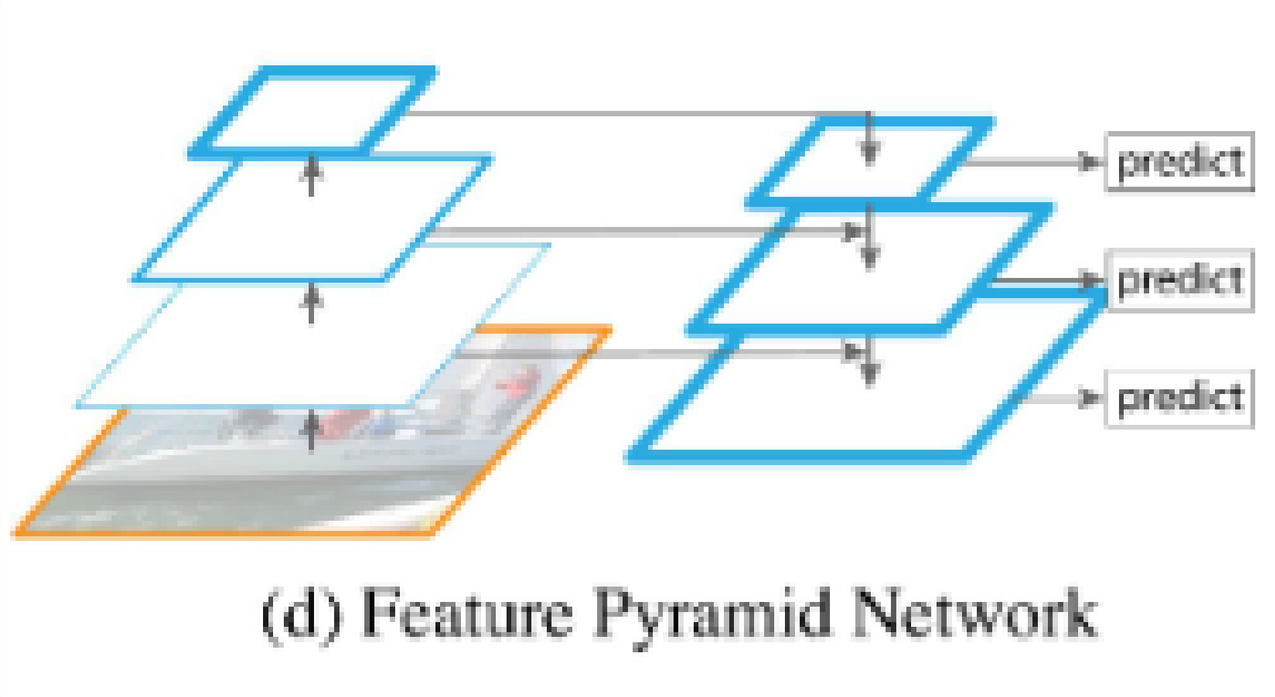
- Mask R-CNN stands for Mask Region Convolutional Network
- Typically for Instant Segmentation tasks
- Evolves through 4 main versions:
  - RCNN → Fast RCNN → Faster RCNN → Mask RCNN
- Main improvement: RoI Align and FCN (Fully Connected Network)

# Mask R-CNN



# Mask R-CNN

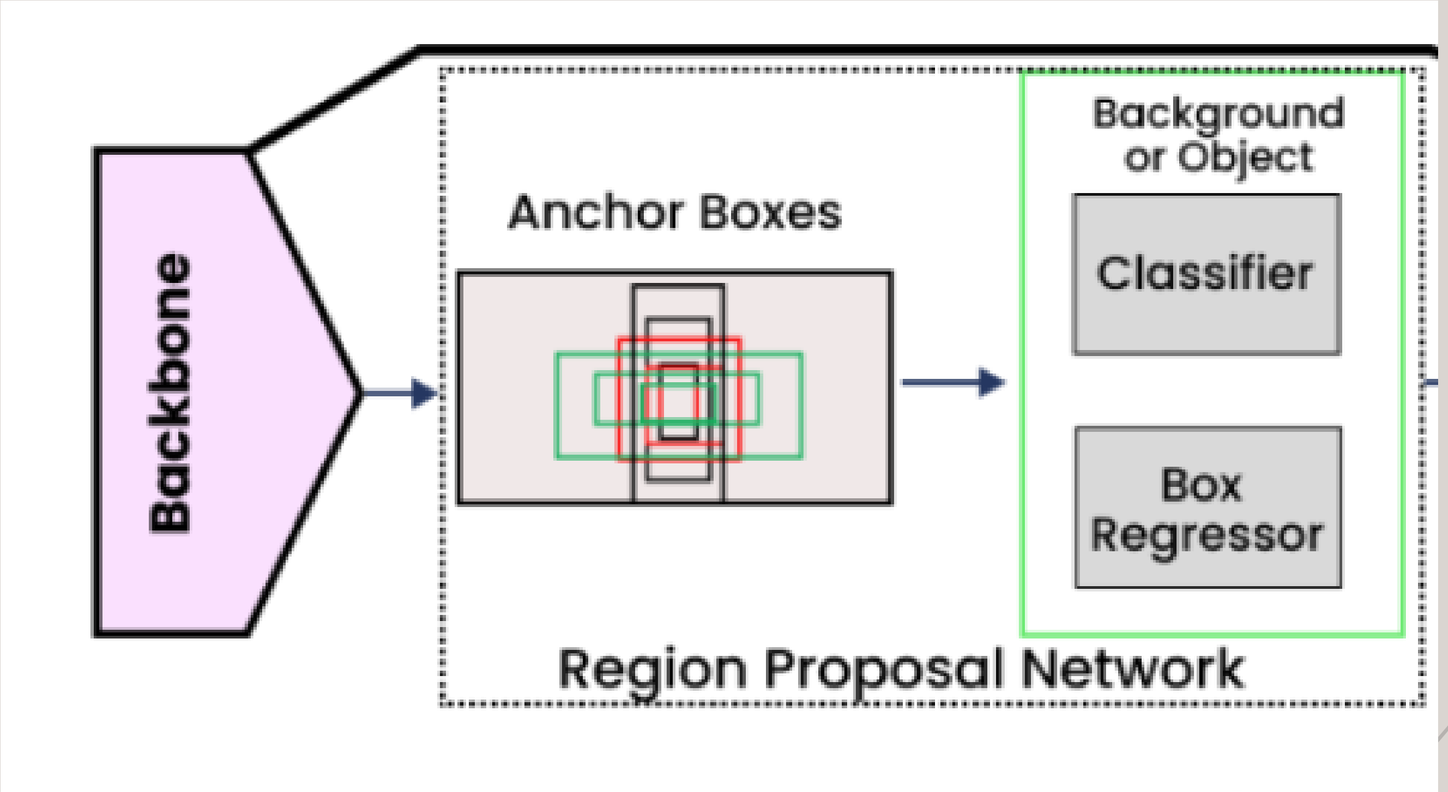
*BackBone*




(d) Feature Pyramid Network

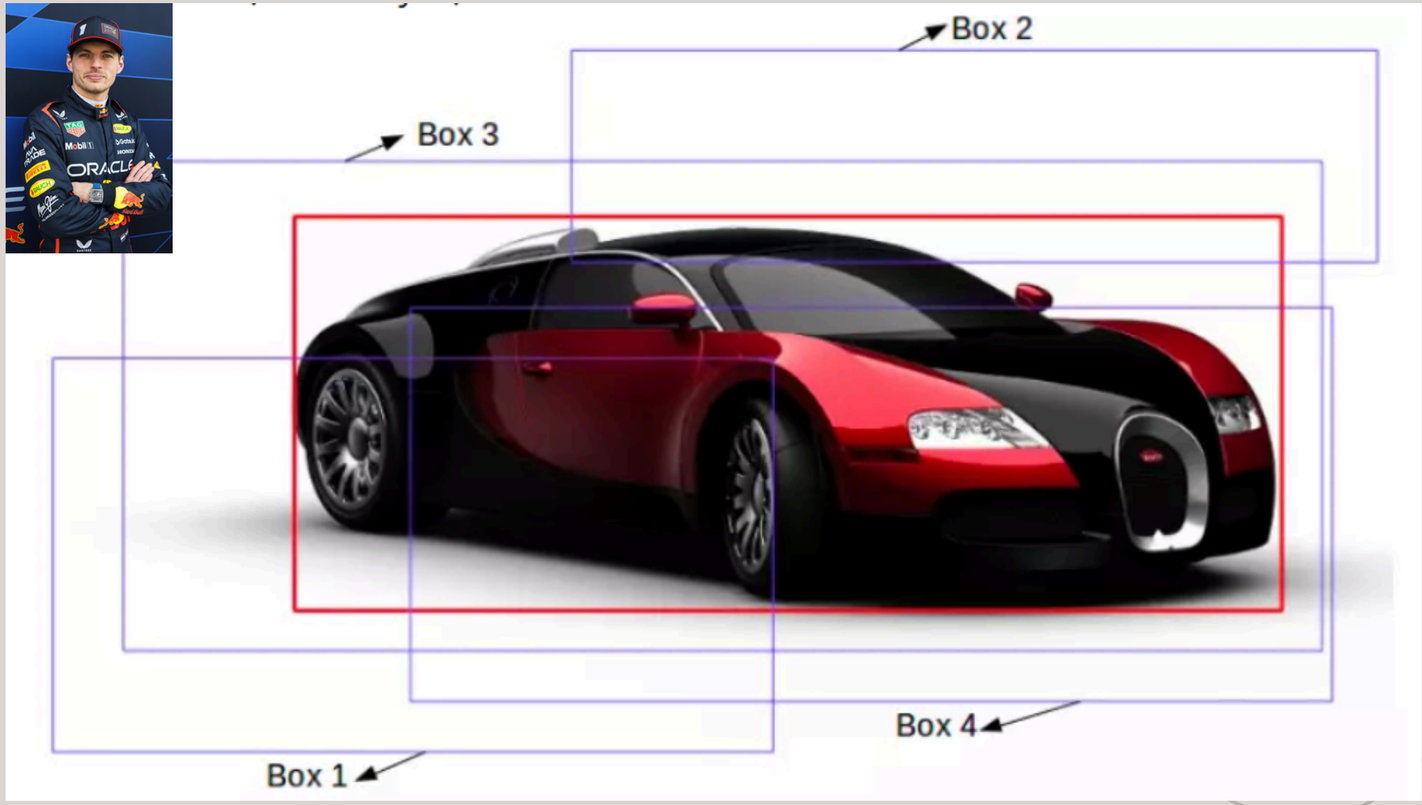
# Mask R-CNN

*Region Proposal Network*



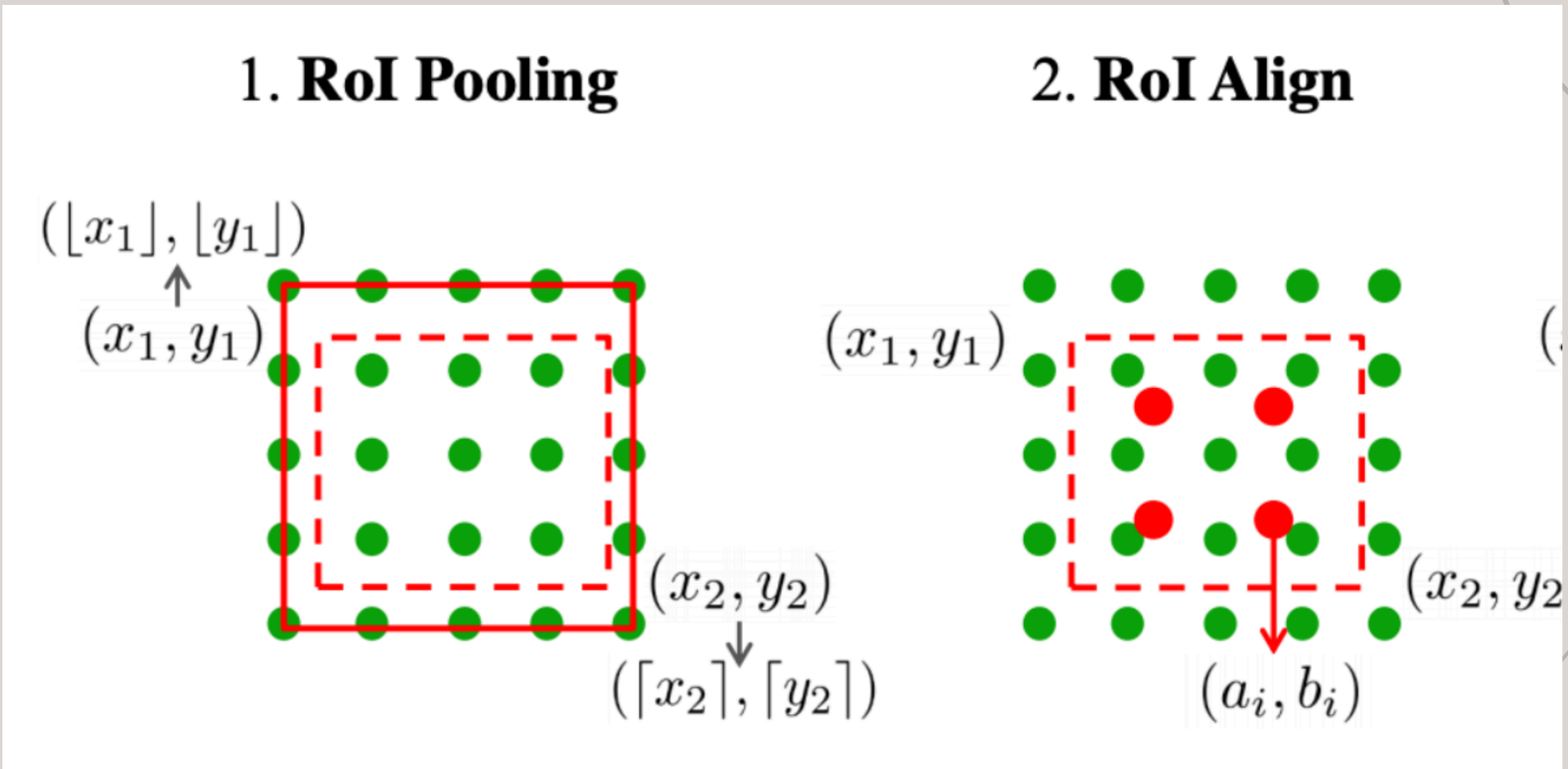
# Mask R-CNN

*RoI (region of interest)*



# Mask R-CNN

*RoI Align*



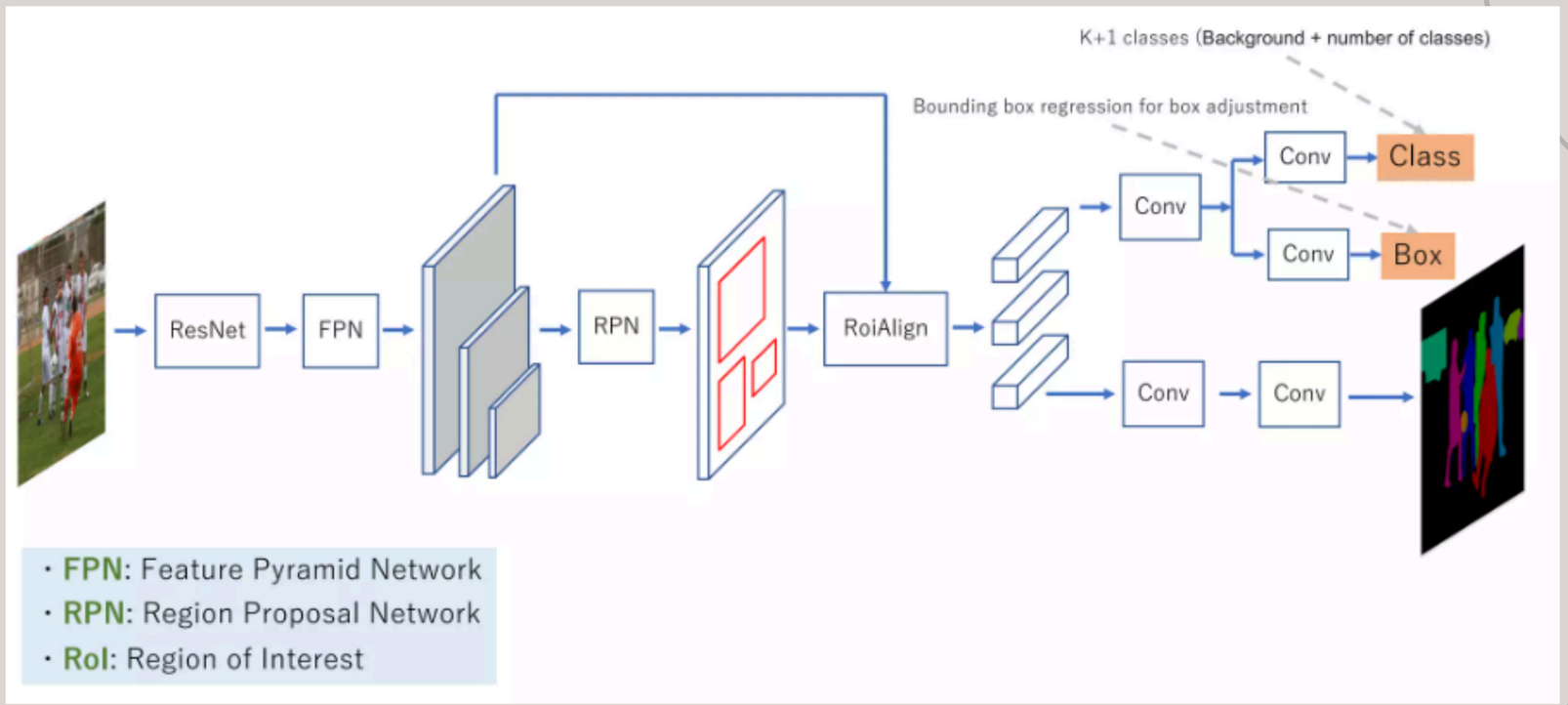
# Mask R-CNN

## *Loss Function*

$$L = L_{cls} + L_{loc} + L_{mask}$$

- Lcls is classification loss
- Lloc is bounding box regression loss
- Lmask is mask loss

# Mask R-CNN


K+1 classes (Background + number of classes)
Bounding box regression for box adjustment
ResNet
FPN
RPN
RoiAlign
Conv
Conv
Conv
Conv
Conv
Class
Box
• FPN: Feature Pyramid Network
• RPN: Region Proposal Network
• RoI: Region of Interest

Ground Truth


Prediction