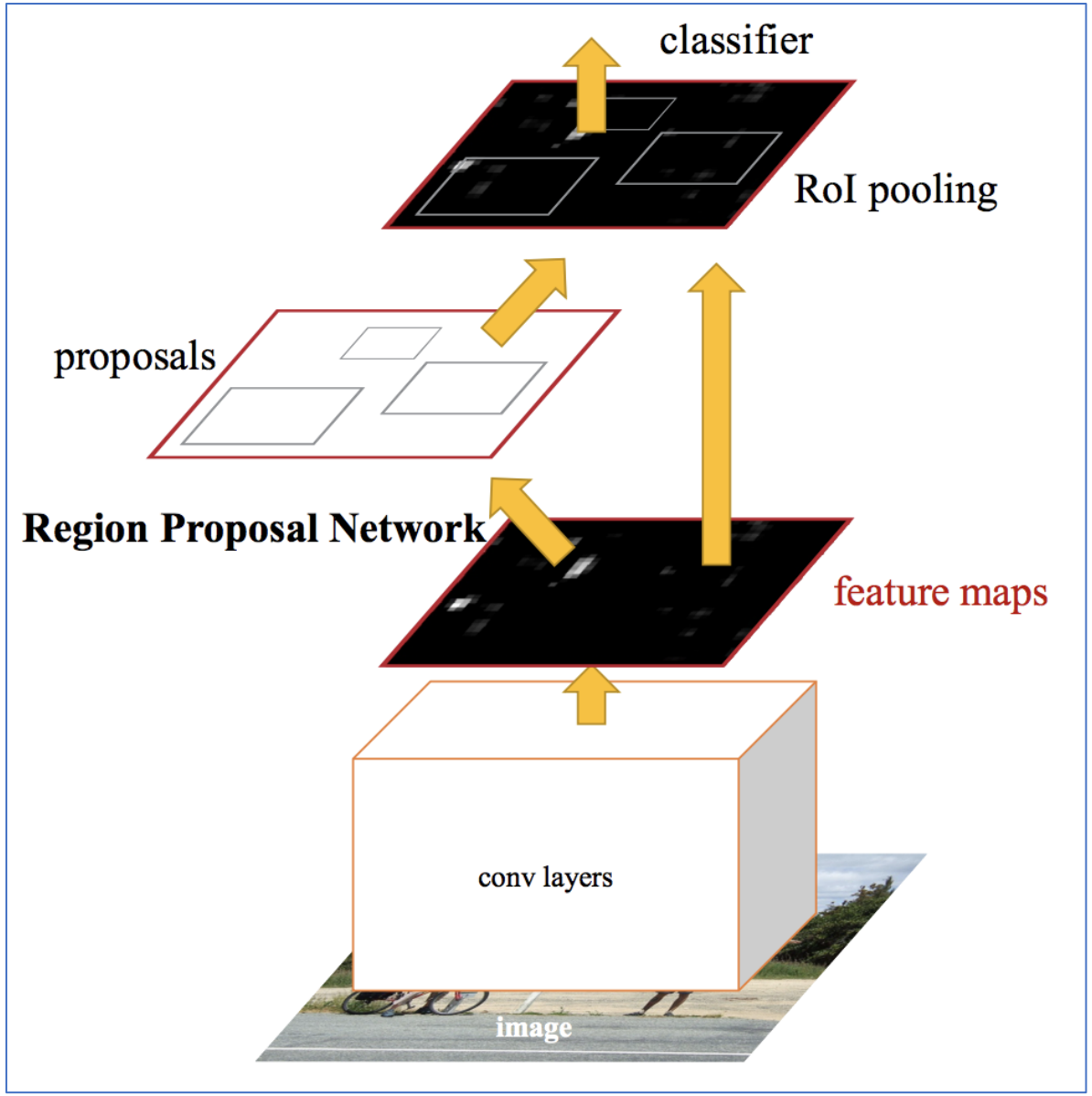
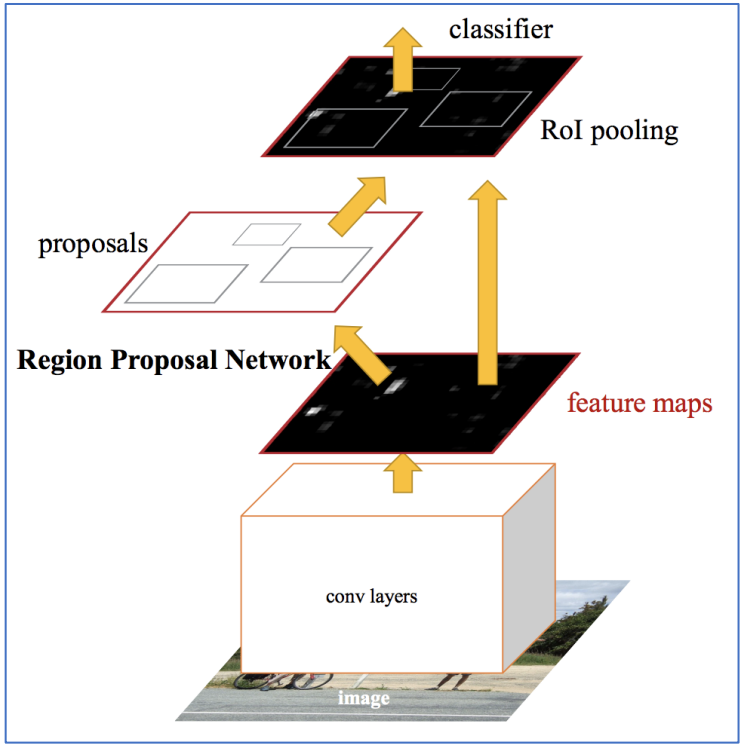
Faster-RCNN

Faster R-CNN

- Although it is a single unified model, the architecture is comprised of two modules:
  - **Module 1: Region Proposal Network**. Convolutional neural network for proposing regions and the type of object to consider in the region.
  - **Module 2: Fast R-CNN**. Convolutional neural network for extracting features from the proposed regions and outputting the bounding box and class labels.



Faster R-CNN



- Both R-CNN & Fast R-CNN, use selective search to find out the region proposals
- Selective search is a slow and time-consuming process affecting the performance of the network
- Therefore, Shaoqing Ren et al. came up with an object detection algorithm that eliminates the selective search algorithm and lets the network learn the region proposals.
- Similar to Fast R-CNN, the image is provided as an input to a convolutional network which provides a convolutional feature map
- Instead of using selective search algorithm on the feature map to identify the region proposals, a separate network, **Region Proposal Network**, is used to predict the region proposals
- The predicted region proposals are then reshaped using a RoI pooling layer which is then used to classify the image within the proposed region and predict the offset values for the bounding boxes
- Faster R-CNN is much faster than it's predecessors. Therefore, it can even be used for real-time object detection.

Faster R-CNN – Region Proposal Network (RPN)

- The RPN works by taking the output of a pre-trained deep CNN, such as VGG-16, and passing a small network over the feature map and outputting multiple region proposals and a class prediction for each
- Region proposals are bounding boxes, based on so-called **anchor boxes** or pre-defined shapes designed to accelerate and improve the proposal of regions
- The class prediction is binary, indicating the presence of an object, or not, so-called "*objectness*" of the proposed region.
- We output the predicted bounding boxes as the proposed regions required by the RoI pooling layer.