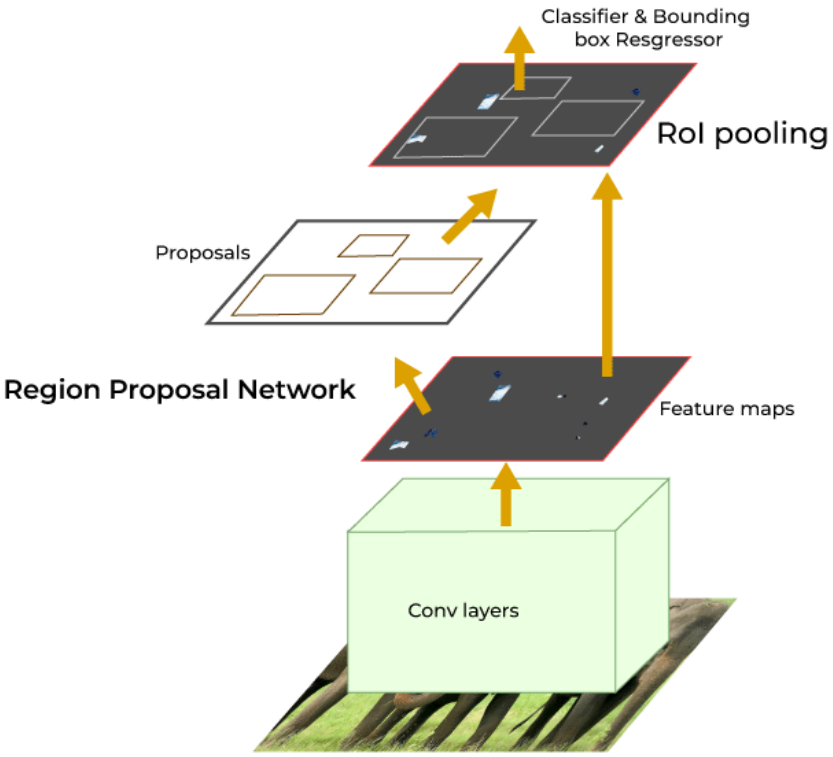
FASTER RCNN

Prior to Faster R-CNN, popular models like R-CNN and Fast R-CNN achieved impressive results but suffered from a critical bottleneck: the region proposal stage which was responsible for identifying potential object locations, was slow and computationally expensive.

Faster R-CNN addressed this bottleneck with two key innovations:

1. Region Proposal Network (RPN):

Replacing slow external algorithms: Traditional object detection models relied on external algorithms like Selective Search to propose regions likely to contain objects. RPN eliminates this dependency, integrating region proposals into the neural network itself.
Simultaneous objectness and bounding box prediction: RPN uses convolutional layers to analyze the image features and simultaneously predicts:
"Objectness" scores for each possible anchor box, indicating the likelihood of an object being present at that location.
Bounding box offsets for each anchor, refining the initial box locations for better accuracy.
Feature sharing and efficiency: RPN and the remaining detection network share convolutional layers, minimizing computational overhead and allowing both tasks to benefit from the same feature extraction.



2. Unified Network Architecture:

End-to-end training and inference: RPN and Fast R-CNN are merged into a single end-to-end network. This unified architecture allows for joint training and optimization, leading to improved performance and smoother inference flow.
ROI Pooling for fixed-size feature extraction: The network extracts features from the entire image but only analyzes specific regions (Regions of Interest) identified by RPN. ROI Pooling extracts a fixed-size feature vector from each region, enabling further classification and bounding box refinement.