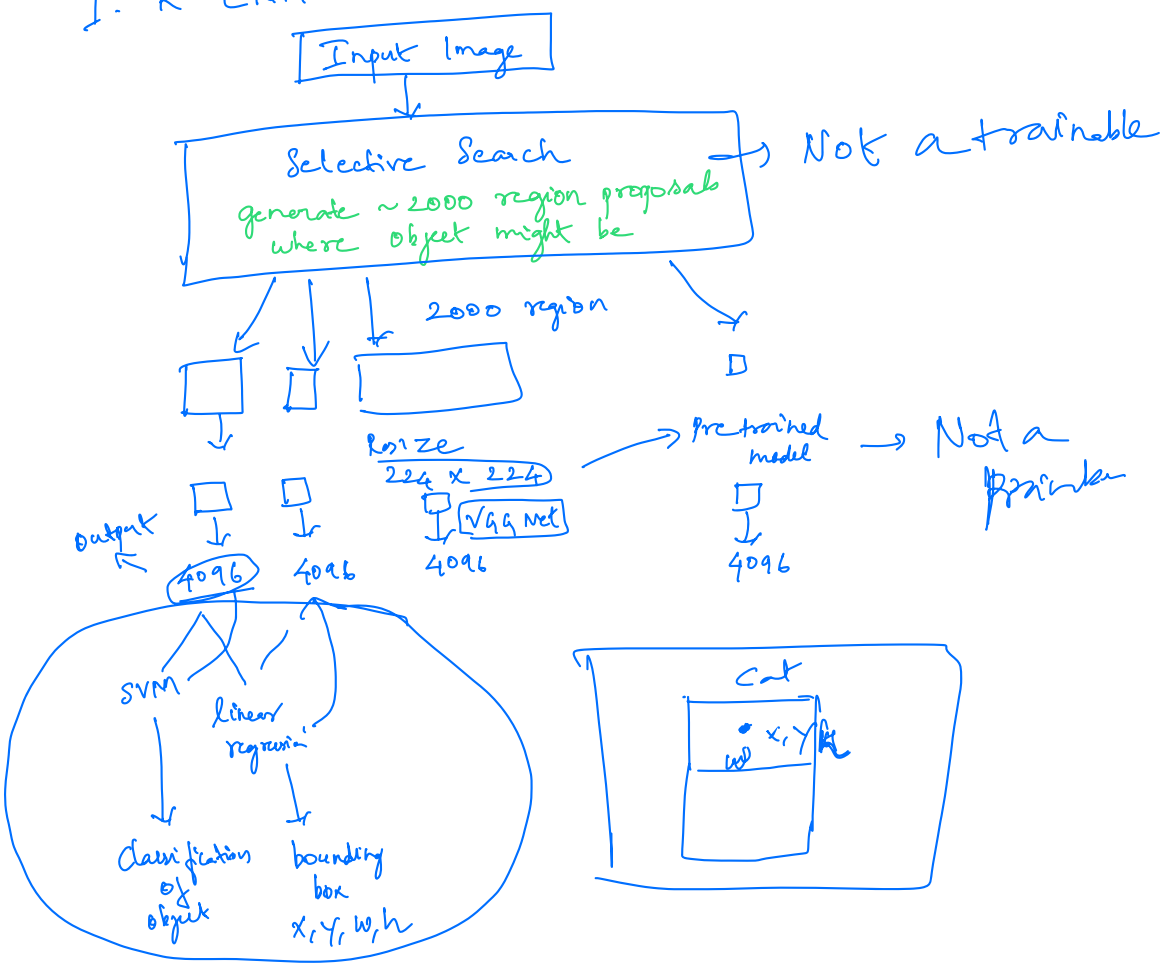
Object Detection

Find what objects are in an image
Locate them using bounding boxes

1. R-CNN — Region Based CNN
2. Fast R-CNN
3. Faster R-CNN

1. R-CNN

2014



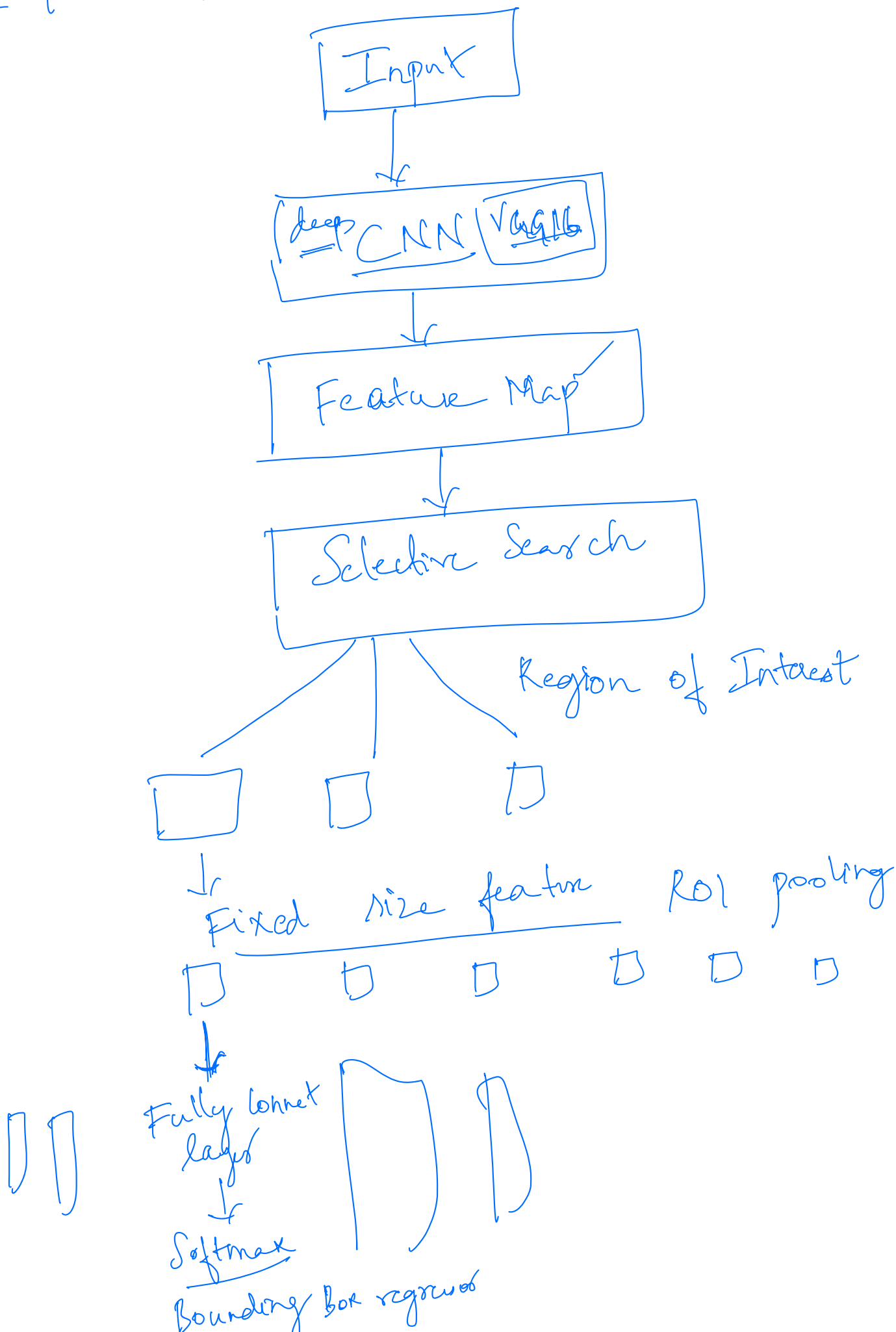
CNN runs 2000 times per image

Not end-end trainable

Pros: First successful deep learning object detector

Cons: very slow, Not end-end trainable, Needs lot of disk space

2. Fast - RCNN


Input
deep CNN
VGG16
Feature Map
Selective Search
Region of Interest
Fixed size feature
RoI pooling
Fully connet layer
Softmax
Bounding Box regressor

Faster than RCNN

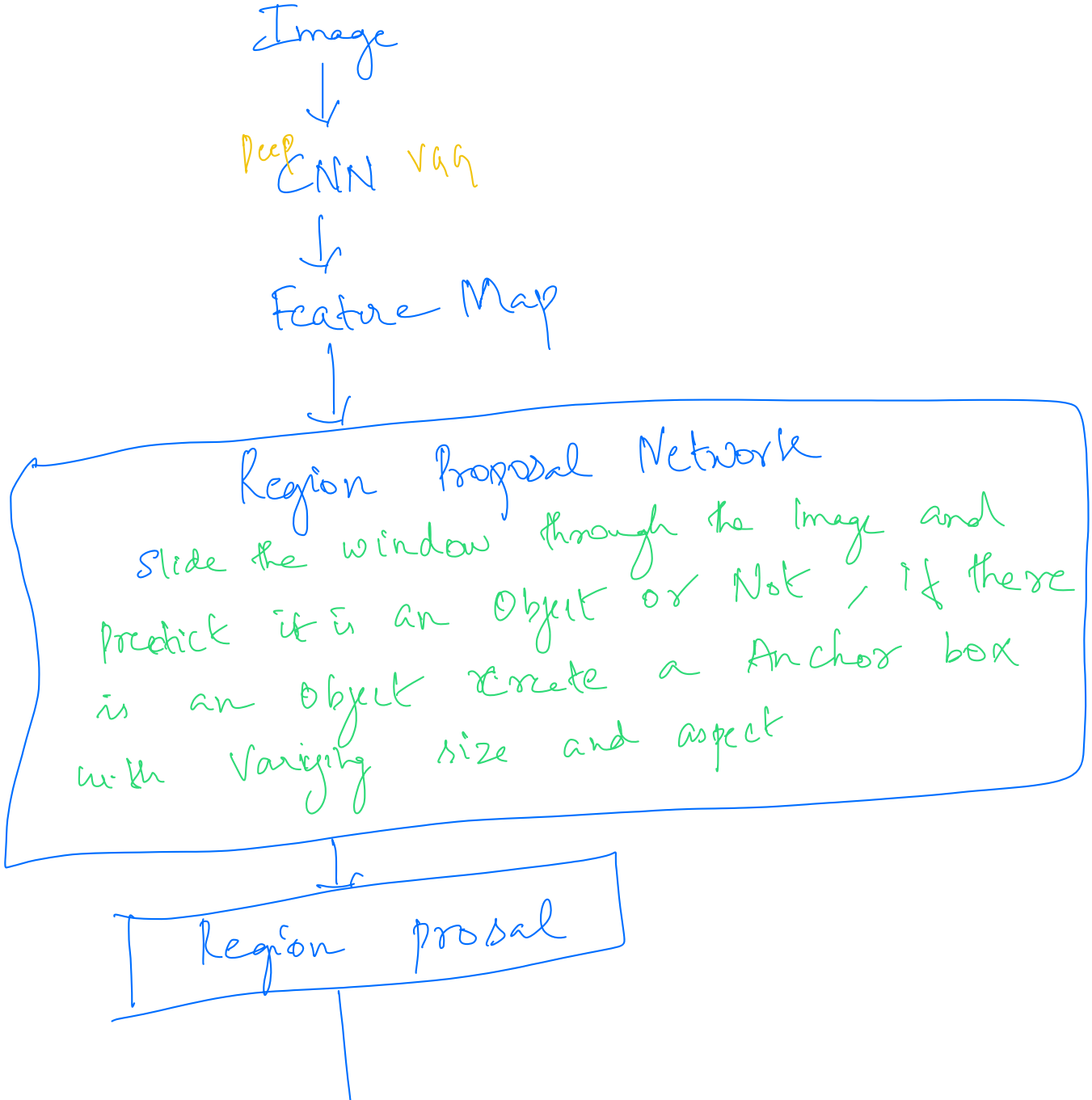
End - end training (excluding Selective Search)

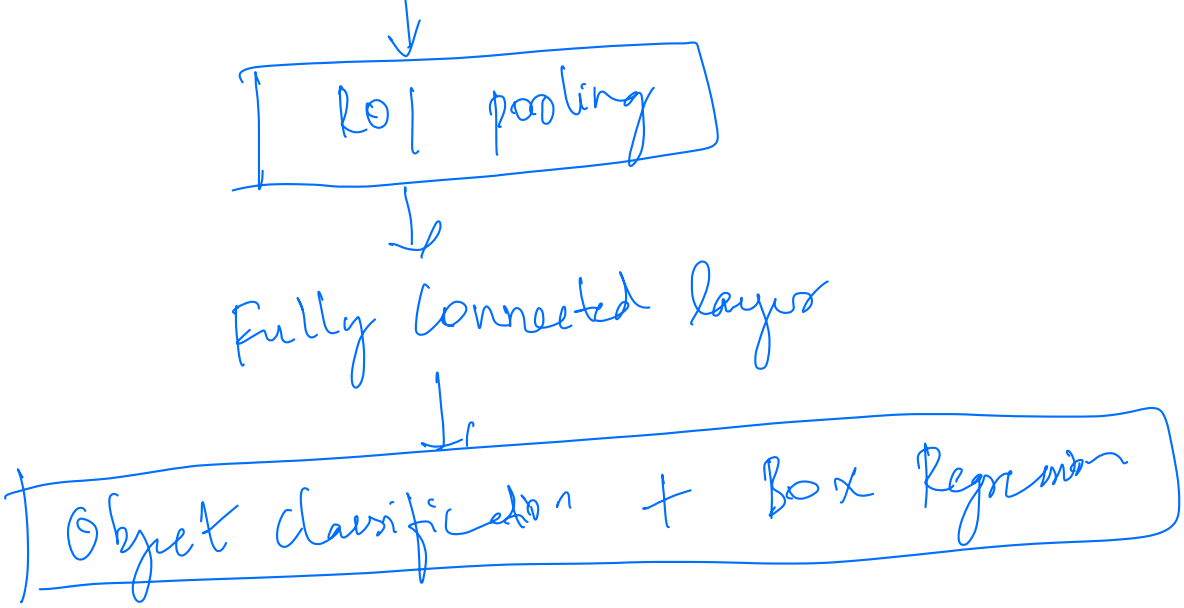
cons

Still depends on slow selective search

Not fully a real - time

Faster RCNN





Pros:

1. End-to-End
2. High Accuracy
3. Faster than old models (Real time almost)

Cons:

1. Not realtime on mobile
2. Complex Architecture

Summary

	Region Proposal method	Speed	Accuracy	End-End?	Additional
R-CNN	Selective Search	Very slow	High	NO	2000 CNN runs per Image
Fast RCNN	Selective Search	Medium	Better	Partially	One CNN + ROI pooling + Selective Search
Faster RCNN	Region Proposal Network	Faster	Very high	Yes	Smart Region proposal + Anchor box