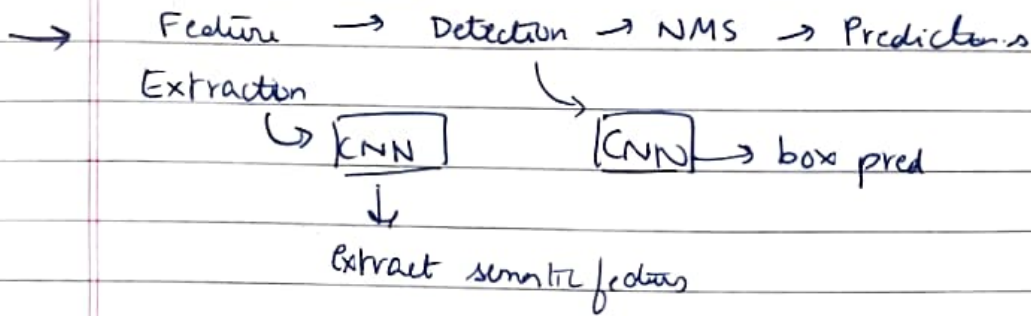


SSD - Single Shot Multibox Detector

Faster RCNN, YOLO, SSD

SSD Architecture



-
1. Multiple layers → handle diff scales
 2. Diff filter box

VGG

→ Feature Extraction has layers of CNN. Also for box prediction we have

$$\text{NN output} = (\beta_x \beta_y \beta_w \beta_h)$$

$$\text{Default box} = (dx \ dy \ dw \ dh)$$

$$\text{Prediction} = (bx \ by \ bw \ bh)$$

$$bx = dx + dw \beta_x$$

$$by = dy + dh \beta_y$$

$$bw = dw e^{\beta_w}$$

$$bh = dh e^{\beta_h}$$