

$2k$ classification

$4k$ regression

k anchor boxes

FC

intermediate layer

3×3 conv

sliding window

feature map

