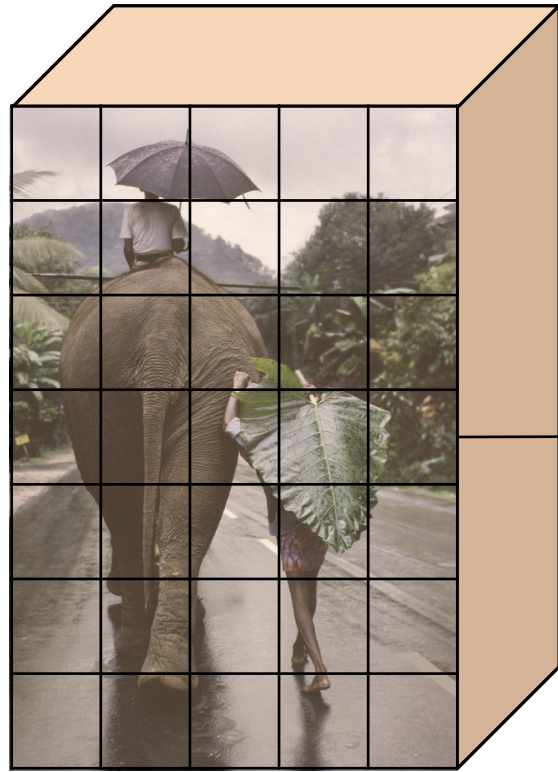




$$3 \times H \times W$$

Convolutional
feature
extractor



$$C_L \times H_L \times W_L$$

3x3 Conv,
 $C_{out}=256$,
 $S=1$, $P=1$
+ ReLU

1x1 Conv,
 $C_{out}=2k$,
 $S=1$, $P=0$

objectness
 $2k \times H_L \times W_L$

1x1 Conv,
 $C_{out}=4k$,
 $S=1$, $P=0$

correction
 $4k \times H_L \times W_L$