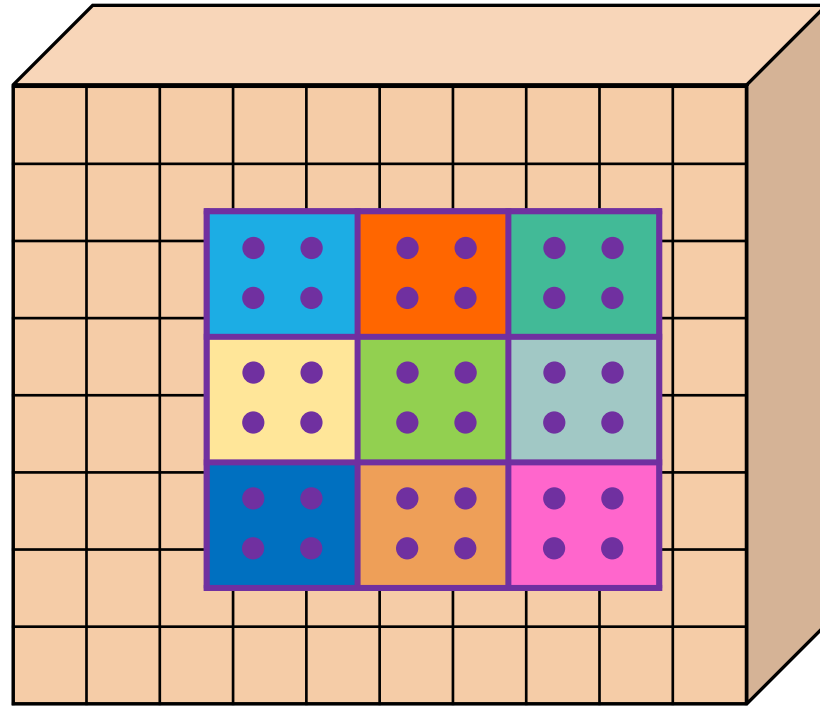


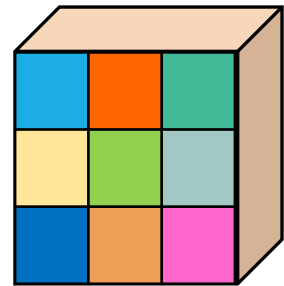
3x600x800

511x634

Convolutional  
feature  
extractor



256x8x10  
 $C_a \times H_a \times W_a$



256x3x3  
 $C_a \times H_o \times W_o$