



$3 \times H \times W$

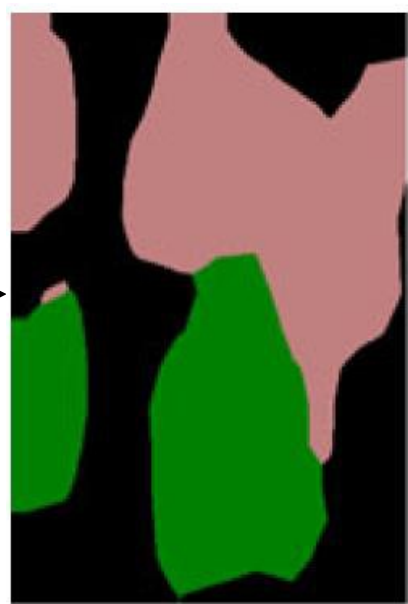
CNN  
backbone with  
total stride 32

$512 \times H/32 \times W/32$

1x1 Conv,  
 $C_{out} = 21$ ,  
 $S=1$ ,  $P=0$

$21 \times H/32 \times W/32$

Bilinear  
Upsampling x32



$21 \times H \times W$