

UCF:(with bn)

对于一个 $h \times w \times d$ 的层 (batchsize为 b) :

| | |
|---|------------------------|
| Input: Values of x over a mini-batch: $B = \{x_1, \dots, x_m\}$; | |
| Parameters to be learned: γ, β | |
| Output: $\{y_i = \text{BN}_{\gamma, \beta}(x_i)\}$ | |
| $\mu_B \leftarrow \frac{1}{m} \sum_{i=1}^m x_i$ | // mini-batch mean |
| $\sigma_B^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2$ | // mini-batch variance |
| $\hat{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}$ | // normalize |
| $y_i \leftarrow \gamma \hat{x}_i + \beta \equiv \text{BN}_{\gamma, \beta}(x_i)$ | // scale and shift |

求BN的时候对 $b \times d \times h \times w$ 的层，分别对 d 层每层的 $h \times w \times b$ 个点求均值和方差

bn的运算量为:

mean: $h \times w \times b \times d$

variance: $h \times w \times b \times d$

normalize: $h \times w \times d$

scale and shift: $h \times w \times d$

Totally: $h \times w \times d \times (b+1) \times 2$

参数量: $d \times 2$ (mean和var)

Conv1: input: $h \times w \times 3$ output: $h/2 \times w/2 \times 64$

cost: $h \times w \times 64 \times 3 \times 3 \times 3 + h \times w \times 64 \times 3 \times 3 \times 64 + h \times w \times 64 + h \times w \times 64 \times (b+1) \times 2 + h \times w \times 64 \times (b+1) \times 2 = 38656 \times h \times w + h \times w \times 256 \times (b+1) = (38912 + 256 \times b) \times h \times w$

Conv2: input: $h/2 \times w/2 \times 64$ output: $h/4 \times w/4 \times 128$

cost: $h/2 \times w/2 \times 128 \times 3 \times 3 \times 64 + h/2 \times w/2 \times 128 \times 3 \times 3 \times 128 + h/2 \times w/2 \times 128 + h/2 \times w/2 \times 128 \times (b+1) \times 2 + h/2 \times w/2 \times 128 \times (b+1) \times 2 = 55328 \times h \times w + 128 \times (b+1) \times h \times w$

Conv3: input: $h/4 \times w/4 \times 128$ output: $h/8 \times w/8 \times 256$

cost: $h/4 \times w/4 \times 256 \times 3 \times 3 \times 128 + h/4 \times w/4 \times 256 \times 3 \times 3 \times 256 \times 2 + h/4 \times w/4 \times 256 + h/4 \times w/4 \times 256 \times (b+1) \times 2 + h/4 \times w/4 \times 256 \times (b+1) \times 2 = 92176 \times h \times w + 96 \times (b+1) \times h \times w$

conv4: input: $h/8 \times w/8 \times 256$ output: $h/16 \times w/16 \times 512$

cost: $h/8 \times w/8 \times 512 \times 3 \times 3 \times 256 + h/8 \times w/8 \times 512 \times 3 \times 3 \times 512 \times 2 + h/8 \times w/8 \times 512 + h/8 \times w/8 \times 512 \times (b+1) \times 2 \times 3 = 92168 \times h \times w + 48 \times (b+1) \times h \times w$

Conv5: input: $h/16 \times w/16 \times 512$ output: $h/32 \times w/32 \times 512$

cost: $h/16 \times w/16 \times 512 \times 3 \times 3 \times 512 \times 3 + h/16 \times w/16 \times 512 + h/16 \times w/16 \times 512 \times (b+1) \times 2 \times 3 = 27650 \times h \times w + 12 \times (b+1) \times h \times w$

deconv5: input: $h/32 \times w/32 \times 512$ output: $h/16 \times w/16 \times 512$

cost: $h/16 \times w/16 \times 512 \times 8 \text{(bilinear interpolation)} + h/16 \times w/16 \times 512 \times 3 \times 3 \times 512 \times 3 + h/16 \times w/16 \times 512 \times (b+1) \times 2 \times 3 = 27664 \times h \times w + 12 \times (b+1) \times h \times w$

deconv4: input: $h/16 \times w/16 \times 512$ output: $h/8 \times w/8 \times 256$

cost: $\frac{h}{8} \cdot \frac{w}{8} \cdot 512 \cdot 8 + \frac{h}{8} \cdot \frac{w}{8} \cdot 512 \cdot 3 \cdot 3 \cdot 512 \cdot 2 + \frac{h}{8} \cdot \frac{w}{8} \cdot 256 \cdot 3 \cdot 3 \cdot 512 + \frac{h}{8} \cdot \frac{w}{8} \cdot 512 \cdot (b+1) \cdot 2 \cdot 3 = 92224 \cdot h \cdot w + 48 \cdot (b+1) \cdot h \cdot w$

deconv3: input: $\frac{h}{8} \cdot \frac{w}{8} \cdot 256$ output: $\frac{h}{4} \cdot \frac{w}{4} \cdot 128$

cost: $\frac{h}{4} \cdot \frac{w}{4} \cdot 256 \cdot 8 + \frac{h}{4} \cdot \frac{w}{4} \cdot 256 \cdot 3 \cdot 3 \cdot 256 \cdot 2 + \frac{h}{4} \cdot \frac{w}{4} \cdot 128 \cdot 3 \cdot 3 \cdot 256 + \frac{h}{4} \cdot \frac{w}{4} \cdot 256 \cdot (b+1) \cdot 2 \cdot 3 = 92288 \cdot h \cdot w + 96 \cdot (b+1) \cdot h \cdot w$

deconv2: input: $\frac{h}{4} \cdot \frac{w}{4} \cdot 128$ output: $\frac{h}{2} \cdot \frac{w}{2} \cdot 64$

cost: $\frac{h}{2} \cdot \frac{w}{2} \cdot 128 \cdot 8 + \frac{h}{2} \cdot \frac{w}{2} \cdot 128 \cdot 3 \cdot 3 \cdot 128 + \frac{h}{2} \cdot \frac{w}{2} \cdot 64 \cdot 3 \cdot 3 \cdot 128 + \frac{h}{2} \cdot \frac{w}{2} \cdot 128 \cdot (b+1) \cdot 2 \cdot 2 = 55552 \cdot h \cdot w + 128 \cdot (b+1) \cdot h \cdot w$

deconv1: input: $\frac{h}{2} \cdot \frac{w}{2} \cdot 64$ output: $h \cdot w \cdot 2$

cost: $h \cdot w \cdot 64 \cdot 8 + h \cdot w \cdot 64 \cdot 3 \cdot 3 \cdot 64 + h \cdot w \cdot 2 \cdot 3 \cdot 3 \cdot 64 + h \cdot w \cdot 64 \cdot (b+1) \cdot 2 \cdot 2 = 38528 \cdot h \cdot w + h \cdot w \cdot 256 \cdot (b+1)$

Totally:

Without BN: $38656 + 55328 + 92176 + 92168 + 27650 + 27664 + 92224 + 92288 + 55552 + 38528 = 612234 \cdot h \cdot w$

BN: $[h \cdot w \cdot 256 \cdot (b+1) + 128 \cdot (b+1) \cdot h \cdot w + 96 \cdot (b+1) \cdot h \cdot w + 48 \cdot (b+1) \cdot h \cdot w + 12 \cdot (b+1) \cdot h \cdot w] \cdot 2 = 1080 \cdot (b+1) \cdot h \cdot w$

Add Up:

$612234 + 1080 \cdot (b+1) = (613314 + 1080 \cdot b) \cdot h \cdot w$

参数量:

conv1: $3 \cdot 3 \cdot 3 \cdot 64 + 3 \cdot 3 \cdot 64 \cdot 64 = 38592$

conv2: $3 \cdot 3 \cdot 64 \cdot 128 + 3 \cdot 3 \cdot 128 \cdot 128 = 221184$

conv3: $3 \cdot 3 \cdot 128 \cdot 256 + 3 \cdot 3 \cdot 256 \cdot 256 + 3 \cdot 3 \cdot 256 \cdot 256 = 1474560$

conv4: $3 \cdot 3 \cdot 256 \cdot 512 + 3 \cdot 3 \cdot 512 \cdot 512 + 3 \cdot 3 \cdot 512 \cdot 512 = 5898240$

conv5: $3 \cdot 3 \cdot 512 \cdot 512 + 3 \cdot 3 \cdot 512 \cdot 512 + 3 \cdot 3 \cdot 512 \cdot 512 = 7077888$

deconv5: $3 \cdot 3 \cdot 512 \cdot 512 + 3 \cdot 3 \cdot 512 \cdot 512 + 3 \cdot 3 \cdot 512 \cdot 512 = 7077888$

deconv4: $3 \cdot 3 \cdot 512 \cdot 512 + 3 \cdot 3 \cdot 512 \cdot 512 + 3 \cdot 3 \cdot 512 \cdot 256 = 5898240$

deconv3: $3 \cdot 3 \cdot 256 \cdot 256 + 3 \cdot 3 \cdot 256 \cdot 256 + 3 \cdot 3 \cdot 256 \cdot 128 = 1474560$

deconv2: $3 \cdot 3 \cdot 128 \cdot 128 + 3 \cdot 3 \cdot 128 \cdot 64 = 221184$

deconv1: $3 \cdot 3 \cdot 64 \cdot 64 + 3 \cdot 3 \cdot 64 \cdot 2 = 38016$

BN: $2 \cdot [64 \cdot 2 + 128 \cdot 2 + 256 \cdot 3 + 512 \cdot 3 + 512 \cdot 3] = 8448$

Totally:

$38592 + 221184 + 1474560 + 5898240 + 7077888 + 7077888 + 5898240 + 1474560 + 221184 + 38016 + 8448 = 29420352 + 8448 = 29428800$

MDF: 假设最小的superpixel数为n, 二级superpixel数为n (每个一级superpixel对应一个二级superpixel), 并且总共有k个superpixel划分方案:

block1: $\frac{h}{4} \cdot \frac{w}{4} \cdot 96 \cdot 11 \cdot 11 \cdot 3 + \frac{h}{8} \cdot \frac{w}{8} \cdot 96 \cdot 3 \cdot 3 = 2191 \cdot h \cdot w$

block2: $\frac{h}{8} \cdot \frac{w}{8} \cdot 256 \cdot 5 \cdot 5 \cdot 96 + \frac{h}{16} \cdot \frac{w}{16} \cdot 256 \cdot 3 \cdot 3 = 9609 \cdot h \cdot w$

block3: $\frac{h}{16} \cdot \frac{w}{16} \cdot 384 \cdot 3 \cdot 3 \cdot 256 = 3456 \cdot h \cdot w$

block4: $\frac{h}{16} \cdot \frac{w}{16} \cdot 384 \cdot 3 \cdot 3 \cdot 384 = 5184 \cdot h \cdot w$

block5: $h/16*w/16*384*3*3*256 + h/32*w/32*256*3*3 = 3458*h*w$
fc6: $h/32*w/32*256*4096 = 1025*h*w$
fc7: $4096*4096 = 16777216$

一个Feature extractor:

$(2191 + 9609 + 3456 + 5184 + 3458 + 1025)*h*w + 16777216 = 24923*h*w + 16777216$

总共的extractor个数:

$(n+1)*(24923*h*w + 16777216)$

一、二级superpixel通过提取n次特征, 三级superpixel(原图)提取1次特征

NN_Layer1: $3*4096 \rightarrow 300$

$3*4096*300 = 3686400$

NN_Layer2: $300 \rightarrow 300$

$300*300 = 90000$

output: $300 \rightarrow 2$

600

NN_layer total:

$3686400 + 90000 + 600 = 3777000$

重复n次

Totally:

$((2n+1)*(24923*h*w + 16777216) + 3777000*n)*k + h*w*k = k*((49846*n + 24924)*h*w + 3777000*n)$

对于k个划分方案做fuse

参数量:

$96*3*11*11 + 256*96*5*5 + 384*256*3*3 + 384*384*3*3 + 384*256*3*3 + h/32*w/32*256*4096 + 4096*4096 + 3*4096*300 + 300*300 + 300*2 = 20523040 + 1024*h*w + 3777000 = 24300040 + 1024*h*w$

RFCN:

block1: $h*w*64*3*3*4 + h*w*64*3*3*64 + h*w*64 = 39232*h*w$

block2: $h/2*w/2*128*3*3*64 + h/2*w/2*128*3*3*128 + h/2*w/2*128 = 55328*h*w$

block3: $h/4*w/4*256*3*3*128 + h/4*w/4*256*3*3*256 + h/4*w/4*256*3*3*256 + h/4*w/4*256 = 92176*h*w$

block4: $h/8*w/8*512*3*3*256 + h/8*w/8*512*3*3*512 + h/8*w/8*512*3*3*512 + h/8*w/8*512 = 92168*h*w$

block5: $h/16*w/16*512*3*3*512 + h/16*w/16*512*3*3*512 + h/16*w/16*512*3*3*512 + h/16*w/16*512 = 27650*h*w$

Fc6: $h/32*w/32*4096*7*7*512 = 100352*h*w$

fc7: $h/32*w/32*4096*4096 = 16384*h*w$

score: $h/32*w/32*60*4096 = 240*h*w$

upscore2: $h/16*w/16*60*4*4*60 = 225*h*w$

Score pool4: $h/16*w/16*60*512 = 120*h*w$

fuse: $h/16*w/16*60 = 0.234375*h*w$

upsample_fused16: $h/8*w/8*60*4*4*60 = 900*h*w$

score_pool3: $h/8*w/8*60*256 = 240*h*w$

score_final: $h/8*w/8*60 = 0.9375*h*w$

bigscore: $h*w*2*16*16*60 = 30720*h*w$

Totally:

$$39232 + 55328 + 92176 + 92168 + 27650 + 100352 + 16384 + 240 + 225 + 120 + 0.234375 + 900 + 240 + 0.9375 + 30720 = 455736.171875 \cdot h \cdot w$$

参数量:

$$\begin{aligned} \text{block1: } & 64 \cdot 4 \cdot 3 \cdot 3 + 64 \cdot 64 \cdot 3 \cdot 3 = 39168 \\ \text{block2: } & 128 \cdot 64 \cdot 3 \cdot 3 + 128 \cdot 128 \cdot 3 \cdot 3 = 221184 \\ \text{block3: } & 256 \cdot 128 \cdot 3 \cdot 3 + 256 \cdot 256 \cdot 3 \cdot 3 + 256 \cdot 256 \cdot 3 \cdot 3 = 1474560 \\ \text{block4: } & 512 \cdot 256 \cdot 3 \cdot 3 + 512 \cdot 512 \cdot 3 \cdot 3 + 512 \cdot 512 \cdot 3 \cdot 3 = 5898240 \\ \text{block5: } & 512 \cdot 512 \cdot 3 \cdot 3 + 512 \cdot 512 \cdot 3 \cdot 3 + 512 \cdot 512 \cdot 3 \cdot 3 = 7077888 \\ \text{fc6: } & 512 \cdot 4096 \cdot 3 \cdot 3 = 18874368 \\ \text{fc7: } & 4096 \cdot 4096 = 16777216 \\ \text{score: } & 4096 \cdot 60 = 2457604 \\ \text{Upscore2: } & 60 \cdot 60 \cdot 4 \cdot 4 = 57600 \\ \text{scorepool4: } & 512 \cdot 60 = 30720 \\ \text{upsample_fused16: } & 60 \cdot 60 \cdot 4 \cdot 4 = 57600 \\ \text{bigscore: } & 60 \cdot 2 \cdot 16 \cdot 16 = 30720 \end{aligned}$$

Totally:

$$39168 + 221184 + 1474560 + 5898240 + 7077888 + 18874368 + 16777216 + 2457604 + 57600 + 30720 + 57600 + 30720 = 52996868$$

DS:

$$\begin{aligned} \text{block1: } & h \cdot w \cdot 64 \cdot 3 \cdot 3 \cdot 3 + h \cdot w \cdot 64 \cdot 3 \cdot 3 \cdot 64 + h/2 \cdot w/2 \cdot 64 \cdot 2 \cdot 2 = 38656 \cdot h \cdot w \\ \text{block2: } & h/2 \cdot w/2 \cdot 128 \cdot 3 \cdot 3 \cdot 64 + h/2 \cdot w/2 \cdot 128 \cdot 3 \cdot 3 \cdot 128 + h/4 \cdot w/4 \cdot 128 \cdot 2 \cdot 2 = 55328 \cdot h \cdot w \\ \text{block3: } & h/4 \cdot w/4 \cdot 256 \cdot 3 \cdot 3 \cdot 128 + h/4 \cdot w/4 \cdot 256 \cdot 3 \cdot 3 \cdot 256 + h/4 \cdot w/4 \cdot 256 \cdot 3 \cdot 3 \cdot 256 + h/4 \cdot w/4 \cdot 256 = 92176 \cdot h \cdot w \\ \text{block4: } & h/8 \cdot w/8 \cdot 512 \cdot 3 \cdot 3 \cdot 256 + h/8 \cdot w/8 \cdot 512 \cdot 3 \cdot 3 \cdot 512 + h/8 \cdot w/8 \cdot 512 \cdot 3 \cdot 3 \cdot 512 + h/8 \cdot w/8 \cdot 512 = 92168 \cdot h \cdot w \\ \text{block5: } & h/16 \cdot w/16 \cdot 512 \cdot 3 \cdot 3 \cdot 512 + h/16 \cdot w/16 \cdot 512 \cdot 3 \cdot 3 \cdot 512 + h/16 \cdot w/16 \cdot 512 \cdot 3 \cdot 3 \cdot 512 + h/16 \cdot w/16 \cdot 512 = 27650 \cdot h \cdot w \\ \text{conv6: } & h/32 \cdot w/32 \cdot 4096 \cdot 7 \cdot 7 \cdot 512 = 100352 \cdot h \cdot w \\ \text{conv7: } & h/32 \cdot w/32 \cdot 4096 \cdot 1 \cdot 1 \cdot 4096 = 16384 \cdot h \cdot w \\ \text{score: } & h/32 \cdot w/32 \cdot 1 \cdot 1 \cdot 1 \cdot 4096 = 4 \cdot h \cdot w \\ \text{deconv: } & h \cdot w \cdot 1 \cdot 63 \cdot 63 \cdot 1 = 3969 \cdot h \cdot w \end{aligned}$$

Totally:

$$38656 + 55328 + 92176 + 92168 + 27650 + 100352 + 16384 + 4 + 3969 = 426687 \cdot h \cdot w$$

参数量:

$$\begin{aligned} \text{block1: } & 64 \cdot 3 \cdot 3 \cdot 3 + 64 \cdot 64 \cdot 3 \cdot 3 = 38592 \\ \text{block2: } & 128 \cdot 64 \cdot 3 \cdot 3 + 128 \cdot 128 \cdot 3 \cdot 3 = 221184 \\ \text{block3: } & 256 \cdot 128 \cdot 3 \cdot 3 + 256 \cdot 256 \cdot 3 \cdot 3 + 256 \cdot 256 \cdot 3 \cdot 3 = 1474560 \\ \text{block4: } & 512 \cdot 256 \cdot 3 \cdot 3 + 512 \cdot 512 \cdot 3 \cdot 3 + 512 \cdot 512 \cdot 3 \cdot 3 = 5898240 \\ \text{block5: } & 512 \cdot 512 \cdot 3 \cdot 3 + 512 \cdot 512 \cdot 3 \cdot 3 + 512 \cdot 512 \cdot 3 \cdot 3 = 7077888 \\ \text{conv6: } & 4096 \cdot 512 \cdot 3 \cdot 3 = 18874368 \\ \text{conv7: } & 4096 \cdot 4096 \cdot 1 \cdot 1 = 16777216 \\ \text{score: } & 4096 \cdot 1 \cdot 1 \cdot 1 = 4096 \\ \text{deconv: } & 63 \end{aligned}$$

Totally:

$$38592 + 221184 + 1474560 + 5898240 + 7077888 + 18874368 + 16777216 + 4096 + 63 = 50366207$$

DCL:

data_conv: $h*w*128*3*3*3 = 3456*h*w$

data_fc: $h*w*128*128 = 16384*h*w$

data_ms_saliency: $h*w*128 = 128*h*w$

conv1_1: $h*w*64*3*3*3 = 1728*h*w$

conv1_2: $h*w*64*3*3*64 = 36864*h*w$

pool1: $h/2*w/2*64*3*3 = 144*h*w$

pool1_conv: $h/2*w/2*128*3*3*64 = 18432*h*w$

pool1_fc: $h/2*w/2*128*128 = 4096*h*w$

pool1_ms_saliency: $h/2*w/2*128 = 32*h*w$

conv2_1: $h/2*w/2*128*3*3*128 = 36864*h*w$

conv2_2: $h/2*w/2*128*3*3*128 = 36864$

pool2: $h/4*w/4*128*3*3 = 72*h*w$

pool2_conv: $h/4*w/4*128*3*3*128 = 9216*h*w$

pool2_fc: $h/4*w/4*128*128 = 1024*h*w$

pool2_ms_saliency: $h/4*w/4*128 = 8*h*w$

block3:

$h/4*w/4*256*3*3*128 + h/4*w/4*256*3*3*256 + h/4*w/4*256*3*3*256 + h/8*w/8*256*3*3 + h/8*w/8*128*3*3*256 + h/8*w/8*128*128 + h/8*w/8*128 = 97170*h*w$

block4:

$h/8*w/8*512*3*3*256 + h/8*w/8*512*3*3*512 + h/8*w/8*512*3*3*512 + h/8*w/8*512*3*3 + h/8*w/8*128*3*3*512 + h/8*w/8*128*128 + h/8*w/8*128 = 101706*h*w$

block5:

$h/8*w/8*512*5*5*512 + h/8*w/8*512*5*5*512 + h/8*w/8*512*5*5*512 + h/8*w/8*512*3*3 + h/8*w/8*4096*8*8*512 + h/8*w/8*4096*4096 + h/8*w/8*4096 = 2666632*h*w$

fuse:

$h*w*6$

Totally:

$3456 + 16384 + 128 + 1728 + 36864 + 144 + 18432 + 4096 + 32 + 36864 + 36864 + 72 + 9216 + 1024 + 8 + 97170 + 101706 + 2666632 + 6 = 3030826*h*w$

参数量:

data: $3*128*3*3 + 128*128*1*1 + 128*1*1*1 = 19968$

block1: $3*64*3*3 + 64*64*3*3 + 64*128*3*3 + 128*128*1*1 + 128*1*1*1 = 128832$

block2: $128*128*3*3 + 128*128*3*3 + 128*128*3*3 + 128*128*1*1 + 128*1*1*1 = 458880$

block3: $128*256*3*3 + 256*256*3*3 + 256*256*3*3 + 256*128*3*3 + 128*128*1*1 + 128*1*1*1 = 1785984$

block4: $256*512*3*3 + 512*512*3*3 + 512*512*3*3 + 512*128*3*3 + 128*128*1*1 + 128*1*1*1 = 6504576$

block5: $512*512*3*3 + 512*512*3*3 + 512*512*3*3 + 512*4096*4*4 + 4096*4096*1*1 + 4096*1*1*1 = 57413632$

Totally:

$19968 + 128832 + 458880 + 1785984 + 6504576 + 57413632 = 66311872$

Mobilenet + deeplabv3:

head_conv: $h/2*w/2*32*3*3 = 216*h*w$

block1: $456*h*w$

block2: $582*h*w + 513*h*w = 1095*h*w$

block3: $308*h*w + 219*h*w + 219*h*w = 746*h*w$

block4: $315*h*w + 822*h*w + 822*h*w + 822*h*w = 2781*h*w$

block5: $1014*h*w + 1809*h*w + 1809*h*w = 4632*h*w$

Block6: $2385*h*w + 4935*h*w + 4935*h*w = 12255*h*w$

block7: $7335*h*w$

ASPP:

Aspp0: $h/8*w/8*256*320 = 1280*h*w$

apss1, 2, 3: $h/8*w/8*256*320*3*3 = 34560*h*w$

Other:

GAP conv:

$256*1*1*320 = 81920$

Concate conv:

$h/8*w/8*256*1280 = 5120*h*w$

Low feature conv:

$h/4*w/4*48*24 = 72*h*w$

Concate conv:

$h/4*w/4*304*1 = 19*h*w$

Totally:

$(216 + 456 + 1095 + 746 + 2781 + 4632 + 12255 + 7335 + 1280 + 34560 + 5120 + 72 + 19)*h*w + 81920 = 70567*h*w + 81920$

参数量:

head_conv: $3*3*3*32 = 864$

block1: 1824

block2: $4704 + 8208 = 12912$

block3: $9360 + 14016 + 14016 = 37392$

block4: $20160 + 52608*3 = 177984$

block5: $64896 + 115776*2 = 296448$

block6: $152640 + 315840*2 = 784320$

block7: 469440

ASPP: $320*1*1*256 + 320*3*3*256*3 = 2293760$

Gap conv: $320*1*1*256 = 81920$

Conv: $1280*256*1*1 = 327680$

conv2: $24*48*1*1 = 1152$

last_conv: $304*1*1*1 = 304$

Totally:

$$864 + 1824 + 12912 + 37392 + 177984 + 296448 + 784320 + 469440 + 2293760 + 81920 + 327680 + 1152 + 304 = 4486000$$

UCF:

Conv1: input: $h*w*3$ output: $h/2*w/2*64$

$$\text{cost: } h*w*64*3*3*3 + h*w*64*3*3*64 + h*w*64 = 38656*h*w$$

Conv2: input: $h/2*w/2*64$ output: $h/4*w/4*128$

$$\text{cost: } h/2*w/2*128*3*3*64 + h/2*w/2*128*3*3*128 + h/2*w/2*128 = 55328*h*w$$

Conv3: input: $h/4*w/4*128$ output: $h/8*w/8*256$

$$\text{cost: } h/4*w/4*256*3*3*128 + h/4*w/4*256*3*3*256*2 + h/4*w/4*256 = 92176*h*w$$

conv4: input: $h/8*w/8*256$ output: $h/16*w/16*512$

$$\text{cost: } h/8*w/8*512*3*3*256 + h/8*w/8*512*3*3*512*2 + h/8*w/8*512 = 92168*h*w$$

Conv5: input: $h/16*w/16*512$ output: $h/32*w/32*512$

$$\text{cost: } h/16*w/16*512*3*3*512*3 + h/16*w/16*512 = 27650*h*w$$

deconv5: input: $h/32*w/32*512$ output: $h/16*w/16*512$

$$\text{cost: } h/16*w/16*512*8(\text{bilinear interpolation}) + h/16*w/16*512*3*3*512*3 = 27664*h*w$$

deconv4: input: $h/16*w/16*512$ output: $h/8*w/8*256$

$$\text{cost: } h/8*w/8*512*8 + h/8*w/8*512*3*3*512*2 + h/8*w/8*256*3*3*512 = 92224*h*w$$

deconv3: input: $h/8*w/8*256$ output: $h/4*w/4*128$

$$\text{cost: } h/4*w/4*256*8 + h/4*w/4*256*3*3*256*2 + h/4*w/4*128*3*3*256 = 92288*h*w$$

deconv2: input: $h/4*w/4*128$ output: $h/2*w/2*64$

$$\text{cost: } h/2*w/2*128*8 + h/2*w/2*128*3*3*128 + h/2*w/2*64*3*3*128 = 55552*h*w$$

deconv1: input: $h/2*w/2*64$ output: $h*w*2$

$$\text{cost: } h*w*64*8 + h*w*64*3*3*64 + h*w*2*3*3*64 = 38528*h*w$$

Totally:

$$38656 + 55328 + 92176 + 92168 + 27650 + 27664 + 92224 + 92288 + 55552 + 38528 = 612234*h*w$$

参数量:

$$\text{conv1: } 3*3*3*64 + 3*3*64*64 = 38592$$

$$\text{conv2: } 3*3*64*128 + 3*3*128*128 = 221184$$

$$\text{conv3: } 3*3*128*256 + 3*3*256*256 + 3*3*256*256 = 1474560$$

$$\text{conv4: } 3*3*256*512 + 3*3*512*512 + 3*3*512*512 = 5898240$$

$$\text{conv5: } 3*3*512*512 + 3*3*512*512 + 3*3*512*512 = 7077888$$

$$\text{deconv5: } 3*3*512*512 + 3*3*512*512 + 3*3*512*512 = 7077888$$

$$\text{deconv4: } 3*3*512*512 + 3*3*512*512 + 3*3*512*256 = 5898240$$

$$\text{deconv3: } 3*3*256*256 + 3*3*256*256 + 3*3*256*128 = 1474560$$

$$\text{deconv2: } 3*3*128*128 + 3*3*128*64 = 221184$$

$$\text{deconv1: } 3*3*64*64 + 3*3*64*2 = 38016$$

Totally:

$$38592 + 221184 + 1474560 + 5898240 + 7077888 + 7077888 + 5898240 + 1474560 + 221184 + 38016 = 29420352$$

PAGRN: input: $h*w*3$

Block1: input: $h*w*3$ output: $h/2*w/2*64$

$$\text{cost: } h*w*64*3*3*3 + h*w*64*3*3*64 + h*w*64 = 38656 * h*w$$

Block2: input: $h/2*w/2*64$ output: $h/4*w/4*128$

$$\text{cost: } h/2*w/2*128*3*3*64 + h/2*w/2*128*3*3*128 + h/2*w/2*128 = 55328 * h*w$$

Block3: input: $h/4*w/4*128$ output: $h/8*w/8*256$

$$\text{cost: } h/4*w/4*256*3*3*128 + h/4*w/4*256*3*3*256*3 + h/4*w/4*256 = 129040 * h*w$$

Block4: input: $h/8*w/8*256$ output: $h/16*w/16*512$

$$\text{cost: } h/8*w/8*512*3*3*256 + h/8*w/8*512*3*3*512*3 + h/8*w/8*512 = 129032 * h*w$$

Block5: input: $h/16*w/16*512$ output: $h/16*w/16*512$

$$\text{cost: } h/16*w/16*512*3*3*512*5 = 46080 * h*w$$

attention5:

$$h/16*w/16*512 + 1*1*512*512 + h/16*w/16*512 + h/16*w/16*512*512 + h/16*w/16*512 + h/8*w/8*512 + h/8*w/8*256*1*1*512 = 265230 * h*w$$

attention4:

$$h/8*w/8*256 + 1*1*256*256 + h/8*w/8*256 + h/8*w/8*256*256 + h/8*w/8*256 + h/4*w/4*256 + h/4*w/4*256*256 = 70684 * h*w$$

attention3:

$$h/4*w/4*256 + 1*1*256*256 + h/4*w/4*256 + h/4*w/4*256*256 + h/4*w/4*256 = 69680 * h*w$$

saliency_map:

$$h/4*w/4*1*1*256 = 16 * h*w$$

totally:

$$38656 * h*w + 55328 * h*w + 129040 * h*w + 129032 * h*w + 46080 * h*w + 265230 * h*w + 70684 * h*w + 69680 * h*w + 16 * h*w = 803746 * h*w$$

参数量:

$$\text{Block1: } 3*3*3*64 + 3*3*64*64 = 38592$$

$$\text{Block2: } 3*3*64*128 + 3*3*128*128 = 221184$$

Block3: $3*3*128*256 + 3*3*256*256 + 3*3*256*256 + 3*3*256*256 = 2064384$

Block4: $3*3*256*512 + 3*3*512*512 + 3*3*512*512 + 3*3*512*512 = 8257536$

Block5: $3*3*512*512 * 5 = 11796480$

attention5: $512*512*2 + 512*256 = 655360$

attention4: $256*256*3 = 65536$

attention3: $256*256*3 = 65536$

saliency_map = 256

totally: $38592 + 221184 + 2064384 + 8257536 + 11796480 + 655360 + 65536 + 65536 + 256 = 23164864$

NLDF: input: $354*354$

Conv1: input: $h*w*3$ output: $h/2*w/2*64$

cost: $h*w*64*3*3*3 + h*w*64*3*3*64 + h*w*64 = 38656*h*w$

Conv2: input: $h/2*w/2*64$ output: $h/4*w/4*128$

cost: $h/2*w/2*128*3*3*64 + h/2*w/2*128*3*3*128 + h/2*w/2*128 = 55328*h*w$

Conv3: input: $h/4*w/4*128$ output: $h/8*w/8*256$

cost: $h/4*w/4*256*3*3*128 + h/4*w/4*256*3*3*256*2 + h/4*w/4*256 = 92176*h*w$

conv4: input: $h/8*w/8*256$ output: $h/16*w/16*512$

cost: $h/8*w/8*512*3*3*256 + h/8*w/8*512*3*3*512*2 + h/8*w/8*512 = 92168*h*w$

Conv5: input: $h/16*w/16*512$ output: $h/32*w/32*512$

cost: $h/16*w/16*512*3*3*512*3 + h/16*w/16*512 = 27650*h*w$

Conv6: input: $h/2*w/2*64$ output: $h/2*w/2*128$

cost: $h/2*w/2*128*3*3*64 = 18432*h*w$

Conv7: input: $h/4*w/4*128$ output: $h/4*w/4*128$

cost: $h/4*w/4*128*3*3*128 = 9216*h*w$

Conv8: input: $h/8*w/8*256$ output: $h/8*w/8*128$

cost: $h/8*w/8*128*3*3*256 = 4608*h*w$

Conv9: input: $h/16*w/16*512$ output: $h/16*w/16*128$

Cost: $h/16*w/16*128*3*3*512 = 2304*h*w$

Conv10: input: $h/32*w/32*512$ output: $h/32*w/32*128$

cost: $h/32*w/32*128*3*3*512 = 576*h*w$

contrast1:

$h/2*w/2*128 + h/2*w/2*128*3*3 = 320*h*w$

contrast2:

$$h/4*w/4*128 + h/4*w/4*128*3*3 = 80*h*w$$

Contrast3:

$$h/8*w/8*128 + h/8*w/8*128*3*3 = 20*h*w$$

Contrast4:

$$h/16*w/16*128*(1+3*3) = 5*h*w$$

Contrast5:

$$h/32*w/32*128*(1+3*3) = 1.25*h*w$$

unpooling5: input: $h/32*w/32*(128+128)$ output: $h/16^2/16*128$

$$\text{cost: } h/16*w/16*128*5*5*256 = 3200*h*w$$

Unpooling4: input: $h/16*w/16*(128^2+128)$ output: $h/8*w/8*256$

$$\text{cost: } h/8*w/8*256*5*5*128*3 = 38400*h*w$$

Unpooling3: input: $h/8*w/8*(128^2+256)$ output: $h/4*w/4*384$

$$\text{cost: } h/4*w/4*384*5*5*128*4 = 307200*h*w$$

Unpooling2: input $h/4*w/4*(128^2+384)$ output: $h/2*w/2*512$

$$\text{cost: } h/2*w/2*512*5*5*128*5 = 2048000*h*w$$

Local: input: $h/2*w/2*(128^2+512)$ output: $h/2*w/2*640$

$$\text{Cost: } h/2*w/2*640*(128^2+512) = 122880*h*w$$

Local score: $h/2*w/2*640$ output: $h/2*w/2^2$

$$\text{cost: } h/2*w/2^2*640 = 320*h*w$$

Global: input: $h/32*w/32*512$ ($11*11*512$) output: $1*1*128$

$$11*11*512 \rightarrow 7*7*128 \rightarrow 3*3*128 \rightarrow 1*1*128$$

$$h/32*w/32*512 \rightarrow (h/32-4)*(w/32-4)*128 \rightarrow (h/32-8)*(w/32-8)*128 \rightarrow (h/32-10)*(w/32-10)*128$$

$$\text{Cost: } 7*7*128*5*5*512 + 3*3*128*5*5*128 + 1*1*128*3*3*128 = 84115456 \text{ (without } h*w)$$

$$\text{Cost: } (h/32-4)*(w/32-4)*128*5*5*512 + (h/32-8)*(w/32-8)*128*5*5*128 + (h/32-10)*(w/32-10)*128*3*3*128 = 2144*h*w - 204800*(h+w) + 26214400 - 102400*(h+w) + 26214400 - 46080*(h+w) + 14745600 = 2144*h*w - 353280*(h+w) + 67174400 \text{ (with } h*w)$$

Global score: input: $1*1*128$ output: $1*1*2$

$$\text{cost: } 1*1*2*1*128 = 256$$

add_score:

$$h/2*w/2^2 = 0.5*h*w$$

totally:

$$38656*h*w + 55328*h*w + 92176*h*w + 92168*h*w + 27650*h*w + 18432*h*w + 9216*h*w + 4608*h*w + 2304*h*w + 576*h*w + 320*h*w + 80*h*w + 20*h*w + 5*h*w + 1.25*h*w + 3200*h*w + 38400*h*w + 307200*h*w + 2048000*h*w + 122880*h*w + 320*h*w + 84115456 + 256 + 0.5*h*w = 2861540.75*h*w + 84115712 \text{ (without } h*w)$$

$$2861540.75*h*w + 256 + 2144*h*w - 353280*(h+w) + 67174400 = 2863684.75*h*w - 353280*(h+w) + 67174656 \text{ (with } h*w)$$

参数量:

$$\text{conv1: } 3*3*3*64 + 3*3*64*64 = 38592$$

conv2: $3*3*64*128 + 3*3*128*128 = 221184$
 conv3: $3*3*128*256 + 3*3*256*256 + 3*3*256*256 = 1474560$
 conv4: $3*3*256*512 + 3*3*512*512 + 3*3*512*512 = 5898240$
 conv5: $3*3*512*512 + 3*3*512*512 + 3*3*512*512 = 7077888$
 conv6: $3*3*64*128 = 73728$
 conv7: $3*3*128*128 = 147456$
 conv8: $3*3*256*128 = 294912$
 conv9: $3*3*512*128 = 589824$
 conv10: $3*3*512*128 = 589824$
 unpooling5: $5*5*256*128 = 819200$
 unpooling4: $5*5*384*256 = 2457600$
 unpooling3: $5*5*512*384 = 4915200$
 unpooling2: $5*5*640*512 = 8192000$
 local: $768*640 = 491520$
 local_score: $640*2 = 1280$
 global: $5*5*512*128 + 5*5*128*128 + 3*3*128*128 = 2195456$
 global_score = $128*2 = 256$

totally:

$38592 + 221184 + 1474560 + 5898240 + 7077888 + 73728 + 147456 + 294912 + 589824$
 $+ 589824 + 819200 + 2457600 + 4915200 + 8192000 + 491520 + 1280 + 2195456 + 256$
 $= 35478720$

DSS:

conv_1:

$h*w*64*3*3*3 + h*w*64*3*3*64 + h*w*64 = 38656*h*w$

conv_2:

$h/2*w/2*128*3*3*64 + h/2*w/2*128*3*3*128 + h/2*w/2*128 = 55328*h*w$

conv_3:

$h/4*w/4*256*3*3*128 + h/4*w/4*256*3*3*256*2 + h/4*w/4*256 = 92176*h*w$

conv_4:

$h/8*w/8*512*3*3*256 + h/8*w/8*512*3*3*512*2 + h/8*w/8*512 = 92168*h*w$

conv_5:

$h/16*w/16*512*3*3*512*3 + h/16*w/16*512 = 27650*h*w$

Conv1_2:

$h*w*128*3*3*64 + h*w*128*3*3*128 + h*w*1*1*128 = 221312*h*w$

conv2_2:

$h/2*w/2*128*3*3*128 + h/2*w/2*128*3*3*128 + h/2*w/2*1*1*128 = 73760*h*w$

Conv3_3:

$h/4*w/4*256*5*5*256 + h/4*w/4*256*5*5*256 + h/4*w/4*1*1*256 = 204816*h*w$

conv4_3:

$$h/8*w/8*256*5*5*512 + h/8*w/8*256*5*5*256 + h/8*w/8*1*256 = 76804*h*w$$

conv5_3:

$$h/16*w/16*512*5*5*512 + h/16*w/16*512*5*5*512 + h/16*w/16*1*512 = 51202*h*w$$

pool5:

$$h/32*w/32*512*7*7*512 + h/32*w/32*512*7*7*512 + h/32*w/32*1*512 = 25088*h*w$$

Short connection:

$$\text{conv4_3: } h/8*w/8*3$$

$$\text{conv3_3: } h/4*w/4*3$$

$$\text{conv2_2: } h/2*w/2*5$$

$$\text{conv1_2: } h*w*5$$

Totally:

$$38656 + 55328 + 92176 + 92168 + 27650 + 221312 + 73760 + 204816 + 76804 + 51202 + 25088 + 6.484375 = 984054.484375*h*w$$

参数量:

$$\text{conv1: } 3*3*3*64 + 3*3*64*64 = 38592$$

$$\text{conv2: } 3*3*64*128 + 3*3*128*128 = 221184$$

$$\text{conv3: } 3*3*128*256 + 3*3*256*256 + 3*3*256*256 = 1474560$$

$$\text{conv4: } 3*3*256*512 + 3*3*512*512 + 3*3*512*512 = 5898240$$

$$\text{conv5: } 3*3*512*512 + 3*3*512*512 + 3*3*512*512 = 7077888$$

$$\text{conv1_2: } 3*3*64*128 + 3*3*128*128 + 1*1*128*1 = 221312$$

$$\text{conv2_2: } 3*3*128*128 + 3*3*128*128 + 1*1*128*1 = 295040$$

$$\text{conv3_3: } 5*5*256*256 + 5*5*256*256 + 1*1*256*1 = 3277056$$

$$\text{conv4_3: } 5*5*512*256 + 5*5*256*256 + 1*1*256*1 = 4915456$$

$$\text{conv5_3: } 5*5*512*512 + 5*5*512*512 + 1*1*512*1 = 13107712$$

$$\text{pool5: } 7*7*512*512 + 7*7*512*512 + 1*1*512*1 = 25690624$$

short connection:

$$\text{conv4_3: } 1*1*3*1 = 3$$

$$\text{conv3_3: } 1*1*3*1 = 3$$

$$\text{conv2_2: } 1*1*5*1 = 5$$

$$\text{conv1_2: } 1*1*5*1 = 5$$

totally:

$$38592 + 221184 + 1474560 + 5898240 + 7077888 + 221312 + 295040 + 3277056 + 4915456 + 13107712 + 25690624 + 3 + 3 + 5 + 5 = 62217680$$