Input $n \times n \times Channel_{in}$

Filter $f \times f \times Channel_{in} \times Number$

Output $(n-f+1) \times (n-f+1) \times Channel_{out}$

(stride=1, padding=0)

$$\left\lfloor \frac{n-f+2p}{s} \right\rfloor + 1 = n-f+1$$

☆ 当 Number < $Channel_{in}$ 时，就实现了降维

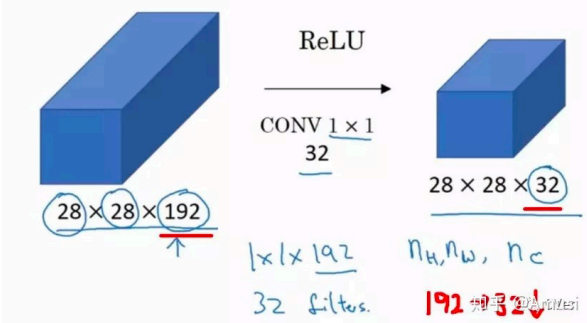
I. $6 \times 6 \times 32$

F. $1 \times 1 \times 32 \times 1$

O. $6 \times 6 \times 1$

降维

## Using 1×1 convolutions




$1 \times 1 \times 192$ $n_H, n_W, n_C$

32 filters. 192 ➡ 32 ✓

2*2

I. $28 \times 28 \times 192$

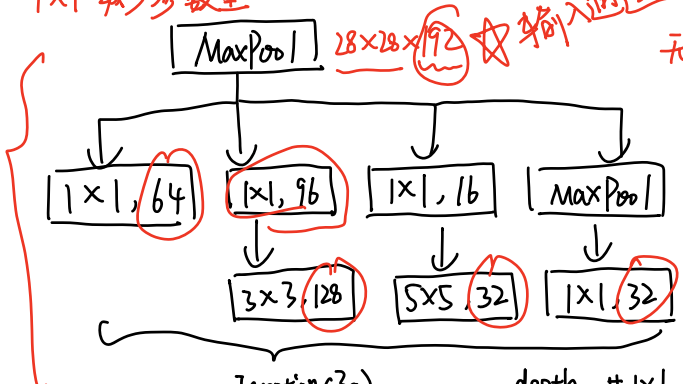
F. $1 \times 1 \times 192 \times 32$

O. $28 \times 28 \times 32$

192 => 32 ✓

channels out (f * f * channels in)

1×1 减少参数量




无1×1卷积

$64 \times 1 \times 1 \times 192 + 128 \times 3 \times 3 \times 192 + 32 \times 5 \times 5 \times 192 = 387072$

有1×1卷积

$64 \times 1 \times 1 \times 192 + 96 \times 1 \times 1 \times 192 + 128 \times 3 \times 3 \times 96 + 16 \times 1 \times 1 \times 192 + 32 \times 5 \times 5 \times 16 = 158048$

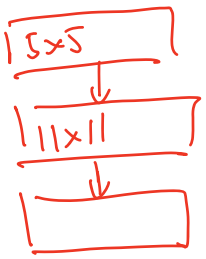
$\frac{158048}{387072} \approx 40\%$

| | Inception(3a) | | depth | #1×1 | | #3×3 | | #5×5 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| convolution | 7×7/2 | 112×112×64 | 1 | | | | | | | 2.7K | 34M |
| max pool | 3×3/2 | 56×56×64 | 0 | | | | | | | | |
| convolution | 3×3/1 | 56×56×192 | 2 | | 64 | 192 | | | | 112K | 360M |
| max pool | 3×3/2 | 28×28×192 | 0 | | | | | | | | |
| inception (3a) | | 28×28×256 | 2 | 64 | 96 | 128 | 16 | 32 | 32 | 159K | 128M |

☆ 最大池化层不影响通道数量，加上1×1卷积降维后可以进一步减少降维后的通道数量（通道数变化）

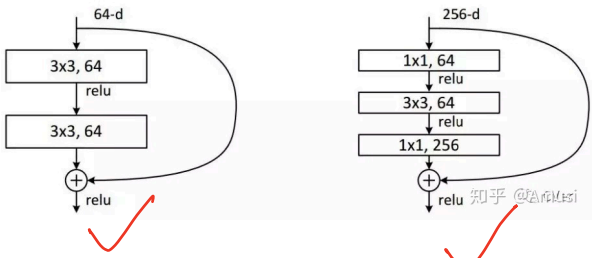
原：64 + 128 + 32 + 192 = 416

现：64 + 128 + 32 + 32 = 256




2012

在上述架构下，即使22层的GoogLeNet参数量也仅为8层AlexNet的$\frac{1}{12}$




resnet

不使用1×1：

256 × 3×3 × 256 × 2 = 1179648

使用1×1：

64 × 1×1 × 256 + 64 × 3×3 × 64 + 256 × 1×1 × 64

= 69632

$\frac{1179648}{69632} \approx 16.94$