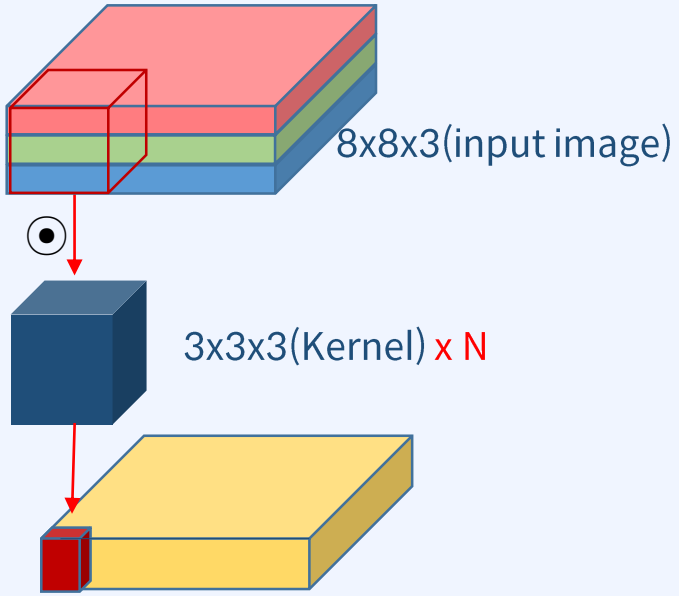
Part 3
Base model

Ch2. MobileNet v2

MobileNetV2: Inverted Residuals and Linear Bottlenecks

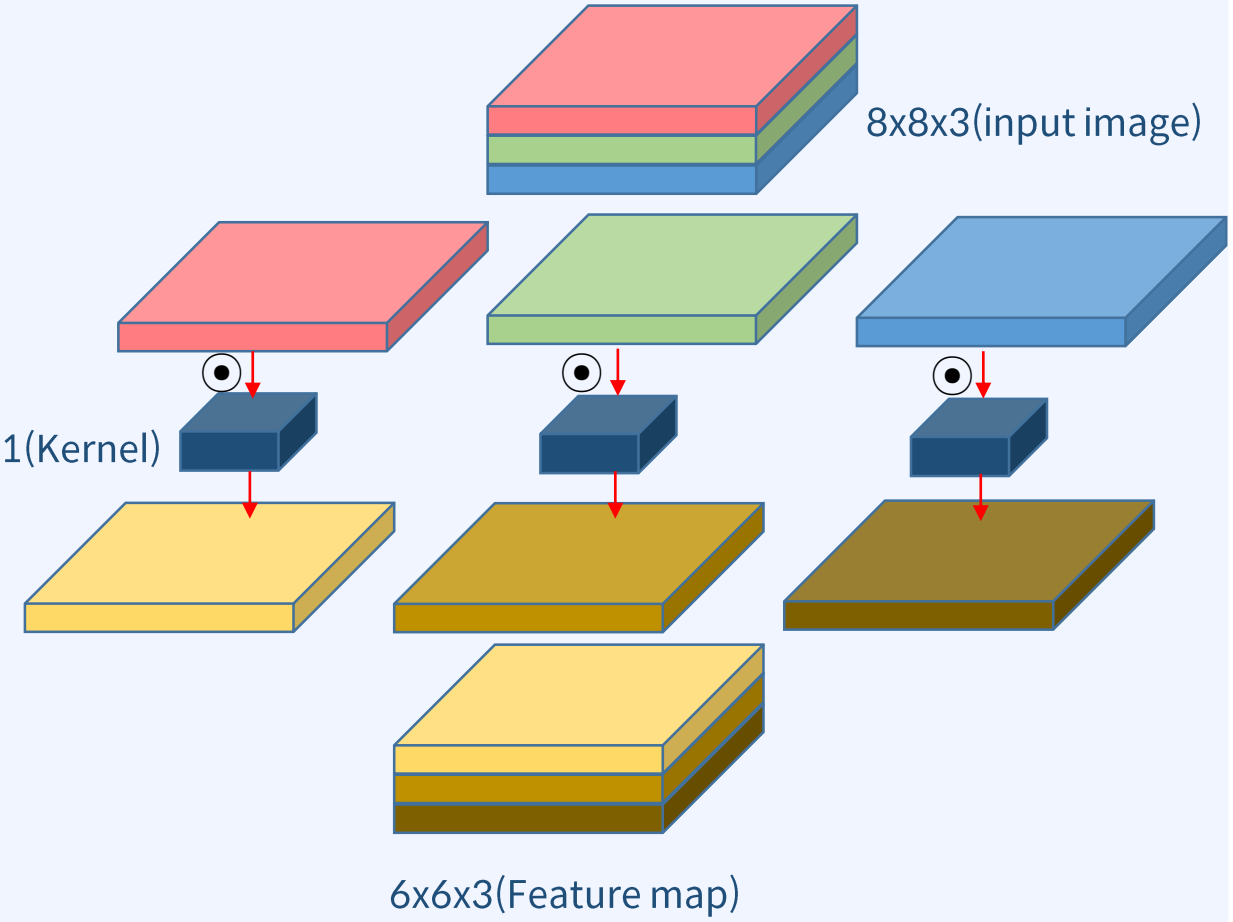
Standard & Depth-wise Convolution

Standard Convolution



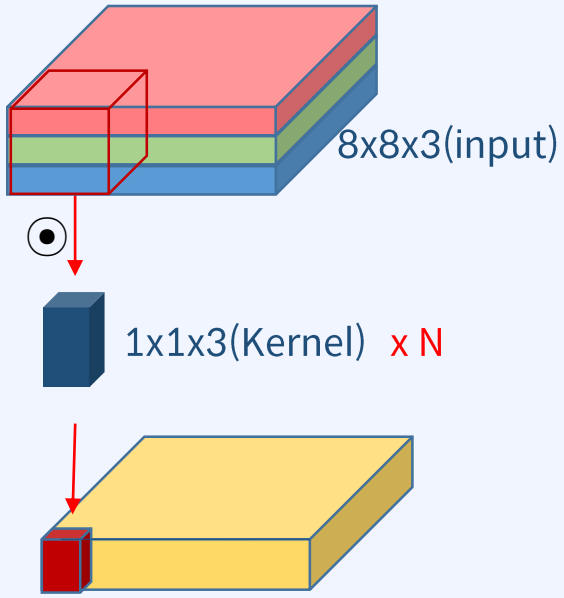
K = Kernel
C = in Channel

Depth-wise Convolution

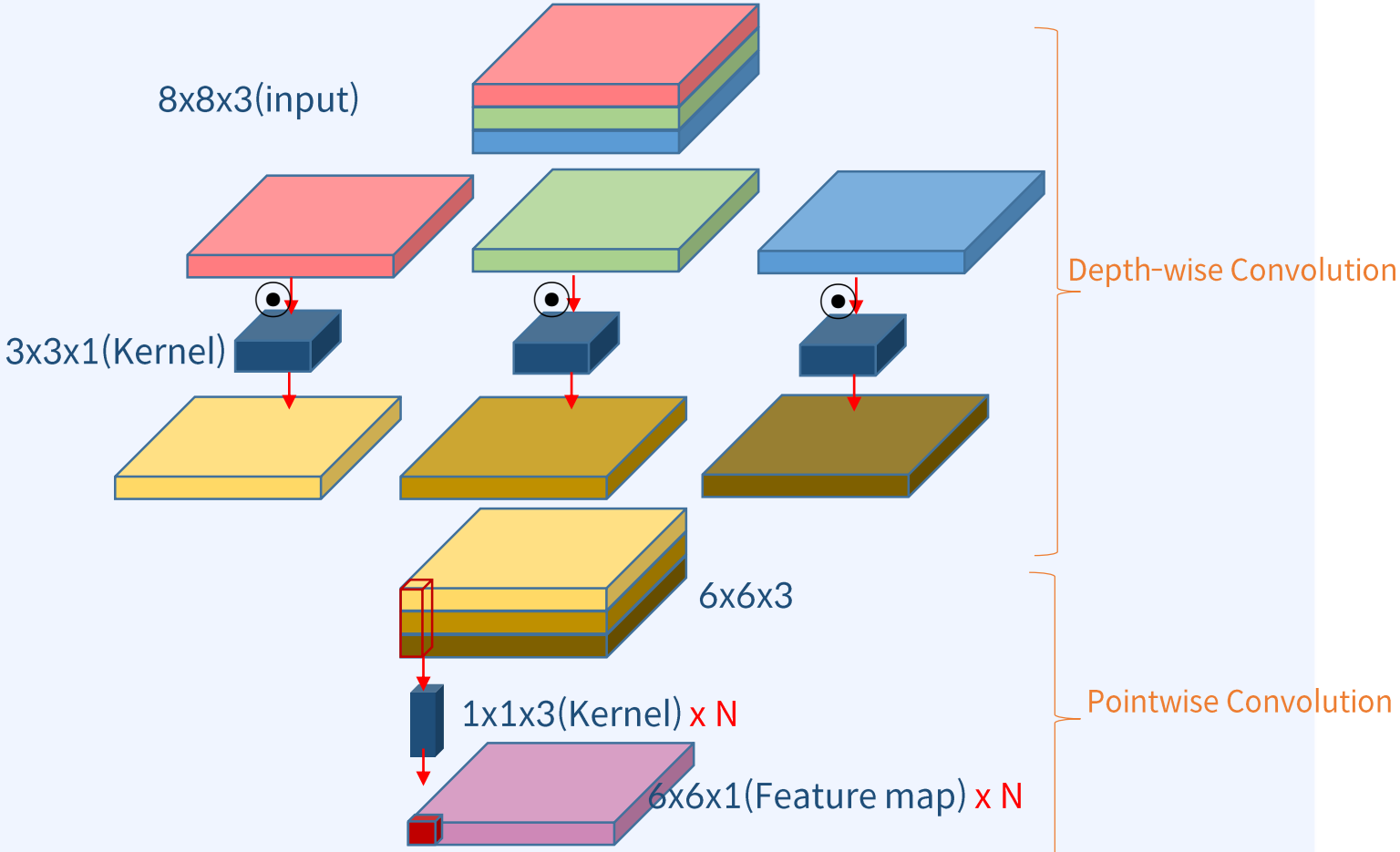


Depthwise Separable Convolutions

Pointwise Convolution



Depthwise **Separable** Convolution



Computation cost

- Standard convolution computational cost
  - D_k x D_k x M x N x D_F x D_F
- Depthwise separable convolution cost
  - D_k x D_k x M x D_F x D_F + M x N x D_F x D_F
- Reduction in computations
  - $$\frac{Depthwise\ sep.\ conv.\ cost}{std.\ conv.\ comp.\ cost} = \frac{1}{N} + \frac{1}{D_k^2}$$

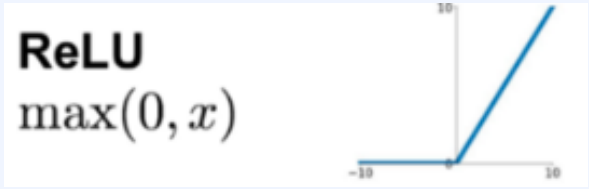
D_K: width/height of kernel
D_F: width/heifth of feature maps
M : number of input channels
N : number of output channels(number of kernel)

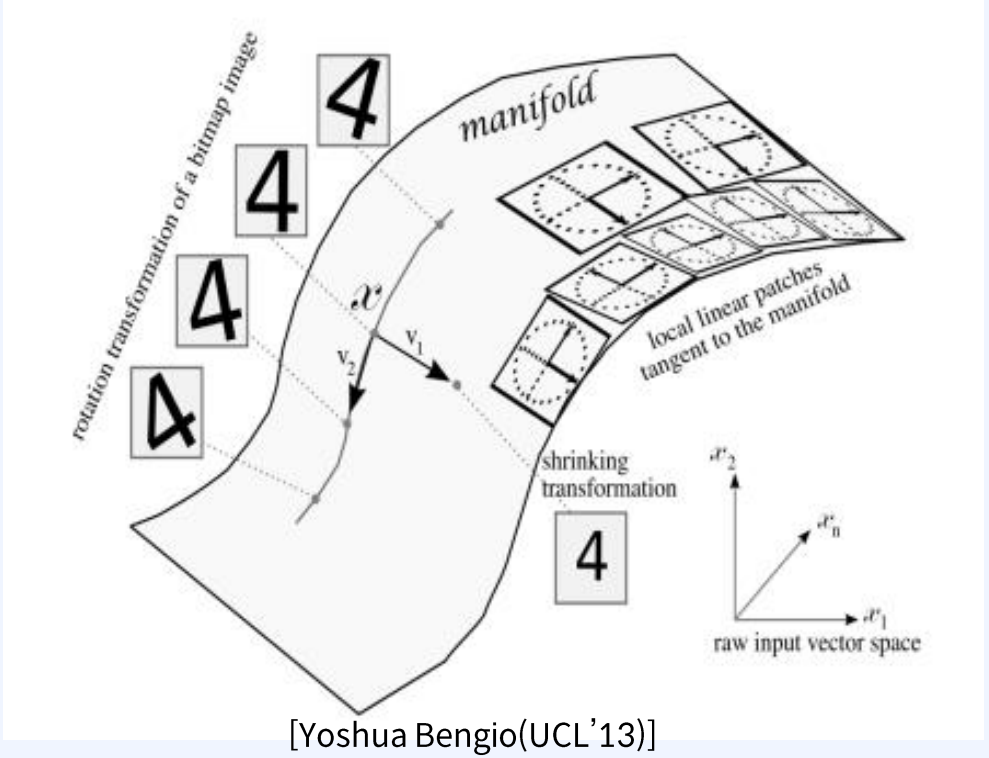
<앞의 예시에서 output channel N=3 인 경우>
Standard convolution computational cost
= 3x3x3x3x6x6 = 2916

Depthwise separable convolutions
= 3x3x3x6x6 + 3x3x6x6 = 1296

Linear Bottlenecks

- projection convolution
  - Linear bottleneck layer(Don't use ReLU)를 만들어서 차원은 줄이되 manifold 상의 중요한 정보들은 그대로 유지

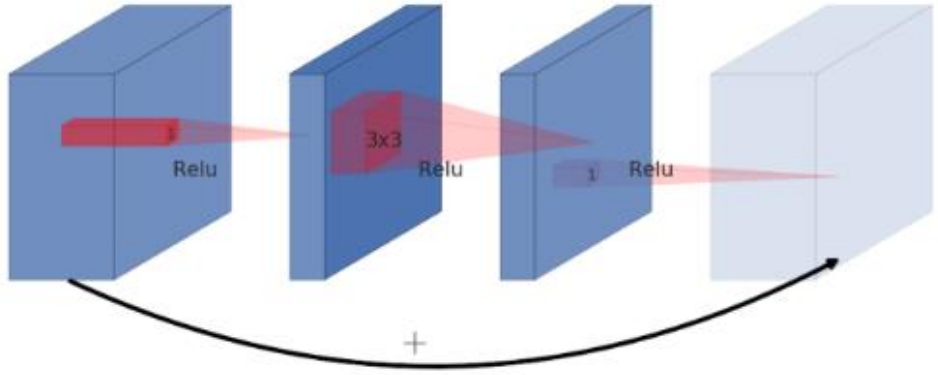




[Yoshua Bengio(UCL'13)]

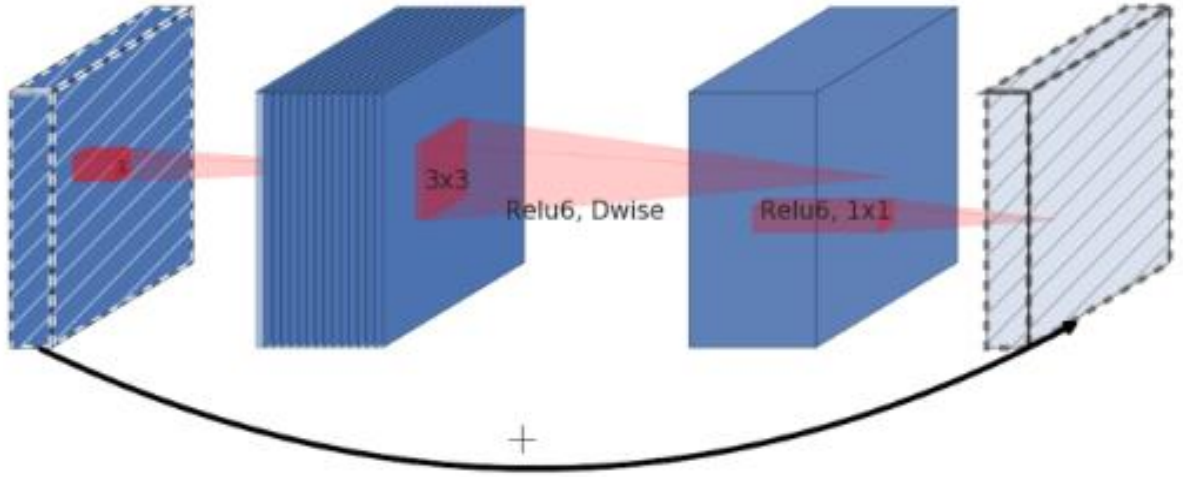
Inverted residuals

Residual block



- wide - narrow - wide 한 형태
- Network가 진행될수록 보통 channel수가 계속해서 증가
- 1x1 conv로 채널을 한번 줄임
  → 3x3 conv 연산→다시 원래의 채널로 돌려놓음
  → skip connection

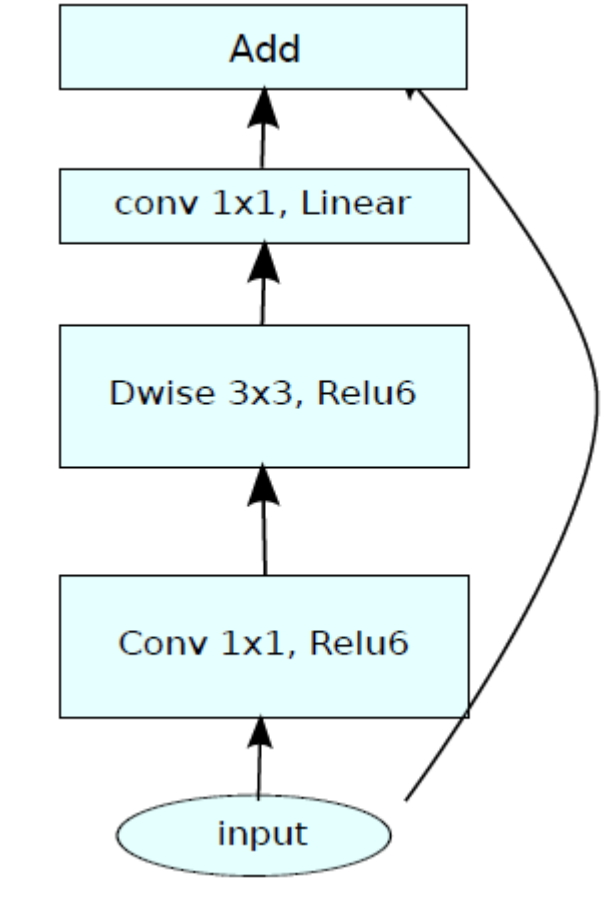
Inverted Residual block



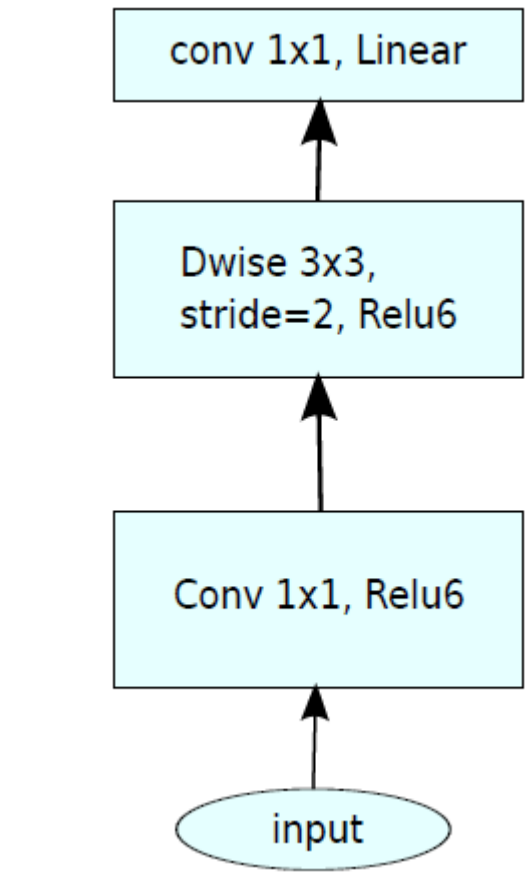
- narrow - wide - narrow한 형태
- narrow에 해당하는 저차원의 layer에는 필요한 정보만 압축되어서 저장되어 있다라는 가정
- skip connection으로 필요한 정보를 더 깊은 layer까지 전달
- 빗금 부분은 linear bottleneck을 의미(ReLU 사용x)
- 연산량 감소가 목적

MobileNet v2 blocks

Stride=1 block

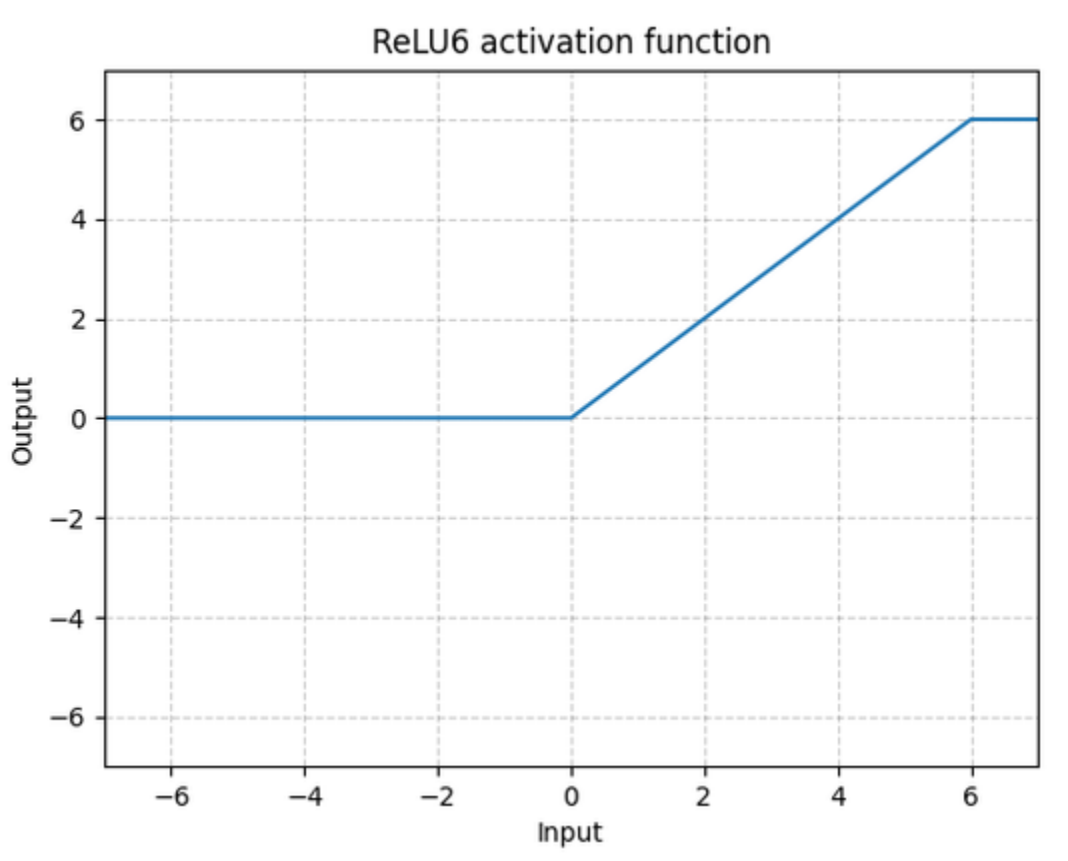


Stride=2 block



ReLU6

- $\mathrm{ReLU6}(x) = \min(\max(0, x), 6)$



Model Architecture

- t=expansion factor
- c=channel
- n=iteration
- s=stride

Input	Operator	t	c	n	s
$224^2 \times 3$	conv2d	-	32	1	2
$112^2 \times 32$	bottleneck	1	16	1	1
$112^2 \times 16$	bottleneck	6	24	2	2
$56^2 \times 24$	bottleneck	6	32	3	2
$28^2 \times 32$	bottleneck	6	64	4	2
$14^2 \times 64$	bottleneck	6	96	3	1
$14^2 \times 96$	bottleneck	6	160	3	2
$7^2 \times 160$	bottleneck	6	320	1	1
$7^2 \times 320$	conv2d 1x1	-	1280	1	1
$7^2 \times 1280$	avgpool 7x7	-	-	1	-
$1 \times 1 \times 1280$	conv2d 1x1	-	k	-	

Summary

- Depthwise separable convolution
- Inverted residual
- Linear Bottleneck
- ReLU6