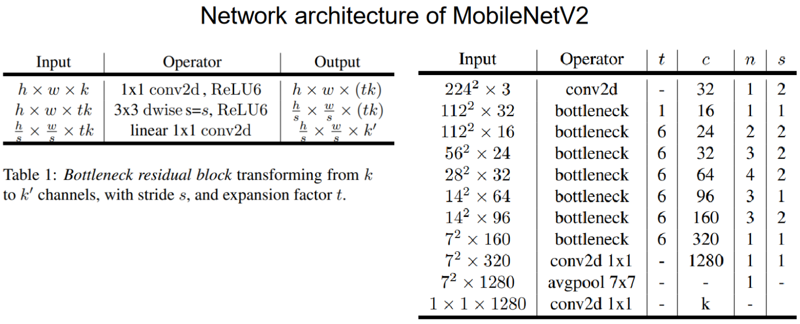
高等積體電路設計 Advanced VLSI **HW#1**

電子所 R10943170 連德宇

1. (40%)

   Calculate the **operation number** of each layer of **MobileNetV2** on ImageNet (1000 classes), including the *conventional convolutional* layer, *depth-wise convolutional* layer, *point-wise convolutional* layer, and *fully-connected* layer.



Network architecture of MobileNetV2

Input	Operator	Output
$h \times w \times k$	1x1 conv2d , ReLU6	$h \times w \times (tk)$
$h \times w \times tk$	3x3 dwise s=s, ReLU6	$\frac{h}{s} \times \frac{w}{s} \times (tk)$
$\frac{h}{s} \times \frac{w}{s} \times tk$	linear 1x1 conv2d	$\frac{h}{s} \times \frac{w}{s} \times k'$

Table 1: *Bottleneck residual block* transforming from k to k' channels, with stride s, and expansion factor t.

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	-	

Sol:

根據講義 p55，

Conventional convolution layer's operation number = $[(Ci \times kw \times kh)+(Ci \times kw \times kh-1)] \times Co \times Wo \times Ho$

$$\cong k^2 \times Ci \times Co \times (output\ size)^2$$

Bottleneck:

Expand layer's operation number = $[Ci + (Ci - 1)] \times Co \times (Wo \times Ho)^2$

$$\cong Ci^2 \times t \times (input\ size)^2$$

Depth-wise convolution layer's operation number = $[(kw \times kh) + (kw \times kh - 1)] \times Ci \times t \times Wo \times Ho$

$$\cong k^2 \times Ci \times t \times (\frac{input\ size}{s})^2$$

Point-wise convolution layer's operation number = $[Ci \times t + (Ci \times t - 1)] \times Co \times Wo \times Ho$

$$\cong Ci \times t \times Co \times (\frac{input\ size}{s})^2$$

the total of the Bottleneck is $\cong [Ci \times t \times (input\ size)^2] \times (Ci + \left(\frac{k^2}{s^2}\right) + \left(\frac{Co}{s^2}\right))$

Layer:

1: $9 * 3 * 32 * 112^2 = 10838016$

2: $[32 \times 1 \times (112)^2] \times \left(32 + \left(\frac{3^2}{1^2}\right) + \left(\frac{16}{1^2}\right)\right) = 22880256$

3: $[16 \times 6 \times (112)^2] \times (16 + \left(\frac{3^2}{2^2}\right) + \left(\frac{24}{2^2}\right)) + 56^2 \times 6 \times 24 \times (24 + 9 + 24) = 54942720$

4: $[24 \times 6 \times (56)^2] \times (24 + \left(\frac{3^2}{2^2}\right) + \left(\frac{32}{2^2}\right)) + 28^2 \times 6 \times 32 \times (32 + 9 + 32) \times 2 = 37443840$

5: $[32 \times 6 \times (28)^2] \times (32 + \left(\frac{3^2}{2^2}\right) + \left(\frac{64}{2^2}\right)) + 14^2 \times 6 \times 64 \times (64 + 9 + 64) \times 3 = 38497536$

6: $[64 \times 6 \times (14)^2] \times (64 + \left(\frac{3^2}{1^2}\right) + \left(\frac{96}{1^2}\right)) + 14^2 \times 6 \times 96 \times (96 + 9 + 96) \times 2 = 58103808$

7: $[96 \times 6 \times (14)^2] \times (96 + \left(\frac{3^2}{2^2}\right) + \left(\frac{160}{2^2}\right)) + 7^2 \times 6 \times 160 \times (160 + 9 + 160) \times 2 = 46560192$

8: $[160 \times 6 \times (7)^2] \times (160 + \left(\frac{3^2}{1^2}\right) + \left(\frac{320}{1^2}\right)) = 23002560$

9: $1^2 \times 320 \times 1280 \times 7^2 = 20070400$

10: $avgpool$

11: $1^2 \times 1280 \times 1000 \times 1^2 = 1280000$

12: fully-connected $1000 \times 1000 = 1000000$

Total = 314619328 operations

2. Now we compare the arithmetic intensity of the the networks.

(a) (10%+10%+10%+10%)

Compare the arithmetic intensity of the *conventional convolutional* layer, *depth-wise convolutional* layer, *point-wise convolutional* layer, and *fully-connected* layer.

∗ Please represent the answer by: C_{in}, C_{out}, W_{in}, W_{out}, H_{in}, H_{out}, k_w, k_h
C_{in} denotes the input depth (or input channel); C_{out} denotes the kernel number (or output channel); $W_{in/out}$ denotes the input/output image width; $H_{in/out}$ denotes the input/output image heights. k_w and k_h denote the width/height of the kernel.

∗ We supposed **all kinds of convolutional layers have the same shape**. That is, C_{in}, C_{out}, W_{in}, H_{in}, W_{out}, and H_{out} are identical. Conventional convolutional layer and depth-wise convolutional layer have the same k_w and k_h. For the fully connected layer, the input dimension is $\mathrm{I} = C_{in} \times W_{in} \times H_{in}$, and the output dimension is $\mathrm{O} = C_{out} \times W_{out} \times H_{out}$.

(b) (20%)

Following (a), based on the comparison result, explain the mechanism to raise the arithmetic intensity.

Sol:

(a)

	Parameters	Operations	Arithmetic intensity
conventional convolutional layer	$kw*kh* Cin*Cout$	$kw*kh*Hout*Wout*Cin*Cout$	$Hout*Wout$
depth-wise convolutional layer	$kw*kh* Cin$	$kw*kh*Hout*Wout*Cin$	$Hout*Wout$
point-wise convolutional layer	$Cin*Cout$	$Cout *Hout*Wout*Cin$	$Hout*Wout$
fully-connected layer	$Cin*Hin*Win*Cout*Hout*Wout$	$Cin*Hin*Win*Cout*Hout*Wout$	1

(b)

For the purpose of increase the arithmetic intensity, we should use conventional convolutional layer, depth-wise convolutional layer and point-wise convolutional layer as more as possible rather than use the fully-connected layer.