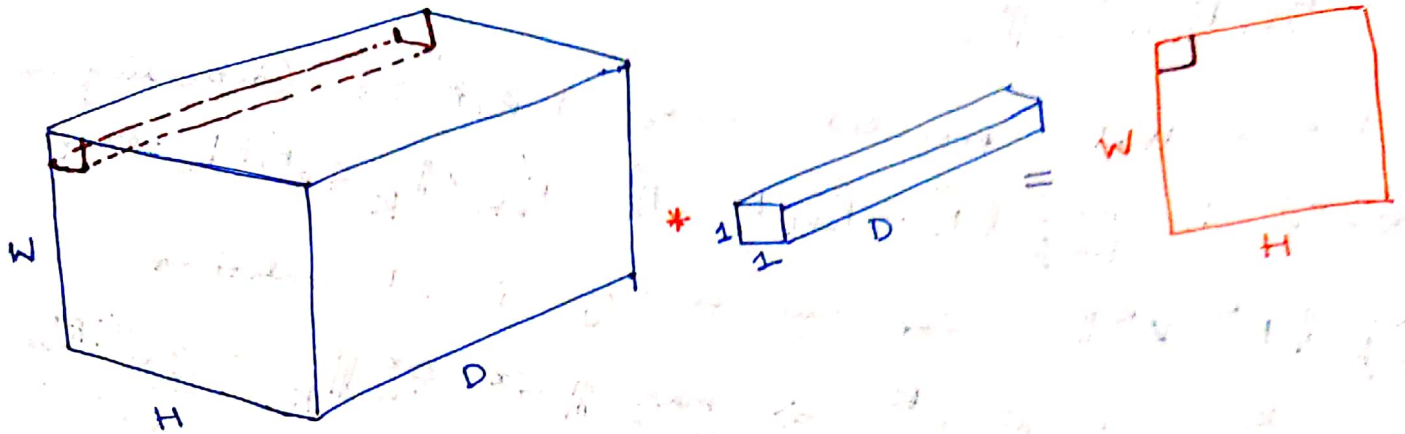## Pointwise Convolutions:
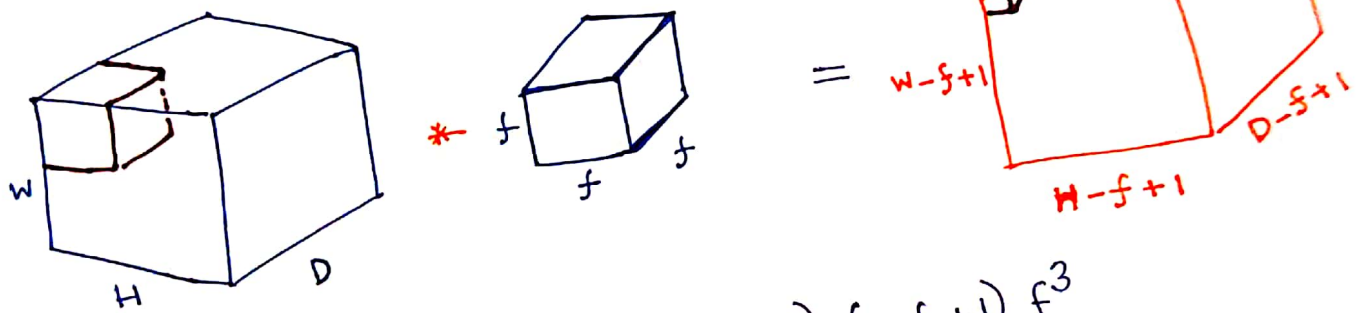

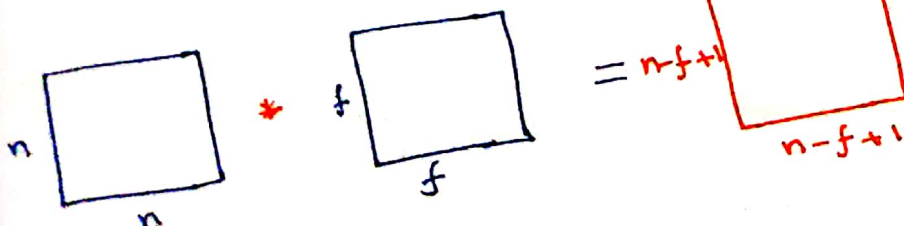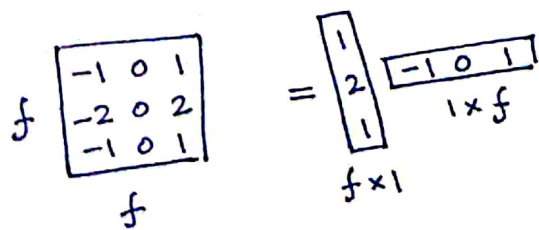Operation count: $W \times H \times D$
(multiplication)

3D Convolutions:

Operation Count: $(W-f+1)(H-f+1)(D-f+1) f^3$

Separable Convolutions:

$$f \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$$

$f \times f \qquad f \times 1 \qquad 1 \times f$

operation count:
$(n-f+1) \times (n-f+1) \times f^2$

$$n \times n \quad * \quad n \times 1 \text{ (f)} \quad = \quad (n-f+1) \times n \quad * \quad 1 \times f$$

Operation count: $(n-f+1) \times n \times f$

$$= \quad (n-f+1) \times (n-f+1)$$

$$+ \; (n-f+1)^2 \times f$$

Comparison: $\dfrac{\text{separable}}{\text{regular}} = \dfrac{(n-f+1)\, f\, (2n-f+1)}{(n-f+1)^2\, f^2} = \dfrac{2n-f+1}{(n-f+1)\, f}$

Assuming $n \gg f$, $\dfrac{\text{separable}}{\text{regular}} \approx \dfrac{2}{f}$ ($f=3 \Rightarrow \frac{1}{3}$ reduction in computation)

Depthwise Separable Convolutions:

Standard Convolution: (Image processing)

$$W \times H \times 3 \quad * \quad f \times f \times 3 \quad \times N \quad = \quad (W-f+1) \times (H-f+1) \times N$$

operation count: $(W-f+1)(H-f+1) N \times 3f^2$

step-1 Depthwise Convolution:

$$W \times H \times 3 \quad * \quad f \times f \times 1 \quad = \quad (W-f+1) \times (H-f+1) \times 3$$

Operation count: $3(W-f+1)(H-f+1) f^2$

Scanned with CamScanner

step-2

Pointwise Convolution :
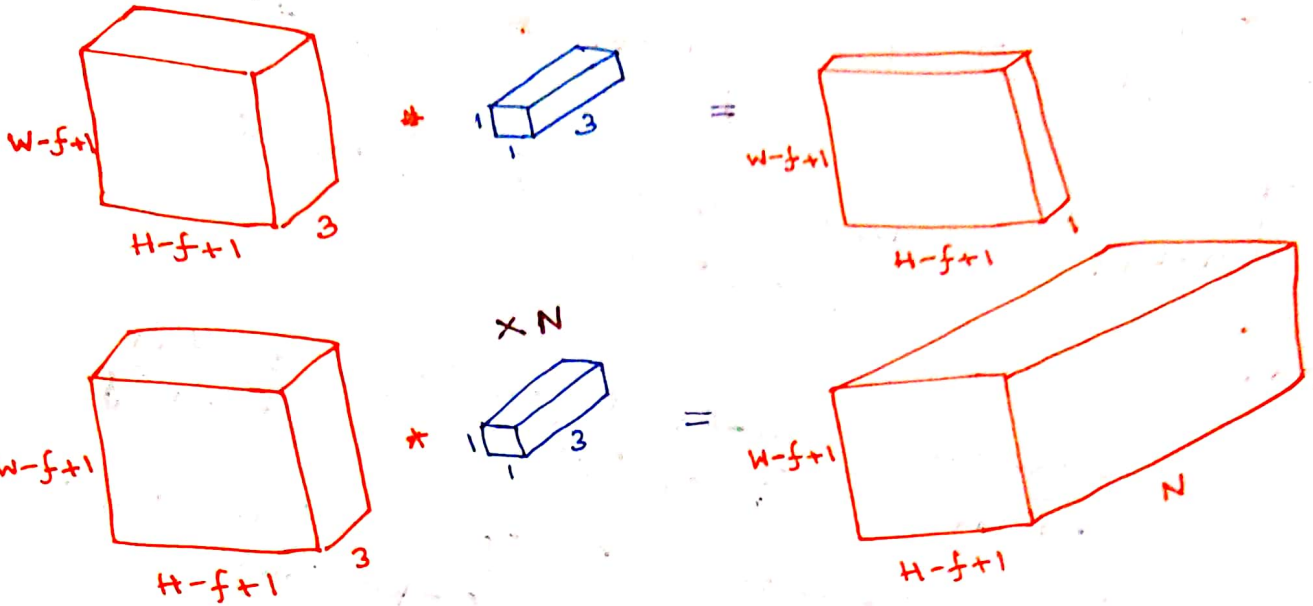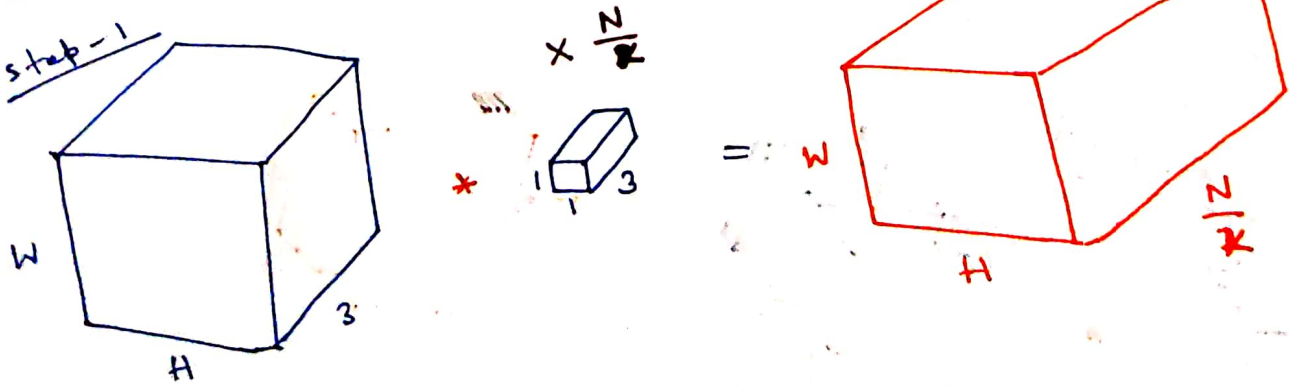

operation Count : $3(W-f+1)(H-f+1)N$

Total Operation Count : $3(W-f+1)(H-f+1)(f^2+N)$

$$\frac{Depth\ Sep}{Regular} = \frac{3(W-f+1)(H-f+1)(f^2+N)}{3(W-f+1)(H-f+1)\,f^2N} = \frac{1}{N}+\frac{1}{f^2}$$

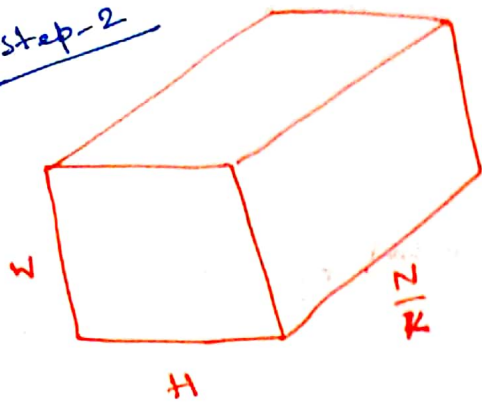reduction in computation

$$N=100, \quad f=3 \quad \Rightarrow \quad \frac{1}{8} \text{ reduction in computation}$$

Modified Depthwise Separable Convolutions :

step-1

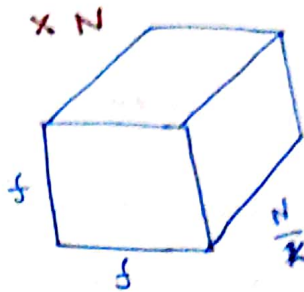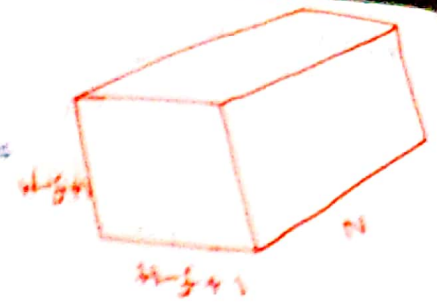
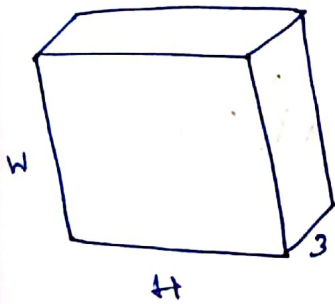operation count : $\frac{3N}{K}\,WH$

<u>step-2</u>

bottleneck layer

operation count: $(W-f+1)(H-f+1) N \times \dfrac{f^2 N}{K}$
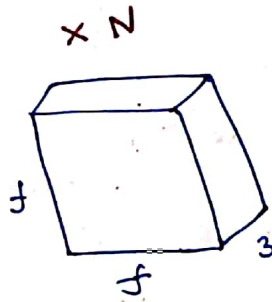
Total operation count: (To simplify computation assume same padding for each step) $WH \times \left(\dfrac{3N}{K} + \dfrac{f^2 N^2}{K}\right) = \dfrac{WHN}{K}\left(3 + \dfrac{f^2 N}{W H}\right)$
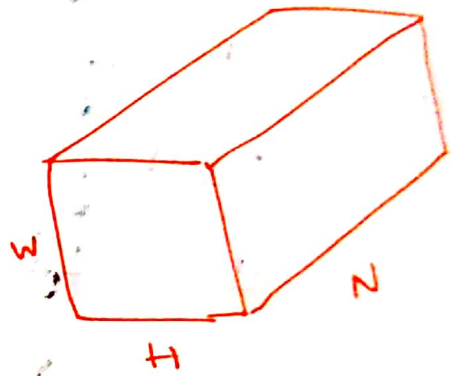
<u>Regular:</u>

operation count: $WHN \times 3f^2$

$\dfrac{\text{Mod dep sep}}{\text{Regular}} = \dfrac{\dfrac{3}{K} + \dfrac{f^2 N}{K}}{3f^2} = \dfrac{1}{Kf^2} + \dfrac{N}{3K}$

XCEPTION:

$299 \times 299 \times 3$ image

Conv 32, 3×3, padding = 1, stride = 2×2

output size = $\left\lfloor \dfrac{299 - 3 + 2 \times 1}{2} + 1 \right\rfloor = 150$

$150 \times 150 \times 32$

Relu
Conv 64, 3×3,

output size = $150 \times 150 \times 64$

Relu

Dep Sep Conv 128, 3×3

output size = $150 \times 150 \times 128$

Relu

Depsep Conv 128, 3×3

Maxpooling 3×3, stride = 2×2

output size = $\left\lfloor \dfrac{150 - 3 + 2 \times 1}{2} + 1 \right\rfloor = 75$

Conv 1×1
stride = 2×2
padding = 0

output size
= $\left\lfloor \dfrac{150-1}{2} + 1 \right\rfloor$
= 75
$75 \times 75 \times 128$

$75 \times 75 \times 128$

+

DepSep Conv 256, 3×3

Relu

DepSepConv 256, 3×3

Conv 1×1
stride = 2×2
padding = 0

Maxpooling 3×3, stride = 2×2

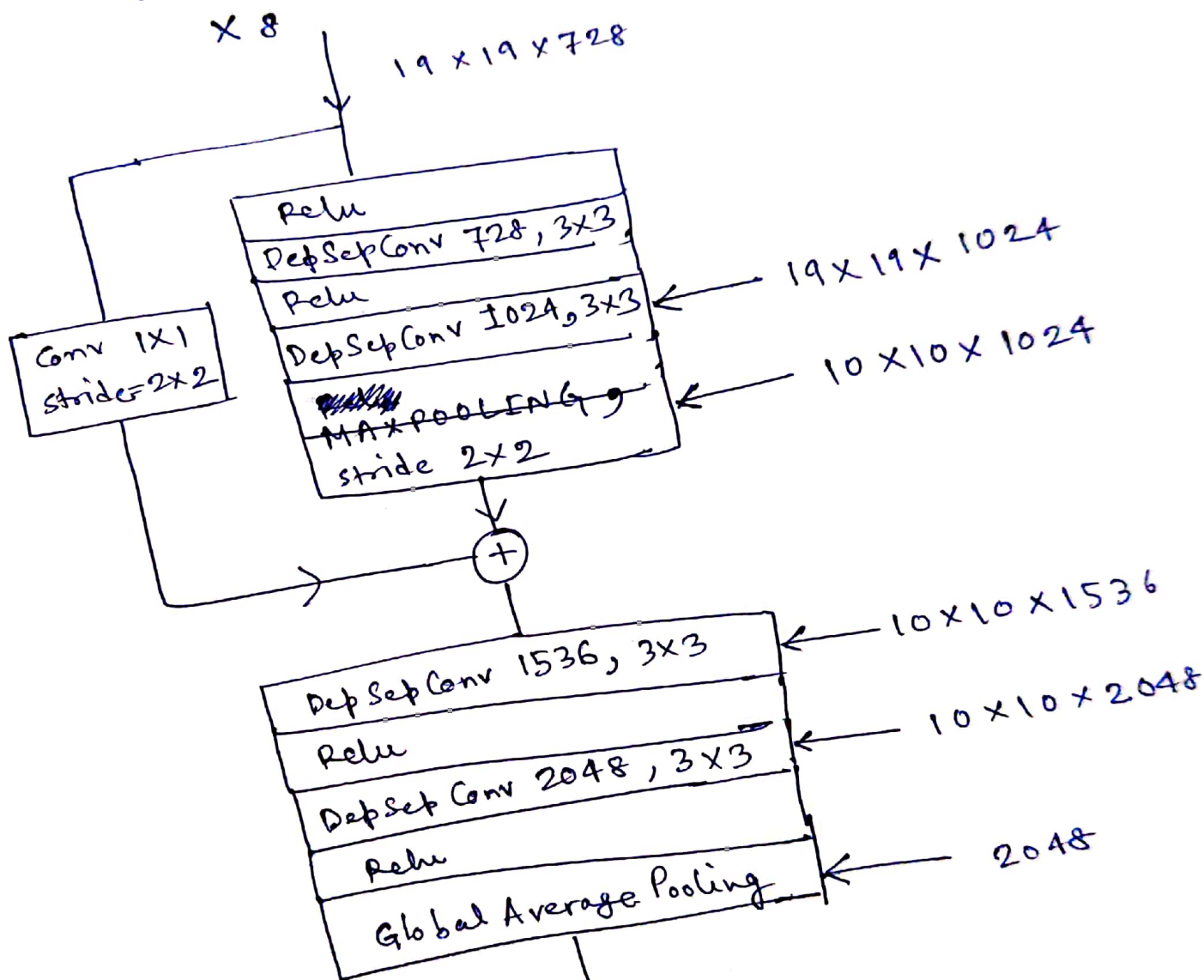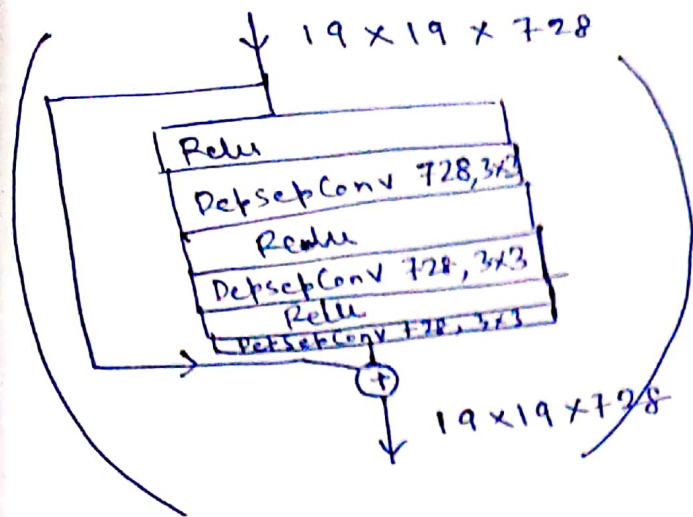output size = $\left\lfloor \dfrac{75 - 3 + 2 \times 1}{2} + 1 \right\rfloor = 38$

+

$38 \times 38 \times 256$

Conv 1×1
stride = 2×2
padding = 0

" , 728

+

output size = $\left\lfloor \dfrac{38 - 3 + 2 \times 1}{2} + 1 \right\rfloor = 19$

$19 \times 19 \times 728$

19 × 19 × 728

Relu
DepSepConv 728, 3×3
Relu
DepSepConv 728, 3×3
Relu
DepSepConv 728, 3×3

19 × 19 × 728

× 8

19 × 19 × 728

Conv 1×1
stride= 2×2

Relu
DepSepConv 728, 3×3
Relu
DepSepConv 1024, 3×3
MAXPOOLING
stride 2×2

19 × 19 × 1024

10 × 10 × 1024

Dep Sep Conv 1536, 3×3
Relu
Dep Sep Conv 2048, 3×3
Relu
Global Average Pooling

10 × 10 × 1536

10 × 10 × 2048

2048

Fully Connected layer 1

Fully ctd layer 2

Fully ctd layer 3

multinomial logistic regression