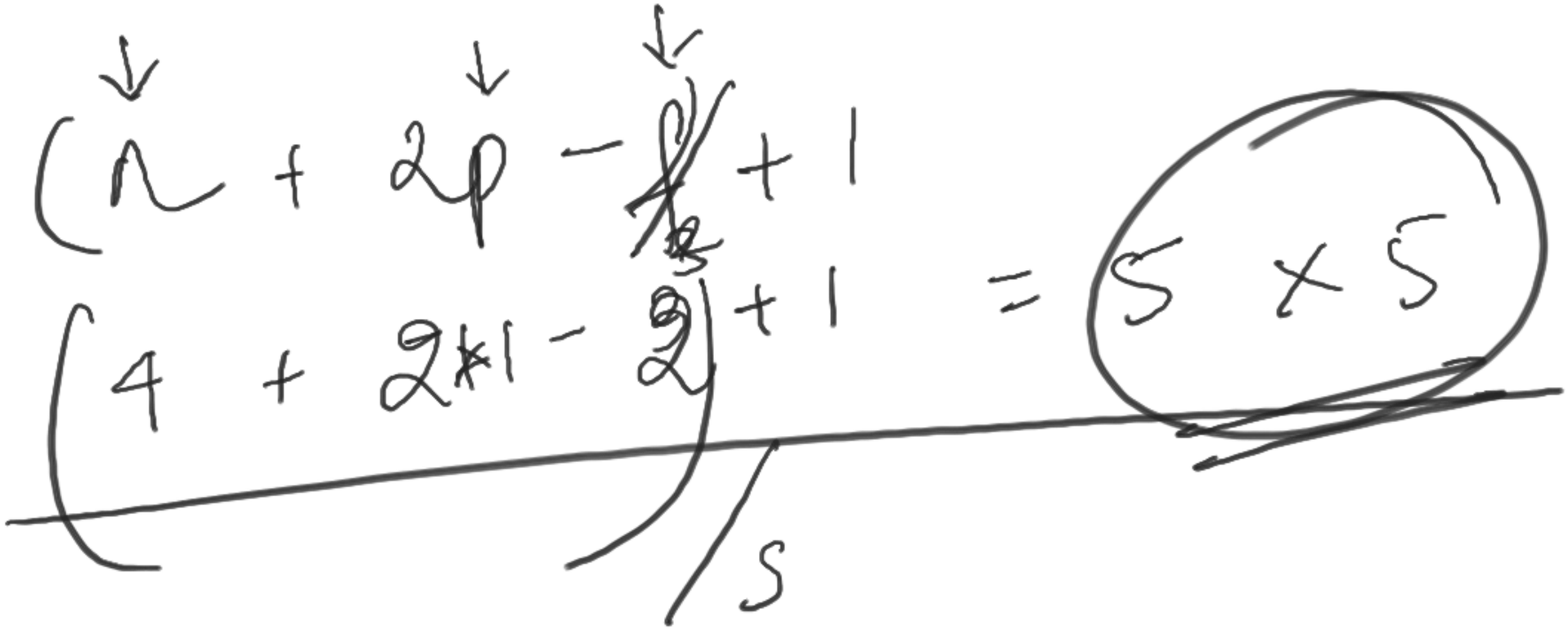


$$\frac{4+1}{2} = 3 \times 3$$

Q i/p=(32, 32, 3)

filter = 3×3

Conv2D1 → MaxPool1 → Conv2D2 → MaxPool2 → Flatten1 → Dense1 → Dense2 → Dense3

For Conv

#wts = #i/p channels * filter height * filter width * #o/p channels

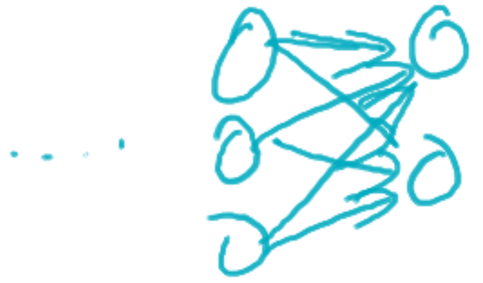
#o/p channels → no. of filters in the next layer

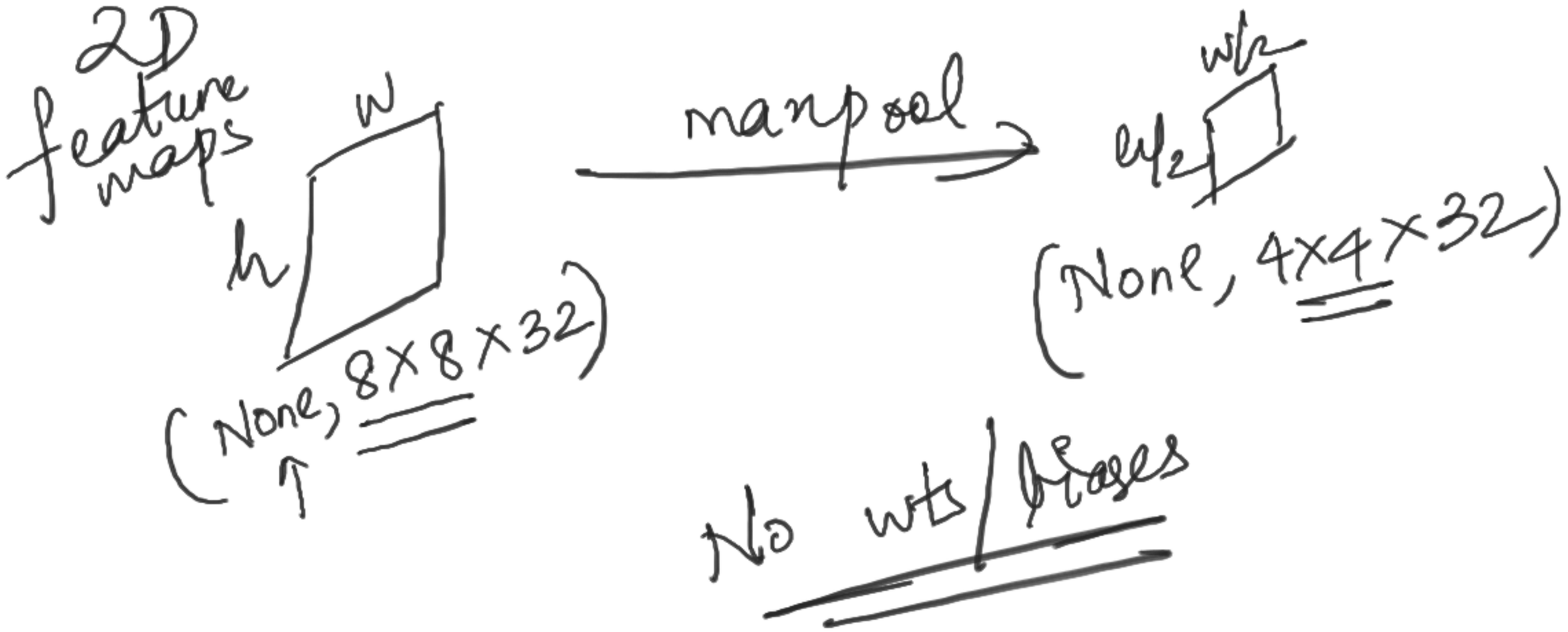
#biases = 1 * #o/p channels → same

o/p shape

For Dense

#wts = #o/p units * #i/p channels




2D feature maps
w
h
maxpool
w/2
h/2
(None, 8×8×32)
(None, 4×4×32)
No wts / biases

32, 32, 3 } padding

3×3

Conv2D | 16

o/p shape = (None, 32, 32, 16)

batch size, i/p H, i/p W, # of o/p channels

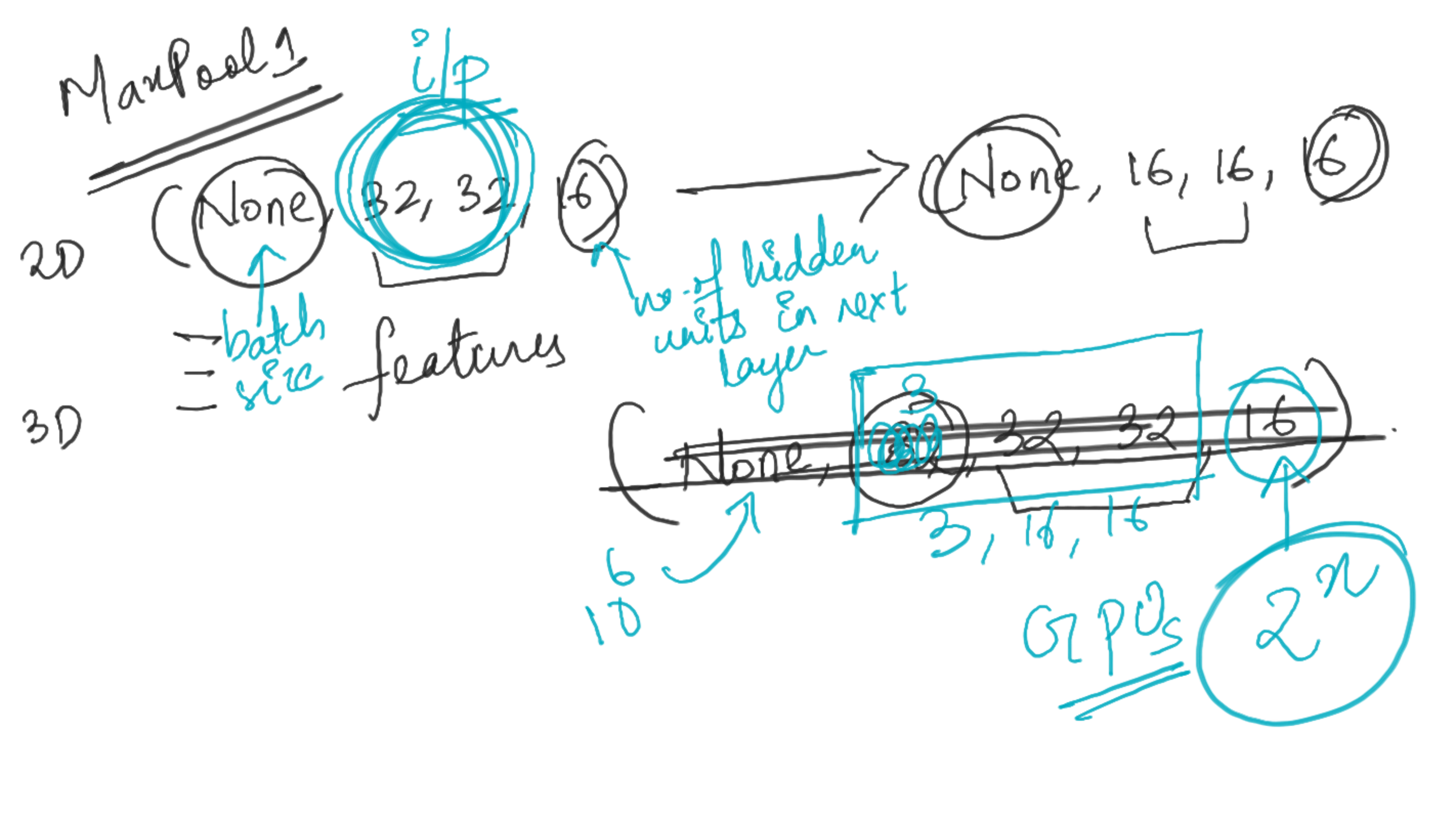
#wts = 3 * 3 * 3 * 16

= 432

16 ↑ no. of filters in this Conv2D layer

432, 16

448


MaxPool 1
i/p
None, 32, 32, 6
2D
batch size
features
3D
no. of hidden units in next layer
None, 16, 16, 6
None, 32, 32, 16
3, 16, 16
6
10
GPUs
2^n

Conv2D 2, 32

i/p shape (None, 16, 16, 16) ⟶ o/p shape (None, 16, 16, 32)

$\#wts = 16 * 3 * 3 * 32 = 4608$

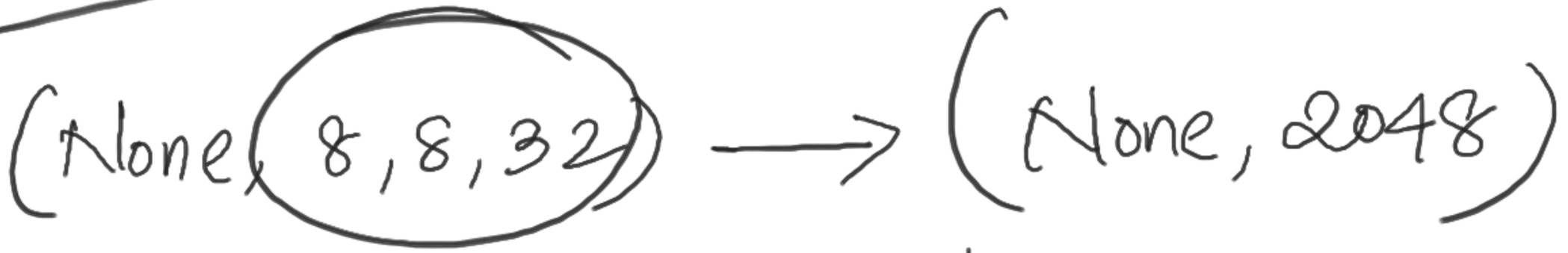
$\#b = 32$

4640

MaxPool2

(None, 16, 16, 32) ⟶ (None, 8, 8, 32)

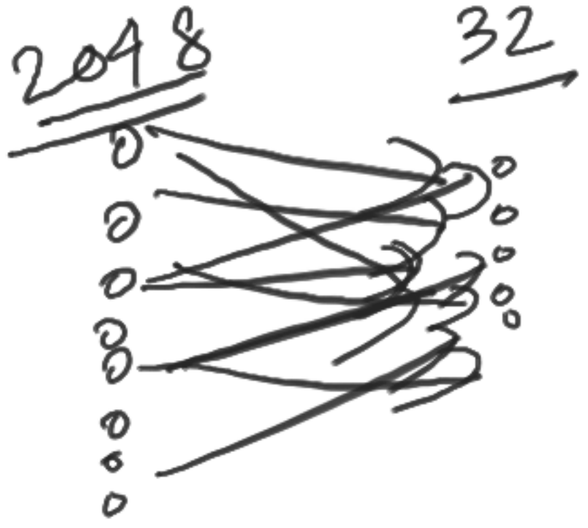
Flatten!



no weights / biases

Dense1 , 32

(None, 2048) ⟶ (None, 32)



#wts = 32 * 2048

= 65536

#b = 32

65568

Dense 2 , 16

(None, 32) ⟶ (None , 16)

32 * 16 = 512 wts

16 b

528

Dense 3 , 10

(None, 16) ⟶ (None, 10)

16 * 10 = 160

10

170