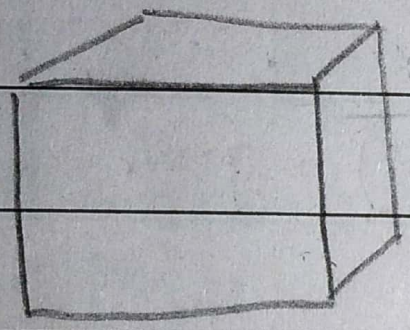
(20k-0310)

* ~~Convolution~~ layer 1:

Input



150 x 150 x 3

filter

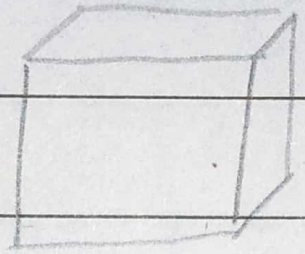
(3x3) x3

filter = 16 x

P = 0

S = 1

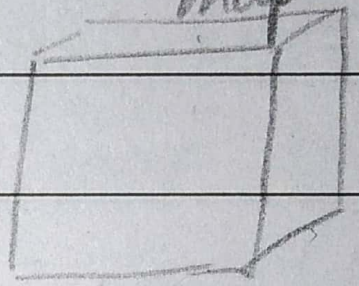
filter



3x3x3

=

resultant feature map



148 x 148 x 16

feature map =

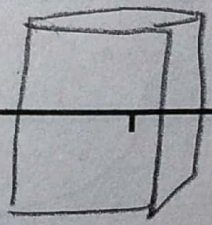
$$\frac{150 + (2(0) - 3) + 1}{1}$$

$$= 148$$

no. of feature = 3x3x3 = 27 + 1

28 x 16 = 448 param

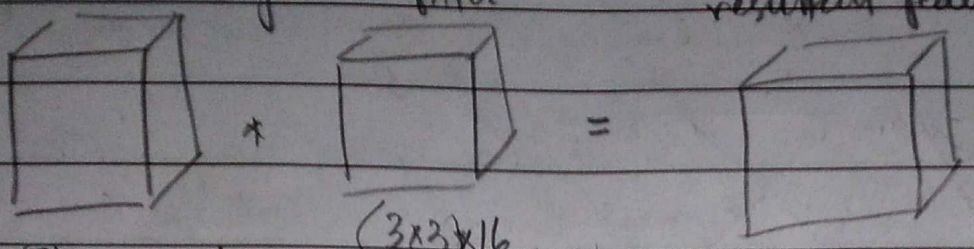
$$\# \text{pooling} = \frac{148 \times 148}{2 \times 2} = 74 \times 74 \times 16$$



URBANE

(20k-0310) Date: M T W T F S S

* convolution layer 2: filter → resultant feature map



32 filters

3x3x16 = filtersize

p=0

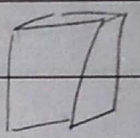
s=1

$\frac{74 + (2(0) - 3)}{1} + 1 = 72$

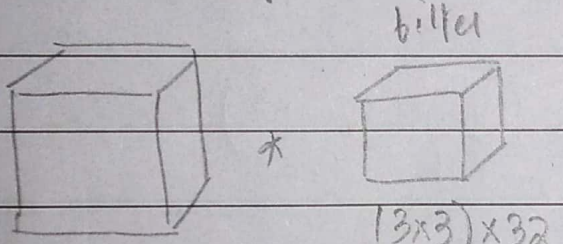
no. of featue = 3x3x16 = 144+1

145 x 32 = 4640 param

pooling = $\frac{72 \times 72}{2 \times 2}$ = 36 x 36 x 32



* convolution layer 3:



64 filters

3x3x32 = filtersize

p=0

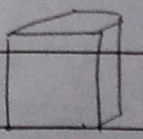
s=1

$\frac{36 + (2(0) - 3)}{1} + 1 = 34$

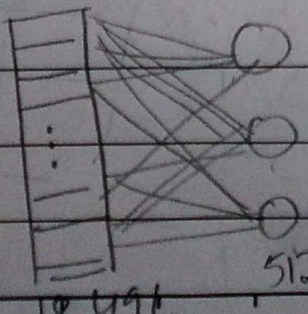
no. of featue = 3x3x32 = 288+1

289 x 64 = 18496 param

pooling = $\frac{34 \times 34}{2 \times 2}$ = 17 x 17 x 64



* flatten the layer:



= (18496+1) x 512 = 9470464