Theo Guidroz

CS 577 As4

$$R = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad G = \begin{bmatrix} 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \\ 4 & 4 & 4 & 4 \end{bmatrix}$$

1) Convolution = $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ - ignore

$$R = \begin{bmatrix} 9 & 9 \\ 9 & 9 \end{bmatrix} \quad G = \begin{bmatrix} 18 & 18 \\ 18 & 18 \end{bmatrix} \quad B = \begin{bmatrix} 18 & 18 \\ 27 & 27 \end{bmatrix}$$

2) Convolution = $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ - zero Padding

$$R \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad G \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 2 & 2 & 2 & 0 \\ 0 & 2 & 2 & 2 & 2 & 0 \\ 0 & 2 & 2 & 2 & 2 & 0 \\ 0 & 2 & 2 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad B \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 2 & 2 & 2 & 2 & 0 \\ 0 & 3 & 3 & 3 & 3 & 0 \\ 0 & 4 & 4 & 4 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$R = \begin{bmatrix} 4 & 6 & 6 & 4 \\ 6 & 9 & 9 & 6 \\ 6 & 9 & 9 & 6 \\ 4 & 6 & 6 & 4 \end{bmatrix} \quad G = \begin{bmatrix} 8 & 12 & 12 & 8 \\ 12 & 18 & 18 & 12 \\ 12 & 18 & 18 & 12 \\ 8 & 12 & 12 & 8 \end{bmatrix} \quad B = \begin{bmatrix} 6 & 9 & 9 & 6 \\ 12 & 18 & 18 & 12 \\ 18 & 24 & 24 & 18 \\ 14 & 21 & 21 & 14 \end{bmatrix}$$

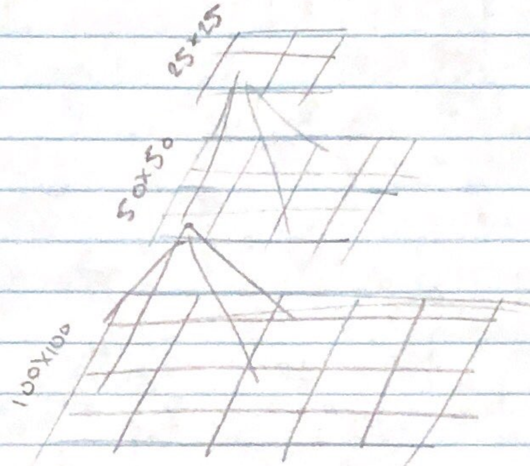
3) $$R = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad G = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 2 & 2 & 2 & 0 \\ 0 & 2 & 2 & 2 & 2 & 0 \\ 0 & 2 & 2 & 2 & 2 & 0 \\ 0 & 2 & 2 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 2 & 2 & 2 & 2 & 0 \\ 0 & 3 & 3 & 3 & 3 & 0 \\ 0 & 4 & 4 & 4 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Convolution = $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ - dilated with rate 2

$$R = \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix} \quad G = \begin{bmatrix} 8 & 8 \\ 8 & 8 \end{bmatrix} \quad B = \begin{bmatrix} 12 & 12 \\ 8 & 8 \end{bmatrix}$$

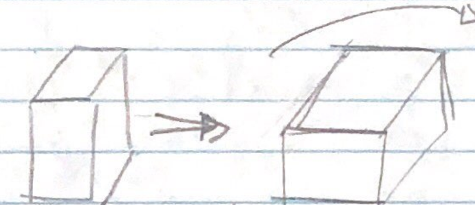
4) It is the process used to search and find the location of a template image, as a larger image. This is done by extracting image patterns and applying convolution filters per location

5)



look at different version of the image at different scale and keep the detection at the same size

6) As spatial dimensions decrease, depth increase to compensate for reduction in coefficients. This helps to keep the same number of coefficients.



Spatial dimensions decrease as depth increase

This is done to avoid information loss.

7) Image = $128 \times 128 \times 32$ $\quad w = 128 \quad k = 3$

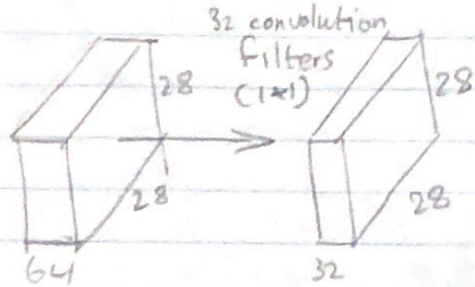
16 conv = $13 \times 3 \times 32$

Output size = $\frac{w-k+2p}{s} + 1 = \frac{128-3}{1} + 1 = 126$ ←

$\therefore$ img = $(126 \times 126 \times 16)$ ←

8) Stride 2, $\frac{w-k+2p}{2} + 1 = \frac{128-3}{2} + 1 = 64$

$\therefore$ img = $(64 \times 64 \times 16)$ ←

9) 

$$\text{dim} = \frac{128-1}{1} + 1 = 128$$

dim stays the same after 1 convolution. Since the convolution has 32 channels the output: (128 − 1 + 1, 128 + 1 − 1, 32)
= (128, 128, 32) reduced channels

10) Early layers extract the more generic featurs of the image such as the contrast while the deeper layers extract special features which uniquely identifies the image.

11) $R = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$ $G = \begin{bmatrix} 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \end{bmatrix}$ $B = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \\ 4 & 4 & 4 & 4 \end{bmatrix}$

$R = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ $G = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$ $B = \begin{bmatrix} 2 & 2 \\ 4 & 4 \end{bmatrix}$

12) Pooling is used to downsample spatial dimensions without changing the depth. Is is done to reduce the number of parameters to be trained which could help prevent overfitting

13) It is the process of altering the existing images properties to help generalizing the model. This technique is useful when there is a small amount of data available for training

14) It is the process in which a pretrained model, which has been trained on a large dataset, is used on a comparable small dataset to achieve better accuracy. It is useful when the dataset consist of general images and not specific.

15) In transfer learning, we have a trained convolution base and a trained classifier. To perform specific training, we replace the trained classifier with a new classifier which is randomly initialized. The model is then trained on the new dataset. To ensure that the trained convolutional base is not affected by noise made by the new classifier, we freeze it. Freezing means to disable changes in some weights in the model. A frozen part is not trainable

16) After training the network, we can unfreeze part of the top layers and retrain to allow the model to fit the data

17) Inception blocks refers to performing convolution in parallel and then doing concatenation with all the outputs. This process increases dimensions but 1x1 convolution is used to couteract. The purpose of inception blocks is to make the model learn spatial and channel features instead of learning them together.

18) Risidual blocks make it easier to learn deviation from identity instead of functions. They skip connections help with vanishing gradients