CNNs (Convolutional Neural Networks)

Convolutions.

⇒ Image is represented as a Matrix.

Grayscale Image 0 — 255 intensity values.

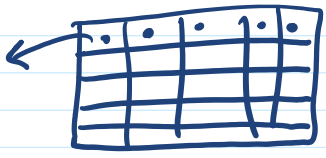
As a function:- $(x, y) \leftarrow I$

A digital image:- discrete (quantized version of this function)

$$\hat{I}(x, y) = I(x, y) + a$$ ← constant

Operations

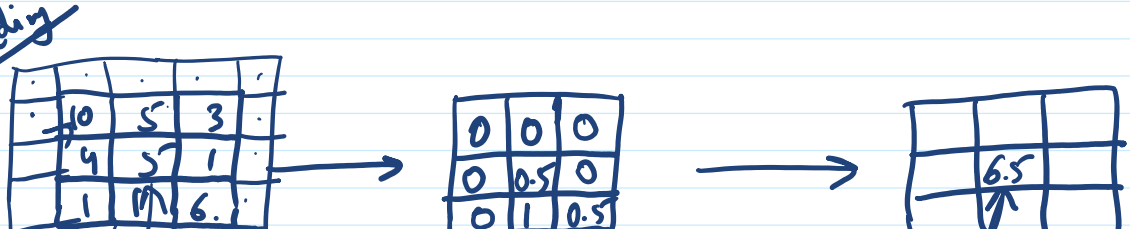
- Point operation.
- Local operation — pxp neighborhood
- Global operation. — NxN

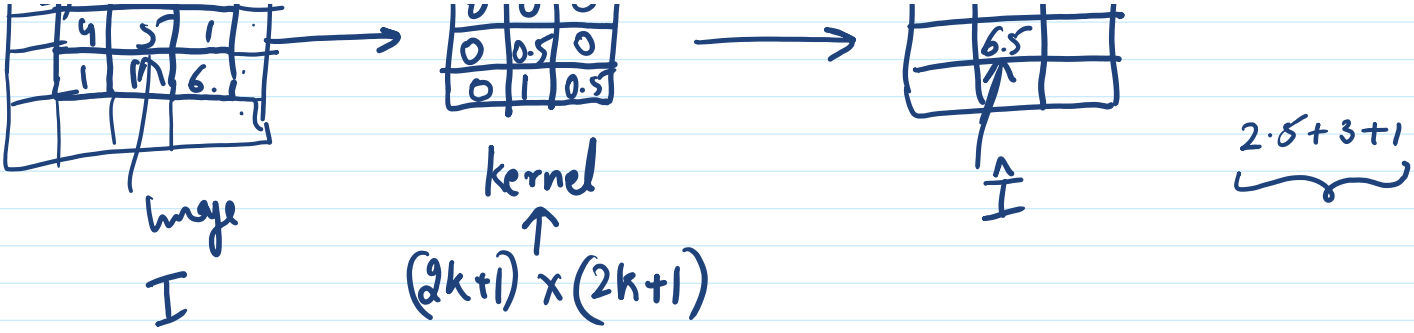


Let I be the image with NxM pixels,
I_{MAX} be the max. intensity value (255)

filter → MxN

padding





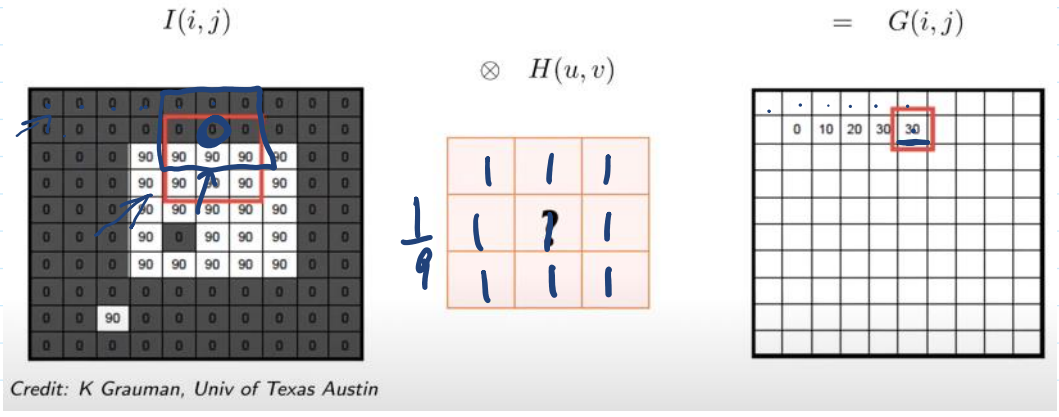
Linear Filter:- Replacing each pixel by weighted sum of neighbors.

Kernel / mask / filter

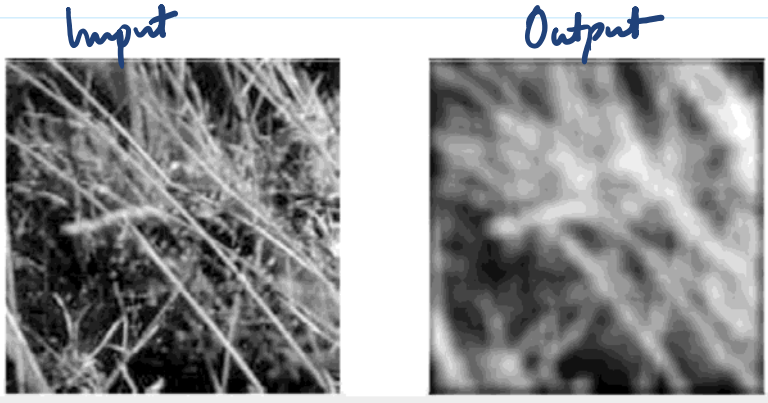
Average_filter_/Box_filter:-

$\frac{1}{9}$ [1 1 1; 1 1 1; 1 1 1]

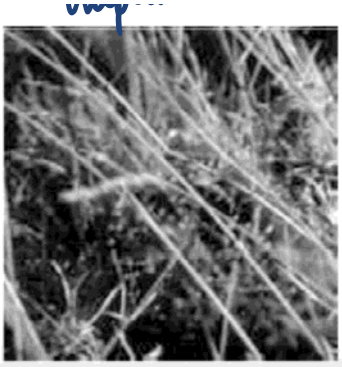
$\frac{1}{9}(90+90+90)$

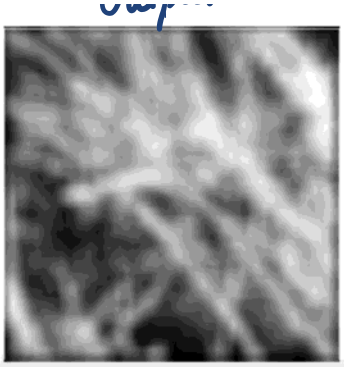


Credit: K Grauman, Univ of Texas Austin



Credit: K Grauman, Univ of Texas Austin

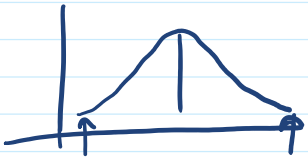




Credit: K Grauman, Univ of Texas Austin

Gaussian Average Filter :-

$$h(u,v) = \frac{1}{2\pi\sigma^2} \exp \frac{-u^2+v^2}{\sigma^2}$$



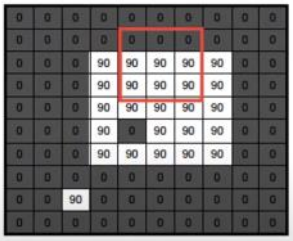
$\frac{1}{16}$

1/16	2/16	1/16
2	4	2
1	2	1

2D gaussian filter.

5x5

$I(i,j)$



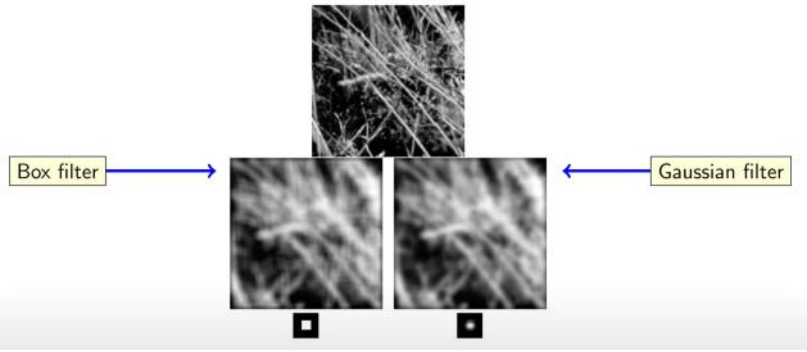
$\otimes \quad H(u,v)$

$\frac{1}{16}$

1	2	1
2	4	2
1	2	1

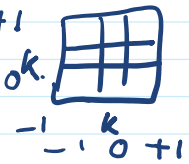
$=$

Credit: K Grauman, Univ of Texas Austin



Given a kernel size $(2k+1) \times (2k+1)$:

$\Rightarrow$ Cross-correlation:-



$$G(i,j) = \frac{1}{(2k+1)^2} \sum_{u=-k}^{k} \sum_{v=-k}^{k} I(i+u, j+v)$$

uniform weights to each pixel

$I(i,j)$: Loop over pixels in considered neighbourhood around $I(i,j)$

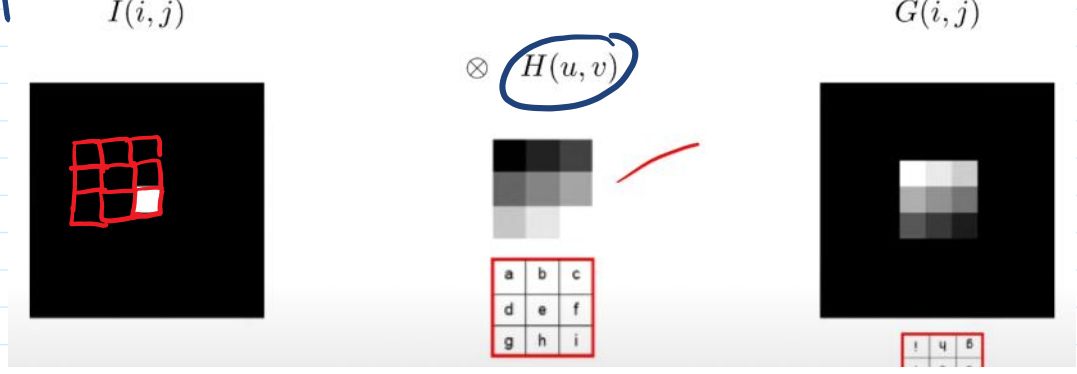
$$G(i,j) = \sum_{u=-k}^{k} \sum_{v=-k}^{k} H(u,v)\, I(i+u, j+v)$$

$H(u,v)$: non-uniform weights / filter coefficients.

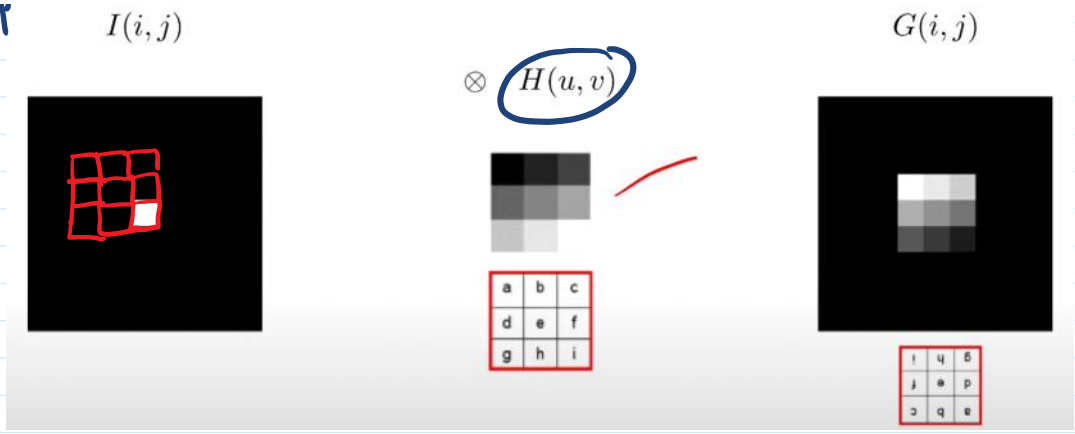
$$G = H \circledast I$$

dot product b/w local neighborhood and kernel for each pixel

Impulse image



Imp



Cross-correlation does not maintain an identity with an impulse function.

Convolution :-

$$G(i,j) = \sum_{u=-k}^{k} \sum_{v=-k}^{k} H(u,v)\, I(i-u, j-v)$$

double flipping the filter so that output turns out to be the filter itself.

$\otimes$ ← correlation.

$$G = H * I$$

↑ convolution.

Cross-correlation is sliding a filter across an image.

Convolution is sliding a flipped filter across an image.