Preclass 15: Image Convolution

[SCS4049] Machine Learning and Data Science

Seongsik Park (s.park@dgu.edu)

School of AI Convergence & Department of Artificial Intelligence, Dongguk University

흑백 이미지 = 2차원 행렬

0 ~ 1 실수.

검은색 흰색.

0 ~ 255 정수.

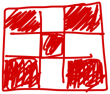
행렬

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

3x3

⇒

이미지



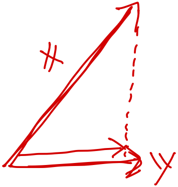
3x3

Color intensity

$$\begin{bmatrix} 1 & 1 & 10 \\ 1 & 9 & 10 \\ 10 & 10 & 9 \end{bmatrix} * \begin{bmatrix} 10 & 1 \\ 10 & 1 \end{bmatrix} = \begin{bmatrix} 102 & 210 & \bigcirc \\ \bigcirc & \bigcirc & \bigcirc \\ \bigcirc & \bigcirc & \bigcirc \end{bmatrix}$$

3x3 이미지 | 2x2 커널 | 3x3 결과이미지

$$1\cdot1 + 10\cdot10 + 9\cdot1 + 10\cdot10 = 210$$



$$x^T y = \sum x_i y_i$$

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}^T \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 1\cdot1 + 2\cdot0 + 3\cdot0$$

$$\begin{bmatrix} 1 & 1 & 10 \\ 1 & 9 & 10 \\ 10 & 10 & 9 \end{bmatrix} * \begin{bmatrix} 10 & 1 \\ 10 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

102

210

3x3

2x2

1 1

3x3

padding = zero

→ heat map
featuremap

$$\underbrace{\begin{matrix} 0 & 1 & 2 & 3 & 0 \\ 0 & 2 & 3 & 4 & 0 \\ 0 & 3 & 4 & 5 & 0 \end{matrix}}_{3\times 3} * \underbrace{\begin{bmatrix} 1 & 0 & 1 \end{bmatrix}}_{1\times 3} = \underbrace{\begin{bmatrix} 2 & 4 & 2 \\ 3 & 6 & 3 \\ 4 & 8 & 4 \end{bmatrix}}_{3\times 3}$$

padding의 크기 ← kernel의 크기가 결정.

이상이면 양옆·위아래로 대칭되는 개수

$$\begin{bmatrix} 1 & 1 & 10 & \boxed{10} \\ 1 & 10 & 10 & \boxed{10} \\ 10 & 10 & 10 & \boxed{10} \end{bmatrix} * \underbrace{\begin{bmatrix} -1 & 1 \end{bmatrix}}_{1\times 2} = \begin{bmatrix} 0 & -9 & 0 \\ -9 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Clamp

① convolution 계산하는 법

② padding 필요, 얼만큼 ← kernel의 크기.

③ padding 종류

④ convolution 결과 → feature map
heat map

Convolution

Given a filter kernel $\mathcal{H}$, the convolution of the kernel with image $\mathcal{F}$ is an image $\mathcal{R}$. The (i, j)-th component of $\mathcal{R}$ is given by

$$R_{ij} = \sum_{u,v} H_{i-u,j-v} F_{uv}. \tag{1}$$

- **kernel** of the filter: the pattern of weights used for a linear filter
- **convolution**: the process of applying the filter

Linear filter and convolution

This operation is called **convolution**

$$R(f) = (h * f) \tag{2}$$

- *commutative*: $(g * h)(x) = (h * g)(x)$
- *associative*: $f * (g * h) = (f * g) * h$
- *distributive*: $f * (g + h) = f * g + f * h$

Padding (border effects)

- **zero**: set all pixels outside the source image to 0
- **constant**: set all pixels outside the source image to a specified border value
- **clamp**: repeat edge pixels indefinitely
- **wrap**: loop "around" the image in a "toroidal" configuration
- **mirror**: reflect pixels across the image edge
- **extend**: extend the signal by subtracting the mirrored version of the signal from the edge pixel value

Padding (border effects)

0 바둑판식 거울모드로 확장

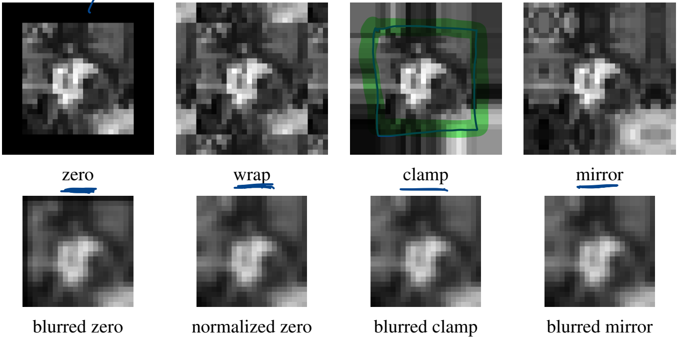


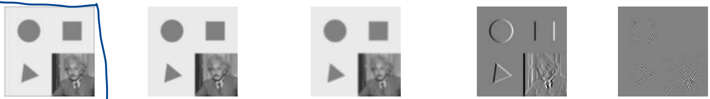
Figure 3.13 *Border padding (top row) and the results of blurring the padded image (bottom row). The normalized zero image is the result of dividing (normalizing) the blurred zero-padded RGBA image by its corresponding soft alpha value.*

Examples of linear filter

3x3

$$\frac{1}{K^2}\begin{array}{|c|c|c|c|}\hline 1 & 1 & \cdots & 1 \\ \hline 1 & 1 & \cdots & 1 \\ \hline \vdots & \vdots & 1 & \vdots \\ \hline 1 & 1 & \cdots & 1 \\ \hline \end{array} \qquad \frac{1}{16}\begin{array}{|c|c|c|}\hline 1 & 2 & 1 \\ \hline 2 & 4 & 2 \\ \hline 1 & 2 & 1 \\ \hline \end{array} \qquad \frac{1}{256}\begin{array}{|c|c|c|c|c|}\hline 1 & 4 & 6 & 4 & 1 \\ \hline 4 & 16 & 24 & 16 & 4 \\ \hline 6 & 24 & 36 & 24 & 6 \\ \hline 4 & 16 & 24 & 16 & 4 \\ \hline 1 & 4 & 6 & 4 & 1 \\ \hline \end{array} \qquad \frac{1}{8}\begin{array}{|c|c|c|}\hline -1 & 0 & 1 \\ \hline -2 & 0 & 2 \\ \hline -1 & 0 & 1 \\ \hline \end{array} \qquad \frac{1}{4}\begin{array}{|c|c|c|}\hline 1 & -2 & 1 \\ \hline -2 & 4 & -2 \\ \hline 1 & -2 & 1 \\ \hline \end{array}$$

$$\frac{1}{K}\begin{array}{|c|c|c|c|}\hline 1 & 1 & \cdots & 1 \\ \hline \end{array} \qquad \frac{1}{4}\begin{array}{|c|c|c|}\hline 1 & 2 & 1 \\ \hline \end{array} \qquad \frac{1}{16}\begin{array}{|c|c|c|c|c|}\hline 1 & 4 & 6 & 4 & 1 \\ \hline \end{array} \qquad \frac{1}{2}\begin{array}{|c|c|c|}\hline -1 & 0 & 1 \\ \hline \end{array} \qquad \frac{1}{2}\begin{array}{|c|c|c|}\hline 1 & -2 & 1 \\ \hline \end{array}$$



(a) box, $K = 5$ (b) bilinear (c) "Gaussian" (d) Sobel (e) corner

Figure 3.14 *Separable linear filters: For each image (a)–(e), we show the 2D filter kernel (top), the corresponding horizontal 1D kernel (middle), and the filtered image (bottom). The filtered Sobel and corner images are signed, scaled up by* $2\times$ *and* $4\times$*, respectively, and added to a gray offset before display.*