Assignment 2

2018年11月8日 20:40

1. $$[f * w](x) = \sum_{y \in \mathbb{Z}^2} \sum_{k=1}^{K} f_k(y) w_k(y-x) \qquad (1)$$

a. $[L_t f](x) = f(x-t)$

$$[[L_t f] * w](x) = \sum_{y \in \mathbb{Z}^2} \sum_{k=1}^{K} f_k(y-t) \cdot w_k(y-x)$$

let $y-t = z$, then $y = t+z$

$$[L_t f * w](x) = \sum_{z \in \mathbb{Z}^2} \sum_{k=1}^{K} f_k(z) w_k(z+t-x)$$

$$[L_t [f * w]](x) = \sum_{y \in \mathbb{Z}^2} \sum_{k=1}^{K} f_k(y) w_k(y-(x-t)) = \sum_{y \in \mathbb{Z}^2} \sum_{k=1}^{K} f_k(y) w_k(y+t-x) = [L_t f * w](x)$$

Q.E.D

b. $$R = \begin{bmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \qquad R^{-1} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

$[L_R f](x) = f(R^{-1}x)$, then

$$[L_R f * w](x) = \sum_{y \in \mathbb{Z}^2} \sum_{k=1}^{K} f_k(R^{-1}y) w_k(y-x)$$

$$= \sum_{z \in \mathbb{Z}^2} \sum_{k=1}^{K} f_k(z) w_k(Rz-x) \qquad [\text{let } z = R^{-1}y]$$

$$L_R[f * L_{R^{-1}} w](x) = \sum_{y \in \mathbb{Z}^2} \sum_{k=1}^{K} f_k(y) w_k[R(y-R^{-1}x)]$$

$$= \sum_{y \in \mathbb{Z}^2} \sum_{k=1}^{K} f_k(y) w_k[Ry-x)$$

$$= [L_R f * w](x) \qquad \text{Q.E.D.}$$

c. $[f * w](g) = \sum_{h \in G} \sum_{k=1}^{K} f_k(h) w_k(g^{-1}h)$, $[L_u f](g) = f(u^{-1}g)$

$$\therefore [L_u f * w](g) = \sum_{h \in G} \sum_{k=1}^{K} f_k(u^{-1}h) w_k(g^{-1}h)$$

$$= \sum_{v \in G} \sum_{k=1}^{K} f_k(v) w_k(g^{-1}uv) \qquad [\text{let } u^{-1}h = v \Rightarrow h = uv]$$

$$L_u[f * w](g) = \sum_{h \in G} \sum_{k=1}^{K} f_k(h) w_k((u^{-1}g)^{-1}h)$$

$$= \sum_{h \in G} \sum_{k=1}^{K} f_k(h) w_k(g^{-1}uh)$$

$$= [L_u f * w](g) \qquad \text{Q.E.D.}$$

+ Group Convolution using traditional convolution:
  - Rotate the input feature map four times by $\{0°, 90°, 180°, 270°\}$, respectively.
  - Rotate the filters by $\{0°, 90°, 180°, 270°\}$ to obtain w_1, w_2, w_3, w_4
  - output is the convolution of these rotated input and filters

Assignment 2

2018年11月8日 21:50

2.b. (d1) + $Var[w_l x_l] = E[w_l^2 x_l^2] - E[w_l x_l]^2$

$= E[w_l^2]E[x_l^2] - E[w_l]^2 E[x_l]^2$ (independe of w_l & x_l)

$= (Var[w_l] + E[w_l]^2) E[x_l^2] - E[w_l]^2 E[x_l]^2$

$= Var[w_l] \cdot E[x_l^2]$. ($E[w_l] = 0$)

From (9), $Var[y_l] = n_l Var[w_l x_l] = n_l Var[w_l] E[x_l^2]$

+ $x_l = ReLU(y_{l-1}) = \begin{cases} y_{l-1}, & y_{l-1} > 0 \\ 0, & \text{otherwise} \end{cases}$

$E[x_l^2] = \int_{-\infty}^{\infty} p(x_l)\, x_l^2\, dx_l$

$= \int_0^{\infty} p(y_{l-1})\, y_{l-1}^2\, dy_{l-1}$ ($y_{l-1} \sim N(0, \sigma^2)$)

$= \frac{1}{2} \int_{-\infty}^{\infty} p(y_{l-1})\, y_{l-1}^2\, dy_{l-1}$

$= \frac{1}{2} E[y_{l-1}^2]$

$= \frac{1}{2} Var(y_{l-1})$. ($E[y_{l-1}] = 0$)

Q.E.D.