◇◇Conv

◇◇

P:padding:(pt,pb,pl,pr)
S:stride
K:kernel
W:weight:(N, C, K , K)

I:input:(C,HI,WI)
O:output:(N,HO,WO)
$\frac{dL}{dO}$:(N,HO,WO)
$\frac{dL}{dI}$:(C,HI,WI)
$\frac{dL}{dW}$:(N,C,K,K)
$\frac{dL}{dB}$:(N,)

◇◇

O:Output I:Input IP:Input with padding

◇◇◇◇

◇◇◇◇

$$O(n,x,y) = \sum_{c=0}^{C-1}\sum_{j=0}^{K-1}\sum_{j=0}^{K-1} F(n,c,i,j) * IP(c, yS+j, xs+i)$$
$$= \sum_{c=0}^{C-1}\sum_{j=0}^{K-1}\sum_{j=0}^{K-1} W(n,c,i,j) * I(c, yS+j-pt, xs+i-pl)$$

◇◇◇◇◇◇◇◇◇◇padding◇◇◇◇◇◇

◇◇◇◇

- stride = 1

$$\frac{dL}{dI(c,y,x)} = \frac{dL}{dO} * \frac{dO}{dI(c,y,x)}$$

$$= \sum_{n=0}^{N-1}\sum_{j=0}^{H-1}\sum_{i=0}^{W-1} \frac{dL}{dO(n,j,i)} * \frac{dO(n,j,i)}{dI(c,y,x)}, \diamond\diamond\diamond\diamond\diamond I(c,y,x)\diamond O\diamond\diamond\diamond$$

$$= \sum_{n=0}^{N-1}\sum_{j=0}^{H-1}\sum_{i=0}^{W-1} \frac{dL}{dO(n,j,i)} * W(n,c,y-j+pt,x-i+pl)$$

$$\diamond j' = y - j + pt, i' = y - i + pl$$

$$= \sum_{n=0}^{N-1}\sum_{j'=0}^{K-1}\sum_{i'=0}^{K-1} \frac{dL}{dO(n,y-j'+pt,x-i'+pt)} * W(n,c,j',i')$$

$$\diamond j = K - 1 - j', i = K - 1 - i'$$

$$= \sum_{n=0}^{N-1}\sum_{j=0}^{K-1}\sum_{i=0}^{K-1} \frac{dL}{dO(n,j+y+(pt-K+1),i+x+(pl-K+1))} * W$$

$$\diamond \frac{dL}{dO} \diamond padding \diamond (K - 1 - pt, HI - HO + pt, K - 1 - pl, WI - WO$$

$$= \sum_{n=0}^{N-1}\sum_{j=0}^{K-1}\sum_{i=0}^{K-1} \frac{dL}{dOP(n,j+y,i+x)} * W(n,c,K-1-j,K-1-i)$$

$$\frac{dL}{dI(n)} = \sum_{n=0}^{N-1} conv(\frac{dL}{dOP(n)}, W^T(n,c))$$

$\diamond \frac{dL}{dI} \diamond \frac{dL}{dOP} \diamond W \diamond\diamond 180° \diamond\diamond\diamond\diamond\diamond\diamond\diamond\diamond\diamond\diamond\diamond\diamond\diamond$

$\diamond\diamond\diamond\diamond$

- $\diamond$bias$\diamond\diamond$

$$\frac{dL}{dB(n)} = \sum_{h}^{HO-1} dO(n,h,w)$$

- $\diamond$weight$\diamond\diamond$

$$\frac{dL}{dW(n,c,y,x)} = \frac{dL}{dO} * \frac{dO}{dW(n,c,y,x)}$$

$$= \sum_{j=0}^{H-1} \sum_{i=0}^{W-1} \frac{dL}{dO(n,j,i)} * \frac{dO(n,j,i)}{dW(n,c,y,x)}$$

$$= \sum_{j=0}^{H-1} \sum_{i=0}^{W-1} \frac{dL}{dO(n,j,i)} * IP(n,c,y,x)$$

◇◇◇

$$\frac{dL}{dW(n,c)} = conv(\frac{dL}{dO(n)}, IP(c))$$