

یادگیری عمیق

تمرین چهارم

نیکی نزاکتی

98522094

$$(w \star f)(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x-s, y-t)$$

$$g(0,0) = 0, g(0,1) = 20$$

$$g(0,2) = 1 \times 0 + 1 \times 0 + 1 \times 0 + 1 \times 0 + 10 \times -8 + 1 \times 0 + 1 \times 0 + 1 \times 10 + 1 \times 0 = -70,$$

$$g(0,3) = 20, g(0,4) = 0,$$

$$g(1,0) = 0, g(1,1) = 30$$

$$g(1,2) = 1 \times 0 + 1 \times 10 + 1 \times 0 + 1 \times 0 + 10 \times -8 + 1 \times 0 + 1 \times 0 + 1 \times 10 + 1 \times 0 = -60,$$

$$g(1,3) = 30, g(1,4) = 0,$$

$$g(2,0) = 0, g(2,1) = 30$$

$$g(2,2) = 1 \times 0 + 1 \times 10 + 1 \times 0 + 1 \times 0 + 10 \times -8 + 1 \times 0 + 1 \times 0 + 1 \times 10 + 1 \times 0 = -60,$$

$$g(2,3) = 30, g(2,4) = 0,$$

$$g(3,0) = 0, g(3,1) = 30$$

$$g(3,2) = 1 \times 0 + 1 \times 10 + 1 \times 0 + 1 \times 0 + 10 \times -8 + 1 \times 0 + 1 \times 0 + 1 \times 10 + 1 \times 0 = -60,$$

$$g(3,3) = 30, g(3,4) = 0,$$

$$g(4,0) = 0, g(4,1) = 20$$

$$g(4,2) = 1 \times 0 + 1 \times 0 + 1 \times 0 + 1 \times 0 + 10 \times -8 + 1 \times 0 + 1 \times 0 + 1 \times 10 + 1 \times 0 = -70,$$

$$g(4,3) = 20, g(4,4) = 0,$$

ماتریس نتیجه:

0	20	-70	20	0
0	30	-60	30	0
0	30	-60	30	0
0	30	-60	30	0
0	20	-70	20	0

این کرنل فیلتر لاپلاسین است که اعمال آن نواحی با تغییر شدت سریع را برجسته می کند و بنابراین اغلب برای تشخیص لبه استفاده می شود. همانطور هم که در ماتریس نتیجه مشاهده می شود خط وسط که لبه محسوب می شد گسترش یافته است.

2.

در مقاله ی Network In Network لایه کانولوشن 1در1 به عنوان Cross Channel Down sampling یا Cross Channel Pooling استفاده شده است. به عبارت دیگر، برای کاهش تعداد کانال ها و در عین حال معرفی non-linearity استفاده شده است. 1X1 Convolution به سادگی به این معنی است که فیلتر دارای اندازه 1X1 است، یعنی یک عدد واحد برخلاف فیلتر ماتریسی. این فیلتر 1X1 بر روی کل تصویر ورودی پیکسل به پیکسل convolve می شود. در این حالت خروجی همان ارتفاع و وزن ورودی را خواهد داشت اما فقط یک کانال خواهد داشت. بنابراین اگر بخواهیم عمق را کاهش دهیم و Hight X Width نقشه های ویژگی را ثابت نگه داریم، می توانیم فیلترهای 1X1 را برای رسیدن به این اثر انتخاب کنیم. این اثر نمونه cross channel down-sampling را "Dimensionality reduction" می نامند. این امر نیازهای محاسباتی را کاهش می دهد و کارایی بیشتری خواهد داشت.

3.

(الف)

Input shape = (28 * 28 * 1)
 32 kernel (3,3) , stride = 2 , padding = 1
 $W2 = (W1 - F + 2P) / S + 1 = (28 - 3 + 2) / 2 + 1 = 14$

Output = (14 * 14 * 32)

(ب)

Input shape = (14 * 14 * 32)
 kernel (2,2) , stride = 2 , padding = 0
 $W2 = (W1 - F) / S + 1 = (14 - 2) / 2 + 1 = 7$

Output = (7 * 7 * 32)

(پ)

Max pool -> flat = $7 * 7 * 32 = (None, 1568)$ -> Dense -> $(1568 + 1) * 32 = (None, 50208)$
 -> output (50208, 5)

$$\frac{\partial L}{\partial y} = 1$$

(4)

o_{11}	o_{12}
o_{21}	o_{22}

$$o_{11} = w_{11}F_{11} + w_{12}F_{12} + w_{21}F_{21} + w_{22}F_{22}$$

$$o_{11} = 2 \times 0 + 3 \times \frac{3}{4} + 3 \times \frac{1}{3} - 2 = 10$$

$$o_{12} = w_{12}F_{11} + w_{13}F_{12} + w_{22}F_{21} + w_{23}F_{22}$$

$$o_{12} = 3 \times 0 + 4 \times \frac{3}{12} + 1 \times \frac{1}{1} - 10 = 3$$

$$o_{21} = w_{21}F_{11} + w_{22}F_{12} + w_{31}F_{21} + w_{32}F_{22}$$

$$o_{21} = 3 \times 0 + 1 \times \frac{3}{3} + 1 \times \frac{4}{4} + 2 = 9$$

$$o_{22} = w_{22}F_{11} + w_{23}F_{12} + w_{32}F_{21} + w_{33}F_{22}$$

$$o_{22} = 5 \times 3 + \frac{1}{15} + 4 = 18$$

$$\Rightarrow$$

10	3
9	18

$$\Rightarrow \text{GAP} = 10$$

$$\frac{\partial L}{\partial w_{11}} = F_{11} * \frac{\partial L}{\partial o_{11}} = 0, \quad \frac{\partial L}{\partial w_{12}} = \frac{1}{2} (3) = \frac{3}{2}$$

$$\frac{\partial L}{\partial w_{13}} = \frac{3}{2}, \quad \frac{\partial L}{\partial w_{21}} = \frac{1}{2} (1) = \frac{1}{2}, \quad \frac{\partial L}{\partial w_{22}} = \frac{1}{2} (-2 + 1 + 3) = 1$$

$$\frac{\partial L}{\partial w_{23}} = \frac{1}{2} (-2 + 3) = \frac{1}{2}, \quad \frac{\partial L}{\partial w_{31}} = \frac{1}{2}, \quad \frac{\partial L}{\partial w_{32}} = \frac{1}{2} (-2 + 3) = \frac{1}{2}$$

$$\frac{\partial L}{\partial w_{33}} = \frac{1}{2} (-2) = -1$$

Resources:

<https://www.geeksforgeeks.org/>

<https://towardsdatascience.com/>