SHARIF UNIVERSITY OF TECHNOLOGY

Deep Learning Assignment 3

Saeedreza Zouashkiani

January 1, 2023

1

1.1

The output of the following image

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

with the following filter:

$$\frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \tag{2}$$

is:

$$\frac{1}{9}\begin{bmatrix} 5 & 4 & 5 & 4 & 5 & 4 \\ 4 & 5 & 4 & 5 & 4 & 5 \\ 5 & 4 & 5 & 4 & 5 & 4 \\ 4 & 5 & 4 & 5 & 4 & 5 \\ 5 & 4 & 5 & 4 & 5 & 4 \\ 4 & 5 & 4 & 5 & 4 & 5 \end{bmatrix} \tag{3}$$

The output using same padding is:

$$\frac{1}{9}\begin{bmatrix} 2 & 3 & 3 & 3 & 3 & 3 & 2 \\ 3 & 5 & 4 & 5 & 4 & 5 & 3 \\ 3 & 4 & 5 & 4 & 5 & 4 & 3 \\ 3 & 5 & 4 & 5 & 4 & 5 & 3 \\ 3 & 4 & 5 & 4 & 5 & 4 & 3 \\ 3 & 5 & 4 & 5 & 4 & 5 & 3 \\ 2 & 3 & 3 & 3 & 3 & 3 & 2 \end{bmatrix} \tag{4}$$

Assuming that the edges remain the same, the output using valid padding is(Vague question!):

$$\frac{1}{9}\begin{bmatrix} 9 & 0 & 9 & 0 & 9 & 0 & 9 & 0 \\ 0 & 5 & 4 & 5 & 4 & 5 & 4 & 9 \\ 9 & 4 & 5 & 4 & 5 & 4 & 5 & 0 \\ 0 & 5 & 4 & 5 & 4 & 5 & 4 & 9 \\ 9 & 4 & 5 & 4 & 5 & 4 & 5 & 0 \\ 0 & 5 & 4 & 5 & 4 & 5 & 4 & 9 \\ 9 & 0 & 9 & 0 & 9 & 0 & 9 & 0 \end{bmatrix} \tag{5}$$

1.2

This filter works as a blurring filter.

2

Layer	Output Dimension	Number of Parameters
Input	$32 \times 32 \times 3$	0
CONV3-10	$32 \times 32 \times 10$	$3 \times 3 \times 10 + 10 = 280$
ReLU	$32 \times 32 \times 10$	0
POOL-2	$16 \times 16 \times 10$	0
CONV3-20(3, 2)	$6 \times 6 \times 20$	$3 \times 3 \times 20 \times 10 + 20 = 3620$
ReLU	$6 \times 6 \times 20$	0
POOL-2	$3 \times 3 \times 20$	0
FLATTEN	180	0
FC-10	10	$180 \times 10 + 10 = 1810$

3

3.1

k, b are the parameters of the model.

3.2

$$\begin{aligned} \frac{\partial L}{\partial \omega_1} &= \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial \omega_1} \\ &= (\hat{y} - y) \frac{\partial \hat{y}}{\partial \omega_1} \\ &= (\hat{y} - y) v_1 \end{aligned} \tag{6}$$

$$\begin{aligned}\frac{\partial L}{\partial \omega_2} &= \frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial \omega_2} \\ &= (\hat{y}-y)\frac{\partial \hat{y}}{\partial \omega_2} \\ &= (\hat{y}-y)v_2\end{aligned} \tag{7}$$

$$\begin{aligned}\frac{\partial L}{\partial a} &= \frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial a} \\ &= (\hat{y}-y)\end{aligned} \tag{8}$$

3.3

$$\begin{aligned}\frac{\partial L}{\partial z_1} &= \frac{\partial L}{\partial v_1}\frac{\partial v_1}{\partial z_1} \\ &= \delta_1 u(z_1-z_2)u(z_1) = \alpha_1\end{aligned} \tag{9}$$

$$\begin{aligned}\frac{\partial L}{\partial z_2} &= \frac{\partial L}{\partial v_1}\frac{\partial v_1}{\partial z_2} + \frac{\partial L}{\partial v_2}\frac{\partial v_2}{\partial z_2} \\ &= \delta_1 u(z_2-z_1)u(z_2) + \delta_2 u(z_2-z_3)u(z_2) = \alpha_2\end{aligned} \tag{10}$$

$$\begin{aligned}\frac{\partial L}{\partial z_3} &= \frac{\partial L}{\partial v_2}\frac{\partial v_2}{\partial z_3} \\ &= \delta_2 u(z_3-z_2)u(z_3) = \alpha_3\end{aligned} \tag{11}$$

Where $u(x) = \begin{cases} 1 & x > 0 \\ 0 & x \le 0 \end{cases}$.

3.4

$$\begin{aligned}\frac{\partial L}{\partial k_1} &= \frac{\partial L}{\partial z_1}\frac{\partial z_1}{\partial k_1} + \frac{\partial L}{\partial z_2}\frac{\partial z_2}{\partial k_1} + \frac{\partial L}{\partial z_3}\frac{\partial z_3}{\partial k_1} \\ &= \alpha_1\frac{\partial z_1}{\partial k_1} + \alpha_2\frac{\partial z_2}{\partial k_1} + \alpha_3\frac{\partial z_3}{\partial k_1} \\ &= \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3\end{aligned} \tag{12}$$

$$\begin{aligned}\frac{\partial L}{\partial k_2} &= \frac{\partial L}{\partial z_1}\frac{\partial z_1}{\partial k_2} + \frac{\partial L}{\partial z_2}\frac{\partial z_2}{\partial k_2} + \frac{\partial L}{\partial z_3}\frac{\partial z_3}{\partial k_2} \\ &= \alpha_1\frac{\partial z_1}{\partial k_2} + \alpha_2\frac{\partial z_2}{\partial k_2} + \alpha_3\frac{\partial z_3}{\partial k_2} \\ &= \alpha_1 x_2 + \alpha_2 x_3 + \alpha_3 x_4\end{aligned} \tag{13}$$

$$\begin{aligned}\frac{\partial L}{\partial k_3} &= \frac{\partial L}{\partial z_1}\frac{\partial z_1}{\partial k_3} + \frac{\partial L}{\partial z_2}\frac{\partial z_2}{\partial k_3} + \frac{\partial L}{\partial z_3}\frac{\partial z_3}{\partial k_3} \\ &= \alpha_1\frac{\partial z_1}{\partial k_3} + \alpha_2\frac{\partial z_2}{\partial k_3} + \alpha_3\frac{\partial z_3}{\partial k_3} \\ &= \alpha_1 x_3 + \alpha_2 x_4 + \alpha_3 x_5\end{aligned} \tag{14}$$

$$\begin{aligned}\frac{\partial L}{\partial b} &= \frac{\partial L}{\partial z_1}\frac{\partial z_1}{\partial b} + \frac{\partial L}{\partial z_2}\frac{\partial z_2}{\partial b} + \frac{\partial L}{\partial z_3}\frac{\partial z_3}{\partial b} \\ &= \alpha_1\frac{\partial z_1}{\partial b} + \alpha_2\frac{\partial z_2}{\partial b} + \alpha_3\frac{\partial z_3}{\partial b} \\ &= \alpha_1 + \alpha_2 + \alpha_3\end{aligned} \tag{15}$$

3.5

$$\begin{aligned}\frac{\partial L}{\partial k_j} &= \sum_{i=1}^{m} \frac{\partial L}{\partial z_i} \frac{\partial z_i}{\partial k_j} \\ &= \sum_{i=1}^{m} \alpha_i \frac{\partial z_i}{\partial k_j}\end{aligned} \tag{16}$$

$$\frac{\partial z_i}{\partial k_j} = x_{i+j-1} \tag{17}$$

$$\frac{\partial L}{\partial k_j} = \sum_{i=1}^{m} \alpha_i x_{i+j-1} \tag{18}$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^{m} \alpha_i \tag{19}$$

4

4.1

$$\lfloor \frac{205-k}{3} \rfloor + 1 = 66 \tag{20}$$

$k = 8, 9, 10$ are all possible values of k.

4.2

Assuming $k = 9$, we have $9 \times 9 \times 10 \times 96 + 96 = 8,064$ learnable parameters.

4.3

We have $9 \times 9 \times 10 = 810$ multiplications for each convolution. We have 65×65 of those convolutions. So we have $810 \times 65 \times 65 = 4,985,250$ multiplications in total.