Homework 4: CNN

(Due Date: Jul 15, 2025)

Question 1: Convolution Operation

Convolution Operation (No Pooling, Stride = 1)

Given the input matrix (7×7 with zero-padding already included):

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 4 & 1 & 0 & 0 \\ 0 & 3 & 1 & 1 & 0 & 1 & 0 \\ 0 & 2 & 4 & 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 5 & 2 & 2 & 0 \\ 0 & 0 & 1 & 3 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and the convolution filter:

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

We compute the output using stride = 1 and valid padding, i.e., the filter slides only where it fits entirely within the input. Here, input size (n)=7, filter size (f)=3, and stride (s)=1.

$$\text{Output size} = \left\lfloor \frac{n-f}{s} \right\rfloor + 1 = \left\lfloor \frac{7-3}{1} \right\rfloor + 1 = 5$$

Resulting output matrix (5×5):

$$\begin{bmatrix} -3 & -2 & 2 & 4 & 1 \\ -7 & -1 & 6 & 4 & 1 \\ -5 & 0 & 3 & 3 & 2 \\ -5 & -5 & 1 & 5 & 4 \\ -1 & -6 & -3 & 5 & 4 \end{bmatrix}$$

For example, the top-left output value is calculated as follows:

$$\text{Position } (0,0): \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} = (0{\cdot}1+0{\cdot}0+0{\cdot}(-1))+(0{\cdot}1+0{\cdot}0+2{\cdot}(-1))+(0{\cdot}1+3{\cdot}0+1{\cdot}(-1)) = -2-1 = -3$$

Convolution Operation (No Pooling, Stride = 2)

Given the 7×7 input matrix and the 3×3 convolution filter, the output using stride = 2 is a 3×3 matrix given by:

$$\begin{bmatrix} -3 & 2 & 1 \\ -5 & 3 & 2 \\ -1 & -3 & 4 \end{bmatrix}$$

For example, let us apply the filter starting at row 0, column 2 of the input

$$\text{Position } (0,1): \quad \begin{bmatrix} 0 & 0 & 0 \\ 2 & 4 & 1 \\ 1 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} = (0{\cdot}1+0{\cdot}0+0{\cdot}(-1))+(2{\cdot}1+4{\cdot}0+1{\cdot}(-1))+(1{\cdot}1+1{\cdot}0+0{\cdot}(-1)) = 1+1 = 2$$

Convolution followed by Max Pooling (3×3 pool size, stride = 2)

The output of the convolution operation with stride = 1 is already calculated in the first part of the problem and is given by:

$$\begin{bmatrix} -3 & -2 & 2 & 4 & 1 \\ -7 & -1 & 6 & 4 & 1 \\ -5 & 0 & 3 & 3 & 2 \\ -5 & -5 & 1 & 5 & 4 \\ -1 & -6 & -3 & 5 & 4 \end{bmatrix}$$

Let's apply 3×3 max pooling on the above result with a stride of 2.

- Region starting at (0,0):

$$\begin{bmatrix} -3 & -2 & 2 \\ -7 & -1 & 6 \\ -5 & 0 & 3 \end{bmatrix} \rightarrow \max = 6$$

- Region starting at (0,2):

$$\begin{bmatrix} 2 & 4 & 1 \\ 6 & 4 & 1 \\ 3 & 3 & 2 \end{bmatrix} \rightarrow \max = 6$$

- Region starting at (2,0):

$$\begin{bmatrix} -5 & 0 & 3 \\ -5 & -5 & 1 \\ -1 & -6 & -3 \end{bmatrix} \rightarrow \max = 3$$

- Region starting at (2,2):

$$\begin{bmatrix} 3 & 3 & 2 \\ 1 & 5 & 4 \\ -3 & 5 & 4 \end{bmatrix} \rightarrow \max = 5$$

Final Output After Pooling (2×2):

$$\begin{bmatrix} 6 & 6 \\ 3 & 5 \end{bmatrix}$$