1. Task

1

R:

$$\text{conv}\left(\begin{array}{c} \text{Image} \\ \begin{vmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 2 & 4 & 2 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{vmatrix} \end{array}, \begin{array}{c} \text{Kernel} \\ \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix} \end{array} \right) =$$

$$= \begin{vmatrix} 1\cdot 0 + 4\cdot 1 & 1\cdot 0 + 2\cdot 1 & 1\cdot 0 + 1\cdot 0 \\ 1\cdot 0 + 1\cdot 2 & 1\cdot 1 + 1\cdot 1 & 2\cdot 1 + 1\cdot 0 \\ 0\cdot 1 + 0\cdot 1 & 2\cdot 1 + 1\cdot 0 & 1\cdot 0 + 4\cdot 1 \end{vmatrix} = \begin{vmatrix} 4 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 4 \end{vmatrix}$$

G:

$$\text{conv}\left(\begin{array}{c} \text{Image} \\ \begin{vmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{vmatrix} \end{array}, \begin{array}{c} \text{Kernel} \\ \begin{vmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{vmatrix} \end{array} \right) =$$

$$= \begin{vmatrix} (-1)\cdot 0 + (-1)\cdot 0 & (-1)\cdot 0 + (-1)\cdot 0 & (-1)\cdot 0 + (-1)\cdot 0 \\ (-1)\cdot 0 + (-1)\cdot 0 & (-1)\cdot 1 + (-1)\cdot 1 & (-1)\cdot 0 + (-1)\cdot 0 \\ (-1)\cdot 0 + (-1)\cdot 0 & (-1)\cdot 0 + (-1)\cdot 0 & (-1)\cdot 0 + (-1)\cdot 0 \end{vmatrix} = \begin{vmatrix} 0 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{vmatrix}$$

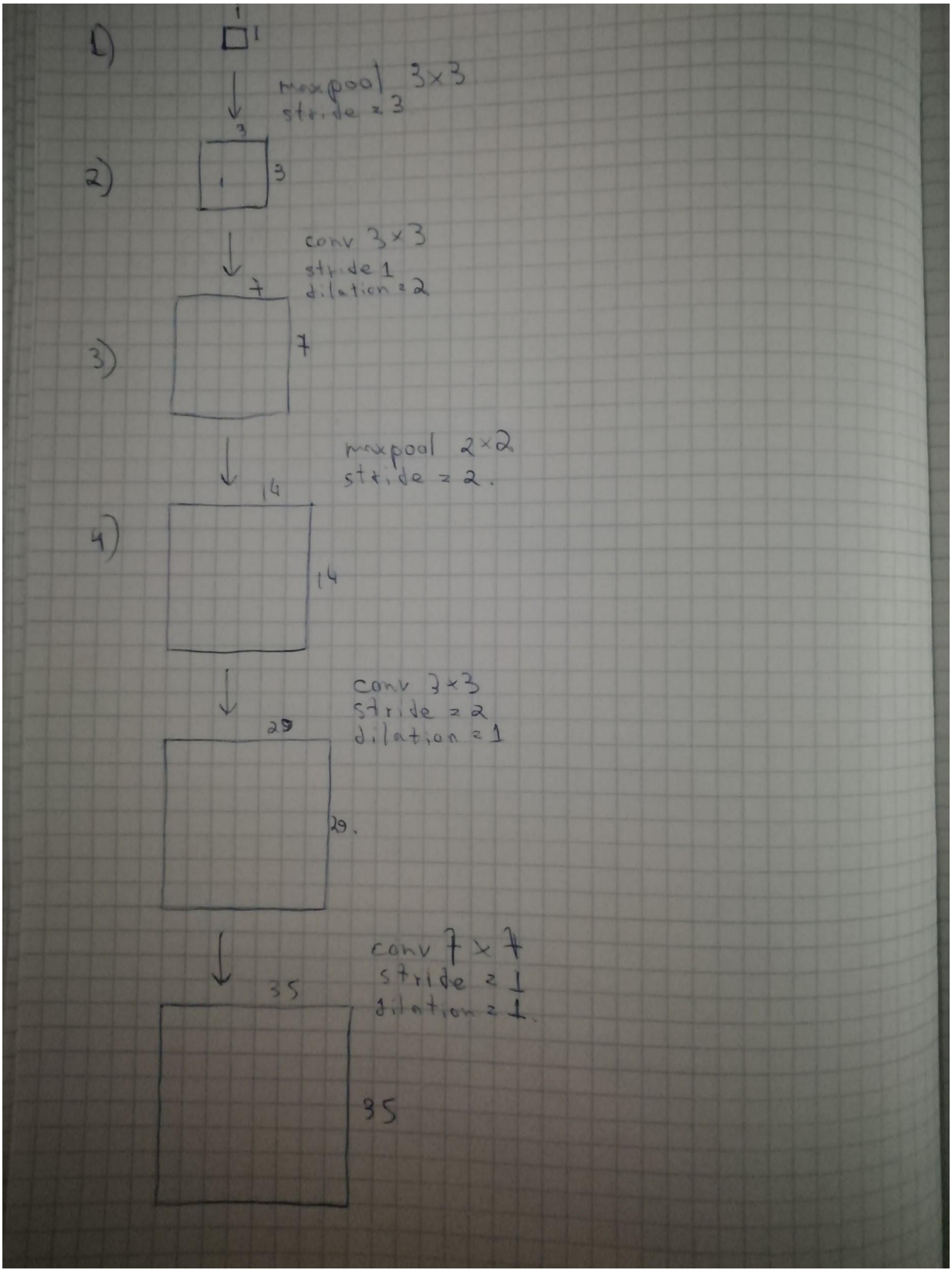
B:

Image

$$\text{conv}\left(\begin{vmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{vmatrix} , \begin{vmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix} \right) = Z$$

$$Z = \begin{vmatrix} 1\cdot 0 + (-1)\cdot 0 & 1\cdot 0 + (-1)\cdot 0 & 1\cdot 0 + (-1)\cdot 0 \\ 1\cdot 0 + (-1)\cdot 1 & 1\cdot 0 + (-1)\cdot 0 & 1\cdot 1 + (-1)\cdot 0 \\ 1\cdot 0 + (-1)\cdot 2 & 1\cdot 1 + (-1)\cdot 1 & 1\cdot 2 + (-1)\cdot 0 \end{vmatrix} = \begin{vmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ -2 & 0 & 2 \end{vmatrix}$$

2. Task



$$receptive\ field = 35 \times 35$$

3. Task

Посчитаем итоговый размер изображения по формуле:

$n_{res} = \lfloor \frac{n_{img} + 2p - d(n_{kern} - 1) - 1}{s} \rfloor + 1$, где

$n_{res} -$ размер итогового изображения,

$n_{img} -$ размер входного изображения,

$n_{kern} -$ размер ядра,

$p -$ размер паддинга,

$d -$ размер $dilation$,

$s -$ размер $stride$.

Найдем итоговый размер изображения:

1) $n_{res} = 224 \times 224,\ conv\, 7 \times 7\ (p = 3,\ s = 2,\ d = 1)$:

   $n_{res} = \lfloor \frac{224 + 2\cdot3 - 1\cdot(7-1) - 1}{2} \rfloor + 1 = 112$
2) $n_{res} = 112 \times 112,\ conv\, 3 \times 3\ (p = 1,\ s = 1,\ d = 1)$:

   $n_{res} = \lfloor \frac{112 + 2\cdot1 - 1\cdot(3-1) - 1}{1} \rfloor + 1 = 112$
3) $n_{res} = 112 \times 112,\ maxpool\, 2 \times 2\ (p = 0,\ s = 2)$:

   $n_{res} = \lfloor \frac{112 + 2\cdot0 - 1\cdot(2-1) - 1}{2} \rfloor + 1 = 56$
4) $n_{res} = 56 \times 56,\ conv\, 3 \times 3\ (p = 1,\ s = 2,\ d = 3)$:

   $n_{res} = \lfloor \frac{56 + 2\cdot1 - 3\cdot(3-1) - 1}{2} \rfloor + 1 = 26$

Ответ: итоговый размер изображения $= 26 \times 26$.

4. Task

функция Dropout выглядит как: $f = \frac{a\cdot mask}{p-1}$, где:

$a -$ входные значения,

$mask -$ бинарная маска, по которой деактивируются нейроны с вероятностью p,

$p -$ вероятность деактивации нейрона.

Посчитаем производную по a:

$$\frac{\partial f}{\partial a} = \frac{\partial(\frac{a \cdot mask}{p-1})}{\partial a} = \frac{mask}{p-1}$$

5. Task

функция Batchnorm1d выглядит как: $f_i = \frac{x_i - EX}{\sqrt{DX + \varepsilon}} \cdot \gamma + \beta$, где:

$x_i -$ входное значение,

$EX -$ выборочное среднее от входных значений,

$DX -$ выборочная дисперсия от входных значений,

$\varepsilon -$ значение, добавляемое к знаменателю, чтобы избежать деления на 0,

$\gamma, \beta -$ гиперпараметры, для обучения модели

Посчитаем производную по x_i:

$$\frac{\partial f_i}{\partial x_i} = \frac{\partial(\frac{x_i - EX}{\sqrt{DX + \varepsilon}} \cdot \gamma + \beta)}{\partial x_i} = \gamma \cdot \frac{\partial(\frac{x_i - EX}{\sqrt{DX + \varepsilon}})}{\partial x_i} =$$

$$= \gamma \cdot \frac{\frac{\partial(x_i - EX)}{\partial x_i} \cdot \sqrt{DX+\varepsilon} - \frac{x_i - EX}{2\sqrt{DX+\varepsilon}} \cdot \frac{\partial(DX+\varepsilon)}{\partial x_i}}{DX+\varepsilon} =$$

$$= \gamma \cdot \frac{\frac{n-1}{n} \cdot \sqrt{DX+\varepsilon} - \frac{x_i - EX}{2\sqrt{DX+\varepsilon}} \cdot \frac{\partial(DX+\varepsilon)}{\partial x_i}}{DX+\varepsilon} =$$

$$= \gamma \cdot \frac{\frac{n-1}{n} \cdot \sqrt{DX+\varepsilon} - \frac{x_i - EX}{2\sqrt{DX+\varepsilon}} \cdot \frac{2(x_i - EX)}{n}}{DX+\varepsilon} =$$

$$= \gamma \cdot \frac{\frac{n-1}{n} \cdot \sqrt{DX+\varepsilon} - \frac{(x_i - EX)^2}{n\sqrt{DX+\varepsilon}}}{DX+\varepsilon} = \gamma \cdot \left(\frac{n-1}{n\sqrt{DX+\varepsilon}} - \frac{(n-1) \cdot (x_i - EX)^2}{n^2 (DX+\varepsilon)^{\frac{3}{2}}}\right)$$