Rotitehtävä 4

(1) a)

$$\ell(w) = \sum \ln\left(1 + \exp(-y_n w^T x_n)\right)$$

$$\frac{d\ell(w)}{dw_1} = \sum_{n=0}^{N-1} \frac{\exp(-y_n w^T x_n)(-y_n)}{1+\exp(-y_n w^T x_n)} \begin{bmatrix} x_1 \\ 0 \end{bmatrix} = \text{HH} \cdot \begin{bmatrix} x_1 \\ x_n \end{bmatrix}$$

$$\frac{d\ell(w)}{dw_2} = \sum_{n=0}^{N-1} \frac{\exp(-y_n w^T x_n)(-y_n)}{1+\exp(-y_n w^T x_n)} \begin{bmatrix} 0 \\ x_1 \end{bmatrix}$$

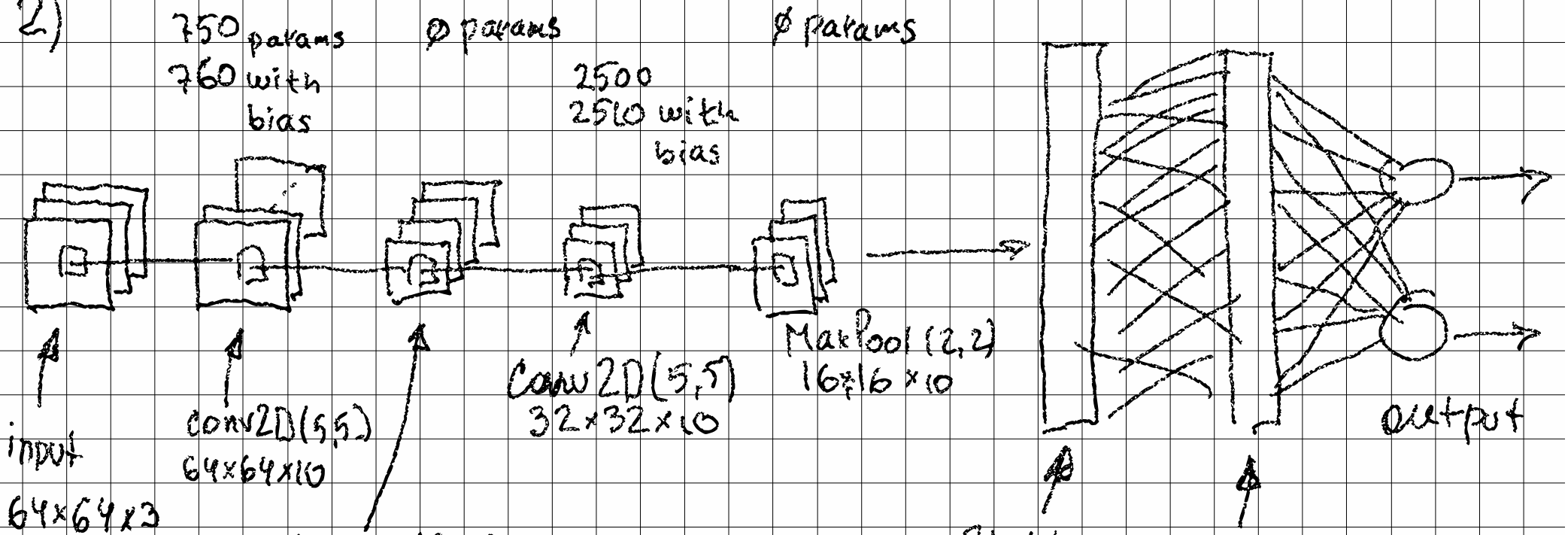
b)

$$\ell(w) = \sum \ln\left(1+\exp(-y_n w^T x_n)\right) + C w^T w$$

$$\frac{d\ell(w)}{dw} = \sum_{n=0}^{N-1} \frac{\exp(-y_n w^T x_n)}{1+\exp(-y_n w^T x_n)} (-y_n) X + 2Cw$$

$$\underset{1\times 3}{[w_1 + w_2 + w_3]} \underset{3\times 1}{\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}} = w_1^2 + w_2^2 + w_3^2 = \begin{matrix} 2w_1 \\ 2w_2 \\ 2w_3 \end{matrix}$$

2)



25×10×3

2560 params [not to be learned]

params 5120

5122 with bias