

FYS5429/9429, Lecture Feb 6, 2025

FYS5429/9429 February 6, 2025

optimization

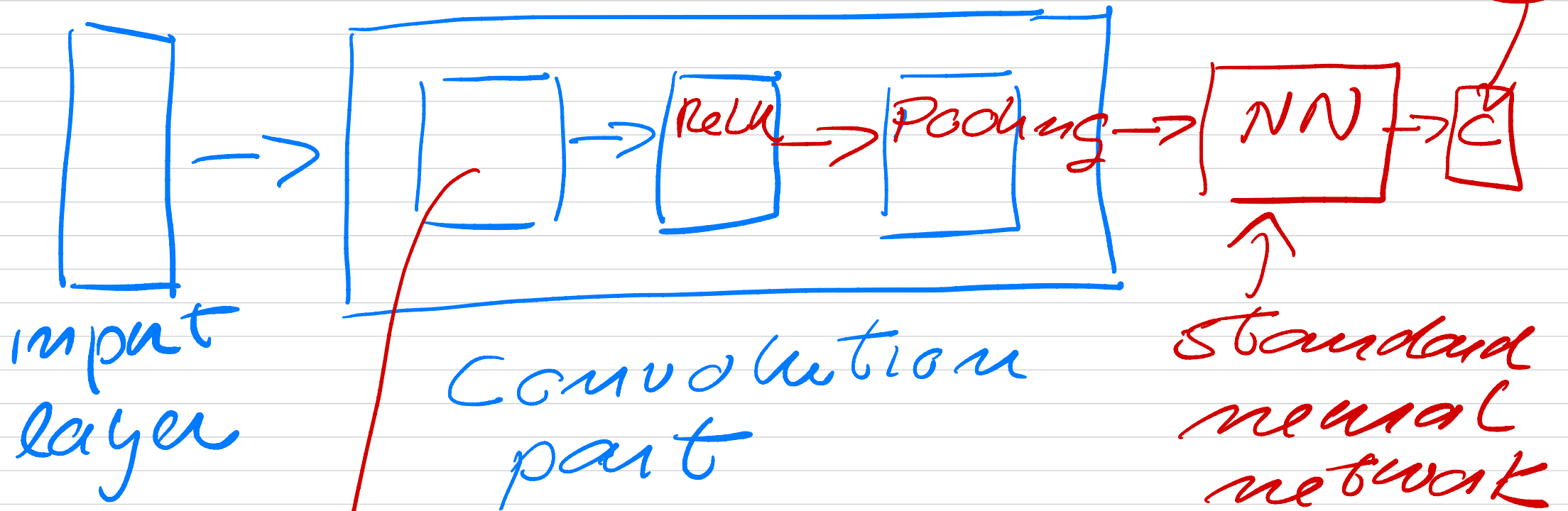
$$\Theta^{(n+1)} = \Theta^{(n)} - \underbrace{\left(\nabla^2 C(\Theta^{(n)}) \right)^{-1}}_{\substack{H(\Theta^{(n)}) \\ \mu^{-1}}} \nabla C(\Theta^{(n)})$$

$(f'(x))^{-1} f(x)$

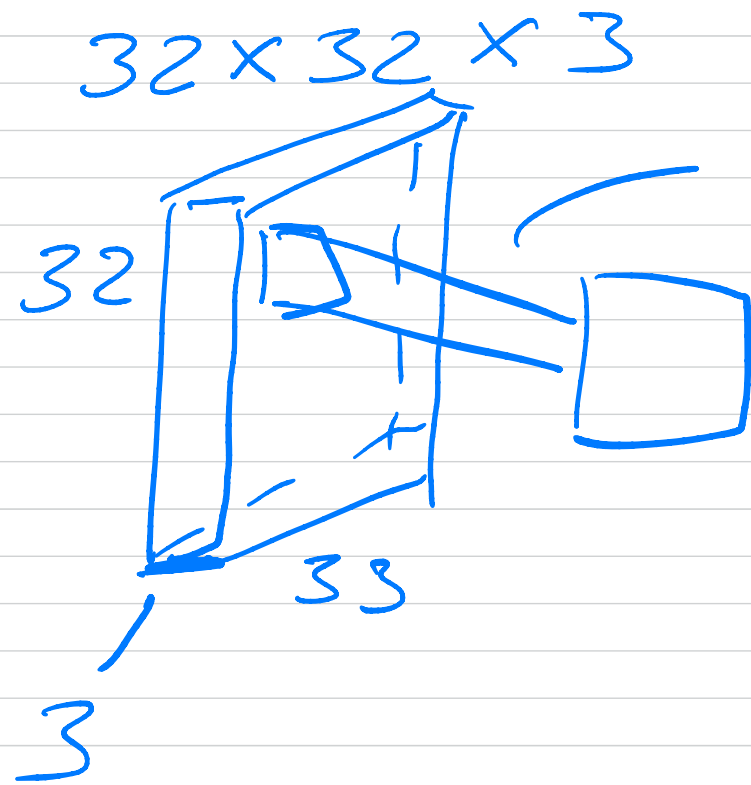
$-\mu^{(n)} \nabla C(\Theta^{(n)})$

$$\mu = [10^{-5}, 10^{-4}, \dots, 10^{-1}]$$

CNNs overarching view



parameters to train [Convolution stage = a set of filters which reduces dim of original data



Filter $5 \times 5 \times 3$

$[]$ $[]$...

10
new
down
scaled
images

Size of filter F

typical values are 3 and 5

Stride S is another

$S = 1$ and 2

padding $P = 1$ and 2
common settings for CNN
filter

$$F = 3$$

$$S = 1$$

$$P = 1$$

$$F = 5$$

$$S = 1$$

$$P = 2$$

$$F = 5$$

$$S = 2$$

$$P = ?$$