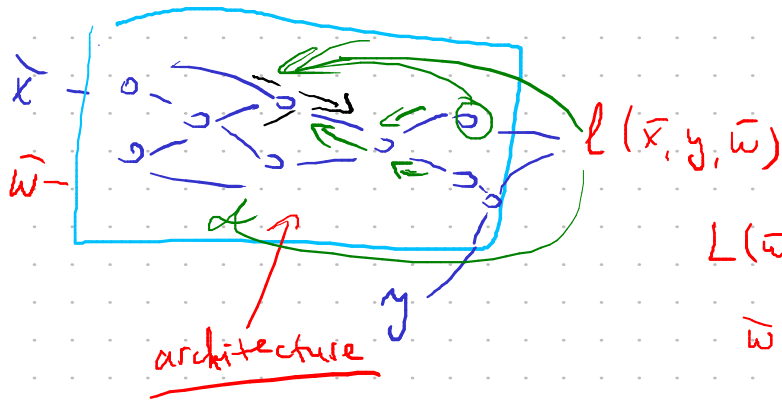
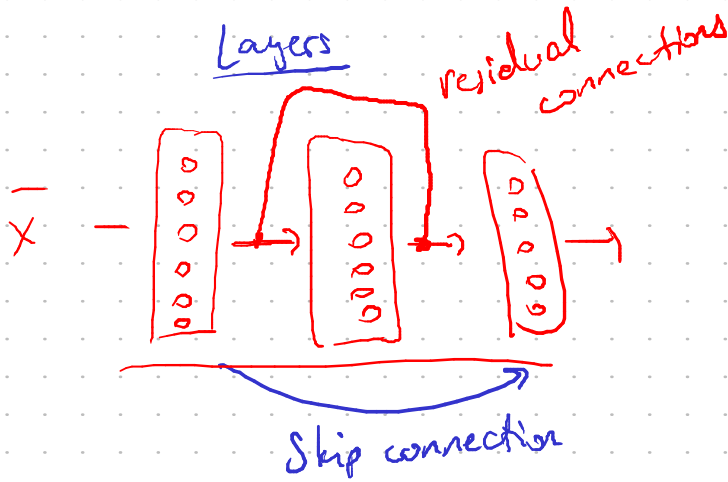


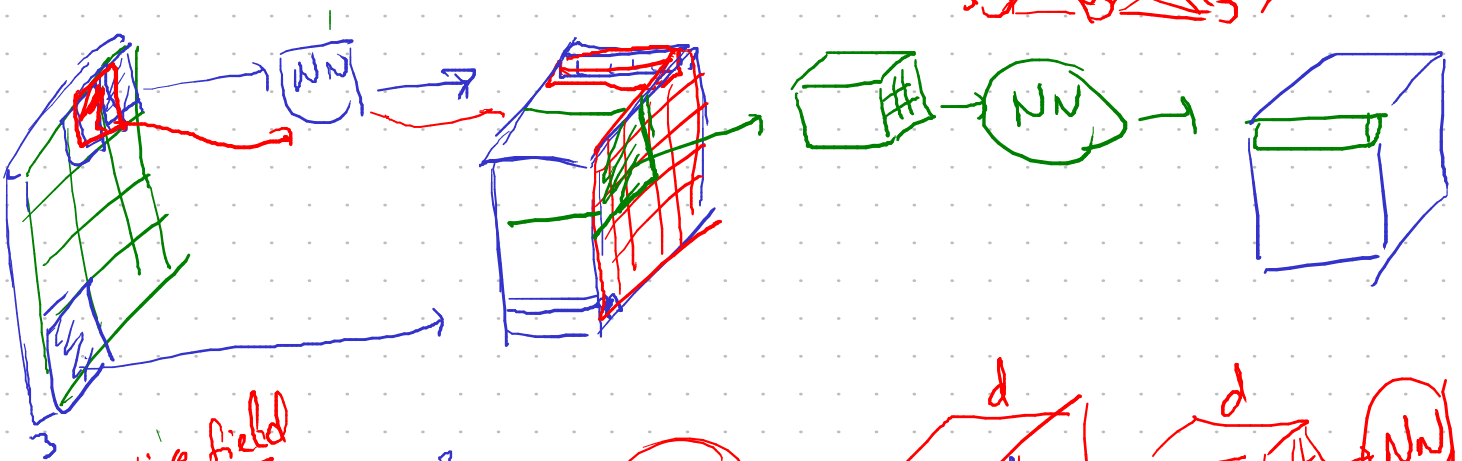
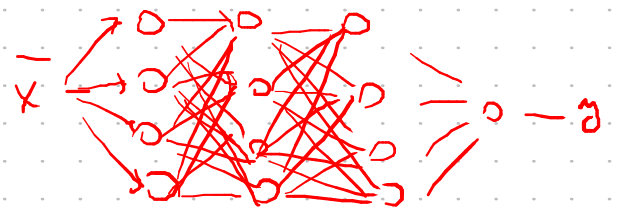


$$L(\bar{w}) = \frac{1}{N} \sum_n \ell(\bar{x}_n, y_n, \bar{w})$$

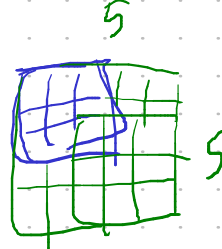
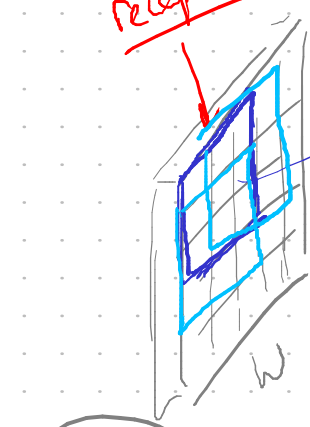
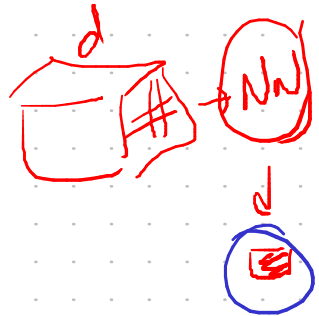
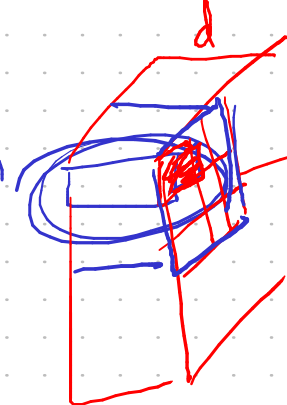
$$\bar{w} := \bar{w} - \alpha \cdot \nabla_w L(\bar{w})$$



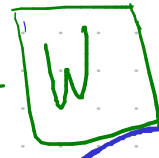
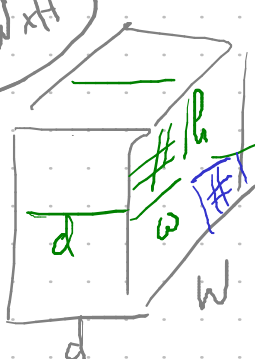
Feedforward
fully connected networks



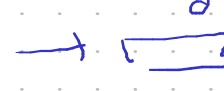
receptive field



$d \times w \times h$



$$\bar{z}W + \bar{a} \in \mathbb{R}^{d'}$$

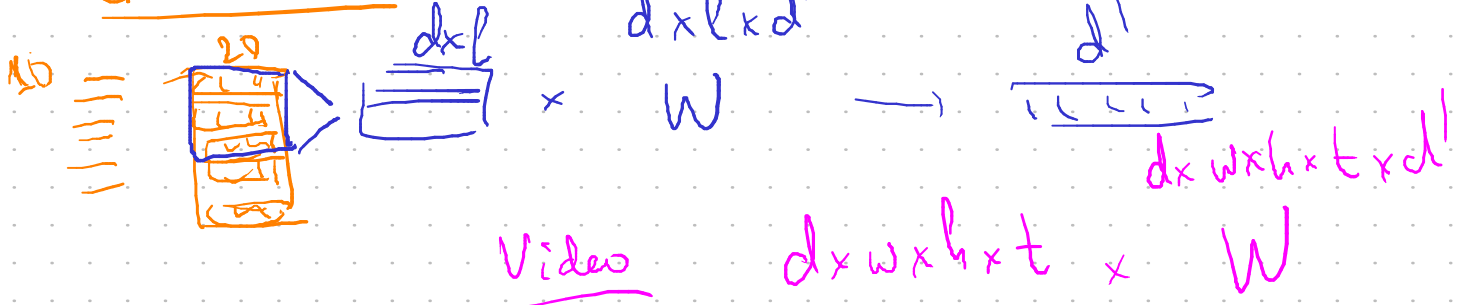




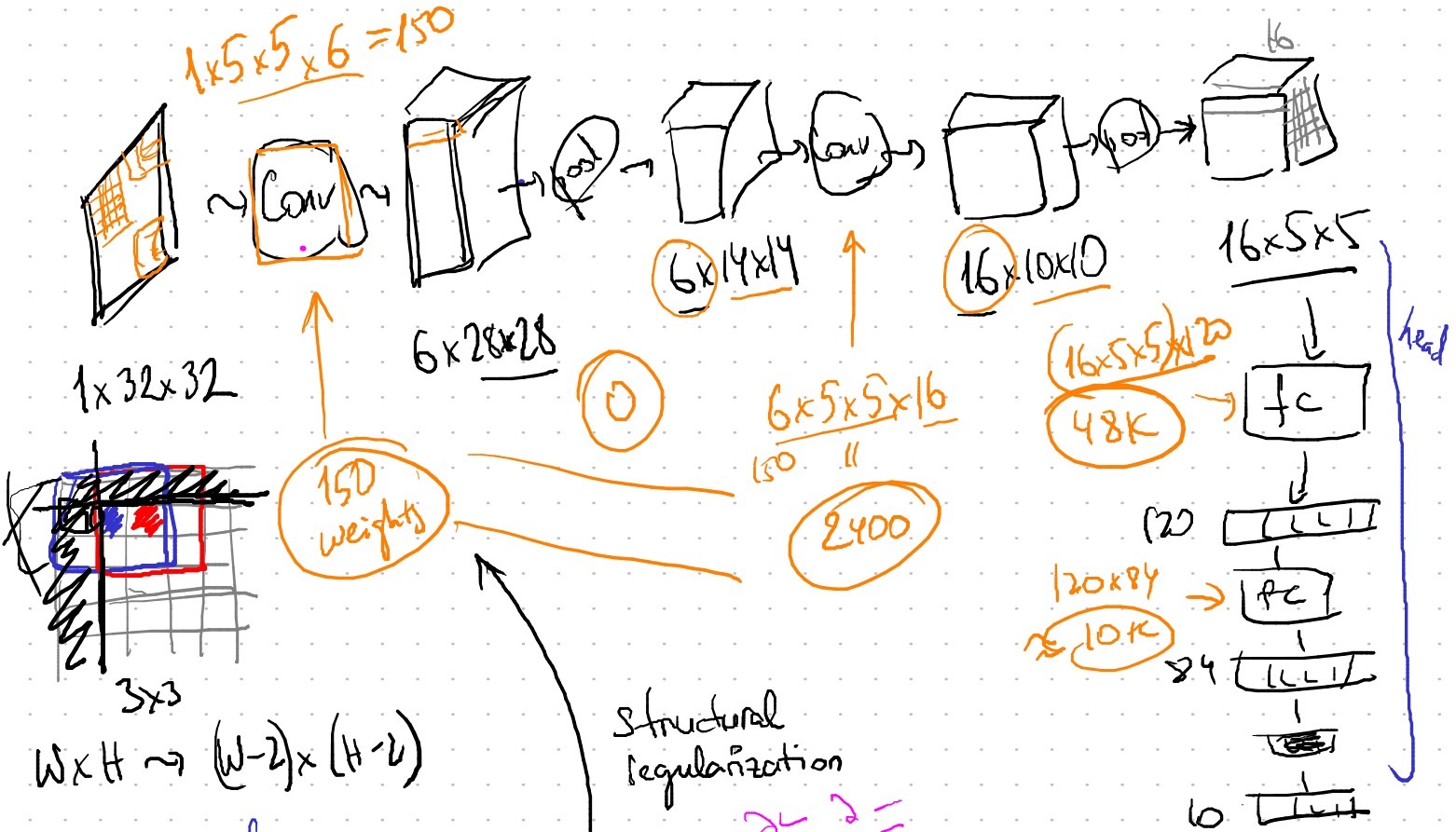
$W:$
 $d \times w \times h \times d'$

CLUT

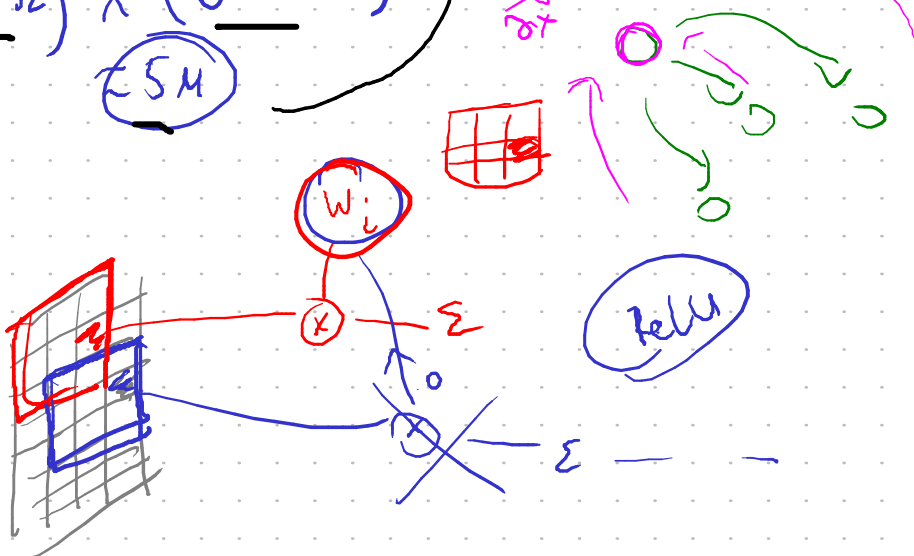
CNNs for text



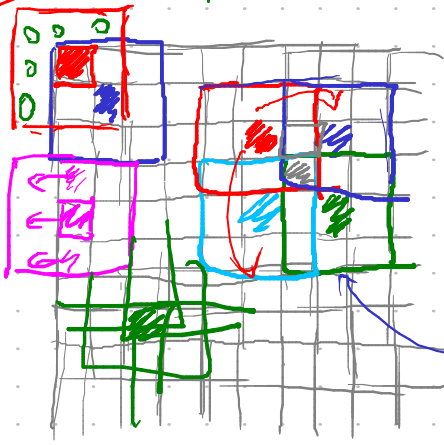
$1 \times 5 \times 5 \times 6 = 150$



$\frac{\partial L}{\partial x} = \sum \frac{\partial L}{\partial x}$



padding zero padding



$$d \times W \times H \rightarrow \text{Conv} \rightarrow d \times (W - k) \times (H - k)$$



stride 2

Conv
Layer:

$d \times w \times h \times d'$
padding

stride (a, b)