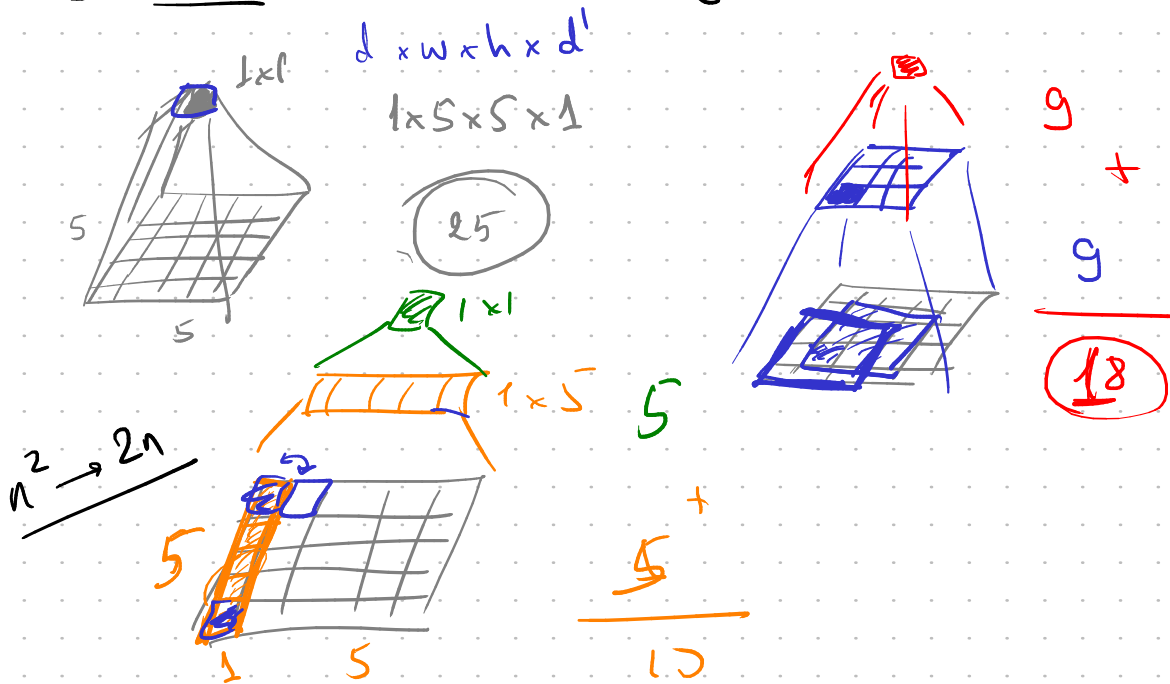
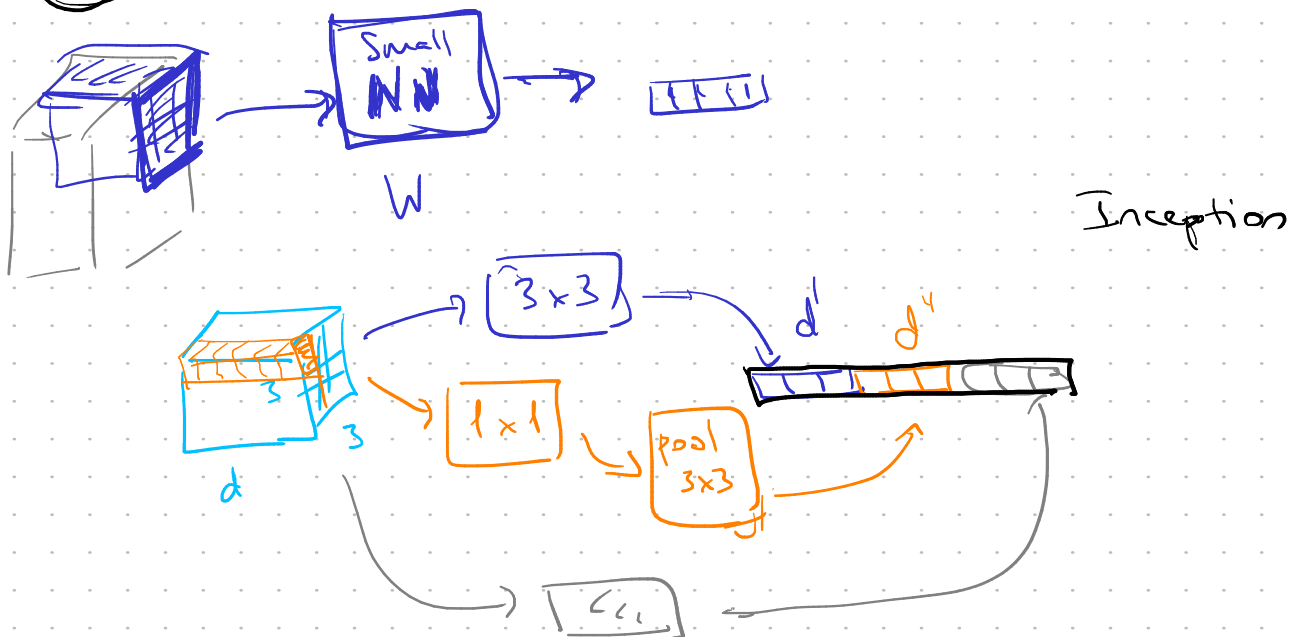


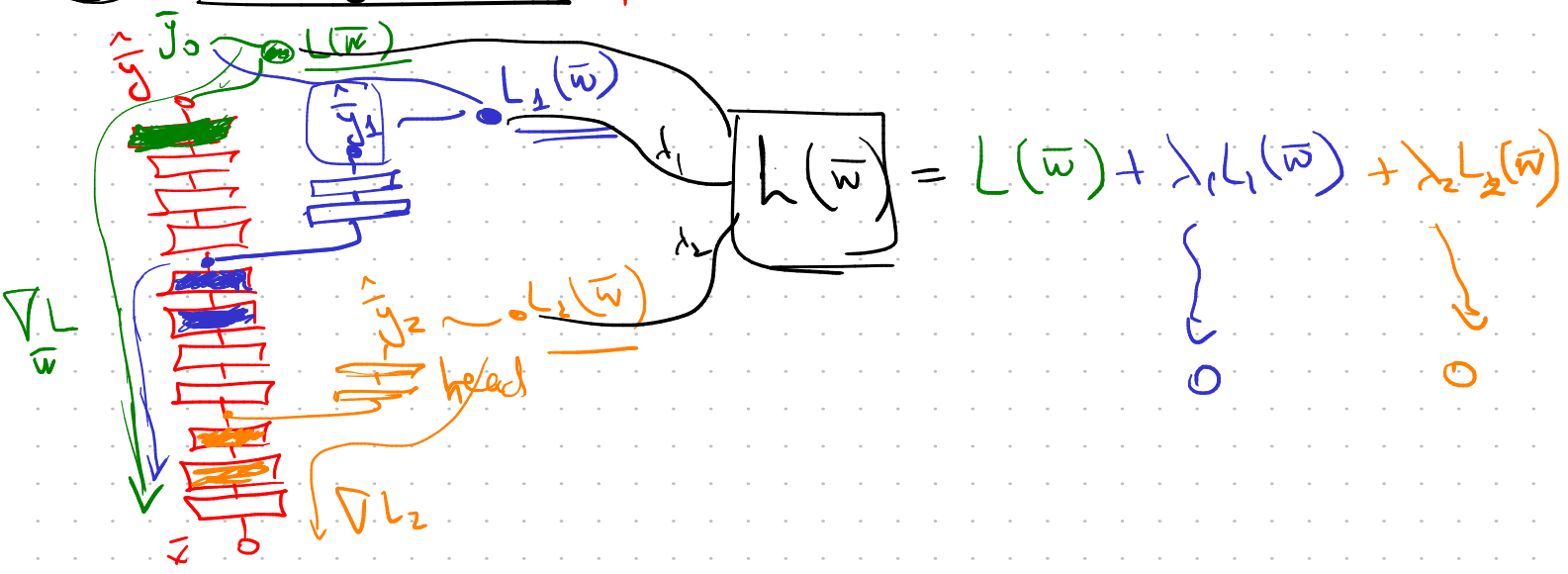
① VGG - Visual Geometry Group / Oxford



② Network in network

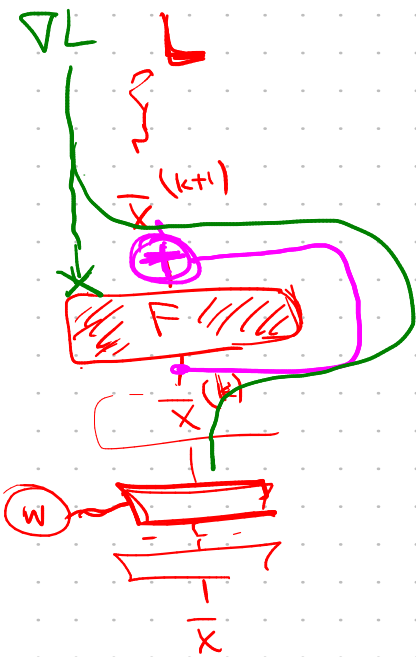


③ Auxiliary classifiers / GoogleNet



④ Residual connections / ResNet

Kaiming He



$$\bar{x}^{(k+1)} = F(\bar{x}^{(k)})$$

$$F \sim \bar{x}^{(k)} \mapsto \bar{x}^{(k+1)}$$

$$\frac{\partial L}{\partial w} = \frac{\partial L}{\partial \bar{x}^{(k+1)}} \cdot \frac{\partial \bar{x}^{(k+1)}}{\partial w} =$$

$$= \frac{\partial L}{\partial \bar{x}^{(k+1)}} \cdot \boxed{\frac{\partial \bar{x}^{(k+1)}}{\partial \bar{x}^{(k)}}} \cdot \frac{\partial \bar{x}^{(k)}}{\partial w}$$

$F' \rightarrow 0$

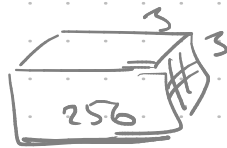
$$\bar{x}^{(k+1)} = F(\bar{x}^{(k)}) + \bar{x}^{(k)}$$

$$F \sim \bar{x}^{(k)} \mapsto \bar{x}^{(k+1)} - \bar{x}^{(k)} \quad \text{-- residue}$$

$$\frac{\partial \bar{x}^{(k+1)}}{\partial \bar{x}^{(k)}} = \cancel{F'} + I$$

$\downarrow \rightarrow 0$

⑤ Bottleneck layers



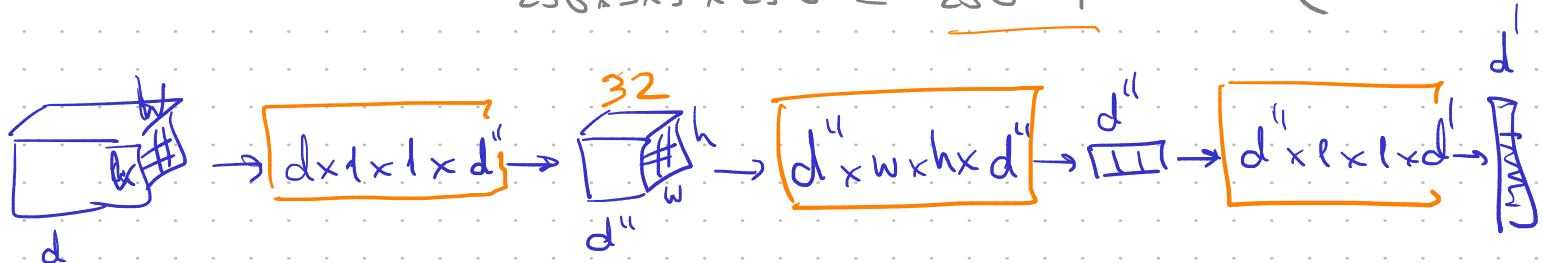
$$d \times w \times h$$

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

$$d'$$

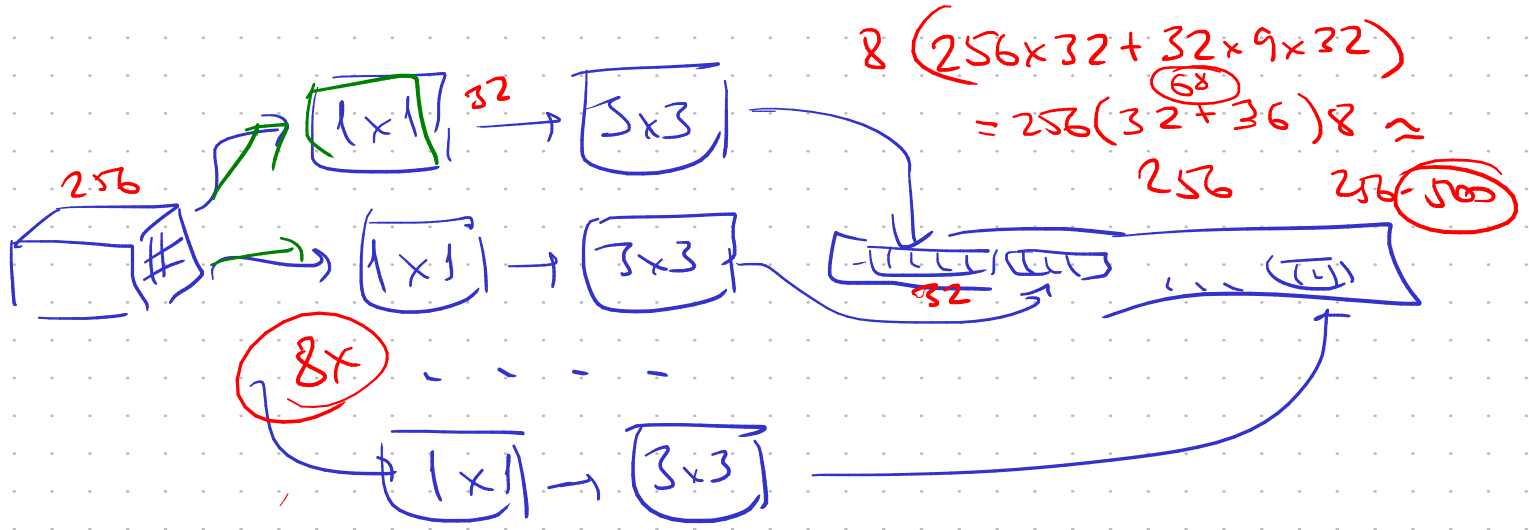
$$2500$$

$$256 \times 3 \times 3 \times 256 = 256^2 \cdot 9 = 256 \cdot (256 \cdot 9)$$



$$256 \times 32 + 32 \times 3 \times 3 \times 32 + 32 \times 256 =$$

$$= 256(32 + 36 + 32) = 256(100)$$



$$\begin{bmatrix} \vdots \\ \vdots \\ \vdots \end{bmatrix} \in \mathbb{R}^{256}$$

Split - transform - merge
ResNeXt

⑥ EfficientNet

NAS - neural architecture search

- depth d / layers $= \alpha^p$
 - width w / channels $= \beta^p$
 - resolution r / input size $= \gamma^p$
- Compound Scaling

$$\boxed{\text{FLOPs}} \propto d \cdot w^2 \cdot r^2 = (\underbrace{\alpha \cdot \beta^2 \cdot \gamma^2}_c)^p \dots \beta_1, \beta_2, \dots$$

NAS :

$$\begin{cases} \alpha, \beta, \gamma \geq 1 \\ \alpha \cdot \beta^2 \cdot \gamma^2 = c \end{cases} \sim (\alpha^*, \beta^*, \gamma^*) \quad \text{BO}$$