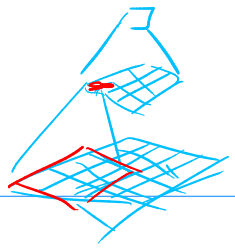
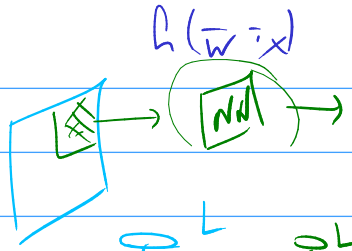
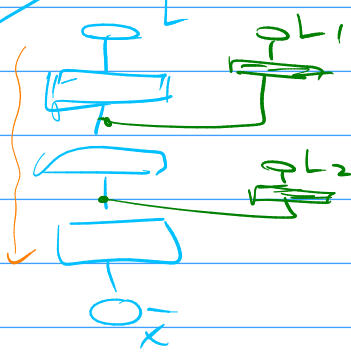
1. VGG



2. Network in network
   GoogleNet



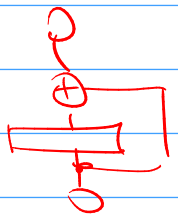
3. Auxiliary classifiers



$h = L + \lambda_1 L_1 + \lambda_2 L_2$

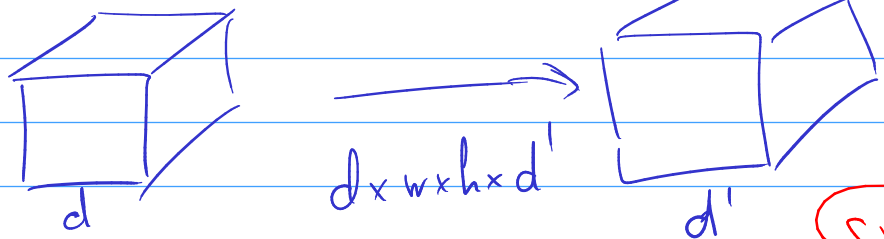
$\lambda_1 \downarrow 0 \quad \lambda_2 \downarrow 0$

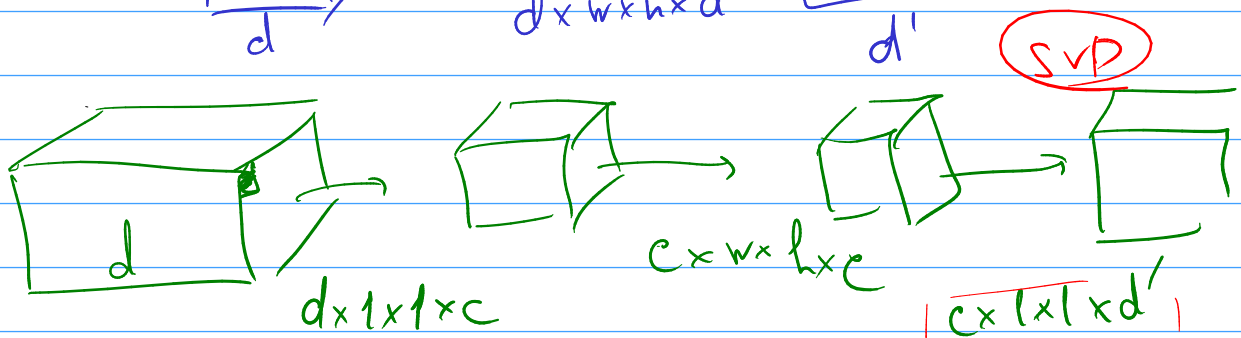
4. Residual connections



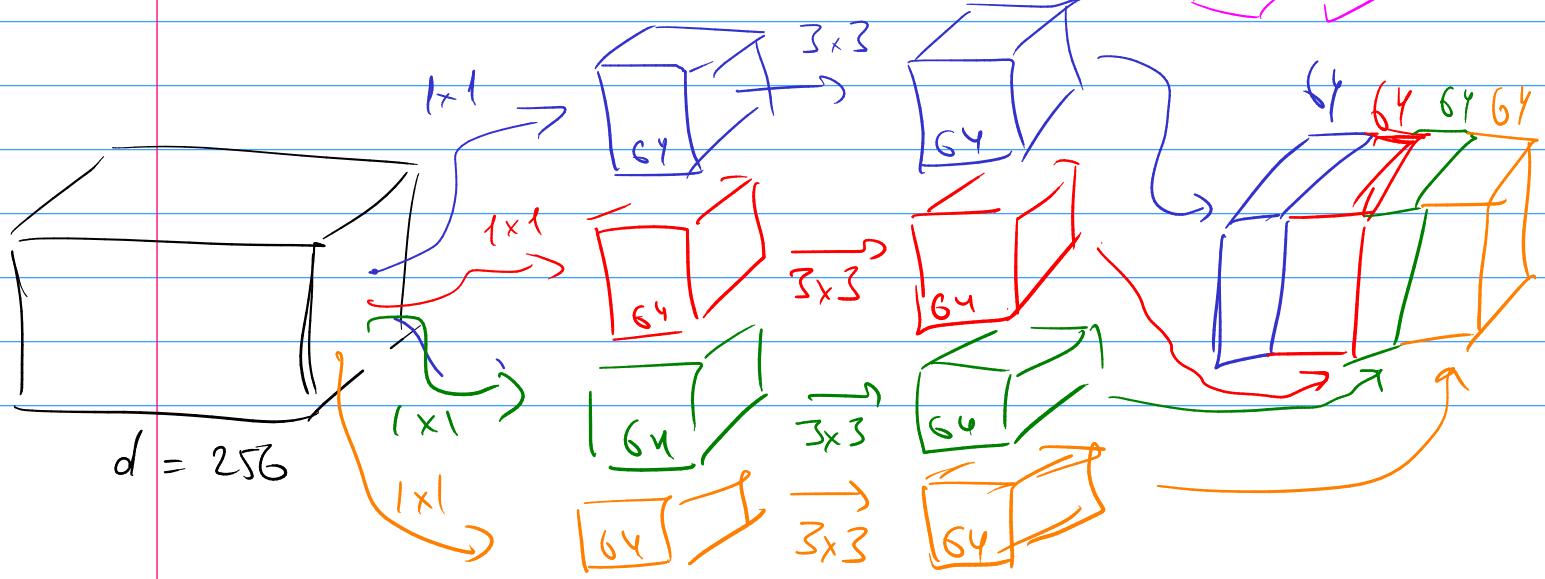
$F(\bar{x}) + \bar{x}$

5. Bottlenecks

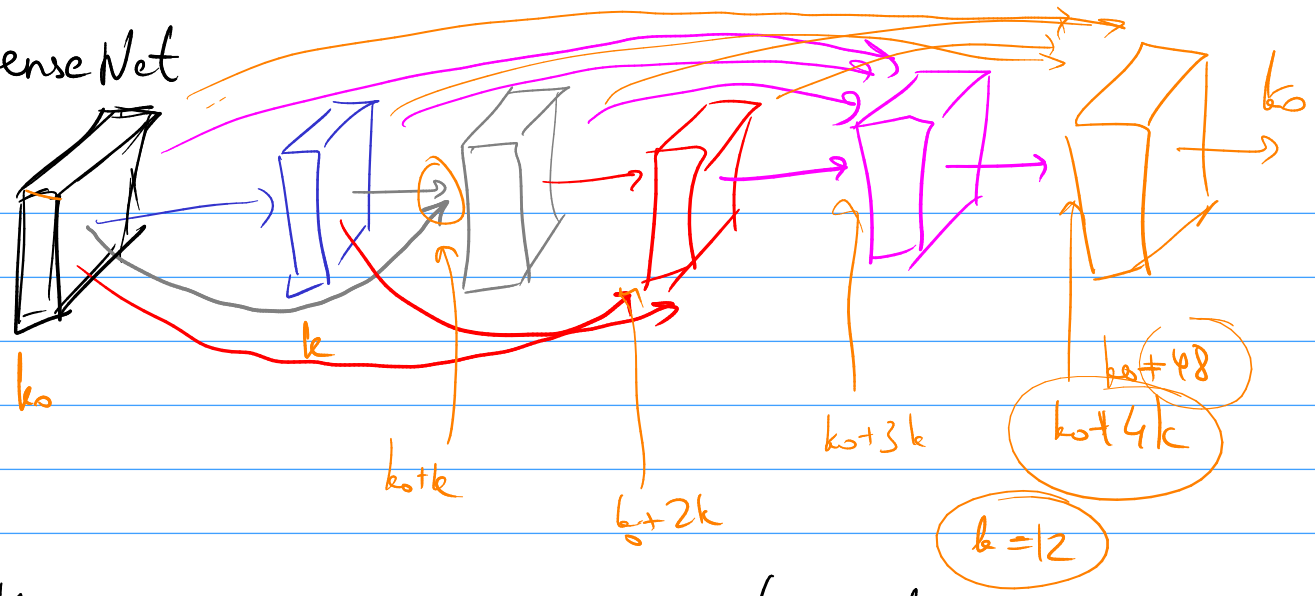




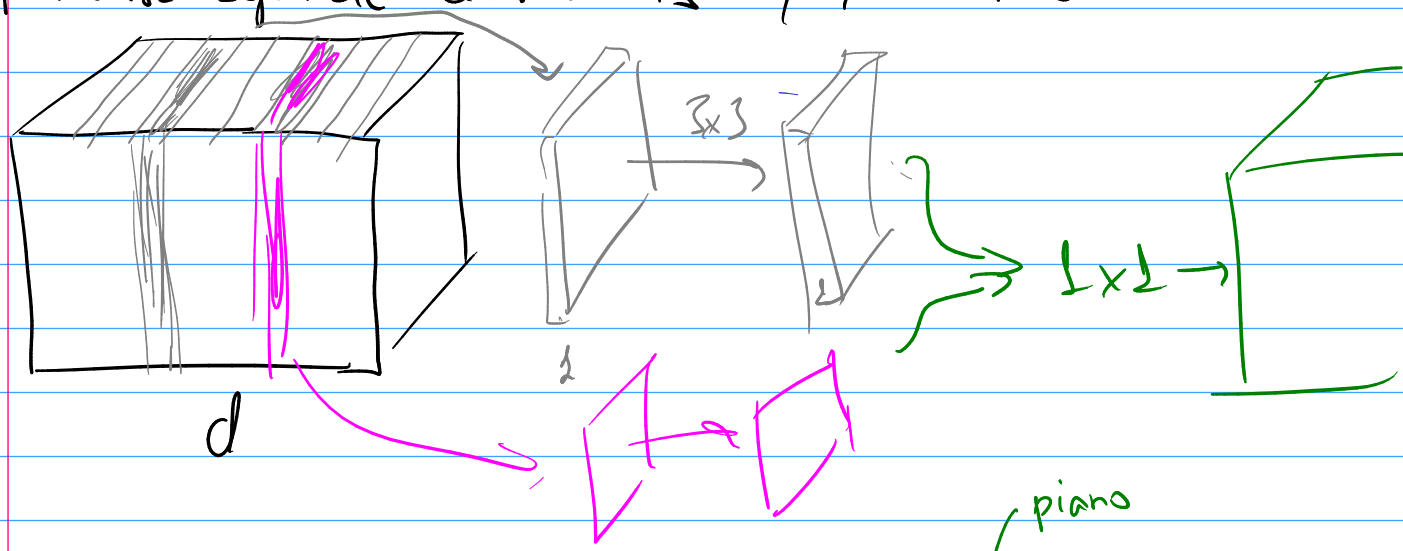
6. Squeeze - then - expand



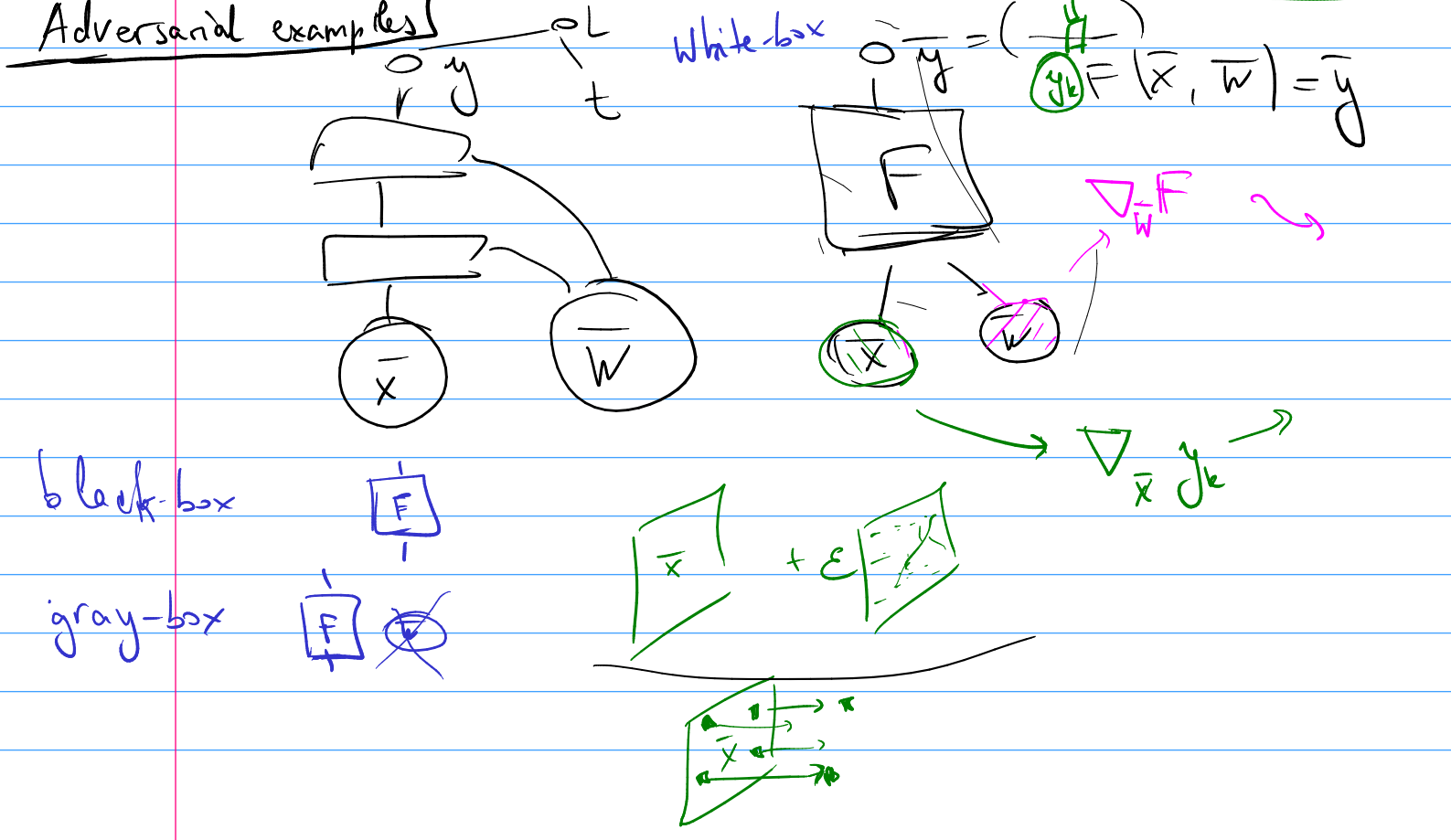
(7) DenseNet

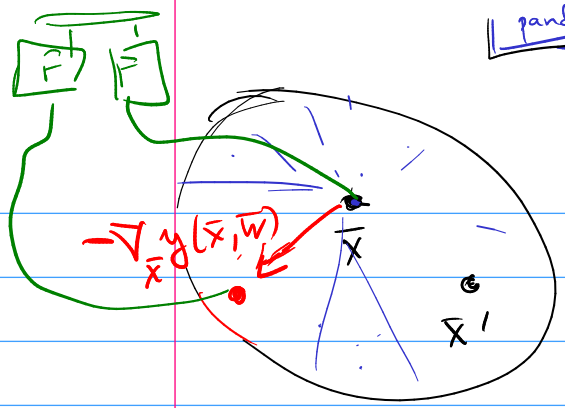


(8) Depthwise separable convolutions / MobileNet



Adversarial examples





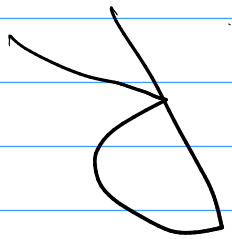
panda

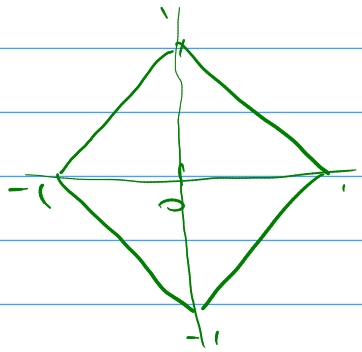
$\|\bar{x} - \bar{x}'\| < \varepsilon$

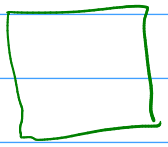
$y(\bar{x}') = \text{"panda"}$

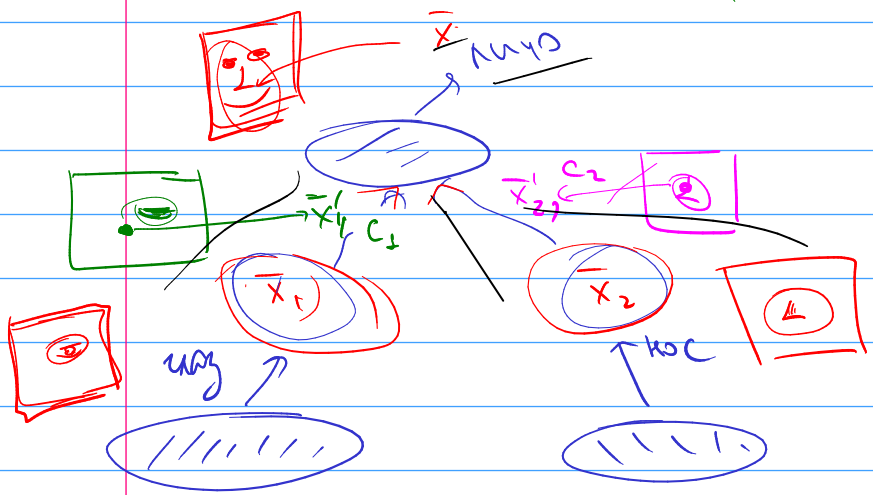
$\bar{x}, \bar{x} + \bar{\varepsilon}_1, x + \bar{\varepsilon}_2$ ↪ random

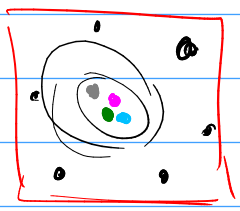
$$h(\bar{w}, \bar{x}, y) = \alpha L(\bar{w}, \bar{x}, y) + (1-\alpha) L(\bar{w}, \bar{x} - \varepsilon \nabla_{\bar{x}} l_y(\bar{w}, \bar{x}, y), y)$$



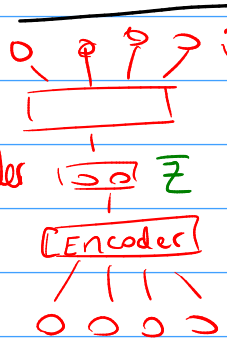




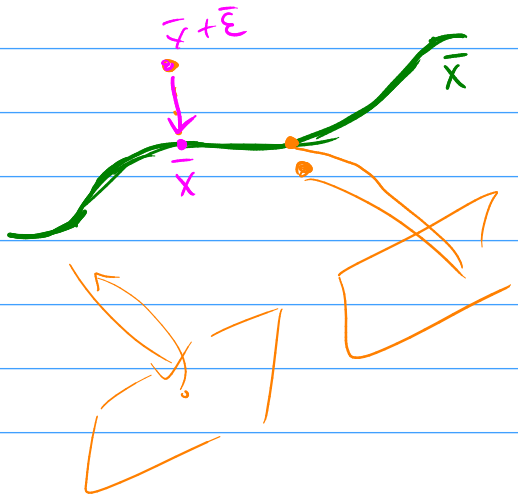




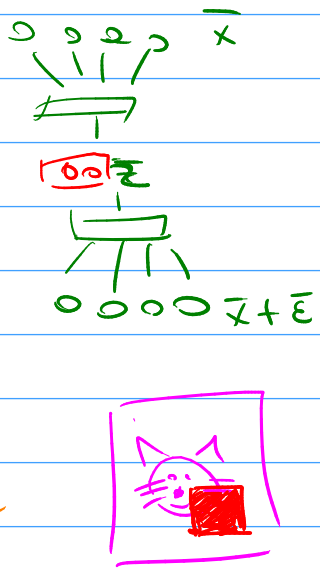
Autoencoder



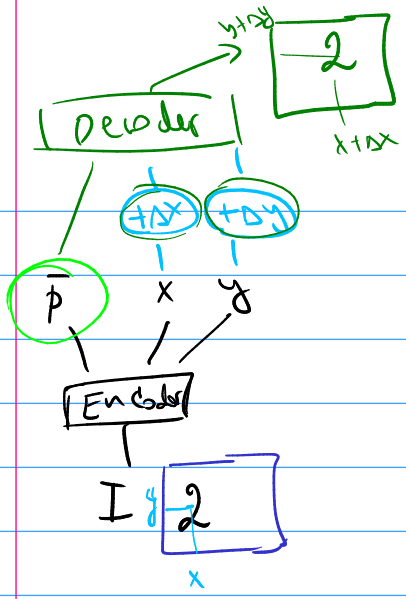
Denoising autoencoders



Disentanglement



VAE


Decoder
2
+Δx
+Δy
p̄
x
y
Encoder
I
2
x