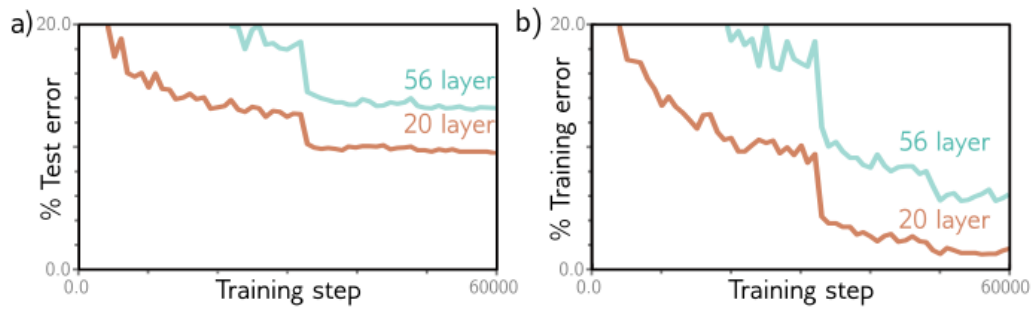
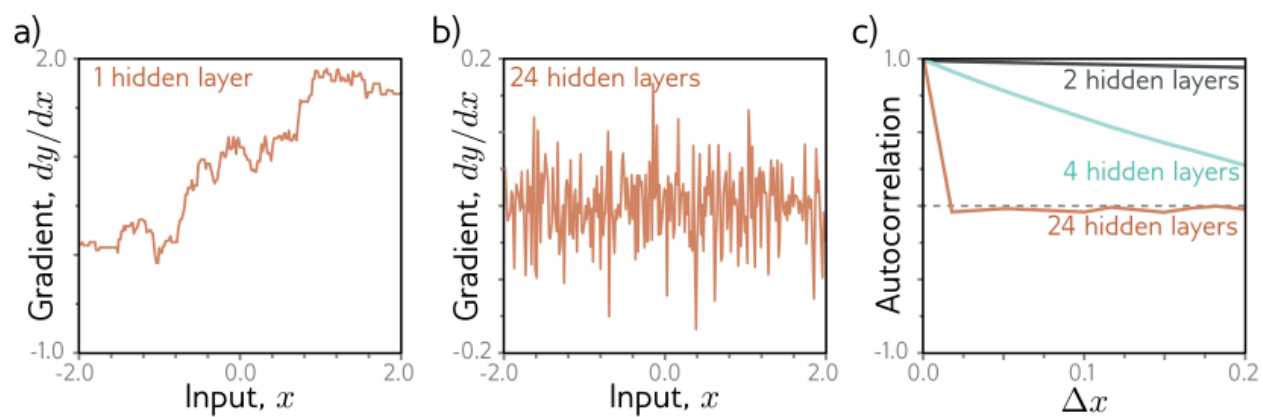


1. Shattered gradients Problem:

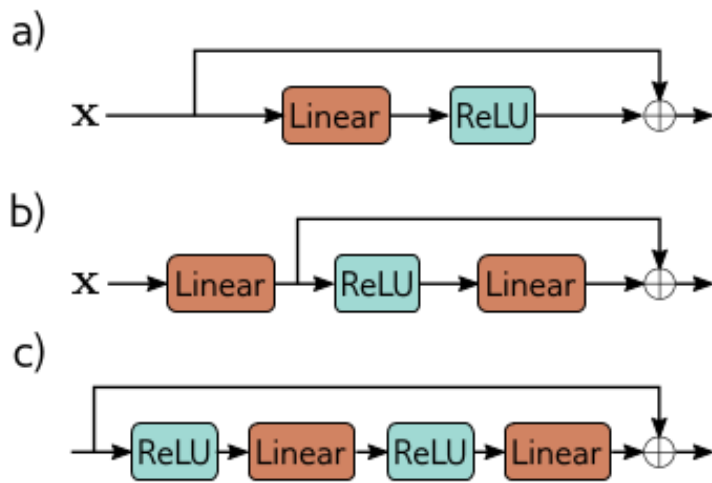


Deep networks have uncorrelated gradients in earlier layers

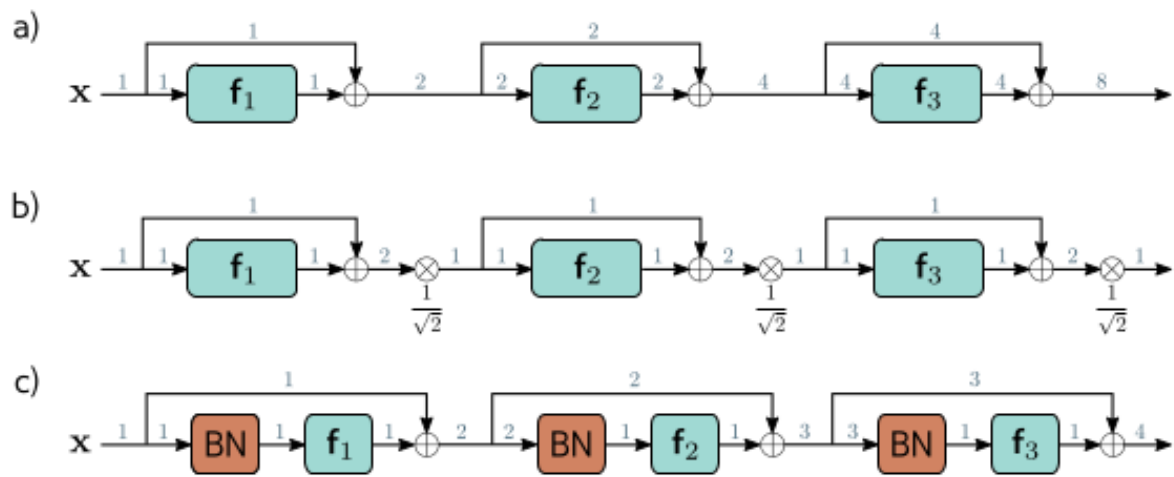


11.2 Residual connections and residual blocks

Deeper networks with residual connections



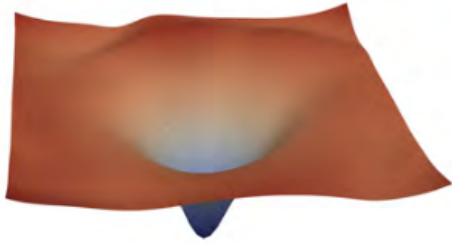
Revisiting He initialization variance propagation



Why do nets with residual connections perform so well?

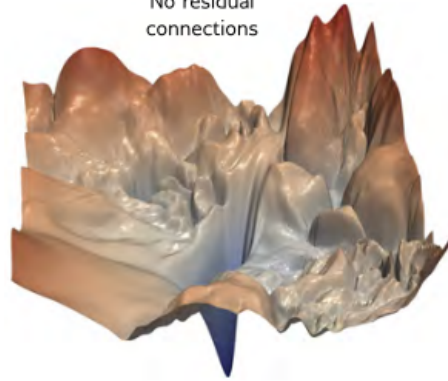
a)

Residual
connections



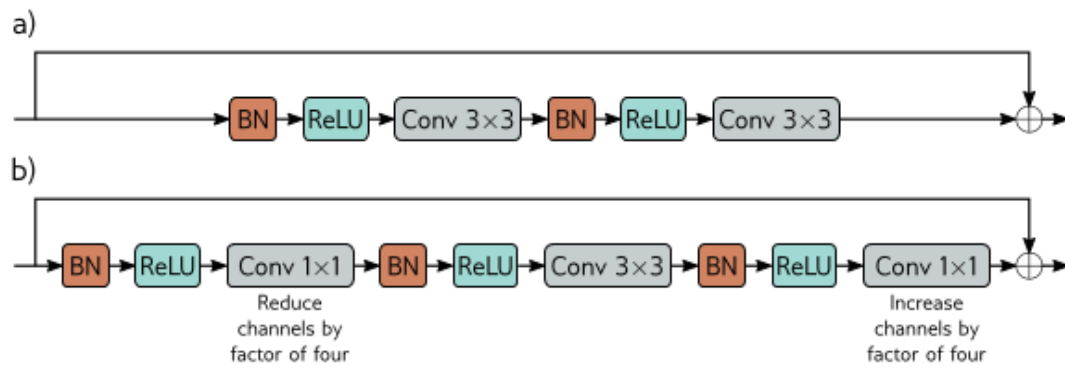
b)

No residual
connections

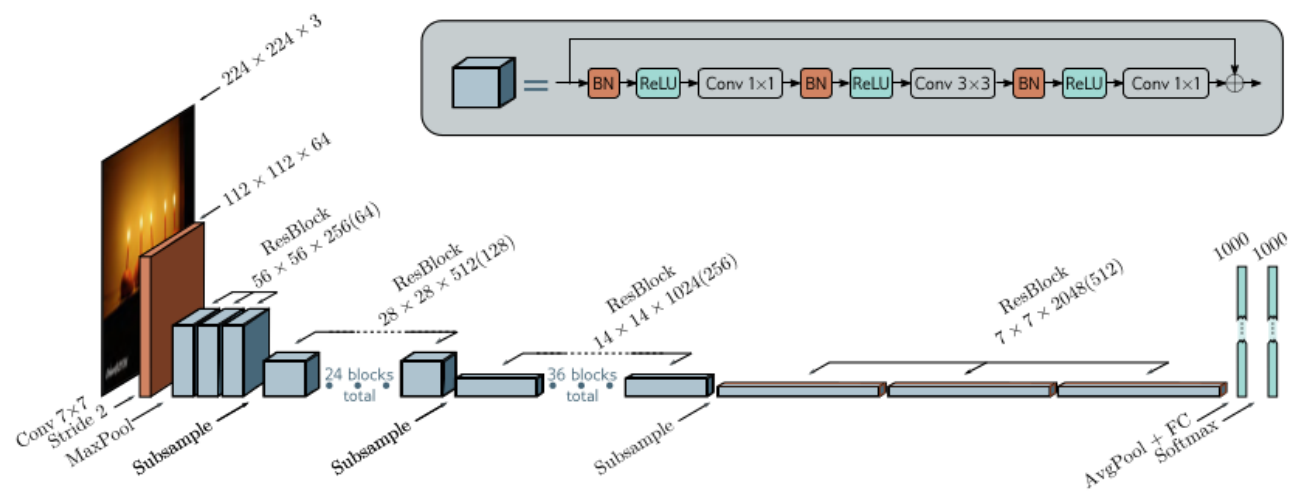


Common Residual Architectures

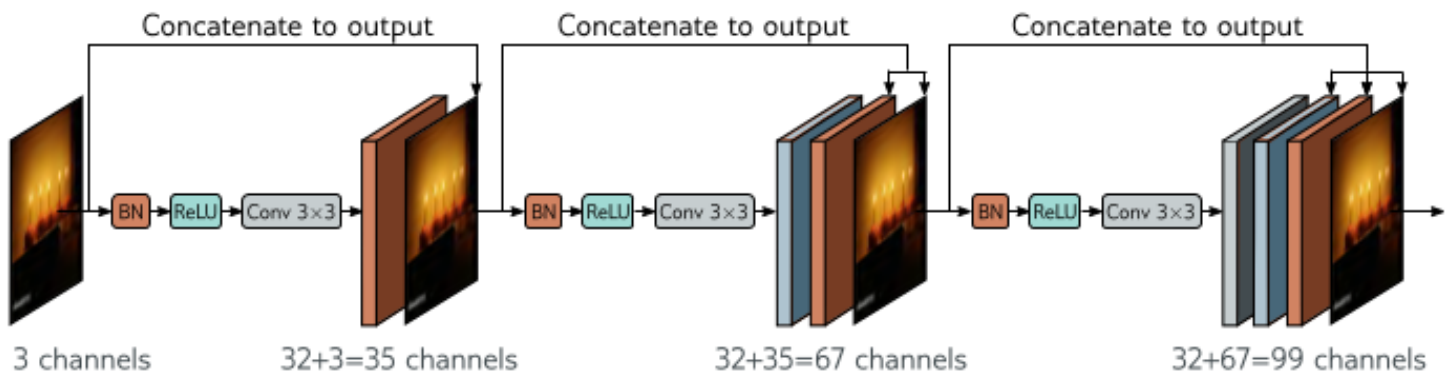
Residual and Bottleneck residual blocks



ResNet 200 Model



DenseNet.



UNet

