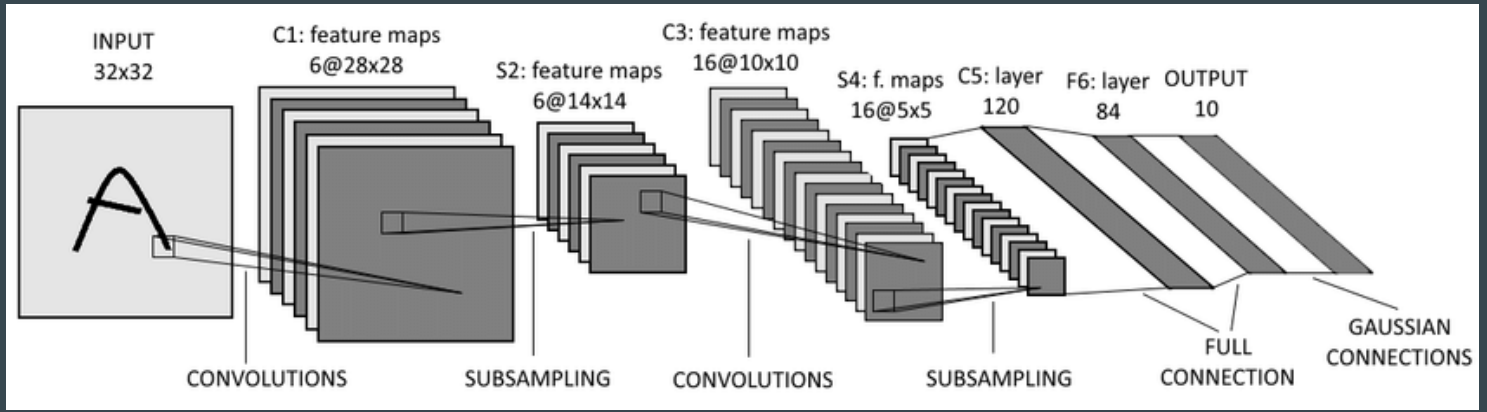
Часть 6:
Архитектуры сверточных сетей
•••

Романов Михаил, Игорь Слинько

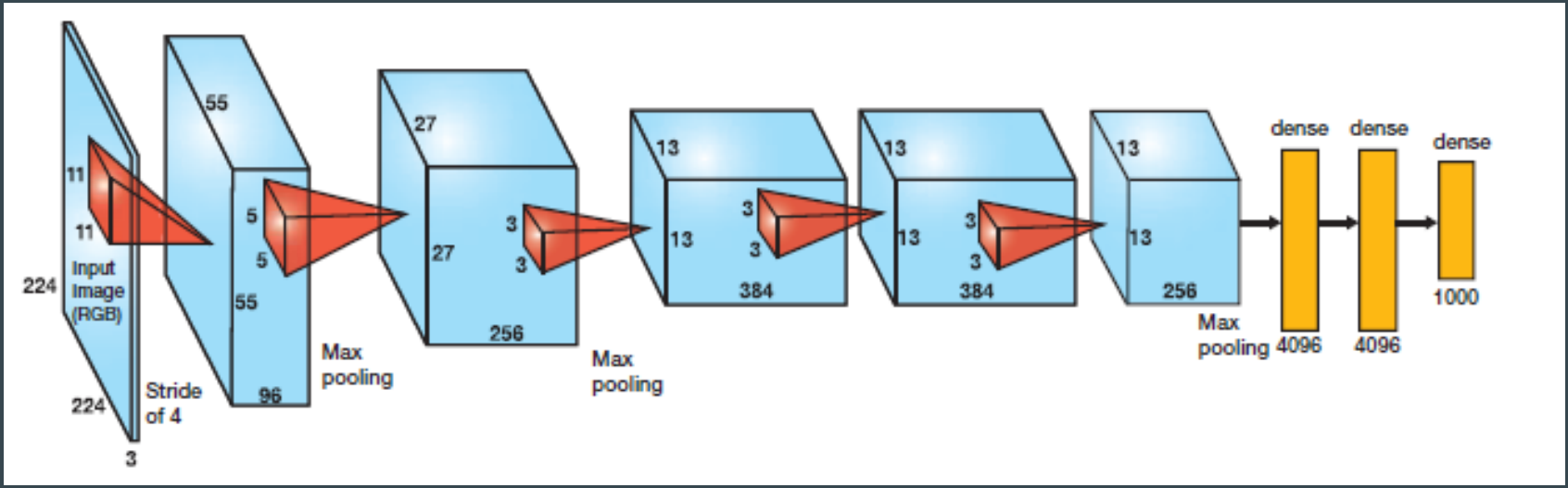
Задача

- ImageNet 1000
- База данных, 15 миллиона реальных изображений 1000 классов

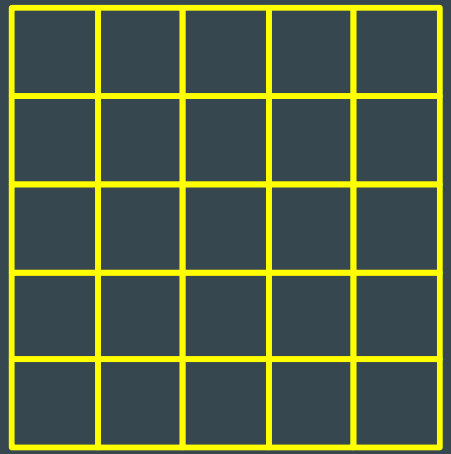
LeNet

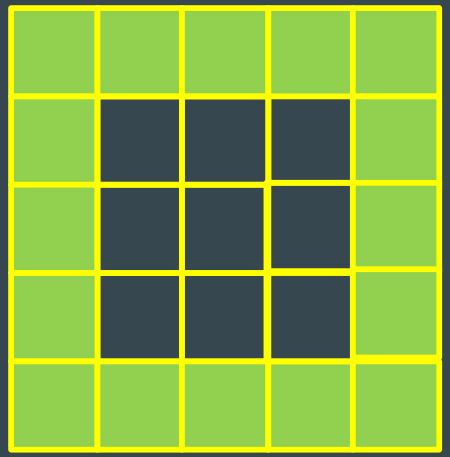

INPUT
32x32
C1: feature maps
6@28x28
S2: feature maps
6@14x14
C3: feature maps
16@10x10
S4: f. maps
16@5x5
C5: layer
120
F6: layer
84
OUTPUT
10
CONVOLUTIONS
SUBSAMPLING
CONVOLUTIONS
SUBSAMPLING
FULL
CONNECTION
GAUSSIAN
CONNECTIONS

AlexNet


224
Input
Image
(RGB)
224
3
11
11
Stride
of 4
55
55
96
5
5
Max
pooling
27
27
256
3
3
Max
pooling
13
13
384
3
3
13
13
384
3
3
13
13
256
Max
pooling
dense
4096
dense
4096
dense
1000

Зачем каскад сверток





Почему ReLU

$$\sigma(x) = \frac{1}{1+e^{-x}} \qquad \sigma' = \sigma(1-\sigma)$$

$$\sigma(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \qquad \sigma' = (1-\sigma)(1+\sigma)$$

Почему ReLU

$$\sigma(x) = \frac{1}{1+e^{-x}} \qquad \sigma' = \sigma(1-\sigma)$$

$$\sigma(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \qquad \sigma' = (1-\sigma)(1+\sigma)$$

$$\mathrm{ReLU}(x) = \begin{cases} 0 & \text{if } x<0 \\ x & \text{if } x>0 \end{cases}$$

Почему ReLU

$$\sigma(x) = \frac{1}{1+e^{-x}} \qquad \sigma' = \sigma(1-\sigma)$$

$$\sigma(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \qquad \sigma' = (1-\sigma)(1+\sigma)$$

$$\mathrm{ReLU}(x) = \begin{cases} 0 & \text{if } x<0 \\ x & \text{if } x>0 \end{cases}$$

$$\mathrm{ELU}(x) = \begin{cases} e^{x}-1 & \text{if } x\le 0 \\ x & \text{if } x>0 \end{cases}$$

Почему ReLU

$$\sigma(x) = \frac{1}{1+e^{-x}} \qquad \sigma' = \sigma(1-\sigma)$$

$$\sigma(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \qquad \sigma' = (1-\sigma)(1+\sigma)$$

$$ReLU(x) = \begin{cases} 0 & \text{if } x<0 \\ x & \text{if } x>0 \end{cases} \qquad L\text{-}ReLU(x) = \begin{cases} \alpha x & \text{if } x \leq 0 \\ x & \text{if } x>0 \end{cases}$$

$$ELU(x) = \begin{cases} e^{x}-1 & \text{if } x \leq 0 \\ x & \text{if } x>0 \end{cases}$$

Почему ReLU

$$\sigma(x) = \frac{1}{1+e^{-x}} \qquad \sigma' = \sigma(1-\sigma)$$

$$\sigma(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \qquad \sigma' = (1-\sigma)(1+\sigma)$$

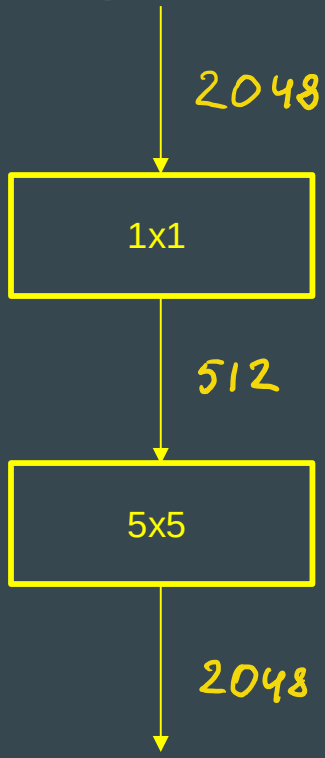
$$ReLU(x) = \begin{cases} 0 & \text{if } x<0 \\ x & \text{if } x>0 \end{cases} \qquad L\text{-}ReLU(x) = \begin{cases} \alpha x & \text{if } x \leq 0 \\ x & \text{if } x>0 \end{cases}$$

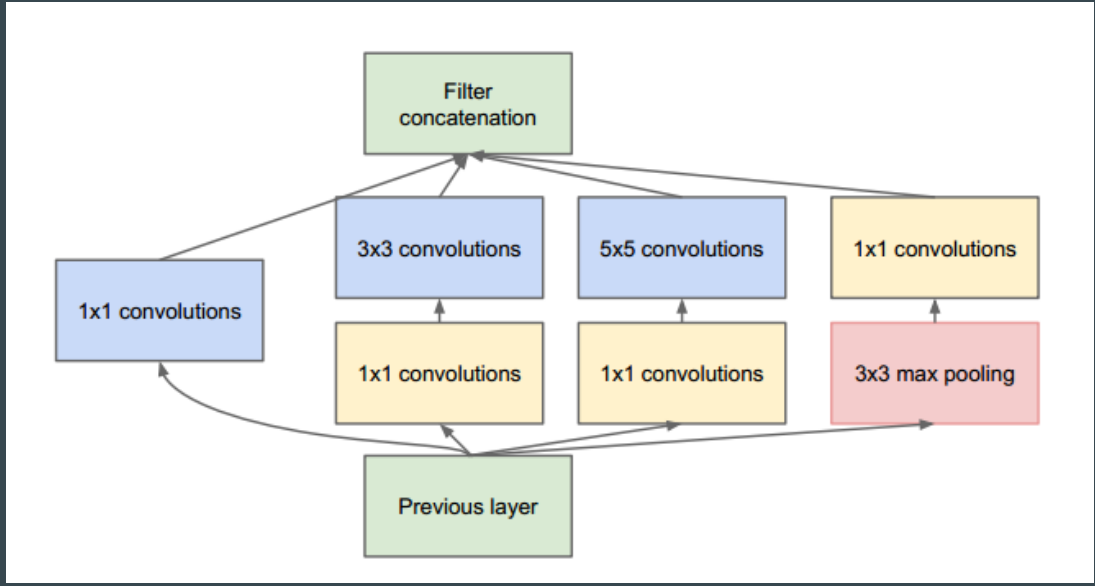
$$ELU(x) = \begin{cases} e^{x}-1 & \text{if } x \leq 0 \\ x & \text{if } x>0 \end{cases} \qquad SeLU = \begin{cases} \alpha e^{-x}-\alpha & \text{if } x \leq 0 \\ x & \text{if } x>0 \end{cases}$$

VGG

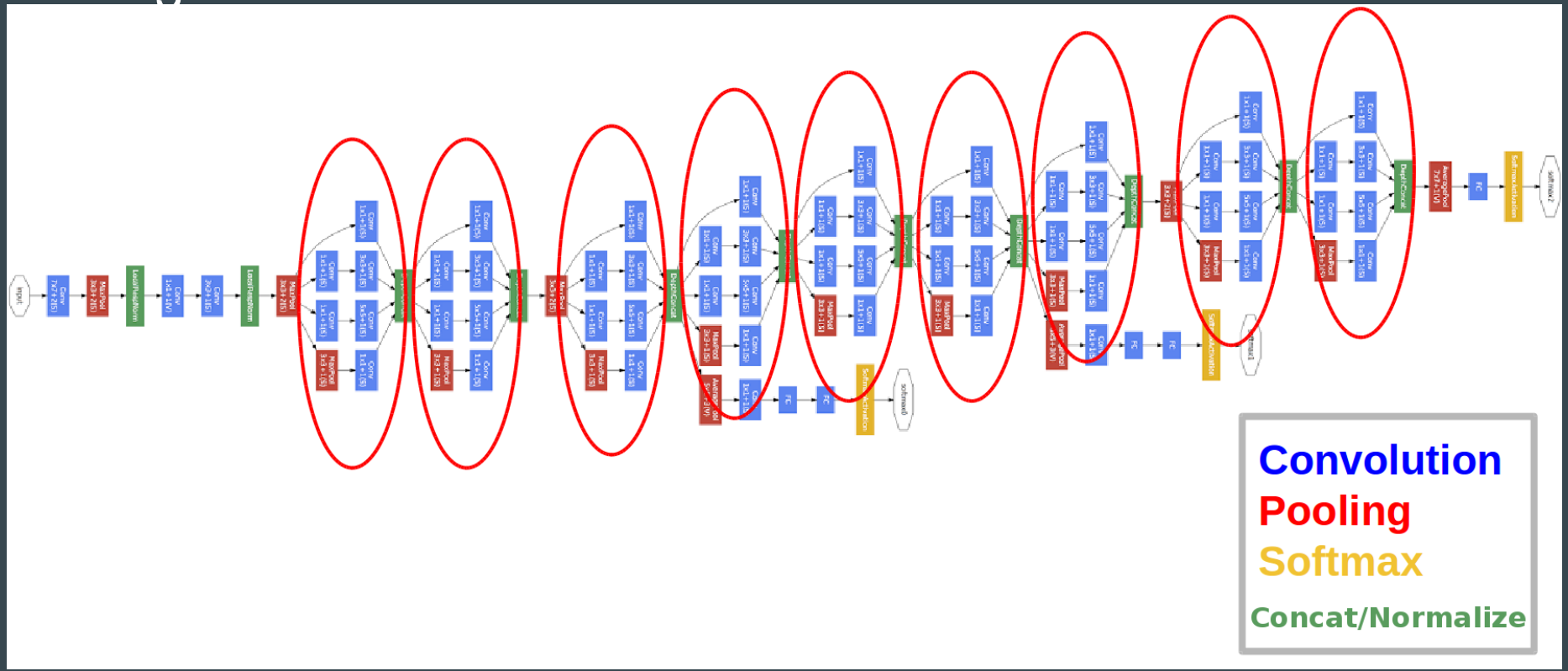
ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Inception Block





GoogLeNet



Residual Block

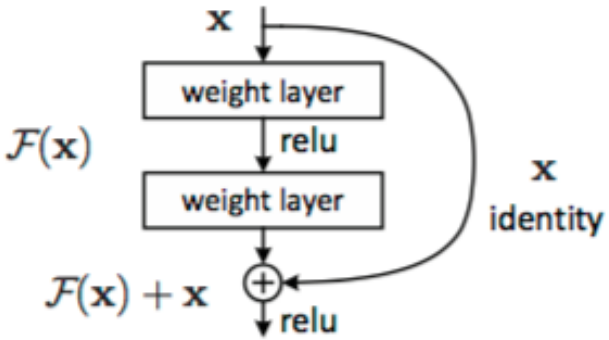


Figure 2. Residual learning: a building block.

$$y = f(x) + x$$

$$y' = f'(x) + 1$$

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

Градиент не затухает

Residual Block

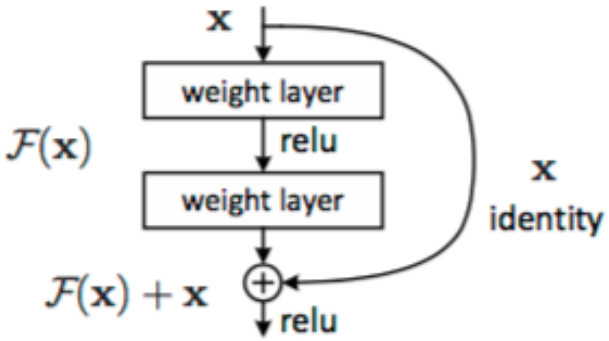


Figure 2. Residual learning: a building block.

$$y = f(x) + x$$

$$y' = f'(x) + 1$$

$$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial y} \cdot \frac{\partial y}{\partial x}$$

Градиент не затухает

Residual Block

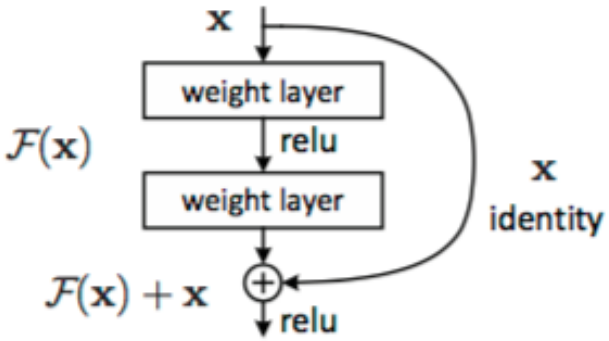


Figure 2. Residual learning: a building block.

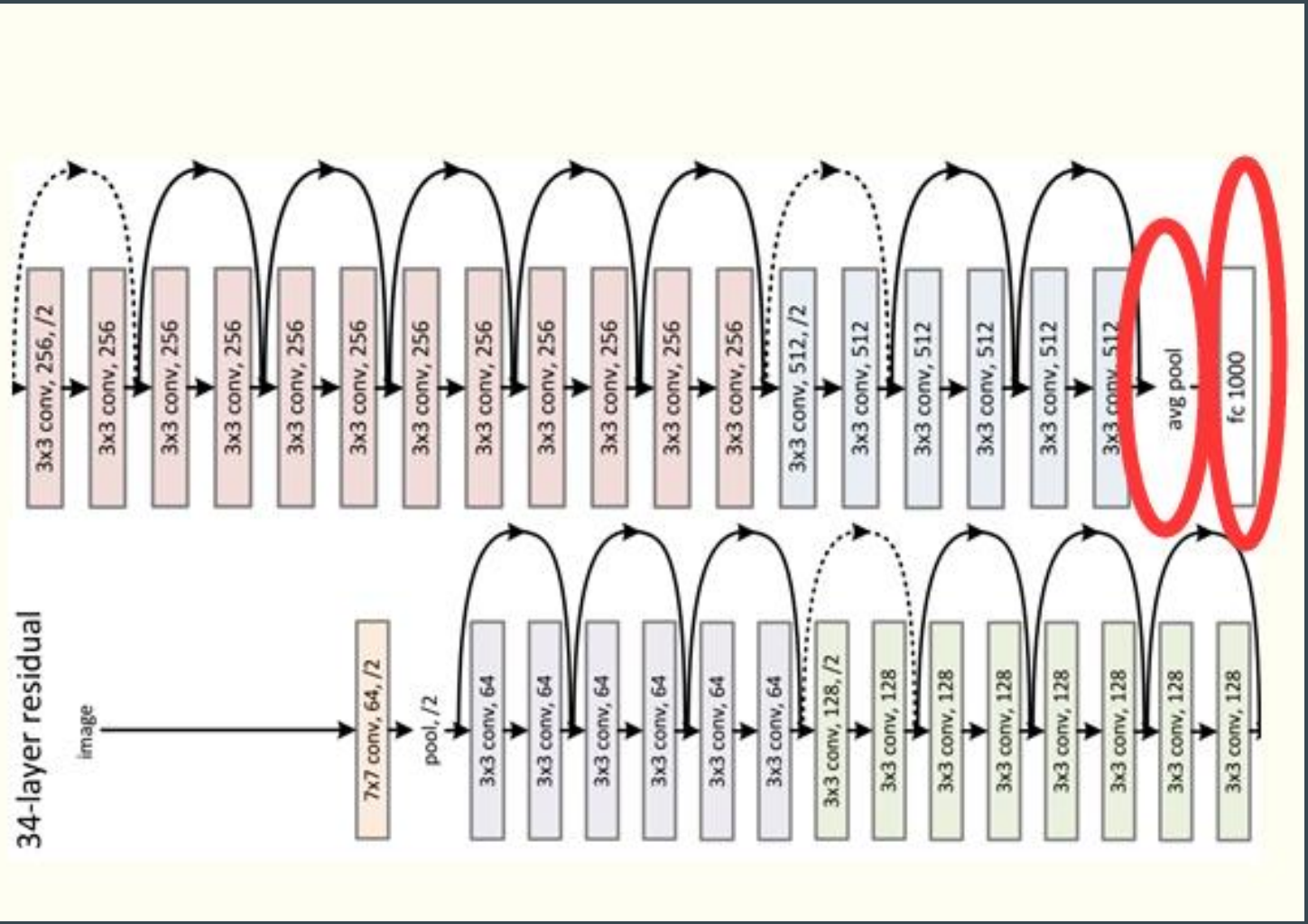
$$y = f(x) + x$$

$$y' = f'(x) + 1$$

$$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial y} \cdot \frac{\partial y}{\partial x} = \frac{\partial L}{\partial y}\left[f'(x) + 1\right]$$

Градиент не затухает

ResNet


34-layer residual
image
7x7 conv, 64, /2
pool, /2
3x3 conv, 64
3x3 conv, 128, /2
3x3 conv, 128
3x3 conv, 256, /2
3x3 conv, 256
3x3 conv, 512, /2
3x3 conv, 512
avg pool
fc 1000

Итоги

- LeNet
- AlexNet и ReLU
- VGG
- Inception
- ResNet