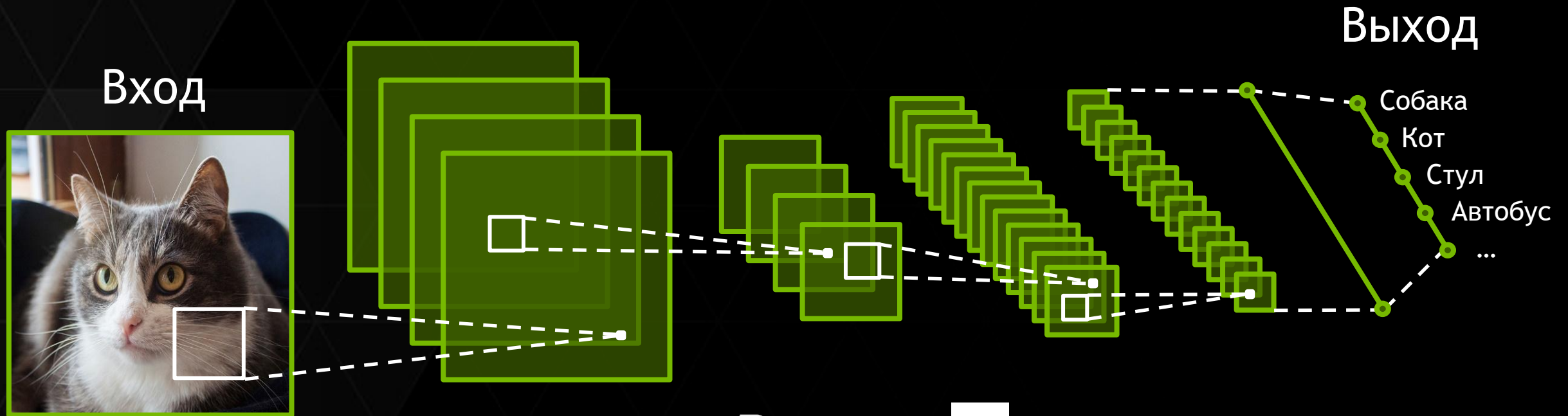




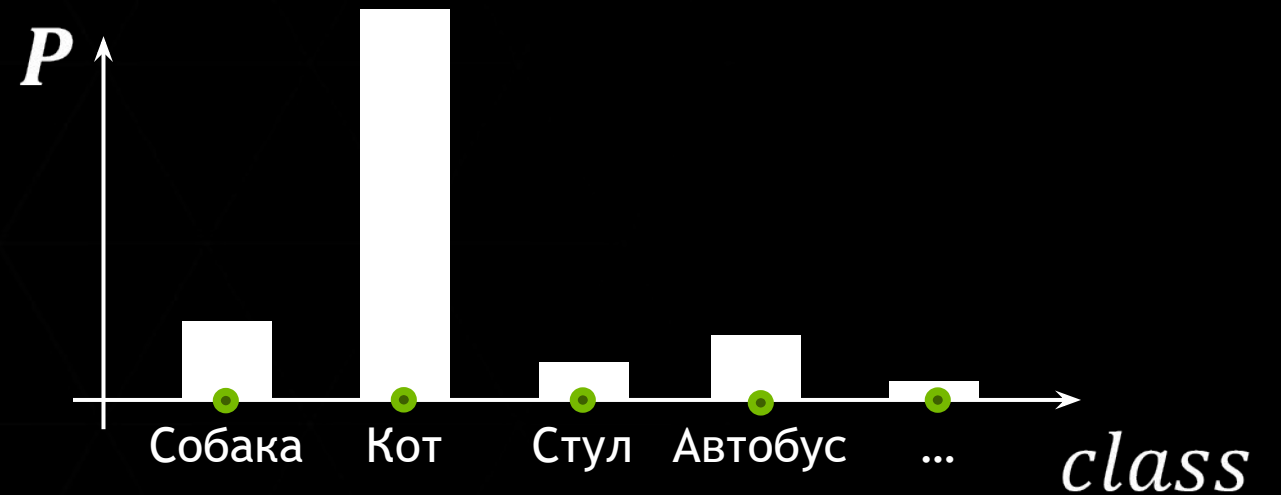
# [3] ПРОДВИНУТЫЕ СВЁРТОЧНЫЕ НЕЙРОСЕТИ

Дмитрий Коробченко

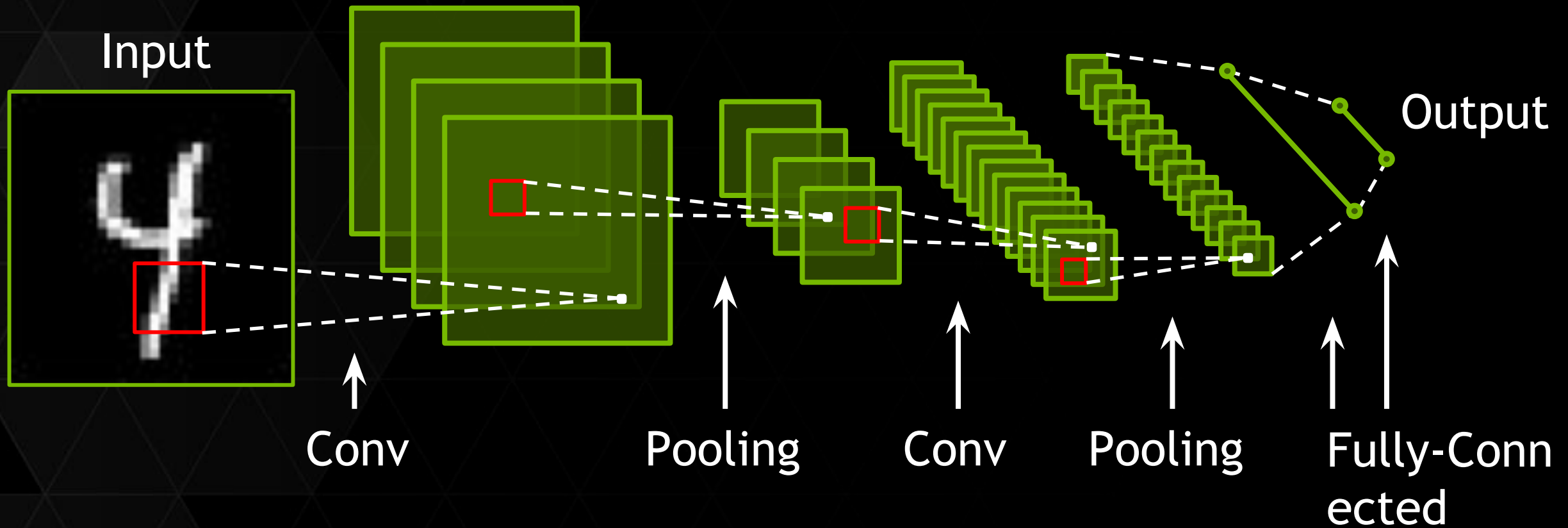
# СВЁРТОЧНАЯ НЕЙРОННАЯ СЕТЬ



Выход сети -  
распределение вероятностей  
принадлежности к классу

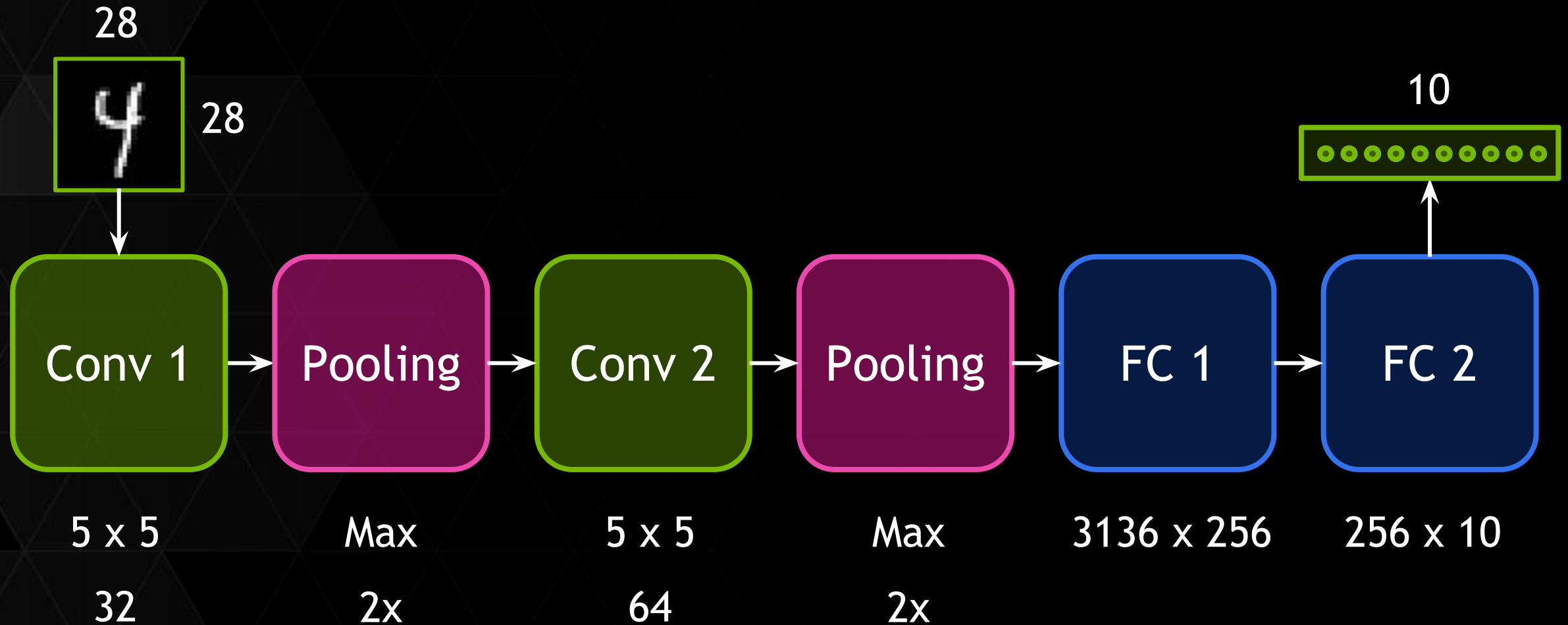


# СВЁРТОЧНАЯ НЕЙРОННАЯ СЕТЬ ДЛЯ MNIST





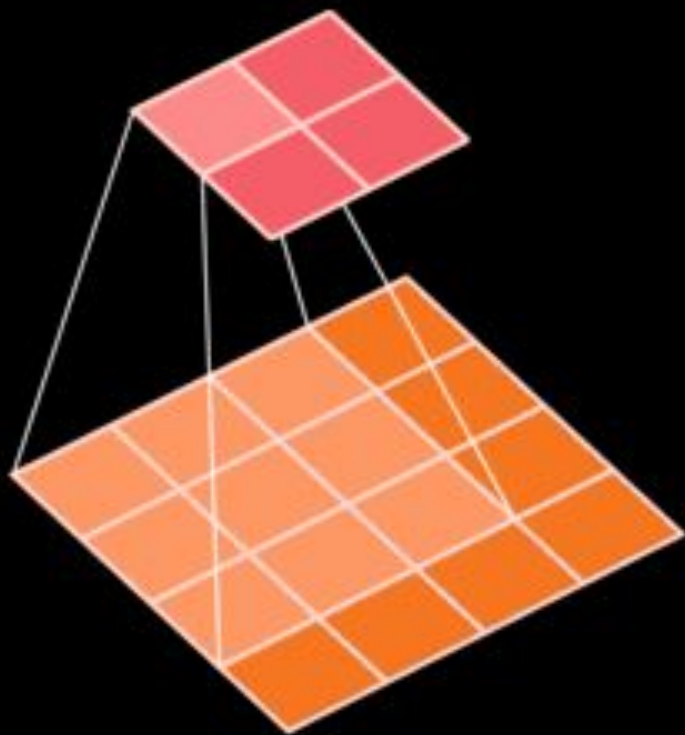
# СВЁРТОЧНАЯ НЕЙРОННАЯ СЕТЬ ДЛЯ MNIST



# СВЁРТКА СО СТРАЙДОМ

Conv

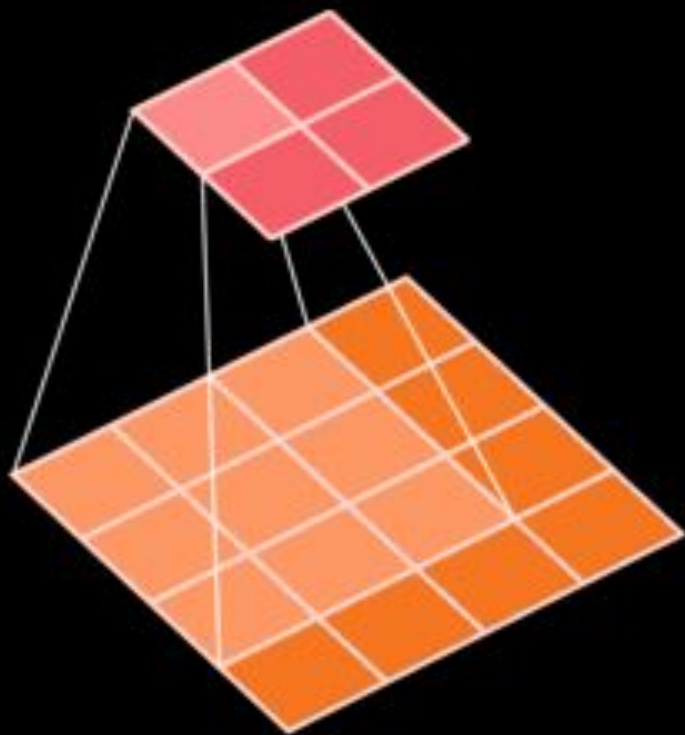
Шаг окна: 1



# СВЁРТКА СО СТРАЙДОМ

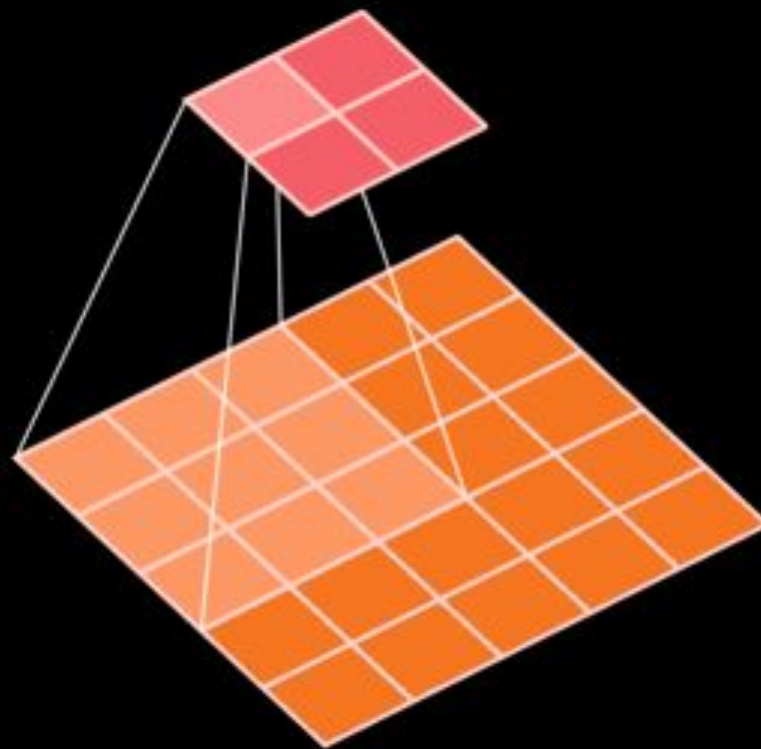
Conv

Шаг окна: 1



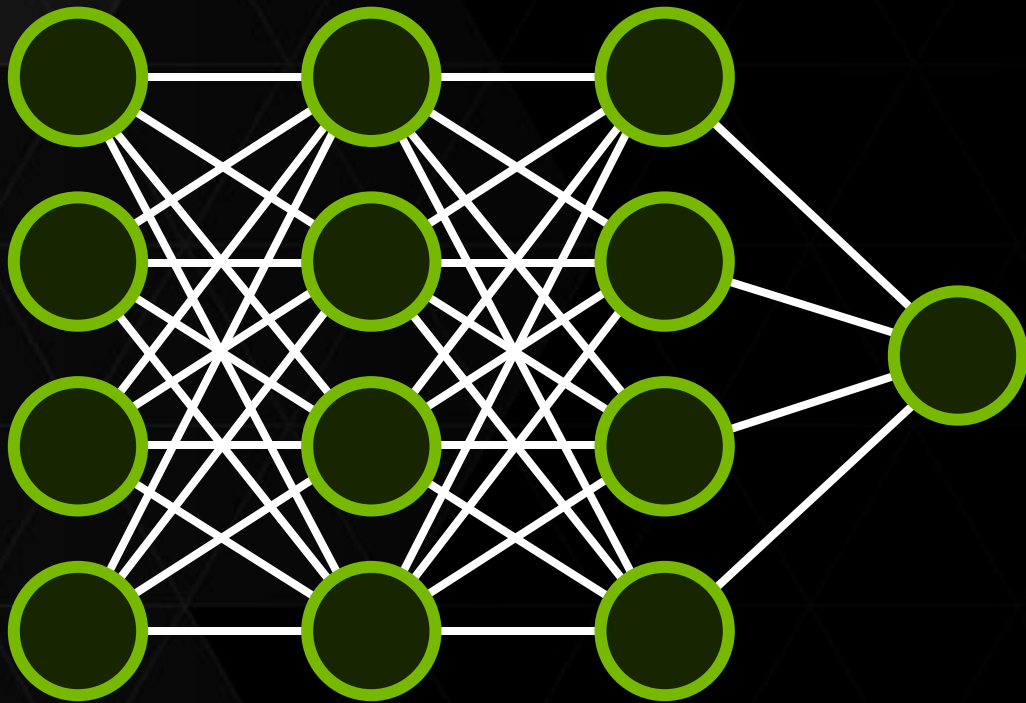
Strided Conv

Шаг окна: 2



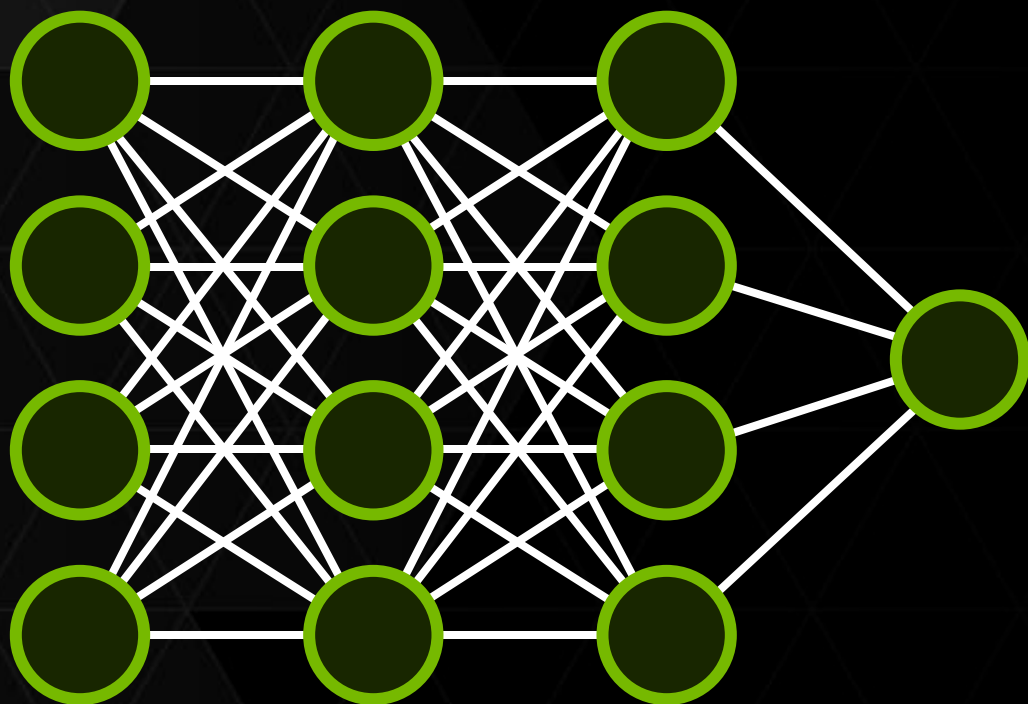
# РЕГУЛЯРИЗАЦИЯ: DROPOUT

Без Дропаута

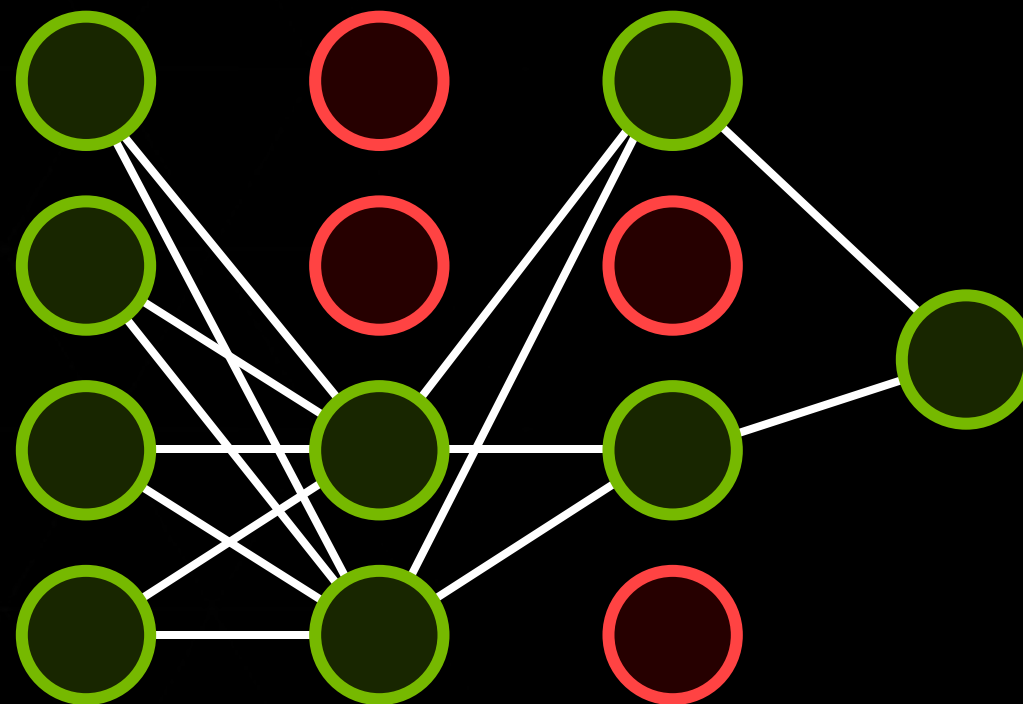


# РЕГУЛЯРИЗАЦИЯ: DROPOUT

Без Дропаута

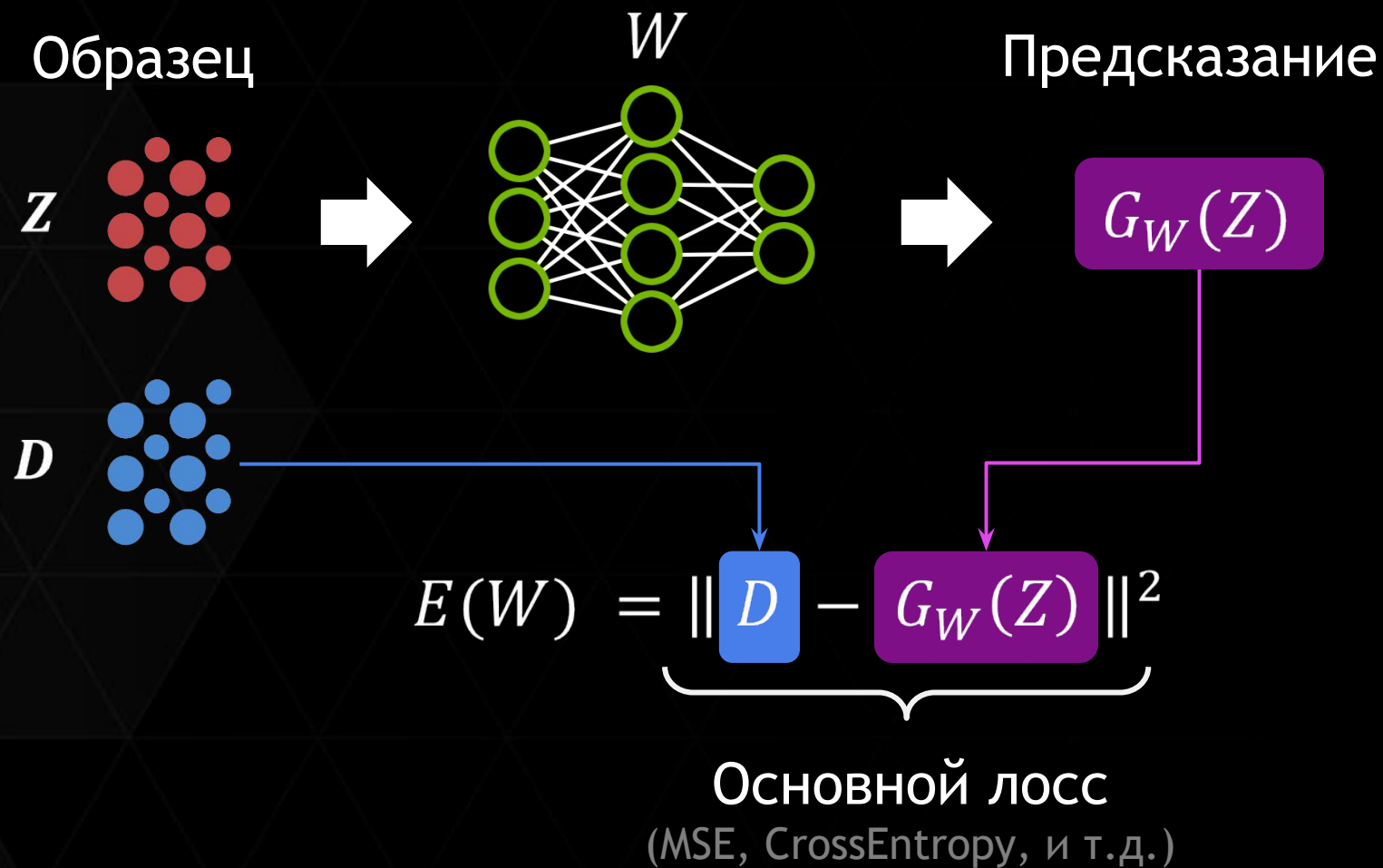


С Дропаутом

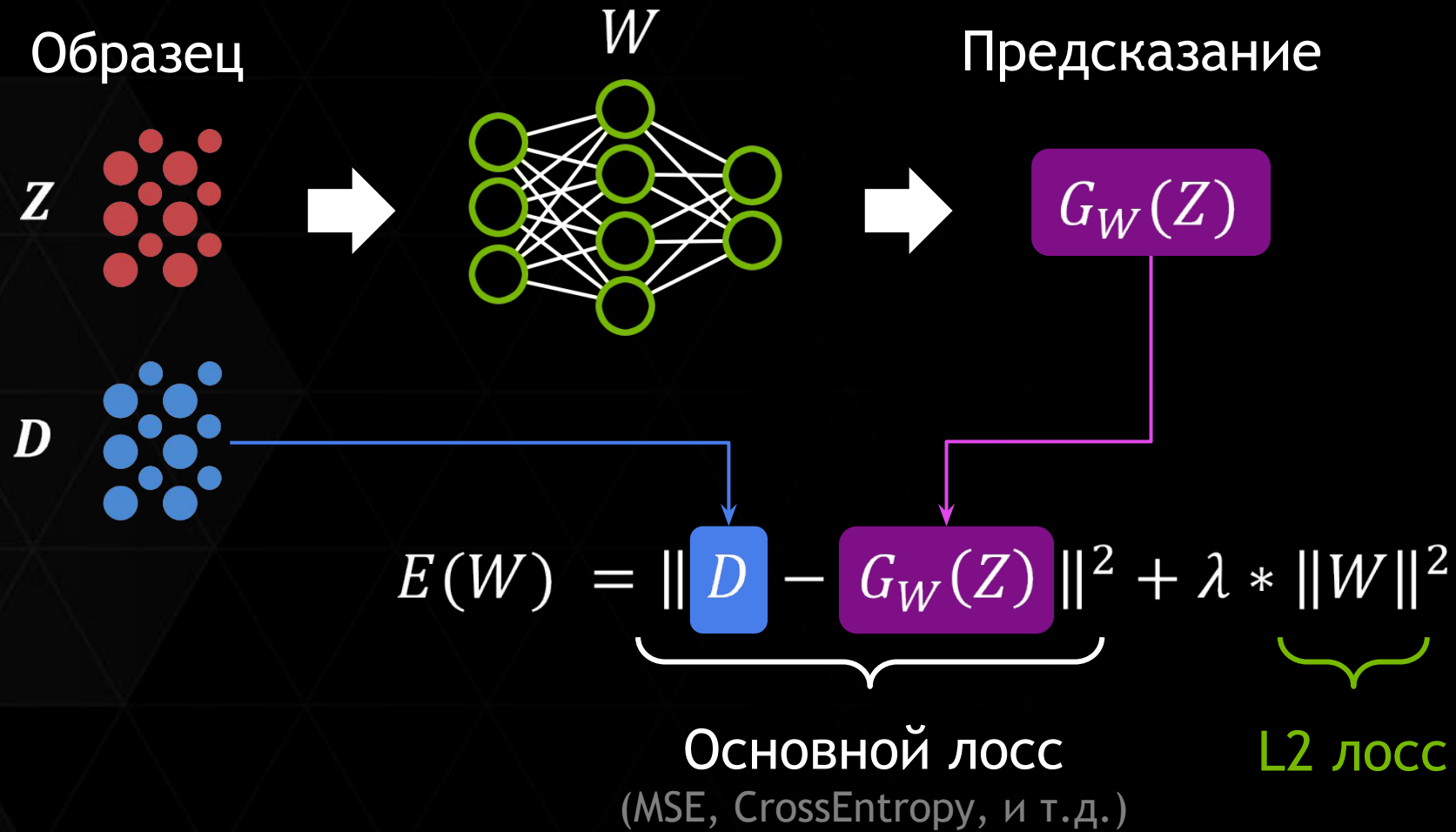




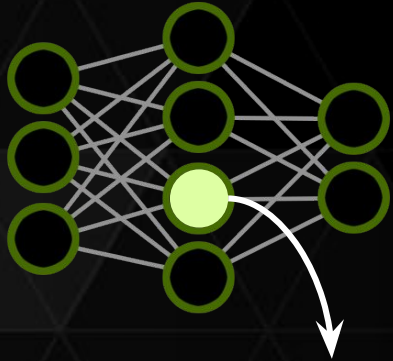
# РЕГУЛЯРИЗАЦИЯ: WEIGHT DECAY (L2)



# РЕГУЛЯРИЗАЦИЯ: WEIGHT DECAY (L2)



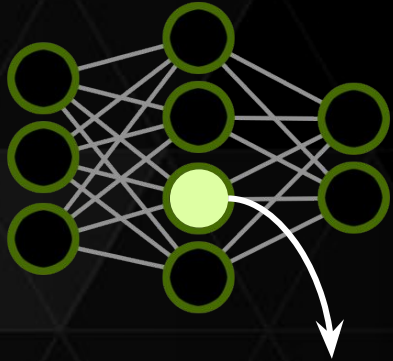
# НОРМАЛИЗАЦИЯ БАТЧА (BATCHNORM)



Значения нейрона  
для батча  $B$

$$(x_1, \dots, x_N)$$

# НОРМАЛИЗАЦИЯ БАТЧА (BATCHNORM)



$(x_1, \dots, x_N)$

$(x_1, \dots, x_N)$

BatchNorm

$(y_1, \dots, y_N)$

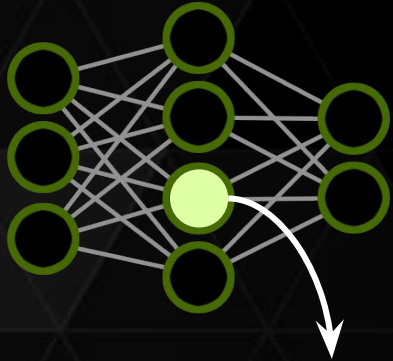
Mean

$$\mu_B = \frac{1}{N} \sum_{i=1}^N x_i$$

Variance

$$\sigma_B^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu_B)^2$$

# НОРМАЛИЗАЦИЯ БАТЧА (BATCHNORM)



Значения нейрона  
для батча  $B$

$(x_1, \dots, x_N)$

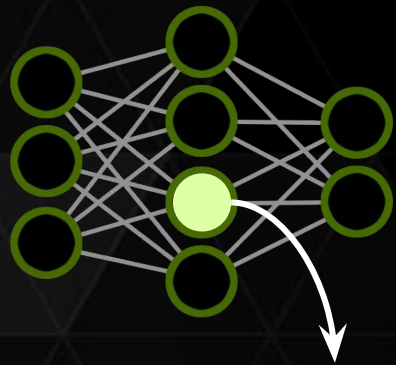


BatchNorm

$(y_1, \dots, y_N)$



# НОРМАЛИЗАЦИЯ БАТЧА (BATCHNORM)



Значения нейрона  
для батча  $B$

$(x_1, \dots, x_N)$



BatchNorm

$(y_1, \dots, y_N)$

Mean

$$\mu_B = \frac{1}{N} \sum_{i=1}^N x_i$$

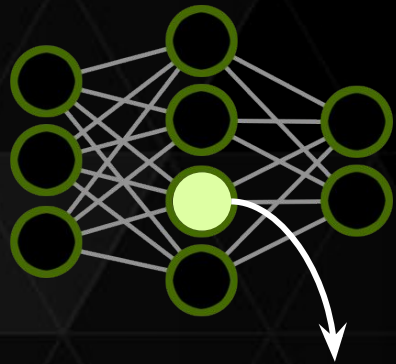
Variance

$$\sigma_B^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu_B)^2$$

Нормализация

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}}$$

# НОРМАЛИЗАЦИЯ БАТЧА (BATCHNORM)



Значения нейрона  
для батча  $B$

$(x_1, \dots, x_N)$

BatchNorm

$(y_1, \dots, y_N)$

Mean

$$\mu_B = \frac{1}{N} \sum_{i=1}^N x_i$$

Variance

$$\sigma_B^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu_B)^2$$

Нормализация

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}}$$

Шкалирование и сдвиг

$$y_i = \gamma \hat{x}_i + \beta$$

# РЕГУЛЯРИЗАЦИЯ ДЛЯ СВЁРТОЧНЫХ НЕЙРОСЕТЕЙ

Dropout



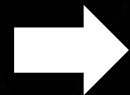
Полносвязные слои

Batch Norm



Свёрточные слои

Weight Decay (L2)



Все слои с параметрами

# ПРОСТЫЕ ИЗОБРАЖЕНИЯ



# СЛОЖНЫЕ ИЗОБРАЖЕНИЯ



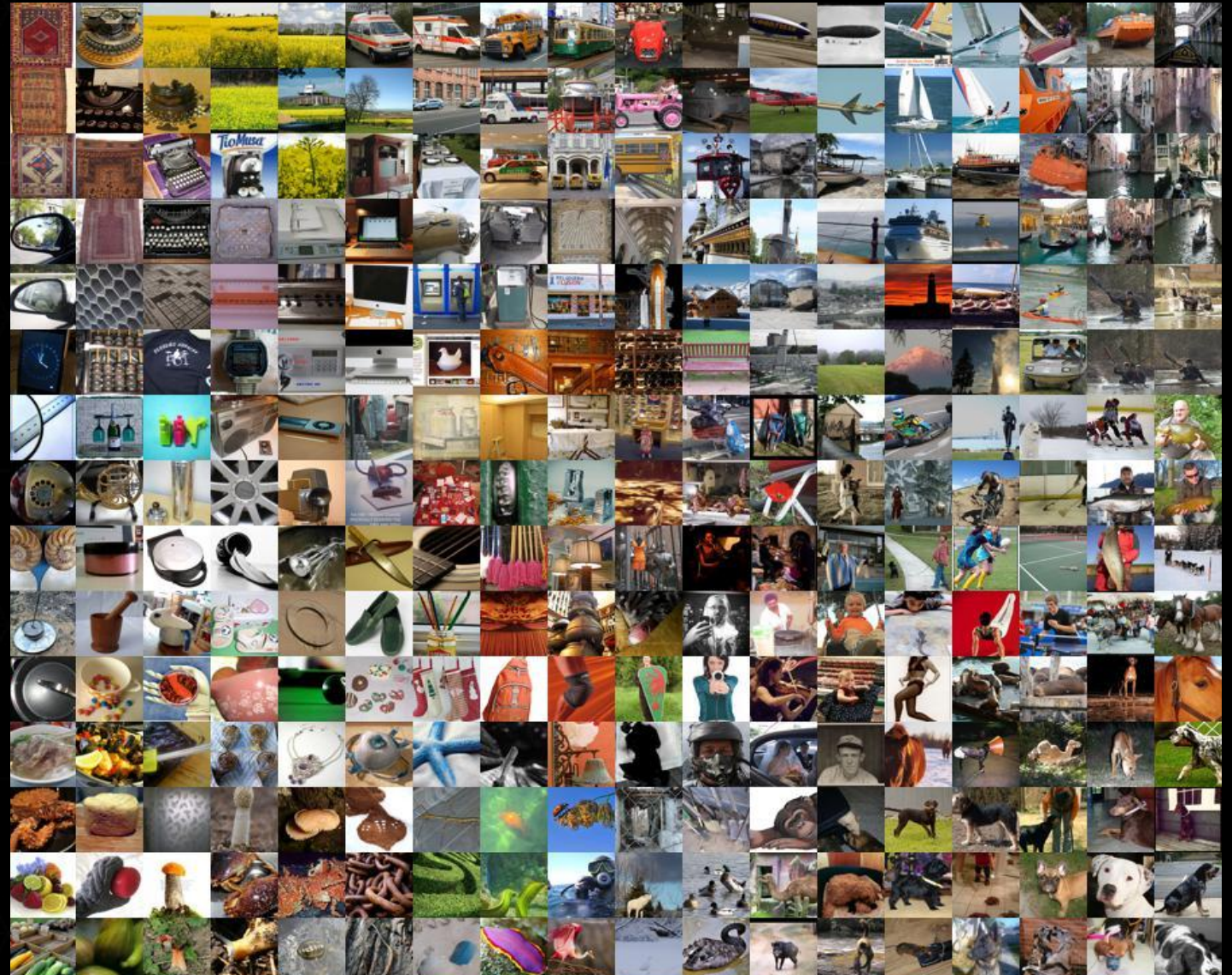


# IMAGENET

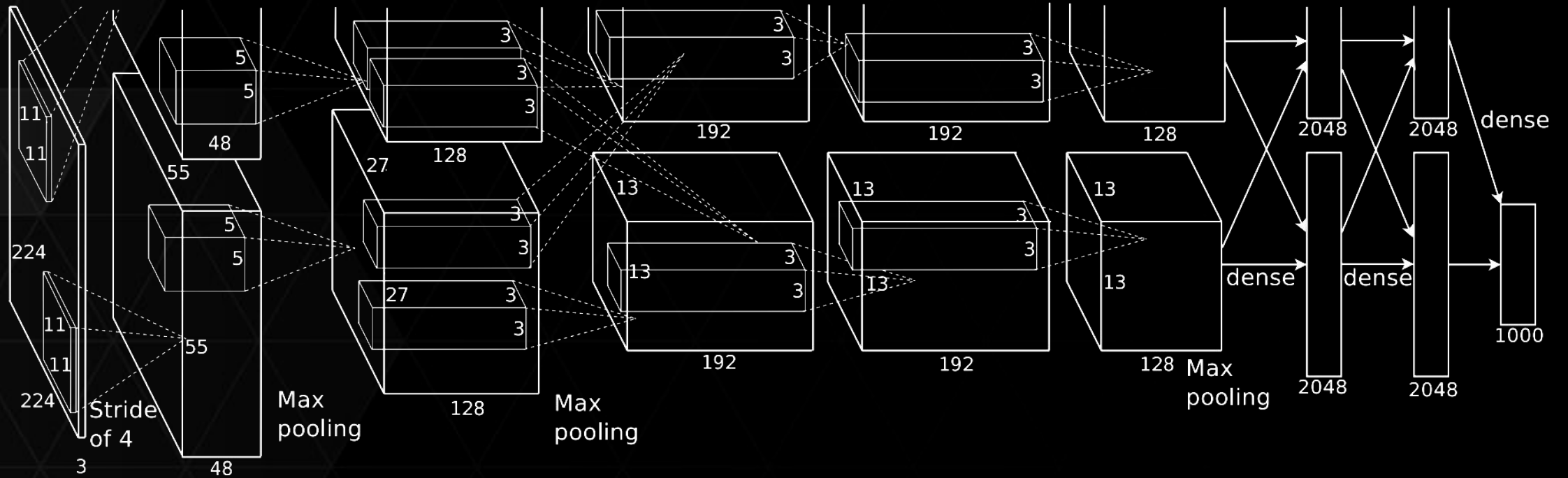
ILSVRC 2012

~ 1'000'000 изображений  
1'000 категорий

- Классификация
- Локализация
- Детектирование объектов



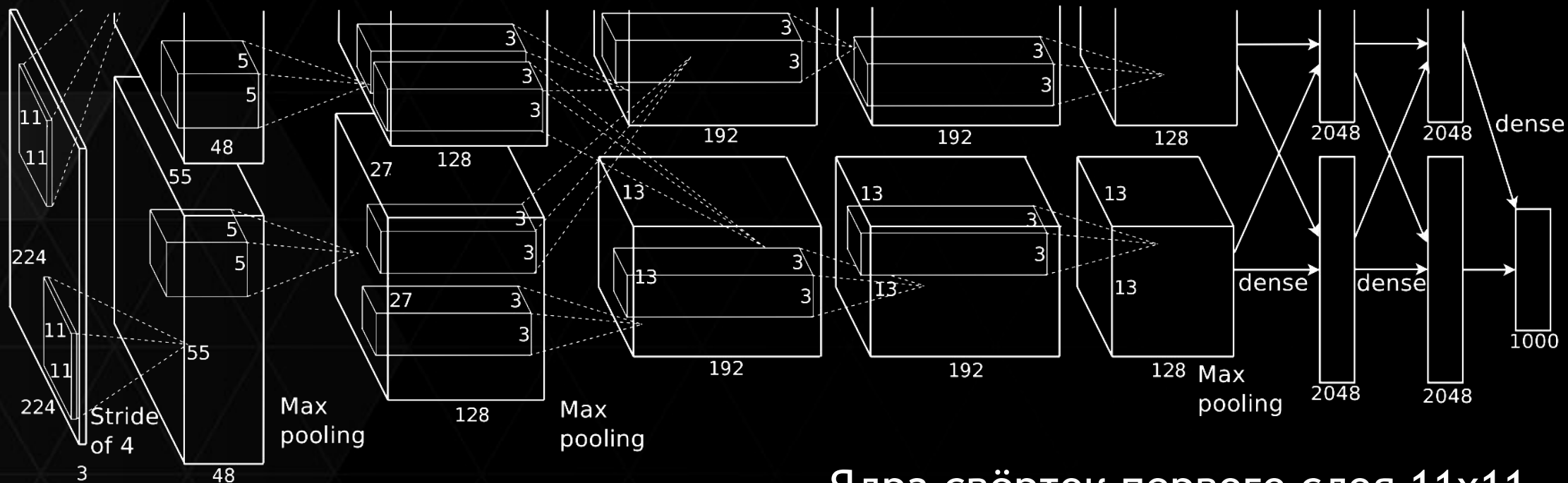
# ALEXNET (2012)



- Архитектура: глубже и шире
- Local Response Normalization
- Dropout
- Аугментация данных
- SGD + Momentum



# ALEXNET (2012)

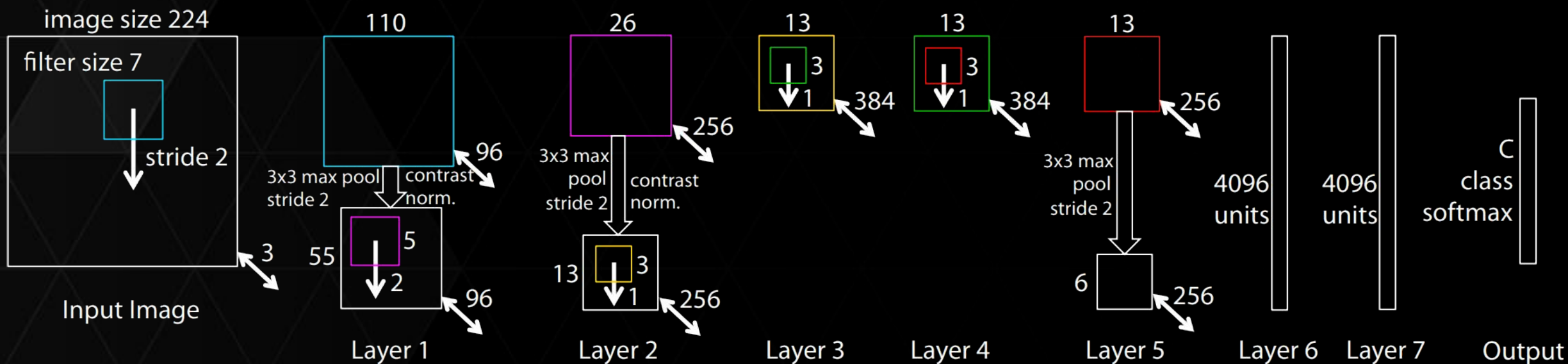


- Архитектура: глубже и шире
- Local Response Normalization
- Dropout
- Аугментация данных
- SGD + Momentum

Ядра свёрток первого слоя 11x11



# ZF Net (2013)



- Архитектура: похожа на AlexNet
- Local Contrast Normalization

# VGG (2014)

- Очень глубокая и широкая сеть
- Группы свёрток 3x3
- Только элементарные слои:
  - Conv
  - Pooling
  - Fully-connected

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 <b>LRN</b>	conv3-64 <b>conv3-64</b>	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 <b>conv3-128</b>	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 <b>conv1-256</b>	conv3-256 conv3-256 <b>conv3-256</b>	conv3-256 conv3-256 conv3-256 <b>conv3-256</b>
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 <b>conv1-512</b>	conv3-512 conv3-512 <b>conv3-512</b>	conv3-512 conv3-512 conv3-512 <b>conv3-512</b>
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 <b>conv1-512</b>	conv3-512 conv3-512 <b>conv3-512</b>	conv3-512 conv3-512 conv3-512 <b>conv3-512</b>
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					



# VGG (2014)

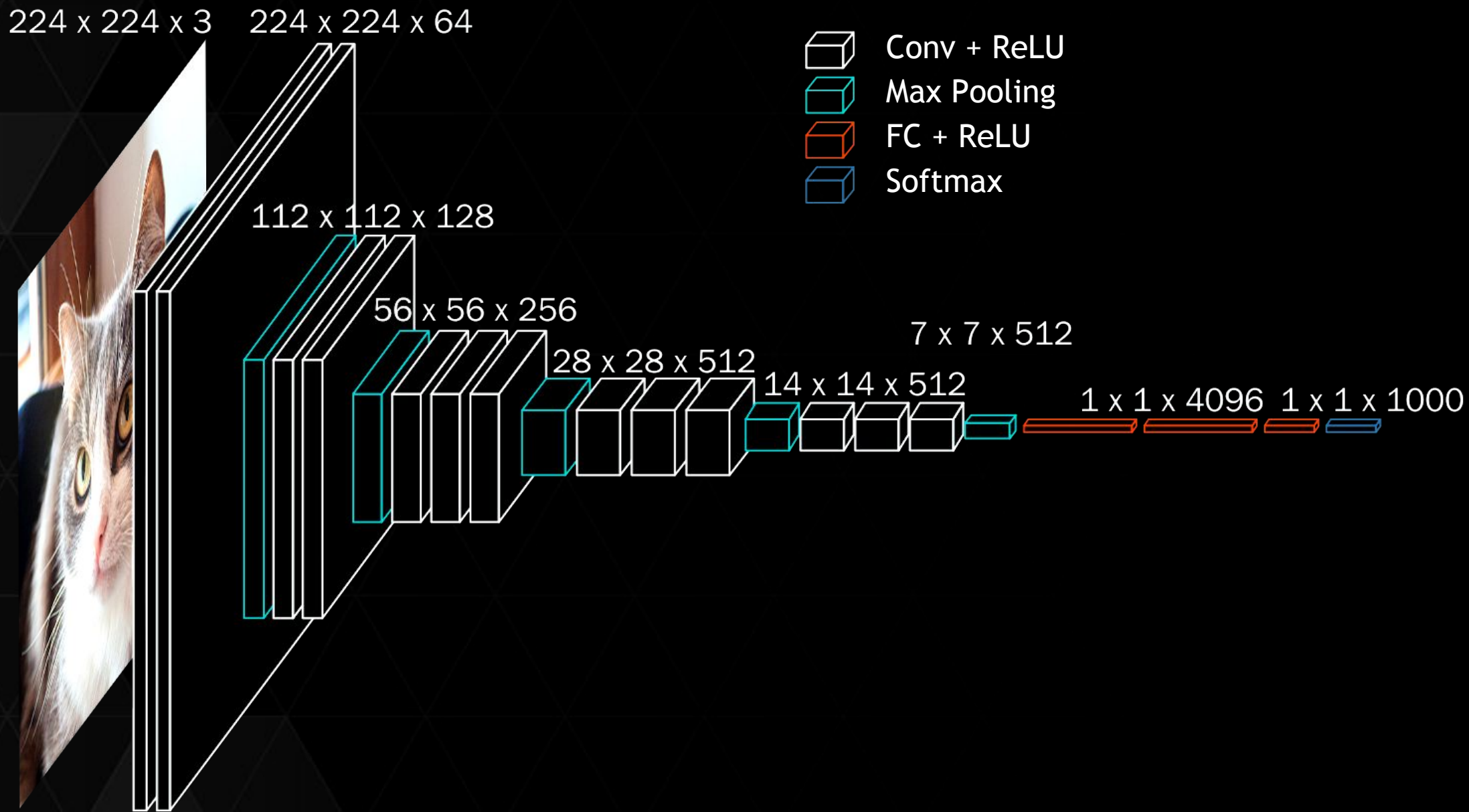
- Очень глубокая и широкая сеть
- Группы свёрток 3x3
- Только элементарные слои:
  - Conv
  - Pooling
  - Fully-connected

VGG-16

VGG-19

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 ( $224 \times 224$ RGB image)					
conv3-64	conv3-64 <b>LRN</b>	conv3-64 <b>conv3-64</b>	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 <b>conv3-128</b>	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 <b>conv1-256</b>	conv3-256 conv3-256 <b>conv3-256</b>	conv3-256 conv3-256 conv3-256 <b>conv3-256</b>
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 <b>conv1-512</b>	conv3-512 conv3-512 <b>conv3-512</b>	conv3-512 conv3-512 conv3-512 <b>conv3-512</b>
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 <b>conv1-512</b>	conv3-512 conv3-512 <b>conv3-512</b>	conv3-512 conv3-512 conv3-512 <b>conv3-512</b>
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

# VGG-16



# SPATIAL PYRAMID POOLING (2014)

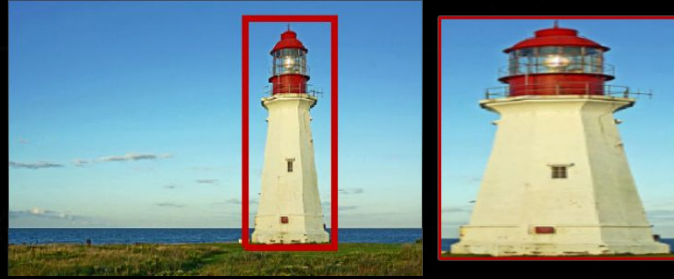


# SPATIAL PYRAMID POOLING (2014)

Crop



Wrap

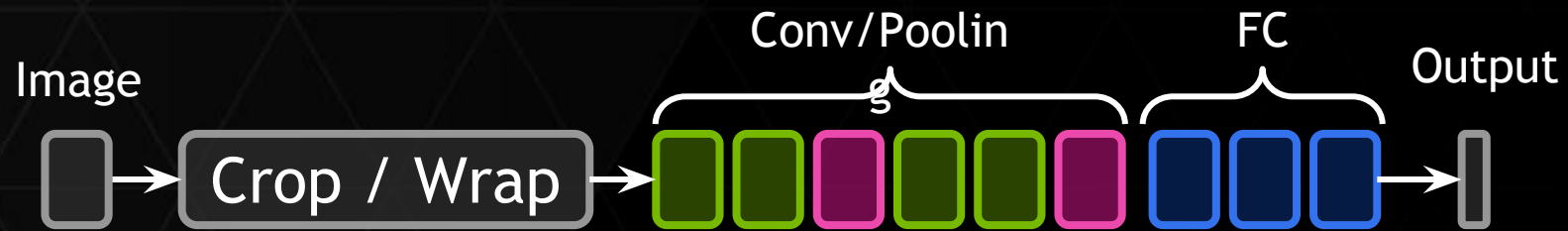


# SPATIAL PYRAMID POOLING (2014)

Crop



Wrap



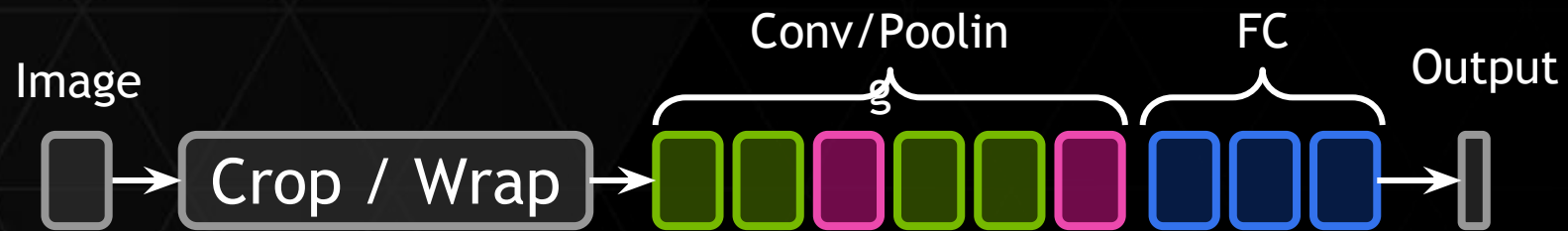


# SPATIAL PYRAMID POOLING (2014)

Crop



Wrap

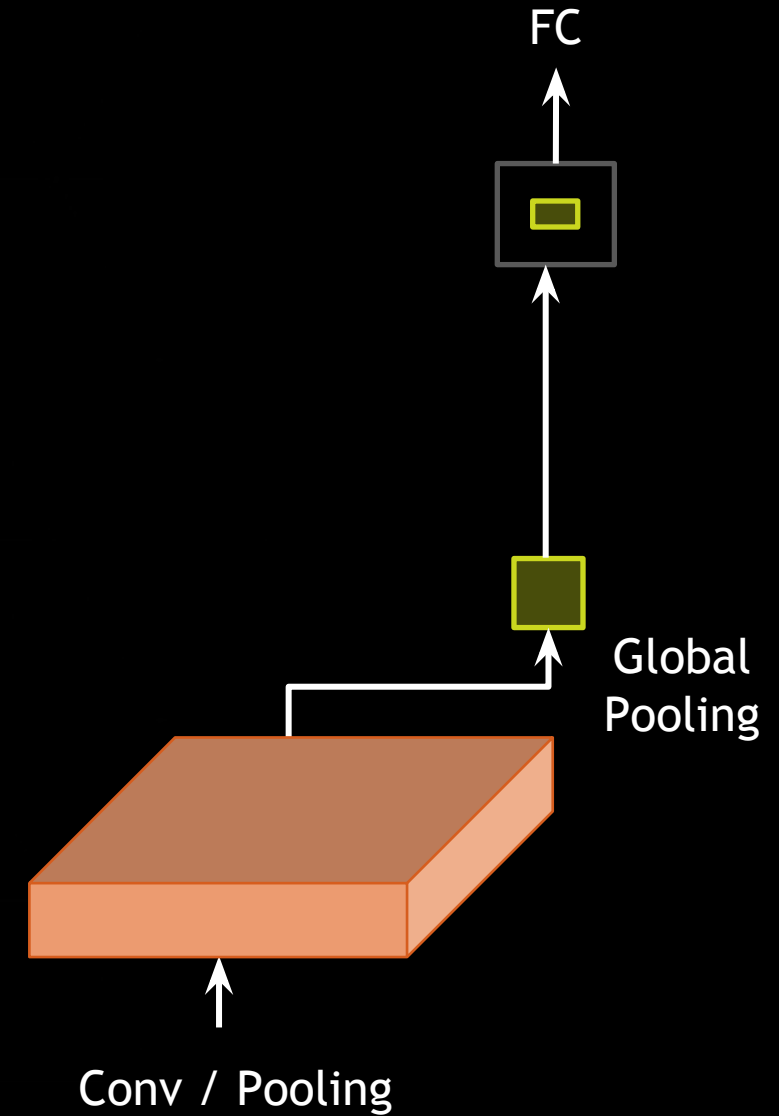
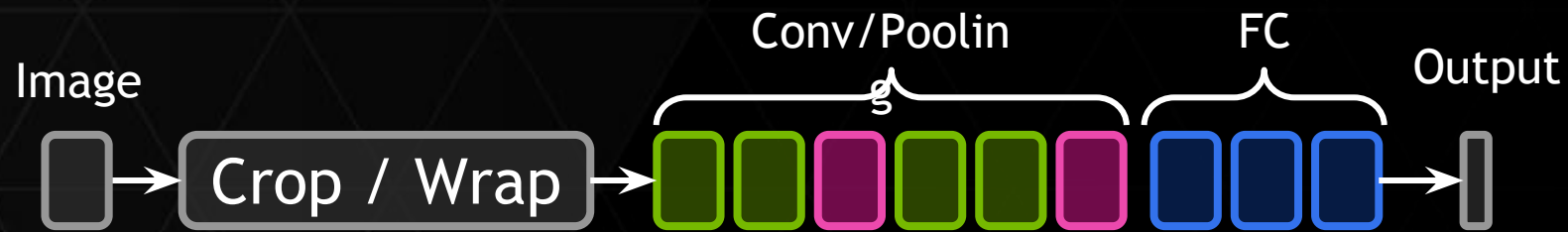


# SPATIAL PYRAMID POOLING (2014)

Crop



Wrap

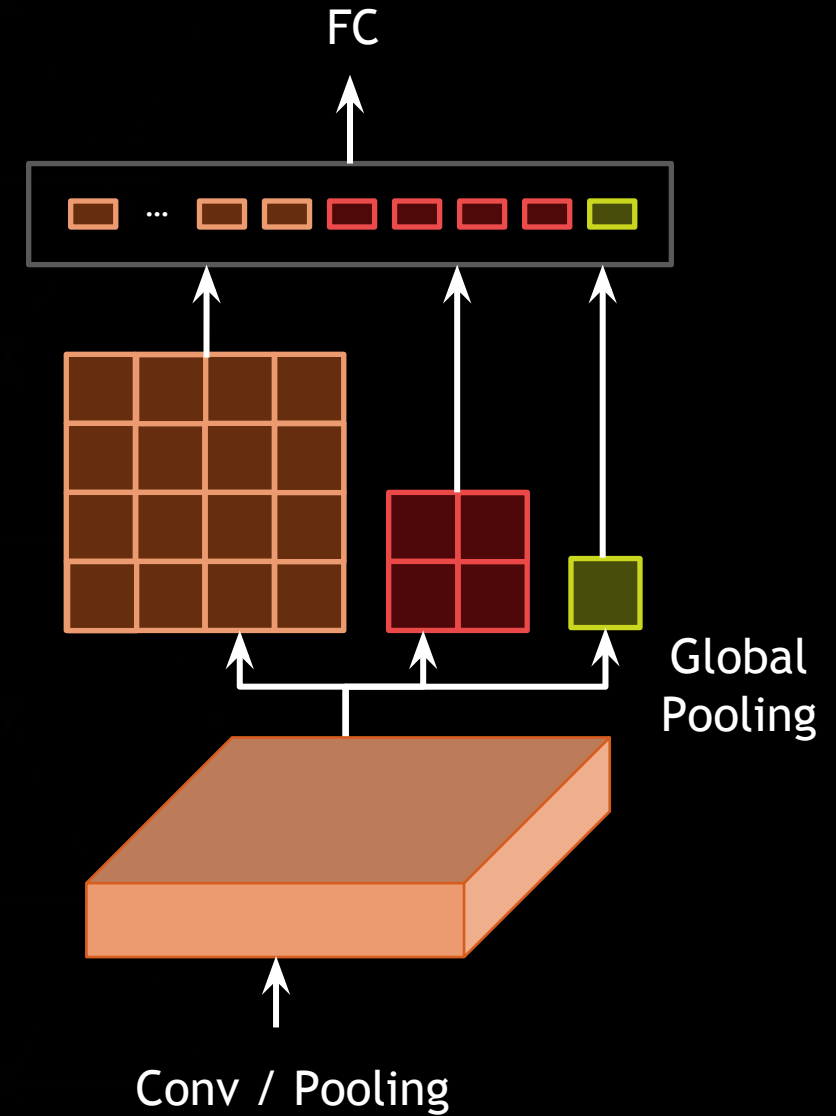
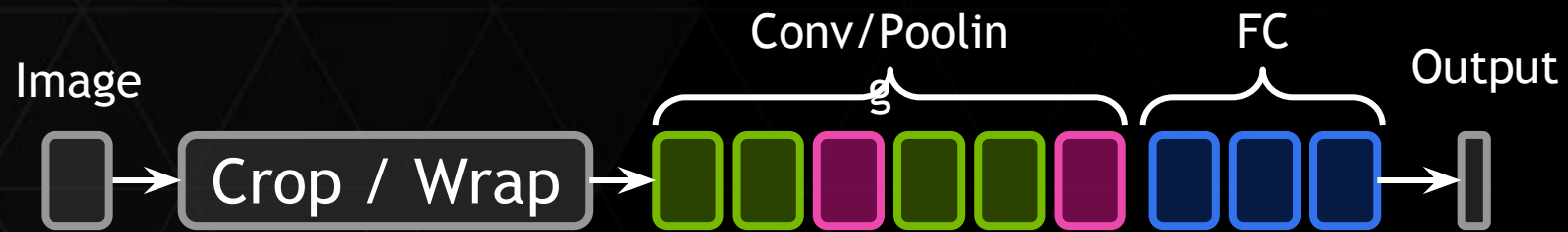


# SPATIAL PYRAMID POOLING (2014)

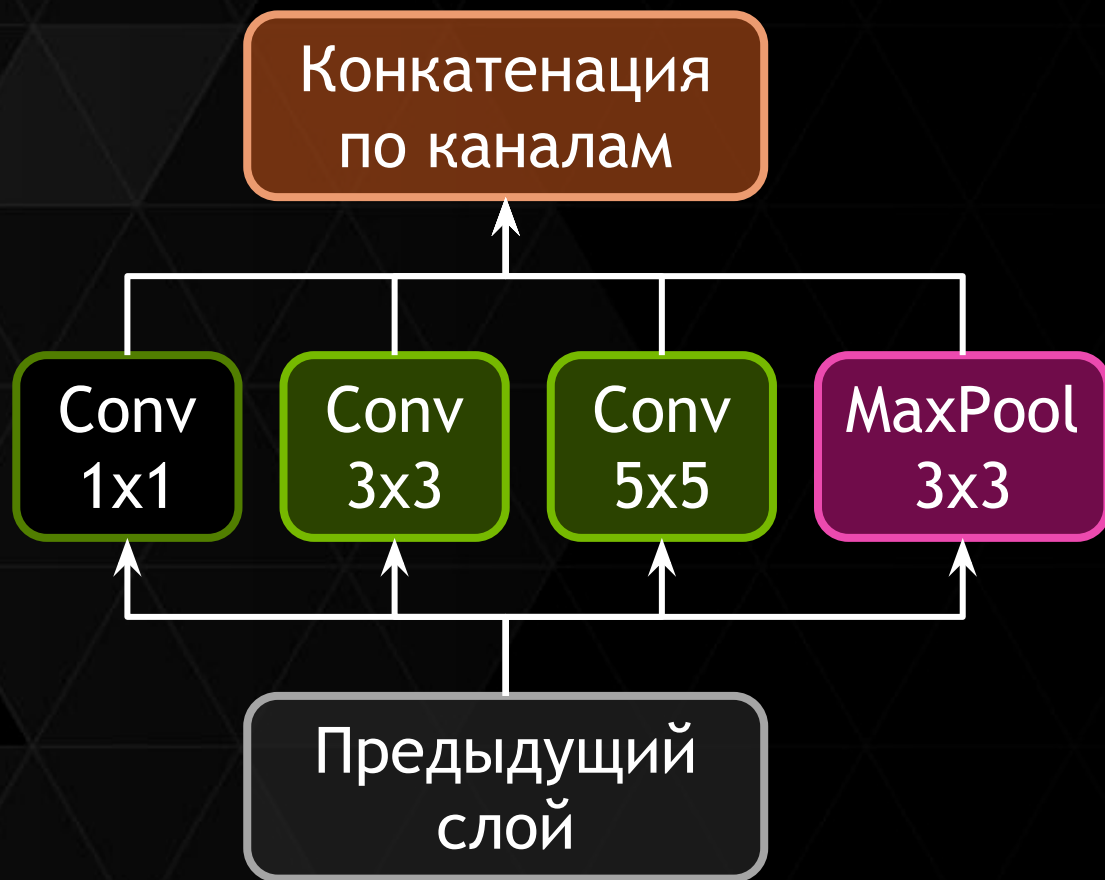
Crop



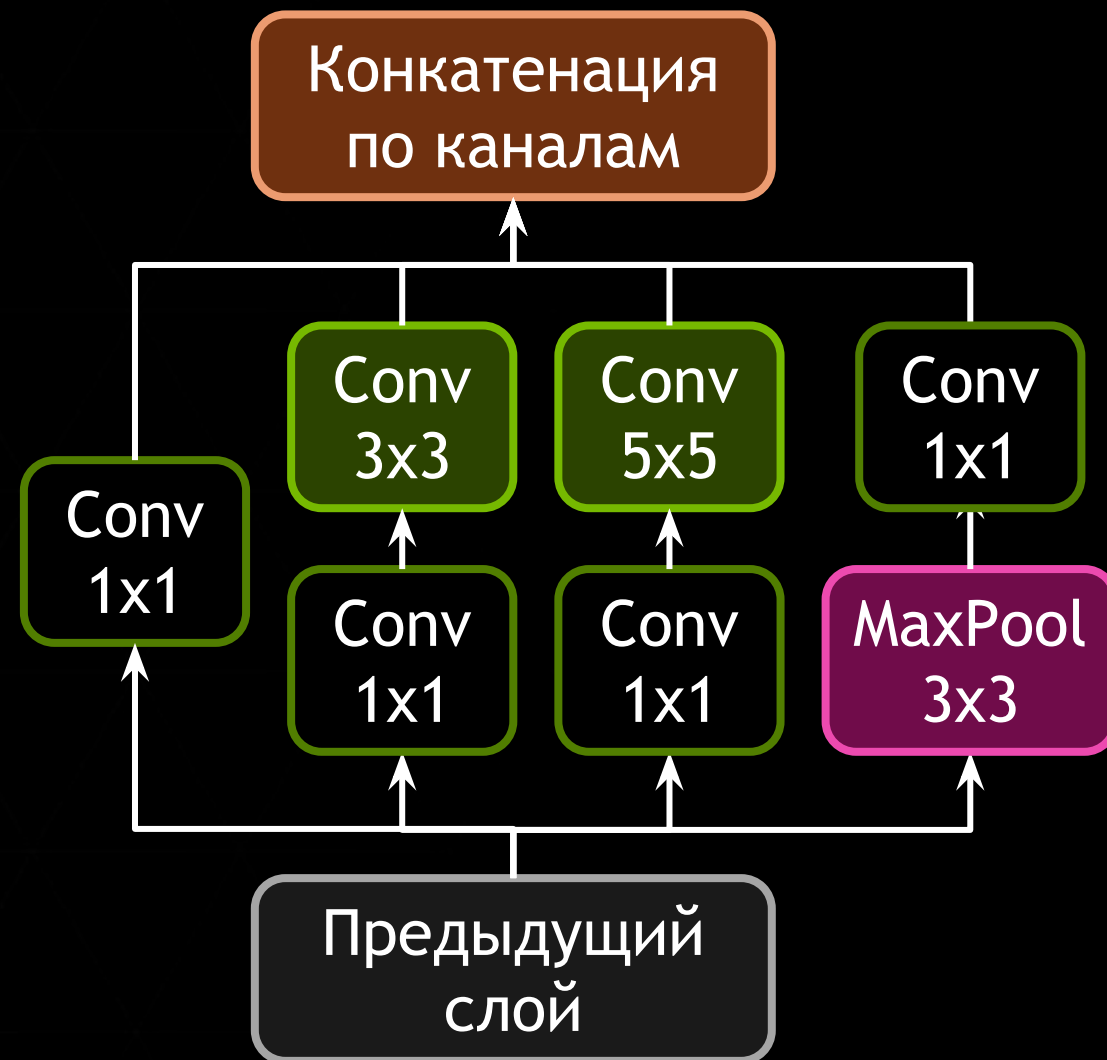
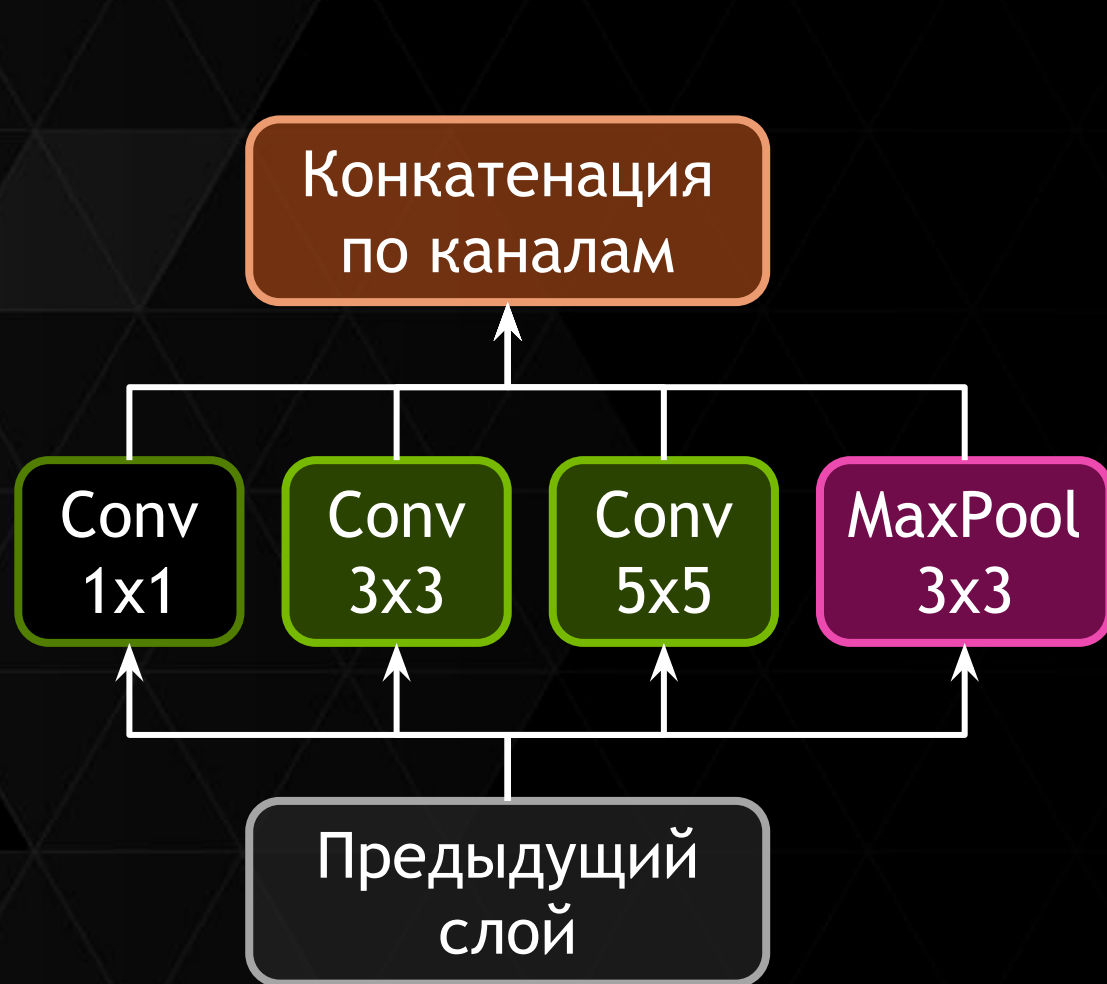
Wrap



# GOOGLNET / INCEPTION (2014)

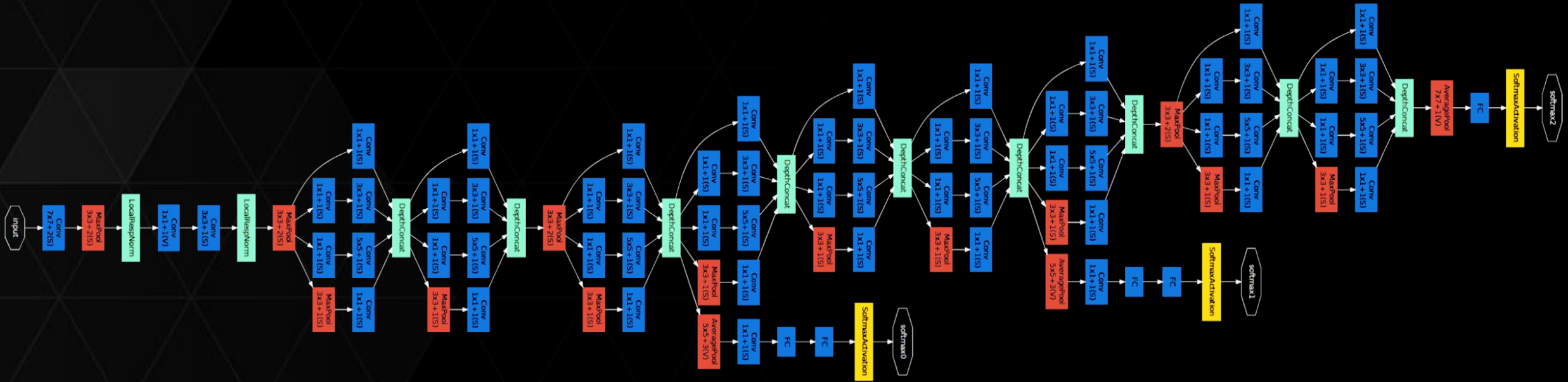


# GOOGLNET / INCEPTION (2014)

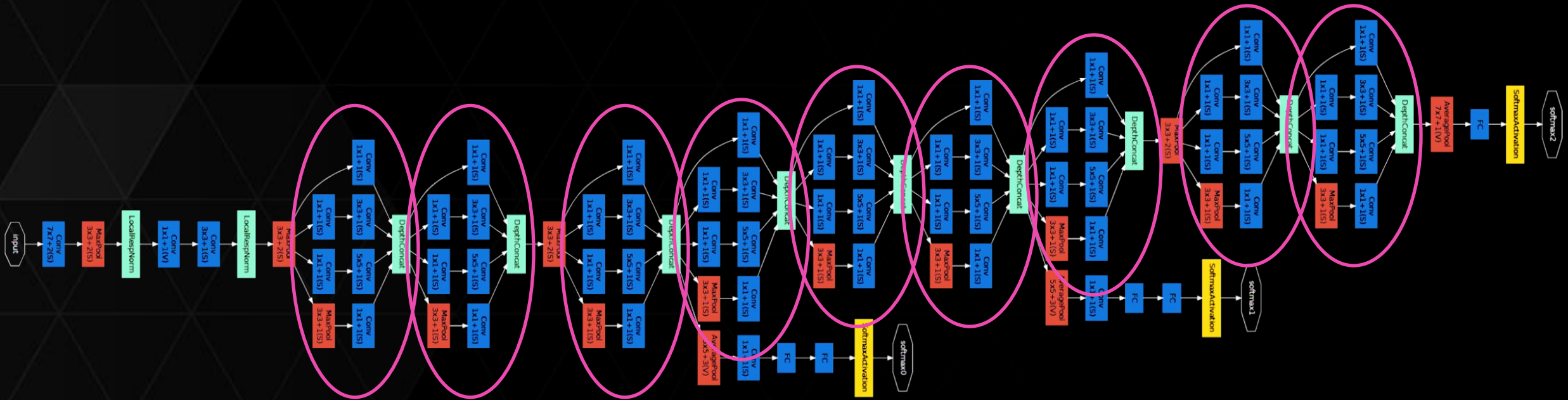




# GOOGLNET / INCEPTION (2014)

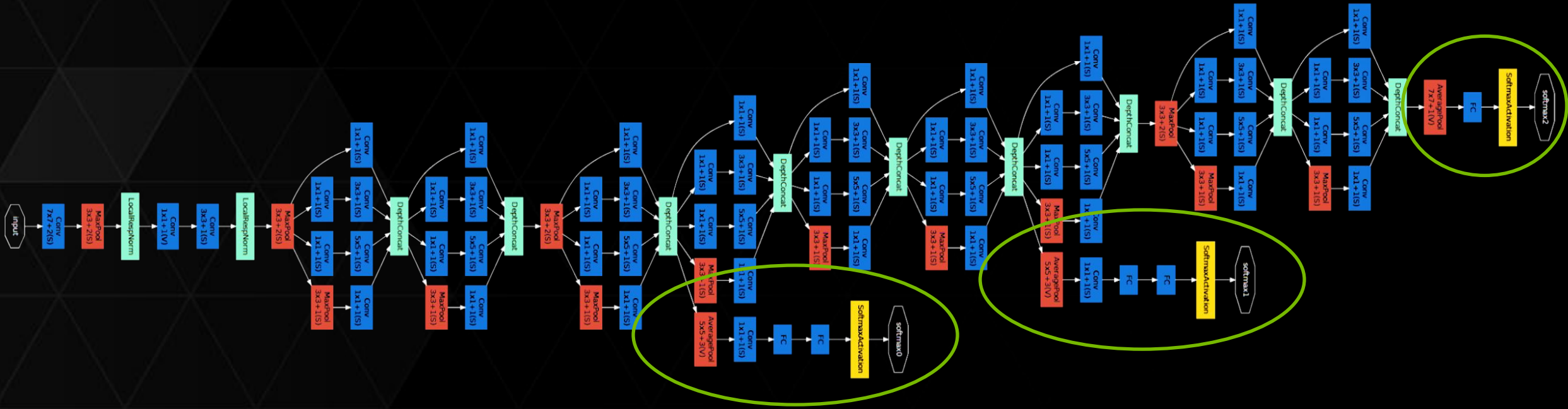


# GOOGLNET / INCEPTION (2014)



- Стек из Inception блоков

# GOOGLNET / INCEPTION (2014)

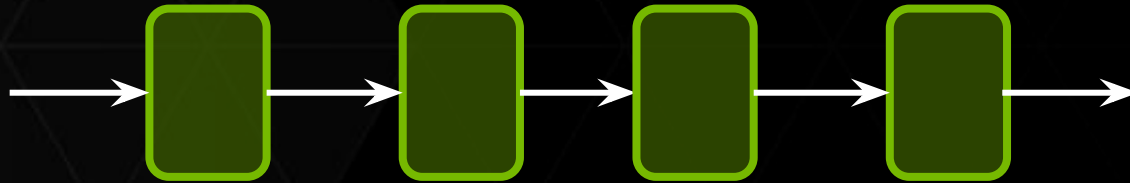
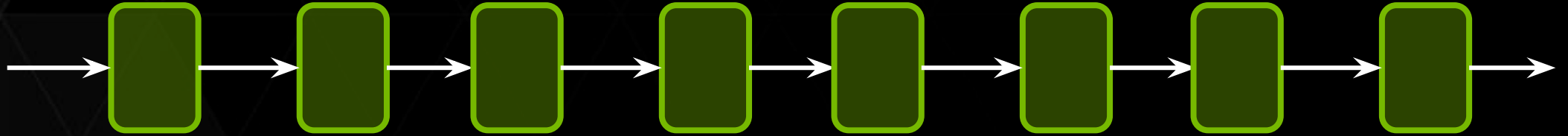


- **Стек из Inception блоков**
- **Несколько выходов на лосс из середины сети**

# INCEPTION-V3

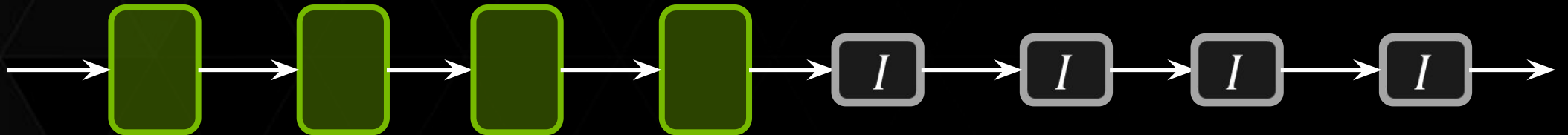
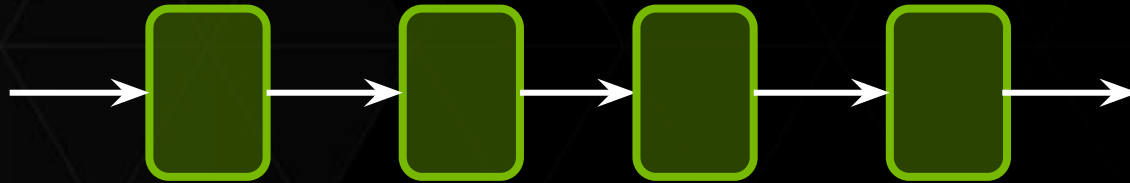
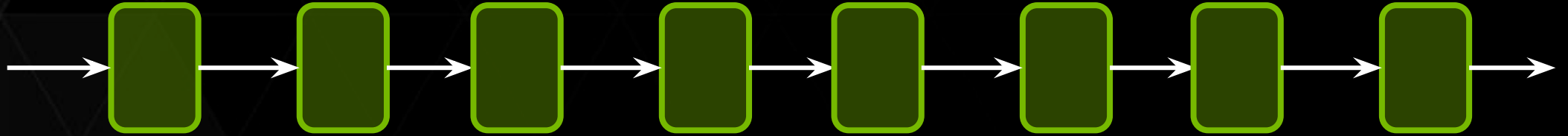


# RESNET (2015)

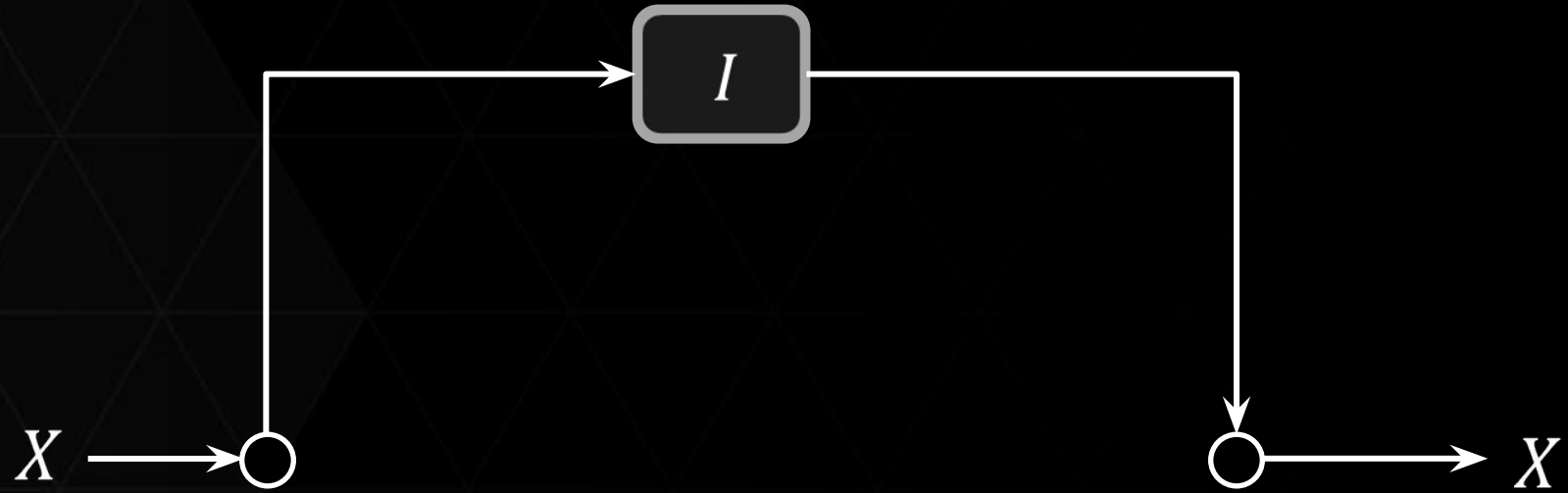




# RESNET (2015)

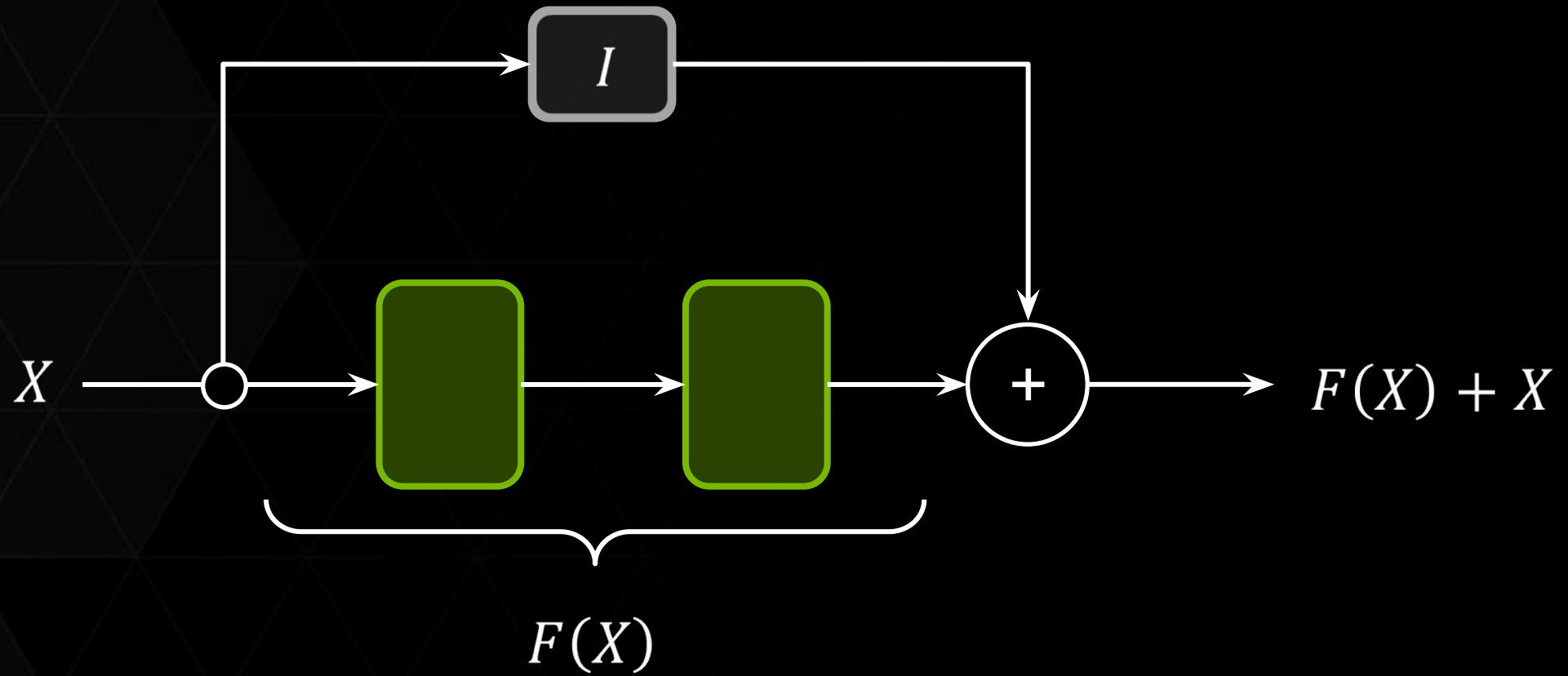


# RESNET (2015)



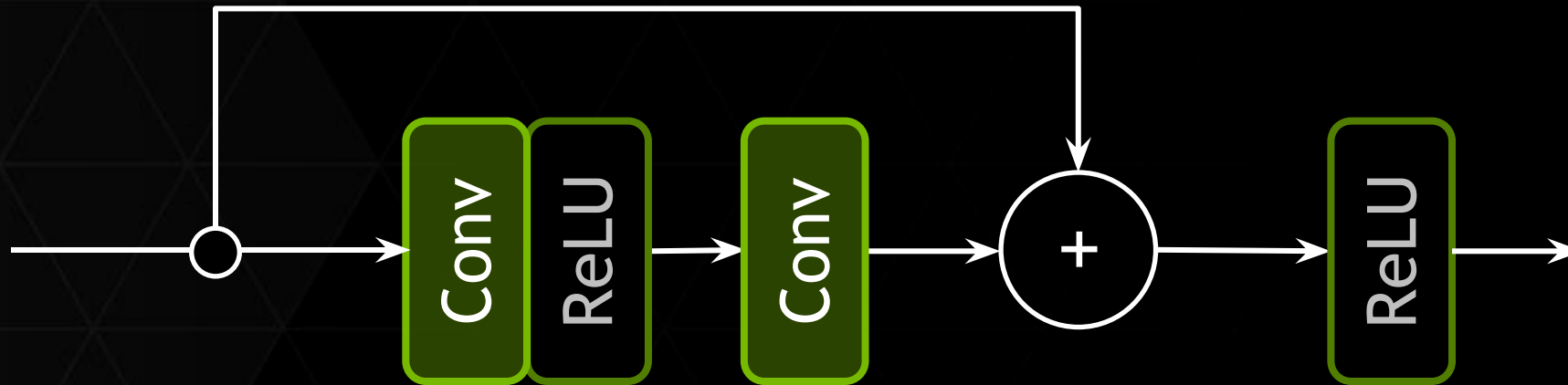
# RESNET (2015)

Residual block



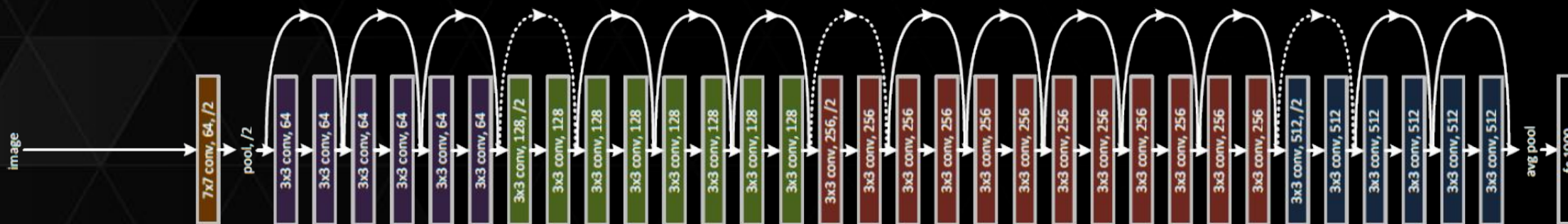
# RESNET (2015)

Residual block

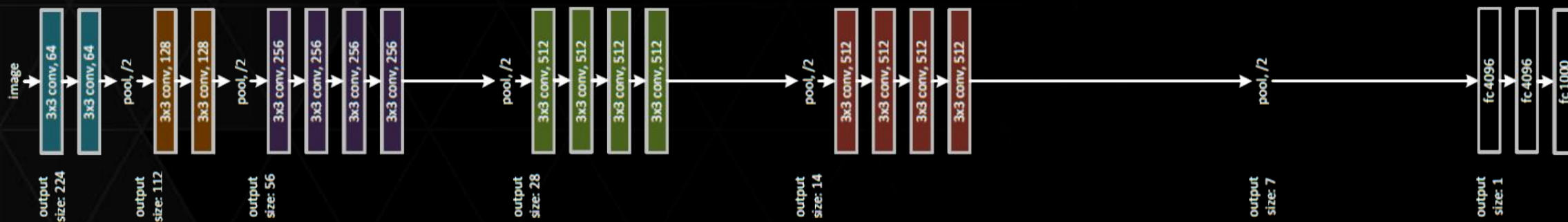


# RESNET (2015)

ResNet-34

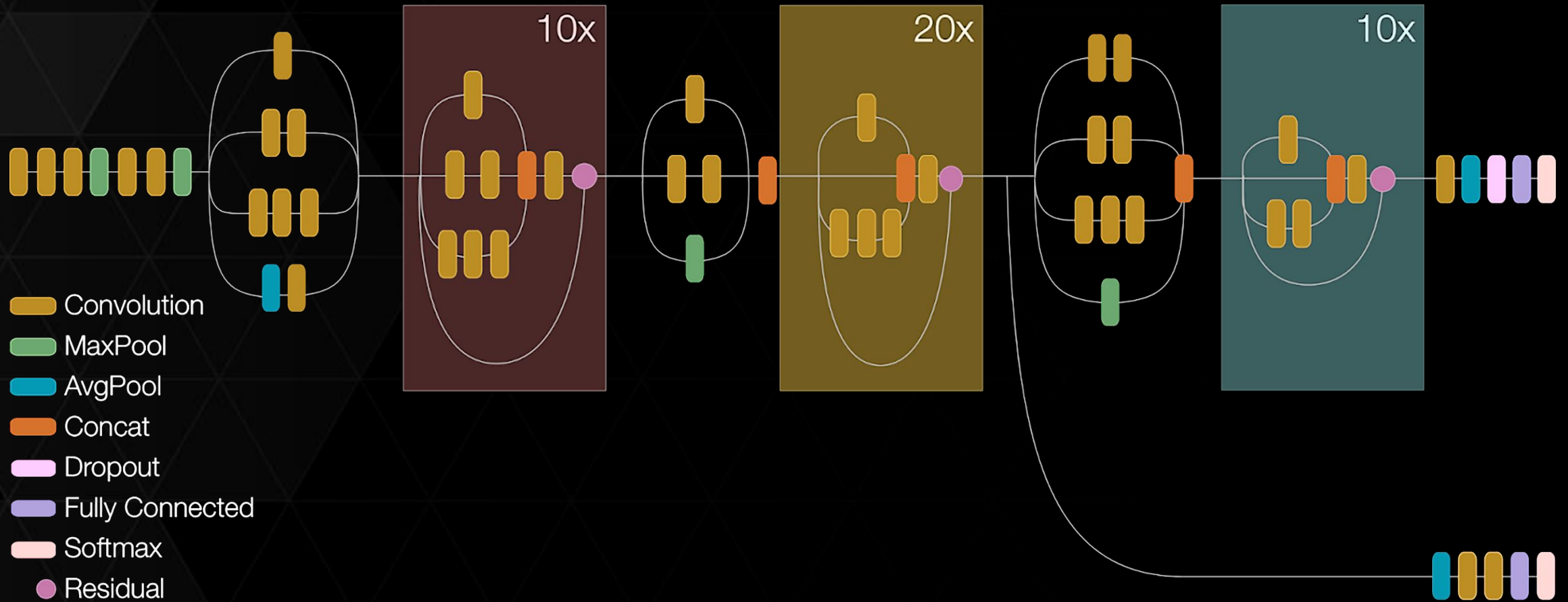


VGG-19

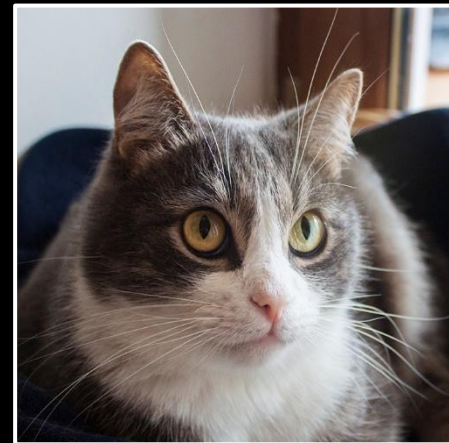
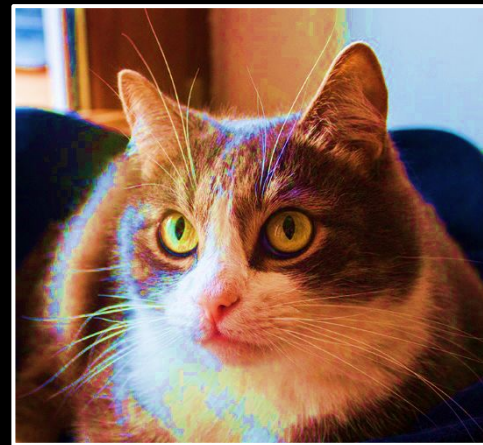




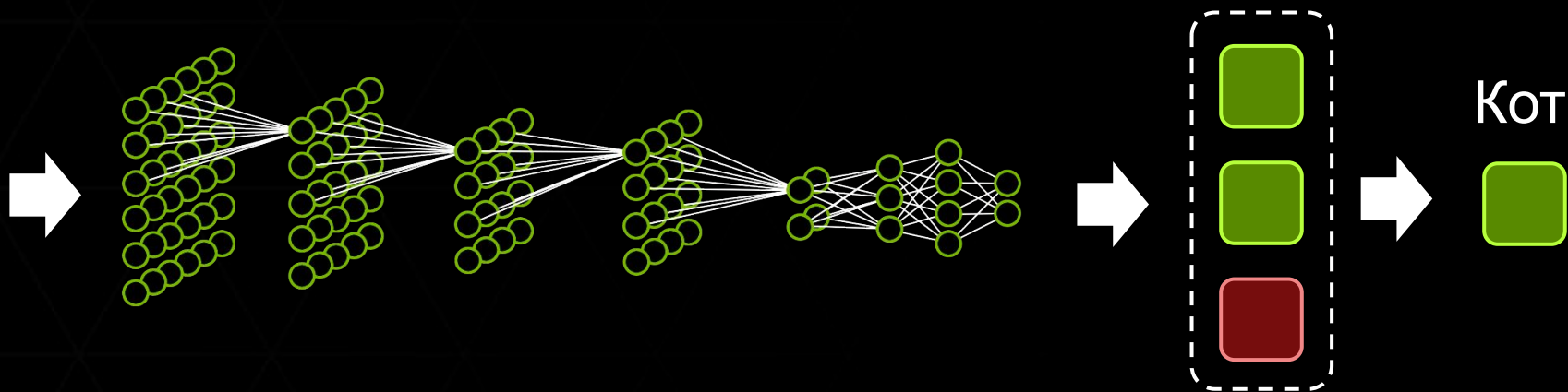
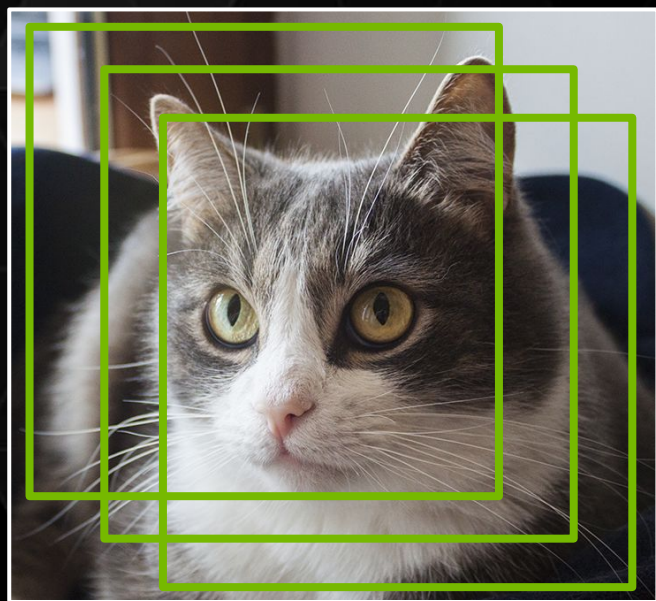
# INCEPTION-RESNET-V2 (2016)



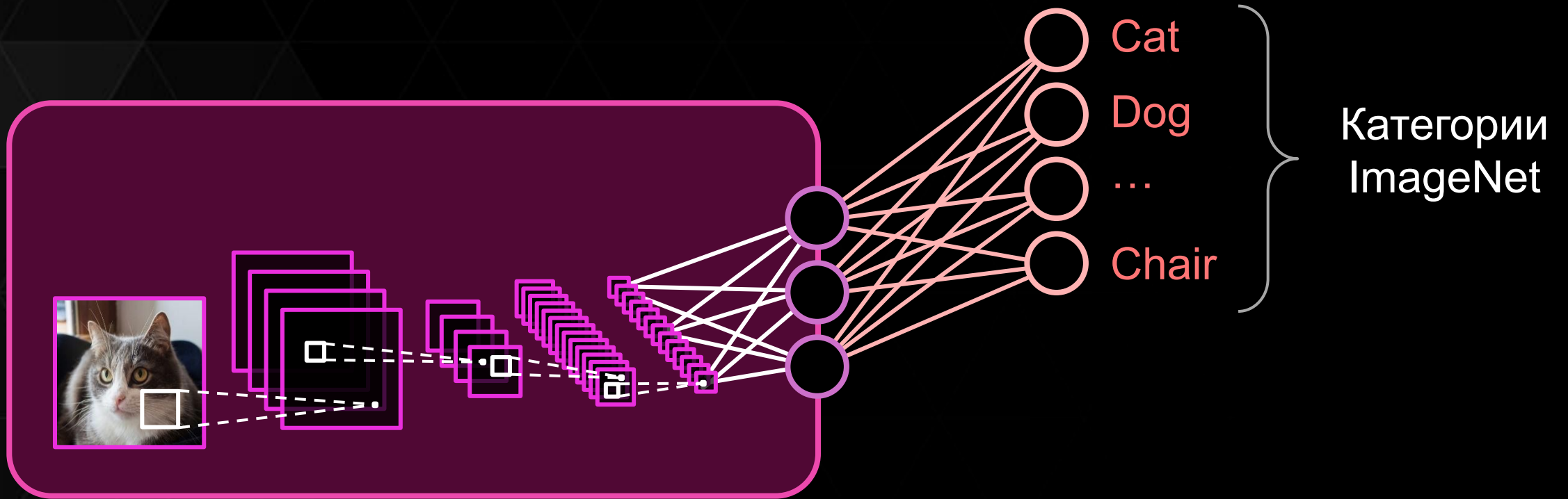
# АУГМЕНТАЦИЯ ДАННЫХ: ОБУЧЕНИЕ



# АУГМЕНТАЦИЯ ДАННЫХ: ИНФЕРЕНС

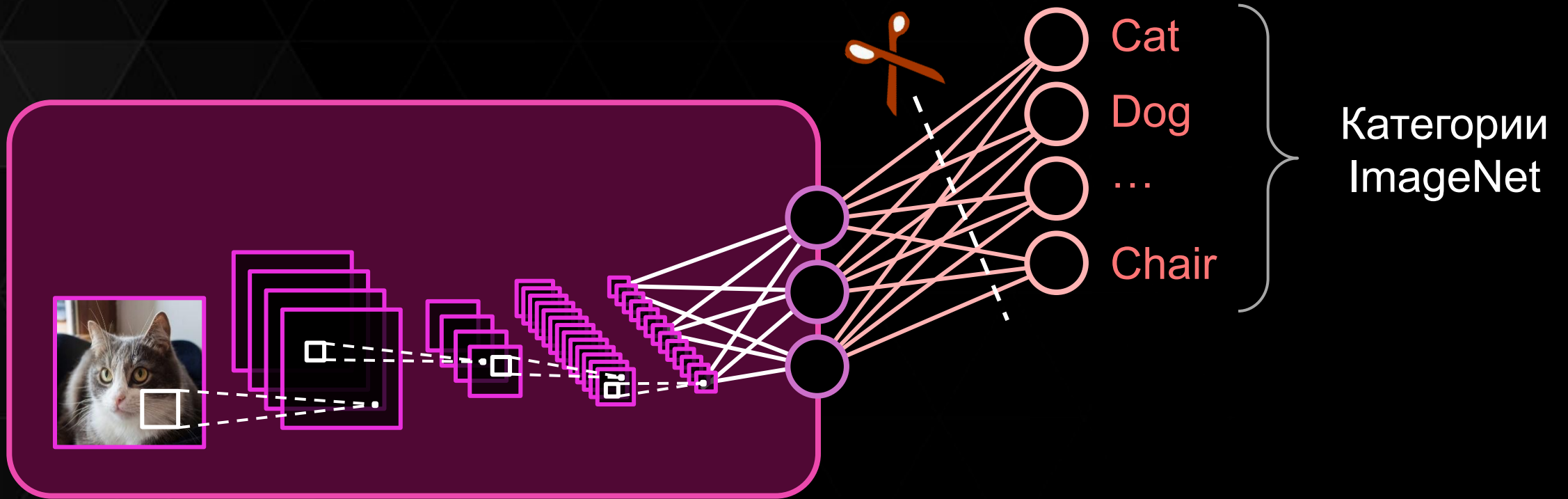


# TRANSFER LEARNING



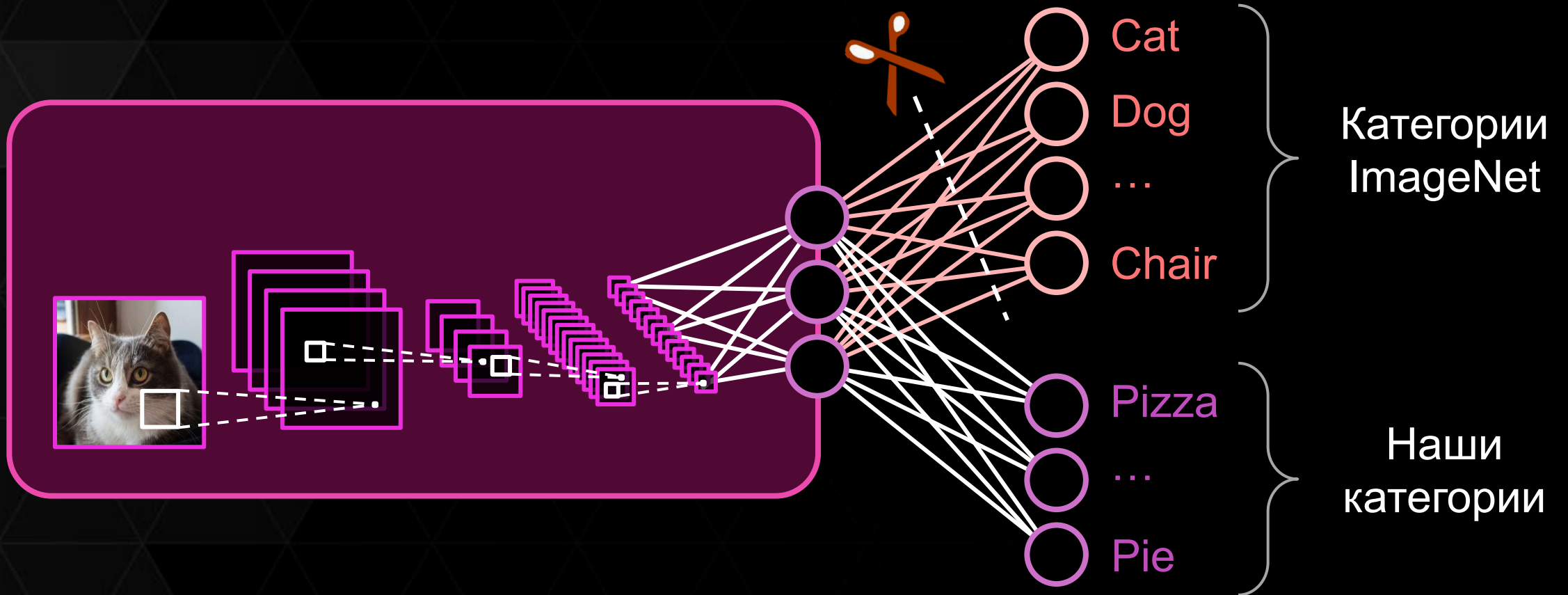


# TRANSFER LEARNING





# TRANSFER LEARNING



The background is a dark blue field with a complex network of thin, light green lines. These lines connect various points, some of which are highlighted as bright green dots. The overall effect is a sense of a digital or neural network. The word "СПАСИБО" is centered in the middle of the image in a white, sans-serif font.

СПАСИБО