



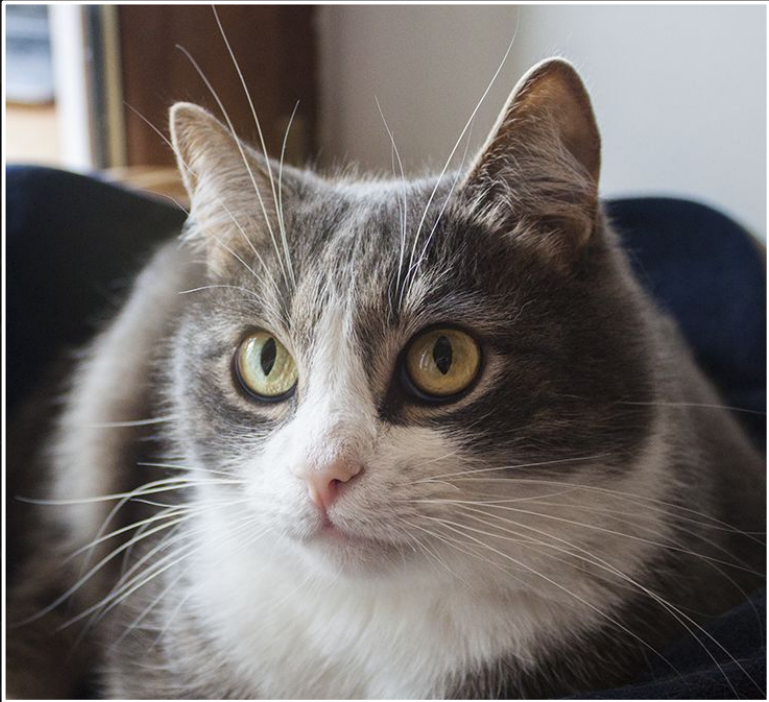
[4] СЕМАНТИЧЕСКАЯ СЕГМЕНТАЦИЯ

Дмитрий Коробченко

The background is a dark blue gradient with a complex network of thin, light green lines crisscrossing across the frame. Several bright green circular nodes of varying sizes are positioned at various points, some acting as endpoints for the lines. The overall effect is a digital or network-like aesthetic.

КОМПЬЮТЕРНОЕ ЗРЕНИЕ

ЗАДАЧИ В КОМПЬЮТЕРНОМ ЗРЕНИИ



ЗАДАЧИ В КОМПЬЮТЕРНОМ ЗРЕНИИ

Классификация

Label



Котик

ЗАДАЧИ В КОМПЬЮТЕРНОМ ЗРЕНИИ

Классификация

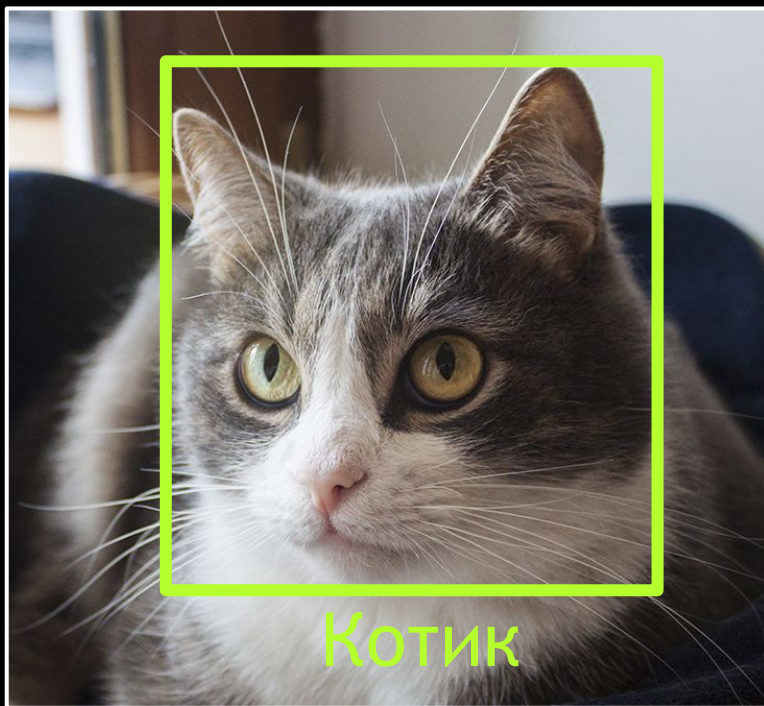
Label



Котик

Локализация [1]

Bounding box + Label



Котик

ЗАДАЧИ В КОМПЬЮТЕРНОМ ЗРЕНИИ

Классификация

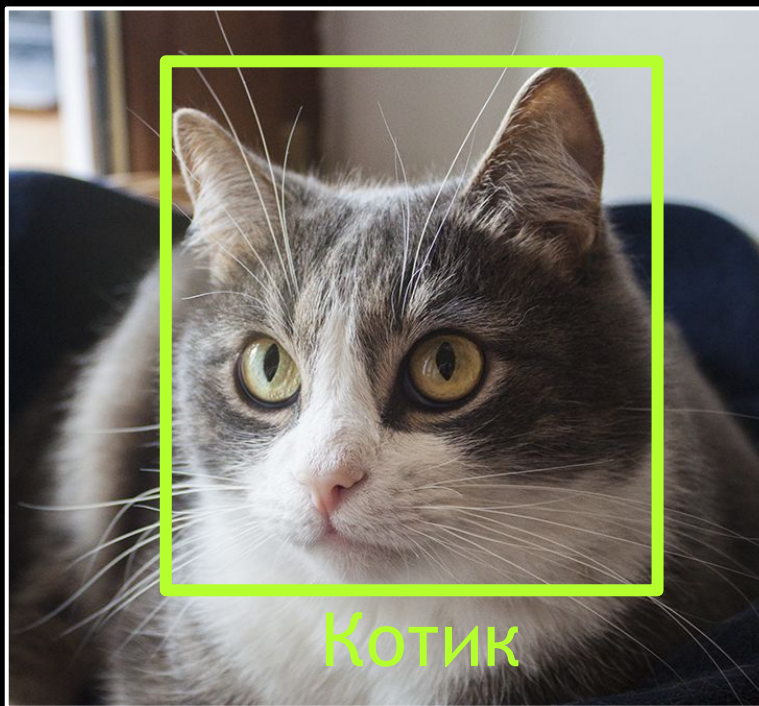
Label



Котик

Локализация [1]

Bounding box + Label



Котик

Локализация [2]

Heatmap

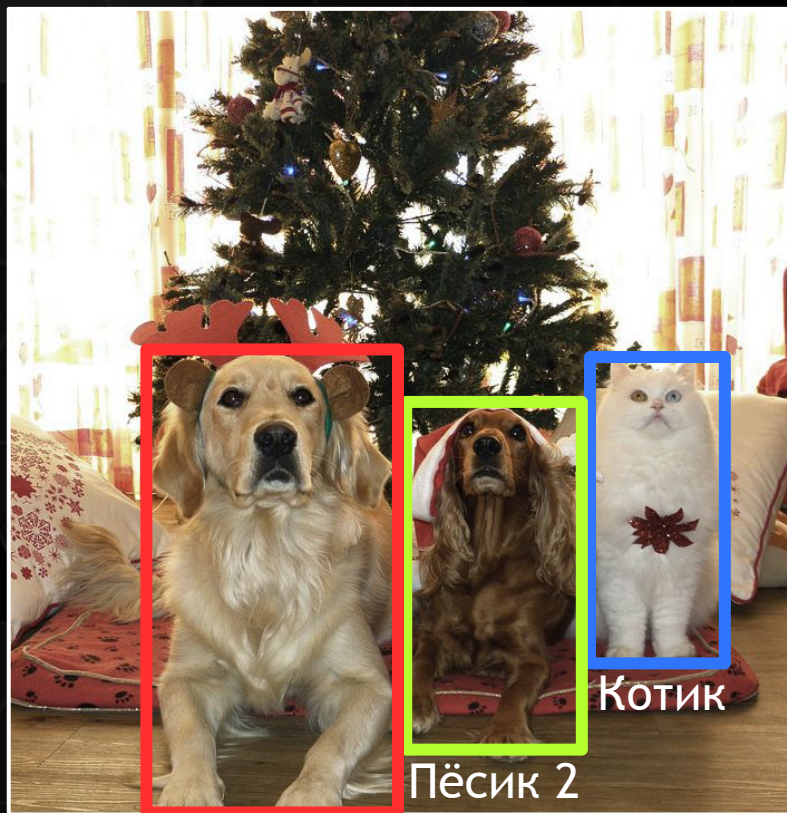


ЗАДАЧИ В КОМПЬЮТЕРНОМ ЗРЕНИИ



ЗАДАЧИ В КОМПЬЮТЕРНОМ ЗРЕНИИ

Детектирование объектов



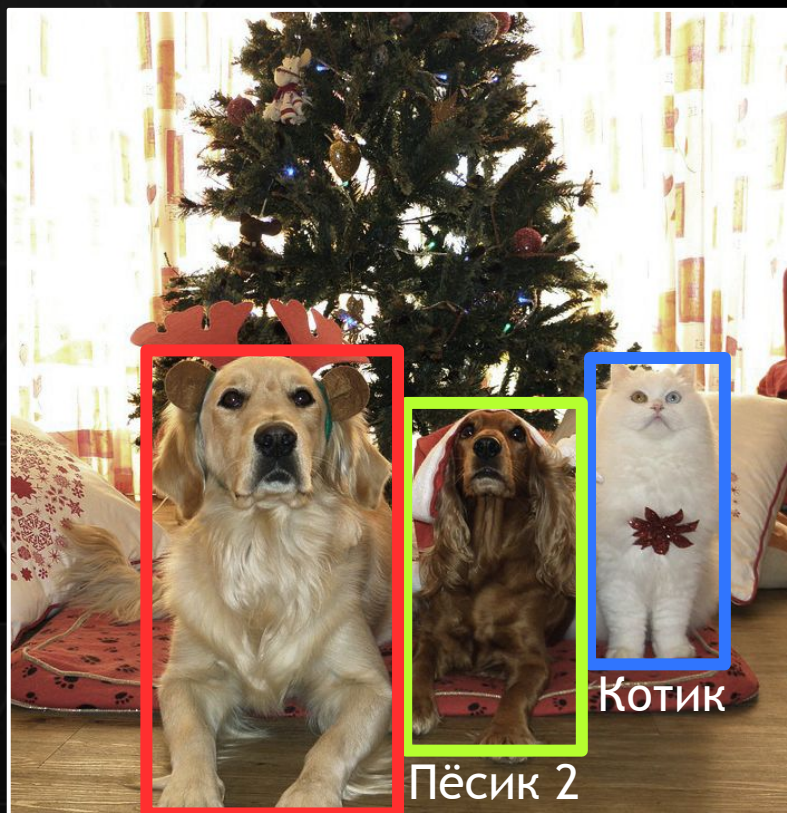
Пёсик 1

Пёсик 2

Котик

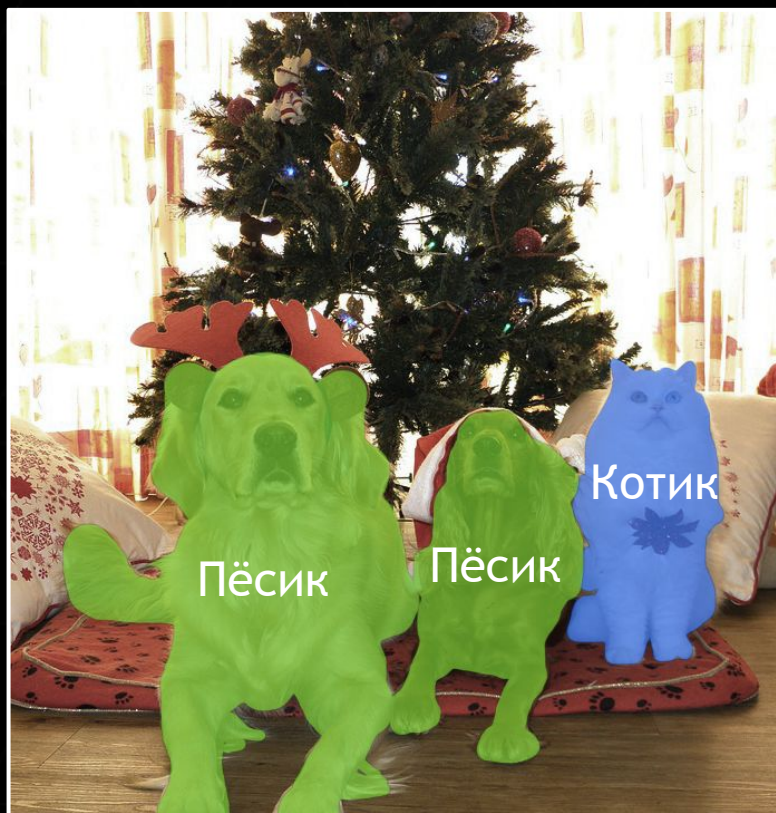
ЗАДАЧИ В КОМПЬЮТЕРНОМ ЗРЕНИИ

Детектирование
объектов



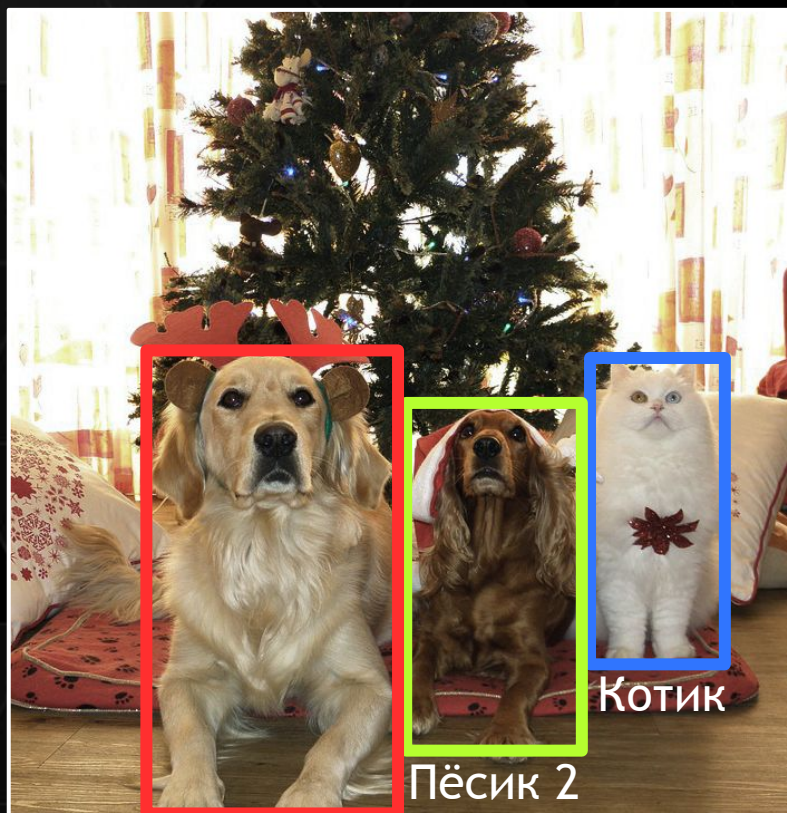
Пёсик 1

Семантическая
сегментация



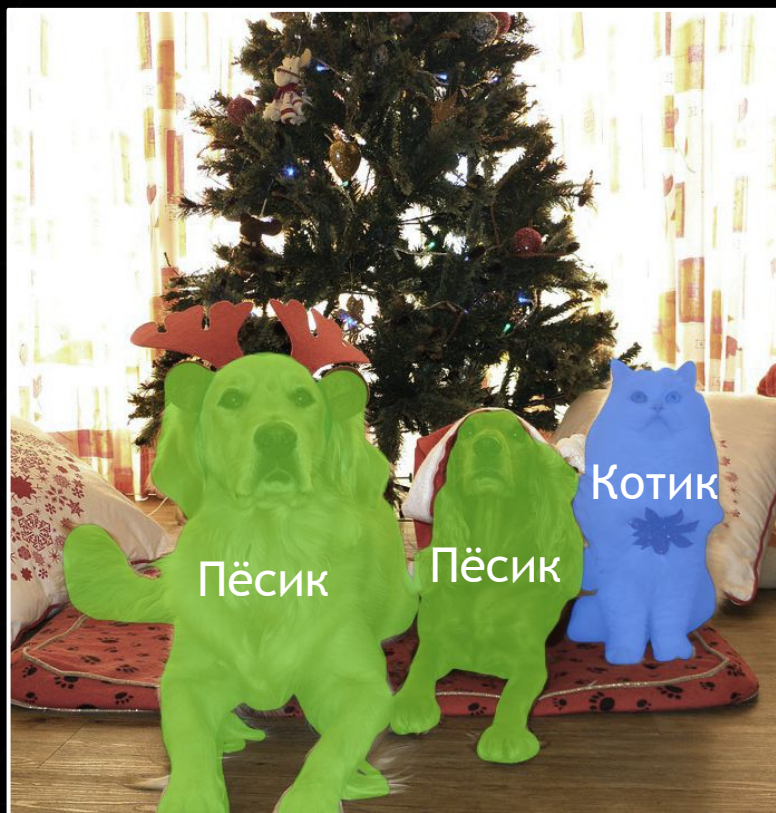
ЗАДАЧИ В КОМПЬЮТЕРНОМ ЗРЕНИИ

Детектирование объектов



Пёсик 1

Семантическая сегментация

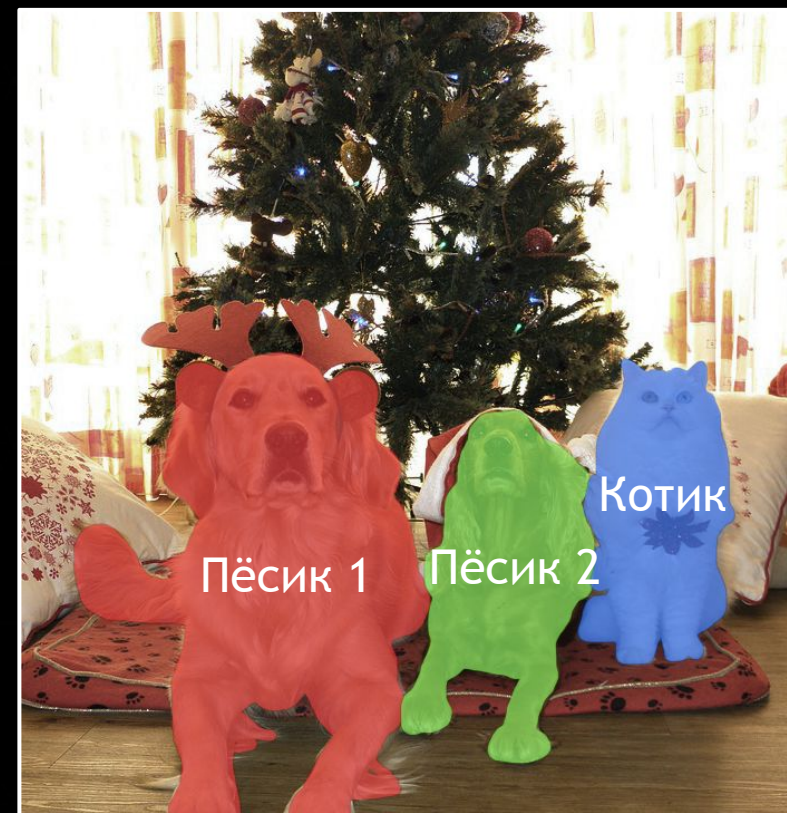


Пёсик

Пёсик

Котик

Сегментация объектов

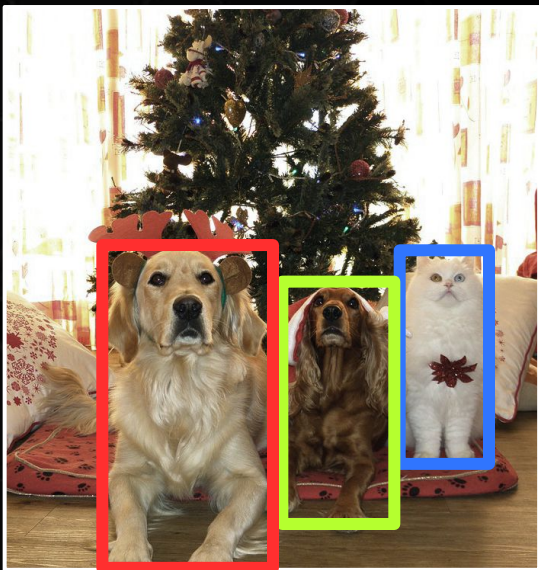
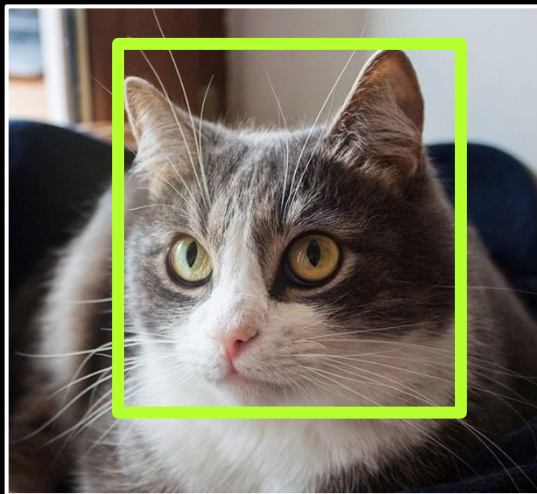


Пёсик 1

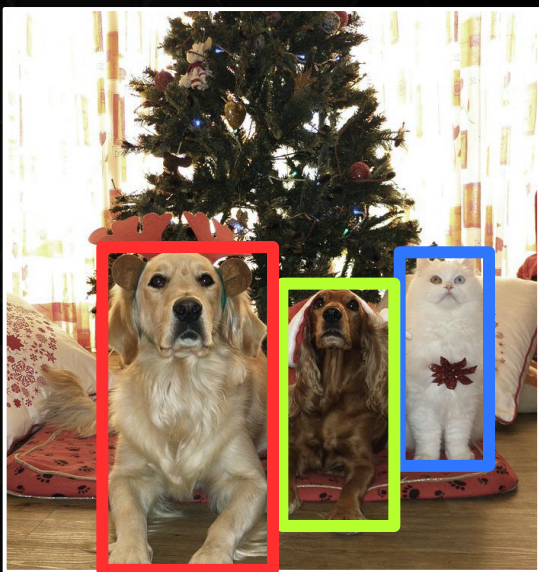
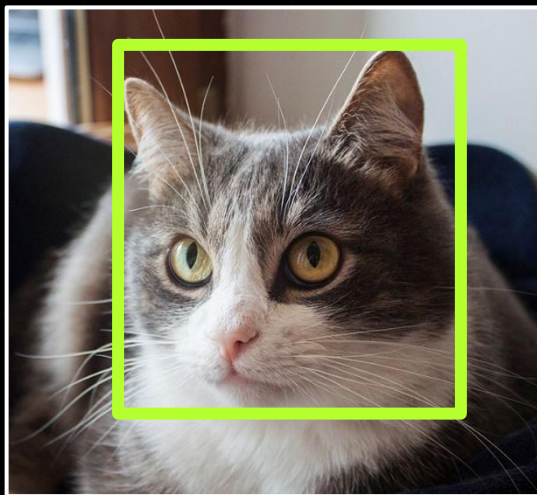
Пёсик 2

Котик

ЗАДАЧИ В КОМПЬЮТЕРНОМ ЗРЕНИИ



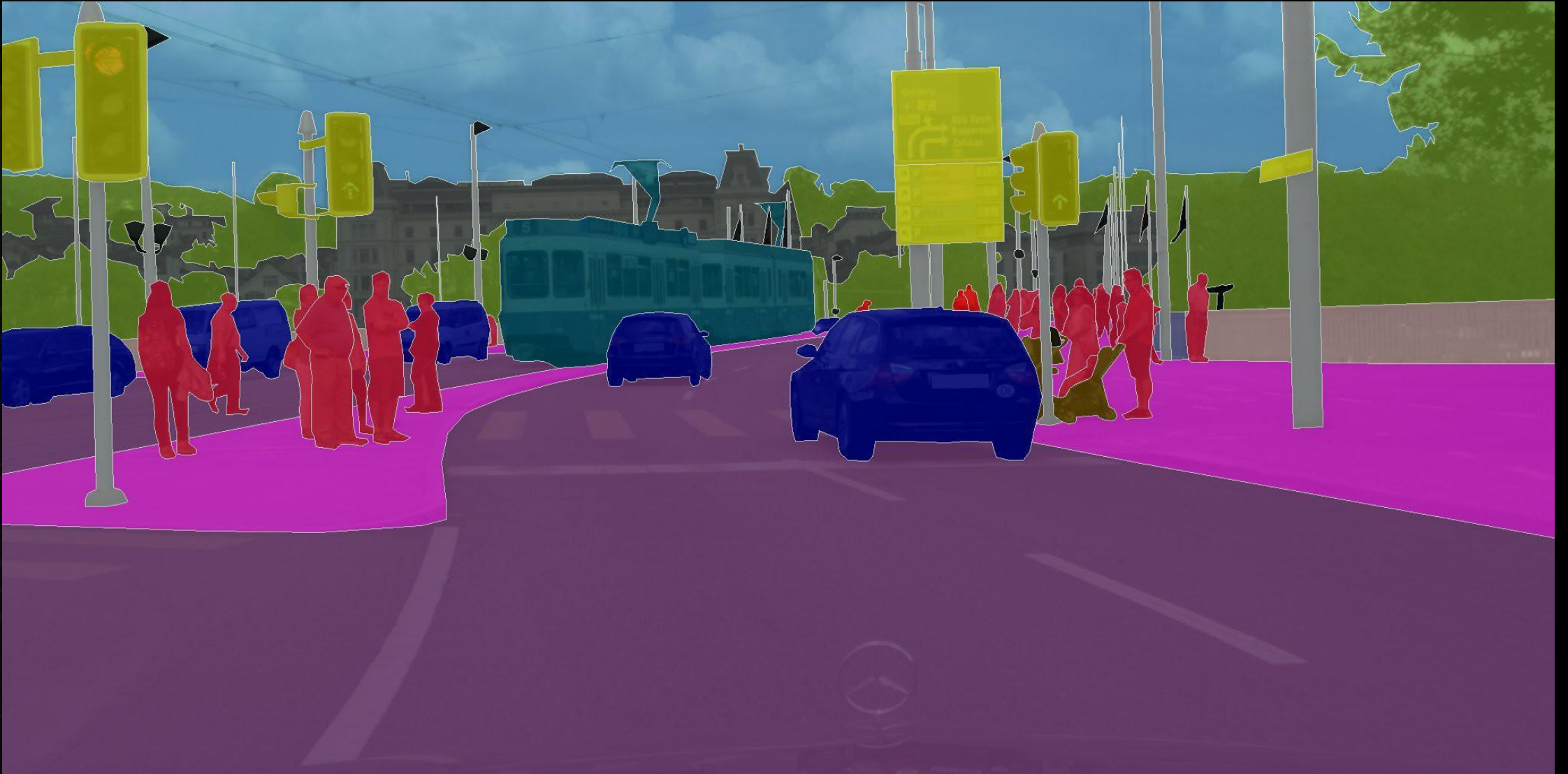
ЗАДАЧИ В КОМПЬЮТЕРНОМ ЗРЕНИИ



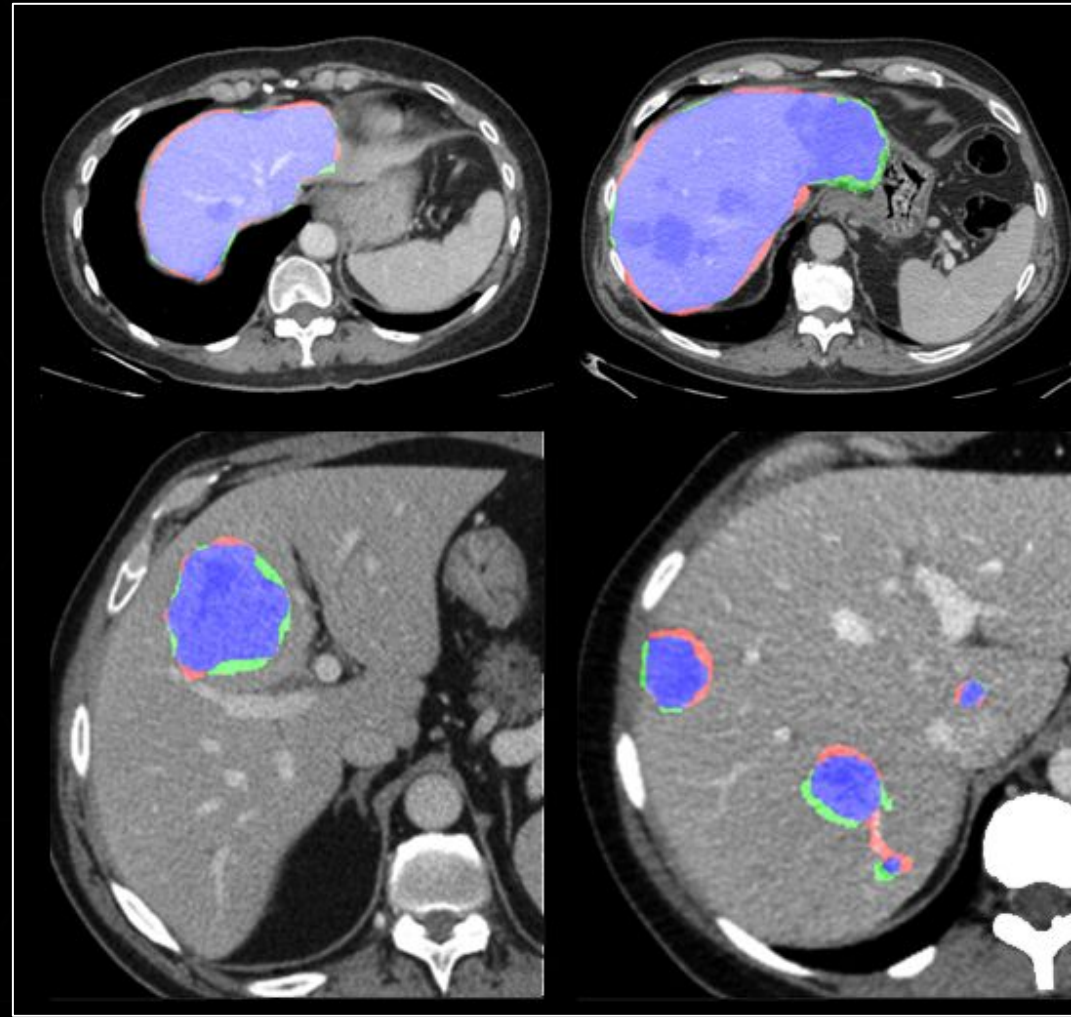
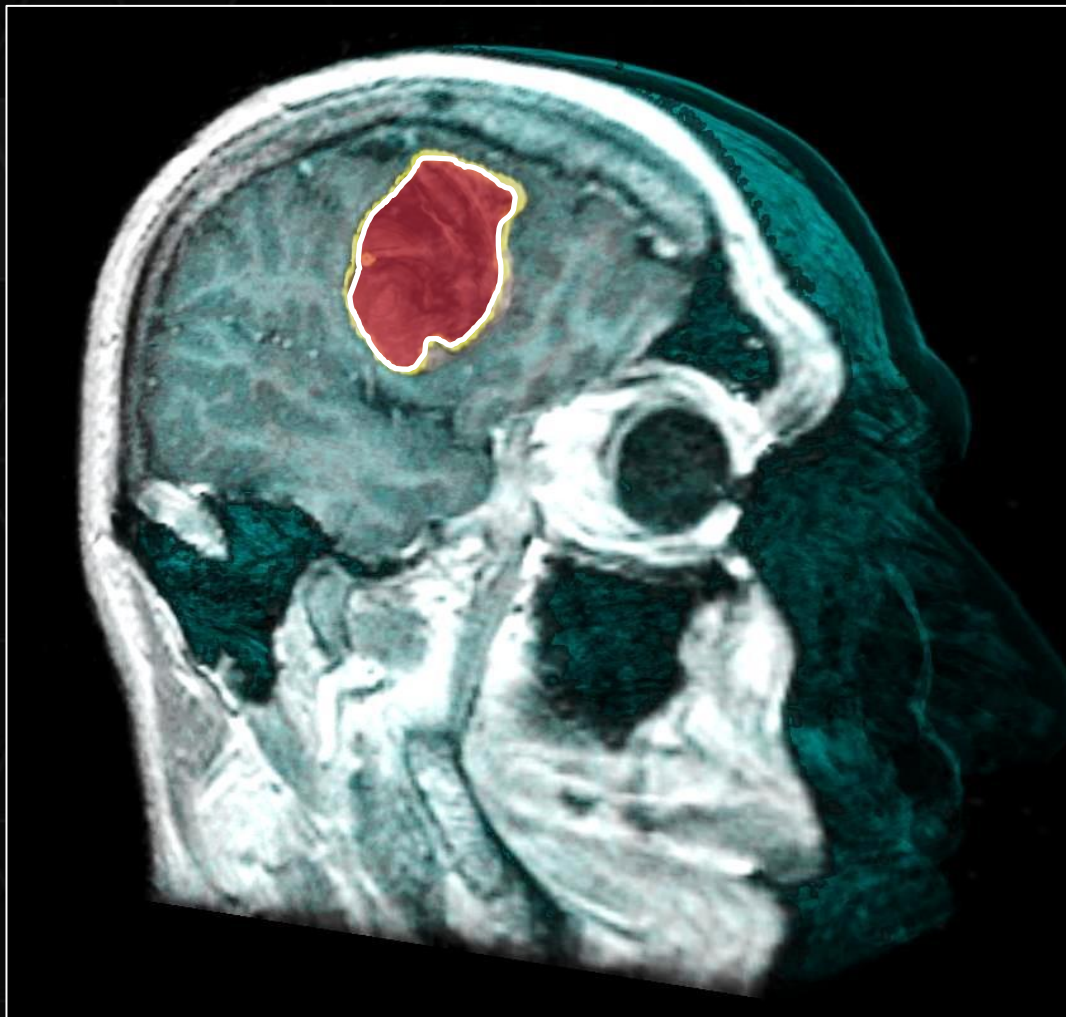
The background is a dark blue field with a network of thin, light green lines connecting various points. Some points are small, bright green dots, while others are larger, fainter blue-green circles. The lines crisscross the frame, creating a complex, web-like pattern.

ПРИЛОЖЕНИЯ СЕГМЕНТАЦИИ

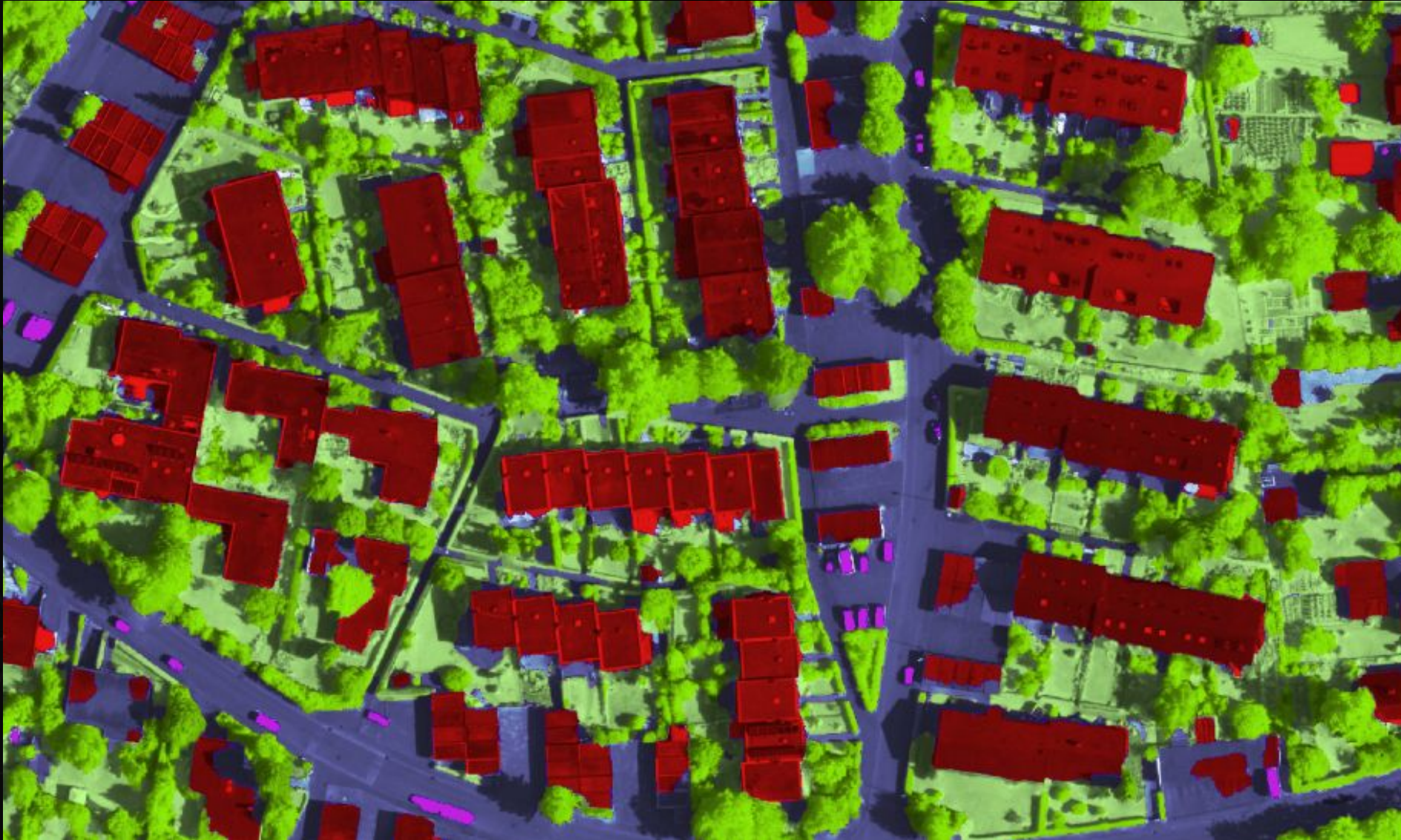
АВТОНОМНЫЕ ТРАНСПОРТНЫЕ СРЕДСТВА



МЕДИЦИНСКИЕ ИЗОБРАЖЕНИЯ



СПУТНИКОВАЯ И АЭРОФОТОСЪЁМКА



СЕГМЕНТАЦИЯ ЛЮДЕЙ



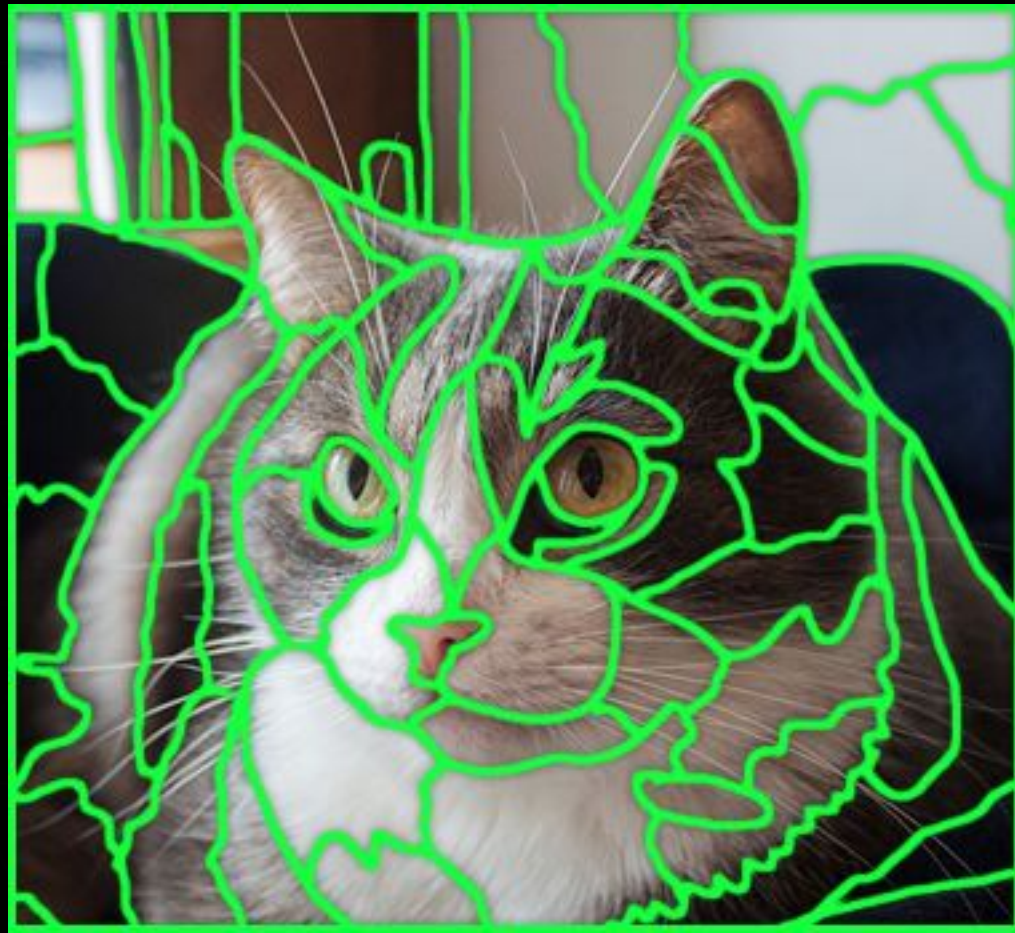
The background is a dark blue gradient with a network of thin, light green lines connecting various points. Some points are small, bright green dots, while others are larger, fainter blue circles. The overall effect is a complex, interconnected web of light against a dark background.

КЛАССИЧЕСКАЯ СЕГМЕНТАЦИЯ

СУПЕРПИКСЕЛИ



СУПЕРПИКСЕЛИ

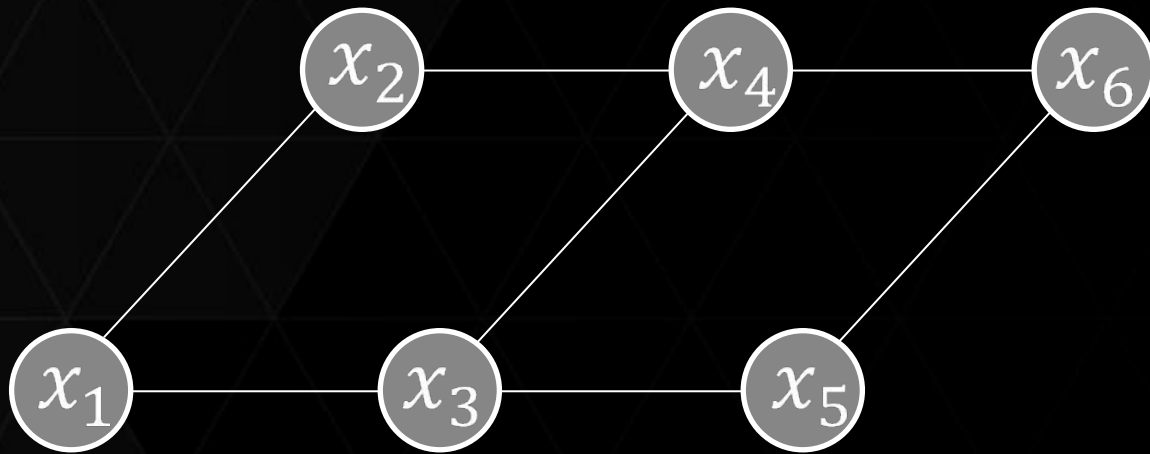


СУПЕРПИКСЕЛИ

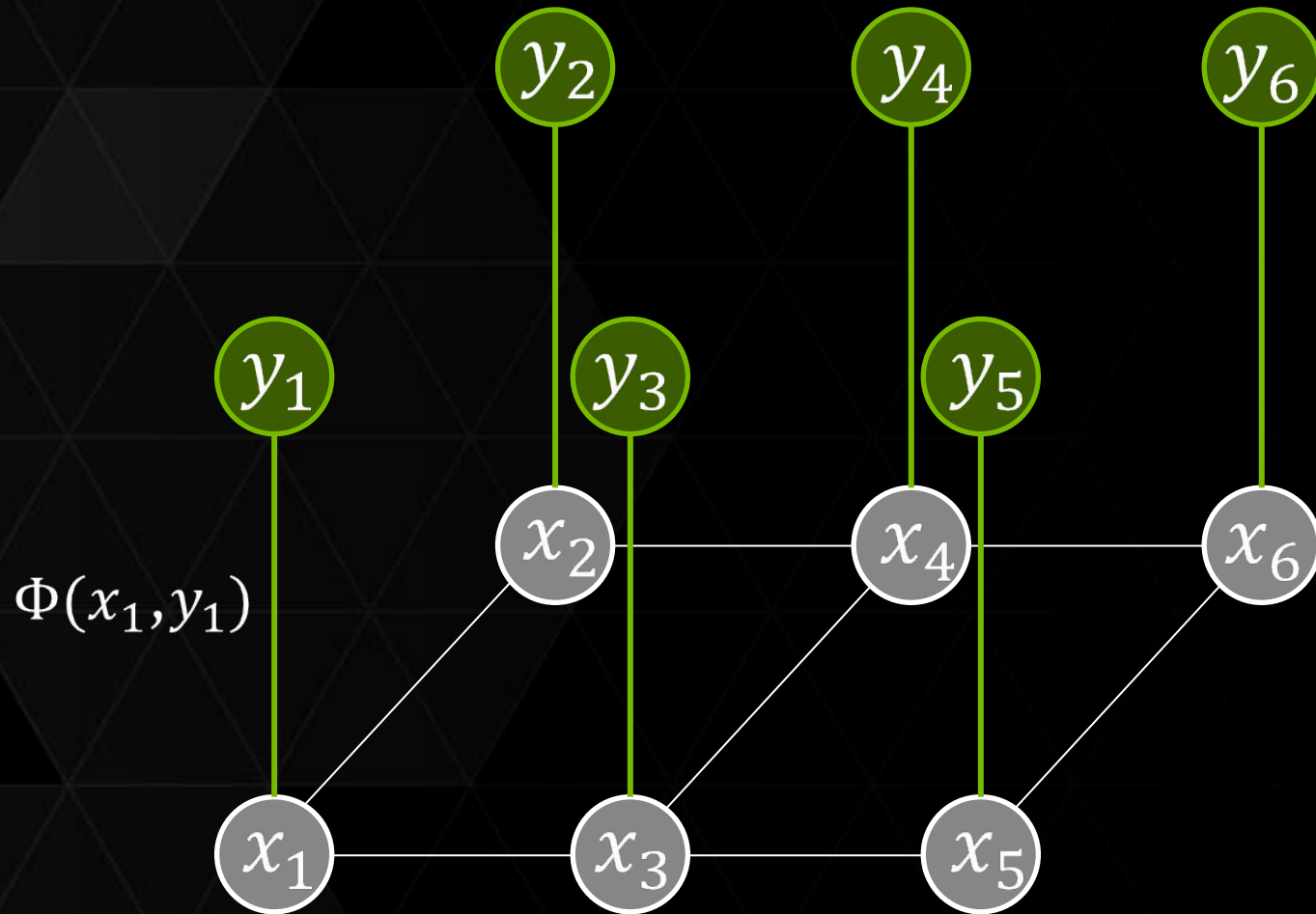


УСЛОВНЫЕ СЛУЧАЙНЫЕ ПОЛЯ (CRF)

- Скрытые значения
(искомые метки)



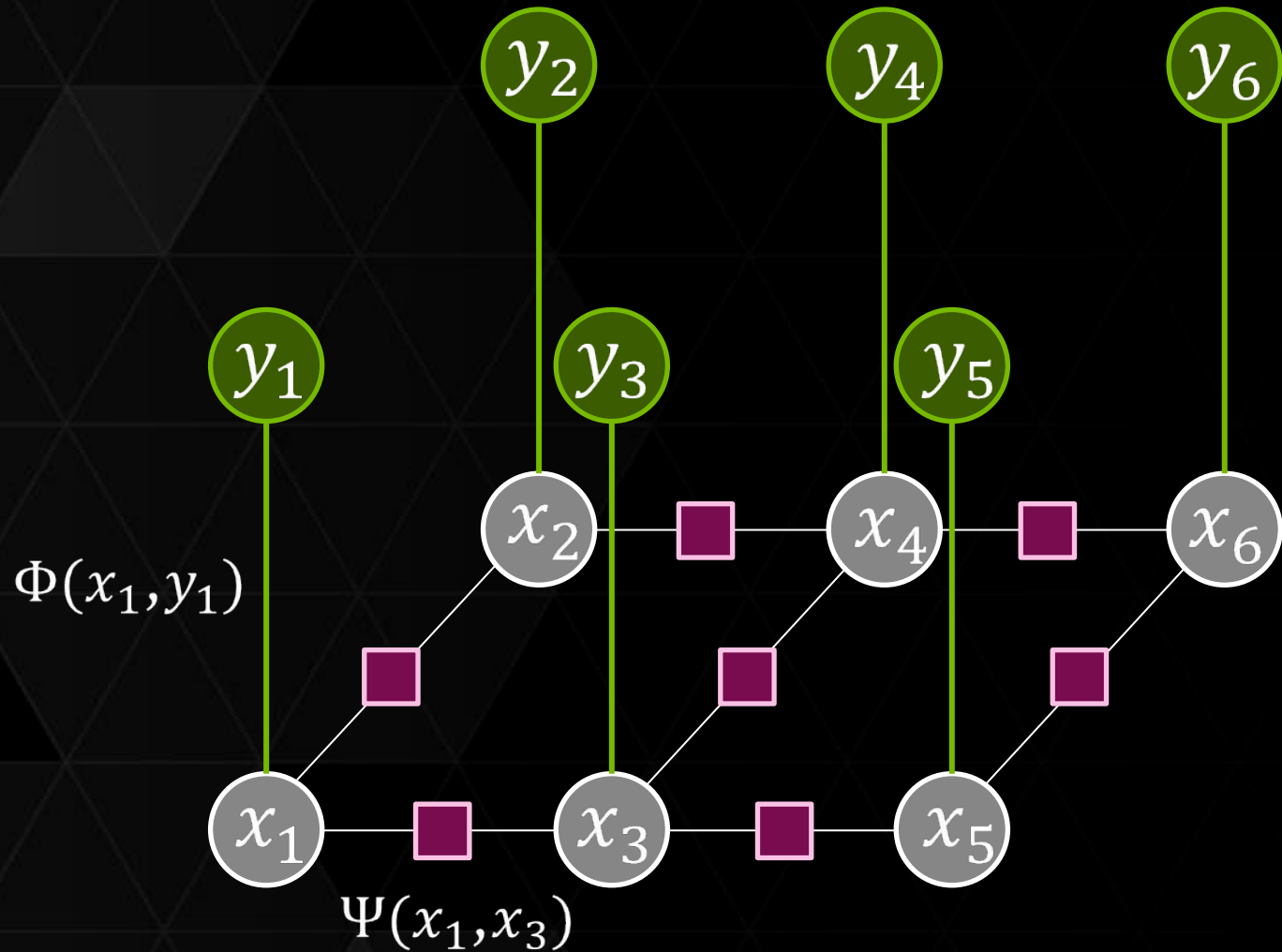
УСЛОВНЫЕ СЛУЧАЙНЫЕ ПОЛЯ (CRF)



○ Скрытые значения
(искомые метки)

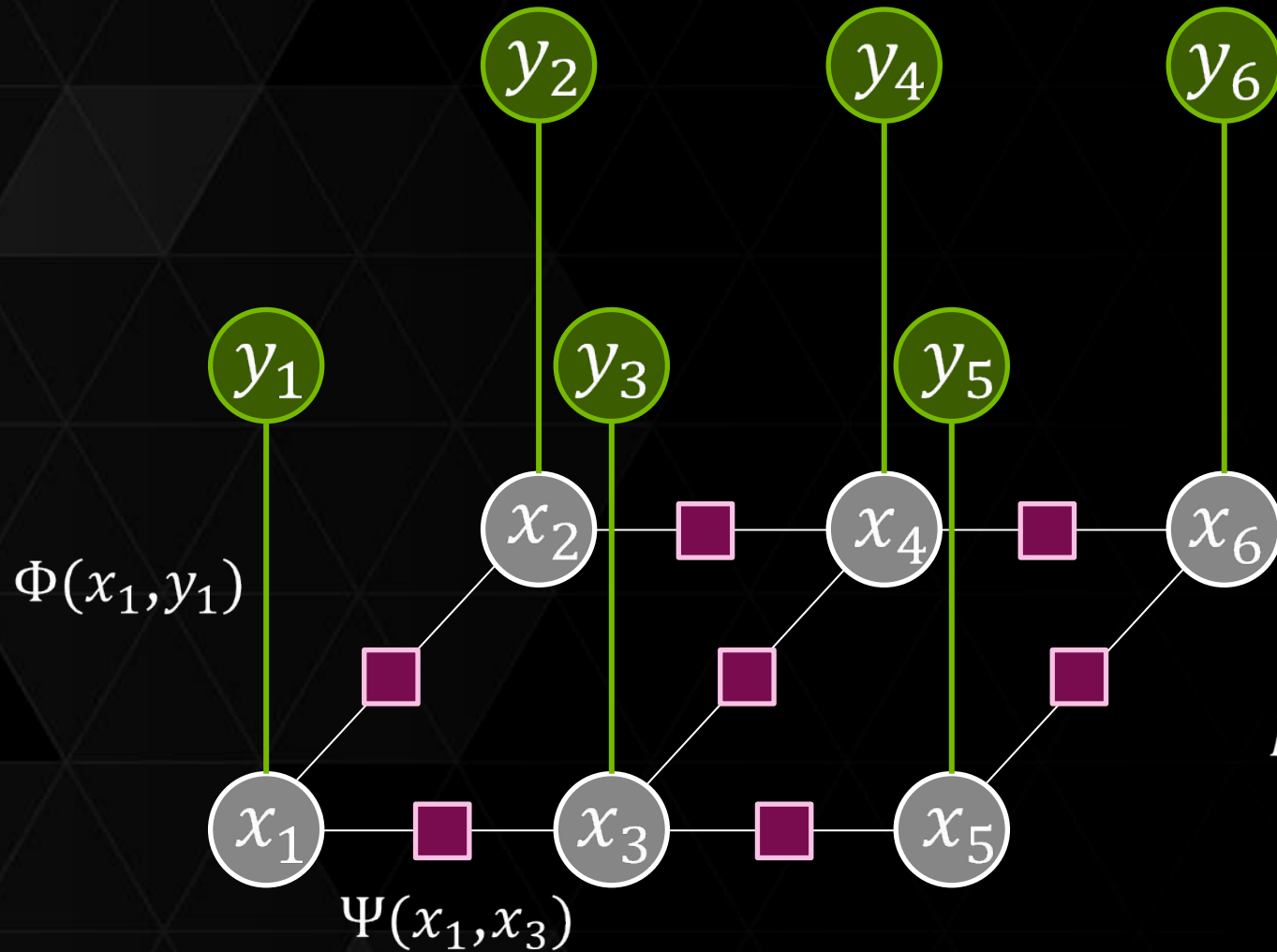
—○ Наблюдаемые значения
(унарный потенциал)

УСЛОВНЫЕ СЛУЧАЙНЫЕ ПОЛЯ (CRF)



- Скрытые значения (искомые метки)
- Наблюдаемые значения (унарный потенциал)
- Связи между элементами (парный потенциал)

УСЛОВНЫЕ СЛУЧАЙНЫЕ ПОЛЯ (CRF)



○ Скрытые значения
(искомые метки)

—○ Наблюдаемые значения
(унарный потенциал)

—■— Связи между элементами
(парный потенциал)

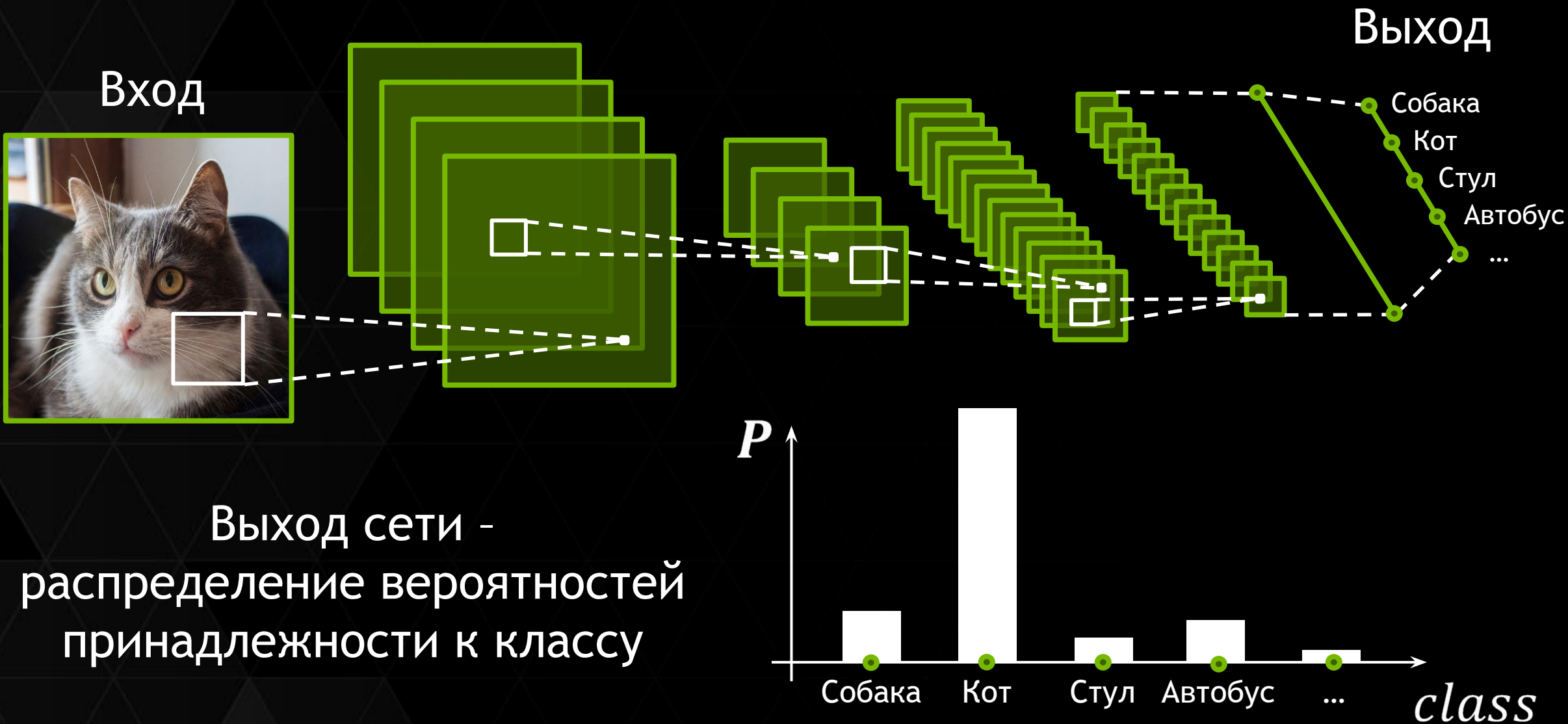
$$E(x, y) = \sum_i \Phi(x_i, y_i) + \sum_{i,j} \Psi(x_i, x_j)$$

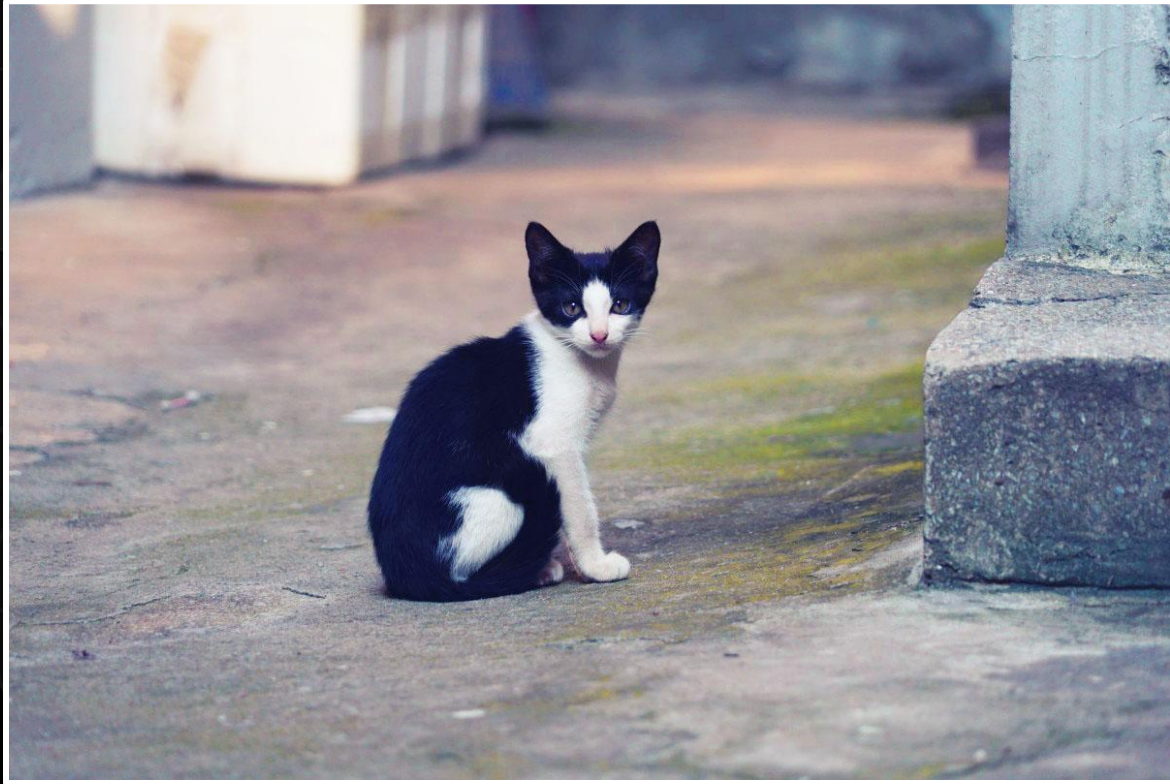
$$x^* = \operatorname{argmax}_x E(x, y)$$

The background is a dark, almost black, field filled with a complex network of thin, glowing green lines. These lines intersect at various points, creating a web-like structure. At many of these intersection points, there are small, bright green circular dots. Some of these dots are slightly larger and more prominent than others. Additionally, there are a few larger, fainter blue circular shapes scattered across the background, giving it a sense of depth and a futuristic or technological feel.

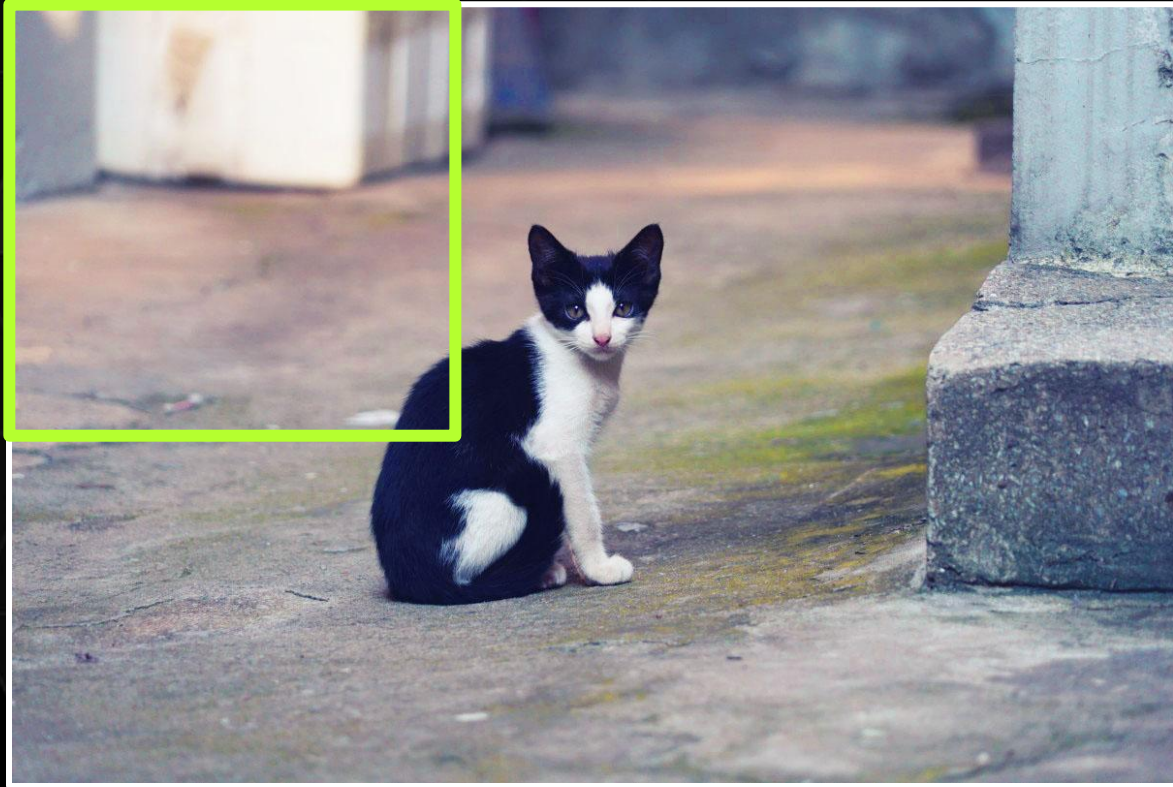
МЕТОДЫ НА ОСНОВЕ ГЛУБОКОГО ОБУЧЕНИЯ

КЛАССИФИКАЦИЯ ИЗОБРАЖЕНИЙ

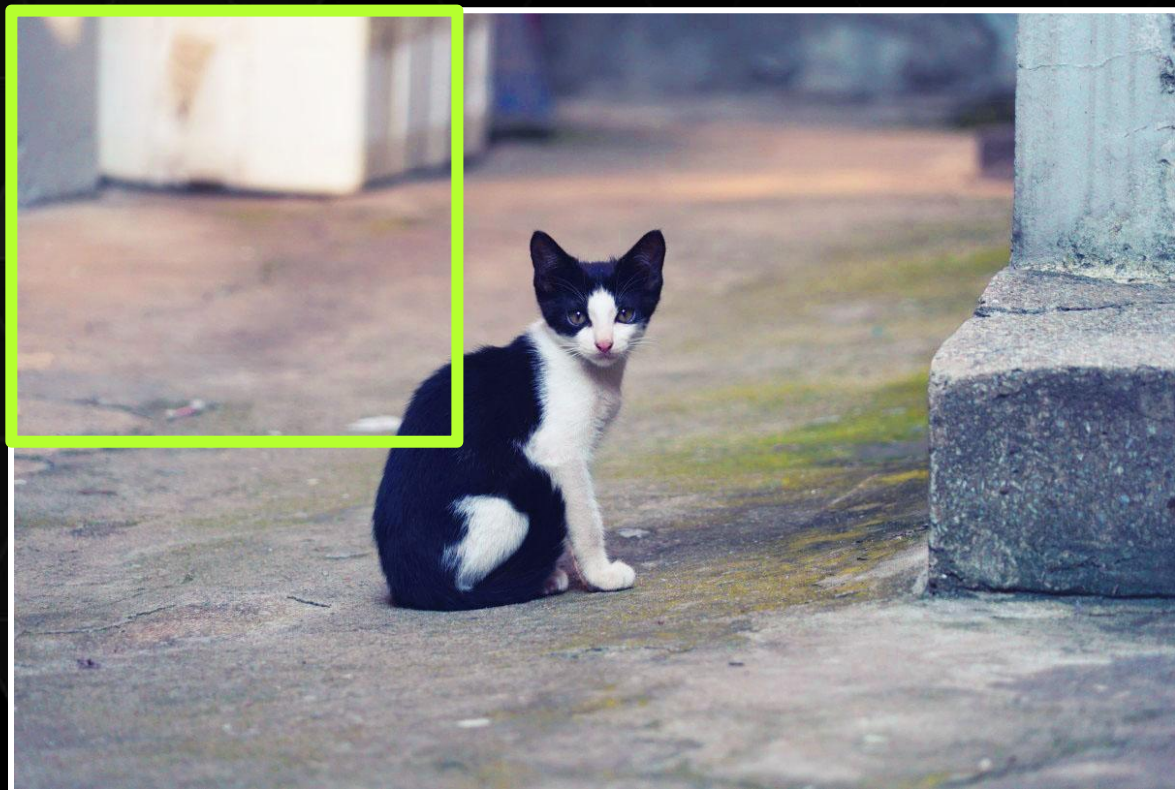




СКОЛЬЗЯЩЕЕ ОКНО



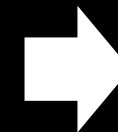
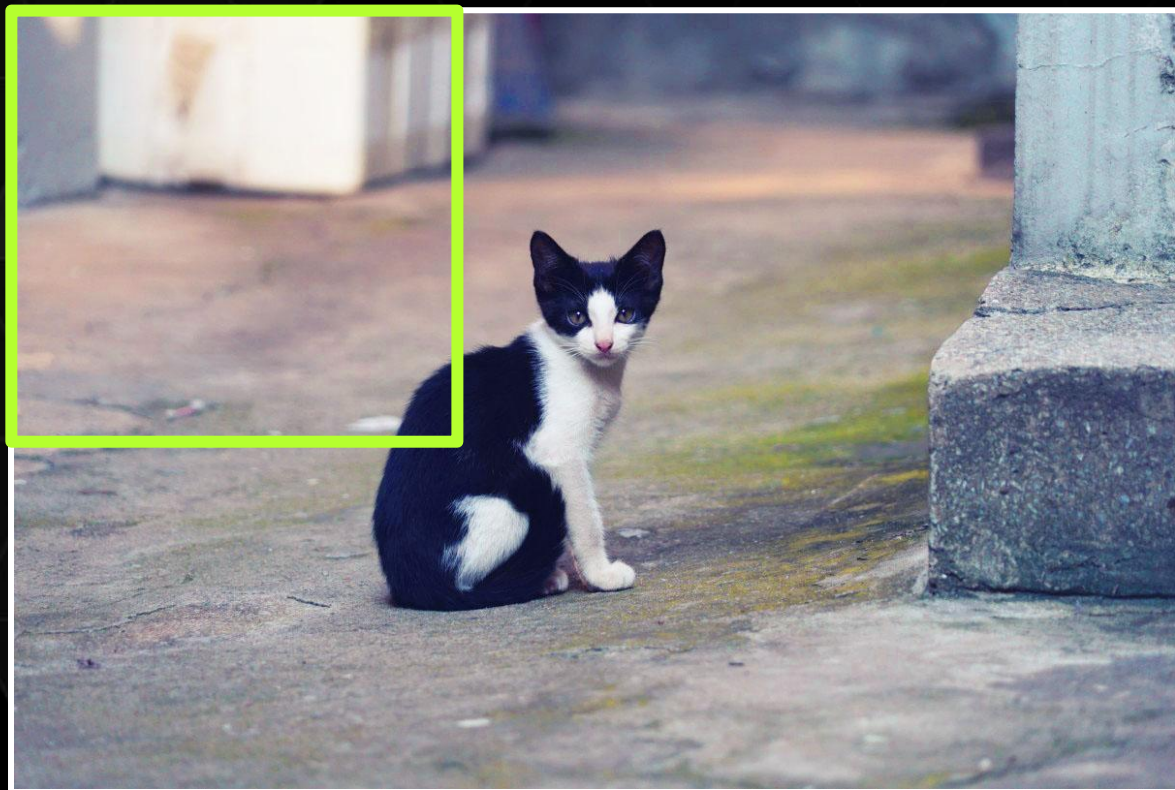
СКОЛЬЗЯЩЕЕ ОКНО



Классификатор

- Собака
- Кот
- Стул
- Автобус
- ...

СКОЛЬЗЯЩЕЕ ОКНО

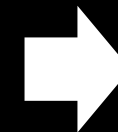
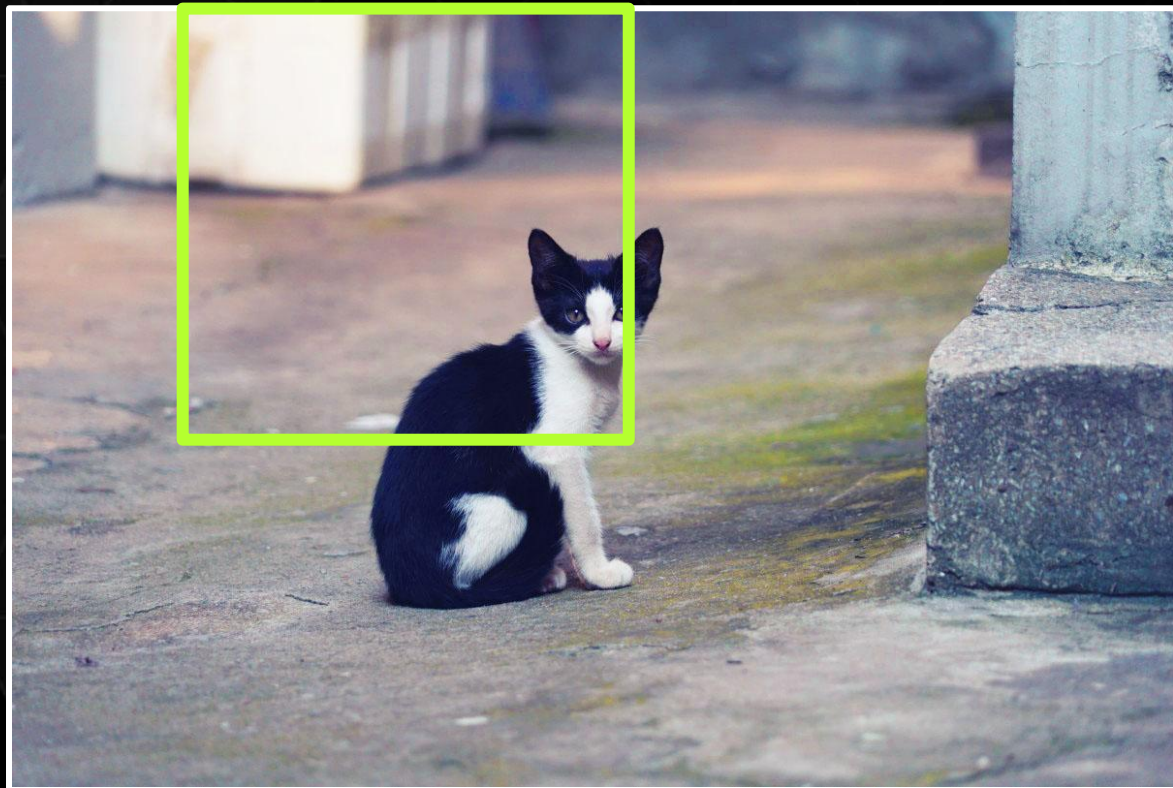


Классификатор

Собака
Кот
Стул
Автобус
...



СКОЛЬЗЯЩЕЕ ОКНО

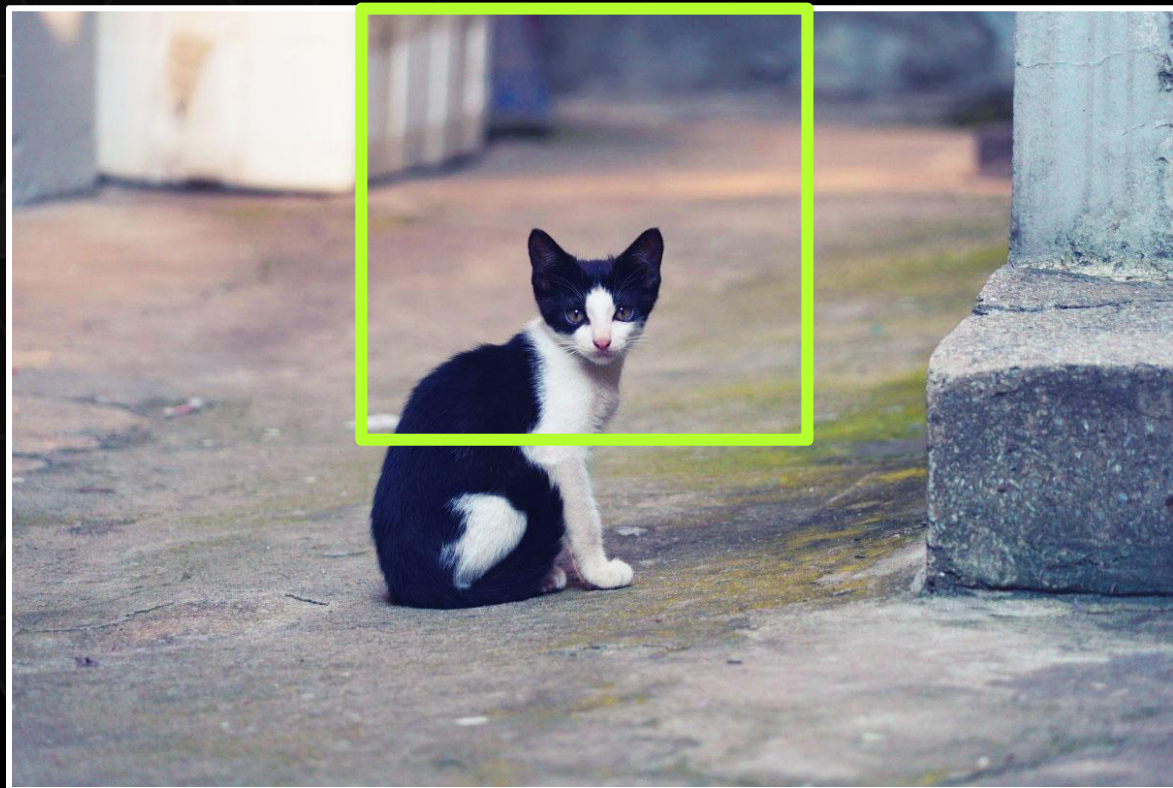


Классификатор

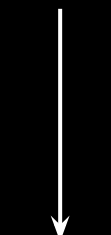
Собака
Кот
Стул
Автобус
...



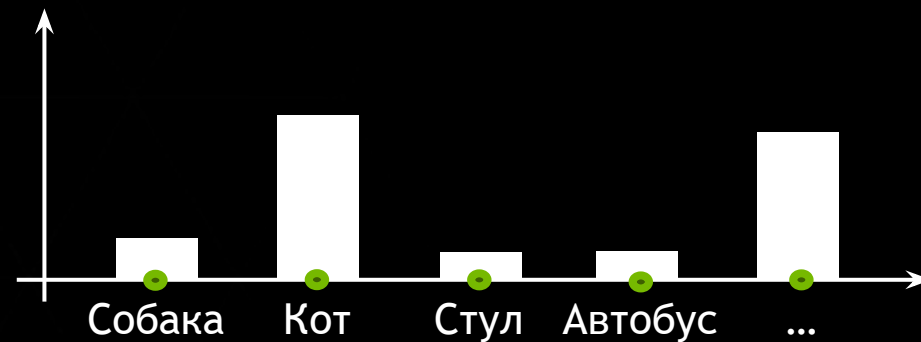
СКОЛЬЗЯЩЕЕ ОКНО



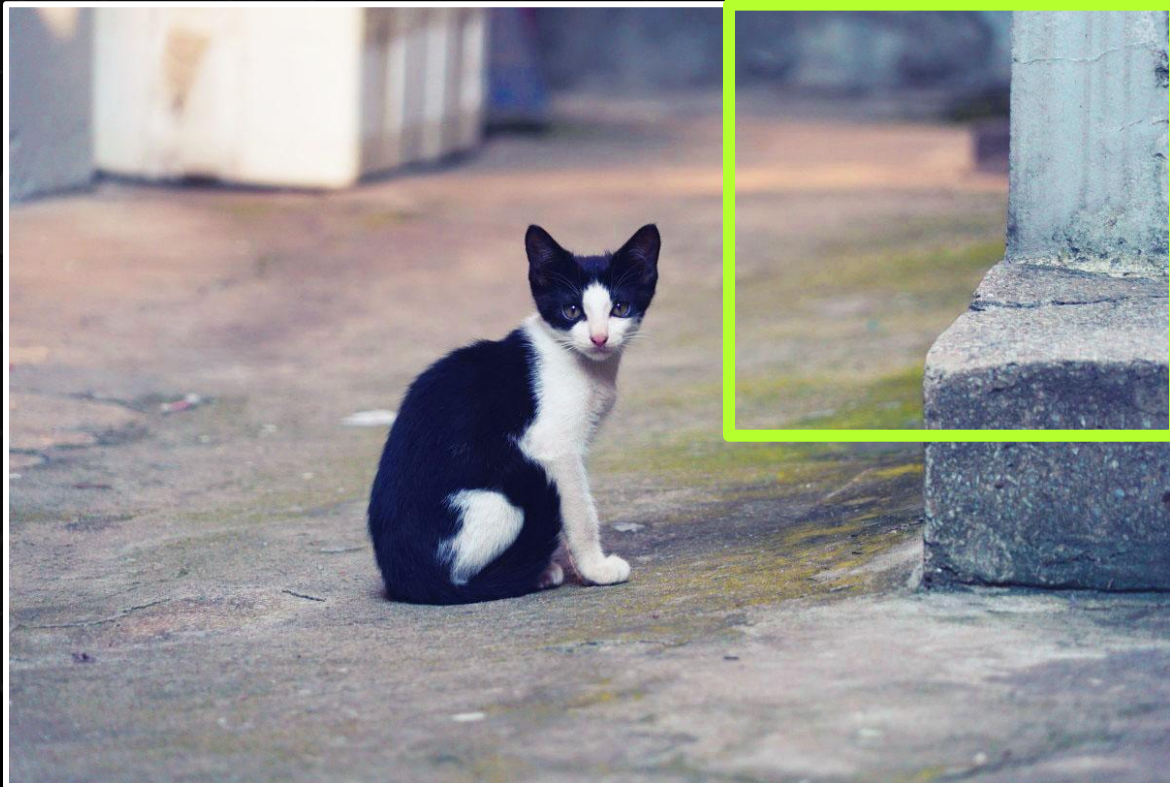
Классификатор



Собака
Кот
Стул
Автобус
...



СКОЛЬЗЯЩЕЕ ОКНО

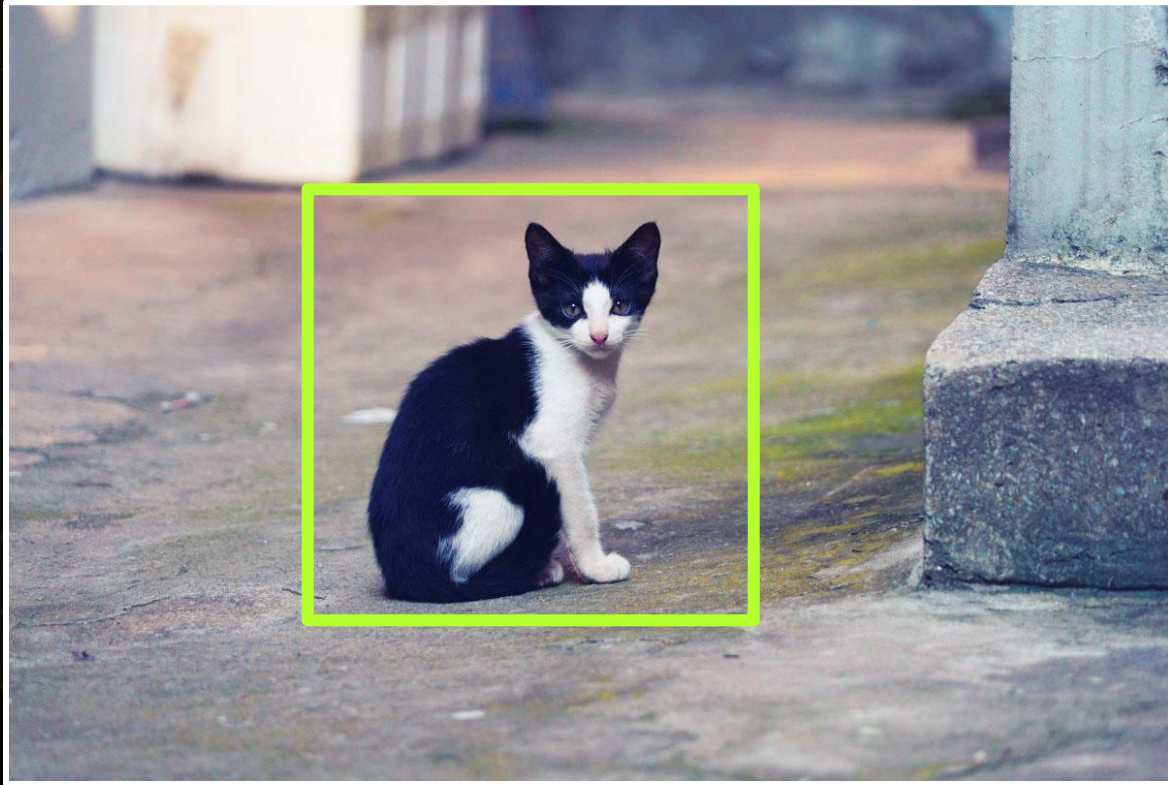


Классификатор

Собака
Кот
Стул
Автобус
...

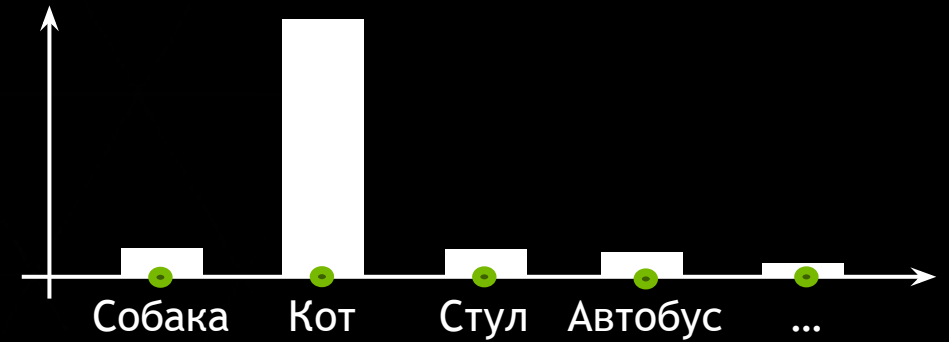


СКОЛЬЗЯЩЕЕ ОКНО

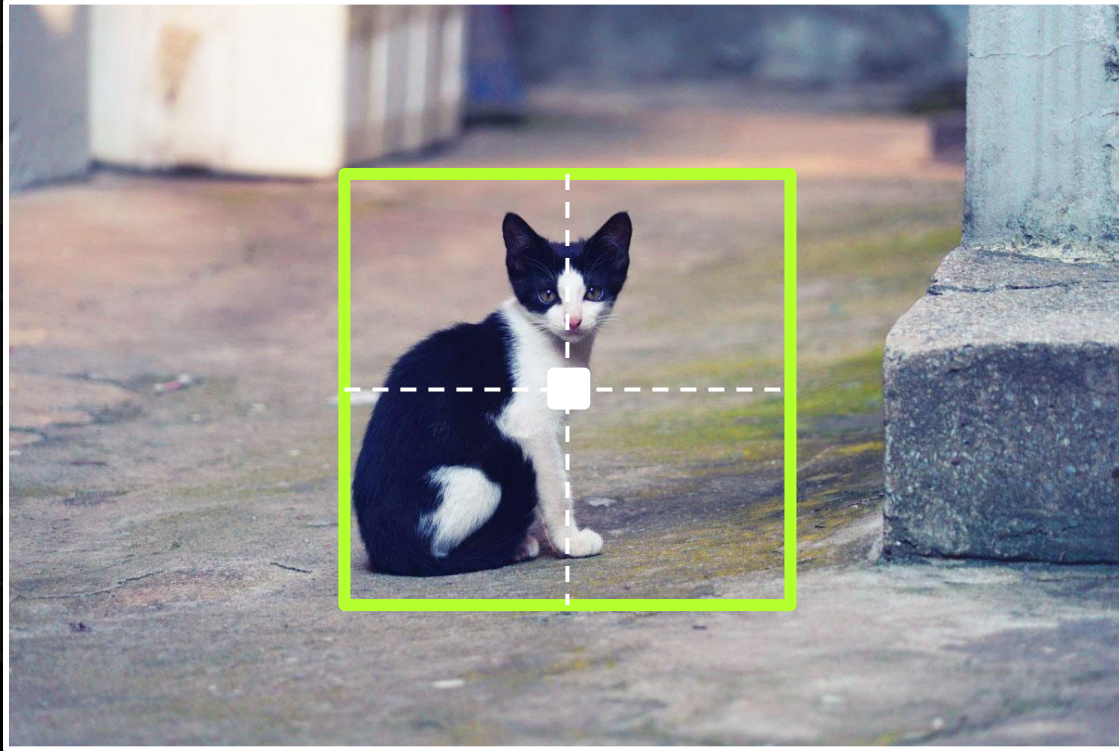


Классификатор

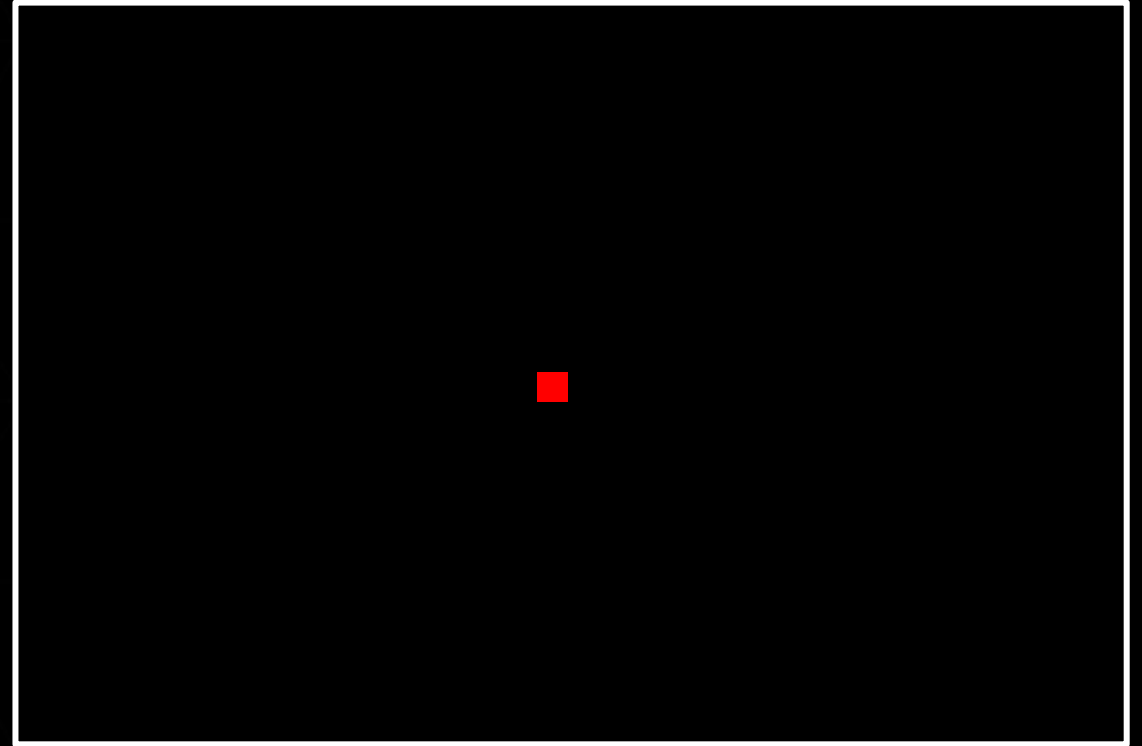
Собака
Кот
Стул
Автобус
...



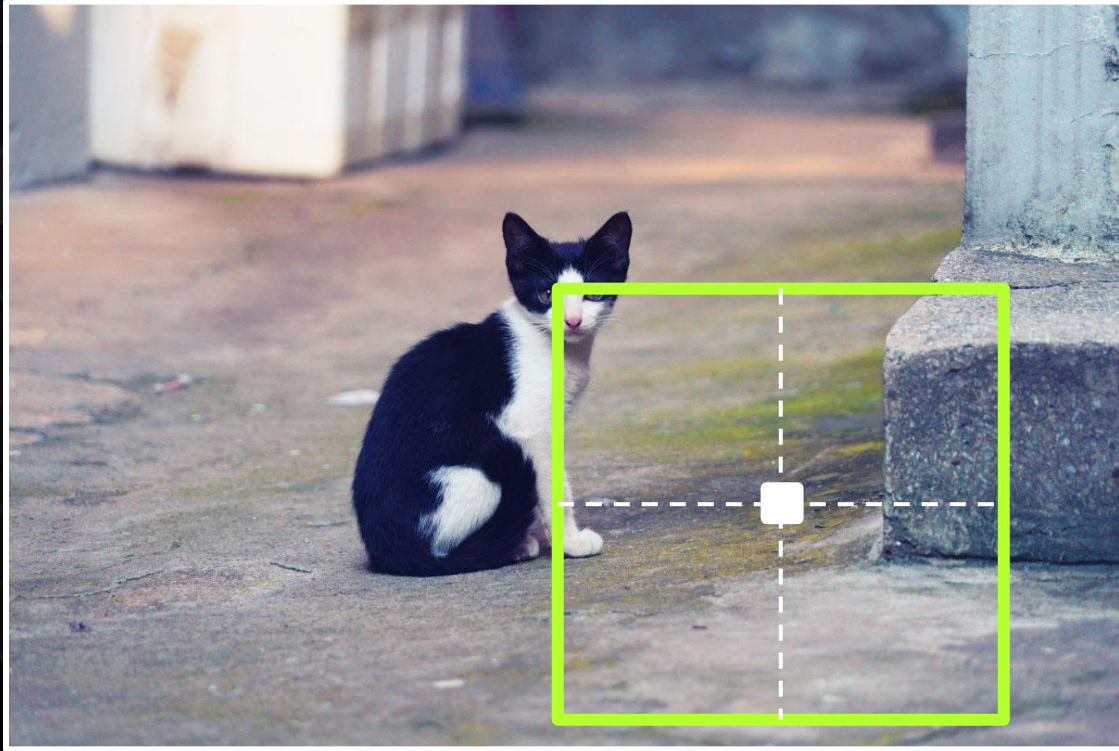
ТЕПЛОВАЯ КАРТА



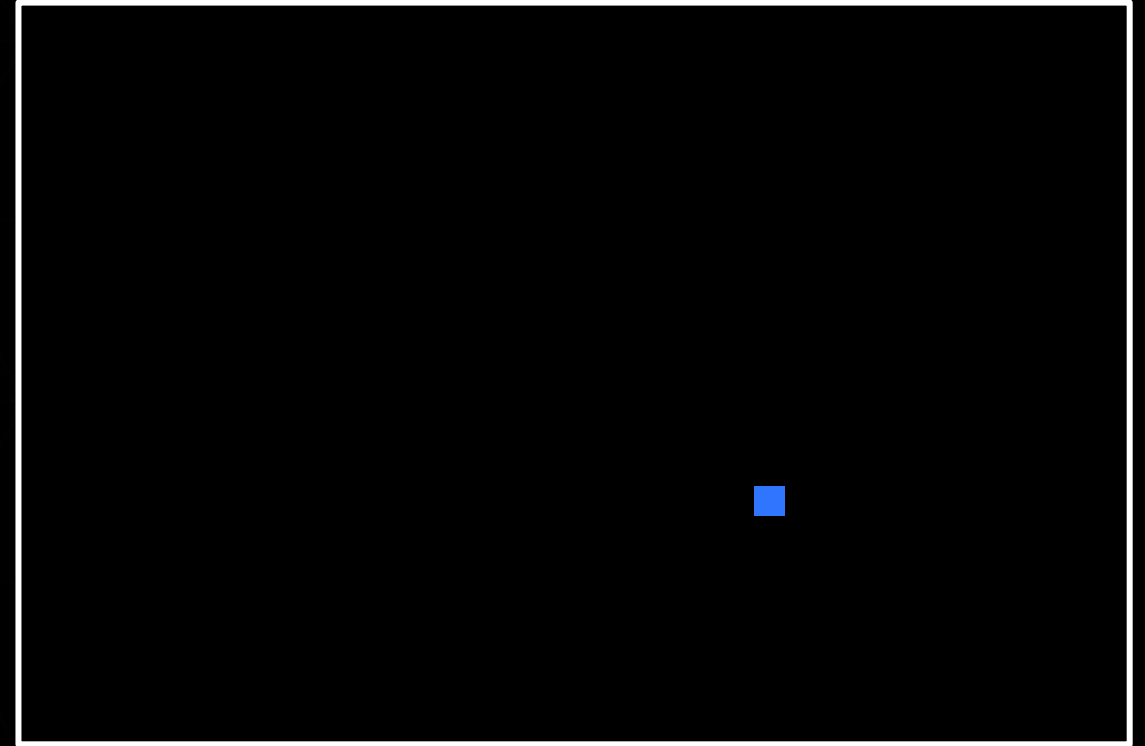
Класс: **Кот**



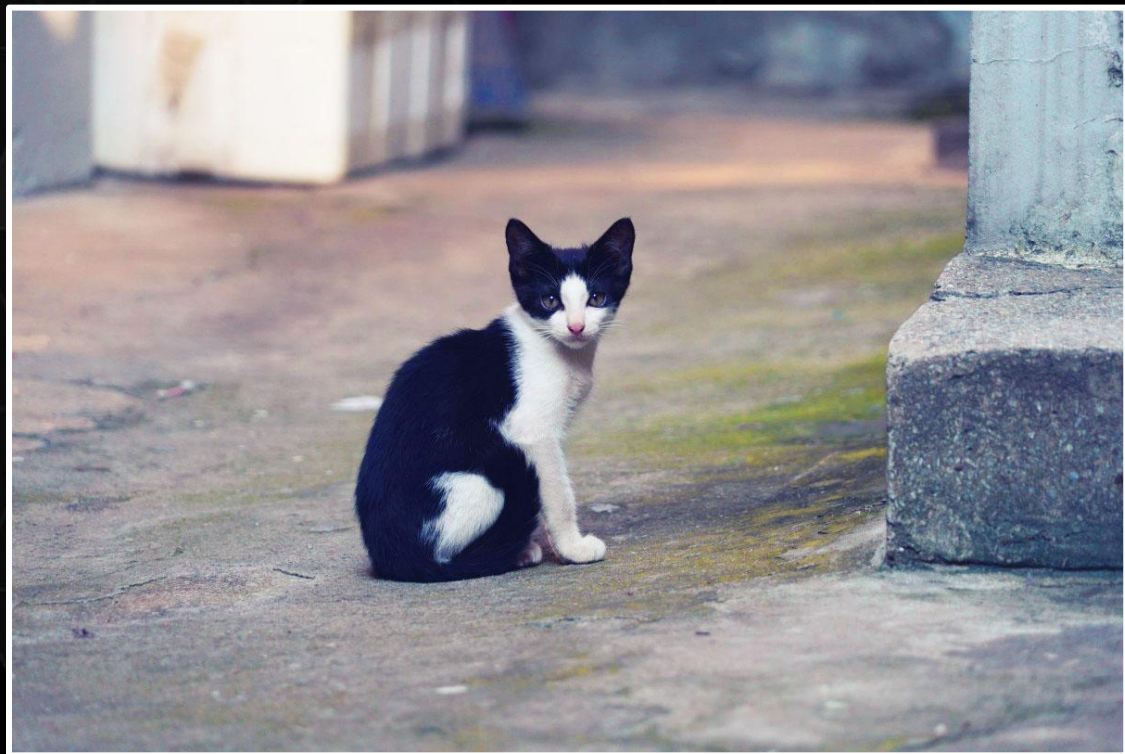
ТЕПЛОВАЯ КАРТА



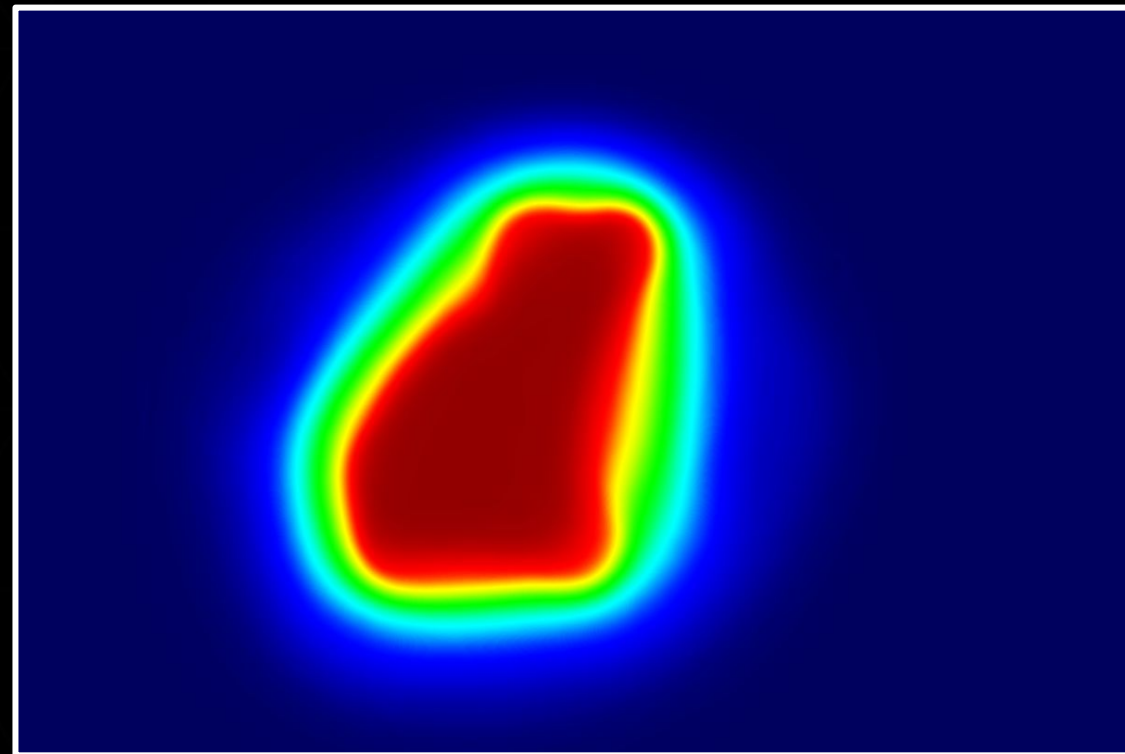
Класс: **Кот**



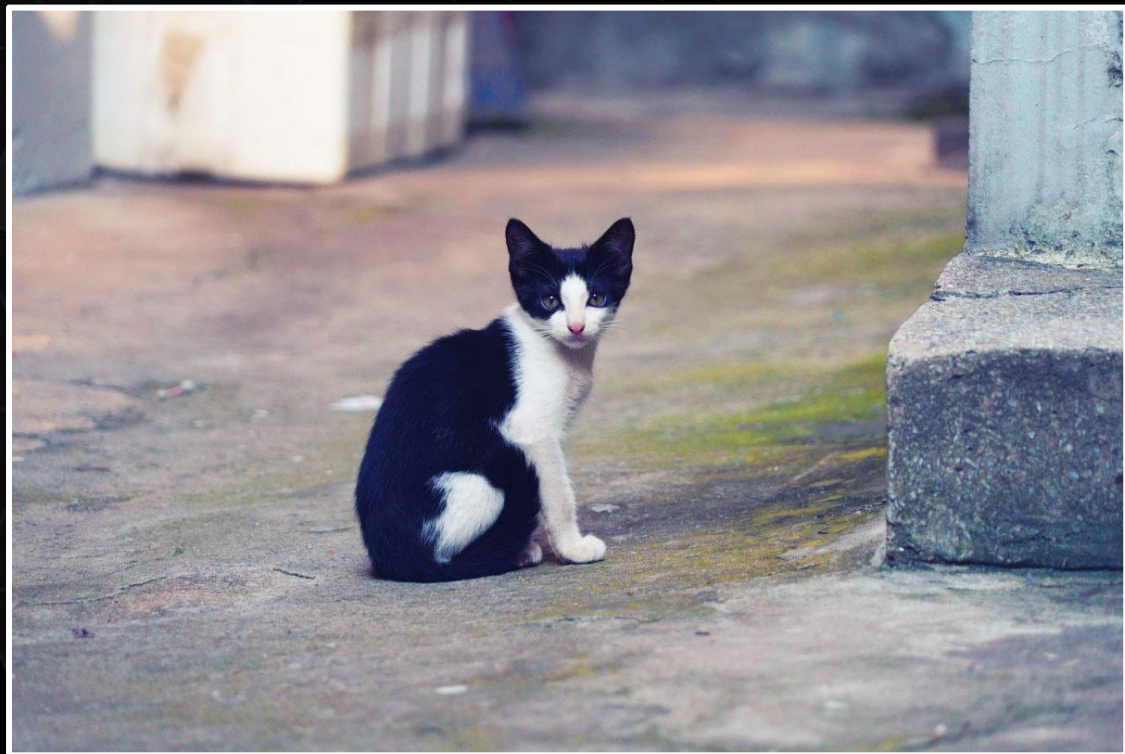
ТЕПЛОВАЯ КАРТА



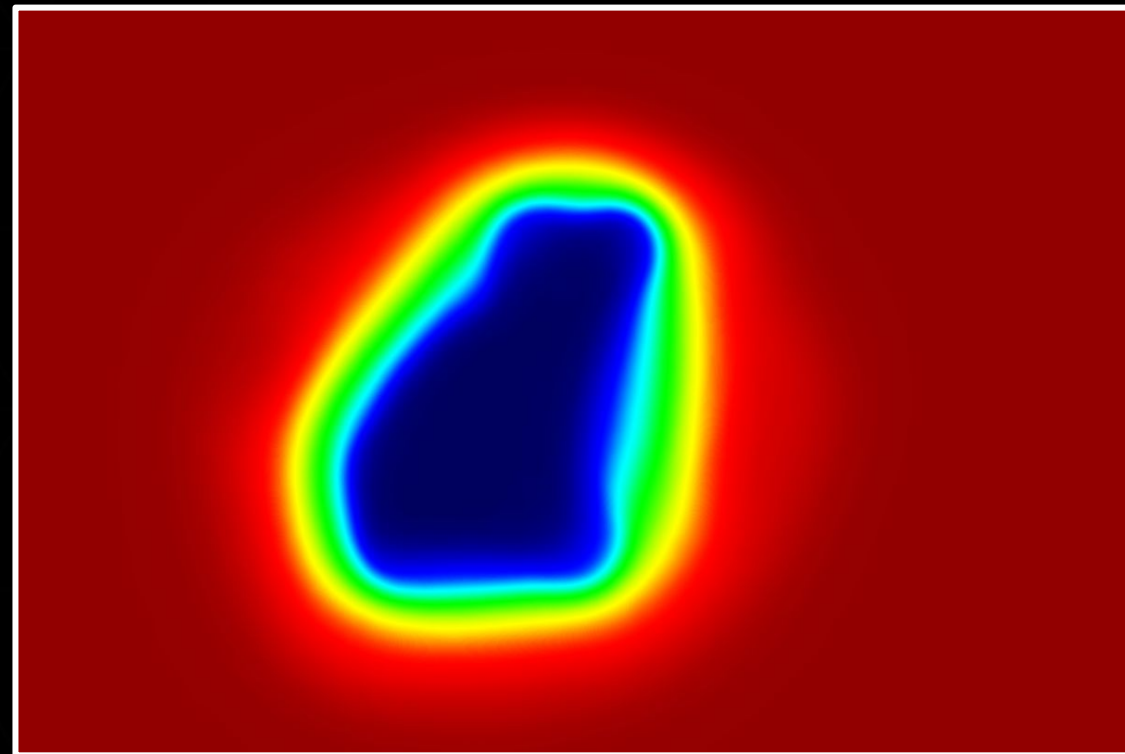
Класс: **Кот**



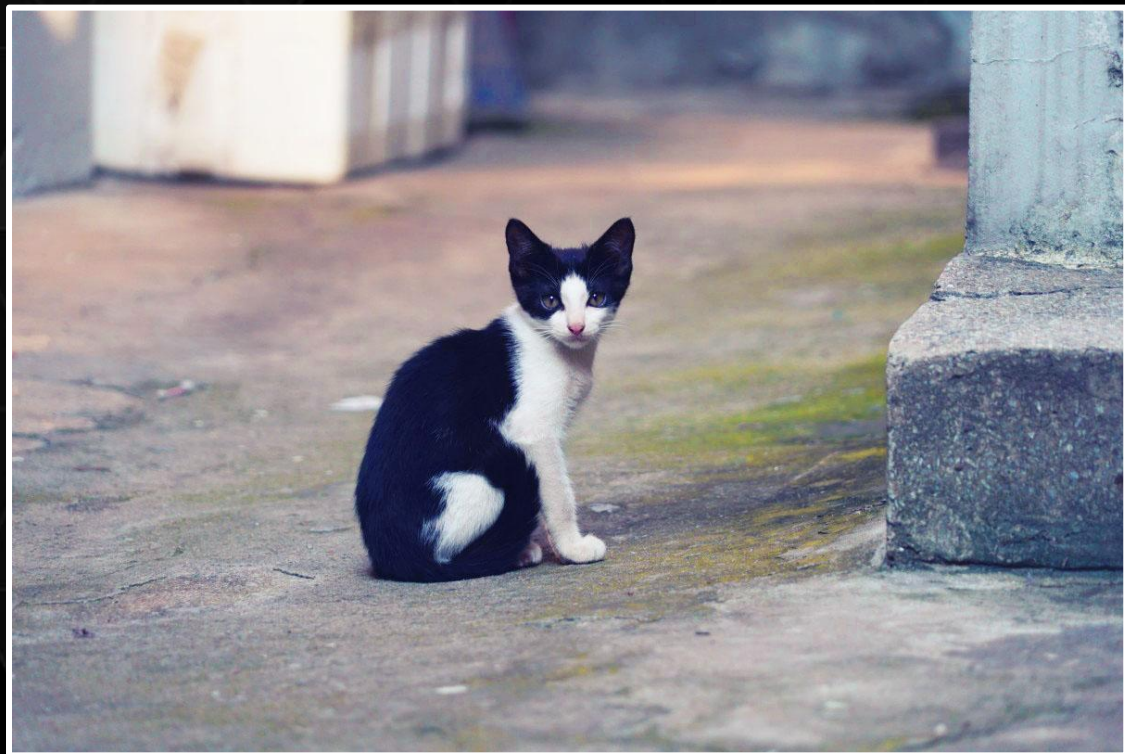
ТЕПЛОВАЯ КАРТА



Класс: Фон



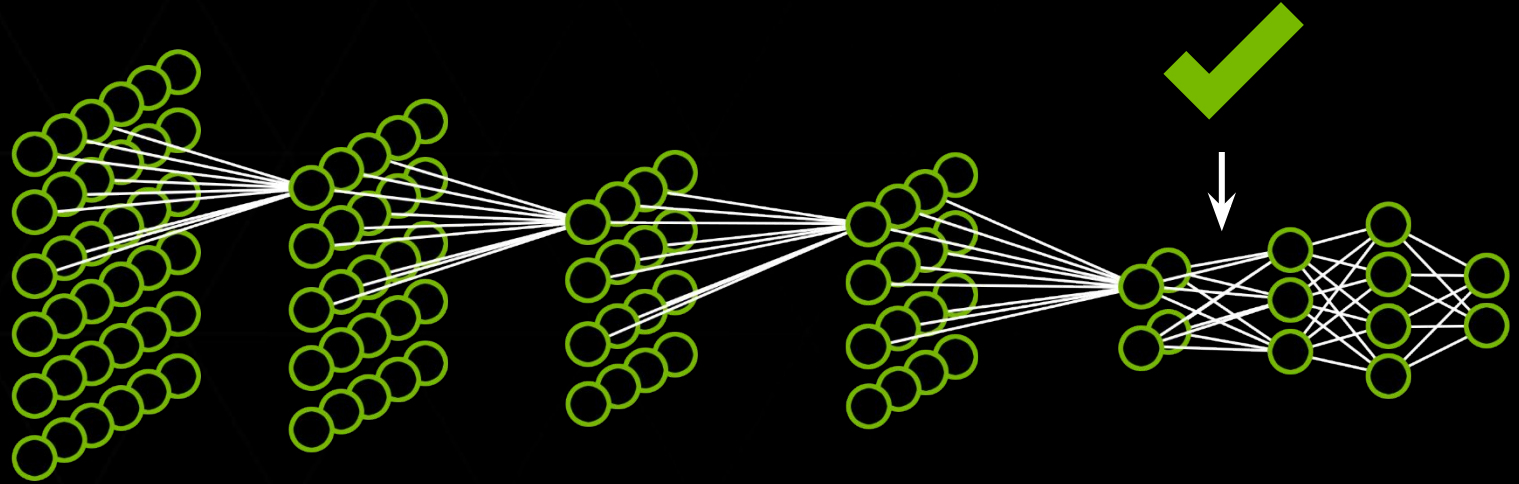
МАСКА КЛАССА



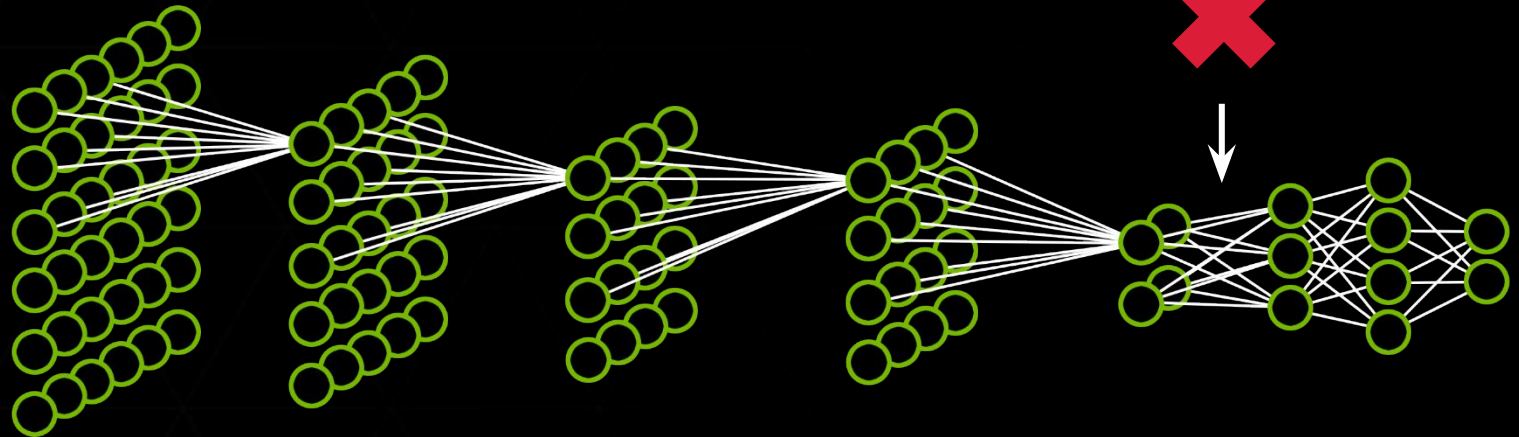
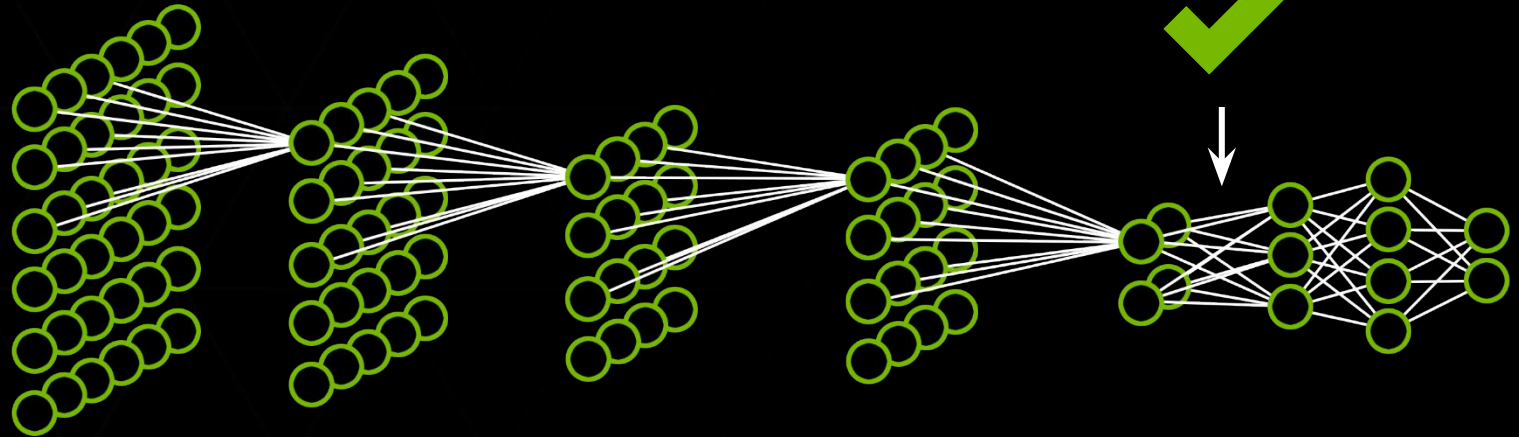
Класс: **Кот**



ФС СЛОИ: ФИКСИРОВАННЫЙ РАЗМЕР

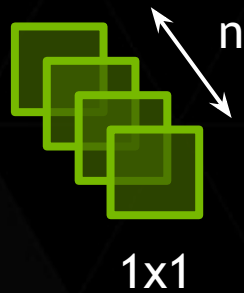


ФС СЛОИ: ФИКСИРОВАННЫЙ РАЗМЕР

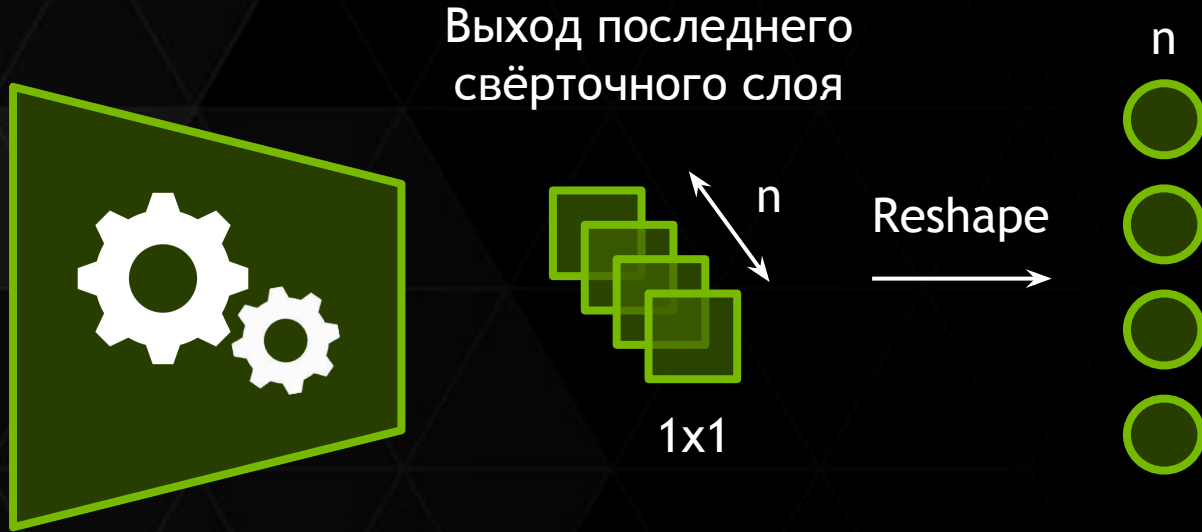


СВЁРТКА 1 X 1

Выход последнего
свёрточного слоя



СВЁРТКА 1 X 1



СВЁРТКА 1 X 1



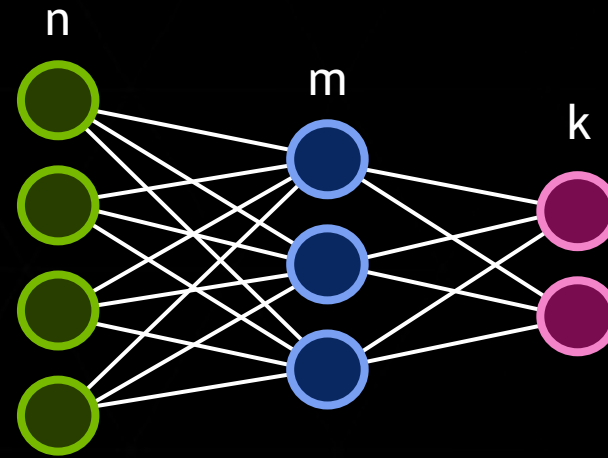
СВЁРТКА 1 X 1



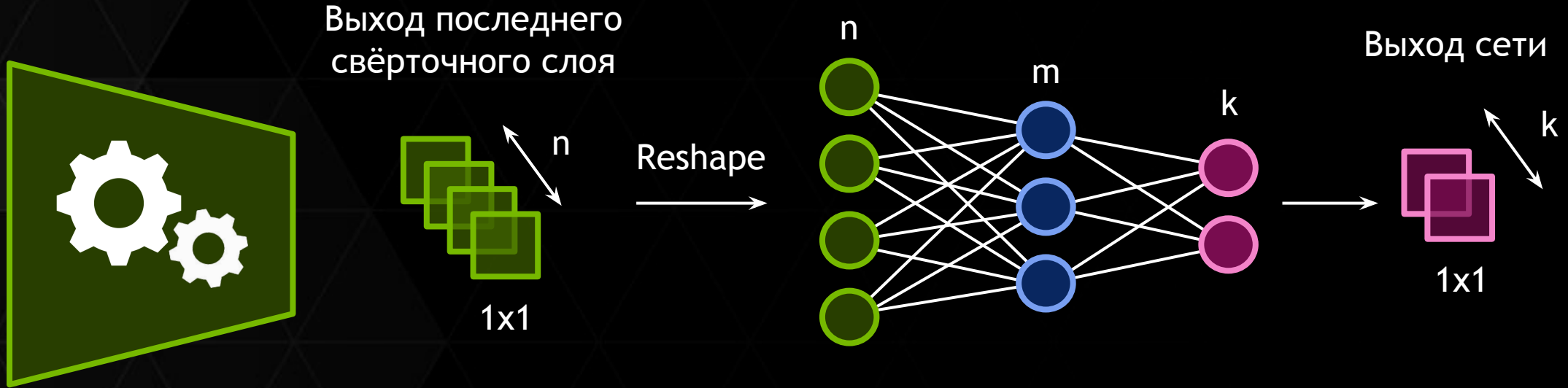
Выход последнего
свёрточного слоя



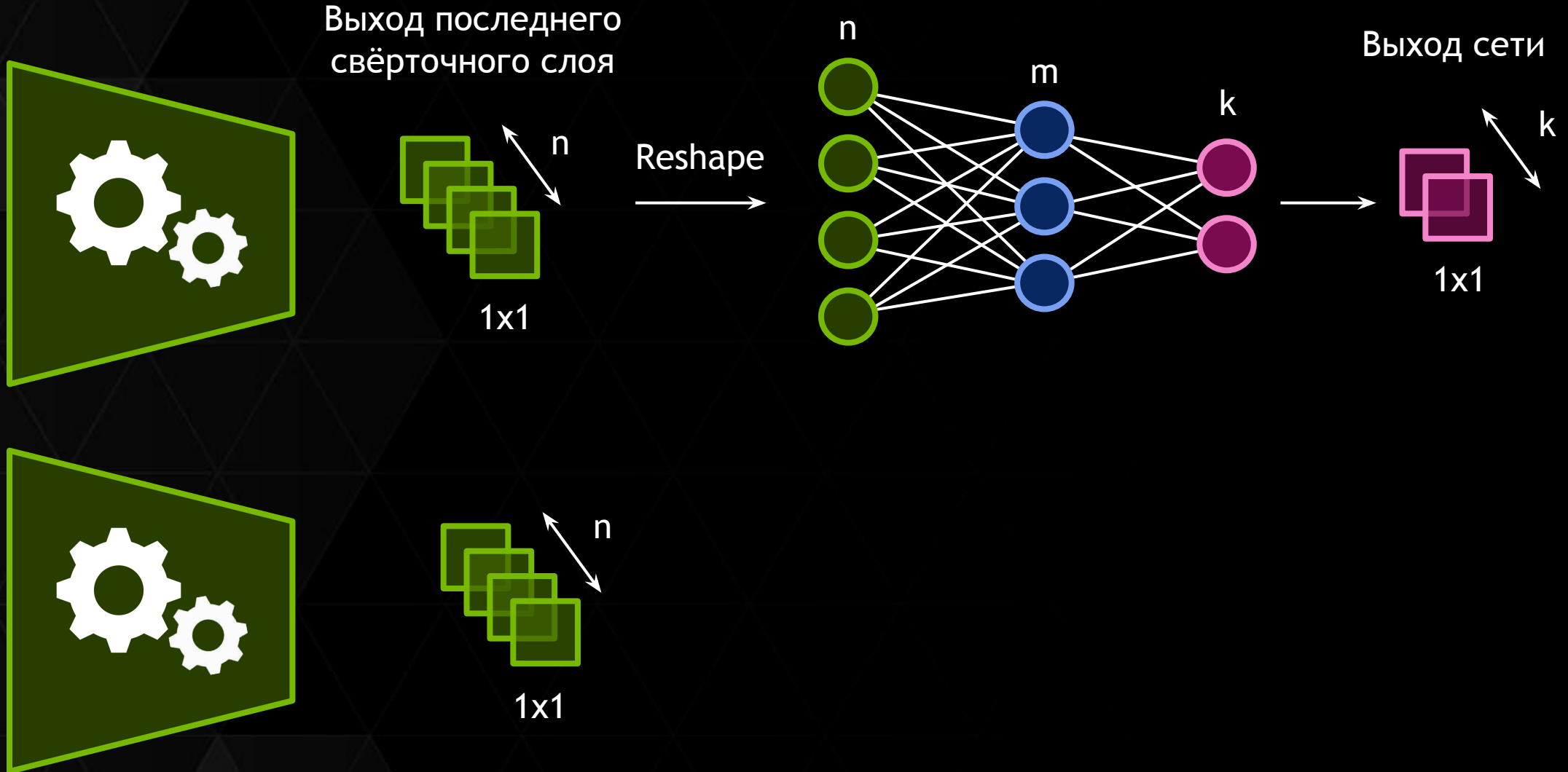
Reshape



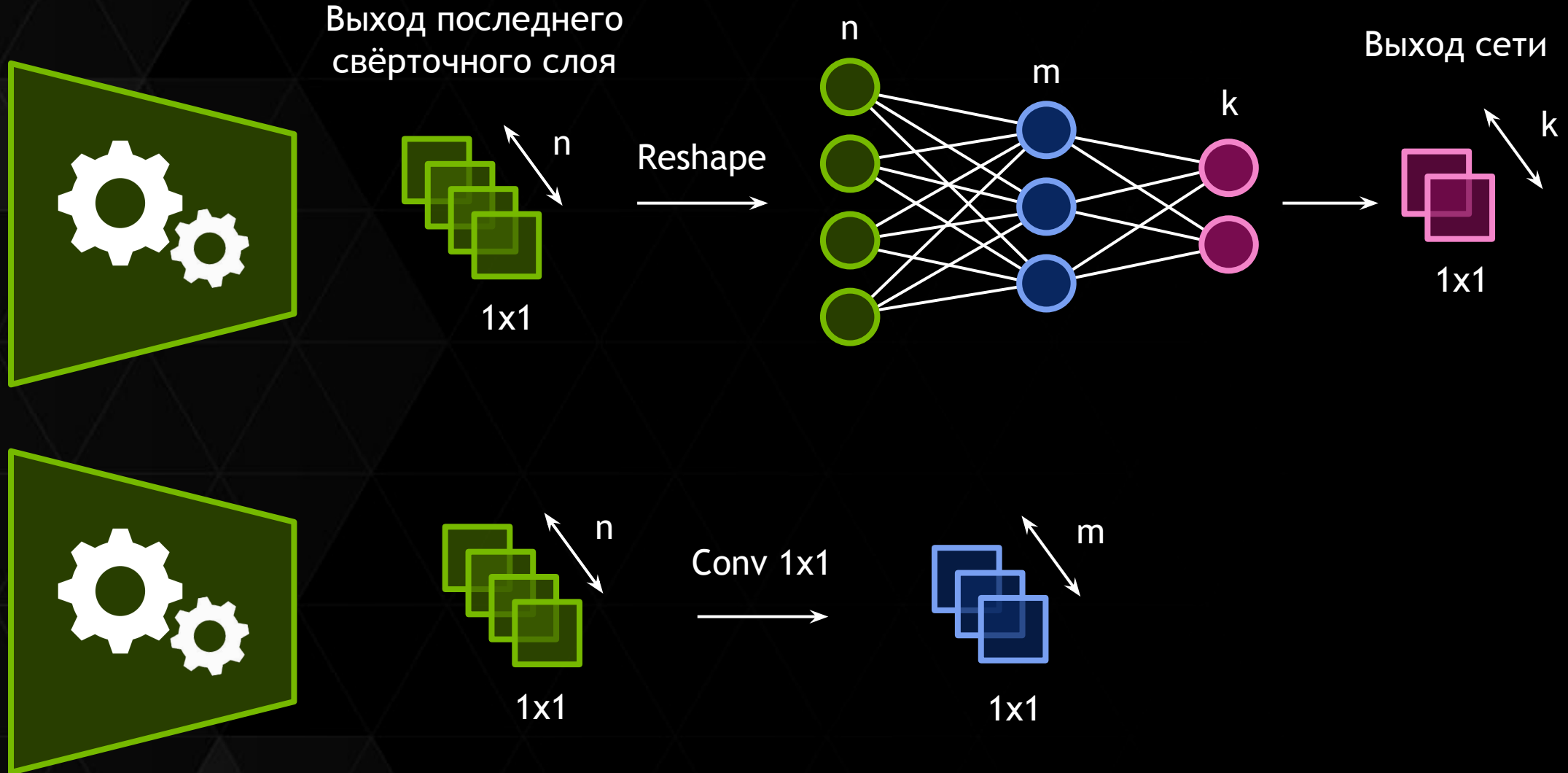
СВЁРТКА 1 X 1



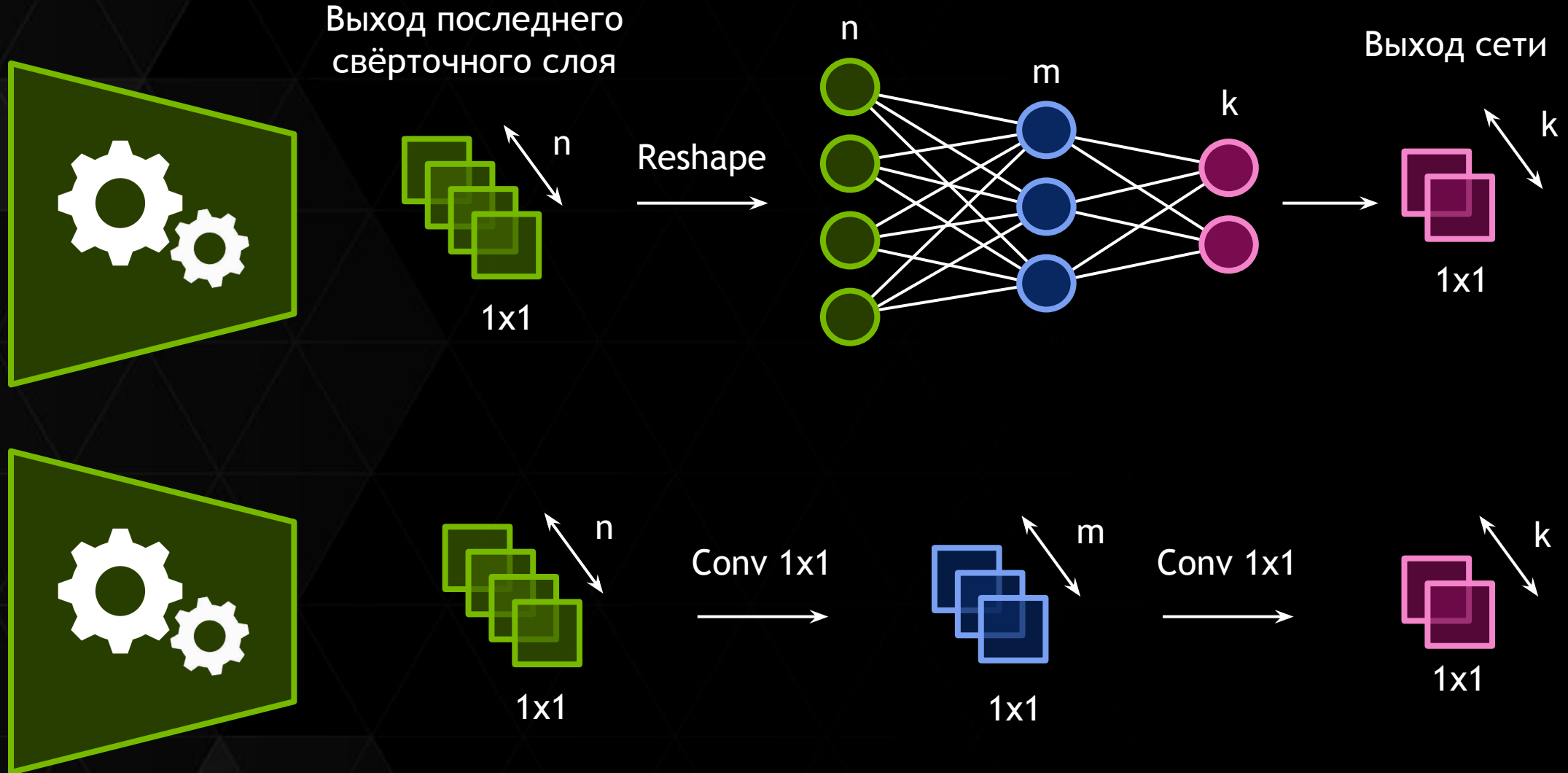
СВЁРТКА 1 X 1



СВЁРТКА 1 X 1

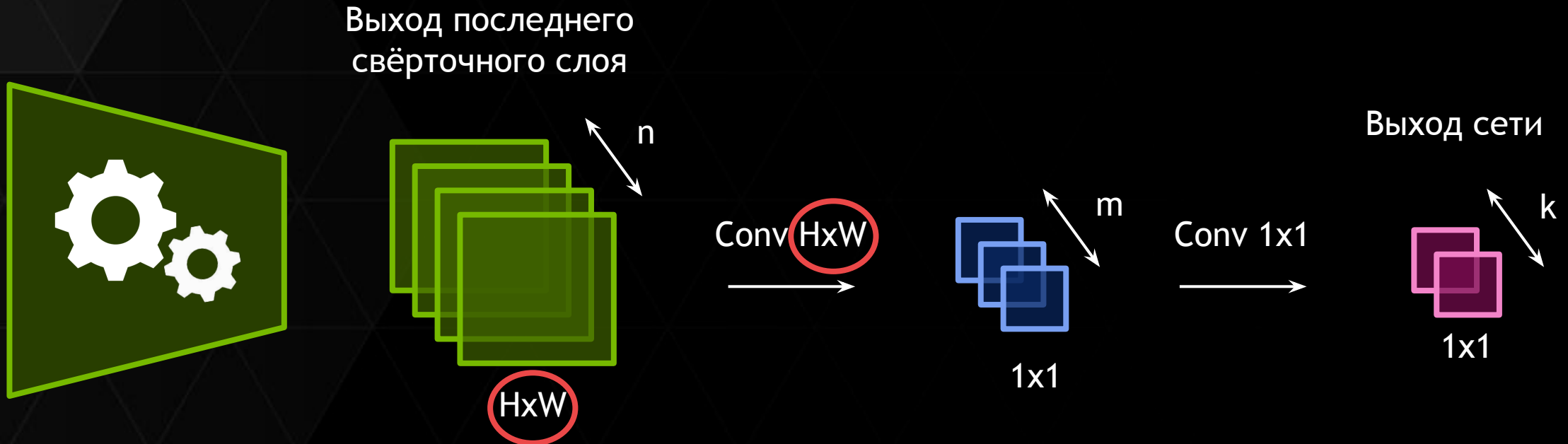


СВЁРТКА 1 X 1

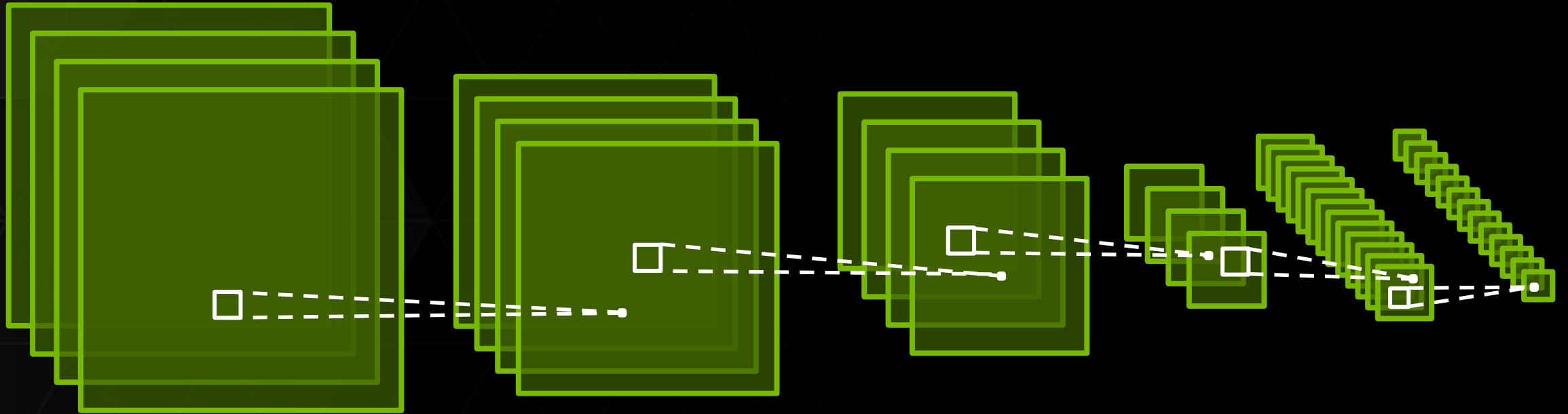


СВЁРТКА 1 X 1

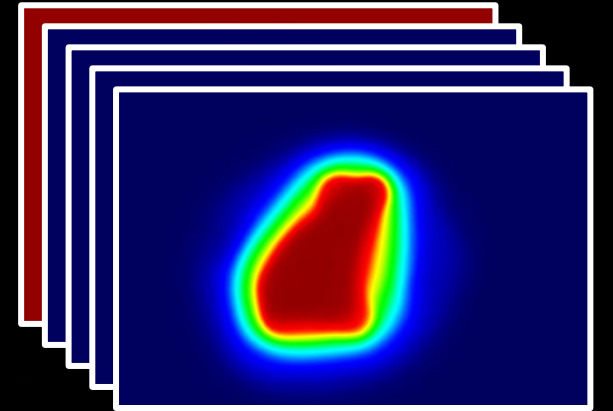
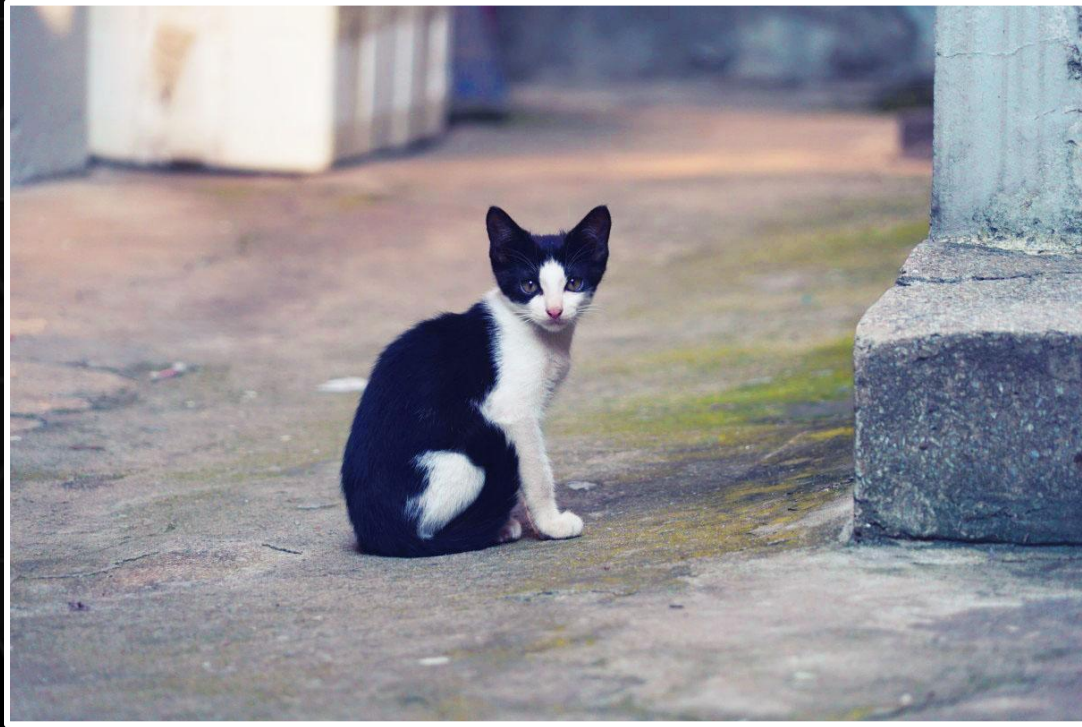
В случае, если выход последнего свёрточного слоя исходной сети имеет размеры $> 1 \times 1$



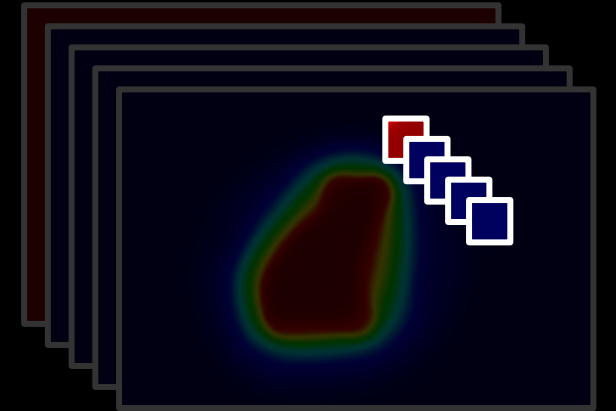
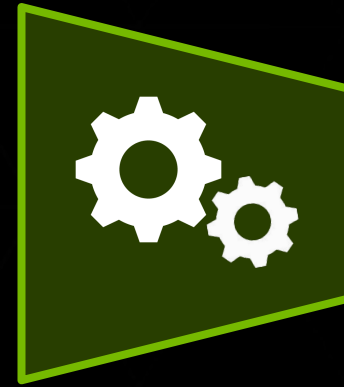
ПОЛНОСВЁРТОЧНАЯ СЕТЬ (FCN)



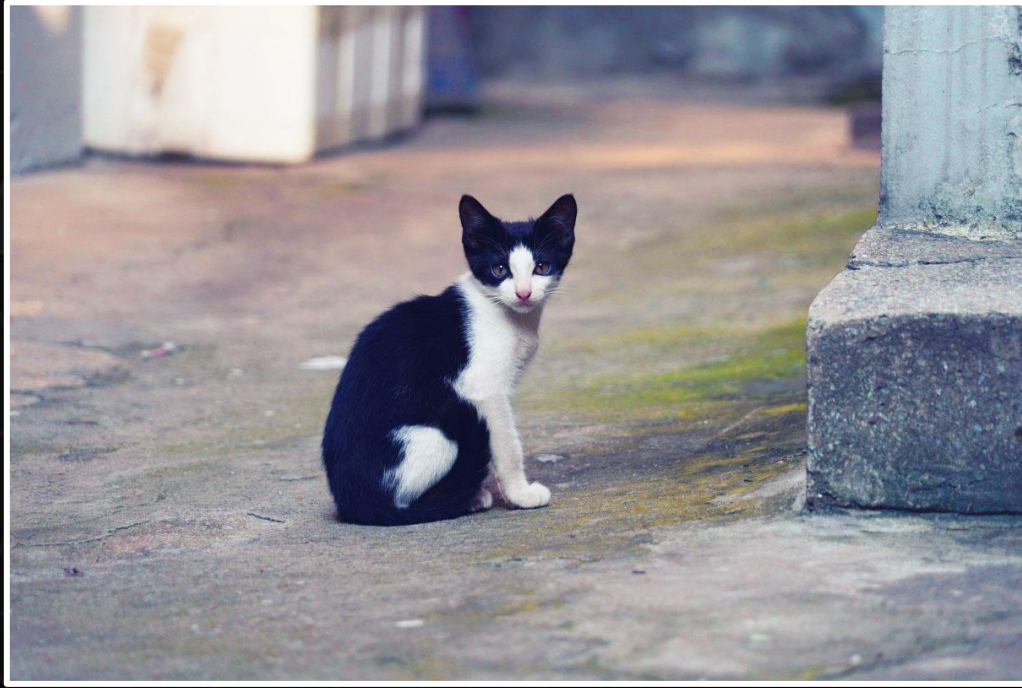
ПОЛНОСВЁРТОЧНАЯ СЕТЬ (FCN)



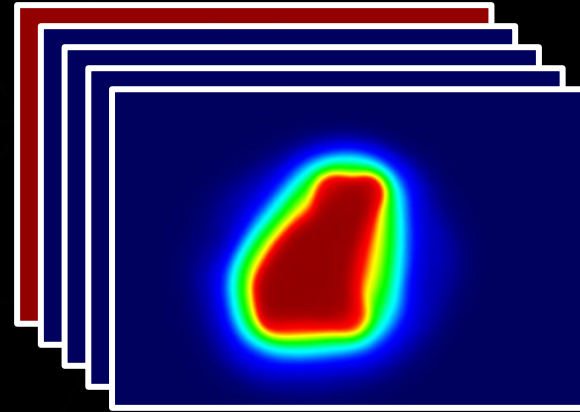
ПОЛНОСВЁРТОЧНАЯ СЕТЬ (FCN)



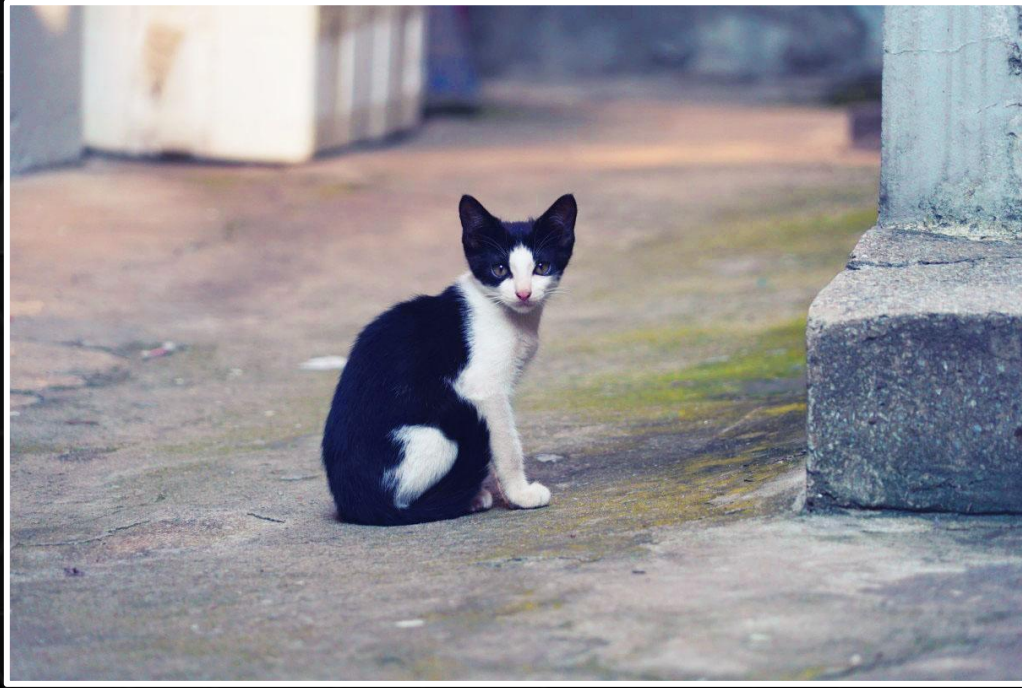
FCN + ARGMAX



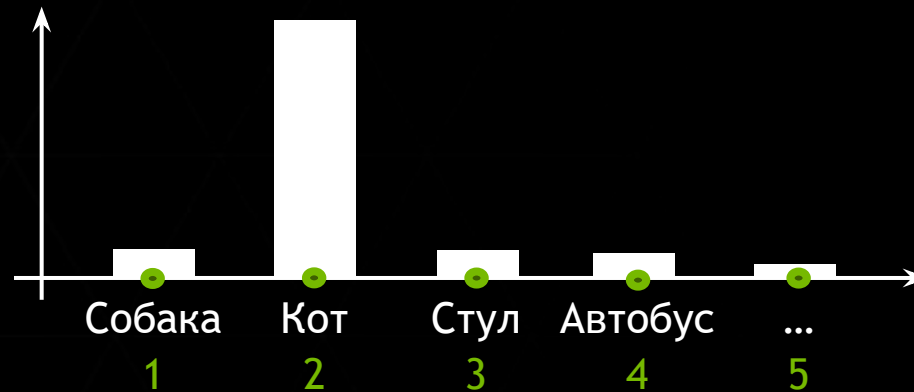
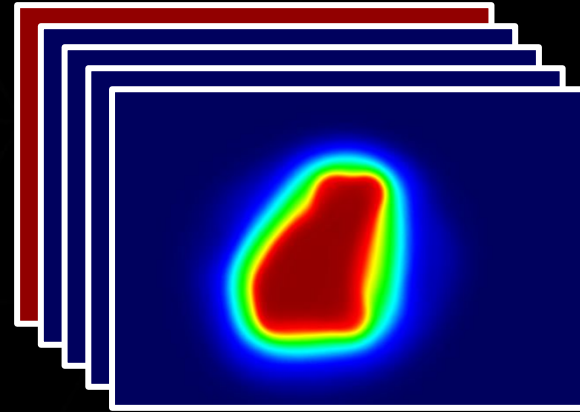
FCN



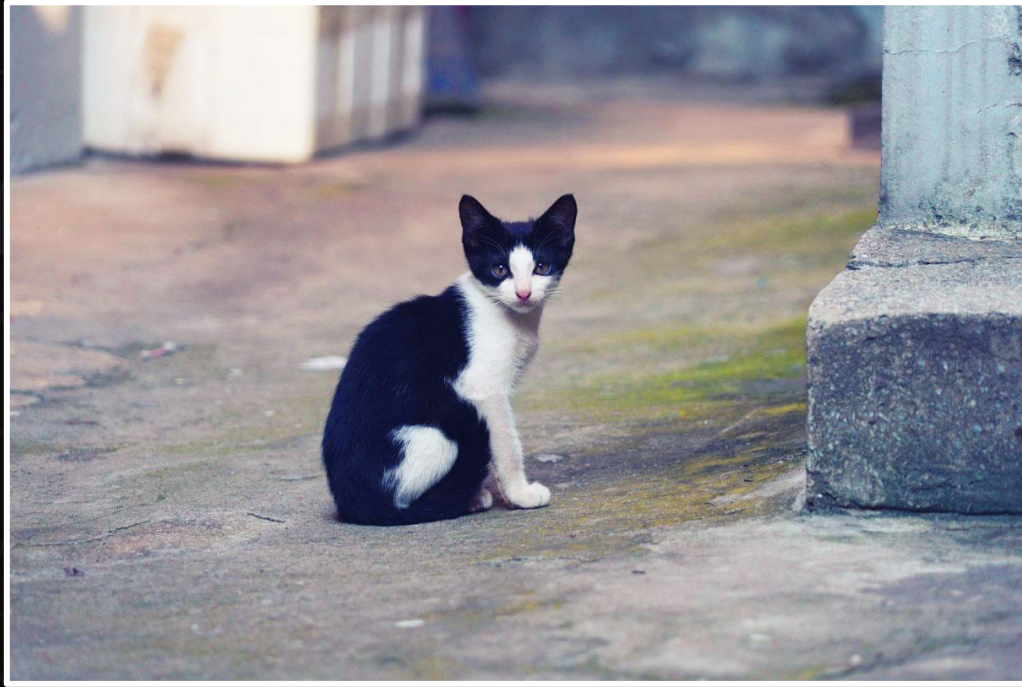
FCN + ARGMAX



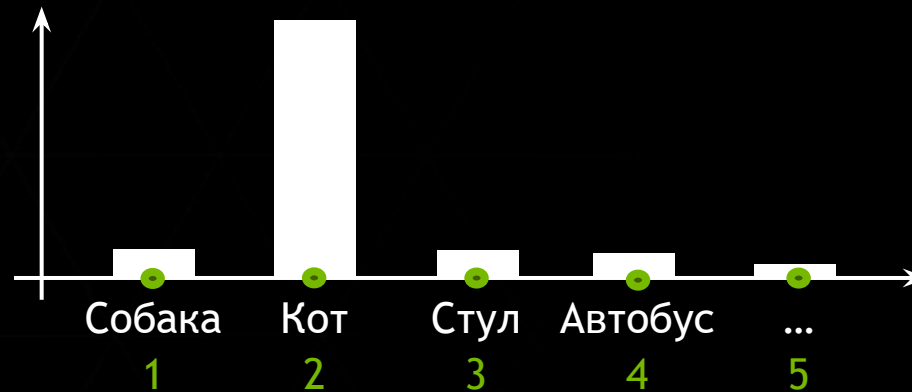
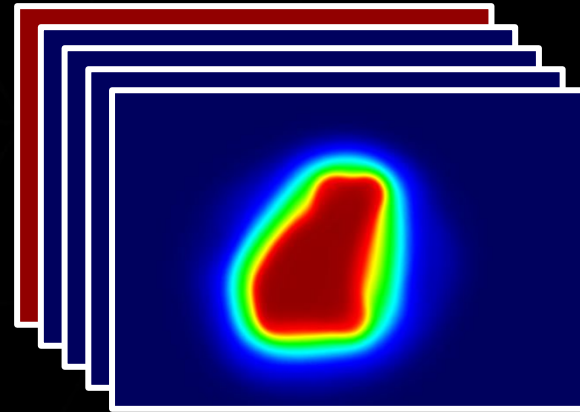
FCN



FCN + ARGMAX



FCN

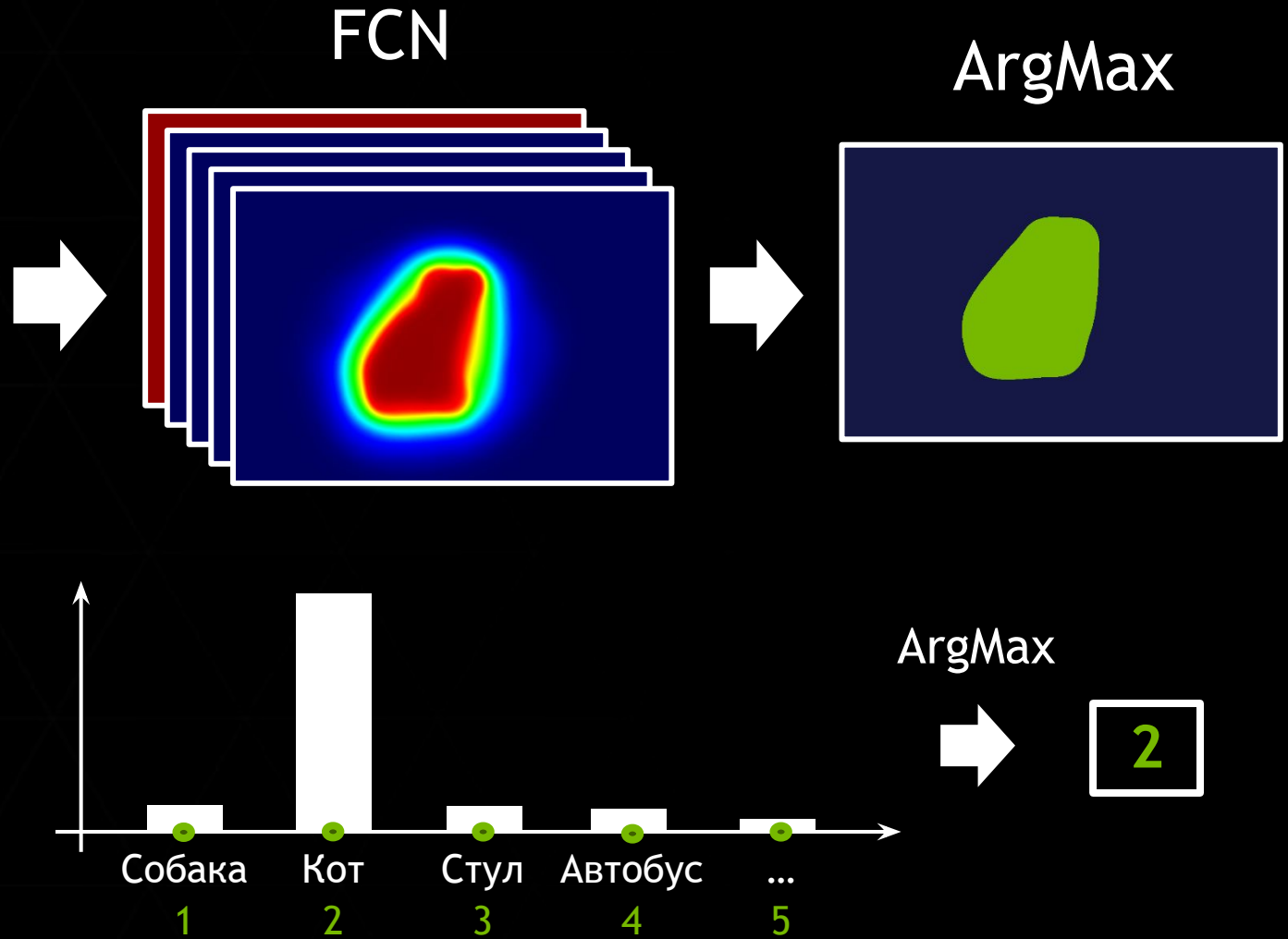
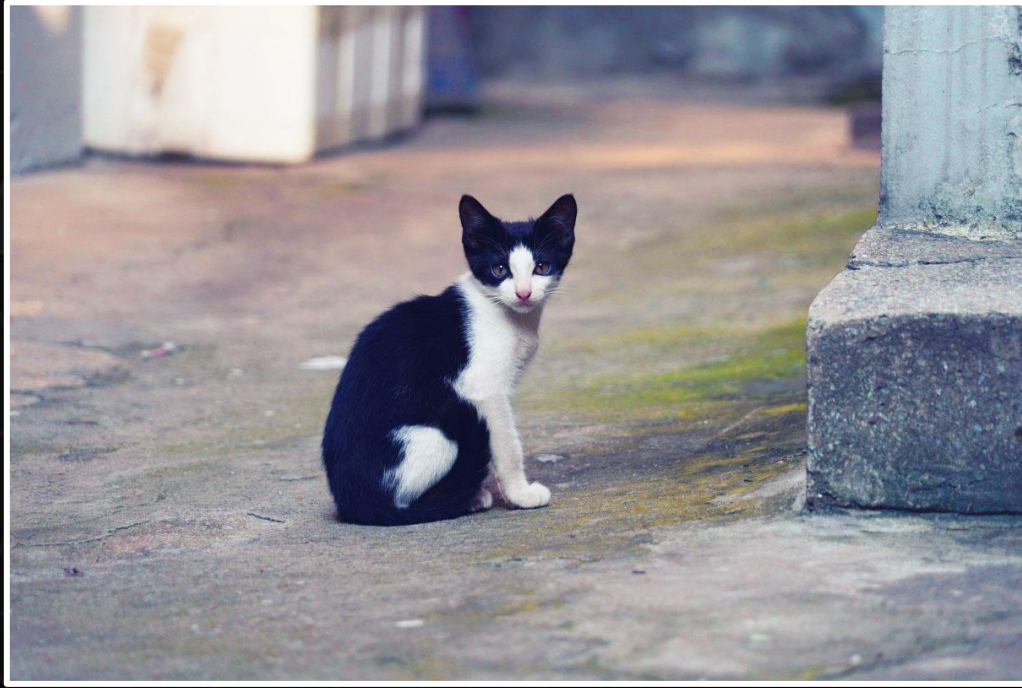


ArgMax



2

FCN + ARGMAX



FCN + UPSCALE + ARGMAX

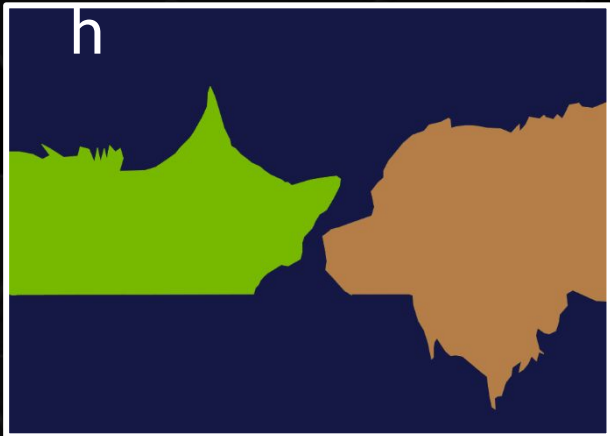


ОБУЧЕНИЕ FCN

Вход X



Ground-Truth Y



ОБУЧЕНИЕ FCN

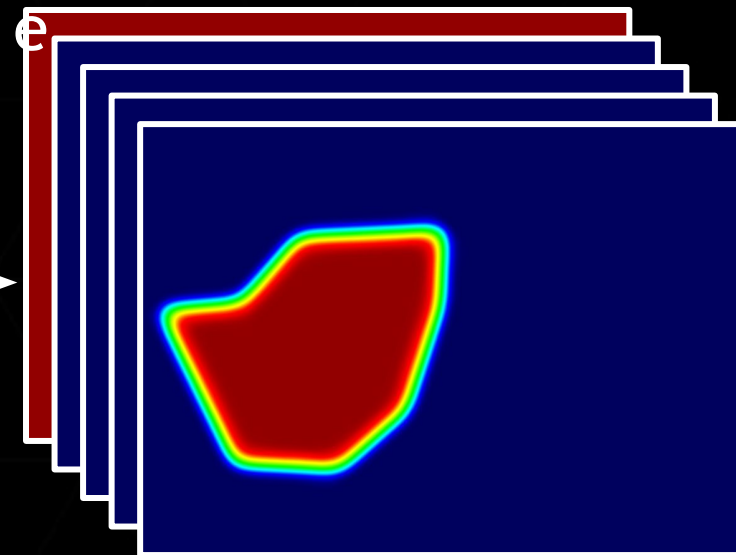
Вход X



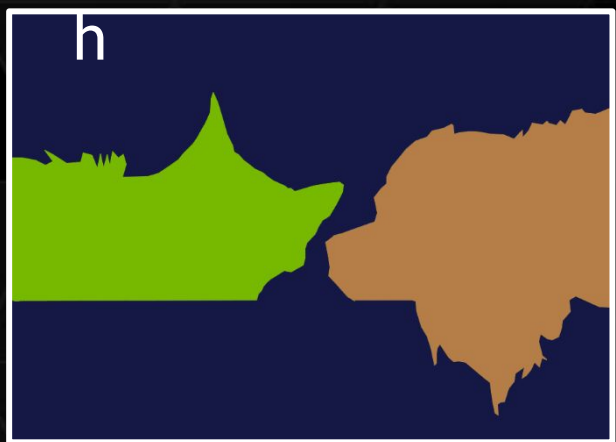
FCN



Предсказания $F(X)$



Ground-Truth Y



ОБУЧЕНИЕ FCN

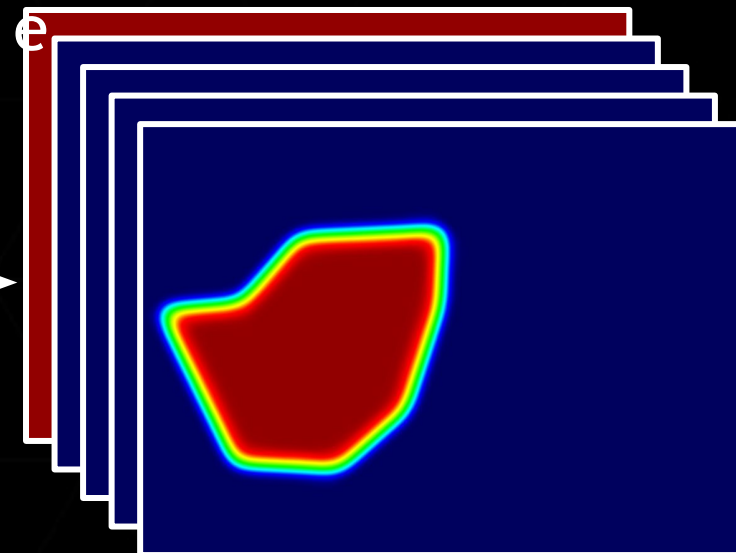
Вход X



FCN



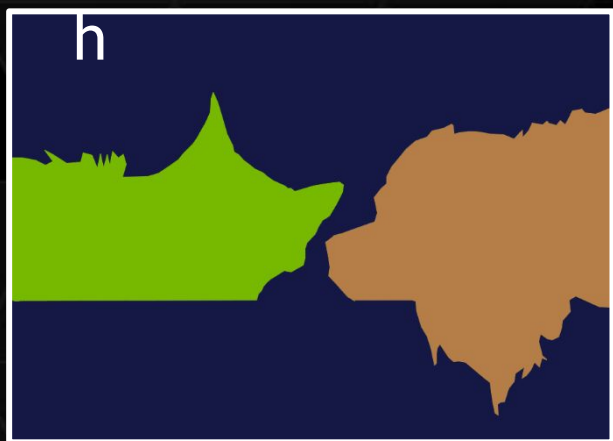
Предсказания $F(X)$



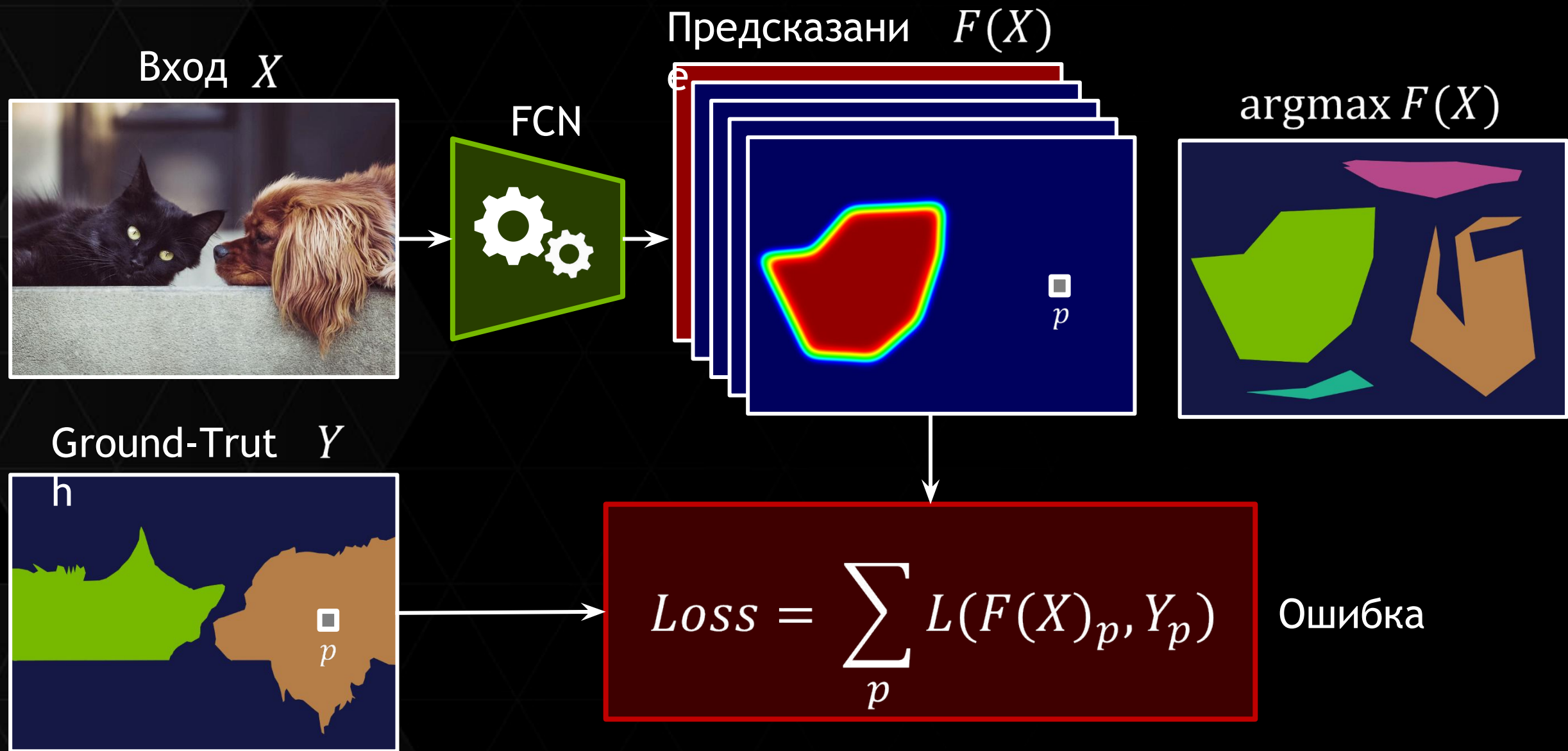
$\operatorname{argmax} F(X)$



Ground-Truth Y



ОБУЧЕНИЕ FCN



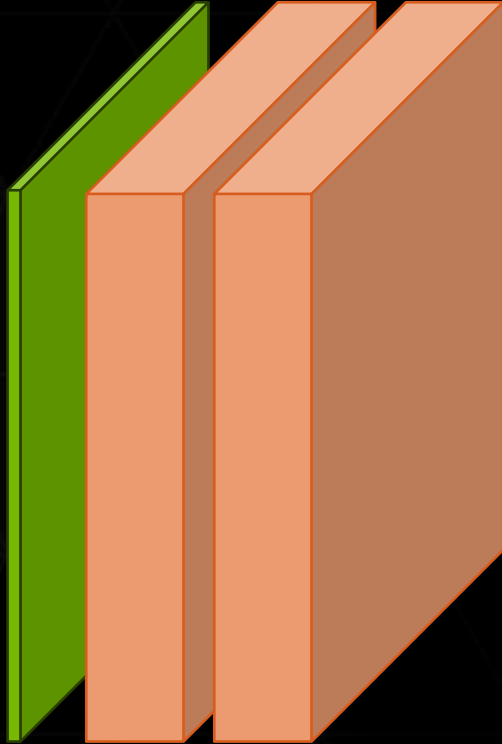


МЕТОДЫ УЛУЧШЕНИЯ FCSN МОДЕЛИ

FCN + UPSCALE

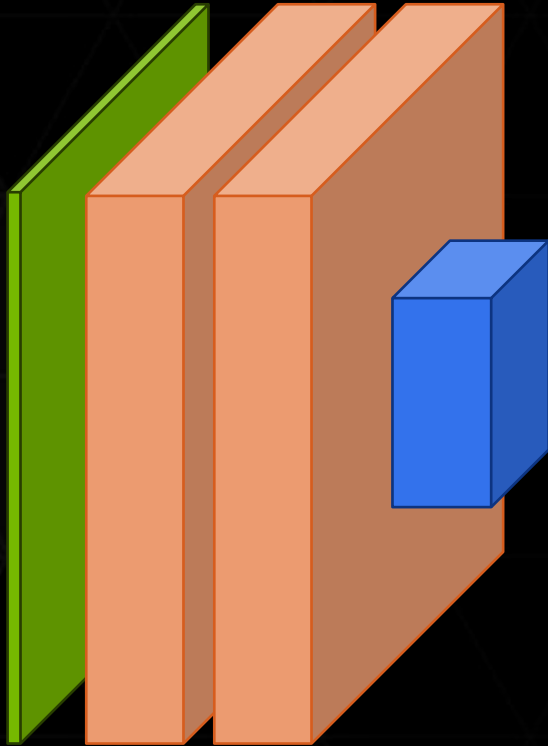


FCN + UPSCALE



Выход Conv слоя

FCN + UPSCALE

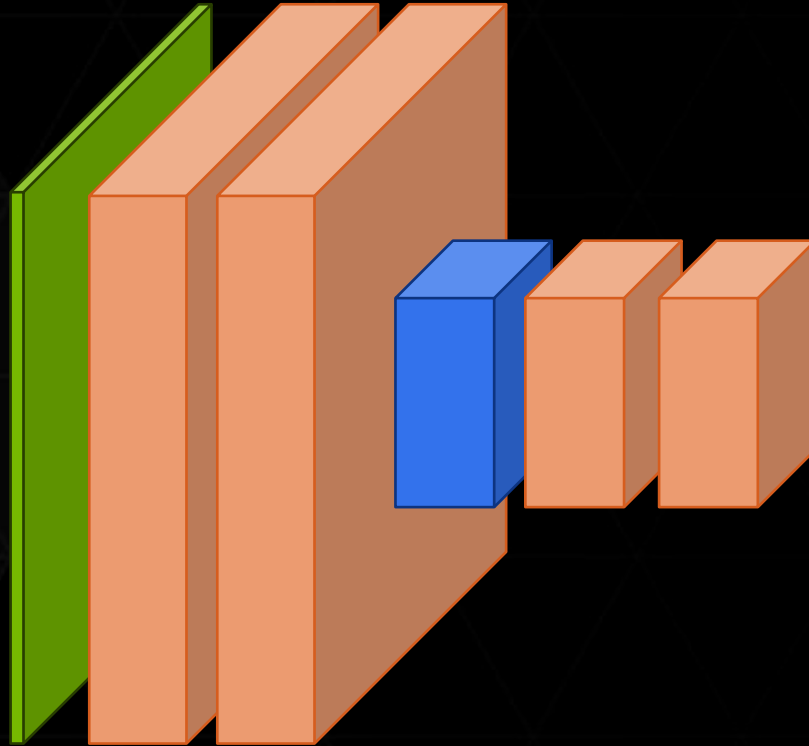


Выход Conv слоя



Выход Pooling слоя

FCN + UPSCALE

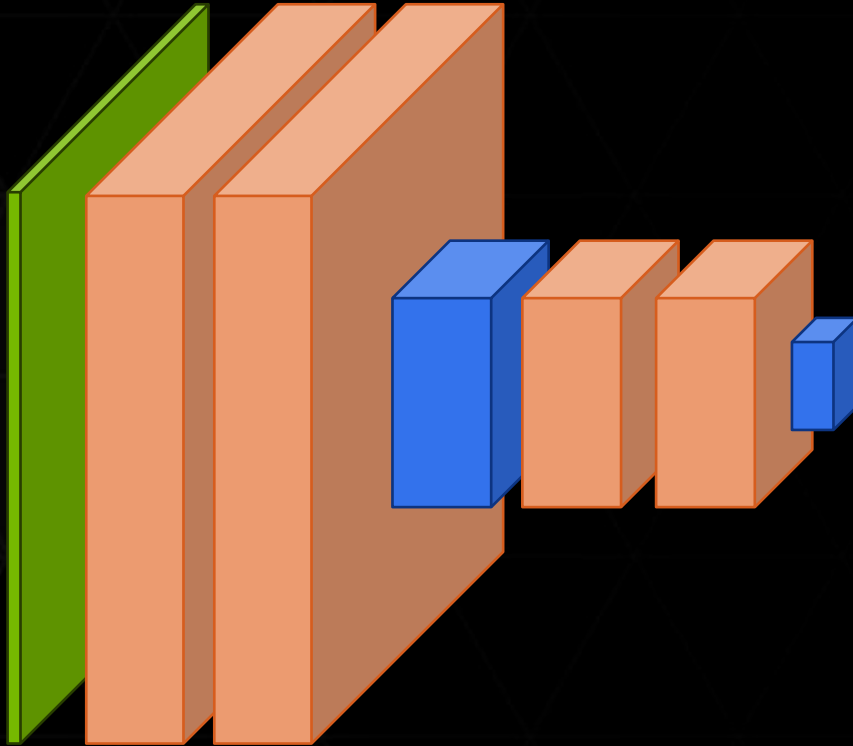




Выход Conv слоя



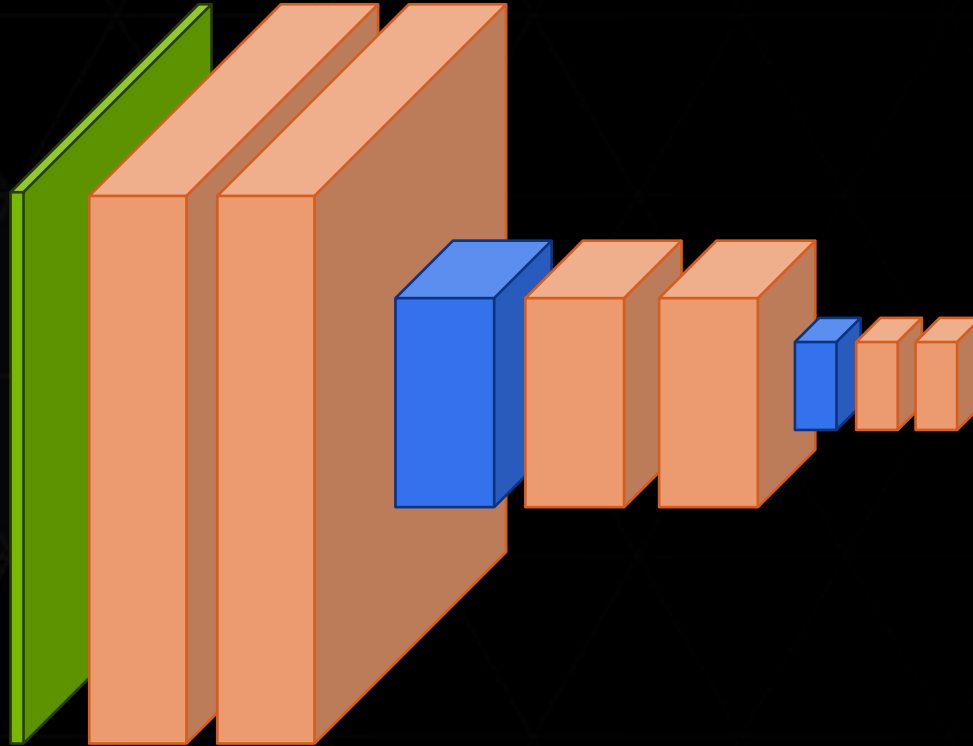
Выход Pooling слоя

FCN + UPSCALE



-  Выход Conv слоя
-  Выход Pooling слоя

FCN + UPSCALE

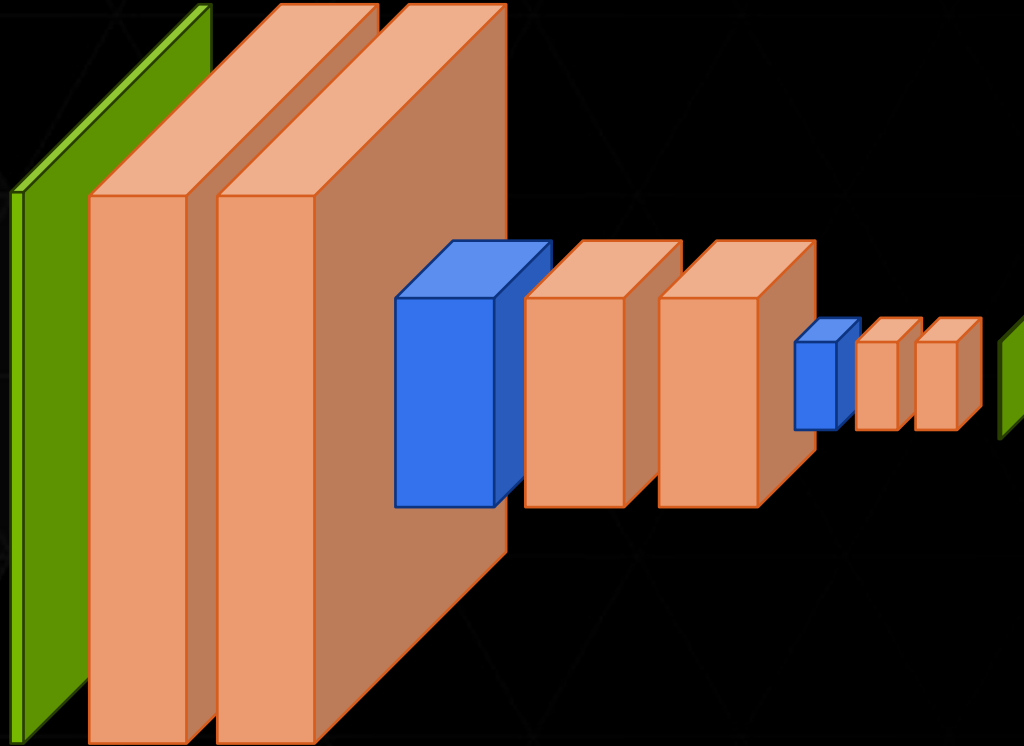




Выход Conv слоя



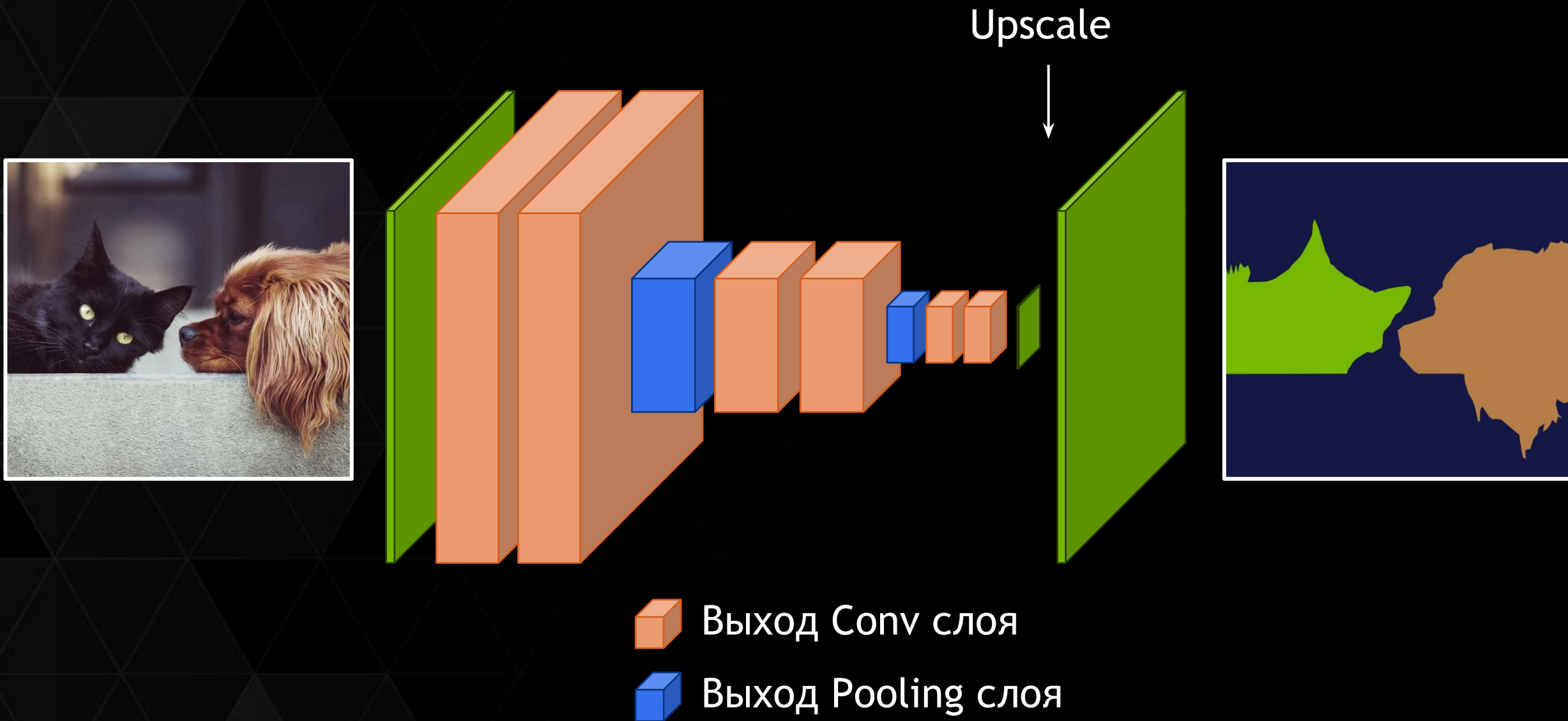
Выход Pooling слоя

FCN + UPSCALE



-  Выход Conv слоя
-  Выход Pooling слоя

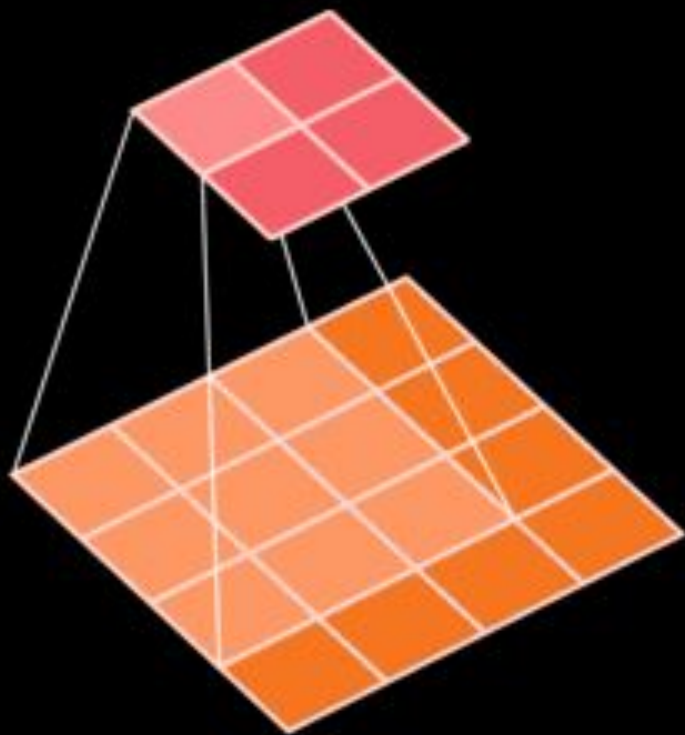
FCN + UPSCALE



СВЁРТКА СО СТРАЙДОМ

Conv

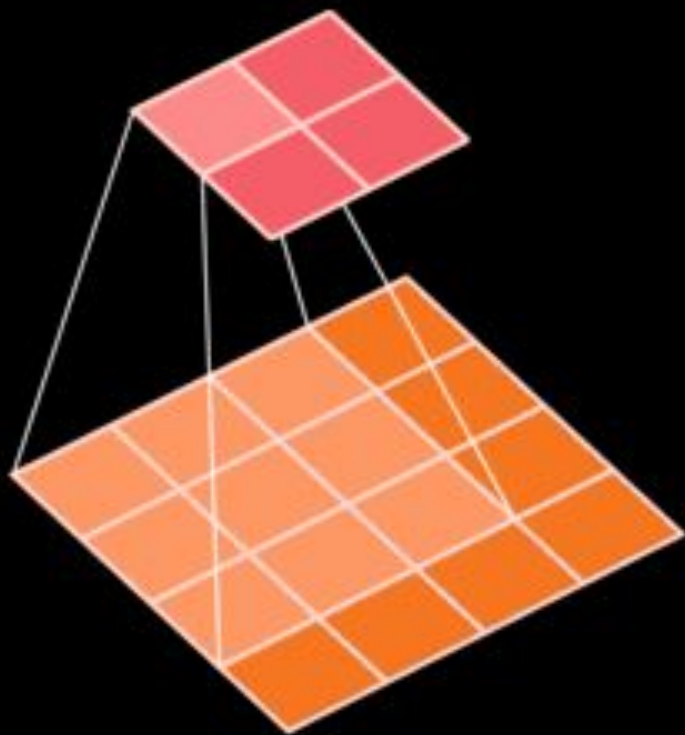
Шаг окна: 1



СВЁРТКА СО СТРАЙДОМ

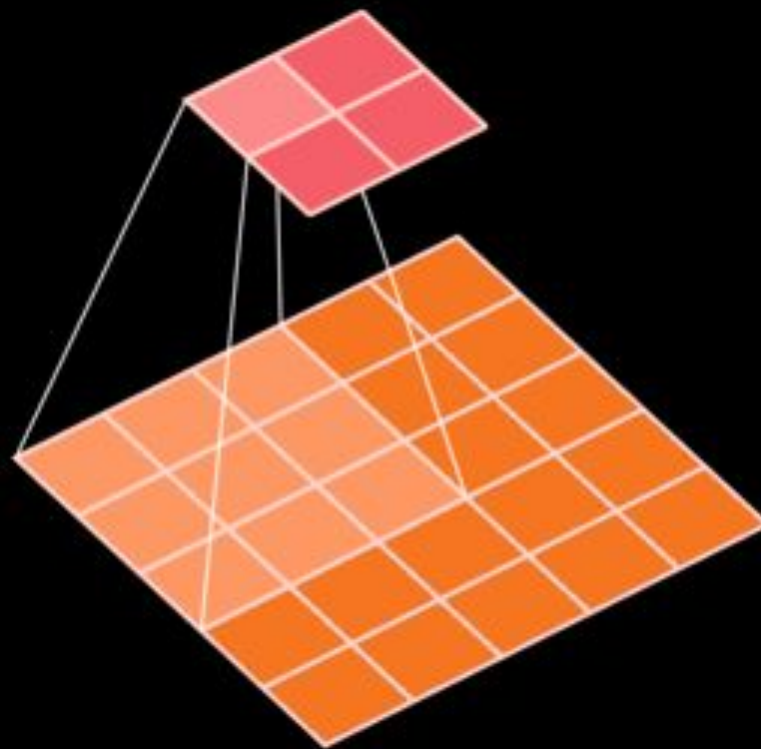
Conv

Шаг окна: 1

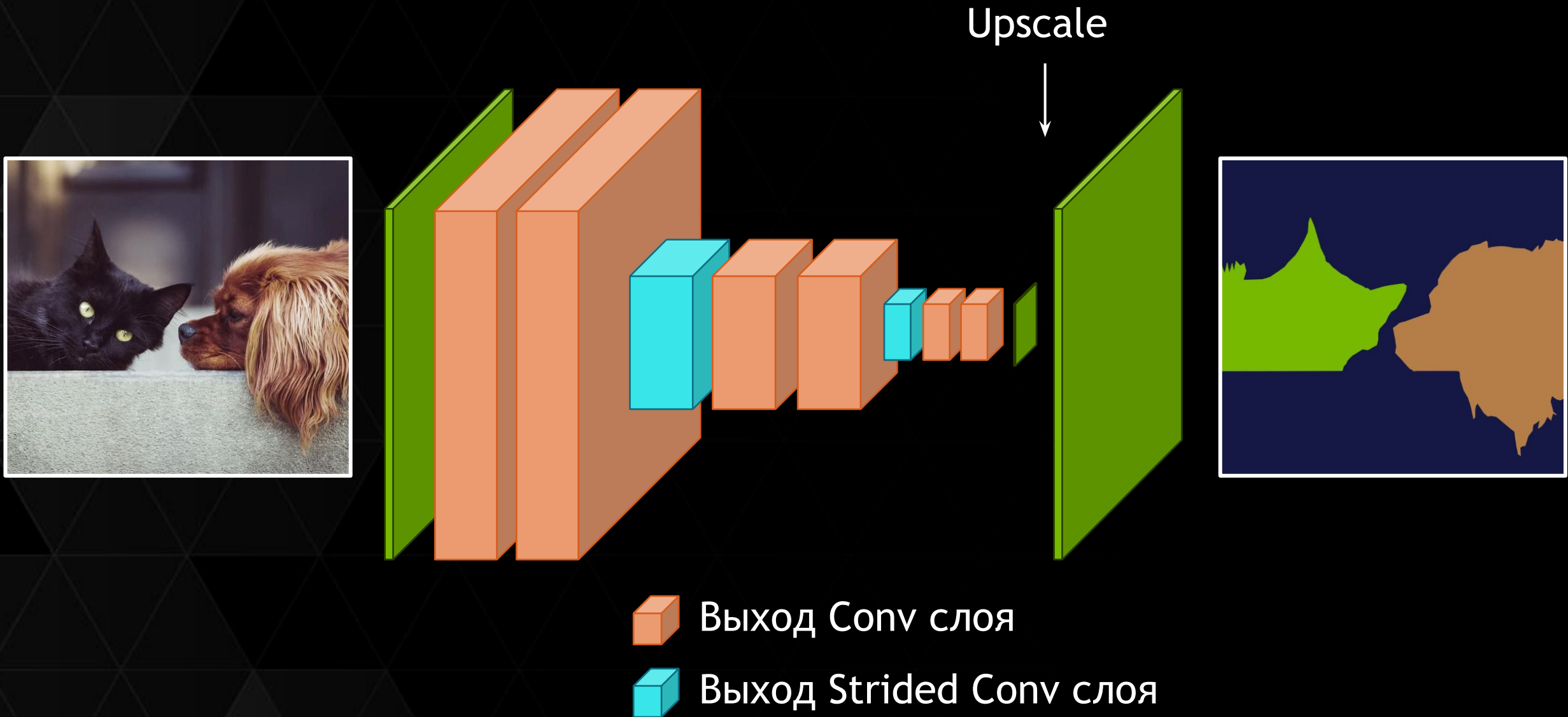


Strided Conv

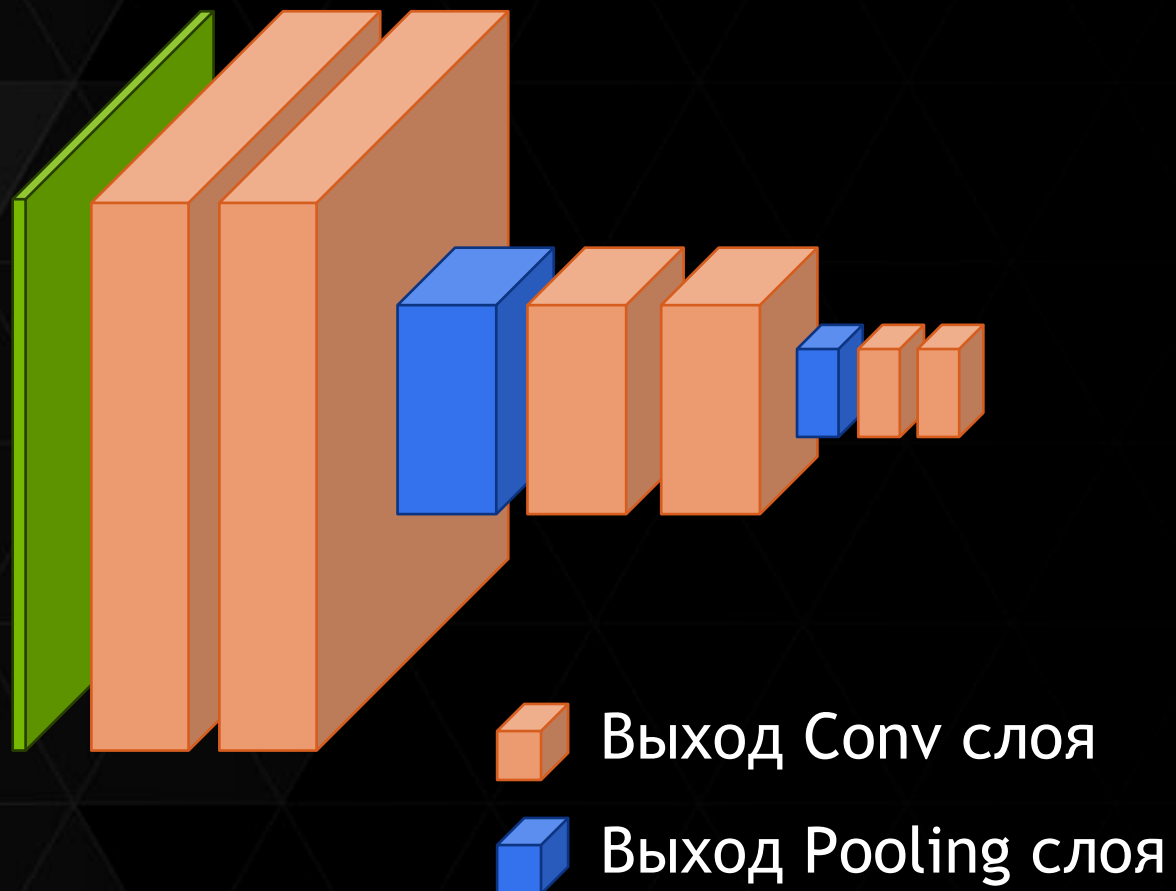
Шаг окна: 2



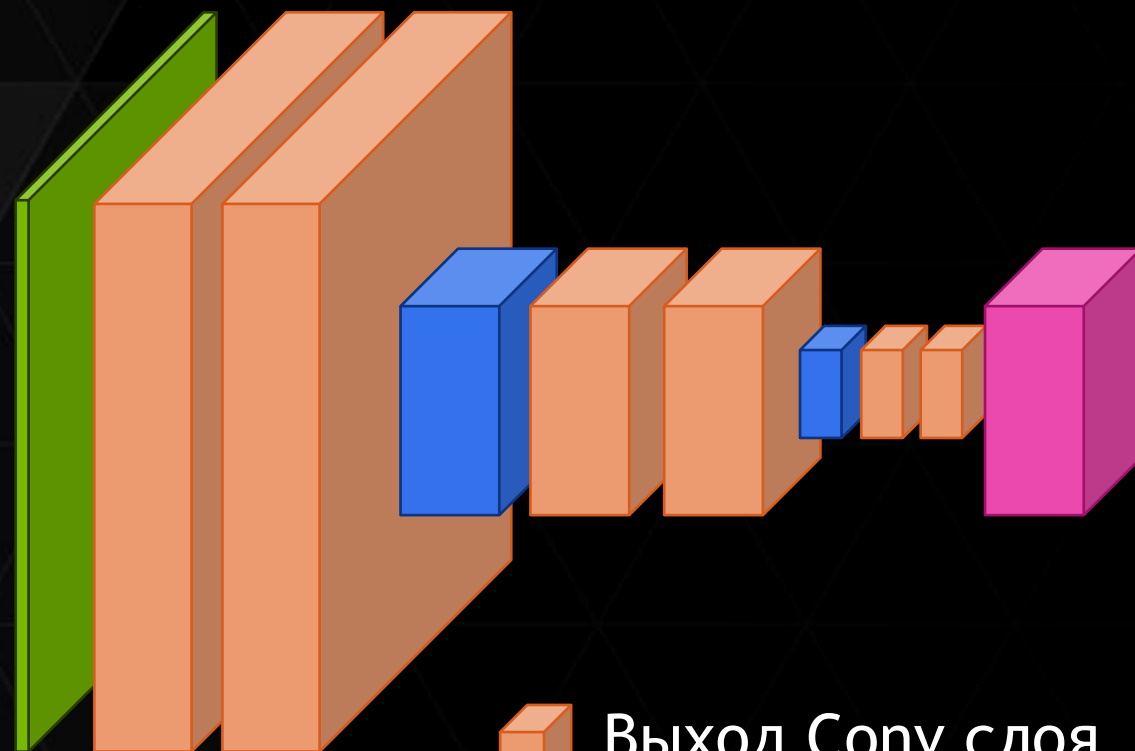
СВЁРТКА СО СТРАЙДОМ



СЛОИ ДЛЯ ПОВЫШЕНИЯ РАЗРЕШЕНИЯ



СЛОИ ДЛЯ ПОВЫШЕНИЯ РАЗРЕШЕНИЯ



Выход Conv слоя

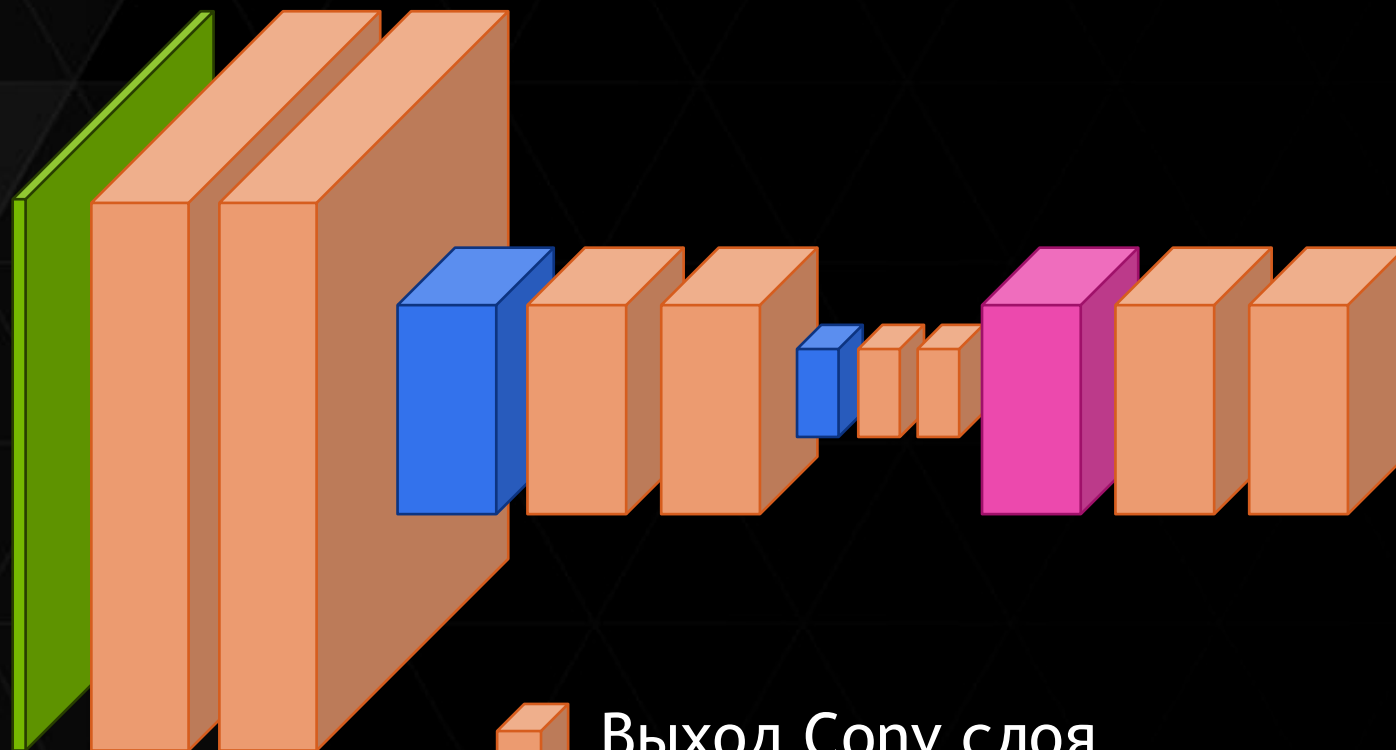


Выход Pooling слоя



Выход **Unpooling** / **Upconv** слоя

СЛОИ ДЛЯ ПОВЫШЕНИЯ РАЗРЕШЕНИЯ



Выход Conv слоя

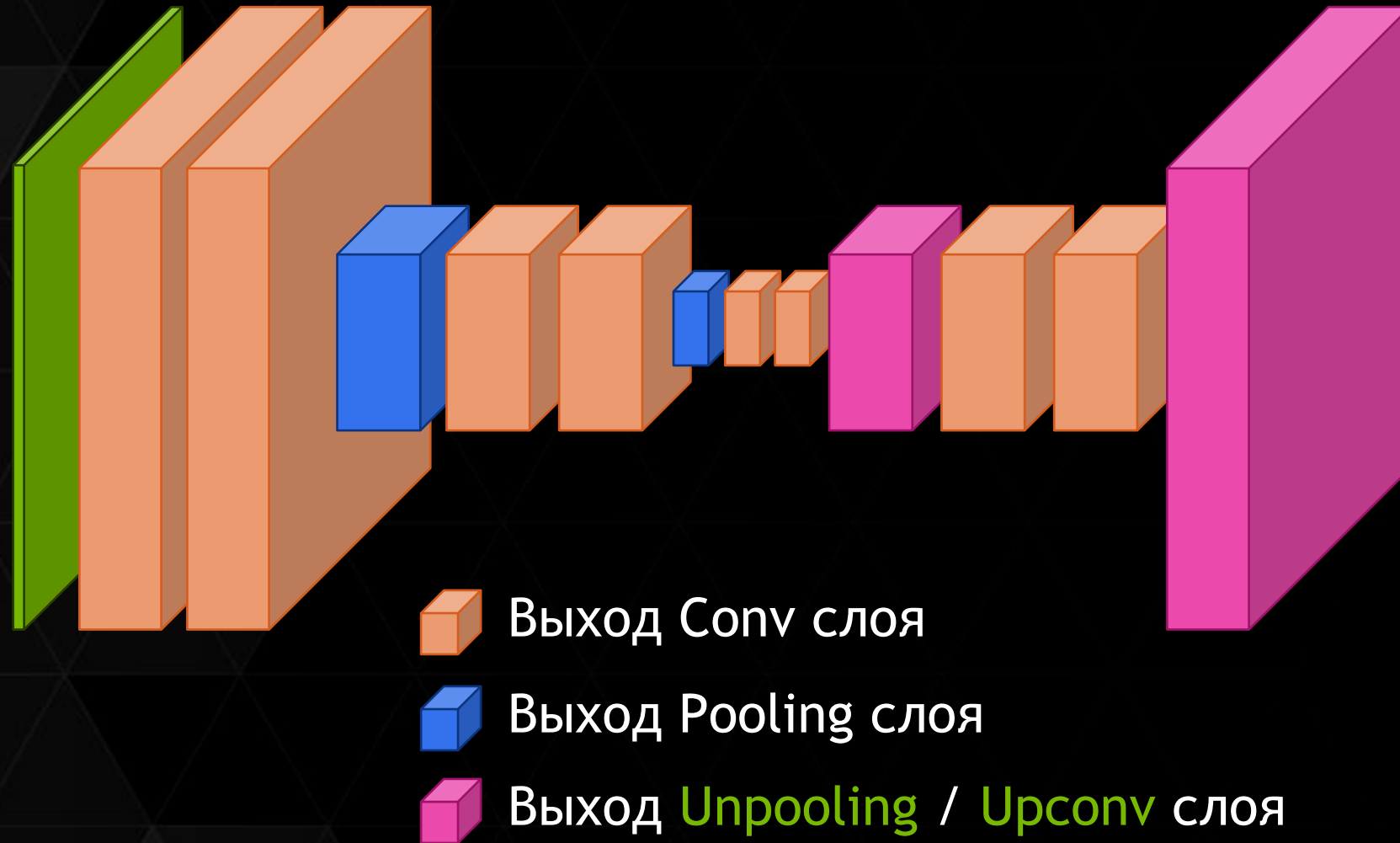


Выход Pooling слоя

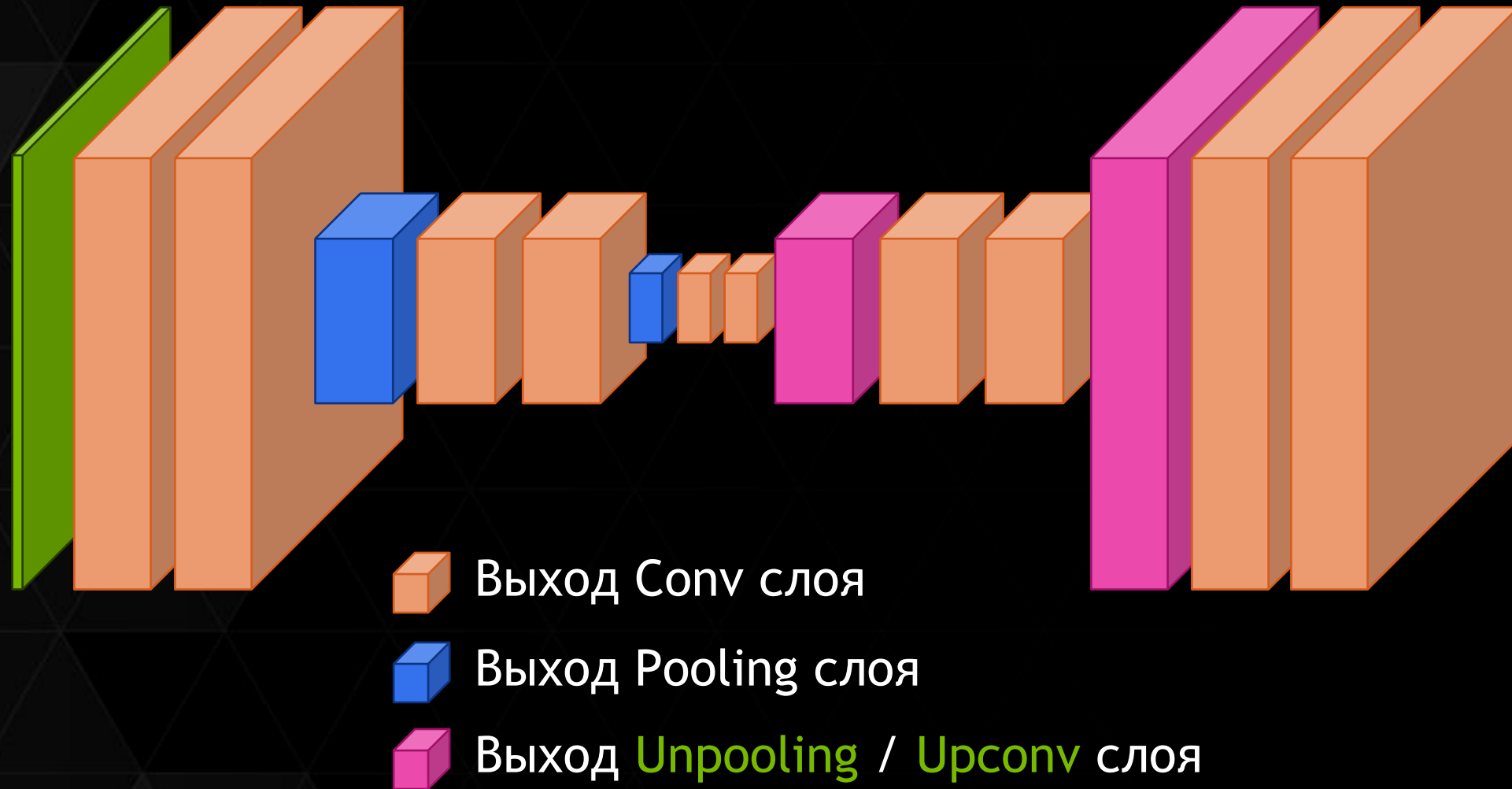


Выход **Unpooling** / **Upconv** слоя

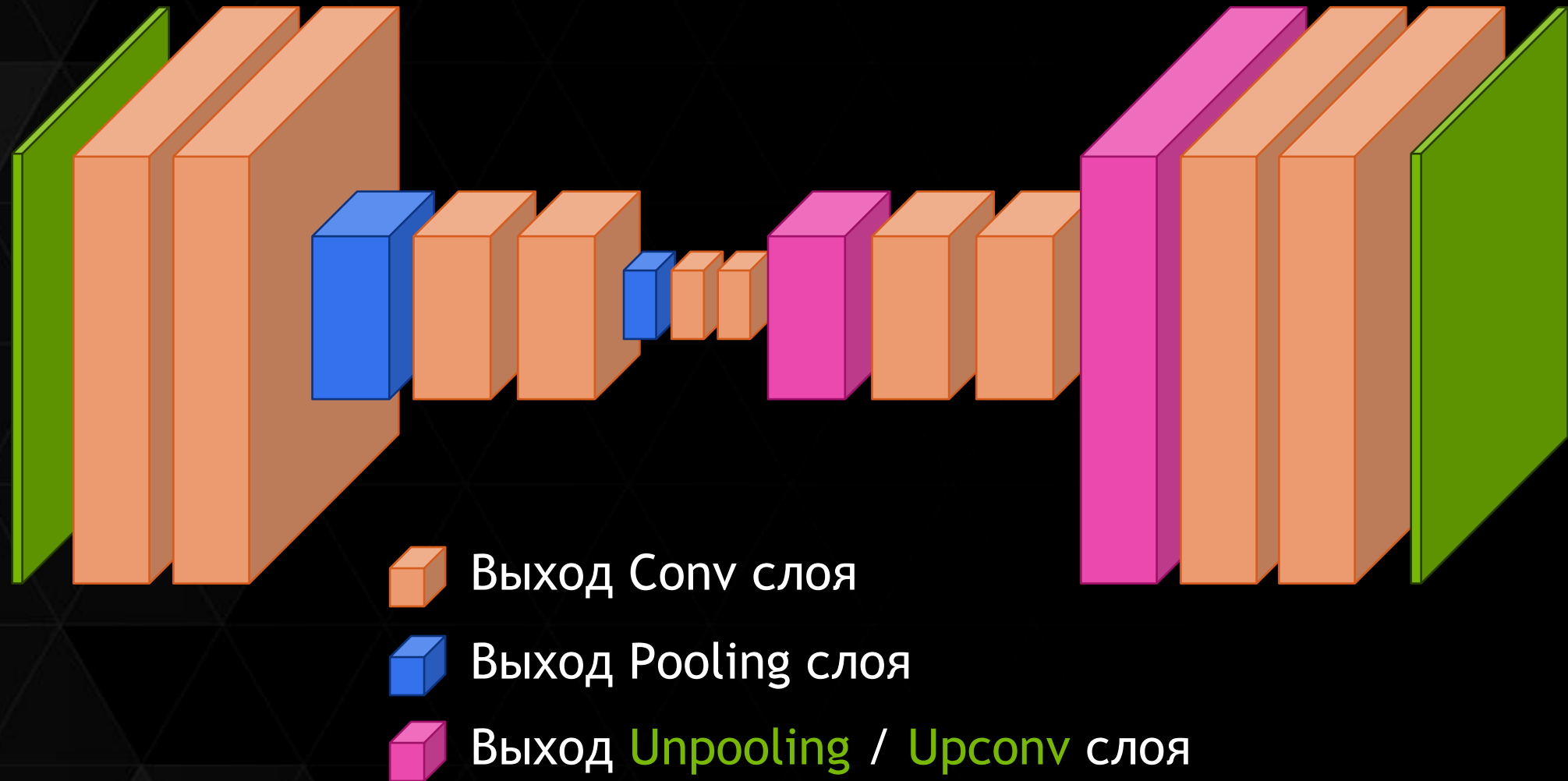
СЛОИ ДЛЯ ПОВЫШЕНИЯ РАЗРЕШЕНИЯ



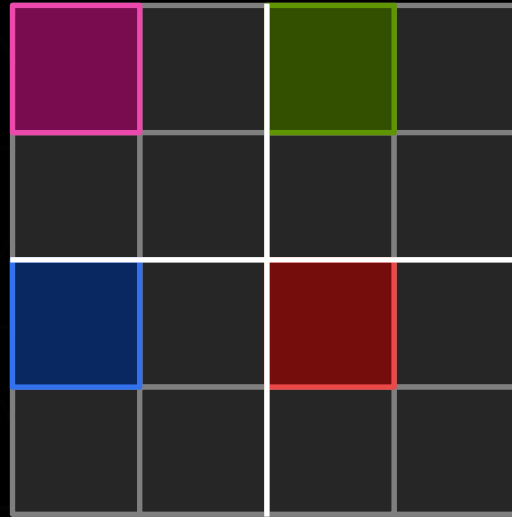
СЛОИ ДЛЯ ПОВЫШЕНИЯ РАЗРЕШЕНИЯ



СЛОИ ДЛЯ ПОВЫШЕНИЯ РАЗРЕШЕНИЯ

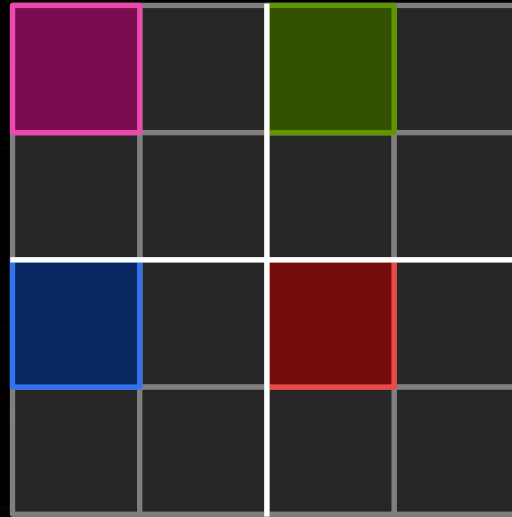
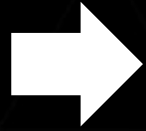


UNPOOLING СЛОЙ



Разреженный
Unpooling

UNPOOLING СЛОЙ

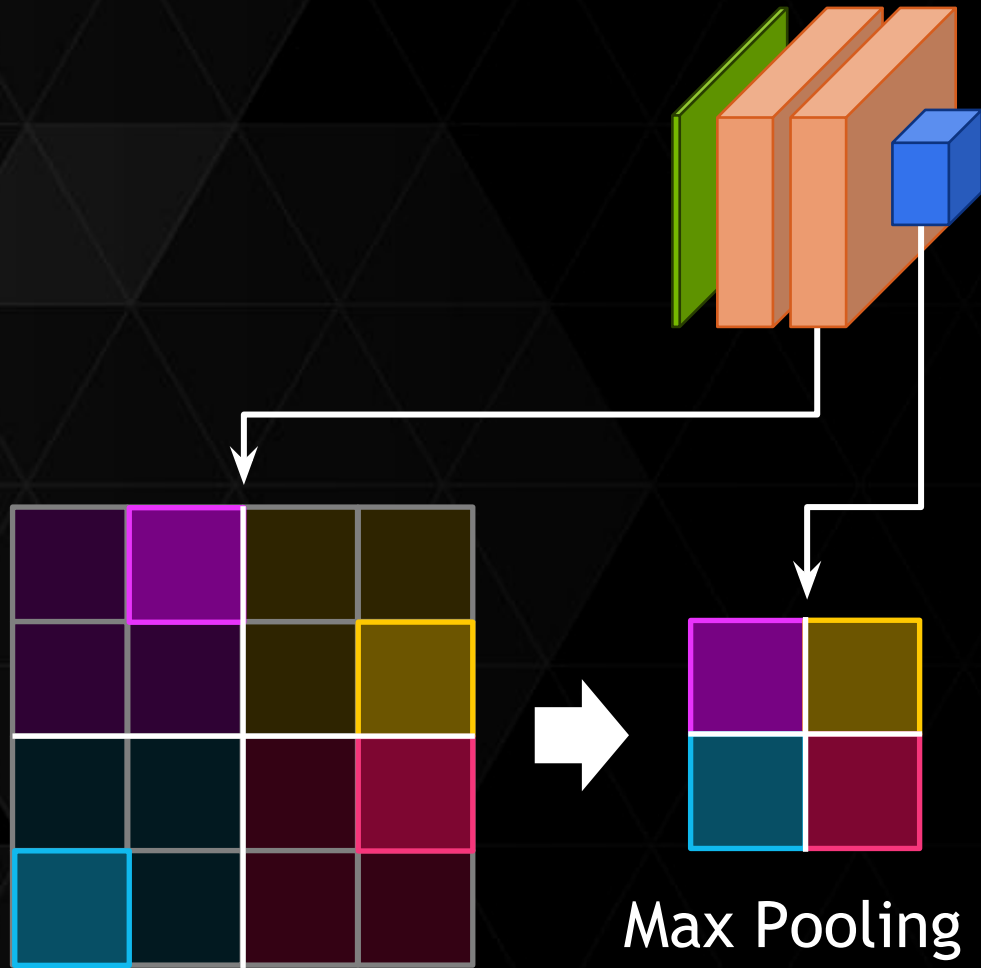


Разреженный
Unpooling

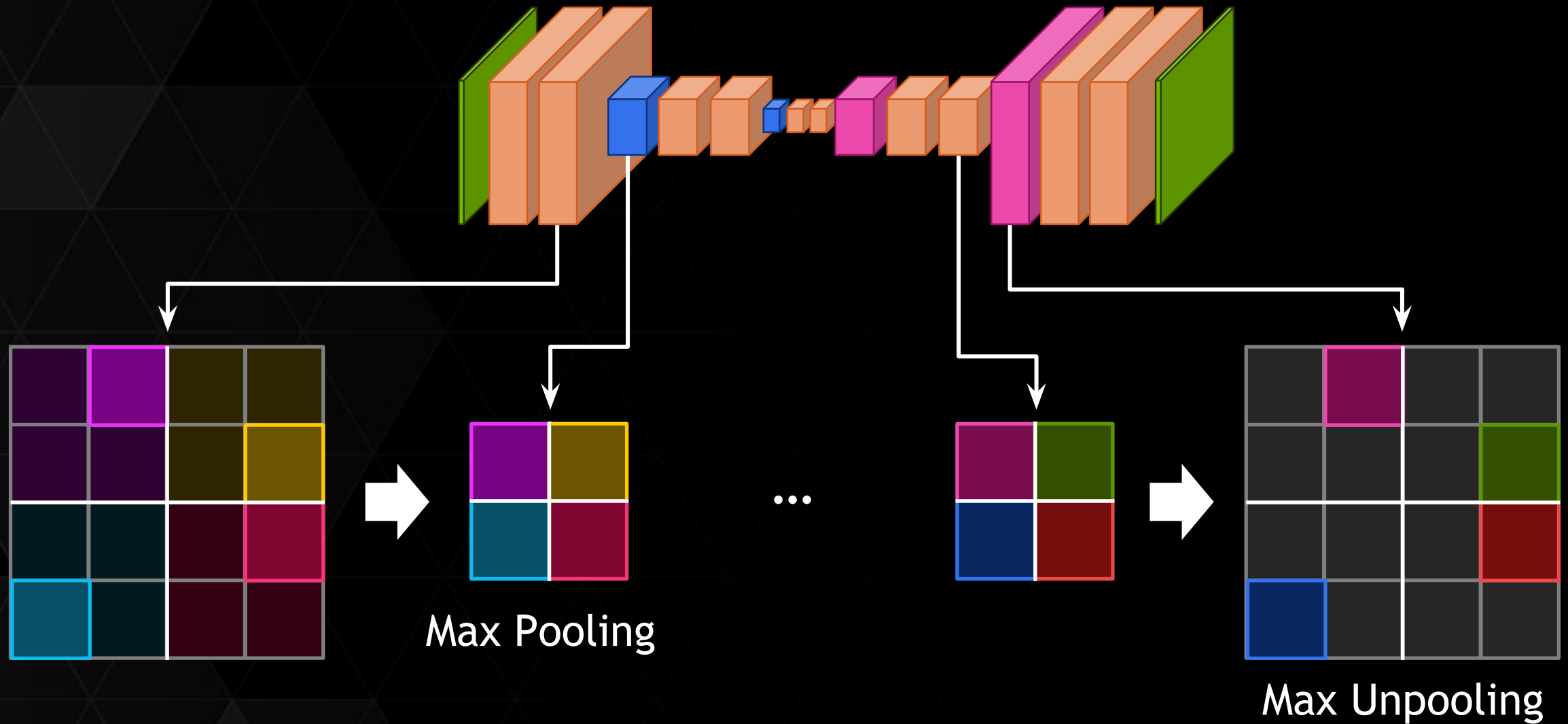


Ближайший
сосед

MAX UNPOOLING

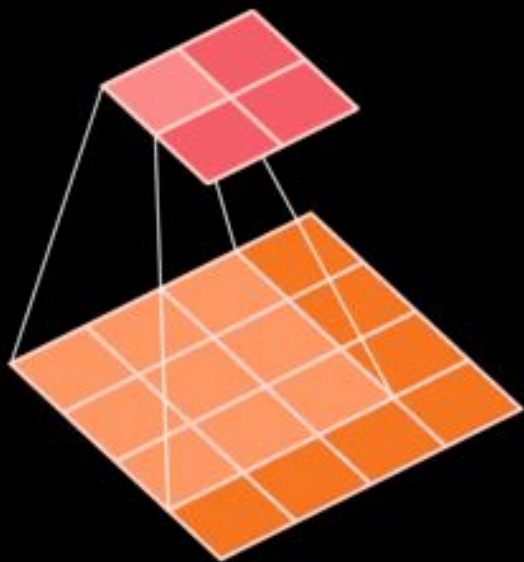


MAX UNPOOLING



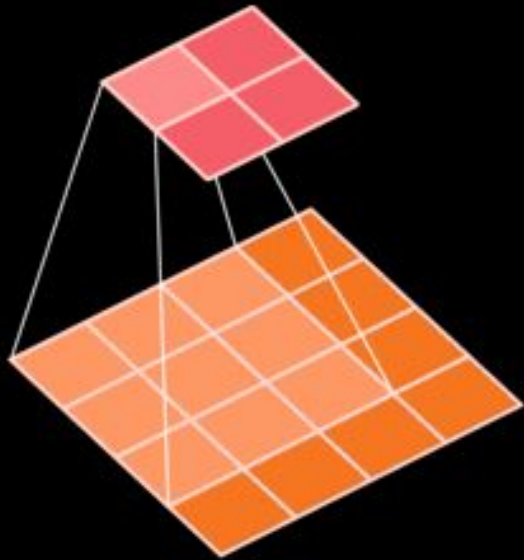
ТРАНСПОНИРОВАННАЯ СВЁРТКА

Обычная свёртка

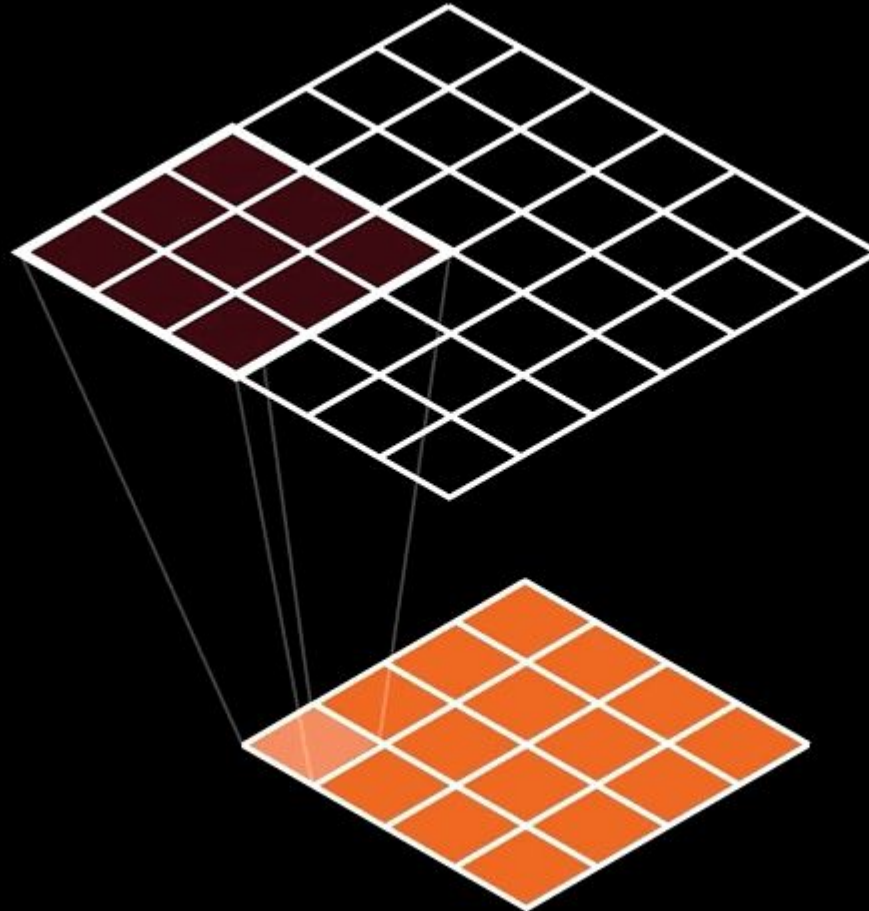


ТРАНСПОНИРОВАННАЯ СВЁРТКА

Обычная свёртка



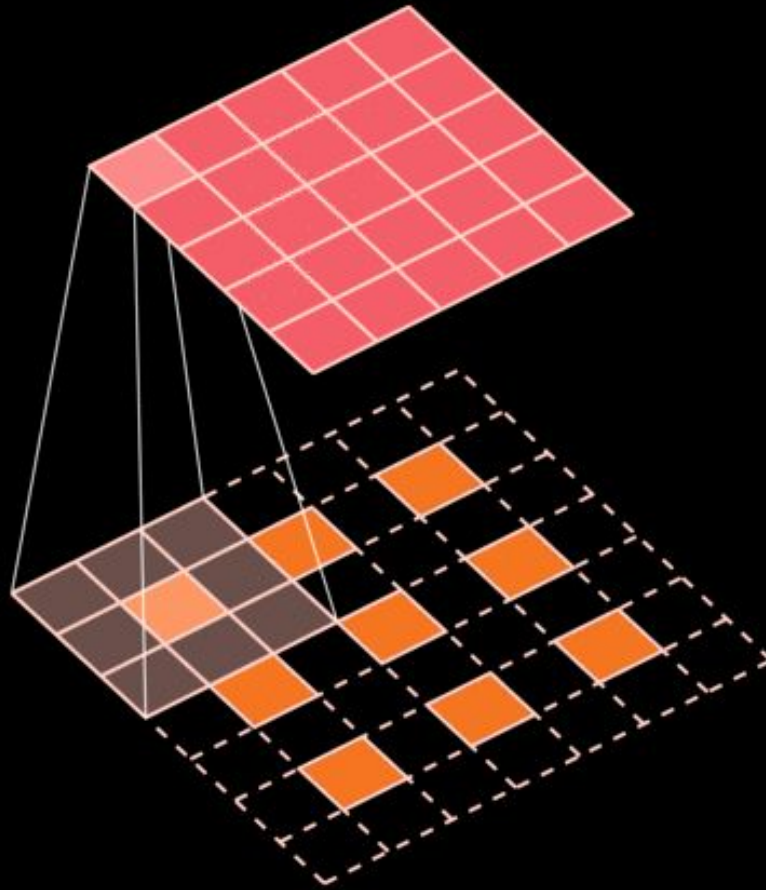
Транспонированная свёртка



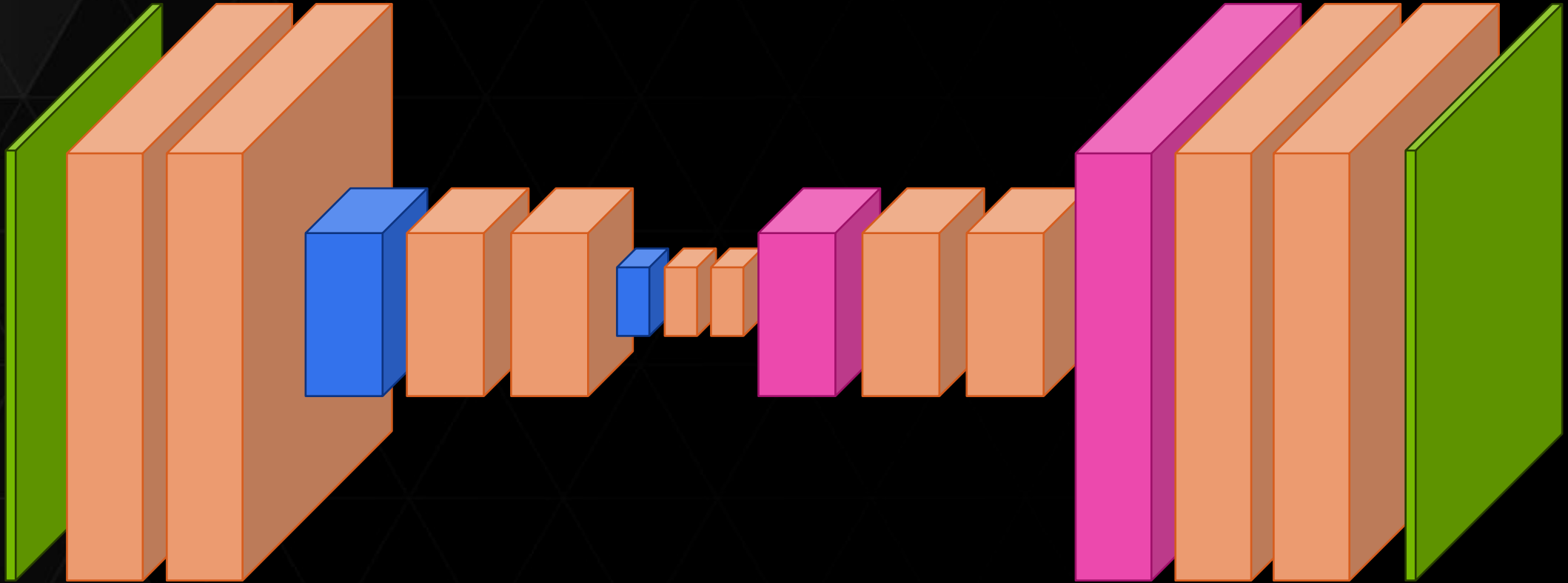
ТРАНСПОНИРОВАННАЯ СВЁРТКА СО СТРАДОМ

Upconv:

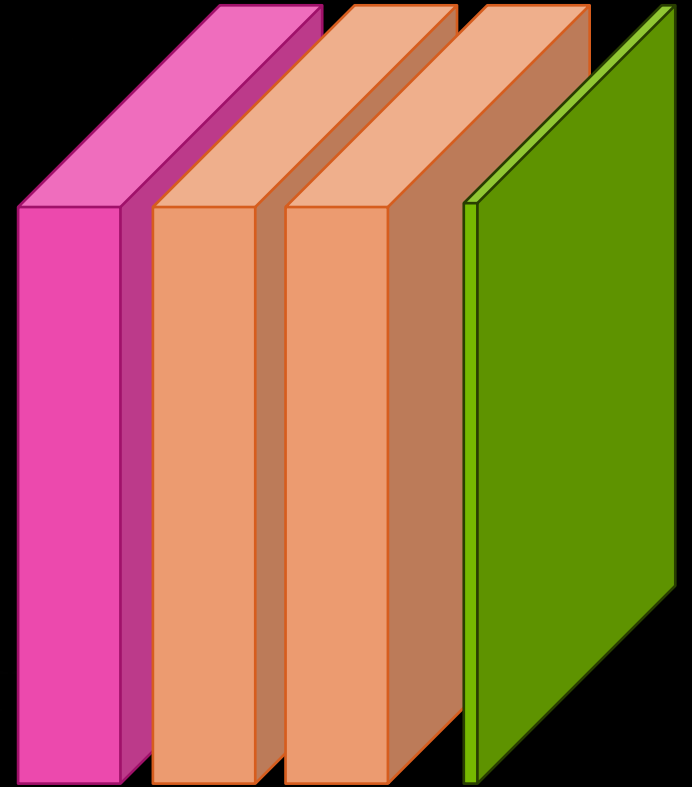
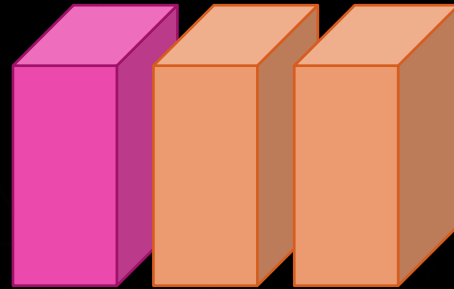
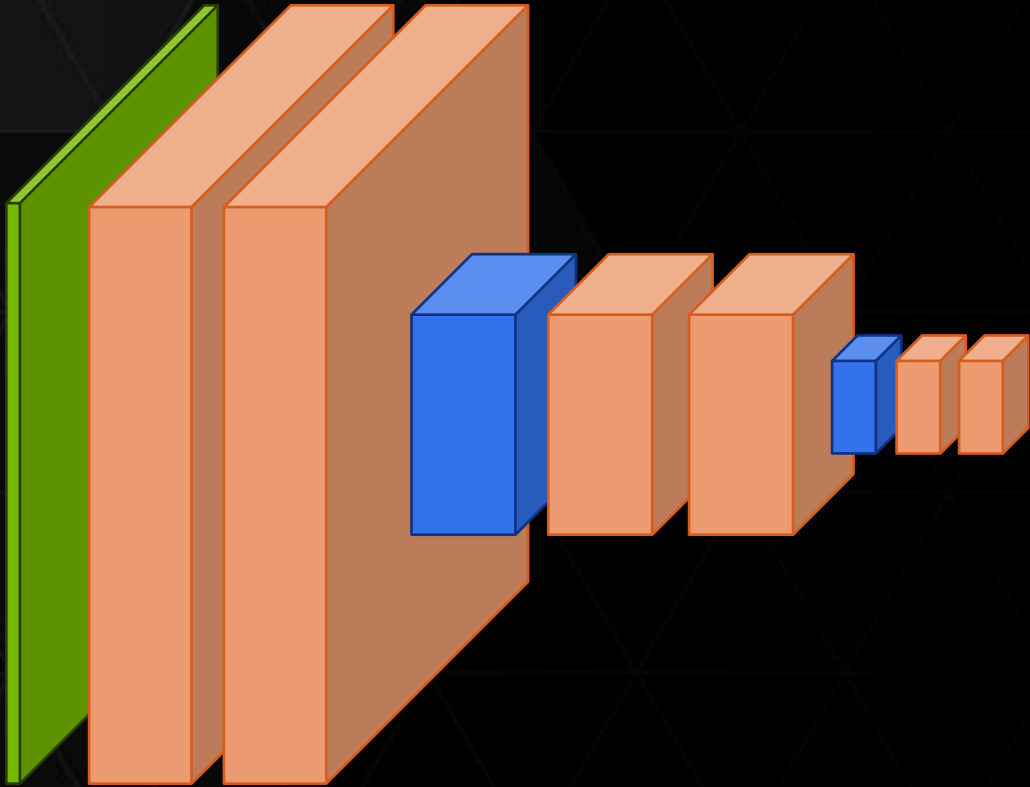
Транспонированная свёртка
stride=2



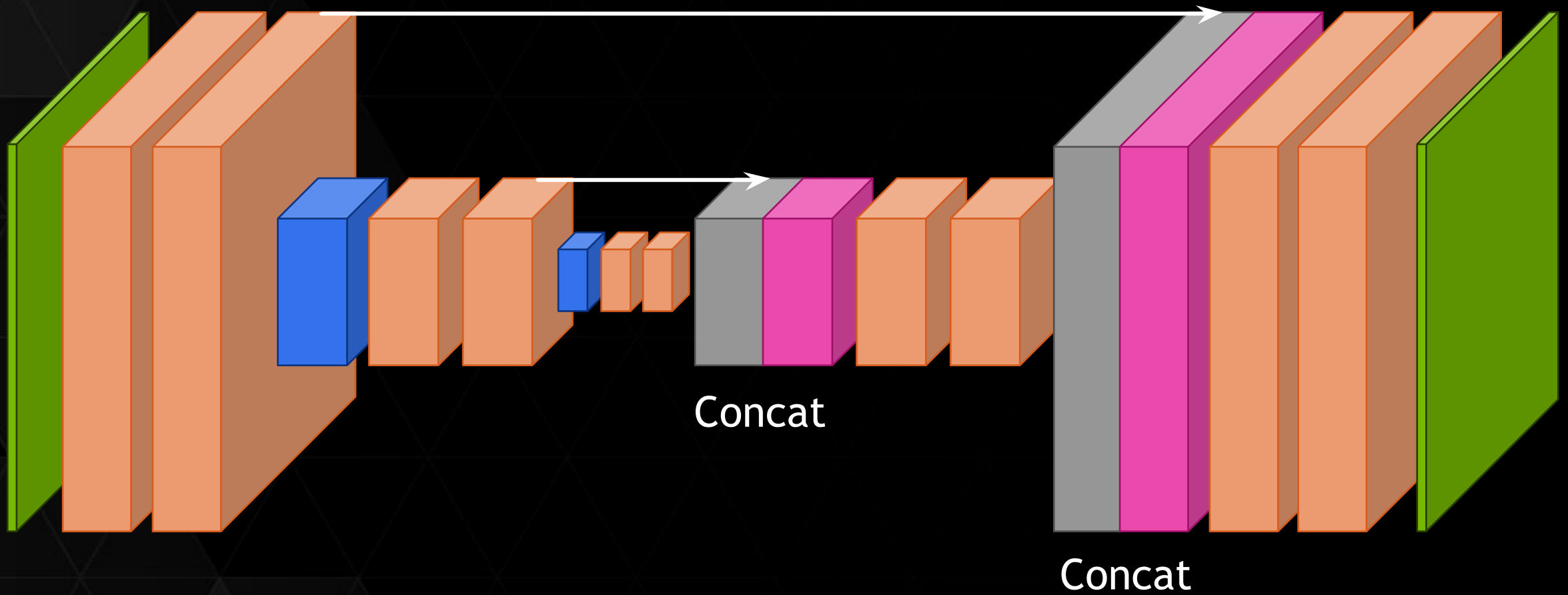
SKIPPED CONNECTIONS



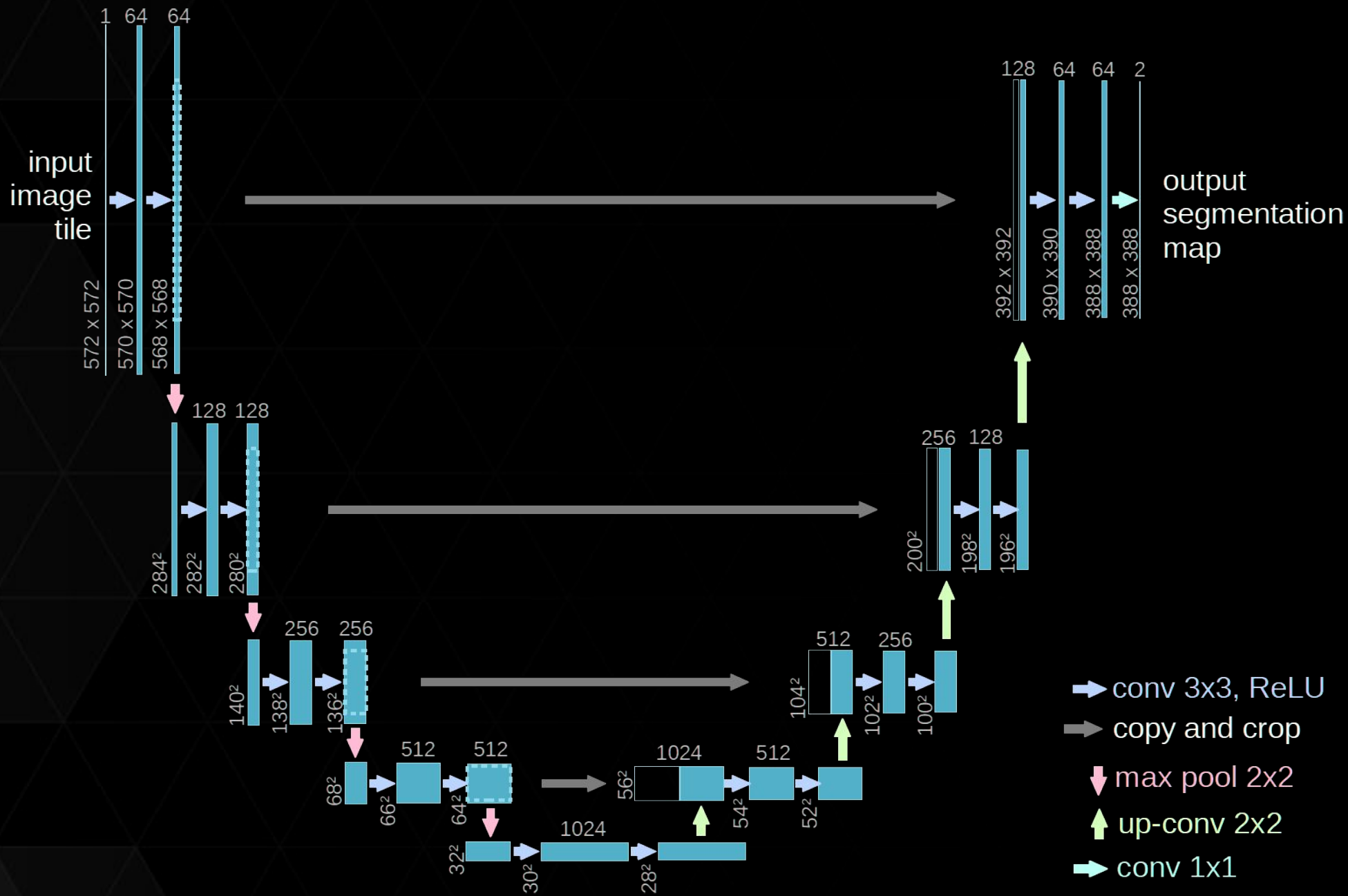
SKIPPED CONNECTIONS



SKIPPED CONNECTIONS



U-NET



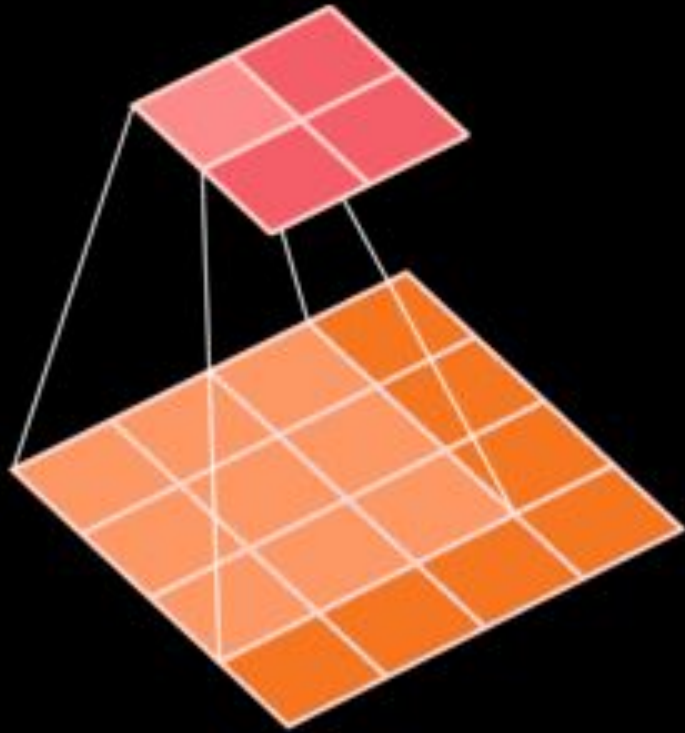
ДИЛАТАЦИОННАЯ СВЁРТКА (DILATED, ATROUS)

Conv

Количество параметров:

3x3

Рецептивное поле: 3x3

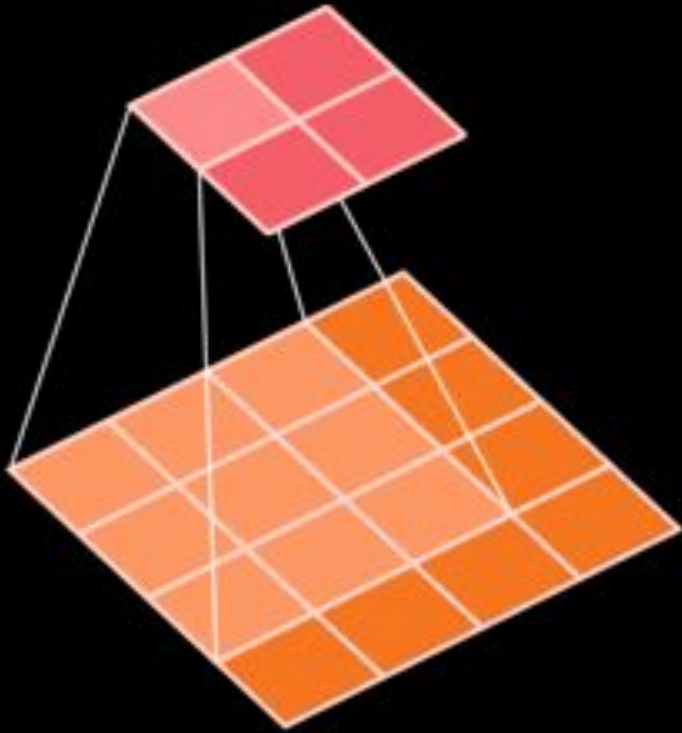


ДИЛАТАЦИОННАЯ СВЁРТКА (DILATED, ATROUS)

Conv

Количество параметров: 3×3

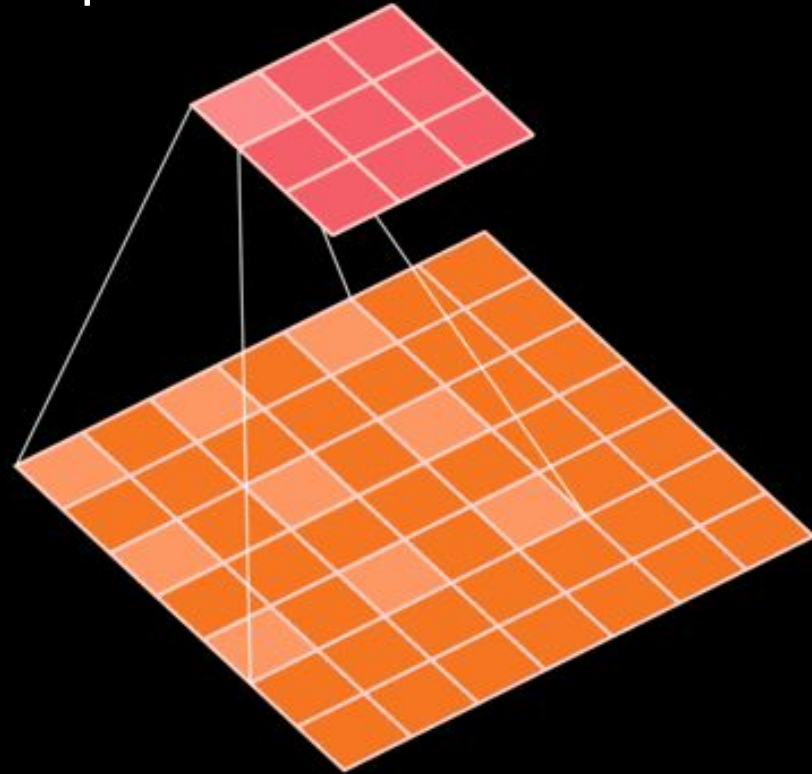
Рецептивное поле: 3×3



Dilated / Atrous Conv

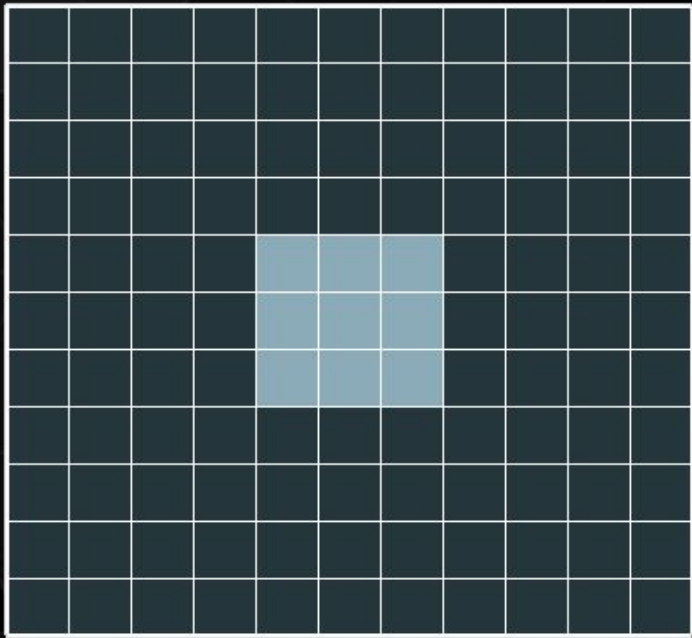
Количество параметров: 3×3

Рецептивное поле: 5×5

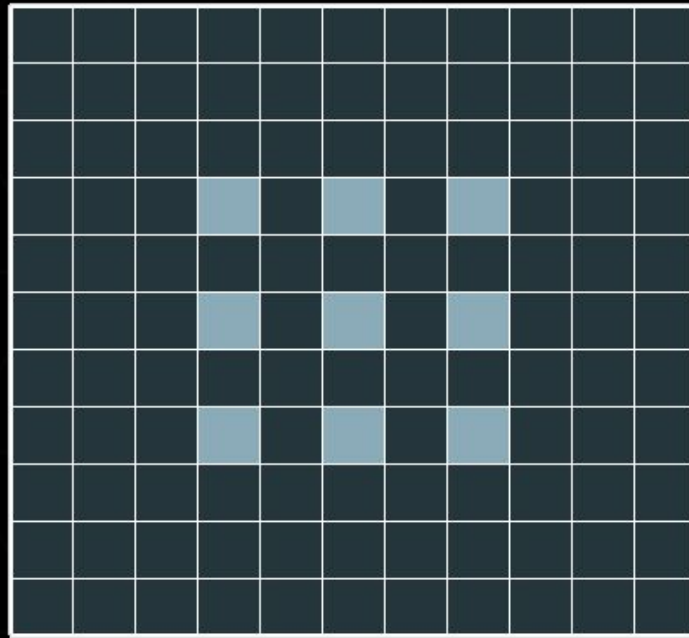


ДИЛАТАЦИОННАЯ СВЁРТКА (DILATED, ATROUS)

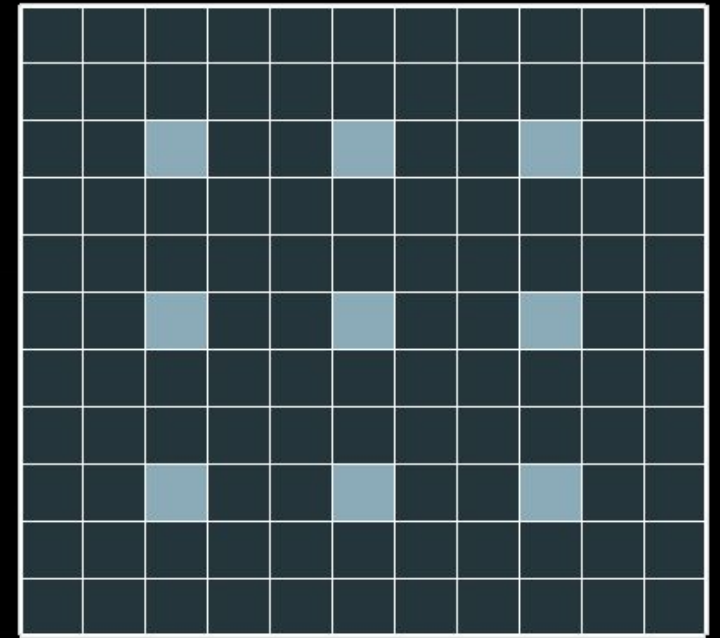
$d = 1$



$d = 2$

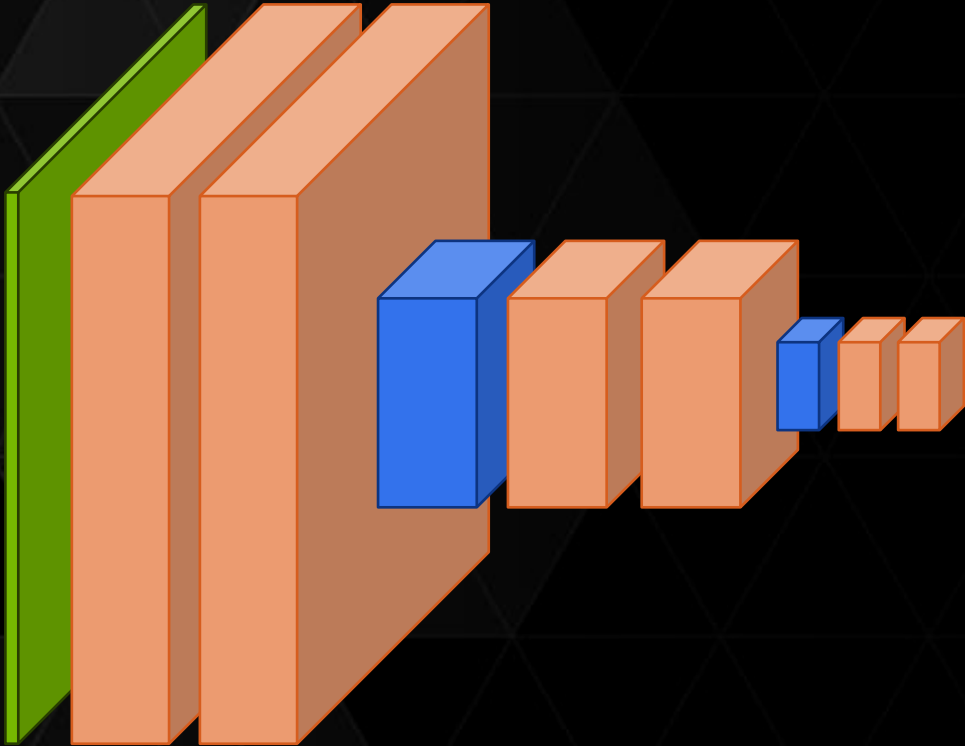


$d = 3$



ATROUS SPATIAL PYRAMID POOLING

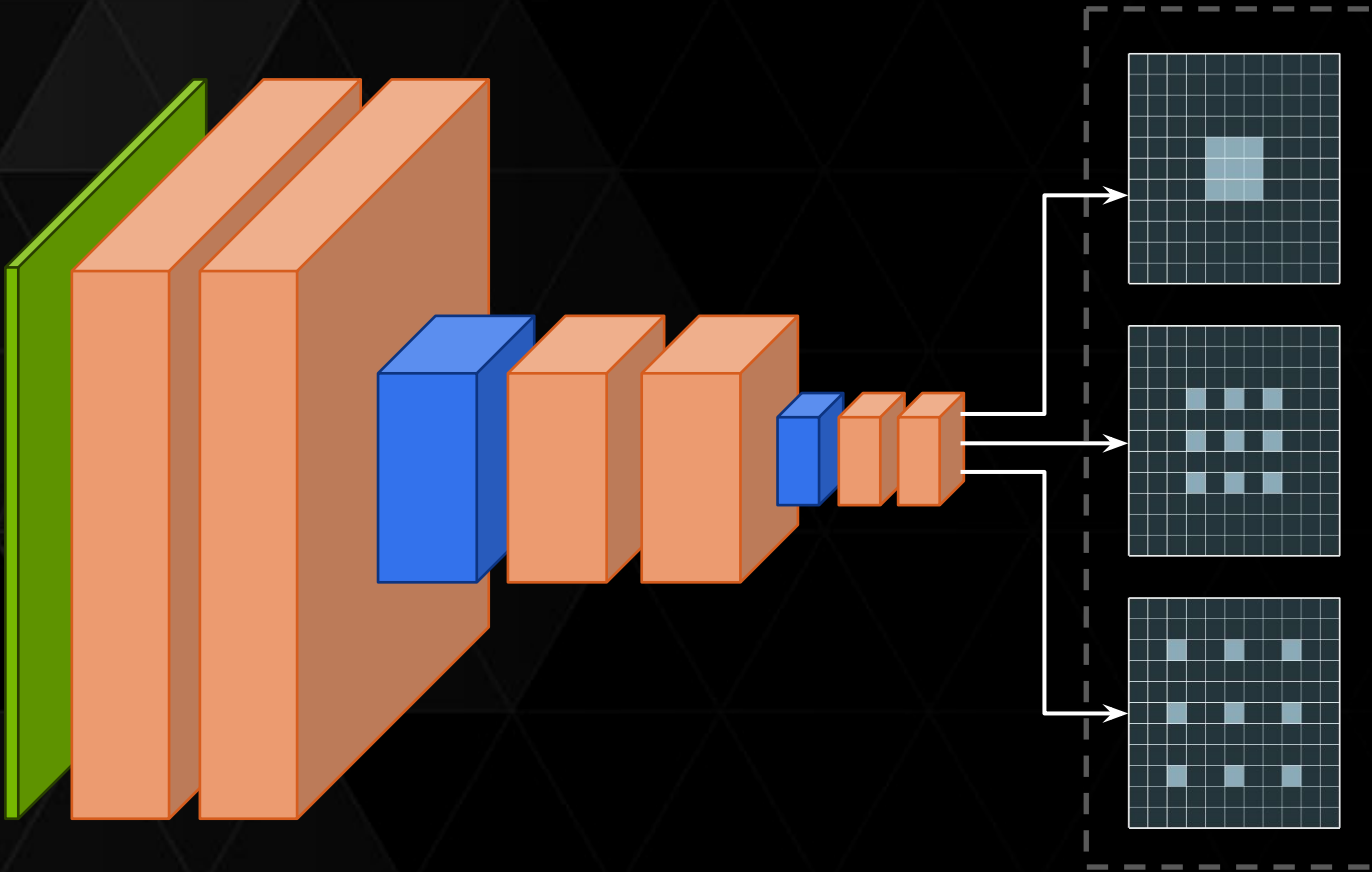
Encoder



ATROUS SPATIAL PYRAMID POOLING

Encoder

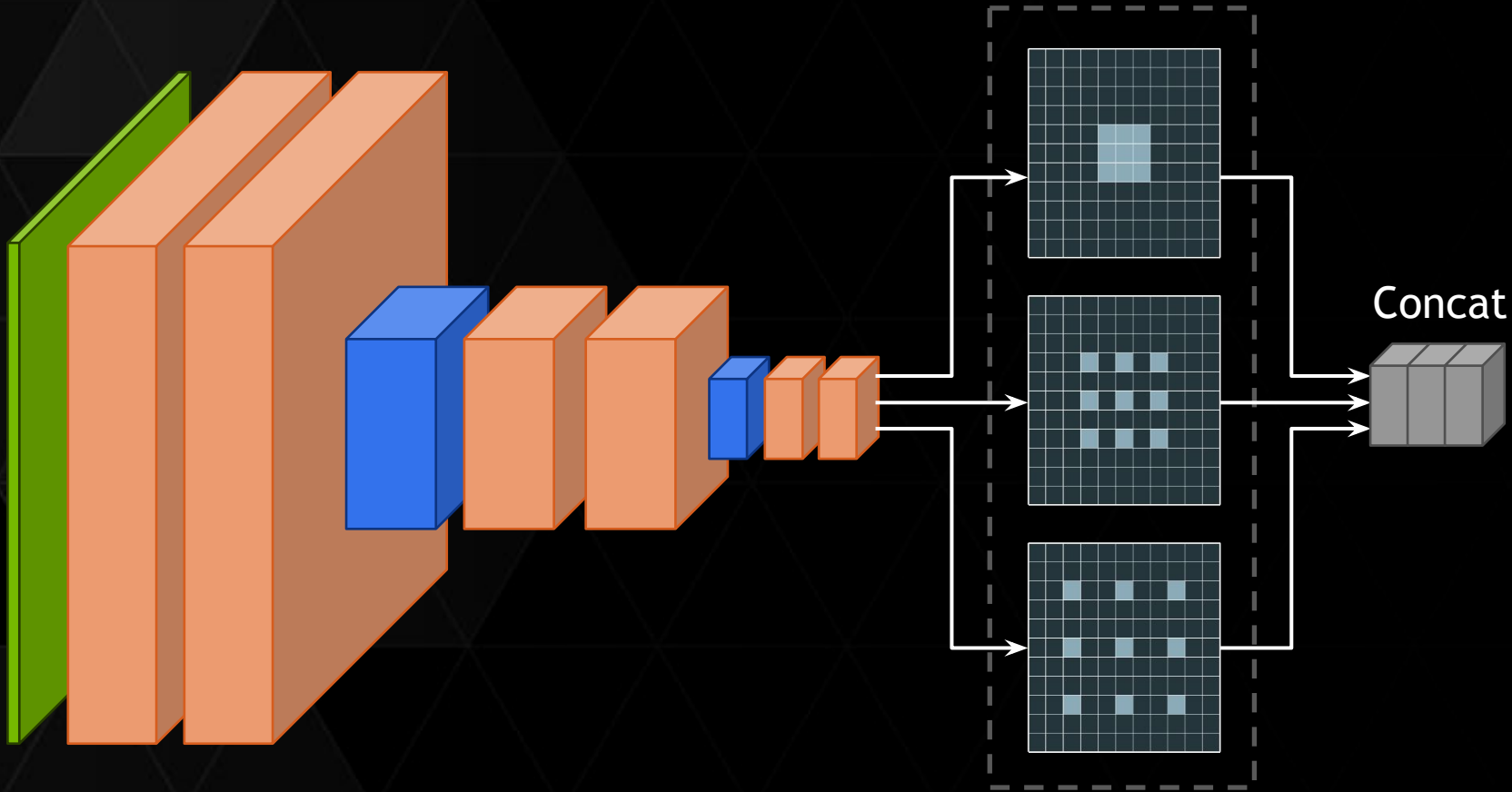
ASPP



ATROUS SPATIAL PYRAMID POOLING

Encoder

ASPP

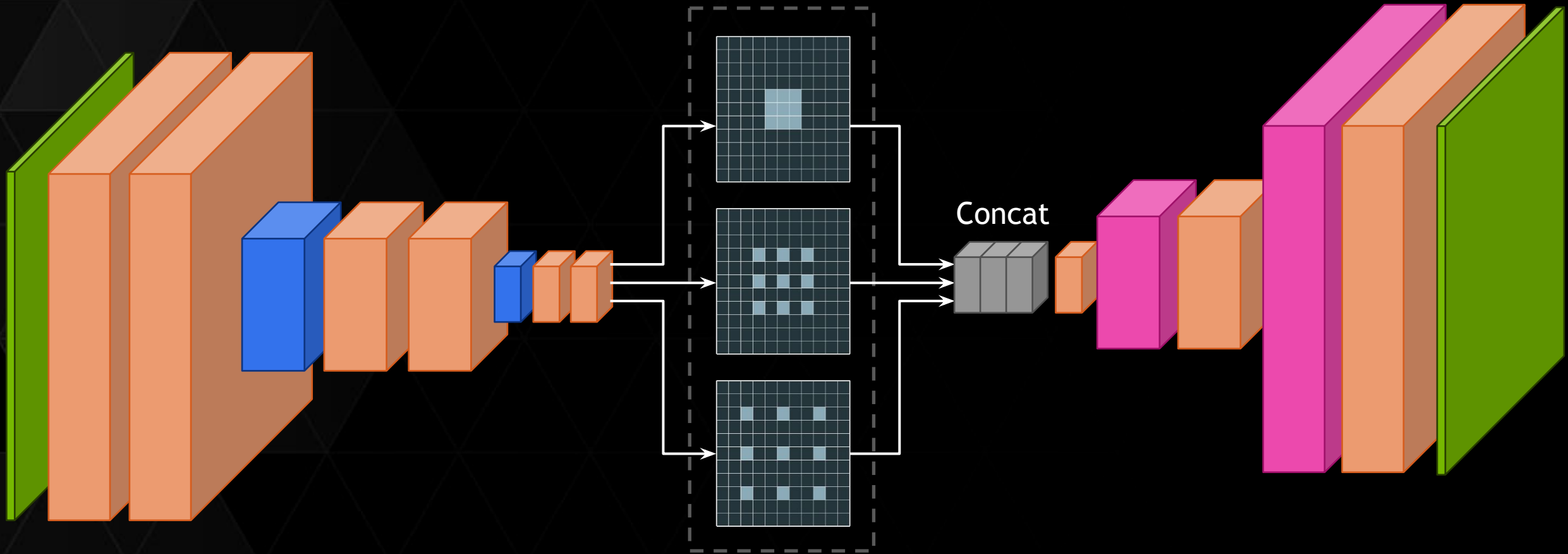


ATROUS SPATIAL PYRAMID POOLING

Encoder

ASPP

Decoder

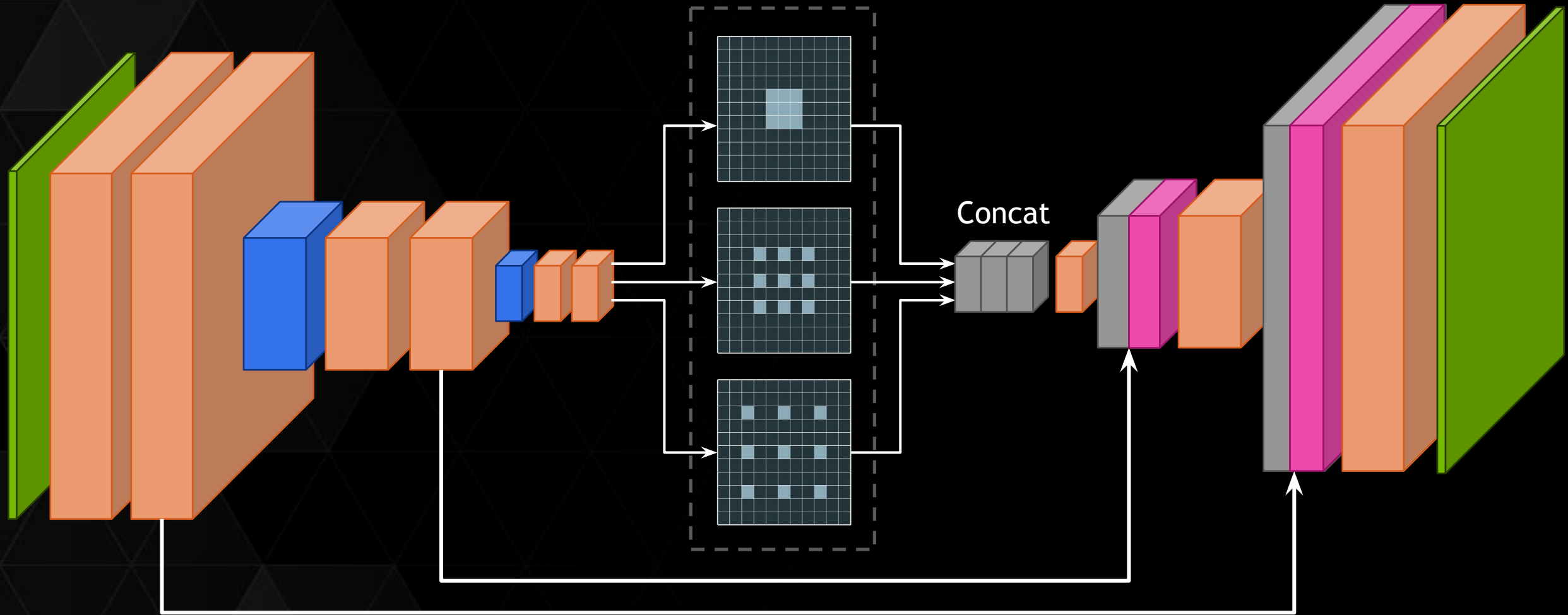


ATROUS SPATIAL PYRAMID POOLING

Encoder

ASPP

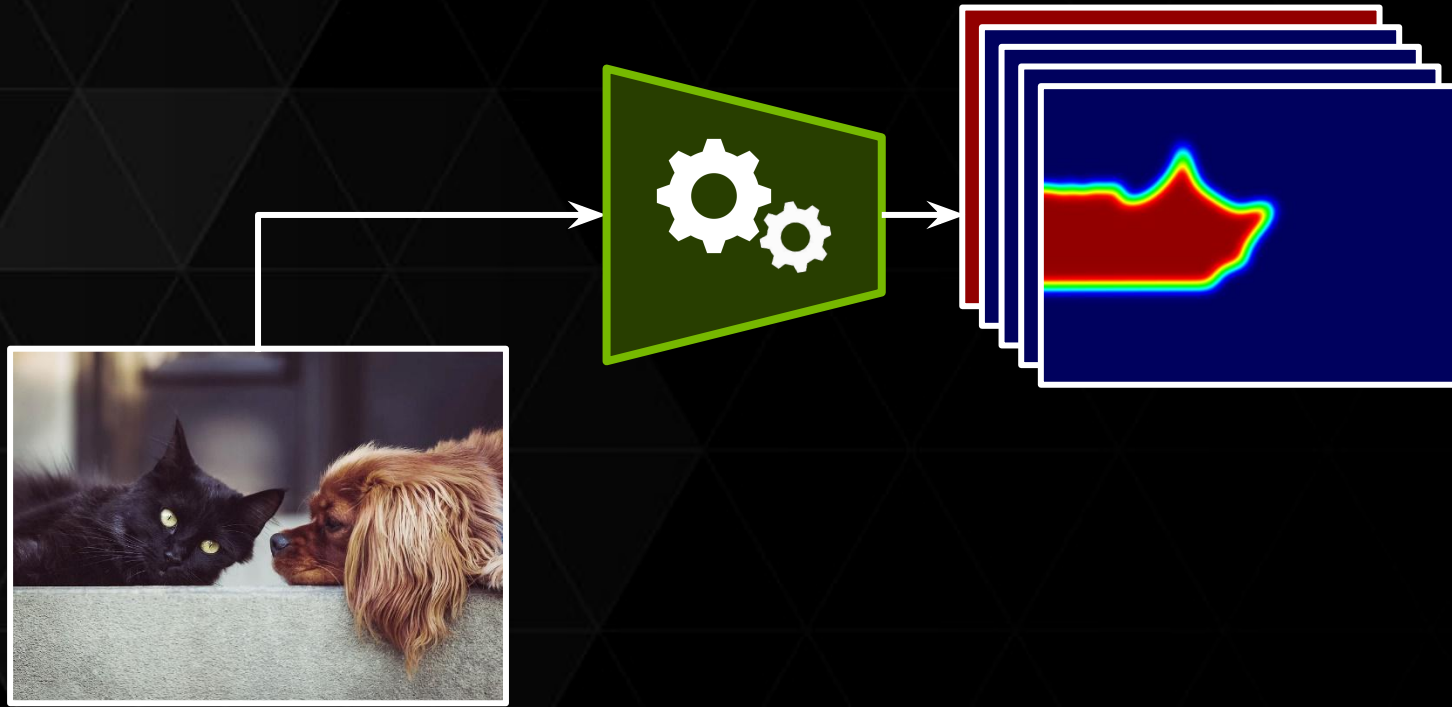
Decoder



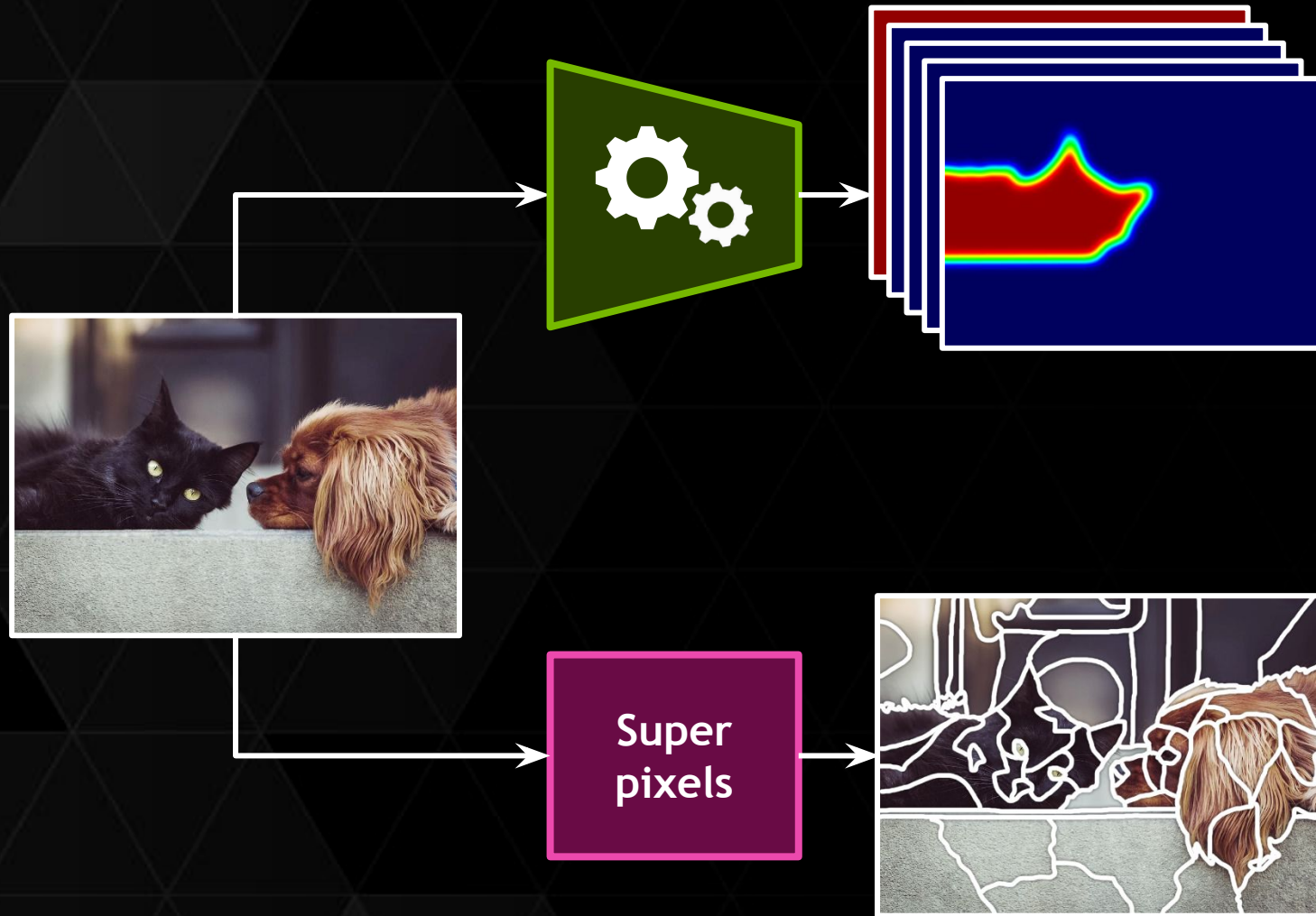
ПОСТОБРАБОТКА: DL + CRF



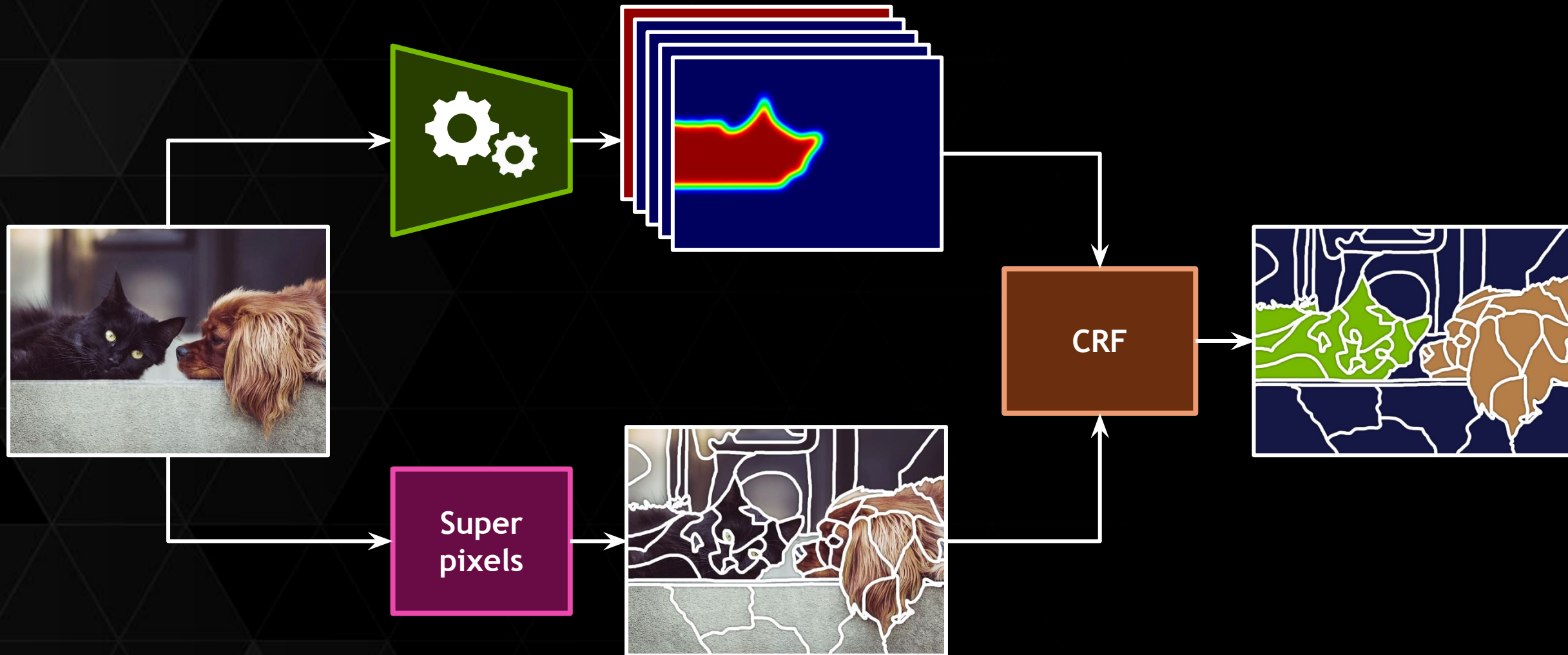
ПОСТОБРАБОТКА: DL + CRF



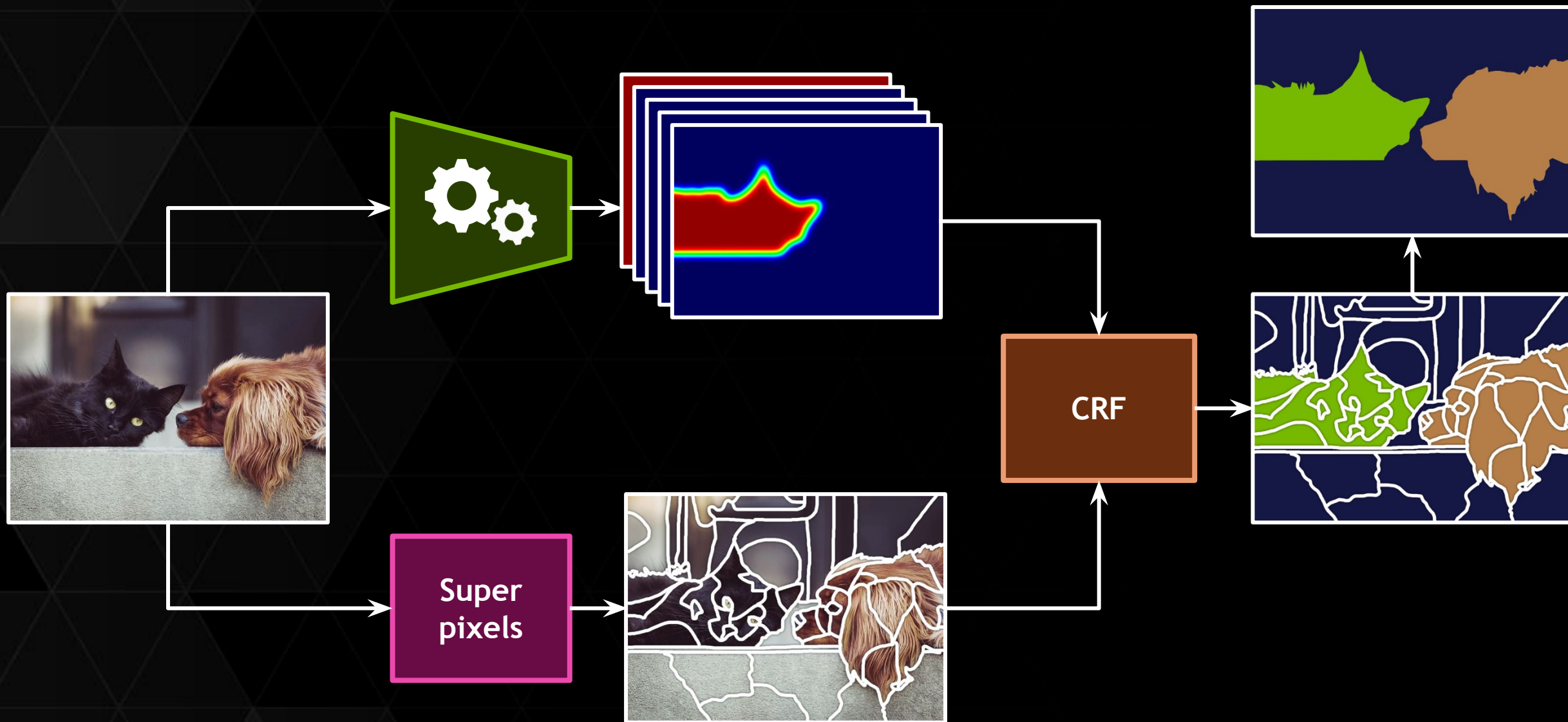
ПОСТОБРАБОТКА: DL + CRF



ПОСТОБРАБОТКА: DL + CRF



ПОСТОБРАБОТКА: DL + CRF



The background is a dark, almost black, field filled with a complex network of thin, light green lines. These lines intersect at various points, creating a web-like structure. At many of these intersection points, there are small, bright green circular dots or nodes. Some of these dots are slightly larger and more prominent than others. The overall effect is one of a dynamic, interconnected system, possibly representing a network or a data visualization. The text "СПАСИБО" is centered in the middle of this abstract pattern.

СПАСИБО