

Глубинное обучение

Практические советы. Обзор задач Computer Vision.

Ирина Сапарина

Как обучать нейросети?

Как обучать нейросети?

Основной принцип - от простого к сложному

Как обучать нейросети?

Основной принцип - от простого к сложному

- Для начала берите простой baseline
 - вы должны знать, какой результат ожидать
 - dummy solution/существующая обученная нейросеть

Как обучать нейросети?

Основной принцип - от простого к сложному

- Для начала берите простой baseline
- Разберитесь с чтением/обработкой данных

Как обучать нейросети?

Основной принцип - от простого к сложному

- Для начала берите простой baseline
- Разберитесь с чтением/обработкой данных
- Выберите метрики и loss

Как обучать нейросети?

Основной принцип - от простого к сложному

- Для начала берите простой baseline
- Разберитесь с чтением/обработкой данных
- Выберите метрики и loss
- Настройте логирование и визуализацию: Tensorboard, Weights&Biases

Как обучать нейросети?

Основной принцип - от простого к сложному

- Для начала берите простой baseline
- Разберитесь с чтением/обработкой данных
- Выберите метрики и loss
- Настройте логирование и визуализацию: Tensorboard, Weights&Biases
- Фиксируйте random seed!

Как обучать нейросети?

Основной принцип - от простого к сложному

- Для начала берите простой baseline
- Разберитесь с чтением/обработкой данных
- Выберите метрики и loss
- Настройте логирование и визуализацию: Tensorboard, Weights&Biases
- Фиксируйте random seed!
- Для начала пробуйте обучаться на простых данных (подвыборка/синтетика)

Можно проверить переобучение на один батч/на маленькую подборку

Как обучать нейросети?

Пользуйтесь гитом, сохраняйте код, модель, ее конфигурацию и результаты

Как обучать нейросети?

Пользуйтесь гитом, сохраняйте код, модель, ее конфигурацию и результаты
Создавайте отдельный environment

Как обучать нейросети?

Одно улучшение за раз!

Как обучать нейросети?

Одно улучшение за раз!

Проверяйте качество на отложенной выборке

Как обучать нейросети?

Одно улучшение за раз!

- Регуляризация при переобучении

Как обучать нейросети?

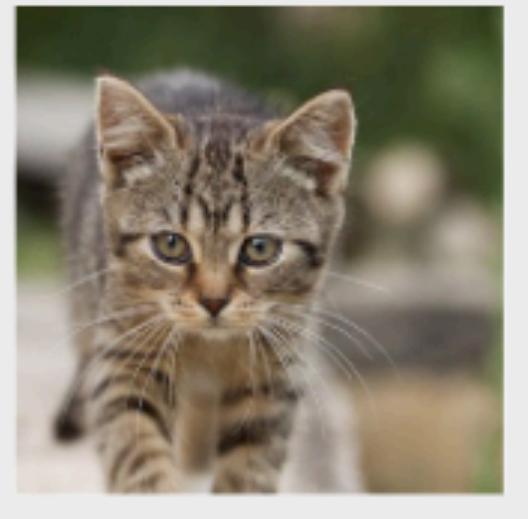
Одно улучшение за раз!

- Регуляризация при переобучении
- Аугментации

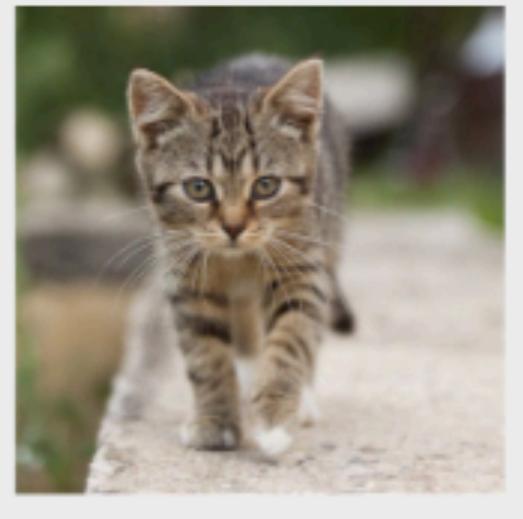
Как обучать нейросети?

Одно улучшение за раз!

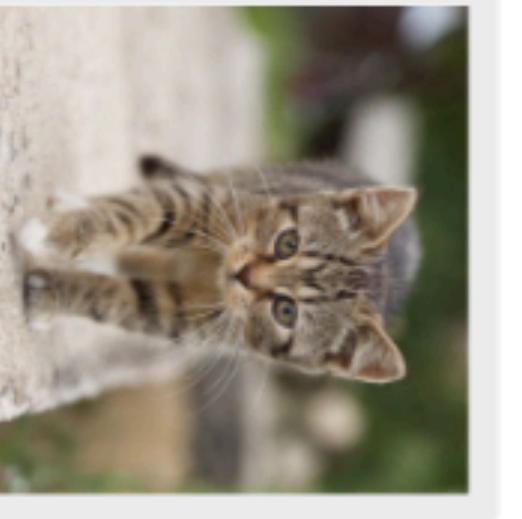
- Регуляризация при переобучении
- Аугментации
 - Random flip/rotate/crop/rescale
 - Изменение цвета
 - Добавление шума



Crop



Symmetry



Rotation



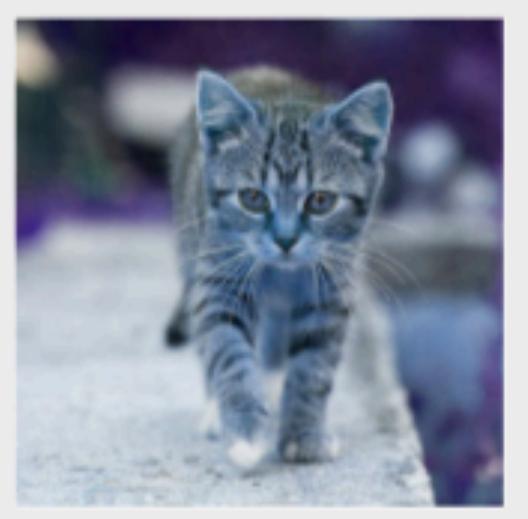
Scale



Original



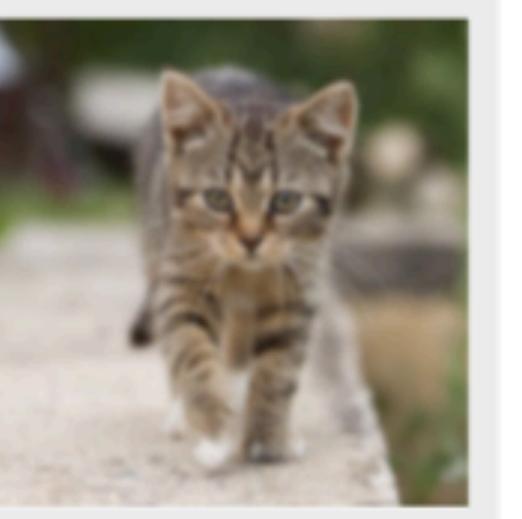
Noise



Hue



Obstruction



Blur

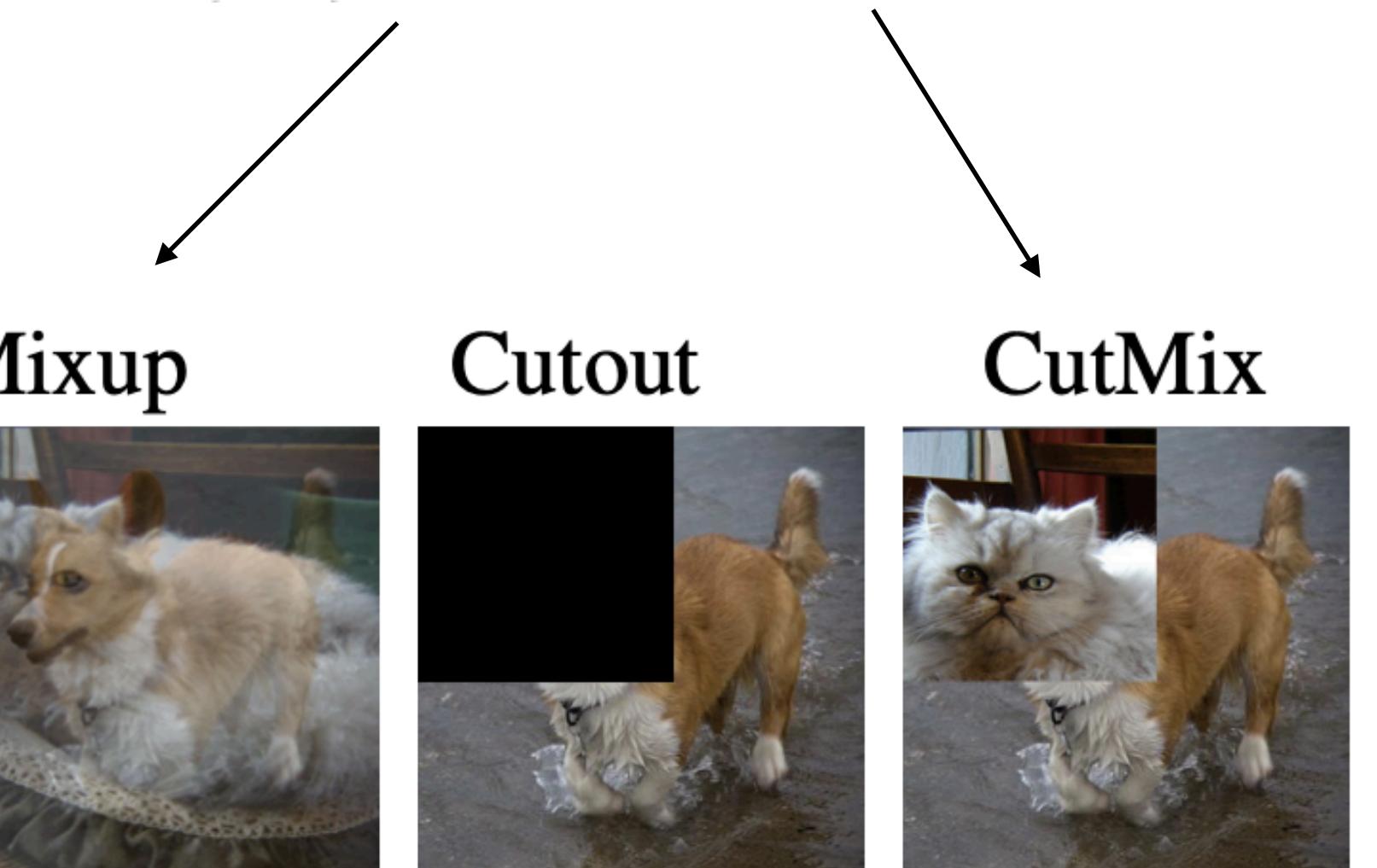
Как обучать нейросети?

Одно улучшение за раз!

- Регуляризация при переобучении
- Аугментации
 - MixUp
 - CutOut
 - CutMix

$$\hat{y} = \lambda y_i + (1 - \lambda)y_j,$$

where $\lambda \in [0, 1]$ is a random number



	ResNet-50	Mixup	Cutout	CutMix
Image				
Label	Dog 1.0	Dog 0.5 Cat 0.5	Dog 1.0	Dog 0.6 Cat 0.4
ImageNet	76.3	77.4	77.1	78.6
Cls (%)	(+0.0)	(+1.1)	(+0.8)	(+2.3)

Image credit

Как обучать нейросети?

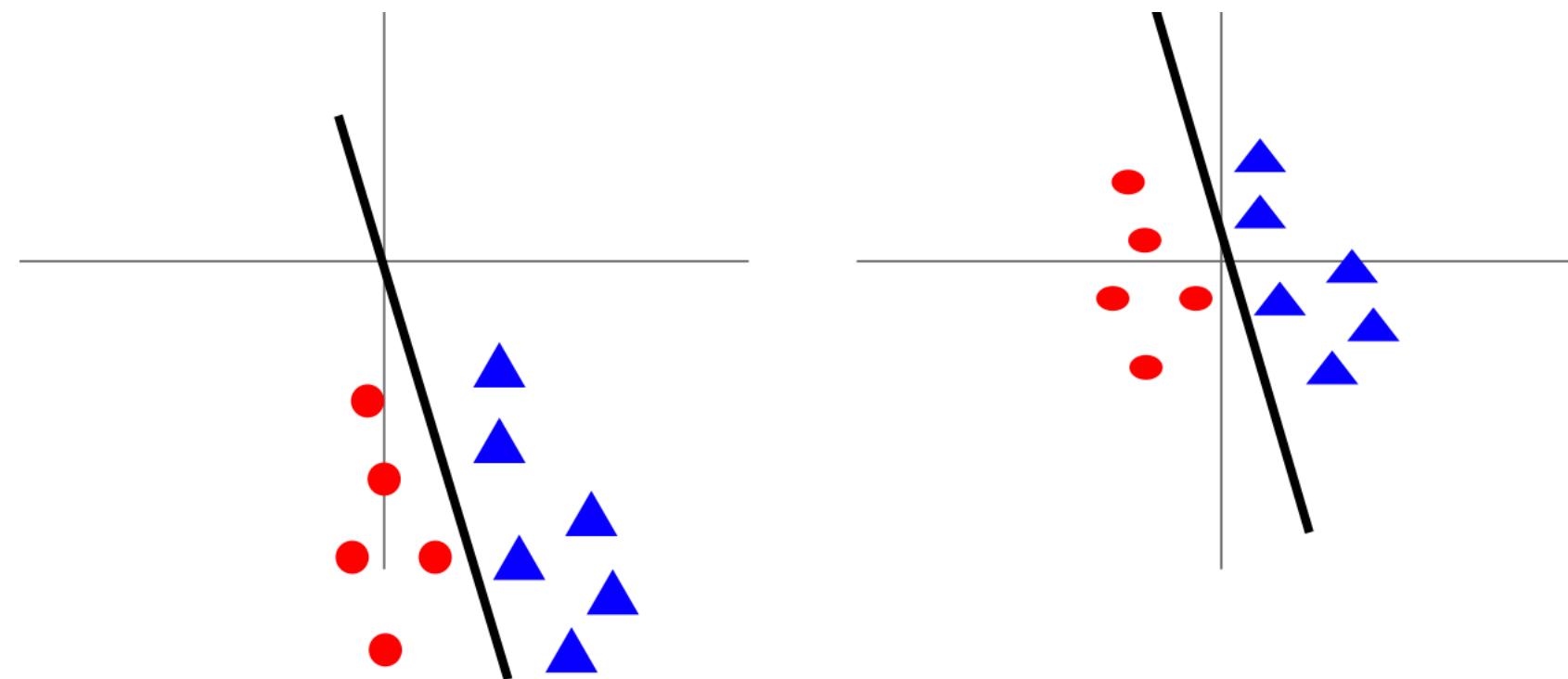
Одно улучшение за раз!

- Регуляризация при переобучении
- Аугментации
- Качественная предобработка

Как обучать нейросети?

Одно улучшение за раз!

- Регуляризация при переобучении
- Аугментации
- Качественная предобработка



Изображения можно нормализовать:

- Вычесть среднее изображение ($H \times W \times C$)
- Вычесть среднее по каналам (C чисел)
- Разделить на std по каналам (C чисел)

Как обучать нейросети?

Одно улучшение за раз!

- Регуляризация при переобучении
- Аугментации
- Качественная предобработка

Изображения можно нормализовать:

- | | | |
|---------|---|---|
| AlexNet | → | • Вычесть среднее изображение ($H \times W \times C$) |
| VGGNet | → | • Вычесть среднее по каналам (C чисел) |
| ResNet | ↖ | • Разделить на std по каналам (C чисел) |

Как обучать нейросети?

Одно улучшение за раз!

- Регуляризация при переобучении
- Аугментации
- Качественная предобработка
- Loss-функция должна соответствовать метрике

Как обучать нейросети?

Одно улучшение за раз!

- Регуляризация при переобучении
- Аугментации
- Качественная предобработка
- Loss-функция должна соответствовать метрике
- Подбор оптимизаторов, lr scheduler, подбор гиперпараметров

Как обучать нейросети?

Одно улучшение за раз!

- Регуляризация при переобучении
- Аугментации
- Качественная предобработка
- Loss-функция должна соответствовать метрике
- Подбор оптимизаторов, lr scheduler, подбор гиперпараметров

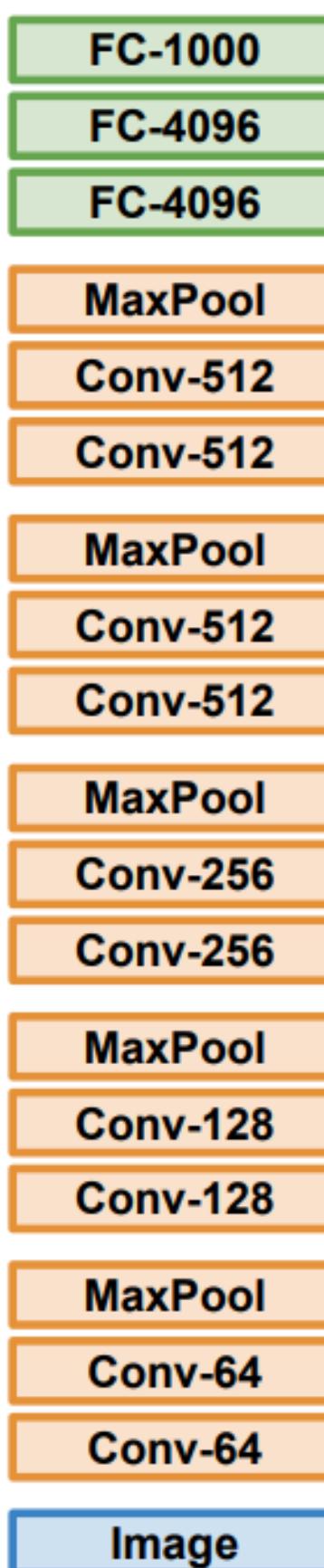
Используйте известные модели и их существующие реализации!

- классические модели для CV - в torchvision

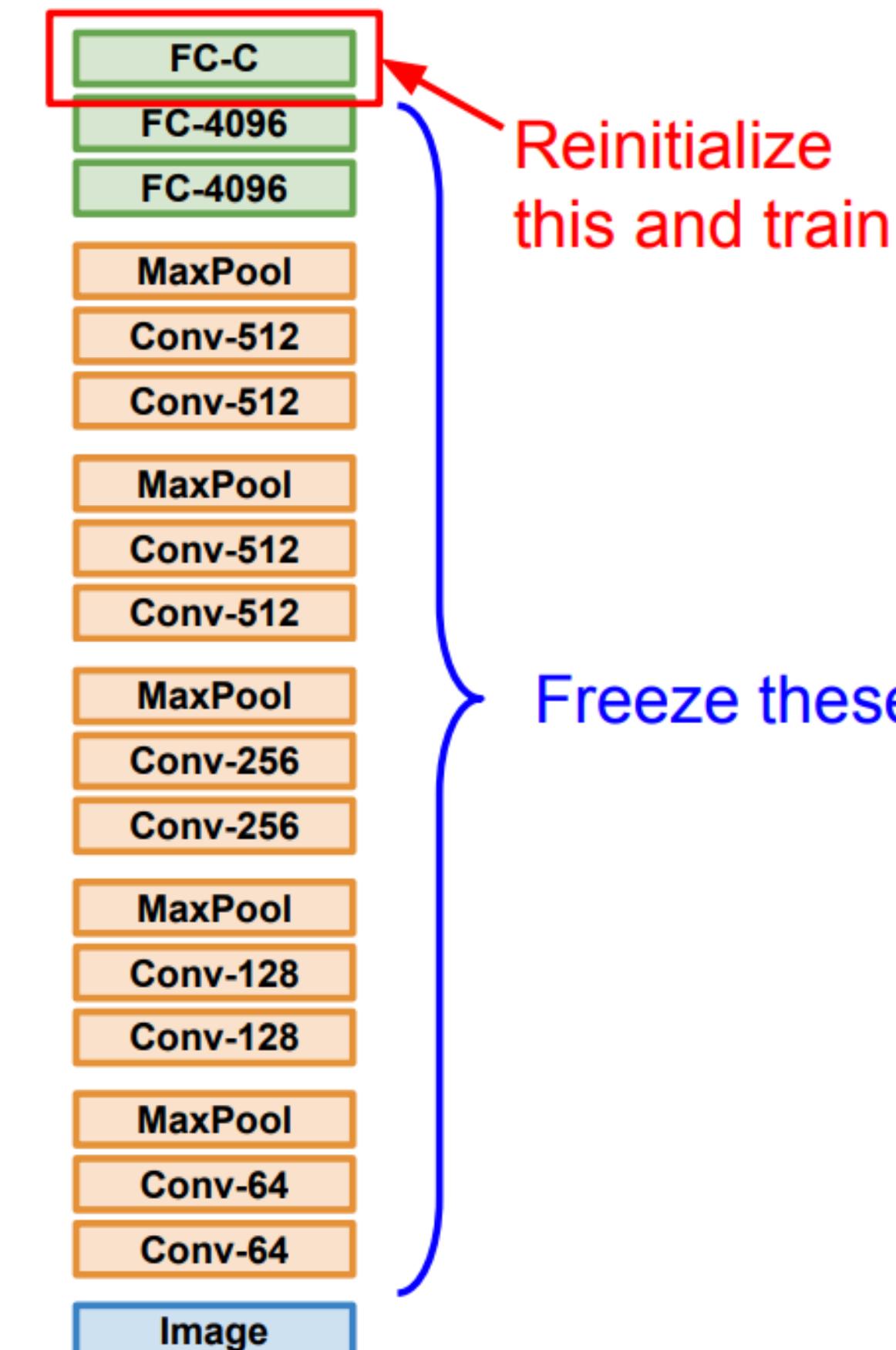
Как обучать нейросети?

Transfer Learning

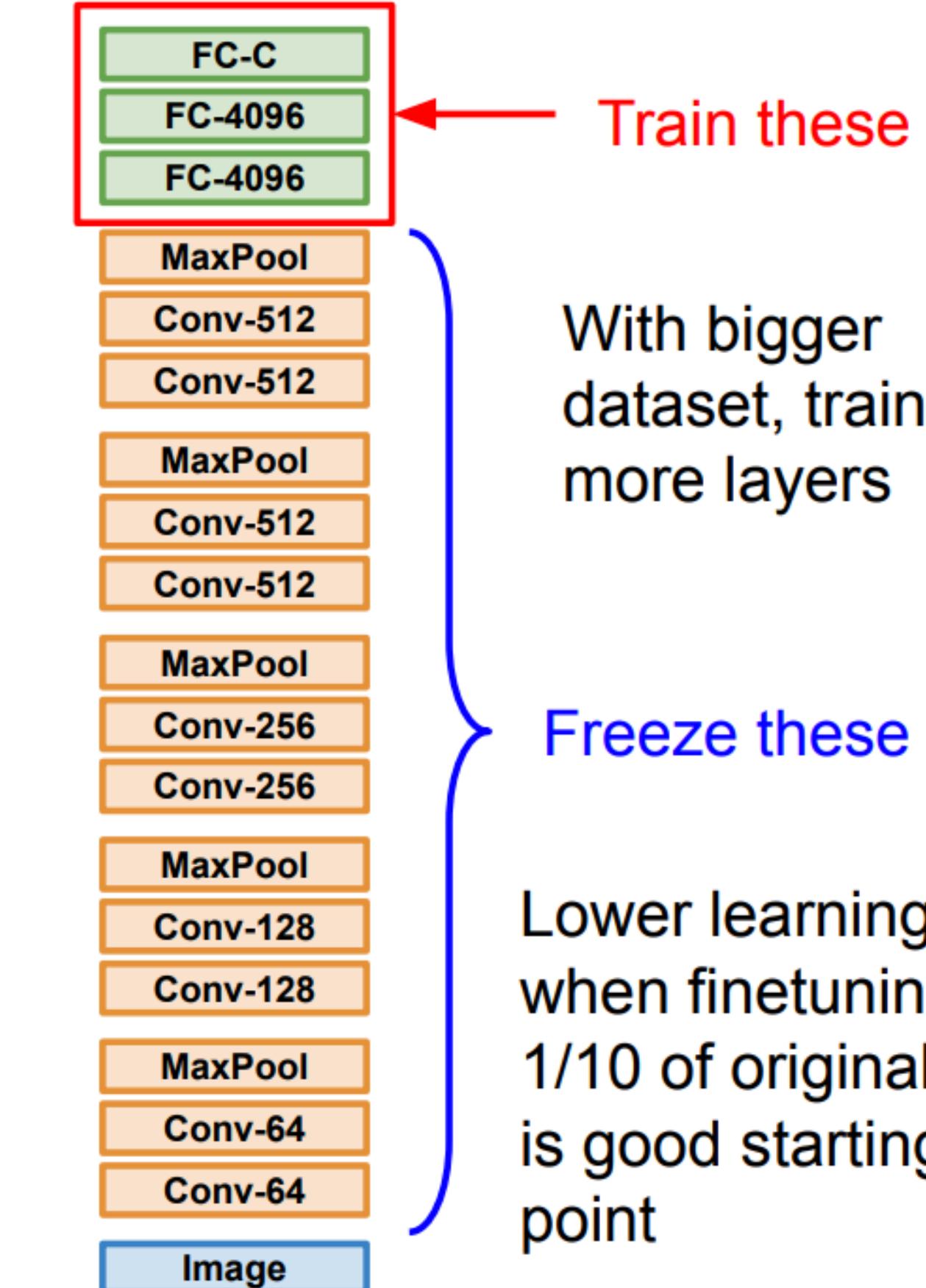
1. Train on Imagenet



2. Small Dataset (C classes)



3. Bigger dataset



Как обучать нейросети?

Transfer Learning

- Быстрее, чем обучение с 0
- Не нужен большой датасет для обучения, много ресурсов
- **Используйте предобученную нейросеть всегда, когда можете**

Обзор задач Computer Vision

Задачи Computer Vision

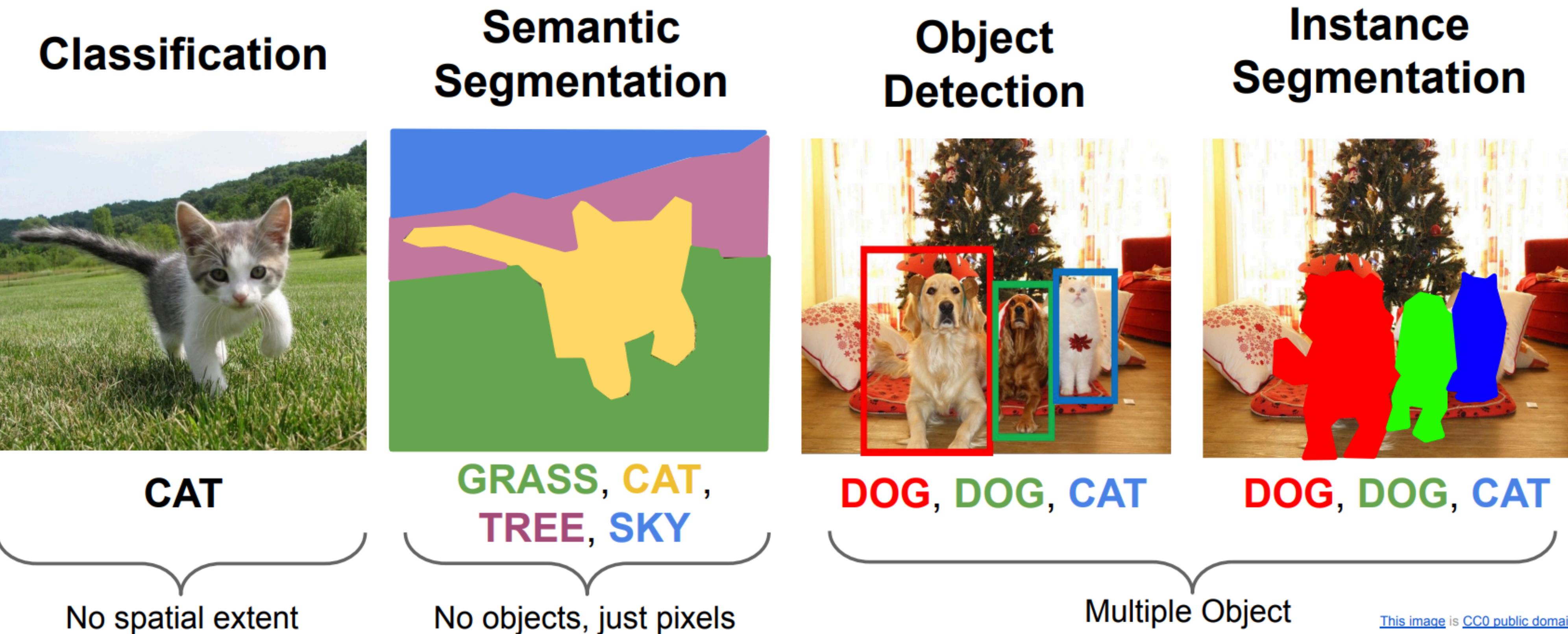
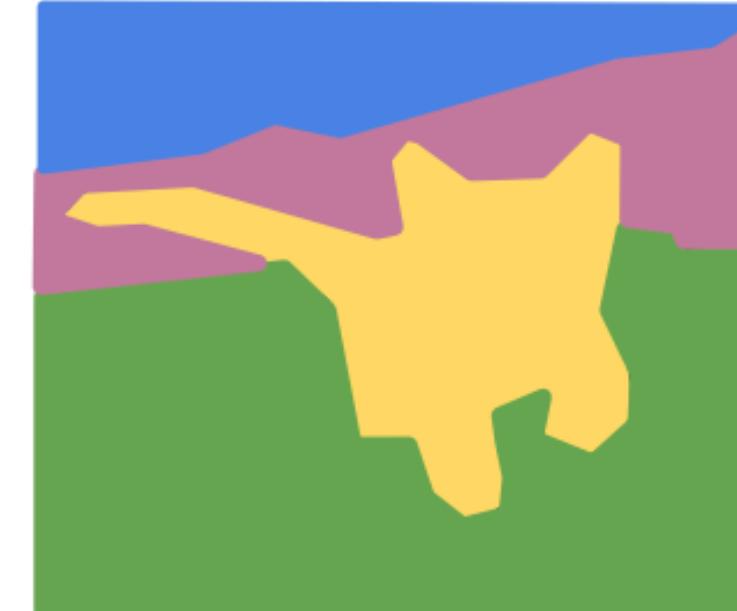


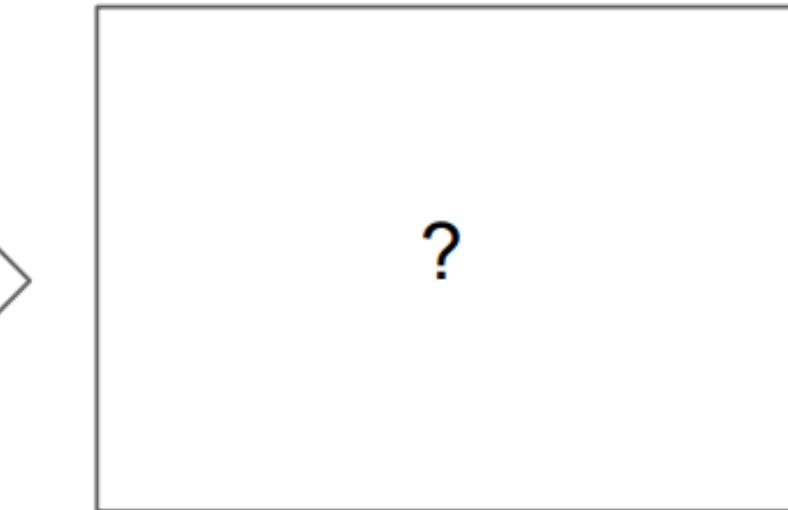
Image credit

Semantic Segmentation

Задача: для каждого пикселя изображения определить его категорию



GRASS, CAT,
TREE, SKY, ...



At test time, classify each pixel of a new image.

[Image credit](#)

Semantic Segmentation

Задача: для каждого пикселя изображения определить его категорию

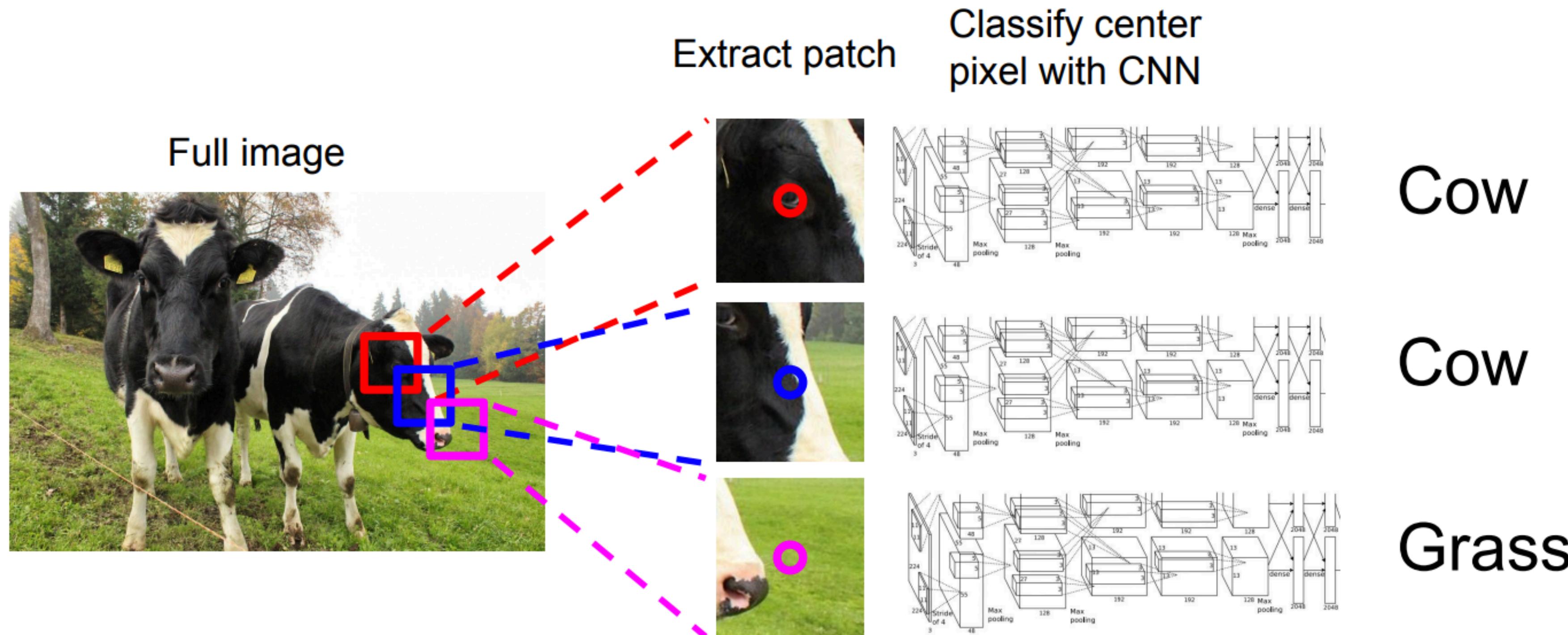
Определить категорию отдельного пикселя без контекста невозможно



[Image credit](#)

Semantic Segmentation

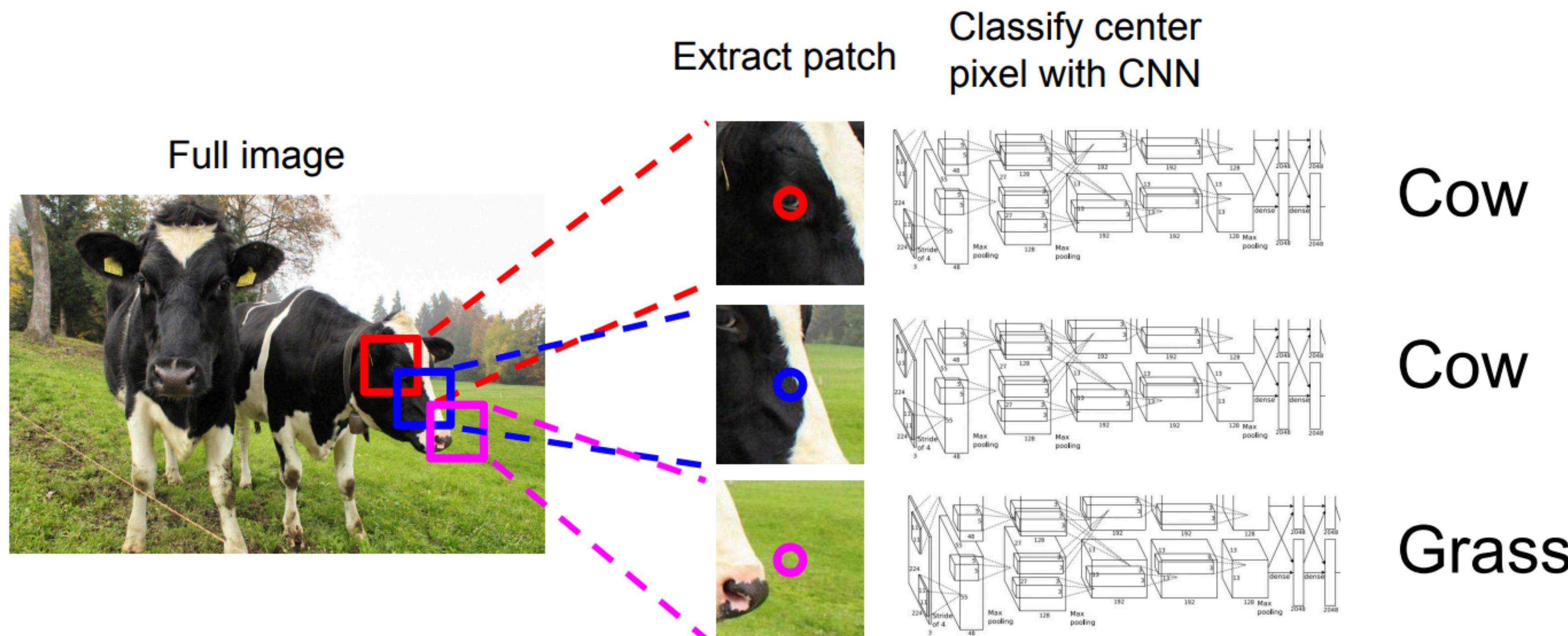
Идея: перемещаться окном по изображению, классифицировать центр



Farabet et al, "Learning Hierarchical Features for Scene Labeling," TPAMI 2013
Pinheiro and Collobert, "Recurrent Convolutional Neural Networks for Scene Labeling", ICML 2014

Semantic Segmentation

Неэфективно!

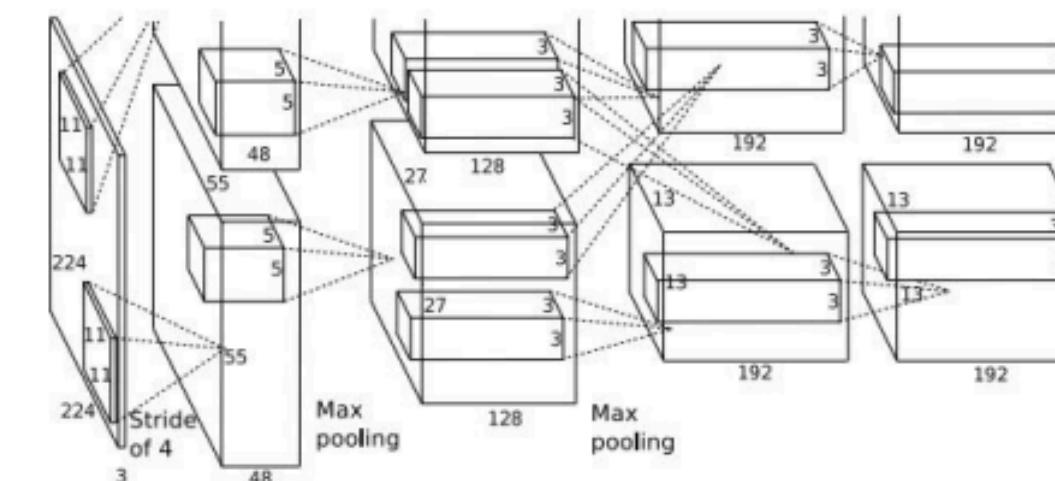


Farabet et al, "Learning Hierarchical Features for Scene Labeling," TPAMI 2013
Pinheiro and Collobert, "Recurrent Convolutional Neural Networks for Scene Labeling", ICML 2014

Semantic Segmentation

Идея: изображение целиком - в CNN (как при классификации), поверх - сегментация

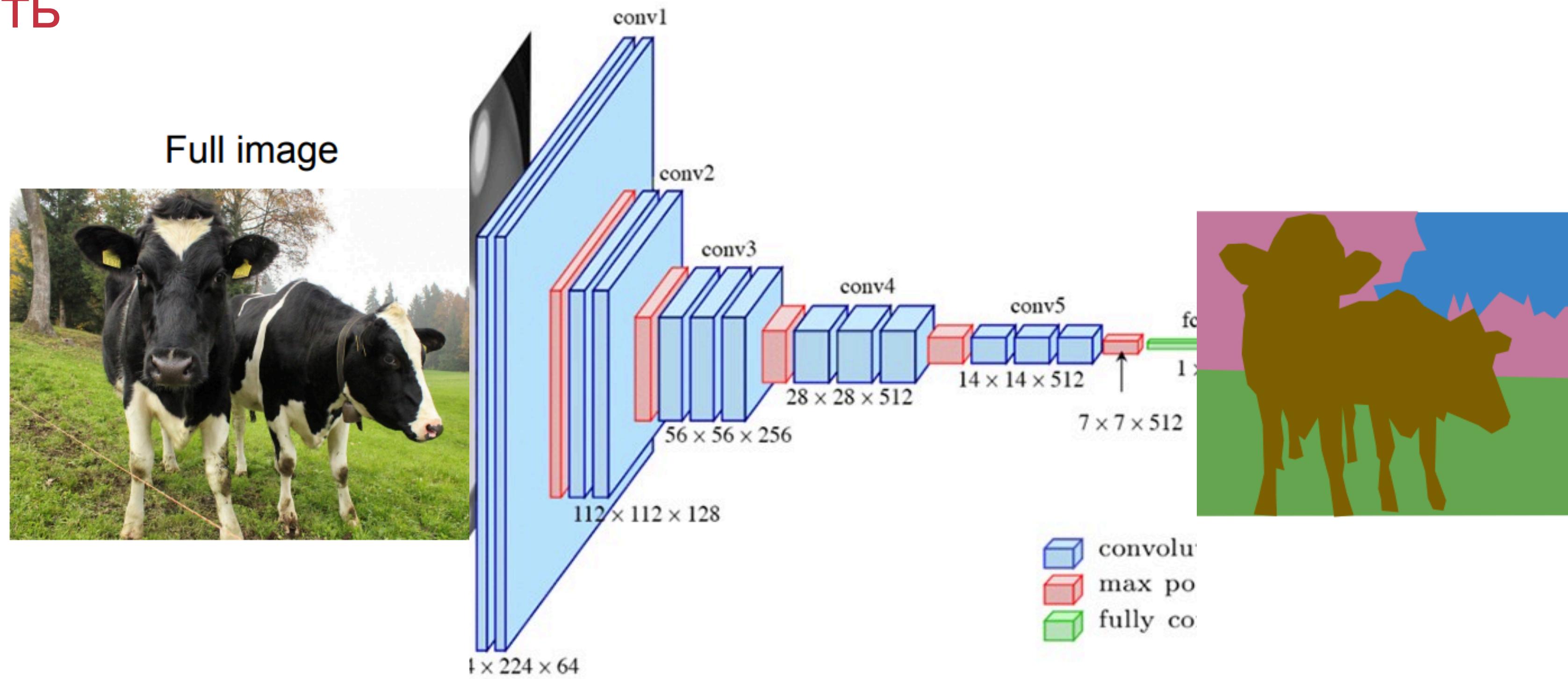
Full image



Semantic Segmentation

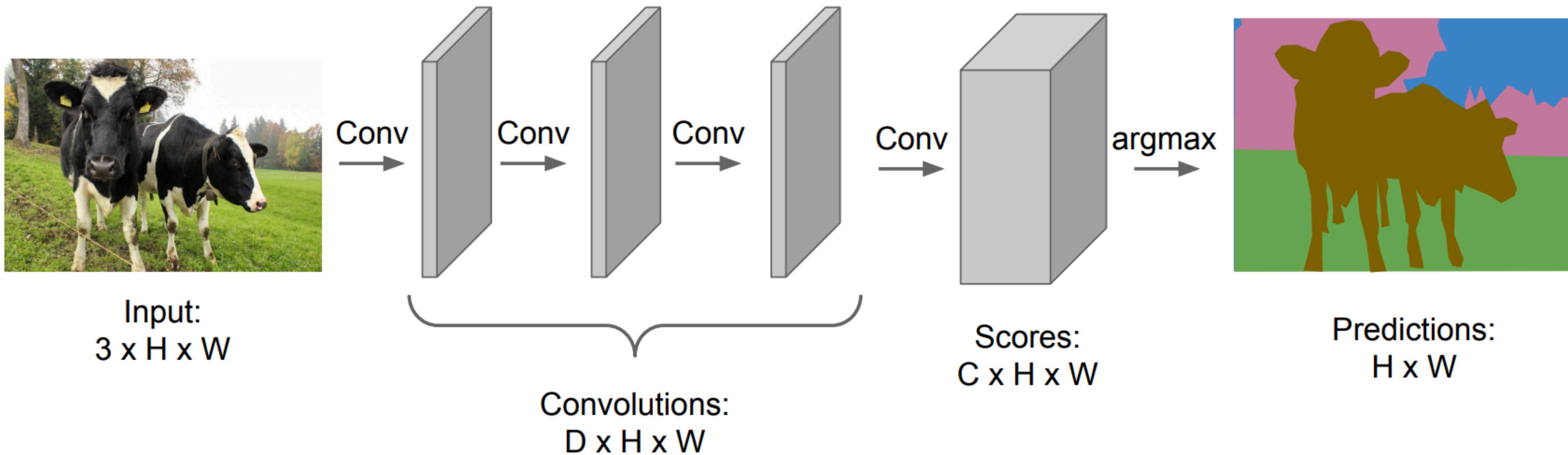
Идея: изображение целиком - в CNN (как при классификации), поверх - сегментация

Проблема: CNN для классификации уменьшают H и W, нам же та же нужна размерность



Semantic Segmentation

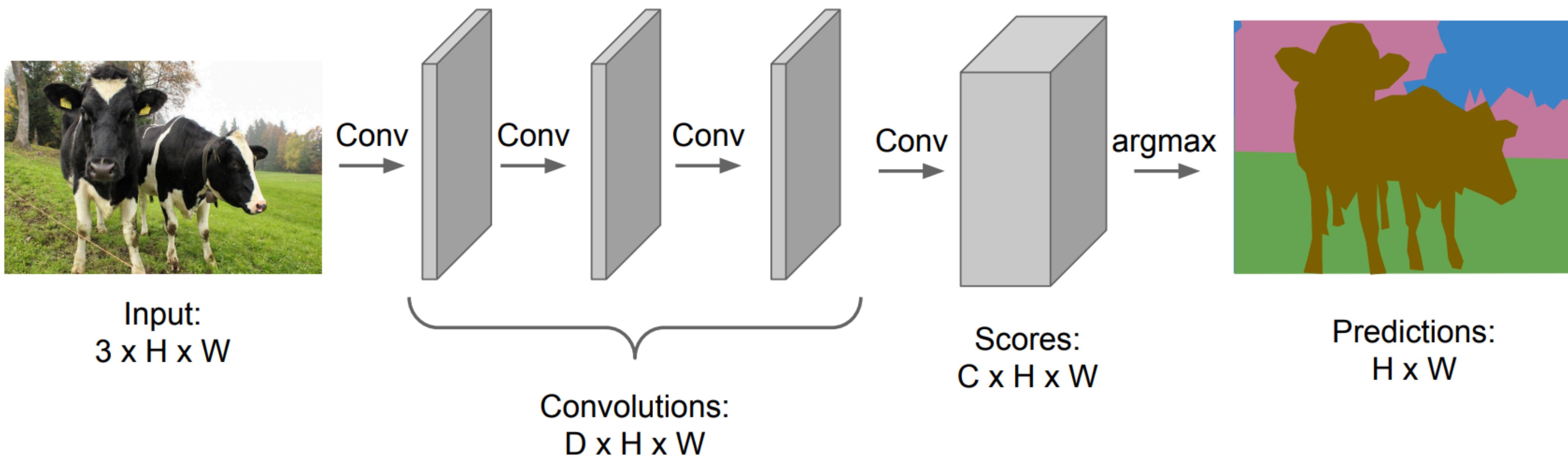
Идея: Fully Convolutional NN - нет pooling операторов, размер остается тот же



Semantic Segmentation

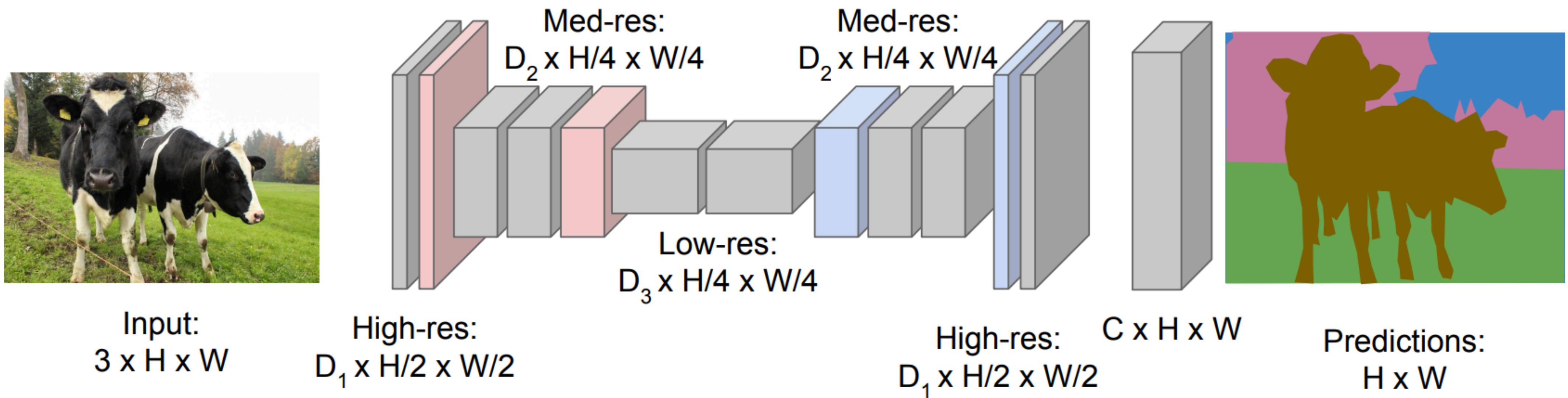
Идея: Fully Convolutional NN - нет pooling операторов, размер остается тот же

Проблема: много операций



Semantic Segmentation

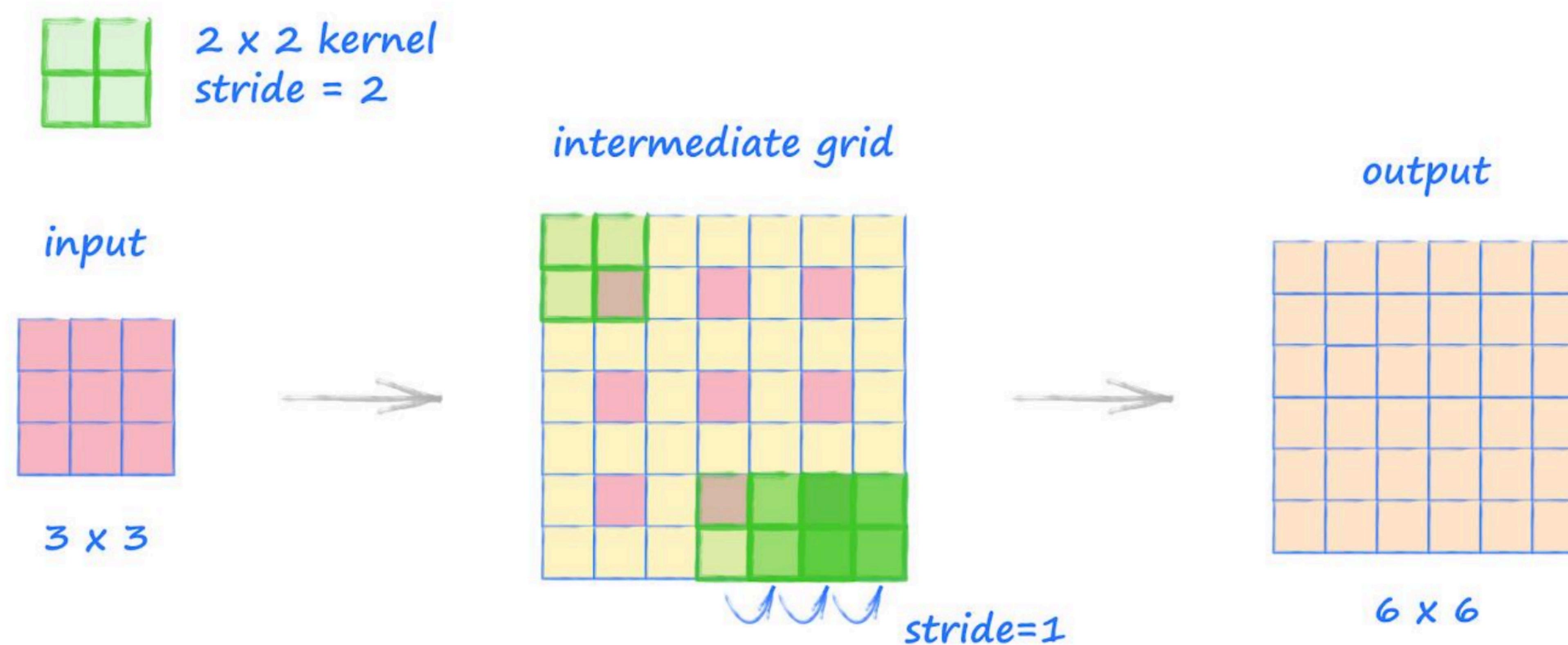
Идея: U-net - downsampling & upsampling



Long, Shelhamer, and Darrell, "Fully Convolutional Networks for Semantic Segmentation", CVPR 2015
Noh et al, "Learning Deconvolution Network for Semantic Segmentation", ICCV 2015

Semantic Segmentation

ConvTranspose, stride=2, no padding



Object Detection

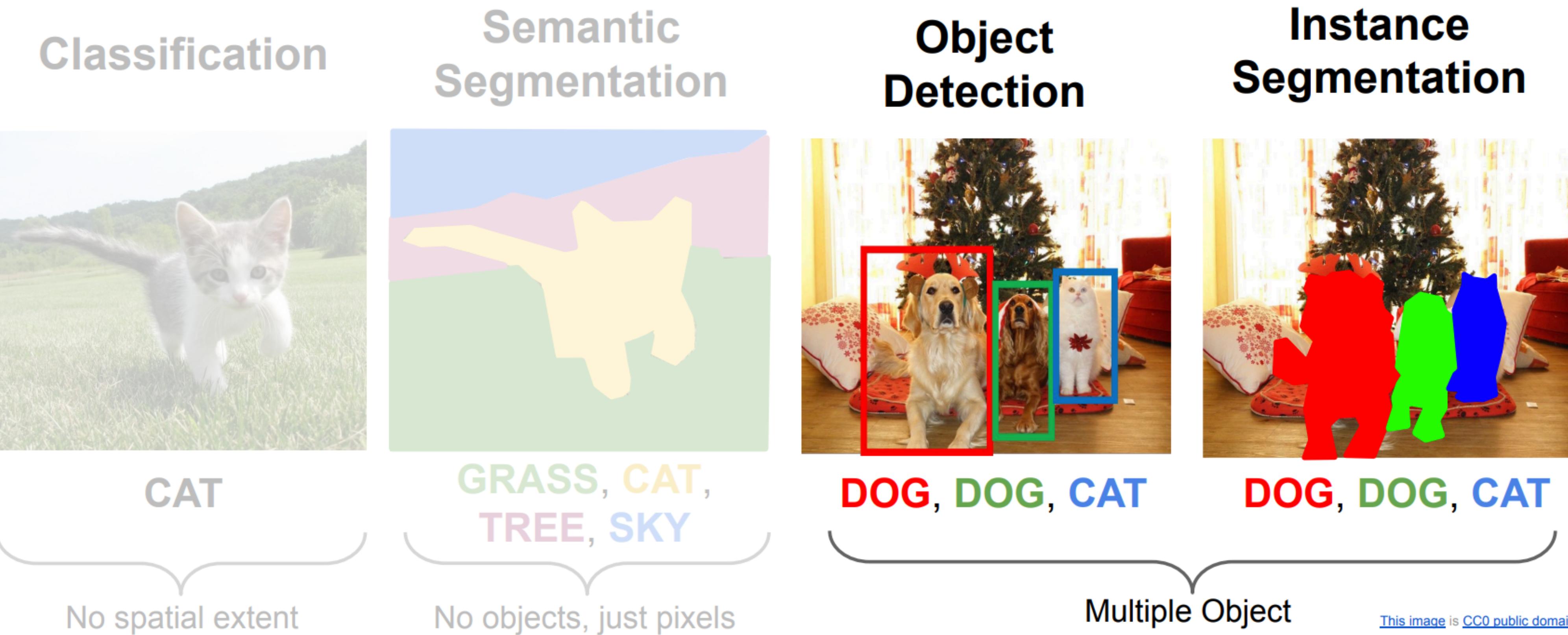
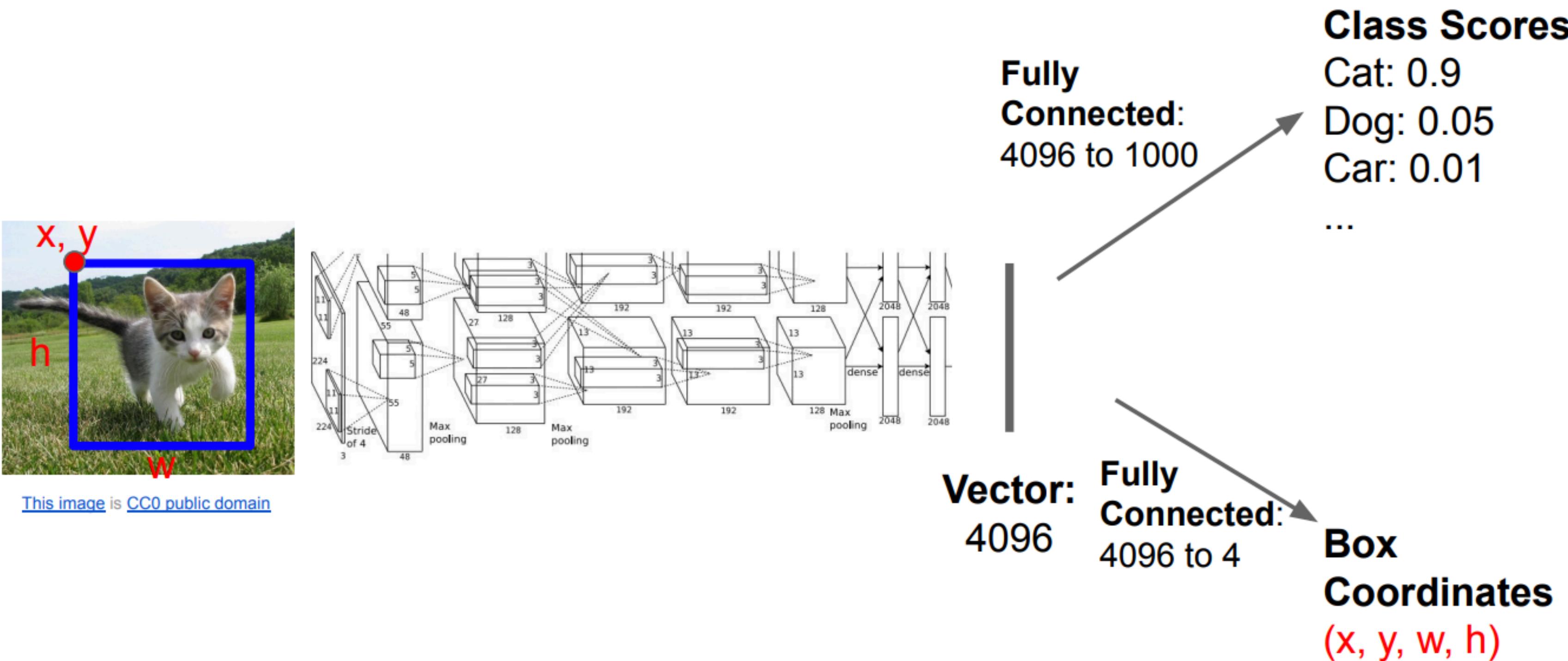


Image credit

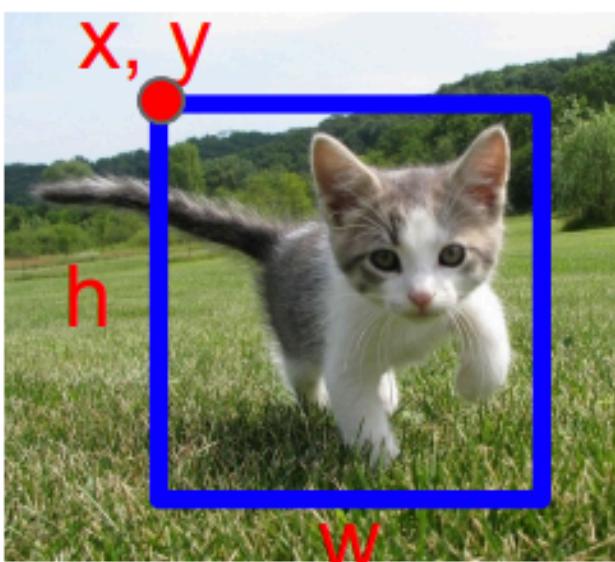
Object Detection

Один объект - классификация и локализация

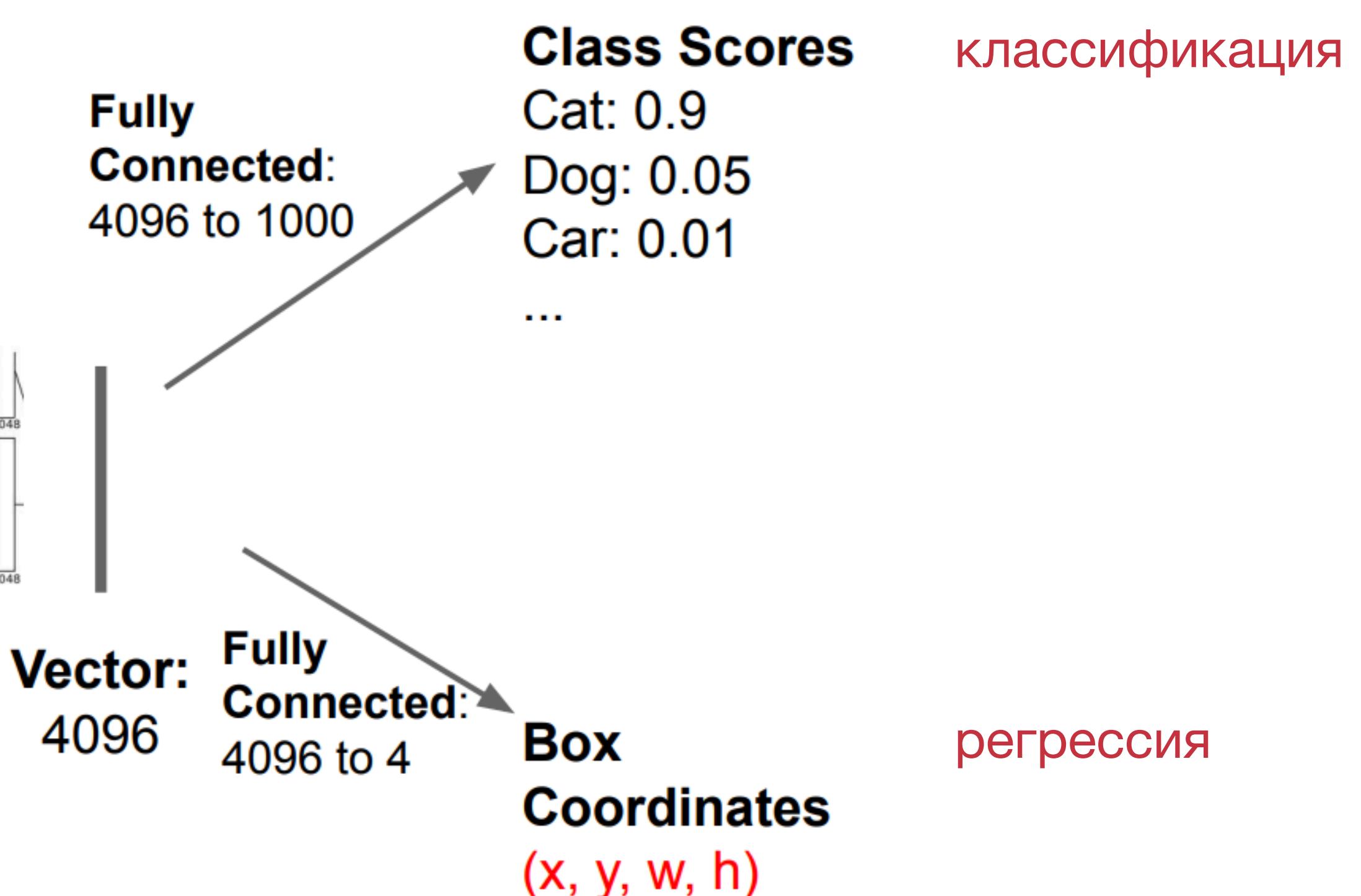
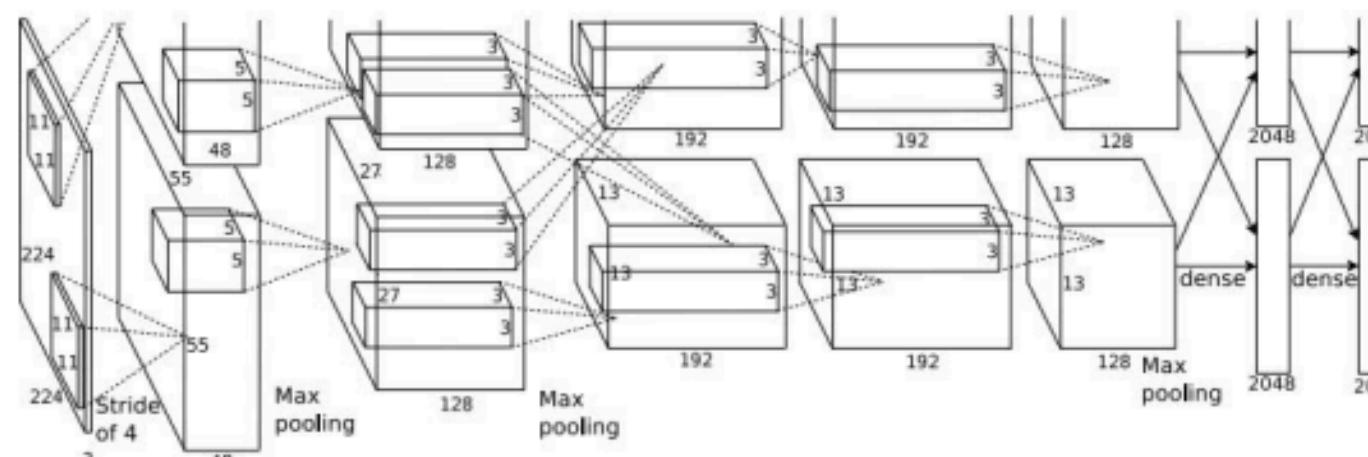


Object Detection

Один объект - классификация и локализация

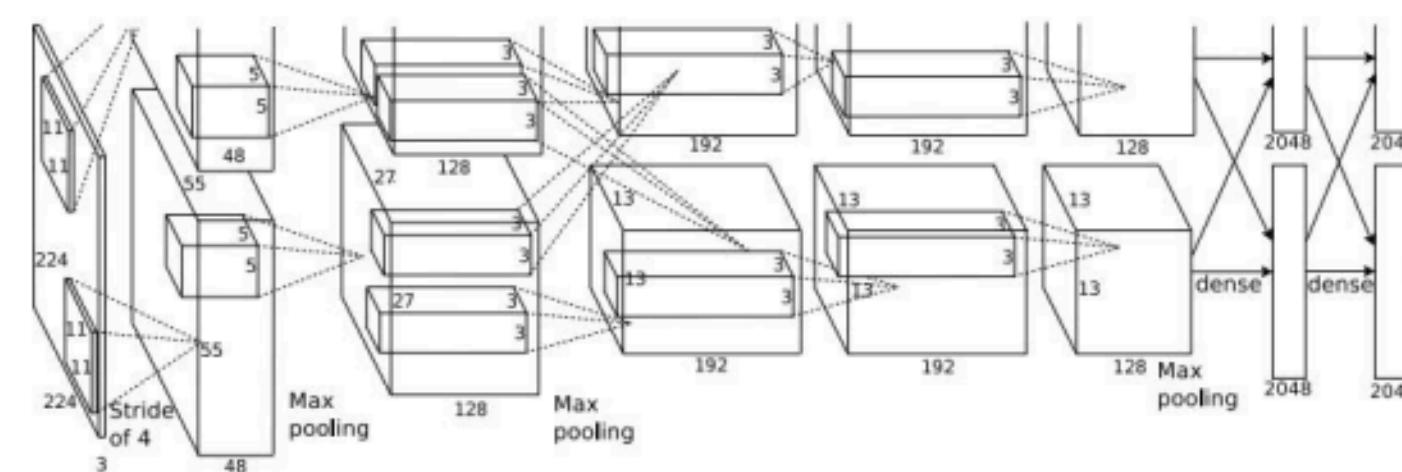
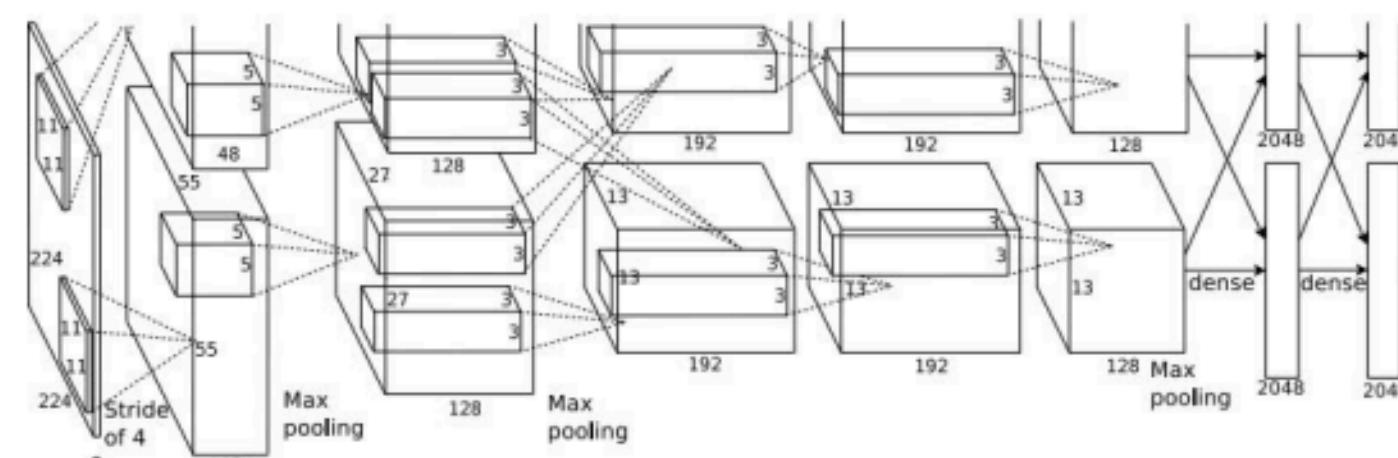
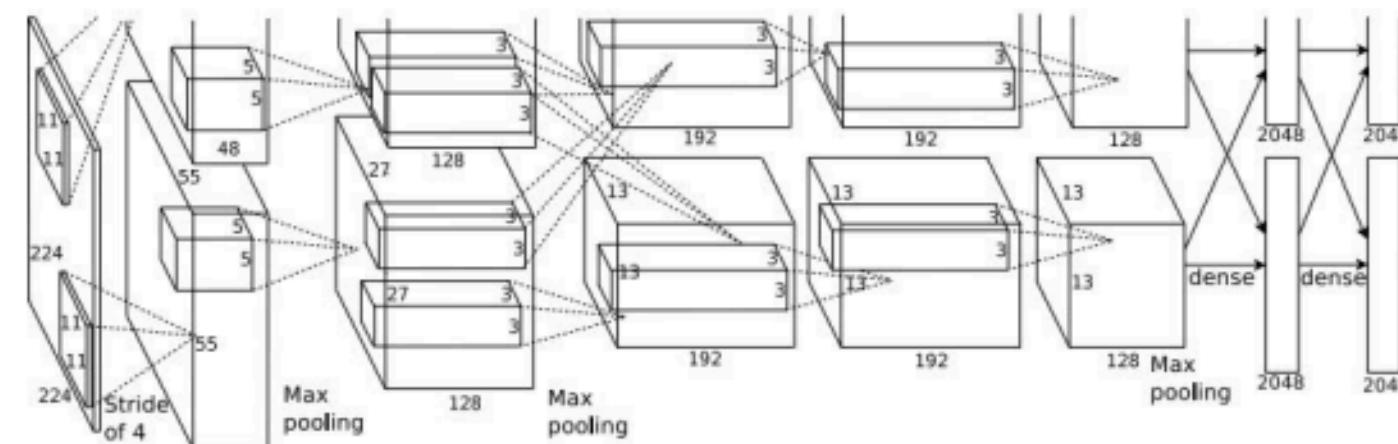


This image is CC0 public domain



Object Detection

Несколько объектов



CAT: (x, y, w, h)

DOG: (x, y, w, h)

DOG: (x, y, w, h)

CAT: (x, y, w, h)

DUCK: (x, y, w, h)

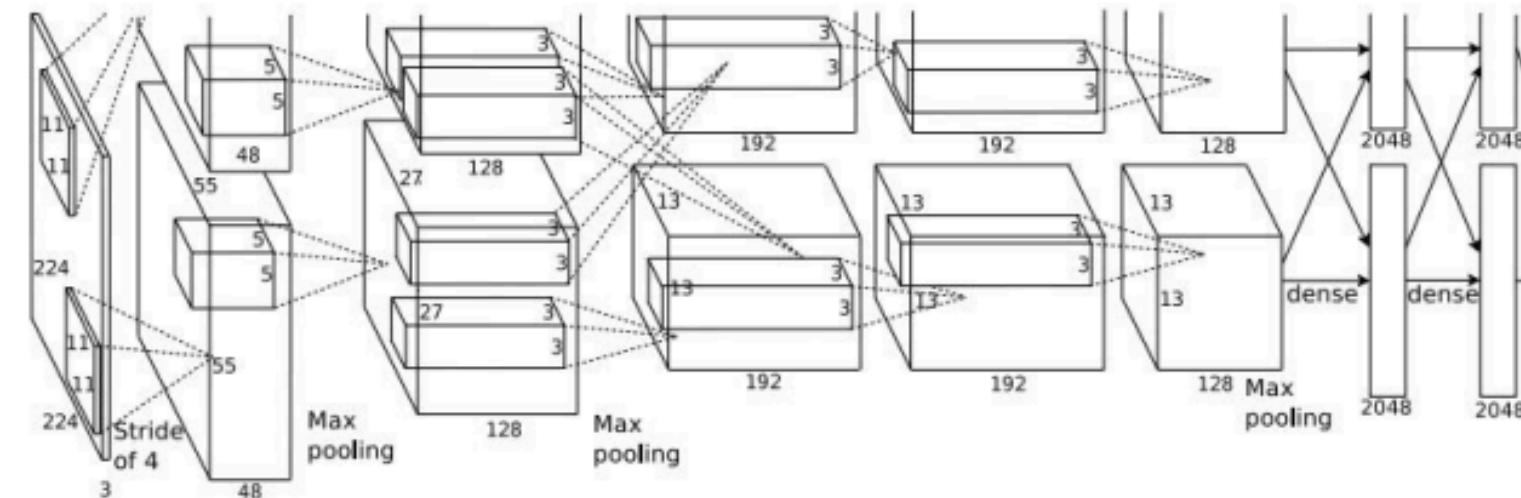
DUCK: (x, y, w, h)

...

Image credit

Object Detection

Несколько объектов - выбираем область изображения и классифицируем

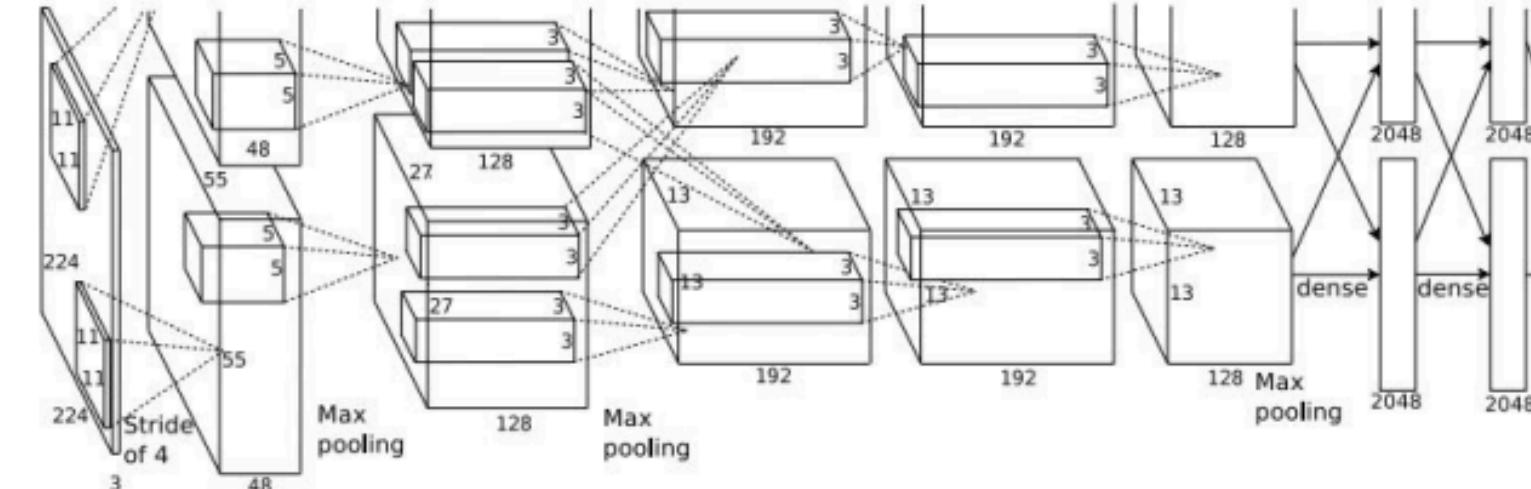
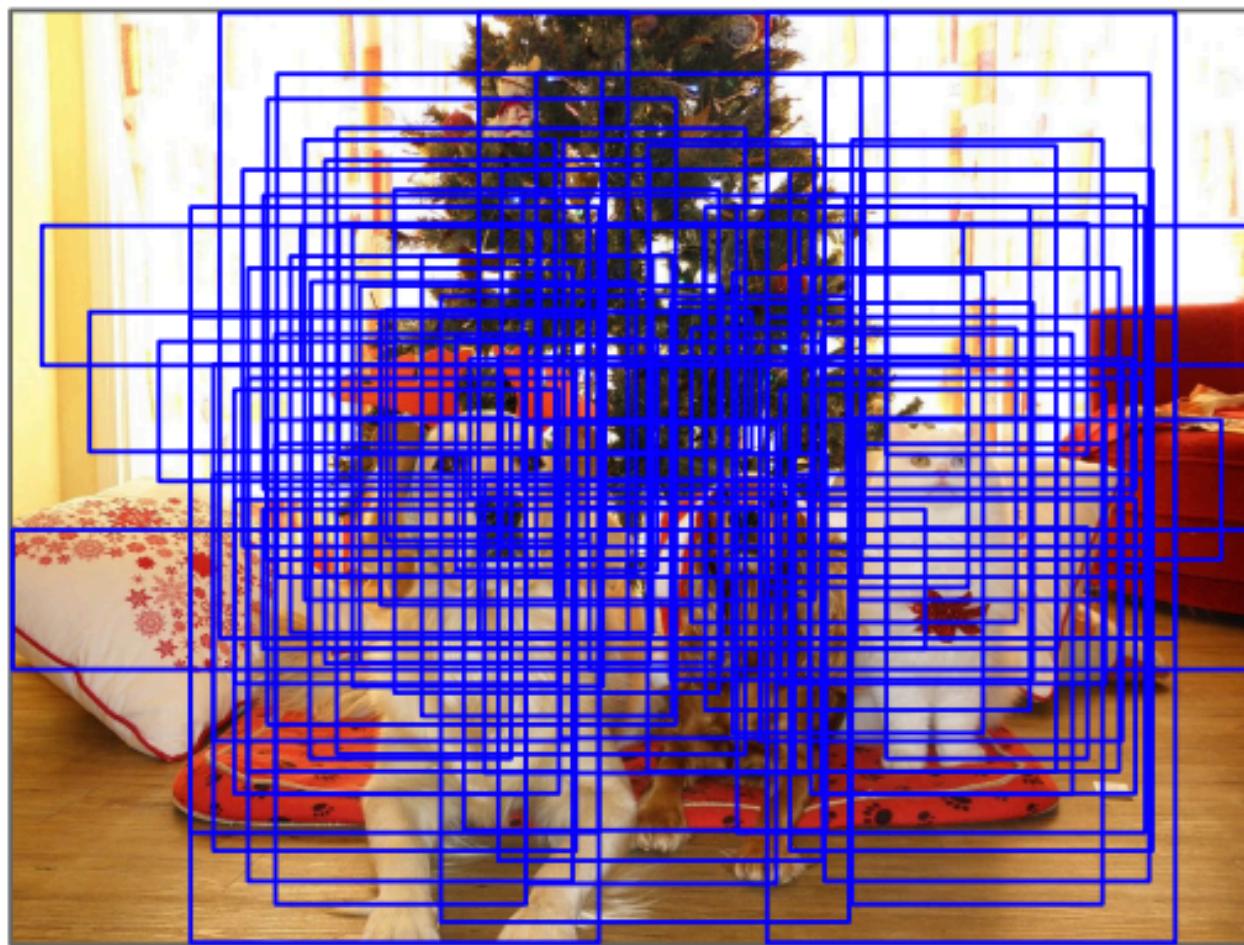


Dog? NO
Cat? NO
Background? YES

Object Detection

Несколько объектов - выбираем область изображения и классифицируем

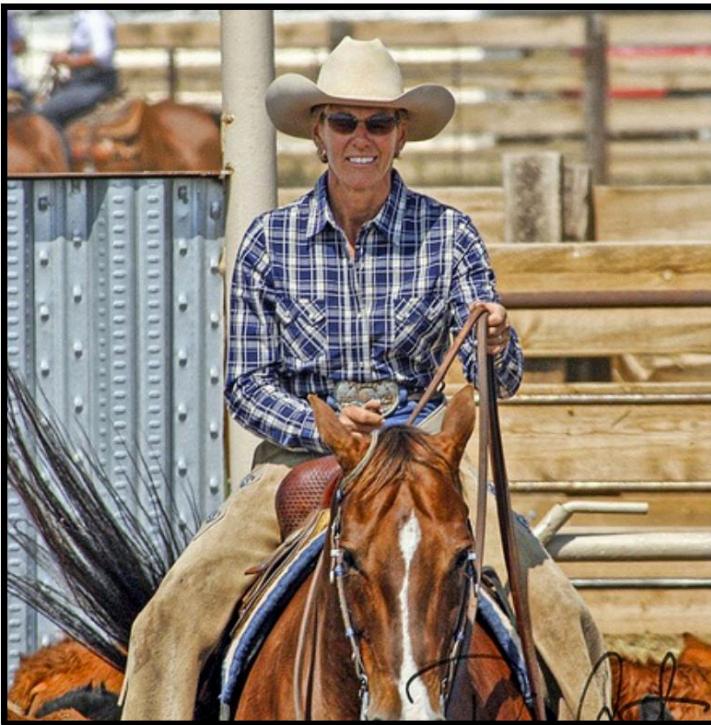
Проблема: вычислительно сложно - нужно очень много раз применять CNN



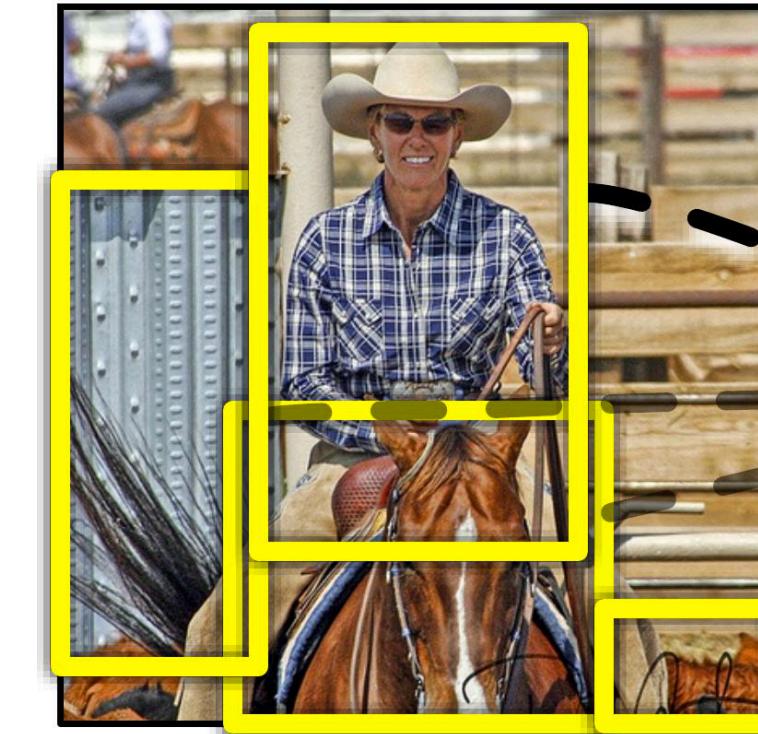
Dog? NO
Cat? NO
Background? YES

Object Detection

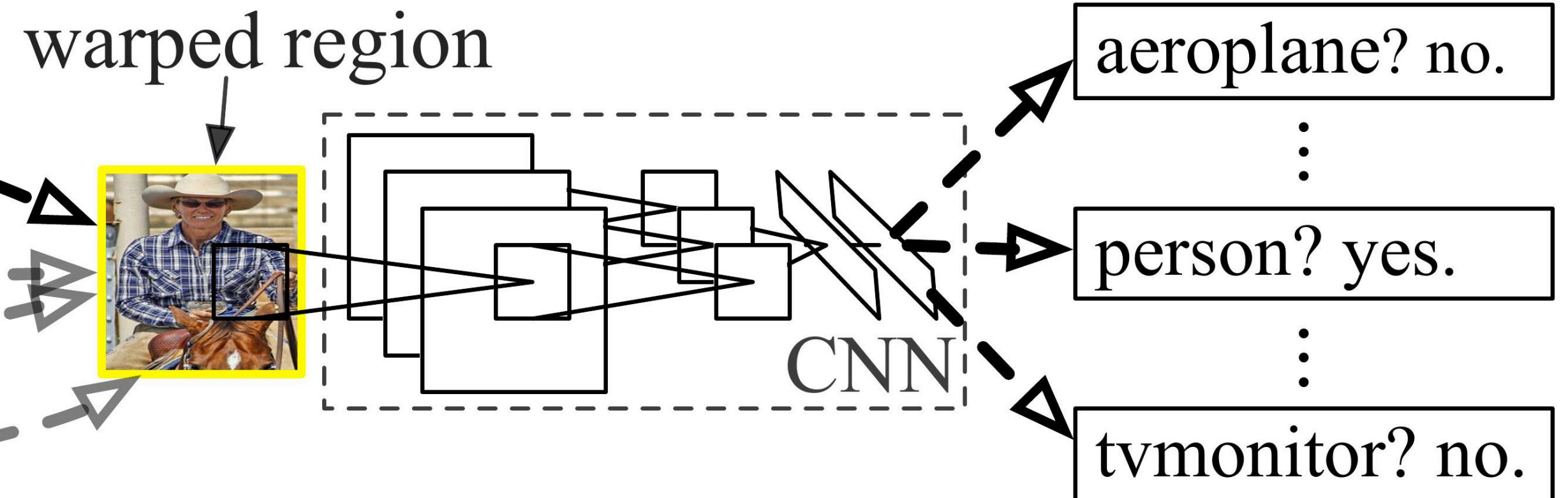
R-CNN: *Regions with CNN features*



1. Input image



2. Extract region proposals (~2k)

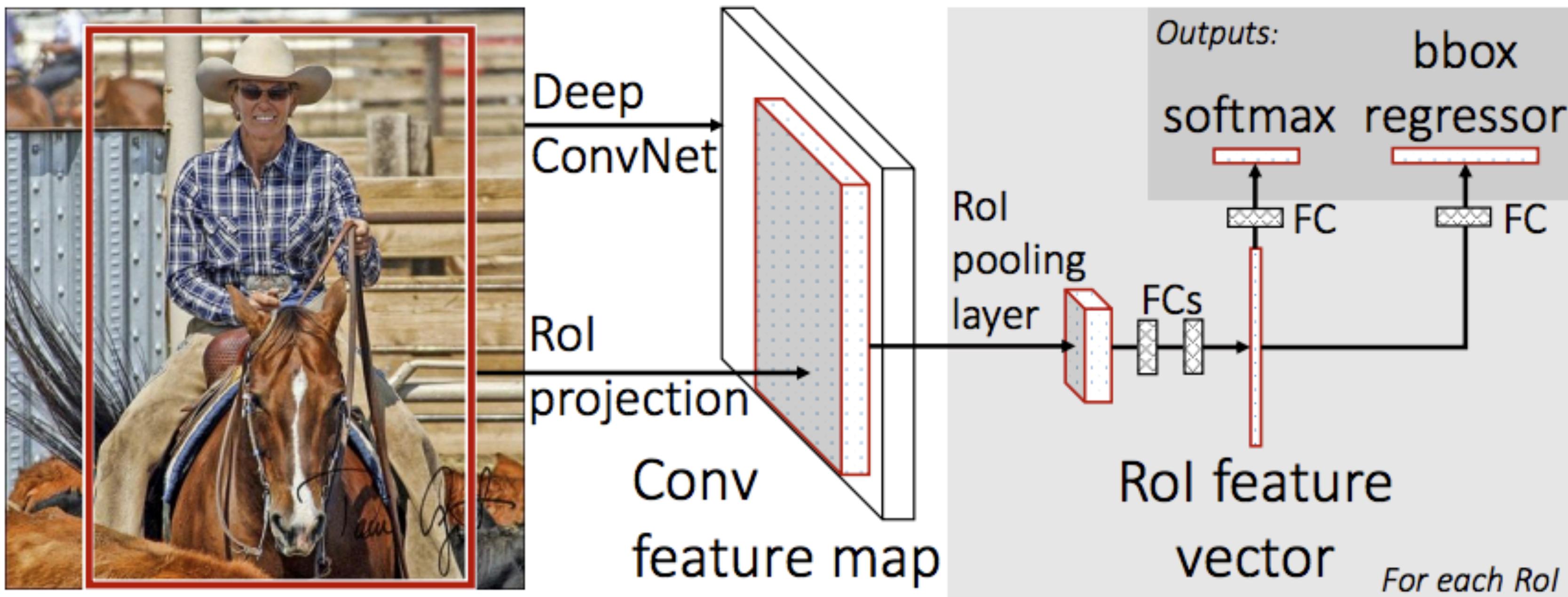


3. Compute CNN features

4. Classify regions

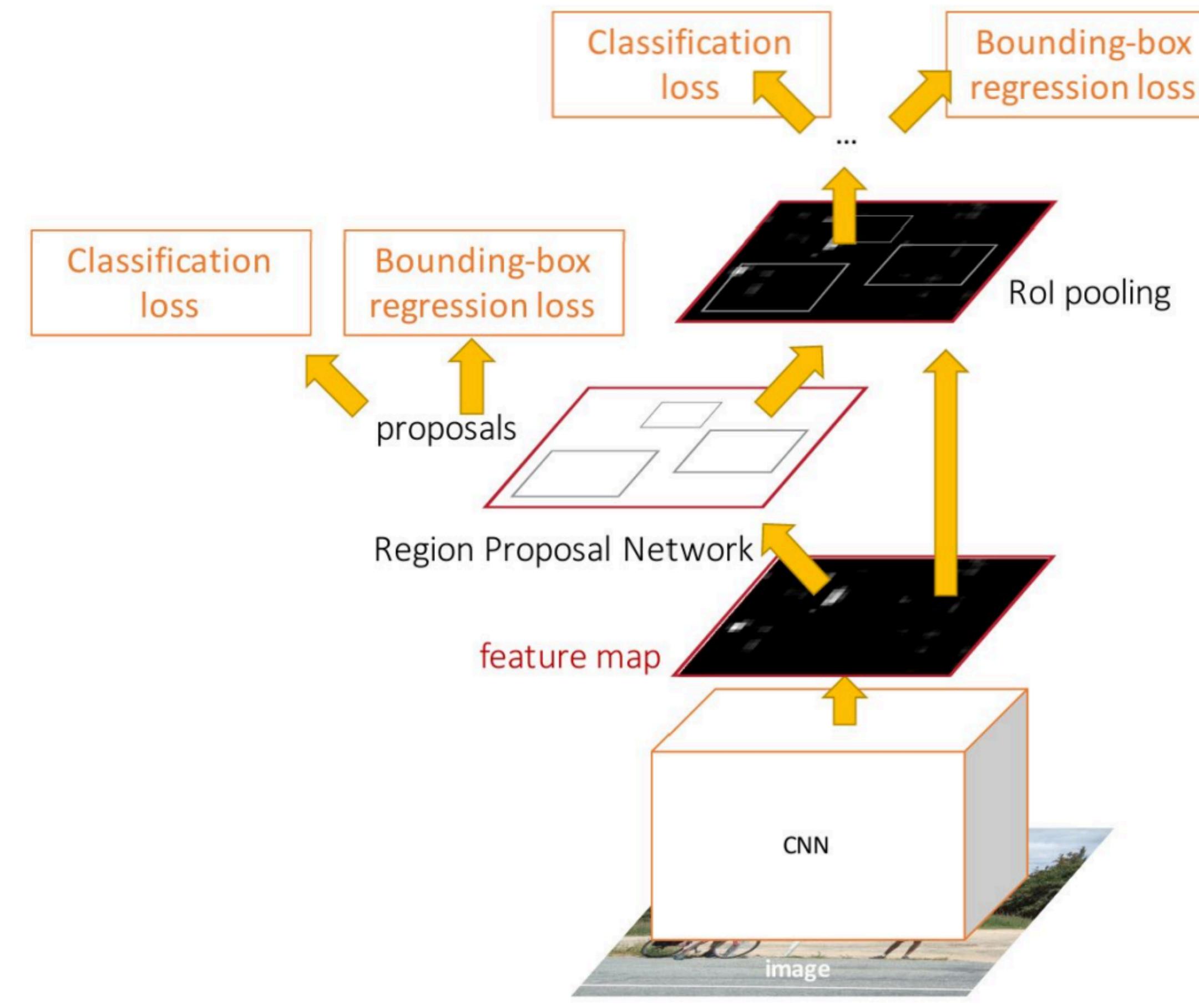
Object Detection

Fast R-CNN: identify RoI from conv feature map

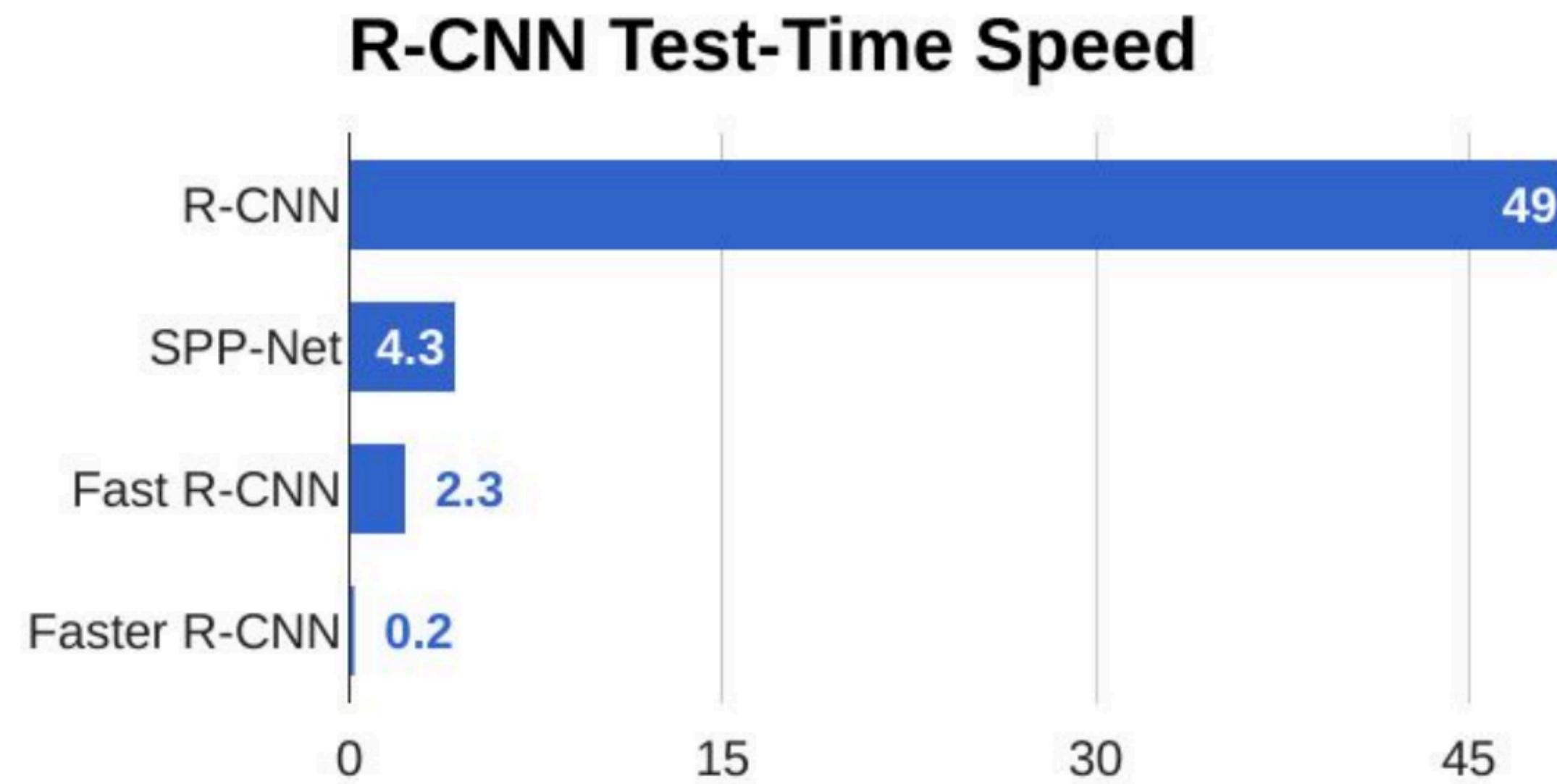


Object Detection

Faster R-CNN: use Region Proposal Network

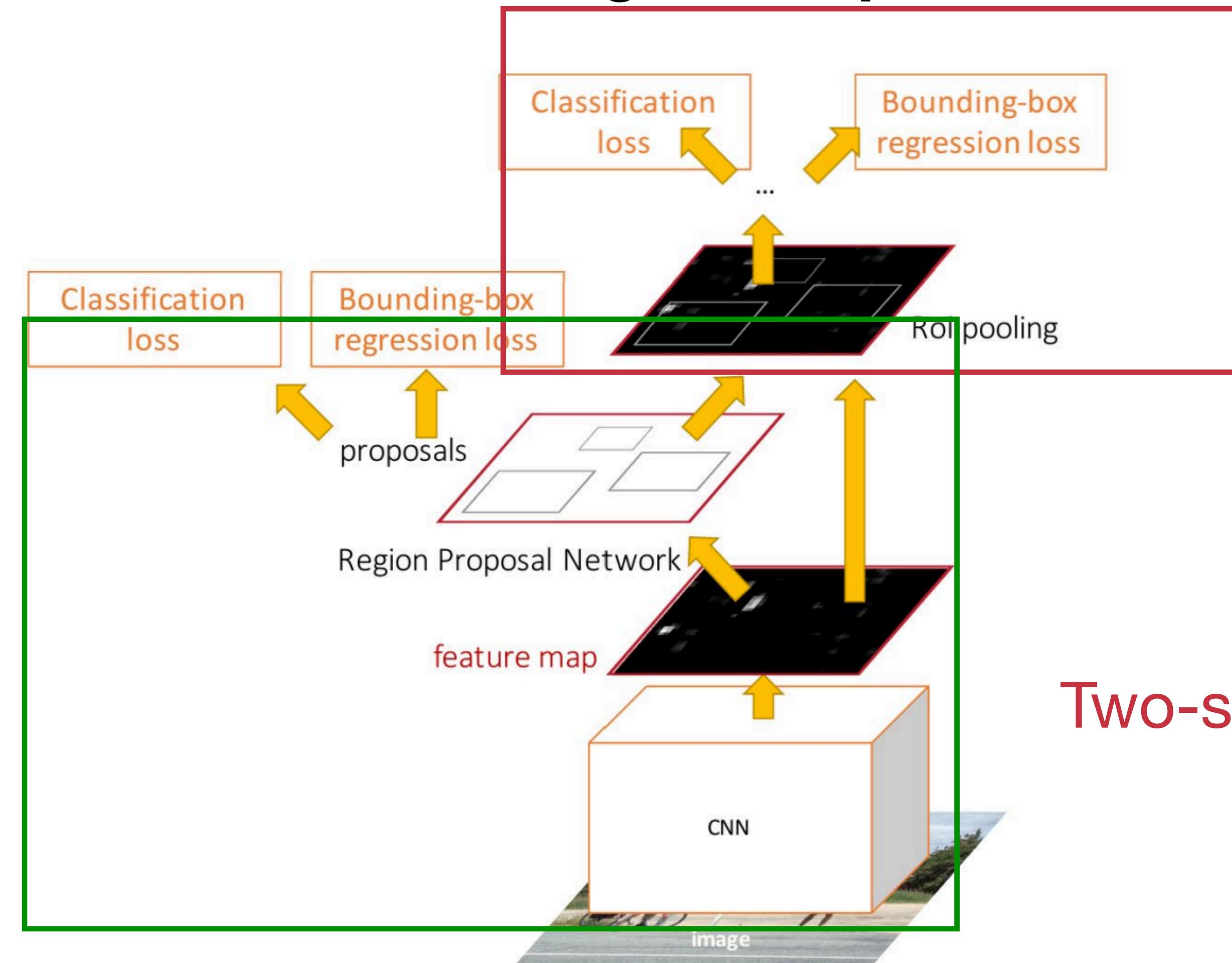


Object Detection



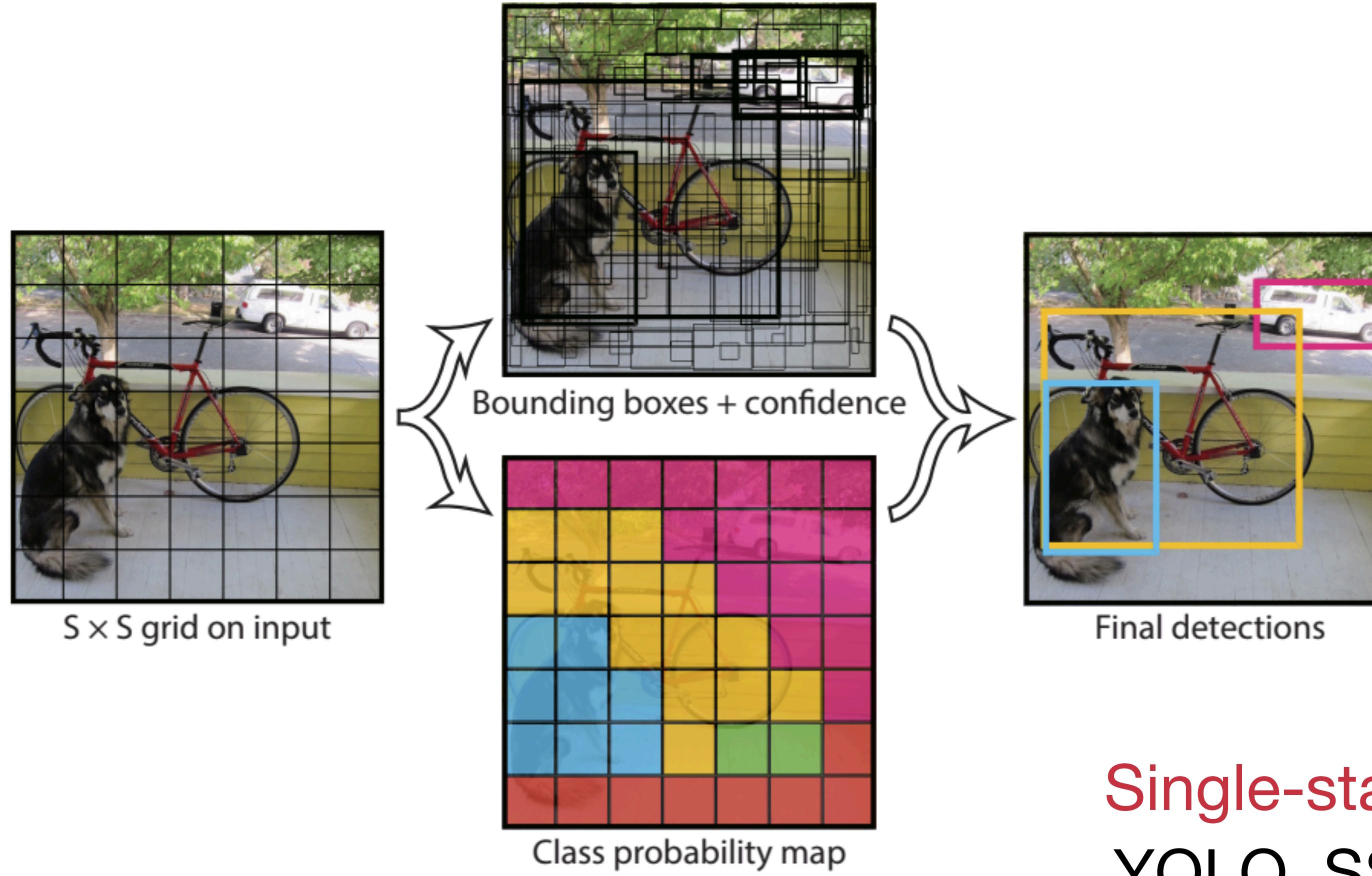
Object Detection

Faster R-CNN: use Region Proposal Network



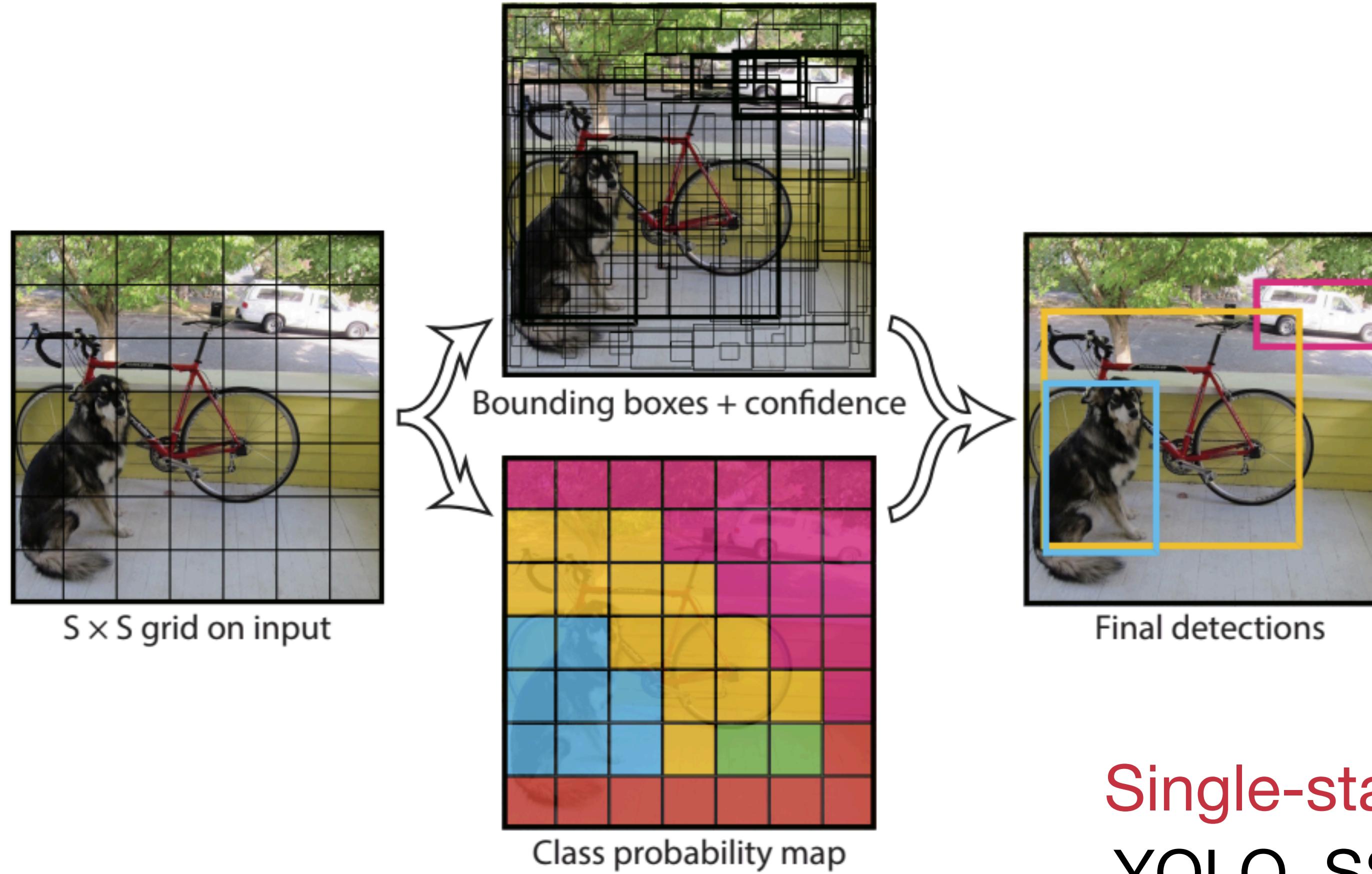
Two-stage Object Detector

Object Detection



Single-stage Object Detector
YOLO, SSD, RetinaNet

Object Detection



Single-stage Object Detector
YOLO, SSD, RetinaNet
менее точные, быстрые

[Image credit](#)

Instance Segmentation

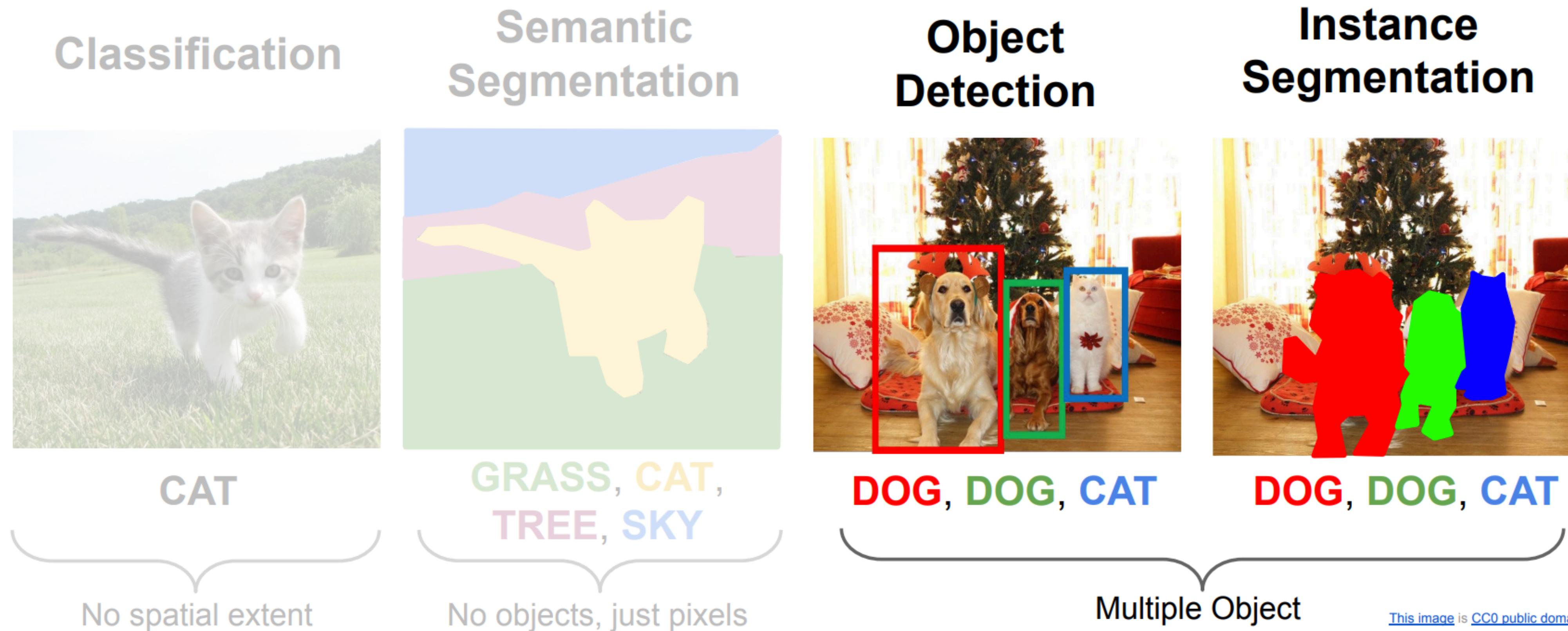


Image credit

Instance Segmentation

Masked R-CNN

