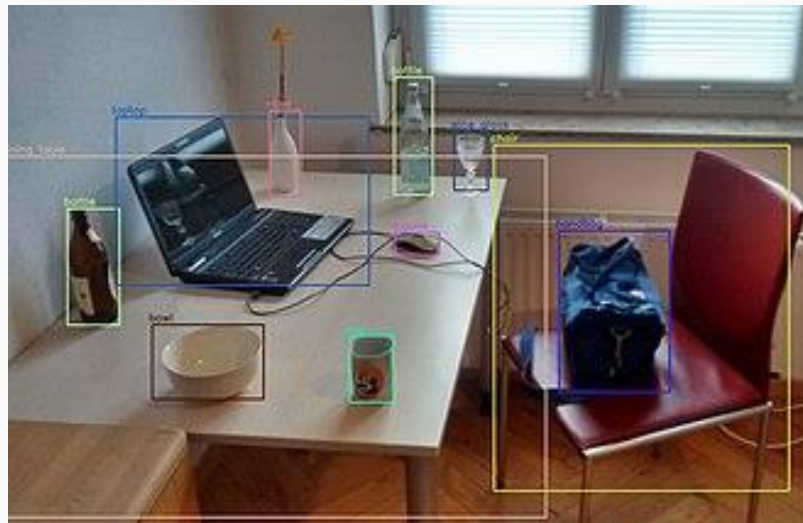


Занятие 35. Object detection

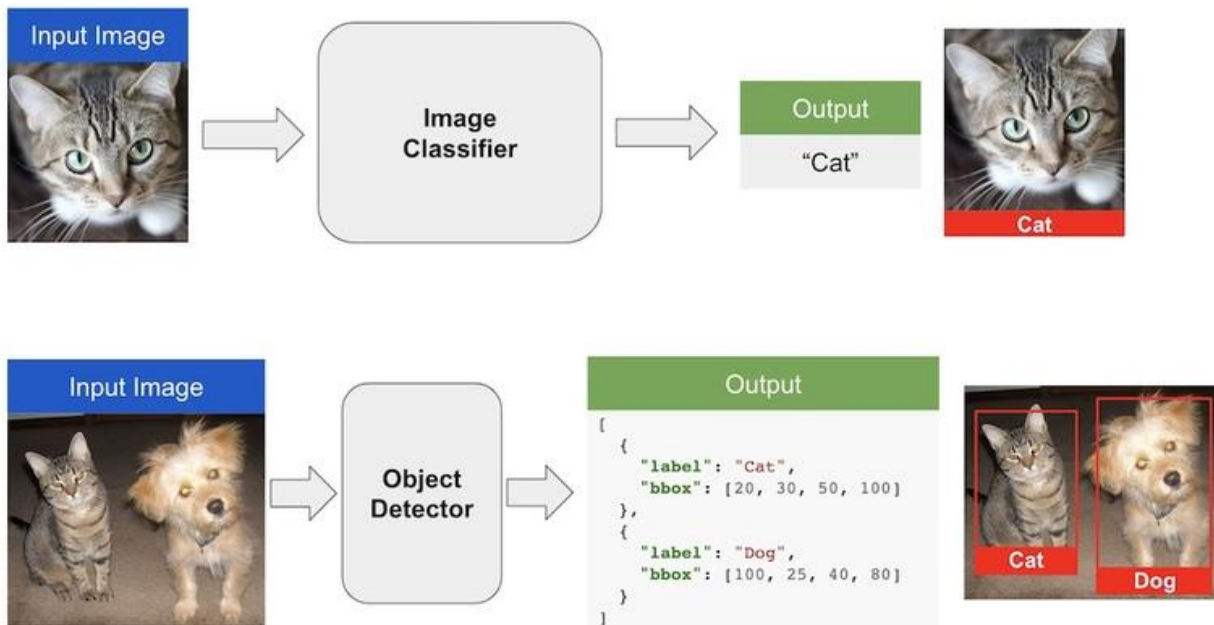
Talk about...

- Viola-Jones, HOG
- Anchor boxes, RPN (Region Proposal Network)
- NMS (Non-maximum suppression)
- Fast R-CNN, Faster R-CNN
- YOLO (You Only Look Once)
- SSD (Single Shot Multibox Detector)



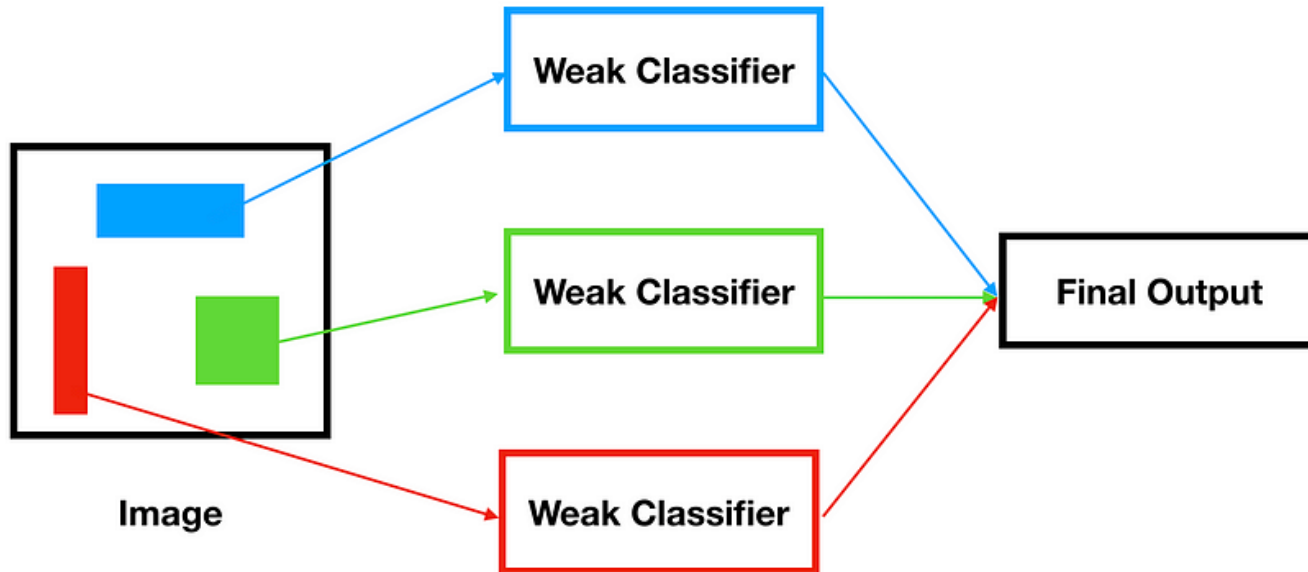
1. Определение и обзор области Object Detection

Различие между классификацией изображений и обнаружением объектов



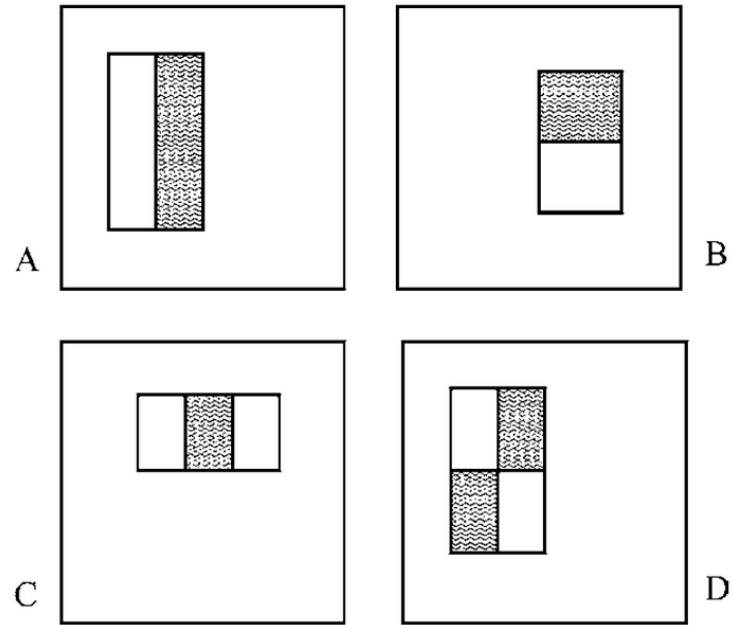
2. Исторический обзор

Viola-Jones



2. Исторический обзор

Viola-Jones

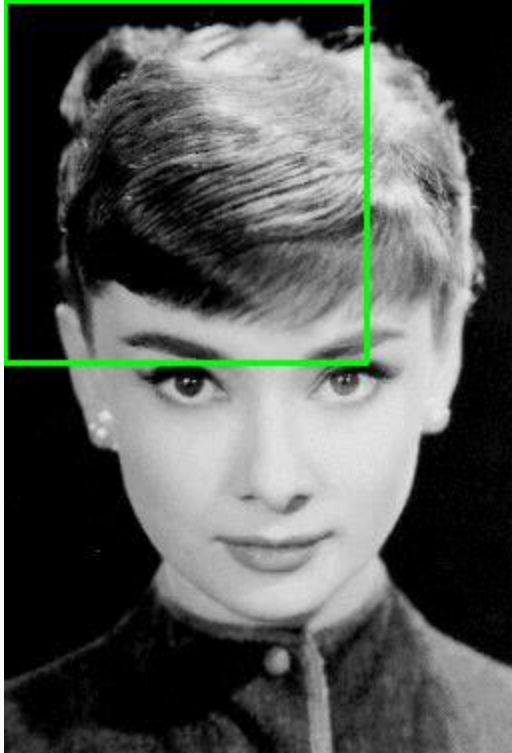


2. Исторический обзор

Histogram of Oriented Gradients and Object Detection

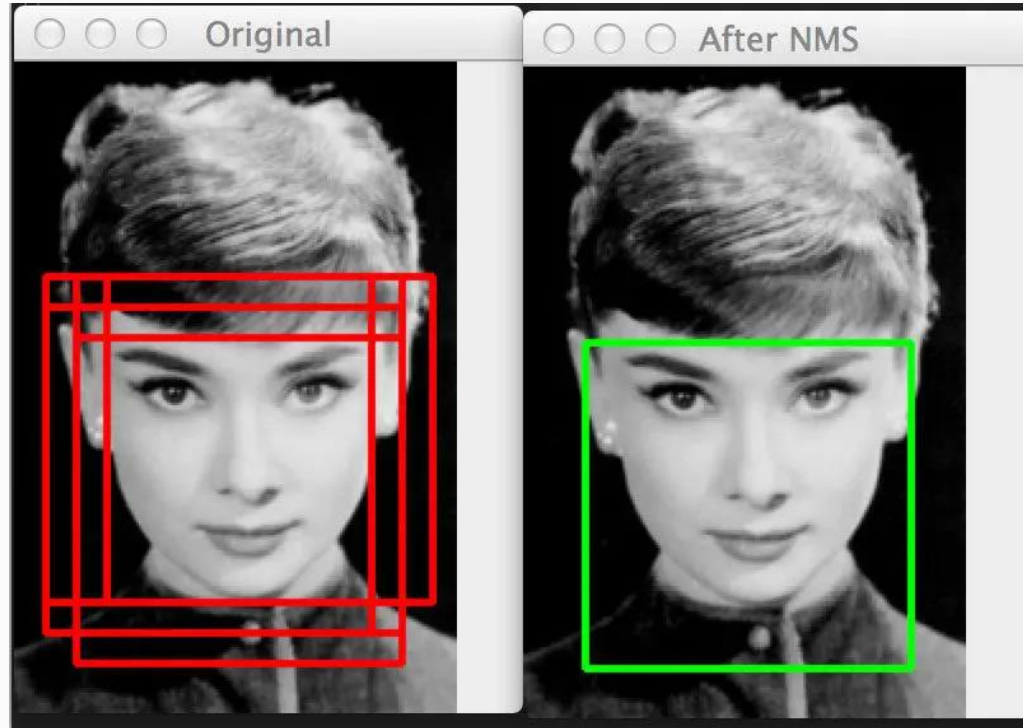


Histogram of Oriented Gradients and Object Detection



Пример подхода к передвижению окна, при котором мы перемещаем окно слева направо и сверху вниз. Примечание. Показан только один масштаб. На практике это окно будет применяться к изображению в нескольких масштабах.

Histogram of Oriented Gradients and Object Detection

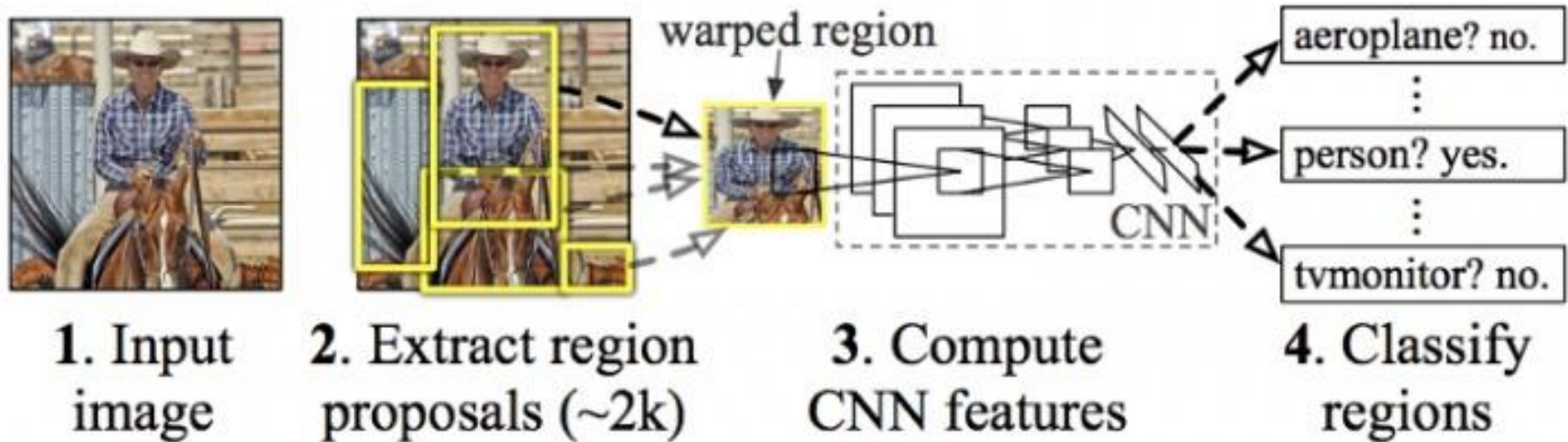


Обнаружение нескольких перекрывающихся ограничивающих рамок вокруг лица, которое мы хотим обнаружить. (Справа) Применение немаксимального подавления для удаления избыточных ограничивающих рамок.

3. Основные компоненты архитектуры сверточных нейронных сетей для Object detection

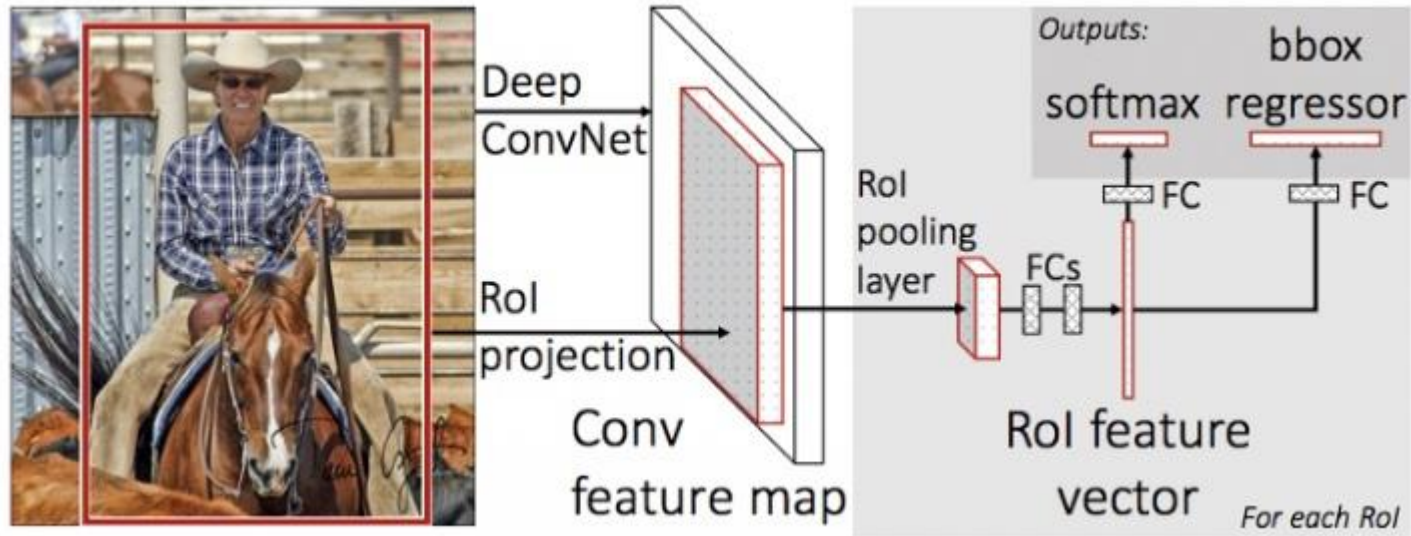
Region Proposals или Object Proposals.

R-CNN: *Regions with CNN features*



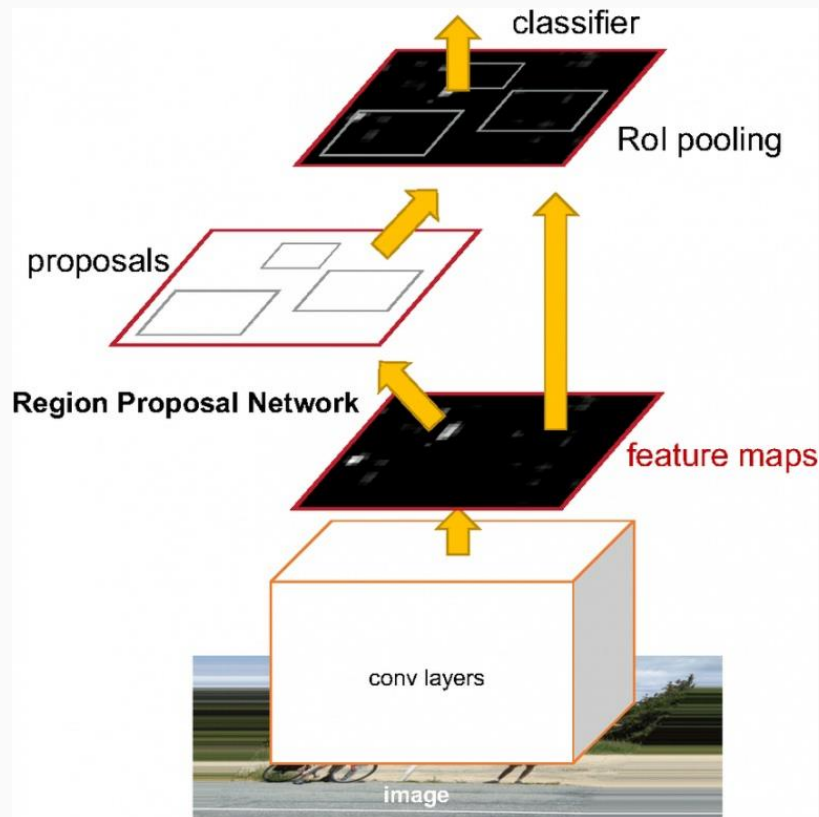
3. Основные компоненты архитектуры сверточных нейронных сетей для Object detection

Fast R-CNN



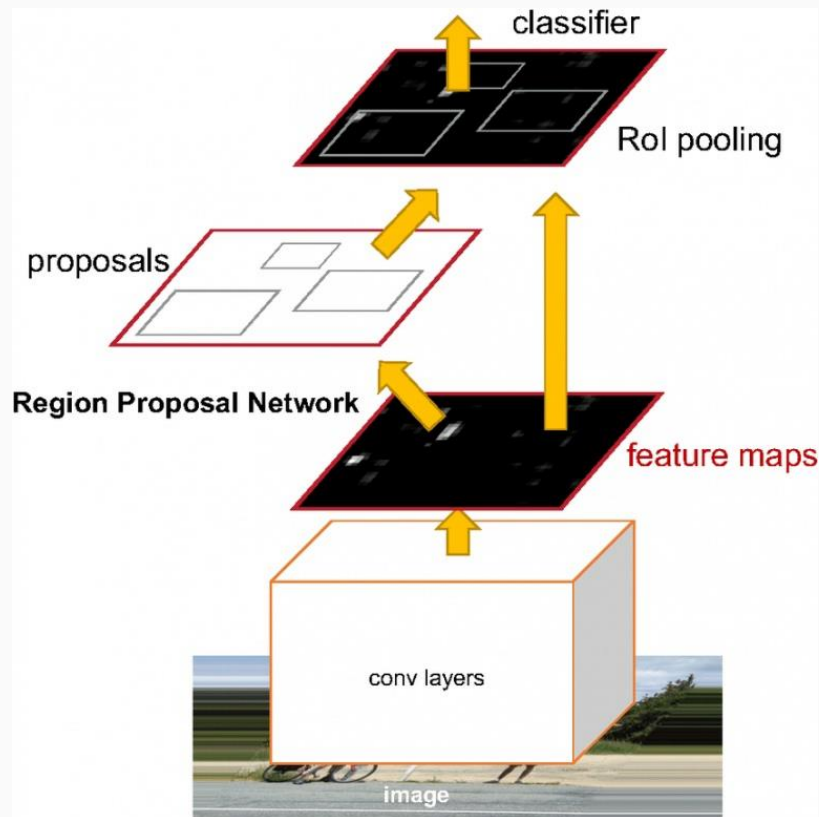
3. Основные компоненты архитектуры сверточных нейронных сетей для Object detection

Fast R-CNN



3. Основные компоненты архитектуры сверточных нейронных сетей для Object detection

Fast R-CNN



3. Основные компоненты архитектуры сверточных нейронных сетей для Object detection

Слои регрессии:

Для задачи обнаружения объектов СНС часто включают слои, предназначенные для предсказания координат ограничивающего прямоугольника объекта.

ROI Pooling:

Используется в некоторых архитектурах (например, Faster R-CNN) для обеспечения фиксированного размера выходных данных из различных регионов интереса на изображении.

Anchors или якоря:

Используются в современных методах обнаружения (например, SSD и Faster R-CNN) для определения базовых размеров и пропорций ограничивающих рамок.

Non-maximum suppression (NMS):

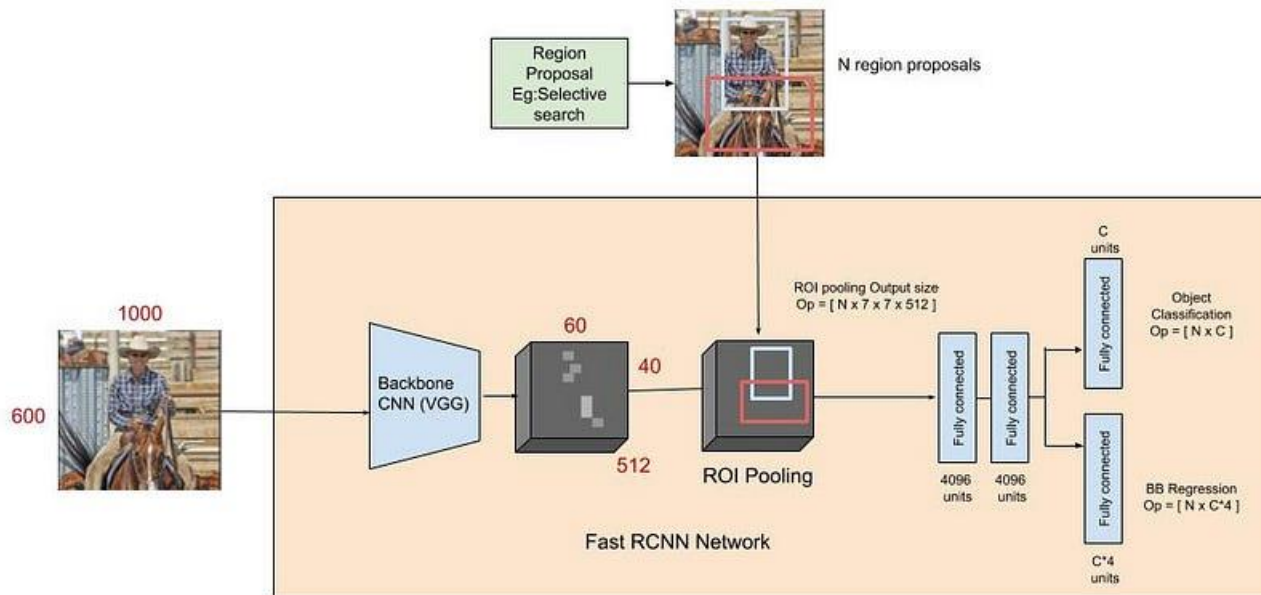
Постобработка, которая удаляет дублирующиеся или перекрывающиеся предсказания, оставляя только наиболее уверенное предсказание для каждого объекта.

Multiscale feature maps:

Особенно в современных архитектурах, таких как RetinaNet или FPN, для обнаружения объектов разных размеров используются признаки из разных уровней сети.

3. Популярные архитектуры для Object detection

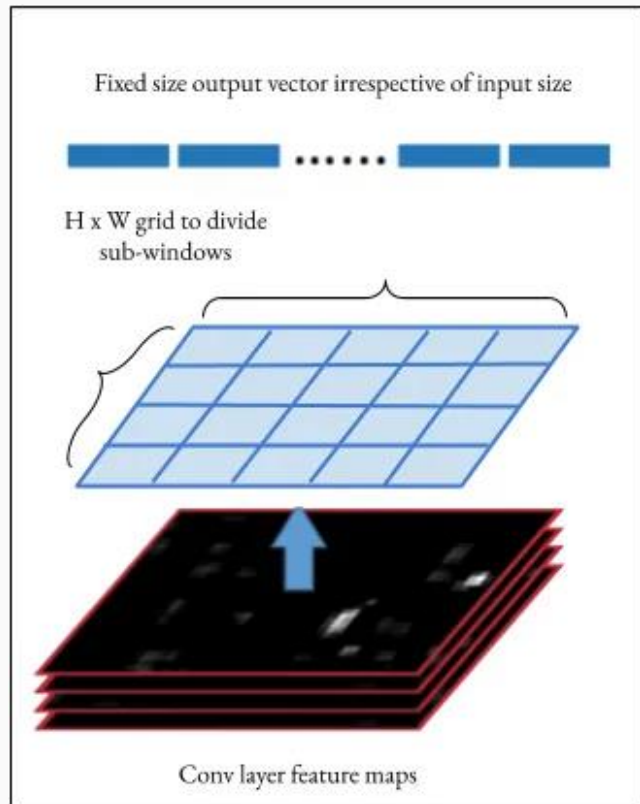
Fast R-CNN:



3. Популярные архитектуры для Object detection

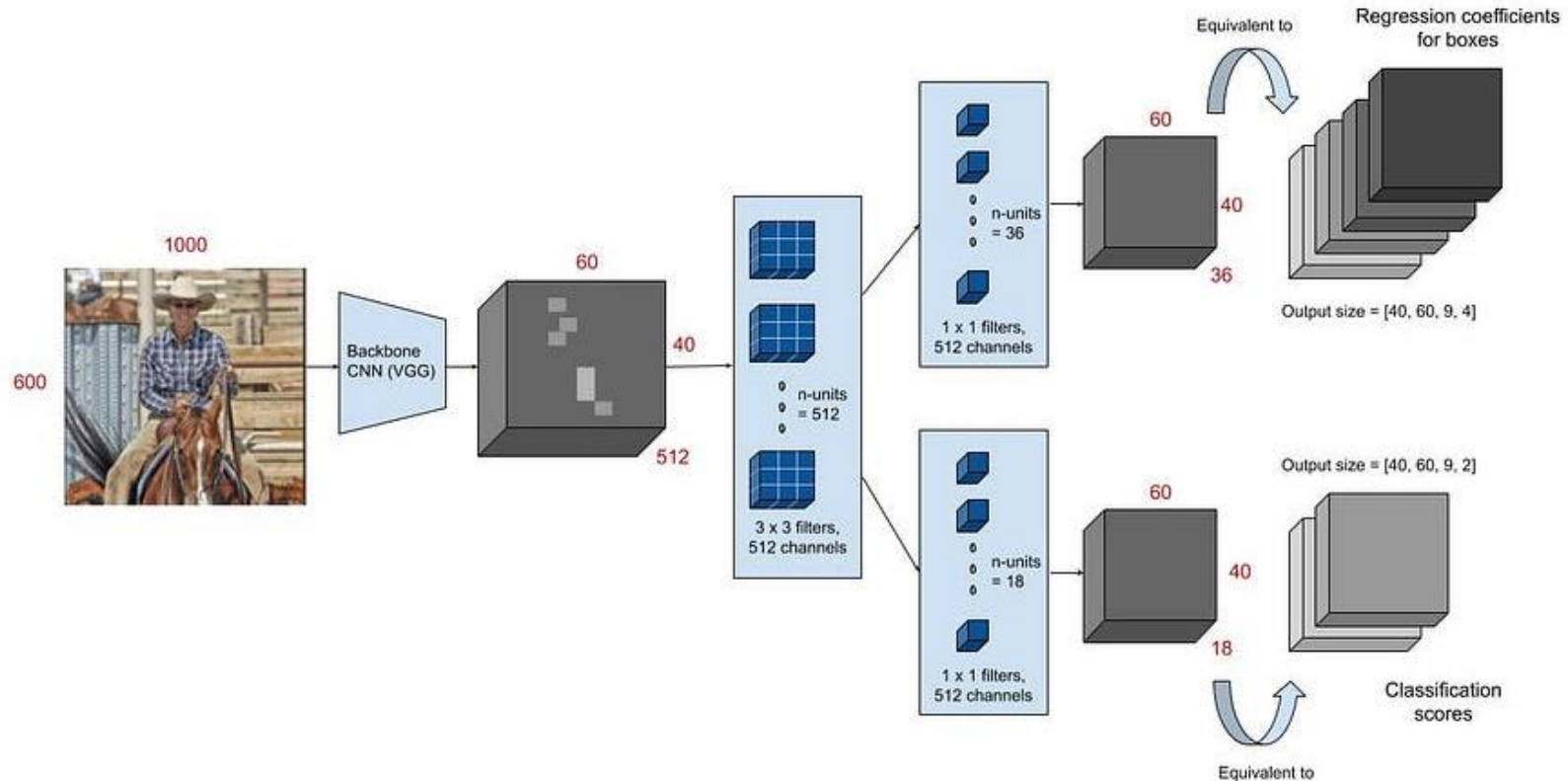
Fast R-CNN:

- **Основное улучшение** по сравнению с исходным R-CNN - это введение метода RoI (Region of Interest) pooling. Это позволяет извлекать признаки разного размера из разных регионов изображения и преобразовывать их в фиксированный размер, чтобы можно было проводить классификацию и регрессию ограничивающего прямоугольника одновременно.
- Быстрее R-CNN, так как избегает необходимости применять классификатор ко всем предложенным регионам отдельно.



3. Популярные архитектуры для Object detection

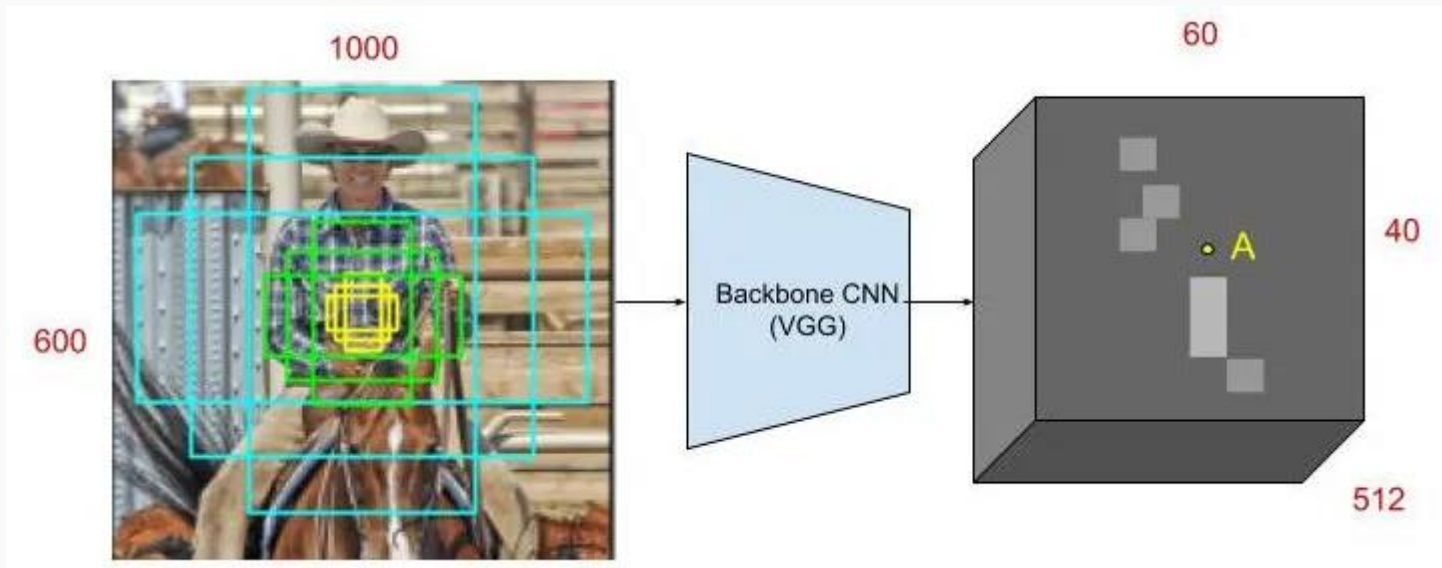
Faster R-CNN:



3. Популярные архитектуры для Object detection

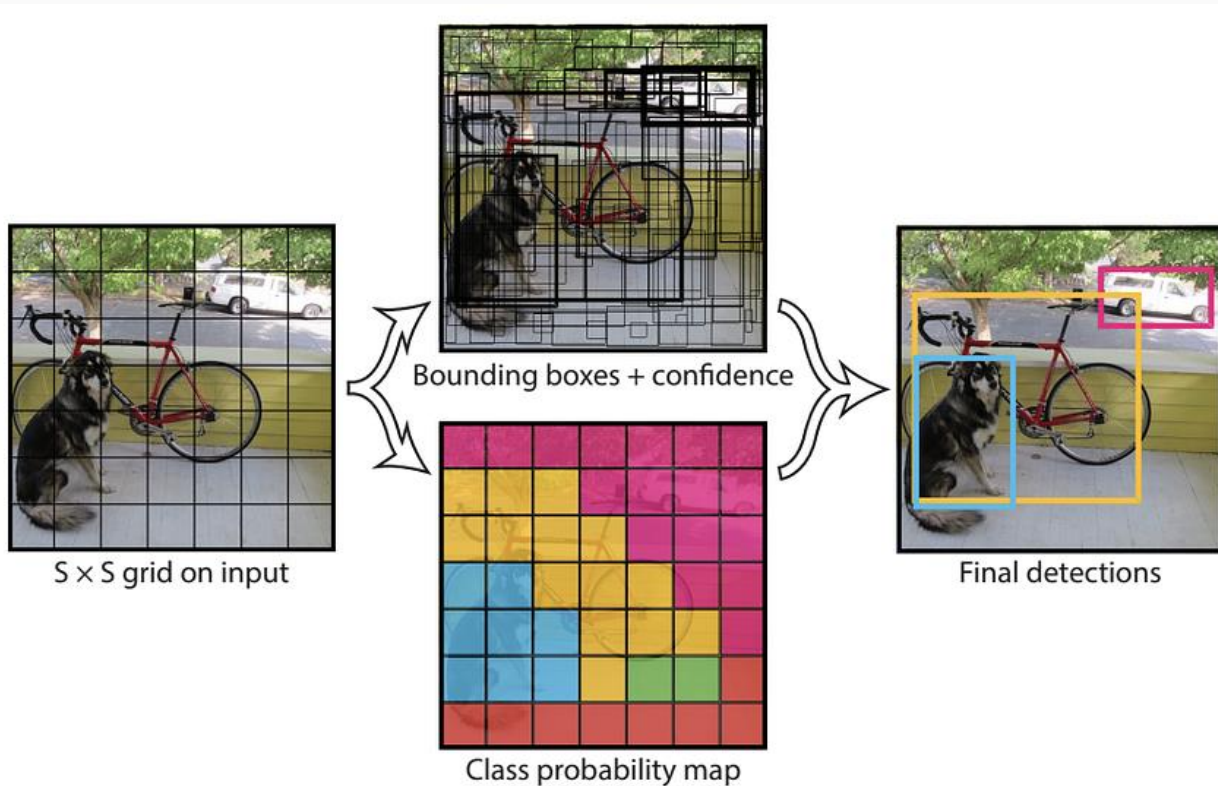
Faster R-CNN:

- **Основное улучшение** Добавляет компонент RPN (Region Proposal Network), который предлагает регионы для обнаружения. Это делает процесс предложения региона более быстрым и интегрированным с процессом обучения.
- Считается одной из наиболее точных архитектур, но может быть довольно медленной в реальном времени.



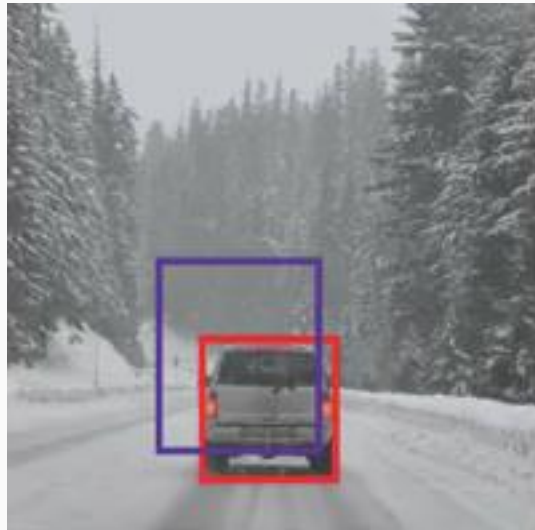
3. Популярные архитектуры для Object detection

YOLO (You Only Look Once)

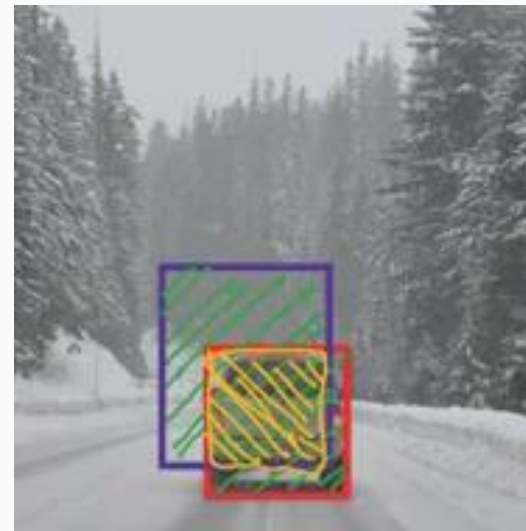


3. Популярные архитектуры для Object detection

YOLO (You Only Look Once)



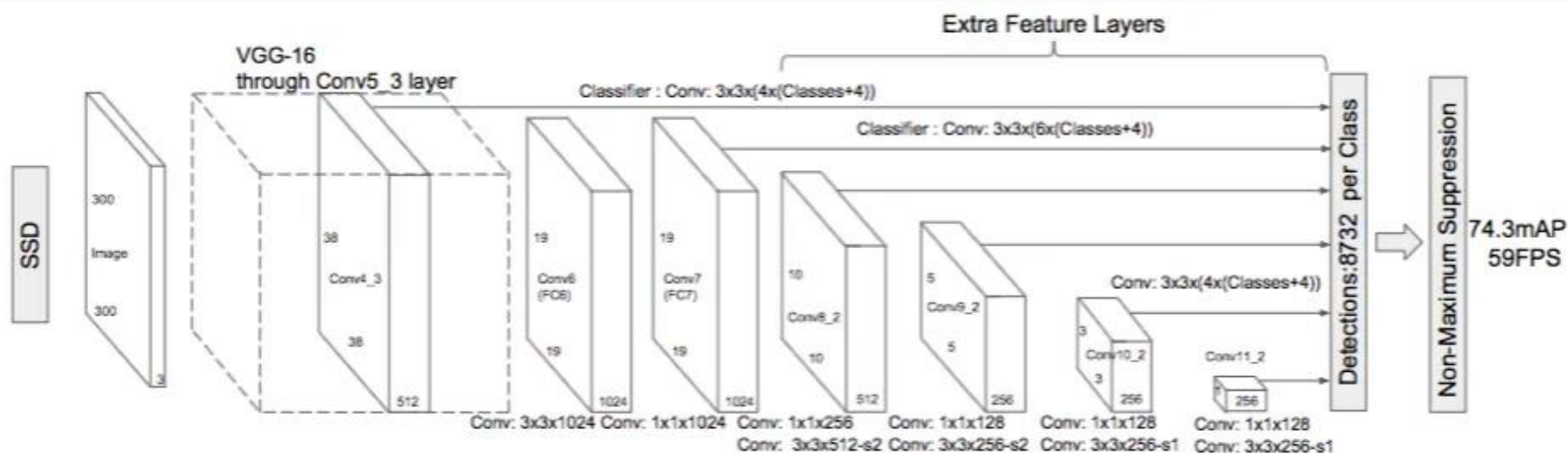
IoU — это то, насколько хорошо прогнозируемая ограничивающая рамка машины совпадает с ограничивающей рамкой фактического объекта.



Не максимальное подавление - NMS

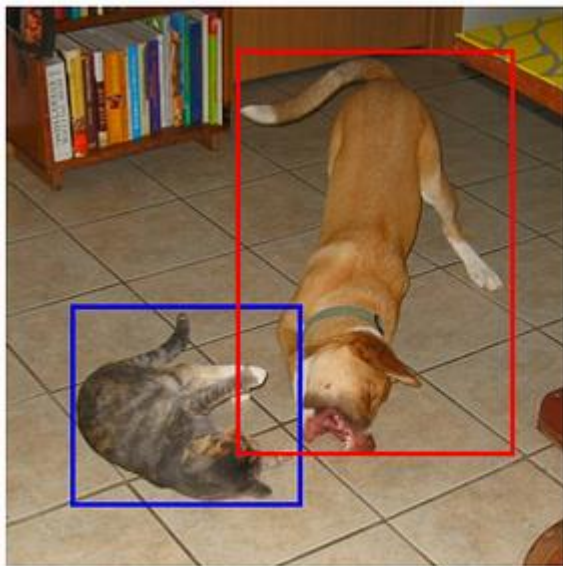
3. Популярные архитектуры для Object detection

SSD MultiBox

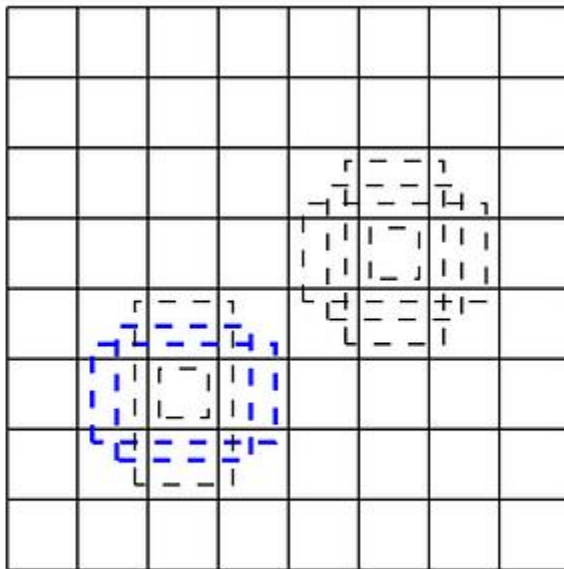


3. Популярные архитектуры для Object detection

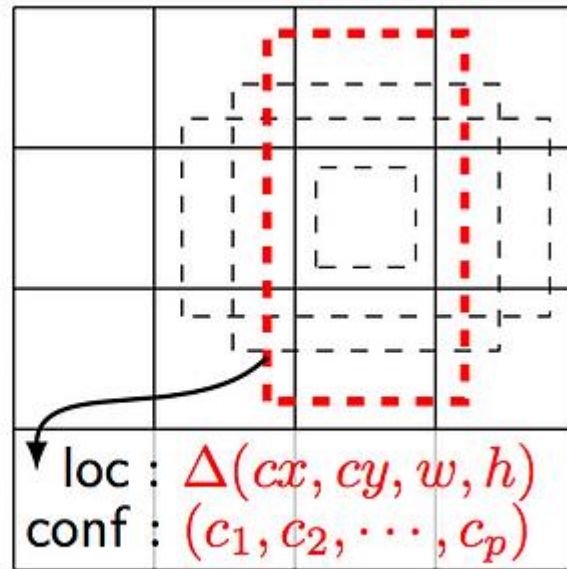
SSD MultiBox



(a) Image with GT boxes



(b) 8×8 feature map



loc : $\Delta(cx, cy, w, h)$
conf : (c_1, c_2, \dots, c_p)

(c) 4×4 feature map

3. Популярные архитектуры для Object detection

SSD MultiBox



Figure 1: We propose a non-maximum suppression convnet that will re-score all raw detections (top). Our network is trained end-to-end to learn to generate exactly one high scoring detection per object (bottom, example result).

3. Популярные архитектуры для Object detection

SSD MultiBox



3. Популярные архитектуры для Object detection

Основные различия и преимущества:

- **Fast R-CNN** и **Faster R-CNN**: улучшенное и более быстрое обнаружение по сравнению с исходным R-CNN; Faster R-CNN интегрирует процесс предложения региона.
- **YOLO**: очень быстрая архитектура с однопроходным обнаружением; может не быть такой точной, как другие методы, особенно при обнаружении мелких или близко расположенных объектов.
- **SSD**: быстрый однопроходный метод с использованием разных масштабов признаков для обнаружения объектов разных размеров; обычно достигает лучшего компромисса между скоростью и точностью по сравнению с Faster R-CNN и YOLO.

Практическое задание

Демонстрация простого примера с использованием предварительно обученной модели.

Участники занятия попробуют обнаружить объекты на своих изображениях с помощью открытых библиотек и инструментов (например, TensorFlow или PyTorch).

Домашнее задание: Решение задачи object detection