



Поиск объектов на изображении. Семантическая сегментация.

Евгений Борисов

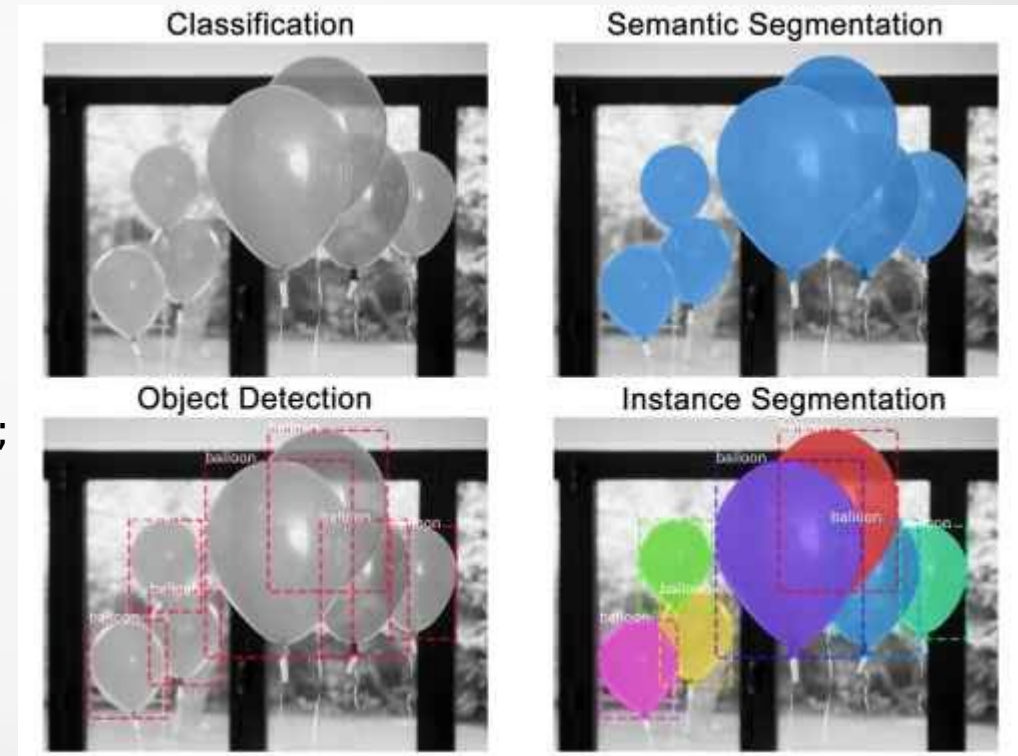
Свёрточные нейросети

Classification — классификация изображения по типу объекта, которое оно содержит;

Semantic segmentation — определение всех пикселей объектов определённого класса или фона на изображении. Если несколько объектов одного класса перекрываются, их пиксели никак не отделяются друг от друга;

Object detection — обнаружение всех объектов указанных классов и определение охватывающей рамки для каждого из них;

Instance segmentation — определение пикселей, принадлежащих каждому объекту каждого класса по отдельности;



ML обработка изображений

Сегментация изображения с помощью кластеризации цветов

гауссовская смесь (GMM) как модель для пространства цветов

применяем EM для сегментации изображения

компонента смеси определяет кластер цвета



ML обработка изображений

О задачах обнаружения и локализации объекта на изображении.

Модель фона - камера неподвижна

Модель объекта - необходимо понимать что хотим найти

ML обработка изображений

Модель фона

Усреднённый фон

Модель фона по Гауссу

Смесь гауссиан в качестве модели фона

ML обработка изображений

Модель фона / Усреднённый фон

Соберём историю из n кадров,
вычислим среднее значение
и будем использовать этот результат как фон

$$B = \frac{1}{n} \sum_{i=1}^n C_i$$

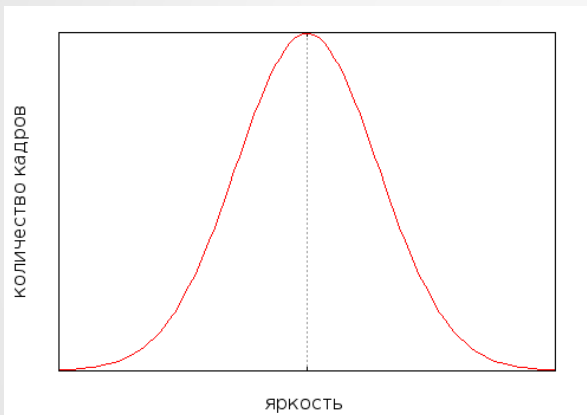


работает удовлетворительно только в условиях
стабильного освещения и отсутствия шума

ML обработка изображений

Модель фона / Модель фона по Гауссу

состояние точек фона находится в окрестности определённого значения
фон задаётся нормальными распределениями яркостей точек



$$\mu = \frac{1}{n} \sum_{i=1}^n x_i ; \quad \sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^2$$

$$p(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right)$$



ML обработка изображений

Модель фона / Смесь гауссиан

фон задаётся смесью нормальных распределений яркостей точек
позволяет формировать несколько кластеров для значений яркости

$$p(x) = \sum_{j=1}^K w_j \varphi_j(x)$$

$$\varphi(x; \mu, \Sigma) = \frac{\exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right)}{\sqrt{(2\pi)^n \det \Sigma}}$$



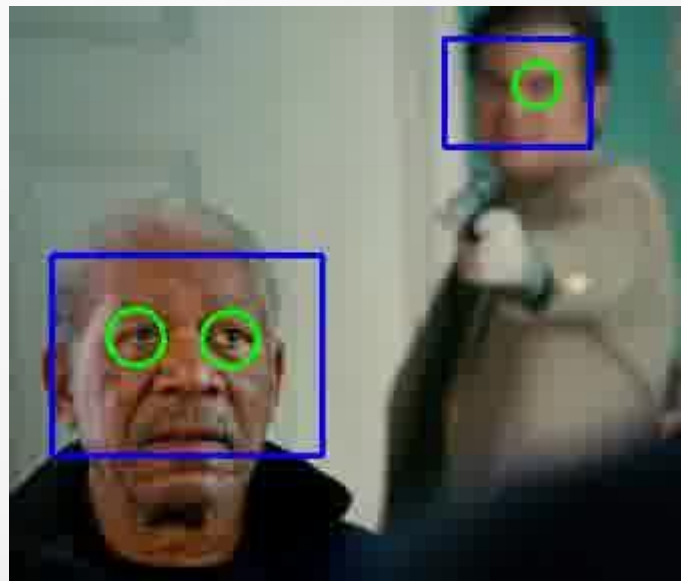
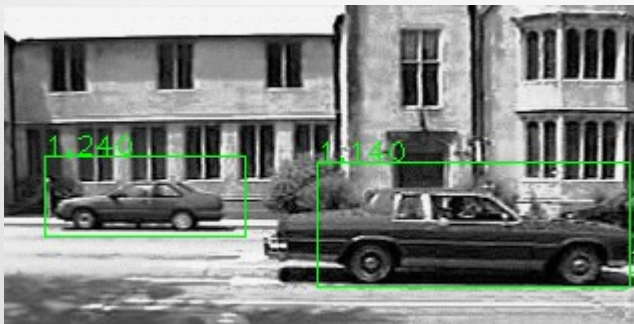
устойчивей к шуму и изменениям освещения чем простые модели

ML обработка изображений

Локализация объектов

метод скользящего окна и кластеризация результатов

1. задать размер окна
2. пройти окном изображение
3. на каждом шаге выполняем классификацию содержимого окна
4. изменить размер окна и повторить процедуру с п.2

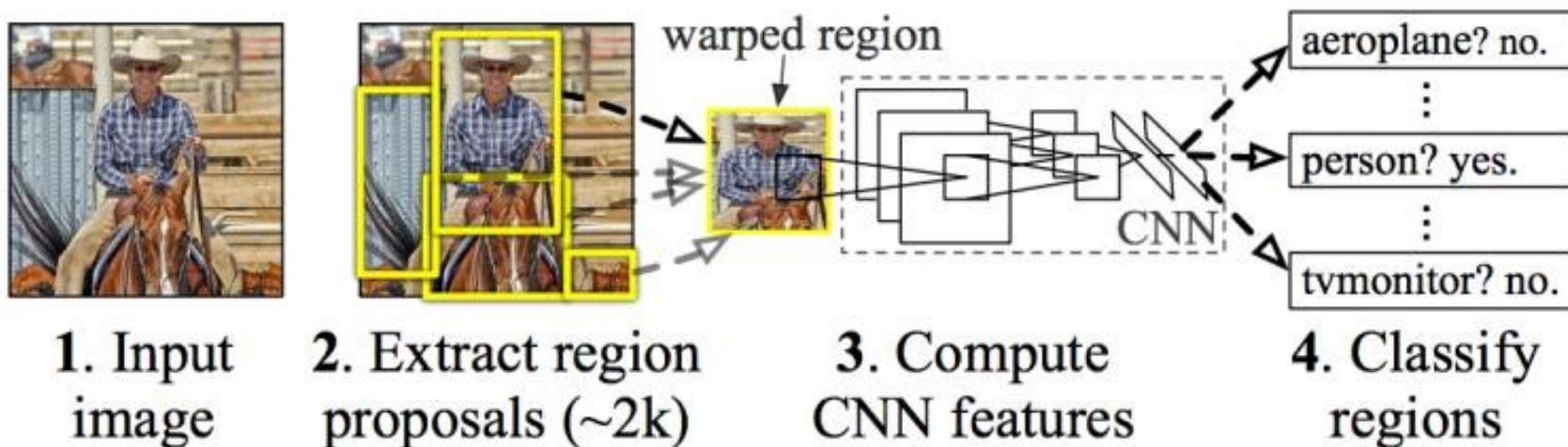


детектор лиц Виолы-Джонса

Свёрточные нейросети

Object detection

R-CNN: *Regions with CNN features*



Region Based Convolutional Neural Networks (R-CNN)

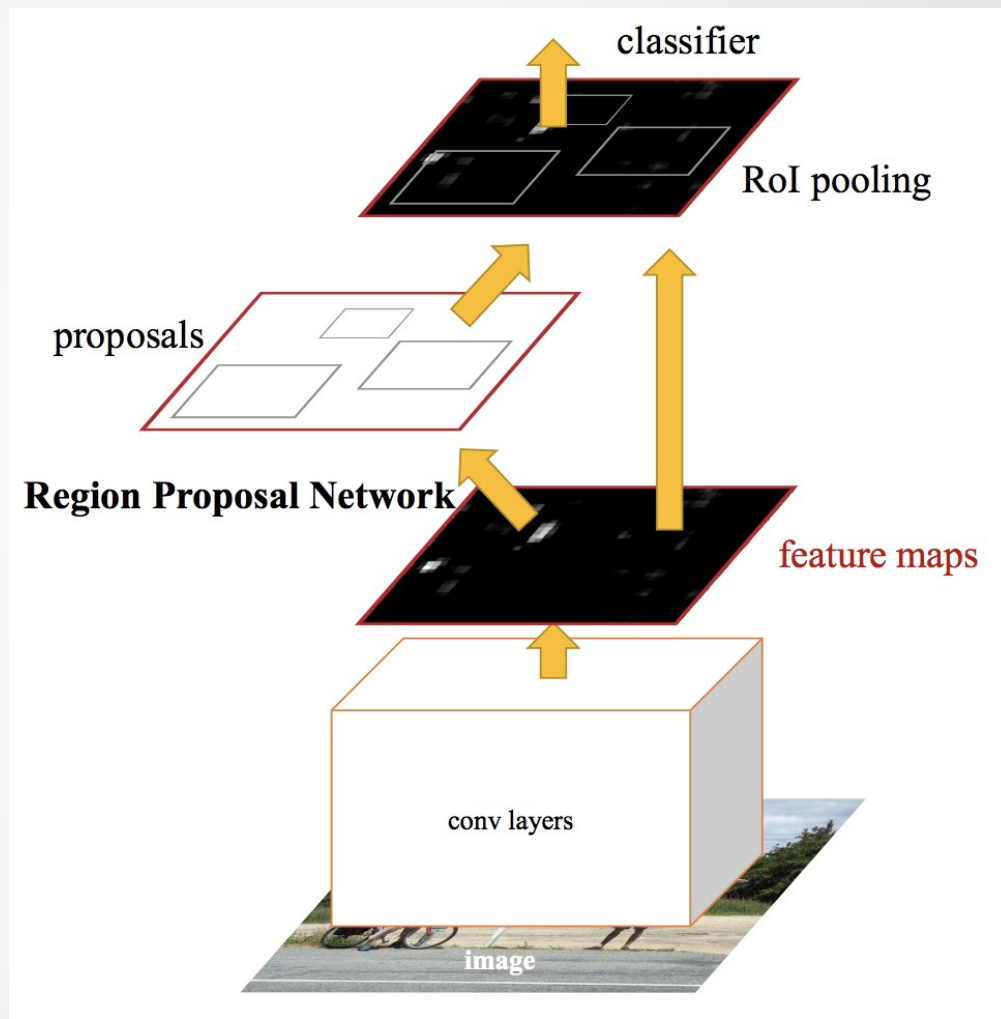
изображение разделяется на части

каждую часть проверяем классификатором

Свёрточные нейросети

Faster-R-CNN

- принимаем картинку на вход
- картинка прогоняется через CNN, формируем feature maps
- определяем регионы-кандидаты (возможно содержащие объекты)
- выделяем эти регионы
- и применяем к ним классификатор картинок

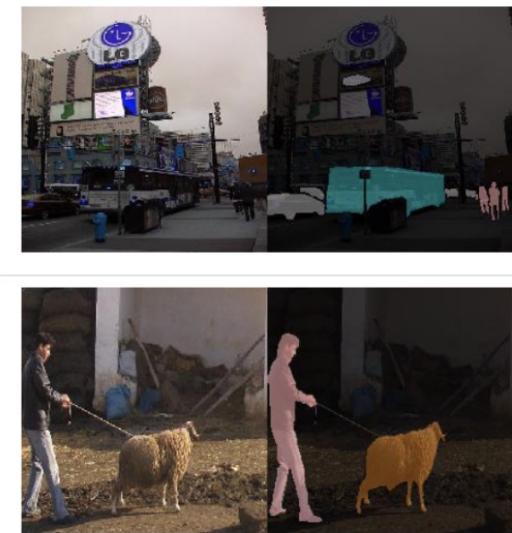
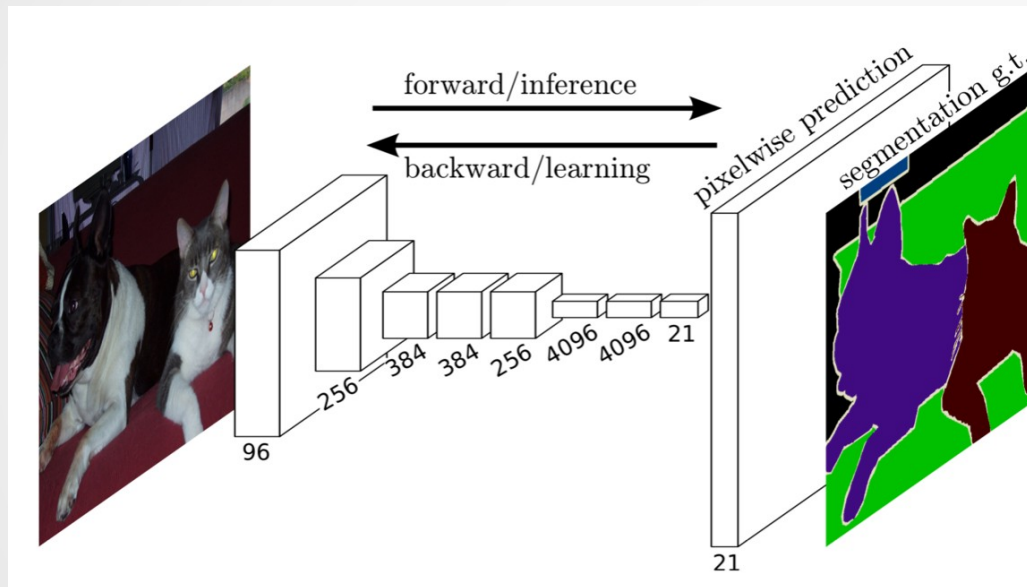


Свёрточные нейросети

Semantic segmentation

FCN: Fully Convolutional Networks

<https://arxiv.org/pdf/1411.4038.pdf>



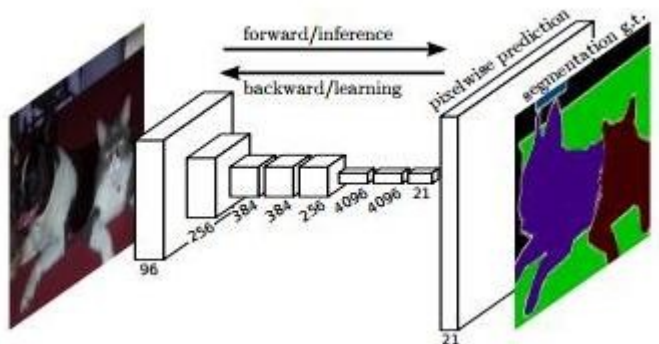
выход — карты поточечной оценки

для каждого класса своя карта

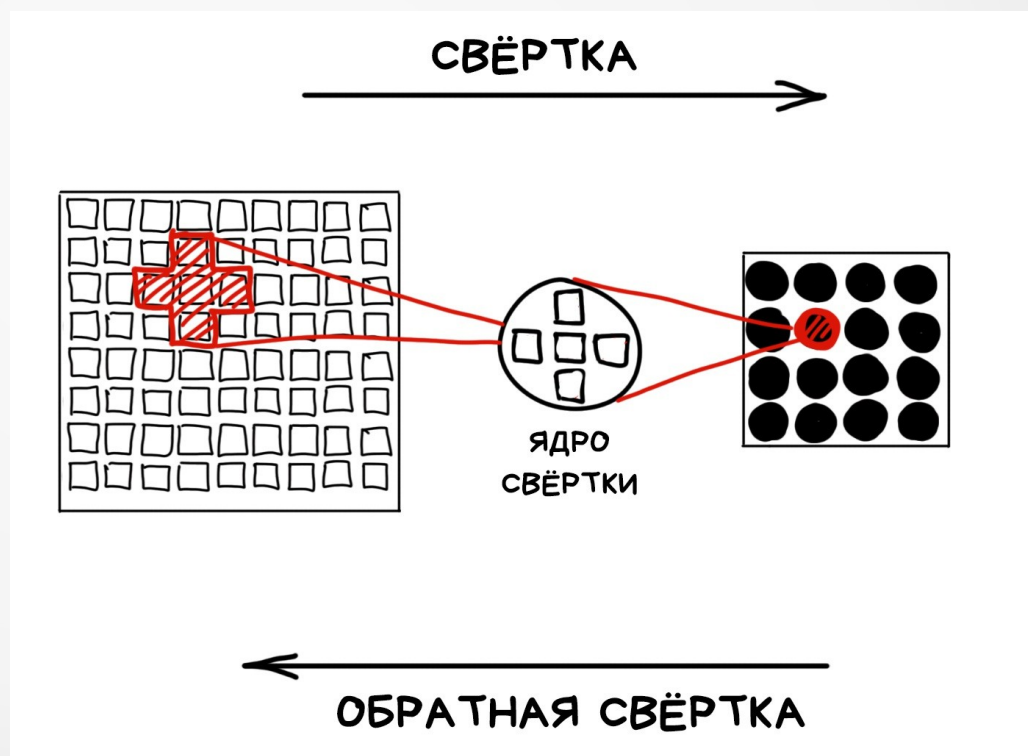
размер входного изображения = размеру входной карты

сравниваем выходные карты поточечно,
для каждой точки определяем карту-победителя

Свёрточные нейросети



изображение обрабатывается свёрточными слоями
на выходе выполняем обратную свёртку



Свёрточные нейросети

FCN: Fully Convolutional Networks

Пример — ищем людей на картинке (датасет Pascal VOC)

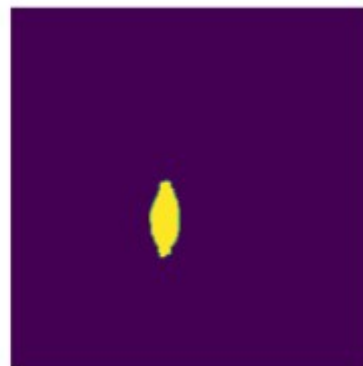
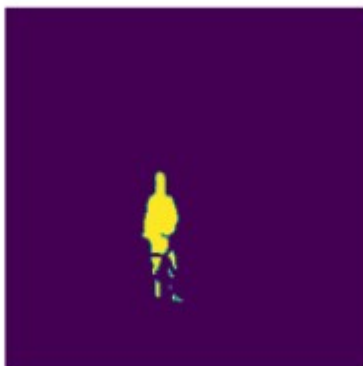
картинка



разметка



результат FCN



Свёрточные нейросети: литература

git clone https://github.com/mechanoid5/ml_lectorium.git

Борисов Е.С. Детектор объектов для неподвижных камер.
<http://mechanoid.su/cv-backgr.html>

Борисов Е.С. О задаче поиска объекта на изображении.
<http://mechanoid.su/cv-image-detector.html>

Борисов Е.С. Классификатор изображений на основе свёрточной сети. -- <http://mechanoid.su/ml-lenet.html>

Fully Convolutional Networks for Semantic Segmentation
<https://arxiv.org/pdf/1411.4038.pdf>

FCN — Fully Convolutional Network (Semantic Segmentation)
<https://towardsdatascience.com/review-fcn-semantic-segmentation-eb8c9b50d2d1>

Vladimir Yuzhikov Восстановление расфокусированных и смазанных изображений
<https://habr.com/ru/post/136853/>