



Государственный научно-исследовательский институт авиационных систем

Семантическое обнаружение объектов на изображениях с использованием ГКНС

*Горбачевич В. С.**

Визильтер Ю. В.

Моисеенко А. С.

One-shot Learning



One-shot Learning – Обучение по одному предъявлению

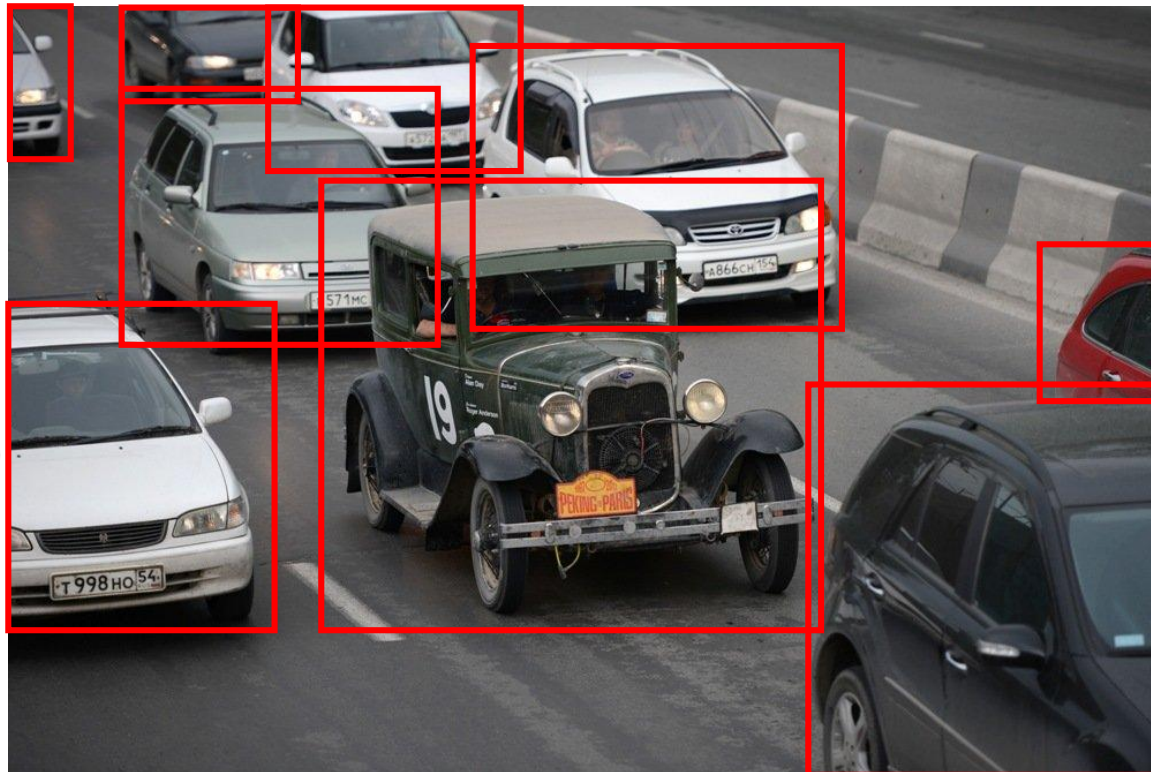


Brenden M. Lake, Ruslan Salakhutdinov, Jason Gross, and Joshua B. Tenenbaum. “One shot learning of simple visual concepts”, Proc. 33th Annual Meeting of the Cognitive Science Society(2011).

Семантическое обнаружение



**Запрос
(эталонное изображение)**



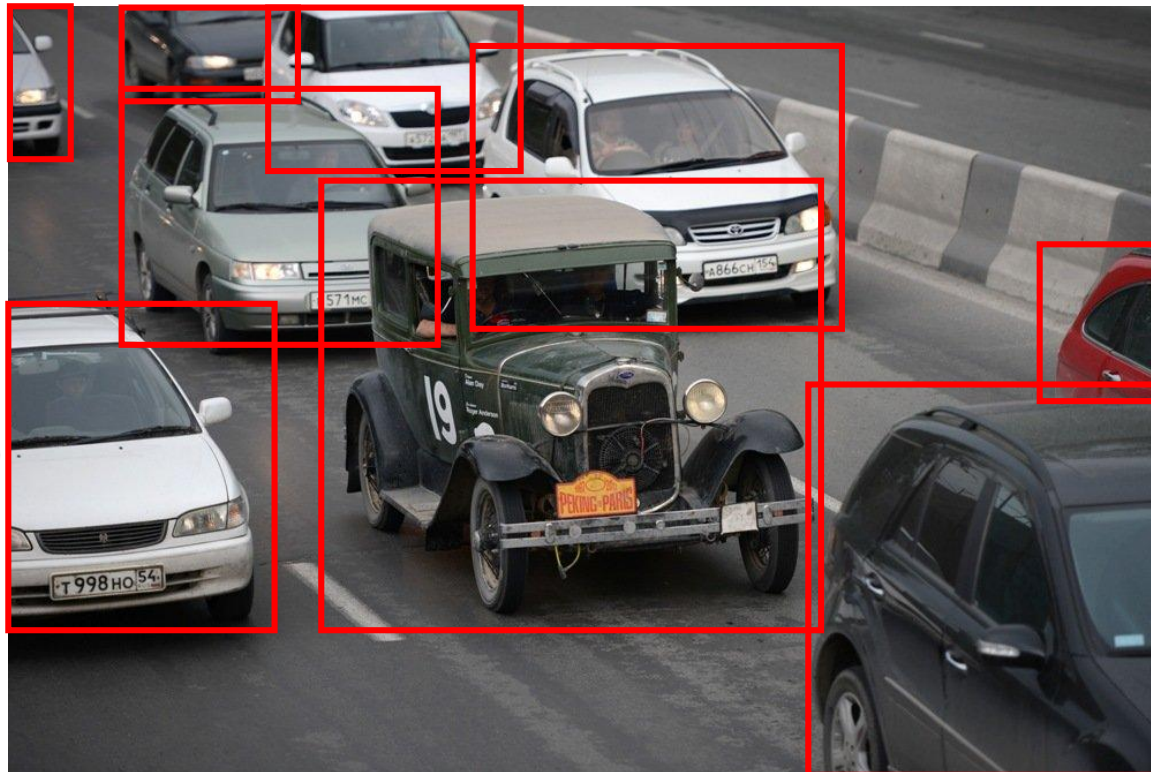
устойчивость к масштабу / устойчивость к ракурсу / устойчивость к условиям освещения

Под семантическим обнаружением объектов на изображении будем понимать поиск и локализацию объектов, соответствующих семантическому классу эталонного изображения

Семантическое обнаружение



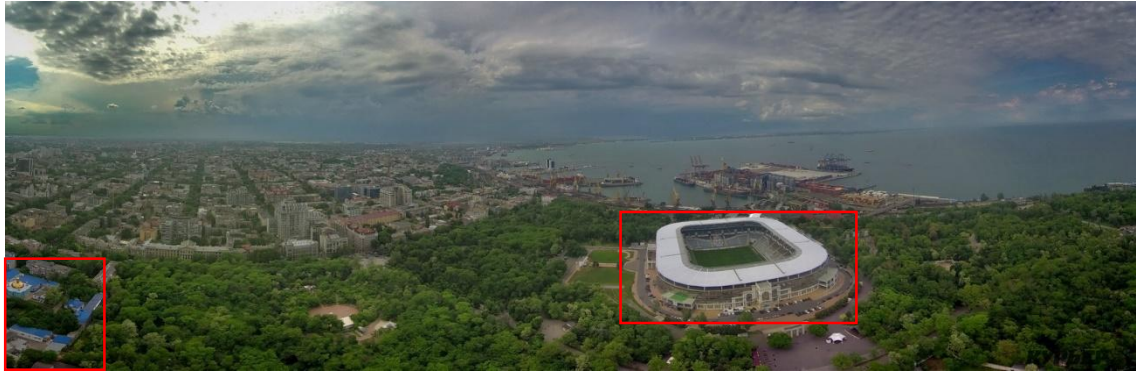
**Запрос
(эталонное изображение)**



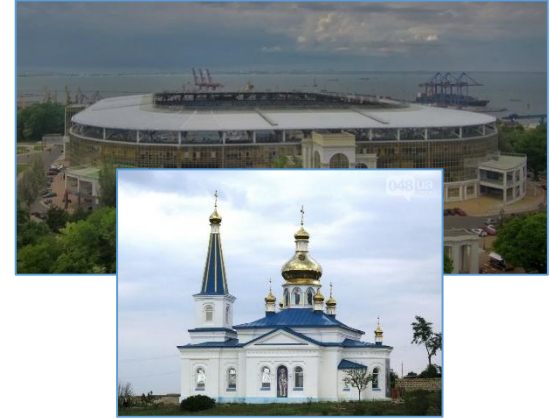
устойчивость к масштабу / устойчивость к ракурсу / устойчивость к условиям освещения

Под семантическим обнаружением объектов на изображении будем понимать поиск и локализацию объектов, соответствующих семантическому классу эталонного изображения

Семантическое обнаружение



Ориентирование

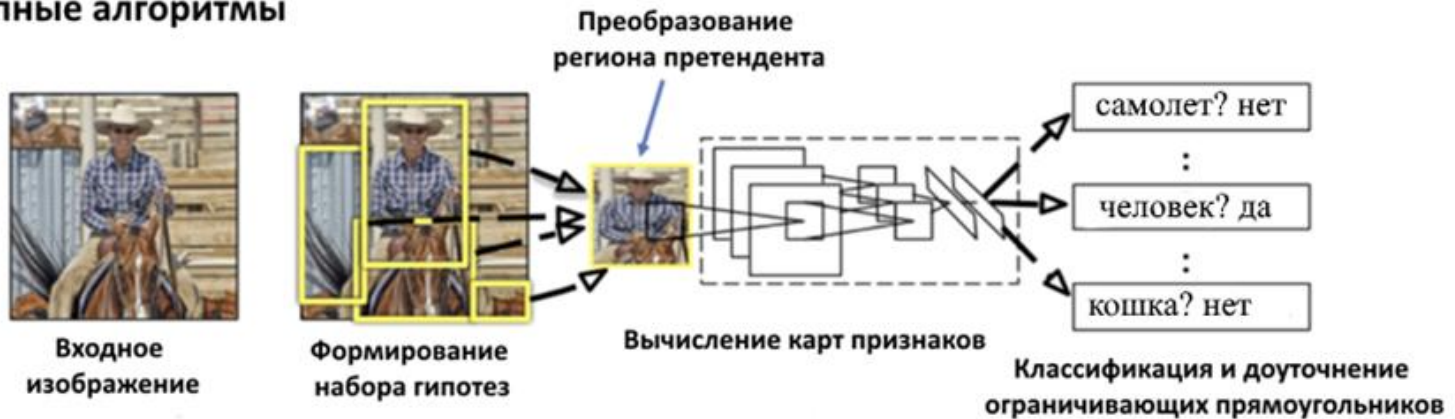


Поиск в коллекциях



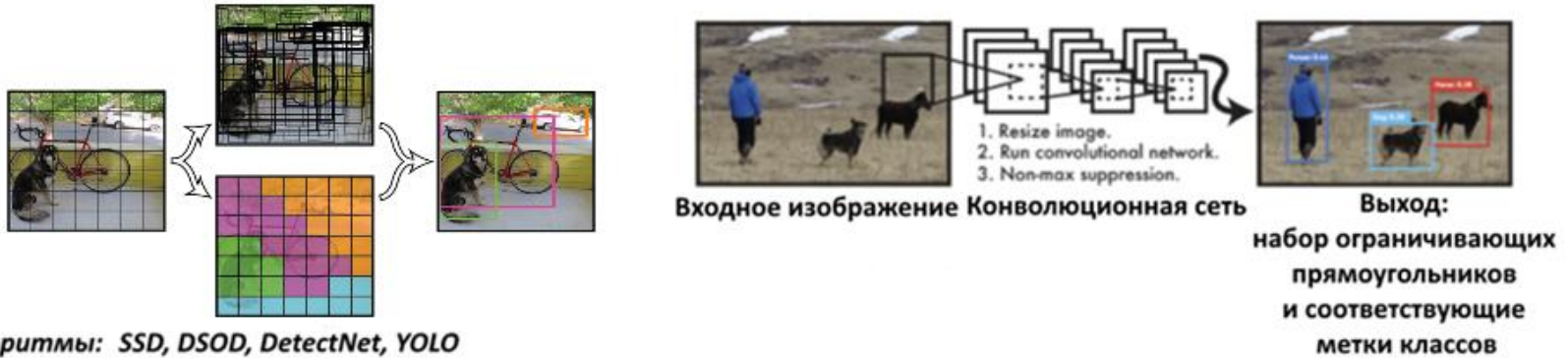
Подходы к обнаружению объектов

Двухэтапные алгоритмы



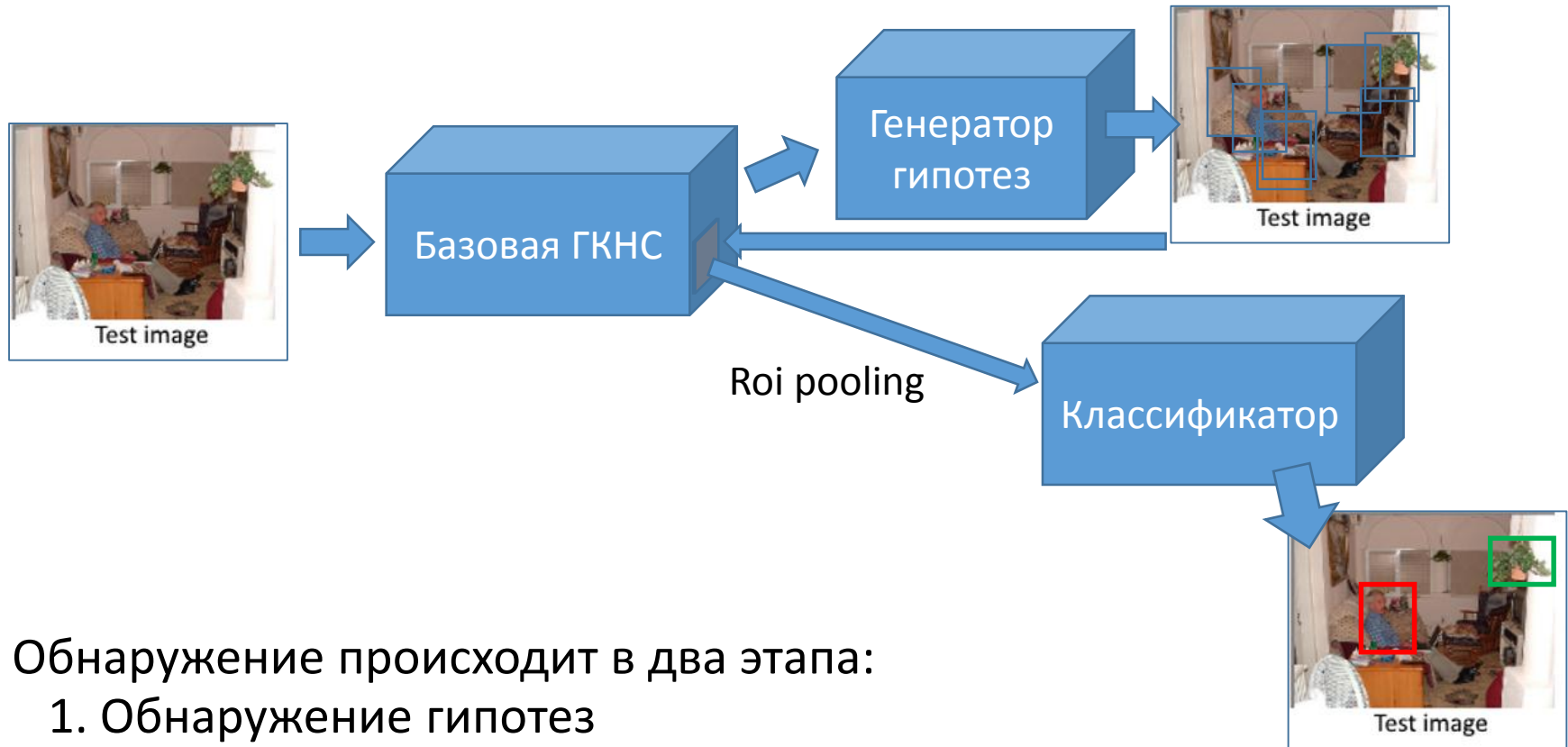
Алгоритмы: *Faster R-CNN, R-FPN, Light-Head R-CNN*

Одноэтапные алгоритмы



Алгоритмы: *SSD, DSOD, DetectNet, YOLO*

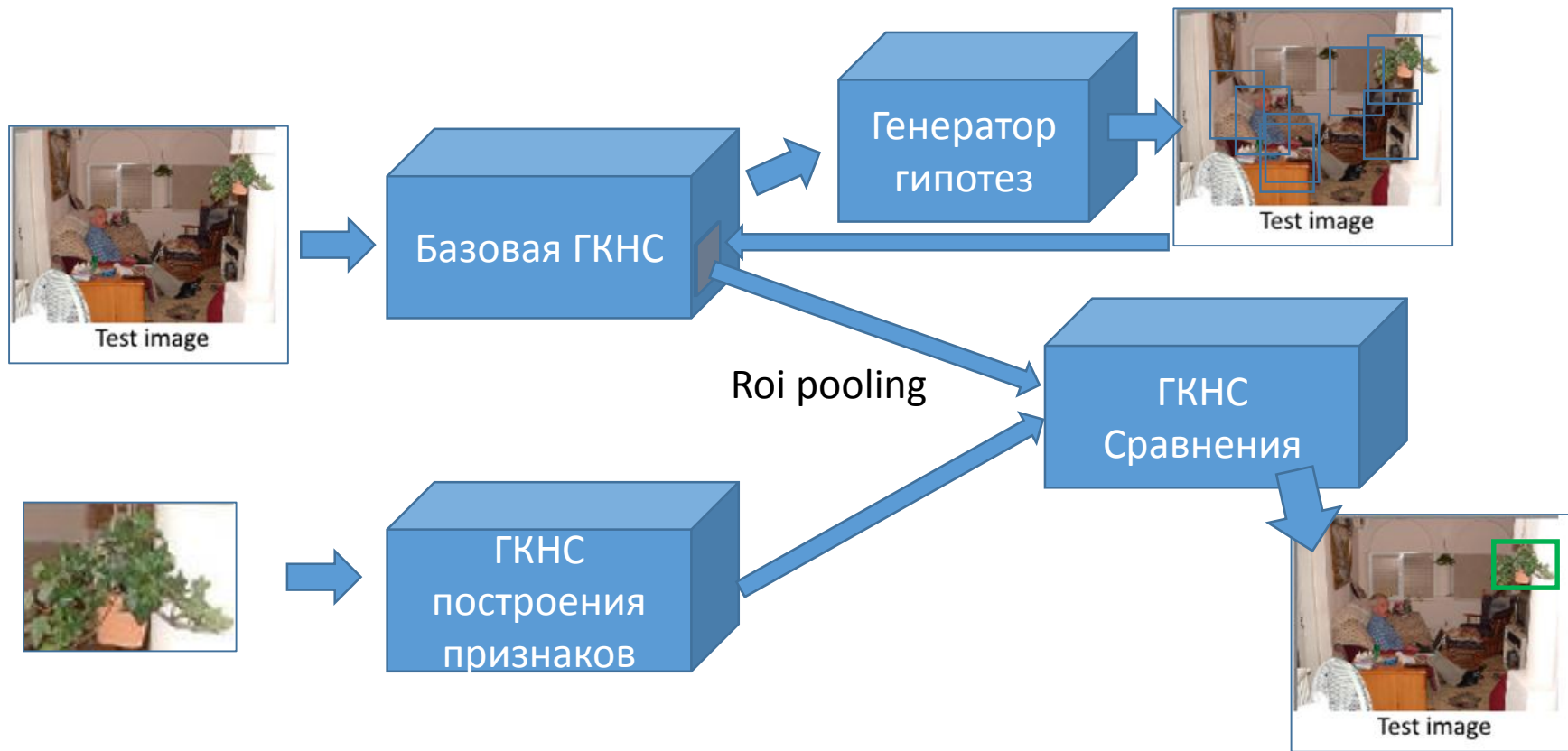
Алгоритм Faster R-CNN



Обнаружение происходит в два этапа:

1. Обнаружение гипотез
2. Получение векторов признаков
3. Классификация и доуточнение

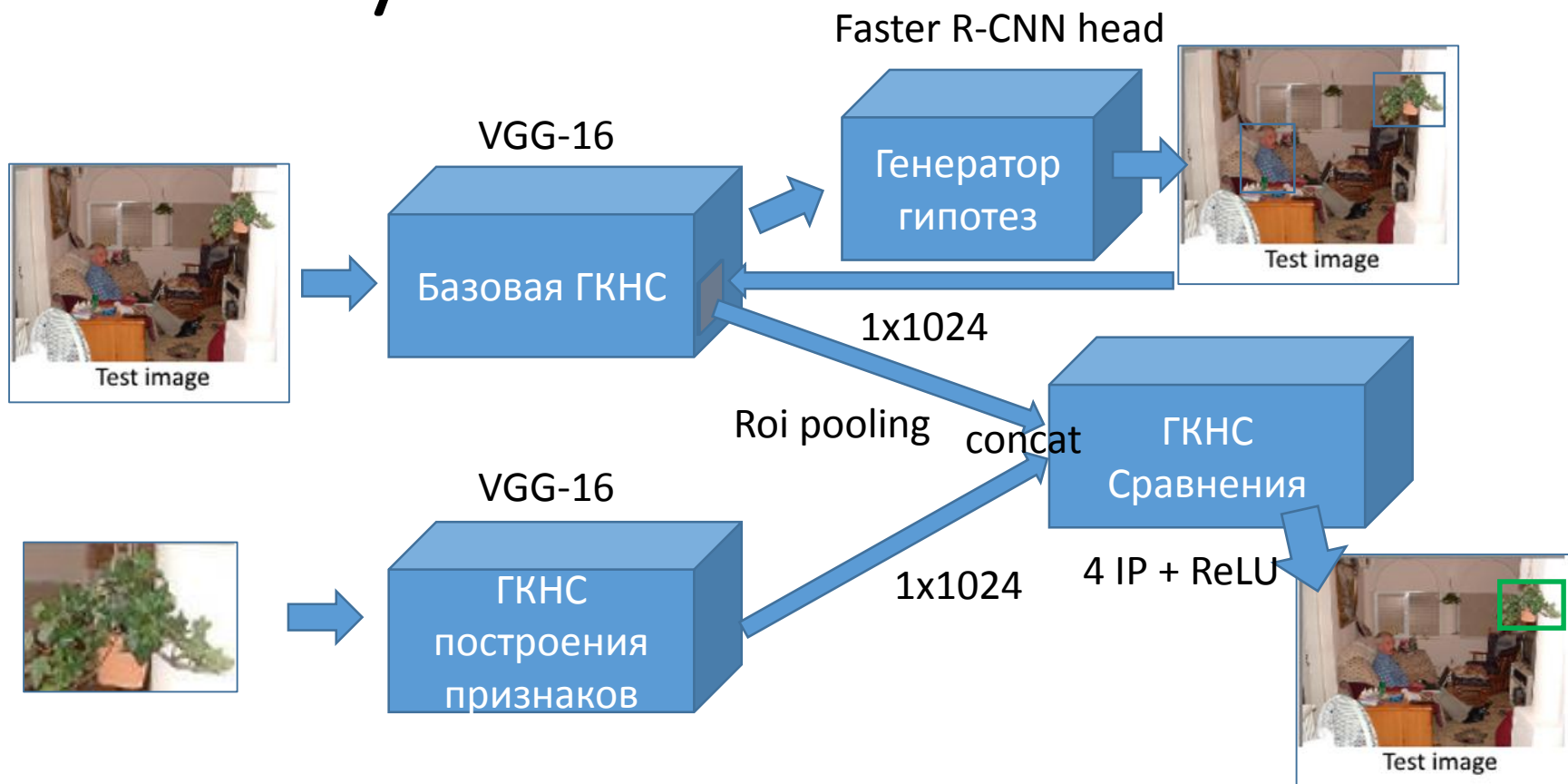
Алгоритм семантического обнаружения



Обнаружение происходит в два этапа:

1. Обнаружение гипотез
2. Генерация вектора признаков “запроса”
2. Получение векторов признаков
3. Сравнение и доуточнение

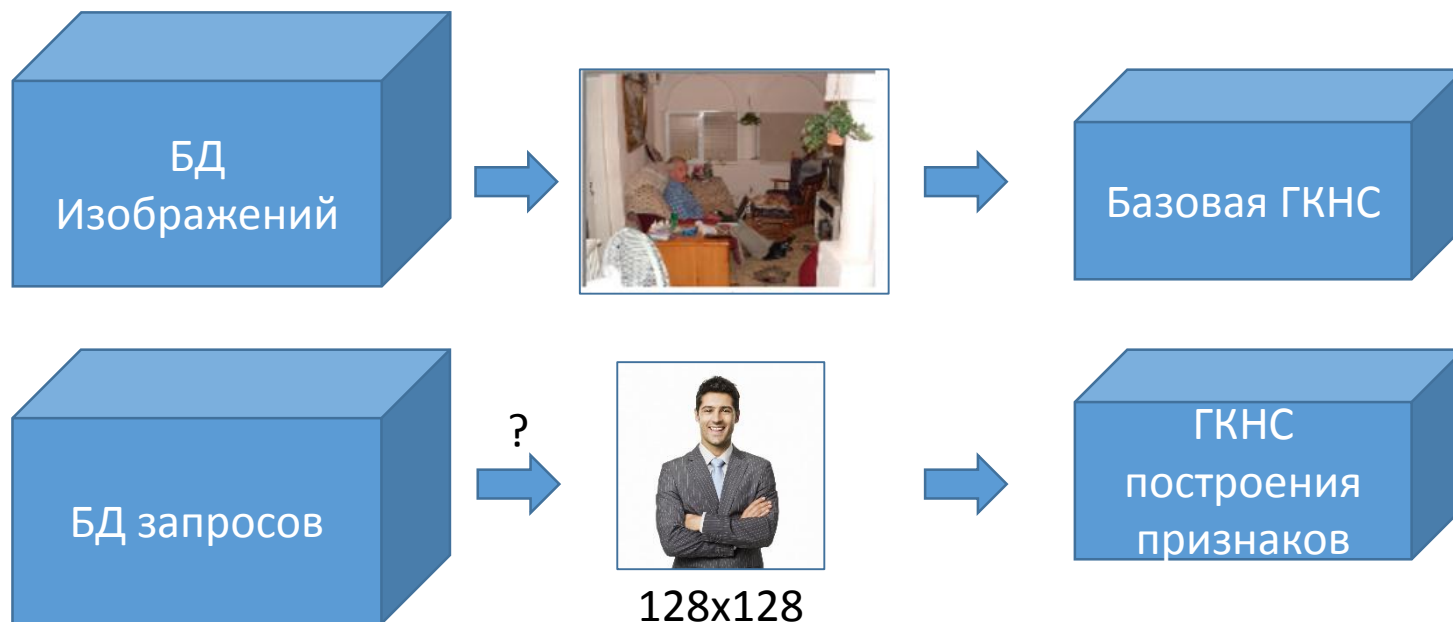
Схема обучения



$$Loss(\{p_i, t_i\}) = \frac{1}{N_{cls}} \sum_i \text{LogLoss}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* \text{smoothL1}(t_i - t_i^*)$$

- i – индекс опорного прямоугольника в мини-батче
- p_i – вероятность того, что i -ый опорный прямоугольник является объектом
- p_i^* – предсказанная вероятность того, что i -ый опорный прямоугольник является объектом
- t_i – параметризованные координаты предсказанных описывающих прям-ков
- t_i^* – заданные описывающие прямоугольники (разметка)
- N_{cls} – размер мини-батча
- N_{reg} – количество положений анкеров
- λ – балансирующий параметр

Схема обучения



БД Pascal VOC 2007:
20 классов
24640 объектов
9963 изображения

Обучение в два этапа:
1. На ограниченной БД запросов
2. На полной базе запросов

Метрики качества

$$\text{Precision} = \frac{TP}{TP + FP}$$

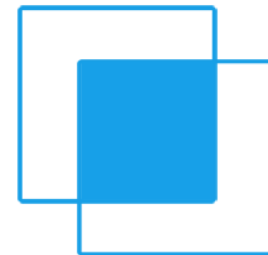
$$\text{Recall} = \frac{TP}{TP + FN}$$

TP – число верных обнаружений

FN – число не обнаруженных объектов

FP – число ложных обнаружений

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$



Если $\text{IoU} > t$ – объект
обнаружен верно ($t = 0.5$)

AP (average precision)

$$AP = \frac{1}{11} \sum_{r \in \{0, 0.1, \dots, 1.0\}} \max_{\tilde{r} \geq r} p(\tilde{r})$$

mAP (mean average precision)

$$mAP = \frac{\sum_{i=0}^{n-1} AP_i}{n}$$

Результаты на БД PASCAL VOC 2007

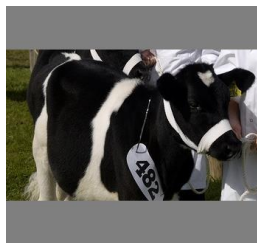
Результаты работы на классах, участвующих в обучении (AP)

bike	bird	bottle	bus	car	chair	table	dog	horse	mbike	person	plant	sheep	sofa	train	tv	mAP(%)
0.72	0.53	0.51	0.73	0.8	0.44	0.59	0.69	0.76	0.71	0.72	0.35	0.68	0.62	0.69	0.59	0.63

Результаты работы на классах, не участвующих в обучении (AP)

aero	boat	cat	cow	mAP(%)
0.25	0.11	0.66	0.73	0.44

128x128 эталонное изображение



Тестовое изображение

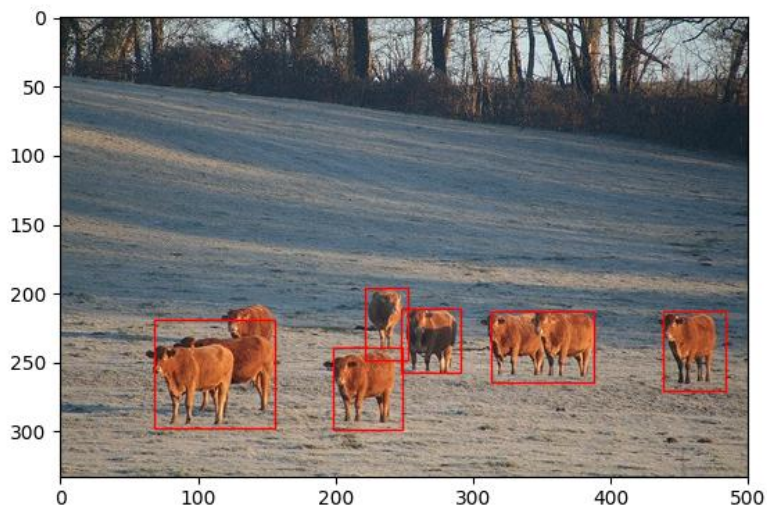
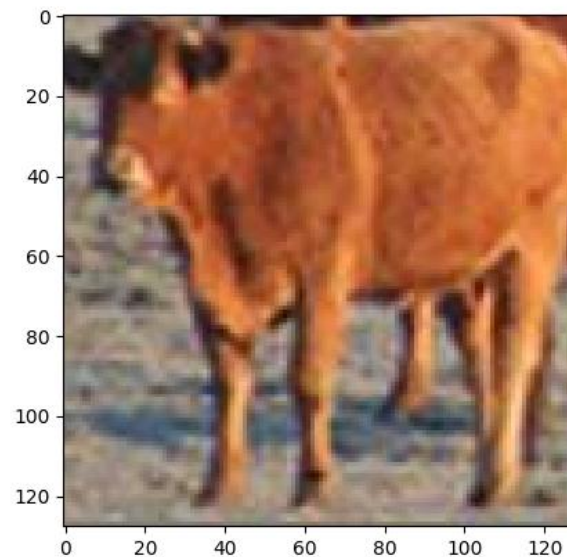
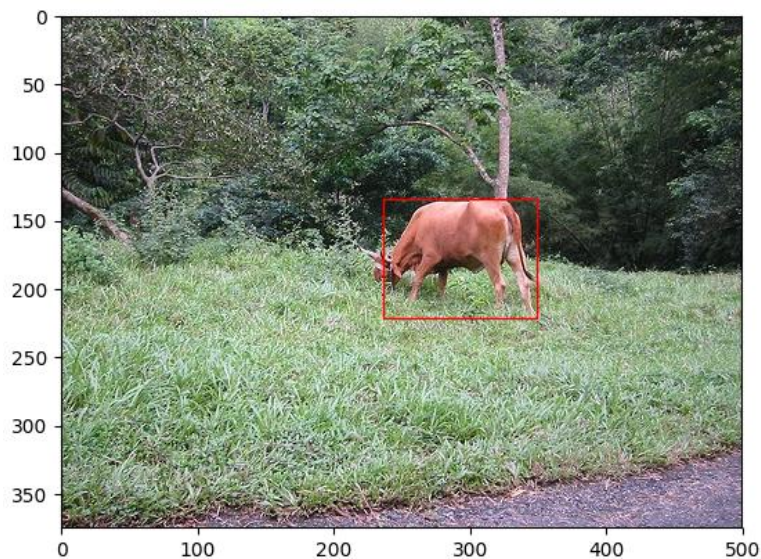


Результаты на БД PASCAL VOC 2007

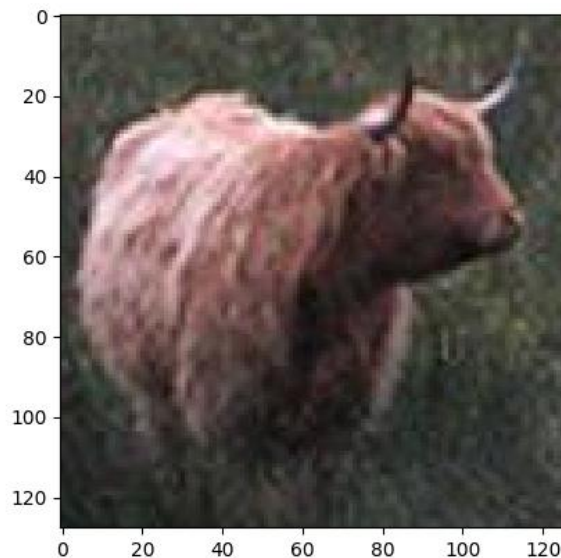
Алгоритм	Seen	Unseen
Our	63	44
SSSM	~	21
SiamRPN	9.6	14.2
SiamFC	15.1	13.3
C-RPN	27.4	32.1
Low Shot(n=1)		
LSTD	70	38
MetaLearner*	~	~20



Примеры работы

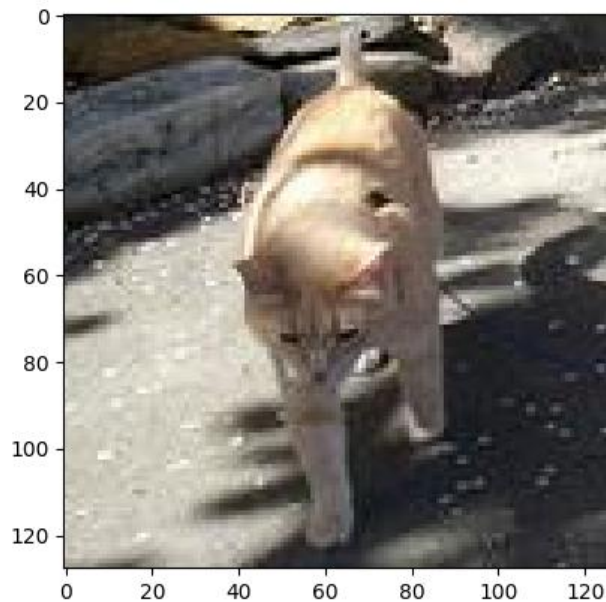
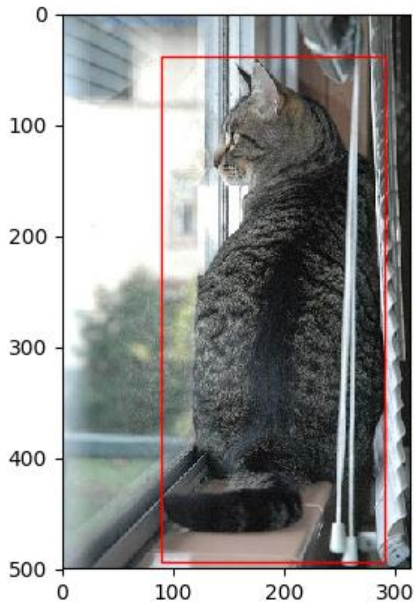
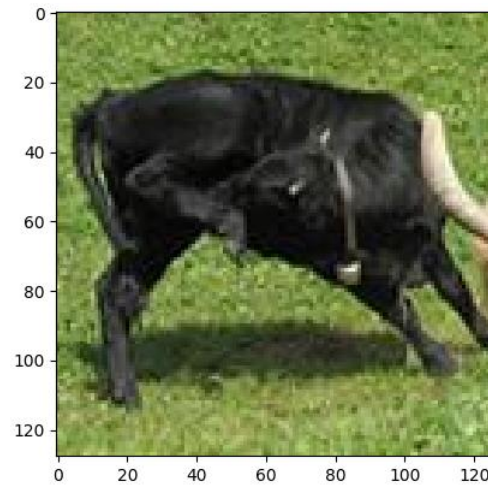
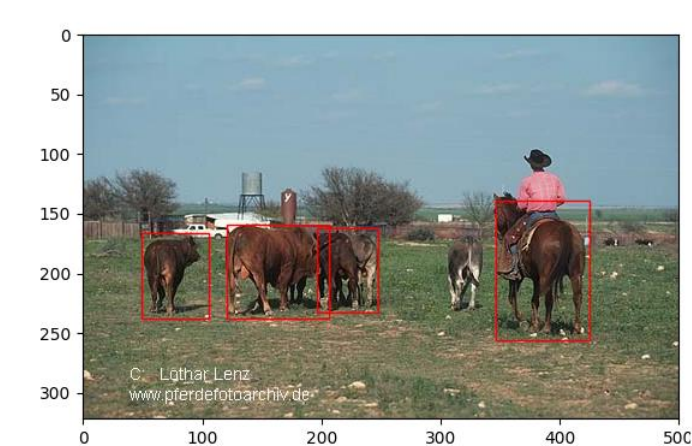


Тестовое
изображение



Эталонное изображение

Примеры работы



Тестовое изображение

Эталонное изображение