



Курсовая работа

Детекция объектов по нескольким примерам без дообучения

Выполнили:

Студенты группы мФТиАД21
Подчезерцев Алексей Евгеньевич
Самоделкина Мария Владимировна

Руководитель:

Приглашенный преподаватель
Озерин Алексей Юрьевич

СОДЕРЖАНИЕ

Актуальность, постановка задачи

Основные понятия

Детекция объектов по произвольному запросу

Демонстрационные материалы

Задача регрессии ключевых точек



Актуальность

Пайплайн обучения детекционной модели



Сбор датасета

Разметка датасета

Обучение модели

- Затраты на сбор датасета
- Удешевление и улучшение эффективности пайплайна



Постановка задачи

Разработка подхода к детекции объектов по произвольному текстовому запросу

- реализация предлагаемого подхода и выбор наиболее оптимальных моделей
- демонстрация результатов работы и улучшение результатов выдачи
- решение задачи регрессии ключевых точек с дообучением по нескольким примерам

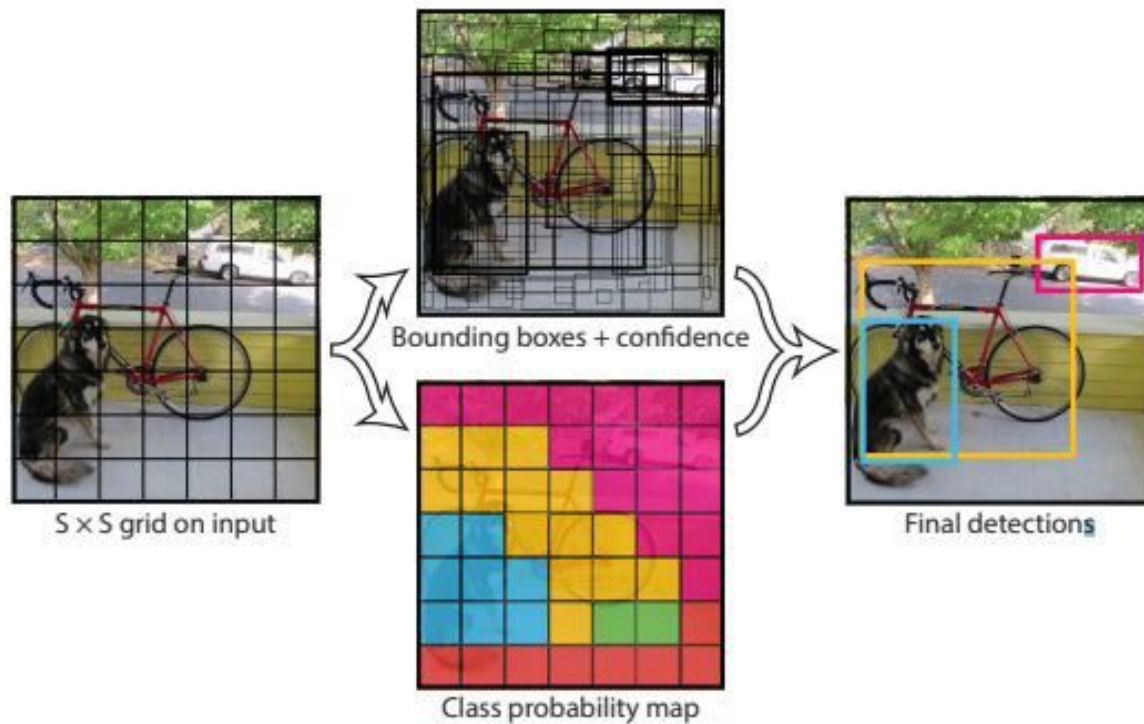


Основные понятия

Задача детекции объектов на изображении

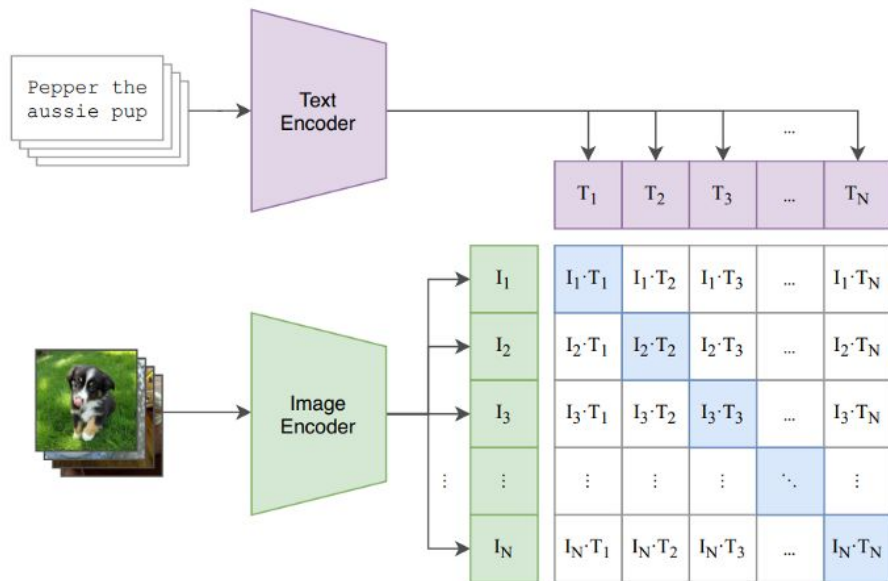
- определение прямоугольной области, ограничивающей объект
- классификация выделенной области
- нахождение ключевых точек

Основные понятия - YOLO

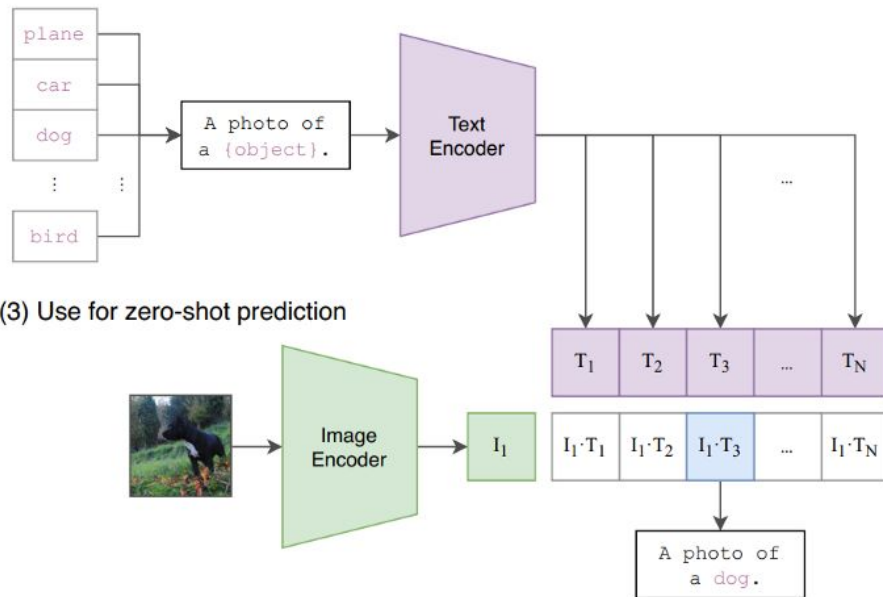


Основные понятия - CLIP

(1) Contrastive pre-training

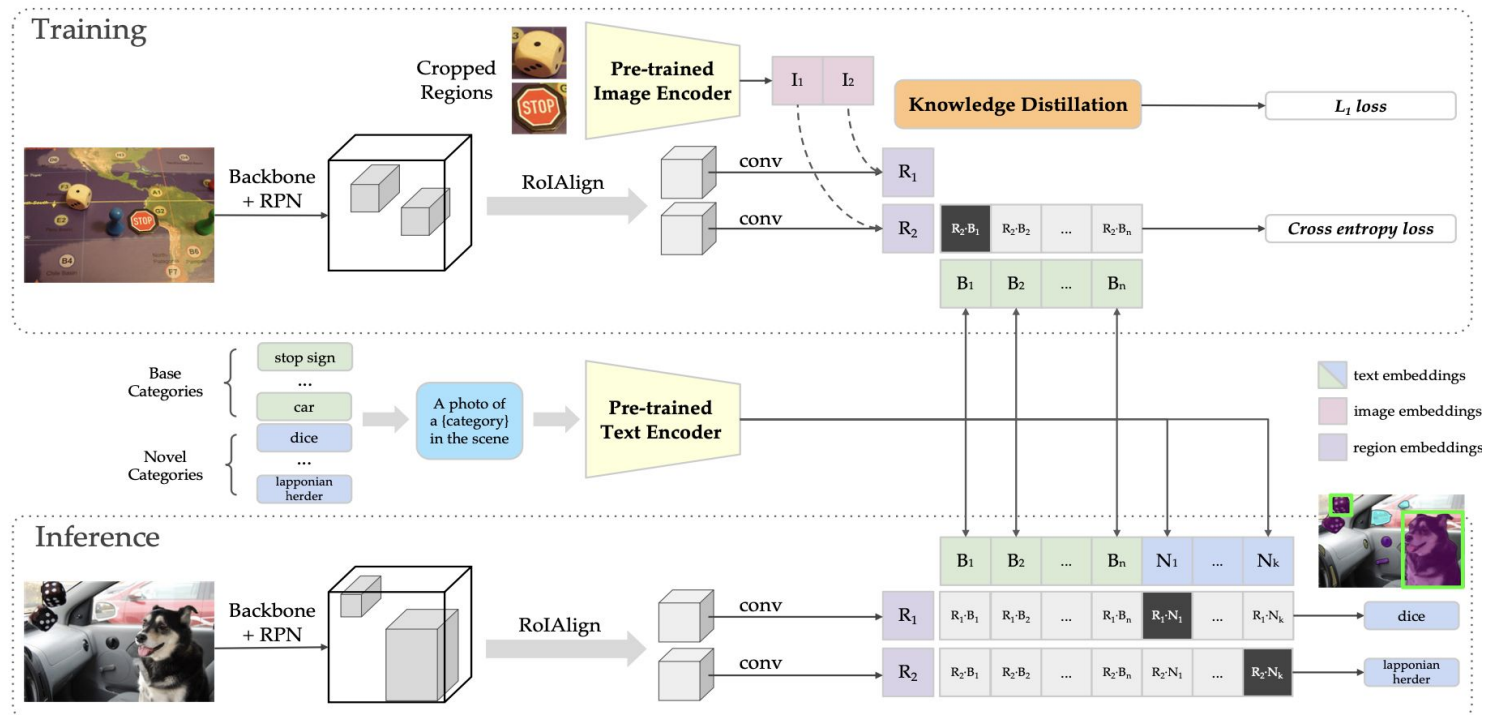


(2) Create dataset classifier from label text



(3) Use for zero-shot prediction

Основные понятия - ViLD



Источник: [Open-vocabulary Object Detection via Vision and Language Knowledge Distillation](#)

Подход к детекции объектов по произвольному запросу



Пайплайн детекции

1) получение регионов для всего доступного множества изображений

- YOLO, ViLD
- модифицированный NMS

2) вычисление векторного представления регионов и текстового запроса

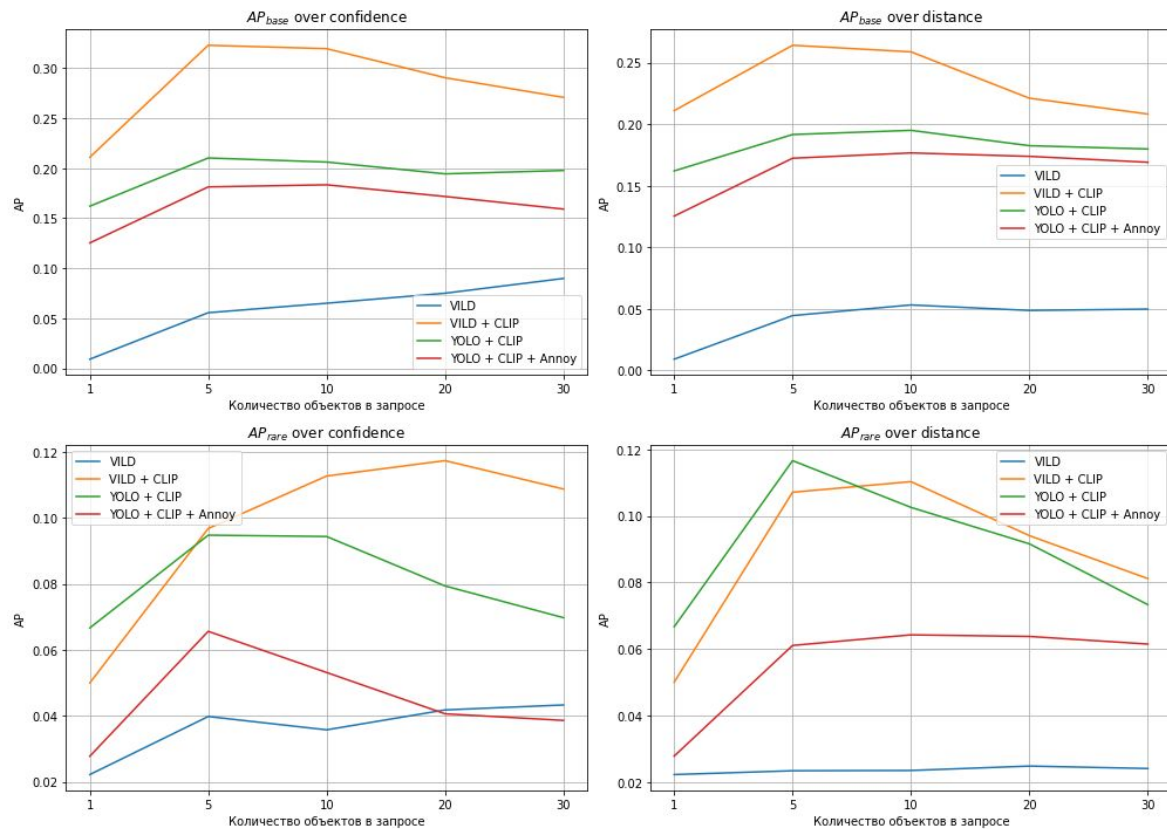
- CLIP, ViLD

3) поиск ближайших регионов к текстовому запросу

- HNSW, Annoy, ScaNN

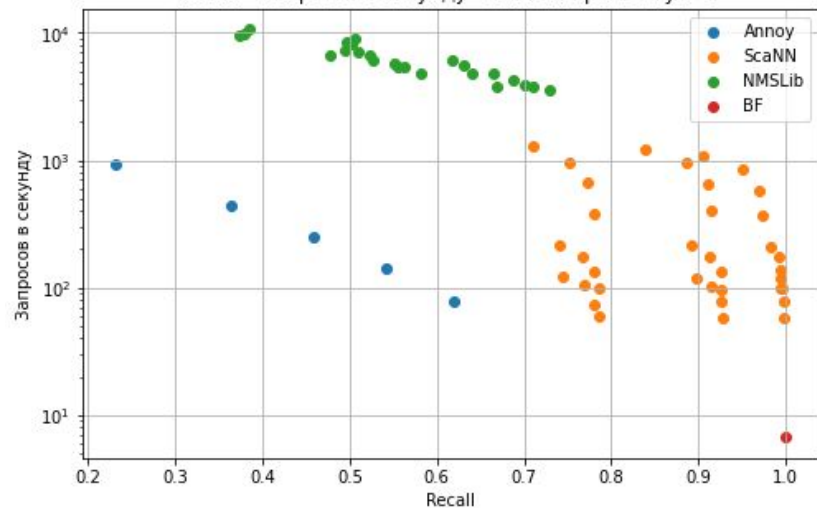
AP в разрезе классов и меры

$$AP = \sum_t (R_t - R_{t-1}) P_t$$

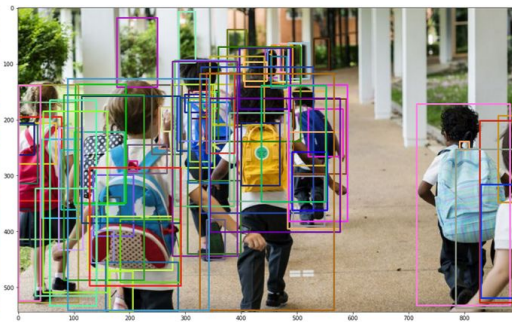


Результаты экспериментов

Recall - запросов в секунду - выше и правее лучше



(a) Результат кластеризации регионов



(b) Результат применения NMS к кластерам

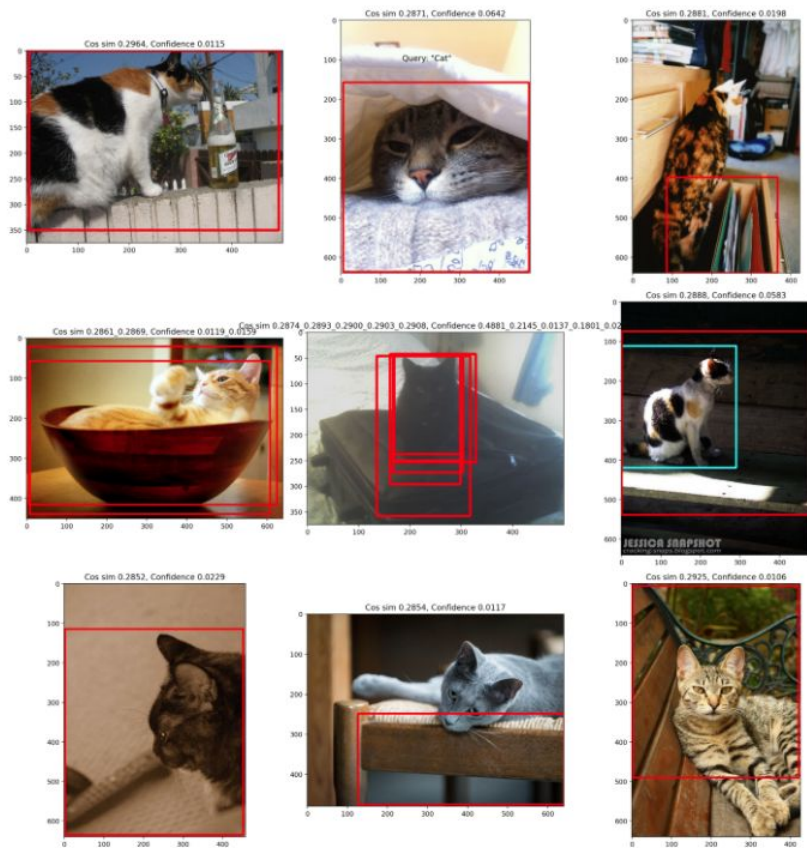
Демо

Page:

Query demo

Query text

Cat





Улучшение результатов выдачи

- Разметка пользователем результатов работы алгоритма
- Формирование датасета, пригодного для обучения модели YOLO
- Дообучение модели на пользовательском датасете

Задача регрессии ключевых точек

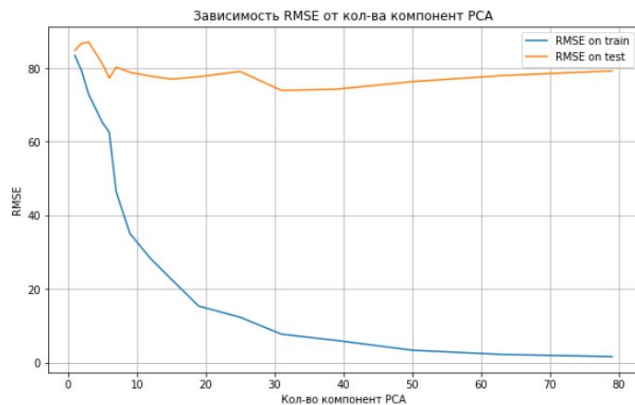


Задача регрессии ключевых точек

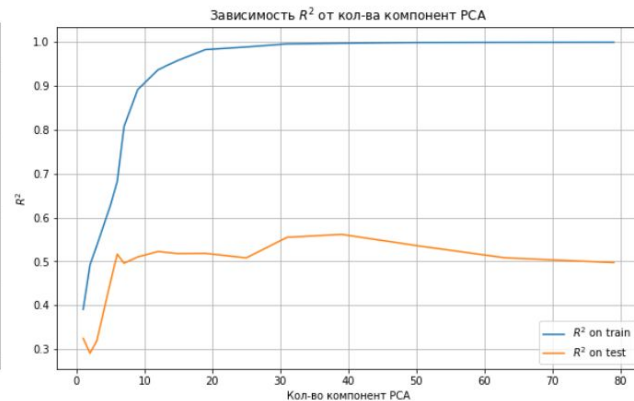
- получение карты признаков с помощью предобученной модели (YOLO)
- обучение нескольких моделей линейной регрессии для нахождения каждой ключевой точки

Проведение экспериментов

- Количество компонент PCA
- Отклонение от центрального вектора карты признаков
- Типы и кол-во аугментаций
 - Цвет, яркость, поворот, кадрирование
- Кол-во изображений в обучающей выборке



(a) Зависимость RMSE от количества компонент PCA

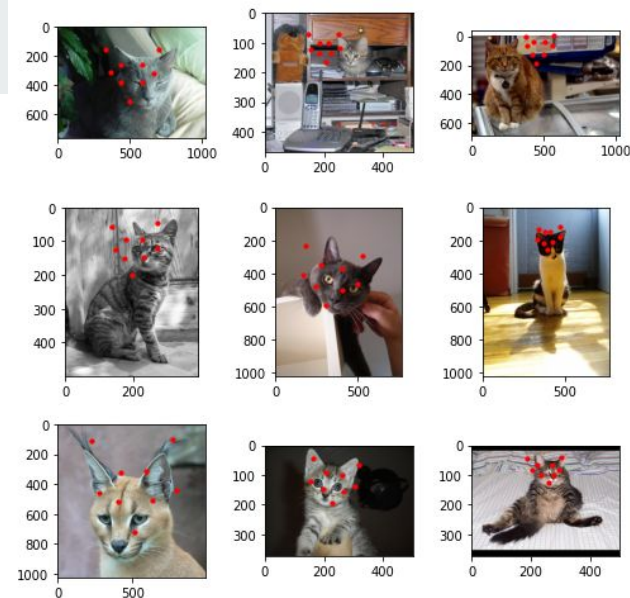


(b) Зависимость R^2 от количества компонент PCA

Результаты экспериментов

Наилучшее качество по метрикам RMSE и R2

- отсутствие аугментаций
- отклонение от центрального вектора - 1
- минимальный размер обучающей выборки - 6



Эксперимент с типами аугментаций

Эксперимент	Медиана PCA	RMSE на обучении	RMSE на тесте	R^2 на обучении	R^2 на тесте
Без аугментаций	9	40±29	84±22	0,81±0,19	0,46±0,20
Аугментации цвета и яркости	7	42±28	84±22	0,80±0,19	0,45±0,20
Аугментации цвета, яркости и поворота	12	39±26	85±22	0,83±0,17	0,44±0,22
Аугментации цвета, яркости, поворота и кадрирования	12	40±24	86±21	0,82±0,17	0,44±0,19

Подведение итогов



Сложности исследования

- Предобработка изображения YOLO
- Итоговое качество зависит от метрик ANN



Результаты

- Реализован прототип для дешевого и быстрого:
 - Сбора датасета для обучения моделей
 - Дообучения модели на полученной разметке
- Проведены эксперименты по регрессии ключевых точек
 - Не удалось добиться заметного успеха



Вклад в работу

Подчерцев Алексей,
aepodchezertsev@edu.hse.ru

- Проведение экспериментов к решению задачи детекции
- Создание демонстрационных материалов: сайт и jupyter-notebook

Самоделкина Мария, mvsamodelkina@edu.hse.ru

- Разработка пайплайна к решению задачи детекции
- Решение задачи регрессии ключевых точек и проведение экспериментов

github.com/AsciiShell/hse-fsod

