Слайд 1.

Здравствуйте, я Илья Белоножко из 241-й группы. Тема моей курсовой работы - Сравнительный анализ методов компьютерного зрения на примере задачи поиска лиц на изображении. Научный руководитель – Казачкова Анна Андреевна

Слайд 2.

Целью работы был сравнительный анализ способов обнаружения лиц на изображении. Я решил сравнить классические методы и один из современных. Задачи проговаривать не буду, далее подробнее расскажу о проделангой работе

Слайд 3.

Сначала была проведена работа по реализации классических методов. Все оные строились по единому принципу: создается метод, который каким-либо образом извлекает из изображения вектор. С помощью этого метода создаются 2 списка с признаками и метками, указывающими на принадлежность к классу лиц. Признаки могут извлекаться долго, поэтому я решил воспользоваться многопоточной обработкой. С помощью ThreadPoolExecutor обрабатываются все входные изображения. И основная модель обучается на входных данных по входным признаком бинарно классифицировать его.

Слайд 4.

Для такого подхода понадобились 2 набора данных: с изображениями, содержащими только лица и с изображениями, содержащие что угодно, кроме лиц. Я выбрал эти, так как они подходят критериям. Во втором датасете было всего 5000 изображений, поэтому пришлось сократить первый. Для искусственного увеличения данных все эти изображения были отзеркалены и итоговый датасет из 20000 изображений был сохранен таким образом.

Слайд 5.

Первая модель использует адаптивный бустинг для классификации и метод Виолы-Джонса для извлечения признаков. Он был создан в 2001 году и основан на расчете разности яркостей в определенных областях. Области называются признаками Хаара. Вот как наглядно он работает –(внимание на слайд)

Слайд 6.

Вторая модель использует построение гистограммы направленных градиентов на входном изображении и классифицирует ее с помощью метода опорных векторов. Подход был предложен в 2005 году. Он вычисляет в каждой небольшой ячейки изображения направление и величину изменения яркости. Для каждой такой ячейки строится гистограмма направлений, которые позже преобразуются в один выходной вектор. С помощью этих векторов SVM и ищет гиперплоскость, чтобы относить изображения к классам.

Слайд 7.

Третий метод предназначен не сколько для обнаружения чего-либо, сколько для поиска ключевых точек изображения. Был предложен в 2004 году. Масштабно-инвариантная трансформация признаков с помощью нескольких сверток находит устойчивые точки, называемые ключевыми. И в их окрестностях строится вектор **признаков, также учитывающий направление и величину граидентов**. Таким образом получится список векторов, имеющий произвольную длину. Для приведения к фиксированному используем метод k-means и дляч каждого изображения строим гистограмму принадлежности к кластерам. Эти гистограммы классифицирует SVM

Слайд 8

Свёрточная нейронная сеть использует множество свёрточных фильтров для извлечения низкоуровневых и высокоуровневых признаков из изображения, таких как края, текстуры и формы. Далее генерируются области интереса, предполагающие местоположения целевых объектов. Для каждой области сеть оценивает вероятность наличия лица и уточняет координаты ограничивающего прямоугольника с помощью регрессии. В процессе обучения корректируются веса сети для повышения точности классификации. Из-за такого метода обучения я выбрал другой датасет, где лица находятся в контексте

Слайд 9.

Из-за большого размера выходных векторов при методе Виолы-Джонса получилось обучить первую модель лишь на 4000 изображений. Для тестирования были найдены несколько изображений, сложность обнаружения лица на них возрастает. Она работает с переменным успехом, видно, что метод очень чувствителен к поворотам головы и падения тени.

Слайд 10

Вторую и третью модели получилось обучить на большем количестве данных. HOG показывает неплохие результаты, но скорее всего из-за малого разнообразия исходного датасета результаты не лучшие.

Слайд 11

Также для честного сравнения я попробовал обучить ее на 4000 данных и она работает явно хуже

Слайд 12

Третий метод работает не так, как я предполагал, скорее всего из-за того, что он предназначен для других задач. Результаты обнаруджения плачевные на 20000 обучающих

Слайд 13

На 4000 данных всё также крайне печально

Слайд 14

А вот свёрточная нейронная сеть даже на 1000 обучающих изображений и 3-х эпохах справилась хорошо, все лица определены верно, возникают проблемы только с множественными прямоугольниками

Слайд 15.

Источники информации, использованные в работе

Слайд 16

Спасибо за внимание, готов ответить на ваши вопросы