**Отчет по тесту для BND LLC**

**1. Решение задачи**

По задачи надо было написать код на python который выполняет дедукцию людей на исходном видео и их отрисовку.

Под словом отрисовка я понял, что нужно просто отрисовать боксы задетекченых людей, а не сигментировать их по контору.

Для решения задачи я воспользовался нейронной сетью YOLOv11 так как она считается удобной для решения задач детекции в CV. При использовании Yolo есть два подхода к решению этой задачи:

* взять уже обученную модель, зашитую в yolo которая умеет распознавать людей;
* обучить свою модель;

Я попробовал сначала обучить свою нейронку на готовом размеченном датасете с людьми. В датасете было 300 фотографий и для обучения я взял 250 эпох. Нейронка получилась не очень так как она находила не всех людей и часто распознавала другие объекты как людей. Предполагаю, что так получилось из-за того, что мало обучал неиросеть и датасет не полностью состоял из записей камер наружного наблюдения, а лишь процентов на 40 (остальной частью датасета были записи с веб-камер, любительские съёмки на телефон и тд.).

Далее я попробовал готовую модель yolo и она справилась лучше но не полностью. Она все еще не распознает людей из далека и не распознает людей, загороженных объектами.

Для решение первой проблемы я нашел патч на гитхабе который увеличивает количество детекций путем увеличения прогона через нейронку каждого кадра и объединения результатов этих прогонов (<https://github.com/Koldim2001/YOLO-Patch-Based-Inference>). Это действительно решило проблему не нахождения людей, но закрытые люди все еще не детектятся, а картины людей стали распознаваться чаще как люди. Также минусом этого решения стало увеличение времени обработки видео (с 37 сек до 4 мин) что критично если подразумевается, что нейронная сеть должна работать в реальном времени.

**2. Результат и пути его улучшения**

По результатам работы я получил 2 решения задачи: первое с использованием патча, а второе без него. В код я завернул решение с использованием патча так как оно лучше находит людей на видео, но для сравнения я оставлю на гите 2 видео (с ним и без него). Без него программа обрабатывает видео за 37 сек.

Полученное решение можно улучшить 2 путями:

Первое это найти хороший датасет с разметкой людей на камерах видеонаблюдения и обучить на нем нейросеть. Возможно посмотреть в сторону других сеток так как проблема нахождения мелких объектов может оказаться болячкой yolo которая уже решена в других моделях.

Второй подход — это найти другие патчи которые решают проблему нахождения мелких и закрытых объектов. Скорее всего они также будут жрать скорость работы сетки, но можно поискать такой патч который будет делать это не сильно.

Третий путь — это сильно подумать и придумать свой способ улучшить модель йолы.