**1 слайд:**

Здравствуйте, уважаемая комиссия, меня зовут Есенов Константин, студент группы СМ7-23М. Сегодня я представляю вам свою НИР на тему: "**Разработка системы детекции для автономного транспорта**".

**2 слайд:**

В прошлом семестре НИР была посвящена теме «Применение системы мультиобъектного трекинга в системах помощи водителю, таких как адаптивный круиз-контроль и автоматическое экстренное торможение».

Согласно плану ВКРМ, целью этого семестра является Разработка системы детекции нескольких объектов в реальном времени для автономного транспорта.

Задачи, решаемые в рамках данной НИР следующие:

1. Необходимо выбрать нейросеть для детекции нескольких объектов, а также датасета, который хорошо подойдет для обучения в сфере автономного вождения.

2. Обучить выбранную модель нейросети на данных с выбранного датасета

3. Проанализировать полученный результат, сделать выводы.

Основным требованием является скорость обработки: она должна составлять не менее 25 кадров в секунду (fps);

**3 слайд:**

По итогам анализа методов детекции в предыдущей НИР, была выбрана нейросеть YOLO (you only look once) за ее высокую скорость обработки видеопотока. Левая диаграмма отражает среднюю точность каждой версии нейросети, в зависимости от категории данных, которые она обрабатывает. Cправа показана средняя точность различных семейств версий YOLO, в зависимости от параметров, и задержки на обработку одного кадра. Основываясь на представленных графиках, остановимся на YOLOv8n, так как она является самой быстрой. Индекс n, означает, что данная модель принадлежит к семейству nano. На графиках также представлены small, medium, large, и extralarge.

**4 слайд:**

Датасет – это важная часть обучения нейросети. Критерии, предъявляемые к датасету следующие:

1. Наличие подробного описания – это сильно облегчает работу датасетом, и, часто, является показателем качественного продукта.

2. Ориентированность на системы автономного вождения

3. Вариативность содержащихся в нем данных – для лучшего обучения.

По выдвинутым критериям была выбран датасет KITTI, который является одним из самых популярных датасетов в сфере автономного вождения, имеет большое количество разнообразных данных, хорошее описание, а также является популярным среди датасетов, что означает наличие широкой аудитории.

Таким образом, было проведено обучение нейронной сети на датасете KITTI. Результаты обучения представлены на следующих слайдах.

**5 слайд:**

Слева представлены оригинальные кадры с ручной разметкой объектов, справа — результаты работы обученной модели.

1. Модель успешно распознаёт и правильно классифицирует большинство объектов, таких как автомобили и велосипедисты.
2. Уровень уверенности в обнаружении объектов высок, указывая на надёжность детекции.
3. Имеются расхождения: модель ошибочно обозначила лишние автомобили на одном из изображений, что может отражать проблемы с точностью или разнообразием данных.

В целом, модель показывает высокую эффективность и точность в детекции объектов, но требует доработки для минимизации ошибок.

**6 слайд:**

Также представлены видео, которые были загружены и обработаны обученной нейросетью. Как можно видеть, большинство объектов модель определяет верно, однако присутствуют и ошибочные срабатывания, а также их отсутствие. Особенно это заметно в темное время суток и при движении объектов на большой скорости.

В случае обработки видео, FPS варьируется в пределах от 29 до 52 (в зависимости от загруженности сцены), что удовлетворяет основному требованию обработки более 25 кадров в секунду.

**7 слайд:**

Теперь перейдем к оценке результатов:

Слева представлен график средней точности по отдельным классам и общая средняя точность по классам (синий цвет).

1. Классы, которые соответствуют наиболее популярным объектам на дороге (автомобиль, микроавтобус, грузовик), показывают высокие значения средней точности – модель хорошо обнаруживает и правильно классифицирует данные объекты.
2. Класс "Сидящий человек" имеет значительно более низкое значение точности и полноты (0.553), что может указывать на то, что этот класс труднее распознать из-за его нечастого появления в датасете.
3. Среднее значение точности для всех классов при пороге полноты в 50% (mAP@0.5) составляет 0.832, что говорит о достаточно высоком общем качестве распознавания по всем классам.

На правом графике представлена кривая Recall-Confidence (Полнота-уверенность) – на нем показано, как меняется полнота (recall) модели при изменении порога уверенности (confidence) для каждого класса.

1. Модель хорошо справляется с обнаружением автомобилей, микроавтобусов, грузовиков, поскольку кривые для этих классов высоки даже при большой уверенности.
2. Для классов с более низкими кривыми, например "Сидящий человек", модель демонстрирует более низкую полноту, что может означать сложность в обнаружении этих классов объектов.
3. Линия "Все классы" показывает совокупную эффективность модели по всем классам. Высокий показатель полноты при нулевом пороге уверенности (0.93 при 0.000) указывает на то, что в целом модель отлично справляется с обнаружением объектов, когда не применяется порог отсечения.

**8 слайд:**

Следующий график показывает кривую F1-Confidence для различных классов объектов. Оценка F1 — это гармоническое среднее между точностью и полнотой, и он стремится к 1, когда и точность, и полнота высоки. Значение F1 = 0.81 при уверенности 0.32 говорит о достаточно высоком качестве обнаружения по всем классам в целом.

Кривая Precision-Confidence – на нем показано, как меняется точность (precision) модели при изменении порога уверенности (confidence) для каждого класса.

Исходя из графика, можно сделать следующие выводы:

1. Классы, которые показывают высокую точность даже при высоком уровне уверенности, что является индикатором надежности модели.
2. Показатель 1.00 при Confidence 0.964 свидетельствует о высоком общем уровне точности для всех классов при высоких порогах уверенности.

В целом, график указывает на то, что результаты модели хорошие, особенно если учитывать высокий уровень Precision при высоких порогах Confidence.

**9 слайд:**

1. **Потери ограничивающей рамки (box loss)**:
   1. Графики показывают, как потери, связанные с ограничивающими рамками объектов, уменьшаются как на тренировочном, так и на валидационном наборе данных по мере обучения.
   2. Снижение этих потерь указывает на то, что модель становится лучше в точном определении местоположения объектов в изображении.
2. **Потери классификации (cls loss)**:
   1. Потери классификации также снижаются, что говорит о том, что модель улучшает свою способность правильно классифицировать объекты.
3. **Dual Focal Loss (dfl loss)**:
   1. Эти потери отражают потери в дисбалансе классов.Уменьшение этих потерь свидетельствует о том, что модель становится более уверенной в своих предсказаниях.
4. **Точность (precision) и полнота (recall)**:
   1. Увеличение этих метрик показывает, что модель становится более точной и полной в своих предсказаниях. Высокая точность означает, что большинство объектов, которые модель правильно.
5. **Средняя точность (mAP)**:
   1. mAP@0.5 и mAP@0.5-0.95 — это средние точности на различных порогах перекрытия. mAP@0.5 измеряет точность при IoU (пересечение по объединению) равном 0.5, а mAP@0.5-0.95 усредняет точность от IoU 0.5 до 0.95.
   2. Улучшение mAP свидетельствует о повышении общей точности модели по всему диапазону порогов IoU.

**10 слайд:**

**По результатам обучения модели можно сделать следующие выводы:**

1. mAP = 0.832 – обучение модели прошло хорошо, средняя точность довольно высока;
2. FPS = 29-52, что удовлетворяет требованию;
3. Ошибки и ложные срабатывания возникают при:

а. Высокой скорости транспортных средств;

б. Высокая загруженность сцены;

в. Недостаточная освещенность.

**Для устранения ошибок и ложных срабатываний необходимо:**

1. Дополнительное обучение с увеличением вариативности данных для повышения точности детекции и уменьшения количества FP результатов;
2. Внедрение механизмов аугментации данных для обучения, чтобы повысить устойчивость модели к изменяющимся условиям окружающей среды;
3. Изменение параметров обучения.

**11 слайд:**

**В ходе НИР были решены все поставленные задачи:**

* Была выбрана нейросеть YOLOv8n в качестве системы детекции объектов в реальном времени для автономного транспорта;
* Выбран датасет KITTI, который удовлетворяет всем предъявленным критериям;
* Проведено обучение модели для дальнейшего применения в сфере автономного транспорта;
* По полученным результатам обучения был проведен анализ и сделаны выводы, а также были представлены возможные способы улучшения показателей.