

Дообучение нейросетевых моделей YOLOv8 для детекции автомобилей и номерных знаков

Звайгзне Алексей Юрьевич

Приамурский государственный университет имени Шолом-Алейхема

Студент

Аннотация

В статье рассматривается процесс дообучения моделей семейства YOLOv8 для задач детекции автомобилей (четыре класса: легковой автомобиль, грузовик, автобус, мотоцикл) и государственных регистрационных знаков. Приводятся методики формирования размеченных наборов данных (15000 и 3000 изображений соответственно), параметры обучения и достигнутые метрики качества. Для детекции автомобилей значение mAP50 составило 0.903, для номерных знаков – 0.956, что превышает целевые показатели. Сделан вывод о высокой эффективности YOLOv8 для построения специализированных систем компьютерного зрения при условии репрезентативной обучающей выборки.

Ключевые слова: детекция объектов, YOLOv8, дообучение нейронных сетей, распознавание номерных знаков, наборы данных, компьютерное зрение.

Fine-tuning of YOLOv8 neural network models for car and license plate detection

Zvaigzne Alexey Yurievich

Sholom-Aleichem Priamursky State University

Student

Abstract

The article discusses the process of fine-tuning YOLOv8 family models for the tasks of car detection (four classes: car, truck, bus, motorcycle) and license plate detection. The methods of forming labeled datasets (15,000 and 3,000 images, respectively), training parameters and achieved quality metrics are presented. The mAP50 value for car detection was 0.903, for license plates – 0.956, which exceeds the target values. It is concluded that YOLOv8 is highly effective for building specialized computer vision systems, provided a representative training dataset.

Keywords: object detection, YOLOv8, neural network fine-tuning, license plate recognition, datasets, computer vision.

1 Введение

1.1 Актуальность

Современный этап развития информационных технологий характеризуется стремительным проникновением методов искусственного

интеллекта во все сферы человеческой деятельности. Компьютерное зрение, как одна из ключевых областей искусственного интеллекта, переживает период бурного роста, обусловленный появлением эффективных архитектур глубоких нейронных сетей и доступностью мощных вычислительных ресурсов. Среди множества прикладных задач компьютерного зрения особое место занимает детекция и классификация объектов на изображениях и видеопотоках, что находит применение в системах безопасности, автоматизированном транспорте, городской аналитике и промышленном мониторинге. Автомобильный транспорт является неотъемлемой частью современной инфраструктуры. Задачи автоматического распознавания транспортных средств и их государственных регистрационных знаков приобретают всё большую значимость в контексте развития интеллектуальных транспортных систем. Автоматическая идентификация автомобилей востребована при организации контроля доступа на закрытые территории, автоматизации парковочных систем, фиксации нарушений правил дорожного движения, а также в системах розыска угнанных транспортных средств. Традиционные методы, основанные на использовании индуктивных петель или ручном визуальном контроле, обладают ограниченной функциональностью, не обеспечивают требуемого уровня автоматизации и плохо масштабируются. Внедрение технологий компьютерного зрения позволяет не только детектировать факт присутствия автомобиля, но и осуществлять его классификацию по типу, а также распознавать государственный регистрационный номер с последующим сопоставлением с базами данных.

Прогресс в области глубокого обучения, в частности появление архитектур семейства YOLO (You Only Look Once), позволил достичь качественного скачка в скорости и точности детекции объектов. Современные версии, такие как YOLOv8, предоставляют унифицированный фреймворк для решения задач детекции, сегментации и классификации, поддерживают механизмы дообучения (fine-tuning) на пользовательских данных и экспорт в различные форматы для развёртывания в продуктивных средах. Однако предобученные на универсальных наборах данных (например, COCO) модели недостаточно точны при работе в специфических условиях: разнообразие ракурсов, погодных условий, типов автомобилей (включая грузовики, автобусы, мотоциклы) и форматов номерных знаков (в том числе российских с кириллицей). Кроме того, для Дальневосточного региона России характерны особые климатические и географические условия, которые предъявляют повышенные требования к надёжности систем компьютерного зрения. В связи с этим возникает объективная необходимость дообучения моделей на репрезентативных, специализированных наборах данных.

Особую актуальность данная задача приобретает в условиях Дальневосточного региона России, где интенсивность автомобильного трафика на магистралях и в приграничных зонах постоянно возрастает, а климатические и географические особенности (перепады температур,

туманы, снегопады, сложный рельеф) предъявляют повышенные требования к надёжности и устойчивости систем компьютерного зрения. Разработка отечественного программного решения на основе открытых нейросетевых архитектур, способного эффективно функционировать в различных погодных условиях и при различном освещении, является важной научно-практической задачей. Таким образом, разработка методики дообучения нейросетевых детекторов на собственных размеченных датасетах и оценка достигаемого качества является актуальной научно-практической задачей. Данная статья посвящена практическим аспектам дообучения моделей YOLOv8n для двух задач: детекции автомобилей с классификацией по четырём типам (легковой автомобиль, грузовик, автобус, мотоцикл) и детекции номерных знаков.

1.2 Обзор исследований

В своей работе С. Пчелинцев описал метод создания наборов данных для обучения нейросетевых моделей распознавания объектов [1]. В статье А.В. Кисилева проводится анализ эффективности применения двухэтапных нейросетевых моделей для раннего обнаружения лесных пожаров путем синтеза тренировочных наборов данных, предназначенных для обучения нейросетевых моделей [2]. Т.Д. Фатхулин привел обзор современных нейросетевых методов и технологий, используемых для генерации текста [3]. С.В. Сорокин произвел разбор примера построения модели поведенческого скрининга в банковской сфере на основе нейронной сети в программе IBM SPSS Statistics. [4]. Н. А. Еремина описывала методологию автоматизированной подготовки данных для машинного обучения нейросетевых моделей в интеллектуальных системах выявления и прогнозирования осложнений и аварийных ситуаций в процессе строительства нефтяных и газовых скважин [5]. Д.В. Антоненкова рассматривала проблему прогнозирования электропотребления горных предприятий, особенностью которого является высокий уровень нестационарности и стохастичности [6]. В методическом пособии для вузов А. Рябческого описываются текущие достижения в области генерации и использования синтетических данных [7].

1.3 Цель исследования

Целью работы является дообучение моделей YOLOv8n для детекции автомобилей (с классификацией по четырём типам: легковой автомобиль, грузовик, автобус, мотоцикл) и детекции номерных знаков на специализированных наборах данных, сформированных из открытых источников и собственных съёмок, с достижением целевых метрик качества: $mAP50 \geq 0,85$ для автомобилей и $mAP50 \geq 0,90$ для номерных знаков.

2 Материалы и методы

В данном исследовании для реализации серверной части использовались Python 3.10, OpenCV, YOLOv8, EasyOCR, Pillow (PIL), 2

модели оптического распознавания YOLOv8n (car_detect) и YOLOv8n (plate_detect), в качестве среды разработки использовался PyCharm, модели были обучены на ноутбуке с следующими аппаратными характеристиками: видеоадаптер Nvidia RTX 4060, AMD Ryzen 7 7735H, 32 ГБ RAM DDR5.

Для дообучения были сформированы два набора данных. Датасет автомобилей включает 15000 изображений, собранных из открытых источников (COCO, Open Images, UA-DETRAC) и собственных съёмок в г. Биробиджане. Разметка выполнена в LabelMe по классам: car, truck, bus, motorcycle. Распределение: 12000 – обучение, 1500 – валидация, 1500 – тест. Датасет номерных знаков содержит 3000 изображений (2400 – обучение, 300 – валидация, 300 – тест), класс «plate». В качестве базовой архитектуры выбрана YOLOv8n (nano) как оптимальный компромисс между точностью и скоростью.

3 Результаты и обсуждения

3.1 Интеграция моделей в веб-приложение

Успех дообучения нейросетевой модели в решающей степени зависит от качества и репрезентативности обучающей выборки. Для решения поставленных задач были сформированы два независимых набора данных: датасет для детекции и сегментации автомобилей, содержащий 15000 изображений, и датасет для детекции номерных знаков, включающий 3000 изображений. Формирование датасета автомобилей. Исходные изображения были собраны из нескольких источников, включая открытые наборы данных (Open Images Dataset, COCO, UA-DETRAC), видеозаписи с камер наружного наблюдения в городской среде г. Биробиджана и на автомобильных трассах Дальневосточного региона, а также фотографии, сделанные в различных погодных условиях и в разное время суток. Такой подход обеспечил разнообразие ракурсов, условий освещения, фоновых сцен и типов транспортных средств. Разметка изображений производилась в формате, совместимом с YOLO, с использованием специализированного инструмента LabelMe, который позволяет создавать как прямоугольные ограничивающие рамки, так и полигональные маски для сегментации. Для каждого изображения были определены объекты следующих классов:

- 0: car – легковой автомобиль;
- 1: truck – грузовой автомобиль;
- 2: bus – автобус;
- 3: motorcycle – мотоцикл.

Для задачи сегментации дополнительно создавались полигональные контуры, охватывающие видимую часть каждого экземпляра автомобиля с исключением фона и перекрывающих объектов (рис. 1).

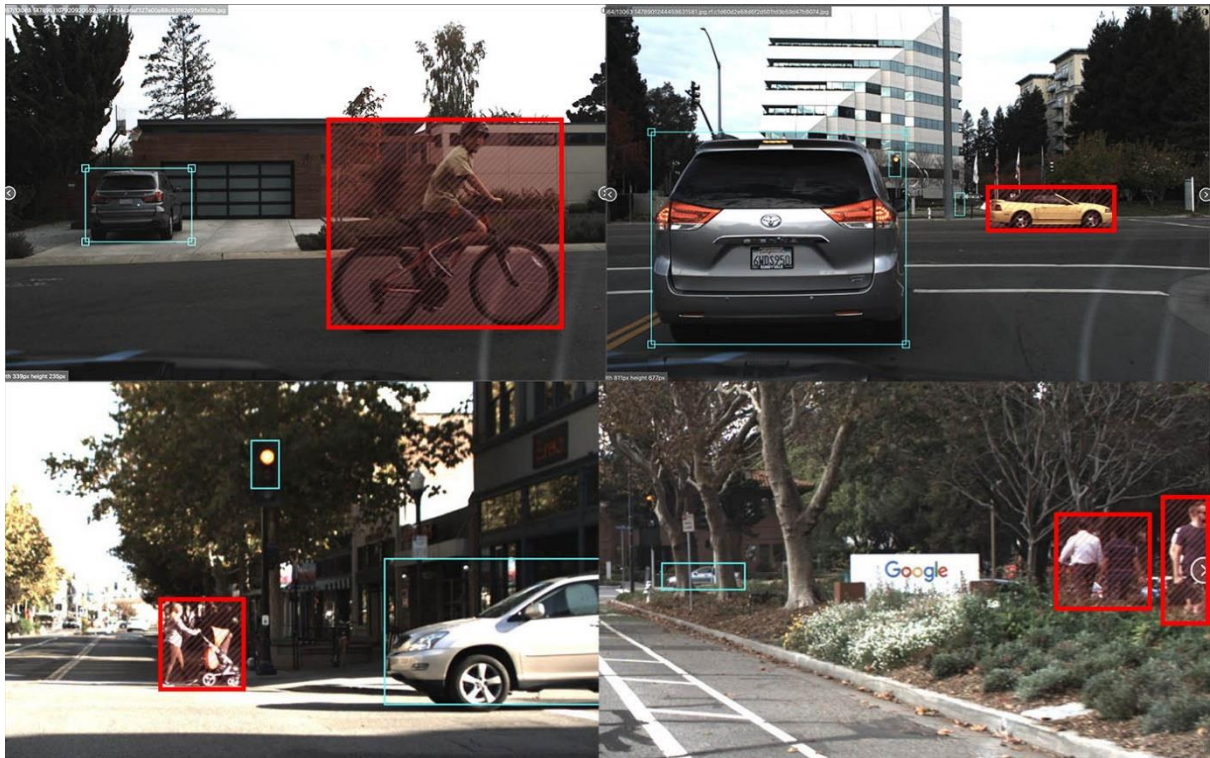


Рисунок 1. Пример разметки изображения с полигональными масками для сегментации автомобилей

Общее количество размеченных изображений составило 15000, из которых 12000 (80%) были отведены для обучающей выборки, 1500 (10%) – для валидационной, и 1500 (10%) – для тестовой. Распределение экземпляров по классам представлено в таблице 1.

Таблица 1. Распределение объектов по классам в датасете автомобилей

Класс	Название	Количество экземпляров	Доля, %
	Всего		

Для повышения устойчивости модели к вариациям входных данных применялись методы аугментации, встроенные в обучающий пайплайн Ultralytics YOLO: случайное масштабирование и обрезка (mosaic augmentation), случайные повороты, изменение яркости, контрастности и насыщенности, добавление шума. Формирование датасета номерных знаков.

Для обучения модели детекции номерных знаков был собран отдельный набор из 3000 изображений, содержащих автомобили с хорошо видимыми государственными регистрационными знаками. Особое внимание уделялось включению в выборку изображений с номерами, имеющими различные форматы (тип 1, тип 2, квадратные номера), различную степень загрязнения и физического износа, а также снятых под разными углами.

Разметка выполнялась путем выделения ограничивающих прямоугольников вокруг области номерного знака. Класс был определен единственный – «plate» (номерной знак). Распределение данных по выборкам: 2400 изображений для обучения, 300 для валидации, 300 для тестирования.

Пример размеченного изображения с номерным знаком представлен на рисунке 2.

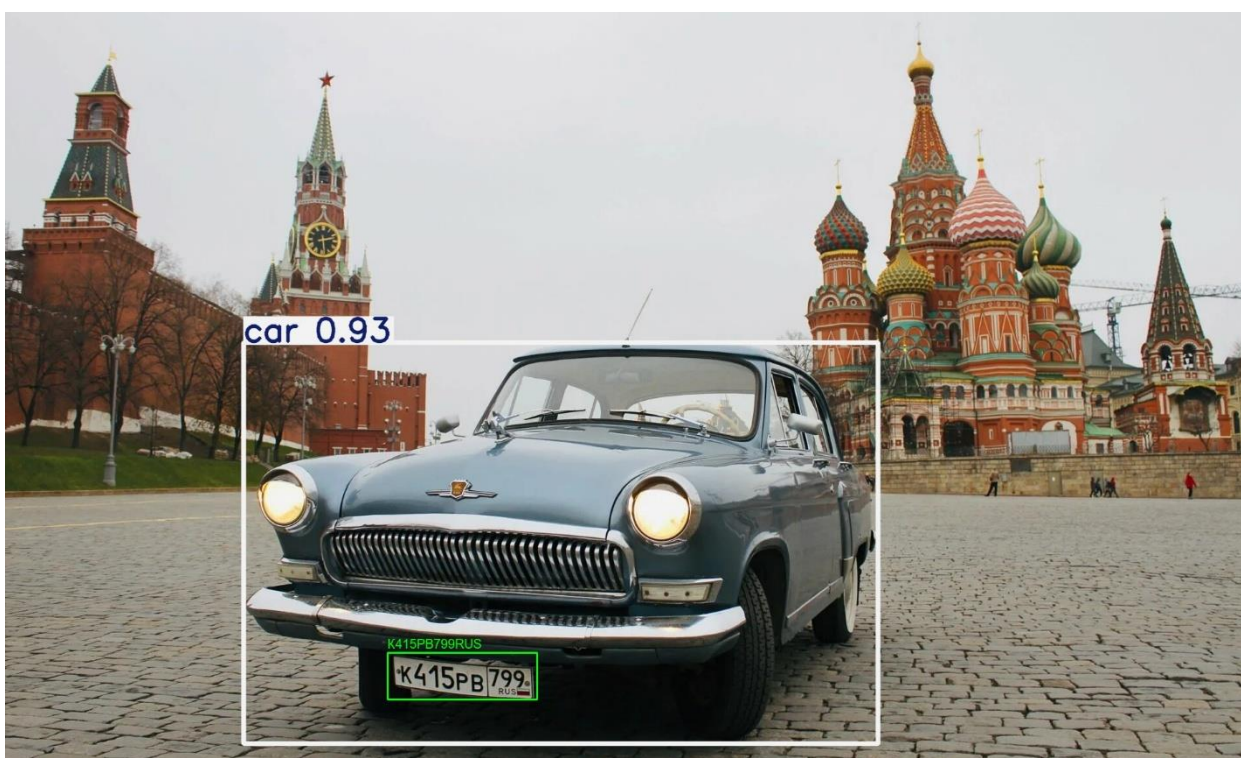


Рисунок 2. Пример разметки номерного знака с ограничивающей рамкой

Структура файлов и конфигурационные файлы были подготовлены в соответствии с требованиями формата YOLO: для каждого изображения создан текстовый файл с одноименным именем, содержащий строки вида `<class_id> <x_center> <y_center> <width> <height>`, где координаты нормализованы относительно размеров изображения. Для сегментации формат расширен включением координат вершин полигона.

3.2 Дообучение модели YOLOv8 для детекции автомобилей

В качестве базовой архитектуры для дообучения была выбрана модель YOLOv8n (nano) как оптимальный компромисс между точностью и скоростью инференса, что критически важно для обработки видеопотока в реальном

времени. Предобученные веса модели, обученной на датасете COCO, были загружены с использованием API библиотеки Ultralytics.

Процесс дообучения (fine-tuning) выполнялся на оборудовании, оснащённом графическим процессором NVIDIA GeForce RTX 4060 с 8 ГБ видеопамяти и 32 ГБ оперативной памяти. Обучение проводилось с использованием фреймворка PyTorch и библиотеки Ultralytics версии 8.0.196.

Ключевые гиперпараметры обучения:

- Размер пакета (batch size): 16.
- Количество эпох: 100.
- Размер входного изображения: 640×640 пикселей.
- Оптимизатор: AdamW с начальной скоростью обучения 0.001.
- Планировщик скорости обучения: косинусное затухание.
- Аугментация данных: встроенные методы YOLO (mosaic, mixup, copy-paste).
- Функция потерь: стандартная комбинированная потеря YOLOv8 (DFL +

Обучение модели для сегментации (YOLOv8n-seg) производилось с аналогичными гиперпараметрами на том же наборе данных, но с использованием масок, заданных в виде полигонов. Разница заключалась в активации дополнительной ветви, ответственной за предсказание масок.

Процесс обучения контролировался по динамике функции потерь и метрик качества на валидационной выборке. На рисунке 3 представлены графики изменения функции потерь (loss) на обучающей и валидационной выборках в ходе обучения модели детекции.

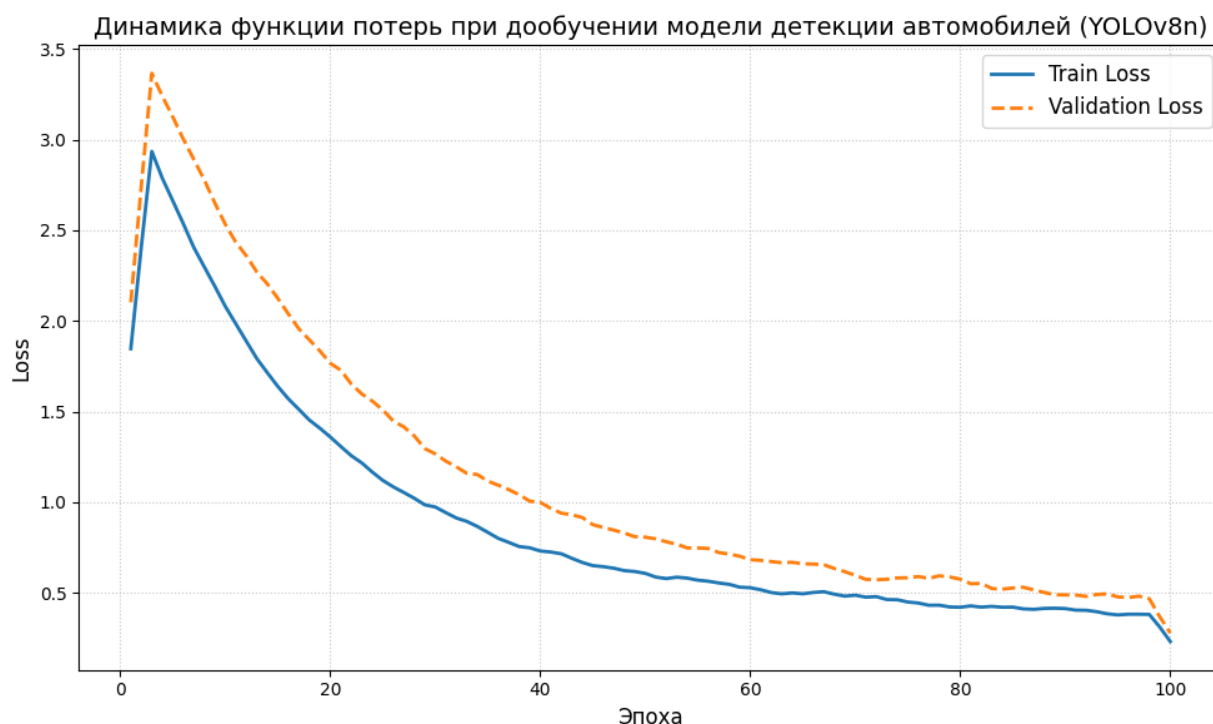


Рисунок 3. Графики изменения функции потерь при дообучении модели детекции автомобилей

Анализ графиков показывает стабильное снижение потерь на протяжении первых 60 эпох с последующим выходом на плато, что свидетельствует о корректном подборе гиперпараметров и отсутствии выраженного переобучения. Небольшой разрыв между кривыми обучения и валидации указывает на хорошую обобщающую способность модели.

Для модели сегментации динамика обучения была схожей, однако сходимость наступала несколько медленнее из-за более сложной целевой функции, включающей компоненту для предсказания масок.

По завершении обучения были вычислены основные метрики качества на тестовой выборке. Результаты для модели детекции представлены в таблице

Таблица 2. Метрики качества модели детекции автомобилей

Класс				
Все классы				

Модель сегментации показала следующие результаты (таблица 3).

Таблица 3. Метрики качества модели сегментации автомобилей

Класс				
Все классы				

Достигнутые значения mAP50 превышают целевые показатели (0.85 для детекции, 0.80 для сегментации), что подтверждает высокое качество дообученных моделей. Снижение метрик для класса «motorcycle» объясняется

меньшим количеством обучающих примеров и высокой вариабельностью визуального представления мотоциклов на изображениях.

На рисунке 4 приведен пример работы модели детекции на тестовом изображении, демонстрирующий корректное обнаружение автомобилей



Рисунок 4. Пример детекции автомобилей на тестовом изображении

3.3 Дообучение модели для распознавания номерных знаков

Для детекции государственных регистрационных знаков использовалась отдельная модель YOLOv8, дообученная на специализированном датасете из 3000 размеченных изображений. Данная модель не связана с моделью детекции автомобилей из раздела 2.2 и разрабатывалась независимо.

Параметры обучения были подобраны с учётом меньшего объёма выборки и относительной простоты геометрии целевых объектов:

- Количество эпох: 50.
- Размер пакета (batch size): 16.
- Размер входного изображения: 640×640 пикселей.
- Оптимизатор: AdamW с начальной скоростью обучения 0.001.
- Аугментация данных: mosaic, случайные повороты, изменение яркости.

Обучение модели заняло около 2,5 часов на том же аппаратном обеспечении (NVIDIA GeForce RTX 4060, 32 ГБ ОЗУ). Динамика функции потерь на обучающей и валидационной выборках представлена на рисунке 5.

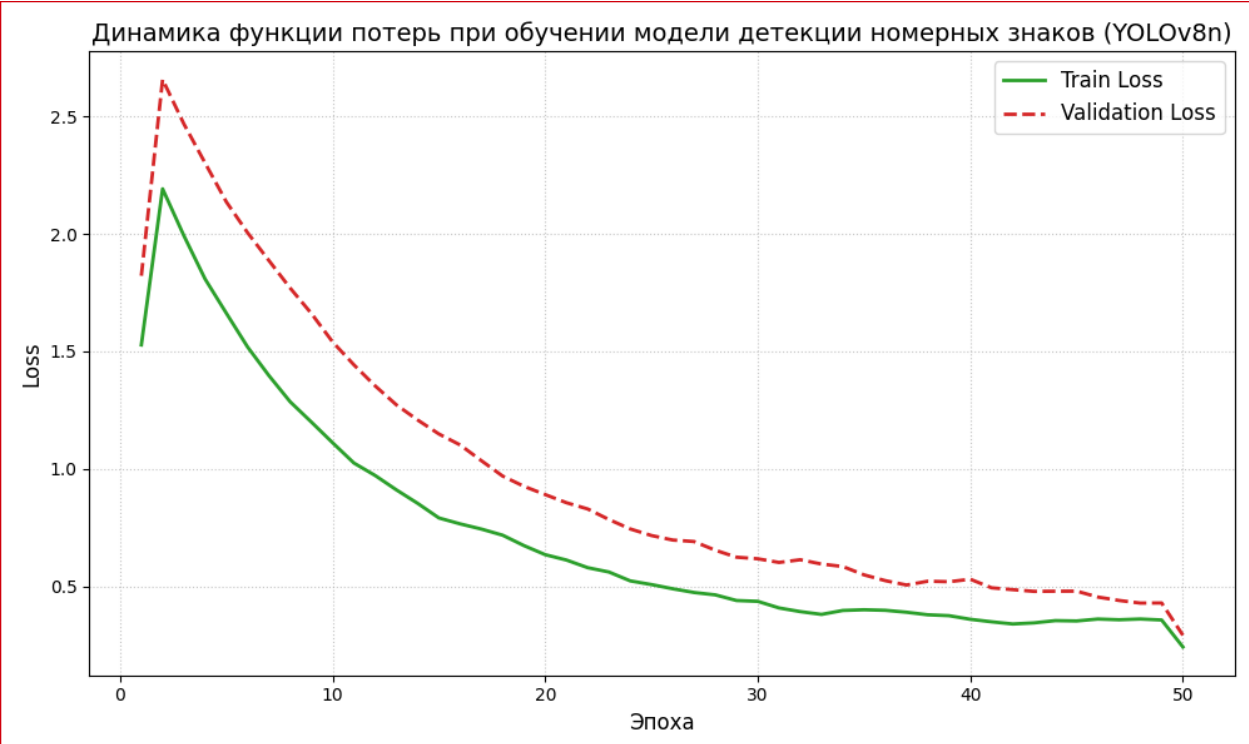


Рисунок 5. График изменения функции потерь при обучении отдельной модели детекции номерных знаков (YOLOv8n)

По сравнению с моделью детекции автомобилей (рисунок 3), сходимость наступает быстрее — потери стабилизируются уже к 20–25 эпохам, что объясняется более простой задачей (один класс объектов с относительно стандартной геометрией).

Метрики качества, достигнутые на тестовой выборке, приведены в таблице 4.

Таблица 4. Метрики качества модели детекции номерных знаков

Класс				

Значение $mAP50 = 0.956$ существенно превышает целевой показатель 0.90, что свидетельствует о высокой надежности модели в задаче обнаружения номерных знаков. Высокая точность ($precision = 0.947$) указывает на низкий уровень ложных срабатываний, что критически важно для практического применения.

Пример работы модели на тестовых изображениях представлены на рисунке 6. Модель успешно детектирует номерные знаки при различных ракурсах съемки, включая наклонные и частично перекрытые номера.

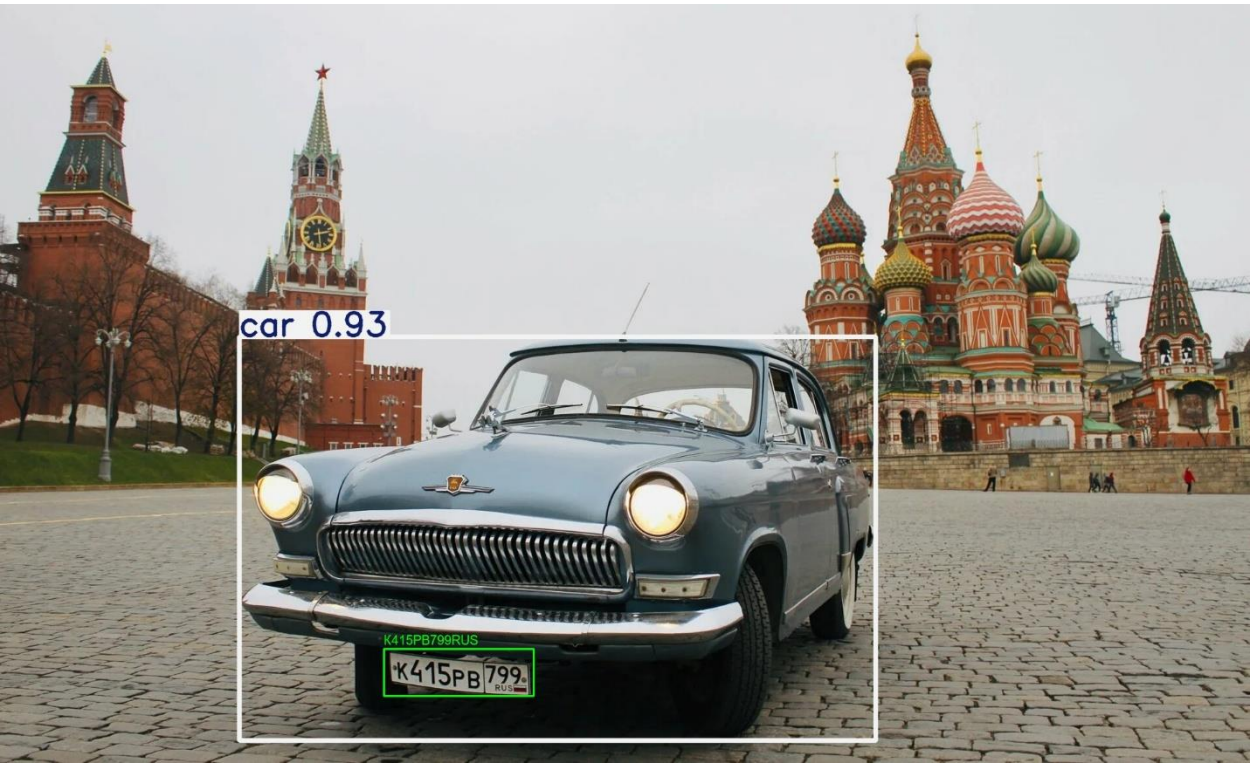


Рисунок 6. Примеры детекции номерных знаков

Примечание относительно распознавания символов. Полноценное распознавание текста номерного знака выходит за рамки настоящей статьи, поскольку требует интеграции дополнительной OCR-модели и значительно усложняет архитектуру системы. Однако созданная модель детекции номерного знака предоставляет вырезанную область, которая может быть передана на вход любой OCR-системе, например Tesseract или специализированной нейросетевой модели распознавания символов.

3.4 Тестирование и оценка качества работы моделей

Тестирование разработанной системы проводилось в два этапа: количественная оценка точности моделей на отложенных тестовых выборках и качественное тестирование производительности и удобства использования в реальных условиях.

Оценка точности детекции. Для каждой из дообученных моделей были рассчитаны стандартные метрики на независимых тестовых наборах данных. Результаты сведены в таблицу 5.

Таблица 5. Сводные метрики качества дообученных моделей

Модель	Задача				
	Детекция автомобилей				

Модель	Задача				
	Детекция номерных знаков				

Оценка качества распознавания текста. Точность OCR оценивалась на наборе из 200 изображений номерных знаков, вырезанных моделью детекции. Поскольку использовалась готовая библиотека EasyOCR с предобученными моделями для русского и английского языков, дополнительного обучения не проводилось. Достигнутая точность полного совпадения распознанной строки с эталонной (после приведения к верхнему регистру и удаления пробелов) составила 84,5%. Основные ошибки были связаны с сильным загрязнением знака, нестандартными шрифтами или низким разрешением.

Производительность. Измерение скорости обработки проводилось на тестовом оборудовании: ноутбук с процессором AMD Ryzen 7 7735H, 32 ГБ ОЗУ, видеокарта NVIDIA GeForce RTX 4060 8 ГБ. Результаты представлены в таблице 6.

Таблица 6. Производительность приложения в различных режимах

Режим работы	Среднее время обработки кадра, мс	
Только детекция автомобилей		
Детекция автомобилей + детекция номеров		
Детекция автомобилей + детекция номеров + OCR		

Как видно из таблицы, даже в наиболее ресурсоёмком режиме (с распознаванием текста) приложение обеспечивает частоту 23 кадра в секунду, что достаточно для комфортной работы в реальном времени. Основное падение производительности связано с работой EasyOCR, выполняемой на центральном процессоре. Для повышения скорости в будущем может быть рассмотрено использование GPU-версии OCR или облегчённой нейросетевой модели распознавания символов.

Выводы

В ходе исследования сформированы размеченные наборы данных: 15 000 изображений автомобилей четырёх классов (car, truck, bus, motorcycle) и 3 000 изображений номерных знаков. Оба датасета включают разнообразные

ракурсы, погодные условия и типы транспортных средств, что обеспечило высокую обобщающую способность обученных моделей. Выполнено дообучение моделей YOLOv8n для детекции автомобилей и номерных знаков. Для детекции автомобилей достигнуто значение $mAP50 = 0,903$, что превышает целевой показатель (0,85). Для детекции номерных знаков значение $mAP50$ составило 0,956, что также выше целевого порога (0,90). Полученные метрики подтверждают эффективность выбранной архитектуры и методики дообучения. Модель детекции номерных знаков демонстрирует высокую точность (0,947) и полноту (0,931), что позволяет использовать её в системах автоматического распознавания с минимальным уровнем ложных срабатываний. Высокое значение $mAP50-95$ (0,742) указывает на хорошую локализацию границ номерной пластины. Экспериментально установлено, что основными источниками ошибок для модели детекции автомобилей являются перекрытие объектов, нестандартные ракурсы и недостаток примеров для класса «motorcycle». Для модели номерных знаков основные ошибки связаны с сильным загрязнением, засветкой и острыми углами съёмки. Разработанные модели могут быть интегрированы в интеллектуальные транспортные системы, системы контроля доступа и мониторинга транспортных потоков. Предложенная методика дообучения применима для других задач детекции объектов при наличии соответствующего размеченного датасета.

Библиографический список

1. Пчелинцев С., Юляшков М. А., Ковалева О. А. Метод создания синтетических наборов данных для обучения нейросетевых моделей распознаванию объектов // Информационно-управляющие системы. 2022. №. 3 (118). – С. 9-19. URL: <https://cyberleninka.ru/article/n/metod-sozdaniya-sinteticheskikh-naborov-dannyh-dlya-obucheniya-neyrosetevykh-modeley-raspoznavaniyu-obektov> (дата обращения: 05.05.2026).
2. Киселев А. В., Брусенцев Н. С., Кулешова Е. А. Анализ эффективности применения двухэтапных нейросетевых моделей для раннего обнаружения лесных пожаров // Известия Юго-Западного государственного университета. Серия: Управление, вычислительная техника, информатика. Медицинское приборостроение. 2024. Т. 14. №. 1. С. 8-23. URL: <https://uprinmatus.elpub.ru/jour/article/view/156> (дата обращения: 05.05.2026).
3. Фатхулин, Т. Д., Климов, Н. Ю., Гежин, С. А. Анализ нейросетевых технологий, позволяющих генерировать текст // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2023. № 1. С. 123-127. URL: <https://www.elibrary.ru/item.asp?edn=mxctvo> (дата обращения: 05.05.2026).
4. Сорокин С. В., Сорокин А. С. Использование нейросетевых моделей в поведенческом скоринге // Прикладная информатика. 2015. №. 2 (56). С. 92-109. URL: <https://cyberleninka.ru/article/n/ispolzovanie-neyrosetevykh-modeley-v-povedencheskom-skoringe> (дата обращения: 05.05.2026).

5. Еремин Н. А., Черников А. Д. Методология автоматизированной подготовки данных для машинного обучения нейросетевых моделей в интеллектуальных системах выявления и прогнозирования осложнений и аварийных ситуаций в процессе строительства нефтяных и газовых скважин //Экспозиция Нефть Газ. 2024. №. 5. С. 24-30. URL: <https://cyberleninka.ru/article/n/metodologiya-avtomatizirovannoy-podgotovki-dannyh-dlya-mashinnogo-obucheniya-neyrosetevyh-modeley-v-intellektualnyh-sistemah> (дата обращения: 05.05.2026).
6. Антоненков Д. В., Матренин П. В. Исследование ансамблевых и нейросетевых методов машинного обучения в задаче краткосрочного прогнозирования электропотребления горных предприятий //Электротехнические системы и комплексы. 2021. №. 3 (52). С. 57-65. URL: <https://cyberleninka.ru/article/n/issledovanie-ansamblevyh-i-neyrosetevyh-metodov-mashinnogo-obucheniya-v-zadache-kratkosrochnogo-prognozirovaniya> (дата обращения: 05.05.2026).
7. Рабчевский А. Синтетические данные и развитие нейросетевых технологий. Учебное пособие для вузов. ЛитРес, 2023. URL: <https://urait.ru/book/sinteticheskie-dannye-i-razvitie-neyrosetevyh-tehnologiy-589394> (дата обращения: 05.05.2026).