

Научная статья
УДК 004.932:625.8

РАЗРАБОТКА МЕТОДИКИ ФОРМИРОВАНИЯ ДАТАСЕТА ДЛЯ ДЕТЕКЦИИ ДЕФЕКТОВ ДОРОЖНОГО ПОКРЫТИЯ С ИСПОЛЬЗОВАНИЕМ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Никита Юрьевич Мокрушин¹, Сергей Александрович Чудинов²

^{1, 2} Уральский государственный лесотехнический университет,
Екатеринбург, Россия

¹ mokrushinnikita621@mail.ru

² chudinovsa@m.usfeu.ru

Аннотация. Описывается методика создания аннотированного датасета для автоматизации мониторинга дорожного покрытия. Этапы включают сбор видеоданных, выделение кадров и их семантическую разметку в CVAT. Результатом является датасет в формате YOLO, готовый для обучения нейросетей для обнаружения дефектов.

Ключевые слова: компьютерное зрение, датасет, дефекты дорожного покрытия

Для цитирования: Мокрушин Н. Ю., Чудинов С. А. Разработка методики формирования датасета для детекции дефектов дорожного покрытия с использованием компьютерного зрения // Научное творчество молодежи – лесному комплексу России = Scientific creativity of youth to the forest complex of Russia : материалы XXII Всероссийской (национальной) научно-технической конференции студентов и аспирантов. Екатеринбург : УГЛТУ, 2026. С. 927–931.

Original article

DEVELOPMENT OF A DATASET CREATION METHODOLOGY FOR ROAD SURFACE DEFECT DETECTION USING COMPUTER VISION

Nikita Yu. Mokrushin¹, Sergey A. Chudinov²

^{1, 2} Ural State Forest Engineering University, Ekaterinburg, Russia

¹ mokrushinnikita621@mail.ru

² chudinovsa@m.usfeu.ru

Abstract. This article describes a methodology for creating an annotated dataset for automated road surface monitoring. The steps include collecting video data, extracting frames, and semantically annotating them using CVAT. The result is a dataset in YOLO format, ready for training neural networks for defect detection.

© Мокрушин Н. Ю., Чудинов С. А., 2026

Keywords: computer vision, dataset, road surface defects

For citation: Mokrushin N. Yu., Chudinov S. A. (2026) Razrabotka metodiki formirovaniya dataseta dlya detekcii defektov dorozhnogo pokry`tiya s ispol`zovaniem komp`yuternogo zreniya [Development of a dataset creation methodology for road surface defect detection using computer vision]. Nauchnoe tvorchestvo molodezhi – lesnomu kompleksu Rossii [Scientific creativity of youth to the forest complex of Russia] : materials of the XXII All-Russian (national) Scientific and Technical Conference of undergraduate and postgraduate students. Ekaterinburg : USFEU, 2026. P. 927–931. (In Russ).

Состояние дорожного покрытия представляет собой критически важный фактор, непосредственно влияющий на безопасность дорожного движения, комфорт вождения и эксплуатационные расходы транспортных средств. Согласно требованиям ГОСТ Р 50597–2017 [1], автомобильные дороги должны соответствовать установленным параметрам эксплуатационного состояния. Существующие традиционные методы инспекции, основанные на визуальном осмотре, демонстрируют очевидные ограничения – высокую трудоемкость, субъективность оценок и недостаточную эффективность при масштабном применении.

Для успешного обучения моделей компьютерного зрения необходим качественный датасет. В данном исследовании разрабатывается методика его создания в соответствии с классификацией дефектов ГОСТ Р 52748–2007 [2].

Исследование началось с определения целевых классов дефектов для детекции. В фокус работы вошли наиболее распространенные и критичные типы повреждений асфальтобетонного покрытия: продольные и поперечные трещины, локальные выбоины и ямы различной конфигурации, продольная колеиность в полосе наката, а также сетка трещин.

Техническая реализация сбора данных обеспечивалась использованием автомобильного видеорежистратора с разрешением 2К (2304×1296 пикселей) и частотой съемки 30 кадров в секунду. Широкоугольная оптика с углом обзора 140° позволила охватывать значительные участки проезжей части. Особое внимание было уделено обеспечению репрезентативности собираемого материала – съемка проводилась в различных погодных условиях и в разное время суток, что соответствует рекомендациям по созданию датасетов для компьютерного зрения [3].

На этапе предобработки видеоматериалы преобразовывались в последовательности отдельных изображений. Данная операция была выполнена в программе Adobe After Effects с использованием функции «Секвенция кадров». Данный тип рендеринга позволяет сохранить каждый кадр видео в виде отдельного изображения (в формате JPEG) с последовательной нумерацией имен файлов. Процесс включал загрузку видео, создание композиции, добавление в очередь рендеринга и настройку модуля вывода на максимальное качество.

После экстракции кадров выполнялся их тщательный ручной анализ и фильтрация. Для формирования финального датасета отбирались изображения двух категорий: с выраженными дефектами дорожного покрытия и без повреждений.

Для аннотирования данных применялась веб-платформа с открытым исходным кодом CVAT [4]. Использовался тип аннотации «Сегментация экземпляров» с применением инструмента «Полигон», что обеспечивает точное выделение контуров дефектов произвольной формы (рис. 1).

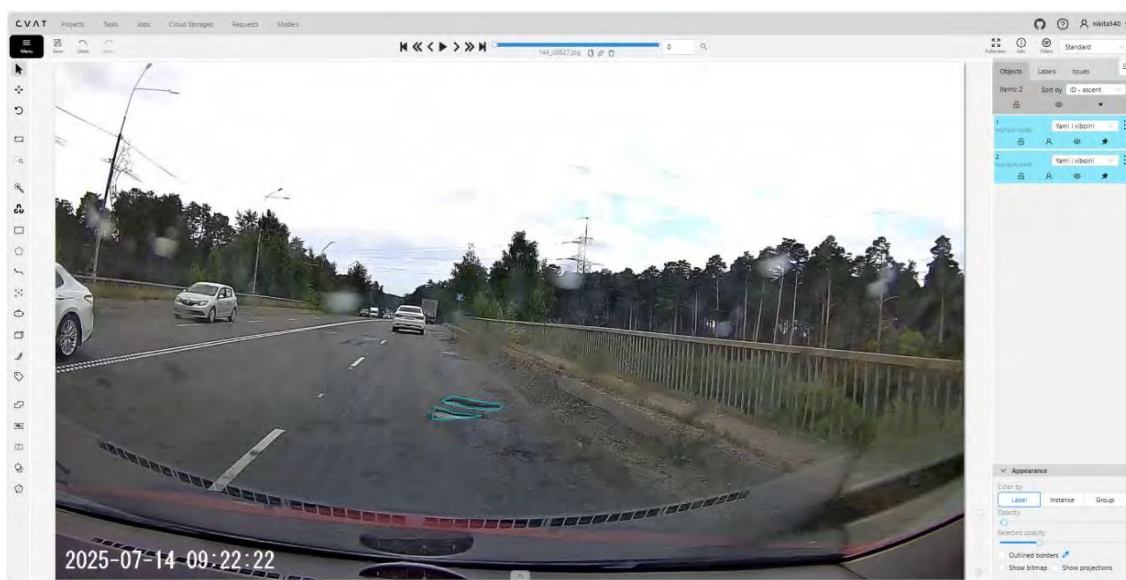


Рис. 1. Размеченное изображение в CVAT

Процесс разметки включал создание проекта с метками, формирование задач с загрузкой изображений и экспорт данных в формате YOLO 1.1.

Перед началом обучения была проведена комплексная подготовительная диагностика вычислительной среды. Для обеспечения стабильной работы на графическом процессоре RTX 3050 с объемом памяти 4 ГБ разработан специализированный скрипт оптимизации, включающий проверку доступности CUDA (Compute Unified Device Architecture – параллельная вычислительная архитектура для ускорения работы на GPU), мониторинг использования видеопамати и настройку гиперпараметров для задач сегментации. Валидация датасета подтвердила корректность структуры разметки в формате YOLO и наличие файлов масок для всех изображений. На этапе предварительного тестирования использовалась модель YOLO11n-seg с оптимизированными параметрами: размер пакета (batch size) = 4, разрешение изображений 512×512 пикселей, mixed precision (AMP) для эффективного использования памяти. Пятиэпохальное обучение в тестовом режиме показало стабильную работу pipeline: потребление памяти не превышало 0,59 ГБ из 4,0 ГБ доступных, что подтвердило корректность выбранной методологии и возможность масштабирования процесса на большее количество эпох.

Для обучения моделей сегментации была разработана комплексная методика на основе фреймворка YOLO. Обучение проводилось на графическом процессоре RTX 3050 с использованием предобученных весов yolo11n-seg.pt. В ходе оптимизации гиперпараметров была применена стратегия увеличения количества эпох до 100 с использованием косинусного планирования скорости обучения (cosine LR scheduler) и mixed precision training (AMP). Для улучшения обобщающей способности модели использовался расширенный набор аугментаций данных, включающий мозаику, изменение цветовых характеристик HSV (Цветовая модель: Hue – тон, Saturation – насыщенность, Value – значение), случайные повороты и отражения. Валидация модели показала значительное улучшение метрик по сравнению с первоначальным обучением: mAP50 увеличился с 0,156 до 0,347, а полнота (recall) выросла с 0,123 до 0,281. Наилучшие результаты продемонстрировал класс «Колейность» (mAP50 = 0,54), тогда как детекция сетки трещин и продольных трещин требует дополнительного расширения датасета (таблица).

Результаты обучения модели сегментации

Класс дефекта	Точность (Precision)	Полнота (Recall)	Map@50	Map@50:95	Количество экземпляров (Val)
Все классы (среднее)	0,421	0,281	0,347	0,152	584
Колейность	0,581	0,492	0,540	0,241	145
Выбоина	0,455	0,321	0,385	0,178	187
Поперечная трещина	0,392	0,234	0,301	0,125	126
Продольная трещина	0,298	0,118	0,187	0,064	78
Сетка трещин	0,359	0,141	0,222	0,102	48

Примечание. Точность (Precision): показывает, насколько модель «была права», когда находила объект. Полнота (Recall): показывает, какую долю реальных дефектов модель смогла обнаружить. mAP (mean Average Precision): сводная метка, объединяющая точность и полноту. mAP@50 – это средняя точность по всем классам при пороге пересечения с реальным объектом (IoU) ≥ 50 %. mAP@50:95 – это усредненное значение mAP при различных порогах IoU от 50 до 95 % с шагом 5 %, что является более строгой метрикой.

На изображение наложены маски обнаруженных дефектов с соответствующими метками классов и значением вероятности (confidence score).

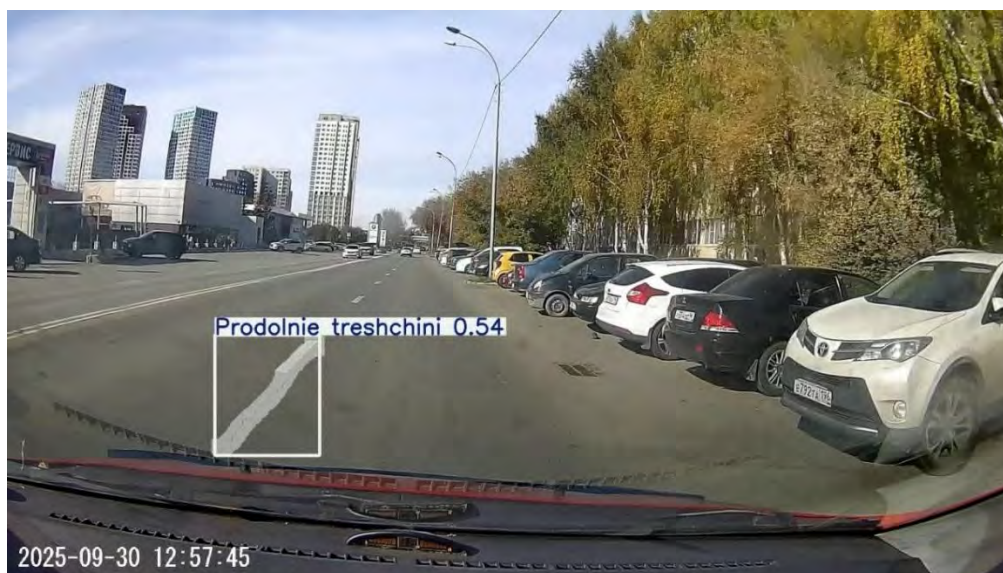


Рис. 2. Визуализация результата работы обученной модели сегментации

Разработанная методология демонстрирует ключевые преимущества: системность, репрезентативность и высокая точность аннотаций. Созданный датасет представляет собой ценный ресурс для разработки систем автоматического мониторинга дорожной инфраструктуры и может быть использован для обучения современных архитектур нейронных сетей [5].

Список источников

1. ГОСТ Р 50597–2017. Автомобильные дороги и улицы. Требования к эксплуатационному состоянию, допустимому по условиям обеспечения безопасности дорожного движения. Введ. 2018–06–01. М. : Стандартинформ, 2017. 25 с.
2. ГОСТ Р 52748–2007. Дороги автомобильные общего пользования. Нормативы допустимых дефектов и повреждений. Введ. 2008–01–01. М. : Стандартинформ, 2007. 15 с.
3. Goodfellow I., Bengio Y., Courville A. Deep Learning. Cambridge : The MIT Press, 2016. 800 p.
4. Computer Vision Annotation Tool (CVAT) [Электронный ресурс] URL: <https://www.cvat.ai/documentation> (дата обращения: 15.10.2025).
5. Mask R-CNN / K. He, G. Gkioxari, P. Dollár, R. Girshick // Proceedings of the IEEE International Conference on Computer Vision (ICCV). Venice, Italy, 2017. P. 2961–2969.