

АННОТАЦИЯ

диссертационной работы Олжаева Олжас Муратұлы «Разработка системы обнаружения повреждений дорог с использованием методов глубокого обучения на основе видеоданных», представленной на соискание степени доктора философии (PhD) по образовательной программе: 8D06105 – «Наука о данных»

Актуальность

Обеспечение эксплуатационной надежности дорожной инфраструктуры следует рассматривать как критический фактор, определяющий не только темпы экономического развития, но и уровень общественной безопасности. Процесс ухудшение дорожного полотна, как правило, инициируется формированием сетки микротрещин, которые при отсутствии своевременного вмешательства неизбежно перерастают в глубокие выбоины. Подобные дефекты создают прямую угрозу для участников движения. Согласно глобальным отчетам Всемирной организации здравоохранения (ВОЗ), ежегодная смертность в результате инцидентов на дорогах достигает 1,19 миллиона человек. При этом диапазон несмертельного травматизма, нередко влекущего за собой стойкую инвалидность, охватывает до 50 миллионов пострадавших. Важно учитывать, что состояние дорожного покрытия детерминирует не менее 10% всех зарегистрированных аварий. В этой связи переход к мониторингу дорожных сетей трансформируется из узкотехнической задачи в приоритетную стратегию в области здравоохранения.

Прямые затраты на ликвидацию последствий ДТП напрямую влияют на финансовый аспект, обусловленный эксплуатацией изношенных дорожных сетей. Глобальная исследовательская практика показывает, что совокупный ущерб от дорожно-транспортного травматизма и деградации покрытия может достигать 3% от валового внутреннего продукта (ВВП) государства. Прямые расходы на медицину, долгосрочное снижение производительности труда, а также форсированный износ инженерных сооружений и транспортных средств формируют этот показатель. И более того, ежегодно расходуются миллиардные издержки на внеплановое сервисное обслуживание из-за эксплуатации автопарка на дефектном полотне, что ведет к преждевременному истощению технического ресурса машин. Учитывая эти условия, когда дорожные ведомства стремятся к оптимизации операционных расходов, переход к стратегиям проактивного мониторинга становится безальтернативным инструментом поддержания активов.

На сегодняшний день популярным методом аудита дорожного хозяйства остается визуальное обследование, делегированное персоналу эксплуатационных служб. Однако потенциальная глубина такой экспертной оценки нивелируется ограниченностью технических ресурсов. В действительности ручной контроль представляет собой крайне инертный процесс, сопряженный с высокими операционными издержками и, что критично для точности данных, существенным риском субъективных искажений. Физические ограничения человеческого фактора делают невозможным обеспечение адекватного пространственного охвата, необходимого для непрерывного мониторинга дорожных сетей в режиме реального времени. На фоне стремительной урбанизации и усложнения топологии транспортных узлов традиционная визуальная инспекция

окончательно утратила свою эффективность, и сдерживает развитие цифровизации технического обслуживания дорог.

Стремительная эволюция компьютерного зрения и алгоритмов глубокого обучения открыла принципиально новые возможности для создания автономных систем мониторинга дорожной инфраструктуры. Прорывные результаты, достигнутые в задачах классификации, объектной детекции и семантической сегментации, сосредоточили исследовательский интерес к автоматизации поиска дорожных дефектов. В рамках данного технологического вектора доминирующим подходом стало использование глубоких сверточных нейронных сетей (CNN) и современных детекторов. Такие архитектуры позволяют экстрагировать высокоуровневые визуальные признаки из снимков высокого разрешения, обеспечивая прецизионную идентификацию повреждений в условиях сложного визуального контекста.

Сфера автоматизированной детекции дорожных дефектов с применением компьютерного зрения на текущие дни считается одной из актуальных задач, и видим экспоненциальный рост интереса в данной сфере. Сделав анализ профильной литературы, наглядно можем видеть этот всплеск научной активности: если в 2015 году числилось лишь около 50 релевантных публикаций, то к 2025 году их количество превысило отметку в 900 работ. Подобная динамика обусловлена тремя факторами: повсеместным внедрением сенсоров высокого разрешения, формированием масштабных репрезентативных датасетов и доступностью высокопроизводительных вычислительных мощностей, необходимых для обучения многослойных нейронных сетей.

Если проанализировать текущие научные работы, можно заметить, что основной упор в них делается на обработку одиночных снимков с регистраторов или смартфонов. Но практика показывает: для реального мониторинга этого мало. Нам нужно видеть состояние полотна в динамике, используя видеопоток. Именно работа с видео добавляет в систему «временную координату», которая помогает точнее распознавать дефекты и отсеивать случайные помехи. Когда камера на автомобиле пишет видео без перерывов прямо во время движения, это позволяет охватить огромные участки дорожной сети. В итоге получается инструмент для диагностики повреждений почти в реальном времени, что критически важно для оперативного обслуживания магистралей.

Несмотря на достигнутые успехи, сегмент автоматизированной детекции дорожных дефектов средствами дистанционного зондирования по-прежнему сталкивается с рядом фундаментальных барьеров. Эффективность существующих систем критически зависит от вариативности внешней среды: динамического диапазона освещенности, наличия глубоких окклюзий (теней), сезонных метаморфоз ландшафта и нестабильности метеоусловий. Так же дополнительно надо учитывать сложную специфику дистантной съемки под острыми углами, что неизбежно ведет к перспективным искажениям объектов. Более того, морфология самих дефектов – их текстура, нелинейные формы и широкий диапазон геометрических размеров – крайне затрудняет процесс их прецизионной дифференциации стандартными методами. Данные вызовы диктуют необходимость перехода к более сложным, иерархическим архитектурам глубокого обучения. В частности, требуется внедрение специализированных модулей экстракции признаков, обладающих высокой

робастностью к визуальному шуму и способных выделять релевантные паттерны повреждений в условиях низкой контрастности сцены.

Целью диссертационного исследования является разработка интеллектуальной системы автоматического обнаружения и анализа повреждений дорожного покрытия в режиме реального времени на основе методов компьютерного зрения и глубокого обучения, обеспечивающей повышение точности выявления дефектов дорожной инфраструктуры и возможность их последующего картографирования.

Были поставлены следующие **задачи** исследования:

- провести сбор видеоданных с использованием камер высокого разрешения и других средств записи;
- выполнить ручную разметку данных, выделяя разнообразные типы повреждений дорожного покрытия, такие как трещины, выбоины и деформации;
- создать высококачественный размеченный датасет для обучения и тестирования моделей глубокого обучения;
- провести анализ визуальных признаков и характеристик изображений дорожного покрытия, влияющих на качество данных и устойчивость моделей глубокого обучения;
- разработать и обучить модель глубокого обучения для классификации и сегментации повреждений дорог;
- провести тестирование разработанных моделей на собранном наборе данных, с целью достижения высокой точности и устойчивости к различным условиям съемки;
- интегрировать обученную модель в приложение для автоматического мониторинга состояния дорог, обеспечив возможность построения интерактивных карт повреждений;

Объектом исследования являются автомобильные дороги и визуальные характеристики повреждений дорожного покрытия, получаемые на основе анализа видеоданных.

Предметом исследования являются методы и алгоритмы компьютерного зрения, глубокого обучения и многозадачных нейронных сетей для обнаружения, классификации и сегментации повреждений дорожного покрытия в режиме реального времени.

Методологическую основу работы составляют методы цифровой обработки изображений, алгоритмы компьютерного зрения и современные методы глубокого машинного обучения. В работе используются сверточные нейронные сети (CNN), трансформерные механизмы внимания, методы многозадачного обучения и анализа видеоданных для построения интеллектуальной системы обнаружения дорожных повреждений. Также применяются методы предварительной обработки данных, аугментации изображений, оптимизации нейронных сетей и экспериментальной оценки качества моделей с использованием метрик детекции и сегментации.

Научные положения, выносимые на защиту:

- разработанная многозадачная архитектура нейронной сети (TCR-RoadNet) для одновременного обнаружения, классификации и сегментации повреждений дорожного покрытия в режиме реального времени;

- разработанный модуль контекстного уточнения признаков на основе кросс-масштабного трансформерного внимания, повышающий точность локализации сложных и фрагментированных дефектов;
- предложенный метод формирования инвариантных визуальных признаков, обеспечивающий устойчивость распознавания дефектов к изменениям освещенности, фоновому шуму и погодным условиям.

Основные результаты исследования:

- разработана интеллектуальная система автоматического обнаружения повреждений дорожного покрытия на основе анализа видеоданных, обеспечивающая непрерывный мониторинг дорожной инфраструктуры в режиме реального времени с использованием методов компьютерного зрения и глубокого обучения;
- предложена многозадачная архитектура нейронной сети TCR-RoadNet, предназначенная для одновременного обнаружения, классификации и сегментации дорожных дефектов. Архитектура включает многомасштабный сверточный блок извлечения признаков, Transformer Context Refinement (TCR) модуль и специализированные ветви обработки, что позволило повысить точность локализации и устойчивость распознавания сложных повреждений;
- разработан и реализован модуль контекстного уточнения признаков на основе трансформерного механизма внимания, обеспечивающий эффективный анализ пространственных зависимостей между объектами различного масштаба. Использование данного модуля позволило улучшить качество сегментации и уменьшить количество ложноположительных срабатываний при наличии теней, бликов и неоднородной текстуры асфальта;
- выполнен сбор, предварительная обработка и разметка собственного набора видеоданных дорожного покрытия, содержащего различные типы повреждений, включая продольные и поперечные трещины, сетчатые разрушения и выбоины, зафиксированные в реальных условиях эксплуатации транспортной инфраструктуры;
- проведена экспериментальная оценка эффективности разработанной модели. Результаты вычислительных экспериментов показали устойчивую сходимость нейронной сети в процессе обучения, высокие показатели точности обнаружения и сегментации дефектов, а также возможность работы системы в режиме реального времени со скоростью до 57 кадров в секунду;
- выполнен сравнительный анализ разработанной архитектуры с существующими моделями глубокого обучения для задач дорожного мониторинга, который показал повышение точности обнаружения дефектов и улучшение качества сегментации при сохранении высокой производительности системы;
- разработано программное обеспечение и веб-интерфейс интеллектуального мониторинга дорожной инфраструктуры, позволяющие автоматически визуализировать обнаруженные повреждения, выполнять их привязку к географическим координатам и отображать результаты анализа на интерактивной карте для дальнейшего использования дорожными службами и коммунальными организациями.

Научная новизна исследования заключается в разработке многозадачной нейросетевой модели TCR-RoadNet, которая одновременно находит, распознает и сегментирует дорожные дефекты в реальном времени. Ключевым элементом новизны является внедрение трансформерного механизма внимания, обеспечивающего анализ глобальных пространственных зависимостей дорожной сцены. Это позволило создать уникальный метод адаптации, благодаря которому система обеспечивает устойчивое распознавание повреждений в любую погоду и при любом освещении, игнорируя тени, лужи и визуальный шум.

Научный вклад:

- разработан и размечен уникальный набор видеоданных (датасет) изображений дорожного покрытия, содержащий различные типы повреждений, зафиксированных в реальных условиях эксплуатации транспортной инфраструктуры.
- определены и алгоритмически учтены дополнительные визуальные признаки, такие как влияние сезонных условий и изменчивость динамического освещения, что позволило существенно повысить устойчивость и робастность модели к сложным условиям съемки.
- разработана комплексная система автоматизированного мониторинга дорожной инфраструктуры, способная обрабатывать непрерывный видеопоток в режиме жесткого реального времени и предоставлять аналитическую отчетность через интегрированный веб-интерфейс.

Теоретическая значимость данной работы связана с применением методов компьютерного зрения и глубокого обучения для решения задачи обеспечения непрерывного мониторинга дорожного анализа. В отличие от большинства актуальных подходов, обрабатывающих статические изображения, данная работа увеличивает научную и методологическую базу для анализа пространственно-временных характеристик видеопоследовательностей. Разработанные методы способствуют алгоритмической поддержке интеллектуальных систем для разработки надежных систем машинного зрения в сложных условиях.

Практическая значимость исследования выражается в создании готового к внедрению программного обеспечения и специализированной информационной системы, которая автоматизирует процесс обнаружения дефектов дорожного покрытия и способствует переходу от ручных инспекций к интеллектуальным системам мониторинга дорожной инфраструктуры. Результатом исследования стала изученная и обоснованная модель нейронной сети для классификации и локализации дефектов дорог (выбоин, трещин) на основе видеомониторинга с оборудования в режиме реального времени, а также специализированное веб-приложение, реализованное в виде интерактивной панели управления. Интеграция обученной модели нейронной сети с данными веб-интерфейса позволяет автоматически привязывать обнаруженные дефекты дорог и отображать их на карте. Полученные результаты и разработанные программные средства могут быть использованы для оперативного реагирования дорожных служб, для объективного определения приоритетного порядка ремонта дорог и для предотвращения расходования бюджета на техническое обслуживание

специализированными коммунальными службами, дорожно-строительными компаниями и министерствами транспорта и дорожной инфраструктуры.

Достоверность полученных результатов обеспечивается использованием современных методов компьютерного зрения и глубокого обучения, применением сверточных нейронных сетей (CNN) и трансформерных механизмов внимания, а также проведением вычислительных экспериментов на размеченном наборе видеоданных дорожного покрытия, собранного в реальных условиях эксплуатации. Достоверность результатов подтверждается сравнительным анализом с существующими архитектурами и алгоритмами обнаружения дорожных дефектов, использованием общепринятых метрик оценки качества (mAP, IoU, Precision, Recall, F1-score), устойчивой сходимостью модели в процессе обучения и воспроизводимостью полученных результатов в различных условиях освещенности, погодных факторов и фонового шума.

Апробация диссертационной работы

Основные результаты работы были представлены и обсуждены на семинарах кафедры «Математическое и Компьютерное Моделирование» АО «МУИТ» (2024–2026 гг.), университете Asia Metropolitan University, Малайзия (2026 г.). По теме диссертации опубликовано 7 публикации, в том числе 5 публикации в рейтинговых научных изданиях, индексируемых в базе Scopus; 2 публикации в материалах международных конференций; получено 1 авторское свидетельство.

1. Olzhayev, O. M., Kulambayev, B. O., & Suliman, A. (2026). TCR-RoadNet: A Transformer-Enhanced Multi-Task Deep Learning Architecture for Real-Time Road Damage Detection and Segmentation. *Computers, Materials & Continua*. <https://doi.org/10.32604/cmc.2026.082618>

2. Kulambayev, B., Olzhayev, O., & Suliman, A. (2026). A multi-scale transformer-enhanced YOLO framework for unified road damage detection and boundary-aware segmentation. *Frontiers in Artificial Intelligence*, 9, 1834179. <https://doi.org/10.3389/frai.2026.1834179>

3. Olzhayev, O. M., Kulambayev, B. O., Suliman, A., Rustem, A., Madiyarova, A., & Omarov, B. (2026). YOLO-CBAM: A Lightweight Attention-Guided Deep Learning Framework for Real-Time Road Damage Detection. *International Journal of Advanced Computer Science & Applications*, 17(5). <https://doi.org/10.14569/IJACSA.2026.0170541>

4. Olzhayev, O. M., Kulambayev, B. O., Sakenkyzy, N., & Belisbek, M. (2026). A Real-Time Multi-Scale Feature Pyramid YOLO Architecture for Accurate and Deployment-Efficient Road Damage Detection. *International Journal of Advanced Computer Science & Applications*, 17(3). <https://doi.org/10.14569/IJACSA.2026.0170350>

5. Kulambayev, B. O., Olzhayev, O. M., Altayeva, A. B., & Zhunisbekova, Z. (2025). A Multi-Scale ROI-Aligned Deep Learning Framework for Automated Road Damage Detection and Severity Assessment. *International Journal of Advanced Computer Science & Applications*, 16(12). <https://doi.org/10.14569/IJACSA.2025.01612107>

6. Olzhayev, O., Kulambayev, B., & Omarov, B. (2025). Real-Time Pixel-Wise Segmentation of Road Surface Damage Using a 2D U-Net Architecture. *Procedia Computer Science*, 269, 131-139. <https://doi.org/10.1016/j.procs.2025.08.266>

7. Kulambayev, B., & Olzhayev, O. (2025). A Mask R-CNN Algorithm for Automated Segmentation of Asphalt Road Cracks. *Procedia Computer Science*, 269, 39-48. <https://doi.org/10.1016/j.procs.2025.08.257>

8. Олжаев О. Свидетельство на право охраны программы для ЭВМ № 69666 Республики Казахстан. Программное обеспечение TCR-RoadNet от 7.04.2026

Во всех перечисленных публикациях соискателю принадлежит ведущая роль и основные результаты исследований, анализы, модели, программы созданы автором, выводы сделаны на основе результатов, полученных от работы и исследования соискателя.

Связь с государственными программами

Диссертационное исследование выполнено в рамках грантового финансирования (ИРН AP23487192) на тему «Разработка системы обнаружения повреждений дорожного покрытия в режиме реального времени с использованием компьютерного зрения и искусственного интеллекта».

Исследование соответствует стратегическим приоритетам развития Республики Казахстан и вносит вклад в реализацию Концепции развития искусственного интеллекта в Республике Казахстан на 2024–2029 годы, утвержденной Постановлением Правительства Республики Казахстан №592 от 24 июля 2024 года, в части разработки и внедрения интеллектуальных систем компьютерного зрения, анализа видеоданных и автоматизированного мониторинга транспортной инфраструктуры. Разработанная многозадачная архитектура нейронной сети TCR-RoadNet и система обнаружения повреждений дорожного покрытия способствуют развитию отечественных технологий искусственного интеллекта, интеллектуальных транспортных систем и цифровизации дорожной инфраструктуры Республики Казахстан.

Основное содержание диссертации

Диссертационная работа фокусируется на проблеме нахождения дефектов по задачам обнаружения, классификации и сегментации повреждений дорожного покрытия на основе анализа видеоданных, где используются передовые методы глубокого обучения.

В первом разделе обосновывается актуальность проблемы деградации дорожной инфраструктуры и необходимость перехода от ресурсоемких ручных инспекций к интеллектуальным системам мониторинга. Формулируются цель и задачи исследования, а также определяются научная новизна и практическая значимость разработанной многозадачной нейросетевой архитектуры и программно-аналитического комплекса.

Во втором разделе проводится масштабный обзор литературы, включающий систематизацию типов дорожных дефектов и анализ эволюции методов их обнаружения. Рассматривается исторический переход от алгоритмов классического компьютерного зрения к современным архитектурам глубокого обучения (сверточные сети, детекторы YOLO, визуальные трансформеры). Выявляются фундаментальные пробелы в существующих исследованиях, такие как сложность обработки непрерывных видеопотоков из-за временной несогласованности, нехватка репрезентативных видеодатасетов и проблема адаптации тяжелых моделей для периферийных вычислений, что алгоритмически обосновывает направление текущей работы.

В третьем разделе подробно описываются материалы и методы исследования. Представляется процесс сбора, предварительной обработки и прецизионного аннотирования уникального набора видеоданных, зафиксированных в реальных условиях эксплуатации. Главное внимание уделяется математическому и структурному проектированию инновационной многозадачной нейросетевой архитектуры TCR-RoadNet. Детально описывается работа многомасштабной сверточной магистрали, модуля контекстного уточнения на основе кросс-масштабного трансформерного внимания (TCR), а также трех специализированных ветвей вывода: модуля отдельного обнаружения, блока уточнения классификации и сегментационной ветви с учетом границ. Дополнительно излагаются стратегия оптимизации и многозадачные функции потерь, многозадачные функции потерь (комбинация CIoU, Focal Loss и BCE+Dice)

В четвертом разделе представлены результаты комплексной вычислительной и экспериментальной оценки разработанной системы. Проводится анализ динамики сходимости модели в процессе обучения и количественная оценка семантической точности с использованием стандартизированных метрик mAP и mIoU. Выполняется детальное абляционное исследование, эмпирически доказывающее архитектурную целесообразность каждого внедренного вычислительного модуля. Сравнительный анализ с передовыми мировыми аналогами (State-of-the-Art) подтверждает, что предложенная модель TCR-RoadNet обеспечивает оптимальный компромисс между высокой точностью детализированного распознавания и способностью работать в режиме жесткого реального времени со скоростью 57 кадров в секунду. Качественный анализ визуализирует способность сети игнорировать фоновый шум и тени.

Пятый раздел посвящен подведению общих итогов диссертационного исследования. Подтверждается успешное достижение поставленной цели, обобщаются главные научные результаты и формулируются выводы о готовности разработанного аппаратно-программного комплекса к практическому внедрению в задачи автоматизированного аудита транспортных сетей.

Общий объем диссертационной работы составляет 104 страниц, включая 28 иллюстраций, 8 таблиц и список литературы из 154 использованных источников.