

Распознавание и отслеживание дефектов дорожного полотна в реальном времени на основе комплексного использования стандартных вычислительных процедур и глубоких нейронных сетей

М.О. Антонов ¹✉, И.О. Темкин ¹

¹ Университет науки и технологий МИСиС, г. Москва, 119049, Россия

Ссылка для цитирования

Антонов М.О., Темкин И.О. Распознавание и отслеживание дефектов дорожного полотна в реальном времени на основе комплексного использования стандартных вычислительных процедур и глубоких нейронных сетей // Программные продукты и системы. 2024. Т. 37. № 3. С. 421–430. doi: 10.15827/0236-235X.142.421-430

Информация о статье

Группа специальностей ВАК: 2.3.1

Поступила в редакцию: 16.05.2024

После доработки: 17.06.2024

Принята к публикации: 27.06.2024

Аннотация. В данной работе рассматриваются вопросы, связанные с распознаванием и отслеживанием дефектов дорожного покрытия в режиме реального времени. Это является важной задачей для современных систем мониторинга, беспилотного транспорта и управления дорожной инфраструктурой. Предложенная система способна выявлять дефекты дорожного покрытия с помощью глубоких нейронных сетей и отслеживать их путем присвоения уникальных идентификаторов. Используемая для распознавания дефектов модель YOLOv8 продемонстрировала высокую точность и скорость в задачах обнаружения объектов. Модель была обучена на составных наборах данных, что позволило добиться надежного обобщения. Примечательной особенностью предлагаемого подхода является отсутствие жестких требований к техническим характеристикам камер и других компонентов системы, это делает решение гибким и адаптируемым к различным условиям эксплуатации. Отслеживание дефектов реализовано с помощью венгерского алгоритма и фильтра Калмана, что обеспечивает высокую точность и надежность процесса наблюдения. Венгерский алгоритм эффективно устанавливает соответствия между обнаружениями в последовательных кадрах, а фильтр Калмана сглаживает траектории и прогнозирует будущие положения дефектов, что имеет решающее значение для точного присвоения уникальных идентификаторов и анализа данных. Представленные параметры обучения моделей, оценки потребления вычислительных ресурсов и производительности системы показывают, что предлагаемое решение имеет приемлемые характеристики для применения в режиме реального времени. Эти измерения помогают оценить затраты на внедрение и определить потенциал интеграции в существующие системы мониторинга. Таким образом, представленное решение обеспечивает высокоэффективное и точное распознавание и отслеживание дефектов дорожного покрытия в режиме реального времени, повышая безопасность и качество дорожной инфраструктуры, а также упрощая управление и планирование работ по техническому обслуживанию.

Ключевые слова: распознавание в реальном времени, глубокие нейронные сети, отслеживание объектов, YOLOv8, венгерский алгоритм, фильтр Калмана, дефекты дорожного полотна

Введение. Национальный проект «Безопасные и качественные автомобильные дороги» ставит своей целью приведение региональных дорог и дорожной сети городских агломераций в соответствие нормам. Реализация этого проекта обусловила увеличение объема контрактов по комплексной диагностике дорог и оценке их технического состояния. Организации, участвующие в проекте, обязаны предоставлять заказчику следующую информацию на основе проведенной диагностики: видеоданные обследуемого участка, ведомость дефектов дорожного покрытия и отчет о комплексной диагностике с оценкой транспортно-эксплуатационного состояния.

За последнее десятилетие было проведено множество исследований, посвященных применению методов искусственного интеллекта для анализа изображений дорожной обстанов-

ки. Учитывая наличие обширного объема качественных видеоданных в дорожной отрасли и стремительное развитие компьютерного зрения, актуальным и перспективным направлением исследований становится выявление дефектов дорожного покрытия не только по статичным изображениям автомобильных дорог, но и в реальном времени. Данная тематика так же актуальна и в области автономных транспортных средств, где точное и оперативное обнаружение дефектов на дороге является ключевым для обеспечения безопасности и эффективности самоуправляемых автомобилей. Достижения в этой области не только способствуют общему улучшению состояния дорог, но и играют решающую роль в расширении возможностей и надежности систем навигации автономных транспортных средств. По мере развития индустрии искусственного интеллек-

та и производительности вычислительных систем в целом обнаружение нарушений как отдельная функция становится недостаточным критерием оценки качества, поскольку сформированные данные требуют дальнейшей обработки путем их классификации по степени критичности того или иного дефекта и выполнения статистического анализа для оперативного принятия управленческих решений.

Научная новизна исследования заключается в разработке универсальной методики для обнаружения и отслеживания дефектов дорожного покрытия, в частности, выбоин и проломов. В отличие от традиционных подходов, которые часто требуют адаптации к различным типам дефектов, в рамках настоящей работы предложено обобщенное решение на базе венгерского алгоритма и фильтра Калмана. Представленный механизм обеспечивает плавную интеграцию в другие сценарии обнаружения и отслеживания, например, деформации, износ поверхности и дорожной разметки, предлагая масштабируемое решение для комплексной диагностики состояния дорог.

Краткий анализ существующих подходов

Традиционно основной формой обнаружения дефектов, определенным образом классифицируемых согласно ГОСТ Р 50597-2017, являлся ручной визуальный осмотр профессиональными инспекторами, однако подобный процесс контроля малоэффективен, затратен материально и по времени. Более того, результаты проверки являются субъективными, поскольку полностью зависят от опыта отдельного специалиста. С другой стороны, технологии мониторинга при помощи специальных датчиков (стационарных и мобильных) имеют ограничения, связанные как с точностью и оперативностью обработки информации, так и с их высокой стоимостью.

Разберем основные известные подходы к их идентификации и анализу.

Подход на основе вибрации. Под основной идеей подхода подразумевается использование GPS, акселерометра и гироскопа для картографирования дорожного покрытия [1, 2]. К достоинствам использования метода можно отнести возможность анализа состояния дорожной сети в реальном времени с достаточно высокой точностью [3]. Стоит отметить, что данные точность и полнота охвата зависят в том числе и от того, был или нет наезд на исследуемый дефект, что является недостатком. К дополнительным не-

достаткам следует отнести ограничения в использовании специфичного устройства ПО [4].

3D-лазер. В данном типе исследования применяется 3D-лазерное сканирование, используя сеточный подход для определения характеристик дефектов [5]. В рамках данного метода формируется облако точек в трехмерном пространстве, что позволяет проводить анализ и классифицировать обнаружения по степени критичности [6]. К недостаткам относятся дороговизна используемого оборудования, а также небольшая дальность обнаружения, что, в свою очередь, накладывает ограничения на применение этого метода в реальном времени на борту автономных транспортных средств [7].

3D-стереовизуализация. Развитием и дополнением предыдущего метода является 3D-стереовизуализация. Метод подразумевает формирование трехмерной поверхности дорожного покрытия на основе набора входных изображений. Например, в [8] авторы представили подход по восстановлению полной трехмерной поверхности дорожного полотна. Методика включала в себя исправление любых наблюдаемых искажений, извлечение особенностей дефектов и 3D-реконструкцию. Метод, создающий набор точек в трехмерном пространстве и позволяющий с высокой точностью отображать дорожное покрытие, рассмотрен в [9]. Используя стереоизображения и технологии обработки изображений, система может идентифицировать различные дорожные дефекты, такие как выбоины, неровности, трещины и т.д. Стоит отметить очень высокую точность анализа дорожного покрытия, обеспечиваемую данным методом. Необходимо упомянуть и требования к производительности оборудования, поскольку трехмерной реконструкции поверхности необходима большая вычислительная мощность.

Искусственные нейронные сети. Данный метод наблюдения использует обработку полученных с камер изображений с последующим обнаружением дефектов средствами сверточных сетей и сетей глубокого обучения, причем в качестве камеры необязательно выступает профессиональное оборудование – в некоторых случаях достаточно мобильного телефона с камерой высокого разрешения [10]. Ряд исследователей предложили использовать модифицированную модель ResNet50-RetinaNet при локализации выбоин по тепловизионным изображениям вместо стандартных средств видеofиксации [11]. В работе [12] исследована модель YOLOv3 на основе собранного для обучения

набора данных, состоящего из 1 500 изображений дорог Индии. К достоинствам подхода можно отнести экономичность, позволяющую при этом определить форму и площадь дефекта с достаточно высокой точностью. Стоит отметить присущий данному методу компромисс между разнообразием набора использованных при обучении данных и точностью результата, временем обработки, а также размером модели в целом, который может являться одновременно с этим существенным недостатком – неполный набор данных и его некачественная разметка сводят возможности методологии на нет.

Принимая во внимание достоинства метода, в рамках настоящей работы представлено решение для обнаружения дефектов на базе более современной архитектуры YOLOv8 в составе с разработанным алгоритмом отслеживания, а приведенные недостатки нивелированы за счет ручного сбора и разметки данных.

Отслеживание дефектов покрытия

Основная идея метода заключается в выделении и сохранении набора идентификаторов, каждый из которых связан с конкретным объектом, и в динамическом обновлении их на основе новых данных по каждому из кадров. Предлагается использовать модель движения в комплексе с алгоритмом целераспределения, где в качестве модели используется фильтр Калмана, а целераспределения – венгерский алгоритм. Представленный метод основан на оценке ряда параметров предсказания в каждом кадре, а именно: текущей скорости движения (по координатам X и Y по отдельности) и размера (высоты и ширины обнаруженного объекта), а значит, может быть применим для различных типов дефектов, если таковые будут обнаружены моделью глубокого обучения на этапе первичной идентификации. Соотношения сторон ограничительной рамки вокруг объекта всегда берутся от самого первого обнаружения. Скорости движения предсказываются при помощи фильтров Калмана, а матрица пересечений ограничительных рамок по IoU-коэффициенту, который отражает степень перекрытия ограничительной рамки детектируемого объекта с предыдущим кадром, а сами детекции назначаются при помощи венгерского алгоритма.

Опишем основные принципы работы метода, представленные в виде последовательности этапов.

Инициализация. Начиная с первого кадра видеоряда каждому обнаруженному объекту

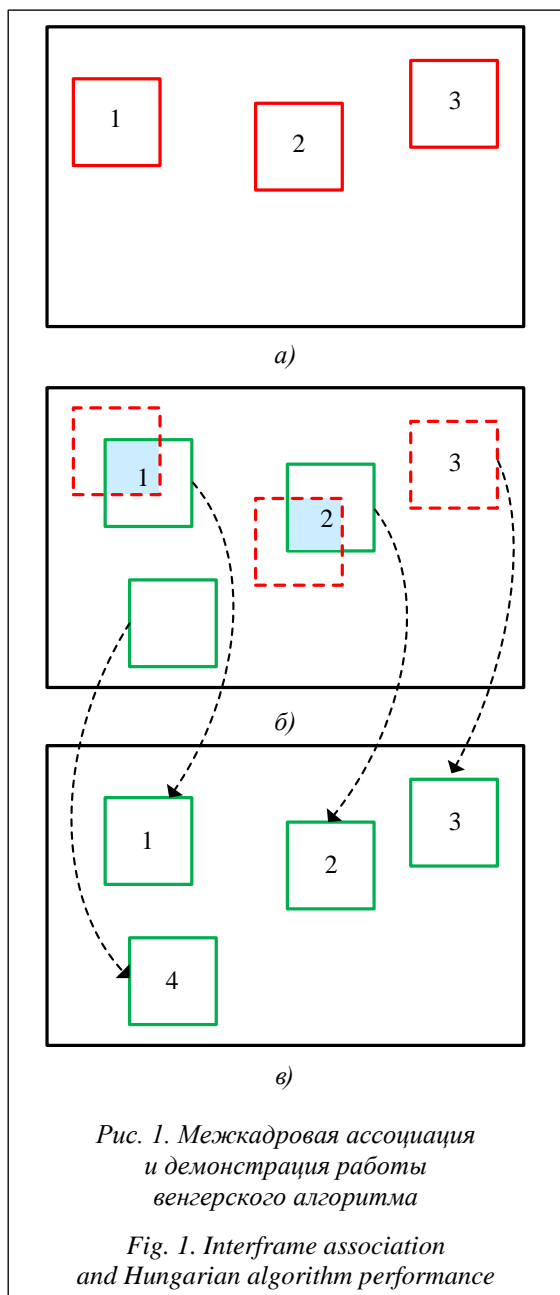
присваивается уникальный идентификатор. В дальнейшем для каждого нового обнаружения вводится новый идентификатор, а отношения между ними при переходе от кадра к кадру реализуются при помощи ассоциации.

Ассоциация. Венгерский алгоритм, также известный как алгоритм Куна–Манкреса, позволяет определить, совпадает ли дефект дорожного покрытия из одного кадра с дефектом из последующего кадра. Сам процесс ассоциации представлен на рисунке 1 на примере трех исходных обнаружений.

Рассмотрим межкадровую ассоциацию и принцип работы на примере двух последовательных кадров живого видеопотока. На первом кадре при инициализации (рис. 1а) обученная модель предсказывает три дефекта дорожного покрытия, в результате чего за счет ассоциации происходит присваивание трех идентификаторов для первого, второго и третьего дефектов соответственно. На втором кадре (рис. 1б) модель идентичным образом предсказывает три дефекта, за исключением третьего, ранее обнаруженного, в результате чего новый объект приобретает четвертое значение идентификатора, а третий подлежит удалению при его отсутствии на последующих кадрах (подробнее в секции «удаление» настоящего раздела), что приводит к состоянию (рис. 1в), позволяющему отслеживать каждое из обнаружений. Стоит отметить, что изменение координат и скорости первого и второго дефектов осуществляется средствами фильтра Калмана.

Предсказание и корректировка. Для формирования оптимальной оценки состояния фильтр Калмана использует динамическую модель системы, известные управляющие воздействия и дальнейшее множество последовательных измерений. Алгоритм на данной стадии состоит из двух повторяющихся фаз – предсказание и корректировка, где в первом случае происходит предсказание состояния в следующий момент времени (с учетом погрешности измерений), а во втором – корректировка за счет поступившей новой информации уже имеющегося предсказанного значения (идентичным образом с учетом погрешности и зашумленности данной информации).

Уравнения для предсказаний и корректировок представлены в матричной форме. Первоначальная оценка параметров осуществляется с помощью вектора средних значений и матрицы ковариаций. Вектор состояния состоит из абсциссы u и ординаты центра области обнаружения объекта v , а также площади области s ,



соотношения сторон r этой области (считается постоянным) и изменения первых трех параметров – скоростей, принимающих нулевое состояние при инициализации. Ковариация – матрица неопределенности в оценке, где большее число означает большую неопределенность. Выбор меньшего значения неопределенности для области обнаружения объекта, а также его площади и соотношения сторон обусловлен наличием погрешности в 10 пикселей, полученной эмпирически на основе среднего размера используемого изображения и размеров детектируемых объектов:

$$x = [u \ v \ s \ r \ \dot{u} \ \dot{v} \ \dot{s}], \quad (1)$$

$$P = \begin{bmatrix} 10 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 10 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 10 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1000 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1000 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1000 \end{bmatrix} \quad (2)$$

Дополняя ассоциацию, осуществляется измерение областей детекции для каждого объекта соответственно (рис. 1а). Одновременно с этим происходит инициализация среднего значения и ковариации в соответствии с выражениями (1) и (2) для каждой области обнаружения. В состоянии б) в фазе предсказания состояния системы определяются области обнаружения дефектов из состояния а) в момент времени б), после чего полученные значения сопоставляются с новыми областями при помощи венгерского алгоритма, а переход в фазу коррекции завершает цикл.

Рассмотрим подробнее последовательность матричных преобразований.

1. Предсказание состояния системы:

$$\hat{x}_k^- = F\hat{x}_{k-1} + Bu_{k-1}, \quad (3)$$

где \hat{x}_k^- – предсказание состояния системы в текущий момент времени; F – матрица перехода между состояниями (динамическая модель системы приведена далее); \hat{x}_{k-1} – состояние системы в прошлый момент времени; B – матрица применения управляющего воздействия (при его отсутствии значение принимается как нулевое); u_{k-1} – управляющее воздействие в прошлый момент времени, которое в рассматриваемой задаче принимается равным нулю.

Матрица перехода F имеет ненулевые и составленные таким образом коэффициенты, чтобы в случае умножения слева на вектор состояния x значения результирующего вектора представляли собой сумму позиционных значений с их производными:

$$F = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (4)$$

2. Предсказание ошибки ковариации:

$$P_k^- = FP_{k-1}F^T + Q, \quad (5)$$

где P_k^- – предсказание ошибки; P_{k-1} – ошибка в прошлый момент времени; Q – ковариация

шума процесса. На основе размерности вектора состояния x , размерности матриц ковариации и шума процесса составят 7×7 для P и Q соответственно.

3. Вычисление усиления Калмана:

$$K_k = P_k^- H^T (H P_k^- H^T + R)^{-1} \quad (6)$$

где K_k – усиление Калмана; H – матрица измерений, отображающая отношения измерений и состояний; R – ковариация шума измерения, которая в случае использования каких-либо радаров задается конструктивно, а в случае использования модели YOLOv8 – в пикселях (по 1 пикселю в центре и по 10 для ширины и высоты соответственно). Размерность матриц измерений H и ковариации шума измерения R , а также выбор коэффициентов обусловлены набором входящих от модели предсказания значений u , v , s и r для центра обнаруженного объекта и его сторон соответственно:

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \quad (7)$$

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 10 \end{bmatrix}. \quad (8)$$

4. Обновление оценки с учетом измерения z_k :

$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - H \hat{x}_k^-), \quad (9)$$

где z_k – измерение в текущий момент времени.

5. Обновление ошибки ковариации:

$$P_k = (I - K_k H) P_k^-, \quad (10)$$

где I – матрица идентичности.

Удаление. Завершаются все преобразования стадией удаления с заданной глубиной хранения объекта, составляющей 10 кадров для настоящего исследования. Принимая значение IoU равным 0.3, достаточно совпадения на 30 % при пересечении областей обнаружения. В случае отсутствия области обнаружения в детекторе (например, вероятность нахождения в кадре или значение IoU ниже пороговых значений) на протяжении 10 кадров подряд равносильно удалению из памяти программы, а значит, на 11-м кадре объект уже получит новый уникальный идентификатор.

Анализ и подготовка наборов данных

В соответствии с поставленными для настоящего изыскания целями были проведены ана-

лиз и подготовка наборов данных для обучения модели, включая наиболее популярные и находящиеся в открытом доступе на момент проведения вычислительных экспериментов изображения с размеченными дефектами дорожного полотна.

В частности, рассмотрев самый популярный из них, авторы [13] представили набор данных из 47 420 изображений дорог разных стран (Японии, Индии, Чешской Республики, Норвегии, США и Китая).

Предварительный анализ показал, что только 26 520 изображений из всего набора аннотированы, иными словами, фактически около половины изображений не содержат заявленных на них дефектов или для них отсутствует разметка. Более глубокое исследование выявило, что из оставшихся 26 520 изображений только 6 544 могут быть классифицированы как содержащие рассматриваемые в настоящей работе выбоины и проломы, а качество их разметки неудовлетворительное (обнаружено множество примеров, где контуры детекции не находятся по границам объекта, вместо дефекта аннотирован, например, светофор и т.п.).

Принимая во внимание этот факт, было принято решение о самостоятельной подготовке набора данных, учитывая опыт и рекомендации других исследователей, но не опираясь на их результаты во избежание попадания некачественных данных в финальную выборку, а разметку осуществить вручную. Результатом стала подборка 30 615 изображений с разрешением не менее $1\,280 \times 1\,280$ только выбоин и проломов, обучение на которых производилось в рамках данного исследования.

При подготовке был сделан акцент на разнообразии сред, в которых запечатлен тот или иной дефект. Например, учитывая потребность применения результатов исследования в среде с большим количеством грунтовых покрытий вместо асфальтовых, а также необходимость размещения камеры на лобовом стекле движущегося транспортного средства вместо создания специализированных видеолaborаторий, выборка была расширена соответствующими, в том числе запечатленными самостоятельно, примерами. Дополнительно набор данных содержит изображения заполненных водой выбоин и проломов (<http://www.swsys.ru/uploaded/image/2024-3/18.jpg>), гарантируя возможность распознавания различных сценариев и повышая надежность результата.

Обучение моделей

В каждой категории YOLOv8 имеются пять моделей для обнаружения, сегментации и классификации, где YOLOv8n самая быстрая и малая по размеру, в то время как YOLOv8x самая точная, но одновременно с этим самая медленная из представленных. Обучение моделей проводилось на графических модулях NVIDIA GeForce RTX 3090 TI в количестве четырех единиц с идентичными параметрами обучения для всех версий модели, представленными в таблице 1.

Таблица 1

Параметры обучения моделей YOLOv8

Table 1

YOLOv8 model training parameters

Параметр	Значение
Эпоха	55
Размер партии	32
Оптимизатор	Adam
Темп обучения	0.000714
Размер изображений	1 280

С целью ускорения обучения был использован модуль DistributedDataParallel (DDP), входящий в состав пакета PyTorch. DDP реализует параллелизм данных на уровне модуля, который может работать на нескольких машинах, а использующие DDP приложения должны порождать несколько процессов и создавать по одному экземпляру DDP для каждого из них. Рекомендуемый способ применения DDP – со-

здать один процесс для каждой реплики модели, которая может охватывать несколько физических устройств. Каждый графический модуль в рамках текущего исследования представляет собой отдельный процесс, и связь между ними осуществляется с помощью стандартного для операционных систем межпроцессного взаимодействия.

На основе прогресса обучения для модели YOLOv8x (рис. 2) можно сделать вывод, что ограничение в виде обозначенных выше 55 эпох необязательно и для достижения лучших результатов обучение может быть продолжено. Дополнительно графики с префиксами train и val (тренировки и валидации соответственно) имеют схожие формы, что свидетельствует о качестве обучения и хорошей работе на данных, которые ранее не демонстрировались. Учитывая, что train loss не приняла нулевое значение, а val loss не искривляются под конец обучения вверх, можно сделать вывод, что модель не была переобучена. Результаты прогресса обучения моделей других размеров носят схожий характер и идентичные для YOLOv8x линии тренда.

После обучения в результате валидации были получены представленные в таблице 2 характеристики моделей. Точность и производительность моделей оценивались на основе параметра mAP50-95 – метрики, представляющей среднюю точность (Average Precision, AP) по нескольким порогам пересечения областей IoU от 0.5 до 0.95. Метрика AP измеряет точ-

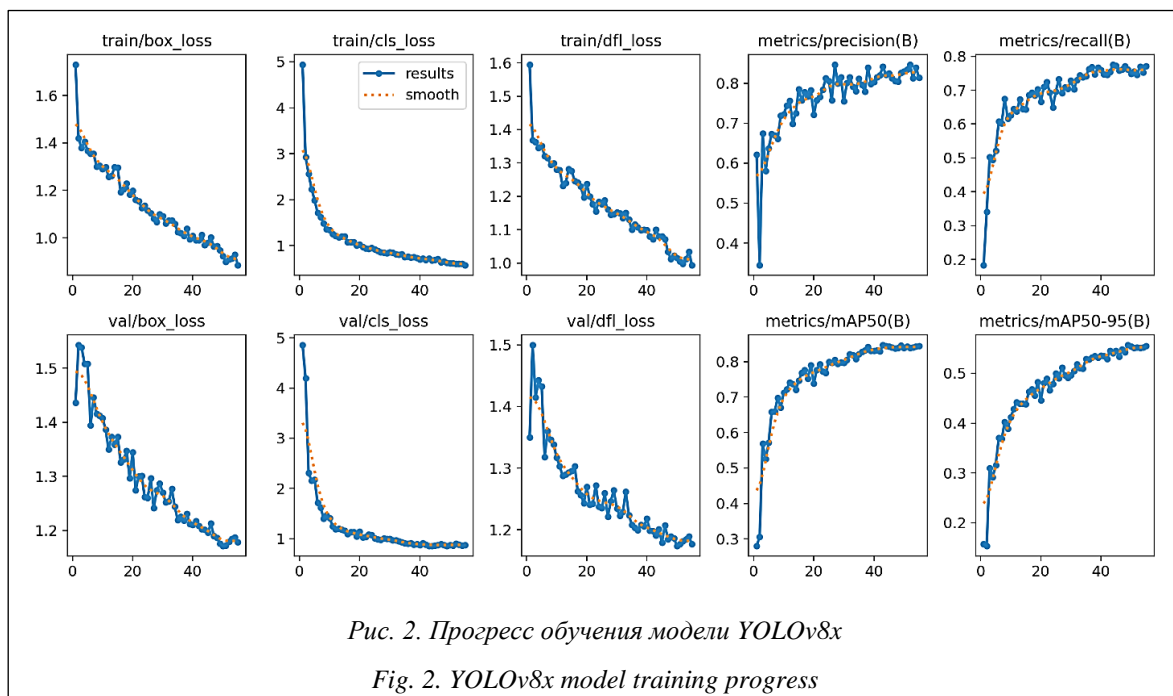


Рис. 2. Прогресс обучения модели YOLOv8x

Fig. 2. YOLOv8x model training progress

Таблица 2

Характеристики обученных моделей YOLOv8

Table 2

Characteristics of trained YOLOv8 models

Модель	Размер, МБ	mAP50-95, ед.	Параметры, млн	FLOPs, млрд
YOLOv8n	6.4	55.8	3.2	8.1
YOLOv8s	22.6	57.6	11.2	28.4
YOLOv8m	52.1	59.9	25.9	78.7
YOLOv8l	87.8	60.3	43.7	164.8
YOLOv8x	136.8	60.6	68.2	257.4

ность модели, суммируя кривую точности-полноты. Усредняя AP по диапазону порогов IoU, mAP50-95 дает всестороннюю оценку производительности модели обнаружения при различных уровнях строгости перекрытия.

Можно наглядно проследить, что наиболее точная, но наименее быстрая модель YOLOv8x достигает значения в 60.6 mAP50-95, в то время как наихудший, но вместе с тем максимально производительный результат достигается моделью YOLOv8n – 55.8 mAP50-95. Одновременно с этим более крупные модели требуют большего количества вычислений, что, в свою очередь, приводит к снижению энергоэффективности. Дополнительно стоит подчеркнуть, что наилучшие характеристики с точки зрения размера, параметров и качества обнаружения у модели YOLOv8m. Здесь и далее будет наглядно прослеживаться тенденция в пользу выбора именно этой версии, а не самой большой YOLOv8x.

Обнаружение дефектов покрытия

По аналогии с приведенными характеристиками для обучения графического модуля тестирование быстродействия моделей осуществляется на NVIDIA GeForce RTX 3090 TI, а в качестве CPU выступают Intel Xeon E5-2696 v4 с частотой 2.20 ГГц в количестве двух единиц и 256 гигабайт RAM.

В таблице 3 представлены результаты потребления вычислительных ресурсов и производительности для каждого типа модели YOLOv8, где под потреблением подразумевается количество занятой видеопамати, а под производительностью – время обработки одного изображения на графическом модуле. Быстрая обработка результатов имеет важное значение для беспилотных транспортных средств, чтобы избежать выбоин. Это делает обязательным минимизацию времени обработки. Учитывая предыдущие результаты точ-

ности, также можно сделать вывод, что модель YOLOv8m обладает наилучшим результатом в том числе по отношению к потреблению и быстродействию, а дальнейшее увеличение скорости работы за счет применения YOLOv8s и YOLOv8n не приносит значительных результатов, а только снижает качество обнаружения в целом. Одновременно с этим модель YOLOv8n обеспечивает исключительное среднее время обработки изображения 8 мс, сохраняя размер модели всего 6.4 МБ, что упрощает ее размещение во встраиваемых системах.

Таблица 3

Потребление и производительность моделей YOLOv8

Table 3

Consumption and performance of YOLOv8 models

Модель	Потребление VRAM, МБ	Быстродействие, мс
YOLOv8n	688	8
YOLOv8s	716	9
YOLOv8m	804	11
YOLOv8l	918	15
YOLOv8x	1086	19

Учитывая, что целью настоящей работы является в том числе исследование возможности применения обычных камер, а не дорогостоящих передвижных видеолaborаторий, на рисунке 3 приведены результаты обнаружения дефектов дорожного полотна из движущегося транспортного средства.

Дополнительно стоит отметить, что результаты тестирования производительности не включают в себя какие-либо методы увеличения быстродействия, например, квантизацию, объединение входящих изображений в группы, конвертацию для поддержки такими сторонними фреймворками, как TensorRT, раскадровку на графическом модуле и т.п. Подобного рода оптимизации находятся вне рамок настоящей работы, но в случае их применения (а ни-



Рис. 3. Работа модели в реальном времени в движущемся транспортном средстве

Fig. 3. Real-time model operation in a moving vehicle

каких ограничений для этого не существует) производительность процесса обнаружения и системы в целом потенциально возрастет.

Рассматривая результаты производительности в составе с упомянутым ранее алгоритмом отслеживания, стоит обратить внимание на представленное в таблице 4 время обработки полного цикла для каждой из моделей. Данное значение является важным показателем для оценки быстродействия и включает в себя суммарно затраченное время на получение кадра из видеопотока, на его предварительную обработку, предсказание моделью, на постобработку, отслеживание и на вывод с формированием соответствующих результатов.

Таблица 4

Затрачиваемое на полный цикл обработки кадра время

Table 4

Time spent on a full frame processing cycle

Модель	Быстродействие, FPS
YOLOv8n	32
YOLOv8s	31
YOLOv8m	29
YOLOv8l	26
YOLOv8x	23

На основании представленных значений можно сделать вывод, что модель YOLOv8m демонстрирует наилучшие по отношению к другим моделям результаты быстродействия. Стоит отметить, что программа выполнялась

на одном ядре центрального процессора Intel Xeon E5-2696 v4 с частотой 2.20 ГГц, а значит, выполнение одноименного кода на более современном оборудовании или применение методов многопоточного программирования вместо текущего однопоточного увеличит производительность системы в целом.

Заключение

Применение обученной в рамках настоящей работы модели YOLOv8 – отличное решение для обнаружения дефектов дорожного полотна, которое обеспечивает баланс между точностью, скоростью и эффективностью использования ресурсов, что позволяет добиться значительных успехов в повышении безопасности дорожного движения и в обслуживании инфраструктуры. Изучение существующих средств диагностики и анализ имеющегося опыта обнаружения дефектов позволяют сделать вывод, что имеющееся техническое обеспечение требует усовершенствования, модернизации и разработки средств диагностики нового поколения. Прослеживается огромный потенциал для использования методов глубокого обучения, что особенно актуально для интеллектуального контроля дорожного движения, где в настоящее время имеется ограниченный объем исследований и данных. Кроме того, исследование подчеркивает важность разнообразного и надежного набора обучающих данных.

Версия модели YOLOv8m продемонстрировала наилучшие результаты, достигнув значе-

ния точности 59.9 mAP при быстродействии в 29 FPS для полного цикла анализа видеопотока в реальном времени на одном ядре центрального процессора. С дальнейшим развитием темы данной работы допустимо применение ее результатов не только для интеллектуального обслуживания городской инфраструктуры, но и для обнаружения и отслеживания иных классов объектов при помощи, например, беспилотных летательных аппаратов.

Таким образом, в рамках настоящего изыскания исследовано обнаружение дефектов дорожного полотна средствами предварительно обученной модели YOLOv8 с последующим отслеживанием объектов при помощи фильтра Калмана и венгерского алгоритма. Критерии оценки качества полученных результатов включали в себя время, размер модели, потребление ресурсов, время обработки полного цикла и надежность в различных условиях.

Список литературы

1. Темкин И.О., Клебанов Д.А., Дерябин С.А., Конов И.С. Метод определения состояния технологических дорог карьера при управлении взаимодействием роботизированных элементов горнотранспортного комплекса // Горный журнал. 2018. № 1. С. 78–82. doi: 10.17580/gzh.2018.01.14.
2. Rana S. Vibration based pavement roughness monitoring system using vehicle dynamics and smartphone with estimated vehicle parameters. Results in Engineering, 2021, vol. 12, art. 100294. doi: 10.1016/j.rineng.2021.100294.
3. Kim Y.-M., Kim Y.-G., Son S.-Y. et. al. Review of recent automated pothole-detection methods. Appl. Sci., 2022, vol. 12, no. 11, art. 5320. doi: 10.3390/app12115320.
4. Kandoi A., Agarwal H., Revadekar A., Chaudhari B.-N. Pothole detection using accelerometer and computer vision with automated complaint redressal. Proc. ICPSC, 2021, pp. 257–264. doi: 10.1109/ICSPC51351.2021.9451762.
5. Issaoui A.E., Feng Z., Lehtomäki M. et al. Feasibility of mobile laser scanning towards operational accurate road rut depth measurements. Sensors, 2021, vol. 21, no. 4, art. 1180. doi: 10.3390/s21041180.
6. Feng Z., Issaoui A.E., Lehtomäki M. et al. Pavement distress detection using terrestrial laser scanning point clouds – Accuracy evaluation and algorithm comparison. ISPRS Open J. of Photogrammetry and Remote Sensing, 2022, vol. 3, art. 100010. doi: 10.1016/j.ophoto.2021.100010.
7. De Blasii M.R., Di Benedetto A., Fiani M. Mobile laser scanning data for the evaluation of pavement surface distress. Remote Sens., 2020, vol. 12, no. 6, art. 942. doi: 10.3390/rs12060942.
8. Fan R., Ozgunalp U., Wang Y., Liu M., Pitas I. Rethinking road surface 3-D reconstruction and pothole detection: From perspective transformation to disparity map segmentation. IEEE Transactions on Cybernetics, 2021, vol. 52, no. 7, pp. 5799–5808. doi: 10.1109/TCYB.2021.3060461.
9. Staniek M. Stereo vision method application to road inspection. The Baltic J. of Road and Bridge Eng., 2017, vol. 12, no. 1, pp. 38–47. doi: 10.3846/bjrbe.2017.05.
10. Maeda H., Sekimoto Y., Seto T., Kashiyama T., Omata H. Road damage detection and classification using deep neural networks with images captured through a smartphone. Computer-Aided Civil and Infrastructure Eng., 2018, vol. 33, no. 12, pp. 1127–1141. doi: 10.1111/mice.12387.
11. Gupta S., Sharma P., Sharma D., Gupta V., Sambual N. Detection and localization of potholes in thermal images using deep neural networks. Multimed Tools Appl., 2020, vol. 79, pp. 26265–26284. doi: 10.1007/s11042-020-09293-8.
12. Dharneeshkar J., Aniruthan S., Karthika R., Parameswaran L. Deep learning-based detection of potholes in Indian roads using yolo. Proc. ICICT, 2020, pp. 381–385. doi: 10.1109/ICICT48043.2020.9112424.
13. Arya D., Maeda H., Ghosh S.K., Toshniwal D., Omata H., Kashiyama T., Sekimoto Y. Crowdsensing-based Road Damage Detection Challenge (CRDDC'2022). Proc. Int. Conf. Big Data, 2022, pp. 6378–6386. doi: 10.1109/Big-Data55660.2022.10021040.

Real-time recognition and tracking of road surface defects based on complex using of standard computational procedures and deep neural networks

Mikhail O. Antonov ¹✉, Igor O. Temkin ¹

¹ University of Science and Technology MISIS, Moscow, 119049, Russian Federation

For citation

Antonov, M.O., Temkin, I.O. (2024) 'Real-time recognition and tracking of road surface defects based on complex using of standard computational procedures and deep neural networks', *Software & Systems*, 37(3), pp. 421–430 (in Russ.). doi: 10.15827/0236-235X.142.421-430

Article info

Received: 16.05.2024

After revision: 17.06.2024

Accepted: 27.06.2024

Abstract. The paper considers several issues related to real-time recognition and tracking of road surface defects. This is an important task for modern monitoring systems, unmanned transportation and road infrastructure management. The pro-

posed system identifies pavement defects using deep neural networks and tracks them by assigning unique identifiers. The YOLOv8 model has demonstrated high accuracy and speed in object detection tasks during defect recognition. Due to training the model on composite datasets, generalization was reliable. A notable feature of this approach is the absence of strict requirements to the technical characteristics of cameras and other system components. This makes the solution flexible and adaptable to various operating conditions. Defect tracking is based on place using the Hungarian algorithm and Kalman filter. This ensures high accuracy and reliability of the observation process. The Hungarian algorithm efficiently establishes correspondences between detections in consecutive frames. The Kalman filter smooths trajectories and predicts future defect positions. This is crucial for accurate assignment of unique identifiers and data analysis. Model training parameters, estimates of computational resource consumption and system performance show that this solution has acceptable performance for real-time applications. These measurements help to estimate implementation costs and determine the potential for integration into existing monitoring systems. Thus, the presented solution provides highly efficient and accurate real-time recognition and tracking of pavement defects. It improves the safety and quality of road infrastructure and simplifies management and maintenance planning.

Keywords: real-time recognition, deep neural networks, object tracking, YOLOv8, Hungarian algorithm, Kalman filter, road surface defects

References

1. Temkin, I.O., Klebanov, D.A., Deryabin, S.A., Konov, I.S. (2018) 'Haul road condition determination under controlled interaction of robotic elements in open pit mining and transport system', *Mining J.*, (1), pp. 78–82 (in Russ.). doi: 10.17580/gzh.2018.01.14.
2. Rana, S. (2021) 'Vibration based pavement roughness monitoring system using vehicle dynamics and smartphone with estimated vehicle parameters', *Results in Engineering*, 12, art. 100294. doi: 10.1016/j.rineng.2021.100294.
3. Kim, Y.-M., Kim, Y.-G., Son, S.-Y. et al. (2022) 'Review of recent automated pothole-detection methods', *Appl. Sci.*, 12(11), art. 5320. doi: 10.3390/app12115320.
4. Kandoi, A., Agarwal, H., Revadekar, A., Chaudhari, B.-N. (2021) 'Pothole detection using accelerometer and computer vision with automated complaint redressal', *Proc. ICPSC*, pp. 257–264. doi: 10.1109/ICSPC51351.2021.9451762.
5. Issaoui, A.E., Feng, Z., Lehtomäki, M. et al. (2021) 'Feasibility of mobile laser scanning towards operational accurate road rut depth measurements', *Sensors*, 21(4), art. 1180. doi: 10.3390/s21041180.
6. Feng, Z., Issaoui, A.E., Lehtomäki, M. et al. (2022) 'Pavement distress detection using terrestrial laser scanning point clouds – Accuracy evaluation and algorithm comparison', *ISPRS Open J. of Photogrammetry and Remote Sensing*, 3, art. 100010. doi: 10.1016/j.ophoto.2021.100010.
7. De Blasiis, M.R., Di Benedetto, A., Fiani, M. (2020) 'Mobile laser scanning data for the evaluation of pavement surface distress', *Remote Sens.*, 12(6), art. 942. doi: 10.3390/rs12060942.
8. Fan, R., Ozgunalp, U., Wang, Y., Liu, M., Pitas, I. (2021) 'Rethinking road surface 3-D reconstruction and pothole detection: from perspective transformation to disparity map segmentation', *IEEE Transactions on Cybernetics*, 52(7), pp. 5799–5808. doi: 10.1109/TCYB.2021.3060461.
9. Staniek, M. (2017) 'Stereo vision method application to road inspection', *The Baltic J. of Road and Bridge Eng.*, 12(1), pp. 38–47. doi: 10.3846/bjrbe.2017.05.
10. Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., Omata, H. (2018) 'Road damage detection and classification using deep neural networks with images captured through a smartphone', *Computer-Aided Civil and Infrastructure Eng.*, 33(12), pp. 1127–1141. doi: 10.1111/mice.12387.
11. Gupta, S., Sharma, P., Sharma, D., Gupta, V., Sambual, N. (2020) 'Detection and localization of potholes in thermal images using deep neural networks', *Multimed Tools Appl.*, 79, pp. 26265–26284. doi: 10.1007/s11042-020-09293-8.
12. Dhameeshkar, J., Aniruthan, S., Karthika, R., Parameswaran, L. (2020) 'Deep learning-based detection of potholes in Indian roads using yolo', *Proc. ICICT*, pp. 381–385. doi: 10.1109/ICICT48043.2020.9112424.
13. Arya, D., Maeda, H., Ghosh, S.K., Toshniwal, D., Omata, H., Kashiyama, T., Sekimoto, Y. (2022) 'Crowdsensing-based Road Damage Detection Challenge (CRDDC'2022)'. *Proc. Int. Conf. Big Data*, pp. 6378–6386. doi: 10.1109/Big-Data55660.2022.10021040.

Авторы

Антонов Михаил Олегович¹, аспирант,
ассистент, amo.mikhail@gmail.com
Темкин Игорь Олегович¹, д.т.н.,
профессор, igortemkin@yandex.ru

Authors

Mikhail O. Antonov¹, Postgraduate Student,
Assistant, amo.mikhail@gmail.com
Igor O. Temkin¹, Dr.Sci. (Engineering),
Professor, igortemkin@yandex.ru

¹ Университет науки и технологий МИСиС,
г. Москва, 119049, Россия

¹ University of Science and Technology MISIS,
Moscow, 119049, Russian Federation