**ФЕДЕРАЛЬНОЕ АГЕНСТВО СВЯЗИ   
ОРДЕНА ТРУДОВОГО КРАСНОГО ЗНАМЕНИ  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ**

**«МОСКОВСКИЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ СВЯЗИ И ИНФОРМАТИКИ»**

Практика по теме: Методы распознавания и обработки изображений. Перевод статьи.

Выполнили студенты 1 курса:

Удалов И.Д. и

Белов С.П.

БВТ1903

Проверила: Мосева М. С.

Москва 2020г.

**Система обнаружения повреждений на дорогах на основе глубоких нейронных сетей для управления физическими активами**

***Аннотация.***

Исследования по выявлению повреждений дорожных покрытий с использованием методов обработки изображений были активно проведены, достигая достаточно высокой точности уровня обнаружения. Пока что большинство исследований было сосредоточено только на обнаружении присутствия или отсутствие повреждений; однако, в реальных сценариях, дорожные менеджерам необходимо четко понимать тип ущерба и его масштабы, чтобы принять эффективные меры заранее или выделять необходимые ресурсы для ремонта. Более того, в настоящее время имеется мало единых и открыто доступных наборов данных о повреждениях дорог, что приводит к отсутствию общего эталона для обнаружения повреждений дорог. Такой набор данных может быть использован для самых разных приложений; здесь он предназначен для использования в качестве компонента приобретения инструмента управления физическими активами, который может помочь правительственным учреждениям в планировании или обслуживании инфраструктуры компании, чтобы они могли осуществлять процедуры прогнозного обслуживания. В этом документе мы делаем два вклада для решения этих проблем. Во-первых, мы представляем крупномасштабный набор данных о повреждениях дорог, который включает в себя более сбалансированный и репрезентативный набор повреждений, которых не было в предыдущих исследованиях. Этот набор данных состоит из 18,0345 изображений повреждений дорог, снятых с помощью смартфона, установленного на автомобиле, с 45 435 случаев повреждения дорожного покрытия (линейные, боковые и аллигаторные трещины(растрескивание дорог), выбоины и различные виды пятен). Для генерации этого набора данных, мы получили изображения из нескольких общедоступных наборов данных и дополнили их краудсорсинговые изображения, которые вручную помечаются для дальнейшей обработки. Изображения были получены при различных погодных условиях с разным освещением. На каждом изображении мы аннотировали (составили) ограничивающую рамку, представляющую местоположение, тип ущерба и его масштаб. Во-вторых, мы разработали различные типы методов универсального обнаружения объекта, как традиционные (каскадный классификатор LBP), так и основанные на глубоком методе обучения, в частности, MobileNet и RetinaNet, которые доступны для встраиваемых и мобильных устройств и реализации с приемлемой производительностью для многих приложений. Мы сравнили точность и время вывода всех этих моделей с другими современными моделями, достигая более высокой точности во всех восьми классов, представленных в наборе данных, предоставленных исследователями из Университета Токио и другими связанными работами с меньшим временем вывода.

1. ***Введение***

Исследования по выявлению повреждений дорожных покрытий с использованием обработки изображений и методы обучения машин были активной областью исследований как в развитых, так и в развивающихся странах [1-4]. Это важный вопрос, так как дороги являются одними из самых важных гражданских инфраструктур в каждой стране, они прямо и косвенно способствуют экономику стран и, что более важно, благосостоянию и безопасности их граждан. Таким образом, содержание дорог имеет первостепенное значение, и во многие страны внедрены механизмы контроля в них имеются стандарты для осуществления этого процесса. Тем не менее, процессы проверки и регистрации дорожных повреждений остаются огромной проблемой, поскольку правительственные учреждения все еще пытаются поддерживать современные базы данных о таких структурных повреждениях, что затрудняет распределение ресурсов на ремонтные работы в информированном порядке. Проблема усугубляется тем, что количество экспертов, которые могут оценить такие структурные повреждения ограничены, и, кроме того, методы, как правило, используемые для сбора данных с мест требуют много времени, затрат, требуют нетривиальный уровень экспертизы, так же очень субъективен и склонен к ошибкам. Следовательно, как академические усилия, так и коммерческие инициативы были проведены для облегчения этого процесс, используя сочетание сложных технологий [5]. Большинство этих подходов объединяют различные датчики (то есть инерционные профилировщики, сканеры), но также методы визуализации, которые продемонстрировали, что они особенно подходят для этой задачи. Информация, собранная этими датчиками и камерами, может быть передана в алгоритмы машинного обучения и объединена с мобильными системами сбора данных и подходами облачных вычислений для автоматизации процесса проверки или для создания комплексных решений. Такие усилия продемонстрировали многообещающие результаты [6], но многие из них имеют ограничения по определенным типам классов повреждения дорог (трещины, пятна, выбоины) и сравнительно небольшие и недостаточно полные наборы данных. Кроме того, традиционные Системы "AI" для классификации дорожных повреждений плохо масштабируются и не достаточно универсальны, чтобы работать в тяжелых ситуациях, которые обычно обнаруживают эксперты дорожной инспекции. Эта ситуация резко изменилось с появлением подходов компьютерного зрения, основанных на архитектуре глубокого обучения, которые продемонстрировали огромные успехи в нескольких задачах компьютерного зрения и постепенно пробились в системы дорожного контроля. Такие системы обнаружения и классификации дорог, основанные на компьютерном зрении, значительно дешевле, чем конкурирующие технологии, и при соответствующей подготовке они могут привести к внедрению сложных и экономически эффективных систем приобретения. Такие Решения могут быть улучшены, если они будут реализованы на мобильных устройствах, которые будут собирать данные для дорог: накопление ущерба и хранение в облачных вычислительных технологиях для дальнейшей обработки и анализ больших данных. Такой подход может привести к созданию гео-локализованных базы данных, содержащих информацию о текущем состоянии дорог и других инфраструктур гражданского строительства, которые могут обновляться и использоваться для целей прогнозирования и планирования. Это идеально согласуется с промышленностью 4.0 и парадигмами цифрового преобразования и может быть легко реализованна как бизнес-модель, которая может быть легко воспроизведена. Например, Великобритания запустила Digital Roads Challenge для достижения этой цели.

В этой статье мы представляем первоначальный подход к инструменту управления цифровыми активами, ориентироваными на управление дорогами и улицами. По сравнению с предыдущими работами, где усилия сосредоточены на конкретных типах повреждений (выбоины [1], трещины [2] и пятна [3]), наше предложение основывается на предыдущей работе в этой области, предлагая большой набор данных различных типов повреждений асфальта (продольные трещины, аллигаторные трещины, выбоины, неровности, пятна и т.д.), насколько нам известно, один из самых полных наборов. Некоторые из структурных повреждений, обычно встречающихся в полевых условиях, изображены на рисунке 1; Эти изображения представляют структурные повреждения, а не повреждения на инфраструктурных знаках и разметке.



**Рис. 1. Примеры структурных повреждений, обычно встречающихся в литературе:**

**a-b) линейные трещины, c) пилинг, d) трещины аллигатора, e) пятна и f) выбоины.**

Предложенный набор данных о повреждениях дорог был использован для обучения различным машин навыкам на основе «детекторов объектов», в которых интересующие объекты представляют собой структурные повреждения в наборе данных. Как будет более подробно описано в следующем разделе, таким системам чтобы быть полезными, нужно работать в режиме реального времени, и поэтому мы проверили набор данных на легких детекторах, таких как каскадные классификаторы LPB, и в более поздних детекторах одиночного выстрела (SSD) с глубоким обучением, таких как MobileNet и RetinaNet. Два последних были разработаны таким образом, чтобы они могли эффективно работать на мобильных устройствах (например, смартфонах) в режиме реального времени, достигая при этом высокой средней точности (mAP). Остальная часть этой статьи организована следующим образом: в разделе 2 мы дополнительно мотивируем потребность в цифровых системах управления активами, особенно для дорог и улиц. Затем в разделе 3 мы обсуждаем современное состояние обнаружения повреждений дорог с использованием различных изображений модальности для обеспечения некоторого контекста, но мы фокусируем наше внимание на этих методах на основе компьютерного зрения и современных подходов машинного обучения. В разделе 4 мы представить предлагаемый набор данных и первоначальный подход к управлению цифровым активами дорог, а также инструменты и методы, использованные для проверки предложения. После этого в разделе 5 мы обсуждаем эти результаты с учетом уровня техники и описываем будущие направления исследований, завершение статьи.

1. ***Мотивация***

Содержание дорог и уличной инфраструктуры является серьезной проблемой, особенно в тропических странах, где асфальт и другие гражданские объекты более подвержены структурным повреждениям. Тем не менее, эта проблема заканчивается этими регионами, и становится распространенной в развитых странах из-за старения их инфраструктуры. Таким образом, чтобы справиться с этими проблемами, правительственные страны разработали конкретные процедуры и стандарты, чтобы помочь экспертам в оценке повреждение асфальта и других материалов, таких как бетон. На местах эксперты традиционно проводят инспекции на основе этих стандартов, придерживаясь двух основных подходов: либо путем непосредственного наблюдения людей, либо путем количественного анализа с использованием методов дистанционного зондирования. Визуальный подход требует опытных дорожных менеджеров, что делает его трудоемким, дорогим и подвержен ошибкам. Кроме того, традиционные методы, как правило, противоречивы и ограничены точечным наблюдением, затрудняющим масштабирование до крупных городов или географических районов, так же нужно создать и обслуживать цифровые корпусы с полученными данными. Проблема усугубляется, поскольку приобретения, сделанные традиционными инспекциями, не хранятся в цифровом виде, и, таким образом, государственные органы не могут использовать возможности цифровой революции (то есть анализ больших данных) для решения этих проблем в более информированном и эффективным способом. В отличие от более поздних количественных систем, основанных на инспекционных системах, использующих так называемые мобильные измерительные системы (MMS) или методы лазерного сканирования, были предложены и внедрены в некоторых странах [5]. MMS может получать высокоточную геопространственную информацию о дороге с помощью мобильного телефона; эта система содержит блок глобальной системы позиционирования (GPS), внутреннюю единицу измерения (IMU), цифровые измеримые изображения, полученные с помощью цифровых камер, лазерный сканер и лидар(технология получения и обработки информации об удалённых объектах с помощью активных оптических систем). Хотя это количественные системы контроля очень точные, они значительно дороже и сложнее в эксплуатации, а также их более трудно приобрести, сложнее управлять и обслуживать небольшие муниципалитеты, в которых отсутствуют финансовые ресурсы. Для того, чтобы облегчить вышеупомянутые проблемы, много исследований были проведены для облегчения и автоматизации описанного процесса ручной проверки выше. Например, область гражданского строительства имеет долгую историю использования методов неразрушающего зондирования и съемки, обычно основанных на компьютерном зрении, таких как 3D-изображения [6] и методы дистанционного зондирования [7] для различных применений. Прежний подход имеет несколько преимуществ по сравнению с традиционными методами контроля, а также что касается других подходов, которые используют исключительно традиционные методы воображения, они лучше справляться с изменениями освещенности и погодных условий.

Хотя такие инструменты предоставляют возможность для частых, всеобъемлющих и количественных обследований транспортных инфраструктур, они дорогостоящие и их труд внедрить и поддерживать, так как такие системы обычно полагаются на транспортные средства, как показано на рисунке 2. Поэтому такие решения просто недоступны многим странам и тем более для отдельных муниципалитетов. Таким образом, исследования ищут другие методы, чтобы уменьшить сложность и затраты, связанные с обследованием и осмотром дорожной инфраструктуры.



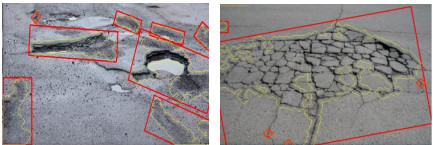
**Рис. 2. Пример системы MMS, разработанной PaveMetrics, проанализированной в [5].**

Как мы увидим в следующем разделе, такие усилия в основном сводятся к методам обработки и компьютерного зрения, как основная информация, используемая инспекторами. Она может быть получена из обычной визуализации при работе только с классификацией проблемы, которые сделали большие успехи в этой области благодаря сочетанию эффективного и легкого универсального объекта, основанного на глубоком обучении детекторы и мобильные вычислительные технологии для целей вывода.

1. ***Связанные работы***

Для автоматической проверки повреждений на дорогах обычно требуются надежные алгоритмы компьютерного зрения с высокой степенью интеллекта, которые могут быть просты в использовании и работать в режиме реального времени. Однако большинство систем, работающих в этом ключе изначально задумывалось как средство помощи экспертам в данной области и было ограничено возможностями алгоритмов ML того времени, которые были не очень надежными и которые было трудно реализовать в портативных системы. Для полного ознакомления с различными методами визуализации читатель направлен на превосходные обзоры в области трехмерной визуализации [7], дистанционного зондирования [8], компьютерного зрения и обработки изображений в целом [9, 10]. Самая большая проблема для автоматизированных систем обнаружения повреждений на дорогах заключается в последовательном достижении высокой производительности, с точки зрения точности, в различных сложных условиях (из-за изменений в освещении, погодных условиях и других проблем). Несмотря на эти проблемы, в литературе было предложено несколько систем для обнаружения отдельных структурных повреждений, как описано в разделе 1. Однако большинство из этих работ не смогли решить более сложные сценарии при разработке своих систем машинного обучения и искусственного интеллекта, из-за отсутствия общепринятых и общедоступных наборов данных. Набор данных Kitti [4] можно использовать для таких целей, поскольку он содержит тысячи примеров дорожных повреждений; однако его основное назначение - автономное вождение и его использование для обнаружения повреждений на дорогах, к сожалению, может оказаться трудным и в конечном итоге бесполезным занятием.

Поэтому академическим сообществом было широко признано, что существует необходимость в обширном наборе данных, специально предназначенном для оценки дорожного ущерба, для двух основных целей. Во-первых, для того, чтобы протестировать и выполнить сравнительные тесты среди конкурирующих решений, это поможет ускорить исследования и разработки в этой области. Во-вторых, гомогенизировать(упростить) процесс сбора данных, поскольку такие наборы данных необходимы для создания надежных решений на основе машинного обучения для обнаружения и проверки дорог и асфальта. Эта проблема была частично решена, поскольку некоторые наборы данных были доступны, но все ещё предстоит проделать большую работу, поскольку дороги в разных странах сильно различаются. Несмотря на эти проблемы, доступные наборы данных использовались в течение многих лет для реализации решений, основанных на машинном обучении, с использованием функций, основанных на компьютерном зрении, и алгоритмов классификации, таких как SVM, для обнаружения трещин [11] и выбоин [12]. По мере того, как наборы данных становились все больше, возрастал интерес к развертыванию подходов, основанных на глубоком обучении (DL). Хорошо известно [13], что алгоритмы DL продемонстрировали впечатляющие результаты во множестве областей, связанных с компьютерным зрением, и стоят за революцией автономного вождения. В этом смысле было предпринято несколько попыток использовать эти возможности в решении проблемы, с которой мы сталкиваемся, например, некоторые результаты были очень впечатляющими для определенных классов повреждений асфальта, таких как трещины [14, 15]. Одно из преимуществ методов глубокого обучения по сравнению с традиционными подходами состоит в том, что для получения обучающих образцов можно использовать более дешевые или менее сложные устройства формирования изображений (то есть смартфоны) вместо гораздо более дорогих платформ MMS, описанных в предыдущем разделе. Подходы, основанные на мобильных устройствах, были исследованы ранее с использованием более традиционных детекторов компьютерного зрения, таких как LBP-каскадные классификаторы [16, 17], но авторы не сравнивали свои результаты с более поздними детекторами, основанными на DL. Кроме того, хотя результаты были похвальными, они использовали пользовательские наборы данных и были ограничены более «статическим анализом» (их нельзя было выполнять в режиме реального времени в очень сложных сценариях), на рис. 3 показана пара примеров обнаружения структурных повреждений с использованием основанный на LPB подход на OpenCV, используя наш набор данных.



**Рис. 3. Обнаружение трещин и выбоин аллигатора с использованием самого основного алгоритма ML, исследуемого в этой работе: каскадный классификатор LBP, обратите внимание на большое количество FP.**

С другой стороны, недавние достижения в общих алгоритмах обнаружения и классификации объектов [18] позволили теперь реализовать очень сложные и ресурсоэффективные алгоритмы DL в устройствах с ограниченными возможностями, таких как мобильные телефоны и интеллектуальные камеры, некоторые из недавних реализаций начали появляться в различных областях и, в частности, в области дорожной инспекции и оценки. В этом смысле теперь также возможно выполнять этап развертывания (обнаружение и обнаружение структурных повреждений) в режиме реального времени, используя также недорогие мобильные устройства [20-21], либо для отдельных инспекций на месте, либо на автомобиле. что делает его привлекательной альтернативой дорогим платформам MMS. Эти успехи стали возможными благодаря разработке двухступенчатых детекторов, таких как Fast R-CNN, и впоследствии одноступенчатых детекторов, таких как YOLO (для подробного обсуждения читатель направлен на превосходный обзор [18]). Первая категория может достигать высокой точности (средняя точность или mAP), но обычно неэффективна, поскольку не может быть развернута на устройствах с ограниченными ресурсами [19]. Последняя категория универсальных детекторов объектов со временем неуклонно совершенствуется: новые архитектуры, основанные на функциональных пирамидальных сетях (т. Е. RetinaNet и MobileNet), обеспечивают удовлетворительную точность для решения сложных задач и требуют относительно небольших занимаемых площадей, что приводит к внедрению в мобильные телефоны для обнаружения повреждений на дорогах. и сбор информации с географической привязкой [19, 21], как показано на рисунке 4 а) и б)



**Рис. 4. Примеры систем обнаружения повреждений дорог на смартфоне. a) Система Университета Токио [21] и б) Решение RoadBotics [22]**

Эти успехи также стали возможными благодаря внедрению крупномасштабных наборов данных, требуемых алгоритмами DL, требующими больших объемов данных, особенно для конкретных задач, таких как повреждения дорог, в качестве набора данных, представленного Токийским университетом [22]. Этот набор данных содержит типы структурных повреждений (трещины, пятна и т. д.), Но он все еще ограничен, поскольку он представляет несбалансированные классы (особенно для выбоин, которые быстро ремонтируются в Японии и, таким образом, не представлены). Таким образом, одним из вкладов этой статьи является дополнение исходного набора данных большим количеством образцов для каждого класса, в частности для классов выбоин и трещин аллигатора, которые очень распространены в тропических регионах и таких странах, как Мексика, и, таким образом, необходимы для большого разнообразия видов и случаев использования.

В следующих разделах мы опишем набор данных, начальные эксперименты, выполненные с различными алгоритмами, и обсудим их сравнение с уровнем техники, описанным выше, а также наметим будущие направления исследований.

1. **Предлагаемое решение**

Учитывая прогресс, достигнутый в этой области, интеграция описанных выше технологий в парадигмы «умных городов», таких как системы управления цифровыми активами, была лишь вопросом времени, и некоторые государственные и частные сервисные компании начали использовать информацию, собранную в базы данных о повреждениях дорог различными способами. Основная идея состоит в том, чтобы использовать прогресс, достигнутый в различных областях (AI, IoT, мобильные и облачные вычисления), для создания систем, которые, с одной стороны, могут облегчить работу дорожных инспекторов, но которые также могут быть использованы для реализации сквозных задач, таких как конечные решения для создания большой базы данных о структурных повреждениях от отдельных дорог до улиц в средних и крупных городах, приводящих к так называемым «цифровым близнецам» (рис. 4 б)).

Эти системы могут помочь улучшить работу муниципалитетов или правительств на федеральном уровне, помогая им управлять своими активами. Примеры платформ с этими возможностями начали развертывать в США (RoadBotics [23]), но их сервис больше ориентирован на более общую оценку качества улиц.

Основная идея нашей работы заключается в использовании возможностей алгоритмов ИИ в сочетании с мобильными и облачными вычислениями, чтобы предложить систему, которая может получать геопривязанные изображения повреждений дорог с целью создания «цифровой карты», которую можно использовать для многих целей, таких как анализ большого числа данных на огромной территории. В этой статье мы сосредоточимся на создании набора данных и обучении системы сбора данных, другие компоненты системы будут описаны в последующей работе.

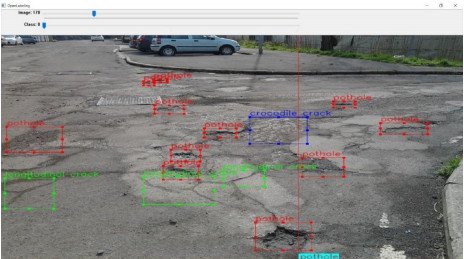
* 1. ***Используемый набор данных***

Как упоминалось ранее, исследования по обнаружению повреждений на дорогах страдали от отсутствия унифицированных наборов данных. Эта проблема была частично решена сообществом с различными наборами данных, различных размеров и характеристик; например, исследователи в Университете ди Рома Тре представили набор данных [16] с тремя основными классами структурных повреждений (линейные трещины и трещины аллигатора) и использовали его для реализации системы сбора данных с помощью мобильного телефона. Более поздние попытки были направлены на включение более широкой категории классов, таких как набор данных Токийского университета [22], который включает в себя 8 типов повреждений (изображения размером 600x600 пикселей), как показано в таблице 1.

**Таблица 1. Типы повреждений дорог в наборе данных, предложенном в [21], и их определения**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Тип повреждения** | | | **Подробности(конкр. детали)** | **Код класса** |
| Трещина | Линейная трещина | Продольная | От колес( трещины, полосы) | **D00** |
| На строительной соед. части | **D01** |
| Поперечная  (боковая) | С равным интервалом | **D10** |
| На строительной соед. части | **D11** |
| Трещины аллигатора | | Неполный участок, Обширный участок | **D20** |
| Другие повреждения | | | Борозды, Выпуклости, Выбоины,  Разложение(разделение) | **D40** |
| **D43** |
| Размытие белой линии/  Пешеходного перехода | **D44** |

Коды в Таблице 1 используются Правительством Японии для классификации дорожных повреждений и ремонтниками для классификации структурных повреждений в их работе, их сложности и стоимости. Этот набор данных является крупнейшим в настоящее время доступным и широкоиспользуемым для реализации систем дорожного контроля с использованием архитектур с глубоким обучением, таких как SSD, таких как RetinaNet [20] и MobileNet [21]. Хотя этот набор данных является относительно большим, некоторые классы или экземпляры повреждений плохо представлены, например, класс выбоин, что связано с тем, что эти повреждения быстро устраняются в Японии.



**Рис. 5. Показан в качестве примера обучающего экземпляра в нашем наборе данных, метки были вручную аннотированы с использованием Open Label Master, обратите внимание на видимость класса выбоин на этом изображении.**

Тем не менее, выбоины на различных этапах разрушения являются одним из наиболее важных структурных повреждений, которые необходимо отслеживать со стороны правительственных органов, и поэтому одним из вкладов нашей работы является включение большего количества экземпляров в классы в наборе данных, представленных авторами должны иметь более сбалансированное представление, как это можно наблюдать в **Таблице 2**. Экземпляр расширенного набора данных, содержащий несколько меток классов, показан **на Рисунке 5**. Мы включили больше экземпляров классов D00, D10 и D11 (продольные трещины различных типов) и для D40 (выбоины), для которых было добавлено несколько сотен примеров, как показано в третьем ряду **Таблицы 2**.

**Таблица 2. Количество экземпляров повреждений в каждом классе для исходного набора данных, представленного в [21], и того, который использовался в этой работе. Обратите внимание, что классы выбоин и трещин лучше представлены.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Набор данных/ Тип** | **D00** | **D01** | **D10** | **D11** | **D20** | **D40** | **D43** | **D44** |
| **Maeda[21]** | 2,678 | 3,789 | 742 | 636 | 2,541 | 409 | 817 | 3,733 |
| **Modified** | 2,978 | 3,789 | **1042** | **1036** | 3,341 | **1,609** | 817 | 3733 |

Используя этот расширенный или модифицированный набор данных, мы провели сравнительный анализ с предыдущими работами в литературе, чтобы оценить, можно ли достигнуть какого-либо выигрыша в производительности, что в большинстве случаев имело место быть. В следующем разделе мы обсудим, как набор данных использовался для обучения набора «универсальных детекторов объектов», которые являются основой нашей системы сбора данных и MMS. Следует отметить, что большинство изображений были получены с «над дорогой» (сверху) с помощью мобильных устройств.

Выбор системы сбора данных был обоснован в разделе 2 с «технической» точки зрения, но следует подчеркнуть, что во многих странах установка камеры вне автомобиля является нарушением закона [22]. Таким образом, было предложено большое количество решений на основе недорогих камер для мобильных телефонов, и, как упоминалось в разделе 3, они небольшого размера и не требуют высокого энергопотребления (производительности) в режиме реального времени, в то же время достигая высокой точности обнаружения и классификации.

Мы уже обсуждали преимущества использования различных недавно разработанных «общих объектов обнаружения», основанных на архитектурах глубокого обучения; в дальнейшем мы обсудим, как мы обучили различные из этих алгоритмов, позже мы сравним их с точки зрения производительности.

* 1. ***Обучение***

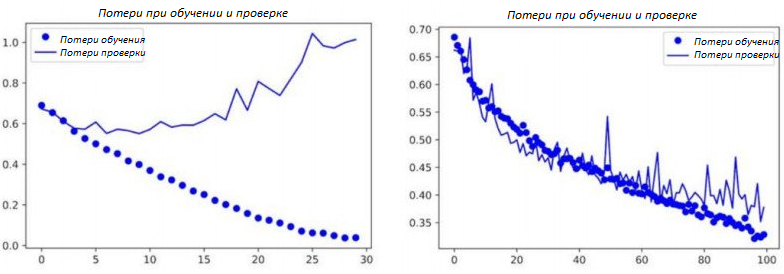
Для полноты картины мы реализовали три типа детекторов объектов, концентрируясь на тех из них, которые можно применять на мобильных и встраиваемых платформах: каскадный классификатор LBP (как в [16]), а также более поздние универсальные детекторы объектов на основе DL, в частности RetinaNet (как в [20]) и MobileNet (как в [22])

В случае RetinaNet, как обсуждалось в [20], эта модель может быть обучена различными магистральными нейронными сетями для целей: извлечения признаков и (RestNet или VGG в его различных конфигурациях) точности в зависимости от времени вывода, получая в целом лучшие результаты по сравнению с другими популярными детекторами одиночного выстрела (SSD), такими как YOLO, в то время как у модели подходящие для встроенных или мобильных реализаций. Например, RetinaNet с VGG19 в качестве магистрали требует 115,7 МБ и может достигать 0,5 с, что достаточно быстро для большинства приложений в этой области.

Мы провели обучение описанным выше моделям глубокого обучения, используя следующую методологию. Помеченный набор данных случайным образом перетасовывается и разделяется на обучающий набор и проверочный набор, 30% процентов изображений каждого класса использовались для обучения, следя за тем, чтобы модель не превышала; для этой цели мы использовали выпадение и увеличение данных по исходному набору данных (используя различные преобразования изображений, такие как обрезка, деформация, вращение и др.)

Увеличение данных - это метод для «синтетического» генерирования большего количества обучающих данных из существующих обучающих примеров, подвергая последние множеству случайных преобразований, которые дают правдоподобные изображения. Процесс увеличения данных может значительно сократить потери при проверке, как показано на рисунке 6, где показаны графики для обучения и проверки для двух идентичных ConvNet, один без использования увеличения данных, а другой с использованием стратегии.

Использование дополнения данных однозначно решает проблему переобучения, которая может наблюдаться в левой части рисунка (это означает, что потери при проверке уменьшаются, а затем снова увеличиваются, тогда как потери при обучении уменьшаются для тренировочного набора). Правая часть рисунка показывает, как можно утратить валидацию при помощи дополнения данных и других стратегий.



**Рис. 6. Потеря достоверности при обучении и проверке для 2 идентичных DNN:**

**а) без увеличения данных и б) увеличение данных. Потеря ясно показывает, как D.A предотвращает переоснащение.**

Описание DNN, а также этап обучения были реализованы при использовании Keras и Tensorflow на Google с графическими процессорами Tesla K80; модели были обучены партией 24 и числом шагов 300, используя Адама в качестве оптимизатора, адаптируя скорость обучения, как обсуждалось в [22]. Как мы увидим в следующем разделе, предлагаемый подход получил улучшения по сравнению с предыдущими работами, особенно для недопредставленных классов в исходном наборе данных о повреждениях дорог.

1. ***Результаты и обсуждение***

В наших экспериментах этапы обучения и тестирования проводились с использованием Google Colab, что позволяет нам иметь доступ к высокопроизводительным вычислениям (графический процессор Tesla K80 и другие ресурсы) и размещать нашу DB в облаке Google, избегая при этом любых вычислительных нагрузок. На этапе оценки мы также использовали этот инструмент, импортируя в обученные модели 30% проверочного набора; Кроме того, мы провели эксперименты на видео в реальном времени, чтобы проверить возможности предлагаемых моделей в различных сценариях.

В таблице 3 показаны полученные результаты для трех «детекторов универсальных объектов», протестированных в наших экспериментах. Как упоминалось ранее, мы решили реализовать LBP-каскадный классификатор, чтобы сравнить наши результаты с результатами, представленными в [16], так как мы использовали часть их набора данных. Мы также сообщаем результаты для подходов глубокого обучения, о которых сообщают в литературе, которые используют RetinaNet [20] и MobileNet [22]. Предложенные модели достигли неизменно лучших результатов, чем те, о которых сообщалось в литературе, с использованием тех же методов ([16], [20] и [22]) соответственно, при одновременном достижении более высокой точности и времени вывода, как мы обсудим позже в этом разделе.

**Таблица 3. Количество экземпляров повреждений в каждом классе для исходного набора данных, представленного в [21], и того, который использовался в этой работе. Обратите внимание, что классы выбоин и трещин представлены лучше.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Измерение/Тип** | D01 (Длинные (в длину) трещины) | D11 (Широкие(в ширину) трещины) | D20(Трещины аллигатора) | D40 (Выбоины) |
| **LBP-каскад** | 0.7080 | 0.7182 | 0.7625 | 0.7476 |
| **MobileNet** | 0.8124 | 0.9032 | 0.9234 | 0.9675 |
| **RetinaNet** | **0.9148** | **0.9511** | **0.9522** | **0.9823** |

В **Таблице 3** мы решили показать результаты только для 4 из наиболее зарегистрированных классов (продольные и боковые трещины, трещины аллигатора и выбоины), так как они являются наиболее распространенными, и это облегчает сравнение с предыдущей работой. Следует отметить, что другие работы, обсуждаемые в данной области техники в **Разделе 3**, получили хорошие результаты для отдельных классов повреждений дорог (даже с точностью до пиксельного уровня), но они не поддаются мобильной или встроенной реализации видения.

Единственная работа в литературе, которая до сих пор проводила всесторонний анализ, - это работа исследователей из Токийского университета [21], которая реализовала несколько общих детекторов объектов с использованием Inception V2 и MobileNet. Авторы получили лучшие результаты с последних, но в литературе были предложены более современные детекторы, такие как RetinaNet, для которых мы получили еще более лучшие результаты, как показано в **Таблице 3**.

Чтобы оценить улучшения, полученные с помощью нашего расширенного набора данных, мы показываем сравнительный анализ между вышеупомянутым методом классификации и обнаружения на основе MobileNet [21] и лучшей из наших моделей, который показан в **Таблице 4**. Детектор и классификатор SSD выполнили свою задачу хорошо, с высокой точностью и отзывом, особенно по сравнению с классами D11 и D40 (боковые трещины и выбоины, соответственно). Это может быть связано с большим количеством обучающих примеров в нашем наборе данных, проблема, о которой сообщили авторы в [21]. Авторы упомянули класс D43 (размытие белой линии) в качестве встречного примера, так как количество примеров ограничено, но мы считаем, что этот класс относительно легче обнаружить и классифицировать, что может объяснить результаты.

**Таблица 4. Количество экземпляров повреждений в каждом классе для исходного набора данных, представленного в [22], и того, который использовался в этой работе. Обратите внимание, что классы выбоин и трещин представлены лучше (выделил их, код D01, D11, D20, D40).**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Измерение** | **Тип дорожных повреждений** | | | | | | | |
| **Maeda и др. [21]** | D00 | **D01** | D10 | **D11** | **D20** | **D40** | D43 | D44 |
| **Память** | 0.40 | **0.89** | 0.20 | **0.05** | **0.68** | **0.02** | 0.71 | 0.85 |
| **Точность** | 0.73 | **0.64** | 0.99 | **0.95** | **0.68** | **0.99** | 0.85 | 0.66 |
| **Правильность** | 0.81 | **0.77** | 0.92 | **0.94** | **0.83** | **0.95** | 0.95 | 0.81 |
| **Лучшая модель (наша)** |  | | | | | | | |
| **Память** | 0.60 | **0.90** | 0.40 | **0.40** | **0.76** | **0.70** | 0.80 | 0.70 |
| **Точность** | 0.87 | **0.70** | 0.89 | **0.92** | **0.92** | **0.88** | 0.87 | 0.82 |
| **Правильность** | 0.91 | **0.81** | 0.92 | **0.95** | **0.95** | **0.98** | 0.95 | 0.84 |

Точности, полученные нашей моделью на основе RetinaNet, в которой, как и в случае авторов из [20], мы использовали VGG19 в качестве магистральной сети. Тем не менее, авторы в этой статье используют одну метрику (mAP 0,8279) для отчета о своих результатах, не выполняя сравнения для каждого класса.

В **Таблице 5** мы сообщаем о некоторых особенностях полученной модели на основе RetinaNet по сравнению с двумя другими современными моделями. Во-первых, размер модели относительно компактен (125,5 МБ), при этом достигается низкое время вывода (0,5 с) и более высокая средняя точность (mAP), что делает наше предложение одним из самых лучших.

**Таблица 5. Сравнение метрик для нашей лучшей модели с современным уровнем техники.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Метрика/Модель** | **Размер модели** | **Время вывода** | **Лучшая mAP** | **Тип устройства** |
| **RetinaNet [20]** | 115.7 MB | 0.5 s | 0.8279 | мобильный |
| **MobileNet [22]** | N/A | 1.5 s | N/A | мобильный |
| **RetinaNet (ours)** | 125.5 MB | 0.5 s | 0.91522 | мобильный |

1. ***Выводы и будущая работа***

В этом исследовании мы представили первый блок для управления цифровыми активами, универсального детектора объектов, обученного распознавать и классифицировать повреждения дорог от неподвижных изображений или видео в реальном времени с высокой точностью и малым временем вывода. Мы протестировали различные модели и подходы для решения этой проблемы, основанные на традиционных компьютерных подходах и более поздних архитектурах глубокого обучения.

Для обучения наших моделей мы развернули относительно большой набор данных, первоначально представленный Токийским университетом, дополнив его изображениями мексиканских и итальянских дорог, чтобы компенсировать дисбаланс, присутствующий при некоторых дорожных повреждениях, таких как выбоины. Полученные результаты ясно показывают, что RetinaNet может превзойти другие современные детекторы дорожных повреждений, он подходит для мобильной или встроенной реализации (например, для приложений ADAS) при работе с приемлемыми скоростями (0,5 секунды в нашей текущей реализации), что было продемонстрировано, что более чем достаточно для транспортного средства со скоростью 40 км / с, чтобы избежать утечки или дублирования информации, как сообщалось в [22] для времени вывода 1500 мс.

В будущем мы планируем интегрировать модель, работающую на мобильном устройстве, в полноценную платформу управления цифровыми активами, работающую на веб-сервисах Amazon. Основная идея состоит в том, чтобы использовать возможности системы облачных вычислений для создания и обслуживания цифрового изображения дорог региона или улиц города, описанного в разделе 2. Эта геолокализованная база данных может использоваться для целей мониторинга и прогнозирования. Использование в сочетании с прогнозирующей моделью приводит к более информированному процессу в распределении ресурсов для технического обслуживания, а также необходимой периодичности

***Подтверждения***

Мы благодарим профессора Benedetto из Университета де Рома Тре (Италия) за любезное предоставление некоторых изображений, используемых для дополнения набора данных из Университета Токио (Япония), наряду с другими изображениями, полученными нашей группой в Гвадалахаре, Мексика. Мы также благодарим сотрудников Vidrona LTD (Эдинбург, Великобритания) за помощь в определении проблемы, рассматриваемой в этом документе, особенно для Shailendra и Ashutosh Natraj.

***Ссылки***

1. 1.Koch, C., Asce, A.M.; Jog, G.M.; and Brilakis, I.: Automated Pothole Distress Assessment Using Asphalt Pavement Video Data, Journal of Computing in Civil Engineering, Volume 27, Issue 4, (2013)
2. 2.Oliveira, H. and Correia, P.L.: Automatic Road Crack Detection and Characterization, in IEEE Transactions on Intelligent Transportation Systems, vol. 14, no. 1, pp. 155-168, 2013.
3. 3.Radopoulou, S.C and Bralakis, I.: Patch Detection for pavement assessment, Automation in Construction, Volume 53, May 2015, Pages 95-104
4. 4.Geiger, A., Lenz, P., Stiller, C., & Urtasun, R.: Vision meets robotics: The KITTI dataset. The International Journal of Robotics Research, 32(11), 1231–1237 (2013)
5. Medina, R., Llamas, J., Zalama E. and Gómez-García-Bermejo, J.: Enhanced automatic detection of road surface cracks by combining 2D/3D image processing techniques, IEEE International Conference on Image Processing (ICIP), Paris, 2014, pp. 778-782 (2014)
6. Seung-Ki, R., Taehyeong, K. and Young-Ro, K.: Image-Based Pothole Detection System for ITS Service and Road Management System, Mathematical Problems in Engineering, (2015)
7. Mathavan, S., Kamal, K., and Rahman, M.: A Review of Three-Dimensional Imaging Technologies for Pavement Distress Detection and Measurements, IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 5, pp. 2353-2362 (2015)
8. Schnebele, E., Tanyu, B. F., Cervone, F., and Waters, G.: Review of remote sensing methodologies for pavement management and assessment, Eur. Transp. Res. Rev. (2015)
9. Koch, C., Giorgieva, K., Kasireddy, V., Akinci, B., and Fieguth, P. A review of computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure, Advanced Engineering Informatics, (2015)
10. Mohan, A., and Poobal, S. Crack detection using image processing: A critical review and analysis, Alexandria Engineering Journal, Volume 57, Issue 2, June 2018, Pages 787-798
11. Hoang, N.D., and Nguyen, Q.L.: A novel method for asphalt pavement crack classification based on image processing and machine learning, Engineering with Computers, April 2018, Volume 35, Issue 2, pp 487–498 (2018)
12. Hoang, N.D.: An Artificial Intelligence Method for Asphalt Pavement Pothole Detection Using Least Squares Support Vector Machine and Neural Network with Steerable FilterBased Feature Extraction, Hindawi Advances in Civil Engineering, (2018)
13. Pouyanfar, S., Sadiq, S., Yan, Y., Tian, H., Tao, Y., Presa-Reyes, M., Shyu, M.L, Chen, S.C, and Iyengar, S.S.: 2018. A Survey on Deep Learning: Algorithms, Techniques, and Applications. ACM Comput. Surv. 51, 5, Article 92, 36 pages, (2018).
14. Zhang, L., Yang, F., Zhang, Y.D., and Zhu, Y.J., "Road crack detection using deep convolutional neural network," 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, 2016, pp. 3708-3712. (2016)
15. Cha, Y.J., Choi, W., and Büyüköztürk, O.: Deep Learning-Based Crack Damage Detection Using Convolutional Neural Network, Computer-Aided Civil and Infrastructure Engineering, Volume 32, Issue 5, May 2017, pp. 361-378 (2017)
16. Tedeschi, A., and Benedetto, F.: A real time pavement crack and pothole recognition system for mobile Android-based devices, Advanced Engineering Informatics, Volume 32, April 2017, Pages 11-25 (2017)
17. Siriborvornratanakul, T.: An Automatic Road Distress Visual Inspection System Using an Onboard In-Car Camera, Hindawi Advances in Multimedia Volume 2018, Article ID 2561953, 10 pages
18. Liu, L., Ouyang, W., Wang, X., Fieguth, P., Liu, X., and Pietikäinen, M.: Deep Learning for Generic Object Detection: A Survey, arXiv:1809.02165v2, 2016
19. Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., Fischer,I. , Wojna, Z., Song, Y., Guadarrama, S., and Murphy, K.: Speed/accuracy trade-offs for modern convolutional object detectors, arXiv:1611.10012v3, 2018
20. Ale, L., Zhang, N., Li, L. : Road Damage Detection Using RetinaNet. International Conference on Big Data 2018: 5197-5200
21. Pereira, V., Tamura, S., Hayamizu, S., and Fukain H.: A Deep Learning-Based Approach for Road Pothole Detection in Timor Leste, 2018 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI), Singapore, 2018, pp. 279-284.
22. Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., and Omata, H: Road Damage Detection Using Deep Neural Networks with Images Captured Through a Smartphone, (2018)
23. RoadBotics, https://www.roadbotics.com/company/, Revised in June 2019-