

Сегментация дефектов дорожного покрытия на основе формирования синтетических выборок с помощью глубоких генеративно-сопоставительных сверточных сетей

И.А. Канаева¹, Ю.А. Иванова¹, В.Г. Спицын^{1,2}

¹ Национальный исследовательский Томский политехнический университет, 634050, Россия, г. Томск, пр. Ленина, д. 30;

² Национальный исследовательский Томский государственный университет, 634050, Россия, г. Томск, пр. Ленина, д. 36

Аннотация

В данной работе рассматривается комплекс задач обнаружения различных дефектов дорожного полотна автомобильных дорог и современные методы их решения. Представленное сравнение общедоступных наборов данных позволяет сделать вывод о сложности и малой разработанности задачи сегментации дефектов дорожного покрытия по изображению общего вида автомобильных дорог. Для решения данной проблемы разработаны алгоритмы генерации синтетического набора данных для сегментации дефектов классов трещин и выбоин на основе методов компьютерной графики и генеративно-сопоставительных сетей. Проведено сравнение точности сегментации дефектов дорожного покрытия полносверточной нейронной сетью U-Net на реальном и комбинированных наборах данных.

Ключевые слова: сегментация, дефекты дорожного покрытия, синтетическая выборка, генеративно-сопоставительные сети, сверточные нейронные сети.

Цитирование: Канаева, И.А. Сегментация дефектов дорожного покрытия на основе формирования синтетических выборок с помощью глубоких генеративно-сопоставительных сверточных сетей / И.А. Канаева, Ю.А. Иванова, В.Г. Спицын // Компьютерная оптика. – 2021. – Т. 45, № 6. – С. 907-916. – DOI: 10.18287/2412-6179-CO-844.

Citation: Kanaeva IA, Ivanova YuA, Spitsyn VG. Deep convolutional generative adversarial network-based synthesis of datasets for road pavement distress segmentation. Computer Optics 2021; 45(6): 907-916. DOI: 10.18287/2412-6179-CO-844.

Введение

С декабря 2018 г. в Российской Федерации принят и активно развивается национальный проект «Безопасные и качественные автомобильные дороги», целью которого является приведение автомобильных дорог регионального значения и дорожной сети городских агломераций в нормативное состояние. Осуществление данного проекта привело к увеличению доли контрактов по комплексной диагностике автомобильных дорог и оценке их технического состояния. Проектные организации в результате проведенной диагностики обязаны предоставлять заказчику следующую информацию: видеоданные обследуемого участка, ведомость дефектов дорожного полотна, отчет о комплексной диагностике с оценкой транспортно-эксплуатационного состояния.

Множество исследований за последнее десятилетие посвящено применению методов искусственного интеллекта к проблеме анализа изображений дорожной сцены. Ввиду наличия большого фонда качественных видеоданных дорожной отрасли и бурного развития области компьютерного зрения актуальным и перспективным направлением исследований является детектирование дефектов дорожного полотна по изображениям автомобильных дорог.

1. Классификация существующего опыта

Существующие методы мониторинга состояния дорожного покрытия подразделяют на три типа в зависимости от регистрирующего оборудования: лазерное 3D-сканирование, вибрационная диагностика и съемка 2D-видео [1].

Наиболее распространенным способом выделения дефектов является подход на основе анализа изображений автомобильных дорог ввиду невысоких затрат по получению таких данных и широкому спектру различных технологий компьютерного зрения. Анализ существующих работ позволяет классифицировать накопленный опыт в данной области по четырем направлениям:

1. Характер обрабатываемого изображения:
 - вид сверху;
 - вид водителя.
2. Тип решаемой задачи:
 - классификация;
 - детектирование;
 - сегментация.
3. Класс выделяемых повреждений покрытия:
 - трещины (одиночные, сетка трещин);
 - разрушения (ямы, выбоины, проломы);
 - износ дорожной разметки;

- выпотевание вяжущего.
4. Степень глубины применяемой модели:
- обработка изображений;
 - машинное обучение;
 - глубокое обучение.

1.1. Характер обрабатываемого изображения

Исторически наиболее развито направление обработки изображений, полученных непосредственно сверху над поверхностью автомобильной дороги. Однако сам процесс получения таких изображений сопряжен с трудностями размещения камер для полного охвата проезжей части, освещения поверхности и высокого разрешения картинки. Работы по созданию подобного набора данных проводят немецкие исследователи с помощью дорогостоящего мобильного картографического комплекса S.T.I.E.R. [2, 3].

В 2018 г. японские ученые в работе [4] представили коллекцию «Road Damage Dataset» из 9053 изображений размером 600×600 пикселей, полученных с помощью смартфона, закрепленного внутри автомобиля на приборной панели и направленного на съемку общего вида, соответствующего водительскому.

1.2. Тип решаемой задачи

В соответствии с анализируемыми данными базовой является задача бинарной классификации – содержит ли изображение дефект или нет. Ее успешное решение на изображениях RGB размером 99×99 пикселей с помощью сверточной нейронной сети позволило Zhang и соавторам [5] перейти к сегментированию трещин на изображениях большего размера методом скользящего окна и построения карты вероятности. В работе [6] описанная идея была успешно применена для выделения выбоин по виду сверху, а в [7] – для выделения трещин бетона.

Модификацией подхода также является обучение классификатора по векторам признаков с участков изображений, полученных предварительным разбиением на регионы-кандидаты [8, 9]. В [10] регионы интереса выделяют как связанные области после применения фильтра границ Канни к области дороги, а затем подают их на сверточную сеть – классификатор, предварительно масштабируя до размеров его входа.

Просмотр последовательно перебираемых в скользящем режиме окон является затратным в вычислительном отношении. Разработка таких архитектур, как полностью сверточные сети FCN, SegNet и U-Net, для задач попиксельной сегментации позволила успешно сегментировать трещины по виду сверху [11–16]. В одной из последних работ по данной тематике [17] предлагается метод под названием ConnCrack, использующий генератор из обученной условной порождающей состязательной сети с расстоянием Вассерштейна (сWGAN) для сегментации трещин. На вход генератору авторы подают изображение трещины, а на выходе получают 8 сегментированных масок ис-

ходного изображения по 8 соседним пикселям. Сеть дискриминатора находит разницу между полученными и эталонными картами связанности.

Развитие глубоких архитектур сверточных нейронных сетей, таких как Faster R-CNN, YOLO, SSD, MobileNet, Inception, позволило сократить вычислительные и временные затраты при определении местоположения дефекта на изображении и его классификации. Ряд работ посвящен применению данного подхода к изображению вида водителя [4]. В недавней работе [18] американские исследователи представили идею совместного применения U-Net для сегментирования дефектов по виду сверху (камера перпендикулярна поверхности) и YOLO v2 для детектирования разрушений по перспективе того же участка (угол наклона камеры к дороге составляет -70°).

Появление архитектуры Mask R-CNN позволило перейти к задаче сегментации экземпляров дефектов по виду водителя [19, 20].

1.3. Класс выделяемых поврежденных покрытия

Подробный анализ современных технических средств и применяемых методов для обнаружения различных типов дефектов дорожного покрытия приведен в работах [21, 22]. Согласно обзору [23], наиболее исследована задача выделения дорожных трещин, а именно сегментация трещины по виду сверху. Задача же выделения таких дефектов, как выкрашивания и выпотевание вяжущего согласно [21] является наиболее сложной, т.к. требует высокого разрешения и хорошего освещения. В наборе данных «Road Damage Dataset» [4] японских исследователей впервые представлен такой тип дефектов, как износ дорожной разметки: износ разметки «пешеходный переход» и износ линий горизонтальной разметки.

1.4. Степень глубины применяемой модели

В работе [24] при проведении анализа существующих подходов к распознаванию трещин выделяют следующие три общие группы методов: обработка изображений – методы пороговой сегментации, методы выделения границ, метод разрастания регионов; машинное обучение – без учителя, с учителем; и глубокое обучение.

Алгоритмы первой группы чаще всего применяются к задаче выделения трещин по виду сверху. Успешное применение методов машинного обучения к данной задаче впоследствии привело к развитию идеи детектирования дефектов по общему виду водителя. Данная проблема является наиболее сложной и актуальной в настоящий момент, для ее решения прибегают к глубокому обучению современных архитектур сверточных нейронных сетей.

2. Анализ открытых наборов данных

В табл. 1 представлено сравнение наиболее популярных и открытых наборов данных для решения задач выделения дефектов дорожного полотна.

Табл. 1. Обзор существующих открытых наборов данных

Наименование	Решаемая задача	Классы дефектов	Количество, штук	Размер, пиксель	Характер изображения
CRACK500 [5]	сегментация	трещина	500	3264 × 2448	вид сверху
CRACKTree 200 [25]	сегментация	трещина	200	800 × 600	вид сверху
CrackForestDataset (CFD) [26]	сегментация	трещина	118	480 × 320	вид сверху
GAPs v2 [3]	классификация	неповрежденная дорога, трещина, заплатка, заплатка с пропиткой, выбоина, открытый шов покрытия	2468	1920 × 1080	вид сверху
FHWA/LTPP [8]	классификация	неповрежденная дорога, трещина	1056	2048 × 3072	вид сверху
Road Damage Dataset 2019 (RDD-2019) [27]	детектирование	трещины: продольные, продольный шов, поперечные, поперечный шов, множественное растрескивание; выбоины, износ разметки пешеходного перехода, износ разметки линий, люк	13135	600 × 600 и 720 × 720	вид водителя
Pavement Image Datasets (PID) [18]	классификация	8 типов трещин, выбоины, выкрашивания	7237	640 × 640	вид водителя и соответствующий вид сверху
CQU-BPDD [28]	детектирование	неповрежденная дорога, 5 типов трещин	60059	1200 × 900	вид сверху

Набор данных RDD-2019 японских ученых позволил исследователям из разных стран оценить применение методов глубокого обучения к задаче обнаружения дефектов покрытия. Основным недостатком набора данных является способ выделения – ограничивающая рамка. Ввиду различных топологий дефектов, в частности диагональных трещин, прямоугольная область выделения может захватывать множество лишней информации. Для целей оценки качества автомобильной дороги наилучшим вариантом является попиксельное выделение с помощью маски, что позволяет не только точно локализовать дефект, но и оценить его площадь. Создание описанного набора данных является трудозатратным и длительным ввиду кропотливого ручного выделения дефектов на изображении. Данная работа посвящена созданию методов генерации синтетических изображений для сегментации дефектов дорожного покрытия.

3. Создание наборов данных

Для создания набора данных для задач сегментации вручную были размечены 475 изображений: 320 изображений из набора RDD-2019 дорог Японии и Чехии и 155 изображений дорог России. Коллекция RDD-2019 включает в себя также дорожную обстановку Индии, которая ввиду значительных отличий по сравнению с российскими дорогами не рассматривалась. В работе [29] положительно оценивается применимость японской модели RDD-2019 для других стран только в случае использования модели как базовой и обязательного дополнения изображениями из самой страны применения. Авторы рекомендуют смешивать коллекции дефектов покрытия разных стран для увеличения количества данных, предот-

ращения переобучения нейросетевых моделей и увеличения их обобщаемости.

Произведенная попиксельная разметка включает в себя класс линейных трещин, сетку трещин («крокодиловая кожа») и выбоины (рис. 1).

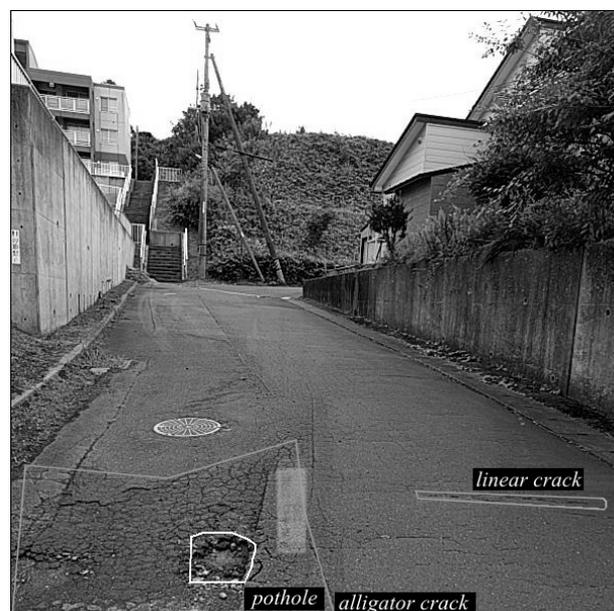


Рис. 1. Пример разметки дефектов на изображении из коллекции RDD-2019

Наиболее простым и часто используемым способом аугментации данных является применение случайных преобразований к исходным изображениям: обрезка, сдвиг, поворот, масштабирование, Гауссово размытие, изменение контраста и цветности.

Более затратным подходом к проблеме аугментации данных является искусственная генерация. Далее будут предложены два метода для генерации синтетических изображений.

тических изображений дефектов покрытия с сегментационной маской на основе метода аппликации существующих дефектов и полученных с помощью генеративно-сопоставительных сетей.

3.1. Компьютерная графика

Для генерации синтетических изображений вида водителя предлагается метод аппликации: добавления реальных изображений дефектов с сегментационной маской в область асфальта целевого изображения. Для этого используются 356 изображений трещин из наборов данных CFD, CRACKTree200, CRACK500 и 280 вручную аннотированных изображений выбоин из набора данных RDD-2019. В ранней реализации данного подхода [30] изображения трещин непосредственно накладывались на область дороги без учета перспективы ввиду их вытянутости. Однако аппликация объектов с единичным соотношением сторон, например, выбоин, без учета перспективных искажений выглядит неправдоподобно. На рис. 2 представлена схема аппликации с преобразованием дефектов от вида сверху к перспективе изображения из коллекции Cityscapes [31] по известным внутренним и внешним параметрам камеры.

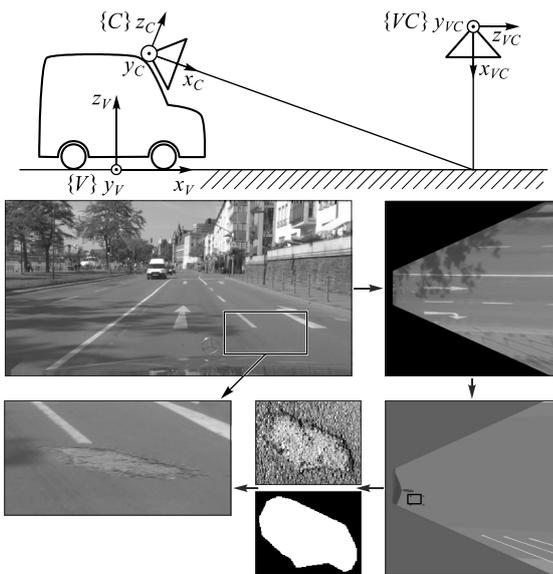


Рис. 2. Аппликация дефектов покрытия с учётом перспективы изображения, где системы координат: V – автомобиля, C – камеры, VC – виртуальной камеры для получения вида сверху

Представленный метод позволяет симитировать перспективные искажения дефектов в отличие от их непосредственного нанесения. При добавлении дефекта на изображение производится его цветокоррекция по методу перцептивного переноса цвета путем вычисления статистик по области асфальта вне маски дефекта. Бесшовное наложение дефекта осуществляется с помощью пуассоновского смешивания.

На рис. 3 представлено изображение, полученное с помощью описанного метода аппликации существующих дефектов.

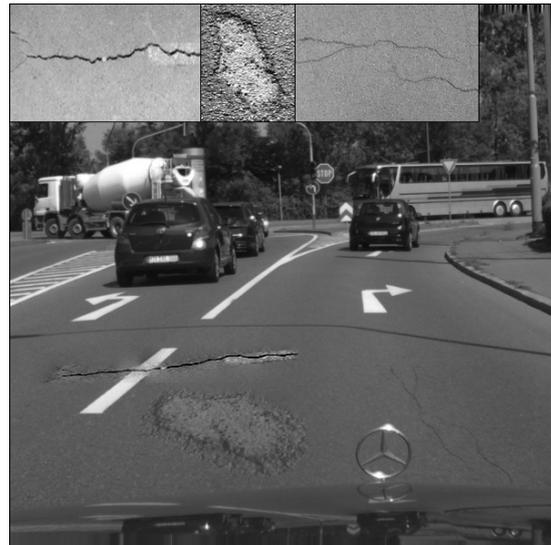


Рис. 3. Пример сгенерированного изображения с помощью компьютерной графики

3.2. Глубокая сверточная генеративно-сопоставительная сеть

Появление сопоставительных сетей (GAN) открыло новые возможности для аугментации данных. В работе [32] немецкие исследователи предложили использовать сеть CycleGAN – порождающую сопоставительную сеть для междоменного преобразования участка асфальта к участку с разрушением и наоборот. Применение прогрессивной генеративно-сопоставительной сети PG-GAN в статье [27] для увеличения числа экземпляров класса выбоин в наборе данных RDD-2019 дорог Японии позволило увеличить качество детектирования на 2–5%.

Приведенные примеры аугментации подходят только для задачи детектирования с помощью ограничивающей рамки. Для генерации синтетических изображений дефектов для задачи сегментации предлагается идея обучения глубокой сверточной генеративно-сопоставительной сети (DCGAN) на 4-канальных изображениях, где 3 канала соответствуют каналам RGB, а в 4 канал записывается бинарная попиксельная маска. На рис. 4 представлена схема работы сети.

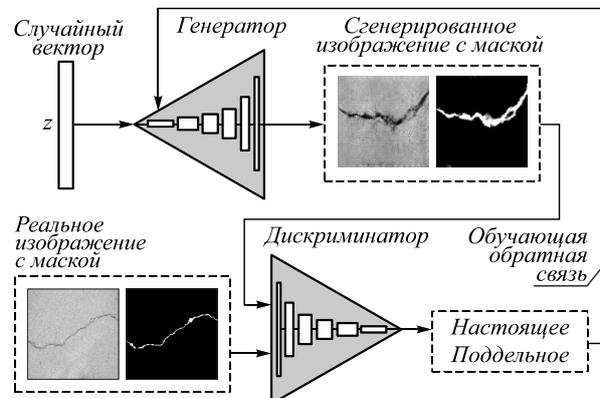


Рис. 4. Схема работы DCGAN для генерации изображения с сегментационной маской

Схожий подход применяли австрийские исследователи в [33] для создания синтетического набора рентгеновских снимков грудной клетки с сегментированным легким.

Конечная целевая функция состязательной сети (1) представляет собой минимаксную игру между дискриминатором и генератором и задается следующим уравнением [34]:

$$\min_G \max_D V(D, G) = E_{x \sim p(x)} \log D(x) + E_{z \sim q(z)} \log(1 - D(G(z))), \quad (1)$$

где $D(x)$ – модель дискриминатора, $G(z)$ – модель генератора, $p(x)$ – плотность реального распределения

данных, $q(z)$ – плотность распределения сгенерированных данных, E – ожидаемый выход.

В табл. 2, 3 представлены архитектуры сети генератора и дискриминатора соответственно.

Для класса трещин использовался набор данных CFD, а для класса выбоин – вручную сегментированные изображения из набора RDD-2019. Изображения масштабировались к размеру 128×128 пикселей. Обучение представленной сети DCGAN производилось на протяжении 3000 эпох с помощью пакетов по 32 изображениям методом Adam (*adaptive moment estimation* – адаптивная оценка момента).

Табл. 2. Архитектура сети генератора

Наименование слоя	Размерность входа	Размерность выхода	Размер ядра свертки	Шаг свертки	Пакетная нормализация	Функция активации	Прореживание
Полносвязный слой	[1 × 128]	[8 × 8 × 128]	-	-	нет	нет	нет
Слой деконволюции	[8 × 8 × 128]	[8 × 8 × 512]	5	1	да	ReLU	да
Слой деконволюции	[8 × 8 × 512]	[16 × 16 × 256]	5	2	да	ReLU	да
Слой деконволюции	[16 × 16 × 256]	[32 × 32 × 128]	5	2	да	ReLU	нет
Слой внимания	[32 × 32 × 128]	[32 × 32 × 128]	-	-	нет	-	нет
Слой деконволюции	[32 × 32 × 128]	[64 × 64 × 64]	5	2	да	ReLU	нет
Слой деконволюции	[64 × 64 × 64]	[128 × 128 × 32]	5	2	да	ReLU	нет
Полносвязный слой	[128 × 128 × 32]	[128 × 128 × 4]	-	-	нет	Tanh	нет

Табл. 3. Архитектура сети дискриминатора

Наименование слоя	Размерность входа	Размерность выхода	Размер ядра свертки	Шаг свертки	Пакетная нормализация	Функция активации	Прореживание
Сверточный слой	[128 × 128 × 4]	[128 × 128 × 64]	5	1	нет	LeakyReLU	да
Сверточный слой	[128 × 128 × 64]	[64 × 64 × 64]	5	2	да	LeakyReLU	нет
Сверточный слой	[64 × 64 × 64]	[32 × 32 × 128]	5	2	да	LeakyReLU	нет
Сверточный слой	[32 × 32 × 128]	[16 × 16 × 256]	5	2	да	LeakyReLU	нет
Плоский слой	[16 × 16 × 256]	[65536]	-	-	-	-	-
Полносвязный слой	[65536]	[1]	-	-	нет	Sigmoid	нет

Как показали исследования, оригинальная архитектура DCGAN создана для получения изображений размера 64×64 пикселей и ее адаптация для изображений разрешением 128×128 пикселей не является простой задачей. Увеличение количества сверточных слоев дискриминатора часто приводит к быстрому достижению режима коллапса ввиду его превосходства над генератором. Представленная в работе архитектура с прореживанием на первом слое позволяет достичь стабильности обучения, однако выдает сильные артефакты в фоновом асфальте, что заметно на сгенерированных изображениях трещин (рис. 5).

Согласно исследованию [35], применение слоев повышения дискретности вместо слоев деконволюции является простым способом решения проблемы неравномерного охвата пикселей в генераторе. Данный подход позволил уменьшить проявление артефактов для класса выбоин, но для трещин подобная замена оказалась неудачной: сглаживалась не только текстура асфальта, но и трещины принимали округлые, нереалистичные очертания.

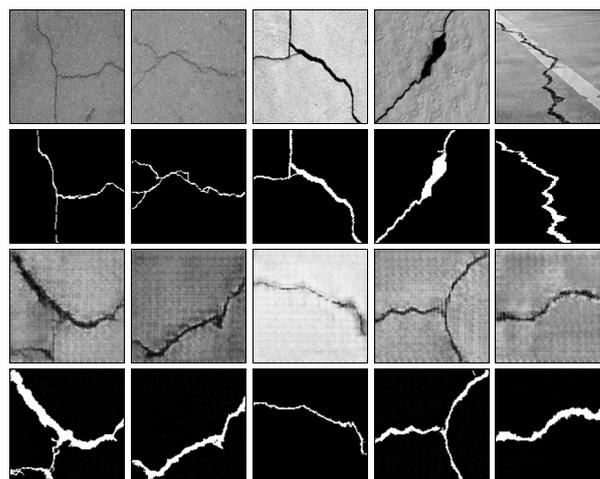


Рис. 5. Пример реальных пар с маской (1–2 строка) и сгенерированных (3–4 строка) для трещин

Ввиду использования сгенерированных изображений для приложения только в области маски самого дефекта полученные результаты генерирования тре-

щин с помощью DCGAN являются удовлетворительными для данной задачи. Примеры сгенерированных выбоин приведены на рис. 6.

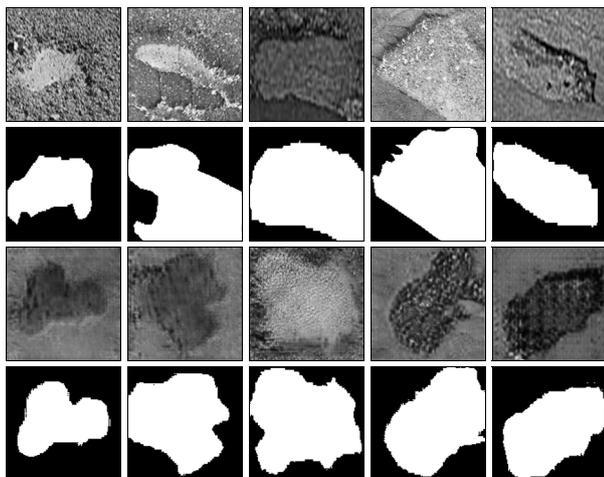


Рис. 6. Пример реальных пар с маской (1–2 строка) и сгенерированных (3–4 строка) для выбоин

В вопросе оценки качества обучения глубоких связательных сетей сверток наиболее эффективными признаны метрики с применением предварительно

обученной начальной сети [36]: начальная оценка (*Inception Score, IS*), оценка начального расстояния Фреше (*the Frechet Inception Distance, FID*), а также их модификации. Чтобы проанализировать распределение между реальными изображениями и сгенерированными с помощью DCGAN, были выбраны по 280 полутоновых изображений каждого класса и визуализированы на рис. 7 с помощью метода нелинейного снижения размерности t-SNE [37] со значением перплексии 30 и числом итераций 1000.

По рис. 7б распределение набора синтетических выбоин в значительной степени перекрывается с распределением реальных. Для трещин (рис. 7а) перекрытие распределений хуже ввиду большей площади асфальта на изображениях, в текстуре которого у сгенерированных изображений наблюдаются артефакты типа «шахматная доска», возникающие из-за применения разреженных слоев в генераторе. В целом, сгенерированные изображения заполняют распределение, охватываемое реальными изображениями.

Аппликация сгенерированных DCGAN дефектов производится способом, представленным на рис. 2, и показана на рис. 8.

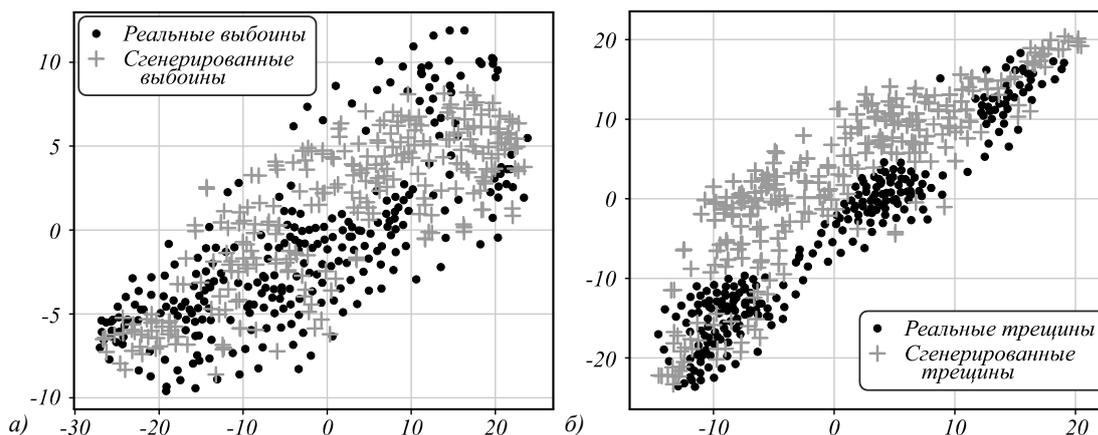


Рис. 7. Визуализация результатов генерации трещин (а) и выбоин (б) с помощью метода t-SNE



Рис. 8. Результат аппликации дефектов, сгенерированных с помощью DCGAN

4. Сегментация дефектов покрытия

В качестве модели для многоклассовой сегментации в работе использовалась популярная полносверточная нейронная сеть U-Net [38], первоначально созданная для сегментации биомедицинских изображений. Архитектура сети U-Net симметрично состоит из комбинации сверток понижающей дискретизации (*pooling*) в первой половине сети (кодирующая) и слоев сверток повышающей дискретизации (*upsampling*) во второй половине для восстановления размеров исходного изображения (декодирующая). Для сохранения информации о пространственном расположении признаков, утрачиваемой при пулинге, используются сквозные соединения между кодирующими и декодирующими уровнями сети, объединяющие карты признаков на соответствующих слоях. Выходной слой сети состоит из k карт признаков, включая фон и

количество классов, по которым категоризируется каждый пиксель.

Для ускорения обучения и улучшения точности модели часто применяют преобразованное обучение (*transfer learning*), заключающееся в адаптации существующей обученной нейронной сети в разрабатываемую архитектуру. С этой целью в качестве кодировщика U-Net использовались начальные слои предварительно обученной на наборе данных ImageNet нейронной сети ResNet50.

Ввиду несбалансированности классов и различия в размерах и топологии масок в качестве функции потерь сети L использовалась сумма (2) обобщенной функции потерь Дайса GDL [39] и фокальной функции потерь FL [40]:

$$GDL(p, y) = 1 - 2 \frac{w_i \sum_i y_i p_i}{w_i \sum_i (p_i + y_i)},$$

$$FL(p, y) = -0.25 \sum_i y_i (1 - p_i)^2 \ln p_i, \quad (2)$$

$$L(p, y) = GDL(p, y) + FL(p, y),$$

где $y_i = 1$, если пиксель принадлежит классу i , иначе $y_i = 0$, p_i – предсказанная вероятность принадлежности пикселя к классу i , а w_i задает вес i -го класса.

Для оценки качества результатов сегментации использовались средние значения по всем классам метрики пересечения по объединению IoU (3) и $F1$ -меры (4), представляющей собой гармоническое среднее между точностью p и полнотой r модели:

$$IoU = \frac{tp}{tp + fp + fn}, \quad (3)$$

$$F1 = 2 \frac{p \cdot r}{p + r}, p = \frac{tp}{tp + fp}, r = \frac{tp}{tp + fn}, \quad (4)$$

где tp – количество правильно классифицированных пикселей, fp – число неправильно классифицированных пикселей, fn – количество пикселей, неверно не отнесенных к классу.

Обучение сети U-Net производилось с помощью метода Adam в течение 30 эпох на изображениях четырех коллекций: реальные, реальные и сгенерированные с помощью компьютерной графики, реальные и синтетические с помощью DCGAN, комбинация всех изображений. В качестве порогового значения выходной сегментации для расчета метрик (3) и (4) использовалось значение 0,5. Тестовый набор данных включал 80 изображений дорог России, результаты сегментации которого приведены в табл. 4.

Результаты проведенных экспериментов по сегментации дефектов дорожного покрытия описанной полносверточной сетью U-Net+ResNet50 показывают целесообразность расширения исходного набора данных с использованием аппликации как реальных изображений дефектов, так и сгенерированных гене-

ративно-состязательной сетью. Для дальнейшего улучшения качества сегментации предлагается также увеличение числа экземпляров класса «крокодиловая кожа», так как он очень разнообразен, часто включает в себя ямы и линейные трещины. Также наиболее часто в качестве дефекта выбоин выделяются разрушения края проезжей части – бордюров и пешеходных дорожек. На рис. 9 представлены результаты сегментации сети, обученной на комбинированном наборе реальных и сгенерированных с помощью КГ и DCGAN изображений.

Табл. 4. Результаты сегментации дефектов

Набор данных	Количество изображений	Среднее IoU	Среднее F1
Реальные	395	0,557	0,598
Реальные + КГ	865	0,581	0,609
Реальные + DCGAN	865	0,619	0,633
Реальные + КГ + DCGAN	1335	0,623	0,647

Заключение

В данной работе проведен аналитический обзор и классификация задач в области обнаружения различных дефектов дорожного покрытия автомобильных дорог, а также рассмотрены современные методы их решения. Представленное сравнение общедоступных наборов данных позволяет сделать вывод о сложности и малой разработанности задачи сегментации дефектов дорожного покрытия по изображению общего вида автомобильной дороги, соответствующего водителю.

Для решения проблемы отсутствия обучающей попиксельно аннотированной коллекции дорожных дефектов были предложены алгоритмы генерации синтетического набора данных для дефектов классов трещин и выбоин на основе методов компьютерной графики и генеративно-состязательных сетей с помощью метода аппликации.

Проведенное сравнение точности сегментации дефектов дорожного покрытия с помощью обучения полносверточной сети U-Net+ResNet50 на реальном, реальном и сгенерированном с помощью компьютерной графики, реальном и синтезированном с помощью DCGAN наборах данных позволяет сделать вывод об улучшении качества сегментации от 3 до 7%, что подтверждает целесообразность применения разработанных методов аугментации.

Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-08-00977 А и в рамках Программы повышения конкурентоспособности ТПУ.

Литература

1. **Azimi, M.** Data-driven structural health monitoring and damage detection through deep learning: State-of-the-art re-

- view / M. Azimi, A. Eslamlou, G. Pekcan // Sensors. – 2020. – Vol. 20, Issue 10. – 2778. – DOI: 10.3390/s20102778.
2. **Eisenbach, M.** How to get pavement distress detection ready for deep learning? A systematic approach / M. Eisenbach, R. Stricker, D. Seichter, K. Amende, K. Debes, M. Sesselmann, D. Ebersbach, U. Stoeckert, H. Gross // 2017 International Joint Conference on Neural Networks (IJCNN). – 2017. – P. 2039-2047. – DOI: 10.1109/IJCNN.2017.7966101.
 3. **Stricker, R.** Improving visual road condition assessment by extensive experiments on the extended GAPs dataset / R. Stricker, M. Eisenbach, M. Sesselmann, K. Debes, H. Gross // 2019 International Joint Conference on Neural Networks (IJCNN). – 2019. – P. 1-8. – DOI: 10.1109/IJCNN.2019.8852257.

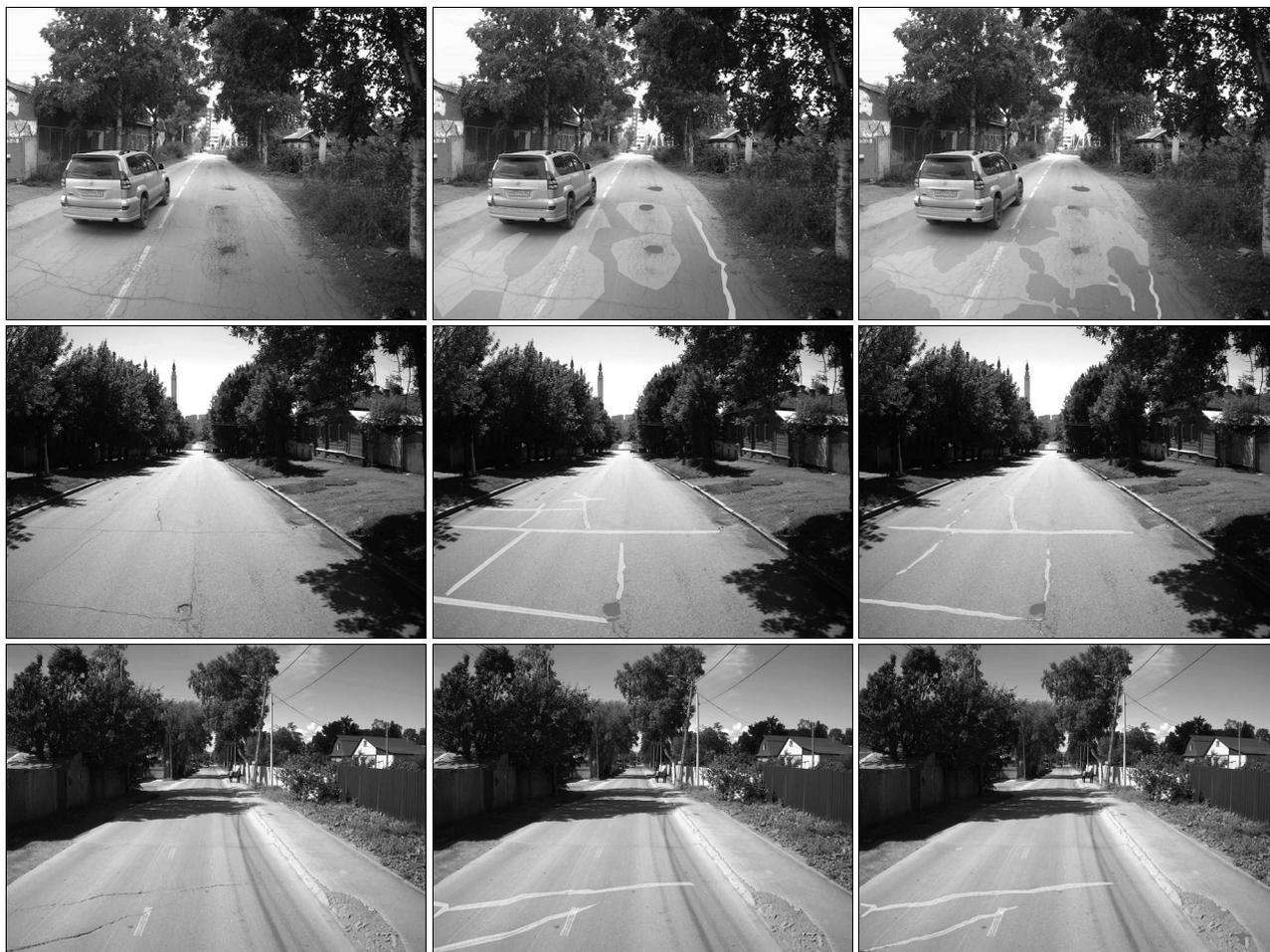


Рис. 9. Исходное изображение, разметка и результат сегментации

4. **Maeda, H.** Road damage detection and classification using deep neural networks with smartphone images / H. Maeda, Y. Sekimoto, T. Seto, T. Kashiya, H. Omata // Computer-Aided Civil and Infrastructure Engineering. – 2018. – Vol. 33, Issue 12. – P. 1127-1141. – DOI: 10.1111/mice.12387.
5. **Zhang, L.** Road crack detection using deep convolutional neural network / L. Zhang, F. Yang, Y.D. Zhang, Y.J. Zhu // 2016 IEEE International Conference on Image Processing (ICIP). – 2016. – P. 3708-3712. – DOI: 10.1109/ICIP.2016.7533052.
6. **Ye, W.** Convolutional neural network for pothole detection in asphalt pavement / W. Ye, W. Jiang, Z. Tong, D. Yuan, J. Xiao // Road Materials and Pavement Design. – 2021. – Vol. 22, Issue 1. – P. 42-58. – DOI: 10.1080/14680629.2019.1615533.
7. **Cha, Y.-J.** Deep learning-based crack damage detection using convolutional neural networks / Y.-J. Cha, W. Choi, O. Büyüköztürk // Computer-Aided Civil and Infrastructure Engineering. – 2017. – Vol. 32, Issue 5. – P. 361-378. – DOI: 10.1111/mice.12263.
8. **Gopalakrishnan, K.** Deep Convolutional Neural Networks with transfer learning for computer vision-based data-driven pavement distress detection / K. Gopalakrishnan, S. Khaitan, A. Choudhary, A. Agrawal // Construction and Building Materials. – 2017. – Vol. 157. – P. 322-330. – DOI: 10.1016/j.conbuildmat.2017.09.110.
9. **Varadharajan, S.** Vision for road inspection / S. Varadharajan, S. Jose, K. Sharma, L. Wander, C. Mertz // IEEE Winter Conference on Applications of Computer Vision. – 2014. – P. 115-122. – DOI: 10.1109/WACV.2014.6836111.
10. **Anand, S.** Crack-pot: autonomous road crack and pothole detection / S. Anand, S. Gupta, V. Darbari, S. Kohli // 2018 Digital Image Computing: Techniques and Applications (DICTA). – 2018. – P. 1-6. – DOI: 10.1109/DICTA.2018.8615819.
11. **Yang, F.** Feature pyramid and hierarchical boosting network for pavement crack detection / F. Yang, L. Zhang, S. Yu, D. Prokhorov, X. Mei, H. Ling // IEEE Transactions on Intelligent Transportation Systems. – 2020. – Vol. 21, Issue 4. – P. 1525-1535. – DOI: 10.1109/TITS.2019.2910595.

12. **Соболь, Б.В.** Модель глубокой сверточной нейронной сети в задаче сегментации трещин на изображениях асфальта / Б.В. Соболь, А.Н. Соловьев, П.В. Васильев, Л.А. Подколзина // Вестник Донского государственного технического университета. – 2019. – Т. 19, № 1. – С. 63-73. – DOI: 10.23947/1992-5980-2019-19-1-63-73.
13. **Zhang, K.** CrackGAN: Pavement crack detection using partially accurate ground truths based on generative adversarial learning / K. Zhang, Y. Zhang, H.-D. Cheng // IEEE Transactions on Intelligent Transportation Systems. – 2021. – Vol. 22, Issue 2. – P. 1306-1319. – DOI: 10.1109/TITS.2020.2990703.
14. **Liu Y.** DeepCrack: A deep hierarchical feature learning architecture for crack segmentation / Y. Liu, J. Yao, X. Lu, R. Xie, L. Li // Neurocomputing. – 2019. – Vol. 338. – P. 139-153. – DOI: 10.1016/j.neucom.2019.01.036.
15. **Dung, C.V.** Autonomous concrete crack detection using deep fully convolutional neural network / C.V. Dung, L.D. Anh // Automation in Construction. – 2019. – Vol. 99. – P. 52-58. – DOI: 10.1016/j.autcon.2018.11.028.
16. **Gao, Z.** Generative adversarial networks for road crack image segmentation / Z. Gao, B. Peng, T. Li, C. Gou // 2019 International Joint Conference on Neural Networks (IJCNN). – 2019. – P. 1-8. – DOI: 10.1109/IJCNN.2019.8851910.
17. **Mei, Q.** A cost effective solution for pavement crack inspection using cameras and deep neural networks / Q. Mei, M. Gül // Construction and Building Materials. – 2020. – Vol. 256. – P. 119397. – DOI: 10.1016/j.conbuildmat.2020.119397.
18. **Majidifard, H.** Pavement image datasets: A new benchmark dataset to classify and densify pavement distresses / H. Majidifard, P. Jin, Y. Adu-Gyamfi, W. Buttlar // Transportation Research Record: Journal of the Transportation Research Board. – 2020. – Vol. 2674, Issue 2. – P. 328-339. – DOI: 10.1177/0361198120907283.
19. **Канаева, И.А.** Методы коррекции цвета и яркости при создании панорамных изображений / И.А. Канаева, Ю.А. Болотова // Компьютерная оптика. – 2018. – Т. 42, № 5. – С. 885-897. – DOI: 10.18287/2412-6179-2018-42-5-885-897.
20. **Dhiman, A.** Pothole detection using computer vision and learning / A. Dhiman, R. Klette // IEEE Transactions on Intelligent Transportation Systems. – 2020. – Vol. 21, Issue 8. – P. 3536-3550. – DOI: 10.1109/TITS.2019.2931297.
21. **Coenen, T.B.J.** A review on automated pavement distress detection methods / T.B.J. Coenen, A. Golroo // Cogent Engineering. – 2017. – Vol. 4, Issue 1. – P. 1374822 (23 p.). – DOI: 10.1080/23311916.2017.1374822.
22. **Sattar, S.** Road surface monitoring using smartphone sensors: A review / S. Sattar, S. Li, M. Chapman // Sensors. – 2018. – Vol. 18, Issue 11. – P. 3845. – DOI: 10.3390/s18113845.
23. **Mohan, A.** Crack detection using image processing: A critical review and analysis / A. Mohan, S. Poobal // Alexandria Engineering Journal. – 2018. – Vol. 57, Issue 2. – P. 787-798. – DOI: 10.1016/j.aej.2017.01.020.
24. **Cao, W.** Review of pavement defect detection methods / W. Cao, Q. Liu, Z. He // IEEE Access. – 2020. – Vol. 8. – P. 14531-14544. – DOI: 10.1109/ACCESS.2020.2966881.
25. **Zou, Q.** CrackTree: Automatic crack detection from pavement images / Q. Zou, Y. Cao, Q. Li, Q. Mao, S. Wang // Pattern Recognition Letters. – 2012. – Vol. 33, Issue 3. – P. 227-238. – DOI: 10.1016/j.patrec.2011.11.004
26. **Shi, Y.** Automatic road crack detection using random structured forests / Y. Shi, L. Cui, Z. Qi, F. Meng, Z. Chen // IEEE Transactions on Intelligent Transportation Systems. – 2016. – Vol. 17, Issue 12. – P. 3434-3445. – DOI: 10.1109/TITS.2016.2552248.
27. **Maeda, H.** Generative adversarial network for road damage detection / H. Maeda, T. Kashiyama, Y. Sekimoto, T. Seto, H. Omata // Computer-Aided Civil and Infrastructure Engineering. – 2021. – Vol. 36. – P. 47-60. – DOI: 10.1111/mice.12561.
28. **Tang, W.** Iteratively optimized patch label inference network for automatic pavement disease detection / W. Tang, Q. Zhao, S. Huang, R. Li, L. Huangfu // arXiv Preprint. – 2021. – URL: <https://arxiv.org/abs/2005.13298> (request date 01.06.2021).
29. **Arya, D.** Transfer learning-based road damage detection for multiple countries / D. Arya, H. Maeda, S.K. Ghosh, D. Toshniwal, A. Mraz, T. Kashiyama, Y. Sekimoto // arXiv Preprint. – 2020. – URL: <https://arxiv.org/abs/2008.13101> (request date 01.06.2021).
30. **Канаева, И.А.** Сегментация изображений трещин дорожного покрытия / И.А. Канаева, Ю.А. Иванова // Труды конференции ГрафиКон-2019. – 2019. – Т. 1. – С. 131-135. – DOI: 10.30987/graphicon-2019-1-131-135.
31. **Cordts, M.** The cityscapes dataset for semantic urban scene understanding / M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, B. Schiele // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2016. – P. 3213-3223. – DOI: 10.1109/CVPR.2016.350.
32. **Kluger, F.** Region-based cycle-consistent data augmentation for object detection / F. Kluger, C. Reinders, K. Raetz, P. Schelske, B. Wandt, H. Ackermann, B. Rosenhahn // 2018 IEEE International Conference on Big Data (Big Data). – 2018. – P. 5205-5211. – DOI: 10.1109/BigData.2018.8622318.
33. **Neff, T.** Generative adversarial network based synthesis for supervised medical image segmentation / T. Neff, C. Payer, D. Stern, M. Urschler // Proceedings of the OAGM&ARW Joint Workshop 2017: Vision, Automation and Robotics. – 2017. – P. 140-145. – DOI: 10.3217/978-3-85125-524-9-30.
34. **Goodfellow, I.J.** GAN (Generative Adversarial Nets) / I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio // Journal of Japan Society for Fuzzy Theory and Intelligent Informatics. – 2017. – Vol. 29, Issue 5. – P. 177. – DOI: 10.3156/jsoft.29.5_177_2.
35. **Odena, A.** Deconvolution and checkerboard artifacts [Electronic Resource] / A. Odena, V. Dumoulin, C. Olah // Distill. – 2016. – Vol. 1, Issue 10. – DOI: 10.23915/distill.00003. – URL: <https://distill.pub/2016/deconv-checkerboard/> (request date 01.06.2021).
36. **Ахирвар, К.** Состязательные сети. Проекты / К. Ахирвар. – ДМК-Пресс, 2018. – 252 с. – ISBN: 978-5-97060-783-1.
37. **van der Maaten, L.** Visualizing data using t-SNE / L. van der Maaten, G. Hinton // Journal of Machine Learning Research. – 2008. – Vol. 9, Issue 86. – P. 2579-2605.
38. **Ronneberger, O.** U-Net: Convolutional networks for biomedical image segmentation / O. Ronneberger, P. Fischer, T. Brox. – In: Medical image computing and computer-assisted intervention – MICCAI 2015 / ed. by N. Navab, J. Hornegger, W.M. Wells, A.F. Frangi. – Cham: Springer, 2015. – P. 234-241. – DOI: 10.1007/978-3-319-24574-4_28.
39. **Carole, H.S.** Generalized Dice overlap as a deep learning loss function for highly unbalanced segmentations /

H.S. Carole, L. Wenqi, T. Vercauteren, S. Ourselin, M.J. Cardoso // Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. – 2017. – P. 240-248.

40. **Lin, T.** Focal loss for dense object detection / T. Lin, P. Goyal, R. Girshick, H. Kaiming, P. Dollar // IEEE International Conference on Computer Vision (ICCV). – 2017. – P. 2999-3007. – DOI: 10.1109/ICCV.2017.324.

Сведения об авторах

Канаева Ирина Александровна, 1994 года рождения, в 2018 окончила магистратуру Томского политехнического университета по направлению «Информатика и вычислительная техника», специализация «Компьютерный анализ и интерпретация данных», обучается в аспирантуре Томского политехнического университета и работает в дорожной организации ООО «СибДор». Область научных интересов: обработка и анализ изображений и видео, сверточные нейронные сети, оценка глубины. E-mail: iap@tpu.ru.

Иванова Юлия Александровна, 1986 года рождения, в 2009 году окончила Томский политехнический университет по специальности «Программное обеспечение вычислительной техники и автоматизированных систем», к.т.н., доцент отделения информационных технологий. Область научных интересов: обработка изображений, распознавание образов, биологически-подобные алгоритмы. E-mail: jbolotova@tpu.ru.

Спицын Владимир Григорьевич, 1948 года рождения, в 1970 году окончил Томский государственный университет по специальности «Радиофизика и электроника», профессор, д.т.н., профессор Национального исследовательского Томского политехнического университета и Национального исследовательского Томского государственного университета. Область научных интересов: нейронные сети, обработка изображений, распространение электромагнитных волн в случайно-неоднородных средах. E-mail: spvg@tpu.ru.

ГРНТИ: 28.23.15

Поступила в редакцию 5 декабря 2020 г. Окончательный вариант – 3 июня 2021 г.

Deep convolutional generative adversarial network-based synthesis of datasets for road pavement distress segmentation

I.A. Kanaeva¹, Yu.A. Ivanova¹, V.G. Spitsyn^{1,2}

¹National Research Tomsk Polytechnic University, 634050, Tomsk, Russia, Lenin Avenue, 30,

²National Research Tomsk State University, 634050, Tomsk, Russia, Lenin Avenue, 36

Abstract

We discuss a range of problems relating to road pavement defects detection and modern approaches to their solution. The presented comparison of publicly available datasets allows one to make a conclusion that the problem of segmentation of road pavement defects in driver wide-view road images is difficult and poorly investigated. To solve this problem, we have developed algorithms for generating a synthetic dataset for cracks and potholes distress based on computer graphics methods and deep convolutional generative adversarial networks. A comparison of the accuracy of road distress segmentation was performed by training a fully convolutional neural network U-Net on real and combined datasets.

Keywords: image segmentation, road pavement distress, synthetic dataset, generative adversarial network, convolutional neural network.

Citation: Kanaeva IA, Ivanova YuA, Spitsyn VG. Deep convolutional generative adversarial network-based synthesis of datasets for road pavement distress segmentation. *Computer Optics* 2021; 45(6): 907-916. DOI: 10.18287/2412-6179-CO-844.

Acknowledgements: The reported study was funded by RFBR according to the research project № 18-08-00977 A and in the framework of Tomsk Polytechnic University Competitiveness Enhancement Program.

References

- [1] Azimi M, Eslamlou A, Pekcan G. Data-driven structural health monitoring and damage detection through deep learning: State-of-the-art review. *Sensors* 2020; 20(10): 2778. DOI: 10.3390/s20102778.
 - [2] Eisenbach M, Stricker R, Seichter D, Amende K, Debes K, Sesselmann M, Ebersbach D, Stoeckert U, Gross H. How to get pavement distress detection ready for deep learning? A systematic approach. *IJCNN* 2017; 2039-2047. DOI: 10.1109/IJCNN.2017.7966101.
 - [3] Stricker R, Eisenbach M, Sesselmann M, Debes K, Gross H. Improving visual road condition assessment by extensive experiments on the extended GAPs dataset. *IJCNN* 2019; 1-8. DOI: 10.1109/IJCNN.2019.8852257.
 - [4] Maeda H, Sekimoto Y, Seto T, Kashiyama T, Omata H. Road damage detection and classification using deep neural networks with smartphone images. *Comput-Aided Civ Inf* 2018; 33(12): 1127-1141. DOI: 10.1111/mice.12387.
 - [5] Zhang L, Yang F, Zhang YD, Zhu YJ. Road crack detection using deep convolutional neural network. *ICIP* 2016; 3708-3712. DOI: 10.1109/ICIP.2016.7533052.
 - [6] Ye W, Jiang W, Tong Z, Yuan D, Xiao J. Convolutional neural network for pothole detection in asphalt pavement. *Road Mater Pavement Des* 2021; 22(1): 42-58. DOI: 10.1080/14680629.2019.1615533.
 - [7] Cha Y-J, Choi W, Büyüköztürk O. Deep learning-based crack damage detection using convolutional neural networks. *Comput-Aided Civ Inf* 2017; 32(5): 361-378. DOI: 10.1111/mice.12263.
 - [8] Gopalakrishnan K, Khaitan SK, Choudhary A, Agrawal A. Deep Convolutional Neural Networks with transfer learning for computer vision-based data-driven pavement distress detection. *Constr Build Mater* 2017; 157: 322-330. DOI: 10.1016/j.conbuildmat.2017.09.110.
 - [9] Varadharajan S, Jose S, Sharma K, Wander L, Mertz C. Vision for road inspection. *IEEE Winter Conf Appl Comput Vis* 2014; 115-122. DOI: 10.1109/WACV.2014.6836111.
 - [10] Anand S, Gupta S, Darbari V, Kohli S. Crack-pot: Autonomous road crack and pothole detection. 2018 DICTA 2018: 1-6. DOI: 10.1109/DICTA.2018.8615819.
 - [11] Yang F, Zhang L, Yu S, Prokhorov D, Mei X, Ling H. Feature pyramid and hierarchical boosting network for pavement crack detection. *IEEE Trans Intell Transp Syst* 2020; 21(4): 1525-1535. DOI: 10.1109/TITS.2019.2910595.
 - [12] Sobol BV, Soloviev AN, Vasiliev PV, Podkolzina LA. Deep convolution neural network model in problem of crack segmentation on asphalt images [In Russian]. *Vestnik Donskogo Gosudarstvennogo Universiteta* 2019; 19(1): 63-73. DOI: 10.23947/1992-5980-2019-19-1-63-73.
 - [13] Zhang K, Zhang Y, Cheng H-D. CrackGAN: Pavement Crack Detection Using Partially Accurate Ground Truths Based on Generative Adversarial Learning. *IEEE Trans Intell Transp Syst* 2021; 22(2): 1306-1319. DOI: 10.1109/TITS.2020.2990703.
 - [14] Liu Y, Yao J, Lu X, Xie R, Li L. DeepCrack: A deep hierarchical feature learning architecture for crack segmentation. *Neurocomputing* 2019; 338: 139-153. DOI: 10.1016/j.neucom.2019.01.036.
 - [15] Dung CV, Anh LD. Autonomous concrete crack detection using deep fully convolutional neural network. *Autom Constr* 2019; 99: 52-58. DOI: 10.1016/j.autcon.2018.11.028.
 - [16] Gao Z, Peng B, Li T, Gou C. Generative adversarial networks for road crack image segmentation. *IJCNN* 2019; 1-8. DOI: 10.1109/IJCNN.2019.8851910.
 - [17] Mei Q, Gül M. A cost effective solution for pavement crack inspection using cameras and deep neural networks. *Constr Build Mater* 2020; 256: 119397. DOI: 10.1016/j.conbuildmat.2020.119397.
 - [18] Majidifard H, Jin P, Adu-Gyamfi Y, Buttlar WG. Pavement image datasets: A new benchmark dataset to classi-
-

- fy and densify pavement distresses. *Transp Res Rec J Transp Res Board* 2020; 2674(2): 328-339. DOI: 10.1177/0361198120907283.
- [19] Kanaeva IA, Bolotova YA. Color and luminance corrections for panoramic image stitching. *Computer Optics* 2018; 42(5): 885-897. DOI: 10.18287/2412-6179-2018-42-5-885-897.
- [20] Dhiman A, Klette R. Pothole detection using computer vision and learning. *IEEE Trans Intell Transp Syst* 2020; 21(8): 3536-3550. DOI: 10.1109/TITS.2019.2931297.
- [21] Coenen TBJ, Golroo A. A review on automated pavement distress detection methods. *Cogent Eng* 2017; 4(1): 1374822. DOI: 10.1080/23311916.2017.1374822.
- [22] Sattar S, Li S, Chapman M. Road surface monitoring using smartphone sensors: A review. *Sensors* 2018; 18(11): 3845. DOI: 10.3390/s18113845.
- [23] Mohan A, Poobal S. Crack detection using image processing: A critical review and analysis. *Alex Eng J* 2018; 57(2): 787-798. DOI: 10.1016/j.aej.2017.01.020.
- [24] Cao W, Liu Q, He Z. Review of pavement defect detection methods. *IEEE Access* 2020; 8: 14531-14544. DOI: 10.1109/ACCESS.2020.2966881.
- [25] Zou Q, Cao Y, Li Q, Mao Q, Wang S. CrackTree: Automatic crack detection from pavement images. *Pattern Recognit Lett* 2012; 33(3): 227-238. DOI: 10.1016/j.patrec.2011.11.004.
- [26] Shi Y, Cui L, Qi Z, Meng F, Chen Z. Automatic road crack detection using random structured forests. *IEEE Trans Intell Transp Syst* 2016; 17(12): 3434-3445. DOI: 10.1109/TITS.2016.2552248.
- [27] Maeda H, Kashiyama T, Sekimoto Y, Seto T, Omata H. Generative adversarial network for road damage detection. *Comput-Aided Civ Inf* 2021; 36: 47-60. DOI: 10.1111/mice.12561.
- [28] Tang W, Zhao Q, Huang S, Li R, Huangfu L. Iteratively optimized patch label inference network for automatic pavement disease detection. *arXiv Preprint* 2021. Source: (<https://arxiv.org/abs/2005.13298>).
- [29] Arya D, Maeda H, Ghosh SK, Toshniwal D, Mraz A, Kashiyama T, Sekimoto Y. Transfer learning-based road damage detection for multiple countries. *arXiv Preprint* 2020. Source: (<https://arxiv.org/abs/2008.13101>).
- [30] Kanaeva I, Ivanova Y. Instance segmentation of road pavement cracks [In Russian]. *Proc Conf "Graphicon-2019"* 2019; 1: 131-135. DOI: 10.30987/graphicon-2019-1-131-135.
- [31] Cordts M, Omran M, Ramos S, Rehfeld T, Enzweiler M, Benenson R, Franke U, Roth S, Schiele B. The cityscapes dataset for semantic urban scene understanding. *CVPR* 2016: 3213-3223. DOI: 10.1109/CVPR.2016.350.
- [32] Kluger F, Reinders C, Raetz K, Schelske P, Wandt B, Ackermann H, Rosenhahn B. Region-based cycle-consistent data augmentation for object detection. *IEEE Int Conf on Big Data (Big Data)* 2018: 5205-5211. DOI: 10.1109/BigData.2018.8622318.
- [33] Neff T, Payer C, Štern D, Urschler M. Generative adversarial network based synthesis for supervised medical image segmentation. *Proc OAGM&ARW Jt Work* 2017 Vision, Autom Robot 2017: 140-145. DOI: 10.3217/978-3-85125-524-9-30.
- [34] Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. GAN (Generative Adversarial Nets). *J Japan Soc Fuzzy Theory Intell Informatics* 2017; 29: 177. DOI: 10.3156/jsoft.29.5_177_2.
- [35] Odena A, Dumoulin V, Olah C. Deconvolution and checkerboard artifacts. *Distill* 2016; 1(10). DOI: 10.23915/distill.00003. Source: (<https://distill.pub/2016/deconv-checkerboard/>).
- [36] Ahrwar K. *Generative adversarial networks projects: Build next-generation generative models using TensorFlow and Keras*. Birmingham: Packt Publishing Ltd; 2018. ISBN: 978-1-78913-667-8.
- [37] van der Maaten L, Hinton G. Visualizing data using t-SNE. *J Mach Learn Res* 2008; 9(86): 2579-2605.
- [38] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In Book: *Navab N, Hornegger J, Wells WM, Frangi AF, eds. Medical image computing and computer-assisted intervention – MICCAI 2015*. Cham: Springer; 2015: 234-241. DOI: 10.1007/978-3-319-24574-4_28.
- [39] Carole HS, Wenqi L, Vercauteren T, Ourselin S, Cardoso MJ. Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations. *Deep Learn Med Image Anal Multimodal Learn Clin Decis Support* 2017: 240-248.
- [40] Lin T-Y, Goyal P, Girshick R, He K, Dollar P. Focal loss for dense object detection. *ICCV* 2017: 2999-3007. DOI: 10.1109/ICCV.2017.324.

Authors' information

Irina Alexandrovna Kanaeva (b. 1994) graduated from Tomsk Polytechnic University in 2018, majoring in Computer Analysis and Data Interpretation. Currently she is postgraduate student of Tomsk Polytechnic University and works at auto-road project company “SibDor”. Research interests include image and video processing and analyzing, convolution neural network, depth estimation. E-mail: iap15@tpu.ru.

Yulia Alexandrovna Ivanova (b. 1986) graduated from Tomsk Polytechnic University in 2009, PhD associated professor at Information Technology, Tomsk Polytechnic University. Her research interests are image processing, object recognition, biologically-inspired models. E-mail: jbolotova@tpu.ru.

Vladimir Grigorievich Spitsyn (b. 1948) graduated from Tomsk State University in 1970, Radio-Physics department. He works as the Professor of Tomsk Polytechnic University and Tomsk State University. His research interests are currently focused on neural networks, image processing, and electromagnetic wave propagation in random discrete media. E-mail: spvg@tpu.ru.