

# СЕГМЕНТАЦИЯ ВЫБОИН С ПОМОЩЬЮ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ YOLOV8

Канаева И.А.<sup>1</sup>, Спицын В.Г.<sup>2</sup>

<sup>1</sup> НИ ТПУ, ИШИТР, ОИТ, зр. А0-39, e-mail: iap15@tpu.ru

<sup>2</sup> НИ ТПУ, ИШИТР, ОИТ, профессор, д.т.н., e-mail: spvg@tpu.ru

## Аннотация

В данной работе рассматривается задача сегментации выбоин на автомобильных дорогах с помощью сверточных нейронных сетей семейства YOLOv8. Актуальность исследования подчеркнута статистическим анализом дорожно-транспортных происшествий в России по вине неудовлетворительного состояния дорог и улиц. Проведен обзор общедоступных наборов данных для задач сегментации. Полученный результат лучшей модели YOLOv8s-seg на тестовой выборке бенчмарка UDTIRI сопоставим с SOTA результатами моделей в работе [1].

**Ключевые слова:** сегментация, дефекты дорожного покрытия, сверточные нейронные сети.

## Введение

На начало 2024 года протяженность дорог в России составляет более 1,6 миллионов километров. Высокая интенсивность и сложные погодные условия часто приводят к деформации и частичному разрушению асфальтобетонных покрытий. Для оценки своевременного мониторинга состояния дорог и оценки качества дорожного покрытия используются передвижные комплексные дорожные лаборатории (КДЛ), оснащенные камерами, giro-инерциальной системой позиционирования, системой глобального позиционирования, лазерным трехмерным сканером и другими датчиками, а также гироскопами. В настоящий момент основная обработка данных, полученных в ходе диагностики, осуществляется инженерами в ручном режиме. Трудоемкость процесса выделения дефектов, а также бурное развитие современных методов компьютерного зрения позволяет судить об актуальности автоматизации процесса обнаружения дефектов. В данной работе будет уделено внимание такому дефекту как выбоины, которые представляют собой углубление на поверхности дороги, возникшее в результате ударов.

## Статистический анализ ДТП

По статистике госавтоинспекции, в дорожно-транспортных происшествиях (ДТП) в Российской Федерации за 2022 год погибло свыше 14 тыс.человек и пострадало более 160 тыс.чел. Более 75 % всех ДТП в РФ происходят в городах, при этом основной или сопутствующей причиной каждого третьего происшествия является неудовлетворительное состояние дорог и улиц [2]. В результате наличия данного отягчающего фактора только в 2022 году погибло свыше 4 тыс.человек (рис. 1).

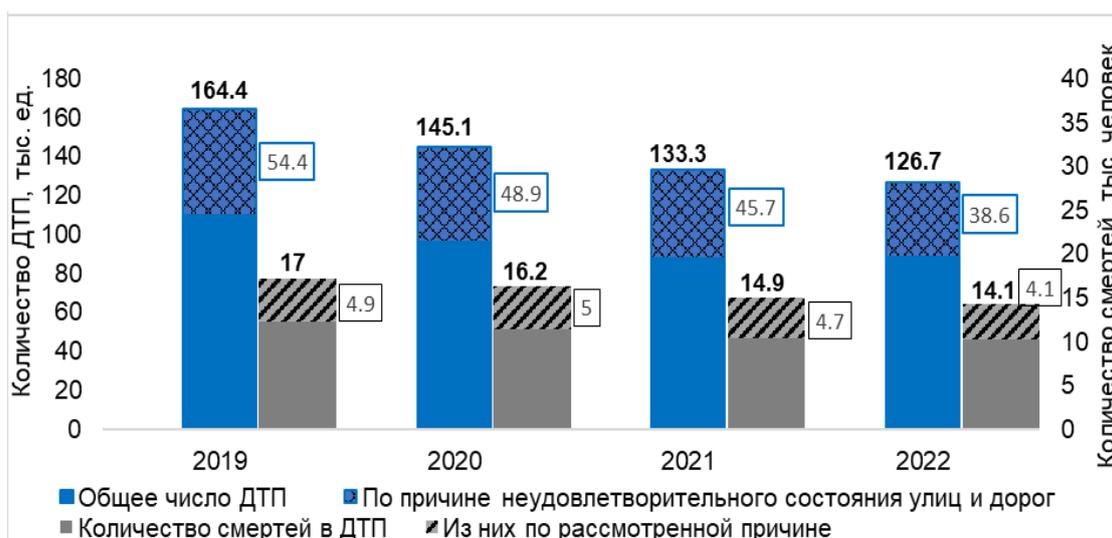


Рис. 1. Статистика ДТП по причине неудовлетворительного состояния улиц и дорог

По параметру смертности в ДТП на 100 тыс.чел населения на 2019 год Россия занимала 65-е место

с показателем в 12 смертей на 100 тысяч человек. При этом данный показатель для стран Евросоюза на 2019 год был в среднем более чем вдвое ниже (5,1), но безусловно показатели в странах западной Европы значительно отличаются от восточной части.

С 2019 года в РФ реализуется национальный проект «Безопасные качественные дороги», целью которого является повышение качества дорог, улучшение доступности инфраструктуры, а также качества самих дорог. Реализация данного проекта позволила значительно сократить общую смертность на дорогах: -17 % в 2022 году относительно 2019 года, а показатель смертности на 100 тысяч населения составил 9,8. Однако не стоит забывать, что на 2022 год 38,6 тысяч ДТП произошло по причине неудовлетворительного состояния улиц и дорог, в которых погибло более 4 тысяч человек и большего количества этих смертей можно было бы избежать при своевременной диагностике и выявления несоответствия нормативным параметрам.

### Обзор существующих открытых наборов данных для задач сегментации

В области обнаружения дефектов дорожного покрытия более распространен подход к детектированию с помощью ограничивающих рамок. Систематизацию данных и знаний в этом направлении провели японские исследователи в ходе серии соревнований по обнаружению дефектов, проводимых с 2018 по 2022 год. Итоговый набор данных RDD2022 содержит 47420 изображений дорог из шести стран: Японии, Индии, Чехии, Норвегии, США и Китая. На изображениях отмечено более 55000 дефектов четырех типов: продольная трещина, поперечная трещина, множественное растрескивание и выбоина [3]. Основные результаты показали успешное применение методов глубокого обучения к задаче обнаружения дефектов покрытия, а использование данных из разных стран повысило надежность алгоритмов.

Ввиду различных топологий дефектов, в частности диагональных трещин, прямоугольная область выделения может захватывать множество лишней информации, что вносит трудности в определение площади поврежденного покрытия. Для целей оценки качества автомобильной дороги наилучшим вариантом является попиксельное выделение дефектов с помощью маски, однако создание описанного набора данных является трудозатратным и длительным ввиду кропотливого ручного выделения дефектов на изображении. В таблице 1 представлены существующие открытые наборы данных для решения задач сегментации разрушений.

Таблица 1

*Свободно-доступные наборы данных для сегментации дефектов дорожного покрытия*

Наименование	Место сбора	Количество изображений, штук	Размер изображений, пиксель	Тип перспективы изображений	Классы дефектов
CRACK500 [4]	США, Пенсильвания	500 (250-50-200)	2000x1500	вид сверху	трещина
CRACKTree 200 [5]	Не известно	206	800x600	вид сверху	трещина
CrackForestDataset (CFD) [6]	Китай, Пекин	118	480x320	вид сверху	трещина
EdmCrack600 [7]	Канада, Эдмонт	600	1920x1080	вид камеры заднего вида	трещина
DNIT [8]	Бразилия	2235	1280x729	вид водителя	трещина, выбоина, полоса движения
Farzad Road Damage Assessment [9]	интернет	780 (720-60)	640x640	различный	выбоина
UDTIRI [1]	интернет	1000 (600-100-300)	различный	различный	выбоина

В конце 2023 года китайские ученые представили первый бенчмарк в области интеллектуальной дорожной инспекции – обнаружение выбоин на дорогах, называемый UDTIRI. Бенчмарк предоставляет собой хорошо аннотированный набор данных из 1000 изображений выбоин, который можно

использовать для обучения и оценки сетей обнаружения объектов, семантической сегментации и сегментации экземпляров. Аннотации к обучающим и проверочным выборкам доступны исследователям, а оценку производительности разработанной модели на тестовом наборе можно получить, отправив результаты через онлайн-платформу тестирования.

### Описание архитектуры YOLOv8-seg

Модель YOLO является одноступенчатой архитектурой, которая обрабатывает изображение в исходном масштабе за один проход, рассматривая задачу детектирования как задачу регрессии ограничивающих рамок и классов объектов. Архитектура YOLO подразделяется на СНС для вычисления признаков изображения, называемую *backbone*, подсеть для обработки признаков – *neck*, и *head* – объединение частей, отвечающих за предсказание ограничивающей рамки.

Модель YOLOv8-seg является расширением модели YOLOv8 для задач семантической сегментации. В качестве подсети выделения признаков (*backbone*) модель YOLOv8-seg использует подсеть CSPDarknet53 с разделением потока градиента от пирамиды признаков *FPN*. За тем следует новый модуль *C2f* вместо традиционной архитектуры YOLO *neck*. За модулем *C2f* следуют два набора *head* слоев сегментации, которые учатся предсказывать маски семантической сегментации для входного изображения. Модель также имеет *head*-слои обнаружения, аналогичные YOLOv8, состоящие из пяти модулей обнаружения и слоя прогнозирования. Схематичное представление описанной архитектуры приведено на рис. 2.

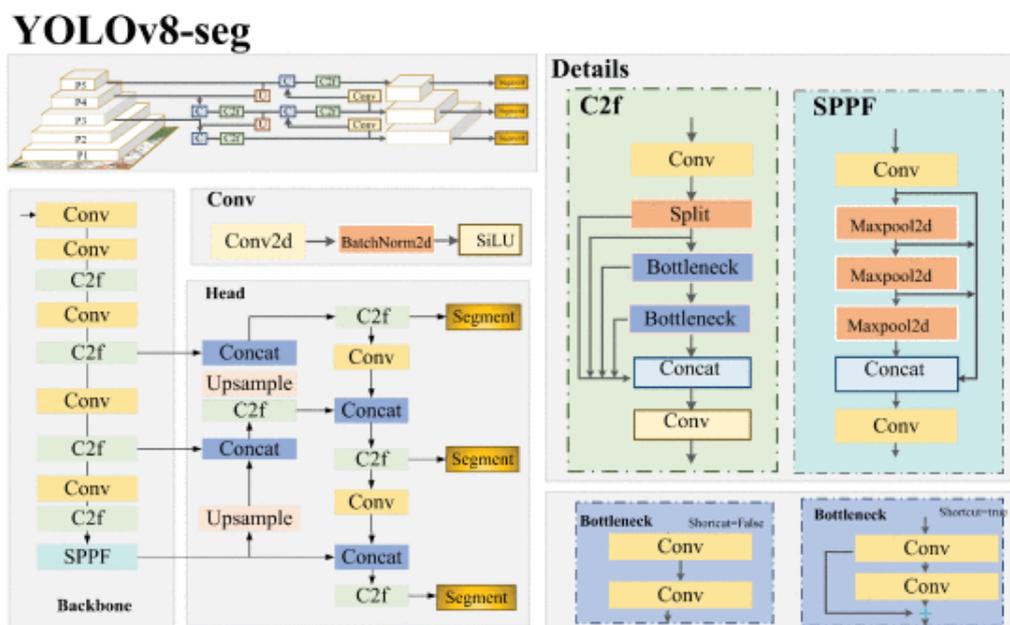


Рис. 2. Архитектура YOLOv8-seg [10]

В [11] показано, что модель YOLOv8-seg позволяет достигать самых современных результатов в различных тестах обнаружения объектов и семантической сегментации при сохранении высокой скорости и эффективности. Авторы фреймворка представили 5 конфигураций модели YOLOv8, по-разному оптимизированных для улучшения скорости и точности, и предварительно обученных на наборе данных COCO [11].

### Формирование датасета и обучение нейросети

Для обучения нейросетевой модели был сформирован датасет на основе наборов данных Farzad и UDTIRI в размере 1320 обучающих и 160 проверочных изображений. Все изображения были масштабированы до размера входа сети 640x640 пикселей. Для ускорения обучения и улучшения точности модели часто применяют преобразованное обучение (transfer learning), заключающееся в адаптации существующей обученной нейронной сети в разрабатываемую архитектуру. Для повышения производительности модели во время обучения YOLOv8 использует мозаичную аугментацию: объединяет четыре изображения в одно для изучения контекстов объектов в разных

местах и на разном фоне. Фреймворк Ultralytics содержит 5 предварительно обученных на наборе данных COCO конфигураций сети YOLOv8-seg: от «крошечной» и быстрой модели «n» с 3,2 млн параметров, до большой и точной «x» с 71,7 млн. Обучение нейросетей производилось в течение 100 эпох с помощью метода AdamW, с начальной скоростью обучения 0,0001 и конечной 0,01, размер пакета batch равен 16. Оценка результатов на тестовой выборке из 300 изображений производилась через загрузку полученных сегментационных масок в онлайн-систему бенчмарка [12] и представлена в таблице 2.

Таблица 2

Точность сегментации моделей YOLOv8 на тестовой выборке бенчмарка UDTIRI

Наименование модели	IoU (%)	F1 (%)	Precision (%)	Recalls (%)	Accuracy (%)
YOLOv8n-seg	0,693	0,790	0,807	0,831	0,957
YOLOv8s-seg	0,697	0,792	0,801	0,834	0,960
YOLOv8m-seg	0,707	0,796	0,828	0,814	0,960
YOLOv8l-seg	0,672	0,767	0,770	0,842	0,950
YOLOv8x-seg	0,650	0,748	0,786	0,789	0,952

Модели «n», «s» и «m» на тестовой выборке по метрикам точности показывают схожие результаты, в то время как большие модели «l» и «x» не дают прироста в точности, ввиду избыточности слоев и весов сетей для обучающего датасета с небольшим числом изображений и объектов.

В таблице 3 представлены сведения о скорости вычисления рассматриваемых моделей СНС путём оценки среднего времени сегментации объектов на одном изображении проверочной выборки.

Таблица 3

Среднее время сегментации объектов на одном изображении

Наименование модели	Preprocess, мс	Inference, мс	Postprocess, мс	Количество параметров, млн	GFLOPs
YOLOv8n-seg	0,5	9,0	4,1	3,2	12,0
YOLOv8s-seg	3,8	11,0	3,9	11,8	42,4
YOLOv8m-seg	2,9	38,9	3,7	27,2	110,0
YOLOv8l-seg	13,8	46,1	5,0	45,9	220,1
YOLOv8x-seg	5,3	49,0	3,9	71,7	343,7

Путем анализа результатов точности и скорости сегментации, для решения поставленной задачи лучшим выбором будет архитектура YOLOv8s-seg. Корректные и ошибочные результаты выделения выбоин данной модели на тестовых изображениях представлены на рис. 3 и 4, соответственно. Выбраны изображения с различной перспективой съёмки, размером и количеством выбоин, времени суток, присутствием погодных явлений (снег, лужи, слякоть).



Рис. 3. Примеры верных результатов сегментации YOLOv8s-seg на тестовых изображениях UDTIRI



Рис. 4. Примеры недостоверных результатов сегментации YOLOv8s-seg на тестовых изображениях UDTIRI

Ввиду обучения модели на наборе данных с разными типами перспектив к дорожному покрытию (вид водителя, вид сверху и так далее), случаются ложные срабатывания в областях, не принадлежащих дороге, что показано на последнем изображении. Также наблюдаются трудности в сегментации выбоин с нетипичным мусором. Так как в обучающем наборе данных присутствуют экземпляры выбоин с лужами и гравием, то в целом модель неплохо справляется с погодными проявлениями на дорожном покрытии.

### Заключение

В работе проведен анализ актуальности темы автоматизации процесса дефектовки дорожного покрытия, выполнен обзор существующих открытых наборов данных для сегментации дефектов, рассмотрена современная архитектура СНС YOLOv8. Обучены и исследованы 5 конфигураций сети YOLOv8-seg. Представленные результаты позволяют сделать вывод об успешности обучения нейросети. Планируется расширение набора данных для минимизации указанных выше недостатков в работе СНС.

### Список использованных источников

1. Guo S. и др. UDTIRI: An Open-Source Road Pothole Detection Benchmark Suite. – 2023. – № 1. – С. 1–12.
2. Дорожно-транспортная аварийность в Российской Федерации за 2022 год: информационно-аналитический обзор. М. : ФКУ «НЦ БДД МВД России». – 2023. – 150 с.
3. Arya D., Maeda H., Ghosh S.K., Toshniwal D., Omata H., Kashiyama T., Sekimoto Y. Crowdsensing-based road damage detection challenge (CRDDC'2022) // Proceedings – 2022 IEEE International Conference on Big Data, Big Data. – 2022. – P. 6378-6386.
4. 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.
5. 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.
6. 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.
7. Mei Q., Gul M. A cost effective solution for pavement crack inspection using cameras and deep neural networks // Constr. Build. Mater. – 2020. – Т. 256. – С. 119397.
8. Passos B. T. и др. Cracks and Potholes in Road Images // Mendeley Data. – V4. – 2020. – URL: <https://biankatpas.github.io/Cracks-and-Potholes-in-Road-Images-Dataset/> (дата обращения: 20.03.2024)
9. Farzad. Pothole\_Segmentation\_YOLOv8 Dataset [Computer software] // Roboflow Universe. – 2023. – URL: [https://universe.roboflow.com/farzad/pothole\\_segmentation\\_yolov8](https://universe.roboflow.com/farzad/pothole_segmentation_yolov8) (дата доступа: 25.03.2024 г.)
10. Bai R., Wang M., Zhang Z., Lu J., Shen F. Automated construction site monitoring based on improved yolov8-seg instance segmentation algorithm // IEEE Access. – 2023. – Vol. 11 – P. 139082-139096.
11. Jocher G., Chaurasia A., Qiu J. Ultralytics YOLO (Version 8.0.0) [Computer software]. – 2023. – URL: <https://github.com/ultralytics/ultralytics>
12. Urban Digital Twins for Intelligent Road Inspection: офиц. сайт. – URL: <https://udtiri.com/> (дата обращения: 25.03.2024)