

УДК 004.8(55)

РАЗРАБОТКА МОДЕЛИ ДЕТЕКЦИИ МОРСКОГО МУСОРА

Достовалов Андрей Андреевич, магистрант, Базовая кафедра «Аналитика больших данных и методы видеоанализа», Уральский федеральный университет

Капустин Савелий Андреевич, магистрант, Базовая кафедра «Аналитика больших данных и методы видеоанализа», Уральский федеральный университет

Аннотация

Работа посвящена разработке модели машинного обучения для автоматической детекции морского мусора на основе анализа изображений. Проведен сравнительный анализ моделей глубокого обучения (Faster R-CNN, SSD, YOLO) с учетом точности, скорости обработки и потребления вычислительных ресурсов. Наиболее перспективной признана модель YOLOv8 благодаря высокой производительности и стабильности. Разработан и подготовлен уникальный датасет, включающий изображения морской поверхности с размеченными объектами мусора. Выполнено обучение модели, достигнуты значительные показатели точности (mAP50-95), полноты и локализации объектов. Создан веб-сервис, позволяющий загружать фотографии, автоматически определять объекты морского мусора и экспортировать результаты в формате JSON. Работа представляет вклад в автоматизацию процессов экологического мониторинга и может быть использована для дальнейших исследований в области борьбы с морским загрязнением. Результаты демонстрируют возможность интеграции модели в системы мониторинга реального времени.

КЛЮЧЕВЫЕ СЛОВА: морской мусор, детекция, мониторинг загрязнений, анализ изображений, бортовые камеры, спутниковый мониторинг, автоматизация, экосистемы, пластиковые отходы, экологическое воздействие, технологии наблюдения, методы детекции.

DEVELOPMENT OF A MARINE DEBRIS DETECTION MODEL

Dostovalov Andrey Andreevich, master student, Dept. of Big Data Analytics and Video Analysis Methods, Ural Federal University

Kapustin Savely Andreevich, master student, Dept. of Big Data Analytics and Video Analysis Methods, Ural Federal University

Abstract

The study focuses on developing a machine learning model for the automatic detection of marine debris based on image analysis. A comparative analysis of deep learning models (Faster R-CNN, SSD, YOLO) was conducted, considering accuracy, processing speed, and computational resource requirements. The YOLOv8 model was identified as the most promising due to its high performance and stability. A unique dataset was developed, featuring annotated images of marine surfaces with debris objects. The model was trained, achieving significant metrics in accuracy (mAP50-95), recall, and object localization. A web service was created, enabling users to upload images, automatically detect marine debris, and export results in JSON format. This work contributes to the automation of environmental monitoring processes and serves as a foundation for further research in combating marine pollution. The results demonstrate the model's potential for integration into real-time monitoring systems.

KEYWORDS: marine debris, detection, pollution monitoring, image analysis, on-board cameras, satellite monitoring, automation, ecosystems, plastic waste, environmental impact, surveillance technologies, detection techniques.

1. Введение

Количество морского мусора в океанах во всем мире увеличивается. Исследователи из Фонда морских исследований Алгалита зафиксировали увеличение количества пластикового мусора в

Центрально-Тихоокеанском круговороте в пять раз в период между 1997 и 2007 годами, где базовый уровень в 1997 году показал, что количество пластиковых частиц превышает количество планктона на поверхности океана в соотношении 6:1 [1]. В связи с этим возникает острая необходимость в разработке и внедрении эффективных методов мониторинга и детекции морского мусора.

Современные подходы к решению этой задачи включают использование различных технологий для автоматизации процесса обнаружения загрязнений, среди которых особое внимание уделяется методам анализа изображений, получаемых с бортовых камер судов, а также спутниковым методам наблюдения.

Ручной поиск — это метод, при котором специалисты физически обследуют морские и прибрежные территории с целью выявления и идентификации мусора. Обычно это включает в себя как визуальный осмотр с судов, так и сбор образцов воды для дальнейшего анализа [2]. Этот метод обладает высокой точностью, так как позволяет экспертам напрямую оценивать типы загрязнений, но ограничен территориально и требует значительных затрат времени и ресурсов (например, для судов и персонала). В дополнение к этому, правовые ограничения, такие как доступ к охраняемым территориям или необходимость получения разрешений для проведения научных исследований, могут усложнить его использование [3].

Спутниковый мониторинг использует данные с орбитальных спутников для анализа загрязнения океанов. Спутниковые данные позволяют получить информацию о масштабе загрязнения в реальном времени, однако они ограничены по точности, особенно в отношении мелких частиц мусора (например, микропластика), и могут быть подвержены атмосферным помехам, таким как облачность или дождь.

Спутниковая съемка активно используется для мониторинга нефтяных разливов и крупных загрязнений. Один из известных спутниковых проектов — это использование данных с аппаратов Sentinel Европейского космического агентства для мониторинга загрязнения в акваториях ЕС [4].

Анализ изображений с бортовых камер судов включает в себя использование камер, установленных на борту судов для съемки поверхности моря и последующего анализа полученных изображений. Технологии обработки изображений, такие как компьютерное зрение и машинное обучение, позволяют эффективно выявлять различные виды загрязнений, включая мелкие частицы пластика. Этот метод обеспечивает высокую точность, особенно при наличии

качественных изображений, но ограничен площадью, которую можно обследовать с судна, а также зависит от погодных условий.

Конечно, использование нейронных сетей для детекции мусора прямоком на собирающем информации судне уже не ново, так например М. Fulton и другие в рамках своей статьи «Robotic Detection of Marine Litter Using Deep Visual Detection Models» обучали модель автоматизации поиска мусора, но, как и почти все другие похожие проекты этот проект был направлен на поиск именно подводного мусора [5].

Новизна данной работы заключается в дополнительном анализе современных методов машинного обучения, их применении в поддержке борьбы с экологической угрозой и возможного развития противодействия вредоносным действиям, а также предоставлением современного актуального бэкграунда для будущих исследований. Исследование может внести вклад в понимание современных тенденций и разработку новых методов в области контроля за экологическим состоянием вод.

Таким образом целью данной работы является готовая модель машинного обучения, способная автоматически обнаруживать и подсчитывать морской мусор на фотографиях морской поверхности. Исходя из цели сформулированы следующие задачи:

- изучение различных моделей машинного обучения и проведение их сравнения;
- изучение методов улучшения качества модели;
- сбор и подготовка датасета для обучения выбранной модели;
- обучение модели на подготовленных данных;
- разработка API модели;
- интерпретация результатов.

2. Материалы и методы

Для задачи детекции рассматривались популярные модели глубокого обучения, такие как Faster R-CNN, SSD и YOLO.

Критерии:

- Точность (Mean Average Precision, MAP): это основной показатель, который используется для сравнения производительности моделей детекции объектов.
- Скорость обработки: Количество кадров в секунду (FPS, frames per second), которое модель может обработать.

- Потребление ресурсов (GPU): Объем вычислительных ресурсов, которые модель требует для работы, включая требования к памяти и графическому процессору.

Таблица 1. Сравнение моделей детекции объектов¹

Критерий	Faster R-CNN	SSD	YOLOv3 (Lu Tan), YOLOv4 (DalmarDakari)
MAP, % (Lu Tan)	87.69	82.41	80.17
FPS (Lu Tan)	7	32	51
MAP, % (DalmarDakari)	61.20	56.80	54.30
FPS (Dalmar Dakari)	8	22	40
GPU, MB (Dalmar Dakari)	5400	3200	2800

На основании сравнения (Таблица 1) наиболее перспективным и эффективным методом детекции морского мусора в реальном времени является использование модели детекции YOLO, потому что модель обеспечивает наиболее высокую скорость обработки и высокую точность, что делает её оптимальным выбором для детекции морского мусора в реальном времени.

Модель SSD также может применяться для задач реального времени и представляет собой более сбалансированный подход между скоростью и точностью, однако её точность и способность детектировать мелкие объекты уступают YOLO.

Faster R-CNN подходит для высокоточных задач, но из-за низкой скорости обработки её использование ограничено в задачах реального времени. Также модель обладает высоким потреблением ресурсов из-за сложной многоступенчатой архитектуры, что делает её требовательной к памяти и вычислительным ресурсам.

Для определения наиболее перспективной и эффективной конкретной вариации YOLO было проведено исследование производительности последних моделей средних размеров (m).

Таблица 2. Сравнение моделей детекции объектов YOLO v8, v9, v10, v11²

Модель	Размер входа (пиксели)	mAPval (50-95), (B)	Параметры (млн.)	FLOPs (млрд)	GPU (ms)
YOLOv8m	640	50.2	25.9	78.9	1.83
YOLOv9m	640	51.4	20.1	76.8	8
YOLOv10m	640	51.1	59.1	59.1	1.84
YOLOv11m	640	51.5	20.1	68.0	4.7 ± 0.1

Объяснение данных:

- Размер входа (pixels): Размер изображения, на котором происходит обучение или детекция.
- mAPval (50-95): Средняя точность модели при различных порогах (от 50 до 95).

¹ Сделан авторами на основании работ [6, 7].

² Сделан авторами на основании исследований [8, 9, 10, 11, 12].

- Параметры (М): Количество параметров модели в миллионах.
- FLOPs (B): Вычислительные затраты модели в миллиардах операций с плавающей запятой.
- GPU (ms): Время, необходимое для выполнения предсказания, зависит от аппаратных условий.

Таким образом, на основе проведённого исследования (Таблица 2) мы пришли к выводу что по показателям лучшим решением является 10 версия, но основываясь на огромном опыте ML-специалистов все же выбрали 8 версию, за счет больших возможностей ее интеграции с другим программным обеспечением, сервисами и инструментами, а также за счет ее большей стабильности в работе и детекции.

В самом начале работы были взяты данные предоставленные ПАО «Сбербанк», которые содержали 16000 фотографий сделанных в реальных условиях с борта судна, но на последующих этапах уже использовали датасеты размещенные другими исследователями на платформе Roboflow, поскольку первоначальный сет содержал всего 8% фотографий с мусором от всего числа, да и мусор был не четко выраженным, поэтому даже использование стратификации и аугментации не дало значительного результата.

3. Результаты

Как и говорилось ранее первоначальное обучение модели проходило на базе реальных фотографий полученных с плаваний круизных лайнеров, торговых кораблей и других судов. После разметки всех фотографий уже предполагалось сомнительное качество модели обучаемой до этого. На рисунках ниже вы можете увидеть пример размеченной фотографии.

На Рисунке 1 показан пример размеченной фотографии и можно заметить что мусор не такой уж и большой на фотографии, хотя можно отметить что это один из самых больших объектов похожих на мусор, но основная проблема заключается в том что искомые объекты зачастую слишком сильно схожи с морской пеной, что в свою очередь сильно затрудняет их детекцию и обучение модели.

Таблица 3. Лучшие 5 эпох согласно оценке YOLO

Эпоха	Точность, (B)	Полнота, (B)	mAP50, (B)	mAP50-95, (B)
46	0,8014	0,2749	0,2912	0,1069
47	0,7865	0,2749	0,2944	0,1043
39	0,7912	0,2567	0,2676	0,1036
44	0,7221	0,2691	0,2890	0,1036
50	0,8362	0,2669	0,2918	0,1033



Рис. 1. Пример размеченной фотографии в ее реальном размере

В Таблице 3 представлены результаты обучения модели на начальном датасете, как видно наилучшими по показателю полноты (recall) стали 46 и 47 эпохи, в которых он составил 27,5%. Это говорит о том, что в валидационном пакете фотография модель смогла найти мусор только на 27,5% фотографий с мусором, следовательно, найден мусор будет только на 1 из 4 фотографий. И это вполне логично поскольку при обучении модель сжимает фотографии до нужного формата (все опирается в вычислительные мощности) и при обучении на фотографиях 1024×1024 пикселя мусор на почти всех фотографиях становится размером с пару пикселей и даже человеку не сказать мусор это или нет. Поэтому следующим шагом мы провели обрезку фотографий для получения фотографий мусора без потери качества.

Использование этого метода в совокупности с методами аугментации не дал значительного прироста, оставив полноту и точность на том же уровне, но увеличив mAP50-95 на 1-2%.

Датасеты полученные с Roboflow же содержали более значимые изображения мусора, по которые значительно увеличили эффективность модели. Сам датасет содержит порядка 3000 фотографий, что больше, чем изначальный датасет.

Как видно на Рисунке 3 на фотографиях с Roboflow гораздо лучше видно мусор при этом она сильно отлична от пены, что сильно увеличило эффективность обучения модели.



Рис. 2. Пример той же самой размеченной фотографии после обрезки

Таблица 4. Лучшие 5 эпох обучения последней версии модели согласно оценке YOLO

Эпоха	Точность, (B)	Полнота, (B)	mAP50, (B)	mAP50-95, (B)
24	0.83629	0.75019	0.8345	0.4456
35	0.84384	0.74397	0.82995	0.4453
33	0.83479	0.69993	0.8111	0.4228
38	0.78569	0.75097	0.80637	0.4224
25	0.81145	0.74163	0.8151	0.4391

В Таблице 4 можно видеть результаты 5 лучших эпох при обучении модели на нашем финальном датасете и как можно видеть лучший результат был достигнут в 24 эпоху, при этом результат значительно превосходил результаты модели, обученной на начальном датасете. Например, mAP50 и mAP50-95 выросли в 4 раза, полнота – в 3.



Рис. 3. Пример размеченной фотографии с Roboflow в ее реальном размере

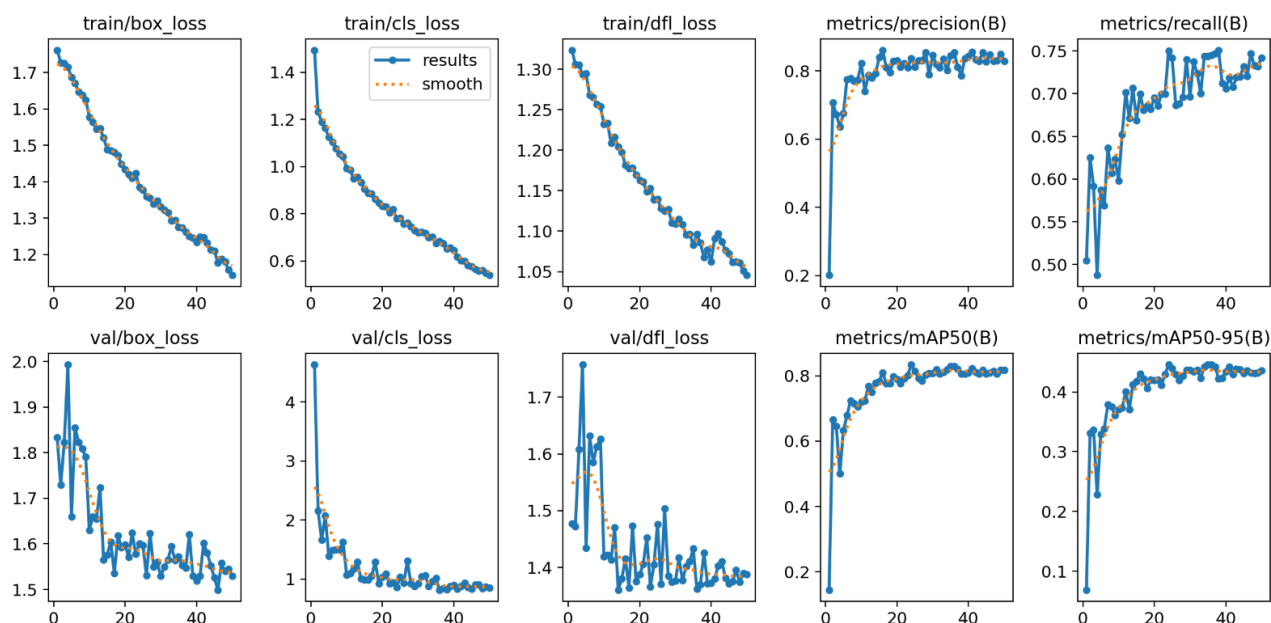


Рис. 4. Все метрики итоговой модели

На Рисунке 4 приведены все метрики, которые получали по мере обучения итоговой модели, на них наглядно видно постепенное уменьшение ошибок локализации (box_loss) и увеличение точности, полноты и mAP50-95, которые показывали ощутимо более значимые результаты чем модели обучаемые на прошлых датасетах.

Следующим этапом стало разработка пользовательского веб-сервиса позволяющего получить доступ к модели. Он позволяет загружать фотографии, а после разметки выводит фотографии с уже размеченным на ней мусором, выделенным боксами и количеством мусора на фотографии. Результаты разметки также можно экспортировать из веб-сервиса в JSON формате.

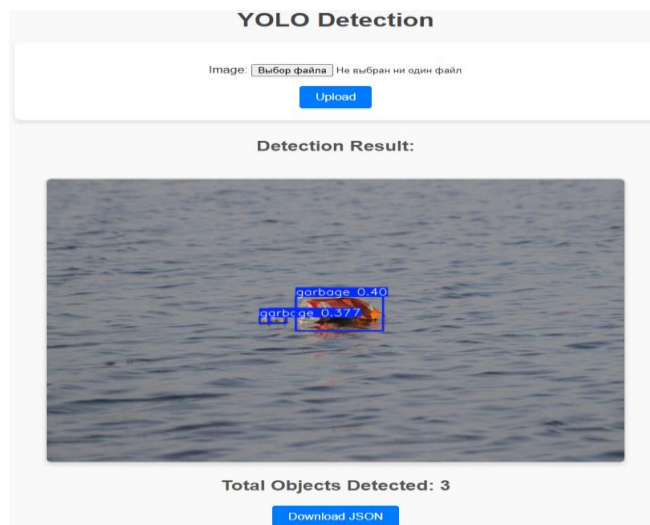


Рис. 5. Разработанный для модели веб-сервис

На Рисунке 5 показан интерфейс пользовательского веб-сервиса, предназначенного для работы с нашей моделью по детекции морского мусора на поверхности.

```
{
  "image_name": "photo_2024-12-09_19-43-00_VIvJhKL.jpg",
  "total_objects_detected": 3,
  "detections": [
    {
      "bbox": [
        385.8293762207031,
        246.02366638183594,
        520.3457641601562,
        319.2868957519531
      ],
      "class": [
        0.0
      ],
      "confidence": [
        0.39541617035865784
      ]
    },
    {
      "bbox": [
        329.7617492675781,
        290.73809814453125,
        343.5577697753906,
        303.389892578125
      ],
      "class": [
        0.0
      ],
      "confidence": [
        0.36832675337791443
      ]
    },
    {
      "bbox": [
        344.1808166503906,
        289.5313415527344,
        370.7807922363281,
        303.28448486328125
      ],
      "class": [
        0.0
      ],
      "confidence": [
        0.26796528697013855
      ]
    }
  ]
}
```

Рис. 6. JSON-файл экспортированный из веб-сервиса

Как видно на рисунке 6 JSON-файл имеет понятную структуру, которая сразу отражает название загруженной фотографии, для более быстрого сопоставления, количество обнаруженных объектов, класс объекта (в модель введен только один класс), координаты

обнаруженного объекта в пикселях и уверенность модели что обнаруженный объект является мусором (где 1 это 100% мусор, а 0 вообще не мусор по мнению модели).

4. Обсуждение и заключение

По итогам работы были проведены сравнительный анализ и количественная оценка различных моделей детекции морского мусора, включая версии YOLO. На основании полученных результатов выбрана оптимальная модель (YOLOv8), которая демонстрирует высокую скорость обработки и точность при интеграции с различным программным обеспечением. Проведен тщательный процесс подготовки данных, включая разметку и аугментацию изображений, что позволило улучшить качество обучения модели.

Разработанный веб-сервис расширяет возможности применения модели, позволяя пользователям удобно загружать и анализировать фотографии с детекцией объектов морского мусора, а также экспортировать результаты. Это решение обеспечивает не только автоматизацию задач мониторинга, но и предоставляет инструмент для дальнейших исследований в области анализа загрязнений.

Для перспективного развития темы возможно проведение более глубокого исследования эффективности других методов детекции и интеграция с системами спутникового мониторинга. Кроме того, комбинирование YOLO с другими алгоритмами глубокого обучения, такими как SSD и Faster R-CNN, может открыть новые возможности повышения точности и надежности детекции. В долгосрочной перспективе идеальным решением станет разработка универсального алгоритма, способного адаптироваться к различным условиям съемки и типам загрязнений с минимальным количеством ложноположительных результатов.

Литература

1. Joshi S., Temgire S. Life below water and the fate of humanity: with special reference to the efforts by Adidas towards reducing the effects of plastic waste on the marine life //Sustainability, Agri, Food and Environmental Research. – 2022. – Т. 10.
2. Ribic C. A., Dixon T. R., Vining I. Marine debris survey manual. – 1992.
3. Agamuthu P. et al. Marine debris: A review of impacts and global initiatives //Waste Management & Research. – 2019. – Т. 37. – №. 10. – С. 987-1002.
4. Blondeau-Patissier D. et al. Detection of marine oil-like features in Sentinel-1 SAR images by supplementary use of deep learning and empirical methods: Performance assessment for the Great Barrier Reef marine park //Marine Pollution Bulletin. – 2023. – Т. 188. – С. 114598.

5. M. Fulton, J. Hong, M. J. Islam and J. Sattar. Robotic Detection of Marine Litter Using Deep Visual Detection Models. 2019 International Conference on Robotics and Automation (ICRA), Montreal, QC, Canada, pp. 5752-5758, doi: 10.1109/ICRA.2019.8793975.
6. Tan L. et al. Comparison of yolo v3, faster r-cnn, and ssd for real-time pill identification. – 2021.
7. Dalmar D., & Daniel C. A Comparative Analysis of Modern Object Detection Algorithms: YOLO vs. SSD vs. Faster R-CNN //ITEJ – 2023 – T.8 – №. 2. – C.96-106.
8. Ultralytics. "YOLOv8 Model Documentation." Retrieved from <https://docs.ultralytics.com/models/yolov8/#citations-and-acknowledgements> Accessed October 10, 2024.
9. Ultralytics. "YOLOv9 Model Documentation." Retrieved from <https://docs.ultralytics.com/models/yolov9/#yolov9-benchmarks> Accessed October 10, 2024.
10. Ultralytics. "YOLOv10 Model Performance." Retrieved from <https://docs.ultralytics.com/models/yolov10/#performance> Accessed October 10, 2024.
11. Ultralytics. "YOLO11 Model Performance Metrics." Retrieved from <https://docs.ultralytics.com/models/yolo11/#performance-metrics> Accessed October 10, 2024.
12. Yaseen M. What is YOLOv9: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector //arXiv preprint arXiv:2409.07813. – 2024.