

# Определение степени заполненности мусорных баков при помощи компьютерного зрения

Я. О. Сениченкова, М. Д. Поляк

Санкт-Петербургский государственный университет аэрокосмического приборостроения

markpolyak@gmail.com

**Аннотация.** В работе рассматривается решение проблемы несвоевременного вывоза мусора с помощью машинного зрения. Предлагается осуществлять фото- и видеofиксацию состояния контейнерных площадок, на которых расположены мусорные баки, после чего, с помощью нейросетевых моделей, определять количество скопившегося мусора. Для решения данной задачи была подготовлена обширная выборка из изображений мусорных баков разных типов и форм. На основе архитектуры ResNet обучена нейронная сеть для классификации мусорных баков. С помощью дообученной модели Yolo осуществляется детекция отдельных баков на изображении.

**Ключевые слова:** ResNet, Yolo, классификация, обработка изображений, экология

## I. ВВЕДЕНИЕ

В последнее десятилетие технологии компьютерного зрения вышли на новый виток развития. Все началось с представления нейронной сети AlexNet [1] на конкурсе ImageNet в 2012 году. Это событие является знаменательным, т. к. в предыдущие года в данном конкурсе побеждали классические алгоритмы машинного обучения, а не нейронные сети. Так, решение, представленное командой NEC-UIUC в 2010 году, является примером подхода к классификации изображений без использования нейронных сетей [2]. Активное использование сверточных нейронных сетей [3] началось после 2015 года, т. к. в этом году нейронная сеть ResNet [4] смогла превзойти человека по качеству классификации изображений на конкурсе ImageNet [5]. В данной работе предлагается использовать технологии компьютерного зрения и глубокого обучения для распознавания уровня заполненности мусорных баков. В качестве решения данной задачи авторы работы [6] использовали сверточную нейронную сеть, которая смогла достичь точности 93 %. Также для решения данной задачи применяются датчики, определяющие уровень заполненности мусорного бака [7]. Однако данное решение является дорогостоящим при внедрении в крупных населенных пунктах. Это связано с тем, что датчики необходимо поместить во все мусорные баки, также есть риски, связанные с вандализмом (оборудование могут воровать) и возможным повреждением датчиков острыми и твердыми отходами.

Автоматизация распознавания наполненности мусорного бака позволит службам, ответственным за вывоз мусора, оперативно получать информацию о том, какие контейнеры переполнены. Информация о наполненности мусорных баков позволит адаптировать график вывоза мусора к ситуативным изменениям, как

единообразным, так и сезонным. Коммунальные службы смогут оптимизировать свои расходы и повысить качество предоставляемых населению услуг.

Для решения поставленной задачи используются YOLOv8 для детекции бака и ResNet18 для классификации наполненности. На вход модели поступает изображение размером 224x224. Далее, часть изображения, попавшая в bounding box, поступает в классификатор, реализованный нейронной сетью ResNet18. Данный классификатор вычисляет вероятности принадлежности изображения к классам «полный бак», «полуполный бак», «пустой бак», «неопределенный бак». Примеры изображений данных классов приведены на рис. 1–3.



Рис. 1. Пример изображения, относящегося к классу «Пустой бак»



Рис. 2. Пример изображения, относящегося к классу «Полный бак»



Рис. 3. Пример изображения, относящегося к классу «Полупустой бак»

Последний класс необходим на случай, если по изображению невозможно определить наполненность контейнера из-за различных факторов (например: угла съемки, освещения). Пример подобных фотографий приведен на рис. 4.



Рис. 4. Примеры фотографий контейнеров, относящихся к классу «Неопределенный бак»

## II. ПОДГОТОВКА ДАННЫХ

Для обучения и тестирования классификатора были собраны фотографии и видео, содержащие мусорные баки. Видео были покадрово нарезаны на изображения. Всего в датасете 8065 фотографий. 933 изображения относятся к классу «Пустой бак», 2863 изображения – к классу «Полный бак», 4165 изображений – к классу «Полупустой бак», 104 изображения – к классу «Неопределенный бак». Выборка была разделена на обучающую, валидационную и тестовую часть. В тестовой выборке содержится 906 изображений. В обучающей выборке – 7258 фотографий, в валидационной – 807.

## III. ОБУЧЕНИЕ КЛАССИФИКАТОРА

Для дообучения нейронной сети ResNet18 использовалась техника Fine tuning, при которой последний слой классификатора заменяется на слой, в котором количество нейронов совпадает с количеством классов изображений. В нашем случае датасет содержит 4 класса, следовательно, последний слой классификатора ResNet18 заменяется слоем, содержащим 4 нейрона. В качестве функции потерь была выбрана Cross Entropy Loss [8]. После дообучения, состоящего из 5 эпох, получились следующие результаты на тестовой выборке: Точность – 0.9837, F1-score – 0.9834, ROC-AUC score – 0,9997. На рис. 5 приведен график изменения значения функции потерь на обучающей выборке, на рис. 6 – график изменения точности.

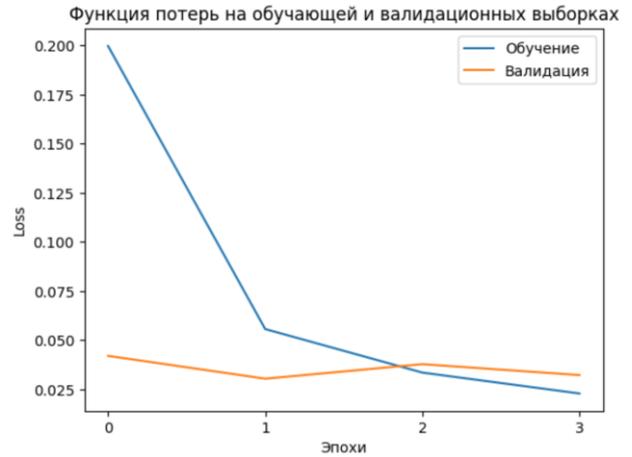


Рис. 5. График изменения значения функции потерь на обучающей и валидационных выборках в процессе обучения классификатора

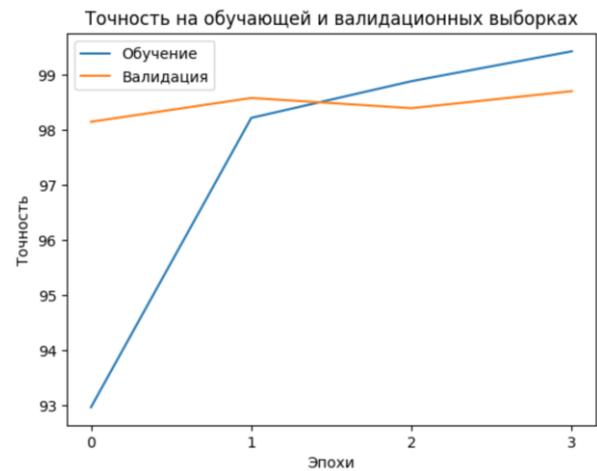


Рис. 6. График изменения значения точности на обучающей и валидационных выборках в процессе обучения классификатора

ТАБЛИЦА I. МАТРИЦА ОШИБОК КЛАССИФИКАТОРА НА ТЕСТОВОЙ ВЫБОРКЕ

| Истинный класс | Предсказанный класс |            |                |            |
|----------------|---------------------|------------|----------------|------------|
|                | Пустой бак          | Полный бак | Полупустой бак | Неизвестно |
| Пустой бак     | 102                 | 1          | 1              | 0          |
| Полный бак     | 0                   | 316        | 1              | 1          |
| Полупустой бак | 5                   | 4          | 453            | 1          |
| Неизвестно     | 0                   | 1          | 0              | 11         |

## IV. ОБУЧЕНИЕ ДЕТЕКТОРА

В качестве детектора была взята нейронная сеть YOLOv8 [9]. По прошествии 10 эпох были получены следующие графики метрик, приведенные на рис. 7.

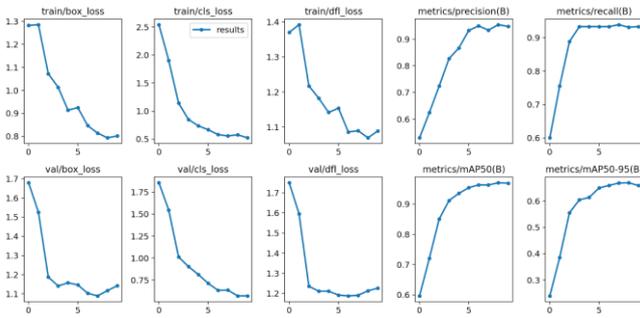


Рис. 7. Графики метрик, полученных в результате обучения YOLOv8

Были получены следующие метрики на тестовой выборке: box loss: 1.142, cls loss: 0.570, dfl loss: 1.23, mAP50: 0.969, mAP50-95: 0.658

На рис. 8 приведены пример работы полученного детектора на изображениях из тестовой выборки. Сверху ограничительной рамки указана вероятность, с которой модель отнесла объект к мусорному баку.



Рис. 8. Примеры изображений, полученных в результате работы обученной

На рис. 9 приведена нормализованная матрица ошибок для каждого класса из тестовой выборки. На диагонали можно увидеть процент верно классифицированных объектов.

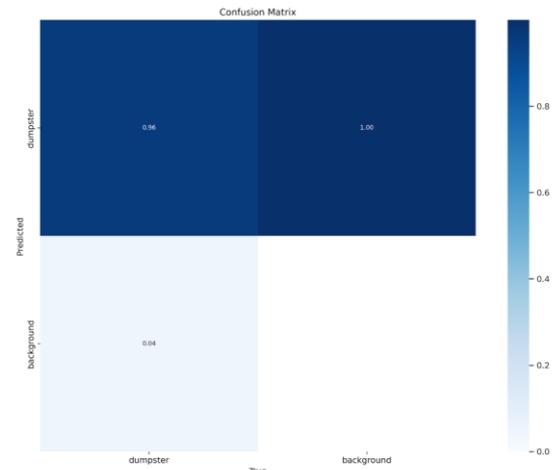


Рис. 9. Нормализованная матрица ошибок для тестовой выборки

## V. РЕЗУЛЬТАТЫ

Визуальный анализ неверно классифицированных изображений позволил выделить следующую закономерность: если внутри мусорного бака есть пятна, то нейронная сеть посчитает, что бак является полупустым. Возможно, это вызвано тем, что пятна ошибочно принимаются за объекты внутри контейнера. Пример данных изображений приведен на рис. 10.



Рис. 10. Примеры изображений, ошибочно отнесенных нейронной сетью к классу «Полупустой бак»

## VI. ЗАКЛЮЧЕНИЕ

В ходе исследования был построен пайплайн, который позволяет определить уровень заполненности мусорного бака. Детектор мусорных баков работает с полнотой 0.969 при значении порога IoU равным 0.5, классификатор определяет уровень наполнения контейнера для мусора с точностью 0.983. В ходе данного исследования было показано, что задача определения наполненности мусорного бака может быть достоверно решена с помощью нейросетевых алгоритмов машинного зрения. Автоматизация данной проблемы позволит коммунальным службам и экологическим сервисам более оперативно вывозить мусор из переполненных контейнеров.

#### БЛАГОДАРНОСТЬ

Авторы выражают благодарность компании ООО «АСТЕРА» за предоставление вычислительных мощностей, и ведущего инженера Бояркина Михаила Игоревича за ценные советы.

#### СПИСОК ЛИТЕРАТУРЫ

- [1] Krizhevsky A., Sutskever I., Hinton G.E. ImageNet classification with deep convolutional neural networks // Communications of the ACM. 2017. Т. 60. №. 6. С. 84-90.
- [2] Zhou X. et al. Image classification using super-vector coding of local image descriptors // Computer Vision–ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part V 11. Springer Berlin Heidelberg, 2010. С. 141-154.
- [3] LeCun Y., Bengio Y. (1995). Convolutional networks for images, speech, and time series. The handbook of brain theory and neural networks, 3361(10), 1995.conference on computer vision and pattern recognition. 2016. С. 770-778.
- [4] He K. et al. Deep residual learning for image recognition // Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. С. 770-778.
- [5] ImageNet 2015. Дата обращения: <https://image-net.org/challenges/LSVRC/2015/> (accessed 01 April 2024)
- [6] Ramirez I., Cuesta-Infante A., Pantrigo J.J., Montemayor A.S., Moreno J.L., Alonso V., Anguita G., Palombarani L. (2020). Convolutional neural networks for computer vision-based detection and recognition of dumpsters // Neural Computing and Applications, 32(17), 13203-13211.
- [7] Draz U., Ali T., Khan J. A., Majid M., Yasin S. (2017, November). A real-time smart dumpsters monitoring and garbage collection system. // In 2017 Fifth International Conference on Aerospace Science & Engineering (ICASE) (pp. 1-8). IEEE.
- [8] Cox D.R. The regression analysis of binary sequences // Journal of the Royal Statistical Society Series B: Statistical Methodology. 1958. Т. 20. №. 2. С. 215-232.
- [9] Han X., Chang J., Wang, K. (2021). You only look once: unified, real-time object detection // Procedia Computer Science, 183(1).