

В. О. Архипов¹✉

Автоматизация распознавания объектов на космических снимках методами глубокого обучения

¹Сибирский государственный университет геосистем и технологий,
г. Новосибирск, Российская Федерация
e-mail: arhipov-vo2024@sgugit.ru

Аннотация. В работе рассматривается задача автоматизации распознавания объектов на космических снимках с использованием современных методов глубокого обучения. Цель работы: разработать автоматизированный подход в решении задачи распознавания объектов на космических снимках с применением методов глубокого обучения. Был определен наиболее эффективный подход к обучению нейросети для решения данной задачи. Для построения датасета использовались изображения с Google Maps, для разметки применялся сервис Roboflow. Было собрано и размечено 244 изображения разрешением 1024×1024. Для сравнения двух подходов использовалась модель YOLOv11 из фреймворка Ultralytics. Модель была обучена на платформе Kaggle. Результаты эксперимента показали, что предобученная модель достигает более высоких показателей точности обнаружения объектов и наложения маски на собранном датасете, чем при обучении «с нуля»: точность обнаружения составила 0,816 против 0,72, а метрика наложения маски – 0,814 против 0,719 соответственно. В дальнейшем планируется масштабирование экспериментов на большие датасеты для уточнения и обобщения выводов.

Ключевые слова: нейронная сеть, python, датасет, ultralytics, yolo, глубокое обучение

V. O. Arkhipov¹✉

Training a neural network to search for people in photos obtained from a UAV

¹Siberian State University of Geosystems and Technologies, Novosibirsk, Russian Federation
e-mail: arhipov-vo2024@sgugit.ru

Abstract. This paper addresses the task of automating object recognition in satellite imagery using modern deep learning methods. The objective of the study is to develop an automated approach for object recognition in satellite images with the application of deep learning techniques. The most effective approach for training the neural network for this task was identified. Google Maps images were used to compile the dataset, and the Roboflow service was employed for annotation. A total of 244 images with a resolution of 1024×1024 were collected and annotated. The YOLOv11 model from the Ultralytics framework was used to compare two approaches. The model was trained on the Kaggle platform. The experimental results showed that the pretrained model achieved higher detection accuracy and mask overlay metrics on the collected dataset than training “from scratch”: detection accuracy reached 0,816 compared to 0,72, and the mask overlay metric was 0,814 compared to 0,719, respectively. In the future, it is planned to scale the experiments to larger datasets to refine and generalize the findings

Keywords: neural network, python, dataset, ultralytics, yolo, deep learning

Введение

Задача анализа объектов на космических снимках актуальна в разных областях, например, распознавание участков поврежденных лесов, мониторинг землепользования, состояния растительности и городского развития, ликвидация последствий стихийных бедствий, мониторинг изменений характеристик поверхности, отслеживание лесных пожаров, создание цифровых карт [2 – 16].

На сегодняшний день одним из самых продвинутых и перспективных методов компьютерного зрения является глубокое обучение и нейронные сети. Применение методов глубокого обучения может позволить более эффективно обрабатывать большие массивы данных и снизить трудозатраты в сравнении с ручной обработкой [1 – 16].

Цель работы: разработать автоматизированный подход в решении задачи распознавания объектов на космических снимках с применением методов глубокого обучения.

Задачи:

- собрать датасет;
- обучить нейросетевую модель;
- разработать подход к автоматизации распознавания объектов на космических снимках.

Выбор фреймворка и модели

Проводилось сравнение нескольких популярных фреймворков и представленных в них моделей на схожих по специфике данных (объекты на снимках с БПЛА): Ultralytics, Detectron, MMDetection. В итоге был выбран Ultralytics, так как с его помощью удалось обучить наиболее точные модели, а также в сравнении с другими фреймворками он выделяется простотой настройки и обширной документацией. Он включает в себя множество различных современных моделей для обнаружения объектов, таких как RT-DETR, YOLO, SAM, и т.д [1, 23].

Модель YOLOv11 – наиболее точная и производительная модель среди представленных (рис. 1).

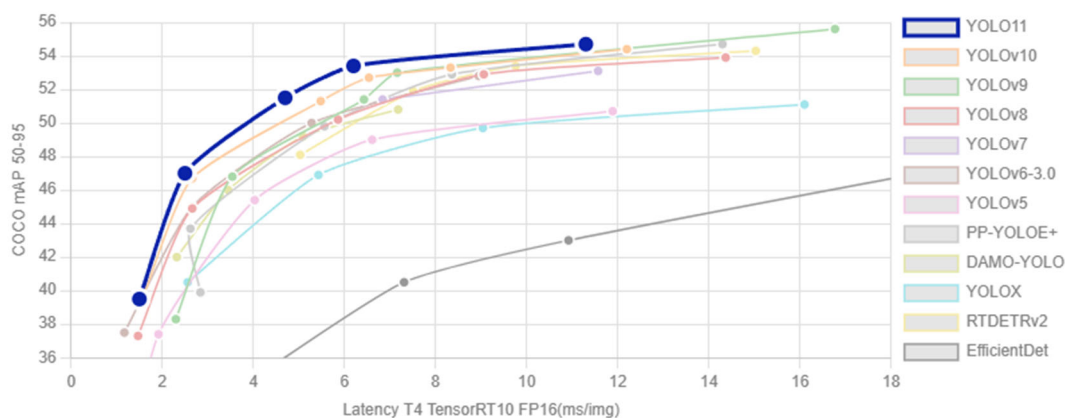


Рис. 1. График сравнения моделей [22]

Архитектура YOLO использует подход, в котором обнаружение объектов рассматривается как единая регрессионная задача, предсказывая ограничительные рамки и вероятности классов непосредственно по полным изображениям за одну оценку. Благодаря этому методу модели YOLO работают значительно быстрее, чем предыдущие двухэтапные детекторы, сохраняя при этом высокую точность [23].

Разметка данных

Существует множество сервисов, предоставляющих доступ к спутниковым снимкам, например, EarthExplorer, Google Earth Engine, Sentinel Hub, Google Maps, Яндекс Карты [17, 18, 22].

Снимки для датасета были взяты с Google Maps: с помощью снимка экрана получены изображения разрешением $\sim 2500 \times 1400$, при этом снимок максимально был приближен таким образом, чтобы здания могли быть распознаны и изображение не потеряло в качестве, т.е. с расчетной высотой обзора ~ 20 -30 метров.

Поскольку здания могут быть разной формы и в разном расположении на снимке (угол поворота, угол съемки), что затрудняет использование прямоугольной разметки, было решено использовать сегментацию для более качественного обучения и точного распознавания зданий нейросетью. Есть два вида сегментации: семантическая и instance-сегментация. Разметка была осуществлена семантической сегментацией, т.е. с использованием одного класса – buildings.

Для разметки применялся сервис Roboflow ввиду удобства его использования, поддержки различных форматов, в том числе и формата yolo, а также наличия разных полезных функций для постобработки [24].

Модель Yolo принимает на вход квадратные изображения с установленным при обучении разрешением, величина которого влияет на размер нейросети и ее производительность, поэтому изображения были поделены на 4 части и сжаты до разрешения 1024×1024 . Размер собранного датасета – 244 изображения (рис. 2).



Рис. 2. Фрагмент размеченного изображения

Обучение модели

Многие популярные фреймворки для обучения нейросетей-детекторов имеют предобученные на больших датасетах (COCO, ImageNet и т.п.) модели и предоставляют возможность их дообучения и обучения «с нуля». Дообучение на целевых данных может лучше сказаться на результатах, однако, что если данные в датасетах сильно различаются? Стандартные датасеты содержат изображения с объектами, снятыми крупным планом в высоком разрешении: люди, животные, растения, транспорт и тд, а эти данные отличаются по специфике от данных со спутниковых снимков, например, зданий, которые представляют собой объекты с низким разрешением. Был проведен эксперимент для выявления самого подходящего подхода из двух [23].

Для обучения использовались:

- Платформа Kaggle с графическим ускорителем P100;
- Фреймворк Ultralytics 8.3.130;
- Pytorch 2.5.1;
- CUDA 12.4;
- Параметры: epochs=100, imgsz=1024, batch=4.

Результаты обучения показывают, что предобученная модель обучилась лучше, точность обнаружения – 0,816 против 0,72, точность наложения маски – 0,814 против 0,719 (табл. 1).

После обучения модели были протестированы на изображении, которого нет в датасете. По результатам распознавания видно, что предобученная модель более уверенно размечает здания, маски точнее, меньше пропусков, однако ложные срабатывания также присутствуют (рис. 3).

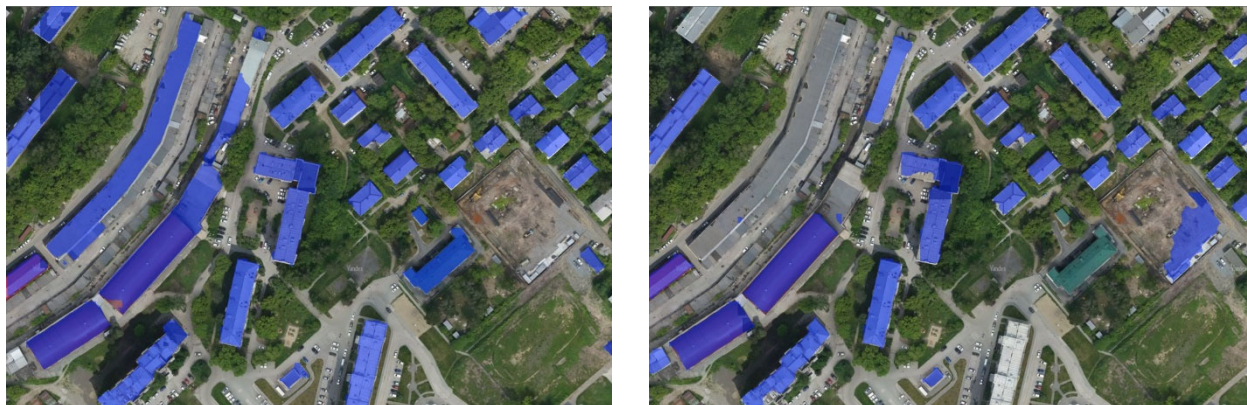


Рис. 3. Результаты работы модели, обученной двумя подходами (слева – предобученная модель, справа – модель, обученная «с нуля»)

Алгоритм автоматизации распознавания объектов

Распознавание объектов на космических снимках можно автоматизировать следующим образом:

- на вход подается космический снимок;

- в зависимости от указанной пользователем расчетной высоты обзора на снимке, он делится на изображения разрешением 1024x1024;
- изображения обрабатываются нейросетью и на выходе получается растровая маска для каждого объекта.

Полученную бинарную маску можно использовать для автоматизации оцифровки зданий, для этого ее можно преобразовать в векторный формат с помощью алгоритмов компьютерного зрения с учетом угла съемки.

Результаты

Таблица 1

Результаты обучения

Модель yolo v11s	Box				Mask			
	P	R	mAP50	mAP50-95	P	R	mAP50	mAP50-95
Обученная «с нуля»	0,805	0,639	0,72	0,448	0,8	0,643	0,719	0,471
Предобученная	0,813	0,75	0,816	0,639	0,86	0,718	0,814	0,617

Заключение

В рамках данной работы была рассмотрена задача автоматизации распознавания объектов на космических снимках, собран датасет, в котором размечены крыши зданий на спутниковых снимках, определен наилучший подход к обучению модели yolo v11 на небольшом датасете.

Экспериментально установлено что наилучшим подходом в данном случае является обучение предобученной модели. Для дальнейшего исследования планируется обучить данную модель на крупном датасете (от 10000 изображений) с применением двух рассмотренных подходов, чтобы более точно интерпретировать результаты эксперимента.

Также был разработан алгоритм автоматизации распознавания объектов, который можно использовать для оцифровки зданий и создания цифровых карт.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Архипов, В. О. Обучение нейросети для поиска людей по фотоснимкам, полученным с БПЛА / В. О. Архипов, А. А. Шарапов // Интерэкспо Гео-Сибирь. – 2024. – Т. 7, № 1. – С. 29-34. – EDN PXEMPJ. URL: https://www.elibrary.ru/download/elibrary_69212979_44018838.pdf (Дата обращения 11.05.2025). Режим доступа: для зарегистрир. пользователей.
2. Белякова, А. Ю. Решение задачи распознавания объектов и инцидентов на фотоматериалах, полученных с беспилотных летательных аппаратов с использованием методов глубокого обучения / А. Ю. Белякова, Ю. Д. Беляков, П. С. Замятин // Инженерный вестник Дона. – 2021. – № 5(77). – С. 210-218. – EDN ZVCEHJ. URL: https://www.elibrary.ru/download/elibrary_46275854_96311013.pdf (Дата обращения 11.05.2025). Режим доступа: для зарегистрир. пользователей.

3. Ваганов, А. В. Алгоритмы обработки изображений и распознавания объектов животного мира на фотоснимках фотовидеорегистраторов методами машинного обучения / А. В. Ваганов, К. С. Печененко, Л. А. Хворова // Известия Алтайского государственного университета. – 2024. – № 1(135). – С. 89-94. – DOI 10.14258/izvasu(2024)1-12. – EDN PBBBLN. URL: https://www.elibrary.ru/download/elibrary_67347961_40912454.pdf (Дата обращения 11.05.2025). Режим доступа: для зарегистрир. пользователей.
4. Ваньков, Е. Э. Алгоритмы машинного обучения для решения задач распознавания / Е. Э. Ваньков, И. К. Радыгин // Мавлютовские чтения : Материалы XVI Всероссийской молодежной научной конференции. В 6-ти томах, Уфа, 25–27 октября 2022 года. Том 5. – Уфа: Уфимский государственный авиационный технический университет, 2022. – С. 1099-1109. – EDN LSTMUV. URL: https://www.elibrary.ru/download/elibrary_50072969_82747018.pdf (Дата обращения 11.05.2025). Режим доступа: для зарегистрир. пользователей.
5. Войтов, Д. Ю. Разработка технологии определения породы деревьев с применением компьютерного зрения / Д. Ю. Войтов, С. Б. Васильев, Д. В. Кормилицын // Лесной вестник. Forestry Bulletin. – 2023. – Т. 27, № 1. – С. 60-66. – DOI 10.18698/2542-1468-2023-1-60-66. – EDN PXEBDF. URL: https://www.elibrary.ru/download/elibrary_50188350_40404336.pdf (Дата обращения 11.05.2025). Режим доступа: для зарегистрир. пользователей.
6. Гарифуллин, Н. Б. Нейросетевые методы обработки изображений в компьютерной графике / Н. Б. Гарифуллин, В. Л. Литвинов // Молодежная школа-семинар по проблемам управления в технических системах имени А.А. Вавилова. – 2024. – Т. 1. – С. 13-15. – EDN HXWDID. URL: https://www.elibrary.ru/download/elibrary_67959860_87708718.pdf (Дата обращения 11.05.2025). Режим доступа: для зарегистрир. пользователей.
7. Использование нейросети для выделения легких на рентгеновских изображениях / Р. Ш. Минязев, А. А. Румянцев, А. А. Баев, Т. Д. Баева // Известия Российской академии наук. Серия физическая. – 2019. – Т. 83, № 12. – С. 1655-1658. – DOI 10.1134/S0367676519120159. – EDN UXFRAZ. URL: https://www.elibrary.ru/download/elibrary_41348118_48368050.pdf (Дата обращения 11.05.2025). Режим доступа: для зарегистрир. пользователей.
8. Как превратить спутниковые снимки в карты. Компьютерное зрение в Яндексе. URL: <https://habr.com/ru/companies/yandex/articles/431108/> (Дата обращения 11.05.2025). Режим доступа: свободный.
9. Каракулов, И. В. Влияние размеров объекта обучающей выборки на качество сегментации методами искусственного интеллекта / И. В. Каракулов, А. В. Ключев // Прикладная математика и вопросы управления. – 2018. – № 3. – С. 89-99. – DOI 10.15593/2499-9873/2018.3.06. – EDN YLLCBV. URL: <https://www.elibrary.ru> (дата обращения 11.05.2025). Режим доступа: для зарегистрир. пользователей.
10. Л.В. Семененко, Е.Н. Кочик А.М. Сорока, В.В. Шкабара. Распознавание объектов на спутниковых снимках с использованием нейронных сетей. URL: <https://elib.bsu.by/bitstream/123456789/315234/1/19-22.pdf> (Дата обращения 11.05.2025). Режим доступа: свободный.
11. Ларионов, Р. В. Использование методов машинного обучения для детектирования объектов на данных дистанционного зондирования земной поверхности / Р. В. Ларионов // Физика, техника и технология сложных систем : Тезисы докладов 3-й Всероссийской с международным участием молодежной научно-практической конференции, Ярославль, 22–30 апреля 2021 года. – Ярославль: Ярославский государственный университет им. П.Г. Демидова, 2024. – С. 170-171. – EDN FVUSNY. URL: https://www.elibrary.ru/download/elibrary_67976510_53630502.pdf (Дата обращения 11.05.2025). Режим доступа: для зарегистрир. пользователей.
12. Нейросетевые технологии поиска целевых объектов на снимках дистанционного зондирования Земли / Н. С. Абрамов, А. А. Талалаев, В. П. Фраленко [и др.] // Сборник трудов ИТНТ-2019, Самара, 21–24 мая 2019 года. Том 2. – Самара: Новая техника, 2019. – С. 26-33. – EDN NBZYZH. URL: <http://repo.ssau.ru/bitstream/Informacionnye-tehnologii-i-nanotehnologii/Neirosetevye->

tehnologii-poiska-celevykh-obektov-na-snimkah-distancionnogo-zondirovaniya-Zemli-76225/1/paper5.pdf (Дата обращения 11.05.2025). Режим доступа: для зарегистрир. пользователей.

13. Обучение нейронной сети обнаружению людей на аэроснимках. URL: <https://habr.com/ru/articles/754338/> (дата обращения: 15.12.2024). Режим доступа: свободный.

14. Поиск пропавших людей на снимках лесного массива, полученных с помощью БПЛА или ещё один разбор задачи Цифрового Прорыва. URL: <https://habr.com/ru/articles/699076/> (дата обращения: 15.12.2024). Режим доступа: свободный.

15. Проект Lastmus: как компьютерное зрение помогает спасти потерявшихся людей. URL: <https://habr.com/ru/companies/ods/articles/483616/> (дата обращения: 15.12.2024). Режим доступа: свободный.

16. Распознавание образов с помощью искусственного интеллекта. URL: <https://habr.com/ru/articles/709432/> (дата обращения: 15.12.2024). Режим доступа: свободный.

17. Яндекс Карты. URL: <https://yandex.ru/maps/> (Дата обращения 11.05.2025). Режим доступа: свободный.

18. EarthExplorer. URL: <https://earthexplorer.usgs.gov/> (Дата обращения 11.05.2025). Режим доступа: для зарегистрир. пользователей.

19. Google Earth Engine. URL: <https://earthengine.google.com/> (Дата обращения 11.05.2025). Режим доступа: для зарегистрир. пользователей.

20. Google Maps. URL: <https://maps.google.ru/> (Дата обращения 11.05.2025). Режим доступа: свободный.

21. Kaggle. URL: <https://www.kaggle.com/> (Дата обращения 11.05.2025). Режим доступа: для зарегистрир. пользователей.

22. Sentinel Hub. URL: <https://www.sentinel-hub.com/> (Дата обращения 11.05.2025). Режим доступа: для зарегистрир. пользователей.

23. Ultralytics Docs. URL: <https://docs.ultralytics.com/> (Дата обращения 11.05.2025). Режим доступа: свободный.

24. Roboflow. URL: <https://app.roboflow.com/> (Дата обращения 11.05.2025). Режим доступа: для зарегистрир. пользователей.

© В. О. Архипов, 2025