Студенческий научный электронный журнал
StudArctic Forum

ПЕТРОЗАВОДСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Т. 10, № 2. С. 40–49 **Компьютерные и информационные науки** **2025**

Научная статья
УДК 004.032.26:591.5(470.22)

| | |
|---|---|
| **БОНДАР Григорий Андреевич** | бакалавриат, Петрозаводский государственный университет (Петрозаводск, Россия), *g4bondar@gmail.com* |
| **КУЛЕЗНЕВ Никита Сергеевич** | бакалавриат, Петрозаводский государственный университет (Петрозаводск, Россия), *kuleznevnik.34@gmail.com* |

# РАЗРАБОТКА СЕРВИСА ДЕТЕКТИРОВАНИЯ ЖИВОТНЫХ ПО ФОТОГРАФИЯМ

**Научный руководитель:**
Димитров Вячеслав Михайлович
**Рецензент:**
Кузьменков Александр Алексеевич
Статья поступила: 02.06.2025;
Принята к публикации: 29.06.2025;
Размещена в сети: 29.06.2025.

**Аннотация.** Проблема автоматизации детектирования животных на фотографиях фотоловушек актуальна для экологических исследований. Цель работы – разработка веб-сервиса для автоматического распознавания представителей фауны Карелии. Методы исследования включают формирование датасета из 6706 изображений животных, обучение нейросетевой модели YOLOv11 и создание веб-интерфейса с интеграцией в платформу «Умные экосистемы. Мониторинг 4.0». В итоге была реализована автоматизированная система обработки данных фотоловушек от получения изображений до веб-визуализации результатов для мониторинга биоразнообразия.

**Ключевые слова:** компьютерное зрение, YOLOv11, детектирование объектов, мониторинг экосистем, фотоловушки, распознавание животных

**Благодарности.** Авторы выражают искреннюю благодарность коллегам, внесшим значительный вклад в реализацию проекта: Ильюшину Антону Павловичу и Митрофанову Аркадию Алексеевичу – за активное участие в разработке сервиса; Титовой Анне Вячеславовне и Севрюкову Роману Юрьевичу – за помощь в сборе и разметке обучающей выборки.

**Для цитирования:** Бондар Г. А., Кулезнев Н. С. Разработка сервиса детектирования животных по фотографиям // StudArctic Forum. 2025. Т. 10, № 2. С. 40–49.

Автоматизация процесса детектирования животных на фотографиях является актуальной задачей для экологических исследований, особенно при использовании фотоловушек. Традиционный подход, предполагающий ручной просмотр и анализ каждого изображения, крайне неэффективен при больших объемах данных. Особую значимость задача имеет для Ботанического сада ПетрГУ, активно использующего фотоловушки для мониторинга биоразнообразия.

Современное состояние проблемы характеризуется активным развитием методов компьютерного зрения на основе глубокого обучения. Семейство архитектур YOLO (You Only Look Once) демонстрирует высокую эффективность в задачах детектирования объектов в реальном времени. При этом специализированные решения для российской фауны остаются недостаточно развитыми, что создает необходимость в разработке локализованных

систем.

Анализ существующих решений показывает наличие нескольких аналогов для детектирования животных. Wildlife Insights представляет собой закрытую систему без локализации для России. MegaDetector не осуществляет классификацию видов животных, а лишь определяет их наличие на изображении. TrapTagger является платным сервисом без возможности интеграции с локальными системами. Данные ограничения определяют необходимость создания специализированного решения для условий российских экосистем.

Целью данной работы является разработка веб-сервиса для автоматического распознавания представителей животного мира Карелии по фотографиям с применением современных технологий машинного обучения.

Для достижения поставленной цели сформулированы следующие задачи исследования:

- разработать и обучить нейросетевую модель для детектирования и классификации животных Карелии на основе архитектуры YOLOv11[1];
- создать веб-платформу для автоматизированной обработки данных фотоловушек с интеграцией в экосистему мониторинга;
- провести экспериментальную оценку эффективности разработанной системы.

В ходе исследования применялись следующие методы: системный подход для создания комплексной архитектуры сервиса; методы машинного обучения, включая обучение нейросетевой модели; методы компьютерного зрения для детектирования и классификации объектов.

## 1. Платформа «Умные экосистемы. Мониторинг 4.0» и ее возможности

Платформа «Умные экосистемы. Мониторинг 4.0» представляет собой комплексную мультиагентную систему интеллектуального мониторинга окружающей среды в режиме реального времени. Ее уникальность заключается в применении комплексного подхода в мониторинге, режиме работы, близком к реальному времени, и автоматической обработке разнонаправленной информации.

Основные функциональные возможности платформы включают взаимодействие с мультиагентными данными со всех типов сенсорных устройств, хранение и анализ больших объемов данных, доступ к данным в режиме реального времени, визуализацию данных, встроенные модули обработки и анализа, встроенные модули искусственного интеллекта.

Разрабатываемый сервис отличается специализацией на фауне Карелии, интеграцией с платформой «Умные экосистемы. Мониторинг 4.0» и возможностью дообучения на локальных данных, что обеспечивает его преимущество перед существующими аналогами.

## 2. Разработка нейросетевой модели

Анализ архитектуры YOLOv11 показал, что данная модель представляет собой значительное улучшение по сравнению с предыдущими версиями семейства YOLO. Основные преимущества включают anchor-free архитектуру, декуплированные головы детектирования и динамическое назначение меток, что обеспечивает более стабильное обучение и высокую точность распознавания [Redmon].

Для обучения нейросетевой модели YOLOv11 была сформирована обучающая выборка, содержащая изображения восьми классов животных, характерных для Карелии: медведь, кабан, лиса, заяц, человек, рысь, лось и волк. Выбор данных классов обоснован результатами зоологических исследований фауны региона.

Сбор данных производился тремя основными методами. Парсинг видеороликов с фотоловушек осуществлялся с помощью разработанного скрипта для автоматического

извлечения кадров из видеофайлов фотоловушек, установленных в лесных зонах, схожих с зонами Карелии. Фотографии из Ботанического сада ПетрГУ формировали основу выборки – использовался архив изображений, полученных с фотоловушек, установленных на территории Ботанического сада. Дополнительные изображения из открытых источников составили небольшую часть (около 5 % выборки) и были включены снимки из свободных онлайн-репозиториев для увеличения разнообразия фонов и ракурсов. При отборе применялись следующие критерии: наличие видимого животного на изображении, разнообразие условий съемки, вариативность ракурсов.

Для увеличения разнообразия обучающей выборки была проведена аугментация данных, включающая несколько преобразований. Все изображения были приведены к единому разрешению 1280x720 пикселей для обеспечения стандартизации входных данных. Изображения были преобразованы в градации серого для повышения устойчивости модели к различным условиям освещения и качеству снимков [Шапиро]. Для симуляции естественных помех, возникающих при работе фотоловушек, к изображениям добавлялся случайный шум, затрагивающий до 0,3 % пикселей. Такая комбинация методов аугментации позволила значительно расширить вариативность обучающих примеров без необходимости сбора дополнительных фотографий. Размеченные данные экспортировались в формате TXT, совместимом с YOLO, где каждый файл содержит строки вида: class_id, x_center, y_center, width, height.

Итоговый набор данных включил в себя 3832 исходных изображения, которые были расширены с помощью аугментаций до 6706 изображений и были разделены на три подмножества: обучающая выборка (Train) – 86 % (5748 изображений); валидационная выборка (Valid) – 6 % (384 изображения); тестовая выборка (Test) – 9 % (574 изображения).

Обучение модели YOLOv11 проводилось на выделенном сервере с использованием GPU NVIDIA GeForce RTX 3060. В процессе обучения использовались предобученные веса для применения техники transfer learning [Redmon]. Основные параметры обучения: количество эпох – 150 (сработала ранняя остановка на 88 эпохе); размер батча – 16; размер входного изображения – 640x640.

Модель обучалась в течение 88 эпох до срабатывания механизма ранней остановки, так как дальнейшее обучение не приводило к улучшению результатов.

В ходе работы были получены следующие метрики:

- точность обнаружения (Precision): 0.943
- полнота обнаружения (Recall): 0.936
- mAP@0.5: 0.966
- mAP@0.5-0.95: 0.791

*Таблица 1*

**Метрики эффективности по классам животных**

| Класс животного | mAP@0.5 | Точность (Precision) |
|---|---|---|
| Медведь | 0.987 | 0.969 |
| Кабан | 0.956 | 0.894 |
| Лиса | 0.993 | 0.982 |
| Заяц | 0.992 | 0.963 |
| Человек | 0.994 | 0.950 |
| Рысь | 0.924 | 0.873 |
| Лось | 0.970 | 0.953 |
| Волк | 0.910 | 0.962 |

Детальный анализ эффективности по классам животных представлен в таблице 1. Наиболее высокие показатели точности продемонстрированы для класса «человек» (mAP@0.5=0.994), что объясняется характерными морфологическими особенностями и позами человека. Хорошие результаты получены для лисы (mAP@0.5=0.993) и зайца (mAP@0.5=0.992), что связано с их выраженными видовыми признаками. Наименьшие значения точности зафиксированы для рыси (mAP@0.5=0.924) и волка (mAP@0.5=0.910). Это объясняется схожестью морфологических признаков данных видов с другими представителями семейства кошачьих и собачьих соответственно, а также относительно меньшим количеством обучающих примеров для этих классов в датасете.

На рисунках 1 и 2 представлены метрики, отражающие эффективность и качество работы модели.

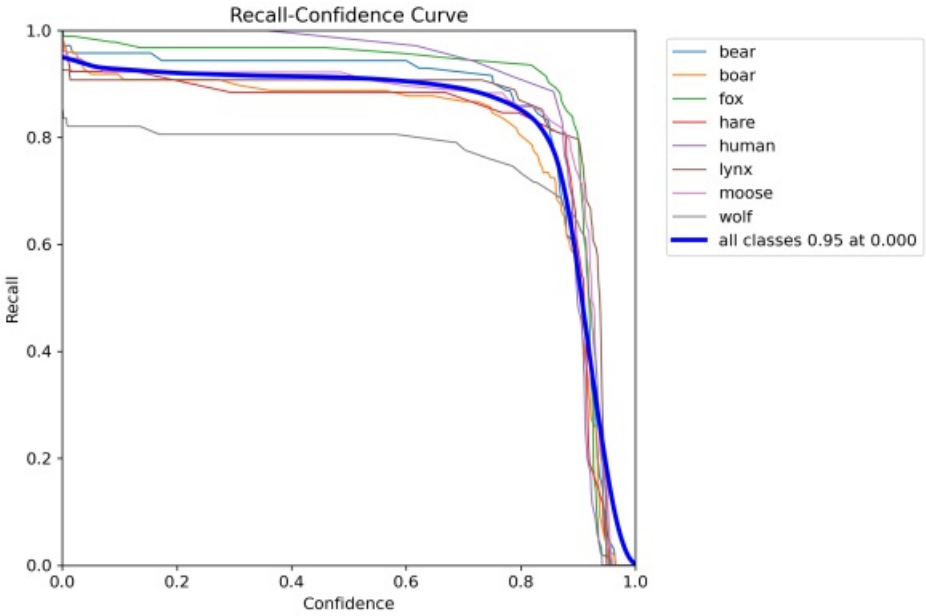


Рис. 1. Кривая точности-полноты для различных классов животных

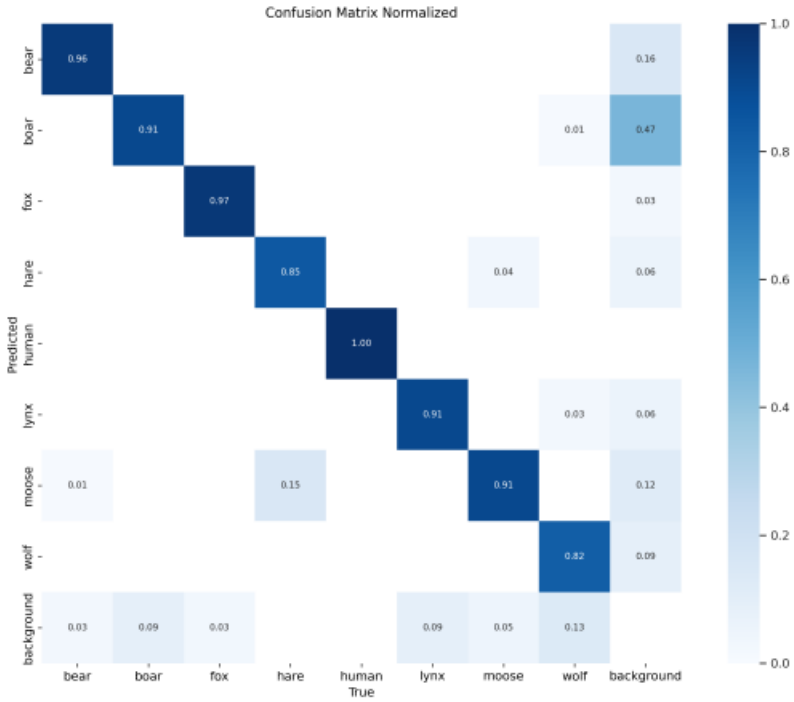


Рис. 2. Матрица ошибок нейросетевой модели

Анализ матрицы ошибок показывает, что основные случаи неправильной классификации происходят между морфологически схожими видами: рысь иногда классифицируется как лиса, волк – как лось при неудачном ракурсе съемки. Данные результаты согласуются с известными сложностями автоматической классификации близкородственных видов [Norouzzadeh].

Достигнутые показатели эффективности (mAP@0.5=0.966) сопоставимы с результатами ведущих исследований в области автоматического распознавания диких животных. Результаты свидетельствуют о высокой точности модели как в определении наличия объектов на изображении, так и в их классификации.

## 3. Разработка веб-сервиса

Для практического применения обученной модели детектирования была разработана комплексная веб-платформа, обеспечивающая автоматизированную обработку данных фотоловушек. Архитектура системы построена на принципах модульности и масштабируемости, что позволяет эффективно обрабатывать большие объемы изображений и обеспечивать стабильную работу в условиях интенсивного использования.

Функциональная архитектура веб-сервиса включает три ключевых компонента, каждый из которых решает специфические задачи автоматизации мониторинга: база данных, обеспечивающая надежное хранение метаданных и результатов анализа; модуль взаимодействия с фотоловушками, отвечающий за прием и первичную обработку поступающих данных; пользовательский интерфейс, предоставляющий инструменты для визуализации и анализа результатов детектирования. Интеграция этих компонентов создает единую экосистему, способную функционировать как автономно, так и в составе более крупных систем экологического мониторинга.

Для хранения данных используется реляционная база данных PostgreSQL[2] со следующими основными таблицами:

1) users – хранит учётные записи пользователей (id, email, статус активности, статус верификации, признак суперпользователя);
2) AccessToken – хранит токены доступа;
3) devices – информация о фотоловушках (id, название, модель, координаты, статус);
4) images – метаданные фотографий (id, устройство, путь к файлу, дата, статус обработки);
5) animal_classes – справочник классов животных;
6) detection_results – результаты детектирования (id, изображение, класс, уверенность, координаты рамки).

Оригинальные и обработанные изображения хранятся в S3-совместимом хранилище с именами вида:

- original_{uuid_image}.jpg – для оригиналов;
- processed_{uuid_image}.jpg – для изображений с распознанными объектами.

Модуль взаимодействия с фотоловушками обеспечивает автоматический сбор и обработку данных в следующем порядке: приём данных осуществляется через систему, где фотоловушка отправляет на почтовый ящик письмо с изображением, временем съёмки и идентификатором устройства; обработка письма, включающая мониторинг почты сервером, идентификацию писем от фотоловушек, извлечение вложений и метаданных; детектирование и сохранение, включающее отправку изображения в модуль YOLOv11 для детектирования, сохранение оригинала в S3-хранилище, сохранение результатов детектирования в базе данных и сохранение изображения с выделенными объектами в S3-хранилище. Этот процесс происходит автоматически после получения письма, обеспечивая

минимальное участие человека в обработке данных с фотоловушек и значительную оптимизацию процесса.

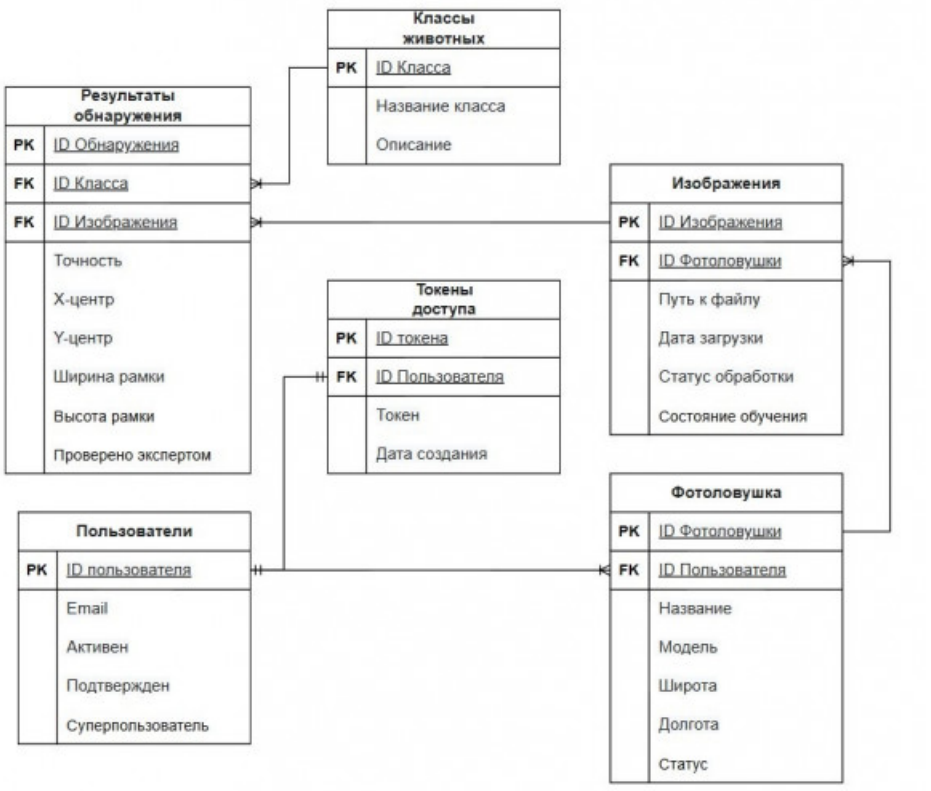


Рис. 3. ER-диаграмма базы данных

Интерфейс сервиса разработан с использованием библиотеки React[3]. В число основных страниц входят:

1) Главная страница – список обработанных изображений с возможностью фильтрации по дате, виду животного и устройству (рис. 4).

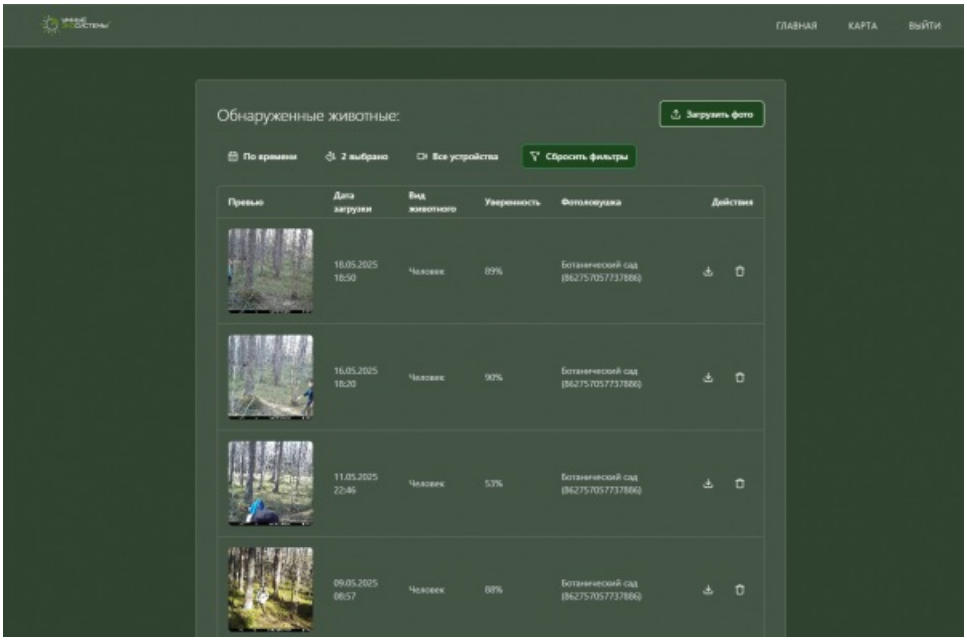


Рис. 4. Интерфейс главной страницы веб-сервиса

2) Страница просмотра изображения – позволяет просмотреть оригинал и обработанное изображение с выделенными объектами (рис. 5).

Рис. 5. Компонент просмотра изображений с результатами детектирования

3) Страница карты – интерактивная карта с размещенными фотоловушками (рис. 6).

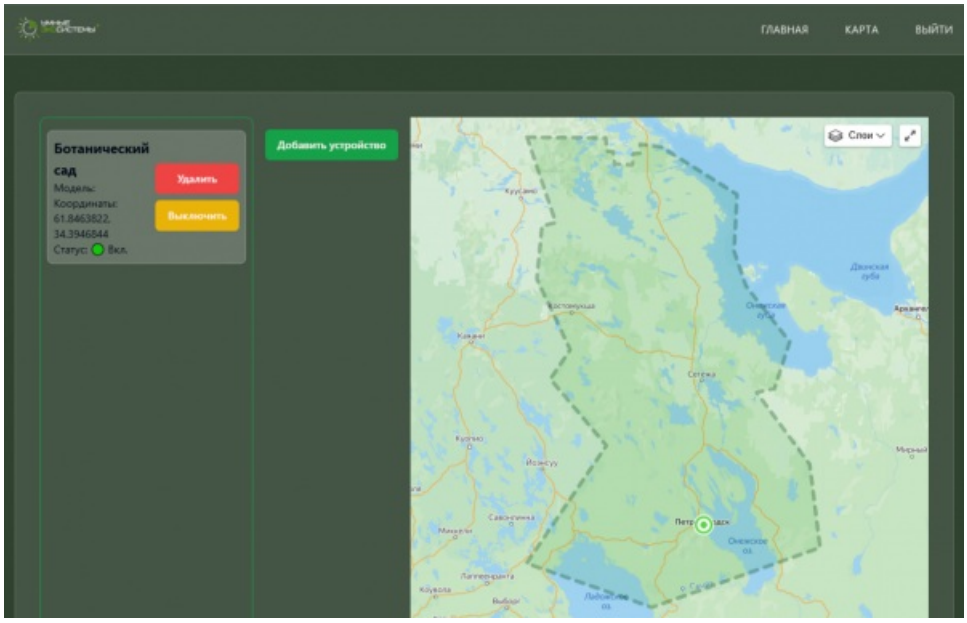


Рис. 6. Страница интерактивной карты с фотоловушками Карелии

Основные возможности пользователей включают просмотр, фильтрацию и скачивание изображений с поддержкой пагинации (по 10 элементов на страницу), управление фотоловушками с возможностью добавления новых устройств, просмотр результатов детектирования с отображением уровня уверенности модели для каждого обнаруженного объекта, и просмотр статистики по обнаруженным животным с группировкой по временным периодам.

Сервис разработан с использованием современных технологий. Серверная часть реализована на FastAPI[4] – асинхронном веб-фреймворке для Python, обеспечивающем быструю обработку запросов и взаимодействие с моделью. Клиентская часть создана с использованием React – библиотеки для создания пользовательских интерфейсов с компонентным подходом. Для хранения данных используется PostgreSQL – надежная СУБД для хранения метаданных и результатов детектирования. Nginx[5] выполняет функции высокопроизводительного веб-сервера и обратного прокси. Docker[6] обеспечивает контейнеризацию приложения, упрощая развертывание. Основой системы детектирования является нейросетевая модель YOLOv11-Large на PyTorch с использованием библиотеки Ultralytics.

* * * * *

В результате выполненного исследования были решены все поставленные задачи. Проведен анализ современных архитектур нейронных сетей для детектирования объектов, показавший преимущества семейства YOLO для задач детектирования в реальном времени. Сформирован специализированный датасет, включающий 6706 изображений восьми классов животных, характерных для фауны Карелии. Проведена качественная разметка данных с использованием современных инструментов аннотации, и применены техники аугментации для повышения качества обучающей выборки.

Обученная модель YOLOv11 продемонстрировала высокие показатели эффективности с точностью детектирования mAP@0.5=0.966, что свидетельствует о способности системы точно определять и классифицировать животных на изображениях.

Разработан и внедрен полнофункциональный веб-сервис с удобным пользовательским интерфейсом, обеспечивающий автоматический прием, обработку и анализ данных, поступающих с фотоловушек. Успешно осуществлена интеграция сервиса с платформой «Умные экосистемы. Мониторинг 4.0», что расширяет возможности системы мониторинга биоразнообразия.

Создана комплексная система хранения данных, включающая реляционную базу данных и объектное хранилище для изображений, что обеспечивает надежность и масштабируемость решения.

Перспективы развития проекта включают дообучение модели на расширенном наборе данных, который будет собираться в ходе дальнейшей работы в Ботаническом саду ПетрГУ.

## Примечания

[1] YOLOv11: руководство пользователя // Ultralytics: сайт. URL: https://docs.ultralytics.com/ru/models/yolo11 (дата обращения: 30.05.2025).

[2] PostgreSQL: Documentation // PostgreSQL: website. URL: https://www.postgresql.org/docs/ (дата обращения: 30.05.2025).

[3] React: The library for web and native user interfaces // React : website. URL: https://react.dev/ (дата обращения: 30.05.2025).

[4] FastAPI // GitHub: website. URL: https://github.com/fastapi/fastapi (дата обращения: 31.05.2025).

[5] NGINX documentation // NGINX: website. URL: https://nginx.org/en/docs/ (дата обращения: 31.05.2025).

[6] Docker documentation // Docker Docs: website. URL: https://docs.docker.com/ (дата обращения: 31.05.2025).

## Список литературы

*Бондар Г.А*. Разработка сервиса детектирования животных по фотографиям / Г.А. Бондар, Н.С. Кулезнев, В.М. Димитров // Цифровые технологии в образовании, науке, обществе: мат-лы XVIII

Всеросс. науч.-практ. конф. (Петрозаводск, 3–5 декабря 2024). Петрозаводск: Изд-во ПетрГУ, 2024. С. 18-21.

*Данилов П.И*. Млекопитающие Карелии / П.И. Данилов, В.В. Белкин. Петрозаводск: Карельский научный центр РАН, 2011. 264 с.

*Шапиро Л*. Компьютерное зрение / Л. Шапиро, Д. Стокман. Москва: БИНОМ. Лаборатория знаний, 2013. 752 с.

*LeCun Y*. Deep learning / Y. LeCun, Y. Bengio, G. Hinton // Nature. 2015. Vol. 521, № 7553. P. 436-444. DOI: 10.1038/nature14539

*Norouzzadeh M.S*. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning / M.S. Norouzzadeh, A. Nguyen, M. Kosmala, et al. // Proceedings of the National Academy of Sciences. 2018. Vol. 115, № 25. P. E5716-E5725. DOI: 10.1073/pnas.1719367115

*Redmon J*. You only look once: unified, real-time object detection / J. Redmon, S. Divvala, R. Girshick, A. Farhadi // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016. P. 779-788. DOI: 10.1109/CVPR.2016.91

| | |
|---|---|
| **Grigory A. BONDAR** | bachelor's degree, Petrozavodsk State University (Petrozavodsk, Russia), *g4bondar@gmail.com* |
| **Nikita S. KULEZNEV** | bachelor's degree, Petrozavodsk State University (Petrozavodsk, Russia), *kuleznevnik.34@gmail.com* |

# DEVELOPING A SERVICE FOR DETECTING ANIMALS IN PHOTOGRAPHS

**Scientific adviser:**
Vyacheslav M. Dimitrov
**Reviewer:**
Aleksandr A. Kuzmenkov
Paper submitted on: 06/02/2025;
Accepted on: 06/29/2025;
Published online on: 06/29/2025.

**Abstract.** Automating the detection of animals in camera trap photographs is a significant challenge in ecological research. This study aims to develop a web service for the automatic identification of fauna representatives native to Karelia. The research methodology involved compiling a dataset of 6,706 images of animals, training the YOLOv11 neural network model for object recognition, and designing a web interface integrated with the "Smart Ecosystems. Monitoring 4.0" platform. The results demonstrate the successful implementation of an automated system for processing camera trap data — from image acquisition to web-based visualization — for efficient biodiversity monitoring.
**Keywords:** computer vision, YOLOv11, object detection, ecosystem monitoring, camera traps, animal recognition

**For citation:** Kuleznev, N. S. Developing a Service for Detecting Animals in Photographs. StudArctic Forum. 2025, 10 (2): 40–49.

## References

Bondar G.A., Kuleznev N.S., et al. Developing a service for detecting animals in photographs. *Digital technologies in education, science, and society: Proceedings of the XVIII all-Russian scientific and practical conference (Petrozavodsk, December 3–5, 2024).* Petrozavodsk, Petrozavodsk State University, 2024, pp. 18-21. (In Russ.)

Danilov P.I., Belkin V.V. *Mammals of Karelia*. Petrozavodsk, Karelian Research Centre of the Russian Academy of Sciences, 2011, 264 p. (In Russ.)

Shapiro L., Stockman G. Computer vision. Moscow, BINOM. Laboratoriya znaniy, 2013, 752 p. (In Russ.)

LeCun Y., Bengio Y., et al. Deep learning. *Nature*, 2015, Vol. 521, No. 7553, pp. 436-444. DOI: 10.1038/nature14539

Norouzzadeh M.S., Nguyen A., et al. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences* , 2018, Vol. 115, No. 25, pp. E5716-E5725. DOI: 10.1073/pnas.1719367115

Redmon J., Divvala S., et al. You only look once: unified, real-time object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Institute of Electrical and Electronics Engineers, 2016, pp. 779-788. DOI: 10.1109/CVPR.2016.91