

Проблемы искусственного интеллекта. 2026. N 1 (40). С. 50-62
Problems of Artificial Intelligence. 2026;1(40):50-62.
Искусственный интеллект и машинное обучение
Научная статья

УДК 004.932.2

doi: 10.24412/2413-7383-2026-1-40-50-62

В. И. Луйкин, Д. В. Лобов, Д. В. Логинов
Федеральное государственное образовательное учреждение
высшего образования «Петрозаводский государственный университет», г. Петрозаводск
185910, г. Петрозаводск, ул. Ленина, 33

РАСПОЗНАВАНИЕ СТВОЛОВ ДЕРЕВЬЕВ С ИСПОЛЬЗОВАНИЕМ ЧЕТЫРЕХ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ

V. I. Lyikin, D. V. Lobov, D. V. Loginov
Federal State Educational Institution of Higher Education «Petrozavodsk state university»
185910, Petrozavodsk, Lenina str, 33

TREE TRUNK RECOGNITION USING FOUR MACHINE LEARNING MODELS

Данная статья посвящена исследованию применимости современных архитектур нейронных сетей (Faster R-CNN, YOLOv11x, DETR, RetinaNet) для автоматического распознавания стволов деревьев в условиях лесов Республики Карелия. В ходе работы собран и размечен специализированный датасет, сочетающий открытые данные и собственные изображения, адаптированные под специфику северных лесных экосистем. Экспериментально установлено, что модель Faster R-CNN демонстрирует наивысшую точность детекции ($mAP@0.5 = 0.659$), в то время как YOLOv11x обеспечивает оптимальный баланс между точностью и скоростью обработки (15.33 FPS), что критически важно для систем реального времени. Результаты исследования подтверждают практическую возможность автоматизации задач инвентаризации и мониторинга лесов с использованием методов компьютерного зрения, указывая на необходимость дальнейшего расширения датасетов для улучшения обобщающей способности моделей.

Ключевые слова: искусственный интеллект, машинное зрение, распознавание объектов, модели машинного обучения, лесной комплекс.

This article investigates the applicability of modern neural network architectures (Faster R-CNN, YOLOv11x, DETR, RetinaNet) for the automatic detection of tree trunks in the forest conditions of the Republic of Karelia. A specialized dataset was compiled and annotated for the study, combining open data and original images adapted to the specifics of northern forest ecosystems. Experiments established that the Faster R-CNN model demonstrates the highest detection accuracy ($mAP@0.5 = 0.659$), while YOLOv11x provides the optimal balance between accuracy and processing speed (15.33 FPS), which is critical for real-time systems. The results confirm the practical feasibility of automating forest inventory and monitoring tasks using computer vision methods and indicate the need for further dataset expansion to improve the models' generalization capability.

Key words: artificial intelligence, machine vision, object recognition, machine learning models, forestry complex.

Введение

Автоматизированная обработка изображений находит широкое применение в различных сферах, таких как промышленность, транспорт, безопасность и, в частности, лесное хозяйство. Среди ключевых задач компьютерного зрения особое место занимает распознавание объектов, которое сегодня активно используется как в научной, так и в прикладной деятельности. Одним из перспективных направлений в этой области является автоматическая идентификация деревьев по их визуальным признакам.

Разработка систем автоматического и автоматизированного распознавания деревьев имеет особую важность для лесных территорий России, включая Республику Карелия, где лесная отрасль играет ключевую роль в экономике [1]. Огромные площади лесных массивов, трудоёмкость их учёта и необходимость постоянного мониторинга обуславливают потребность в технологиях, способных заменить ручной труд. Применение методов компьютерного зрения и машинного обучения для детектирования деревьев может значительно изменить процессы лесозаготовки и инвентаризации. Задача распознавания деревьев по визуальным характеристикам также становится востребованной в связи с развитием беспилотных технологий. В условиях сложного ландшафта и труднодоступных территорий автономные платформы — такие как дроны или роботизированная техника — могут использовать алгоритмы компьютерного зрения для навигации и принятия решений в реальном времени. Например, подобные системы могут быть интегрированы в лесозаготовительные машины, позволяя им автоматически определять деревья для вырубки, прокладывать оптимальные маршруты и снижать риски для человека при выполнении опасных операций.

Кроме того, такие технологии открывают возможности для решения других задач в сфере лесного хозяйства — от мониторинга ресурсов до оперативного отслеживания изменений в лесных экосистемах. Накопление и анализ визуальных данных позволяют создавать обучаемые системы, способные адаптироваться к новым условиям. Это особенно важно при масштабировании технологий на регионы с разным видовым составом деревьев и особенностями рельефа.

Целью данной работы являлась проверка применимости современных моделей нейросетевых архитектур для распознавания стволов деревьев на основе их визуальных характеристик в лесистом регионе России – Карелии (леса занимают более 50% территории).

Анализ текущих исследований в области компьютерного зрения для лесоведения

Впервые в статье 2016 г. [2] нами была сделана попытка применения классических методов компьютерного зрения для автоматического выявления стволов деревьев на статических снимках и в видеопотоке. В отличие от нейросетевых подходов, предлагаемое решение было основано на возможностях, предоставляемых библиотекой OpenCV без предварительного обучения на аннотированных данных. В работе использовались следующие алгоритмы для выделения границ объектов: фильтр Собеля, детектор Канны и поиск по шаблону. В результате проведенных экспериментов выяснилось, что общий коэффициент вероятности детекции (аналог метрики Recall - «полнота») стволов деревьев не превосходил 70%. В данной работе не использовались методы нейросетевых технологий.

За последние десять лет возможности детектирования таких сложных биологических объектов, как стволы деревьев в реальной экосистеме – лесу, значительно усилились за счет появления и масштабного внедрения нейросетевых технологий.

В области автоматической идентификации деревьев также появилось множество научных работ [3], [4], [6-11]. В статье [4] демонстрируется эффективность комбинации многомасштабной (32×32 и 64×64 пикселей) сверточной нейронной сети (CNN) для детектирования масляничных пальм на аэрофотоснимках (вид кроны деревьев сверху). Датасет представлял собой небольшой по размерам набор изображений (больше 700 изображений) с качественно проведенной разметкой. За счет строгой постобработки (порог уверенности больше 0.70) и комбинации двух масштабов CNN удалось добиться значения общей точности, равной 95%. Надо только иметь в виду, что такие высокие результаты были получены за счет того, что анализ проводился с видом сверху и рассматривался простой тип экосистемы (искусственная сельскохозяйственная система – однотипные кроны деревьев). В статье [5] проводились аналогичные эксперименты с кронами масляничных пальм. В качестве основного алгоритма детектирования выступал Faster R-CNN. Была достигнута общая точность (OA) не менее 97% и величина метрики «полнота» (Recall) – около 98%. Также производилось сравнение этого метода с классическими методами «искусственная нейронная сеть» (ANN) и «метод опорных векторов» (SVM). Алгоритм Faster R-CNN значительно превзошел классические методы. Статья [3] посвящена детектированию стволов деревьев и характеризуется крупным, сгенерированным игровым «движком», датасетом (43000 синтетических изображений). В качестве основного алгоритма выступал Mask R-CNN. Был получен очень высокий результат детекции деревьев — средняя точность (mAP@50) около 90%. Однако, в статье указано, что такой высокий результат был получен только для конкретного лесного массива (при тестировании на данных из Португалии фиксировалось снижение точности на 20-40%). В статье [6] проводился анализ классификации видов деревьев в естественных вторичных лесах (точное определение границ кроны отдельных деревьев и их видов). В качестве основных методов использовались виды CNN (1D-CNN, 2D-CNN, 3D-CNN). Показано преимущество 1D-CNN перед 2D/3D-CNN для классификации в смешанных лесах. Сочетание гиперспектральных, лидарных и наземных данных позволило достичь высокой общей точности классификации (OA = 83%).

В отечественной науке также присутствуют статьи с результатами исследований в рассматриваемой области. Например, в статье [7] использовалась сверточная нейронная сеть (CNN) на базе архитектур ResNet и VGG для распознавания и классификации пород деревьев по спектральным изображениям стволов, собранным при наземной лидарной таксации. Набор данных (датасет) состоял из небольшого числа спектральных изображений стволов деревьев четырех пород. В ходе работы было создано приложение для аугментации данных и улучшения точности модели. По результатам работы была достигнута точность более 80% (метрика Accuracy - «доля верных предсказаний»). В работе [8] использовалась библиотека TensorFlow для классификации трех пород деревьев (берёза, ель, сосна) в реальном времени во время лесозаготовок. Датасет включал 24000 цветных изображений с низким разрешением стволов деревьев, зафиксированных камерой в кабине лесозаготовительной техники. Производилась ручная разметка изображений для обучения. В качестве моделей нейронной сети выступали полносвязная нейронная сеть (вероятно, многослойный перцептрон) и метод «случайный лес» (Random Forest). В результате хода работы были получены следующие результаты: точность 62% для полносвязной сети и 90% для модели «случайный лес», т.е. доказана эффективность второго метода для классификации деревьев. Выяснилось, что вариативность данных (все больший размер датасета) уменьшала общую точность, а для второго метода был найден оптимальный размер выборки — более 11000 изображений.

Все эти исследования подтверждают, что глубокое обучение — оптимальный инструмент для автоматизации детектирования деревьев с возможной достигаемой средней точностью от ~70%. При этом он требует больших размеченных наборов данных (датасетов) и страдает от проблем обобщаемости.

Методика эксперимента (методы, формирование набора данных)

Методы

В данной работе были использованы четыре модели нейронных сетей: Faster R-CNN, DETR (Detection Transformer), YOLO и RetinaNet [12-18]. Критерии отбора рассматриваемых моделей включали не только их характеристики производительности, востребованность в области машинного зрения, но и принципиальные различия в архитектурной организации, что обеспечивало методологическую корректность и обобщаемость выводов при проведении сравнительной оценки.

Faster R-CNN — это современная архитектура свёрточной нейронной сети, предназначенная для обнаружения и классификации объектов на изображениях. В отличие от своих предшественников (R-CNN и Fast R-CNN), эта модель объединяет процессы генерации регионов-кандидатов и их классификации в единую сквозную систему. Основная инновация Faster R-CNN заключается в использовании Region Proposal Network (RPN) — специального модуля, который предсказывает потенциальные области с объектами прямо на картах признаков, извлечённых базовой CNN.

Алгоритм данной модели делится на два этапа:

1. Этап поиска областей, где могут быть искомые объекты (Region Proposal Network). На данном этапе применяется свёрточная нейронная сеть (CNN) для выделения (границы) простых и сложных признаков (формы) на изображениях и оценивание правильности совпадения с объектом с помощью RPN.

2. Этап уточнения характеристик объектов. На этом этапе происходит окончательное уточнение границ объектов и их классификация.

Модель детектирования DETR в отличие от Faster R-CNN, использующего сложные конвейеры с генерацией регионов-кандидатов и последующей их обработкой, предлагает «end-to-end» решение на основе механизмов внимания. Эта модель устраняет необходимость в ручных этапах обработки.

Алгоритм данной модели:

1. Извлечение признаков с помощью CNN.
2. Применение трансформера-энкодера для анализа глобального контекста и позиционного кодирования. Результат — решение, какие объекты присутствуют.

3. Применение трансформера-декодера для классификации и выявления границ объектов на основе обучаемых запросов.

4. Применение алгоритма бинарного сопоставления для поиска наилучшего соответствия предсказанных объектов реальным.

5. Фильтрация значимых объектов.

Модель YOLO является быстрым одноэтапным детектором объектов. Анализ в этой модели происходит над всем изображением сразу.

Алгоритм данной модели:

1. Извлечение признаков.

2. Разбиение изображения на ячейки.

3. Предсказание границ объектов и их классификация.

4. Мультимасштабная детекция для поиска объектов любых размеров

5. Фильтрация.

Метрики

В данном исследовании учитывались следующие показатели:

- mAP (mean Average Precision) — метрика, характеризующая общую точность модели;
- IoU (Intersection over Union) — мера соответствия предсказанных и истинных областей объекта, отражающая качество локализации;
- FPS (Frames Per Second) — количество обрабатываемых кадров в секунду; характеристика скорости работы модели, важная при её практическом применении в реальном времени.

Выбор данных метрик определялся тремя основными показателями:

1. Стандартизация: поддержка ведущими фреймворками (TensorFlow, PyTorch, MMDetection, Ultralytics).
2. Сравнимость: объективное сопоставление разных моделей на одних и тех же датасетах.
3. Комплексность: охват точности детекции моделей и их производительности.

Датасет

Для проведения экспериментов был подготовлен собственный набор данных (датасет) для обучения моделей детектирования деревьев в естественной среде, с акцентом на условия северных регионов, таких как Республика Карелия.

Готовые датасеты были проверены и выяснилось, что они не подходят из-за различий в условиях (южные регионы и северные леса) и проблем с качеством разметки. Решено было создать собственный датасет, сочетающий готовые данные и изображения, адаптированные под локальные условия. В качестве источников данных выступали:

- Базовый датасет: Tree Trunks Dataset с платформы Roboflow (663 изображения);
- Собственноручно созданные фотографии (103 изображения) и изображения из интернет (70 изображений) региона Карелии.

Разметка была исправлена и дополнена вручную с использованием CVAT (Computer Vision Annotation Tool). Изображения были приведены к формату JPEG и разрешению 640×640 пикселей. Процесс аугментации (изменение яркости, контраста, наложение шума и др.) увеличил обучающую выборку до 1434 изображений. Тестовая выборка включала 50 изображений.

Выбор библиотек для детекции деревьев

Для решения задачи распознавания деревьев были отобраны современные библиотеки (фреймворки) компьютерного зрения. В исследовании использовались следующие модели из пакета torchvision (фреймворк PyTorch): Faster R-CNN (fasterrcnn_resnet50_fpn) и RetinaNet (retinanet_resnet50_fpn).

Оптимизированные реализации семейства моделей YOLO были взяты из библиотеки Ultralytics. Модель DETR была реализована напрямую из оригинального репозитория FB Research.

Результаты

В рамках данного исследования все вычисления для обучения проводились на графическом процессоре NVIDIA Tesla T4, который доступен в облачной среде Google Colab [20]. Тестирование моделей проводилось на центральном процессоре CPU AMD

Ryzen 5 4600H. Для всех экспериментов обеспечивались одинаковые условия для объективности и воспроизводимости результатов. Тестирование проводилось на едином наборе данных, включающем: разную плотность деревьев; сложные фоны; перекрытия объектов; изменения освещения. Метрики рассчитывались с использованием: COCO API (pyscocotools) для mAP и AR; функций для Mean IoU и FPS.

Таблица 1 – Результаты экспериментов детектирования стволов деревьев

Метрика	Faster R-CNN	YOLOv11x	DETR	RetinaNet
Average Precision (AP) @ [IoU=0.50 area=all maxDets=100]	0.659	0.636	0.621	0.562
Mean IoU	0.530	0.512	0.504	0.472
FPS	2.840	15.330	4.100	11.140

В таблице 1 представлены ключевые метрики: точность $mAP@[0.5]$, показатель качества локализации объектов Mean IoU и метрика производительности FPS. Модель Faster R-CNN показала лучший результат по метрике $mAP@[0.5]$ — значение 66%. Модели YOLOv11x и DETR показывают сравнимые результаты, а RetinaNet специализирована на мелких объектах, т.к. показала наилучший уровень $mAP_{small} = 0.47$. Метрика IoU измеряет точность определения границ объектов через соотношение площади пересечения и объединения предсказанных и истинных границ. Из результатов видно, что объекты в лесных экосистемах сложны для гарантированного детектирования. Модели показали значение $mAP@0.5$ больше 60%, что соответствует промышленным стандартам для подобных задач. Однако, для гарантированного детектирования в реальных условиях необходимо: увеличить датасет (особенно для сложных случаев); найти оптимальный порог уверенности модели (confidence threshold).

Значение Mean IoU — среднее значение точности позиционирования для всех корректных детектирований для объективного сравнения моделей независимо от масштаба объектов. Faster R-CNN продемонстрировала наилучший результат (0.530), что подтверждает преимущества её двухэтапной архитектуры. Высокий показатель свидетельствует о стабильно точном позиционировании объектов в различных условиях. Модели YOLOv11x и DETR показали близкий результат. При сохранении высокой точности локализации они обеспечивают существенно более высокую скорость обработки, что критически важно для практического применения.

Скорость обработки данных — критически важный параметр при выборе модели для систем реального времени, таких как автономные дроны, мобильные роботы, лесозаготовительные машины, работающие в лесных массивах. Для измерения производительности использовалась метрика FPS. Модель YOLOv11x (15.33 FPS) продемонстрировала наилучшие показатели скорости, что делает её идеальным выбором для задач реального времени. Это объясняется эффективной одноэтапной архитектурой, изначально разработанной для высокопроизводительных решений. Остальные модели значительно отстают в быстродействии. Модель DETR (4.10 FPS) показала ожидаемо более низкие результаты, характерные для трансформерных архитектур, требующих значительных вычислительных ресурсов. Модель Faster R-CNN (2.84 FPS) оказалась наименее производительной из-за своей двухэтапной архитектуры, что ограничивает её применение в системах реального времени, несмотря на высокую точность детекции.

Таблица 2 - Результаты работы моделей при различных входных разрешениях

Размер изображения	mAP@0.5	FPS
YOLOv11x		
320x320	0.503 (-21.0%)	37.20 (+142.7%)
640x640	0.636	15.33
1024x1024	0.599 (-5.8%)	4.29 (-72%)
Faster R-CNN		
320x320	0.580 (-11.9%)	4.28 (+50.7%)
640x640	0.659	2.84
1024x1024	0.632 (-4.1%)	1.1 (-61.3%)
DETR		
320x320	0.409 (-34.1%)	9.37 (+128.5%)
640x640	0.621	4.10
1024x1024	0.628 (+1.1%)	1.40 (-65.85%)
RetinaNet		
320x320	0.453 (-19.4%)	28.3 (+154%)
640x640	0.562	11.14
1024x1024	0.508 (-9.6%)	4.83 (-56.6%)

В реальных условиях работы систем компьютерного зрения часто требуется обработка изображений с различным разрешением — как для оптимизации вычислительных ресурсов, так и для повышения детализации. Поведение четырех моделей оценивалось при трех разрешениях: 320×320, 640×640 и 1024×1024 пикселей. В таблице 2 представлен результат работы четырех рассматриваемых моделей.

Из таблицы видно, что для модели YOLOv11x оптимальным разрешением является 640×640. При уменьшении разрешения скорость детекции растет более чем в два раза, но при этом значительно уменьшается точность (на 20%). Увеличение размера изображения приводит к деградации скорости расчета. Модель Faster R-CNN показывает наименьшую чувствительность к уменьшению/увеличению разрешения, сохраняя стабильность значений метрик. Для модели DETR увеличение разрешения не приводит к изменению точности, а уменьшение разрешения значительно деградирует точность детекции. Для модели RetinaNet уменьшение разрешения дает максимальный прирост скорости расчета.

На основании результатов экспериментов можно сказать, что выбор модели должен основываться на конкретных требованиях. Если требуется стабильность в качестве детектирования, то выбор будет в пользу модели Faster R-CNN; если необходима скорость, то модель YOLOv11x; если необходима мелкая детализация, то подойдет модель DETR; модель RetinaNet является неким компромиссным решением.

На рисунке 1 представлено отношение более высокого значения mAP - точности (mAP@[0.5:0.95]) и производительности (FPS) для рассмотренных моделей.

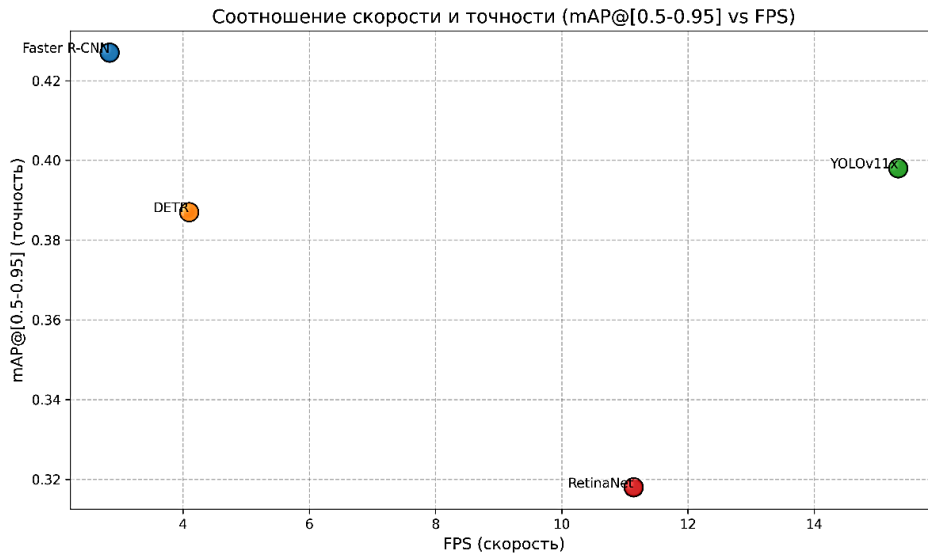


Рисунок 1 – Отношение точности (mAP@[0.5:0.95]) и производительности (FPS) для рассмотренных моделей.

Анализ диаграммы отчетливо показывает лидерство модели YOLOv11x по критерию оптимального баланса между точностью (достоверностью детекции объектов) и скоростью обработки (способностью работать в режиме реального времени).

На рисунке 2 и рисунке 3 показаны реальные результаты детекции стволов деревьев, характерных для лесов региона Карелии.

Faster R-CNN



DETR



YOLOv11x



RetinaNet



Рисунок 2 – Детектирование стволов деревьев летом

Faster R-CNN



DETR



YOLOv11x



RetinaNet



Рисунок 3 – Детектирование стволов деревьев зимой.

Анализ визуальных результатов предсказаний моделей выявил следующие ключевые особенности:

- Модель Faster R-CNN продемонстрировала: идеальную детекцию в ясную погоду и снегопад; точное определение границ даже при перекрытии объектов; незначительное ухудшение качества при плохом освещении.
- Модели DETR и YOLOv11x показали сопоставимые результаты: аналогичное качество определения границ объектов; схожие значения показателя уверенности.
- Модель RetinaNet подтвердила соответствие низким метрическим оценкам: размытые границы объектов; низкие значения уверенности детекции.

Заключение

На основе всех проведенных экспериментов можно сделать вывод оптимальным решением для задачи распознавания деревьев является одноэтапная модель YOLOv11x благодаря:

- Сбалансированной производительности: высокая скорость обработки (15.33 FPS) при достаточной точности ($mAP@0.5=0.636$);
- Практическим преимуществам: реализация работы в реальном времени; устойчивость к различным условиям съемки; эффективное использование ресурсов;
- Техническому превосходству: лучший баланс характеристик среди тестируемых моделей; возможность интеграции в мобильные системы.

В качестве альтернативного решения можно использовать модель Faster R-CNN для тех задач, где требуется максимальная точность и аппаратные средства имеют такое быстроедействие, что скорость обработки не будет являться критичным фактором.

Также были выявлены следующие проблемы детекции при перекрытии объектов: склонность моделей к объединению близко расположенных деревьев; ошибки сегментации при значительном перекрытии контуров выделенных границ объектов; ошибки классификации (ложное распознавание вертикальных конструкций (столбы, мачты) как деревьев, недостаточная селективность при наличии схожих по форме объектов).

Полученные результаты подтверждают аналогичные результаты в научной литературе и дают возможности для совершенствования методов компьютерного зрения в прикладных экологических и лесопромышленных задачах.

Список литературы

1. Шишкин А. И., Костюкевич В. В. Роль лесного комплекса в экономике республики Карелии // Экономика и управление. 2012. №5 (79). URL: <https://cyberleninka.ru/article/n/rol-lesnogo-kompleksa-v-ekonomike-respubliki-karelii>
2. Аблоухов С. И., Егупов Р. Д., Лобов Д. В. Детектирование стволов деревьев на основе алгоритмов библиотеки OpenCV // Инновационная наука. 2016. №5-3 (17). URL: <https://cyberleninka.ru/article/n/detektirovanie-stvolov-dereviev-na-osnove-algoritmov-biblioteki-opencv>
3. Vincent Grondin et al. Tree detection and diameter estimation based on deep learning // Forestry: An International Journal of Forest Research, Volume 96, Issue 2, April 2023, Pages 264–276, URL: <https://doi.org/10.1093/forestry/cpac043>
4. Maciel Zortea et al. Oil-Palm Tree Detection in Aerial Images Combining Deep Learning Classifiers // 2018 IEEE International Geoscience and Remote Sensing Symposium, Spain, pp. 657-660, URL: <https://doi.org/10.1109/IGARSS.2018.8519239>
5. Liu, X., Ghazali, K. H., Han, F., & Mohamed, I. I. 2020. Automatic Detection of Oil Palm Tree from UAV Images Based on the Deep Learning Method. Applied Artificial Intelligence, 35(1), 13–24. URL: <https://doi.org/10.1080/08839514.2020.1831226>
6. Lee, H.Y., Ho, H.W. & Zhou, Y. Deep Learning-based Monocular Obstacle Avoidance for Unmanned Aerial Vehicle Navigation in Tree Plantations // J Intell Robot Syst 101, 2021, URL: <https://doi.org/10.1007/s10846-020-01284-z>
7. Савченко Г. А., Забелина В. А., Терехов В. И. Определение породного состава насаждений по изображению стволов с последующей аугментацией данных // ИИАСУ'22. Искусственный интеллект в автоматизированных системах управления и обработки данных: сборник статей Всероссийской научной конференции, Москва, 27-28 апреля 2022 г. : в 2 т. / МГТУ им. Н. Э. Баумана (национальный исследовательский у-т). – М. : Изд-во МГТУ им. Н. Э. Баумана, 2022. – ISBN 978-5-7038-5910-0. Т. 1. 2022. С. 359-364 URL: <https://www.elibrary.ru/item.asp?id=49219364>
8. Жук К.Д., Угрюмов С.А., Свойкин Ф.В. Распознавание пород деревьев в процессе лесозаготовки с применением методов машинного обучения // Известия Санкт-Петербургской лесотехнической академии, 2023. Вып. 242. С. 167-178. URL: <https://www.elibrary.ru/item.asp?id=50504403>
9. Говядин, И. К. Метод измерения диаметров деревьев на основе технологий искусственного интеллекта / И. К. Говядин, А. Н. Чубинский, А. С. Алексеев // Известия Санкт-Петербургской лесотехнической академии. 2024. – № 249. – С. 177-194. – DOI 10.21266/2079-4304.2024.249.177-194. – EDN RAHSSN.
10. Каримов, Б. М. Искусственный интеллект в лесной отрасли: возможности и перспективы / Б. М. Каримов, И. К. Говядин // Современные тренды управления и цифровая экономика: от регионального развития к глобальному экономическому росту : сборник статей V Международной научно-практической конференции, Екатеринбург, 10 августа 2023 года / Институт цифровой экономики и права. – Екатеринбург: Общество с ограниченной ответственностью "Институт Цифровой Экономики и Права", 2023. С. 334-338. EDN ZNKGZM.
11. Жук, К. Д. Распознавание пород деревьев в процессе лесозаготовки с применением методов машинного обучения / К. Д. Жук, С. А. Угрюмов, Ф. В. Свойкин // Известия Санкт-Петербургской лесотехнической академии. 2023. № 242. С. 167-178. DOI 10.21266/2079-4304.2023.242.167-178. – EDN RZMZUR.
12. S. Ren, K. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. // S. Ren, K. He, R. Girshick, J. Sun / Computer Vision and Pattern Recognition, 2015 URL: <https://doi.org/10.48550/arXiv.1506.01497>

13. PyTorch (Torchvision) Faster-RCNN Documentation: This is the standard, actively maintained implementation in the PyTorch ecosystem. URL: http://docs.pytorch.org/vision/main/models/faster_rcnn.html
14. R. Khanam: YOLOv11: An Overview of the Key Architectural Enhancements // R. Khanam, M. Hussain / Computer Vision and Pattern Recognition, 2024. URL: <https://doi.org/10.48550/arXiv.2410.17725>
15. YOLOv11x Documentation. URL: <https://docs.ultralytics.com/models/yolo11/>
16. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. 2020. End-to-End Object Detection with Transformers. URL: <https://doi.org/10.48550/arXiv.2005.12872>
17. DETR Documentation: URL: https://huggingface.co/docs/transformers/model_doc/detr
18. Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2018). Focal Loss for Dense Object Detection. URL: <https://doi.org/10.48550/arXiv.1708.02002>
19. PyTorch (Torchvision) RetinaNet Documentation: URL: <https://docs.pytorch.org/vision/main/models/retinanet.html>
20. NVIDIA Tesla T4 Tensor Core GPU Product Brief // NVIDIA. [Электронный ресурс]. URL: <https://www.nvidia.cn/content/dam/en-zz/Solutions/Data-Center/tesla-t4/t4-tensor-core-product-brief.pdf>

References

1. Shishkin A. I., Kostyukevich V. V. The Role of the Forestry Complex in the Economy of the Republic of Karelia // *Economy and Management*. 2012. No. 5 (79). URL: <https://cyberleninka.ru/article/n/rol-lesnogo-kompleksa-v-ekonomike-respubliki-karelii>
2. Abloukhov S. I., Egupov R. D., Lobov D. V. Tree Trunk Detection Based on OpenCV Library Algorithms // *Innovative Science*. 2016. No. 5-3 (17). URL: <https://cyberleninka.ru/article/n/detektirovanie-stvolov-dereviev-na-osnove-algoritmov-biblioteki-opencv>
3. Vincent Grondin et al. Tree detection and diameter estimation based on deep learning // *Forestry: An International Journal of Forest Research*, Volume 96, Issue 2, April 2023, Pages 264–276, URL: <https://doi.org/10.1093/forestry/cpac043>
4. Maciel Zortea et al. Oil-Palm Tree Detection in Aerial Images Combining Deep Learning Classifiers // 2018 IEEE International Geoscience and Remote Sensing Symposium, Spain, pp. 657-660, URL: <https://doi.org/10.1109/IGARSS.2018.8519239>
5. Liu, X., Ghazali, K. H., Han, F., & Mohamed, I. I. 2020. Automatic Detection of Oil Palm Tree from UAV Images Based on the Deep Learning Method. *Applied Artificial Intelligence*, 35(1), 13–24. Available at: <https://doi.org/10.1080/08839514.2020.1831226>
6. Lee, H.Y., Ho, H.W. & Zhou, Y. Deep Learning-based Monocular Obstacle Avoidance for Unmanned Aerial Vehicle Navigation in Tree Plantations // *J Intell Robot Syst* 101, 2021, Available at: <https://doi.org/10.1007/s10846-020-01284-z>
7. Savchenko, G.A., Zabelina, V.A., Terekhov, V.I. Determination of tree species composition from trunk images with subsequent data augmentation // *IIASU'22. Artificial Intelligence in Automated Data Control and Processing Systems: Collection of articles from the All-Russian Scientific Conference, Moscow, April 27-28, 2022: in 2 volumes / Bauman Moscow State Technical University (National Research University). – Moscow: Bauman Moscow State Technical University Publishing House, 2022. – ISBN 978-5-7038-5910-0. Vol. 1. – 2022. Pp. 359-364 URL: <https://www.elibrary.ru/item.asp?id=49219364>*
8. Zhuk K.D., Ugryumov S.A., Svoykin F.V. Recognition of tree species during logging using machine learning methods // *Bulletin of the St. Petersburg Forest Engineering Academy*, 2023. – Issue. 242. – P. 167-178. URL: <https://www.elibrary.ru/item.asp?id=50504403>
9. Govyadin, I. K. Method for measuring tree diameters based on artificial intelligence technologies / I. K. Govyadin, A. N. Chubinsky, A. S. Alekseev // *Bulletin of the St. Petersburg Forest Engineering Academy*. - 2024. - No. 249. - P. 177-194. - DOI 10.21266/2079-4304.2024.249.177-194. - EDN RAHSSN.
10. Karimov, B. M. Artificial Intelligence in the Forestry Industry: Opportunities and Prospects / B. M. Karimov, I. K. Govyadin // *Modern Management Trends and the Digital Economy: From Regional Development to Global Economic Growth: Collection of Articles from the V International Scientific and Practical Conference, Yekaterinburg, August 10, 2023 / Institute of Digital Economy and Law. - Yekaterinburg: Limited Liability Company "Institute of Digital Economy and Law", 2023. - Pp. 334-338. - EDN ZNKGZM.*

11. Zhuk, K. D. Recognition of Tree Species in the Logging Process Using Machine Learning Methods / K. D. Zhuk, S. A. Ugryumov, F. V. Svoykin // Bulletin of the St. Petersburg Forest Engineering Academy. - 2023. - No. 242. - Pp. 167-178. – DOI 10.21266/2079-4304.2023.242.167-178. – EDN RZMZUR.
12. S. Ren, K. He, R. Girshick, J. Sun / Computer Vision and Pattern Recognition, 2015 URL: <https://doi.org/10.48550/arXiv.1506.01497>
13. PyTorch (Torchvision) Faster-RCNN Documentation: This is the standard, actively maintained implementation in the PyTorch ecosystem. URL: http://docs.pytorch.org/vision/main/models/faster_rcnn.html
14. R. Khanam: YOLOv11: An Overview of the Key Architectural Enhancements // R. Khanam, M. Hussain / Computer Vision and Pattern Recognition, 2024. URL: <https://doi.org/10.48550/arXiv.2410.17725>
15. YOLOv11x Documentation. URL: <https://docs.ultralytics.com/models/yolo11/>
16. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. 2020. End-to-End Object Detection with Transformers. URL: <https://doi.org/10.48550/arXiv.2005.12872>
17. DETR Documentation: URL: https://huggingface.co/docs/transformers/model_doc/detr
18. Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2018). Focal Loss for Dense Object Detection. URL: <https://doi.org/10.48550/arXiv.1708.02002>
19. PyTorch (Torchvision) RetinaNet Documentation: URL: <https://docs.pytorch.org/vision/main/models/retinanet.html>
20. NVIDIA Tesla T4 Tensor Core GPU Product Brief // NVIDIA. – [Electronic resource]. – URL: <https://www.nvidia.cn/content/dam/en-zz/Solutions/Data-Center/tesla-t4/t4-tensor-core-product-brief.pdf>

RESUME

V. I. Lyikin, D. V. Lobov, D. V. Loginov

Tree trunk recognition Using four machine learning models

The development of automated tree recognition systems is particularly important for Russia's forested areas. The vast expanses of forested areas, the labor-intensive nature of their inventory, and the need for continuous monitoring necessitate technologies that can replace manual labor. Recognizing trees by visual characteristics is also becoming increasingly popular due to the development of unmanned technologies. The results demonstrate that the single-stage YOLOv11x model is the optimal solution for tree recognition due to its balanced performance, practical benefits, and technical superiority.

РЕЗЮМЕ

В. И. Луйкин, Д. В. Лобов, Д. В. Логинов

Распознавание стволов деревьев

с использованием четырех моделей машинного обучения

Разработка систем автоматического и автоматизированного распознавания деревьев имеет особую важность для лесных территорий России. Огромные площади лесных массивов, трудоёмкость их учёта и необходимость постоянного мониторинга обуславливают потребность в технологиях, способных заменить ручной труд. Задача распознавания деревьев по визуальным характеристикам также становится востребованной в связи с развитием беспилотных технологий. Результаты работы демонстрируют, что оптимальным решением для задачи распознавания деревьев является одноэтапная модель YOLOv11x благодаря сбалансированной производительности, практическим преимуществам и техническому превосходству.

Луйкин Владимир Ильич - инженер-программист-разработчик ООО «Неосервис», 185028 г. Петрозаводск, Первомайский пр., д. 54, телефон +79114356032, bondrol786@gmail.com. *Область научных интересов:* системное программирование. ORCID: 0009-0002-5147-6044

Лобов Денис Владимирович - к.ф.-м.н., доцент, ФГБОУ ВО Петрозаводский государственный университет, 185910, г. Петрозаводск, ул. Ленина, 33, телефон +7905 2998059, ldenis@petsu.ru. *Область научных интересов:* физическое материаловедение, компьютерное моделирование физических систем, модели искусственного интеллекта. ORCID: 0009-0007-7981-9383

Логинов Дмитрий Владимирович - к.ф.-м.н., доцент КФТТ, ФГБОУ ВО Петрозаводский государственный университет, 185910, г. Петрозаводск, ул. Ленина, 33, телефон +79535475757, logindm@mail.ru. *Область научных интересов:* физическое материаловедение, компьютерное моделирование физических систем, модели искусственного интеллекта. ORCID: 0000-0003-3255-2603

Vladimir Ilyich Luikin - Software Engineer/Developer, Neoservice LLC, 54 Pervomaysky Prospekt, Petrozavodsk, 185028, phone +79114356032, bondrol786@gmail.com. Research interests: systems programming. ORCID: 0009-0002-5147-6044

Denis Vladimirovich Lobov - PhD, Associate Professor, Petrozavodsk State University, 33 Lenin Street, Petrozavodsk, 185910, phone +7905 2998059, ldenis@petsu.ru. Research interests: physical materials science, computer modeling of physical systems, artificial intelligence models. ORCID: 0009-0007-7981-9383

Loginov Dmitry Vladimirovich - PhD, Associate Professor, Petrozavodsk State University, 33 Lenin St., Petrozavodsk, 185910, phone +79535475757, logindm@mail.ru. Research interests: physical materials science, computer modeling of physical systems, artificial intelligence models. ORCID: 0000-0003-3255-2603

Статья поступила в редакцию 20.02.2026