

DOI: 10.31509/2658-607x-202372-146  
УДК 574.47; 502.72; 004.932.72'1

# АВТОМАТИЧЕСКАЯ СЕГМЕНТАЦИЯ КРОН СОСНОВЫХ ДРЕВОСТОЕВ С ИСПОЛЬЗОВАНИЕМ MASK R-CNN НА RGB-ОРТОФОТОПЛАНАХ БПЛА

© 2024

А. Д. Никитина

*Центр по проблемам экологии и продуктивности лесов РАН  
Россия, 117997 Москва, ул. Профсоюзная, 84/32, стр. 14*

E-mail: nikitina.al.dm@gmail.com

Поступила в редакцию 18.05.2024

После рецензирования: 05.06.2024

Принята к печати: 22.06.2024

В статье представлены результаты применения усовершенствованного метода автоматической сегментации RGB-ортофотопланов, полученных с помощью беспилотных летательных аппаратов (БПЛА) общедоступного сегмента, на основе нейронной сети архитектуры Mask R-CNN. Для работы с геопространственными данными разработаны блоки подготовки и постобработки растровых и векторных файлов. Модель обучена на 7000 кронах, выделенных в сосновых древостоях дренированных местообитаний подзоны хвойно-широколиственных лесов. Обучение выполнено с применением кросс-валидации. Для верификации дополнительно использованы данные 1337 кроны. На этапе последовательной фильтрации по площади, уровню достоверности и повторам сегментов качество итоговых результатов сегментации улучшилось для всех возрастных групп сосновых древостоев. Итоговая средняя точность (precision) равна 0.87, полнота (recall) – 0.81 и F1-мера (F-score) – 0.83. Полученные результаты свидетельствуют о высокой эффективности алгоритма фильтрации для уменьшения избыточности сегментов и увеличения надёжности данных. Метод автоматической сегментации Mask R-CNN представляет собой эффективный инструмент для исследований характеристик сосновых древостоев по RGB-ортофотопланам БПЛА-съёмки, способный с высокой точностью воспроизводить результаты визуального дешифрирования. Метод особенно эффективен при масштабировании исследований на большие территории, где ручное дешифрирование становится трудоёмким.

**Ключевые слова.** *Mask R-CNN, автоматическая сегментация, выделение деревьев, сосновые древостои, RGB-ортофотопланы, БПЛА, экологический мониторинг, дистанционное зондирование*

В контексте глобальных климатических изменений вопросы углеродного баланса и запасов углерода в лесных

экосистемах приобретают особую актуальность. Эффективный мониторинг этих параметров в лесных экосистемах и

управление лесными ресурсами требуют детальной и точной информации о структуре и состоянии лесов (Espíndola, 2023). В этой связи, применение высокодетальной аэрокосмической съёмки открывает новые возможности для экологических исследований, увеличивая эффективность в сборе и анализе данных. Леса с преобладанием сосны обыкновенной (*Pinus sylvestris* L.) широко распространены в умеренных широтах Северного полушария. Особенности вида являются адаптивностью к разнообразным условиям произрастания и высокая устойчивость к экологическим стрессам, таким как засухи и пожары. Высокая экологическая валентность и значительный вклад в углеродный цикл определяют важность изучения характеристик сосновых лесов в экологических исследованиях. Работы (Медведев и др., 2020; Tuominen et al., 2017; Nevalainen et al., 2017; Puliti et al., 2017; Ocer et al., 2020; Diez et al., 2021; Ball et al., 2023; Zhou et al., 2023) подчёркивают потенциал применения БПЛА и нейронных сетей при обработке изображений для точного и эффективного изучения лесов. Однако, в лесах сложной структуры с сомкнутым пологом проведено относительно мало исследований, что указывает на

необходимость дальнейшего изучения и повышения точности методов в данных условиях. Также важно учитывать возможности использования БПЛА массового сегмента, поскольку их доступность и широкое распространение позволяют применять данные методы в прикладной и исследовательской практике более широким кругом пользователей. Целью настоящего исследования является оценка эффективности метода автоматической сегментации с использованием нейронной сети Mask R-CNN для выделения отдельных деревьев в сосновых древостоях разной структуры на основе RGB-ортофотопланов, полученных с использованием БПЛА.

## МАТЕРИАЛЫ И МЕТОДЫ

**Объекты исследования.** Объектами исследования являются сосняки дренированных местообитаний подзоны хвойно-широколиственных лесов западной части Русской равнины на следующих особо охраняемых природных территориях: национальный парк (НП) «Куршская коса», НП «Смоленское Поозерье», государственный природный биосферный заповедник (ГПБЗ) «Брянский лес». В исследуемых лесах выделены три возрастные группы: молодые (10–40 лет), средневозрастные (40–80 лет) и

старовозрастные (старше 80 лет). Исследования (Nezami et al., 2020; Diez et al., 2021) показывают, что алгоритмы сегментации демонстрируют наиболее высокую эффективность обучения на однопородных и структурно простых лесах. Данные, полученные в НП «Куршская коса», позволяют точнее настроить модель, поскольку исследуемые сосновые леса парка преимущественно имеют однопородный состав с формулой древостоя 10С. Молодые сосновые леса, согласно «Определителю типов леса Европейской России» (Определитель типов..., URL: <http://cepl.rssi.ru/bio/forest/index.htm>), относятся к группе типов ксерофитно-зеленомошных и зеленомошно-лишайниковых. Преобладание ксерофитно-зеленомошных сосняков сохраняется и в средневозрастных, и в старовозрастных лесах. Данные, полученные в сосновых лесах «Смоленского Поозерья» и «Брянского леса», улучшают качество разметки для алгоритмов сегментации применительно к древостоям смешанного состава и сложной структуры, делая их более универсальными. В НП «Смоленское Поозерье» на исследуемых участках средне- и старовозрастных лесов преобладают сосняки кустарничково-зеленомошные, молодые леса представлены преимущественно сосня-

ками мелкотравно-зеленомошными, с сомкнутостью 40–90%. В пределах исследуемых участков ГПБЗ «Брянский лес» преобладают средневозрастные кустарничково-зеленомошные сосняки и старовозрастные сосновые леса сложные с липой и дубом, с сомкнутостью 70–80%.

**Аэрофотосъёмка БПЛА.** В исследовании использовались результаты съёмки БПЛА от компании DJI – Phantom 3 Advanced и Mavic Pro. Эти устройства относятся к категории доступного ценового диапазона и оснащены RGB-камерами. Моделирование полётных заданий проводилось в ПО DroneDeploy. Полёты выполнялись на высоте 100-200 м в зависимости от сложности рельефа и высоты лесного полога с продольным и поперечным перекрытием 90% в безветренных условиях (до 10 м/с) с постоянными погодными условиями примерно с 11:00 до 16:00 по местному времени. Площадь съёмки с такими параметрами составляет ~15 га на один аккумулятор.

**Обработка аэрофотосъёмки** проведена в ПО Agisoft Metashape и включала следующие основные этапы: загрузка снимков; процедура выравнивания снимков; построение плотного облака точек; построение цифровой модели местности (ЦММ); построение ортофотоплана; экспорт ЦММ и

ортофотопланов в растровые форматы. Расчётное пространственное разрешение ЦММ варьировалось от 15 до 32 см/пиксель в зависимости от высоты полёта, ортофотопланов – от 2 до 8 см/пиксель. Суммарное количество первоначальных изображений для создания ортофотопланов составило более 20 тыс. со средним расчётным разрешением ортофотоплана 5.9 см/пиксель и ЦММ 23.6 см/пиксель. Итоговое количество ортофотопланов – 55.

**Визуальное дешифрирование.** В процессе дешифрирования вручную обрисовывались границы крон отдельных деревьев в ПО QGIS. Разметка необходима для создания обучающих и проверочных наборов данных. Она выполнена для центральных участков ортофотопланов размером от 20×20 м (для молодняков) до 100×100 м (для средне- и старовозрастных лесов). В результате получены файлы с векторными пространственными данными (.shp) с кронами отдельных деревьев в сосновых древостоях, а также на дополнительных ключевых участках разнопородного состава для расширения обучающей выборки при дальнейшей сегментации. Итоговый набор данных визуальной разметки включал ~8300 отдельных крон, охватывающий 55 ключевых участков. Больше всего крон

было выделено в сосновых древостоях – 6799, из них в средневозрастных лесах – 3330, в старовозрастных – 2325, в молодых – 1144. Вспомогательные участки с другими породами составляют меньшую часть (широколиственные леса – 562, еловые леса – 823, мелколиственные леса – 141).

**Автоматическая сегментация с использованием нейронной сети архитектуры Mask R-CNN.** Для выделения отдельных крон деревьев на аэрофотопланах применяется сверхточная нейронная сеть Mask R-CNN, разработанная в 2017 году (He et al., 2017). Она расширяет возможности нейронной сети Faster R-CNN за счет добавления модуля, который предсказывает маски сегментации для областей интереса (RoI). Этот модуль работает параллельно с классификацией и регрессией ограничивающего прямоугольника (Bounding box). Одной из особенностей Mask R-CNN является пиксельное выравнивание, отсутствующее в Fast / Faster R-CNN.

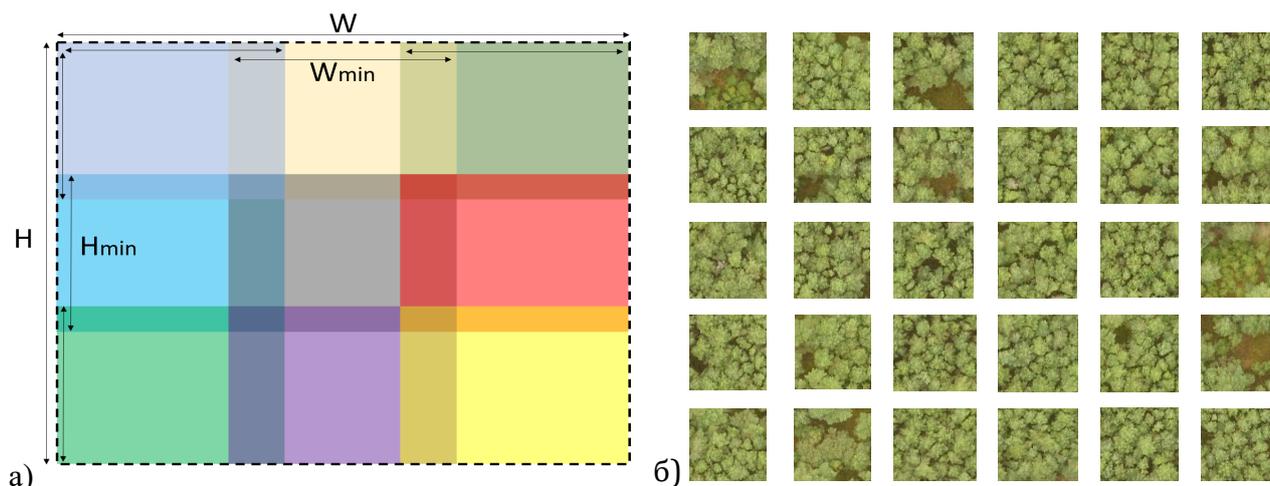
Для создания проекта использовались следующие инструменты, фреймворки и библиотеки: CUDA, Jupyter Notebook, QGIS, Pytorch, *rasterio*, *fiona*, *Matplotlib*. В исследовании применена модель Mask R-CNN, реализованная в фреймворке машинного

обучения *PyTorch* на языке Python, предобученная на наборе данных «СОСО», включающем более 330 000 изображений и 1.5 миллиона объектов. Предобученная модель способна распознавать границы объектов на изображениях. Однако для задачи распознавания крон деревьев необходимо дополнительное обучение. Этот процесс проще и быстрее, чем обучение с нуля, снижает риск застревания в локальных минимумах и уменьшает количество необходимых корректировок. Перенастройка параметров нейронной сети выполнялась на этапах классификации областей интереса, создания ограничивающих прямоугольников и сегментации масок.

*Создание набора данных для обучения модели.* Первым этапом является создание специализированного набора данных. Для этого использовались следующие исходные данные: ортофотопланы БПЛА (формат GeoTIFF .tiff); векторные границы крон деревьев, выделенные вручную (формат .shp); границы исследуемых участков (формат

.shp). Модель обучалась на контурах семи тысяч крон деревьев (84%), полученных при визуальной разметке, со всех исследуемых территорий. Размер валидационного набора составил 1337 крон (16%) с ключевых участков сосновых лесов стратифицировано с учётом объекта и возраста древостоя.

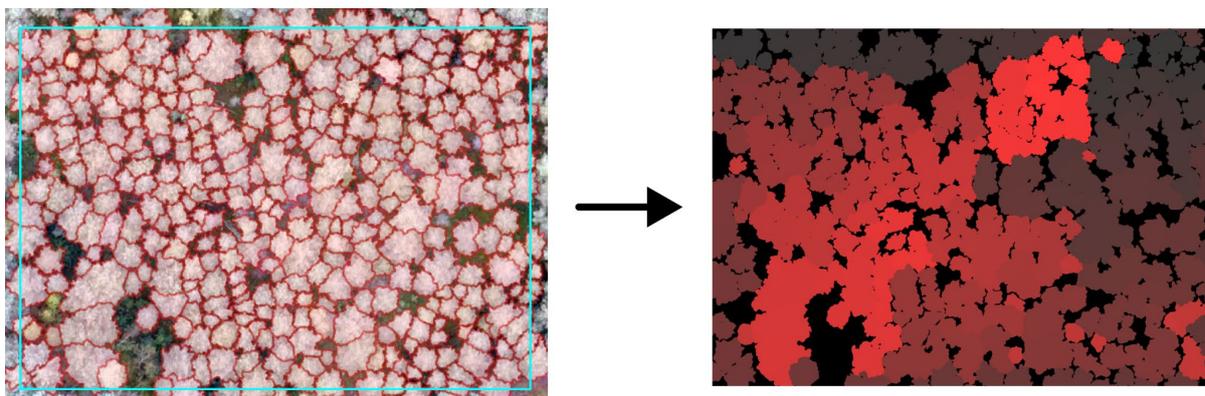
Этапы подготовки набора данных включали проверку и коррекцию геометрии векторных файлов (неправильных форм объектов, самопересечений и т.д.), перепроецирование данных; конвертацию изображений в 24-битный формат; обеспечение правильного совмещения входных данных с использованием единых границ исследуемых участков. Модель принимает на вход изображения в виде матрицы  $[W, H, 3]$  (где  $W$  и  $H$  – ширина и высота, а 3 – число каналов RGB), поэтому исходные изображения были разделены на компоненты с одинаковыми размерами по оптимальной сетке. Ширина и высота каждого компонента определялась как минимальная среди всех изображений (рис. 1).



**Рисунок 1.** (а) Схема разбиения исходных растровых изображений на отдельные компоненты, где  $W$  и  $H$  – ширина и высота, а 3 – число каналов RGB, и (б) отдельные компоненты после разбиения

В качестве разметки использовалось множество бинарных изображений каждо-

го сегмента (кроны), где пиксели имеют значение [0] для фона и [1] для кроны (рис. 2).



**Рисунок 2.** Преобразование входного набора данных в формат многомерного массива

Таким образом, файлы с визуальной разметкой были трансформированы в многомерные массивы, где каждая крона сохранена как отдельное изображение. Итоговый набор данных, принимающийся моделью, включал растровые ортофотопланы (.png), информацию о границах участков с конфигу-

рацией преобразования (.json), маски разметки в виде растров, где каждой кроне соответствует определённое числовое значение пикселя (.png), ортофотопланы, обрезанные по расширенной границе для корректного определения краевых объектов (.png). Дополнительно подгружались векторные данные

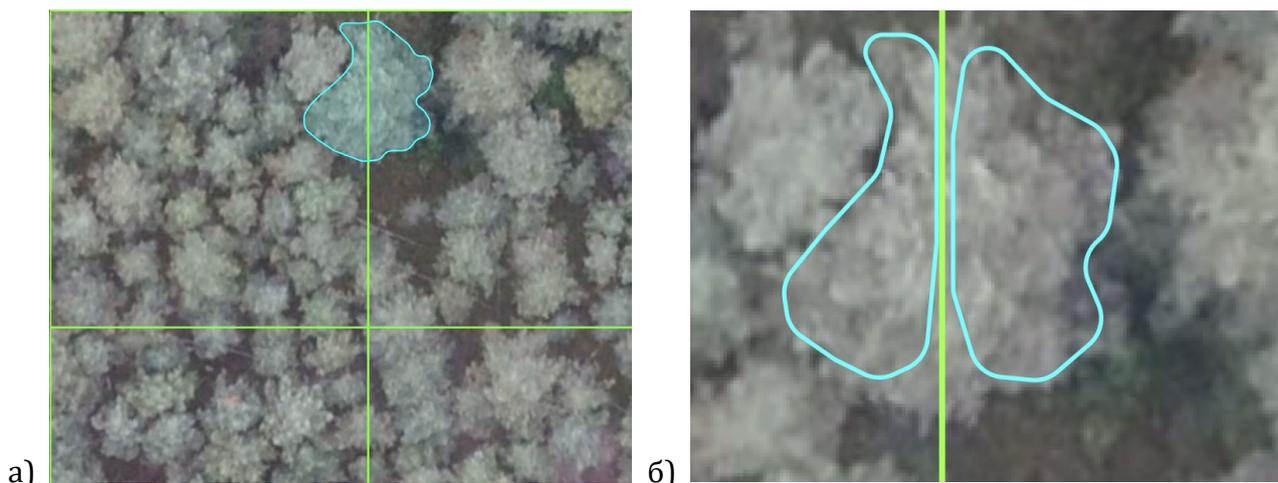
разметки, которые не участвуют в обучении и используются на этапе оценки работы модели.

*Обучение модели нейронной сети Mask R-CNN.* Для обучения модели использовался подготовленный набор данных, который был разделен на тренировочные и валидационные подмножества. Валидационные данные не применялись в процессе обучения. Для тренировочных данных применялась многократная k-блочная кросс-валидация (k=10). Параметры обучения нейронной сети включали начальный коэффициент скорости обучения (learning rate, LR), период изменения LR, фактор изменения LR, коэффициент регуляризации, параметр стохастического градиентного спуска (Stochastic Gradient Descent, SGD). Обучение проводилось в течение 9 эпох (за одну эпоху весь набор данных проходит через нейронную сеть и происходит корректировка весов модели). Преобразования входного изображения для увеличения набора данных включали повороты, изменение контрастности, насыщенности, яркости с общей вероятностью изменения 0.1. Для предотвращения переобучения модель регулярно оценивалась на валида-

ционном наборе данных. Результаты работы нейронной сети – это многомерный стек монохромных изображений каждого объекта с указанием степени уверенности (confidence level) в диапазоне от 0 до 1. Показатель отражает вероятность принадлежности выходного объекта к заданному классу (кроне).

*Обработка данных, полученных моделью.* Для анализа данных обученной нейронной сети передавались ортофотопланы различных пробных площадей в формате .tiff с границами ключевых участков и результатами визуального дешифрирования для последующего расчёта метрик качества модели. Учитывая, что модель работает с изображениями в пиксельных координатах, создавался дополнительный файл метаданных с информацией о системе координат, координатах углов сегментируемой области и географической привязке.

В связи с ограничением модели на обработку не более 100 сегментов (крон) на изображение, необходимо разделить исходный растр на части (патчи) так, чтобы в каждом было менее 100 крон. Однако при делении изображений строго по сетке возможны искажения в сегментах на границах участков (рис. 3).



**Рисунок 3.** Пример а) разбиения изображения на патчи и б) возможного варианта погрешности при сегментации краевых участков

Проблема краевых крон, возникающая при разбиении изображения по сетке, решалась путем создания 50 %-го перекрытия между патчами и определения зоны игнорируемых границ для исключения краевых эффектов. Информация о смещении сохранялась для дальнейшего восстановления координат компонентов в исходном полном размере изображения. Использование пирамиды патчей, где размер компонентов удваивается на каждом уровне, обеспечило адаптацию к разным масштабам объектов. Полученные сегменты преобразовывались в изображения, а затем – в векторный формат с использованием алгоритма маршрутизирующих квадратов (порог уверенности = 0.5), реализованного в библиотеке *Skimage*. Кроны, пересекающие зону игнорируемых границ, удалялись из

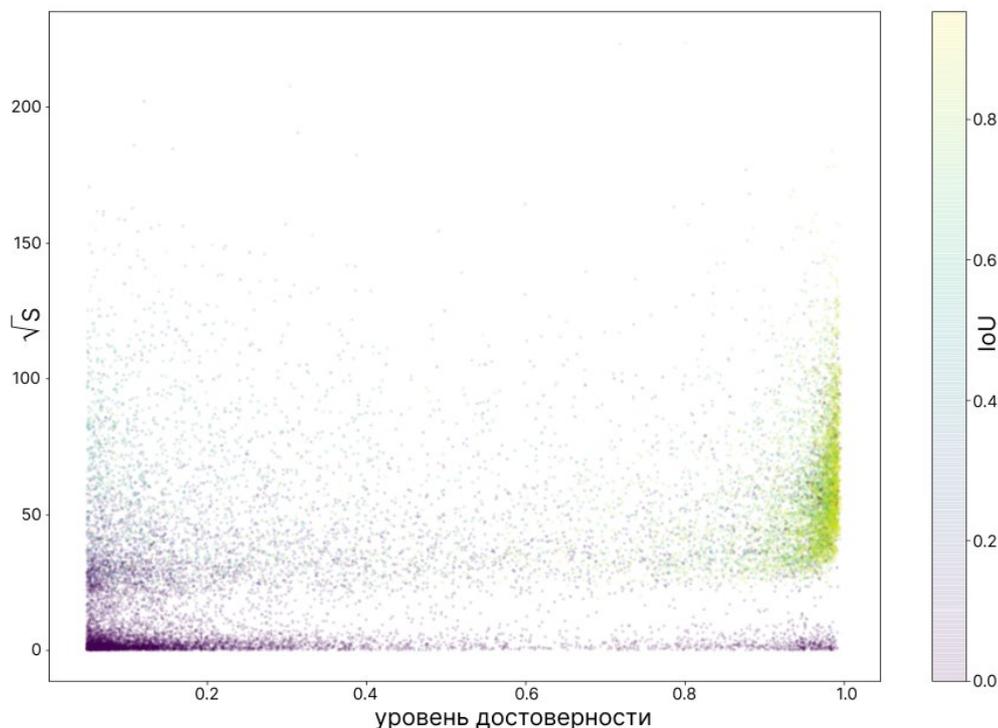
результатов, сокращая число дублированных и краевых крон. Оставшиеся сегменты объединялись в единый набор данных, где для каждой выделенной кроны указывался уровень достоверности.

*Фильтрация данных сегментации.* Фильтрация в контексте обработки данных нейросетью представляет собой важный этап, направленный на улучшение качества результатов. Во время разбиения изображений на отдельные компоненты закладывается большое перекрытие, что не дает модели пропустить краевые кроны деревьев, однако при сборке результатов сегментации в один набор данных возникают дубликаты одной и той же кроны. Помимо дубликатов могут встречаться объекты с неправильной формой, недостоверной площадью или с низкой

степенью уверенности. Всё это требовало тщательной фильтрации, которая включала в себя анализ различных параметров с целью определения оптимальных критериев для удаления нежелательных сегментов.

Вычисление оптимальных параметров для алгоритма фильтрации производилось на основе данных с посчитанной метрикой IoU, которая измеряет степень пересечения между кронами, прогнозируемыми нейронной сетью, и кронами, выделенными визуально. Данные представляли собой множество точек, где каждая точка соответствовала детектированной кроне

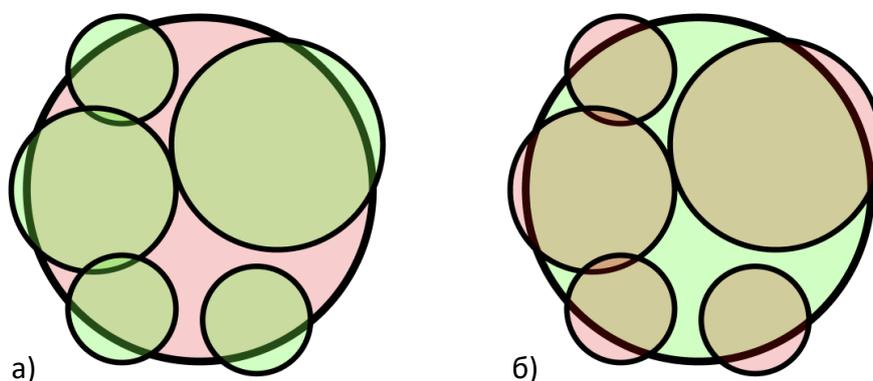
с несколькими параметрами (площадь, уровень достоверности (score), IoU). При анализе диаграммы рассеяния (рис. 4) чётко выделяется порог для фильтрации сегментов с маленькой площадью. Этот этап фильтрации значительно сократил число излишне сегментированных крон (28%) без значительной потери точности. На диаграмме рассеяния также видно, что сегменты с минимальными значениями IoU в большинстве своем также обладают и низким уровнем достоверности, поэтому целесообразно использовать в дальнейшем уровень достоверности как параметр фильтрации.



**Рисунок 4.** Диаграмма рассеяния, показывающая взаимосвязь между площадью сегмента ( $\sqrt{S}$ , пиксели), его уровнем достоверности, и значением степени пересечения сегмента, предсказанного нейронной сетью, с эталонным сегментом разметки (IoU)

Для удаления повторяющихся крон использовалась фильтрация на основе степени пересечения и достоверности данных. Если две кроны сильно пересекаются, выбирается та, которая имеет более высокий уровень достоверности нейронной сети. В случае наличия одного крупного и нескольких мелких перекры-

вающих сегментов, остаётся только крупный, если его достоверность выше. Если же его достоверность ниже, такой сегмент исключается. Так, при высокой уверенности в меньших сегментах, крупная крона отбрасывается (рис. 5а), и наоборот (рис. 5б). Удаление дубликатов также значительно повысило точность.

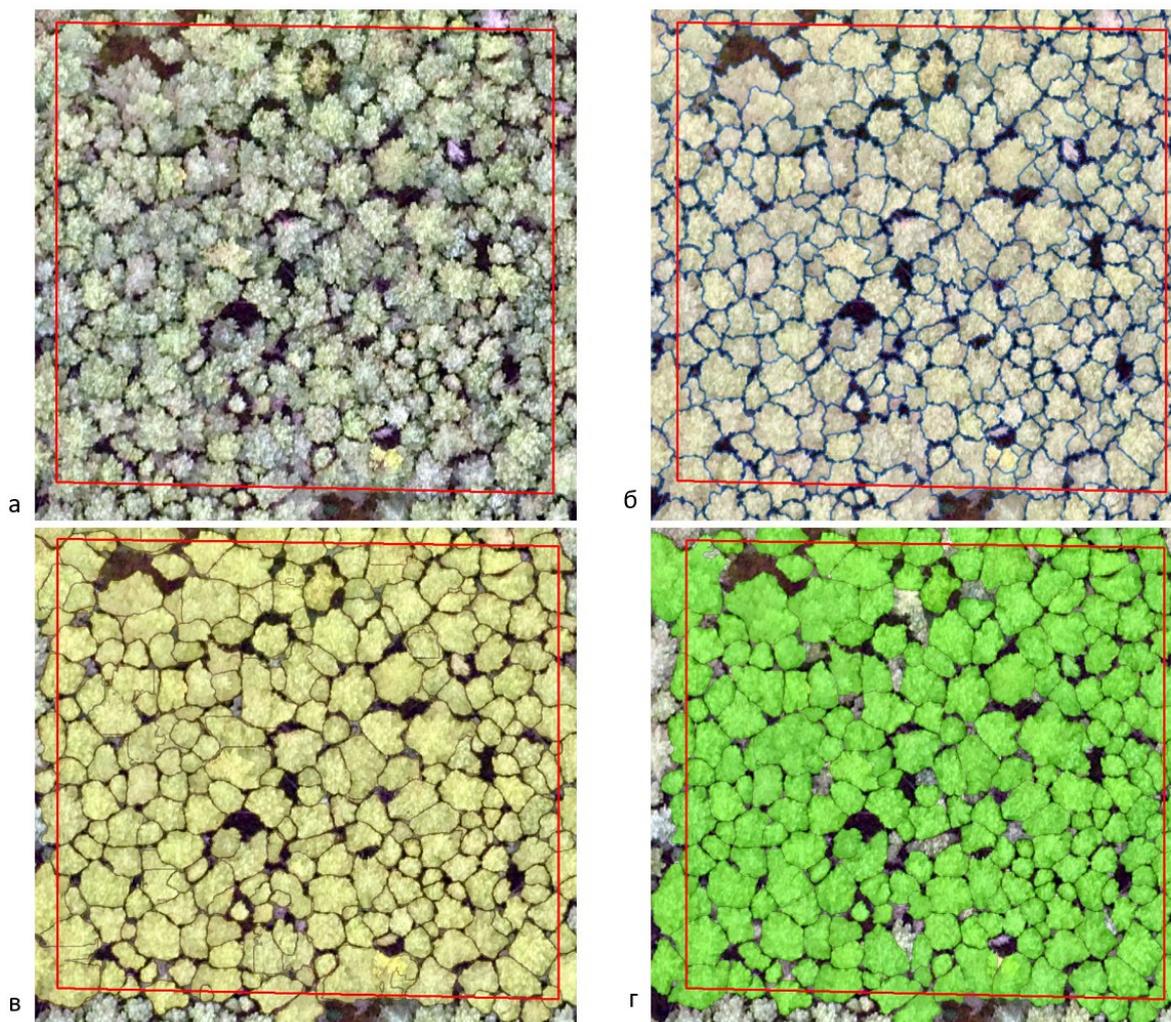


**Рисунок 5.** Варианты фильтрации при перекрытии крон с разной степенью уверенности (зелёные сегменты сохраняются в итоговом списке, красные – удаляются)

Разработанная последовательная фильтрация включала в себя критерии площади и уровня достоверности, а затем повторов сегментов. Это позволило сохранить качественные сегменты, минимизируя потери полноты данных.

*Создание итогового векторного слоя.* Для трансформации координат

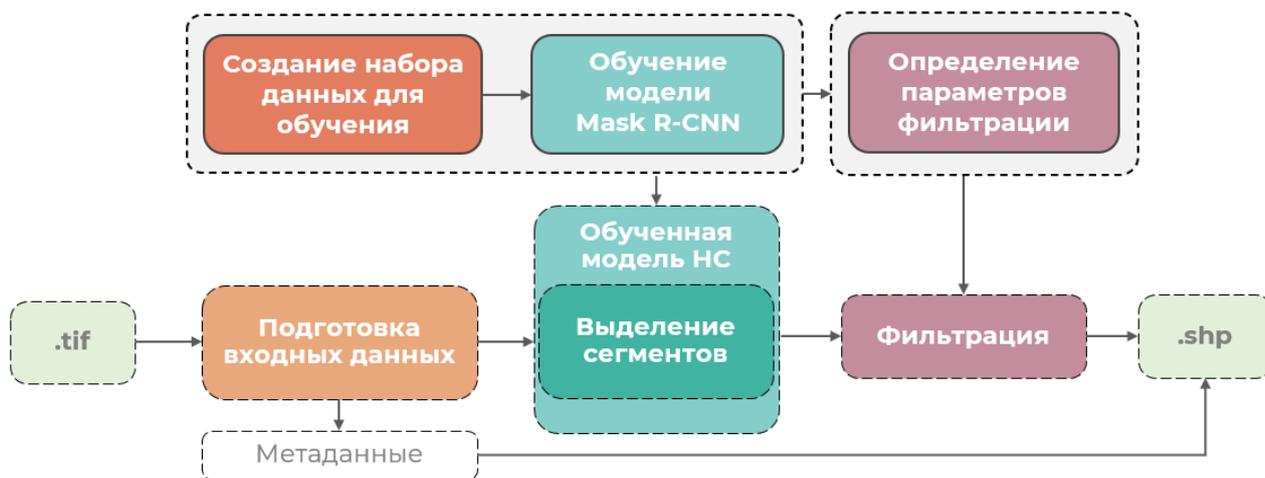
сегментов в географические и создания итогового векторного файла формата .shp использовался файл метаданных, подготовленный на этапе обработки данных. Результаты сегментации крон нейронной сетью представлены на рисунке 6, демонстрируя точность преобразования координат.



**Рисунок 6.** Пример ПП, где (а) – ортофотоплан без разметки, (б) – визуальное дешифрирование, (в) – результат автоматической сегментации без фильтрации, (г) – результат автоматической сегментации после фильтрации

На схеме (рис. 7) представлен итоговый процесс обработки данных для сегментации крон сосновых древостоев с использованием алгоритма Mask R-CNN. Верхняя часть схемы отражает исследовательские этапы, включающие создание набора данных для обучения, непосредственно обучение модели Mask R-CNN и определение параметров фильтрации. Эти этапы выполняются

один раз для настройки модели. Нижняя часть схемы демонстрирует процесс обработки. Он начинается с загрузки растровых файлов, затем происходит подготовка входных данных, последующее выделение сегментов и фильтрация результатов. Итогом является создание векторного файла, который может быть использован для дальнейшего анализа.



**Рисунок 7.** Этапы обработки данных для автоматического выделения крон

*Метрики качества модели.* Для оценки точности нейронной сети в задаче распознавания крон деревьев использовались результаты визуального дешифрирования ортофотопланов. Считалось, что крона детектирована корректно, если IoU превышает 0.5, что является стандартным значением в работах о сегментации (Aubry-Kientz et al., 2019; Hao et al., 2021; Ball et al., 2023). Для анализа качества модели рассчитаны стандартные матрицы ошибок: TP (true positive) – корректное определение кроны; FP (false positive) – неверное определение объекта как кроны; FN (false negative) – неверное исключение кроны; TN (true negative) – равно 0 в задачах сегментации. На основе этих данных рассчитывались ключевые метрики: precision (точность) – доля

правильно идентифицированных крон среди всех распознанных; recall (полнота) – доля корректно идентифицированных крон из всех действительно существующих; F1-мера – гармоническое среднее между точностью и полнотой, обеспечивающее баланс этих показателей.

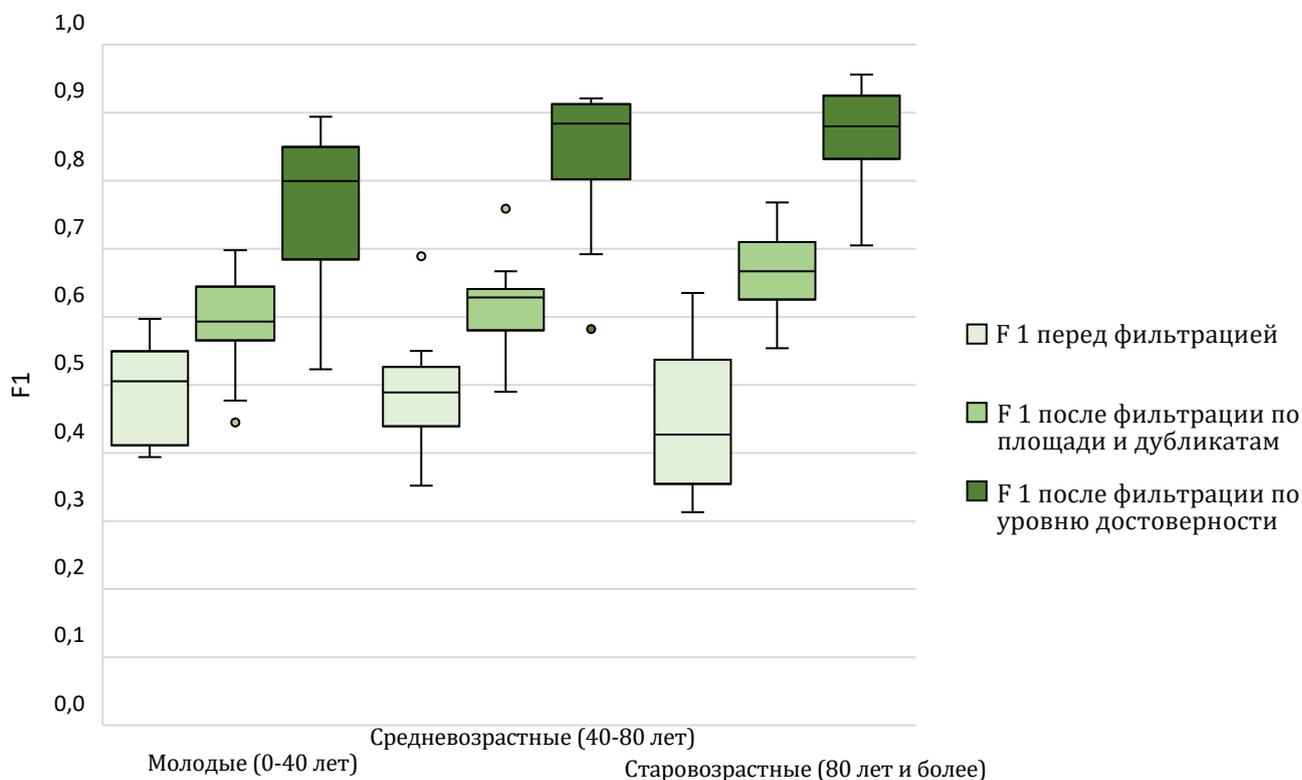
## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

На всех ключевых участках исходные результаты сегментации нейронной сети характеризуются высокими значениями полноты (recall = 0.91 для всех площадок) и низкими показателями точности (precision = 0.31) и F1-меры (0.46), избыточность сегментов видна и по соотношению числа сегментированных крон к данным визуальной разметки. Фильтрация

значительно улучшила итоговые средние показатели точности (0.87) и F1-меры (0.83), незначительно уменьшив итоговую полноту ( $\text{recall} = 0.81$ ).

На графике (рис. 8), показывающем изменение F1-меры на разных этапах фильтрации для разных возрастных групп сосновых древостоев, можно отметить, что после всех этапов фильтрации медианные значения F1-меры увеличиваются для всех

возрастных групп, что указывает на улучшение качества сегментации для всей выборки. Однако для старовозрастных сосновых лесов (от 80 лет) улучшение после фильтрации наиболее выражено. Это указывает на эффективность примененного подхода фильтрации для улучшения качества результатов сегментации, сокращая избыточность и повышая надёжность данных.



**Рисунок 8.** Изменение F1-меры на разных этапах фильтрации для сосновых лесов разных групп возраста

Анализ обучающей и валидационной выборки, используемой для контроля переобучения при настройке модели, показал наибольшие различия в

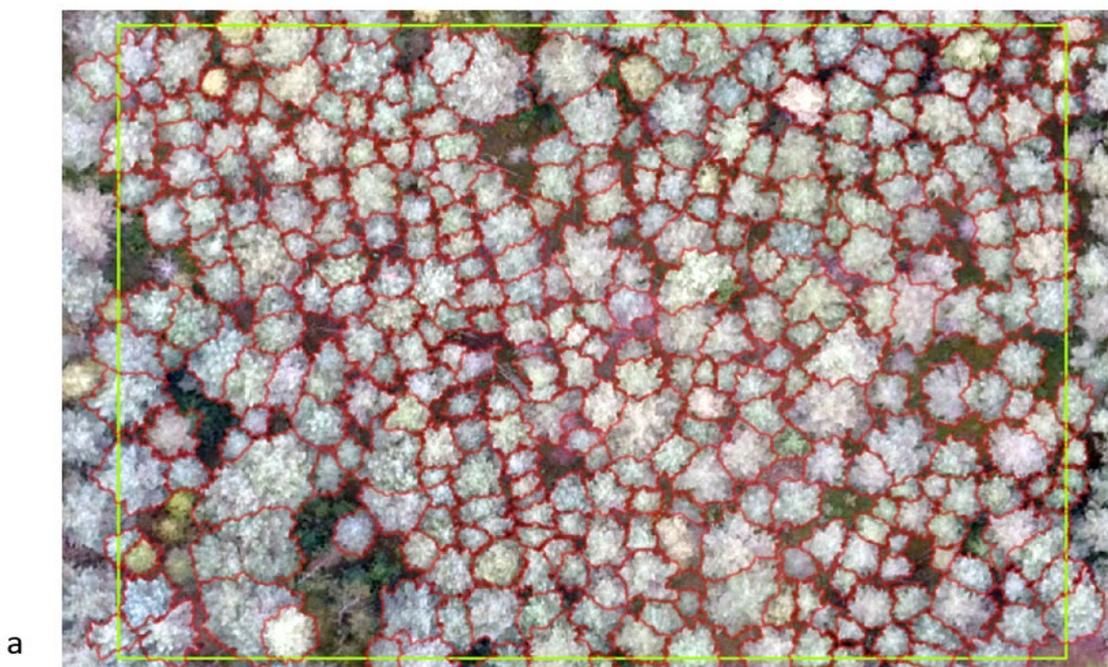
группе молодых сосновых лесов ( $F1_{\text{обучение}} = 0.81$ ,  $F1_{\text{валидация}} = 0.70$ ) с медианным значением 0.8 для всех ключевых участков. Средневозрастные и

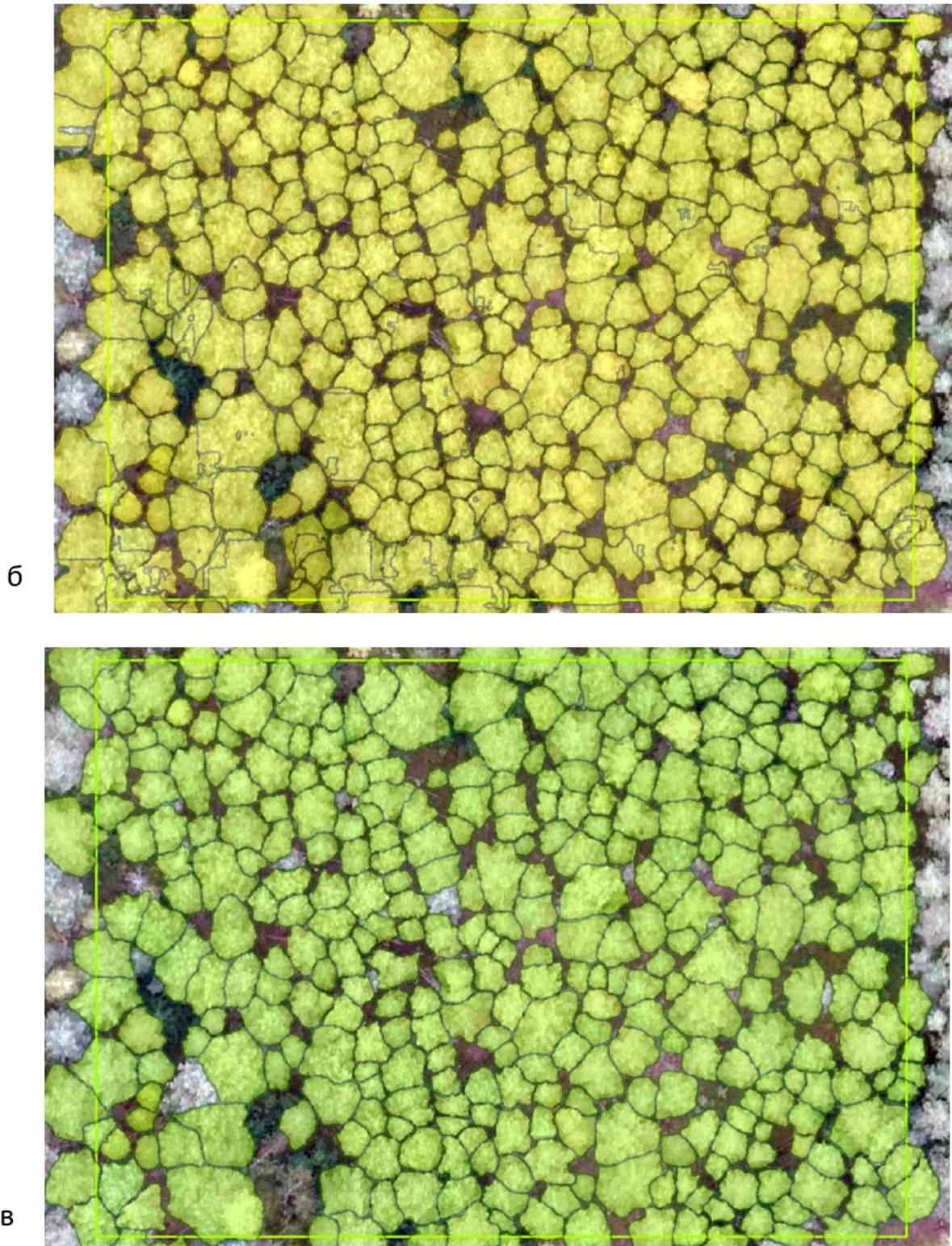
старовозрастные сосновые древостои имеют высокие показатели F1-меры как для обучающего (0.84 и 0.88 соответственно), так и для валидационного набора (0.83 и 0.82 соответственно) участков с медианным значением F1-меры 0.88 для обеих групп.

Разброс значений качества модели в сосновых лесах составил 0.53–0.96 со средним значением 0.83 и медианным значением 0.85. В молодых лесах разброс результатов ( $F1 = 0.53\text{--}0.89$ ) показывает меньшую адаптивность модели к некоторым древостоям этого возраста, однако результаты в среднем высокие ( $F1_{\text{сред.}} = 0.77$ ,  $F1_{\text{мед.}} = 0.8$ ). Это может быть обусловлено низким качеством съёмки (для невысоких древостоев стоит проводить БПЛА-съёмку ниже 120-180 м при ис-

пользовании БПЛА массового сегмента), сложным выделением отдельных деревьев в плотных насаждениях, меньшей обучающей выборкой для ПП молодых сосновых лесов. Более устойчивые результаты ( $F1 = 0.7\text{--}0.96$ ) получены для старовозрастных лесов со средним значением  $F1 = 0.86$  ( $F1_{\text{медиана}} = 0.88$ ).

В исследовании продемонстрировано, что адаптированная модель Mask R-CNN обеспечивает высокую точность результатов в разных возрастных группах сосновых древостоев с разными показателями сомкнутости, так как показатели качества сегментации на ключевых участках остаются высокими для всех наборов данных. Пример результатов сегментации представлен на рисунке 9.





**Рисунок 9** Пример ключевого участка, где (а) – визуальное дешифрирование, (б) – результат сегментации без фильтрации, (в) – результат сегментации после фильтрации

В исследованиях, посвященных сегментации отдельных деревьев в древостоях, важным аспектом является сомкнутость исследуемых древостоев. В работе Н. Е. Осег и соавторов (2020), посвященной выделению отдельных деревьев с использованием Mask R-CNN и пирамиды масштабов (FPN), для трёх тестовых изображений были получены результаты F1-меры в диапазоне 0.82–0.91. Разреженные древостои анализировались в исследовании Н. В. Ивановой с соавторами (Ivanova et al., 2021), где были применены методы водораздела и наращивания областей и продемонстрированы значения F1-меры 0.7–0.9. Более сомкнутые древостои рассмотрены в работе X. Chen и др. (2023), с результатами F1-меры между 0.71 и 0.79. Исследование М. Veloiu и соавторов (2023) посвящено сомкнутым и разнопородным древостоям с результатами F1 от 0.44 до 0.92. Исследования показывают, что эффективность сегментации крон зависит от сомкнутости древостоев, причём с её увеличением точность сегментации становится более изменчивой.

### ЗАКЛЮЧЕНИЕ

Метод автоматической сегментации изображений с использованием нейронной сети архитектуры Mask R-CNN представляет собой эффективный инс-

трумент для исследований сосновых древостоев, способный с высокой точностью воспроизводить результаты визуального дешифрирования. Разбиение RGB-ортофотопланов позволило наиболее полностью учесть отдельные кроны деревьев в сомкнутом пологе. Исходные результаты характеризовались высокими значениями полноты, однако для увеличения точности был разработан блок фильтрации результатов. Фильтрация позволила исключить лишние сегменты и улучшить точность результатов, сохраняя при этом высокую степень распознавания крон. Для всех возрастных групп сосновых лесов наблюдается увеличение значений F1-меры при фильтрации. Итоговая модель демонстрирует стабильно высокое качество сегментации (F1-мера = 0.83) сосновых древостоев.

### ФИНАНСИРОВАНИЕ

Работа выполнена в рамках темы ГЗ ЦЭПЛ РАН «Биоразнообразие и экосистемные функции лесов» (НИОКТР 124013000750-1) при поддержке Российского фонда фундаментальных исследований (проект № 20-34-90152) на базе Лаборатории мониторинга лесных экосистем и молодежной Лаборатории климаторегулирующих функций и биоразнообразия лесов (122111500023-6) ЦЭПЛ РАН.

### СПИСОК ЛИТЕРАТУРЫ

- Медведев А. А., Тельнова Н. О., Кудиков А. В., Алексеенко Н. А.* Анализ и картографирование структурных параметров редкостойных северотаёжных лесов на основе фотограмметрических облаков точек // Современные проблемы дистанционного зондирования Земли из космоса. 2020. Т. 17. № 1. С. 150–163.
- Определитель типов леса Европейской России. URL: <https://cepl.rssi.ru/bio/forest/index.htm> (дата обращения 01.06.2024).
- Agisoft Metashape: официальный сайт. URL: <http://www.agisoft.com> (дата обращения 01.06.2024).
- Aubry-Kientz M., Dutrieux R., Ferraz A., Saatchi S., Hamraz H., Williams J.* A comparative assessment of the performance of individual tree crowns delineation algorithms from ALS data in tropical forests // Remote Sensing. 2019. Vol. 11. No. 9. P. 1086 (1–21).
- Ball J. G., Hickman S. H., Jackson T. D., Koay X. J., Hirst J., Jay W., Coomes D. A.* Accurate delineation of individual tree crowns in tropical forests from aerial RGB imagery using Mask R-CNN // Remote Sensing in Ecology and Conservation. 2023. Vol. 9. No. 5. P. 641–655.
- Beloiu M., Heinzmann L., Rehush N., Gessler A., Griess V. C.* Individual Tree-Crown Detection and Species Identification in Heterogeneous Forests Using Aerial RGB Imagery and Deep Learning // Remote Sensing. 2023. Vol. 15. P. 1463.
- Chen X., Shen X., Cao L.* Tree Species Classification in Subtropical Natural Forests Using High-Resolution UAV RGB and SuperView-1 Multispectral Imageries Based on Deep Learning Network Approaches: A Case Study within the Baima Snow Mountain National Nature Reserve, China // Remote Sensing. 2023. Vol. 15. P. 2697.
- Diez Y., Kentsch S., Fukuda M., Caceres M.L.L., Moritake K., Cabezas M.* Deep Learning in Forestry Using UAV-Acquired RGB Data: A Practical Review // Remote Sensing. 2021. Vol. 13. P. 2837.
- Espíndola R. P., Ebecken N. F. F.* Advances in remote sensing for sustainable forest management: monitoring and protecting natural resources // Revista Caribeña de Ciencias Sociales. 2023. Vol. 12. No. 4. P. 1605–1617.
- Hao Z., Lin L., Post C.J., Mikhailova E.A., Li M., Chen Y. et al.* Automated tree-crown and height detection in a young forest plantation using mask region-based convolutional neural network (Mask R-CNN) // ISPRS Journal of Photog-

- rammetry and Remote Sensing. 2021. Vol. 178. P. 112–123.
- He K., Gkioxari G., Dollár P., Girshick R.* Mask R-CNN // Proceedings of the IEEE International Conference on Computer Vision. 2017. P. 2961–2969.
- Ivanova N. V., Shashkov M. P., Shanin V. N.* Study of pine forest stand structure in the priosko-terrasny state nature biosphere reserve (Russia) based on aerial photography by quadrocopter // Nature Conservation Research. 2021. Vol. 6. No. 4. P. 1–14.
- Nevalainen O., Honkavaara E., Tuominen S., Viljanen N., Hakala T., Yu. X., Hyyppä J., Saari H., Pölönen I., Imai N. N., Tommaselli A. M. G.* Individual tree detection and classification with UAV-based photogrammetric point clouds and hyperspectral imaging // Remote Sensing. 2017. Vol. 9. No. 3. P. 185.
- Nezami S., Khoramshahi E., Nevalainen O., Pölönen I., Honkavaara E.* Tree species classification of drone hyperspectral and RGB imagery with deep learning convolutional neural networks // Remote Sensing. 2020. Vol. 12. No. 7. P. 1070.
- Ocer N. E., Kaplan G., Erdem F., Matci D.K., Avdan U.* Tree extraction from multi-scale UAV images using Mask R-CNN with FPN // Remote Sensing. 2020. Vol. 11. P. 847–856.
- Puliti S., Ene L. T., Gobakken T., Næsset E.* Use of partial-coverage UAV data in sampling for large scale forest inventories // Remote Sensing of Environment. 2017. Vol. 194. P. 115–126.
- Tuominen S., Näsi R., Honkavaara E., Balazs A., Hakala T., Viljanen N., Reinikainen J.* Tree species recognition in species rich area using UAV-borne hyperspectral imagery and stereo-photogrammetric point cloud // International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2017. Vol. XLII-3/W3. P. 185–194.
- Zhou J., Chen X., Li S., Dong R., Wang X., Zhang C., Zhang L.* Multispecies individual tree crown extraction and classification based on BlendMask and high-resolution UAV images // Journal of Applied Remote Sensing. 2023. Vol. 17. No. 1. P. 016503.

## REFERENCES

- Agisoft Metashape, available at: <http://www.agisoft.com> (2024, 01 June).
- Aubry-Kientz M., Dutrieux R., Ferraz A., Saatchi S., Hamraz H., Williams J., A comparative assessment of the performance of individual tree crowns delineation algorithms from ALS data in tropical forests, *Remote Sensing*, 2019, Vol. 11, No 9, pp. 1086 (1–21).

- Ball J. G., Hickman S. H., Jackson T. D., Koay X. J., Hirst J., Jay W., Coomes D. A., Accurate delineation of individual tree crowns in tropical forests from aerial RGB imagery using Mask R-CNN, *Remote Sensing in Ecology and Conservation*, 2023, Vol. 9, No 5, pp. 641–655.
- Beloiu M., Heinzmann L., Rehus N., Gessler A., Griess V. C., Individual Tree-Crown Detection and Species Identification in Heterogeneous Forests Using Aerial RGB Imagery and Deep Learning, *Remote Sensing*, 2023, Vol. 15, p. 1463.
- Chen X., Shen X., Cao L., Tree Species Classification in Subtropical Natural Forests Using High-Resolution UAV RGB and SuperView-1 Multispectral Imageries Based on Deep Learning Network Approaches: A Case Study within the Baima Snow Mountain National Nature Reserve, China, *Remote Sensing*, 2023, Vol. 15, p. 2697.
- Diez Y., Kentsch S., Fukuda M., Caceres M. L. L., Moritake K., Cabezas M., Deep Learning in Forestry Using UAV-Acquired RGB Data: A Practical Review, *Remote Sens.*, 2021, Vol. 13, p. 2837.
- Espíndola R. P., Ebecken N. F. F., Advances in remote sensing for sustainable forest management: monitoring and protecting natural resources, *Revista Caribeña de Ciencias Sociales*, 2023, Vol. 12, No 4, pp. 1605–1617.
- Hao Z., Lin L., Post C.J., Mikhailova E.A., Li M., Chen Y. et al., Automated tree-crown and height detection in a young forest plantation using mask region-based convolutional neural network (Mask R-CNN), *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, Vol. 178, pp. 112–123.
- He K., Gkioxari G., Dollár P., Girshick R., Mask R-CNN, *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 2961–2969.
- <https://cepl.rssi.ru/bio/forest/index.htm> (2024, 1 June).
- Ivanova N.V., Shashkov M.P., Shanin V.N., Study of pine forest stand structure in the priosko-terrasny state nature biosphere reserve (Russia) based on aerial photography by quadrocopter, *Nature Conservation Research*, 2021, Vol. 6, No 4, pp. 1–14.
- Medvedev A. A., Tel'nova N. O., Kudikov A. V., Alekseenko N. A., Analiz i kartografirovanie strukturnykh parametrov redkostojnyh severotajozhnyh lesov na osnove fotogrammetricheskikh oblakov toчек (Use of photogrammetric point clouds for the analysis and mapping of structural variables in sparse northern boreal forests), *Sovremennye problemy distancionnogo zondirovaniya Zemli iz kosmosa*, 2020, Vol. 17, No 1, pp. 150–163.

- Nevalainen O., Honkavaara E., Tuominen S., Viljanen N., Hakala T., Yu X., Hyypä J., Saari H., Pölönen I., Imai N. N., Tommaselli A. M. G., Individual tree detection and classification with UAV-based photogrammetric point clouds and hyperspectral imaging, *Remote Sensing*, 2017, Vol. 9, No 3, p. 185.
- Nezami S., Khoramshahi E., Nevalainen O., Pölönen I., Honkavaara E., Tree species classification of drone hyperspectral and RGB imagery with deep learning convolutional neural networks, *Remote Sensing*, 2020, Vol. 12, No. 7, Article 1070.
- Ocer N. E., Kaplan G., Erdem F., Matci D. K., Avdan U., Tree extraction from multi-scale UAV images using Mask R-CNN with FPN, *Remote Sensing*, 2020, Vol. 11, p. 847–856.
- Puliti S., Ene L. T., Gobakken T., Næsset E., Use of partial-coverage UAV data in sampling for large scale forest inventories, *Remote Sensing of Environment*, 2017, Vol. 194, pp. 115–126.
- Tuominen S., Näsi R., Honkavaara E., Balazs A., Hakala T., Viljanen N., Reinikainen J., Tree species recognition in species rich area using UAV-borne hyperspectral imagery and stereo-photogrammetric point cloud, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2017, Vol. XLII-3/W3, pp. 185–194.
- Zhou J., Chen X., Li S., Dong R., Wang X., Zhang C., Zhang L., Multispecies individual tree crown extraction and classification based on BlendMask and high-resolution UAV images, *Journal of Applied Remote Sensing*, 2023, Vol. 17, No 1, p. 016503.

# AUTOMATIC SEGMENTATION OF TREE CROWNS IN PINE FORESTS USING MASK R-CNN ON RGB IMAGERY FROM UAVS

**A. D. Nikitina**

*Center for Forest Ecology and Productivity of the RAS  
Profsoyuznaya st. 84/32 bldg. 14, Moscow, 117997, Russia*

E-mail: nikitina.al.dm@gmail.com

Received: 18.05.2024

Revised: 05.06.2024

Accepted: 22.06.2024

The article presents the results of applying an improved method for automatic segmentation of RGB imagery obtained using consumer-grade UAVs, based on the Mask R-CNN neural network architecture. Blocks for the preparation and post-processing of raster and vector files have been developed for working with geospatial data. The model was trained on 7000 crowns identified in pine forest of automorphic habitats in the mixed coniferous-broadleaf forest subzone. Training was carried out using cross-validation. Additional data of 1337 crowns were used for verification. During the sequential filtering by area, confidence level, and duplicate segments, the quality of the final segmentation results improved for all age groups of pine forests. The final average precision is 0.87, recall – 0.81, F1-score – 0.83. The results demonstrate the high efficiency of the filtering algorithm in reducing segment redundancy and increasing data reliability. The Mask R-CNN automatic segmentation method is an effective tool for analyzing the characteristics of pine canopies using RGB imagery from UAV surveys. It is capable of replicating the results of visual interpretation with high accuracy. This method particularly advantageous for scaling studies to large areas where manual delineation becomes labor-intensive.

**Keywords:** *Mask R-CNN, automatic segmentation, detection trees, pine forests, RGB imagery, UAVs, ecological monitoring, remote sensing*

**Рецензент:** к.г.н. Малышева Н. В.