

# Segmentation of Olive Trees Using Mask R-CNN Architecture

Anastasia N. Safonova<sup>(1,2)</sup>, Siham Tabik<sup>(2)</sup>, Yuri A. Maglinets<sup>(1)</sup>

<sup>(1)</sup> Institute of Space and Information Technologies.

Siberian Federal University, Krasnoyarsk, Russia

<sup>(2)</sup> Center for Information and Communication Technologies.

University of Granada, Spain

**Abstract:** In this work, we present the results of the instance segmentation of olive trees using high resolution images, NDVI and GNDVI vegetation indices, and modern architecture Mask R-CNN.

*Keywords:* instance segmentation, machine learning, deep neural networks, olive trees, high resolution images.

# СЕГМЕНТАЦИЯ ОЛИВКОВЫХ ДЕРЕВЬЕВ С ИСПОЛЬЗОВАНИЕМ АРХИТЕКТУРЫ MASK R-CNN

Сафонова А.Н.<sup>(1)(2)</sup>, Табик С.<sup>(2)</sup>, Маглинец Ю.А.<sup>(1)</sup>

<sup>(1)</sup> Институт космических и информационных технологий СФУ, г. Красноярск

<sup>(2)</sup> Центр исследований в области информационных технологий и коммуникаций  
Университета Гранады, Гранада, Испания

В этой работе мы представляем результаты сегментации экземпляров оливковых деревьев с использованием изображений высокого разрешения, индексов растительности NDVI и GNDVI и современной архитектуры Mask R-CNN.

*Ключевые слова: сегментация экземпляра, машинное обучение, глубокие нейронные сети, оливковые деревья, снимки высокого разрешения.*

**Введение.** Методы локализации крон деревьев и определения параметров древостоев являются востребованным направлением ДЗЗ и решаются многими научными коллективами. В работе рассматриваются задачи локализации на изображениях земной поверхности деревьев оливы, а также измерения их параметров, с точностью до отдельного дерева. Практическая значимость работы определяется следующими соображениями. Испания занимает лидирующую позицию в мире по производству оливкового масла (44% от общемирового объема производства) [1], площади посадок таковы, что наземные измерения требуют привлечения значительных ресурсов в то время, как своевременное измерение характеристик деревьев позволяет осуществлять прогнозирование урожайности и расчет других экономических параметров.

Методы машинного обучения развиваются бурными темпами. В частности, они демонстрируют высокие результаты в задачах распознавания, классификации и сегментации отдельных объектов на данных дистанционного зондирования Земли [2-4]. В последнее время указанные методы широко используются для решения задач агромониторинга. В настоящей работе представлены результаты применения сверточных нейронных сетей для определения попиксельной маски объектов оливковых деревьев по данным сверхвысокого разрешения.

**Метод исследования.** Одной из ключевых задач мониторинга оливковых полей является задача сегментации изображений, направленной на выявление крон деревьев и других элементов, важных при их распознавании. В работе выполняется сегментация экземпляров (instance segmentation) – задача определения контуров объектов на уровне пикселей анализируемого изображения. Для ее решения мы использовали сверточную нейронную сеть Mask R-CNN (regions with convolutional neural networks) [5]. В данной СНС реализован двухэтапный подход. На первом этапе сканируется входное изображение и генерируются предложения (области, которые могут содержать объект). На втором этапе выполняется классифицирование предложения и формирование ограничивающих рамок и масок.

Mask R-CNN состоит из описанных следующих структурных компонентов.

1. Основа – стандартная СНС ResNet101, на ранних слоях которой обнаруживаются объекты низкого уровня (края и углы), а на более поздних слоях сети обнаруживаются элементы более высокого уровня (дерево, человек, здание). Проходя через нейронную сеть,

изображение преобразуется из матрицы вида  $1024 \times 1024 \times 3$  (RGB) в карту объектов формы  $32 \times 32 \times 2048$ . Эта карта характеристик становится входом для следующих уровней сети. Для улучшения качества извлечения объектов сетью мы использовали метод Feature Pyramid Network (далее FPN), который берет обнаруженные СНС элементы высокого уровня и передает их вниз в нижние слои. Это позволяет функциям на каждом уровне иметь доступ к функциям как нижнего, так и более высокого уровня.

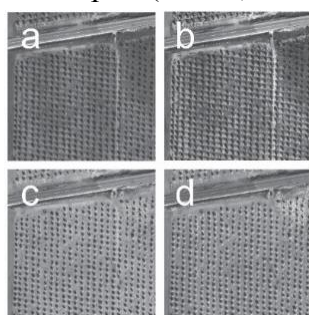
2. Region Proposal Network – СНС, которая сканирует изображение в режиме скользящего окна и находит области, содержащие объекты. В результате сеть генерирует: класс привязки (передний план и фон, где в первом подразумевается присутствие классифицируемого объекта) и ограничительный фрейм с оценкой дельты (процент относительно точности измерения координат, ширины, высоты) для уточнения поля привязки для лучшего соответствия объекту.

3. Классификатор и ограничивающее окно-регрессор. Данный компонент генерирует два выхода для каждого объекта: класс объекта и координаты ограничивающего фрейма, необходимые для дальнейшего уточнения расположения и размера объекта.

Так как для подачи информации на вход классификатора нейронной сети требуются изображения фиксированного размера, мы использовали дополнительный слой подвыборки (субдискретизации), который позволяет обрезать часть карты объекта до фиксированного размера.

4. Сегментные маски – сверточная сеть, которая берет выделенные классификатором области объектов, и генерирует для них маски. Сгенерированные маски имеют низкое разрешение:  $28 \times 28$  пикселей. Во время обучения мы уменьшаем размер маски до  $28 \times 28$  пикселей, чтобы вычислить потери, а во время вывода мы увеличиваем предсказанные маски до размеров ограничивающего фрейма объекта, и они являются окончательными масками, по одной на объект.

**Исходные данные для эксперимента.** Тестовый участок расположен на севере города Гранады, Андалусия, Испания. Объектом исследования выступает дерево *Picular*. Это культивируемый сорт оливок, используемый для изготовления оливкового масла (*Marteño, Nevadillo, Lopereño*). Деревья были высажены в 2006 году. Количество деревьев на испытательном участке составляло 2700 штук, которыми занималась площадь 12 Га, но в общей сложности на плантации произрастает 11000 деревьев (общая площадь 50 Га). Для проведения эксперимента использовались материалы съемки с беспилотного летательного аппарата (далее БПЛА). Мы использовали два БПЛА: PARROT DISCO-PRO AG с многоспектральным датчиком Parrot Sequoia в четырех диапазонах (зеленый (Green), красный (Red), крайний красный (Red\_Edge), ближний инфракрасный каналы (NIR)) и DL-Phantom 4 Pro с RGB камерой (Рис. 1).



а)



б)

Рис. 1. Фрагменты ортоизображений тестового участка. *a* – четырехканальное изображение (*a* – Green, *b* – Red, *c* – Red\_Edge, *d* – NIR), *б* – RGB изображение.

Четырехканальные изображения сверхвысокого разрешения были получены в феврале 2019 года с разрешением 13 см/пиксель, а цветное изображение было получено в июле 2019 года с разрешением 3 см/пиксель. Оба полета дронов производились на высоте 120 метров. Также в качестве эксперимента мы провели расчеты вегетационных индексов (далее ВИ): нормализованный вегетационный индекс растительности NDVI (1) и зеленый нормализованный вегетационный индекс растительности GNDVI (2) по данным четырехканального изображения. GNDVI является показателем фотосинтетической активности растительного покрова, наиболее часто он используется при оценке влагосодержания и концентраций азота в листьях растений. GNDVI более чувствителен к концентрациям хлорофилла по сравнению с индексом NDVI. (Рис. 2) [6].

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (1)$$

$$GNDVI = \frac{NIR - Green}{NIR + Green} \quad (2)$$

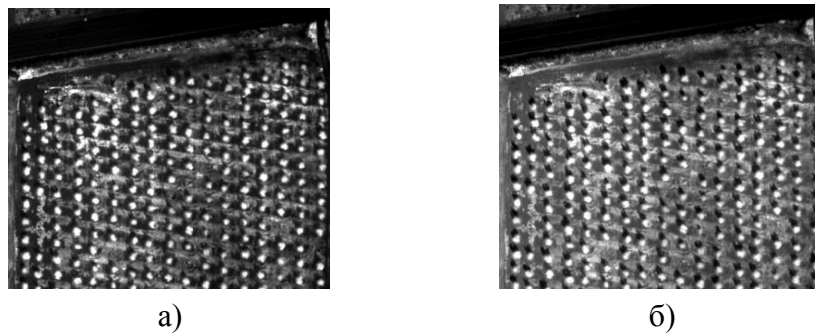


Рис. 2. Фрагмент ортоизображений рассчитанных ВИ NDVI (*a*) и GNDVI (*б*).

Таким образом, для проведения экспериментов мы использовали следующие комбинации каналов: а) RGB изображение, б) Green-Red-NIR, в) NDVI, и г) GNDVI.

На выбранных комбинациях далее велась подготовка набора данных для обучения и тестирования архитектуры Mask R-CNN. Было подготовлено 150 изображений (патчей) из которых 80% было предназначено для обучения модели и 20% для валидации и тестирования обученной модели. Патч представляет собой изображение, на котором представлено от одного до восьми деревьев. Общая схема создания набора данных представлена на Рис. 3.

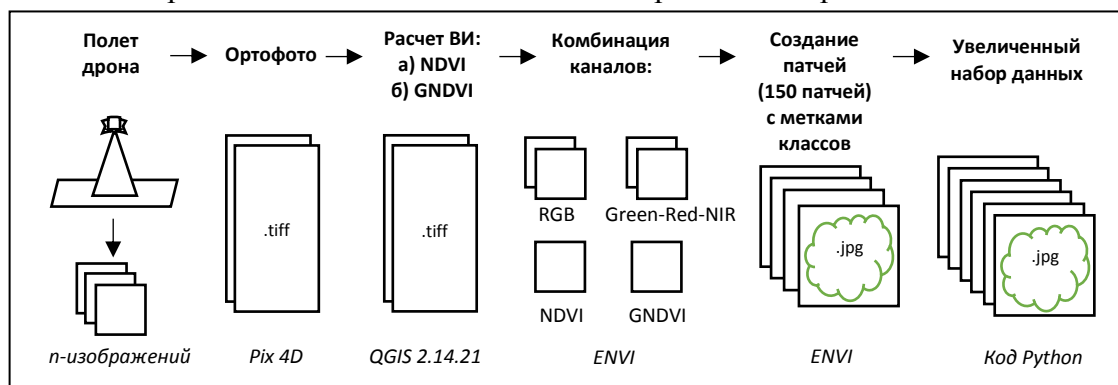


Рис. 3. Процесс получения и создания набора данных для обучения и тестирования модели.

Подготовленный набор патчей был конвертирован в одноканальные изображения в формате .jpg для дальнейшего искусственного увеличения данных с помощью применения таких функций как: поворот изображения по горизонтали на 99%, поворот изображения по вертикали на 99%, улучшение или ухудшение контрастности изображения, увеличение резкости, изменение яркости изображения на 10-150% от исходного изображения. Тестовый набор данных был искусственно увеличен до 1500 изображений для каждой комбинации набора данных.

**Результаты экспериментов.** В данном разделе представлены результаты экспериментальных исследований применения Mask R-CNN в задаче сегментации оливковых деревьев на снимках сверхвысокого разрешения, полученных с помощью БПЛА.

Основная работа проводилась с помощью языка программирования Python и модуля машинного обучения TensorFlow Object Detection API. TensorFlow – программная библиотека с открытым исходным кодом для высокопроизводительных численных вычислений, которая работает в различных гетерогенных системах, включая кластеры распределенных графических процессоров (GPU). Расчеты выполнялись на ПК с процессором Intel Xeon E5-2630v4, ускоренным с помощью графического процессора NVIDIA Titan Xp в качестве платформы для обучения и тестирования предложенной методики. Для подготовки набора изображений были использованы следующие программы: Pix 4D, QGIS 2.14.21, ENVI, VGG Image Annotator. VGG Image Annotator является автономным программным обеспечением для ручного аннотирования изображений, аудио и видео. Создание меток классов осуществлялось в программе VGG Image Annotator с сохранением их в формате JSON (пример ручной отрисовки меток на изображении RGB представлен на Рис. 4).



Рис. 4. Создание меток классов «Оливковые деревья» и «Тень» на изображении RGB.

Далее архитектура Mask R-CNN обучалась в течении 100 эпох для каждой комбинации каналов в течении 60 часов. Результаты сегментации экземпляров оливковых деревьев для каждой комбинации изображений представлены на Рис. 5.

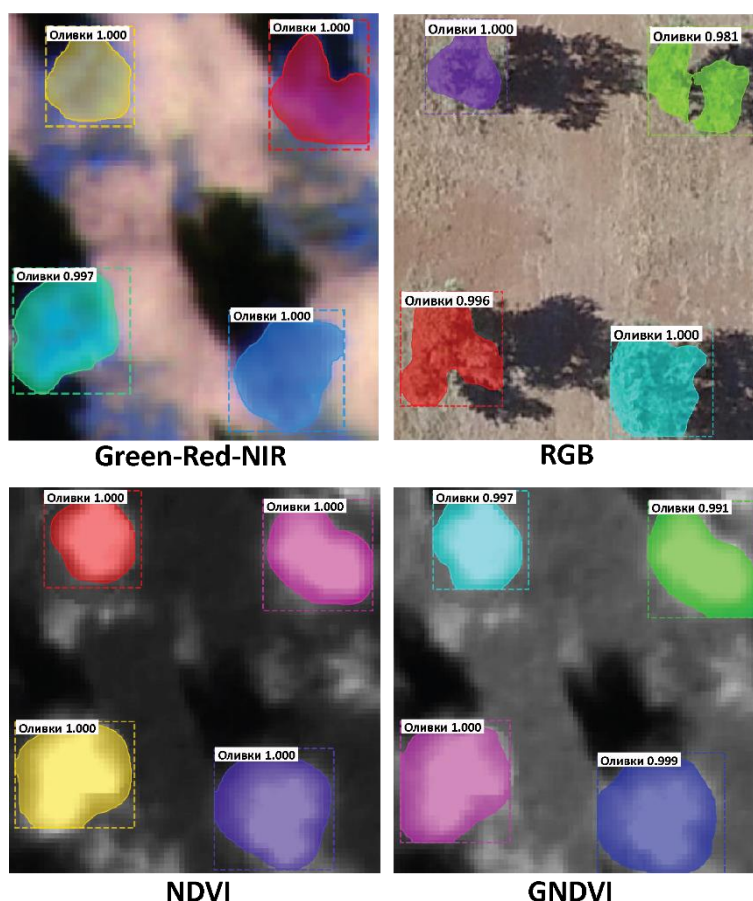


Рис. 5. Результат сегментации класса «Оливковые деревья» с использованием Mask R-CNN на увеличенных наборах данных.

Как видно из Рис. 5, результаты сегментации и построения пиксельных масок для каждого объекта на изображениях дают среднюю точность 99.66%.

В качестве эксперимента нами было проведено дополнительное обучение СНС для сегментации класса «тени оливковых деревьев» и расчета размеров каждого дерева. Результаты выполненных расчетов совпадают с наземными измерениями 10 тестовых деревьев в среднем на 95%. Таким образом, нам удалось рассчитать:

- попиксельно площадь объекта, учитывая, что размер одного пикселя в маске соответствует 3 см для RGB изображений и 13 см для изображения Green-Red-NIR, NDVI, GNDVI;
- высоту дерева ( $h$ ) по размеру тени дерева в метрах ( $l_T$ ) и с учетом высоты солнца над горизонтом в зависимости от географического положения (широты и долготы), даты и времени съемки ( $tg\alpha$ ) (3) [7].

$$h = l_T * tg\alpha \quad (3)$$

- объем кроны дерева, без учета нижней части ствола дерева, на котором нет веток (в среднем до 0.5 метров) (Рис. 6).



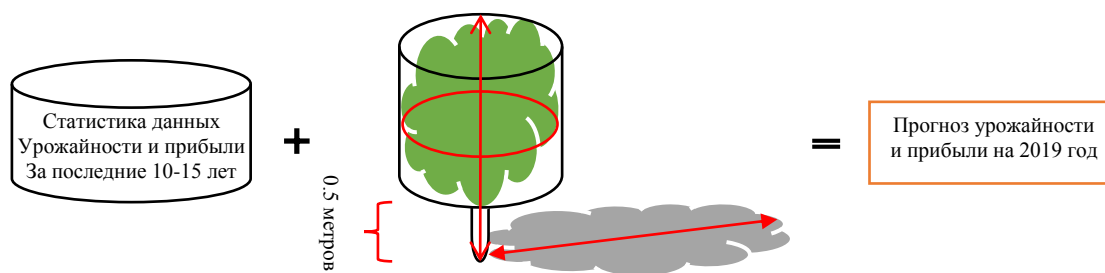


Рис. 6. Графическое представление расчета параметров оливковых деревьев.

Предложенные расчеты возможно использовать для проведения дополнительных полевых измерений и экспериментов по прогнозированию урожайности и прибыли, получаемой с каждого дерева, на основе статистических данных, существующих за последние 10-15 лет.

**Заключение.** Осуществлено исследование применимости одного класса свёрточных нейронных сетей к задачам обнаружения на изображениях земной поверхности деревьев оливы, а также расчету ряда их параметров, с точностью до отдельно стоящего дерева. Полученные результаты показывают, что использование машинного обучения, и, в частности, глубоких нейронных сетей, дает возможность решить поставленные задачи с высокой точностью. Предложенная методика сегментации оливковых деревьев с помощью современной архитектуры Mask R-CNN и модуля TensorFlow продемонстрировала среднюю точность распознавания в 99.66%. Таким образом, данную методику пиксельной сегментации с использованием машинного обучения можно с высокой эффективностью использовать в задачах сельского и лесного хозяйства. В дальнейшем планируется провести тестирование обученных СНС на спутниковых данных среднего разрешения, что представляет наибольший интерес для использования возможных результатов на больших территориях, а также прогнозирования урожайности и прибыли, полученной с оливковых деревьев.

*Работа выполнена при финансовой поддержке стипендии Президента Российской Федерации для проведения научной стажировки за рубежом на 2018/19 учебный год от 07.06.2018 г. № 05-1997 и Российского фонда фундаментальных исследований (номер проекта: 18-47-242002 р\_мк), Правительства Красноярского края, Красноярского краевого фонда науки в рамках научного проекта: «Разработка технологии создания интеллектуальных информационных систем объектно-ориентированного мониторинга территорий по данным дистанционного зондирования».*

#### ЛИТЕРАТУРА

- [1] EU olive oil farms report // European commission. Greece, Italy, Spain. 2012 <http://www.fao.org/family-farming/detail/en/c/326556/> (дата обращения 25.06.2019).
- [2] Зарубин О.А. Применение нейронных сетей для целей анализа данных дистанционного зондирования Земли // Электронный научно-практический журнал «Современные научные исследования и инновации». 2016. №8. <http://web.snauka.ru/issues/2016/08/70887> (дата обращения 25.06.2019).
- [3] Gonzalez-Fernandez I., Iglesias-Otero M.A., Esteki M., Moldes O.A., Mejuto J.C., Simal-Gandara J. A critical review on the use of artificial neural networks in olive oil production, characterization and

authentication // Critical Reviews in Food Science and Nutrition. № 59. 2018. <https://doi.org/10.1080/10408398.2018.1433628> (дата обращения 25.06.2019).

- [4] *Safonova, A., Tabik, S., Alcaraz-Segura, D., Rubtsov, A., Maglinets, Y., & Herrera, F.* Detection of Fir Trees (*Abies sibirica*) Damaged by the Bark Beetle in Unmanned Aerial Vehicle Images with Deep Learning. *Remote Sensing*, 2019. 11(6), 643
- [5] *S. Ren, K. He, R. Girshick, J. Sun*, “Faster R-CNN: Towards real-time object detection with region proposal networks,” in *Neural Information Processing Systems (NIPS)*, 2015.
- [6] Растительность // Геоаналитика.Агро. <http://agro.geoanalitika.com/ru/products/rastitelnost/#> (дата обращения 25.06.2019).
- [7] *Selly H.E.* Tree Heights from shadows. [https://www.asprs.org/wp-content/uploads/pers/1942journal/jun/1942\\_jun\\_100-109.pdf](https://www.asprs.org/wp-content/uploads/pers/1942journal/jun/1942_jun_100-109.pdf) (дата обращения 25.06.2019).