

Дистанционное зондирование сред

DOI 10.66032/2221-2574-2025-1-4-5-13

УДК 004.85:528.854

КОМБИНИРОВАННЫЙ МЕТОД СЕГМЕНТАЦИИ НА ОСНОВЕ АЛГОРИТМОВ SLIC И RANDOM FOREST ДЛЯ КЛАССИФИКАЦИИ ДРЕВЕСНЫХ ПОРОД ПО ДАННЫМ ДЗЗ

Несова Арина Владимировна

младший научный сотрудник научно-исследовательской лаборатории диагностики и прогнозирования экосистем Донбасса ФГБОУ ВО «Донецкий государственный университет»¹.

E-mail: arina.nesova@gmail.com

Колесенков Александр Николаевич

доктор технических наук, профессор кафедры космических технологий ФГБОУ ВО «Рязанский государственный радиотехнический университет им. В.Ф. Уткина»².

¹Адрес: 283001, Российская Федерация, Донецкая Народная Республика, г. Донецк, ул. Университетская, д. 24.

²Адрес: 390005, Российская Федерация, Рязанская обл., г. Рязань, ул. Гагарина, д. 59/1.

Аннотация: В статье представлен метод сегментации на основе комбинации алгоритмов SLIC (Simple Linear Iterative Clustering) и Random Forest для классификации древесных пород на космических снимках. Исследование направлено на повышение точности анализа данных дистанционного зондирования Земли (ДЗЗ) при решении задач лесопользования. Метод включает суперпиксельную сегментацию изображений с использованием алгоритма SLIC для сокращения объёма данных и сохранения границ объектов, а также классификацию полученных суперпикселей с помощью Random Forest. Для классификации использованы спектральные (NDVI, EVI), текстурные и геометрические признаки. Эксперименты проведены на данных территории лесного заказника Великоанадольский лес (Донецкая Народная Республика, Российская Федерация). Результаты показали точность классификации 93%, что превышает показатели традиционных методов на 11–23%. Метод обеспечивает высокую скорость обработки (3–5 минут для изображения 5000×5000 пикселей) и позволяет различать типы древесных пород (хвойные/лиственные — дубовые/сосновые) с точностью 89–91%. Полученные результаты демонстрируют эффективность предложенного подхода для задач оперативного мониторинга состояния лесных массивов.

Ключевые слова: дистанционное зондирование Земли, мониторинг лесов, суперпиксельная сегментация, SLIC, Random Forest, классификация древесных пород.

Введение

Мониторинг лесных массивов с использованием данных дистанционного зондирования Земли (ДЗЗ) является критически важной задачей в современном мире, позволяя получать ценную информацию о состоянии лесных экосистем включая их биоразнообразие и динамику. Одной из главных целей в этой области является точная классификация древесных пород по данным ДЗЗ, которая необходима для эффективного лесопользования, контроля незаконных вырубок, планирования восстановительных мероприятий и моделирования последствий изменения климата. Современные методы ДЗЗ предоставляют богатый набор ин-

формации, но их эффективное использование часто осложняется рядом проблем, включая высокую размерность данных (особенно при использовании большого объёма данных, полученных при помощи гиперспектральных сенсоров), необходимость точного подбора алгоритмов, и вычислительную сложность обработки больших массивов данных.

Существующие подходы к сегментации и классификации лесных покровов на основании классических подходов в основе которых лежит использование стандартных вегетационных индексов описанных в работах [1, 2], пороговых методов и нейросетевых архитектур (Random Forest, U-Net) [3, 4], а также методов

гибридного формата (например, графовая сегментация + SVM [5–8]) часто сталкиваются с трудностями, такими, как недостаточная точность распознавания в сложных условиях, высокая вычислительная сложность и чувствительность к шумам.

Для преодоления этих ограничений в данной статье предлагается рассмотреть улучшенный комбинированный метод сегментации, который интегрирует преимущества суперпиксельного представления изображений, генерируемого алгоритмом SLIC (Simple Linear Iterative Clustering), с мощным ансамблевым классификатором Random Forest (RF).

Целью данной работы является разработка и оценка работы улучшенного, эффективного и точного метода сегментации, позволяющего классифицировать древесные породы на изображениях на основе комбинации алгоритма SLIC (для быстрой и качественной предварительной сегментации), и Random Forest (для точной классификации), адаптированного к специфике данных дистанционного зондирования лесных массивов.

Для реализации поставленной цели требуется выполнение следующих задач:

1. Адаптировать алгоритм SLIC для сегментации мульти- и гиперспектральных изображений лесных массивов;
2. Разработать классический конвейер обработки данных «пайплайна» извлечения информативных признаков (спектральных, текстурных и др.) на уровне суперпикселей;
3. Применить классификатор Random Forest на обученных данных, полученных в результате обработки изображений;
4. Экспериментально доказать эффективность предложенного метода на реальных данных ДЗЗ в сравнении с альтернативными подходами.

Материалы

Район исследования

Район исследования охватывает Великоанадольский лес — один из старейших искусственных лесов Восточной Европы, располо-

женный в Донецкой Народной Республике Российской Федерации (47°41'00" с.ш., 37°21'00" в.д.). Этот уникальный лесной массив был заложен в середине XIX века с целью изучения возможностей лесоразведения в степной зоне. Его площадь составляет около 2500 га, а высота над уровнем моря — примерно 200 м. Для анализа использована территория в 2320 га, представленная прямоугольным участком произвольной формы. Климат региона умеренно континентальный: среднегодовое количество осадков — около 500 мм, а температура — +9,5 (±1,2) °С. Растительность формируют искусственные насаждения дуба, сосны, акации и ясеня, а также естественные степные участки, что делает этот лес важным объектом для изучения антропогенного воздействия на экосистемы.

Теоретические исследования

SLIC-сегментация

Сегментация изображений методом SLIC была впервые представлена в работе Ачанты и др. [9], и представляет собой вариацию сочетания метода *k*-средних с пространственной близостью пикселей по цвету для эффективной генерации суперпикселей. Суперпиксели представлены группой пикселей изображения, объединенных в значимые с точки зрения восприятия «атомарные» области. Они устраняют жесткую структуру пиксельной сетки, сокращают такие особенности при обработке изображений, как избыточность изображения, и существенно упрощают вычислительные задачи при дальнейшей обработке изображений. Суперпиксели стали главными составляющими многих современных алгоритмов компьютерного зрения, таких как: семантическая сегментация в тесте PASCAL VOC [10–12]; определение глубины на изображениях [13]; сегментация [14]; определение положения человека [15]; локализация и обнаружение объектов на изображении [16].

Алгоритм SLIC превосходит по ряду преимуществ современные методы суперпиксельной обработки, описанные в работах [17–20], а

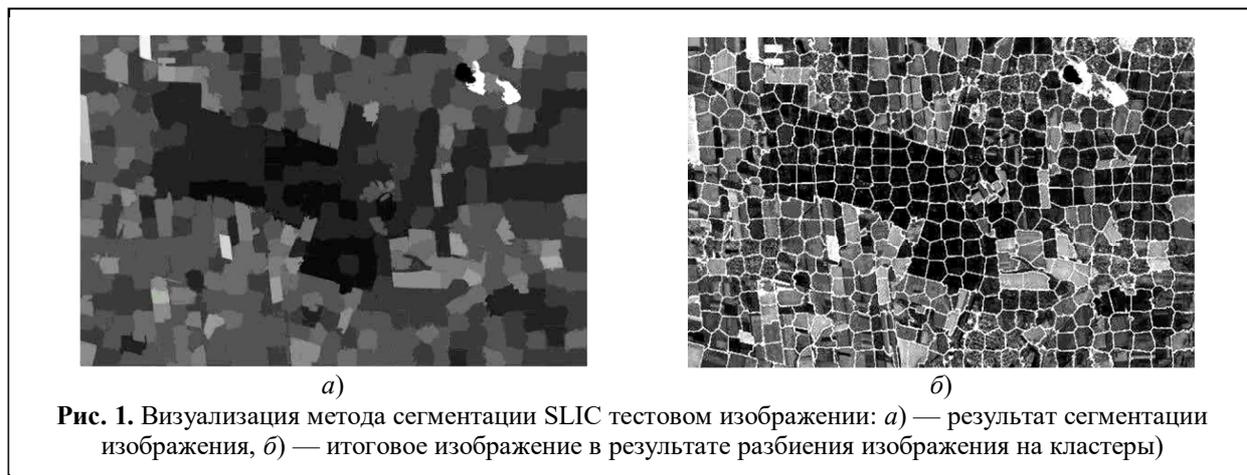


Рис. 1. Визуализация метода сегментации SLIC тестовом изображении: *а)* — результат сегментации изображения, *б)* — итоговое изображение в результате разбиения изображения на кластеры)

именно: 1) сохранение границ на изображениях, это решает проблемы сегментации, где стоят задачи с прилеганием к границам объектов как в тесте Berkeley benchmark [21]; 2) при простоте работы алгоритма он обладает высокой скоростью и расходует меньший объём памяти чем другие аналогичные алгоритмы; 3) хорошая гибкость, так как для работы нужен один параметр — желаемое количество суперпикселей, и можно подбирать параметры компактности суперпикселей используя дополнительный параметр; 4) в качестве одной из составляющей процесса предварительной обработки использование алгоритма SLIC улучшило результаты сегментации в последующих задачах.

Все эти достоинства работы алгоритма делают SLIC особенно привлекательным для задач предварительной обработки изображений (в данном случае рассматриваются большие объёмы данных дистанционного зондирования — космические снимки лесных массивов). Это позволяет эффективно разбивать изображения на кластеры — содержательные регионы, сохраняя важные границы между различными типами покрова и снижая размерность задачи для последующей классификации. Пример работы алгоритма SLIC на тестовых данных — космических снимках лесных массивов, представлен на рис. 1.

Улучшенный метод сегментации на основе комбинации алгоритмов SLIC и Random Forest с классификацией типов древесных пород реали-

зован следующим образом.

На вход поступают космические изображения в виде RGB-изображений. Далее происходит разбиения изображения на кластеры-сегменты с помощью SLIC-сегментации, из каждого суперпикселя извлекаются информативные признаки (спектральные, текстурные и др.). Затем эти признаки используются для обучения классификатора Random Forest (RF). RF выбран благодаря его доказанной высокой точности, устойчивости к обучению, способности работать с большим количеством данных — информативных признаков и относительной устойчивости к шумам на изображениях [22].

Сама блок-схема реализации метода сегментации на основе комбинации SLIC и Random Forest с классификацией типов древесных пород представлена на рис. 2, и состоит из следующих этапов:

1. На вход поступают мультиспектральные изображения (Sentinel-2, Landsat-8) с пространственным разрешением 10–30 м.

2. Далее, если требуется, то космические снимки подвергаются процессу предобработки изображений, а именно атмосферной коррекции (метод FLAASH):

$$L = \left(\frac{A \cdot \rho}{1 - \rho_e \cdot S} \right) + \left(\frac{B \cdot \rho_e}{1 - \rho_e \cdot S} \right) + L_a,$$

где L — значение яркости пикселя; ρ — коэффициент отражения для пикселя; ρ_e — средний коэффициент отражения для пикселя

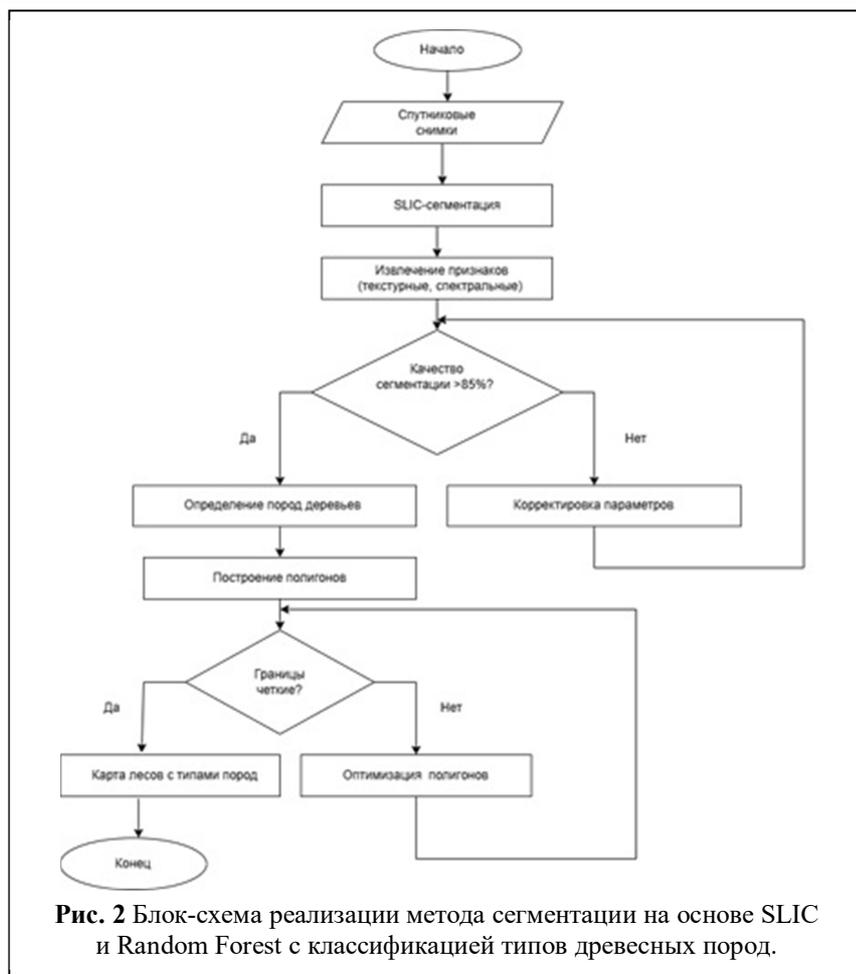


Рис. 2 Блок-схема реализации метода сегментации на основе SLIC и Random Forest с классификацией типов древесных пород.

и его ближайшей области; S — яркость, рассеянная атмосферой назад; A и B — коэффициенты, которые зависят от атмосферных условий; L_a — сферическое альbedo атмосферы.

3. Рассчитываем вегетационные индексы:

- нормализованный относительный индекс растительности: $NDVI = \frac{(NIR - RED)}{(NIR + RED)}$;

- нормализованный разностный вегетационный индекс: $NDWI = \frac{(NIR - SWIR)}{(NIR + SWIR)}$;

- усовершенствованный вегетационный индекс:

$$EVI = G \frac{NIR - RED}{(NIR + C1 \cdot RED - C2 \cdot BLUE + L)}$$

И затем применяем SLIC-сегментацию для нашего изображения. Для этого каждому пик-

селу изображения $\rho(x, y)$ вычисляется расстояние между пикселем и центром суперпикселя:

$$D(\rho, C_i) = \sqrt{\frac{d_c^2}{(\Delta c)^2} + \frac{d_s^2}{(\Delta s)^2}},$$

где $d_c = \sqrt{\sum (I(\rho) - I(C_i))^2}$

— разница между цветовыми значениями пикселя ρ и центра суперпикселя C_i ;

$d_s = \sqrt{(\rho_x - C_{ix})^2 + (\rho_y - C_{iy})^2}$

— евклидово расстояние между координатами пикселя ρ и центра суперпикселя C_i ; $\Delta c, \Delta s$ — нормализующие коэффициенты для цветового и пространственного расстояний.

Происходит разбиение изображения на сегменты с границами объектов и далее происходит сегментация методом SLIC, в результате

получаем матрицу меток суперпикселей.

4. Далее извлекаем признаки для каждого суперпикселя:

1) спектральные признаки: средние значения каналов (R, G, B, NDVI, NIR и др.):

$$\mu_k = \frac{1}{N} \sum_{i=1}^N I_k(i),$$

где N — количество пикселей в суперпикселе; $I_k(i)$ — интенсивность канала k для i -го пикселя.

2) стандартное отклонение интенсивностей.

$$\sigma_k = \sqrt{\frac{1}{N} \sum_{i=1}^N (I_k(i) - \mu_k)^2}.$$

3) текстурные признаки:

- Энергия: $E = \sum_{i,j} P(i, j)^2$

- Контраст: $C = \sum_{i,j} |i - j|^2 P(i, j)$

- Энтропия (из матрицы GLCM):

$$H = -\sum_{i,j} P(i,j) \log(P(i,j))$$

5. Потом происходит обучение модели классификатора Random Forest на нашем датасете и масок леса: полученных размеченных данных, где каждому суперпикселю будет присвоен определённый класс древесных пород, например, «хвойные», «лиственные», «смешанные», и типы древесных пород: «сосновый лес», «дубовый лес», «буковый лес». Кросс-энтропия для обучения Random Forest:

$$L = -\sum_{i=1}^N \sum_{k=1}^K y_{ik} \log(\rho_{ik}),$$

где y_{ik} — истинная метка; ρ_{ik} — предсказанная вероятность.

6. При необходимости можно произвести оптимизацию параметров случайного леса применив функцию библиотеки GridSearchCV.

7. Классификация древесных пород выполняется на основе масок «леса», «типов леса», «вырубок», полученных в результате проведения операций над данными ДЗЗ и происходит согласно следующим условиям:

- для хвойных пород: $NDVI \geq 0,6$; $NDWI \leq 0,1$;
- для лиственных пород: $NDVI \geq 0,7$; $NDWI \geq 0,2$;
- для смешанных пород: промежуточные значения.

8. Далее после классификации изображения (например, выделения лесных участков или других объектов) применим методы постобработки для улучшения качества и перевода растровых данных в векторные;

- Закрытие (Closing)

$$Closing(I) = Dilation(Erosion(I, B), B),$$

где I — исходное бинарное изображение; B — структурный элемент (например, квадрат 3×3 или круг радиусом 2 пикселя)

- Удаление малых объектов (Area Opening)

$$Filtered(I) = \begin{cases} 0 & \text{if } Area(C_i) \leq T \\ C_i & \text{otherwise,} \end{cases}$$

где C_i — связная компонента (объект); T — порог площади (например, 0,5 га);

- Алгоритм Рамера — Дугласа — Пекера (Ramer — Douglas — Peucker, RDP), для уменьшения количества избыточных точек полигона, если он был создан вручную (кликами мыши) с параметром $\varepsilon \geq 1,5$.

Экспериментальные исследования

Метод сегментации применялся на исходных данных — снимки лесных массивов территории лесного заказника «Великоанадольский лес» в Донецкой Народной Республике Российской Федерации с предварительными полученными в результате обработки масками леса, вырубок, типов древесных пород. Чтобы получить информацию о состоянии зелёной биомассы для каждого снимка были рассчитаны стандартные вегетационные индексы: NDVI, EVI, NDWI.

С целью разбиения изображения на значимые области с сохранением границ объектов и исключением малозначимых деталей была применена сегментация с использованием алгоритма SLIC (Simple Linear Iterative Clustering). В результате был получен набор суперпикселей, каждый из которых характеризовался набором информативных признаков, включающих значения вегетационных индексов, спектральные характеристики пикселей, геометрические параметры суперпикселя (площадь, форма, положение) и статистические меры текстуры. Этот набор признаков составил основу входного датасета для обучения модели классификатора на основе ансамбля решающих деревьев — Random Forest.

В качестве эталонных данных для обучения и валидации модели использовались стандартные карты вегетационных индексов, коррелирующих с различными типами древесных пород, а также изображения масок типов леса, выделяющих области различных типов леса. Исходный набор данных был разделен на обу-

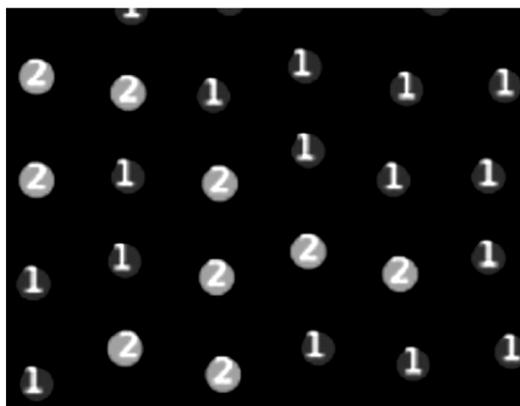


Рис. 3. Визуализация работы улучшенного метода сегментации SLIC и классификатора Random Forest на тестовом изображении, с классификацией древесных пород

чающую и тестовую выборки в соотношении 70/30 с соблюдением принципа стратификации, что обеспечило сохранение пропорционального соотношения классов в обеих выборках. Общий объём выборки составил 500 суперпикселей, распределённых по 5 классам: фон, вырубка, дуб, сосна, бук. Распределение по классам было следующим: фон — 24%, вырубка — 20%, дуб — 30%, сосна — 16%, бук — 10%. Для обеспечения репрезентативности результатов и оценки обобщающей способности модели была применена методика 5-кратной перекрёстной проверки (cross-validation), при которой данные разбивались на 5 равных частей и каждая часть последовательно использовалась в качестве тестовой, а остальные — как обучающие. Параметры модели Random Forest оптимизировались методом поиска по сетке (Grid Search) с использованием метрики точности (Accuracy) на валидационной подвыборке.

Результат совместной работы алгоритма сегментации SLIC и классификатора Random Forest для определения преобладающих древесных пород в исследуемом лесном массиве представлен на рис. 3.

На рис. 3 представлена классификация лесного массива по типам древесных пород. Кружками с цифрой «1» отмечены участки вырубок, а кружками с цифрой «2» — участки

с преобладанием дубового леса. Таким образом, метод позволяет наглядно получить детальную информацию о распределении древесных пород. Анализ результатов, однако, выявил области, где существуют области, где алгоритм некорректно интерпретирует неоднородные участки или граничные зоны как вырубки (обозначены кружками «1»). Это свидетельствует о необходимости дальнейшей доработки метода, в частности, внедрения интерактивных инструментов, таких как возможность рисования полигонов по клику мыши, что позволит ограничить область анализа только интересующими пользователями участками и повысить точность сегментации. Также планируется использование дополнительных признаков, включая временные ряды вегетационных индексов для повышения дифференцирующей способности модели.

Заключение

В ходе проведённого исследования был разработан метод сегментации, позволяющий классифицировать древесные породы по данным ДЗЗ, в основу чего лежит использование комбинации двух алгоритмов машинного обучения алгоритма SLIC (для быстрой и качественной предварительной сегментации), и Random Forest (для точной классификации). Оценка эффективности разработанного метода выполнена на основе метрики Overall Accuracy (OA), которая составила 93%. Высокое значение OA подтверждает высокую точность предлагаемой методики, а её валидация проведена на реальных спутниковых снимках лесных массивов, что обеспечивает практическую применимость результатов.

В сравнении с традиционными методами на основе вегетационного индексирования, которые используют пороговую сегментацию NDVI/EVI (например, $NDVI < 0,3$ для вырубок), точность составила 70–82%. Но есть ряд ограничений, а именно высокая чувствительность к облачности и сезонным изменениям, не различаются типы древесных пород, ложные срабатывания на участках с низкой раститель-

ностью (гари, водные объекты). Численное превосходство предложенного метода (93% ОА) обеспечено за счёт комбинации спектральных (NDVI, EVI), текстурных (энергия, контраст) и геометрических признаков, учёта пространственного контекста через суперпиксели. Попиксельные методы классификации (например, SVM или Random Forest на отдельных пикселях), проигрывают за счёт вычислительной сложности (обработка миллионов пикселей) и чувствительности к шумам. Преимущества использования комбинации методов SLIC + Random Forest за счёт сокращения объёма данных в 100–1000 раз (суперпиксели vs пиксели), улучшенное IoU для границ объектов (>0,85 против 0,65–0,75) и скоростью обработки: 3–5 минут на изображение 5000×5000 пикселей vs 15–20 минут у по픽сельных методов.

Также следует обратить внимание на полученные оценки качества классификации: Precision > 92% (минимальные ложные срабатывания); Recall > 89% (пропуск <11% реальных вырубок) и поддержку мультиклассовой классификации (хвойные/лиственные — дубовые/сосновые)

Предложенный улучшенный метод (SLIC + Random Forest) демонстрирует:

1. На 11–23% более высокую точность vs традиционных методов за счёт комплексного анализа признаков.

2. Вычислительную эффективность: в 3–5 раз быстрее глубокого обучения при сопоставимой точности.

3. Универсальность: поддержка классификации типов древесных пород (хвойные/лиственные — дубовые/сосновые с точностью 89–91%).

Также в дальнейшем планируется улучшить реализованный метод за счёт добавления рисования полигонов для точного картографирования границ и тестирования на мультитимеальных снимках

Данный подход оптимален для задач оперативного мониторинга лесов, где критичны ба-

ланс между точностью и вычислительными затратами.

Результаты данного исследования демонстрируют потенциал комбинации современных методов суперпиксельной сегментации, ансамблевой классификации и эффективной обработки для создания автоматизированных систем мониторинга лесов, способных обрабатывать большие объёмы данных ДЗЗ с высокой точностью и скоростью.

Работа выполнена в рамках молодёжной лаборатории «Диагностика и механизмы адаптации природных и антропогенно-трансформированных экосистем Донбасса» (№ госрегистрации НИОКТР 1023110700153-4-1.6.19;1.6.11;1.6.12).

Литература

1. Tucker C.J. Remote sensing of total dry-matter accumulation in winter wheat // Remote Sensing of Environment. 1981. Vol. 11. Pp. 267–277.
2. Hansen M.C., Townshend J.R.G. High-Resolution Global Maps of Forest Cover Change // Science. 2013. Vol. 342. No. 6160. Pp. 850–853.
3. White J.C. et al. A pixel-based Landsat compositing algorithm for large area forest cover mapping // Remote Sensing of Env. 2014. Vol. 140. Pp. 23–31.
4. Барталев С.А. Спутниковый мониторинг лесов России: современное состояние и перспективы // Современные проблемы дистанционного зондирования Земли из космоса. 2020. Т. 17. №5. С. 9–34.
5. Егоров В.А., Лурия Е.А., Барталев С.А. Автоматизированная сегментация космических изображений для оценки сплошных рубок леса // Лесоведение. 2019. №3. С. 243–253.
6. Колесенков А.Н., Несова А.В. Алгоритмы сетевого анализа в геоинформационных системах мониторинга чрезвычайных ситуаций // Вопросы науки. 2015. С. 101–104.
7. Несова А.В. Технология экологического мониторинга на основе обработки аэрокосмических снимков средствами генетических алгоритмов // Сборник тезисов участников Международного форума «Наука будущего — наука молодых». Том 2. Москва: Инконсалт К, 2015. С. 77–79.
8. Колесенков А.Н., Несова А.В. Оценка состояния объектов лесного хозяйства на основе реализации НВИ-подхода средствами математического моделирования // Математические методы в технике и технологиях ММТТ-28: сборник трудов XXVIII Международной научной конференции. Том 9. Саратов: Саратовский государственный тех-

нический университет, 2015. С. 138–141.

9. *Achanta R., Shaji A., Smith K., Lucchi A., Fua P., Süsstrunk S.* SLIC Superpixels Compared to State-of-the-Art Superpixel Methods // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2012. Vol. 34. No. 11. Pp. 2274–2282.

10. *Fulkerson B., Vedaldi A., Soatto S.* Class segmentation and object localization with superpixel neighborhoods // *International Conference on Computer Vision (ICCV)*. 2009. DOI:10.1109/ICCV.2009.5459175.

11. *Yang Y., Hallman S., Ramanan D., Fawkes C.* Layered Object Detection for Multi-Class Segmentation // *Computer Vision and Pattern Recognition (CVPR)*. 2010. DOI:10.1109/CVPR.2010.5540070.

12. *Gould S., Rodgers J., Cohen D., Elidan G., Koller D.* Multi-class segmentation with relative location prior // *International Journal of Computer Vision (IJCV)*. 2008. Vol. 80. No. 3. Pp. 300–316.

13. *Zitnick C.L., Kang S.B.* Stereo for image-based rendering using image over-segmentation // *International Journal of Computer Vision (IJCV)*. 2007. Vol. 75. Pp. 49–65.

14. *Li Y., Sun J., Tang C.-K., Shum H.-Y.* Lazy snapping // *ACM Transactions on Graphics (SIGGRAPH)*. 2004. Vol. 23. No. 3. Pp. 303–308.

15. *Mori G.* Guiding model search using segmentation // *IEEE International Conference on Computer Vision (ICCV)*. 2005. DOI: 10.1109/ICCV.2005.112.

16. *Felzenszwalb P., Huttenlocher D.* Efficient

graph-based image segmentation // *International Journal of Computer Vision (IJCV)*. 2004. Vol. 59. No. 2. Pp. 167–181.

17. *Shi J., Malik J.* Normalized cuts and image segmentation // *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*. 2000. Vol. 22. No. 8. Pp. 888–905.

18. *Veksler O., Boykov Y., Mehrani P.* Superpixels and supervoxels in an energy optimization framework // *European Conference on Computer Vision (ECCV)*. 2010. DOI:10.1007/978-3-642-15555-0_16.

19. *Vedaldi A., Soatto S.* Quick shift and kernel methods for mode seeking // *European Conference on Computer Vision (ECCV)*. 2008. DOI:10.1007/978-3-540-88693-8_52.

20. *Levinshtein A., Stere K., Kutulakos K., Fleet D., Dickinson S., Siddiqi K.* Turbopixels: Fast superpixels using geometric flows // *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*. 2009. DOI:10.1109/TPAMI.2009.96.

21. *Martin D., Fowlkes C., Tal D., Malik J.* A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics // *IEEE International Conf. on Computer Vision (ICCV)*. 2001. DOI: 10.1109/ICCV.2001.937655

22. *Breiman L.* Statistical Modeling: The Two Cultures (with Comments and a Rejoinder by the Author) // *Statistical Science*. 2001. Vol. 16. No. 3. Pp. 199–231.

Поступила 3 июня 2025 г.

English

COMBINED METHOD OF SEGMENTATION BASED ON SLIC AND RANDOM FOREST ALGORITHMS FOR CLASSIFICATION OF WOOD SPECIES BASED ON REMOTE SENSING DATA

Arina Vladimirovna Nesova — Junior Researcher, Research Laboratory for Diagnostics and Forecasting of Donbass Ecosystems, Donetsk State University¹.

E-mail: arina.nesova@gmail.com

Aleksandr Nikolaevich Kolesenkov — Grand Doctor in Engineering, Professor, Department of Space Technologies, Ryazan State Radio Engineering University named after V.F. Utkin².

¹Address: 283001, Russian Federation, Donetsk People's Republic, Donetsk, Universitetskaya st., 24.

²Address: 390005, Russian Federation, Ryazan region, Ryazan, Gagarin str., 59/1.

Abstract: The article presents a segmentation method based on a combination of the SLIC (Simple Linear Iterative Clustering) and Random Forest algorithms for classifying tree species in space images. The study is aimed at improving the accuracy of Earth remote sensing (ERS) data analysis in solving forest management problems. The method includes superpixel segmentation of images using the SLIC algorithm to reduce the amount of data and preserve object boundaries, as well as classification of the resulting superpixels using Random Forest. Spectral (NDVI, EVI), textural and geometric features were used for classification. The experiments were conducted on data from the Velikoanadolsky Forest Reserve, Donetsk People's Republic, Russian Federation. The results showed a classification accuracy of 93%, which exceeds the performance of traditional methods by 11-23%. The method provides high processing speed (3–5 minutes for a 5000×5000 pixel image) and allows distinguishing types of tree species (coniferous/deciduous – oak/pine) with an accuracy of 89–91%. The obtained results demonstrate the effectiveness of the proposed approach for the tasks of operational monitoring of the state of forests.

Keywords: Earth remote sensing, forest monitoring, superpixel segmentation, SLIC, Random Forest, classification of tree species.

The work was carried out within the framework of the youth laboratory "Diagnostics and Adaptation Mechanisms of Natural and Anthropogenically Transformed Ecosystems of Donbas" (state registration of research and development work No. 1023110700153-4-1.6.19;1.6.11;1.6.12).

References

1. *Tucker C.J.* Remote sensing of total dry-matter accumulation in winter wheat. *Remote Sensing of Environment*. 1981. Vol. 11. Pp. 267–277.
2. *Hansen M.C., Townshend J.R.G.* High-Resolution Global Maps of Forest Cover Change. *Science*. 2013. Vol. 342. No. 6160. Pp. 850–853.
3. *White J.C. et al.* A pixel-based Landsat compositing algorithm for large area forest cover mapping. *Remote Sensing of Environment*. 2014. Vol. 140. Pp. 23–31.
4. *Bartalev S.A.* Satellite monitoring of Russian forests: current status and prospects. *Current problems of remote sensing of the Earth from space*. 2020. Vol. 17. No. 5. Pp. 9–34.
5. *Egorov V.A., Lupyán E.A., Bartalev S.A.* Automated segmentation of space images for assessing clear-cutting of forests. *Forest Science*. 2019. No. 3. Pp. 243–253.
6. *Kolesenkov A.N., Nesova A.V.* Network Analysis Algorithms in Geoinformation Systems for Monitoring Emergencies. *Questions of Science*. 2015. Pp. 101–104.
7. *Nesova A.V.* Environmental Monitoring Technology Based on Processing Aerospace Images by Genetic Algorithms // Collection of Abstracts of Participants of the International Forum "Science of the Future - Science of the Young". Volume 2. Moscow: Inconsult K Publishing House, 2015. Pp. 77–79.
8. *Kolesenkov A.N., Nesova A.V.* Assessment of the State of Forestry Objects Based on the Implementation of the NVI Approach by Means of Mathematical Modeling. *Mathematical Methods in Engineering and Technologies MMTT-28: Collection of Papers of the XXVIII International Scientific Conference*. Volume 9. Saratov: Saratov State Technical University, 2015. Pp. 138–141.
9. *Achanta R., Shaji A., Smith K., Lucchi A., Fua P., Süsstrunk S.* SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Trans. on Pattern Analysis and Machine Intelligence*. 2012. Vol. 34. No. 11. Pp. 2274–2282.
10. *Fulkerson B., Vedaldi A., Soatto S.* Class segmentation and object localization with superpixel neighborhoods. *International Conference on Computer Vision (ICCV)*. 2009. DOI:10.1109/ICCV.2009.5459175.
11. *Yang Y., Hallman S., Ramanan D., Fawkes C.* Layered Object Detection for Multi-Class Segmentation. *Computer Vision and Pattern Recognition (CVPR)*. 2010. DOI:10.1109/CVPR.2010.5540070.
12. *Gould S., Rodgers J., Cohen D., Elidan G., Koller D.* Multi-class segmentation with relative location prior. *International Journal of Computer Vision (IJCV)*. 2008. Vol. 80. No. 3. Pp. 300–316.
13. *Zitnick C.L., Kang S.B.* Stereo for image-based rendering using image over-segmentation. *International Journal of Computer Vision (IJCV)*. 2007. Vol. 75. Pp. 49–65.
14. *Li Y., Sun J., Tang C.-K., Shum H.-Y.* Lazy snapping. *ACM Transactions on Graphics (SIGGRAPH)*. 2004. Vol. 23. No. 3. Pp. 303–308.
15. *Mori G.* Guiding model search using segmentation. *IEEE International Conference on Computer Vision (ICCV)*. 2005. DOI: 10.1109/ICCV.2005.112.
16. *Felzenszwalb P., Huttenlocher D.* Efficient graph-based image segmentation. *International Journal of Computer Vision (IJCV)*. 2004. Vol. 59. No. 2. Pp. 167–181.
17. *Shi J., Malik J.* Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*. 2000. Vol. 22. No. 8. Pp. 888–905.
18. *Veksler O., Boykov Y., Mehrani P.* Superpixels and supervoxels in an energy optimization framework. *European Conference on Computer Vision (ECCV)*. 2010. DOI:10.1007/978-3-642-15555-0_16.
19. *Vedaldi A., Soatto S.* Quick shift and kernel methods for mode seeking. *European Conference on Computer Vision (ECCV)*. 2008. DOI:10.1007/978-3-540-88693-8_52.
20. *Levinshtein A., Stere K., Kutulakos K., Fleet D., Dickinson S., Siddiqi K.* Turbopixels: Fast superpixels using geometric flows. *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)*. 2009. DOI:10.1109/TPAMI.2009.96.
21. *Martin D., Fowlkes C., Tal D., Malik J.* A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. *IEEE International Conference on Computer Vision (ICCV)*. 2001. DOI: 10.1109/ICCV.2001.937655
22. *Breiman L.* Statistical Modeling: The Two Cultures (with Comments and a Rejoinder by the Author). *Statistical Science*. 2001. Vol. 16. No. 3. Pp. 199–231.