

СЕГМЕНТАЦИЯ СПУТНИКОВЫХ СНИМКОВ SENTINEL И LANDSAT: СОВРЕМЕННЫЕ ПОДХОДЫ И ПЕРСПЕКТИВЫ

SEGMENTATION OF SENTINEL AND LANDSAT SATELLITE IMAGES: CURRENT APPROACHES AND PROSPECTS

**A. Polyakov
K. Polyakova**

Summary. Segmentation of Sentinel and Landsat satellite images represents a relevant task in the field of remote sensing of the Earth. Modern machine learning methods open new possibilities for automated processing of large datasets and extracting valuable information. This article provides a systematic analysis of existing segmentation approaches, including the use of convolutional neural networks, clustering algorithms, and ensemble models. Based on a sample of 2500 high-resolution images obtained from Sentinel-2 and Landsat-8 satellites during the period from 2018 to 2022, a comparative testing of 5 different segmentation models was conducted. The best results were demonstrated by an ensemble of the U-Net convolutional neural network and the K-means clustering algorithm, achieving a segmentation accuracy of 94.7 % according to the IoU metric. A new approach was also proposed, based on the pre-trained EfficientNet-B4 model and an original loss function, Focal Tversky Loss, which allows improving accuracy to 96.2 % while reducing computational complexity. The obtained results have high theoretical and practical significance, opening prospects for the development of intelligent next generation geoinformation systems. Further research should be aimed at developing transfer learning methods and adapting models to images from different sources.

Keywords: image segmentation, satellite images, Sentinel, Landsat, machine learning, neural networks, clustering.

Введение

Сегментация спутниковых снимков является одной из ключевых задач компьютерного зрения в сфере дистанционного зондирования Земли (ДЗЗ). Спутники Sentinel и Landsat обеспечивают регулярное покрытие земной поверхности мультиспектральными снимками высокого разрешения, представляющими огромную ценность для мониторинга окружающей среды, управления природными ресурсами, контроля ЧС и др. [1, с. 32]. Однако эффективное использование этих данных требует развития автоматизированных методов их обработки и интерпретации [2, с. 169].

Существующие подходы к сегментации спутниковых снимков можно условно разделить на две категории: ме-

Поляков Артем Николаевич
Инженер, Вычислительный центр, Дальневосточное отделение, Российская академия наук
artem_polyakov@inbox.ru

Полякова Кристина Эдуардовна
Дальневосточный государственный университет путей сообщения
mini-girl2007@mail.ru

Аннотация. Сегментация спутниковых снимков Sentinel и Landsat представляет собой актуальную задачу в сфере дистанционного зондирования Земли. Современные методы машинного обучения открывают новые возможности для автоматизированной обработки больших массивов данных и извлечения ценной информации. В данной статье проведен системный анализ существующих подходов к сегментации, включая использование сверточных нейронных сетей, алгоритмов кластеризации и ансамблевых моделей. На основе выборки из 2500 снимков высокого разрешения, полученных со спутников Sentinel-2 и Landsat-8 в период с 2018 по 2022 гг., проведено сравнительное тестирование 5 различных моделей сегментации. Лучшие результаты показал ансамбль сверточной нейросети U-Net и алгоритма кластеризации K-means, обеспечивший точность сегментации на уровне 94.7 % по метрике IoU. Также предложен новый подход, основанный на предобученной модели EfficientNet-B4 и оригинальной функции потерь Focal Tversky Loss, позволяющий повысить точность до 96.2 % при снижении вычислительной сложности. Полученные результаты имеют высокую теоретическую и практическую значимость, открывая перспективы для создания интеллектуальных геоинформационных систем нового поколения. Дальнейшие исследования целесообразно направить на разработку методов переноса обучения и адаптации моделей к снимкам из разных источников.

Ключевые слова: сегментация изображений, спутниковые снимки, Sentinel, Landsat, машинное обучение, нейронные сети, кластеризация.

тоды на основе классических алгоритмов компьютерного зрения (пороговая обработка, кластеризация, водоразделы и др.) [3, с. 1867] и методы глубокого обучения, использующие нейросетевые архитектуры (полносверточные сети, U-Net, SegNet и др.) [4]. Первые характеризуются относительной простотой и интерпретируемостью, но сильно зависят от выбора гиперпараметров и слабо обобщаются на реальных данных [5, с. 780]. Вторые позволяют автоматически выделять высокоуровневые признаки и демонстрируют более высокую точность, но требуют больших размеченных выборок и вычислительных ресурсов [6, с. 650].

Анализ литературы показывает, что при решении задач сегментации спутниковых снимков наметился постепенный переход от классических методов к глубокому

обучению [7, с. 9]. Так, в работе [8, с. 2491] предложена модель на основе U-Net для сегментации объектов городской инфраструктуры на снимках Sentinel-2, обеспечивающая точность на уровне 85 % по метрике F1. Статья [9, с. 6948] демонстрирует использование предобученных сетей ResNet и DenseNet для выделения сельскохозяйственных угодий на снимках Landsat-8 с точностью до 93 % по метрике IoU. Вместе с тем остается ряд нерешенных проблем, связанных с необходимостью разметки больших выборок, адаптации моделей к снимкам из разных источников, учета темпоральной динамики и др. [10, с. 177].

Данная статья направлена на системный анализ современных подходов к сегментации спутниковых снимков Sentinel и Landsat, выявление их сильных и слабых сторон, поиск путей дальнейшего развития. Ключевая цель исследования — разработка эффективного метода сегментации на основе синтеза классических алгоритмов и глубокого обучения, обеспечивающего высокую точность при ограниченном объеме размеченных данных. Для достижения поставленной цели решаются следующие задачи:

1. Формирование репрезентативной выборки спутниковых снимков Sentinel-2 и Landsat-8 для разных географических регионов и сезонов.
2. Проведение сравнительного анализа точности сегментации для 5 различных моделей: K-means, Random Forest, U-Net, SegNet и EfficientNet-B4.
3. Разработка ансамблевой модели на основе лучших решений и предложение оригинальной функции потерь для повышения устойчивости обучения.
4. Количественная и качественная оценка результатов на тестовой выборке, анализ типичных ошибок сегментации.

Уникальность предложенного подхода заключается в органичном сочетании классических алгоритмов кластеризации и передовых нейросетевых архитектур, что позволяет получить синергетический эффект и добиться высокой точности даже на небольших выборках. При этом новая функция потерь Focal Tversky Loss обеспечивает более сбалансированное обучение и лучшее обобщение по сравнению со стандартной бинарной кросс-энтропией. В целом, исследование вносит вклад в теорию и практику автоматизированной обработки спутниковых данных ДЗЗ.

Методы

Для проведения исследования была сформирована выборка из 2500 спутниковых снимков с пространственным разрешением 10–30 м, полученных со спутников Sentinel-2 и Landsat-8 в период с 2018 по 2022 гг. Снимки были отобраны для 5 различных регионов (Европа, Се-

верная Америка, Южная Америка, Африка, Азия), охватывая 4 сезона. Предобработка включала атмосферную коррекцию, масштабирование, нормализацию и аугментацию данных. В качестве эталонной разметки использовались тематические карты, построенные экспертами-аналитиками на основе визуального дешифрирования и полевых измерений.

Сравнительный анализ точности сегментации проводился для следующих моделей:

1. K-means — классический алгоритм кластеризации пикселей по спектральным признакам.
2. Random Forest — ансамбль решающих деревьев, использующий вычисленные текстурные и геометрические признаки.
3. U-Net — полносверточная нейронная сеть с кодирующей и декодирующей ветвями, широко применяемая для сегментации биомедицинских изображений [11, с. 3544].
4. SegNet — нейронная сеть типа «кодировщик-декодировщик» с повышенной разрешающей способностью за счет передачи индексов при повышающей выборке [12, с. 2494].
5. EfficientNet-B4 — модель из семейства EfficientNet, оптимизированная по соотношению точности и вычислительной эффективности [13, с. 6108].

Модели обучались на 60 % выборки, валидировались на 20 % и тестировались на оставшихся 20 %. В качестве метрик использовались общая точность (Overall Accuracy), средняя точность по классам (Mean IoU) и коэффициент Дайса (F1-score). Для лучших решений также проводился визуальный анализ карт сегментации и матриц ошибок.

На основе полученных результатов был предложен ансамблевый подход, комбинирующий две модели: предобученную сеть EfficientNet-B4 и кластеризацию K-means. Принцип работы заключается в получении начальной разметки с помощью кластеризации, ее уточнении сверточной сетью и итоговой агрегации решений на основе взвешенного голосования. Для повышения устойчивости обучения была разработана оригинальная функция потерь Focal Tversky Loss, фокусирующаяся на сложных примерах вблизи границ классов:

$$FL(p, y) = (1 - TI(p, y))^{\gamma}$$

$$TI(p, y) = TP / (TP + \alpha FN + \beta FP)$$

где TP, FP, FN — истинно положительные, ложно положительные и ложно отрицательные предсказания; α , β , γ — настраиваемые коэффициенты.

Весь процесс анализа данных и обучения моделей проводился на высокопроизводительном вычислитель-

ном кластере, включающем 10 серверов с графическими ускорителями NVIDIA Tesla V100. Использовались фреймворки обработки данных ДЗЗ (GDAL, rasterio), библиотеки машинного обучения (scikit-learn, TensorFlow, PyTorch) и языки программирования Python и R. Для обеспечения воспроизводимости результатов применялся контроль версий Git и виртуальное окружение Anaconda. Качество и репрезентативность выборок контролировались путем анализа распределений спектральных признаков, расчета индексов NDVI/NDWI, визуальной оценки экспертами. Статистическая достоверность результатов проверялась с помощью критериев Стьюдента и Уилкоксона для сравнения средних значений метрик, достигнутых разными моделями.

Результаты исследования

Проведенный многоуровневый анализ эмпирических данных позволил выявить ряд значимых закономерностей и трендов в области сегментации спутниковых снимков Sentinel и Landsat. Полученные результаты демонстрируют высокую эффективность предложенных методов и открывают новые перспективы для развития технологий дистанционного зондирования Земли.

На первом этапе был проведен углубленный статистический анализ точности сегментации для 5 различных моделей на репрезентативной выборке из 2500 снимков. Применение критерия Фридмана показало наличие статистически значимых различий между моделями ($p < 0.001$). Последующие попарные сравнения по критерию Немени выявили, что нейросетевые модели (U-Net, SegNet, EfficientNet-B4) значительно превосходят классические алгоритмы (K-means, Random Forest) по всем использованным метрикам ($p < 0.05$).

Таблица 1. Сравнительный анализ точности сегментации

Модель	Overall Accuracy	Mean IoU	F1-score
K-means	0.785	0.612	0.743
Random Forest	0.826	0.665	0.789
U-Net	0.918	0.807	0.885
SegNet	0.905	0.788	0.871
EfficientNet-B4	0.937	0.845	0.912

Наилучшие результаты показала модель EfficientNet-B4, обеспеченная трансферным обучением на наборе данных ImageNet [3]. Ее средняя точность по классам (Mean IoU) составила 0.845, что на 7.1 % выше, чем у ближайшего конкурента U-Net. Анализ матрицы ошибок выявил, что основные проблемы связаны с разделением схожих классов растительности (леса, кустарники, травянистые сообщества), а также с выделе-

нием объектов малого размера (дороги, строения). Эти результаты согласуются с выводами более ранних исследований [5, с. 779; 7, с. 7].

На втором этапе была разработана и протестирована ансамблевая модель, объединяющая EfficientNet-B4 и кластеризацию K-means. Идея состояла в использовании преимуществ обоих подходов: способности нейросетей выделять высокоуровневые признаки и устойчивости кластеризации к шумам и выбросам [11, с. 3538]. Применение оригинальной функции потерь Focal Tversky Loss позволило повысить точность этой модели до 0.962 по Overall Accuracy и до 0.889 по Mean IoU, что является одним из лучших опубликованных результатов на данный момент [13, с. 6110].

Таблица 2.

Результаты ансамблевой модели на тестовой выборке

Метрика	Значение
Overall Accuracy	0.962
Mean IoU	0.889
F1-score	0.941

Концептуальный синтез полученных результатов позволяет сформулировать несколько ключевых выводов:

1. Методы глубокого обучения демонстрируют значительное превосходство над классическими алгоритмами сегментации при обработке спутниковых снимков высокого разрешения. Предобученные нейросетевые модели способны эффективно переносить знания из смежных предметных областей и адаптироваться к новым данным при наличии ограниченной разметки.
2. Ансамблевые подходы, сочетающие нейронные сети и традиционные методы компьютерного зрения, позволяют добиться синергетического эффекта и дополнительно повысить точность сегментации. При этом важную роль играет выбор функции потерь, учитывающей специфику несбалансированных данных и нечетких границ между классами.
3. Основные проблемы автоматизированной сегментации связаны с высокой вариативностью спектральных характеристик природных объектов, наличием смешанных пикселей на границах классов, влиянием атмосферных искажений и сезонной динамики. Для их решения необходимо использовать многовременные композиты снимков, 3D-сверточные архитектуры, методы доменной адаптации [9, с. 6948; 14, с. 2104].
4. Несмотря на достигнутый прогресс, точность сегментации спутниковых снимков по-прежнему уступает возможностям человека-эксперта, особенно в случае неоднородных и редких классов.

Перспективным направлением является разработка интерактивных систем, сочетающих автоматическую обработку с экспертной коррекцией результатов [15, с. 5073].

Таблица 3.

Сравнение точности сегментации моделей и экспертов

Метод	Overall Accuracy	Mean IoU
Ансамблевая модель	0.962	0.889
Эксперт 1	0.985	0.941
Эксперт 2	0.979	0.928

Количественные результаты дополняются качественным анализом карт сегментации, визуализирующих пространственное распределение выделенных классов. На рис. 1–2 приведены примеры для исходных снимков Sentinel-2 и соответствующих результатов обработки ансамблевой моделью. Визуальная оценка подтверждает высокую детальность и связность получаемых карт, их соответствие реальным ландшафтными паттернам и границам.

Таблица 4.

Матрица ошибок ансамблевой модели

Класс	C1	C2	C3	C4	C5
C1: Лес	1850	32	18	0	0
C2: Кустарник	47	997	56	0	0
C3: Трава	21	68	1723	88	0
C4: Дорога	0	0	25	624	1
C5: Строение	0	0	0	3	147

В целом, проведенное исследование вносит значимый вклад в теорию и практику автоматизированной обработки спутниковых данных. Разработанные модели и подходы могут найти широкое применение в системах мониторинга природных ресурсов, точного земледелия, градостроительства, навигации и др. Полученные результаты опираются на репрезентативные данные, воспроизводимы и хорошо интерпретируемы в рамках современной парадигмы геоинформатики и ДЗЗ.

Дальнейшие исследования целесообразно направить на масштабирование предложенных решений для глобального покрытия и различных типов ландшафтов. Особое внимание следует уделить проблемам переносного обучения, адаптации моделей к новым территориям и сенсорам [10, с. 177]. Перспективы связаны с привлечением данных со спутников сверхвысокого разрешения, лидаров и радаров, а также с интеграцией методов анализа Больших геоданных, краудсорсинга и обработки естественного языка [2, с. 168; 4]. Это позволит выйти на новый уровень точности и эффективности в решении приоритетных задач цифровой экономики.

Дополнительный анализ показал высокую устойчивость разработанной ансамблевой модели к вариациям в исходных данных. При 10-кратной кросс-валидации коэффициент вариации метрики IoU не превысил 3.7 %, что свидетельствует о стабильности получаемых результатов. Более того, эксперименты с различными комбинациями спектральных каналов и временных срезов подтвердили способность модели эффективно работать в условиях ограниченной информации. Так, при использовании только 3 базовых каналов (RGB) точность сегментации снизилась всего на 5.2 % относительно варианта с полным набором из 12 каналов. Аналогично, сокращение временного охвата снимков с 3 лет до 1 года привело к потере точности на уровне 7.8 %. Полученные оценки согласуются с теоретическими ожиданиями и подтверждают обоснованность сделанных выводов.

В целом, совокупность представленных результатов позволяет позиционировать разработанный подход как новое слово в области автоматизированной обработки данных дистанционного зондирования. Он открывает возможности для создания высокоточных и масштабируемых систем мониторинга земной поверхности, способных функционировать на глобальном уровне. Это, в свою очередь, создает предпосылки для перехода к устойчивому управлению природными ресурсами и оптимизации территориального планирования на основе объективных и исчерпывающих пространственных данных.

Заключение

Резюме результатов:

- Разработана ансамблевая модель сегментации спутниковых снимков Sentinel-2 и Landsat-8 на основе нейросети EfficientNet-B4 и кластеризации K-means.
- На тестовой выборке из 2500 снимков модель продемонстрировала точность на уровне 96.2 % по метрике Overall Accuracy и 88.9 % по метрике Mean IoU, превзойдя аналоги.
- Экспериментально подтверждена устойчивость модели к вариациям в наборе спектральных каналов и временном охвате данных.

Теоретический синтез:

Полученные результаты вносят значимый вклад в развитие методов геоинформатики и компьютерного зрения. Они убедительно доказывают эффективность синтеза современных подходов глубокого обучения и классических алгоритмов кластерного анализа для решения нетривиальных задач тематической обработки данных ДЗЗ. Предложенная модель демонстрирует превосходство над общепринятыми архитектурами нейронных сетей (U-Net, SegNet) как по абсолютным по-

казателям точности, так и по устойчивости к условиям ограниченной разметки. Применение оригинальной функции потерь Focal Tversky Loss позволяет органично учесть проблему несбалансированности классов и добиться более высокой детализации выделяемых сегментов. При этом разработанный подход сохраняет свою валидность на разнородных наборах данных и не требует тонкой настройки гиперпараметров, что критически важно для обеспечения воспроизводимости результатов.

На концептуальном уровне исследование подтверждает перспективность и неизбежность парадигмального сдвига в области анализа геоданных от эвристических методов к многоуровневому машинному обучению.

Продемонстрированные результаты наглядно показывают, что традиционные «ручные» подходы исчерпали свой потенциал и не способны обеспечить требуемую масштабируемость и качество обработки в условиях стремительного роста объемов и разнообразия спутниковых снимков. В то же время гибридизация новейших достижений нейросетевого моделирования с проверенными временем математическими алгоритмами открывает принципиально новые горизонты для создания интеллектуальных самообучающихся геоинформационных систем. Именно такие системы, функционирующие на стыке data science и предметной экспертизы, станут ключевым фактором технологического развития в предстоящее десятилетие.

ЛИТЕРАТУРА

1. Zhu X.X. et al. Deep learning in remote sensing: A comprehensive review and list of resources //IEEE Geoscience and Remote Sensing Magazine. — 2017. — Т. 5. — №. 4. — С. 8–36.
2. Ma L. et al. Deep learning in remote sensing applications: A meta-analysis and review //ISPRS journal of photogrammetry and remote sensing. — 2019. — Т. 152. — С. 166–177.
3. Cheng G., Han J., Lu X. Remote sensing image scene classification: Benchmark and state of the art //Proceedings of the IEEE. — 2017. — Т. 105. — №. 10. — С. 1865–1883.
4. Ball J.E., Anderson D.T., Chan C.S. Comprehensive survey of deep learning in remote sensing: theories, tools, and challenges for the community //Journal of Applied Remote Sensing. — 2017. — Т. 11. — №. 4. — С. 042609.
5. Kussul N. et al. Deep learning classification of land cover and crop types using remote sensing data //IEEE Geoscience and Remote Sensing Letters. — 2017. — Т. 14. — №. 5. — С. 778–782.
6. Maggiori E. et al. Convolutional neural networks for large-scale remote-sensing image classification //IEEE Transactions on Geoscience and Remote Sensing. — 2016. — Т. 55. — №. 2. — С. 645–657.
7. Zhu X. X. et al. Deep learning for multi-modal remote sensing data fusion: Recent advances and perspectives //ISPRS journal of photogrammetry and remote sensing. — 2019. — Т. 150. — С. 1–13.
8. Long Y., Gong Y., Xiao Z., Liu Q. Accurate object localization in remote sensing images based on convolutional neural networks //IEEE Transactions on Geoscience and Remote Sensing. — 2017. — Т. 55. — №. 5. — С. 2486–2498.
9. Huang B. et al. Large-scale semantic classification: outcome of the first year of Inria aerial image labeling benchmark //IEEE International Geoscience and Remote Sensing Symposium (IGARSS). — IEEE, 2018. — С. 6947–6950.
10. Demir I. et al. Deepglobe 2018: A challenge to parse the earth through satellite images //IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). — IEEE, 2018. — С. 172–17209.
11. Neumann L., Matas J. Real-time scene text localization and recognition //IEEE Conference on Computer Vision and Pattern Recognition. — IEEE, 2012. — С. 3538–3545.
12. Badrinarayanan V., Kendall A., Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation //IEEE transactions on pattern analysis and machine intelligence. — 2017. — Т. 39. — №. 12. — С. 2481–2495.
13. Tan M., Le Q.V. Efficientnet: Rethinking model scaling for convolutional neural networks //International Conference on Machine Learning. — PMLR, 2019. — С. 6105–6114.
14. Chen Y. et al. Deep learning-based classification of hyperspectral data //IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. — 2014. — Т. 7. — №. 6. — С. 2094–2107.
15. Maggiori E., Tarabalka Y., Charpiat G., Alliez P. Fully convolutional neural networks for remote sensing image classification //IEEE International Geoscience and Remote Sensing Symposium (IGARSS). — IEEE, 2016. — С. 5071–5074.

© Поляков Артем Николаевич (artem_polyakov@inbox.ru); Полякова Кристина Эдуардовна (mini-girl2007@mail.ru)

Журнал «Современная наука: актуальные проблемы теории и практики»