

Применение методов машинного обучения для обработки и анализа данных дистанционного зондирования Земли на платформе ИДСТУ СО РАН

И. В. Бычков, Ю. В. Авраменко, А. К. Попова*, Р. К. Фёдоров

Институт динамики систем и теории управления им. В.М. Матросова СО РАН, 664033, Иркутск, Россия

*Контактный автор: Попова Анастасия Константиновна, e-mail: chudnenko@icc.ru

Поступила 06 июня 2025 г., доработана 19 июня 2025 г., принята в печать 24 июня 2025 г.

Статья посвящена обзору платформы ИДСТУ СО РАН, где представлен комплексный набор инструментов от получения спутниковых снимков до распространения результатов в виде карт. Описаны хранилище мультиспектральных спутниковых снимков и полуавтоматический инструмент создания обучающей выборки для методов машинного обучения. Рассмотрена возможность классификации данных ДЗЗ с помощью моделей на базе JupiterHub. Представлена реализованная технология публикации карт на геопортале платформы, которая ускоряет получение наглядных результатов. Такая платформа не только улучшает рабочий процесс анализа данных ДЗЗ, но и устраняет разрыв между передовыми вычислительными методами и практическими приложениями в области мониторинга окружающей среды.

Ключевые слова: геопортал, данные ДЗЗ, классификация спутниковых снимков, нейронная сеть, платформа обработки спутниковых снимков.

Цитирование: Бычков И.В., Авраменко Ю.В., Попова А.К., Фёдоров Р.К. Применение методов машинного обучения для обработки и анализа данных дистанционного зондирования Земли на платформе ИДСТУ СО РАН. Вычислительные технологии. 2025; 30(4):77–89. DOI:10.25743/ICT.2025.30.4.008.

Вместо предисловия

Спутниковые снимки успешно используют для исследования территорий городов и сельскохозяйственных земель, прогнозирования наводнений, обнаружения пожаров и других задач. Для обработки данных дистанционного зондирования Земли (ДЗЗ) разрабатывают платформенные решения, предоставляющие комплексный набор инструментов от получения изображений до распространения результатов в единой среде. В Сибирском отделении РАН под руководством акад. Ю.И. Шокина инициирована работа по созданию информационно-телекоммуникационной и вычислительной инфраструктуры центра мониторинга состояния окружающей среды на основе анализа данных ДЗЗ. Фактически созданная инфраструктура стала одним из центральных узлов ядра академической специализированной сети данных ДЗЗ, охватывающей всю территорию страны и предоставляющей организациям РАН доступ к необходимым для исследовательской работы оперативным и архивным данным спутниковых наблюдений. На базе этой инфраструктуры акад. Ю.И. Шокиным был создан Центр коллективного пользова-

ния, который взаимодействует с крупнейшими поставщиками спутниковых, авиационных и наземных данных по всей стране. В продолжении этих работ на платформе ИДСТУ СО РАН создано и реализовано хранилище мультиспектральных спутниковых снимков, инструмент создания обучающей выборки для методов машинного обучения. Внедрение JupiterHub дает пользователям возможность создавать свои модели классификации или использовать предварительно обученные модели. Технология публикации карт на геопортале облегчает представление результатов. Такой комплексный подход не только улучшает рабочий процесс анализа данных ДЗЗ, но и устраняет разрыв между передовыми вычислительными методами и практическими приложениями в области мониторинга окружающей среды.

Введение

На протяжении десятилетий с помощью дистанционного зондирования Земли (ДЗЗ) собирались большие объемы данных, которые сложно обрабатывать и анализировать с помощью стандартных программных средств и настольных вычислительных ресурсов. С ростом объема и сложности доступной спутниковой информации становятся более важными платформы для обработки и анализа данных ДЗЗ. В последние годы наблюдается интенсивное развитие технологий, позволяющих интегрировать данные из различных источников для анализа экологических, сельскохозяйственных и управленческих вопросов. Платформы облегчают обработку крупномасштабных данных дистанционного зондирования, расширяя возможности прогнозирования с помощью облачных вычислений, методов искусственного интеллекта и интерактивных инструментов визуализации. Такой подход позволяет повысить качество обработки данных и доступность методов, необходимых для анализа геопространственной информации [1].

Google Earth Engine (GEE) предоставляет доступ к обширным наборам данных, включая снимки Landsat, Sentinel и MODIS, и алгоритмам машинного обучения для решения задач классификации почвенного покрова, гидрологии и анализа климата. Способность обрабатывать данные на больших площадях и за длительный период сделала его популярным инструментом для мониторинга окружающей среды и городского планирования [2, 3]. Платформа GEE поддерживает интерактивные веб-приложения для анализа пространственных данных, таких как NDVI, благодаря интегрированной среде разработки. Эта платформа сокращает время и стоимость разработки, позволяя эффективно объединять для анализа спутниковые снимки с климатическими, топографическими, геофизическими данными [4].

AI Earth, работающая на базе Alibaba Cloud, сочетает в себе высокопроизводительные вычисления и модели глубокого обучения, адаптированные для задач дистанционного зондирования. Она позволяет пользователям обучать собственные модели и внедрять сторонние алгоритмы [5].

Тематические платформы используют методы автоматизации и интеграции геоинформационных систем (ГИС) с данными ДЗЗ. Например, данные, полученные со спутников SPOT-5, успешно применяются для картографирования водорослей в Средиземном море [6]. Система Vega-Science создавалась как средство мониторинга растительного покрова, но расширение состава доступных спутниковых данных (мультиспектральных, радарных, лидарных, гиперспектральных и т. д.) и внедрение инструментов для анализа снимков позволяют проводить с ее помощью мониторинг других природных объектов, включая состояние морей и вулканическую активность [7].

Внедрение в платформы методов машинного обучения играет ключевую роль в анализе данных ДЗЗ, обеспечивая эффективную классификацию сложных наборов данных. Методы искусственного интеллекта трансформировали анализ спутниковых снимков. С ростом доступности мульти-, гиперспектральных и радарных наборов данных на основе алгоритмов машинного обучения создаются мощные инструменты для извлечения значимых закономерностей, автоматизации классификации и улучшения процессов принятия решений. Такой подход позволяет решать проблемы сложности данных и ограниченности размеченных выборок, делая более доступным использование спутниковых снимков в мониторинге окружающей среды, прогнозировании стихийных бедствий и городском планировании.

Для задач классификации данных ДЗЗ широко используются контролируемые методы обучения — деревья решений, ансамблевые методы и подходы на основе ядер. Деревья принятия решений просты, но эффективны для картографирования почвенно-растительного покрова и типов растительности [8]. Ансамблевые методы, такие как Random Forest и Gradient boosted decision trees, повышают точность за счет объединения нескольких моделей [9]. Например, в исследовании, проведенном в зоне разлома Лембанг (Индонезия), для оценки изменений в землепользовании использован метод Random Forest, который позволил обнаружить увеличение городских территорий на 24.96 % за шесть лет [10]. При выделении спектрально близких классов (например, пород деревьев) по мультиспектральным снимкам Random Forest не всегда дает возможность получить результат нужной точности, однако добавление вспомогательных данных (климатических, почвенных, топографических и т. п.) позволяет преодолеть это ограничение [11]. Для оценки температуры поверхности земли в Йезде (Иран) были применены модели градиентного бустинга, что позволило выявить эффект городского теплового острова и поддержать стратегии устойчивости к изменению климата [12].

Алгоритмы на основе ядра (например, метод опорных векторов SVM) лучше справляются с обработкой многомерных данных и сложных границ классов [13], что было продемонстрировано при усовершенствовании картографирования городского почвенно-растительного покрова [14, 15]. Гибридные подходы, например RoF-KOPLS, сочетают методы на основе ядер и ансамблевое обучение для повышения точности классификации гиперспектральных данных [16].

Глубокое и трансферное обучение значительно расширили возможности анализа данных ДЗЗ, обеспечив более точную и эффективную обработку сложных наборов данных. Сверточные нейронные сети (CNN) показывают высокую точность в задачах картографирования почвенно-растительного покрова, обнаружения объектов и выявления изменений благодаря своей способности извлекать иерархические пространственные признаки [17, 18]. Например, модели на основе CNN продемонстрировали значительное повышение производительности при классификации спутниковых изображений в режиме реального времени и атмосферной коррекции с обнаружением облаков [19, 20]. Классификация снимков спутника Sentinel-2 с помощью нейронной сети ResNet-50 позволила выделить 11 классов земной поверхности на территории Иркутской области [21].

Трансферное обучение решает проблему ограниченности размеченных данных, используя предварительно обученные модели из других областей. Такой подход сокращает время обучения и повышает производительность, что видно на примере классификации сельскохозяйственных культур, картирования бедности [22, 23] и др. Тонкая настройка и фиксация слоев оптимизируют адаптацию предварительно обученных моделей к тематическим наборам данных [24]. Однако эти методы сталкиваются с вычис-



Рис. 1. Инфраструктура платформы

Fig. 1. Platform infrastructure

лительными проблемами, включая высокие требования к ресурсам и риск негативного переноса, когда исходная и целевая области исследования значительно различаются [25].

В рамках существующих открытых платформ не всегда можно реализовать полный цикл классификации данных ДЗЗ, от подготовки обучающей выборки до публикации карты результата с подсчетом площадей по классам, которая обеспечит получение результата нужной для экологического мониторинга степени точности. Платформа ИДСТУ СО РАН поддерживает полный цикл обработки спутниковых снимков, результатом которого являются карты классификации земной поверхности по заданному набору классов.

Полный цикл обработки спутниковых снимков для классификации на платформе ИДСТУ СО РАН включает скачивание снимков, размещение их в облачном хранилище (<https://cloud.icc.ru>), атмосферную коррекцию, создание обучающей выборки, классификацию методами машинного обучения с помощью инструментов в среде Jupyter Notebook (<https://lab.icc.ru>) и публикацию на геопортале (<https://geos.icc.ru>) полученной карты (рис. 1).

1. Загрузка и просмотр снимков

На геопортал ИДСТУ СО РАН в автоматизированном режиме регулярно загружаются спутниковые снимки. Для задач природного мониторинга чаще используются мультиспектральные изображения с аппаратов Sentinel-2 с разрешением 10 м. Снимки уровня L1C обрабатываются программой Sen2Cor до уровня L2A, чтобы обеспечить их атмосферную коррекцию. Атмосферная коррекция устраняет искажения, вызванные атмосферой Земли, и обеспечивает представление данных с поправкой на рассеяние и поглощение света атмосферными частицами. Этот шаг повышает точность и надежность данных, что имеет решающее значение для мониторинга окружающей среды.

Каталог снимков можно просматривать, используя доступные фильтры — по наименованию спутника, местоположению снимка, диапазону дат, проценту облачности. Для исследования временной последовательности снимков на определенную область доступен инструмент таймлайн или хронологическая шкала. Пользователь выбирает на карте интересующий объект, в ленте таймлайна — нужный год и месяц (рис. 2). На карте появляется первый снимок, доступный на заданный месяц. Далее пользователь может переходить к более поздним или ранним снимкам на выбранную территорию с помощью кнопок пролистывания изображений в интерфейсе геопортала. Этот инструмент позволяет более подробно изучить область исследования, отследив происходящие на ней изменения во времени. Например, рост сельскохозяйственных насаждений или вос-

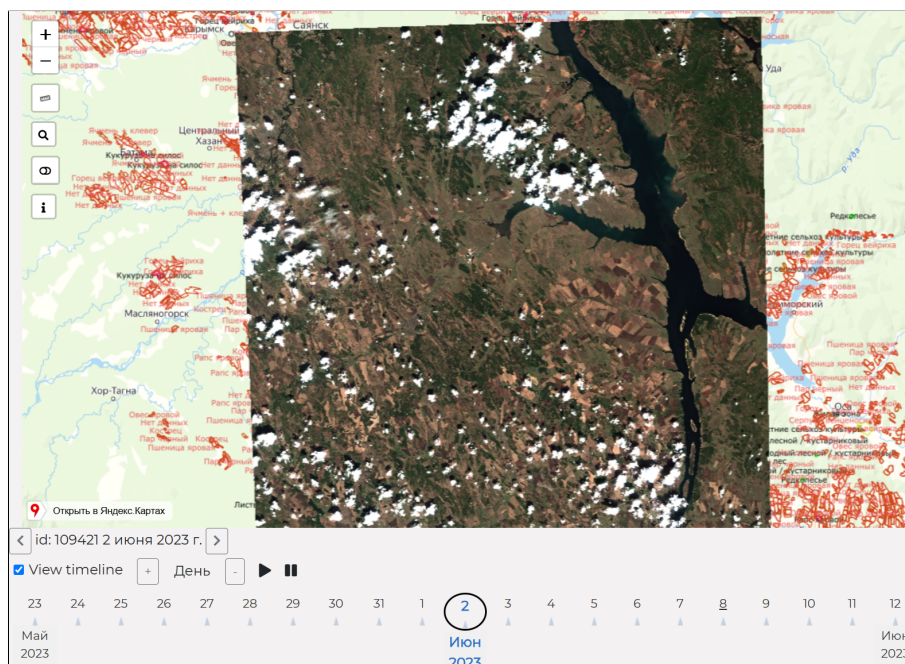


Рис. 2. Таймлайн для просмотра снимков на выбранный участок

Fig. 2. Timeline for viewing images for the selected area

становление лесов после пожара. Полученная информация позволит уточнить будущую разметку снимков для выбора состава классов и расположения полигонов разметки.

2. Создание обучающей выборки и классификация

Для классификации спутниковых снимков методами машинного обучения необходимо создать обучающую выборку, которая содержит достаточное количество образцов каждого класса. Объем и качество такой выборки напрямую влияют на производительность модели классификации. Существуют различные готовые наборы данных для обучения, но при классификации природных территорий пользователи иногда сталкиваются с низкими результатами при использовании таких готовых выборок [21]. Это происходит из-за различия в климатических, почвенных, топографических характеристиках местности, которые влияют на значения спектрального отражения визуально схожих объектов. Для решения данной проблемы рекомендуется создать собственный набор данных, учитывающий особенности исследуемой территории. Различают ручной, автоматический или полуавтоматический методы для создания обучающей выборки.

Для разметки ручным способом используют ГИС, в которой на основе визуального осмотра спутниковых снимков создают векторный слой, на который наносят полигональные объекты для каждого класса на основе экспертных знаний. Этот метод довольно трудоемок, особенно если необходимо разметить снимки на большую площадь уровня региона и выше.

Автоматические методы аннотирования образцов основываются на отборе объектов с использованием пороговых значений [26, 27]. В этом случае применяют глубокое обучение, которое выделяет и анализирует участки снимка для каждого класса, формируя для них набор признаков. Полученные признаки фильтруются на основе пороговых значений, формируя обучающую выборку.

Полуавтоматические методы предлагают предварительное выделение и группировку фрагментов изображения (например, с помощью сегментации) и последующее присвоение им классов пользователем вручную. На геопортале ИДСТУ СО РАН создан Web-интерфейс для полуавтоматической разметки данных ДЗЗ, основанный на выделении объектов с помощью Segment Anything Model (SAM) [28]. SAM — это универсальная модель сегментации, предварительно обученная на 11 млн изображений и 1 млрд масок. Она выделяет объекты на растровых изображениях в соответствии с запросом, в качестве которого может выступать подсказка в виде точки объекта или ограничивающей прямоугольной области. SAM работает в браузере в режиме реального времени и выделяет контур объекта, который средствами геопортала сохраняется в формат WKT и добавляется в базу данных с присвоением метки класса.

С помощью инструмента разметки (рис. 3) сформирована обучающая выборка, содержащая характерные для Иркутской области 11 классов: голая скала, пастбище, вырубки, кустарники, хвойный лес, редколесье, лиственный лес, смешанный лес, пашня, жилая зона, вода. Всего размечены 134 космоснимка Иркутской области и Республики Бурятия, количество полигональных объектов 2246 на площадь более 5900 км². Все снимки летнего периода за 2018–2020 гг.

Для классификации использовалась сверточная нейронная сеть архитектуры ResNet-50. Обучение ResNet-50 проводилось на 120 эпохах по 13 000 образцов в каждом классе без предобученных весов. На вход нейронной сети подается образец, представленный тензором 64×64×12. Размерность тензора определяет окрестность, по которой нейронная сеть принимает решение. В качестве окрестности была выбрана ячейка снимка 64×64 пикселя, размер которой обусловлен архитектурой нейронной сети. Каждый пиксель представлен 12 исходными каналами Sentinel-2 и имеет пространственное разрешение 10 м. На выходе сеть выдает в качестве результата класс центрального пикселя.

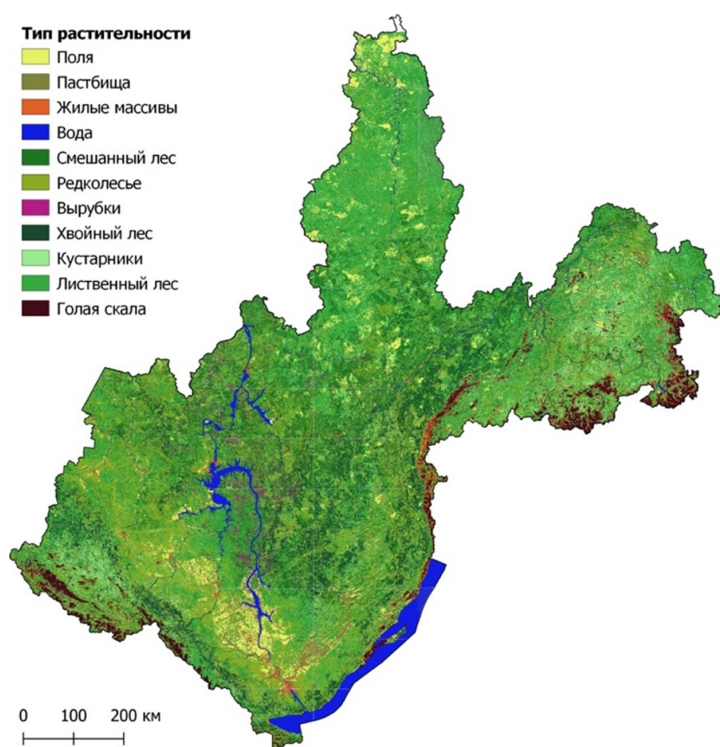


Рис. 3. Результат классификации снимков Иркутской области
Fig. 3. The result of imagery classification for the Irkutsk region

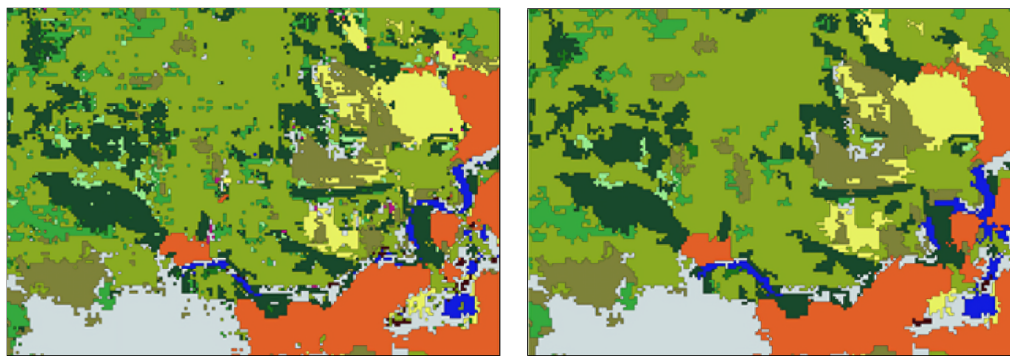


Рис. 4. Карта до и после генерализации

Fig. 4. Map before and after generalization

3. Публикация полученной карты

Для публикации готовой карты классификации на геопортале ИДСТУ СО РАН (рис. 4) используется WMS-сервис, на вход которого подаются растровый файл в формате GeoTIFF или таблица, созданная с помощью сервисов ввода данных, и стили отображения карты. Выделение стиля в отдельный параметр позволяет повторно использовать одинаковые параметры для отрисовки разных карт. Приложение Mapserver осуществляет на стороне сервера генерацию изображений слоев карт в соответствии со стандартом WMS. Модуль MapCache производит кэширование и обеспечивает скоростной доступ к картам. Формирование легенды производится в браузере с помощью специального редактора стилей, в котором легенда формируется для каждого слоя отдельно и хранится в формате JSON. Для отображения слоев с помощью Mapserver стили в формате JSON конвертируются в файл настроек карты (map-файл).

4. Формирование результатов с разбивкой по административному делению

Изначально результат классификации представляет собой растровое изображение в градациях серого, где каждый пиксель означает порядковый номер одного из классов. При мониторинге природных объектов одной из задач выступает подсчет площадей, занятых разными классами на территории определенного района. Важно учитывать топологические свойства и характеристики представленных на карте объектов, что не всегда возможно при ее хранении в растровом формате. Поэтому для эффективного анализа результата решено конвертировать карту в векторный формат. Для получения итоговой карты результат классификации объединяется с векторным слоем административного деления. Здесь каждому пикселю исходного изображения сопоставляется код субъекта административного деления.

Также результат классификации обычно содержит объекты малой площади, например парки и газоны в населенных пунктах, частные дома, трансформаторные будки и т. д. Для решения многих практических задач такая детальность не требуется, кроме того, большое количество мелких объектов увеличивает сложность обработки данных. Сократить их число можно при помощи алгоритма выделения связных компонент.

Применяется алгоритм выделения связных компонент, который преобразует компоненты в соответствии с заданным порядком. Векторизация результата классификации

изображения (размером $10\,980 \times 10\,980$ пикселей) выполняется с помощью библиотеки GDAL. Применяемый метод сокращает количество векторных объектов на 60 % — с исходных 80 тыс. до 26 тыс. объектов. Пример исходного и полученного изображений показан на рис. 4.

Сокращение числа объектов позволяет уменьшить время выполнения пространственных запросов, связанных с отображением карты на геопортале.

Заключение

Классификация данных ДЗЗ является важной задачей в области геопространственного анализа, когда космические снимки классифицируются по заранее определенным типам почвенно-растительного покрова. Она играет важную роль в различных областях, таких как сельское хозяйство, борьба со стихийными бедствиями, экологический надзор, городское планирование и управление природными ресурсами. Методы машинного обучения значительно расширили возможности применения дистанционного зондирования для решения задач мониторинга. Эти методы позволяют эффективно обрабатывать данные и принимать решения в различных экологических и социальных контекстах.

Платформенный подход упрощает и ускоряет весь цикл обработки спутниковых снимков. Предложенная в данном исследовании технология в рамках системы ИДСТУ СО РАН позволяет получать доступ к коллекции снимков, просматривать их для анализа, формировать обучающую выборку, классифицировать методами машинного обучения и публиковать карты на геопортале. Такое комплексное решение помогает расширять мониторинг природных объектов, давая доступ широкому кругу пользователей к результатам в виде наглядных карт, графиков и статистических подсчетов. Созданная платформа позволяет экспертам наблюдать за изменениями земельных участков, выявлять тенденции землепользования и оценивать воздействие на окружающую среду, что делает ее важным инструментом в решении сложных проблем, связанных с изменением климата, городским развитием и защитой природных ресурсов.

Благодарности. Работа поддержана в рамках госзадания Минобрнауки России по проекту “Методы и технологии облачной сервис-ориентированной цифровой платформы сбора, хранения и обработки больших объемов разноформатных междисциплинарных данных и знаний, основанные на применении искусственного интеллекта, модельно-управляемого подхода и машинного обучения” (номер гос. регистрации 121030500071-2). Результаты получены с использованием Центра коллективного пользования “Интегрированная информационная сеть Иркутского научно-образовательного комплекса”.

Список литературы

- [1] Якубайлик О.Э., Кадочников А.А., Токарев А.В. Программные средства для визуализации геопространственных данных в системе оперативного спутникового мониторинга. Материалы конференции “Информационные технологии в дистанционном зондировании Земли — RORSE 2018”. М.; 2019: 264–270. DOI:10.21046/rorse2018.264.
- [2] Amani M., Ghorbanian A., Ahmadi S. Google Earth Engine cloud computing platform for remote sensing big data applications: a comprehensive review. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2020; (13):5326–5350. DOI:10.1109/JSTARS.2020.3021052.

- [3] **Sangeetham R., Reddy N.** A review on Google Earth Engine: an open access cloud analysis platform for planetary scale satellite data. *E3S Web of Conferences*. 2024; (591):09011. DOI:10.1051/e3sconf/202459109011.
- [4] **Habibie I.** The application of machine learning using Google Earth Engine for remote sensing analysis. *Jurnal Teknoinfo*. 2022; 16(2):233. DOI:10.33365/jti.v16i2.1872. Available at: https://www.researchgate.net/publication/361870523_the_application_of_machine_learning_using_google_earth_engine_for_remote_sensing_analysis.
- [5] **Xu H., Man Y., Yang M., Wu J., Zhang Q., Wang J.** Analytical insight of Earth: a cloud-platform of intelligent computing for geospatial big data. *ArXiv*. 2023; abs/2312.16385. DOI:10.48550/arXiv.2312.16385.
- [6] **Новикова А.М., Каширина Е.С., Новиков А.А., Полонский А.Б., Панкеева Т.В.** ГИС в морских исследованиях: мировой опыт и возможности его применения на примере черноморского региона. *Труды Карадагской Научной Станции им. Т.И. Вяземского Природного Заповедника РАН*. 2021; 1(3):54–66. DOI:10.21072/eco.2021.03.05.
- [7] **Лупян Е.А., Прошин А.А., Бурцев М.А., Кашницкий А.В., Балашов И.В.** Система “Вега-Science”: особенности построения, основные возможности и опыт использования. *Современные проблемы дистанционного зондирования Земли из космоса*. 2021; 18(6):9–31. DOI:10.21046/2070-7401-2021-18-6-9-31.
- [8] **Kurdi F.T., Amakhchan W., Gharineiat Z.** Random forest machine learning technique for automatic vegetation detection and modelling in LiDAR data. *International Journal of Environmental Sciences & Natural Resources*. 2021; 28(2):556234. DOI:10.19080/IJESNR.2021.28.556234.
- [9] **Lee S., Hyun Y., Lee S., Lee M.** Groundwater potential mapping using remote sensing and GIS-based machine learning techniques. *Remote Sensing*. 2020; 12(7):1200. DOI:10.3390/rs12071200.
- [10] **Prasetyadi A., Sutyanawan A.G., Nur W.H., Bisri A., Hanifa N.R., Shomim A.F.** Integrating remote sensing and machine learning for monitoring urban growth in seismically active regions: a case study of the Lembang fault zone, West Java, Indonesia. *International Conference on Computer, Control, Informatics and its Applications (IC3INA)*. Bandung, Indonesia; 2024: 208–212. DOI:10.1109/IC3INA64086.2024.10732244.
- [11] **Popova A.** Improving the accuracy of tree species mapping by Sentinel-2 images using auxiliary data — a case study of Slyudyanskoye forestry area near Lake Baikal. *Forests*. 2025; (16):487. DOI:10.3390/f16030487.
- [12] **Mansourmoghaddam M., Roustaei I., Malamiri H.G., Sadeghnejad M., Krzyszczyk J., Ferreira C.S.S.** Modeling and estimating the land surface temperature (LST) using remote sensing and machine learning (case study: Yazd, Iran). *Remote Sensing*. 2024; 16(3):454. DOI:10.3390/rs16030454.
- [13] **Waske B., Fauvel M., Benediktsson J.A., Chanussot J.** Machine learning techniques in remote sensing data analysis. *Kernel Methods for Remote Sensing Data Analysis*. N.Y.: Wiley; 2009: 434. DOI:10.1002/9780470748992.ch1.
- [14] **Rodriguez-Galiano V., Ghimire B., Rogan J., Chica-Olmo M.** An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2012; (67):93–104. DOI:10.1016/J.ISPRSJPRS.2011.11.002.
- [15] **Sheykhmousa M., Mahdianpari M., Ghanbari H.** Support vector machine versus random forest for remote sensing image classification: a meta-analysis and systematic review.

- IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2020; (13):6308–6325. DOI:10.1109/JSTARS.2020.3026724.
- [16] **Chen J., Xia J., Du P., Chanussot J., Xue Z., Xie X.** Kernel supervised ensemble classifier for the classification of hyperspectral data using few labeled samples. *Remote Sensing*. 2016; 8(7):601. DOI:10.3390/rs8070601.
- [17] **Hoeser C., Zhang G.** Trends in image object detection and segmentation with deep neural network on natural and Earth observations. *International Journal of Computer Science and Information Technology*. 2024; 4(3):36–46. DOI:10.62051/ijcsit.v4n3.05.
- [18] **Pilato A., Taggio N., Pompili A.** Deep learning approaches to Earth Observation change detection. *Remote Sensing*. 2021; 13(20):4083. DOI:10.3390/rs13204083.
- [19] **Irtaza G., Latif S., Nadeem R.M., Hussain N., Ijaz H.M.** Real-time satellite image classification using convolutional neural networks for Earth observation and video feed analysis. *Kashf Journal of Multidisciplinary Research*. 2025; 2(03):204–216. DOI:10.71146/kjmr365.
- [20] **Mateo-Garcia G., Aybar C., Acciarini G., Ruzicka V., Meoni G., Longepe N.** Onboard cloud detection and atmospheric correction with deep learning emulators. 2023 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2023). 2023: 1875–1878. DOI:10.1109/IGARSS52108.2023.10282605.
- [21] **Bychkov I.V., Ruzhnikov G.M., Fedorov R.K., Popova A.K., Avramenko Y.V.** On classification of Sentinel-2 satellite images by a neural network ResNet-50. *Computer Optics*. 2023; 47(3):474–481. DOI:10.18287/2412-6179-CO-1216.
- [22] **Gadiraju K.K., Vatsavai R.R.** Remote sensing based crop type classification via deep transfer learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2023; (16):4699–4712. DOI:10.1109/JSTARS.2023.3270141.
- [23] **Xie M., Jean N., Burke M., Lobell D., Ermon S.** Transfer learning from deep features for remote sensing and poverty mapping. *Proceedings of the AAAI Conference on Artificial Intelligence*. Phoenix; 2015; 30(1). DOI:10.1609/aaai.v30i1.9906. Available at: https://www.researchgate.net/publication/282403466_Transfer_Learning_from_Deep_Features_for_Remote_Sensing_and_Poverty_Mapping.
- [24] **Pires de Lima R., Marfurt K.** Convolutional neural network for remote-sensing scene classification: transfer learning analysis. *Remote Sensing*. 2019; 12(1):86. DOI:10.3390/rs12010086.
- [25] **Huber F., Inderka A., Steinhage V.** Leveraging remote sensing data for yield prediction with deep transfer learning. *Sensors*. 2024; 24(3):770. DOI:10.3390/s24030770.
- [26] **Fawzy M., Khodary F., Mostafa Y.** Selecting training samples automatically from VHR satellite images for image classification. *Sohag Engineering Journal*. 2021; 1(1):71–84. DOI:10.21608/SEJ.2021.155929.
- [27] **Zhan Y.** Comparing three methods of selecting training samples in supervised classification of multispectral remote sensing images. *Sensors*. 2023; 23(20):8530. DOI:10.3390/s23208530.
- [28] **Kirillov A., Mintun E., Ravi N., Mao H., Rolland C., Gustafson L.** Segment anything. 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Paris; 2023: 3992–4003. DOI:10.1109/ICCV51070.2023.00371.
-

COMPUTATIONAL TECHNOLOGIES

DOI:10.25743/ICT.2025.30.4.008

Application of machine learning methods for processing and analysis of remote sensing data on ISDCT SB RAS platform

I. V. BYCHKOV, YU. V. AVRAMENKO, A. K. POPOVA*, R. K. FEDOROV

Matrosov Institute for System Dynamics and Control Theory SB RAS, 664033, Irkutsk, Russia

*Corresponding author: Anastasia K. Popova, e-mail: chudnenko@icc.ru*Received June 06, 2025, revised June 19, 2025, accepted June 24, 2025.***Abstract**

Satellite imagery is successfully employed for urban studies, flood forecasting, fire detection, and vegetation classification. Machine learning models analyze and classify these images based on their spatial, spectral, and temporal characteristics. To streamline Earth observation data processing, platform-based solutions have been developed, which integrate a comprehensive set of tools from image acquisition to result dissemination within a unified environment. The platform of the ISDCT SB RAS has introduced a repository of multispectral satellite imagery with an intuitive timeline-based navigation system. This functionality allows researchers to efficiently track environmental and anthropogenic changes in a given study area over time. Additionally, the platform includes a semi-automated training sample generation tool, significantly accelerating the preparation of datasets for machine learning model training. The integration of JupyterHub enables users to develop custom classification models or apply pre-trained models for image analysis. Classification of satellite images allows to identify different types of the land cover on the imagery and track their changes over time. Furthermore, the platform supports the seamless publication of derived geospatial maps on an interactive geoportal, ensuring that results are presented in a decision-ready format. This holistic approach not only optimizes remote sensing data analysis workflows but also bridges the gap between advanced computational methods and practical applications in environmental monitoring.

Keywords: geoportal, remote sensing data, satellite image classification, neural network, platform for satellite image processing.

Citation: Bychkov I.V., Avramenko Yu.V., Popova A.K., Fedorov R.K. Application of machine learning methods for processing and analysis of remote sensing data on ISDCT SB RAS platform. Computational Technologies. 2025; 30(4):77–89. DOI:10.25743/ICT.2025.30.4.008. (In Russ.)

Acknowledgements. The research was supported within the State Assignment of the Ministry of Education and Science of the Russian Federation for the project “Methods and technologies of cloud-based service-oriented platform for collecting, storing and processing large volumes of multi-format interdisciplinary data and knowledge based upon the use of artificial intelligence, model-guided approach and machine learning” (No. 121030500071-2). Some results were obtained using the facilities of both the Centre of collective usage 11Integrated information network of Irkutsk scientific educational complex” and the Centre of collective usage 11Integrated information network of Irkutsk scientific educational complex”.

References

1. **Yakubailik O.E., Kadochnikov A.A., Tokarev A.V.** Software for visualization of geospatial data in the system of operational satellite monitoring. Information Technologies in Remote Sensing of the Earth — RORSE 2018. Moscow; 2019: 264–270. (In Russ.)

2. **Amani M., Ghorbanian A., Ahmadi S.** Google Earth Engine cloud computing platform for remote sensing big data applications: a comprehensive review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2020; (13):5326–5350. DOI:10.1109/JSTARS.2020.3021052.
3. **Sangeetham R., Reddy N.** A review on Google Earth Engine: an open access cloud analysis platform for planetary scale satellite data. *E3S Web of Conferences*. 2024; (591):09011. DOI:10.1051/e3sconf/202459109011.
4. **Habibie I.** The application of machine learning using Google Earth Engine for remote sensing analysis. *Jurnal Teknoinfo*. 2022; 16(2):233. DOI:10.33365/jti.v16i2.1872. Available at: https://www.researchgate.net/publication/361870523_the_application_of_machine_learning_using_google_earth_engine_for_remote_sensing_analysis.
5. **Xu H., Man Y., Yang M., Wu J., Zhang Q., Wang J.** Analytical insight of Earth: a cloud-platform of intelligent computing for geospatial big data. *ArXiv*. 2023; abs/2312.16385. DOI:10.48550/arXiv.2312.16385.
6. **Novikova A.M., Kashirina E.S., Novikov A.A., Polonsky A.B., Pankeeva T.V.** GIS in marine research: world experience and abilities of its application on the example of the Black Sea region. *Trudy Karadagskoy Nauchnoy Stantsii im. T.I. Vyazemskogo Prirodnogo Zapovednika RAN*. 2021; 1(3):54–66. (In Russ.)
7. **Loupian E.A., Proshin A.A., Bourtsev M.A., Kashnitskii A.V., Balashov I.V.** Vega-Science system: design features, main capabilities and usage experience. *Sovremennye Problemy Distantionnogo Zondirovaniya Zemli iz Kosmosa*. 2021; 18(6):9–31. Available at: <http://jr.rse.cosmos.ru/article.aspx?id=2457&lang=eng>. (In Russ.)
8. **Kurdi F.T., Amakhchan W., Gharineiat Z.** Random forest machine learning technique for automatic vegetation detection and modelling in LiDAR data. *International Journal of Environmental Sciences & Natural Resources*. 2021; 28(2):556234. DOI:10.19080/IJESNR.2021.28.556234.
9. **Lee S., Hyun Y., Lee S., Lee M.** Groundwater potential mapping using remote sensing and GIS-based machine learning techniques. *Remote Sensing*. 2020; 12(7):1200. DOI:10.3390/rs12071200.
10. **Prasetyadi A., Sutyawan A.G., Nur W.H., Bisri A., Hanifa N.R., Shomim A.F.** Integrating remote sensing and machine learning for monitoring urban growth in seismically active regions: a case study of the Lembang fault zone, West Java, Indonesia. *International Conference on Computer, Control, Informatics and its Applications (IC3INA)*. Bandung, Indonesia; 2024: 208–212. DOI:10.1109/IC3INA64086.2024.10732244.
11. **Popova A.** Improving the accuracy of tree species mapping by Sentinel-2 images using auxiliary data — a case study of Slyudyanskoye forestry area near Lake Baikal. *Forests*. 2025; (16):487. DOI:10.3390/f16030487.
12. **Mansourmoghaddam M., Rousta I., Malamiri H.G., Sadeghnejad M., Krzyszczyk J., Ferreira C.S.S.** Modeling and estimating the land surface temperature (LST) using remote sensing and machine learning (case study: Yazd, Iran). *Remote Sensing*. 2024; 16(3):454. DOI:10.3390/rs16030454.
13. **Waske B., Fauvel M., Benediktsson J.A., Chanussot J.** Machine learning techniques in remote sensing data analysis. *Kernel Methods for Remote Sensing Data Analysis*. N.Y.: Wiley; 2009: 434. DOI:10.1002/9780470748992.ch1.
14. **Rodriguez-Galiano V., Ghimire B., Rogan J., Chica-Olmo M.** An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2012; (67):93–104. DOI:10.1016/J.ISPRSJPRS.2011.11.002.
15. **Sheykhmousa M., Mahdianpari M., Ghanbari H.** Support vector machine versus random forest for remote sensing image classification: a meta-analysis and systematic review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2020; (13):6308–6325. DOI:10.1109/JSTARS.2020.3026724.
16. **Chen J., Xia J., Du P., Chanussot J., Xue Z., Xie X.** Kernel supervised ensemble classifier for the classification of hyperspectral data using few labeled samples. *Remote Sensing*. 2016; 8(7):601. DOI:10.3390/rs8070601.
17. **Hoeser C., Zhang G.** Trends in image object detection and segmentation with deep neural network on natural and Earth observations. *International Journal of Computer Science and Information Technology*. 2024; 4(3):36–46. DOI:10.62051/ijcsit.v4n3.05.
18. **Pilato A., Taggio N., Pompili A.** Deep learning approaches to Earth Observation change detection. *Remote Sensing*. 2021; 13(20):4083. DOI:10.3390/rs13204083.

19. **Irtaza G., Latif S., Nadeem R.M., Hussain N., Ijaz H.M.** Real-time satellite image classification using convolutional neural networks for Earth observation and video feed analysis. *Kashf Journal of Multidisciplinary Research*. 2025; 2(03):204–216. DOI:10.71146/kjmr365.
20. **Mateo-Garcia G., Aybar C., Acciarini G., Ruzicka V., Meoni G., Longepe N.** Onboard cloud detection and atmospheric correction with deep learning emulators. *2023 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2023)*. 2023: 1875–1878. DOI:10.1109/IGARSS52108.2023.10282605.
21. **Bychkov I.V., Ruzhnikov G.M., Fedorov R.K., Popova A.K., Avramenko Y.V.** On classification of Sentinel-2 satellite images by a neural network ResNet-50. *Computer Optics*. 2023; 47(3):474–481. DOI:10.18287/2412-6179-CO-1216.
22. **Gadiraju K.K., Vatsavai R.R.** Remote sensing based crop type classification via deep transfer learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2023; (16):4699–4712. DOI:10.1109/JSTARS.2023.3270141.
23. **Xie M., Jean N., Burke M., Lobell D., Ermon S.** Transfer learning from deep features for remote sensing and poverty mapping. *Proceedings of the AAAI Conference on Artificial Intelligence*. Phoenix; 2015; 30(1). DOI:10.1609/aaai.v30i1.9906. Available at: https://www.researchgate.net/publication/282403466_Transfer_Learning_from_Deep_Features_for_Remote_Sensing_and_Poverty_Mapping.
24. **Pires de Lima R., Marfurt K.** Convolutional neural network for remote-sensing scene classification: transfer learning analysis. *Remote Sensing*. 2019; 12(1):86. DOI:10.3390/rs12010086.
25. **Huber F., Inderka A., Steinhage V.** Leveraging remote sensing data for yield prediction with deep transfer learning. *Sensors*. 2024; 24(3):770. DOI:10.3390/s24030770.
26. **Fawzy M., Khodary F., Mostafa Y.** Selecting training samples automatically from VHR satellite images for image classification. *Sohag Engineering Journal*. 2021; 1(1):71–84. DOI:10.21608/SEJ.2021.155929.
27. **Zhan Y.** Comparing three methods of selecting training samples in supervised classification of multispectral remote sensing images. *Sensors*. 2023; 23(20):8530. DOI:10.3390/s23208530.
28. **Kirillov A., Mintun E., Ravi N., Mao H., Rolland C., Gustafson L.** Segment anything. *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. Paris; 2023: 3992–4003. DOI:10.1109/ICCV51070.2023.00371.