

Д. А. Стригун¹✉, В. В. Дедкова¹

Разработка алгоритма автоматической сегментации и классификации изображений с применением искусственного интеллекта

¹Сибирский государственный университет геосистем и технологий,
г. Новосибирск, Российская Федерация
e-mail: linkeert@yandex.ru

Аннотация. Статья посвящена разработке алгоритма, способного в автоматическом режиме проводить семантическую сегментацию и выдавать классификацию по сегментированным областям. Целью работы является разработка приложения, использующего алгоритм машинного обучения на основе трансформеров для сегментации и классификации изображений. Основной задачей является классификация и сегментирование земельных покровов на изображениях из большого объема данных, получаемых из различных источников, перед их использованием для планирования и анализа. В рамках данной работы в качестве основы для нейросетевой модели была выбрана архитектура SegFormer, предназначенная для решения задач семантической сегментации изображений. Разработанное веб-приложение построено на основе клиент-серверной архитектуры. В практической части работы приведены результаты реализации приложения, процесс интеграции модели в серверную часть, а также результаты тестирования на реальных данных аэрофотосъемки.

Ключевые слова: веб-приложение, сегментация, классификация

D. A. Strigun¹✉, V. V. Dedkova¹

Development of an algorithm for automatic segmentation and classification of images using artificial intelligence

¹Siberian State University of Geosystems and Technologies, Novosibirsk, Russian Federation
e-mail: linkeert@yandex.ru

Abstract. The article is devoted to the development of an algorithm capable of automatically performing semantic segmentation and classifying segmented areas. The aim of the work is to develop an application using a transformer-based machine learning algorithm for image segmentation and classification. The main task is to classify and segment land cover in images from a large amount of data obtained from various sources before using them for planning and analysis. In the framework of this work, the SegFormer architecture was chosen as the basis for the neural network model, designed to solve the problems of semantic image segmentation. The developed web application is based on a client-server architecture. The practical part of the paper presents the results of the application implementation, the process of integrating the model into the server side, as well as the results of testing on real aerial photography data.

Keywords: web application, segmentation, classification

Введение

Современные технологии анализа изображений находят широкое применение в геодезии и дистанционном зондировании Земли. В последние годы наблюдается значительный рост объема данных, поступающих с различных источников, включая спутники, беспилотные летательные аппараты (БПЛА) и другие средства. Эти данные представляют собой огромный потенциал для анализа, однако их ручная обработка становится все менее эффективной из-за больших объемов и высокой сложности структур. В связи с этим актуальной задачей становится автоматизация процессов анализа изображений, в том числе с применением методов искусственного интеллекта, что позволяет существенно сократить время обработки и повысить ее точность.

Цель исследования заключается в усовершенствовании алгоритмов автоматической сегментации и классификации изображений, полученных методами дистанционного зондирования, с использованием современных технологий искусственного интеллекта.

Для достижения цели должны быть решены следующие задачи:

- аналитический обзор литературных источников по теме исследования;
- усовершенствование алгоритмов автоматической сегментации и классификации изображений с учетом особенностей данных дистанционного зондирования, таких как пространственное и спектральное разрешение;
- адаптация методов глубокого обучения к работе с изображениями, полученными с БПЛА и спутников, включая данные с различной спектральной характеристикой;
- оценка эффективности и точности разработанных алгоритмов на основе эталонных данных, а также выявление факторов, влияющих на их производительность.

Практическая значимость исследования заключается в разработке и внедрении системы, которая способна значительно повысить эффективность и точность обработки данных дистанционного зондирования. Ниже перечислены основные аспекты значимости такой системы.

Методы и материалы

Методы искусственного интеллекта, в частности глубокое обучение, зарекомендовали себя как одни из наиболее перспективных технологий для решения задач обработки изображений. Они позволяют эффективно решать задачи сегментации и классификации, выявляя мелкие детали и сложные взаимосвязи в данных. Применение таких методов в геодезии и дистанционном зондировании открывает новые возможности для анализа территорий, мониторинга изменений в ландшафте, картографирования и управления природными ресурсами. Кроме того, автоматизированные системы находят применение в мониторинге чрезвычайных ситуаций, таких как наводнения, лесные пожары и оползни, обеспечивая своевременную реакцию и принятие решений.

Современными методами основанных на нейронных сетях можно считать классификацию и сегментацию [2].

Семантическая сегментация заключается в том, что нейронная сеть обучается распознавать и выделять на изображении области, принадлежащие различным классам, таким как лес, здание, дорога, озеро и т.д. В отличие от простой классификации, где изображению присваивается один общий класс, сегментация позволяет определить, какие именно части изображения относятся к тому или иному объекту.

Для решения задачи семантической сегментации используется специальный тип нейросетей, таких как SegFormer [1], способных извлекать пространственные и контекстные признаки изображения и точно локализовывать объекты на уровне пикселей. Это позволяет не только идентифицировать присутствие объектов на изображении, но и определить их точное расположение и форму.

В разработке приложения использованы языки программирования Python для реализации серверной части и JavaScript для создания клиентской части. Приложение предназначено для автоматической сегментации и классификации изображений, полученных, например, с беспилотных летательных аппаратов, с применением нейросетевых технологий. Основными фреймворками и библиотеками, использованными при реализации, являются:

- FastAPI - для создания быстрого и удобного веб-сервиса на стороне сервера [6];
- Hugging Face Transformers - для загрузки и использования предварительно обученной модели SegFormer;
- PyTorch - в качестве фреймворка для запуска и возможного дообучения модели;
- NumPy - для работы с числовыми массивами и обработкой результатов сегментации [4];
- OpenCV - для чтения, преобразования и визуализации изображений [3].

Точность сегментации, скорость обработки, адаптированность к веб-приложению, удобство использования и возможность масштабирования - ключевые показатели успешной реализации данного приложения.

В рамках данной работы в качестве основной архитектуры нейронной сети была выбрана SegFormer - современная модель для семантической сегментации, которая сочетает в себе преимущества трансформеров и легких декодеров, обеспечивая высокую точность и эффективность на различных типах изображений. SegFormer способна обрабатывать изображения различного разрешения и точно определять границы объектов, что делает ее особенно полезной для анализа данных дистанционного зондирования.

Результаты

Для подготовки данных использовались аэрофотоснимки, собранные через автоматизированные парсеры с платформ Google Images и Яндекс.Картинки. Это обеспечило широкий спектр изображений различных типов ландшафтов и увеличило разнообразие входных данных.

С целью создания качественного набора для обучения модели, изображения проходили этап предобработки:

- приведение к единому разрешению – для унификации размеров и упрощения дальнейшей обработки;
- нормализация цветовой гаммы – для устранения перепадов яркости и контраста между разными источниками;
- разметка классов объектов – была выполнена вручную на платформе MakeSense.ai с использованием полигональной разметки по классам: лес, вода, дорога, здание, поле.

Приложение было разделено на две основные части:

- фронтенд – пользовательский интерфейс, через который осуществляется загрузка изображений, получение результатов сегментации и классификации;
- бэкенд – серверная логика, обеспечивающая загрузку обученной модели, предварительную обработку изображений, выполнение инференса и передачу результата обратно на клиентскую сторону.

Пользовательский интерфейс веб-приложения был создан с использованием технологий HTML, CSS и JavaScript [5]. На данном этапе реализована базовая версия интерфейса, предназначенная для загрузки изображений, отображения результатов сегментации и представления классификации объектов.

Следует отметить, что текущий внешний вид интерфейса не является окончательным и в дальнейшем может быть доработан и улучшен для повышения удобства взаимодействия и визуальной привлекательности.

Внешний вид разработанного на данном этапе интерфейса представлен на рис. 2.

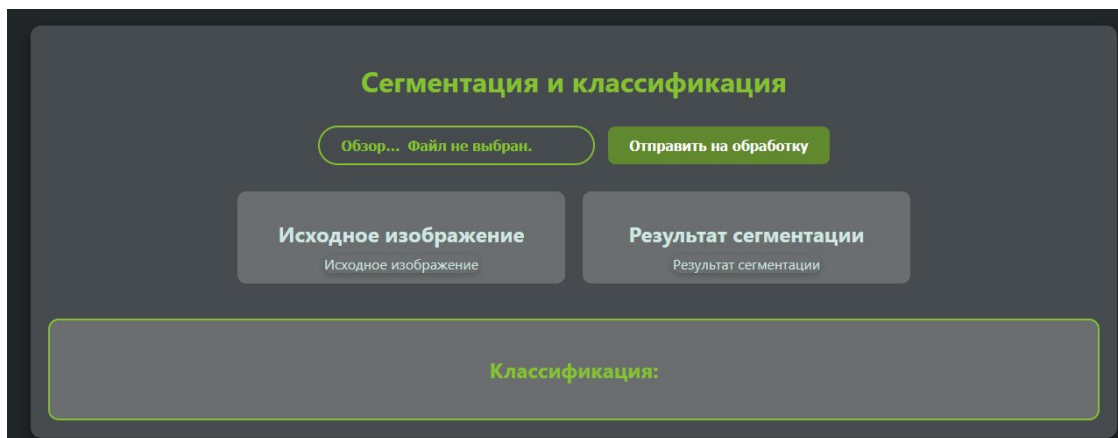


Рис. 2. Интерфейс приложения

После успешного дообучения модели SegFormer на специально подготовленном наборе данных, следующим этапом стало ее интегрирование в серверную часть веб-приложения. Этот процесс обеспечил связующее звено между пользовательским интерфейсом и аналитическим ядром на основе нейронной сети.

При получении запроса от пользователя, содержащего изображение для обработки, сервер инициирует последовательность действий:

- загруженное изображение декодируется и преобразуется в формат массива NumPy, который необходим для работы модели;
- выполняется предобработка изображения: приведение к требуемому размеру и нормализация значений пикселей, аналогично условиям обучения модели;
- происходит передача подготовленного изображения в модель SegFormer, которая осуществляет сегментацию и классификацию земельных покровов.

На клиентской стороне используется технология AJAX для асинхронной отправки изображений на сервер и получения результатов без перезагрузки страницы. После обработки пользователю отображается исходное изображение, результат сегментации, а также сводная классификация с процентным распределением обнаруженных классов.

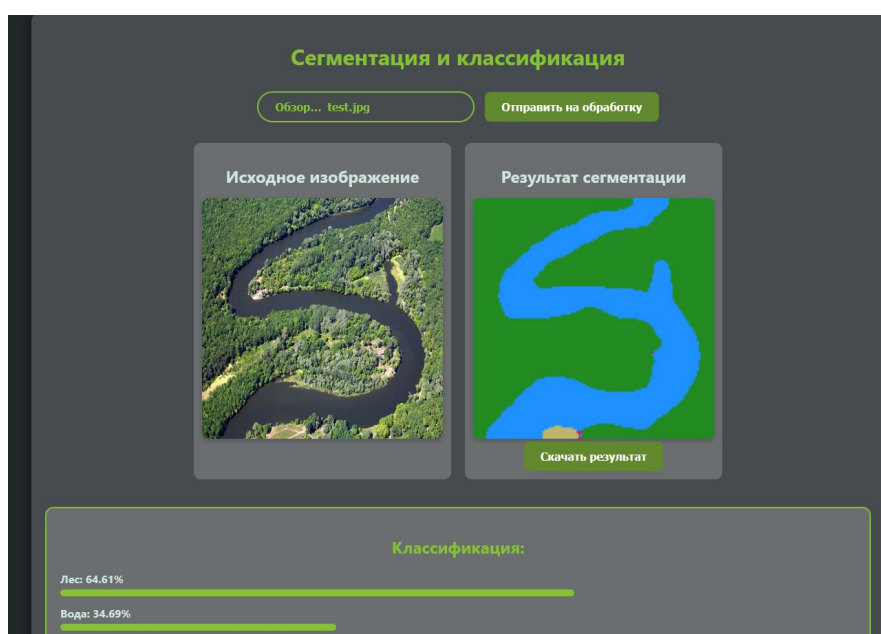


Рис. 2. Результат работы приложения

Обсуждение

Учитывая масштабы и разнообразие данных, которые необходимо анализировать в геодезии и дистанционном зондировании, традиционные методы обработки уже не способны удовлетворять текущим требованиям. Возрастающая потребность в высокоточной и оперативной обработке изображений делает внедрение автоматизированных методов, основанных на искусственном интеллекте, критически важным. Это особенно актуально в контексте задач, связанных с глобальным изменением климата, урбанизацией, а также необходимостью решения задач управления территориальными ресурсами и кадастрового учета.

Заключение

В рамках данной работы было разработано веб-приложение для автоматической сегментации и классификации земельных покровов на основе методов

искусственного интеллекта. Проект построен на архитектуре клиент-сервер, где для реализации серверной части использован язык программирования Python совместно с фреймворком FastAPI, а пользовательский интерфейс разработан с использованием технологий HTML, CSS и JavaScript.

Практическая значимость разработанного приложения заключается в возможности автоматизированной обработки аэрофотоснимков и оперативного выявления типов земельных покровов. Это позволяет значительно сократить временные затраты на ручную интерпретацию изображений, повысить точность классификации, а также минимизировать влияние субъективного человеческого фактора.

Результаты работы могут быть востребованы в таких сферах, как геодезия, дистанционное зондирование Земли, сельское хозяйство, мониторинг окружающей среды и управление земельными ресурсами. Внедрение подобных решений способствует улучшению качества картографической продукции, оптимизации планирования территорий и принятию обоснованных решений на основе данных дистанционного зондирования.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Власов П.Е., Романов А.В. Опыт использования модели SegFormer для анализа растительности по данным БПЛА // Геоинформатика, 2023 №2.
2. Киселев М.А. Сравнительный анализ архитектур нейронных сетей для задач семантической сегментации ландшафтов // Программные продукты и системы, 2022.
3. Освоение OpenCV с помощью Python: Полное руководство по обработке изображений и компьютерному зрению // Machine learning: – 2023. – URL: <https://vc.ru/u/1389654-machine-learning/661520-osvoenie-opencv-s-pomoshyu-python-polnoe-rukovodstvo-po-obrabotke-izobrazhenii-i-kompyuternomu-zreniyu> (дата обращения: 22.01.2025).
4. Обработка изображений с помощью Python, NumPy // Обработка изображений: – 2020. – URL: <https://note.nkmk.me/en/python-numpy-image-processing/> (дата обращения: 05.02.2024).
5. Мандел, Т. Разработка пользовательского интерфейса / Мандел Т. - Москва: ДМК Пресс. - 416 с. (Серия "Для программистов") - ISBN 5-94074-069-3. - Текст: электронный // ЭБС "Консультант студента": [сайт]. - URL: <https://www.studentlibrary.ru/book/ISBN5940740693.html> (дата обращения: 07.04.2024).
6. FastAPI Руководство пользователя: [сайт]. – URL: <https://fastapi.tiangolo.com/> (дата обращения: 09.04.2024). – Текст: электронный.

© Д. А. Стригун, В. В. Дедкова, 2026