

М. В. Коломеец¹, А. А. Колесников^{1}*

Оценка влияния количества информации на качество автоматизированного дешифрирования спутниковых снимков

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

Аннотация. Статья посвящена выявлению закономерности влияния количества информации на качество автоматизированного дешифрирования спутниковых снимков. Проведен обзор алгоритмов классификации спутниковых снимков на основе искусственного интеллекта, а также обработки космических снимков методами машинного обучения, сверточных нейронных сетей и геоинформационных систем. Кроме того, проведены экспериментальные работы с использованием модуля Semi-Automatic Classification Plugin, с открытым исходным кодом для QGIS, который позволяет выполнять полуавтоматическую классификацию изображений дистанционного зондирования. В ходе экспериментов было выявлено, что количество информации, содержащейся на спутниковых снимках, имеет прямую зависимость с качеством автоматизированного дешифрирования. Наличие информации на снимках и их разрешение являются ключевыми факторами, которые оказывают значительное влияние на качество автоматизированного дешифрирования. Эти результаты могут быть полезными для исследователей и специалистов в области обработки космических снимков, которые могут использовать эти знания для улучшения точности и скорости дешифрирования.

Ключевые слова: спутниковые снимки, автоматизированное дешифрирование, данные ДЗЗ, количество информации

M. V. Kolomeets¹, A. A. Kolesnikov^{1}*

Assessment of the influence of the amount of information on the quality of automated interpretation of satellite images

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

Abstract. The article is devoted to revealing the regularity of the influence of the amount of information on the quality of automated interpretation of satellite images. A review of algorithms for classifying satellite images based on artificial intelligence, as well as processing space images using machine learning methods, convolutional neural networks and geographic information systems, was carried out. In addition, experimental work was carried out using the Semi-Automatic Classification Plugin, an open source module for QGIS, which allows you to perform semi-automatic classification of remote sensing images. During the experiments, it was found that the amount of information contained in satellite images is directly dependent on the quality of automated decoding. The presence of information in images and their resolution are key factors that have a significant impact on the quality of automated interpretation. These results may be useful for researchers and specialists in the field of satellite image processing, who can use this knowledge to improve the accuracy and speed of interpretation.

Keywords: satellite images, automated interpretation, remote sensing data, amount of information

Введение

Современные технологии позволяют собирать огромные объемы информации с помощью спутниковых снимков аэрокосмического наблюдения. Однако обработка и дешифрирование этих снимков являются трудоемкой задачей для человека. В связи с этим, активно ведется работа по созданию систем автоматизированного дешифрирования, которые могут обрабатывать большое количество информации в короткие сроки.

Одним из главных факторов, который влияет на качество автоматизированного дешифрирования, является количество информации. Оно может зависеть от многих факторов, например, разрешения снимка, качества изображений, спектрального диапазона и т.д. Новые методы и алгоритмы, разрабатываемые и внедряемые исследователями, позволяют улучшать точность и скорость обработки данных, и способствуют развитию этой технологии в будущем.

Методы и материалы

Алгоритмы классификации спутниковых изображений на основе искусственного интеллекта являются инновационным подходом к обработке данных, полученных с помощью спутникового обзора Земли. Они позволяют автоматически определять типы покрытия земной поверхности на основе анализа изображений. Алгоритмы классификации используются для многих приложений, включая географические информационные системы, прогнозирование изменений климата, мониторинг природных ресурсов и многое другое. Эти алгоритмы обеспечивают быструю и точную обработку больших объемов данных, что делает их важным инструментом для исследований и приложений в области землеустройства и окружающей среды. Они основаны на использовании нейронных сетей, которые обучаются распознавать определенные характеристики объектов на изображениях, такие как размер, форма, цвет и текстура. Примеры применения алгоритмов классификации спутниковых изображений на основе искусственного интеллекта включают мониторинг лесных пожаров, оценку рисков природных катастроф, подготовку картографических и садово-парковых работ.

Некоторые из наиболее распространенных алгоритмов классификации, которые могут быть использованы для дешифрирования спутниковых снимков, включают в себя:

- деревья решений: алгоритм деревьев решений используется для построения дерева решений на основе обучающей выборки. Это дерево может потом быть использовано для классификации новых объектов на основе атрибутов объектов;

- k-ближайших соседей: алгоритм k-ближайших соседей используется для нахождения k количество ближайших соседей для новых объектов и на основе этого определения класса;

- метод опорных векторов: метод опорных векторов используется для проекции точек на гиперплоскости и классификации на основе разделения пространства на классы.

Методы сверточных нейронных сетей (Convolutional Neural Networks - CNNs), которые являются наиболее распространенным типом нейронных сетей для обработки изображений.

Для реализации сегментации изображений в сверточной нейронной сети используется сверточный слой (Convolutional Layer), который фильтрует изображение через каскад слоев. Каждый слой анализирует определенные признаки изображения, такие как цвет, текстура и форма, и передает информацию своим последующим слоям. Эта информация дополнительно обрабатывается и переносится на следующий слой, который анализирует более высокоуровневые признаки.

Результаты

Сегодня существует широкий спектр программных продуктов для обработки данных, полученных с ДЗЗ. Для работы с аэрокосмической информацией наиболее популярным выбором является использование современных растровых и интегрированных ГИС-пакетов, которые предоставляют широкие возможности для обработки снимков. Все программы для обработки снимков содержат типичный набор процедур, включая предварительную коррекцию, трансформирование, классификацию и многое другое. В целом, программные продукты для обработки изображений достаточно похожи друг на друга по возможностям цифровой обработки.

При выполнении процессов предварительной обработки и визуализации результатов предполагается использовать QGIS и модуль Semi-Automatic Classification Plugin. Для создания, оценки и анализа влияния методов предварительной обработки, на качество математических моделей автоматизированного дешифрирования будут использованы Scikit-learn и PyTorch.

Модуль Semi-Automatic Classification Plugin (SCP) – это плагин, который используется для автоматической классификации изображений и картографических данных. Он предоставляет инструменты для создания классификационных моделей, обучения алгоритмов машинного обучения и применения этих моделей к новым данным.

SCP поддерживает несколько алгоритмов классификации, включая мультиспектральную классификацию, классификацию на основе объектов и классификацию на основе текстуры. Плагин также обеспечивает интеграцию с открытыми источниками данных, такими как Google Earth Engine и Sentinel Hub.

Принцип работы модуля SCP заключается в следующих шагах:

- ввод и подготовка данных – подготовка входных данных, которые должны иметь одинаковую пространственную разрешающую способность и проекцию, где входные данные могут быть растровыми или векторными;
- извлечение признаков – процесс извлечения информации из входных данных, такой как текстуры, цвета, формы и т. д.;
- обучение классификатора – создание модели классификатора на основе обучающих данных, при котором обучающие данные это изображения, где объ-

екты имеют реальные значения классификации, которые используются для обучения классификатора;

- классификация – применение обученной модели классификатора к новым наборам изображений;

- постобработка – обработка результатов классификации с использованием различных алгоритмов для удаления шума и улучшения качества классификации;

- проверка классификации – оценка точности классификации путем сравнения результатов классификации с реальными значениями классов.

С целью экономии времени, для оценки влияния количества информации на автоматизированное дешифрирование был использован пробный участок территории Новосибирска.

В Semi-Automatic Classification Plugin можно непосредственно осуществлять поиск исходных данных для экспериментальных исследований. При помощи ключевых запросов, таких как координаты и дата съемки местности, а также даты доступности снимков и уровень облачности, можно подобрать наиболее оптимальный результат, не выходя из программы.

Далее необходимо выполнить преобразование уже имеющихся изображений, в том числе настроить длину волны, необходимую для анализа спектральных характеристик, и отобразить композицию. Предоставление каналов в цветовой композиции позволяет лучше понимать, какие части земли отображаются в каждом канале и как они соотносятся друг с другом. Это упрощает выбор оптимального канала, чтобы получить наилучшие результаты в дальнейшем исследовании.

Данное исследование базируется на использовании ярко выраженных водных и лесных участках, таких как Новосибирское водохранилище и Караканский бор. Оставшаяся часть снимка заполнена пашнями. На рис. 1 показаны примеры фрагментов снимков с различным набором каналов (4-3-2 и 3-2-1), которые использовались для сравнения точности дешифрирования.

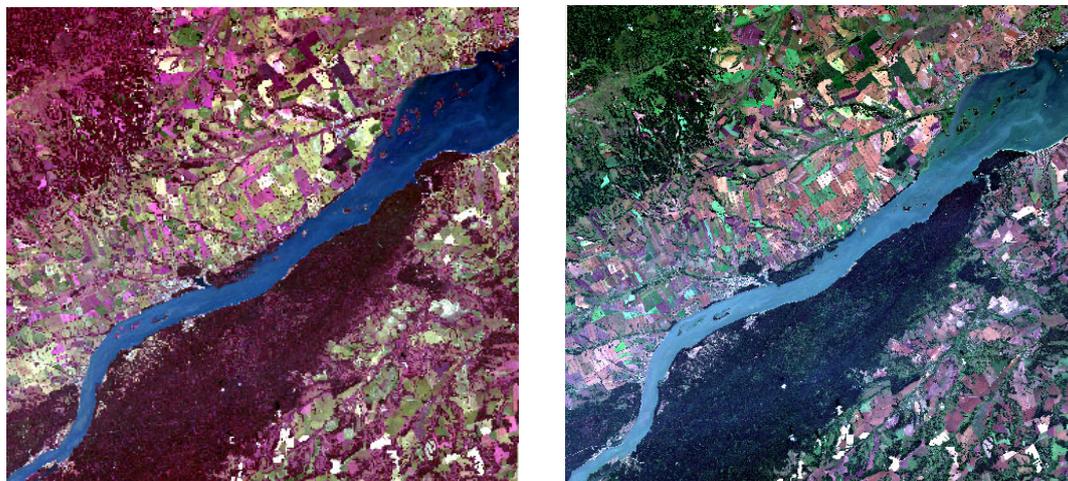


Рис. 1. Фрагментов снимков Sentinel-2 с различным набором каналов (4-3-2 и 3-2-1)

Для автоматизированной обработки данных необходимо определить несколько областей интереса (ROI), которые будут относиться к определенным классам и макроклассам. Каждая область интереса связана с идентификатором класса (С ID) и назначается классу земной поверхности с помощью идентификатора макрокласса (МС ID). Макроклассы состоят из нескольких образцов, их спектральные характеристики могут отличаться. Для достижения лучших результатов классификации необходимо создавать различные образцы, которые будут помогать разделять спектральные характеристики разных примеров, относящихся к одному макроклассу. Кроме того, нужно создать несколько областей интереса для каждого макрокласса (используя одинаковый идентификатор МС ID, но различивая каждой области интереса разные идентификаторы С ID).

Для получения более яркого представления результатов автоматической классификации необходимо назначить соответствующие цвета для каждого макрокласса, например, использование синего для водных объектов, зеленого для лесов, желтого для равнин и пашен и серого для населенных пунктов. Результат классификации выглядит следующим образом – рис. 2.

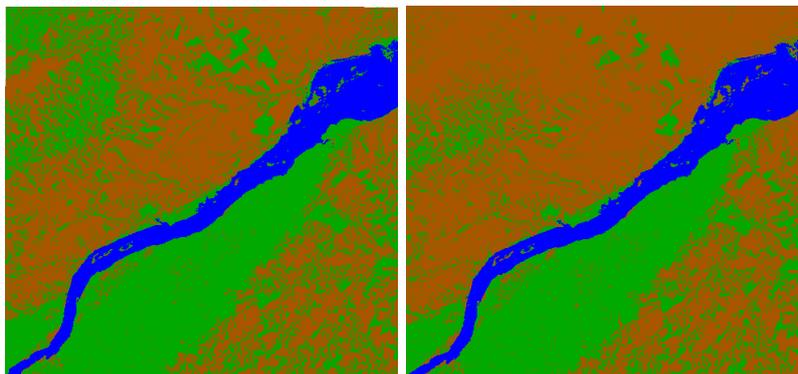


Рис. 2. Примеры результатов классификации наборов каналов (4-3-2 и 3-2-1)

В ходе проведения экспериментальных исследований на предмет достоверности полученных результатов использовался метод определения точности Accuracy.

В результате анализа формируется следующий вывод, что наилучшим для дешифрирования растительности и большинства других объектов является следующий порядок назначения цвета: Red - первая, Green - вторая, Blue - третья. При данном порядке назначения цвета лесная растительность выделяется пурпурным цветом, вода - голубым и темно-синим, пашни и поля – оттенками зеленого.

Проведенный анализ комплексной методики расшифровки космических снимков, которая позволяет использовать несколько алгоритмов и определять различные типы объектов. В отличие от методов, основанных на спектральных признаках, существует другая группа методов, которая использует структурные признаки. Структурные методы имеют меньшую зависимость от внешних фак-

торов, так как они основаны не на спектральной яркости изображения объектов местности, а на других, более устойчивых характеристиках. Эти методы позволяют выделить информацию на снимках, которую невозможно обнаружить, используя только спектральные свойства. Результаты сравнения вариантов сочетаний каналов приведены в табл. 1.

Таблица 1

Сводная таблица результатов точности для различных комбинаций каналов

Исследуемый набор каналов	Вычисления для водных участков	Вычисления для лесных участков	Вычисления для равнинных участков	Общая вычисленная Accuracy
весь набор каналов Sentinel-2	96,4	30,9	84,4	68,6 %
RGB (4-3-2)	99,4	72,5	97,3	89,2 %
RGB и ближний инфракрасный (4-3-2 + 8A)	99,9	98,1	87,8	95,3 %
RGB и дальний инфракрасный (4-3-2 + 11);	99,8	94,8	63,2	85,9 %
спектральная характеристика NDVI;	90,2	67,9	31,7	61,5 %
RGB и спектральная характеристика NDVI (4-3-2 + NDVI)	89,9	75,9	33,8	69,2 %

Заключение

Уже по представленным расчётам, точность правильной классификации лесных массивов довольно высока для большей части исследований. Наиболее низкая доля неверного определения пикселя как пикселя массива леса достигается при классификации всего набора каналов Sentinel-2 и составляет 30,9%, а высокая доля распознавания пришлась на набор RGB и ближнего инфракрасного и составила 98,1%.

Точность правильной классификации пикселей пашни составляет от 31,7 до 97,26%. Наибольшие значения (97,3%) характерны для комбинаций 4, 3 и 2 набора каналов. Худший результат представлен в результате работы с характеристикой NDVI. Точность правильной классификации пикселей объектов гидрографии высока и не опускается ниже отметки 89%. В ходе экспериментальных работ выяснилось, что распознавание объектом местности может иметь прямую зависимость от набора каналов для области интереса.

Кроме того, важную роль играет разрешение снимков. Чем выше разрешение снимка, тем выше качество дешифрирования, так как объекты на них более детализированы и точны. Но при этом также возрастает количество информации, которую нужно обработать, что может вызвать трудности у автоматизированных систем дешифрирования.

Таким образом, наличие информации на спутниковых снимках и разрешение этих снимков оказывают значительное влияние на качество автоматизированного дешифрирования.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Алтынцев, М.А. Разработка методик автоматизированного дешифрирования многозональных космических снимков высокого разрешения для мониторинга природно-территориальных комплексов: автореферат диссертации на соискание ученой степени кандидата технических наук // Сибирская государственная геодезическая академия. – Новосибирск: 2011. – 25 с.
2. Головастова, Е. С. Автоматическое дешифрирование природных объектов по материалам дистанционного зондирования Земли // Тенденции, направления и перспективы развития экономических отношений в современных условиях хозяйствования / Крымский федеральный университет им. В.И. Вернадского. – Симферополь: 2019. – Материалы IV межрегиональной с международным участием научно-практической конференции: сборник трудов. – С. 24-27.
3. Гордиенко, А.С. Разработка методики многоступенчатого дешифрирования космических снимков / А.С. Гордиенко, М.А. Алтынцев, С.А. Арбузов. // Геодезия и аэрофотосъемка. – 2011. – №2. – С. 29-33.
4. Интеллектуальная система поддержки принятия решений при автоматическом дешифрировании видовой информации космического и воздушного мониторинга / Ю. Г. Веселов, А. А. Киселев, Е. Г. Симаков. // Состояние и перспективы развития современной науки по направлению «Техническое зрение и распознавание образов» / Федеральное государственное автономное учреждение "Военный инновационный технополис «ЭРА». - Анапа:2019. - Сборник тезисов докладов научно-технической конференции. – С. 182-186.
5. Augmenting a deep-learning algorithm with canal inspection knowledge for reliable water leak detection from multispectral satellite images / Jiawei Chen, Pingbo Tang, Todd, Michael Patrick, Xiran Zhou. - 2020.
6. Learn to Optimize Panchromatic Imagery for Accurate Building Extraction / Yuxuan Wang, Guangming Wu, Chuyao Feng, Ryosuke Shibasaki. – 2019.
7. Comparison of multi-source satellite images for classifying marsh vegetation using DeepLabV3 Plus deep learning algorithm Man / Liu, Bolin Fu, Shuyu Xie, Hongchang He, Feiwu Lan, Yuyang Li, Peiqing Lou, Donglin Fan. – 2021.

© М. В. Коломеец, А. А. Колесников, 2023