

ОСОБЕННОСТИ ИСПОЛЬЗОВАНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ПРИ РАБОТЕ С ПРОСТРАНСТВЕННЫМИ ДАННЫМИ

© 2022 А. Г. Субботин, А. В. Линкина

Воронежский институт высоких технологий (Воронеж, Россия)

В статье рассматриваются некоторые вопросы, связанные с характеристиками ГИС. Отмечены особенности решения задач с пространственными, временными, графическими и текстовыми данными. Приведены особенности использования искусственного интеллекта в геовизуализации. Описываются возможности разработки математических методов анализа многомерных данных с использованием машинного обучения. Дано сравнительное описание алгоритмического подхода и подходов с использованием обучения нейронных сетей.

Ключевые слова: геовизуализация, пространственные данные, искусственный интеллект, ГИС, машинное обучение, предиктивная аналитика.

Повсеместное внедрение цифровых технологий способствует ускорению научно-технического прогресса, расширению экономических связей, повышению качества жизни граждан. Инструменты информационных систем постоянно совершенствуются, активно развивается машинное обучение, которое позволяет частично или полностью автоматизировать решение наиболее сложных задач в самых разнообразных сферах человеческой деятельности. Особенностью использования искусственного интеллекта является необходимость работы с большим количеством данных.

Задачи, возникающие в геоинформационных системах (ГИС) на современном этапе, связаны с необходимостью обработки самых разнообразных типов данных-изображений, пространственных и временных данных, текстовой информации, вычислений. Развитие информационных технологий привело к значительному расширению возможностей по хранению, обработке и представлению пространственных данных. Это дало новую веху развития такому направлению картографии и геоинформатики как геовизуализация.

Интерактивность и динамика стали главными отличительными чертами современных картографических произведений, а

разработки в области картографического дизайна теперь распространяются на проблемы взаимодействия человека и компьютера для обеспечения более успешного проведения анализа геоданных и выработки пространственных решений [1].

Одной из основных задач ГИС, которую ставят перед собой исследователи, является получение новых знаний, представлений о природе пространственных данных. В то же время пользователи иногда недооценивают возможности ГИС в области поддержки принятия решений, которые эти системы могут обеспечивать, уделяя основное внимание главным образом представлению, в частности визуализации данных. Ценность географической информации в системах поддержки принятия решений становится особенно значимой, когда в ГИС включаются программные средства, базирующиеся на технологиях и методах искусственного интеллекта. Поэтому их широкое внедрение в геоинформационные технологии становится все более актуально и осуществляется по следующим направлениям: расширение функциональной полноты традиционных методов, технологий и программных средств пространственного анализа в ГИС, за счет использования возможностей развитых математических методов анализа многомерных данных; развитие

Субботин Андрей Геннадьевич – Воронежский институт высоких технологий, магистрант, e-mail: anna_linkina@rambler.ru.

Линкина Анна Вячеславовна – Воронежский институт высоких технологий, старший преподаватель, e-mail: anna_linkina@rambler.ru.

новых методов, основанных на интеллектуальных вычислительных технологиях, как базы для создания следующего поколения удобных и более сильных инструментальных программных средств анализа геоданных в условиях все возрастающих объемов первичной информации; создание новых моделей данных, информационных технологий и программных средств, специально предназначенных для многомерного анализа данных, моделирования и прогноза территориально распределенных процессов и обеспечивающих интеграцию с традиционными ГИС.

За последние несколько лет исследователи разработали целый класс статистических и адаптивных методов анализа многомерных данных, получивших название нейросетевых методов. Нейросетевые методы применяются не только для анализа данных, но и, что существенно, для построения моделей процессов, разворачивающихся в многомерных пространствах. Уже сегодня предлагаются интересные классы нелинейных моделей, построенных на основе статистического анализа первичных данных. При этом средства информационных технологий используются для организации доступа и предобработки первичных данных, хранящихся в ГИС и базах данных (БД). Статистические и адаптивные методы анализа геоданных позволяют улучшить качество исходной информации и построить нейросетевую модель, адекватную как по назначению и качеству исходных данных, так и по суждениям экспертов и задачам исследователей [2]. Общую структуру организации данных можно увидеть на рисунке 1.



Рисунок 1. Общая схема анализа геоданных

Пространственные данные – наиболее устойчивые и наименее меняющиеся, поэтому в этой группе следует искать основу для интеграции. Среди пространственных данных наиболее устойчивыми (наименее изменчивыми) являются координаты. Именно они являются основой для объединения различных данных, т. е. основой для интеграции [3].

В настоящее время во многих сферах все чаще возникают задачи по интеллектуальной обработке и анализу изображений, в том числе снимков, используемых в геоинформационных системах. Одним из интеллектуальных методов для решения такого рода задач является машинное обучение, как одно из средств искусственного интеллекта. В программных продуктах для обработки геоинформационных данных начинают появляться инструменты для интеллектуального анализа данных.

Например, в программном продукте ArcGIS разработаны модули, основанные на современных алгоритмах для принятия решений и анализа данных. Модули включают в себя инструментарий, позволяющий реализовать модели машинного обучения путем включения пространственных данных и алгоритмов, выполнять классификацию данных дистанционного зондирования.

В качестве входных данных для нейронной сети используются пространственные данные, в том числе пространственные данные ГИС. Нейросетевые алгоритмы динамически модифицируют характеристики существующих объектов, отображенных в ГИС, и могут создавать новые объекты и даже новые слои карт.

Основными инструментами для обучения нейросети являются языки программирования Python и TensorFlow – открытая программная библиотека для машинного обучения, разработанная компанией Google для решения задач построения и тренировки нейронной сети с целью автоматического нахождения и классификации образов, достигая качества человеческого восприятия.

Изначально необходимо выполнить обучение нейронной сети на основе подготовленного набора данных, для выполнения семантической сегментации растрового изображения (рис. 2).

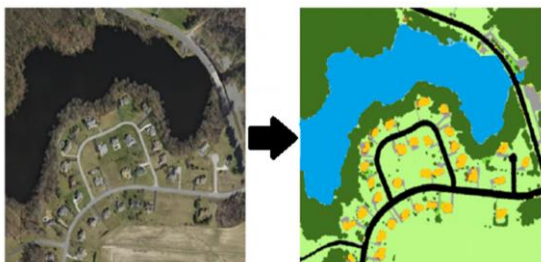


Рисунок 2. Результат обработки изображения

Семантическая сегментация изображений – это разделение изображения на отдельные группы пикселей, области, соответствующие одному объекту с одновременным определением типа объекта в каждой области. Именно при помощи сегментации изображения нейронная сеть определяет принадлежность того или иного пикселя изображения к искомой группе объектов, в данном случае зданиям и сооружениям. Огромное значение при обучении играет качество подготовки данных, их полнота и условия, в которых они были получены. Например, если фотоснимок был получен в зимнее время года, то нейронная сеть не определит здания на снимке, или определит с большой погрешностью. Для устранения этой проблемы требуется дополнительное обучение на основе снимков, которые были получены в зимнее время года. В целом, точность построения модели напрямую зависит от качества обучения нейронной сети. Результат обучения сети приведен на рисунке 3.



Рисунок 3. Результат обучения нейронной сети

Чем больше исходных данных, тем выше точность определения объектов на изображениях нейронной сетью. В отличие от общеизвестных классических методов, которые основаны на математических вычислениях, методы, работающие средствами нейронных сетей способны адаптироваться в

процессе обучения и построить математическое обоснование цифровой модели местности [4].

В заключении отметим, что внедрение искусственного интеллекта в инструментальной работе с геопространственными данными является крайне эффективным инструментом для решения производственных задач. Несмотря на то, что многие операции можно выполнять не за счет применения машинного обучения, а за счет алгоритмизации, возникает проблема необходимости подготовки датасетов. Очень часто данные являются неполными и разнородными, что усложняет поиск оптимального решения. При использовании нейронных сетей эта проблема, как отмечалось выше, достаточно просто решается. Кроме того, несмотря на более затратное по количеству времени и ресурсов обучение сетей, в дальнейшем возможно использовать методы предиктивной аналитики на основе полученных данных.

СПИСОК ИСТОЧНИКОВ

1. Балтыжакова Т. И. Основы геоинформатики: учебное пособие / Т. И. Балтыжакова. – Москва Ай Пи Ар Медиа, 2022. – 114 с. – ISBN 978-5-4497-1591-3. – Текст: электронный // IPR SMART: [сайт]. – URL: <https://www.iprbookshop.ru/119109.html> (дата обращения: 20.11.2022). – Режим доступа: для авторизир. пользователей.
2. Глотов А. А. Технологические барьеры развития геоинформационных технологий как предпосылки интеллектуализации / А. А. Глотов // Интеллектуализация геоинформационных систем: подходы и направления – 2015. [Электронный ресурс] режим доступа: <https://sovzond.ru/upload/iblock/651/18-24.pdf> (дата обращения 05.11.2022 г.)
3. Дэвид М. Путеводитель по цифровому будущему: отрасли, организации и профессии / М. Дэвид – М.: Альпина Паблишер, 2020 – 215 с.
4. Колесниченко О. Ю. Data Science (наука о данных) в становлении информационного общества: учебное пособие / О. Ю. Колесниченко. – Москва: Прометей, 2021. – 52 с. – ISBN 978-5-00172-110-9. – Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. – URL: <https://www.iprbookshop.ru/125600.html> (дата обращения: 09.11.2022). Режим доступа: для авторизир. пользователей.

5. Кольер, Р. Машинное обучение в Elastic Stack / Р. Кольер, К. Монтонен, Б. Азарми; перевод В. С. Яценкова. – Москва: ДМК Пресс, 2022. – 380 с. – ISBN 978-5-93700-107-8. – Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. – URL: <https://www.iprbookshop.ru/126320.html> (дата обращения: 05.12.2022). – Режим доступа: для авторизир. пользователей.

6. Линкина А. В. Применение методов искусственного интеллекта при работе с геопространственными данными / А. В. Линкина // Научно-технический прогресс в сельскохозяйственном производстве. Материалы XVII Международной научно-практической конференции. – Великие Луки, 2022. – С. 196-201.

7. Линкина А. В. Информационное обеспечение цифровых технологий в агропромышленном комплексе / А. В. Линкина, И. Ю. Богданчиков // Вестник Воронежского института высоких технологий. – 2021. – № 2 (37). – С. 25-27.

8. Львович А. И. Алгоритмизация процесса визуально-экспертного моделирования при оптимизации управления развитием организационных систем с использованием мониторинговой информации / А. И. Львович, А. П. Преображенский. // Моделирование, оптимизация и информационные технологии. – 2022. – Т. 10. – № 1 (36). – [Электронный ресурс]: <https://moitvvt.ru/ru/journal/article?id=1154> (дата обращения 10.09.2022)

9. Мельникова Т. В. Моделирование обработки больших массивов данных в рас-

пределенных информационно-телекоммуникационных системах / Т. В. Мельникова, М. В. Питолин, Ю. П. Преображенский // Моделирование, оптимизация и информационные технологии. – 2022. – Т. 10. №1 (36). – [Электронный ресурс]: <https://moitvvt.ru/journal/article?id=1117> (дата обращения: 10.09.2022).

10. Платонов А. В. Машинное обучение: учебное пособие для вузов / А. В. Платонов. – Москва: Издательство Юрайт, 2023. – 85 с. – (Высшее образование). – ISBN 978-5-534-15561-7. – Текст: электронный // Образовательная платформа Юрайт [сайт]. – URL: <https://urait.ru/bcode/520544> (дата обращения: 15.10.2022)

11. Федин Ф. О. Анализ данных. Часть 2. Инструменты Data Mining: учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва: Московский городской педагогический университет, 2012. – 308 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprbookshop.ru/26445.html> (дата обращения: 13.10.2022). – Режим доступа: для авторизир. пользователей.

12. Чубукова И. А. Data Mining: учебное пособие / И. А. Чубукова. – 3-е изд. – Москва, Саратов: Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. – 469 с. — ISBN 978-5-4497-0289-0. – Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. – URL: <https://www.iprbookshop.ru/89404.html> (дата обращения: 19.11.2022). – Режим доступа: для авторизир. пользователей.

FEATURES OF USING ARTIFICIAL INTELLIGENCE WHEN WORKING WITH SPATIAL DATA

© 2022 A. G. Sybbotin, A. V. Linkina

Voronezh Institute of High Technologies (Voronezh, Russia)

The article discusses some issues related to the characteristics of GIS. The features of solving problems with spatial, temporal, graphical and textual data are noted. The features of the use of AI in geovisualization are given. A description is given of the possibilities for developing mathematical methods for analyzing multidimensional data using machine learning. A comparative description of the algorithmic approach and approaches using neural network training is given.

Keywords: geovisualization, spatial data, artificial intelligence, GIS, machine learning, predictive analytics.