

A. C. Tumor^{1}*

Кластеризация изображений с применением нейронных сетей

¹ Томский государственный университет систем управления и радиоэлектроники,
г. Томск, Российская Федерация

* e-mail: office@tusur.ru

Анотация. В данной статье рассматривается проблема диагностирования уровня загрязненности экологической среды промышленными отходами нефтегазового комплекса. Одним из факторов, не позволяющих проведение диагностики обычными методами, является труднодоступность рассматриваемых регионов. Применение данных дистанционного зондирования позволит нивелировать влияние представленного фактора, а также уменьшить временные и экономические затраты на проведение диагностических исследований. Предполагается использование спутников для получения данных дистанционного зондирования требуемой территории, а также последующей их обработки. Наиболее актуальным методом обработки изображений является кластеризация с использованием нейросетевых алгоритмов. Одним из таких алгоритмов являются самоорганизующиеся карты Кохонена. Предлагается разработка методики обработки данных дистанционного зондирования с помощью такого алгоритма. Результатом выполнения разработанной методики будет являться сравнительный анализ загруженных данных за несколько лет, что позволит диагностировать неблагоприятное воздействие промышленных предприятий на выбранный регион.

Ключевые слова: нейросетевые алгоритмы, самоорганизующиеся карты Кохонена, спутниковые снимки

A. S. Titov^{1}*

Image clustering using neural networks

¹ Tomsk State University of Control Systems and Radioelectronics, Tomsk,
Russian Federation

* e-mail: office@tusur.ru

Abstract. This article discusses the problem of diagnosing the level of environmental pollution by industrial waste from the oil and gas complex. One of the factors that do not allow diagnostics by conventional methods is the inaccessibility of the regions under consideration. The use of remote sensing data will make it possible to level the influence of the presented factor, as well as to reduce the time and economic costs of conducting diagnostic studies. It is planned to use satellites to obtain remote sensing data of the required territory, as well as their subsequent processing. The most relevant method of image processing is clustering using neural network algorithms. One such algorithm is Kohonen's self-organizing maps. It is proposed to develop a technique for processing remote sensing data using such an algorithm. The result of the developed methodology will be a comparative analysis of the downloaded data for several years, which will allow diagnosing the adverse impact of industrial enterprises on the selected region.

Keywords: neural network algorithms, self-organizing Kohonen maps, satellite images

Введение

Изучение характера воздействия производственных предприятий и предприятий нефтегазового комплекса на природную среду является сложной и многокомпонентной задачей, актуальность которой только возрастает. Процессы разведки, добычи, транспортировки и переработки углеводородного сырья являются источниками негативного влияния на окружающую среду. В настоящее время разработано большое число подходов, методик и мониторинговых систем диагностики состояния природных экосистем с целью обнаружения аварийных ситуаций и оценки экологической обстановки в нефтедобывающих регионах [1]. Для сохранения и улучшения состояния природных экосистем, подверженных техногенному воздействию нефтегазодобывающих предприятий, необходимо регулярно проводить мониторинговые исследования и оценку состояния атмосферы, почвы, растительного покрова, поверхностных и подземных вод. Перспективным и экономически оправданным методом оценки экологического состояния отдаленных территорий является применение методики диагностирования экологических проблем с использованием данных дистанционного зондирования Земли (ДДЗ). Применение ДДЗ обусловлено их пространственным охватом исследуемых территорий, временным и спектральным разрешением изображений, актуальностью (своевременной доступностью) для исследователей и возможностью их бесплатного использования [2].

Методика (алгоритм) кластеризации спутниковых снимков с использованием самоорганизующихся карт Кохонена

В процессе обработки и анализа ДДЗ используют классификацию изображений для идентификации различных типов почвенно-растительного покрова. Важным методом неконтролируемой классификации является кластеризация, которая группирует набор паттернов или векторов в различные кластеры на основе внутренней однородности и внешнего разделения.

Кластеризация – это процесс объединения схожих объектов в кластеры по идентичным или схожим признакам. Внутри каждого кластера оказываются объекты со схожими признаками. Кластеризация от классификации отличается так, что в кластеризации не известно количество кластеров, количество которых определяется в процессе. Одной из важнейших задач при анализе больших массивов данных на определение скрытых закономерностей является кластеризация. Список областей применения практически неограничен, начиная от обработки изображений, анализа текстов и заканчивая маркетингом.

Самоорганизующиеся карты — это один из вариантов организации алгоритмов нейронных сетей. Одним из основных отличий данного метода организации от нейросетей, обучаемых по алгоритму обратного распространения, является отсутствие так называемого «учителя», то есть результат обучения зависит только от расположения и индексирования изначальных данных.

В последние годы для классификации изображений дистанционного зондирования широко используются искусственные нейронные сети, что обусловлено их неотъемлемыми преимуществами перед традиционными подходами.

Для анализа дистанционных материалов необходимы современные подходы и методы их обработки. Изучение спектральных характеристик почвенно-растительного покрова, полученных из спутниковых снимков, является одним из таких подходов. Одним из таких подходов является изучение спектральных характеристик почвенно-растительного покрова, полученных по спутниковым снимкам, с целью проведения автоматизированной классификации.

Представленная в работе методика, позволяет классифицировать почвенно-растительный покров на основе результатов обработки GeoTIFF-снимков. Из цветовых значений формируется результирующее значение интенсивности каждого пикселя. Полученные таким образом значения интенсивности задаются как входные параметры для самоорганизующихся карт Кохонена, после чего применяется алгоритм кластеризации, разделяющий пиксели изображения на 12 сегментов, позволяющих точно классифицировать выбранную местность и определить уровень загрязнения почвы.

После выполнения кластеризации и классификации полученных результатов выполняется сравнительный анализ за выбранный промежуток времени (от 1 года до 5 лет).

На основании полученных результатов можно сделать выводы о влиянии развития городской и промышленной инфраструктуры на загрязненность региона.

Заключение

В статье описана одна из проблем влияния производственных предприятий нефтегазового комплекса на природную среду. Одним из способов диагностики и предотвращения повышенного уровня загрязнения в экологической среде является предлагаемая методика диагностики растительного покрова территории с помощью нейронных сетей. Система, разрабатываемая в рамках данной работы, способна автоматически обрабатывать загружаемые изображения, а также сообщать пользователю об ухудшении или улучшении экологической ситуации в регионе.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Конторович А.Э., Эдер Л.В. Новая парадигма стратегии развития сырьевой базы нефтедобывающей промышленности Российской Федерации // Минеральные ресурсы России. Экономика и управление. 2015. № 5. С. 8-17.
2. Терехин Э.А. Сезонная динамика проективного покрытия растительности агроэкосистем на основе спектральной спутниковой информации / Э.А. Терехин // Современные проблемы дистанционного зондирования Земли из космоса. 2019. Т.16, № 4. С. 111-123.

© А. С. Тутов, 2023