
ОЦЕНКА БИОРАЗНООБРАЗИЯ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННЫХ СЕТЕЙ

Андрей Шелестов, Екатерина Насуро, Сергей Скакун

Резюме: Рассматривается роль обработки данных дистанционного зондирования Земли в задаче оценки биоразнообразия. Предлагается модификация метода оценки биоразнообразия за счет применения нейросетевой классификации подстилающей поверхности.

Ключевые слова: биоразнообразие, данные ДЗЗ, нейронные сети.

ACM Classification Keywords: F.1.1 Models of Computation (F.4.1) - Self-modifying machines (e.g., neural networks)

Введение

Основными признаками здоровья, благополучия и жизнеспособности биосферы Земли является количество и биоразнообразие экосистем, составляющих биосферу. Экосистемы обеспечивают ресурсы, необходимые для поддержания жизни на Земле и влияют на жизнедеятельность и процветание всех биологических видов, включая и человека [1].

Управление изменениями в биоразнообразии и понимание процессов, происходящих в экосистемах Земли возможно при наличии оперативной, достоверной и систематизированной информации. Научное изучение экосистем, своевременное информирование мировой общественности, консолидация усилий по защите окружающей среды позволит правильно интерпретировать знания об изменениях, происходящих в окружающей среде в масштабах всего земного шара, изучить роли экосистем и их взаимодействие.

"Система систем" GEOSS

Решением проблем глобализации и объединения усилий по сохранению здоровья биосферы Земли может стать "система систем" глобального наблюдения земли GEOSS. Она ориентирована на объединение существующих и будущих систем наблюдения Земли, развивающихся во многих странах мира. Глобальность системы позволяет привлечь наибольшее количество стран и наблюдать процессы, происходящие в масштабах, значительно больших, чем масштаб отдельно взятой страны. С другой стороны такая система позволит объединить усилия на более низком уровне, привлекая к решению задач функционирующие региональные системы [2].

Разработка высококачественной, современной и всесторонней глобальной системы мониторинга биоразнообразия, обеспечит данными все заинтересованные стороны. Среди которых: участники многосторонних соглашений по окружающей среде, правительства разных стран, органы по управлению природными ресурсами, научные деятели, структуры по экологии, сельскому хозяйству, здравоохранению, чрезвычайным ситуациям и гидрометео службы.

Среди ближайших приоритетов в создании GEOSS называют создание условий и координирование непрерывного наблюдения на всех уровнях, начиная от глобального мониторинга поверхности земли, до локальных измерений и обеспечение к полученным данным свободного доступа для любого пользователя системы. Применять данные дистанционного наблюдения Земли для описания, картографирования и мониторинга глобальных защищенных зон, включая зоны Всемирного наследия, природные области,

области большого культурного, геологического и археологического значения. Использовать данные наблюдения Земли и другие пространственные данные для обеспечения оконтуривания и корректировки границ защищенных зон.

Роль спутников ДЗЗ в программах по сохранению биоразнообразия

Необходимость сохранить хрупкие экосистемы и их биоразнообразие усложнена их распределением по всему земному шару, удаленностью и национальными границами. В таких случаях важную роль играют спутники дистанционного зондирования Земли. Такие данные предоставляют различную информацию, необходимую для принятия природоохранных решений. Проекты по дистанционному зондированию Земли используют возможности мультиспектральных оптических и радарных приборов спутников для анализа изменений поверхности земли, определения массы растительности, распознавания типов растительности и обнаружения и картографирования отдельных биологических видов. Эти данные полезны для составления графиков, диаграмм, таблиц основных факторов окружающей среды. Анализ такой информации позволит сформировать элементы и параметры биоразнообразия: развитие растений, их среду обитания и климат.

Биоразнообразие

Биоразнообразие – это разнообразие видов жизни на Земле. По мнению авторов [2] биоразнообразие имеет структуру, состоящую из трех основных уровней организации и трех различных аспектов разнообразия. Уровни биоразнообразия делятся на генетический, видовой уровень и уровень экосистем. Аспектами разнообразия называются композиция, структура и функции (см. рис.1).

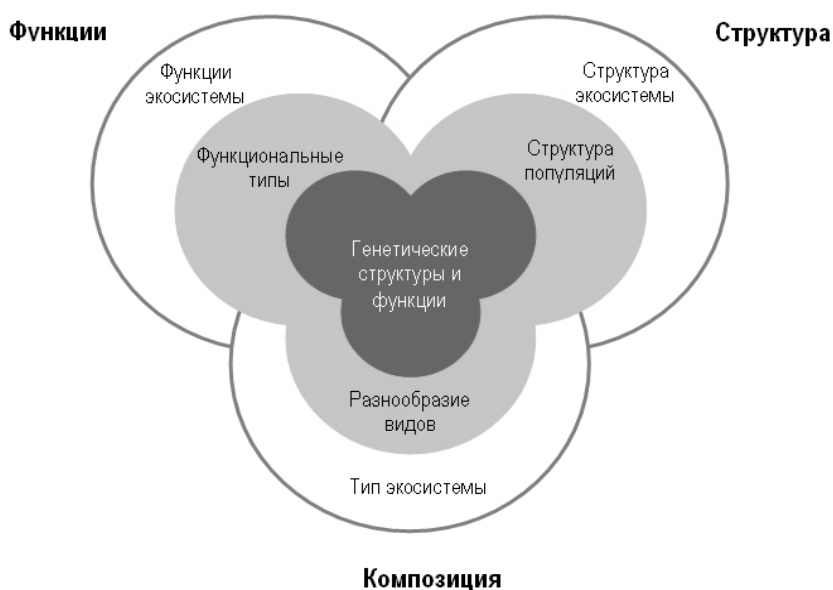


Рисунок 1. Концептуальная диаграмма, отображающая структурные уровни биоразнообразия

Биоразнообразие необходимо для сохранения жизни на земле в целом. Объединение различных сервисов охраны окружающей среды позволит осознать важность биоразнообразия, степень и последствия его истощения, построить глобальную, локальные и национальные системы, принимать правильные решения по сохранению биоразнообразия и успешно использовать результаты охранных мер.

Методика вычисления биоразнообразия

Успешное сбережение биоразнообразия невозможно без разработки четкой научной концепции его оценивания и мониторинга. Перспективы изучения и картографирования биоразнообразия тесно связаны с использованием материалов ДЗЗ. Важным заданием во время реализации этого подхода является определение показателей биоразнообразия, которые могут быть получены с использованием данных ДЗЗ [3].

Биоразнообразии часто ассоциируется с количеством видов, как в конкретной группе, так и по всему ландшафту в целом. Видовое разнообразие состоит из двух компонентов. Во-первых, видовое богатство, т.е. общее число существующих видов. Во-вторых, равномерность распределения видов, которая базируется на определении положения вида в структуре доминирования на основе мер значимости [4].

При оценивании видового разнообразия широко используется индекс Шеннона, который учитывает не только богатство видов, но и равномерность распределения этих видов по любому из признаков [5]

$$B = \sum_{i=1}^N P_i \log_2 P_i \quad (1)$$

где N – количество видов, P_i – значимость вида i .

Использование данных дистанционного зондирования Земли и современных цифровых технологий обработки аэрокосмических изображений предоставляет ряд преимуществ в процессе расчета индекса видового разнообразия по Шеннону. Например, позволяет значительно расширить территорию исследований, отображать реальную информацию о пространственном распределении групп растений и предоставить возможность постоянного обновления этих данных. Кроме этого, представление данных в цифровом виде позволит обеспечить их эффективную обработку и визуализацию.

Для построения методики оценки биоразнообразия используется экосистемный подход, предложенный А.Н.Киселевым [6]. Его главный принцип состоит в том, что оценка биоразнообразия может опираться на любые интегральные характеристики, которые обеспечивают "представительство" всех биокомпонентов. Основной идеей экосистемного подхода является следующее: элементы живой и неживой природы находятся в тесной связи и взаимодействии друг-с другом, то есть создают эко- и геосистемы различных уровней и длительности. В упрощенном виде принцип взаимодействия элементов этих систем выглядит следующим образом: условия обитания влияют на процессы заселения территории видами растительности, что в свою очередь обеспечивает условия для развития определенных видов животных.

Для связи биоразнообразия и ряда влияющих на него факторов используется модель, основанная на нечеткой логике [7]. Она позволяет описать классические динамические связи с неопределенными или слабоопределенными параметрами и статистические зависимости с неизвестными законами распределения, а так же формализовать экспертные представления про характер процессов в исследуемых экосистемах.

Если известны количественные значения $F_i, i = 1..n$ факторов, которые влияют на оценку биоразнообразия B , где n – общее количество факторов, для каждого из m классов ландшафта земного покрытия можно определить потенциально возможное биоразнообразие $B_0^{(j)}, j = 1..n$, которое при использовании индекса Шеннона (1) определяется количеством характерных биологических видов данного класса $N_0^{(j)}$ по формуле:

$$B_0^{(j)} = \log_2 N_0^{(j)} \quad (2)$$

Базой для оценки биоразнообразия поверхности служит классификация покрытий ландшафта. Ее можно получить по данным ДЗЗ и наземным измерениям. Именно по классификации выполняется оценка

потенциального биоразнообразия для каждого класса земной поверхности с использованием формулы (2). Если возможности космической съемки позволяют выполнить видовую классификацию растительности, то есть возможность получить непосредственную оценку видимого биоразнообразия территории исследования, установить зависимость между видимым и скрытым биоразнообразием и вычислить полное биоразнообразие для каждого участка ландшафта.

Факторы, которые влияют на биоразнообразие и могут быть определены по материалам ДЗЗ, приведены в таблице 1. Сейчас в мире существует большое количество центров, которые предоставляют не только аэрокосмические снимки, но и информационные продукты более высоких уровней. Примером может служить международный научный продукт ДЗЗ EOS, в рамках которого предоставляется достаточно широкий спектр информационных продуктов, которые могут использоваться при оценке биоразнообразия.

Таблица 1. Факторы, которые влияют на биоразнообразие

Фактор	Показатель	Диапазон значений	Оптимальное значение
Рельеф местности	Среднесуточное облучение солнечным излучением	0 – 300 Вт/м ²	максимум
Вода	Влажность почв	0 – 100 %	максимум
Тепло	Дневная и ночная температура	253 – 323 К	293 К
Осадки	Среднегодовое количество осадков	0 – 200 мм	максимум
Растительность	Вегетационный индекс	-1 .. +1	максимум
Жизненная активность	Относительное количество усвоенного растительностью излучения	-1 .. +1	максимум
Жизненная эффективность	Чистая продуктивность растительности	0 – 600 кг карбона/м ² в год	максимум

В настоящее время существует множество программ, направленных на создание продуктов по классификации земной поверхности, которые отличаются системами классификации, покрытием, используемыми данными, технологиями построения и методами обработки. Среди них следует выделить европейский проект Corine [8] и американский проект National Land Cover Data (NLCD) [9], которые в качестве основы используют данные спутника Landsat, а также MODIS Land Cover and Land-Cover Change [10], который основан на данных прибора MODIS спутника Terra и использует глобальную систему классификации программы International Geosphere-Biosphere Program (IGBP) [11]. Стоит отметить, что для территории Украины доступны лишь данные по глобальной классификации земной поверхности с разрешением не менее 1 км. Кроме того, самые свежие данные классификации поверхности для Украины датируются 2000 годом. Поэтому актуальной является задача создания более эффективных методов классификации покрытий ландшафта.

Классификация данных дистанционного зондирования Земли

Применение нейронных сетей для классификации данных ДЗЗ

В настоящее время при выборе подхода к классификации земного покрова по данным дистанционного зондирования земли все чаще отдается предпочтение нейросетевым методам. Нейронные сети обеспечивают адаптивный и устойчивый подход к анализу и обобщению информации, который не требует априорных знаний о статистическом распределении данных. Существующее разнообразие различных

нейросетевых архитектур, а также методов обучения нейронных сетей, позволяют выбрать наиболее оптимальную для поставленной задачи конфигурацию, найти компромисс между скоростью, размерностью сети и эффективностью работы. В то же время такое разнообразие возможностей приводит к необходимости проводить большое количество экспериментов для получения достоверных и наиболее подходящих к задаче вариантов решения.

Наши исследования, проведенные в работах [12, 13, 14] раскрывают результаты применения различных нейросетевых архитектур, методов обучения нейронных сетей первого и второго порядка, использования сетей на основе радиальных базисных функций и сетей на основе теории адаптивного резонанса для обработки данных спутника Landsat. Полученные результаты показывают, что для разных классов наибольшее значение верно классифицированных образов достигалось для разных методов, что объясняется сложностью топологии исходных данных.

Для нейронных сетей с обучением с учителем, необходимо определить набор эталонных данных. Эталоном могут быть результаты замеров на местности, приведенные к нужному виду. Кроме этого, данные о спектральных свойствах объектов, можно получить в лаборатории или из библиотек спектральных кривых. В таком случае данные необходимо приводить к виду, информации, получаемой из космоса, в зависимости от числа каналов прибора и диапазонов измерения, а так же с учетом функций отклика прибора [15]. Кроме перечисленных методов доступны данные международных программ, например, проект Corine по классификации наземных объектов [8].

Сравнение результатов работы различных нейросетевых парадигм

Для сравнительного анализа работы различных нейронных сетей с обучением с учителем было использовано изображение, полученное мультиспектральным прибором ETM+ (Enhanced Thematic Mapper Plus) спутника Landsat-7 [16]. Прибор предоставляет данные с пространственным разрешением 30м. Для эксперимента были взяты каналы 1-5 и 7. Размерность входных векторов составляла 6 элементов. На выходе был представлен вектор из 8 элементов – по количеству определяемых классов. Для определения эталонных данных использовалась информация проекта Corine по классификации земной поверхности. На тестовом снимке были определены следующие 8 классов (в скобках указан код классификации согласно проекту Corine): лиственный лес (311), хвойный лес (312), смешанный лес (313), пашня (211), пастбище (231), вода (51), искусственные материалы (1), открытая местность с разреженной растительностью (33) [12, 13].

В процессе работы были рассмотрены и протестированы следующие нейросетевые парадигмы: многослойный перцептрон, обученный методом обратного распространения ошибки, EDBD, квази-Ньютоновским методом и методом сопряженных градиентов; нейронная сеть на основе радиальных базисных функций; нейронная сеть, основанная на теории адаптивного резонанса.

Выполненные эксперименты были направлены на определение оптимальной архитектуры нейронной сети, предназначенной для классификации мультиспектральных спутниковых снимков; анализ методов обучения нейронных сетей и их влияние на результаты работы сети; анализ раделимости данных и применимости нейронных сетей для решения задач такого класса. Сравнение результатов работы различных нейросетевых архитектур приведено в таблице 2.

Таблица 2. Сравнение результатов работы различных нейросетевых архитектур

Классы	Архитектура сети					
	Многослойный перцептрон				Радиальная базисная функция	Нейронные сети ARTMAP
	Обратного распространения	EDBD	Квази-Ньютоновский	Сопряженных градиентов		

		нения ошибки					
лиственный лес	1	99,6%	98,9%	100,0%	99,9%	99,9%	99,3%
хвойный лес	2	94,6%	85,7%	98,3%	96,3%	88,8%	80,2%
смешанный лес	3	53,4%	67,4%	0,5%	35,2%	57,7%	68,7%
пашня	4	91,2%	89,6%	87,2%	92,7%	97,1%	82,0%
пастбище	5	90,5%	91,4%	0,0%	0,0%	77,8%	75,8%
вода	6	97,8%	97,7%	0,0%	0,0%	92,6%	96,2%
искусственные материалы	7	81,4%	83,6%	58,8%	67,2%	70,8%	77,8%
открытая местность	8	89,6%	96,5%	0,0%	0,0%	79,4%	97,3%
Всего		87,9%	88,9%	58,5%	64,35%	85,6%	83,7%

Анализируя результаты проведенных экспериментов видно, что для разных классов наилучший результат достигался разными методами. Это можно объяснить сложностью входных данных и сложными свойствами разделимости. В среднем по всей выборке данных наилучшие результаты дает многослойный перцептрон, обученный методом EDBD. Кроме того, обучение этим методом занимает значительно меньше времени, по сравнению с другими методами [14].

Выводы и направления дальнейшей деятельности

В статье предложен метод оценки биоразнообразия на основе использования результатов нейросетевой классификации земного покрова, который позволяет выполнить видовую классификацию участка ландшафта по данным дистанционного зондирования.

Сложное разбиение поверхности земли на классы и подклассы влечет за собой значительное усложнение архитектуры данных. При увеличении количества классов усложняется архитектура нейронной сети, что приводит к замедлению процесса обучения и снижению точности классификации. Наиболее оптимальной с точки зрения разбиения данных на классы является нейронная сеть, которая разбивает предложенные образцы на два класса. Для улучшения классификации спутниковых снимков целесообразно проанализировать применение модульных нейронных сетей со сложной архитектурой. Разные модули такой сети, используя различные методы и архитектуры, проводят вначале грубое, а потом более точное разбиение данных на классы.

В докладе будет рассмотрена разработка структуры модульной нейронной сети, каждый модуль которой проводит разбиение на два класса, а результаты такой классификации являются входными данными для следующей нейронной сети, которая проведет более точную классификацию. Это даст возможность оценить преимущества такого подхода и применимость его для решения задач классификации спутниковых данных.

Список литературы

[1] Earth Observation Handbook. Committee on Earth Observation Satellites. www.eohandbook.com

- [2] Global Earth Observation System of Systems GEOSS. Group on Earth Observation. ESA Publication Divisions, Netherlands, February 2005.
- [3] Багатоспектральні методи дистанційного зондування Землі в задачах природокористування / За ред. В.І. Лялько та М.О. Попова.- Київ: Наукова думка, 2006.- 360 с.
- [4] Одум Ю. Экология.- Т.2.- Пер. с англ.– М.: Мир, 1986.– 376 с.
- [5] Протасов А.А. Биоразнообразие и его оценка: Концептуальная диверсикоология.- К.: Институт гидробиологии, 2002.- 108 с.
- [6] Киселев А.Н. Оценка и картографирование биологического разнообразия (на примере Приморья) // Геоботаническое картографирование, 2000.- Т.3.- С.3–15.
- [7] Козлова А.А., Станкевич С.А. Оценка и картографирование биоразнообразия Северо-Причерноморского региона Украины на основании многоспектральных космоснимков и геоинформационных технологий // Сборник тезисов Шестой украинской конференции по космическим исследованиям.- Евпатория: ИКИ НАН Украины, 2006.- С.117.
- [8] NASA Landsat 7, <http://landsat.gsfc.nasa.gov>
- [9] National Land Cover Data (NLCD) [<http://www.epa.gov/mlrc/nlcd.html>]
- [10] Strahler Alan, et al. MODIS Land Cover and Land-Cover Change, Algorithm Theoretical Basis Document (ATBD), Version 5.0, 1999. – 72 p.
- [11] Belward A.S., Estes J.E., Kline K.D. The IGBP-DIS Global 1-km Land-Cover Data Set DISCover: A Project Overview// Photogram. Eng. Remote Sens. — 1999. — 65. — P. 1013-1020.
- [12] N. Kussul, S. Skakun, O. Kussul. Comparative Analysis of Neural Networks and Statistical Approaches to Remote Sensing Image Classification// International Scientific Journal of "Computing". – 2006. – Vol. 5, Issue 2. – P 93-99.
- [13] Куссуль Н.Н., Скакун С.В., Куссуль О.М. Анализ нейросетевых и статистических подходов к классификации изображений спутника Landsat-7 // Интелектуальні системи прийняття рішень і прикладні аспекти інформаційних технологій: Матеріали науково-практичної конференції. Том 1. – Херсон: Видавництво Херсонського морського інституту, 2006. – С. 117-120.
- [14] Скакун С.В., Насуро Е.В., Лавренюк А.Н., Куссуль О.М. Анализ применимости нейронных сетей для классификации спутниковых данных // Проблемы управления и информатики. — 2007. — №2. — С. 105-117.
- [15] Лавренюк А.Н., Яровая Е.В., Скакун С.В. Нейросетевая классификация спектральных характеристик наземных объектов // Кибернетика и вычислительная техника, выпуск 148, 2005. – С. 21-28.
- [16] Landsat-7 Science Data User's Handbook. – http://ftpwww.gsfc.nasa.gov/las/handbook/handbook_toc.html

Информация об авторах

Шелестов Андрей Юрьевич - Институт космических исследований НАН и НКА Украины, кандидат технических наук, старший научный сотрудник, докторант; проспект Академика Глушкова 40, 03680 Киев, Украина; e-mail: inform@ikd.kiev.ua.

Насуро Екатерина Валериевна - Национальный авиационный университет, аспирантка; проспект Комарова 1, 03058 Киев, Украина; e-mail: mail_aspir@ukr.net.

Скакун Сергей Васильевич - Институт космических исследований НАН и НКА Украины, кандидат технических наук, научный сотрудник; проспект Академика Глушкова 40, 03680 Киев, Украина; e-mail: inform@ikd.kiev.ua.