



Научная статья

МЕТОД ФОРМИРОВАНИЯ КЛАССОВ ПРОСТРАНСТВЕННО РАСПРЕДЕЛЕННЫХ ОБЪЕКТОВ С ЦЕЛЬЮ ИХ КЛАССИФИКАЦИИ И РАСПОЗНАВАНИЯ НА МНОГОСПЕКТРАЛЬНЫХ КОСМИЧЕСКИХ ИЗОБРАЖЕНИЯХ ВЫСОКОГО РАЗРЕШЕНИЯ

Рашад. Г. Нематзаде^{ID}, Гюнель. З. Рзаева^{ID}, Рауф. М. Рагимов^{ID} PhD, Фарид. Р. Самедов^{ID}

Научно- Исследовательский Институт Аэрокосмической Информатики, Баку, Азербайджан; Rashad.Nematzade@nasa.az (РГН), Gunel.Rzayeva@nasa.az (ГЗР), rauffmahmud@yahoo.com (РМР), farid.samedov@hotmail.com (ФРС)

*Автор корреспонденции: Фарид Р. Самедов, farid.samedov@hotmail.com

КЛЮЧЕВЫЕ СЛОВА

классификация,
объект,
методика,
космический снимок,
виды,
растительность

АБСТРАКТ

В работе рассматриваются методы автоматизированной классификации типов растительности и почвенного покрова на основе аэрокосмических данных высокого пространственного разрешения и геоинформационных технологий. Исследование выполнено с использованием космического снимка IKONOS и включает формирование обучающих и тестовых выборок для 12 классов растительности и почв, характерных для исследуемого региона. Проведён статистический анализ спектральных характеристик кластеров и оценка их репрезентативности и разделяемости. Показано, что неравномерное распределение обучающих примеров и перекрытие спектральных признаков отдельных классов снижают устойчивость классификации. Для повышения качества распознавания предложен подход, основанный на объединении спектрально близких классов и формировании альтернативных классификационных схем. Эффективность метода оценена с использованием статистического классификатора максимального правдоподобия и нейросетевого классификатора типа многослойного перцептрона. Полученные результаты подтверждают целесообразность оптимизации структуры классификационной схемы и состава обучающих выборок при решении задач аэрокосмического экологического мониторинга.

По статье:

Получено: 03.07.2025

Пересмотрено: 22.12.2025

Принято: 22.12.2025

Опубликовано: 30.12.2025

1. ВВЕДЕНИЕ

Аэрокосмические методы мониторинга экологической обстановки в исследуемом регионе предполагают регулярное отслеживание изменений ключевых индикаторов, включая климатические параметры, распределение и деградацию растительных сообществ, а также химические характеристики почвенного покрова. При выявлении неблагоприятных изменений в природно-экологическом равновесии становится необходимым установить первопричины этих процессов – как с точки зрения их влияния на окружающую среду, так и с позиции пространственно-временного распределения.

К традиционным сферам применения данных и методов дистанционного зондирования относятся, прежде всего, исследования природных ресурсов, мониторинг экологической обстановки и оценка антропогенного воздействия на окружающую среду, мониторинг для оценки состояния и роста культивируемой растительности, исследования, связанные с оценкой сейсмической и оползневой активностью в зоне прокладки магистральных нефте-газо проводов и т.д. Однако, в связи с появлением космических снимков высокого пространственного разрешения и высокоточных космических данных, изменились характер и акценты исследований задач в области

Для цитирования:

Нематзаде Р., Рзаева Г., Рагимов Р., Самедов Ф. Метод формирования классов пространственно распределенных объектов с целью их классификации и распознавания на многоспектральных космических изображениях высокого разрешения // Гидрометеорология и экология, 120 (5), 2025, 66-78.

дистанционного зондирования. Кроме того, космоснимки метрового и субметрового разрешения стали применяться в таких областях, где они раньше не применялись или применялись в очень ограниченном виде, например, создание крупномасштабных электронных карт, поиск и обнаружение новых месторождений нефти и газа, прогнозирование урожайности сельскохозяйственных культур, экологический мониторинг геологической среды и т.д. [1...2].

Существенные изменения претерпели формы и подходы применения данных дистанционного зондирования в связи с разработкой и внедрением геоинформационных технологий, когда ГИС технологии и технологии дистанционного зондирования, фактически интегрировались. При этом подходе, абстрактная постановка основной задачи дистанционного зондирования – классификация, распознавание и дешифрирование объектов – трансформируется в задачу исследования геораспределенных ресурсов, где из них выбираются распознаваемые объекты-индикаторы, которые характеризуют состояние и особенности распространения объектов, могут адекватно отображать ситуацию.

На начальной стадии анализа определяются координаты географических точек, в которых зафиксировано присутствие определённых видов растительности. Эти точки подвергаются геокодированию на основе спутниковых снимков, после чего выделяются участки территории, используемые в качестве обучающих и тестовых площадок для последующей работы алгоритма классификации.

В таблице 1 и рисунке 1 представлены наименования отдельных видов растительности, характерных для рассматриваемого региона.

Таблица 1*Первоначальный набор классов и типов растительности и почв*

Номер класса	Полное название растения
Класс 1	Болотистая тростниковая растительность
Класс 2	Болотистый кустарниковый тамариск (Tamarix)
Класс 3	Прибрежная зона: полупустынная растительность
Класс 4	Тростник ложный австралийский (Phragmites australis)
Класс 5	Солянка древовидная (Salsola ericoides)
Класс 6	Солянка гористая (Salsola nodulosa)
Класс 7	Солянка гористая (Salsola nodulosa) / Полынь Лерха (Artemisia lerchiana)
Класс 8	Солянка гористая (Salsola Nodulosa) / Трава
Класс 9	Полупустынная растительность - поташник каспийский, (Kalidium capsicum)
Класс 10	Полупустынная растительность с доминированием Верблюжьей колючки (Alhagi pseudoalhagi)
Класс 11	Голая почва
Класс 12	Солянка гористая (Salsola nodulosa) / голая почва

2. МАТЕРИАЛЫ И МЕТОДЫ

Возможности, проектируемой ГИС для распознавания и классификации объектов позволяют не только автоматизировать процессы извлечения тестовых примеров из космического снимка в границах выделенных участков, но и использовать для этого различные алгоритмы.

В первом случае используется алгоритм, который учитывает все пиксели и имеет незначительное перекрытие с выделенным участком. Во втором случае пиксели, пересекающиеся с границей участка, добавляются в набор только в том случае, если площадь перекрытия составляет больше половины площади одного пикселя (1 м², в данном случае).

В таблице 2 приведены количественные данные по разным классам растительного покрова, которые сформированы по космическому снимку исследуемого региона со спутника IKONOS. В таблицу добавлены вычисленные статистические характеристики, которые позволяют получить предварительные оценки по репрезентативности кластеров [3...5]. Другие комбинации объектов классов во втором случае. Поэтому количественно отличаются друг от друга.



Растительность болотного типа с преобладанием тростника



Заболоченные участки с кустарниковыми формами тамариска (Tamarix)



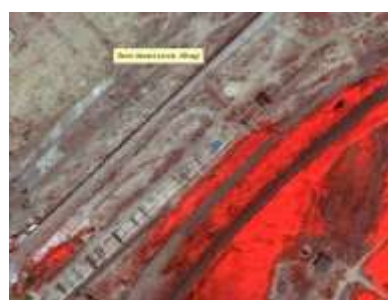
Горный тип солянки (Salsola nodulosa)



Древовидная форма солянки (Salsola ericoides)



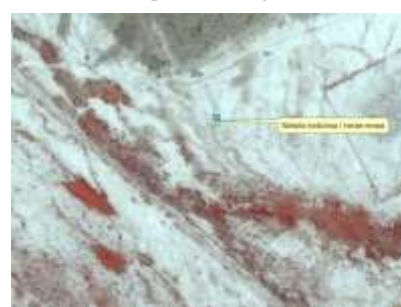
Полупустынные сообщества с доминированием поташника каспийского (Kalidium capsicum)



Полупустынные растительные сообщества, где преобладает Верблюжья колючка (Alhagi pseudoalhagi)



Обнажённый почвенный покров



Горная солянка (Salsola nodulosa) и открытый почвенный покров

Рисунок 1. Некоторые виды растительности на исследуемой территории

Из таблицы видно, что распределение примеров по классам неравномерно. Это связано, как с общим количеством примеров, извлекаемых из разных классов растительности и типов почв, представленных на сцене, так и с характером фрагментарности ареалов распространения различных типов растений. В первую очередь рассмотрим оценки по их общему количеству:

$$N_T = 5N_w$$

$$N_T = 30pN_0$$

$$N_T = 10N_w$$

$$N_T = 30N_1(N_1^{+1})$$

$$N_w = N_H(N_1 + N_0)$$

где, N_1 – количество исходных нейронов в рассматриваемом случае соответствует числу спектральных каналов спутникового изображения; N_0 – число выходных нейронов, соответствующих распознаваемым классам объектов; N_H – количество нейронов, размещённых в скрытом уровне нейросети; N_w – количество весовых коэффициентов; p – число нейронов или элементов, приходящихся на один класс, варьируется в диапазоне от 3 до 8.

Таблица 2

Размеры тренировочных и тестовых образцов, определенные для выделенных 12 классов растительности и почвы.

Идентификатор класса	Количество примеров. Процедура 1	Количество примеров. Процедура 2	Количество примеров. Процедура 1	Количество примеров. Процедура 2
Класс 1	1215	1113	1487	1368
Класс 2	3181	2955	2087	1893
Класс 3	97	67	63	33
Класс 4	1891	1748	2055	1859
Класс 5	234	186	279	176
Класс 6	2393	1992	2867	2658
Класс 7	690	564	478	388
Класс 8	200	182	147	124
Класс 9	52	35	49	30
Класс 10	40	33	28	21
Класс 11	474	242	553	287
Класс 12	680	604	579	472
Общее	11147	9721	10672	9309
Математическое ожидание	928,9167	810,0833	889,3333	775,75
Среднеквадратическое отклонение	1036,354	950,3847	976,7583	917,4441

Используя приведенные формулы, определены минимальное и максимальное значение необходимого количества нейронов классификатора:

$$N_{Hmin}=57$$

$$N_{Hmax}=120$$

$$N_{wmin}=8(4+12)=684$$

$$N_{wmax}=120(8+12)=2400$$

На основе проанализированных данных выявлен значительный диапазон значений, что создает определённые сложности при установлении оптимального порогового значения для выбора необходимого объема обучающих примеров. Наиболее обоснованными считаются те оценки, в которых учитывается число весовых параметров модели. В целях повышения точности анализа, произведена оценка распределения обучающих примеров по категориям объектов с использованием интервалов, определённых на основе крайних значений – минимального и максимального количества доступных примеров.

$$CI = [N_{av} - \alpha N_{msd}, N_{av} + \alpha N_{msd}]$$

$$N_{av} = (N_T / N_0)$$

где, N_{av} - математическое ожидание, N_T - общее количество примеров для обучения или тестирования; N_0 – количество классов, α – действительное число в интервале чисел

от одного до трех, N_{msd} – среднеквадратическое отклонение количества примеров по различным классам [6, 7].

3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Рассмотрены статистические тесты на репрезентативность и разделяемость образцов (таблица 3).

Из таблицы 3 следует, что некоторые кластеры или пересекаются или частично совпадают. В статье были рассмотрены различные комбинации гистограмм по различным вариантам комбинаций классифицируемых классов растительности. Для удобства проведения анализа распознаваемые классы были представлены в различных цветах.

Таблица 3

Статистические характеристики обучающих кластеров из «Основного набора» по четырем каналам.

Название класса	Красный канал		Зеленый канал		Синий канал		Инфракрасный канал	
	Минимум-Максимум	Среднее±Стан.откл	Минимум-Максимум	Среднее±Стан.откл	Минимум-Максимум	Среднее±Стан.откл	Минимум-Максимум	Среднее±Стан.откл
Класс 1	241–655	363,5 ± 40,5	398–725	497,9 ± 31,4	356–553	418,0 ± 20,0	466–1132	820,5±115,8
Класс 2	264–738	497,8 ± 65,9	391–779	574,4 ± 53,2	354–605	468,6 ± 33,9	396–821	631,8 ± 52,6
Класс 3	555–826	704,7 ± 60,4	648–878	772,5 ± 56,5	511–655	586,1 ± 36,7	481–759	649,4 ± 57,4
Класс 4	216–697	380,7 ± 44,5	362–737	498,7 ± 38,0	329–553	424,7 ± 22,4	416–1394	806 ± 289,4
Класс 5	544–981	769,6 ± 95,3	653–1008	839,7 ± 77,2	493–724	619 ± 50,1	481–941	713,7±100,8
Класс 6	395–1058	743,2 ± 89,7	523–1051	799,2 ± 75,9	409–760	609,3 ± 42,8	362–1000	698,1 ± 87,8
Класс 7	554–930	729 ± 55,5	631–973	780,1 ± 50,7	493–691	583,7 ± 30,1	534–896	700,4–50,2
Класс 8	597–972	784,1 ± 63,5	694–994	835,8 ± 51,2	517–715	618,6 ± 33,2	541–928	738,3 ± 65,7
Класс 9	639–761	689,0 ± 30,6	684–781	722,4 ± 24,4	528–598	553 ± 16,1	642–783	698 ± 30,2
Класс 10	389–870	657,1±139,6	520–914	742,8 ± 115,7	424–672	561,4 ± 70,2	432–909	696,8±137,9
Класс 11	760–1104	941,4 ± 64,6	841–1141	1001,8 ± 60,2	607–800	709,8 ± 38,0	673–1001	851,5 ± 61,1
Класс 12	613–1047	854,0 ± 75,9	715–1078	912 ± 63,3	530–768	660,6 ± 40,7	523–975	780,0 ± 76,4

В таблице 4 представлена цветовая характеристика класса по типам растительности и почвенных категорий.

Таблица 4

Исходная классификация типов растительности и почвенных категорий

Номер класса	Цветовая характеристика класса
Класс 1	Красный
Класс 2	Оранжевый
Класс 3	Зеленый
Класс 4	Светло-зеленый
Класс 5	Темно-зеленый
Класс 6	Желтый
Класс 7	Светло-оранжевый
Класс 8	Синий
Класс 9	Темно-синий
Класс 10	Желтый
Класс 11	Зеленый
Класс 12	Оранжевый

На рисунке 2 рассмотрены совместные гистограммы кластеров по отдельным спектральным каналам.

Приведенный тест характеризует один из возможных вариантов модификации исходной классификационной схемы. Например, рассмотрен алгоритм классификационной схемы, в которой дополнительный набор состоит из 7-ми классов типов растительности [8].

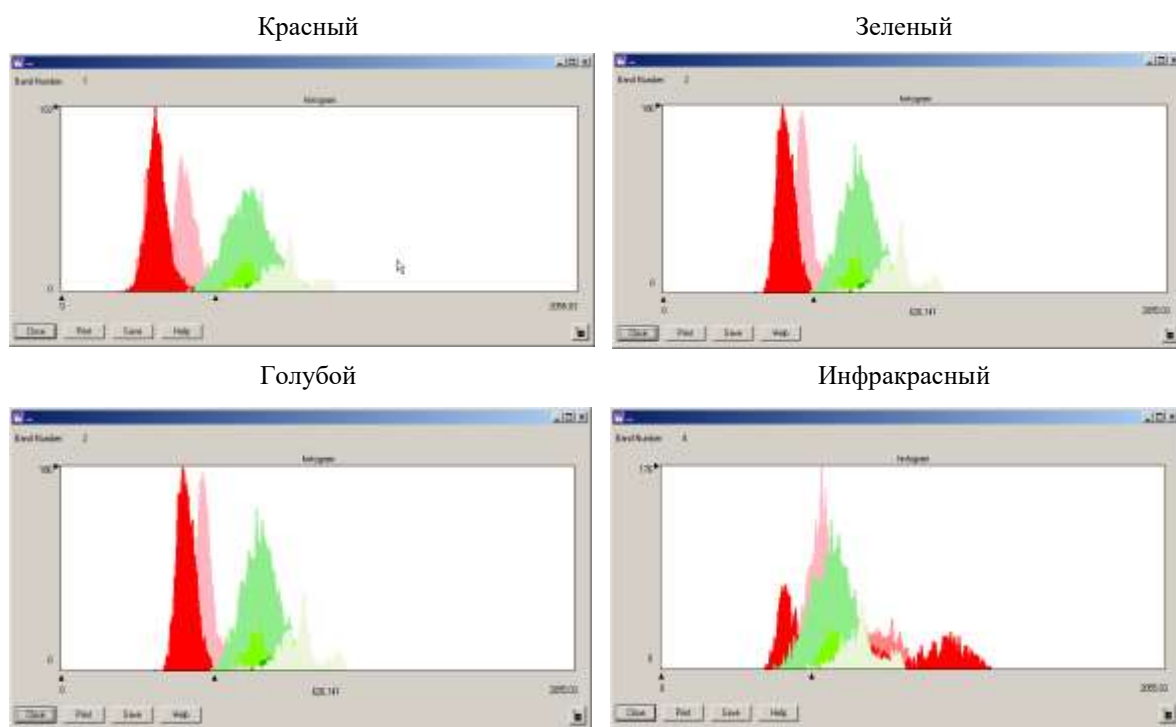


Рисунок 2. Гистограммы кластеров

Вычисление таблиц неточностей с использованием правил решения, базирующихся на статистических характеристиках.

После проведения статистических тестов и оценок, мы из исходного 12-и кластерового набора создали два дополнительных набора, состоящих, соответственно, из 7-и и 5-и кластеров (таблицы 5...6). Дополнительные наборы были составлены до начала процедуры обучения нейронного классификатора с целью определения границ возможных изменений классификационной схемы.

Таблица 5.

Таблица неточностей, полученных при применении статистического правила максимального правдоподобия для примеров из «Максимального» набора обучающих примеров - (12 классов).

Класс	1	2	3	4	5	6	7	8	9	10	11	12	Кол. прим	Кол. прав. откл. прим.
1	88,89	0,00	0,00	9,57	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1261	1080
2	0,66	82,62	0,00	9,94	0,00	0,00	0,87	0,00	0,00	0,00	0,00	0,00	2830	2628
3	0,00	7,48	79,38	0,05	8,12	2,05	5,65	1,50	0,00	0,00	0,00	0,29	428	77
4	10,45	5,16	0,00	80,43	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1812	1521
5	0,00	0,03	0,00	0,00	83,33	0,82	6,96	4,00	0,00	0,00	0,00	1,32	520	195
6	0,00	0,00	1,03	0,00	0,43	82,45	0,00	0,00	0,00	0,00	0,00	0,00	1975	1973
7	0,00	3,87	14,43	0,00	2,14	2,13	64,06	13,50	0,00	0,00	0,00	0,00	662	442
8	0,00	0,00	4,12	0,00	4,70	1,76	17,39	81,00	0,00	0,00	0,00	0,00	339	162
9	0,00	0,22	0,00	0,00	0,00	0,00	0,00	0,00	0,00	98,08	0,00	0,00	58	51
10	0,00	0,63	1,03	0,00	0,00	0,00	2,90	0,00	0,00	1,92	100,0	0,00	82	40
11	0,00	0,00	0,00	0,00	0,00	0,00	0,43	0,00	92,62	0,00	0,00	5,74	481	439
12	0,00	0,00	0,00	0,00	1,28	0,79	1,74	0,00	7,38	0,00	0,00	92,65	699	630
Итого	1215	3181	97	1891	234	2393	690	200	474	52	40	680	11147	11147

$$\text{ОУК} = 100\% * (\text{КПКП} / \text{ОКП}) = 100\% * (9238 / 11147) = \mathbf{82.87\%}$$

Примечание: ОУК – Оценка Усредненной Корректности; КПКП - Количество Правильно Классифицируемых Примеров; ОКП - Общее Количество Примеров.

Полученные результаты позволяют прогнозировать возможность качественного обучения нейронного классификатора по данной классификационной схеме, на примерах из данного набора.

Таблица 6.

Таблица неточностей, полученных в ходе обучения статистического классификатора максимального правдоподобия с помощью тестовых образцов из «Максимального набора» - (12 классов).

Класс	1	2	3	4	5	6	7	8	9	10	11	12	Кол. прим.	Кол. прав. откл. прим.
1	90,32	1,25	0,00	11,92	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1614	1343
2	5,78	90,42	0,00	10,36	0,00	0,00	0,21	0,00	0,00	0,00	0,00	0,00	2187	1887
3	0,00	0,00	73,02	0,00	8,24	1,88	2,09	4,76	0,00	0,00	0,00	0,35	142	46
4	0,90	4,55	0,00	7,71	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1750	1597
5	0,00	0,00	15,87	0,00	74,19	5,16	3,35	0,68	0,00	0,00	0,00	5,70	415	207
6	0,00	0,14	0,00	0,00	2,51	65,47	0,63	3,40	0,00	0,00	0,00	0,35	1897	1877
7	0,00	3,59	11,11	0,00	0,72	4,12	83,05	1,36	0,00	0,00	0,00	0,00	594	397
8	0,00	0,00	0,00	0,00	4,30	18,63	7,74	89,80	0,00	0,00	0,00	0,00	715	132
9	0,00	0,00	0,00	0,00	0,00	0,00	0,42	0,00	95,66	0,00	0,00	7,25	54	49
10	0,00	0,05	0,00	0,00	0,00	0,00	0,84	0,00	0,00	100,00	0,00	0,00	28	28
11	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100,00	0,00	573	529
12	0,00	0,00	0,00	0,00	10,04	4,74	1,67	0,00	4,34	0,00	0,00	86,36	696	500
Итого	1487	2087	63	2055	279	2867	478	147	553	49	28	579	10672	10665

Чтобы минимизировать риски, связанные с неопределенностью процедуры обучения, необходимо сформировать репрезентативные множества обучающих примеров. Перед началом процесса обучения нейронных классификаторов, наборы множеств обучающих примеров тестируются на основе вычисления таблиц неточностей с использованием правил решения, базирующихся на статистических характеристиках. В работе составлены таблицы неточностей вычисленные по правилу максимального правдоподобия на кластерах из максимального 12-ти классового набора примеров. Каждый набор имеет две разновидности: наборы обучающих примеров и примеры для тестирования называемые тренинговыми и тестовыми кластерами. В полном варианте статьи приводятся таблицы по двум разновидностям наборов, а также таблица неточностей классификации примеров из тестового набора.

Предварительный анализ таблиц позволяет сделать следующие выводы:

1. Результаты тестирования, при использовании того же множества примеров, что и для обучения классификатора, по обеим разновидностям кластеров, составляют порядка 80 %;

2. Результаты тестирования при использовании для тестирования тестовых кластеров, а для обучения классификатора тренинговых кластеров, составляет порядка 70 %;

3. Практическое совмещение результатов тестирования классификаторов выявило несовпадение результатов по отдельным классам, что указывает на наличие проблем процесса обучения.

Применяемый алгоритм формирует новые кластеры путём выборки примеров из одного или нескольких ранее созданных кластеров, основанных на исходной классификационной структуре. Основная сложность на начальном этапе формирования обучающей выборки заключается в неравномерном распределении примеров между кластерами. Теоретически возможно перераспределить примеры так, чтобы они были более равномерно распределены между классами.

В данной работе предложена методика, предполагающая изменение постановки задачи, связанной с определением структуры классификационной схемы. Как правило, этот подход демонстрирует более высокую эффективность при решении задач управления природными ресурсами, а также в задачах контроля и аэрокосмического мониторинга. При его применении приоритет отдаётся выбору распознаваемых индикаторов, необходимых для оценки ситуации. По результатам первичных тестов был предложен вариант объединения классов с близкими характеристиками и формирования новых классификационных схем. Это позволяет повысить качество обучения классификатора и улучшить результаты распознавания объектов.

Предлагаемый подход, объединяющий методы современных информационных технологий и модели данных, обеспечивает возможность подбора оптимального набора индикаторов и эффективного решения поставленной задачи.

После проведения статистических тестов и оценок из исходного 12-кластерного набора были сформированы два дополнительных набора, содержащие соответственно 7 и 5 кластеров. До начала обучения нейронного классификатора, с целью определения границ возможных изменений классификационной схемы, были составлены дополнительные наборы данных. Для минимизации рисков, связанных с неопределённостью процедуры обучения, необходимо формировать репрезентативные множества обучающих примеров. Перед началом обучения нейронных классификаторов эти множества проходят тестирование на основе вычисления таблиц неточностей с использованием правил решения, основанных на статистических характеристиках [9...11].

В данной работе таблицы неточностей рассчитаны по правилу максимального правдоподобия для кластеров исходного 12-классового набора примеров. Каждый набор представлен в двух вариантах: обучающие примеры (тренировочные кластеры) и примеры для тестирования (тестовые кластеры). В полном тексте статьи приведены таблицы для обеих разновидностей наборов, а также таблица неточностей классификации примеров из тестового набора. [12...13].

Процесс активного обучения классификатора Многослойный Перцептрон зависит от разных параметров, значениями которых, в зависимости от условий эксперимента, можно в разной степени варьировать.

Основными параметрами классификатора являются:

- Размер тренировочных областей;
- Количество итераций;
- Использование разного количества исходных нейронов;
- Параметры скорости обучения и импульс;
- Количество скрытых слоев и количество в них нейронов;
- Вид и тип передаточной функции.

Рассмотрим набор эвристических правил, используемых для уменьшения времени обучения и, в целом, для улучшения качества и производительности классификатора:

- при определении количества примеров необходимо иметь больше примеров для тренировочных образцов, чем весов. Можно предположить, что качество МСП, при

использовании тестового набора, будет ограничено отношением $N > W/\epsilon$, где N – количество итераций обучения, W – количество весов, ϵ – ошибка в качестве обучения классификатора минимизируется, пока среднеквадратическая ошибка не будет меньше, чем $\epsilon/2$;

- тренинговые данные необходимо нормализовать;
- использовать функцию гиперболического тангенса вместо других сигмоидов;
- нормализовать ожидаемый сигнал, выбрав его значение в интервале (0, 1);
- установить соответствующий размер шага, например, для одного скрытого слоя Многослойный Персептрон (МСП), установить размер шага на 0,05 выше в синапсе между входным и скрытым слоями и 0,01 в синапсе между скрытым и выходным слоями;
- инициализировать сетевые веса в линейной части нелинейности.

В результате проведенных экспериментов было составлено оптимальное по разным наборам множество значений для всех перечисленных параметров классификатора МСП, кроме числа итераций, значения которых варьировалось с целью достижения наилучшей производительности. При этом состав тренинговых образцов не изменялся, учитывая, что размеры образцов, которые использовались для обучения статистических классификаторов, превосходили оптимальные значения, вычисленные согласно существующим правилам. [14...15].

Таким образом, во всех экспериментах был использован следующий набор постоянных значений параметров, для обучения классификатора МСП:

1. Исходные веса, сгенерированные программой, это случайные значения в интервале [0; 0,05];
2. Количество исходных нейронов – 4, равное количеству спектральных каналов;
3. Количество скрытых слоев – 1;
4. Количество нейронов в скрытом слое – 25 для «основного набора» и 50 - для «расширенного набора». Увеличение количества примеров не приводит к увеличению количества весов, варьирование которыми позволяет обучать классификатор, следовательно, для того, чтобы сохранить порядок соотношения кол-во примеров/кол-во весов мы увеличиваем количество нейронов, а значит, и количество весов;
5. Параметры скорости обучения или размер шага между входным и скрытым слоями, η – 0,5;
6. Момент между входным и скрытым слоями, α – 0,7. Параметры скорости обучения или размер шага между скрытым входным и выходным слоями, η – 0,25;
7. Момент между входным и скрытым слоем, α – 0,7;
8. Вид и тип передаточной функции - гиперболический тангенс;
9. Порог ошибки, ϵ_{thresh} – 0,01;
10. Качество классификатора определяется, начиная с 2500-ой итерации с проверкой через каждые 2500 итераций, вплоть до достижения порога установленной ошибки - (0,001). При этом использовались следующие параметры качества:
 - MSE – среднеквадратическая ошибка. При достижении двойной кратности порога ошибки (установленное значение ошибки- 0,001), обучение останавливалось;
 - r - коэффициент корреляции между изменениями направлений реальных выходов сети, к ожидаемым. Чем ближе эта величина к 1, тем лучше классификатор обучения;
 - $\%err$ – процент погрешности, который определяет погрешность на единицу элемента, по которой вычисляются ошибки. Чем ближе значение этой величины к 0,1...0,2, тем выше качество обучения классификатора.

4. ЗАКЛЮЧЕНИЕ

В настоящей работе рассмотрены возможности применения аэрокосмических данных высокого пространственного разрешения и геоинформационных технологий для решения задач распознавания и классификации типов растительности и почвенного покрова в условиях сложной природной среды. На основе космического снимка IKONOS

и наземной информации сформирован исходный набор из 12 классов растительности и почв, отражающих основные типы природных сообществ исследуемого региона.

Проведённый анализ показал, что ключевым фактором, влияющим на качество автоматической классификации, является репрезентативность обучающих выборок и их распределение между классами. Установлено, что при неравномерном распределении примеров по классам существенно возрастает вероятность перекрытия спектральных характеристик и снижается разделяемость кластеров, что негативно отражается на результатах обучения как статистических, так и нейросетевых классификаторов.

На этапе предварительного анализа выполнена оценка статистических характеристик кластеров по всем спектральным каналам, что позволило выявить частичное совпадение и пересечение спектральных признаков отдельных классов. Результаты тестирования статистического классификатора максимального правдоподобия показали, что при использовании одних и тех же выборок для обучения и тестирования общая точность классификации составляет порядка 80–83 %, тогда как при раздельном использовании тренировочных и тестовых наборов точность снижается до уровня около 70 %. Это указывает на наличие переобучения и подтверждает необходимость оптимизации структуры классификационной схемы и состава обучающих данных.

В работе предложен методический подход, основанный на изменении постановки задачи классификации за счёт объединения спектрально близких классов и формирования альтернативных классификационных схем. В результате из исходного 12-классового набора сформированы дополнительные наборы, содержащие 7 и 5 укрупнённых классов, что позволило снизить влияние неравномерности обучающих выборок и повысить устойчивость классификаторов. Проведённые статистические тесты показали целесообразность такого подхода на этапе предварительной оценки данных до обучения нейронных моделей.

В целом результаты исследования подтверждают, что эффективное применение нейросетевых и статистических методов классификации в задачах аэрокосмического мониторинга требует не только выбора алгоритма, но и обоснованного формирования классификационной схемы и системы индикаторов. Предлагаемый подход, основанный на интеграции данных дистанционного зондирования, ГИС-технологий и методов интеллектуального анализа данных, может быть использован при решении широкого круга задач экологического мониторинга, управления природными ресурсами и оценки состояния природных территорий.

Перспективы дальнейших исследований связаны с расширением временных рядов космических данных, использованием дополнительных спектральных и текстурных признаков, а также внедрением методов активного и ансамблевого обучения для повышения устойчивости и точности классификации в условиях высокой природной неоднородности.

ВКЛАД АВТОРОВ

Концептуализация - РГН; создание программного обеспечения - ГЗР; подготовка и редактирование текста - ФРС; проведение исследования - РМР, ФРС; разработка методологии - РГН, ГЗР; проведение исследований и статистического анализа - ФРС; Проведение исследований, визуализация - ФРС; ресурсы - РМР, ФРС.

СПИСОК ЛИТЕРАТУРЫ

1. Якушев В.П., Захарян Ю.Г., Блохина С.Ю. Состояние и перспективы использования дистанционного зондирования Земли в сельском хозяйстве // Современные проблемы дистанционного зондирования Земли из космоса. - 2022. - Т. 19. - №1. - С. 287–294.
2. Нематзаде Р., Рзаева Г. Теоретические основы проектирования модели Геоинформационной системы для решения задач ДЗ по исследованию пространственно-распределённых объектов и ресурсов // Scientific Research International Scientific Journal. – 2024. – Vol. 4, Issue 10/99-105. e-ISSN: 2789-6919.
3. Груздов В.В., Колковский Ю.В., Криштопов А.В., Кудря А.И. Новые технологии дистанционного зондирования Земли из космоса. - Москва: ТЕХНОСФЕРА, 2018. – 482 с. ISBN 978-5-94836-502-2

4. Грекусис Дж. Методы и практика пространственного анализа. Описание, исследование и объяснение с использованием ГИС / пер. с англ. А.Н. Киселева. – Москва: ДМК Пресс, 2021.
5. Пузаченко М.Ю., Черненко Т.В., Беляева Н.Г. Совместный анализ наземных и дистанционных данных при оценке структуры и состава лесов на примере западной части Подмоскovie // Вестник СПбГУ. Науки о Земле. – 2020. – Т. 65. – Вып.2 – С.303-313
6. Ragimov R., Isgenderzade E., Ramazanov R., Samadov F., Jahidzada Sh. (2023). An innovative technology for aerospace monitoring of geotechnical systems based on the use of unmanned aerial vehicles (UAVs). Azerbaijan National Aerospace Agency 74th International Astronautical Congress Baku, Azerbaijan 2-6 October 2023
7. Ban H., Ahn J., Lee B. (2019). Assimilating MODIS data-derived minimum input data set and water stress factors into CERES-Maize model improves regional corn yield predictions. PLOS ONE, [e-journal] 14(2), e0211874. DOI: 10.1371/journal.pone.0211874.
8. Николаева О.В. Алгоритм обнаружения водных объектов на многоспектральных снимках // Современные проблемы дистанционного зондирования земли из космоса. – 2023. – Т. 20. – Вып. 3. – С 9-18.
9. Нематзаде Р.Г., Рзаева Г.З., Рагимов Р.М., Самедов Ф.Р. Выбор оптимального маршрута прокладки магистральных нефтегазопроводов на основе дешифрирования космических снимков высокого разрешения // Земля Беларуси. – 2025 (2). – С.53-58. https://belzeminfo.by/images/archive/2025/ZB_2025_2.pdf
10. Zhao H., Wang W., Zou X., Chen M., Pan Zh. (2025). Low-level and high-level features co-directed weakly supervised instance segmentation for optical remote sensing image interpretation. International Journal of Remote Sensing, Volume 46, Issue 13, P. 4959-4980.
11. PIMENTEL J.F.F., ARENAS R.D., SANTILLÁN Sh. M. Sh., APARICIO P.E.G., PIMENTEL D.E.F. (2022). Application Of Remote Sensing In Environmental Studies: A Theoretical Review. International Journal of Environmental, Sustainability, and Social Science. Vol. 3 No. 1.
12. Tan Y., Xu X., You H., Zhang Y., Chen M. (2025). Automated registration of forest point clouds from terrestrial and drone platforms using structural features. SPRS Journal of Photogrammetry and Remote Sensing. Volume 223. P. 28-45
13. Feigl J., Frey J., Seifert T., Koch B. (2025) Close-Range Remote Sensing of Forest Structure for Biodiversity Assessments: A Systematic Literature Review. Current Forestry Reports 11, 18. <https://doi.org/10.1007/s40725-025-00251-x>
14. Čorňák A., Delina R. (2022). Application of Remote Sensing Data in Crop Yield and Quality: Systematic Literature Review. Quality Innovation Prosperity. Vol. 26, No.3. p. 22-36. <https://doi.org/10.12776/qip.v26i3.1708>
15. Eze E., Girma A., Zenebe A., Kourouma J., Zenebe G. (2020). Exploring the possibilities of remote yield estimation using crop water requirements for area yield index insurance in a data-scarce dryland. Journal Of Arid Environments, [e-journal] 183, 104261. DOI: 10.1016/j.jaridenv.2020.104261.

REFERENCES

1. Yakushev V. P., Zakharyan Y. G., Blokhina S. Y. (2022). Sostoyaniye i perspektivy ispol'zovaniya distancionnogo zondirovaniya Zemli v sel'skom hozyajstve [Current state and prospects of using Earth remote sensing in Agriculture]. Current Problems in Remote Sensing of the Earth from Space, 19(1), 287–294. [in Russian]
2. Nematzade R., Rzaeva G. (2024). Teoreticheskie osnovy proektirovaniya modeli Geoinformacionnoy sistemy dlya resheniya zadach DZ po issledovaniyu prostranstvenno-raspredeennykh ob'ektov i resursov [Theoretical foundations for designing a geoinformation system model for solving remote sensing problems of spatially distributed objects and resources]. Scientific Research International Scientific Journal, 4(10), 99–105. e-ISSN 2789-6919. . [in Russian]
3. Gruzov V. V., Kolkovsky Y. V., Krishtopov A. V., Kudrya A. I. (2018). Novye tekhnologii distancionnogo zondirovaniya Zemli iz kosmosa [New technologies of Earth remote sensing from Space]. Moscow, Russia: Technosphere. ISBN 978-5-94836-502-2. [in Russian]
4. Grekousis G. (2021). Metody i praktika prostranstvennogo analiza. Opisanie, issledovanie i ob'yasnenie s ispol'zovaniem GIS [Methods and practice of spatial analysis: Description, exploration, and explanation using GIS] (A. Kiseleva, Trans.). Moscow, Russia: DMK Press. (Original work published in English)
5. Puzachenko M. Y., Chernenkova T. V., Belyaeva N. G. (2020). Sovmestnyy analiz nazemnykh i distancionnykh dannykh pri ocenke struktury i sostava lesov na primere zapadnoy chasti Podmoskov'ya [Integrated analysis of ground-based and remote sensing data in assessing forest structure and composition: A case study of western Moscow Region]. Vestnik of Saint Petersburg University. Earth Sciences, 65(2), 303–313. . [in Russian]
6. Ragimov R., Isgenderzade E., Ramazanov R., Samadov F., Jahidzada S. (2023). An innovative technology for aerospace monitoring of geotechnical systems based on the use of unmanned aerial vehicles (UAVs). In Proceedings of the 74th International Astronautical Congress (October 2–6, 2023). Baku, Azerbaijan.
7. Ban H., Ahn J., Lee B. (2019). Assimilating MODIS data-derived minimum input data set and water stress factors into the CERES-Maize model improves regional corn yield predictions. PLOS ONE, 14(2), Article e0211874. <https://doi.org/10.1371/journal.pone.0211874>
8. Nikolaeva O. V. (2023). Algoritmy obnaruzheniya vodnykh ob'ektov na mnogosppektral'nykh snimkakh [Algorithm for detecting water bodies in multispectral images]. Current Problems in Remote Sensing of the Earth from Space, 20(3), 9–18. . [in Russian]
9. Nematzade R. G., Rzaeva G. Z., Ragimov R. M., Samedov F. R. (2025). Vybory optimal'nogo marshruta prokladki magistral'nykh neftegazoprovodov na osnove deshifirovaniya kosmicheskikh snimkov vysokogo razresheniya [Selection of the optimal route for main oil and gas pipelines based on interpretation of high-resolution satellite imagery]. Land of Belarus, 2, 53–58. https://belzeminfo.by/images/archive/2025/ZB_2025_2.pdf . [in Russian]
10. Zhao H., Wang W., Zou X., Chen M., Pan Zh. (2025). Low-level and high-level features co-directed weakly supervised instance segmentation for optical remote sensing image interpretation. International Journal of Remote Sensing, Volume 46, Issue 13, P. 4959-4980.
11. PIMENTEL J.F.F., ARENAS R.D., SANTILLÁN Sh. M. Sh., APARICIO P.E.G., PIMENTEL D.E.F. (2022). Application Of Remote Sensing In Environmental Studies: A Theoretical Review. International Journal of Environmental, Sustainability, and Social Science. Vol. 3 No. 1.
12. Tan Y., Xu X., You H., Zhang Y., Chen M. (2025). Automated registration of forest point clouds from terrestrial and drone platforms using structural features. SPRS Journal of Photogrammetry and Remote Sensing. Volume 223. P. 28-45

13. Feigl J., Frey J., Seifert T., Koch B. (2025) Close-Range Remote Sensing of Forest Structure for Biodiversity Assessments: A Systematic Literature Review. *Current Forestry Reports* 11, 18. <https://doi.org/10.1007/s40725-025-00251-x>
14. Čornák A., Delina R. (2022). Application of Remote Sensing Data in Crop Yield and Quality: Systematic Literature Review. *Quality Innovation Prosperity*. Vol. 26, No.3. p. 22-36. <https://doi.org/10.12776/qip.v26i3.1708>
15. Eze E., Girma A., Zenebe A., Kourouma J., Zenebe G. (2020). Exploring the possibilities of remote yield estimation using crop water requirements for area yield index insurance in a data-scarce dryland. *Journal Of Arid Environments*, [e-journal] 183, 104261. DOI: 10.1016/j.jaridenv.2020.104261

ЖОҒАРЫ АЖЫРАТЫМДЫЛЫҚТАҒЫ КӨП СПЕКТРЛІ ҒАРЫШТЫҚ БЕЙНЕЛЕРДЕ ОЛАРДЫ ЖІКТЕУ ЖӘНЕ ТАҢУ МАҚСАТЫНДА КЕҢІСТІКТЕ БӨЛІНГЕН ОБЪЕКТИЛЕР КЛАСТАРЫН ҚАЛЫПТАСТЫРУ ӘДІСІ

Рашад. Г. Нематзаде, Гюнель. З. Рзаева, Рауф. М. Рагимов PhD, Фарид. Р. Самедов

Аэроғарыштық Информатика Ғылыми-Зерттеу Институты, Баку, Әзірбайжан; Rashad.Nematzade@nasa.az, Gunel.Rzayeva@nasa.az, rauffmahmud@yahoo.com, farid.samedov@hotmail.com

*Автор корреспонденция: Фарид Р. Самедов, farid.samedov@hotmail.com

ТҮЙІН СӨЗДЕР

жіктеу,
нысан,
әдістеме,
ғарыштық сурет,
көріністер,
өсімдік жамылғысы

Мақала жайында:

Жіберілді: 03.07.2025
Қайта қаралды: 22.12.2025
Қабылданды: 22.12.2025
Жарияланды: 30.12.2025

АБСТРАКТ

Жұмыста кеңістіктік ажыратымдылығы жоғары аэроғарыштық деректер мен геоақпараттық технологиялар негізінде өсімдік жамылғысы мен жер жамылғысының түрлерін автоматтандырылған жіктеу әдістері қарастырылады. Зерттеу IKONOS ғарыштық суретін қолдану арқылы жүзеге асырылады және зерттелетін аймаққа тән өсімдіктер мен Топырақтардың 12 класы үшін оқыту және сынақ үлгілерін қалыптастыруды қамтиды. Кластерлердің спектрлік сипаттамаларына статистикалық талдау және олардың өкілдігі мен бөлінуін бағалау жүргізілді. Оқыту мысалдарының біркелкі бөлінбеуі және жеке сыныптардың спектрлік белгілерінің қабаттасуы классификацияның тұрақтылығын төмендететіні көрсетілген. Таңу сапасын арттыру үшін спектрлік жақын сыныптарды біріктіруге және балама жіктеу схемаларын қалыптастыруға негізделген тәсіл ұсынылады. Әдістің тиімділігі максималды ықтималдық статистикалық жіктеуші мен көп қабатты перцептрон типіндегі нейрондық желілік жіктеушінің көмегімен бағаланады. Алынған нәтижелер аэроғарыштық экологиялық мониторинг міндеттерін шешу кезінде жіктеу схемасының құрылымын және оқыту үлгілерінің құрамын оңтайландырудың орындылығын растайды.

METHOD OF FORMING CLASSES OF SPATIALLY DISTRIBUTED OBJECTS FOR THE PURPOSE OF THEIR CLASSIFICATION AND RECOGNITION ON HIGH-RESOLUTION MULTISPECTRAL SPACE IMAGES

Rashad Nematzade, Gunel Rzayeva, Rauf Ragimov PhD, Farid Samedov

Research Institute of Aerospace Informatics, Baku, Azerbaijan; Rashad.Nematzade@nasa.az, Gunel.Rzayeva@nasa.az, rauffmahmud@yahoo.com, farid.samedov@hotmail.com

* Corresponding author: Farid Samedov, farid.samedov@hotmail.com

KEY WORDS

classification,
object,
methodology,
satellite image,
types,
vegetation

ABSTRACT

The paper discusses methods for automated classification of vegetation types and soil cover based on high-resolution aerospace data and geoinformation technologies. The study was conducted using IKONOS satellite imagery and includes the formation of training and test samples for 12 classes of vegetation and soils characteristic of the region under study. A statistical analysis of the spectral characteristics of clusters and an assessment of their representativeness and separability were performed. It was shown that the uneven distribution of training examples and the overlap of spectral features of individual classes reduce the stability of classification. To improve recognition quality, an approach based on combining spectrally similar classes and forming alternative classification schemes was

About article:

Received: 03.07.2025

Revised: 22.12.2025

Accepted: 22.12.2025

Published: 30.12.2025

proposed. The effectiveness of the method is evaluated using a maximum likelihood statistical classifier and a multilayer perceptron neural network classifier. The results confirm the feasibility of optimizing the structure of the classification scheme and the composition of training samples when solving aerospace environmental monitoring problems.

Примечание издателя: заявления, мнения и данные во всех публикациях принадлежат только автору (авторам), а не журналу "Гидрометеорология и экология" и/или редактору (редакторам).