

Жарикова Е. П., Григорьев Я. Ю., Григорьева А. Л.
E. P. Zharikova, Ya. Yu. Grigoriev, A. L. Grigorieva

АНАЛИЗ ИЗМЕНЕНИЯ СПЕКТРАЛЬНЫХ ХАРАКТЕРИСТИК ВОДНОЙ ПОВЕРХНОСТИ РЕКИ АМУР

ANALYSIS OF CHANGES IN SPECTRAL CHARACTERISTICS OF THE AMUR RIVER WATER SURFACE

Жарикова Евгения Павловна – аспирант Комсомольского-на-Амуре государственного университета (Россия, Комсомольск-на-Амуре).

Evgeniia P. Zharikova – Postgraduate Student, Komsomolsk-na-Amure State University (Russia, Komsomolsk-on-Amur).

Григорьев Ян Юрьевич – кандидат физико-математических наук, доцент, проректор по учебной работе Комсомольского-на-Амуре государственного университета (Россия, Комсомольск-на-Амуре).

Yan Yu. Grigoriev – PhD in Physics and Mathematics, Associate Professor, Vice Rector for Academic Affairs, Komsomolsk-na-Amure State University (Russia, Komsomolsk-on-Amur).

Григорьева Анна Леонидовна – кандидат физико-математических наук, доцент, заведующая кафедрой «Прикладная математика» Комсомольского-на-Амуре государственного университета (Россия, Комсомольск-на-Амуре).

Anna L. Grigorieva – PhD in Physics and Mathematics, Associate Professor, Head of Applied Mathematic Department, Komsomolsk-na-Amure State University (Russia, Komsomolsk-on-Amur).

Аннотация. Важный вопрос экологии связан с загрязнением водоёмов. Своевременная оценка возможных изменений водной среды, связанных с воздействием негативных факторов, не всегда возможна и требует дополнительных исследований качества воды. Для своевременного реагирования на нештатное изменение состава воды предлагается применение мультиспектрального анализа водной поверхности методами машинного обучения. Химический состав водных объектов зависит от большого числа факторов, а загрязнения имеют различную природу, поэтому в работе ограничивается область анализа: рассматривается поверхность реки Амур за период с 2020 по 2022 годы.

Summary. Pollution of water bodies is an important environmental problem. The cause of environmental abnormal situations is not always obvious, requiring additional water quality investigations. In order to respond to abnormal changes in water body composition in a timely manner, the authors propose using a multispectral water surface analysis approach using machine learning methods. Since the chemical compositions of water bodies vary as well as their pollution, it is necessary to limit the scope of research. This paper reflects the results of work on the surface analysis of the Amur River for the period from 2020 to 2022.

Ключевые слова: кластеризация, понижение размерности, искусственный интеллект, нейронные сети, спектральный анализ, машинное зрение, машинное обучение.

Key words: clustering, dimensionality reduction, artificial intelligence, neural networks, spectral analysis, machine vision, machine learning.

Исследование выполнено за счёт средств гранта Министерства образования и науки Хабаровского края № 41С/2023.

УДК 004.932.2

Проведённые исследования [2–8] показывают возможность детектирования загрязнений водных объектов при помощи мультиспектральных снимков. Гипотеза состоит в том, что на основе полученных спутниковых изображений можно провести анализ динамики состояния водной поверхности и определить время изменения химического состава вод реки Амур.

В работах [4–6] решается задача классификации загрязнений, в рассматриваемом исследовании не берутся в расчёт знания о возможных загрязнениях. Рассматривается задача анализа нескольких областей в течение продолжительного периода времени и выявления на основе алгоритмов кластеризации изменений водной поверхности реки Амур.

На основе исследований загрязнений водной поверхности [7] и спектральных каналов, применяемых при классическом анализе воды [10], определяются требования к набору данных, описанные в табл. 1. Дополнительным требованием является время получения снимка, позволяющее оценить динамику изменений.

Таблица 1

Требования к спектральным каналам в обозначениях аппарата Sentinel-2LA

Название	Центральная длина волны, нм
1. Coastal aerosol	442.7
2. Blue	492.4
3. Green	559.8
4. Red	664.6
5. Vegetation red edge	704.1
6. Vegetation red edge	740.5
7. Vegetation red edge	782.8
8. NIR	832.8
9. Narrow NIR	864.7
10. SWIR – Cirrus	1373.5
11. SWIR	1613.7
12. SWIR	2202.4

Для каждой области обрабатывается четыре трёхканальных изображения, включающих все вышеописанные каналы.

На первом этапе производится разметка данных: определяются области, подлежащие анализу, содержащие только воду (см. рис. 1).



Рис. 1. Пример разметки изображения

На втором этапе формируются векторы признаков на основе изображений. Отдельный вектор представляет собой набор из 12 значений спектральных каналов для каждого выделенного пикселя.

На третьем этапе определяются матрицы корреляции спектральных каналов (см. рис. 2).

Полученные коэффициенты корреляции Пирсона свидетельствуют о высокой степени корреляции каналов ближнего инфракрасного спектра, что указывает на необходимость уменьшения количества спектральных каналов для построения конечной модели.

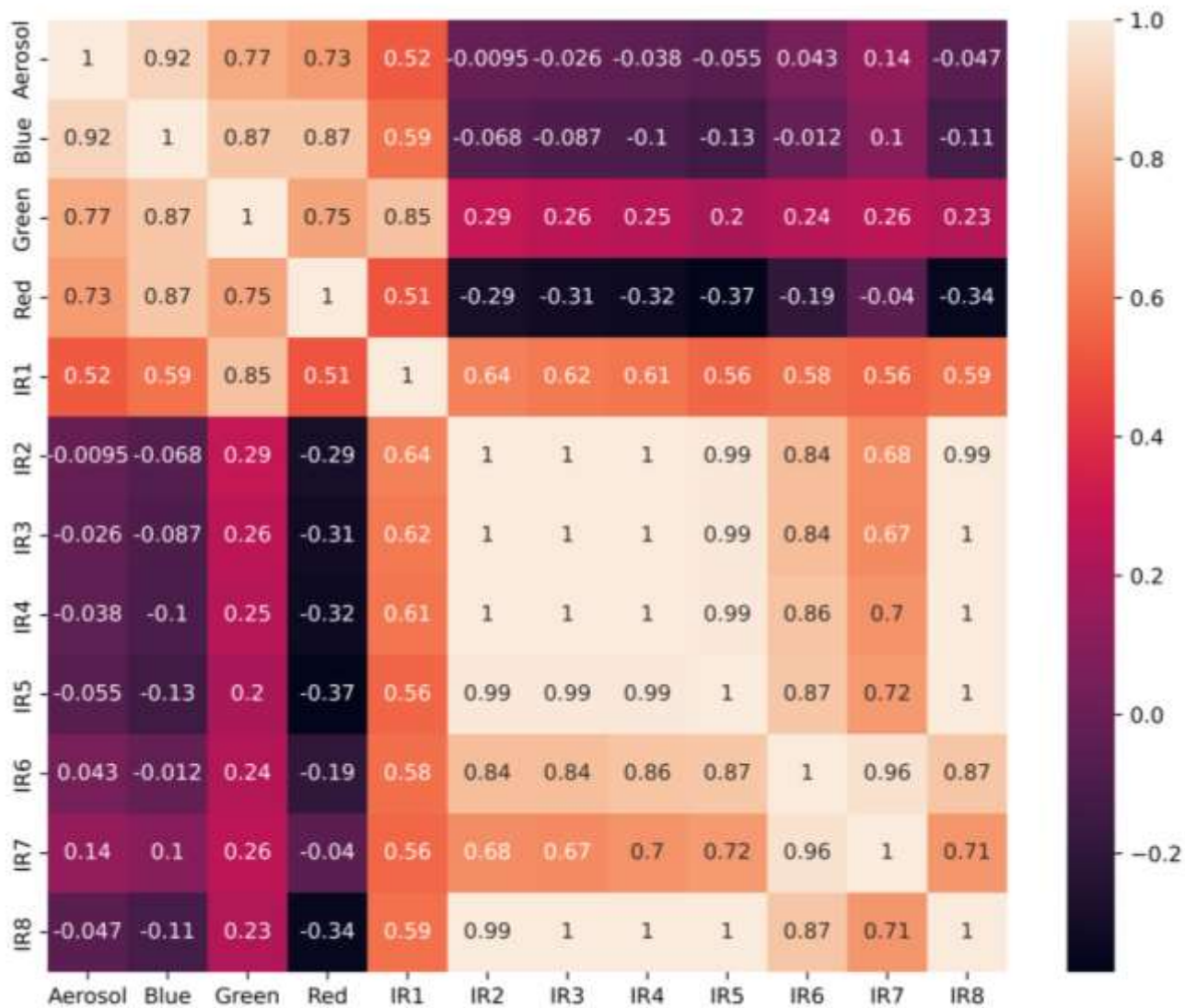


Рис. 2. Матрица корреляций спектральных каналов

Графики распределения значений рассматриваемых каналов, отражённые на рис. 3, имеют схожий вид для значений каналов с высокой корреляцией.

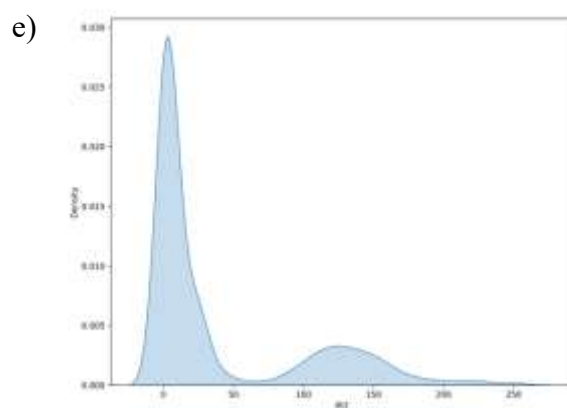
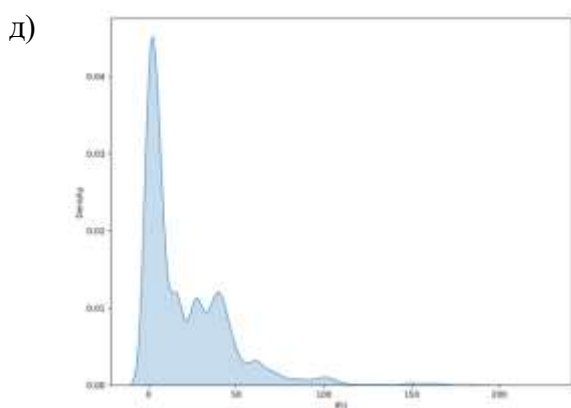
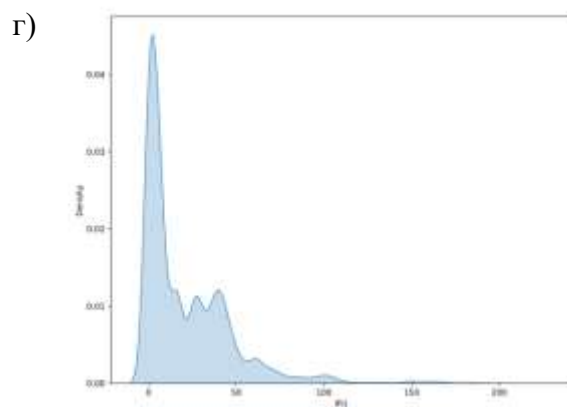
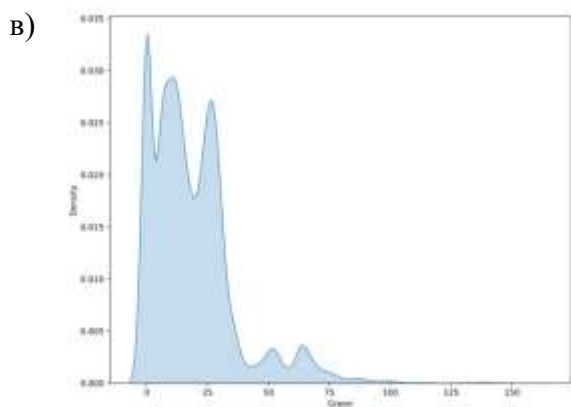
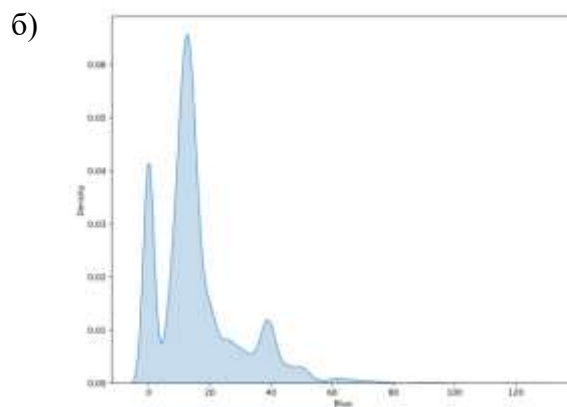
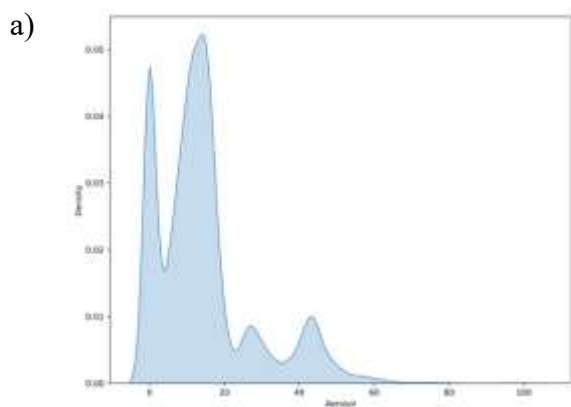


Рис. 3. Графики распределения значений каждого спектрального канала (начало):
 а – Coastal aerosol – 442.7 нм; б – Blue – 492.4 нм; в – Green – 559.8 нм; г – Red – 664.6 нм;
 д – Vegetation red edge – 704.1 нм; е – Vegetation red edge – 740.5 нм; ж – Vegetation red
 edge – 782.8 нм; з – NIR – 832.8 нм; и – Narrow NIR – 864.7 нм; к – SWIR – Cirrus – 1373.5 нм;
 л – SWIR – 1613.7 нм; м – SWIR – 2202.4 нм

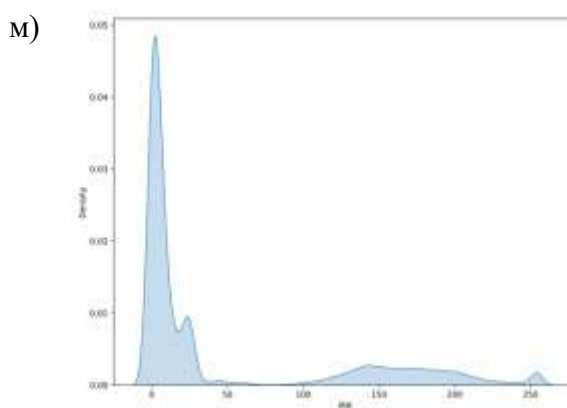
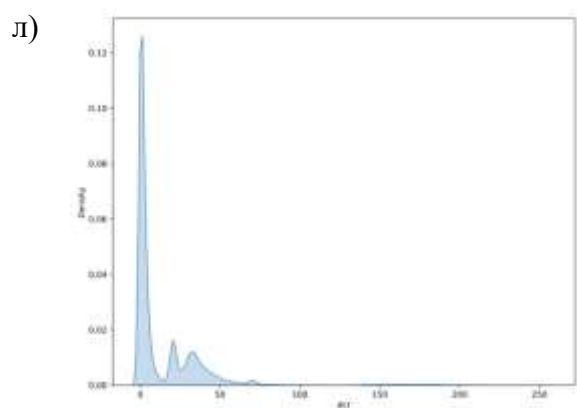
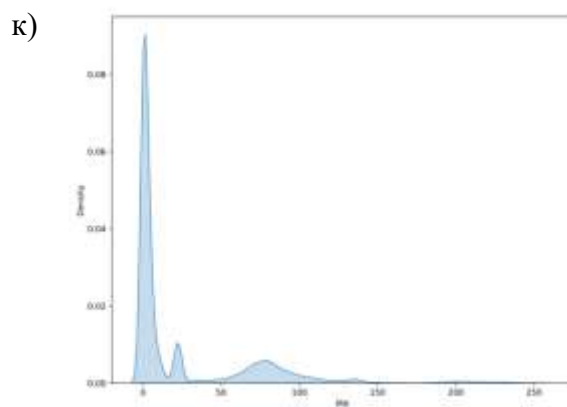
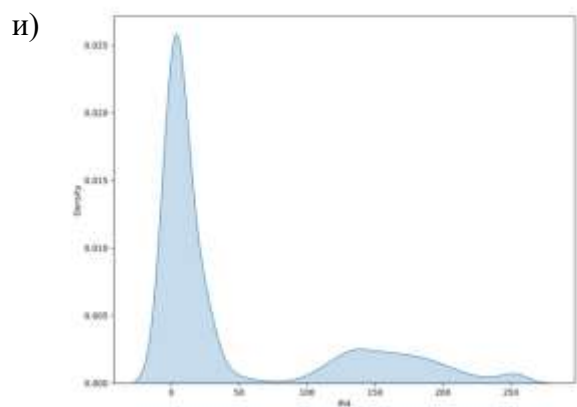
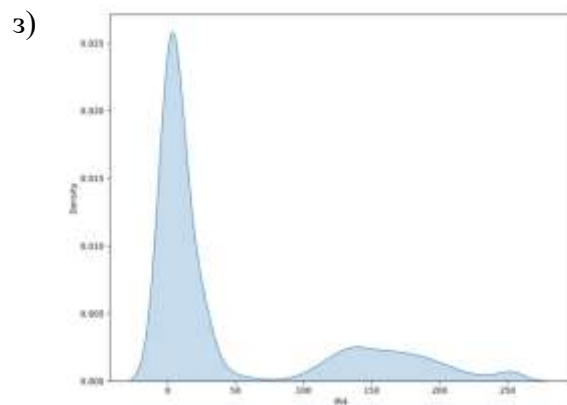
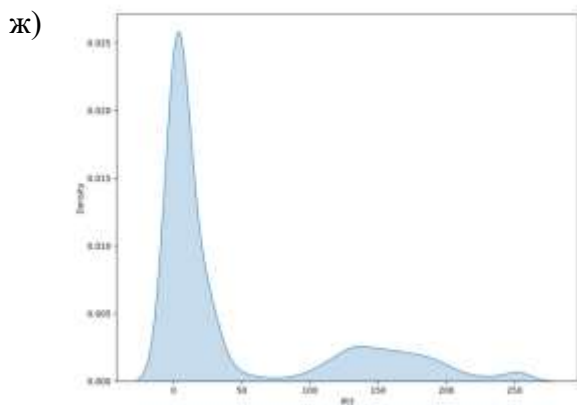


Рис. 3. Графики распределения значений каждого спектрального канала (продолжение):
 а – Coastal aerosol – 442.7 нм; б – Blue – 492.4 нм; в – Green – 559.8 нм; г – Red – 664.6 нм;
 д – Vegetation red edge – 704.1 нм; е – Vegetation red edge – 740.5 нм; ж – Vegetation red
 edge – 782.8 нм; з – NIR – 832.8 нм; и – Narrow NIR – 864.7 нм; к – SWIR – Cirrus – 1373.5 нм;
 л – SWIR – 1613.7 нм; м – SWIR – 2202.4 нм

Несмотря на высокую корреляцию между спектральными каналами, первая кластеризация проводится со всеми спектральными каналами. Результат кластеризации для количества кластеров, равного 5 (предполагаемое количество кластеров), отражён на рис. 4.

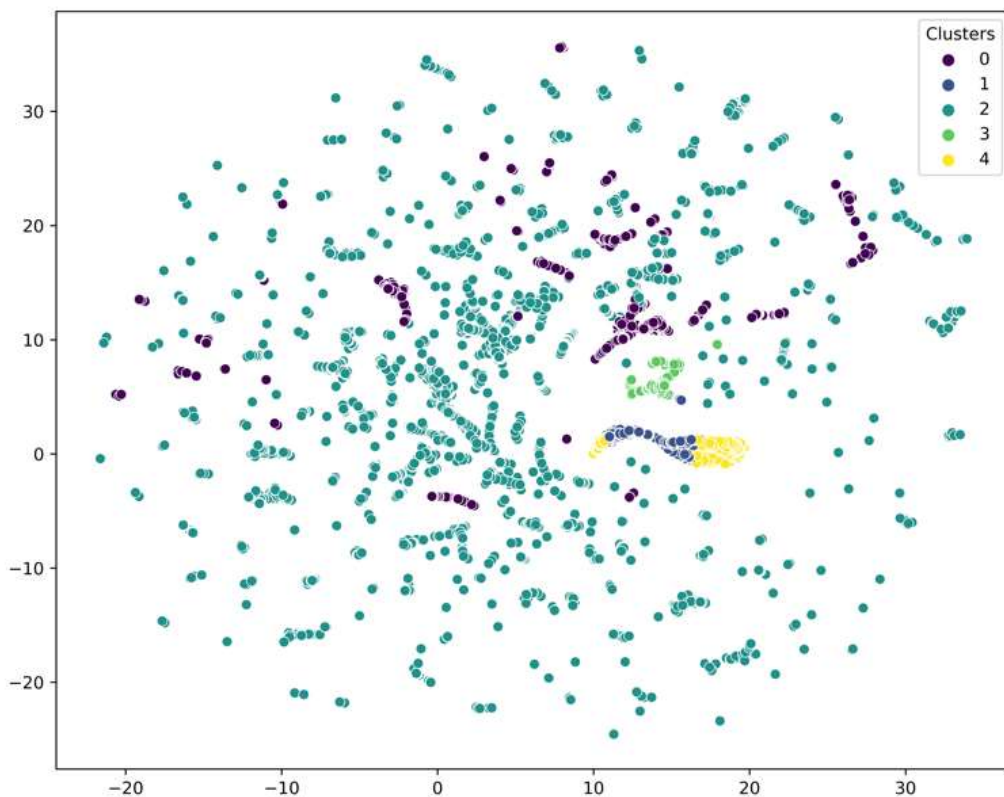


Рис. 4. Результат кластеризации для всего набора данных

Для визуализации и анализа полученных данных используется алгоритм нелинейного снижения размерности Uniform Manifold Approximation and Projection (UMAP). На изображении определяются общая масса и явно выделяющиеся кластеры, но требуется проверка достоверности полученных результатов. Обученная модель применяется к спутниковым снимкам двух областей реки Амур для визуальной оценки (на рис. 5, *a* – первая область; на рис. 5, *б* – вторая область).

a)



Рис. 5. Результаты обработки изображения обученным алгоритмом (начало):
а – первая исследуемая область; б – вторая исследуемая область

б)



Рис. 5. Результаты обработки изображения обученным алгоритмом (продолжение):
а – первая исследуемая область; б – вторая исследуемая область

Для обучения алгоритма кластеризации выбираются спектральные диапазоны: Aerosol, Green, Red, R1, R2, R7. Результаты кластеризации представлены на рис. 6.

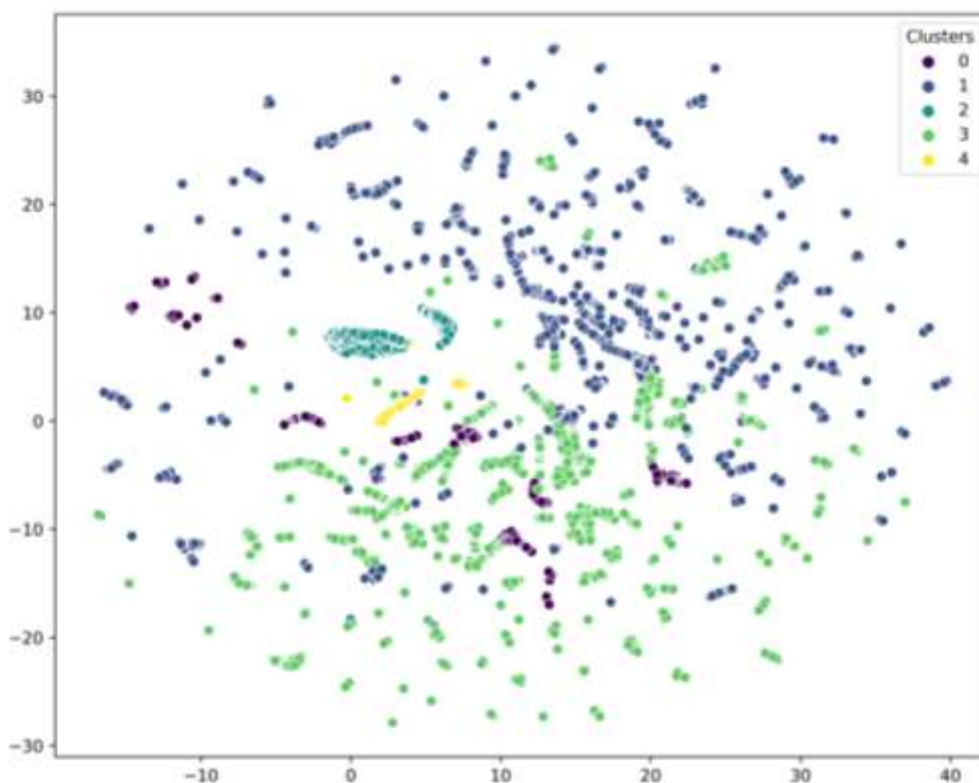


Рис. 6. Результат кластеризации для каналов Aerosol, Green, Red, R1, R2, R7

Новое разбиение набора данных на кластеры отличается от предыдущего результата, но также содержит обособленные кластеры, что может указывать на их схожую природу. Результаты работы обученного алгоритма с явным разделением на кластеры представлены: на рис. 7, *a* – для первой исследуемой области, на рис. 7, *б* – для второй исследуемой области.

Кластеризация с меньшим количеством спектральных каналов показывает более однородное разделение.

Анализ полученных результатов для первой области показывает, что изображения в видимом спектре значительно отличаются по цвету. Несмотря на визуальное изменение цвета поверхности акватории, рассматривается предположение о незначительных изменениях состава воды реки Амур в данной области, изменение цвета в видимом диапазоне связано с астрономическим временем съёмки и погодными условиями.

Рассмотрение работы алгоритма для другого участка реки указывает на явные различия в спектральных характеристиках. Изменения поверхности воды на рис. 7, *б* (нижнее изображение) по своей форме напоминают течения, что свидетельствует о некоторых закономерностях, выявленных алгоритмом. Наиболее часто встречающийся кластер на рис. 7, *б* (нижнее изображение) также определён на рис. 7, *б* (среднее изображение) и может свидетельствовать об илистой природе речного дна. Ещё один определённый кластер на рис. 7, *б* (нижнее изображение) также встречается и на суше, где есть растительность (см. рис. 7, *a* (верхнее изображение)). На рис. 7, *б* (верхнее изображение) также определяется данный кластер, где присутствует зелёный цвет, что может свидетельствовать о том, что данный кластер можно интерпретировать как растительность. Очередной кластер встречается только на водной поверхности (см. рис. 7, *a* (верхнее и среднее изображения), рис. 7, *б* (верхнее и нижнее изображения)), однако в рамках данного исследования сложно определить, какими свойствами обладает выделенная водная поверхность.

a)



Рис. 7. Результаты обработки изображения обученным алгоритмом с явным разделением на кластеры (начало):
а – первая исследуемая область; б – вторая исследуемая область

б)

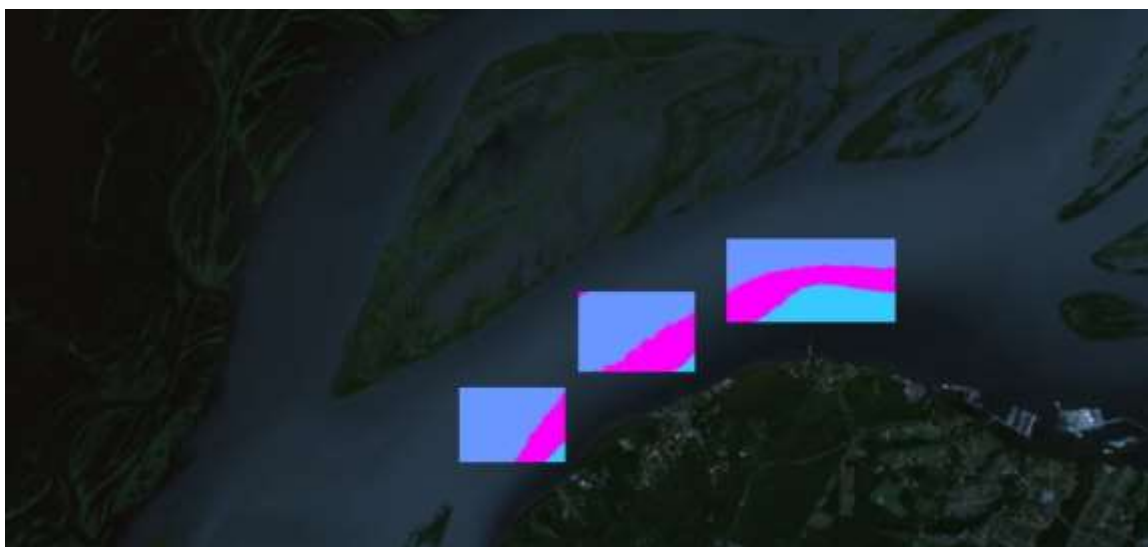


Рис. 7. Результаты обработки изображения обученным алгоритмом с явным разделением на кластеры (продолжение):
а – первая исследуемая область; б – вторая исследуемая область

Результаты исследования не опровергают гипотезу о возможности определения изменений отдельных областей водной поверхности без начальных представлений об их возможной природе. Сопоставление полученных данных с информацией о нештатных ситуациях на указанных территориях акваторий может способствовать предположению о нормальном составе воды либо отклонении от него. Дальнейшее исследование состоит в построении классификаторов для данных областей, размеченных на основе текущих результатов работы алгоритмов кластеризации. Полученные результаты должны позволять масштабировать исследования за счёт большего охвата площадей водных поверхностей реки Амур и выявлять закономерности спектральных характеристик составов водной поверхности для более глубокого анализа.

ЛИТЕРАТУРА

1. Gitelson A. The peak near 700 nm on radiance spectra of algae and water: Relationships of its magnitude and position with chlorophyll concentration. *Int. J. Remote Sens.*, 1992., 13, p. 3367-3373.
2. Жарикова, Е. П. Модели расчёта и фильтрации вегетационных индексов / Е. П. Жарикова, Я. Ю. Григорьев // Научно-техническое творчество аспирантов и студентов: сборник материалов 4-й науч.-техн. конф. студентов и аспирантов. – Комсомольск-на-Амуре: ФГБОУ ВО «КнАГУ», 2018. – С. 162-164.
3. Жарикова, Е. П. Сравнение вегетационных индексов в задачах оценки поверхностей / Е. П. Жарикова, Я. Ю. Григорьев // Научно-техническое творчество аспирантов и студентов: сборник материалов 4-й науч.-техн. конф. студентов и аспирантов. – Комсомольск-на-Амуре: ФГБОУ ВО «КнАГУ», 2018. – С. 165-167.
4. Модели компьютерного зрения в задачах дистанционного зондирования / Е. П. Жарикова, И. А. Трещев, Я. Ю. Григорьев, А. Л. Григорьева // Постулат. Приамурский государственный университет им. Шолом-Алейхема. – 2018. – № 4 (30). – С. 73.
5. Zharikova, E. P. Artificial intelligence methods for detecting water pollution / E. P. Zharikova, J. Y. Grigoriev, A. L. Grigorieva // *IOP Conference Series: Earth and Environmental Science*. Ser. «International Science and Technology Conference «Earth Science», ISTC EarthScience 2022 – Chapter 1». 2022. – P. 022082.
6. Формирование набора данных для решения задач детектирования нештатных ситуаций в экологии / И. Н. Альхименко, Е. П. Жарикова, М. А. Чепурных, А. И. Базюк, Я. Ю. Григорьев // Молодёжь и наука: актуальные проблемы фундаментальных и прикладных исследований: сборник материалов 5-й Всероссийской национальной научной конференции молодых учёных. – Комсомольск-на-Амуре: ФГБОУ ВО «КнАГУ», 2022. – С. 360-362.
7. Жарикова, Е. П. Применение искусственного интеллекта в задачах анализа состояния акваторий / Е. П. Жарикова, Я. Ю. Григорьев, А. Л. Григорьева // *Морские интеллектуальные технологии*. – 2021. – № 2-2 (52). – С. 129-133.
8. Applications of computer vision in cross-sectoral tasks / Zharikova E. P., Grigoriev Y. Y., Grigorieva A. L. *Lecture Notes in Networks and Systems*. 2021. T. 200. P. 415-426.
9. Surface state detection / Zharikova E. P., Grigoriev Ya. Yu., Grigoryeva A. L. // 2019 International Multi-Conference on Industrial Engineering and Modern Technologies, FarEastCon 2019. 2019. P. 8934205.
10. Титоренко, А. В. Разработка системы автоматизированного анализа состояния водных объектов с применением технологий дистанционного зондирования / А. В. Титоренко, Я. Ю. Григорьев // *Материалы II Всероссийской национальной научной конференции студентов, аспирантов и молодых учёных*. – Комсомольск-на-Амуре: ФГБОУ ВО «КнАГУ», 2019. – С. 480.
11. Амосов, О. С. Быстродействующие численные нейросетевые и нечёткие методы стохастического оценивания состояния динамических систем / О. С. Амосов, С. Г. Баена // *Информатика и системы управления*. – Комсомольск-на-Амуре: ФГБОУ ВО «КнАГУ», 2014. – С. 118-129.