

doi: 10.17586/2226-1494-2022-22-1-82-92

УДК 004.932

Классификация объектов на изображениях с учетом искажений на основе двухэтапного топологического анализа

Сергей Владимирович Еремеев^{1✉}, Артём Владимирович Абакумов²

^{1,2} Муромский институт (филиал) ФГБОУ ВО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых», Муром, 602252, Российская Федерация

¹ sv-eremeev@yandex.ru✉, <https://orcid.org/0000-0001-8482-1479>

² artem210966@yandex.ru, <https://orcid.org/0000-0001-5784-7147>

Аннотация

Предмет исследования. Предложен метод автоматической классификации пространственных объектов на изображениях в условиях ограниченного набора данных. Исследована устойчивость метода к искажениям, которые часто возникают на изображениях в силу природных явлений и частичном перекрытии объектов городской инфраструктуры. Для получения высокой точности классификации с помощью существующих подходов требуется большая обучающая выборка, включая наборы данных с искажениями, что значительно увеличивает вычислительную сложность. **Метод.** Предложен метод двухэтапного топологического анализа изображений. Топологические признаки первоначально извлекаются при анализе изображения в диапазоне яркостей от 0 до 255, а затем от 255 до 0. Эти признаки дополняют друг друга и отражают топологическую структуру объекта. При определенных деформациях и искажениях структура объекта в виде извлеченных признаков сохраняется. Преимущество метода — небольшое количество эталонов, что снижает вычислительную нагрузку при обучении по сравнению с нейросетевым подходом. **Основные результаты.** Выполнено исследование и сравнение предложенного метода с современным нейросетевым подходом. Исследование проведено на наборе данных DOTA (Dataset for Object deTection in Aerial images), содержащем снимки пространственных объектов нескольких классов. При отсутствии искажений на изображении нейросетевой подход показал точность классификации свыше 98 %, а при использовании предложенного метода около 82 %. Также в работе применены искажения изображений: поворот на 90°, сужение и обрезание края на 50 %, а также их комбинации. Предложенный метод показал свою устойчивость и превзошел нейросетевой подход. В самой сложной комбинации теста снижение точности классификации нейросетевого подхода составило 46 %, в то время как для предложенного метода только 12 %. **Практическая значимость.** Предложенный метод целесообразно использовать при большой вероятности появления искажений на изображениях. Такие искажения возникают в сфере геоинформатики при анализе объектов разных масштабов, при различных погодных условиях, частичном перекрытии одного объекта другим, при наличии тени и других факторов. Возможно использование предложенного метода в системах технического зрения промышленных предприятий при автоматической классификации типа деталей наложенных объектов.

Ключевые слова

топологический анализ, персистентная гомология, искажение изображений, классификация объектов, нейронные сети

Благодарности

Исследование выполнено в рамках Программы развития ЯрГУ, проект № П2-К-1-Г-3/2021.

Ссылка для цитирования: Еремеев С.В., Абакумов А.В. Классификация объектов на изображениях с учетом искажений на основе двухэтапного топологического анализа // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 1. С. 82–92. doi: 10.17586/2226-1494-2022-22-1-82-92

Classification of objects in images with distortions based on a two-stage topological analysis

Sergey V. Ereemeev¹, Artyom V. Abakumov²

^{1,2} Murom Institute (Branch) of Vladimir State University, Murom, 602252, Russian Federation

¹ sv-eremeev@yandex.ru, <https://orcid.org/0000-0001-8482-1479>

² artem210966@yandex.ru, <https://orcid.org/0000-0001-5784-7147>

Abstract

The authors propose a method for automatic classification of spatial objects in images under conditions of a limited data set. The stability of the method to distortions appearing in images due to natural phenomena and partial overlap of urban infrastructure objects is investigated. High classification accuracy, when using existing approaches, requires a large training sample, including data sets with distortions, which significantly increases computational complexity. The paper proposes a method for a two-step topological analysis of images. Topological features are initially extracted by analyzing the image in the brightness range from 0 to 255, and then from 255 to 0. These features complement each other and reflect the topological structure of the object. Under certain deformations and distortions, the object preserves its structure in the form of extracted features. The advantage of the method is a small number of patterns, which reduces the computational complexity of training compared to neural networks. The proposed method is investigated and compared with the modern neural network approach. The study was performed on a DOTA dataset (Dataset for Object deTectiOn in Aerial images) containing images of spatial objects of several classes. In the absence of distortion in the image, the neural network approach showed a classification accuracy of over 98 %, while the proposed method achieved about 82 %. Further distortions such as 90 degree rotation, 50 % narrowing and 50 % edge truncation and their combinations were applied in the experiment. The proposed method showed its robustness and outperformed the neural network approach. In the most difficult combination of the test, the decrease in classification accuracy of the neural network was 46 %, while the described method showed 12 %. The proposed method can be applied in cases with a high probability of distortion in the images. Such distortions arise in the field of geoinformatics when analyzing objects of various scales, under different weather conditions, partial overlap of one object with another, in the presence of shadows, etc. It is possible to use the proposed method in vision systems of industrial enterprises for automatic classification of the parts that belong to superimposed objects.

Keywords

topological analysis, persistent homology, image distortion, object classification, neural networks

Acknowledgements

The reported study was funded by the YSU Programme (the research project No. П12-К-1-Г-3/2021).

For citation: Ereemeev S.V., Abakumov A.V. Classification of objects in images with distortions based on a two-stage topological analysis. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 1, pp. 82–92 (in Russian). doi: 10.17586/2226-1494-2022-22-1-82-92

Введение

Проблема классификации объектов на изображениях в настоящее время является актуальной в самых различных сферах. Под классификацией изображений понимают соотнесение объектов на этих изображениях к одному из классов. Интерес исследователей к этой проблеме с каждым годом возрастает.

В первую очередь это связано с тем, что увеличивается сложность исходной информации, объекты на изображении подвергаются различным искажениям. Это могут быть шумовые факторы, включая освещение, наличие теней, погодные условия [1]. Также могут быть различные аффинные преобразования, получение изображений с разных масштабов и углов съемки в 3D-пространстве, частичное перекрытие объектов на картах городской инфраструктуры и др. Таким образом, увеличение сложности классифицируемых объектов на изображении влечет за собой совершенствование и разработку новых подходов для решения проблемы их классификации.

Цель работы заключается в разработке метода для извлечения топологических особенностей объекта на изображении, а также в исследовании устойчивости этих признаков при классификации с учетом внесения различных искажений в проверочную выборку.

Существующие методы и подходы

В мировом сообществе сформировалась общая концепция для классификации данных на изображениях, которая включает в себя следующие основные этапы: предварительная обработка изображений, извлечение признаков и непосредственно классификация по этим признакам. Каждый этап играет важную роль и вносит свой определенный вклад в итоговый результат.

В настоящее время существует большое количество разработок в области классификации объектов на изображениях, которые используют самые передовые технологии. При этом их можно условно разделить на две большие группы: традиционные и нейросетевые подходы. Последние исследования показывают, что нейросетевые подходы опережают другие методы.

В работе [2] отмечено, что в данном направлении также возникают следующие проблемы:

- переобучения;
- для каждой ситуации требуется своя выборка;
- увеличение вычислительной сложности.

Для решения этих проблем перспективным считается направление GAN (Generative Adversarial Networks), основная суть которого заключается в том, что одна сеть обучает другую.

Рассмотрим традиционные подходы, которые наиболее успешно справляются с наличием искажений при классификации изображений.

Наиболее популярный метод для извлечения признаков — SIFT (Scale Invariant Feature Transform), который позволяет выделить ключевые точки на изображении [3]. Набор признаков из этих точек инвариантен к аффинным преобразованиям. Автоматический поиск на изображении оптимальных ключевых точек является сложной задачей, одно из ее решений показано в работе [4].

Также популярны геометрические подходы, которые позволяют сформировать из ключевых точек графовые модели объекта [5]. При этом образуются топологические связи, которые сохраняются, например, при движении человека на видеопоследовательности. Еще один смежный, широко используемый подход — вычисление геодезических расстояний [6] между объектами интереса на некоторой поверхности, а также использование дифференциальной геометрии [7].

Текстурные признаки также содержат много полезной информации, которая может быть использована для учета различных искажений. К таким методам относятся подходы на основе локальных бинарных шаблонов [8], а также близкие к этой тематике новые направления в виде знаковых представлений изображений [9]. Эти подходы показали свою устойчивость к шуму. Также широко используются методы на основе гистограммы направленных градиентов (Histogram of Oriented Gradients, HOG) [11], которые сохраняют устойчивость признаков при разном освещении. Кроме того, часто для анализа текстурных признаков изображений используют Wavelet [10] и Фурье-преобразование [11], где обработка информации ведется в пространстве низких и высоких частот.

В отдельную группу можно выделить методы, которые используют семантическую информацию о близко расположенных объектах. Так, например, в работе [12] показан подход для контекстного анализа объектов на изображении. При этом заранее создаются правила, которые описывают, как объекты могут быть связаны между собой. В частности, площадка находится напротив дома и около дороги. Такой подход является достаточно перспективным для классификации объектов, особенно в условиях пересечения объектов (например, дерево закрывает часть дома). Однако создание правил является очень трудоемкой операцией, и количество этих правил будет резко увеличиваться при появлении новых типов объектов.

Непосредственно для классификации применяют уже ставшие классическими следующие широко используемые подходы: метод опорных векторов (Support Vector Machine, SVM) [13], дерево решений (Random Forest) [14], сверточные нейронные сети (Convolutional Neural Network, CNN) [15]. Зачастую, для методов машинного обучения требуется вектор признаков заданного размера, что вызывает определенные трудности при неопределенном количестве признаков.

В работе [16] выполнен обширный обзор самых современных методов для извлечения признаков. Вместе с этим продемонстрировано использование различных

классификаторов в сочетании с полученными признаками. На разных наборах данных наилучшие результаты показали нейронные сети, однако без учета искажений.

На взгляд авторов настоящей работы, для классификации объектов с учетом их искажений целесообразно использовать топологические методы, так как они сохраняют структуру объектов при различных поворотах, небольших деформациях, изменениях масштаба, а также при появлении шума.

Топология для анализа объектов применяется уже достаточно давно. Сначала это был подсчет количества компонент связности и дыр, которые используются для анализа бинарных изображений. Каждая компонента связности представляет собой отдельный объект из черных пикселей. Дыра характеризуется множеством белых пикселей, которые расположены внутри объекта. Сложный объект может содержать в себе несколько дыр. Относительно недавно появились новые методы алгебраической топологии [17], что вызвало интерес у многих научных лабораторий по всему миру. Основная идея топологического анализа данных заключается в вычислении и анализе глобальных структур данных, которые образуются множеством точек в n -мерном пространстве. В отличие от кластерного анализа учитывается продолжительность существования каждой такой структуры, что отражается в виде топологической характеристики, называемой баркодом [18]. Баркод представляет собой множество параллельных отрезков разной длины, которые содержат информацию о всех сформированных структурах, а также об их объединении. Отдельное направление занимает топологический анализ изображений [19, 20], где также можно анализировать топологические особенности. В этом случае баркод хранит информацию о компонентах связности и дырах, которые формируются на каждом уровне яркости.

Методология топологического анализа данных

В работе использована персистентная гомология, которая является подразделом топологического анализа данных. Построение топологических признаков из множества точек на плоскости с помощью персистентной гомологии заключается в следующем. Сначала соединяются наиболее близкие точки, затем расстояние между точками постепенно увеличивается. К топологическим характеристикам относят компоненты связности и дыры. Изначально каждая отдельная точка представляет собой компоненту связности, далее они объединяются в более крупные компоненты связности.

Наибольший интерес представляет информация о дырах, которые формируются при соединении точек между собой и образуют замкнутый контур. Если все точки контура будут соединены между собой, то дыра исчезает. Персистентная гомология фиксирует начало и окончание существования каждой дыры, а также систематизирует эту топологическую информацию в виде баркода.

Описанный принцип построения баркодов придает топологическим признакам устойчивость к изменению

объекта. В качестве множества точек можно использовать ключевые точки, выделенные на изображении, как, например, показано в работе [21].

Таким образом, сравнив два объекта разной целостности, можно узнать степень их отличия.

В настоящей работе рассмотрен каждый пиксел на изображении. Для формирования признаков постепенно будет увеличиваться не расстояние между объектами, а яркость, которая для изображений в градациях серого изменяется от 0 до 255. Для цветных изображений необходимо анализировать три составляющих: R (Red), G (Green) и B (Blue), каждая из которых имеет диапазон изменения яркостей от 0 до 255.

Сначала анализируется бинарное изображение с порогом 0, затем с порогом 1, 2, ..., 255. На каждом бинарном изображении вычисляются топологические особенности.

Под компонентой связности считается любой связанный между собой набор пикселов. Для образования компоненты связности, как и в случае с множеством точек, достаточно всего одного пиксела. Компоненты могут объединяться. В итоге после анализа всех бинарных изображений останется лишь одна компонента связности.

Определение дыр сложнее, так как они начинают существовать, если три рядом стоящие точки соединяются в треугольник. Один несвязанный ни с кем пиксел

или два рядом стоящих пиксела считаются потенциальными дырами. Если добавляется пиксел, находящийся рядом с дырой, то он присоединяется к ней. Если же дыра появляется у границы изображения, то она исчезает. Таким образом, не останется ни одной дыры.

Под началом существования дыры понимается яркость пиксела, после добавления которого образуется дыра, а конец существования — яркость пиксела, после добавления которого дыра исчезает.

На рис. 1 показан пример процесса формирования двух дыр при изменении яркости от максимального пиксела к минимальному. Также продемонстрировано поглощение одной дыры другой (рис. 1, e).

Первая дыра образована при яркости 5 (рис. 1, b), на котором также отдельно выделены две потенциальные дыры. Первая дыра поглощает вторую при яркости 2, тем самым вторая дыра заканчивает свое существование. В итоге топологическая информация о дырах содержится в баркоде, который для приведенного примера состоит из двух отрезков и показывает сведения о появлении и исчезновении каждой дыры (рис. 1, g).

Отметим, что на практике компоненты связности являются более слабой характеристикой, чем дыры. Это следует из того, что у всех баркодов, построенных по компонентам связности, всегда будет один большой отрезок, протяженностью во весь баркод.

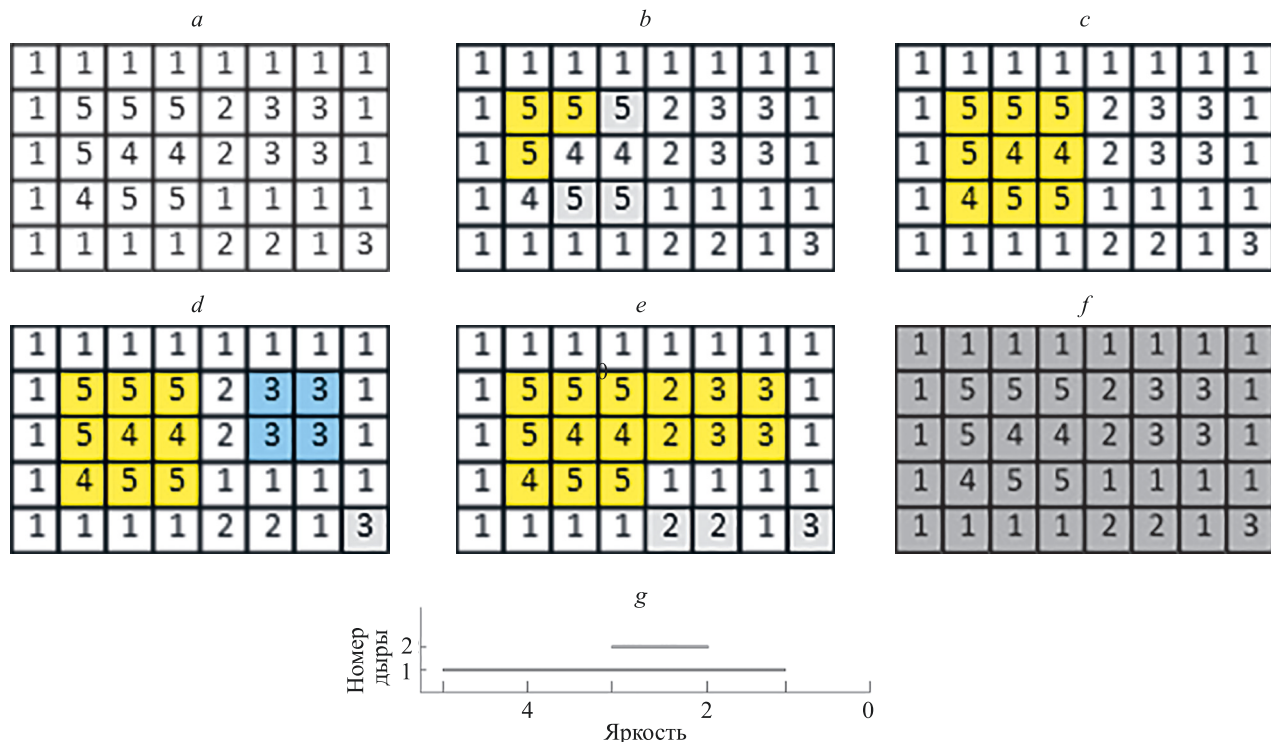


Рис. 1. Процесс формирования дыр: исходное изображение (a); выделение всех пикселов, которые больше или равны: 5 (b), 4 (c), 3 (d), 2 (e), 1 (f); систематизация информации о дырах в виде баркода (g).

Серым цветом обозначены потенциальные дыры, желтым и синим — сформированные дыры, темно-серым на рис. 1, f — исчезающая дыра

Fig. 1. Holes formation process: the source image (a); selection of all pixels that are greater than or equal to: 5 (b), 4 (c), 3 (d), 2 (e), 1 (f); systematization of information about holes is shown in the barcode (g).

Potential holes are marked in gray, formed holes are marked in yellow and blue, disappeared holes are marked in dark gray

Предлагаемый подход

В работе рассмотрен двухэтапный анализ изображения для извлечения топологических признаков в виде информации о дырах. Схема предлагаемого подхода показана на рис. 2.

$$\text{Пусть } I = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix} \text{ — исходное изобра-}$$

жение в градациях серого размером $m \times n$.

На первом шаге координаты точек изображения сортируются в порядке возрастания значений их яркостей, т. е. $I(x_{\alpha_1}, y_{\beta_1}) \leq I(x_{\alpha_2}, y_{\beta_2}) \leq \dots \leq I(x_{\alpha_{m \times n}}, y_{\beta_{m \times n}})$, где $\alpha_1, \alpha_2, \dots, \alpha_k, \dots, \alpha_{m \times n}$ и $\beta_1, \beta_2, \dots, \beta_k, \dots, \beta_{m \times n}$ — номера координат точек изображения I по строкам и столбцам соответственно; $I(x_{\alpha_k}, y_{\beta_k})$ — значение яркости в точке с координатами (α_k, β_k) .

На втором шаге формируются топологические признаки по дырам при анализе изображений в диапазоне яркостей от 0 до 255, а затем от 255 до 0. При этом для каждого из двух этапов анализа изображений создается свой баркод, представляющий собой набор отрезков, каждый из которых описывает поведение отдельной дыры. Для первого этапа имеем баркод $B_0^{255} = \{h_1, h_2, \dots, h_s\}$, где s — количество дыр. В свою очередь, $h_i = (b_i, l_i)$, где b_i ($i = 1, 2, \dots, s$) — яркость, на которой появляется дыра; l_i — длина отрезка, который соответствует периоду существования дыры. Для второго этапа при анализе изображения в диапазоне яркостей от 255 до 0 — баркод: $B_{255}^0 = \{h'_1, h'_2, \dots, h'_t\}$. В результате получим следующий набор топологических признаков: $V = \{h_1, h_2, \dots, h_s, h'_1, h'_2, \dots, h'_t\}$.

На этапе предобработки баркодов выполним отсеив шумов — коротких отрезков на баркоде, экспериментально полученные длины которых составляют 10 значений яркости. Это означает, что те отрезки h_i и h'_j , которые имеют длину $l_i < 10$ и $l'_j < 10$ соответственно, отсекаются и не несут существенной информации о структуре объекта на изображении. Это позволяет исключить влияние шума на точность классификации.

Приведем примеры исходных баркодов и после отсеивания шума (рис. 3).

Отфильтрованный набор признаков V поступает на заключительный этап классификации объектов. Сравнение объектов происходит в пространстве то-

пологических признаков. Для соотнесения объекта к определенному классу использован принцип наложения одного баркода на другой [22]. Пусть даны два набора топологических признаков $V_1 = \{h_1, h_2, \dots, h_{s1}, h'_1, h'_2, \dots, h'_{t1}\}$ и $V_2 = \{g_1, g_2, \dots, g_{s2}, g'_1, g'_2, \dots, g'_{t2}\}$ таких, что длина каждого отрезка $h \in V_1$ и $g \in V_2$ больше или равна 10. Сначала вычисляется длина общей части при пересечении двух самых больших по длине отрезков из подмножеств $\{h_1, h_2, \dots, h_{s1}\}$ и $\{g_1, g_2, \dots, g_{s2}\}$ с учетом параметра яркости, на которой появилась каждая из дыр. Затем вычисляется пересечение следующих по максимальной длине двух дыр и т. д. Тем самым определяется вклад в общую схожесть каждой пары отрезков. Аналогичная процедура выполняется для подмножеств $\{h'_1, h'_2, \dots, h'_{t1}\}$ и $\{g'_1, g'_2, \dots, g'_{t2}\}$. Чем больше суммарная площадь пересечения, тем больше схожи баркоды и, значит, больше схожи объекты между собой.

Для каждого класса используется набор из эталонных объектов. Для них также вычисляется набор топологических признаков. Проверка на схожесть баркодов происходит с каждым эталоном. После процедуры сравнения выдается результат о принадлежности объекта к определенному классу.

Программный код для классификации объектов на основе двухэтапного топологического анализа разработан на языке C++¹.

Результаты экспериментов и их обсуждение

Проведем ряд экспериментов и сравним результаты классификации предлагаемого подхода с одной из нейронных сетей. Для сравнения точности классификации и устойчивости подходов выполним тестирование на наборе данных DOTA v1.5².

DOTA представляет собой набор спутниковых снимков и файл с зонами интереса конкретных объектов. В дополнении к координатам зоны интереса, запись об объекте содержит показатель сложности (0 — не сложно, 1 — сложно).

¹ [Электронный ресурс]. Режим доступа: <https://github.com/Noremos/barclassifier> (дата обращения: 18.12.2021).

² Официальная страница набора данных DOTA [Электронный ресурс]. Режим доступа: <https://captain-whu.github.io/DOTA/dataset.html> (дата обращения: 18.12.2021).

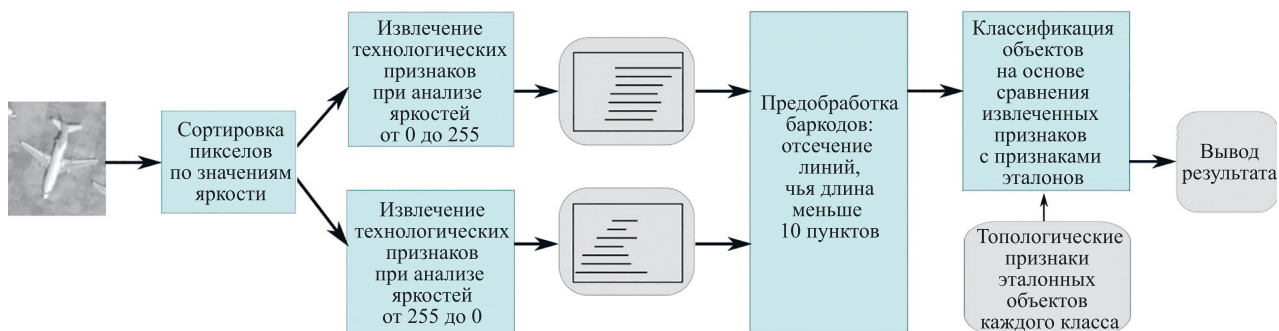


Рис. 2. Схема двухэтапного топологического анализа для классификации объектов на изображении

Fig. 2. Scheme for a two-step topological analysis to classify objects in an image

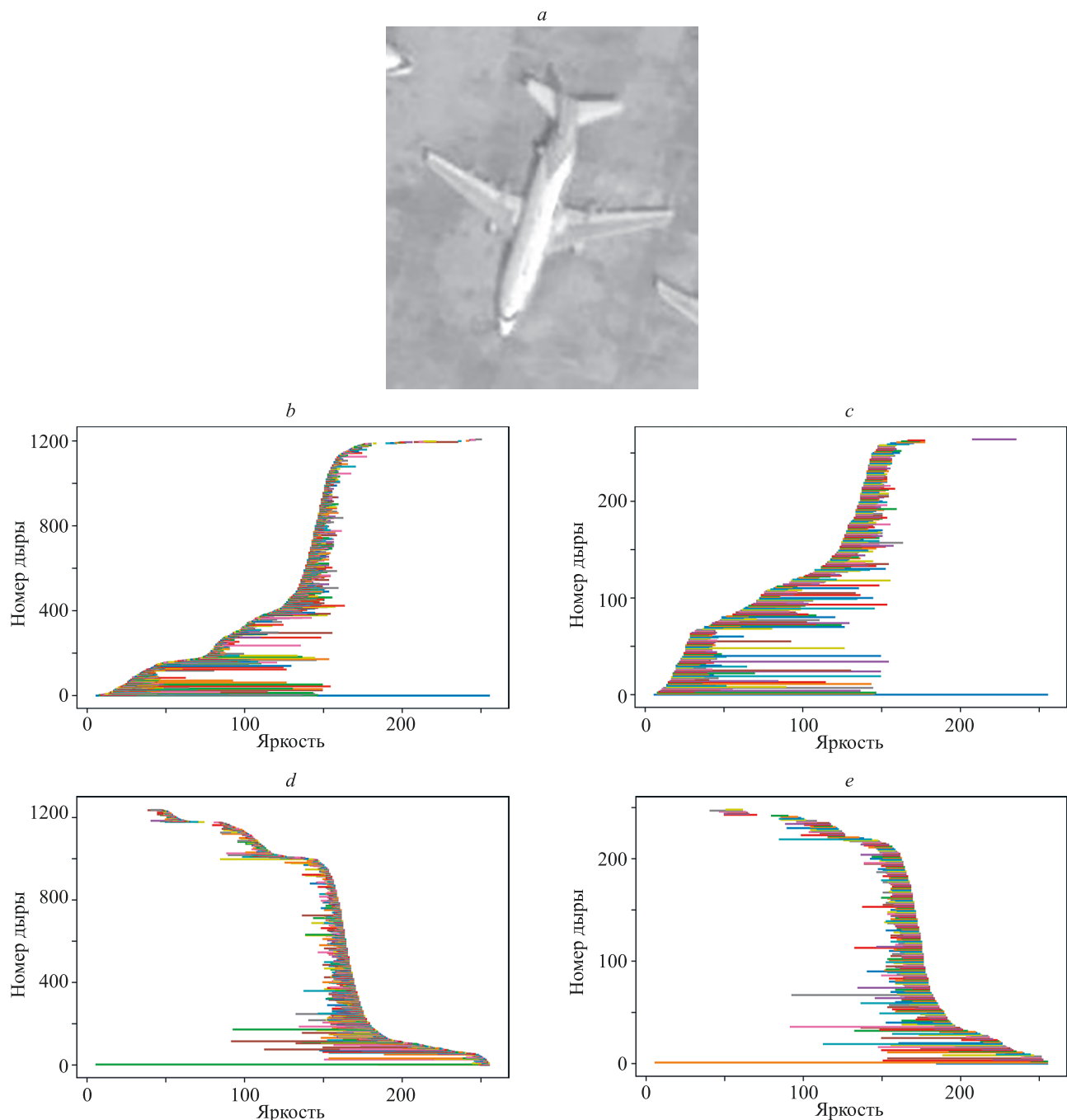


Рис. 3. Топологические признаки изображения в виде баркодов: исходное изображение (a); баркод B_0^{255} при анализе изображения в диапазоне яркостей от 0 до 255 с шумом (b) и без шума (c); баркод B_{255}^0 при анализе изображения в диапазоне яркостей от 255 до 0 с шумом (d) и без шума (e)

Fig. 3. Topological features of the image in the form of barcodes: the original image (a), barcode B_0^{255} after image analysis with brightness from 0 to 255 with noise (b) and without noise (c), barcode B_{255}^0 after image analysis with brightness from 255 to 0 with noise (d) and without noise (e)

Главное ограничение при использовании нейронных сетей – необходимость иметь большой набор изображений, нужных для обеспечения максимальной адаптивности. При этом нет гарантии, что нейронная сеть будет корректно работать на новых, отличных от исходного набора данных. Отметим, что чем больше данных было при обучении, тем сложнее будет дообучить сеть.

Сравнение с помощью баркодов работает иначе. Схожесть двух изображений выражается в процентном

соотношении. Для использования баркодов в задаче классификации изображений необходимо отобрать ряд эталонов для каждого класса и сравнивать каждый из них с классифицируемым объектом. Причем качество эталонов гораздо важнее количества, поэтому для корректной работы достаточно отобрать несколько наиболее уникальных объектов каждого класса.

По сравнению с новой областью топологического анализа данных, нейронные сети сильно развиты, осо-

бенно в области обработки изображений. Речь идет о CNN. Они способны показывать хорошие результаты при низком количестве и качестве обучаемых данных. Возможно провести сравнение с классической архитектурой нейронной сети, но больший интерес представляет сравнение баркодов с CNN, архитектура которой показана на рис. 4.

Из набора данных для исследования выбраны изображения со следующими классами: самолет, корабль, резервуар для хранения, теннисный корт, баскетбольная

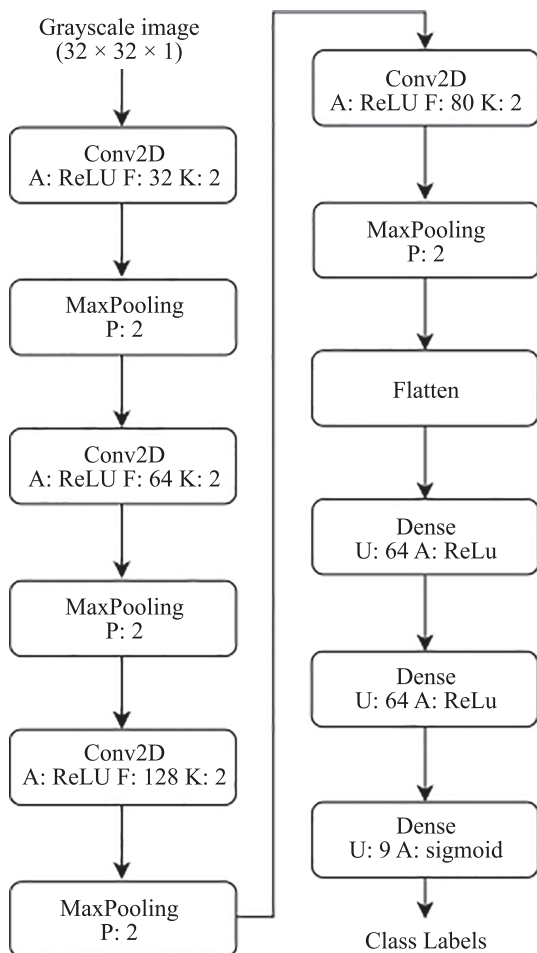


Рис. 4. Архитектура сверточной нейронной сети. Conv2D — слой свертки с параметрами: A (activation function) — функция активации, F (filters) — количество выходных фильтров в свертке, K (kernel size) — ядро окна свертки; Max Pooling — слой подвыборки, где P (pool size) — уменьшает входные данные в P раз; Flatten — преобразует входную матрицу в один массив; Dense — полносвязный слой, где U (units) — размерность выходного пространства; Class Labels — метки

Fig. 4. The architecture of the convolutional neural network. Conv2D is a convolution layer with the following parameters: A is the activation function, F is a number of output filters in the convolution, K is a kernel size; Max Pooling is a subsample layer, where P (pool size) reduces the input data by a factor of P; Flatten is a layer that converts the input matrix into a single array; Dense is a deeply connected layer, where U (units) is the dimensionality of the output space; Class Labels are the output labels

площадка, мост, небольшое транспортное средство, вертолет и контейнерный кран. Параметр сложности: 0.

Классификатор для оценки баркодов построен по простому принципу: для каждого класса вычисляется набор баркодов, построенных по каждому из изображений, принадлежащих этому классу. Во время валидации по входному изображению строится баркод, который поочередно сравнивается с каждым баркодом из каждого набора, где выбирается пара с наибольшим процентом схожести.

Помимо обычного набора данных для валидации в работе выполнено исследование устойчивости к искажениям. Главная особенность персистентной гомологии в том, что она в большинстве случаев устойчива именно к искажениям, в то время как нейронная сеть не может распознать искажение объекта, которого не было в исходной выборке.

Во время экспериментов в качестве искажений изображений применены три параметра: поворот — α ; сужение — β ; обрезание края — γ .

Нейронная сеть обучена на выборке из 87 266 экземпляров. Без внесения искажений тест валидации, состоящий из 26 249 объектов, для нейронной сети дал точность 98,95 %. Базовая точность классификатора по баркодам на наборе для валидации составила около 82 %. Для вычисления точности (Precision) используем формулу:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}},$$

где TP и FP — число правильно и неправильно сопоставленных классов.

Заметим, что классификатору по баркодам не нужно много однородных изображений, достаточно поместить уникальные случаи. Однако для чистоты эксперимента этот классификатор принимает ту же обучающую выборку, что и нейронная сеть.

Внесем искажения в изображения, которые используются для теста валидации. Примеры полученных искажений показаны в табл. 1.

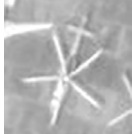
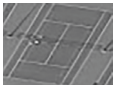
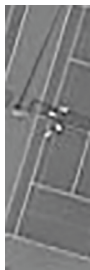
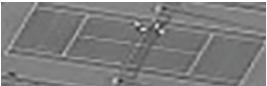
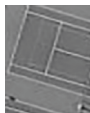
Результаты экспериментов сведены в табл. 2. Высокое значение точности нейронной сети может означать наличие переобучения, поэтому была проведена повторная проверка с меньшим числом эпох.

Из полученных результатов видно, что количество эпох практически не влияет на результаты, поэтому подробное сравнение можно опустить. Заметим, что больше всего на нейронные сети влияет обрезание. Из табл. 2 видно, что предложенный метод превосходит нейронную сеть CNN для этого типа искажений.

Приведем результаты, показывающие снижение точности классификации при внесении различных искажений (табл. 3).

Из табл. 3 видно, что при сужении или обрезании края изображения предложенный метод превосходит результаты, чем классификатор на основе нейронной сети, где наихудший результат проявляется при $\gamma = 50$. При добавлении поворота точность у предложенного метода (строки с параметрами 0, 0, 50 и 90, 0, 50 в табл. 2 и 3) несколько возрастает, что может показаться несколько необычным. Однако эта погрешность

Таблица 1. Примеры исходных и искаженных изображений
Table 1. Examples of original and distorted images

Параметры искажения			Визуальное искажение объекта					
$\alpha, ^\circ$	$\beta, \%$	$\gamma, \%$	Машина	Вертолет	Теннисный корт	Самолет	Лодка	Резервуар
0	0	0						
90	0	0						
0	50	0						
0	0	50						
90	50	0						
90	0	50						
90	50	50						

возникает из-за изменившегося порядка анализа однотонных пикселей, так как они проверяются с левого верхнего края до правого нижнего построчно. Потому эта погрешность является лишь особенностью реализации.

Подводя итоги результатов экспериментов, можно сказать, что если решается задача, в которой есть вероятность получить на вход данные с непредсказуемыми искажениями, а набор эталонов сильно ограничен, то имеет смысл обратить внимание на классификатор

Таблица 2. Сравнение результатов точности классификации объектов с учетом искажений предложенным методом и с помощью нейронной сети CNN

Table 2. Comparison of the results of classification accuracy object with distortion using the proposed method and the CNN neural network

Параметры искажения			Предложенный метод	Сверточная нейронная сеть (количество эпох обучения)			
α , °	β , %	γ , %		16	10	8	5
0	0	0	0,8193	0,9895	0,9885	0,9830	0,9755
90	0	0	0,8069	0,9169	0,9139	0,9136	0,9058
0	50	0	0,7790	0,9199	0,9244	0,9111	0,9252
0	0	50	0,7124	0,5851	0,6282	0,5520	0,5754
90	50	0	0,7818	0,8654	0,8633	0,8609	0,8654
90	0	50	0,7203	0,5966	0,5870	0,5875	0,5945
90	50	50	0,7203	0,5290	0,5364	0,5445	0,5369

Таблица 3. Снижение точности классификации объектов предложенным методом и с помощью нейронной сети CNN при внесении искажений

Table 3. Decreasing the accuracy of object classification using the proposed method and the CNN neural network when introducing distortions

Параметры искажения			Предложенный метод, %	CNN, 16 эпох обучения, %
α , °	β , %	γ , %		
90	0	0	-1,51	-7,34
0	50	0	-4,91	-7,03
0	0	50	-13,05	-40,87
90	50	0	-4,57	-12,54
90	0	50	-12,07	-39,71
90	50	50	-12,07	-46,54

с помощью топологического анализа данных. Если же задача подразумевает анализ и классификацию изображений без искажений, то следует использовать нейросетевые подходы.

Заключение

Разработан метод двухэтапного топологического анализа для классификации объектов на изображениях.

Основу метода составляет персистентная гомология, которая позволяет извлекать топологические признаки и хранить их в виде баркода. Особенность полученных признаков заключается в том, что они устойчивы к искажениям. Проведено исследование работы предложенного метода с нейросетевым подходом. Для этого был выбран набор данных DOTA, который содержит снимки пространственных объектов нескольких классов. Без искажений нейронная сеть показала очень высокие результаты, точность классификации составила свыше 98 %, у предложенного метода около 82 %. Для экспериментов выбраны следующие искажения изображений: поворот на 90°, сужение и обрезание края на 50 %, а также их комбинации. В результате предложенный метод показал свою устойчивость. Особенно резко у нейронной сети теряется точность классификации

при обрезании края изображения. В самой сложной комбинации теста снижение точности классификации для нейросетевого подхода составило 46 %, а для предложенного метода 12 %. Тем самым эксперименты подтвердили, что при заданных комбинациях искажений предложенный метод превосходит нейронную сеть CNN. Это достигается за счет того, что признаки, полученные с помощью двухэтапного топологического анализа, сохраняют общую структуру объекта даже при частичной потере исходной информации.

Данный метод целесообразно использовать для классификации пространственных объектов на спутниковых снимках, так как на них часто присутствуют рассмотренные в работе искажения. Также можно применить предложенный метод в системах технического зрения промышленных предприятий для автоматической классификации изделий и в смежных областях.

Исходя из проведенных исследований, можно сказать, что персистентная гомология является перспективной областью для изучения и разработки на ее основе новых методов, в том числе для классификации объектов на изображениях, особенно с учетом искажений. Дальнейшее развитие может открыть новые способы применения топологического анализа и те области, где они наиболее эффективны.

1. Ярошевич П.В., Богущ Р.П. Алгоритм классификации изображений парковочных мест автостоянки на основе гистограмм ориентированных градиентов и метода опорных векторов // Компьютерная оптика. 2017. Т. 41. № 1. С. 110–117. <https://doi.org/10.18287/2412-6179-2017-41-1-110-117>
2. Краснов Ф.В., Буторин А.В., Ситников А.Н. Обзор подходов к анализу пространственных изображений высокого разрешения для применения в геофизике // Cloud of Science. 2019. Т. 6. № 1. С. 127–143.
3. Fu J., Rui Y. Advances in deep learning approaches for image tagging // APSIPA Transactions on Signal and Information Processing. 2017. V. 6. P. e11. <https://doi.org/10.1017/ATSIP.2017.12>
4. Ященко А.В., Беликов А.В., Петерсон М.В., Потапов А.С. Дистилляция нейросетевых моделей для детектирования и описания ключевых точек изображений // Научно-технический вестник информационных технологий, механики и оптики. 2020. Т. 20. № 3. С. 402–409. <https://doi.org/10.17586/2226-1494-2020-20-3-402-409>
5. Blyumin S., Pogodaev A., Khabibullina E. Graph-structural modeling of some special organizational systems // Proc. of the 2nd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA). 2020. P. 279–283. <https://doi.org/10.1109/SUMMA50634.2020.9280724>
6. Handrich S. Al-Hamadi A. Localizing body joints from single depth images using geodetic distances and random tree walk // Proc. 24th IEEE International Conference on Image Processing (ICIP). 2017. P. 146–150. <https://doi.org/10.1109/ICIP.2017.8296260>
7. Karimova L., Terekhov A., Makarenko N., Rybintsev A. Methods of computational topology and discrete Riemannian geometry for the analysis of arid territories // Cogent Engineering. 2020. V. 7. N 1. P. 1808340. <https://doi.org/10.1080/23311916.2020.1808340>
8. Ojala T., Pietikäinen M., Mäenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002. V. 24. N 7. P. 971–987. <https://doi.org/10.1109/TPAMI.2002.1017623>
9. Мясников В.В. Реконструкция функций и цифровых изображений по их знаковым представлениям // Компьютерная оптика. 2019. Т. 43. № 6. С. 1041–1052. <https://doi.org/10.18287/2412-6179-2019-43-6-1041-1052>
10. Luo Z., Chen J., Takiguchi T., Arika Y. Neutral-to-emotional voice conversion with cross-wavelet transform F0 using generative adversarial networks // APSIPA Transactions on Signal and Information Processing. 2019. V. 8. P. 1–11. <https://doi.org/10.1017/ATSIP.2019.3>
11. Ruckdeschel P., Kohl M. General purpose convolution algorithm in S4 classes by means of FFT // Journal of Statistical Software. 2014. V. 59. N 4. P. 1–25. <https://doi.org/10.18637/jss.v059.i04>
12. Касимов Д.Р. Моделирование рассуждений при поиске и описании объектов на космоснимках // Компьютерная оптика. 2020. Т. 44. № 5. С. 772–781. <https://doi.org/10.18287/2412-6179-CO-716>
13. Sahu S.K., Pujari A.K., Kagita V.R., Kumar V., Padmanabhan V. GP-SVM: Tree structured multiclass SVM with greedy partitioning // Proc. of the 14th International Conference on Information Technology (ICIT). 2015. P. 142–147. <https://doi.org/10.1109/ICIT.2015.24>
14. Zhou Q., Lan W., Zhou Y., Mo G. Effectiveness evaluation of anti-bird devices based on random forest algorithm // Proc. of the 7th International Conference on Information, Cybernetics, and Computational Social Systems (ICCSS). 2020. P. 743–748. <https://doi.org/10.1109/ICCSS52145.2020.9336891>
15. Xu H., Chen Y., Lin R., Kuo J. Understanding convolutional neural networks via discriminant feature analysis // APSIPA Transactions on Signal and Information Processing. 2018. V. 7. P. e20. <https://doi.org/10.1017/ATSIP.2018.24>
16. Рюмина Е.В., Карпов А.А. Аналитический обзор методов распознавания эмоций по выражениям лица человека // Научно-технический вестник информационных технологий, механики и оптики. 2020. Т. 20. № 2. С. 163–176. <https://doi.org/10.17586/2226-1494-2020-20-2-163-176>
17. Edelsbrunner H., Letscher H., Zomorodian A. Topological persistence and simplification // Discrete and Computational Geometry. 2002. V. 28. N 4. P. 511–533. <https://doi.org/10.1007/s00454-002-2885-2>
1. Yarashevich P.V., Bohush R.P. Classification algorithm of parking space images based on a histogram of oriented gradients and support vector machines. *Computer Optics*, 2017, vol. 41, no 1, pp. 110–117. <https://doi.org/10.18287/2412-6179-2017-41-1-110-117>. (in Russian)
2. Krasnov F.V., Butorin A.V., Sitnikov A.N. Review of approaches to the analysis of high-resolution spatial images for Geophysics. *Cloud of Science*, 2019, vol. 6, no. 1, pp. 127–143. (in Russian)
3. Fu J., Rui Y. Advances in deep learning approaches for image tagging. *APSIPA Transactions on Signal and Information Processing*, 2017, vol. 6, pp. e11. <https://doi.org/10.1017/ATSIP.2017.12>
4. Yashchenko A.V., Belikov A.V., Peterson M.V., Potapov A.S. Distillation of neural network models for detection and description of image key points. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2020, vol. 20, no. 3, pp. 402–409. (in Russian). <https://doi.org/10.17586/2226-1494-2020-20-3-402-409>
5. Blyumin S., Pogodaev A., Khabibullina E. Graph-structural modeling of some special organizational systems. *Proc. 2nd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA)*, 2020, pp. 279–283. <https://doi.org/10.1109/SUMMA50634.2020.9280724>
6. Handrich S. Al-Hamadi A. Localizing body joints from single depth images using geodetic distances and random tree walk. *Proc. 24th IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 146–150. <https://doi.org/10.1109/ICIP.2017.8296260>
7. Karimova L., Terekhov A., Makarenko N., Rybintsev A. Methods of computational topology and discrete Riemannian geometry for the analysis of arid territories. *Cogent Engineering*, 2020, vol. 7, no. 1, pp. 1808340. <https://doi.org/10.1080/23311916.2020.1808340>
8. Ojala T., Pietikäinen M., Mäenpää T. Multiresolution grayscale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, vol. 24, no. 7, pp. 971–987. <https://doi.org/10.1109/TPAMI.2002.1017623>
9. Myasnikov V.V. Reconstruction of functions and digital images using sign representations. *Computer Optics*, 2019, vol. 43, no. 6, pp. 1041–1052. (in Russian). <https://doi.org/10.18287/2412-6179-2019-43-6-1041-1052>
10. Luo Z., Chen J., Takiguchi T., Arika Y. Neutral-to-emotional voice conversion with cross-wavelet transform F0 using generative adversarial networks. *APSIPA Transactions on Signal and Information Processing*, 2019, vol. 8, pp. 1–11. <https://doi.org/10.1017/ATSIP.2019.3>
11. Ruckdeschel P., Kohl M. General Purpose Convolution Algorithm in S4 Classes by Means of FFT. *Journal of Statistical Software*, 2014, vol. 59, no. 4, pp. 1–25. <https://doi.org/10.18637/jss.v059.i04>
12. Kasimov D.R. Searching and describing objects in satellite images on the basis of modeling reasoning. *Computer Optics*, 2020, vol. 44, no. 5, pp. 772–781. (in Russian). <https://doi.org/10.18287/2412-6179-CO-716>
13. Sahu S.K., Pujari A.K., Kagita V.R., Kumar V., Padmanabhan V. GP-SVM: Tree structured multiclass SVM with greedy partitioning. *International Conference on Information Technology (ICIT)*, 2015, pp. 142–147. <https://doi.org/10.1109/ICIT.2015.24>
14. Zhou Q., Lan W., Zhou Y., Mo G. Effectiveness evaluation of anti-bird devices based on random forest algorithm. *Proc. of the 7th International Conference on Information, Cybernetics, and Computational Social Systems (ICCSS)*, 2020, pp. 743–748, <https://doi.org/10.1109/ICCSS52145.2020.9336891>
15. Xu H., Chen Y., Lin R., Kuo J. Understanding convolutional neural networks via discriminant feature analysis. *APSIPA Transactions on Signal and Information Processing*, 2018, vol. 7, pp. e20. <https://doi.org/10.1017/ATSIP.2018.24>
16. Ryumina E.V., Karpov A.A. Analytical review of methods for emotion recognition by human face expressions. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2020, vol. 20, no. 2, pp. 163–176. (in Russian). <https://doi.org/10.17586/2226-1494-2020-20-2-163-176>
17. Edelsbrunner H., Letscher H., Zomorodian A. Topological persistence and simplification. *Discrete and Computational Geometry*, 2002, vol. 28, no. 4, pp. 511–533. <https://doi.org/10.1007/s00454-002-2885-2>
18. Carlsson E., Carlsson G., de Silva V. An algebraic topological method for feature identification. *International Journal of Computational*

18. Carlsson E., Carlsson G., de Silva V. An algebraic topological method for feature identification // *International Journal of Computational Geometry and Applications*. 2006. V. 16. N 4. P. 291–314. <https://doi.org/10.1142/S021819590600204X>
19. Макаренко Н.Г., Уртвев Ф.А., Князева И.С., Малкова Д.Б., Пак И.Т., Каримова Л.М. Распознавание текстур на цифровых изображениях методами вычислительной топологии // *Современные проблемы дистанционного зондирования Земли из космоса*. 2015. Т. 12. № 1. С. 131–144.
20. Gonzalez-Diaz R., Jimenez M.-J., Medrano B. Spatiotemporal barcodes for image sequence analysis // *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2015. V. 9448. P. 61–70. https://doi.org/10.1007/978-3-319-26145-4_5
21. Еремеев С.В., Андрианов Д.Е., Титов В.С. Алгоритм совмещения пространственных объектов разномасштабных карт на основе топологического анализа данных // *Компьютерная оптика*. 2019. Т. 43. № 6. С. 1021–1029. <https://doi.org/10.18287/2412-6179-2019-43-6-1021-1029>
22. Еремеев С.В. Абакумов А.В. Программный комплекс для обнаружения и классификации природных объектов на основе топологического анализа // *Программные продукты и системы*. 2021. Т. 34. № 1. С. 201–208. <https://doi.org/10.15827/0236-235X.133.201-208>
- Geometry and Applications*, 2006, vol. 16, no. 4, pp. 291–314. <https://doi.org/10.1142/S021819590600204X>
19. Makarenko N.D., Urtiev F.A., Knyazeva I.S., Malkova D., Pak I.T., Karimova L.M. Texture recognition in digital images by computational topology methods. *Sovremennye Problemy Distantionnogo Zondirovaniya Zemli iz Kosmosa*, 2015, vol. 12, no. 1, pp. 131–144. (in Russian)
20. Gonzalez-Diaz R., Jimenez M.-J., Medrano B. Spatiotemporal barcodes for image sequence analysis. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2015, vol. 9448, pp. 61–70. https://doi.org/10.1007/978-3-319-26145-4_5
21. Eremeev S.V., Andrianov D.E., Titov V.S. An algorithm for matching spatial objects of different-scale maps based on topological data analysis. *Computer Optics*, 2019, vol. 43, no. 6, pp. 1021–1029. (in Russian). <https://doi.org/10.18287/2412-6179-2019-43-6-1021-1029>
22. Eremeev S.V., Abakumov A.V. Software complex for detection and classification of natural objects based on topological analysis. *Software & Systems*, 2021, vol. 34, no. 1, pp. 201–208. (in Russian). <https://doi.org/10.15827/0236-235X.133.201-208>

Авторы

Еремеев Сергей Владимирович — кандидат технических наук, доцент, доцент, Муромский институт (филиал) ФГБОУ ВО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых», Муром, 602252, Российская Федерация, <https://orcid.org/0000-0001-8482-1479>, sv-eremeev@yandex.ru

Абакумов Артём Владимирович — студент, Муромский институт (филиал) ФГБОУ ВО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых», Муром, 602252, Российская Федерация, <https://orcid.org/0000-0001-5784-7147>, artem210966@yandex.ru

Authors

Sergey V. Eremeev — PhD, Associate Professor, Associate Professor, Murom Institute (Branch) of Vladimir State University, Murom, 602252, Russian Federation, <https://orcid.org/0000-0001-8482-1479>, sv-eremeev@yandex.ru

Artyom V. Abakumov — Student, Murom Institute (Branch) of Vladimir State University, Murom, 602252, Russian Federation, <https://orcid.org/0000-0001-5784-7147>, artem210966@yandex.ru

Статья поступила в редакцию 07.12.2021
Одобрена после рецензирования 20.12.2021
Принята к печати 26.01.2022

Received 07.12.2021
Approved after reviewing 20.12.2021
Accepted 26.01.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»