Интеллектуальные системы управления, анализ данных

© 2023 г. А.М. МИХАЙЛОВ, канд. техн. наук (alxmikh@gmail.com),
М.Ф. КАРАВАЙ, д-р. техн. наук (mkaravay@yandex.ru),
В.А. СИВЦОВ (TheDeGe@yandex.ru)
(Институт проблем управления им. В.А. Трапезникова РАН, Москва),
М.А. КУРНИКОВА (mish2109@yandex.ru)
(Национальный медицинский исследовательский центр детской гематологии, онкологии и иммунологии им. Дмитрия Рогачева, Москва)

МАШИННОЕ ОБУЧЕНИЕ ДЛЯ ДИАГНОСТИКИ ЗАБОЛЕВАНИЙ ПО ПОЛНОМУ ПРОФИЛЮ ЭКСПРЕССИИ ГЕНОВ

Рассматривается использование машинного обучения для диагностики заболеваний, основанной на анализе полного профиля экспрессии генов, что отличает данную работу от других подходов, где необходимо проведение предварительного этапа, на котором производится поиск ограниченного числа релевантных генов (десятки и сотни генов). Проведены эксперименты с полными профилями генетической экспрессии (20 531 генов), полученными в результате обработки транскриптомов 801 пациента с известными онкологическими диагнозами (онкология легких, почек, молочной железы, простаты и толстой кишки). Использование индекстрона (индексной системы мгновенного обучения) по новому назначению, т.е. для обработки полных профилей экспрессии, обеспечило точность диагностирования, которая на 99,75 процентов совпала с результатами гистологической верификации.

Ключевые слова: распознавание образов, машинное обучение, обратные образы, профили экспрессии генов, диагностика заболеваний.

DOI: 10.31857/S000523102307005X, **EDN:** FDMPHU

1. Введение

Общим знаменателем существующих подходов к диагностике онкологических заболеваний по профилям экспрессии генов, см., например, методологические и обзорные работы [1–6], является использование ограниченного набора генов. При таком подходе также необходимо проводить предварительный анализ и выделение отдельных генов или комбинаций генов, экспрессионная активность которых наиболее характерна в случае тех или иных заболеваний. При этом достигнутая точность диагностики не превышает 95%, на наборах, включающих до 90 генов [4], что связано, по-видимому, с ограниченным числом используемых генов. Так, в [5] это число снижается даже до 10 генов.

Таким образом, работа с данными проходит в два этапа. Например, в [6] на первом этапе используется анализ принципиальных компонент, в результате которого отбираются 103 гена. Окончательная диагностика производится на втором этапе, когда в результате использования байесовской нейронной сети была обеспечена точность в 93,66%, глубокая нейронная сеть обеспечила 93,41%, а логистическая регрессия — 92,82%. При этом возможно использование третьего этапа, когда отбрасываются сомнительные результаты, определяемые порогом принятия решения, что позволяет почти полностью исключить ошибки диагностирования [6] ценой того, что часть случаев не диагностируется, что считается более приемлемым, чем ошибочный диагноз.

В настоящей работе рассматривается использование машинного обучения для диагностики заболеваний, основанной на анализе полного профиля экспрессии генов, т.е. активности всех 20 531 известных генов, что упрощает работу, так как при этом отпадает необходимость предварительного отбора релевантных генов. В статье рассматриваются подробности работы этой использованной системы машинного обучения.

В биологии уровень экспрессии оценивает транскрипционную активность гена как количество произведенной им матричной РНК (мРНК). Получение профиля экспрессии генов — это малоинвазивная или полностью не инвазивная процедура, как, например, взятие образца слюны, чем определяется комфортность такой процедуры для пациента. При этом качество диагностирования зависит от полноты профиля, т.е. от количества рассмотренных генов. Однако использование полных профилей экспрессии генов — это, скорее, будущее медицинской диагностики, так как в настоящее время дешевое оборудование для массового применения с целью получения таких профилей отсутствует. В настоящее время профили могут быть найдены, в частности, с помощью следующих технологий [7]:

— ПЦР-тесты;

— ДНК-микрочипы (microarrays);

— секвенирование.

Если речь идет об анализе уровня экспрессии относительно небольшого количества генов, может использоваться доступный и сравнительно недорогой метод количественной ПЦР в реальном времени (qPCR). Вторая технология — ДНК-микрочипы, дает более качественную оценку экспрессии генов. Однако когда речь идет о экспрессионном профиле всех генов, то на сегодняшний день для этого необходимо использование секвенирования. Одной из наиболее востребованных платформ для высокопроизводительного секвенирования является оборудование компании Illumina [8]. Система HiSeq 2500 применяет технологию нового поколения SBS, которая поддерживает массово-параллельное секвенирование, используя метод флуоресцентно-меченных нуклеотидов, который позволяет прочтение отдельных оснований по мере их включения в растущие нити ДНК. На рис. 1 представлен вид оборудования, использующего этот метод.

Рис. 1. Система HiSeq 2500.

Illumina HiSeq выполняет две функции, такие как чтение генома и определение уровня экспрессии. Последнее находится путем прочтения не ДНК, а мРНК-кода. При этом по количеству копий мРНК судят об уровне экспрессии. С учетом того, что стоимость РНК-секвенирования падает, в перспективе возможно создание более простого и дешевого оборудования для установки в таких широко распространенных лабораториях, как INVITRO. При массовой доступности такого оборудования речь могла бы идти о перспективной диагностике, поскольку использование полного профиля экспрессии автоматически учитывает все возможные **активные** комбинации генов, которые трудно предугадать и которые влияют на возникновение различных заболеваний. Кроме того, профили экспрессии генов служат важным материалом не только при диагностике, но и в разнообразных научных и клинических исследованиях.

Вторая составляющая рассматриваемой методологии — это машинное обучение, позволяющее автоматически, без анализа генетических комбинаций, обучаться на полных биологических профилях с целью их последующей классификации и диагностирования. Традиционно для машинного обучения используются искусственные нейронные сети [9], являющиеся популярным методом, применяемым для решения множества различных задач предсказания и распознавания образов. Однако обучение таких сетей занимает много времени, требуя больших вычислительных ресурсов, что связано с необходимостью вычисления огромного количества коэффициентов для адаптации многослойной сетевой архитектуры к конкретной задаче. Тем не менее обучение можно намного ускорить и сделать чуть ли не мгновенным, если воспользоваться системой индексации образов [10, 11], напоминающей поисковики типа Google [12], где инкрементное обучение сводится к индексированию новых документов. В 1998 г. был введен термин индекстрон [13] — термин, который используется только как название устройства индексации образов, но не как название метода индексации образов. Отличие индекстрона от поисковиков состоит в том, что вместо инвертирования текстовых документов инвертируются числовые данные. Специфика инвертирования числовых образов рассмотрена в разделе 3.

Обратим внимание, что подход, использованный в [10], имеет общие моменты с предложенным позднее методом TF-IDF [14], разработанным для поиска документов путем вычисления обратных частот слов документов. Метод

TF-IDF используется в поисковых системах, где вычисляются частоты имен документов в полностью инвертированных файлах. Вместе с тем текстовые поисковые системы не будут работать с числовыми данными, поскольку документная частота слова может измениться в зависимости от имеющегося в данный момент набора документов, но и само ключевое слово может полностью исчезнуть даже из-за небольшого шума.

Отметим, что индексный подход к решению задач распознавания образов практически не используется в системах машинного обучения, где большинство методов используют итеративное обучение, градиентный спуск и значительное количество адаптивных коэффициентов, что, собственно, и приводит к медленному обучению. В то же время мозг человека может запомнить новые зрительные образы с одного взгляда. В настоящей статье для обеспечения практически моментального обучения при работе с большими объемами данных используется не метод итеративного обучения, а метод индексации числовых данных. Предыдущие работы по индексации не текстовых, а числовых данных можно найти в [15], где такие быстрые методы использованы для распознавания изображений в кинофильмах. Но в [15] распознавание зашумленных числовых образов, представляющих изображения, сводится к конвертации числовых образов в текстовую форму с дальнейшим использованием уже стандартных методов для поиска текстовой информации. Подход, использованный в данной статье, заключается в применении метода индексного распознавания [10, 11], который делает возможным индексирование зашумленных числовых образов без их промежуточного преобразования в текстовую форму. Отметим также, что целью данной статьи является не разработка нового метода машинного обучения, а использование индексного метода распознавания по новому назначению, а именно для практически мгновенного обучения при работе с большими базами биологических данных.

2. Исходные данные

В эксперименте для сравнения эффективности обучения и классификации данных нейронной сетью и индекстроном был использован набор данных Gene expression cancer RNA-Seq [16]. Этот набор данных содержит 801 строчку, каждая из которых содержит 20 531 число с плавающей точкой (см. табл. 1). Каждое число с координатами профиль/ген — это уровень активности соответствующего гена, измеряемой в условных единицах от 0 (нулевая

Таблица 1. Активность генов пациентов

Профиль	Ген 0	Ген 1	Ген 2	...	Ген 20 529	Ген 20 530	Класс (диагноз)
0	0,0	2,017	3,266	...	5,287	0,0	3
1	0,0	0,592	1,588	...	2,094	0,0	1
...	...	...	...	...	...	...	...
800	0,0	2,325	3,806	...	4,551	0,08	5

активность или отсутствие гена) до максимальной активности 15. При этом числа каждой n-й строчки представляют уровни активности соответствующих генов n-го пациента. Диагнозы всех пациентов в наборе данных [15] известны и были определены другими клиническими методами. В эксперименте использовали 401 нечетную строку для обучения, в процессе которого системе сообщался диагноз, связанный с соответствующей строчкой. В процессе тестирования использовались 400 четных строк, которые классифицировались путем их отнесения к одному из пяти классов, а найденный класс сопоставлялся с известным диагнозом соответствующего пациента.

3. Формальная постановка задачи распознавания и метод ее решения

В настоящей статье индексный метод распознавания, предложенный и рассмотренный в [10] и усовершенствованный в [11] и в других работах, используется по новому назначению, а именно для диагностирования заболеваний путем классификации профилей экспрессии генов. Ниже рассмотрены особенности применения этого метода в данной задаче. Пусть все переменные являются целыми числами и пусть даны N образов, где каждый образ представлен K-мерным вектором признаков

$$\mathbf{x_n} = (x_{n,1}, \ldots, x_{n,k}, \ldots, x_{n,K}), \quad n = 1, \ldots, N. \tag{1}$$

Здесь расстояние Чебышева между любыми двумя образами $\mathbf{x_p}$ и $\mathbf{x_q}$ $(p, q \leqslant N)$ больше некоторого заданного числа R , т.е., $|\mathbf{x_p} - \mathbf{x_q}| > R$, а переменная $x_{n,k}$ $(0 \leqslant x_{n,k} < X)$ является значением k-го признака вектора, представляющего n-й образ. Неравенство $|\mathbf{x_p} - \mathbf{x_q}| > R$ означает, что для векторов $\mathbf{x_p}$, $\mathbf{x_q}$ существует хотя бы одно измерение k, такое что $|x_{p,k} - x_{q,k}| > R$. В этом случае каждый n-й вектор представляет n-й класс, к которому относятся все векторы $\mathbf{x}$, такие что $|\mathbf{x} - \mathbf{x_n}| \leqslant R$. Величина R определяет размер класса и называется радиусом обобщения.

Задача классификации. Для заданного вектора $\mathbf{x} = (x_1, \ldots, x_k)$ найти класс n, такой что $|\mathbf{x} - \mathbf{x_n}| \leqslant R$. Если такой класс не существует, то список (1) дополняется новым вектором $\mathbf{x_{N+1}} = \mathbf{x}$, а число классов увеличивается на 1.

Эту задачу, очевидно, можно решать путем сравнения заданного вектора $\mathbf{x}$ с векторами, представляющими классы, т.е. путем полного перебора классов. Однако решение задачи классификации с помощью метода обратных образов позволяет значительно ускорить поиск, особенно при большом числе классов.

Рассмотрим решение задачи классификации с помощью обратных образов. При этом методе неизвестный образ $\mathbf{x}$ относится к классу m такому, что

$$m : H_R(m|\mathbf{x}) = \max_{n=1}^{N} H_R(n|\mathbf{x}).$$

Здесь $H_R(m|\mathbf{x})$ — это гистограмма имен классов, содержащихся в обратных образах признаков вектора $\mathbf{x}$. Таким образом, $H_R(m|\mathbf{x})$ — это условная гистограмма классов, поскольку она зависит от вектора $\mathbf{x}$.

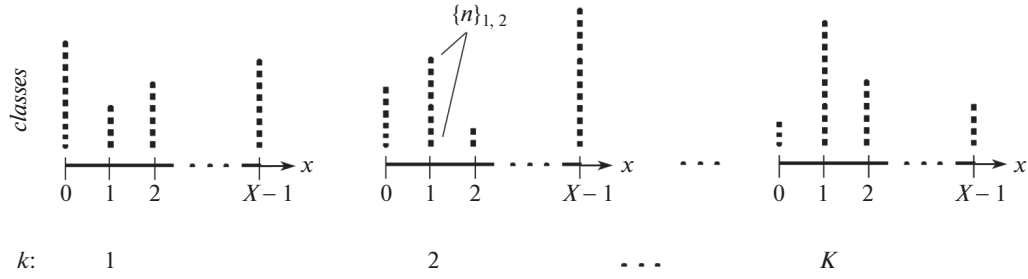

Рис. 2. Обратные образы показаны как K групп колонок.

Заметим, что если радиус обобщения R равен 0, то число классов будет равно числу N векторов списка (1). Если число реально существующих внешне заданных классов меньше N, то используется табличная функция $class(n)$, $n = 1, \ldots, N$, устанавливающая соответствие между классами списка (1) и внешними классами, которые при обучении с учителем всегда известны. Обратные образы признаков определяются в [10, 11] как некоторые множества имен классов. Такие множества индексируются с помощью двухмерных индексов. Каждый двухмерный индекс определяется парой чисел (x, k), где x — значение признака и k — номер измерения. Множество имен $\{n\}$ классов в левой части равенства (2) является обратным образом признака x из измерения k.

$$\{n\}_{x,k} = \{n : x_{n,k} = x\}, \quad k = 1, \ldots, K, \quad x = 0, 1, \ldots, X-1. \tag{2}$$

Понятие обратных образов проиллюстрировано графически с помощью рис. 2, где они представлены колонками точек, высоты которых равны мощностям соответствующих обратных образов. Как отмечено выше, элементы колонок — это имена классов. Поскольку обратные множества имеют двухмерные индексы, то все колонки разбиты на K групп, где каждая группа k может содержать максимум X колонок, а каждая колонка — максимум N классов. Здесь X — это диапазон значений признаков. Для решения задачи классификации с помощью обратных образов остается найти гистограмму классов $H_R(n|\mathbf{x})$. Пусть входной образ представлен вектором $\mathbf{x} = (x_1, \ldots, x_k)$. Тогда гистограмма классов, содержащихся в обратных образах, может быть найдена с помощью следующего алгоритма:

$$\forall n \in \{n\}_{x(k)+r,k}, \quad H_R(n|\mathbf{x}) = H_R(n|\mathbf{x}) + 1. \tag{3}$$

Здесь $r = -R, \ldots, -1, 0, 1, \ldots, R,\ k = 1, \ldots, K,\ x(k+r), k = x_{k+r,k}$. Таким образом, представленный выше критерий классификации — это гистограмма классов, положение максимума которой требуется найти. Для понимания алгоритма необходимо использование понятия “обратные образы”. Оно определяется выражением (2), снабженным необходимыми комментариями. Обратные образы, содержащие множества классов образов, служат входной информацией цикла (3). Этот цикл гистограммирует классы, выдавая на выходе

Таблица 2. Нормализованные данные

Профиль	Ген 0	Ген 1	Ген 2	⋯	Ген 20 529	Ген 20 530	Класс (диагноз)
0	0	82	137	...	112	0	3
1	0	24	66	...	36	0	1
...	...	...	...	...	...	...	...
800	0	95	160	...	95	0	5

искомую гистограмму, положение максимума которой соответствует искомому классу. В терминах программирования — это цикл по макроколонкам r, измерениям k и классам n, находящимся в обратных образах.

При обучении индекстрона если максимум гистограммы при входе $\mathbf{x}$ меньше выбранного некоторого порога, то число классов увеличивается на единицу, $N = N + 1$, и это новое число записывается в соответствующие обратные образы

$$\{n\}_{x(k),k} = \{n\}_{x(k),k} \bigcup N, \qquad k = 1, \ldots, K.$$

Начальным условием всегда служит значение $N = 0$. Однако гистограмма классов (3) не вычисляется, если в процессе обучения использовано нулевое значение радиуса обобщения $R = 0$. В то же время при каждой итерации число классов увеличивается на единицу. Графическая иллюстрация данного алгоритма приведена в [17].

При распознавании если максимум гистограммы при входе $\mathbf{x}$ меньше некоторого порога, то входной образ остается нераспознанным.

Заметим, что максимум гистограммы равен размерности K образа $\mathbf{x}$. Можно показать, что это следует из следующих свойств обратных образов: колонки каждой группы

- содержат строго N разных имен классов: $\sum\limits_{x=0}^{X-1} |\{n\}_{x,k}| = N$,
- не пересекаются: $\{n\}_{x,k} \cap \{n\}_{y,k} = \varnothing$.

При обучении и распознавании образа значение его признаков x_k, $k = 1, \ldots, K$, используется в качестве адреса колонки, расположенной в k-м измерении. При этом реальные значения x признаков приводятся к целочисленному диапазону [0–255]:

$$x = 255(x - x_{\min})/(x_{\max} - x_{\min}).$$

Опыт показывает, что такая дискретизация непрерывных значений, как правило, не снижает точности классификации. В табл. 2 приведен пример нормализации данных табл. 1.

Таким образом, признаки каждого поступившего K-мерного образа выделяют K колонок — по одной колонке в каждой группе. Проблема состоит в том, что, как правило, для числовых признаков пересечение колонок оказывается пустым, так как погрешности измерения и деформации образа искажают

адреса колонок. Поэтому вводятся макроколонки, т.е. наряду с колонкой с адресом x рассматриваются соседние колонки с адресами $x \in [-R, R]$. Возможность использования макроколонок вытекает из приведенных выше свойств обратных образов. Как уже отмечено выше, пересечения колонок находятся путем вычисления гистограммы, определяющей частоты появления классов. При этом входной образ относится к классу, встречающемуся наиболее часто.

4. Программно-аппаратная реализация индекстрона

Распараллеленная версия алгоритма была написана на языке Python и запущена на видеокарте Nvidia GeForce GTX 1660 Super. Эта видеокарта имеет 44 мультипроцессора и 1408 ядер, что позволило распараллелить один процесс на 1408 параллельных подпроцессов. При этом обучение в рассматриваемой задаче может быть распараллелено на 20 531 процессов по числу всех генов профиля, а максимально возможное распараллеливание — это $20\,531 * (2R + 1)$ процессов, так как для каждого гена возможно параллельно проверить все значения в пределах радиуса обобщения.

Распараллеливание было произведено с помощью модуля *cuda*, входящего в библиотеку *numba*. Для обработки потоков (*threads*) мультипроцессорами видеокарты потоки были разделены на блоки по формуле:

$$blocks\_per\_grid = \\ = (number\_of\_iterations| + (threads\_per\_block - 1))//threads\_per\_block,$$

где $number\_of\_iterations$ — это количество параллельных итераций.

Перед выполнением программы массивы исходных данных были скопированы из оперативной памяти в память видеокарты с помощью функции *cuda.to_device*(), и по завершении выполнения программы результаты выполнения были скопированы обратно в оперативную память с помощью функции *device_array.copy_to_host*().

Для распараллеливания функций записи и чтения использовался декоратор. Внутри распараллеливаемых функций вызывалась функция, возвращающая индекс потока, на котором должна осуществляться соответствующая параллельная итерация.

5. Результаты

Точность и количество операций при обучении индекстрона и нейронных сетей в задаче диагностики заболеваний по данным, полученным от 801 пациента, представлены в табл. 3.

Таблица 3. Сравнительные результаты

	Нейронная сеть	Индекстрон
Точность (%)	99,5	99,75
Количество операций	46 млрд операций сложения/умножения	8 млн операций записи в память

Время обучения индекстрона на одноядерном ноутбуке, 1,6 Ггц, составило 0,43 сек. Время обучения четырехслойной нейронной сети на аналогичной аппаратуре увеличивается в соответствии с ростом количества операций в $46 * 109/8 * 106 = 5750$ раз. Заметим, что архитектура индекстрона идеально подходит для распараллеливания, что позволило достичь практически мгновенного обучения за 75 миллисекунд в рассматриваемой задаче. Радиусы обобщения R в режимах обучения и распознавания составляют соответственно 0 и 84, т.е. 33% от диапазона (0–255) изменения признака.

6. Заключение

1. Простота алгоритма индекстрона, в результате которой обучение каждому образу требует только записи в память K целых чисел, а классификация образа имеет порядок сложности $O(hK)$ операций чтения и суммирования, позволяет создавать большие базы экспрессии генов для диагностирования самых разнообразных заболеваний (h — это средняя высота колонок, где $h = N/X$).

2. При широкой доступности оборудования для нахождения профилей генной экспрессии становится возможным создание поисковой диагностической системы типа Google для массового пользования, в которой запросы имеют не текстовую, а численную форму.

3. Проведенные эксперименты подтверждают наличие такой возможности, поскольку быстрое обучение при пополнении базы новыми данными даже на программном уровне занимает всего $0{,}43/800 = 0{,}00054$ сек, а точность диагностирования пяти типов онкологических заболеваний составила 99,75%. Однако диагностика многих других заболеваний потребует создания большого числа различных баз данных.

СПИСОК ЛИТЕРАТУРЫ

1. *Khan J., Wei J., Ringner M. et al.* Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. // Nat Med. (2001). June 7(6): 673–9. https://doi.org/10.1038/89044
2. *Kumar A., Halder A.* Greedy fussy vaguely quantified rough approach for cancer relevant gene selection from gene expression data // Soft Comput. 2022. V. 26. P. 13567–13581. https://doi.org/10.1007/s00500-022-07312-4
3. *Houssein E., Hassan H., Mustafa al-sayed et. al.* Gene Selection for Microarray Cancer Classification based on Manta Rays Foraging Optimization and Support Vector Machines // Arabian Journal for Science and Engineering. 2022. V. 47. P. 2555–2572. https://doi.org/10/1007/s13369-021-06101-8
4. *Zheng Y., Sun Y., Kuai Y. et al.* Gene expression profiling for the diagnosis of multiple primary malignant tumors // Cancer Cell Int. 2021. V. 21, Article no. 47. https://doi.org/10.1186/s12935-021-01748-8
5. *Ye Q., Wang Q., Qi P. et. al.* Development and validation of a 90-gene real-time PCR assay for tumor origin identification // Symposium MXW, 2018.

6. *Joshi P., Dhar R. EpICC:* A Bayesian neural network model with uncertainty correction for a more accurate classification of cancer // Sci. Rep 12, (2022). Article no. 14628. https://doi.org/10.1038/s41598-022-18874-6
7. *Steiling K., Christenson S.* Tools for genetics and genomics: Gene expression profiling // UpToDate.(2021). Retrieved from https://www.uptodate.com/contents/tools-for-genetics-and-genomics-gene-expression-profiling
8. СПбГУ Научный парк. Система высокопроизводительного полногеномного секвенирования, 2023. https://researchpark.spbu.ru/equipment-biobank-rus/equipment-biobank-genom-rus/equipment-biobank-ngsseq-rus/1762-biobank-hiseq-2500-sequencing-system-rus
9. IBM. What are neural networks? // Retrieved from https://www.ibm.com/cloud/learn/neural-networks
10. *Mikhailov A., Pok Y.M.* Artificial Neural Cortex // Smart Engineer. Syst. Design. 2001. V. 11. ASME PRESS. N. Y. P. 113–120.
11. *Mikhailov A., Karavay M.* Pattern Inversion as a Pattern Recognition Method for Machine Learning // Cornell University. 2021. Retrieved from https://arxiv.org/abs/2108.10242
12. *Brin S., Page L.* The Anatomy of a large-scale hypertextual web search engine // Comput. Networks ISDN Syst. 1998. V. 30. Iss. 1–7. Stanford University, Stanford, CA, 94305, USA. Retrieved from https://doi.org/10.1016/S069-7552(98)00110-X
13. *Mikhailov A.* Indextron // Artificial Neural Networks in Engineering Conf. (ANNIE 1998), St. Louis, Missouri, Nov. 4–7, 1998. Proceedings Vol. 8: ANNIE 1998, Publisher: ASME Press, ISBN: 0791800822
14. *Jones K.* A statistical interpretation of term specificity and its application in retrieval // J. Document.: MCB Univer.: MCB Univer. Press, 2004. V. 60. No. 5. P. 493–502. ISSN 0022-0418
15. *Sivic J., Zisserman A.* Efficient visual search of videos cast as text retrieval // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2009. V. 31. Issue 4. https://doi.org/10.1109/TPAMI.2008.111
16. UCI. Machine learning repository // Retrieved from https://archive.ics.uci.edu/ml/datasets/gene+expression+cancer+RNA-Seq
17. *Mikhailov A., Karavay M.* Indextron // Proceedings of the 10th International Conference on Pattern Recognition Application and Methods, 4–6 Feb 2021, Vienna, V.1-978-989-758-486-2. P. 143–149. https://doi.org/10.5220/0010180301430149

Статья представлена к публикации членом редколлегии О.Н. Граничиным.

Поступила в редакцию 19.07.2022

После доработки 21.03.2023

Принята к публикации 30.03.2023