



ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР
«ИНФОРМАТИКА И УПРАВЛЕНИЕ»
РОССИЙСКОЙ АКАДЕМИИ НАУК

ОТДЕЛЕНИЕ МАТЕМАТИЧЕСКИХ НАУК
РОССИЙСКОЙ АКАДЕМИИ НАУК

НАЦИОНАЛЬНЫЙ КОМИТЕТ
РОССИЙСКОЙ АКАДЕМИИ НАУК
ПО РАСПОЗНАВАНИЮ ОБРАЗОВ И АНАЛИЗУ ИЗОБРАЖЕНИЙ

Математические методы распознавания образов

21-я Всероссийская конференция с международным
участием

Москва, 2023

УДК 004.85+004.89+004.93+519.2+519.25+519.7

ББК 22.1:32.973.26-018.2

И 73

Математические методы распознавания образов: Тезисы докладов 21-й Всероссийской конференции с международным участием, г. Москва, 2023 г. — М.: Российская академия наук, 2023. — 256 с.

ISBN 978-5-907645-52-3

В сборнике представлены тезисы докладов 21-й Всероссийской конференции с международным участием «Математические методы распознавания образов», проводимой Федеральным исследовательским центром «Информатика и управление» Российской академии наук, Отделением математических наук Российской академии наук и Национальным комитетом Российской академии наук по распознаванию образов и анализу изображений.

Конференция ММРО с 1983 года является ведущим форумом исследователей и профессионалов, работающих в области интеллектуального анализа данных, площадкой для обсуждения, распространения и продвижения передовых идей, достижений и разработок. Конференция призвана способствовать обмену идеями между представителями науки и индустрии. Конференция организована представителями российской научной школы машинного обучения и нацелена на расширение взаимодействия между российскими и зарубежными исследователями и представителями высокотехнологичного ИТ бизнеса.

Сайт конференции — <https://mmro.ru>.

ISBN 978-5-907645-52-3

© Авторы докладов, 2023

Оргкомитет

- Председатель:** Соколов Игорь Анатольевич, *акад. РАН*
Заместитель председателя: Матвеев Иван Алексеевич, *д. т. н.*
Ученый секретарь: Докукин Александр Александрович, *к. ф.-м. н.*
Члены комитета: Грабовой Андрей Валериевич, *к. ф.-м. н.*
Инякин Андрей Сергеевич, *к. ф.-м. н.*
Рейер Иван Александрович, *к. т. н.*
Чехович Юрий Викторович, *к. ф.-м. н.*
Борисова Татьяна Игоревна
Громов Андрей Николаевич

Программный комитет

- Председатель:** Соколов Игорь Анатольевич, *акад. РАН*
Заместитель председателя: Воронцов Константин Вячеславович, *д. ф.-м. н.*
Члены комитета: Семенов Алексей Львович, *акад. РАН*
Сойфер Виктор Александрович, *акад. РАН*
Шананин Александр Алексеевич, *акад. РАН*
Петров Игорь Борисович, *чл.-корр. РАН*
Хачай Михаил Юрьевич, *чл.-корр. РАН*
Гимади Эдуард Хайрутдинович, *д. ф.-м. н.*
Горнов Александр Юрьевич, *д. т. н.*
Громова Ольга Алексеевна, *д. м. н.*
Двоенко Сергей Данилович, *д. ф.-м. н.*
Дюкова Елена Всеволодовна, *д. ф.-м. н.*
Краснопрошин Виктор Владимирович, *д. т. н.*
Лазарев Александр Алексеевич, *д. ф.-м. н.*
Местецкий Леонид Моисеевич, *д. т. н.*
Пытьев Юрий Петрович, *д. ф.-м. н.*
Сенько Олег Валентинович, *д. ф.-м. н.*
Стрижов Вадим Викторович, *д. ф.-м. н.*
Чуличков Алексей Иванович, *д. ф.-м. н.*
Ватолин Дмитрий Сергеевич, *к. ф.-м. н.*
Гнеушев Александр Николаевич, *к. ф.-м. н.*
Копылов Андрей Валерьевич, *к. т. н.*
Середин Олег Сергеевич, *к. ф.-м. н.*

Рецензенты

Адуенко А. А.
Анциперов В. Е.
Бахтеев О. Ю.
Визильтер Ю. В.
Володин С. Е.
Воронцов К. В.
Гасников А. В.
Гнеушев А. Н.
Гончаров А. В.
Грабовой А. В.
Двоенко С. Д.
Докукин А. А.
Драгунов Н. А.
Ивахненко А. А.

Инякин А. С.
Исаченко Р. Г.
Катруца А. М.
Копылов А. В.
Кушнир О. А.
Ланге М. М.
Майсурадзе А. И.
Масляков Г. О.
Матвеев И. А.
Местецкий Л. М.
Мурашов Д. М.
Неделько В. М.
Нейчев Р. Г.
Новик В. П.

Одиноких Г. А.
Панов М. Е.
Потапенко А. А.
Рейер И. А.
Сафонов И. В.
Сенько О. В.
Середин О. С.
Стрижов В. В.
Сулимова В. В.
Торшин И. Ю.
Турдаков Д. Ю.
Хачай М. Ю.
Чуличков А. И.
Янина А. О.

Краткое оглавление

| | |
|---|-----|
| Интеллектуальный анализ данных | 6 |
| Машинное обучение | 18 |
| Прогнозирование и рекомендательные системы | 51 |
| Аналитика больших данных | 63 |
| Нейронные сети и глубокое обучение | 68 |
| Методы оптимизации для интеллектуального анализа данных | 87 |
| Вычислительная сложность и приближенные методы | 106 |
| Обработка и анализ изображений, компьютерное зрение | 111 |
| Обработка и анализ сигналов | 148 |
| Информационный поиск и анализ текстов | 159 |
| Индустриальные приложения науки о данных | 175 |
| Анализ биомедицинских данных, биоинформатика | 229 |

Подходы к выбору содержательных аспектов при анализе сложных выборок

*Рязанов Владимир Васильевич*¹

rvcas@mail.ru

Виноградов Александр Петрович^{1*}

vngccas@mail.ru

¹Москва, Федеральный исследовательский центр «Информатика и Управление»
Российской академии наук

Любая цель анализа и структурирования сложных выборок неизменно сводится к выявлению скрытых закономерностей, управляющих поведением или выборки в целом, или ее частей. Как известно, общее понятие закономерности является сложным и не обладает устоявшимся статусом, поэтому каждое частное определение важно. В последние годы в ФИЦ ИУ РАН был выполнен ряд исследований, связанных с поиском скрытых закономерностей в прикладных данных различной природы [1–6].

Центральным объектом исследований было понятие обобщенного прецедента (ОП). ОП представляет собой цифровую модель закономерности, заданную в терминах ограниченного числа наблюдаемых или комбинированных переменных. Строились параметризованные гипотезы о наличии в данных того или иного варианта закономерности, статистическая значимость и типичные параметры которого оценивались апостериори по результатам анализа вторичного распределения ОП в пространствах пониженных размерностей, с использованием методов обработки изображений (ОИ) и анализа сцен (АС). ОП такого вида являются конкретизированным представлением того или иного аспекта данных.

В ходе исследований выявились и другие возможности. Прямое задание числовой зависимости является сужением поиска, обнаруживаемые закономерности не являются в полном смысле слова новыми или ранее неизвестными. В то же время, накопленное в области ОИ и АС обилие методов анализа, которые эффективны как раз в малых размерностях, позволяет обнаруживать вполне новые явления при анализе вторичного распределения даже в специализированных пространствах ОП. Примеры такого рода описаны в [4], но обнаружение неизвестных закономерностей оказалось здесь лишь побочным продуктом.

Поэтому в настоящее время исследуется некоторая ослабленная версия ОП, где не требуется явной формулировки числовых связей. Так, работоспособность подхода с ОП показана в предположениях, регламентирующих лишь класс функций для искомой зависимости. В частности, исследовалась адекватность предположения о гладкости закономерностей естественного происхождения. Ослабленные таким образом ОП ограничивают варианты решений с неприемлемой суммарной кривизной. Эффективность двух названных конкретизаций ОП может возрастать, если при поисках используются их суперпозиции с более традиционными алгоритмами структурирования выборок, в том числе, иерархическими: по размерностям, кластерам, уровням или слоям представ-

лений [7–10], поскольку локально может наблюдаться выраженное проявление искомой закономерности.

Пусть подмножество $X \subset R^N$ — выборка оцифрованных сложных данных. Исследуется случай, когда эксперт указывает среди наблюдаемых $x = (x_1, \dots, x_N)$, $x \in R^N$, ограниченный набор тех из них $y = (y_1, \dots, y_M)$, $M < N$, которые оказывают, по его мнению, существенное влияние на поведение выборки X в целом. Если задача эксперта заключается в изучении взаимного влияния каких-то общепринятых параметров или характеристик F_1, \dots, F_M , которые представимы через наблюдаемые величины: $y_m = F_m(x_1, \dots, x_N)$, $m = \overline{1, M}$, то тем самым формируется пространство Y для весьма простых ОП в виде $y = (y_1, \dots, y_M)$, $y \in R^M$, выше имеет место тривиальный частный случай $y_n = y_{n(m)}$, $Y \subset R^M$.

Цель состоит в выделении числовых взаимосвязей элементов набора $y = (y_1, \dots, y_M)$ в явном виде, где побочные влияния редуцированы. Функции $y_m = F_m(x_1, \dots, x_N)$, $m = \overline{1, M}$, задают слоение пространства R^N , и набор векторов $\partial F_m(x_1, \dots, x_N)$, $m = \overline{1, M}$, определяет базис касательного пространства \mathbf{F}_x слоя в точке x . Наилучший случай реализуется, когда объекты $x = (x_1, \dots, x_N)$, $x \in X$, концентрируются вблизи конкретного M -мерного слоя, который в окрестности точки x адекватно проецируется на \mathbf{F}_x , что может соответствовать целям эксперта.

Здесь мы не даем готовых решений, а лишь обращаем внимание на возможности использования структурных и иерархических методов для выделения окрестности нужного слоя при прогрессивной детализации данных, а также для выделения нужного аспекта, в частности, подмножества индексов из M . В общем случае, при дискретизации представления, искомые S уравнений, $S < M$, могут задавать такую сетку ячеек на гиперплоскости \mathbf{F}_x , что почти в каждой из них проявления выборки представительны.

Пусть $M = 2$, тогда $S = 1$, и отыскивается зависимость в виде гладкой монотонной функции f . Как известно, сумма $\sum_i^I p(i, f(i)) \log p(i, f(i))$, где $p(i, f(i))$ — эмпирическая плотность ОП в ячейке $(i, f(i))$, во многих случаях отражает степень информативности факта обнаружения объекта f . Для кривизны аналогом служит сумма $\sum_i^I c(i, f(i))$, где $c(i, f(i))$ — кривизна графика функции f в точке $(i, f(i))$.

При конкуренции этих двух критериев, приоритеты для них по важности и естественности может выбрать только пользователь или эксперт, исходя из той или иной конкретной постановки задачи структуризации или поиска.

Разумеется, вариантов конкурирующих критериев такого рода существует множество. Несложно предъявить более размерные и содержательные версии данной ситуации для различных уровней детализации. Так, для некоторых вариантов представления как эмпирической плотности, так и параметра кривизны, построены не просто иерархии, а шкалированные пространства [10, 11]. Так или иначе, при некотором S параметр суммарной гладкости (кривизны, эла-

стичности) для конкретной $(M - S)$ -мерной гиперповерхности в составе x -слоя может оказаться решающим при отсеве вариантов, в том числе, более значимых статистически, чем финальная сглаженная версия.

- [1] *Lapin Yu. P., Nelyubina E. A., Ryazanov V. V., Vinogradov A. P.* Shape of Basic Clusters: Using Analogues of Hough Transform in Higher Dimensions // Pattern Recognition and Image Analysis. — Pleiades Publishing, 2018. — Vol. 28, No. 4. — P. 653–658.
- [2] *Naumov V. A., Nelyubina E. A., Ryazanov V. V., Vinogradov A. P.* Analysis and Prediction of Hydrological Series Based on Generalized Precedents // Book of abstracts 12-th Int. Conf. Intelligent Data Processing (IDP-12), Gaeta, Italy, 2018. — P. 178–179.
- [3] *Nelyubina E. A., Ryazanov V. V., Vinogradov A. P.* Shape of Basic Clusters: Finding Coherent ELR-2s via Hough-type Transform // Proceedings of ICPRAI 2018 — International Conference on Pattern Recognition and Artificial Intelligence, Montréal, Canada: CENPARMI, Concordia University, 2018. — P. 702–706.
- [4] *Zhuravlev Yu. I., Ryazanov V. V., Sen'ko O. V., Dokukin A. A., Vinogradov A. P., Nelyubina E. A., Stefanovskii D. V.* Using Hough-Like Transforms for Extracting Relevant Regularities from Big Applied Data // Pattern Recognition and Image Analysis. — Pleiades Publishing, 2021. — Vol. 31, No. 4. — P. 699–709.
- [5] *Ryazanov V. V., Vinogradov A. P.* Analogues of Image Analysis Tools in the Problems of Finding Latent Regularities in Big Applied Data // Pattern Recognition and Image Analysis, Pleiades Publishing, 2022. — Vol. 32, No. 3. — P. 639–644.
- [6] *Nelyubina E. A., Ryazanov V. V., Vinogradov A. P.* Analogs of Image Analysis Tools in the Search for Latent Regularities // Lecture Notes in Computer Science 13644, Pattern Recognition, Computer Vision, and Image Processing. ICPR 2022, International Workshops and Challenges, Proceedings, Part II, 2023. — P. 529–541.
- [7] *Davies E. R., Turk M. A., eds.* Advanced Methods and Deep Learning in Computer Vision. — AP, 2022. — 569 p.
- [8] *Nixon M. S., Aguado A. S.* Feature Extraction for Image Processing and Computer Vision. — AP, 2020. — 626 p.
- [9] *Сойфер В. А., ред.* Методы компьютерной обработки изображений. — М.: ФИЗМАТЛИТ, 2003. — 784 с.
- [10] *de Berg M., van Kreveld M., Overmars M., Schwarzkopf O.* Computational Geometry // 2nd revised ed., Chapter 14: Quadrees. — P. 291–306.
- [11] *Sai A. C., Tamilarasan M., Arjun P.* A Study on Curvature Scale Space // Int. J. Innovative Research in Computer and Communication Engineering. — 2014. — Vol. 2, Spec. Iss. 3. — P. 168–174.

О метрических свойствах множеств решений некоторых задач анализа целочисленных данных

Дюкова Елена Всеволодовна¹

edjukova@mail.ru

Дюкова Анастасия Петровна^{1*}

anastasia.d.95@gmail.com

¹Москва, ФИЦ ИУ РАН

Представляемый доклад посвящен обзору исследований, касающихся изучения метрических (количественных) свойств множеств решений вычислительно сложных задач анализа целочисленных данных. Рассматриваемые задачи возникают в ряде прикладных областей, в частности при поиске ассоциативных правил в базах данных и на этапе обучения логических процедур классификации по прецедентам.

Введем основные понятия. Исследуется множество объектов M . Известно, что каждый объект из M может быть представлен в виде числового вектора, полученного на основе наблюдения или измерения ряда его характеристик, называемых атрибутами. Предполагается, что каждый атрибут имеет ограниченное множество допустимых значений, которые кодируются целыми числами.

Пусть $X = \{x_1, \dots, x_n\}$ — множество всех атрибутов; $H = \{x_{j_1}, \dots, x_{j_r}\}$ — набор из r различных атрибутов; $\sigma = (\sigma_1, \dots, \sigma_r)$ — набор, в котором σ_i — допустимое значение атрибута x_{j_i} , $i = \overline{1, r}$. Пара (σ, H) называется элементарным фрагментом (ЭФ) длины r . Объект $S = (a_1, \dots, a_n)$ из M (здесь a_j , $j \in \{1, 2, \dots, n\}$, — значение атрибута x_j для объекта S) содержит ЭФ (σ, H) , если $a_{j_i} = \sigma_i$ при $i = \overline{1, r}$.

Дана некоторая совокупность объектов D из M , и задано число p , $1 \leq p \leq |D|$, где $|D|$ — число объектов в D . ЭФ (σ, H) называется (p, D) -частым, если не менее p объектов из D содержат ЭФ (σ, H) . ЭФ (σ, H) длины r называется правильным, если (σ, H) — (r, D) -частый.

ЭФ (σ, H) называется нечастым, если ни один объект из D не содержит (σ, H) . Нечастый ЭФ (σ, H) называется минимальным нечастым, если любой ЭФ (σ', H') такой, что $\sigma' \subset \sigma$, $H' \subset H$, является $(1, D)$ -частым.

При поиске ассоциативных правил выборку D называют базой данных, а каждый объект базы D — транзакцией. Ассоциативное правило устанавливает зависимость между двумя частыми ЭФ, согласно которой один частый ЭФ (посылка) с некоторой «достоверностью» влечет другой частый ЭФ (следствие). При этом посылка и следствие порождаются одним частым ЭФ [1].

Логическая классификация целочисленных данных предполагает наличие нескольких непересекающихся выборок D_1, \dots, D_l , $l \geq 2$, объектов из M , каждая из которых представляет отдельный класс объектов. Объекты, содержащиеся в этих выборках, называются прецедентами, а атрибуты из X — признаками. На этапе обучения для каждой выборки D , $D \in \{D_1, \dots, D_l\}$, ищутся такие $(1, D)$ -частые ЭФ, которые одновременно являются нечастыми в \overline{D} , $\overline{D} = \left(\bigcup_{j=1}^l D_j\right) \setminus D$. Найденные ЭФ позволяют различать прецеденты из разных

классов и называются логическими закономерностями или представительными элементарными классификаторами [2, 3]. По их наличию в распознаваемом объекте решается вопрос о его классификации.

Могут накладываться некоторые дополнительные условия на вид искомого ЭФ (в зависимости от рассматриваемой модели классификатора). Например, ищутся так называемые тупиковые представительные элементарные классификаторы, т. е. такие $(1, D)$ -частые ЭФ, $D \in \{D_1, \dots, D_l\}$, которые являются минимальными нечастыми в \bar{D} . При этом возникает необходимость рассматривать труднорешаемую дискретную задачу построения тупиковых покрытий целочисленной матрицы [4], строками которой являются прецеденты из \bar{D} . В [5] предложена модель логического классификатора, базирующаяся на первоначальном поиске в выборке D , $D \in \{D_1, \dots, D_l\}$, правильных ЭФ и последующем отборе тех из них, которые являются нечастыми в \bar{D} . Эта модель демонстрирует существенное преимущество по скорости счета перед классической моделью, основанной на поиске в \bar{D} минимальных нечастых ЭФ, не уступая последней в качестве классификации.

Представляет интерес изучение метрических (количественных) свойств множества правильных ЭФ. Имеется в виду получение асимптотических оценок (при $n \rightarrow \infty$) типичного числа правильных ЭФ и типичной длины правильного ЭФ. В [5–7] требуемые оценки получены для двух следующих случаев: 1) $|D|$ существенно меньше n ; 2) $n \leq |D|$. Следует отметить, что аналогичные свойства множества минимальных нечастых ЭФ изучались ранее в связи с задачей построения асимптотически оптимальных алгоритмов перечисления тупиковых покрытий целочисленной матрицы (см. например, в [4]). Сравнение оценок числа правильных ЭФ с аналогичными оценками минимальных нечастых ЭФ свидетельствует о целесообразности (в плане сокращения временных затрат) применения методов поиска частых ЭФ для синтеза логических классификаторов и согласуется с полученными в [5] результатами экспериментов на случайных модельных данных.

- [1] *Aggarwa C., Jiawei H.* Frequent Pattern Mining. Springer, 2014. — 469 p.
- [2] *Журавлев Ю. И., Рязанов В. В., Сенько О. В.* Распознавание. Математические методы. Программная система. Практические применения. — М.: ФАЗИС, 2006. — 159 с.
- [3] *Дюкова Е. В., Песков Н. В.* Поиск информативных фрагментов описаний объектов в дискретных процедурах распознавания // Ж. вычисл. матем. и матем. физ.. — 2002. — Т. 42. № 5. — С. 741–753.
- [4] *Дюкова Е. В., Журавлев Ю. И.* Дискретный анализ признаков описаний в задачах распознавания большой размерности // Ж. вычисл. матем. и матем. физ.. — 2000. — Т. 40. № 8. — С. 1264–1278.
- [5] *Dragunov N., Djukova E., Djukova A.* Supervised Classification and Finding Frequent Elements in Data // VIII Intern. Conf. Information Technology and Nanotechnology (ITNT-2022). Samara, Russian Federation: IEEE, 2022. — P. 1–5.

-
- [6] Дюкова Е. В., Дюкова А. П. О сложности обучения логических процедур классификации // Информатика и ее применения. — 2022. — Т. 16. № 4. — С. 57–62.
- [7] Дюкова А. П., Дюкова Е. В. О числе решений некоторых специальных задач логического анализа целочисленных данных // Известия РАН. Теория и системы управления, 2023. — № 5. — С. 57–66.

Исследование процессов торможения пассажирских высокоскоростных поездов

Матковский Владислав Андреевич

vladislavmatkovskiy2015@mail.ru

Козлова Маргарита Геннадьевна

art-inf@mail.ru

Лукьяненко Владимир Андреевич

art-inf@yandex.ru

Симферополь, Крымский федеральный университет им. В. И. Вернадского

Особую значимость в моделировании процессов торможения имеет система аварийного торможения поездов, обеспечивающая безопасность пассажиров, экстренное торможение поезда при возникновении аварийной ситуации. Так как торможение зависит от многих факторов, возникают различные тормозные последовательности, оценка качества которых представляет актуальную математическую задачу.

В работе приведен подход, основанный на интеллектуализации обработки больших данных, характеризующих процесс торможения поездов метрополитена, и позволяющий уточнять разработку систем управления поездов в автоматическом режиме, прогнозировать качество и точность остановки. Для этого необходимо реализовать следующие этапы решения задачи интеллектуализированной обработки данных: определение модели торможения поезда, извлечение информации из данных (статистическая обработка, многокритериальное исследование, ранжирование) и кластеризация данных.

Математические модели, описывающие процесс торможения объекта (поезда метро), зависят от сложности компонент системы объекта. Статистические методы и методы кластерного анализа позволяют извлечь знания о процессе торможения из данных измерений, таких как координаты станций метро, на основе изучения гистограмм и применения алгоритмов кластеризации к координатам остановок поездов. Зная измеренные датчиками показания, можно восстановить математическую модель. В общем случае задача восстановления модели — это обратная задача (задача восстановления дифференциального уравнения по данным измерений). Она является некорректной. Для практических целей оценки качества торможения достаточно предположить, что такие модели описываются системой обыкновенных дифференциальных уравнений первого порядка [1]. Следовательно, по измеряемым данным можно находить статистические характеристики процесса торможения и наглядно оценивать качество торможения по набору этих статистик и сравнения сглаженных траекторий (расстояние — скорость).

Приведем алгоритм расчета основных показателей торможения, в результате которого могут быть получены характерные траектории торможений и значения параметров кинематик движения.

Входные данные: номер поезда $k \in K$; номер последовательности $i \in I_k$; отсчеты t_j^{ki} , j -го времени для k -го поезда i -й последовательности; $x_j^{ki} = x(t_j^{ki})$ —

абсцисса k -ого поезда i -й последовательности в момент времени t_j^{ki} ; $x_j^{ki} = v_j^{ki}$ — величина скорости в i -й последовательности k -ого поезда.

Выход: основные характеристики движения.

1) Вычислить:

$$1.1) \text{ ускорение } a_j^{ki} = \ddot{x}_j^{ki} = \dot{v}_j^{ki} = \frac{v_{j+1}^{ki} - v_j^{ki}}{t_{j+1}^{ki} - t_j^{ki}};$$

$$1.2) \text{ торможение в момент времени } t_j^{ki};$$

$$1.3) \text{ среднее торможение deceleration за период от начала торможения } t_0^{ki} \text{ до остановки } t_n^{ki}: d \equiv deceleration = \mathbf{a}^{ki} = \frac{v_n^{ki} - v_0^{ki}}{t_n^{ki} - t_0^{ki}};$$

$$1.4) \text{ равнодействующие силы: } g_j^{ki} = \dot{x}_j^{ki} \frac{\dot{x}_{j+1}^{ki} - \dot{x}_j^{ki}}{x_{j+1}^{ki} - x_j^{ki}};$$

$$1.5) \text{ фазовые траектории: } \frac{g_j^{ki}}{\dot{x}_j^{ki}} = \frac{\ddot{x}_j^{ki}}{\dot{x}_j^{ki}}.$$

2) На основании данных из последовательностей базы данных и шага 1 построить графики (в зависимости от времени) расстояния (м), скорости (м/с) и ускорения (торможения) (м/с²), траектории движения.

После разделения последовательностей на положительные (координата растет со временем) и отрицательные (координата убывает со временем), получены характерные графики траекторий положительного направления и отрицательного (см. рис. 1).

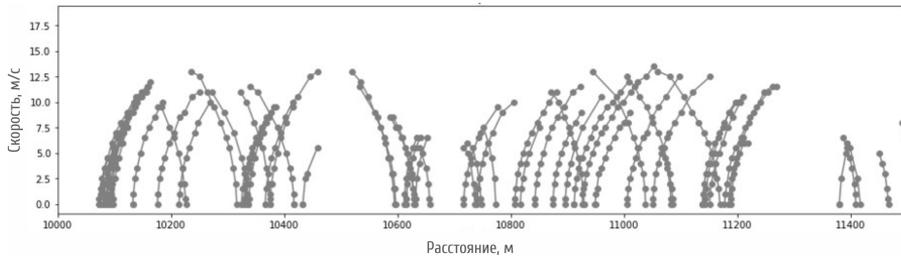


Рис. 1. График зависимости скорости движения от расстояния в положительном и отрицательном направлениях

Статистические данные движения поездов метро, полученные из них траекторные данные и визуализированные представления являются большими данными (Big Data, BD), из которых извлекаются знания о качестве торможения поездов (Data Mining, DM), а также представляется возможным определение класса торможения.

Интеллектуализированная система обработки таких данных должна сочетать статистические методы машинного обучения (Machine learning, ML) и ней-

ронные сети. Структура такой интеллектуализированной системы обработки больших данных (ИСОБД) представлена на рис. 2.



Рис. 2. Структура ИСОБД

Работа поддержана Министерством науки и высшего образования Российской Федерации, соглашение № 075-02-2023-1799.

- [1] Лукьяненко В. А., Португальская А. А., Абдулаева Д. М., Павлюков В. В. Моделирование процесса торможения поездов // Математика, информатика, компьютерные науки, моделирование, образование: сб. науч. тр. научно-практической конференции МИКМО-2018. — Симферополь: ИП Корниенко А. А., 2020.

Применение критериев проверки статистических гипотез для анализа параметров атмосферы

Кузин Кирилл Андреевич^{1*}

kirill.kuzin456@yandex.ru

Газарян Варвара Арамовна^{1,2}

varvaragazaryan@yandex.ru

*Чуличков Алексей Иванович*¹

achulichkov@gmail.com

Шапкина Наталья Евгеньевна^{1,3}

neshapkina@mail.ru

*Голубцов Петр Викторович*¹

pgolubtsov@gmail.com

¹Москва, МГУ имени М. В. Ломоносова, физический факультет

²Москва, Финансовый университет при правительстве РФ

³Москва, ИТПЭ РАН

Изменение климата — угроза для человечества, т. к. оно негативно влияет на экосистемы и экономическое развитие. Изучение различных метеорологических показателей, в особенности температуры воздуха приземного слоя атмосферы, актуально для анализа климатических изменений [1, 2].

В данной работе проводится статистический анализ временных рядов метеорологических показателей в Тверской области. Для более детального анализа динамики метеорологических показателей рассмотрены следующие параметры: среднесуточная, среднедневная, максимальная и минимальная суточные температуры воздуха; а также соответствующие параметры, связанные с концентрацией CO_2 .

Изучение характера климатических изменений, вызванных динамикой температуры воздуха и концентрации CO_2 , проводилось с помощью различных критериев проверки статистических гипотез. Для оценки изменения изучаемых параметров за длительный период наряду с годовыми промежутками времени рассматривались также сезонные: зимние, летние и весенние периоды этих лет. Проводилась также проверка гипотез об изменениях параметров атмосферы в пятилетние периоды. Так как некоторые критерии требуют независимости выборочных наблюдений, был проведен анализ зависимости значений параметров атмосферы в различные промежутки времени.

Для проверки гипотезы о том, что температурные параметры и параметры концентрации CO_2 являются независимыми в разные периоды наблюдений, применялись критерии Пирсона, Спирмена и Кенделла. В результате проверки гипотезы о независимости температурных параметров оказалось, что значения температуры воздуха в одни и те же сезонные периоды разных лет независимы. При сравнении временных рядов температуры за более длительные периоды: длиною в 1 год и 5 лет, наблюдается сильная связь данных при использовании всех трех критериев.

При проверке гипотезы о независимости концентрации CO_2 в разные периоды наблюдений были получены иные результаты. Данные оказались независимыми для всех периодов наблюдения. Таким образом, при анализе динамики

концентрации CO_2 нельзя применять критерии, требующие статистическую зависимость данных, такие как критерий Вилкоксона и критерий знаков.

В силу независимости рядов температуры за временные промежутки, не превышающие 1 год, для проверки гипотезы о равенстве средних значений температуры воздуха и концентрации CO_2 в указанных временных интервалах, применены критерии Манна-Уитни (М-У), Колмогорова-Смирнова (К-С), Вальда-Вольфовица (В-В) и Вилкоксона.

В таблице 1 представлены результаты проверки альтернативных гипотез об увеличении различных температурных параметров. Из таблицы 1 видно, что

Таблица 1. Результаты проверки гипотез

| Параметр | Критерии | | | |
|--|------------|------------|-----------|-------------|
| | М-У | К-С | В-В | Вилкоксона |
| Параметр зимой 2019–2020 г. больше, чем зимой 2000–2001 г. | | | | |
| Сред. темп. | 0 | 0 | 0 | 0 |
| Мат. ожид. среднесут. темп. | $9,03E-13$ | $2,63E-12$ | $1,85E-5$ | $2,40E-10$ |
| Параметр летом 2020 г. больше, чем летом 2000 г. | | | | |
| Сред. темп. | $8,75E-54$ | $8,55E-47$ | 0 | $3,64E-61$ |
| Мат. ожид. среднесут. темп. | 0,002 | 0,002 | 0,657 | 0,006 |
| Параметр в 2020 г. больше, чем в 2000 г. | | | | |
| Сред. темп. | $1,06E-26$ | $1,29E-65$ | 0 | $4,86E-181$ |
| Мат. ожид. среднесут. темп. | 0,048 | 0,018 | 0,297 | $3,00E-6$ |

на основании всех статистических критериев средняя температура воздуха зимой 2020-го года выше, чем температура воздуха зимой 2000-го, а средняя температура воздуха летом 2020-го года больше, чем температура воздуха летом 2000-го.

Аналогично были проверены гипотезы о равенстве средних значений концентрации CO_2 . Из результатов тестов видно, что средняя концентрация CO_2 зимой 2020-го года выше, чем средняя концентрация CO_2 зимой 2000-го, а средняя концентрация CO_2 летом 2020-го года ниже, чем средняя концентрация CO_2 2000-го.

Основным результатом данной работы является статистически обоснованный вывод о значимом увеличении температуры воздуха в 2020 году по сравнению с каждым годом, начиная с 2000-го и заканчивая 2007-м. Гипотеза об увели-

чении концентрации CO_2 за указанные промежутки времени не подтвердилась на уровне значимости 0,01. По данным долгосрочных наблюдений можно сделать вывод о том, что тенденция статистически значимого увеличения температуры воздуха в рассматриваемом регионе наблюдается с 1987-го года, в то время как рост концентрации CO_2 в атмосфере не является существенным.

- [1] Газарян В. А., Курбатова Ю. А., Овсянников Т. А., Шапкина Н. Е. Современные климатические изменения на юго-западе Валдайской возвышенности: статистический анализ многолетней динамики температуры воздуха // Вестник Московского университета. Серия 3. Физика. Астрономия. — 2015. — № 5. —
- [2] Barker T., Bashmakov I., Bernstein L. et al. IPCC. Climate Change 2007. — N.Y., 2007.

Топологическая теория для анализа сложных данных, включающих разнородные признаковые описания

Торшин Иван Юрьевич¹

tiy135@yahoo.com

¹Москва, ФИЦ ИУ РАН

Факторизация вкладов различных переменных при анализе разнородных признаков описаний — насущная задача интеллектуального анализа сложных данных. В работе предложено развитие решеточного формализма топологической теории анализа данных, в рамках которого получены новые способы порождения параметрических оценок и метрик на решетках, образованных над топологиями множеств объектов. Формализм был апробирован на задаче формирования множеств прецедентов для проведения хеомикробиомного анализа.

В биомедицинских исследованиях объектом служит формализованное описание состояния пациента, включающее булевы (диагнозы, прием лекарств и др.), числовые (лабораторные анализы) и категориальные (показатели демографии и др.) переменные, графы (формулы лекарств), временные ряды и изображения (результаты обследования пациента аппаратными методами). При формализации таких задач важно выделить независимые вклады переменных-признаков в целевую переменную (отклик) таким образом, чтобы получить высокое качество работы алгоритмов распознавания/классификации. Топологический анализ данных, развиваемый в рамках алгебраического подхода к распознаванию научной школы Ю. И. Журавлева и К. В. Рудакова [1], позволяет систематически исследовать возможные решения этой задачи.

В настоящей работе данный подход апробирован на задаче формирования множества прецедентов для проведения хеомикробиомного анализа лекарств. В этой прикладной задаче (оценка влияния лекарств на микробиоту) факт использования лекарства, как правило, описан булевым признаком, а откликом служит уровень той или иной бактерии микробиома (числовая переменная). Для выделения независимых вкладов булевых переменных в числовые целевые переменные предложен системный подход к порождению параметризованных оценок на решетке (решеточных термов) и соответствующих метрик. Такой подход представляет собой теоретическое обобщение «расщепления» эмпирической функции распределения булевым признаком и тесно связан с порождением метрических функций расстояния и с проблематикой т. н. «оценок информативности» (точнее, весовых функций признаков), используемых в комбинаторной теории разрешимости.

Формализм был апробирован на задаче формирования множеств прецедентов для проведения хеомикробиомного анализа. В исходной таблице признаков описаний представлены булевы переменные, соответствующие 122 лекарствам, которые влияют на 1533 числовые переменные, соответствующие уровням отдельных бактерий или их групп. Для квантификации вклада лекарства в изменение уровней заданного типа бактерии среди разработанных оценок ис-

пользовались уклонение Колмогорова со знаком для учета направления изменения уровней бактерий.

По тесту Колмогорова–Смирнова вычислялось значение статистической достоверности $p(D)$ для отсеечения наименее достоверных ассоциаций «лекарство — эффект». Также использовались более очевидные подходы: коэффициенты линейной многопараметрической регрессии (с отбором признаков по модели «лассо»), разность числовых значений переменных, доля разности и др.

Лекарства с известной химической структурой описывались на основе теории хемографов; в качестве множества начальных информаций использовалось множество хемоинвариантов над алфавитом элементарных меток [1]. Алгоритмы строились в виде композиций вложенных корректирующих функций нижнего уровня (т. е. порождения синтетических признаков) для фиксированного числа моделей.

Результаты вычислительных экспериментов показывают, что порождение множества исходных информаций на основе регрессионных коэффициентов и разности значений материала обучения соответствует крайне низкой обобщающей способности настраиваемых алгоритмов. При этом ужесточение порога на значение $p(D)$, давая снижение числа различных типов хеомикробиомной активности, не приводило к увеличению качества алгоритмов.

Тогда как порождение множества исходных информаций на основе регрессионных коэффициентов и разности значений материала обучения соответствовало крайне низкой обобщающей способности настраиваемых алгоритмов (коэффициент корреляции на контроле — $0,32 \pm 0,20$), использование предлагаемых оценок для порождения множеств прецедентов в задачах хеомикробиомики позволило существенно повысить обобщающую способность соответствующих алгоритмов (коэффициент корреляции на контроле — $0,79 \pm 0,21$).

Работа выполнена при поддержке гранта РФ (проект 23-21-00154).

- [1] *Торшин И. Ю.* О формировании множеств прецедентов на основе таблиц разнородных признаков описаний методами топологической теории анализа данных // Информатика и ее применения. — 2023. — Т. 17, № 3. — С. 2–7.

Новый двухуровневый метод машинного обучения для оценивания вещественных характеристик объектов

Сенько Олег Валентинович^{1*}

senkoov@mail.ru

Докукин Александр Александрович¹

dalex@ccas.ru

¹Москва, Федеральный исследовательский центр «Информатика и управление»
Российской академии наук

Ансамблевые методы имеют весьма продолжительную историю и являются существенной частью технологий машинного обучения, используемых при решении задач обучения по прецедентам: классификации или предсказания числовых переменных. Под ансамблевым методом обычно понимается метод, вычисляющий решения в два этапа: отдельными алгоритмами ансамбля и коллективным алгоритмом.

Среди ансамблевых технологий наибольшее распространение получили метод случайных лесов [1] и метод градиентного бустинга [2]. Они широко применяются при решении разнообразных прикладных задач, демонстрируя во многих случаях высокую эффективность. Причем ни один из них не является бесспорным лидером — на практике встречаются задачи, где то один, то другой из этих методов демонстрируют превосходство, иногда со значительным отрывом. Вместе с тем можно предположить, что два указанных метода не исчерпывают все возможности достижения высокой эффективности ансамблевых решений. Возникает даже желание объединить эти подходы, что в какой-то мере реализуется в предлагаемом методе.

В методе случайных лесов ансамбли генерируются случайно, что, очевидно, не обеспечивает их оптимальности. В то же время градиентный бустинг не является в полной мере ансамблевым методом — это, несомненно, сумма алгоритмов, но каждый в отдельности аппроксимирует не целевую переменную, а производные функции потерь. В нашем методе предлагается совместить лучшее из двух подходов. С одной стороны, будем строить ансамбль из разнообразных элементов, каждый из которых в определенной степени случаен, и аппроксимирует целевую переменную. При этом сами элементы будут строиться путем оптимизации специального функционала.

Введем некоторые обозначения. Целевую переменную будем обозначать Y , признаковое описание — $x = (X_1, \dots, X_n)$, а их значения для i -го объекта — (y_i, x_i) . Обучающую выборку, т.е. набор объектов с известным целевым значением, обозначим S . Конечная цель состоит в построении алгоритма A , такого что $A(x) \approx y$ для произвольного объекта (y, x) .

Рассмотрим теперь итерационную процедуру построения ансамбля. Обозначим через $B_i(x)$ произвольный алгоритм: как дерево, так, например, и линейную комбинацию деревьев, получаемый на i -м шаге. К шагу k ансамбль будет состоять из всех алгоритмов, построенных на предыдущих шагах, а ансамблевым

решением $A_k(x)$ будем считать среднее этих алгоритмов:

$$A_1(x) = O(x),$$

$$A_k(x) = \frac{1}{k-1} \sum_{i=1}^{k-1} B_i(x), \quad k > 1,$$

где $O(x)$ — тождественный ноль.

В качестве функционала ошибки нас будет интересовать среднее квадратичное отклонение:

$$L(S, A_k) = \frac{1}{m} \sum_{j=1}^m l(y_j, A_k(x_j)) = \frac{1}{m} \sum_{j=1}^m (y_j - A_k(x_j))^2.$$

После преобразований, подробно описанных в [3], его можно преобразовать к следующему виду:

$$L(S, A_k) = \frac{1}{km} \sum_{j=1}^m \sum_{i=1}^k (y_j - B_i(x_j))^2 - \frac{1}{km} \sum_{j=1}^m \sum_{i=1}^k (A_k(x_j) - B_i(x_j))^2.$$

Таким образом, оптимальность ансамбля может достигаться за счет двух факторов: хорошей аппроксимации связи Y с переменными X_1, \dots, X_n на обучающей выборке и высокой дисперсии прогнозов обучающих объектов. Эти рассуждения позволяют выбрать функционал для оптимизации нового слагаемого ансамбля

$$Q(B, A_k, \mu) = \frac{1}{m} \sum_{j=1}^m (y_j - B(x_j))^2 - \mu \frac{1}{km} \sum_{j=1}^m \sum_{i=1}^k (A_k(x_j) - B(x_j))^2. \quad (1)$$

В качестве B_k может быть использовано регрессионное дерево T_k , обучаемое по выборке $\{(B^*(x_1), x_1), \dots, (B^*(x_m), x_m)\}$. Предположим, что за первые $k-1$ шагов получен ансамбль B_1, \dots, B_{k-1} . Сгенерируем бутстрэп репликацию S_k^b выборки S в проекции на случайное подпространство. По репликации S_k^b построим регрессионное дерево T_k . Алгоритм B_k ищется в виде суммы $B_k = T_k + t_k$, где t_k — корректирующее дерево, которое строится, исходя из условия минимизации $Q(B, T_k + t, \mu) = Q(t, \mu)$. Для этого ищется действительный вектор $t^* = (t_1^*, \dots, t_m^*)$, компонентами которого являются оптимальные смещения прогнозов, вычисляемых деревом T_k , т.е. $t^* = \arg \min Q(t, \mu)$. Минимальное значение функционала $Q(t, \mu)$ по t_j достигается при $\partial Q(t, \mu) / \partial t(x_j) = 0$ или при

$$t(x_j) = t_j^* = \frac{k}{k-\mu} y_j - T_k(x_j) - \frac{\mu}{k-\mu} A_k(x_j), \quad j = \overline{1, m}.$$

Таким образом, описана рекурсивная процедура построения ансамбля алгоритмов. Начиная с пустого ансамбля и тождественного нуля в качестве коллективного решения, затем пополним его суммами пар деревьев, где первое аппроксимирует целевую переменную напрямую, а второе аппроксимирует поправку, минимизирующую функционал $Q(t, \mu)$. Построение ансамбля завершается, если k достигает задаваемого пользователем порогового значения N . Ансамбли, которые строятся с помощью представленной выше процедуры, назовем декорелированными.

Наряду со способом генерации ансамбля важную роль для достижения высокой эффективности ансамблевого алгоритма играет процедура, вычисляющая коллективное решение. Хотя ансамбль строится, исходя из соображений минимизации ошибки усредненного ответа алгоритмов, использование такой схемы коллективного решения оказалось недостаточно эффективно. По этой причине в методе реализованы несколько вариантов коллективного решения: усреднение и стекинг с методом градиентного бустинга и случайным лесом, а также их выпуклой комбинацией. Также для улучшения качества метод дополнен различными эвристиками.

Результаты использования метода для прогнозирования свойств химических элементов представлены в [4]. Примеры решения других задач из области медицины и экономики представлены в [3]. Там же приводится подробное обоснование метода.

Работа выполнялась с использованием инфраструктуры Центра коллективного пользования «Высокопроизводительные вычисления и большие данные» (ЦКП «Информатика») ФИЦ ИУ РАН (г. Москва).

- [1] *Breiman L.* Random forests // *Machine Learning*. — N. Y.: Springer, 2001. — Vol. 45, No. 1. — P. 5–32.
- [2] *Schapire R. E., Freund Y.* *Boosting: Foundations and Algorithms*. — Cambridge, Massachusetts, London: MIT Press, 2012. — 543 p.
- [3] *Dokukin A. A., Sen'ko O. V.* A New Two-Level Machine Learning Method for Evaluating the Real Characteristics of Objects // *J. Computer and Systems Sciences International*, — Pleiades Publishing, Ltd., 2023. — Vol. 62, No. 4. — P. 607–614.
- [4] *Senko O. V., Dokukin A. A., Kiselyova N. N., Dudarev V. A., Kuznetsova Yu. O.* New Two-Level Ensemble Method and Its Application to Chemical Compounds Properties Prediction // *Lobachevskii J. Mathematics*. — Pleiades Publishing, Ltd., 2023. — Vol. 44, No. 1. — P. 188–197.

ДСМ-классификатор над производением частичных порядков

*Дюкова Елена Всеволодовна*¹

edjukova@mail.ru

*Масляков Глеб Олегович*¹

gleb-mas@mail.ru

Янаков Дмитрий Спартакович^{2*}

dmitriyyanakov@gmail.com

¹Москва, ФИЦ ИУ РАН

²Москва, НИУ ВШЭ

Рассматривается задача классификации по прецедентам. Под прецедентной (обучающей) информацией понимается совокупность примеров изучаемых объектов, в которой каждый объект представлен в виде числового вектора, полученного на основе измерения или наблюдения ряда его параметров или характеристик, называемых признаками. Каждый пример (обучающий объект или прецедент) приписан к определенному классу объектов. Предполагается, что каждый признак имеет ограниченное множество допустимых значений, которые кодируются целыми числами, и признаковые описания любых двух обучающих объектов, принадлежащих разным классам, не совпадают. Требуется уметь классифицировать новые, не входящие в обучающую выборку, объекты.

Исследуется логический подход к рассматриваемой задаче, позволяющий прогнозировать редкие события. Обучение классификатора сводится к поиску в исходных данных информативных фрагментов описаний прецедентов. Такие фрагменты позволяют различать объекты из разных классов и, как правило, имеют содержательное описание в терминах той прикладной области, в которой решается задача. По их наличию в описании распознаваемого объекта решается вопрос о его классификации. Корректность обучения позволяет безошибочно распознавать прецеденты. Наиболее эффективен логический подход в случае целочисленной информации низкой значности, особенно бинарной. В случае большого числа признаков возникают вычислительные трудности, связанные с необходимостью решать сложные дискретные задачи.

Одними из основных направлений логической классификации являются процедуры корректного голосования (Correct Voting Procedures или CVP) и анализ формальных понятий (Formal Concept Analysis или FCA). В [1] приведена общая схема работы логического классификатора, включающая оба направления CVP и FCA.

Фундаментальную роль в создании методов CVP сыграли работы С. В. Яблонского, Ю. И. Журавлева, М. М. Бонгарда и М. Н. Вайнцвайга. В дальнейшем направление CVP развивалось отечественными и зарубежными учеными и наиболее существенное развитие получило в работах представителей школы академика РАН Ю. И. Журавлева.

Основополагающие идеи FCA принадлежат Р. Вилле (1981 г.). В России методы FCA представлены в работах В. К. Финна, С. О. Кузнецова, М. И. Забейко, Д. И. Игнатова и Д. В. Виноградова. В [2] В. К. Финном пред-

ложен так называемый метод автоматического порождения гипотез (или ДСМ-метод), который позднее был адаптирован автором и его учениками для задач машинного обучения.

Пусть $X = \{x_1, \dots, x_n\}$ — заданное множество признаков; H — набор из r различных признаков вида $H = \{x_{j_1}, \dots, x_{j_r}\}$; $\sigma = (\sigma_1, \dots, \sigma_r)$ — набор, в котором σ_i — допустимое значение признака x_{j_i} , $i = \overline{1, r}$. Пару (σ, H) назовем элементарным классификатором (ЭК).

Пусть $S = (a_1, \dots, a_n)$ — объект из M (здесь a_j , $j = \overline{1, n}$, — значение признака x_j для объекта S). Будем говорить, что ЭК (σ, H) содержится в S , если $a_{j_i} = \sigma_i$ при $i = \overline{1, r}$. ЭК (σ, H) называется представительным для класса K , если (σ, H) содержится в хотя бы одном прецеденте из K и не содержится ни в каком прецеденте не из K .

Понятие представительного ЭК введено в направлении СVP [3] и используется в [1] при описании общей схемы работы логического классификатора. Согласно данной схеме каждый такой классификатор на этапе обучения задает некоторый частичный порядок на специальном подмножестве представительных ЭК класса K и ищет максимальные относительно заданного порядка элементы этого подмножества.

В СVP наиболее информативными считаются тупиковые представительные ЭК (σ, H) класса K , обладающие свойством: любой ЭК (σ', H') класса K такой, что $\sigma' \subset \sigma$, $H' \subset H$ не является представительным ЭК класса K . Решающее правило основано на процедуре «голосования», в которой участвуют найденные тупиковые представительные ЭК класса K [1, 3].

В ДСМ-методе искомое множество $P(K)$ представительных ЭК класса K порождается так называемыми положительными для класса K гипотезами. Каждый ЭК (σ, H) из $P(K)$ обладает следующим свойством: любой представительный ЭК (σ', H') класса K такой, что $\sigma \subset \sigma'$, $H \subset H'$, содержится в меньшем числе прецедентов. Распознаваемый объект S относится к классу K , если в S содержится хотя бы один ЭК из $P(K)$ и не содержится ни одного ЭК из $P(K')$, $K' \neq K$. В противном случае происходит отказ от классификации. Основная цель представляемой работы — улучшение качества работы рассматриваемого классификатора за счет изменения слишком строгого решающего правила и модификация усовершенствованной версии на случай, когда на множествах значений признаков заданы частичные порядки.

На базе ДСМ-метода разработан классификатор ДСМ-Г с использованием решающего правила, применяемого в СVP. Новый классификатор модифицирован для работы с описаниями объектов, представляющими собой элементы декартова произведения конечных частично упорядоченных множеств. В указанной модификации ДСМ-метода, названной в работе классификатором ДСМ-Г+, используются идеи, предложенные ранее в [4] при создании аналогичных алгоритмов направления СVP.

Авторами настоящей работы проведено экспериментальное исследование построенных классификаторов на модельных данных различной размерности, а также на реальных задачах с линейными порядками на множествах значений признаков, полученными на основе анализа встречаемости отдельного значения признака в описаниях прецедентов. Благодаря специальной бинарной кодировке данных процедура обучения сводилась к классическому виду. Для перечисления ДСМ-гипотез применялся алгоритм Close By One [5]. В тестировании также участвовали алгоритмы из SVM и классические алгоритмы машинного обучения, такие как логистическая регрессия, случайный лес и градиентный бустинг над решающими деревьями. На рассмотренных реальных задачах классификатор ДСМ-Г+ показал более высокое качество по сравнению с классическим вариантом ДСМ-метода, а на некоторых из этих задач превзошел все тестируемые алгоритмы. Однако в случае задач большой размерности из-за большого числа порождаемых ДСМ-гипотез классификаторы ДСМ-Г и ДСМ-Г+ работают существенно дольше, чем другие методы.

- [1] Дюкова Е. В., Масляков Г. О., Дюкова А. П. Логические методы корректной классификации данных // Информатика и ее применения. — 2023. — Т. 17, № 3. — С. 64–70.
- [2] Финн В. К. О возможности формализации правдоподобных рассуждений средствами многозначных логик // Всесоюз. симп. по логике и методологии науки, Киев: Наукова думка. — 1976. — С. 82–83.
- [3] Дюкова Е. В., Песков Н. В. Поиск информативных фрагментов описаний объектов в дискретных процедурах распознавания // Ж. вычисл. матем. и матем. физ. — 2002. — Т. 42, № 5. С. 741–753.
- [4] Дюкова Е. В., Масляков Г. О., Прокофьев П. А. О логическом анализе данных с частичными порядками в задаче классификации по прецедентам // Ж. вычисл. матем. и матем. физ. — 2019. — Т. 59, № 9. — С. 1605–1616.
- [5] Кузнецов С. О. Быстрый алгоритм построения всех пересечений объектов из нижней полурешетки // Научная и техническая информация. — 1993. — Сер. 2, № 1. — С. 17–20.

Частично формализованный способ отбора значимых признаков для линейной бинарной классификации (на примере диагностики гипоксически-ишемических поражений головного мозга пациентов детского возраста по данным МР-морфометрии)

*Ерохин Михаил Владимирович*¹

erokhin_m_v@mail.ru

*Прошин Владимир Иванович*¹

proshin.v@list.ru

Ерохин Владимир Иванович^{2*}

vka@mail.ru

¹Санкт-Петербург, Санкт-Петербургский государственный педиатрический медицинский университет

²Санкт-Петербург, Военно-космическая академия имени А. Ф. Можайского

Отбор подмножества значимых признаков (а также исключения избыточных признаков) является важным этапом построения модели в задачах машинного обучения. Соответствующие методы и алгоритмы интенсивно развиваются, им посвящена обширная литература. Мотивация соответствующих исследований хорошо понятна специалистам: «Часто наборы данных, с которыми приходится работать, содержат большое количество признаков, число которых может достигать нескольких сотен и даже тысяч. При построении модели машинного обучения не всегда понятно, какие из признаков действительно для нее важны (т.е. имеют связь с целевой переменной), а какие являются избыточными (или шумовыми). Удаление избыточных признаков позволяет лучше понять данные, а также сократить время настройки модели, улучшить ее точность и облегчить интерпретируемость. Иногда эта задача и вовсе может быть самой значимой, например, нахождение оптимального набора признаков может помочь расшифровать механизмы, лежащие в основе исследуемой проблемы...» (<https://habr.com/ru/articles/550978/>).

Одной из простейших, но в то же время базовых задач машинного обучения, требующих отбора признаков, является задача линейной бинарной классификации. Указанная задача может возникать в двух принципиально разных случаях:

- задача классификации является частной задачей автоматического (без прямого управления человеком) функционирования некоторой сложной технической системы, возможно работающей в режиме реального времени и обрабатывающей большие объемы информации;
- задача классификации является «разовой» частной задачей обработки результатов некоторого научного исследования, выполняется человеком «в ручном режиме» для сравнительно небольшого объема данных.

В докладе будет рассмотрен метод отбора значимых признаков для модели линейной бинарной классификации, в большей степени применимый ко второму случаю. Для его апробации использованы данные реального медицинского исследования гипоксически-ишемических поражений (ГИП) головного мозга у детей раннего возраста методами магнитно-резонансной (МР) морфометрии [1–3].

Рассматриваемую задачу можно отнести к задачам малой размерности (43 пациента с признаками ГИП, 50 пациентов без признаков ГИП, исходное количество признаков — 71). Наличие избыточных признаков в исследуемом наборе данных было выявлено ранее в рамках исследования [2].

Предлагаемый метод выделения значимых признаков (и удаления избыточных признаков) основан на результатах работы [4]. Идея метода заключается в следующем:

Рассматриваются два конечных множества точек $\mathbf{P} = \{p_i\}_{i=1}^{m_1}$ и $\mathbf{Q} = \{q_i\}_{i=1}^{m_2}$ в пространстве \mathbb{R}^n . Указанным множествам соответствуют матрицы

$$P = \begin{bmatrix} p_1 \\ \vdots \\ p_{m_1} \end{bmatrix} = (p_{ij}) \in \mathbb{R}^{m_1 \times n}, \quad Q = \begin{bmatrix} q_1 \\ \vdots \\ q_{m_2} \end{bmatrix} = (q_{ij}) \in \mathbb{R}^{m_2 \times n}.$$

Классическая задача линейной бинарной классификации, сводящаяся к решению системы линейных неравенств

$$Pw \leq 1, \quad Qw \geq 1, \tag{1}$$

где $w \in \mathbb{R}^n$ — вектор неизвестных коэффициентов левой части уравнения разделяющей гиперплоскости (правая часть уравнения нормирована к 1), «погружается» в задачу линейной бинарной классификации с интервальной неопределенностью данных. В указанной задаче вместо точных значений матриц P и Q известны их нижние и верхние границы $\underline{P} = (\underline{p}_{ij}) \in \mathbb{R}^{m_1 \times n}$, $\overline{P} = (\overline{p}_{ij}) \in \mathbb{R}^{m_1 \times n}$, $\underline{Q} = (\underline{q}_{ij}) \in \mathbb{R}^{m_2 \times n}$, $\overline{Q} = (\overline{q}_{ij}) \in \mathbb{R}^{m_2 \times n}$, такие, что следующие неравенства выполнены поэлементно:

$$\underline{P} \leq P \leq \overline{P}, \quad \underline{Q} \leq Q \leq \overline{Q} \Leftrightarrow (P_c - P_r) \leq P \leq (P_c + P_r), \quad (Q_c - Q_r) \leq Q \leq (Q_c + Q_r), \tag{2}$$

где $P_c = (\underline{P} + \overline{P})/2$, $Q_c = (\underline{Q} + \overline{Q})/2$, $P_r = (\overline{P} - \underline{P})/2$, $Q_r = (\overline{Q} - \underline{Q})/2$.

Как было показано в [4], справедливо

Утверждение 1. *Задача линейной бинарной классификации данных с интервальной неопределенностью, заданная условиями (1)–(2), разрешима тогда и только тогда, когда имеет решение система линейных неравенств*

$$\begin{bmatrix} \overline{P} \\ -\underline{Q} \end{bmatrix} w^+ - \begin{bmatrix} \underline{P} \\ -\overline{Q} \end{bmatrix} w^- \leq \begin{bmatrix} 1_p \\ -1_q \end{bmatrix}, \quad w^+, w^- \geq 0, \tag{3}$$

где $w^+, w^- \in \mathbb{R}^n$, 1_p — вектор размерности m_1 , составленный из единиц, -1_q — вектор размерности m_2 , составленный из минус единиц. При этом $w = w^+ - w^-$, $|w| \leq w^+ + w^-$.

Пусть $v = w^+ + w^-$, $\alpha > 0$ — параметр, регулирующий ширину «разделительной полосы» между множествами \mathbf{P} и \mathbf{Q} . Выполнив очевидные преобразования

системы линейных неравенств (3), запишем задачу линейного программирования:

$$\begin{bmatrix} P_c \\ -Q_c \end{bmatrix} w + \begin{bmatrix} P_r \\ Q_r \end{bmatrix} v \leq \begin{bmatrix} 1_p(\alpha + 1) \\ 1_q(\alpha - 1) \end{bmatrix}, \quad v \geq 0, \quad \sum_{j=1}^n v_j \rightarrow \min. \quad (4)$$

Задача (4) и составляет основу предлагаемого (частично формализованного) метода отбора значимых признаков (и исключения избыточных).

Результаты и особенности практического применения указанного метода при анализе реальных медицинских данных, заключающиеся в «ручной» настройке ширины интервалов, ширины разделяющей полосы и точечной «ручной» корректировке состава отбираемых признаков на основе экспертного мнения врача-специалиста по лучевой диагностике, предполагается обсудить в данном докладе.

- [1] *Ерохин М. В., Тащилкин А. И.* О возможностях МР-морфометрии при диагностике гипоксически-ишемических поражений головного мозга у детей раннего возраста // *Детская медицина Северо-Запада.* — 2018. — Т. 7, № 1. — С. 116.
- [2] *Ерохин М. В., Плоткин А. В.* Анализ объема церебральных структур пациентов с гипоксически-ишемической энцефалопатией // *Математические методы распознавания образов: Тез. докл. 19-й Всеросс. конф. с междунар. участием.* — М.: Российская академия наук, 2019. — С. 272.
- [3] *Ерохин М. В. и др.* Возможности МР-морфометрии головного мозга у детей с гипоксически-ишемической энцефалопатией // *Визуализация в медицине.* — 2020. — Т. 2, № 1. — С. 32–39.
- [4] *Ерохин В. И., Кадочников А. П., Сотников С. В.* Линейная бинарная классификация данных с интервальной неопределенностью // *Искусственный интеллект и принятие решений.* — 2023. — № 3. — С. 76–83.

Об одном робастном варианте метода главных компонент

Шибзухов Заур Мухадинович¹*

intellimath@mail.ru

¹Москва, Московский педагогический государственный университет

Классический метод главных компонент первоначально рассматривался как задача наилучшей аппроксимации конечного множества точек прямыми и плоскостями [1]. В машинном обучении метод главных компонент часто используется как метод сокращения размерности входных данных. Однако, если часть данных существенно искажена, то неизбежно также существенно искажаются результаты работы классического метода. В настоящей работе предлагается один робастный вариант [2] постановки задачи поиска главных компонент. Он основан на применении дифференцируемых оценок среднего значения, нечувствительных или мало чувствительных к выбросам [3,4]. На реальных данных [5] наглядно демонстрируется эффективность предложенного подхода к построению робастных вариантов метода главных компонент по сравнению с классическим вариантом.

В классическом варианте решается задача поиска вектор центра a_0 и ортонормированного базис a_1, \dots, a_m m -мерной гиперплоскости в \mathbb{R}^n , такой что сумма квадратов Евклидовых расстояний до нее от точек x_1, \dots, x_N минимальна:

$$\sum_{k=1}^N \left(\|x_k - a_0\|^2 - \sum_{j=1}^m (x_k - a_0, a_j)^2 \right) \rightarrow \min.$$

Решение этой задачи редуцируется к цепочке задач, в которых последовательно находятся векторы a_0, a_1, \dots, a_m .

Робастный вариант постановки задачи поиска центра a_0 имеет вид:

$$a_0 = \arg \min_{a \in \mathbb{R}^n} M\{\|x_1 - a\|^2, \dots, \|x_N - a\|^2\},$$

где $M\{z_1, \dots, z_N\}$ — дифференцируемая оценка среднего значения, нечувствительная или малочувствительная к выбросам. Она сводится к решению уравнения

$$a = \sum_{k=1}^N \frac{\partial M\{\|x_1 - a_0\|^2, \dots, \|x_N - a_0\|^2\}}{\partial z_k} x_k,$$

которое можно решить, используя следующую итерационную процедуру:

$$a^{t+1} = \sum_{k=1}^N v_k^t x_k, \quad v_k^t = \frac{\partial M\{\|x_1 - a^t\|^2, \dots, \|x_N - a^t\|^2\}}{\partial z_k}.$$

После нахождения a_0 также осуществляется центрирование: $x_k \rightarrow x_k - a_0$.

Робастный вариант задачи поиска a_j принимает следующий вид:

$$a_j = \arg \min_{\|a\|=1} \mathbf{M} \{ \|x_1\|^2 - (a, x_1)^2, \dots, \|x_N\|^2 - (a, x_N)^2 \}.$$

При помощи метода итеративного перевзвешивания ее решение также можно свести к цепочке задач:

$$a_j^{t+1} = \arg \min_{\|a\|=1} \sum_{k=1}^N v_k^t (\|x_k\|^2 - (a, x_k)^2)$$

со следующими значениями весов точек:

$$v_k^t = \frac{\partial \mathbf{M} \{ \|x_1\|^2 - (a^t, x_1)^2, \dots, \|x_N\|^2 - (a^t, x_N)^2 \}}{\partial z_k}.$$

Матрица ковариации принимает вид: $S^t = X^\top \text{diag}(v_1^t, \dots, v_N^t) X$.

Вектор a_j^{t+1} является решением системы $S^t a = \lambda a$ при условии $\|a\|^2 = 1$, т. е. является ортонормированным собственным вектором матрицы S^t , а λ_j^{t+1} — ее собственным числом.

Для поиска a_j применяется итерационная процедура:

$$a^{t+1} = \frac{1}{\lambda^t} (S^t a^t), \quad \lambda^t = \frac{(a^t)^\top S^t a^t}{(a^t, a^t)}.$$

После нахождения a_j осуществляется проецирование: $x_k \rightarrow x_k - a_j(a_j, x_k)$.

- [1] *Pearson K.* On lines and Planes of Closest Fitto Systems of Points in Space // *Philosophical Magazine.* — 1901. — Vol. 2. — P. 559–572.
- [2] *Rousseeuw P. J., Leroy A. M.* Robust Regression and Outlier Detection. — New York: John Wiley and Sons, 1987. — 341 p.
- [3] *Shibzukhov Z. M.* Machine Learning Based on the Principle of Minimizing Robust Mean Estimates // *Brain-Inspired Cognitive Architectures for Artificial Intelligence: BICA*AI 2020.* — Springer International Publishing, 2020. — P. 472–477.
- [4] *Shibzukhov Z. M.* Minimizing Robust Estimates of Sums of Parameterized Functions // *J. Mathematical Sciences.* — Springer Science and Business Media LLC, 2022. — Vol. 260. — P. 249–264.
- [5] Dataset Hipparcos. https://www.astrostatistics.psu.edu/datasets/HIP_star.html

Технология кластеризации парных сравнений

*Двоенко Сергей Данилович*¹

sergedv@yandex.ru

¹Тула, Тульский государственный университет

Рассмотрена технология кластеризации, которая заключается в выполнении обработки данных, представленных только парными сравнениями похожести или различий элементов множества: вычисление дисперсий кластеров, представления начала координат как нового элемента множества, представление линейной комбинации элементов множества, определения центра множества, перенос начала координат, определение центра подмножества [1].

Указанные операции позволяют построить новые и модифицировать известные алгоритмы кластер-анализа и машинного обучения для обработки матриц парных сравнений при отсутствии традиционной матрицы данных, представляющей собой измерения характеристик объектов.

На примере алгоритма *k*-средних показаны его варианты для парных сравнений расстояний и близостей как положительных скалярных произведений. Показаны условия, при которых результат обработки парных сравнений не отличается от результата работы для матрицы данных. Разработаны новые версии в виде т. н. перестановочных алгоритмов. Показана связь задачи кластеризации парных сравнений с известными задачами агрегирования произвольных матриц связей, факторного анализа, экстремальной группировки параметров.

Показано, что наличие технологии предполагает выполнения ряда действий для получения корректного результата. В задаче кластеризации парных сравнений рассмотрены технологические этапы для построения алгоритмов кластер-анализа как полных аналогов таких алгоритмов для матрицы данных (*k*-средних, Форель, алгоритм Козинца и проч.). Перестановочные процедуры образуют новый класс алгоритмов *k*-средних. Они применяются в задачах улучшения качества кластеризации, где обычные версии этого алгоритма не работают.

Работа поддержана грантом РФФИ № 20-07-00055.

- [1] *Dvoenko S. D.* Clustering of Data Represented by Pairwise Comparisons // Control and Cybernetics. — 2022. — Vol. 51, No. 3. — P. 343–387.

Нижняя граница погрешности для оценок параметров распределений при заданном количестве информации

Ланге Михаил Михайлович¹*

lange_mm@mail.ru

Ланге Андрей Михайлович¹

lange_am@mail.ru

¹Москва, Федеральный исследовательский центр «Информатика и управление» Российской академии наук

Традиционным критерием качества оценок случайных параметров распределений по выборкам независимых наблюдений является отклонение средней квадратичной погрешности от минимального значения, которое определяется средней по выборкам дисперсией апостериорной плотности распределения параметров. Указанная средняя дисперсия не зависит от количества информации в выборках, которая используется при построении оценок. В настоящей работе исследуется наименьшая средняя квадратичная погрешность для оценок вектора случайных параметров плотности распределения при фиксированном среднем количестве информации, содержащейся в выборках относительно множества оценок параметров. Зависимость наименьшей средней погрешности от среднего количества информации строится в форме обращения монотонной зависимости наименьшей средней взаимной информации от значений погрешности. Построенная зависимость дает двухфакторный критерий качества оценивания параметров, который получен на основе обобщения известной в теории информации функции «скорость-погрешность» (rate distortion function) [1, 2] для модели кодирования непрерывных сообщений с допустимой погрешностью.

Формализация задачи. Пусть $p_{X|\Theta^m}(x|\theta^m)$ — условная плотность распределения случайной величины x на множестве значений X , где $\theta^m = (\theta_1, \dots, \theta_m)$ — вектор неизвестных случайных параметров с априорной плотностью $p_{\Theta^m}(\theta^m) = \prod_{k=1}^m p_{\Theta_k}(\theta_k)$, $\theta_k \in \Theta_k$, на ансамбле множеств $\Theta^m = (\Theta_1, \dots, \Theta_m)$. Будем считать, что вектор оценок параметров $\hat{\theta}^m = (\hat{\theta}_1, \dots, \hat{\theta}_m)$ строится по выборке $x^n = (x_1, \dots, x_n)$ независимых n наблюдений со среднеквадратичной погрешностью $(1/n) \sum_{k=1}^m (\theta_k - \hat{\theta}_k)^2$. Выборки наблюдений образуют множество X^n с условной плотностью $p_{X^n|\Theta^m}(x^n|\theta^m) = \prod_{i=1}^n p_{X_i|\Theta^m}(x_i|\theta^m)$, $x^n \in X^n$, $\theta^m \in \Theta^m$. Заданы априорные плотности $p_{\Theta_k}(\theta_k)$, $\theta_k \in \Theta_k$, $k = \overline{1, m}$, которые на ансамбле Θ^m дают плотность распределения $p_{\Theta^m}(\theta^m) = \prod_{k=1}^m p_{\Theta_k}(\theta_k)$. Множества оценок параметров образуют ансамбль $\hat{\Theta}^m = \hat{\Theta}_1 \dots \hat{\Theta}_m$, на котором вводится свободная условная плотность распределения $q_{\hat{\Theta}^m|X^n}(\hat{\theta}^m|x^n)$, $\hat{\theta}^m \in \hat{\Theta}^m$, $x^n \in X^n$. Тогда с учетом марковости тройки Θ^m , X^n , $\hat{\Theta}^m$ вводятся средние на компоненту вектора параметров погреш-

ность и взаимная информация

$$E_{q_{\widehat{\Theta}^m|X^n}}(X^n; \widehat{\Theta}^m) = (1/m) \int_{X^n} p_{X^n}(x^n) dx^n \int_{\widehat{\Theta}^m} q_{\widehat{\Theta}^m|X^n}(\widehat{\theta}^m|x^n) d\widehat{\theta}^m \int_{\Theta^m} \sum_{k=1}^m (\theta_k - \widehat{\theta}_k)^2 p_{\Theta^m|X^n}(\theta^m|x^n) d\theta^m, \quad (1)$$

$$I_{q_{\widehat{\Theta}^m|X^n}}(X^n; \widehat{\Theta}^m) = (1/m) \int_{X^n} p_{X^n}(x^n) dx^n \int_{\widehat{\Theta}^m} q_{\widehat{\Theta}^m|X^n}(\widehat{\theta}^m|x^n) \ln \frac{q_{\widehat{\Theta}^m|X^n}(\widehat{\theta}^m|x^n)}{q_{\widehat{\Theta}^m}(\widehat{\theta}^m)} \widehat{\theta}^m d\widehat{\theta}^m \quad (2)$$

как функционалы от условной плотности $q_{\widehat{\Theta}^m|X^n}(\widehat{\theta}^m|x^n)$. Здесь $p_{X^n}(x^n)$ и $q_{\widehat{\Theta}^m}(\widehat{\theta}^m)$ — безусловные плотности на множествах X^n и $\widehat{\Theta}^m$, $p_{\Theta^m|X^n}(\theta^m|x^n)$ — апостериорная плотность на ансамбле Θ^m .

Для значений $\varepsilon > 0$ функционалы (1) и (2) позволяют определить функцию

$$R_{mn}(\varepsilon) = \min_{q_{\widehat{\Theta}^m|X^n} : E_{q_{\widehat{\Theta}^m|X^n}}(X^n; \widehat{\Theta}^m) \leq \varepsilon} I_{q_{\widehat{\Theta}^m|X^n}}(X^n; \widehat{\Theta}^m), \quad (3)$$

где минимум берется по всевозможным условным плотностям $q_{\widehat{\Theta}^m|X^n}(\widehat{\theta}^m|x^n)$ при заданном ε -ограничении на среднюю погрешность. Функция (3) аналогична функции «скорость-погрешность», предложенной в [1] для модели кодирования независимых непрерывных сообщений, переданных по каналу с аддитивным гауссовым шумом. Задача состоит в построении монотонно убывающей нижней границы $\underline{R}_{mn}(\varepsilon) \leq R_{mn}(\varepsilon)$. Тогда обратная функция $\underline{R}_{mn}^{-1}(I)$ дает нижнюю границу средней погрешности при любом фиксированной значении средней взаимной информации I .

Основной результат. Применяя технику вычисления функции «скорость-погрешность» [2], получена нижняя граница

$$\underline{R}_{mn}(\varepsilon) = \frac{1}{m} \sum_{k=1}^m h(p_{\Theta_{k_n}}) - \frac{1}{2} \ln(2\pi e(\varepsilon - \varepsilon_{\min})), \quad \varepsilon_{\min} \leq \varepsilon \leq \varepsilon_{\max}. \quad (4)$$

Здесь $h(p_{\Theta_{k_n}}) = - \int_{\Theta_{k_n}} p_{\Theta_{k_n}}(\theta_{k_n}) \ln p_{\Theta_{k_n}}(\theta_{k_n}) d\theta_{k_n}$ — дифференциальная энтропия от плотности распределения оценок $\theta_{k_n}(x^n) = \int_{\Theta_k} \theta_k p_{\Theta_k|X^n}(\theta_k|x^n) d\theta_k$ на множестве Θ_{k_n} ; $p_{\Theta_k|X^n}(\theta_k|x^n)$ — апостериорная плотность на множестве Θ_k с дисперсией $\sigma_{k_n}^2(x^n) = \int_{\Theta_k} (\theta_k - \theta_{k_n}(x^n))^2 p_{\Theta_k|X^n}(\theta_k|x^n) d\theta_k$ и $\varepsilon_{\min} = (1/m) \sum_{k=1}^m \int_{X^n} p_{X^n}(x^n) \sigma_{k_n}^2(x^n) dx^n$. $\underline{R}_{mn}(\varepsilon_{\min}) \rightarrow \infty$ и $\underline{R}_{mn}(\varepsilon_{\max}) = 0$, т. е. граница (4) монотонно убывает. В [3] аналогичный подход использован для построения нижней границы вероятности ошибки классификации в пространстве представлений объектов с заданной метрикой.

Пример границы $\underline{R}_{mn}(\varepsilon)$ при $m = 1$. Пусть $p_X(x|\theta) = N(\theta, \sigma^2)$ — гауссова плотность с неизвестным средним значением θ и заданной дисперсией σ^2 . Предполагается, что априорная плотность $p_\Theta(\theta) = N(\theta_0, \sigma_0^2)$ является нормальной со средним θ_0 и дисперсией σ_0^2 . В этом случае апостериорное распределение также является нормальным с плотностью $p_{\Theta|X^n}(\theta|x^n) = N(\theta_n, \sigma_n^2)$, в которой $\theta_n = (\sigma_0^2/(n\sigma_0^2 + \sigma^2)) \sum_{k=1}^n x_k + (\sigma^2/(n\sigma_0^2 + \sigma^2)) \theta_0$ и $\sigma_n^2 = \sigma_0^2 \sigma^2 / (n\sigma_0^2 + \sigma^2)$ [4]. Тогда нормальное распределение $N(\theta_0, \sigma_0^2 + \sigma^2/n)$ выборочного среднего $(1/n) \sum_{k=1}^n x_k$ дает дифференциальную энтропию $h(p_{\Theta_n}) = (1/2) \ln(2\pi e n \sigma_0^4 / (n\sigma_0^2 + \sigma^2))$ на множестве оценок $\Theta_n = \{\theta_n\}$. Подстановка $h(p_{\Theta_n})$ в (4) дает границу $\underline{R}_{mn}(\varepsilon)$ при $m = 1$, в которой $\varepsilon_{\min} = \sigma_n^2 = \sigma_0^2 \sigma^2 / (n\sigma_0^2 + \sigma^2)$ и $\varepsilon_{\max} = \sigma_0^2$. Поведение полученной границы иллюстрирует кривая 2 на рисунке 1. Кривые 1 и 3 соответствуют известным функциям «скорость-погрешность» для модели кодирования независимых гауссовых величин θ по наблюдениям x на выходе канала с аддитивным гауссовым шумом $x - \theta$ [1], и для модели непосредственного кодирования величин θ [2]. Следует отметить, что выборки гауссовых наблюдений допускают вычисление границы $\underline{R}_{mn}(\varepsilon)$ при $m = 2$, когда неизвестными параметрами являются среднее значение θ и дисперсия σ^2 .

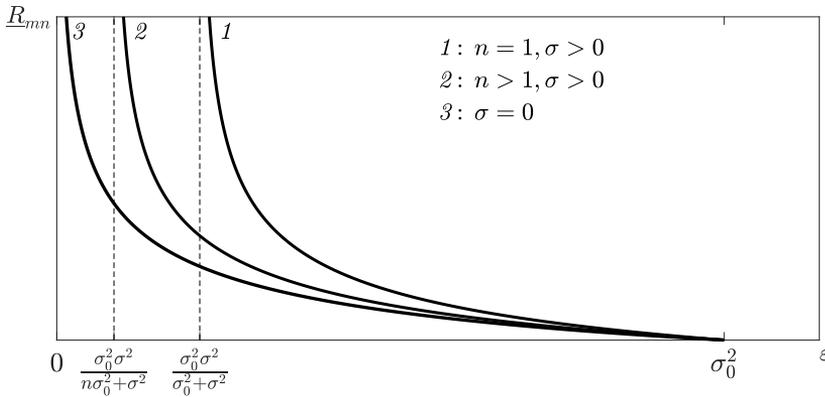


Рис. 1. Поведение границы $\underline{R}_{mn}(\varepsilon)$ для гауссовой модели (случай $m = 1$).

- [1] *Dobrushin R. L., Tsybakov B. S.* Information Transmission with Additional Noise // IRE Trans. Inform. Theory. — 1962. — Vol. 8, No. 5. — P. 293–304.
- [2] *Berger T.* Rate Distortion Theory: A Mathematical Basis for Data Compression. — Englewood Cliffs, New Jersey: Prentice-Hall Inc., 1971.
- [3] *Lange M. M., Lange A. M.* Information-Theoretic Lower Bounds to Error Probability for the Models of Noisy Discrete Source Coding and Object Classification // Pattern Recognition and Image Analysis. — 2022. — Vol. 32, No. 3. — P. 570–574.
- [4] *Duda R. O., Hart P. E., Stork D. G.* Pattern Classification, 2nd ed. — New York: Wiley & Sons, 2001.

Математическая модель эффекта обратной связи в системах искусственного интеллекта

Веприков Андрей Сергеевич^{1*}

veprikov.as@phystech.edu

Афанасьев Александр Петрович²

apa@isa.ru

Хританков Антон Сергеевич^{1,3}

akhritankov@hse.ru

¹Долгопрудный, МФТИ

²Москва, ИППИ РАН

³Москва, НИУ ВШЭ

Внедрение машинного обучения в общественном масштабе требует понимания его долговременных эффектов, включая надежность, смещенность и соответствие требованиям безопасности. В этой работе вводится процесс многократного машинного обучения, который отражает эффекты петель обратной связи (feedback loop), усиления ошибок (error amplification) и связанные с этим явления. Процесс многократного машинного обучения характеризует сценарий, при котором входные данные для алгоритма машинного обучения зависят от предыдущих предсказаний, сделанных моделью.

Иллюстрацией [1] процесса многократного машинного обучения может служить алгоритм прогнозирования цен на жилье, который опирается на покупки, рекомендованные моделью. Аналогичными проблемами являются эхо-камеры (echo chambers) и пузыри фильтров (filter bubbles) [2,3]. Ensign и др. [4] исследовали эффект петли обратной связи, когда система прогнозирования расписания для полиции влияла на преступность.

В этой работе исследуется динамическая система, моделирующая процесс многократного машинного обучения и находится предельное множество распределений вероятностей для режимов работы системы с положительной и отрицательной петлей обратной связи. Этот процесс охватывает весь цикл получения данных, обучения модели и выдачи прогнозов конечным пользователям в рамках единой математической модели.

Определим следующую дискретную динамическую систему:

$$f_{t+1}(x) = D_t(f_t)(x), \forall x \in \mathbb{R}^n, t \in \mathbb{N} \text{ и } D_t \in \mathbb{D}, \quad (1)$$

где D_t обычно называется оператором эволюции, $f_t(x)$ — функции плотности вероятности распределения данных, а начальная функция $f_0(x)$ задана.

Теорема 1 (Предельное множество системы (1)). Для любой функции плотности вероятности $f_0(x)$, $x \in \mathbb{R}$ и дискретной динамической системы (1), если существует функция $g(x)$ из $L_1(\mathbb{R})$ и последовательность $\psi_t \geq 0$ такие, что $f_t(x) \leq \psi_t \cdot |g(\psi_t \cdot x)|$ для всех $t \in \mathbb{N}$ и $x \in \mathbb{R}$. Тогда, если $\psi_t \rightarrow +\infty$, то плотность $f_t(x)$ стремится к дельта-функции Дирака, $f_t(x) \xrightarrow{t \rightarrow +\infty} \delta(x)$ слабо. Если ψ_t сходится к нулю, то плотность $f_t(x)$ стремится к нулевому распределению, $f_t(x) \xrightarrow{t \rightarrow +\infty} \zeta(x)$ слабо.

В данной работе под нулевым распределением понимается предел функций плотности распределения белого шума. То есть нулевое распределение можно представить как равномерное распределение на интервале $(-R; R)$ при R стремящемся к $+\infty$.

Теорема 1 дает достаточные условия для возникновения петли обратной связи в системе (1). Согласно Теореме 1 можно предположить, что операторы D_t имеют вид

$$D_{1,t}(f_0)(x) = \psi_t \cdot f_0(\psi_t \cdot x), \quad \forall x \in \mathbb{R}^n \text{ и } \forall t \in \mathbb{N}, \quad (2)$$

где $D_{1,t}(\cdot) := D_t(D_{t-1}(\dots D_1(\cdot)\dots))$. Для операторов (2) можно сформулировать критерий автономности системы (1).

Теорема 2 (Критерий автономности). *Если операторы эволюции D_t динамической системы (1) имеют форму (2), то система автономна тогда и только тогда, когда*

$$\psi_{\tau+\varkappa} = \psi_\tau \cdot \psi_\varkappa, \quad \forall \tau, \varkappa \in \mathbb{N}. \quad (3)$$

Автономность — важное свойство любой динамической системы. Критерий (3) легко проверить, поскольку он означает, что $\psi_t = e^{kt}$ для некоторого $k \in \mathbb{R}$, т. е. $f_t(0) = C \cdot e^{kt}$ для некоторых $k, C \in \mathbb{R}$. В данной работе на языке Python разработан стенд проведения вычислительных экспериментов, симулирующий процесс многократного машинного обучения. Постановка эксперимента *скользящее окно* идентична [1], схема *обновление выборки* отличается тем, что на каждом шаге случайно выбирается элемент из всего набора данных с возвратом. По построению схема *скользящее окно* не является автономной, т. к. шаг t ограничен размером исходного датасета. Постановка *обновление выборки автономная*, т. к. на каждом шаге делаются одинаковые действия.

В эксперименте проверяется условие (3) для последовательности ψ_t — график $\ln(f_t(0))$ должен быть близок к прямой. Для примера решалась задача регрессии с линейной моделью без регуляризации, оптимизированная с помощью SGD с максимальным числом итераций — 50. Такая задача и модель выбраны для лучшей демонстрации эффекта многократного машинного обучения. Задача регрессии решается на синтетическом наборе данных (\mathbf{X}, \mathbf{y}) , где \mathbf{X} нормально распределены, и $\mathbf{y} = \mathbf{X} \cdot \theta^* + \varepsilon$, где $\varepsilon \sim \mathcal{N}(0, 1)$ — добавленный шум.

Как видно из рисунка 1, в случае включения всех предсказаний без шума (слева сверху) в выборку процесс обновления разбивается на два линейных сегмента (зеленый и оранжевый). Если же добавляемый при включении шум *adherence* велик (справа сверху) мы имеем один (зеленый) автономный сегмент. В постановке *скользящее окно* (внизу) заметна прямая линия только при t от 100 до 600.

Гипотезы об автономности двух систем оказались частично выполнены. Для схемы *обновление выборки* на графике (слева сверху) присутствуют две области автономности с разным показателем степени k . Постановка *скользящее окно* не автономна при $t > 600$. Проблема многократного машинного обучения важна,

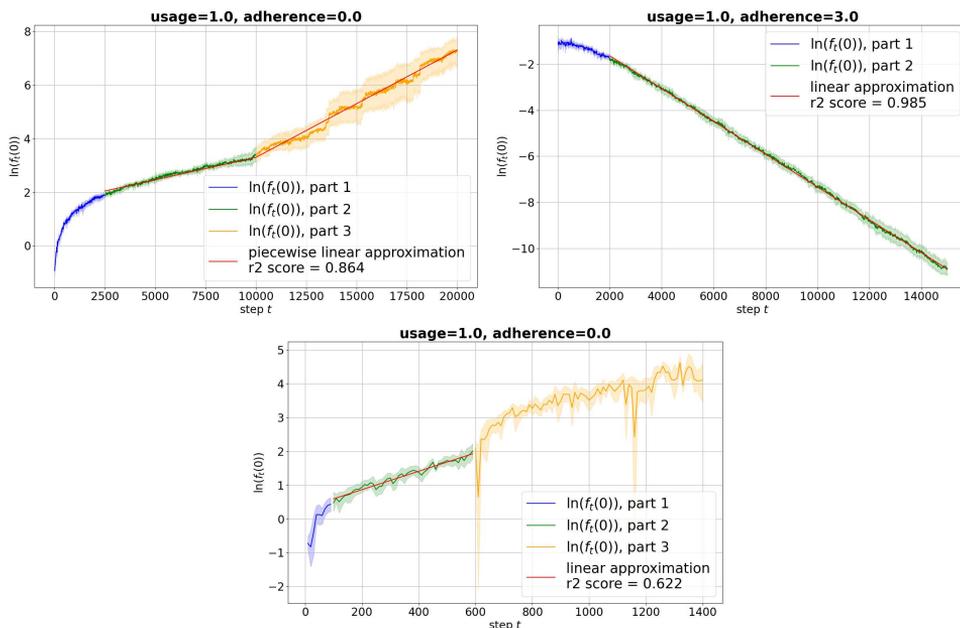


Рис. 1. Тестирование двух систем на автономность: обновление выборки (верхняя часть), скользящее окно (нижняя часть)

т. к. алгоритмы машинного обучения обычно функционируют в составе больших систем и взаимодействуют со своим окружением. В данной работе применяется аппарат дискретных динамических систем и математического анализа для построения математической модели. Все полученные результаты строго доказаны и проверены в вычислительном эксперименте.

- [1] *Khritankov A.* Positive Feedback Loops Lead to Concept Drift in Machine Learning Systems // *Applied Intelligence*. — Springer, 2023. — Vol. 53, No. 19. — P. 1–19.
- [2] *Davies H. C.* Redefining Filter Bubbles as (Escapable) Socio-Technical Recursion // *Sociological Research Online*. — 2018. — Vol. 23, No. 3. — P. 637–654.
- [3] *Spohr D.* Fake News and Ideological Polarization: Filter Bubbles and Selective Exposure on Social Media // *Business Information Review*. — London, England: SAGE Publications Sage UK, 2017. — Vol. 34, No. 3. — P. 150–160.
- [4] *Ensign D., Friedler S. A., Neville S., Scheidegger C., Venkatasubramanian S.* Runaway Feedback Loops in Predictive Policing // *Conf. Fairness, Accountability and Transparency*. — 2018. — P. 160–171.

Графовый метод полуконтролируемого обучения в задаче классификации периодических решений динамических систем в режиме реального времени

*Авдюшина Анна Евгеньевна*¹

avdiushina@itmo.ru

*Авдюшина Елена Владимировна*²

elena.v.a.2023@mail.ru

Ручкин Константин Анатольевич^{3*}

construchk@gmail.com

¹Санкт-Петербург, Национальный исследовательский университет ИТМО

²Донецк, Донецкий Государственный Университет

³Донецк, Донецкий Национальный Технический Университет

В данной статье рассматривается задача классификации периодических решений динамических систем. В предыдущих исследованиях [1] для этого используется графо-аналитический метод сечений Пуанкаре и его компьютерная реализация для построения траекторной структуры фазового пространства динамической системы на двумерных многообразиях. Эта двумерная структура состоит из периодических кривых и/или облаков точек. Построение траектории происходит из множества точек численного решения системы дифференциальных уравнений, для которой задаются начальные условия в режиме реального времени. Этот процесс ограничен во времени из-за ошибки накопления, которая является основной проблемой для полной автоматизации процесса классификации решений. В результате фиксированного временного периода на двумерных сечениях Пуанкаре получаются траекторные потоки, которые имеют разные свойства и должны быть классифицированы на разные классы.

Для решения этой задачи — классификации проекций траекторных потоков на двумерных многообразиях — в работе предлагается использовать графовый метод полуобучения на основе Лапласового подхода называемого Label Propagation [2]. Графовые методы полуобучения являются современным аппаратом для решения различных задач классификации, как двух так и более классов. Эта техника используется тогда, когда нельзя использовать методы полной классификации и структура пространства данных может быть описана в виде графа, что хорошо подходит для нашего случая.

В полуконтролируемом обучении на основе графов граф строится на основе размеченных и неразмеченных данных, использующих геометрические свойства структуры данных. Цель полуконтролируемого обучения на основе графов — получить хорошую производительность для классификатора, используя гораздо меньше маркированных данных. Одним из популярных методов полуконтролируемого обучения является использование уравнения Лапласа. Обучение по Лапласу означает, что у нас есть метки для первых значений, а затем эти метки распространяются дальше, используя гармоническую функцию, описанную уравнениями Лапласа. Уравнения Лапласа имеют вид

$$L(u) = 0, u = y$$

или

$$\sum_{j=1}^n w_{ij}[u(x_i) - u(x_j)] = 0, \quad u(x_i) = y(x_i), \quad i = \overline{1, m},$$

где x_i — вершины неориентированного взвешенного графа, $y(x_i)$ — начальные метки вершин графа, $u(x_i)$ — функция меток вершин графа, w_{ij} — вес ребра (x_i, x_j) , n — количество вершин графа, первые m из которых полагаются размеченными.

Лапласово обучение имеет и вариационную формулировку, приводящую задачу обучения к минимизации функции энергии Дирихле

$$E(u) = \frac{1}{2} \sum_{i,j=1}^n w_{ij}[u(x_i) - u(x_j)]^2.$$

Итерационное решение этой задачи может быть получено с помощью метода Якоби и сводится к виду

$$u^{k+1}(x_i) = \frac{1}{d_i} \sum_{j=1}^n w_{ij} u^k(x_j), \quad d_i = \sum_{j=1}^n w_{ij}.$$

В теории полубучения обучения это решение называется случайным блужданием на графах и описывает диффузию на графах. Оно имеет как преимущества, так и недостатки. Так, одним из недостатков является скорость его сходимости.

В данной работе предложена модификация итерационного решения случайного блуждания для ускорения процесса обучения. С помощью данного подхода и модификации задача классификации периодических решений динамической системы достаточно точно решается даже в режиме реального времени с учетом ошибки накопления и временных ограничений.

- [1] *Ruchkin C.* The General Conception of the Intellectual Investigation of the Regular and Chaotic Behavior of the Dynamical System Hamiltonian Structure // Applied Non-Linear Dynamical Systems. Springer Proc. Mathematics and Statistics. — Cham: Springer, 2014. <https://api.semanticscholar.org/CorpusID:115968327>
- [2] *Xiaojin Z.* Semi-Supervised Learning with Graphs. PhD thesis, Carnegie Mellon University, 2005.

Оценивание ошибки классификации в последовательном алгоритме распознавания объектов с учетом погрешностей исходных данных

Гагарин Юрий Евгеньевич¹*

gagarin_ye@bmstu.ru

Гагарина Светлана Николаевна²

g Ug@mail.ru

¹Калуга, КФ МГТУ им. Баумана

²Калуга, КГУ им. К. Э. Циолковского

В задачах распознавания образов решение о принадлежности распознаваемого объекта к соответствующему классу принимается после измерения всей совокупности признаков этого объекта. В последовательном алгоритме с получением очередного признака решается задача распознавания на основе данных об измеренных к текущему моменту признаках неизвестного объекта, и в зависимости от результатов сравнения полученного решения с некоторыми установленными заранее границами либо измеряется очередной признак объекта, либо прекращается дальнейшее накопление информации об этом объекте.

Последовательный алгоритм распознавания объектов заключается в том, что решение о принадлежности объекта к какому-либо классу принимается после измерения некоторого набора признаков. При распознавании множества объектов, разделенных на два класса: ω_1 и ω_2 , на каждом шаге последовательного алгоритма определяется отношение условных плотностей распределения вероятностей (УПВ):

$$\lambda_n = f_1(x_i)/f_2(x_i),$$

где x_i — набор признаков $i = \overline{1, n}$, $f_j(x_i)$ — УПВ распознаваемых объектов. Значение λ_n сравнивается с границами A и B , называемыми верхним и нижним порогами, которые определяются, исходя из значений ошибок первого и второго рода — Q_1 и Q_2 , соответственно. При этом наблюдения производятся до тех пор, пока выполняется условие $B < \lambda_n < A$. Наблюдения прекращаются, и принимается решение о принадлежности распознаваемого объекта классу ω_1 , если $\lambda_n \geq A$ и решение о принадлежности классу ω_2 , если $\lambda_n \leq B$. Верхний и нижний пороги A и B определяются из неравенств:

$$A \geq ((1 - Q_1)/Q_2), \quad B \leq (Q_1/(1 - Q_2)).$$

Исходные данные в задачах идентификации объектов являются результатами конкретных экспериментов и содержат случайные ошибки, которые необходимо учитывать. Учет погрешностей исходных данных приводит к необходимости учета погрешностей параметров УПВ и оцениванию неопределенности $\Delta f_j(x)$ [1]. С учетом этих неопределенностей отношение УПВ примет вид:

$$\lambda_n \pm \Delta \lambda_n = \frac{f_1(x_i) \pm \Delta f_1(x_i)}{f_2(x_i) \pm \Delta f_2(x_i)}$$

Погрешности УПВ окажут влияние и на ошибки первого и второго рода

$$Q_1 \pm \Delta Q_1 = \int_{G_2} (f_1(x_i) \pm \Delta f_1(x_i)) dx,$$

$$Q_2 \pm \Delta Q_2 = \int_{G_1} (f_2(x_i) \pm \Delta f_2(x_i)) dx,$$

где G_1 и G_2 — области в признаковом пространстве, соответствующие классам ω_1 и ω_2 .

При этом следует заметить, что области G_1 и G_2 также определяются по УПВ, а с учетом погрешностей УПВ и области в признаковом пространстве будут иметь некоторую неопределенность $G_1 \pm \Delta G_1$ и $G_2 \pm \Delta G_2$ [2]. В результате граница между классами не будет четко определена, а будет иметь место зона неопределенности, в которой нельзя достоверно идентифицировать объекты.

Формулы для определения ошибок первого и второго рода с учетом погрешностей УПВ и границ разделения классов будут иметь вид:

$$Q_1 \pm \Delta Q_1 = \int_{G_2 \pm \Delta G_2} (f_1(x_i) \pm \Delta f_1(x_i)) dx,$$

$$Q_2 \pm \Delta Q_2 = \int_{G_1 \pm \Delta G_1} (f_2(x_i) \pm \Delta f_2(x_i)) dx.$$

Определение интервальных оценок порогов и значения λ_n приводит к тому, что зона неопределенности, которая появилась в результате учета погрешностей областей признакового пространства, может увеличиться и для принятия решения потребуются дополнительный набор признаков. Однако, в то же время, учет неопределенностей значений признаков приводит к повышению достоверности принятия решений.

- [1] Гагарин Ю. Е. Гагарина С. Н. Прогнозирование показателей деятельности предприятий с учетом неопределенности исходных данных // Вестник университета. — М.: ГУУ, 2019. — С. 94–99.
- [2] Гагарин Ю. Е. Никитенко У. В. Степович М. А. Алгоритм последовательного распространения вероятностей с учетом интервальных оценок исходной информации // Современные методы теории функций и смежные проблемы. — Воронеж: ВГУ, 2021. — С. 89–90.

Применение генеративно-сопоставительного подхода с частичным привлечением учителя для решения задачи классификации с отрицательным классом

Максимова Александра Юрьевна¹

maximova.alexandra@mail.ru

¹Донецк, ФГБНУ «Институт прикладной математики и механики»

В работе рассматривается метод решения задачи классификации с отрицательным классом. Особенностью такой задачи является отсутствие явно заданного отрицательного класса, для которого нет примеров в обучающей выборке, однако на этапе эксплуатации встречаются объекты, которые не принадлежат ни одному из известных классов. Этим обусловлено отсутствие разметки у некоторых объектов в обучающей выборке.

Пусть задан алфавит известных классов Ω и обучающая выборка как множество объектов X в N -мерном признаковом пространстве. Каждому классу $\omega \in \Omega$ соответствует некоторое подмножество описаний $X_\omega \subseteq X$. Введем новый класс O , который будет соответствовать множеству признаков описаний $X_O = X \setminus \cup_{\omega \in \Omega} X_\omega$ и назовем его отрицательным классом. Таким образом, отрицательный класс является теоретико-множественным дополнением положительного класса до универсального множества допустимых признаков описаний.

Практическим примером задачи, формализуемой как классификация с отрицательным классом, является задача идентификационной экспертизы бензинов в лаборатории контроля качества нефтепродуктов. Необходимо определить производителя и марку исследуемого образца. Обучающее множество состоит из образцов бензинов разных марок и производителей, при этом для части примеров информация о принадлежности к классу отсутствует. Среди таких неразмеченных образцов могут присутствовать как образцы, принадлежащие известным классам, так и фальсификаты, которые можно рассматривать как примеры отрицательного класса.

Для решения данной задачи предлагается использовать модифицированную генеративно-сопоставительную архитектуру нейронных сетей GAN-SSL [1], в которой генератор G используется для генерации отрицательных примеров, отсутствующих в обучающей выборке. В качестве дискриминатора D выступает классификатор с $|\Omega| + 1$ выходами. Последний выход соответствует отрицательному классу и для примеров, которые не похожи ни на один из известных классов, его значение стремится к единице.

Наличие достаточно большого количества неразмеченных данных в обучающей выборке позволяет использовать обучение с частичным привлечением учителя (полуконтролируемое обучение). Обучающие примеры, для которых не известна принадлежность к классу, все же содержат полезную информацию об общем распределении данных в признаковом пространстве задачи.

Предложенная классификационная модель реализована на языке python с использованием библиотеки pytorch. При выполнении экспериментов в качестве архитектуры генератора и дискриминатора используются многослойные перспетроны. На вход генератор получает шум z из равномерного распределения p_z на интервале $[0, 1]$, а на выходе — примеры отрицательного класса, совпадающие по размерности с обучающими данными из распределения p_g . Используются нормализация по пакетам обучающих данных и нормализация весов перед выходным слоем. На вход дискриминатора поступают данные из распределения p_{data} . Выходами дискриминатора являются логиты, которые преобразуются в вероятность принадлежности классу функцией softmax.

Генератор и дискриминатор состоят из трех и шести слоев соответственно. Количество нейронов на скрытом слое каждой из нейронных сетей зависит от размерности задачи. В работе предлагается использовать один гиперпараметр N_{base} , по которому рассчитывается количество нейронов каждого слоя. Это означает, что количество нейронов каждого слоя генератора равно N_{base} , а количество нейронов скрытых слоев дискриминатора — $(2N_{base}, 2N_{base}, N_{base}, N_{base}, N_{base})$, соответственно, т.к. мощность дискриминатора в генеративно-сопоставительной архитектуре всегда больше.

Модель обучается методом стохастического градиентного спуска. Функция ошибки дискриминатора p_{model} является суммой двух функций ошибки для обучения с учителем и обучения без учителя, где выполнена декомпозиция стандартной функции кросс-энтропии на стандартную функцию обучения с учителем $L_{supervised}$ и обучения без учителя $L_{unsupervised}$:

$$L_{supervised} = -E_{x,y \sim p_{data}} \log[p_{model}(x, y < K + 1)],$$

$$L_{unsupervised} = -E_{x \sim p_{data}} \log[1 - p_{model}(y = K + 1|x)] - E_{x \sim p_g} \log[p_{model}(y = K + 1|x)],$$

где $K = |\Omega|$.

Эксперименты с моделью выполнялись на реальных данных задачи идентификационной экспертизы и на модельных данных. В результате выполненных экспериментов получены модели с качеством 94 % на реальной прикладной задаче и 97 % на модельной задаче. Модель хорошо и быстро сходится для задач со сбалансированными классами, представленными нормальным распределением. Для каждой задачи модель обучалась при разных значениях гиперпараметра N_{base} , в результате чего были выбраны лучшие значения: $N_{base} = 280$ и $N_{base} = 80$ для задачи идентификационной экспертизы и модельной задачи, соответственно. Эксперименты показали, что как слишком маленькие, так и слишком большие значения этого параметра приводят к падению качества распознавания.

Преимуществом предложенного подхода является возможность обучить классификационную модель в задачах, когда процент размеченных данных достаточно мал, количество размеченных данных может опускаться до 10 % от общего количества примеров.

- [1] *Salimans T., Goodfellow I., Zaremba W. et al.* Improved Techniques for Training GANs // NIPS'16: Proc. 30th Int. Conf. Neural Information Processing Systems December, 2016. — P. 2234–2242.

Сравнение оценок out-of-bag и скользящего экзамена

Неделько Виктор Михайлович¹★

nedelko@math.nsc.ru

¹Новосибирск, Институт математики им. С. Л. Соболева

Получение наиболее точных оценок качества решений досих пор является открытой проблемой машинного обучения [1, 2]. Основным инструментом оценивания качества решений является кроссвалидация, различные варианты которой, как правило, показывают сопоставимую точность [1]. При этом есть, как минимум, два способа построить итоговое решение по полной выборке: обучить модель заново и усреднить модели, построенные в процессе кроссвалидации. Вопрос, какой из способов лучше, до настоящего времени малоизучен.

Для простоты рассмотрим задачу бинарной классификации. Пусть X — пространство значений переменных, используемых для прогноза, $Y = \{0, 1\}$ — пространство значений прогнозируемой переменной, и $D = X \times Y$. Предположим, что все переменные являются случайными величинами с некоторым неизвестным совместным распределением $P(\cdot)$. Решающей функцией (алгоритмом классификации) называется соответствие $\lambda: X \rightarrow Y$.

Качество принятого решения оценивается заданной функцией потерь $\mathcal{L}: Y^2 \rightarrow [0, \infty)$. Будем использовать $\mathcal{L}(y, y') = \#(y \neq y')$ — индикатор ошибочной классификации. Под риском будем понимать средние потери:

$$R(\lambda) = E\mathcal{L}(y, \lambda(x)) = \int_D \mathcal{L}(y, \lambda(x)) P(dx, dy),$$

$x \in X, y \in Y$.

Пусть $Q: D^N \rightarrow \Lambda$ — метод (алгоритм) построения решающих функций, D^N — множество всех выборок объема N , Λ — заданный класс решающих функций.

Пусть $V = ((x^i, y^i) \in D \mid i = \overline{1, N})$ — случайная независимая выборка из распределения $P(\cdot)$, $V \in D^N$. Выражение $\lambda_{Q, V}$ будет обозначать функцию, построенную по выборке V методом Q .

Функционал скользящего экзамена определяется как

$$\check{R}(V, Q) = \frac{1}{N} \sum_{i=1}^N \mathcal{L}(y^i, \lambda_{Q, V'_i}(x^i)),$$

где $V'_i \subset V$, $(x^i, y^i) \notin V'_i$, — выборка, получаемая из V удалением i -го наблюдения (и, возможно, еще некоторых). Чаще всего скользящий экзамен используется в виде K -fold кроссвалидации, когда обучающая выборка разбивается на K непересекающихся подмножеств (фолдов) и при обучении тестируемый объект удаляется из выборки вместе со всем фолдом, которому он принадлежит.

Как правило, после проведения скользящего экзамена проводится обучение метода уже на всей выборке, и полученная решающая функция используется в качестве итогового решения. Однако в качестве итогового решения можно использовать решающие функции, уже построенные в ходе кроссвалидации, попросту объединив их в ансамбль. В обоих случаях оценкой качества выступает одно и то же выражение — функционал скользящего экзамена. Однако решающие функции, качество которых он оценивает, разные. Поэтому с формальной точки зрения это разные оценки. И во втором случае будем называть этот функционал оценкой *out-of-bag*. Таким образом, поставленное целью работы сравнение оценок *out-of-bag* и скользящего экзамена фактически означает сравнение двух способов построения итогового решения (а функционал качества в обоих случаях одинаковый). Заметим, что первоначально оценка *out-of-bag* [3] появилась в методе бэггинга (RandomForest), где она выглядит несколько иначе, чем мы определили, однако основная идея та же.

Чтобы оценить свойства рассмотренных оценок, требуется провести моделирование на большом количестве выборок. Это затруднительно сделать на реальных данных, поэтому зададим вероятностную модель. Пусть X_1, \dots, X_n — независимы и имеют равномерное распределение на интервале $[-1, 1]$. Функцию условной вероятности зададим как

$$P(y = 1 | x) = \varkappa + (1 - 2)\varkappa \cdot \#(x_1^2 + x_1x_2 + x_2^2 < 0,5),$$

где \varkappa — байесовский уровень ошибки (изменяемый параметр). В этой модели информативными являются две первые переменные, остальные — «шумовые».

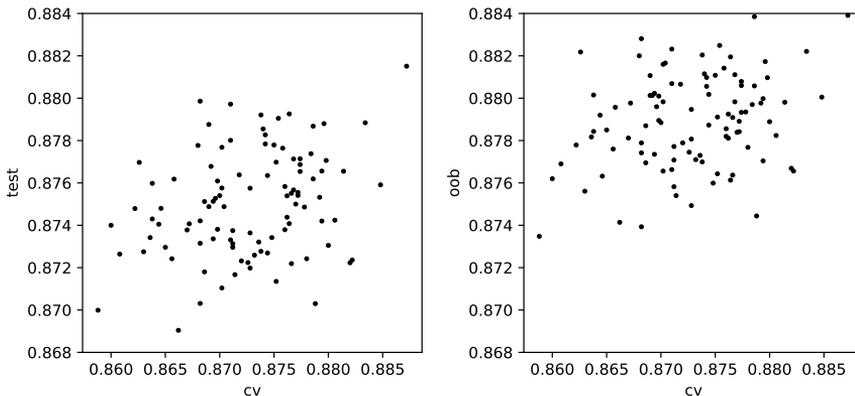


Рис. 1. Результаты экспериментов

На рисунке 1 приведен пример результата эксперимента на 100 выборках объема 5000 для метода градиентного бустинга (200 деревьев глубины 7, ско-

рость обучения — 0,05), $\varkappa = 0,1$, $n = 5$. По оси абсцисс отложены значения функционала скользящего экзамена, по оси ординат — значения риска (т.е. фактической точности на распределениях) для двух вариантов финального решения: путем нового обучения по всей выборке (слева) и путем ансамблирования решений из кроссвалидации (справа).

Объем публикации не позволяет привести таблицы результатов экспериментов, поэтому отметим только выводы. В среднем точность решения на основе ансамблирования получается несколько выше, чем при обучении заново. Различия статистически значимо. Любопытно, что корреляция риска с оценкой скользящего экзамена в обоих вариантах практически одинакова. Согласно проведенным экспериментам ансамблирование решений по кроссвалидации выглядит более привлекательным по сравнению с «классическим» вариантом обучения по всей выборке.

Работа выполнена в рамках госзадания Института математики им. С. Л. Соболева (проект № FWNF-2022-0015).

- [1] *Nedel'ko V. M.* On the Accuracy of Cross-Validation in the Classification Problem // Известия Иркутского государственного университета. Серия Математика. — 2021. — Т. 38. — С. 84–95.
- [2] *Nedel'ko V. M.* On the Properties of Bias-Variance Decomposition for kNN Regression // Известия Иркутского государственного университета. Серия Математика. — 2023. — Т. 43. — С. 110–121.
- [3] *Breiman L.* Out-of-bag estimation. — Citeseer, 1996. — 13 p.

Новая F -мера качества бинарных классификаторов

Копылов Андрей Валериевич

and.kopylov@gmail.com

Середин Олег Сергеевич*

oseredin@yandex.ru

Тула, Тульский государственный университет

Оценка качества работы различных классификаторов является одной из критических задач машинного обучения. В случае распознавания с учителем можно опираться на имеющиеся данные об «истинных» значениях классов объектов обучающей или тестовой выборки. Наиболее универсальным средством изучения статистических связей является таблица сопряженности, которая в применении к задачам машинного обучения принимает вид матрицы спутанности (Confusion Matrix). Для двухклассовой задачи — это матрица размера 2×2 (таб. 1). Попыткой свести все четыре элемента матрицы вместе и харак-

Таблица 1. Матрица спутанности

| | | Predicted | |
|--------|----------|-----------|----------|
| | | Positive | Negative |
| Actual | Positive | TP | FN |
| | Negative | FP | TN |

теризовать качество классификатора одним числом является широко известная мера F [1], представляющая собой гармоническое среднее между такими характеристиками как точность, понимаемая как доля верно распознанных объектов среди всех объектов, распознанных как положительные, и полнота, представляющая собой долю верно распознанных объектов в общем количестве позитивных объектов. Мера F часто подвергается справедливой критике по нескольким основным направлениям [2, 3]. Во-первых, данная мера не зависит от количества верно распознанных объектов негативного класса, и, во-вторых, мера зависит от того, какой из классов назван позитивным, а какой негативным, хотя это может быть оправданным далеко не во всех задачах. Ситуация еще более усугубляется при несбалансированных наборах данных, когда количество объектов позитивного и негативного класса значительно отличается друг от друга.

В данной работе мы предлагаем новый вариант меры F , который не зависит от того, какой из классов назван позитивным, а какой негативным, и способен давать корректные результаты при несбалансированной выборке.

Мы используем значения точности Pr и полноты Re как для случая, когда первый класс назван позитивным, так и для случая, когда позитивным назван другой класс, что дает матрицу спутанности представленную в таблице 2, используя обозначения $Pr^{(Inv)}$ и $Re^{(Inv)}$ соответственно. Таким образом, качество классификатора будет оценено четырьмя числами, выраженными через элемен-

Таблица 2. Матрица спутанности в случае, когда положительным назван второй класс

| | | Predicted | |
|--------|-------------------------------------|-------------------------------------|-------------------------------------|
| | | Positive ^(Inv) =Negative | Negative ^(Inv) =Positive |
| Actual | Positive ^(Inv) =Negative | $TP^{(Inv)} = TN$ | $FN^{(Inv)} = FP$ |
| | Negative ^(Inv) =Positive | $FP^{(Inv)} = FN$ | $TN^{(Inv)} = TP$ |

ты исходной матрицы спутанности:

$$Pr = \frac{TP}{TP + FP}, Re = \frac{TP}{TP + FN}, Pr^{(Inv)} = \frac{TN}{TN + FN}, Re^{(Inv)} = \frac{TN}{TN + FP}.$$

Мы придерживаемся идеи гармонического усреднения так же, как и в оригинальной мере F , и предлагаем в качестве скорректированной меры использовать среднее гармоническое значение четырех оценок. В результате получим следующее выражение:

$$F^{eq} = \frac{4}{\left(\frac{1}{Pr} + \frac{1}{Re} + \frac{1}{Pr^{(Inv)}} + \frac{1}{Re^{(Inv)}}\right)}.$$

В терминах исходной матрицы спутанности значение предлагаемой меры выражается следующим образом:

$$F^{eq} = \frac{4}{\left(\frac{1}{TP} + \frac{1}{TN}\right)(FP + FN) + 4}.$$

Заметим, что слагаемые в знаменателе представляют собой отношения общего количества ошибок классификации к количеству верно распознанных объектов положительного класса и к количеству верно распознанных объектов отрицательного класса. Теперь значение меры зависит от всех четырех параметров матрицы спутанности, включая TN и переименование положительного класса в отрицательный и наоборот не изменяет значения меры, что важно для задач в которых не оказывается предпочтения ни одному из классов. На рисунке 1) представлено сравнение меры F , предлагаемой меры F^c , а также двух альтернатив: меры F^* [4], представляющей собой монотонное преобразование меры F , и AGF [5] — F -меры, скорректированной с целью учета результатов классификации малонаполненного класса и TN .

В тех задачах, где можно выделить единственный представляющий интерес класс, использование меры F оказывается вполне оправданным, т. к. количество правильно распознанных объектов другого класса оказывается не важным, с тем замечанием, что классификатор, называющий любой объект объектом целевого (положительного) класса, будет иметь достаточно высокую оценку F , зависящую, впрочем, от сбалансированности выборки. Так, при одинаковом количестве объектов обоих классов $F = 0,667$, а если объектов положительного класса в два раза больше, чем отрицательного, то $F = 0,8$.

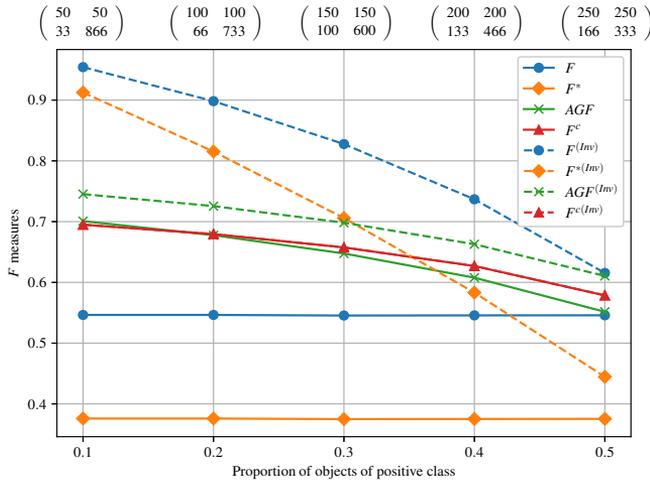


Рис. 1. Значения мер качества при распознавании 1000 объектов с $Pr = 0,6$, $Re = 0,5$ и изменяющейся долей объектов положительного класса для матриц спутанности, когда положительным назван первый класс (сплошные линии), и когда положительным назван второй класс (пунктирные линии)

Предложенная мера сочетает в себе несколько уникальных свойств с точки зрения задач классификации: позволяет учесть несбалансированность выборки и не отдает предпочтение ни одному из классов. В дополнение к этому данная мера обладает хорошей интерпретируемостью, а также простотой вычисления.

- [1] *Van Rijsbergen J. C.* Information Retrieval: Theory and Practice // Proceedings of the Joint IBM/University of Newcastle upon Tyne Seminar on Data Base Systems, 1979. — Vol. 79. — P. 1–14.
- [2] *Christen P., Hand D. J., Kirielle N.* A Review of the F-Measure: Its History, Properties, Criticism, and Alternatives // ACM Computing Surveys. — 2023. — Vol. 56, No. 3. — P. 1–24.
- [3] *Foody G. M.* Challenges in the Real World Use of Classification Accuracy Metrics: From Recall and Precision to the Matthews Correlation Coefficient // Plos one. — 2023. — Vol. 18, No. 10. — P. 1–27.
- [4] *Hand D. J., Christen P., Kirielle N.* F*: An Interpretable Transformation of the F-measure // Machine Learning. — 2021. — Vol. 110, No. 3. — P. 451–456.
- [5] *Maratea A., Petrosino A., Manzo M.* Adjusted F-Measure and Kernel Scaling for Imbalanced data Learning // Information Sciences. — 2014. — Vol. 257. — P. 331–341.

Многомерная модель динамического материального баланса в задаче мониторинга системы заводнения нефтяных месторождений

Тырсин Александр Николаевич^{1,2}

at2001@yandex.ru

Кащеев Станислав Евгеньевич^{3*}

kashcheevs@susu.ru

¹Екатеринбург, Научно-инженерный центр «Надежность и ресурс больших систем и машин» УрО РАН

²Екатеринбург, Уральский федеральный университет

³Челябинск, Южно-Уральский государственный университет

В настоящее время наблюдается тенденция переход большинства разрабатываемых нефтяных месторождений на завершающую стадию разработки. Это делает актуальной проблему поддержания приемлемых уровней добычи нефти путем оперативного управления заводнением. Огромные неопределенности гидродинамических моделей, связанные с отсутствием знаний о строении межскважинного пространства, с проявлением масштабных эффектов при переносе керновых данных на ячейки гидродинамической модели, а также с особенностями численного решения дифференциальных уравнений, описывающих пластовые процессы приводят к необходимости использования более простых математических моделей [1].

Одним из перспективных направлений сопровождения разработки нефтяных месторождений является использование емкостно-резистивных аналитических моделей CRM (Capacitance Resistive Model), представляющих собой аналитическое решение уравнения динамического материального баланса в виде временного ряда [2, 3]. Сопровождение разработки нефтяных месторождений включает в себя решение задач мониторинга системы заводнения и прогнозирования нефтедобычи. Для этого необходимо по имеющимся данным о приемистости нагнетательных и дебите добывающих скважин определить коэффициенты взаимовлияния и постоянные времени, и затем по найденным параметрам и планируемым значениям приемистости нагнетательных скважин сделать прогноз дебита жидкости добывающих скважин.

Дебит жидкости добывающих скважин в k -й период определяется [1] как:

$$q_l(k) = q_l(k-1) \exp\{-1/\tau_l\} + (1 - \exp\{-1/\tau_l\}) \sum_{j=1}^M f_{jl} I_j(k) + \varepsilon(k), \quad (1)$$

где $\tau_l > 0$ — постоянная времени; f_{jl} — коэффициент взаимовлияния нагнетательной скважины I_j и добывающей скважины P_l ; $I_j(k)$ — значение приемистости нагнетательной скважины I_j в k -й период; $\varepsilon(k)$ — компонента, описывающая случайные погрешности; k — номер периода (месяца); $l = \bar{1}, \bar{L}$, $j = \bar{1}, \bar{M}$, $k = \bar{1}, \bar{N}$. Выражение (1) является рекуррентным представлением модели CRM

в варианте CRMP (объем дренирования определяется в области «одна добывающая скважина — все влияющие нагнетательные скважины»).

В [4] описан алгоритм определения коэффициентов взаимовлияния скважин постоянных времени добывающих скважин. Анализ этого алгоритма показал его ограниченные возможности для решения реальных многомерных задач. Это вызвано тем, что объемы выборок ограничены 15–20 наблюдениями при достигающем сотен числе нагнетательных и добывающих скважин. Данный алгоритм рассчитан на обучающие выборки объема не менее 30 наблюдений и числа скважин порядка нескольких десятков (в рассмотренном примере было 7 нагнетательных и 18 добывающих скважин).

В предложенном решении удалось снизить объем обучающей выборки вдвое и довести число нагнетательных и добывающих скважин до 60 и 160, соответственно. Дополнительно удалось повысить точность определения коэффициентов взаимовлияния и постоянных времени, заменив приведенные значения приемистости нагнетательных скважин фактическими значениями.

Данных результатов удалось достичь благодаря задаче оценивания CRM модели. Задача оценивания состоит из двух этапов. Сначала производилась оценка параметров CRM модели для каждой добывающей скважины и с помощью энтропийного показателя взаимосвязи уточнялось множество взаимовлияющих скважин. Затем определялись параметры CRM модели для всей системы заводнения.

В работе реализовано несколько методов аппроксимации CRM модели. Данные решения реализованы на языке Python с использованием глобально сходящегося метода перемещения асимптот (ММА) для локальной оптимизации на основе градиента из библиотеки нелинейной оптимизации NLOPT [5]. Метод перемещения асимптот (ММА) является алгоритмом нелинейного программирования из семейства консервативных выпуклых сепарабельных аппроксимаций CCSA, который зарекомендовал себя в области структурной оптимизации, где вычисления функций и градиентов занимают много времени. Данный алгоритм обладает глобальной сходимостью и высокой скоростью работы [6]. Суть метода состоит в том, что в каждой точке $X(k)$ генерируется локальная аппроксимация с помощью градиента функции, условий ограничения и квадратичного штрафного слагаемого, чтобы сделать аппроксимацию консервативной.

Алгоритмы оказались устойчивы к случайным погрешностям ε . Так, при стандартном отклонении $\sigma_\varepsilon = 0,3$ и стандартном отклонении дебита скважин, равном 5,1, максимальное значение среднеквадратичной ошибки коэффициентов взаимовлияния составило $2 \cdot 10^{-4}$, а максимальное значение среднеквадратичной ошибки постоянных времени составило $5 \cdot 10^{-6}$.

В рассматриваемых примерах в качестве приемистостей нагнетательных скважин использованы фактические значения на одной из систем заводнения нефтяного месторождения.

- [1] *Степанов С. В., Бекман А. Д., Ручкин А. А., Поспелова Т. А.* Сопровождение разработки нефтяных месторождений с использованием моделей CRM. — Тюмень: ИПЦ «Экспресс», 2021. — 300 с.
- [2] *Данько М. Ю., Бриллиант Л. С., Завьялов А. С.* Применение метода динамического материального баланса и CRM-метода к подсчету запасов ачимовских и баженских коллекторов // Недропользование XXI век. — 2019. — Т. 80, № 4. — С. 76–85.
- [3] *Афанаскин И. В., Крыганов П. В., Глушаков А. А., Ялов П. В.* Использование CRM-моделей интерференции скважин для оценки фильтрационно-емкостных свойств пласта по данным разработки // Успехи кибернетики. — 2020. — Т. 1. № 1. — С. 17–27.
- [4] *Тырсин А. Н., Степанов С. В., Ручкин А. А., Бекман А. Д.* Повышение достоверности моделирования взаимовлияния скважин для анализа эффективности системы заводнения // Математическое моделирование. — 2020. — Т. 35, № 6. — С. 63–80.
- [5] *Johnson S. G.* The NLOpt Nonlinear-Optimization Package // URL:<http://github.com/stevengj/nlopt>, 2007.
- [6] *Svanberg K.* A Class of Globally Convergent Optimization Methods Based on Conservative Convex Separable Approximations // SIAM J. Optim. — 2002. — Vol. 12, No. 2. — P. 555–573.

Восстановление пропусков данных в многомерных временных рядах геофизических параметров методами статистики и математического моделирования.

Алешновский Валентин Сергеевич¹*

alesh-valentin@yandex.ru

Макеев Даниил Александрович¹

makeev05.06daniil2003@mail.ru

Газарян Варвара Арамовна¹

varvaragazaryan@yandex.ru

Чуличков Алексей Иванович¹

achulichkov@gmail.com

Шапкина Наталья Евгеньевна¹

neshapkina@mail.ru

Голубцов Петр Викторович¹

golubtsov@physics.msu.ru

¹Москва, МГУ имени М. В. Ломоносова, физический факультет

В контексте исследования многомерных временных рядов геофизических параметров ключевым аспектом является создание математической модели для изучения их структуры, предсказания будущих значений на основе имеющихся данных, и управления соответствующими процессами. Такое моделирование необходимо для понимания динамики процессов и адекватной интерпретации полученных результатов моделирования [1, 2].

В ходе работы измерительных устройств нередко возникают технические неисправности, из-за чего в ряде данных возможны пропуски на определенных интервалах времени, что негативно сказывается на точности построения математической модели и последующем анализе временных рядов. В связи с этим авторами предложены два метода восстановления отсутствующих данных и проведен сравнительный анализ для оценки эффективности этих методов [3].

Первый метод использует авторегрессионные алгоритмы *ARIMA* и ее расширенную версию *ARIMAX*, которая учитывает воздействие экзогенных переменных на временной ряд [3]:

$$\Delta^D y_t = \sum_{i=1}^p \varphi_i \Delta^D y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \sum_{m=1}^M \beta_m X_{m,t} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2), \quad (1)$$

где p — порядок авторегрессии, D — порядок интегрируемости, q — порядок скользящего среднего, ε_t — нормально распределенные остатки, φ_i , θ_j , β_m — параметры модели, Δ^D — оператор разности временного ряда порядка D , а последняя сумма учитывает M внешних параметров. Параметры модели подбираются из условий максимального согласия результатов наблюдения с моделью.

Во втором подходе к восстановлению пропущенных значений в многомерных временных рядах используются методы машинного обучения, а именно, рекуррентные нейронные сети (*RNN*), специально адаптированные для работы с последовательностями данных: *LSTM* и *GRU*, которые способны обучаться долгосрочным зависимостям [5].

На графике на рисунке 1 представлены исходный временной ряд концентрации углекислого газа на высоте 30 метров над уровнем земли и значения, вычисленные на основе моделей.

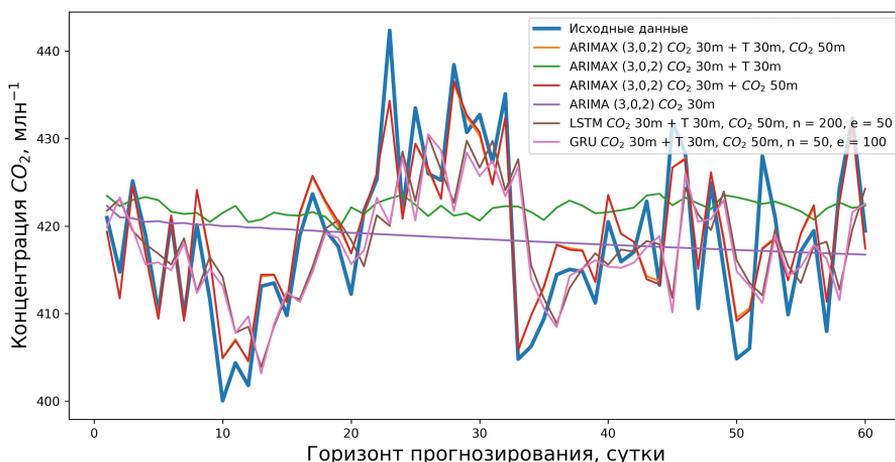


Рис. 1. Временные ряды

| Модель | MSE | R^2 | T , с |
|---|-------|-------|---------|
| ARIMAX (3,0,2) CO_2 30m + T 30m, CO_2 50m | 11,21 | 0,87 | 2,3 |
| ARIMAX (3,0,2) CO_2 30m + T 30m | 97,24 | -0,13 | 3,3 |
| ARIMAX (3,0,2) CO_2 30m + CO_2 50m | 11,41 | 0,87 | 1,13 |
| ARIMA (3,0,2) CO_2 30m | 90,03 | -0,05 | 2,9 |
| LSTM CO_2 30m + T 30m, CO_2 50m, $n = 200$, $e = 50$ | 72,20 | 0,16 | 388 |
| GRU CO_2 30m + T 30m, CO_2 50m, $n = 50$, $e = 100$ | 70,63 | 0,18 | 372 |

Таблица 1. Анализ метрик различных моделей для восстановления ряда CO_2 на высоте 30 метров.

Для сравнения моделей используются следующие критерии: среднеквадратическая ошибка (MSE), коэффициент детерминации (R^2) и время восстановления T .

Комплексный анализ всех моделей позволяет прийти к выводу, что при наличии сильной корреляции нескольких компонент многомерного временного ряда наиболее эффективно восстанавливает данные модель ARIMAX, при условии использования в качестве дополнительного параметра компоненты с высокой

корреляцией. В случае, когда нет возможности использовать подобные связи, хорошей альтернативой является модель *GRU*, основанная на рекуррентных нейронных сетях.

- [1] *Kurbatova J., Tatarinov F., Molchanov A. et al.* Partitioning of Ecosystem Respiration in a Paludified Shallow-Peat Spruce Forest in the Southern Taiga of European Russia // *Environ. Res. Lett.* — 2013. — Vol. 8, No. 4. — 045028.
- [2] *Box G., Jenkins G.* Time Series Analysis: Forecasting and Control. — San Francisco: Holden-Day, 1970.
- [3] *Hocke K., Kämpfer N.* Gap Filling and Noise Reduction of Unevenly Sampled Data by Means of the Lomb-Scargle Periodogram // *Atmos. Chem. Phys.* — 2009. — No. 9. — P. 4197–4206.
- [4] *Pirani M., Thakkar P., Jivrani P. et al.* A Comparative Analysis of ARIMA, GRU, LSTM and BiLSTM on Financial Time Series Forecasting // *IEEE Int. Conf. Distributed Computing and Electrical Circuits and Electronics.* — IEEE, 2022. — P. 1–6.
- [5] *Atabay F. V., Pagkalinawan R. M., Pajarillo S. D. et al.* Multivariate Time Series Forecasting Using ARIMAX, SARIMAX, and RNN-based Deep Learning Models on Electricity Consumption // *3rd Int. Informatics and Software Engineering Conf.* — IEEE, 2022. — P. 1–6.

Интеграция рекомендательных алгоритмов и решения задачи оптимизации для улучшения опыта туристов

Шекунов Михаил Андреевич¹*

baozorp@gmail.com

¹Москва, Институт проблем управления имени В. А. Трапезникова РАН

Многие посетители музеев сталкиваются с ограничением времени и не всегда могут насладиться всеми экспонатами в полной мере. Для решения этой проблемы разрабатывается современный подход, который объединяет три различных алгоритма рекомендации [1]:

- коллаборативная фильтрация — алгоритм анализирует предпочтения пользователей на основе исторических данных и находит схожие интересы между посетителями, после чего предлагает наиболее интересные объекты для каждого посетителя [2];
- рекомендательная система, основанная на сходстве объектов — алгоритм анализирует характеристики самих объектов (экспонатов) и находит другие объекты, схожие по характеристикам;
- рекомендательная система, основанная на расстоянии между объектами — метод оценивает схожесть объектов, исходя из их физического расположения в музее.

После получения рекомендаций от каждого из этих алгоритмов их результаты объединяются и используются в оптимизации. На этом этапе используется алгоритм VNS [3]. Он включает в себя две фазы: фазу «встряхивания» (shaking phase) и фазу локального поиска (local-search phase). В фазе встряхивания выбирается случайное решение из набора окрестностей. Основная цель — максимизировать количество посещенных экспонатов.

Этот подход комбинирует методы рекомендации с решением задачи оптимизации. В результате получается рекомендательная система, которая лучше адаптируется к потребностям и интересам посетителей музея.

В будущем планируется совершенствование данной системы и разработка дополнительных эвристических методов для улучшения точности и эффективности в решении общей задачи оптимизации.

- [1] *Burke R.* Hybrid Recommender Systems: Survey and Experiments // User Modeling and User-Adapted Interaction. — 2002. — Vol. 12. — P. 331–370.
- [2] *Su X., Khoshgoftaar T. M.* A Survey of Collaborative Filtering Techniques // Advances in Artificial Intelligence, 2009 [Электронный ресурс]. <https://downloads.hindawi.com/archive/2009/421425.pdf> (проверено: 11.11.2023).
- [3] *Campbell A. M., Gendreau M., Thomas B. W.* The Orienteering Problem with Stochastic Travel and Service Times // Annals of Operations Research. — 2011. — Vol. 186, No. 1. — P. 61–81.

Прогнозирование временных рядов с добавлением информации о стационарности в функционал потерь

Кравцова Ольга Анатольевна¹★

oakravts@gmail.com

¹Москва, МГУ им. М. В. Ломоносова

Анализ временных рядов имеет большое значение для проведения исследования данных в ретроспективе и в режиме реального времени, а также для прогнозирования. Временной ряд представляет собой последовательность измерений, сделанных в разные моменты времени. Например, в эпидемиологии, обычно измеряется количество новых случаев заболевания или нормализованный показатель заболеваемости.

Информация о стационарности играет важную роль в анализе временных рядов. Стационарность означает, что статистические свойства временного ряда остаются постоянными во времени. Если ряд стационарен, то его характеристики, такие как среднее значение и дисперсия, не зависят от времени. Информация о стационарности временного ряда является важным компонентом его анализа и использования в прогнозировании, моделировании и статистическом анализе [1].

Основной результат работы заключается в том, что при условии стационарности остатков модель является правильно специфицированной. Иными словами, если модель допускает ошибки, то эти ошибки в основном вызваны шумовой компонентой данных, которую невозможно предсказать с помощью математических методов [2].

При построении стандартных деревьев решений и их ансамблей, функционал потерь может быть изменен путем включения меры стационарности, которая представляет собой оценку статистики остатков, в процессе создания деревьев.

Зададим нестационарный временной ряд $Y = y_1, y_2, \dots, y_n$, прогнозируемый ряд $F = f_1, f_2, \dots, f_n$ и ряд остатков $R = R_1, R_2, \dots, R_n$, где $R_i = y_i - f_i$. Тогда остатки ряда в момент t запишем как: $R_t = \beta R_{(t-1)} + \varepsilon_t$, где ε — белый шум.

Пусть $f(Y) = F$ — регрессионный лес, с помощью которого строится прогноз. При генерации выборки для очередного дерева в ансамбле используется выборка с возвращением. В данных временных рядов наблюдения не являются независимыми друг от друга, т. е. текущие наблюдения зависят от их прошлых значений. Поэтому при работе с временными рядами используются блочные методы бутстреп.

Аналогично будем строить ансамбль LightGBM [3]. Наша задача состоит в том, чтобы минимизировать β . Это необходимо для того, чтобы наш ряд остатков R_i приближался к белому шуму.

В исследовании показывается, что итоговый функционал ошибок для модели Random Forest и ансамбля LightGBM с использованием следующего функцио-

нала ошибки:

$$F = \sum_{i=1}^n \frac{R_i^2}{n} + C \times |\alpha + 1|,$$

где $\alpha = \text{cov}(\Delta R, R)/D(R)$, $\Delta R_t = R_t - R_{t-1} = \alpha R_t + \varepsilon$, улучшает прогноз стандартной модели.

Для анализа использовались данные о заболеваемости COVID-19 в Российской Федерации за период с марта 2020 года по май 2023 года. Также для проверки утверждения использовались сгенерированные временные ряды. Величина прогнозного периода составляла от одной до трех недель, качество прогноза измерялось с помощью среднеквадратичной ошибки (MSE, Mean Square Error). Модель с использованием информации о стационарности остатков позволяет улучшить качество прогноза в сравнении со стандартной моделью в среднем на 15 % для Random Forest, и на 9 % для ансамбля LightGBM для данных о заболеваемости. Для случайных сгенерированных рядов ошибку прогноза относительно стандартной модели удалось уменьшить в среднем на 5 % для Random Forest и на 3 % для ансамбля LightGBM.

- [1] *Dickey D. A., Fuller W. A.* Distribution of the Estimators for Autoregressive Time Series with a Unit Root // J. American Statistical Association. — USA: Taylor & Francis, Ltd, 1979. — P. 427–431.
- [2] *Кравцова О. А.* Использование критериев стационарности для настройки моделей при прогнозировании временных рядов // Информатика и ее применения. — Москва: Информатика и ее применения, 2022. — С. 11–28.
- [3] *Ke G., Meng Q., Finley T. et al.* LightGBM: A Highly Efficient Gradient Boosting Decision Tree // Advances in Neural Information Processing Systems 30 (NIPS 2017). — P. 3149–3157.

Экспериментальная проверка вычислительной эффективности одной модели транспортных потоков

*Старожилец Всеволод Михайлович*¹★

starvsevol@gmail.com

*Чехович Юрий Викторович*¹

chehovich@forecsys.ru

¹Москва, Вычислительный центр им. А. А. Дородницына ФИЦ ИУ РАН

Целью данной работы является демонстрация применимости математической модели, описанной авторами в [1] для моделирования больших транспортных сетей за существенно меньшее в сравнении с микроскопическими моделями, на примере модели разумного водителя [2], время. Для достижения данной цели проведены эксперименты по моделированию конкретной транспортной магистрали — Московской кольцевой автомобильной дороги (МКАД). Ввиду того, что модель разумного водителя является одной из классических моделей, проводится сравнение результатов предложенной авторами модели с результатами микроскопической модели, а также показывается значительное преимущество предложенной модели по скорости вычислений.

На сегодняшний день моделирование крупных транспортных сетей представлено в [3, 4] в виде примеров применения существующих программных пакетов таких, как SUMO (Simulation of Urban Mobility), iTETRIS (“An Integrated Wireless and Traffic Platform for Real-Time Road Traffic Management Solutions”) и др. Детальное описание подхода к моделированию автомагистрали в данных пакетах зачастую отсутствует.

Моделирование транспортных потоков на автомагистрали тесно связано с задачей оптимизации светофорного управления в транспортной сети. Однако, в большинстве работ, посвященных светофорному управлению на перекрестке, не ставится задача обеспечения максимальной пропускной способности выделенной автомагистрали, и как следствие все дороги считаются равнозначными. В данной же работе ставится как раз задача обеспечения наиболее свободного проезда на выделенной ключевой магистрали.

В [1] авторами предложена модель, позволяющая компенсировать недостатки макроскопического и микроскопического моделирования, свойства которой были исследованы на модельных элементарных фрагментах транспортной сети. Предложенная модель оперирует в качестве объектов моделирования группами автомобильных транспортных средств (АТС). Скорость групп автомобилей рассчитывается с помощью фундаментальной диаграммы поток-плотность на магистрали [5] на основе комплексирования данных с дорожных датчиков и GPS-треков. Данные предположения привели нас к оригинальному мезоскопическому подходу, в котором мы оперируем группами АТС как элементарными объектами моделирования с помощью фундаментальных диаграмм поток-плотность. Такой подход позволяет быстро обчислять достаточно большие транспортные сети, в том числе такую магистраль, как МКАД, что необходимо для решения оптимизационных задач, для которых проводится моделирование.

Данный подход отличается от двух классических направлений к моделированию транспортных потоков, представленных микроскопическим подходом, основанным на моделировании движения каждого отдельного автомобиля [6], и макроскопическим подходом, опирающемся на сходство движения АТС с жидкостью или газом [7].

В данной работе авторы провели два модельных эксперимента. В первом был взят прямой участок МКАД, закрытый с двух сторон дорожными датчиками, проводится моделирование на основе данных с датчика на въезде на участок и проверяются на датчике на съезде. Результаты моделей сравниваются как между собой так и с данными с дорожных датчиков. Во втором — авторами построена модель одной из сторон МКАД с крупными въездами и съездами с нее. На основе имеющихся данных с дорожных датчиков на некоторых из въездов и статистических данных Центра организации дорожного движения генерируются модельные данные на въездах на автомагистраль двух типов: с утренней пиковой загрузкой и с вечерней, что соответствует большому потоку автомобилей в Москву и из Москвы. Промоделировано поведение автомагистрали с различными данными на въездах, и результаты сравниваются с результатами работы микроскопической модели разумного водителя посредством расчета средней абсолютной процентной ошибки (MAPE) между результатами моделирования на всей автомагистрали. Окончательно представлено время моделирования для предложенной мезоскопической модели и рассматриваемой микроскопической модели. Пример такого сравнения изображен на рисунке 1.

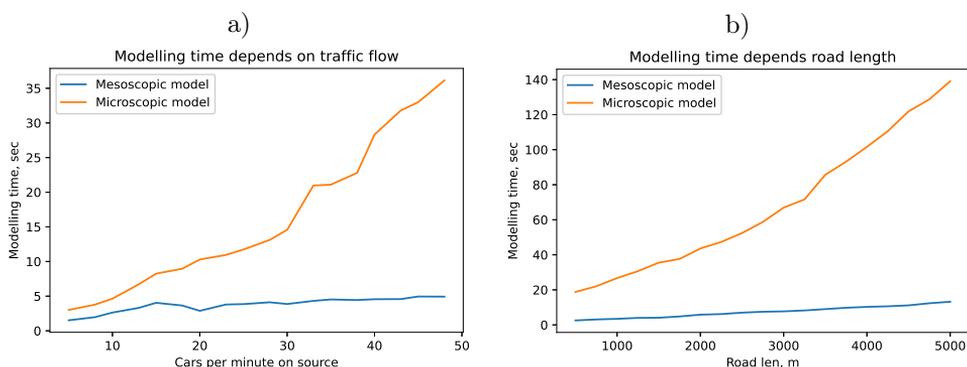


Рис. 1. а) Время моделирования 1 км магистрали в зависимости от потока АТС на въезде в диапазоне от 5 до 48 АТС/мин, б) Время моделирования в зависимости от длины моделируемого участка магистрали от 500 м до 5 км при входном потоке 45 АТС/мин

- [1] Старожилец В. М., Чехович Ю. В. Об одном подходе к статистическому моделированию транспортных потоков // Журнал Вычислительной математики и математической физики. — М.: РАН, 2021. — Т. 61, № 7. — P. 114–134.
- [2] Treiber M., Hennecke A., Helbing D. Congested Traffic States in Empirical Observations and Microscopic Simulation // Physical review E. — 2000. — Vol. 62. — P. 1805–1824.
- [3] Yuta A., Nobuyasu I., Hajime I., Tetsuo I., Uchitane T. Traffic Simulation of Kobe-City // Proc. Int. Conf. Social Modeling and Simulation, Plus Econophysics Colloquium 2014. — Berlin: Springer, 2015. — Vol. 229. — P. 255–264.
- [4] Bieker L., Krajzewicz D., Morra A., Michelacci C., Cartolano F. Traffic Simulation for All: A Real World Traffic Scenario from the City of Bologna // Modeling Mobility with Open Data. — Berlin: Springer, 2015. — Vol. 229. — P. 47–60.
- [5] Алексеенко А. Е., Холодов Я. А., Холодов А. С., Горева А. И., Васильев М. О., Чехович Ю. В., Мишин В. Д., Старожилец В. М. Разработка, калибровка и верификация модели движения трафика в городских условиях. Ч. I // Компьютерные исследования и моделирование. — Ижевск: Институт компьютерных исследований, 2015. — Т. 7, № 6. — С. 1185–1203.
- [6] Гасников А. В. и др. Введение в математическое моделирование транспортных потоков. — М.: Litres, 2015. — 89 с.
- [7] Whitham J. B. Linear and nonlinear waves. — Wiley, 1974. — 656 p.

Классификация геомагнитно-индуцированных токов в нейтрали автотрансформатора с применением машинного обучения

Аксенович Татьяна Валерьевна^{1*}

t.aksenovich@ksc.ru

*Селиванов Василий Николаевич*¹

v.selivanov@ksc.ru

¹Апатиты, Центр физико-технических проблем энергетики Севера — филиал Федерального государственного бюджетного учреждения науки Федерального исследовательского центра «Кольский научный центр Российской академии наук» (ЦЭС КНЦ РАН)

Во время геомагнитных бурь, являющихся следствием взаимодействия солнечного ветра с магнитосферой Земли, в протяженных электрических сетях на поверхности Земли начинают протекать геомагнитно-индуцированные токи (ГИТ). ГИТ становятся причиной целого ряда негативных явлений, влияющих на нормальную работу всей электроэнергетической системы. Примерами отрицательных эффектов ГИТ являются генерация четных и нечетных гармоник, увеличение потерь реактивной мощности, появление локальных перегревов обмоток и элементов конструкции трансформатора, генерация вибраций сердечника и обмоток трансформатора, нежелательное срабатывание защитных устройств [1,2]. Обычно длительность протекания токов не превышает нескольких минут, но из-за повторяющегося воздействия ГИТ на трансформаторы они могут выйти из строя. В прошлом было зафиксировано несколько случаев, когда во время сильных геомагнитных бурь протекание ГИТ привело к отключениям высоковольтных энергетических систем в Канаде и Швеции [3,4].

Для исследования природы этого явления и его возможных эффектов на трансформаторы, в 2011-м году была создана система мониторинга ГИТ, устройства которой были установлены на пяти подстанциях Кольско-Карельского транзита на Северо-Западе России [5]. С начала работы системы мониторинга ГИТ было создано более 16 000 файлов, содержащих данные токов в нейтральных автотрансформаторов. Общий размер этих файлов составляет около 27 ГБ на начало 2023-го года. Непрерывно растущий объем данных, содержащих записи токов в нейтрали, не может быть проанализирован вручную в разумные сроки.

Для автоматического анализа 12-летних данных, накопленных системой мониторинга ГИТ на Северо-Западе России, была разработана модель сверточной нейронной сети (СНС). Для обучения модели СНС был создан набор данных, содержащий 800 изображений скейлограмм двух классов: ГИТ и геомагнитно «спокойные» часы. В результате сравнения производительностей четырех СНС с разной архитектурой, была выбрана модель, которая показала высокую эффективность классификации ГИТ на валидационной выборке (100.00 % точность и 0,0115 — значение функции потери бинарной перекрестной энтропии).

В состав предложенной модели СНС, помимо основных слоев, вошли слои предобработки и слой дропаута.

Работа поддержана грантом РНФ № 22-29-00413.

- [1] *Molinski T. S.* Why Utilities Respect Geomagnetically Induced Currents // *J. Atmospheric and Solar-Terrestrial Physics*. — 2002. — P. 1765–1778.
- [2] *Rajput V. N., Boteler D. H., Rana N., Saiyed M., Anjana S., Shah M.* Insight into Impact of Geomagnetically Induced Currents on Power Systems: Overview, Challenges and Mitigation // *Electric Power Systems Research*. — 2021. — Vol. 192. — P. 106927.
- [3] *Guillon S., Toner P., Gibson L., Boteler D.* A Colorful Blackout: The Havoc Caused by Auroral Electrojet Generated Magnetic Field Variations in 1989 // *IEEE Power and Energy Magazine*. — 2016. — Vol. 14, No. 6. — P. 59–71.
- [4] *Pulkkinen A., Lindahl S., Viljanen A., Pirjola R.* Geomagnetic Storm of 29–31 October 2003: Geomagnetically Induced Currents and Their Relation to Problems in the Swedish High-Voltage Power Transmission System // *Space Weather*. — 2005. — Vol. 3.
- [5] *Barannik M. B., Danilin A. N., Kat'kalov Yu. V., Kolobov V. V., Sakharov Ya. A., Selivanov V. N.* A System for Recording Geomagnetically Induced Currents in Neutrals of Power Autotransformers // *Instruments and Experimental Techniques*. — 2012. — Vol. 55. — P. 110–115.

Новый параллельный алгоритм кластеризации больших данных Big-means с конкурентной стохастической оптимизацией размера выборки

*Мусабаев Рустам*¹

rustam@iict.kz

Мусабаев Равиль^{1*}

ravmus@uw.edu

*Кульдеев Нурсултан*¹

¹Алматы, Казахстан, Satbayev University

Кластеризация является ключевым компонентом анализа данных и машинного обучения, особенно в области обучения без учителя. Среди различных методов кластеризации широко используется критерий кластеризации с минимальной суммой квадратов (КМСК), также известный как критерий K-means. Этот критерий направлен на разделение данных на K кластеров для минимизации суммы квадратов евклидовых расстояний между точками данных и их центрами кластеров.

Критические препятствия возникают при применении алгоритмов КМСК к большим данным. Во-первых, многие алгоритмы КМСК не масштабируются с большими наборами данных, становясь непрактичными для них из-за слишком большой вычислительной сложности. Например, алгоритм K-means является вычислительно затратным в случае больших данных. Во-вторых, существует потребность в полном использовании современных высокопроизводительных вычислений, включая многоядерные процессоры и распределенные вычислительные платформы, для разработки параллельных алгоритмов КМСК, способных эффективно обрабатывать большие наборы данных. Однако создание таких алгоритмов включает преодоление проблем с разделением данных, балансировкой нагрузки и управлением коммуникационными издержками. Наконец, чувствительность алгоритмов КМСК к начальным конфигурациям центроидов часто приводит к субоптимальным решениям, особенно с учетом сложности и разнообразия больших данных. Существует явная потребность в новом алгоритме КМСК, который был бы эффективным, масштабируемым, способным обрабатывать большие данные, использовал бы современные вычислительные инфраструктуры и предлагал бы надежные решения для кластеризации независимо от начальной конфигурации центроидов. Такой алгоритм значительно улучшил бы нашу способность извлекать информацию из больших наборов данных в различных областях.

Основная цель данной работы — создать новый алгоритм, который преодолел бы все проблемы, возникающие при кластеризации больших данных в контексте кластеризации с минимальной суммой квадратов (КМСК). Наш новый алгоритм называется «Параллельный алгоритм кластеризации Big-means с конкурентной стохастической оптимизацией размера выборки». Этот алгоритм сочетает в себе конкурентную оптимизацию и использование стохастической выборки, реализованные на аппарате параллельных вычислений.

Алгоритм Big-means [1], являющийся ключевым звеном данной методологии, специально разработан для решения задач КМСК большого масштаба. Его упрощенный псевдокод представлен в алгоритме 1. Он работает, получая случайную выборку из набора данных на каждой итерации и кластеризуя ее с использованием метода K-means, при этом начальная кластеризация для первой выборки осуществляется с использованием K-means++. Последующие выборки инициализируются с использованием наилучшего решения, найденного на данный момент. Алгоритм неявно включает «процедуру встряхивания», которая соответствует генерации новой случайной выборки на каждой итерации, внося изменчивость и предотвращая застой в результатах кластеризации. Этот итерационный процесс продолжается до выполнения заранее определенного условия остановки, такого как лимит по времени обработки или максимальное количество обработанных выборок. Окончательным результатом является набор центроидов, которые достигли наилучшего значения целевой функции на протяжении всех итераций.

Algorithm 1: Высокоуровневый псевдокод кластеризации Big-means

```

1  Инициализировать центроиды  $C$  как пустые и установить  $\hat{f} \leftarrow \infty$ ;
2  Установить  $t = 0$ ;
3  while  $t < T$  do
4      Получить случайную выборку  $S$  фиксированного размера  $s$  из  $X$ ;
5      Реинициализировать вырожденные центроиды в  $C$  с использованием K-means++
        на  $S$ ;
6       $C_{\text{new}} \leftarrow$  K-means на  $S$ ;
7      Обновить  $C$  и  $\hat{f}$ , если новые центроиды  $C_{\text{new}}$  улучшают  $f$ ;
8      Увеличить  $t$ ;
9  end
10 return  $C$ ;
```

Масштабируемость алгоритма Big-means проистекает из его способности обрабатывать меньшие подмножества данных на каждой итерации, что значительно снижает его вычислительные требования. Использование случайных подмножеств и периодическая реинициализация центроидов для вырожденных кластеров помогает избежать попадания в субоптимальные решения. Экспериментальные результаты [1] свидетельствуют о том, что этот новый алгоритм более эффективен и лучше масштабируется с большими наборами данных по сравнению с традиционным методом K-means и его всевозможными вариациями, предлагая многообещающий подход для кластеризации больших данных. В настоящее время алгоритм Big-means является лучшим в области кластеризации больших данных [1].

Конкурентная оптимизация, которая является ключевым компонентом предложенного алгоритма, включает в себя множество решений или агентов (в данном случае, параллельных процессов или алгоритмов), конкурирующих за достижение наилучших результатов. Это соревнование создает динамичный и адаптивный процесс, постоянно стремящийся к лучшим решениям. В пред-

ложенном алгоритме различные параллельные процессы, называемые «работниками», соревнуются друг с другом. Каждый работник выбирает случайное подмножество точек из всего набора данных для выполнения кластеризации методом K-means. Размер оптимальной выборки динамически регулируется на основе производительности отдельных работников, что приводит к более эффективному процессу кластеризации. Краткий псевдокод предложенного подхода показан в алгоритме 2.

Algorithm 2: Высокоуровневый псевдокод параллельного алгоритма кластеризации Big-means с конкурентной стохастической оптимизацией размера выборки

```

1  Инициализировать параметры и переменные для каждого параллельного работника;
2   $L \leftarrow$  Пустой список;
3  while работники не достигли предела итераций do
4    for каждый параллельный работник  $w$  do
5       $s_w \leftarrow$  Случайное целое число между  $s_{min}$  и  $s_{max}$ ;
6      Выполнить кластеризацию Big-means на  $p$  итерациях с размером выборки  $s_w$ ;
7      if целевая функция выборки улучшается then
8        Обновить наилучшее решение на данный момент для работника  $w$ ;
9        Добавить  $s_w$  в список  $L$ ;
10     end
11   end
12 end
13 Анализировать распределение размеров в  $L$  для нахождения оптимального  $s_{opt}$ ;
14 foreach работник do
15   | Пересчитать целевую функцию выборки с  $s_{opt}$ ;
16 end
17  $C \leftarrow$  Центроиды от лучшего работника;
18 return  $C, s_{opt}$ ;

```

Эксперименты были проведены на 23 реальных больших наборах данных с использованием предложенного алгоритма и Big-means. Результаты показали, что предложенный алгоритм достиг значительно более высокой точности кластеризации и меньшего времени вычислений в целом по сравнению с Big-means. Это демонстрирует способность предложенного алгоритма балансировать качество кластеризации и эффективность конкурентной оптимизации и стохастической выборки.

Используя принципы параллельных вычислений, конкурентной оптимизации и стохастической выборки, предложенный алгоритм достиг передовых результатов в области кластеризации больших данных. Дальнейшие исследования в этом направлении могут помочь изучить расширения, такие как улучшенное исследование (exploration) и использование (exploitation) размеров выборок во время итераций.

Это исследование было финансировано Комитетом науки Министерства науки и высшего образования Республики Казахстан по гранту № BR21882268.

- [1] *Mussabayev R., Mladenovic N., Jarboui B., Mussabayev R.* How to Use K-means for Big Data Clustering? // Pattern Recognition. — 2023. — Vol. 137. — 109269.

Дистилляция знаний для увеличения устойчивости сверточных сетей к сдвигам изображений

Михайлапов Денис Иванович^{1*}

d.mikhailapov@g.nsu.ru

*Бериков Владимир Борисович*¹

berikov@math.nsc.ru

*Алямкин Сергей Анатольевич*²

s.alyamkin@enot.ai

¹Новосибирск, Институт математики им. С. Л. Соболева СО РАН

²Новосибирск, Механико-математический факультет НГУ

В настоящее время активно используется съемка с дронов и стационарных камер. Получаемые видеоданные необходимо обрабатывать в реальном времени. Однако такие данные сопровождаются шумами и сдвигами объектов на изображениях. Это меняет локальную картину изображения и затрудняет предсказание сверточной сети в задачах классификации и детекции.

Существующие подходы [1–3] опираются на добавление дополнительных слоев сети для обработки более глобального контента, что нивелирует локальные изменения на изображении. К примеру, в [1] авторы модифицируют сверточные блоки, добавляя в них гауссов фильтр с ядром фиксированного размера (по умолчанию размер ядра 3 на 3), названный BlurPool, что размывает резкие локальные изменения пикселей и увеличивает точность. Однако это увеличивает время работы сети, что затрудняет внедрение в системы реального времени.

Дистилляция знаний [4] — метод, позволяющий более эффективно обучать нейронные сети. В классической постановке модель-ученик (менее затратная, но менее точная модель) настраивается копировать ответы обученной модели-учителя (более затратной, но более точной модели). Данный подход часто используется для увеличения точности моделей и уменьшения затрат на вычисления.

В данной работе используется метод само-дистилляции, когда отдельная модель-учитель не требуется. Нейронная сеть обучается выдавать одинаковый ответ на аугментированных вариантах одних и тех же изображений, что улучшает устойчивость к искажениям данных.

Цель работы — увеличить устойчивость модели к сдвигам объектов при классификации изображений, не модифицируя исходную архитектуру. Для достижения данной цели были определены следующие задачи:

1. реализовать кратное контролируемое увеличение объема обучающих данных с использованием случайных преобразований (аугментаций);
2. реализовать метод само-дистилляции для уменьшения разницы предсказаний модели на аугментированных вариантах одной и той же картинке;
3. использовать предложенный метод для увеличения точности и устойчивости модели к сдвигам при разных параметрах и сравнить с существующими подходами.

В качестве архитектуры модели для классификации изображений используется ResNet18. Для определения качества метода использовались следующие метрики: *Consistency*, *Accuracy* и *Stability* [1].

- *Consistency* — метрика, измеряющая процент совпадения предсказаний модели на *двух случайно сдвинутых* вариантах одного и того же изображения.
- *Accuracy* — метрика, показывающая точность предсказаний модели.
- *Stability* — метрика, измеряющая процент совпадения предсказаний модели на *сдвинутых по диагонали* изображениях.

Разработанный метод позволяет увеличить точность модели (accuracy) на 2,99 % и устойчивость к случайным сдвигам (consistency) на 2,7 % на наборе данных CIFAR-100 в сравнении с обученной моделью ResNet18 без дистилляции знаний. Так же данный подход достиг показателей метрик, сравнимых с методом, использующим BlurPool.

Ниже представлена таблица иллюстрирующая зависимость метрик от разной конфигурации обучения модели. Здесь n — коэффициент кратного увеличения объема данных, λ_{distil} — коэффициент силы дистилляции.

Таблица 1. Влияние увеличения объема данных и дистилляции на показатели метрик

| | $n = 1$ $\lambda_{distil} = 0$ | $n = 2$ $\lambda_{distil} = 0$ | $n = 2$ $\lambda_{distil} = 3$ | $n = 3$ $\lambda_{distil} = 15$ |
|---------------|-----------------------------------|-----------------------------------|-----------------------------------|------------------------------------|
| consistency % | $85,98 \pm 0,29$ | $86,62 \pm 0,13$ | $87,53 \pm 0,17$ | $88,7 \pm 0,3$ |
| accuracy % | $76,11 \pm 0,17$ | $77,93 \pm 0,09$ | $78,32 \pm 0,08$ | $79,1 \pm 0,15$ |
| stability % | $69,26 \pm 0,24$ | $70,25 \pm 0,2$ | $70,11 \pm 0,11$ | $71,6 \pm 0,16$ |

Первый столбец — модель, обученная без применения дополнительных методов обучения, второй столбец — обученная с увеличением объема данных в 2 раза, третий столбец — обученная с применением дистилляции и увеличением объема данных в 2 раза и четвертый столбец — обученная с применением дистилляции и увеличением объема данных в 3 раза.

Видно, что увеличение объема обучающих данных улучшает показатели метрик итоговой модели. Так же это справедливо при применении дистилляции.

Результаты сравнения предложенного метода с методом BlurPool представлена в следующей таблице.

Таблица 2. Сравнение предложенного метода с методом BlurPool

| | BlurePool $n = 1$ | ResNet18 $n = 2$ $\lambda_{distil} = 0$ | ResNet18 $n = 2$ $\lambda_{distil} = 3$ | ResNet18 $n = 2$ $\lambda_{distil} = 15$ |
|---------------|-----------------------------------|---|---|--|
| consistency % | $87,91 \pm 0,25$ | $86,62 \pm 0,13$ | $87,53 \pm 0,17$ | $88,36 \pm 0,54$ |
| accuracy % | $77,61 \pm 0,18$ | $77,93 \pm 0,09$ | $78,32 \pm 0,08$ | $78,24 \pm 0,15$ |
| stability % | $70,39 \pm 0,1$ | $70,25 \pm 0,2$ | $70,11 \pm 0,11$ | $69,52 \pm 0,16$ |

Первый столбец — модель, модифицированная с помощью BlurPool и обученная без применения предложенного метода, второй столбец — обученная с увеличением объема данных в 2 раза, третий столбец — обученная с увеличением объема данных и применением дистилляции ($\lambda_{distil} = 3$) и четвертый столбец — с применением дистилляции ($\lambda_{distil} = 15$).

Обучение модели с использованием дистилляции позволило достичь сравнимых показателей метрики точности относительно модифицированной модели.

Работа выполнена в рамках государственного задания ИМ СО РАН (проект FWNF-2022-0015).

- [1] *Zhang R.* Making Convolutional Networks Shift-Invariant Again // Int. Conf. Machine Learning. — PMLR, 2019. — P. 7324–7334.
- [2] *Zou X., Fanyi X., Zhiding Yu, Yuheng L., and Yong J. L.* Delving Deeper into Anti-Aliasing in ConvNets // Int. J. Computer Vision, 2022.
- [3] *Anadi C., Dokmanic I.* Truly Shift-Invariant Convolutional Neural Networks // Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition. — 2021. — P. 3773–3783.
- [4] *Hinton G., Vinyals O., Dean J.* Distilling the Knowledge in a Neural Network // arXiv. — 2015. <https://arxiv.org/abs/1503.02531>

Влияние дистилляции моделей при обучении на дистиллированных данных

Баринов Никита Андреевич^{1*}
Филатов Андрей Викторович²

barinov.na@phystech.edu
andrei.filatov@skoltech.ru

¹Долгопрудный, МФТИ

²Москва, Сколковский институт науки и технологий

Дистилляция моделей (knowledge distillation) — способ обучения нейросетевых моделей машинного обучения, направленный на передачу знаний от масштабной модели-учителя к ощутимо более компактной модели-ученику. Первой статьёй, в которой можно встретить дистилляцию знаний, является [1], где предложили технику сжатия ансамбля моделей в одну модель.

Еще одной техникой по передаче знаний является дистилляция данных (dataset distillation) [2] — процедура создания небольшой синтетической выборки, которая агрегирует полезную информацию, хранящуюся в данных, и тем самым модель, обученная на синтетических данных, будет соответствовать точности модели, обученной на полном наборе.

В данной работе предлагается новый подход: одновременная дистилляция моделей и данных. Для этого мы при обучении маленькой модели на дистиллированных изображениях добавили дистилляцию от большой заранее обученной на полной выборке модели, тем самым частично скомпенсировав потерю информации при дистилляции данных.

Дистилляция данных. Пусть $\mathcal{D}_{real} = \{(\mathbf{x}_i, y_i)\}_{i=1}^N$ — исходная выборка. Наша задача — создать меньшую выборку $\mathcal{D}_{syn} = \{(\hat{\mathbf{x}}_i, \hat{y}_i)\}_{i=1}^M$, где $M \ll N$ и такую, что качество модели, обученной на ней, было бы сопоставимо с качеством при обучении на исходных данных. Таким образом, оптимизационная задача, к которой сводится дистилляция данных, выглядит так:

$$\mathcal{D}_{syn} = \arg \min_{\mathcal{D} \in \mathcal{D}_M} \mathcal{L}_{DD}(\mathcal{D}, \mathbf{f}, \mathcal{D}_{real}), \quad (1)$$

где \mathcal{D}_{syn} — дистиллированная выборка; а \mathcal{L}_{DD} — функция потерь дистилляции данных, зависящая от способа дистилляции; $\mathbf{f} \in \mathcal{F}$ — модель, используемая при получении дистиллированных изображений.

Дистилляция моделей. Предсказания обученной модели-учителя используются через дивергенцию Кульбака–Лейблера при обучении меньшей модели-ученика согласно формуле:

$$\begin{aligned} \mathcal{L}_{KD} &= \frac{1}{N} \sum_{i=1}^N \left(- \sum_{j=1}^K y_{ij} \log p_{ij} + \lambda D_{KL}(\mathbf{p}_i || \mathbf{q}_i) \right) = \\ &= \frac{1}{N} \sum_{i=1}^N \left(- \sum_{j=1}^K y_{ij} \log p_{ij} + \lambda \sum_{j=1}^K p_{ij} \log \frac{p_{ij}}{q_{ij}} \right), \end{aligned} \quad (2)$$

где \mathcal{L}_{KD} — функция потерь для дистилляции моделей; N — число объектов, а K — классов, представленных в обучающей выборке; $y_{ij} = 1$, если i -й объект принадлежит j -му классу, иначе — 0; p_{ij} — вероятности классов, предсказанные моделью-учеником, а q_{ij} — моделью-учителем. Коэффициент λ позволяет настраивать влияние учителя. Перед взятием D_{KL} сглаживаем распределения учителя и ученика с помощью температуры T :

$$p_i^T = \text{softmax}\left(\frac{z_i}{T}\right), \quad (3)$$

где z_i — логиты классов, предсказанные моделью-учеником. Формула для q_i^T выглядит аналогично.

Дистилляция моделей и данных. Мы предлагаем к обучению на дистиллированных данных небольшой модели (модели-ученика) добавить дистилляцию знаний от модели-учителя. Она обучилась на полной выборке и мы предполагаем, что знания, переданные большой моделью при обучении меньшей на дистиллированных данных, помогут достичь лучшего качества.

Пусть $\mathcal{L}(\mathcal{D}_{real})$ — функция потерь при обучении модели-учителя на полной выборке, тогда пусть \mathcal{F} — семейство функций модели-учителя, тогда обучение модели-учителя есть решение оптимизационной задачи:

$$\mathbf{f}^* = \arg \min_{\mathbf{f} \in \mathcal{F}} \mathcal{L}(\mathbf{f}, \mathcal{D}_{real}). \quad (4)$$

Теперь \mathbf{f}^* используется при обучении на дистиллированных изображениях согласно формуле (2). Пусть $\mathcal{L}_{KD} = \mathcal{L}_{KD}(\mathbf{f}^*, \mathbf{g}, \mathcal{D})$, где $\mathbf{g} \in \mathcal{G}$ — модель-ученик из семейства функций \mathcal{G} , \mathcal{D} — выборка, на которой происходит обучение. Тогда итоговая оптимизационная задача при добавлении дистилляции моделей к обучению на дистиллированных данных выглядит так:

$$\mathbf{g}^* = \arg \min_{\mathbf{g} \in \mathcal{G}} \mathcal{L}_{KD}(\mathbf{f}^*, \mathbf{g}, \mathcal{D}_{syn}), \quad (5)$$

где \mathbf{f}^* — модель-учитель, полученная по выражению (4), которая в режиме валидации предсказывает сглаженное распределение вероятностей классов.

В эксперименте использовались данные, дистиллированные методами MTT [3] и DATM [4], а также случайные подмножества полной выборки (RANDOM) размером 1, 10 и 50 изображений на класс (IPC). В качестве модели-учителя использовался ResNet18, в качестве модели-ученика — ConvNet. Модель-учитель училась так, что при валидации на тестовой подвыборке каждой выборки она показывала 78,8 % на CIFAR10 и 74,4 % на CIFAR100. При обучении на дистиллированных изображениях с добавлением дистилляции учителя использовался оптимизатор SGD и 50 эпох градиентного спуска. Температура и вес дистилляции были следующими: $T = 1$, $\lambda = 0,25$. Результаты, усредненные по 6 запускам, представлены в таблице 1. «—» означает отсутствие значения).

Таблица 1. Результаты обучения ConvNet с дистилляцией моделей и без нее (Acc_d и Acc_s , соответственно)

| Выборка | IPC | MTT | | DATM | |
|-----------|-----|----------------|----------------------------------|----------------|----------------|
| | | Acc_s | Acc_d | Acc_s | Acc_d |
| CIFAR-10 | 1 | $41,8 \pm 0,2$ | $43,2 \pm 0,3$ | — | — |
| | 10 | $58,7 \pm 0,3$ | $59,8 \pm 0,4$ | $42,3 \pm 0,1$ | $44,1 \pm 0,3$ |
| | 50 | $66,3 \pm 0,1$ | $67,1 \pm 0,2$ | $54,9 \pm 0,4$ | $56,6 \pm 0,2$ |
| CIFAR-100 | 1 | $20,2 \pm 0,2$ | $21,2 \pm 0,1$ | — | — |
| | 10 | $33,5 \pm 0,1$ | $34,8 \pm 0,3$ | $27,2 \pm 0,1$ | $28,1 \pm 0,2$ |
| | 50 | $39,5 \pm 0,2$ | $40,4 \pm 0,1$ | $38,3 \pm 0,2$ | $39,1 \pm 0,2$ |
| Выборка | IPC | RANDOM | | | |
| | | Acc_s | Acc_d | | |
| CIFAR-10 | 1 | $13,1 \pm 0,2$ | $14,1 \pm 0,4$ | | |
| | 10 | $26,9 \pm 0,3$ | $28,1 \pm 0,2$ | | |
| | 50 | $42,0 \pm 0,2$ | $42,8 \pm 0,2$ | | |
| CIFAR-100 | 1 | $11,0 \pm 0,4$ | $12,3 \pm 0,2$ | | |
| | 10 | $24,2 \pm 0,3$ | $25,5 \pm 0,4$ | | |
| | 50 | $33,8 \pm 0,3$ | $36,1 \pm 0,3$ | | |

Выводы. В этой работе мы добавили дистилляцию моделей к обучению на дистиллированных данных. Как показывают результаты, наблюдается заметный прирост качества в 1,5–2 %, причем с уменьшением числа классов он возрастает. Это говорит об эффективности нашего подхода и о возможности его применения к другим задачам машинного обучения.

[1] *Hinton G., Vinyals O., Dean J.* Distilling the Knowledge in a Neural Network // arXiv. — 2015. <https://arxiv.org/abs/1503.02531>

[2] *Wang T., Zhu J.-Y., Torralba A., Efros A. A.* Dataset Distillation // arXiv. — 2018. <https://arxiv.org/abs/1811.10959>

[3] *Cazenavette George, Wang T., Torralba A., Efros A. A., Zhu J.-Y.* Dataset Distillation by Matching Training Trajectories // Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2022. — P. 4750–4759.

[4] *Guo Z., Wang K., Cazenavette G., Li H., Zhang K., You Y.* Towards Lossless Dataset Distillation via Difficulty-Aligned Trajectory Matching // arXiv. — 2023. <https://arxiv.org/abs/2310.05773>

Разработка метода визуального внимания на основе нейронных сетей на графах

Захаров Алексей Александрович^{1*}

aa-zaharov@ya.ru

*Шамшин Максим Николаевич*¹

maxim5579@mail.ru

*Жизняков Аркадий Львович*¹

lvovich1975@mail.ru

¹Муром, Муромский институт (филиал) федерального государственного бюджетного образовательного учреждения высшего образования «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых»

Визуальное внимание представляет собой способность системы компьютерного зрения быстро находить нужную область в наблюдаемой сцене. Поток визуальных данных размером 100–1000 млн бит поступает в зрительную систему человека каждую секунду. Основная цель визуального внимания — использование наименее возможного количества информации для решения сложных высокоуровневых задач распознавания объектов [1]. В настоящее время методы глубокого обучения в области визуального внимания на многих наборах данных превосходят зрительные возможности человека. Однако результаты современных подходов сильно ухудшаются в реальных условиях наблюдения, а также при использовании ограниченного количества обучающих данных.

Разработан метод визуального внимания на основе нейронных сетей на графах и данных с неточной маркировкой. Структурные подходы в области компьютерного зрения основаны на описании объектов в терминах их частей и отношений между частями. Основная идея структурных методов — использование совокупности внутренних отношений в анализируемой сцене, которые позволяют рассматривать группу объектов как единое целое.

Графы являются удобной формой представления информации для задач компьютерного зрения. Уникальные возможности графов позволяют анализировать структурные отношения между объектами сцен. Это дает возможность собирать больше информации по сравнению с локальным анализом данных. Одним из основных направлений исследований является определение сверток графов на основе спектральной теории. Свертка графа может быть определена в спектральной области на основе преобразования Фурье.

Структурное описание сцены на основе нейронных сетей на графах позволяет восполнить недостаток аннотированных данных и, таким образом, повышает точность обнаружения значимых объектов при меньших затратах на создание обучающих наборов изображений.

Исследование выполнено за счет гранта Российского научного фонда № 23-21-00486, <https://rscf.ru/project/23-21-00486/>.

- [1] *Захаров А. А., Титов Д. В., Жизняков А. Л., Титов В. С.* Метод визуального внимания на основе ранжирования вершин графа по разнородным признакам изображений // Компьютерная оптика. — 2020. — Т. 44, № 3. — С. 427–435.

Применение нейронных сетей для определения меры отражательной симметрии фигуры на бинарных растровых изображениях

*Середин Олег Сергеевич*¹

oseredin@yandex.ru

Ляхов Даниил Викторович^{1*}

liakhov.daniil@mail.ru

*Ломов Никита Александрович*¹

nikita-lomov@mail.ru

¹Тула, Тульский государственный университет

Задача отражательной симметрии исследована достаточно хорошо. Представлено большое количество математических алгоритмов для поиска параметров отражательной симметрии [1–3]. Однако, при всей популярности подходов на основе глубокого обучения попыток использовать нейронные сети для оценки меры симметрии на бинарных растровых изображениях до сих пор не предпринималось.

В задаче автоматической оценки меры симметрии можно выделить две подзадачи: 1) математическая формализация меры симметрии, которая бы хорошо согласовывалась с интуитивным представлением человека о симметрии; 2) разработка эффективных алгоритмов для вычисления меры симметрии на основе заданной формализации и получение параметров преобразования исходной фигуры (например, отражения или вращения).

Мы будем рассматривать задачу оценки меры симметрии плоской фигуры A . В этом случае одной из общепотребимых оценок меры симметрии является оценка на основе меры Жаккара:

$$J(A_\gamma) = \frac{|A \cap T_\gamma(A)|}{|A \cup T_\gamma(A)|}.$$

где $T_\gamma(A)$ — преобразование из множества допустимых преобразований T , заданное набором параметров γ из соответствующего множества Γ . Требуется найти преобразование, обеспечивающее максимальное значение меры, которое и будет считаться итоговой мерой симметрии фигуры:

$$\hat{J}(A) = J(A_{\hat{\gamma}}), \quad \hat{\gamma} = \arg \max_{\gamma \in \Gamma} J(A_\gamma).$$

Таким образом, ключевым пунктом постановки задачи является задание множества допустимых преобразований T и способа их параметризации.

В качестве такого преобразования мы будем использовать отражение фигуры относительно прямой, заданной уравнением $kx + b$, где $k = \tan(\varphi)$, φ — угол между прямой и осью X .

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos^2 \varphi - \sin^2 \varphi & 2 \cos \varphi \sin \varphi & -2b \cos \varphi \sin \varphi \\ 2 \cos \varphi \sin \varphi & \sin^2 \varphi - \cos^2 \varphi & b + b(\cos^2 \varphi - \sin^2 \varphi) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Таким образом, задачу можно разделить на поиск оптимальной прямой, отражение фигуры относительно которой максимизирует меру симметрии, и определение максимальной меры симметрии. В данной работе мы будем использовать нейронные сети только для определения максимальной меры симметрии. Возможность определения оптимальной прямой с помощью нейронных сетей будет рассмотрено в следующих работах.

Мы будем следовать сложившемуся и уже ставшему вполне стандартным подходу к современному анализу изображений с использованием ИНС:

1. Выбор корпуса изображений.
2. Экспертная разметка изображений.
3. Выбор функции потерь, архитектуры сверточной нейронной сети.
4. Проведение экспериментов с целью подбора параметров ИНС.
5. Статистическая оценка качества результатов.

В качестве корпуса данных для обучения ИНС и проведения экспериментов мы используем 4 базы данных. Первая — широко известная база листьев растений FLAVIA. База содержит 1907 изображений размером 800 на 600 пикселей. Вторая база данных — база MPEG-7 CE Shape-1 Part B, которая содержит 1401 изображение. Третья база данных — база Kimia Conv, которая содержит 220 изображений. Наконец, четвертая база — база Chicken Pieces Silhouettes Database, содержащая 446 изображений. На всех изображениях пиксели, принадлежащие фону, кодируются нулями, а принадлежащие фигуре — единицами.

В качестве оценки степени симметричности каждого изображения мы используем процедуру полного перебора прямых, проходящих через все пары точек контура фигуры, с построением отраженной копии фигуры и вычисления меры Жаккара. Все изображения из разных баз отличаются по своим оригинальным размерам. Обеспечить обработку изображений оригинальных размеров с помощью ИНС оказалось невозможным. Поэтому все изображения были приведены к размерам 200 на 150 пикселей.

Каждое изображение в корпусе наделяется меткой — действительным числом из интервала $[0, 1]$, что приводит к классу задач по восстановлению регрессионных зависимостей.

В качестве функции потерь мы выбрали среднеквадратичную ошибку MSE, однако контроль значений производили на основе средней абсолютной величин отклонений MAE. Такая метрика более привычна с точки зрения ее восприятия человеком.

Мы использовали следующую архитектуру сверточной нейронной сети.

1. Конволюционный слой для двумерных данных (32 фильтра с функцией активации Relu с ядром размера 5 на 5).
2. Слой пулинга MaxPooling для двумерных данных (окно просмотра размера 4 на 4, с шагом перемещения равным окну просмотра).

3. Конволюционный слой для двумерных данных (32 фильтра с функцией активации Relu с ядром размера 10 на 10).
4. Слой пулинга MaxPooling для двумерных данных (окно просмотра размера 4 на 4, с шагом перемещения равным окну просмотра).
5. Конволюционный слой для двумерных данных (32 фильтра с функцией активации Relu с ядром размера 20 на 20).
6. Слой пулинга MaxPooling для двумерных данных (окно просмотра размера 4 на 4, с шагом перемещения равным окну просмотра).
7. Слой преобразования многомерных данных в одномерные (Flatten).
8. Полносвязный слой (16 нейронов с функцией активацией Relu).
9. Полносвязный слой (1 нейрон с функцией активации SoftSign).

Для 1-го, 3-го, 5-го, 8-го и 9-го слоев используется инициализация весов Ксавье для исключения стохастического результата обучения сети при нескольких запусках. Для слоев 1–6 мы используем рамку накраях изображения, заполненную нулями.

Для проведения экспериментов весь корпус данных был разделен на два набора (для обучения и для теста) в соотношении 9 к 1, а также на 25 наборов для проведения кросс-проверки. В результате проверки в эксперименте с двумя наборами оценка MAE на обучающем множестве составила 0,00191, на тестовом множестве — 0,03513. При проведении кросс-проверки такая оценка составила 0,03383.

В докладе приводится подробное описание результатов экспериментов, а также алгоритм аугментации исходных данных и результаты экспериментов на аугментированной базе.

Дальнейшее исследование связано с повышением точности и решением задачи о нахождении оси симметрии. Остается открытым вопрос об использовании ассиметричной функции потерь, существенно сильнее штрафующей превышение предсказанного значения над эталонным.

Исследование выполнено при поддержке гранта Российского научного фонда № 22-21-00575, <https://rscf.ru/project/22-21-00575/>.

- [1] *Kushnir O., Fedotova S., Seredin O., Karkishchenko A.* Reflection Symmetry of Shapes Based on Skeleton Primitive Chains // 5th Int. Conf. AIST 2016, Yekaterinburg, Russia, April 7–9, 2016, Revised Selected Papers, CCIS. — Springer International Publishing Switzerland, 2017. — Vol. 661. — P. 293–304.
- [2] *Lomov N., Seredin O., Kushnir O.* Detection of the Optimal Reflection Symmetry Axis with the Jaccard Index and the Radon Transform // International Russian Automation Conference (RusAutoCon), Sochi, Russian Federation, 2022. — P. 489–498.
- [3] *Ломов Н. А., Середин О. С., Кушнир О. А., Ляхов Д. В.* Определение оптимальной оси отражательной симметрии с точки зрения меры Жаккара методом проекций // Труды Международной конференции по компьютерной графике и зрению «Графикон». — 2022. — № 32. — С. 715–727.

Применение причинно-следственных графов при тестировании инвариантами

Якушева Софья Федоровна^{1*}

yakusheva.sf@phystech.edu

Хританков Антон Сергеевич^{1,2}

akhritankov@hse.ru

¹Долгопрудный, Московский физико-технический институт

²Москва, НИУ ВШЭ

В наше время системы доверенного искусственного интеллекта становятся все популярнее. Такие системы напрямую влияют на своих пользователей, например, давая рекомендации, и поэтому нуждаются в тщательной проверке на удовлетворение требованиям. Однако формализовать требования к системе не всегда просто. Кроме того, для таких систем характерно наличие проблемы тестового оракула [1]. Она состоит в том, что невозможно по тем или иным причинам получить функцию для определения правильности ответа программы. В области машинного обучения эта проблема часто возникает из-за дороговизны сбора и разметки тестовых данных.

Одним из эффективных методов тестирования при наличии проблемы тестового оракула является тестирование инвариантами (metamorphic testing) [2]. При таком тестировании не проверяется правильность каждого конкретного ответа программы, вместо этого проверяется выполнение тестового инварианта. Тестовый инвариант — это функция от нескольких входов и соответствующих им выходов программы

$$R(x_1, x_2, \dots, x_n, f(x_1), f(x_2), \dots, f(x_n)) \longrightarrow \{0, 1\}, \quad (1)$$

где $n \geq 2$ — общее количество входных значений, x_i — i -е входное значение, а $f(x_i)$ — i -е выходное значение.

Одним из недостатков тестирования инвариантами является трудность обнаружения местонахождения ошибки в коде, особенно в случае многокомпонентных систем.

Объект тестирования. В [3] методом тестирования инвариантами была протестирована многокомпонентная система искусственного интеллекта для генерации персонализированных стикеров. Персонализированный стикер — это тематическая картинка, содержащая лицо человека, декоративные элементы и хвалебные надписи. Система принимает на вход фотографию человека и текстовое описание тематики, а возвращает готовый стикер или ничего, если создать стикер не удалось.

Данная система состоит из глубокой нейронной сети для сегментации лица (MediaPipe), нейронной сети для определения темы текстового описания (SpaCy), дополнительных инструментов для вставки объектов на картинку, мультипликации (AnimeGAN), добавления очков и переноса стиля. Тематические фоны и декоративные элементы хранятся в облачном хранилище.



Рис. 1. Пример исходного изображения и тематического стикера по теме «программирование»

Проверялось, что система удовлетворяет следующему требованию: пользователь должен узнавать свое лицо на стикере. Это требование было сформулировано как тестовый инвариант: *если по фотографии удалось сгенерировать стикер, то из этого стикера всегда можно получить новый стикер* (будем называть его *повторным*). Этот инвариант почти всегда нарушался, однако причина этого была не ясна.

Метод. Для выявления причин тех или иных событий применяются причинно-следственные графы [4]. Была поставлена задача проверить применимость этого метода для выявления причин нарушения тестовых инвариантов в многокомпонентной системе.

На основании эмпирического опыта и архитектуры системы был построен граф причинно-следственных связей (рис. 14). Граф включает в себя преобразования, применяемые в процессе получения стикера и повторного стикера, которые потенциально могут привести к искажению лица на стикере. К таким преобразованиям были отнесены мультипликация (*cartoon*), перенос стиля (*style*), добавление очков (*glasses*). Кроме того, рассматривалось влияние темы (*sphere*), которая является контролируемым входным параметром этого графа.

Параметры *face-detected* и *face-detected2* отвечают за факт распознавания лица на входном изображении и стикере, соответственно, и являются наблюдаемыми. При этом правильность каждого распознавания проверяется в силу проблемы тестового оракула. Выполнение инварианта *mr* зависит от *face-detected* и *emph*.

По графу была построена байесовская сеть для оценки влияния факторов.

Результаты. В результате вычислительного эксперимента была показана разная степень влияния преобразований, входных параметров и других факторов на итоговый результат. Это также свидетельствует о применимости метода

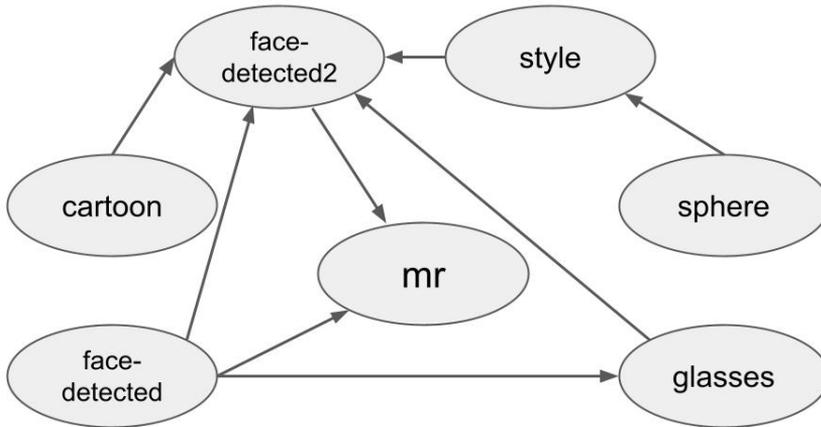


Рис. 2. Причинно-следственный граф системы.

причинно-следственных графов к определению источника ошибок при тестировании инвариантами.

- [1] Barr E. T., Harman M., McMinn P., Shahbaz M., Yoo S., The Oracle Problem in Software Testing: A Survey // IEEE Transactions on Software Engineering. — IEEE, 2015. — Vol. 41, No. 5. — P. 507–525.
- [2] Chen T. Y., Kuo F.-C., Liu H., Poon P.-L., Towey D., Tse T. H., Zhou Z. Q., Metamorphic Testing: A Review of Challenges and Opportunities // New York: Association for Computing Machinery. — 2018. — Vol. 51, No. 1. — P. 1–27.
- [3] Якушева С. Ф., Хританков А. С., Горбаченок Д. И. Тестирование инвариантами в применении к генеративным системам искусственного интеллекта // Системная информатика. — 2023. — № 22. — P. 37–44.
- [4] Domshlak C., Brafman R. I. Structure and Complexity in Planning with Unary Operators // Proc. Sixth Int. Conf. Artificial Intelligence Planning and Scheduling (AIPS 2002). — AAAI Press, 2002. — P. 34–43.

Обобщенная ортогональная регуляризация сверточных нейронных сетей

Григорьев Алексей Дмитриевич^{1*}

grigorev.ad@phystech.edu

Гнеушев Александр Николаевич^{1,2}

gneushev@ccas.ru

¹Москва, Московский физико-технический институт (НИУ)

²Москва, Федеральный исследовательский центр «Информатика и управление» РАН

Способность к обобщению на новые данные является ключевым свойством любой модели глубокого обучения для ее применимости в реальных приложениях. Регуляризация является общепринятым методом повышения обобщающей способности модели [1] путем наложения дополнительных ограничений на ее параметры, например, условия ортогональности весов.

В работе рассматривается задача регуляризации параметров нейронной сети с целью уменьшения избыточности ее параметров и повышения устойчивости модели. Ортогональная регуляризация является эффективным методом обеспечения обобщаемости модели. Однако полагается, что ортогонализация препятствует эффективной оптимизации параметров, налагая чрезмерно сильное ограничение, которое может быть ослаблено. В данной работе предлагается подход к обобщению ортогональной регуляризации, который основывается на построении фрейма в пространстве параметров модели и позволяет упростить обучение модели и повысить ее обобщающую способность. Фреймовая регуляризация вводится для увеличения разнообразия параметров модели. Показано, что ортогональная регуляризация является частным случаем предложенного метода регуляризации на основе функциональных фреймов.

Вычислительные эксперименты, проведенные на выборках CIFAR10, CIFAR100 и ImageNet, показали, что предложенный метод превосходит альтернативные подходы к регуляризации с точки зрения качества классификации и устойчивости к смене домена. Проведен анализ свойств параметров модели, индуцируемых различными методами регуляризации и указывающих на превосходство предложенного подхода к регуляризации. Выделены перспективные направления дальнейших исследований.

- [1] *Гнеушев А. Н., Григорьев А. Д., Матвеев И. А.* Фреймовая регуляризация сверточной нейронной сети в задачах классификации изображений // Известия РАН. Теория и системы управления. — 2022. — № 6. — С. 150–160.

Применение технологии динамического отслеживания и семантической сегментации для визуальной навигационной системы мобильных беспилотных объектов

Шэнь Синь^{1*}

Shenxin@yandex.ru

*Цзинь Цзюнь*¹

tszints@student.bmstu.ru

*У Го*¹

ug@student.bmstu.ru

*Лукьянов Вадим Викторович*¹

vdmlknv@yandex.ru

¹Москва, Московский государственный технический университет имени Н. Э. Баумана

В работе приведено описание модифицированного алгоритма SLAM, позволяющего определять параметры траектории автомобиля в сцене с большим количеством автомобилей. Приведены экспериментальные результаты тестирования алгоритма ORB-SLAM. Представлена модифицированная версия алгоритма ORB-SLAM, адаптированная для работы в условиях, когда в сцене присутствуют движущиеся объекты: пешеходы, машины и т. д. Была создана и предложена полноценная визуальная навигационная система для беспилотных мобильных устройств, использующая RGB-камеру в качестве основного датчика для сбора изображений или видео окружающей среды автономного транспортного средства.

В современном мире активно развиваются технологии, связанные с контролем за автономными транспортными системами. Основной проблемой для создания систем управления беспилотными автомобилями является разработка автономной навигации. Одним из актуальных направлений развития наземной мобильной робототехники является совершенствование управления мобильными роботами, включая навигацию в условиях неопределенной и динамически меняющейся внешней среды [1]. Для достижения высокого уровня автономности необходимо максимально эффективно использовать все доступные информационные каналы о внешнем мире при решении задач навигации, в том числе и семантическое представление данных. Одним из многообещающих подходов к решению этой проблемы является использование алгоритмов визуального SLAM (одновременная локализация и картографирование), которые опираются на данные с видеокамер для определения местоположения автономного транспортного средства.

В алгоритме SLAM происходит оптимизация с использованием различных алгоритмов для минимизации функции ошибок [2]. Это обеспечивает более точное позиционирование и картографирование. Обнаружение объектов работает с использованием рамок для выявления целей, однако внутри этих рамок могут быть исключены и другие пиксели, что приводит к потерям информации. В отличие от этого, семантическая сегментация обеспечивает более тонкую и точную обработку, выполняя разделение на уровне отдельных пикселей между перед-

ним планом и фоном, что позволяет более эффективно удалять динамические объекты от набора ключевых точек для систем визуальной навигации.

TUM [1] — это стандартный набор данных, широко используемый для исследований в области одновременной локализации и построения карт (SLAM). Особенно важной частью является динамический сегмент данных (Dynamic), который фокусируется на динамических средах, где присутствуют движущиеся объекты, например, ходящие люди. Поэтому было решено сначала провести проверку на этом наборе данных.

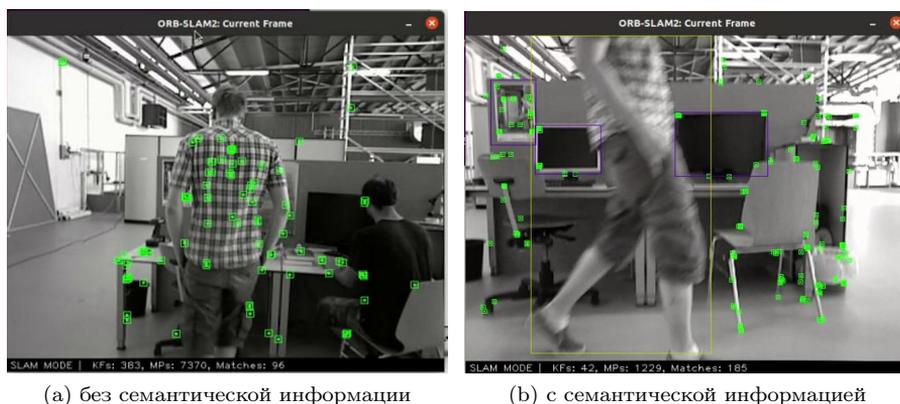


Рис. 1. Визуальная навигационная система

На рисунках 1a и 1b представлены результаты работы ORB-SLAM без и с участием технологии глубокого обучения семантической сегментации. Заметное отличие заключается в том, извлекаются ли особенности движущихся людей на изображении для построения карты.

Данные экспериментов показывают, что модифицированный алгоритм ORB-SLAM может работать в режиме реального времени на знаменитом продукте NVIDIA — классическом встраиваемом устройстве Nano. Веса семантической сегментации обучались на другой системе — Ubuntu 20.04 с графическим процессором Nvidia GTX 3060Ti и настольном компьютере, оснащённом процессором i7. Цель эксперимента заключалась в качественном и количественном анализе улучшения точности системы навигации и позиционирования SLAM после добавления модуля семантической сегментации, а также в оценке реального времени работы этой системы.

Сравнение среднеквадратичной ошибки (RMSE) траектории движения набора данных TUM показано в таблице 1. Результаты эксперимента оцениваются с использованием абсолютной погрешности траектории (ATE) и относительной погрешности траектории (RPE). Чем меньше RMSE у ATE и RPE, тем ближе

Таблица 1. Результаты эксперимента

| | ORB-SLAM2 | DS-SLAM | DynaSLAM | Наша Система |
|-----|-----------|---------|----------|--------------|
| ATE | 0,7521 | 0,0247 | 0,015 | 0,01757 |
| RPE | 0,4124 | 0,0333 | 0,2215 | 0,02324 |

траектория движения к реальной траектории, т. е. тем выше точность позиционирования мобильного робота.

ORB-SLAM2 является одним из классических проектов визуального SLAM и часто используется исследователями в качестве объекта для сравнения. Однако, как показывают результаты в таблице 1, его точность сильно ухудшается в сценах с динамическими объектами, способными к движению. Нашими основными объектами сравнения являются два известных проекта визуального SLAM, ориентированные на динамические среды: DS-SLAM и DynaSLAM. Как видно из таблицы, наша точность выше, чем у DS-SLAM. Что касается DynaSLAM, этот проект использует глубокую нейросеть MaskRCNN, работает достаточно медленно, потребляет значительные вычислительные ресурсы и не является программой для реального времени, поэтому наши встроенные устройства не имеют возможности запускать этот проект в реальном времени. Он может быть испытан только на компьютерах с высокой производительностью. Это скорее алгоритм для постобработки данных карты. Таким образом, хотя наша точность немного ниже, чем у DynaSLAM (но все же сравнима), наш проект может работать в реальном времени, что имеет большое практическое значение.

- [1] *Yu Chao, et al.* DS-SLAM: A Semantic Visual SLAM towards Dynamic Environments // 2018 IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS). — IEEE, 2018. — P. 1168–1174.
- [2] *Mur-Artal R., Martinez Montiel J. M., Tardos J. D.* ORB-SLAM: a Versatile and Accurate Monocular SLAM System // IEEE Transactions on Robotics. — IEEE, 2015. — Vol. 31, No. 5. — P. 1147–1163.

Распознавание рукописного текста на основе реконструкции траектории пера по цифровому изображению

Афанасьев Глеб Ильич^{1,2*}

afanasyev-8g@yandex.ru

Местецкий Леонид Моисеевич^{1,2}

mestlm@mail.ru

¹Москва, МГУ

²Москва, ВШЭ

Задачи распознавания рукописного текста можно разделить в зависимости от типа входных данных на два класса: off-line и on-line. К первому относятся задачи, входом для которых является статическое изображение рукописного текста, представленное в виде матрицы пикселей. Т.к. в этом случае процесс письма предшествует процессу распознавания, то задачи такого рода получили название задачи off-line распознавания. Системы, решающие такие задачи, получают на вход изображение и выдают текст, представленный в виде кодов символов.

Ко второму классу (on-line) относятся задачи, в которых входной информацией является последовательность пикселей, соответствующая траектории движения пера при написании текста. Обладая такой информацией, можно осуществить распознавание непосредственно в момент письма, поэтому задачи такого рода получили название on-line распознавания. Они основываются на траектории движения пера при написании текста. Технологии, решающие задачи on-line распознавания, нашли широкое применение во всевозможных электронных записных книжках, карманных компьютерах и везде, где возможен непосредственный ввод информации при письме.

В связи с этим, возникает идея преобразования off-line сигнала к on-line сигналу, т.е. восстановления динамической информации о траектории пера по цифровому изображению рукописного текста, чтобы провести редукцию задачи off-line распознавания к задаче on-line распознавания. В данной работе предлагается подход к реконструкции траектории пера на основе штриховой сегментации рукописного текста [1] и распознавания текста по этой реконструированной траектории.

Наиболее популярные методы извлечения информации из рукописного текста основаны на использовании нейронных сетей, получающих на вход признаковое представление текста и выдающих необходимую информацию. В данной работе предлагается использовать реконструированную траекторию пера в качестве признакового представления. Задачу можно сформулировать следующим образом: по входному бинаризованному изображению текста нужно восстановить траекторию пера во время его написания. После чего проверить, возможно ли на основании траектории пера решить задачу извлечения информации из рукописного текста более эффективно, чем это предлагают существующие методы.

Для проверки гипотезы рассматриваются две задачи: распознавание рукописного текста и идентификация автора рукописного текста.

В качестве архитектуры нейросети для второй задачи была выбрана сверточная нейросеть из 3 сверточных блоков, а для первой задачи — сверточно-рекуррентная сеть из 6 сверточных слоев, выполняющих функцию кодировщика, и двунаправленной рекуррентной нейронной сети в качестве декодировщика. Эксперименты ставились на датасете IAM [2,3], который содержит 13 353 строки текста на английском языке от 657 различных авторов.

Для первой задачи после 25 эпох обучения CER на датасете IAM final составила 12,1. Этот результат немного уступает рекордам существующих методов распознавания [4]. Дальнейшее обучение сети позволит его улучшить. Но он подтверждает хорошее качество реконструкции траектории пера по штриховому представлению рукописного текста.

Для второй задачи результат получился более высокий. После 100 эпох обучения, занявших примерно 5 минут на ресурсах системы google colab, мера ассигасы на тесте достигла 0,99. Это показывает что реконструированная траектория пера может быть эффективно использована для идентификации автора.

Работа поддержана грантом РФФИ № 22-68-00066.

- [1] *Местецкий Л. М.* Штриховая сегментация рукописного текста // Математические методы распознавания образов: Тезисы докладов 21-й Всероссийской конференции с международным участием, г. Москва 2023 г, М.: Российская академия наук, 2023. — С. 118–119.
- [2] *Marti U., Bunke H.* The IAM-database: An English Sentence Database for Off-line Handwriting Recognition // Int. J. Document Analysis and Recognition. — 2002. — Vol. 5. — P. 39–46.
- [3] *Liwicki M., Bunke H.* IAM-OnDB—an on-line English Sentence Database Acquired from Handwritten Text on a Whiteboard // Eighth Int. Conf. Document Analysis and Recognition (ICDAR'05), Seoul, Korea (South). — 2005. — Vol. 2. — P. 956–961.
- [4] *Hernandez Diaz D., Qin S., Ingle R., Fujii Ya., Bissacco A.* Rethinking Text line Recognition Models // arXiv. — 2021. <https://arxiv.org/abs/2104.07787>

MDM-алгоритм и задача Сильвестра

Малозёмов Василий Николаевич¹

v.malozemov@spbu.ru

Соловьёва Наталья Анатольевна²

4vinyo@gmail.com

Тамасян Григорий Шаликович^{3,4}★

grigoriytamasjan@mail.ru

¹ Санкт-Петербургский государственный университет

² Санкт-Петербургский государственный экономический университет

³ Военно-космическая академия им. А. Ф. Можайского

⁴ Институт проблем машиноведения РАН

В пространстве \mathbb{R}^n с евклидовой нормой заданы m точек a_1, a_2, \dots, a_m . Задача Сильвестра ставится так: *найти шар наименьшего объема, содержащий все точки $a_i, i \in 1 : m$.*

Запишем формализацию этой задачи

$$\max_{i \in 1:m} \left\{ \frac{1}{2} \|a_i - x\|^2 \right\} \rightarrow \min_{x \in \mathbb{R}^n}. \quad (1)$$

Здесь x — центр шара. Задача (1) имеет решение x^* и оно единственно.

Обозначим через A матрицу со столбцами a_1, a_2, \dots, a_m и через b — вектор из \mathbb{R}^m с компонентами $b[i] = \frac{1}{2} \|a_i\|^2$. Рассмотрим вспомогательную задачу квадратичного программирования

$$\begin{aligned} Q(u) &:= \frac{1}{2} \langle A^T A u, u \rangle - \langle b, u \rangle \rightarrow \min, \\ \sum_{i=1}^m u[i] &= 1, \quad u[i] \geq 0 \text{ при всех } i \in 1 : m. \end{aligned} \quad (2)$$

Множество ее планов обозначим через U . Очевидно, что у задачи (2) существует оптимальный план. Отметим, что целевая функция $Q(u)$ задачи (2) на множестве планов U допускает представление

$$Q(u) = -\frac{1}{2} \sum_{i=1}^m u[i] \cdot \|a_i - x\|^2, \quad \text{где } x = A u = \sum_{i=1}^m u[i] a_i.$$

Теорема 1. *Если u^* — оптимальный план задачи (2), то вектор $x^* = A u^*$ будет решением задачи (1). При этом $\max_{i \in 1:m} \left\{ \frac{1}{2} \|a_i - x^*\|^2 \right\} = -Q(u^*)$.*

Таким образом, задача Сильвестра (1) сводится к задаче квадратичного программирования (2). Для решения задачи (2) существуют известные конечные методы (например, метод Данцига). Они хорошо работают при небольших значениях m и n . Эффективность конечных (базисных) методов связана с использованием формул пересчета параметров при переходе к очередной итерации. Однако при больших m и n последовательный пересчет параметров приводит к критическому накапливанию погрешности вычислений. В связи с этим

разрабатываются простые итеративные методы, бесконечные в общем случае, с гарантированной сходимостью. В данном докладе предлагается такой метод для решения задачи (2). Он аналогичен MDM-алгоритму [1], разработанному для нахождения точки из выпуклой оболочки конечного множества, ближайшей к началу координат.

Получим нестандартный вариант критерия оптимальности для задачи (2). Для этого введем *оценку плана* u — величину

$$\Delta(u) = \max_{i \in I^+(u)} \{ \langle a_i, x \rangle - b[i] \} - \min_{i \in 1:m} \{ \langle a_i, x \rangle - b[i] \}.$$

Здесь $x = Au$ и $I^+(u) = \{i \in 1 : m \mid u[i] > 0\}$ — носитель плана u . Величина $\Delta(u)$ при всех $u \in U$ неотрицательна.

Теорема 2. *План u задачи (2) будет оптимальным тогда и только тогда, когда $\Delta(u) = 0$.*

С помощью $\Delta(u)$ можно оценить близость вектора $x = Au$ к решению x^* задачи (1).

Теорема 3. *Пусть u — план задачи (2) и $x = Au$. Тогда $\|x - x^*\|^2 \leq \Delta(u)$.*

Оценке $\Delta(u)$ при $u \in U$ можно дать эквивалентное определение в терминах исходной задачи (1). А именно, $\Delta(u) = \frac{1}{2} (\max_{i \in 1:m} \|a_i - x\|^2 - \min_{i \in I^+(u)} \|a_i - x\|^2)$.

Теорема 4. *План u задачи (2) будет оптимальным тогда и только тогда, когда для вектора $x = Au$ при всех $s \in I^+(u)$ выполняются равенства*

$$\|a_s - x\| = \max_{i \in 1:m} \|a_i - x\|.$$

Теорему 4 можно переформулировать в геометрических терминах:

план u задачи (2) будет оптимальным тогда и только тогда, когда точки a_s при всех $s \in I^+(u)$ лежат на поверхности шара с центром в точке $x = Au$ и радиусом R , равным $\max_{i \in 1:m} \|a_i - x\|$. Такие точки a_s будем называть опорными.

Опишем, как применяется MDM-алгоритм к решению задачи Сильвестра. В качестве начального приближения возьмем произвольный план u_0 задачи (2). Ему соответствует вектор $x_0 = Au_0$.

Пусть уже имеется k -е приближение $u_k \in U$. Ему соответствует вектор $x_k = Au_k$. Проверим план u_k на оптимальность. Для этого выберем индексы $i'_k \in 1 : m$ и $i''_k \in I^+(u_k)$ из условий

$$\begin{aligned} \langle a_{i'_k}, x_k \rangle - b[i'_k] &= \min_{i \in 1:m} \{ \langle a_i, x_k \rangle - b[i] \}, \\ \langle a_{i''_k}, x_k \rangle - b[i''_k] &= \max_{i \in I^+(u_k)} \{ \langle a_i, x_k \rangle - b[i] \}. \end{aligned}$$

Вычислим оценку $\Delta(u_k) = \langle a_{i''_k} - a_{i'_k}, x_k \rangle - (b[i''_k] - b[i'_k])$. Если $\Delta(u_k) = 0$, то u_k — решение задачи (2), а x_k — решение задачи (1). Процесс закончен.

Пусть $\Delta(u_k) > 0$. В этом случае будем говорить, что план u_k порождает k -ю итерацию. Вычислим $\tilde{\delta}_k = \frac{\Delta(u_k)}{\|a_{i'_k} - a_{i''_k}\|^2}$ и положим

$$\delta_k = \begin{cases} \tilde{\delta}_k, & \text{если } \tilde{\delta}_k \leq u_k[i'_k], \\ u_k[i''_k], & \text{если } \tilde{\delta}_k > u_k[i'_k]. \end{cases}$$

Очевидно, что $0 < \delta_k \leq \tilde{\delta}_k$. Если $\delta_k = \tilde{\delta}_k$, то k -ю итерацию будем называть *неусеченной*. При $\delta_k < \tilde{\delta}_k$ получаем *усеченную итерацию*. Формулы

$$u_{k+1} = u_k + \delta_k (e_{i'_k} - e_{i''_k}), \quad x_{k+1} = x_k + \delta_k (a_{i'_k} - a_{i''_k}),$$

где e_i — i -й орт в пространстве \mathbb{R}^m , определяют очередное $(k+1)$ -е приближение. Описание алгоритма завершено.

Теорема 5. *Количество идущих подряд усеченных итераций конечно.*

MDM-алгоритм может содержать конечные блоки идущих подряд усеченных итераций. При этом существует бесконечная последовательность неусеченных итераций.

Из последовательности $\{u_k\}$ исключим все планы, порождающие усеченные итерации. Останется бесконечная подпоследовательность планов $\{u_{k_j}\}$, порождающих неусеченные итерации.

Теорема 6. *Для векторов $x_{k_j} = Au_{k_j}$ справедливо предельное соотношение $\lim_{j \rightarrow \infty} x_{k_j} = x^*$, где x^* — решение задачи (1).*

Теорема 7. *Справедливо предельное соотношение $\lim_{k \rightarrow \infty} Q(u_k) = Q(u^*)$, где u^* — некоторое решение задачи (2).*

Замечание 1. Эффективность MDM-алгоритма решения задачи Сильвестра зависит от величины $|I^+(u^*)|$. Наиболее благоприятный случай, когда $|I^+(u^*)| = n + 1$.

Замечание 2. Если план u_k порождает усеченную итерацию, то при переходе от u_k к u_{k+1} мощность $|I^+(u_k)|$ при $u_k[i'_k] > 0$ уменьшается на единицу и остается прежней при $u_k[i'_k] = 0$.

Численные эксперименты выполнялись в Институте проблем машиноведения РАН при финансовой поддержке Российского научного фонда (проект № 23-41-00060). В [2] представлено подробное изложение полученных результатов.

- [1] Митчелл Б. Ф., Демьянов В. Ф., Малоземов В. Н. Нахождение ближайшей к началу координат точки многогранника // Вестник ЛГУ. — 1971. — № 19. — С. 38–45.
- [2] Малозёмов В. Н., Соловьёва Н. А., Тамасян Г. Ш. MDM-алгоритм и задача Сильвестра. II // Семинар «O & ML». Избранные доклады. 28 сентября 2023 г. (<http://oml.cmlaboratory.com/rep23.shtml#0928>)

Оптимизационный подход к проблеме определения поля скоростей в задачах обработки изображений

Котина Елена Дмитриевна^{1*}

e.kotina@spbu.ru

Овсянников Дмитрий Александрович¹

d.a.ovsyannikov@spbu.ru

Харченко Данил Сергеевич¹

st098139@student.spbu.ru

¹Санкт-Петербург, Санкт-Петербургский государственный университет

В работе представлен новый оптимизационный подход к задаче построения поля скоростей. Оптимизационный подход при построении оптического потока используется многими авторами. Наиболее известные — это методы Хорна-Шанка [1], Лукаса-Канаде [2], Фарнебака [3] и их модификации. В данной работе рассматривается принципиально новый подход, который основан на методах теории управления и оптимизации ансамблей траекторий [4]. Данный подход не исключает возможности, что яркость вдоль траекторий движения может изменяться. Это позволяет строить направленные методы оптимизации для определения оптических и неоптических потоков. Поле скоростей задается как некоторая функция, зависящая от вектора произвольных параметров, которые определяются в результате минимизации функционала, заданного на ансамбле траекторий, определяемом этим полем скоростей [5].

Предполагается, что перемещения описываются следующим уравнением

$$\dot{x} = f(t, x, u), \quad i = \overline{1, n}. \quad (1)$$

Здесь t — время, $t \in [0, T]$, x — вектор пространственных координат, $x \in R^n$, u — вектор параметров, $u \in U$, U — компакт в R^r , $f = f(t, x, u)$ — n -мерная вектор-функция.

Для системы (1) задаются начальные условия

$$x(0) = x_0 \in M_0, \quad i = \overline{1, n}.$$

Уравнение переноса, которое при заданной вектор-функции $f(t, x, u)$ определяет изменение функции плотности в фазовом пространстве с течением времени, имеет вид

$$\frac{\partial \rho(t, x)}{\partial t} + \frac{\partial \rho(t, x)}{\partial x} f(t, x, u) + \rho(t, x) \operatorname{div}_x f(t, x, u) = 0, \quad (2)$$

где

$$\operatorname{div}_x f(t, x, u) = \sum_{i=1}^n \frac{\partial f_i(t, x, u)}{\partial x_i}.$$

При этом задается начальное условие

$$\rho(0, x) = \rho_0(x), \quad x \in M_0, \quad (3)$$

где $\rho_0(x)$ — заданная функция.

Обозначим через $M_{T,u}$ образ множества M_0 вдоль траекторий системы (1) в момент времени T . Задача состоит в нахождении вектора параметров u такого, что в момент времени T плотность, вычисленная в силу уравнения (2) с условием (3), совпала бы с заданной плотностью $\rho(x)$ на $M_{T,u}$, т. е.

$$\rho(T, x) = \hat{\rho}(x), \quad x \in M_{T,u}, \quad i = \overline{1, n}. \quad (4)$$

Сформулируем оптимизационную задачу. С этой целью введем функционал

$$J(u) = \int_{M_{t,u}} g(x_T, \rho(T, x_T)) dx_T. \quad (5)$$

Здесь $g(x_T, \rho(T, x_T))$ — неотрицательная, непрерывно-дифференцируемая по своим аргументам функция, характеризующая в том или ином виде условие (4).

Ставится задача минимизации функционала (5). Предполагаем, что момент времени T здесь не фиксирован и также будет варьироваться. Это позволяет исследовать задачу построения поля скоростей не только при малых смещениях. Решая задачу минимизации функционала (5) и определяя параметры вектора u , мы решаем задачу восстановления функций $f(t, x, u)$, $i = \overline{1, n}$, т. е. определяем поле скоростей, задаваемое формулой (1). Появляется возможность не только построения поля скоростей, но и восстановления динамики исследуемого объекта, что отличает данный подход от вышеупомянутых подходов.

В работе построены алгоритмы определения поля скоростей на основе полученных аналитических представлений вариаций функционала (5) по параметрам и по времени. Использование вариации по времени позволяет ускорить процесс нахождения искомых параметров. В алгоритме предполагается возможность разбиения изображения на подобласти на основе различных методов сегментации. Предлагаемый метод можно использовать для решения различных задач. Рассматриваются тестовые и медицинские радионуклидные изображения для которых, в частности, актуально использование неоптических потоков, т. к. для радионуклидных изображений яркость определяется плотностью распределения радиофармпрепарата, который перераспределяется во времени. Радионуклидные диагностические изображения могут быть получены методами ОФЭКТ (однофотонная эмиссионная компьютерная томография) и ПЭТ (позитронная эмиссионная томография) [6].

- [1] *Horn B. K. P., Schunck B. G.* Determining Optical Flow // *Artificial Intelligence*. — 1981. — No. 17 — P. 185–203.
- [2] *Lucas B. D., Kanade T.* An Iterative Image Registration Technique with an Application to Stereo Vision // *Proc. Imaging Understanding Workshop*. — 1981. — P. 121–130.
- [3] *Farneback G.* Two-Frame Motion Estimation Based on Polynomial Expansion // *Lecture Notes in Computer Science*. — 2003. — Vol. 2749. — С. 363–370.

- [4] *Овсянников Д. А.* Математические методы управления пучками. — Л.: Изд-во Ленингр. ун-та, 1980. — 228 с.
- [5] *Bazhanov P. V., Kotina E. D., Ovsyannikov D. A., Ploskikh V. A.* Optimization Algorithm of the Velocity Field Determining in Image Processing // *Cybernetics and Physics*. — 2018. — Vol. 7, No. 4. — P. 174-181.
- [6] *Kotina E., Ploskikh V., Shirokolobov A.* Digital Image Processing in Nuclear Medicine // *Phys. Part. Nuclei*. — 2022. — Vol. 53. — P. 535-540.

Метод попарных сравнений для задачи коммивояжера. Специальный случай выпуклой оболочки

Лемтюжникова Дарья Владимировна¹

darabbt@gmail.com

Шушко Никита Игоревич¹

shushko.ni@phystech.edu

Красоткин Семён Александрович^{1*}

krasotkin.sa@phystech.edu

Барашов Егор Борисович¹

barashov.eb@gmail.com

¹Москва, Институт проблем управления им. В. А. Трапезникова РАН

Задача коммивояжера — это задача дискретной оптимизации, в которой необходимо определить кратчайший путь обхода всех вершин (городов), посещая каждый только один раз. В общем случае, эта задача является NP-трудной и не имеет полиномиального алгоритма. Однако существуют специальные случаи, для которых можно найти оптимальное решение за полиномиальное время.

Для расширения применимости данных алгоритмов и изучения структуры NP-трудных задач в [1] предлагается *метод попарного сравнения*. Данный подход предлагает сравнивать два примера задачи A и B по некоторому критерию и задавать *функцию различия*, значение которой равно целевой функции A при использовании в качестве решения полиномиального (псевдополиномиального) точного решения задачи B . Таким образом, можно оценить насколько эвристика, дающая точное решение для подкласса примеров задачи, может быть применима для более широкого класса примеров задачи. В нашей работе описано применение метода попарного сравнения для задачи коммивояжера.

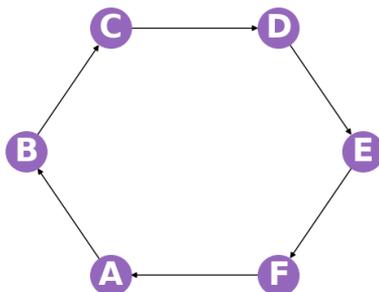


Рис. 1. Оптимальный маршрут в случае, когда все вершины расположены на евклидовой плоскости и образуют границу выпуклого многоугольника.

В работе рассматривается специальный случай, когда все вершины расположены на евклидовой плоскости и образуют границу выпуклого многоугольника [2], т. е. лежат на выпуклой оболочке множества вершин. В таком случае оптимальный маршрут будет проходить по границе многоугольника (см. рис. 1) и решение можно найти за полиномиальное время $O(n)$, где n — количество заданных вершин.

Для перехода от исходной задачи к специальному случаю были изучены два подхода: рассмотрение исходной задачи как набора вложенных выпуклых оболочек и как набора непересекающихся выпуклых оболочек, что является кластеризацией вершин на выпуклые оболочки.

Для каждого варианта перехода предложена функция попарных сравнений и описаны алгоритмы получения допустимого решения. Проведено сравнение этих подходов, а также сравнение кластеризации на выпуклые оболочки с классической *k*-means кластеризацией.

Работа поддержана грантом РФФИ № 22-71-10131.

- [1] *Lemtyuzhnikova D., Chebotarev P., Goubko M., Shushko N., Somov M.* Pairwise Similarity Estimation for Discrete Optimization Problems // *Advances in Systems Science and Applications*. — 2023. — Vol. 23. — No. 2. — P. 164–177.
- [2] *Burkard R. E. et al.* Well-Solvable Special Cases of the Traveling Salesman Problem: A Survey // *SIAM review*. — 1998. — Vol. 40, No. 3. — P. 496–546.

Hybrid Algorithms Based on Nature-Inspired and Local Descent Methods for Solving Multiextremal Optimization Problems

*Sorokovikov Pavel*¹

pavel@sorokovikov.ru

¹Irkutsk, Matrosov Institute for System Dynamics and Control Theory of SB RAS

The problem of searching for a global optimum of multiextremal objective function remains one of the most difficult and relevant in theory and applications of mathematical programming and optimizing dynamic systems. As a rule, reliable procedures of nonlocal optimization are based on the balance between the global study of the search space (“exploration”) and the local improvement of the decisions obtained (“exploitation”). Accordingly, one iteration of the algorithm for finding an absolute extremum should include two stages: a global search in an admissible set and a local clarification by gradient methods in areas where the presence of a global optimum is likely. The paper proposes an approach that uses the advantages of nature-inspired algorithms to study the admissible set and gradient methods for local optimization, which allows you to build computational schemes that underlie effective methods for solving the global search problems.

The proposed approach to the numerical study of the tasks of searching for a global extremum of multimodal objective functions is based on the use of genetic search [1], differential evolution [2], biogeography, harmony search, particle swarm, cuckoo search [3] and L-BFGS algorithms. Six two-method computing schemes for solving multiextremal optimization problems were proposed and implemented in the form of algorithms library.

The developed algorithms were investigated in the set of multiextremal test problems characterized by a different level of complexity. All algorithms based on combinations with the L-BFGS local search method showed significant improvements compared to the initial nature-inspired algorithms. The results of computational experiments are given. The resulting research results confirm the performance of the developed algorithms.

This research is funded by the grant from the Ministry of Education and Science of Russia within the framework of the project “Theory and methods of research of evolutionary equations and controlled systems with their applications” (state registration number 121041300060-4).

- [1] *Whitley D.* A Genetic Algorithm Tutorial // Statistics and Computing. — 1994. — Vol. 4. — P. 65–85.
- [2] *Storn R., Price K.* Differential Evolution—a Simple and Efficient Heuristic for Global Optimization Over Continuous Spaces // Journal of Global Optimization. — 1997. — Vol. 11. — P. 341–359.
- [3] *Devika G., Karegowda A. G.* Bio-Inspired Optimization: Algorithm, Analysis and Scope of Application // Swarm Intelligence—Recent Advances and Current Applications. — 2023. — P. 1–37.

Модификация протокола локального голосования для оптимизации балансировки загрузки ресурсов распределенной системы

Володина Елена Александровна¹*

el.al.volodina@gmail.com

Ерофеева Виктория Александровна¹

eva@ipme.ru

¹Санкт-Петербург, Институт проблем машиноведения РАН

Современная разработка программного обеспечения постепенно переходит от монолитной архитектуры к микросервисной, становясь в результате распределенной [1]. В то же время рабочая нагрузка, приходящая на сетевые системы, постоянно растет из-за растущего спроса на облачные вычисления и машинное обучение. Это приводит к необходимости в балансировщиках нагрузки. Наиболее распространены централизованные балансировщики нагрузки, но в распределенных сетях они могут вызывать сетевые задержки, поэтому возникает необходимость в создании новых балансировщиков, работающих в распределенных децентрализованных системах.

Возможные структурные изменения в коммуникационной сети (добавление и удаление узлов, потеря связи), наличие шумов и задержек, а также неравномерное распределение входящей рабочей нагрузки между узлами требуют использования динамической балансировки нагрузки. Кроме того, появление новых и отсутствие информации о трудоемкости существующих задач приводят к разбалансировке системы. В результате задача становится нестационарной, также называемой задачей трекинга. Протокол локального голосования, предложенный в [2], позволяет достичь субоптимальной балансировки нагрузки в случае задач трекинга.

В статье предлагается новый ускоренный децентрализованный протокол балансировки нагрузки, основанный на современных теоретических разработках в области нестационарной оптимизации, консенсуса и техник ускорения, в частности ускорения по Нестерову для задач трекинга [3].

Рассматривается сетевая система, состоящая из n узлов (серверов). Система выполняет балансировку нагрузки децентрализованным способом и пытается распределить входящие задачи между узлами. Сетевая топология задана графом $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, где $\mathcal{N} = \overline{1, n}$ — набор узлов, каждый из которых обслуживает входящие запросы в порядке очереди и $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ — множество граней. Пусть $(j, i) \in \mathcal{E}$ если есть ребро между узлами j и i , иными словами узлы j и i способны обмениваться данными. Графу ставится в соответствие взвешенная матрица смежности $B = [b^{ij}]$, $b^{ij} > 0$ если $(j, i) \in \mathcal{E}$ и $b^{ij} = 0$ в ином случае.

Задачи поступают в систему в случайный момент времени и принимаются различными узлами системы. В любой момент времени $k = 0, 1, \dots$, состояние узла i ($i \in \mathcal{N}$), описывается тремя характеристиками: $q_k^i(u_k^i)$ — количество элементарных атомарных задач узла i в момент времени k , учитывающее поступившие задачи z_k^i и перераспределенные задачи u_k^i ; p — производительность

узлов; $x_k^i(u_k^i) = q_k^i(u_k^i)/p$ — загрузка или время, необходимое узлу i для выполнения всех задач в очереди. При оптимальном распределении задач $x^i = x^j \quad \forall i, j \in \mathcal{N}$ [2].

Необходимо найти протокол перераспределения $\mathbf{u}_k = [u_k^1, \dots, u_k^n]^T \in \mathbb{R}^n$ для достижения одинаковой загруженности системы. Это делается путем минимизации потенциала Лапласа (см. [4]):

$$F_k(\mathbf{u}_k) = \sum_{i,j \in \mathcal{N}} \left(\frac{b^{ij}}{2} \left(x_{k+1}^i(u_k^i) - x_{k+1}^j(u_k^j) \right)^2 + \frac{1}{p^2} b^{ij} u_k^j u_k^i \right) \rightarrow \min_{\mathbf{u}_k}. \quad (1)$$

Обозначим через L — константу Липшица функции F_k , μ — константу строгой выпуклости функции F_k . В процессе оптимизации в момент времени k узел i по полученным от других узлов зашумленным данным о загрузке может посчитать градиент, вычисляемый по протоколу локального голосования:

$$f_k^i(u_k^i) = - \left(\sum_{j \in \mathcal{N}} \frac{b^{ij}}{p} \left(x_{k+1}^{i,i}(u_k^i) - x_{k+1}^{i,j}(u_k^j) - \frac{1}{p} u_k^j \right) \right) + \xi_k^i, \quad k = 0, 1, \dots, \quad (2)$$

где $\xi_k^i \in \mathbb{R}$ — независимые одинаково распределенные случайные величины.

В начале работы алгоритма устанавливаются начальные значения используемых переменных — начальных оценок $\hat{\mathbf{u}}_0 \in \mathbb{R}^n$ и других вспомогательных величин: $\gamma_0 > 0$, $h > 0$, $\eta \in (0, \mu)$, $H_1 = h - \frac{h^2 L}{2}$, $v_0^i = \hat{u}_0^i$. На каждом шаге $k \geq 0$ выбирается $\alpha_k \in (0, 1)$, удовлетворяющее неравенству $H_1 - \frac{\alpha_k^2}{2\gamma_{k+1}} > 0$, где $\alpha_k \in [\alpha_x, 1)$.

На каждом узле $i \in \mathcal{N}$ выполняются следующие шаги:

1. Вычисляется вспомогательный параметр алгоритма

$$\gamma_{k+1} = (1 - \alpha_k)\gamma_k + \alpha_k(\mu - \eta).$$

2. Выбирается точка наблюдения

$$z_k^i = \frac{1}{\gamma_k + \alpha_k(\mu - \eta)} \left(\alpha_k \gamma_k v_k^i + \gamma_{k+1} \hat{u}_k^i \right)$$

и вычисляется значение протокола локального голосования $f_k^i(z_k^i)$.

3. Обновляется текущая оценка протокола перераспределения

$$\hat{u}_{k+1}^i = z_k^i - h f_k^i(z_k^i).$$

4. Пересчитывается вспомогательный параметр алгоритма

$$v_{k+1}^i = \frac{1}{\gamma_k} \left[(1 - \alpha_k)\gamma_k v_k^i + \alpha_k(\mu - \eta) z_k^i - \alpha_k f_k^i(z_k^i) \right].$$

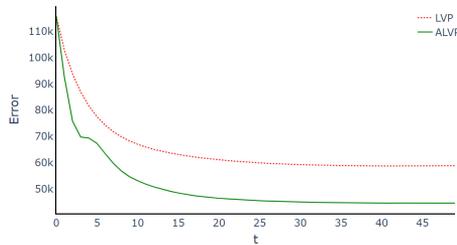


Рис. 1. Усредненные по 100 экспериментам величины ошибок: сплошная линия — предложенный алгоритм с ускорением, пунктирная линия — версия без ускорения

Для сравнения работы ускоренного протокола локального голосования с исходной версией была проведена серия экспериментов по балансировке нагрузки сети. Была смоделирована работа распределенной системы, состоящей из 30 серверов. В каждый момент времени t новые задачи распределялись между серверами случайным образом. Доступ к информации о статистических свойствах задач у серверов отсутствовал. В каждом эксперименте строилась случайная топология сети. При моделировании установлены следующие параметры алгоритма: $h = 0,2$, $\eta = 0,8$, $\gamma_0 = 0,07$, $\alpha_x = 0,09$, $L = 0,4$, $\mu = 0,02$. На рисунке 1 представлены результаты сравнения предлагаемого ускоренного алгоритма с неускоренной версией.

Случайные величины (сложность задач и их распределение между серверами, шумы и топология сети) установлены идентичным образом для обоих алгоритмов. Как видно из результатов моделирования, предлагаемый ускоренный протокол локального голосования сходится быстрее исходного [2].

В докладе представлен ускоренный распределенный алгоритм балансировки нагрузки сети. Эмпирические результаты показывают более быструю скорость сходимости по сравнению с предшествующим решением в случае оптимизации строго выпуклых функций при наличии неизвестного, но ограниченного шума.

Работа проводилась при поддержке гранта Президента Российской Федерации для молодых российских ученых (МК-372.2022.4)

- [1] Newman S. Building microservices. — O'Reilly Media, Inc., 2021.
- [2] Amelina N., Fradkov A., Jiang Y., Vergados D. J. Approximate Consensus in Stochastic Networks with Application to Load Balancing // IEEE Transactions on Information Theory. — IEEE, 2015. — Vol. 61, No. 4. — P. 1739–1752.
- [3] Kosaty D., Vakhitov A., Granichin O., Yuchi M. Stochastic Fast Gradient for Tracking // 2019 American Control Conference (ACC). — IEEE, 2019. — P. 1476–1481.
- [4] Olfati-Saber Reza, Murray R. M. Consensus Problems in Networks of Agents with Switching Topology and Time-Delays // IEEE Transactions on automatic control. — IEEE, 2004. — Vol. 49, No. 9. — P. 1520–1533.

Метаэвристические алгоритмы для задачи упаковки в контейнеры

*Лемтюжникова Дарья Владимировна*¹

darabbt@gmail.com

*Посыпкин Михаил Анатольевич*²

mposypkin@frccsc.ru

Барашов Егор Борисович^{1*}

barashov.eb@gmail.com

¹Москва, Институт проблем управления им. В. А. Трапезникова РАН

²Федеральный исследовательский центр «Информатика и управление» РАН

Задача упаковки в контейнеры — это классическая задача, которая широко изучалась на протяжении десятилетий. Она включает в себя упаковку предметов различных размеров в ограниченное количество контейнеров так, чтобы общее количество используемого пространства было минимальным. Приведем постановку одномерной задачи упаковки в контейнеры. Имеются различные целочисленные размеры предметов w_1, \dots, w_n , которые необходимо упаковать в конечное количество контейнеров, у каждого из которых ограниченная целочисленная вместимость C , таким образом, чтобы количество использованных контейнеров было минимально. Предполагается, что $w_i \leq C$ для любых i . В этой статье совместно с редукцией рассматривается 2 метаэвристических алгоритма, которые хорошо себя зарекомендовали на других задачах [1, 2].

Решение задачи разбивается на 2 части. Сначала применяется редукция, алгоритм которой представлен в [3]. С его помощью без потери общности сокращается исходная задача, часть предметов помещается в контейнеры. Затем оставшаяся задача решается при помощи двух метаэвристических алгоритмов. 41,11 % экземпляров были размещены на этапе редукции, остальные — на этапе сокращения.

Производительность алгоритма Shuffled Frog Leaping Algorithm в значительной степени зависит от выбора параметров. SFLA включает в себя пять ключевых параметров, а именно: количество мемплексов (m), количество лягушек в мемплексе (n), количество лягушек в подмемплексе (q), количество шагов эволюции или заражения в мемплексе между двумя последовательными перетасовками (N) и максимальный размер шага (S_{max}), допустимый во время эволюционного шага. В настоящее время не существует установленной теоретической базы для выбора подходящих значений этих параметров в контексте данной метаэвристики.

Генетические алгоритмы представляют собой метаэвристический подход, позволяющий итеративно искать решения сложных задач. Представление хромосомы, также известное как генетическое кодирование, определяет способ представления решений-кандидатов. В нашем исследовании мы используем групповое представление хромосом, известное также как хромосома в стиле Фалькенауэра, поскольку оно продемонстрировало значительное превосходство над другими представлениями с точки зрения вычислительной и результативной эффективности для задачи упаковки в контейнеры. Это представле-

ние имеет два ключевых преимущества. Во-первых, оно позволяет различным решениям-кандидатам иметь хромосомы разного размера, что удобно для случаев, когда целевое количество контейнеров неизвестно. Во-вторых, это представление позволяет устранить важное узкое место в работе генетического алгоритма, а именно вырождение, которое приводит к неэффективному покрытию пространства поиска, когда одна и та же конфигурация групп исследуется многократно.

Таблица 1. Описание набора данных

| bin1data | |
|------------------|--------------------------------|
| Число предметов | 50, 100, 200, 500 |
| Вместимость бина | 100, 120, 150 |
| Весы предметов | [1, 100], [20, 100], [30, 100] |

Для проведения экспериментов был использован следующий набор данных (bin1data) — 36 различных классов, в каждом по 20 примеров (<https://www.wiwi.uni-jena.de/Entscheidung/binpp/bin1dat.htm>), описание которых представлено в таблице 1. В данном наборе содержится большое число различных значений вместимости контейнеров, количества и распределения весов предметов.

Таблица 2. Результат работы алгоритмов

| | Genetic | Memetic |
|---|---------|---------|
| Процент примеров, решенных оптимально | 72,78 % | 96,11 % |
| Стандартное отклонение результатов от оптимума | 0,55 | 0,2 |
| Стандартное отклонение результатов от нижних границ | 0,66 | 0,35 |

В таблице 2 показано, что с помощью эвристик удалось получить оптимум максимум в 96,11 % случаев. Генетическим алгоритмом только в 72,78 % случаев был получен оптимум. Также по показателям стандартного отклонения и времени работы из таблицы 3 меметический алгоритм сильно опережает по качеству генетический.

Работа поддержана грантом РФФИ № 22-71-10131.

- [1] *Maarof B. B., Rashid T. A., Abdulla J. M., Hassan B. A.* Current Studies and Applications of Shuffled Frog Leaping Algorithm: A Review // Archives of Computational Methods in Engineering. — 2022. — Vol. 29, No. 5. — P. 3459–3474.

Таблица 3. Время работы алгоритмов

| | Genetic | Memetic |
|---|---------|---------|
| Медианное время работы алгоритма для 50 предметов, с | 1,399 | 0,667 |
| Медианное время работы алгоритма для 100 предметов, с | 4,304 | 1,546 |
| Медианное время работы алгоритма для 200 предметов, с | 12,146 | 2,872 |
| Медианное время работы алгоритма для 500 предметов, с | 34,823 | 5,05 |

- [2] *Bai L., Li H., Gao W., Xie J.* A Cooperative Genetic Algorithm Based on Extreme Learning Machine for Data Classification // *Soft Computing*. — 2022. — Vol. 26, No. 17. — P. 8585–8601.
- [3] *Martello S., Toth P.* Knapsack Problems: Algorithms and Computer Implementations. — John Wiley & Sons, Inc, 1990.

Ближкие задачи маршрутизации типа многих коммивояжеров

Макаров Олег Олегович*

fantom2.00@mail.ru

Козлова Маргарита Геннадьевна

art-inf@mail.ru

Лукьяненко Владимир Андреевич

art-inf@yandex.ru

Симферополь, Крымский федеральный университет им. В. И. Вернадского

В работе ставится задача описания NP -трудных задач, близких к полиномиально разрешимым или к практически неразрешимым. Приближенное решение класса NP -трудных задач прикладной многоагентной маршрутизации ($NPmTSP$, Multiple Traveling Salesman Problem, задача многих коммивояжеров) будем связывать с процедурой сравнения класса исследуемых задач $Z \subset NPmTSP$ с близкими к ним задачами \tilde{Z} из подкласса полиномиально разрешимых задач ($\tilde{Z} \in PmTSP \subset NPmTSP$) или с представителями явно полиномиально неразрешимых из $NPmTSP$.

Близость двух задач определяется по взвешенной норме разности наборов метрических характеристик графовых структур сети; близости гистограммы распределения дуг и др. В каждом конкретном случае выбираются качественные и количественные показатели и метрики для сравнений [1].

Наиболее общей является модель в виде NP -полной задачи на условный экстремум (комбинаторной оптимизации, целочисленного программирования, псевдодволевой условной оптимизации и др.), которая является формализацией некоторой экстремальной задачи на графе (сложной сети). Близкими математическими моделями будем называть такие модели, которым отвечают близкие операторы задач на условный экстремум, решения которых будут близки в некоторой метрике. Например, целевые функции совпадают, а соответствующие операторы ограничений близки по норме; ограничения совпадают, а целевые функции отличаются; для задач псевдодволевого программирования с дизъюнктивными ограничениями изменение параметров сети (добавление вершин или запрет на прохождение дуги) приводит к задачам реоптимизации и др. (см. [1]).

Выбор метаэвристических алгоритмов решения задачи многоагентной маршрутизации может быть осуществлен в виде интеллектуализированной системы. В работе представлен этап формирования баз данных для ее обучения, основанный на близости задач по метаинформации.

Подход, основанный на решении близких задач, опирается на систему «задача — близкая задача — алгоритм». В работе описаны этапы формирования решения оригинальной задачи TSP по близкой TSP_j (Traveling Salesman Problem, задача коммивояжера). В качестве близости задач понимается близость их математических моделей и близость участвующих в решении фрагментов сложной сети. Метод определения близости фрагментов (кластеров C_j) состоит в нахождении взвешенного метрического расстояния между векторами метаэвристических параметров соответствующих графов.

Выбор подходящей метаэвристики для решения локальной задачи TSP_j становится важным, поскольку эти алгоритмы позволяют эффективно решать задачи оптимизации всех маршрутов. Отдельной задачей становится выбор оптимального алгоритма и метапараметров для каждой конкретной задачи.

Ранее авторами был приведен детальный обзор способов решения близкой к $mTSP$ задачи $VRPTW$ (Vehicle Routing Problem with Time Windows, задача маршрутизации транспортных средств с временными окнами), где за основу субоптимального решения считали сумму маршрутов коммивояжеров на каждом из кластеров C_j [2]. Т.е. решение TSP_j применяется при решении более сложных задач оптимизации $mTSP$. В [3] рассматривался алгоритм иерархической кластеризации для $HmTSP$ (Hierarchical multiple Traveling Salesman Problem, иерархическая задача многих коммивояжеров), сравнивались результаты решателя Concorde и алгоритма имитации отжига для решения задач коммивояжера. Получено, что многоуровневая кластеризация для задач большой размерности, с последующим применением решателей TSP , способствует снижению сложности решения $mTSP$.

В данной работе приведен эксперимент, подтверждающий гипотезу, что векторы метрических характеристик у близких задач находятся на небольшом расстоянии друг от друга. Также описаны вспомогательные эксперименты, показывающие максимальное допустимое различие между графами, при котором будем считать их близкими. Приведен эксперимент, подтверждающий гипотезу, что метаэвристический алгоритм, работающий оптимально для определенной задачи, будет оптимально работать и для близкой.

Структура интеллектуализированной системы по выбору метаэвристик может быть представлена следующим образом (см. рис. 1):

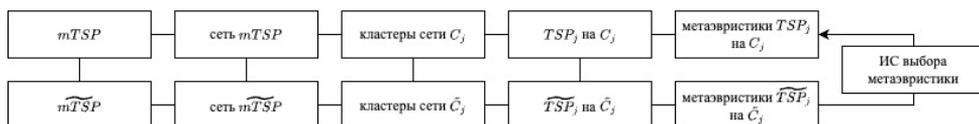


Рис. 1. Схема решения $mTSP$ по близкой задаче

Основным элементом схемы (рис. 1) является интеллектуализированная система по выбору метаэвристик (ИСМЭ). Выбор метаэвристик задачи TSP_j на кластере C_j , $j = \overline{1, m}$, осуществляется на основании близости задачи (TSP_j, C_j) к задаче (TSP_j, C_j) по набору метрических характеристик сети $mTSP$ и соответствующих TSP_j . Одним из факторов близости задач является гистограмма распределения дуг соответствующего графа сети. Управление выбором метаэвристики в этом случае осуществляется на основе близости гистограмм. На рисунке 2 представлена структура ИСМЭ.

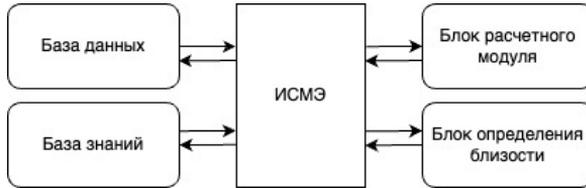


Рис. 2. Схема ИСМЭ

В системе ИСМЭ наиболее трудоемким этапом является наполнение базы метаэвристик, что, в свою очередь, связано с проведением экспериментов на большом наборе примеров графов сети. Рассматривается этап выбора метаэвристик для задачи TSP_j на кластере сети C_j , $j = \overline{1, m}$.

В решении многоагентных задач типа коммивояжера большой размерности производится согласованная декомпозиция на локальные кластерные задачи маршрутизации. На локальном уровне выбираются подходящие алгоритмы.

Алгоритмы решения задачи многоагентной маршрутизации типа коммивояжера (многих коммивояжеров) опираются на метаэвристические алгоритмы решения TSP . Таким образом, в соответствии со схемой на рисунке 1, предлагаемая модель применима для $mTSP$, а также является основой для выбора алгоритмов решения $VRPTW$, описанной в [2].

Приведены результаты экспериментов по решению задач коммивояжера с использованием различных метаэвристик на данных из TSPLIB. Определены эффективные алгоритмы нахождения маршрутов. Эксперимент показал, что большинство примененных метаэвристик позволяют находить приближенные или оптимальные решения. Приведен список лучших по отношению точности алгоритмов, которые используются в разработке интеллектуальной системы выбора метаэвристик, с учетом особенностей задачи (структура и сложность сети, точность, время). Следующим этапом является комбинирование метаэвристик, которое может привести к оптимальным результатам.

Работа поддержана Министерством науки и высшего образования Российской Федерации, соглашение № 075-02-2023-1799.

- [1] Лемтюжникова, Д. В., Лукьяненко, В. А. Проблематика исследования труднорешаемых задач // Интеллектуализация обработки информации: Тезисы докладов 14-й Международной конференции, г. Москва 2022 г. — М.: Российская академия наук, 2022. — С. 437–442.
- [2] Лукьяненко В. А., Лемтюжникова Д. В., Макаров О. О., Козлова М. Г. Модели и алгоритмы многоагентной иерархической маршрутизации с временными окнами // Известия РАН. Теория и системы управления. — 2023. — № 5. — С. 103–126.
- [3] Козлова М. Г., Лукьяненко В. А., Макаров О. О. Построение многоагентных маршрутов в сети с иерархией вершин // Вестник ВГУ. Системный анализ и информационные технологии. — 2023. — № 3. — С. 32–50.

Алгоритмы предоптимизационного и постоптимизационного анализа для невыпуклых функций

*Горнов Александр Юрьевич*¹*

gornov@icc.ru

*Аникин Антон Сергеевич*¹

anikin@icc.ru

*Зароднюк Татьяна Сергеевна*¹

tz@icc.ru

¹Иркутск, Институт динамики систем и теории управления им. В. М. Матросова Сибирского отделения РАН

В докладе обсуждаются алгоритмы и вычислительные технологии, которые позволяют, вообще говоря, не используя методы оптимизации, оперативно получать содержательную информацию об исследуемой задаче (как новые, так и разработанные ранее (см., напр., [1–3])). Алгоритмы предоптимизационного анализа нацелены на получение информации о свойствах целевой функции на всем допустимом множестве. К таким алгоритмам мы относим алгоритмы оценки затратности оракула и его градиентов, точности разностных оценок градиентов, верификации аналитических формул для градиентов, оценки степени выпуклости и вогнутости функции, оценки констант Липшица до четвертой степени и др. Имеют место также несколько алгоритмов оценки числа экстремумов и степени овражности функции. Для большинства указанных подпроблем удалось придумать не один, а несколько алгоритмов, что позволяет верифицировать выдаваемые ими результаты путем сравнения друг с другом.

Алгоритмы постоптимизационного анализа предназначены для изучения свойств найденных точек, «подозрительных на решение». В этом множестве алгоритмы оценки овражности экстремума, оценки точности удовлетворения условиям оптимальности, оценки градиентов по направлениям, оценки чувствительности экстремума (прямой и двойственный локальный поиск) и др.

Конструкции большинства предложенных алгоритмов позволяют ограничивать их работу при достижении лимита затраченного процессорного времени, что позволяет быстро получать искомые оценки. Накопленный к настоящему времени опыт эксплуатации обсуждаемых методик, в том числе для функций с миллионами переменных, позволил убедиться в их полезности. Приводятся результаты вычислительных экспериментов.

Работа выполнена в рамках проекта «Теория и методы исследования эволюционных уравнений и управляемых систем с их приложениями» (№ гос. регистрации: 121041300060-4).

- [1] Горнов А. Ю., Зароднюк Т. С. Вычислительная технология оценки степени выпуклости многоэкстремальной функции // Машинное обучение и анализ данных. — 2014. — Т. 1, № 10. — С. 1345–1353.
- [2] Горнов А. Ю. Алгоритмы предоптимизационного анализа невыпуклой функции многих переменных // М-лы конф. «Ляпуновские чтения». — 2015. — С. 18–19.
- [3] Горнов А. Ю. Алгоритмы постоптимизационного анализа функции многих переменных // М-лы конф. «Ляпуновские чтения». — 2016. — С. 20.

О сложности вычисления элементарных алгебраических функций

Карацуба Екатерина Анатольевна¹*

ekaratsuba@gmail.com

¹Москва, Федеральный исследовательский центр «Информатика и управление» Российской Академии Наук (ФИЦ ИУ РАН)

Алгебраическими функциями называют функции, которые в окрестности каждой точки области определения могут быть заданы алгебраическими уравнениями. Элементарными алгебраическими функциями называют степенную функцию и рациональную функцию.

Считаем, что числа записаны в двоичной системе счисления, знаки которой 0 и 1 называются битами.

Запись знаков 0, 1, плюс, минус, скобка; сложение, вычитание и умножение двух битов назовем одной битовой операцией (далее — просто операцией).

Пусть элементарная алгебраическая функция $y = f(z)$ задана в некоторой ограниченной области $D \in E$, $y = f(z)$ не имеет в \bar{D} особенностей и ограничена вместе со своей производной. Переменная и значение функции записываются последовательностями

$$z = (\tilde{x}_{-j}, \dots, \tilde{x}_0, \tilde{x}_1, \dots); \quad y = (\tilde{y}_{-j}, \dots, \tilde{y}_0, \tilde{y}_1, \dots), \quad (1)$$

где \tilde{x}_i, \tilde{y}_j равны 0 или 1.

Вычислить функцию $y = f(z)$ в точке $z = z_0 \in D$ с точностью до n знаков, значит найти такое число S_n , что

$$|f(z_0) - S_n| \leq c2^{-n}, \quad (2)$$

где постоянная c не зависит от n .

Вычисление функции $f(z)$ в т. $z = z_0$ можно заменить ее вычислением в т. $z = z_n$, где $|z_0 - z_n| < 2^{-n}$, поскольку $f'(z)$ ограничена в \bar{D} . Таким образом, все участвующие в вычислениях числа можно записывать в виде конечных последовательностей, причем из (1), (2) легко видеть, что z и y можно представить в виде целой части и C_0n двоичных знаков после запятой, $C_0 = const$. Поскольку целые части $[z]$, $[y]$ из (1) являются фиксированными величинами, а основным растущим параметром является точность вычисления n , $n \rightarrow +\infty$, то действия фактически производятся над числами, имеющими порядка n знаков.

Сложностью умножения двух n -значных чисел $M(n)$ называется число операций, достаточное для вычисления произведения двух n -значных чисел.

О первом быстром методе — методе умножения и его усовершенствованиях см. [1]. Далее предполагаем, что для сложности умножения двух n -значных чисел справедлива оценка

$$M(n) = O\left(n \log^C n\right),$$

где C — константа (см. алгоритмы из [2]).

Количество операций, достаточное для вычисления функции $f(z)$ в точке z_0 с точностью до n знаков посредством данного алгоритма, называется сложностью вычисления $f(z)$ в точке z_0 и обозначается $s_f(n) = s_{f,z_0}(n)$.

Будем называть быстрыми такие методы и алгоритмы вычисления функции f , что для них

$$s_f(n) = O\left(n \log^K n\right), \quad (3)$$

где K — константа.

Авторский метод БВЕ (Быстрого Вычисления Е-функций, см. [3]) был создан для вычисления простейших и высших трансцендентных функций со сложностью (3). Это второй после АГС Гаусса метод быстрого вычисления простейших трансцендентных функций при любом аргументе. При этом, в отличие от АГС, с помощью БВЕ со сложностью близкой к оптимальной можно вычислить также некоторые высшие трансцендентные функции для алгебраических значений аргумента и параметров.

В [4] были представлены первые алгоритмы быстрого вычисления элементарных алгебраических функций, основанные на методе Ньютона. Например, самый простой алгоритм деления числа a на число b заключается в вычислении методом Ньютона обратной величины $1/b$ с точностью до n знаков с последующим «быстрым умножением» на a .

Ранее в БВЕ-алгоритмах, в промежуточных вычислениях, всегда предполагалось использование метода Ньютона для вычисления элементарных алгебраических и обратных функций. В то же время, возникает вопрос о существовании других быстрых алгоритмов для вычисления этих функций. Такие алгоритмы с применением БВЕ-конструкций впервые были построены в [5]. При этом иногда там использовались также конструкции метода А. А. Карацубы от 1960 г. (см. [1]). Цель настоящего доклада — представить и обсудить эти алгоритмы и сложность вычисления элементарных алгебраических функций.

- [1] Карацуба А. А. Сложность вычислений // Тр. МИАН, 211. — М.: Наука, Физматлит, 1995. — С. 186–202.
- [2] Fürer M. Faster Integer Multiplication // SIAM J. Computing, 2009. — Vol. 39, No. 3. — P. 979–1005.
- [3] Карацуба Е., А. Быстрые вычисления трансцендентных функций // Пробл. передачи информ. — 1991. — Т. 27, № 4. — С. 76–99.
- [4] Cook S. A. On the Minimum Computation Time of Functions // Thesis. — Harvard: Harvard University, 1966.
- [5] Карацуба Е. А. Быстрые алгоритмы вычисления элементарных алгебраических и обратных функций с применением БВЕ // Пробл. передачи информ. — 2022. — Т. 58, № 3. — С. 90–103.

Повышение скорости работы сверточных сетей с использованием алгоритма Винограда

Бобков Александр Валентинович¹

Alexander.Bobkov@bmstu.ru

Ду Кэжао^{1*}

duhuhaizi@gmail.com

¹Москва, МГТУ им. Н. Э. Баумана

В данной работе рассматриваются алгоритмы, позволяющие ускорить работу сверточной сети за счет ускорения вычислений базовой операции — свертки.

Современные успехи в применении нейронных сетей в задачах распознавания не в последнюю очередь связаны с использованием очень глубоких сетей с большим количеством сверточных слоев. Однако при реализации нейронных сетей на бортовых вычислителях реализовать сеть большого размера просто невозможно. Ситуация еще более осложняется, когда требуется обеспечить работу большой сети на маломощном вычислителе еще и в режиме реального времени. Поэтому разработка методов повышения производительности нейросетей и отдельных ее компонентов представляется весьма актуальной.

В 1980 году Шмуэль Виноград предложил алгоритм [1] для уменьшения объема вычислений КИХ-фильтра. Было показано, что для КИХ-фильтра с размером входа m и числом параметров r вместо $m * r$ умножений требуется всего лишь $u(F(m, r)) = m + r - 1$.

В 2015 году Эндрю Лавин (Andrew Lavin) обнаружил [2], что для ускорения процесса расчета свертки можно использовать алгоритм Винограда. При этом некоторые операции умножения преобразуются в более быстрые операции сложения. Так, при одномерной свертке теоретическая эффективность увеличивается в 1,5 раза, при двумерной — в 2,25 раза. Данный подход нашел широкое применение.

В качестве примера рассмотрим одномерную свертку $F(2, 3)$, входной сигнал

$$d = [d_0, d_1, d_2, d_3]^T$$

и ядро свертки

$$g = [g_0, g_1, g_2]^T.$$

Тогда свертка может быть записана в виде матричного умножения следующим образом:

$$F(2, 3) = \begin{bmatrix} d_0 & d_1 & d_2 \\ d_1 & d_2 & d_3 \end{bmatrix} \begin{bmatrix} g_0 \\ g_1 \\ g_2 \end{bmatrix} = \begin{bmatrix} r_0 \\ r_1 \end{bmatrix}.$$

Если в процессе расчета используется стандартное матричное умножение, требуется всего 6 умножений и 4 сложения. Однако можно обнаружить, что матрица, полученная путем преобразования входного сигнала в операции свертки,

не является произвольной матрицей и содержит большое количество регулярно повторяющихся элементов, таких как d_1 и d_2 . Виноград сделал следующие преобразования:

$$F(2, 3) = \begin{bmatrix} d_0, d_1, d_2 \\ d_1, d_2, d_3 \end{bmatrix} \begin{bmatrix} g_0 \\ g_1 \\ g_2 \end{bmatrix} = \begin{bmatrix} m_1 + m_2 + m_3 \\ m_2 - m_3 - m_4 \end{bmatrix},$$

где

$$\begin{aligned} m_1 &= (d_0 - d_2)g_0, \\ m_2 &= (d_1 + d_2)(g_0 + g_1 + g_2)/2, \\ m_3 &= (d_2 - d_1)(g_0 - g_1 + g_2)/2, \\ m_4 &= (d_1 - d_3)g_2. \end{aligned}$$

После завершения обучения сети элементы ядер свертки будут фиксированы, поэтому значения, связанные с g в приведенной выше формуле можно рассчитать заранее при инициализации модели. Весь этап вывода необходимо рассчитать только один раз, поэтому его можно игнорировать. Таким образом, здесь требуется 4 умножения и 8 сложений. Запишем приведенный выше процесс расчета в виде матрицы следующим образом:

$$Y = A^T[(Gg) \odot (B^T d)],$$

где \odot — поэлементное умножение (произведение Адамара); g — ядро свертки; d — карта входных объектов; G — матрица ядра свертки размера $(u + k - 1) * k$; B^T — входная матрица преобразования размера $(u + k - 1) * (u + k - 1)$; A^T — выходная матрица преобразования размера $(u + k - 1) * u$; u — размер вывода; k — размер ядра свертки. Конкретные значения каждой матрицы, следующие:

$$B^T = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix},$$

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 \\ 0 & 0 & 1 \end{bmatrix},$$

$$A^T = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & -1 \end{bmatrix},$$

$$d = [d_0, d_1, d_2, d_3]^T,$$

$$g = [g_0, g_1, g_2]^T.$$

В данной работе исследовалась сверточная нейронная сеть для распознавания рукописных букв на основе алгоритма Винограда, скорость работы которой по сравнению с исходной увеличилась в 12 раз. В работе рассматриваются результаты экспериментов и предпринята попытка проанализировать причины, по которым теоретическое значение ускорения не было достигнуто.

- [1] *Winograd S.* On Multiplication of Polynomials Modulo a Polynomial // *SIAM J. Computing.* — 1980. — P. 225–229.
- [2] *Lavin A., Gray S.* Fast Algorithms for Convolutional Neural Networks // *Proc. IEEE Conf. Computer Vision and Pattern Recognition.* — IEEE, 2016. — P. 4013–4022.

Обнаружение и оценка нарушений целостности межпанельных швов зданий на изображении

Кабанова Вероника Валерьевна^{1*}

nika_homework@mail.ru

*Логунова Оксана Сергеевна*¹

logunova66@mail.ru

*Наркевич Михаил Юрьевич*¹

narkevich_mu@mail.ru

*Корниенко Владимир Дмитриевич*¹

volodya.kornienko2319@yandex.ru

¹Магнитогорск, Магнитогорский государственный технический университет им. Г. И. Носова

Исследование посвящено разработке программного модуля для обнаружения и оценки нарушений целостности межпанельных швов зданий на изображении. Разработанный программный модуль внедрен в веб-приложение для контроля территорий, технического состояния зданий и сооружений и их частей на опасных производственных объектах металлургического производства (ПАО «ММК») с использованием беспилотных воздушных судов. В настоящее время веб-приложение находится на стадии тестирования. Модуль также можно применять для визуального обследования жилых панельных зданий на изображениях.

За последние несколько лет в схожих по тематике доклада исследований учеными были разработаны и описаны методы обработки изображений для обнаружения трещин в бетонных конструкциях, кирпичной кладке, но в данных работах не уделяется должного внимания обнаружению разрушений соединительных швов панельных зданий. Несвоевременный контроль за состоянием межпанельных швов может привести к дискомфорту проживающих (жилые здания) и к поломке оборудования из-за проникновения влаги, возникновения коррозии (промышленные предприятия). Также описанные алгоритмы обнаружения трещин с применением нейросетей и без искусственного интеллекта имеют ряд ограничений к входным данным и требований к процессу обработки, что только подчеркивает потребность в проведении дополнительных исследований и необходимость доработки существующих алгоритмов обработки изображений или создания пользовательских алгоритмов. Основным ограничением к входным данным является наличие однородного фона на изображении, трещина должна быть четко видна на фасаде без других дефектов. Обычно в качестве примеров обнаружения трещин приводятся небольшие фрагменты фасадов зданий на изображении, но для того, чтобы автоматизировать процесс визуального обследования здания, требуется обрабатывать изображение, на котором запечатлена крупная часть фасада здания вместе с неинформативными областями (деревья, окна, лестницы и т. п.).

Особенностью разработки, представленной в исследовании, является возможность обработки крупных частей фасадов зданий на изображениях, полученных с беспилотного летательного аппарата за счет передачи в программный модуль вектора исключенных областей на изображении. Обеспечивается

это тем, что перед обработкой эксперт исключает неинформативные области из изображения с высоким разрешением [1]. Также на фасаде здания могут быть другие дефекты и элементы, например, косметический ремонт, разрушения панелей. Наличие группы параметров позволяет корректировать результат обработки.

В работе предлагается методика обнаружения и оценки нарушений целостности соединительных швов панельных зданий, основанная на композиции методов изменения контраста изображения, размытия по Гауссу, адаптивного порога, преобразовании Хафа (Hough Transform) и разработанных в процессе исследования алгоритмах для группировки линий, найденных алгоритмом Хафа, и определения местоположения межпанельных швов и их разрушений на изображении в виде прямоугольников с учетом исключенных фрагментов изображения и возможностью подсчета площади швов и разрушений [2]. В этап предварительной обработки исходного изображения включены следующие методы: приведение изображения к полутоновому, размытие изображения, применение морфологических операций, пороговая обработка. На этапе распознавания межпанельных швов осуществляется поиск прямых линий, отбор найденных линий, построение швов-прямоугольников, разделение в соответствии с исключенными областями и подсчет площади межпанельных швов. При обнаружении разрушений межпанельных швов этап предобработки изображения аналогичен. Для выполнения процесса обработки изображения требуется установить значения или применить значения по умолчанию следующих параметров: контраст изображения, размер окрестности для вычисления порогового значения, гауссовский размер ядра, минимальная длина линии, окрестность обнаружения вертикального (горизонтального) шва, минимальное количество вертикальных (горизонтальных) линий в шве-прямоугольнике. Выходными данными программного модуля является процентная доля разрушений швов, а также два изображения, на которых выделены цветом обнаруженные межпанельные швы здания и разрушения.

Программный модуль реализован на языке программирования C++ в виде динамически подключаемой библиотеки и Python в виде ру-модуля с использованием библиотеки OpenCV [3]. Проведено тестирование программного модуля на изображениях производственных объектов, расположенных на территории ПАО «ММК», в частности, подвергался обследованию фасад здания главного корпуса ЛШЦ № 8 и здания склада № 54. Исходные изображения получены с помощью беспилотного летательного аппарата. Визуальный осмотр выходных изображений с обнаруженными швами и разрушениями позволил сделать вывод, что швы выделяются верно, и процентная доля разрушений швов сопоставима с количеством найденных межпанельных швов и разрушений. Для отладки программного модуля реализовано приложение для настольных компьютеров (desktop-приложение) с применением фреймворка Qt. Также разра-

ботано мобильное приложение (фреймворк Flutter) с возможностью обработки изображения фасада здания, полученного с камеры телефона или из галереи.

Работа выполняется в МГТУ им. Г.И.Носова по договору № 247715 «Разработка и применение методик контроля территорий, зданий и сооружений ПАО «ММК» с использованием беспилотных воздушных судов (БВС)» от 05.07.2021 г.

- [1] *Наркевич М. Ю., Логунова О. С., Корниенко В. Д., и др.* Интеллектуальная система принятия решений при оценке качества зданий и сооружений на опасных производственных объектах: определение траектории движения беспилотного летательного аппарата // Вестник Магнитогорского государственного технического университета им. Г.И.Носова. — 2022. — Т. 20, № 1. — С. 50–60.
- [2] *Кабанова В. В., Логунова О. С., Наркевич М. Ю., Корниенко В. Д.* Программное обеспечение для автоматизированного обнаружения и оценки разрушений соединительных швов зданий // Программные продукты и системы. — Тверь: АО НИИ ЦПС, 2023. — № 3. — С. 474–485.
- [3] *Кабанова В. В., Наркевич М. Ю., Корниенко В. Д., Логунова О. С., и др.* Automated building joint failure estimator: программа для ЭВМ // Свидетельство о гос. регистрации № 2022681316 от 22.11.2022.

Цифровая подпись на основе инвариантов геометрических преобразований

Козлов Вадим Никитович¹

vnkozlov@mail.ru

¹МГУ им. М. В. Ломоносова

В докладе представлен способ защиты изображения от подделки и проверки подлинности полученной информации на основе аффинных преобразований. Способ предназначен для работы с любыми изображениями (потенциально, в том числе, и с трехмерными), или с информацией, приводимой к изображениям (например, со звуком). Этот способ может быть использован для защиты канала связи между отправителем информации в виде изображения и получателем от попыток третьих сторон отправить получателю информацию под видом информации от отправителя, а также аутентификации полученного изображения.

Способ является аналогом известной электронной цифровой подписи, но на иной принципиальной базе. В основе принципа — теорема, доказанная более двадцати лет назад, и потому работу можно считать примером превращения довольно абстрактного математического результата в полезную технологию. Подробности — в [1]. Способ запатентован (два патента).

- [1] *Kozlov V. N. An Analog of the Electronic Digital Signature Based on Codes Defining Images with Accuracy up to Affine Transforms // Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications. — Road Town, United Kingdom: Pleiades Publishing, Ltd, 2023. — Vol. 33, No. 2. — P. 203–207.*

Генерация оптимальных приближений цветного изображения

Харинов Михаил Вячеславович¹

khar@iias.spb.su

¹Санкт-Петербург, СПб ФИЦ РАН

В докладе описывается новое направление исследований по аппроксимации изображений кусочно постоянными приближениями наилучшего качества, которое достигается, по крайней мере, в пределах визуального восприятия. Тема доклада связана с автоматизацией извлечения из изображения объективной первичной информации о сцене без использования априорных сведений об изображении и обучения, что требует модернизации классических методов кластерного анализа, разработанных до появления компьютеров и перехода к обработке больших данных. Текущее состояние работ отражено в [1–4]. Новизна заключается в том, что в данном докладе впервые представляется компьютерная программа, которая для входного цветного изображения генерирует на выходе последовательность оптимальных приближений в установленном диапазоне числа цветов, например, от 1 до 100.

Оптимальным считается кусочно-постоянное приближение изображения из N пикселей с минимальной суммарной квадратичной ошибкой $E = 3N\sigma^2$ при данном числе цветов, где σ — среднеквадратичная ошибка (RMSE).

Проблема практического вычисления оптимальных приближений является NP-трудной [5], что не исключает ее практического приближенного решения в приложении к установленному домену данных. Если в кластерном анализе для эффективного решения необходимо ограничить множество входных многомерных векторов, то в современном компьютерном зрении, наоборот, — объединить неформально разделяемые предметные области космических, медицинских и др. изображений и унифицировать методы распознавания за счет формализации понятий изображения, объектов на изображении, а также элементов тех и других, которые именуются *суперпикселями*.

Изображение описывается последовательностью оптимальных приближений в $g = 1, 2, \dots$ цветах. Предполагается, что ошибки E_g оптимальных приближений изображения составляют выпуклую последовательность, что проверяется экспериментально. При этом изображение является *полиерархической* структурой, которая, в общем случае, не является иерархической, но неоднозначно аппроксимируется выпуклыми иерархическими последовательностями приближений с заведомо ограниченными ошибками E_g .

Объекты описываются иерархическими последовательностями приближений, содержащими хотя бы по одному оптимальному приближению изображения. Каждое оптимальное приближение изображения в g_0 цветах порождает некоторую иерархию структурированных объектов, которая вычисляется по определенному алгоритму. При этом объекты из конкретной иерархии аппроксимируются объединениями или частями g_0 кластеров пикселей оптимального приближения изображения.

Суперпиксели задаются иерархией приближений изображения, которая порождается пересечением ряда из $g = 1, 2, \dots$ перекрывающихся оптимальных приближений изображения. Каждое из указанных приближения изображения позволяют воспроизвести без искажений ряд из соответствующего числа оптимальных приближений посредством слияния суперпикселей.

Перечисленные определения являются основой математической модели, которая не ограничивает контент изображения и не ориентирована на выделение и распознавание только заранее известных объектов (автомобильных номеров, дефектов печатных плат и пр.). При построении модели мы, с определенными оговорками, руководствуемся принципами, сформулированными в [6] для сегментации изображения, для которой требуются: а) единый алгоритм; б) численный критерий выбора лучшего из нескольких вариантов; в) инвариантность относительно масштабирования, поворота, сдвига, отражения и др. геометрических преобразований. При этом в качестве критерия мы рассматриваем суммарную квадратичную ошибку E приближений изображения, которая согласно кластерному анализу является стандартным критерием качества разбиения и соответствующего приближения изображения. Поскольку целевыми считаются не отдельные экземпляры, а кластеры объектов, вместо сегментации изображения мы выполняем более общую кластеризацию пикселей, тем более, что отождествление кластеров пикселей с сегментами изображения препятствует эффективной минимизации ошибки E . По той же причине при реальной минимизации суммарной квадратичной ошибки E исключается использование обучающих и иных априорных сведений об объектах интереса.

В качестве основного метода кластеризации и сегментации через кластеризацию вслед за [5] и ряда последующих работ по кластерному анализу применяется метод Уорда, который обеспечивает генерацию иерархии приближений в $g = 1, 2, \dots$ цветах, описываемых выпуклой последовательностью E_g суммарных квадратичных ошибок. В [5] метод Уорда применяется в сочетании с методом К-средних, который был придуман для упрощенных вычислений до изобретения компьютеров и продуцирует вычислительные ошибки в виде пустых кластеров (подробнее о К-средних см. в [1]). В нашей версии метод Уорда модернизируется для обработки *больших данных* в виде полииерархических структур. Хотя метод Уорда не учитывает геометрического положения пикселей, он характеризуется выраженной неустойчивостью значений E_g для приближений с малым числом цветов из-за многоитерационных вычислений. Результирующие приближения в нескольких цветах заметно меняются, например, при повороте изображения на 90 градусов, его незначительной модификации при сжатии с потерями и др. изменениях входных данных, а также зависят от алгоритма перебора пар пиксельных кластеров. Обнаруженная неустойчивость метода Уорда трактуется как свойство вариативности, которое обеспечивает стабильную генерацию оптимальных приближений за счет отбора для каждого числа цветов прибли-

жений с минимальной ошибкой E при многократной кластеризации по Уорду в сочетании с методом Монте-Карло.

Поскольку в базах видеоданных отсутствуют примеры оптимальных приближений, таковые получались посредством: а) укрупнения пикселей в первоначальные сегменты, содержащие в среднем не более 1,5–2 пикселей; б) кластеризацией сегментов по Уорду и минимизацией ошибки E в установленном диапазоне числа цветов [1, 4]. Чрезмерная вычислительная сложность кластеризации сотен тысяч начальных сегментов преодолевалась за счет параллельных вычислений. В презентуемой программе удалось ускорить вычисления в 1000 и более раз за счет применения метода Уорда по частям изображения, в котором: а) множество пикселей разбивается на p подмножеств, обрабатываемых по Уорду как самостоятельные изображения; б) выполняется откат вычислений от $g = p$ до $g = r$ кластеров из условия выпуклости последовательности ошибок E_g ; в) выполняется кластеризация по Уорду при различных значениях числа r начальных кластеров и отбором иерархий с минимальными значениями ошибки E для каждого числа цветов в установленном диапазоне.

Как подтверждает эксперимент [2,3], генерация оптимальных и соответствующих иерархических последовательностей приближений изображения коммутирует с масштабированием и др. геометрическими преобразованиями, а также устойчиво относительно преобразования цветного изображения в полутоновое. При этом устойчивость относительно изменения контента может обеспечиваться за счет автоматической настройки параметров. Таким образом, генерация оптимальных приближений изображения обеспечивает инвариантную раскраску изображения, которая упрощает детектирование и распознавание объектов.

Исследование выполнено при поддержке государственного бюджета, номер проекта FFZF-2022-0006.

- [1] *Khanykov I., Nenashev V. and Kharinov M.* Algebraic Multi-Layer Network: Key Concepts // J. Imaging. — 2023. — No. 9. — P. 146.
- [2] *Kharinov M.* About Computer Vision using Optimal Image Approximations // Proc. of the 16th Int. Conf. PRIP'2023, Belarus, Minsk, Oct. 17–19, 2023. — P. 305–311.
- [3] *Харинов М. В.* Полииерархическая структура изображения и инвариантное детектирование объектов // ГрафиКон 2023: 33-я Межд. конф. по компьютерной графике и машинному зрению, 19–21 сентября 2023 г., Москва (в печати).
- [4] *Nenashev V., Khanykov I. and Kharinov M.* A Model of Pixel and Superpixel Clustering for Object Detection // J. Imaging. — 2022. — No. 8. — P. 274.
- [5] *Murtagh F. and Pierre L.* Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? // J. Classif.. — 2014. — No. 31 — P. 274–295.
- [6] *Koepfler G.* Segmentation by Minimizing Functionals and the Merging Methods // World Congress of Nonlinear Analysts'92, Proc. of the First World Congress of Nonlinear Analysts, Tampa, FL, USA, 19–26 August 1992; Lakshmikantham, V., Ed.. — Berlin, Germany; Boston, MA, USA: DeGruyter, 1996. — P. 1933–1942.

Штриховая сегментация рукописного текста

Местецкий Леонид Моисеевич^{1,2*}

mestlm@mail.ru

¹Москва, МГУ

²Москва, ВШЭ

Рассматривается задача структурного анализа рукописного текста по цифровому бинарному изображению. Предлагается новый подход к анализу структуры письма, состоящий в представлении изображения текста в виде совокупности каллиграфических элементов — штрихов. Разработан метод штриховой сегментации рукописного текста, основанный на построении непрерывного скелетного графа бинарного изображения рукописного документа [1]. Штрихи представляют собой циклы и цепи — подграфы скелетного графа. Предлагается метод разрезания скелета на такие подграфы. Каждый штрих рассматривается как след от движения пера. Для этого определяется начальная и конечная точки штриха и направление движения пера. В качестве приложения рассматривается задача распознавания рукописного текста на основе штриховой сегментации. Предложен метод решения, стоящий в реконструкции всей траектории пера по штриховому представлению с последующим распознаванием этой траектории с помощью нейросети.

Штриховая сегментация представляет собой реконструкцию траектории пишущего инструмента — пера с выделением в этой траектории отдельных фрагментов, соответствующих каноническому набору базовых штрихов. Понятие штриха иллюстрируется изображениями в прописях, которые используются для обучения детей в начальной школе. Это понятие включает в себе как сами элементы письма, так и маршрут движения пера при их написании. Узнаваемость этих элементов при чтении напрямую связана с правильными каллиграфическими навыками пишущего. Читатель распознает их, поскольку у него те же навыки письма, что и у писателя.

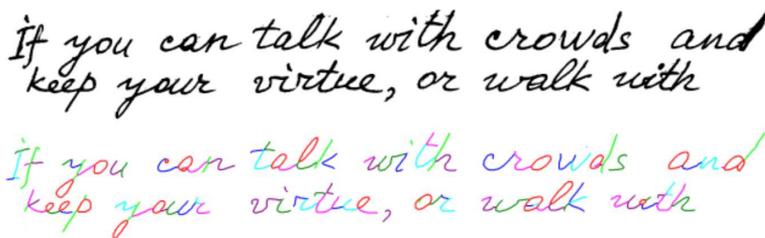


Рис. 1. Исходное растровое изображение рукописного текста и его штриховое представление

Таким образом, задача штриховой сегментации рукописного текста выглядит следующим образом (рис. 1). Исходными данными для нее являются циф-

ровые изображения, полученные в результате сканирования и бинаризации рукописных документов. Рукописный текст на них представлен черными пикселями, а фон — белыми. Выходом должно быть упорядоченное множество штрихов в виде линий с указанием начала, конца и направления движения пера при их написании.

В статье представлен подход к построению штрихового представления и пример практического приложения применительно к решению задачи расшифровки рукописных текстов.

Предлагаемое решение основано на следующих принципах.

1. Штриховая сегментация осуществляется на основе непрерывного медиального представления цифрового бинарного изображения. Медиальное представление состоит из скелета, образованного срединными осями, и радиальной функции, описывающей ширину объекта относительно точек скелета. Медиальное представление имеет вид нагруженного геометрического графа.
2. Штрихи формируются в виде подграфов скелетного графа. Выделяются штрихи-кольца и штрихи-ветви, представляющие собой циклы и цепи графа соответственно.
3. Для каждого штриха определяются его начальная и конечная точки. Для штрихов-ветвей эта пара точек задает направление движения пера. В штрихах-кольцах направление по часовой стрелке или против вычисляется в зависимости от положения штрихов относительно базовых линий текста.
4. Все выделенные штрихи упорядочиваются в последовательности их написания.

Непрерывный скелет, полученный на основе аппроксимации изображения текста многоугольными фигурами, дает возможность получить штриховое представление рукописи в виде кортежа канонических каллиграфических элементов. При этом штрихи выделяются на основе корректной математической модели как подграфы скелетного графа. Вычислительные эксперименты показывают реализуемость и эффективность подхода. В дальнейших исследованиях предполагается применить предложенный метод в задачах идентификации почерка, для навигации в больших архивах рукописных документов, а также в других задачах, связанных с анализом и распознаванием рукописных текстов.

Работа поддержана грантом РФФИ № 22-68-00066.

- [1] Местецкий Л. М. Непрерывная морфология бинарных изображений. — М., Физматлит, 2009.

3D-реконструкция поверхности Марса на основе SFM

*Бобков Александр Валентинович*¹

Дай Ифань^{1*}

Alexander.Bobkov@bmstu.ru

daiyifan1997@outlook.com

¹Москва, МГТУ им. Н. Э. Баумана

В данной работе рассматривается задача трехмерной реконструкции поверхности по видеоизображению для планирования маршрута автономного робота-марсохода. Современный этап освоения космического пространства тесно связан с исследованием ближайших планет солнечной системы при помощи роботов с той или иной степенью автономности. Удаленность планет от Земли порождает существенные задержки в распространении сигналов, что усложняет использование телеуправления и делает перспективным использование именно автономных аппаратов.

Одна из наиболее важных и сложных проблем здесь связана с необходимостью планирования и отработки безопасного маршрута движения робота-марсохода.

Трехмерные (3D) модели поверхности важны для миссий по исследованию планет и научных исследований, особенно для таких миссий, как выбор места посадки [1], геоморфологические исследования [2], морфометрический анализ [3] и геологические исследования [4]. Существующее оборудование 3D-съемки и реконструкции, такое как камеры глубины и лазерные системы, дорогое и сложное в эксплуатации. 3D-реконструкция на основе изображений часто требует некоторой предварительной информации, такой как положение камеры и априорные данные 3D-точек, последовательно записанные изображения и т. д.

Решить вышеуказанные проблемы может метод восстановления наблюдаемой поверхности по видеоизображению (structure from motion, SFM). Этот метод позволяет одновременно восстанавливать положение камеры и параметры трехмерных точек из перекрывающихся изображений, тем самым восстанавливая разреженную трехмерную структуру сцены. Метод SFM изначально был разработан для восстановления небольших сцен, записанных по фотографиям небольшого размера с низким пространственным разрешением. Непосредственное использование SFM для реконструкции марсианской поверхности может потребовать много вычислительного времени, и это может повлиять на эффект реконструкции.

SFM состоит из трех основных этапов: извлечение признаков одного изображения, сопоставление признаков каждой пары изображений и расчет параметров в рамках итеративного выполнения (bundle adjustment, BA).

На первом этапе, затраты времени линейно связаны с количеством изображений, и решение обычно состоит в использовании аппаратного ускорения графического процессора или оптимизации детектора функций.

На втором этапе, когда для описания ключевой точки используются локальные признаки, такие как алгоритм масштабно-инвариантного преобразования

признаков SIFT [5] и Approximate Nearest Neighbours (ANN), по которым производится сопоставление, неправильное сопоставление всегда неизбежно. Классическим методом геометрической верификации совпавших признаков является RANSAC (согласованность случайной выборки) [6], однако он имеет временные затраты, которые увеличиваются экспоненциально с увеличением процента найденных признаков, и его производительность резко снижается. Чтобы повысить эффективность и надежность реконструкции, необходимо решить проблему надежного обнаружения и фильтрации локальных признаков.

С другой стороны, количество совпадающих пар также повлияет на временные затраты на их сопоставление. Когда перекрытие изображений слишком велико и генерируется слишком много совпадающих пар, вычислительные затраты на их сопоставление существенно возрастает. В этом случае эффективной стратегией ускорения является отбор совпадающих пар [7].

На третьем этапе для определения параметров положения камеры и структуры сцены итеративное выполнение ВА приводит к высоким вычислительным затратам в дополнительном рабочем процессе SFM. Традиционный инкрементальный алгоритм SFM чувствителен к выбору начальной пары изображений. В случае небольших задач ВА легко сходится к глобально оптимальному решению. Тем не менее, ВА может занять очень много времени и с трудом поддается сходимости в крупномасштабных задачах, если не доступна хорошая первоначальная оценка.

Методы восстановления поверхности достаточно сложны, и для автономных роботов они сейчас не применяются ни на Земле, ни на Луне. Для земных задач методы восстановления поверхности не используют, потому что здесь имеется разветвленная сеть дорог, по которым можно передвигаться, не сильно анализируя их качество. Поэтому данные темы мало проработаны, и надежные алгоритмы, способные работать в реальном времени, на данный момент практически отсутствуют. Луна расположена достаточно близко от нас, сигнал до Луны и обратно доходит за три секунды, и поэтому есть возможность реализовать простое и надежное прямое телеуправление, что и было реализовано на всех луноходах. Методы восстановления поверхности здесь оказались не востребованы.

С Марсом ситуация иная. Марс имеет сложную поверхность с разнообразными, зачастую труднопреодолимыми формами рельефа, а сигналы до Марса идут десятки минут. Здесь телеуправление становится слишком неудобным и медленным: марсоходом приходится управлять очень маленькими, аккуратными и медленными шажочками. Это порождает необходимость разработки новых быстродействующих и надежных алгоритмов визуального восприятия среды именно для марсианских программ. В качестве решения указанных проблем в данной работе предлагается метод 3D-реконструкции поверхности Марса с использованием SFM на основе MVE (Multi-View Environment). Для сильно перекрывающихся изображений Марса используются четыре критерия для получения надежных начальных совпадающих пар, а также фильтруются точки,

находящиеся слишком далеко в трехмерном пространстве, и точки с большими ошибками перепроецирования. Экспериментальные результаты показывают, что надежная фильтрация совпадающих пар может ускорить провести реконструкцию и эффективно реконструировать поверхность Марса.

- [1] *Wu B., Li F., Hu H. et al.* Topographic and Geomorphological Mapping and Analysis of the Chang'E-4 Landing Site on the Far Side of the Moon // Photogrammetric Engineering & Remote Sensing. — 2020. — Vol. 86, No. 4. — P. 247–258.
- [2] *Hynek B. M., Beach M.* Updated Global Map of Martian Valley Networks and Implications for Climate and Hydrologic Processes // J. Geophysical Research: Planets. — 2010. — Vol. 115.
- [3] *Brooker L. M., Blame M. R., Conway S. J. et al.* Clastic Polygonal Networks around Lyot Crater, Mars: Possible Formation Mechanisms from Morphometric Analysis // Icarus. — 2018. — Vol. 302. — P. 386–406.
- [4] *Head J. W., Fassett C. I., Kadish S. J. et al.* Global Distribution of Large Lunar Craters: Implications for Resurfacing and Impactor Populations // Science. — 2010. — Vol. 329. — P. 1504–1507.
- [5] *Lowe G.* Sift-the Scale Invariant Feature Transform // Int. J. Computer Vision. — 2004. — Vol. 60, No. 2. — P. 91–110.
- [6] *Derpanis K. G.* Overview of the RANSAC Algorithm // Image Rochester NY. — 2010. — Vol. 4, No. 1. — P. 2–3.
- [7] *Jiang S., Jiang C., Jiang W.* Efficient Structure from Motion for Large-Scale UAV Images: A Review and a Comparison of SfM Tools // ISPRS J. Photogrammetry and Remote Sensing. — 2020. — Vol. 167. — P. 230–251.

Некоторые вопросы снижения функционала летательного аппарата, оснащенного корреляционно-экстремальной навигационной системой в условиях возможных помех

Мостовой Дмитрий Николаевич^{1*}

dmitriy.mostovoi@mail.ru

*Алчинов Александр Иванович*¹

alchinov46@mail.ru

¹Москва, Институт проблем управления им. В. А. Трапезникова РАН

В качестве альтернативы спутниковым навигационным системам (СНС), которые требуют для своего функционирования непрерывной связи между навигационной аппаратурой потребителя и созвездием навигационных спутников, используются корреляционно-экстремальные навигационные системы (КЭНС), работающие автономно. В условиях, когда СНС могут не работать или они выведены из строя, КЭНС, основанные на обработке текущей и эталонной информации становятся единственным способом точного определения местоположения ЛА. Ретроспективный анализ истории развития КЭНС [1], проведенный современными учеными, показывает, что до конца 1990-х гг. исследования и разработки по различным аспектам развития КЭНС велись интенсивно. На разных этапах решение проблем управления движением и навигацией было связано с поиском общих принципов синтеза, которые бы при конкретных обстоятельствах позволяли автоматически синтезировать параметры управления движущимися объектами. Были определены направления дальнейшего развития КЭНС:

- новые принципы построения бортовых алгоритмов, их интеллектуализация и самоорганизация;
- применение новых типов съемочных систем и их комплексирование;
- развитие параллельных процессоров, включая и специализированные процессоры для реализации алгоритмов с единой параллельной структурой.

Уровень развития информационных технологий и достижимые характеристики бортовых вычислителей тогда сдерживали практическую реализацию указанных выше направлений. Однако в настоящее время появился потенциал для их реализации. В связи с интенсивным развитием БПЛА различного назначения, оснащенных современными сенсорными системами и мощными бортовыми вычислителями, появилась необходимость ускоренного развития КЭНС, в том числе путем моделирования. Для маскировки объекта от датчиков КЭНС разработано большое количество методов, средств и алгоритмов, однако отсутствуют достаточно удобные средства моделирования условий применения различного рода помех, которые позволили бы планировать и оптимизировать применение средств стрессового воздействия, в частности, средств маскировки.

В связи с этим требуется разработать имитационную модель работы КЭНС в районе применения, анализирующую изображения попиксельно, локально, зонально и глобально, использующую соответствующее математическое обеспечение. В рамках этого направления был реализован алгоритм обработки изобра-

жений, основанный на минимизации корреляции между изображением до применения средств маскировки и изображением того же участка местности после их применения. В результате его работы формируется карта помеховой обстановки (КПО), предоставляющая информацию о необходимом изменении характеристик яркости участков местности маскируемого объекта. Опираясь на эту информацию, проектировщик маскировки будет делать выбор в пользу того или иного средства или метода маскировки. На основе КПО, например, может быть рассчитана структура и состав пенных маскирующих покрытий, поскольку толщина слоя пены и ее состав определяют свойства покрытия в разных спектрах длин волн ЭМИ. Наряду с этим разработанное программное обеспечение позволяет синтезировать изображения для тестирования помехоустойчивости алгоритмов корреляционно-экстремальной навигации и предоставлять информативно и наглядно требуемое изменение яркости объекта, который необходимо замаскировать от систем технического зрения, использующих корреляционно-экстремальные алгоритмы.

В настоящее время в мире разрабатываются многоспектральные средства маскировки [2] с использованием новых композитных материалов из нескольких слоев, которые могут быть использованы при управлении значениями контрастов фоно-целевой обстановки в видимом, тепловом, радиотепловом и радиолокационном диапазонах длин волн электромагнитного излучения. Установлено, что многоспектральные средства маскировки более эффективно изменяют фоно-целевую обстановку участка местности в момент работы датчика КЭНС. Это целиком исключает возможность решения навигационной задачи. В данной работе функционирование алгоритма продемонстрировано на примере маскировки в видимом диапазоне длин волн, однако такой же подход применим и к другим диапазонам ЭМИ и материалам, используемым при маскировке. Таким образом, предложенный способ имитационного моделирования КЭНС может обеспечить получение условий невыполнения КЭНС своей задачи в районе применения за счет определения требуемых изменений фоно-целевой обстановки и обоснования необходимых параметров средств маскировки и выбора соответствующих средств маскировки для практического применения.

- [1] *Алимов А. И., Гороховский И. Н.* Концепция создания прикладной географической информационной системы моделирования поисковых корреляционно-экстремальных систем автономной навигации // Проблемы управления. — 2022. — № 1. — С. 54–66.
- [2] *Данилочев Д. В., Павлушенко М. И., Волохов В. И.* Патент № 2704147 С1 Российская Федерация, МПК F41H 3/00. Способ маскировки объектов от много- и гиперспектральных средств разведки: № 2018139313; заявл. 08.11.2018; опубл. 24.10.2019 / Д. В. Данилочев, М. И. Павлушенко, В. И. Волохов; заявитель — Федеральное государственное казенное военное образовательное учреждение высшего образования «Военная академия Ракетных войск стратегического назначения имени Петра Великого» МО РФ. — EDN WGHRTJ.

Метод ассоциации данных семантического SLAM на основе моделей вероятностного вывода

Цзинь Цзюнь¹*

Шэнь Синь¹

У Го¹

Веселовская Ольга Александровна¹

tszints@student.bmstu.ru

shs20ua057@student.bmstu.ru

ug@student.bmstu.ru

volga@bmstu.ru

¹Москва, Московский государственный технический университет им. Н. Э. Баумана

Точное позиционирование и картографирование являются основой для надежного автономного выполнения роботами навигационных задач. Традиционные алгоритмы визуального SLAM (одновременная локализация и картографирование) более подвержены ошибкам сопоставления в случае, если геометрические особенности неочевидны или среда существенно меняется. По сравнению с традиционными алгоритмами SLAM, semantic SLAM использует семантические ориентиры, полученные камерами, для оценки и оптимизации поз робота. Алгоритм обнаружения ориентиров глубокой нейронной сетью используется для получения информации о категории и местоположении семантических ориентиров. Однако обнаружение семантических ориентиров не обеспечивает никакой связи данных между изображениями. Роботу необходимо было бы установить ассоциацию данных об обнаружении ориентиров на изображениях с разных ракурсов. Хорошая ассоциация данных обеспечивает сходимость алгоритма SLAM. Исследования ассоциации данных обычно основаны на предположении о низком уровне шума и статических помехах. Значительные шумы наблюдения, вызванные помехами от освещения, теней и движущихся объектов в реальных сценах, приводят к большим ассоциативным ошибкам при ассоциации смысловых ориентиров.

В статье предлагается улучшенный метод непараметрического вывода, основанный на процессе Дирихле для ассоциации семантических ориентиров, когда шум наблюдения велик, а количество семантических ориентиров неизвестно. Апостериорное распределение вероятностей семантических ориентиров получено на основе априорной вероятности Дирихле и вероятности наблюдения. Во время ассоциации апостериорное распределение вероятностей итеративно оптимизируется с использованием метода Гиббса для получения оптимальной ассоциации и оптимизации позы робота. На основе описанного выше метода непараметрического вывода предложен алгоритм семантического SLAM.

Были проведены эксперименты для демонстрации эффективности предлагаемого подхода с помощью смоделированных и реальных данных. Набор модельных данных содержит 1098 наблюдений, случайным образом сгенерированных по 15 семантическим ориентирам. Шум наблюдения соответствует гауссову распределению $N(0, 0.2^2)$. Результаты работы нашего алгоритма с моделируемым набором данных показаны на рисунке 1. Чтобы проверить надежность алгоритма, мы провели сравнительные эксперименты с алгоритмом DP-SLAM,



Рис. 1. Результаты ассоциации данных с набором данных моделирования

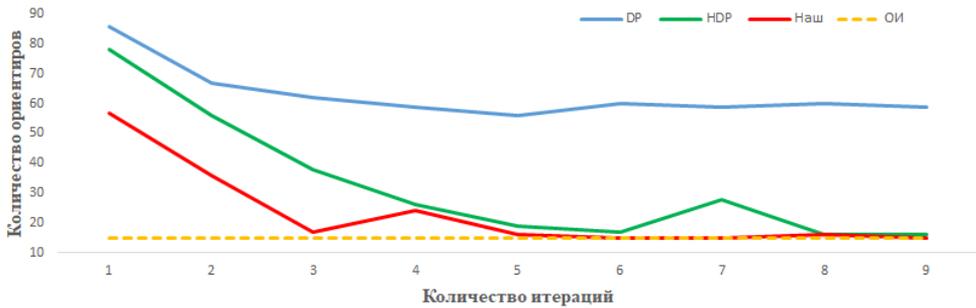


Рис. 2. Результаты сходимости нашего алгоритма, DP и HDP с шумом

предложенным в [1], и алгоритмом HDP-SLAM, предложенным в [2]. Результаты показаны на рисунке 2. Видно, что наш алгоритм сходится быстрее и обладает более высокой надежностью.



Рис. 3. Результат обнаружения семантического ориентира

В экспериментах на реальных сценах набор данных монокуляра KITTI [3] использовался для проверки точности позиционирования робота. Выбранная дорожная сцена содержит достаточное количество семантических объектов. Транспортные средства, появляющиеся в сцене, используются в качестве семантических ориентиров для ассоциации данных. Обнаружение YOLO используется для извлечения информации о семантических ориентирах (см. рис. 3. Сравнение среднеквадратичной ошибки (RMSE) траектории движения набора данных KITTI показано в таблице 1. Алгоритмами, используемыми для сравнительных экспериментов, являются алгоритм DP-SLAM и алгоритм ORB SLAM 2 [4]. Чем меньше RMSE, тем ближе траектория движения к реальной траектории, т. е. тем выше точность позиционирования робота. Экспериментальные результаты, приведенные в таблице 1, показывают, что среднеквадратичная ошибка алгоритма, предложенного в этой статье, улучшилась на наборе данных, подтверждая, что алгоритм обладает более высокой согласованностью траектории и точностью позиционирования.

Таблица 1. Сравнение ошибок траектории в наборе данных KITTI

| ряд | наш | DP-SLAM | ORB-SLAM2 |
|-----|------------------|------------------|-----------|
| 00 | 10,061578 | 11,684571 | 80,164613 |
| 05 | 7,631594 | 8,436478 | 40,168613 |
| 07 | 7,980654 | 9,136484 | 19,016113 |
| 08 | 17,064585 | 18,419821 | 46,168461 |
| 09 | 14,256484 | 19,648911 | 67,164761 |
| 10 | 11,094617 | 10,365161 | 29,164614 |

- [1] *Mu B., Liu S.* SLAM with objects using a nonparametric pose graph // IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). — IEEE, 2016. — P. 4602–4609.
- [2] *Zhang J., Gui M.* Hierarchical Topic Model Based Object Association for Semantic SLAM // IEEE Transactions on Visualization and Computer Graphics. — IEEE, 2019. — P. 3052–3062.
- [3] *Fritsch J., Kühnl T.* A New Performance Measure and Evaluation Benchmark for Road Detection Algorithms // 16th International IEEE Conference on Intelligent Transportation Systems. — IEEE, 2013. — P. 1693–1700.
- [4] *Mur-Artal R., Tardós J.* ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras // IEEE Transactions on Robotics. — IEEE, 2017. — P. 1255–1262.

Применение беспризнакового подхода к задаче распознавания активностей человека на основе базисной совокупности скелетных моделей

Сурков Егор Эдуардович^{1*}

eg-su@mail.ru

*Середин Олег Сергеевич*¹

oseredin@yandex.ru

*Копылов Андрей Валериевич*¹

and.kopylov@gmail.com

¹Тула, Тульский государственный университет

Область распознавания активностей человека в настоящее время становится все более актуальной, что обусловлено совершенствованием нейросетевых технологий, вычислительных ресурсов, а также ростом количества данных для обучения сетей. Использование скелетной модели в задачах, связанных с распознаванием активности человека, является общепринятым стандартом, одной из причин чего является сохранение приватности представления данных, снижающее беспокойство людей к системам мониторинга. Разработка системы признаков для представления скелетной модели не является тривиальной задачей и имеет большое количество методов ее решения [1]. Методы делятся на четыре группы: первые методы используют положения вершин скелетной модели, вторые — формируют геометрические характеристики скелетной модели (расстояние между двумя точками, площадь между тремя точками), третья группа использует не вершины, а соединения вершин, соответствующие частям тела человека. В нашей предыдущей работе [1] предложен метод, относящийся к четвертой группе методов — представлению скелетных моделей подходящей мерой их попарного сходства или различия. Таким образом, следуя идее беспризнакового распознавания образов, предлагается использовать вторичное признаковое пространство для представления скелетных моделей и работать только с парами скелетов, а именно с мерой их несходства.

В [1] описан метод решения задачи детектирования падений на основе базисной совокупности скелетных моделей с применением беспризнакового подхода к распознаванию образов. А также предложен набор базисной совокупности эталонных скелетных моделей, мера непохожести двух скелетных моделей, и определено понятие карты активности, которая представляет собой матрицу стекированных векторов значений функции непохожести двух скелетных моделей (рис. 1). Учитывая, что объекты базисной совокупности являются элементами метрического пространства, а структура карты активностей зависит от порядка объектов базисной совокупности [1], базисная совокупность упорядочена методом поиска кратчайшего незамкнутого пути между объектами, предложенным в [2]. Размерность карты активности определяется количеством объектов базисной совокупности и размером скользящего окна [1]. В работе выполнено обучение и тестирование сверточной нейронной сети ResNet50V2 [3] на базе данных TST Fall Detection v2 [4]. Качество распознавания вычислено для каждого актера отдельно согласно методу leave-out-person-out [1]. Усредненная оценка

ассигнату для всей базы данных зафиксирована как 0,954. Указанная величина является отношением правильно классифицированных видеопоследовательностей к общему количеству видео в базе данных.

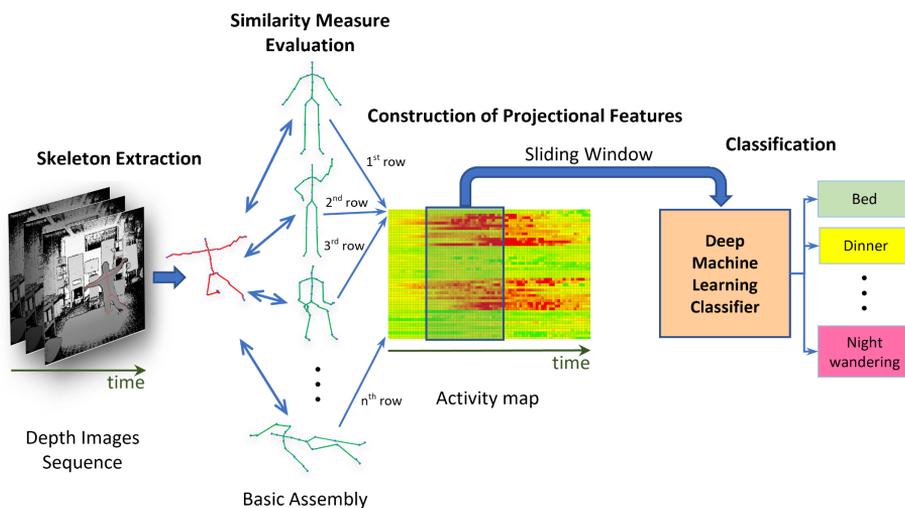


Рис. 1. Представление скелетных моделей для распознавания активности человека на основе анализа карты активности

Высокий показатель качества решения задачи детектирования падений человека на основе базисной совокупности и карты активности человека стимулирует к применению такого метода к решению более общей задачи — распознаванию активности человека.

Экспериментальная часть в работе проводилась на базе данных TST Fall Detection dataset v2 [4]. База данных содержит 12 активностей, разделенных на две группы Fall и ADL. В работе выполнена экспертная покadroвая разметка базы данных и выделены следующие активности для решения задачи классификации: «walking», «staying», «sitting», «falling», «grasp», «ungrasp», «post-fall-laying», «laying», «noclass». Меткой «noclass» обозначены кадры, которые находятся внутри переходных процессов из одного класса в другой, например sit-to-stand.

В работе применяется скорректированная архитектура сверточной нейронной сети с пробрасыванием связей ResNet50V2 [3], и с механизмом внимания Squeeze-and-Excitation (SE-block), предложенным в [5]. SE-block применяется после блоков свертки для обучения модели выявлению межканальных зависимостей и селекции наиболее важных признаков. Модель обучается при помощи оптимизации функции категориальной кросс-энтропии при помощи оптимизации

тора Adam с параметром $\text{learning rate} = 0,001$. Входной тензор имеет размер $32 \times 32 \times 1$. Размер тензора обусловлен количеством отобранных базисных скелетных моделей и размером скользящего окна [1].

Обучающая выборка разделена на стратифицированные батчи, каждый из которых содержит 128 карт активностей. Выборка и батчи перемешиваются каждую эпоху, количество эпох 115. Каждая карта активностей имеет соответствующий целевой вектор, в котором закодирована принадлежность объекта к меткам классов методом one-hot.

В результате экспериментов произведена покадровая оценка качества многоклассовой задачи классификации на 9 классах и вычислены оценки Accuracy = 0,83, F1 = 0,747. Показатели качества демонстрируют хорошее решение, полученное на картах активности. В дальнейшем планируется разработать технику аугментации на уровне скелетных представлений, применение одномерной свертки, провести эксперименты с техникой Transfer Learning, предварительно обучив модель на более объемной базе данных NTU 120 dataset, а также применить модель машинного обучения Трансформер.

Работа выполнена при финансовой поддержке Министерства науки и высшего образования РФ в рамках государственного задания FEWG-2021-0012.

- [1] *Surkov E. E., Seredin O. S., Kopylov A. V., Huang S.-C.* The Basic Assembly of Skeletal Models in the Fall Detection Problem // *Computer Optics*. — 2023. — Vol. 47, No. 2. — P. 323–334.
- [2] *Surkov E. E., Seredin O. S., Kopylov A. V.* Locally Optimal Solutions in the Shortest Unclosed Path Search Problem // *IEEE Ural-Siberian Conf. Biomedical Engineering, Radioelectronics and Information Technology (USBREIT)*, 2023, Yekaterinburg, Russian Federation. — P. 221–224.
- [3] *He K., Zhang X., Ren S.* Identity Mappings in Deep Residual Networks // *Computer Vision—ECCV 2016: 14th European Conf., Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14*, Springer International Publishing, 2016. — P. 630–645.
- [4] *Gasparri S., Cippitelli E., Gambi E., Spinsante S., Wahlen J., Orhan I., Lindh T.* Proposal and Experimental Evaluation of Fall Detection Solution Based on Wearable and Depth Data Fusion // *Braverman Readings in Machine Learning. Key Ideas from Inception to Current State*. — Stockholm: Springer, 2016. — P. 99–108.
- [5] *Hu J., Shen L., Sun G.* Squeeze-and-Excitation Networks // *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2018. — P. 7132–7141.

Дескриптор формы бинарной фигуры на основе профиля вращения

*Середин Олег Сергеевич*¹

oseredin@yandex.ru

*Ломов Никита Александрович*¹

nikita-lomov@mail.ru

*Ляхов Даниил Викторович*¹

liakhov.daniil@mail.ru

Митюгов Никита Сергеевич^{1*}

nikita.mityugov.2001@mail.ru

*Кушнир Олеся Александровна*¹

kushnir-olesya@rambler.ru

*Копылов Андрей Валериевич*¹

and.kopylov@gmail.com

¹Тула, Тульский государственный университет

В области компьютерного зрения и обработки изображений задача анализа формы играет ключевую роль в распознавании объектов. Форма объекта является важным аспектом его визуальной характеристики, содержащим информацию о его структуре и природе. Для достижения точности и эффективности в анализе формы, необходимы специализированные способы описания, называемые дескрипторами формы. Дескриптор формы должен с одной стороны обеспечивать точное извлечение и компактное представление информации о форме, а с другой стороны обеспечивать инвариантность относительно возможных преобразований, например, таких как масштабирование, сдвиг и поворот, что делает построение дескрипторов сложной задачей.

В работе предлагается новый дескриптор формы, основанный на профиле вращения [1]. Он представляет собой последовательность из некоторого фиксированного набора чисел, которые были получены путем поворота фигуры относительно самой себя на угол 180 градусов с некоторым шагом. В данной работе мы использовали шаг в один угловой градус. Для каждого угла поворота от 1 до 180 градусов этой последовательности рассчитывается мера Жаккара, которая отражает степень сходства между исходной и повернутой фигурами. Таким образом, предлагаемый дескриптор содержит 180 чисел в замкнутом интервале от 0 до 1. Профиль вращения позволяет описать зависимость меры Жаккара от угла поворота в заданной точке при выбранной дискретизации по углам. Дополнительным свойством дескриптора является то, что мы можем получить информацию о том, насколько симметрична фигура при повороте на разные углы относительно определенной точки [2, 3].

Предложенный дескриптор обладает инвариантностью к смещению, повороту, отражению и масштабированию, что делает его эффективным в анализе объектов различных форм. При этом стоит отметить, что в данной работе рассматривается дескриптор на бинарном изображении, хотя возможны обобщения на случай полутоновых изображений. Дополнительно, представлена эффективная алгоритмическая реализация вычисления профиля вращения, которая дает в среднем трех-пятикратное ускорение по сравнению с экстенсивной базовой процедурой. На рисунке 1 представлены профили, принадлежащие листьям березы и липы.

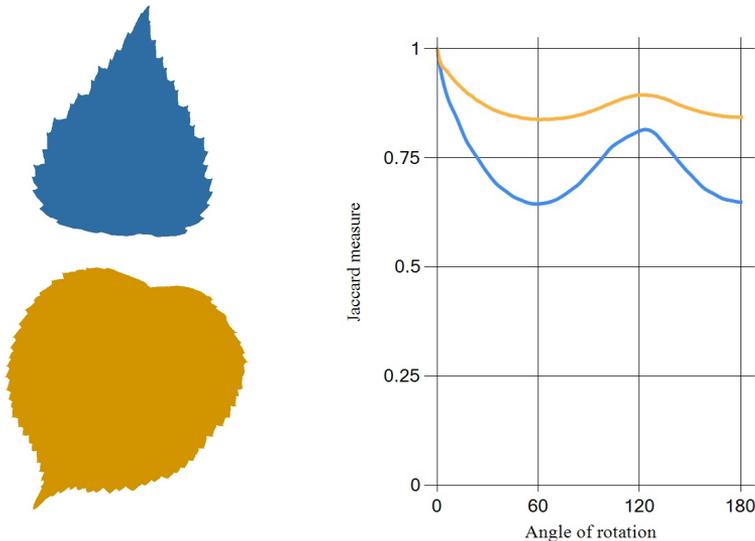


Рис. 1. Сравнение профилей, принадлежащих листьям березы и липы, отмеченных синим и оранжевым цветом соответственно

Исследование профиля вращения было произведено на четырех задачах классификации изображений. На базе растений, включающей сканированные изображения листьев деревьев двух распространенных видов *Tilia cordata* (липа) и *Betula pendula* (береза) [4], решались первые две задачи: отделение листьев липы от листьев березы и отделение листьев растений одного и того же вида, которые растут в разных по экологическому состоянию точках.

Третьей задачей оценки классификации с помощью профиля вращения является задача с детектированием кистей рук на цифровых изображениях из базы http://lda.tsu.tula.ru/papers/TulaSU_HandsDetDB.zip, описанная в [5].

Для четвертой задачи классификации использовалась известная база бинарных растровых изображений листьев растений FLAVIA. В базе содержится 1907 изображений листьев размером 800 на 600 пикселей, тридцати двух различных классов. Среднее количество объектов в классе — 50–70.

Для каждой из задач были применены классификаторы на основе метода опорных векторов (SVM), градиентного бустинга (XGBoost) и глубоких нейронных сетей, обучающие данные для которых представляли собой рассчитанные профили вращения с соответствующими индексами классов. Для сравнения использовались непосредственно цветные изображения и их бинаризованные силуэты, а в качестве классификаторов применялись сверточные нейронные сети. Таким образом, мы оценивали влияние предложенного дескриптора на каче-

ство работы разных классификаторов в сравнении с другими классификаторами, обученными непосредственно на цветных и бинарных изображениях. Стоит отметить, что качество обучения оценивалось с помощью кросс-валидации. В каждом случае обучающее множество было поделено на десять частей. Исходя из полученных результатов, профиль вращения не занимал первых мест, однако в каждом случае разница от наилучшего результата составляет около 1–3 %. Таким образом, можно сделать вывод о высокой эффективности и устойчивости предложенного дескриптора формы в различных контекстах классификации, а также значительном потенциале данного дескриптора для задач анализа формы в области компьютерного зрения.

Исследование выполнено при поддержке гранта Российского научного фонда № 22-21-00575, <https://rscf.ru/project/22-21-00575/>.

- [1] Ломов Н. А., Ляхов Д. В., Середин О. С. Применение вычислительно эффективных альтернатив поворота изображения в задаче поиска вращательной симметрии бинарных фигур // Интеллектуализация обработки информации: Тезисы докладов 14-й Международной конференции, г. Москва 2022 г. — М.: Российская академия наук, 2022. — С. 243–246.
- [2] Seredin O., Liakhov D., Kushnir O., Lomov N. Jaccard Index-Based Detection of Order 2 Rotational Quasi-Symmetry Focus for Binary Images // Pattern Recognition and Image Analysis. — 2022. — Vol. 32, No. 3. — P. 672–681.
- [3] Lomov N., Seredin O., Kushnir O., Liakhov D. Search for Rotational Symmetry of Binary Images via Radon Transform and Fourier Analysis // In Proc. Int. Joint Conf. Computer Vision, Imaging and Computer Graphics Theory and Applications VISAPP. — 2023. — Vol. 4. — P. 280–289.
- [4] Семин А. Н., Местецкий Л. М., Турас Х. П. Извлечение признаков формы для классификации экологического состояния по изображениям листьев // Интеллектуализация обработки информации: Тезисы докладов 14-й Международной конференции, г. Москва 2022 г. — М.: Российская академия наук, 2022. — С. 232–234.
- [5] Kopylov A., Seredin O., Kushnir O., Gracheva I., Larin A. Background-Invariant Robust Hand Detection based on Probabilistic One-Class Color Segmentation and Skeleton Matching // ICPRAM. — 2018. — P. 503–510.

Применение механизмов внимания и увеличения набора данных для задач обнаружения мелких объектов

Шэнь Синь^{1*}

shs20ua057@student.bmstu.ru

*У Го*¹

ug@student.bmstu.ru

*Цзинь Цзюнь*¹

tszints@student.bmstu.ru

*Лукьянов Вадим Викторович*¹

vdmlknv@yandex.ru

¹Москва, Московский государственный технический университет имени Н. Э. Баумана

Важность обработки данных мобильных роботов обусловлена их применением в различных областях, где мульти-модальные данные открывают новые возможности для исследований [1]. В области компьютерного зрения и глубокого обучения, которое становится все более популярным, были изучены основные аспекты, включая последние достижения, такие как YOLOv8 — передовой однопоточный детектор.

Проблема, которую изучает наша команда, возникла в результате открытого конкурса компьютерного зрения, проводимого в Китае. Набор открытых данных представляет собой данные, собранные двумя разными датчиками: камерами SAR (радар с синтезированной апертурой) и RGB. Объекты, которые необходимо идентифицировать и разметить на изображении, — это различные типы транспортных средств, припаркованных на земле. Поскольку датчик установлен на роботе и находится далеко от земли, размер объектов на снимке очень мал и представляет собой типичный мелкий объект. Из-за большого размера изображений SAR трудно различить конкретные типы транспортных средств, поэтому существует только единственная категория: автомобили. Тогда как в области видимого света можно точно определить тип транспортного средства, например: легковой автомобиль, внедорожник, фургон, пикап и т. д.

В архитектуре YOLO обычно присутствуют три слоя или масштаба, составляющих «шею», что позволяет учитывать различные перспективы и масштабы. YOLOv8 включает три слоя обнаружения: P3, P4, P5 для эффективной работы с объектами разного масштаба: маленький, средний, большой. Эти слои позволяют модели с точностью обнаруживать как крупные, так и мелкие объекты. Для обнаружения мелких объектов введен специализированный слой P2. Слой обнаружения мелких объектов P2 специально разработан для улучшения способности модели распознавать и локализовать мелкие объекты на изображении. Он работает в меньшем масштабе, что необходимо для обнаружения мелких, часто сложных целей. Кроме того, в качестве основной структуры в «шею» была интегрирована BiFPN (Двунаправленная полностью сверточная сеть, от англ. Bidirectional Fully Convolutional Network), улучшающая использование и слияние семантической и локализационной информации, что критически важно для точного обнаружения объектов.

Механизм внимания (от англ. attention mechanism) часто используется как техника усовершенствования в задачах обнаружения объектов. Интеграция механизма внимания позволяет модели нейронной сети улучшить способность оценивать значимость различных сегментов входной последовательности, что приводит к более точным и контекстуально осведомленным прогнозам. В исследовании были проанализированы различные механизмы внимания для данных видимого света (VI) и синтетической апертуры радара (SAR) через абляционные эксперименты, подчеркивая их важность в процессе извлечения признаков и вклад в эффективность предложенного подхода.

Аугментация данных является ключевым элементом проекта, поскольку увеличение разнообразия и богатства набора данных необходимо для создания устойчивых моделей глубокого обучения [1]. Искусственное введение изменений в данные фактически повышает способность модели к обобщению с обучающего набора на новые данные. Общие техники аугментации, такие как добавление шума, горизонтальное отражение (вертикальное отражение может быть двусмысленным в контексте изображений) и аффинные преобразования (включающие вращение, перемещение и масштабирование в трех измерениях), служат нескольким целям: имитируют внутренние неопределенности реальных данных и вводят вариации в размер, положение и ориентацию объектов. В совокупности они повышают способность модели к обобщению, делая ее более универсальной и адаптируемой к реальным сценариям.

Благодаря экспериментальной проверке новая структура сети глубокого обучения, которую мы разработали, достигла лучших результатов, чем оригинальная YOLOv8 в этом сценарии и задаче. Для набора данных VI обучение проводилось на Ubuntu 20.04 с GPU NVIDIA GeForce RTX 4060 на ноутбуке с 8 ГБ памяти. Для набора данных SAR, обучение проходило на сервере с двумя GPU NVIDIA GeForce RTX 4090, каждая с 24 ГБ памяти.

В данной работе мы используем набор данных SAR-изображений, размер которых составляет 4096×4096 пикселей в формате TIF. Исходный набор данных содержит 60 изображений, и был значительно за счет методов аугментации данных, при этом размер отдельных изображений был уменьшен, что сделало их более подходящими для использования с улучшенным детектором на базе YOLO. Данные были разделены на обучающую, тестовую и валидационную выборки в соотношении 8 : 1 : 1, насчитывая 11 290, 1412 и 1411 изображений, соответственно. Также использовался набор данных, собранный с помощью RGB-камеры, включающий 6000 изображений, собранных роботом с различных ракурсов, с размером каждого изображения 640×512 пикселей.

Модели сравнивались по следующим показателям:

1. Точность (Precision) — мера, которая показывает, какой процент обнаруженных объектов действительно соответствует реальным объектам:

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

где TP (True Positives) — это количество правильно обнаруженных объектов, а FP (False Positives) — количество ошибочно обнаруженных объектов.

2. Полнота (Recall) — мера, которая показывает, какой процент реальных объектов был обнаружен:

$$\text{Recall} = \frac{TP}{TP + FN}.$$

Наша сеть стремится достичь оптимального баланса между точностью и полнотой, что позволяет ей эффективно работать в разнообразных сценариях обнаружения объектов.

3. mAP@.5 — средняя точность пересечения по объединению при пороге 0,5 (средняя точность при IoU = 0,5).
4. mAP@.5:.95 — средняя точность пересечения по объединению в диапазоне от 0,5 до 0,95 (средняя точность при IoU от 0,5 до 0,95).

| модель | Precision | Recall | mAP@.5 | mAP@.5:.95 |
|----------------------|-----------|--------|--------|------------|
| YOLOv8s+BiFPN+P2 | 91,5 % | 68,4 % | 76,4 % | 45,1 % |
| YOLOv8s+BiFPN+P2+MCA | 92,1 % | 72,2 % | 79,1 % | 47,9 % |

Таблица 1. Сравнение двух моделей по данным VI

| модель | Precision | Recall | mAP@.5 | mAP@.5:.95 |
|--------------------------------|-----------|--------|--------|------------|
| YOLOv8x+BiFPN+P2 | 99,0 % | 98,6 % | 99,1 % | 95,6 % |
| YOLOv8x+BiFPN+ +P2+BiFormer | 99,3 % | 98,8 % | 99,5 % | 93,5 % |

Таблица 2. Сравнение двух моделей по данным SAR

- [1] *Oyku S., Ozer S.* Yolodrone: Improved Yolo Architecture for Object Detection in Drone Images // 2021 44th Int. Conf. Telecommunications and Signal Processing (TSP). — IEEE, 2021. — P. 361–365.

Ускоренные алгоритмы формирования сегментов из суперпикселей

Мурашов Дмитрий Михайлович^{1*}

d_murashov@mail.ru

Мурашова Елизавета Дмитриевна¹

mur-lizon@mail.ru

¹Москва, ФИЦ ИУ РАН

Рассматривается задача сегментации цифровых изображений. В [1] был предложен метод обеспечения качества сегментации на основе принципа минимизации меры информационной избыточности на картах сегментации, полученных с помощью суперпиксельного алгоритма SLIC [2] и процедуры объединения суперпикселей в сегменты при разных значениях параметра. Из множества разбиений V_q , $q = \overline{1, Q}$ оригинального изображения U с разным числом сегментов выбиралось разбиение, которое минимизировало меру избыточности $R(U, V_q) = 1 - I(U; V_q)/H(V_q)$, где $I(U; V_q)$ — средняя взаимная информация между входом и выходом алгоритма сегментации, $H(V_q)$ — энтропия выхода. Этот алгоритм достаточно трудоемкий, и возникает необходимость ускорить получение оптимального разбиения.

Целью представляемой работы является создание ускоренных алгоритмов объединения суперпикселей в сегменты, которые представляют объекты, зафиксированные на изображении. Предлагаются две модификации алгоритма, описанного в [1], и три алгоритма на основе «жадной» стратегии, которые отличаются условиями принятия решения об объединении суперпикселей. Проводится сравнение предлагаемых алгоритмов на тестовых изображениях.

Базовый алгоритм объединения суперпикселей [1] включает следующие операции. Для выбранного значения параметра выполняется перебор центров всех суперпикселей изображения, полученного на выходе алгоритма SLIC, и проверяется условие $d(C_i, C_j) \leq ct$, где C_i, C_j — векторы координат центров соседних суперпикселей в выбранном цветовом пространстве, d — евклидово расстояние между центрами суперпикселей с номерами i и j , c — константа, t — параметр. Формируется матрица M , элементы которой $m_{ij} = 1$, если центры близки, и $m_{ij} = 0$ в противном случае. После завершения формирования матрицы выполняется объединение суперпикселей, которые соответствуют ненулевым элементам этой матрицы, в сегменты. Полученное разбиение V_q оригинального изображения сохраняется, выбирается новое значение параметра t , повторяется сканирование суперпиксельного изображения, и т. д. Далее из полученного множества разбиений выбирается разбиение V_{qmin} , на котором минимизируется мера избыточности $R(U, V_q)$. В предлагаемых модификациях описанного базового алгоритма используются условия объединения, которые включают требование невозрастания меры избыточности $R(U, V_q)$: $R_{l+1}(U, V_q) \leq R_l(U, V_q)$, где l — номер шага объединения суперпикселей. В этом случае не требуется формировать большого числа разбиений для всего диапазона значений параметра t

Таблица 1. Значения информационных мер качества для разбиений, полученных разными алгоритмами

| Алгоритм / Условие объединения | $R(U, V_{qmin})$ | $VI(U, V_{qmin})$ | $VI(V_{GT}, V_{qmin})$ |
|---|------------------|-------------------|------------------------|
| Базовый / $d(C_i, C_j) \leq ct$ | 0,4366 | 0,7532 | 0,2991 |
| Базовый / $d(C_i, C_j) \leq ct$; $R_{l+1}(U, V_q) \leq R_l(U, V_q)$ | 0,4219 | 0,7392 | 0,5449 |
| Базовый / $R_{l+1}(U, V_q) \leq R_l(U, V_q)$ | 0,4219 | 0,7392 | 0,5449 |
| Жадный / $d(C_i, C_j) \leq ct$ | 0,1924 | 0,7017 | 0,3249 |
| Жадный / $d(C_i, C_j) \leq ct$; $R_{l+1}(U, V_q) \leq R_l(U, V_q)$ | 0,1724 | 0,6547 | 0,5051 |
| Жадный / $R_{l+1}(U, V_q) \leq R_l(U, V_q)$ | 0,1111 | 0,6541 | 0,3681 |

и выбирать оптимальное. Указанное обстоятельство обусловило более высокое быстродействие новых алгоритмов.

В «жадных» алгоритмах объединение суперпикселей выполняется непосредственно в процессе сканирования суперпиксельного изображения. Суперпиксель с номером i объединяется с соседним суперпикселем j , если выполняется некоторое условие. Далее полученный новый суперпиксель объединяется с соседним и т. д. Если условие объединения не выполняется для всех соседних суперпикселей, то переходим к следующему, и т. д. до достижения стабильности или минимально допустимого количества сегментов. Как и в описанных выше модификациях базового алгоритма используются три различных условия объединения. В первом из предлагаемых алгоритмов применяется условие, наложенное на расстояние в цветовом пространстве между объединяемыми объектами. Во втором алгоритме необходимо, чтобы в результате слияния суперпикселей не возрастала мера $R(U, V_q)$. В третьем алгоритме условием объединения является комбинация этих двух условий.

Проведен вычислительный эксперимент на изображениях из базы BSDS500. Алгоритм SLIC разбивал изображение на 651 суперпиксель. Далее суперпиксели объединялись в сегменты с помощью представленных выше алгоритмов. Результаты эксперимента для одного из тестовых изображений приведены в таблицах 1 и 2. В таблице 1 представлены значения меры избыточности для разбиений V_{qmin} оригинального изображения, полученных разными алгоритмами, меры информационного различия (вариации информации [3]) $VI(U, V_{qmin})$ оригинального изображения U и полученного разбиения V_{qmin} , а также меры информационного различия $VI(V_{GT}, V_{qmin})$ полученного разбиения V_{qmin} и эталонного разбиения V_{GT} .

Из данных, приведенных в таблице 1, следует, что, во-первых, на разбиениях, порождаемых «жадными» алгоритмами, достигаются меньшие значения ме-

Таблица 2. Число сегментов в разбиениях тестового изображения и время работы алгоритмов

| Алгоритм / Условие объединения | Число сегментов | Время работы алгоритма, с |
|---|-----------------|---------------------------|
| Базовый / $d(C_i, C_j) \leq ct$ | 28 | 151,86 |
| Базовый / $d(C_i, C_j) \leq ct$; $R_{l+1}(U, V_q) \leq R_l(U, V_q)$ | 113 | 99,79 |
| Базовый / $R_{l+1}(U, V_q) \leq R_l(U, V_q)$ | 113 | 21,34 |
| Жадный / $d(C_i, C_j) \leq ct$ | 5 | 20,78 |
| Жадный / $d(C_i, C_j) \leq ct$; $R_{l+1}(U, V_q) \leq R_l(U, V_q)$ | 2 | 22,59 |
| Жадный / $R_{l+1}(U, V_q) \leq R_l(U, V_q)$ | 2 | 2,81 |

ры избыточности и информационного отличия от оригинального изображения. Во-вторых, алгоритмы, в которых при объединении суперпикселей/сегментов используется условие невозрастания величины меры избыточности, позволяют получить меньшее значение меры $R(U, V_{qmin})$ для разбиения V_{qmin} и меньшую величину информационного различия $VI(U, V_{qmin})$ по сравнению с алгоритмами, в которых суперпиксели объединяются только из условия, наложенного на расстояние между их центрами в цветовом пространстве. В-третьих, разбиения, полученные алгоритмами с условием объединения $d(C_i, C_j) \leq ct$, имеют наибольшее сходство с эталонным разбиением из базы BSDS500 (т.е. меньшее значение вариации информации $VI(V_{GT}, V_{qmin})$). При этом наибольшее сходство обеспечивает базовый алгоритм.

В таблице 2 приведены данные о количестве сегментов разбиений тестового изображения и времени, за которое были получены эти разбиения. Результаты показывают, что «жадная» стратегия объединения суперпикселей позволяет существенно ускорить выполнение операции сегментации изображений. Наиболее быстро работает «жадный» алгоритм, использующий условие объединения только на основе меры избыточности, однако, выявлено, что этот алгоритм не для всех входных изображений позволяет получить приемлемое разбиение. Поэтому целесообразно в алгоритмах сегментации комбинировать критерии.

- [1] *Murashov D. M.* An Information Model for Digital Image Segmentation // Pattern Recognition and Image Analysis. — 2021. — Vol. 31, No. 4, — P. 632–645.
- [2] *Achanta R., Shaji A., Smith K., Lucchi A., Fua P., Susstrunk S.* SLIC Superpixels Compared to State-of-the-Art Superpixel Methods // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2012. — Vol. 34, No. 11, — P. 2274–2282.
- [3] *Meila M.* Comparing Clusterings: An Axiomatic View // Proc. 22nd International Conference on Machine Learning (ICML 2005). — 2005. — P. 577–584.

Оптимизация разрешения в сканирующем туннельном микроскопе: математические подходы и экспериментальные результаты

*Зырянов Константин Сергеевич*¹* zyryanov.konstantin.sergeevich@gmail.com

*Орешкин Андрей Иванович*¹ oreshkin@spmlab.phys.msu.su

*Чуличков Алексей Иванович*¹ achulichkov@gmail.com

¹Москва, МГУ имени М. В. Ломоносова

Одним из эффективных методов исследования наномасштабных объектов на поверхности материалов, биологических образцов и др., является метод сканирующей зондовой микроскопии (СЗМ) рельефа. Регистрация информации о рельефе поверхности получается при помощи сканирования поверхности образца. В процессе регистрации эта информация оказывается искаженной вследствие неидеальности системы, наличия шумов и др. В том числе одной из проблем СЗМ является компенсация искажений. В данном исследовании был применен ряд классических методов обработки и анализа изображений для компенсации искажений в оценках рельефа, полученных при помощи СЗМ.

Имеется набор данных D , полученных из сканирующего туннельного микроскопа (СТМ). Эти данные представляют собой измерения туннельного тока, преобразованные в изображения. Обозначим каждое изображение как I_i , где i — индекс изображения. Основной целью этого исследования является улучшение качества изображений и оценка точности результатов применения различных методов. Исследование состояло из следующих этапов: преобразование данных туннельного тока в изображение, вычитание фона, выравнивание данных, фильтрация и выравнивание строк [1, 2].

Каждое изображение I_i может быть представлено как матрица пикселей. Пусть $I_i(x, y)$ — интенсивность пикселя на позиции (x, y) [3]. Вводится оператор улучшения изображения F , где $I_{enhanced} = F(I_i)$, который может включать в себя различные методы обработки изображений. Оценка качества изображения Q может быть формализована как функция, оценивающая насколько лучше становится видно атомы после применения оператора: $Q(I_{enhanced}) > Q(I_i)$.

На основе данных D и цели улучшения точности обработки изображений, ставится задача выбора оператора обработки изображений F , который максимизирует оценку качества Q на наборе изображений:

$$F^* = \arg \max_F \left[\sum_{i=1}^N Q(F(I_i)) \right].$$

Таким образом, задача формулируется как поиск оптимального оператора F , который повышает качество изображений I_i в соответствии с выбранной оценкой качества Q .

- Отношение сигнал-шум: измеряет соотношение между сигналом (полезной информацией) и шумом на изображении.

$$Q_{SNR}(I) = \frac{\frac{1}{N} \sum_{j=1}^N x_j}{\frac{1}{N} \sqrt{\sum_{j=1}^N (x_j - \frac{1}{N} \sum_{k=1}^N x_k)^2}},$$

где N — количество пикселей изображения.

- Усредненная дисперсия по строкам измеряет изменение интенсивности вдоль строк изображения и усредняет его по всем строкам:

$$Q_{RAV}(I) = \frac{1}{M} \sum_{j=1}^M \frac{1}{N} \sum_{i=1}^N (x_{ij} - \frac{1}{N} \sum_{k=1}^N x_{kj})^2,$$

где M — количество строк, N — количество пикселей в каждой строке, x_{ij} — значение пикселя в i -м столбце j -й строки.

- Усредненная дисперсия по окнам оценивает изменение интенсивности в локальных областях изображения:

$$Q_{WAV}(I) = \frac{1}{K} \sum_{k=1}^K \frac{1}{N^k} \sum_{i=1}^{N^k} (x_{ik} - \frac{1}{N^k} \sum_{j=1}^{N^k} x_{jk})^2,$$

где K — количество окон, N^k — количество пикселей в k -м окне, x_{ik} — значение пикселя в i -м окне.

Как показали эксперименты, наиболее эффективными алгоритмами обработки СЗМ изображений из каждой группы методов: вычитание фона, выравнивание строк и фильтрация, оказались соответственно медианное выравнивание, выравнивание строк на основе усеченных срединных разностей соседних строк и применение сглаживающего фильтра Гаусса. Методы из группы выравнивания данных не показали удовлетворительных результатов ни по одной из метрик. Ниже предоставлены результаты вычислительных экспериментов и оценка ключевых метрик для исходного изображения и для каждого из вышеописанных методов соответственно:

Таблица 1. Результаты отдельных методов

| Q | Исходное изображение | Медианное выравнивание | Усеченные срединные разности | Фильтр Гаусса |
|-----|----------------------|------------------------|------------------------------|------------------|
| SNR | 0,004 | 0,016 | 0,075 | 0,007 |
| RAV | $358 \cdot 10^6$ | $176 \cdot 10^6$ | $84 \cdot 10^6$ | $152 \cdot 10^6$ |
| WAV | $191 \cdot 10^6$ | $131 \cdot 10^6$ | $31 \cdot 10^6$ | $25 \cdot 10^6$ |

Также проведены эксперименты по совмещению наилучших алгоритмов из каждой группы методов обработки СЗМ изображений для достижения наилучшего результата:

Таблица 2. Результаты совмещения методов

| Q | Исходное изображение | Комбинация методов |
|-----|----------------------|--------------------|
| SNR | 0,004 | 0,094 |
| RAV | $358 \cdot 10^6$ | $47 \cdot 10^6$ |
| WAV | $191 \cdot 10^6$ | $7 \cdot 10^6$ |

Данное исследование представляет собой значимый шаг в направлении улучшения процесса обработки и анализа СЗМ изображений. Классические методы показали определенные положительные результаты, также удалось определить наиболее эффективные и позволяющие получить высокую точность методы обработки СЗМ изображений. Однако в будущем исследовании стоит обратить внимание на перспективы использования методов глубокого обучения, таких как сверточные нейронные сети, чтобы улучшить точность обработки данных СЗМ изображений и обеспечить более высокое качество по введенным метрикам.

- [1] *Миронов В. Л.* Основы сканирующей зондовой микроскопии. — РАН, Институт физики микроструктур, 2004.
- [2] *Anguiano E.; Aguilar M.* A Cross-Measurement Procedure (CMP) for Near Noise-Free Imaging in Scanning Microscopes // Ultramicroscopy. — 1999. — Vol. 76, No. 1–2. — P. 39–47.
- [3] *Пытьев Ю. П.; Чуличков А. И.* Методы морфологического анализа изображений. — М.: ФИЗМАТЛИТ, 2010.

Мера сходства элементов штрихового представления рукописного текста на основе Фурье-дескриптора

Феоктистов Дмитрий Дмитриевич^{1*}

feoktistovdd@my.msu.ru

Местецкий Леонид Моисеевич^{1,2}

mestlm@mail.ru

¹Москва, МГУ

²Москва, ВШЭ

В работе рассматривается подход к распознаванию русскоязычных рукописных текстов, основанный на штриховой сегментации цифровых изображений текстовых документов. Штриховая сегментация состоит в представлении изображения текста в виде последовательности каллиграфических элементов письма, называемых штрихами: овалов, палочек, крючков и т.п. Штриховая сегментация осуществляется методами, основанными на непрерывном медиальном представлении изображения рукописного текста [1].

Распознавание текста предлагается свести к количественной оценке сходства и различия отдельных штрихов, а также кортежей штрихов. Такой подход основывается на следующей гипотезе: все почерки являются вариацией некоторого эталонного «правильного» варианта написания каллиграфических элементов. Действительно, обучение письму производится с помощью прописей, в которых в течение многих лет, на протяжении жизни нескольких поколений, не меняются правила написания штрихов, из которых строятся буквы и слова. В основе подхода лежит разработка меры сходства и различия штрихов, полученных в результате штриховой сегментации. Целью является построение дискриминативной метрики на пространстве штрихов [2]. В нашей работе исследуется метод построения такой метрики на основе так называемых Фурье-дескрипторов [3, 4] штрихов.

Алгоритм, описанный в [1], представляет каждый штрих в виде ломаной $L = \{(x_j, y_j)\}_{j=1}^l$. На нее можно посмотреть, как на последовательность точек в комплексной плоскости $\{x_j + iy_j\}_{j=1}^l$, т.е. как на сигнал, для получения векторного представления которого предлагается использовать дискретное преобразование Фурье. Для первичного определения того, является ли предлагаемое векторное представление дискриминативным была рассмотрена задача классификации штрихов набора панграмм, изображенного на рисунке 1.

В первую очередь была произведена проверка того, как ведет себя отношение внутриклассового расстояния к межклассовому в евклидовой метрике в зависимости от количества используемых коэффициентов Фурье. Как видно из рисунка 2, при всех значениях наблюдается выполнение гипотезы компактности, при этом при увеличении количества коэффициентов наблюдается деградация разделимости классов, т.к. появляются коэффициенты, отвечающие за шумы. Из этого можно сделать вывод, что для описания штрихов достаточно использовать небольшое количество коэффициентов Фурье.

аэрофотоъёмка ландшафта уже
 вывела земли богачей и
 процветающих крестьян.
 всё ускоряющаяся эволюция
 компьютерных технологий
 предъявила жесткие требования
 к производителям как собственно
 вычислительной техники, так и
 периферийных устройств.
 завершен ежегодный съезд
 эрудированных школьников, мечтающих
 глубоко проникнуть в тайны
 физических явлений и химических
 реакций.
 съези же ещё этих легких
 французских булок да выпей чаю.
 флегматичная эта верблюдица
 искуёт у подвезда засыхающий
 горячий шиповник.
 мифробальница попросту забыла
 ряд ключевых миксителей
 и тэлов.

Рис. 1. Изображение текста, для штрихов которого рассматривалась задача классификации

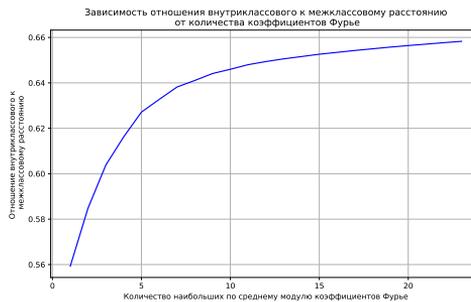


Рис. 2. Результаты эксперимента, проверяющего гипотезу компактности для различного количества коэффициентов Фурье

Для проверки предложенной гипотезы решалась задача классификации штрихов методом k ближайших соседей, качество измерялось на кросс валидации с пятью разбиениями выборки с помощью F1-масо метрики, т. к. она учитывает дисбаланс классов. Для каждого количества коэффициентов Фурье выбиралось лучшее значение k по результатам кросс валидации, после чего изучались результаты этого алгоритма. Как видно из рисунка 3 существенной разницы между использованием 5–23 коэффициентов Фурье нет: и значение метрики, и ее устойчивость примерно одинаковы. Значит, мы можем использовать

малое количество признаков, например, 5 в случае евклидова расстояния и 7 в случае косинусного. Отметим, что идею с выбором расстояния следует развить, т. к. косинусное расстояние дает большие значения F1-масго, хоть и менее устойчивые. Это говорит о том, что может существовать более дискриминативное расстояние, чем рассмотренные. Для этой цели предлагается использовать алгоритмы обучения метрик.

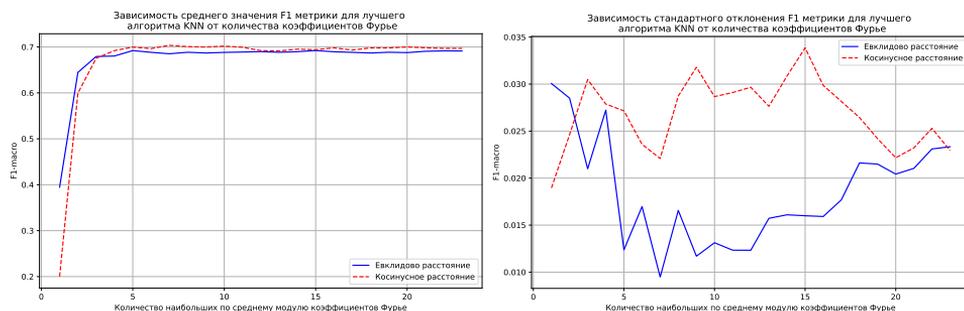


Рис. 3. Результаты эксперимента, проверяющего гипотезу компактности для различного количества коэффициентов Фурье

Таким образом, в работе предложен способ векторизации штрихов с помощью дискретного преобразования Фурье. Экспериментально показано, что для такого представления штрихов выполнена гипотеза компактности, также показано, что для оптимального описания достаточно 5–7 коэффициентов. При дальнейшем развитии работы будет исследована возможность применения алгоритмов обучения метрик для получения расстояния более дискриминативного, чем евклидово и косинусное. После чего найденное расстояние будет применено к задаче поиска ключевых слов в рукописном контексте.

Работа поддержана грантом РФФ № 22-68-00066.

- [1] Местецкий Л. М. Штриховая сегментация рукописного текста // Математические методы распознавания образов: Тезисы докладов 21-й Всероссийской конференции с международным участием, г. Москва 2023 г., М.: Российская академия наук, 2023. — С. 118–119.
- [2] Пронина Н. М., Местецкий Л. М. Классификация штрихов рукописного текста на основе расстояния Фреше (Устный) // 33-я Международная конференция по компьютерной графике и машинному зрению «ГрафиКон 2023», г. Москва, ИПУ РАН, 19–21 сентября 2023.
- [3] Zahn C. T., Roskies R. Z. Fourier Descriptors for Plane Closed Curves // IEEE Transactions on Computers. — 1972. — Vol. 21, No. 3. — P. 269–281.
- [4] Yip R., Tam P., Leung D. Application of Elliptic Fourier Descriptors to Symmetry Detection under Parallel Projection // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 1994. — Vol. 16. — P. 277–286.

Обнаружение лиц с помощью гибридных детекторов

*Веселов Дмитрий Иванович*¹

veselow@yandex.ru

*Фан Анатолий Чонг*¹

acfanfan@fa.ru

*Гаврилин Павел Сергеевич*¹

pav.gavr5@icloud.com

Андрянов Никита Андреевич^{1*}

naandriyanov@fa.ru

¹Москва, Финансовый университет при Правительстве Российской Федерации

Одной из наиболее интересных задач в наше время является задача распознавания лиц. С ростом числа приложений, в которых используются такие технологии (системы доступа, умные магазины, оплата в метро), растут и требования к качеству. В статье рассмотрено применение новых архитектурных гибридных решений для задачи идентификации лиц на сложных данных. Под сложными данными подразумеваются изображения, на которых присутствует множество лиц разного размера с различными мешающими факторами. Такие лица, например, находятся в бенчмарке Wider Face [1]. В работе показано, что гибридная архитектура Detectron_v2 [2], которая оперирует трансформером SWIN-T [3], дообученным в настоящей работе для идентификации лиц, лучше справляется с поставленной задачей по сравнению с моделями MTCNN [4] и Haar Cascades [5].

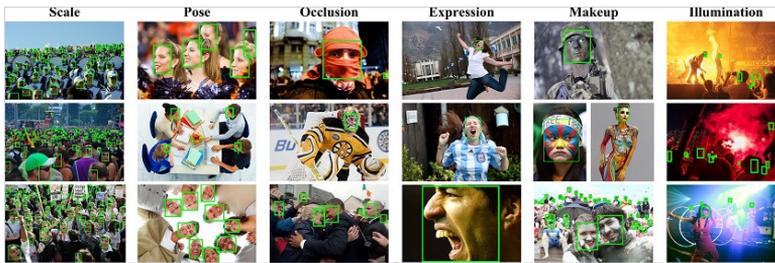


Рис. 1. Примеры изображений со сложными условиями обнаружения лиц

Задача обнаружения и идентификации лиц, с одной стороны, считается одной из тех задач, которые уже достаточно хорошо решены сегодня, однако, с другой стороны, когда дело доходит до усложняющих обстоятельств, то точность распознавания значительно падает. Это связано с увеличением числа лиц на изображении, с маленькими размерами лиц и т. д. Для детекции лиц в таких условиях требуются модели высокой точности. Классические детекторы работают по методу Виола-Джонса [6] и позволяют обнаруживать лица в режиме близком к реальному времени. Однако их скорость и точность существенно проигрывает на данных датасета Wider Face (рис. 1). То же самое можно сказать и про сверточные детекторы, но в меньшей степени.

Мы выбрали 3 категории изображений Wider Face: чирлидеры, собрания и пикники, с целью исследования различных архитектур для детекции. Также можно использовать дополнительно генеративные модели [8], чтобы усложнить процесс обнаружения, однако в этой работе такой подход не применяется. На этих данных использовали несколько подходов — классический Хаар, сверточные сети, трансформерные сети и гибридные модели. В таблице 1 показаны основные результаты экспериментов. В качестве основной метрики выбрана метрика средней точности mAP.

Таблица 1. Сравнительный анализ детекторов лиц

| Модель | mAP |
|----------|------|
| Haar | 0,21 |
| MTCNN | 0,64 |
| DETR [7] | 0,69 |
| SWIN-T | 0,71 |

Анализ данных в таблице показывает, что наилучшую точность обеспечивает гибридная модель свертки и трансформера на базе архитектуры SWIN-T из «зоопарка» Detectron. Выигрыш по сравнению с трансформерами составляет порядка 2 %. Таким образом, был предложен алгоритм детектирования лиц, отличающейся высокой точностью в сложных условиях.

- [1] Онлайн-набор данных. <https://shuoyang1213.me/WIDERFACE/> (дата обращения: 12.11.2023).
- [2] Open Source репозиторий модели детектирования. <https://github.com/facebookresearch/detectron2> (дата обращения: 20.11.2023).
- [3] Open Source репозиторий модели детектирования. https://github.com/xiaohu2015/SwinT_detectron2 (дата обращения: 23.11.2023).
- [4] *Xiang J., Zhu G.* Joint Face Detection and Facial Expression Recognition with MTCNN // 4th Int. Conf. Information Science and Control Engineering (ICISCE), Changsha, China, 2017. — P. 424–427.
- [5] *Андрянов Н. А., Орлов Е. А.* Разработка модели машинного обучения для оценки состояния глаз водителя // Инженерный вестник Дона. — 2022. — Т. 5, № 89. — С. 142–159.
- [6] *Viola P., Jones M.* Rapid Object Detection Using a Boosted Cascade of Simple Features // Proc. 2001 IEEE Computer Society Conf. Computer Vision and Pattern Recognition. CVPR 2001, Kauai, HI, USA, 2001. — P. I–I.
- [7] *Carion N., Massa F., Synnaeve G., Usunier N., Kirillov A., Zagoruyko S.* End-to-End Object Detection with Transformers // arXiv, 2005. — <https://arxiv.org/abs/2005.12872>
- [8] *Андрянов Н. А., Куличенко Я. В.* Применение генеративных моделей изображений для аугментирования данных обучения детектора лиц // Нейрокомпьютеры: разработка, применение. — М.: Радиотехника, 2023. — Т. 25, № 5. — С. 7–15.

Бейесов синтез стохастической системы методом вейвлет канонических разложений

*Синицын Игорь Николаевич*¹

sinitstin@dol.ru

*Синицын Владимир Игоревич*¹

vsinitstin@ipiran.ru

*Корепанов Эдуард Рудольфович*¹

ekorepanov@ipiran.ru

Конашенкова Татьяна Дмитриевна^{1*}

tkonashenkova64@mail.ru

¹Москва, Федеральный исследовательский центр «Информатика и управление» Российской академии наук

В работе дано решение задачи синтеза оптимальной по функциональному Бейесовому критерию (ФБК) нестационарной нелинейной многомерной стохастической системы (СтС) методом вейвлет канонических разложений (ВЛКР).

Система описывается уравнениями В. С. Пугачева с аддитивными нормально распределенными случайными помехами. На вход системы поступает n_Z -мерный действительный стохастический процесс (СтП) $Z(t) = \Phi(t, U) + X(t)$, наблюдаемый в течение промежутка времени T . На выходе СтС требуется получить n_W -мерный действительный СтП $W(s) = \Psi(s, U) + Y(s)$, ($s \in S$). Здесь $\Phi(t, U) = \|\Phi_i(t, U)\|_{i=1}^{n_Z}$ и $\Psi(s, U) = \|\Psi_i(s, U)\|_{i=1}^{n_W}$ — известные, в общем случае, нелинейные функции времени $t \in T$ или $s \in S$ и n_U -мерной случайной величины (СВ) U с заданной плотностью вероятности $f(u)$. Случайные помехи $X(t)$ и $Y(s)$ имеют нулевые математические ожидания и не зависят от СВ U . Рассмотрим случай, когда случайные помехи $X(t)$ и $Y(s)$ зависимы. Ковариационная матрица $K^X(t_1, t_2) = \|K_{i_1 i_2}^X(t_1, t_2)\|_{i_1, i_2=1}^{n_Z}$ помехи $X(t)$ известна, и ее элементы принадлежат пространству $L^2(T \times T)$. Совместная ковариационная матрица $K^{XY}(t, s) = \|K_{i_1 i_2}^{XY}(t, s)\|_{i_1, i_2=1}^{n_Z, n_W}$ помех $X(t)$ и $Y(s)$ известна, и $K_{i_1 i_2}^{XY}(t, s) \in L^2(T \times S)$. Функции $\Phi_i(t, u)$, ($i = \overline{1, n_Z}$) принадлежат пространству $L^2(T)$ относительно переменной t , считая переменной u параметром.

Требуется найти оптимальную СтС с оператором A_t , выходной n_W -мерный СтП которой $W^*(s) = A_t(Z)$ обеспечивал бы минимум математического ожидания функции потерь: $M[l(W, W^*)] = \min$. ФБК задается функцией потерь $l(W, W^*)$ в виде линейного функционала F_s от функции $\sigma(s, W, W^*)$ переменной s : $l(W, W^*) = F_s(\sigma(s, W, W^*))$. Функция $\sigma(s, W, W^*)$ в общем случае может зависеть от сигнала W , его оценки W^* и их производных любых порядков. Для сокращения обозначений не указываем производные функций W и W^* в числе аргументов функции.

В [1] показано, что для решения этой задачи достаточно найти такой оператор, который обеспечивает минимум условного математического ожидания функции потерь для каждой реализации $z(t)$ СтП $Z(t)$, наблюдаемого на промежутке времени T : $M[l(W, W^*)|Z] = \min$. Эту величину часто называют условным риском $\rho(Z, W^*) = M[l(W, W^*)|Z]$. Так как линейный функционал и операцию математического ожидания можно менять местами, то для условного риска справедлива формула: $\rho(Z, W^*) = F_s(M[\sigma(s, W, W^*)|Z])$. Очевидно, что

для нахождения условного риска надо определить условную плотность вероятности $f_W(w|z)$ требуемого выходного СтП W , или СВ U и случайной помехи Y , относительно наблюдаемого СтП Z . Качество оценки W^* будем определять для каждой данной реализации СтП Z соответствующим значением условного риска $\rho(Z, W^*)$. Решение задачи состоит из 4 шагов.

Шаг 1. Построение ВЛКР векторных помех $X(t)$ и $Y(s)$. ВЛКР векторной помехи $X(t) = \|X_\theta(t)\|_{\theta=1}^{nz}$ имеет вид: $X(t) = \sum_{\nu=1}^L V_\nu^X x_\nu(t)$, $X_\theta(t) = \sum_{\nu=1}^L V_\nu^X x_{\nu\theta}(t)$, ($\theta = \overline{1, n_Z}$), где V_ν^X — независимые скалярные СВ с нулевыми математическими ожиданиями и дисперсиями D_ν^X , $x_\nu(t) = \|x_{\nu\theta}(t)\|_{\theta=1}^{nz}$ — координатные функции. ВЛКР помехи $Y(s) = \|Y_\theta(s)\|_{\theta=1}^{nw}$ построим так, чтобы оно содержало только СВ V_ν^X из ВЛКР $X(t)$: $Y(s) = \sum_{\nu=1}^L V_\nu^X y_\nu(s)$, $Y_\theta(s) = \sum_{\nu=1}^L V_\nu^X y_{\nu\theta}(s)$, $y_\nu(s) = \|y_{\nu\theta}(s)\|_{\theta=1}^{nw}$ — координатные функции. В пространстве $L^2(T)$ задается ортонормированный базис вейвлетов с компактными носителями [2] вида $\{\varphi_{00}(t), \psi_{jk}(t)\}$, где $\varphi_{00}(t) = \varphi(t)$ — масштабирующая функция, $\psi_{00}(t) = \psi(t)$ — материнский вейвлет, $\psi_{jk}(t)$ — вейвлеты уровня j для $j = \overline{1, J}$, $k = \overline{0, 2^j - 1}$, J — максимальный уровень вейвлет разрешения. Для удобства вейвлет-базис можно представить в виде: $g_1(t) = \varphi_{00}(t)$, $g_2(t) = \psi_{00}(t)$, $g_\nu(t) = \psi_{jk}(t)$, если $\nu = 2^j + k + 1$ для $j = \overline{1, J}$, $k = \overline{0, 2^j - 1}$. Тогда число базисных функций $g_\mu(t)$ равно $L = 2^{J+1}$, ($\mu = \overline{3, L}$). В пространстве $L^2(T \times T)$ задается двумерный ортонормированный вейвлет-базис: $\Phi^A(t_1, t_2) = \varphi_{00}(t_1)\varphi_{00}(t_2)$, $\Psi^H(t_1, t_2) = \varphi_{00}(t_1)\psi_{00}(t_2)$, $\Psi^B(t_1, t_2) = \psi_{00}(t_1)\varphi_{00}(t_2)$, $\Psi_{jkn}^D(t_1, t_2) = \psi_{jk}(t_1)\psi_{jn}(t_2)$, где $j = \overline{1, J}$, $k, n = \overline{0, 2^j - 1}$. Параметры ВЛКР выражаются через коэффициенты вейвлет разложения (ВЛР) функций $K_{i_1 i_2}^X(t_1, t_2)$ по ортонормированному двумерному базису вейвлетов пространства $L^2(T \times T)$. Алгоритм построения ВЛКР векторной СВ подробно изложен в [3]. Координатные функции $y_{\nu\theta}(s)$ представляют собой линейные комбинации базисных вейвлетов пространства $L^2(S)$ с коэффициентами, выраженными через коэффициенты ВЛР элементов совместной ковариационной функции $K_{i_1 i_2}^{XY}(t, s)$ по двумерному вейвлет-базису пространства $L^2(T \times S)$.

Шаг 2. Построение ВЛКР входного СтП $Z(t)$. ВЛКР СтП $Z(t) = \|Z_\theta(t)\|_{\theta=1}^{nz}$ строится в виде: $Z(t) = \sum_{\nu=1}^L V_\nu^Z x_\nu(t)$, $Z_\theta(t) = \sum_{\nu=1}^L V_\nu^Z x_{\nu\theta}(t)$, ($\theta = \overline{1, n_Z}$), где СВ $V_\nu^Z = \alpha_\nu(U) + V_\nu^X$, $\alpha_\nu(U) = \sum_{i=1}^{n_Z} \int_T a_\nu(\tau) \Phi_i(\tau, U) d\tau$. Здесь $a_1(t) = g_1(t)$, $a_\mu(t) = \sum_{\lambda=1}^{\mu-1} d_{\mu\lambda} g_\lambda(t) + g_\mu(t)$, ($\mu = \overline{2, L}$). Параметры $d_{\mu\lambda}$ определяются при построении ВЛКР $X(t)$. Функции $\Phi_i(t, u)$ относительно t , считая u параметром, имеют ВЛР: $\Phi_i(t, u) = a^{\Phi_i}(u)\varphi_{00}(t) + \sum_{j=0}^J \sum_{k=0}^{2^j-1} d_{jk}^{\Phi_i}(u)\psi_{ik}(t)$, $a^{\Phi_i}(u) = \int_T \Phi_i(\tau, u)\varphi_{00}(\tau) d\tau$, $d_{jk}^{\Phi_i}(u) = \int_T \Phi_i(\tau, u)\psi_{jk}(\tau) d\tau$. Обозначим через $a^{\Phi_i}(u) = c_1^{\Phi_i}(u)$, $d_{jk}^{\Phi_i}(u) = c_\nu^{\Phi_i}(u)$ для $j = \overline{0, J}$, $k = \overline{0, 2^j - 1}$, $\nu = 2^j + k + 1$, тогда $\alpha_1(u) = \sum_{i=1}^{n_Z} c_1^{\Phi_i}(u)$, $\alpha_\nu(u) = \sum_{i=1}^{n_Z} \left[\sum_{\lambda=1}^{\nu-1} d_{\nu\lambda} c_\lambda^{\Phi_i}(u) + c_\nu^{\Phi_i}(u) \right]$, ($\nu =$

$= \overline{2, L}$). Выразив СВ V_ν^X через СВ V_ν^Z и $\alpha_\nu(U)$, получим ВЛКР случайной помехи $Y(s) = \sum_{\nu=1}^L (V_\nu^Z - \alpha_\nu(U))y_\nu(s)$ и выходного СтП $W(s) = \Psi(s, U) + \sum_{\nu=1}^L (V_\nu^Z - \alpha_\nu(U))y_\nu(s)$.

Шаг 3. *Нахождение условной плотности $f_W(w|z)$.* Условная плотность $f_W(w|z)$ тождественно совпадает с условной плотностью СВ U относительно совокупности СВ V_ν^Z и определяется после наблюдения СтП Z на промежутке времени T [1]: $f_W(w|z) = \varkappa(z)f(u) \exp \left[\sum_{\nu=1}^L \frac{v_\nu^Z \alpha_\nu(u)}{D_\nu^X} - \frac{1}{2} \sum_{\nu=1}^L \frac{\alpha_\nu^2(u)}{D_\nu^X} \right]$,

$$\text{где } \varkappa(z) = \left[\int_{-\infty}^{+\infty} f(u) \exp \left[\sum_{\nu=1}^L \frac{v_\nu^Z \alpha_\nu(u)}{D_\nu^X} - \frac{1}{2} \sum_{\nu=1}^L \frac{\alpha_\nu^2(u)}{D_\nu^X} \right] du \right]^{-1}.$$

Шаг 4. *Определение вида оператора оптимальной системы.* Зная плотность $f_W(w|z)$, можно вычислить условный риск: $\rho(Z, W^*) = F_s \left(\int_{-\infty}^{+\infty} \sigma(s, \Psi(s, u) + \sum_{\nu=1}^L (v_\nu^Z - \alpha_\nu(u))y_\nu(s), W^*) f_W(w|z) du \right)$. Для решения задачи необходимо определить оптимальный выходной СтП $W^* = W^*(s)$ из условия минимума функционала F_s . Функционал F_s следует рассмотреть как функцию величины $P = W^*$ при фиксированных значениях параметров v_ν^z и времени s . Значение параметра $P = P_0(s, v_1^z, \dots, v_L^z)$, при котором функционал F_s достигает наименьшего значения, определяет оптимальный оператор по ФБК. Замена в $P_0(s, v_1^z, \dots, v_L^z)$ переменные v_1^z, \dots, v_L^z , соответственно, СВ V_1^z, \dots, V_L^z , получаем вид оператора ФБК оптимальной системы: $W^*(s) = A_t(Z) = P_0(s, V_1^z, \dots, V_L^z)$.

В качестве примера можно привести критерий минимума усредненного по области S среднего квадрата ошибки для функции потерь $l(W, W^*) = \int |W - W^*|^2 ds$. В этом случае $W^*(s) = M[W(s)|V_1^z, \dots, V_L^z]$, или $W^*(s) = \int_{-\infty}^{+\infty} \left(\Psi(s, u) + \sum_{\nu=1}^L (v_\nu^Z - \alpha_\nu(u))y_\nu(s) \right) f_W(w|z) du$.

Применение: Обработка сигналов в сложных системах информатики и управления в условиях нестационарных помех, в том числе ударных.

- [1] Синицын И. Н. Канонические представления случайных функций. Теория и приложения. Изд. 2-е. — М.: ГОРСУС-ПРЕСС, 2023. — 816 с.
- [2] Добеши И. Десять лекций по вейвлетам. Изд. 2-е. — Москва-Ижевск: НИЦ «Регулярная и хаотическая динамика», 2004. — 464 с.
- [3] Синицын И. Н., Сергеев И. В., Корепанов Э. Р., Конашенкова Т. Д. Инструментальное программное обеспечение анализа и синтеза стохастических систем высокой доступности (V) // Системы высокой доступности. — М.: Радиотехника, 2018. — Т. 14, № 1. — С. 59–70.

Интерпретируемость обучения в системе обработки сигналов

*Сенько Олег Валентинович*¹

senkoov@mail.ru

Чучупал Владимир Яковлевич^{1*}

v.chuchupal@mail.ru

*Кузнецова Анна Викторовна*²

azforum@yandex.ru

*Окулов Николай Васильевич*³

stereoperm@yandex.ru

*Докукин Александр Александрович*¹

dalex@ccas.ru

¹Москва, Федеральный исследовательский центр «Информатика и управление»
Российской академии наук

²Москва, Институт биохимической физики им. Н. М. Эмануэля Российской академии наук

³Пермь, ООО «Правильное Пчеловодство»

Существенным недостатком большинства методов современного машинного обучения является непрозрачность полученных с их помощью алгоритмов автоматической классификации в смысле сложности понимания пользователем их работы, а также процесса обучения [1]. Одним из возможных способов повышения прозрачности служит дополнительное использование методов интеллектуального анализа данных, позволяющих представить работу сложного многофакторного алгоритма в виде совокупности относительно простых, достоверных и интерпретируемых эффектов. Такими возможностями обладает метод Оптимальных достоверных разбиений (ОДР), позволяющий находить разнообразные нелинейные взаимосвязи с оценкой их статистической значимости [2]. Предполагается, что высокая статистическая значимость может обеспечить достоверность эффекта.

Для оценки статистической значимости выявленных закономерностей применяется комбинация технологии перестановочных тестов и принципа бритвы Оккама. Перестановочные тесты во многих случаях обладают несомненными преимуществами по сравнению с традиционными параметрическими и непараметрическими критериями и прочно вошли в современный статистический анализ. Принцип бритвы Оккама используется для оптимизации сложности моделей. В этом смысле цели применения совпадают с целями известных информационных критериев. Однако их недостатком является невозможность по отдельности оценить статистическую значимость каждого элемента модели, что осуществляется с помощью рассматриваемого варианта принципа бритвы Оккама, ранее предложенного в [3]. Основные цели рассматриваемого варианта принципа — резкое сокращение множества генерируемых моделей с сохранением только значимых эффектов, что существенно упрощает и облегчает анализ пользователем выявленных закономерностей.

Для решения задач компьютерной диагностики была разработана программная система, состоящая из пакета прикладных программ автоматической классификации и интеллектуального анализа данных Data Master Azforum и программы, генерирующей векторные признаковые описания. Пакет Data Master

содержит набор методов автоматической классификации, методы кластерного анализа, методы статистического анализа, включая метод ОДР.

Метод ОДР включает поиск оптимальных разбиений интервалов значений отдельных переменных и двумерных областей значений пар переменных по обучающей выборке стандартного вида $\tilde{S}_0 = \{(y_1, x_1), \dots, (y_m, x_m)\}$. Границы разбиений подбираются таким образом, чтобы достичь наилучшего разделения групп с различными значениями целевой переменной Y . При этом качество разделения для разбиения признакового пространства $r = q_1, \dots, q_k$ описывается функционалом

$$Q(r, \tilde{S}_0) = \frac{1}{\nu_0(1 - \nu_0)} \sum_{i=1}^k m_i(\nu_0 - \nu_i)^2,$$

где ν_0 — доля объектов с $y_j = 1$ во всей обучающей выборке \tilde{S}_0 , ν_i — доля объектов с $y_j = 1$ внутри области q_i , m_i — полное число объектов обучающей выборки внутри области q_i . Оптимальные разбиения ищутся внутри простых семейств разбиений с одной или двумя границами.

Статистическая значимость закономерностей, описываемых оптимальными разбиениями, может быть оценена с помощью перестановочного теста и датчика случайных чисел. Пусть $\tilde{f}_g = \{f(j)\}$ представляет собой набор перестановок чисел из множества $\{\overline{1, m}\}$, сгенерированных с помощью генератора случайных чисел. Вычисление p -значений тогда приближенно производится по формуле

$$p = \frac{|\{f \mid P_{pv}(\tilde{S}_p(f)) = 1, f \in \tilde{f}_g\}|}{|\tilde{f}_g|},$$

где $P_{pv}(\tilde{S})$ — предикат: $F[M_b(\tilde{S}), \tilde{S}] \geq F[M_b(\tilde{S}_0), \tilde{S}_0]$, F — функционал, характеризующий качество аппроксимации Y от X_1, \dots, X_n , $M_b(\tilde{S})$ — оптимальная модель, найденная по выборке \tilde{S} внутри множества моделей \tilde{M} . Преимуществом перестановочного теста является отсутствие априорных предположений о вероятностных распределениях и возможность использования при произвольных размерах выборок.

Рассматриваемый программный комплекс может применяться для распознавания состояния разнообразных биологических или технических систем по генерируемым ими сигналам различной природы. В частности, в задачах медицинской диагностики с помощью электрокардиограмм, электроэнцефалограмм, фотоплетизмограмм и т.д.; в задачах технической диагностики по акустическим и иным сигналам. Например, приведенные подходы использовались для решения задачи автоматической диагностики деменции по данным электроэнцефалографии [4].

Другим интересным примером практического приложения представленных методов явилась задача автоматической классификации ульев по признаку наличия или отсутствия в них пчелиной матки [5]. В иерархии многих видов соци-

альных насекомых, в частности пчел, центральное место занимает матка, которая производит на свет всех членов сообщества. При ее отсутствии сообщество быстро погибает. В случае пчел, если это произошло во второй половине лета, потеря матки ведет к гарантированному вымиранию пчелиной семьи. Поэтому важно вовремя уметь находить улья, в которых пропала матка. Для промышленных пасек с десятками и сотнями ульев — это трудоемкая операция. Существенно ее удешевить и облегчить в целом иные процедуры оценки состояния ульев может использование компьютерных методов, в частности интеллектуальный анализ звуковых сигналов, издаваемых пчелами.

Данные для анализа собирались с помощью приложения для Android на смартфонах Dexp. Всего для анализа использовались 108 звуковых профилей продолжительностью 3 мин, включая 55 профилей, соответствующих ульям без матки, и 53 профиля, соответствующих ульям с маткой. В экспериментах применялось признаковое описание в амплитудно-частотной области. Каждый m -й сегмент аудиозаписи описывался 60-мерным вектором, в котором содержатся 30 логарифмов энергий сигнала в частотных полосах и столько же их первых разностей, а вся запись представляется как последовательность из таких векторов, длина которой зависит от длительности аудиозаписи. Полученные спектральные показатели анализировались с помощью метода ОДР. Подробно методология и результаты описаны в [5].

Работа выполнялась с использованием инфраструктуры Центра коллективного пользования «Высокопроизводительные вычисления и большие данные» (ЦКП «Информатика») ФИЦ ИУ РАН (г. Москва).

- [1] *Haibe-Kains B., Adam G. A., Hosny A. et al.* Transparency and Reproducibility in Artificial Intelligence // *Nature*. — 2020. — No. 586. — E14–E16.
- [2] *Kuznetsova A. V., Kostomarova I. V., Senko O. V.* Modification of the Method of Optimal Valid Partitioning for Comparison of Patterns Related to the Occurrence of Ischemic Stroke in Two Groups of Patients // *Pattern Recognition and Image Analysis*. — Pleiades Publishing, Ltd, 2014. — Vol. 24, № 1. — P. 114–123.
- [3] *Senko O. V., Dzyba D. S., Pigarova E., Rozhinskaya L. Ya., Kuznetsova A. V.* A Method for Evaluating Validity of Piecewise-linear Models // *Intern. Conf. on Knowledge Discovery and Information Retrieval (KDIR-2014), Rome*. — 2014. — P. 437–443.
- [4] *Доровских И. В., Сенько О. В., Чучупал В. Я., Докужкин А. А., Кузнецова А. В.* Исследование возможности диагностики деменции по сигналам ЭЭГ с помощью методов машинного обучения // *Мат. биология и биоинформатика*. — 2019. — Т. 14, № 2. — С. 543–553.
- [5] *Dokukin A. A., Kuznetsova A. V., Okulov N. V., Senko O. V., Chuchupal V. Ya.* Interpretability of Learning in a Signal Processing System // *J. Computer and Systems Sciences International, Pleiades Publishing, Ltd., 2024.* (in print).

A New Method for Tackling Overlapping Speech to Improve Speaker Diarization and Target Speaker Voice Activity Detection

Popov Dmitrii¹★

dr.dmitrii2000@yandex.ru

Maysuradze Archil¹

maysuradze@cs.msu.ru

¹Moscow, Lomonosov Moscow State University

Currently, technologies for processing and analyzing voice on audio records are rapidly developing. For our further discussion, let's define two substantial tasks. Speaker diarization task is referring to determination of multiple speakers from an audio track as well as distinguishing multiple speech segments where each speaker is active. Target speaker voice activity detection task is determining the continuous speech segments on an audio track where the target speaker, whose voice sample is known beforehand, is active. A range of methods have been proposed in the scholarly literature to address these tasks [1]; nevertheless, several concomitant problems arise for the majority of them.

The first issue is the need for streaming audio processing. In that case, the algorithm does not have access to the full context of the audio recording, which makes the implementation of most sequence processing algorithms unfeasible. Another problem is the overlapping of multiple speakers' voices. In overlapping speech segments the algorithm must work correctly and define the tasks of diarization and target speaker voice activity detection in a way that allows simultaneous activities of multiple voices. It is worth noting that the overlapping speech problem is addressed in several modern data sets [2, 3]. However, the solutions proposed for this problem involve post-processing of the algorithm's output and show low effectiveness on test datasets. The third issue lies in the voice activity processing tasks, where background noise interferes with the extraction of the speakers' feature representation and can cause significant errors in such systems.

The aim of this research is to develop a method capable of effectively handling speaker diarization and target speaker voice activity detection tasks under various levels of background noise, overlapping speech situations, while ensuring the possibility of use in streaming processing conditions.

A review of the literature reveals that methods for solving the mentioned essential tasks can be divided into two groups. The first group includes algorithms that explicitly break the problem down into several sequential stages. In this subject area such an approach is commonly called cascading. The second group consists of a set of so-called "end-to-end" algorithms, which do not perform explicit splitting of the problem into component blocks and train the model to solve the final task directly.

In all prevalent algorithms related to the cascading method, the following stages are typically identified: audio segmentation, extraction of speaker embeddings, clustering, and possibly post-processing. It is essential to note that in such a sequence of stages, the overlapping speech issue is only addressed at the post-processing step,

which involves the search and accurate redefining speaker labels at a given time. In this case, the problem of overlapping voices, which is more substantively related to the audio segmentation stage, is addressed too late, and errors from previous stages may not always be corrected.

Several changes were proposed in the model architecture:

In this study, a cascading method for solving the set tasks is proposed. We suggest using new problem formulations for the existing stages, explicitly taking into account the overlap of speech segments of multiple people.

In the segmentation stage, the classical problem statement requires continuous segments of one person's speech to be identified. We propose to add the overlapping speech detection task to this step.

During the extraction of embeddings, the classical approach relies on a set of segments obtained from the previous step. We suggest using the information about overlapping speech segments and calculating the speaker's embedding only from the part of the audio signal without overlapping speech. For segments with overlapping speech, two embeddings are computed—where the given segment is concatenated with adjacent speech segments. The post-processing function for overlapping speech is transferred to the audio segmentation stage to detect such segments in advance and assign them to different speakers. Furthermore, this approach avoids the requirement of access to the full context at the post-processing stage, allowing the method to be applied in streaming mode.

In the clustering stage, we can also specially process the overlapping speech segment, assigning it to two clusters simultaneously. In this work, an agglomerative hierarchical clustering algorithm was used, calculating the distance between clusters as the average distance between the embeddings of these clusters. At the same time, the embeddings of speech segments with overlapping speech were not taken into account in this calculation. Such segments are assigned to the corresponding cluster using the nearest neighbors search method when assigning a cluster.

To address the streaming processing problem at the clustering stage, an iterative clustering algorithm is proposed, slightly degrading the quality compared to the clustering algorithm with full context available. It is also suggested to use agglomerative hierarchical clustering and, with each new speaker embedding, dynamically recalculate the centroids of the speaker clusters. At the same time, a complete re-clustering of all incoming segments is performed every several iterations.

The following was proposed in the training method:

Obtaining data in the necessary form for model training with accurate labeling can be difficult. However, datasets containing clean voice samples from different people are available, and using synthetic data is a known approach in this situation. A method for synthesizing data containing overlapping speech and background noise is proposed to address the given problems.

For clustering algorithm training, both for streaming processing and full-context processing, a synthetically generated data sample was used.

The proposed system was compared with existing solutions [4, 5] on well-known datasets [2, 3], both at the audio segmentation stage and when compared with complete diarization systems. The high quality of the proposed method is shown, with improved performance on benchmarks containing more overlapping speech, aligning with state-of-the-art solutions. In streaming mode, it outperforms current works by 3.5 percentage points in diarization error.

The research is supported by Scientific and educational school of Moscow State University “Brain, cognitive systems, artificial intelligence”, research work of Moscow State University 5.1.21.

- [1] *Dimitriadis D. et al.* A Review of Speaker Diarization: Recent Advances with Deep Learning // Computer Speech & Language. — 2021. — Vol. 72. — P. 101317.
- [2] *Carletta J.* Announcing the AMI Meeting Corpus // The ELRA Newsletter. — 2006. — P. 3–5.
- [3] *Ryant N. et al.* The Third DIHARD Diarization Challenge // Interspeech 2021. — 2020. — P. 3570–3574.
- [4] *Zhang A. et al.* Fully supervised speaker diarization // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). — 2019. — P. 6301–6305
- [5] *Bredin H. et al.* Pyannote.Audio: Neural Building Blocks for Speaker Diarization // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). — 2020. — P. 7124–7128

Анализ и распознавание техногенных шумов с использованием нейросетевого подхода

*Копылова Оксана Андреевна*¹

oksana@org.sssc.ru

¹Новосибирск, Институт вычислительной математики и математической геофизики СО РАН

Проблема геоэкологического мониторинга окружающих техногенных шумов в связи с их возрастающим воздействием на социальную среду приобретает все более высокую актуальность. Целью работы и основной задачей является разработка методов, алгоритмов и программ для анализа, обнаружения и распознавания техногенных сейсмических и акустических шумов на низких и инфранизких частотах, наиболее уязвимых для человека. В качестве источников транспортных шумов рассматриваются железнодорожный, авиационный транспорт, тяжелая гусеничная и колесная техника. В качестве промышленных источников с повышенным уровнем воздействия на окружающую среду рассматриваются мощные низкочастотные вибраторы и техногенные взрывы.

Решение задачи распознавания движущихся транспортных средств по записям сейсмических сигналов может быть построено на сравнении значений текущего вектора информативных признаков каждого объекта с эталонами, полученными статистическими методами на этапе предварительного обучения [1, 2]. При этом выбор адекватного вектора на этапе анализа данных является ключом к успеху решения задачи распознавания. В [3] распознавание объектов транспортных средств проводилось с помощью вероятностной нейронной сети. В [4] предлагалось использовать трехслойную нейронную сеть прямого распространения и метод k -ближайших соседей. В [5] использовался метод опорных векторов. В [6] распознавание основывалось на сверточной нейронной сети.

Отличительная особенность настоящей работы увязывается с комплексов вопросов обнаружения, анализа и распознавания на расстояниях, превосходящих ранее полученные в условиях воздействия внешних шумов.

Решение задачи выделения вектора информативных признаков техногенных шумов осуществляется на основе использования спектрального, спектрально-временного анализа, вейвлет-преобразования. Предложенные здесь подходы используются для решения задач повышения помехоустойчивости обнаружения и распознавания техногенных источников по их шумам.

Повышение помехоустойчивости обнаружения полезных сигналов на фоне шумов осуществляется на основе предварительной цифровой фильтрации. Выбор методов фильтрации основан на апостериорной информации об особенностях сейсмоакустических сигналов от источников с учетом критерия помехоустойчивости «сигнал/шум». Для выделения широкополосного сигнала используется фильтр Баттерворта. Для выделения узкополосных колебаний используется квадратурный алгоритм фильтрации. Обнаружение источников колебаний происходит на основе порогового алгоритма вычисления энергетических

характеристик на отфильтрованных данных с применением скользящего окна сглаживания [7].

Решение задачи распознавания основано на использовании сверточной нейронной сети с учетом спектральных характеристик техногенных колебаний [8]. Классификация происходит по четырем типам источников колебаний: электропоезд, грузовой поезд, тяжелый колесный и гусеничный транспорт. Отдельно рассматриваются вопросы распознавания в условиях меняющегося уровня внешних шумов и пространственного положения транспорта по отношению к пунктам регистрации колебаний.

Рассмотренные в работе алгоритмы прошли апробацию на реальных данных полевых экспериментов, проведенных в ИВМиМГ СО РАН при участии автора. Получены численные оценки точности решения задач. Показано, что метеоусловия оказывают существенное влияние на получаемые результаты.

Работа выполнена в рамках госзадания FWNM–2022–0004.

- [1] *Левковская Т. В., Козлов Э. В., Мурашко Н. И.* Обработка сейсмических сигналов в интеллектуальных системах пассивной локации // Информатика. — 2010. — Т. 3, № 27. — С. 89–96.
- [2] *Хайретдинов М. С., Авроров С. А.* Обнаружение и распознавание взрывных источников // Вестник НЯЦ РК. — 2012. — № 12. — С. 17–24.
- [3] *Алямкин С. А.* Классификация объектов в сейсмической системе обнаружения с учетом параметров их движения. — Автореф. дис. кан. тех. наук, Новосибирск, 2014. — 21 с.
- [4] *Khunarsal P., Lursinsap C., Raicharoen T.* Very Short Time Environmental Sound Classification Based on Spectrogram Pattern Matching // Information Sciences. — 2013. — Vol. 243. — P. 57–74.
- [5] *Zhou Q., Yao X., Wang C., Hu J., Liu P., Lin J.* Adaptive Moving Ground-Target Detection Method Based on Seismic Signal // IEEE Geoscience and Remote Sensing Letters. — IEEE, 2022. — Vol. 19. — P. 1–5.
- [6] *Wang Y., Cheng X., Li X., Li B., Yuan X.* Powerset Fusion Network for Target Classification in Unattended Ground Sensors // IEEE Sensors Journal. — IEEE, 2021. — Vol. 21, No. 12. — P. 13466–13473.
- [7] *Копылова О. А., Хайретдинов М. С.* Анализ и распознавание транспортных шумов с использованием нейросетевого подхода // Тр. Международной конференции «Знания — Онтологии — Теории» (ЗОНТ-2023). — Новосибирск: Институт математики им. С. Л. Соболева СО РАН, Новосибирский государственный университет, 2023. — С. 188–194.
- [8] *Добrorодный В. И., Копылова О. А., Хайретдинов М. С.* Обнаружение и пеленгование транспортных объектов в сейсмоакустических системах наблюдения // Вестник СибГУТИ. — 2023. — Т. 17, № 1. — С. 3–17.

Трансформерные модели BERT, взаимное сходство смыслов коротких текстов и их ранжирование по близости эталону

Михайлов Дмитрий Владимирович^{1*}

mdv74@list.ru

*Емельянов Геннадий Мартинович*¹

Gennady.Emelyanov@novsu.ru

¹Великий Новгород, Россия, НовГУ

Составление подборки публикаций по заданной теме требует не только анализа релевантности словаря каждой публикации интересующей пользователя теме, но и учета конечной цели самого пользователя (т.е., для решения каких именно задач делается подборка). При подготовке электронного учебного материала это означает поиск оптимального порядка работы с первоисточниками от более общего к более специфическому в целях формирования индивидуальной образовательной траектории обучаемого (студента). В идеальном случае получаем оценку взаимной смысловой зависимости текстов относительно наиболее рациональных (эталонных) вариантов описания представляемых ими фрагментов знаний [1]. В настоящей работе задача подобного ранжирования текстов решается на основе анализа их взаимной смысловой близости с применением семейства нейросетевых языковых моделей BERT (от англ. Bidirectional Encoder Representations from Transformers). Модели данного семейства основаны на архитектуре Transformer [2] и предварительно обучаются на больших текстовых коллекциях. С помощью указанных моделей предложения отображаются в многомерные векторы («эмбединги»). Содержательно каждый такой вектор показывает встречаемость заданного предложения в определенном контексте, также возможно их построение для любого законченного текстового фрагмента, например, слова или параграфа. При этом оценка смысловой близости (т.е. «силы» смысловой связи) анализируемых текстовых фрагментов может быть формально определена через меру близости соответствующих им векторов, например, на основе косинусного расстояния. Из известных моделей семейства BERT в решаемой нами задаче наибольший интерес представляют модели типа SciBERT [3], обучаемые на корпусах научных текстов.

Основная идея предлагаемого решения состоит в том, что «точкой входа» в формируемой траектории работы пользователя с первоисточниками будет та публикация в составе ранжируемой коллекции, которая максимально связана по смыслу с остальными работами коллекции. При этом среднеквадратическое отклонение оценки «силы» смысловой связи здесь должно быть минимальным. Как и в [1] анализируемыми фрагментами публикаций являются их аннотации вместе с заголовками как отражающие основное содержание каждой из работ и наиболее значимые результаты без излишних методологических деталей.

Параллельно с оцениванием «силы» смысловой связи публикации с остальными работами в составе коллекции ее аннотация проходит оценку на смысловую связность. Смысловая связность аннотации здесь предполагает то, что входящие в нее предложения должны быть максимально связаны друг с дру-

гом по смыслу. При этом по каждому предложению Ts_j аннотации Ts_i для отвечающего ему вектора-эмбединга вычисляется массив значений Cs_j косинусной близости аналогичным векторам остальных предложений аннотации и выбирается предложение Ts_{\max} с максимальным суммарным значением близости до остальных предложений. Назовем далее предложение Ts_{\max} центром масс анализируемой аннотации относительно смысловой связности. Сама смысловая связность аннотации Ts_i может быть формально определена как $\max(Cs_{\max}) / (1.0 + \text{std}(Cs_{\max}))$, где $\text{std}(Cs_{\max})$ — среднеквадратическое отклонение значения косинусной близости предложения Ts_{\max} остальным предложениям аннотации, $\max(Cs_{\max})$ — максимальное из значений, представленных в массиве Cs_{\max} . При этом для «силы» смысловой связи публикации с другими работами коллекции используются две не зависящие друг от друга оценки: для полных текстов аннотаций публикаций и для центров масс аннотаций. Данные оценки содержательно близки оценке смысловой связности аннотации и имеют сходные с ней расчетные формулы. В случае оценки «силы» смысловой связи относительно центров масс аннотаций в роли массива Cs_{\max} будет массив значений косинусной близости вектора центра масс анализируемой аннотации аналогичным векторам центров масс аннотаций остальных публикаций коллекции. При оценке «силы» смысловой связи относительно полных текстов аннотаций указанный массив будет состоять из значений косинусной близости эмбединга для текста анализируемой аннотации и соответствующих эмбедингов аннотаций остальных публикаций. Результирующий рейтинг публикации, который ассоциируется авторами с близостью ее аннотации эталону, определяется произведением оценки «силы» смысловой связи публикации с остальной коллекцией и оценки смысловой связности аннотации анализируемой публикации.

Программная реализация предложенного решения на Python 3.10 (включая блокнот Jupyter Notebook, исходные данные и результаты эксперимента) представлена по адресу: <https://portal.novsu.ru/file/1986783>. Были задействованы четыре наиболее известные модели трансформеров предложений, работающие с русским языком, а именно: *bert-base-nli-mean-tokens*, *sberbank-ai/ruscibert*, *sentence-transformers/distiluse-base-multilingual-cased-v1*, *sentence-transformers/all-MiniLM-L6-v2* (более подробное их описание представлено на портале <https://huggingface.co/>). Для вычисления косинусной близости эмбедингов центров масс была задействована функция *cosine_similarity* библиотеки *sklearn.metrics.pairwise* [4], а для полных текстов аннотаций — аналогичная функция *pytorch_cos_sim* из библиотеки *sentence_transformers.util* [5].

Для формирования оптимального порядка работы пользователя с публикациями уже в ранжированной коллекции для каждой из работ находится наиболее близкая ей по смыслу на основе косинусной близости соответствующих векторов-эмбедингов. При этом траектория навигации пользователя по коллекции строится «сверху вниз» от публикации с большим рейтингом к публикации с меньшим рейтингом, наиболее близкой ей по смыслу.

В качестве ранжируемых были взяты коллекции, из которых производился отбор статей в работе [1]. Отметим, что статья из коллекции «Статистическая теория обучения» сборника трудов конференции ММРО-15, получившая максимальную близость эталону в [1], получила наибольший результирующий рейтинг и в настоящей работе относительно моделей *sentence-transformers/distiluse-base-multilingual-cased-v1* (центры масс и полные тексты аннотаций) и *bert-base-nli-mean-tokens* (полные тексты аннотаций). Относительно модели *bert-base-nli-mean-tokens* (центры масс аннотаций) данная статья оказалась наиболее близкой статье с наибольшим результирующим рейтингом на основе косинусной близости соответствующих эмбедингов. Также для данной статьи статья, наиболее близкая ей по косинусной близости эмбедингов относительно моделей *sentence-transformers/distiluse-base-multilingual-cased-v1* и *sentence-transformers/all-MiniLM-L6-v2* (центры масс и полные тексты аннотаций), а также *bert-base-nli-mean-tokens* и *sberbank-ai/ruscibert* (полные тексты аннотаций), получила с ней же наибольшее значение дополняемости по смыслу в работе [1] на основе взаимосвязи смысловых эталонов. В этом плане предложенное решение дополняет разработанную в [1] методику иерархизации текстов.

Следует отметить, что при предлагаемом ранжировании публикаций предположение об отражении аннотацией основного содержания работы и ее результатов без излишних деталей может не выполняться, например, если не учитывается когнитивная сложность текста. Поэтому при существенном расхождении оценок по разным моделям трансформеров целесообразно расширить анализируемый текстовый материал, добавив к аннотациям вводные и заключительные разделы сравниваемых статей. Отдельного исследования при этом заслуживает связь указанного расхождения и близости текста смысловому эталону.

Работа поддержана грантом РФФИ № 19-01-00006.

- [1] *Mikhaylov D. V., Emelyanov G. M.* Hierarchization of topical texts based on the estimate of proximity to the semantic pattern without paraphrasing // Pattern Recognition and Image Analysis. — 2020. — Vol. 30, No. 3. — P. 440–449. <https://rdcu.be/b7fei>.
- [2] *Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser L., Polosukhin I.* Attention is All you Need // Arxiv.org. — 2017. <https://doi.org/10.48550/arXiv.1706.03762>.
- [3] *Герасименко Н. А., Чернявский А. С., Никифорова М. А., Воронцов К. В.* Трансформерная языковая модель ruSciBERT для векторизации и обработки научных текстов на русском языке // Тез. докл. 14-й Междунар. конф. ИОИ-2022. — М.: Российская академия наук, 2022. — С. 150–151.
- [4] *Pedregosa F. et al.* Scikit-learn: Machine Learning in Python // J. Machine Learning Research. — 2011. — Vol. 12. — P. 2825–2830.
- [5] *Reimers N., Gurevych I.* Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks // Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing, — 2019. <https://doi.org/10.48550/arXiv.1908.10084>.

Решение задачи кореференции с использованием токенизации по подсловам

Кудинов Илья Дмитриевич¹*

ilja@kdsli.ru

¹Москва, Институт проблем управления имени В. А. Трапезникова РАН

Стандартом решения многих задач обработки текстов, связанных с классификацией слов в тексте, является использование языковых моделей, проводящих токенизацию входного текста по словам. Это следует из того что интуитивно проще классифицировать слово, представленное только одним токеном.

Ввиду большого числа слов в используемом словаре, такие языковые модели требуют значительных затрат памяти и вычислительных ресурсов на работу. Для языков с богатой морфологией модели должны хранить информацию о каждой возможной словоформе каждого слова, что увеличивает размер словаря в среднем в два десятка раз. В качестве альтернативного подхода используется токенизация текста по наборам подряд стоящих символов, называемых подсловами (subword, word pieces). Это позволяет модели оперировать словарем ограниченных размеров [1]. Однако, при использовании такой стратегии токенизации необходимы дополнительные механизмы объединения векторных представлений нескольких токенов, представляющих одно слово [2].

Задача кореференции — это задача обработки естественных языков. В заданном тексте устанавливаются именные группы, ссылающиеся на один объект [3]. Предполагается, что задача может быть решена более точно при использовании токенизации по подсловам. Задача усложняется необходимостью классифицировать не слова, а именные группы — группы подряд стоящих слов [4].

В настоящей работе предлагается способ решения задачи кореференции при использовании токенизации по подсловам путем вычисления для каждой пары токенов двух оценок. Первая оценка выражает вероятность того, что два токена входят в одну именную группу. Вторая оценки выражает вероятность того, что пара токенов входят в две разные кореферентные именные группы. Совмещение двух оценок позволяет получить модель решение задачи кореференции, наследующей все преимущества моделей с токенизацией по подсловам: меньший размер модели, более точной работой с языками с богатой морфологией.

- [1] *Mielke S. J., et al.* Between Words and Characters: A Brief History of Open-Vocabulary Modeling and Tokenization in NLP // arXiv. — 2021. <https://arxiv.org/abs/2112.10508>
- [2] *Ács J., Kádár Á., Kornai A.* Subword Pooling Makes a Difference // arXiv. — 2021. <https://arxiv.org/abs/2102.10864>
- [3] *Liu R., Mao R., Luu A. T., Cambria E.* A Brief Survey on Recent Advances in Coreference Resolution // Artificial Intelligence Review. — 2023. — Vol. 56, No. 12. — P. 1–43.
- [4] *Joshi M., Levy O., Weld D. S., Zettlemoyer L.* BERT for Coreference Resolution: Baselines and Analysis // arXiv. — 2019. <https://arxiv.org/abs/1908.09091>

Анализ информационной согласованности области «Объяснимый искусственный интеллект»

Садиев Сабрина Аймановна^{1*}

sad.sabrina.d@yandex.ru

*Рогов Александр Александрович*¹

rogov@petrsu.ru

¹Петрозаводск, Петрозаводский государственный университет

Область объяснимого искусственного интеллекта (explainable AI, XAI) является относительно новой в исследованиях, связанных с применением искусственного интеллекта (ИИ). Новизна области порождает отсутствие общепринятой терминологии и расширяет используемые в работах ключевые слова. Это приводит к проблемам при отслеживании и анализе литературы [1], а также к различным подходам к интерпретированию и объяснению алгоритмов. Таким образом, хотя частично решен вопрос существования методов интерпретации и сформирована отдельная область исследований, не существует общего подхода, который позволял бы интерпретацию алгоритма сделать задачей, осуществляемой каждым исследователем с наименьшим возможным затраченным временем. В работе приводятся результаты анализа 6471 статьи, собранной по ключу «explainable AI» с ресурсов arxiv.org и pubmed, и предлагается возможное упрощение внедрения шага в объяснении модели в рабочий процесс.

Существует классификация области XAI верхнего уровня, делящая алгоритмы на объяснимые (explainable) и интерпретируемые (interpretable). С точки зрения смыслового различия, интерпретируемость моделей преследует цель создания или использования изначально понятной модели, тогда как объяснимость преследует цель рассмотрения скрытой взаимосвязи между входными и выходными данными у модели, изначально не являющейся объяснимой. В дополнение, имеются разные более глубокие классификация методов и алгоритмов. Они связаны с разнообразием архитектур моделей и разнообразием модальностей данных, к которым на данный момент применяется искусственный интеллект. Например, можно выделить классификации по (i) глубине интерпретации относительно архитектуры модели, (ii) количеству обучающих примеров, охватываемых методом, (iii) типу или типам алгоритмов, к которым метод может быть применен и др. Кроме того, можно встретить название области также со словами «governance», «ethics», «responsibility», «trustworthiness» и «fairness» [2]. При помощи модели Word2Vec, примененной к ранее описанным собранным статьям, также были обнаружены слова «reliable», «transparency» и «causability». Вариативность терминологии порождает различные разные классификации одних и тех же частных методов интерпретации и различные подходы к проектированию «прозрачных» решений от источника к источнику.

На основе обнаруженной вариативности была выдвинута гипотеза о том, что при поиске релевантных исследований использования основных для области ключевых слов (interpretable и explainable) может быть недостаточно. В первых, из 4706 статей, собранных с ресурса pubmed, 1262 (27 %) не были разме-

чены ключевыми словами. Чтобы проверить правильность поиска по ключевым словам, 176 статей с ресурсов arXiv (50 статей) и PubMed (126 статей) были проанализированы вручную и на основе логического классификатора. Логический классификатор был построен следующим образом: присваивал статье флаг 1, если она содержит в аннотации одно из 9 ключевых слов («explainable ai», «explainable artificial intelligence», «xai», «explainable deep learning», «explainable machine learning», «interpretable artificial intelligence», «interpretable machine learning», «interpretable deep learning») и 0 — иначе. Аналогичная разметка проводилась вручную. Далее было проведено сравнение количества релевантных статей, отобранных вручную и классификатором. В ходе ручной проверки с ресурса arXiv.org было отсеяно 3 статьи, а с ресурса PubMed — 62. Классификатором было предложено отсеять 30 и 82, соответственно.

Таким образом, было установлено, что обилие терминологии усложняет задачу приведения исследователя к наиболее релевантным методам. Чтобы достичь цели упрощения выбора метода интерпретации или объяснения, предлагается классификация методов при помощи интерактивной таблицы на основе запросов исследователя. Для этого в ходе исследования была выделена таксономия, которая позволяет разделить методы между собой и выбрать хотя бы начальный метод интерпретации или объяснения модели. В таксономии были использованы 3 разделяющие, и 1 подразделяющая категории. Из разделяющих были выделены (i) Microscope AI, (ii) Model understating, (iii) Model debugging [3, 4]. В качестве подразделяющей был выбран вид выходного значения метода: коэффициенты или визуальный выход. При создании таксономии также была учтена практическая доступность методов, на основании предшествующей работы [5]. Полученный результат показал, что задача приведения исследователя к конкретному методу может быть сведена к выбору по интерактивной таблице, что может ускорить процесс практического выбора метода. Демо-версия интерактивной таблицы доступна по ссылке [6].

Исследование поддержано грантом УМНИК-22 в республике Карелия от «Фонда содействия инновациям».

- [1] Rudin C., Chen C., Chen Z., Huang H., Semenova L., Zhong C. Interpretable Machine Learning: Fundamental Principles and 10 Grand Challenges // *Statistics Surveys*. — 2022. — No. 16. — P. 18–21.
- [2] Alsaigh R., Mehmood R., Katib I. AI Explainability and Governance in Smart Energy Systems: A review // *In Frontiers in Energy Research*. — 2023. — Vol. 11.
- [3] Rauker T., Ho A., Casper S., Hadfield-Menell D. Toward Transparent AI: A Survey on Interpreting the Inner Structures of Deep Neural Networks // *Proc. 2023 IEEE Conf. Secure and Trustworthy Machine Learning, SaTML 2023*. — IEEE, 2023. — P. 464–483.
- [4] Choo J., Liu S. Visual Analytics for Explainable Deep Learning // *IEEE Computer Graphics and Applications*. — IEEE, 2018. — Vol. 38, No. 4. — P. 84–92.
- [5] Caduex C. A. Анализ возможности обоснования работы алгоритмов искусственного интеллекта // *Процессы управления и устойчивость. Труды 54-й международной*

научной конференции аспирантов и студентов. — СПб.: Publishing House Fedorova G.V., 2023. — Т. 10, № 1. — С. 331–339.

- [6] *Садиев С. А.* Automatic Selection of the Explanation Method for Your Task // streamlit.app, 2023. <https://sadsabrina-streamlit-xai-table-app-5uq64s.streamlit.app> (дата обращения: 10.11.2023).

Генерация и поиск искусственно сгенерированных текстовых фрагментов в домене научных работ

Грицай Герман Михайлович^{1,2}

gritsai@ap-team.ru

Грабовой Андрей Валериевич^{1,2}★

grabovoy@ap-team.ru

*Кильдяков Александр Сергеевич*²

kildyakov@ap-team.ru

Чехович Юрий Викторович^{2,3}

chehovich@ap-team.ru

¹Москва, Московский физико-технический институт

²Москва, Антиплагиат

³Москва, Федеральный исследовательский центр «Информатика и управление» РАН

В докладе рассматривается задача поиска искусственно сгенерированных фрагментов в научных документах. Актуальность задачи обосновывается распространением моделей генерации текстовых последовательностей в открытом доступе, способных в том числе генерировать ответы на вопросы из научного домена. Для решения задачи поиска искусственных фрагментов используются подходы, которые основаны на применении моделей глубокого обучения, порожденных архитектурой трансформер. Улучшение методов генерации текстов одновременно влечет за собой необходимое улучшение методов выявления искусственных фрагментов. В работе совместно со способами детектирования предлагается алгоритм генерации научных публикаций с искусственными фрагментами при помощи языковых моделей [1].

Задача поиска сгенерированных фрагментов рассматривается как задача классификации текстовых последовательностей, где научный документ разбивается на непересекающиеся фрагменты, которые полностью покрывают исходный документ. Затем над каждым фрагментом проводится бинарная классификация $\{0, 1\}$, где 0 — текст написан человеком, 1 — текст сгенерирован.

Для повышения качества работы алгоритма поиска в докладе рассматривается проблема множественных сравнений, которая возникает из-за наличия этапа фрагментации в схеме алгоритма. С увеличением количества проверяемых фрагментов повышается риск ложноположительных срабатываний. Предложено использовать методы, которые позволяют взять контроль над уровнем значимости и достичь высокого уровня мощности, не накапливая ошибку отклонения верной гипотезы. Помимо контроля уровня значимости приняты во внимание особенности решения задачи детекции сгенерированных фрагментов, которые были выявлены ранее: зависимость качества детектирования фрагментов от длины входной последовательности, применение аугментаций и очистки данных от неинформативных словосочетаний, использование различных тематик документов на этапе дообучения [2, 3].

В вычислительном эксперименте проводится анализ качества работы алгоритма поиска сгенерированных фрагментов для русского и английского языков. Для получения качественной модели классификации текстовых фрагментов в работе проведено несколько экспериментов с разными подходами дообуче-

ния моделей: порождение ручных признаков, мультиязычное дообучение, переводы и перефразирование текстов. Наборы данных с различными тематиками были выбраны из открытого доступа, в которых средняя длина текстов приблизительно равна длине потенциального фрагмента из описанного алгоритма.

Таблица 1. Сводная таблица результатов вычислительного эксперимента

| Язык | Эксперимент | F1-score | Precision | Recall |
|------|----------------------------|--------------|--------------|--------------|
| ru | ручные признаки | 0,960 | 0,962 | 0,959 |
| | мультиязычное обучение | 0,964 | 0,964 | 0,966 |
| | перевод текстов 25 % | 0,958 | 0,961 | 0,958 |
| | перевод текстов 50 % | 0,966 | 0,968 | 0,966 |
| | парафраз предложений 100 % | 0,968 | 0,970 | 0,968 |
| | парафбраз предложений 50 % | 0,964 | 0,965 | 0,963 |
| en | ручные признаки | 0,801 | 0,856 | 0,807 |
| | мультиязычное обучение | 0,823 | 0,867 | 0,828 |
| | перевод текстов 25 % | 0,821 | 0,865 | 0,826 |
| | перевод текстов 50 % | 0,825 | 0,868 | 0,830 |
| | парафраз предложений 100 % | 0,822 | 0,866 | 0,827 |
| | парафраз предложений 50 % | 0,816 | 0,862 | 0,817 |

Результаты исследования показали, что переведенные и перефразированные тексты позволяют повысить качество и обобщающую способность у моделей детекции машинно-сгенерированных фрагментов. Показан подход к детекции в домене научных работ, а именно: разбиение исходного документа по разделам, фрагментация наиболее значимых и последующая их подача на вход дообученному классификатору.

- [1] *Грицай Г. М., Грабовой А. В., Кильдяков А. С., Чехович Ю. В.* Поиск искусственно сгенерированных текстовых фрагментов в научных документах // Доклады Российской Академии Наук. Математика, информатика, процессы управления. — 2023. — Т. 541. — С. 1–10 (в печати).
- [2] *Gritsay G., Grabovoy A., Chekhovich Y.* Automatic Detection of Machine Generated Texts: Need More Tokens // 2022 Ivannikov Memorial Workshop (IVMEM). — IEEE, 2022. — P. 20–26.
- [3] *Gritsay G., Grabovoy A., Kildyakov A., Chekhovich Y.* Automated Text Identification: Multilingual Transformer-Based Models Approach // Proceedings of the Iberian Languages Evaluation Forum, 2023.

Анализ диалогов для определения эмоций персонажей в художественных текстах

Гаврилов Максим Сергеевич¹

cobraj@yadnex.ru

¹Москва, Институт проблем управления им. В. А. Трапезникова

Определение эмоций персонажей в тексте может применяться для решения задачи автоматического анализа текста, т. е. для определения основного сюжета произведения, его персонажей и взаимодействий между ними [1, 2]. Потребность в автоматическом анализе текста есть в любой сфере деятельности, где человеку необходимо знакомиться с большим количеством художественных текстов, а также он может применяться для анализа тенденций в литературе различных эпох. К тому же, такая технология позволит автоматизировать процесс разметки текстов для обучения искусственных нейронных сетей.

Цель работы — получить модель, способную определять эмотивную тональность речи персонажей в диалоге на основе модели эмоций Келлермана-Плутчика [3]. Для достижения поставленной цели были поставлены и решены следующие задачи:

- создать алгоритм извлечения диалога из текста, способный устанавливать соответствия персонаж — реплика;
- для этого найти решение задачи разрешения кореференции в диалоге;
- создать анализатор эмотивной тональности по Келлерману-Плутчику;
- для этого сформировать и разметить набор данных текстов или предложений и присущей им эмотивной тональности по Келлерману-Плутчику, либо найти существующий датасет такого рода;
- сравнить различные методы машинного обучения применительно к задаче классификации текстов и предложений по их эмотивной тональности;
- собрать из реализованных в прошлых пунктах частей модель, способную отслеживать эмотивную тональность реплик определенных персонажей в диалоге.

Для извлечения диалога из текста использовался алгоритм на основе синтаксического парсера проекта *natasha*, применяемый к токенизированному тексту. Результатом работы алгоритма должен быть список референтов и список высказываний. Между референтами и высказываниями должна быть установлена связь «один ко многим». Каждый референт представляет собой персонажа, произносившего реплику в тексте [4].

Для выбора архитектуры детектора эмоций производилось сравнение следующих методов машинного обучения:

- наивный байесовский классификатор;
- наивный словарный классификатор;
- классификатор на основе модели DeepPavlov RuBERT-base-cased [5, 6];

- классификатор на основе модели DeepPavlov RuBERT-base-cased, дополнительно предобученной на датасете cedr;
- классификатор на основе модели cointegrated rubert-tiny 2, дополнительно предобученной на датасете cedr.

Сравнение этих алгоритмов выполнялось с использованием набора данных из 900 размеченных в соответствии с моделью Келлермана-Плутчика предложений. По результатам сравнения было принято решение использовать классификатор на основе модели DeepPavlov RuBERT, предобученной на наборе данных cedr.

В ходе выполнения данной работы была создана модель на основе нейронных сетей глубокого обучения, способная определять эмоции персонажей в художественном тексте посредством анализа диалогов. Для этого было проведено сравнение различных алгоритмов машинного обучения применительно к задаче определения эмоций в тексте и выбран алгоритм, показавший наилучшие результаты, которым оказался ансамбль классификаторов на основе модели BERT. Также был создан алгоритм извлечения диалогов из текста.

Можно сделать вывод, что при дальнейшей разработке данной темы необходимо сделать фокус на сборе и разметке данных для обучения. Также можно рассмотреть переход на более полную модель эмоций.

- [1] *Куратов Ю. М.* Специализация языковых моделей для применения к задачам обработки естественного языка. Дисс. канд. наук. — 2020.
- [2] *Acheampong F. A. Wenyu C. Nunoo-Mensah H.* Text-Based Emotion Detection: Advances, Challenges, and Opportunities // *Engineering Reports*. — 2020. — Vol. 2, No. 7. — P. e12189.
- [3] *Mahalakshmi G.* Emotion Models: A Review // *Int. J. Control Theory and Applications*. — 2017. — Vol. 10, No. 8. — P. 651–657.
- [4] *Бодрова А. А.* Разрешение кореференции методом кластеризации // *Магистерская диссертация, СБГУ*. — 2016.
- [5] *Poria S.* Emotion Recognition in Conversation: Research Challenges, Datasets, and Recent Advances // *IEEE Access*. — IEEE, 2019. — Vol. 7. — P. 100943–100953.
- [6] *Graterol W. O.* Emotion Detection for Social Robots Based on NLP Transformers and an Emotion Ontology // *Sensors*. — 2021. — Vol. 21, No. 4. — P. 1322.

Structure Extractor: Multilingual Extraction of Sections from Scientific Documents

Kopanichuk Iliia^{1,2*}

kopanichuk@ap-team.ru

*Chashchin Artem*¹

chashchin@ap-team.ru

*Ochneva Inga*¹

ochneva@ap-team.ru

*Ogaltsov Aleksandr*¹

ogaltsov@ap-team.ru

Grabovoy Andrey^{1,2}

grabovoy@ap-team.ru

*Kildyakov Aleksandr*¹

kildyakov@ap-team.ru

*Chekhovich Yury*¹

chekhovich@ap-team.ru

¹Moscow, Antiplagiat

²Moscow, Moscow Institute of Physics and Technology

In this paper we present a method to extract the structure of the scientific documents. A scientific article usually has a good structure. The structure helps to guide both readers and journal editors. It also allows differentiated assessment of text reuse occurring in the different sections of the article. Considering the wide use of plagiarism detectors in scientific practice, the task of automatic structure extraction from scientific articles becomes relevant in the plagiarism detection process. Most of the published articles and theses consist of consequent sections: title, contents, introduction, methods, results and discussion, conclusions, bibliography, and appendixes [1].

Our solution processes formatted documents, extracts the text layer and the layout from them. To identify section borders we use histogram-based gradient boosting trees. The algorithm detects title, contents, bibliography section [2], appendix and parts of the IMRAD structure: introduction, methods, and conclusions.

In order to solve the classification problem we obtain tokens and layout of each page of the document by DevExpress PDF API. The tokens and their text are sorted according to their relevance to the lines of the document. Geometric parameters and text content of document lines are converted into line features. Page features are generated for each page of the document.

In this study we collect and label 3 datasets. IMRAD Dataset contains 3132 documents in 36 languages with marked up IMRAD parts and headers. DocStructure

Table 1. Final metrics on test data for each section.

| Section | Precision | Recall | F1 |
|---------------------|-----------|--------|------|
| Title | 0.98 | 0.87 | 0.92 |
| Contents | 0.99 | 0.74 | 0.85 |
| Introduction | 0.79 | 0.72 | 0.75 |
| Methods | 0.77 | 0.49 | 0.60 |
| Conclusions | 0.88 | 0.85 | 0.86 |
| Bibliography | 0.88 | 0.95 | 0.91 |
| Appendix | 0.92 | 0.93 | 0.93 |

Dataset contains 840 documents in 36 languages with marked up title, contents, bibliography, appendix and headers. DocFullStructure Dataset consists of 70 documents in 30 languages with all sections marked up.

We train HistGradientBoosting [3] model on the page features from DocStructure dataset to predict if a page belongs to the title, contents or other sections. Another HistGradientBoosting model is trained on line features from IMRAD Dataset to predict if a document line is a candidate for any section header. Then we classify the candidates into headers based on the presence of the section keywords in them. The text after the section header is classified as the section until we reach the next section header or the limit of unmarked candidates is exceeded.

We tune both models using hyperopt package [4] with the negative F-measure as the optimization target. Due to a huge imbalance in distribution of candidates we remove all the pages without any candidates from IMRAD train dataset.

By using histogram based gradient boosting and unsupervised regex-based approach we have made a model that extracts sections from well-structured scientific documents. The solution achieves 0.89 average precision and 0.79 average recall per section on the dataset, see table 1. The solution is multilingual and can be scaled to support more languages by an unsupervised approach.

This research is funded by Antiplagiati Company.

- [1] *Sollaci L., et al.* The Introduction, Methods, Results, and Discussion (IMRAD) Structure: A Fifty-Year Survey // J. Medical Library Association. — Medical Library Association, 2004. — Vol. 92, No. 3. — P. 364.
- [2] *Ogaltsov A* A Language-Free Regular Expression Search of Document's References // Recent Trends in Analysis of Images, Social Networks and Texts. — Springer International Publishing, 2022. — Vol. 92, No. 3. — P. 45–54.
- [3] *Guolin K., et al.* LightGBM: A Highly Efficient Gradient Boosting Decision Tree // Advances in Neural Information Processing Systems. — Curran Associates, Inc., 2017. — Vol. 30. — P. 364.
- [4] *Bergstra J., et al.* Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures // Proc. 30th Int. Conf. Machine Learning. — PMLR, 2013. — Vol. 28, No. 1. — P. 115–123.

Разработка методики для автоматического выделения структуры в текстах научных публикаций с использованием методов машинного обучения и эвристических алгоритмов

*Лемтюжникова Дарья Владимировна*¹

darabtb@gmail.com

*Блашкун Михаил Романович*¹

mblashkun@gmail.com

*Латипов Айзат Рафаилович*¹

latipov257@gmail.com

Кирьянов Павел Александрович^{1*}

lovecoldreams@gmail.com

¹Москва, Институт проблем управления им. В. А. Трапезникова РАН

Задача автоматического выделения структуры научных публикаций возникает при необходимости систематизировать и нормализовать накопленные данные с разными целями: формирование базы данных публикаций с возможностью поиска по ней, построение графов цитирования по библиографическим ссылкам, использование размеченных данных в качестве данных для обучения языковых моделей.

Одной из главных проблем для решения этой задачи является большая разрозненность в структуре самих публикаций. Это может быть и разная последовательность структурных элементов, к которым относятся идентификаторы, заголовок статьи, авторы, аннотация и т. д., и отсутствие каких-либо из них, и разный формат написания в рамках одного структурного элемента.

Очевидным решением кажется использование эвристических алгоритмов, например, регулярных выражений, основанных на предположениях о расположении и формате нужных данных в тексте. Однако, для обеспечения высокой эффективности таких методов требуется разработка множества правил, учитывающих все возможные особенности каждого типа структурных элементов.

Другие доступные методы автоматического выделения структуры в большинстве случаев ориентированы исключительно на использование технологий машинного обучения. Важно отметить, что эти подходы не всегда обеспечивают высокую точность результата, которая также может зависеть от языка, на котором написана соответствующая публикация [1, 2].

В [3] авторы разработали подход по извлечению метаданных из заголовков документов с кириллическими символами. Подход включает в себя: создание датасета (Набор данных CORE Dataset предоставляет данные научных публикаций, включающие метаданные и полные тексты в машинно-обрабатываемом формате, <https://core.ac.uk/services/dataset>), извлечение текста из PDF с помощью утилиты pdfMiner с последующей токенизацией и обучение моделей grobid [4] и BiLSTM для сравнения результатов. Последняя модель показала лучшие результаты. Датасет состоит из 15 553 документов, полученных после фильтрации всех кириллических исходных данных по языку, удаления дубликатов, отбрасывания тех документов, которые не являются научными публикациями.

Рассматриваемая работа демонстрирует эффективность подхода машинного обучения, однако на практике остается недостаточной для полной автоматизации процесса, в том числе в силу ограниченности исходных данных для обучения модели.

В нашей работе решалась задача автоматического выделения структуры научных публикаций с помощью машинного обучения и эвристических алгоритмов. Совмещение двух этих методов позволяет избавиться от недостатков друг друга. Эмпирические алгоритмы, основанные на регулярных выражениях, используются для выделения таких структурных элементов, которые постоянны в публикациях. Машинное обучение — для тех структурных элементов, что могут различаться по своему объему, содержанию, расположению в самом тексте публикации. Изначально была выбрана модель BiLSTM, но в результате тестирования и сравнения с более простой LSTM было выявлено, что первая модель не приносит значительного прироста в точности, при этом требуя значительно больше ресурсов и времени вычислений (время обучения и работы BiLSTM превышает время обучения и работы LSTM примерно на 10–15 %). Поэтому в качестве основы была выбрана модель LSTM.

Набор подготовленных данных был разделен на обучающую и валидационную выборки в соотношении 99 : 1. Для обучения был использован алгоритм оптимизации ADAM (Adaptive Moment Estimation) с параметром `learning_rate` — $1e-3$. В качестве функции потерь была выбрана функция NLL loss (The negative log likelihood loss). Обучение проводилось в течение 10 эпох.

В качестве данных для обучения был создан набор размеченных текстовых данных. Каждая единица данных представляет собой файл формата `txt`, содержащий первые 1000 слов статьи, а также файл формата `json` с соответствующей разметкой на основе данных из файла `txt`. Всего получилось 628 пар файлов `txt` и `json`. Разметка включает следующие структурные элементы: код DOI, код УДК, заголовок, ФИО авторов, адрес электронной почты авторов, научная степень авторов, должность в организации, адрес, аффилиации, ключевые слова, аннотация.

Некоторые структурные элементы, такие как код DOI, код УДК, адрес электронной почты авторов, научная степень авторов, ключевые слова, определялись в автоматическом режиме с помощью регулярных выражений. В ходе обучения за элементарную единицу данных было взято слово. На этом уровне были проведены операции преобразования слов в нижний регистр, лемматизация (приведение слов к нормальной форме), удаление стоп-слов (слова, знаки, символы, которые самостоятельно не несут никакой смысловой нагрузки), подготовка общего словаря для всех документов, преобразование слов в вектора (с помощью фреймворка `pytorch`), с которыми умеет работать нейронная сеть.

В ходе подготовки данных для обучения из них были удалены предложения, которые удовлетворяют условию: предложение содержит в себе только те слова, которые не имеют меток структурных элементов, а также перед ним и после

него идут такие же предложения. С помощью этого удалось решить проблему разреженности данных. В нашей работе под разреженностью данных понимаются те слова, которые не попали ни в одну из категорий структурных элементов.

Показатель F1 рассчитывается как гармоническое среднее между точностью и полнотой и используется для оценки качества бинарной классификации. Для валидационного датасета он составляет от 0,97 до 1, что указывает на то, что модель обеспечивает как высокую точность, так и высокую полноту в определении отношения слова к тому или иному структурному элементу.

С учетом выявленных потребностей в доработке, мы планируем активно продолжать обучение моделей, сосредотачиваясь, в первую очередь, на тех, которые продемонстрировали менее удовлетворительные результаты. При этом мы также рассматриваем возможные улучшения в методологии, направленные на повышение эффективности обучения и достижение более высоких показателей точности. В контексте общей значимости исследования следует подчеркнуть его вклад в область обработки текста и структурирования научных статей на русском языке. Наша работа направлена на решение актуальных проблем в этой области, в том числе повышение точности и эффективности инструментов автоматической обработки текста на русском языке, что может значительно облегчить работу исследователей и обеспечить более точные результаты в анализе научных материалов.

- [1] *Gomes Júnior A. d. A., Schramm V. B.* Problem Structuring Methods: A Review of Advances Over the Last Decade // *Syst Pract Action Res.* — 2022. — No. 35. — P. 55–88.
- [2] *Tkaczyk D., Szostek P., Fedoryszak M. et al.* CERMINE: Automatic Extraction of Structured Metadata from Scientific Literature // *Int. J. Document Analysis and Recognition (IJ DAR).* — 2015. — Vol. 18, No. 4. — P. 317–335.
- [3] *Krause J., Shapiro I., Saier T., Farbe M.* Bootstrapping Multilingual Metadata Extraction: A Showcase in Cyrillic // *Proc. Second Workshop on Scholarly Document Processing.* — 2021. — P. 66–72.
- [4] GROBID [Электронный ресурс]. <https://github.com/kermitt2/grobid> (дата обращения: 22.11.2023).

Структурно-стохастические вычислительные модели контроля технического состояния космических средств

*Копкин Евгений Вениаминович*¹

vka@mil.ru

Иваню Анна Юрьевна^{1*}

vka@mil.ru

¹Санкт-Петербург, Военно-космическая академия им. А. Ф. Можайского

В современном состоянии развития космической техники, которая в настоящее время становится все более сложной и постоянно расширяющей сферу своего применения, особую актуальность имеют вопросы оперативного контроля и диагностирования технического состояния (ТС) космических средств (КСр), поскольку это влияет на качественное выполнение возложенных на КСр функциональных задач. Особо необходимо отметить важность оценивания ТС (говоря языком теории распознавания образов, классификации образов как видов ТС) в условиях наличия недостоверной измерительной информации при формировании признакового пространства. А это значит, что одним из главных условий применения вычислительных моделей (ВМ) в составе процедур оценивания ТС является недостаточная полнота описания априорных сведений о распознаваемых объектах. Задача построения ВМ распознавания в таких условиях может быть решена известными из теории статистическими методами. Однако, в рамках статистического распознавания ТС в силу неполноты априорной информации о распознаваемых объектах и слабой формализуемости существующих сведений довольно затруднительно получать необходимые данные для формирования различного рода статистических величин, описывающих вероятностные зависимости (статистики) между признаками распознаваемых объектов и классами, к которым эти объекты относятся.

Другим важным условием, при котором приходится создавать и использовать ВМ, является наличие самого разного рода возможных искажений входных данных — как значений признаков для распознавания [1]. Это накладывает дополнительные сложности на процесс распознавания ТС, также связанные с учетом названных выше вероятностных зависимостей.

Все это является основанием для разработки специализированных ВМ, позволяющих наиболее полно учитывать неопределенность априорных сведений и наличие искажений в процессе распознавания ТС. При реализации таких процессов используется эвристический и статистический подходы. Это позволяет самым эффективным образом использовать не только всю имеющуюся априорную информацию, но и организовать рекуррентную процедуру дообучения модели при обработке дополнительной обучающей информации, содержащей в себе искажения значений признаков параметров.

Для описания сложно формализуемых ВМ и процесса получения оценок ТС, возможно использовать так называемые G-модели, которые были разработаны в рамках концептуального моделирования и программирования в ограничениях.

Синтезируемые на основе G-модели программные комплексы позволяют оценивать состояния сложных технических объектов с применением методов теории распознавания образов. В рамках традиционной прикладной интерпретации задачи оценивания ТС КСр распознаваемый объект представляется в виде набора значений его телеметрируемых параметров (ТМП), а ТС — это класс, к которому необходимо отнести распознаваемый объект.

Для решения этой задачи предлагается использовать комбинированные ВМ [2], включающие как структурную, так и стохастическую составляющие. Синтез перечисленных выше методов обеспечит наибольший эффект от применения процедуры распознавания, т. к. для исследуемой предметной области необходимым является учет неопределенности исходной телеметрической информации (неоднородность, избыточность, разнообразие смыслового содержания), который реализуется в рамках стохастических моделей.

Основой для предлагаемых моделей служат структурные ВМ, которые учитывают иерархичность как в рамках построения самого контролируемого объекта, так и по отношению к положению контролируемого объекта относительно отдельных элементов структуры одного или нескольких уровней иерархии. Каждая структурная ВМ описывается некоторым счетным множеством формальных грамматик и языков.

Структурно-стохастическая ВМ является обобщением структурной ВМ за счет замены в ней формальных грамматик на стохастические грамматики. Таким образом, каждый стохастический язык в структурно-стохастической модели представляет собой такое множество слов в алфавите, которые соответствуют единственному (в случае наличия фактор-отображения) значению выходного параметра со своей стохастической мерой (показателем достоверности).

Показатель достоверности представляет собой условную вероятность принятия в качестве оценки вычисления (распознавания) выходного параметра x_j значения $\bar{x}_j^k \in D_{x_j}$ из множества возможных значений выходного параметра $D_{x_j} = \{x_1, x_2, \dots, x_k, \dots\}$, $k = \overline{1, |D_{x_j}|}$, по структурно-стохастической ВМ при условии, что терминальной сентенцией, сформированной в алфавите значений входных параметров этой ВМ из стохастического языка является α^S [3]:

$$P(x_j = \bar{x}_j^k | \alpha^S) = d_j^k.$$

Таким образом, реализация структурно-стохастической ВМ заключается в вычислении показателя достоверности по входному слову α^S и выборе в качестве оценки значения выходного параметра x_j такой величины, которая соответствует максимальному значению показателя достоверности d_j^k .

Использование структурно-стохастических ВМ распознавания ТС КСр позволит совместить достоинства структурных моделей по возможностям априорного описания предметной области для оценивания ТС КСр и их экономичности к ресурсам с достоинствами стохастических моделей, позволяющих наиболее полно учитывать влияние возможных искажений значений ТМП.

Кроме того, предлагается организовать многошаговый процесс обучения ВМ, при котором на каждом шаге производится дообучение на основе новой порции обучающих данных. Это позволит самым эффективным образом использовать не только всю имеющуюся априорную информацию, но и организовать рекуррентную процедуру дообучения — при обработке дополнительной обучающей информации, содержащей в себе искажения значений ТМП, что важно при оценивании ТС в реальном масштабе времени.

Реализация предложенной технологии оценивания ТС в составе автоматизированной системы информационного обеспечения в контуре автоматизированной системы управления КСр позволила обеспечить достижение необходимого уровня качества процессов оперативного контроля и диагностирования ТС КСр. При этом основными преимуществами использования структурно-стохастических ВМ при распознавании ТС являются полнота априорного описания взаимосвязей ТМП (учитывается семантика этих взаимосвязей), достоверность (учет возможных искажений значений ТМП), оперативность и точность получаемых результатов.

- [1] *Охтилев М. Ю., Каргин В. А., Абалаков А. Н., Николаев Д. А.* Обработка структурно-неоднородной телеметрической информации семейства ракет-носителей «СОЮЗ-2» на активном участке полета // *Информация и Космос*. — СПб.: Институт телекоммуникаций, 2018. — № 2. — С. 157–162.
- [2] *Иваню А. Ю., Копкин Е. В.* Применение технологий искусственного интеллекта в процессах контроля и диагностирования бортовых систем космических средств // *Авиакосмическое приборостроение*. — М.: Научтехлитиздат, 2022. — № 7. — С. 42–49.
- [3] *Иваню А. Ю., Копкин Е. В.* Структурно-стохастическая вычислительная модель контроля технического состояния бортовых систем космических аппаратов // *Известия Тульского государственного университета «Технические науки»*. — Тула: ТулГУ, 2022. — № 10. — С. 9–14.

Решение задачи гендерной и расовой классификации по биометрическому вектору признаков лица человека

Исмагилов Радик Рустемович^{1,2*}

rri10@tpu.ru

Андреев Сергей Юрьевич²

sergA@td.tomica.ru

¹Томск, Национальный исследовательский Томский политехнический университет

²Томск, ООО «Научно-производственная компания «Техника дела»

Технологии распознавания и идентификации лиц долгое время используются в самых различных сферах — от охранных систем и криминалистики до цифрового маркетинга и социальных сетей. В настоящее время образцовым решением подобных задач являются остаточные сверточные нейронные сети (ResNet), обученные с помощью метода ArcFace [1].

На вход модели поступает изображение лица человека, выравненное относительно ключевых точек (центр глаз, кончик носа, уголки губ). Результат работы модели — биометрический вектор, лежащий на поверхности N -мерной сферы. Задачи идентификации и верификации лиц решаются с помощью вычисления косинусного расстояния между биовекторами. Однако возникает вопрос, какую еще информацию о человеке можно получить с помощью биовектора.

Целью данной работы является решение задачи определения гендера и расы человека по биовектору лица с помощью методов машинного обучения.

Результаты решения задачи гендерной классификации по биовектору были приведены в [2, 3]. Так, в [2] лучший результат показал классификатор, обученный методом опорных векторов с гауссовым ядром, в [3] наивысшую точность показал алгоритм k -ближайших соседей. Отметим следующие недостатки приведенных исследований: тренировочные и тестовые выборки были взяты из одного набора данных, что могло заведомо улучшить результаты обучения; среднее время вычислений с помощью предложенных моделей не позволяет решать задачу в режиме реального времени.

В качестве тренировочной и валидационной выборки для обучения моделей мы использовали набор данных FairFace [4], содержащий 108501 изображение лиц людей, размеченных по гендеру (мужчины, женщины), семи расово-этническим группам (европеиды, негроиды, индийцы, представители Восточной Азии, Юго-Восточной Азии, Среднего Востока и Латинской Америки). Для тестирования моделей был выбран набор данных UTKFace [5], содержащий более 20000 изображений лиц, размеченных по гендерному признаку (мужчины, женщины) и пяти расово-этническим группам. Для упрощения решения задачи расовой классификации были выделены следующие группы популяций: европеиды, негроиды, монголоиды (только представители Восточной Азии) и индийцы. Представители Юго-Восточной Азии, Среднего Востока и Латинской Америки были объединены в класс «другие», т. к. на визуальном уровне они сочетают внешние черты других расовых типов, что может затруднить решение задачи.

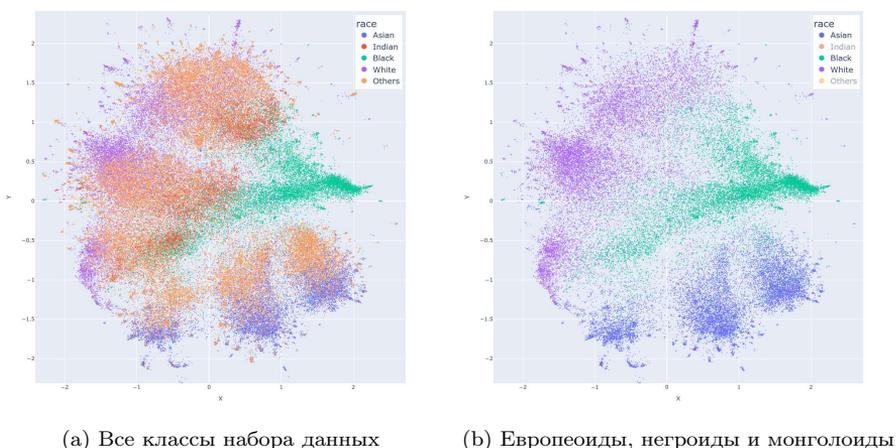


Рис. 1. Распределение биовекторов в пространстве по расовому признаку

Биовекторы изображений лиц были получены с помощью модели ResNet50, обученной с помощью метода ArcFace, на наборе данных WebFace600k.

На рисунке 1 приведена проекция биовекторов из набора данных FairFace на двумерную плоскость с помощью алгоритма UMAP. На графиках можно заметить, что представители классов «индийцы» и «другие» не выделены в отдельный кластер, а распределены на различных участках концентрации трех других групп популяций (рис. 1a). В то время как представители европеоидной, монголоидной и негроидной рас выделены в три кластера (рис. 1b). Это подтверждает гипотезу о том, что остальные расово-этнические группы весьма сложно классифицировать из-за сочетания различных внешних черт трех часто выделяемых основных рас.

Таким образом, было принято решение проводить классификацию в рамках трех расовых групп: европейцы, негроиды и монголоиды. Если вероятность принадлежности человека к каждой из этих групп окажется ниже некоторого порогового значения, то ему будет присвоена метка класса «другие».

Для проведения экспериментов были рассмотрены следующие методы машинного обучения: логистическая регрессия, метод опорных векторов (SVM с линейным и гауссовым ядром), алгоритм k -ближайших соседей (5 и 40 соседей), а также нейронная сеть с одним скрытым слоем (300 нейронов) и слоем Dropout. Нейронная сеть была обучена одновременно на двух задачах классификации — гендерной и расовой, в то время, как классические модели были обучены для каждой задачи отдельно. В таблице 1 приведены средние значения F-меры для каждой из рассматриваемых моделей на валидационной и тестовой выборке.

Таблица 1. Среднее значение F-меры различных классификаторов на валидационной (FairFace) и тестовой (UTKFace) выборках

| Модель | Гендер | | Pasa | |
|-------------------------|----------|---------|----------|---------|
| | FairFace | UTKFace | FairFace | UTKFace |
| Логистическая регрессия | 0,83 | 0,56 | 0,85 | 0,45 |
| SVM, линейное ядро | 0,83 | 0,57 | 0,86 | 0,46 |
| SVM, Гауссово ядро | 0,94 | 0,91 | 0,96 | 0,88 |
| KNN, 5 соседей | 0,91 | 0,91 | 0,95 | 0,89 |
| KNN, 40 соседей | 0,93 | 0,95 | 0,96 | 0,93 |
| Двуслойный перцептрон | 0,94 | 0,91 | 0,97 | 0,88 |

Полученные результаты позволяют сделать вывод, что наилучшей обобщающей способностью для решения задачи классификации лиц по гендерному и расовому признакам обладает классификатор на базе алгоритма k-ближайших соседей с числом соседей, равным 40. Линейные модели показали слабые результаты на тестовой выборке, что говорит о нелинейной зависимости класса от компонент биометрического вектора.

Однако алгоритм k-ближайших соседей обладает высокой вычислительной сложностью, поэтому не применим в биометрических системах. Средняя скорость вычислений модели на базе алгоритма k-ближайших соседей составила 220 мс. Наиболее оптимальной моделью по совокупным показателям качества и скорости работы можно считать перцептрон с одним скрытым слоем — его средняя скорость работы составила 0,2 мс. SVM с гауссовым ядром имеет близкие к перцептронку показатели качества, но среднее время вычислений составило 17 мс. Вычисления проводились на восьмиядерном процессоре AMD FX(tm)-8320 с максимальной тактовой частотой 3,5 ГГц.

- [1] *Deng J., Guo J., Xue N. and Zafeiriou S.* ArcFace: Additive Angular Margin Loss for Deep Face Recognition // Proc. IEEE Conf. on Computer Vision and Pattern Recognition. — 2019. — P. 4685–4694.
- [2] *Farzaneh M.* ArcFace Knows the Gender, Too! // arXiv. — 2021. <https://arxiv.org/abs/2112.10101>
- [3] *Swaminathan A.* Gender Classification Using Facial Embeddings: a Novel Approach // Procedia Computer Science. — 2020. — Vol. 167. — P. 2634–2642.
- [4] *Kärkkäinen K., Joo J.* Fairface: Face Attribute Dataset for Balanced Race, Gender, and Age // arXiv. — 2019. <https://arxiv.org/abs/1908.04913>
- [5] *Zhang Z., Song Y., Qi H.* Age Progression/Regression by Conditional Adversarial Autoencoder // Proc. IEEE Conf. Computer Vision and Pattern Recognition. — 2017. — P. 5810–5818.

Позиционирование человека без нательных датчиков внутри помещений на основе информации о состоянии канала связи

Астафьев Александр Владимирович^{1*}

Alexandr.Astafiev@mail.ru

*Астафьева Олеся Сергеевна*¹

Olesia_Astafieva@mail.ru

¹Муром, Муромский институт (филиал) ФГБОУ ВО «Владимирский государственный университет имени А. Г. и Н. Г. Столетовых»

В современном мире информация играет все большую и большую роль в реализации сложных технических систем. Совершенствование этих систем предъявляет все более и более высокие требования к точности, скорости получения и подробности информации. Примером может являться информация о местоположении человека в пространстве в режиме реального времени. В настоящее время существует огромное количество алгоритмов для организации позиционирования, но самыми сложными из них являются алгоритмы, позволяющие определить местоположение человека без использования нательных датчиков. Целью работы является разработка алгоритма позиционирования человека без нательных датчиков внутри помещений на основе информации о состоянии канала связи (channel state information, CSI). CSI — это информация, которая описывает как сигнал от одного радиоустройства распространяется до другого. Состав оборудования для реализации системы включает в себя набор из радиоустройств, поддерживающих технологию MIMO и OFDM модуляцию. В качестве таких устройств могут выступать WiFi-роутеры. Процесс позиционирования можно представить следующей последовательностью шагов:

1. Передающее радиоустройство передает сигналы на принимающее.
2. Принимающее радиоустройство в момент приема сигнала извлекает CSI с физического уровня самого устройства.
3. Из пакета CSI извлекаются значения фаз и амплитуд по всем парам антенн (технология MIMO) и подортогональным поднесущим (OFDM модуляция).
4. Полученные данные формируют пакет, поступающий на вход нейронной сети, отвечающей за позиционирование.
5. Выходом нейронной сети являются координаты квадрата текущего местоположения человека с заданной точностью.

При использовании стандартов IEEE 802.11n можно извлечь матрицу CSI размером $2 \times 2 \times 56$ (с 2 передающих антенн на 2 принимающие антенны по 56 подортогональным поднесущим), каждое значение которой будет содержать значения фазы и амплитуды. Так, при использовании на входе нейронной сети всех значений, максимальный размер входного вектора может составить 448 значений. В работе предлагается организация нейронной сети с внутренними полносвязными слоями и функцией активации softmax на выходном слое. Выходной слой показывает вероятность нахождения человека в конкретном квадрате наблюдаемой территории. Для проведения экспериментальных исследований бы-

ла размечена территория размером 15×4 метра. Размер квадратной области составил $0,5 \times 0,5$ метра. Схема разметки приведена на рисунке 1.

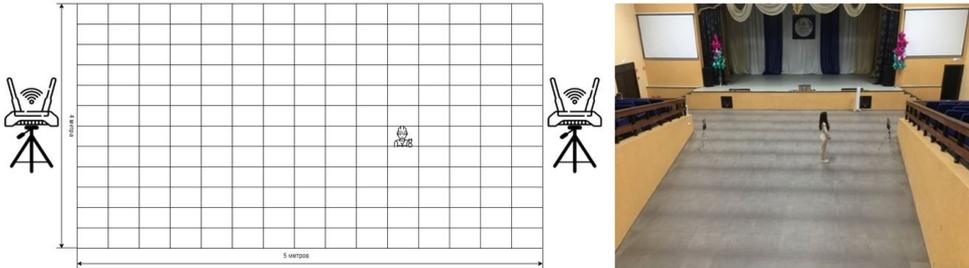


Рис. 1. Схема разметки наблюдаемой территории и ее внешний вид

Величина ошибки определения расстояния варьируется в диапазоне 0,3–8,5 %. Средняя ошибка позиционирования человека без нательных датчиков на наблюдаемой территории составила 14 сантиметров. Результаты позиционирования приведены в таблице 1.

Таблица 1. Результаты работы алгоритма позиционирования

| Исходные данные (IEEE 802.11n) | Архитектура сети | Точность позиционирования |
|---------------------------------------|---|---------------------------|
| Амплитуда 1 пары антенн (56 значений) | 1 скрытый слой, 64 нейронов | 55–76,1 % |
| Амплитуда 2 пар антенн (112 значений) | 1 скрытый слой, 128 нейронов | 73,1–87,8 % |
| Амплитуда 3 пар антенн (168 значений) | 1 скрытый слой, 256 нейронов 2 скрытых слоя, 128 нейронов | 89–98,3 % |
| Амплитуда 4 пар антенн (224 значения) | 1 скрытый слой, 64 нейронов | 91,5–99,7 % |

Работа выполнена при поддержке гранта РФФ № 24-21-00360.

- [1] Астафьев А. В., Астафьева О. С. Локализация радиоустройств внутри помещений на основе информации о состоянии канала связи с использованием ограниченной машины Больцмана // XXX Юбилейная Санкт-Петербургская международная конференция по интегрированным навигационным системам: сборник материалов конференции, Санкт-Петербург, 29–31 мая 2023 года. — СПб.: «Концерн «Центральный научно-исследовательский институт «Электроприбор», 2023. — С. 78–84.

Разработка модели компьютерного зрения для автоматизации контроля действий оператора на линии инспекции

Редько Дмитрий Александрович^{1,3*}

dmitryredko444@gmail.com

Семенов Михаил Евгеньевич^{1,2}

semenov.me@talantiushch.ru

*Турбов Артур Олегович*³

ao.turbob@severstal.com

¹Томск, Томский политехнический университет

²п. г. т. Сириус, Университет «Сириус»

³Череповец, ПАО «Северсталь»

Автоматизация и внедрение современных технологий играют важную роль в промышленности. В ПАО «Северсталь» активно внедряют современные технологии и инновации для оптимизации производственных процессов и повышения качества продукции. Для снижения влияния человеческого фактора в производственных цехах внедряют модели компьютерного зрения. Такую модель можно использовать в том числе для контроля действий сотрудника при осмотре поверхности стальной полосы после оцинкования. В данном случае важно соблюдать требования по технологии при проведении контрольных операций, таких как «брускование» и «осмотр», чтобы обеспечить соответствие продукции установленным стандартам качества.

Цель работы — разработать модель компьютерного зрения для автоматической классификации действий мастера на производстве.

В докладе обсудим выбор модели, процесс сбора и предобработки набора данных, а также оптимизацию гиперпараметров.

Мы собрали набор данных с камер видеонаблюдения, состоящий из видеофрагментов разной длительности (от 16 до 240 секунд) с частотой кадров 12 fps. Затем исходные видеофрагменты были разделены на три класса: «брускование», «осмотр» и «другое». Последний класс включает в себя действия мастера, не связанные с классами «брускование», «осмотр» или статичные изображения.

Для анализа и классификации действий мастера мы провели предобработку данных: каждый видеофрагмент был преобразован в видео с разной частотой кадров: 2, 4 и 8 fps. Затем видеофрагменты были разбиты на кадры продолжительностью 16 секунд каждый, что позволило нормализовать длину видеофрагментов. Окончательно, мы обрезали видеофрагменты, используя рамку с параметрами: верхний отступ — 100 px, левый отступ — 200 px, высота — 1000 px, ширина — 1600 px. Разделение видеофрагментов на три класса позволило создать сбалансированные наборы данных, что важно для обучения модели и получения качественных результатов: «брускование» — 274, «осмотр» — 292, «другое» — 287 видеофрагментов соответственно.

В качестве базовой модели для обучения мы выбрали модель R3D_18 [1], которая является модификацией классической модели ResNet [2] и учитывает временные зависимости.

Таблица 1. Результаты численных экспериментов

| fps | Epochs | Learning Rate | Loss | | Precision | |
|-----|--------|------------------|-------|-------|-----------|-------|
| | | | Train | Test | Train | Test |
| 2 | 15 | 0,0003 | 0,086 | 0,135 | 0,979 | 0,846 |
| 2 | 15 | 0,0003 | 0,060 | 0,058 | 0,964 | 0,931 |
| 4 | 15 | 0,0003 | 0,058 | 0,105 | 0,971 | 0,873 |
| 8 | 20 50 | 0,0003 0,00015 | 0,030 | 0,013 | 0,930 | 0,966 |

Для обучения модели данные были преобразованы с использованием случайно инициализированных весов и адаптивной свертки к размеру видеокладов 112×112 px.

Для оценки эффективности выбранной модели и оптимизации ее гиперпараметров, были проведены серии вычислительных экспериментов. Вначале, исходный набор данных мы разделили на обучающую и валидационную выборки в соотношении 70/30 и установили значения гиперпараметров: количество эпох (epochs) — 15, скорость обучения (learning rate, LR) — 0,003, размер партии (batch-size) — 5, критерий потерь (CR) — CrossEntropyLoss.

Эксперименты с видеофайлами (2 fps) позволили нам оценить производительность модели и выявить тенденции на валидационных данных. Мы установили, что с ростом количества эпох метрики на валидационной выборке скачкообразно изменяются.

Следующая серия экспериментов была связана с использованием в качестве критерия потерь FocalCELoss [3], который разработан для задач с трудно классифицируемыми примерами:

$$\text{FocalCELoss} = -\frac{1}{N} \sum_{i=1}^N \alpha(1 - p_i)^\gamma \log(p_i),$$

где N — количество обучающих примеров, y_i — истинная метка класса для примера i (0 или 1), p_i — предсказанная вероятность принадлежности примера i к классу 1, α — коэффициент для учета дисбаланса классов, γ — параметр, который регулирует влияние «трудных» примеров.

Для достижения требования Precision $> 0,95$ мы использовали данные с частотой кадров 4 и 8 fps, увеличив тем самым сложность задачи и объем данных для обучения. Это позволило нам дополнительно оценить обобщающую способность модели.

Итоговая модель, обученная на 8 fps с увеличением до 50 количеством эпох и корректировкой LR, продемонстрировала прирост в метриках. Корректировка начального значения LR = 0,0003 заключалась в следующем: после 20 эпох значение LR было уменьшено в два раза (таб. 1).

Метрики указывают (рис. 1), что модель не была дообучена до максимума. Эксперименты с увеличением количества эпох более 60 приводят к переобучению, которое можно объяснить недостаточным объемом обучающих данных.

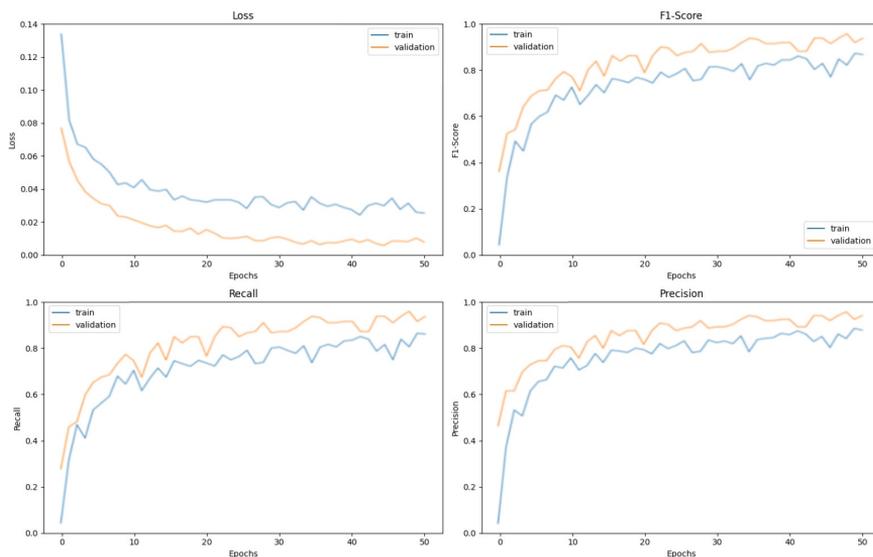


Рис. 1. Результаты обучения при $lr = 0,003$, $batchsize = 3$, $epochs = 50$, $fps = 8$

В работе были проведены вычислительные эксперименты для разработки и оптимизации модели компьютерного зрения, предназначенной для классификации действий мастера на производстве при контроле стальной полосы после оцинкования.

Результаты проведенной работы открывают новые перспективы для дальнейших исследований. В будущем мы планируем расширить задачу классификации и добавить 2 класса: «мелование», «замер ширины», что позволит более полно охватить разнообразие действий мастера на производстве.

- [1] Tran D., Wang H., Torresani L., Ray J., LeCun Y., Paluri M. A Closer Look at Spatiotemporal Convolutions for Action Recognition // arXiv. — 2018. <https://arxiv.org/abs/1711.11248>
- [2] He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition // arXiv. — 2015. <https://arxiv.org/abs/1512.03385>
- [3] Lin T., Goyal P., Girshick R., He K., Dollár P. Focal Loss for Dense Object Detection // arXiv. — 2017. <https://arxiv.org/abs/1708.02002>

Описание подходов к проектированию работающего макета автоматической оптической инспекции

Ходатаева Татьяна Сергеевна^{1*}

tkhodataeva@gmail.com

*Аверина Александра Ивановна*¹

irene75@inbox.ru

*Гурьянов Артем Евгеньевич*¹

artem-guryanov-00@mail.ru

*Русинов Григорий Алексеевич*¹

rusinoff.grigory1@yandex.ru

*Каширин Николай Владимирович*¹

kachnic@mail.ru

¹Йошкар-Ола, ФГБОУ ВО Марийский государственный университет

В Марийском государственном университете в сотрудничестве с ведущим предприятием по производству печатных плат «ТЕХНОТЕХ» проводится исследование, ориентированное на поиск наиболее передовых, современных методов обработки изображений и распознавания образов для создания рабочего макета автоматической оптической инспекции (АОИ), предназначенного для поиска дефектов проводящего рисунка на печатных платах.

Технология АОИ для раннего обнаружения дефектов является приоритетным направлением развития при производстве печатных плат. Системы АОИ необходимы для печатных плат высокой плотности и сверхмалых компонентов.

К настоящему времени авторами изготовлен рабочий макет системы автоматического обнаружения наиболее распространенных дефектов проводящего рисунка печатной платы — это короткие замыкания (short), обрывы (open), выступы (spur), медные включения (островки, copper), микроотверстия (проколы, pin-hole) и отклонение толщины проводников (mousebite).

Аппаратное обеспечения включает модуль освещения, модуль получения изображения, блок управления движением, компьютер, устройство графического дисплея и операционный блок. Программное обеспечение для бесконтактного обнаружения дефектов печатных плат включает в себя модуль обработки изображений, который получает данные с оптического датчика и анализирует их; модуль фильтрации, который устраняет шумы и искажения на изображении; модуль сегментации, который разделяет изображение на объекты и фон; модуль классификации на основе сверточной нейронной сети архитектуры YOLO, определяющий, является ли объект дефектом или нет; и модуль отображения, который выводит результаты на пользовательский интерфейс.

При разработке аппаратного обеспечения было уделено особое внимание модулю освещения для получения качественных изображений. Наилучшее качество изображения для компьютерной обработки было получено при использовании рассеянного освещения красными и синими диодами мощностью 16 мВт (27 мА) с длиной волны 400 ± 10 нм. Интенсивность освещения подбирается отдельно для каждой партии плат, которые проходят АОИ, т. к. принимаются во внимание основной состав печатной платы и свойства отражающих поверхностей, состоящих из линий, отверстий и подложек.

Для обработки изображений особенно важен выбор алгоритма снижения шума без потери особенностей изображения — краев, углов и других «резких» структур, и сохранение мелких деталей изображения. Важно сохранять в изображении мелкие дефекты, чтобы после сглаживания они не были удалены и потеряны. При окончательном выборе метода обработки изображений печатной платы, полученных цифровой камерой, авторы основывались на исследованиях представленных в [1], где сделана попытка определить общую математическую и экспериментальную методологию для сравнения и классификации классических алгоритмов шумоподавления изображений и предложен алгоритм нелокальных средних (NL-means), обеспечивающий сохранение структуры изображения. Вычислительная сложность алгоритма зависит от размера изображения, что делает его дорогостоящим для прямого применения. Для увеличения скорости вычисления окна подобия между двумя пикселями исследовались методы известные как таблицы суммированных площадей [2, 3] и быстрое преобразование Фурье [4].

Для обнаружения дефектов проводящего рисунка печатной платы мы остановились на подходах, которые используют глубокое машинное обучение на основе сверточной нейронной сети (CNN). Основной идеей использования CNN при реализации обнаружения дефектов проводящего рисунка является быстрое нахождение областей, имеющих дефект, сравнение этих областей изображения с соответствующими областями гербер файла и проверки искажения проводящего рисунка печатной платы на допуски для уменьшения количества ложных срабатываний.

Выбор нейронной сети основывался на критериях точности и эффективности, что привело к использованию архитектуры YOLO (You Only Look Once). Для конкретных экспериментов была выбрана архитектура YOLOv4 [5] четвертой версии, которая доступна по открытой лицензии, написана на языке C с использованием библиотеки CUDA и поддерживает вычисления с использованием универсального и графического процессоров. Для решения проблемы затухающего градиента к сверточным слоям была добавлена стратегия CSP (Cross Stage Partial). Использование слоя SPP (Spatial Pyramid Pooling) позволило увеличить рецептивное поле и выделить наиболее значимые контекстные функции. Измененная сеть PANet (Path Aggregation Network) способствовала улучшению процесса сегментации экземпляров за счет сохранения пространственной информации. Также применялись подходы CutMix и Mosaic, регуляризация DropBlock, сглаживание меток класса, использование функции активации Mish. В совокупности все это позволило получить высокую точность при обнаружении маленьких объектов.

Для обучения нейронной сети использовалась сжатая версия YOLOv4 — YOLOv4-tiny, предобученная на наборе данных Microsoft COCO (Common Objects in Context). Обучение YOLOv4-tiny на собственном наборе данных из 60 изображений размерности 416×416 пикселей с двумя классами дефектов

дало хорошие результаты на тестовых данных. Для классов Mouse Bit $ap = 96,21\%$ и Open Circuit $ap = 95,10\%$; $mAP@0,50 = 95,65\%$. Обучение на GPU NVIDIA Quadro P6000 заняло 20 часов процессорного времени.

Авторами также проводились эксперименты с сетью YOLOv8, также предварительно обученной на наборе данных Microsoft COCO и дополнительном наборе данных из 1500 изображений размерностью 640×640 пикселей, имеющих 6 классов дефектов. Набор данных PCBData был взят с платформы Kaggle и адаптирован к формату YOLO. Результаты обучения для 6 классов дефектов: pin-hole, $ap = 98,88\%$; open, $ap = 87,20\%$; short, $ap = 91,80\%$; mousebite, $ap = 90,00\%$; spur, $ap = 89,8\%$; copper, $ap = 97,8\%$; $mAP@0,50 = 92,60\%$. Обучение на GPU NVIDIA Quadro P6000 заняло около получаса.

Наборы данных, которые использовались для обучения CNN создавались вручную, для внесения нужных искажений в изображения использовались графические редакторы. По оценкам для обучения нейронной сети необходимо до 700 изображений на каждый потенциально выявляемый дефект. Было показано, что на точность предсказаний CNN отрицательный эффект оказывает несбалансированность данных.

Отсутствие в открытом доступе больших наборов данных для обучения сети подтвердило актуальность задачи получения реальных изображений дефектов.

На следующих этапах работы планируется проведение сравнения результатов точности обнаружения дефектов с импортными промышленными аналогами АОИ, которые используются в данный момент на отечественных предприятиях.

Работа выполняется в рамках государственного задания на оказание государственных услуг (выполнение работ) № 075-01252-22-03 от 26.10.2022.

- [1] Buades A., Coll B., Morel J. M. A Non-Local Algorithm for Image Denoising // Proc. 2005 IEEE Computer Society Conf. Computer Vision and Pattern Recognition, CVPR, 2005. — P. 60–65.
- [2] Crow F. C. Summed-Area Tables for Texture Mapping // Proc. SIGGRAPH, Minneapolis, 1984. — P. 207–212.
- [3] Viola P., Jones M. J. Robust Real-Time Object Detection // 2nd Int. Workshop on Statistical and Computational Theories of Vision — Modeling, Learning, Computing, and Sampling, Vancouver, Canada Jul. 13, 2001.
- [4] Гонсалес Р., Вудс Р. Цифровая обработка изображений. — 3-е изд., перераб. и доп. — М.: ТЕХНОСФЕРА, 2012. — 1104 с.
- [5] Bochkovskiy A., Wang C.-Y., Mark Liao H.-Y. YOLOv4: Optimal Speed and Accuracy of Object Detection // arXiv. — 2020. <https://arxiv.org/abs/2004.10934>

Применение машинного обучения к задаче выбора сигнально-кодовой конструкции

Бобров Евгений Александрович¹

eugenbobrov@ya.ru

¹Московский государственный университет им. М. В. Ломоносова

В настоящей работе проводится сравнение алгоритмов машинного обучения по метрике ROC-AUC для решения задачи выбора сигнально-кодовой конструкции (Modulation and Coding Scheme – MCS) (рис. 2–4) [1] в системе беспроводной связи ММО (Multiple-input Multiple-output) (рис. 1) при различных скоростях движения пользователей (рис. 5–7).

Спектральная эффективность с учетом ошибок передачи информации во времени $t = \overline{1, T}$ измеряется в бит/с/Гц и вычисляется по формуле:

$$\sum_{t=1}^T SE(mcs_t)ack_t(mcs_t, sinr_t) = \widetilde{SE}(\{mcs_t, ack_t, sinr_t\}_{t=1}^T), \quad (1)$$

где функция $ack_t = ack_t(mcs_t, sinr_t)$ – это индикатор успешной передачи данных (Acknowledgement – ACK) при условии выбранного MCS и заданного отношения сигнала к интерференции и шуму (Signal-to-Interference-plus-Noise-Ratio – SINR). Величина SINR измеряется на базовой станции с помощью пилотных сигналов. Задача настоящей работы заключается в максимизации функции \widetilde{SE} (1) посредством выбора оптимального MCS при заданном SINR.

Рассмотрим выбор MCS по следующей формуле:

$$\widehat{mcs}(sinr_t) = \arg \max_{mcs} \{p(ack|mcs, sinr_t)SE(mcs)\}, \quad (2)$$

где ищется точка максимального произведения убывающей по MCS функции вероятности $p(ack|mcs, sinr)$ и возрастающей табличной функции $SE(mcs)$.

Замечание 1. О решении задачи оптимального выбора MCS.

Для функции (1) метод (2) является оптимальным по мат. ожиданию:

$$\begin{aligned} \mathbb{E}_{(ack_1, \dots, ack_T)} \widetilde{SE}(\{\widehat{mcs}(sinr_t), ack_t, sinr_t\}_{t=1}^T) &= \\ &= \max_{(mcs_1, \dots, mcs_T)} \mathbb{E}_{(ack_1, \dots, ack_T)} \widetilde{SE}(\{mcs_t, ack_t, sinr_t\}_{t=1}^T). \end{aligned} \quad (3)$$

Таким образом, метод (2) является оптимальным для функции (1).

Исследуется система ММО с одной базовой станцией и одним пользователем, оснащенными многими антеннами. Пользователь движется вокруг базовой станции по случайной траектории с постоянной скоростью на плоскости. Алгоритм в онлайн режиме выбирает сигнально-кодовую конструкцию (MCS), обеспечивающую максимальную спектральную эффективность (SE) системы (рис. 1).

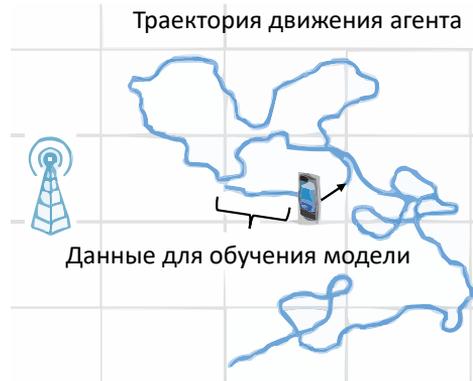


Рис. 1. Описание сценария задачи выбора сигнално-кодовой конструкции

Для оценки вероятности успеха передачи сигнала (АСК) из формулы (2) применяется алгоритм машинного обучения — $p_{\theta}(ack|mcs, sinr)$ с параметрами θ , так что $p_{\theta} \sim p$. В качестве признаков используется вектор значений SINR на каждой антенне пользователя и значение MCS, а в качестве меток — успех передачи данных (АСК). Алгоритм выбирает оптимальный MCS по формуле (2). Параметры модели θ настраиваются в процессе обучения с помощью градиентного спуска по собранным в процессе движения данным.

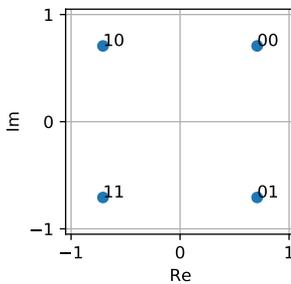


Рис. 2. QAM-4

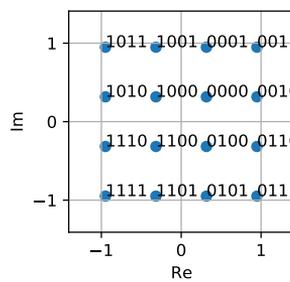


Рис. 3. QAM-16

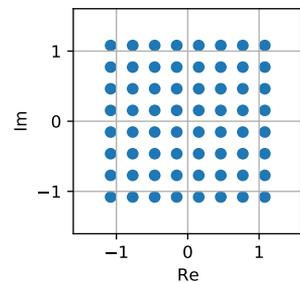


Рис. 4. QAM-64

Сигнално-кодová конструкция (MCS) переводит бинарную последовательность цифрового сигнала в комплексные числа для передачи аналогового сигнала и обратно с помощью квадратурной амплитудной модуляции (QAM — Quadrature-Amplitude-Modulation). Большая модуляция соответствует более плотному коду, позволяющему передать большее количество информации за один промежуток времени, но при этом налагает более высокие требования к значению отношения сигнала к интерференции и шуму (SINR).

Приводится сравнение методов машинного обучения по метрике ROC-AUC, а также исследуется устойчивость методов к изменению скорости движения пользователей, когда алгоритм обучается на одной скорости движения, а применяется на другой (рис. 5–7). Несмотря на то, что нейросетевая модель усту-

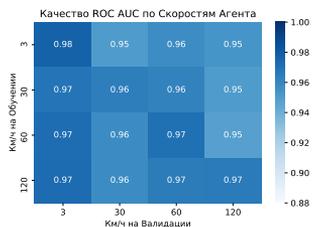


Рис. 5. Нейросеть

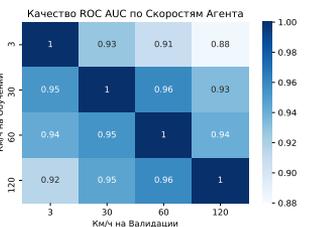


Рис. 6. Случайный лес

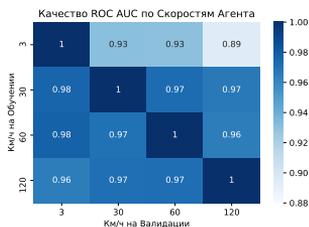


Рис. 7. Град. бустинг

пает по качеству градиентному бустингу (таб. 1), она имеет возможность динамического дообучения в онлайн режиме и показывает лучшую обобщающую способность в терминах стандартного отклонения по различным сценариям, что делает ее более устойчивой к перемене скоростей движения пользователей. Также нейросетевая модель на этапе предсказания для данной задачи по вычислительной эффективности превосходит метод градиентного бустинга [2], а также превосходит по качеству алгоритм случайного леса [3] и метод опорных векторов с гауссовым ядром.

Таблица 1. Среднее значение и стандартное отклонение ROC-AUC между различными скоростям пользователей и возможность дообучения методов.

| Метод маш. обучения | Возможность дообучения | Среднее ROC-AUC | Отклонение ROC-AUC |
|------------------------|------------------------|-----------------|--------------------|
| Метод опорных векторов | + | 0,945 | 0,035 |
| Нейросетевая модель | + | 0,962 | 0,009 |
| Случайный лес | +/- | 0,951 | 0,034 |
| Градиентный бустинг | - | 0,968 | 0,028 |

[1] Bobrov E., Kropotov D., Lu H., Zaev D. Massive MIMO Adaptive Modulation and Coding Using Online Deep Learning Algorithm // IEEE Communications Letters. — IEEE, 2021. — Vol. 26, No. 4. — P. 818–822.

[2] Alamgir M., Sultana M. N., Chang K. Link Adaptation on an Underwater Communications Network Using Machine Learning Algorithms: Boosted Regression Tree Approach // IEEE access. — IEEE, 2020. — Vol. 8. — P. 73957–73971.

[3] Wang A., Wan G., Cheng Z., Li S. An Incremental Extremely Random Forest Classifier for Online Learning and Tracking // 16th IEEE Int. Conf. Image Processing (ICIP). — IEEE, 2020. — P. 1449–1452.

Особенности применения сверточных нейронных сетей для распознавания аневризм сосудов головного мозга

Кружалов Алексей Сергеевич^{1}*

alexkruzhalov@gmail.com

Филиппович Андрей Юрьевич¹

a.y.philippovich@mospolytech.ru

¹Москва, Московский политехнический университет

В докладе рассматриваются особенности применения сверточных нейронных сетей при решении задачи распознавания аневризм сосудов головного мозга с точки зрения влияния используемых методов формирования наборов обучающих примеров и гиперпараметров алгоритма обучения на точность получаемых результатов распознавания.

Интерес исследователей к проблематике разработки автоматизированных систем распознавания патологий на медицинских изображениях обусловлен все более широким внедрением цифровых технологий в сферу медицины и необходимостью повышения эффективности работы специалистов. На сегодняшний день уже накоплен определенный опыт применения искусственных нейронных сетей (ИНС) для решения задачи распознавания церебральных аневризм. В большинстве исследований последних лет рассматриваются различные разновидности сверточных нейронных сетей. При их обучении необходимо учитывать, что ангиограммы сосудов головного мозга представляют собой трехмерные массивы достаточно больших размеров, и поэтому использовать их целиком при обучении нейронной сети, как правило, не представляется возможным. Эта проблема решается путем разбиения исходного изображения на отдельные фрагменты. При этом возникает проблема выбора оптимального размера фрагментов изображений для обучения нейронной сети, способов семплирования фрагментов, методов предварительной обработки, исходя из особенностей решаемой задачи и используемой архитектуры нейронной сети.

В данной работе предложена методика распознавания аневризм сосудов головного мозга, основанная на обобщении накопленного опыта применения сверточных нейронных сетей для решения поставленной задачи. Было проведено исследование влияния размера используемого контекста (фрагмента изображения), способов стохастического семплирования фрагментов, а также методов предварительной обработки изображений на эффективность процесса обучения нейронной сети.

Для проведения вычислительных экспериментов были использованы данные, опубликованные в рамках соревнования Cerebral Aneurysm Detection (CADA). В этом наборе данных представлено 110 изображений (результатов ротационной ангиографии сосудов головного мозга).

В качестве алгоритма распознавания используется сверточная нейронная сеть, являющаяся обобщением широко известной архитектуры U-Net. Для обучения нейронной сети был выбран алгоритм Adam. В качестве целевой функции использовалась сумма взвешенной бинарной перекрестной энтропии и DiceLoss.

В докладе приводятся результаты 3 серий вычислительных экспериментов, направленных на изучение влияния на процесс обучения размеров извлекаемых из изображений фрагментов, способов семплирования фрагментов и методов предварительной обработки. Для оценки полученных результатов был использован коэффициент Дайса-Серенсена (DSC). В качестве статистических методов обработки результатов проведенных экспериментов были использованы тест Фридмана, критерий Уилкоксона (для попарных сравнений) и метод Бенджамини-Хохберга для контроля ожидаемой доли ложных отклонений гипотез (англ. false discovery rate, FDR).

В первой серии экспериментов нейронные сети обучались на фрагментах изображений различного размера. Фрагменты извлекались случайным образом, но с условием, что по центру должна располагаться область интереса (аневризма) или область, соответствующая здоровому сосуду. Для выделения сосудов был использован метод треугольника. При семплировании фрагментов вероятности выбора были заданы таким образом, что фрагменты с областью интереса выбирались в 4 раза чаще.

Во второй серии экспериментов было проведено исследование влияния различных методов предварительной обработки данных и их комбинаций на точность результатов обучения нейронных сетей. Структура нейронной сети и параметры алгоритма обучения остались прежними, варьировались только методы предобработки. Было рассмотрено влияние стандартизации, ресемплинга и различных вариаций преобразований с использованием порогового значения, которые задаются представленными ниже формулами:

$$x'_{k,l,m} = \begin{cases} 1, & \text{если } x_{k,l,m} > t; \\ 0, & \text{иначе;} \end{cases}$$

$$\mathbf{X}' = \mathbf{X} \odot G(LCA([\mathbf{X} > t]));$$

$$x'_{k,l,m} = \frac{1}{1 + e^{-(x_{k,l,m} - t)}}, \quad (1)$$

где $x_{k,l,m}$ — значение вокселя в позициях k, l, m до предобработки; $x'_{k,l,m}$ — значение вокселя в позиции k, l, m после предобработки; t — значение порога, вычисленное по методу треугольника; \mathbf{X} — многомерный массив изображения до предобработки; \mathbf{X}' — многомерный массив изображения после предобработки; LCA — функция извлечения наибольшей связной области; G — функция размытия с фильтром Гаусса; \odot — произведение Адамара (покомпонентное произведение).

В третьей серии экспериментов было проведено исследование методов генерации карт вероятностей семплирования фрагментов и формирования наборов

обучающих примеров. Исследовано влияние учета при семплировании положения аневризмы (ROI), сосудов (выделяемых по методу Оцу и методу треугольника), а также соотношения вероятностей выбора фрагментов, соответствующих различным областям.

Во всех сериях проведенных экспериментов были получены статистически значимые результаты ($\alpha = 0,05$). Первая серия экспериментов показала, что использование фрагментов небольшого размера в комбинации с увеличением размера пакета обучающих примеров позволяет получить значимо более высокую точность распознавания [1]. По результатам второй серии экспериментов можно сделать вывод о том, что применение к изображению в совокупности стандартизации, ресемплинга и сигмоидного преобразования (формула (1)) позволяет значимо снизить среднее количество ложно положительных прогнозов. Третья серия экспериментов показала важность учета при семплировании фрагментов не только области интереса, но и здоровых сосудов, т.к. важно предъявлять нейронной сети при обучении не только положительные примеры, но и отрицательные (не содержащие область интереса), причем отрицательные примеры не должны быть слишком простыми (относиться к «пустым» областям изображения, не содержащим сосуды). При этом значимое влияние на получаемые результаты оказывает также и используемый метод выделения сосудов [2]. Наилучшая точность распознавания среди всех проведенных экспериментов составила $DSC = 0,85$.

- [1] Кружалов А. С. Методика обучения сверточной нейронной сети по фрагментам медицинских изображений в задаче распознавания церебральных аневризм // Моделирование, оптимизация и информационные технологии. — 2023. — Т. 11, № 2(41). — С. 3–4.
- [2] Kruzhalov A., Philippovich A. Analysis of Thresholding Methods for the Segmentation of Brain Vessels // International Conference on Analysis of Images, Social Networks and Texts. — Cham : Springer International Publishing, 2021. — P. 85–95.

Трекинг и идентификация игроков в командных видах спорта

*Витвицкий Антон Александрович*¹★

darkalert2517@gmail.com

*Потапов Сергей Алексеевич*¹

scorp321@yandex.ru

*Гаро Вадим Вячеславович*¹

vadim27garo@gmail.com

¹Москва, ООО «АВТОМАТИЗАЦИЯ ВИДЕОАНАЛИЗА»

Спортивная аналитика предоставляет ценную информацию об эффективности команд и отдельных игроков, которая широко используется клубами, тренерами, аналитиками и скаутами. На текущий момент в этой области работает несколько иностранных крупных компаний, однако они предоставляют аналитику на основе данных, собираемых преимущественно вручную. Автоматизация же этого процесса классическими методами требует дорогостоящего оборудования (3д-камер), которое могут себе позволить не все клубы. Учитывая степень развития технологий искусственного интеллекта, решения, основанные на использовании нейросетевых моделей, кажутся очень перспективными. В этой работе мы представляем нашу попытку разработать полностью автоматизированное решение для бизнеса, которое позволяет отслеживать и идентифицировать игроков в командных видах спорта на основе видеоданных.

Типичный трекер состоит из шагов детекции, сопоставления этих детекций с треками, и алгоритма создания и удаления треков. В качестве детектора мы используем YOLOv8 [1], который обучается на данных из спортивного домена, и может детектировать игроков, голкиперов и арбитров. На наших тестах детектор достигает mAP 0,868 и 0,881 для футбола и баскетбола, соответственно. Сопоставление треков и детекций происходит на основе предсказания движения игроков при помощи фильтра Калмана, и сравнения этих предсказаний с фактическими детекциями. Дополнительно, мы используем стадию реидентификации, в которой для игроков вычисляются векторы признаков при помощи модели OSNet [1], однозначно описывающие игроков. В качестве алгоритмов создания треков и их сопоставления с детекциями мы используем те, что были предложены в StrongSORT [2]. На наших тестах трекер достигает НОТА 77,732 и 80,214 для футбола и баскетбола, соответственно.

Для подсчета персональных статистических данных необходимо однозначно идентифицировать игроков, для чего достаточно знать их команду и игровой номер. Задача определения команды может быть решена путем обучения классификатора, однако такой классификатор не будет универсальным, т. к. будет способен классифицировать только те команды, которые были в обучающей выборке. Кроме того, в некоторых играх присутствуют игроки, чья форма отличается от основной формы команды (вратари в футболе, или либеро в волейболе). Мы используем подход, который не требует стадии обучения, а именно — метод кластеризации k-means, который может разделять игроков на заданное число кластеров. В качестве векторов признаков для k-means, мы используем

те же векторы из OSNet. Точность нашего классификатора составляет 90–95 % в зависимости от вида спорта.

Второй элемент идентификации — номера на форме. Эта информация также помогает корректировать ошибочные треки на этапе постпроцессинга. Мы используем детектор YOLOv8 для детекции номеров, и TrOCR [3] для распознавания найденных номеров. Точность распознавания на наших тестах составляет 86–90 % в зависимости от вида спорта и качества видео.

В результате работы трекера предсказывается большее число треков, чем реальное число игроков. Иногда треки могут прерываться из-за перекрытий игроков другими игроками, или в следствие пропадания игрока из кадра на некоторое время. Такие треки могут быть довольно короткими и не всегда могут быть идентифицированными по номеру/команде. Поэтому мы используем дополнительную стадию постобработки, в которой треки одного и того же игрока могут объединяться либо разделяться в случае ошибочного слияния на предыдущих этапах. В качестве векторов признаков мы также используем векторы из OSNet, при помощи которых происходит сопоставление идентифицированных треков (известны номер и команда) с не идентифицированными. Данный этап дает прирост точности от 5 до 10 %.

Чтобы использовать предсказанные траектории игроков для различных статистических показателей, таких как скорость, ускорение, пройденный путь, распределение по зонам и т. п., необходимо перейти от относительных координат кадра к абсолютным координатам поля/корта. Используя тот факт, что игроки бегают по плоскости, для преобразования координат можно использовать гомографию. В случае статичной панорамной камеры оценить гомографию достаточно просто по известным ключевым точкам. Однако, при использовании PTZ-камер (видеотрансляции/технические камеры), необходимо оценивать гомографию в каждом кадре. Для этого мы разработали следующую модель: на вход энкодер-декодера поступает игровой кадр и предсказывается сегментационная маска корта/поля, далее эта маска поступает на вход регрессионной сети, которая предсказывает матрицу гомографии 3×3 . Такой подход обладает большей устойчивостью, чем подходы, основанные на вычисление ключевых точек и алгоритмах DLT и RANSAC. Ошибка репроекции нашей модели составляет от 26 до 49 см в зависимости от вида спорта и размера игровой площадки.

Для комплексного тестирования всех моделей, которое бы отражало степень решения реальных бизнес-задач, мы разработали метрику, которая показывает распределение треков игроков по игровому времени и оценивает, какой процент времени каждый из игроков был отслежен корректно (предсказанные координаты трека а также номер и команда совпадают с реальными данными), а значит, может быть использован в подсчете статистики. Мы получили следующие оценки: для стритбола (баскетбол трое на трое) 89 %, для баскетбола — 81 %, для футбола — 59 %. Как можно заметить, точность предсказаний напрямую зави-

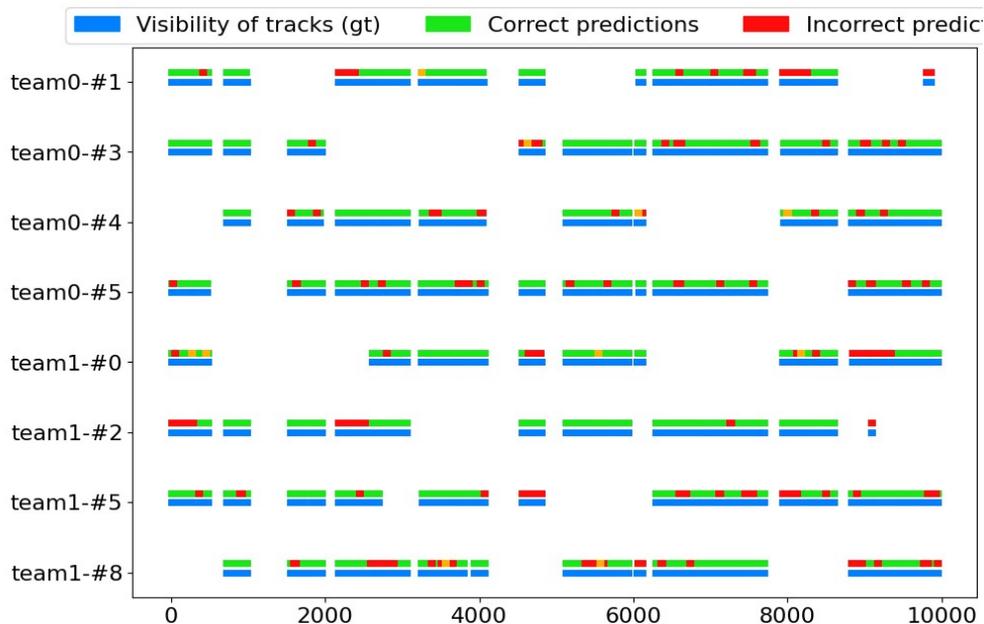


Рис. 1. Пример распределение треков для баскетбола трое на трое

сит от сложности игры (чем больше игроков и чем больше игровая площадка, тем сложнее трекинг). На рисунке 1 показан пример распределения треков.

Работа поддержана грантом Старт-1 2022.3.

- [1] Jocher G., Chaurasia A., Qiu J. YOLO by Ultralytics (Version 8.0.0) [Computer software], 2023. <https://github.com/ultralytics/ultralytics>
- [2] Zhou K. et al. Omni-Scale Feature Learning for Person Re-Identification // arXiv. — 2019. <http://arxiv.org/abs/1905.00953>
- [3] Du Y., et al. StrongSORT // IEEE Transactions on Multimedia. — 2023. — P. 1–14.
- [4] Li M. et al. TrOCR // arXiv. — 2021. <https://arxiv.org/abs/2109.10282>

Интеллектуальный программно-аппаратный комплекс с оцифровкой жестов для дошкольного образования

Малинина Алина Эдуардовна^{1*}

alinamalinina512@gmail.com

*Яворский Ростислав Эдуардович*²

ryavorsky@gmail.com

¹Иваново, Ивановский государственный энергетический университет

²Астана (Казахстан), GBL Academy

Данная работа является частью проекта по разработке программно-аппаратного комплекса для организации занятий с детьми дошкольного возраста в форме образовательной ситуации [1]. Образовательная ситуация строится на основе сценария взаимодействия педагогов и обучающихся по поводу определенного содержания образования и обычно включает три этапа:

- этап 1 — создание образовательной напряженности, уточнение образовательного объекта, конкретизация задания;
- этап 2 — решение ситуации, работа с культурно-историческими аналогами, систематизация полученных знаний;
- этап 3 — рефлексия, самооценка детьми образовательной ситуации, учительская оценка ответов.

Использование информационных технологий и цифровых игрушек нового поколения позволяет существенно повысить производительность и эффективность образовательного процесса на каждом из перечисленных этапов.

Разрабатываемое решение включает центральный вычислительный узел, к которому подключаются различные цифровые компоненты, в том числе устройство для оцифровки жестов. Как видно на рисунке 1 небольшие размеры платы дают много разных возможностей для внешнего вида этого модуля.

Разработка модуля оцифровки жестов требует использования технологий определения траектории перемещения объекта в многомерном пространстве в режиме реального времени и распознавания 3Д почерка. Основные методы можно разделить на следующие группы: акустические, радиочастотные, магнитные, оптические, инерциальные, гибридные [2]. Наибольший интерес представляют системы трекинга, основанные на использовании инерциальных приборов — акселерометров и датчиков угловых скоростей. Такие датчики отличаются относительно небольшими размерами и могут быть жестко закреплены непосредственно на объекте. Инерциальные системы трекинга (IMU) успешно применяются в VR-тренажерах, а также технологиях отслеживания тела человека и визуальной имитации [3]. Например, в [4] описаны методы расчета угла наклона сустава на основе данных инерциальных измерений в контексте анализа движения человека. Развитие методов машинного обучения также нашло широкое применение в решении задач распознавания жестов, см. [5–8]

Прототип устройства для оцифровки жестов описан в [6]. В качестве платы формы нами выбрана плата ESP32-S3 [9], к которой подключен 6-осевой датчик положения MPU-6050 (GY-521), объединяющий акселерометр и гироскоп

и использующийся в любительской робототехнике для определения положения в пространстве [10] (см. фотографию на рис. 1).

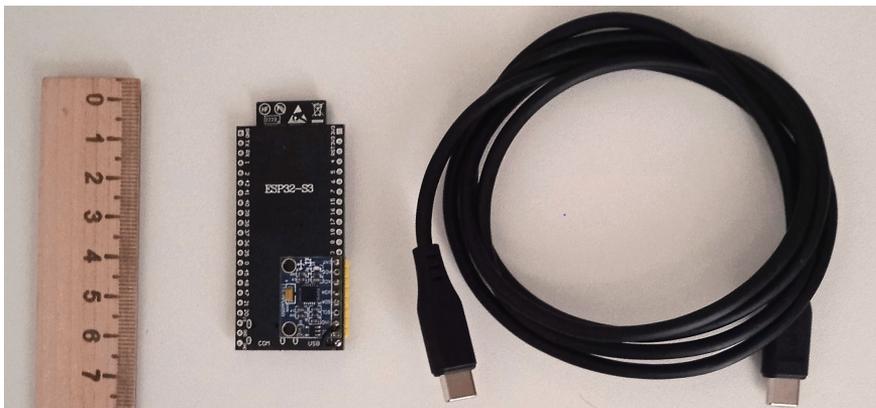


Рис. 1. Прототип устройства для оцифровки жестов

Модуль оцифровки жестов планируется использовать для занятий, развивающих моторику и чувство ритма у детей дошкольного возраста, а также для интерактивных обучающих игр, повышающих заинтересованность и вовлеченность детей.

Развитие координации движений (крупная моторика) — это активные движения телом, в которых задействованы руки, ноги, туловище, голова. Хотя взрослые пользуются этими навыками, не задумываясь, детям приходится их постепенно осваивать. Разрабатываемый модуль будет использован для оценки уровня развития этих навыков у детей дошкольного возраста.

Чувство ритма — это способность активно переживать музыку, чувствовать эмоциональную выразительность музыкального ритма и точно его воспроизводить. Известно, что ребенок с развитым чувством ритма говорит увереннее, выразительнее, умело расставляет паузы и акценты в предложении, лучше усваивает информацию, что в итоге отражается в его обучении. Разрабатываемое устройство по сути является цифровой дирижерской палочкой, которая может быть использована в музыкальных играх для оценки и развития чувства ритма у детей. В [11] описан опыт использования аналогичного устройства для визуальных упражнений, направленных на адаптивную коррекцию тонкой моторики кистей рук в условиях невесомости. Одной из целей нашей работы является применение данной технологии для задач дошкольного образования.

[1] Землянская Е. Н. Образовательные ситуации как единица событий образовательного процесса // Наука и школа. — 2020. — № 3. — С. 83–92.

- [2] Кармадонов В. Ю. Методы отслеживания положения в виртуальной реальности // *Academy*. — 2019. — Т. 12, № 51. — С. 19–22.
- [3] Кручинина А. П., Латонов В. В., Чертополохов В. А. Обзор технологий визуальной имитации в тренажерных системах // *Пилотируемые полеты в космос*. — 2019. — № 3. — С. 89.
- [4] Seel T., Raisch J., Schauer T. IMU-Based Joint Angle Measurement for Gait Analysis // *Sensors*. — 2014. — Vol. 14, No. 4. — P. 6891–6909.
- [5] Девятков В. В., Алфимцев А. Н. Распознавание манипулятивных жестов // *Вестник Московского государственного технического университета им. Н. Э. Баумана. Серия «Приборостроение»*, 2007. — № 3. — С. 56–74.
- [6] Хельвас А. В., Беляйкина Н. Г., Гиля-Зетимов А. А., Черникова Д. Д., Шабунин В. М., Япрынцева Е. О. Распознавание жестов с помощью нейронной сети и применение этого подхода для создания игровых гаджетов нового поколения // *Труды Московского физико-технического института*. — 2017. — Т. 9, № 2 (34). — С. 164–175.
- [7] Чуйков А. В., Вульфим А. М. Система распознавания жестов на основе нейросетевых технологий // *Вестник Уфимского государственного авиационного технического университета*. — 2017. — Т. 21, № 3 (77). — С. 113–122.
- [8] Муратов Е. Р., Никифоров М. Б., Скачков А. М. Искусственный интеллект в задачах распознавания динамических жестов // *Современные информационные технологии и ИТ-образование*. — 2020. — Т. 16, № 4. — С. 883–892.
- [9] Cameron N. ESP32 Microcontroller // *In ESP32 Formats and Communication: Application of Communication Protocols with ESP32 Microcontroller*. — Berkeley, CA: Apress, 2023. — P. 1–54.
- [10] Fedorov D. S., Ivoilov A. Y., Zhmud V. A., Trubin V. G. Using of Measuring System MPU6050 for the Determination of the Angular Velocities and Linear Accelerations // *Automatics & Software Engineering*. — 2015. — Vol. 11, No. 1. — P. 75–80.
- [11] Демин А. А., Власов А. И., Шахнов В. А. Методы и средства визуальных упражнений для адаптивной коррекции тонкой моторики кистей рук в условиях невесомости // *Вестник Московского государственного технического университета им. Н. Э. Баумана. Серия «Приборостроение»*. — 2015. — Т. 3, № 102. — С. 23–38.

Алгоритм детекции и сегментации осетровых

Тевяшов Глеб Константинович¹*

glebtevyashov96@yandex.ru

Русаков Константин Дмитриевич¹

rusakov.msk@yandex.ru

Мещеряков Роман Валерьевич¹

mrv@ipu.ru

¹Москва, Институт проблем управления им. В. А. Трапезникова РАН

В сфере аквакультуры и рыбных хозяйств для решения задач сохранения и восполнения популяций активно внедряются киберфизические системы автоматического определения параметров воды в режиме реального времени, систем технического зрения для решения «визуальных» задач, а также систем автоматического кормления [1,2]. Одной из наиболее актуальных задач на рыбных хозяйствах, в частности, осетровых, является определение средней навески в бассейне. Это необходимо для того, чтобы оценить проводимые с рыбой мероприятия: кормление кормом с различным составом, изменение условий их содержания и влияние других факторов на рост особей [3].

Для измерения длины рыбы достаточно производить съемку в одной плоскости. Разработан экспериментальный стенд, который устанавливается на бассейн осетрового хозяйства (рис.1).

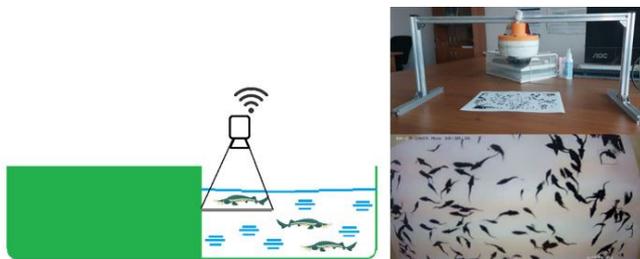


Рис. 1. Разработанный макет

В нашем экспериментальном проекте по детекции и сегментации осетровых на изображении с использованием модели YOLOv8-seg мы составили датасет, основанный на материалах Филиала по пресноводному рыбному хозяйству ФГБНУ «ВНИРО». Этот датасет включает примерно 2000 изображений, демонстрирующих широкий спектр видов осетровых в различных условиях окружающей среды. Каждое изображение тщательно аннотировано прямоугольными рамками и масками на уровне пикселей для точного отделения осетровых от фона.

Обучающий набор данных состоит из 1800 изображений, каждое из которых аннотировано для задач детекции (прямоугольная рамка) и сегментации (маска). Валидационный набор включает 200 изображений. Разнообразие углов съемки, условий освещения и фонов на изображениях обеспечивает надежное обучение и валидацию, помогая модели хорошо обобщать данные в разных сце-

нариях. Этот датасет имеет ключевое значение для улучшения нашего понимания поведения и среды обитания осетровых, способствуя разработке более эффективных практик управления рыбным хозяйством.

В результате, для каждой обнаруженной особи создается маска, которая точно повторяет контуры рыбы. Маска представляет собой бинарное изображение, где пиксели, принадлежащие осетровой, обозначены единицами, а фон — нулями. Это позволяет точно измерить длину рыбы, поскольку все пиксели, относящиеся к особи, могут быть отслежены и измерены линейно.

На представленных графиках (рис. 2) демонстрируется изменение ключевых мер в процессе обучения и валидации модели машинного обучения для задачи детекции и сегментации рыбы. Мера Accuracy отражает общую точность модели, достигая значения 0,98; Precision, показывающая точность распознавания рыбы, достигает уровня 0,97; Recall, или полнота, достигает значения 0,965.

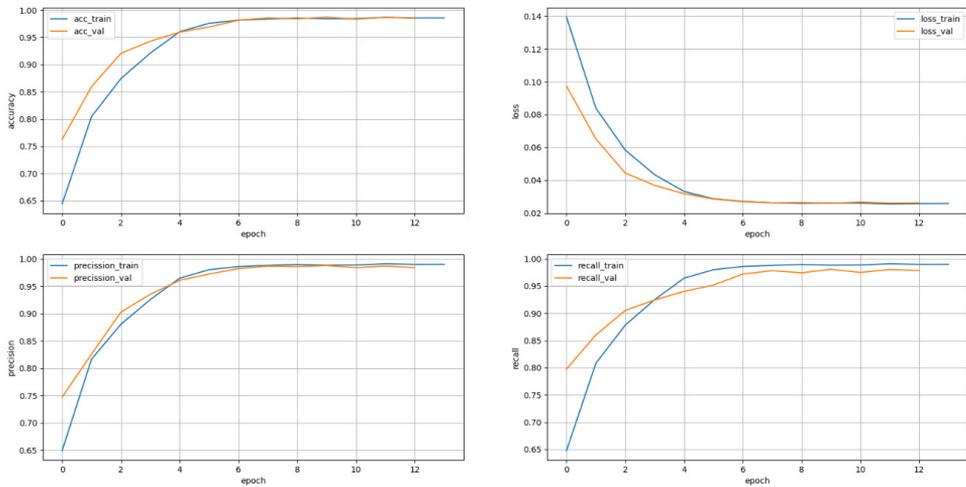


Рис. 2. Графики обучения модели

По полученным результатам можно сделать вывод, что модель YOLOv8-seg проявляет адекватные результаты в задачах детекции и сегментации осетровых, обеспечивая надежное и точное определение положения и размеров рыбы в резервуарах. Высокие значения мер измерений указывают на то, что система может быть эффективно внедрена в реальные условия работы рыбных хозяйств, способствуя повышению их производительности и устойчивости. Ниже на рисунке 3 предоставлен пример работы системы с камеры, установленной над резервуаром.

- [1] Пешкин К. С. Назначение и перспективы развития киберфизических систем управления технологическими процессами // Вестник Самарского государственного тех-



Рис. 3. Пример работы алгоритма

- нического университета. Серия: Технические науки. — 2023. — Т. 31, № 1. — С. 47–59.
- [2] *Meshcheryakov R. V., Ananiev P. P., Anannieva A. V., Belyakov K. O., Nevredinov A. V., Tevyashov G. K., Mikhailova M. V.* Using the Digital Reef to Monitor and Enhance Biological Resources in Aquaculture // *Agriculture Digitalization and Organic Production*. — St Petersburg, Russia: Springer, 2022. — P. 425–437.
- [3] *Тевяшов Г. К.* Применение нейросетевого алгоритма для мониторинга динамики роста осетровых в УЗВ // *Материалы 3-й Международной научно-практической конференции «Цифровизация агропромышленного комплекса»* (Тамбов, 2022). — Тамбов: Издательский центр ФГБОУ ВО «ТГТУ», 2022. — Т. 1. — С. 223–225.
- [4] *Садиев С. А.* Анализ подхода к построению системы оценки численности рыб, основанной на нейронной сети YOLOV4 // *StudArctic Forum*. — 2022. — Т. 7, № 3. — С. 40–44.

Автоматизация процесса подготовки исходных данных для гидрогеологического моделирования

Ильина Софья Андреевна^{1,2*}

iliinasa@mail.ru

Семенов Михаил Евгеньевич^{2,3}

semenov.me@talantiuspeh.ru

¹Москва, АО «ВНИПИПромтехнологии»

²Томск, Томский политехнический университет

³п. г. т. Сириус, Университет «Сириус»

При создании компьютерных гидрогеологических моделей важным фактором является качество исходных данных, которые могут иметь различные формы представления. Так, одним из вариантов получения исходной информации являются изображения геологических разрезов интересующего участка с отмеченными скважинами. Такие разрезы содержат в себе информацию о литологическом составе слоев, о положении скважин относительно азимута, их глубине и уровне их устья. В зависимости от целей построения гидрогеологических разрезов они также могут содержать дополнительную информацию, например, уровень подземных вод на участке или области оруднения с определенным процентом содержания некоторого вещества. Однако в процессе извлечения информации из исходных данных такого формата возникают проблемы: отсутствие масштабной линейки на изображениях разрезов увеличивает время их обработки, при этом полученная информация может содержать большие погрешности и ошибки.

В связи с этим, в данной работе предлагается метод автоматизации данного процесса, основанный на компьютерной обработке изображений. Предлагаемый алгоритм получает на вход полное изображение геологического разреза в виде растрового файла (скан, фотография, снимок экрана и др.), а результатом его работы является электронная таблица (*.xlsx), содержащая информацию по каждой скважине, представленной на входном изображении, а именно: номер скважины, абсолютные отметки ее устья и уровня подземных вод, глубины подошвы и (или) кровли и описание литологического состава для каждого слоя, границы областей оруднений.

Т. к. на вход поступает полное изображение разреза, первым этапом алгоритма является обнаружение вертикальных проекций скважин на разрезе (Рис. 1). Поиск скважин осуществляется путем поиска их номеров методами оптического распознавания символов (Optical Character Recognition, OCR) вблизи линии азимута. Азимут всегда располагается ниже разреза, а отметки скважин вдоль азимута определяют положение забоя и устья проекции скважины на разрезе по оси абсцисс. Зная положение устья скважины по оси абсцисс, ее положение по оси ординат определяется обнаружением верхней границы разреза; вблизи устья скважины методами OCR распознаются ее номер и абсолютная отметка устья. Далее алгоритм анализирует пиксели вдоль линии скважины для обнаружения пересечений слоев. Для анализа выделяются квадраты размером 8×8

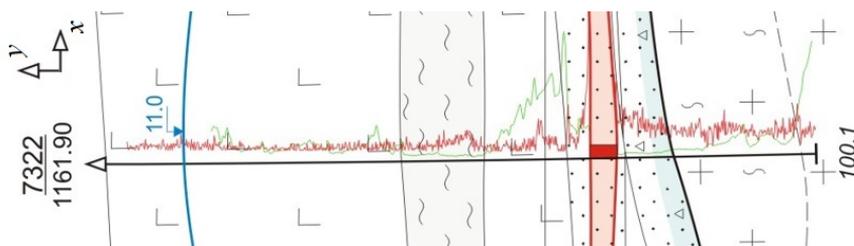


Рис. 1. Изображение разреза скважины 7322, 1161.90 — абсолютная отметка устья, 11.0 — уровень подземных вод, 100.1 — глубина скважины, область красного цвета — урановое оруднение

пикселей, бинаризируются методом Оцу [1] и проверяются на наличие на них вертикальной линии, соответствующей линии скважины и линии, ориентированной горизонтально, соответствующей разделу слоев, уровню подземных вод или границе оруднения, в зависимости от цвета линии на исходном изображении.

Таким образом, анализируется область вблизи линии скважины до тех пор, пока не будет обнаружена отметка ее забоя — короткая горизонтальная черта, ниже которой отсутствует линия скважины. Ниже отметки забоя расположено значение его глубины, которое распознается методами OCR. Зная глубину скважины, абсолютный уровень ее устья и количество пикселей между двумя этими отметками, можно вычислить абсолютные и относительные глубины разделов слоев, границ областей оруднения. Для определения масштаба использована формула:

$$m_{pp} = \frac{deep_p - wellhead_p}{deep_m},$$

где $deep_p$, $wellhead_p$ — положение забоя и устья скважины по оси ординат в пикселях, $deep_m$ — указанная глубина скважины в метрах.

В процессе анализа скважины, при обнаружении линий раздела слоев (черные, ориентированные горизонтально линии), вырезается фрагмент изображения не менее 30 пикселей в высоту и 100 пикселей в ширину, содержащий условные обозначения слоя, расположенного между двумя линиями раздела. Эти фрагменты поступают на вход классифицирующей нейронной сети, которая определяет литологический состав, соответствующий данным условным обозначениям. В случае обнаружения уровня подземных вод (синяя, горизонтально ориентированная линия), вырезается фрагмент изображения правее и выше обнаруженной точки, содержащий глубину уровня вод в метрах, которая извлекается из данного фрагмента методами OCR.

Описанный алгоритм реализован средствами языка программирования Python. Для работы с изображениями использовались возможности библиотеки OpenCV [2], такие как функции бинаризации, изменение цветовой схемы, чте-

ние и запись изображений. Методы OCR реализованы при помощи механизма TesseractOCR [3], использованы такие операции как выделение текста на изображении с учетом его ориентации и шрифта (Orientation and script detection, OSD), распознавание текста при помощи LSTM (Long short-term memory) нейронной сети, обученной на наборе символов русского и английского алфавитов, ограничение допустимого набора символов при распознавании (для распознавания чисел допустимыми символами являются только цифры от 0 до 9 и точка). Определение литологического состава производится при помощи классифицирующей нейронной сети линейной конфигурации с двумя полносвязными слоями, использующей функции активации ReLU (Rectified linear unit) и многопеременные логистические функции (Softmax). На данный момент реализуется классификация на 42 класса, в последствии возможно увеличение числа классов.

Применение всех перечисленных технологий позволило реализовать безошибочное извлечение числовых значений из изображения разреза, которые определяют номер скважины, абсолютную отметку устья, глубину забоя и уровень подземных вод. Также получено достаточно точное определение линий раздела слоев и границ оруднений. Обусловленная алгоритмом погрешность определения точки пересечения границы слоя (или оруднения) с линией скважины не превышает 8 пикселей. В проведенных экспериментах погрешность не превышала 3 пикселей. В переводе на метры погрешность будет зависеть от глубины скважины. Однако возможны случаи ложного обнаружения линии, т. к. некоторые условные обозначения литологического состава слоя могут восприниматься алгоритмом как линии раздела слоев. Таким образом, могут появиться лишние линии раздела внутри одного слоя, которые можно впоследствии исключить, объединяя соседние слои с одинаковым составом. Случаи не обнаружения разделов слоев не наблюдались, бинаризация методом Оцу увеличивает контраст, явно выделяя все контуры изображения. Основной проблемой на данном этапе является высокая ошибка классификации при определении литологического состава слоев (около 54 % на 56 примерах), определяемая формулой: $acc_{рас} = P/N$, где P — количество верно классифицированных слоев, N — общее количество слоев в выборке. Возможными вариантами уменьшения ошибки являются увеличение обучающей выборки, в том числе путедополнительной аугментации, или изменение конфигурации нейронной сети.

- [1] Otsu N. A Threshold Selection Method from Gray-Level Histograms // IEEE Transactions on Systems, Man, and Cybernetics. — IEEE, 1979. — Vol. 9, No. 1. — P. 62–66.
- [2] OpenCV [Электронный ресурс]. — Режим доступа: <https://github.com/opencv/opencv/>, свободный (дата обращения: 10.11.2023).
- [3] TesseractOCR [Электронный ресурс]. — Режим доступа: <https://github.com/tesseract-ocr/tesseract/>, свободный (дата обращения: 10.11.2023).

Детекция игровых событий в командных видах спорта на основе видео

Коншин Артем Максимович^{1}*

a.m.konshin@yandex.ru

Витвицкий Антон Александрович¹

darkalert2517@gmail.com

¹Москва, ООО «АВТОМАТИЗАЦИЯ ВИДЕОАНАЛИЗА»

Под игровыми событиями подразумеваются различные события, происходящие во время спортивных матчей и влияющие на ход игры. Например, в футболе таким событием может быть удар мячом в ворота или передача паса от одного игрока другому. Эта информация может применяться для подсчета различных статистических показателей и широко используется в спортивной аналитике. Однако, на текущий момент сбор таких данных осуществляет преимущественно вручную и является крайне трудозатратным, т. к. количество различных игровых событий в одном матче может достигать нескольких тысяч [1].

Нейросетевые методы, применяемые для решения задач распознавания действий, могут быть задействованы и для решения данной задачи. Однако, поиск событий на длинных видео (полный игровой матч) может быть осложнен следующими факторами. Во-первых, необходимо разработать решение, которое сможет улавливать как пространственный контекст события, так и временной. Во-вторых, разные типы событий требуют разного временного контекста из-за разной продолжительности событий. К примеру удар по воротам в среднем занимает не более 0,2 секунд, а единоборство может занимать до 3–4 секунд.

Еще одной сложностью является необходимость обрабатывать последовательности изображений высокого разрешения. Связано это с тем, что игроки занимают лишь малый процент от общей картинки, а нейросетевые методы, традиционно используемые для обработки последовательности изображений, обычно не превышают стандартного разрешения 224×224 , что усложняет их прямое использование для последовательностей изображений в высоком разрешении из-за крайней низкой скорости работы.

Мы провели анализ различных современных нейросетевых подходов к поиску событий на длинных видео. Такие подходы могут быть условно разделены на двухстадийные, в которых сперва применяются модели и алгоритмы, строящие векторное представление объектов на изображениях, а затем — алгоритмы, которые определяют тип игрового события на основе их векторного представления [2]. Противоположностью этому — одностадийные (end-to-end) подходы, которые позволяют одновременно обучать сеть энкодера (которая строит векторное представление), и головной части сети, выполняющей финальный поиск и классификацию событий.

Разработанный нами подход основан на использовании 2д и 3д сетей, которые совместно обеспечивают баланс между скоростью и качеством работы (см. рис. 1). Мы используем комбинацию из нескольких нейронных сетей, входными данными для которых являются последовательности кадров (видеокли-

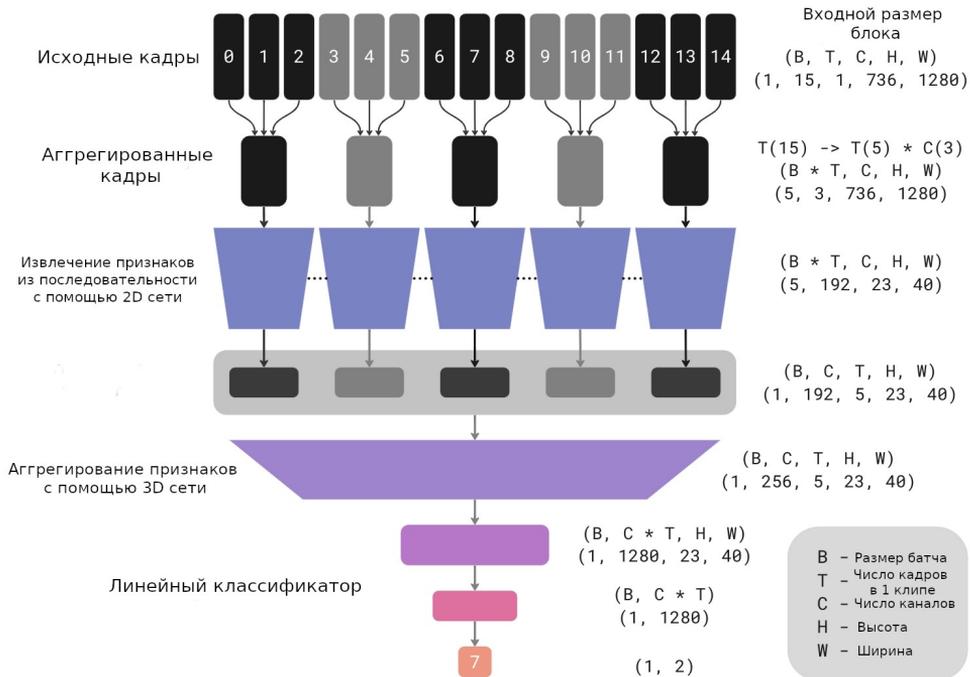


Рис. 1. Схема работы предложенного решения

пов) из целого видео, а выходными данными — предсказанная временная метка и метка класса, которая соответствует одному из целевых событий. На вход модели поступает последовательность кадров в градациях серого, после чего кадры агрегируются и поступают на вход 2д сети (например, EfficientNet или ViT), на выходе из которой каждый агрегированный кадр представлен в виде числового вектора. Затем эти векторы подаются на вход финальной 3д сети, которая и определяет метку класса для данного клипа. При прохождении скользящим окном по всем кадрам видео мы получаем агрегированное представление данных, в котором каждому кадру исходного видео соответствует некоторая метка класса и ее вероятность. Далее при помощи функции Гаусса полученные предсказания преобразовываются и определяются пики сигнала, которые и представляют собой предсказания финальных временных меток игровых событий.

На наших тестах предложенная модель достигает mAP 79,3 и 80,2 для футбола и баскетбола, соответственно. При этом скорость обработки одного матча составляет несколько часов на машине оснащенной GPU NVIDIA RTX 4080.

Работа поддержана грантом Старт-1 2022.3.

-
- [1] *Cioppa A. et al.* SoccerNet 2023 Challenges Results // arXiv. — 2023. <https://arxiv.org/abs/2309.06006>
 - [2] *Vahdani E. et al.* Deep Learning-based Action Detection in Untrimmed Videos: A Survey // arXiv. — 2021. <https://arxiv.org/abs/2110.00111>

Формирование системы осмысленных признаков и ее применение для определения упорядоченности расположения наночастиц на снимках с электронного микроскопа

Курбаков Михаил Юрьевич^{1*}

muwsik@mail.ru

*Сулимова Валентина Вячеславовна*¹

vsulimova@yandex.ru

*Середин Олег Сергеевич*¹

oseredin@yandex.ru

*Копылов Андрей Валериевич*¹

and.kopylov@gmail.com

¹Тула, Тульский Государственный Университет

Электронная микроскопия позволяет осуществлять увеличение объектов в миллионы раз и фиксировать результат на изображениях, в результате чего находит активное применение для исследования микроскопических объектов [1]. В частности, одна из актуальных задач анализа изображений, полученных с электронного микроскопа, связана с анализом количества, размеров и взаимного расположения металлических наночастиц, наносимых на поверхность исследуемых углеродных материалов. Полученная в результате такого анализа информация используется для определения значимых характеристик исследуемых материалов и выявления их особенностей, например, для выявления на поверхности катализаторов скрытых дефектов [2, 3], влияющих на их свойства, но не поддающихся обнаружению другими методами [4]. При этом для исследования только одного образца материала могут формироваться сотни и даже тысячи изображений, поэтому их анализ является серьезной проблемой, требующей автоматизации ее решения.

Предшествующие работы [5,6] были сосредоточены на автоматизации детектирования отдельных наночастиц без исследования более глобальных свойств материала, например, анализа взаимного расположения наночастиц и определения их общего порядка. Предложенный же ранее нейросетевой подход [4] представляет собой модель «черного ящика» и не позволяет дать обоснование полученного решения, что является крайне важным для дальнейшего исследования свойств катализаторов.

В связи с этим в [7] предлагается новый подход к определению порядка расположения наночастиц на поверхности углеродных материалов. Метод использует данные визуализации наночастиц и формализует концепцию их упорядоченного расположения с учетом специфики человеческого восприятия. В результате выявляется ряд легко интерпретируемых признаков, которые в дальнейшем используются при объяснимом анализе.

Анализ был проведен на базе изображений [8], полученных с электронного сканирующего микроскопа для нескольких различных материалов, содержащей 750 снимков с упорядоченным и 250 снимков с неупорядоченным расположением наночастиц, среди которых имеются изображения разного масштаба с варьирующейся степенью плотности расположения наночастиц и содержащие

различные видимые особенности поверхности материалов, не имеющие отношения к упорядоченности расположения наночастиц.

В результате анализа факта упорядоченности линейный классификатор (SVM), использующий всего 7 интерпретируемых признаков, показывает более высокое качество, чем нейросетевой подход (ResNet34), использующий 21 миллион абстрактных параметров.

Авторы благодарят Научную школу акад. В. П. Ананикова за тематику применения машинных методов в нанотехнологиях, полезные дискуссии и предоставленные экспериментальные данные.

Работа выполнена при финансовой поддержке Министерства науки и высшего образования РФ в рамках государственного задания FEWG-2021-0012.

- [1] *Илюшин А. С., Орешко А. П.* Введение в дифракционный структурный анализ // Учебное пособие. — М.: физический факультет МГУ, 2008. — 335 с.
- [2] *Liu X., Astruc D.* Development of the Applications of Palladium on Charcoal in Organic Synthesis // *Advanced Synthesis and Catalysis*. — 2018. — Vol. 360, No. 19. — P. 3818–3818.
- [3] *Felpin F., Ayad T., Mitra S.* Pd/C: An Old Catalyst for New Applications—Its Use for the Suzuki–Miyaura Reaction // *European Journal of Organic Chemistry*. — 2006. — Vol. 2006, No. 12. — P. 2679–2690.
- [4] *Boiko D. A., Pentsak E. O., Cherepanova V. A., Gordeev E. G., Ananikov V. P.* Deep Neural Network Analysis of Nanoparticle Ordering to Identify Defects in Layered Carbon Materials // *Chem. Sci.* — 2021. — No. 12. — P. 7428–7441.
- [5] *Kurbakov M. Y., Sulimova V. V.* High-Performance Two-Level Parallel Computing Scheme for Nanoparticles Detection in SEM Images // *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* — 2023. — P. 145–150.
- [6] *Kopylov A. V., Kurbakov M. Y., Seredin O. S., Sulimova V. V., Boko D. A., Cherepanova V. A., Pentsak E. O., Ananikov V. P.* Automated Recognition of Nanoparticles in SEM images of Pd/C Catalysts // *Nanomaterials*. — 2022. — Vol. 12, No. 21. — P. 3914.
- [7] *Kopylov A. V., Kurbakov M. Y., Seredin O. S., Sulimova V. V., Boiko D. A., Pentsak E. O., Ananikov V. P.* Determining the Orderliness of Carbon Materials with Pd-Nanoparticle Imaging and Explainable Machine Learning // *Nanomaterials* (в печати).
- [8] *Boiko D. A., Pentsak E. O., Cherepanova V. A., Ananikov V. P.* Electron Microscopy Dataset for the Recognition of Nanoscale Ordering Effects and Location of Nanoparticles // *Sci Data*. — 2020. — Vol. 8, No. 101.

Модифицированный сегментатор дефектов на бетонных изделиях

*Дементьев Виталий Евгеньевич*¹

dve@ulstu.ru

*Ким Александр Леонидович*²

alkim@fa.ru

*Коновалов Матвей Алексеевич*²

matveikon8287@gmail.com

Андрянов Никита Андреевич^{2*}

naandriyanov@fa.ru

¹Ульяновск, Ульяновский государственный технический университет

²Москва, Финансовый университет при Правительстве Российской Федерации

Сегодня все большее и большее применение получают методы компьютерного зрения в таких областях, как неинвазивный мониторинг или контроль изделий. Действительно, традиционные этапы мониторинга строительных конструкций зачастую связаны с их визуальным анализом. При этом необходимо привлечение специалистов, обеспечение их выезда на объекты. Кроме того, точность работы человека со временем снижается, приходится делать перерывы. Оптимизация задачи мониторинга бетонных изделий возможна с применением систем поддержки принятия решений на базе алгоритмов компьютерного зрения. Следует отметить, что при анализе дефектов на бетоне необходимо не только их обнаружить, но и выделить полностью, поскольку это позволит принять верные будущие решения. Таким образом, необходимо решение задачи сегментации [1]. В этой задаче необходимо классифицировать каждый пиксель изображения, чтобы найти пиксели, относящиеся к дефекту. Мерой качества в таких задачах обычно выступает отношение площади пересечения прогноза дефекта и реального дефекта к площади их объединения (критерий Дайса, Dice-score).

В статье предложен новый метод для сегментации дефектов на изображениях. Используется предварительная обработка изображения с помощью подсветки краев, а также аугментации изображений с разметкой для детекции. Реализация подсветки происходит в соответствии со следующим алгоритмом. Сначала выполняется медианная фильтрация исходного изображения. Затем применяется детектор границ Канни. Наконец, выполняется фильтрация с максимизацией. При этом ядра фильтров во всех случаях составляют размеры 5×5 . Двухэтапный алгоритм сначала выполняет детектирование трещин, а после уточняет их с помощью модели сегментации. Такой подход позволил повысить метрику Dice-Score на 6,8 % за счет применения сильного детектора и аугментирования.

Задача обнаружения дефектов на железных и бетонных изделиях в автоматическом режиме [2] является актуальной в связи с необходимостью мониторинга большого числа конструкций неинвазивными методами. Технологии компьютерного зрения позволяют лишь по данным с камеры выполнять анализ структуры строительных конструкций, железнодорожных путей и т. д. В этой работе, в отличие от [2], особое внимание уделяется методам сегментации бетонных изделий, а также предлагаются новые методы решения общей задачи

сегментации. Для сравнения использовался алгоритм на базе сверточной нейронной сети Mask R-CNN [3, 4].

В качестве исходного датасета был выбран набор данных RoboFlow по задаче сегментации дефектов на бетонных изделиях [5]. На рисунке 1 приведен пример изображений с дефектами из данной базы.

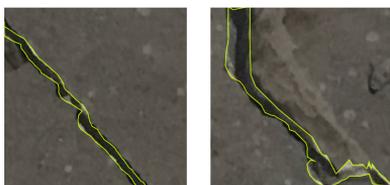


Рис. 1. Пример размеченных данных

Особенностью датасета является малое количество данных. Так, обучающая выборка составляет 59 изображений, валидационная — 16 изображений, а тестовая — 14. Поэтому было предложено, во-первых, использовать аугментацию изображений, но для этого выборка была переразмечена в датасет детекции, для которого проще проводить преобразования разметки. Кроме того, было предложено использовать пайплайн предобработки: медианный фильтр, выделение краев, фильтр максимума. Это позволило улучшить отличие дефектных зон от нормальных (рис. 2).



Рис. 2. Подсветка краев

В итоге с помощью детекционных преобразований обучающая выборка была увеличена в 3 раза, а затем на детектированных участках снимка применялась модель SegFormer. Для детекции был выбран и дообучен трансформер DINO. Тестовая выборка была одинаковой. В результате экспериментов с различными моделями были получены следующие значения метрики Dice, представленные в таблице 1.

Таким образом, предложенный пайплайн позволил превзойти результаты модели на базе Mask R-CNN более, чем на 6 %. В целом, также помогают аугментация и предобработка изображений, обеспечивающая выделение краев. Это позволяет соотнести предложенную модель с известными и показывает преимуще-

Таблица 1. Сравнение моделей

| Модель | Аугментация | Выделение краев | Dice |
|----------------|-------------|-----------------|-------|
| Mask R-CNN | - | - | 82,68 |
| Mask R-CNN | - | + | 84,38 |
| DINO+SegFormer | - | - | 83,82 |
| DINO+SegFormer | + | + | 89,43 |

щество применения трансформерных архитектур. Недостатком остается меньшая скорость трансформера по сравнению со сверточной сетью. В связи с этим в будущем планируется рассмотреть методы квантизации полученной модели.

Работа поддержана грантом РФФ №23-21-00249.

- [1] *Andriyanov N., Dementiev V., Suetin M.* Neural Networks Combinations for Detecting and Highlighting Defects in Steel and Reinforced Concrete Products // International KES Conference on Intelligent Decision Technologies, 2023. — P. 293–301.
- [2] *Андрьянов Н. А., Волменко А. А., Дементьев В. Е.* Мониторинг состояния металлических изделий на основе систем компьютерного зрения // Нейрокомпьютеры: разработка, применение. — М.: Радиотехника, 2023. — Т. 25, № 5. — С. 50–57.
- [3] *Wei T., Xu C., Qin L., Chen Y., Fangfang G., Yinyin C.* Meso-Structure Segmentation of Concrete CT Image Based on Mask and Regional Convolution Neural Network // Materials and Design. — 2021. — Vol. 208. — P. 109919.
- [4] *Andriyanov N. A., Dementiev V. E., Tashlinskiy A. G.* Detection of Objects in the Images: From Likelihood Relationships towards Scalable and Efficient Neural Networks // Computer Optics. — 2022. — Vol. 46, No. 1. — P. 139–159.
- [5] Онлайн-набор данных. https://universe.roboflow.com/-akkub/crack-detection_p (дата обращения: 12.11.2023).

Сегментация изображений рук и инструментов в коллаборативной робототехнической ячейке

У Го^{1*}

Шэнь Синь¹

Цзинь Цзюнь¹

Серебрянный Владимир Валерьевич¹

ug@student.bmstu.ru

shs20ua057@student.bmstu.ru

tszints@student.bmstu.ru

vsereb@bmstu.ru

¹Москва, МГТУ им. Н. Э. Баумана

В данной работе представлена разработанная система компьютерного зрения для промышленных сценариев применения коллаборативных роботов. Основной акцент сделан на задаче сегментации изображений рук и инструментов. В коллаборативной робототехнической ячейке роботы должны понимать и адаптироваться к действиям человека, поэтому взаимодействие руки с инструментом стало актуальной темой исследований. Традиционные методы обнаружения объектов демонстрируют определенные ограничения в коллаборативных средах. Обнаружение объектов рассматривает весь объект или руку как целое и не может предоставить детали тонкого взаимодействия между рукой и инструментом. В статье [1] использовался модуль механизма внимания для улучшения детектора YOLOv5. Несмотря на улучшение различных показателей, в условиях сильного перекрытия все еще нет идеального результата в режиме реального времени. В случае общего рабочего пространства требуется более подробная и точная информация о взаимодействии руки с инструментом, чтобы помочь роботам избегать столкновений с руками и инструментами и обеспечить безопасность процесса сотрудничества.

Для решения этой проблемы данное исследование фокусируется на сегментации изображений при взаимодействии руки с инструментом. Сегментация изображений — это задача в компьютерном зрении, направленная на выделение различных объектов на изображении и присвоение уникального идентификатора каждому экземпляру. Через сегментацию изображений робот может лучше понимать взаимное положение, отношение перекрытия и статус взаимодействия руки и инструмента. На данный момент не были выполнены работы по сегментации рук и инструментов в коллаборативной робототехнической ячейке, существуют только наборы данных для взаимодействия рук и объектов в повседневной жизни.

Создан специализированный набор данных, ориентированный на задачу сегментации изображений рук и инструментов. Набор данных WorkingHands [3] идеально подходит для задач взаимодействия руки с инструментами, содержит 13 видов инструментов, но предоставляет только маски для семантической сегментации. Всего в набор данных включено 7906 RGB-изображений, из которых часть получена с камеры Kinect V2, а часть создана с помощью программного обеспечения Blender, при этом размер каждого изображения составляет 1920 × 1080 пикселей для Kinect V2 и 864 × 656 — для Blender. С использованием

платформы RoboFlow была выполнена разметка данного набора данных, в результате чего получены маски рук и инструментов. Поскольку между рукой и инструментом существует контакт, при разметке важно максимально перекрыть границы. Данные были разделены на обучающую, тестовую и валидационную выборки в соотношении 8 : 1 : 1. Перед обучением были применены техники аугментации данных, включая размытие по Гауссу, зеркальное отображение и изменение яркости.

Исследованы существующие нейросети для сегментации изображений, а также рассмотрено использование нейросети PointRend [2]. По сравнению с Mask R-CNN, в сети PointRend многослойный перцептрон используется для получения точечного представления признаков изображения на краю объекта, что позволяет обновить начальное обнаружение. В результате получаются более тонкие края объектов, таким образом, маски руки и инструмента максимально перекрываются друг с другом. Была использована предварительно обученная на наборе данных COCO модель. Все методы обучались с использованием основной сети ResNet50-FPN. Обучение проводилось на системе Ubuntu 20.04 с видеокартой NVIDIA GeForce RTX 4060, установленной в ноутбуке с 8 ГБ оперативной памяти. При обучении были заданы следующие гиперпараметры моделей: размер пакета (batch size) — 8, начальная скорость обучения — 0,001 и количество эпох — 200, для остальных параметров использовались значения по умолчанию.

Таблица 1. Результаты оценки набора данных WorkHands

| Нейросеть | mAP рук | mAP инструментов |
|------------|---------|------------------|
| Mask R-CNN | 59,0 | 33,8 |
| PointRend | 71,2 | 35,8 |

Производительность моделей оценивается с использованием показателя mAP. mAP рассчитывается как средняя точность модели при различных пороговых значениях IoU (пересечение над объединением). В таблице 1 показаны результаты валидации двух моделей для сегментации рук и инструментов. Значение показателя mAP для сегментации инструментов после обучения нейросети PointRend на 2,0 выше, чем у Mask R-CNN. Значение показателя mAP для сегментации рук после обучения нейросети PointRend на 12,2 выше, чем у Mask R-CNN. Результаты экспериментов показывают, что нейросеть PointRend более подходит для задач сегментации рук и инструментов при взаимодействии человека и робота.

Работа поддержана грантом Минобрнауки России № FSFS-2021-0004.

- [1] *Wu Guo, Shen Xin, Serebrenny Vladimir.* Attention Guided In-Hand Mechanical Tools Recognition in Human-Robot Collaborative Process // Int. Conf. Interactive Collaborative Robotics. — Springer, 2023. — P. 1–12.

-
- [2] *Kirillov A., Wu Y., He K., Girshick R.* Pointrend: Image segmentation as rendering // Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition. — IEEE, 2020. — P. 9799–9808.
- [3] *Shilkrot Roy, Narasimhaswamy S., Vazir Saif, Hoi M.* WorkingHands: A Hand-Tool Assembly Dataset for Image Segmentation and Activity Mining // Proc. British Machine Vision Conference. — BMVA, 2019. — P. 258–269.

Технология интеллектуального автоматизированного отбора информативных признаков натуральных гиперспектральных изображений

Хотилин Максимилиан Ильич¹

turbomax_94@yandex.ru

¹Самара, Самарский университет

Анализ гиперспектральных изображений является одной из популярных тематик, а автоматизация процесса анализа и процесса поиска информативных признаков является актуальной задачей в области обработки изображений и компьютерного зрения [1, 2]. В рамках данной работы рассматривается набор методов, включающий метод поиска информативных признаков и сверточные нейронные сети, для задачи классификации гиперспектральных изображений.

Обработка изображений, содержащих большое количество слоев, например гиперспектральных, требует значительных вычислительных ресурсов, ввиду ряда особенностей. Например, для классификации одной области гиперспектрального изображения размером 10×10 пикселей количество необходимых совокупных яркостных и текстурных признаков может составлять десятки тысяч. Актуальность данной работы заключается в сочетании методов, позволяющем автоматизировать процесс поиска и значительно снизить временные и аппаратные ресурсы, используемые в процессе анализа гиперспектральных изображений.

Весь процесс обработки можно разделить на последовательно выполняемые этапы. На первом этапе выбирается набор рассматриваемых изображений. Далее посредством использования модуля предобработки происходит отображение каждого из рассматриваемых изображений на набор значений своих текстурных и яркостных признаков [3, 4].

Следующим этапом проводится обработка полученного массива данных с целью снижения его размерности. Производится поиск признаков, являющихся информативными. Также из рассмотрения убираются константные признаки и признаки, не имеющие значения для отдельных изображений.

Далее, используя метод снижения размерности, описанный ранее в [5], производится дальнейшее снижение размерности. Добавляя метку класса к каждому из экземпляров отображения, получаем набор для проведения дальнейшей классификации. После этого переходим к поиску информативных признаков, посредством применения метода последовательного добавления признаков. Используя полученные на предыдущих шагах данные и информативные признаки, производится обучение сверточной нейронной сети, позволяющей по набору входных параметров, в качестве вывода предоставлять необходимое количество искоемых информативных признаков.

В качестве исходных данных рассматривался набор изображений HSI Dataset v1.3, состоящий из изображений листьев различных растений. Размер каждого из изображений — 512×512 пикселей с количеством спектральных каналов — 237. Было выбрано 4 класса для рассмотрения: листья яблони, кар-

тофеля, травы, клубники. Посредством Python был реализован модуль предобработки, позволяющий найти ряд признаков изображения, таких как текстурные, гистограммные, морфологические и др. Общее количество признаков для каждого изображения составило 62 152 539.

Далее было проведено 3 серии экспериментов с различными наборами данных и классификаторами. В качестве исходных данных рассматривались: исходный полный набор признаков; фильтрованный набор, из которого были удалены «выбросы» — изображений, которые большинство классификаторов или ни один из них не смогли верно классифицировать; признаки, отобранные посредством корреляционного анализа.

Каждая новая серия экспериментов позволяла значительно улучшать результаты классификации. Например, при использовании корреляционного анализа, при удалении константных, N/A признаков, и варьировании коэффициента пороговой корреляции удалось значительно снизить размерность рассматриваемых данных. Рассматривались пороговые коэффициенты корреляции от 0,5 до 0,99 с шагом в 0,01. При этом количество признаков составило от 28 (при коэффициенте в 0,5) до 2364 (при 0,99).

В качестве вспомогательных использовались алгоритмы классификации: линейный дискриминантный анализ, метод опорных векторов, логистическая регрессия, метод случайного леса, алгоритм Voruta [6] совокупно с методом случайного леса, а также многослойный перцептрон (MLPClassifier) с различными параметрами. В качестве наиболее эффективного по точности классификации и затраченным ресурсам, были выбран LDA и Voruta.

Следующим шагом является поиск информативных признаков из набора, оставшегося в результате снижения размерности. В итоге получаем паттерн обработки гиперспектральных изображений, обладающий меньшими ресурсными требованиями, по сравнению с классическими методами. В результате работы данного паттерна мы получаем данные, которые могут являться обучающими для нейронной сети. Иными словами, зная входные (изображения) и выходные (признаки) данные, можно обучить нейронную сеть не только классифицировать входные изображения, но и определять информативные признаки данных изображений.

В рамках проверки данной гипотезы, в качестве нейронной сети была выбрана сверточная CFR-сеть, использующая алгоритм MCCFR. В настоящее время ведется работа над обучением и настройкой данной нейросети.

Поиск признаков, определяющих однозначно принадлежность изображений к классу, является одной из значимых задач обработки изображений. В случае рассмотрения гиперспектральных изображений необходимо провести некоторые преобразования для оптимизации используемых ресурсов. В качестве одного из современных подходов к сокращению требуемых для вычислений ресурсов можно выделить использование сверточных нейронных сетей, ввиду ряда их свойств, что потенциально позволяет использовать описанные в статье набор

методов на портативных устройствах, например носимых устройствах или беспилотных летательных аппаратах.

- [1] *Zimichev E. A.* Spectral-Spatial Classification with k-Means++ Particional Clustering // Computer Optics. — 2014. — Vol. 38, No. 2. — P. 281–286.
- [2] *Kazanskiy N. L.* Simulation of Hyperspectrometer on Spectral Linear Variable Filters // Computer Optics. — 2014. — Vol. 38, No. 2. — P. 256–270.
- [3] *Khotilin M. I.* Classification of Objects of Natural Hyperspectral Images // 2020 International Conference on Information Technology and Nanotechnology (ITNT), 2020. — P. 1–3.
- [4] *Goncharova E. F.* Greedy Algorithms of Feature Selection for Multiclass Image Classification // CEUR Workshop Proc. (IPERS-ITNT 2018–Proc. Int. Conf. Information Technology and Nanotechnology–Session: Image Processing and Earth Remote Sensing), 2018. — P. 38–46.
- [5] *Хотилин М. И.* Технология автоматизированного интеллектуального отбора информативных признаков для задачи классификации областей натуральных гиперспектральных изображений // Информационные технологии и нанотехнологии (ИТНТ-2023) : сб. тр. по материалам IX Междунар. конф. и молодеж. шк. (г. Самара, 17–23 апр. 2023 г.): в 6 т. / М-во науки и высш. образования Рос. Федерации, Самар. нац. исслед. ун-т им. С. П. Королева (Самар. ун-т), Ин-т систем обраб. изобр. РАН — Фил. Федер. науч.-исслед. центра «Кристаллография и фотоника» Рос. акад. наук. — Самара : Изд-во Самар. ун-та, 2023. — Т. 5: Науки о данных / под ред. Е. В. Гошина. — С. 38–39.
- [6] *Синотова С. Л., Лимановская О. В., Плаксина А. Н., Макутина В. А.* Сравнение эффективности различных методов отбора признаков для решения задачи бинарной классификации предсказания наступления беременности при проведении экстракорпорального оплодотворения // Моделирование, оптимизация и информационные технологии. — 2020. — Т. 8, № 3. — С. 1–16.

-

Автоматизированное установление размеров клеток живой водоросли *Chlorella vulgaris* по цифровым фотографиям

Торопкина Маримьяна Андреевна^{1*}

marimiana.toropkina@gmail.com

Рюмин Александр Георгиевич²

a.ryumin@spbu.ru

Соболев Иван Сергеевич³

i@v4nchelli.ru

¹ Санкт-Петербург, ЗИН РАН

² Санкт-Петербург, СПбГУ

³ Санкт-Петербург

Обнаружение клеток водоросли на цифровой фотографии, установление их размеров (диаметр, площадь) важно для оценки состояния культуры живого организма при воздействии на него внешних факторов (света, питательных веществ, токсикантов и пр.).

В последнее время автоматизированный подсчет клеток находится на пике своей популярности. В первую очередь из-за трудоемкости прямого учета клеток с помощью микроскопа. Использование экспрессных методов существенно ускоряет подсчет, но при этом автоматизированные методы, основанные на обработке косвенных данных (возмущение электрического слоя или удельное сопротивление частиц, содержание хлорофилла и других пигментов) уступают в точности классическому «методу прямого счета» [1]. Существует различное ПО для работы с изображениями, но часто оно перегружено функциями, сложно для освоения, требует снимков высокого качества, может быть достаточно дорогим.

Мы изучали влияние раствора препарата гуминовой кислоты на водоросль *Chlorella vulgaris*. Для оценки действия препарата мы использовали следующие характеристики — оптическую плотность суспензии водоросли, интенсивность флуоресценции, размер и численность клеток живой культуры и др. Так же нами предпринята попытка использовать лазерный анализатор (дифрактометр) размера частиц Shimadzu SALD-2201 для оценки распределения клеток водоросли по размерам. Для установления численности и размеров клеток мы использовали микроскоп с цифровой окуляр-приставкой Levenhuk C510 и счетную камеру Горяева. Количество клеток в 1 мл исследуемой суспензии вычисляли по формуле: $M = a10^3/Sh$, где M — число клеток в 1 мл; a — среднее число клеток в квадрате сетки; S — площадь квадрата сетки, мм²; h — глубина камеры, мм.

В одном нашем эксперименте даже при неполной обработке камеры Горяева мы получаем около 10 тыс. клеток, которые нужно обработать, что сделать вручную очень трудоемко. Т.к. попытки использовать существующее ПО дали неудовлетворительный результат, то подсчет клеток и их размеров на цифровых фотографиях вели с помощью специально написанной на языке Python программы. Для анализа изображений использовали библиотеку компьютерного зрения OpenCV. Для оценки адекватности работы программы проводили

визуальный контроль правильности результатов, т. к. иногда за объект интереса ошибочно принимаются пузырьки газа, различные примеси, скопления клеток.

Работа программы состоит из следующих шагов. В директорию *input* помещаются фотографии для обработки, и для каждого файла выполняются следующие действия. Загружается изображение и увеличивается его яркость через изменение параметра *value* цветовой модели HSV. Изображение преобразуется в оттенки серого. Применяется пороговая обработка для получения бинарного изображения. Производится удаление шумов на изображении. Выполняется дилатация и морфологическое закрытие для определения заднего плана. Выполняется преобразование расстояния для определения переднего плана. Находятся контуры на изображении. Анализируются найденные контуры на условия координат (для исключения краевых областей снимка), площади и радиуса. Контур добавляется в итоговый список в зависимости от выполнения условий. Подсчитывается количество «хороших» и «плохих» контуров. «Плохие» контуры — чаще всего это объекты очень маленького (менее 1 мкм) или большого (более 10 мкм) размера. «Хорошие» контуры отмечаются (обводятся) на изображении. Результирующие фотографии, а также протокол в формате CSV размещаются в каталоге *output*.

В протоколе работы программы указывается имя обработанного файла, номер найденного объекта (клетки), которая отмечена на фотографии. Программа записывает радиус окружности, в которую вписана клетка и площадь, которую она занимает. Площадь клетки — характеристика более надежная, чем радиус для описания размеров, т. к. клетки имеют близкую к сферической, но все же неправильную форму. Измерение площади клетки вручную мы не проводили, но для выделения контура клетки неправильной формы потребуется кратно больше времени, по сравнению с измерением только диаметра.

В поле зрения камеры микроскопа в среднем попадает от 20 до 40 клеток водоросли. На ручное измерение в этом случае уходит примерно 10–20 минут, куда входит поиск клетки, измерение на фотографии диаметра, протоколирование, отметка обнаруженной клетки на фотографии. С помощью программы производится обработка серии фотографий в одной папке (обычно 10 шт.), что занимает около 30 с. За это время программа успевает выполнить описанные выше процедуры.

Качество работы программы в сильной степени зависит от характеристик исходного изображения. Количество объектов, которые диагностируются программой как «плохие» примерно в 7–10 раз больше, чем «хорошие». Так, в одном эксперименте на 804 объекта, которые были опознаны как клетки водоросли, пришлось 6980 объектов, которые не попали в финальный протокол. Программа пропускает единичные клетки водоросли на фотографиях, что очевидно при визуальном контроле, но, по нашему мнению, это существенно не сказывается на итоговом распределении размеров. Использование программы позволило подойти к решению вопроса поиска клеток на изображениях с двух сторон: под-

бора оптимальных условий работы программы и подбора параметров получения изображений с микроскопа.

В дополнение к указанному методу опыты по изучению поведения культуры водоросли при помощи лазерного дифрактометра позволили гораздо быстрее получать данные о размерах клеток, а точнее — их распределении в пробе. Помимо экспрессности получения результатов удастся сразу обработать относительно большой, т. е. более репрезентативный объем пробы. Но вместе с этим применимость его к биологическим объектам требует дополнительного изучения, т. к. одним из явных недостатков является невозможность получить численность клеток в объеме пробы и высокое влияние «загрязнений» пробы разными частицами на результаты.

Проделанные опыты показали, что результаты распределения клеток (по модальному значению) для метода лазерной дифрактометрии дают примерно в 2 раза меньший диаметр клеток водоросли по сравнению с данными прямого счета (около 2 и 4 мкм, соответственно). При это сам характер распределения близок к нормальному в обоих вариантах. Таким образом, использование метода лазерной дифрактометрии требует дополнительного изучения возможности применения к биологическому объекту, и, в данном случае, целесообразно дополнять прямым анализом пробы с помощью камеры Горяева.

По результатам данной работы отмечено следующее:

- использование машинного зрения позволило оценить изменение размера клеток водоросли, существенно сократило время на обработку изображений;
- оно зрения снижает влияние человеческого фактора при измерении объектов, чьи границы на изображении часто не высококонтрастны;
- показано, что распределения размеров клеток, рассчитанные на основе прямого учета и лазерной дифрактометрии хоть и близки, но все же отличаются; это связано как с принципиально разными принципами анализа, так и, возможно, с объемами анализируемых проб;
- использование обоих методов усложняется тем, что регистрируются примеси — чаще всего они сильно отличаются от типичных размеров клеток водоросли, но в некоторых случаях нужны дополнительные исследования проб с визуальным контролем с помощью микроскопа;
- наиболее целесообразным является совмещение двух методов: лазерной дифрактометрии для оценки распределения клеток по размерам и прямого счета для установления количества клеток в суспензии.

Работа выполнена при поддержке РЦ СПбГУ «Культивирование микроорганизмов», проект № 116-7485.

- [1] *Задереев Е. С.* и др. Сравнение экспрессных методов определения численности, размерной структуры и видового состава микроводорослей // Журн. Сиб. федер. ун-та. Биология. — Красноярск: Издательство Сибирского федерального университета, 2021. — Т. 14, № 1. — С. 5–27.

Применение методов машинного обучения для анализа упругих характеристик многослойных композиционных материалов

*Абгарян Каринэ Карленовна*¹

kristal83@mail.ru

*Загордан Надежда Леонидовна*¹

zagordann@gmail.com

Мочалова Юлия Дмитриевна^{1*}

juliamochalova96@gmail.com

¹Москва, Федеральный исследовательский центр «Информатика и управление» РАН

В настоящее время переход к передовым цифровым, интеллектуальным, производственным технологиям, новым материалам и способам их конструирования является одним из важнейших направлений в развитии нашей страны. Модернизация промышленного производства требует создания новых материалов с заданными свойствами с целью их применения в разных областях, включая авиационно-космическую отрасль, микроэлектронику, медицину и др. Разработка информационных систем, базирующихся на применении оригинальных методов программной инженерии, дает возможность автоматизировать процесс научных и инженерных исследований, наиболее эффективно применять методы многомасштабного компьютерного моделирования и машинного обучения для достижения целей опережающего развития в области создания новых материалов.

Предсказание характеристик многослойных полимерных композиционных материалов (ПКМ), а также решение обратной задачи по подбору нового материала с заданными свойствами являются актуальными проблемами при создании изделий из ПКМ, в связи с тем, что ремонт агрегатов в зонах выявленных дефектов, в том числе и с неоригинальными компонентами, требуется конструкциям на протяжении всего жизненного цикла [1]. Однако, задача расчета характеристик композиционного материала в области ремонта является нетривиальной даже при знании характеристик монослоев ПКМ. Применение методов компьютерного моделирования с использованием программной системы, при помощи которой возможно проводить расчеты упругих характеристик многослойных ПКМ, дает возможность автоматизировать научные и инженерные расчеты и с наименьшими временными затратами подбирать оптимальные компоненты и методологии их укладки по заданным характеристикам полимерных композиционных материалов.

Для формирования информационной базы данных был реализован программный комплекс, позволяющий моделировать структурные свойства композиционного материала на основе многомасштабного подхода [2]. В данной работе класс композиционных материалов, моделируемых комплексом [2], был дополнен многослойным материалом, состоящим из разноориентированных монослоев однонаправленного волокна, пропитанного связующим. Для расчета структурных свойств нового материала исходные данные по каждому из монослоев передавались на мезоуровень, где рассчитывались упругие характеристики го-

могенизированного монослоя. Вычисленные свойства монослоев поступали в базу данных свойств материалов программного комплекса [2] и использовались в качестве входных данных для моделирования эффективных упругих характеристик многослойного композиционного материала с разноориентированной структурой армирования. Для их вычисления был программно реализован специализированный расчетный модуль [1].

В целях верификации результатов расчета эффективной упругости многослойного композиционного материала было проведено сравнение экспериментальных данных [3–6] с результатами численного моделирования. Также результаты расчета сравнивались с данными, полученными в ходе численного моделирования упругих характеристик в программном комплексе AnsysCompositePrePost (АСР), позволяющем в явном виде задавать свойства монослоев, структуру укладки многослойных композитных материалов, рассчитывать и визуализировать их упруго-жесткостные свойства.

В настоящее время система машинного обучения использует накопленную с помощью проведенных расчетов базу данных о свойствах композитных материалов с разных масштабных уровней. Применение программных решений, позволяющих автоматизировать процесс расчета упруго-жесткостных свойств многослойных композитных материалов с дальнейшей передачей результатов в базу данных для хранения, дает возможность существенно ускорить процесс накопления многоуровневых данных для их анализа с помощью системы машинного обучения. В дальнейшем планируется решение обратной задачи, в рамках которой для заданных упругих свойств многослойного композита будет возможно системно решать задачу по определению области значений свойств базовых компонентов монослоев, выбору варианта их укладки и других особенностей с целью подбора нового многослойного композитного материала, обладающего требуемыми характеристиками.

Основным ограничением при применении машинного обучения является неполнота валидации полученных результатов. Это связано с отсутствием открытой информации о полном цикле производства композитного материала, включая особенности процесса формовки и характеристик базовых компонентов, используемых для получения композитных материалов.

Работа выполнена в рамках Крупного научного проекта № 075-15-2020-799 Министерства науки и высшего образования Российской Федерации 2022–2023 гг.

- [1] *Абгарян К. К., Загордан Н. Л., Мочалова Ю. Д.* Метод компьютерного моделирования упругих характеристик многослойных композиционных материалов // Системы и средства информатики. — 2023. — Т. 33, № 4. (в печати)
- [2] *Абгарян К. К., Гаврилов Е. С.* Программный комплекс для многомасштабного моделирования структурных свойств композиционных материалов // Информатика и ее применения. — 2022. — Т. 16, № 1. — С. 99–97.

- [3] *Алфутов Н. А., Зиновьев П. А., Попов Г. Б.* Расчет многослойных пластин и оболочек из композиционных материалов // *Машиностроение*, 1984. — С. 264.
- [4] *Peel L. D., Hyer W. M., Shuart M. J.* Compression Failure of Angle-Ply Laminates // *Interim Report*, Blacksburg: Virginia Tech, 1991. — Vol. 16, No. 1. — P. 87–162.
- [5] *Finckenor J. L.* Determination of Significant Composite Processing Factors by Designed Experiment // *MSFS Center Director's Discretionary Fund Final Report*. — Alabama, 2003. — P. 68.
- [6] *Barbero E. J.* *Introduction to Composite Materials Design*. — CRC Press, 2011. — 508 p.

Модели представления экологических объектов по данным гиперспектральной съемки

Мурынин Александр Борисович^{1,2}

Рихтер Андрей Александрович^{2*}

*Козуб Владимир Александрович*²

*Гвоздев Олег Геннадьевич*²

AMurynin@bk.ru

urfin17@yandex.ru

postbox-kozub@ya.ru

olegggs@yandex.ru

¹Москва, ФИЦ ИУ РАН

²Москва, НИИ АЭРОКОСМОС

В работе описывается метод построения спектральных моделей экологических объектов по данным гиперспектральной съемки, выполняемой аэрокосмическими средствами. Спектральная модель описывает области изображения, соответствующие данному классу и содержит статистические характеристики спектров этого класса. Спектральная модель математически описывает класс, воспроизводя всевозможные сигнатуры пикселей этого класса с учетом их статистических характеристик [1].

В целом, спектральная модель строится по результатам обработки некоторой выборки изображений, отобранных по определенным условиям. Только по координатам и яркости отдельной точки нельзя утверждать о принадлежности к определенному классу, поэтому необходимо знать окружение анализируемой точки (координаты и яркости ее окрестности).

Выделение полезной информации и получение моделей классов на гиперспектральных изображениях осуществляется с помощью методов кластеризации, морфологических операций над изображениями, визуального или автоматизированного сопоставления распределения яркости с эталоном. Большое значение имеет разработка также методики определения оптимальных комбинаций спектральных каналов.

Для конкретного случая классификации объектов выбирают ту или иную меру с тем, чтобы с одной стороны уменьшить вычислительную сложность задачи классификации, а с другой — обеспечить наилучшее разбиение на классы. Евклидово расстояние между средними значениями по классам удобно использовать в тех случаях, когда классы достаточно однородны по яркости, т. е. образуют небольшие компактные области в пространстве признаков. Расстояние Махаланобиса определяет точки с одинаковой плотностью вероятностей, которые образуют гиперэллипсоиды в многомерном пространстве. Расстояние Джеффриса-Матуситы пригодно практически для любых ситуаций, однако оценивает соотношения между вероятностями по классам более грубо.

При кластерном анализе учитываются два процесса на многомерном статистическом многообразии — дивергентность (способность исследуемых областей к расхождению на классы) и конвергентность (обратный процесс, способность областей к схождению в один класс). Дивергенция представляет собой функцию, устанавливающую расстояние от одного распределения вероятностей

до другого на статистическом многообразии. В дивергенции учитываются отношения между вероятностями появления признака в каждом из классов, поэтому их используют при классификации статистическими методами — по расстоянию Махаланобиса или максимуму правдоподобия.

Меры разделимости могут быть внутренними и внешними. Первые характеризуют разделяемость области на подобласти, принадлежащие разным классам, вторые характеризуют возможность соединять области по принадлежности их к одному классу.

При построении спектральных моделей могут использоваться алгоритмы кластеризации *k-means* и *isodata*, однако они имеют множество недостатков в отношении кластеризации спектральных сигнатур. Для решения задачи лучше подходят более совершенные методы, такие как метод DBSCAN и его иерархическое расширение HDBSCAN, которое автоматически определяет оптимальное количество кластеров и может работать с кластерами различной плотности.

При обработке большого объема обучающей выборки спектров целесообразно использовать такие механизмы, как тайлирование, агрегирование, уменьшение размерности, кодирование информации.

Значительное внимание следует уделить визуализации построенных спектральных моделей. Спектральный рисунок (изображение) — иллюстрация спектральной модели, получаемая с использованием различных программных средств. Как и спектральная сигнатура, спектральный рисунок может характеризовать пиксель, объект или класс. Также, спектральный рисунок может иллюстрировать и соотношение сигнатур объектов или классов.

- [1] *Мурынин А. Б., Рихтер А. А., Шахраманьян М. А.* Выбор информативных признаков для выделения областей размещения отходов по космическим изображениям высокого пространственного разрешения // Исследование Земли из космоса. — 2019. — № 2. — С. 29–44.

Оценка относительной площади сечения сонной артерии в динамике на ультразвуковых изображениях

Александров Илья Константинович ilya_alexandrov_volgatech@mail.ru

Хафизов Динар Гафиятулович hafizovdg@volgatech.net

Йошкар-Ола, Поволжский Государственный Технологический Университет

Оценка изменения площади поперечного сечения сонной артерии позволяет оценить силу пульсовой волны. Врач, работая с такими данными может сделать выводы о состоянии сонной артерии и предположить развитие возможных проблем в кровообращении, таких как атеросклероз, гипертония, стеноз артерий.

Главная особенность ультразвуковых изображений в наличии сильной зашумленности, поэтому в качестве первичной обработки применяется низкочастотный фильтр:

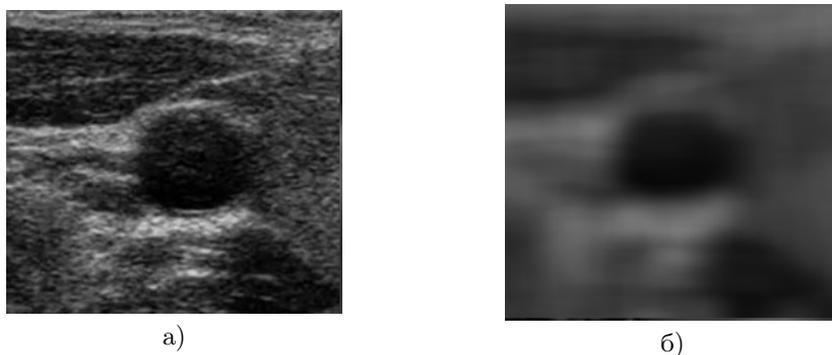


Рис. 1. Изображение ультразвукового исследования сонной артерии: а) исходное; б) после низкочастотной фильтрации

Поиск границ сонной артерии осуществляется через определение точек наибольшего перепада яркости в пределах срезов яркости. На первом этапе поиска рассматривается срез яркостей взятых вдоль луча, формирование которого происходит следующим образом: вручную определяется примерный центр сонной артерии, формируются точки, равноудаленные от центра, и обнаруживаются все позиции между ними и центром по следующей формуле:

$$x_i = x_c + R * \cos\left(\frac{\pi * \alpha * i}{180}\right),$$

$$y_i = y_c + R * \sin\left(\frac{\pi * \alpha * i}{180}\right), \quad i = \overline{0, (360/\alpha) - 1},$$

где R — расстояние от центра до точки, α — угол между соседними точками, x_c и y_c — координаты центра окружности.

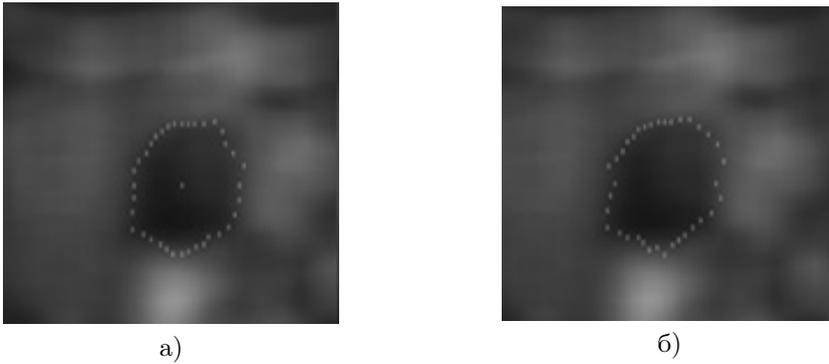


Рис. 2. Изображение с найденными точками перепада яркости: а) без использования критерия Стьюдента; б) с уточненным положением по критерию Стьюдента

Для нахождения перепада в рамках каждого луча используется корреляционная функция данных яркости к функции перепада:

$$K_i = \frac{\sum_{j=0}^{2Nf-1} S_{j+i} * P_j}{\sqrt{\sum_{j=0}^{2Nf-1} S_{j+i}^2 + \sum_{j=0}^{2Nf-1} P_{j+i}^2}},$$

где S — выборка значений яркостного луча, P — выборка значений модели перепада, Nf — размер апертуры сглаживающего фильтра.

На втором этапе поиска необходимо оценить области около найденных точек через определение максимального значения критерия Стьюдента:

$$t = \frac{|M_1 - M_2|}{\sqrt{\frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2}}},$$

где M_1, M_2 — среднее арифметическое, σ_1, σ_2 — среднеквадратичное отклонение, N_1, N_2 — объем выборки.

Для получения примерного контура границы сонной артерии полученные точки объединяются в контур, который сглаживается [1]. Рассчитав площадь контура и повторив все вышеуказанные операции на некоторой последовательности снимков, можно воссоздать колебания сердечной мышцы и отобразить их на графике:

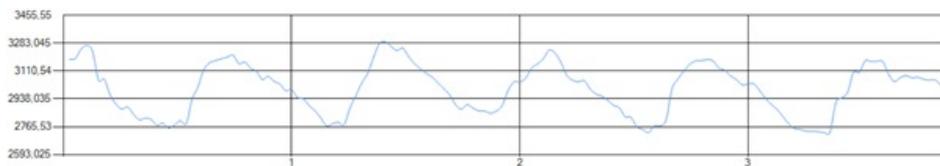


Рис. 3. Изображение графика зависимости изменения площади сечения сосуда от времени

Данные полученные в ходе данной работы визуально соответствуют схематическому отображению пульсовой волны [2]:



Рис. 4. Схематическое изображение пульсовой волны, где: 1 — анакорта; 2 — катарта; 3 — дикротический подъем

- [1] *Фурман Я. О.* Введение в контурный анализ и его приложения к обработке изображений и сигналов / Фурман Я. А., Кревецкий А. В., Передреев А. К., Роженцов А. А., Хафизов Р. Г. и др.; под ред. Я. А. Фурмана. — М.: Физматлит, 2002.
- [2] *Охотников С. А., Ким Е. В., Хафизов Д. Г., Хафизов Р. Г.* Метод выделения контура и оценки площади сечения сосудов шеи при ультразвуковом исследовании // Медицинская техника. — 2022. — № 4. — С. 33–35.

Разработка и применение метода обнаружения фазовых сдвигов в парных биомедицинских сигналах для дифференциальной диагностики болезни Паркинсона и эссенциального тремора

| | |
|---|---|
| <i>Сушкова Ольга Сергеевна</i> ^{1*} | <code>o.sushkova@mail.ru</code> |
| <i>Морозов Алексей Александрович</i> ¹ | <code>morozov@cplire.ru</code> |
| <i>Кершнер Иван Андреевич</i> ¹ | <code>ivan.kershner@gmail.com</code> |
| <i>Хохлова Маргарита Николаевна</i> ¹ | <code>margokhokhlova@gmail.com</code> |
| <i>Габова Александра Васильевна</i> ² | <code>agabova@yandex.ru</code> |
| <i>Карбанов Алексей Вячеславович</i> ³ | <code>doctor.karabanov@mail.ru</code> |
| <i>Чигалейчик Лариса Анатольевна</i> ³ | <code>chigalei4ick.lar@yandex.ru</code> |
| <i>Иллариошкин Сергей Николаевич</i> ³ | <code>snillario@gmail.com</code> |

¹Москва, Институт радиотехники и электроники им. В. А. Котельникова РАН

²Москва, Институт высшей нервной деятельности и нейрофизиологии РАН

³Москва, ФГБНУ «Научный центр неврологии»

Разработан новый метод разведочного анализа данных, направленный на обнаружение закономерностей в парных биомедицинских сигналах. Метод позволяет выявить диапазоны частот, в которых мгновенная фаза исследуемых сигналов демонстрирует нейрофизиологические закономерности, имеющие диагностическую ценность. Кроме того, метод позволяет найти оптимальное соотношение между вкладом фазовых и амплитудных характеристик сигнала при поиске нейрофизиологических закономерностей. Разработанный метод позволил выявить новые нейрофизиологические закономерности в электромиографических сигналах (ЭМГ) мышц-антагонистов пациентов с болезнью Паркинсона (БП) и пациентов с эссенциальным тремором (ЭТ). В частности, обнаружено, что характеристики всплесков в ЭМГ-сигналах мышц-антагонистов демонстрируют одновременно несколько эффектов противоположной направленности, что объясняет противоречивые результаты, опубликованные в нейрофизиологической литературе. Обнаруженные закономерности могут быть использованы для разработки высокочувствительных методов дифференциальной диагностики БП и ЭТ. Разработанный метод является универсальным и может быть применен для анализа других видов биомедицинских сигналов.

Исследование выполнено за счет гранта Российского научного фонда № 22-75-10079, <https://rscf.ru/project/22-75-10079/>.

- [1] *Sushkova O. S., Morozov A. A., Gabova A. V., Karabanov A. V., Illarioshkin S. N.* A Statistical Method for Exploratory Data Analysis Based on 2D and 3D Area Under Curve Diagrams: Parkinson's Disease Investigation // *Sensors*. — Switzerland: MDPI, 2023. — P. 4700.

Сравнительный анализ влияния длины ЭКГ фрагмента на качество распознавания опасных аритмий

*Манило Людмила Алексеевна*¹

lmanilo@yandex.ru

*Немирко Анатолий Павлович*¹

apn-bs@yandex.ru

Евдакова Екатерина Геннадьевна^{1*}

kat355@mail.ru

¹ Санкт-Петербург, Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»

Известно, что риск развития сердечно-сосудистых заболеваний связан с наличием у пациентов сердечных аритмий. Выявление этих нарушений во время ЭКГ-мониторинга позволит предотвратить возможные серьезные последствия. Ключевым моментом в распознавании опасных аритмий является длительность фрагмента ЭКГ, по которому может быть выявлен опасный класс нарушений. В случае кардиологического наблюдения в условиях стационара время выявления опасных нарушений, а также их предвестников является достаточно критичным. Сокращение его позволит применить реанимационные меры как можно скорее, тем самым обеспечивая пациенту максимально возможную помощь в критическом состоянии. Поэтому актуальной задачей становится не только разработка алгоритмов, способных по коротким фрагментам сигналов ЭКГ классифицировать сердечные ритмы, но и задача определения длительности этого сигнала.

В данной работе рассматривались основные нарушения ритма сердца, и были выделены 6 классов [1]. Такая классификация записей ЭКГ позволяет решить сложную задачу, а именно, выявить не только опасную желудочковую фибрилляцию (1-й класс), но и распознавать предвестники аритмии, такие как пируэтная форма желудочковой тахикардии (2-й класс) и желудочковая тахикардия с высокой частотой сердечных сокращений (3-й класс). Остальные классы: 4-й, 5-й и 6-й представляют собой альтернативу для группы опасных аритмий, сгруппированных с учетом возможного риска для жизни пациента.

В исследовании использовались две базы данных: содержащая двухсекундные фрагменты ЭКГ (2с) и новая база данных, содержащая четырехсекундные фрагменты ЭКГ (4с). Принцип создания базы данных был описан в [2]. Классификация проводилась по спектральным характеристикам фрагментов сигнала. Ожидалось, что более длинный фрагмент сигнала будет иметь более устойчивый спектр и тем самым обеспечит лучшую точность классификации [3]. Цель исследования — провести сравнительный анализ качества классификации опасных и неопасных аритмий по двухсекундным и четырехсекундным фрагментам ЭКГ.

Задачи исследования:

1. Классификация на два класса: I — группа опасных нарушений (классы 1, 2, 3); II — группа неопасных нарушений (классы 4, 5, 6).

2. Классификация группы опасных аритмий (классы 1, 2, 3) на два класса: I — группа крайне опасных нарушений (классы 1, 2), требующих применения дефибрилляции; II — класс угрожающих аритмий (класс 3).

Данные представляли собой нормированные спектральные описания ЭКГ фрагментов: а) база 2с, количество гармоник — 15, и база 4с, количество гармоник — 15; шаг 1,0 Гц. б) база 2с, количество гармоник — 15, и база 4с, количество гармоник — 30; суммируется одинаковое число компонент (2 компоненты), но шаг по оси частот разный: 1,0 Гц (для 2с) и 0,5 Гц (для 4с). Т. е. используя 4с, мы получаем более подробное описание спектра.

Методы классификации: метод ближайших соседей (kNN) с оценкой средневзвешенного расстояния; линейный дискриминантный анализ (LDA); метод опорных векторов (SVM) с разными ядрами (линейное, квадратичное, кубическое ядро); случайный лес (RF). Использовался скользящий контроль качества классификации, 9/10 — обучение и 1/10 — тестирование. Деление проводилось пропорционально объемам классов. В качестве основной метрики качества использовалась F1-мера.

Результаты исследования. Задача 1. Данные — случай а. Лучше всего себя показала модель SVM (class weight: None, degree: 2, kernel: 'rbf'). F1-мера для двухсекундных фрагментов — 0,942, для 4с — 0,941. Тест показал:

Таблица 1. Результаты классификации, задача 1, случай а

| t | model | f1 | accuracy | recall | precision | balanced_acc |
|----|-------|-------|----------|--------|-----------|--------------|
| 2с | SVM | 0,921 | 0,91 | 0,932 | 0,922 | 0,91 |
| 4с | SVM | 0,959 | 0,954 | 0,952 | 0,967 | 0,954 |

Данные — случай б. Лучше всего себя показала модель kNN. Параметры лучшей модели для двухсекундных фрагментов: (metric: 'manhattan', n_neighbors: 10, weights: 'distance'). Параметры лучшей модели для четырехсекундных фрагментов: (metric: 'manhattan', n_neighbors: 2, weights: 'distance'). F1-мера для 2с — 0,947, для 4с — 0,942. Тест показал:

Таблица 2. Результаты классификации, задача 1, случай б

| t | model | f1 | accuracy | recall | precision | balanced_acc |
|----|-------|-------|----------|--------|-----------|--------------|
| 2с | KNN | 0,93 | 0,92 | 0,91 | 0,94 | 0,92 |
| 4с | KNN | 0,964 | 0,958 | 0,968 | 0,960 | 0,957 |

Задача 2. Данные — случай а. Лучше всего себя показала модель RF. Параметры лучшей модели для двухсекундных фрагментов: (class_weight: None, n_estimators: 100). Параметры лучшей модели для четырехсекундных фрагментов: (class_weight: None, n_estimators: 150). F1-мера для 2с — 0,865, для 4с — 0,935. Тест показал:

Таблица 3. Результаты классификации, задача 2, случай а

| t | model | f1 | accuracy | recall | precision | balanced_acc |
|----|-------|-------|----------|--------|-----------|--------------|
| 2с | RF | 0,891 | 0,838 | 0,940 | 0,848 | 0,764 |
| 4с | RF | 0,922 | 0,887 | 0,943 | 0,902 | 0,847 |

Данные — случай б. Результаты (лучшая модель и параметры) аналогичны предыдущему пункту. F1-мера для 4с — 0,916. Тест показал:

Таблица 4. Результаты классификации, задача 2, случай б

| t | model | f1 | accuracy | recall | precision | balanced_acc |
|----|-------|-------|----------|--------|-----------|--------------|
| 2с | RF | 0,891 | 0,838 | 0,940 | 0,848 | 0,764 |
| 4с | RF | 0,916 | 0,879 | 0,932 | 0,901 | 0,841 |

Заключение. Модели, обученные на данных 4с, дают значения метрик на тесте лучше, чем модели, обученные на 2с. В задаче 1, классификации опасных/неопасных аритмий, наибольшую точность показали данные 4с, количество гармоник — 30, шаг по частоте — 0,5 Гц. Лучшая модель — kNN ($k = 2$, взвешенное расстояние, расстояние — манхеттен), F1 = 0,964 (против F1 = 0,93 для 2с). В задаче 2, распознавания классов 1 и 2 по отношению к 3-му классу, наилучшее предсказание обеспечивает модель RF ($n = 150$), F1 = 0,922 (против F1 = 0,891 для двухсекундного). Этот результат получен для 4с, количество гармоник — 15, шаг по частоте — 1 Гц. Предположение о том, что четырехсекундный фрагмент сигнала обладает более устойчивым спектром, не было отвергнуто. Прирост в метриках качества составляет около 5 процентов. В то же время следует учитывать, что увеличение длительности рассматриваемого фрагмента в 2 раза соответствующим образом приведет к росту задержки при оказании медицинской помощи в случае детектирования опасных форм нарушений ритма.

Исследование выполнено за счет гранта Российского научного фонда № 23-21-00215, <https://rscf.ru/project/23-21-00215/>.

- [1] *Nemirko A. P., Manilo L. A., Tatarinova A. A., Alekseev B. E., Evdakova E. G.* ECG Fragment Database for the Exploration of Dangerous Arrhythmia // *PhysioNet*. — 2022. <https://doi.org/10.13026/kpfg-xs25>
- [2] *Manilo L., Nemirko A. P., Evdakova E. G., Tatarinova A. A.* ECG Database for Evaluating the Efficiency of Recognizing Dangerous Arrhythmias // *IEEE CSGB*. — 2021. — P. 120–125. <https://ieeexplore.ieee.org/document/9496029>
- [3] *Manilo L., Nemirko A. P., Evdakova E. G.* Recognition of Dangerous Rhythm Disturbances from Short ECG Fragments // *IEEE USBEREIT*. — 2021. — P. 41–44. <https://ieeexplore.ieee.org/document/9455071>

Мета-анализ и метод классификации в условиях отсутствия табличной обучающей выборки

*Кривуля Полина Юрьевна*¹*

polina_krivulya@mail.ru

Сенько Олег Валентинович^{1,2,3}

senkoov@mail.ru

*Воронин Евгений Михайлович*³

voronin@cmd.su

*Плоскирева Антонина Александровна*³

Хлытовка Юлия Николаевна^{3,4}

¹Москва, МГУ

²Москва, ФИЦ ИУ РАН

³Москва, ЦНИИЭ Роспотребнадзора

⁴Москва, ММКЦ Коммунарка

Основным методом решения задач компьютерной диагностики является обучение алгоритма машинного обучения на основе выборки, которая состоит из описания отдельных пациентов. Однако такая информация часто ограничена рамками одного медицинского учреждения и имеет небольшой объем. Для достижения более точных результатов необходимо использовать информацию из многих источников.

Информация из литературных источников представима в виде общего числа пациентов и числа заболевших, имеющих определенный симптом. Основными этапами настоящей работы являются анализ таких данных и поиск методов машинного обучения, решающих задачу диагностики с помощью выборки, полученной из большого количества различных медицинских статей. Целью работы является диагностика вирусных инфекций по набору симптомов.

Рассматривается база данных, собранных из медицинских статей, с характеристиками четырех вирусных инфекций. Для каждой статьи имеется информация о суммарном числе заболевших и о количестве людей, имеющих один из девяти симптомов. В некоторых статьях отсутствует информация о части симптомов. Для различных вирусов количество рассматриваемых статей различно.

Для получения априорной вероятности $P(K_i)$ наличия вируса K_i в определенный сезон использована отдельная таблица, отражающая заболеваемость каждым из вирусов по датам в течение нескольких лет.

Для каждого заболевания определено количество зафиксированных симптомов и суммарное число заболеваний (по всем статьям для рассматриваемого типа вируса). Для i -го типа вируса выполняется подсчет вероятности $p_i(x_j)$ иметь j -й симптом как среднее по статьям. При этом в подсчете такого среднего не учитываются статьи, где данный симптом не упомянут.

В дополнение к проявлению симптомов, рассматривается сезон. В результате применения точного критерия Фишера [1] установлено, что сезон является статистически значимым дополнительным признаком, а также то, что один из рассматриваемых симптомов не является значимым.

Важным вопросом мета-анализа является оценка гетерогенности включаемых исследований. Рассматривается значение I^2 , как наиболее широко применяемый показатель оценки гетерогенности исследований [2]. Анализ показал, что данные, соответствующие одному типу инфекции, являются в значительной степени гетерогенными. Тем не менее, дальнейшее исследование показывает, что такая гетерогенность не мешает работе алгоритма с использованием наивного байесовского классификатора.

Для оценки принадлежности объекта к конкретному типу вируса на основании сезона и набора симптомов применяется алгоритм наивного байесовского классификатора в предположении, что все признаки являются независимыми. Статистически не значимый симптом не рассматривается.

Используем формулу Байеса для наивного байесовского классификатора:

$$P(K_l|x) = \frac{P(K_l) \prod_{i=1}^n p_l(x_i)}{\sum_{j=1}^L P(K_j) \prod_{i=1}^n p_j(x_i)}, \quad (1)$$

где K — тип вируса, n — число рассматриваемых симптомов, $x \in \{0, 1\}^n$ — набор симптомов, $P(K_l)$ — априорная вероятность наличия вируса K_l в текущем сезоне, $p_j(x_i)$ — вероятность наличия (или отсутствия) симптома x_i у вируса j , L — число рассматриваемых вирусов.

У одного из вирусов отсутствует информация о наличии одного из рассматриваемых симптомов. Обозначим этот признак x_i . Вирусы разделяются на две группы: G_1 , состоящую из вирусов K_1, K_2, K_3 , для которых известны все признаки, и G_2 , состоящую из вируса K_4 . Далее вероятность наличия некоторого вируса при заданной симптоматике будем называть вероятностью вируса.

Предлагаемый в данном исследовании алгоритм выглядит следующим образом:

Шаг 1. Вычисление $P(G_1)$ и $P(G_2)$. Рассматривается подмножество признаков X , известных для всех вирусов. По этим признакам с использованием формулы (1) вычисляются вероятности наличия каждого из вирусов. Вероятность $P(G_1)$ вычисляется как сумма полученных вероятностей вирусов K_1, K_2, K_3 ; вероятность $P(G_2)$ — как полученная вероятность вируса K_4 .

Шаг 2. Вычисление $P(K_1), P(K_2), P(K_3)$. Рассматриваются вирусы, для которых известны признаки из подмножества X и признак x_i . С помощью формулы (1) с использованием всех признаков вычисляются вероятности $P(K_1), P(K_2), P(K_3)$. Полученные вероятности домножаются на значение $P(G_1)$, рассчитанное на шаге 1.

Шаг 3. Значение $P(K_4)$. Для вируса K_4 с отсутствующим симптомом x_i остается вероятность $P(G_2)$, вычисленная на шаге 1 с использованием формулы 1: $P(K_4) = P(G_2)$.

Качество построенного алгоритма оценивается с помощью сгенерированной выборки и анализа ROC-кривой. Полученное качество по метрике ROC-AUC превышает 0,87, а для отдельных типов вирусов достигает 0,98.

Работа алгоритма рассматривается и для исследования гетерогенности данных. Для каждого вируса случайным образом выбирается медицинская статья и исключается, при этом для остальных вирусов статистики остаются неизменными. В результате данного эксперимента получено, что качество работы алгоритма не ухудшается, что говорит о том, что различия между вирусами более существенны, чем различия между статьями для одного вируса.

Таким образом, построен алгоритм, позволяющий решать задачу классификации типа вируса по данным, собранным из медицинских статей. Преимуществом метода является его актуальность в случаях, когда наличие полноценной обучающей выборки невозможно.

- [1] Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Часть I. Критерии типа хи-квадрат. // Рекомендации по стандартизации. — М.: Изд-во стандартов, 2002. — 87 с.
- [2] Белов Ю. В., Салагаев Г. И., Лысенко А. В., Леднев П. В. Мета-анализ в медицине // Хирургия. Журнал им. Н. И. Пирогова. — М.: Медиа Сфера, 2018. — № 3. — С. 4–15.

Linking EEG Channels with Spatial Metadata to Impact the Interpretability and Generalization Ability of Graph Neural Networks in Neurophysiological Applications

*Sidorov Leonid*¹★

leon.sidorov@gmail.com

Maysuradze Archil^{1,2}

maysuradze@cs.msu.ru

¹Moscow, Lomonosov Moscow State University, Faculty of Computational Mathematics and Cybernetics

²Moscow, Lomonosov Moscow State University, MSU Institute for Artificial Intelligence

We present a novel technique and a specialized machine learning model architecture designed to assist researchers across diverse domains in automatically and explicitly identifying functional patterns in multivariate time series derived from a series of experiments. Our approach formalizes the original task within the realm of machine learning, eliminating the need for domain-specific exploration. The proposed architecture leverages information about the recording device, encoding the initial electrode positions as a graph for transmission to the corresponding model. The efficacy of our technique has been demonstrated in neurophysiology, specifically for data where the P300 pattern is known to exist. Additionally, we applied the same technique to emotion recognition using the SEED dataset, achieving notable results. Our experiments highlight the substantial impact of the coordinate graph employed in the algorithm. As part of future research, we aim to extend our proposed technique to other domains, such as data from sensors on production lines or banking transactions.

The first dataset under consideration was collected using the *oddball paradigm* method, an experimental design widely used in psychological and neurophysiological research. This paradigm involves presenting subjects with sequences of repetitive stimuli, occasionally interrupted by non-standard exposure. The subject's reaction to this "odd" stimulus is recorded and analyzed, following the process detailed in [1]. The dataset, recorded using a character matrix and brain-computer interface (BCI), is part of the third BCI competition [2] and offers diverse experimental results with various ML models. While the duration of each fragment varies, the presence of the *jwave P300j* pattern can be reliably detected within one second.

For additional architecture validation, we employed the SEED dataset [3] for an emotion recognition task, utilizing the same raw EEG signal input data. The SEED dataset, designed for discrete emotion models, comprises data from 15 subjects participating in three experiments spaced approximately one week apart. Each experiment involves subjects viewing 15 film clips, with EEG acquisition using a cap with 62 electrodes based on the international standard 10-20 system.

In contrast to recent papers neglecting the physical structure of BCI interfaces, our approach emphasizes creating a dense graph representing the actual shape of the recording device. This graph is constructed based on electrode neighborhood, where edges connect vertices if and only if they are neighbors in terms of coordinates.

Two common solutions for graph construction were considered: *Delaunay triangulation* [4], producing a planar graph consisting of triangles, and the classical *k nearest neighbors* algorithm, employing Euclidean distance with a selected parameter value of $k = 9$.

The proposed model family comprises three types of modules: a spatial processing block, a temporal processing block, and a predictor block (a multi-layer perceptron). The research focuses on the capabilities of Graph Convolutional Networks (GCN), making the architectures relatively simple. These architectures can represent a wide range of existing EEG processing neural networks.

The base model consists of 1×1 convolution as spacial processing block and one-dimensional convolution as temporal processing block. The predictor block is identical for all approaches and consists in multiplying the object vector by the matrix of the parameters being trained. Such an operation is defined only for vectors, but it will be applied at the last stage of object transformation, when both the graph in our problem and the matrix in the original one will be reduced to a one-dimensional embedding vector.

An alternative approach involves adding a graph convolution layer from GCN before the 1×1 convolution in the spatial processing block. This addition aims to potentially enhance model performance by allowing the capture of more complex patterns of brain activity.

Using the BCI competition dataset, the graph-based architecture not only improved model quality but also provided more interpretable results. As stated in [5], the most informative electrodes in the recording device are Fz, Cz, P3, P4, Po7, Po8, Pz and Oz. In fig. 1 we can notice that the basic architecture in our approach focused its attention on only three electrodes Pz, Po7 and Po8. While the graph neural network in fig. 2 was able to recognize all of the above electrodes, except Fz and Cz. The SEED dataset's weighted electrodes also revealed the model's deep understanding, aligning with sources of emotional impulses. The proposed neural network architecture, with its alternating application of filters to channels and time intervals, demonstrated the ability to independently detect the P300 wave and classify emotions within the SEED dataset.

The introduction of a graph filter to the original architecture enhanced the model's generalization capability, as evidenced by spatial filter weights. The incorporation of Graph Convolutional Networks (GCN) further improved the model's performance on EEG data by capturing intricate patterns of brain activity. The model excelled in solving binary classification and character classification tasks on the BCI Competition dataset and achieved competitive accuracy levels on the SEED dataset compared to current best approaches, as reported in [6].

A key finding from this research underscores the significant influence of the graph type on the GCN model's recognition capabilities. Specifically, for problem classes like emotion recognition, the planar Delaunay graph consistently outperformed the dense KNN graph.

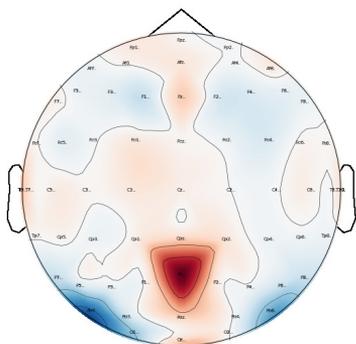


Fig. 1. Baseline model at characters recognition.

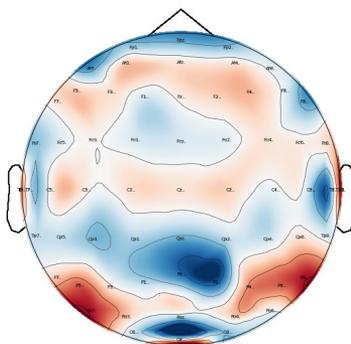


Fig. 2. Graph model at characters recognition.

The research is supported by Scientific and educational school of Moscow State University “Brain, cognitive systems, artificial intelligence”, research work of Moscow State University 5.1.21.

- [1] *Squires NK.* Two Varieties of Long-Latency Positive Waves Evoked by Unpredictable Auditory Stimuli in Man // *Electroencephalography and Clinical Neurophysiology.* — Amsterdam: Elsevier, 1975. — P. 387–401.
- [2] *Schalk G.* BCI2000: A General-Purpose Brain-Computer Interface (BCI) System // *IEEE Transactions on Biomedical Engineering.* — Piscataway: IEEE, 2004. — P. 1034–1043.
- [3] *Zheng WL.* Investigating Critical Frequency Bands and Channels for EEG-Based Emotion Recognition with Deep Neural Networks // *IEEE Transactions on Autonomous Mental Development.* — Piscataway: IEEE, 2015. — P. 162–175.
- [4] *Delaunay B.* Sur la Sphere Vide // *Otdelenie Matematicheskikh i Estestvennykh Nauk,* Moscow: Izv. Akad. Nauk SSSR, 1934. — P. 1–2.
- [5] *Nomenclature SE.* American Electroencephalographic Society Guidelines for Standard Electrode Position Nomenclature // *J. Clinical Neurophysiology.* — Philadelphia: Lippincott Williams & Wilkins, 1991. — P. 200–202.
- [6] *Gong L.* EEG Emotion Recognition Using Attention-Based Convolutional Transformer Neural Network // *Biomedical Signal Processing and Control.* — Amsterdam: Elsevier, 2023. — P. 104835.

Идентификация пациентов по рентгенологическим изображениям

Новик Владимир Петрович¹*

novikvp1@zdrav.mos.ru

Матвеев Иван Алексеевич²

matveev@frccsc.ru

Арзамасов Кирилл Михайлович¹

arzamasovkm@zdrav.mos.ru

¹Москва, ГБУЗ «НПКЦ ДиТ ДЗМ»

²Москва, ФИЦ ИУ РАН

Сканирование головы человека с помощью компьютерного томографа (КТ) является важным методом диагностики. В информационной системе каждому изображению сопоставлены идентификационные данные пациента. Из-за ошибок оператора исследование может быть приписано другому пациенту, что может привести к серьезным ошибкам при постановке диагноза. Цель данной работы — создание метода идентификации пациента по рентгенологическому изображению для обнаружения таких ошибок, в том числе для очищения (дубликации) баз данных.

Материалом исследования служат данные Единого радиологического информационного сервиса Единой медицинской информационно-аналитической системы города Москва (ЕРИС ЕМИАС). Использованы 150 КТ изображений, содержащих анатомическую область головы, взятые от 50 пациентов.

Два этапа работы метода — создание шаблона из исходных данных и вычисление расстояния между двумя шаблонами. Создание шаблона (вектора признаков облака точек) $T(i)$ для каждого исследования $i \in [1, N]$ включает в себя последовательные шаги:

- вычисление координат краниометрических точек;
- нормализация положения головы;
- сегментирование костных тканей черепа;
- создание трехмерного облака точек;
- разреженное описание облака точек.

Вычисление расстояния $D(i, j) = D(T(i), T(j))$ между облаками точек изображений i и j :

- точное совмещение облаков точек итеративным алгоритмом [1];
- вычисление расхождения поверхностей двух шаблонов $D(i, j)$ [2].

Значение $D(i, j)$ далее сравнивается с порогом для принятия решения. Величина порога устанавливается экспериментально.

Созданный метод формирования и сравнения шаблонов позволил достичь полной разделимости на данных, построенных по 150 образцам 50 пациентов.

[1] Besl P. J., McKay N. D. A Method for Registration of 3-D Shapes // IEEE Trans. PAMI. — 1992. — Vol. 14, No. 2. — P. 239–256.

[2] Дышкант Н. Ф., Местецкий Л. М. Сравнение 3D портретов при распознавании лиц // Материалы ММРО-2007. — С. 314–316.

Содержание

| | |
|---|----|
| Интеллектуальный анализ данных | 6 |
| <i>Рязанов В. В., Виноградов А. П.</i> Подходы к выбору содержательных аспектов при анализе сложных выборок | 6 |
| <i>Дюкова Е. В., Дюкова А. П.</i> О метрических свойствах множеств решений некоторых задач анализа целочисленных данных | 9 |
| <i>Матковский В. А., Козлова М. Г., Лукьяненко В. А.</i> Исследование процессов торможения пассажирских высокоскоростных поездов | 12 |
| <i>Кузин К. А., Газарян В. А., Чулмичков А. И., Шапкина Н. Е., Голубцов П. В.</i> Применение критериев проверки статистических гипотез для анализа параметров атмосферы | 15 |
| Машинное обучение | 18 |
| <i>Торшин И. Ю.</i> Топологическая теория для анализа сложных данных, включающих разнородные признаковые описания | 18 |
| <i>Сенько О. В., Докукин А. А.</i> Новый двухуровневый метод машинного обучения для оценивания вещественных характеристик объектов | 20 |
| <i>Дюкова Е. В., Масляков Г. О., Янаков Д. С.</i> ДСМ-классификатор над произведением частичных порядков | 23 |
| <i>Ерохин М. В., Прошин В. И., Ерохин В. И.</i> Частично формализованный способ отбора значимых признаков для линейной бинарной классификации (на примере диагностики гипоксически-ишемических поражений головного мозга пациентов детского возраста по данным МР-морфометрии) | 26 |
| <i>Шибзухов З. М.</i> Об одном робастном варианте метода главных компонент | 29 |
| <i>Двоенко С. Д.</i> Технология кластеризации парных сравнений | 31 |
| <i>Ланге М. М., Ланге А. М.</i> Нижняя граница погрешности для оценок параметров распределений при заданном количестве информации | 32 |

| | |
|--|-----------|
| <i>Веприков А. С., Афанасьев А. П., Хританков А. С.</i> Математическая модель эффекта обратной связи в системах искусственного интеллекта | 35 |
| <i>Авдюшина А. Е., Авдюшина Е. В., Ручкин К. А.</i> Графовый метод полуконтролируемого обучения в задаче классификации периодических решений динамических систем в режиме реального времени | 38 |
| <i>Гагарин Ю. Е., Гагарина С. Н.</i> Оценивание ошибки классификации в последовательном алгоритме распознавания объектов с учетом погрешностей исходных данных . . . | 40 |
| <i>Максимова А. Ю.</i> Применение генеративно-сопоставительного подхода с частичным привлечением учителя для решения задачи классификации с отрицательным классом | 42 |
| <i>Неделько В. М.</i> Сравнение оценок out-of-bag и скользящего экзамена | 45 |
| <i>Копылов А. В., Середин О. С.</i> Новая F-мера качества бинарных классификаторов | 48 |
| Прогнозирование и рекомендательные системы | 51 |
| <i>Тырсин А. Н., Кащеев С. Е.</i> Многомерная модель динамического материального баланса в задаче мониторинга системы заводнения нефтяных месторождений | 51 |
| <i>Алешновский В. С., Макеев Д. А., Газарян В. А., Чулмичков А. И., Шапкина Н. Е., Голубцов П. В.</i> Восстановление пропусков данных в многомерных временных рядах геофизических параметров методами статистики и математического моделирования. | 54 |
| <i>Шекунов М. А.</i> Интеграция рекомендательных алгоритмов и решения задачи оптимизации для улучшения опыта туристов | 57 |
| <i>Кравцова О. А.</i> Прогнозирование временных рядов с добавлением информации о стационарности в функционал потерь | 58 |
| <i>Старожилец В. М., Чехович Ю. В.</i> Экспериментальная проверка вычислительной эффективности одной модели транспортных потоков | 60 |
| Аналитика больших данных | 63 |

| | |
|--|-----------|
| <i>Аксенович Т. В., Селиванов В. Н.</i> | |
| Классификация геомагнитно-индуцированных токов в нейтрали автотрансформатора с применением машинного обучения | 63 |
| <i>Мусабаев Рус., Мусабаев Рав., Кульдеев Н.</i> | |
| Новый параллельный алгоритм кластеризации больших данных Big-means с конкурентной стохастической оптимизацией размера выборки | 65 |
| Нейронные сети и глубокое обучение | 68 |
| <i>Михайлапов Д. И., Бериков В. Б., Алямкин С. А.</i> | |
| Дистилляция знаний для увеличения устойчивости сверточных сетей к сдвигам изображений | 68 |
| <i>Баринов Н. А., Филатов А. В.</i> | |
| Влияние дистилляции моделей при обучении на дистиллированных данных | 71 |
| <i>Захаров А. А., Шамшин М. Н., Жизняков А. Л.</i> | |
| Разработка метода визуального внимания на основе нейронных сетей на графах | 74 |
| <i>Середин О. С., Ляхов Д. В., Ломов Н. А.</i> | |
| Применение нейронных сетей для определения меры отражательной симметрии фигуры на бинарных растровых изображениях | 75 |
| <i>Якушева С. Ф., Хританков А. С.</i> | |
| Применение причинно-следственных графов при тестировании инвариантами | 78 |
| <i>Григорьев А. Д., Гнеушев А. Н.</i> | |
| Обобщенная ортогональная регуляризация сверточных нейронных сетей | 81 |
| <i>Шэнь С., Цзинь Ц., У Г., Лукьянов В. В.</i> | |
| Применение технологии динамического отслеживания и семантической сегментации для визуальной навигационной системы мобильных беспилотных объектов | 82 |
| <i>Афанасьев Г. И., Местецкий Л. М.</i> | |
| Распознавание рукописного текста на основе реконструкции траектории пера по цифровому изображению | 85 |
| Методы оптимизации для интеллектуального анализа данных | 87 |
| <i>Малозёмов В. Н., Соловьёва Н. А., Тамасян Г. Ш.</i> | |
| MDM-алгоритм и задача Сильвестра | 87 |
| <i>Котина Е. Д., Овсянников Д. А., Харченко Д. С.</i> | |
| Оптимизационный подход к проблеме определения поля скоростей в задачах обработки изображений | 90 |

| | |
|--|------------|
| <i>Лемтюжникова Д. В., Шушко Н. И., Красоткин С. А., Барашов Е. Б.</i> Метод попарных сравнений для задачи коммивояжера. Специальный случай выпуклой оболочки | 93 |
| <i>Sorokovikov P.</i> Hybrid Algorithms Based on Nature-Inspired and Local Descent Methods for Solving Multiextremal Optimization Problems | 95 |
| <i>Володина Е. А., Ерофеева В. А.</i> Модификация протокола локального голосования для оптимизации балансировки загрузки ресурсов распределенной системы | 96 |
| <i>Лемтюжникова Д. В., Посыпкин М. А., Барашов Е. Б.</i> Метаэвристические алгоритмы для задачи упаковки в контейнеры | 99 |
| <i>Макаров О. О., Козлова М. Г., Лукьяненко В. А.</i> Близкие задачи маршрутизации типа многих коммивояжеров | 102 |
| <i>Горнов А. Ю., Аникин А. С., Зароднюк Т. С.</i> Алгоритмы предоптимизационного и постоптимизационного анализа для невыпуклых функций | 105 |
| Вычислительная сложность и приближенные методы | 106 |
| <i>Карацуба Е. А.</i> О сложности вычисления элементарных алгебраических функций | 106 |
| <i>Бобков А. В., Ду К.</i> Повышение скорости работы сверточных сетей с использованием алгоритма Винограда | 108 |
| Обработка и анализ изображений, компьютерное зрение | 111 |
| <i>Кабанова В. В., Логунова О. С., Наркевич М. Ю., Корниенко В. Д.</i> Обнаружение и оценка нарушений целостности межпанельных швов зданий на изображении | 111 |
| <i>Козлов В. Н.</i> Цифровая подпись на основе инвариантов геометрических преобразований | 114 |
| <i>Харинов М. В.</i> Генерация оптимальных приближений цветного изображения | 115 |
| <i>Местецкий Л. М.</i> Штриховая сегментация рукописного текста | 118 |
| <i>Бобков А. В., Дай И.</i> 3D-реконструкция поверхности Марса на основе SFM | 120 |

Мостовой Д. Н., Алчинов А. И.

Некоторые вопросы снижения функционала летательного аппарата, оснащенного корреляционно-экстремальной навигационной системой в условиях возможных помех 123

Цзинь Ц., Шэнь С., У Г., Веселовская О. А.

Метод ассоциации данных семантического SLAM на основе моделей вероятностного вывода 125

Сурков Е. Э., Середин О. С., Копылов А. В.

Применение беспризнакового подхода к задаче распознавания активностей человека на основе базисной совокупности скелетных моделей . . 128

Середин О. С., Ломов Н. А., Ляхов Д. В., Митюгов Н. С., Кушнир О. А., Копылов А. В.

Дескриптор формы бинарной фигуры на основе профиля вращения . . 131

Шэнь С., У Г., Цзинь Ц., Лукьянов В. В.

Применение механизмов внимания и увеличения набора данных для задач обнаружения мелких объектов 134

Мурашов Д. М., Мурашова Е. Д.

Ускоренные алгоритмы формирования сегментов из суперпикселей . . 137

Зырянов К. С., Орешкин А. И., Чулчиков А. И.

Оптимизация разрешения в сканирующем туннельном микроскопе: тематические подходы и экспериментальные результаты 140

Феоктистов Д. Д., Местецкий Л. М.

Мера сходства элементов штрихового представления рукописного текста на основе Фурье-дескриптора 143

Веселов Д. И., Фан А. Ч., Гаврилин П. С., Андриянов Н. А.

Обнаружение лиц с помощью гибридных детекторов 146

Обработка и анализ сигналов 148

Синицын И. Н., Синицын В. И., Корепанов Э. Р., Конашенкова Т. Д.

Бейесов синтез стохастической системы методом вейвлет канонических разложений 148

Сенько О. В., Чучупал В. Я., Кузнецова А. В., Окулов Н. В., Докучкин А. А.

Интерпретируемость обучения в системе обработки сигналов 151

Роров Д., Maysuradze А.

A New Method for Tackling Overlapping Speech to Improve Speaker Diarization and Target Speaker Voice Activity Detection 154

Копылова О. А.

Анализ и распознавание техногенных шумов с использованием нейросетевого подхода 157

| | |
|---|-----|
| Информационный поиск и анализ текстов | 159 |
| <i>Михайлов Д. В., Емельянов Г. М.</i> | |
| Трансформерные модели BERT, взаимное сходство смыслов коротких текстов и их ранжирование по близости эталону | 159 |
| <i>Кудинов И. Д.</i> | |
| Решение задачи кореференции с использованием токенизации по под- словам | 162 |
| <i>Садиев С. А., Рогов А. А.</i> | |
| Анализ информационной согласованности области «Объяснимый ис- кусственный интеллект» | 163 |
| <i>Грицай Г. М., Грабовой А. В., Кильдяков А. С., Чехович Ю. В.</i> | |
| Генерация и поиск искусственно сгенерированных текстовых фрагмен- тов в домене научных работ | 166 |
| <i>Гаврилов М. С.</i> | |
| Анализ диалогов для определения эмоций персонажей в художествен- ных текстах | 168 |
| <i>Koranichuk I., Chashchin A., Ochneva I., Ogaltsov A., Grabovoy A., Kildyakov A., Chekhovich Y.</i> | |
| Structure Extractor: Multilingual Extraction of Sections from Scientific Documents | 170 |
| <i>Лемтюжникова Д. В., Блашкун М. Р., Латипов А. Р., Кирьянов П. А.</i> | |
| Разработка методики для автоматического выделения структуры в текстах научных публикаций с использованием методов машинного обучения и эвристических алгоритмов | 172 |
| Индустриальные приложения науки о данных | 175 |
| <i>Копкин Е. В., Иваню А. Ю.</i> | |
| Структурно-стохастические вычислительные модели контроля техни- ческого состояния космических средств | 175 |
| <i>Исмагилов Р. Р., Андреев С. Ю.</i> | |
| Решение задачи гендерной и расовой классификации по биометриче- скому вектору признаков лица человека | 178 |
| <i>Астафьев А. В., Астафьева О. С.</i> | |
| Позиционирование человека без нательных датчиков внутри помеще- ний на основе информации о состоянии канала связи | 181 |
| <i>Редько Д. А., Семенов М. Е., Турбов А. О.</i> | |
| Разработка модели компьютерного зрения для автоматизации контроля действий оператора на линии инспекции | 183 |

| | |
|--|-----|
| <i>Ходатаева Т. С., Аверина А. И., Гурьянов А. Е., Русинов Г. А., Каширин Н. В.</i> Описание подходов к проектированию работающего макета автоматической оптической инспекции | 186 |
| <i>Бобров Е. А.</i> Применение машинного обучения к задаче выбора сигнально-кодовой конструкции | 189 |
| <i>Кружалов А. С., Филиппович А. Ю.</i> Особенности применения сверточных нейронных сетей для распознавания аневризм сосудов головного мозга | 192 |
| <i>Витвицкий А. А., Потанов С. А., Гаро В. В.</i> Трекинг и идентификация игроков в командных видах спорта | 195 |
| <i>Малмина А. Э., Яворский Р. Э.</i> Интеллектуальный программно-аппаратный комплекс с оцифровкой жестов для дошкольного образования | 198 |
| <i>Тевляшов Г. К., Русаков К. Д., Мещеряков Р. В.</i> Алгоритм детекции и сегментации осетровых | 201 |
| <i>Ильина С. А., Семенов М. Е.</i> Автоматизация процесса подготовки исходных данных для гидрогеологического моделирования | 204 |
| <i>Коньшин А. М., Витвицкий А. А.</i> Детекция игровых событий в командных видах спорта на основе видео | 207 |
| <i>Курбаков М. Ю., Сулимова В. В., Середин О. С., Копылов А. В.</i> Формирование системы осмысленных признаков и ее применение для определения упорядоченности расположения наночастиц на снимках с электронного микроскопа | 210 |
| <i>Дементьев В. Е., Ким А. Л., Коновалов М. А., Андриянов Н. А.</i> Модифицированный сегментатор дефектов на бетонных изделиях | 212 |
| <i>У Г., Шэнь С., Цзинь Ц., Серебряный В. В.</i> Сегментация изображений рук и инструментов в коллаборативной робототехнической ячейке | 215 |
| <i>Хотимин М. И.</i> Технология интеллектуального автоматизированного отбора информативных признаков натуральных гиперспектральных изображений | 218 |
| <i>Торопкина М. А., Рюмин А. Г., Соколов И. С.</i> Автоматизированное установление размеров клеток живой водоросли <i>Chlorella vulgaris</i> по цифровым фотографиям | 221 |
| <i>Абгарян К. К., Загордан Н. Л., Мочалова Ю. Д.</i> Применение методов машинного обучения для анализа упругих характеристик многослойных композиционных материалов | 224 |

| | |
|--|------------|
| <i>Мурынин А. Б., Рихтер А. А., Козуб В. А., Гвоздев О. Г.</i> | |
| Модели представления экологических объектов по данным гиперспектральной съемки | 227 |
| Анализ биомедицинских данных, биоинформатика | 229 |
| <i>Александров И. К., Хафизов Д. Г.</i> | |
| Оценка относительной площади сечения сонной артерии в динамике на ультразвуковых изображениях | 229 |
| <i>Сушкова О. С., Морозов А. А., Кершнер И. А., Хохлова М. Н., Габова А. В., Карabanов А. В., Чигалейчик Л. А., Иллариошкин С. Н.</i> | |
| Разработка и применение метода обнаружения фазовых сдвигов в парных биомедицинских сигналах для дифференциальной диагностики болезни Паркинсона и эссенциального тремора | 232 |
| <i>Манило Л. А., Немирко А. П., Евдакова Е. Г.</i> | |
| Сравнительный анализ влияния длины ЭКГ фрагмента на качество распознавания опасных аритмий | 233 |
| <i>Кривуля П. Ю., Сенько О. В., Воронин Е. М., Плоскирева А. А., Хлытовка Ю. Н.</i> | |
| Мета-анализ и метод классификации в условиях отсутствия табличной обучающей выборки | 236 |
| <i>Sidorov L., Maysuradze A.</i> | |
| Linking EEG Channels with Spatial Metadata to Impact the Interpretability and Generalization Ability of Graph Neural Networks in Neurophysiological Applications | 239 |
| <i>Новик В. П., Матвеев И. А., Арзамасов К. М.</i> | |
| Идентификация пациентов по рентгенологическим изображениям | 242 |
| Содержание | 243 |
| Авторский указатель | 251 |

Авторский указатель

- А**
Абгарян К. К., 224
Авдюшина А. Е., 38
Авдюшина Е. В., 38
Аверина А. И., 186
Аксенович Т. В., 63
Александров И. К., 229
Алешновский В. С., 54
Алчинов А. И., 123
Алямкин С. А., 68
Андреев С. Ю., 178
Андрянов Н. А., 146, 212
Аникин А. С., 105
Арзамасов К. М., 242
Астафьев А. В., 181
Астафьева О. С., 181
Афанасьев А. П., 35
Афанасьев Г. И., 85
- Б**
Барашов Е. Б., 93, 99
Баринов Н. А., 71
Бериков В. Б., 68
Блашкун М. Р., 172
Бобков А. В., 108, 120
Бобров Е. А., 189
- В**
Веприков А. С., 35
Веселов Д. И., 146
Веселовская О. А., 125
Виноградов А. П., 6
Витвицкий А. А., 195, 207
Володина Е. А., 96
Воронин Е. М., 236
- Г**
Габова А. В., 232
Гаврилин П. С., 146
Гаврилов М. С., 168
- Гагарин Ю. Е., 40
Гагарина С. Н., 40
Газарян В. А., 15, 54
Гаро В. В., 195
Гвоздев О. Г., 227
Гнеушев А. Н., 81
Голубцов П. В., 15, 54
Горнов А. Ю., 105
Грабовой А. В., 166
Григорьев А. Д., 81
Грицай Г. М., 166
Гурьянов А. Е., 186
- Д**
Дай И., 120
Двоенко С. Д., 31
Дементьев В. Е., 212
Докукин А. А., 20, 151
Ду К., 108
Дюкова А. П., 9
Дюкова Е. В., 9, 23
- Е**
Евдакова Е. Г., 233
Емельянов Г. М., 159
Ерофеева В. А., 96
Ерохин В. И., 26
Ерохин М. В., 26
- Ж**
Жизняков А. Л., 74
- З**
Загордан Н. Л., 224
Зароднюк Т. С., 105
Захаров А. А., 74
Зырянов К. С., 140
- И**
Иваню А. Ю., 175

Иллариошкин С. Н., 232
 Ильина С. А., 204
 Исмагилов Р. Р., 178

К

Кабанова В. В., 111
 Карабанов А. В., 232
 Карацуба Е. А., 106
 Каширин Н. В., 186
 Кащеев С. Е., 51
 Кершнер И. А., 232
 Кильдяков А. С., 166
 Ким А. Л., 212
 Кирьянов П. А., 172
 Козлов В. Н., 114
 Козлова М. Г., 12, 102
 Козуб В. А., 227
 Конашенкова Т. Д., 148
 Коновалов М. А., 212
 Коньшин А. М., 207
 Копкин Е. В., 175
 Копылов А. В., 48, 128, 131, 210
 Копылова О. А., 157
 Корепанов Э. Р., 148
 Корниенко В. Д., 111
 Котина Е. Д., 90
 Кравцова О. А., 58
 Красоткин С. А., 93
 Кривуля П. Ю., 236
 Кружалов А. С., 192
 Кудинов И. Д., 162
 Кузин К. А., 15
 Кузнецова А. В., 151
 Кульдеев Н., 65
 Курбаков М. Ю., 210
 Кушнир О. А., 131

Л

Ланге А. М., 32
 Ланге М. М., 32
 Латишов А. Р., 172
 Лемтюжникова Д. В., 93, 99, 172

Логунова О. С., 111
 Ломов Н. А., 75, 131
 Лукьяненко В. А., 12, 102
 Лукьянов В. В., 82, 134
 Ляхов Д. В., 75, 131

М

Макаров О. О., 102
 Макеев Д. А., 54
 Максимова А. Ю., 42
 Малинина А. Э., 198
 Малозёмов В. Н., 87
 Манило Л. А., 233
 Масляков Г. О., 23
 Матвеев И. А., 242
 Матковский В. А., 12
 Местецкий Л. М., 85, 118, 143
 Мещеряков Р. В., 201
 Митюгов Н. С., 131
 Михайлапов Д. И., 68
 Михайлов Д. В., 159
 Морозов А. А., 232
 Мосговой Д. Н., 123
 Мочалова Ю. Д., 224
 Мурашов Д. М., 137
 Мурашова Е. Д., 137
 Мурьнин А. Б., 227
 Мусабаев Рав., 65
 Мусабаев Рус., 65

Н

Наркевич М. Ю., 111
 Неделько В. М., 45
 Немирко А. П., 233
 Новик В. П., 242

О

Овсянников Д. А., 90
 Окулов Н. В., 151
 Орешкин А. И., 140

П

Плоскирева А. А., 236

Посыпкин М. А., 99
 Потапов С. А., 195
 Прошин В. И., 26

Р

Редько Д. А., 183
 Рихтер А. А., 227
 Рогов А. А., 163
 Русаков К. Д., 201
 Русинов Г. А., 186
 Ручкин К. А., 38
 Рюмин А. Г., 221
 Рязанов В. В., 6

С

Садиех С. А., 163
 Селиванов В. Н., 63
 Семенов М. Е., 183, 204
 Сенько О. В., 20, 151, 236
 Серебрянный В. В., 215
 Середин О. С., 48, 75, 128, 131, 210
 Сеницын В. И., 148
 Сеницын И. Н., 148
 Соболев И. С., 221
 Соловьёва Н. А., 87
 Старожилец В. М., 60
 Сулимова В. В., 210
 Сурков Е. Э., 128
 Сушкова О. С., 232

Т

Тамасян Г. Ш., 87
 Тевяшов Г. К., 201
 Торопкина М. А., 221
 Торшин И. Ю., 18
 Турбов А. О., 183
 Тырсин А. Н., 51

У

У Г., 82, 125, 134, 215

Ф

Фан А. Ч., 146

Феоктистов Д. Д., 143
 Филатов А. В., 71
 Филиппович А. Ю., 192

Х

Харинов М. В., 115
 Харченко Д. С., 90
 Хафизов Д. Г., 229
 Хлышовка Ю. Н., 236
 Ходатаева Т. С., 186
 Хотилин М. И., 218
 Хохлова М. Н., 232
 Хританков А. С., 35, 78

Ц

Цзинь Ц., 82, 125, 134, 215

Ч

Чехович Ю. В., 60, 166
 Чигалейчик Л. А., 232
 Чуличков А. И., 15, 54, 140
 Чучупал В. Я., 151

Ш

Шамшин М. Н., 74
 Шалкина Н. Е., 15, 54
 Шекунов М. А., 57
 Шибзухов З. М., 29
 Шушко Н. И., 93
 Шэнь С., 82, 125, 134, 215

Я

Яворский Р. Э., 198
 Якушева С. Ф., 78
 Янаков Д. С., 23

С

Chashchin A., 170
 Chekhovich Y., 170

G

Grabovoy A., 170

K

Kildyakov A., 170
Kopanichuk I., 170

M

Maysuradze A., 154, 239

O

Ochneva I., 170
Ogaltsov A., 170

P

Popov D., 154

S

Sidorov L., 239
Sorokovikov P., 95

Научное издание

МАТЕМАТИЧЕСКИЕ МЕТОДЫ
РАСПОЗНАВАНИЯ ОБРАЗОВ

Тезисы докладов
21-й Всероссийской конференции с международным участием

Подписано в печать 29.11.2023 г.

Формат 60×84 1/8

Усл.-печ. л. 29,76. Уч.-изд. л. 12,9.

Тираж 100 экз

Издатель — Российская академия наук

Публикуется в авторской редакции

Отпечатано в экспериментальной цифровой типографии РАН

Издается по решению Научно-издательского совета
Российской академии наук (НИСО РАН) от 31.03.2023 г.
и распространяется бесплатно