



Факультет компьютерных наук
Образовательная программа «Программная инженерия»
Выпускная квалификационная работа

**Библиотека для решения задач компьютерного зрения
методами топологического анализа данных**

**A Library to Solve Various Computer Vision Problems Using
Topological Data Analysis**

Исследовательская ВКР

Выполнил студент группы БПИ213
образовательной программы
09.03.04 «Программная инженерия»
Абрамов Александр Сергеевич

Руководитель:
Доцент департамента математики
факультета экономических наук НИУ ВШЭ,
кандидат физико-математических наук
Михайлец Екатерина Викторовна

Соруководитель:
Руководитель группы разработки контроля
качества ООО «Яндекс.Такси Технологии»
Кубарев Александр Сергеевич

Москва, 2024



ОСНОВНЫЕ ПОНЯТИЯ, ОПРЕДЕЛЕНИЯ, ТЕРМИНЫ

Генерация – одна из задач комп. зрения, заключающаяся в создании уникальных изображений по их описаниям.

Глубинное обучение – набор методов машинного обучения, основанных на применении нейронных сетей.

Детекция – одна из задач комп. зрения, заключающаяся в поиске объектов на изображениях.

Классификация – одна из задач комп. зрения, в которой требуется установить принадлежность объекта к одному или нескольким заранее установленным классам.

Машинное обучение – класс алгоритмов, способных обучаться некоторой задаче при решении сходных задач.

Метрика качества – численное значение, показывающее, насколько полученные в результате применения алгоритма машинного обучения предсказания соответствуют правильным ответам.

Признак – численное значение, описывающее некоторый объект.

Признаковое описание – набор (вектор) признаков некоторого объекта.

Пространство признаков – векторное (линейное) пространство, в котором лежат векторы признаков.

Распознавание лиц – одна из задач комп. зрения, направленная на разделение изображений разных людей.

Сегментация – одна из задач комп. зрения, заключающаяся в поиске границ отдельных объектов на изображениях.

Фотоконтроль – система для дистанционной проверки документов и контроля качества перевозок партнёрами сервиса “Яндекс.Такси”

ПРЕДМЕТНАЯ ОБЛАСТЬ КОМПЬЮТЕРНОЕ ЗРЕНИЕ

Компьютерное зрение - набор задач, направленных на получение, обработку, анализ и понимание изображений автоматизированной системой.

Традиционный подход к решению – **глубинное обучение и нейросетевые технологии**.

Классификация

Генерация

Сегментация

Детекция

Чтение текста

Распознавание
лиц

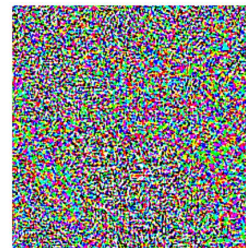


x

“panda”

57.7% confidence

$+ .007 \times$



$\text{sign}(\nabla_x J(\theta, x, y))$

“nematode”

8.2% confidence

$=$



$x +$

$\epsilon \text{sign}(\nabla_x J(\theta, x, y))$

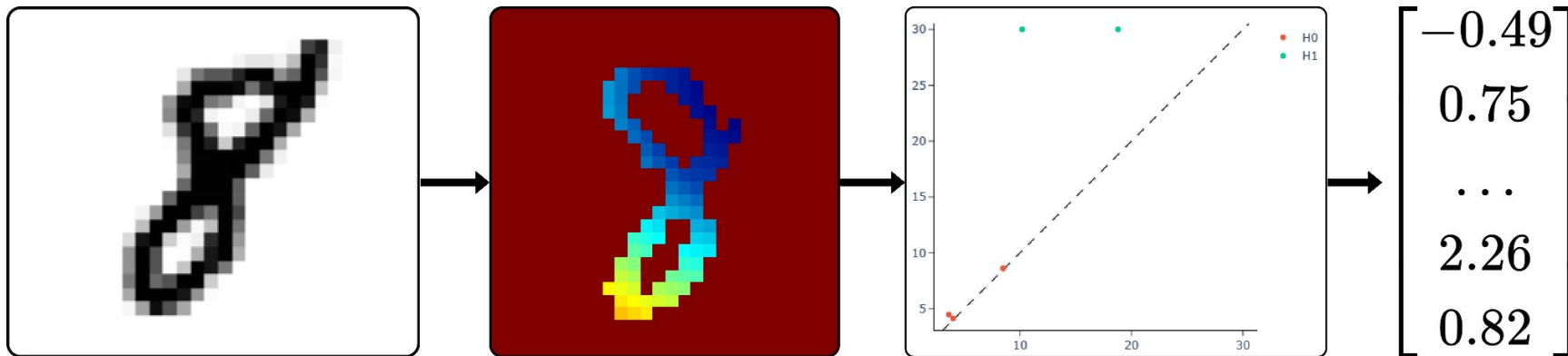
“gibbon”

99.3 % confidence

[3] I. J. Goodfellow et al. Explaining and Harnessing Adversarial Examples. doi: 10.48550/arXiv.1412.6572

ПРЕДМЕТНАЯ ОБЛАСТЬ ТОПОЛОГИЧЕСКИЙ АНАЛИЗ ДАННЫХ

Топологический анализ данных – современный подход к анализу данных различного рода, основанный на понижении размерности без значительной потери информации путём изучения пространственных характеристик данных (формы, структуры и др.) методами алгебраической топологии.



[5] G. Tauzin et al. giotto-tda: A Topological Data Analysis Toolkit for Machine Learning and Data Exploration. doi: 10.48550/arXiv.2004.02551



ПРОБЛЕМА И АКТУАЛЬНОСТЬ РАБОТЫ



Наука о формах объектов должна “работать” особенно хорошо в задачах, связанных с анализом и преобразованием изображений



Не требует большого количества вычислительных ресурсов для обучения



Плохая интерпретируемость нейронных сетей нередко блокирует их использование на практике; топологический анализ данных не обладает этой проблемой



Представлено много исследований отдельных подходов на конкретных данных



Отсутствие единого, обобщённого подхода к обработке изображений методами топологического анализа данных



ТЕКУЩЕЕ СОСТОЯНИЕ ПРОБЛЕМЫ

На практике создание специализированных алгоритмов сложно, долго и дорого; необходима простая в использовании модель, позволяющая быстро сделать вывод о целесообразности дальнейших исследований и примерно оценить качество возможных результатов.

Требуется универсальная модель, позволяющая оценить принципиальную применимость топологического анализа данных к решению задачи

При разработке моделей машинного обучения необходимы агрегированные знания о наиболее эффективных типичных подходах и их особенностях

Высокую теоретическую и практическую ценность имеют любые результаты при решении нетипичных задач: распознавание лиц, чтение текста и др.



ОБЪЕКТ, ПРЕДМЕТ И МЕТОДЫ ИССЛЕДОВАНИЙ

Объект исследования – методы топологического анализа данных, применяемые при работе с изображениями для решения различных задач компьютерного зрения.

Предмет исследования – оценка универсальности методов топологического анализа данных и получаемых таким образом признаков для создания базовой модели, показывающей приемлемое для дальнейших исследований качество на большом ряде задач.

Методы исследования:

1. Анализ поведения существующих подходов на различных открытых наборах данных
2. Реализация и комбинирование существующих методов для создания более универсальных
3. Отбор признаков и применение алгоритмов понижения размерности данных
4. Выбор наиболее эффективных подходов и разработка модульной, гибкой библиотеки
5. Сравнение качества полученной реализации с существующими моделями на реальных данных

АНАЛИЗ СУЩЕСТВУЮЩИХ РЕШЕНИЙ

➤ Классификация

- Рукописных цифр методами TDA [2]
- Комбинирование TDA и глубинного обучения [a33]
- Анализ цветowych каналов фотографий птиц [a7]
- Родинок для диагностики раковых заболеваний [a1]
- Изучение срезов тканей человека (WSI) [a24]

➤ Инструменты

- Регуляризация моделей машинного обучения [a9]
- Сравнение топологий пространств данных [a5]

➤ Понижение размерности

- UMAP: Аппроксимация многообразия [a27]
- Топологический автокодировщик [a39]

➤ Сегментация

- Общий подход к многомерным изображениям [a4]
- Оптимизация запросов к Segment Anything [a16]
- UNET с топологическими признаками [a35]
- Обработка, изменение изображений; задача детекции [a41]

ЦЕЛЬ И ЗАДАЧИ ВКР

Цель работы:

Разработка универсального алгоритма признакового описания изображений методами топологического анализа данных, который будет применим к большинству задач компьютерного зрения в качестве базовой модели с приемлемым для дальнейших исследований качеством.



Изучить и реализовать существующие методы



Собрать открытые наборы размеченных данных



Выбрать наиболее эффективные подходы



Разработать универсальный алгоритм



Спроектировать и реализовать библиотеку



Собрать реальные данные системы “Фотоконтроль”



Сравнить качество подхода с существующими



Оценить целесообразность внедрения модели

НАБОРЫ ДАННЫХ

MNIST [a23]



Labeled faces in the wild [a18]



MIDV-500 [a3]



ImageNet [a11]



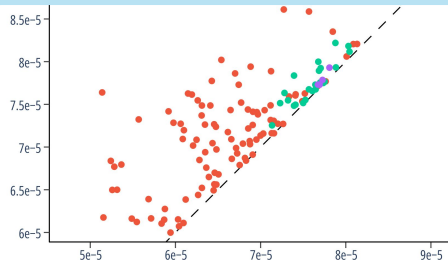
CelebFaces [a26]



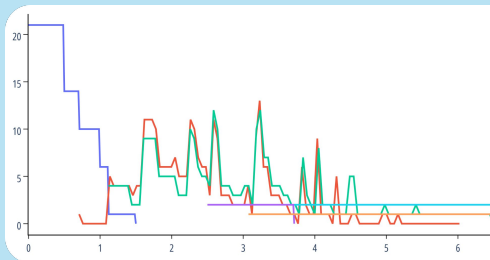
Фотоконтроль



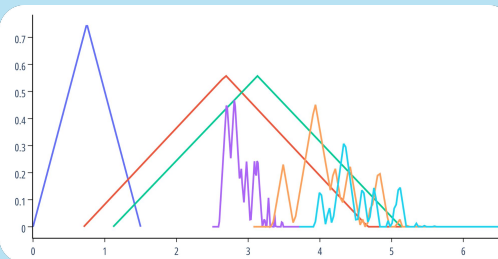
Диаграмма устойчивости



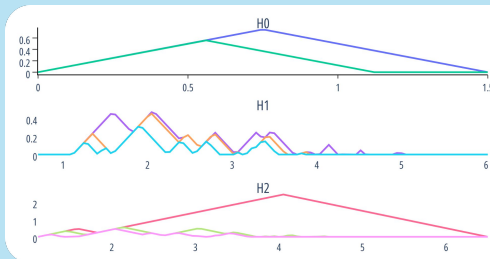
Числа и кривые Бетти



Силуэт устойчивости



Ландшафт устойчивости



Другие алгоритмы

- Симплициальный комплекс
- Кубический комплекс
- Кубические гомологии
- Амплитуды устойчивости
- Энтропия устойчивости
- Метод главных компонент
- Аппроксимация и проекция многообразия

Языки программирования и другие инструменты



Топологический анализ данных



[5] G. Tauzin et al. – 2021. –
doi: 10.48550/arXiv.2004.02551



[a12] Dlotko, P. – 2020. –
<https://gudhi.inria.fr/doc/3.1.1>

Машинное и глубинное обучение

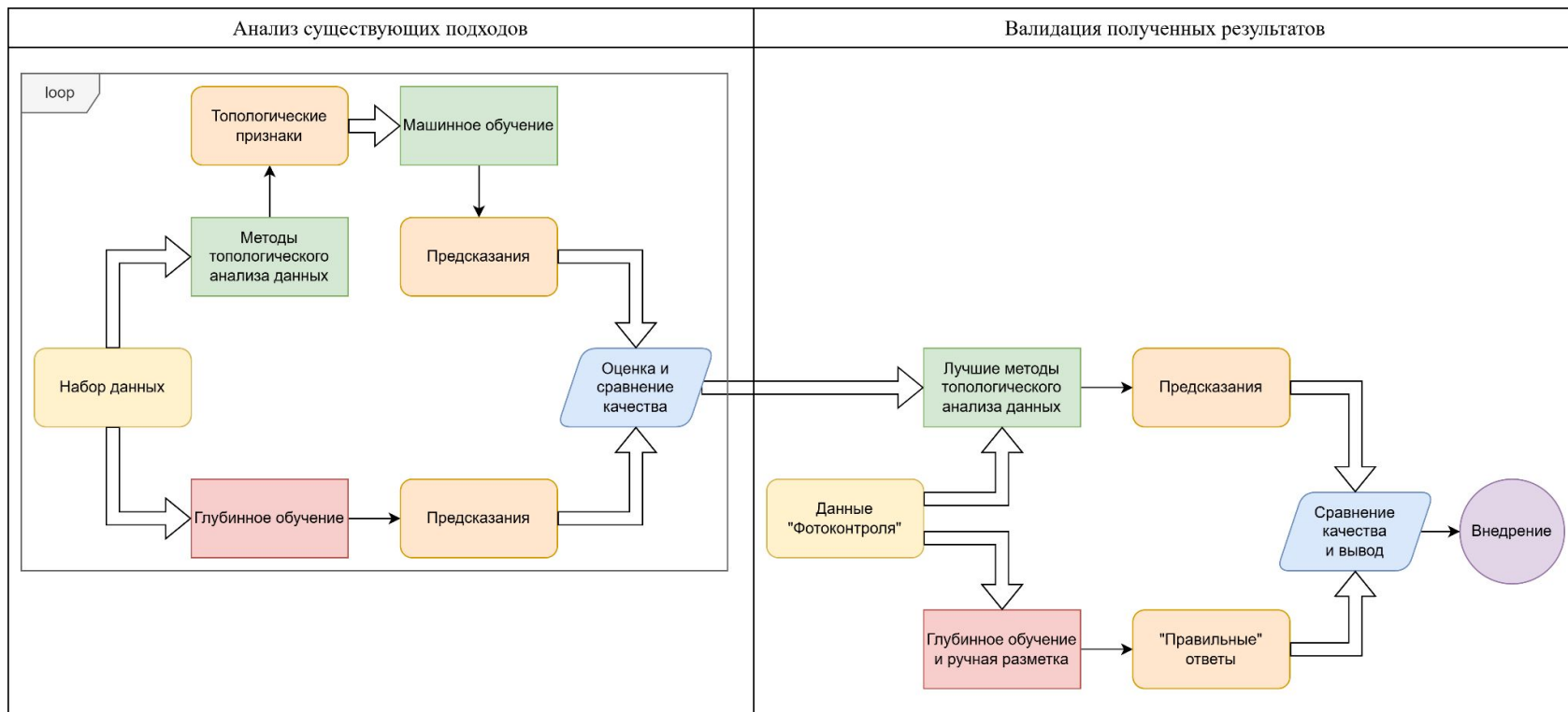


[a30] A. Paszke et al. – 2019. –
doi: 10.48550/arXiv.1912.01703



[a32] F. Pedregosa et al. – 2011. –
<https://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html>

СХЕМА ЭКСПЕРИМЕНТА



ОЖИДАЕМЫЕ РЕЗУЛЬТАТЫ

➤ Результаты



Библиотека с программной реализацией алгоритма и примерами его использования

Вывод о применимости топологического анализа данных к задачам комп. зрения

➤ Научная новизна



Попытка обобщения топологического подхода применительно к изображениям

Основа для “переиспользования” знаний от одних данных с другими данными

➤ Теоретическая значимость



Возможность эффективно проектировать модели, основанные на тополог. анализе

Создание базовой реализации модели, демонстрирующей приемлемое качество

➤ Практическая значимость



Возможность внедрения в проверку изображений “Фотоконтроля” в Яндекс.Такси

Снижение нагрузки и затрат на ручную разметку данных с помощью ассессоров



РЕЗУЛЬТАТЫ, ПОПАДАЮЩИЕ ПОД NDA



Набор реальных данных пользователей, собранных из системы “Фотоконтроль” в “Яндекс.Такси” на заключительном этапе проекта



При этом реализация библиотеки и все метрики качества разработанного алгоритма будут опубликованы в полном объёме



СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1) Cohen-Steiner, D. Stability of persistence diagrams [Электронный ресурс] / D. Cohen-Steiner, H. Edelsbrunner, J. Harer. – Discrete & Computational Geometry: электрон. журн. – vol. 37. – С. 103-120. – Springer, 12 Dec 2006. – URL: <https://doi.org/10.1007/s00454-006-1276-5>. (дата обращения: 09.11.2024).
- 2) Garin, A. A Topological "Reading" Lesson: Classification of MNIST using TDA [Электронный ресурс] / A. Garin, G. Tauzin. – arXiv, 22 Oct 2019. – URL: <https://doi.org/10.48550/arXiv.1910.08345>. (дата обращения: 09.11.2024).
- 3) Goodfellow, I. J. Explaining and Harnessing Adversarial Examples [Электронный ресурс] / I. J. Goodfellow, J. Shlens, C. Szegedy. – arXiv, 20 Mar 2015. – URL: <https://doi.org/10.48550/arXiv.1412.6572>. (дата обращения: 09.11.2024).
- 4) Kaczynski, T. Computational Homology [Электронный ресурс] / T. Kaczynski, K. Mischaikow, M. Mrozek. – Springer, 2004. – URL: <https://doi.org/10.1007/b97315>. (дата обращения: 09.11.2024).
- 5) Tauzin, G. giotto-tda: A Topological Data Analysis Toolkit for Machine Learning and Data Exploration [Электронный ресурс] / G. Tauzin, U. Lupo, L. Tunstall, J. B. Pérez, M. Caorsi, W. Reise, A. Medina-Mardones, A. Dassatti, K. Hess. – arXiv, 5 Mar 2021. – URL: <https://doi.org/10.48550/arXiv.2004.02551>. (дата обращения: 09.11.2024).
- 6) Zomorodian, A. Computing persistent homology [Электронный ресурс] / A. Zomorodian, G. Carlsson. – Discrete & Computational Geometry: электрон. журн. – vol. 33. – С. 249-274. – Springer, 19 Nov 2004. – URL: <https://doi.org/10.1007/s00454-004-1146-y>. (дата обращения: 09.11.2024).



Факультет компьютерных наук
Образовательная программа «Программная инженерия»
Выпускная квалификационная работа

Библиотека для решения задач компьютерного зрения методами топологического анализа данных

A Library to Solve Various Computer Vision Problems Using Topological Data Analysis

Исследовательская ВКР

Выполнил студент группы БПИ213
образовательной программы
09.03.04 «Программная инженерия»
Абрамов Александр Сергеевич

Руководитель:
Доцент департамента математики
факультета экономических наук НИУ ВШЭ,
кандидат физико-математических наук
Михайлец Екатерина Викторовна

Соруководитель:
Руководитель группы разработки контроля
качества ООО «Яндекс.Такси Технологии»
Кубарев Александр Сергеевич