

# DSPA:

## Вопросы применения цифровой обработки сигналов

№2

2024



## СОДЕРЖАНИЕ

<b>Гадасин Д.В., Бессолицын А.Д., Гадасин Д.Д. ОЦЕНКА КАЧЕСТВА ДАННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ</b>	<b>4</b>
<b>Голованова А.С. АНАЛИЗ МЕТОДОВ ТРЕХМЕРНОЙ РЕКОНСТРУКЦИИ ОБЪЕКТОВ, НЕПОДВИЖНЫХ ОТНОСИТЕЛЬНО СКАНИРУЮЩИХ УСТРОЙСТВ</b>	<b>13</b>
<b>Макаров А.А., Большаков А.С. МЕНЕДЖЕР ПАРОЛЕЙ С ГЕОЛОКАЦИЕЙ</b>	<b>21</b>
<b>Медведев Е.Ю., Воронова Л.И., РАСПОЗНАНИЕ ТВЁРДЫХ КОММУНАЛЬНЫХ ОТХОДОВ С ИСПОЛЬЗОВАНИЕМ АЛГОРИТМА RT-DETR</b>	<b>28</b>
<b>Палагушин А.Д., Воронов В.И. ПРОЕКТИРОВАНИЕ ПОДСИСТЕМЫ ВНЕШНЕГО СИТУАТИВНОГО УПРАВЛЕНИЯ РОБОТОТЕХНИЧЕСКИМ КОМПЛЕКСОМ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА</b>	<b>36</b>
<b>Синева И.С., Головченко В.Е. ПРИМЕНЕНИЕ МЕТОДОВ МНОГОМЕРНОГО СТАТИСТИЧЕСКОГО АНАЛИЗА И NLP ДЛЯ КЛАССИФИКАЦИИ НАУЧНЫХ ПУБЛИКАЦИЙ</b>	<b>44</b>
<b>Сотников А.Ю. ПОВЫШЕНИЕ ТОЧНОСТИ РАСПОЗНАВАНИЯ НА СОБСТВЕННОМ НАБОРЕ ДАННЫХ ПРИ ОБУЧЕНИИ С ИСПОЛЬЗОВАНИЕМ МОДЕЛИ YOLOv8 NANO</b>	<b>52</b>
<b>Шматко М.А., Василевич А.Д., Завидова Е.А., Завада А.Р., Осипов А.В. РАЗРАБОТКА АЛГОРИТМА ПОИСКА АССОЦИАТИВНЫХ ПРАВИЛ СЛАБОСТРУКТУРИРОВАННЫХ ДАННЫХ</b>	<b>70</b>

## ОЦЕНКА КАЧЕСТВА ДАННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ

**Гадасин Денис Вадимович,**

*МТУСИ, доцент кафедры СИТиС, к.т.н., Москва, Россия*

[dengadiplom@mail.ru](mailto:dengadiplom@mail.ru)

**Бессолицын Антон Дмитриевич,**

*МТУСИ, Москва, Россия*

[besstoni2100@gmail.com](mailto:besstoni2100@gmail.com)

**Гадасин Даниил Денисович,**

*МТУСИ, Москва, Россия*

[gadasin115@gmail.com](mailto:gadasin115@gmail.com)

### **Аннотация**

*С увеличением объемов данных и их важности для принятия стратегических решений в различных областях, вопрос об эффективном управлении качеством данных становится критическим. После того, как были определены требования к качеству данных необходимо определить модель управления данными. Для данной модели детерминируются критерии оценки и выявляются группы пользователей для которых проводится упорядочивание критериев исходя из оценок важности. Итогом является параметр соответствующей оценки качества данных системы.*

**Ключевые слова:** *Инженерия данных, качество данных, точность, полнота, согласованность, достоверность, своевременность, качество данных.*

### **Введение**

В современном информационном обществе объемы данных неуклонно растут, предоставляя огромный потенциал для принятия взвешенных решений в бизнесе, науке и государственном управлении. Однако, с увеличением объемов данных возрастает их разнообразие, усложняется передача и обработка [15-34], и, следовательно, растет риск низкого качества информации. На начальном этапе работы с данными необходимо провести их обработку в соответствии с решаемыми задачами [1-4]. После проведения обработки на первый план выходит управление качеством данных, что становится приоритетной задачей для организаций, стремящихся максимизировать свою эффективность и конкурентоспособность.

Целью данной работы – провести анализ существующих методов управления качеством данных и разработать модель, которая позволит эффективно оценивать, управлять и улучшать качество данных в организациях. В работе будут рассмотрены ключевые аспекты управления качеством данных, включая сбор, хранение, обработку и использование информации. В результате исследования предлагается модель управления качеством данных, интегрированная в организационные процессы, которая может служить эффективным инструментом для совершенствования информационной инфраструктуры и повышения достоверности данных.

В работе «Определение системы оценки качества данных в процессах инженерии данных», было определено, что качественные данные можно обозначить, как точные, полные, последовательные, надежные и своевременные, которые способствуют получению значимой и достоверной информации. Была получена система для оценки качества данных. Данная система представлена в формуле 1.

В работе «Определение системы оценки качества данных в процессах инженерии данных», было определено, что качественные данные можно обозначить, как точные, полные, последовательные, надежные и своевременные, которые способствуют получению значимой и достоверной информации. Была получена система для оценки качества данных. Данная система представлена в формуле 1.

$$Q_{\text{данных}} = \begin{cases} \text{Точность} \\ \text{Полнота} \\ \text{Согласованность} \\ \text{Достоверность} \\ \text{Своевременность} \end{cases} = \begin{cases} \frac{T_{\text{ип}}}{T_{\text{ип}} + T_{\text{ип}}} \\ \frac{T_{\text{ип}}}{T_{\text{ип}} + T_{\text{ло}}} \\ \cos_{\text{sim}(A,B)} = \frac{(A \cdot B)}{\|A\| \cdot \|B\|} \\ \frac{(M_1 - M_2)}{\sqrt{(m_1^2 + m_2^2)}} \\ t_{\text{реагирования}} \cdot v_{\text{загрузки}} \end{cases} \quad (1)$$

### Модель управление качеством данных

После того, как были определены требования к качеству данных необходимо определить процесс управления данными. Начинается процесс разработки модели, в которой пользователи будут управлять качеством данных в зависимости от потребностей, например от количества информации, которая содержится в единичном фрагменте [5-8]. Другими словами, необходима пользовательская подсистема для управления качеством данных на основе пяти параметров (точность, полнота, последовательность, достоверность и своевременность). Опишем ее следующим образом:

1. Создание модели:

- Для каждого из пяти критериев (точность, полнота, последовательность, достоверность и своевременность) создается тумблер, который позволяет задать величину в диапазоне от 0 до 1.
- Значение 0 на тумблере означает, что соответствующий критерий не будет учитываться при оценке качества данных.
- Значение 1 на тумблере означает максимальное учет данного критерия при оценке качества данных.

2. Оценка качества данных:

- Для каждого из пяти критериев применяется соответствующая модель для оценки качества данных.
- Значения на тумблерах используются для взвешивания значимости каждого критерия.
- Результаты оценки каждого критерия приводятся к шкале от 0 до 1.

3. Общая оценка качества данных:

- Вычисляется средневзвешенное значение результатов оценки каждого критерия.
- Если хотя бы один из критериев имеет значение 0, общая оценка качества данных будет 0.
- В противном случае, общая оценка качества данных будет отлична от 0 и отражать соответствие данных требуемым стандартам.

4. Управление качеством данных:

- Пользователь может изменять значения на тумблерах для каждого из критериев в зависимости от требований и приоритетов.
- При изменении значений на тумблерах, система пересчитывает общую оценку качества данных и предоставляет информацию о текущем уровне качества данных.

Такая подсистема позволит пользователям гибко управлять качеством данных, опираясь на заданные ими требования и приоритеты для каждого из пяти критериев.

Представим критерии качественных данных в виде схемы изображенной на рисунке 1.

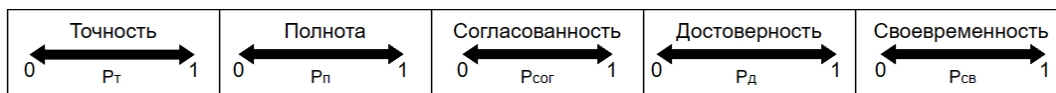
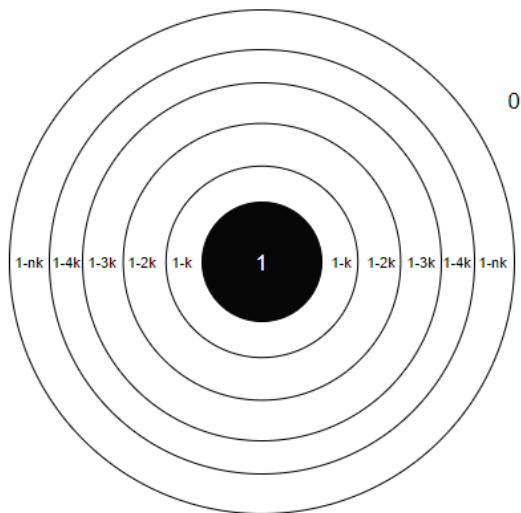


Рис. 1. Схема модели управления качеством данных

У каждого из критериев в рассматриваемой модели предусмотрен "тумблер", который позволяет установить величину в диапазоне от 0 до 1. Такой подход позволяет более гибко настраивать параметры и учитывать различные аспекты при оценке качества данных. Оценка качественных данных в данной модели вычисляется по формуле:

$$P = P_T \cdot P_{II} \cdot P_{COG} \cdot P_D \cdot P_{CB} \quad (2)$$

Оценка может иметь любое значение от 0 до 1 включительно. Если хотя бы один из критериев принимает значение 0, оценка качества данных считается нулевой. Это свидетельствует о том, что даже недостаточный учет любого критерия может привести к отсутствию возможности достичь высокого качества данных. Стоит подчеркнуть, что каждый критерий оказывает влияние на значения других критериев, что делает систему взаимосвязанной и требующей комплексного подхода к оценке.



**Рис. 2.** Модель для управления точностью

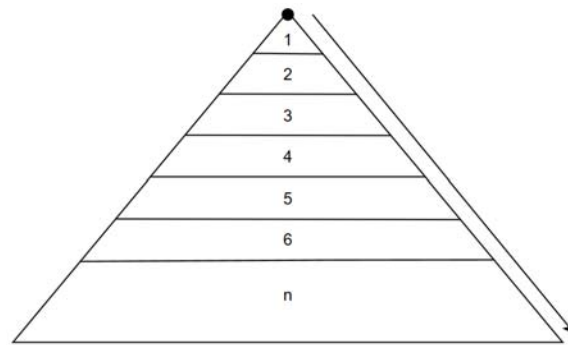
Для более полного понимания данной модели рассмотрим изменение критериев на конкретных иллюстративных моделях. Это позволит нам выявить, как изменение величины каждого критерия влияет на общую оценку качества данных.

Для иллюстрации изменения точности воспользуемся аналогией с мишенью, рассматривая ее как наилучшее представление точности в контексте качественных данных. Представим, что точность в центре мишени равна 1, что соответствует идеально точным данным. Отдаляясь от центра, точность падает, а за границей мишени оценивается как 0. Такая оценка характерна для кластеров, где центр мишени соответствует аттрактору, а чем дальше от него, тем «притяжение», оценка становится меньше [9-12].

Для того чтобы быть ближе к центру, необходимо затрачивать большее количество ресурсов, а также необходимо учитывать внешние факторы, которые могли повлиять на точность данных. Поэтому критерий точности всегда стремится к 1, но достигнуть такой величины, в реальных условиях, крайне проблематично.

На рисунке 2 изображена визуальная модель, предназначенная для изменения точности данных.

Рассмотрим следующий критерий данных, а именно – полноту. Полнота данных напрямую зависит от количества данных, чем больше критерий полноты, тем больше данных. Полноту данных можно представить в виде пирамидальной модели. Для начала необходимо представить всю предметную область в виде точки. От точки вниз пойдут уровни. На каждом уровне будет 3 подуровня, на верхнем подуровне будут располагаться сущности, ниже их атрибуты и в самом низу значения атрибутов. Чтобы опуститься на 1 уровень ниже представим атрибуты нашей сущности, как отдельные сущности. Тогда на уровне ниже также будут 3 подуровня, но сущности 2 уровня будут отображением атрибутов 1 уровня. Схема данной модели изображена на рисунке 2.



**Рис. 3.** Пирамидальная модель критерия «Полнота»

Для примера рассмотрим предметную область животное. На 1 уровне сущность будет животное, атрибуты – это класс животного, кличка, хозяин и т.д. На 2 уровне сущностью станет класс животного, а атрибутами подкласс, вид, подвид и т.д. На 3 уровне сущностью уже будет вид животного, а атрибутами его рост, вес, конечности и части тела. Потом сущностью может стать наличие конечности, а атрибутами кости, кожа и так может продолжаться пока мы не дойдем до атомарных сущностей. В этом и заключается полнота нашей предметной области, спускаясь на уровень ниже величина нашей полноты постоянно растет. И тут возникает вопрос, а какой-же все-таки уровень необходим? Рассмотрим на данном примере, что данные о животных необходимы для продажи костюмов для них, будет достаточно 3 уровня данной модели, так как при пошиве костюмов нет необходимости знать ничего подробнее веса, роста и особенностей строения животного.

Следующий рассматриваемый критерий – достоверность. Для представления достоверности данных необходимо вспомнить алгоритм "дырявого ведра". Это достаточно интересный иллюстративный способ. В данном контексте, "дырявое ведро" может отражать степень достоверности данных, где каждая дыра представляет потенциальное искажение или ошибка в информации.

Рассмотрим пункты данного алгоритма и то, что эти пункты могут продемонстрировать.

1. Заполнение ведра:

Ведро начинается с заполнения данными, представляющими исходную, достоверную информацию.

2. Уровень воды (достоверность):

Уровень воды в ведре отражает степень достоверности данных.

Дыры в ведре представляют потенциальные ошибки или неопределенности.

3. Процесс утечек (ошибки искажения):

Дыры в ведре могут представлять ошибки, искажения или неопределенности в данных.

Чем больше дыр в ведре, тем выше вероятность наличия ошибок в информации.

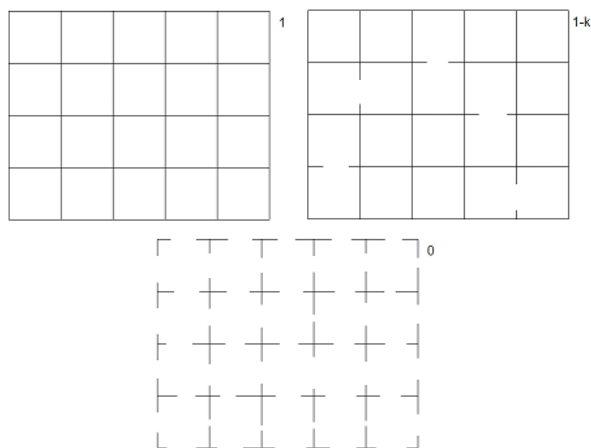
4. Управление дырами (управление достоверностью):

Оптимизация и устранение дыр может быть рассмотрено как процесс повышения достоверности данных.

Механизмы управления, такие как проверки на достоверность и коррекция ошибок, направлены на уменьшение дыр в ведре.

Возвращаясь к модели с тумблерами, тумблер в данном случае представляет собой инструмент для регулировки параметров, влияющих на достоверность данных. Регулировка тумблера может изменять размер дыр в ведре, что влияет на степень допустимых искажений в данных.

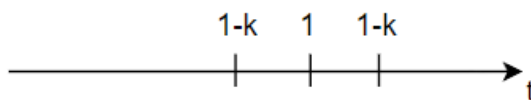
Такой алгоритм позволяет визуализировать понятие достоверности данных, подчеркивая наличие потенциальных ошибок и неопределенностей в информации. Управление дырами и регулировка параметров через тумблеры предоставляют инструменты для улучшения достоверности данных и контроля качества информации.



**Рис. 4.** Модель управления согласованностью

Согласованность данных можно представить в виде сетки, где каждая ячейка представляет собой взаимосвязь между различными аспектами данных. Если в сети нет обрывов, то все данные согласованы, значение параметра для данного случая примем за 1. Чем больше обрывов в сетки, тем ниже значение параметра. Если в сети произошел полный обрыв, тогда значение параметра равно 0. Согласованность будет снижаться, когда будут повышаться значения остальных параметров. Для получения критерия согласованности необходимо проводить постоянные тестирования сети. Пример согласованности данных изображен на рисунке 4, на нем представлены несколько вариантов сети.

Заключительным критерием для рассмотрения осталась своевременность данных. Ее проще всего представить в виде временной линии, в которой отмечена точка, чем дальше ты отдаляешься от точки, тем ниже значение параметра для критерия. Если данные пришли ровно в тот момент, где отмечена точка, то значение параметра будет равно 1. Если данные вообще не поступили тогда параметр критерия равен 0. На рисунке 5 изображена схема данной иллюстрационной модели.



**Рис. 5.** Модель управления своевременностью

Для достижения более качественных данных необходимо предпринять определенные усилия. Каждый критерий, задаваемый через тумблер, является частью сложной системы взаимосвязанных параметров, определяющих общее качество данных. Исследование этих параметров на конкретных моделях позволяет выявить оптимальные настройки для достижения высокой точности данных.

Важно отметить, что эта модель предоставляет возможность учесть различные аспекты при оценке качества данных, что делает ее гибкой и применимой в различных областях. Такой подход к оценке данных открывает новые перспективы для улучшения точности и достоверности информации в различных приложениях.

### Группы пользователей и качество данных для них

После моделирования системы по управлению качеством данных, разберем на настоящих бизнес примерах, для чего же она нужна. Были определены несколько групп пользователей для базы данных и описано по каждой группе, какие параметры качества данных им необходимы на основе пяти основных характеристик: точность, полнота, согласованность, достоверность и своевременность. Стоит помнить, чем больше значение у той или иной характеристики, тем больше нагрузка на нашу систему, поэтому удобно управлять качеством данных, в зависимости от потребностей.

#### 1. Группа Менеджмента:

- **Точность:** Важна для принятия будущей стратегии компании, но могут допускаться не критичные расхождения. Выставим данному параметру значение 0,75.



- **Полнота:** Для определения будущих стратегий, нет необходимости углубляться в происхождении данных, достаточно будет данных с верхних уровней. Выставим данному параметру значение 0,5.
- **Согласованность:** Необходима для единства взгляда на бизнес-процессы. Необходима для связи данных из различных источников. Выставим данному параметру значение 0,8.
- **Достоверность:** Данные должны быть достоверными для надежных анализов. Необходимо соблюдать стабильность данных в течении времени. Выставим данному параметру значение 0,75.
- **Своевременность:** Можно обновлять раз в неделю. Выставим данному параметру значение 0,3.

#### 2. Группа Аналитиков:

- **Точность:** Критична для точных аналитических отчетов. Выставим данному параметру значение 0,9.
- **Полнота:** Важна для комплексного анализа всех аспектов данных. Здесь нам уже надо спуститься на более низкие уровни, для поиска взаимосвязей различных факторов. Выставим данному параметру значение 0,8.
- **Согласованность:** Гарантируется согласованность данных между разными источниками. Выставим данному параметру значение 0,85.
- **Достоверность:** Обеспечивает надежность результатов анализа. Необходимо соблюдать еще более высокую стабильность данных в течении времени. Выставим данному параметру значение 0,9.
- **Своевременность:** Позволяет быстро реагировать на изменения. Необходимо обновлять раз в сутки. Выставим данному параметру значение 0,6.

#### 3. Группа Специалистов по обслуживанию клиентов:

- **Точность:** Важна для предоставления точной информации клиентам. Установим значение, равное 0,75.
- **Полнота:** Гарантирует полноту данных о клиентах и их истории. Необходимы данные из различных уровней доступа, установим значение 0,75.
- **Согласованность:** Обеспечивает единое представление о клиентах. Для того чтобы точно сопоставить данные клиента из различных источников, необходимо установить значение 0,95.
- **Достоверность:** Важна для доверия клиентов к предоставленной информации. Ошибка для этой группы менее критична, поэтому значение будет 0,65.
- **Своевременность:** Позволяет оперативно реагировать на запросы клиентов. Необходимо обновлять раз в 3 часа. Поставим данному параметру значение 0,8.

#### 4. Маркетологи:

- **Точность:** Необходима для эффективного таргетирования рекламы. Установим значение 0,5.
- **Полнота:** Важна для анализа всех аспектов рынка и потребителей. Установим значение 0,5.
- **Согласованность:** Гарантирует согласованность данных между маркетинговыми каналами. Установим значение 0,75.
- **Достоверность:** Обеспечивает надежность данных для разработки маркетинговых стратегий. Установим значение 0,6.
- **Своевременность:** Позволяет маркетологам быстро реагировать на изменения в рынке. Обновляется раз в сутки. Установим значение 0,6.

Эти примеры демонстрируют, как разные группы пользователей могут требовать разные уровни качества данных в зависимости от их конкретных потребностей и задач. На практике, реализация модели регулирования качества, закладывается в информационные системы и должны проходить тщательное, а лучше избыточное тестирование [13].

### Заключение

В данной работе была проведена аналитика в области управления качеством данных, а также предложена модель, призванная оптимизировать данный процесс в системах, связанных с обработкой трехмерной графики, где на разных этапах требуется разное качество изображения [14]. Основываясь на результатах исследования, можно сделать несколько ключевых выводов.

Во-первых, управление качеством данных становится неотъемлемой частью эффективного бизнес-управления в условиях информационного общества. Высококачественные данные являются основой принятия обоснованных решений, обеспечивают точность аналитики и способствуют повышению конкурентоспособности организаций.

Во-вторых, модель управления качеством данных, представленная в данной работе, представляет собой систематический и комплексный подход к решению проблемы. Интегрированная в организационные процессы, она предоставляет инструментарий для постоянного мониторинга, оценки и поддержания качества данных.

Третий вывод связан с практической применимостью модели. Рекомендации, предложенные в рамках данной работы, могут быть использованы организациями при внедрении систем управления качеством данных. Это позволит им минимизировать риски, связанные с недостоверной информацией, и повысить эффективность своей деятельности. Были представлены несколько групп пользователей, и какой уровень качества данных необходим им для работы.

Так же в процессе создания модели управлением качеством данных была получена формула оценки качества данных (формула 2).

Определим оценку качества данных из системы, которую удалось получить в работе «Определение системы оценки качества данных в процессах инженерии данных» (формула 1) получим, что оценка качества данных определенной системы равна:

$$P = \frac{T_{ин}}{T_{ин} + T_{ли}} \cdot \frac{T_{ин}}{T_{ин} + T_{ло}} \cdot \frac{(A \cdot B)}{\|A\| \cdot \|B\|} \cdot \frac{(M_1 - M_2)}{\sqrt{(m_1^2 + m_2^2)}} \cdot t_{реализации} \cdot v_{загрузки} =$$

$$= \frac{T_{ин} \cdot (A \cdot B) \cdot (M_1 - M_2) \cdot t_{реализации} \cdot v_{загрузки}}{(T_{ли} + T_{ин}) \cdot \|A\| \cdot \|B\| \cdot \sqrt{(m_1^2 + m_2^2)}} \quad (3)$$

В целом, данная работа позволяет регулировать качество данных для информационных систем путем изменения значений пяти основных параметров. Последующие исследования должны состоять в более глубоком анализе конкретных отраслей и дальнейшем совершенствовании модели с учетом изменяющейся информационной среды.

## Литература

1. Золотарева П.Ю., Гадасин Д.В., Маклачков К.А. Методы обработки информации в распределенных информационных системах // Тенденции развития Интернет и цифровой экономики: Труды VI Международной научно-практической конференции, Симферополь-Алушта, 01-03 июня 2023 года. Симферополь: ИП Зуева, 2023. С. 187-189. EDN LGONZK.
2. Гадасин Д.В., Бессолицын А.Д. Виды и методы структурирования данных из различных информационных систем: анализ и применение // Актуальные проблемы и перспективы развития экономики, Симферополь - Гурзуф, 12-14 октября 2023 года. Симферополь: ИП Зуева Т. В., 2023. С. 202-204. EDN UGZRXL.
3. Гадасин Д.В., Шведов А.В., Савкин Д.И. Проектирование единого информационного пространства в рамках курса лабораторных работ по дисциплине «Принципы построения систем управления базами данных и знаний» // Методические вопросы преподавания инфокоммуникаций в высшей школе. 2023. Т. 12, № 2. С. 14-21. EDN FLOPUM.
4. Гадасин Д.В., Пак Е.В., Коровушкина В.М., Мелькова Е.К. Предобработка текстовой информации на основе термов естественного языка // REDS: Телекоммуникационные устройства и системы. 2022. Т. 12, № 1. С. 4-11. EDN PDGAVP.
5. Гадасин Д.В., Комкова М.Г., Пантелеева К.А., Гадасин Д.Д. Связь между величиной энтропии и количеством информации при представлении предметной области разными лингвистическими единицами // DSPA: Вопросы применения цифровой обработки сигналов. 2023. Т. 13, № 2. С. 12-21. EDN WRHCSW.
6. Gadasin D.V., Shvedov A.V., Vakurin I.S. Determination of Semantic Proximity of Natural Language Terms for Subsequent Neural Network Training // 2022 Systems of Signals Generating and Processing in the Field of on Board Communications, SOSG 2022 - Conference Proceedings, Moscow, 15-17 марта 2022 г. DOI 10.1109/IEEECONF53456.2022.9744290. EDN LASMDY
7. Гадасин Д.В., Шведов А.В., Вакурин И.С., Трemasова Л.А. Семантический и вероятностный векторы в поисковых запросах // REDS: Телекоммуникационные устройства и системы. 2023. Т. 13, № 2. С. 19-32. EDN HUBVNI.
8. Гадасин Д.В., Шведов А.В., Пантелеева К.А., Гадасин Д.Д. Определение порога количества информации для возможности структурирования данных // Телекоммуникационные и вычислительные системы: Юбилейный сборник трудов тридцатого международного научно-технического форума, Москва, 12-15 декабря 2022 года. М.: Издательство МБА, 2022. С. 125-130. EDN MYMHUP.

9. *Гадасин Д.В., Шведов А.В., Мелькова Е.К.* Структурирование данных исходя из центра масс // Актуальные проблемы и перспективы развития экономики: Труды XXI Международной научно-практической конференции, Симферополь-Гурзуф, 20-22 октября 2022 г. Симферополь: Крымский федеральный университет им. В.И. Вернадского, 2022. С. 266-268. EDN RFCCST.
10. *Гадасин Д.В., Шведов А.В., Вакурин И.С.* Определение семантической близости текстов с использованием алгоритма сравнения сущности графов // REDS: Телекоммуникационные устройства и системы. 2022. Т. 12, № 4. С. 11-19. EDN PVJKQJ.
11. *Маклачкова В.В., Шведов А.В., Шульпина П.Д., Гадасин Д.В.* Сравнительный анализ алгоритмов обратного распространения ошибки и имитации отжига // REDS: Телекоммуникационные устройства и системы. 2023. Т. 13, № 1. С. 26-32. EDN ORDSXE.
12. *Gadasin D.V., Koltsova A.V., Gadasin D.D.* Algorithm for Building a Cluster for Implementing the 'Memory as a Service' Service in the IoT Concept // 2021 Systems of Signals Generating and Processing in the Field of on Board Communications, Conference Proceedings, Moscow, 16-18 марта 2021 года. P. 9416112. DOI 10.1109/IEEECONF51389.2021.9416112. EDN VRPCFG.
13. *Zolotukhin P.A., Melkova E.K., Gadasin D.V., Korovushkina V.M.* Using Intelligent Testing as a Tool to Improve the Quality of Transport Information Systems // 2022 Systems of Signals Generating and Processing in the Field of on Board Communications, SOSG 2022 - Conference Proceedings, Moscow, 15-17 марта 2022 г. DOI 10.1109/IEEECONF53456.2022.9744348. EDN NOMJLX.
14. *Гадасин Д.В., Шведов А.В., Кузин И.А.* Трехмерная реконструкция объекта по одному изображению с использованием глубоких сверточных нейронных сетей // Т-Comm: Телекоммуникации и транспорт. 2022. Т. 16, № 7. С. 29-35. DOI 10.36724/2072-8735-2022-16-7-29-35. EDN YT
15. *Гадасин Д.В., Вакурин И.С.* Кластерное проектирование сетей Wi - Fi с высокой плотностью абонентов // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2020. № 1. С. 138-144. EDN EHXMFJ.
16. *Гадасин Д.В., Юдина А.А.* Кластеризация в крупномасштабных сетях // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2020. № 1. С. 19-26. EDN OYSXON.
17. *Гадасин Д.В., Кольцова А.В., Полякова А.Н.* Модель построения кластера для пограничных вычислений // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2020. № 1. С. 86-92. EDN NOUODG.
18. *Гадасин Д.В., Смальков Н.А., Кузин И.А.* Использование метода роя частиц для балансировки нагрузки в сетях Интернета вещей // Системы синхронизации, формирования и обработки сигналов. 2022. Т. 13, № 2. С. 17-23. EDN LIUWNT.
19. *Докучаев В.А., Кальфа А.А., Мытенков С.С., Шведов А.В.* Анализ технических решений по организации современных центров обработки данных // Т-Comm: Телекоммуникации и транспорт. 2017. Т. 11. № 6. С. 16-24. EDN ZAOJVB.
20. *Shvedov A.V., Gadasin D.V., Alyoshintsev A.V.* Segment routing in data transmission networks // Т-Comm: Телекоммуникации и транспорт. 2022. Vol. 16. No 5. P. 56-62. DOI 10.36724/2072-8735-2022-16-5-56-62. EDN VAYLJQ.
21. *Гадасин Д.В., Мисевич И.М.* Использование материалов наглядного информирования в воспитательной работе студентов // Методические вопросы преподавания инфокоммуникаций в высшей школе. 2022. Т. 11. № 3. С. 4-12. EDN TXHBXF.
22. *Гадасин Д.В., Мисевич И.М.* Выполнение курса лабораторных работ по предмету "сетевые технологии в информационных системах и сервисах" // Методические вопросы преподавания инфокоммуникаций в высшей школе. 2022. Т. 11. № 4. С. 4-11. EDN PVRNZW.
23. *Гадасин Д.В., Маклачкова В.В., Шведов А.В.* Особенности организации и проведения вузовского чемпионата по стандартам Worldskills по компетенции "сетевое и системное администрирование" // Методические вопросы преподавания инфокоммуникаций в высшей школе. 2021. Т. 10. № 4. С. 9-18. EDN FCPLCH.
24. *Назаров М.Д., Шведов А.В.* Корреляция атрибутов соглашения об уровне обслуживания с основными параметрами QoS в корпоративных сетях // Телекоммуникации и информационные технологии. 2020. Т. 7. № 2. С. 73-79. EDN VQHDTJ.
25. *Шведов А.В., Гадасин Д.В., Клыгина О.Г.* Организация взаимодействия туманных вычислений и сегментной маршрутизации для предоставления сервисов IOT в smart grid // Системы синхронизации, формирования и обработки сигналов. 2022. Т. 13. № 3. С. 40-49. EDN TRRYZN.
26. *Kalmykov N.S., Dokuchaev V.A.* Segment routing as a basis for software defined network // Т-Comm: Телекоммуникации и транспорт. 2021. Т. 15. № 7. С. 50-54.
27. *Dokuchaev V.A., Maklachkova V.V., Statev V.Yu.* Classification of personal data security threats in information systems // Т-Comm: Телекоммуникации и транспорт. 2020. Т. 14. № 1. С. 56-60.
28. *Докучаев В.А., Маклачкова В.В., Статьев В.Ю.* Цифровизация субъекта персональных данных // Т-Comm: Телекоммуникации и транспорт. 2020. Т. 14. № 6. С. 27-32.

29. *Pavlov S.V., Dokuchaev V.A., Mytenkov S.S.* Model of a fuzzy dynamic decision support system // Т-Comm: Телекоммуникации и транспорт. 2020. Т. 14. № 9. С. 43-47.
30. *Pavlov S.V., Dokuchaev V.A., Maklachkova V.V., Mytenkov S.S.* Features of supporting decision making in modern enterprise infocommunication systems // Т-Comm: Телекоммуникации и транспорт. 2019. Т. 13. № 3. С. 71-74.
31. *Гадасин Д.В., Кольцова А.В., Гадасин Д.Д., Полякова А.Н.* Оценка вероятности формирования виртуального кластера // Системы синхронизации, формирования и обработки сигналов. 2021. Т. 12. № 1. С. 4-12.
32. *Кузин И.А., Гадасин Д.В.* Модель контейнера данных для минимизации трафика при передаче субъективных характеристик объектов на изображении трехмерной сцены // Телекоммуникации и информационные технологии. 2021. Т. 8. № 2. С. 96-100.
33. *Усачева Д.И., Шишкин М.О., Гадасин Д.В., Гузев А.В.* Применение OLAP-технологий для анализа многомерных данных в контакт-центре // Телекоммуникации и информационные технологии. 2019. Т. 6. № 1. С. 142-149.
34. *Гадасин Д.В., Кузин И.А.* Модель представления цветовых и глубинметрических характеристик объектов на изображении // DSPA: Вопросы применения цифровой обработки сигналов. 2021. Т. 11. № 1. С. 31-38.

# АНАЛИЗ МЕТОДОВ ТРЕХМЕРНОЙ РЕКОНСТРУКЦИИ ОБЪЕКТОВ, НЕПОДВИЖНЫХ ОТНОСИТЕЛЬНО СКАНИРУЮЩИХ УСТРОЙСТВ

Голованова Анна Сергеевна,  
МТУСИ, аспирант, Москва, Россия  
[ann-august@yandex.ru](mailto:ann-august@yandex.ru)

## Аннотация

Современные технологии требуют точных и эффективных методов трёхмерной реконструкции объектов для широкого спектра приложений, от компьютерного зрения до виртуальной реальности. В данной статье представлен подробный обзор и сравнение трех бинокулярных методов, используемых для создания трёхмерных моделей изображаемых объектов.

**Ключевые слова:** Трёхмерное восстановление, бесконтактные методы, параллакс, карта глубины, компьютерное зрение

## Введение

Возможности трёхмерной реконструкции объектов так развились за последнее время, что количество вариантов для получения восстановленных трёхмерных данных объекта стало значительно большим, однако нет универсального метода для поставленной задачи т.к. и объект восстановления может быть сложен, так и условия могут быть недостаточно хорошими для получения качественного результата [8-22]. Среди множества современных методов построения трёхмерной модели объекта существуют те, что используют два изображения с бинокулярной пары для восстановления трёхмерной информации их называют бинокулярными. В данной статье рассмотрены такие подходы, а также приведены их описания и основные шаги реализации. Описанные преимущества и недостатки позволят оценить и сделать выводы для работы с каждым из них в зависимости от поставленной задачи.

## Методы

Бинокулярное построение трёхмерной модели является методом, использующим два изображения с бинокулярной пары для воссоздания трёхмерной информации. Существует несколько методов для реализации этого процесса. Давайте рассмотрим и сравним некоторые из них:

- Стереопара метод
- Спектральный метод
- Адаптивные метод.

Рассмотрим каждый из них более полно.

### Стереопарный метод построения трёхмерной модели

Стереопарный метод трёхмерного восстановления объекта использует пару изображений, полученных с разных точек обзора, чтобы определить глубину каждой точки объекта и создать трёхмерное представление.

Для каждой точки на изображении первой камеры (левого изображения) ищется соответствующая точка на изображении второй камеры (правого изображения). Это можно сделать путем сопоставления особых признаков, таких как углы и границы. После нахождения соответствующих точек, используется триангуляция для определения глубины каждой точки объекта.

Триангуляция основана на принципе подобия треугольников. Если известны расстояния между камерами, положение камер и соответствующие точки на изображениях, то можно вычислить глубину каждой точки объекта. Глубина вычисляется как отношение расстояния между камерами к разности соответствующих координат точек на изображениях.

Этот процесс повторяется для всех точек объекта на изображениях, что позволяет создать трёхмерное представление объекта.

Однако стереопарный метод имеет свои ограничения. Например, он требует точного сопоставления особых признаков на изображениях для определения соответствующих точек. Также необходимо знать расстояние между камерами и положение камер относительно объекта.

Кроме того, этот метод может иметь проблемы с неоднородным освещением и текстурами объекта, что может затруднить точное определение соответствующих точек [1].

В целом, стереопарный метод трехмерного восстановления объекта позволяет получить достаточно точные трехмерные модели объектов, но требует аккуратной калибровки камер и точного сопоставления особых признаков на изображениях.

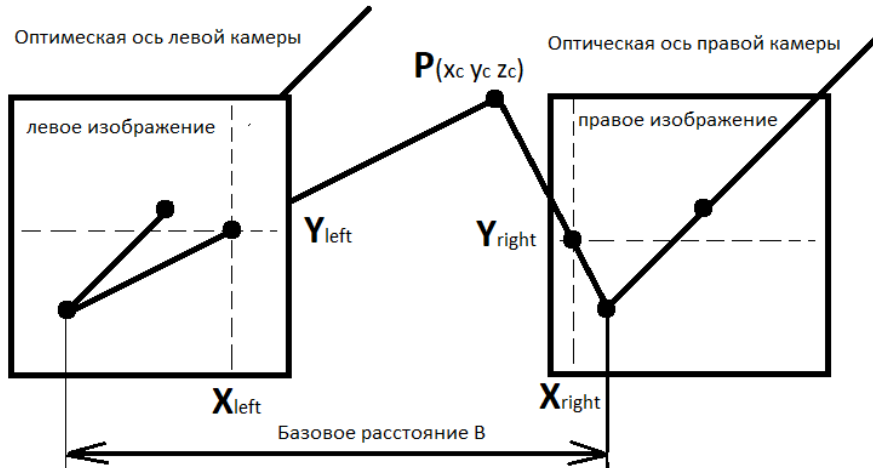


Рис. 1. Принцип стереопарного метода

На рисунке выше показана принципиальная схема простого бинокулярного стереоизображения с наклоном вверх. Расстояние между центрами проекции двух камер является базовым расстоянием  $B$ . Обе камеры одновременно наблюдали одну и ту же характерную точку  $P$  объекта пространства-времени и получали изображения точки  $P$  на «левом глазу» и «правом глазу» соответственно, и их координаты были  $P_{left} = (X_{left}, Y_{left})$ ;  $P_{right} = (X_{right}, Y_{right})$ . Если изображения двух камер находятся в одной плоскости, координаты  $Y$  координат изображения особой точки  $P$  должны быть одинаковыми, то есть  $Y_{left} = Y_{right} = Y$ . Следующее отношение может быть получено из треугольного геометрического отношения:

$$X_{left} = f \frac{x_c}{z_c}$$

$$X_{right} = f \frac{x_c - B}{z_c}$$

$$Y = f \frac{y_c}{z_c}$$

Параллакс (Disparity) =  $X_{left} - X_{right}$ . Из этого можно рассчитать трехмерные координаты характерной точки  $P$  в системе координат камеры:

$$x_c = \frac{B \cdot X_{left}}{\text{Disparity}}$$

$$y_c = \frac{B \cdot Y}{\text{Disparity}}$$

$$z_c = \frac{B \cdot f}{\text{Disparity}}$$

Следовательно, до тех пор, пока любая точка на плоскости изображения левой камеры может находить соответствующую точку совпадения на плоскости изображения правой камеры, трехмерные координаты точки могут быть полностью определены.[1] Этот метод представляет собой операцию «точка-точка». Пока во всех точках на плоскости изображения имеется соответствующая точка сопоставления, вы можете участвовать в вышеуказанной операции, чтобы получить соответствующие трехмерные координаты.

Описание метода:

Стереопарный метод использует пару изображений, захваченных с двух камер с небольшим горизонтальным сдвигом между ними. Основная идея заключается в том, чтобы использовать этот сдвиг для расчета глубины объектов в сцене.

Шаги стереопарного метода:

1. Съемка стереопары:
  - Две камеры захватывают изображения одной и той же сцены с небольшим горизонтальным сдвигом между ними.
2. Предварительная обработка:
  - Изображения могут быть выровнены и приведены в соответствие, чтобы обеспечить точное соответствие пикселей между ними.
3. Вычисление сдвига:
  - Определение сдвига между соответствующими пикселями на двух изображениях. Это может быть выполнено с использованием методов корреляции или оптимизационных алгоритмов.
4. Триангуляция:
  - Применение триангуляции для определения глубины каждой точки в сцене. Это позволяет построить трехмерную модель.

Преимущества стереопарного метода:

- Простота реализации: требует всего лишь пары камер и базовой обработки изображений.
- Эффективность для статичных сцен: хорошо работает в статичных условиях, где сцена остается неизменной.
- Простота восприятия: подходит для восприятия глубины человеческим зрением и стереоскопических отображений.

Недостатки стереопарного метода:

- Чувствительность к изменениям в освещении: Изменения в освещении могут влиять на точность расчетов глубины.
- Точная калибровка камер: требует точной калибровки камер для правильного соответствия пикселей.
- Ограничения на динамичные сцены: не так эффективен в случае динамичных сцен, где объекты движутся.

Применение стереопарного метода:

- Робототехника: для навигации роботов в неизвестных средах.
- Медицинская область: в реконструкции трехмерных моделей для хирургического планирования.
- Компьютерное зрение: для распознавания объектов и обнаружения глубины в различных приложениях.

Стереопарный метод является эффективным и распространенным способом бинокулярного построения трехмерной модели, который находит широкое применение в различных областях, несмотря на свои ограничения [2, 3].

### **Спектральный метод построения трёхмерной модели**

Спектральный метод трехмерного восстановления объекта – это математический метод, который основывается на использовании спектральных данных для определения трехмерной структуры объекта.

Для начала необходимо получить спектральные данные объекта. Это может быть сделано с помощью специального оборудования, такого как спектрометры или спектральные камеры. Спектральные данные представляют собой набор значений интенсивности излучения в различных длинах волн.

Затем производится обработка и анализ спектральных данных. Это может включать в себя различные методы сегментации изображений, фильтрацию шума, выделение контуров объектов и другие операции обработки изображений.

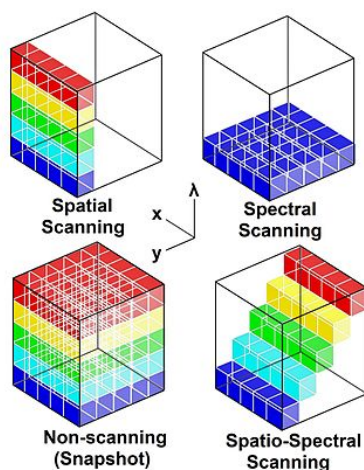
Далее производится восстановление трехмерной структуры объекта на основе спектральных данных. Операция восстановления может включать в себя использование алгоритмов компьютерного

зрения, обработку данных методами стереозрения, обработку изображений в глубину и другие методы трехмерной реконструкции.

В результате получается трехмерная модель объекта, которая может быть использована для дальнейшего анализа и визуализации. Трехмерная модель может быть представлена в виде облака точек, поверхности или других геометрических форматов.

Спектральный метод трехмерного восстановления объекта имеет широкий спектр применения в различных областях, включая медицину, археологию, промышленность и другие отрасли. Он позволяет получить детальное представление о трехмерной структуре объектов и обнаружить скрытые детали или особенности, которые не могут быть видны на двухмерных изображениях.

При использовании спектрального метода каждый двухмерный выходной сигнал датчика представляет собой монохроматическую ("одноцветную") пространственную  $(x, y)$  карту сцены. Устройства HSI для спектрального сканирования обычно основаны на оптических полосовых фильтрах (настраиваемых или фиксированных). Спектральный метод восстановления сцены осуществляется путем замены одного фильтра за другим, пока платформа остается неподвижной. В таких "пристальных" системах сканирования по длине волны при наличии движения в кадре может возникать размытие спектра, что делает недействительной спектральную корреляцию/обнаружение [4]. Тем не менее, преимущество заключается в возможности выбирать спектральные диапазоны и иметь прямое представление двух пространственных измерений сцены [2, 4]. Если система визуализации используется на движущейся платформе, такой как самолет, полученные изображения на разных длинах волн соответствуют разным участкам сцены. Пространственные особенности на каждом из изображений могут быть использованы для выравнивания пикселей.



**Рис. 2.** Методы получения гиперспектральной визуализации, визуализируемые в виде участков гиперспектрального куба данных с его двумя пространственными измерениями  $(x, y)$  и одним спектральным измерением (лямбда) [5]

Описание метода:

Спектральный метод использует анализ спектральных характеристик изображений для построения трехмерной модели сцены. Основная идея заключается в том, чтобы анализировать различия в частотах между изображениями для определения глубины объектов.

Шаги спектрального метода:

1. Получение изображений:
  - Две камеры захватывают изображения одной и той же сцены.
2. Преобразование Фурье:
  - Каждое изображение подвергается преобразованию Фурье для перехода из пространства пикселей в частотное пространство.
3. Анализ различий в частотах:
  - Вычисление различий в частотах между соответствующими частотными областями двух изображений.
4. Определение глубины:



- Использование полученных различий для определения глубины каждой точки в сцене.

Преимущества спектрального метода:

- Устойчивость к изменениям в освещении: метод менее чувствителен к изменениям в освещении по сравнению с некоторыми другими методами.
- Применимость к текстурированным объектам: эффективен в случае объектов с разнообразной текстурой.
- Возможность работы с динамичными сценами: может быть применен к динамичным сценам, где объекты могут изменять свое положение.

Недостатки спектрального метода:

- Большие вычислительные затраты: преобразование Фурье и анализ частот требуют значительных вычислительных ресурсов.
- Менее точный для монотонных текстур: может быть менее точен в случае монотонных текстур или однородных областей.

Применение спектрального метода:

- Медицинская диагностика: для реконструкции трехмерных моделей органов на основе медицинских изображений.
- Геодезия и картография: в реконструкции рельефа местности.
- Компьютерное зрение: в области распознавания объектов, особенно там, где важна текстурная информация.

Спектральный метод предоставляет эффективный подход к бинокулярному построению трехмерной модели, особенно в условиях, где важны текстурные характеристики объектов и изменения в освещении.

### **Адаптивные методы построения трёхмерной модели**

Адаптивный бинокулярный метод трехмерного восстановления объекта основан на использовании двух камер, расположенных на некотором расстоянии друг от друга, и анализе изображений, полученных от них.

Для начала необходимо калибровать камеры, то есть определить их внутренние параметры (фокусное расстояние, коэффициенты дисторсии и т.д.) и внешние параметры (положение и ориентацию каждой камеры в пространстве). Это делается с помощью специального калибровочного процесса, который требует съемки объекта с известной геометрией [6, 7].

Затем проводится съемка объекта с помощью бинокулярной системы. Каждая камера формирует изображение объекта, которое затем анализируется для получения информации о его глубине.

Алгоритм адаптивного бинокулярного метода трехмерного восстановления объекта состоит из следующих шагов:

1. Детекция и сопоставление особенных точек на обоих изображениях. Это может быть сделано с использованием алгоритма SIFT (Scale-Invariant Feature Transform) или других подобных методов.
2. Вычисление диспаратности для каждой пары сопоставленных точек. Диспаратность – это разница в горизонтальных координатах точек на обоих изображениях. Она пропорциональна глубине объекта и может быть вычислена с использованием различных алгоритмов, таких как SSD (Sum of Squared Differences) или SAD (Sum of Absolute Differences).
3. Создание глубинной карты путем сопоставления диспаратностей с конкретными пикселями на изображении.
4. Адаптивное уточнение глубинной карты. В этом шаге используется информация о границах объектов на изображении для улучшения точности восстановления глубины. Например, можно применить метод графовых срезов для нахождения наиболее правдоподобной глубинной карты, учитывая граничные условия.
5. Преобразование глубинной карты в трехмерное представление объекта. Это может быть достигнуто путем применения обратного процесса проекции, используя калибровочные параметры камеры и формулы из геометрии.

Таким образом, адаптивный бинокулярный метод трехмерного восстановления объекта позволяет получить трехмерное представление объекта с использованием информации, полученной от двух камер и алгоритмов компьютерного зрения.

Принцип данного метода схож со стереопарным, однако включает в себя адаптивные алгоритмы, поэтому они приведены далее.

Адаптивные методы построения трёхмерной модели используют алгоритмы, способные адаптироваться к различным условиям и характеристикам сцены. Они представляют собой группу методов, которые активно изменяют свои параметры или стратегии в зависимости от условий съёмки или объектов, чтобы повысить точность и надёжность получаемых результатов. Вот несколько типов адаптивных методов:

1. Машинное обучение и нейронные сети:
  - Алгоритмы машинного обучения и нейронные сети могут адаптироваться к особенностям сцен и улучшать точность оценки глубины или реконструкции трёхмерных моделей. Они могут обучаться на больших объёмах данных для улучшения обобщения и способности обрабатывать различные условия.
2. Адаптивные фильтры и алгоритмы коррекции:
  - Эти методы могут изменять свои параметры на основе информации, полученной в процессе обработки изображений. Например, они могут динамически корректировать фильтры для устранения шума, улучшения качества изображения и точности определения глубины.
3. Динамическая настройка параметров:
  - Этот подход позволяет алгоритмам изменять параметры в зависимости от условий освещения, текстуры объектов, разрешения изображений и других факторов, чтобы достичь оптимальной точности и стабильности в построении трёхмерных моделей.
4. Адаптивная сегментация:
  - Используется для выделения объектов или регионов на изображениях, где различные части сцены могут требовать разных методов или параметров для точного восстановления их трёхмерной структуры.

Преимущества адаптивных методов включают способность работать в различных условиях, лучшую устойчивость к изменениям и способность обрабатывать динамические сцены. Однако они могут потребовать больше вычислительных ресурсов и тщательной настройки параметров для достижения оптимальных результатов.

Важно отметить, что выбор конкретного адаптивного метода зависит от конкретных требований задачи, характеристик сцены и доступных ресурсов для обработки и анализа изображений.

### **Заключение**

Принцип стереопарного метода основан на триангуляции для определения глубины объектов и использует пару изображений с небольшим горизонтальным сдвигом. Прост в реализации и показывает высокую эффективность для статических сцен, однако чувствителен к освещению и требует точную калибровку камер. Подходит для стереоскопического восприятия.

Спектральный метод использует спектральные характеристики изображений и анализирует различия в частотах для определения глубины. Устойчив к изменениям в освещении и может использоваться для динамических сцен, что характерно отличает его от предыдущего, но требует более сложных вычислений и имеет худшую точность в случае текстурированных объектов. Эффективен при различной текстуре объектов.

Адаптивные методы используют алгоритмы, способные адаптироваться к изменяющимся условиям и могут включать в себя машинное обучение для улучшения результатов. Более гибкие и способные работать в различных условиях и могут обрабатывать динамические сцены, но потребуют более высокой вычислительной мощности, а также менее стабильны в сложных сценах без должной настройки. Подходит для сцен с изменяющимися условиями освещения и динамикой объектов.

Выбор метода зависит от конкретных требований задачи. Если необходима простота и стабильность, стереопарный метод может быть предпочтителен. В случае изменяющихся условий и необходимости работы с динамическими сценами, адаптивные методы или спектральные методы могут быть более подходящими. Важно также учесть вычислительные требования и возможности калибровки оборудования.

Общие соображения:

- Выбор метода зависит от конкретных требований задачи.
- Стереопарный метод подходит для простых статичных сцен.
- Спектральный метод обеспечивает устойчивость к освещению, но может быть менее точным.
- Адаптивные методы предоставляют гибкость, но требуют высокой вычислительной мощности и подходят для динамичных сцен.

Важно учитывать:

- Вычислительные требования.
- Способность калибровки оборудования.
- Сложность реализации и поддержки.

Выбор метода должен зависеть от конкретных потребностей задачи, уровня сложности сцены, и доступных ресурсов для реализации и поддержки выбранного метода.

## Литература

1. *Силантьева А.С., Власюк И.В.* Анализ определения стереобазиса распространенными методами для решения прикладных задач фотограмметрии // Телекоммуникационные и вычислительные системы. Юбилейный сборник трудов тридцатого международного научно-технического форума. 2022. С. 349-352.
2. *Лу Гуолань, Фэй, Баовэй.* SPIE – Журнал биомедицинской оптики – Медицинская гиперспектральная визуализация: обзор // Журнал биомедицинской оптики. 2014. №19 (1): 010901. Bibcode:2014JBO....19a0901L. doi:10.1117/1.JBO.19.1.010901. PMC 3895860. PMID 24441941
3. *Силантьева А.С.* Классификация и сравнительный анализ методов трехмерной реконструкции объектов // Системы синхронизации, формирования и обработки сигналов. 2023. Т. 14. № 4. С. 52-60.
4. *Шурмер Дж.Х.* Исследовательские лаборатории BBC, Технологические горизонты. Декабрь 2003.
5. *Dosovitskiy A., Fischer P., Ilg E., Hausser P., Hazirbas C., Golkov V., Van Der Smagt P., Cremers D., Brox T.* FlowNet: Learning optical flow with convolutional networks. In: ICCV. 2015.
6. *Flynn J., Neulander I., Philbin J., Snavely N.* DeepStereo: Learning to predict new views from the world's imagery. In: CVPR. 2016.
7. *Gu X., Fan Z., Zhu S., Dai Z., Tan F., Tan P.* Cascade cost volume for high-resolution multi-view stereo and stereo matching. In: CVPR. 2020. <https://blog.csdn.net/shenziheng1/article/details/52883536/#commentsedit>
8. *Федоров В.Д., Кудряшов В.В., Власюк И.В.* Методология комплексного измерения пространственно-динамических характеристик преобразователей свет-сигнал // Системы синхронизации, формирования и обработки сигналов. 2023. Т. 14, № 2. С. 39-46.
9. *Степанов Н.С., Матуа Д.Д., Мазин В.А., Вотяков С.Ю., Винецкий В.В., Власюк И.В.* Анализ текущих алгоритмов вычисления области регионов интереса пользователей при потоковой передаче видеоконтента // Т-Comm: Телекоммуникации и транспорт. 2023. Т. 17. № 2. С. 27-32.
10. *Романов С.Г., Власюк И.В.* Методика расчета параметров анти-алайсинговых фильтров для коррекции спектральных характеристик в зависимости от используемых структур дискретизации массивов светофильтров // Т-Comm: Телекоммуникации и транспорт. 2023. Т. 17. № 5. С. 4-13.
11. *Safonov I., Vyatkin M., Hizhenkova O., Nezhivleva K., Vlasuyk I.* Threat analysis and security of personal data using user devices with eye-trackers // Т-Comm. 2023. Т. 17. № 8. С. 56-62.
12. *Власюк И.В., Узеев А.А., Пахомова Е.А.* Исследование методов коррекции изображений с расширенным динамическим диапазоном для воспроизведения на устройствах с ограниченными параметрами отображения // Телекоммуникации и информационные технологии. 2023. Т. 10. № 1. С. 135-144.
13. *Mozhaeva A., Vashenko E., Selivanov V., Potashnikov A., Vlasuyk I., Streeter L.* Analysis of current video databases for quality assessment // Т-Comm. 2022. Т. 16. № 2. С. 48-56.
14. *Vyatkin M., Potashnikov A., Selivanov V., Vlasuyk I., Nezhivleva K., Mozhaeva A.* Method of preventing leakage of personal data through eyetracking modules of user devices // Т-Comm. 2022. Т. 16. № 7. С. 44-51.
15. *Valitskaya N.S., Vlasuyk I.V., Potashnikov A.M.* Video compression method on the basis of discrete wavelet transform for application in video information systems with non-standard parameters // Т-Comm. 2020. Т. 14. № 3. С. 47-53.
16. *Поташиников А.М., Власюк И.В.* Метод построения равноконтрастного цветового пространства для данной системы отображения информации и условий контроля // Т-Comm: Телекоммуникации и транспорт. 2020. Т. 14. № 4. С. 15-22.
17. *Ревина Д.С., Левченко Р.Р., Маликов А.Ю., Власюк И.В., Егоров Д.А.* Анализ эффективности безреференсных метрик оценки качества видео последовательности // Системы синхронизации, формирования и обработки сигналов. 2022. Т. 13. № 2. С. 41-53.

18. Можяева А., Власюк И., Неживлева К., Комина А. Методология создание стимула для количественной оценки эффекта мерцания при сжатии видео // Системы синхронизации, формирования и обработки сигналов. 2022. Т. 13. № 6. С. 29-36.

19. Власюк И.В., Пашиковская А.Р., Мясникова В.С., Никольская Д.И., Можяева А.И. Анализ стоимости создания современных баз данных видеопоследовательностей с субъективной оценкой качества // DSPA: Вопросы применения цифровой обработки сигналов. 2022. Т. 12. № 1. С. 4-11.

20. Иванчев В.В., Власюк И.В. Метод распределения выборки цветовых оттенков в цветовом пространстве CIELCH / Телекоммуникации и информационные технологии. 2022. Т. 9. № 2. С. 108-119.

21. Власюк И.В., Киселева А.С. Анализ эффективности безреференсных метрик применительно к оценке качества видео при потоковой передаче // Телекоммуникации и информационные технологии. 2022. Т. 9. № 2. С. 65-74.

22. Егоров Д.А., Федоров В.Д., Лейман В.В., Власюк И.В. Методика оценки пространственно-частотной характеристики камер на основе генеративных случайных последовательностей // DSPA: Вопросы применения цифровой обработки сигналов. 2021. Т. 11. № 1. С. 47-53.

## МЕНЕДЖЕР ПАРОЛЕЙ С ГЕОЛОКАЦИЕЙ

**Макаров Арсений Александрович**  
МТУСИ, студент, Москва, Россия,  
[makarov.arsenii02@mail.ru](mailto:makarov.arsenii02@mail.ru)

**Большаков Александр Сергеевич,**  
МТУСИ, доцент кафедры ИБ, к.т.н., Москва, Россия  
[alexbol57@mail.ru](mailto:alexbol57@mail.ru)

### Аннотация

*В данном исследовании представлена разработка менеджера паролей, который интегрирует в себя преимущества геолокационной аутентификации для повышения безопасности и удобства управления паролями в современном цифровом мире. Работа акцентирует внимание на актуальных проблемах управления паролями и предлагает решения, включая геолокационную защиту и новые методы восстановления пароля. Это исследование подчеркивает важность развития программных продуктов, способных обеспечивать надежную кибербезопасность и удовлетворять потребности современных пользователей в безопасности и удобстве.*

**Ключевые слова:** менеджер паролей, геолокация, безопасность, управление паролями, двухфакторная аутентификация, геозависимые настройки, синхронизация, защита от потери, шифрование данных, геолокационные данные.

### Введение

В эпоху цифровых технологий, где сохранение конфиденциальности играет ключевую роль, эффективное управление паролями становится неотъемлемой частью нашей повседневной жизни. Пароли представляют собой первую линию защиты от несанкционированного доступа к нашей ценной информации, и обеспечение их безопасности становится крайне важным аспектом. В данном исследовании представлены результаты создания менеджера паролей, в котором ключевым элементом является технология геолокации. Этот продукт совмещает современные методы обеспечения безопасности, предоставляя пользователям возможность более надежного управления своими паролями с учетом их географического положения.

**Менеджер паролей** [1] - программное или аппаратное средство, которое предназначено для хранения, генерации и управления паролями. Основная цель менеджера паролей - обеспечить безопасное и удобное управление множеством паролей, которые используются для доступа к различным онлайн-сервисам, приложениям, веб-сайтам и другим системам.

### Проблематика

В современном мире все больше внимания уделяется проблеме безопасности управления паролями. Менеджеры паролей, как важное средство для обеспечения безопасности данных и доступа к онлайн-сервисам, также сталкиваются с рядом проблем и вызовов.

Несмотря на то, что менеджеры паролей призваны обеспечить безопасное управление паролями, они сами могут стать точкой уязвимости. Каждый менеджер паролей имеет мастер-пароль, который является ключом ко всем хранящимся в нем данным.

Совместимость и ограничения, накладываемые некоторыми веб-сайтами и приложениями на использование менеджеров паролей, также являются факторами, мешающими полноценному использованию таких инструментов.

Забывчивость или утеря мастер-пароля может привести к блокировке доступа ко всем сохраненным паролям, что ставит пользователей в неприятное положение.

Существуют проблема аутентификации. Например, ситуация, когда у пользователя нет доступа к дополнительным средствам аутентификации, что может привести к блокировке доступа. Кроме того, некоторые пользователи могут считать процесс аутентификации неудобным и отказываться от его использования, создавая дополнительные риски для безопасности.

### Актуальность

Менеджер паролей с функцией геолокации весьма актуален, поскольку обеспечивает дополнительный уровень безопасности. Он позволяет адаптировать уровень защиты в зависимости от местоположения пользователя, предоставляя более гибкие и персонализированные настройки безопасности. Интеграция геолокационных данных усиливает общую безопасность цифровых идентификаторов и минимизирует риски несанкционированного доступа к учетным записям, а также менеджеры паролей представляют собой неотъемлемый элемент обеспечения кибербезопасности и соответствия требованиям, установленным Федеральной службой по техническому и экспортному контролю (ФСТЭК) [2]. В контексте регулирования и нормативных документов, созданных ФСТЭК, менеджеры паролей выполняют несколько ключевых функций.

Во-первых, ФСТЭК устанавливает строгие стандарты безопасности информации [4], политику паролей, включая требования к их сложности, длине и периодической смене. Менеджеры паролей служат средством для соблюдения этих стандартов, помогая пользователям создавать и хранить пароли, соответствующие требованиям ФСТЭК.

Во-вторых, ФСТЭК требует проведения аудита безопасности [3], чтобы обнаруживать аномальную активность и нарушения в системе. Менеджеры паролей предоставляют инструменты аудита, которые позволяют оценивать безопасность паролей и контролировать их использование.

Кроме того, ФСТЭК разрабатывает рекомендации и руководства по безопасности информации [4], включая методы защиты паролей и управления ими. Менеджеры паролей могут служить средством для реализации этих рекомендаций, помогая организациям следовать стандартам безопасности.

Проведено сравнение программ, представленное в таблице 2, указывающее на отсутствие функционала геолокации и seed-фразы, что является общим недостатком рассмотренных менеджеров паролей. В таблице 1, представлен функционал выбранных менеджеров паролей.

Таблица 1

Название	Функционал
1Password	Сервис 1Password [6] обеспечивает защиту данных пользователей с применением 256-битного AES шифрования и предоставляет возможность создания нескольких хранилищ, таких как хранилище для паролей и хранилище для банковских карт. Важной характеристикой данного сервиса является поддержка двухфакторной аутентификации (2FA) и возможность сканирования баз данных на предмет украденных паролей.
Dashlane	Сервис Dashlane [7], в свою очередь, предоставляет функциональность, аналогичную 1Password, при этом отличительной особенностью является наличие встроенной VPN и системы оповещения о слабых паролях.
RoboForm	Сервис RoboForm [8] различается от вышеупомянутых сервисов тем, что в нем предусмотрено защищенное хранилище для закладок. Это позволяет пользователям сохранять и синхронизировать свои закладки браузера между компьютером и другими устройствами, на которых установлен RoboForm.

Таблица 2.

Название	Наличие seed-фразы	Наличие аутентификации	Наличие локального хранилища	Наличие облачного хранилища	Наличие геолокации
1Password	-	+	+	+	-
Dashlane	-	-	-	+	-
RoboForm	-	+	-	+	-

### Цель работы

Целью данной работы является не только разработка, но и внедрение функциональных усовершенствований в менеджер паролей, сфокусированных на обеспечении высокого уровня безопасности и максимального комфорта пользователей.

Одним из ключевых инновационных аспектов улучшения классического менеджера паролей является внедрение геолокационной защиты. Этот подход обеспечивает дополнительный уровень безопасности, позволяя системе автоматически блокировать доступ к аккаунту при попытке несанкционированного входа с незнакомого местоположения. Одновременно, чтобы учесть ситуации перемещения пользователя между городами или регионами, система предоставляет возможность предварительно указать планируемые даты и местоположение. Этот подход гарантирует безопасность доступа, предотвращая потенциальные несанкционированные вторжения.

Дополнительным важным изменением является отсутствие опции восстановления пароля через электронную почту. Вместо этого пользователю предоставляется seed-фраза [9], которую необходимо сохранить в безопасном месте. Это решение эффективно устраняет возможность взлома аккаунта через доступ к электронной почте, подчеркивая важность ответственного хранения уникальной фразы для дополнительного слоя безопасности.

### Принцип работы программы

1. Геолокационная защита: Основным элементом повышения безопасности пользователя является функция определения местоположения. В случае несанкционированного доступа к аккаунту, система автоматически блокирует аккаунт, поскольку пользователь всегда будет ассоциирован с определенным местоположением. Однако, для учета случаев перемещения пользователя между городами или регионами, система позволяет пользователю предварительно указать планируемые даты и местоположение, чтобы обеспечить безопасный доступ. На рисунке 1 представлена диаграмма последовательности аутентификации пользователя сети при использовании менеджера паролей с дополнительным атрибутом аутентификации в виде его местоположения, которое он выбирает при регистрации. Местоположение представляет собой область координат, который пользователь указывает вручную или выбирает на карте область, в которой он будет использовать менеджер паролей. На рисунке 2 представлен принцип работы менеджера паролей, когда местоположение не подтверждено.

2. Отсутствие восстановления пароля по электронной почте. В рамках разработки менеджера паролей не будет предоставляться возможность восстановления пароля посредством электронной почты. Вместо этого, пользователю будет выдана seed-фраза, которую необходимо сохранить в безопасном месте, так как потеря фразы приведет к невозможности восстановления данных.

3. Дополнительные функции включают:

- Хранение не только паролей, но и пользовательских конфиденциальных данных.
- Оценка силы и безопасности текущих паролей пользователя, а также предложение более надежных альтернативных паролей.
- Слежение за историей изменений паролей, чтобы избежать использования ранее использованных паролей.
- Автоматическое уничтожение паролей после определенного количества попыток ввода (например, 8 попыток).
- Встроенный генератор надежных паролей.
- Двухфакторная аутентификация (2FA) для повышения безопасности входа [5].
- Функция SOS, которая способствует изменению всех паролей для дополнительной защиты пользователя.

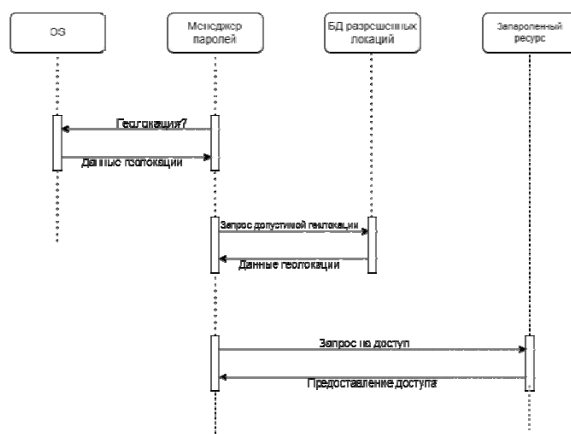


Рис. 1. Диаграмма последовательности в случае успешной аутентификации

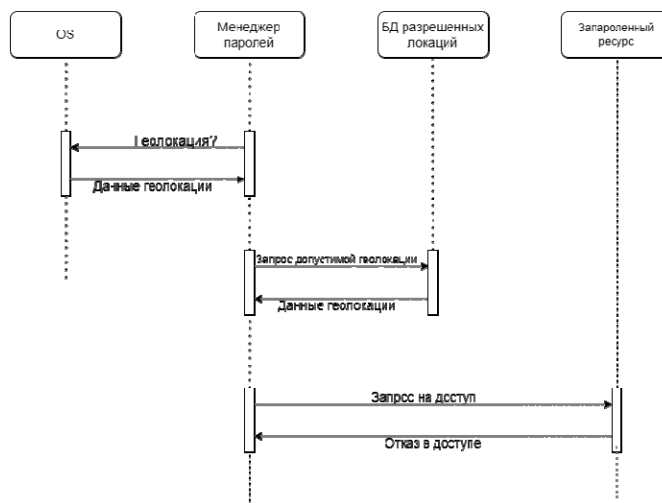


Рис. 2. Диаграмма последовательности в случае отказа в аутентификации

### Алгоритм работы

#### 1. Геолокационная защита:

- При регистрации пользователя программа требует указать местоположение, набор координат, которые пользователь указывает сам, либо выбирает на карте область.
- Во время каждого входа в аккаунт система записывает текущее местоположение пользователя.
- Если пользователь пытается войти в аккаунт из незнакомой точки (области координат), система автоматически блокирует доступ и отправляет уведомление на заранее указанный контакт.
- Для учета перемещений между городами или регионами, система предоставляет возможность пользователю предварительно указать планируемые даты и область координат, чтобы обеспечить безопасный доступ с учетом изменения геолокации.

#### 2. Отсутствие восстановления пароля по электронной почте:

- При регистрации пользователь создает мастер-пароль, система предоставляет пользователю seed-фразу, которую пользователь обязан сохранить в безопасном месте.
- В случае утери или забывания пароля, пользователь обращается к системе.
- Вместо сброса пароля через электронную почту, пользователь указывает seed-фразу.
- Seed-фраза служит для восстановления доступа к аккаунту в случае утери пароля.

#### 3. Дополнительные функции:

- Программа позволяет хранить не только пароли, но и конфиденциальные данные, такие как пин-коды, номера кредитных карт и другие секретные информационные данные.
- Программа оценивает силу пароля, базируясь на различных параметрах, таких как длина, использование разнообразных символов, чисел и букв разного регистра. Она также анализирует безопасность текущих паролей пользователя и, при необходимости, предлагает более надежные альтернативы.
- Ведется слежение за историей изменений паролей, чтобы предотвратить использование ранее использованных паролей и повторное их использование.
- Программа устанавливает автоматическое уничтожение паролей после заданного количества попыток ввода (8 попыток), что снижает риски брутфорс-атак.
- Встроенный генератор надежных паролей помогает пользователям создавать сложные и надежные пароли.
- Двухфакторная аутентификация (2FA) предоставляет дополнительный уровень безопасности при входе в аккаунт.
- Функция SOS позволяет пользователю изменить все пароли для дополнительной защиты в случае обнаружения угрозы или компрометации аккаунта.



- Функционал регистрации пользователя и процессы создания или добавления новых местоположений осуществляются в режиме онлайн. Это предполагает взаимодействие с центральным сервером, где регистрируются учетные записи, а также осуществляется сохранение и обновление геолокационных данных.
- Программа при входе требует подключение к интернету, после успешной авторизации, приложением можно пользоваться офлайн.

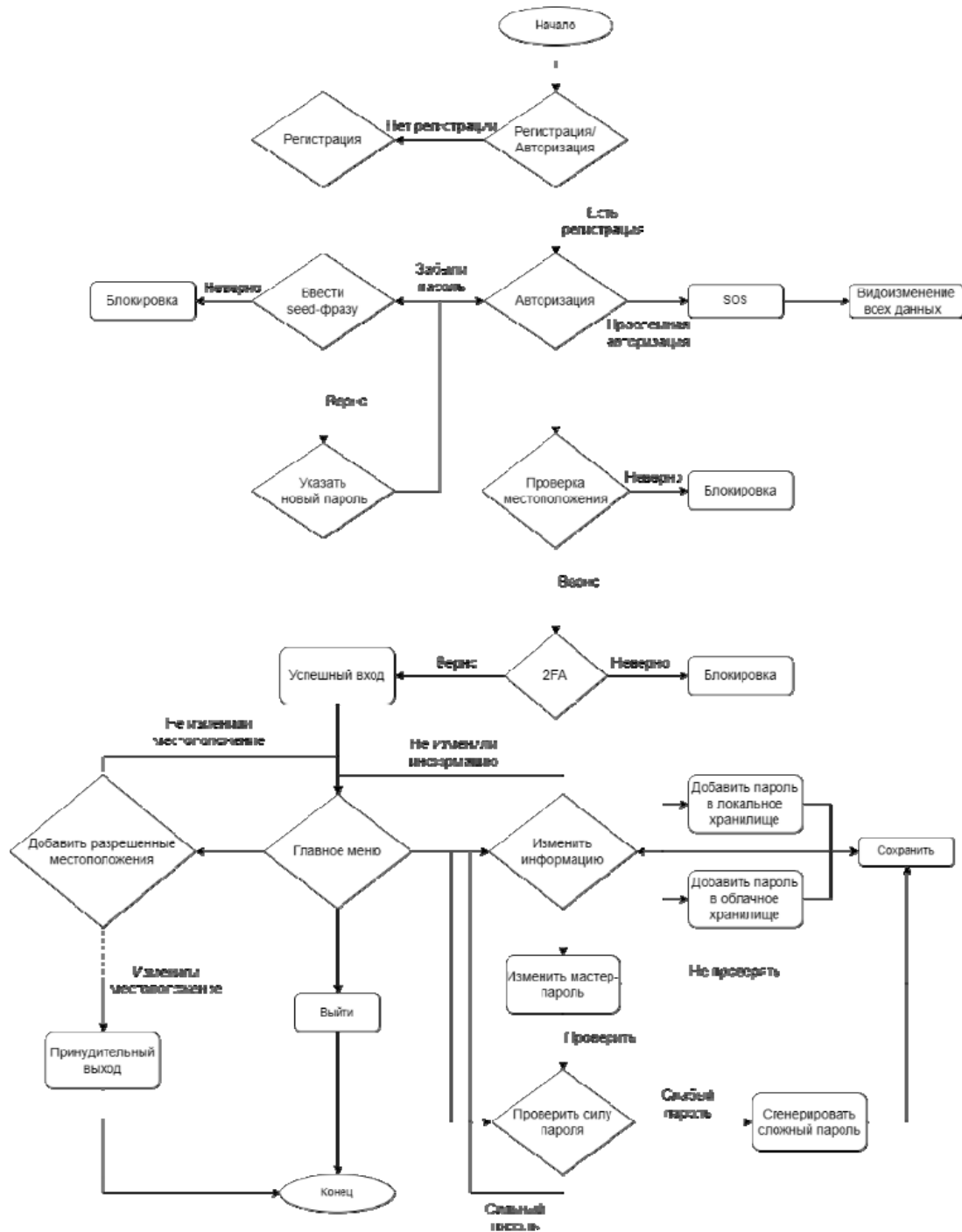


Рис. 3. Алгоритм работы программы

### Регистрация, авторизация, аутентификация пользователя

На рисунке 4 представлена блок-схема процессов регистрации, авторизации и аутентификации доступа пользователя к информационному ресурсу. При запуске приложения пользователю предоставляется выбор между двумя опциями: зарегистрироваться или авторизоваться. Если пользователь еще не имеет учетной записи, то при выборе регистрации, ему предлагается создать имя профиля и установить мастер-пароль [10]. При этом система также генерирует seed-фразу для дополнительной безопасности. Кроме того, система записывает и запоминает местоположение пользователя. В завершающей части регистрации включается двухфакторная аутентификация (2FA).

Если у пользователя уже есть учетная запись, то он вводит свой логин и пароль. После этого система проводит проверку местоположения пользователя. Если местоположение не соответствует ожидаемому, учетная запись блокируется, что помогает предотвратить несанкционированный доступ. После успешной проверки местоположения, пользователю необходимо пройти двухфакторную аутентификацию (2FA) для завершения процесса входа в систему.

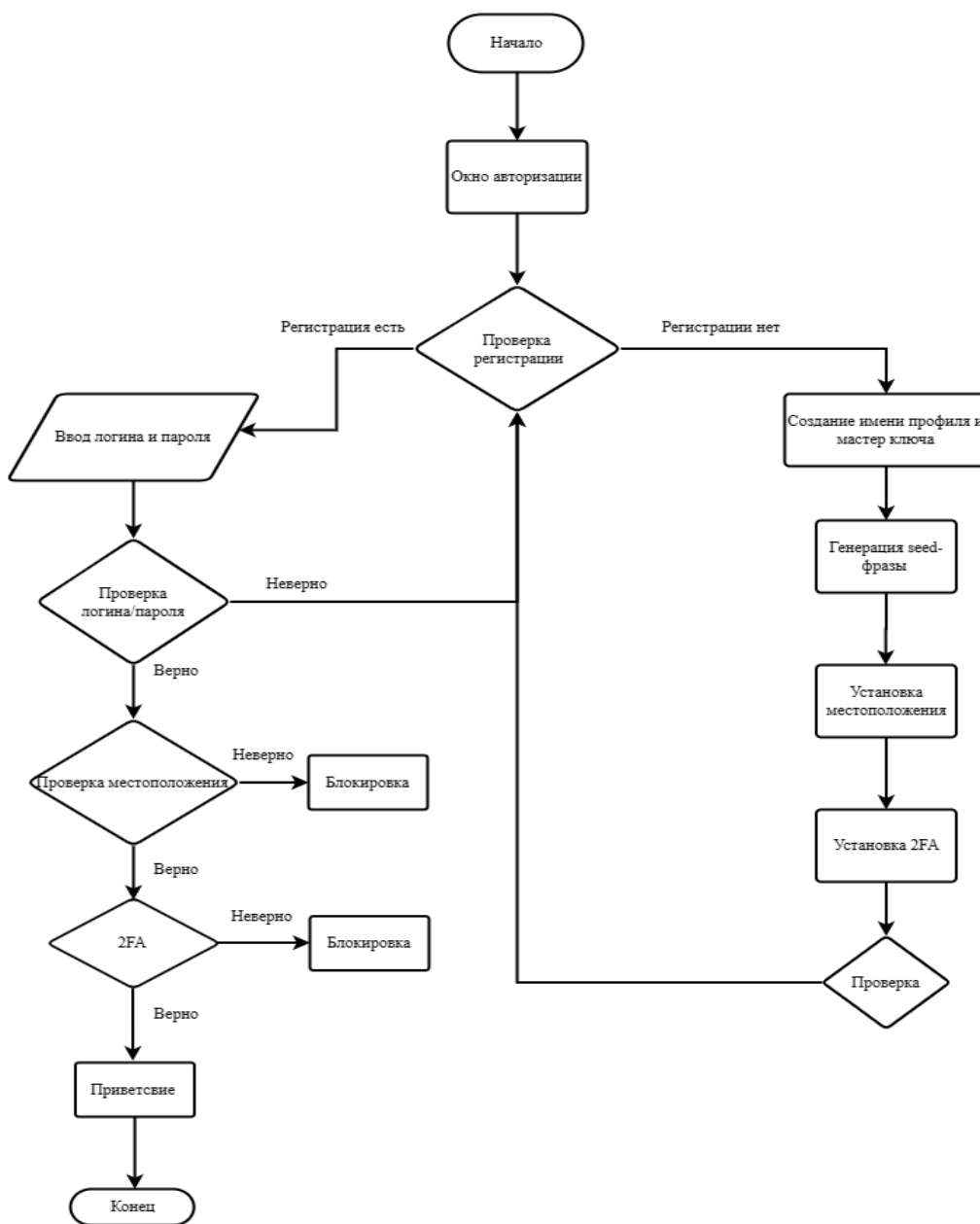


Рис. 4. Алгоритм работы процессов регистрации, авторизации, аутентификации пользователя

### Заключение

В современном цифровом мире, обеспечение безопасности данных становится неотъемлемой частью нашей жизни. В рамках постоянно развивающегося цифрового мира, где угрозы безопасности становятся все более сложными, управление паролями выходит на первый план как критически важный аспект обеспечения безопасности. Эта работа нацелена на трансформацию современного менеджера паролей, внедряя инновационные функции, которые не только повышают уровень безопасности, но и делают процесс управления паролями более интуитивным и удобным для конечного пользователя. Одним из ключевых направлений является внедрение геолокационной защиты. Теперь ПО будет активно адаптироваться к местоположению, обеспечивая дополнительный барьер для неправомерного доступа. Это решение позволяет системе реагировать на потенциальные угрозы в зависимости от физического расположения пользователя. Функции, представленные в данном ПО, не только повышают уровень безопасности данных, но и делают процесс управления паролями более инновационным и интуитивным.

### Литература

1. Словарь Академик Онлайн-версия словаря, доступная на сайте Академик (academic.ru). <https://dic.academic.ru/dic.nsf/ruwiki/1388435>
2. Федеральная служба по техническому и экспортному контролю (ФСТЭК). Официальный веб-сайт ФСТЭК (fstec.ru).
3. "Аудит безопасности" Банк данных угроз безопасности информации. [Интернет-ресурс]. <https://bdu.fstec.ru/threat-section/defenses/%D0%90%D0%A3%D0%94.1.1?ysclid=lnpvi1a8mj560087791>
4. "Требования по безопасности" Официальный веб-сайт ФСТЭК (fstec.ru). [Интернет-ресурс]. <https://fstec.ru/dokumenty/vse-dokumenty/spetsialnye-normativnye-dokumenty/trebovaniya-po-bezopasnosti-informatsii-utverzhdenny-prikazom-fstek-rossii-ot-14-aprelya-2023-g-n-64?ysclid=lnpvn3b5xq967850296>
5. Немного о 2FA: Двухфакторная аутентификация. <https://habr.com/ru/companies/1cloud/articles/277901/>
6. 1Password Официальный веб-сайт - <https://1password.com/ru>
7. Dashlane Официальный веб-сайт - <https://www.dashlane.com/>
8. RoboForm Официальный веб-сайт - <https://www.roboform.com/ru>
9. Что такое сид-фраза? <https://www.okx.com/ru/help/11177929096333?ysclid=lnuj5b869f465754053>
10. "Шифрование с мастер-паролем [Интернет-ресурс]. <https://yandex.ru/support/browser-passwords-crypto/with-master.html?ysclid=lnujg11xfc103550788>

## РАСПОЗНАНИЕ ТВЁРДЫХ КОММУНАЛЬНЫХ ОТХОДОВ С ИСПОЛЬЗОВАНИЕМ АЛГОРИТМА RT-DETR

**Медведев Евгений Юрьевич,**

Московский Технический Университет Связи и Информатики, магистрант, Москва, Россия  
[gn.medvedev@gmail.com](mailto:gn.medvedev@gmail.com)

**Воронова Лилия Ивановна,**

Московский Технический Университет связи и Информатики, д.ф.-м.н., профессор, Москва, Россия  
[voronova.lilia@yandex.ru](mailto:voronova.lilia@yandex.ru)

### Аннотация

В статье описывается процесс проектирования и обучения нейронной сети на основе SOTA алгоритма RT-DETR для классификации бытового мусора по шести классам. Рассматривается процесс обучения сети, формирование обучающих наборов данных и тестирование полученной модели. На основе анализа полученных результатов сделано заключение о возможности использования разработанной нейросети для промышленного применения в автоматизированных мусоросортировочных комплексах.

**Ключевые слова:** Компьютерное зрение, распознавание объектов, нейронная сеть, DETR, RT-DETR, машинное обучение, бытовой мусор, Transfer Learning, дообучение модели

### Введение

Переработка и утилизация огромного количества твёрдых коммунальных отходов превратились на сегодняшний день в глобальную общемировую проблему [1]. С каждым годом потребление ресурсов растёт, а вместе с ним растёт и количество производимого мусора [2]. В разных странах мира, используются различные варианты решения этой проблемы, от полной переработки на вторсырьё до захоронения на мусорных полигонах всего производимого мусора.

Внедрение методов технического зрения в сферу автоматического распознавания, классификации и роботизированной сортировки мусора, является в сложившейся в мире ситуации весьма перспективной идеей [3]. На данный момент первичная сортировка отходов производится зачастую вручную, что не может не сказаться на трудоёмкости и точности результата. Именно поэтому необходимо переложить эту задачу на программное обеспечение, использующее алгоритмы машинного обучения. Исследования в этом направлении ведутся и на кафедре ИСУиА МТУСИ [4-6, 13-16].

В данной статье описывается процесс разработки нейросети для комплексного обнаружения объектов мусора и использованием State-Of-The-Art алгоритма DETR [7] и двух наборов данных.

### Теоретическое описание алгоритма DETR

В последнее время, модели нейросетей использующие сквозные детекторы на основе трансформеров (DEtection TRansformers) (DETR) достигли выдающихся результатов и зачастую по производительности некоторых своих версий, они превосходят иные популярные модели, применяемые для обучения нейросетей, например, YOLO или Faster-RCNN. Архитектура моделей DETR представлена на рисунке 1.

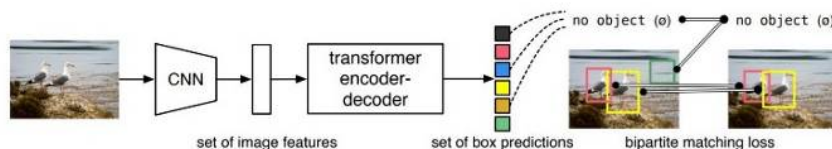


Рис. 1. Архитектура DETR [7]

### Преимущества DETR

- **Сквозное обнаружение объектов:** DETR предлагает комплексное решение для обнаружения объектов, устраняющее необходимость в отдельных сетях предложений регионов и этапах пост-

обработки. Это упрощает общую архитектуру и оптимизирует конвейер обнаружения объектов.

- *Параллельная обработка:* DETR прогнозирует классы объектов и ограничивающие рамки для всех объектов изображения одновременно благодаря архитектуре Transformer. Эта параллельная обработка приводит к более быстрому выводу по сравнению с последовательными методами.
- *Эффективное использование механизмов «self-attention»:* позволяет ему улавливать сложные отношения между объектами и их пространственным контекстом. Это приводит к повышению точности обнаружения объектов, особенно в сценариях с переполненными или перекрывающимися объектами.

#### Недостатки DETR

- *Высокие вычислительные ресурсы:* обучение и использование DETR может потребовать больших вычислительных ресурсов, особенно для больших моделей и изображений с высоким разрешением. Это может ограничить его доступность для исследователей и практиков, не имеющих доступа к мощному оборудованию.
- *Фиксированное количество запросов к объектам:* DETR требует заранее указать количество запросов к объектам, что может быть ограничением при работе со сценами, содержащими различное количество объектов. Неправильное количество запросов может привести к пропущенным обнаружениям или снижению эффективности.

Проблема высоких вычислительных затрат не была решена до недавнего времени, что ограничивало практическое применение моделей этого семейства и не позволяло им в полной мере использовать преимущества отсутствия постобработки, таких как «не максимальное подавление» (NMS) [8]. Однако в 2023 году, компанией Vaidi была разработана новая версия модели, под названием RT-DETR, в которой эта проблема была решена [9]. Она превосходит в скорости работы и точности другие детекторы объектов в реальном времени, что хорошо видно на графике, изображённом на рисунке 2, где показано сравнение модели RT-DETR и последних моделей YOLO при работе с тестовым набором данных COCO.

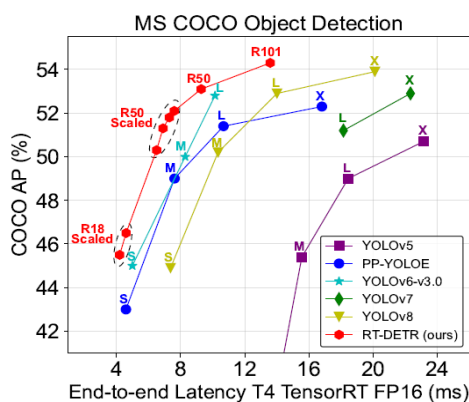


Рис. 2. Архитектура DETR в сравнении с YOLO [9]

### Проектирование и реализация нейронной сети на основе модели RT-DETR

#### 1. Подготовка наборов данных разрабатываемой нейросети.

Для обучения и последующего тестирования нейросети, были использованы два набора фотографий.

Первый датасет под названием «Trashnet» был взят из публичного хранилища наборов данных huggingface.co [10]. Этот набор данных состоит из 2467 неразмеченных изображений, включающих в себя шесть основных классов бытовых отходов, а именно: пластик, стекло, металл, картон, бумага и несортируемые отходы, разделённых в следующем соотношении:

- Cardboard 393 изображения
- Glass 491 изображение
- Metal 400 изображений
- Paper 584 изображений
- Plastic 472 изображения
- Trash 127 изображений

На рисунке 3 представлен коллаж из нескольких случайных элементов обучающего набора данных.



**Рис. 2.** Пример изображений классов из датасета «Trashnet»

Второй обучающий датасет под названием «My\_domestic\_trash» был сформирован вручную из фотографий типичных российских бытовых отходов. Этот набор данных состоит из 3300 неразмеченных изображений. Его структура повторяет предыдущий набор данных и также состоит из шести основных классов бытовых отходов. Пример изображений из каждого класса изображён на рисунке 4.



**Рис. 3.** Пример изображений классов из датасета «My\_domestic\_trash»

Для решения задачи обучения нейросети, было проведено аннотирование изображений обоих наборов данных с помощью коммерческого продукта Roboflow [12].

После аннотирования, следующим шагом является аугментация имеющихся изображений в проектах. Она применяется для увеличения исходного количества изображений, путём создания дополнительных изображений из уже имеющихся. Из-за небольшого количества исходных фотографий мусора, применяются одновременно сразу пять методов, а именно:

- Переворот по горизонтали
- Переворот по вертикали
- Поворот на 90° по часовой стрелке
- Поворот на 90° против часовой стрелки
- Поворот на произвольный градус в промежутке между -45° и +45°

Архитектура RT-DETR использует аналогичный YOLO разметочный формат в виде TXT файлов, с тем же именем, что и у соответствующего ему изображения в наборе данных и содержит запись вида <object-class> <x> <y> <width> <height>.

Дальнейшим шагом, после разметки и аугментации обоих датасетов, было проведено разделение всего объёма изображений на три набора: учебный, проверочный и тестовый в пропорциях 70%/20%/10%

2. Архитектура разрабатываемой сети

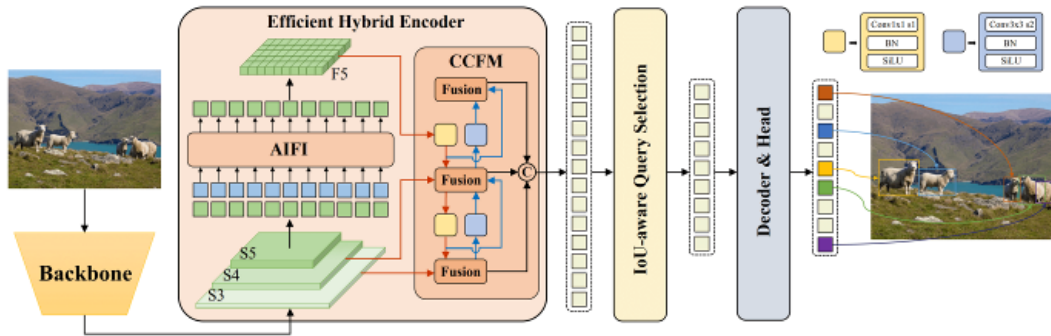


Рис. 4. Архитектура RT-DETR [9]

RT-DETR состоит из магистрали (backbone), гибридного шифратора (efficient hybrid encoder) и декодера-трансформера со вспомогательными головками обнаружения (decoder & auxiliary detection head). Обзор архитектуры модели показан на рисунке 4. Выходные характеристики трех последних этапов магистрали {S3, S4, S5} используются в качестве входных данных для шифратора. Гибридный шифратор преобразует многомасштабные признаки в последовательность признаков изображения путем внутримасштабного взаимодействия и кросс-масштабного слияния. Затем используется IoU-осознанный отбор запросов для выбора фиксированного числа признаков изображения из выходной последовательности шифратора, которые служат начальными объектными запросами для декодера. Наконец, декодер со вспомогательными головками обнаружения итеративно оптимизирует запросы к объектам для получения значений и оценок точности.

3. Описание процесса обучения и исследование результатов работы

Первое обучение проводилось на основе набора данных Trashnet, используя подход «переводного обучения» (transfer learning) на стандартную модель RT-DETR-L, в течение 50 эпох, а весь процесс обучения занял примерно 4 часа. После завершения процессов обучения и валидации обученной модели, были построены несколько графиков с основными метриками, которые позволяют оценить качество получившейся модели.

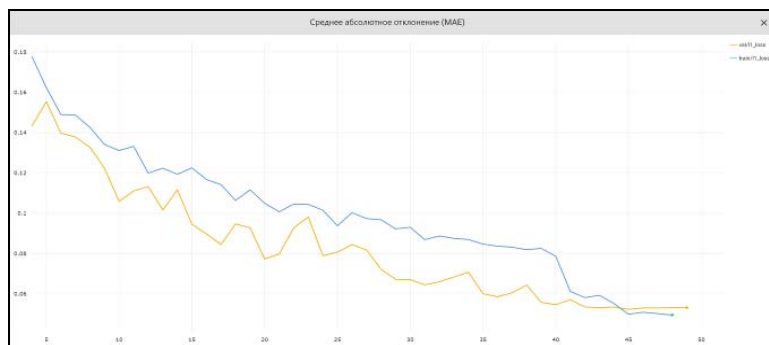


Рис. 5. Графики зависимостей значений MAE от эпохи обучения

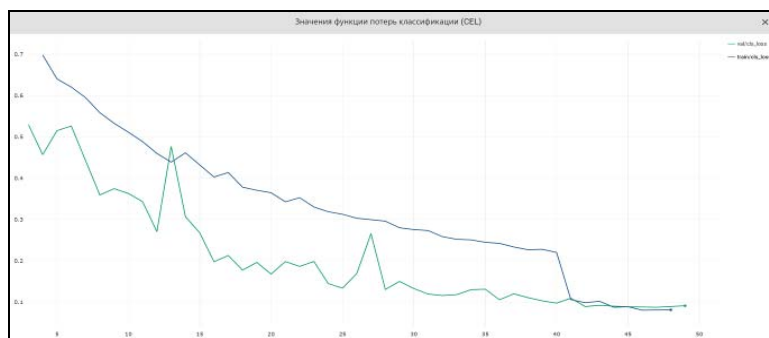


Рис. 6. Графики зависимости значений CEL от эпохи обучения

Графики значений среднего абсолютного отклонения (MAE) и функции потерь классификации (CEL) в результате обучения и валидации модели представлены на рисунках 5 и 6, где по осям X – эпохи обучения, а по осям Y – значения функций потерь. Согласно графикам, на 49 эпохе, значения среднего абсолютного отклонения (MAE) составили 0,04 для обучения и 0,05 для валидации. Значения функции потерь классификации (CEL) на 49 эпохе составили 0,08 для обучения и 0,09 для валидации.

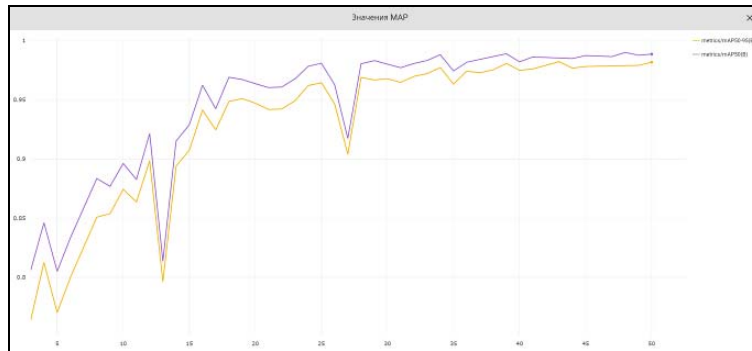


Рис. 7. Графики зависимости значений MAP от эпохи обучения

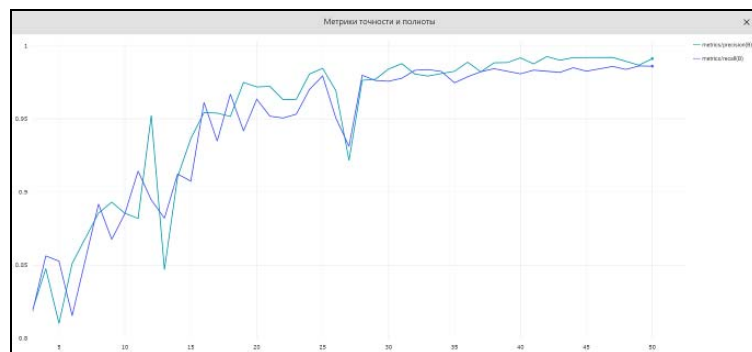


Рис. 8. Графики зависимости точности и полноты от эпохи обучения

Графики значений усреднённого значения средней точности (mAP) приведены на рисунке 7, где по оси X – эпохи обучения, а по оси Y – значения «mAP 50» и «mAP50-95». Максимальные значения для обеих метрик были достигнуты на 50 эпохе и составили 0,98 для обоих параметров.

Графики значений точности и полноты приведены на рисунке 8. Максимальные значения для обеих метрик были достигнуты на 50 эпохе и составили 0,99 для обоих параметров.

Далее была построена нормализованная матрица ошибок [12], чтобы сравнить реальные значения с прогнозами нейронной сети (на тестовом наборе). Данное действие показано на рисунке 9.

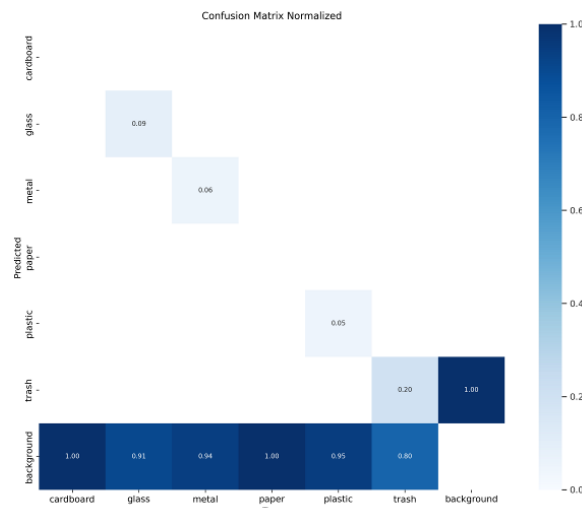


Рис. 9. Нормализованная матрица ошибок обученной модели



После анализа данной матрицы, можно сделать два вывода:

1. Хотя основные метрики обучения нейросети и показали неплохие результаты обучения, но фактически сеть почти со 100% вероятностью воспринимает все объекты мусора, как фоновые изображения и, следовательно, для продуктивной работы данная нейросеть не подходит.
2. Для устранения этого недостатка требуется обязательное дообучение полученной нейросети вторым набором данных «My\_domestic\_trash» для повышения качества распознавания объектов.

Дообучение полученной ранее модели, происходило с использованием аналогичных подходов и методов. Основой для обучения была взята ранее обученная модель, обученная на наборе данных «Trashnet», но в этом случае модель дообучалась на основе датасета «My\_domestic\_trash».

В результате нейросеть была обучена в течение 50 эпох, а весь процесс обучения занял примерно 4 часа. После завершения процессов обучения и валидации обученной модели, были построены несколько графиков с основными метриками, которые позволят оценить качество получившейся модели.

Графики значений среднего абсолютного отклонения (MAE) и функции потерь классификации (CEL), полученные по результатам дообучения и валидации ранее обученной модели представлены на рисунках 10 и 11, где по осям X – эпохи обучения, а по осям Y- значения функций потерь. Согласно графикам, на 49 эпохе, значения среднего абсолютного отклонения (MAE) составили 0,16 для обучения и 0,12 для валидации. Значения функции потерь классификации (CEL) на 47 эпохе составили 0,35 для обучения и 0,50 для валидации.

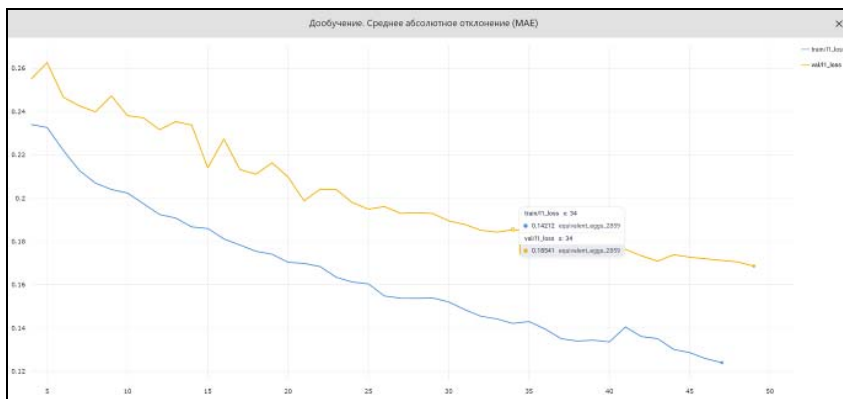


Рис. 10. Графики зависимости MAE от эпохи второго обучения

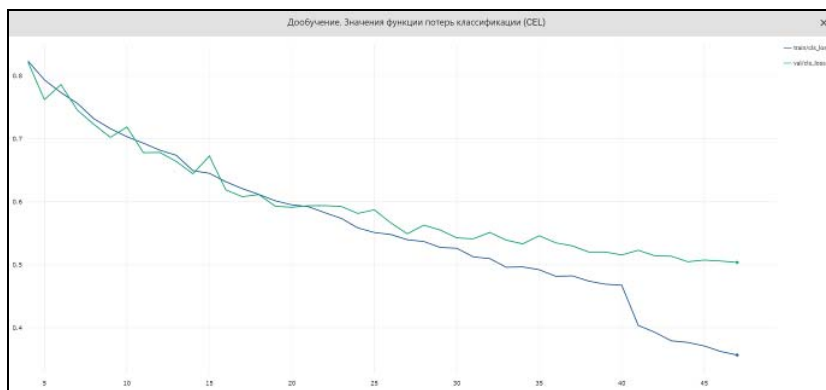


Рис. 11. Графики зависимости CEL от эпохи второго обучения

Графики значений усреднённого значения средней точности (mAP) приведены на рисунке 12, где по оси X – эпохи обучения, а по оси Y– значения «mAP 50» и «mAP50-95». Максимальные значения для обеих метрик были достигнуты на 47 эпохе и составили 0,88 для метрики mAP50 и 0,67 для метрики mAP50-95.

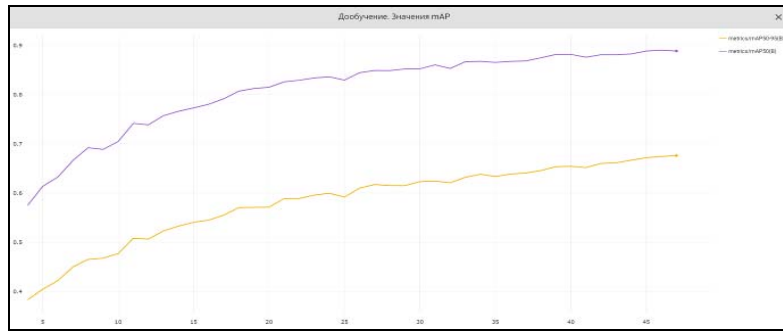


Рис. 12. Графики зависимости MAP от эпохи второго обучения

Графики значений точности и полноты приведены на рисунке 13. Максимальные значения для обеих метрик были также достигнуты на 47 эпохе и составили 0,90 для точности, и 0,88 для полноты дообученной модели.

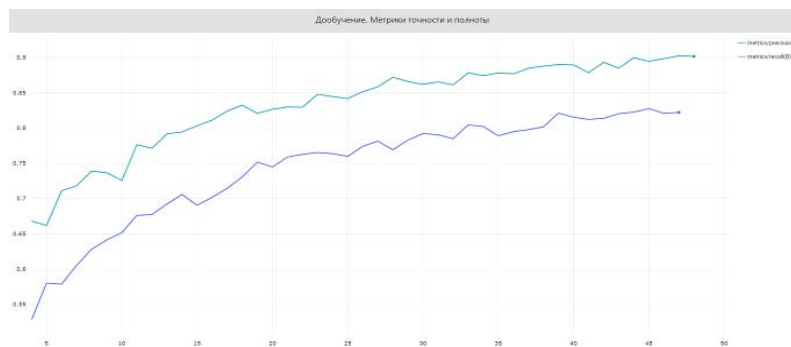


Рис. 13. Графики точности и полноты при дообучении модели

Далее, по результатам второго обучения, была построена матрица ошибок для дообученной модели, чтобы сравнить реальные значения с прогнозами нейронной сети (на тестовом наборе). Нормализованная матрица показана на рисунке 14.

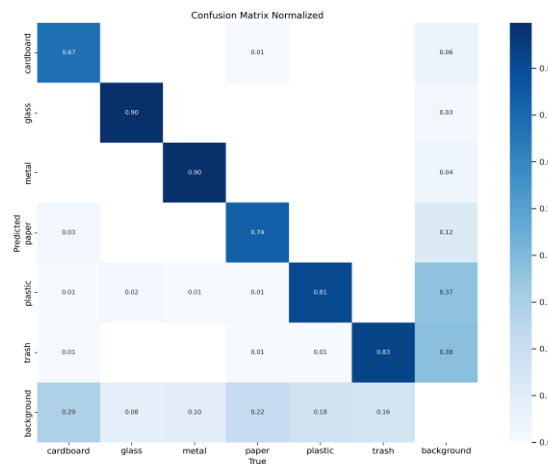


Рис. 14. Нормализованная матрица ошибок дообученной модели

После сравнительного анализа метрик, полученных после первого обучения и метрик, полученных после второго обучения, можно сделать вывод, что процедура дообучения позволяет кардинально улучшить обобщающие способности нейросети, а также повысить её точность, путём исправления потенциальных ошибок, допущенных в исходном обучении. Также дополнительные данные позволяют НС адаптироваться к новым условиям работы, т.к. объекты бытового мусора крайне вариативны по своей форме и структуре.

## Выводы

В статье приведён пример разработки и реализации нейросети по выбранной авторами предметной области, в частности выполнена реализация нейросети для распознавания и классификации объектов бытового мусора. Выбрана и рассмотрена наиболее современная, быстрая и точная архитектура сети из семейства моделей, использующих сквозные детекторы на основе трансформеров (DEtection TRansformers) (DETR) – RT-DETR. Подготовлены два набора учебных, проверочных и тестовых данных для разрабатываемой нейросети. Общее число изображений в исходных наборах данных составило более 5700 снимков.

Итогом данной работы является спроектированная и корректно работающая нейронная сеть, способная с высокой точностью распознавать и классифицировать объекты бытового мусора используя техническое зрение. Полученной точности модели хватает для решения поставленных перед ней задач.

В рамках проведения данных работ удалось добиться, при достижении количества эпох равному 50 итоговой точности модели равной 0.88 (88%) для метрики mAP50 и 0,67 (67%) для метрики mAP50-95. Увеличить точность нейросети путём увеличения наборов данных, применяемых при регулярных дополнительных обучениях модели. Следует учитывать, что максимальную эффективность будут иметь данные полученные непосредственно в реальных условиях на реальном мусоросортировочном комплексе.

Разработанная авторами нейронная сеть полностью оправдывает возможность своего дальнейшего промышленно применения в сфере роботизированной сортировки твёрдых коммунальных отходов, т.к. показала высокую точность работы и потенциал к увеличению своей точности.

### Литература

1. *Jérémie Cavé, Mathieu Durand, Adeline Pierrat.* Less matter, less waste: Taking our cue from the Global South // *Ecologie & politique.* Editions Le Bord de l'eau, 2020. С. 121-135.
2. *Ежеченко А.В.* Проблема мусора в современном мире // XXXV международные Плехановские чтения. М.: Российский экономический университет имени Г.В. Плеханова, 2022. С. 64-68.
3. *Менибаев М.Р.* Обзор и анализ современных систем роботизированной сортировки мусора // Лучшая студенческая статья 2023. 2023. С. 34-38.
4. *Лешин Е.В., Репинский В.Н., Воронова Л.И.* Распознавание изображений на основе технологии Tensorflow // Сборник трудов XV Международной отраслевой научно-технической конференции «Технологии информационного общества». 2021. С. 321-323.
5. *Вовик А.Г.* и др. Программный комплекс для дистанционного управления промышленным роботом kuka с применением алгоритмов компьютерного зрения.
6. *Шишканова Е.О., Воронова Л.И.* Применение нейронных сетей для распознавания продукции молочно-производства // Технологии информационного общества. 2022. С. 327-329.
7. *Dai Z. et al.* Up-detr: Unsupervised pre-training for object detection with transformers // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021. С. 1601-1610.
8. *Neubeck A., Van Gool L.* Efficient non-maximum suppression // 18th international conference on pattern recognition (ICPR'06). IEEE, 2006. Т. 3. С. 850-855.
9. *Ly W. et al.* Detsr beat yolos on real-time object detection // arXiv preprint arXiv:2304.08069. 2023.
10. *garythung / trashnet* // Dataset of images of trash; Torch-based CNN for garbage image classification URL: <https://github.com/garythung/trashnet> (дата обращения: 26.01.2024).
11. Everything you need to build and deploy computer vision models. // Roboflow.com URL: <https://roboflow.com/> (дата обращения: 26.01.2024).
12. *Marom N.D., Rokach L., Shmilovici A.* Using the confusion matrix for improving ensemble classifiers // 2010 IEEE 26-th Convention of Electrical and Electronics Engineers in Israel. IEEE, 2010. С. 000555-000559.
13. *Мохаммад Н., Воронова Л.И., Воронов В.И.* Разработка имитационной модели использования роя беспилотных летательных аппаратов в сельском хозяйстве // Научные технологии в космических исследованиях Земли. 2022. Т. 14. № 3. С. 55-61.
14. *Воронов В.И., Быков А.Д., Воронова Л.И.* Проектирование подсистемы детектирования лиц и интерфейса работы с базой данных в программно-аппаратном комплексе биометрической идентификации на основе нейросетевого распознавания лиц // Т-Comm: Телекоммуникации и транспорт. 2020. Т. 14. № 4. С. 31-38.
15. *Voronov V.I., Zharov I.A., Bykov A.D., Trunov A.S., Voronova L.I.* Designing a neural network identification subsystem in the hardware-software complex of face recognition // Т-Comm. 2020. Т. 14. № 5. С. 69-76.
16. *Трунов А.С., Воронов В.И., Воронова Л.И.* Интеграция унаследованных приложений в экосистему научных исследований "ЭНИ" // Т-Comm: Телекоммуникации и транспорт. 2020. Т. 14. № 8. С. 33-41.

# ПРОЕКТИРОВАНИЕ ПОДСИСТЕМЫ ВНЕШНЕГО СИТУАТИВНОГО УПРАВЛЕНИЯ РОБОТОТЕХНИЧЕСКИМ КОМПЛЕКСОМ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

**Палагушин Александр Дмитриевич,**  
МТУСИ, магистрант, Москва, Россия  
[a.palagushin@yandex.ru](mailto:a.palagushin@yandex.ru)

**Воронов Вячеслав Игоревич,**  
МТУСИ, доцент, к.т.н., Москва, Россия  
[v.i.voronov@mtuci.ru](mailto:v.i.voronov@mtuci.ru)

## Аннотация

В статье описывается проектирование подсистемы внешнего ситуативного управления робототехнического комплекса на базе шестиосевого промышленного манипулятора, выполняющей следующие задачи: детектирование и идентификация объекта в рабочей зоне робототехнического комплекса с применением системы компьютерного зрения, определение местоположения объекта в трехмерном пространстве рабочей зоны робототехнического комплекса, подведение захвата манипулятора к объекту с учетом реальных координат объекта; захват и перемещение детали с помощью манипулятора.

**Ключевые слова:** компьютерное зрение, подсистема внешнего ситуативного управления, робототехнический комплекс, нейронная сеть, промышленная автоматизация.

## Введение

В настоящее время компьютерное зрение набирает популярность, как средство решения широкого круга задач. Это подтверждается большим количеством научных работ, в том числе проводимых на кафедре ИСУиА МТУСИ [1, 2, 12-15].

В промышленности компьютерное зрение может успешно использоваться в составе робототехнических комплексов (РТК), а именно в подсистемах внешнего ситуативного управления. Подсистема внешнего ситуативного управления позволяет взаимодействовать непосредственно с системой управления РТК по различным каналам связи и выполнять команды, полученные из какой-либо подсистемы принятия решений, которая в свою очередь, может базироваться на методах искусственного интеллекта (ИИ). Компьютерное зрение использует в своем составе набор методов искусственного интеллекта (ИИ), а наиболее продвинутые решения основываются на нейронных сетях. Задача исследования - проектирование подсистемы ситуативного управления РТК с использованием методов искусственного интеллекта. В процессе решения задачи создана архитектура проектируемого программно-аппаратного комплекса, а также разработана структура управляющего программного обеспечения.

## Архитектура программно-аппаратного комплекса подсистемы внешнего ситуативного управления РТК

На рисунке 1 показана архитектура разрабатываемого программно-аппаратного комплекса в нотации UML в виде диаграммы развертывания (Deployment Diagram) [3, 4].

Разработанный программно-аппаратный комплекс подсистемы внешнего ситуативного управления РТК состоит из системы технического зрения и подсистемы принятия решений. Система управления сообщается с устройством управления РТК посредством локальной сети (протокол TCP/IP) [5].

Система технического зрения в подсистеме управления выполняет функции детекции объекта в общей сцене и включает в себя следующие компоненты:

- Аппаратная часть:
  - Вычислительное устройство (ЭВМ);
  - Стереокамера;
- Программная часть:
  - Модуль захвата изображения;
  - Модуль предварительной обработки изображения;
  - Модель ИИ.

Вычислительное устройство в системе технического зрения представляет собой компьютер на базе процессора с архитектурой x86-64 или Arm64 (aarch64) с тактовой частотой не менее 2 ГГц, ОЗУ не менее 8 Гб, а также видеоадаптером Nvidia с поддержкой CUDA. Такая конфигурация рекомендована разработчиком нейронной сети для оптимальной и быстрой работы модели ИИ [6]. На вычислительном устройстве должна быть установлена 64-разрядная ОС с необходимой поддержкой требуемых программных компонентов, а именно Microsoft Windows 10 (WSL) и новее или UNIX-подобные системы [8]. Такая конфигурация обеспечит поддержку необходимых фреймворков для работы с техническим зрением и моделью ИИ.

Стереокамера в системе технического зрения должна иметь разрешение не менее HD 720p, но не более FullHD 1080p 30 кадров в секунду для получения достаточно детальных изображений с необходимой скоростью. Учитывая вышеуказанные требования к вычислительной мощности, а также примеры использования аналогичных аппаратных мощностей [8] предварительно данный аппаратный стек должен обеспечивать производительность более 2 детекций в секунду.

Программная часть подсистемы технического зрения представляет собой исходный код на языке Python, а также набор библиотек и фреймворков в том числе для работы с моделями на основе машинного обучения.

Модуль захвата изображения осуществляет захват двух изображений со стереокамеры, из которого кадры поступают на модуль предварительной обработки изображения, где производится коррекция изображений в соответствии с заранее проведенной калибровкой каждой из камер в стереокамере и друг относительно друга (стереоректификация) [9]. Также, при необходимости, возможна коррекция изображений в части цветности (перевод в монохромную цветовую гамму) и яркости и контрастности.

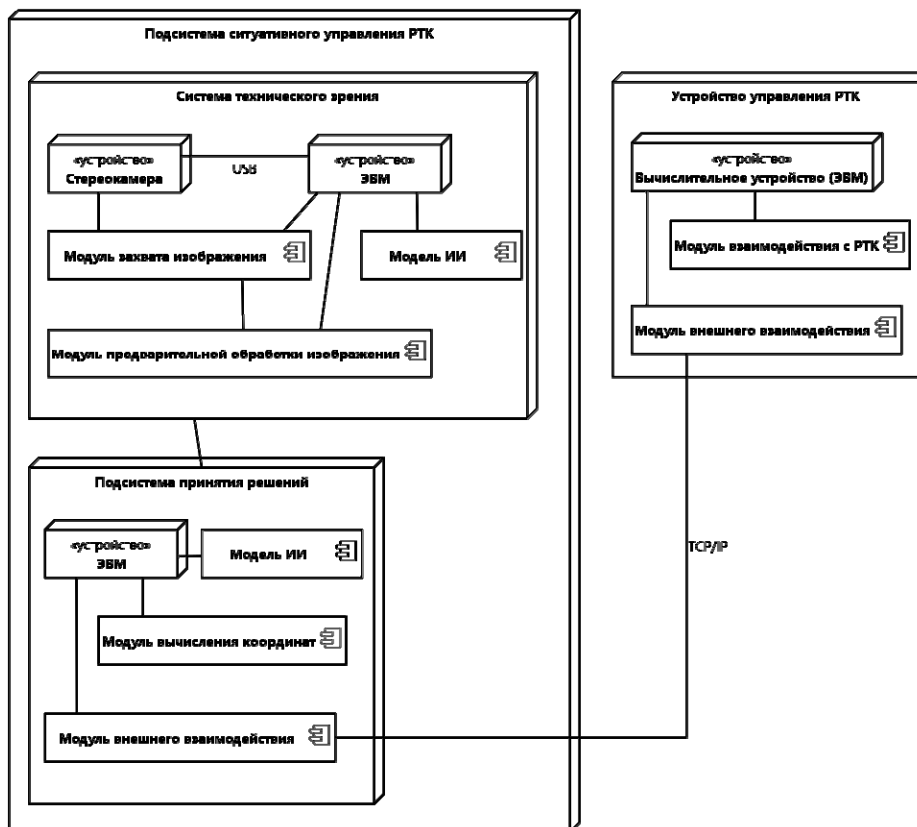


Рис. 1. Архитектура программно-аппаратного комплекса подсистемы внешнего ситуативного управления РТК (нотация UML, диаграмма развертывания)

После захвата и предварительной обработки кадры поступают на вход модели ИИ, основанной на сверточной нейросети, которая в свою очередь позволяет определить, какая из деталей сейчас находится на конвейере и выполнить сегментацию изображения для последующей детальной обработки в подсистеме принятия решений.

Подсистема принятия решений отвечает за вычисление координат точки захвата детали хватом манипулятора РТК и включает в себя следующие компоненты:

- Аппаратная часть:
  - Вычислительное устройство (ЭВМ);
- Программная часть:
  - Модель ИИ;
  - Модуль вычисления координат;
  - Модуль внешнего взаимодействия.

Вычислительное устройство для подсистемы принятия решений может быть тем же, что и для системы технического зрения. В таком случае необходимо в дополнение к ранее заявленным требованиям необходимо иметь ОЗУ не менее 16 Гб и более одного ядра CPU, поскольку при данных условиях на одной ЭВМ вторая модель ИИ.

Модель ИИ в программной части подсистемы принятия решений основана на сверточной нейронной сети и позволяет выделить опорные грани по сегментированному изображению детали, полученной из системы технического зрения.

По опорным граням, полученным из модели ИИ в модуле вычисления координат, происходит подгонка трехмерной модели детали к текущему положению детали на конвейере, что позволяет вычислить координаты, а также углы поворота модели по осям. Кроме того, в данном модуле предусмотрено определение точки захвата детали хватом манипулятора, что в совмещении с вычисленными координатами детали позволяет получить координаты и углы поворота для точки захвата.

Совместно с координатами из модуля вычисления координат в устройство управления РТК посредством модуля внешнего взаимодействия передаются команды по перемещению хвата манипулятора, захвата и перемещения детали в пространстве. Обмен данными осуществляется через локальную сеть по протоколу TCP/IP.

В устройстве управления РТК также предусмотрен модуль внешнего взаимодействия, который предоставляет возможность использовать команды для управления РТК из внешних подсистем. В данной работе модуль позволяет связать систему управления РТК с подсистемой внешнего ситуативного управления РТК. В свою очередь в устройстве управления РТК решаются задачи по траектории перемещения манипулятора с учетом прямоугольной системы координат, результаты которых используются в модуле взаимодействия с РТК для непосредственного управления манипулятором.

Структура программного обеспечения подсистемы внешнего ситуативного управления РТК

Структура программного обеспечения разрабатываемой подсистемы внешнего ситуативного управления описана с помощью нотации С4 [10]. Предполагается реализация ПО подсистемы внешнего ситуативного управления РТК на языке программирования Python, как наиболее распространённого на настоящий момент языка с точки зрения работы с данными и машинным обучением [11].

В обобщенном виде с точки зрения взаимодействия с окружающей средой разрабатываемая подсистема представлена на рисунке 2 в виде диаграммы контекста.

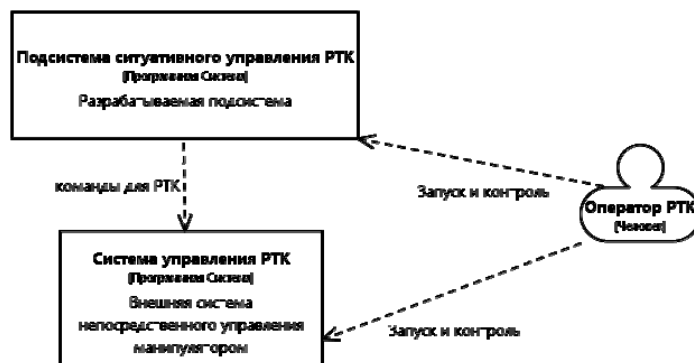


Рис. 2. Взаимодействие разрабатываемой подсистемы с окружающей средой (Диаграмма контекста)

Как видно из рисунка 2, участие человека в системе сведено к минимуму. Поскольку парадигмой разрабатываемой подсистемы ситуативного управления РТК является использование гибких производственных систем, задачи оператора в данном производственном цикле сводятся к запуску системы управления РТК и подсистемы ситуативного управления РТК и контроль работы данных систем.

Результатом декомпозиции схемы, указанной на рисунке 2, является следующий уровень – диаграмма контейнеров (рис. 3), на которой представлены модули, аналогичные модулям, представленным на рисунке 1.

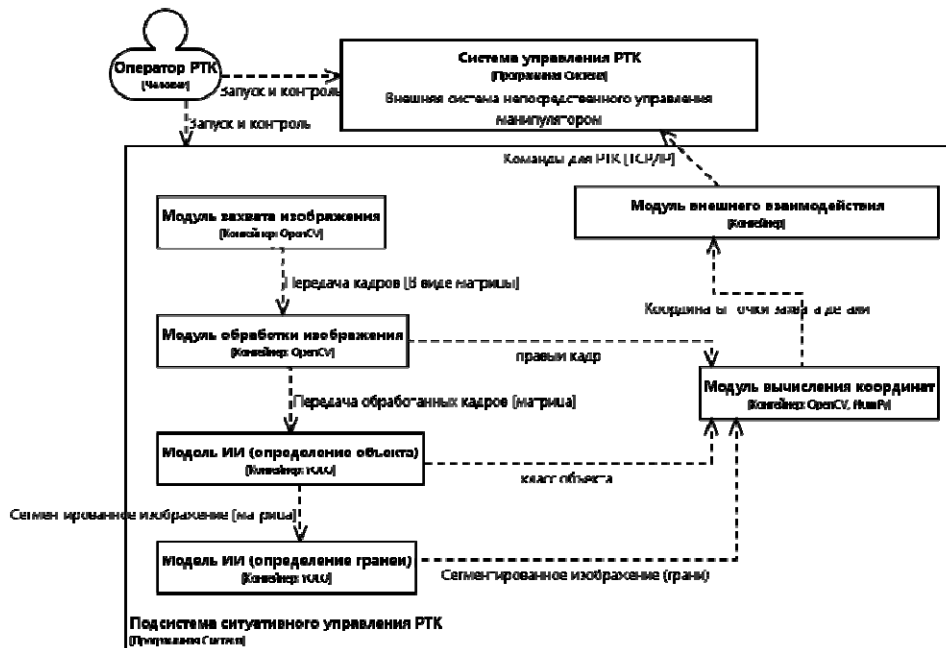


Рис. 3. Структура программного обеспечения (Диаграмма контейнеров)

Компоненты подсистемы внешнего ситуативного управления РТК:

- модуль захвата изображения;
- модуль обработки изображения;
- модель ИИ для детекции объекта;
- модель ИИ для детекции граней объекта;
- модуль вычисления координат;
- модуль внешнего взаимодействия.

Модуль захвата изображения отвечает за получение изображений со стереокамеры и состоит из считывателей изображений на основе функций библиотеки OpenCV, которые захватывают изображения с левой и правой камер и передают их далее в модуль обработки изображения (Рис. 4).

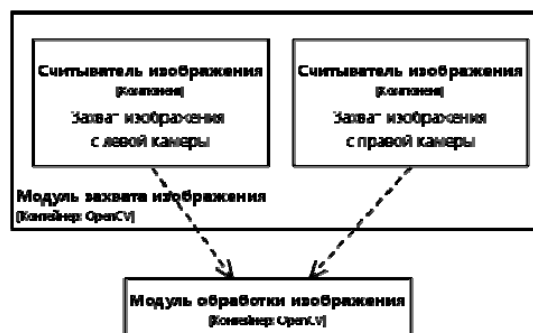


Рис. 4. Структура модуля захвата изображения (Диаграмма компонентов)

В модуле обработки изображения (рис. 5) на основе OpenCV происходит преобразование изображений в монохромные цвета для ускорения обработки, а далее кадры попадают в корректор изображения, где изображения корректируются с учетом заранее подготовленной калибровочной матрицей для каждой из камер стереопары, что позволяет устранить дисторсию изображения, связанную с особенностями объективов стереопары. После преобразований и коррекции правый кадр со стереопары отправляется в модуль вычисления координат, а левый кадр со стереопары поступает на вход модели ИИ для детектирования объекта.

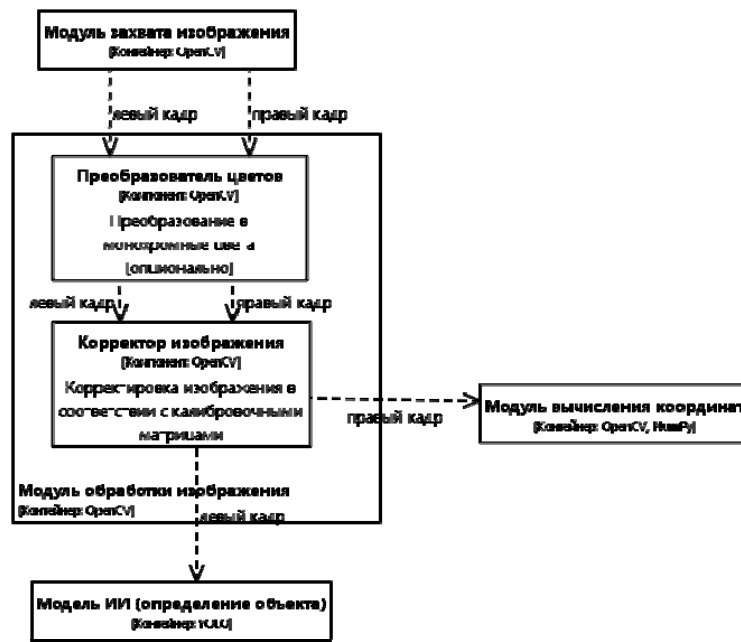


Рис. 5. Структура модуля обработки изображения (Диаграмма компонентов)

В модели ИИ для детектирования объекта поступивший кадр из модуля обработки изображения поступает на вход детектора объектов на основе нейросети YOLOv8 (рис. 6). В детектор объектов с помощью загрузчика параметров из файла подгружаются параметры (веса) модели, обученной на специфическом датасете. Детектор объектов позволяет получить класс объекта по изображению, а также сегментированное изображение для дальнейшего детектирования граней объекта и вычисления координат объекта. Сегментированное изображение и название класса передается далее в модель ИИ для детектирования граней, в модуль вычисления координат из данной модели поступает только класс объекта.

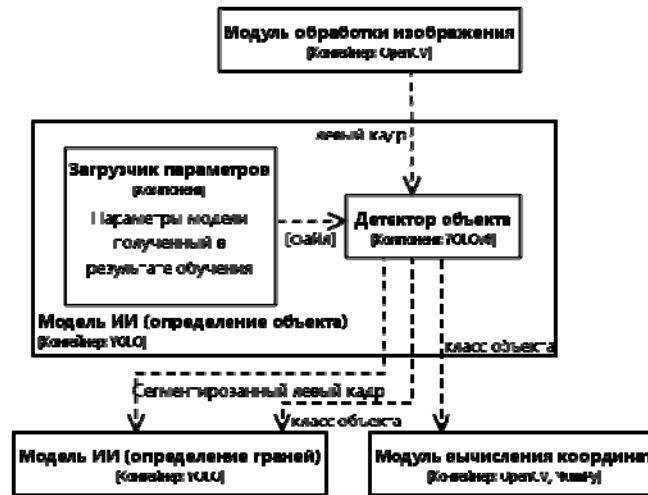


Рис. 6. Схема модели ИИ для детектирования объекта (Диаграмма компонентов)

В модели ИИ для детектирования граней на основе YOLOv8 (рис. 7) на поступившем сегментированном изображении в детекторе граней выделяются грани объекта. В детектор граней с помощью загрузчика параметров, на который из предыдущей модели поступает информация о классе объекта, из файла загружаются определенные веса модели (по заранее обученной модели) в соответствии с определенным классом объекта. Например, если это шестерня, то будут загружены параметры модели детектирования граней для шестерни. По результатам работы детектора граней на изображении выделяются грани объекта и определяется их класс.



Далее, после детектирования граней объекта сегментированное изображение с информацией о гранях и класс граней объекта передается в модуль вычисления координат.



Рис. 7. Схема модели ИИ для детектирования граней объекта (Диаграмма компонентов)

В модуле вычисления координат (Рис. 8) происходит определение координат захвата объекта манипулятором. Данные с модуля обработки изображения (левый и правый кадры со стереокамеры) поступают на генератор облака точек. В генераторе облака точек происходит дополнительная корректировка изображений на основании калибровочной матрицы, полученной заранее в результате стереоректификации камер и поступающей из загрузчика параметров, и далее создается облако точек объекта полного кадра, которое в свою очередь передается на вход калибратора 3D модели. В калибратор 3D модели также поступает облако точек из загрузчика 3D модели, сгенерированное на основании трехмерной модели объекта.

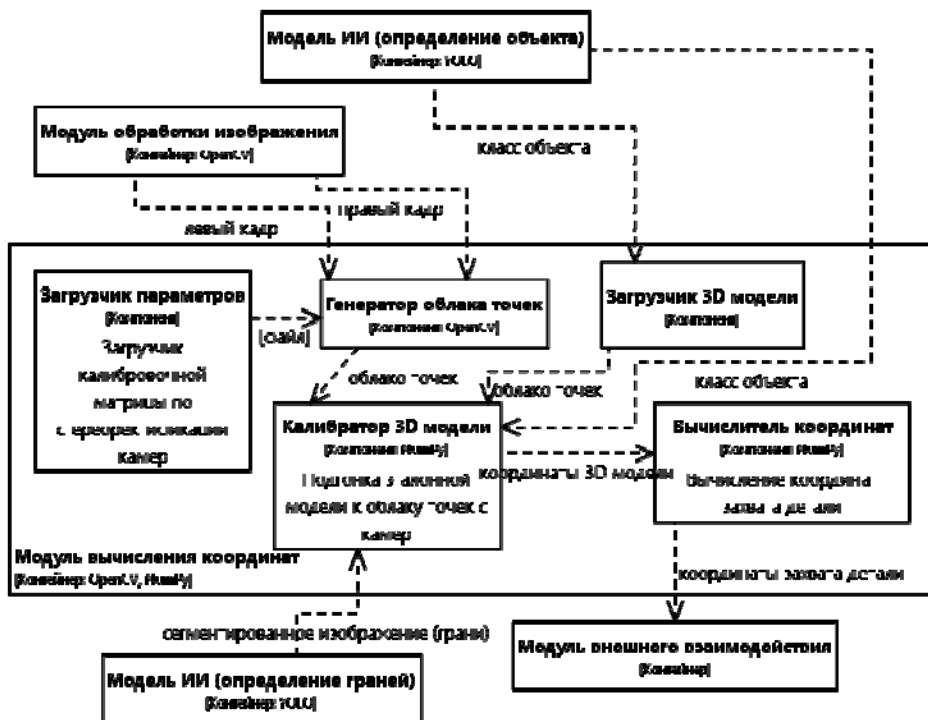


Рис. 8. Структура модуля вычисления координат (Диаграмма компонентов)

Загрузчик 3D модели в свою очередь принимает информацию о классе от модели ИИ, отвечающей за определение объекта и на основе этих данных загружает трехмерную модель в соответствии с классом. После получения всех необходимых данных, включая информацию о гранях объекта, в калибраторе 3D модели запускается процесс подгонки эталонной модели к облаку точек, полученных с камер. В результате выполнения данной процедуры возвращаются относительные координаты 3D модели, которые далее передаются в вычислитель координат. Одновременно в вычислитель координат с загрузчика 3D модели передаются предустановленные данные о точке захвата детали. На основании полученных данных вычислитель координат возвращает относительные координаты захвата детали, которые затем передаются на вход модуля внешнего взаимодействия.

В модуле внешнего взаимодействия (рис. 9) на основании поступивших данных из модуля вычисления координат происходит сопоставление полученных координат с СК системы управления РТК, которая в свою очередь выполняется в корректоре координат. Корректор координат возвращает скорректированные с учётом СК манипулятора координаты захвата детали и передает их на вход генератора команд РТК. В генераторе команд РТК на основании полученных координат точки захвата детали создается программа движения, захвата и перемещения детали манипулятором в виде команд для СУ РТК. Далее указанные команды поступают в передатчик команд РТК, откуда уже по каналу связи TCP/IP поступают в систему управления РТК.

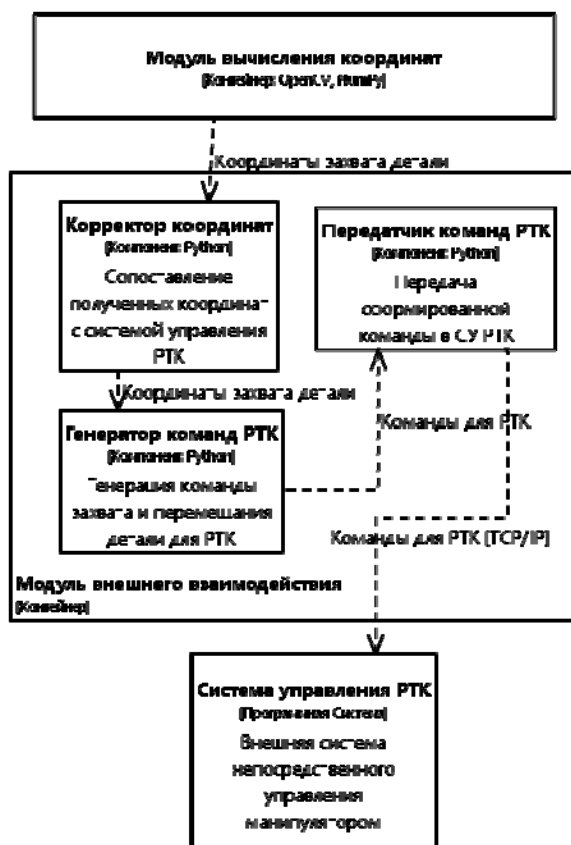


Рис. 9. Структура модуля внешнего взаимодействия (Диаграмма компонентов)

### Заключение

В исследовании приведены результаты проектирования подсистемы внешнего ситуативного управления на базе системы компьютерного зрения. Проектируемая подсистема построена на базе стереокамеры с использованием методов искусственного интеллекта, а также математического алгоритма, позволяющего вычислить местоположение объекта в рабочей зоне робототехнического комплекса. Реализация подсистемы осуществляется на базе языка Python с использованием библиотек с открытым исходным кодом.

Проектируемая подсистема внешнего ситуативного управления робототехническим комплексом позволит решить задачу правильной ориентации и захвата объекта манипулятором, а также его перемещение согласно заранее заданной программы.

Такой подход может быть полезен в задачах сортировки предметов, когда объекты в сортировочной корзине могут располагаться с произвольным наклоном относительно базовой плоскости (плоскости конвейера или рабочей зоны манипулятора). В настоящее время подсистема реализуется на экспериментальном стенде в лаборатории робототехники кафедры ИСУиА МТУСИ.

## Литература

1. Палагушин А.Д., Воронов В.И. Нейросетевое распознавание водоупоров по литологическим описаниям геологических слоев // DSPA: Вопросы применения цифровой обработки сигналов. 2022. Т. 12, № 2. С. 23-31.
2. Палагушин А.Д., Воронов В.И. Анализ возможности применения 3D-технического зрения в работе робототехнического комплекса // Технологии информационного общества : Сборник трудов XVII Международной отраслевой научно-технической конференции, Москва, 02-03 марта 2023 г. М.: Издательский дом Медиа Паб-лишер, 2023. С. 294-297.
3. ISO/IEC 19505-1:2012 Information technology - Object Management Group Unified Modeling Language (OMG UML) - Part 1: Infrastructure.
4. ISO/IEC 19505-2:2012 Information Technology - Object Management Group Unified Modelling Language (OMG UML) - Part 2: Superstructure.
5. Eddy W. (ed.). RFC 9293: Transmission Control Protocol (TCP). 2022.
6. Frequently Asked Questions (FAQ) - Ultralytics YOLOv8 Docs. URL: <https://docs.ultralytics.com/help/FAQ/>
7. Download Python | Python.org. URL: <https://www.python.org/downloads/>
8. Varma V. et al. Real time detection of speed hump/bump and distance estimation with deep learning using GPU and ZED stereo camera // Procedia computer science. 2018. Т. 143. С. 988-997.
9. Чиров А.Н., Сыроева Л.П., Архипов П.К. Калибровка стереоскопической системы // Актуальные проблемы авиации и космонавтики. 2021. С. 60-62.
10. Brown S. The C4 model for software architecture // Updated August. 2018. Т. 1.
11. Мамедов В.С. Использование языка программирования python для создания искусственных нейронных сетей // Современные условия взаимодействия науки и техники. 2017. С. 95-98.
12. Мохаммад Н., Воронова Л.И., Воронов В.И. Разработка имитационной модели использования роя беспилотных летательных аппаратов в сельском хозяйстве // Научные технологии в космических исследованиях Земли. 2022. Т. 14. № 3. С. 55-61.
13. Воронов В.И., Быков А.Д., Воронова Л.И. Проектирование подсистемы детектирования лиц и интерфейса работы с базой данных в программно-аппаратном комплексе биометрической идентификации на основе нейросетевого распознавания лиц // Т-Comm: Телекоммуникации и транспорт. 2020. Т. 14. № 4. С. 31-38.
14. Voronov V.I., Zharov I.A., Bykov A.D., Trunov A.S., Voronova L.I. Designing a neural network identification subsystem in the hardware-software complex of face recognition // Т-Comm. 2020. Т. 14. № 5. С. 69-76.
15. Трунов А.С., Воронов В.И., Воронова Л.И. Интеграция унаследованных приложений в экосистему научных исследований "ЭНИ" // Т-Comm: Телекоммуникации и транспорт. 2020. Т. 14. № 8. С. 33-41.

# ПРИМЕНЕНИЕ МЕТОДОВ МНОГОМЕРНОГО СТАТИСТИЧЕСКОГО АНАЛИЗА И NLP ДЛЯ КЛАССИФИКАЦИИ НАУЧНЫХ ПУБЛИКАЦИЙ

**Синева Ирина Сергеевна,**

*Московский технический институт связи и информатики, доцент, к.ф.-м.н., Москва, Россия*

[iss@mtuci.ru](mailto:iss@mtuci.ru)

**Головченко Владислав Евгеньевич,**

*Московский технический институт связи и информатики, Москва, Россия*

[vlgv.py@yandex.ru](mailto:vlgv.py@yandex.ru)

## **Аннотация**

*Данная работа посвящена использованию методов статистического анализа и моделей обработки естественного языка (NLP) для решения задачи классификации тематики публикаций на основе анализа их аннотаций. Показано эффективное использование для предобработки текстов его очистки (удаление знаков препинания, стоп-слов), токенизации, векторизации. Для классификации отобраны алгоритмы  $k$ -means и иерархический на основе метода Ward'a. Подбор числа классов выполнен на основе анализа силуэтов. Полученные классы носят контекстно-осмысленный характер.*

**Ключевые слова:** классификация текстов, NLP, TF-IDF, закон Ципфа–Мандельброта, метод силуэта

## **Введение**

Развитие и использование моделей обработки естественного языка (Natural Language Processing, NLP) генерирует, в том числе, большое количество докладов на конференциях и статей в научных журналах. Плюс к этому сам научный процесс в силу своей универсальности требует открытого обмена идеями и результатами. Ориентироваться в этом массиве больших данных становится все сложнее. Нужны инструменты отбора материалов по интересующей проблематике, при этом с предметной областью все обстоит проще. Сложнее ориентироваться в методике и инструментариях, поскольку полезные перспективные подходы могут все шире применяться в одних задачах и быть совершенно не известны в других, несмотря на свою универсальность в целом. Данная работа посвящена использованию методов статистического анализа и NLP для решения задачи кластеризации публикаций на основе анализа их аннотаций.

Классификация и выявление тематик в научных публикациях - одна из важнейших задач наукометрии, поскольку она имеет большое значение для ее практического применения. Точное определение тематик необходимо для нормализации цитирований, построения журнальных метрик, таких как квартили, и любых осмысленных сопоставлений ученых, организаций и журналов. В этой области существует множество наработок, но на практике в стандартных наукометрических инструментах, таких как Web of Science или Scopus, используется наиболее простой подход: тематики приписываются редакторами баз журналам и другим источникам в момент начала их индексации, а затем все публикации в этих источниках приписываются к этим же тематикам. Полноценная классификация на уровне отдельных статей существует в этих базах только для некоторых мультидисциплинарных журналов, а сами наборы тематик устарели и почти не меняются.

Один из современных подходов к определению тематик основан на анализе ключевых слов и фраз в научных публикациях. Этот подход использует автоматизированные методы обработки текста для извлечения и анализа ключевых слов и фраз из названий, аннотаций и полного текста статей. Ключевые слова и фразы затем используются для определения тематики публикации в соответствии с заранее определенной классификацией.

Другой подход к выявлению тематик основывается на анализе цитирований. Этот подход использует информацию о том, какие публикации цитируются в данной публикации, для определения ее тематики. Предполагается, что публикации, которые цитируются в данной публикации, имеют схожую тематику с ней.

Существует ряд преимуществ использования анализа ключевых слов и фраз и анализа цитирований для определения тематик. Эти методы являются автоматизированными, что позволяет обрабаты-

вать большие объемы данных. Есть и проблема с объективностью результатов, поскольку ключевые слова выделяет автор. Основная претензия к последним состоит в том, что авторы самостоятельно решают, в какой поисковый запрос должны попасть их работа. В итоге часто можно встретить в публикации примерно такой набор ключевых слов: «artificial intelligence, intelligent systems, machine learning, big data, smart city, data analysis». То есть многие самые общие и «модные» характеристики и ничего или почти ничего о самом содержании работы.

Предлагается использовать современные методы статистики и NLP для решения задачи классификации на основе анализа текстов публикаций, точнее, их аннотаций. Общий анализ текстов может быть проведен по той же методике, но будет более требовательным к вычислительным ресурсам.

[3] data.head(10)

	title	authors	groups	keywords	topics	abstract
0	Kernelized Bayesian Transfer Learning	Mehmet Gonen and Adam A. Margolin	Novel Machine Learning Algorithms (NMLA)	cross-domain learning\ndomain adaptation\нкern...	APP: Biomedical / Bioinformatics\nNMLA: Bayesi...	Transfer learning considers related but distin...
1	"Source Free" Transfer Learning for Text Class...	Zhongqi Lu, Yin Zhu, Sinno Pan, Evan Xiang, Yu...	AI and the Web (AIW)\nNovel Machine Learning A...	Transfer Learning\nAuxiliary Data Retrieval\nT...	AIW: Knowledge acquisition from the web\nAIW: ...	Transfer learning uses relevant auxiliary data...
2	A Generalization of Probabilistic Serial to Ra...	Haris Aziz and Paul Stursberg	Game Theory and Economic Paradigms (GTEP)	social choice theory\nvoting\nfair division\ns...	GTEP: Game Theory\nGTEP: Social Choice / Voting	The probabilistic serial (PS) rule is one of L...
3	Lifetime Lexical Variation in Social Media	Liao Lizi, Jing Jiang, Ying Ding, Heyan Huang ...	NLP and Text Mining (NLPMT)	Generative model\nSocial Networks\nAge Prediction	AIW: Web personalization and user modeling\nNL...	As the rapid growth of online social media att...
4	Hybrid Singular Value Thresholding for Tensor ...	Xiaoqin Zhang, Zhengyuan Zhou, Di Wang and Yi Ma	Knowledge Representation and Reasoning (KRR)\n...	tensor completion\nlow-rank recovery\nhybrid s...	KRR: Knowledge Representation (General/Other)...	In this paper, we study the low-rank tensor co...
5	Locality Preserving Hashing	Kang Zhao, Hongtao Lu and Jincheng Mei	Vision (VIS)	Similarity Search\nApproximate Nearest Neighbo...	VIS: Image and Video Retrieval	Hashing has recently attracted considerable at...
6	Discovering Better AAAI Keywords via Clusterin...	Kelly Moran, Byron Wallace and Carla Brodley	Machine Learning Applications (MLA)	constraint-based clustering\nmachine learning\...	MLA: Applications of Unsupervised Learning	Selecting good conference keywords is importan...
7	Online Classification Using a Voted RDA Method	Tianbing Xu, Jianfeng Gao, Lin Xiao and Amelia...	Machine Learning Applications (MLA)\nNLP and M...	Online Classification\nVoted Dual Averaging Me...	MLA: Machine Learning Applications (General/ot...	We propose a voted dual averaging method for o...
8	Fraudulent Support Telephone Number Identifica...	Xin Li, Yiqun Liu, Min Zhang and Shaoping Ma	AI and the Web (AIW)	Fraudulent Support Telephone Number\nCo-occurr...	AIW: Enhancing web search and information retr...	"Fraudulent support phones" refers to the mistl...
9	Supervised Hashing for Image Retrieval via Ima...	Rongkai Xia, Yan Pan, Hanjiang Lai, Cong Liu a...	Novel Machine Learning Algorithms (NMLA)\nVisi...	supervised hashing\napproximate near neighbor ...	NMLA: Neural Networks/Deep Learning\nVIS: Imag...	Hashing is a popular approximate nearest neigh...

Рис.1. Фрагмент исходного датасета

[4] data.groupby('groups').count()

	title	authors	keywords	topics	abstract
groups					
	AI and the Web (AIW)	18	18	18	18
	AI and the Web (AIW)\nApplications (APP)	4	4	4	4
	AI and the Web (AIW)\nApplications (APP)\nHeuristic Search and Optimization (HSO)\nPlanning and Scheduling (PS)	1	1	1	1
	AI and the Web (AIW)\nApplications (APP)\nHuman-Computation and Crowd Sourcing (HCC)\nMachine Learning Applications (MLA)\nNovel Machine Learning Algorithms (NMLA)	1	1	1	1
	AI and the Web (AIW)\nApplications (APP)\nMachine Learning Applications (MLA)	1	1	1	1
	...	...	...	...	...
	Reasoning under Uncertainty (RU)\nRobotics (ROB)	1	1	1	1
	Robotics (ROB)	2	2	2	2
	Robotics (ROB)\nVision (VIS)	1	1	1	1
	Search and Constraint Satisfaction (SCS)	13	13	13	13
	Vision (VIS)	8	8	8	8

125 rows x 5 columns

Рис. 2. Примеры композиционных тематических групп

### Исходные данные и их предобработка

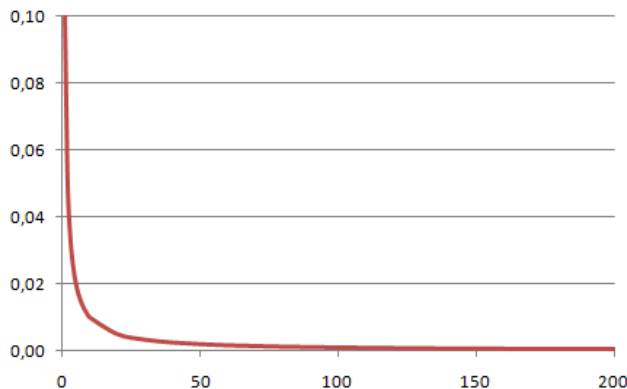
Для анализа использован датасет публикаций докладов Twenty-Eighth Conference on Artificial Intelligence [1], который представляет собой массив текстовых данных. На эту конференцию было представлено 1406 статьи, приняты к выступлению и публикации 398 (28,3%), общий объем отобранных докладов 3282 страницы. Датасет англоязычный, содержит наименование работы, авторов, секцию конференции, ключевые слова, тематику в рамках секции и аннотацию (от 1500 знаков). Пример исходных данных представлен на рис.1.

Примеры композиционных тематических групп, отраженных в датасете, приведены на рис. 2. Дисбаланс групп очевиден уже на этом фрагменте и это будет дополнительной проблемой для дальнейшего анализа.

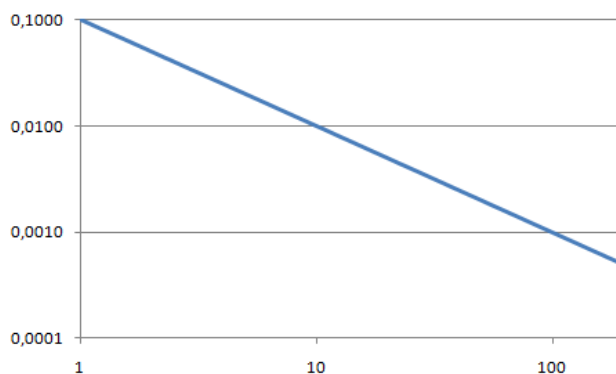
Вторая проблема датасета в том, что в нем имеются пропущенные значения, у двух статей отсутствует атрибут group, у шести – topic. Для реальных массивов это типичная картина. Вариантов дальнейшего шага в этом случае два – игнорирование соответствующих строк (работ) либо аппроксима-

ция пропущенных значений на основе похожих работ. Но в данном случае цель как раз и состоит в получении классификации, т.е. групп похожих работ, поэтому второй подход не релевантен.

В обработке естественных языков широко используется закон Ципфа-Мандельброта, связывающий между собой частоту слова и его ранг. Графически закон представлен на рисунках 3 и 4.



**Рис. 3.** Закон Ципфа в естественной шкале



**Рис. 4.** Закон Ципфа в логарифмических шкалах

Частотный подход к анализу текстов интуитивно был понятен, полезен и долгое время упирался только в технологические ограничения получения соответствующих статистик [2-4]. При этом нужно было убирать самые частые слова (English stop-words), поскольку это служебные части речи (предлоги, союзы и т.п.), некоторые наречия и др. Составление словарей такого словесного «мусора» позволяет фильтровать исходные тексты (мешки слов, bag-of-words) быстро и эффективно. Основным недостатком частотного подхода в задаче извлечения содержания в том, что он не учитывает важность слов, поэтому нужно вводить весовые коэффициенты.

Для извлечения признаков из текстовых данных в данной работе использовался метод TF-IDF (Term Frequency-Inverse Document Frequency). TF-IDF вычисляет важность каждого слова в документе относительно количества его употреблений в данном документе и во всей коллекции текстов. Этот метод позволяет выделить значимые слова и понять, какие слова имеют больший вес для определенного документа в контексте всей коллекции.

### Результаты классификации аннотаций научных статей

Для классификации полученного массива требовался метод обучения без учителя, жадный, не жесткий, масштабируемый с легко интерпретируемыми результатами [5]. Поэтому были отобраны метод k-means и иерархический агломеративный алгоритм.

#### Метод k-means.

При всех своих достоинствах обладает двумя недостатками – чувствительностью к выбросам и требованием задания k. Первый фактор будет нивелирован генерацией нескольких процедур для каждого фиксированного значения k, второй – выполнением классификации при различных значениях числа классов и анализом полученных результатов. Для оценки результатов при различных k использовался метод силуэта (Silhouette).

Значение силуэта — это мера того, насколько объект похож на собственный класс по сравнению с другими классами. Силуэт находится в диапазоне  $[-1,1]$ , где высокое значение силуэта указывает на то, что объект хорошо соответствует собственному классу и плохо соответствует другим классам. Если большинство объектов имеют высокое значение силуэта, конфигурация классификации может быть принята. Если многие объекты имеют низкие или отрицательные значения силуэта – в конфигурации слишком много или слишком мало классов.

Анализ силуэтов показывает, что обосновано остановиться на 10 классах в методе k-means. При этом было обнаружено, что распределение исходных разделов по генерализованным классам обладает рядом хороших свойств (рис. 5). Нумерация классов – от 0 до 9.

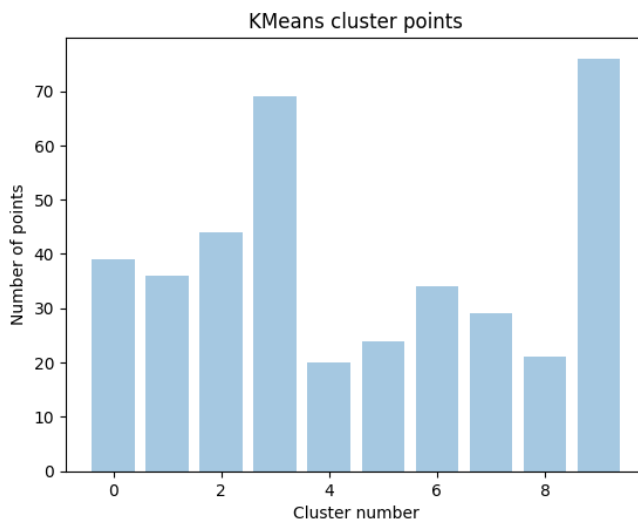


Рис. 5. Гистограмма численности классов,  $k=10$

1. Классы неплохо сбалансированы, гистограмма на рисунке 5.

2. Выделенные чисто математически классы несут в себе знания по обобщению контента, каждый класс имеет свои смысловые отличительные особенности. При этом, конечно, ввиду общей тематики конференции (Artificial Intelligence, AI) классы будут частично накладываться, но при этом облако исходных аннотаций оказывается линейно упорядоченным в пространстве первых двух главных компонент (рис. 6).

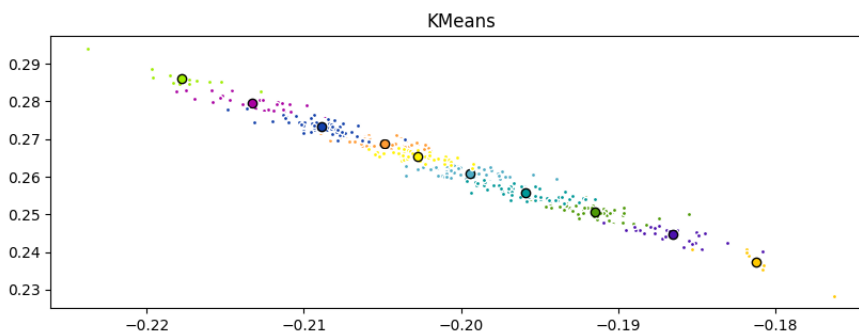


Рис. 6. Распределение 10 классов и их центры

Также удалось выделить некоторые самостоятельные классы, связанные с изображениями, компьютерным зрением, теорией игр, экономикой, планированием, новыми алгоритмами.

Описание выделенных классов по аннотациям приведено в таблице 1. Из результатов таблицы 1 следует, что средствами кластерного анализа были выделены новые закономерности в компоновке документов, при этом они получены из анализа текста аннотаций, а не списка ключевых слов. Поэтому обработка текста информативнее и точнее соответствует как классификации, так и релевантности поисковых запросов.

Таблица 1

Результаты классификации

Класс	Секции	Число работ	Общая характеристика
0	AI and the Web (AIW)	5	В большей степени связан с веб
	AI and the Web (AIW) Applications (APP)	3	
	AI and the Web (AIW) Novel Machine Learning Algorithms (NMLA)	4	
	Machine Learning Applications (MLA)	2	
	NLP and Text Mining (NLPTM)	2	
	Novel Machine Learning Algorithms (NMLA)	4	
1	Game Theory and Economic Paradigms (GTEP)	9	В большей степени связан с экономикой и теорией игр
	Game Theory and Economic Paradigms (GTEP) Multiagent Systems (MAS)	2	
	Knowledge Representation and Reasoning (KRR)	4	
	Multiagent Systems (MAS)	2	
	Search and Constraint Satisfaction (SCS)	3	
2	AI and the Web (AIW)	3	В большей степени связан с областью изображений и компьютерным зрением
	Applications (APP)	2	
	Knowledge Representation and Reasoning (KRR)	2	
	Machine Learning Applications (MLA) Vision (VIS)	2	
	NLP and Machine Learning (NLPML)	2	
	Novel Machine Learning Algorithms (NMLA)	7	
	Novel Machine Learning Algorithms (NMLA) Vision (VIS)	2	
	Vision (VIS)	7	
3	Heuristic Search and Optimization (HSO)	2	В большей степени связан с новыми алгоритмами ML
	Knowledge Representation and Reasoning (KRR)	3	
	Machine Learning Applications (MLA)	3	
	Machine Learning Applications (MLA) Novel Machine Learning Algorithms (NMLA)	5	
	NLP and Machine Learning (NLPML) Novel Machine Learning Algorithms (NMLA)	2	
	Novel Machine Learning Algorithms (NMLA)	16	
4	Computational Sustainability and AI (CSAI)	3	В большей степени связан с вычислительными системами
	Computational Sustainability and AI (CSAI) Game Theory and Economic Paradigms (GTEP) Multiagent Systems (MAS)	2	
	Game Theory and Economic Paradigms (GTEP) Multiagent Systems (MAS)	2	
	Search and Constraint Satisfaction (SCS)	2	
5	Game Theory and Economic Paradigms (GTEP)	7	В большей степени связан с экономикой и представлением знаний
	Game Theory and Economic Paradigms (GTEP) Multiagent Systems (MAS)	3	
	Knowledge Representation and Reasoning (KRR)	3	
	Multiagent Systems (MAS)	2	
6	AI and the Web (AIW)	2	В большей степени связан с планированием, составлением графиков
	Heuristic Search and Optimization (HSO) Planning and Scheduling (PS)	2	
	Knowledge Representation and Reasoning (KRR) Planning and Scheduling (PS)	3	
	Machine Learning Applications (MLA) Novel Machine Learning Algorithms (NMLA)	3	
	Planning and Scheduling (PS)	5	
7	AI and the Web (AIW)	2	В большей степени связан с эвристическим поиском, оптимизацией
	Heuristic Search and Optimization (HSO)	3	
	Heuristic Search and Optimization (HSO) Planning and Scheduling (PS)	2	
	Heuristic Search and Optimization (HSO) Search and Constraint Satisfaction (SCS)	2	
	Novel Machine Learning Algorithms (NMLA)	3	
	Reasoning under Uncertainty (RU)	3	
8	Game Theory and Economic Paradigms (GTEP) – 7	7	В большей степени связан с теорией игр
	Game Theory and Economic Paradigms (GTEP) Multiagent Systems	3	



Класс	Секции	Число работ	Общая характеристика
	(MAS)		
9	AI and the Web (AIW)	3	В большей степени связан с новыми алгоритмами ML и их применением
	Applications (APP) Machine Learning Applications (MLA)	2	
	Game Theory and Economic Paradigms (GTEP)	3	
	Heuristic Search and Optimization (HSO) Search and Constraint Satisfaction (SCS)	2	
	Knowledge Representation and Reasoning (KRR)	4	
	Machine Learning Applications (MLA)	5	
	Machine Learning Applications (MLA) Novel Machine Learning Algorithms (NMLA)	3	
	NLP and Text Mining (NLPTM)	2	
	Novel Machine Learning Algorithms (NMLA)	16	
	Search and Constraint Satisfaction (SCS)	4	

Таблица 2 иллюстрирует разделимость классов на основе использованного метода. Приведены наиболее часто использованные термины для каждого класса в порядке убывания частоты, словоформы одного и того же термина были исключены из списка.

Таблица 2

Значимые термины найденных классов

Класс	Самые частые термины
0	user, model, social, behavior
1	problems, algorithm, counting, number
2	image, visual, semantic, learning
3	domain, data, target, learning, feature
4	energy, materials, vehicles, preferences
5	agent, mechanism, design, exchange
6	planning, hybrid, translation, qualitative
7	quality, search, bound, heuristic, gbfs
8	game, form, strategy, mechanics
9	data, algorithm, clustering, learning

Иерархический алгоритм

Иерархический алгоритм опирается на агломеративный метод Ward с квадратичной функцией потерь, поскольку он обладает лучшей группирующей и разделяющей способностью на таких массивах. Процесс агломерации представлен на рисунке 7. На том же рисунке отмечена линия отсечения 10 классов.

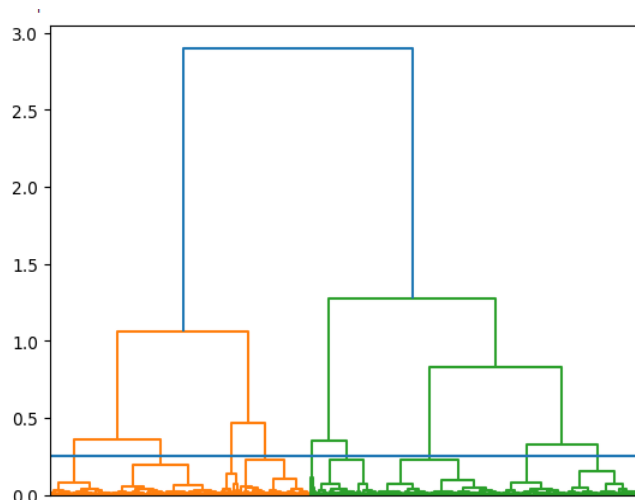


Рис. 7. Дендрограмма иерархического алгоритма

Анализ структуры классов в целом совпадает с полученным ранее и приведенным в таблице 1.

При анализе дендрограммы возникает предположение, что, возможно, применение 15 классов будет также релевантно. Это предположение было проверено алгоритмом k-means. Результаты приведены на рисунках 8 и 9.

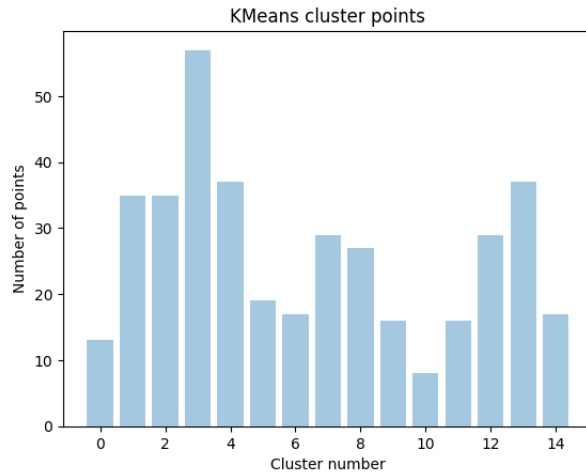


Рис. 8. Гистограмма численности классов, k=15

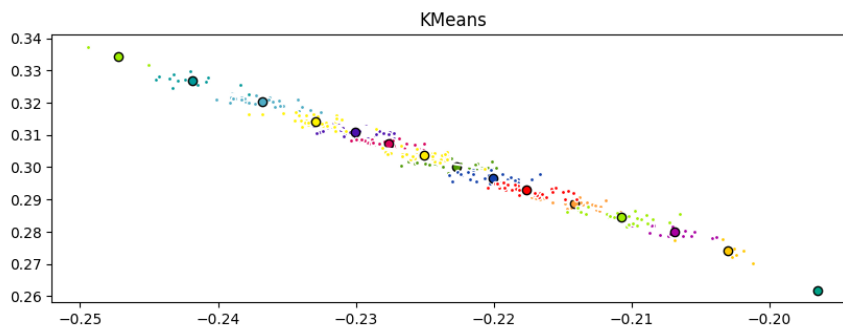


Рис. 9. Распределение 15 классов и их центроиды

Возвращаясь к методу k-means, для k=15 получаем результаты хуже, чем при k=10: классы разбалансированы (рис. 8), сильно перекрываются (рис. 9.). Эта проверка показывает, что массив аннотаций конференции AAAI-14, действительно, может быть отнесен к 10 классам. Причем ввиду их перекрытия получалась нечеткая (fuzzy) классификация на базе использования методов NLP [6-7].

### Заключение

Обработка и классификация текстов на основе естественного языка широко используется в поисковых системах. При этом она требует предварительной разметки обучающего массива вручную. Причем для научных публикаций задача осложняется тем, что терминологически классификация выглядит неоднозначно. Например, термин «узел» имеет множественные толкования в таких разных научных областях, как математика, физика, ботаника, техника, навигация. Еще примеры: «потенциал», «ядро», «функция» и множество других. Причем в естественно-научных и технических публикациях один и тот же термин может употребляться в разных значениях в пределах одной работы. Алгоритмы на основе законов Ципфа, Манделброта и их взвешенные модификации позволяют точнее выделять терминологические особенности и сопоставлять публикации.

В данной работе продемонстрирован математический аппарат классификации научных текстов по близкой тематике – искусственному интеллекту. При общем объеме работ 3282 с. для классификации использованы только аннотации (общим объемом около 700 000 знаков). Использованы два метода – k-means и агломеративный иерархический анализ на базе метода Ward'a и квадратичной функции потерь. Для нахождения количества классов были протестированы разные классификации, их сравнение проведено на основе силуэтов. Наилучшим результатом оказалось разбиение публикаций на 10 классов, для каждого класса обнаружены особенности контента.

Таким образом, математическими алгоритмами показана эффективность применения статистического подхода к задаче выделения смысловых компонент и классификации научных публикаций по их содержательным аннотациям. Можно предположить, что использование не сокращенных аннотаций, а полных текстов работ даст более точные результаты, но это потребовало бы огромных вычислительных ресурсов, что вряд ли оправданно.

### Литература

1. *Brodley C.* AAAI 2014 Accepted Papers. UCI Machine Learning Repository. URL: <https://doi.org/10.24432/C5S31F> (дата обращения 20.11.2023)
2. *Филиппович Ю.Н., Прохоров А.В.* Семантика информационных технологий: опыты словарно-тезаурусного описания. Серия «Компьютерная лингвистика». М.: МГУП, 2002. 368 с.
3. *Мандельброт Б.* Теория информации и психоллингвистика: теория частот слов // Математические методы в социальных науках / Сб. статей под ред. П. Лазарсфельда и Н. Генри. М.: Прогресс, 1973. С. 316-337.
4. *Вьюгин В.В.* Математические основы машинного обучения и прогнозирования. Электронное издание М.: МЦНМО, 2014. 304 с.
5. *Gelman A., Hill J.* Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge: Cambridge University Press, 2007. 625 p.
6. *Li Mao, Qidong Chen, Jun Sun.* Construction and Optimization of Fuzzy Rule-Based Classifier with a Swarm Intelligent Algorithm: электрон. версия ст. <https://doi.org/10.1155/2020/9319364> // Mathematical Problems in Engineering, vol. 2020, Article ID 9319364, 12 pages, 2020 (дата обращения 20.11.2023).
7. *El M., Safwat S., Hegazy O.* Big Data Classification using Fuzzy K-Nearest Neighbor // International Journal of Computer Applications. 2015. Vol. 132. No. 10, pp. 8-13.

# ПОВЫШЕНИЕ ТОЧНОСТИ РАСПОЗНАВАНИЯ НА СОБСТВЕННОМ НАБОРЕ ДАННЫХ ПРИ ОБУЧЕНИИ С ИСПОЛЬЗОВАНИЕМ МОДЕЛИ YOLOv8 NANO

Сотников Александр Юрьевич,

Московский технический университет связи и информатики, магистрант, Москва, Россия,  
[sotner@mail.ru](mailto:sotner@mail.ru)

## Аннотация

В статье представлены результаты способов повышения точности распознавания на собственном наборе данных при обучении с использованием модели YOLOv8 NANO. Обучение происходило на видеокарте GeForce GTX 1050 Ti. В ходе испытаний, набор данных динамически меняется от 8665 до 8855 изображений. Набор данных состоял из 7 классов: «White», «Brown», «Dirt», «Mark», «Shellbroken», «Ovoscope», «Ovospecrack».

**Ключевые слова:** Машинное обучение, YOLO, YOLOv8, модель nano, Python, Машинное зрение, Техническое зрение, Метрики обнаружения объектов, Object Detection Metrics, Классовые метрики, Class-wise Metrics, chicken eggs, куриные яйца, овоскопирование, овоскоп, обучение на CUDA, GeForce GTX 1050 Ti, Nvidia, Anaconda, PyTorch, Roboflow, PyCharm Professional, 300 epochs, 300 эпох, 400 epochs, 400 эпох, применение фоновых изображений, нулевая аннотация, Окская птицефабрика (АО «Окское»), птицефабрика Волжанин (ОАО «Волжанин»)

## Введение

В данной статье описывается процесс ручной аннотации изображений в системе «Roboflow», экспорт и адаптация данных осуществлялась на основе выбранной модели YOLOv8 [1], а также проводилась установка библиотек на локальной ЭВМ для дальнейшего обучения. Приводится обучение набора данных на модели YOLOv8 Nano. Показан результат с запуска веб-камеры в реальном времени. В исследовании используются фотографии собственного набора данных, где под собственным набором данных понимается изображения куриных яиц категории C1, отсортированные на семь классов. Представлены метрики обнаружения объектов, используемые аргументы при обучении. Показано сравнение точности обучения при использовании следующих способов повышения: нулевая аннотация с использованием фоновых изображений и увеличение количества эпох.

## Результаты исследований

Результаты исследований представлены с использованием модели обнаружения YOLOv8n.pt (Nano). Это наномодель самая легкая и компактная. Она имеет меньшее количество слоев и меньшую размерность фильтров, что делает ее идеальной для использования на системах с ограниченными вычислительными ресурсами, таких как мобильные устройства или встроенные системы.

Аннотация изображений была проведена вручную в онлайн сервисе «ROBOFLOW» [2].

Для обучения на YOLOv8 необходимо 640x640 пикселей. В сервисе «ROBOFLOW» при генерировании версии датасета задавалась данная настройка. В ходе сборки версии была выбрана авто ориентация – «Applied», размер 640x640. Аугментация была отключена.

Распространенными способами повышения скорости точности распознавания изображений являются: добавление большого количества данных, добавление количества слоев, изменение размера изображений, добавление количества эпох, уменьшение цветового канала [3].

В нашем цикле испытаний добавлялись изображения, количество эпох – от 300 до 400. Также проверялся еще один способ – это добавление фоновых изображений с нулевой аннотацией.

Особенно сложной проблемой может быть определение изображений, на которых отсутствуют аннотации – случайно или намеренно.

**Отсутствующая аннотация** возникает, когда на изображении есть объекты, которые не помечены, хотя должны быть. Это проблематично, поскольку ваша модель будет обучаться на ложноотрицательных изображениях ваших объектов [4].

**Нулевая аннотация** возникает, когда на изображении нет объектов, и, следовательно, нет необходимости записывать ограничивающие рамки. Это не обязательно проблематично – на самом деле, для обучения модели может быть желательно, чтобы объекты не всегда присутствовали в кадре [4].

В процессе исследования проводились ниже представленные испытания.

«FONOFF» - испытание № 1 - набор данных без фоновых изображений (300 эпох).

«FONMOD» - испытание №2 с фоновыми изображениями с нулевой аннотацией (300 эпох).

«FONON400» - испытание №3 с фоновыми изображениями (нулевая аннотация) и дополнительные изображения классов (400 эпох).

Таблица 1

Описание экспериментов

№ испытания	Кодовое имя	Количество эпох	Количество изображений/аннотаций
1	«FONOFF»	300	8665/12372
2	«FONMOD»	300	8785/12372
3	«FONON400»	400	8855/12588

Для отслеживания жизненного цикла модели во время обучения использовали платформу «COMET ML» [5].

### Классовые метрики

**Class:** обозначает имя класса объектов, например "человек", "автомобиль" или "собака".

**Images:** этот показатель показывает количество изображений в наборе проверки, содержащих класс объекта.

**Instances:** это позволяет подсчитать, сколько раз класс появляется во всех изображениях в наборе проверки.

**P (точность):** точность обнаруженных объектов, указывающая, сколько обнаружений было правильным.

**R (отзыв):** способность модели идентифицировать все экземпляры объектов на изображениях.

**mAP50:** средняя точность, рассчитанная при пересечении порога превышения union (IoU) 0,50. Это показатель точности модели, учитывающий только "простые" обнаружения.

**mAP50-95:** среднее значение средней точности, рассчитанное при различных пороговых значениях IoU в диапазоне от 0,50 до 0,95. Это дает полное представление о производительности модели на разных уровнях сложности обнаружения.

**mAp:** подходит для широкой оценки характеристик модели.

**IoU:** необходима, когда решающее значение имеет точное местоположение объекта.

**Precision:** важна, когда приоритетом является минимизация ложных обнаружений.

**Recall:** жизненно важно, когда важно обнаруживать каждый экземпляр объекта.

**F1 Score:** полезно, когда необходим баланс между точностью и запоминанием [6].

### Аргументы установленные вручную перед обучением

Таблица 2

Ручные аргументы

№ испытания	Кодовое имя	Аргументы	Версия Ultralytics
1	«FONOFF»	epochs=300, batch=4, patience=25, imgsz=640	8.0.235
2	«FONMOD»	epochs=300, batch=4, imgsz=640	
3	«FONON400»	epochs=400, batch=4, imgsz=640, save=True, save_period=10	8.1.1

### Автоматические аргументы при обучении

engine\trainer: task=detect, mode=train, model=yolov8n.pt, data=data.yaml, time=None, cache=False, device=None, workers=8, project=None, name=NAME, exist\_ok=False, pretrained=True, optimizer=auto, verbose=True, seed=0, deterministic=True, patience=50, single\_cls=False, rect=False, cos\_lr=False, close\_mosaic=10, resume=False, amp=True, fraction=1.0, profile=False, freeze=None, multi\_scale=False, overlap\_mask=True, mask\_ratio=4, dropout=0.0, val=True, split=val, save\_json=False, save\_hybrid=False, conf=None, iou=0.7, max\_det=300, half=False, dnn=False, plots=True, source=None, vid\_stride=1, stream\_buffer=False, visualize=False, augment=False, agnos-

```

tic_nms=False, classes=None, retina_masks=False, embed=None, show=False, save_frames=False, save_txt=False,
save_conf=False, save_crop=False, show_labels=True, show_conf=True, show_boxes=True, line_width=None, for-
mat=torchscript, keras=False, optimize=False, int8=False, dynamic=False, simplify=False, opset=None, workspace=4,
nms=False, lr0=0.01, lrf=0.01, momentum=0.937, weight_decay=0.0005, warmup_epochs=3.0,
warmup_momentum=0.8, warmup_bias_lr=0.1, box=7.5, cls=0.5, dfl=1.5, pose=12.0, kobj=1.0, label_smoothing=0.0,
nbs=64, hsv_h=0.015, hsv_s=0.7, hsv_v=0.4, degrees=0.0, translate=0.1, scale=0.5, shear=0.0, perspective=0.0,
flipud=0.0, fliplr=0.5, mosaic=1.0, mixup=0.0, copy_paste=0.0, auto_augment=randaugment, crasing=0.4,
crop_fraction=1.0, cfg=None, tracker=botsort.yaml, save_dir=runs\detect\NAME
Overriding model.yaml nc=80 with nc=7
optimizer: 'optimizer=auto' found, ignoring 'lr0=0.01' and 'momentum=0.937' and determining best 'optimizer', 'lr0' and
'momentum' automatically...
optimizer: SGD(lr=0.01, momentum=0.9) with parameter groups 57 weight(decay=0.0), 64 weight(decay=0.0005), 63
bias(decay=0.0) [7]
    
```

### Информация о собственном наборе данных. Испытание № 1 – без фоновых изображений (FONOFF)

Набор данных состоит из 8665 изображений, 12372 аннотаций и делится на 7 классов: «White», «Brown», «Dirt», «Mark», «Shellbroken», «Ovoscope», «Ovoscopecrack».

«TRAIN SET» - тренировочный набор – 6898 изображений (~80%), «VALID SET» - валидационный набор – 1752 изображения (~20%), «TEST SET» - тренировочный набор – 15 изображений (23 аннотации).

В «VALID SET» каждый класс равен 350 аннотациям (суммарно 2450 аннотаций).

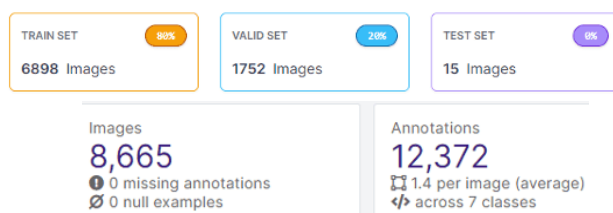


Рис. 1. Количество изображений и аннотаций. Разбиение набора данных. Испытание № 1.



Рис. 2. Количество аннотаций в «all». Испытание № 1.



Рис. 3. Количество аннотаций в «train». Испытание № 1.



Рис. 4. Количество аннотаций в «valid». Испытание № 1.



Рис. 5. Количество аннотаций в «test». Испытание № 1.

Таблица 3

Описание классов

Класс	Описание	Кол-во аннот./изобр.
«White»	Чистые белые яйца птицефабрики АО «Окское»	1614/900
«Brown»	Чистые коричневые яйца птицефабрики ОАО «Волжанин»	1644/793
«Dirt»	Белые и коричневые яйца с грязью, пометом, пигментными пятнами, в том числе и с применением овоскопа.	2268/857
«Mark»	Маркировки Российских птицефабрик АО «Окское» и АО «Волжанин».	1456/1452
«Shellbroken»	Разбитая белая и коричневая скорлупа (без желтка и белка).	1478/1476
«Ovoscope»	Целые белые и коричневые яйца. Фотографии сделаны с боковой и тупой сторон яйца (применялся овоскоп) Белые яйца при просвечивании овоскопом имеют желтый и бело-прозрачные оттенки, коричневые – красноватый.	2430/2428
«Ovoscopecrack»	Скрытые дефекты на белых и коричневых яйцах - микротрещины (применялся овоскоп).	1482/1477

**Результаты обучения на модели Nano. Испытание № 1 – без фоновых изображений (FONOFF)**

Установленные вручную аргументы: epochs=300, batch=4, patience=25, imgsz=640. На обучение ушло 300 эпох (CUDA).

```

Epoch   GPU_mem  box_loss  cls_loss  dfl_loss  Instances  Size
300/300  0.675G  0.2117   0.1667   0.8142    2          640: 100% ██████████ 1725/1725 [05:38<00:00, 5.10it/s]
Class   Images  Instances  Box(P)   R        mAP50  mAP50-95: 100% ██████████ 219/219 [00:20<00:00, 10.92it/s]
all     1752    2450      0.869    0.879    0.872  0.824

300 epochs completed in 30.912 hours.
Optimizer stripped from runs\detect\V12NANO\weights\last.pt, 6.3MB
Optimizer stripped from runs\detect\V12NANO\weights\best.pt, 6.3MB

Validating runs\detect\V12NANO\weights\best.pt...
Ultralytics V1.0.0 Python-3.11.5 torch-2.1.1 CUDA:0 (NVIDIA GeForce GTX 1050 Ti, 4096MiB)
Model summary (fused): 168 layers, 3007013 parameters, 0 gradients, 8.1 GFLOPs
Class   Images  Instances  Box(P)   R        mAP50  mAP50-95
all     1752    2450      0.867    0.88     0.872  0.824
Brown  1752    350       0.995    0.977    0.995  0.99
Dirt    1752    350       0.333    0.306    0.21   0.078
Mark    1752    350       0.757    0.886    0.92   0.781
Ovoscope 1752    350       0.996    0.997    0.995  0.995
Ovoscopecrack 1752    350       0.994    0.997    0.995  0.995
Shellbroken 1752    350       0.998    1        0.995  0.935
White   1752    350       0.998    1        0.995  0.995

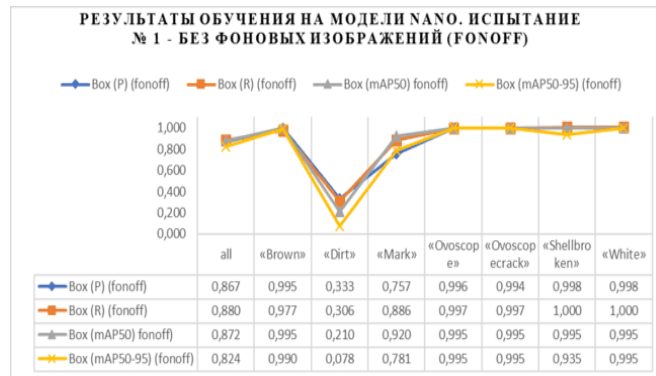
Speed: 0.3ms preprocess, 7.4ms inference, 0.0ms loss, 0.8ms postprocess per image
Results saved to runs\detect\V12NANO
    
```

Рис. 6. Успешное завершение обучения на модели Nano. Испытание № 1.

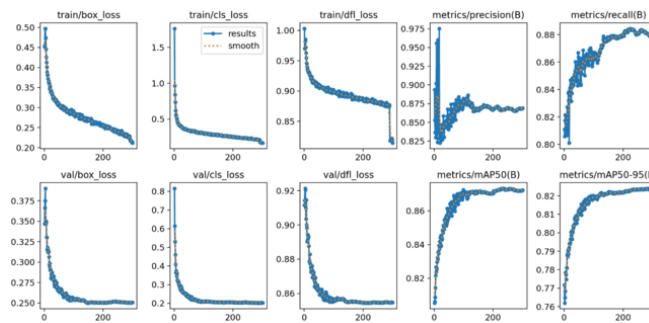
Таблица 4

Результаты. Испытание № 1

Класс	Images	Instances	Box (P)	Box (R)	Box (mAP50)	Box (mAP50-95)
all	1752	2450	0.867	0.88	0.872	0.824
«Brown»	1752	350	0.995	0.977	0.995	0.99
«Dirt»	1752	350	0.333	0.306	0.21	0.078
«Mark»	1752	350	0.757	0.886	0.92	0.781
«Ovoscope»	1752	350	0.996	0.997	0.995	0.995
«Ovoscopecrack»	1752	350	0.994	0.997	0.995	0.995
«Shellbroken»	1752	350	0.998	1	0.995	0.935
«White»	1752	350	0.998	1	0.995	0.995



**Рис. 7.** График метрик «P», «R», «mAP50», «mAP50-95» после завершения обучения на модели Nano. Испытание № 1.

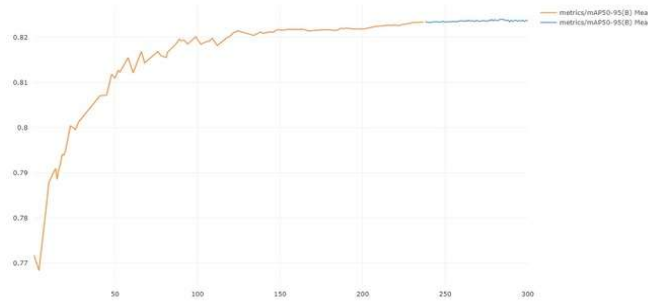


**Рис. 8.** Графики «train/box\_loss», «train/cls\_loss», «train/dfl\_loss», «metrics/precision(B)», «metrics/recall(B)», «val/box\_loss», «val/cls\_loss», «val/dfl\_loss», «metrics/mAP50(B)», «metrics/mAP50-95(B)» после завершения обучения на модели Nano. Испытание № 1.

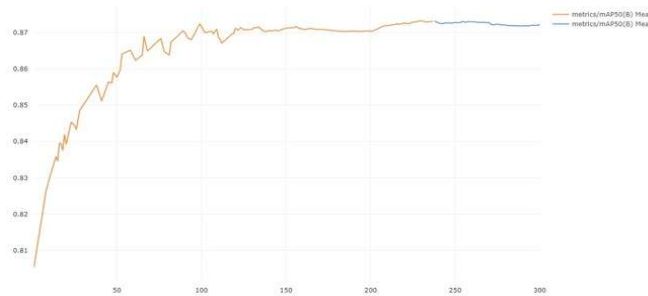
**Результаты метрик с «СОМЕТ ML». Испытание № 1.**

В ходе испытания произошёл технический сбой.

Прерывание обучения было на 237 эпохе. Возобновление началось с 238 эпохи («last.pt»). С 1 по 237 эпоху – оранжевый график, с 238 по 300 эпоху – синий график.



**Рис. 9.** «metrics\_mAP50-95(B)». Испытание № 1.



**Рис. 10.** «metrics\_mAP50(B)». Испытание № 1.



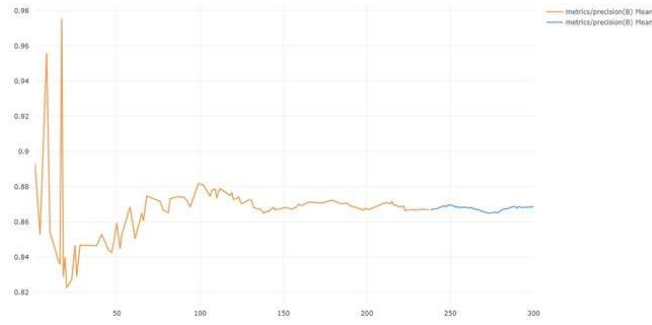


Рис. 11. «metrics\_precision(B)». Испытание № 1.

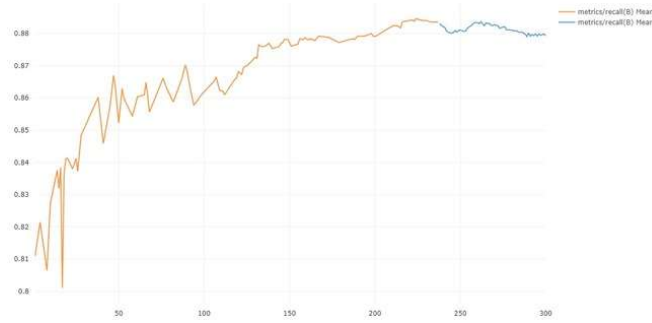


Рис. 12. «metrics\_recall(B)». Испытание № 1.

Brown	99.7	0	0.29	0	0	0	0	
Dirt	0	32.07	0	0	0	0	67.92	
Mark	0.23	0.46	74.47	0	0	0	24.82	
Ovoscope	0	0	0	99.42	0.28	0	0.28	
Ovoscopecrack	0	0	0	0.28	99.43	0	0.28	
Shellbroken	0	0	0	0	0	100	0	
White	0	0	0	0	0	0	100	
background	2.64	86.41	10.56	0.37	0	0	0	
	Brown	Dirt	Mark	Ovoscope	Ovoscopec...	Shellbroken	White	background

Рис. 13. «Матрица ошибок» в процентах после завершения обучения на модели Nano. Испытание № 1.

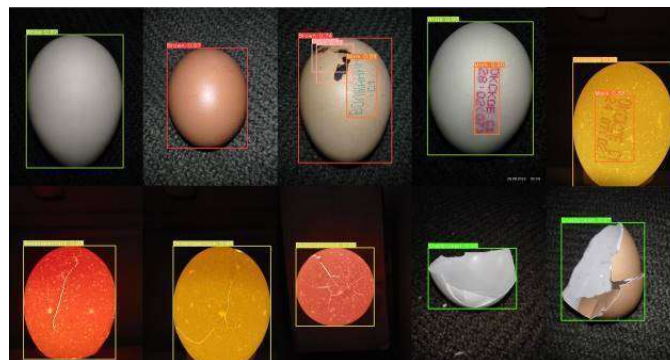


Рис. 14. Результаты правильного распознавания. Испытание № 1

### Вывод по испытанию № 1 «FONOFF»

Общая точность всех классов «all»: P=0.867, R=0.88, mAP50=0.872, mAP50-95=0.824.

Данный эксперимент будет считаться точкой отсчета для испытаний № 2 и 3.

Считается что метрика «mAP50» - единая первичная метрика для сравнения, которая охватывает как точность, так и повторяемость модели. [8]

Поэтому выберем «mAP50» как общую метрику точности, в данном случае эксперимент № 1 = 82,4%

Затраченное время на обучение: 30:54:43 часов.

### Информация о собственном наборе данных. Испытание № 2 – с фоновыми изображениями (FONMOD)

Набор данных состоит из 8785 изображений с 12372 аннотациями и делится на 7 классов: «White», «Brown», «Dirt», «Mark», «Shellbroken», «Ovoscope», «Ovoscopecrack».

В отличие от испытания №1, здесь были добавлены фоновые изображения в количестве 120 изображений в «train set». Аннотация не применялась. Пустые лотки для яиц – 16 шт., чёрный фон – 9 шт., фон без вспышки – 25 шт., фон со вспышкой – 70 шт.

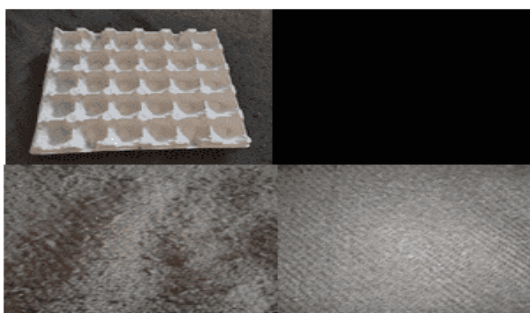


Рис. 15. Фоновые изображения

Установленные аргументы: epochs=300, batch=4, patience=25, imgsz=640. На обучение ушло 300 эпох (CUDA).

```

Epoch  GPU_mem  box_loss  cls_loss  dfl_loss  Instances  Size
300/300  0.6736   0.2084   0.1565   0.8163     1          640: 100% ██████████ 1755/1755 [05:44:00:00, 5.091t/s]
Class   Images  Instances  Box(P)   R        mAP50  mAP50-95: 100% ██████████ 220/220 [00:20:00:00, 10.931t/s]
all     1753    2450      0.88     0.879    0.88   0.827

300 epochs completed in 30.853 hours.
Optimizer stripped from runs\detect\V9NANO\weights\last.pt, 6.3MB
Optimizer stripped from runs\detect\V9NANO\weights\best.pt, 6.3MB

Validating runs\detect\V9NANO\weights\best.pt...
Ultralytics YOLOv8.0.235 Python-3.11.5 torch-2.1.1 CUDA:0 (NVIDIA GeForce GTX 1050 Ti, 4096MiB)
Model summary (fused): 168 layers, 3007013 parameters, 0 gradients, 8.1 GFLOPs
Class   Images  Instances  Box(P)   R        mAP50  mAP50-95: 100% ██████████ 220/220 [00:18:00:00, 11.991t/s]
all     1753    2450      0.88     0.877    0.88   0.827
Brown   1753    350       0.994    0.985    0.995   0.992
Dirt    1753    350       0.419    0.294    0.259   0.086
Mark    1753    350       0.768    0.869    0.927   0.793
Ovoscope 1753    350       0.995    0.997    0.994   0.994
Ovoscopecrack 1753    350       0.998    0.997    0.995   0.995
Shellbroken 1753    350       0.998    1        0.995   0.933
White   1753    350       0.986    1        0.995   0.995

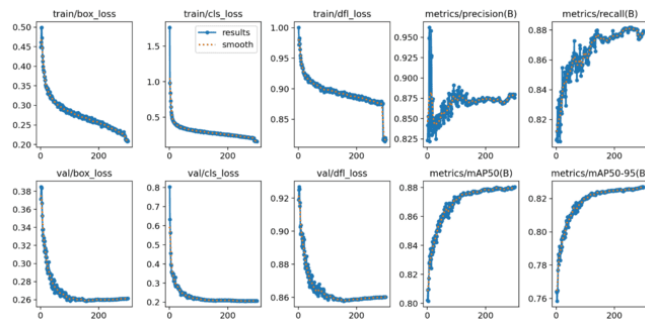
Speed: 0.3ms preprocess, 7.4ms inference, 0.0ms loss, 0.8ms postprocess per image
Results saved to runs\detect\V9NANO
    
```

Рис. 16. Успешное завершение обучения на модели Nano. Испытание № 2.

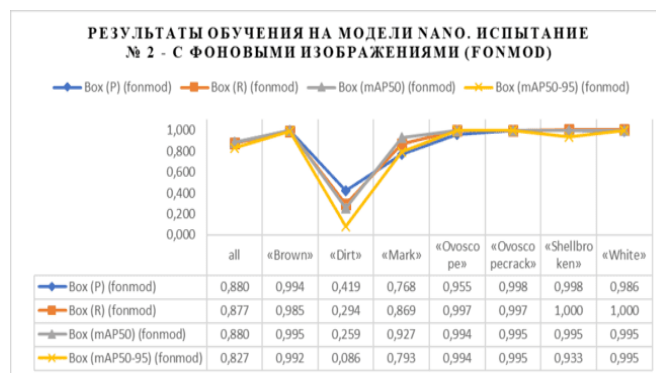
Таблица 5

### Результаты. Испытание № 2

Класс	Images	Instances	Box (P)	Box (R)	Box (mAP50)	Box (mAP50-95)
all	1753	2450	0.88	0.877	0.88	0.827
«Brown»	1753	350	0.994	0.985	0.995	0.992
«Dirt»	1753	350	0.419	0.294	0.259	0.086
«Mark»	1753	350	0.768	0.869	0.927	0.793
«Ovoscope»	1753	350	0.995	0.997	0.994	0.994
«Ovoscopecrack»	1753	350	0.998	0.997	0.995	0.995
«Shellbroken»	1753	350	0.998	1	0.995	0.933
«White»	1753	350	0.986	1	0.995	0.995

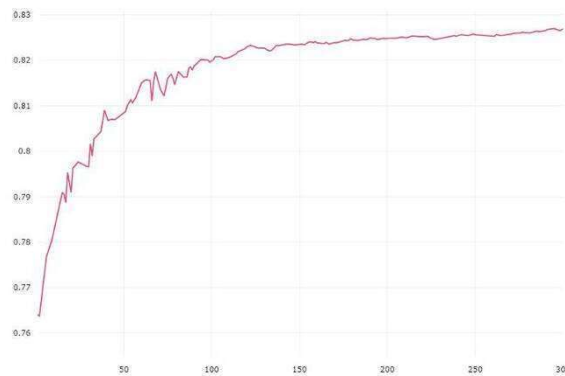


**Рис. 17.** Графики «train/box\_loss», «train/cls\_loss», «train/df\_l\_loss», «metrics/precision(B)», «metrics/recall(B)», «val/box\_loss», «val/cls\_loss», «val/df\_l\_loss», «metrics/mAP50(B)», «metrics/mAP50-95(B)» после завершения обучения на модели Nano. Испытание № 2.

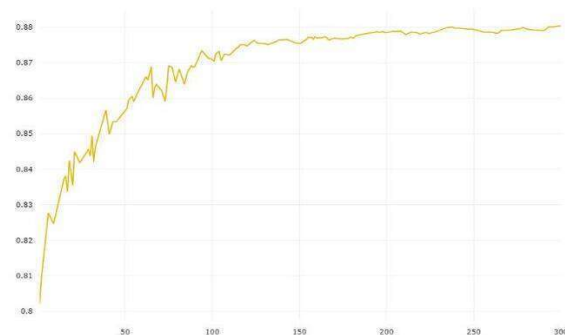


**Рис. 18.** График метрик «P», «R», «mAP50», «mAP50-95» после завершения обучения на модели Nano. Испытание № 2.

**Результаты метрик с «СОМЕТ ML». Испытание № 2.**



**Рис. 19.** «metrics\_mAP50-95(B)». Испытание № 2.



**Рис. 20.** «metrics\_mAP50(B)». Испытание № 2.

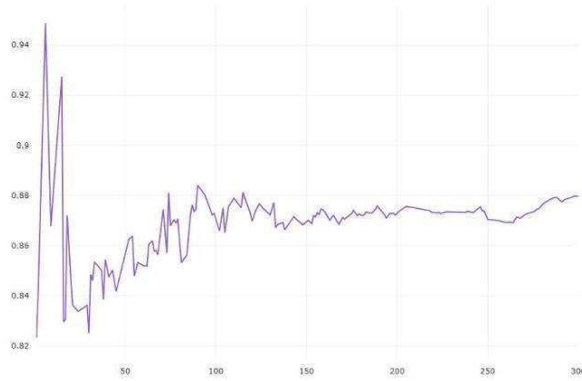


Рис. 21. «metrics\_precision(B)». Испытание № 2.

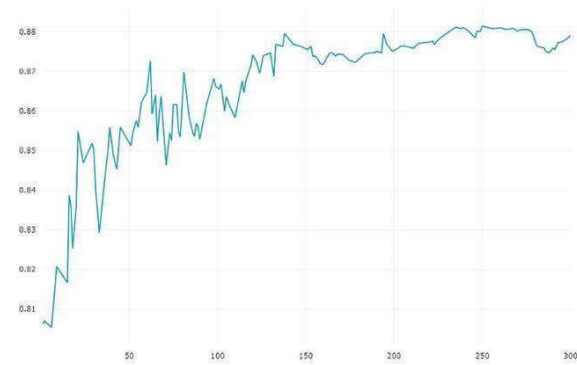


Рис. 22. «metrics\_recall(B)». Испытание № 2.

Brown	99.42	0	0.28	0	0	0	0.28	
Dirt	0	40.26	0	0	0	0	59.73	
Mark	0	0.23	74.18	0	0	0	25.58	
Ovoscope	0	0	0	99.71	0.28	0	0	
Ovoscopecrack	0	0	0	0	100	0	0	
Shellbroken	0	0	0	0	0	100	0	
White	0	0	0	0	0	0	98.59	
background	1.88	86.41	11.32	0.37	0	0	0	
	Brown	Dirt	Mark	Ovoscope	Ovoscopec...	Shellbroken	White	background

Рис. 23. «Матрица ошибок» в процентах после завершения обучения на модели Nano. Испытание № 2.

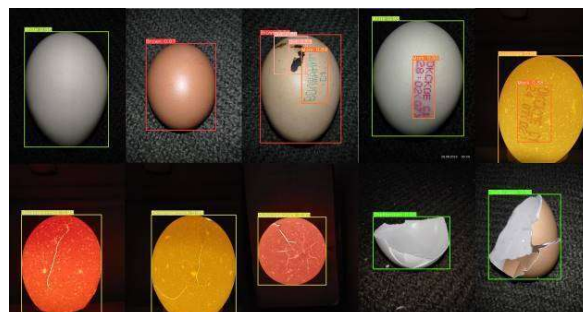


Рис. 24. Результаты правильного распознавания. Испытание № 2.

**Вывод по испытанию № 2 «FONMOD»**

Общая точность всех классов «all»: P=0.88, R=0.877, mAP50=0.88, mAP50-95=0.827.  
 Точность испытания № 2 = 88%  
 Затраченное время на обучение: 30:51:10 часов.

**Сравнение между испытаниями № 1 «FONOFF» и № 2 «FONMOD»**

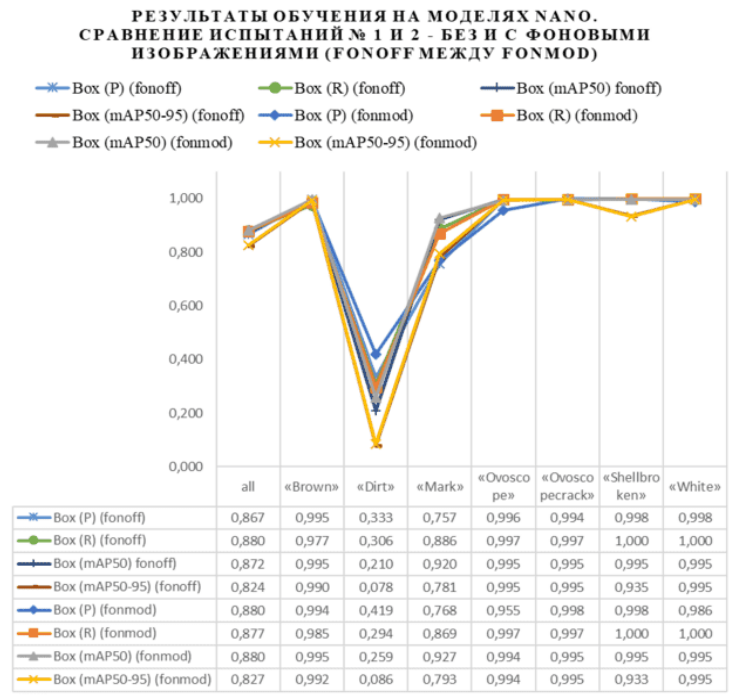


Рис. 25. Сравнение Испытаний № 1 и 2 - без и с фоновыми изображениями (fonoff между fonmod)

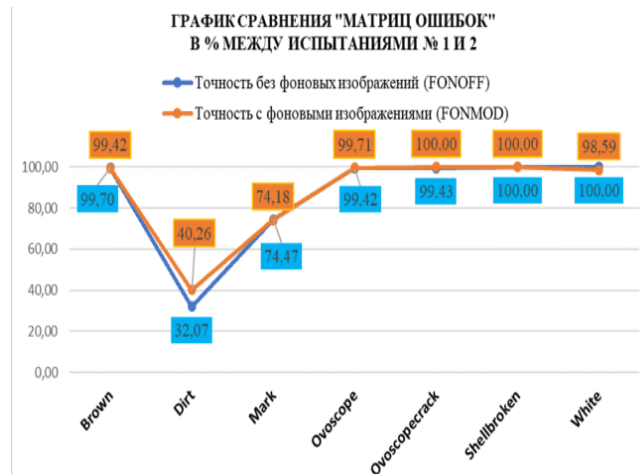


Рис. 26. График сравнения "матриц ошибок" в % между испытаниями № 1 и 2



Рис. 27. График показателей разности точности в % "FONMOD" относительно "FONOFF"

### Результаты сравнения между испытаниями № 1 «FONOFF» и № 2 «FONMOD»

Фоновые изображения в количестве 120 изображений с нулевой аннотацией повысили точность класса «Dirt» на 8,19%, что является отличным результатом.

Основной проблемой данного класса была путаница распознавания грязи, которая в основном имеет темный оттенок. Хотя, аннотация класса «Dirt» была полностью сделана вручную, это не помогло правильно распознать данный класс. Трудность заключалась в том, что класс «Dirt» сравнивала автоматически с созданным классом «Background» (рис. 13), в которых не было изначально фоновых изображений.

#### Информация о собственном наборе данных.

#### Испытание № 3 – с фоновыми изображениями и добавлением фото (FONON400)

Набор данных состоит из 8855 изображений с 12588 аннотациями и делится на 7 классов: «White», «Brown», «Dirt», «Mark», «Shellbroken», «Ovoscope», «Ovoscopecrack».

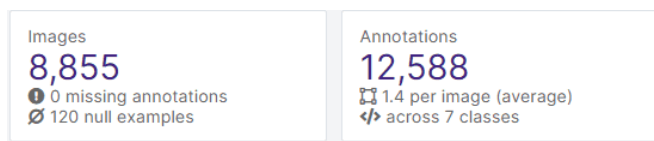


Рис. 28. Количество изображений и аннотаций. Испытание № 3.

В ходе анализа испытаний № 1 и 2 по «матрицам ошибок» было решено добавить фотографии в классы «Brown», «White» и «Mark», так как после добавления фоновых изображений, показания точности ухудшились у данных классов.

В отличие от испытания №2, здесь были добавлены изображения в количестве 70 штук в «train set» с аннотацией. Коричневые яйца с маркировкой («Brown», «Mark») – 60 шт., белые яйца с маркировкой («White», «Mark») – 10 шт. Увеличилось количество аннотаций на 216 штук и количество эпох – 400. Остальные параметры остались без изменений.



Рис. 29. Пример добавления изображений с аннотациями класс «White», «Mark» в «FONON400». Испытание № 3.



Рис. 30. Количество аннотаций в «all». Испытание № 3.



Рис. 31. Количество аннотаций в «train». Испытание № 3.

Установленные аргументы: epochs=400, batch=4, imgsz=640. На обучение ушло 400 эпох (CUDA).

```

Epoch   GPU_mem  box_loss  cls_loss  dfl_loss  Instances  Size
400/400  0.635G   0.2069   0.1811   0.816   4   440:1000 | 1772/1772 [05:46:00:00, 6.111it/s]
Class   Images  Instances  Box(P)   R      mAP50  mAP50-95): 100% | ██████████ | 219/219 [00:19:00:00, 10.991it/s]
all     1752    2450     0.889   0.873  0.882  0.83

299 epochs completed in 27:259 hours.
Optimizer stripped from runs\detect\vi44nand\weights\last.pt, 4.3MB
Optimizer stripped from runs\detect\vi44nand\weights\best.pt, 4.3MB

Validating runs\detect\vi44nand\weights\best.pt...
 Ultralytics YOLOv8.1.1 | Python-3.11.8 torch-2.1.1 CUDA-0 (NVIDIA GeForce RTX 1050 Ti, 4096MiB)
Model summary (fused): 246 layers, 2007012 parameters, 0 gradients, 8.1 GFLOPs
Class   Images  Instances  Box(P)   R      mAP50  mAP50-95): 100% | ██████████ | 219/219 [00:18:00:00, 12.861it/s]
all     1752    2450     0.889   0.873  0.882  0.83
Brown  1752    350      0.997   0.961  0.993  0.992
Dirt   1752    350      0.422   0.303  0.271  0.0996
Mark   1752    350      0.815   0.854  0.932  0.802
Ovoscope 1752    350      0.995   0.997  0.994  0.994
Ovoscopecrack 1752    350      0.997   0.997  0.995  0.995
Shellbroken 1752    350      0.998   1       0.995  0.934
White  1752    350      0.996   1       0.995  0.995
Speed: 0.3ms preprocess, 7.3ms inference, 0.0ms loss, 0.8ms postprocess per image
Results saved to runs\detect\vi44nand
    
```

Рис. 32. Успешное завершение обучения на модели Nano. Испытание № 3.

Таблица 6

Результаты. Испытание № 3

Класс	Images	Instances	Box (P)	Box (R)	Box (mAP50)	Box (mAP50-95)
all	1752	2450	0.889	0.873	0.882	0.83
«Brown»	1752	350	0.997	0.961	0.993	0.992
«Dirt»	1752	350	0.422	0.303	0.271	0.0996
«Mark»	1752	350	0.815	0.854	0.932	0.802
«Ovoscope»	1752	350	0.995	0.997	0.994	0.994
«Ovoscopecrack»	1752	350	0.997	0.997	0.995	0.995
«Shellbroken»	1752	350	0.998	1	0.995	0.934
«White»	1752	350	0.996	1	0.995	0.995

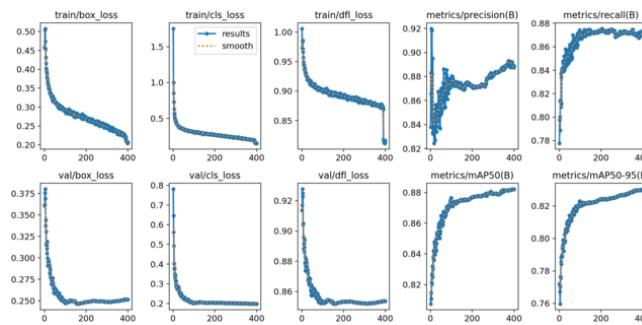


Рис. 33. Графики «train/box\_loss», «train/cls\_loss», «train/dfl\_loss», «metrics/precision(B)», «metrics/recall(B)», «val/box\_loss», «val/cls\_loss», «val/dfl\_loss», «metrics/mAP50(B)», «metrics/mAP50-95(B)» после завершения обучения на модели Nano. Испытание № 3.

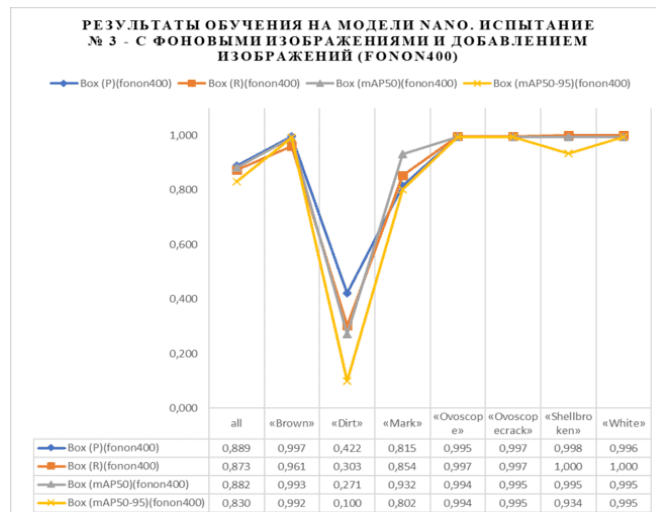


Рис. 34. График метрик «P», «R», «mAP50», «mAP50-95» после завершения обучения на модели Nano. Испытание № 3.

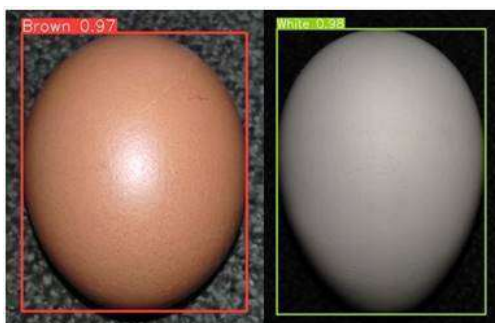


Рис. 35. Правильное распознавание классов «Brown» и «White»

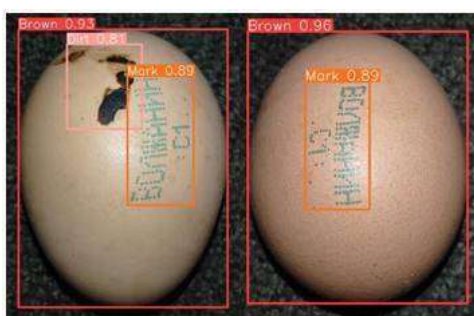


Рис. 36. Правильное распознавание классов «Brown», «Dirt», «Mark»

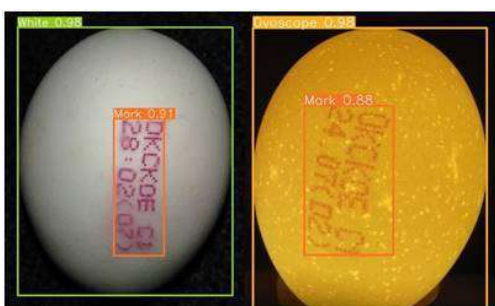


Рис. 37. Правильное распознавание классов «White», «Mark», «Ovoscope»

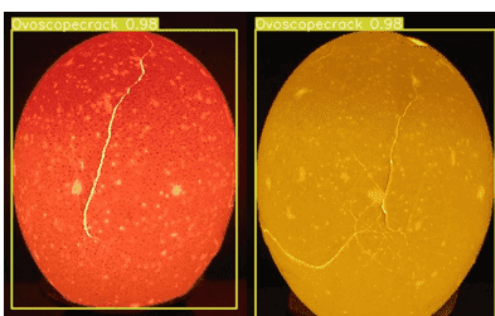


Рис. 38. Правильное распознавание класса «Ovoscopecrack»

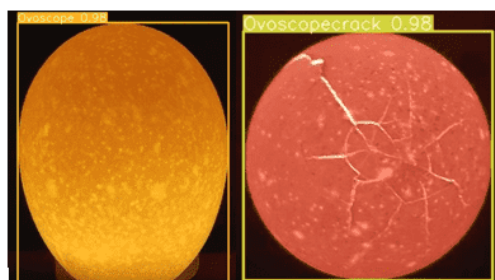
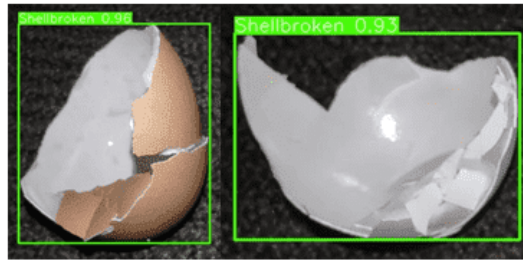


Рис. 39. Правильное распознавание классов «Ovoscope» и «Ovoscopecrack»





**Рис. 40.** Правильное распознавание класса «Shellbroken»

**Вывод по испытанию № 3 «FONON400»**

Общая точность всех классов «all»: P=0.889, R=0.873, mAP50=0.882, mAP50-95=0.83.

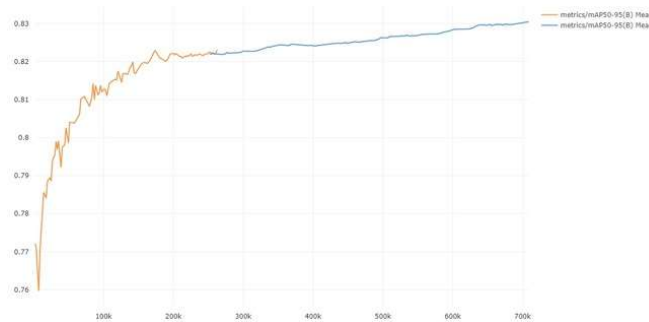
Точность испытания № 3 = 88,2%

Затраченное время на обучение: 42:59:27 часа.

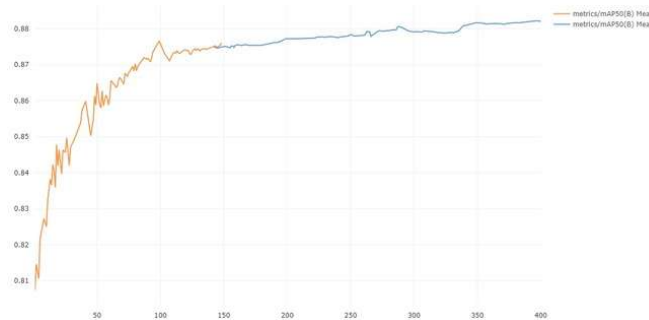
**Результаты метрик с «COMET ML». Испытание № 3.**

В ходе испытания № 3 произошёл технический сбой.

Прерывание обучения было на 148 эпохе. Возобновление началось с 142 эпохи по причине того, что файл «last.pt» был повреждён. Благодаря ручному установленному аргументу, сохранение было каждые 10 эпох. С 1 по 148 эпоху – оранжевый график, с 142 по 400 эпоху – синий график.



**Рис. 41.** «metrics\_mAP50-95(B)». Испытание № 3.



**Рис. 42.** «metrics\_mAP50(B)». Испытание № 3.



**Рис. 43.** «metrics\_precision(B)». Испытание № 3.

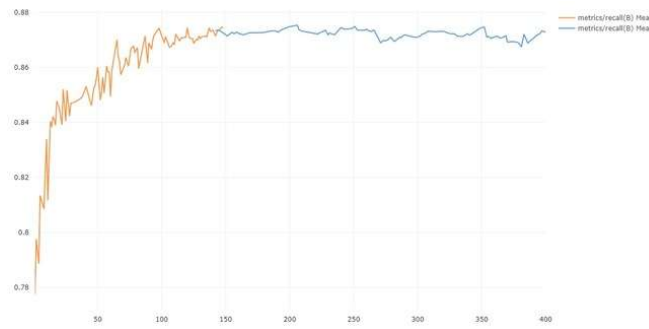


Рис. 44. «metrics\_recall(B)». Испытание № 3.

		Predicted Category							
		Brown	Dirt	Mark	Ovoscope	Ovoscopecrack	Shellbroken	White	background
Actual Category	Brown	99.7	0	0.29	0	0	0	0	0
	Dirt	0.37	44.15	0	0	0	0	0	55.47
	Mark	0.25	0	78.42	0	0	0	0	21.31
	Ovoscope	0	0	0	99.71	0.28	0	0	0
	Ovoscopecrack	0	0	0	0	99.71	0	0	0.28
	Shellbroken	0	0	0	0	0	99.71	0	0.28
	White	0	0	0	0	0	0.28	99.71	0
	background	4.19	81.46	13.98	0.34	0	0	0	0

Рис. 45. «Матрица ошибок» в процентах после завершения обучения на модели Nano. Испытание № 3.

### Результаты обучения на модели Nano. Сравнение испытаний № 1 (FONOFF), 2 (FONMOD) и 3 (FONON400)

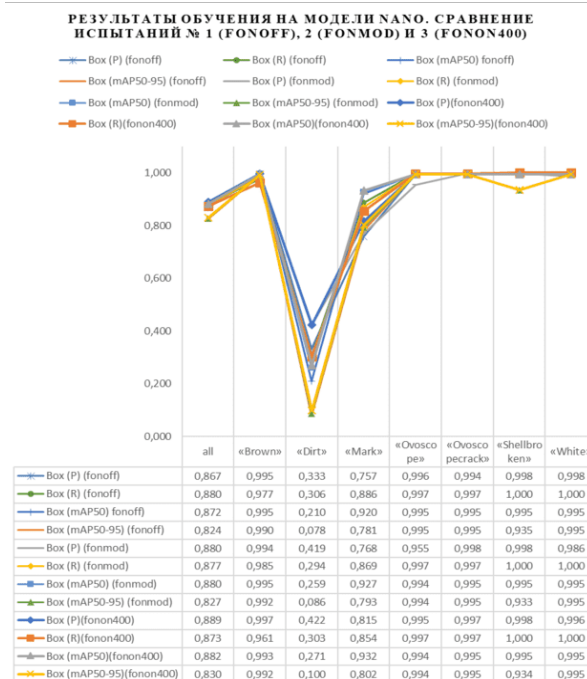


Рис. 46. Сравнение испытаний № 1 «FONOFF», № 2 «FONMOD» и № 3 «FONON400»

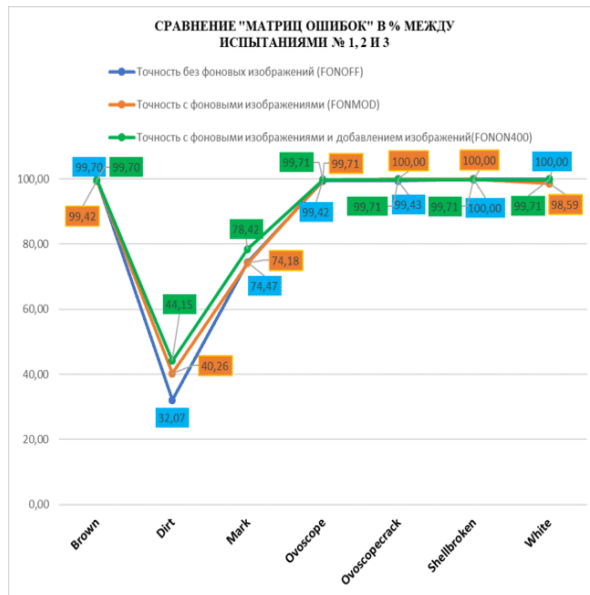


Рис. 47. График сравнения "матриц ошибок" в % между испытаниями № 1, 2 и 3

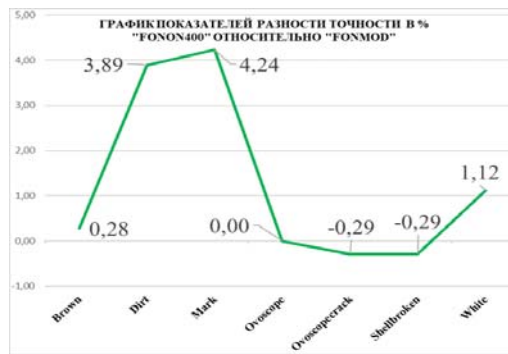


Рис. 48. График показателей разности точности в % "FONON400" относительно "FONMOD"

### Заключение

В результате испытаний видно, что фоновые изображения с нулевой аннотацией повысили точность класса «Dirt» на 8,19% (рис. 26 и 27), однако некоторые классы немного ухудшились.

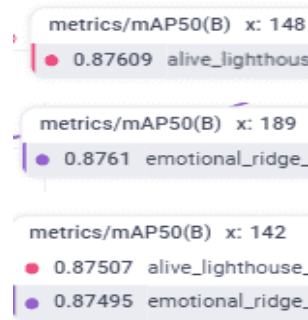
Добавление изображений с аннотациями действительно повысил показатели точности классов «Brown» на 0,28%, «Dirt» на 3,89%, «Mark» на 4,24%, «White» на 1,12%. Остальные ухудшились на 0,29% (рис. 47 и 48).

Также оказалось, что прерывание обучения и дальнейшее возобновление – это не самое лучшее решение. Так, в испытании № 3 был поврежден файл, поэтому возобновлять обучение пришлось на 6 эпох раньше, что негативно сказалось на дальнейшем обучении. Системе понадобилось 41 эпоха, чтобы достигнуть прежних показателей (рис. 49 и 50).

До прерывания, на 148 эпохе была точность mAP50 = 87,609%, на 142 эпохе = 87,507%. После возобновления 142 эпоха = 87,495%, что уже хуже на 0,012%. Улучшение появилось только на 189 эпохе=87,61%.

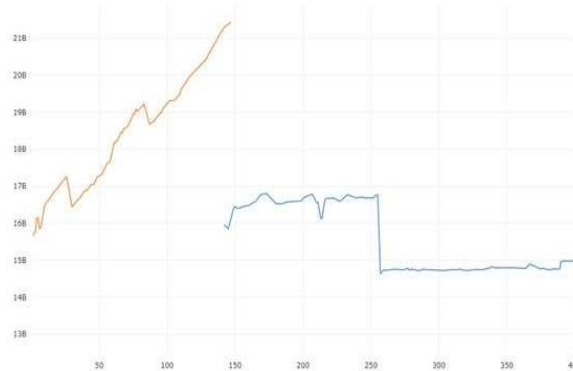


Рис. 49. Фрагмент графика «mAP50» прерывания обучения и его дальнейшее возобновление. Эксперимент № 3.



**Рис. 50.** Показания метрик «mAP50» на 142, 148, 189 эпохах. Эксперимент № 3.

Причиной прерывания явилось повышенное использование RAM на 147 эпохе – 21,4 Гб (рис. 51). Ошибка прерывания – нехватка RAM. Изначально в локальной системе ЭВМ установлено 32 Гб RAM.



**Рис. 51.** Причина прерывания обучения и его дальнейшее ручное возобновление. Использование RAM. Эксперимент № 3.

В эксперименте № 1 тоже был сбой, но возобновление обучения было с «last.pt» - файл был не поврежден, который не занижил точность, а продолжил с того прерванного момента.

Отсюда вывод, при помощи аргументов «save=True» и «save\_period=1» желательно сохранять большие массивы данных каждую эпоху.

В ходе анализа эпох графиков «mAP50» трех испытаний видно, что наибольший рост во время обучения наблюдался в диапазоне от 38 до 99 эпохи. В диапазоне от 100 до 200 эпохи – могут быть как ухудшения, так и улучшения показателей. В диапазоне от 201 до 300 эпохи, в нашем случае – стабильный рост от 0,083% до 0,445%.

В диапазоне от 301 до 400 эпохи, в нашем случае – стабильный рост на 0,164%.

В итоге испытаний сделан вывод о том, что минимальное обучение нейросети должно быть 100 эпох, так как на эпохах от 1 до 100 можно увидеть активную и мгновенную динамику улучшения всех показателей.

Таблица 7

Показатели «mAP50» на разных эпохах

№ испытания	Кодовое имя	Кол-во эпох	Показатели «mAP50»
1	«FONOFF»	300	38 эпоха = 85,556%, 99 эпоха = 87,247%, 230 эпоха = 87,33%
2	«FONMOD»	300	39 эпоха = 85,661%, 94 эпоха = 87,344%, 198 эпоха = 87,879%, 300 эпоха = 88,042%
3	«FONON400»	400	41 эпоха = 85,983%, 99 эпоха = 87,659%, 288 эпоха = 88,057%, 397 эпоха = 88,221%

### Литература

1. Главная - Документация Ultralytics YOLOv8. URL: [https://docs.ultralytics.com/ru/#\\_1](https://docs.ultralytics.com/ru/#_1) (дата обращения: 18.01.2024).
2. Introduction - Roboflow Docs. URL: <https://docs.roboflow.com/> (дата обращения: 18.01.2024).
3. 7 методов, которые могут повысить скорость точности распознавания изображений - Русские Блоги [Электронный ресурс]. URL: <https://russianblogs.com/article/84073812444/> (дата обращения: 18.01.2024).
4. The Difference Between Missing and Null Annotations [Электронный ресурс]. URL: <https://blog.roboflow.com/missing-and-null-image-annotations/> (дата обращения: 18.01.2024).
5. Comet ML - Build better models faster [Электронный ресурс]. URL: <https://www.comet.com> (дата обращения: 18.01.2024).
6. Ultralytics YOLOv8 Docs. Performance Metrics Deep Dive. URL: <https://docs.ultralytics.com/guides/yolo-performance-metrics/> (дата обращения: 18.01.2024).
7. Configuration - Ultralytics YOLOv8 Docs. URL: <https://docs.ultralytics.com/usage/cfg/#train> (Дата обращения: 18.01.2024).
8. Mean Average Precision (mAP) in Object Detection [Электронный ресурс]. URL: <https://blog.roboflow.com/mean-average-precision/> (Дата обращения: 18.01.2024).

## РАЗРАБОТКА АЛГОРИТМА ПОИСКА АССОЦИАТИВНЫХ ПРАВИЛ СЛАБОСТРУКТУРИРОВАННЫХ ДАННЫХ

**Шматко Мария Алексеевна,**  
*МТУСИ, студент, Москва, Россия*

**Василевич Анастасия Дмитриевна,**  
*МТУСИ, студент, Москва, Россия*

**Завидова Елизавета Александровна,**  
*МТУСИ, студент, Москва, Россия*

**Завада Анастасия Руслановна,**  
*МТУСИ, студент, Москва, Россия*

**Осипов Алексей Викторович,**  
*МТУСИ, доцент кафедры «Бизнес-информатика», к.ф.-м.н., Москва, Россия*  
[a.v.osipov@mtusi.ru](mailto:a.v.osipov@mtusi.ru)

### **Аннотация**

*Эта статья подробно рассматривает ассоциативные правила, важный инструмент в машинном обучении и анализе данных, используемый для выявления связей между элементами в больших наборах данных. Статья демонстрирует, как ассоциативные правила могут использоваться для анализа готовности студентов к участию в научных конференциях. В результате чего в процентах были представлены результаты текущего опроса.*

**Ключевые слова:** ассоциативные правила, анкетирование, научная деятельность, машинное обучение.

### **Введение**

Ассоциативные правила были впервые предложены в 1993 году для анализа рыночных корзин. Они позволяют выявлять связи и закономерности между элементами набора данных, например, между товарами в корзине покупок.

Ассоциативные правила широко применяются в различных сферах, включая розничную торговлю, финансовую аналитику, медицинскую диагностику и образование. В розничной торговле ассоциативные правила используются для оптимизации размещения товаров на полках и создания персонализированных рекомендаций покупателям. В финансовой аналитике они помогают выявлять связи между финансовыми событиями и предсказывать риски. В медицинской диагностике ассоциативные правила применяются для выявления связей между симптомами и болезнями, что помогает в диагностике и принятии решений о лечении. В образовательном контексте анализ ассоциативных правил помогает выявить связи между различными переменными, такими как методы обучения, успеваемость и факторы, влияющие на успех студентов. Это позволяет улучшить процессы обучения, адаптировать программы и предоставить персонализированную поддержку для студентов [7].

Ассоциативные правила, зародившиеся в 1993 году для анализа рыночных корзин, стали ключевым элементом в мире машинного обучения и анализа данных. Их применение в образовании, в частности, при анализе готовности первокурсников к участию в научных конференциях, открывает новые перспективы. Давайте более детально рассмотрим, как ассоциативные правила могут преобразовать подход к тестированию студентов.

**Ассоциативные правила** - установление закономерностей между связанными событиями.

Впервые эта задача была предложена для нахождения типичных шаблонов покупок, совершаемых в супермаркетах, поэтому ее еще называют анализом рыночной корзины (market basket analysis).

Ассоциация имеет место в том случае, если несколько событий связаны друг с другом. Например, исследование, проведенное в компьютерном супермаркете, может показать, что 55 % купивших компьютер берут также и принтер или сканер, а при наличии скидки за такой комплект принтер приобре-

тают в 80 % случаев. Располагая сведениями о подобной ассоциации, менеджерам легко оценить, насколько действенна предоставляемая скидка.

Если существует цепочка связанных во времени событий, то говорят о последовательности. Так, например, после покупки дома в 45 % случаев в течение месяца приобретается и новая кухонная плита, а в пределах двух недель 60 % новоселов обзаводятся холодильником [1].

**Целью работы** было исследование мотивации студентов к участию в научных конференциях, а также выявление препятствий, которые могут ограничивать их участие. **Предметом исследования** являлись факторы, влияющие на решение студентов участвовать в конференциях, такие как выбор тематики, возможность представления своих работ, интерес к международной аудитории и другие. Анализ данных с использованием метода поиска ассоциативных правил позволил выделить ключевые факторы и выявить корреляции между ними, что помогает формировать более эффективные программы и стимулировать активное участие студентов в научной жизни.

#### **Ассоциативное правило**

Ассоциативное правило представляет собой метод машинного обучения на базе правил обнаружения интересующих нас связей между переменными в большой базе данных [3].

Оно помогает находить закономерности в результате проведенного анализа произошедшего технологического события или явления.

Задача нахождения ассоциативных правил разбивается на две подзадачи [10]:

1. Нахождение всех наборов элементов, которые удовлетворяют порогу  $\text{minsupport}$ . Такие наборы элементов называются часто встречающимися.
2. Генерация правил из наборов элементов, найденных согласно п.1. с достоверностью, удовлетворяющей порогу  $\text{minconfidence}$ .

#### **Практическое применение метода поиска ассоциативных правил для выявления потенциальных участников научных конференций**

Для изучения метода «Ассоциативные правила» было проведено анкетирование 150 первокурсников из разных групп. Анкета состояла из 15 вопросов, на каждый из которых было 3-5 вариантов ответа.

Пример вопроса для выявления потенциальных участников конференций:

Занимаетесь ли вы самообразованием?

Варианты ответов:

- a) Подписан на курсы
- b) Читаю научную литературу
- c) Хожу на дополнительные занятия
- d) Занимаюсь программированием
- e) Не занимаюсь

Рассмотрим применение метода «Ассоциативные правила» на примере двух вопросов:

1. Какое из утверждений наиболее близко описывает ваше отношение к участию в научных конференциях?
  - a) Я активно ищу возможности для участия
  - b) Мне интересно, но я еще не участвовал
  - c) Мне это не интересно
  - d) Я не знаю, что такое научные конференции
2. В какой тематике вы бы хотели участвовать в научной конференции?
  - a) Технологии и инновации
  - b) Гуманитарные науки
  - c) Естественные науки
  - d) Социальные науки
  - e) Любая тема интересует

Из ответов студентов на эти вопросы мы выделили следующую ассоциацию:

{a} Я активно ищу возможности для участия} => {a} Технологии и инновации} (Поддержка: 0.25, Достоверность: 0.8)

В контексте ассоциативных правил, поддержка относится к вероятности того, что второй элемент (в данном случае, "Технологии и инновации") будет выбран, если первый элемент (в данном случае,

"Я активно ищу возможности для участия") также выбран. То есть, это вероятность или частота встречаемости второго элемента при условии, что присутствует первый.

Достоверность же оценивает степень правдоподобия или уверенности в правиле. Она указывает на то, насколько правдоподобно, что второй элемент появится, если первый также присутствует, основываясь на общем количестве встречаемости обоих элементов и их сочетания [6].

Таким образом, получается:

- Поддержка 0.25 означает, что из всех раз, когда "Я активно ищу возможности для участия" было упомянуто, "Технологии и инновации" появлялись в 25% случаев.

- Достоверность 0.8 указывает на то, что в 80% случаев, когда "Технологии и инновации" были упомянуты, они были связаны с тем, что было сказано о "Я активно ищу возможности для участия".

После проведения анализа всех ответов студентов были получены следующие выводы:

#### 1. Отношение к участию в научных конференциях

Студенты, активно ищущие возможности для участия и заинтересованные в обмене опытом и знаниями, чаще выражают желание принять участие в конференциях. Они чаще выбирают тематику технологий и инноваций, а также видят важность участия для первокурсников.

#### 2. Мотивация для участия

Мотивацией для участия становятся возможность представить свою работу и интерес к теме конференции. Студенты, выбирающие участие за границей, чаще мотивируются желанием представить свои исследования международной аудитории.

#### 3. Препятствия для участия

Факторы, мешающие участию, чаще всего связаны с недостатком времени и финансовых затрат. Также недостаток опыта или знаний может быть значимым фактором для некоторых студентов.

#### 4. Преимущества и выбор специализации

Студенты склонны участвовать как в конференциях своей специализации, так и в других областях, что свидетельствует о широком интересе к разнообразию научных направлений.

#### 5. Самообучение и мотивация для обучения

Самообучение играет важную роль: студенты, занимающиеся самообразованием и саморазвитием, чаще проявляют интерес к участию в научных конференциях.

### **Методика извлечения ассоциативных правил.**

Тест, разработанный для определения готовности первокурсников к участию в конференции, представляет собой набор данных, содержащий информацию о студентах и их ответах на вопросы. Эти данные могут быть анализированы с использованием ассоциативных правил для выявления интересных закономерностей и взаимосвязей [9].

Ассоциативные правила применяются к тесту следующим образом [4]:

1. Подготовка данных – приведение данных теста в подходящий формат для анализа.
2. Генерация правил – использование алгоритма для выявления часто встречающихся вместе ответов на вопросы и преобразования их в ассоциативные правила.
3. Оценка правил – использование метрик, таких как поддержка, уверенность и подъем, для определения значимости ассоциаций, а также учет инновационности.
4. Интерпретация правил – выявление взаимосвязей между вопросами и понимание их влияния на готовность студента.

Один из первых алгоритмов, эффективно решающих подобный класс задач, – это алгоритм APriori. Кроме него были разработаны и другие алгоритмы: DHP, Partition, DIC и другие [2].

### **Алгоритм Apriori**

Один из наиболее распространенных методов извлечения ассоциативных правил – это алгоритм Apriori. Этот метод основан на подсчете частоты встречаемости элементов в наборе данных. Алгоритм состоит из следующих шагов [5]:

1. На первом шаге анализируются часто встречающиеся элементы в наборе данных.
2. Затем создаются кандидаты для ассоциативных правил.
3. На последних шагах проводится фильтрация правил на основе минимальной поддержки и доверительности.



Другой эффективный алгоритм для извлечения ассоциативных правил – это FP-growth. Он работает более эффективно, чем Apriori, особенно при больших объемах данных. Этот алгоритм использует структуру данных FP-дерево для представления и анализа набора данных. Пример алгоритма Apriori приведен ниже.

На первом этапе формируем 1-элементных кандидатов и отбираем часто встречающихся кандидатов, чья поддержка не меньше минимальной.



Рис. 1. Алгоритм Apriori

Далее сканируем базу данных и формируем 3-элементных кандидатов. Отбираем часто встречающихся.

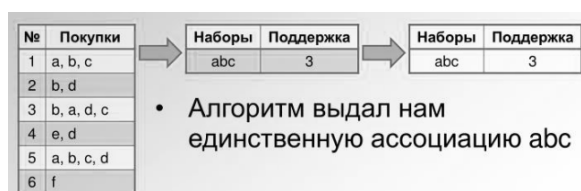


Рис. 2. Алгоритм Apriori

### Алгоритм FP-growth

Рассмотрим пример использования алгоритма FP-growth на языке Python. Набор данных содержит список транзакций ночного магазина. Для каждой транзакции у нас есть список продуктов, которые были куплены во время транзакции.

Transaction No	Products
1	beer, wine, cheese
2	beer, potato chips
3	eggs, flour, butter, cheese
4	eggs, flour, butter, beer, potato chips
5	wine, cheese
6	potato chips
7	eggs, flour, butter, wine, cheese
8	eggs, flour, butter, beer, potato chips
9	wine, beer
10	beer, potato chips
11	butter, eggs
12	beer, potato chips
13	flour, eggs
14	beer, potato chips
15	eggs, flour, butter, wine, cheese
16	beer, wine, potato chips, cheese
17	wine, cheese
18	beer, potato chips
19	wine, cheese
20	beer, potato chips

Рис. 3. Список транзакций

Идея алгоритма FP Growth состоит в том, чтобы найти наиболее часто встречающиеся наборы элементов в наборе данных, будучи быстрее, чем алгоритм Apriori. Алгоритм Apriori в основном обращается к набору данных и проверяет совпадение продуктов в наборе данных.

На первом шаге нужно ввести данные в виде списка транзакций. Каждая транзакция представляет собой список предметов [8].

```

transactions = [
    ["beer", "wine", "cheese"],
    ["beer", "potato chips"],
    ["eggs", "flower", "butter", "cheese"],
    ["eggs", "flower", "butter", "beer", "potato chips"],
    ["wine", "cheese"],
    ["potato chips"],
    ["eggs", "flower", "butter", "wine", "cheese"],
    ["eggs", "flower", "butter", "beer", "potato chips"],
    ["wine", "beer"],
    ["beer", "potato chips"],
    ["butter", "eggs"],
    ["beer", "potato chips"],
    ["flower", "eggs"],
    ["beer", "potato chips"],
    ["eggs", "flower", "butter", "wine", "cheese"],
    ["beer", "wine", "potato chips", "cheese"],
    ["wine", "cheese"],
    ["beer", "potato chips"],
    ["wine", "cheese"],
    ["beer", "potato chips"],
]

```

Далее нужно закодировать его с помощью кодировщика, сопоставимого с One-Hot Encoder.

```

import pandas as pd
from mlxtend.preprocessing import TransactionEncoder

my_transactionencoder = TransactionEncoder()
my_transactionencoder.fit(transactions)
encoded_transactions = my_transactionencoder.transform(transactions)
encoded_transactions_df = pd.DataFrame(encoded_transactions, columns=my_transactionencoder.columns_)
encoded_transactions_df

```

Полученный датафрейм выглядит следующим образом

	beer	butter	cheese	eggs	flower	potato chips	wine
0	True	False	True	False	False	False	True
1	True	False	False	False	False	True	False
2	False	True	True	True	True	False	False
3	True	True	False	True	True	True	False
4	False	False	True	False	False	False	True
5	False	False	False	False	False	True	False
6	False	True	True	True	True	False	True
7	True	True	False	True	True	True	False
8	True	False	False	False	False	False	True
9	True	False	False	False	False	True	False
10	False	True	False	True	False	False	False
11	True	False	False	False	False	True	False
12	False	False	False	True	True	False	False
13	True	False	False	False	False	True	False
14	False	True	True	True	True	False	True
15	True	False	True	False	False	True	True
16	False	False	True	False	False	False	True
17	True	False	False	False	False	True	False
18	False	False	True	False	False	False	True
19	True	False	False	False	False	True	False

Рис. 4. Полученный датафрейм

Следующим шагом является вычисление частых наборов элементов. Вы можете использовать функцию `fpgrowth` из `mlxtend`, как показано ниже:

```
min_support = 7/len(transactions)
from mlxtend.frequent_patterns.fpgrowth import fpgrowth
frequent_itemsets = fpgrowth(encoded_transactions_df, min_support=min_support,
use_colnames = True)
frequent_itemsets
```

В качестве последнего шага нам нужно использовать функцию `association_rules` для преобразования этих частых наборов элементов в правила ассоциации. Это можно сделать с помощью следующего кода:

```
from mlxtend.frequent_patterns import association_rules
association_rules(frequent_itemsets, metric="confidence", min_threshold=0.7)
```

Окончательные правила ассоциации, вычисленные с использованием FP Growth:

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
0	(wine)	(cheese)	0.40	0.40	0.35	0.875000	2.187500	0.1900	4.800
1	(cheese)	(wine)	0.40	0.40	0.35	0.875000	2.187500	0.1900	4.800
2	(beer)	(potato chips)	0.55	0.45	0.45	0.818182	1.818182	0.2025	3.025
3	(potato chips)	(beer)	0.45	0.55	0.45	1.000000	1.818182	0.2025	inf

Рис. 5. Правила ассоциации

### Заключение

В проведенном анализе были получены следующие выводы:

1. 25% студентов, активно ищущих возможности для участия, выбирают участие в конференциях по тематике "Технологии и инновации".
2. Достоверность связи между активным поиском возможностей для участия и выбором тематики "Технологии и инновации" составляет 80%.
3. Более 40% студентов мотивируются возможностью представления своих работ при участии в конференциях.
4. Более 30% студентов заинтересованы в участии в конференциях за границей с целью представления своих исследований международной аудитории.
5. Недостаток времени и финансовых затрат являются препятствием для участия в конференциях для более 40% студентов.
6. Более 30% студентов выбирают участие в конференциях не только в своей специализации, но и в других областях.
7. Самообучение и саморазвитие стимулируют интерес к участию в научных конференциях для более 30% студентов.

Студенты проявляют интерес и мотивацию к участию в научных конференциях, особенно если они активно ищут возможности для участия и занимаются самообучением. Однако, препятствия в виде недостатка времени и финансовых затрат могут ограничивать их участие.

Использование метода поиска ассоциативных правил на основе ответов студентов на вопросы анкеты об участии в научных конференциях позволяет выделить ключевые факторы, мотивацию и барьеры, которые влияют на их решение участвовать. Это помогает формировать более индивидуализированные и адаптированные программы, способствуя активному участию студентов в научной жизни и исследовательской деятельности.

Анализ данных с помощью метода поиска ассоциативных правил не только помогает понять предпочтения студентов, но и стимулирует дальнейшее исследование и адаптацию образовательных программ под их потребности и интересы.

Такой анализ предлагает глубокое понимание мотивации студентов и их отношения к участию в научных конференциях, что позволяет образовательным учреждениям лучше подготовиться к созданию стимулирующей среды для активной учебной деятельности.

### Литература

1. *Agrawal Rakesh, Srikant Ramakrishnan*. Fast algorithms for mining association rules in large databases. Чили, 1994. С. 487-499
2. *Charu C. Agrawal, Jiawei Han*. Frequent Pattern Mining. Швейцария: Springer, 2014.
3. *Han J., Kamber M*. Data Mining: Concepts and Techniques. США, 2011.
4. *Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И.* Методы и модели анализа данных: OLAP и DataMining. СПб: БХВ-Петербург, 2008.
5. *Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И.* Технологии анализа данных. Data Mining, Visual Mining, Text Mining, OLAP. СПб: БХВ-Петербург, 2007 г.
6. *Паклин Н.Б., Орешков В.И.* Бизнес-аналитика: от данных к знаниям. СПб: Изд. Питер, 2009.
7. *Antti Syvajarvi*. Data Mining in Public and Private Sectors: Organizational and Government Applications (Premier Reference Source) - Information Science Reference, 2010.
8. *Senthil Kumar A.V*. Knowledge Discovery Practices and Emerging Applications of Data Mining: Trends and New Domains - IGI Global, 2010.
9. *Paolo Giudici*. Applied Data Mining. Wiley, 2009.
10. *Чубукова И.А.* DataMining: учебное пособие. М.: Интернет-университет информационных технологий: БИНОМ: Лаборатория знаний, 2006.