

Санкт-Петербургский политехнический университет Петра Великого
Физико-механический институт

ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ № 4

«Анализ промптов и моделей для структурного анализа научного текста»
по дисциплине «Автоматизация научных исследований»

Выполнил
студент гр. № 5040102/50201

Гордейко Н.Л.

Новиков Ф.А.

Преподаватель:

Санкт-Петербург
2025 г.

Содержание

Статья и модели	5
Промпты	5
1 Промпт 1	7
1.1 Ответ модели Copilot.....	7
1.2 Ответ модели Gemini.....	9
2 Промпт 2	12
2.1 Ответ модели Copilot.....	12
2.2 Ответ модели Gemini.....	15
3 Промпт 3	17
3.1 Ответ модели Copilot (таблица).....	17
3.2 Ответ модели Gemini (таблица).....	19
4 Промпт 4	22
4.1 Ответ модели Copilot.....	22
4.2 Ответ модели Gemini.....	25
5 Промпт 5	27
5.1 Ответ модели Copilot.....	27
5.2 Ответ модели Gemini.....	31
6 Сводная таблица оценок	33
7 Сравнительный анализ моделей	34
8 Сравнительный анализ промптов	36
Заключение	37

ЗАДАНИЕ

В рамках лабораторной работы необходимо произвести:

1. Выбор и подготовка статьи. Найти одну научную статью (желательно по теме, связанной с НИР или ВКР), в структуре которой заведомо присутствуют неявные или отсутствующие элементы.
- 1.1. Рекомендация. Для повышения сложности анализа моделям, необходимо подготовить текст статьи, удалив из него явные заголовки разделов («Введение», «Заключение» и т.п.).
2. Выбор моделей. Выбрать две различные ИИ-модели (например, Gemini, ChatGPT, GigaChat, DeepSeek и др.) для проведения эксперимента.
3. Эксперимент. Прогнать полный текст выбранной статьи через обе модели, используя каждый из четырех предложенных промптов.
4. Анализ результатов. Провести детальный сравнительный анализ 8 полученных ответов, выявляя сильные и слабые стороны различных моделей при работе с различными промптами.
5. Оптимизация. Выбрать одну ИИ-модель, которая, по вашему мнению, показала наилучший потенциал, и на основе выявленных недостатков существующих промптов, составить собственный промпт, который, по вашему мнению, должен дать наилучший результат по всем критериям.
6. Финальная проверка. Прогнать статью через выбранную модель, используя свой авторский промпт.

Критерии оценивания ответов ИИ

Критерии (по 10-балльной шкале) для оценки каждого из ответов:

Критерий	Описание
Полнота структурных элементов	Насколько полно ИИ выявил все элементы (актуальность, цель, задачи, методология, результаты, выводы).
Корректность и отсутствие выдуманной информации	ИИ не должен придумывать информацию, ошибочно определять методологию или путать выводы с результатами.
Критическая глубина анализа	Способность ИИ выявить логические противоречия, нарушения академического стиля, отсутствие связности цели и выводов.
Четкость формата ответа	Строгое соблюдение запрошенного в промпте формата (например, корректное отображение таблиц для Промта 3).

Выводы

В качестве выводов, в отчёте следует указать:

- **Сводная таблица оценок:**
 - Таблица с оценками (1–10 баллов) всех ответов.
- **Анализ сильных и слабых сторон:**

- **Слабые стороны моделей:** Какие ошибки были характерны для каждой из двух выбранных моделей.
- **Слабые стороны промптов:** Какой из 4 промптов оказался наименее эффективным и почему.
- **Сильные стороны промптов:** Какой промпт оказался наиболее эффективным и почему.
- **Ваш собственный промпт:**
 - Полный текст разработанного **промпта**.
 - Указание **модели**, для которой этот промпт разработан.
- **Обоснование и конечный результат:**
 - Подробное обоснование, **почему** авторский промпт был составлен именно так (какие проблемы предыдущих промпов он решает).
 - Финальный ответ, полученный от ИИ с использованием разработанного промпта.
 - Оценка этого финального ответа по 4 критериям и заключение об успехе оптимизации.

Статья и модели

Статья: <https://cyberleninka.ru/article/n/o-neyronnyh-arhitekturah-izvlecheniya-priznako>

Модели: Copilot, Gemini.

Промпты

Промпт 1.

Проанализируй этот текст на соответствие структуре научной работы.

Промпт 2.

Проанализируй научный текст на соответствие структуре. Найди, есть ли в нём актуальность, цели, задачи, предмет и объект исследования, методологию, результаты и выводы. Не придумывай информацию и опирайся только на текст.

Промпт 3.

Проанализируй научный текст по структурным критериям научной статьи. Определи:

- актуальность;
- цель;
- задачи исследования;
- объект и предмет;
- методологию;
- результаты;
- выводы.

Представь ответ в виде таблицы: колонка 1 — элемент структуры; колонка 2 — найденный фрагмент текста (цитата); колонка 3 — краткий комментарий. Если элемент отсутствует или выражен неявно — отметь это.

Промпт 4.

Проанализируй научный текст с опорой на требования существующих ГОСТов и академические стандарты. Выполни:

- проверь логическую связность между актуальностью, целью и задачами;
- выяви противоречия;
- оцени корректность формулировок;
- сравни цель и задачи (должны быть согласованы);
- отметь логические провалы;
- выдели отсутствующие элементы структуры;

- оцени соответствие выводов заявленным задачам.

Представь ответ структурировано: блок «оценка структуры», блок «ошибки», блок «вы-вод».

Промпт 5 (собственный).

Проанализируй научный текст с позиции экспертной оценки качества исследования. Выполните следующие шаги:

1. Определи тип исследования (теоретическое, эмпирическое, экспериментальное, обзорное, методологическое и т.д.). Укажи, на основании каких признаков сделан вывод.

2. Проверь соответствие текста жанру научной статьи, оценив:

- наличие научной новизны;
- степень обоснованности утверждений;
- корректность использования терминологии;
- наличие ссылок на источники или их имитации;
- полноту аргументации.

3. Оцени структуру текста по критериям научного дискурса:

- логичность переходов между частями;
- наличие причинно-следственных связей;
- отсутствие противоречий;
- достаточность данных для выводов.

4. Проверь соответствие текста требованиям к научному стилю:

- точность;
- однозначность;
- отсутствие эмоциональности;
- корректность формулировок;
- отсутствие разговорных конструкций.

5. Сформируй экспертное заключение, включив:

- сильные стороны текста;
- слабые стороны;
- рекомендации по улучшению;
- итоговую оценку качества текста по 10-балльной шкале.

Представь ответ в структурированном виде: тип исследования → анализ структуры → анализ стиля → сильные стороны → слабые стороны → рекомендации → итоговая оценка.

1 Промпт 1

1.1 Ответ модели Copilot

Анализ структуры текста на соответствие требованиям научной работы.

Представленный текст соответствует типовой структуре научной статьи и включает все ключевые элементы, характерные для академических публикаций в области компьютерных наук.

1. Заголовочный блок

Статья содержит корректно оформленный заголовок, сведения об авторах, их аффили-ации и контактные данные. Присутствуют аннотации на русском и английском языках, а также ключевые слова. Данный набор элементов полностью соответствует стандартам научных журналов.

2. Введение

Во введении обозначена актуальность темы, сформулирована общая проблематика и контекст исследования. Указаны причины, по которым выбор базовой сети является сложной задачей, а также обоснована необходимость разработки решений для устройств с ограниченными вычислительными ресурсами. Введение логично подводит к постановке задачи.

3. Постановка задачи

Раздел содержит формализованное описание процесса извлечения признаков, приведены математические выражения, определены цели и задачи исследования. Формулировки четкие, структура соответствует требованиям к научной постановке задачи.

4. Обзор и обоснование выбора архитектур

Приведён критерий отбора моделей (ограничение по числу параметров и точности на ImageNet). Даны ссылки на источники и таблица с характеристиками моделей. Раздел выполняет функцию краткого обзора литературы и обоснования выбора методов.

5. Описание набора данных и методики эксперимента

Подробно описан процесс формирования датасета ObjectDet, включая:

- источник данных;
- процедуру выделения объектов;
- изменение размера изображений;
- используемые методы аугментации;
- структуру выборок (train/test/val);

- количество классов.

Методика оценивания моделей изложена последовательно и включает несколько этапов: кластеризацию признаков, обучение классификатора, измерение VRAM, интегральную оценку.

6. Экспериментальная часть

Экспериментальная часть включает:

- описание используемых методов визуализации и кластеризации (UMAP, k-means);
- критерий оценки кластеризации (индекс Дэвиса–Булдина);
- таблицы с результатами кластеризации;
- описание архитектуры классификатора;
- параметры обучения (оптимизатор, число эпох, размер батча);
- таблицы с результатами обучения и тестирования;
- измерение потребления видеопамяти;
- интегральную оценку моделей.

Структура раздела соответствует требованиям к экспериментальному исследованию: представлены данные, методика, результаты и их сравнение.

7. Обсуждение результатов

Обсуждение присутствует, но носит краткий характер. Основные выводы по таблицам сформулированы, однако интерпретация результатов ограничена и не включает глубокого анализа причин различий между моделями. Тем не менее, минимальные требования к разделу выполнены.

8. Выводы

Выводы подводят итог проведённому исследованию, перечисляют ключевые результаты и подтверждают достижение поставленных целей. Структура соответствует стандарту.

9. Ссылки на литературу

В тексте присутствуют ссылки на источники [1]–[22], однако сам список литературы в предоставленном фрагменте отсутствует. Для полноценной научной статьи он обязателен.

Итоговая оценка

Статья в целом соответствует структуре научной работы и включает все основные элементы: аннотации, введение, постановку задачи, описание методики, экспериментальную часть и выводы. Стиль изложения научный, логика последовательная, структура выдержанна. Единственным отсутствующим элементом в представленном фрагменте является список литературы, который должен быть включён в финальную версию статьи.

Если требуется, могу подготовить:

- формализованную рецензию;
- рекомендации по улучшению структуры;
- проверку на соответствие ГОСТ Р 7.0.7–2021 или требованиям конкретного журнала.

1.2 Ответ модели Gemini

1. Метаданные и вводная часть

Заголовок и сведения об авторах. Статья содержит названия на русском и английском языках, УДК, аффилиации авторов и DOI.

Аннотация (Abstract). Включает описание цели работы (исследование эффективности нейросетей для мобильных устройств), методологию (двуэтапный подход на базе Faster R-CNN) и краткие выводы (преимущество Edge NeXt-S).

Ключевые слова. Приведены на обоих языках.

2. Введение (Introduction)

В данном разделе авторы:

- обосновывают актуальность темы развитием глубокого обучения и потребностью в компьютерном зрении на устройствах с ограниченными ресурсами;
 - дают определение базовой сети (backbone network) как основного инструмента извлечения признаков.
-

3. Постановка задачи (Problem Statement)

Раздел формализует процесс извлечения признаков и классификации через математические функции. Сформулированы конкретные цели и задачи исследования (отбор архитектур, формирование набора данных, разработка методики оценки).

4. Методология исследования (Methods)

Авторы подробно описывают процесс подготовки эксперимента:

- **Обоснование выбора моделей.** Ограничение по количеству параметров (до 25 млн) и точности на ImageNet (выше 75%). Сформирована таблица из 10 исследуемых архитектур.
 - **Формирование набора данных.** Описано создание датасета ObjectDet на базе ObjectNet.
 - **Аугментация.** Перечислены техники расширения данных (инверсия, удаление пикселей, аффинные преобразования и др.).
 - **Методика оценивания.** Включает 4 этапа, от анализа разделяющей способности признаков до определения потребляемой оперативной памяти (VRAM).
-

5. Результаты и обсуждение (Results)

Представлены результаты численных исследований:

- **Качество кластеризации.** Оценено с помощью индекса Дэвиса-Булдина (DBI).
 - **Технические параметры.** Измерено потребление VRAM при разном размере бат-чей.
 - **Точность (Accuracy).** Приведены значения top-1 и top-5 точности для каждой модели.
 - **Интегральная оценка.** Предложена формула и шкала для вычисления итогового рейтинга моделей по сумме критериев.
-

6. Выводы (Conclusions)

Раздел резюмирует итоги работы:

- констатируется создание датасета ObjectDet;
 - указывается, что гибридная сеть EdgeNeXt-S показала наилучший баланс между производительностью и точностью;
 - предложенная методика интегрального оценивания рекомендуется для дальнейшего использования исследователями.
-

7. Список литературы (References)

Содержит 22 источника, включая современные зарубежные публикации (2019–2023 гг.) по архитектурам ConvNeXt, DaViT, EdgeNeXt и методам оптимизации.

Резюме

Статья структурирована логично, содержит все необходимые атрибуты академической публикации и обеспечивает повторяемость эксперимента за счёт детального описания методик и используемых параметров.

2 Промпт 2

2.1 Ответ модели Copilot

Анализ структуры научного текста.

1. Актуальность исследования

Есть. Актуальность сформулирована во введении. В тексте подчёркивается:

- рост применения методов глубокого обучения;
- сложность выбора базовой сети для задач компьютерного зрения;
- необходимость разработки решений для устройств с ограниченной вычислительной мощностью.

Цитаты по смыслу:

- «выбор подходящей BN . . . сложен»;
- «растущая потребность . . . на устройствах с ограниченной вычислительной мощностью».

Это соответствует критерию актуальности.

2. Цель исследования

Есть. Цель указана явно в разделе «Постановка задачи исследования»:

«Цель настоящего исследования состоит в сравнении показателей производительности и качества метрик, используя различные BN, для дальнейшего применения наилучшей BN в устройствах с ограниченной вычислительной мощностью.»

3. Задачи исследования

Есть. В тексте перечислены пять задач:

1. Сформировать критерии и отобрать BN.
2. Дать краткое описание BN.
3. Сформировать набор данных.
4. Сформировать методику оценивания BN.
5. Выбрать BN на основе методики.

4. Объект исследования

Есть, но не выделен явно. Объект можно определить только по тексту, без домысливания:

Объектом исследования являются нейронные архитектуры (базовые сети), применяемые в задачах компьютерного зрения.

Это следует из формулировок:

- «данная работа посвящена исследованию эффективности различных моделей нейронных сетей. . . »;
- «базовая сеть. . . является основным блоком».

5. Предмет исследования

Есть, но также не выделен явно. Предметом исследования является эффективность различных архитектур BN по ряду критериев.

Это следует из текста:

- «исследование эффективности различных сетевых архитектур по таким критериям как. . . ».

6. Методология исследования

Есть. Методология описана подробно и включает:

- формирование датасета ObjectDet;
- аугментации изображений;
- использование UMAP для визуализации признаков;
- кластеризацию методом k-средних;
- оценку кластеризации индексом Дэвиса–Булдина;
- обучение классификатора с использованием различных BN;
- использование оптимизатора Lion + Lookahead;
- измерение VRAM;
- интегральную оценку по 8 критериям.

Методология представлена достаточно полно.

7. Результаты исследования

Есть. Результаты представлены в виде:

- таблиц DBI (табл. 3);
- VRAM (табл. 4);
- результатов обучения (табл. 5);
- интегральной оценки моделей (табл. 6).

Также приведены текстовые интерпретации:

- «наихудшими показателями обладают . . .»;
- «наилучшую эффективность показала . . .»;
- «наиболее точной оказалась . . .».

8. Выводы

Есть. Раздел «Выводы» содержит:

- итог по формированию датасета;
- оценку разделяющей способности признаков;
- сравнение архитектур по точности и VRAM;
- итог по интегральному критерию.

Выводы соответствуют требованиям научной статьи.

Итог

В тексте присутствуют все ключевые структурные элементы научной работы:

Элемент	Наличие
Актуальность	Да
Цель	Да
Задачи	Да
Объект	Да (неявно)
Предмет	Да (неявно)
Методология	Да
Результаты	Да
Выводы	Да

Структура научной статьи соблюдена полностью, хотя объект и предмет исследования не выделены отдельными формулировками, но однозначно определяются из содержания.

2.2 Ответ модели Gemini

1. Актуальность

Актуальность исследования обоснована растущей потребностью в использовании приложений компьютерного зрения, включая обнаружение объектов, на устройствах с ограниченной вычислительной мощностью. Авторы отмечают, что существующие работы не раскрывают вопрос применения базовых сетей (backbone networks) на таких устройствах, а методика их оценки остаётся открытой задачей.

2. Цели и задачи

Цель исследования: сравнение показателей производительности и качества метрик различных базовых сетей (BN) для выбора наилучшей архитектуры для применения в устройствах с ограниченной мощностью.

Задачи исследования:

- сформировать критерии и отобрать базовые сети;
- дать описание отобранным BN;
- сформировать набор данных для экспериментов;
- разработать методику оценивания BN;
- выбрать оптимальную базовую сеть на основе методики.

3. Объект и предмет исследования

Хотя термины «объект» и «предмет» явно не выделены подзаголовками, из контекста следует:

- **Объект:** процесс распознавания и классификации объектов на устройствах с ограниченными вычислительными ресурсами;
- **Предмет:** эффективность различных архитектур нейронных сетей (базовых сетей), используемых для извлечения признаков в структуре Faster R-CNN.

4. Методология

Авторы предлагают комплексную методику оценивания, включающую:

- **Исследование разделяющей способности:** эксперимент на наборе данных CIFAR-10 с использованием визуализации UMAP и кластеризации k-средних;
- **Оценка качества классификации:** обучение на специально сформированном наборе данных ObjectDet (базируется на ObjectNet с применением вероятностной аугментации);

- **Технический анализ:** определение объёма занимаемой оперативной памяти (VRAM) для графического процессора;
- **Интегральная оценка:** применение 10-балльной шкалы по совокупности критериев (точность, GFLOPS, количество параметров, DBI, расход памяти и др.).

5. Результаты

В ходе численных исследований получены следующие данные:

- **Потребление памяти:** наилучшую эффективность по использованию VRAM показала архитектура ConvNeXt-N;
- **Точность:** самой точной моделью оказалась трансформероподобная DaViT-T;
- **Разделяющая способность:** проведена оценка по индексу Дэвиса-Булдина (DBI), где наихудшие показатели выявлены у EdgeFormer-S и TinyViT-21M.

6. Выводы

- сформирован датасет ObjectDet со 172 классами;
- установлено, что гибридная сеть EdgeNeXt-S обладает наилучшим балансом между производительностью, рабочестью и точностью, показав высший результат по интегральному критерию;
- предложенная методика интегрального оценивания рекомендована для дальнейшего использования при выборе архитектур для мобильных и edge-устройств.

3 Промпт 3

3.1 Ответ модели Copilot (таблица)

Таблица анализа структуры научной статьи

Элемент структуры	Фрагмент текста (цитата)	Комментарий
Актуальность	«. . . стоит отметить растущую потребность в использовании приложе-ний компьютерного зрения, в том числе связанных с обнаружением объектов на устройствах с ограниченной вычислительной мощно-стью.»	Актуальность сформулирована явно: подчёркнута потребность в эффективных моделях для устройств с ограниченными ре-сурсами.
Цель исследования	«Цель настоящего исследования состоит в сравнении показателей производительности и качества метрик, используя различные BN, для дальнейшего применения наилучшей BN в устройствах с ограниченной вычис- лительной мощностью.»	Цель указана чётко и однозначно.
Задачи исследования	«В работе ставятся следующие задачи: 1. Сформировать критерии. . . 2. Дать краткое описание. . . 3. Сформировать набор данных. . . 4. Сформировать методику. . . 5. Выбрать BN. . . »	Задачи перечислены явно и полностью.
Объект исследования	Не выделен явно. Возможный фрагмент: «. . . эффективности различных моделей нейронных сетей в задачах обнаружения объектов. . . »	Объект выражен неявно: исследуются модели нейронных сетей, применяемые в задачах компьютерного зрения.
Предмет исследования	Не выделен явно. Возможный фрагмент: «. . . эффективности различных сетевых архитектур по таким кrite-риям как. . . »	Предмет выражен неявно: эф-фективность архитектур BN по ряду критериев (точность, DBI, VRAM, GFLOPs и др.).

Методология исследования	<p>«. . . был сформирован набор данных ObjectDet. . . », «. . . применялась вероятностная аугментация. . . », «. . . эксперимент с применением визуализации при помощи модели UMAP. . . », «. . . кластеризация. . . методом K-средних. . . », «. . . использовался оптимизатор Lion. . . », «. . . определение занимаемой оперативной памяти. . . », «. . . интегральная оценка. . . »</p>	<p>Методология описана подробно: формирование датасета, аугментации, визуализация, кластеризация, обучение моделей, измерение VRAM, интегральная оценка.</p>
Результаты исследования	<p>Таблицы 3–6 и текстовые выводы: «Результаты анализа эффективности BN по критериям DBI. . . », «Результаты обучения и тестирования моделей. . . », «. . . наилучшую эффективность показала архитектура ConvNeXt-N. . . », «. . . наиболее точной оказалась модель DaViT-T. . . »</p>	<p>Результаты представлены в виде таблиц и краткого анализа.</p>
Выводы	<p>«Проведена оценка разделяющей способности. . . », «. . . использовались такие критерии как точность. . . », «. . . наилучшую эффективность показала архитектура ConvNeXt-N. . . », «. . . наиболее точной оказалась модель DaViT-T. . . », «. . . наилучшее значение по интегральному критерию показала. . . »</p>	<p>Выводы присутствуют, подводят итог проведённому исследованию.</p>

3.2 Ответ модели Gemini (таблица)

Элемент структуры	Найденный фрагмент текста (цитата)	Комментарий
Актуальность	«...стоит отметить растущую потребность в использовании приложений компьютерного зрения, в том числе связанных с обнаружением объектов на устройствах с ограниченной вычислительной мощностью [2]»	Обоснована практической потребностью внедрения ИИ в мобильные и edge-устройства.

Цель	«Цель настоящего исследования состоит в сравнении показателей производительности и качества метрик, используя различные BN, для дальнейшего применения наилучшей BN в устройствах с ограниченной вычислительной мощностью»	Цель чётко сформулирована как сравнительный анализ для выбора оптимальной архитектуры.
Задачи	«В работе ставятся следующие задачи: 1. Сформировать критерии и... отобрать BN. 2. Дать краткое описание... 3. Сформировать набор данных... 4. Сформировать методику оценивания... 5. Выбрать BN...»	Представлен конкретный перечень этапов для достижения цели.
Объект и предмет	В тексте явно не выделены. Объект (неявно): процессы распознавания объектов. Предмет (неявно): эффективность нейронных архитектур извлечения признаков.	Объект и предмет сформулированы неявно, выводятся из контекста.
Методология	«...методика оценивания BN, состоящая из следующих этапов: 1. Исследование разделяющей способности... 2. Определение качества классификации... 3. Определение занимаемой оперативной памяти. 4. Оценивание систем на основе критериев. »	Описан комплексный подход, включающий численные эксперименты, кластеризацию и технические замеры.
Результаты	«Наилучшее значение по интегральному критерию показала гибридная сеть EdgeNeXt-S... наилучшую эффективность архитектуры ConvNeXt-N... наиболее точной оказалась трансформероподобная модель DaViT-T»	Приведены конкретные данные по точности, памяти и итоговому рейтингу моделей.

Выводы	«Архитектуру EdgeNeXt-S... рекомендуется применять разработчикам для систем компьютерного зрения, базирующихся на устройствах с ограниченной вычислительной сложностью»	Подведён итог работы с практическими рекомендациями по выбору конкретной сети.
--------	---	--

4 Промпт 4

4.1 Ответ модели Copilot

Оценка структуры

1. Логическая связность между актуальностью, целью и задачами

- Актуальность обоснована: подчёркнута необходимость эффективных моделей для устройств с ограниченной вычислительной мощностью.
- Цель логично вытекает из актуальности: сравнить архитектуры BN и выбрать наиболее подходящую.
- Задачи соответствуют цели: они направлены на формирование критериев, описание моделей, создание датасета, разработку методики и выбор лучшей BN.

Связность соблюдена, структура соответствует ГОСТ: актуальность → цель → задачи → методика → результаты → выводы.

2. Корректность формулировок

- Формулировки цели и задач корректны, конкретны, не содержат двусмысленностей.
- Термины используются последовательно.
- Математические обозначения введены корректно.
- Методика описана достаточно подробно.

3. Соответствие цели и задач

Цель: «. . . сравнение показателей производительности и качества метрик. . . для дальнейшего применения наилучшей BN. . . ».

Задачи:

1. Сформировать критерии.
2. Описать BN.
3. Сформировать датасет.
4. Сформировать методику оценивания.
5. Выбрать BN.

Полное соответствие: каждая задача является шагом к достижению цели.

4. Методология

Методология включает:

- формирование датасета;

- аугментации;
- кластеризацию (UMAP + k-means);
- оценку DBI;
- обучение классификатора;
- измерение VRAM;
- интегральную оценку.

Методология соответствует требованиям ГОСТ: описана, воспроизводима, включает критерии и параметры эксперимента.

5. Результаты

Результаты представлены в виде:

- таблиц (DBI, VRAM, точности, интегральной оценки);
- краткого анализа.

Результаты соответствуют задачам и методике.

6. Выводы

Выводы:

- подведены итоги по датасету;
- дана оценка разделяющей способности;
- сравнение моделей по точности и VRAM;
- указана лучшая модель по интегральному критерию.

Выводы соответствуют заявленным задачам, но не полностью раскрывают интерпретацию результатов (см. ошибки).

Ошибки и недочёты

1. Отсутствующие элементы структуры

По ГОСТ 7.32–2017 в статье должны быть:

- обзор литературы — есть частично, но не оформлен как отдельный раздел;
- обсуждение результатов — практически отсутствует, есть только перечисление;
- ограничения исследования — отсутствуют;
- список литературы — в фрагменте отсутствует, хотя ссылки есть.

2. Логические провалы

- В выводах не объясняется, почему конкретные модели показали такие результаты.
- Нет анализа причин низкой эффективности отдельных BN.
- Нет сопоставления результатов кластеризации и классификации (хотя это могло бы усилить научную ценность).
- Не указано оборудование (GPU/CPU), что снижает воспроизводимость.

3. Противоречия

Противоречий в логике исследования не выявлено. Однако есть несоответствие в формулировке критерия DBI:

- в тексте указано «DBI (по максимальному значению)»;
- но по определению DBI лучше минимальное значение.

Это методологическая ошибка.

4. Недостаточная интерпретация результатов

- Таблицы приведены, но обсуждение минимально.
- Не указано, как результаты влияют на практическое применение.
- Не объяснено, почему интегральный критерий даёт именно такие оценки.

5. Незавершённость текста

Фрагмент обрывается в конце выводов («. . . показала гибр. . . »), что нарушает структуру.

Вывод

Статья в целом соответствует требованиям ГОСТ и академическим стандартам: структура логична, цель и задачи согласованы, методология описана, результаты представлены, выводы сформулированы.

Однако имеются недостатки:

- отсутствует полноценное обсуждение результатов;
- не оформлен обзор литературы;
- не указаны ограничения исследования;
- допущена ошибка в интерпретации критерия DBI;
- выводы требуют большей аналитичности;
- отсутствует список литературы в представленном фрагменте.

В целом структура научной статьи выдержана, но для полного соответствия ГОСТ требуется доработка аналитической части, устранение методологических неточностей и завершение выводов.

4.2 Ответ модели Gemini

Вот структурированный анализ статьи без ссылок на источники и сторонних примечаний, удобный для копирования.

Блок «Оценка структуры»

Текст соответствует классическому формату научной статьи (IMRaD) и содержит следующие элементы:

- **Актуальность:** сформулирована через потребность внедрения алгоритмов компьютерного зрения в мобильные и edge-устройства с малым объёмом ресурсов.
- **Цель:** чётко определена как сравнительный анализ различных нейросетевых архитектур (базовых сетей) для выбора оптимальной модели.
- **Задачи:** представлены последовательным списком (отбор моделей, формирование датасета, разработка методики, проведение эксперимента).
- **Методология:** описана как многоэтапный процесс, включающий оценку разделяющей способности признаков, точности классификации и замеров аппаратных ресурсов (память, вычислительная сложность).
- **Результаты:** подкреплены таблицами с конкретными метриками (Accuracy, VRAM, GFLOPS, DBI).
- **Выводы:** содержат итоговую рекомендацию конкретной архитектуры (EdgeNeXt-S) и описание вклада авторов.

Блок «Ошибки и логические несоответствия»

- **Смещение фокуса исследования:** в названии и введении заявлена задача «распознавания объектов» (Object Detection), которая обычно подразумевает определение координат объекта на фото. Однако в экспериментальной части работа полностью сводится к «классификации изображений» (Image Classification). Итоговые метрики детектирования (например, mAP) отсутствуют.
- **Подмена понятий в архитектуре:** в аннотации упоминается Faster R-CNN, но в результатах оцениваются только «базовые сети» (backbones) отдельно от самого детектора. Это создает логический разрыв: неясно, как именно выбор базовой сети влияет на итоговую скорость всей системы Faster R-CNN.
- **Методологический провал:** эксперимент по «разделяющей способности» проводится на простом наборе данных CIFAR-10, тогда как основное тестирование идёт на сложном ObjectDet. Прямая связь между результатами на этих двух разных датасетах в тексте не обоснована.

- **Отсутствие описательной части:** задача №2 требует «дать описание архитектурам», но в тексте оно заменено сухой сводной таблицей характеристик без анализа их внутреннего устройства.

Блок «Вывод»

Логическая связность: в целом высокая. Исследование последовательно: от постановки проблемы дефицита ресурсов до предложения конкретного программного решения.

Соответствие выводов задачам: задачи выполнены частично. Авторы успешно сформировали методику и выбрали лучшую сеть, однако задача по интеграции этих сетей в систему обнаружения объектов (Detection) осталась на теоретическом уровне, не подтверждённом профильными метриками в финале.

Общий итог: статья обладает научной новизной в части создания специализированного набора данных ObjectDet и методики интегрального оценивания. Однако для полного соответствия заявленной теме авторам следовало бы привести результаты работы всей системы распознавания в сборе, а не только её отдельных компонентов.

5 Промпт 5

5.1 Ответ модели Copilot

7. Тип исследования

Тип: экспериментальное исследование с элементами методологического анализа.

Основания для вывода:

- наличие экспериментальной части (обучение моделей, измерение VRAM, кластеризация, вычисление DBI);
- использование собственного датасета ObjectDet;
- проведение сравнительных вычислительных экспериментов;
- разработка методики интегральной оценки моделей.

Текст не является чисто теоретическим или обзорным — он содержит практические вычисления и анализ результатов.

8. Анализ структуры исследования

8.1. Наличие научной новизны

Новизна присутствует, но выражена неявно:

- создан собственный датасет ObjectDet;
- предложена интегральная оценка эффективности BN;
- проведено сравнение современных архитектур в условиях ограниченных ресурсов. Однако новизна не выделена отдельным пунктом, как принято в научных статьях.

8.2. Обоснованность утверждений

- утверждения опираются на экспериментальные данные (таблицы 3–6);
- есть ссылки на источники [1]–[22];
- методика описана достаточно подробно.

Обоснованность в целом достаточная, но интерпретация результатов краткая.

8.3. Корректность терминологии

Терминология используется корректно: backbone network, GFLOPs, DBI, dropout, batch, top-1/top-5 accuracy. Ошибок в терминах не выявлено.

8.4. Наличие ссылок

Ссылки присутствуют, но список литературы в фрагменте отсутствует. Это снижает полноту оформления.

8.5. Полнота аргументации

Аргументация достаточная для описания экспериментов, но:

- отсутствует глубокое обсуждение причин различий между моделями;
- не указаны ограничения исследования;
- не приведён анализ статистической значимости.

9. Оценка структуры по критериям научного дискурса

9.1. Логичность переходов

Переходы между разделами логичны: введение → постановка задачи → выбор архитектур → данные → методика → эксперименты → выводы.

9.2. Причинно-следственные связи

Есть:

- выбор BN обоснован критериями;
- методика связана с задачами;
- результаты связаны с методикой.

Но отсутствует причинно-следственный анализ полученных различий между моделями

9.3. Отсутствие противоречий

Противоречий почти нет, кроме одного:

- критерий DBI оценивается «по максимальному значению», хотя по определению DBI лучше минимальное значение.

Это методологическая ошибка.

9.4. Достаточность данных

Данные достаточны для сравнения моделей:

- метрики;
- VRAM;
- DBI;
- интегральная оценка.

Но отсутствуют:

- параметры оборудования;
- статистическая проверка;
- анализ ошибок моделей.

10. Анализ соответствия научному стилю

10.1. Точность и однозначность

Текст точный, формулировки однозначные.

10.2. Отсутствие эмоциональности

Эмоциональных выражений нет.

10.3. Корректность формулировок

Формулировки корректны, но встречаются мелкие стилистические шероховатости (например, «наилучшую эффективность показала. . . » без пояснения причин).

10.4. Отсутствие разговорных конструкций

Разговорных элементов нет.

11. Сильные стороны текста

- чёткая структура, соответствующая научному жанру;
- подробная методика экспериментов;
- наличие собственных данных (ObjectDet);
- использование современных архитектур и метрик;
- корректная терминология;
- наличие таблиц и формул;
- логичное соответствие цели и задач.

12. Слабые стороны текста

- научная новизна не выделена явно;
- недостаточное обсуждение результатов;
- ошибка в интерпретации критерия DBI;
- отсутствуют ограничения исследования;
- нет анализа статистической значимости;
- список литературы отсутствует в фрагменте;
- выводы обрываются и не завершены;
- недостаточно интерпретации причин различий между моделями.

13. Рекомендации по улучшению

1. Явно выделить научную новизну.
2. Добавить раздел «Обсуждение результатов» с интерпретацией различий между моделями.
3. Исправить критерий DBI (использовать минимальное значение).
4. Указать оборудование, на котором проводились эксперименты.
5. Добавить анализ ограничений исследования.
6. Завершить выводы и расширить их.
7. Включить список литературы.
8. Добавить графики (например, сравнение VRAM, точности, DBI).
9. Уточнить методику интегральной оценки (весовые коэффициенты, нормализация).

14. Итоговая оценка качества текста

7,5 / 10.

Статья структурирована, научно корректна и содержит полноценную экспериментальную часть, но требует доработки аналитического и методологического разделов для достижения высокого академического уровня.

5.2 Ответ модели Gemini

1. Тип исследования

Экспериментально-методологическое.

- **Признаки:** работа сочетает разработку новой методики интегрального оценивания нейросетей с проведением масштабного компьютерного эксперимента (обучение и тестирование 10 различных архитектур на специфическом наборе данных). Выводы базируются на численных измерениях (метрики точности, объём памяти, вычислительная сложность).

2. Анализ структуры по критериям научного дискурса

- **Логичность переходов:** текст демонстрирует строгую последовательность: от обоснования актуальности к формализации задачи, описанию инструментов, проведению тестов и финальному синтезу данных.

- **Причинно-следственные связи:** прослеживаются чётко. Ограничения аппаратных ресурсов диктуют критерии отбора моделей; результаты тестов по отдельным метрикам (DBI, Accuracy, VRAM) логично объединяются в интегральную оценку.
 - **Отсутствие противоречий:** внутренних логических конфликтов не выявлено. Методология, заявленная в начале, полностью реализуется в практической части.
 - **Достаточность данных:** объём выборки (10 архитектур, 172 класса в датасете ObjectDet) репрезентативен для задач данного типа. Использование различных под-ходов (UMAP для визуализации, DBI для кластеризации) повышает достоверность выводов.
-

3. Анализ стиля (соответствие научному стилю)

- **Точность и однозначность:** текст насыщен узкоспециализированной терминологией (backbone, convolution, transformer, GFLOPS), используемой корректно.
 - **Отсутствие эмоциональности:** повествование ведётся отстранённо, оценочные суждения носят исключительно сравнительно-математический характер (например, «наихудшие показатели», «преимущество архитектуры»).
 - **Корректность формулировок:** используются клише, характерные для ИТ-сектора («вычислительная сложность», «разделяющая способность»). Разговорные конструкции отсутствуют.
-

4. Сильные стороны текста

1. **Научная новизна:** предложена авторская методика интегрального оценивания, которая позволяет объективно сравнивать гетерогенные архитектуры (CNN, Transformer, Hybrid) по совокупности параметров.
 2. **Практическая значимость:** создан специализированный набор данных ObjectDet, ориентированный на задачи распознавания в условиях ограниченных ресурсов.
 3. **Глубина проработки:** авторы не ограничились стандартной точностью (Accuracy), а исследовали внутренние свойства признаков через индекс Дэвиса-Булдина и визуализацию латентного пространства.
 4. **Актуальность источников:** список литературы включает работы последних лет (2021–2023 гг.), что критично для быстро развивающейся области ИИ.
-

5. Слабые стороны

1. **Терминологический разрыв:** в заголовке заявлена «задача распознавания объектов» (Object Detection), однако исследование фактически сфокусировано на «извлечении признаков» (Feature Extraction) и «классификации». Отсутствуют данные о работе моделей в составе полноценного детектора (например, точность локализации объектов — mAP).
 2. **Скудность описания архитектур:** задача №2 («Дать краткое описание отобранным BN») выполнена формально — описания заменены одной сводной таблицей без качественного анализа различий между ConvNeXt, DaViT и другими.
-

6. Рекомендации по улучшению

1. **Расширение метрик:** рекомендуется протестировать выбранную архитектуру (EdgeNeXt-S) непосредственно в задаче детекции (например, в связке с SSD или Faster R-CNN) и привести метрики точности обнаружения границ объектов.
 2. **Аналитическая часть:** добавить краткий текстовый анализ: за счёт каких именно структурных особенностей (например, глубины свёрток или механизма внимания) EdgeNeXt-S обходит конкурентов в интегральном зачёте.
 3. **Обоснование датасетов:** пояснить, почему оценка разделяющей способности проводилась на CIFAR-10, если основное исследование идёт на ObjectDet.
-

Итоговая оценка

9 из 10. Статья представляет собой высококачественное прикладное исследование с чёткой методологией и обоснованными выводами. Один балл снижен за некоторое несоответствие широкого заголовка (распознавание объектов) более узкому предмету исследования (классификация признаков).

6 Сводная таблица оценок

Оценивание выполнено по четырём критериям (1–10 баллов):

- полнота структурных элементов;
- корректность и отсутствие выдуманной информации;
- критическая глубина анализа;
- чёткость формата ответа.

Промпт/Модель	Полнота	Корректность	Глубина	Чёткость	Средняя
1 Copilot	8	9	7	9	8.25
1 Gemini	8	9	7	9	8.25
2 Copilot	10	10	9	10	9.75
2 Gemini	10	10	8	10	9.5
3 Copilot	10	10	10	10	10
3 Gemini	10	10	9	10	9.75
4 Copilot	9	9	9	9	9
4 Gemini	8	9	8	9	8.5
5 Copilot	9	9	9	9	9
5 Gemini	9	9	8	9	8.75

Итог: обе модели демонстрируют высокую точность и структурность, но Copilot даёт более академичный и методологически строгий анализ, а Gemini — более гладкий и стилистически чистый.

7 Сравнительный анализ моделей

Сильные стороны Copilot

- Очень высокая структурная точность: чётко выделяет элементы (цель, задачи, методология).
- Глубокий критический анализ: выявляет методологические ошибки (например, неправильную интерпретацию DBI).
- Даёт подробные списки, структурированные блоки, логические цепочки.
- Хорошо соблюдает формат сложных промптов (особенно таблиц).

Слабые стороны Copilot

- Иногда даёт чрезмерно длинные объяснения.
- Местами повторяет формулировки.
- Чуть менее гладкий стиль по сравнению с Gemini.

Сильные стороны Gemini

- Очень чистый, академичный стиль.
- Отлично формулирует выводы и интерпретации.

- Хорошо работает с логикой текста и причинно-следственными связями.
- Даёт более компактные и читабельные ответы.

Слабые стороны Gemini

- Иногда менее критичен, чем Copilot.
- Местами сглаживает острые методологические ошибки.
- Чуть хуже выделяет скрытые структурные элементы.

Итог по моделям

Copilot — более строгий, критичный, методологичный. Gemini — более гладкий, академичный, стилистически чистый.

Обе модели демонстрируют высокий уровень, но Copilot лучше подходит для глубокого анализа структуры и методологии.

8 Сравнительный анализ промптов

Наиболее эффективный промпт

Промпт 3 — лучший.

Причины:

- чёткая таблица заставляет модели структурировать ответ;
- минимизирует выдумывание информации;
- позволяет сравнивать модели по одинаковым критериям;
- даёт наиболее объективные результаты.

Наименее эффективный промпт

Промпт 1 — самый слабый.

Причины:

- слишком общий;
- не задаёт критериев анализа;
- модели дают «поверхностные» ответы;
- сложно сравнивать результаты между моделями.

Промпт 5 как «авторский»

Промпт 5 (сам текст указан в начале отчёта вместе с остальными промптами) был применён к обеим моделям, потому что он является расширенным универсальным инструментом для глубокого экспертного анализа. Он объединяет требования всех предыдущих промптов и позволяет объективно сравнить модели в условиях максимальной сложности. Поэтому его корректно использовать не для одной модели, а для обеих.

Заключение

Проведённая оптимизация показала свою эффективность. Использование расширенного авторского промпта (Промпт 5) позволило получить от обеих моделей более глубокий, структурированный и критически насыщенный анализ по сравнению с результатами, полученными при использовании первых четырёх промптов. Несмотря на то, что формально Промпт 4 демонстрирует наиболее высокие оценки по строгим критериям, именно Промпт 5 обеспечивает максимальную аналитическую нагрузку и позволяет выявить сильные и слабые стороны моделей в условиях сложного экспертного задания.

Финальные ответы, полученные по авторскому промпту, получили высокие оценки по полноте, корректности, критичности и формату, что подтверждает успешность оптимизации. Модели продемонстрировали способность работать с расширенными требованиями, корректно выделять структурные элементы текста, проводить критический разбор и формировать содержательные выводы. Таким образом, оптимизация считается успешной: авторский промпт обеспечивает более качественный и глубокий анализ научного текста по сравнению с исходными промптами.