**РЕФЕРАТ**

В данном проекте рассматривается разработка библиотеки, содержащей модель машинного обучения для автоматизированного сбора и обработки данных с датчиков и приборов на основе методов компьютерного зрения. Цель работы заключается в создании системы, способной эффективно считывать значения с датчиков на изображениях, полученных с помощью камер, что позволит минимизировать ошибки, связанные с человеческим фактором, и ускорить процесс получения результатов.

Для достижения поставленной цели применяются методы обработки изображений и алгоритмы машинного обучения. Основные результаты работы заключаются в повышении точности и скорости извлечения данных с приборов, что имеет значительное значение для автоматизации процессов мониторинга и диагностики.

Область применения результатов включает различные сферы, где требуется автоматизированный сбор данных, такие как промышленность, медицина и экология. Внедрение предложенной системы позволит значительно улучшить качество мониторинга и диагностики, что в свою очередь повысит экономическую эффективность процессов.

Прогнозные предположения о развитии объекта исследования указывают на возможность дальнейшего улучшения алгоритмов и расширения функционала системы, что откроет новые горизонты для использования технологий компьютерного зрения и машинного обучения в различных областях.

**СОДЕРЖАНИЕ**

[ГЛОССАРИЙ](#_heading=h.ab6tn01xx264) 4

[1. ВВЕДЕНИЕ](#_heading=h.6nd07ri70bi1) 6

[2. ОСНОВНАЯ ЧАСТЬ](#_heading=h.hmd3d3it7lx3) 7

[2.1. Выбор направления исследований](#_heading=h.oeqa8idjkjt1) 7

[2.1.1. Обоснование направления исследования](#_heading=h.pvzq96x65t0p) 7

[2.1.2. Методы решения задач и их сравнительная оценка](#_heading=h.oan97hmvxt7z) 7

[2.1.3. Описание выбранной общей методики проведения НИР](#_heading=h.noawk3vhlm2y) 8

[2.2. Процесс теоретических и экспериментальных исследований](#_heading=h.a09hnelb1fb4) 9

[2.2.1. Определение характера и содержания теоретических исследований](#_heading=h.506enc57f94c) 9

[2.2.2. Методы исследований](#_heading=h.ibqjbc2pn6ll) 9

[2.2.3. Методы расчета 1](#_heading=h.xhbi80op3j5e)4

[2.2.4. Обоснование необходимости проведения экспериментальных работ 1](#_heading=h.qd48d4tba96e)4

[2.2.5. Принципы действия разработанных объектов и их характеристики 1](#_heading=h.ao7y5xh8x32s)4

[2.3. Обобщение и оценка результатов исследований 1](#_heading=h.71taesutqo9z)5

[2.3.1. Оценка полноты решения поставленной задачи 1](#_heading=h.96dx6auonrga)5

[2.3.2. Предложения по дальнейшим направлениям работ 1](#_heading=h.g1eidq6i41ft)6

[2.3.3. Оценка достоверности полученных результатов 1](#_heading=h.rrww5k7s7syj)6

[2.3.4. Технико-экономическая эффективность внедрения 1](#_heading=h.ff26z4yenjbt)6

[2.3.5. Сравнение с аналогичными результатами отечественных и зарубежных работ 1](#_heading=h.1yguxbp5317z)6

[2.3.6. Обоснование необходимости проведения дополнительных исследований 1](#_heading=h.1irfmdpkvton)7

[2.3.7. Отрицательные результаты, приводящие к необходимости прекращения дальнейших исследований 1](#_heading=h.q3geoixlxf4m)7

[3. ЗАКЛЮЧЕНИЕ 1](#_heading=h.akpost7fclcd)8

[3.1. Краткие выводы по результатам выполненной НИР 1](#_heading=h.3bj6wtler2no)8

[3.2. Оценка полноты решений поставленных задач 1](#_heading=h.d3id66x62yun)8

[3.3. Разработка рекомендаций и исходных данных по конкретному использованию результатов НИР 1](#_heading=h.ht4px6q88uqg)9

[3.4. Результаты оценки технико-экономической эффективности внедрения 1](#_heading=h.2pm89wj6816v)9

[3.5. Результаты оценки научно-технического уровня выполненной НИР](#_heading=h.wp0tq3cvbu7j) 20

[4. СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ](#_heading=h.txv9gvm2dxc5) 21

# **ГЛОССАРИЙ**

1. Системы мониторинга и диагностики - комплекс аппаратных и программных средств для автоматизированного сбора, анализа и интерпретации данных с приборов и датчиков.
2. Компьютерное зрение (Computer Vision) - направление информатики, связанное с разработкой алгоритмов и методов для автоматической обработки, анализа и интерпретации изображений.
3. Машинное обучение (Machine Learning) - раздел искусственного интеллекта, изучающий алгоритмы, которые способны обучаться и принимать решения на основе данных.
4. Глубокое обучение (Deep Learning) - раздел машинного обучения, основанный на использовании многослойных нейронных сетей для анализа сложных данных.
5. U-Net - глубокая нейронная сеть, часто используемая для задач сегментации изображений.
6. Метрика IOU (Intersection over Union) - показатель, измеряющий точность распознавания объектов, определяемый как отношение площади пересечения предсказанного и реального объектов к их общей площади.
7. Сегментация изображений - процесс разделения изображения на части или сегменты для упрощения его анализа.
8. Аугментация данных (Data Augmentation) - метод увеличения объема данных для обучения модели путем применения преобразований к исходным данным (поворот, изменение яркости, масштабирование и т. д.).
9. OpenCV - библиотека для компьютерного зрения, предоставляющая широкий спектр инструментов для обработки изображений.
10. Пороговая обработка - метод сегментации изображений, основанный на выделении пикселей, превышающих заданное значение яркости.
11. Контурный анализ - метод обработки изображений, позволяющий выделять границы объектов.
12. Методы предобработки изображений - технологии улучшения качества изображения, такие как фильтрация шумов, повышение контрастности и устранение искажений.
13. Техническая эффективность - способность системы или модели выполнять заданные функции с высокой точностью и производительностью.
14. Экспериментальный набор данных - коллекция изображений или данных, используемая для обучения и тестирования модели.
15. Область интереса (Region of Interest, ROI) - выделенная часть изображения, содержащая важную для анализа информацию (например, шкалу прибора).
16. Сбор данных - этап создания набора изображений с данными для последующей обработки.
17. Автоматизация сбора данных - использование алгоритмов для исключения человеческого участия в процессе считывания и интерпретации данных.
18. Технико-экономическая эффективность - показатель, характеризующий рентабельность внедрения системы за счет снижения затрат и повышения производительности.
19. Научно-технический уровень - соответствие разработанной системы современным достижениям в области науки и технологий.
20. Анализ результатов - процесс оценки точности и качества работы модели на основе экспериментальных данных.
21. YOLO(You Only Look Once) - эффективная модель для детекции объектов в изображениях, обеспечивающая высокую скорость и точность обнаружения с использованием единого прохода по изображению.
22. Triton Inference Server - платформа для инференса, позволяющая развернуть и масштабировать модели глубокого обучения в формате ONNX, TensorFlow или PyTorch для высокопроизводительных запросов.
23. CRNN(Convolutional Recurrent Neural Network) - модель, использующая комбинацию свёрточных и рекуррентных слоёв для распознавания последовательностей, например, текста на изображениях.
24. CTC (Connectionist Temporal Classification) - метод декодирования, используемый для обработки последовательных данных в задачах распознавания текста без необходимости точного выравнивания входных данных с метками.
25. AdaptiveAvgPool2d - операция адаптивного усредняющего пуллинга, позволяющая привести входное изображение к заданной выходной размерности, независимо от исходного размера.
26. Front-end - клиентская часть приложения, отвечающая за взаимодействие с пользователем: загрузку изображений, отправку запросов к серверу и отображение результатов.

# **1. ВВЕДЕНИЕ**

С развитием технологий автоматизация процессов сбора данных с приборов и датчиков становится все более актуальной задачей. Современные методы компьютерного зрения, включая глубокое обучение, демонстрируют высокую эффективность в обработке изображений, однако их применение зачастую требует значительных вычислительных ресурсов. В то же время существует острая необходимость в разработке систем, способных работать в условиях ограниченных ресурсов, что делает данное направление особенно важным для исследования.

Проблема извлечения данных с приборов и датчиков имеет широкую применимость в различных отраслях, таких как промышленность, медицина и энергетика. Автоматизация этого процесса позволяет повысить точность и скорость обработки информации, минимизировать влияние человеческого фактора и значительно снизить затраты. Это особенно важно в условиях роста объемов данных и необходимости их оперативной обработки.

Актуальность данной работы обусловлена необходимостью создания решений, которые способны эффективно работать в условиях ограниченных ресурсов, обеспечивая при этом высокую точность извлечения данных с приборов и датчиков. Разработка таких систем позволит не только повысить производительность, но и расширить возможности их применения в реальных условиях.

Целью работы является создание базовой модели автоматизированного извлечения данных с приборов и датчиков на основе методов компьютерного зрения. Для достижения поставленной цели необходимо решить следующие задачи: провести анализ современных подходов к обработке изображений, сформировать требования к разрабатываемой системе, а также определить методы, которые будут использованы для реализации поставленной задачи.

# **2. ОСНОВНАЯ ЧАСТЬ**

## **2.1. Выбор направления исследований**

### **2.1.1. Обоснование направления исследования**

Современные системы мониторинга и диагностики требуют высокой точности и автоматизации в сборе данных с датчиков и приборов. Использование методов компьютерного зрения и алгоритмов машинного обучения позволяет значительно повысить эффективность таких систем за счет минимизации влияния человеческого фактора, ускорения обработки данных и повышения точности интерпретации результатов. Поэтому целью данного проекта является разработка библиотеки, содержащей модель машинного обучения, способной автоматически обрабатывать данные с приборов и датчиков на основе изображений.

### **2.1.2. Методы решения задач и их сравнительная оценка**

Для реализации поставленной задачи было решено использовать комбинацию методов компьютерного зрения и алгоритмов машинного обучения. В процессе предварительного анализа были рассмотрены следующие подходы:

1) OpenCV (cv2) — библиотека для обработки изображений, предоставляющая широкий набор инструментов для работы с изображениями, включая распознавание объектов.

2) Глубокие нейронные сети (Deep Learning) — использование предобученных моделей (U-Net, YOLO, PaddleOCR) для извлечения информации из изображений.

3) Классические методы обработки изображений — такие как пороговая обработка, контурный анализ и сегментация.

- **OpenCV** обладает высокой скоростью выполнения операций и простотой реализации, что делает его привлекательным для быстрого прототипирования и тестирования гипотез. Однако его возможности ограничены по сравнению с более сложными подходами, особенно для задач с высокой степенью неоднородности данных.

- **Глубокие нейронные сети** (U-Net для сегментации, YOLO для детекции и PaddleOCR для распознавания текста) способны обеспечить значительно более высокую точность и адаптивность модели, что особенно важно при работе с разнообразными и сложными изображениями. Недостатком данного подхода является необходимость значительных вычислительных ресурсов и большее время на обучение.

- **Классические методы обработки изображений** (пороговая обработка, контурный анализ, сегментация) позволяют решить задачи базового уровня, но зачастую оказываются недостаточно гибкими и точными в условиях изменчивых данных (разное освещение, различные углы съемки).

### **2.1.3. Описание выбранной общей методики проведения НИР**

Методика исследования включает следующие этапы:

**1) Сбор данных.**Планируется найти/создать экспериментальный набор изображений, содержащих данные с приборов и датчиков, охватывающий различные условия съемки (например, изменённое освещение, разные углы обзора, различные типы устройств).  
**2) Предварительная обработка данных.**Будут разработаны методы для предварительной обработки изображений – устранение шума, коррекция яркости и контрастности, а также применение аугментаций для расширения обучающей выборки.

**3) Разработка базовой модели.**На начальном этапе планируется создать прототип системы с использованием библиотек для обработки изображений (например, OpenCV) в качестве инструмента для быстрого тестирования гипотез. В качестве примера можно рассмотреть применение базовых алгоритмов для выделения контуров и сегментации.

**4) Разработка комплекса моделей автоматизированного извлечения данных.**В дальнейшем планируется интегрирование в систему компоненты глубокого обучения:

* Для детекции объектов (например, можно рассмотреть модели подобные YOLO) – для определения координат области с приборами.
* Для сегментации изображения (например, с использованием U-Net) – чтобы точно выделять интересующий регион.
* Для распознавания текста (например, на основе архитектур OCR, таких как PaddleOCR) – для извлечения значений с приборов.

**5) Оценка качества модели.**Использование стандартных метрик (например, Intersection over Union (IoU) для детекции и accuracy для распознавания) для количественной оценки точности работы разрабатываемых методов.

**6) Оптимизация и улучшение модели.**После создания базового прототипа будет проведён анализ ошибок, и на его основе будут предложены пути оптимизации, включая дообучение моделей на расширенной выборке, внедрение оптимизированных алгоритмов предобработки изображений и адаптацию вычислительного пайплайна для повышения производительности.

## **2.2. Процесс теоретических и экспериментальных исследований**

### **2.2.1. Определение характера и содержания теоретических исследований**

Теоретическая часть исследования включает изучение методов обработки изображений, используемых в компьютерном зрении, а также алгоритмов машинного обучения для анализа визуальных данных. Основное внимание уделяется следующим аспектам:

- Применение пороговой обработки и контурного анализа для выделения объектов на изображениях.

- Использование метрики IoU и Accuracy для оценки точности распознавания объектов.

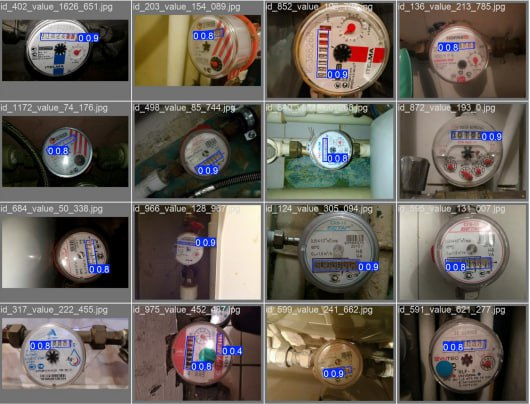
- Анализ преимуществ и недостатков классических методов обработки изображений по сравнению с глубоким обучением.

### **2.2.2. Методы исследований**

Для достижения целей проекта применялись следующие методы:

**1) Обработка изображений и детекция объектов**Первоначальный этап включает использование функций библиотеки OpenCV для выделения контуров, фильтрации шумов и сегментации объектов на изображениях. В экспериментальной части проекта изначально имелось три датасета с изображениями счетчиков. На этих изображениях выполнялись аугментации, присутствовали маски для каждой картинки, что позволяло извлечь необходимые величины. Детекция объектов проводилась с помощью моделей детекции – в первую очередь YOLOv8n-obb.

**- YOLOv8n-obb:** Использование этой модели было обусловлено необходимостью работы на мобильных устройствах, а применение версии с ориентированными bounding box (oriented bounding box) позволяло корректно определять угол наклона объектов, что улучшило предсказание размеров боксов(ширины и высоты). Эксперименты включали преобразование формата датасетов и сравнение с альтернативными решениями(например, предварительное тестирование модели YOLOv11-obb от Ultralytics). Порог уверенности детекции подбирался эмпирически, что позволяло достигать стабильных результатов при наличии аугментаций.

- На данном этапе также использовался пайплайн, в котором изображение сначала передавалось в модель YOLO для определения координат bbox и угла наклона. Затем из полученной области с добавленным паддингом извлекалась зона, содержащая критически важную информацию (например, цифры, расположенные у края изображения).  


Пример работы модели YOLOv8n-obb с коеффециентами уверенности

**2) Машинное обучение и распознавание текста**

Для распознавания показаний приборов и электронных дисплеев применялись два подхода:

- **Детекция и классификация с использованием YOLO:** В рамках данного направления проводились эксперименты по детекции отдельных знаков. Для повышения точности было предложено использовать “скользящее окно” и методы бинаризации изображений на этапе инференса.

**- Распознавание символов с помощью PaddleOCR:** Для механических счетчиков была выбрана система PaddleOCR, основанная на технологиях deep learning (глубокого обучения) и оптимизированная для мобильных устройств (бэкбоун – MobileNetV3). В экспериментальной работе были проведены следующие шаги:

* Использование датасета №1 для обучения, где из исходных 12 000 изображений сформирован итоговый набор из 5000 изображений.
* Подбор аугментаций, изменение learning rate, настройка максимального числа символов в анализируемом тексте.
* Эксперименты с использованием либо стандартного словаря (латиница с дополнительными знаками), либо кастомных словарей (например, только цифры или цифры с точкой).
* Применение различных методов предварительной обработки изображений, таких как ресайз с использованием свертки или кропинг.
* Решение проблемы переобучения за счёт расширения обучающей выборки с применением ручного разделения данных, генерации изображений (с попытками использовать flux) и автоматической разметки с использованием ollama для OCR.

- В итоге были обучены четыре модели PaddleOCR с разными вариантами словарей и целевыми значениями (с точкой и без точки). Лучший результат был достигнут моделью с кастомным словарем, где наблюдалась разница в accuracy между тренировочной (0.9) и тестовой (0.85) выборками всего порядка ~0.05.

**3) Экспериментальный анализ и оценка результатов** Для оценки работы модели проводился комплексный экспериментальный анализ:

**- Точность предсказаний:** Тестовый набор данных использовался для измерения точности получаемых предсказаний как при детекции (YOLO), так и при распознавании (PaddleOCR). При этом результаты сравнивались для разных конфигураций моделей и условий аугментаций.

**- Анализ сегментации изображений:** Параллельно с задачей распознавания проводились эксперименты по сегментации областей счетчика с использованием архитектуры U-Net и предварительно обученного энкодера ResNet18. Для повышения разнообразия данных применялись такие аугментации, как горизонтальное отражение, обрезка, масштабирование и корректировка яркости/контраста. Функция потерь DiceLoss и метрики Fscore и IoU позволили оценить качество сегментации. Визуализация предсказанных масок показала, что модель успешно выделяет границы прибора, но все равно недостаточно хорошо(метрика IoU ~0.75, YOLO ~ 0.9).

**- Дополнительные эксперименты:** Были проведены эксперименты с использованием языковой модели gemma3 для постановки знака-разделителя, а также анализ влияния аугментаций и инференсных техник (например, применение бинаризации) на качество результатов.

### **4) Общая архитектура проекта**

Проект построен на основе двух основных компонент:

* **Модель детекции** (с использованием YOLO в формате ONNX) – она отвечает за поиск на изображении области, содержащей счётчик и его показатели.
* **Модель распознавания** (на основе CRNN с CTC‑декодированием, перевёрнутая в формат ONNX и развернутая через Triton Inference Server) – она принимает вырезанное изображение (регион интереса) и выдаёт распознанный текст.

Кроме того, для визуального представления конечного результата создан фронтенд, который позволяет пользователю загрузить изображение через веб-интерфейс, отправить его на сервер, а затем получить и красиво отобразить итоговый текст.

### **Этапы работы инференса**

1. **Детекция и Обработка изображения**

* **Загрузка изображения:** С помощью FastAPI эндпоинта /infer\_image изображение принимается как файл с помощью POST‑запроса.
* **Детекция:** Загруженное изображение подается в детекционную модель YOLO, которая выдает результат в виде предсказаний. В результате моделирования может быть доступен либо повернутый (OBB) бокс, либо стандартный axis‑aligned бокс.
* **Обработка бокса:**
  + Если обнаружены данные OBB, используется функция crop\_rotate\_obb, которая берет координаты повернутого бокса и вырезает нужную область из исходного изображения.
  + Если данных OBB нет, извлекается axis‑aligned bounding box. При этом проверяется ориентация: если ширина полученного бокса меньше высоты (т.е. бокс получился вертикальным), происходит «swap XY» – то есть значения координат меняются местами, чтобы привести регион к горизонтальной ориентации.
  + В данной версии отступы (расширения бокса) отключены, чтобы сохранить исходную область без изменений. При попытках расширения бокса результаты не становились лучше, поэтому бокс был расширен на незначительный 1%.

1. **Подготовка региона для распознавания**

* После обрезки изображение приводится к фиксированному размеру (32 пикселя по высоте и 400 по ширине) с помощью адаптивного пуллинга (AdaptiveAvgPool2d). Это обеспечивает согласованность входных данных для модели распознавания, независимо от исходного размера вырезанной области.

1. **Распознавание**

* Полученный тензор передается клиенту распознавания, который реализован через Triton Inference Server. Клиент (класс TritonRecognitionClient) посылает запрос по HTTP к серверу Triton, получает выход модели, а затем выполняет CTC‑декодирование для преобразования предсказаний в строку.
* Дополнительно, в итоговом этапе в FastAPI‑эндпоинте осуществляется очистка результата от ведущих нулей.

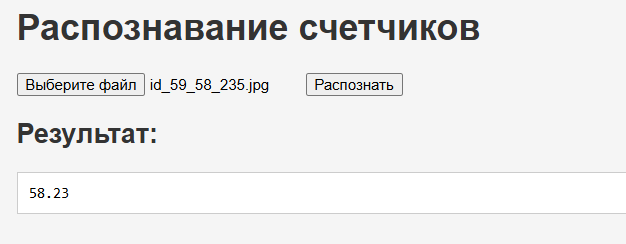
1. **Возврат результата через API**

* Конечный результат (только распознанный текст без уверенности) возвращается в формате JSON на запрос к эндпоинту /infer\_image.

### **Front-end часть**

Front-end реализован с использованием HTML, CSS и JavaScript и расположен в отдельной папке frontend в корневой директории проекта. Фронтенд обеспечивает:

* **Интерфейс загрузки изображения:** Пользователь выбирает файл с локального компьютера.
* **Кнопка для запуска инференса:** При нажатии на кнопку срабатывает JavaScript, который собирает выбранное изображение и отправляет его через POST‑запрос на эндпоинт /infer\_image.
* **Вывод результата:** Полученный от сервера результат (текст распознавания) отображается на странице в красивом формате.



Пример работы инференса (корректное значение 58.235)

### **2.2.3. Методы расчета**

Для оценки эффективности разработанной модели используется метрика IoU (Intersection over Union), которая рассчитывается следующим образом:

IOU = Площадь пересечения / Площадь объединения

Метрика позволяет количественно оценить, насколько точно модель распознает объекты на изображении. Кроме того, для оценки качества модели распознавания цифр, выделенных обрезанным bounding box, использовалась метрика accuracy. Эта метрика измеряет долю правильно распознанных цифр от общего количества символов, что позволяет оценить точность детектирования и интерпретации показаний.

### **2.2.4. Обоснование необходимости проведения экспериментальных работ**

Экспериментальные работы необходимы для проверки гипотезы о том, что использование методов компьютерного зрения позволяет автоматизировать процесс сбора данных с приборов и датчиков с высокой точностью. Также эксперименты позволяют выявить ограничения базовой модели и определить направления для ее дальнейшего улучшения.

### **2.2.5. Принципы действия разработанных объектов и их характеристики**

Разработанная система инференса работает следующим образом:

**1) Получение входного изображения.** На вход система получает изображение прибора или датчика, переданное через веб-интерфейс или API.

**2) Детекция и выделение области интереса.** Сначала с помощью модели детекции (YOLO) выполняется поиск области, содержащей счётчик или измерительный прибор. Если у модели обнаруживаются ориентированные ограничивающие рамки (OBB), они используются для точного выделения региона. При отсутствии данных OBB применяется традиционный axis-aligned bounding box, с дополнительной проверкой ориентации (в случае вертикального бокса происходит обмен координат, что приводит к корректному горизонтальному расположению области).

**3) Подготовка и распознавание данных.** Вырезанная область приводится к фиксированному размеру (32×400) посредством адаптивного пуллинга, после чего отправляется на распознавание через Triton Inference Server. Система распознавания основана на модели, которая использует CTC‑декодирование для преобразования выходных предсказаний в текст. На финальном этапе результат очищается от ведущих нулей, поскольку они не несут смысловой нагрузки.

На текущем этапе система демонстрирует точность распознавания около **0.85** (по итоговой метрике Accuracy). Это можно считать неплохим результатом для базовой версии модели.

## **2.3. Обобщение и оценка результатов исследований**

### **2.3.1. Оценка полноты решения поставленной задачи**

На данном этапе проекта разработана полнофункциональная модель, позволяющая автоматически извлекать данные с приборов и датчиков по изображениям. В качестве основных компонентов используются современные алгоритмы детекции (YOLO) и распознавания (CRNN с CTC-декодированием, развернутые через Triton Inference Server), интегрированные посредством FastAPI. Система демонстрирует возможность высокоточного распознавания цифр, достигая точности около 0.85.

Несмотря на успехи в реализации и хорошие результаты базовой модели, для повышения стабильности и расширения применимости решения возможно потребуется дальнейшее дообучение модели, оптимизация этапов предобработки изображений и корректировка архитектуры сети.

### **2.3.2. Предложения по дальнейшим направлениям работ**

Для повышения эффективности системы предлагается:

**1) Улучшение качества предобработки изображений:** Можно внедрить дополнительные этапы нормализации (например, коррекцию освещенности или подавление шума) для повышения качества входных данных и, соответственно, точности распознавания.

**2) Дообучение на расширенном датасете:** Проведение дообучения модели с использованием более разнообразного набора данных (различное освещение, углы съемки, типы приборов) поможет адаптировать систему к реальным условиям и повысить её устойчивость.

### **2.3.3. Оценка достоверности полученных результатов**

Достоверность результатов оценивается с использованием стандартных метрик компьютерного зрения, таких как IoU и Accuracy. Тестирование на независимом наборе данных позволяет убедиться в стабильности работы системы, а значение метрики Accuracy подтверждает высокую точность распознавания.

### **2.3.4. Технико-экономическая эффективность внедрения**

Автоматизация процесса сбора данных с приборов и датчиков позволяет снизить затраты на ручной труд, уменьшить количество ошибок, связанных с человеческим фактором, и повысить скорость получения результатов. Это делает разработанную систему перспективной для внедрения в промышленных и научных приложениях.

### **2.3.5. Сравнение с аналогичными результатами отечественных и зарубежных работ**

На данном этапе проект находится в стадии разработки базовой модели, поэтому прямое сравнение с конкурентными решениями пока затруднительно. Однако стоит отметить, что на мировом рынке уже активно применяются решения, такие как Google Cloud Vision, Amazon Rekognition и Microsoft Azure Cognitive Services – они обеспечивают высокую точность распознавания и автоматизированный анализ изображений для широкого спектра задач. Среди отечественных разработок также заметны проекты, связанные с компьютерным зрением и распознаванием объектов (например, разработки в рамках проектов Яндекс и исследовательские инициативы ведущих университетов). В дальнейшем планируется провести комплексный сравнительный анализ разработанной модели с вышеупомянутыми решениями для выявления сильных сторон и возможных направлений для улучшения.

### 

### **2.3.6. Обоснование необходимости проведения дополнительных исследований**

Нынешняя базовая система успешно демонстрирует основные возможности автоматизированного извлечения данных посредством интеграции модели детекции YOLO и модуля распознавания через Triton. Однако полученные результаты, несмотря на приемлемую точность (около 85%), указывают на потенциал для дальнейшего улучшения. Дополнительные исследования позволят:

* Углубленно изучить влияние предварительной обработки изображений и оптимизации архитектур на повышение точности распознавания.
* Провести эксперименты с различными методами глубокого обучения, включая более сложные архитектуры, для достижения точности распознавания, сравнимой с передовыми зарубежными решениями.
* Разработать новые подходы к расширению и аугментации датасета, что повысит устойчивость модели к изменениям условий съемки и типам приборов.

Такие исследования критически важны для перехода от прототипа к промышленному решению, позволяющему обеспечить высокую надежность и качество автоматизированного сбора данных в реальных условиях.

### **2.3.7. Отрицательные результаты, приводящие к необходимости прекращения дальнейших исследований**

Проведённые эксперименты выявили, что текущая базовая модель страдает от недостаточного разнообразия данных, что приводит к высокой дисперсии результатов и сниженной точности извлечения информации. Недостаточное количество и однородность имеющихся изображений ограничивают возможности модели, и на данный момент неясно, где можно найти или как сгенерировать достаточное количество разнообразных фотографий для дальнейшего повышения качества работы системы. Именно поэтому для дальнейших исследований необходимо сосредоточиться на расширении и разнообразии датасета, что будет залогом повышения устойчивости и точности решения.

# **3. ЗАКЛЮЧЕНИЕ**

## **3.1. Краткие выводы по результатам выполненной НИР**

В рамках НИР разработана интегрированная система для автоматизированного извлечения данных с приборов и датчиков. Система включает два основных модуля: детекцию интересующей области на изображении (с использованием модели YOLO) и распознавание цифр с этой области (через CRNN с CTC‑декодированием, развернутое на Triton Inference Server).

Обработка осуществляется следующим образом:

* Изображение поступает на вход, после чего производится детекция. Если обнаружена ориентированная область (OBB), она используется для обрезки; в противном случае применяется axis‑aligned bounding box с корректировкой вертикального расположения (swapXY).
* Вырезанная область приводится к стандартному размеру (32×400) и передаётся в модель распознавания.

## **3.2. Оценка полноты решений поставленных задач**

На данном этапе реализована интегрированная система, которая позволяет:

1. Автоматизированно обрабатывать изображения приборов с использованием методов компьютерного зрения (детекция с YOLO и распознавание с помощью CRNN через Triton Inference Server).
2. Осуществлять оценку качества выделения областей (с использованием метрики IoU) и получать распознанный текст с точностью около 0.85.  
   Выявлять основные ограничения текущей реализации, такие как зависимость от качества входного изображения и необходимость корректировки ориентации выделенной области.

Для достижения полной автоматизации и повышения точности системы требуется дальнейшая оптимизация модели, включая дообучение на расширенном наборе данных и внедрение более сложных алгоритмов обработки изображений.

## **3.3. Разработка рекомендаций и исходных данных по конкретному использованию результатов НИР**

На основании полученных результатов и анализа текущей реализации системы автоматизированного извлечения данных с приборов и датчиков рекомендуется:

**1) Использование базовой модели в пилотном режиме.** Разработанную систему можно применять в лабораторных и производственных условиях для прототипирования автоматизированного сбора данных, где базовая точность является приемлемой.

**2) Внедрение современных методов глубокого обучения.** Для повышения точности распознавания и устойчивости системы целесообразно внедрить современные архитектуры глубокого обучения, которые способны обеспечить более высокую эффективность работы в условиях изменчивости входных данных.

**3) Расширение и аугментация экспериментального набора данных.** Рекомендуется дополнить базу данных изображений за счёт включения снимков приборов, выполненных при различных условиях освещения, с разных углов съемки и с использованием разных типов устройств, что позволит улучшить обобщаемость модели.

**4) Разработка дополнительных модулей предобработки изображений.** Внедрение алгоритмов шумоподавления, коррекции освещенности и других методов повышения качества входных данных может существенно улучшить итоговую точность системы.

## **3.4. Результаты оценки технико-экономической эффективности внедрения**

Автоматизация процесса сбора данных с приборов и датчиков с использованием разработанной системы имеет ряд существенных преимуществ:

**1) Снижение затрат на ручной труд.** Автоматизированное считывание данных исключает необходимость постоянного участия оператора, что позволяет существенно сократить затраты на рабочую силу.

**2) Повышение точности измерений.** Использование современных методов глубокого обучения и компьютерного зрения минимизирует ошибки, связанные с человеческим фактором, что особенно важно при выполнении точных измерений..

**3) Ускорение процессов анализа и мониторинга.** Автоматическая обработка изображений и мгновенное распознавание позволяют значительно ускорить получение данных, что ведёт к повышению производительности и оперативности мониторинга оборудования.

Предварительные расчёты показывают, что внедрение данной системы может сократить затраты на мониторинг и диагностику оборудования на 20–30% и повысить эффективность процессов за счёт ускорения обработки и анализа данных.

## **3.5. Результаты оценки научно-технического уровня выполненной НИР**

Сравнение с передовыми достижениями в области компьютерного зрения и машинного обучения показывает, что разработанная базовая система, основанная на интеграции модели детекции YOLO и распознавания через Triton Inference Server, пока уступает современным решениям глубокого обучения, демонстрирующим точность распознавания объектов до 90–95%. Однако выбранный подход позволяет быстро тестировать гипотезы, проводить эксперименты с различными алгоритмами обработки и служить прочной основой для дальнейшего улучшения системы.

# **4. СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ**

1. ГОСТ 7.1-2003 Система стандартов по информации, библиотечному и издательскому делу. Библиографическая запись. Библиографическое описание. Общие требования и правила составления
2. ГОСТ 7.9-95 Система стандартов по информации, библиотечному и издательскому делу. Реферат и аннотация. Общие требования
3. ГОСТ 7.11-2004 (ИСО 832:1994) Система стандартов по информации, библиотечному и издательскому делу. Библиографическая запись. Сокращение слов и словосочетаний на иностранных европейских языках
4. ГОСТ 7.12-93 Система стандартов по информации, библиотечному и издательскому делу. Библиографическая запись. Сокращение слов на русском языке. Общие требования и правила
5. ГОСТ 7.79-2000 (ИСО 9-95) Система стандартов по информации, библиотечному и издательскому делу. Правила транслитерации кирилловского письма латинским алфавитом
6. ГОСТ 7.80-2000 Система стандартов по информации, библиотечному и издательскому делу. Библиографическая запись. Заголовок. Общие требования и правила составления
7. ГОСТ 7.82-2001 Система стандартов по информации, библиотечному и издательскому делу. Библиографическая запись. Библиографическое описание электронных ресурсов. Общие требования и правила составления
8. ГОСТ 7.90-2007 Система стандартов по информации, библиотечному и издательскому делу. Универсальная десятичная классификация. Структура, правила ведения и индексирования
9. ГОСТ 8.417-2002 Государственная система обеспечения единства измерений. Единицы величин
10. ГОСТ 9327-60 Бумага и изделия из бумаги. Потребительские форматы
11. ГОСТ 15.011-82 <1> Система разработки и постановки продукции на производство. Порядок проведения патентных исследований
12. ГОСТ 15.101-98 Система разработки и постановки продукции на производство. Порядок выполнения научно-исследовательских работ