

Федеральное государственное бюджетное образовательное  
учреждение высшего образования  
«Российский химико-технологический университет имени Д.И. Менделеева»

*На правах рукописи*



Корчагин Валерий Дмитриевич

МОДЕЛЬ И АЛГОРИТМЫ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ИЗОБРАЖЕНИЙ  
В DISOM-ПОДОБНЫХ ДАННЫХ ПРИ ПОДДЕРЖКЕ ПРИНЯТИЯ РЕШЕНИЙ

Специальность 2.3.1. – Системный анализ, управление и обработка информации,  
статистика

ДИССЕРТАЦИЯ

на соискание ученой степени

кандидата технических наук

Научный руководитель:  
доктор технических наук, профессор  
Ковшов Евгений Евгеньевич

Москва 2026

## Оглавление

ВВЕДЕНИЕ.....	4
Глава 1. АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ И ПОСТАНОВКА ЗАДАЧИ ИССЛЕДОВАНИЯ.....	15
1.1 Стандарты DICOM, DICONDE, DICOS .....	15
1.2 Алгоритмы поиска изображений и их ограничения .....	20
1.3 Методы комплексирования при поиске изображений .....	30
1.4 Вывод по главе.....	32
Глава 2. ФОРМАЛИЗАЦИЯ ЗАДАЧИ КОМПЛЕКСИРОВАННОГО ПОИСКА СХОЖИХ ИЗОБРАЖЕНИЙ В DICOM-ПОДОБНЫХ ФОРМАТАХ.....	33
2.1 Формализация DICOM-подобных данных и пространства признаков .....	33
2.2 Токенизация и стандартизация метаданных .....	38
2.3 Формирование визуальных признаков и архитектура «DICONDE-YOLO».....	41
2.4 Выводы по главе.....	52
Глава 3. ВЫЧИСЛИТЕЛЬНЫЕ ЭКСПЕРИМЕНТЫ. ВЕРИФИКАЦИЯ МОДЕЛЕЙ И АЛГОРИТМОВ.....	53
3.1 Описание аппаратной платформы и воспроизводимости вычислительных экспериментов .....	54
3.2 Экспериментальные данные и методика подготовки датасета .....	55
3.3 Сравнительный анализ архитектур YOLOv8-YOLOv12 .....	59
3.4 Выводы по главе.....	72
Глава 4. ПРАКТИЧЕСКАЯ РЕАЛИЗАЦИЯ ВСТРАИВАЕМОСТИ В ЕДИНОЕ ИНФОРМАЦИОННОЕ ПРОСТРАНСТВО. СРАВНЕНИЕ С ПРОТОТИПОМ СППР И АНАЛИЗ ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ .....	73
4.1 Требования к интеграции алгоритмов интеллектуального анализа изображений в единое информационное пространство СППР.....	74
4.2 Архитектура программной реализации и место разработанного алгоритма в системе .....	76
4.3 Реализация прототипа и описание базового решения для сравнительного анализа.....	79
4.4 Анализ результатов опытной эксплуатации и сравнительная оценка эффективности разработанной модели и алгоритмов в составе СППР .....	82

ЗАКЛЮЧЕНИЕ .....	94
СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ.....	97
СЛОВАРЬ ТЕРМИНОВ.....	99
СПИСОК ЛИТЕРАТУРЫ.....	103
ПРИЛОЖЕНИЕ А. ПСЕВДОКОД КОМПЛЕКСНОГО АЛГОРИТМА ПОИСКА ИЗОБРАЖЕНИЙ И СТРУКТУРНЫХ СОСТАВЛЯЮЩИХ.....	116
ПРИЛОЖЕНИЕ Б. СВИДЕТЕЛЬСТВО О РЕГИСТРАЦИИ ПРОГРАММЫ ДЛЯ ЭВМ «РАЗМЕТКА DICONDE-ИЗОБРАЖЕНИЙ» .....	121
ПРИЛОЖЕНИЕ В. ПРИМЕРЫ ДЕТЕКЦИИ ДЕФЕКТОВ С ПРИМЕНЕНИЕМ YOLOV8 И «DICONDE-YOLO» .....	122
ПРИЛОЖЕНИЕ Г. ИНТЕРФЕЙС РАЗРАБОТАННОГО ПРОТОТИПА СППР .....	123
ПРИЛОЖЕНИЕ Д. АКТ О ВНЕДРЕНИИ В АО «НИКИМТ-АТОМСТРОЙ».....	127
ПРИЛОЖЕНИЕ Е. СПРАВКА ОБ ИСПОЛЬЗОВАНИИ В РХТУ ИМ. Д.И. МЕНДЕЛЕЕВА.....	128
ПРИЛОЖЕНИЕ Ж. АНАЛИЗ ПОТРЕБНОСТИ СМЕЖНЫХ ПРОМЫШЛЕННЫХ ОТРАСЛЕЙ В РЕАЛИЗАЦИИ ФУНКЦИОНАЛЬНЫХ ВОЗМОЖНОСТЕЙ СППР.....	129

## ВВЕДЕНИЕ

*Актуальность работы.* В условиях интенсивного развития технологий искусственного интеллекта и компьютерного зрения наблюдается устойчивая тенденция к интеграции передовых решений производственные, диагностические и инспекционные процессы в рамках концепции Индустрии 4.0 [1-3]. Цифровая трансформация прикладных систем анализа визуальной информации ориентирована на повышение эффективности, качества и безопасности принимаемых решений за счет внедрения интеллектуальных алгоритмов обработки данных. Существенный рост вычислительных возможностей и объемов накапливаемой визуальной информации формирует предпосылки для широкого применения интеллектуальных методов анализа изображений в системах поддержки принятия решений.

В современных информационных системах значительный объем данных представлен изображениями, сопровождаемыми структурированными метаданными, описывающими параметры оборудования, условия получения и характеристики исследуемых объектов. Для стандартизированного хранения и передачи подобных данных применяются форматы семейства DICOM, а также их специализированные расширения — DICONDE и DICOS. Указанные стандарты обеспечивают унифицированное представление визуальной и фактографической информации и позволяют рассматривать DICOM-подобные форматы как универсальные носители гетерогенных данных вне зависимости от предметной области.

Рост объемов информации, представленной в DICOM-подобных форматах, актуализирует задачу автоматизированного поиска схожих изображений, направленного на выявление аналогичных объектов, структурных элементов или аномалий. Применение интеллектуальных моделей и алгоритмов позволяет снизить нагрузку на эксперта, сократить время анализа и повысить объективность и воспроизводимость результатов. Наиболее активно данные технологии используются

в медицинской диагностике, где интеллектуальные системы выполняют функции вспомогательного инструмента для поддержки принятия решений для врача-специалиста. Успешный опыт их применения [7-12] подтверждает целесообразность адаптации и развития аналогичных подходов для смежных прикладных областей, включая НК, в которых также используются стандартизированные DICOM-подобные форматы хранения изображений.

Вместе с тем практическая реализация моделей и алгоритмов поиска схожих изображений в условиях реальных данных сталкивается с рядом фундаментальных проблем. Визуальная информация характеризуется вариативностью параметров ее получения, наличием шумов и артефактов, различием модальностей и многоканальной структурой данных. Дополнительную сложность представляет необходимость корректной интерпретации пиксельных значений, физический смысл которых в форматах DICOM, DICONDE и DICOS определяется в совокупности с фактографическими данными. Использование алгоритмов, основанных исключительно на работе с визуальными признаками или на отдельных классах методов, как правило, не обеспечивает требуемой устойчивости и универсальности результатов поиска в условиях изменяющихся параметров контроля, диагностики и неполноты данных.

В этой связи особую актуальность приобретает задача разработки моделей и алгоритмов комплексированного поиска схожих изображений, предполагающий совместное использование визуальных и фактографических данных в рамках единого формализованного подхода. Комплексирование позволяет компенсировать ограничения отдельных методов, повысить устойчивость и релевантность результатов поиска, а также обеспечить масштабируемость и приемлемое время отклика системы при существенном росте объемов данных. Реализация таких подходов является особенно значимой для прикладных СППР, функционирующих в условиях реального промышленного производства.

Таким образом, актуальность настоящей диссертационной работы обусловлена необходимостью повышения эффективности выполнения поисковых запросов в прикладных СППР на основе разработки и комплексного применения моделей и алгоритмов интеллектуального анализа изображений в визуальных и фактографических датасетах, представленных в форматах семейства DICOM и их специализированных расширениях. Решение поставленной задачи имеет существенное практическое значение для широкого круга прикладных отраслей, использующих стандартизированные форматы хранения изображений, и направлено на формирование инвариантных и масштабируемых подходов к интеллектуальному анализу визуальной информации.

*Степень разработанности проблемы.* Теоретические и прикладные аспекты анализа визуальных данных рассматриваются в работах отечественных и зарубежных авторов, включая исследования, посвященные применению классических алгоритмов обработки изображений и современных архитектур искусственных нейронных сетей (ИНС) (J. Redmon, R.V. Girshick, S. Nagraj, А.В. Пролетарский, Л.Г. Гагарина, Дорогов В.Г., Меньшутин Н. В., Кольцова Э. М., Глебов М. Б., Демидова Л.А., Пылькин А.Н., Белов В.В. и другие). Вместе с тем вопросы реализации алгоритмов поиска изображений в составе интегрированных программно-алгоритмических комплексов, ориентированных на поддержку принятия решений в условиях промышленной эксплуатации, остаются недостаточно проработанными.

Вопросам разработки и применения СППР для промышленной отрасли активно посвящены работы таких исследователей, как Ильясов Б.Г., Куликов Г.Г., Буракова Е.А., Конев К.А., Родионов Д.Г., Денисов Д.Ю., Демидова Л.А., Пруцков А.В. и других. В их работах подчеркивается важность интеграции имитационного моделирования и нейросетевых технологий для создания адаптивных и высокоточных систем, учитывающих специфику производственных процессов, включая особенности НК сварных соединений металлических конструкций.

Несмотря на значительный объем исследований в области СППР и современных методов и алгоритмов обработки изображений, проводимых как отдельными исследовательскими группами, так и ведущими научно-исследовательскими институтами и университетами Российской Федерации, такими, как МАИ, МГТУ им. Баумана, ИТМО, МИФИ, МФТИ, МИРЭА, НИИ «Фотон» (РГРТУ), ФГУП «ГосНИИАС», а также – компаниями-разработчиками программного обеспечения ООО «ИнТехКонтроль», ООО «Микромаш», ООО «К-СИСТЕМС ГРУПП», большинство из актуальных решений характеризуются узкой специализацией или ограниченной применимостью к широкому спектру задач НК. Данный факт ограничивает возможности выбора оптимальных решений и увеличивает затраты, связанные с приобретением, адаптацией и интеграцией разрозненных программных компонентов в масштабируемой производственной среде. Комплексные решения, обеспечивающие гибкость, модульность и адаптивность в контексте современного процесса неразрушающего контроля (НК), остаются недостаточно изученными в отечественной научной среде.

Выявленные «пробелы» в опубликованных результатах исследований и возрастающая потребность промышленности в цифровизации процессов НК обусловили *актуальность* настоящего диссертационного исследования, определив его цели и задачи.

Объектом исследования диссертационной работы являются многоканальные растровые изображения DICOM-подобного формата, рассматриваемые в контексте обработки результатов радиационного НК.

Предметом диссертационной работы является комплексирование программно-алгоритмических методов, направленных на улучшение процесса поиска схожих изображений в СППР.

Целью диссертационной работы является повышение эффективности и устойчивости выполнения поисковых запросов в прикладной системе поддержки принятия решений на основе разработки модели и алгоритмов интеллектуального

анализа изображений в DICOM-подобных данных, обеспечивающих комплексное использование визуальных и фактографических признаков в условиях больших объемов информации и распределенной сетевой архитектуры.

Для достижения поставленной цели сформулированы и решены следующие научные задачи:

1. Проведен системный анализ существующих алгоритмов поиска и интеллектуального анализа изображений, включая классические алгоритмы, методы на основе глубокого обучения и подходы к ANN-поиску, а также выявлены ограничения использования исключительно визуальных признаков в DICOM-подобных форматах хранения данных;

2. Разработана формальная модель комплексированного представления изображений, предусматривающая интеграцию визуальных эмбеддингов и токенизированных фактографических метаданных в едином семантическом пространстве признаков;

3. Разработана модифицированная архитектура ИНС «DICONDE-YOLO», обеспечивающая интеграцию метаданных на уровне внутренних представлений и адаптацию к специфике DICOM-подобных форматов хранения данных;

4. Разработан алгоритм комплексированного поиска схожих изображений с использованием ANN-индексации, обеспечивающий масштабируемость и снижение вычислительной сложности при росте объема данных;

5. Реализован программный прототип интеллектуальной СППР, интегрирующий разработанные модель и алгоритм в единое информационное пространство и обеспечивающий промышленную масштабируемость;

6. Проведена экспериментальная и эксплуатационная оценка эффективности предложенных решений в сравнении с базовыми архитектурами, основанными исключительно на визуальных признаках.

Научная новизна диссертационной работы заключается в следующих положениях:

1. Разработана модель комплексированного интеллектуального анализа изображений в DICOM-подобных данных, основанная на интеграции визуальных эмбеддингов и токенизированных метаданных в едином признаковом пространстве, что обеспечивает повышение релевантности поиска и устойчивости к внутриклассовой вариативности по сравнению с методами, использующими только визуальные признаки на 33% (*п. 12 — «Визуализация, трансформация и анализ информации на основе компьютерных методов обработки информации.»*)

2. Разработана модифицированная архитектура ИНС «DICONDE-YOLO», отличающаяся механизмом интеграции фактографических данных на уровне внутренних слоев сети и возможностью адаптивной реконфигурации с использованием AutoML-контуров, что обеспечивает повышение показателей mAP на 35%, Recall на 45%, NDCG на 33% и расширяет применимость модели в составе СППР (*п. 9 — «Разработка проблемно-ориентированных систем управления, принятия решений и оптимизации технических объектов.»*);

3. Разработан алгоритм масштабируемого комплексированного поиска схожих изображений с использованием ANN-индексации и двухуровневого механизма хранения и кэширования (векторный индекс + обратный индекс), для которого выполнено математическое обоснование экстраполяции вычислительной сложности при росте объема данных, что обеспечивает снижение времени отклика на 62% и повышение пропускной способности системы на 44% (*п. 12 — «Визуализация, трансформация и анализ информации на основе компьютерных методов обработки информации.»*);

4. Разработана архитектура распределенной интеллектуальной СППР, реализующая интеграцию сервисов обработки изображений, векторного поиска, текстового индексирования и мониторинга, обеспечивающая масштабируемость, воспроизводимость и инвариантность по отношению к структуре DICOM-подобных данных (*п. 9 — «Разработка проблемно-ориентированных систем управления, принятия решений и оптимизации технических объектов.»*).

*Методы исследования.* При выполнении диссертационных исследований использовались методы системного анализа, математической статистики, экспертных оценок, компьютерного моделирования, машинного и глубокого обучения, методы векторной индексации и приближенного поиска соседей (ANN).

*Соответствие паспорту научной специальности.* Диссертационное исследование выполнено в соответствии с пунктами 9, 12 паспорта специальностей научных работников 2.3.1. – «Системный анализ, управление и обработка информации, статистика».

Достоверность и обоснованность научных положений, теоретических выводов и рекомендаций обеспечена корректным применением математического аппарата и интеллектуальных методов на основе машинного и глубокого обучения. Достоверность подтверждается результатами вычислительных экспериментов, экспериментальной апробации и практическим применением результатов диссертационной работы.

Теоретическая значимость диссертационного исследования заключается в развитии моделей и методов интеллектуального анализа растровых изображений в DICOM-подобных форматах на основе принципов системного анализа и комплексирования разнородной информации. В работе сформирована модель интеграции визуальных и токенизированных фактографических данных в едином семантическом пространстве признаков, что расширяет теоретические представления о построении гибридных признаковых пространств для задач поиска и ранжирования изображений.

Предложенный подход углубляет теоретические основы построения СППР, ориентированных на обработку больших объемов гетерогенных данных, за счет формализации механизмов комплексирования визуальной и фактографической информации, а также математического обоснования масштабируемого ANN-поиска в условиях роста объема данных. Полученные результаты развивают положения системного анализа применительно к задачам интеллектуального поиска в

распределенных информационных средах и формируют основу для построения адаптивных архитектур СППР.

Практическая значимость работы заключается в разработке и внедрении программно-алгоритмического комплекса интеллектуального анализа изображений в DICOM-подобных форматах хранения данных, интегрированных в единое информационное пространство прикладной СППР. Реализованная система обеспечивает комплексированный поиск схожих изображений на основе объединения визуальных и фактографических признаков, что позволяет повысить релевантность выдачи, устойчивость к вариативности данных и скорость обработки запросов в условиях промышленной эксплуатации.

Разработанный программный комплекс автоматизирует процедуру интеллектуального поиска и аналитической интерпретации результатов визуального контроля, снижает влияние субъективных факторов при принятии решений и обеспечивает сокращение времени анализа при одновременном росте пропускной способности системы. Архитектура решения ориентирована на масштабируемую производственную среду и допускает интеграцию в существующие информационные контуры предприятий, использующих DICOM-подобные форматы хранения данных.

*Реализация и внедрение результатов работы.* Полученные результаты и разработанный программно-алгоритмический комплекс внедрены и используются в Учебно-аттестационном центре АО «НИКИМТ-Атомстрой» при VR симуляции операций контроля качества сварных соединений и основного металла, что подтверждается актом о внедрении результатов работы, а также – в учебном процессе и прикладных исследованиях Передовой инженерной школы химического инжиниринга и машиностроения РХТУ им. Д.И. Менделеева.

На защиту выносятся:

1. Модель комплексированного алгоритма интеллектуального поиска схожих изображений в DICOM-подобных данных, основанная на интеграции визуальных

эмбедингов и токенизированных фактографических метаданных в едином семантическом пространстве признаков;

2. Модифицированная архитектура ИНС «DICONDE-YOLO», обеспечивающая формирование информативных представлений с учетом специфики DICOM-подобных форматов и интеграцию метаданных на уровне внутренних слоев модели;

3. Архитектура распределенной интеллектуальной СППР, реализующая взаимодействие сервисов обработки изображений, векторной индексации, текстового поиска и мониторинга в условиях масштабируемой производственной среды;

4. Программный прототип прикладной СППР, реализующий разработанную модель и алгоритм комплексированного поиска схожих изображений;

5. Результаты вычислительных экспериментов и опытной эксплуатации, подтверждающие повышение показателей качества поиска, устойчивости работы системы и эффективности поддержки принятия решений по сравнению с базовыми решениями, основанными исключительно на визуальных признаках.

*Апробация работы.* Основные положения диссертационной работы докладывались и обсуждались на научном семинаре «Системный анализ и обработка данных в психологии и образовании» (РАО) (Москва, 2024), XVI Международной научной конференции «Интеллектуальные системы» (INTELS'24) (Москва, 2024), Международной научно-технической конференции молодых ученых «Новые материалы, оборудование и технологии в промышленности» (Могилев, 2024), VIII Всероссийской научно-практической конференции с международным участием «Программная инженерия: современные тенденции развития и применения (ПИ-2024)» (Курск, 2024), II Межрегиональной научно-практической конференции с международным участием «Цифровые, компьютерные и информационные технологии в науке и образовании» (Брянск, 2024), VIII Международном научно-техническом форуме «Современные технологии в науке и образовании» (Рязань, 2025), III Межрегиональной научно-практической конференции «Цифровые, компьютерные и информационные технологии в науке и образовании» (Брянск, 2025),

), «Фундаментальные исследования и инновационные технологии в машиностроении – FRITME» (Москва, 2025), а также – на научных семинарах Передовой инженерной школы химического инжиниринга и машиностроения РХТУ им. Д.И. Менделеева и расширенных заседаниях секции №8 НТС Научно-исследовательского и конструкторского института монтажной технологии АО «НИКИМТ-Атомстрой» (Москва, 2024-2026).

*Личный вклад автора.* Представленные к защите результаты теоретических и экспериментальных исследований получены лично соискателем или при его непосредственном участии. В частности, соискателем выполнена формализация задачи комплексированного поиска схожих изображений в DICOM-подобных данных с учетом необходимости интеграции визуальной и фактографической информации в рамках единого семантического пространства признаков. Разработана модель комплексированного алгоритма поиска, включающая этапы предобработки, токенизации метаданных, формирования эмбеддингов и ANN-индексации, а также определено место алгоритма в структуре прикладной СППР. Разработана и обоснована модифицированная архитектура ИНС «DICONDE-YOLO», отличающаяся интеграцией токенизированных метаданных на уровне внутренних представлений и адаптацией к специфике DICOM-подобных форматов. Проведено экспериментальное исследование базовых и современных нейросетевых архитектур в задачах детекции и поиска дефектов, выполнен сравнительный анализ с прототипом, основанным исключительно на визуальных признаках. Соискателем спроектирована архитектура распределенного программного прототипа СППР, включая структуру внутреннего и межсетевого взаимодействия компонентов, механизмы векторного поиска, кэширования, балансировки нагрузки и мониторинга. Реализован программный прототип, интегрированный в единое информационное пространство предприятия.

Автором проведены вычислительные эксперименты, выполнена оценка качественных и эксплуатационных характеристик СППР, проанализированы показатели точности, релевантности и масштабируемости, а также обоснована

экстраполяция производительности ANN-индекса при росте объема данных. Полученные результаты подтвердили достижение цели исследования и эффективность предложенных моделей и алгоритмов в составе прикладной СППР.

*Объем и структура диссертации.* Диссертационная работа изложена на 129 страницах машинописного текста и состоит из введения, четырех глав, заключения, списка литературы и семи приложений. Список литературы включает в себя 101 источник.

*Публикации.* По теме исследования опубликовано 16 работ, в том числе 3 статьи в журналах, индексируемых в международной наукометрической базе данных Scopus и 7 статей в изданиях, рекомендованных ВАК Минобрнауки России, 8 работ в материалах международных и всероссийских конференций, 1 свидетельство о государственной регистрации программы для ЭВМ.

## **Глава 1. АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ И ПОСТАНОВКА ЗАДАЧИ ИССЛЕДОВАНИЯ**

Проведен анализ предметной области, охватывающей вопросы обработки данных в формате DICOM [5], а также производных форматов, основанных на DICOM [6, 7]. Проводится детализированное рассмотрение структуры формата DICOM, особенности хранения, а также методов обработки данных.

Особое внимание уделяется выявлению ограничений, присущих существующим решениям в области обработки DICOM-изображений, что определяет границы их применимости. Анализируются сложности, возникающие при разработке инвариантного подхода к обработке DICOM-подобных данных, учитывающего неоднородность форматов хранения. Анализ основывается на современных научных публикациях и передовом опыте соответствующих отраслей, связанных с хранением и обработкой данных в DICOM-подобной структуре. Это обеспечивает необходимую теоретическую основу для дальнейших исследований.

Результаты проведенного анализа служат основой для формулировки целей и задач диссертационного исследований, а также определения критериев оценки эффективности предлагаемых решений и разрабатываемых модели и алгоритмов комплексного поиска схожих изображений.

### **1.1 Стандарты DICOM, DICONDE, DICOS**

Семейство стандартов DICOM [5] предназначено для структурированной передачи, обработки изображений и сопутствующей информации. На его основе разработаны специализированные расширения, включая DICONDE [6],

адаптированный под задачи хранения, обработки и передачи информации о результатах НК, а также DICOM [7], применяемый в системах радиационной и досмотровой диагностики. Несмотря на различие предметных областей, указанные форматы реализуют единый принцип представления данных — объединение визуальной и фактографической информации в рамках унифицированной структуры [5-7].

Каждый файл DICOM-подобного формата содержит растровые данные (визуальные данные) и набор метаданных (фактографические данные), организованных в виде иерархической системы тегов. Такая структура обеспечивает формализованное описание условий получения изображения, параметров оборудования и характеристик объекта исследования, а также корректную интерпретацию визуальной составляющей.

Визуальная составляющая представлена многомерными массивами интенсивностей с глубиной квантования 8–16 бит и канальностью 1–N. Значения пикселей не обладают однозначным физическим смыслом без учета соответствующих метаданных, что делает невозможным корректный анализ изображения исключительно на основе пиксельной информации. Параметры формирования изображения, масштабирование, динамический диапазон и интерпретация интенсивностей определяются атрибутами файла.

Фактографическая составляющая представлена структурированными атрибутами и описывает условия получения изображения. В задачах поиска схожих изображений такие данные могут использоваться как семантические ограничения, позволяющие сократить пространство поиска и повысить релевантность результатов. При этом метаданные обладают большей устойчивостью к шумам по сравнению с визуальной информацией, однако не содержат сведений о пространственной структуре объектов.

Таким образом, ни визуальная, ни фактографическая информация по отдельности не обеспечивают достаточной полноты анализа. Использование

исключительно визуальных признаков приводит к снижению устойчивости результатов при вариативности условий съемки, тогда как применение только метаданных не позволяет выявлять локальные структурные особенности изображений. Это обуславливает необходимость разработки формальной модели интеграции разнородных признаков.

Особенностью визуальных данных в DICOM-подобных форматах является отсутствие жесткой привязки пиксельных значений к конкретному физическому смыслу без учета метаданных [5-7]. Значения интенсивности могут быть представлены в различных шкалах, иметь различную интерпретацию и требовать дополнительного преобразования для корректного анализа. Это обстоятельство накладывает ограничения на применение универсальных методов обработки изображений без учета контекста получения данных.

В задачах поиска схожих изображений визуальная информация используется для извлечения признаков, отражающих локальные и глобальные характеристики изображения. К таким признакам относятся текстурные показатели, статистические характеристики распределения интенсивностей, геометрические структуры, а также высокоуровневые дескрипторы, получаемые с использованием методов машинного обучения. Несмотря на высокую информативность визуальных признаков, их использование в изоляции от контекста приводит к снижению устойчивости результатов поиска, особенно в условиях вариативности параметров съемки и наличия шумов.

Дополнительным фактором, влияющим на эффективность визуального анализа, является полихроматичность данных. Многоканальные изображения требуют согласованной обработки каналов и адаптации процедур извлечения признаков, поскольку простое объединение информации по каналам не всегда обеспечивает корректное отражение структурных особенностей объекта. В то же время одноканальные изображения, характерные для ряда радиационных и

томографических методов, представляют собой частный случай данной модели и также требуют учета параметров представления данных.

Визуальные данные в стандартах DICOM, DICONDE и DICOS обладают высокой информативностью и являются важным компонентом в задачах анализа и поиска изображений. Однако их эффективное использование невозможно без учета условий получения и интерпретации, которые неявно представлены непосредственно в пиксельных данных. Это обстоятельство обуславливает необходимость привлечения дополнительной информации, содержащейся в структуре форматов, для повышения точности и устойчивости принимаемых решений.

Основные характеристики форматов DICOM, DICONDE и DICOS приведены в таблице 1.1 и отражены на рисунке 1.

Таблица 1.1 – Основные характеристики DICOM-подобных форматов

<b>Характеристика</b>	<b>DICOM</b>	<b>DICONDE</b>	<b>DICOS</b>
Основная область применения	Здравоохранение	Неразрушающий контроль	Досмотровая и радиационная инспекция
Тип изображения	2D, 3D, 4D	2D, многоканальные сенсоры	2D
Канальность	1 – N	1 – N	1 – N
Глубина квантования	8 – 16 бит	8 – 16 бит	8 – 16 бит
Метаданные	Полный набор для медицины	Ограниченный для NDE	Ограниченный для инспекции
Структура хранения	Файл – серия – исследование – метаданные / изображение	Файл – серия – исследование – метаданные / изображение	Файл – серия – исследование – метаданные / изображение

## Продолжение таблицы 1.1

Характеристика	DICOM	DICONDE	DICOS
Возможность предварительной фильтрации	Да	Да	Да
Тип фактографических данных	Стандартизированные теги, семантика медицины	Теги для оборудования, материала	Теги для оборудования, условий досмотра

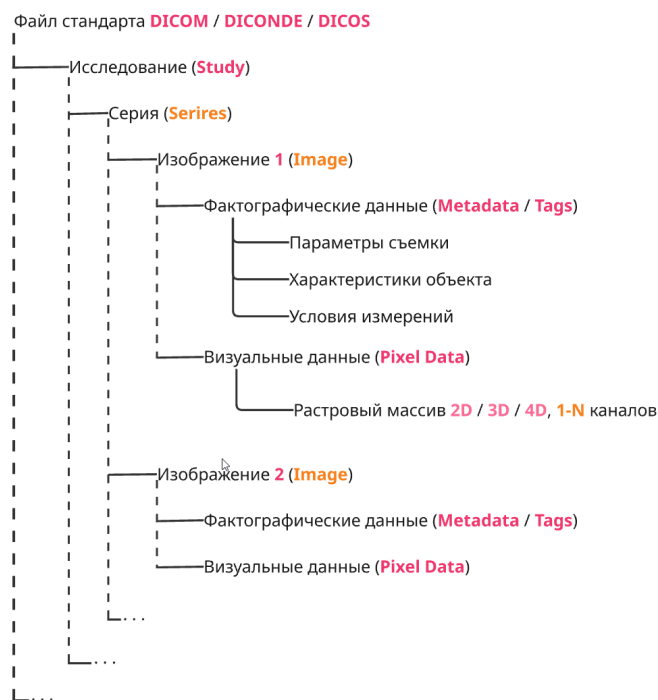


Рис. 1. Обобщенная структура DICOM-подобных форматов

Таким образом, стандарты семейства DICOM задают единое пространство фактографических данных, что обеспечивает возможность создания алгоритмов поиска, инвариантных к специфичной предметной области, и формирует основу для последующего проектирования модели и алгоритма поиска изображений с интеграцией визуальных и фактографических признаков.

Проведенный анализ стандартов, представленный в таблице 1.1 демонстрирует, что данные форматы обеспечивают структурированное хранение визуальной и фактографической информации. Однако само наличие данных не определяет способ их применения в задачах интеллектуального поиска схожих изображений. Это обуславливает необходимость разработки формальной модели совместного учета визуальных и фактографических признаков, которая будет представлена во второй главе диссертационной работы.

## **1.2 Алгоритмы поиска изображений и их ограничения**

Извлечение информативных признаков из визуальных и фактографических данных, рассмотренных в разделе 1.1, является ключевым этапом построения СППР. Основная задача поиска схожих изображений заключается в выявлении объектов или областей, обладающих близкими визуальными характеристиками, при одновременном обеспечении минимального времени отклика системы и снижении вероятности ложных совпадений. Алгоритмы, применяемые для решения данной задачи, условно подразделяются на классические методы обработки изображений и методы, основанные на машинном и глубоком обучении, представленные нейронными сетями. Каждый из указанных подходов обладает собственными преимуществами и ограничениями [21, 22], которые необходимо учитывать при решении основной задачи исследования.

Классические методы извлечения признаков ориентированы на анализ локальных и глобальных характеристик изображений и широко применялись в задачах сопоставления и поиска задолго до появления нейросетевых подходов. Значительную группу таких методов составляют локальные дескрипторы, включая SIFT [23], SURF [24] и ORB [25], основанные на выделении устойчивых ключевых

точек и формировании их описаний, инвариантных к масштабированию, повороту и изменениям освещенности. Данные методы позволяют выявлять характерные элементы изображения и выполнять сопоставление на уровне локальных структур. Пример сравнения на основе ключевых точек при использовании алгоритма SIFT представлен на рисунке 2.

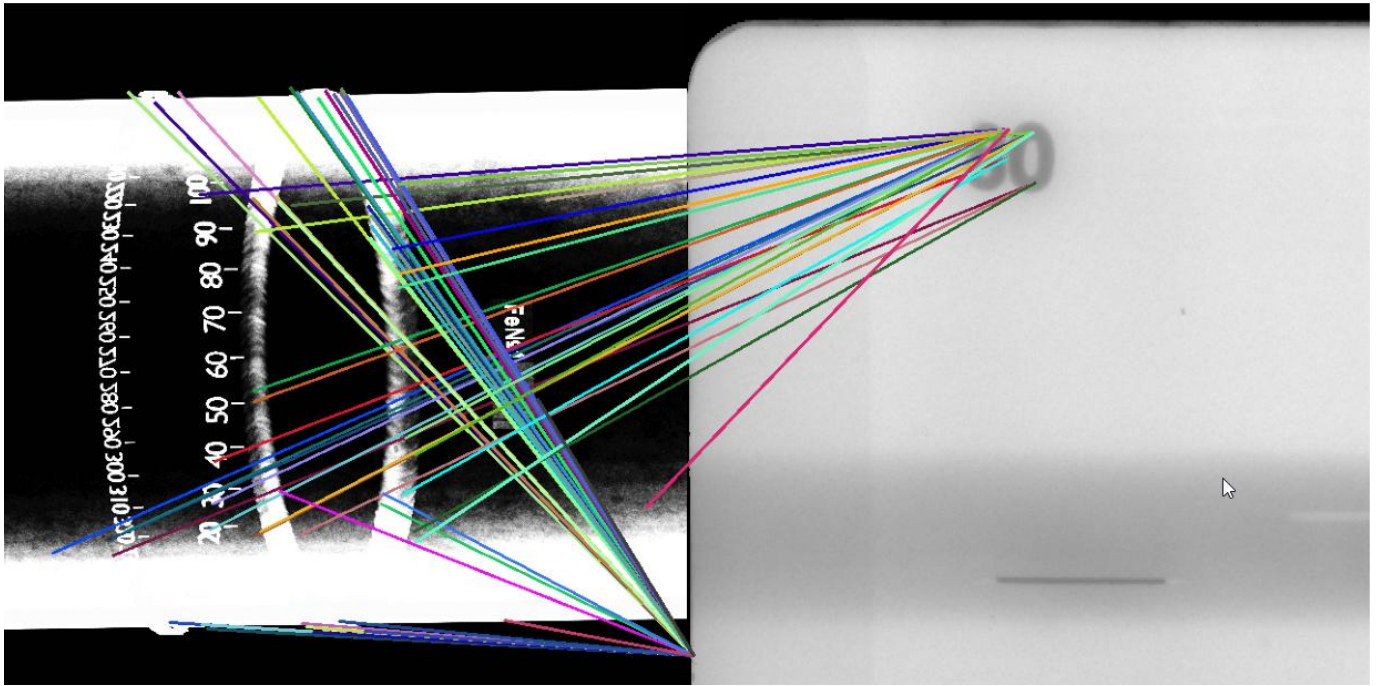


Рис. 2. Сравнение на основе ключевых точек с помощью алгоритма SIFT [23]

Вместе с тем их практическое применение в задачах поиска по крупным БД ограничено высокой вычислительной сложностью, а также чувствительностью к шумам и артефактам, характерным для изображений после проведения НК и досмотровой, либо радиационной инспекции.

Другой класс подходов представлен методами анализа распределений градиентов и текстур, к которым относятся HoG [26], LBP [27] и другие. Эти методы кодируют локальные изменения интенсивности и структуры изображения, что позволяет описывать текстурные свойства поверхностей и материалов. Их достоинством является вычислительная простота и интерпретируемость, однако они слабо отражают глобальную форму объектов и плохо масштабируются при увеличении размерности признакового пространства [28, 29]. Пример сравнения на

основе распределения текстур с использованием алгоритма LBP представлен на рисунке 3. Верхняя пара отражает абсолютно одинаковые изображения, нижняя пара – различающиеся варианты.

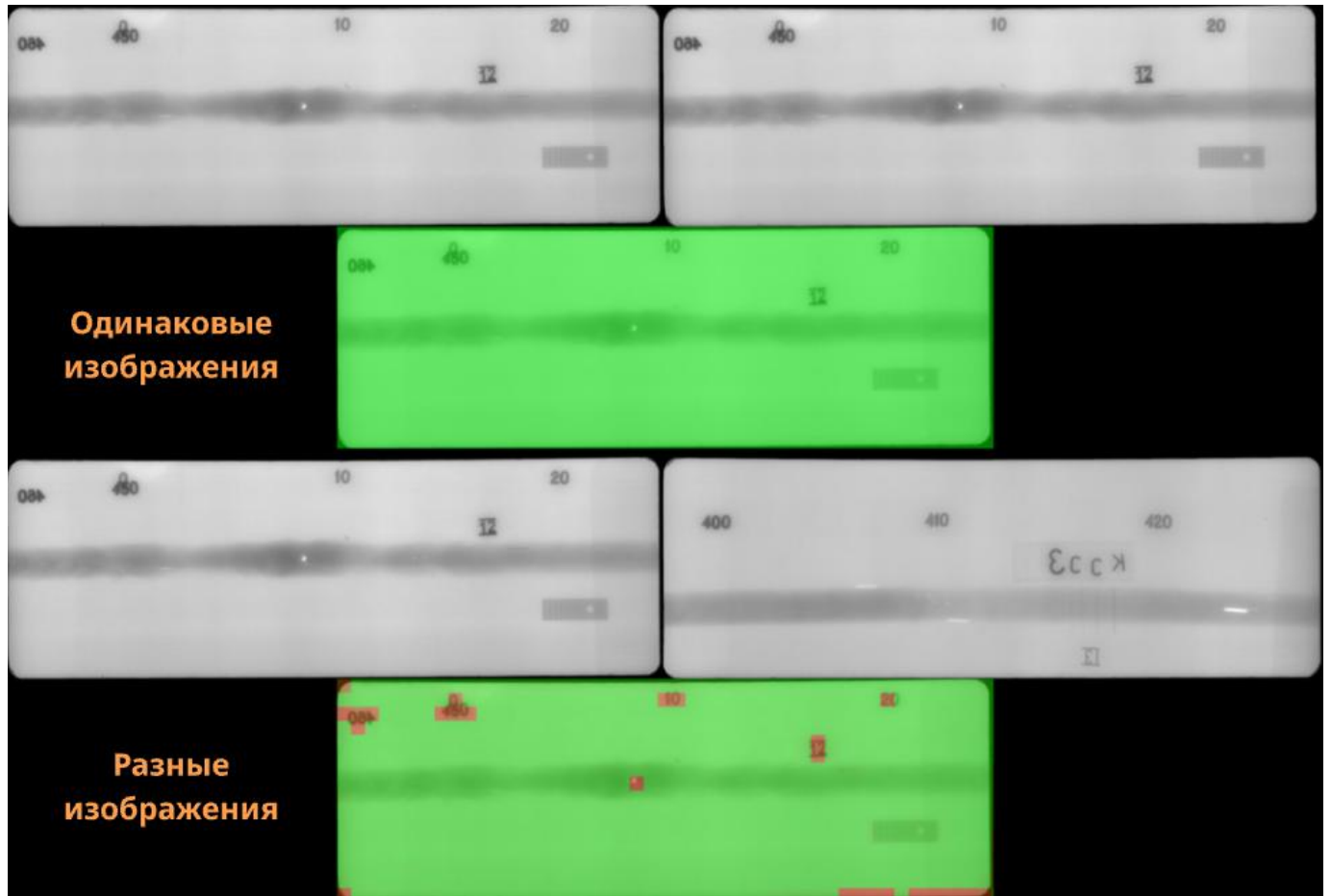


Рис. 3. Сравнение изображения с помощью алгоритма LBP

Для формирования более компактного глобального представления изображения используются методы агрегации локальных признаков, в частности подходы типа BoVW [30]. В рамках данного подхода локальные дескрипторы преобразуются в статистическое описание, что позволяет упростить сравнение изображений и повысить скорость поиска. Вместе с тем подобная агрегация приводит к потере пространственной информации и снижению точности при анализе сложных структурированных объектов. Пример сравнения изображений на основе алгоритма BoVW представлен на рисунке 4. В качестве инструмента извлечения признаков выступал алгоритм SIFT.

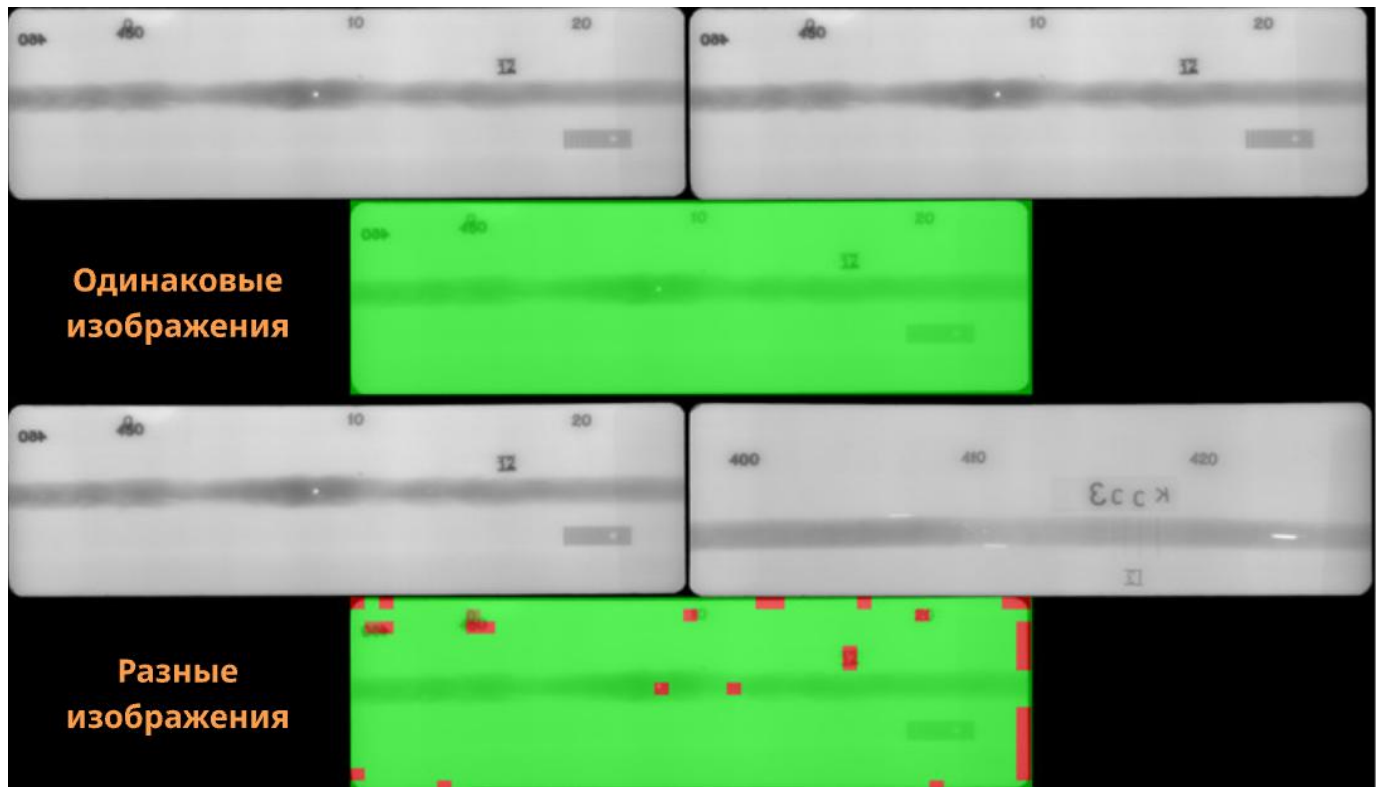


Рис. 4. Сравнение изображений с помощью алгоритма VoVW

В задачах анализа формы и геометрии объектов применяются методы выделения границ и сегментации, включая операторы Канни [31], Собеля и Лапласа [32]. После выделения контуров вычисляются геометрические характеристики объектов, такие как площадь, периметр, соотношение сторон и углов. Эти признаки хорошо отражают форму и структуру дефектов, однако отличаются высокой чувствительностью к шуму, ориентации и масштабированию изображений, что ограничивает их универсальность. Пример сравнения изображений на основе оператора Канни представлен на рисунке 5. В качестве инструмента поиска ключевых точек после применения оператора выступал алгоритм ORB.

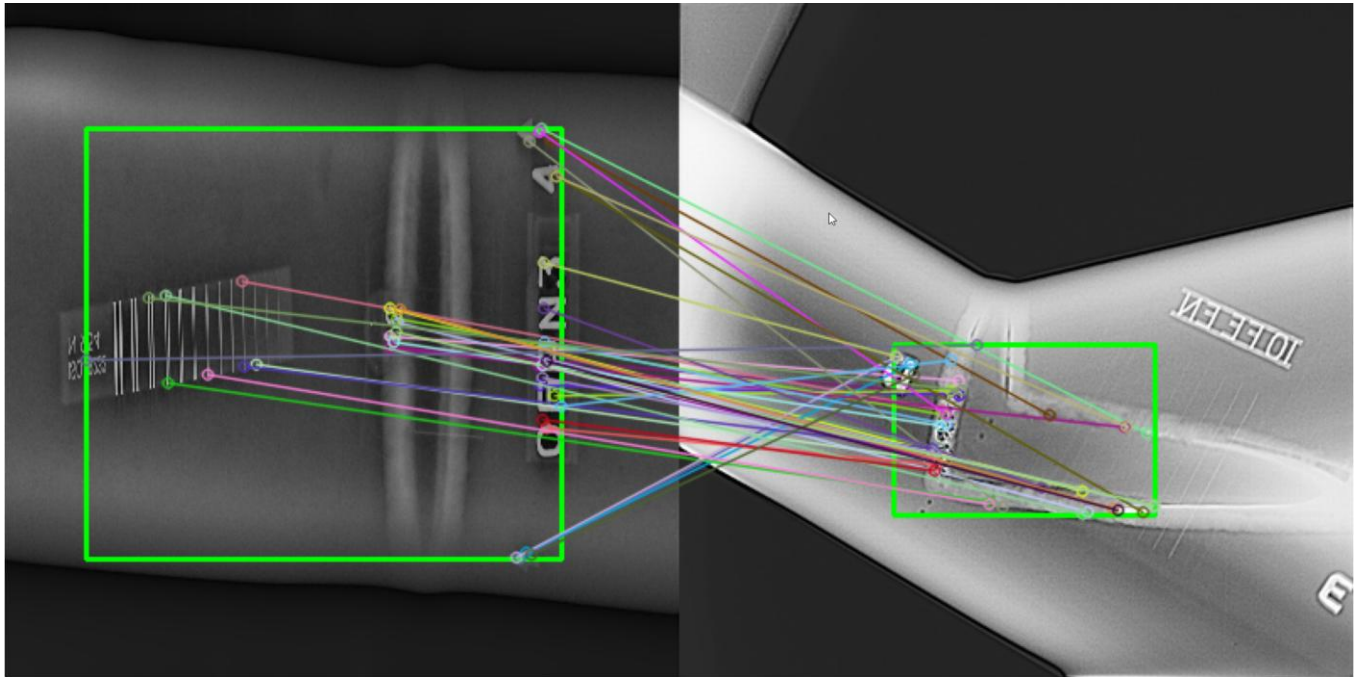


Рис. 5. Пример сравнения изображений на основе оператора Канни

Частотные методы, включая дискретное косинусное преобразование, быстрое преобразование Фурье [33] и вейвлет-преобразование [34], позволяют анализировать изображения в частотной области и выявлять повторяющиеся структурные особенности. Они демонстрируют устойчивость к локальным изменениям освещенности и шуму, однако сопровождаются потерей локальной пространственной информации и требуют дополнительных процедур при обработке многоканальных данных. Пример сравнения на основе Фурье и Вейвлет-преобразований иллюстрируется на рисунке 6.

Современные методы поиска схожих изображений в значительной степени основаны на машинном обучении и глубоких нейронных сетях, обеспечивающих формирование высокоуровневых признаков. В частности, СНС осуществляют последовательную обработку изображения с помощью сверточных и субдискретизирующих операций, формируя карты признаков, отражающие информацию различных уровней абстракции. Полученные карты преобразуются в векторы признаков и сравниваются с использованием метрик сходства, таких как евклидово или косинусное расстояние, что позволяет выявлять сложные визуальные

аналогии между изображениями. Основным преимуществом данного подхода является высокая устойчивость к шуму, масштабным изменениям и частичной окклюзии объектов. В то же время нейросетевые модели требуют значительных объемов обучающих данных и вычислительных ресурсов, а также могут быть чувствительны к специфике предметной области.

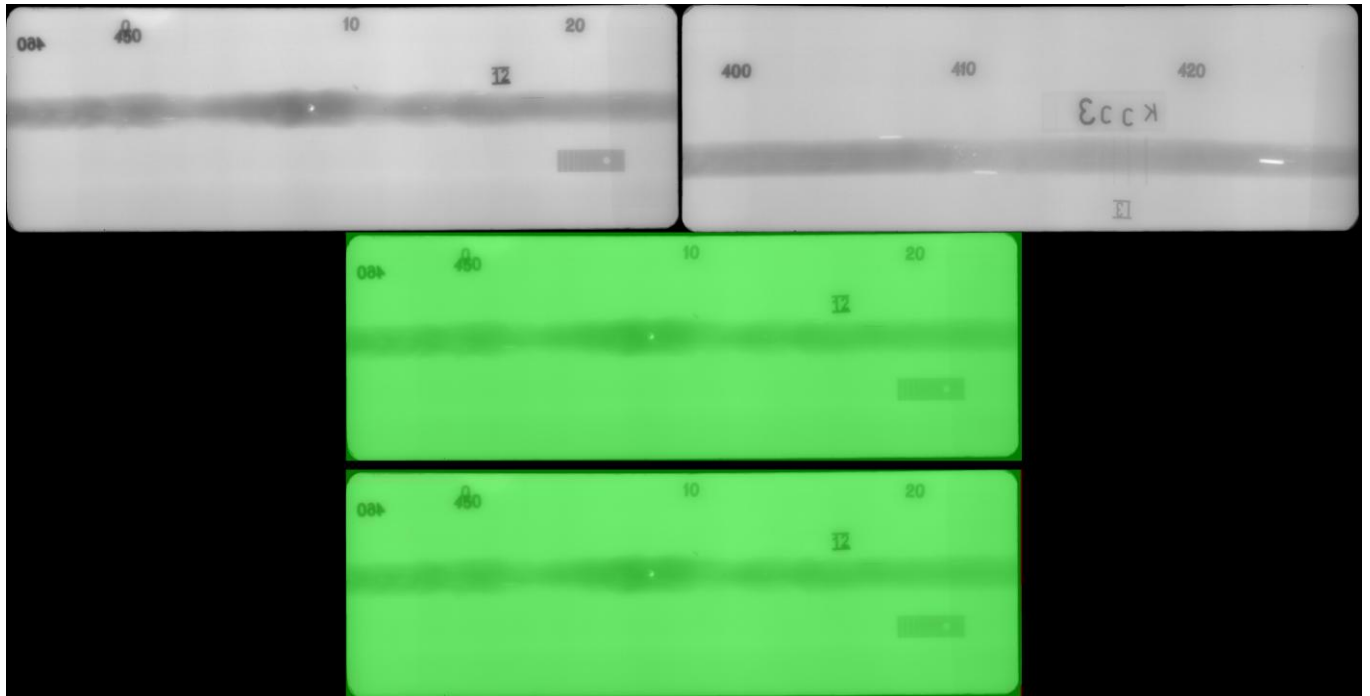


Рис. 6. Пример сравнения на основе Фурье и Вейвлет-преобразований

В задачах, где требуется локализация значимых областей изображения, применяются архитектуры детекции объектов, такие как YOLO [35] и EfficientDet [36]. Эти модели позволяют автоматически выделять области интереса, соответствующие потенциальным дефектам или структурным элементам, извлекать визуальные признаки на разных уровнях представления, что позволяет выполнять поиск схожих изображений на уровне отдельных фрагментов. Такой подход повышает интерпретируемость результатов анализа, однако требует интеграции с методами глобального сравнения и усложняет архитектуру системы.

Помимо YOLO и EfficientDet, для задач локализации объектов и выделения областей интереса применяются архитектуры семейства R-CNN, включая Fast R-CNN,

Faster R-CNN и Mask R-CNN [37], а также детекторы с одним проходом, такие как SSD. Эти модели ориентированы на идентификацию и классификацию отдельных объектов или дефектов на изображении, что позволяет создавать репрезентативные локальные дескрипторы и повышать точность поиска схожих объектов. Классическая R-CNN сначала генерирует кандидаты регионов (Region Proposals), после чего каждый из них проходит через СНС для формирования признаков и классификации. Fast R-CNN улучшает данный подход, объединяя обработку всех регионов в едином проходе по CNN, что сокращает время вычислений и улучшает точность обучения. Далее Faster R-CNN вводит Region Proposal Network, позволяющую генерировать кандидаты регионов непосредственно в нейронной сети, что значительно ускоряет процесс детекции. Mask R-CNN расширяет Faster R-CNN, добавляя возможность сегментации объектов, что позволяет не только локализовать, но и точно выделять форму дефекта, включая сложные структурные элементы. Пример сравнения изображений на основе обученной модели YOLO 8-й версии представлен на рисунке 7.

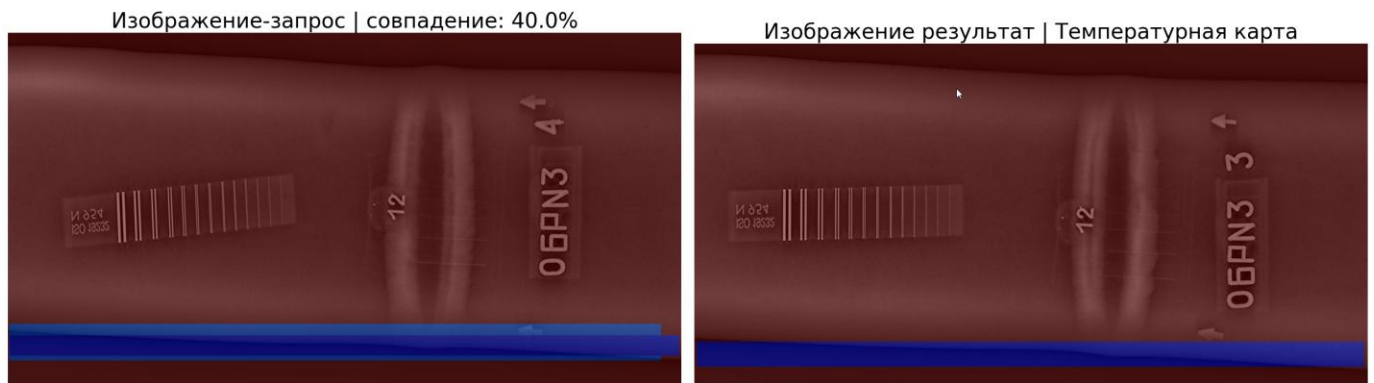


Рис. 7. Пример сравнения на основе модели YOLO 8-й версии

SSD [38], в свою очередь, выполняет детекцию объектов в одном проходе, сразу предсказывая координаты ограничивающих рамок и классы объектов на различных масштабах изображения, что делает его пригодным для систем с ограниченными ресурсами и требованиями к реальному времени. Все перечисленные методы обладают высокой точностью локализации и позволяют работать с множеством

объектов на изображении, что особенно важно для задач анализа в DICOM-подобных форматах.

Однако применение нейросетевых архитектур сопряжено с определенными ограничениями. Высокая вычислительная сложность требует мощных аппаратных ресурсов и увеличивает время отклика системы. Для обучения моделей необходимы размеченные датасеты с точной локализацией дефектов, что в прикладных областях НК и инспекции часто затруднительно. Кроме того, модели чувствительны к специфике домена данных, и их адаптация к особенностям изображений требует дополнительных процедур нормирования и согласования представлений в каждом цветовом канале.

Отдельное направление развития составляют модели на основе трансформеров, включая ViT [39] и Swin Transformer [40]. В отличие от СНС, данные архитектуры обрабатывают изображение как последовательность патчей и позволяют учитывать глобальные зависимости между различными областями изображения. Это обеспечивает более целостное представление структуры объектов и повышает точность анализа сложных сцен. Ограничениями таких моделей являются высокая вычислительная сложность и необходимость значительных объемов данных для обучения, существенно превышающих объемы, используемые в сверточных аналогах.

Дополнительно в задачах масштабируемого поиска по крупным БД применяются методы нейросетевого хэширования, включая CNNH [41], DPSH [42], UDQH-IQ [43] и SPL-UDH [44]. CNNH формирует бинарные коды через СНС, минимизируя различия между схожими изображениями, что ускоряет поиск и снижает требования к памяти. DPSH интегрирует дискретное хэширование с оптимизацией по парным связям, повышая точность на больших БД. UDQH-IQ использует квантизирующие операции с учетом распределения признаков, а SPL-UDH сочетает селективное обучение с дискретными хэш-кодами, обеспечивая баланс между компактностью представления и устойчивостью к шуму. Эти методы существенно ускоряют поиск, но требуют репрезентативных обучающих выборок и

точной настройки архитектуры ИНС. Пример сравнения на основе метода SPL-UDH представлен на рисунке 8.

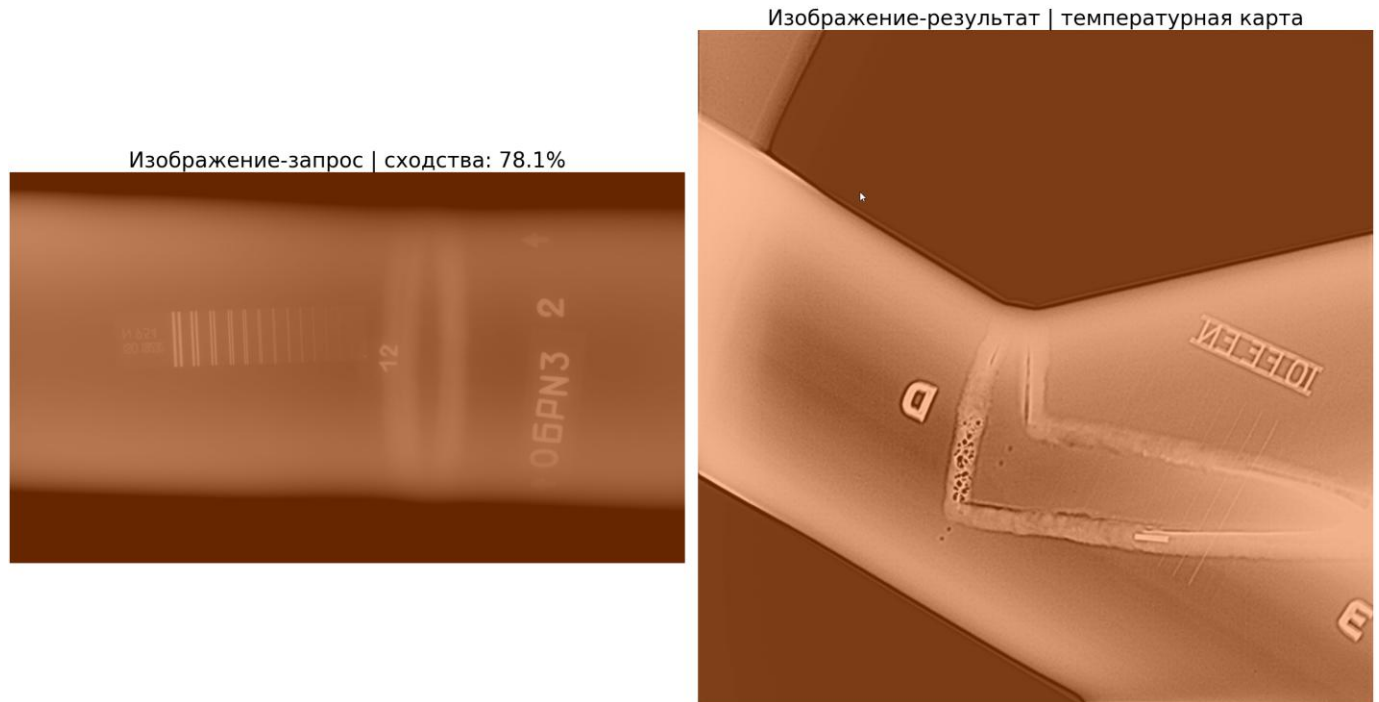


Рис. 8. Пример сравнения на основе метода нейросетевого хэширования SPL-UDH

При обработке многоканальных изображений, характерных для форматов DICOM, DICONDE и DICOS, анализ канальной информации рассматривается как совместная обработка в виде единого тензора, так и отдельное извлечение признаков с последующим объединением. Такой подход позволяет учитывать спектральные и структурные особенности данных, однако увеличивает вычислительную нагрузку и предъявляет дополнительные требования к операциям нормализации.

Анализ существующих алгоритмов поиска схожих изображений показывает, что подходы, основанные исключительно на визуальных признаках, демонстрируют ограниченную устойчивость в условиях высокой визуальной схожести. В данном контексте под подходами понимаются сами алгоритмы. Кроме того, ни один из существующих подходов не обеспечивает универсального и устойчивого решения задачи. Применение данных подходов по отдельности выявляет ряд принципиальных ограничений. Классические алгоритмы демонстрируют ограниченную устойчивость

к вариативности параметров съемки и шумам, а варианты, основанные на локальных дескрипторах, плохо масштабируются при работе с большими БД. Алгоритмы, основанные на агрегации глобальных признаков, сопровождаются потерей пространственного контекста, что снижает точность анализа сложных структур.

Нейросетевые алгоритмы позволяют формировать более выразительные признаки, однако требуют значительных объемов размеченных данных и чувствительны к смещению домена, характерному для прикладных задач НК. Дополнительную сложность представляет обработка многоканальных изображений, где универсальные подходы не всегда корректно отражают физический смысл данных.

На основе проведенного анализа, была составлена обобщенная таблица 1.2, отражающая усредненные количественные показатели подходов к поиску схожих изображений.

Таблица 1.2 Количественные показатели подходов к поиску схожих изображений

<b>Подход</b>	<b>mAP@50</b>	<b>Recall</b>	<b>Время инференса</b>
Классический	0.45-0.55	0.5-0.6	Среднее
CNN	0.65-0.75	0.63-0.71	Низкое
ViT	0.75-0.85	0.71-0.83	Высокое
CNN-хэширование	0.6-0.7	0.6-0.65	Высокое

Таким образом, существующие алгоритмы поиска схожих изображений демонстрируют высокую эффективность в рамках отдельных классов задач, однако обладают рядом принципиальных ограничений, связанных с устойчивостью, масштабируемостью, вычислительной сложностью и интерпретируемостью результатов. Подобные ограничения обуславливают необходимость перехода к подходам, основанным на совместном использовании визуальных и фактографических данных в рамках единого признакового пространства, что рассматривается в последующих разделах диссертационного исследования.

### 1.3 Методы комплексирования при поиске изображений

Ограничения отдельных подходов, перечисленные в предыдущем разделе, обуславливают необходимость перехода к комплексированию алгоритмов поиска изображений, которое представляет собой системное объединение нескольких алгоритмов анализа визуальных и фактографических данных, позволяющее компенсировать слабые стороны отдельных подходов за счет сильных сторон других с целью повышения точности, устойчивости и вычислительной эффективности поиска схожих изображений. В условиях работы с DICOM-подобными форматами, где данные имеют многоканальную структуру и включают как визуальные, так и фактографические компоненты, комплексирование становится важным этапом для обеспечения корректности выходных результатов и минимизации числа ложных совпадений.

Концептуально комплексирование может осуществляться на трех взаимосвязанных уровнях [45-47].

Первый уровень, который принято называть комплексированием на уровне данных, заключается в объединении разнородных источников информации еще до извлечения признаков. Это позволяет создавать единое, согласованное представление объекта, включающее визуальные данные различных каналов и спектральных диапазонов, а также фактографические атрибуты метаданных, такие как параметры съемки, условия контроля и характеристики исследуемого материала. Такой подход обеспечивает максимальное использование исходной информации, но требует корректного нормирования данных и согласования их масштабов и типов.

Второй уровень, известный как комплексирование на уровне признаков, предполагает извлечение данных различными методами из одной и той же исходной информации с последующим объединением в единый вектор или дескриптор. Например, локальные дескрипторы, такие как SIFT или ORB, могут сочетаться с текстурными характеристиками HoG и LBP, а также с признаками, извлеченными

СНС. Такой подход позволяет учитывать как локальные, так и глобальные свойства объектов и обеспечивает более информативное признаковое представление. Основным ограничением данного уровня является увеличение размерности признакового пространства и рост вычислительной нагрузки при обработке больших БД.

Третий уровень именуется комплексированием на уровне решений. Реализуется после получения отдельных оценок сходства от нескольких методов, когда результаты каждого из них агрегируются для формирования окончательного решения. Агрегация может осуществляться с использованием взвешенного голосования, ранжирования или обучаемых моделей объединения. Этот подход минимизирует влияние ошибок отдельных методов и повышает стабильность результатов, однако является менее информативным по сравнению с комплексированием на уровне признаков, так как теряется часть детальной информации о структуре объектов.

Все три уровня комплексирования взаимосвязаны и часто применяются совместно. На схематическом уровне концептуальная модель комплексирования может быть представлена следующим образом. Исходные визуальные и фактографические данные поступают на этап предварительной обработки и нормирования, после чего происходит извлечение признаков разными методами. Полученные признаки объединяются в единое представление, которое затем используется для вычисления метрик сходства. Итоговые оценки могут дополнительно агрегироваться на уровне решений для формирования окончательного ранжирования изображений по степени соответствия запросу. Такой многоуровневый подход обеспечивает баланс между точностью, скоростью поиска и устойчивостью к шумам и вариативности условий съемки.

В рамках настоящей диссертационной работы комплексирование рассматривается как способ повышения устойчивости и релевантности поиска схожих изображений в DICOM-подобных форматах. Особое внимание уделяется согласованной интеграции визуальных и фактографических данных, что позволяет

учитывать как пространственные характеристики объектов, так и контекст их получения. Такой подход формирует основу для построения комплексированного алгоритма поиска, обеспечивающего инвариантность к условиям съемки и масштабируемость при работе с крупными БД.

#### **1.4 Вывод по главе**

В первой главе выполнен анализ предметной области интеллектуального поиска изображений в DICOM-подобных форматах, рассмотрены особенности стандартов хранения визуальных и фактографических данных, а также существующие алгоритмы интеллектуального анализа изображений. Показано, что использование исключительно визуальных, либо исключительно фактографических признаков не позволяет обеспечить требуемое качество поиска и поддержки принятия решений в условиях высокой вариативности параметров контроля и неполноты данных.

Установлено, что ключевым направлением повышения эффективности интеллектуального анализа является формализованная интеграция визуальных и фактографических данных в рамках единого признакового пространства. Это определяет необходимость строгой формализации структуры данных, пространства признаков и критериев сходства, а также разработки модели и комплексированного алгоритма поиска схожих изображений на основе совместного применения визуальных и фактографических данных, что является предметом рассмотрения второй главы диссертационной работы.

## **Глава 2. ФОРМАЛИЗАЦИЯ ЗАДАЧИ КОМПЛЕКСИРОВАННОГО ПОИСКА СХОЖИХ ИЗОБРАЖЕНИЙ В DICOM-ПОДОБНЫХ ФОРМАТАХ**

Цель данной главы заключается в разработке формальной модели поиска, описывающей совместный учет визуальных и фактографических признаков, согласованную обработку многоканальных изображений, а также устойчивость и воспроизводимость результатов при изменяемых условиях проведения НК. Указанная задача рассматривается в контексте инвариантного применения в промышленной СППР, где критически важными являются как точность выдачи релевантных результатов, так и скорость обработки поисковых запросов при масштабировании объемов данных.

Достижение поставленной цели требует строгой формализации исходных данных, способов их представления, а также алгоритмов преобразования и сравнения визуальной и фактографической информации. В текущей главе последовательно вводятся математические модели, описывающие структуру DICOM-подобных форматов, пространство признаков и критерии сходства, модель ИНС «DICONDE-YOLO» после чего формализуется комплексированный алгоритм поиска схожих изображений, интегрирующий визуальную информацию и фактографические сведения, извлекаемые из DICOM-подобных изображений.

### **2.1 Формализация DICOM-подобных данных и пространства признаков**

Рассматривается абстрактная БД изображений НК для формализации задачи исследования. Здесь под БД понимается совокупность цифровых изображений и соответствующих им метаданных, формируемая различными система регистрации,

методиками проведения НК, а также – конфигурациями технологического оборудования. Конкретная реализация системы хранения, формат физического представления данных, а также способ их размещения в хранилище не являются фиксированными и не оказывают влияния на разрабатываемую математическую модель. Введение данной абстракции позволяет провести описание алгоритма обработки и поиска в обобщенной форме, сохраняя инвариантность применения полученных результатов к широкому классу DICOM-подобных форматов хранения данных.

Рассмотрим БД изображений НК, представленную в виде множества:

$$\mathcal{D} = \{D_1, D_2, \dots, D_N\}, \quad (2.1)$$

Каждый элемент БД представляется в виде упорядоченной пары:

$$D_i = (X_i, M_i), \quad (2.2)$$

где  $X_i$  – тензор данных изображения, а  $M_i$  – множество метаданных, представленных в виде набора тегов DICOM-подобного стандарта.

Визуальная составляющая  $X_i$  описывается как многомерный тензор:

$$X_i \in \mathbb{R}^{H_i \times W_i \times C_i}, \quad (2.3)$$

где  $H_i$  и  $W_i$  – пространственные размеры изображения, а  $C_i$  – число каналов, определяемое используемой модальностью и параметрами регистрации. Под каналом в данном контексте подразумевается отдельное измерение интенсивности, соответствующее вариативным цветовым диапазонам, сенсорам или способам реконструкции изображений. Одноканальные изображения рассматриваются как частный случай модели при  $C_i = 1$ .

Фактографическая составляющая  $M_i$  представляет собой конечное множество атрибутов:

$$M_i = \{m_{i1}, m_{i2}, \dots, m_{iK}\}, \quad (2.4)$$

где каждый атрибут  $m_{ij}$  соответствует отдельному тегу DICOM-подобного формата и описывает параметры получения изображения, характеристики исследуемого объекта или условия проведения измерений.

Пространство метаданных задается в виде декартового произведения:

$$\mathcal{M} = \mathcal{M}_{num} \times \mathcal{M}_{cat} \times \mathcal{M}_{text}, \quad (2.5)$$

где  $\mathcal{M}_{num}$  – числовые параметры (напряжение, расстояние, толщина),  $\mathcal{M}_{cat}$  – категориальные признаки (тип контроля, оборудование),  $\mathcal{M}_{text}$  – символьные и строковые описания.

В таблице 2.1 представлен фрагмент перечня используемых метаданных и их потенциальное применение при обучении ИНС.

Таблица 2.1 Фрагмент используемых метаданных

<b>Тег</b>	<b>Описание</b>	<b>Потенциальное применение при обучении ИИ</b>
(0008, 1030) Study Description	Описание исследования	Применяется для извлечения семантического контекста
(0008,0060) Modality	Метод контроля	Категориальный признак для учета различий в способе контроля при обучении
(0010,0010) Component Name	Наименование объекта контроля	Применяется для учета типа изделия и формирования контекстных признаков
(0010,2160) Material Name	Материал объекта	Категориальный признак, влияющий на визуальные характеристики дефектов
(0014,0030) Material Thickness	Толщина материала	Числовой признак для учета влияния геометрии на проявления дефектов

Продолжение таблицы 2.1

<b>Тег</b>	<b>Описание</b>	<b>Потенциальное применение при обучении ИИ</b>
(0014,0032) Material Pipe Diameter	Диаметр трубы	Применяется для нормализации геометрических характеристик и учета при визуальном анализе
(0014,0034) Material Isolation Diameter	Диаметр изоляции	Применяется для учета дополнительных конструктивных особенностей объекта
(0014,0042) Material Grade	Марка материала	Категориальный признак для уточнения физико-механических свойств объекта
(0018,1155) Radiation Setting	Параметры излучения	Применяется для учета условий съемки и повышения устойчивости модели
(0014,201A) Indication Type	Тип дефекта	Применяется как целевая переменная при обучении модели

Продолжение таблицы 2.1

Тег	Описание	Потенциальное применение при обучении ИИ
(0014,3022) Sensor Name	Наименование датчика	Применяется для повышения инвариантности к различным типам оборудования
(0014,3028) Sensor Temperature	Температура датчика	Числовой признак, влияющий на качества сигнала и уровень шума

Поскольку атрибуты метаданных DICOM-подобных форматов характеризуются различными физическими размерностями, существенно различающимися диапазонами значений, а также наличием возможной вариативности между сериями измерений, комплексное их использование с визуальными признаками в модели ИНС приводит к численному доминированию отдельных весовых коэффициентов и ухудшению обобщения информации, а также выходных результатов поиска. Поэтому для всех числовых атрибутов вводится процедура стандартизации. На рисунке 9 представлена схема токенизации, демонстрирующая последовательность преобразований исходных метаданных в числовой вектор  $z_i$ , используемый в дальнейшем для извлечения и комплексированного поиска схожих изображений с учетом вектора визуальных признаков.

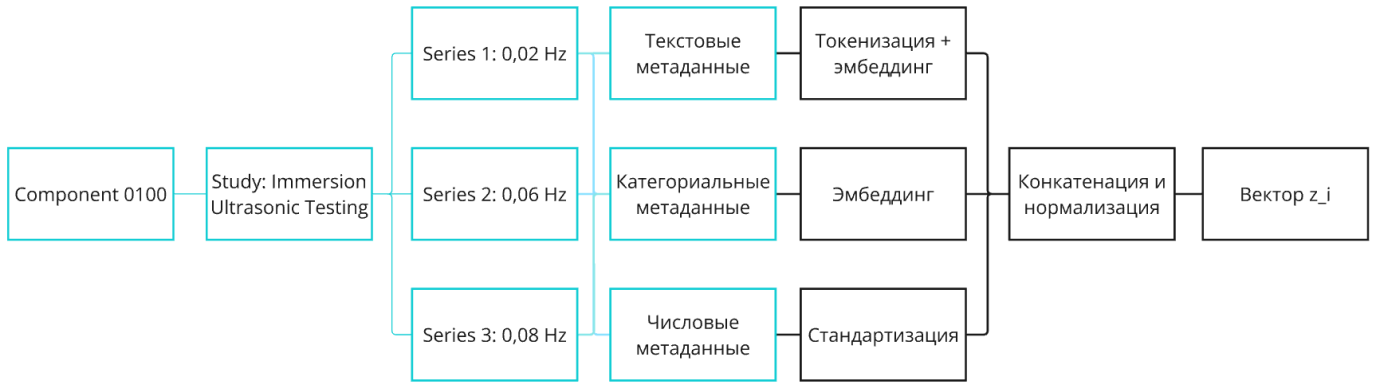


Рис. 9. Схема токенизации метаданных

Таким образом, данные DICOM-подобного формата формализуются в виде объектов, объединяющих многоканальные визуальные тензоры и неоднородные фактографические признаки. Дальнейшая задача заключается в приведении метаданных к унифицированному числовому представлению, допускающему совместную обработку с визуальными частью в рамках нейросетевой архитектуры.

## 2.2 Токенизация и стандартизация метаданных

Пусть  $m_{ij} \in M_{num}$  –  $j$ -й числовой атрибут  $i$ -го изображения. Для каждого атрибута в выборке вычисляются:

$$\bar{m}_j = \frac{1}{N} \sum_{i=1}^N m_{ij}, \quad \sigma_j = \sqrt{\frac{1}{N} \sum_{i=1}^N (m_{ij} - \bar{m}_j)^2}, \quad (2.6)$$

Показатели  $\bar{m}_j$  и  $\sigma_j$  фиксируются на обучающей выборке и не обучаются на этапе эксплуатации системы.

Нормализованное значение атрибута определяется как:

$$\tilde{m}_{ij} = \frac{m_{ij} - \bar{m}_j}{\sigma_j}, \quad (2.7)$$

где  $\bar{m}_j$  – математическое ожидание  $j$ -го параметра,  $\sigma_j$  – среднее квадратичное отклонение,  $\tilde{m}_{ij}$  – стандартизированное значение [48, 49].

Введение стандартизации обеспечивает:

- численную устойчивость при комплексной обработке визуальных тензоров;
- инвариантность модели к абсолютным масштабам физических параметров;
- корректную работу AutoML-механизма [50, 51] при сравнении архитектур-кандидатов.

Важно отметить, что использование min-max нормализации в рамках описываемой задачи менее предпочтительно, так как диапазоны физических параметров обладают вариативностью при расширении БД, что нарушает воспроизводимость модели.

Пусть  $\psi(\cdot)$  – функция токенизации, отображающая отдельный атрибут  $m_{ij}$  в вектор признаков фиксированной или переменной размерности и описываемая следующим образом:

$$\psi(M_{ij}) = \begin{cases} \frac{m_{ij} - \bar{m}_j}{\sigma_j}, & m_{ij} \in \mathbb{R} \\ \text{Embedding}(m_{ij}), & m_{ij} \in E \end{cases} \quad (2.8)$$

Формализация метаданных описывается отображением:

$$\Psi : M_i \rightarrow \mathbb{R}^{d_i}, \quad (2.9)$$

Общая функция токенизации определяется как:

$$z_i = \Psi(M_i) = \bigoplus_{j=1}^{K_i} \psi_j(m_{ij}), \quad (2.10)$$

где  $\bigoplus$  – операция упорядоченной конкатенации с применением нормализации, описываемая в рамках итогового представления как:

$$z_i = \omega \left( \begin{bmatrix} M_{num_i} \\ M_{cat_i} \\ M_{text_i} \end{bmatrix} \right), \quad (2.11)$$

где  $\omega$  – операция векторной нормализации (например, L2-нормализация).

Тогда вектор токенизированных метаданных определяется как:

$$z_i = \psi(M_i) = \text{Concat}(\psi(m_{i1}), \psi(m_{i2}), \dots, \psi(m_{iK_i})), \quad (2.12)$$

где операция конкатенации формирует единое представление массива фактографических признаков изображения с помощью объединения нескольких тензоров по размерности каналов и описываемая как:

$$X_{concat} = \text{Concat}(X_1, X_2, \dots, X_n) \rightarrow [C_1 + C_2 + \dots + C_n], \quad (2.13)$$

в отличие от классических подходов, где фактографические данные используются исключительно для предварительной фильтрации, в данной работе вектор  $z_i$  рассматривается как итоговый входной сигнал для модели ИНС, участвующий в формировании визуальных признаков.

Размерность итогового токена вычисляется как сумма размерностей отдельных токенов:

$$d_i = \sum_{j=1}^{K_i} d_j,$$

где  $d_j$  – размерность токена  $j$ -го атрибута. Позволяет зафиксировать размерность  $z_i$  для последующей интеграции с визуальными признаками в модели ИНС.

Число атрибутов  $K_i$  не является фиксированным и может варьироваться в зависимости от типа и производителя оборудования, используемой методики получения изображений, а также конфигурации модуля регистрации данных [52].

Рассмотрим изображение  $i$  с набором атрибутов  $M_i = \{m_{i1}, m_{i2}, m_{i3}\}$ , где  $m_{i1}$  – “slag” – категориальный атрибут,  $m_{i2} = 12.5$  – числовой атрибут температуры поверхности  $m_{i3}$  – “СТ” – текстовый атрибут, отвечающий за тип исследования.

Фиксируются параметры стандартизации обучающей выборки на значениях:  $m_{i2} = 15$ ,  $\sigma_2 = 4.2$ . Для категориального атрибута задается словарь  $E_{defect} = \{\text{"inclusion"}, \text{"slag"}, \text{"crease"}, \text{"oil\_spot"}\}$  с однопроходным кодированием размерности  $d=4$ . Тогда токенизация отдельных атрибутов рассчитывается следующим образом:

$$\psi_1(\text{"slag"}) = [0, 1, 0, 0] \in \mathbb{R}^4, \quad (2.14)$$

$$\psi_2(12.5) = \frac{12.5-15}{4.2} = -0.595 \in \mathbb{R}^1, \quad (2.15)$$

$$\psi_3("CT") = [0,0,1] \in \mathbb{R}^3 \quad (2.16)$$

Итоговый вектор метаданных до нормализации:

$$z'_i = \text{Concat}([0,1,0,0], [-0.595], [0,0,1]) = [0,1,0,0, -0.595,0,0,1] \in \mathbb{R}^8 \quad (2.17)$$

После применения L2-нормализации  $\omega(\cdot)$ :

$$z_i = \frac{z'_i}{\|z'_i\|_2} = [0.354,0.354,0,0, -0.211,0,0,0.354] \quad (2.18)$$

Введение процедуры стандартизации и токенизации метаданных позволяет перейти от разнородного фактографического описания изображений к унифицированному числовому пространству признаков, что позволяет осуществить совместную обработку с визуальными признаками в рамках модели ИНС и обеспечивает устойчивость результатов поиска при масштабировании БД и изменении диапазонов физических параметров объекта.

### 2.3 Формирование визуальных признаков и архитектура «DICONDE-YOLO»

Для обеспечения комплексной обработки визуальных и фактографических данных вводится операция токенизации метаданных, проецирующая элементы множества  $M_i$  в числовое представление.

Задача поиска схожих изображений формализуется следующим образом. Пусть задан запрос:

$$D_q = (X_q, M_q), \quad (2.19)$$

соответствующий изображению, для которого требуется найти визуально и контекстно близкие аналоги. Тогда требуется определить подмножество:

$$\mathcal{R}_K \subset \mathcal{D}, |\mathcal{R}_K| = K, \quad (2.20)$$

такое, что элементы  $D_i \in \mathcal{R}_K$  максимизируют меру сходства с запросом  $D_q$  при одновременном минимуме времени обработки запроса.

Для этого вводится отображение:

$$\Phi : (X_i, z_i) \rightarrow f_i, \quad (2.21)$$

где  $f_i \in \mathbb{R}^d$  – вектор визуальных признаков изображения, формируемый СНС. Вектор признаков  $f_i$  включает как информацию о пространственной структуре изображения, так и контекст, связанный с условиями его получения.

Для удобства последующего описания алгоритмов и формализации поиска схожих изображений в данной главе собраны все основные переменные и обозначения в таблице 2.2.

Таблица 2.2 – Описание используемых переменных

Переменная	Тип	Размерность / диапазон	Описание
D	Множество	$\{D_1, D_2, \dots, D_N\}$	База данных изображений
$D_i$	Пара (тензор, метаданные)	$(X_i, M_i)$	i-ое изображение базы данных
$X_i$	Тензор	$\mathbb{R}^{H_i \times W_i \times C_i}$	Визуальные данные i-го изображения
$M_i$	Множество	$\{m_{i1}, \dots, m_{iK_i}\}$	Фактографические данные i-го изображения
$z_i$	Вектор	$\mathbb{R}^{d_i}$	Токенизированное представление метаданных $M_i$

Продолжение таблицы 2.2

Переменная	Тип	Размерность / диапазон	Описание
$f_i$	Вектор	$\mathbb{R}^d$	Визуальные признаки, формируемые модифицированной ИНС
$D_q$	Пара (тензор, метаданные)	$(X_q, M_q)$	Запрос: изображение и метаданные
$X_q$	Тензор	$\mathbb{R}^{H_q \times W_q \times C_q}$	Визуальные данные запроса
$M_q$	Множество	$\{m_{q1}, \dots, m_{qK_q}\}$	Фактографические данные запроса
$z_q$	Вектор	$\mathbb{R}^{d_q}$	Токенизированное представление метаданных запроса
$f_q$	Вектор	$\mathbb{R}^d$	Вектор признаков запроса
$R_K$	Множество	Размер K	Подмножество K наиболее релевантных изображений
$s(f_q, f_i)$	Функция	—	Метрика сходства в пространстве признаков
$A^*$	Агент архитектуры	—	Выбранная конфигурация нейросети (AutoML)
$N_s$	Целое	—	Число серверов для горизонтального масштабирования
$L_j$	Действительно е	—	Текущая нагрузка сервера j

Продолжение таблицы 2.2

Переменная	Тип	Размерность / диапазон	Описание
$P(s_j)$	Действительное	[0, 1]	Вероятность назначения изображения серверу $j$

Визуальные признаки формируются с помощью модифицированной архитектуры ИНС YOLOv8 [53] и описываются следующим образом:

$$f_i = \Phi_{\text{YOLO}}(X_i, z_i, A^*), \quad (2.22)$$

где  $A^*$  – конфигурация архитектуры, выбранная на основе AutoML подхода,  $\Phi_{\text{YOLO}}$  – модифицированная архитектура СНС.

Рассматривая ключевые архитектурные изменения оригинальной ИНС YOLO, следует выделить замену блока SPPF [54] на расширенную модификацию ASPP [55]. Блок SPPF предназначен для агрегирования информации с нескольких пространственных масштабов без фиксации размера входного изображения. Его работа заключается в последовательном применении операции MaxPooling [56] с различными вариантами ядер к входному тензору признаков.

Пусть входной тензор признаков равен:

$$X \in \mathbb{R}^{H \times W \times C}, \quad (2.23)$$

где  $H, W, C$  – высота, ширина и число каналов тензора признаков

SPPF применяет три последовательных max-pooling с ядрами  $k = 5, s = 1$  и конкатенирует результаты с исходным тензором:

$$P_1 = \text{MaxPool}(X, k = 5, s = 1), \quad (2.24)$$

$$P_2 = \text{MaxPool}(P_1, k = 5, s = 1), \quad (2.25)$$

$$P_3 = \text{MaxPool}(P_2, k = 5, s = 1), \quad (2.26)$$

$$X_{\text{SPPF}} = \text{Concat}(X, P_1, P_2, P_3), \quad (2.27)$$

где MaxPool – операция конкатенации по каналной размерности.

Таким образом, SPPF позволяет объединить локальные и глобальные пространственные признаки, сохраняя вычислительную эффективность за счет последовательно применения операции max-pooling с небольшими ядрами.

ASPP расширяет концепцию SPPF за счет применения расширенных сверток, позволяющих получать более широкий контекст без увеличения числа параметров. Основное отличие данной модификации в использовании последовательных пропусков, что увеличивает рецептивное поле. В отличие от SPPF, ASPP формирует неоднородное рецептивное поле, адаптивное к масштабу объектов.

Пусть входной тензор остается неизменным. Тогда для ASPP применяется несколько операций свертки с различными коэффициентами расширения  $r_j$ :

$$F_j = \sigma \left( \text{Conv}(X; W_j, k = 3, s = 1, r_j) \right), \quad (2.28)$$

где  $W_j$  – веса свертки для  $j$ -го канала расширения,  $k$  – размер ядра свертки,  $s$  – шаг свертки,  $r_j$  – коэффициент расширения,  $\sigma$  – функция активации, Conv – операция свертки, применяющая фильтры, именуемые ядрами свертки к входному тензору для извлечения признаков (контуров, текстур, локальных паттернов) и формально описываемая как:

$$Y = \text{Conv}(X; W, k, s), \quad (2.29)$$

где  $X$  – входной тензор,  $W$  и  $k$  – фильтр размером  $k * k$  с шагом  $s$ ,  $Y$  – выходной тензор после проведения операции свертки.

Итоговая конкатенация результатов описывается следующим образом:

$$X_{\text{ASPP}} = \text{Concat}(F_1, F_2, \dots, F_n), \quad (2.30)$$

где  $n$  – число параллельных сверток с разными коэффициентами расширения.

Дополнительно включается операция получения глобального контекста GAP:

$$F_{\text{global}} = \text{Upsample} \left( \text{Conv}(\text{GAP}(X)) \right), \quad (2.31)$$

где Upsample [57] – операция интерполяции, увеличивающая пространственное разрешение тензора и формально описываемая как:

$$\text{Upsample}(X) \rightarrow X', \tag{2.32}$$

где тензор  $X'$  формируется с обновленным размером  $(H', W')$ .

И интегрируется в операцию конкатенации:

$$X_{\text{ASPP-final}} = \text{Concat}(F_1, \dots, F_n, F_{\text{global}}), \tag{2.33}$$

Использование глубинных и канальных механизмов внимания позволяет компенсировать увеличение контекстного окна ASPP за счет селективного усиления информативных каналов, что предотвращает деградацию локальных признаков.

Наглядное описание структуры и сравнение ключевых изменений ASPP по сравнению с SPPF представлено на рисунке 10 и в таблице 2.3.

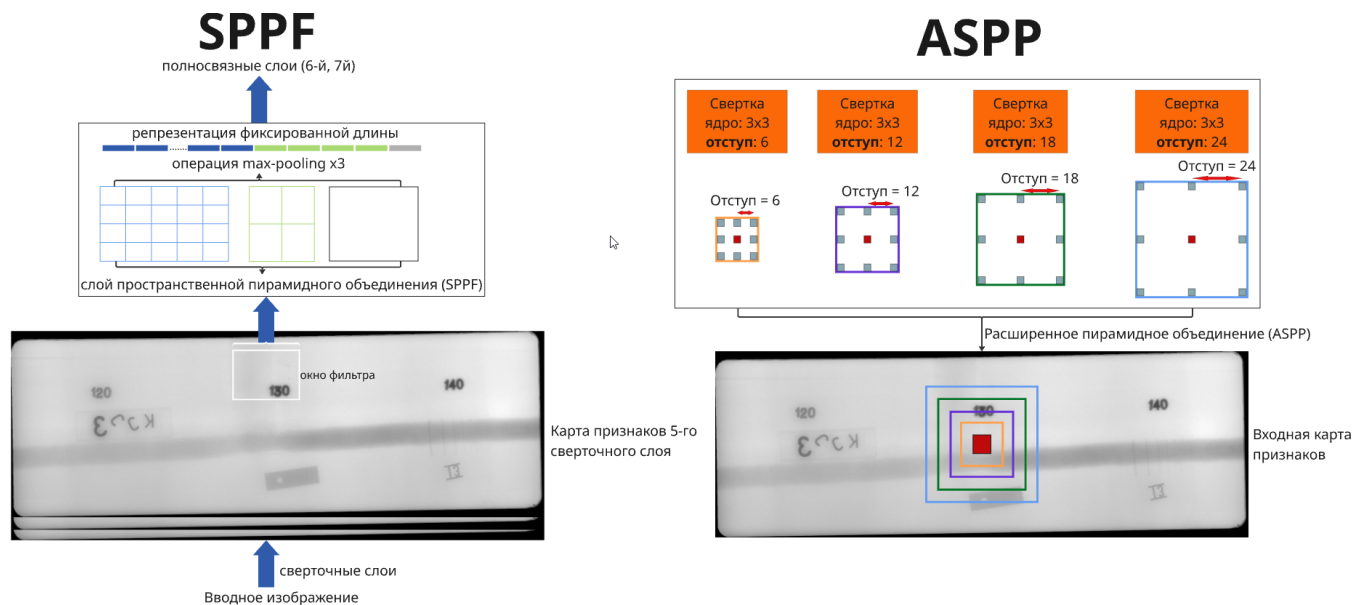


Рис. 10. Описание структуры SPPF и ASPP блоков

Таблица 2.3 – Сравнение ключевых характеристик SPPF и ASPP блоков

Характеристика	SPPF	ASPP
Основной механизм	Max-pooling	Расширенные свертки (dilated conv)

Продолжение таблицы 2.3

Характеристика	SPPF	ASPP
Контекст	Локальный + средний	Локальный + широкий + глобальный
Параметры	Отсутствуют	Свертки с обучаемыми весами
Эффективность	Быстрота операций, малое количество параметров	Более информативное представление, больше параметров, сохранение вычислительной сложности
Применение	Локально-ориентированная агрегация параметров	Расширенный контекст для малых объектов

Интеграция DC-блока [58, 59] глубинной свертки, разделяющийся на две параллельные ветви, а также варианта с добавлением параллельной операции MaxPool [60], описываемые следующим образом:

$$F_1 = \text{Conv}_{3 \times 3, s=1} \left( \text{Conv}_{1 \times 1, s=1} (X_{in}) \right), \quad (2.34)$$

$$F_2 = \text{DWConv}_{3 \times 3, s=1} \left( \text{Conv}_{1 \times 1, s=1} (X_{in}) \right), \quad (2.35)$$

$$X_{DC} = \text{Conv}_{1 \times 1, s=1} \left( \text{Concat}(F_1, F_2) \right), \quad (2.36)$$

$$FF_1 = \text{Conv}_{3 \times 3, s=2} \left( \text{Conv}_{3 \times 3, s=1} (X_{in}) \right), \quad (2.37)$$

$$FF_2 = \text{DWConv}_{3 \times 3, s=2} \left( \text{Conv}_{3 \times 3, s=1} (X_{in}) \right), \quad (2.38)$$

$$X_{MDC} = \text{Conv}_{1 \times 1, s=1} \left( \text{Concat}(FF_1, FF_2) \right), \quad (2.39)$$

Интеграция модуля ECA [61], описываемого следующим образом при учете сохранения математического представления входного тензора:

$$z = \text{GAP}(X_{in}) \in \mathbb{R}^C, X_{in} = \mathbb{R}^{H \times W \times C}, \quad (2.40)$$

Тогда адаптивный сверточный блок описывается как:

$$s = \sigma(\text{Conv1d}(z, c, k)), \quad (2.41)$$

$$X_{out} = X_{in} \odot s, \quad (2.42)$$

где  $c$  – число каналов,  $k$  – адаптивный размер ядра,  $\odot$  – скалярное произведение каналов (векторов), Conv1d – свертка с ядром  $1 \times 1$ .

Обобщенная схема объединения визуальных и фактографических признаков представлена на рисунке 11.



Рис. 11. Обобщенная схема объединения визуальных и фактографических признаков

Мера сходства между запросом и элементами БД определяется как функция:

$$\text{Sim}(D_q, D_i) = s(f_q, f_i), \quad (2.43)$$

где  $s(\cdot)$  – метрика в пространстве признаков:

$$s(f_q, f_i) = \begin{cases} 1 - \frac{f_q \cdot f_i}{\|f_q\| \|f_i\|} & \text{— косинусная метрика} \\ \|f_q - f_i\| & \text{— евклидова метрика} \end{cases} \quad (2.44)$$

Тогда задача поиска описывается как задача оптимизации:

$$\hat{i} = \underset{i \in \mathcal{D}'}{\operatorname{argmin}} s(D_q, D_i), \quad (2.45)$$

при ограничении на среднее время отклика системы:

$$T_{resp} \leq T_{avg\_max} \quad (2.46)$$

Для снижения вычислительной сложности и повышения масштабируемости системы поиска схожих изображений вводится предварительная фильтрация на основе фактографических данных.

Пусть  $\mathcal{F}(M_q)$  – оператор фильтрации, формирующий подмножество кандидатов:

$$\mathcal{D}' = \mathcal{F}(M_q) \subseteq \mathcal{D}, \quad (2.47)$$

включающее только те изображения, метаданные которых удовлетворяют заданным условиям сопоставимости. Поиск визуального сходства выполняется исключительно в пространстве  $\mathcal{D}'$ , что уменьшает асимптотическую сложность с  $O(N)$  до  $O(|\mathcal{D}'|)$ .

При учете горизонтального масштабирования системы балансировка нагрузки определяется следующим образом. Пусть имеется  $N_s$  серверов, каждая выборка  $F_i$  распределяется по серверу  $s_j$  с вероятностью:

$$P(s_j) = \frac{L_{avg} - L_j}{\sum_{k=1}^{N_s} (L_{avg} - L_k)}, \quad (2.48)$$

$$L_{avg} = \frac{1}{N_s} \sum_{j=1}^{N_s} L_j, \quad (2.49)$$

где  $L_j$  – текущая нагрузка сервера,  $L_{avg}$  – средняя нагрузка по всем серверам,  $N_s$  – общее количество серверов.

Тогда целевая функция описывается как:

$$\max Q_{\Phi, \theta} = \alpha \cdot \operatorname{Rel}(R_K, D_q) - \beta \cdot T_{resp},$$

где  $\operatorname{Rel}$  – релевантность,  $T_{resp}$  – время отклика,  $\alpha, \beta$  – веса

Иллюстрация оригинальной архитектуры YOLOv8 представлена на рисунке 12. Для обеспечения ясности, модифицированная архитектура YOLOv8 именуется «DICONDE-YOLO».

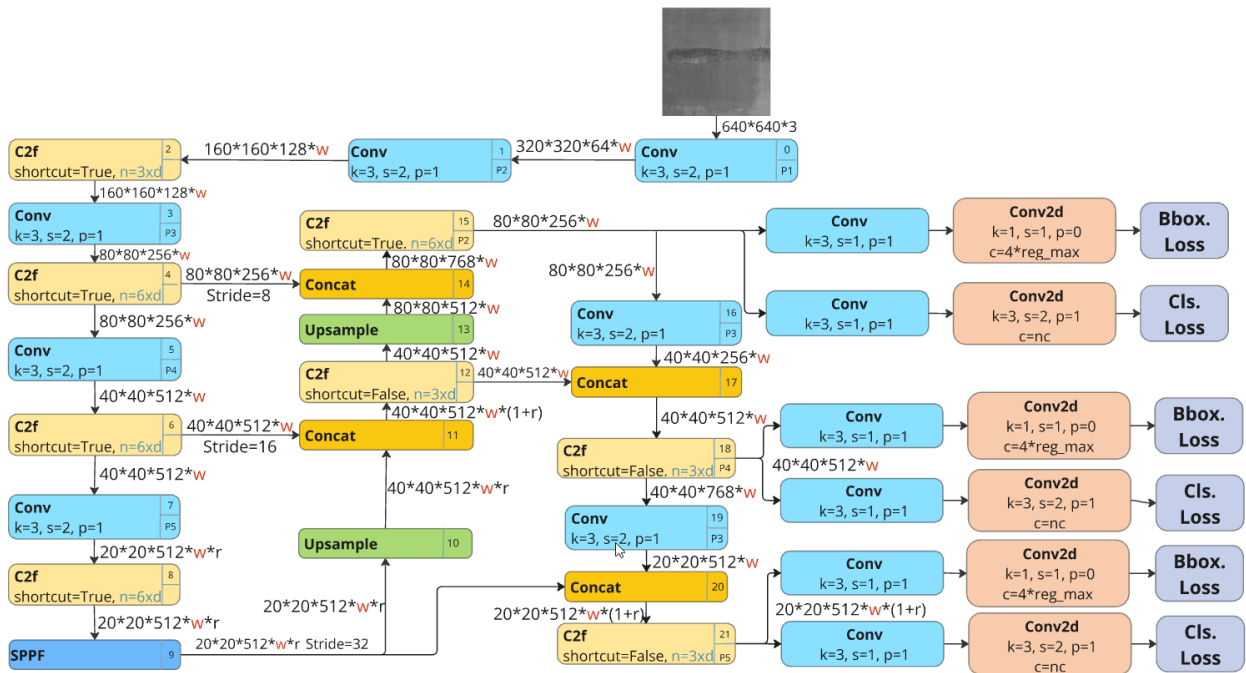


Рис. 12. Структура оригинальной архитектуры YOLOv8

Модифицированная архитектура «DICONDE-YOLO» представлена на рисунке 13.

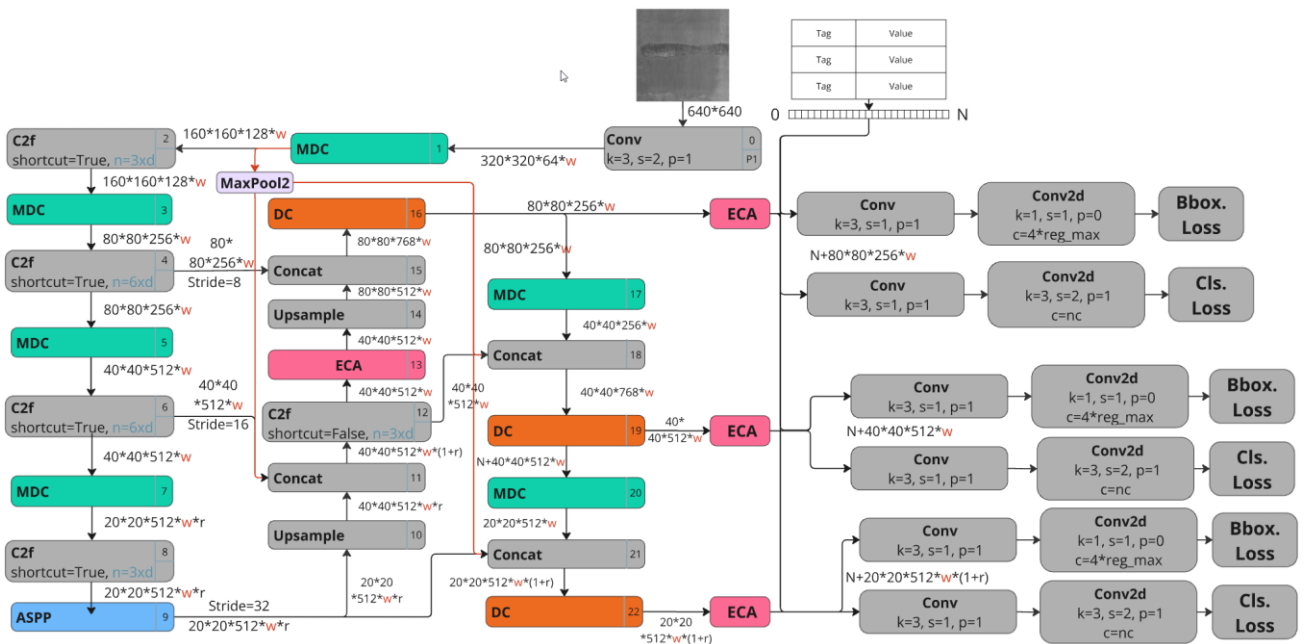


Рис. 13. Структура модифицированной архитектуры «DICONDE-YOLO»

Расширенное описание структурных блоков архитектуры «DICONDE-YOLO» представлено на рисунке 14.

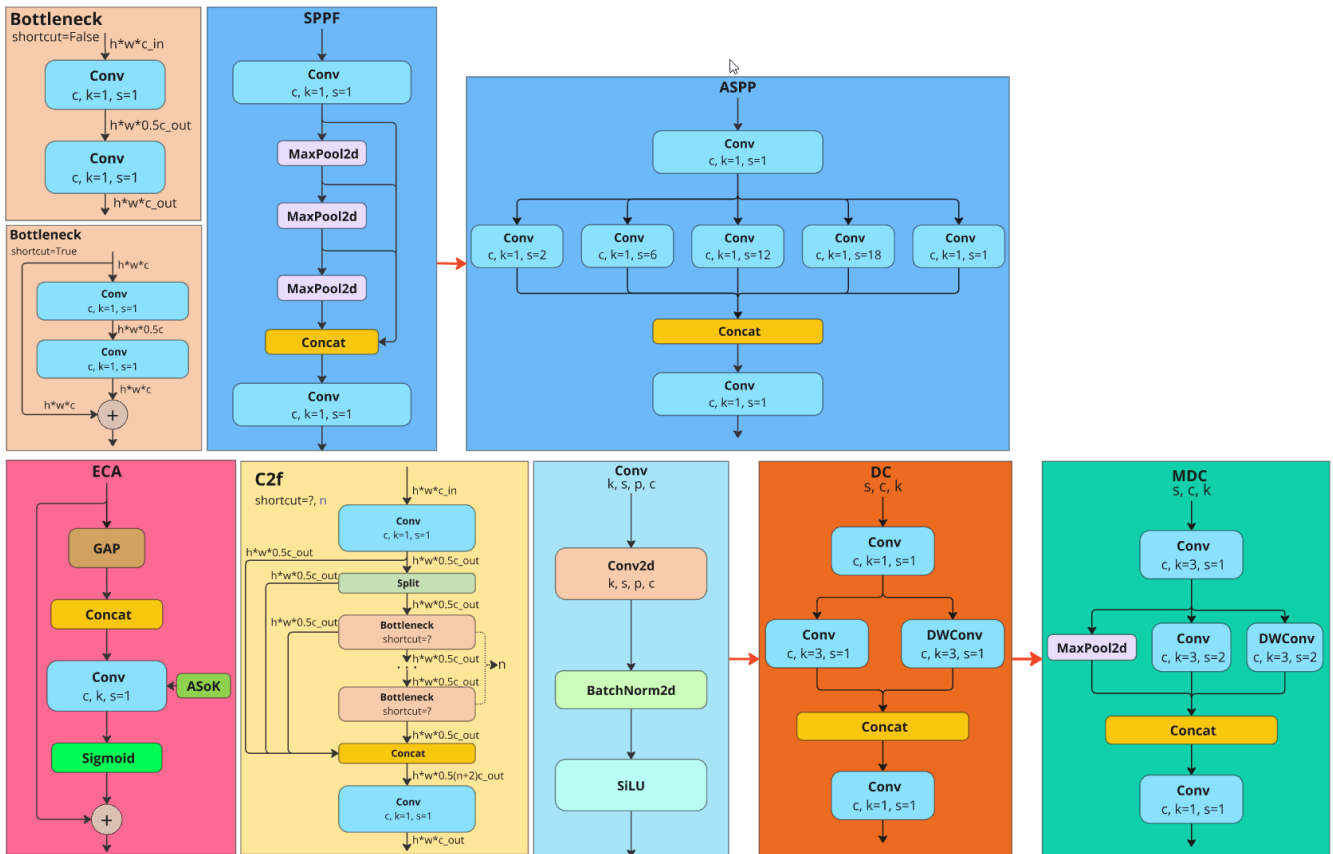


Рис. 14. Структурные составляющие оригинальных блоков и внедренных модификаций

Таким образом, «DICONDE-YOLO» обеспечивает извлечение устойчивых и информативных визуальных признаков, учитывающих как локальную структуру изображений, так и широкий пространственный контекст. Интеграция DC/MDC и ECA-модулей позволяет усилить значимые каналы и локальные признаки, предотвращая деградацию информации при увеличении рецептивного поля, что критически важно для последующего поиска схожих изображений с учетом визуальных и фактографических данных.

Проведенная формализация структуры данных, пространства признаков и архитектуры модели ИНС позволяет перейти от математической модели к алгоритмической реализации комплексированного поиска. Ввиду многоэтапного характера обработки данных, включающего токенизацию метаданных, предварительную фильтрацию, извлечение визуальных признаков, векторный поиск

и балансировку нагрузки — дальнейшее изложение целесообразно представить в виде псевдокода, отражающего последовательность выполнения операций и представленного в приложении А.

## **2.4 Выводы по главе**

В данной главе разработана формальная модель комплексированного алгоритма поиска схожих изображений, сочетающего визуальные и фактографические признаки в едином пространстве представлений. Выполнена формализация структуры данных, введен механизм токенизации и стандартизации метаданных, предложена модифицированная архитектура «DICONDE-YOLO» для извлечения признаков. Кроме того, в рамках формализации описаны механизмы предварительной фильтрации, векторного сопоставления и балансировки нагрузки, обеспечивающие масштабируемость и снижение времени отклика системы. Полученные результаты служат базой для проведения вычислительных экспериментов с целью оценки эффективности предложенного подхода, представленных в следующей главе.

### **Глава 3. ВЫЧИСЛИТЕЛЬНЫЕ ЭКСПЕРИМЕНТЫ. ВЕРИФИКАЦИЯ МОДЕЛЕЙ И АЛГОРИТМОВ**

Экспериментальная часть диссертационной работы направлена на верификацию разработанной модели и комплексированного алгоритма поиска схожих изображений, ориентированного на применение в СППР. Необходимость проведения вычислительных экспериментов обусловлена тем, что предложенный алгоритм основан на интеграции разнородных источников информации — визуальных признаков изображений и фактографических метаданных — а также на использовании модифицированной архитектуры СНС «DICONDE-YOLO» для извлечения признаков. Подобное комплексирование алгоритмов требует экспериментального подтверждения своей практической состоятельности, устойчивости и вычислительной эффективности в условиях, приближенных к реальным сценариям эксплуатации.

Проверяется гипотеза о том, что совместное использование визуальных и фактографических признаков в едином пространстве представлений позволяет повысить обобщающую способность модели без существенного увеличения вычислительных затрат. Экспериментальные исследования направлены на подтверждение корректности формальной модели, разработанной во второй главе, а также на количественную оценку вклада отдельных архитектурных и алгоритмических компонент, включая модифицированную архитектуру извлечения визуальных признаков, механизмы токенизации и стандартизации метаданных и процедуры ранжирования результатов поиска.

### 3.1 Описание аппаратной платформы и воспроизводимости вычислительных экспериментов

Для обеспечения воспроизводимости результатов все вычислительные эксперименты проводились на фиксированной аппаратной и программной платформе. Конфигурация оборудования и программной среды приведена в таблице 3.1. Использование единой вычислительной платформы, а также фиксация случайного седа позволяют исключить влияние аппаратных факторов на результаты и обеспечить корректное сравнение различных архитектур и алгоритмов.

Таблица 3.1 – Аппаратная конфигурация платформы для экспериментов

Аппаратный компонент	Характеристика
CPU	AMD Ryzen 7 3700X, 8 ядер, 16 потоков, 4 ГГц
RAM	32 ГБ DDR4, 3800 MHz
GPU	NVIDIA RTX 3080 10 ГБ VRAM
CUDA / cuDNN	CUDA 11.8, cuDNN 8.6
ОС	Windows 10 23H2
Фреймворк	PyTorch 2.2.0

В рамках экспериментов использовался фреймворк PyTorch версии 2.2.0 с поддержкой CUDA 11.8 и cuDNN 8.6, что обеспечивает возможность обучения и инференс СНС. Ограничения по объему видеопамати графического ускорителя определили максимальный размер батча и размерность эмбедингов, используемых в экспериментах, что учитывается при анализе результатов. Все эксперименты проводились в идентичных условиях без изменения аппаратной конфигурации и версий программных библиотек. Это позволило исключить влияние стохастических факторов на результаты и обеспечить корректность сравнительного анализа архитектур и алгоритмических модификаций.

### 3.2 Экспериментальные данные и методика подготовки датасета

Экспериментальная проверка алгоритма выполнялась на специализированных датасетах общей численностью более 160 000 изображений, представленных в DICOM-подобных форматах и сопровождаемых структурированными метаданными.

Каждый объект датасета рассматривается как совокупность визуальной информации и фактографических признаков. Визуальные данные предварительно приводятся к унифицированному формату, пригодному для использования в качестве входного тензора для ИНС. На данном этапе выполняется нормализация интенсивностей, масштабирование пространственного разрешения и преобразование изображений в тензорное представление. Это позволяет снизить влияние различий в условиях получения изображений и обеспечить стабильность формирования визуальных эмбедингов.

Фактографические данные подвергаются процедурам стандартизации и токенизации. Числовые признаки нормализуются, а категориальные и дискретные параметры преобразуются в последовательности токенов с последующим отображением в векторное пространство фиксированной размерности. Такой подход обеспечивает возможность совместной обработки визуальных и фактографических признаков в рамках единого признакового пространства.

Для проведения вычислительных экспериментов датасеты были разделены на обучающую, валидационную и тестовую выборки в соотношении 70 % / 15 % / 15 %. Разделение выполнялось с учетом предотвращения пересечений по идентификаторам объектов и временным меткам, что исключало утечку информации между выборками. Тестовая выборка использовалась исключительно для итоговой оценки качества поиска и не участвовала в процессе обучения и настройки гиперпараметров.

Подобное распределение данных позволяет обучить модель на значительном объеме информации, одновременно сохраняя возможность валидации и тестирования

на независимых выборках. Основная часть данных использовалась для обучения модели закономерностям и зависимостям, представленным в датасете, а валидационная и тестовая выборки обеспечивают надежную оценку производительности на новых данных, что согласуется с классическим подходом, рассмотренным Жаворонковой О. Г. [62].

Основные характеристики используемых датасетов приведены в таблице 3.2, что позволяет оценить масштаб и полноту экспериментальных данных.

Таблица 3.2 – Характеристики датасетов, используемых в экспериментальных исследованиях

<b>Датасет</b>	<b>Кол-во изображений</b>	<b>Формат данных</b>	<b>Наличие метаданных</b>	<b>Тип метаданных</b>	<b>Размерность визуального эмбединга</b>
Dataset A	12 000	DICOM	Да	Числовые, категориальные	512
Dataset B	45 000	DICOM	Да	Числовые, категориальные, текстовые	512
Dataset C	110 000	DICOM	Частично	Числовые	512

Следует отметить, что используемые экспериментальные данные характеризуются умеренным дисбалансом классов, обусловленным неравномерным представлением отдельных типов дефектов и сценариев. Разметка эталонных множеств релевантных изображений выполнялась экспертным образом, что обеспечивает семантическую корректность оценки, но накладывает ограничения, связанные с субъективностью экспертных решений и трудоемкостью масштабирования разметки. Данные особенности учитываются при интерпретации экспериментальных результатов.

Процесс формирования эталонных множеств релевантных изображений включал анализ визуального сходства, а также сопоставление ключевых фактографических атрибутов. Экспертная разметка выполнялась с использованием специализированного ПО, обеспечивающего просмотр изображений и связанных метаданных, что позволяло минимизировать субъективные ошибки и повысить согласованность оценок. Свидетельство о регистрации продукта представлено в приложении Б. Процедура экспертной разметки была организована в виде формализованного итерационного цикла, обеспечивающего воспроизводимость и согласованность формирования эталонных множеств релевантных изображений. На первом этапе для каждого целевого изображения формировалось предварительное подмножество кандидатов на основе визуального сходства, вычисляемого в пространстве базовых эмбедингов. Это позволяло существенно сократить объем данных, подлежащих экспертной оценке, и сфокусировать внимание экспертов на потенциально релевантных примерах.

На втором этапе эксперты осуществляли визуальный анализ изображений-кандидатов с одновременным учетом сопутствующих фактографических данных. Решение о включении изображения в эталонное множество принималось на основе совокупной оценки визуального сходства и семантической согласованности ключевых параметров, что соответствовало целевому сценарию применения алгоритма в СППР.

На заключительном этапе результаты экспертной оценки агрегировались и использовались для уточнения эталонных множеств релевантных изображений с маркировкой конкретных типов дефектов, а также для формирования обучающих и тестовых выборок. Итерационный характер процедуры позволял корректировать критерии релевантности и повышать согласованность разметки при появлении новых данных или уточнении требований к задаче поиска. Такая организация цикла экспертной разметки обеспечивает баланс между семантической корректностью эталонов и практической реализуемостью процесса разметки при работе с

крупномасштабными специализированными датасетами. Схема процедуры экспертной разметки представлена на рисунке 15.

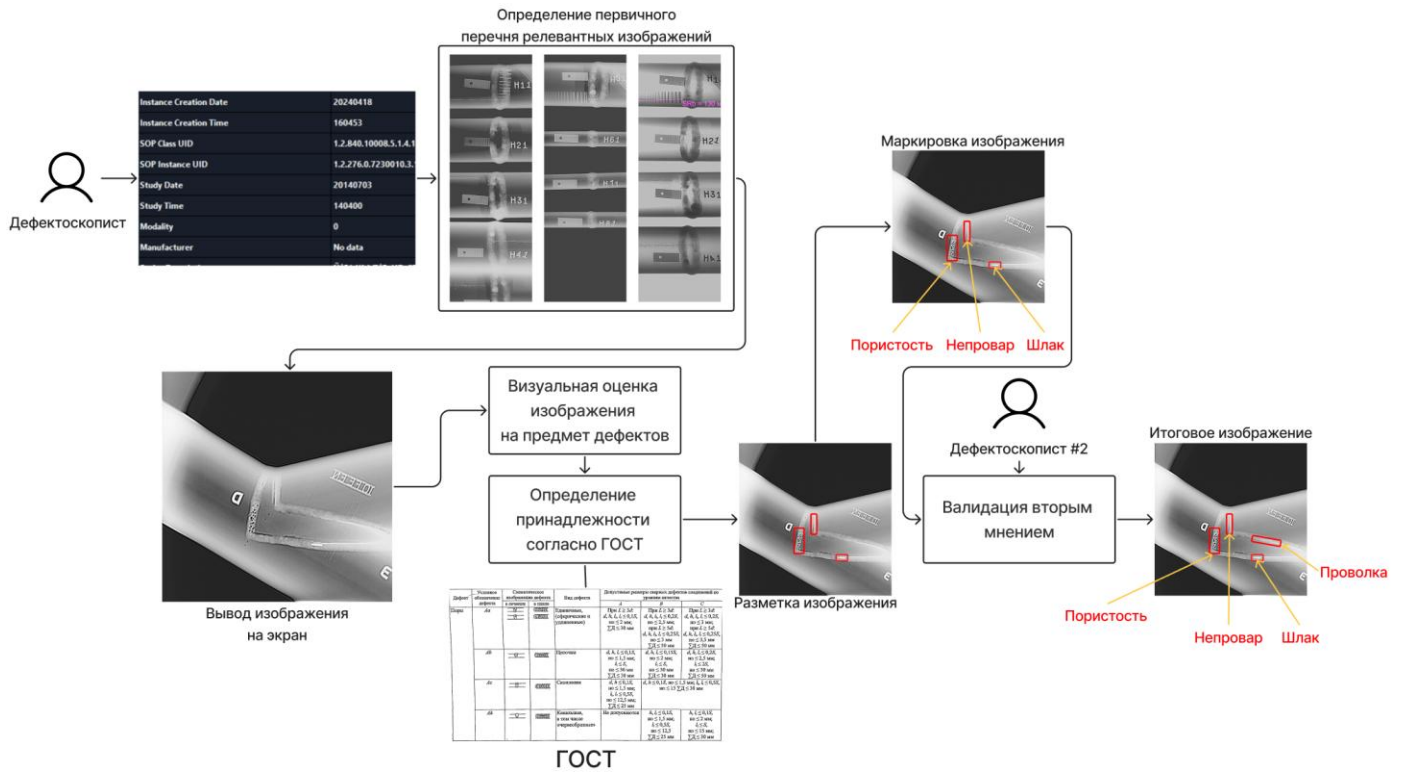


Рис. 15. Схема процедуры экспертной разметки

Интерфейс специализированного ПО для проведения разметки представлен на рисунке 16.

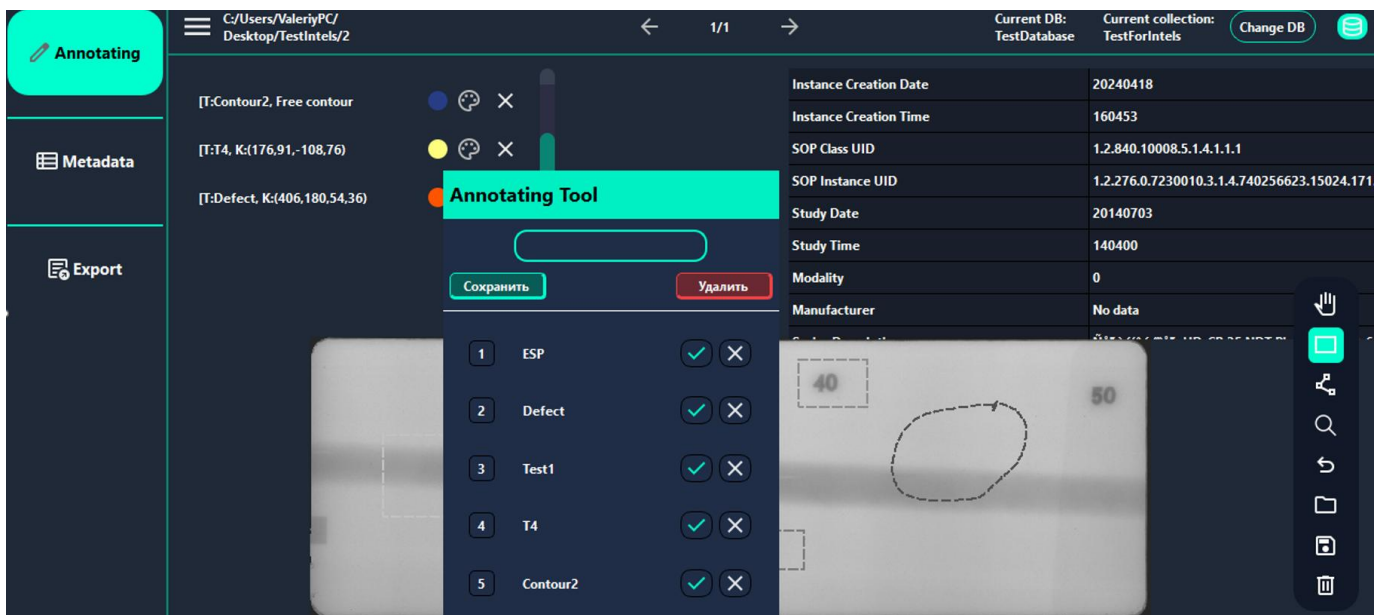


Рис. 16. Интерфейс специализированного ПО для разметки изображений

### 3.3 Сравнительный анализ архитектур YOLOv8-YOLOv12

Для обоснования выбора базовой архитектуры ИНС с целью дальнейшей модификации, был проведен сравнительный анализ нескольких передовых версий YOLO на момент проведения исследования, начиная с восьмой и заканчивая двенадцатой [53, 63-66]. Все версии обучались в идентичных условиях с фиксированными гиперпараметрами и конфигурацией аппаратного обеспечения, что обеспечивало корректность сравнений.

В качестве основных критериев сравнения использовались показатели Precision, Recall, mAP50, метрика локализации CIOU, а также время инференса и объем используемой видеопамяти. Использование метрики CIOU обусловлено спецификой решаемой задачи, поскольку качество локализации объектов напрямую влияет на формирование визуальных эмбеддингов. Более точное определение пространственных границ дефектов способствует снижению шума в признаковом представлении и, как следствие, повышению качества поиска схожих изображений в пространстве эмбеддингов. Сводные результаты сравнительного анализа представлены в таблице 3.3.

Таблица 3.3 – Сравнительный анализ версий YOLOv8-YOLOv12

<b>Версия YOLO</b>	<b>Precision</b>	<b>Recall</b>	<b>mAP50</b>	<b>CIOU</b>	<b>Время инференса, мс</b>	<b>VRAM, ГБ</b>
YOLOv8	0.63	0.65	0.67	0.58	36	6.4
YOLOv10	0.66	0.62	0.54	0.47	52	9.5
YOLOv11	0.66	0.65	0.65	0.57	48	8.9
YOLOv12	0.66	0.64	0.67	0.58	50	9.2

Следует отметить, что версия YOLOv9 не была включена в итоговое сравнительное исследование. Предварительные тесты показали отсутствие

статистически значимого улучшения показателей точности и обобщающей способности по сравнению с YOLOv8 при сопоставимых условиях обучения. При этом наблюдалось увеличение нестабильности обучения и отсутствие прироста эффективности, что согласуется с результатами, опубликованными в оригинальной работе авторов архитектуры. В связи с этим использование YOLOv9 не рассматривалось как целесообразное для дальнейшей модификации и интеграции в комплексированный алгоритм поиска изображений.

Динамика обучения модели YOLOv8, включая изменение функции потерь и значения mAP50 на обучающей и валидационной выборках представлен на рисунке 17. Представленные графики означают следующее:

`train/box_loss` – оценка потерь локализации на тренировочном датасете. Иллюстрирует точность положения и размер предсказанных ограничивающих рамок (далее – боксов) для объектов интереса. Чем меньше показатель, тем точнее модель определяет положение дефекта;

`train/cls_loss` – оценка потерь классификации на тренировочном датасете. Иллюстрирует точность классификации объектов внутри предсказанных боксов;

`train/df_l_loss` – оценка потерь фокального распределения на тренировочном датасете. Является вспомогательным показателем, применяемым для повышения точности локализации малых и пропорционально сложных объектов;

`val/box_loss`, `val/cls_loss`, `val/df_l_loss` – аналогичные показатели, применяемые на валидационном датасете;

`metrics/precision(B)` – точность для предсказанных ограничивающих рамок. Отражает долю предсказанных боксов, совпавших с реальными объектами;

`metrics/recall(B)` – полнота для боксов. Отражает долю корректно обнаруженных объектов;

`metrics/mAP50` – расчет основной метрики качества детекции. Отражает среднюю точность по всем классам при условии, что  $\text{IoU} \geq 0.5$ ;

metrics/mAP50-95(B) – расчет более строгой метрики, учитывающей насколько модель точно локализует объекты. Отражает устойчивость при разных показателях точности боксов. Расчет происходит на основании показателя mAP при IoU от 0.5 до 0.95 с шагом 0.05.

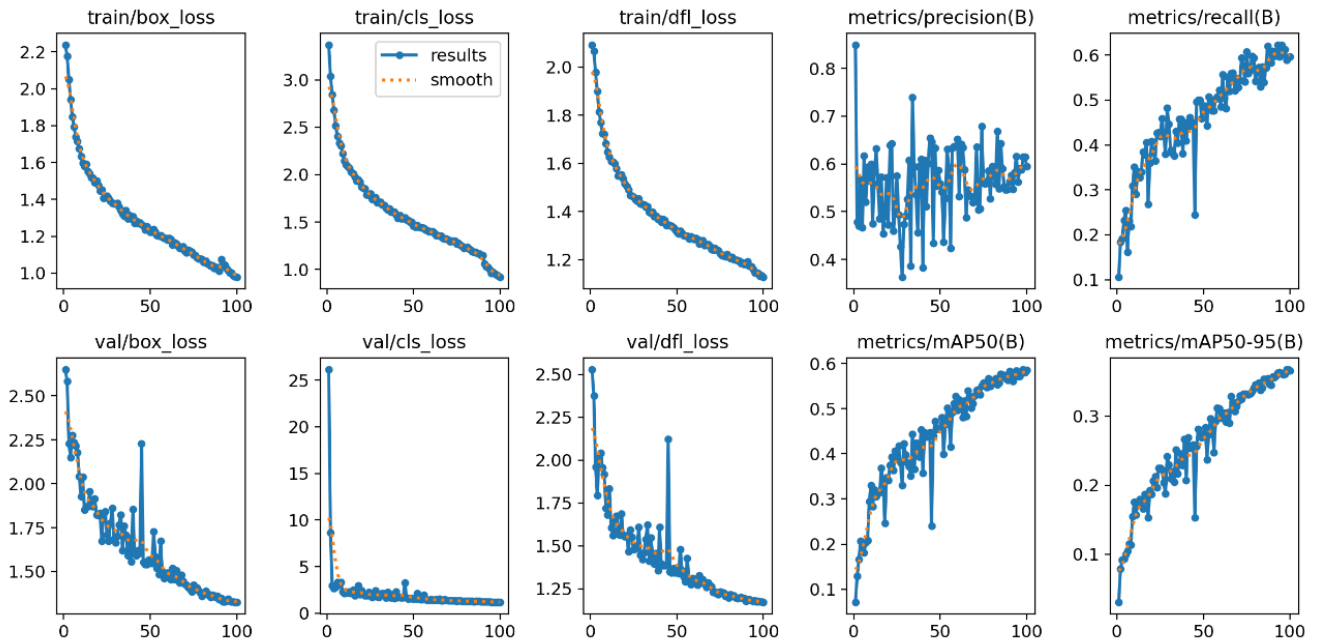


Рис. 17. Динамика обучения модели YOLOv8

Аналогичная динамика обучения по для модифицированной архитектуры «DICONDE-YOLO» представлена на рисунке 18.

Проиллюстрированные кривые подтверждают устойчивость процесса обучения и отсутствие выраженного переобучения для архитектуры YOLOv8. Кроме того, диаграммы сравнения моделей YOLO, представленные на рисунке 18, 19 иллюстрируют, что 8-я версия обеспечивает наилучшее сочетание точности детекции и качества локализации дефектов по сравнению с более поздними версиями при сопоставимых условиях обучения, а также обладает минимальными вычислительными затратами при сохранении высокой точности, что делает ее наиболее пригодной для использования в составе СППР.

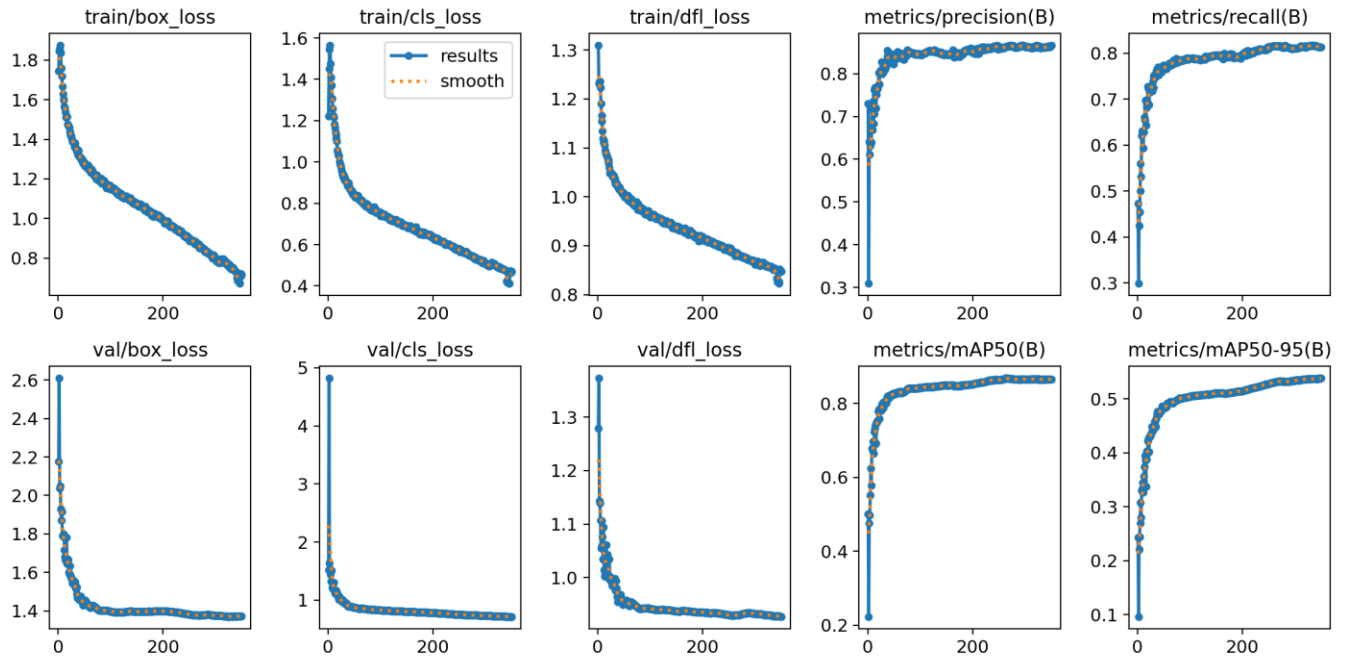


Рис. 18. Динамика обучения модели «DICONDE-YOLO»

Полученные результаты подтверждают, что увеличение глубины и вычислительной сложности архитектур YOLOv11 и YOLOv12 не приводит к пропорциональному росту качества визуальных эмбеддингов в условиях ограниченного объема специализированных данных. YOLOv10 была исключена из дальнейших исследований вследствие высокой вычислительной нагрузки и слабой обобщающей способности, тогда как YOLOv11 и YOLOv12 не продемонстрировали существенного прироста качества при значительном увеличении времени инференса.

На основе сравнительного анализа таблицы 3.3 и рисунков 17 и 18 делается вывод, что версия YOLOv8 представляет собой оптимальный баланс между обобщающей способностью и вычислительной эффективностью модели, что делает ее предпочтительной основой для дальнейшей архитектурной модификации в рамках комплексированного алгоритма поиска схожих изображений.

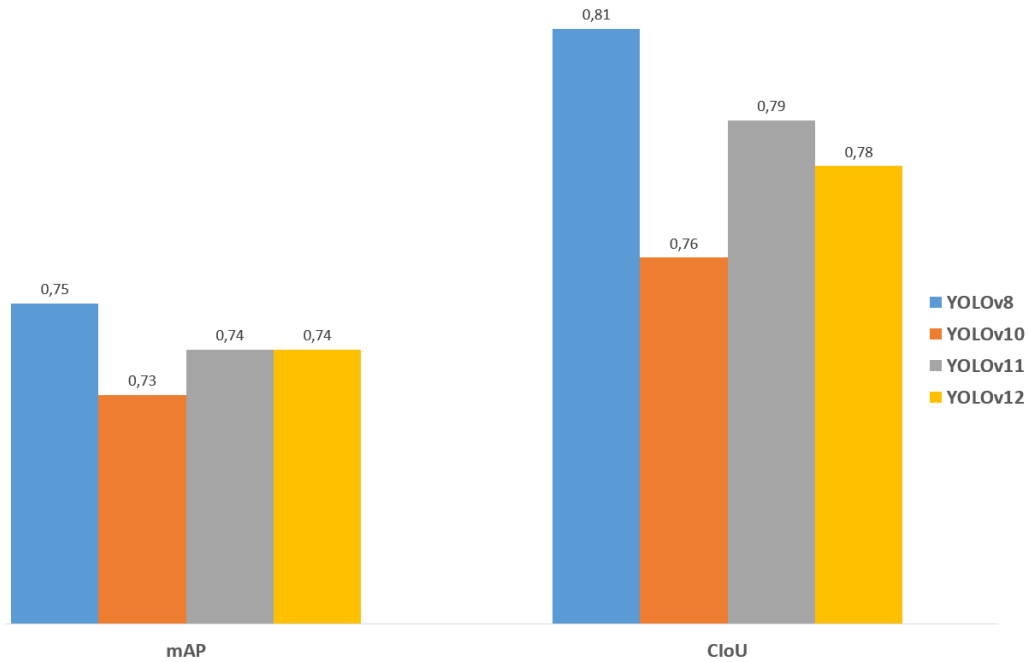


Рис. 19. Сравнение моделей YOLO по метрикам mAP и CIoU

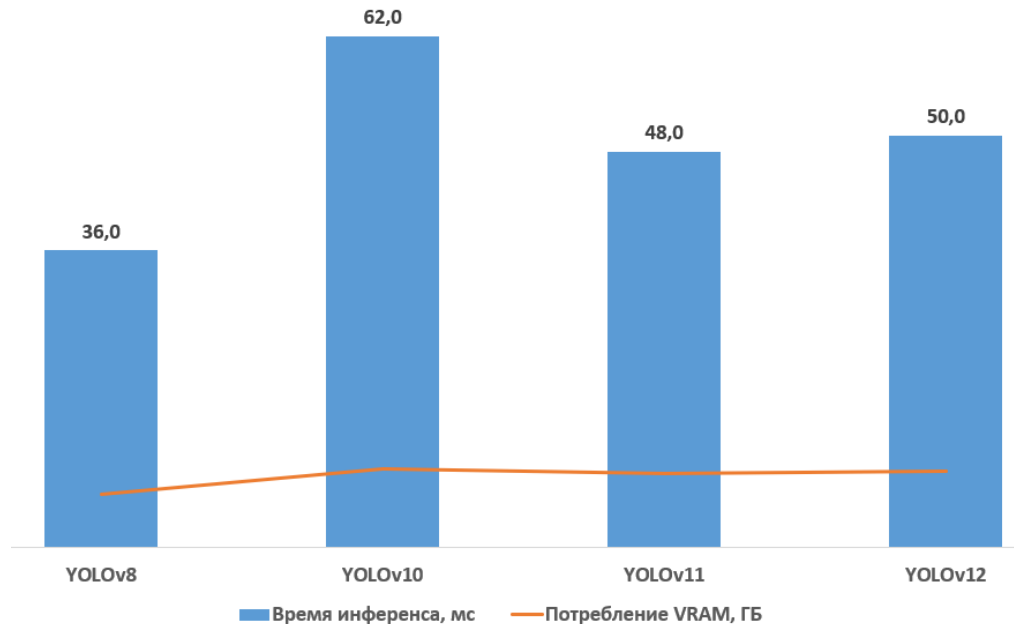


Рис. 20. Сравнение моделей YOLO по метрикам быстродействия и потребления VRAM

Для оценки вклада отдельных компонент модифицированной архитектуры «DICONDE-YOLO» были проведены абляционные эксперименты, в рамках которых поочередно исключались блок ASPP, модули DC и MDC, механизм эффективного канального внимания ECA, а также эмбединги метаданных.

Оценка качества поиска изображений выполняется с использованием метрик Precision, Recall и mAP50, отражающих как точность, так и полноту выдачи. Дополнительно используется метрика NDCG, позволяющая учитывать качество ранжирования результатов с учетом их позиций в списке выдачи. Для оценки точности локализации дефектов применяется метрика CIoU, что особенно важно в контексте использования модифицированной архитектуры «DICONDE-YOLO».

Для анализа вычислительных характеристик используются показатели времени обработки запроса, латентности поиска ближайших соседей и потребления памяти. Совокупность указанных метрик позволяет провести оценку эффективности алгоритма в условиях практического применения. Результаты абляционных экспериментов представлены в таблице 3.4.

Таблица 3.4 – Результаты абляционных экспериментов

<b>Конфигурация модели</b>	<b>Визуальные признаки</b>	<b>Метаданные</b>	<b>Токенизация метаданных</b>	<b>mAP50</b>	<b>NDCG</b>
Полная модель	+	+	+	0.822	0.893
Полная модель	+	+	-	0.756	0.792
Полная модель	+	-	+	0.781	0.824
Полная модель	+	-	-	0.724	0.742
Без ASPP	+	+	+	0.787	0.841
Без ECA	+	+	+	0.771	0.826
Без DC/MDC	+	+	+	0.801	0.858

Анализ абляционных экспериментов подтверждает, что наибольший вклад в повышение качества ранжирования вносит интеграция фактографических метаданных и их токенизация, что отражается в существенном росте метрики NDCG. Модули канального внимания ECA оказывают стабилизирующее влияние на формирование визуальных эмбедингов, тогда как блоки DC и MDC в большей степени влияют на устойчивость модели к вариативности входных данных. Для

оценки устойчивости полученных результатов обучение и тестирование моделей выполнялись в нескольких независимых запусках с фиксированными, но различающимися начальными значениями случайного седа. Разброс значений основных метрик качества не превышал 2–3 %, что свидетельствует о статистической стабильности наблюдаемых эффектов и воспроизводимости выводов абляционного анализа. Полученные результаты согласуются с теоретическими положениями, изложенными во второй главе, и экспериментально подтверждают целесообразность предложенной архитектурной модификации.

На рисунке 21 приведен пример выявления дефектов, характеризующихся низкой визуальной контрастностью и высокой внутриклассовой вариативностью. В подобных случаях использование только визуальных признаков приводит к снижению качества поиска, тогда как интеграция фактографических данных позволяет повысить устойчивость и семантическую корректность результатов. Кроме того, была произведена оценка влияния размерности эмбедингов токенизированных данных на качество поиска изображений. Полученные результаты отражены в таблице 3.5.

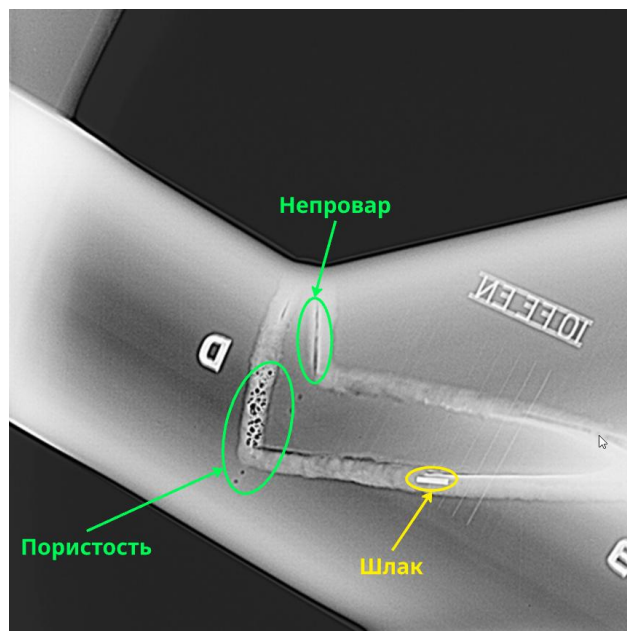


Рис. 21. Пример выявления дефектов с низкой визуальной контрастностью и высокой внутриклассовой вариативностью

Таблица 3.5 – Влияние размерности эмбедингов на качество поиска

<b>Размерность эмбединга</b>	<b>Precision</b>	<b>Recall</b>	<b>mAP50</b>	<b>Среднее время запроса, мс</b>
128	0.74	0.68	0.61	18
256	0.79	0.73	0.68	24
512	0.83	0.78	0.74	37
1024	0.84	0.79	0.75	65

Анализ влияния размерности эмбедингов показывает, что увеличение численного значения сверх 512 приводит к незначительному приросту качества поиска при существенном росте вычислительных затрат. Таким образом, размерность 512 была выбрана в качестве рационального компромисса между точностью и производительностью.

Помимо интегральных метрик точности, был выполнен анализ матриц путаницы, построенных для базовых архитектур YOLOv8–YOLOv12, а также для предложенной модифицированной модели «DICONDE-YOLO» при обучении на тестовой выборке. Матрицы ошибок позволяют выявить характерные классы ошибок детекции, которые не всегда очевидны при рассмотрении агрегированных показателей mAP или CIoU, и тем самым служат инструментом качественной интерпретации полученных результатов. Качестве классов использовались дефекты, представленные в таблице 3.6.

Таблица 3.6 – Классы дефектов и их описание

<b>Класс</b>	<b>Описание</b>
crease	склада
crescent_gap	серповидный зазор
inclusion	включение
oil_spot	масляное пятно
punching_hole	дефект штамповки

Продолжение таблицы 3.6

Класс	Описание
rolled_pit	прокатная выемка
silk_spot	волосовина
waist_folding	клиновидный дефект
water_spot	водяное пятно
welding_line	линия сварного шва

Анализ матриц путаницы для YOLOv8–YOLOv12, представленных на рисунках 22 и 23, демонстрирует, что при ограниченной репрезентативности обучающих данных наблюдается систематическое смещение визуально близких классов дефектов, а также рост числа ложных срабатываний для редко представленных категорий. Во всех случаях в качестве модели для обучения применялись базовые модификации моделей типа small. Данный эффект особенно выражен у более сложных версий моделей, где увеличение числа параметров не сопровождается пропорциональным ростом обобщающей способности.

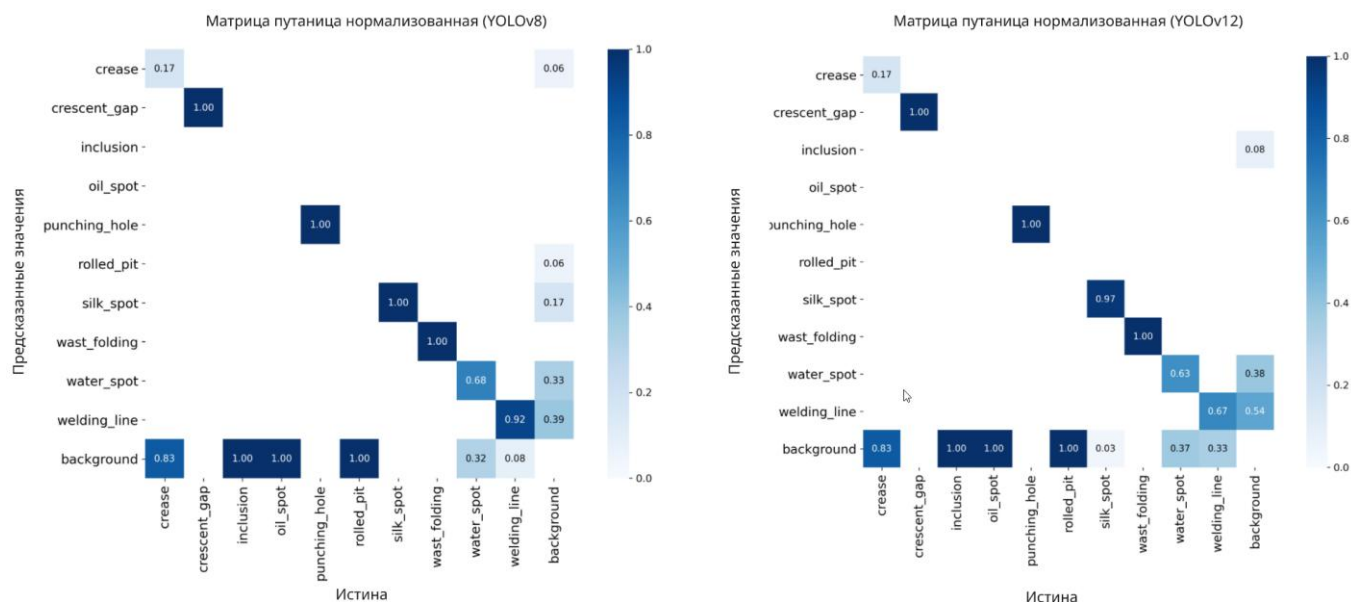


Рис. 22. Матрицы путаницы YOLOv8, YOLOv12

В то же время матрица путаницы для архитектуры «DICONDE-YOLO», представленная на рисунке 24, иллюстрирует снижение доли межклассовых ошибок

и более устойчивое распознавание редких и слабо различимых дефектов. Это подтверждает, что комплексирование визуальных признаков с токенизированными фактографическими метаданными, а также использование модифицированных блоков агрегации признаков, способствует формированию более дискриминативного пространства представлений. В частности, наблюдается уменьшение ошибок путаницы между классами, различающимися контекстными и технологическими характеристиками, что невозможно достичь при использовании исключительно визуальной информации. Отдельные аспекты проведения экспериментов, включая процедуру экспертной разметки, характеристики используемых датасетов, а также условия обучения и тестирования моделей, были апробированы и обсуждались на Всероссийских и международных научных конференциях [67–70], а также отражены в публикациях автора [71–80].

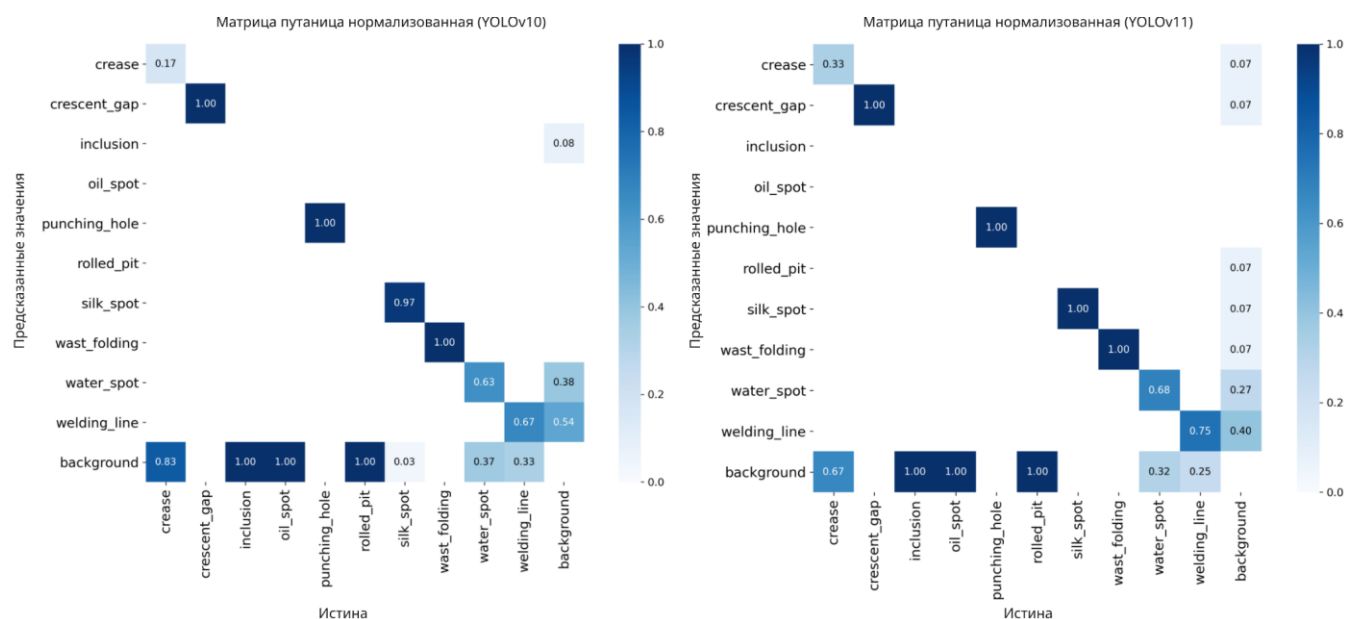


Рис. 23. Матрицы путаницы YOLOv10, YOLOv11

Таким образом, анализ матриц ошибок дополняет количественные результаты и наглядно подтверждает эффект от внесенных архитектурных и алгоритмических изменений, демонстрируя повышение устойчивости модели к дефициту данных и улучшение интерпретируемости принимаемых решений в задаче разработки комплексированного алгоритма поиска схожих изображений.

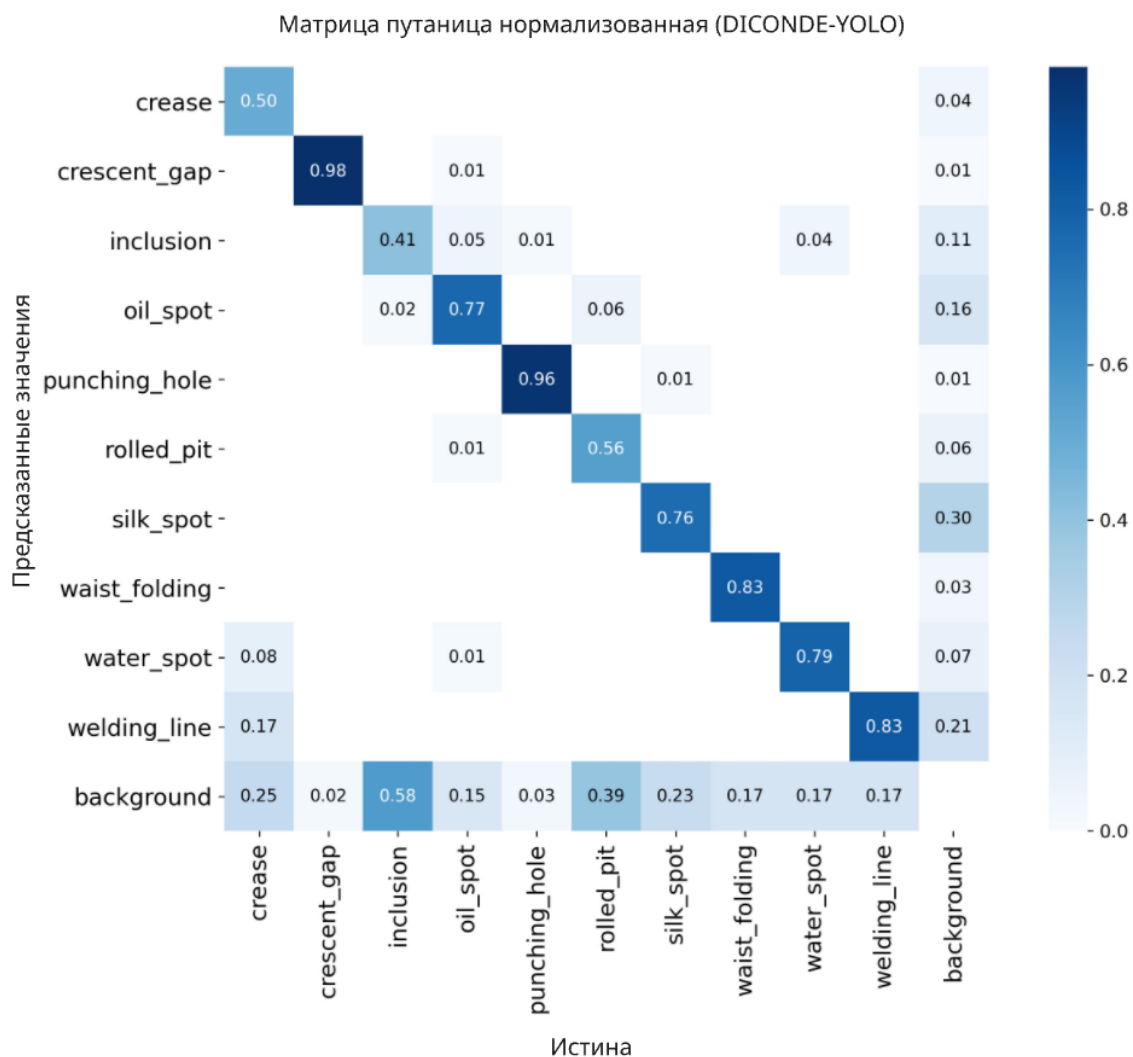


Рис. 24. Матрица путаницы «DICONDE-YOLO»

Для оценки практической масштабируемости системы проведены измерения времени отклика и использования памяти при увеличении объема БД изображений. До  $10^6$  объектов использовались прямые измерения, а для диапазона  $10^6$ – $10^8$  показателей применена экстраполяция с учетом особенностей ANN-индексов, используемых для поиска по эмбедингам. Полученные результаты представлены в таблице 3.7 и на рисунке 25. Принцип экстраполяции ANN-индекса заключается в следующем:

Пусть  $\chi = \{x_i \in \mathbb{R}^d\}$  – множество эмбедингов изображений фиксированной размерности  $d$ , полученных с использованием архитектуры «DICONDE-YOLO», а  $N = |\chi|$  – размер БД. Поиск ближайших соседей осуществляется с применением ANN-

индекса, реализующего навигацию в индексной структуре при фиксированных параметрах, определяемых в процессе AutoML-оптимизации.

Согласно теоретическим оценкам для современных ANN-структур, используемых в задачах высокоразмерного поиска, математическое ожидание времени выполнения одного поискового запроса имеет логарифмическую зависимость от объема БД и может быть представлено в виде:

$$T_{ANN}(N) = a * \log N + b, \quad (3.1)$$

где коэффициенты  $a$  и  $b$  зависят от параметров индекса, размерности пространства признаков и допустимой вероятности приближенного поиска, но не зависят от объема данных  $N$ .

Аналогичным образом обосновывается экстраполяция требований к памяти. Потребление памяти ANN-индекса определяется линейной зависимостью от числа хранимых эмбеддингов и параметров индексной структуры и может быть описано выражением:

$$M(N) = N * (d + M_{conn} + \alpha), \quad (3.2)$$

где  $M_{conn}$  – среднее число связей на элемент индекса, а  $\alpha$  – накладные расходы, связанные с хранением служебной информации.

Таблица 3.7 – Оценка масштабируемости системы

<b>Кол-во изображений в БД</b>	<b>Время отклика, мс</b>	<b>Precision</b>	<b>Использование RAM, ГБ</b>
$10^4$	34.3	0.84	1.2
$10^5$	43.1	0.83	12
$10^6$	56.3	0.82	127.3
$10^7$	68.2	0.81	1200
$10^8$	74.4	0.79	12000

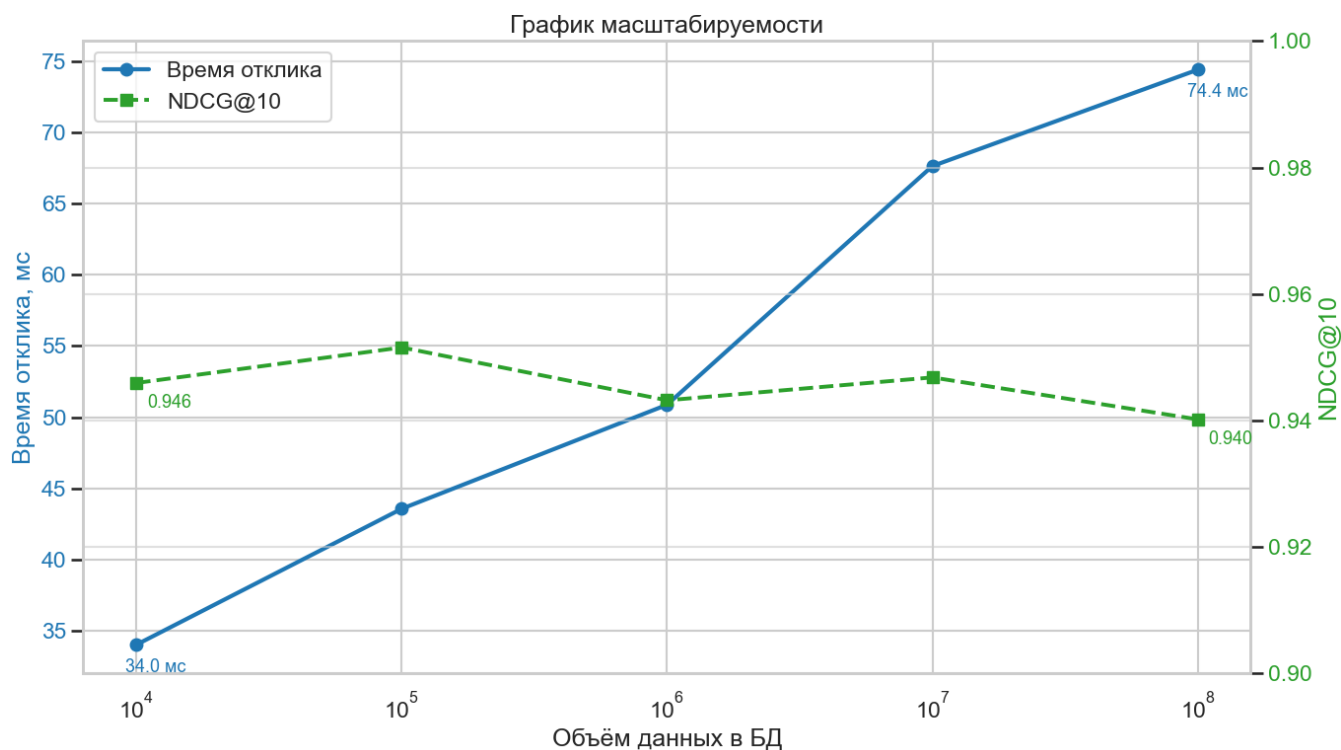


Рис. 25. График масштабируемости при увеличении объема БД

Полученные характеристики масштабируемости соответствуют требованиям реальных СППР, где поиск осуществляется в интерактивном режиме и время отклика не должно превышать сотен миллисекунд. Представленные результаты позволяют рассматривать предложенный алгоритм как практическое решение для интеграции в СППР и эксплуатации в производственных условиях.

Анализ результатов, представленных в таблице 3.7 и на рисунке 25, показывает, что предложенный комплексированный алгоритм демонстрирует устойчивый рост времени отклика при увеличении объема данных, при этом показатель точности поиска NDCG остается стабильным, вне зависимости от объема данных. Наблюдаемое увеличение потребления памяти обусловлено ростом ANN-индекса и увеличением числа хранимых эмбедингов.

Полученные зависимости свидетельствуют о масштабируемости алгоритма и подтверждают его применимость в условиях реальных СППР, где объем хранимых данных может достигать сотен тысяч и более объектов.

### 3.4 Выводы по главе

В результате проведенных вычислительных экспериментов подтверждена эффективность предложенного комплексированного алгоритма поиска схожих изображений. Экспериментально показано, что интеграция визуальных и фактографических признаков в едином пространстве эмбедингов повышает качество поиска и улучшает ранжирование без существенного роста вычислительной нагрузки.

Сравнительный анализ версий YOLO подтвердил выбор YOLOv8 в качестве базовой архитектуры, а абляционные эксперименты выявили значимость блока ASPP, модулей ECA, DC/MDC и токенизации метаданных для повышения устойчивости и обобщающих способностей модели.

Масштабируемость системы подтверждена измерениями и экстраполяцией до объемов  $10^8$  изображений. Сделан вывод, что время отклика растет сублинейно, а снижение точности поиска носит умеренный характер, что демонстрирует практическую применимость алгоритма в условиях крупных СППР.

Ограничения работы связаны с объемом экспертной разметки и спецификой датасетов, что определяет направления дальнейших исследований, включая автоматизацию разметки и улучшение индексации эмбедингов. Полученные результаты вычислительных экспериментов создают прочную основу для внедрения модели и алгоритма в производственные СППР.

#### **Глава 4. ПРАКТИЧЕСКАЯ РЕАЛИЗАЦИЯ ВСТРАИВАЕМОСТИ В ЕДИНОЕ ИНФОРМАЦИОННОЕ ПРОСТРАНСТВО. СРАВНЕНИЕ С ПРОТОТИПОМ СППР И АНАЛИЗ ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ**

Текущая глава диссертационной работы посвящена практической реализации, внедрению и анализу результатов применения разработанной модели и алгоритмов интеллектуального анализа изображений в DICOM-подобных данных в составе прикладной СППР. В отличие от предыдущих глав, ориентированных на формализацию и экспериментальное обоснование предложенного подхода, в данной главе основной акцент сделан на интеграцию программно-алгоритмических решений в единое информационное пространство, оценке их эффективности и сопоставление с базовыми решениями.

Рассматривается архитектура программной реализации, особенности встраивания алгоритмов в существующие информационные контуры, а также результаты опытной эксплуатации разработанного программного прототипа. Особое внимание уделяется оценке качества интеллектуального анализа изображений, скорости обработки и устойчивости работы системы в условиях, приближенных к реальным сценариям использования.

Рассматриваются требования к интеграции алгоритмов интеллектуального анализа изображений в СППР, архитектура программной реализации, описание реализованного прототипа и базового решения для сравнения, а также результаты сравнительной оценки эффективности в условиях опытной эксплуатации.

Кроме того, в главе отражены результаты внедрения диссертационной работы на промышленном предприятии, а также ее применение в образовательном процессе при подготовке студентов РХТУ имени Д.И. Менделеева, что подтверждает практическую значимость и прикладную направленность выполненной работы.

#### **4.1 Требования к интеграции алгоритмов интеллектуального анализа изображений в единое информационное пространство СППР**

Современные СППР, функционирующие в условиях обработки изображений в DICOM-подобных форматах, характеризуются высокой степенью распределенности и гетерогенностью информационных источников. В рамках таких систем визуальные данные, сопровождаемые метаинформацией, формируются, хранятся и анализируются в различных подсистемах, что требует наличия единого информационного пространства, обеспечивающего согласованный доступ как к данным, так и результатам интеллектуального анализа.

Под единым информационным пространством понимается совокупность взаимосвязанных программно-аппаратных компонентов, включающая хранилища изображений, базы структурированных метаданных, вычислительные модули интеллектуального анализа и пользовательские интерфейсы СППР, взаимодействующие на основе унифицированных протоколов и форматов представления данных. Тогда разработанные алгоритмы интеллектуального анализа изображений должны рассматриваться не как изолированные программные модули, а как функциональные сервисы, встроенные в общий процесс поддержки принятия решений.

Интеграция разработанной модели и алгоритмов интеллектуального анализа изображений в единое информационное пространство предъявляет ряд функциональных требований. К ним относятся возможность обработки изображений и сопутствующих метаданных в различных DICOM-подобных представлениях, поддержка пакетной и запросной обработки, а также обеспечение детерминированного формирования признаков представлений, пригодных для последующего поиска и ранжирования. Существенным является требование инвариантности алгоритмов к вариативности структуры метаданных и различиям в

параметрах визуальных данных, обусловленных условиями их получения и формирования.

Практическая значимость сформулированных требований подтверждена в ходе внедрения разработанных решений в информационную систему производственного предприятия, где интеллектуальный анализ изображений используется как часть контура поддержки принятия решений при контроле и анализе визуального ряда изображений. Реализация данных требований позволила обеспечить согласованную работу алгоритмов анализа изображений с существующими системами хранения и обработки данных без нарушения их функциональной целостности.

Наряду с функциональными, ключевое значение имеют нефункциональные требования, определяющие практическую применимость алгоритмов в составе СППР. К таким требованиям относятся масштабируемость при росте объема хранимых изображений, ограничение времени отклика при выполнении поисковых запросов, воспроизводимость результатов интеллектуального анализа, а также совместимость с существующими информационными и аналитическими компонентами системы. Выполнение указанных требований является необходимым условием использования разработанных алгоритмов в реальных сценариях эксплуатации СППР, ориентированных на поддержку экспертных решений.

В программной реализации разработанного прототипа указанные требования были реализованы за счет микросервисной организации модулей интеллектуального анализа, использования унифицированных форматов обмена данными и разделения вычислительных и пользовательских контуров. Это позволило обеспечить масштабируемость системы при росте объема данных и воспроизводимость результатов анализа в условиях промышленной эксплуатации.

Таким образом, разработка и внедрение алгоритмов интеллектуального анализа изображений в DICOM-подобных данных предполагает не только достижение высоких показателей качества анализа, но и обеспечение их корректной

архитектурной интеграции в единое информационное пространство систем поддержки принятия решений.

#### **4.2 Архитектура программной реализации и место разработанного алгоритма в системе**

С учетом требований к устойчивости, воспроизводимости и прикладной применимости интеллектуального анализа изображений в составе СППР, программная реализация разработанных моделей и алгоритмов построена на основе многоуровневой модульной архитектуры. Ключевой особенностью данной архитектуры является явное выделение этапа предобработки DICOM-подобных данных, предшествующего использованию нейросетевых моделей, а также применение модифицированной архитектуры «DICONDE -YOLO».

Этап предобработки выполняет функции нормализации визуальных данных, приведения изображений к унифицированному пространственному и каналному представлению, а также синхронизации визуальной информации с сопутствующими фактографическими метаданными. Данный этап обеспечивает устойчивость последующего нейросетевого анализа к вариативности входных данных, обусловленной различиями в условиях формирования и хранения DICOM-подобных изображений.

Непосредственно извлечение визуальных признаков осуществляется с использованием разработанной модифицированной нейросетевой архитектуры «DICONDE -YOLO», представляющей собой развитие базовой модели YOLOv8. В рамках «DICONDE -YOLO» реализованы механизмы адаптации к специфике DICOM-подобных изображений, а также расширенные возможности формирования

информативных эмбеддингов, пригодных для задач интеллектуального поиска и аналитической обработки.

Полученные визуальные представления интегрируются с токенизированными фактографическими признаками в рамках комплексированного алгоритма интеллектуального анализа изображений. Формирование единого признакового пространства обеспечивает повышение релевантности поиска и устойчивость работы СППР в условиях неполной представленности метаданных.

Архитектура системы дополнена алгоритмами балансировки нагрузки и автоматизированного подбора параметров моделей на основе подходов AutoML [59]. Балансировка нагрузки реализована на основе NGINX, что обеспечивает равномерное распределение вычислительных задач между узлами кластера, а AutoML — адаптацию параметров «DICONDE-YOLO» и механизмов комплексирования признаков к изменяющимся характеристикам входных данных.

Представленная на рисунке 26 архитектура принципиально отличается от традиционных решений, в которых метаданные используются на этапе постобработки или внешнего ранжирования. В разработанной модели «DICONDE-YOLO» извлечение и токенизация метаданных реализованы в виде отдельного блока, результаты работы которого интегрируются с визуальными признаками на уровне последних слоев нейронной сети.

Такое решение обеспечивает формирование единого признакового представления, в котором визуальная и фактографическая информация взаимно усиливают друг друга. Интеграция на уровне представлений позволяет учитывать контекстные и технологические характеристики объектов уже на этапе формирования признаков, а не на стадии принятия решений, что положительно сказывается на устойчивости и интерпретируемости результатов интеллектуального анализа.

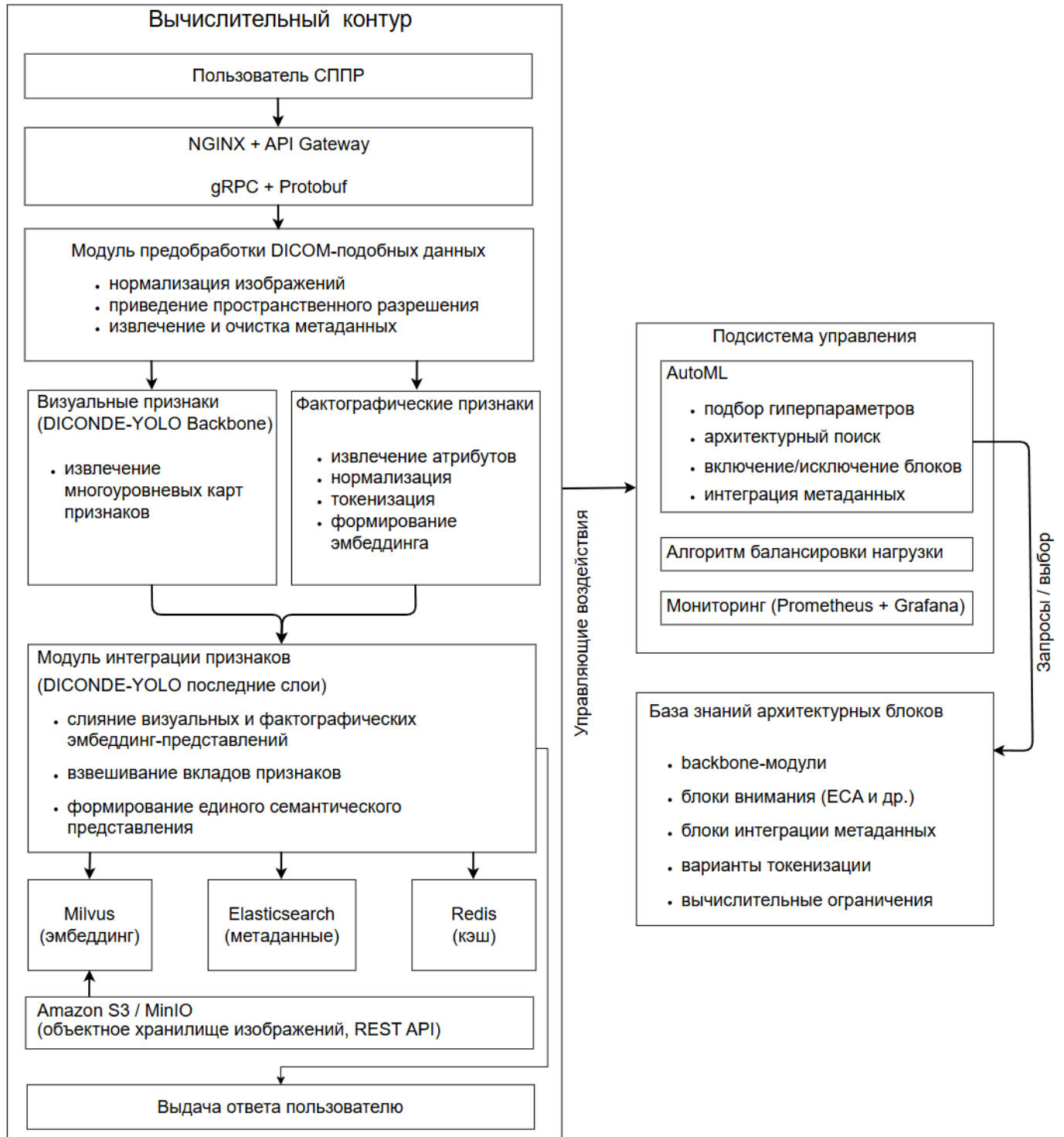


Рис. 26. Архитектура интеллектуальной СППР с управляющим AutoML-контуром

В отличие от классических систем, в которых архитектура нейронной сети фиксируется на этапе проектирования, в разработанной системе «DICONDE-YOLO» реализован управляющий контур AutoML, обеспечивающий автоматизированный

параметрический и архитектурный поиск. AutoML использует сведения о доступных нейросетевых блоках, механизмах интеграции метаданных и вычислительных ограничениях, хранящиеся в БЗ, и на их основе формирует модифицированные конфигурации модели.

Управляющий контур получает ответ от подсистем мониторинга на основе Prometheus [81] и Grafana [82] и распределяет нагрузку в соответствии с имеющимися вычислительными мощностями, что позволяет адаптировать архитектуру модели под особенности данных, объем обучающей выборки и требования к вычислительным ресурсам. Таким образом, архитектура «DICONDE-YOLO» является адаптивной и формируется как результат оптимизационного процесса, а не как заранее заданная структура.

Предложенный подход обеспечивает масштабируемость архитектуры, так как обработка фактографических и визуальных данных может выполняться асинхронно, а механизмы AutoML и балансировки нагрузки позволяют адаптировать систему к изменяющимся условиям эксплуатации и вариативности входного набора данных.

Таким образом, предложенная архитектура обеспечивает согласованную работу этапов предобработки, нейросетевого анализа и комплексирования признаков в составе распределенного информационного пространства СППР и подтверждает технологическую готовность разработанного решения к практическому внедрению.

#### **4.3 Реализация прототипа и описание базового решения для сравнительного анализа**

Для оценки практической эффективности предложенной модели и алгоритмов интеллектуального анализа изображений был реализован программный прототип системы, интегрированный в единое информационное пространство СППР. Прототип обеспечивает обработку DICOM-подобных изображений совместно с фактографическими метаданными и построен с использованием микросервисной

архитектуры, поддерживающей масштабирование, отказоустойчивость и высокую производительность.

Прототип СППР реализован на Python 3.11, с использованием программных библиотек PyTorch 2.1 [83] и OpenCV 4.13 [84], используемых для предобработки изображений и формирования визуальных эмбеддингов. Для высокопроизводительной коммуникации между сервисами, обрабатывающими визуальные и фактографические данные, используется связка protobuf [85] + gRPC [86], а для интеграции с внешними клиентами — REST API + JSON [87, 88]. В качестве инструмента локальной контейнеризации использовался Docker [89]. Промышленная интеграция предполагает замену данного инструмента на масштабируемый аналог, поддерживающий возможность оркестрации – Kubernetes [90].

Хранение и поиск визуальных данных осуществляется через Milvus [91], использующий объектное хранилище MinIO [91] для долговременного хранения изображений. Это позволяет эффективно управлять большим объемом визуальной информации, обеспечивая быстрый доступ к эмбеддингам и масштабируемость при росте данных. Кэширование промежуточных результатов и ускорение поиска реализованы с помощью программного решения Redis [92], а асинхронная обработка задач — посредством Apache Pulsar [93] или NATS [94], в зависимости от требуемого масштаба.

Для балансировки нагрузки используется NGINX, который распределяет запросы между узлами обработки, поддерживая высокую доступность и отказоустойчивость. Мониторинг производительности, сбор логов и визуализация метрик обеспечиваются Prometheus [81] + Grafana [82], а структурированные метаданные индексируются на основе обратного индекса, что позволяет выполнять оперативное получение данных с помощью Elasticsearch [95, 96].

Прототип реализует комплексированное формирование признаков на основе модифицированной архитектуры «DICONDE-YOLO», объединяя визуальные эмбеддинги с токенизированными фактографическими данными. Это позволяет

создавать единое семантическое пространство признаков, обеспечивая высокую релевантность поиска и устойчивость работы СППР при частичной или неполной представленности метаданных. Применение Milvus совместно с Redis обеспечивает ANN-поиск и кэширование на двух уровнях, что положительно сказывается на времени отклика и масштабируемости системы. Асинхронная обработка запросов через Pulsar/NATS и балансировка нагрузки через NGINX [97] обеспечивают поддержание стабильного времени отклика и точности анализа даже при высоких нагрузках.

Для сравнительного анализа использовалась базовая модель YOLOv8, ориентированная исключительно на извлечение визуальных признаков без интеграции фактографических данных и распределенных сервисов. Данное решение отражает распространенную практику интеллектуального анализа изображений и служит репрезентативным «эталоном» для количественной оценки улучшений, достигнутых с применением комплексирования признаков и распределенной архитектуры. Сравнение с базовым решением показало значительное улучшение ключевых показателей качества поиска и ранжирования, а также повышение устойчивости к вариативности данных и неполноте метаданных.

В качестве базового решения для сравнительного анализа использовалась архитектура, реализующая типовой подход к интеллектуальному анализу изображений, основанный исключительно на визуальных признаках, извлекаемых с помощью модели СНС YOLOv8 без модификаций и без интеграции фактографических данных. Данное решение отражает распространенную практику применения нейросетевых моделей для задач анализа изображений и поиска схожих объектов и рассматривается репрезентативный прототип базового решения.

Следует отметить, что базовое решение, основанное исключительно на визуальных признаках, не учитывает контекстную и технологическую информацию, содержащуюся в метаданных, что ограничивает его применимость в задачах поддержки принятия решений. В условиях визуальной неоднозначности и высокой

внутриклассовой вариативности это приводит к снижению устойчивости поиска и росту числа семантически некорректных результатов.

Сравнение на основе результатов вычислительных экспериментов с базовым решением обеспечило количественную оценку вклада предложенных архитектурных и алгоритмических модификаций, включая интеграцию фактографических признаков, механизмы их токенизации и формирование единого признакового пространства, рассматриваемого в разделе 4.4 настоящей диссертационной работы. Такой подход позволяет обоснованно судить о преимуществах разработанной модели и алгоритмов интеллектуального анализа изображений в условиях их практического применения в СППР.

Описанная архитектура реализована в виде программного прототипа и внедрена в опытную эксплуатацию на промышленном предприятии, где алгоритмы интеллектуального анализа изображений в виде программных приложений интегрированы в существующую информационную инфраструктуру. В рамках внедрения разработанные модель и комплексированный алгоритм функционируют в качестве сервиса интеллектуального анализа, предоставляющего результаты поиска и ранжирования изображений другим подсистемам СППР.

#### **4.4 Анализ результатов опытной эксплуатации и сравнительная оценка эффективности разработанной модели и алгоритмов в составе СППР**

В данном разделе представлен анализ результатов опытной эксплуатации разработанной модели и комплексного алгоритма интеллектуального анализа изображений на основе «DICONDE-YOLO», ориентированного на обработку DICOM-подобных данных в составе прикладной СППР. Основной акцент сделан на количественную и качественную оценку эффективности предлагаемых модели и

алгоритма, сравнении с базовым решением в виде YOLOv8, использующим только визуальные признаки, а также анализе эксплуатационных характеристик системы в условиях, приближенных к реальным сценариям промышленной эксплуатации. По мнению автора, такое решение позволяет подтвердить эффективность комплексированного подхода и обосновать практическую значимость для задач поддержки принятия решений.

Важно отметить, что при обучении модели ИНС «DICONDE-YOLO» применялся расширенный набор классов, который отличается от применяемого при проведении сравнительного анализа архитектур YOLOv8-YOLOv12. Описание расширенного набора представлено в таблице 4.1.

<b>Класс</b>	<b>Описание</b>
crease	складка
crescent_gap	серповидный зазор
inclusion	включение
oil_spot	масляное пятно
punching_hole	дефект штамповки
rolled_pit	прокатная выемка
silk_spot	волосовина
waist_folding	клиновидный дефект
water_spot	водяное пятно
welding_line	линия сварного шва
pore	пора
clustered_pores	скопление пор
faulty_fusion	непровар
slag	шлак
slag_cluster	кластер шлака

Количественная оценка качества интеллектуального анализа проводилась с использованием совокупности метрик, характеризующих точность обнаружения и поиска объектов, полноту результатов и релевантность ранжирования. В качестве ключевых показателей применялись mAP50, Precision, Recall, CIoU и NDCG [98, 99], что обеспечило комплексную оценку как точности локализации, так и качества формирования ранжированной выдачи результатов поиска. Сводные данные сравнительной оценки качества интеллектуального анализа изображений на основе базовой модели YOLOv8 и разработанного комплексного алгоритма в сочетании с моделью на основе архитектуры «DICONDE-YOLO» приведены в таблице 4.2.

Таблица 4.2 – Сравнительная данные оценки качества интеллектуального анализа изображений

<b>Модель / алгоритм</b>	<b>mAP50</b>	<b>Precision</b>	<b>Recall</b>	<b>CIoU</b>	<b>NDCG</b>
YOLOv8	0.67	0.63	0.65	0.58	0.61
Комплексированный алгоритм + «DICONDE-YOLO»	0.94	0.85	0.83	0.84	0.81
Изменение	+40%	+34%	+27%	+44%	+33%

Анализ полученных результатов демонстрирует устойчивое улучшение качества поиска и ранжирования изображений при интеграции визуальных признаков и токенизированных фактографических данных. Наиболее существенный прирост наблюдается при оценке показателя NDCG, характеризующего релевантность первых позиций выдачи, показавший улучшение результатов на 33%. Данный результат имеет принципиальное значение для СППР, поскольку корректность и информативность наиболее релевантных результатов поиска напрямую влияют на

качество принимаемых экспертных решений. Повышение показателей mAP50 и Recall свидетельствуют о способности СППР на основе разработанной модели и алгоритма корректно выявлять и идентифицировать объекты как при стандартных ярко выраженных признаках, так и в условиях высокой внутриклассовой вариативности изображений и неполноты представленных фактографических сведений.

Оценка эксплуатационных характеристик системы проводилась на основе обработки 1000 случайно отобранных изображений с многократным повторением измерений. В качестве основных показателей рассматривались время инференса, время отклика на пользовательский запрос, устойчивость работы при неполных метаданных и стабильность получаемых результатов. Сводные эксплуатационные показатели системы при использовании YOLOv8 и «DICONDE-YOLO» представлены в таблице 4.3.

Таблица 4.3 – Эксплуатационные показатели системы интеллектуального анализа изображений

<b>Показатель</b>	<b>YOLOv8</b>	<b>DICONDE-YOLO</b>
Время инференса, мс	20.18	18.73
Время отклика, мс	65.3	59.4±7.7
Устойчивость при неполных метаданных	Средняя	Высокая
Стабильность результатов	Средняя	Высокая

Полученные результаты показывают, что модель на основе архитектуры «DICONDE-YOLO» демонстрирует улучшение как по времени отклика, так и по устойчивости и стабильности результатов по сравнению с базовым вариантом

YOLOv8. Незначительное снижение времени инференса объясняется меньшей вычислительной сложностью при формировании признаковых представлений. Использование токенизированных фактографических данных совместно с визуальными признаками обеспечивает корректное ранжирование объектов даже при частичной или неполной представленности метаданных, что является характерной особенностью промышленных DICOM-подобных данных.

Качественный анализ результатов дополняет количественные оценки и проводится на примерах визуально схожих изображений, но различающихся по классам дефектов и контекстным характеристикам. На рисунке 27 представлен случай визуальной неоднозначности, при которой изображения имеют близкие визуальные признаки, но относятся к различным классам дефектов.



Рис. 27. Пример визуально близких признаков при разных классах дефектов

Результаты работы прототипа на основе базовой модели YOLOv8 как инструмента при поиске схожих изображений и представленные на рисунках 28, 29,

демонстрируют наличие топ-5 выдачи изображений, не соответствующих фактическому контексту изображения-запроса.

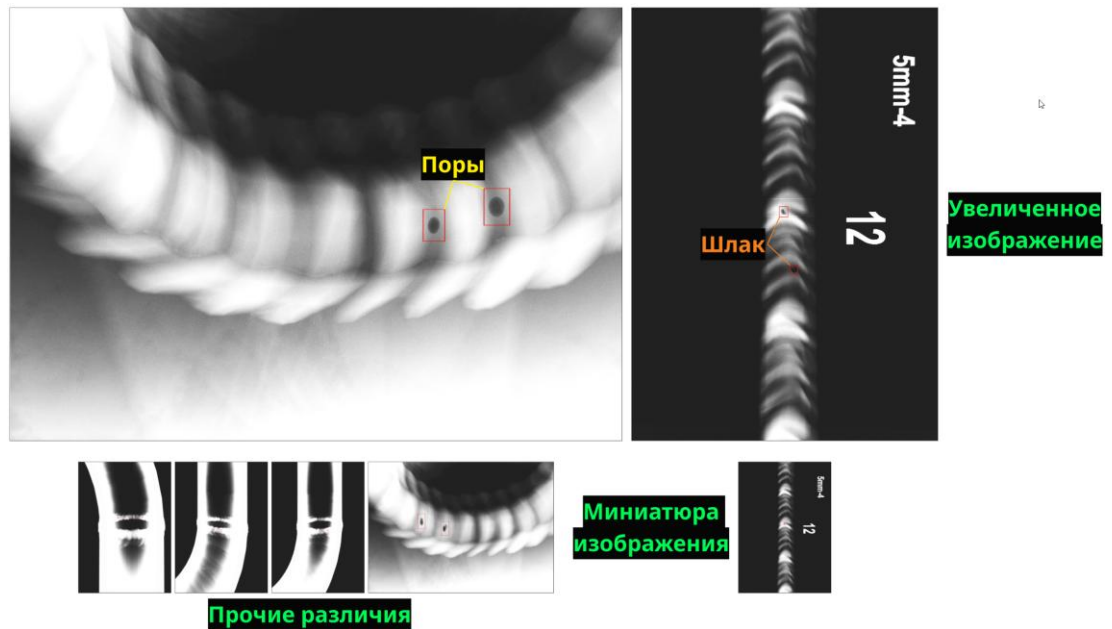


Рис. 28. Результаты работы прототипа на основе архитектуры YOLOv8

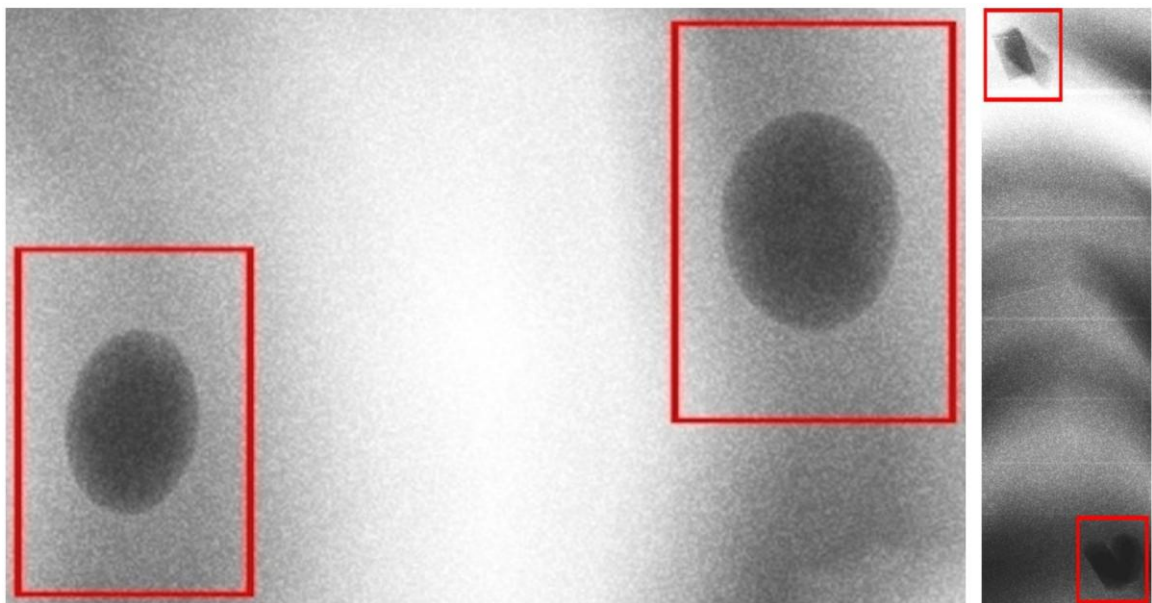


Рис. 29. Пример различий (увеличено)

В то же время применение комплексированного алгоритма поиска совместно с моделью на основе архитектуры «DICONDE-YOLO» демонстрирует полностью соответствующую фактическим характеристикам изображению-запросу выдачу, что иллюстрируют рисунки 30, 31 и подтверждает улучшение обобщающей способности

модели на основе интеграции визуальных и фактографических признаков на уровне эмбедингов. Дополнительные примеры детекции дефектов с использованием указанных моделей приведены в приложении В.

Разработанная модель и комплексированный алгоритм поиска функционируют в составе распределенной СППР совместно с алгоритмами балансировки нагрузки и автоматизированного подбора гиперпараметров и архитектурных конфигураций на основе AutoML. Интерфейс разработанного прототипа СППР и примеры работы представлены в приложении Г. Автоматизированный управляющий контур обеспечивает не только оптимизацию гиперпараметров, но и перестройку внутренней структуры нейронной сети на основе сведений о доступных вычислительных блоках и данных мониторинга качества. Это позволяет адаптировать архитектуру «DICONDE-YOLO» к изменяющимся характеристикам входного набора данных и требований к производительности.

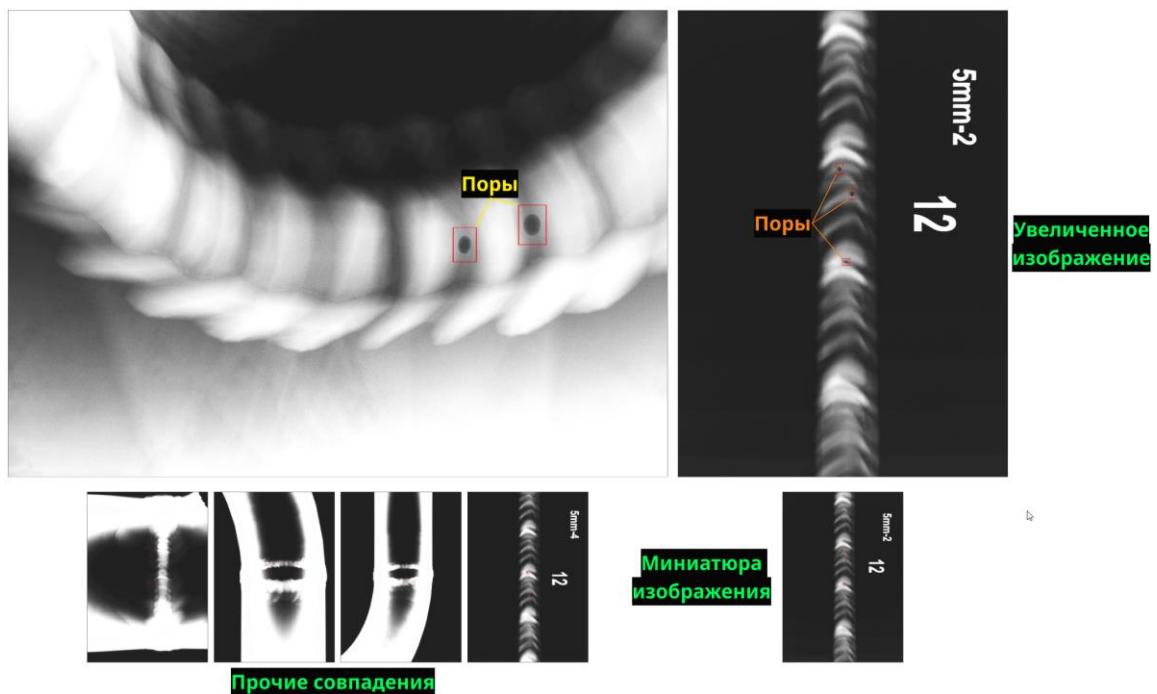


Рис. 30. Результаты поиска схожих изображений при использовании комплексированного алгоритма поиска и «DICONDE-YOLO»

Балансировка нагрузки обеспечивает равномерное распределение вычислительных задач, что позволяет сохранять стабильное время отклика и точность анализа при росте интенсивности запросов.

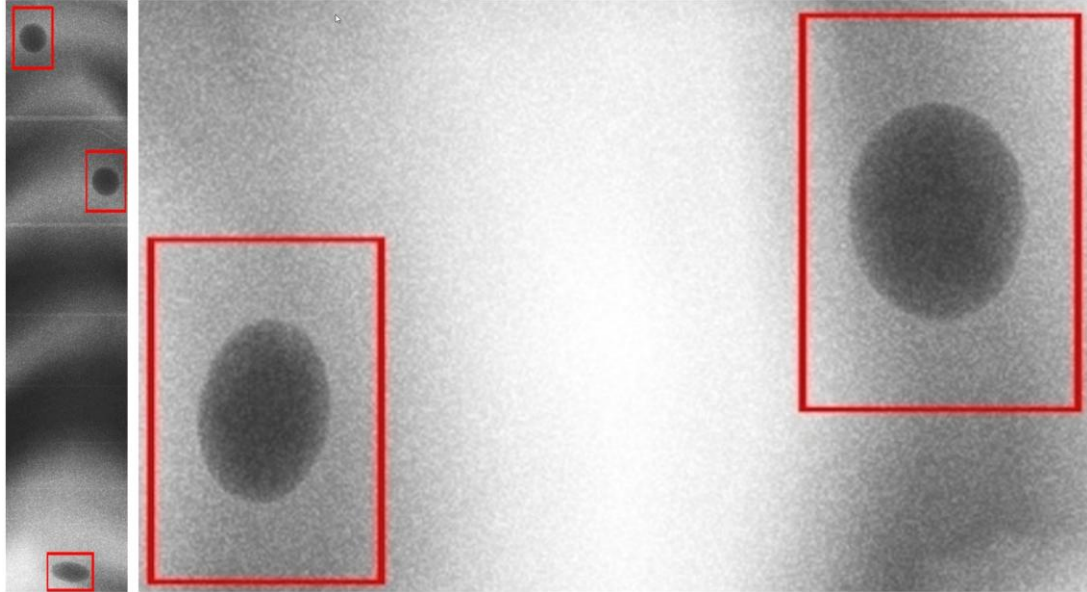


Рис. 31. Пример совпадений (увеличено)

Кроме того, для подтверждения инвариантности применения предложенного решения, был проведен дополнительный эксперимент на основе обучения на данных, относящихся к иной предметной области – медицинскими изображениям. Результат представлен на рисунке 32.

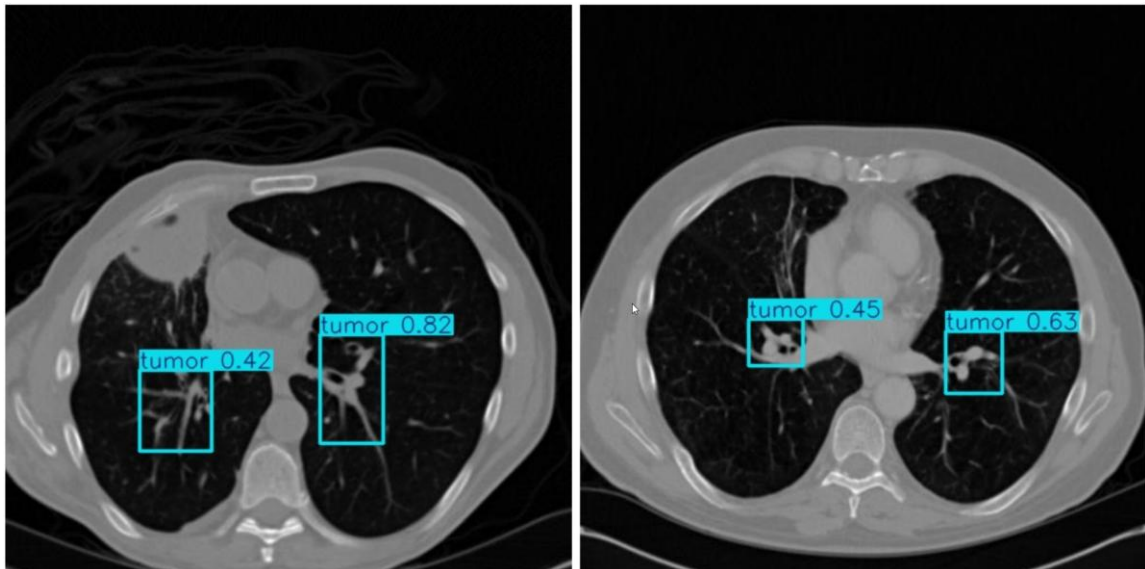


Рис. 32. Пример поиска совпадений на медицинских изображениях

Сравнительные показатели задержек, времени отклика и пропускной способности при различной нагрузке приведены в таблице 4.4.

Таблица 4.4 – Сравнительные результаты основных метрических показателей системы

Кол-во запросов	Latency (D), мс	Latency (Y), мс	Response (D), мс	Response (Y), мс	Throughput (D), зап/сек	Throughput (Y), зап/сек
100	10.9±2.1	19.7±4.7	46.1±7.1	71.9±13.4	1587±23	1185±18
500	13.7±2.2	23.8±5.2	58.9±8.9	93.1±19.7	1508±25	1102±27
1000	16.0±3.6	25.1±5.8	65.3±10.7	103.2±21.5	1379±42	1043±31
3000	18.7±3.7	29.3±7.2	77.5±14.3	117.1±28.9	997±40	747±24
5000	20.0±4.3	32.7±9.7	84.3±15.5	130.5±33.6	750±30	559±26
10000	24.3±5.3	37.9±9.9	97.0±19.9	147.1±42.2	434±21	322±14

График масштабируемости и изменения времени отклика в зависимости от количества запросов для СППР на основе архитектуры «DICONDE-YOLO» с применением комплексированного алгоритма поиска в сравнении с YOLOv8 представлен на рисунке 33.

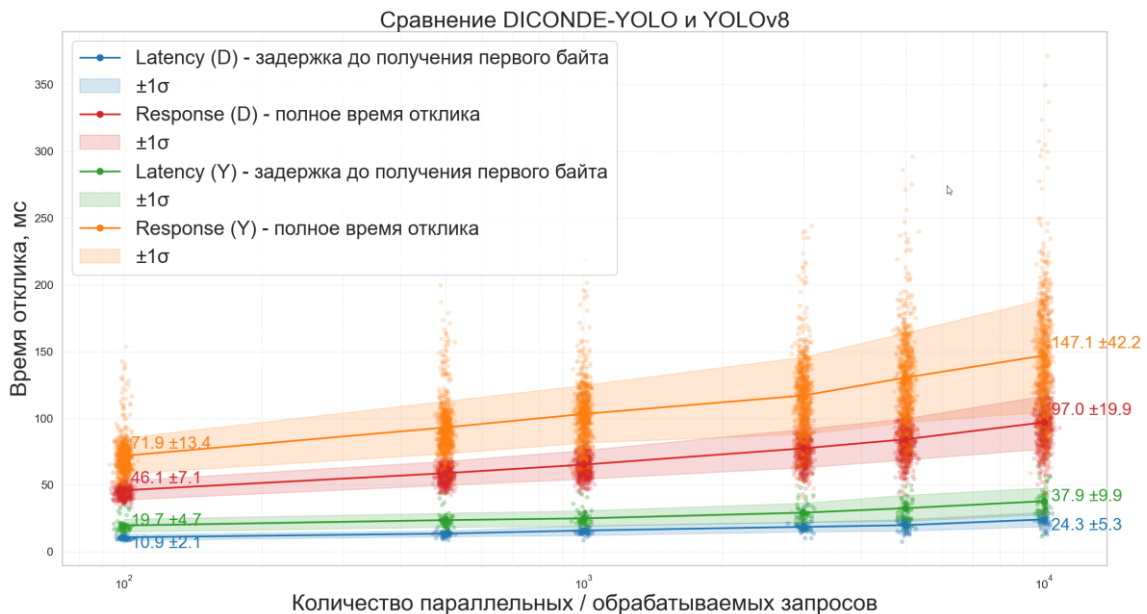


Рис. 33. График времени отклика и задержек YOLOv8 и «DICONDE-YOLO» в зависимости от количества запросов

График показателя пропускной способности системы при использовании YOLOv8 и DICONDE-YOLO с комплексированным алгоритмом поиска представлен на рисунке 34.

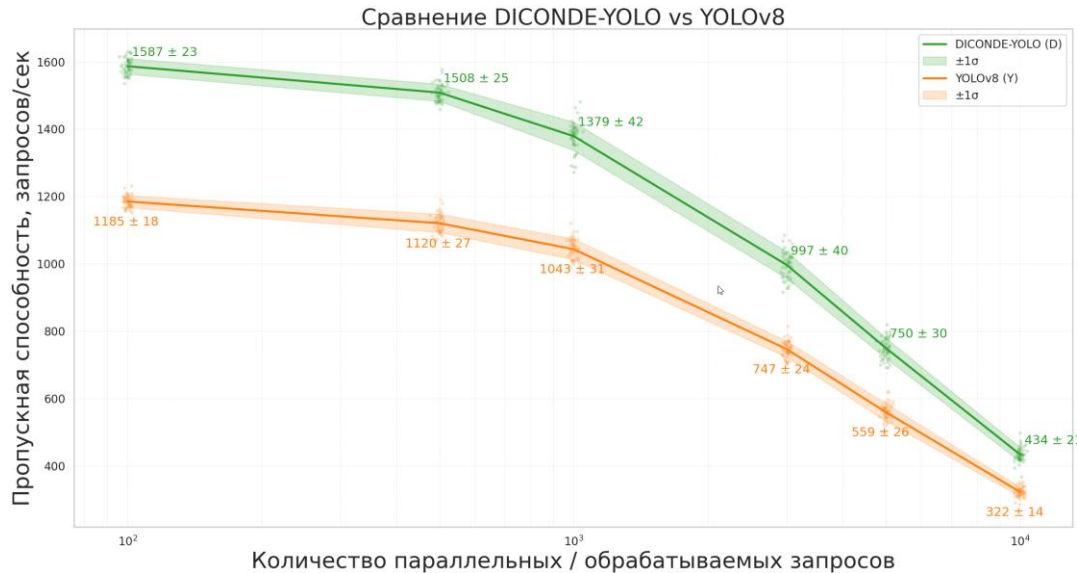


Рис. 34. График пропускной способности YOLOv8 и «DICONDE-YOLO» в зависимости от количества запросов

Опытная эксплуатация и тестовые испытания прототипа на промышленном предприятии (далее – предприятие) показала, что внедрение комплексированного алгоритма поиска изображений в сочетании с моделью на основе архитектуры «DICONDE-YOLO» приводит к существенному сокращению времени анализа визуальных данных, снижению числа ложных тревог и уменьшению нагрузки на экспертов.

В рамках опытной эксплуатации и тестовых испытаний использовался архив визуальных данных DICOM-подобного формата общим объемом порядка 160 000 изображений, сформированных в ходе производственного контроля за предыдущие периоды. Ежедневно обрабатывалось в среднем от 550 до 650 новых изображений.

Система функционировала в формате микросервисной архитектуры с использованием протоколов gRPC для внутреннего взаимодействия и REST-интерфейсов для интеграции с внешними модулями предприятия. В качестве

основной единицы вычислительного контура выступал сервер со следующими аппаратными характеристиками:

- CPU: I7-13700K;
- RAM: 32GB DDR4 3200MHz;
- GPU: RTX 4060 8GB.

Внедрение результатов исследований подтверждены актом, представленным в приложении Д, а основные количественные результаты опытной эксплуатации приведены в таблице 4.5. Полученные данные свидетельствуют о повышении эффективности работы специалистов и росте пропускной способности системы без ухудшения качества принимаемых решений.

Использование разработанного прототипа в образовательном процессе при подготовке студентов подтвердило ее дидактическую ценность и устойчивость работы алгоритмов, позволяя демонстрировать современные подходы к интеллектуальному анализу изображений с практической интеграцией в СППР, что подтверждено справкой об использовании результатов, приведенной в приложении Е. Таблица 4.5 – Результаты опытной эксплуатации системы на основе разработанных решений

<b>Показатель</b>	<b>До внедрения</b>	<b>После внедрения</b>	<b>Изменение</b>
Среднее время анализа дефекта	90 мин	35 мин	–61%
Среднее количество ложных обнаружений	12%	5%	–58%
Уровень нагрузки эксперта	4.7 / 5	2.9 / 5	–38%
Количество обработанных изображений в смену	30	75	+150%

Результаты количественного и качественного анализа, а также данные опытной эксплуатации подтверждают, что разработанная модель и алгоритмы интеллектуального анализа изображений обеспечивают значимое повышение качества поддержки принятия решений, превосходят базовые решения по ключевым показателям и обладают высокой практической значимостью.

Полученные результаты исследований демонстрируют существенный научный и практический вклад результатов диссертационной работы в развитие методов интеллектуального анализа изображений в DICOM-подобных данных.

## ЗАКЛЮЧЕНИЕ

В результате проведенного исследования решена актуальная научно-техническая задача повышения эффективности выполнения поисковых запросов в СППР за счет разработки модели и комплексного применения визуальных и фактографических признаков, представленных в DICOM-подобных форматах.

Проведенный анализ современных моделей и алгоритмов компьютерного зрения, поиска изображений и построения распределенных информационных систем показал, что использование исключительно визуальных признаков не обеспечивает требуемой устойчивости и релевантности результатов в условиях высокой внутриклассовой вариативности и неполноты метаданных. Установлено, что решающим фактором повышения качества поиска является не увеличение глубины нейросетевой архитектуры как таковой, а формирование дискриминативного признакового пространства, учитывающего контекстные и технологические характеристики объектов, а также корректная организация инфраструктуры хранения и индексации данных.

На основе системного анализа разработана модифицированная модель ИНС «DICONDE-YOLO» и комплексированный алгоритм поиска схожих изображений, предусматривающий предобработку DICOM-подобных данных, извлечение и токенизацию фактографической информации, формирование визуальных эмбеддингов, их интеграцию в единое семантическое пространство. Предложенный подход обеспечивает согласованное использование визуальной и фактографической информации и ориентирован на функционирование в составе распределенной СППР.

Таким образом, в результате проведенных исследований:

1. Решена актуальная научно-техническая задача повышения эффективности выполнения поисковых запросов в СППР за счет разработки модели и комплексного применения визуальных и фактографических признаков, представленных в DICOM-подобных форматах.

2. Разработана архитектура ИНС «DICONDE-YOLO» на основе модификации YOLOv8, отличающаяся интеграцией токенизированных метаданных на уровне внутренних представлений нейронной сети и адаптацией к специфике DICOM-подобных форматов. Модификация структуры базовой модели позволила повысить дискриминативность эмбедингов, улучшить показатели локализации и ранжирования и обеспечить устойчивость к вариативности входных данных.

3. Реализована распределенная микросервисная архитектура программного прототипа СППР, включающая алгоритмы векторного поиска, кэширования, балансировки нагрузки, мониторинга и автоматизированной адаптации параметров модели. Прототип интегрирован в единое информационное пространство производственного предприятия и продемонстрировал масштабируемость при росте объема данных и интенсивности запросов без ухудшения качества поиска.

4. Проведена экспериментальная оценка результатов вычислительных экспериментов и опытной эксплуатации разработанной модели и комплексного алгоритма в составе СППР на промышленном предприятии. В условиях опытной эксплуатации зафиксировано сокращение среднего времени анализа дефекта на 61%, снижение доли ложных тревог и значительное уменьшение нагрузки на эксперта при одновременном росте производительности обработки изображений на 150%.

Полученные результаты свидетельствуют о том, что интеграция визуальных и фактографических, использование ANN-индексации и реализация адаптивной архитектуры с управляющим контуром AutoML позволяют сформировать масштабируемую и инвариантную к предметной области модель интеллектуального поиска. Разработанный подход может быть использован при построении вычислительных кластеров и центров обработки данных, ориентированных на анализ больших массивов стандартизированных визуальных данных [100, 101]. Кроме того, на основе проведенного анализа делается вывод о адаптивности разработанных решений к смежным по типу хранимых данных отраслей. Результаты проведенного анализа представлены в приложении Ж.

Таким образом, цель диссертационного исследования достигнута, поставленные научные задачи решены, а разработанные модели, алгоритмы и программная реализация обладают как научной новизной, так и подтвержденной практической значимостью в задачах интеллектуального анализа изображений и поддержки принятия решений.

Дальнейшие рекомендации по результатам выполненных исследований заключаются в широком внедрении комплексированного алгоритма поиска схожих изображений в существующие системы НК, позволяющих сократить время анализа и изображений, уменьшить долю ложных срабатываний за счет учета метаданных и обеспечить воспроизводимость результатов независимо от квалификации персонала. При модернизации систем НК перейти на микросервисную архитектуру с использованием ANN-индексации для хранения и поиска изображений, что обеспечит возможность горизонтального масштабирования на кластерах с использованием инструментов контейнеризации. Кроме того, рекомендуется продолжить исследования в области гибридных признаков пространств, объединяющих визуальные и токенизированные метаданные, что обеспечит дискриминативность представлений объектов при высокой внутриклассовой вариативности и переносимость между предметными областями с использованием стандартизированных форматов. Разработать стандарты обмена данными между СППР различных предметных областей на основе DICOM-подобных форматов, что обеспечит интероперабельность систем НК в рамках единой экосистемы Индустрии 4.0, а также возможность кросс-доменного обучения моделей на разнородных датасетах.

**СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ**

ANN	Approximate Nearest Neighbors
AP	Average Precision
API	Application Programming Interface
ASPP	Atrous Spatial Pyramid Pooling
AutoML	Automated Machine Learning
BoVW	Bag of Visual Words
CIoU	Complete Intersection over Union
CNNH	Convolutional neural network hashing
CPU	Central Processing Unit
DICOM	Digital Imaging and Communications in Medicine
DICONDE	Digital Imaging and Communications for Non- Destructive Evaluation
DICOS	Digital Imaging and Communications in Security
DPSH	Deep pairwise-supervised hashing
ECA	Efficient Channel Attention
ERP	Enterprise Resource Planning
GAP	Global Average Pooling
GPU	Graphics Processing Unit
HoG	Histogram of Oriented Gradients
LBP	Local Binary Patterns
mAP	Mean Average Precision
NDCG	Normalized Discounted Cumulative Gain
NDE	Nondestructive Evaluation
ORB	Oriented FAST and Rotated BRIEF

RAM	Random Access Memory
R-CNN	Region-based Convolutional Neural Networks
SIFT	Scale-Invariant Feature Transform
SPL-UDH	Soft-pseudo-label-based unsupervised deep hashing
SPPF	Spatial Pyramid Pooling Fast
SURF	Speeded-Up Robust Features
UDQH-IQ	Deep Quadruplet Hashing with Isometric Quantization
VRAM	Video Random Access Memory
YOLO	You Only Look Once
БД	База данных
БЗ	База знаний
ВБД	Векторная база данных
ИНС	Искусственная нейронная сеть
ПО	Программное обеспечение
СНС	Сверточная нейронная сеть
СППР	Система поддержки принятия решений
ЭВМ	Электронная вычислительная машина

### СЛОВАРЬ ТЕРМИНОВ

Backbone	<p>Основная архитектура ИНС, используемая для извлечения признаков из входных данных, чаще всего изображений. В контексте задач компьютерного зрения Backbone отвечает за формирование высокоуровневых эмбедингов объектов, которые затем могут быть использованы для детекции, сегментации или классификации.</p>
Аугментация данных	<p>Совокупность методов искусственного расширения обучающего набора данных за счет применения к исходным объектам контролируемых преобразований, сохраняющих их семантическую принадлежность. В контексте машинного обучения и компьютерного зрения аугментация используется для повышения обобщающей способности моделей, снижения переобучения и повышения устойчивости к вариациям входных данных. К типичным операциям аугментации относятся геометрические преобразования, изменения яркости и контрастности, добавление шума, а также вероятностные и параметрические модификации, выполняемые в процессе обучения модели.</p>
Батч	<p>Подмножество обучающих примеров, передаваемых одновременно в нейронную сеть для одной итерации обучения. Использование батчей позволяет балансировать между скоростью обучения и стабильностью градиентов. Малые батчи обеспечивают более частые обновления весов, но с большей</p>

	<p>дисперсией градиентов. Большие батчи уменьшают шум обновлений, но требуют больше памяти и могут замедлять сходимость.</p>
<p>Визуальные признаки</p>	<p>Числовое представление графических данных изображения, извлекаемое из исходного тензора пиксельных значений <math>X_i</math> с помощью сверточной нейронной сети. Визуальные признаки могут варьироваться по числу цветовых каналов, а также формироваться на разных уровнях абстракции и использоваться для описания пространственных структур, текстур, а также контуров изображения.</p>
<p>Гиперпараметры</p>	<p>Настраиваемые параметры алгоритма машинного обучения, значения которых задаются до начала процесса обучения и не оптимизируются непосредственно в ходе обновления весов модели. Гиперпараметры определяют структуру модели и режим ее обучения, включая, в частности, скорость обучения, размер батча, количество эпох, архитектурные параметры нейронной сети, параметры регуляризации и методы аугментации данных. Корректный выбор гиперпараметров оказывает существенное влияние на сходимость обучения, обобщающую способность модели и итоговое качество интеллектуального анализа изображений, что обуславливает необходимость их обоснованного подбора и экспериментальной валидации в рамках диссертационного исследования.</p>

Датасет	Структурированная совокупность данных, используемая для обучения, тестирования и верификации алгоритмов машинного обучения и компьютерного зрения. В контексте обработки изображений каждый элемент датасета представляет собой пару «изображение – метаданные» или «входные данные – эталонная разметка», включающую визуальные признаки, а также при необходимости фактографические, категориальные или текстовые характеристики.
Инференс	Этап функционирования модели машинного обучения или нейронной сети, заключающийся в применении уже обученных параметров модели для получения выходных предсказаний по новым, ранее не использовавшимся входным данным.
Патч	Локальный фрагмент исходного изображения фиксированного размера, выделяемый на этапе предварительной обработки в моделях на основе визуальных трансформеров. Изображение разбивается на неперекрывающиеся или частично перекрывающиеся патчи, каждый из которых преобразуется в вектор признаков и рассматривается как отдельный элемент последовательности, аналогичный токену в трансформерных моделях обработки естественного языка. Такое представление позволяет моделировать глобальные и локальные зависимости между различными областями изображения с использованием механизмов самовнимания (self-attention mechanism),

	обеспечивая формирование компактного и информативного признакового пространства для последующего анализа и принятия решений.
Развертывание приложения	Процесс подготовки, установки и интеграции программного решения в целевую вычислительную среду с целью обеспечения его штатной эксплуатации, включающий настройку аппаратных и программных ресурсов, конфигурацию зависимостей, запуск сервисов инференса и обеспечение взаимодействия с внешними компонентами системы.
Случайный сид	Начальное значение, используемое для инициализации генераторов псевдослучайных чисел в вычислительных алгоритмах и программных средах. Фиксация случайного седа обеспечивает детерминированность стохастических процессов, таких как инициализация весов нейронных сетей, формирование батчей, аугментация данных и порядок выборки, что позволяет воспроизводить результаты вычислительных экспериментов при повторных запусках при идентичной конфигурации аппаратного и программного обеспечения.
Токенизация	Процедура преобразования разнородных фактографических данных $M_i$ в унифицированный числовой вектор $z_i$ фиксированной или переменной размерности, пригодный для совместной обработки с визуальными признаками.

Фреймворк	Программная платформа, представляющая собой набор библиотек, инструментов и правил организации кода, предназначенная для упрощения разработки, тестирования и развертывания приложений. В контексте нейронных сетей фреймворк обеспечивает стандартизированные механизмы построения моделей, обработки данных, обучения, оптимизации и инференса.
Эмбединг	Числовое представление дискретного или категориального объекта в виде вектора фиксированной размерности, формируемое с целью сохранения семантической или структурной близости объектов в соответствующем векторном пространстве.

### СПИСОК ЛИТЕРАТУРЫ

1. Saboonchi, H. Advancements in radiographic evaluation through the migration into NDE 4.0 / H. Saboonchi, D. Blanchette, K. Hayes // Journal of nondestructive evaluation. – 2021. – Vol. 40, Is. 1. – P. 17.
2. Meyendorf, N. NDE 4.0: Progress, promise, and its role to industry 4.0 / N. Meyendorf, N. Ida, J. Vrana, J. Singh // NDT & E International. – 2023. – Vol. 140. – P. 102957.
3. Trampus, P. Symbiosis of industry 4.0 and non-destructive evaluation / P. Trampus // Buletinul Științific. – Kishinev, 2023. – P. 43-51
4. Larobina, M. Thirty years of the DICOM standard / M. Larobina // Tomography. – 2023. – Vol. 9, Is. 5. – P. 1829-1838.
5. ASTM E2339-21: Standard practice for digital imaging and communication in nondestructive evaluation (DICONDE) [Электронный ресурс]. – West Conshohocken,

- РА : ASTM International, 2021. – URL: <https://www.astm.org/e2339-21.html> (дата обращения: 16.02.2026).
6. DICONDE [Электронный ресурс] // Scholarly Community Encyclopedia. – 2023. – URL: <https://encyclopedia.pub/entry/28374> (дата обращения: 16.02.2026).
7. Skrzypkowiak, S. DICOS: part of the solution for an interoperable, open airport architecture / S. Skrzypkowiak // Anomaly Detection and Imaging with X-Rays (ADIX) VI : Proceedings of SPIE. – Bellingham : SPIE, 2021. – Vol. 11738. – P. 24-33.
8. Михайлов, С. С. Искусственный интеллект и его применение в медицине [Электронный ресурс] / С. С. Михайлов // Современные инновации. – 2017. – № 2. – URL: <https://cyberleninka.ru/article/n/iskusstvennyy-intellekt-i-ego-primenenie-v-medicine> (дата обращения: 16.02.2026).
9. Алексеева, М. Г. Искусственный интеллект в медицине / М. Г. Алексеева, А. И. Зубов, М. Ю. Новиков // Международный научно-исследовательский журнал. – 2022. – № 7-2 (121). – С. 10-13.
10. Фершт, В. М. Современные подходы к использованию искусственного интеллекта в медицине / В. М. Фершт, А. П. Латкин, В. Н. Иванова // Территория новых возможностей. Вестник Владивостокского государственного университета. – 2020. – Т. 12, № 1. – С. 121-130.
11. Хусанов, У. А. У. Искусственный интеллект в медицине / У. А. У. Хусанов [и др.] // Science and Education. – 2023. – Т. 4, № 5. – С. 772-782.
12. Рынок нейронных сетей глубокого обучения (DNN) в Азиатско-Тихоокеанском регионе – тенденции отрасли и прогноз до 2028 года [Электронный ресурс] // Data Bridge Market Research. – URL: <https://www.databridgemarketresearch.com/ru/reports/asia-pacific-deep-learning-neural-networks-dnns-market> (дата обращения: 16.02.2026).
13. Ильясов, Б. Г. Структурный подход к построению методологии исследования систем / Б. Г. Ильясов [и др.] // Вестник Самарского государственного технического университета. Серия: Технические науки. – 2017. – № 2 (54). – С. 24-33.

14. Куликов, Г. Г. Формальная модель производственного процесса для организации проектного и производственного менеджмента / Г. Г. Куликов [и др.] // Вестник Казанского государственного технического университета им. А. Н. Туполева. – 2014. – № 4. – С. 175-186.
15. Буракова Е. А. Система поддержки принятия решений при производстве катализатора синтеза углеродных нанотрубок / Е. А. Буракова [и др.] // Вестник ТГТУ. – 2023. – Т. 29, – №2. – С. 187-198.
16. Буракова, Е. А. Поддержка принятия решений при управлении производством углеродных нанотрубок с заданными параметрами / Е. А. Буракова // Вестник Воронежского государственного технического университета. – 2023. – Т. 19, № 2. – С. 61-67.
17. Буракова, Е. А. Информационная поддержка принятия решений при производстве катализатора синтеза углеродных нанотрубок / Е. А. Буракова // Прикаспийский журнал: управление и высокие технологии. – 2021. – № 3 (55). – С. 9-15.
18. Конев, К. А. Машинное обучение для поддержки принятия решений в сфере качества на промышленном предприятии / К. А. Конев // Экономика. Информатика. – 2023. – Т. 50, № 3. – С. 689-703.
19. Родионов, Д. Г. Основные функциональные и технические требования информационно-аналитических систем поддержки принятия решений в инновационной деятельности малых промышленных предприятий / Д. Г. Родионов, Д. А. Алферьев // Вестник Алтайской академии экономики и права. – 2020. – № 9-2. – С. 294-299.
20. Денисов, Д. Ю. Развитие систем поддержки принятия решений на основе искусственного интеллекта в менеджменте российских компаний / Д. Ю. Денисов // Экономические системы. – 2021. – Т. 14, № 4. – С. 29-36.
21. Денисов, Д. Ю. Развитие систем поддержки принятия решений на основе искусственного интеллекта в менеджменте российских компаний / Д. Ю. Денисов // Экономические системы. – 2021. – Т. 14, № 4. – С. 29-36.

22. Лебедев, А. А. Исследование нейросетевых алгоритмов обнаружения объектов на видеоизображениях в медицинских системах прикладного телевидения : дис. ... канд. техн. наук / А. А. Лебедев; [место защиты: Владимир. гос. ун-т им. А. Г. и Н. Г. Столетовых]. – Ярославль, 2022. – 131с.
23. Hossein-Nejad, Z. Detailed review of the scale invariant feature transform (SIFT) algorithm; concepts, indices and applications / Z. Hossein-Nejad, H. Agahi, A. Mahmoodzadeh // Journal of Machine Vision and Image Processing. – 2020. – Vol. 7, Is. 1. – P. 165-190.
24. Bay, H. SURF: Speeded up robust features / H. Bay, T. Tuytelaars, L. Van Gool // European Conference on Computer Vision, Graz, Austria, May 7-13, 2006 : proceedings. – Berlin ; Heidelberg : Springer, 2006. – Pt. 1. – P. 404-417.
25. Aglave, P. Implementation of high performance feature extraction method using oriented fast and rotated brief algorithm / P. Aglave, V. S. Kolkure // International Journal of Research in Engineering and Technology. – 2015. – Vol. 4, Is. 2. – P. 394-397.
26. Ilas, M. E. A new method of histogram computation for efficient implementation of the HOG algorithm / M. E. Ilas, C. A. Ilas // Computers. – 2018. – Vol. 7, Is. 1. – P. 18.
27. Pietikäinen, M. Two decades of local binary patterns: A survey / M. Pietikäinen, G. Zhao // Advances in Independent Component Analysis and Learning Machines / edited by E. Bingham [et al.]. – Amsterdam : Academic Press, 2015. – P. 175-210.
28. Половинченко, М. И. Алгоритмы распознавания лиц / М. И. Половинченко, В. С. Елисеев // Journal of Advanced Research in Technical Science. – 2021. – № 25. – С. 51-55.
29. Чирышев, Ю. В. Распознавание круглых лесоматериалов с помощью случайных решающих деревьев и гистограммы направленных градиентов / Ю. В. Чирышев, А. С. Атаманова // Фундаментальные исследования. – 2017. – № 1. – С. 124-128.
30. Faraji, M. Bag-of-visual-words, its detectors and descriptors; a survey in detail / M. Faraji, J. Shanbehzadeh // Advances in Computer Science: an International Journal. – 2015. – Vol. 4, Is. 2. – P. 8-20.

31. Agrawal, H. Canny edge detection: A comprehensive review / H. Agrawal, K. Desai // International Journal of Technical Research & Science. – 2024. – Vol. 9, Is. Spl. – P. 27-35.
32. Dharampal, V. M. Methods of image edge detection: A review / V. M. Dharampal // Journal of Electrical and Electronics Systems. – 2015. – Vol. 4, Is. 2. – P. 2332-0796.
33. Кошелева, Д. Д. Преобразование Фурье и быстрое преобразование Фурье / Д. Д. Кошелева, А. В. Доронина // Инновации. Наука. Образование. – 2021. – № 38. – С. 626-632.
34. Jiang, P. A Review of YOLO algorithm developments / P. Jiang, D. Ergu, F. Cai, B. Ma // Procedia Computer Science. – 2022. – Vol. 199. – P. 1066-1073.
35. Tan, M. EfficientNet: Rethinking model scaling for convolutional neural networks / M. Tan, Q. Le // International Conference on Machine Learning, Long Beach, CA, USA, 9-15 June 2019 : proceedings. – 2019. – P. 6105-6114.
36. Tan, M. EfficientDet: Scalable and efficient object detection / M. Tan, R. Pang, Q. V. Le // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2020. – P. 10781-10790.
37. Bharati, P. Deep learning techniques—R-CNN to mask R-CNN: a survey / P. Bharati, A. Pramanik // Computational Intelligence in Pattern Recognition : proceedings of the International Conference (CIPR 2019). – Singapore : Springer, 2019. – P. 657-668.
38. Liu, W. SSD: Single shot multibox detector / W. Liu, D. Anguelov, C. Szegedy [et al.] // European Conference on Computer Vision, Amsterdam, The Netherlands, October 11-14, 2016 : proceedings. – Cham : Springer International Publishing, 2016. – Pt. 1. – P. 21-37.
39. Sharir, G. An image is worth 16x16 words, what is a video worth? [Electronic resource] / G. Sharir, A. Noy, L. Zelnik-Manor // arXiv preprint arXiv:2103.13915. – 2021. – URL: <https://arxiv.org/abs/2103.13915> (дата обращения: 16.02.2026).
40. Liu, Z. Swin transformer: Hierarchical vision transformer using shifted windows / Z. Liu, Y. Lin, Y. Cao [et al.] // Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). – 2021. – P. 10012-10022.

41. Cai, Y. Medical image retrieval based on convolutional neural network and supervised hashing / Y. Cai, Y. Li, C. Qiu, J. Ma, X. Gao // IEEE Access. – 2019. – Vol. 7. – P. 51877-51885.
42. Li, W. J. Feature learning based deep supervised hashing with pairwise labels [Electronic resource] / W. J. Li, S. Wang, W. C. Kang // arXiv preprint arXiv:1511.03855. – 2015. – URL: <https://arxiv.org/abs/1511.03855> (дата обращения: 16.02.2026).
43. Qin, Q. Unsupervised deep quadruplet hashing with isometric quantization for image retrieval / Q. Qin, L. Huang, Z. Wei, J. Nie, K. Xie, J. Hou // Information Sciences. – 2021. – Vol. 567. – P. 116-130.
44. Hu, Q. Pseudo label based unsupervised deep discriminative hashing for image retrieval / Q. Hu, J. Wu, J. Cheng, L. Wu, H. Lu // Proceedings of the 25th ACM International Conference on Multimedia. – New York : ACM, 2017. – P. 1584-1590.
45. Курочкин, А. Г. Схема комплексирования данных для системы управления мобильным роботом / А. Г. Курочкин [и др.] // Extreme Robotics : сборник статей. – [Б. м.], 2018. – Т. 1, № 1. – С. 465-471.
46. Потапов, А. А. Анализ способов комплексирования изображений, формируемых многодиапазонными радиолокационными станциями с синтезированной апертурой / А. А. Потапов, В. А. Кузнецов, Е. А. Аликулов // Известия высших учебных заведений России. Радиоэлектроника. – 2021. – Т. 24, № 3. – С. 6-21.
47. Карташов, В. М. Комплексирование изображений при обнаружении беспилотных летательных аппаратов / В. М. Карташов [и др.]. – [Б. м.], 2020.
48. Старовойтов, В. В. Нормализация данных в машинном обучении / В. В. Старовойтов, Ю. И. Голуб. – [Б. м.], 2021.
49. Хузяханов, Е. И. Эффективные методы предобработки данных: как подготовить данные для машинного обучения / Е. И. Хузяханов, С. С. Шарко, П. М. Поляков // Современные научные исследования и инновации. – 2024. – № 11 (163). – URL: <https://web.snauka.ru/issues/2024/11/102920> (дата обращения: 16.02.2026).

- 50.Karmaker, S. K. AutoML to date and beyond: challenges and opportunities / S. K. Karmaker, M. M. Hassan, M. J. Smith, L. Xu // ACM Computing Surveys (CSUR). – 2021. – Vol. 54, Is. 8. – P. 1-36.
- 51.Weaver, M. AutoML for multi-label classification: overview and empirical evaluation / M. Weaver, A. Tornede, F. Mohr, E. Hullermeier // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2021. – Vol. 43, Is. 9. – P. 3037-3054.
- 52.Коваленко, Л. А. Универсальная модель архитектуры краудсорсинговой системы разметки и подготовки медицинских данных / Л. А. Коваленко [и др.] // Научно-технический вестник информационных технологий, механики и оптики. – 2025. – Т. 25, № 5. – С. 844-855.
- 53.Sohan, M. A review on YOLOv8 and its advancements / M. Sohan, R. T. Sai, C. V. Rami Reddy // International Conference on Data Intelligence and Cognitive Informatics (ICDICI 2023) : proceedings. – Singapore : Springer, 2024. – P. 529-545.
- 54.He, K. Spatial pyramid pooling in deep convolutional networks for visual recognition / K. He, X. Zhang, S. Ren, J. Sun // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2015. – Vol. 37, Is. 9. – P. 1904-1916.
- 55.Ahmed, K. R. DSteelNet: a real-time parallel dilated CNN with atrous spatial pyramid pooling for detecting and classifying defects in surface steel strips / K. R. Ahmed // Sensors. – 2023. – Vol. 23, Is. 1. – P. 544.
- 56.Gholamalinezhad, H. Pooling methods in deep neural networks, a review [Electronic resource] / H. Gholamalinezhad, H. Khosravi // arXiv preprint arXiv:2009.07485. – 2020. – URL: <https://arxiv.org/abs/2009.07485> (дата обращения: 16.02.2026).
- 57.Kolarik, M. Upsampling algorithms for autoencoder segmentation neural networks: a comparison study / M. Kolarik, R. Burget, K. Riha // International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), Dublin, Ireland, 2019 : proceedings. – IEEE, 2019. – P. 1-5.
- 58.Li, W. DC-YOLO: an improved field plant detection algorithm based on YOLOv7-tiny / W. Li, Y. Zhang // Scientific Reports. – 2024. – Vol. 14, Is. 1. – P. 26430.

59. Wang, H. DC-YOLO: A dual channel YOLO for small object defect detection of circuit boards / H. Wang, X. Song // International Conference on Computer Vision, Image and Deep Learning (CVIDL), 2024 : proceedings. – IEEE, 2024. – P. 1292-1296.
60. Pei, S. Detection of hydrophobicity grade of composite insulators based on MDC-YOLO algorithm / S. Pei, W. Wang, C. Hu, H. Sun, K. Li, M. Wu, B. Lan // Energy Science & Engineering. – 2025. – Vol. 13, Is. 6. – P. 3360-3370.
61. Wang, Q. ECA-Net: Efficient channel attention for deep convolutional neural networks / Q. Wang, B. Wu, P. Zhu, W. Zuo, P. Li, Q. Hu // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2020. – P. 11534-11542.
62. Жаворонкова, О. Г. Подготовка данных и оптимизация обучения нейросетей в условиях многопоточности / О. Г. Жаворонкова // Информационные технологии в науке, промышленности и образовании : сборник статей. – [Б. м.], 2025.
63. Yaseen, M. What is YOLOv9: An in-depth exploration of the internal features of the next-generation object detector [Electronic resource] / M. Yaseen // arXiv preprint arXiv:2409.07813. – 2024. – URL: <https://arxiv.org/abs/2409.07813> (дата обращения: 16.02.2026).
64. Alif, M. A. R. YOLOv1 to YOLOv10: A comprehensive review of YOLO variants and their application in the agricultural domain [Electronic resource] / M. A. R. Alif, M. Hussain // arXiv preprint arXiv:2406.10139. – 2024. – URL: <https://arxiv.org/abs/2406.10139> (дата обращения: 16.02.2026).
65. Khanam, R. YOLOv11: An overview of the key architectural enhancements [Electronic resource] / R. Khanam, M. Hussain // arXiv preprint arXiv:2410.17725. – 2024. – URL: <https://arxiv.org/abs/2410.17725> (дата обращения: 16.02.2026).
66. Jegham, N. YOLO evolution: A comprehensive benchmark and architectural review of YOLOv12, YOLO11, and their previous versions [Electronic resource] / N. Jegham, C. Y. Koh, M. Abdelatti, A. Hendawi // arXiv preprint arXiv:2411.00201. – 2024. – URL: <https://arxiv.org/abs/2411.00201> (дата обращения: 16.02.2026).

67. Корчагин, В. Д. Поиск дефектов в сварных металлических соединениях при помощи технологий компьютерного зрения / В. Д. Корчагин // Новые материалы, оборудование и технологии в промышленности : материалы Международной научно-технической конференции молодых ученых, Могилев, 24–25 октября 2024 года. – Могилев : Белорусско-Российский университет, 2024. – С. 197-198.
68. Корчагин, В. Д. DICONDE как инструмент повышения обобщающих способностей в нейросетевом обучении / В. Д. Корчагин // Программная инженерия: современные тенденции развития и применения (ПИ-2024): сборник материалов VIII Всероссийской научно-практической конференции с международным участием, Курск, 17 октября 2024 года. – Курск: ЗАО «Университетская книга», 2024. – С. 171-176.
69. Корчагин, В. Д. Automl как практико-ориентированный инструмент адаптации моделей обработки данных неразрушающего контроля / В. Д. Корчагин // Актуальные проблемы современной науки и производства: Материалы IX Всероссийской научно-технической конференции «Рязанский государственный радиотехнический университет имени В.Ф. Уткина», 27–29 ноября 2024 года. – Рязань: Рязанский государственный радиотехнический университет им. В.Ф. Уткина, 2024. – С. 230-236.
70. Корчагин, В. Д. Разметка DICONDE-изображений для обучения искусственных нейронных сетей / В. Д. Корчагин // Цифровые, компьютерные и информационные технологии в науке и образовании: Сборник статей II Межрегиональной научно-практической конференции с международным участием, Брянск, 14–15 ноября 2024 года. – Брянск: Брянский государственный университет им. акад. И.Г. Петровского, 2025. – С. 104-109.
71. Корчагин, В. Д. Качество обучения искусственных нейронных сетей для обработки радиографических изображений / В. Д. Корчагин // Современные технологии в науке и образовании - СТНО-2025: Сборник трудов VIII Международного научно-технического форума. В 10-ти томах, Рязань, 04–06 марта 2025 года. – Рязань:

Рязанский государственный радиотехнический университет им. В.Ф. Уткина, 2025.  
– С. 125-131.

72. Кувшинников, В. С. Практика применения цифровых двойников при симуляции операций контроля в промышленной радиографии / В. С. Кувшинников, Е. Е. Ковшов, А. В. Сударев, В. Д. Корчагин // *Фундаментальные исследования и инновационные технологии в машиностроении : научные труды IX Международной научной конференции, посвященной 135-летию со дня рождения профессора Хрущова Михаила Михайловича, Москва, 11–13 ноября 2025 года.* – Москва : ИМАШ РАН, 2025. – С. 173-176.
73. Корчагин, В. Д. Синтез технологий виртуальной реальности и компьютерного зрения в симуляторе промышленной радиографии / В. Д. Корчагин, В. С. Кувшинников, Е. Е. Ковшов // *Известия Юго-Западного государственного университета. Серия: Управление, вычислительная техника, информатика. Медицинское приборостроение.* – 2024. – Т. 14, № 4. – С. 98-115. – DOI 10.21869/2223-1536-2024-14-4-98-115.
74. Корчагин, В. Д. Анализ современных SOTA-архитектур искусственных нейронных сетей для решения задач классификации изображений и детекции объектов / В. Д. Корчагин // *Программные системы и вычислительные методы.* – 2023. – № 4. – С. 73-87. – DOI 10.7256/2454-0714.2023.4.69306.
75. Корчагин, В. Д. Критериальный анализ моделей обработки данных радиационного неразрушающего контроля / В. Д. Корчагин, В. С. Кувшинников, Е. Е. Ковшов // *International Journal of Open Information Technologies.* – 2024. – Т. 12, № 4. – С. 23-31.
76. Корчагин, В. Д. Концепция разработки и применения проблемно-ориентированной системы поддержки принятия решений в промышленном производстве / В. Д. Корчагин // *Автоматизация. Современные технологии.* – 2024. – Т. 78, № 7. – С. 291-296. – DOI 10.36652/0869-4931-2024-78-7-291-296.

- 77.Свидетельство о государственной регистрации программы для ЭВМ № 2024688377 Российская Федерация. Разметка DICONDE-изображений / В. Д. Корчагин, В. С. Кувшинников ; правообладатель ФГБОУ ВО «Рязанский государственный радиотехнический университет им. В.Ф. Уткина». – № 2024684754 ; заявл. 15.10.2024 ; опубл. 27.11.2024.
- 78.Корчагин, В. Д. Модификация нейронной сети для автоматического выявления дефектов сварных соединений на цифровых радиографических изображениях / В. Д. Корчагин, А. В. Сударев // Цифровые, компьютерные и информационные технологии в науке и образовании: сборник статей III Межрегиональной научно-практической конференции с международным участием (Брянск, 14 ноября 2025 года). – Брянск : РИСО БГУ, 2026. – Т. 3. – С. 70-75.
- 79.Сударев, А. В. Повышение качества обучающих выборок результатов радиационного неразрушающего контроля / А. В. Сударев, В. Д. Корчагин // Цифровые, компьютерные и информационные технологии в науке и образовании: сборник статей III Межрегиональной научно-практической конференции с международным участием (Брянск, 14 ноября 2025 года). – Брянск : РИСО БГУ, 2026. – Т. 2. – С. 137-140.
- 80.Korchagin, V., Kovshov, E., Kuvshinnikov, V. (2026). Artificial Neural Networks in Non-Destructive Testing and Evaluation: A Novel Approach to Radiographic Inspection. In: Diveev, A., Fomichev, V., Ilin, A., Zelinka, I., Sofronova, E. (eds) Intelligent Systems. INTELS 2024. Communications in Computer and Information Science, vol 2603. Springer, Cham. [https://doi.org/10.1007/978-3-032-04758-8\\_7](https://doi.org/10.1007/978-3-032-04758-8_7)
- 81.Grafana [Электронный ресурс] // Grafana Labs. – URL: <https://grafana.com/> (дата обращения: 16.02.2026).
- 82.Prometheus [Электронный ресурс] // The Linux Foundation. – URL: <https://prometheus.io/> (дата обращения: 16.02.2026).
- 83.PyTorch [Электронный ресурс] // The Linux Foundation. – URL: <https://pytorch.org/> (дата обращения: 16.02.2026).

84. OpenCV [Электронный ресурс] // OpenCV team. – URL: <https://opencv.org/> (дата обращения: 16.02.2026).
85. Protocol Buffers [Электронный ресурс] // Google. – URL: <https://protobuf.dev/> (дата обращения: 16.02.2026).
86. gRPC [Электронный ресурс] // Cloud Native Computing Foundation. – URL: <https://grpc.io/> (дата обращения: 16.02.2026).
87. Jin, B. Designing Web APIs: Building APIs That Developers Love / B. Jin, S. Sahni, A. Shevat. – Sebastopol : O'Reilly Media, 2018. – 230 p.
88. Rose, O. JSON & APIs: A Developer's Guide to Data Exchange / O. Rose. – [Б. м.] : [б. и.], 2025. – 282 p.
89. Docker Docs [Электронный ресурс] // Docker Inc. – URL: <https://docs.docker.com/> (дата обращения: 16.02.2026).
90. Kubernetes Documentation [Электронный ресурс] // The Linux Foundation. – URL: <https://kubernetes.io/docs/home/> (дата обращения: 16.02.2026).
91. Milvus [Электронный ресурс] // LF Projects, LLC. – URL: <https://milvus.io/> (дата обращения: 16.02.2026).
92. Redis [Электронный ресурс] // Redis Ltd. – URL: <https://redis.io/docs/latest/> (дата обращения: 16.02.2026).
93. Apache Pulsar [Электронный ресурс] // The Apache Software Foundation. – URL: <https://pulsar.apache.org/> (дата обращения: 16.02.2026).
94. NATS Docs [Электронный ресурс] // CNCF. – URL: <https://docs.nats.io/> (дата обращения: 16.02.2026).
95. Меджидов, Р. Г. Анализ многоколоночных индексов баз данных / Р. Г. Меджидов // Актуальные проблемы прикладной математики, информатики и механики : сборник трудов Международной научной конференции, Воронеж, 17-19 декабря 2019 г. – Воронеж : Издательство "Научно-исследовательские публикации", 2019. – С. 420-422. — Примечание: Выходные данные уточнены по сайту конференции.

96. Elasticsearch [Электронный ресурс] // Elasticsearch B.V. – URL: <https://www.elastic.co/elasticsearch> (дата обращения: 16.02.2026).
97. NGINX Documentation [Электронный ресурс] // F5, Inc. – URL: <https://docs.nginx.com/> (дата обращения: 16.02.2026).
98. Кутырев А. И. Сверточная нейронная сеть для сегментации цветков яблони на изображениях // Аграрная наука Евро-Северо-Востока. – 2024. – Т. 25. – №. 5. – С. 949-961.
99. Семенихин С. В., Денисова Л. А. Метод машинного обучения ранжированию на основе модифицированного генетического алгоритма для метрики NDCG // Омский научный вестник. – 2016. – №. 3 (147). – С. 100-104.
100. Kuvshinnikov V., Kovshov E., Korchagin V. Digital Educational Resources in Industrial Radiography // 2024 7th International Conference on Information Technologies in Engineering Education (Inforino). – IEEE, 2024. – С. 1-5. – DOI: 10.1109/Inforino60363.2024.10551923.
101. Kuvshinnikov V., Kovshov E., Korchagin V. 3D Modeling for a Flexible Detector in VR Radiography Simulator // 2024 7th International Conference on Information Technologies in Engineering Education (Inforino). – IEEE, 2024. – С. 1-6. – DOI: 10.1109/Inforino60363.2024.10551886

## ПРИЛОЖЕНИЕ А. ПСЕВДОКОД КОМПЛЕКСНОГО АЛГОРИТМА ПОИСКА ИЗОБРАЖЕНИЙ И СТРУКТУРНЫХ СОСТАВЛЯЮЩИХ

```

function ComplexSearchAlgorithm(D, D_q, K, H)
    # AutoML выбор оптимальной архитектуры
    A* ← AutoML_Search(D_train, D_val, H)
    # z_q — токенизация метаданных запроса D_q
    z_q ← Tokenize(M_q)
    # f_q — извлечение визуальных признаков запроса с учетом
    токенизированных метаданных
    f_q ← FeatureExtractor (X_q, z_q, A)
    # D_d — предварительная фильтрация базы по метаданным запроса
    D_d ← FilterByMetadata (D, M_q)
    # Распределение изображений по серверам для балансировки нагрузки
    foreach D_i in D_d do
        s_j ← AssignServer(D_i, N_s) # s_j — выбранный сервер для D_i
        P(s_j) = (L_avg - L_j) / sum((L_avg - L_k)) # вероятность
    закрепления сервером
    end for
    # Извлечение признаков для всех кандидатов после фильтрации
    foreach D_i in D_d do
        z_i ← Tokenize(M_i) # токенизация метаданных
        f_i ← FeatureExtractor (X_i, z_i, A_i) # извлечение признаков
        InsertToVectorIndex(f_i) # добавление в векторный индекс для
    поиска
    end for
    # Поиск K наиболее релевантных изображений
    R_K ← VectorSearch(f_q, K)
    return R_K

```

**end function**

Псевдокод выбора оптимальной архитектуры AutoML:

```

function AutoML_Search(D_train, D_val, H):
    #Инициализация лучшей метрики
    best_score ← -∞
    #Инициализация лучшей архитектуры
    best_arch ← None
    #Перебор всех конфигураций гиперпараметров
    for h in H:
        #Построение модели с заданной конфигурацией
        model ← Build_DICONDE_YOLO(h)
        #Обучение модели на тренировочной выборке
        Train(model, D_train)
        #Извлечение признаков из валидационной выборки
        f_val ← ExtractFeatures(D_val, model)
        #Вычисление метрики качества поиска по метрике NDCG
        score ← Evaluate(f_val, D_val)
        #Если текущее решение лучше, то обновляются лучшие варианты
        if score > best_score:
            best_score ← score
            best_arch ← h
    #Возврат оптимальной архитектуры
    return best_arch
end function

```

Псевдокод построения модели по гиперпараметрам Build\_DICONDE\_YOLO:

```

function Build_DICONDE_YOLO(h):

```

```

# Создание базовой структуры YOLO
model ← YOLOv8_Base()
# Настройка ASPP блоков
model.ASPP ← h.ASPP_count
# Настройка DC/MDC блоков
model.DC_MDC ← h.DC_MDC_kernel
# Настройка ECA внимания
if h.ECA_enabled:
    model.ECA ← True
else
    model.ECA ← False
# Настройка размерности эмбединга метаданных
model.MetaEmbeddingSize ← h.EmbeddingDim
return model
end function

```

Псевдокод обучения модели Train:

```

function Train(model, D_train):
    # Разделение данных на батчи
    Batches ← CreateBatches(D_train)
    # Итеративное обучение
    for epoch = 1 in MaxEpochs do
        for batch in Batches do
            X_batch, M_batch ← batch
            # Токенизация метаданных
            z_batch ← Tokenize(M_batch)
            # Прямой проход
            f_batch ← model.Forward(X_batch, z_batch)

```

```

# Вычисление функции потерь
loss ← ComputeLoss(f_batch, X_batch.labels)
# Обратное распространение ошибки
Backpropagate(loss)
# Обновление весов
UpdateWeights(model)
end for
end for
return model
end function

```

Псевдокод вычисления метрики качества Evaluate:

```

function Evaluate(f_val, D_val):
  #Индексирование признаков
  Index ← BuildVectorIndex(f_val)
  #Вычисление NDCG для каждого запроса
  Scores ← []
  for D_q in D_val do
    f_q ← f_val[D_q]
    R_K ← VectorSearch(f_q, K=10)
    score ← ComputeNDCG(R_K, D_q.labels)
    Append(Scores, score)
  end for
  #Среднее значение по всем запросам
  mean_score ← Mean(Scores)
  return mean_score
end function

```

Псевдокод преобразования метаданных в вектор Tokenize:

```
function Tokenize(M_i):  
  z ← []  
  for m in M_i do  
    if m.type == "numerical":  
      z_m ← Standardize(m)  
    else if m.type == "categorical":  
      z_m ← OneHotEncode(m)  
    else if m.type == "text" then  
      z_m ← TextEmbedding(m)  
    end if  
    Append(z, z_m)  
  end for  
  return Concatenate(z)  
end function
```

**ПРИЛОЖЕНИЕ Б. СВИДЕТЕЛЬСТВО О РЕГИСТРАЦИИ ПРОГРАММ ДЛЯ ЭВИ «РАЗМЕТКА DICONDE-ИЗОБРАЖЕНИЙ»**

РОССИЙСКАЯ ФЕДЕРАЦИЯ



**СВИДЕТЕЛЬСТВО**

о государственной регистрации программы для ЭВМ

**№ 2024688377**

**«Разметка DICONDE-изображений»**

Правообладатель: **Корчагин Валерий Дмитриевич (RU)**

Авторы: **Корчагин Валерий Дмитриевич (RU), Кувшинников Владимир Сергеевич (RU)**



Заявка № **2024684754**

Дата поступления **15 октября 2024 г.**

Дата государственной регистрации

в Реестре программ для ЭВМ **27 ноября 2024 г.**

Руководитель Федеральной службы  
по интеллектуальной собственности

**Ю.С. Зубов**

## ПРИЛОЖЕНИЕ В. ПРИМЕРЫ ДЕТЕКЦИИ ДЕФЕКТОВ С ПРИМЕНЕНИЕМ YOLOV8 И «DICONDE-YOLO»

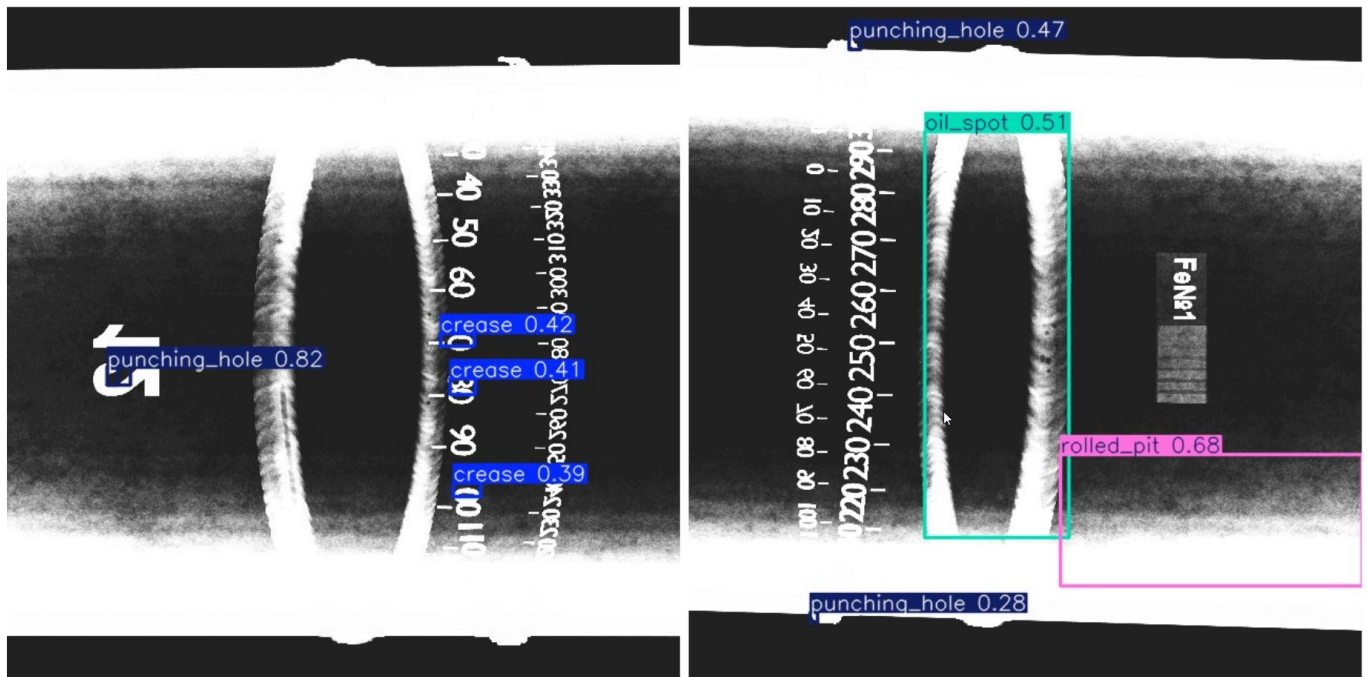


Рис. В.1 Примеры детекции дефектов с помощью YOLOv8

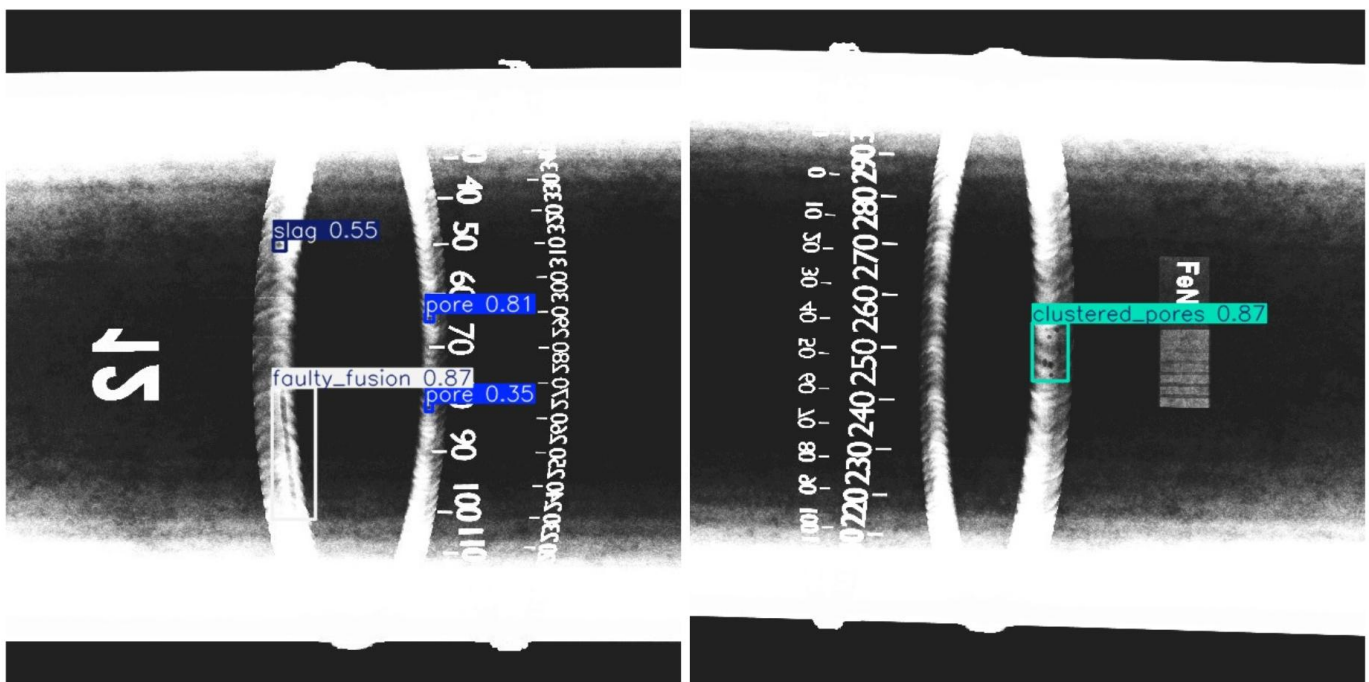




Рис. В.2 Примеры детекции дефектов с помощью «DICONDE-YOLO»


# ПРИЛОЖЕНИЕ Г. ИНТЕРФЕЙС РАЗРАБОТАННОГО ПРОТОТИПА СППР



## СППР DICONDE

>  Пример работы с Redis-кэшем


DICONDE Viewer
🔍 Анализ Признаков
💬 Ассистент и Поиск



## Виртуальный Ассистент

Задайте вопрос о материале/изображении:

Спросить Ассистента




## Поиск Похожих Случаев


Найти похожие случаи в Milvus




Найдено 4 похожих случаев.

	Расстояние	Файл	Класс дефекта	Время ответа
0	0.9999	Обр7-4Ф.dcm	crease	85.2 мс
1	0.8801	12ВХТ2-29.dcm	water_spot, crease, welding_line	48.0 мс
2	0.8735	12ВХТ3-25.dcm	oil_spot, crease, rolled_pit	42.6 мс
3	0.8695	12ВХТ1-12.dcm	oil_spot, crease, water_spot, water_spot	77.5 мс

Рис. Г.1 Пример поиска схожих изображений в БД

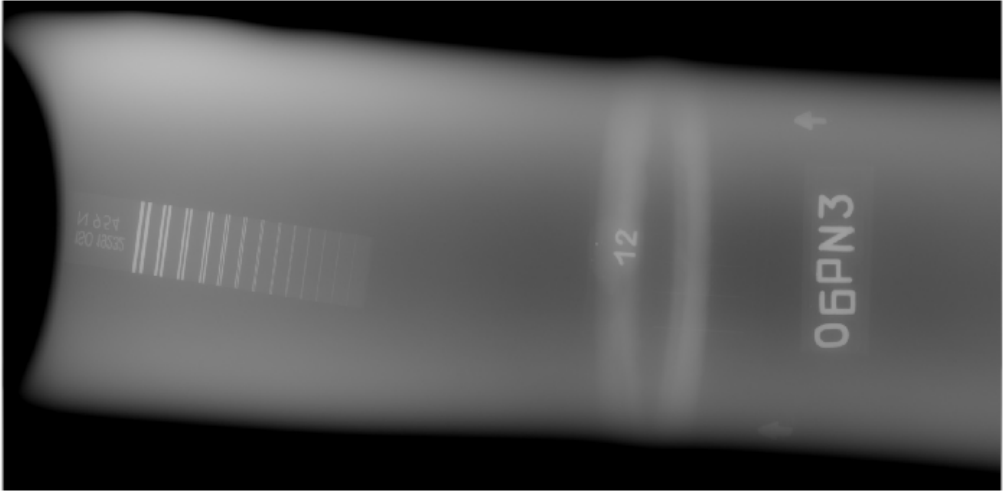
 **СППР DICONDE**

>  Пример работы с Redis-кэшем

 DICONDE Viewer  Анализ Признаков  Ассистент и Поиск

## Просмотр и Редактирование файла: IM00010

### Исходное DICONDE Изображение



**Редактировать существующие теги**

Patient's Name:  
\*\*\*\*\*

Patient ID:  
\*\*\*\*\*

Patient's Birth Date:  
\*\*/\*\*/\*\*

Patient's Sex:  
?

Рис. Г.2 Пример функционала просмотра DICONDE-изображений

**СППР DICONDE**

Пример работы с Redis-кэшем

Сохранить тестовые данные в кэш

Получить тестовые данные

✓ Данные получены. Время ответа с кэшем и фильтром: 43 мс

✓ Данные получены. Время ответа без кэша и фильтра: 193 мс

Удалить тестовые данные

DICONDE Viewer   Анализ Признаков   Ассистент и Поиск

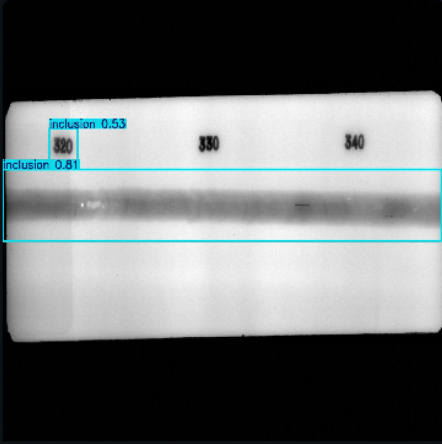
## Результаты Обнаружения Признаков (DICONDE-YOLO)

Анализ завершен. Найдено 2 признаков.

### Сводка Обнаруженных Объектов

Класс	Уверенность	BBox (x1,y1,x2,y2)
Inclusion	0.8145	[4, 467, 1273, 673]
Inclusion	0.5321	[234, 211, 287, 366]

### Визуализация признаков



Обнаруженные признаки

✓ Добавить этот случай в базу данных (Milvus)

Рис. Г.3 Пример функционала анализа признаков и времени получения тестовых данных

**Настройки Системы**

Порог уверенности DICONDE-YOLO: 0,50

Порог уверенности DICONDE-YOLO: 0.50

Порог сходства: 0.00

CUDA недоступен. Используется CPU.

Модель DICONDE-YOLO загружена. Устройство: cpi

Автоматически выбран слой 159 для извлечения векторов признаков.

**DB Statuses**

✓ Milvus Connected (Dim: 1024)

**Загрузка DICONDE**

Загрузите DICONDE файл(ы)

Drag and drop files here  
Limit 200MB per file

Browse files

IM00770  
0.5MB

???

Study Description: \*\*\*\*\*

Series Description: \*\*\*\*\*

Manufacturer: \*\*\*\*\*

Instance Creation Date: \*\*/\*\*/\*\*

Instance Creation Time: \*\*.\*.\*\*

SOP Class UID: \*\*\*\*\*

SOP Instance UID: \*\*\*\*\*

Сохранить изменения

**Редактирование Метаданных**

**Добавить новый тег**

Имя нового тега:

Patient's Name

Patient ID

Patient's Birth Date

Patient's Sex

Modality

Study Description

Series Description

Manufacturer: Choose options

Удалить выбранные теги

Рис. Г.4 Пример блока настроек и редактирования метаданных

**ПРИЛОЖЕНИЕ Д. АКТ О ВНЕДРЕНИИ В АО «НИКИМТ-АТОМСТРОЙ»**

НИКИМТ-АТОМСТРОЙ  
РОСАТОМ

ОРГАНИЗАЦИЯ АО АСЭ

**Акционерное общество  
«Научно-исследовательский и  
конструкторский институт  
монтажной технологии - Атомстрой»  
АО «НИКИМТ-Атомстрой»**

Алтуфьевское шоссе, д. 43, стр. 2  
Москва, 127410  
Телефон (495) 411-65-50, факс (495) 411-65-52

E-mail: [post@atomrus.ru](mailto:post@atomrus.ru)  
ОКПО 08621486, ОГРН 5087746235836  
ИНН 7715719854, КПП 771501001

14.11.2025 № б/н

"На № от

**«УТВЕРЖДАЮ»**

Начальник учебно-аттестационного  
центра НИКИМТ



/Е.Ю. Аленичева/

**АКТ ВНЕДРЕНИЯ  
результатов диссертационной работы  
на соискание учёной степени кандидата технических наук  
Корчагина Валерия Дмитриевича**

Алгоритмы, модели и методы, предложенные в диссертационной работе Корчагина Валерия Дмитриевича, связанные с решением научной задачи разработки комплексного алгоритма поиска и анализа радиографических изображений в системе поддержки принятия решений, рекомендованы к применению для оценки результатов радиационного контроля сварных соединений и использования в Учебно-аттестационном центре НИКИМТ, профессиональная деятельность которого связана, в том числе, с разработкой и внедрением образовательных программ, методик для обучения специалистов неразрушающего контроля.

Проведённые тестовые испытания и опытная эксплуатация в составе прикладного программного обеспечения для расшифровки результатов радиационного неразрушающего контроля подтвердили корректность работы всех программных модулей и их совместимость с соответствующими программами дополнительного профессионального образования, реализуемые учебно-аттестационным центром.

Автоматизация процесса поиска и семантического анализа изображений позволила повысить качество подготовки обучаемых, а разработанный комплексный алгоритм и сверточная нейронная сеть демонстрируют высокую эффективность при детекции дефектов сварных соединений на всех этапах их жизненного цикла.

Главный специалист,  
кандидат химических наук

В.Н. Соловьёв

## ПРИЛОЖЕНИЕ Е. СПРАВКА ОБ ИСПОЛЬЗОВАНИИ В РХТУ ИМ. Д. И. МЕНДЕЛЕЕВА



МИНОБРАЗОВАНИЯ РОССИИ  
Федеральное государственное  
бюджетное образовательное  
учреждение высшего образования  
**«Российский химико-технологический  
университет имени Д.И. Менделеева»**

Миусская пл., д. 9, Москва, 125047  
Тел.: +7 (499) 978-86-60; Факс: +7 (495) 609-29-64  
E-mail: pochta@muctr.ru; https://www.muctr.ru  
ОКПО 02066492; ОГРН 1027739123224  
ИНН/КПП 7707072637/770701001

17.10.2025 № 8/Н  
на № \_\_\_\_\_ от \_\_\_\_\_

### Справка об использовании результатов диссертационной работы

Теоретические решения, разработанные в рамках диссертационного исследования Корчагина Валерия Дмитриевича, реализованы в составе прототипа программно-алгоритмического комплекса «СПИР DICONDE», предназначенного для комплексирования алгоритмов обработки DICONDE-изображений в системе поддержки принятия решений. Разработанные решения успешно апробированы и находят практическое применение в учебном процессе и IT-исследованиях Передовой инженерной школы химического инжиниринга и машиностроения РХТУ им. Д.И. Менделеева, в том числе, по направлению подготовки 09.03.02 – «Информационные системы и технологии», профиль подготовки – «Программная инженерия и наука о данных». Результаты диссертационного исследования используются при проведении лекций, практических занятий и выполнении лабораторных работ в рамках учебных дисциплин: «Введение в инженерную деятельность», «Основы алгоритмизации и программирования», «Объектно-ориентированное программирование на языке C++», «Индустриальное программирование».

Директор ПИШ ХИМ

И.С. Сиротин



**ПРИЛОЖЕНИЕ Ж. АНАЛИЗ ПОТРЕБНОСТИ СМЕЖНЫХ ПРОМЫШЛЕННЫХ ОТРАСЛЕЙ В  
РЕАЛИЗАЦИИ ФУНКЦИОНАЛЬНЫХ ВОЗМОЖНОСТЕЙ СППР**

Отрасль	Критерий									
	Детекция объектов	Поиск по сходству	Масштабируемость системы	Отказоустойчивость системы	Применение нейросетей	Микросервисная архитектура	Балансировка нагрузки	Векторные БД	Обработка потоковых изображений	Оценка
Цифровое производство	Есть	Частично	Есть	Есть	Есть	Есть	Есть	Нет	Есть	15
Симуляторы и обучающие программы	Частично	Частично	Есть	Частично	Есть	Редко	Частично	Нет	Есть	11
Робототехника	Есть	Нет	Частично	Частично	Нет	Редко	Частично	Нет	Есть	8
VR-системы	Частично	Нет	Частично	Нет	Есть	Редко	Есть	Нет	Есть	9
Промышленные системы управления (АСУ ТП)	Нет	Нет	Есть	Критично	Частично	Частично	Частично	Нет	Частично	9
Системы электронного документооборота (СЭД)	Нет	Есть	Есть	Есть	Частично	Есть	Есть	Нет	Нет	11
Химическое производство	Частично	Нет	Частично	Критично	Частично	Редко	Частично	Нет	Нет	8
Автомобильная промышленность	Есть	Частично	Есть	Есть	Есть	Есть	Есть	Частично	Есть	16
Прикладное ПО для НК	Критично	Критично	Есть	Частично	Критично	Есть	Есть	Есть	Критично	21