ЯРОСЛАВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ им. П.Г. ДЕМИДОВА

На правах рукописи

Хрящев Владимир Вячеславович

СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ ВРАЧЕБНЫХ РЕШЕНИЙ НА ОСНОВЕ АНАЛИЗА ЭНДОСКОПИЧЕСКИХ ВИДЕОИЗОБРАЖЕНИЙ С ПРИМЕНЕНИЕМ МЕТОДОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Специальность:

2.2.12. Приборы, системы и изделия медицинского назначения (технические науки)

ДИССЕРТАЦИЯ на соискание ученой степени доктора технических наук

> Научный консультант: доктор технических наук, профессор Приоров Андрей Леонидович

Ярославль – 2025

оглавление

ВВЕДЕНИЕ	7
ГЛАВА 1. АНАЛИЗ СОСТОЯНИЯ ВОПРОСА И ФОРМУЛИРОВКА	
ЗАДАЧ ИССЛЕДОВАНИЯ	22
1.1 Скрининговые эндоскопические исследования ЖКТ	22
1.2 Врачебные ошибки	26
1.2.1 Возможные врачебные ошибки при эндоскопии ЖКТ	28
1.3 Система поддержки принятия врачебных решений	31
1.3.1 Структура систем поддержки принятия врачебных решений	34
1.3.2 Классификация механизмов рассуждения в системах поддержки	
принятия врачебных решений	37
1.3.3 Требования к системам поддержки принятия врачебных решений	41
1.4 Современные эндоскопическая техника	42
1.5 Системы поддержки принятия врачебных решений в эндоскопической	
диагностике	50
1.6 Нейронные сети как инструмент глубокого машинного обучения	53
1.7 Использование нейронных сетей для анализа медицинских	
изображений	58
1.7.1 Сверточные нейронные сети	59
1.7.2 Другие современные архитектуры для систем глубокого обучения	63
1.8 Современные методы и алгоритмы обработки и анализа	
эндоскопических видеоизображений	64
1.8.1 Классификация задач обработки эндоскопических видеоизображений	65
1.8.2 Оценка качества эндоскопических видеоизображений	67
1.8.3 Детектирование объектов интереса нейронными сетями	70
1.8.4 Трекинг объектов интереса	73
1.8.5 Сегментация объектов интереса	79
1.9 Метрики оценки качества работы нейросетевых алгоритмов	81
1.10 Выводы и постановка задач исследования	83

3	
ГЛАВА 2. КОНЦЕПЦИЯ ПОСТРОЕНИЯ СППВР В ЭНДОСКОПИИ НА	
ОСНОВЕ АЛГОРИТМОВ АНАЛИЗА ВИДЕООИЗОБРАЖЕНИЙ И	
МЕТОДОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА	86
2.1. Особенности построения систем искусственного интеллекта для	
анализа медицинских изображений	86
2.2. Разработанная концепция построения СППВР в эндоскопии	88
2.3. Разработанная методология подготовки и разметки эндоскопических	
видеоданных	92
2.3.1. Описание предложенных процессов физической разметки	
и физической аугментации	92
2.3.2. Разработанная методика полуавтоматической разметки изображений	95
2.3.3. Используемое программное обеспечение для разметки изображений	99
2.3.4. Методология для контроля качества разметки эндоскопических	
изображений	103
2.4. Построение ИИ-инфраструктуры для задач дообучения	
эндоскопических СППВР	107
2.5. Краткие выводы	112
ГЛАВА 3. РАЗРАБОТКА И ИССЛЕДОВАНИЕ МЕТОДА КОНТРОЛЯ	
КАЧЕСТВА КОЛОНОСКОПИЧЕСКОГО ИССЛЕДОВАНИЯ	114
3.1. Актуальность задачи обнаружения купола слепой кишки	114
3.2. Описание собранной базы видеоизображений	117
3.3 Исследование базового нейросетевого алгоритма детектирования	
купола на базе моделей SSD и YOLO	120
3.4 Разработка улучшенного алгоритма детектирования купола слепой	
кишки с использованием анализа видеопоследовательности	127
3.4.1 Алгоритм детектирования купола с использованием трекинга	127
3.4.2 Алгоритм детектирования купола с использованием анализа	
оптического потока	130

3.4.3 Алгоритм расчета оптического потока на основе метода Хорна-	
Шунка	132
3.4.4 Нейросетевой алгоритм дополнительной классификации данных	136
3.5. Оценка качества колоноскопических изображений	141
3.5.1. Описание набора колоноскопических изображений	141
3.5.2 Анализ использования известных неэталонных алгоритмов оценки	
качества изображений	142
3.5.3. Разработка алгоритмов на основе признаков NSS	145
3.5.4. Разработка алгоритма на основе сверточной нейронной сети	145
3.5.5. Сравнительный анализ работы алгоритмов оценки качества	147
3.6. Краткие выводы	148
ГЛАВА 4. РАЗРАБОТКА И ИССЛЕДОВАНИЕ АЛГОРИТМОВ	
ДЕТЕКТИРОВАНИЯ И СЕГМЕНТАЦИИ КОЛОРЕКТАЛЬНЫХ	
ПОЛИПОВ	150
4.1. Актуальность задачи обнаружения полипов	150
4.2. Описание используемых наборов изображений полипов	154
4.2.1 Стандартные базы для задач детектирования и сегментации	155
4.2.2. Собственная база изображений полипов OnkoYar-SEG для задач	
детектирования сегментации	156
4.3. Разработка и исследование алгоритмов детектирования полипов	158
4.4. Разработка и исследование алгоритмов сегментации полипов	164
4.4.1. Исследование классических методов сегментации	164
4.4.2. Исследование нейросетевых алгоритмов сегментации	167
4.5. Краткие выводы	171
ГЛАВА 5. РАЗРАБОТКА И ИССЛЕДОВАНИЕ АЛГОРИТМОВ	
ДЕТЕКТИРОВАНИЯ И КЛАССИФИКАЦИИ ПАТОЛОГИЙ ЖЕЛУДКА	172
5.1. Актуальность задачи	172
5.2. Описание созданной базы изображений желудка с патологиями	177
5.2.1. Классы патологий желудка	178

-	
5.2.2 Фильтрация изображений перед этапом разметки	183
5.2.3 Разметка изображений желудка	185
5.2.4. Итоговая статистика по базе изображений желудка	186
5.3 Результаты моделирования нейронных сетей на основе архитектур	
SSD и RetinaNet со стандартными параметрами	188
5.3.1 Исследования для алгоритмов на базе нейронной сети SSD со	
стандартными параметрами	190
5.3.2. Результаты для алгоритма на базе нейронной сети RetinaNet	
со стандартными параметрами	192
5.4. Оптимизация параметров нейросетевых алгоритмов	194
5.4.1. Исследования по выбору политики обучения	195
5.4.2. Исследования по выбору типа аугментации	204
5.4.3 Исследования по устранению эффекта чересстрочной развертки	209
5.4.4 Выбор функции потерь	212
5.5 Сравнение результатов работы нейросетевых алгоритмов на основе	
SSD и RetinaNet со стандартными и оптимизированными параметрами	213
5.5.1 Результаты для алгоритмов на базе нейронной сети SSD с	
оптимизированными параметрами	214
5.5.2 Результаты для алгоритма на базе нейронной сети RetinaNet с	
оптимизированными параметрами	217
5.6 Результаты моделирования дополнительных архитектур нейронных	
сетей	219
5.6.1 Алгоритмы на основе нейронной сети SSD	219
5.6.2 Алгоритм на основе нейронной сети EfficientDet	231
5.6.3 Сравнительный анализ результатов для дополнительных архитектур	
нейронных сетей	233
5.7 Исследование алгоритмов детектирования на базе нейронной сети	
YOLO с базовыми параметрами	236
5.8 Оптимизация параметров для нейронной сети YOLO	239
5.8.1 Исследования по выбору политики обучения	239

241
250
253
256
257
259
262
263
263 266
263 266 273
263 266 273 275
263 266 273 275 279
263 266 273 275 279 284

введение

Актуальность темы исследования

В 2023 г. в Российской Федерации (РФ) выявлено 674 587 случаев злокачественных новообразований (ЗНО). Ведущие локализации в общей структуре онкологической заболеваемости включают: кожа (кроме меланомы) – 12,7%, молочная железа – 12,3%, трахея, бронхи и легкие – 9,2%, предстательная железа – 7,7%, ободочная кишка – 7,1%, желудок – 5,3%, прямая кишка, ректосигмоидное соединение и анус – 5,0%, лимфатическая и кроветворная ткани – 4,6%, тело матки – 4,5%, почки – 3,9%, поджелудочная железа – 3,1%, мочевой пузырь – 2,6%, шейка матки – 2,6%, яичники – 2,3%. Таким образом, онкологическая патология желудочно-кишечного тракта (ЖКТ) составляет 17,4% и является абсолютным лидером в структуре заболеваемости ЗНО. У мужчин ЗНО ЖКТ составляет 21,7%, у женщин – 15,6% [1, 2].

Одним из важнейших факторов в борьбе с ЗНО является раннее выявление опухолей. Ранняя диагностика онкологических заболеваний приводит к уменьшению летальности, существенной экономии на лечении, сохранению трудовых ресурсов, снижению процента людей с инвалидностью, уменьшению психологической нагрузки на семьи заболевших, повышению доверия населения к системе здравоохранения, т.е. оказывает существенный социально-экономический эффект на государственном уровне [3-9].

К сожалению, по статистике большинство 3HO ЖКТ в России диагностируются на поздних (III-IV) стадиях заболевания. Так, рак желудка выявляется на III-IV стадии в 57,5% случаев, а рак ободочной кишки – в 48,7%. Ранняя диагностика на I и II стадиях остается недостаточной [3, 4]. Эти данные необходимость профилактических подчеркивают усиления программ (скрининговых) эндоскопических исследований. Во многих развитых странах (США, Япония, Южная Корея, Германия, Нидерланды и др.) принимаются национальные программы скрининга колоректального рака и рака желудка,

7

которые рекомендуются людям, начиная с возраста 50 лет, с постепенным снижением этой возрастной планки до 45 лет [5-9].

В Приказе Министерства здравоохранения Российской Федерации от 27.04.21 № 404н «Об утверждении Порядка проведения профилактического медицинского осмотра и диспансеризации определенных групп взрослого населения» термин «скрининг» употребляется как синоним первого этапа диспансеризации [10-11]. В РФ, несмотря на высокую заболеваемость и смертность от онкопатологии ЖКТ, нет федеральных скрининговых программ. В отдельных регионах РФ проводились пилотные проекты, однако они не приобрели массовости в том числе из-за неподготовленности эндоскопических служб и населения [12-20].

Эндоскопический метод диагностики широко применяется с целью поиска предраковых изменений и раннего рака органов пищеварительного тракта в программах диспансеризации, скрининга отдельных групп населения, а также при диспансерном наблюдении пациентов с хроническими заболеваниями ЖКТ. Следует отметить, что эндоскопические исследования ЖКТ относятся к исследованиям экспертного уровня и отличаются высокой трудоемкостью, что связано с высокой вариативностью изображений слизистых оболочек ЖКТ и многочисленными артефактами на изображениях, получаемых непосредственно с эндоскопа. Из-за сложности интерпретации эндоскопических видеоизображений даже опытные специалисты пропускают до 20-25% ранних патологий [21-24]. Кроме того, недостаток эндоскопистов высокой квалификации делает массовый скрининг населения практически невозможным [3, 12, 17].

Выходом может стать внедрение систем искусственного интеллекта (ИИ), способных анализировать эндоскопические видеоданные в режиме реального времени (PPB). Такие системы решают две задачи: контроль качества исследования (например, оценка полноты осмотра слизистой) и поддержка принятия врачебных решений в PPB. Алгоритмы ИИ, обученные на больших медицинских данных, автоматически обнаруживают, сегментируют и классифицируют патологии, минимизируя влияние человеческого фактора. Использование ИИ-систем по некоторым данным увеличивает достоверность диагностики до 20-30% и позволяет расширить число скрининговых исследований даже при ограниченном числе специалистов экспертного уровня [25-30].

Создание и исследование современных систем ИИ, способных анализировать эндоскопические видеоданные, является важной теоретической и прикладной задачей в сфере разработки систем поддержки принятия врачебного решения (СППВР). Помимо поиска оптимальной архитектуры нейронной сети (НС) и процедуры ее обучения, имеются сложности, связанные с обработкой исходных видеоданных: их разметкой, аннотированием, настройкой методов размножения (аугментации) для расширения наборов данных. Также требуют решения вопросы разработки алгоритмов для обнаружения, сегментации и классификации патологий, а также пред- и постобработки результатов с учетом особенностей эндоскопических видеоизображений – таких как низкая контрастность, наличие артефактов и динамические изменения в кадре [31-35].

Применение систем СППВР в эндоскопии ЖКТ потенциально позволяет повысить точность диагностики, уменьшить влияние человеческого фактора на качество исследований, снизить стоимость и временные затраты на их проведение как в случае клинических, так и в случае скрининговых исследований. Кроме того, подобные системы могут успешно использоваться для обучения и повышения квалификации врачей-эндоскопистов [3, 23, 28].

Значительный вклад в развитие области цифровой обработки и анализа изображений И видеопоследовательностей внесли отечественные ученые М.И. Кривошеев, Ю.Б. Зубарев, В.П. Дворкович, А.В. Дворкович, Ю.В. Визильтер, А.С. Крылов, А.С. Конушин, Д.С. Ватолин, М.К. Чобану, Б.А. Алпатов, Ю.С. Бехтин, П.В. Бабаян, Д.П. Николаев, М.Н. Фаворская, Ю.С. Радченко, Е.В. Медведева, С.С. Садыков, А.Л. Жизняков, А.А. Орлов, S. Mitra, R. Gonzalez, R. Woods, Z. Wang, A. Bovik, H. Sheikh, E. Simoncelli, J. Astola, K. Egiazarian, R. Szeliski, R. Lukac и др. [36-45].

Важнейшие результаты в области построения нейросетевых моделей и методов глубокого машинного обучения (МО) для анализа одномерных и многомерных данных получили следующие ученые: А.И. Галушкин, А.Н. Горбань,

Ю.И. Журавлев, В.А. Сойфер, И.В. Оселедец, Е.В. Бурнаев, С.С. Марков, С.И. Николенко, А.А. Сирота, G. Hinton, J. Hopfield, Y. LeCun, Y. Bengio, I. Goodfellow [46-57].

В области применения методов и алгоритмов автоматического анализа и обработки эндоскопических видеоизображений следует отметить работы Н.А. Обуховой, А.А. Мотыко, А.А. Поздеева, С.А. Кашина, Е.Л. Никонова, Д.В. Завьялова, Р.О. Куваева, О.А. Степановой, А.А. Лебедева, R. Haidry, M. Hafner, D. Jha, Y. Shin, S. Ali, N.K. Tomar, H. Johansen, D. Johansen, J. Rittscher и др. [3, 18, 19, 24, 58-62].

Таким образом, анализ научно-технических источников показывает, что актуальной научной проблемой, имеющей важное социально-экономическое значение в области разработки приборов, систем и изделий медицинского усовершенствование назначения является аппаратно-программных видеоэндоскопических комплексов, которое приводит к повышению заболеваний достоверности диагностики онкологических И других при скрининговых И клинических исследованиях ЖКТ. Под повышением достоверности здесь понимаются следующие ключевые аспекты: снижение доли пропущенных патологий. снижение влияния человеческого фактора. стандартизация заключений и контроль качества исследований. Решением указанной проблемы может стать использование методов ИИ для построения алгоритмов и аппаратно-программных комплексов, составляющих основу для СППВР. Это позволит существенно повысить качество диагностики различных патологий желудка и кишечника и способствовать выявлению онкологических и других заболеваний на ранних стадиях их развития.

Целью работы является повышение достоверности клинических и скрининговых эндоскопических исследований желудочно-кишечного тракта за счет применения методов глубокого машинного обучения для обработки и анализа видеоизображений. Для достижения поставленной цели необходимо решить следующие задачи:

- Провести критический анализ современного состояния в области систем автоматического анализа видеоизображений эндоскопических исследований ЖКТ с помощью методов и алгоритмов искусственного интеллекта.
- Разработать концепцию построения СППВР для клинических и скрининговых исследований в эндоскопии на основе методов глубокого машинного обучения в условиях практической ограниченности входного набора аннотированных видеоизображений.
- 3. Усовершенствовать архитектуру системы искусственного интеллекта, работающей в качестве программного ядра для СППВР в эндоскопии ЖКТ.
- Создать физически и алгоритмически расширенные базы эндоскопических изображений и видеоданных с экспертной разметкой, аннотированием и врачебной валидацией областей патологий.
- Разработать алгоритмические решения для всех этапов обработки изображений – от предобработки и оценки качества до постобработки данных, учитывающие особенности эндоскопических видеопоследовательностей.
- Разработать метод контроля качества скринингового колоноскопического исследования и алгоритмы анализа видеопоследовательностей его реализующего.
- 7. Разработать алгоритмы детектирования и сегментации колоректальных полипов на видеоизображениях колоноскопических исследований.
- Разработать алгоритмы детектирования и классификации патологий желудка на видеоизображениях гастроскопических исследований.
- Создать, провести тестирование, опытную эксплуатацию и внедрение специализированного программного обеспечения для контроля качества, визуализации и сопоставления результатов эндоскопической диагностики.
- 10. Создать, провести тестирование, опытную эксплуатацию и внедрение аппаратно-программных комплексов для захвата, разметки, обработки и хранения видеоизображений эндоскопических исследований.

Методы исследования. Для решения поставленных задач использовались методы цифровой обработки сигналов, изображений и видеопоследовательностей, компьютерного зрения, математического моделирования, аппарат теории нейронных сетей и глубокого машинного обучения.

Объектом исследования являются цифровые эндоскопические видеосистемы для проведения гастроскопических и колоноскопических исследований.

Предметом исследования являются концепция построения СППВР на основе системы искусственного интеллекта клинических скрининговых для И эндоскопических исследований, методика построения аннотированных наборов видеоизображений, метод контроля качества скринингового колоноскопического видеоизображений обработки исследования, алгоритмы И анализа исследований, гастроскопических аппаратно-программные комплексы для обработки эндоскопических видеоданных в режиме реального времени.

Научная новизна полученных результатов

- Разработана концепция построения СППВР в эндоскопии ЖКТ на основе методов и алгоритмов глубокого машинного обучения, отличающаяся оригинальными методами цифровой обработки изображений и видеопоследовательностей, которая позволяет расширить функциональные возможности ее применения в условиях практической ограниченности входного набора аннотированных видеоданных.
- 2. Усовершенствована архитектура системы искусственного интеллекта, работающей в качестве программного ядра для СППВР, что позволило создать дополнительные модули и алгоритмы, результаты работы которых по точности обнаружения и классификации патологий на видеоданных, сопоставимы со средним результатом, показываемым врачом во время эндоскопической процедуры.
- Предложена методология расширения баз эндоскопических изображений и видеоданных для обучения, валидации и тестирования систем искусственного интеллекта, позволяющая снизить нормированные трудозатраты врачейэндоскопистов.

- 4. Усовершенствованы с учетом специфики видеоэндоскопических данных алгоритмы обработки и анализа видеоизображений, учитывающие априорную информацию об объекте исследования и явление межкадровой корреляции видеоданных.
- 5. Разработан метод для контроля качества колоноскопического исследования за счет реализации оригинальных алгоритмов детектирования купола слепой кишки, отличающийся использованием алгоритмов сопровождения областей интереса и анализа оптического потока.
- Созданы робастные алгоритмы детектирования и сегментации полипов на колоноскопических видеоданных, работающие в условиях ограниченности входного набора аннотированных видеоизображений.
- 7. Разработаны робастные алгоритмы детектирования и классификации аномальных областей (рак, ранний рак, иные патологии) на гастроскопических видеоизображениях желудка, позволяющие повышать достоверность диагностики в стандартном режиме и в режиме с оптическим увеличением.

Практическая значимость полученных результатов

- 1. Разработаны и апробированы в лечебных учреждениях системы на базе обеспечения зарегистрированного программного «EndoscopyVA.RT программа для анализа эндоскопических видеоизображений в реальном времени на основе методов искусственного интеллекта» И «EndoscopyDSS.Hub визуализации _ программа для результатов эндоскопических исследований с учетом анализа видеоданных методами искусственного интеллекта» для СППВР в эндоскопических исследованиях желудка и кишечника.
- Созданы и зарегистрированы размеченные базы изображений кишечника и желудка – «IntestinesImages» и «StomachImages», которые могут служить основой для дальнейших прикладных научных исследований и НИОКР в области автоматического анализа эндоскопических видеоизображений ЖКТ.
- 3. Для поддержки принятия врачебных решений в гастроскопии желудка решены следующие задачи: поиск областей рака/раннего рака, детектирование других

патологических областей, контроль качества проведения исследования. Соответствующий аппаратно-программный комплекс внедрен в деятельность эндоскопического отделения Ярославской областной клинической онкологической больницы.

- 4. Разработана методика создания баз эндоскопических видеоизображений с аннотацией патологических областей, позволяющая снизить временные врачей-эндоскопистов затраты на процедуру разметки на 35-45%. Методика протестирована на выборке из более 100 исследований и внедрена В национальном медико-хирургическом центре имени Н.И. Пирогова.
- 5. Созданы производительные алгоритмы детектирования и сегментации колоректальных полипов на эндоскопических изображениях и видеоданных, работающие с точностью свыше 80%. Предложенное аппаратно-программное решение внедрено в национальном медико-хирургическом центре имени Н.И. Пирогова и используется для частичной автоматизации диагностики, контроля качества исследований, а также обучения и повышения квалификации врачей-эндоскопистов.
- 6. Разработан и внедрен в разработки ООО «А-Вижн» алгоритм детектирования купола слепой кишки на видеоданных колоноскопических исследований, работающий с точностью порядка 90%. Использование данного алгоритма в Ярославской областной клинической онкологической больнице позволяет осуществлять контроль качества проведенного колоноскопического исследования.
- Осуществлена интеграция разработанных нейросетевых алгоритмов обработки и анализа эндоскопических видеоизображений с внешней специализированной эндоскопической медико-информационной системой, разрабатываемой ООО «ЭМИС».
- 8. Предложены эффективные по точности итогового решения реализации алгоритмов искусственного интеллекта на базе открытой архитектуры нейронной сети, которые по мере расширения аннотированных баз

видеоизображений могут использоваться в аппаратно-программных эндоскопических комплексах, в том числе в проектах по импортозамещению соответствующего оборудования.

Достоверность результатов подтверждается тем, что при разработке СППВР общепризнанном корректно использованы положения. основанные на математическом аппарате. Вводимые допущения и ограничения моделей и соответствуют известным В эндоскопии ЖКТ алгоритмов положениям. Обоснованность подходов построению алгоритмов обработки К видеоизображений, выбору параметров нейросетевых моделей и способов СППВР построения подтверждается положительными результатами экспериментальных исследований, выполненных на современном сертифицированном эндоскопическом оборудовании, а также их согласованием с аналогичными результатами, известными из российских и зарубежных научнотехнических источников.

На защиту выносятся:

- Предложенная концепция построения системы поддержки принятия врачебных решений в эндоскопии ЖКТ на основе анализа цифровых видеоизображений с использованием методов глубокого машинного обучения, повышающая достоверность скрининговых и клинических исследований в задаче ранней диагностики онкологических и других заболеваний.
- 2. Усовершенствованная архитектура системы искусственного интеллекта, работающей в качестве программного ядра для системы поддержки принятия врачебных решений, позволяющая получать практически значимые результаты в условиях ограниченности входного набора аннотированных эндоскопических видеоизображений.
- Разработанная методика создания расширенных баз эндоскопических изображений и видеоданных с экспертной разметкой, аннотированием и врачебной валидацией областей обнаружения патологий ЖКТ, позволяющая

снизить трудозатраты медицинского персонала на этапах обучения и дообучения систем искусственного интеллекта.

- 4. Метод контроля качества скринингового колоноскопического исследования с применением комбинированного алгоритма обнаружения области купола слепой кишки на колоноскопических видеоданных, учитывающего информацию с нейросетевого детектора, алгоритма сопровождения области интереса и алгоритма анализа оптического потока.
- 5. Предложенные алгоритмы детектирования и сегментации колоректальных полипов на эндоскопических видеоданных толстого кишечника, построенные на открытой нейросетевой архитектуре и учитывающие явление межкадровой корреляции видеоизображений.
- 6. Разработанные робастные алгоритмы обнаружения и классификации аномальных областей (рак, ранний рак, иные патологии) на эндоскопических видеоизображениях желудка, построенные на открытой нейросетевой архитектуре и предоставляющие поддержку в принятии врачебных решений как в режиме с оптическим увеличением, так и в стандартном режиме.
- 7. Система поддержки принятия врачебных решений в скрининговых колоноскопических исследованиях, позволяющая проводить анализ видеоизображений соответствующего исследования как в режиме реального времени, так и в режиме анализа сохраненной видеопоследовательности.
- Система поддержки принятия врачебных решений в клинических гастроскопических исследованиях желудка, позволяющая обнаруживать патологические области и проводить их классификацию.

Реализация и внедрение результатов исследования

Результаты работы внедрены в ФГБУ «НМХЦ им. Н.И. Пирогова» Минздрава России (Пироговский центр, г. Москва), в ГБУЗ ЯО «Клиническая онкологическая больница» (г. Ярославль), в Медицинский центр диагностики и профилактики (г. Ярославль), а также в перспективные разработки компаний ООО «А-Вижн» (г. Ярославль), ООО «ЭМИС» (г. Ярославль). Отдельные результаты диссертационной работы внедрены в научноисследовательские работы Ярославского государственного университета им. П.Г. Демидова при выполнении гранта РФФИ «Исследование применимости технологий глубокого обучения и сверточных нейронных сетей для распознавания патологий желудка по видеоизображениям эндоскопического исследования» (№19-37-90153, 2019-2021 гг.) и хоздоговорных работ с ООО «СМАРТ АНАЛИЗ» по теме «Обучение и тестирование моделей нейронных сетей для сегментации, распознавания и классификации эндоскопических исследований» (№ 2338-х/д, 2020-2021 гг.).

Апробация работы. Результаты работы обсуждались на 49 научных конференциях, среди которых: «Цифровая обработка сигналов и ее применение» (DSPA), Москва, ИПУ РАН, 2005-2025; «Телевидение: передача и обработка изображений», Санкт-Петербург, ЛЭТИ, 2016-2018; Методы, средства И технологии получения и обработки измерительной информации («Шляндинские чтения»), Пенза, 2023; «Новые информационные технологии и системы», Пенза, 2020; «Проблемы информатики в образовании, управлении, экономике и технике», Пенза, 2022-2024; Биотехнические, медицинские и экологические системы, измерительные устройства и робототехнические комплексы («Биомедсистемы»), Рязань, 2024; Научная сессия, посвященной Дню радио, Москва, 2019-2025; Международная конференция по компьютерной графике и машинному зрению «ГрафиКон», Москва, МГУ, 2010-2023; «Нейроинформатика», Москва, МИФИ, 2023; «Нейрокомпьютеры и их применение», Москва, 2018; Научно-техническая конференция Санкт-Петербургского НТО РЭС им. А.С. Попова, посвященная Дню радио, Санкт-Петербург, 2022; «Физика и радиоэлектроника в медицине и экологии» (ФРЭМЭ), Владимир, 2006; «Высокие технологии в эндоскопии пищеварительной системы», Ярославль, 2018-2022; Image Processing, Computer Vision and Pattern Recognition (IPCV), Jac-Berac, 2010; Image Processing (ICIP), Генуя, 2005; EUROCON, Белград / Санкт-Петербург, 2005, 2009; International Conference of Signal and Image Engineering, Лондон, 2011; International MultiConference of Engineers and Computer Scientists, Гонконг, 2012; Conference of Open Innovations Association (FRUCT), Хельсинки, 2017; International Conference on Machine Vision (ICMV), Амстердам, 2019; IEEE East-West Design & Test Symposium, Варна / Батуми, 2020-2021; Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO), Псков, 2023.

Публикации. По теме диссертации опубликовано 107 научных работ, из которых: 27 публикаций в журналах из перечня ВАК при Минобрнауки России, 3 публикации в журналах из перечня ВАК при Минобрнауки России по другим научным специальностям, 25 публикаций с индексацией в международной системе Scopus, 45 публикаций в сборниках трудов конференций (РИНЦ), 2 патента на изобретение РФ, 3 свидетельства о государственной регистрации программы для ЭВМ, 2 свидетельства о государственной регистрации баз данных.

Личный вклад. Теоретические и практические результаты, относящиеся к концепции построения системы поддержки принятия врачебных решений в гастроскопии и колоноскопии, к методам и средствам обнаружения и визуализации областей интереса при эндоскопических исследованиях, получены непосредственно автором, под его руководством или при его решающем участии. В работах, написанных совместно, автору принадлежат методы, алгоритмы и программные средства: обнаружения и визуализации областей интереса при эндоскопических исследованиях; определения оптимальных параметров нейросетевых моделей; выбора того или иного метода предварительной фильтрации, оценки качества и постобработки видеоизображения; построения методики полуавтоматической разметки эндоскопических изображений объекта интереса; структуры построения программно-аппаратного комплекса для обработки эндоскопического потока в режиме реального времени.

Единолично опубликована 21 научная работа, в том числе 4 статьи в изданиях из перечня ВАК. Автор обладает опытом руководства НИР по тематике диссертации в рамках выполнения грантов РФФИ, являлся научным руководителем и исполнителем по программам Фонда содействия инновациям по тематике создания СППВР в эндоскопии. С 2017 года Хрящев В.В. является научным руководителем центра искусственного интеллекта и цифровой экономики ЯрГУ им. П.Г. Демидова. Центр оснащен высокопроизводительными суперкомпьютерами NVIDIA DGX-1 и A-VSN-1, что позволило провести глубокое обучение современных моделей нейронных сетей в рамках темы диссертационной работы. В 2017-2023 гг. автор работы являлся руководителем стратегических проектов по программе развития ЯрГУ в рамках министерских программ «Опорный ВУЗ» и «Приоритет-2030». Хрящев В.В. в 2023 году награжден почетной грамотой губернатора Ярославской области многолетний за добросовестный труд, высокий профессионализм и значительные научные достижения, способствующие социально-экономическому развитию Ярославской области.

О высокой квалификации соискателя также свидетельствует то, что он в течении последних десяти лет является членом программного комитета и рецензентом профильных международных конференций Conference of Open Innovations Association FRUCT (https://fruct.org/), IEEE East-West Design & Test Symposium (https://conf.ewdtest.com/), SPIE International Conference on Machine Vision (https://icmv.org/), International Conference on Machine Vision and Applications (http://icmva.org/), International Conference on Signal Processing and Machine Learning (http://spml.net/), на которых неоднократно выступал в качестве руководителя секций по направлениям «искусственный интеллект» и «обработка и анализ изображений».

Структура и объем работы. Диссертация состоит из введения, шести глав, заключения, списка сокращений, списка литературы, содержащего 315 наименований, и одного приложения. Она изложена на 326 страницах машинописного текста, содержит 98 рисунков и 75 таблиц.

Проблемы и задачи, решенные в диссертации, соответствуют паспорту специальности 2.2.12. Приборы, системы и изделия медицинского назначения (технические науки) в части следующих направлений исследований:

 - 2. Приборы, системы и аппаратно-программные комплексы для оценки текущего состояния, скринингового обследования, мониторинга, прогнозирования и диагностики состояния здоровья человека.

- 10. Технические средства и системы, обеспечивающие повышение точности медицинской диагностики, воспроизводимости и сопоставимости результатов биомедицинских исследований.
- 14. Методы, модели и алгоритмы, включая распознавание образов, для медицинских информационных и интеллектуальных систем, обеспечивающих повышение эффективности медико-биологических исследований и врачебных решений.
- 19. Методы и средства регистрации, анализа и интерпретации медицинских изображений.
- 20. Системы поддержки принятия врачебных решений и медикотехнологических процессов, экспертные, информационные и управляющие системы медицинского назначения, обеспечивающие повышение качества медицинского обслуживания населения.

Благодарности

Выражаю благодарность научному консультанту, доктору технических наук, профессору Приорову Андрею Леонидовичу за помощь на всех этапах написания данной работы.

Выбор научного направления произошел благодаря доверию и неизменной поддержке ректора ЯрГУ им. П.Г. Демидова в 2005-2022 гг. Русакова Александра Ильича и руководителя управления научных исследований ЯрГУ – Мазалецкой Анны Леонидовны.

Высокий уровень теоретических и экспериментальных исследований достигнут благодаря работе в области разработки и внедрения систем искусственного интеллекта в медицине на кафедре цифровых технологий и машинного обучения ЯрГУ, в центре искусственного интеллекта и цифровой экономики ЯрГУ, а также в малых инновационных предприятиях совместно с Брюхановым Юрием Александровичем, Чистяковым Михаилом Валерьевичем, Ганиным Александром Николаевичем, Апальковым Ильей Владимировичем, Саутовым Евгением Юрьевичем, Назаровским Александром Евгеньевичем, Карелиным Павлом Валентиновичем, Голубевым Максимом Николаевичем, Волковым Денисом Борисовичем, Матвеевым Дмитрием Вячеславовичем, Лебедевым Антоном Александровичем, Островской Анной Александровной, Степановой Ольгой Анатольевной, Ивановским Леонидом Игоревичем, Стефаниди Антоном Федоровичем, Седовым Анатолием Геннадьевичем, Плиссом Романом Евгеньевичем, Капрановым Виктором Алексеевичем, Киселевым Виктором Ивановичем, Лукашевичем Юрием Александровичем, Волоховым Владимиром Андреевичем, Топниковым Артемом Игоревичем, Средняковой Анастасией Сергеевной, Андержановой Анастасией Сергеевной, Ларионовым Романом Васильевичем, Котовым Никитой Валерьевичем, Тихомировым Александром Андреевичем.

Отдельная благодарность за сотрудничество и помощь в постановке экспериментов разработанными аппаратно-программными совместных с отделения эндоскопии Ярославской областной комплексами сотрудникам клинической онкологической больницы: Кашину Сергею Владимировичу, Завьялову Дмитрию Вячеславовичу, Куваеву Роману Олеговичу, Тарасовой Екатерине Львовне, Гусейновой Сабине Руслановне. Также за помощь в экспериментальном тестировании разработанных решений благодарю сотрудников ФГБУ «НМХЦ им. Н.И. Пирогова» Минздрава России: Маады Аяса Сергеевича, Пензина Олега Владимировича, Субботина Сергея Александровича.

Глубокую благодарность хотелось бы выразить близким – жене и родителям – за возможность заниматься научной деятельностью и неизменную поддержку во время написания данной работы.

ГЛАВА 1 АНАЛИЗ СОСТОЯНИЯ ВОПРОСА И ФОРМУЛИРОВКА ЗАДАЧ ИССЛЕДОВАНИЯ

1.1 Скрининговые эндоскопические исследования ЖКТ

Скрининговые эндоскопические исследования желудочно-кишечного тракта играют ключевую роль в профилактике и ранней диагностике онкологических и других заболеваний. К основным методам диагностики относятся: гастроскопия (ФГДС) – осмотр пищевода, желудка и двенадцатиперстной кишки и колоноскопия исследование толстой кишки ДЛЯ диагностики колоректального рака. Современные способы седации и тонкие эндоскопы минимизируют дискомфорт пациента [63]. Важной альтернативой для труднодоступных участков тонкой кишки является капсульная эндоскопия [64]. Эндоскопическая диагностика позволяет обнаружить рак желудка, пищевода, двенадцатиперстной или толстой кишки на ранних стадиях, когда лечение наиболее эффективно. Также важным является обнаружение предраковых состояний – полипов, атрофического гастрита, язвы и болезни Барретта [31-32].

Целевыми группами скрининговых исследований являются люди с 45-50 лет (колоректальный рак) или с 40-45 лет (рак желудка в регионах с высокой заболеваемостью). Факторами риска являются наследственные синдромы, хронические заболевания – гастрит, язва, воспалительные болезни кишечника, образ жизни – курение, алкоголь, ожирение, диета с избытком красного мяса. Рекомендуемая регулярность скрининга для колоноскопии составляет каждые 10 лет (при отсутствии патологий), а для гастроскопии – каждые 3–5 лет (в группах риска). В странах с развитыми программами скрининга (например, Япония для рака желудка) смертность снизилась на 40-60%. Колоноскопия, выполняемая каждые 10 лет, снижает риск летальных исходов от колоректального рака на 50% [4, 5, 23, 29].

В РФ, несмотря на высокую заболеваемость и смертность от онкопатологии ЖКТ, нет федеральных скрининговых программ. В отдельных регионах РФ проводились пилотные, в том числе корпоративные проекты [12-16], однако они не приобрели массовости, в том числе из-за неподготовленности эндоскопических служб и населения [3, 17].

Внедрение методов ИИ в программы скрининговой эндоскопии ЖКТ способно существенно повысить эффективность борьбы с онкологическими заболеваниями в России [65-66]. Особенно значимым это может стать в свете существующих системных проблем здравоохранения. По данным Всемирной организации здравоохранения (ВОЗ), применение современных технических обеспечивающих повышение средств И систем, точности медицинской диагностики для скрининга, позволяет значительно понизить затраты на лечение поздних стадий рака. Для РФ это открывает перспективу оптимизации распределения ресурсов в сфере здравоохранения и, тем самым, повышения качества медицинского обслуживания населения. Использование приборов, систем и аппаратно-программных комплексов, основанных на методах и алгоритмах ИИ, потенциально способно решить следующие проблемы в области скрининговой и клинической эндоскопической диагностики ЖКТ [67-76]:

- Кадровый дефицит. В РФ постоянно наблюдается дефицит эндоскопистов, при этом не все они обладают экспертной квалификацией для выявления ранних форм онкологических заболеваний. Это делает массовый скрининг практически нереализуемым [64, 72, 75]. Алгоритмы ИИ, обученные на миллионах изображений патологий (полипы, язвы, ранние опухоли и др.), могут стать «виртуальными ассистентами» для врачей-эндоскопистов. Они формируют «второе мнение», повышая точность врачебных решений до 80–90%, и позволяют менее опытным специалистам проводить исследования на уровне экспертов, что критически важно для регионов РФ с дефицитом врачебных кадров.
- 2. Контроль качества процедур. Во время эндоскопической диагностики существует проблема неполного осмотра: до 25% слизистой может остаться

неисследованной из-за человеческого фактора (усталость, спешка). Алгоритмический мониторинг в реальном времени способен, например, анализировать покрытие зоны осмотра, контролировать скорость выведения эндоскопа, анализировать качество подготовки кишечника [23, 77-79].

3. Масштабирование скрининга. Даже в успешных региональных инициативах охват редко превышает 15–20% целевой группы из-за ограниченности ресурсов. Интеллектуальные СППВР позволят обрабатывать эндоскопические видеоданные удаленно, автоматизировать формирование протоколов, сокращая время на одну процедуру на 20–30%, проводить «предварительную фильтрацию» – пациенты с высоким риском по результатам ИИ-анализа направляются к экспертам, остальные – в наблюдение. Это позволит повысить эффективность медицинских исследований и врачебных решений.

Для успешного внедрения интеллектуальных систем поддержки врачебных решений в программы скрининговой эндоскопии необходимо обратить внимание на следующие аспекты [3, 17, 21, 23, 30]:

- Преодоление системных барьеров. Для работы систем ИИ необходима 1. цифровизация эндоскопических отделений полная _ переход на цифровым видеоэндоскопы с выводом данных, интеграция с соответствующими медико-информационными системами (МИС). Внедрение методов ИИ сопровождаться И средств должно соответствующими квалификации врачей по программами повышения эндоскопической диагностике с использованием СППВР.
- 2. Этические и правовые аспекты. Статус рекомендаций, даваемых СППВР, требует юридического закрепления (например, окончательно врач подтверждает предложенный системой ИИ диагноз). Необходимы просветительские кампании, чтобы избежать страха перед «роботамиврачами». Анонимизация видеоархивов И соблюдение 152-ФЗ «О персональных данных» – обязательное условие для обучения нейросетевых алгоритмов.

Создание эндоскопических центров с централизацией потока ЛИЦ повышенного риска развития опухолей ЖКТ является решением вопроса эффективной реализации скрининга ЗНО органов пищеварения. В настоящее время в регионах РФ для расширения доступа к онкоскринингу создаются подобные эндоскопические центры (например, высокопотоковые центры эндоскопического скрининга колоректального рака при ГБКУЗ ЯО «Центральная городская [3, 17, 20]. больница») Направление пациентов осуществляется врачами, оказывающими первичную медико-санитарную помощь, а также специалистами узкого профиля (гастроэнтерологами и онкологами) при наличии показаний к проведению онкологического поиска. После выполнения эндоскопического исследования при выявлении патологии пациент должен быть направлен к специалисту для определения тактики дальнейшего лечения и наблюдения (рисунок 1.1) [3].



Рисунок 1.1 – Маршрутизация пациентов в эндоскопических центрах

Под термином «оппортунистический скрининг» здесь понимается подход, при котором обследование на выявление заболеваний или факторов риска проводится у пациентов, обратившихся за медицинской помощью в лечебно-профилактическое учреждение (ЛПУ) по другим причинам. В отличие от плановых скрининговых программ, которые направлены на массовое обследование условно

здорового населения, оппортунистический скрининг использует возможность уже состоявшегося контакта пациента с системой здравоохранения для дополнительной диагностики [3]. Такой подход особенно актуален в условиях нехватки медицинских ресурсов и кадров, так как позволяет эффективно использовать существующие возможности ЛПУ. В РФ возможна интеграция такого подхода в этап диспансеризации, на котором пациенты получают направление на эндоскопические исследования при наличии показаний [17].

Таким образом, скрининговые эндоскопические исследования представляют собой мощный инструмент сохранения здоровья населения. Они не только спасают жизни за счет ранней диагностики, но и сокращают расходы на лечение поздних стадий онкологических болезней. Использование систем ИИ в данном способе диагностики не приводит к замене врачей-эндоскопистов, а предоставляет инструмент для преодоления дисбаланса между потребностью в скрининге и существующими ресурсами системы здравоохранения. Для РФ такие системы предоставляют шанс сократить смертность от рака ЖКТ (по аналогии с опытом Японии и Южной Кореи); создать синергетический эффект, объединив федеральные медицинские программы с коммерческими И научными ИИ-разработками; стать передовым регионом по внедрению интеллектуальных систем в медицину, привлекая инвестиции в соответствующие стартапы. Успех использования систем ИИ в эндоскопической диагностике зависит от системного подхода: обновления стандартов, финансирования пилотных проектов И формирования «культуры цифровых данных» в медицинском сообществе.

1.2 Врачебные ошибки

В области настоящее время В медицины широко используются диагностические процедуры, в ходе которых возникает необходимость в анализе визуальных технологий: данных, полученных с помощью различных ультразвуковой диагностики, рентгеновских исследований, магнитно-резонансной терапии, компьютерной томографии, эндоскопии и др. [3, 31, 32, 42]. Проведение такого анализа с ростом разрешающей способности датчиков, использованием

26

мультимодальных данных и внедрением новых технологий обработки информации в медицинскую технику становится все более сложной и трудоемкой задачей. В такой ситуации ключевую роль при соответствующей диагностике играет человеческий фактор. Уровень подготовки специалиста, его физическое и эмоциональное состояние, возможное утомление и многие другие не всегда хорошо предсказуемые аспекты могут оказывать значительное влияние на качество диагностики, в том числе и на вероятность возникновения врачебных ошибок [80-82].

Под термином «врачебная ошибка» можно объединить ряд терминов, наиболее часто использующихся для описания результатов неправильного оказания медицинских услуг: недосмотр, непреднамеренные действия работников, неблагоприятный медицинских диагноз, послеоперационные осложнения, нарушения при оказании медицинской помощи, которые могут нанести существенный вред пациенту [81]. Врачи и медицинские центры не всегда охотно сообщают о таких ошибках из-за страха юридических последствий или репутационных рисков. Это приводит к недостоверной статистике («культура замалчивания»). Однако отдельные страны и международные организации публикуют данные, которые позволяют оценить масштаб проблемы [82].

ВОЗ, каждый десятый пациент в мире страдает из-за По оценке предотвратимых врачебных ошибок. В США причине ПО неправильно поставленных диагнозов в год умирает от 210 тысяч до 440 тысяч пациентов лечебных учреждений. В отчете Европейского союза указано, что 8-12% пациентов лечебных учреждений испытывает последствия врачебных ошибок. В странах с низким уровнем дохода соответствующие риски выше из-за нехватки ресурсов и слабой медицинской инфраструктуры [79].

Официальной статистики по количеству врачебных ошибок в РФ нет, но по данным различных организаций, средняя смертность от последствий неправильно принятых решений медицинскими специалистами составляет порядка 50 тысяч человек ежегодно. Уровень ошибочной диагностики онкологических заболеваний составляет 30-40% от общего числа поставленных диагнозов. Эксперты в области

здравоохранения утверждают, что около 70% врачебных ошибок можно предотвратить при более тщательном подходе к оказанию медицинских услуг [81].

Медицинские ошибки могут возникать на разных этапах процесса: диагностика (12%), лечение (38%), администрирование (43%), выписка (5%), прочее (2%) [80]. Несмотря на то, что частота ошибок при проведении диагностических процедур не так велика, именно этот этап является особенно важным для формирования дальнейшего плана лечения пациента, и поэтому крайне важно минимизировать подобные случаи.

1.2.1 Возможные врачебные ошибки при эндоскопии ЖКТ

При проведении эндоскопии в полость ЖКТ вводится видеогастроскоп (оптический прибор для визуализации пищевода, желудка и двенадцатиперстной кишки) или видеоколоноскоп (оптический прибор для визуализации толстого кишечника). Это высокотехнологичный многофункциональный инструмент, позволяющий не только производить малоинвазивный осмотр и диагностику ЖКТ, но и осуществлять взятие биопсии для гистологического анализа из образцов тканей, подозрительных на наличие неоплазии, выполнять удаление неоплазий слизистой оболочки и подслизистого слоя, останавливать кровотечения и др. Видеоизображение, получаемое при проведении исследования, выводится на монитор, расположенный перед врачом-эндоскопистом. Специалисту требуется обнаружить патологии на слизистой оболочке пищевода, желудка или кишечника путем анализа визуальных данных [31]. Кроме того, данные с эндоскопической системы могут параллельно транслироваться удаленному эксперту, в том числе для врачебного консилиума. Большинство формирования современных эндоскопических систем поддерживает возможность формирования видеоданных с высоким разрешением, обеспечивающих возможность качественного осмотра органов ЖКТ [32].

Диагностическая эндоскопия – один из ключевых инструментов современной медицины, от результатов проведения которой зависят дальнейшие действия врачей других специализаций (хирургии, онкологии и т.д.), а, соответственно, и

28

здоровье пациента. Эндоскопия ЖКТ способна диагностировать предраковые состояния и рак, язвенные болезни, болезни Крона, целиакии и ряд других заболеваний [32]. Ранняя диагностика снижает затраты на лечение поздних стадий указанных заболеваний. Например, по данным ВОЗ стоимость лечения рака желудка на ранней стадии (эндоскопическая резекция) в 10 раз ниже, чем на поздней (хирургия + химиотерапия).

Общий вид эндоскопического оборудования от Санкт-Петербургской компании ООО «ПК Прометей» приведен на рисунке 1.2 [83]. Эндоскопическая диагностика сокращает расходы на лечение и постоянно совершенствуется благодаря как новым аппаратным, так и программным технологиям обработки «больших медицинских данных».

Врачебные ошибки врача-эндоскописта заключаются в неверной интерпретации патологических областей ЖКТ. Их можно разделить на следующие виды [79, 81]:

- пропуск патологии, когда специалист не смог обнаружить область интереса (полип, рак, язву и т.д.);
- неверное определение патологической области, когда, например, специалист определяет раковую опухоль как язву или наоборот и т.д.;
- неверное определение характеристик обнаруженной патологии, например, области расположения или размеров.

Особо опасными для постановки дальнейшего диагноза являются ошибки, связанные с пропуском / неверной интерпретацией такой патологии как онкологическое новообразование, которая требует незамедлительного лечения или хирургического вмешательства во избежание дальнейшего развития заболевания [32, 82]. В этом случае происходит формирование неверного предположения об отсутствии патологии (ошибки первого рода), и, как следствие, отсутствие своевременного лечения. В то же время ошибки, связанные с предположением о наличии патологии того или иного вида, которой на самом деле нет в организме пациента (ошибки второго рода), могут привести к неблагоприятным последствиям из-за проведения интенсивного лечения или процедур, вредящих здоровому организму [33-35], а также подрыву доверия к данному виду обследования.





Рисунок 1.2 – Эндоскопическая стойка (а); эндоскопический видеопроцессор VSE-100 (б); видеоколоноскоп C-SVI100 (в)

B)

Использование достижений в области информационных и интеллектуальных систем и технологий может способствовать повышению качества медицинского обслуживания населения. Современные методы и технологии ИИ и анализа больших данных позволяют разработать СППВР для различных областей медицины [54-56]. Рассмотрим их более подробно.

1.3 Системы поддержки принятия врачебных решений

Системы, которые не только предоставляют информацию, но и участвуют в действиях по принятию решений в организации какого-либо процесса, известны как системы поддержки принятия решений (СППР) [84-87]. Они представляют собой информационную систему, которая поддерживает деятельность организаций, учреждений или предприятий, связанную каким-либо образом с принятием решений. Они особенно востребованы, когда требуется предсказать будущие ситуации, условия или возможные состояния на основе анализа уже имеющихся данных [88].

Одной из наиболее важных сфер применения подобных информационных систем являются поддержка принятия врачебных (клинических) решений [89]. Можно считать, что любая компьютерная система, имеющая дело с врачебными данными или знаниями, предназначена для поддержки принятия врачебных решений. Такие системы являются ярким примером возможного применения методов ИИ и анализа больших данных в государственных и частных учреждениях здравоохранения. Они помогают в принятии клинических решений об отдельных пациентах, т.е. являются активными системами знаний, которые используют два или более элемента данных пациента для выработки рекомендаций по каждому конкретному случаю. Подобные системы также все чаще интегрируются в клиническую практику телемедицины [90].

Причина необходимости использования частичной автоматизации в ходе оказания медицинских услуг заключается в следующем. В общем случае, медицинский специалист способен самостоятельно справиться с конкретным клиническим случаем, если обладает высоким уровнем знаний по анализируемому

вопросу. Общие медицинские знания часто не используются врачом в момент оказания медицинской помощи, в неотложных ситуациях или в ходе проведения сложной диагностики. Интеллектуальные системы содержат данные о пациенте и применимую, хорошо структурированную и актуальную базу знаний; руководящие принципы, основанные на фактических данных, которые помогают врачу совершенствовать процесс принятия решений. Такие системы позволяют акцентировать внимание специалиста на важных для работы с пациентом характеристиках, которые в обычных условиях могут быть пропущены по причине недостатка знаний, невнимательности, усталости, нетипичности случая. Такая поддержка основных когнитивных процессов, связанных с медицинским мышлением, в некоторой степени облегчает работу врачу и предоставляет ему новые, лучше сформированные методы для наиболее эффективного обследования пациента [86]. Следует еще раз отметить, что в современной концепции такие системы не заменяют мышление конкретного врача, а только усиливают его когнитивные способности, позволяют иметь доступ к огромной базе обработанной и структурированной информации, акцентируют внимание на определенных областях, позволяют предотвратить ошибки, вызванные усталостью и другими факторами.

В настоящее время СППВР нашли применение в разных медицинских сферах [85, 86, 88, 89]:

- профилактические процедуры (использование помощи для подбора мер по повышению иммунитета пациента, направлению его на диагностические процедуры);
- дифференциальная диагностика;
- назначение лечения (подбор необходимых лекарственных средств и лечебных процедур в соответствии с данными о пациенте, помощь в наблюдении за пациентом в процессе лечения);
- наблюдение за пациентом по окончании лечения (выявление результативности оказанных медицинских услуг, уведомление об ухудшении положения пациента на учете).

Потенциальные преимущества использования информационных СППВР в клинической практике рассмотрены на рисунке 1.3 [88, 90].



Рисунок 1.3 – Преимущества использования СППВР в клинической практике

Их использование приводит к уменьшению медицинских ошибок при осмотре пациентов, диагностике, назначении лечения. Они предупреждают о потенциальных рисках: аллергических реакциях, взаимодействии лекарств, противопоказаниях; автоматически предлагают врачам рекомендации, основанные на актуальных клинических руководствах и доказательной медицине. Например, система может подсказать оптимальную дозировку препарата или необходимость дополнительной диагностики, что повышает качество лечения и минимизирует субъективность.

В своей работе СППВР должны получать информацию из электронных медицинских карт, которые объединяют историю болезней, результаты анализов, снимки, назначения и другие данные о пациенте в одном месте. Это ускоряет доступ к информации, снижает риск ошибок из-за неполных данных и позволяет врачам принимать решения на основе полной картины состояния пациента. Персонализация медицины заключается в учете генетических особенностей, образа жизни и анамнеза пациента, что позволяет подбирать индивидуальные схемы лечения.

информационные СППВР Таким образом, не только повышают эффективность И безопасность медицинской помощи, НО И становятся прогностической, инструментом для перехода к превентивной И персонализированной медицине. Однако их успешное внедрение требует внимания к защите данных, обучению медицинского персонала и интеграции с существующими рабочими процессами (РП).

1.3.1 Структура систем поддержки принятия врачебных решений

Многие характеристики СППВР связаны с клинической эффективностью, функциональностью, предотвращением ошибок, возможностью интеграции в рабочий процесс врача, экономической эффективностью и т.д., поэтому важно полностью понимать архитектуру и различные способы построения таких систем [87]. Наиболее эффективным способом использования СППВР является ее имеющуюся лечебном внедрение уже В заведении медицинскую В информационную систему (МИС). Данное решение позволяет облегчить медицинскому персоналу процесс адаптации к новому программному обеспечению (ПО).

Работа с СППВР происходит следующим образом: система осуществляет обработку данных, их классификацию и сегментацию, а затем предоставляет медицинскому специалисту возможные решения анализируемой проблемы и набор возможных результатов. Далее врач отбирает полезную информацию и отклоняет неверные предложения системы. Предоставляемая поддержка принятия решений основана на процессах сложной оценки результатов и алгоритмах, которые используют базы знаний [46-48, 87].

Между рассуждениями врача и СППВР существует сходство, хотя система может прийти к аналогичному выводу без использования той модели рассуждений, которую использовал врач. В клинической практике часто используется некоторая вероятностная информация при принятии решений. Например, врач может поставить диагноз гриппа чаще зимой, когда грипп более распространен (вероятен), чем летом. Тем не менее, врачи используют эту информацию неформальным образом; другими словами, они не используют точные расчеты и формулы для принятия диагностических решений.

Еще одной особенностью принятия клинических решений в реальной жизни является то, что врачам не требуется полная информация для принятия решения.

Большинство врачей способно принимать решения на основе неполной или противоречивой информации. Напротив, СППВР полагаются на четко определенные численные методы для обоснования своих аргументов, и им требуется достаточно информации для завершения анализа данных. Такие информационные системы (ИС) не очень подходят для ситуаций, в которых точные данные неизвестны.

Чтобы понять, как работают такие ИС, и при каких условиях они наиболее эффективно функционируют, важно оценить общую модель СППВР (рисунок 1.4). У такой системы есть вход и выход, механизм рассуждений (умозаключений) и база знаний [87]. Существуют архитектуры систем, которые не следуют этой модели.



Рисунок 1.4 – Стандартная структура СППВР, основанной на знаниях

Пользователь вводит соответствующие данные (то есть термины из контролируемого словаря для представления клинических данных), а система предоставляет выходные данные (например, дифференциальный диагноз, расположение патологии или предложения по терапии) [88, 89]. Ввод данных пациентов в компьютерные системы может осуществляться медицинскими работниками вручную или путем автоматической расшифровки речевой информации. Другой вариант – автоматическое извлечение информации из электронных медицинских карт (ЭМК) пациентов. К входным данным могут относиться:

- Электронные медицинские карты: история болезней, диагнозы, назначения, результаты обследований.
- Лабораторные и инструментальные данные: анализы, рентген, МРТ, КТ, УЗИ.

- Внешние базы знаний: клинические рекомендации (например, NCCN, WHO),
 фармакологические справочники, научные публикации.
- Данные с датчиков и с носимых устройств (пульс, давление, глюкоза).
- Генетическая информация (результаты секвенирования, биомаркеры).

Механизм рассуждений применяет детерминированные или вероятностные цепочки действий к входным данным и часто опирается на дополнительные факты, закодированные в базе знаний системы. База знаний для работы системы в общем случае должна включать, например, следующие типы информации:

- Клинические руководства: стандарты лечения (например, схемы терапии при гипертонии).
- Фармакологические данные: взаимодействие лекарств, дозировки, противопоказания.
- Медицинские классификаторы (МКБ-10, SNOMED CT).

Поддержание актуальности базы знаний является важнейшим процессом при эксплуатации таких систем, поскольку база знаний должна расширяться и обновляться по мере роста медицинских знаний [90].

Для работы с большими медицинскими данными используются: базы данных, как централизованные хранилища структурированных данных (SQL/NoSQL); «озера данных» (data lake) для неструктурированных данных (текстовые заметки, изображения); интерфейсы интеграции (API), позволяющие проводить подключение к сторонним системам (лаборатории, страховые компании и др.).

Таким образом, СППВР – сложная экосистема, объединяющая данные, аналитику и интерфейс. Её эффективность зависит от качества интеграции компонентов, актуальности клинических знаний использования передовых механизмов рассуждений и удобства взаимодействия для врачей. Рассмотрим типы механизмов рассуждений более подробно.
1.3.2 Классификация механизмов рассуждений в системах поддержки принятия врачебных решений

Классификация механизмов рассуждений в СППВР позволяет понять, как системы анализируют данные, интерпретируют информацию и формируют рекомендации. Эти механизмы варьируются от жестких логических правил до адаптивных алгоритмов МО. Существует несколько основных механизмов рассуждений, которые могут быть положены в основу СППВР [87-89]:

- 1. Системы на основе правил. Используют заранее заданные правила формата «ЕСЛИ-ТО», основанные на клинических протоколах и экспертных знаниях. Такая система является примером эвристического подхода, в котором отдельные логические утверждения в форме правил получают путем наблюдения за врачами-экспертами, а затем объединяют в попытке подражать процессам рассуждения экспертов. Этот подход впервые использовался в системе MYCIN в 1976 году для диагностики бактериальных инфекций и подбора антибиотиков [42]. Плюсами таких систем являются: прозрачность и интерпретируемость решений, а также относительная простота во внедрении стандартизированных протоколов. Недостатки связаны с тем, что такие системы не справляются с неопределенностью и сложными взаимосвязями, а также требуют постоянного обновления правил.
- (байесовские 2. Вероятностные методы сети, марковские модели). Учитывают вероятность событий и причинно-следственные связи между ними. Типичной системой принятия решений, основанной на знаниях, является байесовская сеть. Она показывает вероятностные отношения между наборами переменных – заболеваний и симптомов на основе условной вероятности, согласно теореме Байеса [55, 56]. Это сеть с явным требованием, чтобы отношения были причинно-следственными. Такие методы способны работать с неполными данными и неопределенностью, моделировать сложные медицинские сценарии. Например, они могут помогать моделировать прогресс заболевания во времени. Тем не менее, большим препятствием к их широкому внедрению является то, что в медицине иногда трудно напрямую определить,

что является следствием, а что – причиной. Недостатками такого подхода являются также требования к точности данных для обучения и сложность интерпретации при большом количестве переменных.

- 3. Рассуждения на основе прецедентов. Анализируют похожие случаи из базы данных, чтобы предложить решение для текущего пациента. К ним относятся, например, поиск пациентов с аналогичными симптомами и успешными схемами лечения. Такой подход учитывает клинический опыт и нетипичные случаи и подходит для ситуаций, где нет четких правил. С другой стороны, он критически зависит от качества и объема базы прецедентов.
- 4. Нейросетевые модели и глубокое машинное обучение. Используют современные модели НС для выявления скрытых зависимостей («паттернов») в больших объемах данных. Достоинством данного подхода является высокая точность в задачах классификации и прогнозирования. Однако современные НС требуют огромных вычислительных ресурсов и размеченных больших баз медицинских данных [54, 55], что не всегда реализуемо на данном этапе развития МИС.
- 5. Генетический алгоритм. Такой алгоритм не основан на знаниях, он использует итерационные процессы для перестройки себя и обеспечения оптимального решения на основе данных пациента.
- 6. Гибридные системы. Комбинируют несколько подходов для повышения точности и гибкости. Например, использование правил для первичного скрининга и НС для уточнения диагноза. Или интеграция рассуждений на основе прецедентов с байесовскими сетями для оценки рисков. Такие методы нивелируют слабые стороны отдельных методов и подходят для комплексных медицинских задач. Однако их практическое использование осложняется ввиду трудностей разработки и настройки подобных систем.
- 7. Онтологии и семантические модели. Используют формализованные знания (онтологии) для структурирования медицинских понятий и их взаимосвязей. Примерами их использования могут служить классификация заболеваний по SNOMED CT или МКБ-10, автоматическое извлечение симптомов из

текстовых записей с помощью технологий автоматической обработки текстовой информации (NLP). Достоинствами таких систем являются унификация терминологии и улучшение взаимодействия между системами, поддержка междисциплинарного анализа. В качестве недостатков обычно отмечают трудоемкость создания и поддержки онтологий.

Выбор того или иного механизма рассуждений зависит в большей степени от типа анализируемой медицинской задачи. Для диагностических задач хорошо подходят глубокие НС и рассуждения на основе прецедентов, для выбора способа лечения – системы на основе правил, для прогнозирования развития заболевания – байесовские сети и классические алгоритмы МО. Другим важным фактором является доступность данных. В случае наличия больших медицинских данных целесообразно использование методов глубокого МО, в случае дефицита данных предпочтительней использование правил и онтологий [54, 55].

Также следует заметить, что в случае жестких юридических требований на интерпретацию результатов механизма рассуждений должны использоваться логические, а не глубокие вероятностные системы, выступающие во многих случаях, как «черный ящик».

В рамках данной работы основой для разрабатываемой системы анализа эндоскопических видеоизображений выбраны глубокие НС. Предоставляемая поддержка принятия решений основана на процессах сложной оценки результатов и алгоритмах, использующих базы размеченных данных и ИИ (рисунок 1.5). Важным фактором является также интеграция СППВР в рабочие процессы медицинских специалистов.



Рисунок 1.5 – Структура СППВР с ядром на основе нейронной сети

Известно, что HC меняет свое поведение на основе предыдущих паттернов, т.е. обучается. Другой особенностью HC является отсутствие необходимости понимать сложные отношения между входными переменными; сеть изучает эти отношения по мере изменения связей между ее узлами. В HC связи устанавливаются по мере развития сети, часто на основе процесса обучения, без учета патофизиологических фактов.

В настоящее время одним из наиболее эффективных подходов, используемых в основе СППР на основе анализа изображений, является использование сверточных нейронных сетей (СНС) – особого типа НС, предложенного Я. Лекуном [55]. Данный подход копирует ряд особенностей зрительной коры головного мозга. Структура СНС является многослойной и включает в себя чередование сверточных слоев, осуществляющих свертку входного сигнала с ядром (матрицей) свертки, а также слоев подвыборки (например, maxpooling), осуществляющих снижение размерности. Пример архитектуры СНС приведен на рисунке 1.6. Она состоит из нескольких видов слоев [57]:

- входной слой;
- сверточный слой;
- субдискретизирующий (подвыборочный) слой;
- полносвязный слой;
- выходной слой.



Рисунок 1.6 – Пример архитектуры сверточной нейронной сети

Под обучением СНС понимают поиск оптимальных значений весов слоев, при которых достигается минимум ошибки, выдаваемой НС на обучающей выборке. Чаще всего для обучения используется алгоритм обратного распространения ошибки [46, 50, 53].

Существующим ограничением нейросетевого подхода является то, что, в отличие от других рассматриваемых систем, «правила», используемые сетью, не следуют определенной логике и в большинстве случаев не являются полностью понятными для специалистов. В этом смысле НС ведет себя как «черный ящик», что в определенной степени тормозит проникновение таких технологий в медицинскую практику. Для решения этой проблемы развивается отдельный класс интерпретируемых HC [55].

Современные нейросетевые модели показывают высокое качество обучения на больших массивах данных. Однако проблема с использованием НС связана с необходимостью сбора и экспертной разметки огромных баз медицинских данных (порядка нескольких тысяч изображений и более на класс или патологию), которые часто недоступны, а во многих случаях и вовсе не существуют. В связи с этим возникает необходимость поиска подходов к построению робастных нейросетевых моделей на базах данных относительно малого размера. Зачастую эта проблема может быть решена путем предобучения используемой нейросетевой модели (HM) на изображениях других, в том числе немедицинских объектов. Кроме того, имеющиеся для обучения данные могут быть размножены (аугментированы) путем различных математических преобразований.

1.3.3 Требования к системам поддержки принятия врачебных решений

Понимание основных причин, которые приводят либо к успеху, либо к провалу, может помочь повысить эффективность разработки и развертывания СППВР в повседневной практике. Сбои могут возникать на разных этапах разработки и внедрения: неспособность удовлетворить техническим требованиям, неудача в принятии системы врачами и сложность в интеграции системы в информационную или организационную среду. Основными требованиями, позволяющими повысить эффективность использования таких систем, являются [89, 90]:

- Система должна автоматически обеспечивать поддержку принятия решений в рамках РП врача. Внедрения, в которых врачи должны обращаться за советом вручную, признаны неуспешными.
- Поддержка принятия решений должна предоставляться в режиме реального времени (PPB). Если врач должен прервать обычную процедуру ухода за пациентом, чтобы перейти на отдельную рабочую станцию или следовать сложным, длительным процедурам запуска, маловероятно, что такая система будет востребована для оптимизации лечебных и диагностических процессов.
- Ядро системы должно быть включено в более крупную ИС, которая уже является частью профессиональной деятельности врачей.
- Системы должны иметь возможность получать новые данные автоматически.
 Из-за необходимости прикладывать усилия для ввода новых данных о пациентах врачи избегают этого процесса, который необходим для дообучения систем.
- Интерфейс системы должен быть удобным для навигации и использования.
- Должна присутствовать периодическая обратная связь по вопросу использования врачом СППР.

Таким образом, при разработке СППВР существуют факторы, помимо программного обеспечения и данных, которые необходимо учитывать. Важными вопросами также являются приемлемая стоимость и доступность аппаратного обеспечения, техническая поддержка и обучение работе с системой, уровень интеграции системы в клинический рабочий процесс, а также актуальность и своевременность предоставляемых сообщений-подсказок. Рассмотрим вопросы интеграции СППВР с эндоскопическими системами.

1.4 Современная эндоскопическая техника

Современные эндоскопы используют белый свет (white light imaging, WLI) и CMOS (Complementary Metal-Oxide-Semiconductor, комплементарный металло-

оксидный полупроводник, КМОП) или ССD (Charge-Coupled Device, прибор с зарядовой связью, ПЗС) матрицы для детализации слизистой. Современные КМОП-матрицы практически сравнялись с технологией ПЗС по качеству изображения, но превосходят их в скорости, энергосбережении и стоимости. Для большинства клинических задач КМОП считается более перспективным выбором, особенно в комбинации с улучшенными алгоритмами обработки видеосигнала. Однако в отдельных приложениях, где требуется абсолютная точность в условиях экстремально низкой освещенности, ПЗС все еще может сохранять преимущество. Разрешение видеоизображений в современной практике варьируется от SD (720х576 пикселей) до Full HD (1920×1024 пикселя) и 4К (3840×2160 пикселей) с перспективой увеличения до 8К (7680×4320 пикселей). Повышение разрешения позволяет проводить высокую детализацию опухолей, язв, эрозий и других патологических участков. Сравнение типичных разрешений для видеосигналов приведено на рисунке 1.7 [36].



Рисунок 1.7 – Сравнение различных разрешений видеоизображений

Также в современной эндоскопии ЖКТ используется ряд инновационных технологий, направленных на улучшение качества диагностики. Одним из примеров является режим NBI (Narrow Band Imaging), позволяющий производить

осмотр тканей в узкоспектральном диапазоне видимого света. При его использовании ткани освещаются светом с двумя длинами волн: синий свет с длиной волны 415 нм проникает только в поверхностные слои слизистой оболочки и хорошо поглощается гемоглобином в поверхностных капиллярах, а зеленый свет с длиной волны 540 нм способен проникать чуть глубже и поглощаться более глубокими сосудистыми структурами (рисунок 1.8). Таким образом, фильтры сужают спектр света до синего и зеленого и патологические участки (например, неоплазии, воспаления и др.) становятся более контрастными, так как их сосудистая сеть отличается от здоровой ткани [32, 91-94].



Рисунок 1.8 – Узкоспектральная визуализации режима NBI

Узкоспектральный режим повышает выявляемость следующих заболеваний: раннего рака желудка, пищевода, толстой кишки; плоских аденом и дисплазий (например, при барреттовом пищеводе). Технология помогает точно определить края опухолей или зон дисплазии, что критично для планирования биопсии, хирургического или эндоскопического удаления новообразования. Использование NBI доброкачественные режима позволяет отличить изменения OT злокачественных, поскольку сосудистый рисунок при опухолях имеет хаотичную структуру, а поверхность слизистой при дисплазии часто неоднородна. При колоноскопии режим NBI помогает отличить гиперпластические полипы от без биопсии Таким аденоматозных с высокой точностью. образом, узкоспектральный режим – широко используемая технология в современной эндоскопии, повышающая точность диагностики, снижающая количество способствующая ложноотрицательных результатов и раннему выявлению онкологических заболеваний. Однако эффективность ее использования зависит от опыта врача и качества оборудования.

Примеры изображений [95] одного и того же участка слизистой в режиме NBI и при осмотре в классическом белом свете приведены на рисунок 1.9.



Рисунок 1.9 – Пример использования режима NBI для диагностики патологии в желудке по микрососудистому рисунку: а) кадр в режиме осмотра в белом свете; б) кадр в режиме NBI (стрелки указывают на область с патологией)

Другим примером технологии, широко используемой в современной эндоскопической диагностике, является режим с оптическим увеличением, достигающим 150 крат, для изучения ямочного рисунка слизистой. При подобном хорошо визуализируется микрососудистый увеличении рисунок И микроархитектоника поверхности слизистой оболочки, что приближает данную методику к оптической микроскопии [32, 92]. Особое значение данный режим имеет при осмотре пациентов с множественными участками, подозрительными на наличие патологии. В таких случаях невозможно осуществить отбор тканей путем биопсии из всех патологических очагов ввиду травматичности данной процедуры, и использование эндоскопии с сильным увеличением позволяет снижать количество биопсий на 80-90% [93-98]. При этом дополнительное использование режима NBI способствует лучшей визуализации. Пример изображений [95] в режиме с увеличением приведен на рисунке 1.10.



Рисунок 1.10 – Увеличительная эндоскопия: а) режим работы без увеличения; б) увеличительная эндоскопия (115 крат) с использованием режима NBI

Еще одним способом улучшения визуализации при проведении эндоскопии является нанесение безопасных красителей на слизистую оболочку ЖКТ – хромоэндоскопия [32, 94]. Данный метод основан на том, что здоровые и больные клетки по-разному поглощают краситель, что позволяет выявлять небольшие поражения и неоплазии на ранней стадии их формирования. Среди красителей

широко используются индигокармин, конго красный, фенол красный, раствор уксусной кислоты (ацетохромоскопия). Пример изображений [95] при использовании хромоэндоскопии приведен на рисунке 1.11.





Рисунок 1.11 – Хромоэндоскопия с использованием раствора уксусной кислоты: а) эндоскопическое изображение без использования красителя; б) эндоскопическое изображение после орошения слизистой оболочки раствором уксусной кислоты

В таблице 1.1 приводится сравнение двух технологий, из которого следует, что хромоэндоскопия не требует дорогостоящего оборудования – достаточно красителей и стандартного эндоскопа, при этом предоставляя разумную альтернативу для клиник без доступа к оборудованию с NBI или аналогичным цифровым технологиям.

Таблица 1.1 – Сравнение	технологий NBI и х	ромоэндоскопии
-------------------------	--------------------	----------------

Параметр	NBI	Хромоэндоскопия
Стоимость	Высокая	Низкая (красители дешевы)
	(спецоборудование)	
Время процедуры	Не меняется	Увеличено на 5-10 минут
Риск аллергии	Нет	Есть
Детализация	Сосуды+поверхность	Поверхностные структуры

Следующей технологий, повышающей качество диагностики, является эндоскопическая ультрасонография. Она представляет собой сочетание эндоскопа и ультразвукового датчика и позволяет визуализировать не только поверхность слизистых оболочек, но и глубокие слои тканей, а также окружающие органы и структуры. На конце гибкого эндоскопа устанавливается ультразвуковой датчик, который генерирует высокочастотные акустические волны (5–20 МГц). Он может быть радиальным, создавая круговое изображение (360°), подходящее для общей оценки, или линейным, обеспечивая изображение в одной плоскости, что удобно для биопсии. Такое оборудование позволяет повысить достоверность для диагностики патологий поджелудочной железы, желчных протоков и раннего рака ЖКТ. Следует отметить, что эндоскопическая ультрасонография требует высокой квалификации врача и возможности закупки дорогостоящего специализированного оборудования [32].

Еще одна современная технология – конфокальная лазерная эндомикроскопия. В этом случае миниатюрный конфокальный микроскоп интегрирован в кончик эндоскопа. Слово «конфокальная» происходит от латинского con-focus («с общим фокусом»). Это означает, что в системе используются точечный источник света (лазер), фокусирующийся на конкретной глубине ткани, и детектор с «пинхолом» (маленьким отверстием), улавливающий только свет, отраженный из той же фокальной плоскости. Принцип работы такой системы, следующий: лазер сканирует ткань послойно, фокусируясь на определенной глубине. Рассеянный свет из других слоев блокируется пинхолом, что устраняет «шум» и повышает четкость изображения. В результате формируется высококонтрастное изображение с разрешением до 1 микрона (на уровне отдельных клеток). Эта технология позволяет получать микроскопические изображения тканей в реальном времени на клеточном уровне во время эндоскопических процедур. Она позволяет получать изображения, сравнимые с гистологическими срезами, но без удаления ткани. На сегодняшний день подобные аппаратные комплексы дороги для клиник, поэтому не являются широко распространёнными [32, 91].

Используется также капсульная эндоскопия [99, 100], основанная на следующем принципе: пациент проглатывает капсулу с миниатюрной видеокамерой, передающей изображение на внешнее записывающее устройство. Изображения слизистой записываются, пока капсула естественным образом продвигается по ЖКТ. Она позволяет визуализировать тонкий кишечник – участок, который невозможно осмотреть при гастроскопии или колоноскопии. Параметры современных капсул: разрешение 256×256 пикселей, скорость съемки 3 кадра/сек, продолжительность работы до 8 часов (рисунок 1.12). Капсула сохраняет порядка 50 000 кадров за процедуру, фиксируя малейшие изменения ЖКТ. Однако ее факторами: применение ограничено следующими относительно высокой биопсии стоимостью процедуры, невозможностью ИЛИ терапевтических манипуляций, риском застревания, ограниченностью диагностики в желудке и толстой кишке, низким качеством сохраненных изображений. Капсульная эндоскопия – ценный, особенно для исследования труднодоступной тонкой кишки, но все-таки вспомогательный метод. Она не заменяет стандартную гастроскопию и колоноскопию, особенно когда требуется биопсия или лечение.



Рисунок 1.12 – Внешний вид капсул для эндоскопии

В заключение отметим, что проведение эндоскопической диагностики с помощью современного оборудования требует от врача высокого уровня подготовленности [22, 23, 35, 92]:

 наличия знаний по таким областям медицины как гастроэнтерология, морфология, гистология, онкология и др.;

- умения пользоваться методиками узкоспектральной и увеличительной эндоскопии;
- умения интерпретировать результаты капсульной эндоскопии;
- наличия опыта использования указанных знаний и умений в реальной клинической практике.

Перечисленные факторы повышают актуальность разработки эффективных СППВР в помощь врачу при эндоскопической диагностике.

1.5 Системы поддержки принятия врачебных решений в эндоскопической диагностике

Даже среди опытных врачей существует риск пропуска особо значимых областей во время эндоскопической диагностики, и, как следствие, принятие неверного решения при постановке диагноза пациенту [22, 23]. Внедрение СППВР при интерпретации видеоданных, полученных с эндоскопа, позволяет повысить эффективность и точность проводимой диагностики, снизить количество пропусков патологий, дать врачу дополнительную уверенность касательно принимаемых им решений [88]. Заметным преимуществом таких систем является то, что при таком подходе не требуется вносить принципиальные изменения в процесс проведения самой процедуры эндоскопии.

Введение в работу врачей-эндоскопистов СППВР способствовало появлению направления автоматизированной эндоскопии (computer-aided endoscopy), которое в настоящее время активно развивается. Для гибкой эндоскопии первые СППВР появились в конце прошлого века, для капсульной эндоскопии – в 2004 г. Основная направленность подобных систем – помощь специалистам в поиске разных видов патологий ЖКТ (злокачественные изменения, доброкачественные изменения, опухоли, полипы, пищевод Баррета, болезнь Крона и др.) [32].

Как уже отмечалось, одним из важнейших критериев успешности внедрения СППВР является возможность ее интеграции в уже существующую МИС, с которой на постоянной основе работают специалисты конкретного медицинского учреждения. Так, в настоящее время в эндоскопическом отделении ГБУЗ ЯO «Клиническая больница» (ЯОКОБ) онкологическая активно используется эндоскопическая медицинская информационная система (ЭМИС), разработанная специалистами ООО «ЭМИС» и разработчиками из Ярославского государственного университета им. П.Г. Демидова. Она позволяет не только работать с медицинской информацией пациентов, полученной в ходе диагностики, оптимизировать работу специалистов но эндоскопического отделения, И осуществлять контроль качества эндоскопических исследований, обрабатывать и сохранять видеоданные на выходе эндоскопов [101].

Необходимо, чтобы создаваемая СППВР могла взаимодействовать с ЭМИС так, как это показано на рисунке 1.13. Передаваемые по аналоговому или цифровому интерфейсу в режиме реального времени в компьютер специалиста видеофайлы имеют то же качество, что и исходные видеоданные, которые специалист анализирует на мониторе эндоскопа в ходе исследования. По окончании исследования итоговый видеофайл сохраняется в системе ЭМИС в папку в электронной базе данных, соответствующую определенному пациенту.



Рисунок 1.13 – Взаимодействие СППВР с ЭМИС

Помимо видеозаписей исследований в подобных папках хранится и другая персонализированная предпроцедурная, интрапроцедурная и постпроцедурная информация о пациенте и проведенной процедуре эндоскопии. Обычно она включает: подписанные согласия пациента, результаты анализов, информацию о состоянии здоровья в течение стационарного пребывания в эндоскопическом отделении, поставленном диагнозе, предыдущих обследованиях и диагнозах, медикаментозных препаратах, проведенных биопсиях и т.д. Также специалист осуществляет фотофиксацию участков слизистой оболочки ЖКТ, которые, на его взгляд, могут содержать патологию. Это в значительной степени экономит время при осуществлении последующего анализа результатов эндоскопической диагностики или при возникновении спорных ситуаций, когда требуется повторно просмотреть видеозапись эндоскопического исследования.

Далее в настоящей работе предполагается, что видеопоток поступает на СППВР, собой которая представляет аппаратно-программный комплекс, включающий алгоритмы ИИ. Они обрабатывают эндоскопические изображения, извлекаемые из видео, и определяют наличие патологии. Результат обработки отображается на дополнительном мониторе в режиме реального времени. При этом обнаруженные патологии области ЖКТ заключаются в цветные рамки или сопровождаются звуковыми сигналами. Пример работы врача ЯОКОБ с подобной системой представлен на рисунке 1.14. При соблюдении высоких значений параметров скорости работы нейросетевого алгоритма и его точности подобная система может эффективно дополнить существующую процедуру проведения эндоскопической диагностики.



Рисунок 1.14 – Использование СППВР при выполнении гастроскопии желудка

Таким образом, использование СППВР в эндоскопии позволяет снизить временные и финансовые расходы на проведение исследований, обеспечивает повышение точности и эффективности диагностики, а также может использоваться для обучения начинающих специалистов и повышения квалификации в области эндоскопического анализа ЖКТ.

1.6 Нейронные сети как инструмент глубокого машинного обучения

Под HC подразумевается совокупность нейронов, расположенных последовательно друг за другом и соединенных определенными связями – синапсами [46-50]. Нейрон принимает на вход совокупность сигналов $x_1, x_2, ..., x_n$, которые после прохождения синапсов суммируются, будучи умноженными на коэффициенты (веса) $w_1, w_2, ..., w_n$ согласно формуле:

$$S = \sum_{i}^{n} x_{i} w_{i}.$$

Результирующее выходное значение получается после применения к величине *S* функции активации φ:

$$Y = \varphi(S).$$

Обучаются НС на больших наборах данных, содержащих пары «вход-выход». В процессе обучения система автоматически подбирает оптимальные весовые коэффициенты. При подаче входных данных они последовательно преобразуются слоями сети (путем умножения на веса). Сравнивая полученный результат с ожидаемым выходным значением, НС корректирует свои веса так, чтобы минимизировать разницу между предсказанием и истинным значением (рисунок 1.15) [55, 56, 95].

Из-за огромного числа параметров (сотни тысяч или миллионы) ручная настройка весов НС невозможна. Для решения этой задачи вводится функция

потерь (loss function). Она сравнивает выход сети с эталонным значением и вычисляет численную меру их расхождения (рисунок 1.16), которая и используется для оптимизации весов.



Рисунок 1.15 – Подбор весов нейронной сети в процессе обучения



Рисунок 1.16 – Принцип работы функции потерь

Качество работы HC напрямую оценивается через значение функции потерь. Эта функция количественно измеряет расхождение (ошибку) между предсказаниями сети и истинными (эталонными) значениями для предоставленных обучающих данных. Чем точнее предсказания HC, тем меньше величина ошибки, и, соответственно, тем ниже значение функции потерь. Поэтому фундаментальная задача этапа обучения HC – оптимизация (минимизация) функции потерь.

Поскольку ручной подбор параметров (весов) для модели в данном случае неосуществим, для автоматизации и эффективного решения задачи минимизации используются специальные алгоритмы – оптимизаторы [56]. Эти алгоритмы (например, Stochastic Gradient Descent – SGD, Adam, RMSProp) определяют направление и величину корректировки каждого веса на основе вычисленного градиента функции потерь относительно параметров модели.

Обучение может начинаться со случайной инициализации весов сети. Поскольку эти начальные значения обычно не соответствуют оптимальным, первоначальное значение функции потерь оказывается высоким, отражая значительную ошибку предсказаний. Процедура обучения представляет собой итеративный циклический процесс. За один цикл (эпоху) модель последовательно обрабатывает обучающие данные (или их пакеты – батчи). На каждой итерации:

- Данные проходят через сеть (прямое распространение forward pass).
- Вычисляется значение функции потерь для текущих предсказаний.
- С помощью алгоритма обратного распространения ошибки (backpropagation)
 [57, 102] рассчитываются градиенты функции потерь по всем весам сети, показывающие, как именно нужно изменить каждый вес, чтобы уменьшить общую ошибку.
- Оптимизатор использует эти градиенты для обновления (корректировки) весов в направлении, минимизирующем потери.

Этот процесс повторяется многократно – от десятков до сотен или даже тысяч эпох. С каждой итерацией значение функции потерь постепенно уменьшается, а точность предсказаний сети на обучающих данных повышается, приближаясь к истинным значениям. Когда дальнейшее значительное снижение потерь становится невозможным или достигается требуемый уровень точности, модель считается обученной или достигшей сходимости.

Весь описанный процесс, где HC обучается на наборе данных с известными правильными ответами (парами «вход-целевой выход»), является классическим примером обучения с учителем (supervised learning) (рисунок 1.17).



Рисунок 1.17 – Использование оптимизатора в обучении нейронной сети

Ошибка, совершенная НС, в ходе обработки входных данных чаще всего определяется выражением:

$$E(\lbrace w_{i,j}\rbrace) = \frac{1}{2}\sum_{k \in OUT} (t_k - y_k)^2,$$

где E – среднеквадратическая ошибка; $w_{i,j}$ – значение весов между i, j; OUT – совокупность выходов HC; t_k – истинное значение на выходе, $k \in OUT$; y_k – предсказание сети. При наличии нескольких классов объектов возможно использование выражения:

$$E(\lbrace w_{i,j}\rbrace) = \frac{1}{k} \sum_{k \in OUT} (-\log (p_k)),$$

где E – кросс-энтропия, p_k – вероятность принадлежности объекта X_k к классу T_k .

Корректировка весов чаще всего осуществляется путем использования методов градиентного спуска [54, 55]:

$$\Delta w_{i,j} = -\mu \frac{\partial E}{\partial w_{i,j}},$$

где E – дифференцируемая функция, а параметр $0 < \mu < 1$, характеризует скорость движения. Суть метода градиентного спуска – перемещение по пространству весов против направления градиента. Это позволяет последовательно уменьшать значение функции потерь, приближаясь к ее минимуму.

Глубокое обучение – фундаментальное и активно развивающееся направление МО. Его ключевая особенность и отличие от классических подходов заключается в автоматическом построении признаков (features) исходных данных. Это достигается не путем их ручного конструирования, что характерно для методов классического МО, а через последовательное преобразование информации при прохождении через каскад слоев НС. Каждый последующий слой в этой цепочке выявляет все более сложные и абстрактные закономерности, основываясь на представлениях, сформированных предыдущим слоем [55]. Например, в компьютерном зрении начальные слои могут обнаруживать простые границы и текстуры, промежуточные – собирать их в части объектов, а глубокие слои – распознавать целые объекты (полипы, опухоли).

Термин «глубина» в этом случае является строго техническим понятием. Он не подразумевает более глубокое философское понимание данных системой, а однозначно указывает на количество последовательных уровней преобразования (слоев) в модели [55, 95]. Каждый слой выполняет специфические преобразования (линейные операции, за которыми следуют нелинейные функции активации), а совокупность слоев позволяет модели аппроксимировать чрезвычайно сложные функции. Современные модели (такие как глубокие сверточные нейронные сети или трансформеры) включают десятки и сотни таких слоев, что было вычислительно невозможно представить всего 10-15 лет назад.

Глубокое обучение является наиболее современной областью МО, которая, в свою очередь, представляет собой одну из ключевых ветвей ИИ. Глубокие модели достигли революционных успехов в задачах, где данные обладают высокой структурной сложностью (изображения, речь, текст) [55, 56, 95]. Взаимосвязь и вложенность этих областей знаний наглядно иллюстрируется на Рисунке 1.18.



Рисунок 1.18 – Взаимосвязь терминов искусственного интеллекта, машинного и глубокого обучения

1.7 Использование нейронных сетей для анализа медицинских изображений

Современные модели HC, которые могут использоваться для обработки и анализа эндоскопических видеоизображений активно развиваются, сочетая как уже ставшие классическими архитектуры CHC, так и постоянно разрабатываемые новые структуры [103-106]. Рассмотрим кратко ключевые современные нейросетевые архитектуры, их достоинства и недостатки.

1.7.1 Сверточные нейронные сети

Наиболее часто в задачах обработки изображений используются СНС [103]. В них осуществляется обучение с учителем [7]. Обучение происходит с помощью алгоритма обратного распространения ошибки путем минимизации функции потерь относительно неизвестных весов *W* [55, 107-109]:

$$L = -\frac{1}{|X|} \sum_{i}^{|X|} \ln\left(p(y^{i}|X^{i})\right),$$

где |X| определяет количество изображений в обучающем наборе, X^i – i-е изображение в обучающем наборе, которому соответствует метка y^i , $p(y^i|X^i)$ – условная вероятность, при которой X^i правильно классифицировано. Стохастический градиентный спуск используется для минимизации этой функции стоимости, когда стоимость всего обучающего набора аппроксимируется с затратами на мини-наборы (батчи) данных. Если W_l^t характеризует веса *l*-го сверточного слоя на итерации *t*, а \hat{L} – стоимость по мини-батчу размером *N*, тогда обновленные веса в следующей итерации вычисляются следующим образом:

$$W_l^{t+1} = W_l^t + V_l^{t+1},$$

где $\gamma^{t} = \gamma^{\left\lfloor \frac{tN}{|X|} \right\rfloor}$, $V_{l}^{t+1} = \mu V_{l}^{t} - \gamma^{t} \alpha_{l} \frac{\partial \hat{L}}{\partial W_{l}}$, α_{l} – скорость обучения *l*-го слоя, μ – импульс, указывающий на вклад предыдущего обновления веса в текущей итерации, γ – скорость планирования, уменьшающая скорость обучения α в конце каждой эпохи.

Для структуры СНС характерно наличие (чередование) слоев свертки и слоев субдискретизации, образованных картами признаков [56, 110-113].

При распространении изображения от входа к выходу НС осуществляется операция математической свертки этого изображения с определенным набором фильтров. Фильтры, или ядра свертки, представляют собой набор весов нейронов скрытого слоя (рисунок 1.19). Операция свертки подразумевает вычисление скалярных произведений соответствующих пикселей ядра свертки (матрицы весов) и входного изображения, а также организацию локального поля для каждого нейрона карты признаков.



Рисунок 1.19 – Процедура обработки пикселей изображения на сверточном слое

Свертка – специализированная линейная операция, которая широко используется в различных областях, включая обработку изображений, статистику, физику [45, 53, 54]:

$$S(i,j) = \sum_{m} \sum_{n} I(m,n)k(i-m,j-n).$$

Множество фильтров с разными весовыми коэффициентами представляет множество свойств, которые могут быть идентифицированы в процессе такой обработки. В качестве свойств первого сверточного слоя могут выступать прямые линии, кривые линии, простые цвета. Полученная в результате данной процедуры карта признаков характеризует наличие определенного признака в той или иной части изображения. Чем большее количество фильтров используется (больше количество рассматриваемых свойств), тем глубже становится результирующая карта признаков, и соответственно, тем больше информации имеется об изображении на входе сети.

Увеличение в дальнейшем числа сверточных слоев приводит к усложнению искомых свойств (например, поиск геометрических фигур, частей тела человека и т.д.), что в дальнейшем способствует распознаванию изображения [55, 114-117].

Несмотря на потерю информации о точном расположении определенного признака, сохраняется и в дальнейшем учитывается информация о взаимном расположении разных признаков.

В некоторых архитектурах СНС за сверточным слоем следует слой активации, когда скалярный результат каждой свертки попадает на нелинейную функцию [54, 95]. Распространенной функцией активации является ReLU, принцип действия которой изображен на рисунок 1.20.



Рисунок 1.20 – Процедура обработки пикселей изображения на слое активации с функцией ReLU

Основной целью данного слоя является повышение инвариантности данных к деформациям и смещениям, путем обеспечения устойчивости к трансляции локальных признаков в пределах одного региона [54]. Операции, выполняемые данным слоем, осуществляются на входных данных, в качестве которых выступают данные на выходе предыдущего сверточного слоя.

Обеспечение инвариантности к искажениям в пространстве реализуется следующим образом. Происходит рассмотрение регионов фиксированного размера $k \times k$ и агрегация значений на выходе сверточного слоя. Тогда, если один и тот же признак сдвинут на значение ($\Delta x, \Delta y$), не превышающие k, соответствующий нейрон, содержащий локальный признак изображения, останется активным. В большинстве архитектур агрегирующей функцией является функция нахождения максимального значения (рисунок 1.21). Выбор величины k – размера локального рецептивного поля нейрона – является важным этапом создания сети. Выбор малого размера окна даст возможность находить малые признаки, но в случае присутствия на изображении тех же признаков в большем масштабе могут

возникнуть ситуации пропуска. Полученные данные поступают в виде вектора значений на полносвязный слой из N нейронов, где N – число возможных классов объектов на изображении. Вычисление произведения весов полносвязного слоя с результатом работы предыдущих слоев позволяет определить вероятности принадлежности к каждому классу. Выбор класса происходит по наибольшей вероятности [54-56, 95].



Рисунок 1.21 – Процедура обработки пикселей изображения на слое субдискретизации

К достоинствам СНС можно отнести способность выявления схожих признаков у изображений, относящихся к одному классу, инвариантность к расположению и размеру объекта на изображении за счет использования свертки, инвариантность К пространственным искажениям (поворот, смещение, масштабирование), возможность использования CHC для классификации изображений с характеристиками разными использования за счет субдискретизации.

К недостаткам СНС относятся фиксированный размер входных изображений, необходимость наличия большого количества обучающих данных и высокая длительность обучения [55, 118].

Примерами современных архитектур СНС являются: SSD, ResNet, EfficientNet, YOLO [110-113]. Они будут рассмотрены ниже при описании нейросетевых детекторов объектов на изображении.

1.7.2 Другие современные архитектуры для систем глубокого обучения

Кратко рассмотрим другие варианты архитектуры HC, использующихся для анализа и интерпретации медицинских изображений.

Трансформеры, изначально разработанные для задач обработки текста (NLP), медицинской визуализации адаптированы для через архитектуру Vision Transformers (ViT). К достоинствам трансформера обычно относят высокую точность за счет глобального контекста: алгоритм «видит» связи между удаленными участками изображения [119-121]. В отличие от СНС, ограниченных окрестностями, механизм внимания (self-attention) вычисляет локальными взаимовлияние всех участков изображения (патчей), даже пространственно удаленных от текущей анализируемой области. Такая архитектура эффективна для обнаружения редких патологий (например, ранних опухолей размером <5 мм). К недостаткам трансформера относятся требование огромных объемов данных для обучения (что особенно сложная проблема для медицины) и повышенная по сравнению с СНС вычислительная сложность обработки одного видеокадра.

Гибридные архитектуры (СНС + трансформер) комбинируют логику СНС для извлечения локальных признаков и трансформера для глобального контекста. Они обычно достаточно сложны в практической реализации и предъявляют высокие требования к объему памяти. Кроме того, существует риск переобучения на сравнительно небольших медицинских наборах видеоданных [122, 123]. Пример использования гибридной архитектуры будет приведен в главе 3 настоящей работы.

Специализированные архитектуры обработки для медицинских изображений. Они оптимизированы для специфических задач, например, обнаружение кровотечений или оценка качества подготовки кишечника, а также устойчивы К шумам И помехам, характерным для эндоскопических видеоизображений. К их недостаткам относят сложную адаптацию для других типов патологий и ограниченную доступность открытых реализаций. Пример

использования специализированной архитектуры будет приведен в главе 4 настоящей работы.

Подводя итог краткому рассмотрению современных архитектур HC для анализа эндоскопических изображений, следует отметить, что их использование в практических задачах позволяет в ряде случаев достичь точности, сопоставимой с врачами-эндоскопистами. Выбор конкретной модели зависит от конкретных технических требований. К ним часто относятся возможность работы в режиме реального времени на имеющемся оборудовании, обеспечение максимальной точности, возможность интерпретируемости результатов, возможность проводить обучение при ограниченном наборе видеоданных.

1.8 Современные методы и алгоритмы обработки и анализа эндоскопических видеоизображений

Эндоскопические видеоизображения обладают уникальными особенностями, связанными с условиями их получения [32]. Во-первых, съемка ведется в узких полостях тела с ограниченным освещением, что приводит к низкому уровню яркости, неравномерной засветке и появлению бликов от влажных поверхностей. Во-вторых, изображения часто содержат шумы и искажения (шум матрицы, артефакты сжатия) и динамические искажения из-за движения эндоскопа или сокращений органов (например, перистальтики кишечника). Кроме того, цветопередача может искажаться из-за особенностей тканей (кровь, слизь) и источников света, что затрудняет визуальную диагностику [31, 35, 124-126].

Соответствующие алгоритмы цифровой обработки изображений (ЦОИ) должны быть адаптированы к работе в РРВ, часто в условиях ограниченных аппаратных ресурсов; робастны к вариативности данных – изменению освещения, углов обзора, наличию препятствий (пузыри, инструменты) и других артефактов. Также важным является сохранение диагностически значимых деталей объектов при ЦОИ [127-131]. При использовании алгоритмов фильтрации или сжатия нельзя потерять мелкие структуры (сосуды, микроопухоли), а цветокоррекция не должна искажать естественные оттенки тканей.

1.8.1 Классификация задач обработки эндоскопических видеоизображений

Алгоритмы обработки и анализа изображений используются на разных этапах эндоскопической диагностики [132] (до процедуры, во время процедуры, после процедуры), как показано на рисунок 1.22. Рассмотрим некоторые классы алгоритмов более подробно.



Рисунок 1.22 – Классификация задач в области обработки и анализа эндоскопических видеоизображений

Алгоритмы предобработки эндоскопических изображений включают следующие категории, охватывающие основные задачи подготовки видеоданных для последующего этапа визуального анализа [36, 39, 124, 132-137]:

- Алгоритмы улучшения изображений. К ним можно отнести: алгоритмы компенсации неравномерного освещения, адаптивную коррекцию гистограмм, повышение контрастности, устранение засветок и бликов, улучшение цветопередачи.
- Восстановление изображений. Здесь следует отметить алгоритмы устранения размытия (blind deconvolution), коррекции геометрических искажений, алгоритмы для заполнения дефектных областей (inpainting).

- Оценка качества изображений. Детекция размытия, расфокусировки или компрессионных артефактов, оценка контраста и освещенности, расчет метрик резкости.
- Алгоритмы фильтрации и подавления шумов. Пространственная фильтрация, частотная фильтрация, в том числе в области вейвлетпреобразования, адаптивные фильтры, нелокальное сглаживание (non-local means), фильтры на основе НС [41, 52, 138-143].
- Сегментация области интереса (ROI). Классические пороговые методы, методы на основе активных контуров, нейросетевые алгоритмы сегментации областей.
- Детектирование области интереса (ROI). Нейросетевые алгоритмы выделения сосудов, опухолей, полипов.
- Компенсация движения. Анализ оптического потока (ОП) для коррекции дрожания камеры, совмещения последовательных кадров.

Алгоритмы PPB используются непосредственно для поддержки исполняемой процедуры и делятся на [37, 40, 126-128, 132]:

- Улучшение визуализации в реальном времени. Алгоритм коррекции СLAHE для выделения сосудов, динамическая коррекция цвета (баланс белого), фильтрация шумов в реальном времени, устранение бликов и засветок.
- Алгоритмы поддержки диагностики. Детекция и классификация полипов, опухолей, язв в реальном времени, классификация тканей для определения злокачественности/доброкачественности, определение глубины инвазии для злокачественных опухолей.
- Алгоритмы поддержки хирургии. Сопровождение областей инструментов и тканей, алгоритмы дополненной реальности, оценка расстояния до стенок органа, детекция кровотечений.

Алгоритмы постобработки, применяемые после эндоскопической процедуры, фокусируются на анализе, архивации и подготовке данных для последующего

использования (диагностики, обучения, детальных исследований). Их основные классы включают [124, 126, 128, 132, 144-147]:

- Контроль качества данных. Обнаружение дефектов (размытие, блики, шум), детектирование пропущенных кадров, валидация полноты данных, оценка диагностической ценности, проверка освещения и контраста.
- Аннотирование и разметка. Сегментация патологий, текстовое аннотирование, добавление метаданных, привязка к медицинским стандартам через онтологии.
- Сжатие видеоданных. Использование стандарта JPEG 2000 с ROI, динамическое управление битрейтом, преобразование RAW-данных в стандартизованные форматы (DICOM).
- Хранение и управление данными. Автоматическая категоризация по диагнозам, датам, локализации; удаление персональных данных из метаданных; шифрование; интеграция с медицинскими системами (PACS).
- Анализ и генерация отчетов. Статистический анализ, оценка динамики (например, рост полипа за год), автозаполнение отчетов, построение пространственных или объемных визуализаций толстого кишечника или желудка по серии кадров.

Описанные выше алгоритмы ЦОИ обеспечивают долгосрочное использование видеоданных для аудита, обучения систем ИИ и улучшения клинических практик. Рассмотрим некоторые из рассмотренных задач ЦОИ более подробно.

1.8.2 Оценка качества эндоскопических видеоизображений

Технические особенности современных эндоскопических систем таковы, что они имеют некоторые недостатки, в определенной степени осложняющие анализ полученных с их помощью колоноскопических изображений, причем это характерно как для работы врачей-экспертов, так и при использовании алгоритмов обработки информации методами ИИ. К ним можно отнести [32, 94, 124]:

 Оптические артефакты, связанные с физической конструкцией эндоскопа (ограничение глубины резкости, наличие пространственно-неоднородного размытия на полученных изображениях, отсутствие автофокусировки, локальные блики, вызванные коаксиальным расположением источника освещения и приемника).

- Динамические артефакты съемки (нелинейное размытие движением из-за турбулентного перемещения эндоскопа, интерлейсинговые артефакты, возникающие в системах с чересстрочной разверткой).
- Нестабильность фотометрических характеристик (вариации освещенности, яркости, контрастности вследствие неоднородного рассеяния на тканях и появлении локальных теней при изменении угла падения света).
- Низкое отношение сигнал/шум, вызванное шумом ПЗС/КМОП-матриц
 в условиях низкой освещенности; артефактов сжатия видеоизображений (блочности/размытия) при архивации данных.

Указанные недостатки приводят к необходимости дополнительной процедуры контроля качества сохраняемых изображений и видеопоследовательностей, которые в дальнейшем используются как входные данные для систем ИИ. В большинстве практических случаев после проведения обследования врачэндоскопист должен сохранить набор изображений согласно соответствующему регламенту в МИС. С большой вероятностью произвольно взятый из видеопотока колоноскопического исследования кадр не будет достаточно информативным. Следовательно, на данном этапе развития систем анализа изображений в эндоскопии актуален вопрос о контроле качества сохраненных кадров в МИС с использованием алгоритмических критериев оценивания качества изображений.

Оценка качества – важная часть процесса исследования при разработке систем обработки и анализа изображений [43, 148-149]. Проводить ее можно как путем усреднения экспертных оценок, так и путем разработки соответствующих алгоритмов. Кажется естественным использовать для указанных целей наиболее распространенные эталонные алгоритмы оценки качества изображений (PSNR, SSIM, UQI и др.) [150-153]. Однако во многих реальных задачах эталонное изображение, необходимое для их работы, отсутствует, поэтому приходится использовать класс неэталонных алгоритмов [148]. Они подразделятся на два

68

больших класса по признаку использования либо неиспользования априорной информации о виде искажения анализируемого изображения (рисунок 1.23).



Рисунок 1.23 – Классификация неэталонных алгоритмов оценки качества изображений

К неэталонным алгоритмам оценки качества изображений, учитывающим априорную информацию, относятся те, в которых определенным образом измеряется уровень соответствующего типа этого искажения. Примерами могут служить различные виды шумов, размытие, артефакты при сжатии по соответствующим стандартам (чаще всего – это стандарты сжатия статических изображений JPEG и JPEG2000) [43, 148, 154-158].

На современном этапе развития визуальных технологий значительное внимание разработчиков уделяется созданию и исследованию таких неэталонных алгоритмов оценки качества, работа которых уже не ограничивается только одним типом искажения изображений. К таким методам и алгоритмам, например, относятся метрики, базирующиеся на статистике естественных изображений (NSS – Natural Scene Statistics). Важно отметить, что указанная статистика может рассчитываться в пространственной области [36, 45], в области трансформант вейвлет-преобразования, а также дискретного косинусного преобразования [41, 159], которые традиционно широко используются в ЦОИ. Для неэталонной оценки качества также могут применяться различные методы как классического, так и глубокого МО [55-56]. Обучение нейросетевых алгоритмов традиционно осуществляется на предварительно собранной базе изображений с различными типами искажений и усредненными оценками экспертов MOS (Mean Opinion Score). В практических приложениях часто используют такие общедоступные наборы изображений, как LIVE [160], TID2013 [161] и др. При использовании методов классического MO для решения рассматриваемой задачи часто применяют рандомизированные деревья, машину опорных векторов, а также различные нейросетевые архитектуры. Использование методов глубокого MO в таких задачах также возможно, но требует сбора и оценивания большой базы изображений, релевантных решаемой практической задаче.

1.8.3 Детектирование объектов интереса нейронными сетями

Рассмотрим основные этапы разработки нейросетевой системы детекции объектов интереса в эндоскопии (рисунок 1.24). На этапе сбора и аннотирования данных накапливается набор записанных видеопоследовательностей, включающий в себя как нормальные участки ЖКТ, так и патологии различных типов и размеров. Патологии должны быть тщательно размечены и аннотированы квалифицированными эндоскопистами. Аннотации обычно включают в себя ограничивающие рамки (Bbox – bounding boxes) вокруг патологий или сегментацию (выделение пикселей, принадлежащих области интереса). Подробно процесс разметки описан в главе 2 настоящей работы.

Выбор архитектуры HC зависит от типа решаемой задачи (детектирование или сегментация). Для первой из них чаще всего используются следующие однопроходные архитектуры детекторов: SSD, YOLO, Retina Net [107-110] (разработана для решения проблемы дисбаланса классов). Для задач сегментации часто используют архитектуры: U-Net, Mask R-CNN, DeepLab [111]. Также, начиная с 2020 года, для решения и задач детектирования и задач сегментации стали использоваться архитектуры типа трансформеров, которые потенциально способны давать более высокие показатели по метрикам качества, но требуют

больше размеченных данных для обучения и большей вычислительной мощности на этапах обучения и применения (инференса).

 Разметка для задач детектирования или сегментации Валидация разметки 	
 Детектирование: SSD, RetinaNet, YOLO Сегментация: U-Net, R-CNN, DeepLab Трансформеры 	
 Изменение разрешения Нормализация значения пикселей Аугментация данных 	
 Обучающая выборка Валидационная выборка Тестовая выборка 	
 Выбор функции потерь Процедура оптимизации Регуляризация 	
 Алгоритмические метрики Сравнение с результатами эндоскопистов 	
 Режим реального времени Контроль качества 	

Рисунок 1.24 – Процесс реализации нейросетевого детектора

Предобработка данных состоит из изменения разрешения изображений для соответствия требованиям выбранной архитектуры сети, нормализации для повышения устойчивости процесса обучения, аугментации данных путем поворотов, масштабирований, отражений, сдвигов, изменений яркости и др.

Разделение данных предполагает их дробление на обучающую, валидационную и тестовую выборки, которое происходит обычно в пропорции 80:10:10 [56].

На этапе обучения НМ происходит выбор функции потерь, значение которой будет минимизироваться в итерационном цикле. Для задач сегментации типовым выбором является функция бинарной кросс-энтропии (Binary Cross-Entropy Loss), которая оценивает разницу между предсказанными и фактическими масками

сегментации. Также в задачах детектирования и сегментации активно используется метрика IoU (Intersection over Union). Для решения задачи дисбаланса классов используется метрика Focal Loss, которая увеличивает вес сложных примеров в обучающей выборке. Для обновления весов HC используется алгоритм оптимизации (например, Adam или SGD). На этом этапе часто возникает эффект переобучения, для предотвращения которого обычно используются методы регуляризации (например, L1, L2, Dropout) [54].

Этап оценки производительности заключается в расчете численных метрик, описанных далее в этой главе. Также важным моментом является сравнение результатов системы ИИ с аналогичными результатами работы опытных эндоскопистов.

Интеграция в МИС предполагает либо работу в режиме реального времени, предоставляя эндоскописту информацию о найденных патологиях во время процедуры, либо в режиме «после проведения процедуры», например, представляющая рассмотренную в главе 3 процедуру контроля качества колоноскопического исследования.

К современным задачам и проблемам в области нейросетевых методов детектирования объектов интереса в эндоскопии следует отнести следующие требования [55, 117, 118, 162, 163]:

- сокращение количества ложноположительных результатов (когда система ошибочно определяет нормальную ткань как патологию);
- разработка специализированных нейросетевых архитектур, специально разработанных для обнаружения конкретных патологий;
- применение архитектур на основе трансформеров для улучшения моделирования глобального контекста и зависимостей между пикселями;
- использование мультимодальной информации, такой как данные узкоспектральной или увеличительной эндоскопии, для улучшения точности обнаружения;

- возможность интерпретации результатов работы нейросетевого детектора.
1.8.4 Трекинг объектов интереса

Задача сопровождения объектов (трекинг) заключается в назначении соответствующих меток одним и тем же объектам в различных кадрах видеоданных. Классическим подходом к решению задач трекинга является алгоритм Лукаса-Канаде [39]. Он применим для функций любой размерности n и предполагает, что смещение между двумя кадрами мало и одинаково в пределах соседей некоторой точки p по определению. Пусть x – особенность первой функции F, необходимо найти такую точку x + h функции G, что разность окрестностей этих точек по мере минимальна.

Расстояние между окрестностями представляется следующим образом:

$$E = \sum \left[F(x+h) - G(x) \right]^2$$

здесь F(x), G(x) – рассматриваемые функции.

Используя приближенное разложение функции F(x) в ряд Тейлора с точностью до линейного члена разложения, получим:

$$F(x+h) = F(x) + h \frac{\partial F(x)}{\partial x},$$

где $\frac{\partial}{\partial x} = \left[\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n}\right]^T$ – обозначение градиента.

С использованием этого разложения найдем минимум расстояния *E*, приравняв к нулю соответствующую производную:

$$0 = \frac{\partial}{\partial x} E \approx \frac{\partial}{\partial x} \sum_{x} \left[F(x) + h \frac{\partial F(x)}{\partial x} - G(x) \right]^{2} =$$
$$= \sum_{x} 2 \frac{\partial F}{\partial x} \left[F(x) + h \frac{\partial F(x)}{\partial x} - G(x) \right].$$

Из этого выражения рассчитаем смещение h

$$h = \left[\sum_{x} \left(\frac{\partial F}{\partial x}\right)^{T} \left[G(x) - F(x)\right]\right] \left[\sum_{x} \left(\frac{\partial F}{\partial x}\right)^{T} \left(\frac{\partial F}{\partial x}\right)\right]^{-1}.$$

Метод Лукаса-Канаде требует предварительного сглаживания входного видеоизображения из-за его чувствительности к шумам и локального характера анализа. Дополнительные ограничения включают предположения алгоритма о неизменности яркости точек и локальной однородности движения, которые часто нарушаются в эндоскопических видеоданных.

Другим популярным методом решения такой задачи является подход, известный как фильтр Калмана [39]. Он решает задачу оценки состояния линейных динамических систем при наличии шумов измерений. Фильтр реализует оптимальный (в смысле минимума среднеквадратической ошибки) рекурсивный алгоритм обработки зашумленных данных, выдавая сглаженные и более точные оценки.

Пусть дана последовательность $\mathbf{y}_0, \mathbf{y}_1, ..., \mathbf{y}_{t-1}$ зашумленных наблюдений линейной динамической системы. Цель алгоритма – оценить внутреннее состояние системы в момент t, $\mathbf{x}_t = [\mathbf{x}_{1,t}, \mathbf{x}_{2,t}, ..., \mathbf{x}_{n,t}]^T$. Ошибку оценки следует минимизировать (т.е., заглянуть «на один шаг вперед»).

Обозначим $\hat{\mathbf{x}}_{t1|t2}$ оценку состояния \mathbf{x}_{t1} , основанную на перечисленной ниже информации, имеющейся в момент t_2 .

Обозначим $P_{t1|t2}$ матрицу дисперсий ошибки предсказания $x_{t1} - \hat{x}_{t1|t2}$. Цель – минимизировать $P_{t|t}$ в некотором математически строгом смысле. Решаем задачу предсказания в момент t, имея следующий набор данных:

- 1. Матрица перехода состояний **F**, которая применяется к («хорошо известному») предыдущему состоянию **x**_{t-1}.
- Управляющая матрица В, которая применяется к вектору управления u_t, если в данную систему вообще встроен механизм управления.

- Понимание характера шума системы wt (например, моделируемого многомерным нормальным распределением), выраженное в виде матрицы дисперсий Qt и математических ожиданий µi,t =E[wi,t] = 0 для i = 1, 2,..., n.
- 4. Вектор наблюдений \mathbf{y}_t состояния \mathbf{x}_t .
- 5. Матрица наблюдения **H** («Как наблюдать y_t »?).
- Понимание характера шума наблюдения vt (например, моделируемого многомерным нормальным распределением), выраженное в виде матрицы дисперсий Rt и математических ожиданий µi,t =E[vi,t] = 0 для i = 1, 2,..., n.

На первой фазе фильтра вычисляется предсказание матрицы дисперсий, зная матрицу перехода состояний **F** и управляющую матрицу **B**. Также применяется матрица дисперсий шума системы Q_t:

$$\begin{split} \hat{\mathbf{x}}_{t|t-1} &= \mathbf{F} \hat{\mathbf{x}}_{t-1|t-1} + \mathbf{B} \mathbf{u}_t, \\ \mathbf{P}_{t|t-1} &= \mathbf{F} \mathbf{P}_{t-1|t-1} \mathbf{F}^{\mathrm{T}} + \mathbf{Q}_t. \end{split}$$

На второй фазе вычисляется вектор невязки измерений $\tilde{\mathbf{z}}_t$ и матрица невязки дисперсий \mathbf{S}_t , для чего используется входящая в модель матрица наблюдения **H**. Также применяется матрица дисперсий шума наблюдения \mathbf{R}_t с целью улучшить матрицы шума:

$$\begin{split} \widetilde{\mathbf{z}}_{t} &= \mathbf{y}_{t} - \mathbf{H} \widehat{\mathbf{x}}_{t|t-1}, \\ \mathbf{S}_{t} &= \mathbf{H} \mathbf{P}_{t|t-1} \mathbf{H}^{\mathrm{T}} + \mathbf{R}_{t}. \end{split}$$

Чтобы вычислить скорректированный вектор оценки состояния (т.е. предсказание решения в момент t), рассматривается также шаг коррекции фильтра в момент t:

$$\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t \hat{\mathbf{z}}_t.$$

Матрица

$$\mathbf{K}_{\mathrm{t}} = \mathbf{P}_{\mathrm{t}|\mathrm{t}-1}\mathbf{H}^{\mathrm{T}}\mathbf{S}_{\mathrm{t}}^{-1}$$

минимизирует среднеквадратичную ошибку E[$(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t})^2$], что эквивалентно минимизации следа (суммы элементов на главной диагонали) матрицы **P**_{t|t}.

Матрица **K**_t называется оптимальной матрицей коэффициентов усиления Калмана и определяет линейный фильтр Калмана. Для него нужна также скорректированная оценка матрицы дисперсий шума системы

$$\mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_{t|t-1}$$

для подготовки к фазе экстраполяции в момент t + 1. Матрицу дисперсий **P**₀₁₀ необходимо инициализировать в начала процесса фильтрации.

На практике использование классических методов трекинга часто приводит к неудовлетворительным результатам. Это вызвано необходимостью учета множества факторов в эндоскопических видеоданных, таких как:

- изменение условий съемки (освещенность, шум камеры, артефакты изображения);
- деформация формы объектов;
- временное исчезновение объекта;
- ошибки детекции на предыдущих этапах обработки;
- требование работы в РРВ.

Для решения этих проблем разработаны современные алгоритмы трекинга [164-166], которые можно классифицировать следующим образом (таблица 1.2).

Способ	Примеры алгоритмов трекинга			
классификации				
1. По методу	1.1. Классические методы:			
обработки данных	 Фильтр Калмана – предсказывает положение объекта на 			
	основе предыдущих состояний, эффективен для линейных			
	движений.			
	 Партикл-фильтры – используют набор частиц для 			
	моделирования нелинейных движений и шумов.			
	 Корреляционные методы – сравнивают шаблоны объекта 			
	между кадрами.			
	1.2. Методы на основе глубокого МО:			
	– На основе СНС.			
	– На основе рекуррентных НС (учитывают временные			
	зависимости между кадрами).			
	 На основе трансформеров (моделируют глобальные 			
	контекстные связи).			
2. По стратегии	2.1. Трекинг с обнаружением. Объекты сначала детектируются в			
трекинга	каждом кадре, затем связываются между кадрами (например,			
	DeepSORT, FairMOT).			
	– Преимущество: устойчивость к перекрытиям.			
	 Недостаток: зависит от качества детектора. 			
	2.2. Трекинг без обнаружения			
	Использует начальное положение объекта и отслеживает его без			
	повторного детектирования (например, Siamese Networks).			
	 Преимущество: Высокая скорость. 			
	– Недостаток: Сложность при длительных перекрытиях.			
3. По режиму	3.1. Онлайн-трекинг (режим реального времени). Обрабатывает			
работы	кадры последовательно, без использования будущих данных.			
	3.2. Оффлайн-трекинг (постобработка).			
	Использует информацию из будущих кадров для уточнения			
	траекторий.			

Таблица 1.2 – Классификация алгоритмов трекинга

Продолжение таблицы 1.2

4. По способу	4.1. Методы с разделением и объединением.			
обработки	Разделяют траектории при перекрытии и объединяют их после			
перекрытий	восстановления видимости объекта.			
	4.2. Трекинг с учетом взаимодействий.			
	Моделируют социальные или физические взаимодействия между			
	объектами (например, для толпы людей).			
5. По уровню	5.1. Алгоритмы, требующие размеченных данных для обучения			
использования	(большинство современных методов на основе СНС			
данных	и трансформеров).			
	5.2. Алгоритмы, обучающиеся на неразмеченных данных,			
	используя внутренние свойства видеоданных (например,			
	CycleTrack).			

Использование алгоритмов трекинга совместно с рассмотренными выше алгоритмами детектирования области интереса позволяет улучшить характеристики системы обнаружения патологий, стабилизировать ее работу на последовательности видеокадров (рисунок 1.25) [167].



a)



б)

Рисунок 1.25 – Визуализация работы алгоритмов на последовательности кадров: без трекинга (а); с трекингом (б) На нем изображено сравнение работы алгоритма детектирования колоректального полипа без процедуры трекинга и с его использованием. Видно, что использование трекинга существенно увеличивает стабильность определения области интереса, что является важным для ее последующей классификации.

1.8.5 Сегментация объектов интереса

Задача сегментации объектов области патологий на эндоскопических изображениях заключается в точном выделении границ аномальных участков (полипов, опухолей, язв и др.) для последующей диагностики, хирургического планирования или анализа динамики заболевания. На практике ее решение требует учета множества факторов, таких как:

- Низкая контрастность и неоднородность тканей, затрудняющая различение патологий и здоровых областей.
- Артефакты изображения (блики, засветы, пузыри, кровь), маскирующие или имитирующие патологические структуры.
- Вариабельность форм и размеров аномалий даже в рамках одного типа заболеваний (например, полипы могут быть плоскими или выпуклыми).
- Динамические изменения во время процедуры: движение эндоскопа, перистальтика органов, деформация тканей инструментами.

Для решения этих проблем разработаны современные алгоритмы сегментации [111, 117, 125, 168-170], которые можно классифицировать следующим образом (таблица 1.3).

Таким образом, для решения задачи сегментации применяются гибридные подходы, комбинирующие методы глубокого обучения и классического компьютерного зрения. Ключевой проблемой остается недостаток размеченных медицинских данных, что может компенсироваться трансферным обучением и синтетическими данными.

-			
Таблица	 3 – Классифин 	сация алгоритмс	в сегментации

Способ	Примеры алгоритмов сегментации				
классификации					
1. По типу	1.1. Классические методы, основанные на анализе				
подхода	интенсивности пикселей и геометрических свойствах изображений.				
	– Пороговая обработка (разделение пикселей по				
	интенсивности).				
	– Методы на основе объединения пикселей в регионы по				
	схожести.				
	– Поиск границ через фильтры (Canny, Sobel).				
	1.2. Алгоритмы, основанные на классическом МО, требующие				
	ручного выделения признаков и классификации.				
	– Кластеризация (k-means).				
	– Активные контуры.				
	– Случайные леса, машины опорных векторов.				
	1.3. Современные архитектуры, автоматически извлекающие				
	признаки на основе глубокого МО.				
	– CHC (U-Net, Mask R-CNN).				
	 Трансформеры (Vision Transformer, Swin Transformer). 				
	– Генеративно-состязательные сети (GAN).				
2. По типу задачи	2.1. Семантическая сегментация. Каждому пикселю присваивается				
	класс объектов.				
	2.2. Инстанс-сегментация. Разделение объектов одного класса на				
	отдельные экземпляры.				
	2.3. Комбинация семантической и инстанс-сегментации.				
3. По разметке	3.1. Методы, требующие размеченных данных (масок).				
данных	3.2. Алгоритмы с использованием частичных меток (bounding				
	boxes, точки).				
	3.3. Алгоритмы, работающие без размеченных данных.				
	Кластеризация (K-means), автоэнкодеры.				
4. По типу	4.1. 2D-изображения.				
входных данных	4.2. 3D-данные (КТ, МРТ).				
	4.3. Видеоданные.				

Для анализа патологий на эндоскопических изображениях наиболее исследованными и популярными остаются следующие методы [111]: U-Net (стандарт в медицинской сегментации), Mask R-CNN (для сложных случаев с множественными объектами), классические методы (как основа для гибридных подходов). Эти алгоритмы обеспечивают баланс между точностью, скоростью и интерпретируемостью, что важно в клинической практике.

1.9 Метрики оценки точности работы нейросетевых алгоритмов

Для обработки видеоизображений в составе СППВР чаще всего применяются нейросетевые алгоритмы детектирования, сегментации и классификации объектов. Для воспроизводимости И сопоставимости результатов исследований используются стандартные метрики [54-57], такие как точность (Precision, P), полнота или чувствительность (Recall, R) и средняя точность (Average Precision). Еще одной часто используемой метрикой является точность (Accuracy) показывающая долю правильных предсказаний модели относительно всех прогнозов [56]. Она может использоваться для начальной оценки качества нейросетевой модели (например, после обучения) и оптимальна для задач классификации, когда классы распределены равномерно 50% (например, доброкачественных новообразований и 50% злокачественных):

$$P = \frac{TP}{TP + FP} , \qquad R = \frac{TP}{TP + FN} , \qquad AP = \frac{\sum_{r=1}^{N} P}{N}$$
$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}.$$

где TP – число истинно положительных результатов, FP – число ложноположительных результатов, TN – число верных отрицательных результатов, FN – число ложноотрицательных результатов, P_r – значение точности, соответствующее определенному значению полноты, а N – множество различных значений полноты. Дополнительно используется метрика среднего гармонического точности и полноты (F1):

$$F1 = 2\frac{PR}{P+R} = \frac{TP}{TP + \frac{1}{2}(FP + FN)}$$

Для проведения совокупной оценки точности работы модели необходимо рассчитать среднюю точность по всем *C* классам (mean Average Precision, *mAP*):

$$mAP = \frac{\sum_{i=1}^{C} P_i}{N}$$

Кроме того, для оценки влияния не только факта наличия или отсутствия предсказания, но и его качественных характеристик в зависимости от геометрических характеристик, найденной и искомой областей, используется такая метрика как отношение пересечения к объединению (Intersection over Union, *IoU*).

$$IoU = \frac{S_{int}}{S_{union}}$$

где S_{int} – площадь фигуры-пересечения экспертной разметки с рамкой фигуры из предсказания алгоритма, S_{union} – площадь фигуры-объединения экспертной разметки с рамкой фигуры из предсказания алгоритма. Ее геометрическая интерпретация для задачи сегментации полипа приведена на рисунке 1.25 [167].



Рисунок 1.25 – Визуализация метрики *IoU*: а) площадь пересечения, б) площадь объединения

Метрика *mAP* по показателю IoU, превышающего порог пересечения в 50% (или *mAP*@0,50) является основной метрикой для определения точности работы алгоритма детектирования объектов. Позднее ввели вспомогательную метрику mAP@[0,50...0,95], которая рассчитывается как усреднение показателей *mAP* для предсказания по *IoU* от 50% до 95% с шагом в 5%.

Основными метриками сравнения качества работы алгоритмов сегментации объектов на медицинских изображениях являются [111, 125]:

- *IoU* основная метрика для тестирования алгоритмов в рамках большого открытого датасета PASCAL Visual Object Classes (VOC). Чувствительна к форме границ.
- Коэффициент Дайса менее строгая к дисбалансу классов, чем *IoU*. Широко используется в медицинской визуализации.
- Hausdorff Distance отвечает за точность контуров. Устойчива к шумам.
- Чувствительность (Sensitivity) доля правильно обнаруженных пикселей объекта. Критична для медицинских задач, где пропуск объекта недопустим (медицина).

1.10 Выводы и постановка задач исследования

Проведенный анализ в области возможного применения методов ИИ и анализа видеоизображений применительно к построению СППВР для скрининговых (исследование толстой кишки) и клинических (исследование желудка) эндоскопических процедур позволяет сделать следующие основные выводы:

– Внедрение ИИ в скрининговые эндоскопические программы ЖКТ может стать существенным ускоряющим фактором для России в борьбе с онкологическими заболеваниями. Использование СППВР, основанных на методах и алгоритмах ИИ потенциально способно решить следующие проблемы: кадровый дефицит, контроль качества процедур, масштабирование скрининговой диагностики.

- Врачебные ошибки являются одной из основных проблем общественного здравоохранения и рассматриваются как угроза безопасности пациентов. Использование достижений в области ИТ может способствовать повышению качества медицинского обслуживания пациентов. Использование СППВР потенциально приводит к уменьшению врачебных ошибок при осмотре пациентов, диагностике, назначении лечения.
- Современные НМ показывают высокое качество обучения на больших массивах видеоданных. Однако проблема с использованием НС связана с необходимостью сбора и экспертной разметки огромных баз медицинских данных (порядка нескольких тысяч изображений на класс), которые часто недоступны, а во многих случаях и вовсе не существуют. В связи с этим возникает необходимость поиска подходов к построению робастных НМ на базах данных относительного малого размера.
- Эндоскопические видеоизображения обладают уникальными особенностями, связанными с условиями их получения. Соответствующие алгоритмы ЦОИ должны быть адаптированы к работе в реальном времени в условиях ограниченных аппаратных ресурсов и робастны к вариативности данных – изменению освещения, углов обзора, наличия размытия, бликов и иных артефактов. Также важным является сохранение диагностически значимых деталей слизистой при ЦОИ.
- К основным проблемам при разработке СППВР в эндоскопии ЖКТ относят: недостаток аннотированных видеоизображений, вариабельность данных в зависимости от оборудования, качество подготовки пациента, наличие маленьких и плоских новообразований, необходимость проведения клинических испытаний и сертификации.

Таким образом, проведенный анализ позволил сформулировать следующие основные задачи диссертации:

 Провести критический анализ современного состояния в области систем автоматического анализа видеоизображений эндоскопических исследований ЖКТ с помощью методов и алгоритмов ИИ.

- Разработать концепцию построения СППВР для клинических и скрининговых исследований в эндоскопии на основе методов глубокого МО в условиях практической ограниченности входного набора аннотированных видеоизображений.
- 3. Усовершенствовать архитектуру системы ИИ, работающей в качестве программного ядра для СППВР в эндоскопии ЖКТ.
- Создать физически и алгоритмически расширенные базы эндоскопических изображений и видеоданных с экспертной разметкой, аннотированием и врачебной валидацией областей патологий.
- Разработать алгоритмические решения для всех этапов обработки изображений – от предобработки и оценки качества до постобработки данных, учитывающие особенности эндоскопических видеопоследовательностей.
- Разработать метод контроля качества скринингового колоноскопического исследования и алгоритмы анализа видеопоследовательностей его реализующего.
- 7. Разработать алгоритмы детектирования и сегментации колоректальных полипов на видеоизображениях колоноскопических исследований.
- 8. Разработать алгоритмы детектирования и классификации патологий желудка на видеоизображениях гастроскопических исследований.
- Создать, провести тестирование, опытную эксплуатацию и внедрение специализированного ПО для контроля качества, визуализации и сопоставления результатов эндоскопической диагностики.
- 10. Создать, провести тестирование, опытную эксплуатацию и внедрение аппаратно-программных комплексов для захвата, разметки, обработки и хранения видеоизображений эндоскопических исследований.

ГЛАВА 2

КОНЦЕПЦИЯ ПОСТРОЕНИЯ СППВР В ЭНДОСКОПИИ НА ОСНОВЕ АЛГОРИТМОВ АНАЛИЗА ВИДЕОИЗОБРАЖЕНИЙ И МЕТОДОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

2.1 Особенности построения систем искусственного интеллекта для анализа медицинских изображений

Построение систем ИИ для анализа медицинских изображений связано с рядом особенностей, которые отличают эту задачу от аналогичных его применений во многих других областях. Медицинские изображения (рентген, МРТ, КТ, эндоскопические видеоизображения, гистологические снимки и др.) обладают высокой сложностью для анализа и интерпретации и требуют детального изучения экспертами [42, 124, 131]. Сбор больших аннотированных наборов медицинских изображений (датасетов) усложнен ввиду целого факторов: ряда конфиденциальности, присутствия редких патологий (несбалансированность данных) и высоких трудозатрат медицинского персонала на процессы сбора и разметки данных. Для решения проблемы сбора больших медицинских датасетов используются следующие подходы: трансферное обучение НС, аугментация данных, генерация синтетических обучающих данных, которая на данном этапе вызывает серьезные этические сомнения у ведущих медицинских специалистов [25-27].

Ошибки в диагностике могут напрямую влиять на жизнь и здоровье пациента, поэтому нейросетевые модели должны достигать высокой точности (95-99%). Ложноположительные или ложноотрицательные результаты работы алгоритма могут привести к крайне нежелательным последствиям (перегрузка внимания врача и пропуск опухоли) [27]. Алгоритмы и СППВР должны обладать свойством робастности, т.е. быть устойчивыми к вариации данных (например, нейросетевая модель, обученная на изображениях из Азии, эксплуатируется в Европе), замене производителя аппаратуры в рамках определенного класса (например, замена эндоскопа Olympus на Pentax), условиям съемки и другим факторам. Важно избегать смещений в данных, особенно при работе с разнородными популяциями пациентов.

Перспективным направлением развития является мультимодальная обработка медицинских данных, представленных в виде изображений, видео и аудиоданных, текстовых отчетов и результатов лабораторных анализов [43, 45, 162, 163, 171]. Мультимодальность для эндоскопических систем ИИ означает интеграцию результатов анализа видеоданных с результатами биопсии, лабораторными анализами или предыдущими исследованиями пациента.

Во время взаимодействия с СППВР врачи должны понимать логику работы нейросетевой модели, она не должна представлять собой «черный ящик». В этом смысле важным этапом развития такого интерфейса взаимодействия стало появление и развитие с 2025 года больших текстовых нейросетевых моделей (LLM) с механизмом рассуждений [119].

Построенные СППВР должны проходить клиническую валидацию, связанную с проведением тестирования на реальных репрезентативных клинических случаях и сравнением полученных результатов с мнением экспертов. После проведения испытаний проводится сертификация соответствующих систем (FDA в США, CE Mark в ЕС, Росздравнадзор в РФ) [66]. Разработан ряд ГОСТов в рассматриваемой области [65]. Важным вопросом также остается соблюдение норм при работе с персональными данными (HIPAA, GDPR, ФЗ № 152 «О персональных данных»). Для этого данные анонимизируются, собираются соответствующие соглашения с пациентами, в ряде случаев отдается приоритет локальной обработке данных без передачи в облачные сервисы.

Система ИИ должна иметь возможность интеграции в клиническую практику, например, быть встроенной в существующие сертифицированные процессы обмена данными: DICOM-изображения с метаданными, системами PACS (Picture Archiving and Communication System) и EHR (Electronic Health Records) [84, 85].

При разработке систем ИИ для анализа эндоскопических видеоизображений возникают специфичные особенности, связанные с высокой динамикой

отображаемого процесса разнообразием оборудования, В видеоданных, клиническими требованиями и техническими особенностями. Видеоданные содержат временную последовательность кадров, что дает возможность использования нейросетевых моделей, учитывающих временные зависимости. В этом плане особые надежды возлагаются на относительно недавно появившуюся для анализа изображений нейросетевую архитектуру – трансформер, которая успешно применяется для построения больших текстовых моделей [119].

Типичными артефактами эндоскопической видеосъемки являются: блики, пузыри, кровь, неравномерность освещения, чересстрочная развертка, дрожание камеры. Они существенно усложняют автоматический анализ видеопоследовательности. Еще одним затрудняющим фактором является то, что сами видеопоследовательности могут длится от 10 до 30 минут и более, а критически важные для постановки диагноза моменты занимать 1–5% от всего объема обследования.

Таким образом, особенностью создания и использования систем ИИ для анализа медицинских (эндоскопических) изображений является необходимость соблюдения баланса технической между сложностью, регуляторными требованиями и практической полезностью для врача. Успешная практическая реализация требует не только эффективных алгоритмов, но и глубокого понимания медицинского контекста, а также тесного сотрудничества с врачами. Ответственность за диагноз в итоге остается за врачом, а система ИИ представляет собой инструмент поддержки его решений, а не замену специалиста. Для развития таких систем также следует регламентировать юридические аспекты, т.е. четко определить, кто несет ответственность за возможную ошибку (разработчик, врач, больница). Эти и другие важные задачи предстоит решить в рамках мультидисциплинарных исследований в ближайшие годы и десятилетия.

2.2 Разработанная концепция построения СППВР в эндоскопии

Предлагаемая структура построения СППВР в эндоскопии [172] на основе алгоритмов анализа видеоизображений и методов глубокого машинного обучения

88

в условиях практической ограниченности входного набора аннотированных (размеченных) видеоизображений приведена на рисунке 2.1.



Рисунок 2.1 – Концепция построения СППВР в эндоскопии

При ее реализации аппаратная эндоскопическая стойка соединяется с высокопроизводительным сервером, оснащенном картой видеозахвата. На этом сервере работает обученная под конкретные эндоскопические задачи глубокая НС. Врач, производящий визуальный анализ видеоизображений время BO исследования, взаимодействует с первой системой, работающей в режиме реального времени (СППВР-РРВ). Поддержка в принятии решения для врача выводится на мониторе в виде обведенных (раскрашенных) областей, либо с помощью звуковых подсказок, фокусирующих его внимание. В работе СППВР-РРВ особенно важным является минимизировать ложноположительные ошибки, перегружающие внимания врача во время исследования. В случае их наличия, как показывает практика, врачи отказываются использовать СППВР-РРВ в работе и активно участвовать в сборе/разметке материала для дообучения систем ИИ [173].

В процессе эндоскопического исследования сохраняется полная видеопоследовательность, а также набор статических изображений (кадров) по выбору врача, отражающих значимые области интереса. Врачи дополняют эти визуальные данные текстовой аннотацией (в том числе получаемой путем автоматического преобразования речи в текст) и лабораторными данными, получаемыми из гистологических исследований. Данные загружаются В используемую медико-информационную систему и взаимодействуют со второй системой в режиме постобработки данных (СППВР-РПД). Она работает в режиме офлайн и позволяет аннотировать проведенное исследование, проводить контроль качества медицинской процедуры в соответствии с ее регламентом, отбирать изображения с соответствующими областями интереса и др.

Нейросетевая система контроля качества работает на основе регламента соответствующей процедуры, что особенно актуально для проведения массовых скрининговых эндоскопических исследований. Работа системы СППВР-РПД также позволяет дополнительно выявить видеозаписи исследований с высокой вероятностью наличия патологий, в том числе онкологических заболеваний ЖКТ на ранней стадии. В этом случае такая видеозапись отсылается в медицинский центр экспертного уровня для дополнительного анализа и, в случае необходимости, проводится еще одно клиническое исследование пациента.

В процессе работы СППВР накапливают данные в локальном или защищенном облачном хранилище данных. С определенной очередностью проводится дообучение их нейросетевых моделей на основе дополнительно накопленных верифицированных данных. Процесс дообучения представляет собой корректировку коэффициентов HC, что улучшает параметры работы систем с учетом накопленного опыта.

Рассмотрим более подробно структуру системы ИИ, составляющую основу для систем СППВР-РРВ и СППВР-РПД [174-177]. Важными этапами создания эндоскопической ИИ-системы являются подготовка и разметка визуальных данных, процесс обучения и тестирования HC, а также этап применения (инференса) нейросетевой модели (рисунок 2.2). При разработке нейросетевой модели критически важно учитывать, что на вход подается видеопоследовательность, где область патологии обычно занимает лишь 1-5% от ее общей длительности.



Рисунок 2.2 – Этапы построения системы ИИ для СППВР в эндоскопии

По сравнению с другими системами ИИ, применяемыми в различных областях, следует более подробно рассмотреть вопросы подготовки и разметки эндоскопических видеоданных, а также проблемы построения ИИ-инфраструктуры, включающей модули дообучения НС, приведенные на рисунке 2.1. Этим вопросам посвящены следующие два пункта настоящей главы. Вопросы разработки конкретных алгоритмов детектирования, сегментации и классификации областей интереса с учетом специфики эндоскопических задач в области ЖКТ рассматриваются в главах 3-6 настоящей работы.

2.3 Разработанная методология подготовки и разметки эндоскопических видеоданных

Для того чтобы снизить огромные трудозатраты эндоскопистов экспертного класса и специалистов в области разметки изображений предложена оптимизированная методика разметки областей интереса для последующего обучения и/или тестирования нейросетевых моделей. Рассмотрим ее составляющие более подробно.

2.3.1 Описание предложенных процессов физической разметки и физической аугментации

На этапе подготовки данных с учетом специфики задачи применены два новшества. Первое состоит в использовании устройства для «физической разметки данных», выполненного в форм-факторе педалей [174]. Такой выбор обусловлен занятостью рук врача-эндоскописта во время проведения процедуры диагностики. Во время визуального анализа видеопотока, выполняемого врачом в режиме реального времени, в случае нахождения патологии он нажимает на педаль, что приводит к появлению отметки в сохраняемом видеофайле (рисунок 2.3). Вторая отметка фиксируется в момент отпускания педали. Это позволяет физически размечать видеопоследовательность непосредственно в процессе осмотра пациента и не требует от врачей значительных дополнительных усилий.

Альтернативным способом являлось сохранение видеопоследовательности и ее дальнейший просмотр с дополнительной текстовой аннотацией. Однако, как показывает практический опыт, возвращение врача к просмотру и анализу видеоданных во второй раз за короткое время является психологически некомфортным. Разметка с помощью педалей апробирована в эндоскопическом отделении Пироговского центра г. Москва (ПРГВ) в 2022 году во время опытной эксплуатации системы детектирования колоректальных полипов для ускорения создания обучающего набора изображений. Подробней эта задача будет рассмотрена в главе 4 настоящей работы.



Рисунок 2.3 – Процесс физической разметки с помощью педали

Второе новшество состоит в блоке «физической аугментации», приведенном на рисунке 2.2, и поясняется более подробно на рисунке 2.4. Технически это достигается управляемого перемещения фокусировки за счет И оборудования процедуры, обеспечивая видеоэндоскопического во время мультимодальную съемку (разные ракурсы, увеличения, временные точки) в рамках осмотра конкретной области патологии. Это позволяет более эффективно наполнять соответствующие базы видеоизображений [176, 177].



Рисунок 2.4 – Процесс физической аугментации

Процесс физической аугментации позволяет получать дополнительные области видеоизображения интереса физическим путем, а не путем алгоритмической аугментации, который широко применяется для обучения современных архитектур НС. Особенно это актуально для часто возникающих ситуаций, где размноженные алгоритмически изображения могут в итоговом обучающем датасете составлять доминирующую часть. В отличие от алгоритмических методов, физическая аугментация гарантирует получение несинтезированных естественными вариациями, реальных, данных с ЧТО критически важно для обучения робастных нейросетевых моделей. Процесс физической аугментации использован для фиксации области купола слепой кишки в Ярославской областной клинической онкологической больнице (ЯОКОБ) в 2024 году, что привело к расширению соответствующего датасета. Подробней эта задача будет рассмотрена в главе 3 настоящей работы.

2.3.2 Разработанная методика полуавтоматической разметки изображений

Для обучения и тестирования нейросетевых моделей требуется создание больших баз аннотированных изображений. В рамках данной работы они создаются из видеопоследовательностей эндоскопических исследований. Самым затратным по времени этапом при их создании является разметка, которую в идеале должны выполнять врачи экспертного уровня, что является практически недостижимой целью. Чтобы значительно сократить такой тип трудозатрат, полуавтоматической предложена методика разметки эндоскопических изображений с целью выделения на них соответствующих областей интереса. Полуавтоматический способ разметки известный также, как Human-in-the-Loop (HITL) представляет собой гибридный подход к аннотированию данных, в котором предобученные НС взаимодействуют с экспертом в цикле обработки данных. Цель такого подхода состоит в совмещении скорости машинной обработки с точностью человеческого контроля, минимизируя при этом ручной труд без ущерба для качества. Нейросетевая модель может выбирать, какие данные требуют приоритета для проверки экспертом (например, кадры с низкой уверенностью предсказаний). Таким образом, эксперты тратят время только на сложные случаи (спорные диагнозы), а не на рутинную многочасовую разметку [178-181].

Структурные схемы для классической процедуры разметки эндоскопических видеоданных и полуавтоматической разметки, предложенной в диссертации [167] для разметки колоректальных полипов, приведены на рисунке 2.5a и рисунке 2.5б, соответственно.

Роли «разметчик» и «ведущий разметчик» отличаются опытом процесса разметки изображений эндоскопического класса, а также скоростью работы в соответствующем программном обеспечении для разметки.

Как показывает практический опыт разметки больших объемов эндоскопических видеоданных, проведенный с 2019 по 2024 год в центре ИИ ЯрГУ им. П.Г. Демидова и партнёрских малых инновационных предприятиях [174, 176], час работы обычного разметчика после уплаты всех налогов стоит в г. Ярославле в 2025 году порядка 250 рублей, час работы ведущего разметчика порядка 500 рублей (x2), час работы врача экспертного класса для валидации разметки порядка 1500 рублей (x6). Типовое соотношение между количеством ведущих и обычных разметчиков составляет 1:10, но для больших медицинских датасетов оно может составлять 1:5 ввиду трудоемкости разметки.



Рисунок 2.5 – Процедура ручной разметки (а); полуавтоматической разметки (б)

Дальнейший процесс оптимизации процедуры разметки заключается в минимизации нормированных трудозатрат, т.е. снижению занятости для позиций «ведущий разметчик» и «врач». Предлагаемая методика приведена на рисунке 2.6.

Видеопоток поступает на этап разметки уже с помеченными маркерами интервалов нахождения областей интереса, соответствующими времени нажатия педали. Далее происходит машинная разметка указанных интервалов с помощью предобученной HC. За ней следует автоматический этап выбора информативных данных, подразумевающий контроль качества изображений (например, удаление размытых и изображений с бликами), проверка на согласованность данных внутри определенного трека. Далее разметчик просматривает полученные последовательности кадров с автоматической разметкой и удаляет ошибочные предсказания алгоритма. Результаты работы валидируют врач и ведущий разметчик, формируя финальный датасет [172].



Рисунок 2.6 – Разработанная методика разметки

Представленный подход допускает поэтапное применение для аннотирования обширных массивов видеоданных. На каждом этапе можно размечать ограниченную порцию данных с последующим обновлением нейросетевой модели, проводящей первичную разметку, сокращая число ошибок, а, следовательно, и ручной труд на следующих итерациях.

Сравнение разных способов разметки приведено в таблице 2.1. Здесь проставлены экспертные оценки по пятибальной шкале, полученные путем анкетирования группы специалистов компании «Точка зрения» [182],

специализирующейся на создании систем машинного зрения с использованием методов ИИ. Критерий «гибкость» означает возможность адаптивной перестройки от увеличения автоматизации в цикле до полностью ручной проработки данных. Балл «2» соответствует оценке «неудовлетворительно», а балл «5» – «отлично». Результаты показывают преимущество полуавтоматических методов разметки, предоставляющих высокую точность при приемлемом объеме трудозатрат.

	Ручная	Полная	Полуавтоматическая	Предлагаемая
Критерий	разметка	автоматизация	разметка	методика
	(рисунок 2.5а)	(без человека)	(рисунок 2.5б)	(рисунок 2.6)
Точность	5	2	5	5
Скорость	2	5	4	4
Затраты	2	5	3	4
Гибкость	2	2	3	4
Итого	11	14	15	17

Таблица 2.1 – Экспертная оценка методов разметки

Таким образом, установлено, что полуавтоматические методы разметки представляют оптимальный баланс между автоматизацией и контролем человека. Для медицинских (и в особенности для эндоскопических) данных, где цена ошибки крайне высока, такой подход является эффективным: НС берут на себя рутинные операции, а врачи-эксперты фокусируются на сложных кейсах [173, 174].

Итоговая структурная схема, иллюстрирующая разработанную методологию подготовки и разметки эндоскопических видеоданных, приведена на рисунке 2.7. Она сочетает в себе физическую разметку с помощью педали и использование предложенной оптимизированной методики полуавтоматической разметки видеоданных.

Согласно экспериментальным оценкам, использование предлагаемого подхода позволяет снизить нормированные трудозатраты на создание больших баз размеченных эндоскопических видеоданных при сохранении точности разметки на

35-45% по сравнению с классической ручной методикой, приведенной на рисунке 2.5а [174, 177].



Рисунок 2.7 – Структурная схема процесса разметки видеоизображений с использованием педали и полуавтоматической процедуры

2.3.3 Используемое программное обеспечение для разметки изображений

Программное обеспечение для разметки эндоскопических видеоизображений играет важную роль в развитии соответствующих систем ИИ. Оно позволяет структурировать «сырые» видеоданные в аннотированные датасеты, необходимые для обучения и валидации нейросетевых моделей. Без точной разметки патологий (полипы, язвы, опухоли) невозможно создать алгоритмы поддержки принятия решений, способные обнаруживать аномалии с клинически значимой точностью. Без использования специализированного ПО процесс ручной разметки кадров врачами-экспертами становится чрезвычайно трудоемким и непрактичным. Специализированные программные инструменты ускоряют аннотирование (например, с помощью интерполяции меток между ключевыми кадрами), обеспечивают единые стандарты описания и гарантируют согласованность данных для многопользовательской работы, что в итоге напрямую влияет на качество, надежность и скорость внедрения ИИ-решений в реальную клиническую практику.

Для разметки эндоскопических изображений использован программный продукт LabelMe [183]. Это графический инструмент для аннотирования изображений с открытым исходным кодом, доступным на ресурсе GitHub. Интерфейс программы LabelMe на примере задачи разметки эндоскопического изображения желудка изображен на рисунке 2.8.



Рисунок 2.8 – Интерфейс программы для разметки изображений LabelMe

Выбор в пользу программы LabelMe на текущем этапе эволюции разработки программных решений для разметки (2024 г.) обусловлен рядом факторов: бесплатный доступ ко всем функциям; удобный и понятный интерфейс (низкое

обучения новых разметчиков); время возможность структурированного отображения баз изображений; широкий выбор инструментов для аннотирования; возможность внедрения пользовательских надстроек в интерфейс. Среди других программных продуктов для разметки медицинских изображений следует отметить: Label Studio, CVAT, VGG Image Annotator (VIA), RectLabel, Supervisely, Roboflow, Make Sense, LabelImg и др. [181]. Следует отметить, что многие из них предполагают загрузку изображений на сторону сервера, что вызывает дополнительные вопросы с точки зрения хранения и обмена персональными данными. В программе LabelMe разметка и хранение изображений производится на компьютере разметчика.

Последовательность действий при разметке эндоскопических видеоизображений в LabelMe следующая:

- 1. Открытие изображения. Пользователем выбирается нужное изображение или папка с изображениями для последующей аннотации.
- Процедура разметки. Изображение просматривается в программе и выделяются объекты интереса ограничивающей рамкой (для задач детектирования) или полигоном (для задач сегментации).
- Добавление метаданных. Пользователь добавляет метаданные к выделенному объекту в появившимся окне, такие, например, как наименование класса объекта (показано на рисунке 2.9).
- 4. Экспорт данных. В папке, где находится размеченное изображение, создается файл формата JSON с тем же названием.

Полученный JSON-файл хранит в себе информацию о размеченных на снимках полигонах в виде координат каждой выделенной области, наименования класса, вида полигона. Кроме того, в файле хранится информация о версии программы, размерах изображения, а также другие дополнительные данные. На рисунке 2.10 изображены примеры разметки для задач детектирования и сегментации патологий желудка.



Рисунок 2.9 – Добавление метаданных к выделенному объекту



Рисунок 2.10 – Разметка изображений для задач детектирования (а) и сегментации (б)

Помимо разметки может использоваться и более детальная процедура аннотирования эндоскопических изображений (разметка + текстовое описание объекта интереса), как показано на рисунке 2.11. Это нетривиальная задача, которая требует экспертного уровня квалификации специалиста. Для того чтобы разметку могли осуществлять специалисты, не имеющие соответствующей квалификации, используется следующий подход: изображения для базы данных выбираются из видеопоследовательностей, каждая из которых состояла из определенного количества сцен. Врачам-экспертам предоставляется по несколько кадров из каждой сцены, для которых они осуществляют разметку и при необходимости текстовое описание. Далее файл с таблицей, в которой доступны сведения о названии проекта, патологии в кадре и местоположении рамки для данного изображения предоставлялся разметчикам, осуществляющим аннотирование всех интересующих классов видеопоследовательности.

08.07.24 earlycancer Кадр с оптическим увеличением. Образование, приподнятый участок с отчетливой границей, с нерегулярным микрососудистым рисунком, разрушенным поверхностным рисунком. Ранний рак в зоне пищеводножелудочного перехода. Проксимальнее неоплазии определяется плоский эпителий пищевода бледно-розового цвета, дистальнее более красный эпителий желудка.



Рисунок 2.11 – Пример подробного описания кадра, выполненного врачом

2.3.4 Методология для контроля качества разметки эндоскопических изображений

Контроль качества разметки эндоскопических видеоизображений критически важен для создания СППВР на основе методов ИИ. Даже большой объем плохо размеченных видеоданных не поможет нейросетевой модели научиться корректно решать задачи эндоскопической диагностики (качество данных в данном случае важнее, чем их количество). Поскольку медицинские ИИ-системы должны проходить обязательную сертификацию, то на этом этапе могут потребоваться доказательства качества обучающих данных и отчеты о валидации разметки врачами-экспертами. Кроме того, контроль на этапе разметки снижает затраты на дообучение системы.

Ошибки в разметке (например, неточное выделение патологии) приводят к обучению нейросетевой модели на некорректных данных. Это может вызвать появление ошибок 1-го и 2-го рода и, следовательно, риски для пациента при использовании системы в клинической практике. Например, если модель обучится на видеоданных, где кровь ошибочно была помечена как патология, она будет давать неверные предупреждения во время исследования, что будет подрывать доверие врачей к конкретной системе и в целом к использованию методов ИИ в медицинской диагностике.

Другой типичной проблемой является согласованность экспертных данных. Например, врачи могут по-разному интерпретировать одни и те же визуальные структуры. В задачах классификации часто используют функционал качества Каппа Коэна (Cohen's Kappa) для проверки согласованности двух экспертов и Каппа Флейса (Fleiss' Kappa) – для трех и более экспертов [53].

Функционал качества Каппа Коэна (k) – статистическая метрика, используемая для оценки согласованности между двумя разметчиками, каждый из которых классифицирует N объектов по C взаимоисключающим категориям. Она учитывает не только фактическое согласие, но и вероятность случайного совпадения, что делает её более надежной, чем простая точность (Accuracy) в задачах с несбалансированными классами (редкие патологии в эндоскопических видеоданных). Формула для расчета метрики выглядит следующим образом:

$$k=\frac{p_0-p_e}{1-p_e},$$

где p_o – относительное наблюдаемое согласие между разметчиками, а p_e – вероятность случайного согласия, рассчитанная на основе наблюдаемых данных для определения вероятности случайного выбора каждой категории

каждым наблюдателем. Значение k варьируется от -1 (полное несогласие) до +1 (идеальное согласие).

Схема «слепого» эксперимента по контролю качества разметки приведена на рисунке 2.12 [184]. Врач-эксперт с помощью ведущего разметчика подготавливает эталонную выборку. Показателем качества является согласованность разметки между конечным разметчиком и экспертом. В проведенных «слепых» экспериментах по оценке согласованности мнений разных пар врач-разметчик показатель *k* превышал уровень 0,7, что означало хорошую согласованность между экспертными оценками.



Рисунок 2.12 – «Слепая» проверка качества разметки

Также в процессе разметки эндоскопических данных использовался метод «ловушек» (honeypots) [184]. Они представляют собой заранее подготовленные образцы данных с нужным распределением, которые подмешиваются во входные данные разметчиков (рисунок 2.13). Сравнивая ответы разметчиков на примерахловушках с известными правильными метками, можно объективно оценить их точность. Такой подход выявляет недобросовестных или невнимательных

разметчиков. Он особенно полезен для больших баз изображений, где изменения в технической постановке задачи происходят не слишком часто.



Рисунок 2.13 – Проверка качества разметки с помощью метода ловушек

Также использовались перекрестные проверки (кросс-валидация), при которых аннотаторы проверяли разметку друг друга. В спорных ситуациях изображения размечались несколькими людьми, и итоговая метка определялась большинством голосов. Итоговую валидацию размеченных данных проводили врачи-эндоскописты. После этого анализировались изображения, которые были некорректно размечены по итогам «слепой эталонной проверки».

Для поддержания высокого качества разметки необходим системный подход: регулярный контроль ошибок, их анализ командой при участии ведущего разметчика и коллективное обсуждение проблем. Это не только повышает согласованность в работе, но и косвенно улучшает качество и объем обучающих данных за счёт фокусировки на «слабых местах» – примерах с повышенной частотой ошибок.

Если разбор противоречивых случаев (где мнения участников расходятся) помогает выявить системные ошибки или пробелы в понимании задач, то работа с

некачественными аннотациями изображений предполагает индивидуальную обратную связь: ведущий разметчик детально разбирает ошибки с конкретным исполнителем, чтобы скорректировать подход. Такой баланс между командным анализом и персональной работой позволяет устранять как общие, так и частные проблемы в процессе разметки.

В еще более сложных случаях, например, при попиксельной разметке области патологии желудка использовалась разработанная методология разметки, предполагающая работу консилиума из 3, 5 или 7 врачей (рисунок 2.14) [173, 174].



Рисунок 2.14 – Попиксельная разметка изображений консилиумом врачей

Сохраненные в таком процессе веса для каждого пикселя из размеченной области могут использоваться для обучения нейросетевых моделей. Такой способ является наиболее точным, но и наиболее трудоемким.

2.4 Построение ИИ-инфраструктуры для задач дообучения эндоскопических СППВР

Важным процессом для создания СППВР остаются вопросы хранения архивных видеоданных и реализации процессов дообучения нейросетевых моделей. В случае эндоскопии создание и поддержка соответствующей ИИ-инфраструктуры является особенно дорогостоящей ввиду большого объема видеоданных в высоком разрешении.

Для реализации методов глубокого машинного обучения, как правило, требуются сравнительно большие объемы данных, что может усложнять процесс обучения нейросетевой модели для отдельных пользователей. Для решения этой проблемы данные от многих пользователей можно агрегировать в дата-центрах, как это делается в модели централизованного обучения (рисунок 2.15). Однако это негативно сказывается на вопросах безопасности персональных данных пользователей. Для решения этой проблемы используются методы федеративного обучения [185-190].

Федеративное обучение – метод создания ИИ-инфраструктуры, при котором данные остаются на стороне владельца и не передаются внешним подрядчикам. Метод соответствует ФЗ № 152 «О персональных данных», его рекомендуется применять для работы с конфиденциальной информацией, в частности для медицинских приложений. Федеративное обучение работает следующим образом. Модель тренируется на локальном сервере владельца данных (клиенте), который является частью распределенной системы обучения. Данные остаются у владельца, а на центральный федеративный сервер отправляются только обновленные параметры модели, где они объединяются с параметрами, полученными от других владельцев (рисунок 2.16).



Рисунок 2.15 – Модель централизованного дообучения

108


Рисунок 2.16 – Модель федеративного дообучения

При таком подходе данные не обязательно должны полностью совпадать по структуре. В разных медицинских организациях может накапливаться разнородная информация по пациентам. С помощью специальных алгоритмов технология федеративного обучения позволяет обновлять нейросетевые модели даже на таких разнородных данных [185, 186].

Сценарии применения технологии федеративного обучения в медицине могут быть Во-первых, разными. технология хорошо подходит ДЛЯ ИИ-проектов, когда модель нужно обучить на конфиденциальных данных, которые еще и сильно распределены географически. Например, ЭТО удобно ДЛЯ международных медицинских научных проектов в области ИИ, когда данные нельзя передать В другую страну из-за особенностей национального законодательства или других требований.

Во-вторых, федеративное обучение подходит тогда, когда нужно средство для безопасного предоставления доступа к своим данным многим участникам проекта. Например, если отдельный регион, ведомство или организация накопили большой массив медицинских данных, но не могут предоставить к нему доступ, так как данные очень чувствительные. Такой подход, например, выбрал Сеченовский университет для предоставления доступа разработчикам из своего биобанка [187]. С помощью облачных сервисов специалисты организовали платформу

медицинских данных, которую можно использовать для разработки новых технологий машинного обучения – например, для изготовления лекарств и исследования новых методик лечения.

Анализ моделей централизованного и федеративного дообучения в сочетании с реальным состоянием ИТ и ИИ-инфраструктуры в эндоскопических отделениях современных российских медицинских центров привел к разработке гибридной модели дообучения (рисунок 2.17) [173, 174].



Рисунок 2.17 – Предлагаемая гибридная модель дообучения

Предложенная гибридная модель сочетает сильные стороны централизованного и федеративного обучения, адаптируя их под специфику медицинских данных и законодательные ограничения. Вот ключевые преимущества такого подхода:

- Соблюдение законодательства РФ (ФЗ № 152). Передача данных в центральный кластер в рамках российской юрисдикции позволяет соблюдать требования закона о персональных данных. Если кластеризация происходит внутри РФ, риски несанкционированного доступа или утечек минимизируются.
- Экономическая эффективность для малых эндоскопических отделений.
 Небольшие клиники и отделения при использовании гибридной модели не

обязаны закупать GPU-серверы – их данные обрабатываются на центральном кластере. Это снижает капитальные затраты и упрощает внедрение ИИ-системы. Крупные медицинские центры с собственными вычислительными мощностями могут участвовать в федеративном обучении напрямую, не перегружая дополнительный кластер.

- Масштабируемость и гибкость. Модель позволяет вовлекать в обучение даже удаленные и малоресурсные медицинские учреждения, что критично для редких патологий или региональных особенностей заболеваний ЖКТ. Региональный кластер аккумулирует информацию из множества источников, включая малые отделения, что обогащает датасет и снижает риски переобучения.
- Возможность международного сотрудничества. Центральный сервер может выступать посредником между российскими данными и зарубежными федеративными сетями. Например, агрегированные обновления от российских отделений могут объединяться с иностранными, не нарушая локальных законов.
- Баланс между конфиденциальностью и эффективностью. Федеративное обучение сохраняет анонимность данных, но гибридный подход дополняет его возможностью глубокой обработки сложных случаев на региональном кластере (например, аномалий в видео).
- Упрощение контроля и стандартизации. Кластер позволяет проводить предобработку данных, валидацию разметки и выявление артефактов перед их использованием в обучении. Все участники системы работают по согласованным алгоритмам разметки и обучения, что критично для медицинских приложений, где ошибки недопустимы.

Таким образом, предложенная гибридная модель решает ряд важных проблем: для малых эндоскопических отделений – экономит ресурсы и предоставляет доступ к системам ИИ; для качества нейросетевой модели – объединяются разнородные данные и методы обучения; гарантируется соблюдение законодательства в части защиты персональных данных. Это делает такой подход оптимальным для медицинских систем с высокими требованиями к конфиденциальности и разнородной инфраструктурой участников.

2.5 Краткие выводы

Основные выводы и результаты, полученные в данной главе, можно сформулировать следующим образом:

- 1. Разработана СППВР ЖКТ. концепция построения В эндоскопии Ее использование в клинической и скрининговой диагностике повышает достоверность эндоскопических исследований за счет автоматического обнаружения патологий, снижения влияния человеческого фактора, стандартизации формы представления результатов, контроля качества исследования.
- 2. Предложены усовершенствования системы ИИ, работающей в качестве программного ядра для СППВР в эндоскопии ЖКТ в условиях практической ограниченности входного набора аннотированных видеоизображений. Особенностью системы являются дополнительные программно-аппаратные модули, позволяющие получать результат по точности обнаружения и локализации патологий на видеоданных, сопоставимый со средним результатом, показываемым врачом-эндоскопистом во время диагностики.
- 3. Разработаны и апробированы методики аппаратной аугментации, физической разметки и полуавтоматической разметки базы эндоскопических видеоданных для обучения нейросетевых моделей. Их использование снижает нормированные трудозатраты при создании больших размеченных баз видеоизображений на 35-45%.
- 4. Предложена и апробирована методология контроля качества процесса разметки эндоскопических видеоданных. В рамках нее предложен точный метод разметки, предполагающий работу консилиума из 3, 5 или 7 врачей, который может использоваться при попиксельной разметке областей патологии ЖКТ в задачах сегментации.

5. Разработана гибридная модель процесса дообучения эндоскопических СППВР, сочетающая сильные стороны централизованного и федеративного обучения, адаптируя их под специфику эндоскопических видеоданных, законодательные ограничения и разнородную ИТ-инфраструктуру медицинских центров.

ГЛАВА З

РАЗРАБОТКА И ИССЛЕДОВАНИЕ МЕТОДА КОНТРОЛЯ КАЧЕСТВА КОЛОНОСКОПИЧЕСКОГО ИССЛЕДОВАНИЯ

Рассмотрим задачу контроля качества колоноскопического исследования. Для ее решения нужно разработать алгоритм детектирования купола слепой кишки (КСК) на видеопоследовательности колоноскопического исследования, а также алгоритм определения некачественных изображений (размытых, с артефактами сжатия, неинформативных), сохраненных в медико-информационной системе в рамках проведенного скринингового колоноскопического исследования.

3.1 Актуальность задачи обнаружения купола слепой кишки

Область интереса, рассмотренная в данном разделе, – проксимальная часть слепой кишки (лат. Caecum – начальный отдел толстой кишки), которая находится в начале толстого кишечника (в нижней правой области живота). Ее можно представить как «тупик» толстой кишки, куда впадает тонкий кишечник через специальный илеоцекальный клапан. Другими словами, это своеобразная «финишная черта» при выполнении процедуры колоноскопии: если врач провел эндоскоп в эту анатомическую область, значит, он полностью осмотрел толстую кишку (рисунок 3.1) [191-194].

Слепая кишка имеет мешкообразную форму длиной от 4 до 8 см и самый большой из всех кишок человека поперечный размер, достигающий от 6 до 11 см. От медиальной поверхности слепой кишки отходит червеобразный отросток (аппендикс), имеющий внешний диаметром от 0,5 до 1 см и длину от 2 до 20 см. Купол слепой кишки визуализируется как мешкообразное расширение в правой подвздошной ямке с характерными утолщенными складками, сходящимися в одной точке [32, 33]. При проведении колоноскопии достижение КСК является ключевой контрольной точкой, которое подтверждает, что обследование

проведено в полном объеме. Если не осмотреть эту область, то можно пропустить полипы, опухоли или воспаления в этой зоне.



Рисунок 3.1 – Строение толстого кишечника человека

Обнаружение области КСК не является тривиальной задачей ввиду следующих основных причин: кишки толстого кишечника имеют изгибы и складки, внешне напоминающие КСК; часто осмотру мешает содержимое кишечника (остатки кала, слизь); существуют вариации в расположении купола в анатомии разных людей (ближе-дальше). Степень достижения купола слепой кишки (Cecal Intubation Rate, CIR) представляет собой ключевой показатель качества колоноскопического исследования, отражающий долю процедур, при которых врач-эндоскопист успешно достигает КСК. Значение показателя CIR характеризует полноту обследования и снижает риск пропущенных патологий (полипы, рак и др.) [191, 195].

Согласно международным стандартам ASGE и ESGE рекомендуемый минимум в части достижения показателя CIR для скрининговых колоноскопических исследований составляет ≥ 90%. Средние значения в реальных исследованиях составляют 92-95% для колоноскопий в крупных клиниках с опытными эндоскопистами и 80-90% для скрининговой практики [196, 197]. К факторам, снижающим значение CIR, обычно относят: ожирение (индекс массы

тела >30); наличие спаек или перенесенные абдоминальные операции; плохая подготовка кишечника; низкий опыт эндоскописта [198].

Методами, способствующими повышению значений CIR в рамках эндоскопического отделения медицинского учреждения, являются [199-204]:

- Повышение квалификации и опыта врачей. Важную роль в этом играют физические симуляторы процесса колоноскопии. Например, в рамках ежегодного Ярославского эндоскопического симпозиума (YES – Yaroslavl Endoscopy Symposium) традиционно проводится конкурс «Отработка и совершенствование навыков манипуляций эндоскопом и координации глазрука в неклинических условиях».
- Использование ИИ-ассистентов (алгоритмы для навигации и детекции купола способны повысить CIR на 3-5%).
- Качественная подготовка кишечника пациента (оптимизация диеты и режима очистки кишечника повышает СІК до 94-96%).

Проведенная неполностью процедура колоноскопии, когда КСК не достигнут, снижает эффективность скрининга колоректального рака (КРР) и может привести к пропуску полипов в правой половине толстой кишки. В странах с внедренными системами контроля качества (США, Япония, ЕС) показатель CIR служит одним из индикаторов в процессе аккредитации соответствующих клиник [205-209].

Алгоритмы детектирования купола слепой кишки (АДК) с использованием методов ИИ представляют собой перспективный инструмент для автоматической оценки полноты колоноскопии, помощи эндоскопистам в навигации по толстой кишке и повышения показателя CIR. Таким образом, внедрение результатов работы АДК в стандартизированные протоколы колоноскопии используется для повышения качества процедуры.

В задаче разработки АДК могут использоваться различные алгоритмические подходы. Классические алгоритмы МО могут быть основаны на детектировании анатомических ориентиров, таких как илеоцекальный клапан (ileocecal valve, ICV), отверстие аппендикса (appendiceal orifice), гаустры (haustra) и слизистая оболочка слепой кишки. Такие алгоритмы используют стандартные, известные из научно-

технической литературы фиксированные признаки (например, SIFT, HOG) для обучения различных классификаторов (SVM, Random Forest) [56].

Современные алгоритмы, основанные на глубоком МО также можно использовать в качестве основы для решения задач детектирования КСК, сегментации области КСК, анализа локальных окрестностей видеоданных (например, с помощью рекуррентных НС), классификации изображений с помощью трансформерных архитектур [119].

Оценка производительности АДК включает в себя использование как стандартных для МО метрик, рассмотренных в главе 1, так и специфических показателей CIR и TCI (Time to Cecal Intubation – время, достижения купола слепой кишки). Важным аспектом остается сравнение метрик для АДК с аналогичными результатами, полученными в исследованиях опытных эндоскопистов.

Разработка и внедрение АДК сталкивается с определенными проблемами. К ним относятся: недостаток аннотированных изображений с четко идентифицированным куполом слепой кишки и ICV; плохая подготовка кишечника, затрудняющая визуализацию купола слепой кишки и ICV; большое число ложноположительных результатов нейросетевого детектора; необходимость проведения испытаний для оценки эффективности АДК в реальных клинических условиях [211-213].

Таким образом, анализ научно-технической литературы показывает, что разработка АДК является актуальным направлением, позволяющим повысить качество процедуры, снизить риски неполных колоноскопий и повысить значение ключевого параметра CIR. Современные исследования демонстрируют значительный прогресс в разработке эффективных и робастных подходов для АДК. Однако для широкого внедрения в клиническую практику необходимо решить ряд алгоритмических проблем. Их рассмотрению и посвящена настоящая глава.

3.2 Описание собранной базы видеоизображений

При использовании методов ЦОИ и глубокого МО к анализу видеоданных колоноскопических исследований следует разделять два класса задач. Первый из

них связан непосредственно со статическими изображениями с разметкой объектов интереса (полипы, купол слепой кишки, новообразования и др.). Здесь решаются задачи детектирования, классификации, сегментации на фиксированной базе изображений (датасете). К сожалению, размер доступных баз таких изображений на практике является весьма ограниченным и не соответствует уровню развития нейросетевых технологий на 2025 г. Тем не менее возможность тестировать алгоритмы на небольшом валидированном датасете открывает широкие возможности для научных исследований в данной области [178, 181].

Второй класс задач связан непосредственно с анализом видеоданных колоноскопических исследований. Длительность записи одного исследования составляет обычно около 10 минут. Объект интереса всегда представлен на нескольких (чаще всего десятках и даже сотнях) кадров, что потенциально позволяет улучшить результаты его обнаружения и классификации. Однако и обработка таких видеоданных требует повышенных вычислительных мощностей и более сложных алгоритмов обработки, учитывающих корреляцию между кадрами.

В связи с ограничениями, связанными с объемом видеоданных, а также с существующими ограничениями в области персональных данных, сбор и накопление таких баз на федеральном и международном уровнях на сегодняшний день представляется затруднительным. Поэтому пока не существует открытых унифицированных наборов таких видеопоследовательностей, что сильно затрудняет сравнение результатов, полученных разными исследователями [178].

Для развития систем ИИ в колоноскопии важным представляется решение задач как первого, так и второго классов. При этом отдельное внимание следует уделять вопросам оценки качества входного изображения, не допуская для автоматического анализа НС неинформативные кадры исследования, процент которых в реальном колоноскопическом исследовании очень велик в связи с турбулентным характером движения колоноскопа.

Для обучения нейросетевого детектора необходимо произвести сбор и разметку базы видеоданных. На базе ЯОКОБ проведены исследования и сохранены соответствующие видеоматериалы, из которых удалялись персональные данные.

118

С помощью манипулятора типа «педали» врачи разметили области появления КСК купола слепой кишки. Собранный набор видеоданных обладает следующими характеристиками [214, 215]:

- общее число пациентов 103;
- средняя длительность видеопоследовательностей 17 минут 28 сек;
- средняя длительность нахождения купола на записи 1 минута 19 сек;
- количество кадров в секунду (fps) 25;
- разрешение кадра в процессе обучения (после его обрезки) 640×640 пикселей.

Стоит отметить, что при формировании базы видеозаписей врачами намеренно завышен показатель средней длительности времени наличия купола в кадре, т.е. проводилась описанная в главе 2 процедура физической аугментации данных. В рамках реального исследования этот показатель варьируется от 1 до 15 секунд в зависимости от квалификации и реакции эндоскописта. Это позволило собрать материал для формирования обучающей (8 436 шт.), валидационной (1 024 шт.) и тестовой (1 024 шт.) выборок изображений (рисунок 3.2), что позволяет обучать современные алгоритмы глубокого МО. Создана также дополнительная тестовая база колоноскопических обследований пациентов, которая позволила проводить независимую оценку точности работы АДК. Она будет рассмотрена в главе 6.

База содержит цифровые изображения, полученные со следующих видеоэндоскопических систем: FUJIFILM ELUXEO: серия 7000; OLYMPUS EXERA III: серии 180 и 190; PENTAX EPK-i7010.

Однако количество пациентов и объем уникальных данных в базе изображений не является достаточным для уверенного выделения НС ряда общих признаков между различными видами КСК, что отражается на качестве работы обученного нейросетевого детектора. Данная проблема обусловлена анатомическими особенностями строения кишечника, а именно тем, что снятый под разными углами КСК по визуальным параметрам может почти не отличаться от складки или бугра на стенке кишечника, как показано на рисунке 3.2.





Рисунок 3.2 – Примеры изображений из базы: с куполом слепой кишки (а-в); слизистой оболочки кишечника (г-е)

Эта схожесть приводит к повышению частоты ложноположительных срабатываний (принятие складки за купол) и ложноотрицательных результатов (пропуск истинного купола), снижая клиническую полезность алгоритма. Таким образом, ключевая сложность заключается в низкой специфичности визуальных маркеров купола в условиях меняющейся эндоскопической картины и ограниченных ракурсов обзора, что делает его надежную автоматическую детекцию без достаточного разнообразия видеоданных достаточно трудной задачей. Рассмотрим базовый алгоритм детектирования области КСК с использованием современных нейросетевых архитектур.

3.3 Исследование базового нейросетевого алгоритма детектирования купола на базе моделей SSD и YOLO

Для решения задачи детектирования купола на видеоизображениях в качестве первичных архитектур HC выбраны SSD-300 и SSD-512 на базе сети VGG-16

120

(подробней описаны в главе 5 настоящей работы) и семейство сетей YOLO в виде YOLOv5s, YOLOv7, YOLOv8s, YOLOv8n [112, 113]. Выбор данных архитектур, приведенных на рисунке 3.3. обусловлен их высокой производительностью в области детектирования объектов при сохранении высокой точности.



Рисунок 3.3 – Архитектуры нейронных сетей SSD (a), YOLOv7 (б)

изображений Описанная база разделялась обучающую, выше на валидационную и тестовую. Стоит отметить, что данные, разнесённые по выборкам, взяты из разных исследований, жестко закреплены за выборкой и не имеют пересечений. Так как медицинские и в том числе эндоскопические изображения имеют крайне высокую вариативность, этот шаг позволит произвести не только независимое тестирование, но и понять, как нейросетевой алгоритм может обрабатывать новые данные, которые не участвовали в обучении. Результаты тестирования **P-Presicion**, по метрикам R-Recall, F1 и mAP приведены в таблице 3.1 и таблице 3.2.

Нейронная сеть	Р	R	F1	mAP@0,50	mAP@[0,500,95]
YOLOv5s	0,988	0,974	0,981	0,988	0,598
YOLOv7	0,983	0,987	0,985	0,993	0,658
YOLOv8s	0,998	0,996	0,997	0,995	0,753
YOLOv8n	0,991	0,988	0,989	0,994	0,753
SSD300(VGG16)	0,991	0,987	0,989	0,906	0,549
SSD512(VGG16)	0,994	0,994	0,994	0,908	0,602

Таблица 3.1 – Исследование детекторов КСК на валидационной выборке

Таблица 3.2 – Исследование детекторов КСК на тестовой выборке

Нейронная сеть	Р	R	F 1	mAP@0,50	mAP@[0,500,95]
YOLOv5s	0,368	0,432	0,397	0,331	0,292
YOLOv7	0,392	0,348	0,369	0,374	0,232
YOLOv8s	0,411	0,387	0,399	0,353	0,215
YOLOv8n	0,415	0,401	0,408	0,389	0,273
SSD300(VGG16)	0,322	0,367	0,343	0,292	0,287
SSD512(VGG16)	0,343	0,354	0,348	0,217	0,296

Приведенные результаты показывают, что детекторы на базе семейства архитектур YOLO показывают лучшие результаты, по сравнению с подходом на базе детектора SSD. Единственным исключением является тестирование по метрике mAP@[0,50...0,95], показывающее сопоставимые результаты. Кроме того, обучение алгоритмов на базе YOLO занимает в несколько раз меньше времени при тех же вычислительных ресурсах [214].

Тем не менее, несмотря на высокие значения метрик для валидационного тестирования, при проведении исследования на сторонних данных возникало большое количество ложноположительных примеров детектирования купола с высоким уровнем показателя вероятности. Эта ситуация вынуждала в значительной степени поднимать порог допуска, при этом количество ложноотрицательных примеров резко возрастало, что негативно влияло на работу блока принятия решений и приводило к постановке неверных заключений о качестве проведения обследования [215]. При детальном анализе выявлено, что данная особенность алгоритма обусловлена высокой вариативностью данных в сочетании с относительно небольшой обучающей выборкой. Система давала сбой, поскольку обучающая выборка содержала в себе изображения КСК под разными углами, кроме того, встречались примеры полного или частичного перекрытия области купола его стенками (рисунок 3.4).

Описанная выше особенность в свою очередь приводила к тому, что система ИИ, оценивая только визуальную составляющую, принимала складки стенок кишечника за КСК, так как НС не имела иных примеров в процессе обучения, характеризующих складки и бугры. С учетом этого проведен дополнительный этап тестирования с добавлением в обучающий набор кадров с фоном и данных, полученных в результате сбора примеров ложноположительных срабатываний нейросетевых детекторов [216]. Для формирования выборки использовался алгоритм на основе нейросетевой модели YOLOv8n, которая предварительно обучалась на 8 436 изображениях, полученных ранее.

На вход НС поочередно подавались изображения из базы фонов. При положительном срабатывании НС происходило сохранение кадра в отдельную директорию. Изображения в базу фонов также нарезались с использованием временных интервалов, выделенных врачами с использованием педали во время исследования.

Для определения наиболее эффективной пропорции добавления ложноположительных срабатываний в обучающую выборку использовалось разное количество фоновых изображений в пропорциях 1:2, 1:1 и 3:2, что эквивалентно 4 200, 8 400 и 12 600 кадрам соответственно. В дополнении к предыдущему шагу в ходе тестирования алгоритма применялся подход изменения порога срабатывания блока принятия решений.



Рисунок 3.4 – Пример работы АДК на тестовых данных: кадр без купола, без детектирования (а); кадры с ложным срабатыванием (б-г); кадры с куполом (д, е)

e)

д)

Совместное применение этих методов (добавление фоновых изображений и вариация порога) дало возможность оценить их комплексное влияние на точность алгоритма. Обучение нейронной сети YOLOv8n производилось при стандартных

124

параметрах, а именно:

- структура разметки: прямоугольники BBox;
- размер батча: 16 (объем данных или количество изображений, одновременно поступающих на вход видеокарты для проведения обучения HC);
- алгоритм оптимизации: Adam (описан в главе 5 настоящей работы);
- начальная скорость обучения: 10⁻³.

Результаты тестирования АДК на основе YOLOv8n после обучения на базе с ложноположительными примерами приведены в таблице 3.3. Из полученных данных видно, что указанный алгоритм показывает наибольшую точность mAP@0,50, равную 0,616, и mAP@[0,50...0,95], равную 0,459, при количестве ложноположительных примеров в размере 8 400 штук и порогах срабатывания 0,5 и 0,75, соответственно.

Результаты тестирования базового нейросетевого алгоритма детектирования КСК, приведенные в таблицах 3.2 и 3.3 показывают, что максимальное значение метрики F1 достигает 0,439. При добавлении фонов в обучающую выборку удается существенно увеличить метрику Presicion (до 0,996), но при этом пропорционально уменьшается метрика Recall (до 0,22). Повышение метрики Precision до 0,9 при падении Recall после добавления ложноположительных примеров означает, что модель стала гораздо осторожнее в своих предсказаниях. Ранее (при P=R=0,4) алгоритм допускал много ошибок обоих типов: он часто пропускал настоящие КСК (низкий Recall) и принимал за КСК то, что им не является (низкий Precision). Добавление примеров фона научило модель лучше отличать КСК от похожих структур, резко сократив количество ложных срабатываний – отсюда увеличение метрики Precision. Однако эта «осторожность» привела к обратному эффекту: модель теперь чаще пропускает настоящие КСК, что снизило метрику Recall. Требуется поиск баланса модели, подходящего для рассматриваемой задачи, т.к. в медицинской диагностике метрика Recall критически важна.

Кол-во фонов	Порог	Р	R	F1	mAP@0,50	mAP@[0,500,95]
	0,5	0,620	0,340	0,439	0,501	0,324
	0,75	0,768	0,290	0,395	0,550	0,382
0	0,8	0,854	0,263	0,369	0,571	0,408
	0,85	0,921	0,233	0,339	0,582	0,427
	0,9	0,923	0,030	0,057	0,476	0,348
	0,5	0,916	0,245	0,387	0,590	0,432
	0,75	0,979	0,233	0,372	0,608	0,452
4 200	0,8	0,996	0,225	0,361	0,612	0,457
	0,85	0,996	0,203	0,332	0,601	0,450
	0,9	0,996	0,003	0,005	0,501	0,451
8 400	0,5	0,980	0,247	0,395	0,616	0,444
	0,75	0,996	0,233	0,376	0,611	0,459
	0,8	0,996	0,233	0,376	0,611	0,459
	0,85	0,996	0,180	0,304	0,590	0,450
	0,9	-	_	_	-	-
12 600	0,5	0,912	0,233	0,371	0,582	0,430
	0,75	0,989	0,225	0,361	0,608	0,454
	0,8	0,996	0,220	0,354	0,610	0,455
	0,85	0,996	0,193	0,319	0,596	0,450
	0,9	0,996	0,013	0,025	0,506	0,405

Таблица 3.3 – Результаты тестирования АДК на основе YOLOv8n при обучении с фонами

Рассмотрим возможные способы совершенствования базового алгоритма детектирования купола, предполагающие работу с последовательностью кадров видеопоследовательности колоноскопического исследования. Для этого предполагается использовать два класса алгоритмов: трекинга и анализа оптического потока (ОП) [39].

3.4 Разработка улучшенного алгоритма детектирования купола слепой кишки с использованием анализа видеопоследовательности

Применение алгоритмов трекинга объектов и анализа ОП позволяет использовать временную согласованность видеоданных, что принципиально улучшает детектирование по сравнению с анализом одиночных кадров. Современные алгоритмы трекинга связывают обнаружения купола между последовательными кадрами, отфильтровывая случайные ложные срабатывания – если объект не отслеживается стабильно во времени, это вероятно артефакт. Одновременно анализ ОП (движения пикселей между кадрами) помогает определить точку перегиба в движении эндоскопа, наблюдаемую в зоне купола.

Такой подход потенциально дает два ключевых преимущества:

- Повышение метрики Recall система «додумывает» купол в кадрах, где он частично скрыт или плохо виден, опираясь на его положение в соседних кадрах.
- Повышение метрики Precision ложные срабатывания подавляются за счет требования временной устойчивости и анализа паттернов движения.

Таким образом, анализ видеопоследовательности способен компенсировать ограничения детекции, используя контекст, недоступный при обработке статичных изображений [39, 216].

3.4.1 Алгоритм детектирования купола с использованием трекинга

Разрабатываемый алгоритм детектирования купола с трекингом будем обозначать, как АДКТ. Алгоритмы трекинга предназначены для того, чтобы собирать статистику по нескольким последовательным предсказаниям объекта и при наличии достаточного количества данных в рамках небольшого промежутка времени объединять их в обобщенные сущности, называемые треком. Современные подходы к решению задачи трекинга активно развиваются, сочетая методы глубокого МО с вычислительной оптимизацией для работы в PPB. Однако типовые проблемы, такие как обработка перекрытий и устойчивость к искажениям и шумам, остаются актуальными [217-220].

Блок-схема разработанного алгоритма приведена на рисунке 3.5.



Рисунок 3.5 – Блок-схема алгоритма детектирования КСК с применением трекинга

Алгоритм анализирует видеофрагмент последовательно, кадр за кадром, используя вероятностный подход и временной контекст для обнаружения КСК. Для каждого кадра сначала рассчитывается вероятность нахождения КСК. Если на предыдущем кадре КСК уже был обнаружен, система пересчитывает текущую вероятность с учетом предыдущих данных, повышая устойчивость детекции.

После обработки всего фрагмента алгоритм проверяет, обнаружены ли интервалы с КСК. Если да – выбирается интервал с максимальной совокупной вероятностью (например, фрагмент, где КСК появлялся стабильно), и на выходе формируется временной отрезок его присутствия в видео. Если ни в одном интервале КСК не найден –возвращается статус «КСК не обнаружен».

Основная особенность предлагаемого алгоритма – решение о нахождении КСК на видеопоследовательности принимается не по одному кадру, а на основе анализа временной корреляции (связи между кадрами) и агрегации вероятностей, что снижает ложные срабатывания и пропуски.

Для трекинга предсказаний КСК использовались алгоритмы OSTrack, ByteTrack и BoT-Sort [218-220]. В качестве показателя, отвечающего за сходство детекций и позволяющего отнести обнаружение к одному из возможных вариантов (*TP*, *FP* или *FN*), применяется стандартная метрика отношения пересечения к объединению (Intersection over Union, *IoU*):

$$IoU = \frac{S_{int}}{S_{union}},$$

где S_{int} и S_{union} – площади пересечения и объединения между экспертной разметкой и обнаружением, предсказанным алгоритмом (введена ранее в пункте 1.9)

Далее требуется определить наиболее вероятный временной интервал из набора треков. Блок принятия решений производит его выбор согласно следующему выражению [214]:

$$T_p = T_n k + S_n (1 - k), k \in [0, 1],$$

где T_p – наиболее вероятный интервал времени, T_n – промежуток времени трека, нормированный на максимальный среди набора треков, S_n – средняя точность трека, k=0,3 – подобранный эмпирически нормировочный коэффициент (при $k\rightarrow 0$ получаем приоритет по средней точности, при $k\rightarrow 1$ – по длительности отдельного трека). Однако для итоговой системы обработки видеопотока процесс выбора

единственной временной области купола из нескольких возможных все еще остается затруднительным, поскольку приведенное выше выражение не может учесть все возможные случаи и не обладает достаточной гибкостью.

Таким образом, алгоритм трекинга связывает данные о предсказании наличия купола слепой кишки в различных областях видеофрагмента и формирует связанные треки, компенсируя при этом ложноположительные выбросы [221, 222]. Кроме того, блок принятия решений предлагает на выходе сбалансированную оценку, которая характеризуется как длительностью нахождения предполагаемого купола в кадре, так и вероятностью его правильного обнаружения.

3.4.2 Алгоритм детектирования купола с использованием анализа оптического потока

После проведения экспериментов по тестированию АДК с модулем трекинга (АДКТ) было принято решение внедрить дополнительный модуль для обработки ОП между соседними видеокадрами и определения направления движения эндоскопа согласно данным, полученным от этого блока. Разработанный алгоритм детектирования купола представлен на рисунке 3.6. Он включает в себя непосредственно нейросетевой детектор, работающий в связке с алгоритмом трекинга, алгоритм анализа ОП и алгоритм ситуационной логики, увязывающий логику работы всех алгоритмов.

обрабатывает Алгоритм последовательно эндоскопическую видеопоследовательность, комбинируя нейросетевые методы с анализом соседних кадров для надежного обнаружения КСК. Сначала модель YOLOv8 выделяет области-кандидаты в каждом кадре, после чего алгоритм трекинга связывает эти области между кадрами, формируя устойчивые треки объектов. Параллельно анализ ОП оценивает характер движения эндоскопа: определение точки изменения направления происходит посредством расчета среднего арифметического направлений интервале исследований предсказанных В заданном (для тестировались интервалы длительностью 10, 20 и 30 сек.): движение вперед – (+1), остановка на месте -(0), движение назад -(-1).



Рисунок 3.6 – Блок-схема алгоритма детектирования области КСК с применением методов анализа оптического потока

Затем нейросетевой классификатор верифицирует кандидатов, используя данные ОП для расчета вероятности присутствия КСК в каждом временном интервале. На этапе постобработки короткие/нестабильные треки отфильтровываются, а алгоритм ситуационной логики определяет итоговый результат: если найден интервал с максимальной и стабильной вероятностью КСК – система сохраняет ключевые кадры, записывает метки временного интервала в МИС; если устойчивые признаки КСК отсутствуют – фиксирует в МИС статус «КСК не обнаружен». Запись в МИС при любом исходе исследования гарантирует документирование результата для клинической практики [215].

Рассмотрим механизм работы алгоритма анализа ОП более подробно.

3.4.3 Алгоритм расчета оптического потока на основе метода Хорна-Шунка

В основу блока классификации направления движения эндоскопа лег модифицированный алгоритм Хорна-Шунка по вычислению ОП [39, 40, 223, 224]. Он рассчитывает движение между кадрами видеопоследовательности, взятыми в моменты t и $t+\Delta t$ в позиции каждого пикселя. Эти методы базируются на разложении функции, описывающей сигнал изображений I(x, y, t), в трехмерный ряд Тейлора с сохранением только слагаемых первого порядка:

$$I(x + \delta x, y + \delta y, t + \delta t) =$$

= $I(x, y, t) + \delta x \frac{\partial I}{\partial x}(x, y, t) + \delta y \frac{\partial I}{\partial y}(x, y, t) + \delta t \frac{\partial I}{\partial t}(x, y, t).$

В качестве δt берется разница между моментами съемки двух последовательных кадров: I(x, y, t) и I(x, y, t + 1). С помощью δx и δy оценивается перемещение пикселя в интервале времени от t до t + 1. Предполагаем, что сигнал изображения остается постоянным при бесконечно малых сдвигах по трехмерным координатам, т.е. $I(x + \delta x, y + \delta y, t + \delta t) = I(x, y, t)$, тогда это выражение принимает следующий вид:

$$0 = \delta x \frac{\partial I}{\partial x}(x, y, t) + \delta y \frac{\partial I}{\partial y}(x, y, t) + \delta t \frac{\partial I}{\partial t}(x, y, t)$$

Разделив на δt получим:

$$0 = \frac{\delta x}{\delta t} \frac{\partial I}{\partial x}(x, y, t) + \frac{\delta y}{\delta t} \frac{\partial I}{\partial y}(x, y, t) + \frac{\partial I}{\partial t}(x, y, t).$$

В итоге, уравнение ОП можно записать в более компактном виде:

$$I_x V_x + I_y V_y + I_t = 0,$$

где I_x , I_y , I_t – производные функции сигнала изображения при (x, y, t) в соответствующих направлениях, V_x и V_y – компоненты ОП для I(x, y, t). Для решения уравнения ОП часто используется итерационная схема. Использовались следующие аппроксимации для частных производных I_x , I_y , I_t :

$$\begin{split} I_x(x,y,t) &= \frac{1}{4} [I(x+1,y,t) + I(x+1,y,t+1) + I(x+1,y+1,t) \\ &+ I(x+1,y+1,t+1)] - \\ &- \frac{1}{4} [I(x,y,t) + I(x,y,t+1) + I(x,y+1,t) + I(x,y+1,t+1)], \\ I_y(x,y,t) &= \frac{1}{4} [I(x,y+1,t) + I(x y + 1,t+1) + I(x + 1,y + 1,t) \\ &+ I(x+1,y+1,t+1)] - \\ &- \frac{1}{4} [I(x,y,t) + I(x,y,t+1) + I(x + 1,y,t) + I(x + 1,y,t+1)], \\ I_t(x,y,t) &= \frac{1}{4} [I(x,y,t+1) + I(x,y+1,t+1) + I(x + 1,y,t+1)] - \\ &- \frac{1}{4} [I(x,y,t) + I(x,y+1,t) + I(x+1,y,t) + I(x + 1,y,t+1)]. \end{split}$$

В результате работы алгоритма для каждого пикселя получаем вертикальное смещение У в диапазоне [-1; 1] и горизонтальное смещение Х в диапазоне [-1; 1].

Для того чтобы изобразить оба смещения на одном изображении наиболее удобным является переход из декартовой системы координат в полярную, где модуль вектора ρ показывает величину смещения, а угол φ определяет его направление:

$$\rho = \sqrt{x^2 + y^2}, \quad \mathrm{tg}\varphi = \frac{y}{x},$$
$$\cos \varphi = \frac{x}{\sqrt{x^2 + y^2}}, \quad \sin \varphi = \frac{y}{\sqrt{x^2 + y^2}}.$$

В качестве цветовой модели, визуализирующей полярные координаты, выбрана HSV (Hue – цветовой тон, Saturation – насыщенность, Value – значение цвета) [36, 45]. Для перехода установлено, что значение цвета (V) отражено величиной модуля вектора ρ , цветовой тон (H) определяется углом ϕ , а параметр насыщенности (S) постоянный и равен максимальному значению. Такой выбор позволяет однозначно изобразить результаты работы математической модели определения направления смещения ОП на одном изображении (рисунок 3.7).



Рисунок 3.7 – Использование цветовой модели HSV: визуальное представление (a); представление в полярных координатах (б); визуализация движения вглубь кадра (в); визуализация движения из кадра (г)

Для определения направления движения кадра, необходимо определить направление смещения на границах объектов. Для этого удобно использовать следующую цветовую логику. Если объект окрашен по контуру как в случае, показанном на рисунке 3.7в, это значит, что его границы стремятся к краям кадра, следовательно, камера приближается к нему. Аналогично, если контур объекта окрашен по схеме, приведенной на рисунке 3.7г, то это значит, что камера удаляется от него и движется назад. Примеры визуализаций кадров при движении вперед, назад и останове эндоскопа приведены на рисунке 3.8.

Исходя из того, какой цвет преобладает в той или иной области изображения, можно сделать вывод об итоговом направлении смещения камеры в данный момент времени (движение вперед, назад или останов). Приведенный визуальный пример говорит о том, что человек способен по изображению определить тип движения эндоскопа, а значит решение такой задачи способна делать и глубокая HC.





Рисунок 3.8 – Визуализации смещения для алгоритма оптического потока: движение вперед (а-б); движение назад (в-г); останов (д-е)

3.4.4 Нейросетевой алгоритм дополнительной классификации данных

Для классификации движения эндоскопа использовался дополнительный алгоритм классификации данных, основанный на глубоком обучении. В качестве исследуемых нейросетевых моделей выбраны три архитектуры: сверточная HC MobileOne, трансформер FastViT и гибридная модель на базе сверточных слоев и трансформера CaFormer [119-122]. Собранная база изображений включает в себя три класса: «нет движения», «движение вперед» и «движение назад». Статистика для разбиения базы приведена в таблице 3.4. Входные изображения приведены к одному разрешению 224х224 пикселя. Представленные сети обучены на протяжении 400 эпох с батчем (фрагментом) – 256.

Выборка Класс	Обучающая, шт.	Тестовая, шт.	Валидационная, шт.
Нет движения	68 010	3 580	3 479
Движение вперед	22 479	1 184	1 195
Движение назад	15 230	802	793
Всего	105 719	5 566	5 467

Таблица 3.4 – Набор изображений для классификации движения эндоскопа

Зависимости изменения функции потерь при обучении для всех трех алгоритмов представлены на рисунках 3.9-3.11. Они показывают, что наибольшая скорость и степень сходимости наблюдается для моделей FastViT-sa12 и CAFormer-s18. При обучении сети MobileOne-s0 функция потерь падает незначительно. Количество эпох, необходимых для процесса обучения, во всех трех случаях превышает 300. Для оценки точности процесса классификации на валидационной выборке использовалась метрика Accuracy, введенная в главе 1. Результаты исследования приведены на рисунках 3.12-3.14. Они показывают более медленную сходимость метрики для модели CAFormer-s18 к итоговому значению, по сравнению с аналогами.



Рисунок 3.9 – Зависимость изменения функции потерь для нейронной сети MobileOne-s0 от номера эпохи на обучающей выборке



Рисунок 3.10 – Зависимость изменения функции потерь для нейронной сети FastViT-sa12 от номера эпохи на обучающей выборке



Рисунок 3.11 – Зависимость изменения функции потерь для нейронной сети CAFormers18 от номера эпохи на обучающей выборке



Рисунок 3.12 – Зависимость изменения точности для нейронной сети MobileOne-s0 от номера эпохи на валидационной выборке



Рисунок 3.13 – Зависимость изменения точности для нейронной сети FastViT-sa12 от номера эпохи на валидационной выборке



Рисунок 3.14 – Зависимость изменения точности CAFormer-s18 от номера эпохи на валидационной выборке

Для более точного сравнения трех моделей, рассчитаны метрики Precision, Recall и F1 на полном наборе имеющихся изображений. Результаты приведены в таблицах 3.5–3.7.

Класс	Р	R	F1
Движения нет	0,903	0,934	0,919
Движение вперед	0,860	0,792	0,825
Движение назад	0,819	0,789	0,804
Среднее	0,861	0,839	0,849

Таблица 3.5 – Точность классификации для модели MobileOne-s0 на полном наборе данных

Таблица 3.6 – Точность классификации для модели FastViT-sa12 на полном наборе данных

Класс	Р	R	F1
Движения нет	0,916	0,938	0,927
Движение вперед	0,876	0,823	0,849
Движение назад	0,829	0,815	0,822
Среднее	0,874	0,859	0,866

Таблица 3.7 – Точность классификации для модели CAFormer-s18 на полном наборе данных

Класс	Р	R	F1
Движения нет	0,903	0,934	0,918
Движение вперед	0,862	0,798	0,829
Движение назад	0,824	0,791	0,807
Среднее	0,863	0,841	0,851

Анализ результатов моделирования показывает, что наилучшую точность работы по всем метрикам показывает алгоритм на базе трансформера FastViT-sa12, где средняя точность по метрике F1 по всем классам составляет 0,866. Результаты для двух других моделей проигрывают незначительно, что позволяет использовать любую из них для практической реализации (с учетом вычислительных затрат и сложности в реализации) [214, 216]. С точки зрения результатов для разных классов – самая устойчивая идентификация достигается для класса «движения нет», на котором метрика F1 превышает значение 0,9.

3.5 Оценка качества колоноскопических изображений

Рассмотрим особенности построения неэталонных алгоритмов оценки качества изображений, рассмотренных выше в главе 1 на наборе статических кадров с колоноскопического исследования.

3.5.1 Описание набора колоноскопических изображений

Собранная совместно с врачами ЯОКОБ база состоит из 1000 статических изображений, 10 видеопоследовательностей взятых ИЗ различных колоноскопических исследований. Четыре видеопоследовательности получены с чересстрочным форматом изображения, на остальных – эффект чересстрочности (ЭЧ) отсутствует. Статические изображения нарезались из видеоданных путем сохранения из них каждого десятого кадра. Затем вручную отбрасывались последовательные имеющие кадры, не заметных видимых различий. Из оставшихся кадров случайным образом выбрано 1000 изображений. Изображения с ЭЧ в итоге составили около половины от всех собранных в наборе [225, 226].

Далее происходила процедура субъективного (экспертного) оценивания качества собранного набора изображений по оценкам MOS (Mean opinion score). Группа из пяти экспертов по обработке изображений поставила каждому изображению оценки от «1» до «5». Оценка «1» ставилась, если изображение являлось полностью не информативным. Оценка «5» ставилась для идеального информативного изображения. Для изображений с оценкой «3» допускалось наличие зашумленных или размытых областей, при условии четкой визуальной интерпретации полезных областей. Промежуточные оценки «2» и «4» ставились, если изображения логически сложно соотнести с установленными критериями для оценок «1», «3» или «5».

На рисунке 3.15 показаны примеры колоноскопических изображений с усредненными экспертными оценками качества MOS.



a) MOS = 1,6
 b) MOS = 2,8
 c) MOS = 4
 Pисунок 3.15 – Примеры изображений с усредненными экспертными оценками MOS

На этом этапе дополнительно рассчитывался ряд статистических показателей. Установлено, что коэффициент линейной корреляции оценок экспертов друг с другом составил от 0,6 до 0,68. Экспертам дополнительно предлагалось повторно оценить 50 изображений. Корреляция собственных оценок в двух случаях составила от 0,75 до 0,84. Рассчитывалось также среднее арифметическое пяти экспертных оценок MOS. Среднее квадратичное отклонение полученной оценки составило 0,72.

3.5.2 Анализ использования известных неэталонных алгоритмов оценки качества изображений

Популярный алгоритм расчета метрики BRISQUE, приведенный на рисунке 3.16, сопоставляет каждому изображению вектор, состоящий из 36 признаков [148, 227].



Рисунок 3.16 - Структурная схема вычисления метрики BRISQUE

Из них 18 признаков вычисляется для исходного изображения, а другие 18 – для изображения, уменьшенного в 2 раза. Итоговое значение метрики

142

вычисляется с использованием предварительно обученного на базе экспертных оценок алгоритма МО. Для расчета признаков алгоритм BRISQUE вычисляет распределение коэффициентов, отвечающих за распределение статистики в естественных изображениях (NSS) (рисунок 3.17) для оцениваемого изображения. Признаки представляют собой коэффициенты симметричного (GGD) и ассиметричного (AGGD) нормальных распределений, наилучшим образом аппроксимирующих наблюдаемую статистику данных.



Рисунок 3.17 – Структурная схема вычисления признаков NSS

Метрика NIQE [228] оценивает различие между распределением коэффициентов NSS для изображения и приближением того же самого распределения с использованием многомерного нормального распределения. Для своего расчета она не использует методы МО на экспертных оценках.

Известно, что традиционные метрики оценки качества плохо решают проблему контекста: они не способны анализировать содержимое изображения, хотя оно влияет на субъективную оценку качества [229-232]. Для решения этой проблемы исследователи прибегают к созданию алгоритмов, использующих СНС и другие алгоритмы глубокого МО. К ним, например, относятся метрики TOPIQ [233] и PaQ-2-PiQ [234]. Область их применения на практике ограничена наличием большого (желательно от 10 000 штук) обучающего набора изображений с экспертными оценками MOS.

В ходе проведенного эксперимента рассчитан коэффициент линейной корреляции между экспертной оценкой MOS и описанными выше известными алгоритмическими метриками неэталонной оценки качества:

$$\rho_{x,y} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}} ,$$

где x_i , y_i – экспертные и алгоритмические оценки *i*-го изображения, а \bar{x} , \bar{y} – средние значения оценок на всем множестве изображений.

Исследование проводилось: на всем наборе изображений, только на изображениях с ЭЧ, только на изображениях без ЭЧ. Результаты приведены в таблице 3.8. Вычисление значений метрик произведено с использованием библиотеки IQA-PyTorch [235]. Метрики BRISQUE, NIQE, TOPIQ оценивают степень зашумленности, поэтому обратно коррелируют с качеством изображения, по этой причине в таблице 3.9 приведены абсолютные значения коэффициента корреляции. Установлено, что наилучшие результаты на всем наборе демонстрируют метрики BRISQUE и NIQE, использующие признаки NSS и два пространственной пирамиды (используется как исходное, уровня так И уменьшенное в два раза изображение).

Метрики	Все изображения	Изображения с ЭЧ	Изображения без ЭЧ
BRISQUE	0,46	0,71	0,08
NIQE	0,50	0,67	0,37
TOPIQ	0,29	0,18	0,36
PaQ-2-PiQ	0,18	0,38	0,36

Таблица 3.8 – Значения коэффициента линейной корреляции для разных метрик

Одним из этапов вычисления признаков NSS является нахождение распределений разностей значений между соседними пикселями. В частности, вычисляются разности с соседями по горизонтали и вертикали. У изображений с ЭЧ эти распределения будут отличаться, но только на первом уровне пространственной пирамиды. На ее втором уровне изображение уменьшается в два раза, и характерные ЭЧ пропадают. Эти особенности могут объяснять высокую корреляцию оценок метрик BRISQUE и NIQE для изображений с ЭЧ.
3.5.3 Разработка алгоритмов на основе признаков NSS

Необходимо разработать алгоритм МО на наборе колоноскопических изображений, не изменяя способов вычисления коэффициентов NSS. При этом рассматривались методы линейной регрессии (LRG) и использование персептрона (PR) с одним скрытым слоем (100 нейронов в скрытом слое, функция активации ReLU). В качестве функции потерь использована среднеквадратичная ошибка (*MSE*):

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (z_i - \bar{z}_i)^2,$$

где Z_i и $\overline{Z_i}$ – исходная и предсказанные оценки для *i*-го изображения. Набор изображений делился на обучающее и тестовое множество в соотношении 7:3. В таблице 3.9 приведены результаты, полученные на тестовом множестве.

Алгоритм	Все изображения	Изображения с ЭЧ	Изображения без ЭЧ
LRG	<i>MSE</i> = 0,20	<i>MSE</i> = 0,10	<i>MSE</i> = 0,20
PR	MSE = 0,14	<i>MSE</i> = 0,11	<i>MSE</i> = 0,18

Таблица 3.9 – Значения среднеквадратичной ошибки

3.5.4 Разработка алгоритма на основе сверточной нейронной сети

На субъективную оценку качества изображения среди прочих факторов отрицательно влияют низкая освещенность и наличие бликов отраженного света. Кроме того, движение эндоскопа в органах кишечника приводит к частому появлению артефактов размытия на полученных изображениях. Напротив, если на изображении отчетливо видны стенки внутренних органов, то это положительно влияет на субъективную оценку. Выявление перечисленных выше признаков должно быть под силу алгоритму, базирующемуся на небольшой СНС. Одна из первых попыток использования СНС для оценки качества изображений приведена в работе [236].

В предлагаемой архитектуре (рисунок 3.18) несколько сверточных слоев возвращают на выходе тензор размерности 17х17х50. Затем, как и в оригинальной

работе, в каждом из 50 каналов извлекаются максимальное и минимальное значение. Кроме того, добавлено вычисление среднего значения в каждом канале. После конкатенации получается вектор из 150 элементов. С использованием нескольких полносвязных слоев вычисляется оценка качества. Далее в работе будем называть эту модель MMA-CNN (min+max+average – convolutional neural network).



Рисунок 3.18 – Блок-схема алгоритма оценки качества изображения на основе сверточной нейронной сети

В результате обучения среднеквадратичная ошибка на тестовом множестве составила 0,2, как и при использовании линейной регрессии по признакам NSS. Также визуальный анализ значений в последнем канале сверточного слоя позволяет сделать следующие наблюдения:

- Слишком темные и слишком яркие области отчетливо выделились в отдельных каналах. Один канал имеет отклики только для изображений с ЭЧ.
- Области «высокого качества» не имеют отдельных каналов.
- Примерно на четверти всех изображений виден зонд эндоскопа. Зачастую он неподвижен на подверженном размытию движения изображении или наоборот, размыт или засвечен на переднем плане в целом качественного изображения. Изображение зонда не представляет никакой ценности и не должно влиять на субъективную оценку, однако извлекаемые признаки слишком низкоуровневые, чтобы НС научилась его игнорировать.

3.5.5 Сравнительный анализ работы алгоритмов оценки качества

На новом тестовом множестве из 300 изображений вычислен коэффициент линейной корреляции (ρ) с экспертными оценками MOS. Результаты приведены в таблице 3.10. Они позволяют сделать вывод о преимуществе использования алгоритмов расчета метрик PR и MMA-CNN, обученных на колоноскопических изображениях, по сравнению с универсальной неэталонной метрикой типа BRISQUE.

Метрика	Коэффициент линейной корреляции			
BRISQUE	0,43			
PR	0,88			
MMA-CNN	0,83			

Таблица 3.10 – Сравнительный анализ рассматриваемых алгоритмов

На рисунке 3.19 приведены примеры изображений с экспертными оценками MOS и оценками, полученными рассмотренными выше алгоритмами. Видно, что оба алгоритма, базирующихся на методах MO (PR и MMA-CNN), в целом

справляются с оцениванием качества изображений, при этом несколько занижая среднюю экспертную оценку. Для практической реализации в системе контроля качества колоноскопического исследования выбран алгоритм MMA-CNN, обладающий невысокой вычислительной сложностью [237].



a) MOS=1,8; 6) MOS=2,8; b) MOS=4,2; BRISQUE=131,24; PR=1,4; BRISQUE=20,64; PR=3,03; BRISQUE=43,45; PR=3,72; MMA-CNN =1,5 MMA-CNN =2,6 MMA-CNN =3,6 Рисунок 3.19 – Сравнение оценок экспертов (MOS) и алгоритмических метрик

Таким образом, создан тестовый набор колоноскопических изображений, проведена его экспертная оценка, оценена корреляция ряда неэталонных алгоритмов с экспертными значениями, предложено 2 оригинальных алгоритма, использующих методы машинного обучения. В исследовании рассмотрены модели, использующие только низкоуровневые признаки или их статистику распределения. Такие модели менее подвержены переобучению на небольшом наборе эндоскопических изображений, который чаще всего и имеется в распоряжении исследователей.

3.6 Краткие выводы

Проведенные в третьей главе исследования позволяют сделать следующие основные выводы:

 Создана оригинальная база видеоданных колоноскопических исследований из ЯОКОБ, которая может использоваться для последующих НИР и НИОКР в области дальнейшего улучшения алгоритмов для оценки воспроизводимости и сопоставимости результатов исследований с использованием методов ИИ.

- Разработан метод контроля качества колоноскопического исследования, реализующий: оценку полноты процедуры колоноскопии (CIR) на базе алгоритма детектирования КСК на видеопоследовательности; аудит изображений, сохраненных в МИС в рамках проведенного скринингового или клинического исследования на наличие технических дефектов (размытых, с артефактами сжатия, неинформативных).
- Разработан алгоритм детекции КСК на основе нейронной сети YOLOv8
 с дополнительными алгоритмами трекинга и анализа оптического потока.
 Проверка его эффективности для автоматизации оценки полноты проведенной колоноскопии будет проведена в главе 6 настоящей работы.
- Включение модуля трекинга и добавление ложноположительных примеров в обучающую выборку существенно снижают количество ложных детекций КСК, вызванных визуальной схожестью области купола со складками кишечника.
- Алгоритм классификации направления движения эндоскопа на основе анализа оптического потока и классификатора на базе трансформера FastViT (F1=0,866) успешно определяет тип движения «вперед», «назад», «останов», интегрируясь в алгоритм детекции КСК для повышения точности.
- Внедрение алгоритма детекции КСК в клиническую практику потенциально позволит повысить показатель полноты колоноскопии (CIR) на 3-5%, снижая риск пропуска патологий в правой части толстой кишки.
- Разработаны и валидированы метрики оценки качества колоноскопических изображений (PR и MMA-CNN), показавшие высокую корреляцию (0,83-0,88) с экспертными оценками MOS и превосходящие универсальные метрики (BRISQUE, NIQE).

ГЛАВА 4 РАЗРАБОТКА И ИССЛЕДОВАНИЕ АЛГОРИТМОВ ДЕТЕКТИРОВАНИЯ И СЕГМЕНТАЦИИ КОЛОРЕКТАЛЬНЫХ ПОЛИПОВ

4.1 Актуальность задачи обнаружения полипов

Полипы – аномальные разрастания ткани, выступающие над слизистой оболочкой. Они могут возникать в любом месте ЖКТ, но чаще обнаруживаются в колоректальной области. Хотя большинство полипов являются доброкачественными, некоторые из них (особенно аденоматозные) могут со временем переродиться в колоректальный рак (КРР). Размер колоректальных полипов может существенно различаться: от совсем небольших и сложных в обнаружении (≤5 мм) до весьма значительных (≥20 мм). Своевременное и точное детектирование полипов кишечника, а также их удаление (полипэктомия – малоинвазивная эндоскопическая операция) являются важнейшим элементом в профилактике КРР [238-240].

Наиболее эффективным методом, позволяющим визуально осмотреть всю толстую кишку и при необходимости удалить полипы, по-прежнему остается колоноскопическая диагностика. Однако в среднем по имеющейся статистике при колоноскопических обследованиях врачами пропускается порядка 20% полипов. Новые эндоскопические устройства и диагностические инструменты увеличивают частоту обнаружения полипов, но проблема их пропуска остается по-прежнему актуальной [22-23]. Риски образования колоректальных полипов многофакторны и включают в себя: генетическую предрасположенность; возраст старше 50 лет; диету в виде низкого потребления клетчатки или высокого потребления жиров; ожирение; курение; чрезмерное употребление алкоголя; малоподвижный образ заболевания воспалительные кишечника; перенесенная ранее жизни: полипэктомия [32].

Процесс трансформации полипа в злокачественную опухоль (малигнизация) занимает в среднем 5-10 лет. Своевременное удаление полипов прерывает этот процесс и предотвращает развитие рака. Раннее выявление КРР также значительно повышает шансы на успешное лечение и выживаемость. Скрининг с частотой раз в 5 лет, начиная с 45-летнего возраста (включающий колоноскопию и другие методы), доказанно снижает смертность от этого заболевания [3, 5, 7]. Классификация методов обнаружения полипов кишечника приведена на рисунке 4.1.



Рисунок 4.1 – Классификация методов выявления полипов кишечника

Современные методы лиагностики такие как хромоколоноскопия, узкоспектральная эндоскопия (NBI) и другие позволяют улучшить визуализацию полипов и повысить точность их выявления. Метод КТ-колонографии является менее инвазивным и основывается на компьютерной томографии. Он требует особо тщательной подготовки кишечника и в дальнейшем может потребовать проведения обычной колоноскопии для удаления обнаруженных полипов. Осмотр нижней части толстой кишки (сигмовидной и прямой кишки) возможен путем гибкой Капсульная обладает сигмоидоскопии. колоноскопия меньшей чувствительностью и обычно стоит существенно дороже для пациента, однако такой метод пользуется популярностью в ряде регионов (более подходит с точки зрения культурных установок в обществе) и удобен для начальных скрининговых обследований [99-100]. Также для обнаружения полипов могут использоваться анализ кала на скрытую кровь (FOBT/FIT) – относительно простое исследование, которое может выявить наличие крови в кале, что может указывать на наличие полипов или рака, и анализ ДНК кала (Cologuard) – более чувствительный тест, который обнаруживает ДНК мутации, связанные с полипами и раком. В случае положительных результатов и в том и другом случае пациенту рекомендуется прохождение колоноскопии [32].

Примеры изображений, полученных из видеоданных колоноскопического исследования с отмеченными алгоритмом ИИ колоректальными полипами [167], приведены на рисунке 4.2.





Рисунок 4.2 – Примеры изображений колоректальных полипов

Существует несколько классификаций полипов кишечника, основанных на различных критериях [32, 35, 241]. Наиболее важные и часто используемые классификации учитывают гистологический тип, морфологию (форму) и размер полипов (таблица 4.1).

Таблица 4.1 – Классификация полипов кишечника

Тип классификации	Представители				
1. Гистологическая	1.1. Аденоматозные полипы (аденомы) – самый распрост-				
(по типу ткани):	раненный тип полипов, считающийся предраковым. Они				
	составляют около 70% всех полипов толстой кишки. Аденомы				
	подразделяются на несколько подтипов (тубулярные,				
	ворсинчатые, тубуло-ворсинчатые).				
	1.2. Гиперпластические полипы – обычно небольшие (менее 5 мм)				
	полипы, расположенные в прямой и сигмовидной кишке.				
	Считаются доброкачественными и не имеющими значительного				
	риска малигнизации, хотя некоторые исследования показывают,				
	что крупные гиперпластические полипы (>10 мм),				
	расположенные в правой половине толстой кишки, могут иметь				
	небольшой риск развития КРР.				
	1.3. Воспалительные полипы, которые образуются в результате				
	воспаления слизистой оболочки толстой кишки, например, при				
	язвенном колите или болезни Крона. Не имеют риска				
	малигнизации.				
	1.4. Зубчатые полипы – этот тип полипов включает несколько				
	подтипов, некоторые из которых имеют повышенный риск				
	развития КРР, особенно если они расположены в правой половине				
	толстой кишки.				
	1.5. Другие типы полипов (гамартомные, липомы, лимфоидные				
	полипы, карциноидные опухоли).				
2. Морфологическая	Наиболее распространенная – Парижская эндоскопическая				
(по форме):	классификация, принятая в 2002 г. и используемая для описания				
	полипов в ЖКТ, включая толстую кишку [241].				
3. Размер:	 Миниатюрные полипы: ≤ 5 мм 				
	– Маленькие полипы: 6-9 мм				
	– Средние полипы: 10-19 мм				
	– Большие полипы: ≥ 20 мм				

Продолжение таблицы 4.1

4. По локализации:	4.1. Правая половина толстой кишки: включает слепую кишку,		
	восходящую ободочную кишку и правую половину поперечной		
	ободочной кишки.		
	4.2. Левая половина толстой кишки: включает левую половину		
	поперечной ободочной кишки, нисходящую ободочную кишку и		
	сигмовидную кишку.		
	4.3. Прямая кишка: самый нижний отдел толстой кишки.		

Детектирование полипов в колоноскопии с помощью СППВР является активно развивающейся областью исследований, направленной на повышение точности и эффективности ранней диагностики КРР. Использование ИИ в этом контексте призвано помочь эндоскопистам обнаруживать полипы, особенно небольшие и плоские, которые могут быть пропущены во время стандартной скрининговой процедуры [242-245]. Пропущенные полипы увеличивают риск интервального рака (рак, развивающийся между колоноскопиями). Продолжительные процедуры колоноскопии могут привести к усталости эндоскописта и снижению внимания, что увеличивает вероятность пропуска полипов. Следует отметить также, что процесс обнаружения и классификации полипов является отчасти субъективным, т.е. его результаты в значительной квалификации степени зависят ОТ опыта И врача-эндоскописта. Bce вышеперечисленное делает задачу разработки ИИ-решений для обнаружения колоректальных полипов актуальной и своевременной [246-253].

4.2 Описание используемых наборов изображений полипов

Рассмотрим сначала существующие открытые баз эндоскопических изображений и видеоданных, полученных с помощью эндоскопа в режиме визуализации в белом свете.

4.2.1 Стандартные базы для задач детектирования и сегментации

В открытом доступе находятся базы изображений Kvasir-SEG, ETIS-Larib и CVC-ClinicDB, которые содержат изображения колоноскопических исследований и содержат данные ручной разметки полипов; базы Kvasir, Nerthus и HyperKvasir [167, 254-256] – содержат изображения всего ЖКТ. Еще одна база KvasirCapsule [257] представляет собой 4 820 739 изображений и 118 видеопоследовательностей всего ЖКТ, полученных с помощью капсульной эндоскопии и содержит разметку 13 различных классов аномалий ЖКТ для задачи классификации изображений.

Для исследования использована самая распространенная открытая база эндоскопических изображений, полученных с помощью эндоскопа Olympus в режиме визуализации в белом свете, Kvasir-SEG [255, 258, 259]. Общее количество полипов в ней составляет 1 072 штуки. На рисунке 4.3 показаны изображения, бинарные маски и ограничивающие рамки для полипов из набора Kvasir-SEG.



Рисунок 4.3 – Примеры изображений из набора данных Kvasir-SEG:

а) исходные изображения; б) изображения с ограничивающими рамками; в) аннотированные бинарные маски

Изображения и их маски могут использоваться для решения задачи сегментации полипов, а ограничивающие рамки для решения задачи их детектирования. Разрешение изображений в этом наборе данных варьируется от 332×487 до 1920×1072 пикселей. Он включает изображения 700 «больших полипов» (>160×160 пикселей), 323 полипов «среднего размера» (>64×64 пикселей и ≤ 160×160 пикселей) и 48 «маленьких полипов» (≤64×64 пикселей).

Следует отметить, что такое число обучающих примеров, которое содержится в базе Kvasir-SEG на современном этапе развития нейросетевых технологий обнаружения объектов является недостаточным. С учетом этого необходимым этапом в построении алгоритма обнаружения полипов является этап аугментации данных, который должен производиться с учетом специфики эндоскопических изображений [167, 260].

4.2.2 Собственная база изображений полипов OnkoYar-SEG для задач детектирования и сегментации

В ходе проведения исследований собрана собственная база из 127 видеофайлов исследований и соответствующих шаблонов попиксельного выделения области полипов. Изображения для создания этого набора данных отобраны из видеопоследовательностей, предоставленных ЯОКОБ. Разрешение изображений составляло 626х532 пикселя. Примеры изображений приведены на рисунке 4.4. Из изображений перед разметкой удалена служебная информация, содержащая персональные данные [261-263].

Разметка с выделением полипов на изображениях осуществлялась при поддержке опытных врачей-эндоскопистов. Для выполнения данной задачи использовалась программа Labelme. Выполнялся процесс контроля качества разметки на основе методологии, описанной в главе 2. На каждое изображение наносилась разметка, которая указывала точную локализацию полипа, как показано на рисунке 4.4. Файл с разметкой сохранялся в формате JSON. Статистические данные по базе OnkoYar-SEG после проведения процедуры аугментации приведены в таблице 4.2.











Рисунок 4.4 – Примеры размеченных изображений полипов из базы OnkoYar-SEG



Обучающий набор						
Количество изображений, шт	227 582					
Количество полипов, шт	135 247					
Средняя высота полипов, пиксели	191,31					
Средняя ширина полипов, пиксели	185,66					
Валидаци	юнный набор					
Количество изображений, шт	40 853					
Количество полипов, шт	19 026					
Средняя высота полипов, пиксели	196,63					
Средняя ширина полипов, пиксели	193,41					
Тесто	Тестовый набор					
Количество изображений, шт	49 562					
Количество полипов, шт	14 238					
Средняя высота полипов, пиксели	145,60					
Средняя ширина полипов, пиксели	136,30					

Таблица 4.2 – Набор эндоскопических изображений OnkoYar-SEG

4.3 Разработка и исследование алгоритмов детектирования полипов

Исследования нейросетевых детекторов полипов происходило в три этапа. На первом из них, описанном в диссертации [167], проводилось сравнение нейросетевых алгоритмов обнаружения объектов (НАОО), построенных на основе семейства сверточных нейронных сетей YOLO5 с известными аналогами, опубликованными в [247, 251, 254]. На втором этапе [264, 265] проводилось исследование точности работы алгоритма, построенного на базе семейства сетей YOLO6 и YOLO7. Результаты первого и второго этапов, полученные на базе Кvasir-SEG, суммируются в таблице 4.3. Для алгоритмов на основе архитектур YOLO7 использовались следующие параметры обучения: размер батча (последовательности изображений) составлял 16 изображений, в качестве алгоритма оптимизации использовался алгоритм Adam (Adaptive Moment Estimation) с начальной скоростью обучения – 10^{-3} и общим числом эпох обучения – 250.

Для оценивания качества работы алгоритмов использовались метрики средней точности (*mAP*), рассчитанные при значениях порога *IoU* 0,25; 0,50 и 0,75. Они обозначены как *mAP*@0,25, *mAP*@0,50 и *mAP*@0,75 соответственно. Кроме того, для комплексной оценки по множеству порогов *IoU* использовалась метрика mAP@[0,25..0,75], которая рассчитывается как среднее 11-ти значений *mAP*, рассчитанных при 11 различных порогах *IoU* в пределах от 0,25 до 0,75 с шагом 0,05.

Алгоритм	mAP@[0,250,75]	mAP@0,25	mAP@0,5	,5 <i>mAP</i> @0,75			
Известные аналоги							
EfficientDet-D0	47,6	68,5	50,5	22,8			
Faster R-CNN	78,7	89,5	84,2	56,6			
YOLOv3+spp	81,0	88,6	85,3	75,9			
YOLOv4	85,1	91,2	82,3	75,9			
ColonSegNet	80,0	90,0	81,7	67,1			
Исследованные алгоритмы							
YOLOR	86,8	90,2	90,2 87,2				
YOLOv6	85,2	90,0	87,1	84,1			
YOLOv7	87,6	91,5	88,3	80,4			
YOLOv7X	88,9	92,4	89,9	85,2			

Таблица 4.3 – Тестирование алгоритмов детектирования полипов

Среди представленных алгоритмов детектирования лучшие результаты показал алгоритм на основе архитектуры YOLO7X. Усредненная по диапазону порогов метрика mAP@[0,25..0,75] для него составляет 88,9. Это говорит о высокой способности детектирования объектов в широком диапазоне порогов детектирования. Заметим, что методология построения тестовой базы для расчета метрик в таблице 4.3 отличалась от приведенной в работе [167], что не позволяет полностью корректно сравнить результаты. Применяемая в настоящей работе методология описана в разделе 3 при тестировании алгоритмов детектирования КСК.

На 3-м этапе исследования проводилось сравнение результатов алгоритмов, построенных на базе сети YOLO7 и трансформерной архитектуры D-FINE-L [123]. Рассмотрим его более подробно. Исследование проводилось на собственной базе изображений OnkoYar-SEG. В процессе обучения глубоких моделей МО использован следующий набор процедур аугментаций:

- Случайные яркостные изменения. Меняется яркость изображения в пределах 12,5% и контраст в пределах 50%. Также в цветовом пространстве HSV меняется тон и насыщенность в пределах 5% и 50%, соответственно.
- Случайные изменения масштаба в пределах 20% в большую и меньшую сторону. При уменьшении оставшиеся пустые области заполняются нулями.
- Случайная обрезка изображения. Остается до 80% от оригинала. Если размер ограничивающей рамки в результате обрезки получился меньше 2х2 пикселей, то рамка удаляется.
- Зеркальные горизонтальные и вертикальные отражения.
- Случайные повороты на 180 градусов.
- Случайные повороты на малые углы до 10 градусов.

Все аугментации делались в режиме реального времени с вероятностью 0,5. Такая процедура не увеличивает размер используемого датасета.

Особенность эндоскопических изображений заключается в том, что одновременно полипов на нем, как правило, один или (иногда) несколько единиц. В связи с этим, для увеличения количества ограничивающих рамок в процессе обучения использовались аугментации типа «мозаика» и «смешивание», которые отключались на последние 15 эпох:

- Мозаика. Выбираются случайные 4 изображения. Для каждого из них осуществляется уменьшение размера до 320х320 пикселей, затем случайная обрезка до 50%. Далее в случайном порядке собирается мозаика. Модель получает на входе значительно большее количество рамок разных размеров, что положительно сказывается на итоговом качестве работы.
- Смешивание. Происходит смешивание случайных изображений друг с другом на уровне не только пикселей, но и эталонных рамок (аналогия: мир не черно-

белый, а есть множество оттенков серого). Модель в процессе обучения видит не только «абсолютные» полипы и фон, но и что-то среднее, то есть лучше обучается находить разницу между ними.

На входе НС принимают квадратные изображения размером 640х640 пикселей. В силу того, что имеющиеся изображения меньше и имеют размер 532х626 пикселей – дополнительно проводилась процедура «паддинга», т.е. заполнения недостающих областей справа и снизу нулями до необходимого размера.

Структура трансформерной сети D-FINE-L приведена на рисунке 4.5.



Рисунок 4.5 – Архитектура трансформера D-FINE-L

Базовый блок (backbone) извлекает признаки из входного изображения. Его выходом является карта признаков с пространственным разрешением 20×20 (для входа 640×640 пикселей). Кодировщик выполняет функцию агрегации глобального контекста изображения. Он представляет собой трансформерный блок, обрабатывающий признаки. Слои декодера генерируют объектные запросы и начальные предсказания bbox. Модуль FDR (Fine-grained Distribution Refinement) итеративно уточняет распределения смещений границ bbox. Это ключевой преобразующий компонент D-FINE-L, регрессию bbox В итеративную оптимизацию вероятностных распределений. Модуль GO-LSD (Global Optimal Localization Self-Distillation) проводит самодистилляцию знаний о локализации от более глубоких слоев к менее глубоким. На выходе получаем финальные предсказания классов и bbox [123].

Результаты тестирования приведены в таблице 4.4. Метрика Precision у алгоритма D-FINE-L ниже на 0,025, а Recall выше на 0,04, то есть трансформер создает большее количество ложных срабатываний, при этом пропускает меньше полипов. В среднем для медицинских приложений можно считать, что такой подход работает лучше, чем YOLO. Другим важным преимуществом модели D-FINE-L является ее лицензионная открытость, что позволяет использовать ее в конечных аппаратно-программных комплексов для СППВР [123].

Метрика	YOLO7	D-FINE-L
mAP@0,50	0,833	0,840
Р	0,764	0,739
R	0,810	0,850
F1	0,786	0,791

Таблица 4.4 – Сравнение результатов YOLO7 и D-FINE-L на базе OnkoYar-SEG

Графики зависимость метрик Precision и Recall от порога и кривые «Precision-Recall» для сетей YOLO7 и D-FINE-L приведены на рисунке 4.6. Зависимости Precision и Recall от порога анализируются одновременно и используются для подбора порога и баланса между ошибками 1-го и 2-го рода. Общее качество модели определяется по кривой «Precision-Recall». Площадь под ней есть *mAP*@0,5, то есть кривая для лучшей модели (в исследуемом случае D-FINE-L) должна в среднем идти чуть выше [265].









в)

Рисунок 4.6 – Исследование алгоритмов YOLO7 и D-FINE-L на базе OnkoYar-SEG: a) зависимость Precision от порога; б) зависимость Recall от порога; в) зависимость Precision от Recall

4.4 Разработка и исследование алгоритмов сегментации полипов

Одним из наиболее часто используемых подходов к обнаружению объектов является попиксельное детектирование (сегментация). Изображение разбивается на области, представляющие различные классы объектов, в рассматриваемой задаче – это полипы и фон. Существующие подходы можно условно разделить на классические методы и методы с использованием НС [266-269]. Классические методы обычно уступают место нейросетевым алгоритмам, которые показывают более высокие точность и эффективность в последние несколько лет. С другой стороны, результаты классических методов хорошо интерпретируются, а нейросетевых методов представляют собой «черный ящик». Чтобы подчеркнуть это, консервативно настроенные специалисты иногда используют термин «физически обоснованные» признаки. С учетом важности рассматриваемой медицинской задачи требуется реализовать и классические, и нейросетевые методы, а также провести сравнительный анализ результатов [261-262].

4.4.1 Исследование классических методов сегментации

Одним из наиболее известных классических алгоритмов сегментации является метод кластеризации k-means [270]. Он минимизирует суммарное квадратичное отклонение точек, принадлежащих кластерам, от их центроидов. Метод k-means делит множество пикселей изображений на k кластеров, стремясь к тому, чтобы точки внутри кластера были максимально близки к его центроиду. Процесс кластеризации осуществляется путем итеративного перемещения центроидов до достижения их стабилизации.

Исследования по практическому применению метода k-means при разных значениях числа кластеров k показали его низкую эффективность при работе с реальными колоноскопическими изображениями полипов. Результаты сегментации по методу k-means приведены на рисунке 4.7.



Рисунок 4.7 – Результаты алгоритма сегментации на базе метода k-means: кадры колоноскопического исследования (а, б); бинарные шаблоны (в, г); результаты сегментации при k=3 (д, е); результаты сегментации при k=4 (ж, з)

Другим популярным классическим методом сегментации является метод Оцу (Otsu), основанный на анализе гистограммы интенсивности пикселей изображения [271]. Гистограмма представляет собой графическое отображение распределения интенсивностей пикселей, где ось абсцисс соответствует значениям интенсивности в диапазоне [0, 255], а ось ординат – количеству пикселей для каждой интенсивности. Метод Оцу представляет пороговое значение на гистограмме, разделяющее все пиксели на 2 класса. Для этого минимизируется внутриклассовая дисперсия, что означает, что пиксели внутри каждого класса должны быть максимально однородными. Таким образом, пороговое значение обеспечивает наиболее четкому выделению фона и объекта.

Однако при практическом применении данного алгоритма к задаче сегментации полипов на колоноскопических изображениях результаты оказываются неудовлетворительными (рисунок 4.8).



166



д)

г)

Рисунок 4.8 – Примеры сегментации полипов методом Оцу: полученные этим методом маски (а-в); результаты наложения масок на исходные изображения (г-е)

e)

Проведенный анализ подтверждает несостоятельность классических алгоритмов в контексте сегментации области полипов. Это обусловлено следующими основными факторами [262]:

- Эндоскопические изображения характеризуются наличием различных шумов, что существенно затрудняет работу классических алгоритмов сегментации, которые демонстрируют высокую чувствительность к подобным артефактам.
- Классические методы сегментации не принимают во внимание пространственные и контекстуальные особенности изображений, в то время как при анализе сложных изображений важно учитывать форму и границы объектов, а также их взаиморасположение в контексте общего изображения.
- Традиционные алгоритмы опираются на цветовую информацию изображений, однако полипы могут обладать цветом и текстурой, схожей с окружающими их обычными тканями толстого кишечника, что приводит к некорректной сегментации.

4.4.2 Исследование нейросетевых алгоритмов сегментации

Современные HC, в частности архитектуры на основе U-Net и её модификаций (ResUNet++, PraNet, ColonSegNet), стали де-факто стандартом для семантической и инстанс-сегментации медицинских изображений благодаря их способности автоматически выявлять сложные морфологические паттерны, не доступные классическим методам. Ключевыми вопросами построения нейросетевых алгоритмов являются [261, 268, 269]:

- Устойчивость к вариативности данных адаптация к изменениям освещенности, слизи, кровотечениям за счет аугментаций (MixUp, CutMix) и модифицированных сверток.
- Точность границ механизмы внимания минимизируют ошибки типа ложных отрицаний на границах опухолей.
- Возможность оптимизации для работы в режиме реального времени в системах СППВР (fps>30) для интраоперационной навигации.

Сохраняются и ряд проблем: зависимость от размеченных датасетов, сложность интерпретируемости предсказаний и риски ложных срабатываний при дисплазиях низкой степени [266-267].

В ходе исследования проведено обучение и тестирование нейросетевых алгоритмов на базе следующих архитектур: SSformer-S, SSformer-L, U-Net, U-Net++, DUCK-Net-17, DUCK-Net-34, Meta-Polyp [272-277]. Для обучения и тестирования использовались описанные выше базы изображений Kvasir-SEG и OnkoYar-SEG. Наборы делились на обучающую, валидационную и тестовую выборки.

В проведенных программных экспериментах использовались следующие параметры обучения: функция потерь ВСЕ; оптимизатор Adam с параметрами $\beta 1 = 0.9$; $\beta 2 = 0.999$; $\varepsilon = 10^{-8}$; число эпох обучения – 200; постоянная скорость обучения – 0,001. Фиксация этих гиперпараметров позволяет судить о возможности использования полученных результатов для оценки эффективности каждой из рассматриваемых нейросетевых архитектур.

Для сравнения качества работы алгоритмов сегментации использовался коэффициент Дайса (*Dice*) [56, 272]. Он принимает значения от 0 до 1 и определяется как отношение двойного размера пересечения множеств (например, A и B) к их сумме:

$$Dice = \frac{2 \cdot |A \cap B|}{|A| + |B|} = \frac{2 \cdot TP}{2 \cdot TP + FP + FN},$$

где A – множество пикселей, предсказанных алгоритмом; B – множество пикселей эталонной маски; $|A \cap B|$ – количество элементов, присутствующих в обоих множествах; TP (true positive) – истинно положительный результат, количество правильно обнаруженных пикселей; FP (false positive) – ложноположительный результат, показывающий количество пикселей, определенных алгоритмом ошибочно как объект патологии, FN (false negative) – ложноотрицательный результат, показывающий количество пикселей, относящихся к объектам интереса, хотя алгоритм их к таковым не относит.

В качестве основы итогового алгоритма сегментации взята архитектура Meta-Polyp (MP), показавшая наилучшие результаты тестирования и представленная на рисунке 4.9a. Вход сети имеет размерность $X \in R^{Width \times Height \times 3}$, $X \in R^{\frac{W}{2^{i+1}} \times \frac{H}{2^{i+1}} \times F_i}$. где $F_i \in \{64; 128; 320; 512\}; i \in \{1; 2; 3; 4\}$ – фильтр на шаге *i*-го этапа кодера и декодера. На этапе декодера признак декодируется дважды. Затем он объединяется с признаком на *i*+2-м шаге для улучшения и повышения частоты дискретизации. Декодер генерирует маску и преобразует карту признаков в 1-й фильтр. Проведена модификация сверточно-трансформерного блока (СТБ), который выполняет последовательность операций, показанных на рисунке 4.96. Сеть С модифицированным СТБ-блоком получила название Modified Meta-Polyp (MMP) и представляет собой архитектуру на базе U-Net, но заточенную под конкретный медицинский класс объектов для сегментации [262].



Рисунок 4.9 – Архитектура нейронной сети Meta-Polyp: общая схема сети (a); схема модифицированного СТБ-блока (б)

Результаты тестирования нейросетевых алгоритмов по метрике Дайса для двух рассматриваемых наборов изображений приведены в таблице 4.5. Наилучшие результаты на наборе Kvasir-SEG достигнуты с применением архитектуры SSformer-S, которая продемонстрировала коэффициент Дайса 0,96. Наихудшие результаты получены с использованием архитектуры U-Net++, коэффициент Дайса которой составил 0,81.

Набор	Нейросетевая архитектура							
	CCformon C	CCformon I	U-	U-	DUCK-	DUCK-	MD	MMD
данных	551011161-5	SSIOIMEI-L	Net	Net++	Net-17	Net-34	1411	1011011
Kvasir-	0.060	0.030	0.812	0.810	0.801	0.007	0.052	0.057
SEG	0,900	0,939	0,012	0,010	0,891	0,907	0,952	0,957
OnkoYar-	0.882	0.802	0.715	0.702	0.711	0.757	0.855	0.002
SEG	0,002	0,892	0,715	0,702	0,711	0,757	0,855	0,702
OnkoYar-								
SEG +	0.008	0.007	0.750	0.751	0.805	0.822	0.022	0.045
Kvasir-	0,908	0,907	0,739	0,751	0,805	0,833	0,935	0,943
SEG								
Среднее	0.917	0.913	0.762	0.751	0.754	0.832	0.913	0.935
значение	0,917	0,715	0,702	0,751	0,754	0,052	0,715	0,755

Таблица 4.5 – Значения коэффициента Дайса

Дополнительно проведены тесты на собственной базе данных изображений OnkoYar-SEG. Здесь наилучший результат достигнут с использованием архитектуры MMP с коэффициентом Дайса 0,902. Наихудший результат снова получен с использованием архитектуры U-Net++, где коэффициент Дайса составил всего 0,702.

Также проведены тесты на объединенной базе данных, включающей 200 изображений из базы OnkoYar-SEG и 100 изображений из базы Kvasir-SEG, не входившие в первые два набора. Наилучший результат достигнут с использованием архитектуры Meta-Polyp с коэффициентом Дайса 0,945. Среднее значение коэффициента Дайса для трех экспериментов получилось наибольшим также при использовании сети MMP.

На основании полученных результатов можно сделать вывод, что архитектуры SSformer-S и SSformer-L, а также MP и MMP эффективно справляются с задачей сегментации полипов. Они демонстрируют высокую точность в определении границ и контуров полипов, превосходя классическую для задач сегментации объектов на изображениях архитектуру U-Net в среднем на 19-27% по

коэффициенту Дайса [261-262]. Вопросы вычислительной сложности рассмотренных алгоритмов сегментации будут рассмотрены в главе 6.

4.5. Краткие выводы

Исследования, проведенные в данной главе, позволяют сделать следующие выводы:

- Полипы толстой кишки могут привести к развитию колоректального рака.
 Своевременное обнаружение и удаление полипов критически важно для профилактики рака, учитывая риск малигнизации некоторых типов полипов (особенно аденоматозных) и значительный процент (около 20%) пропусков при проведении скрининговой колоноскопии.
- Разработка алгоритмов для детекции и сегментации полипов является актуальным направлением, способным повысить точность диагностики, снизить пропуск полипов (особенно мелких и плоских), минимизировать влияние усталости врача, что в итоге снизит риск интервального рака и вызванную им смертность.
- Нейросетевые алгоритмы, особенно основанные на архитектурах семейства YOLO7 и на базе трансформерных архитектур, демонстрируют высокую эффективность в автоматическом детектировании полипов на эндоскопических изображениях. Среди рассмотренных алгоритмов детектирования лучшие результаты показал алгоритм на основе открытой трансформерной архитектуры D-FINE-L. Метрика mAP@[0,5] – для него составляет 0,84.
- Традиционные алгоритмы (k-means, Оцу) показали низкую эффективность в задаче сегментации полипов из-за чувствительности к шумам, игнорирования контекста/формы и схожести цвета/текстуры полипов со здоровой тканью.
- Современные нейросетевые архитектуры для сегментации (SSformer-S/L, MMP) значительно превосходят как классические методы, так и традиционные нейросетевые U-Net/U-Net++ (на 19-27%), достигая коэффициента Дайса до 0,96 на базе Kvasir-SEG и 0,945 на объединенной базе данных.

ГЛАВА 5 РАЗРАБОТКА И ИССЛЕДОВАНИЕ АЛГОРИТМОВ ДЕТЕКТИРОВАНИЯ И КЛАССИФИКАЦИИ ПАТОЛОГИЙ ЖЕЛУДКА

5.1 Актуальность задачи

Одним из наиболее распространенных видов онкологических заболеваний является рак желудка. До 1980-х годов он занимал первое место как в структуре заболеваемости раком, так и среди причин смертности от онкологических заболеваний [278]. В настоящее время данный вид рака по-прежнему находится в пятерке наиболее распространенных видов онкозаболеваний с высокими показателями смертности, однако существуют некоторые тенденции к улучшению сложившейся ситуации [279]. Высокая заболеваемость раком желудка отмечается в Восточной Азии (Япония, Корея), Восточной Европе и Южной Америке. В России ежегодно регистрируется около 35 тыс. новых случаев. Рак желудка представляет собой злокачественное новообразование, развивающееся из клеток слизистой оболочки желудка. Наиболее распространённая форма аденокарцинома (90-95% случаев), которая возникает из железистых клеток. Основные причины возникновения рака желудка, представлены на рисунке 5.1.



Рисунок 5.1 – Основные причины возникновения рака желудка

Бактерия Helicobacter pylori вызывает воспаление, повреждает ДНК клеток и запускает предраковые изменения (метаплазия, дисплазия). Ее появление в организме увеличивает риск развития рака в 3-6 раз [280]. К вредным пищевым привычкам относят избыток солёной, копчёной, маринованной пищи, низкое потребление овощей, фруктов, клетчатки, злоупотребление красным мясом и продуктами с консервантами. Риск возникновения рака желудка повышается в 2-3 раза при наличии аналогичного заболевания у близких родственников. Также его возникновение может быть вызвано другими хроническими заболеваниями желудка: атрофический гастрит, пернициозная анемия, кишечная метаплазия, полипы желудка (адематозные типы). Дополнительными факторами риска также являются вредные привычки (курение, алкоголь), ожирение, возраст от 50 лет, мужской пол и длительная работа на вредных производствах [281].

Обычно рак желудка выявляется с помощью комплекса инструментальных и лабораторных методов (рисунок 5.2). Среди причин снижения уровня смертности от него важное место занимает повышение качества диагностики данного заболевания, связанное с появлением и постепенным распространением, начиная с 60-70-х годов прошлого века, эндоскопической техники. Современные эндоскопы оснащены технологиями, повышающими точность визуализации предраковых изменений и микроопухолей. Тем не менее, стоит отметить, что ряд вопросов, связанных с раком желудка, в том числе вопросы его ранней диагностики, все еще далеки от окончательного решения [279-281].



Рисунок 5.2 – Методы диагностики рака желудка

На сегодняшний день эндоскопическое исследование играет значительную роль в диагностике и лечении рака желудка [32]. Его обнаружение на начальной

стадии требует высокой квалификации врача-эндоскописта. Особенно сложной задачей является дифференциальная диагностика раннего рака желудка с прочими дегенеративными изменениями слизистой оболочки. Она имеет наиболее сложную микроархитектонику, поскольку в отличие от других отделов ЖКТ, при требуется равной эндоскопическом осмотре В степени осмотр как микрососудистого рисунка, так и рисунка микроструктуры поверхности эпителия [281]. Более того, слизистая оболочка желудка имеет три типа желез, имеющих различное строение, а большой спектр возможных патологических процессов определяет разнообразие изменений в ее структуре. С учетом сложности наблюдаемого объекта увеличительная И узкоспектральная эндоскопия, значительно повышающие уровень диагностики, остаются методиками экспертного уровня, а обучение начинающих специалистов в данной области представляет собой длительный процесс, что замедляет внедрение новых технологий в широкую клиническую практику [21].

Таким образом, задача повышения достоверности диагностики при проведении эндоскопических исследований желудка является актуальной. Применение СППВР позволит эффективней применять самые современные методики диагностики и решать сложные клинические задачи в условиях ограниченности временных ресурсов, недостатка опыта, в отсутствии возможности привлечения врачей экспертного уровня.

Технологии ИИ усиливают эндоскопическую диагностику, повышая точность выявления предраковых состояний (на 20-30% по некоторым оценкам), позволяя выявлять ранние формы рака желудка (опухоли размером до 5 мм) [282, 283]. К основным направлениям применения таких технологий относятся (рисунок 5.3): видеоизображений В реальном времени, улучшение анализ визуально воспринимаемого качества изображения, прогноз развития опухоли и предраковых состояний, мультимодальная интеграция с результатами гистологических исследований и др.

Сегментация опухолей состоит в попиксельном построении соответствующей маски на изображении для возможного планирования дальнейшего хирургического

вмешательства. Мультимодальные модели совмещают данные с эндоскопии, КТ и биопсии для оценки риска распространения рака, предсказывая метастазы в лимфоузлах. Также НС могут использоваться для анализа гистологических срезов, классифицируя клетки и определяя степень злокачественности. Их использование сокращает время анализа биопсии до 40% [95, 168-170].



Рисунок 5.3 – Возможные применения нейронных сетей при эндоскопической диагностике рака желудка

Как показывает опыт практического применения глубоких HC, они способны выявлять, в том числе предраковые состояния (метаплазия, дисплазия), которые часто пропускаются при визуальном осмотре врачами-эндоскопистами. В ряде случаев использование систем ИИ устраняет вариабельность между мнениями разных экспертов. Кроме того, СППВР позволяют решать задачу обнаружения микроопухолей, в том числе в узкоспектральных режимах, выделяя аномальные сосудистые сети и текстуру слизистой [284-286].

К проблемам применения HC в задаче обнаружения рака желудка обычно относят те же факторы, что отмечались выше для колоноскопии: нехватка обучающих данных, необходимость в интерпретации работы HC, зависимость работы системы ИИ от качества разметки видеоданных и др. В отличие от колоноскопических исследований для исследований желудка на сегодняшний день не существует широко известных открытых верифицированных баз изображений с экспертной аннотацией [287-288].

Исследования в области построения СППВР для эндоскопических исследований желудка проводятся в коллаборации ЯрГУ с ЯОКОБ уже более

10 лет. На первом этапе они проводились под руководством австрийского профессора Герберта Эдельсбруннера и заключались в использовании новейших математических методов, основанных на аппарате вычислительной алгебраической топологии [283]. Такой подход отличался от традиционного на тот момент использования методов классического машинного обучения, опирающихся на многочисленные статистические предположения, которые часто нарушаются на реальных видеоизображениях желудка. Была предложена система устойчивых топологических признаков, описывающих локальную структуру эндоскопического изображения (центральные моменты диаграммы персистентности). Однако практическое тестирование комплекса программ, реализующих алгоритмы сегментации и классификации изображений желудка, показало большое число ложных срабатываний системы, ЧТО делало ее постоянное клиническое использование невозможным.

На втором этапе работ под руководством Хрящева В.В. в 2017-2022 гг. использовались появившиеся на тот момент глубокие НС первого поколения (SSD, RetinaNet, EfficientDet), показавшие хорошие результаты при тестировании на ограниченных базах статических изображений, но допускавшие ошибки при тестировании на реальном видеопотоке [289-299].

На третьем этапе работ также под руководством Хрящева В.В. в 2023-2025 гг. для детектирования и классификации патологий использовались глубокие НС второго поколения (YOLO, трансформерные архитектуры), что позволило улучшить точность результатов как на изображениях, так и на видеоданных [300-303]. Их дальнейшее совершенствование в совокупности с увеличением объемов размеченных баз изображений позволяет надеяться на широкое клиническое использование таких СППВР для исследований желудка. Использование лицензионно открытых нейросетевых архитектур также будет усиливать проекты по импортозамещению эндоскопического оборудования.

5.2 Описание созданной базы изображений желудка с патологиями

Для создания базы изображений использовано 127 видеопоследовательностей, на которых представлены исследования желудка, проведенные в 2020-2021 гг. в эндоскопическом отделении ЯОКОБ. При сборе базы использовались видеоданные, полученные при помощи эндоскопов GIF-HQ190 и GIF-Q180 системы OLYMPUS EXERA III.

Использовались следующие минимальные требования к собираемым эндоскопическим изображениям [304]:

- 1) Минимальное разрешение 624х528 пикселей.
- Формат файла изображения ВМР (без сжатия), 24 бита/пиксель, цветовое пространство – RGB.
- Входной контроль качества изображения отсутствие сильной размытости, артефактов сжатия, сильных шумов на изображениях.
- 4) В обучающую выборку должны входить не менее 1000 изображений на каждый из классов (желательно приближаться к цифре 10 000 изображений на класс): нормальный тип слизистой, раковый тип слизистой, ранний рак, другие патологии.

Типовая диагностическая процедура выглядит следующим образом: врачэндоскопист с помощью видеогастроскопа анализирует состояние слизистой оболочки желудка пациента на предмет наличия доброкачественных и злокачественных патологий. При прохождении эндоскопом участка, потенциально содержащего область интереса, врач просматривает его в двух режимах: без увеличения и с увеличением. Второй режим позволяет принять решение по особо сложным случаям путем изучения слизистой оболочки желудка, в том числе ее микрососудистого рисунка, который также может отличаться от нормального при наличии патологии того или иного вида. Таким образом, для того чтобы СППВР на базе методов ИИ могла успешно находить раковые патологии, необходимо собрать базу изображений, содержащую примеры патологий как в режиме с увеличением, так и в режиме без него. Важнейшим моментом для построения эффективного нейросетевого классификатора является четкая формализация классов наблюдаемых объектов [303-304].

5.2.1 Классы патологий желудка

Среди наиболее часто встречающихся патологий слизистой оболочки желудка врачи определяют следующие основные классы (рисунок 5.4) [31, 32, 278]:

- ранний рак;
- рак тела желудка;
- полип;
- кишечная метаплазия слизистой оболочки желудка;
- ксантома;
- эрозии;
- гиперплазия;
- аденома;
- язва;
- нейроэндокринные опухоли.



Рисунок 5.4 – Примеры изображений слизистой желудка: нормальное состояние слизистой (аб); рак тела желудка (в-г); ранний рак в режиме NBI с увеличением (д-е); полипы желудка (з); ксантома в режиме NBI с увеличением (ж)

Отдельный класс при проведенном исследовании составляло также нормальное состояние слизистой оболочки желудка без патологий.

Важно, чтобы разрабатываемая система обнаружения онкологических заболеваний желудка обучалась как на большом количестве изображений с раковой патологией (положительных примерах), так и на большом количестве отрицательных примеров, где не представлена патология, либо патология является иной (отрицательные примеры).

Созданы следующие группы изображений:

- «рак», в состав которой вошли изображения с разными видами злокачественных раковых образований (ранний рак желудка, рак тела желудка);
- «иная патология», в состав которой вошли изображения, представляющие нераковые доброкачественные патологии (ксантома, кишечная метаплазия, эрозия, аденома, полип, язва, гиперплазия, карциноиды);
- изображения нормальной слизистой оболочки желудка без патологий.

В ходе проведения эндоскопических исследований врач осуществляет фотофиксацию патологических областей, и полученные снимки сохраняются в базе данных ЭМИС совместно с видеоданными исследования и другой информацией о пациенте. Таким образом, из этих данных можно сформировать основу для базы изображений патологий желудка. Однако количество статических изображений, соответствующих одному исследованию, при таком способе сбора базы очень мало (в среднем, от 5 до 10), несмотря на их высокое качество. Более эффективным способом с точки зрения сбора базы является извлечение изображений непосредственно из видеоданных. Достоинства данного подхода заключаются в следующем [299]:

- Потенциально большой объем видеоизображений патологий, выдаваемых с высокой кадровой частотой.
- Одна и та же область может быть представлена с разных ракурсов, что повышает устойчивость системы к различным ситуациям расположения патологии.

Патология в кадре увеличивается в размерах по мере приближения к ней эндоскопа, когда речь идет о режиме без увеличения. Акцентирование внимания специалиста на этапе приближения эндоскопа к патологии является важным качеством разрабатываемой системы. Наполнение обучающей базы разнообразными примерами с точки зрения расположения области интереса повысит способность ее обнаружения.

Недостатком такого способа сбора базы данных является более низкое качество изображений при их извлечении из видеоданных. В данной работе происходило извлечение каждого 10-го кадра из видео при кадровой частоте 25 кадров/секунда. Для большого количества полученных изображений можно отсутствие четких особенностей, отметить размытости И ПО которым врачи-эндоскописты определяют вид патологии. Особенно данное явление ухудшает информационную составляющую снимков, полученных с эндоскопа в режиме с оптическим увеличением. В случае же, когда формирование снимков происходит врачом в процессе эндоскопического исследования, в базе данных сохраняются кадры, обладающие лучшим качеством.

Другим недостатком такой процедуры значительная является ee Процедуре областями трудозатратность. извлечения с интереса ИЗ видеопоследовательности предшествовал длительный процесс создания описаний к видеоданным врачами-эндоскопистами: тип патологии, описание патологии и ее признаков, внешних отличительных временные интервалы видеопоследовательности, наличие/отсутствие режима увеличения при проведении исследования, в пределах которых патология присутствует в кадре. Снижение трудозатрат врачей-эндоскопистов происходило с использованием методики «физической разметки» с помощью педалей, описанной в главе 2.

Тем не менее, данный подход позволяет значительно увеличить количество изображений, приходящихся на каждую патологию, что является ключевым фактором повышения производительности систем на основе методов глубокого машинного обучения [299]. Итоговое количество статических изображений зависит от длительности исследования патологической области на видеоданных.
Так, минимальное количество изображений, полученное для одного участка патологии на видеозаписи, составило 2 шт., а максимальное – 424 шт.

Итоговое количество образцов в базе эндоскопических изображений с раком желудка, собранных путем извлечения из видеоданных, составило 1 357 шт. Соотношение между разными типами собранных раковых патологий представлено в таблице 5.1. Разрешение исходных кадров до обработки составляло 760х576 пикселей [304].

Таблица 5.1 – Статистика по изображениям с раком желудка

Тип патологии	Количество изображений, шт.
Ранний рак желудка	346
Рак тела желудка	1 011
ИТОГО	1 357

Итоговое количество изображений, содержащих нераковые патологии и нормальную слизистую в сумме составило 12 778 шт. Больше всего собрано изображений нормальной слизистой оболочки желудка, так как именно она занимает большую часть видеоданных эндоскопического исследования даже в том случае, если у пациента обнаруживается патология желудка. Наименьшее количество изображений пришлось на доброкачественную патологию типа «карциноид» – 29 шт. Сведения о количественном соотношении изображений различных типов классов отображены в таблице 5.2.

Классы	Количество изображений, шт.
Полип	538
Кишечная метаплазия	645
Ксантома	171
Эрозия	275
Гиперплазия	34
Аденома	53
Язва	50
Карциноид	29
Нормальная слизистая оболочка желудка	10 953
ИТОГО	12 778

Таблица 5.2 – Иные классы, представленные в базе эндоскопических изображений

В эндоскопических исследованиях, из которых извлекались кадры для формирования обучающей базы изображений, использовались режимы без оптического увеличения, а также с оптическим увеличением в 1,5 и 65 раз. Использование изображений, полученных с использованием увеличительной эндоскопии для обеих групп изображений (таблица 5.3), позволяет обучить систему способности анализировать данные на выходе эндоскопа независимо от наличия режима оптического увеличения.

Таблица 5.3 – Количество изображений для групп «рак» и «иная патология», полученных в режимах с увеличением и без увеличения

Класс изображений	Увеличение	Количество изображений, шт.
«Рак»	Без увеличения	758
	С увеличением	599
«Иная патология»	Без увеличения	10 359
	С увеличением	2 419

Соотношение количества изображений, полученных с помощью разных эндоскопов, представлено в таблице 5.4. Заметно, что количество изображений с

раком, полученное с использованием обоих эндоскопов, примерно равно в случае использования режима без увеличения.

Модель эндоскопа	«Рак» без увеличения, шт.	«Рак» с увеличением, шт.	«Иная патология» без увеличения, шт.	«Иная патология» с увеличением,
				шт.
OLYMPUS				
EXERA III	397	63	7 123	1 106
GIF-Q180				
OLYMPUS				
EXERA III	361	536	3 216	1 341
GIF-HQ190				

Таблица 5.4 – Количество изображений для групп «рак» и «иная патология» в базе эндоскопических изображений в зависимости от использованной модели эндоскопа

5.2.2 Фильтрация изображений перед этапом разметки

Изображения, получаемые в ходе эндоскопического исследования желудка, имеют ряд особенностей [95, 283, 290]. Во-первых, для них зачастую характерно размытие, что связано с движением эндоскопа внутри полости органа, а также с сокращениями мышечных волокон стенок желудка (перистальтикой), как показано рисунке 5.5а-г. При размытии эндоскопического изображения без на использования оптического увеличения теряется часть информации о структуре ткани, однако в большинстве случаев область с патологией в кадре может быть обнаружена. Анализ изображений в режиме с увеличением показывает, что при размытии практически полностью теряется информация о микроархитектонике слизистой оболочки и структуре сосудистого рисунка. Таким образом, изображения в режиме с увеличением при наличии размытия становятся практически неинформативными для диагностических мероприятий. Вследствие этого из собранной базы подобные изображения удалялись. Для их автоматической фильтрации в том числе использовался модифицированный алгоритм оценки качества, разработанный в главе 3 настоящей работы.

Другой сложностью при анализе изображений желудка является наличие большого числа бликов, появление которых связано с практически соосным расположением объектива и источника света (рисунок 5.5 д-е).





Рисунок 5.5 – Изображения с размытием в режиме без увеличения (а-б); с размытием в режиме с 65-кратным увеличением (в-г); с бликами (д-е)

Решение об удалении изображений с бликами принималось на основании возможности однозначно выделить патологический участок на эндоскопическом изображении врачом-эндоскопистом.

Из базы данных удалялись также изображения с прочими артефактами, такими как значительные искажения, практически полное отсутствие информации в кадре, наличие в кадре посторонних предметов, закрывающих обзор и т.п. Кроме того, на каждом изображении присутствует служебная информация о враче-эндоскописте, проводящем исследование, и пациенте. Такая текстовая информация на изображении закрывалась черными прямоугольниками. Разметка эндоскопических изображений желудка проводилась в соответствии с методикой, рассмотренной в главе 2 настоящей работы.

5.2.3 Разметка изображений желудка

Аннотирование базы эндоскопических изображений осуществлялось специалистами по разметке, не имеющими медицинского образования, согласно рекомендациям и описанию сцен, предоставленному врачами-эндоскопистами. При таком подходе неизбежно могут возникать ошибки в разметке, связанные с некорректным положением обрамляющей рамки для конкретной патологии. Решение данной проблемы выполнялось врачами на этапе валидации размеченных данных.

Выделение патологий на изображениях желудка связано, в первую очередь, с определением положения демаркационной линии, которая разграничивает здоровые ткани и ткани с патологической анатомией. В общем случае демаркационная линия формирует сложный полигон. Для упрощения задачи в ряде случаев сложные полигоны описывались прямоугольниками (рисунок 5.6). На них демаркационной линией зеленого цвета ограничивался участок тканей, содержащий патологию. Далее участок с патологией описывался красным обрамляющим прямоугольником. Кроме того, использование прямоугольных областей для выделения патологических участков тканей удобно с точки зрения оценки качества работы большинства детекторов патологий, так как детектирование осуществляется также прямоугольником [181, 183].

В случае использования алгоритмов сегментации (попиксельного выделения опухоли) использовалась трудоемкая разметка консилиумом врачей, описанная в главе 2. Ее использование позволяет точно описать демаркационную линию на патологии и потенциально разработать более эффективный алгоритм.



Рисунок 5.6 – Демаркационная линия и обрамляющий ее прямоугольник

В результате извлечения кадров из видеопоследовательностей эндоскопических исследований желудка получена общая база изображений размером 14 424 шт. Она содержит изображения, на которых нет патологических объектов, относящихся к рассмотренным выше классам, но представлен участок слизистой оболочки желудка (фон). Такие изображения определены как «норма», при этом разметка объектов на таких изображениях не производилась [304].

5.2.4 Итоговая статистика по базе изображений желудка

Предложено три варианта баз эндоскопических изображений желудка, описание которых представлено в таблице 5.5 [299, 300]:

- 1. База «А», включающая только изображения класса «рак».
- 2. База «В», включающая в себя изображения классов «рак», «ранний рак» и «иная патология».
- 3. База «С», включающая классы «рак», «ранний рак», «иная патология», а также фоновые изображения (нормальная слизистая оболочка желудка).

186

|--|

Характеристики	База «А»	База «В»	База «С»
	(«только рак»)	(без «норма»)	(с «норма»)
Классы	1 класс: «рак»	3 класса: «рак»,	3 класса: «рак»,
изображений		«ранний рак», «иная	«ранний рак», «иная
		патология»	патология» +
			фоновые («норма»)
Общий размер базы	1 335	3 246	14 424
изображений			
Деление базы на	1 166 шт.	2 769 шт.	10 983 шт. (обуч.,
обучающую /	(обуч., 87%),	(обуч., 86%),	76%), 669 шт.
валидационную /	70 шт.	174 шт.	(валид., 5%),
тестовую выборки	(валид., 5%),	(валид., 5%),	2 772 шт.
	99 шт.	303 шт.	(тест., 19%)
	(тест., 8%)	(тест., 9%)	
Изображения класса	Нет	Да, в соответствии с	Да, в соответствии с
«иная патология» в		таблицей 5.2.	таблицей 5.2.
составе базы			
Изображения без	Нет	Нет	8 214 шт. (обуч.),
патологии в составе			495 шт. (валид.),
базы			2 469 шт. (тест.)

Таблица 5.5 – Описание баз изображений желудка

Изображения в каждой базе распределялись между обучающей, валидационной и тестовой выборками, которые не пересекались между собой. Формирование таких выборок для баз данных происходило следующим образом:

- Для тестовой выборки брались наборы изображения, соответствующих исследованиям отдельных пациентов (так, чтобы присутствовали изображения с двух эндоскопических аппаратов, в отношении примерно 1:1).
- Из оставшихся данных выделялась валидационная и обучающая выборки. В валидационную выборку входило примерно 5% последних кадров видеопоследовательности (но всегда минимум 1 снимок включался в валидационный набор) для исследования каждого пациента. Такой метод позволил разделить выборки таким образом, чтобы данные в них максимально

отличались друг от друга, и результат работы сети максимально не зависел от изображений в обучающей выборке.

5.3 Результаты моделирования нейронных сетей на основе архитектур SSD и RetinaNet со стандартными параметрами

Рассмотрим алгоритмы детектирования патологий на изображениях желудка на основе нейросетевых архитектур SSD и RetinaNet. Такие алгоритмы работают за один проход (single shot), что дает выигрыш в производительности по сравнению двухпроходными архитектурами [110]. Для построения первой версии алгоритма обнаружения патологий слизистой оболочки желудка при проведении эндоскопического исследования выбрана архитектура SSD размер входного слоя которой составил 300х300 (рисунок 5.7). Ряд начальных слоев сети SSD (базовая сеть) представляет собой стандартную архитектуру сети VGG16, предназначенной для анализа изображений высокого качества (при этом отсекаются слои классификации). В качестве базовой сети также может быть выбрана другая сеть.



Рисунок 5.7 – Архитектура сети SSD-300 на основе базовой сети VGG-16

Сеть SSD генерирует огромное количество (десятки тысяч) предсказаний ограничивающих рамок и классов объектов на разных масштабах и соотношениях сторон из различных слоев VGG16, а также баллы за наличие объекта в рамке, после чего эти данные подаются на вход алгоритма подавления немаксимумов (Non-Maximum Suppression – NMS) для окончательного детектирования

188

патологического объекта в кадре. Многие из этих предсказанных рамок для одного и того же объекта будут сильно перекрываться (иметь высокий IoU – Intersection over Union). Без слоя NMS на выходе модели для одного объекта было бы множество почти идентичных рамок, что практически неприемлемо [109].

В связи с тем, что разрешение кадра видеоданных в записанных гастроскопических исследованиях составляло 626х532 пикселей, помимо стандартной архитектуры SSD-300 в исследовании использовалась также архитектура SSD с базовой сетью VGG16, размер входного слоя которой составляет 512х512. При ее использовании в качестве базовой сети появляется еще один дополнительный масштаб карт признаков для детектирования.

Разрабатываемое аппаратно-программное решение для СППВР должно быть с одной стороны недорогим для конечной реализации, с другой стороны – проводить обработку видеопотока в режиме реального времени. С учетом этого, дополнительно реализована сеть SSD, в которой в качестве базовой сети использовалась часть архитектуры сети MobileNet. Особенностью архитектуры MobileNet является ее «легковесность» относительно других сетей. Экономия способами: вычислительных ресурсов происходит тремя основными использование более простых вычислений (раздельные свертки), уменьшение числа каналов, возможность работы с изображениями низкого разрешения. Благодаря этим особенностям архитектура MobileNet часто используется на мобильных и портативных устройствах, оставаясь при этом достаточно «продвинутой» для многих типовых задач распознавания [299].

Еще одной сетью, выбранной в качестве базовой сети для алгоритма на основе SSD, является сеть ResNet50 [250]. Данная сеть показывает хорошие результаты в задачах анализа изображений и также является более «легковесной», чем сеть VGG. Пятой архитектурой к детектированию патологий желудка на изображениях выбрана архитектура сверточной нейронной сети – RetinaNet также с базовой сетью ResNet50.

Таким образом, для дальнейших исследований в первой итерации выбраны следующие архитектуры сверточных нейронных сетей:

189

- 1) SSD с базовой сетью VGG16, размер входного слоя 300х300;
- 2) SSD с базовой сетью VGG16, размер входного слоя 512х512;
- 3) SSD с базовой сетью MobileNetV1, размер входного слоя 300х300;
- 4) SSD с базовой сетью ResNet50, размер входного слоя 300х300;
- 5) RetinaNet с базовой сетью ResNet50.

Ввиду различий архитектур на основе SSD и RetinaNet, опишем результаты обучения соответствующих нейросетевых алгоритмов при использовании стандартных параметров раздельно в следующих двух пунктах.

5.3.1 Исследования для алгоритмов на базе нейронной сети SSD

со стандартными параметрами

Для проведения исследования работы алгоритма детектирования на основе четырех из описанных выше архитектур на основе HC SSD использовались стандартные параметры запуска. В качестве оптимизатора применялся алгоритм Adam с параметрами: beta_1 = 0,9, beta_2 = 0,999, epsilon = 10^{-8} , в качестве аугментации – безгоризонтный тип со стандартными настройками. Исследования проводились на базе изображений желудка с устраненным эффектом чересстрочной развертки. Более подробно эти параметры обсуждаются ниже в пункте 5.4.

Обучение НС происходило на суперкомпьютере NVIDIA DGX-1 в центре искусственного интеллекта ЯрГУ. В глубоком обучении данные (в данном случае изображения) редко подаются в НС по одному. Вместо этого их группируют в пакеты. Размер батча (пакета) составлял 8 изображений, которые НС «видит» и обрабатывает за один проход. Нейронная сеть выполняет свои вычисления (прямой проход – forward pass) над всеми 8 изображениями в батче одновременно. После обработки батча происходит запуск процедуры обучения [54]:

- Расчет усредненной ошибки в батче (loss).
- Обратный проход (backward pass): на основе этой усредненной ошибки батча вычисляются градиенты, которые показывают, как нужно изменить каждый

параметр (вес) сети, чтобы уменьшить ошибку именно на этом батче из 8 изображений.

 Обновление весов: оптимизатор (например, SGD, Adam) использует вычисленные градиенты, чтобы обновить веса сети. Это обновление происходит один раз за обработку одного батча.

Весь этот цикл (прямой проход 8 изображений – расчет усредненной ошибки – обратный проход – обновление весов) называется одним шагом обучения, итерацией или одной эпохой.

Проводилась нормировка данных. Из каждого цветового канала в системе RGB осуществлялся вычет, который составлял 123, 117 и 104 для красного, зеленого и синего каналов, соответственно. Затем осуществлялась перестановка цветовых каналов в обратном порядке: $RGB \rightarrow BGR$.

В исследовании использовалась следующая стратегия изменения скорости обучения: базовая (начальная) скорость обучения выбиралась равной 10^{-4} , затем с 80 по 99 эпоху скорость обучения принимала значение 10^{-5} , а с 100 по 120 эпоху – значение, равное 10^{-6} . Обучение алгоритма останавливалось заранее (EarlyStopping), если значение метрики mAP на валидационной выборке не увеличивалось в течение пяти эпох. Далее перечисленные выше условия запуска процесса обучения будем называть «стандартными параметрами запуска».

Результаты обучения на базе изображений желудка «В» для различных архитектур на базе сети SSD со стандартными параметрами запуска представлены в таблице 5.6. В качестве метрик оценки качества работы алгоритмов на основе SSD использовались значения функции потерь на обучающей сети И валидационной выборках (loss и val_loss, соответственно), а также значения стандартных метрик F1, mAP, Precision, Recall. Они вычислялись на валидационной выборке, поэтому перед названием каждой из них в таблице присутствует приставка «val_». Кроме того, для каждого типа оптимизатора указывался номер последней эпохи (last_epoch) при обучении алгоритма, на которой происходило увеличение значения метрики тАР.

Нейронная сеть	loss	val_F1	val_ loss	val_ mAP	val_ Precision	val_ Recall	last_ epoch
SSD + VGG16-300	4,375	0,711	4,812	0,644	0,706	0,632	18
SSD + VGG16-512	4,190	0,658	4,440	0,624	0,686	0,632	13
SSD + MobileNet-300	4,611	0,561	4,852	0,475	0,842	0,421	13
SSD + ResNet50	5,500	0,638	5,349	0,347	0,710	0,579	13

Таблица 5.6 – Исследования для сети SSD со стандартными параметрами запуска

Результаты моделирования показывают, что наилучшие показатели согласно метрике mAP, как основной мере качества работы детекторов при запуске сетей со стандартными параметрами, соответствуют использованию сети SSD с базовой частью VGG16-300.

5.3.2 Исследования для алгоритма на базе нейронной сети RetinaNet со стандартными параметрами

Для проведения исследования работы алгоритма детектирования на основе нейронной сети RetinaNet с базовой частью ResNet50 также использовались стандартные параметры запуска. Исследования, как и в предыдущем случае, проводились на базе изображений «В» с устраненным эффектом чересстрочной развертки. Результаты представлены в таблице 5.7. Для сравнения в таблицу добавлены результаты для алгоритма на основе сети SSD с базовой частью VGG16-300, показавшего лучшие результаты в предыдущем исследовании.

Нейронная сеть	val_F1	val_mAP	val_Precision	val_Recall	last_epoch
RetinaNet	0,567	0,791	0,660	0,496	25
SSD + VGG16-300	0,711	0,644	0,706	0,632	18

Таблица 5.7 – Исследования для сети RetinaNet со стандартными параметрами запуска

Сеть RetinaNet при запуске со стандартными параметрами значительно превосходит алгоритм SSD с базовой частью VGG16-300 согласно значению метрики mAP (на 0,147). Однако согласно всем остальным метрикам (F1, Precision, Recall) наблюдается преимущество сети SSD с базовой частью VGG16-300.

В таблице 5.8 приведены значения метрики AP, вычисленные для HC RetinaNet и SSD с базовой частью VGG16-300 для различных классов, присутствующих в исследовании на базе изображений «В» при запуске обучения с параметрами по умолчанию.

		~ ~ ~ ~	
$\mathbf{h} = \mathbf{h} = $			\mathcal{A}
1 a 0 1 0 1 a 0 0 - 1 0 3 0 1 b a b 1 1 3	назличных классов при с	лоучений на оа	13C ИЗООНАЖСНИИ «D»
i worning work i togris i wish gasi			

Нейронная	AP	AP	AP
сеть	(«рак»)	(«ранний рак»)	(«иная патология»)
RetinaNet	0,873	0,976	0,524
SSD + VGG16-300	0,642	0,937	0,453

Видно, что алгоритм на базе сети RetinaNet превосходит по значениям метрики AP подход на основе сети SSD для всех классов, рассматриваемых в данном исследовании. Тем не менее, текущее исследование выполнялось при запуске HC со стандартными параметрами, поэтому делать вывод об однозначном преимуществе какой-либо архитектуры неверно. Значения метрик, полученные в данном исследовании для всех HC, являются достаточно невысокими, что говорит

о необходимости настройки параметров алгоритма под рассматриваемую в данном исследовании задачу обнаружения патологий желудка.

5.4 Оптимизация параметров нейросетевых алгоритмов

Целью оптимизации является поиск значений весов (weights) и смещений (biases), вместе называемых параметрами нейросетевой модели, при которых функция потерь (loss function) достигает своего минимума (или хотя бы достаточно малого значения) на обучающих данных. Это означает, что нейросетевая модель делает предсказания, максимально близкие к истинным значениям.

Математически задача обучения НС сводится к задаче минимизации скалярной функции потерь *L* в пространстве очень высокой размерности (миллионы или миллиарды параметров). Ввиду того, что пространство параметров огромно – перебор всех возможных комбинаций их значений невозможен [54]. особенно Алгоритмы оптимизации, основанные на градиентном спуске, предоставляют эффективный итерационный метод поиска хорошего минимума. Алгоритм обратного распространения ошибки (backpropagation) позволяет точно и эффективно вычислить, как каждый отдельный параметр влияет на общую ошибку [102]. Оптимизаторы используют эту информацию для целенаправленного обновления параметров. Без направленного обновления на основе градиента (или его оценок) модель либо никогда не научится (параметры останутся случайными), либо обучение займет неоправданно много времени.

Важным практическим моментом также является выбор наилучшей процедуры аугментации данных с учетом специфики эндоскопических изображений. Другим важным моментом является использование той или иной функции потерь и рассмотрение различных вариаций архитектур HC.

Таким образом, необходимость подбора оптимизированных параметров для большого числа различных составляющих алгоритма детектирования патологий делает задачу повышения точности и робастности нейросетевого алгоритма весьма сложной. Рассмотрим исследования данного процесса более подробно.

5.4.1 Исследования по выбору политики обучения

Выбор политики обучения заключается в подборе типа оптимизатора для рассматриваемых нейросетевых архитектур. Для компьютерных экспериментов выбраны три популярных типа оптимизаторов. Рассмотрим кратко основные принципы их работы.

1) Стохастический градиентный спуск

Наиболее известным алгоритмом оптимизации является стохастический градиентный спуск (stochastic gradient descent – SGD). Он позволяет добиться оценки градиента за счет его усреднения по мини-пакету нескольких обучающих примеров.

Перед началом работы алгоритма задаются скорость обучения ε_k и начальные значения параметров θ . Далее стохастический градиентный спуск реализуется следующим образом до тех пор, пока не происходит остановка:

- из обучающего набора выбирается m примеров $\{x_{(1)}, ..., x_{(m)}\}$, каждому из которых соответствует своя метка $y_{(i)}$;
- производится оценка градиента $\hat{g} \leftarrow +(1/m) \nabla_{\theta} \sum_{i} L(f(x_{(i)}; \theta), y_{(i)}),$ где L выбранная функция потерь;
- рассчитываются новые значения параметров: $\theta \leftarrow \theta \varepsilon \hat{g}$.

Скорость обучения в данном алгоритме является главным параметром, с ростом количества итераций она уменьшается линейно до определенной итерации τ . На всех последующих итерациях скорость постоянна и равна $\varepsilon_k = (1 - \alpha)\varepsilon_0 + \alpha\varepsilon_{\tau}$, где $\alpha = k/\tau$.

Главное достоинство алгоритма SGD заключается в том, что время, необходимое на одно обновление, не возрастает с ростом количества примеров в обучающей выборке данных. Основная сложность данного метода заключается в выборе начальной скорости обучения.

Многочисленные исследования показывают, что функция потерь может обладать разной чувствительностью по отношению к разным параметрам HC [54]. В связи с этим, особо популярными являются алгоритмы оптимизации

с адаптивной скоростью обучения, где присутствует возможность ее выбора отдельно для каждого параметра и автоматического регулирования таких скоростей на протяжении всего обучения. К таким алгоритмам относят алгоритм среднеквадратичного распространения (RMSProp – root mean square propagation) и алгоритм с адаптивными моментами (adaptive moments – Adam).

2) Алгоритм RMSProp представляет собой усовершенствованную модификацию адаптивного градиентного алгоритма AdaGrad, который позволял подстраивать скорость обучения каждого из параметров путем их умножения на величину, обратно пропорциональную квадратному корню из всех предыдущих значений градиента в квадрате. Особенность, характерная для алгоритма RMSProp, заключается в том, что вместо квадрата градиентов всех предыдущих значений применяется экспоненциально затухающее среднее значение: скорость сходимости при обнаружении выпуклой впадины повышается, будто по достижении ее алгоритм инициализируется заново.

Алгоритмически RMSProp выглядит следующим образом [54]. Сначала задаются некоторые начальные параметры: ε – скорость обучения, ρ – скорость затухания, θ – начальные значения параметров, δ – малая константа. Инициализируется переменная для агрегированного градиента – r.

Далее до тех пор, пока алгоритм не будет остановлен, повторяется следующий цикл:

- из обучающего набора выбирается m примеров $\{x_{(1)}, \dots, x_{(m)}\}$, каждому из которых соответствует своя метка $y_{(i)}$;
- производится оценка градиента $g \leftarrow (1/m) \nabla_{\theta} \sum_{i} L(f(x_{(i)}; \theta), y_{(i)});$
- объединяются квадраты градиента как $r \leftarrow \rho r + (1 \rho)g^2$;
- рассчитываются новые значения параметров: $\Delta \theta \leftarrow -\frac{\varepsilon}{\sqrt{\delta+r}}g$, где $\frac{\varepsilon}{\sqrt{\delta+r}}$ применяется к каждому элементу;
- производится обновление параметров: θ ← θ + $\Delta\theta$.
- 3) Алгоритм Adam представляет собой комбинацию алгоритма RMSProp и импульсного метода (Momentum) с небольшими изменениями. Импульсный

метод в оптимизации HC – техника, заимствованная из физики, которая помогает ускорить сходимость и улучшить устойчивость градиентного спуска. Ее суть – добавить «инерцию» процессу обновления весов. Так, в Adam импульс включен в виде оценки первого момента градиента. Он применяется в сочетании с масштабированным градиентом RMSProp. Кроме того, в Adam есть поправка на смещение в оценке первых и вторых моментов для учета их инициализации в начале координат.

Процедура оптимизации Adam выглядит следующим образом [55]. Устанавливаются значения параметров: $\varepsilon = 0,001 - \text{шаг}$, коэффициенты экспоненциального затухания для оценки моментов $\rho_1 = 0,9$ и $\rho_2 = 0,999$, δ – малая константа, θ – начальные значения параметров. Инициализируются переменные s = 0 для первого момента и r = 0 – для второго момента, а также шаг t = 1 в нулевой момент времени. Далее, до тех пор, пока алгоритм не будет остановлен, повторяется следующий цикл:

- из обучающего набора выбирается m примеров $\{x_{(1)}, ..., x_{(m)}\}$, каждому из которых соответствует своя метка $y_{(i)}$;
- производится оценка градиента $g \leftarrow (1/m) \nabla_{\theta} \sum_{i} L(f(x_{(i)}; \theta), y_{(i)});$
- $-t \leftarrow t+1;$
- производится обновление смещенной оценки первого момента: $s \leftarrow \rho_1 s + (1 \rho_1) g;$
- производится обновление смещенной оценки второго момента: $r \leftarrow \rho_2 r + (1 \rho_2) g^2;$
- корректируется смещение первого момента: $\hat{s} \leftarrow \frac{s}{1-\rho_1^t}$;
- корректируется смещение второго момента: $\hat{r} \leftarrow \frac{r}{1-\rho_2^t}$;
- вычисляется обновление весов: $\Delta\theta$ ← $-\varepsilon \frac{\hat{s}}{\sqrt{\hat{r}}+\delta}$;
- производится обновление параметров: $\theta \leftarrow \theta + \Delta \theta$.

Используемые при компьютерном моделировании параметры оптимизаторов приведены в таблице 5.9.

Тип оптимизатора	Параметры
SGD	Momentum = 0,9 (ускорение в направлении градиентного спуска) Nesterov = False (момент Нестерова не включен в расчет)
RMSProp	rho = 0,9 (коэффициент забывания градиентов с предыдущих шагов) epsilon = 10 ⁻⁷ (малая величина, для численной устойчивости)
Adam	beta_1 = 0,9 beta_2 = 0,999 (beta_1, beta_2 – коэффициенты забывания градиентов) epsilon = 10 ⁻⁸

Таблица 5.9 – Параметры оптимизаторов

Для проведения исследования по выбору оптимизатора необходимо зафиксировать все другие параметры и составляющие алгоритма. В данном случае такая оценка осуществлялась фиксацией архитектуры SSD с базовой сетью VGG16 и размером входного слоя 300х300. В качестве способа размножения данных использовался стандартный тип аугментации.

Результаты использования различных типов оптимизаторов при обучении на базе изображений «А» представлены в таблице 5.10.

Тип	loss	val_F1	val_loss	val_mAP	val_ Precision	val_ Recall	last_ epoch
SGD	6,048	0,870	6,182	0,851	0,896	0,845	18
RMSProp	4,498	0,871	4,440	0,861	0,842	0,901	8
Adam	4,022	0,944	4,095	0,909	0,944	0,944	20

Таблица 5.10 – Исследование по выбору оптимизатора при обучении на базе «А»

Наибольшее число итераций обучения алгоритма детектирования потребовалось при использовании алгоритма Adam. Однако результаты, которые он продемонстрировал, значительно превосходят аналогичные показатели для двух

других типов оптимизаторов. Значение функции потерь на обучающей и валидационной выборках для оптимизатора Adam минимальное, что говорит о лучшей сходимости этого алгоритма по сравнению с другими.

Динамика значений функции потерь, вычисленных на обучающей и валидационной выборках (loss и val_loss, соответственно), а также значений метрик F1, mAP, Precision, Recall, вычисленных на валидационной выборке, от числа эпох обучения для оптимизатора Adam при обучении алгоритма на базе изображений «А» приведены на рисунок 5.8.

Видно, что значения функции потерь почти монотонно падают, а значения метрик F1, mAP, Precision, Recall почти монотонно возрастают в процессе обучения. При этом метрики с небольшими осцилляциями приближаются к своему максимально возможному значению – единице.

Рассмотрим результаты обучения алгоритма детектирования при использовании различных типов оптимизаторов при обучении на более сложной базе изображений «В» (таблица 5.11).

Для обучения алгоритма с использованием разных типов оптимизаторов на базе изображений «В» потребовалось от 16 до 20 итераций. При этом минимальное значение функции потерь на обучающей выборке показал оптимизатор Adam, а на валидационной выборке – оптимизатор RMSProp. Наибольшие значения метрик F1 и mAP на валидационной выборке также получены при использовании оптимизатора Adam, а наибольшие значения метрик Precision и Recall – при использовании оптимизаторов SGD и RMSProp, соответственно. Таким образом, анализируя совокупность показателей, представленных в таблице 5.11, можно сделать вывод о том, что для базы изображений «В» также предпочтительным является использование оптимизатора Adam, но это преимущество не является таким однозначным, как в случае более простой базы изображений «А».



Рисунок 5.8 – Динамика значений метрик от числа эпох обучения на базе изображений «А» для оптимизатора Adam

200

Тип	loss	val_F1	val_loss	val_mAP	val_	val_	last_
					Precision	Recall	epoch
SGD	6,940	0,688	6,982	0,575	0,846	0,579	16
RMSProp	4,510	0,667	4,578	0,620	0,842	0,901	20
Adam	4,375	0,711	4,812	0,643	0,706	0,632	18

Таблица 5.11 – Исследование по выбору оптимизатора при обучении на базе «В»

Рассмотрим зависимости значений показателей качества работы алгоритма детектирования при обучении на базе «В» от числа эпох обучения при использовании оптимизатора Adam (рисунок 5.9).



Рисунок 5.9 – Динамика значений метрик от числа эпох обучения на базе изображений «В» для оптимизатора Adam



Рисунок 5.9 – Динамика значений метрик от числа эпох обучения на базе изображений «В» для оптимизатора Adam

Анализ приведённых графиков показывает, что значения метрик F1, mAP, Precision, Recall подвержены существенно большим осцилляциям в процессе обучения, что подчеркивает увеличение сложности задачи при рассмотрении базы изображений «В».

Результаты использования различных типов оптимизаторов при обучении на базе изображений «С» представлены в таблице 5.12.

Тип	loss	val_F1	val_loss	val_mAP	val_	val_	last_
					Precision	Recall	epoch
SGD	7,411	0,546	7,236	0,495	0,480	0,632	17
RMSProp	5,146	0,438	4,820	0,380	0,539	0,368	17
Adam	4,451	0,688	4,580	0,623	0,846	0,579	29

Таблица 5.12 – Исследование по выбору оптимизатора при обучении на базе «С»

При использовании оптимизаторов SGD и RMSProp для обучения алгоритма детектирования потребовалось по 17 итераций, а при использовании оптимизатора Adam – 29. Однако данный тип оптимизатора снова превзошел остальные подходы почти по всем метрикам (минимальные значения функции потерь на обеих выборках, наибольшие значения большинства метрик), кроме Recall (наибольшее значение здесь получено при использовании оптимизатора SGD).

Рассмотрим зависимости значений показателей качества работы алгоритма детектирования при обучении на базе «С» от числа эпох обучения при использовании оптимизатора Adam (рисунок 5.10). Из характера приведенных зависимостей видно, что значения метрик увеличиваются с ростом числа эпох обучения, но при этом достигаются более низкие значения, чем при обучении алгоритма на базе изображений «А». Причиной снижения значений метрик, как и в случае обучения на базе изображений «В», является высокая сложность обучающих данных, в частности, наличие значительной внутриклассовой изменчивости для класса патологий, не относящихся к раковым.



Рисунок 5.10 – Динамика значений метрик от числа эпох обучения на базе изображений «С» для оптимизатора Adam



Рисунок 5.10 – Динамика значений метрик от числа эпох обучения на базе изображений «С» для оптимизатора Adam

Подводя итоги проведенного исследования, отметим, что, исходя из результатов обучения алгоритма для всех трех баз, лучший результат показал оптимизатор Adam. Таким образом, он и выбран для использования в дальнейших исследованиях как основной алгоритм оптимизации.

5.4.2. Исследования по выбору типа аугментации

При работе с медицинскими изображениями важным шагом является этап размножения или аугментации данных [105]. Он позволяет из имеющейся размеченной выборки изображений получить базу существенно большего размера путем применения к изображениям различных типов преобразований. В текущем исследовании для сети SSD рассматривались 3 варианта аугментации – стандартный, безгоризонтный и промежуточный.

При стандартной аугментации (СА), для того чтобы сделать модель более устойчивой к различным размерам и формам входных объектов, к каждому изображению обучающей выборки случайным образом применяется один из следующих вариантов преобразований:

- используется все исходное входное изображение;

204

- из изображения вырезается фрагмент (патч), для которого минимальное пересечение с объектом согласно мере Жаккара [42] составляет 0,1; 0,3; 0,5; 0,7 или 0,9;
- из изображения вырезается фрагмент (патч) случайным образом.

Размер каждого патча составляет от 10% до 100% от размера исходного изображения, а соотношение сторон находится в диапазоне от 1/2 до 2 к 1.

После выполнения описанного этапа производится изменение размеров каждого патча до некоторого фиксированного значения и осуществляется отзеркаливание части патчей по горизонтали, а также выбирается с вероятностью 0,5 одна из двух последовательностей фотометрических преобразований, приведенных в таблице 5.13.

Последовательность I	Последовательность ІІ
Рандомизация яркости	Рандомизация яркости
Рандомизация контраста	Рандомизация насыщенности (для формата
Рандомизация насыщенности (для формата	HSV)
HSV)	Рандомизация оттенка (для формата HSV)
Рандомизация оттенка (для формата HSV)	Рандомизация контраста
Случайная перестановка цветовых каналов	Случайная перестановка цветовых каналов
(отключена)	(отключена)
Случайное масштабирование	Случайное масштабирование
Случайная обрезка изображения	Случайная обрезка изображения
Отзеркаливание по горизонтали	Отзеркаливание по горизонтали
Изменение размера	Изменение размера

Таблица 5.13 – Последовательности фотометрических преобразований

При использовании безгоризонтной аугментации (БА) дополнительно используются преобразования, при которых учитывается тот факт, что рассматриваемые в исследовании эндоскопические изображения не имеют «верха» и «низа». В частности, в отличие от стандартной аугментации, в данном варианте рассматривается также отзеркаливание изображений по вертикали. В целом, при

безгоризонтного использовании варианта аугментации выполняется изображения преобразование ИЗ цветового пространства RGB в HSV, рандомизация яркости, контрастности, насыщенности и оттенка в заданных пределах, осуществляется случайный поворот изображения на 90, 180 и 270 градусов с вероятностью 0,5, отзеркаливание по горизонтали и вертикали, вырезание из изображения патчей случайным образом и изменение их размера с выбором различного соотношения сторон.

Третий вид размножения данных, применяемый в данном исследовании, – промежуточная аугментации (ПА). Она представляет собой набор преобразований, применяемый обычно для датасетов, в которых изображения имеют разный размер. Данная особенность не характерна для рассматриваемых изображений, однако использование данного вида аугментации также может дать хорошие результаты. При использовании промежуточного варианта аугментации к изображению применяется следующая последовательность преобразований: рандомизация яркости, рандомизация контрастности, рандомизация насыщенности и тона (для системы HSV), случайное зеркальное отражение по горизонтали, случайный выбор патчей, случайное масштабирование.

При проведении исследования по выбору наилучшего типа аугментации использовалась сеть SSD с базовой частью VGG16, вход которой составлял 300х300. В качестве оптимизатора использовался алгоритм Adam со стандартными параметрами.

Результаты исследования для алгоритма детектирования патологий желудка, обученного на базе изображений «А» при использовании различных типов аугментации, представлены в таблице 5.14.

Видно, что наименьшее значение функции потерь на обучающей и валидационной выборках получено при использовании стандартной аугментации. Также при использовании данного типа аугментации получены наибольшие значения метрик F1, mAP, Recall. Стоит отметить, что при использовании безгоризонтного варианта размножения данных получено наибольшее значение метрики Precision, а также достаточно большое значение mAP, очень близкое к аналогичному показателю, полученному для стандартной аугментации.

Вид	loss	val_F1	val_loss	val_mAP	val_ Precision	val_ Recall	last_
аутментации					1 100151011	Recall	epoen
СА	4,022	0,944	4,095	0,909	0,944	0,944	20
БА	4,198	0,934	4,176	0,905	0,970	0,901	16
ПА	4,109	0,929	4,303	0,888	0,942	0,916	15

Таблица 5.14 — Исследование по выбору типа аугментации при обучении на базе «А»

Рассмотрим аналогичные исследования на базе изображений «В» (таблица 5.15). Видно, что лучшие результаты при обучении алгоритма на базе изображений «В» согласно значениям функции потерь и метрикам F1 и mAP показал подход на основе безгоризонтной аугментации. Стоит также отметить, что наивысшие значения метрик Precision и Recall получены для промежуточной аугментации.

Вид	loss	val_F1	val_loss	val_	val_Pre-	val_Re-	last_
аугментации				mAP	cision	call	epoch
СА	4,375	0,711	4,812	0,643	0,706	0,632	18
БА	4,002	0,757	4,253	0,680	0,778	0,737	31
ПА	4,514	0,667	4,812	0,621	0,833	0,789	13

Таблица 5.15 – Исследование по выбору типа аугментации при обучении на базе «В»

В таблице 5.16 приведены исследования по выбору типов аугментации при использовании для обучения базы изображений «С».

Вид	loss	val_F1	val_loss	val_mAP	val_	val_	last_
аугментации					Precision	Recall	epoch
CA	4,451	0,688	4,580	0,623	0,846	0,579	29
БА	4,780	0,632	4,770	0,503	0,632	0,632	21
ПА	4,655	0,647	4,488	0,604	0,733	0,579	20

Таблица 5.16 – Исследование по выбору типа аугментации при обучении на базе С

Наилучшие результаты при обучении алгоритма на базе изображений «С» согласно значению функции потерь на обучающей выборке и почти всем используемым метрикам (кроме val_loss и Recall) показал подход на основе стандартного типа аугментации.

Проведенные дополнительные исследования по настройке параметров позволили сделать окончательный выбор в пользу стандартной аугментации, позволяющей получить наилучшее качество работы алгоритма детектирования на собранной базе эндоскопических изображений. При этом наилучшие значения метрик достигаются при следующих настройках алгоритма:

- рандомизация яркости, диапазон: (-16, 16);

- рандомизация контрастности, диапазон: (0,3; 1,5);
- рандомизация насыщенности, диапазон: (0,5; 1,8);
- рандомизация тона, значение сдвига: 18;
- случайное перемешивание цветовых каналов: выключено.

Результаты на базах изображений «А», «В», «С» при использовании стандартной аугментации с этими значениями параметров приведены в таблице 5.17.

Из анализа данных в таблицах 5.14–5.17 видно, что настройка параметров стандартной аугментации позволила повысить значения метрик для всех баз изображений, используемых в исследовании.

Базы	loss	val_F1	val_loss	val_mAP	val_	val_	last_
					Precision	Recall	epoch
А	4,056	0,935	4,093	0,891	0,956	0,916	16
В	3,841	0,718	4,185	0,728	0,728 0,700		32
С	4,380	0,706	4,639	0,610	0,800	0,632	31

Таблица 5.17 – Результаты при использовании СА с оптимальными значениями параметров

5.4.3 Исследования по устранению эффекта чересстрочной развертки

Чересстрочная развертка (ЧР) – метод телевизионной развертки, при котором каждый кадр разбивается на два полукадра, составленные из строк, выбранных через одну. В первом полукадре развертываются и воспроизводятся нечетные строки, во втором – четные, располагающиеся в промежутках между строками первого полукадра. Чересстрочная развертка является компромиссом между критической частотой заметности мельканий человеческим глазом и шириной полосы частот, занимаемой видеосигналом. Применение такой технологии в прикладном телевидении позволяет устранить избыточность частоты передаваемых кадров [44].

Чересстрочная развертка, применяемая в телевидении и некоторых форматах видеозаписи, имеет один существенный недостаток: из-за последовательной передачи полукадров при выводе телевизионного изображения на устройства отображения с прогрессивной развёрткой, такие как жидкокристаллические телевизоры, мультимедийные проекторы и др., на движущихся объектах появляется эффект «гребёнки» или мерцание изображения (эффект «интерлейсинга»). Чтобы избавиться от этого эффекта, применяется процесс конвертации видео в прогрессивный формат, который носит название процедуры деинтерлейсинга [36].

Прогрессивная развертка – метод телевизионной развертки, при котором для отображения, передачи или хранения движущихся изображений все строки

каждого кадра отображаются последовательно. Изображение с прогрессивной разверткой требует вдвое большей полосы частот для передачи по сравнению с чересстрочной. Таким образом, применение алгоритмов деинтерлейсинга к базе видеоизображений позволяет увеличить ее информативность вдвое.

Для проведения исследования по оценке влияния устранения ЧР на качество работы алгоритма детектирования необходимо зафиксировать все прочие параметры и составляющие алгоритма. В данном случае исследование проводилось с использованием нейронной сети SSD с базовой сетью VGG16 с размером входного слоя 300х300. В качестве оптимизатора использовался Adam со стандартными параметрами запуска, в качестве аугментации – стандартный тип с оптимизированными параметрами, выбранными в ходе предыдущего этапа. Исследования проводились на базах изображений «В» и «С». Процедура деинтерлейсинга в данном исследовании фактически играла роль предобработки эндоскопических изображений желудка. Рассмотрим ее влияние на качество работы алгоритма детектирования (таблица 5.18).

Набор изображений	loss	val_F1	val_loss	val_mAP	val_ Precision	val_ Recall	last_ epoch
База «В» с ЧР	3,841	0,718	4,185	0,728	0,700	0,737	32
База «В» с устраненной ЧР	3,384	0,833	4,046	0,737	0,882	0,790	29
База «С» с ЧР	4,380	0,706	4,639	0,610	0,800	0,632	31
База «С» с устраненной ЧР	3,777	0,824	3,854	0,678	0,933	0,737	35

Таблица 5.18 – Влияние деинтерлейсинга на метрики для баз изображений «В» и «С»

Установлено, что устранение ЧР позволяет повысить качество работы алгоритма детектирования согласно значениям абсолютно всех метрик, используемых в данном исследовании. При этом на базе изображений «В» с устраненной чересстрочной разверткой наблюдается относительно небольшое увеличение значений метрики mAP (на 0,009) относительно результатов, полученных на той же базе с наличием чересстрочной развертки. На более сложной базе изображений «С» с устраненной ЧР наблюдается более заметное улучшение качества работы алгоритма (на 0,068 по метрике mAP).

В таблице 5.19 представлены результаты анализа качества работы алгоритма детектирования при использовании деинтерлейсинга в качестве этапа предобработки для различных классов при обучении на базе изображений «В». Для двух классов из трех наблюдается улучшение средней точности, особенно заметное для изображений класса «рак».

Набор изображений	AP	AP	AP	
	(«рак»)	(«ранний рак»)	(«иная патология»)	
База В с ЧР	0,688	0,953	0,543	
База В с устраненной ЧР	0,737	0,955	0,519	

Таблица 5.19 – Влияние деинтерлейсинга при обучении на базе изображений «В»

Таким образом, введение этапа устранения чересстрочной развертки позволило повысить качество работы алгоритма на базах эндоскопических изображений «В» и «С», в частности, улучшить качество детектирования областей с раком. Полученные результаты позволяют рекомендовать использование процедуры деинтерлейсинга в качестве этапа предобработки для увеличения размеров обучающей выборки.

5.4.4 Выбор функции потерь

Важной составляющей процесса обучения НС является выбор функции потерь. Она используется для оценки расстояния между выходом сети, полученным для поданного на вход обучающего примера, и эталонным значением В лля данного примера. данном исследовании функции потерь ДЛЯ рассматриваемых нейросетевых архитектур выбирались на основе данных, представленных в исследованиях и научно-технической литературе по данной теме [53-55, 57]. Рассмотрим функции потерь, которые использовались для каждой из используемых архитектур.

Первая группа представлена архитектурами SSD, для которых варьируется базовая часть (VGG16-300, VGG16-512, MobileNetV1-300, ResNet50-300).

Функция потерь, выбранная для использования в нейронной сети SSD, состоит из двух величин: *L_{conf}* и *L_{loc}*

$$L(x,c,l,g) = \frac{1}{N} \Big(L_{conf}(x,c) + \alpha L_{loc}(x,l,g) \Big),$$

где *N* – количество сравниваемых блоков.

Сопоставляемые блоки определяются как:

$$\begin{split} L_{loc}(x, l, g) &= \sum_{i \in Pos \ m \in \{cx, cy, w, h\}}^{N} \sum x_{ij}^{k} \ smooth_{L1}(l_{i}^{m} - \hat{g}_{j}^{m}), \end{split}$$
где $\hat{g}_{j}^{cx} &= \left(g_{j}^{cx} - d_{i}^{cx}\right)/d_{i}^{w}, \ \hat{g}_{j}^{cy} &= \left(g_{j}^{cy} - d_{i}^{cy}\right)/d_{i}^{h}, \ \hat{g}_{j}^{w} = \log\left(\frac{g_{j}^{w}}{d_{i}^{w}}\right), \end{cases}$ $\hat{g}_{j}^{h} = \log\left(\frac{g_{j}^{h}}{d_{i}^{h}}\right), \end{split}$

где L_{loc} – характеризует потерю локализации, которая представляет собой плавную потерю L1 между параметрами прогнозируемого прямоугольника (l) и базового истинного значения (g). Эти параметры включают смещения для центральной точки (cx, cy), ширины (w) и высоты (h) ограничительной рамки

$$L_{conf}(x,c) = -\sum_{i\in Pos}^{N} x_{ij}^{p} \log(\hat{c}_{i}^{p}) - \sum_{i\in Neg}^{N} \log(\hat{c}_{i}^{0}),$$

где $\hat{c}_{i}^{p} = \frac{exp(c_{i}^{p})}{\sum_{p} exp(c_{i}^{p})},$

а L_{conf} характеризует потерю достоверности, которая представляет собой потерю в доверительных отношениях нескольких классов (*c*) (α устанавливается равным 1 при перекрестной проверке), $x_{ij}^p = \{1, 0\}$ – индикатор для сопоставления *i*-го блока с *j*-м истинным блоком категории *p*.

Второй тип архитектуры – сеть RetinaNet. В ней предлагается использование особой функции потерь – Focal Loss [57]. Выражение для ее расчета имеет следующий вид:

$$FL(\rho_t) = -\alpha_t (1 - \rho_t)^{\gamma} \log (\rho_t),$$

где (1 − *ρ*_t) – модулирующий фактор, *γ* ≥ 1 – параметр фокусировки, *α* – вес, присвоенный редкому классу, введенный для балансировки,

$$p_t = \begin{cases} p, если \ y = 1 \\ 1 - p, иначе \end{cases}$$

где *p* – оцененная моделью вероятность присутствия объекта в рамке-кандидате, которая в общем случае равна значению на выходе HC, приведенному к значению из промежутка [0, 1].

В процессе обучения сети RetinaNet функция потерь вычисляется для всех рассматриваемых ориентаций рамок-кандидатов со всех уровней масштабирования изображения. Ее итоговое значение вычисляется как сумма значений функции для всех рамок-кандидатов, нормализованное по количеству рамок, содержащих искомые объекты.

5.5 Сравнение результатов работы нейросетевых алгоритмов на основе SSD и RetinaNet со стандартными и оптимизированными параметрами

В рамках исследований в предыдущем пункте подобраны оптимальные параметры различных составляющих алгоритма детектирования патологий на изображениях желудка. Рассмотрим сравнительные результаты обучения нейросетевых алгоритмов при использовании стандартных и оптимальных параметров для архитектур на основе SSD и RetinaNet.

5.5.1 Результаты для алгоритмов на базе нейронной сети SSD с оптимизированными параметрами

Рассмотрим результаты анализа точности работы алгоритма детектирования на основе архитектур SSD. Для проведения исследования качества работы алгоритма детектирования использовались оптимизированные параметры запуска. В качестве оптимизатора использовался алгоритм Adam со стандартными аугментации параметрами запуска, В качестве _ стандартный ТИП С оптимизированными параметрами. Исследования проводились базе на изображений «В» с устраненным эффектом чересстрочной развертки. Результаты архитектур базе сети SSD различных на co стандартными ДЛЯ И таблице 5.20. оптимизированными параметрами запуска представлены В Результаты показывают, что использование оптимизированных параметров позволило улучшить показатели для всех архитектур на основе сети SSD.

При этом согласно метрике mAP как основной мере качества работы алгоритма детектирования:

- качество работы сети SSD + VGG16-300 удалось улучшить на 0,127;
- качество работы сети SSD + VGG16-512 удалось улучшить на 0,074;
- качество работы сети SSD + MobileNetV1-300 удалось улучшить на 0,115;
- качество работы сети SSD + ResNet50-300 удалось улучшить на 0,183.

Установлено, что наилучшие значения метрик при обучении алгоритма детектирования с оптимизированными параметрами достигнуты при использовании нейронной сети SSD с базовой частью VGG16-300 (значение метрики mAP составило 0,771).

Зависимости значений функции потерь, вычисленных на обучающей и валидационной выборках (loss и val_loss, соответственно), а также значений метрик F1, mAP, Precision, Recall, вычисленных на валидационной выборке, от числа эпох обучения для сети SSD с базовой частью VGG16-300 с использованием оптимизированных параметров на базе изображений «В» приведены на рисунке 5.11.

Таблица 5.20 – Сравнение результатов для алгоритма на основе сети SSD со стандартными и оптимизированными параметрами запуска при обучении на базе изображений «В»

Сеть, параметры запуска	loss	val_F1	val_loss	val_mAP	val_ Precision	val_ Recall	last_ epoch
SSD + VGG16-300, стандартные	4,375	0,711	4,812	0,644	0,706	0,632	18
SSD + VGG16-300, оптимизированные	3,081	0,842	4,209	0,771	0,842	0,842	30
SSD + VGG16-512, стандартные	4,190	0,658	4,440	0,624	0,686	0,632	13
SSD + VGG16-512, оптимизированные	3,989	0,747	4,226	0,698	0,757	0,737	17
SSD + MobileNetV1- 300, стандартные	4,611	0,561	4,852	0,475	0,842	0,421	13
SSD + MobileNetV1- 300, оптимизированные	3,525	0,821	4,540	0,590	0,800	0,842	48
SSD +ResNet50-300, стандартные	5,500	0,638	5,349	0,347	0,710	0,579	13
SSD + ResNet50-300, оптимизированные	4,498	0,788	4,644	0,530	0,929	0,684	31

Приведенные зависимости для значений функции потерь показывают, что алгоритм сходится как на обучающей, так и на валидационной выборках в районе 20-25 эпох. При этом достигаются относительно малые значения функции потерь в обоих случаях. Зависимости для метрик F1 и mAP показывают, что значение метрики F1 устанавливается в районе 0,8, а метрики mAP в районе 0,7 в районе 25-30 эпох обучения. Эти результаты соответствуют данным, приведенным в таблице 5.20. Зависимости для метрик Precision и Recall показывают, что их значения колеблются на уровне 0,85 и 0,8, соответственно, начиная с 25 эпохи

215

обучения. Эти результаты также соответствуют данным, приведенным в таблице 5.20. В целом зависимости, приведенные на рисунке 5.11, показывают, что значения метрик F1, mAP, Precision и Recall увеличиваются с ростом числа эпох обучения, достигая при этом достаточно высоких значений [289-291].



Рисунок 5.11 – Динамика значений метрик от числа эпох обучения для сети SSD + VGG16-300 с использованием оптимизированных параметров
5.5.2 Результаты для алгоритма на базе нейронной сети RetinaNet с оптимизированными параметрами

Для проведения исследования точности работы алгоритма детектирования на основе сети RetinaNet с базовой частью ResNet50 использовались следующие параметры запуска: в качестве оптимизатора использовался алгоритм Adam со стандартными параметрами запуска, в качестве способа размножения данных – тип аугментации, соответствующий стандартному варианту аугментации с оптимизированными параметрами. Исследования проводились базе на изображений «В» с устраненным эффектом чересстрочной развертки.

Сравнительные результаты для нейронной сети RetinaNet с базовой частью ResNet50 с оптимизированными и стандартными параметрами запуска представлены в таблице 5.21.

Сеть, параметры запуска	loss	val_F1	val_mAP	val_ Precision	val_ Recall	last_ epoch
RetinaNet, стандартные	0,238	0,567	0,791	0,660	0,496	25
RetinaNet, оптимизированные	0,307	0,583	0,808	0,739	0,482	19

Таблица 5.21 – Сравнение результатов для сети RetinaNet со стандартными и оптимизированными параметрами при обучении на базе изображений «В»

Использование оптимизированных параметров для сети RetinaNet позволило увеличить качество работы алгоритма детектирования патологий на эндоскопических изображениях на 0,017 согласно метрике mAP. Значения метрик F1 и Precision также увеличились. Стоит отметить, что при использовании оптимизированных параметров для нейронной сети RetinaNet достигнуто не столь заметное улучшение качества работы алгоритма, как в случае с архитектурами на основе сети SSD.

В таблице 5.22 приведены значения метрики AP, вычисленные для сети RetinaNet для различных классов, присутствующих в исследовании, при запуске со стандартными и оптимизированными параметрами. Значения метрики AP оказались в данном случае близкими для всех трех классов, рассматриваемых в данном исследовании. При этом при запуске алгоритма на основе сети RetinaNet с оптимизированными параметрами наблюдается незначительное ухудшение качества работы для классов «рак» и «ранний рак» и улучшение для наиболее «сложного» класса – «иная патология».

Таблица 5.22 – Результаты сети RetinaNet со стандартными и оптимизированными параметрами для различных классов при обучении на базе изображений «В»

Сеть, параметры запуска	AP («рак»)	АР («ранний рак»)	АР («иная патология»)
RetinaNet, стандартные	0,873	0,976	0,524
RetinaNet, оптимизированные	0,869	0,968	0,586

Из всех пяти рассмотренных на первом этапе нейросетевых алгоритмов детектирования патологий на изображениях желудка наиболее высокое значение метрики mAP получено при использовании сети RetinaNet (0,808). Однако при использовании данного подхода значения метрик F1 и Recall оказались значительно ниже, чем при использовании сети SSD с базовой частью VGG16-300 с оптимизированными параметрами запуска.

Дальнейшее совершенствование нейросетевых алгоритмов возможно либо путем наращивания размеченных баз обучающих примеров с патологиями желудка (что является трудоемкой операцией), либо путем использования более современных архитектур HC. На практике желательно использовать два этих подхода одновременно [292-295].

5.6 Результаты моделирования дополнительных архитектур нейронных сетей

Развитие методов глубокого машинного обучения привело к совершенствованию архитектур НС. В связи с этим проведены дополнительные компьютерные эксперименты по следующим направлениям:

- исследование алгоритмов на основе сверточной нейронной сети SSD с различными вариациями базовой части VGG19, MobileNetV2, Xception [110];
- исследование алгоритма на основе модели сверточной нейронной сети EfficientDet на базе EfficientNetB0 [300].

Для обучения и тестирования использовалась база эндоскопических изображений «В», описанная выше. Рассмотрим полученные результаты компьютерного моделирования.

5.6.1 Алгоритмы на основе нейронной сети SSD

Сверточная нейронная сеть VGG19 содержит 19 слоев: 16 сверточных и 3 полносвязных слоя (рисунок 5.12) [110]. В ней используются сверточные фильтры 3×3 с шагом 1, а также слои макспулинга (maxpool) 2x2 с шагом 2. Это операция уменьшения размерности в сверточных HC, которая выделяет наиболее значимые признаки из входных данных. Другими словами макспулинг – «грубое сжатие» данных, оставляющее только самые сильные сигналы в локальных областях для эффективности и устойчивости модели.

В модели VGG-19 содержится 138 миллионов параметров, она обучается на подмножестве базы данных ImageNet27 и может классифицировать изображения по 1000 категориям объектов [104]. В данной работе усеченная сеть VGG19 использована в архитектуре SSD как базовая часть для выделения признаков из изображений желудка.



Рисунок 5.12 – Архитектура сверточной нейронной сети VGG19

По аналогии с пунктом 5.5 проводилась оптимизация параметров нейронной сети SSD + VGG19. Зависимости значений функции потерь, вычисленных на обучающей и валидационной выборках (loss и val_loss, соответственно), а также значений метрик F1, mAP, Precision, Recall, вычисленных на валидационной выборке, от числа эпох обучения для данного алгоритма с использованием стандартных и оптимизированных параметров запуска приведены на рисунках 5.13–5.14.

Как видно из кривых зависимости функции потерь от числа эпох обучения, как на валидационной, так и на обучающей выборках алгоритм хорошо сходится, при этом при запуске с оптимизированными параметрами для обучения алгоритма требуется большее число эпох. Значения метрик F1, mAP, Precision и Recall достигают достаточно средних значений при использовании оптимизированных параметров запуска. Следует отметить высокую осцилляцию метрик Precision и Recall в процессе обучения. Число эпох, требуемое для сходимости алгоритма обучения в этом случае выше, чем у нейронной сети SSD+VGG16.



Рисунок 5.13 – Динамика значений метрик от числа эпох обучения для нейронной сети SSD + VGG19 с использованием стандартных параметров



Рисунок 5.14 – Динамика значений метрик от числа эпох обучения для нейронной сети SSD + VGG19 с использованием оптимизированных параметров

Нейронная сеть MobileNetV2 – второе поколение нейросетевых алгоритмов семейства MobileNet. Архитектуры данного семейства являются очень

легковесными благодаря использованию особых типов свертки – DS-свертки (Depthwise Convolution) и 1×1-свертки (Pointwise Convolution). Основная идея состоит в том, чтобы заменить полный сверточный оператор факторизованной версией, которая разбивает свертку на два отдельных слоя. DS-свертка представляет собой следующую последовательность операций:

- разделение входных данных и фильтра по каналам;
- выполнение стандартной свертки для каждого канала;
- объединение выходных данных.

Обычная свертка принимает входной тензор L_i размером $h_i \times w_i \times d_i$ и применяет сверточное ядро $K \in \mathbb{R}^{k \cdot k \cdot d_i \cdot d_j}$, после чего на выходе получается тензор L_j размером $h_i \times w_i \times d_j$. Стандартные сверточные слои требуют больших вычислительных затрат: $h_i \cdot w_i \cdot d_i \cdot d_j \cdot k \cdot k$.

DS-свертки, используемые в семействе MobileNet, являются эффективной заменой стандартных сверточных слоев. Эмпирически они работают почти так же хорошо, как обычные свертки. При этом их вычислительные затраты ниже и составляют $h_i \cdot w_i \cdot d_i (k^2 + d_i)$.

Вторая модификация представляет собой свертку 1×1, которая «проходит через каждый пиксель». Она отвечает за создание новых признаков путем вычисления линейных комбинаций входных каналов. Ядро 1×1-свертки имеет глубину, равную количеству каналов входного изображения.

В данном исследовании часть сети MobileNetV2, а также ряд достроенных слоев использовались в качестве базовой части в архитектуре сверточной нейронной сети SSD.

Зависимости значений функции потерь, вычисленных на обучающей и валидационной выборках (loss и val_loss, соответственно), а также значений метрик F1, mAP, Precision, Recall, вычисленных на валидационной выборке, от числа эпох обучения для данного алгоритма с использованием стандартных и оптимизированных параметров запуска приведены на рисунках 5.15–5.16.



Рисунок 5.15 – Динамика значений метрик от числа эпох обучения для сети MobileNetV2 с использованием стандартных параметров



225

Рисунок 5.16 – Динамика значений метрик от числа эпох обучения для сети MobileNetV2 с использованием оптимизированных параметров

Как видно из кривых зависимости функции потерь от числа эпох обучения, алгоритм хорошо сходится на обучающей выборке, при этом при запуске с

оптимизированными параметрами кривая носит менее гладкий характер. Худшая сходимость достигается при запуске на валидационной выборке: кривые имеют характер, не типичный для подобного рода зависимостей. При этом, наоборот, при запуске с оптимизированными параметрами алгоритм сходится быстрее. В целом, из анализа рисунках 5.15–5.16 следует, что значения метрик оценки качества при использовании сети MobileNetV2 в качестве базовой части модели SSD оказываются ниже, чем при использовании сети VGG19. Также оптимизированные параметры запуска в данном случае не дают выигрыша, например, по метрике mAP.

Нейронная сеть Xception – компактная глубокая сверточная HC, предложенная создателем библиотеки Keras Франсуа Шолле [56]. В описании противоположные по смыслу слова «компактная» и «глубокая» присутствуют неслучайно: в данной архитектуре удалось соединить эти две противоположности, используя факторизацию обычного сверточного слоя.

В обычном сверточном слое происходит обучение фильтров в трехмерном пространстве: с двумя пространственными размерностями (ширина и высота) и размерностью, отвечающей за число каналов. Таким образом, одному ядру свертки ставится задача одновременного отображения межканальной информации (корреляции точек разных каналов) и пространственной информации (корреляции соседних точек внутри одного канала).

Архитектура Xception имеет 36 сверточных слоев, образующих базу выделения признаков сети. Эти слои структурированы в 14 модулей, каждый из которых имеет линейные остаточные связи вокруг них, кроме первого и последнего модулей. Иными словами, архитектура Xception – линейный стек разделимых по глубине сверточных слоев с остаточными связями. В данной работе от архитектуры Xception брались 6 достроенных слоев.

Зависимости значений функции потерь, вычисленных на обучающей и валидационной выборках (loss и val_loss, соответственно), а также значений метрик F1, mAP, Precision, Recall, вычисленных на валидационной выборке, от числа эпох обучения для данного алгоритма с использованием стандартных и оптимизированных параметров запуска приведены на рисунках 5.17–5.18.



Рисунок 5.17 – Динамика значений метрик от числа эпох обучения для нейронной сети Хсерtion с использованием стандартных параметров



Рисунок 5.18 – Динамика значений метрик от числа эпох обучения для нейронной сети Хсерtion с использованием оптимизированных параметров

Как видно из кривых зависимости функции потерь от числа эпох обучения, алгоритм хорошо сходится на обучающей выборке, при этом при запуске с

оптимизированными параметрами кривая носит менее гладкий характер. Худшая сходимость получается при запуске на валидационной выборке, особенно при запуске с оптимизированными параметрами.

Из приведенных графиков видно, что при тестировании алгоритма на основе сети SSD с базовой частью Хсерtion достигаются достаточно высокие значения всех метрик оценки качества. Также следует отметить, что по данным графикам сложно сделать однозначный вывод преимуществе использования 0 оптимизированных или стандартных параметров запуска алгоритма. Рассмотрим сравнительные результаты для различных базовых сетей модели SSD. исследованных в этом пункте. Наибольшие значения метрик качества, полученные среди всех эпох обучения, для разработанных алгоритмов приведены в таблицах 5.23–5.24. Для сравнения в них добавлены значения метрик для наилучшего алгоритма на базе сети SSD300 + VGG16, полученные ранее.

Нейронная	loss	val_F1	val_loss	val_	val_	val_	last_
сеть				mAP	Precision	Recall	epoch
SSD300+ VGG19	4,547	0,571	5,029	0,514	0,889	0,421	47
SSD300+ MobileNetV2	4,045	0,720	5,334	0,459	0,730	0,711	22
SSD300+ Xception	2,519	0,762	4,068	0,669	0,696	0,842	63
SSD300+ VGG16	4,375	0,711	4,812	0,644	0,706	0,632	18

Таблица 5.23 – Результаты тестирования алгоритмов на основе модели SSD

с использованием различных базовых сетей со стандартными параметрами запуска

Как видно из данных таблицы 5.21, лучшие результаты со стандартными параметрами запуска получены при использовании сети Хсерtion в качестве

базовой части согласно метрикам F1, mAP и Recall. Также видно, что для обучения данной модели потребовалось самое большое число итераций.

В таблице 5.24 приведены результаты тестирования разработанных алгоритмов при использовании оптимальных параметров запуска. Использование оптимизированных параметров позволило значительно повысить значения практически всех метрик качества для всех алгоритмов.

val F1 Нейронная loss val loss val_ val_ val last mAP Precision Recall epoch сеть SSD300+ 0,849 4,249 5,207 0,653 0,886 0,816 61 VGG19 SSD300+ 5,099 3,452 0,743 0,511 0,813 0,684 38 MobileNetV2 SSD300+ 0,842 0,842 2,679 0,842 4,731 0,654 31 Xception SSD300+ 4,209 3,081 0,842 0,771 0,842 0,842 30 VGG16

Таблица 5.24 – Результаты тестирования алгоритмов на основе модели SSD с использованием различных базовых сетей с оптимизированными параметрами запуска

Следует отметить, что преимущества модели Xception перед остальными подходами, как при использовании стандартных параметров запуска, уже не наблюдается. Более того, она суммарно немного проигрывает в точности алгоритмам SSD300 + VGG16 и SSD300 + VGG19.

5.6.2 Алгоритм на основе нейронной сети EfficientDet

В 2020 году предложено новое семейство нейросетевых детекторов EfficientDet. Этот алгоритм использует базовую сеть EfficientNetB0 (предобучена на базе ImageNet) для извлечения признаков и эффективный алгоритм агрегации признаков на базе двунаправленной пирамиды (BiFPN – bi-directional feature pyramid network). Сверточные HC семейства EfficientDet строятся с использованием идеи одновременного масштабирования базовой части, пирамиды признаков и подсетей детектирования и классификации. Архитектура сверточной HC данного типа приведена на рисунке 5.19 [300-302].



Рисунок 5.19 – Архитектура сверточной нейронной сети EfficientDet

В данном исследовании сеть EfficientDet использована в качестве основы алгоритма детектирования патологий на изображениях желудка. Проведена аналогичная процедура оптимизации параметров запуска. Зависимости значений функций потерь (подсети классификации, подсети регрессии), вычисленных на обучающей выборке, а также значений метрик F1, mAP, Precision, Recall, вычисленных на валидационной выборке, от числа эпох обучения для данного алгоритма приведены на рисунке 5.20.

Как видно из приведенных зависимостей функций потерь для сети EfficientDet, алгоритм успешно сходится и достигаются малые значения функций потерь. Значения метрик для разработанного алгоритма очень высокие и во всех случаях стремятся к единице.

Наибольшие значения метрик, полученные среди всех эпох обучения, для разработанного алгоритма приведены в таблице 5.25.



Рисунок 5.20 – Динамика значений метрик от числа эпох обучения для нейронной сети EfficientDet с использованием оптимизированных параметров

Таблица 5.25 – Наилучшие значения метрик для сети архитектуры EfficientDet

Нейронная сеть	loss	val_F1	val_	val_	val_	last_
			mAP	Precision	Recall	epoch
EfficientDet+	0,218	0,872	0,973	0,898	0,848	24
EfficientNetB0						

5.6.3 Сравнительный анализ результатов для дополнительных архитектур нейронных сетей

Для сравнения работы нейросетевых алгоритмов патологий желудка использовались 4 варианта алгоритмов детектирования, построенных на базе нейросетевой архитектуры SSD300, сети RetinaNet и алгоритма на основе сети EfficientDet. В таблице 5.26 приведены значения средней метрики точности (AP), вычисленной для трех классов, имеющихся в размеченном наборе изображений «В», при использовании нейронных сетей SSD и EfficientDet с оптимизированными параметрами запуска.

Таблица 5.26 – Сравнение нейронной сети EfficientDet с алгоритмами на базе архитектуры SSD300

Нейронная сеть	AP	AP	AP
	(«рак»)	(«ранний рак»)	(«иная патология»)
SSD300+VGG19	0,805	0,873	0,284
SSD300+MobileNetV2	0,625	0,625	0,282
SSD300+Xception	0,704	0,810	0,445
SSD300+VGG16	0,707	0,896	0,409
RetinaNet	0,869	0,968	0,586
EfficientDet+ EfficientNetB0	0,889	0,911	0,902

Лучшие результаты по метрике точности показала HC EfficientDet. Для вариантов широко распространенной нейросетевой архитектуры SSD результаты по средней точности значительно ниже, особенно для класса «иная патология».

Важным моментом для практической работы всех нейросетевых алгоритмов, включая и алгоритмы детектирования патологий, является обеспечение их устойчивости к различным негативным факторам (робастности). Для тестирования по этому критерию использовался следующий подход. Нейросетевые модели обучались только на изображениях с эндоскопа OLYMPUS EXERA III GIF-HQ190, а тестирование проводилось только на выборке изображений с более ранних моделей эндоскопов OLYMPUS 160, 170 и 180 серий. Результаты такого эксперимента приведены в таблице 5.27. В ячейках таблицы отражены результаты метрики AP и уровень этих показателей относительно результатов (показатель деградации алгоритма), приведенных в таблице 5.26. Также приведен средний показатель деградации алгоритма для трех указанных классов.

Таблица 5.27 — Анализ робастности алгоритмов при обучении и тестировании на разных поколениях эндоскопического оборудования

Нейронная сеть	AP	AP	AP	Средний
	(«рак»)	(«ранний	(«иная	показатель
		рак»)	патология»)	деградации
	0,722	0,786	0,239	88.00/
22D200+70013	89,7%	90,0%	84,2%	88,0%
SSD200 MabilaNatV2	0,560	0,566	0,241	<u> </u>
SSD300+MobileNetv2	89,6%	91,4%	85,4%	88,8%
SSD200 Vegetion	0,628	0,732	0,361	86.004
SSD300+Aception	89,2%	90,4%	81,1%	80,9%
	0,644	0,824	0,351	80.6%
330300+70010	91,1%	92,0%	85,8%	89,070
PotineNot	0,794	0,843	0,516	00 5%
Retifiainet	91,3%	92,2%	88,1%	90,3%
EfficientDet+	0,797	0,806	0,815	80.5%
EfficientNetB0	89,7%	88,5%	90,4%	09,370

Приведенные результаты показывают, что рассчитанный показатель робастности для разных моделей НС находится в диапазоне 86,9% – 90,5%. Некоторое преимущество по этому показателю показывает архитектура RetinaNet,

опережая рассматриваемую сеть EfficientDet в среднем на 1% для трех классов. Однако абсолютные значения метрики AP для нейронной сети EfficientDet остаются наилучшими для такого «стресс-теста», хотя и показывают несколько более высокий уровень деградации результатов для классов «рак» и «ранний рак» [299, 300].

Время обработки одного кадра эндоскопической видеопоследовательности является еще одним важным практическим параметром работы нейросетевого алгоритма детектирования патологий желудка. Для подсчета времени обработки одного кадра реализован программный эксперимент по тестированию алгоритмов наборах обучающих выборок Для батчах _ ИЗ входных данных. на общее рассматриваемых нейросетевых алгоритмов рассчитано время тестирования, затраченное на обработку набора за 22 итерации. Также вычислено среднее время обработки одного видеокадра (отношение времени тестирования к произведению числа изображений в одном наборе и числа итераций). Результаты указанного программного эксперимента приведены в таблице 5.28.

Таблица 5.28 — Сравнение результатов работы исследуемых алгоритмов по скорости обработки одного кадра

Нейронная сеть	Общее время тестирования	Среднее время обработки
	(число итераций за 1 с)	одного кадра, с
SSD300+VGG19	13,485 c	0,038 c
	(1,61 итер/с)	
SSD300+ MobileNetV2	13,727 c	0,039 c
	(1,620 итер/с)	
SSD300+Xception	12,644 c	0,036 c
	(1,74 итер/с)	
SSD300+VGG16	13,566 c	0,039 c
	(1,6 итер/с)	
EfficientDet+	17,720 c	0,050 c
EfficientNetB0	(1,24 итер/с)	

Учитывая приведенные значения, можно сделать вывод о том, что наилучшим подходом по скорости работы среди перечисленных моделей НС можно считать модель SSD с базовой сетью Хсерtion. Следует отметить, что значения, приведенные в таблице 5.28 с практической точки зрения различаются между собой незначительно. Хотя некоторое отставание в процессе обработки кадра (число итераций в секунду) алгоритмом на основе архитектуры EfficientDet составляет порядка 30%, но это не является критичным, что подтверждается расчетом среднего времени обработки кадра [301].

5.7 Исследование алгоритмов детектирования на базе нейронной сети YOLO с базовыми параметрами

В 2020 году, на момент появления ранних версий нейросетевого детектора YOLO, их производительность (в части точности, робастности и/или скорости в определенных сценариях) часто уступала устоявшимся решениям на базе рассмотренной выше нейронной сети SSD. В связи с этим, для первичных исследований и разработки начальных прототипов СППВР выбран алгоритм на базе сети SSD как более надежный и проверенный вариант. Однако, начиная примерно с 2022 года, благодаря значительным архитектурным улучшениям в новых версиях YOLO (таких как YOLOv5, YOLOv6, YOLOv7, YOLOv8), эти детекторы стали демонстрировать превосходство над SSD по ключевым метрикам (точность mAP, скорость FPS, эффективность использования аппаратных ресурсов) [109, 113].

Параллельно с этим, актуальным направлением стали архитектуры, таких, трансформерах, как DETR (DEtection основанные на например, TRansformer) и его производные (например, DINO, Deformable DETR) [119-123]. Эти модели показывают хорошие результаты на крупных бенчмарках благодаря своей способности моделировать глобальный контекст сцены. Хотя традиционно для эффективного обучения таких моделей требуются очень большие объемы размеченных данных, современные методы аугментации данных и техники обучения с использованием «слабой» разметки позволяют существенно снизить

эту зависимость. Новые методы аугментации изображений, такие как RandAugment, MixUp, CutMix, Mosaic и их комбинации эффективно увеличивают разнообразие и размер обучающего набора, улучшая обобщающую способность модели даже при ограниченных исходных данных [105].

Учитывая значительный прогресс как в развитии YOLO-детекторов, так и в области трансформерных архитектур, а также доступность новых методов для аугментации данных, актуальной задачей становится проведение сравнительного тестирования. Необходимо оценить производительность не только новых YOLO-детекторов (например, YOLOv7), но и современных трансформерных моделей в задаче детектирования патологий на изображениях желудка.

Результаты тестирования HC на базе SSD, сети EfficientDet с базовой частью EfficientNetB0 и базовой модели YOLOv7 без использования аугментации на расширенной базе изображений патологий желудка приведены в таблицах 5.29–5.32. Они показывают, что наилучшие результаты среди моделей на базе SSD, как и ранее в этой главе показывает нейронная сеть SSD + VGG16-300, однако эти результаты существенно уступают аналогичным, получаемым при использовании более современной сети EfficientDet. Ряд ключевых метрик (F1, mAP, Recall) для сети YOLOv7 оказываются также заметно ниже, что ставит вопрос о необходимости оптимизации параметров соответствующего алгоритма.

Нейронная сеть	loss	val_F1	val_ loss	val_ mAP	val_ Precision	val_ Recall	last_ epoch
SSD + VGG16-300	4,375	0,711	4,812	0,644	0,706	0,632	18
SSD + VGG16-512	4,190	0,658	4,440	0,624	0,686	0,632	13
SSD + MobileNet-300	4,611	0,561	4,852	0,475	0,842	0,421	13
SSD + ResNet50	5,500	0,638	5,349	0,347	0,710	0,579	13

Таблица 5.29 – Тестирование нейронных сетей на базе SSD на расширенной базе изображений

Нейронная	loss	val_F1	val_	val_	val_	last_
сеть			mAP	Precision	Recall	epoch
EfficientDet+	0.219	0.872	0.802	0.949	0.949	24
EfficientNetB0	0,210	0,072	0,095	0,040	0,040	24

Таблица 5.30 – Тестирование нейронной сети EfficientDet на расширенной базе изображений

Таблица 5.31 – Тестирование нейронныой сети YOLOv7 на расширенной базе изображений

Нейронная	val_loss	val_F1	val_	val_	val_Re-call	last_
сеть			mAP	Precision		epoch
YOLOv7						
базовая без	0,102	0,793	0,811	0,891	0,717	56
аугментации						

Таблица 5.32 – Значения метрики АР для различных классов на расширенной базе изображений

Нейронная сеть	AP	AP	AP
	(«рак»)	(«ранний рак»)	(«иная патология»)
RetinaNet	0,873	0,976	0,524
SSD + VGG16-300	0,642	0,937	0,453
EfficientDet+ EfficientNetB0	0,872	0,860	0,815
YOLOv7 базовая без аугментации	0,906	0,857	0,671

5.8 Оптимизация параметров для нейронной сети YOLO

Проведем исследования по настройке параметров нейронной сети YOLOv7, аналогичные проведенным выше для сети SSD в п. 5.4. Эксперименты будут касаться исследованиям по выбору политики обучения и исследованию параметров аугментации.

5.8.1 Исследования по выбору политики обучения

Рассмотрим обучение базовой модели YOLOv7 при использовании трех различных оптимизаторов, описание которых приведено в п. 5.4. Результаты в таблице 5.33 показывают, что метрики для алгоритмов стохастического градиентного спуска (SGD) и Adam практически идентичны, а для алгоритма RMSProp сильно ниже. С учетом этого для дальнейших исследований выбран алгоритм SGD. Графики динамики метрик процесса обучения для него приведены на рисунок 5.21. Они показывают сходимость всех рассматриваемых метрик в районе 40-50 эпохи обучения.

Тип	loss	val_F1	val_loss	val_mAP	val_	val_	last_epoch
оптимиза-					Precision	Recall	
тора							
SGD	0,030	0,793	0,094	0,82	0,873	0,728	44
RMSProp	0,078	0,538	0,291	0,64	0,456	0,662	72
Adam	0,020	0,793	0,102	0,811	0,891	0,717	56

Таблица 5.33 – Исследование по выбору оптимизатора при обучении модели YOLOv7



Рисунок 5.21 – Динамика значений метрик от числа эпох обучения для нейронной сети YOLOv7 при использовании оптимизатора SGD



Рисунок 5.21 – Динамика значений метрик от числа эпох обучения для нейронной сети YOLOv7 при использовании оптимизатора SGD

5.8.2 Исследования по выбору типа аугментации

В базовой модели YOLOv7 могут использоваться следующие ключевые настройки аугментации данных в процессе обучения [105]:

Аугментация в цветовых координатах HSV:

- hsv_h – изменение тона (диапазон: 0,0–0,1);

- hsv_s изменение насыщенности (диапазон: 0,0–0,9);
- hsv_v изменение яркости (диапазон: 0,0–0,9).

Геометрические трансформации:

- degrees вращение изображения (например, ±10 градусов);
- translate сдвиг (например, 0,1 = 10% от размера изображения);
- scale масштабирование (например, $0.5 = \pm 50\%$ изменения размера);
- shear наклон (например, ± 2 градуса);
- perspective искажение перспективы (диапазон: 0,0–0,001).

Отражения:

- flipud вертикальное отражение (вероятность: 0,0–1,0);
- fliplr горизонтальное отражение (вероятность: 0,5 по умолчанию).

Мозаичные преобразования (Mosaic):

Представляют собой комбинирование 4 изображений в одно (вероятность 1,0 включена по умолчанию). Такое преобразование улучшает детекцию мелких объектов и контекстное обучение.

Смешивание (МіхUр):

Смешивание двух изображений с линейной интерполяцией (вероятность по умолчанию 0,1 задается в гиперпараметрическом файле hyp.scratch.yaml).

Рассмотрим влияние некоторых из этих параметров аугментации на метрики работы нейросетевого детектора. Исследования при аугментации в цветовых координатах HSV приведены в таблицах 5.34–5.37. Результаты позволили выбрать рекомендуемые значения сдвигов с учетом специфики эндоскопических изображений: hsv_h=0,02; hsv_s=0,7; hsv_v=0,3. Результаты по метрике AP для различных классов показывают увеличение значений для классов «ранний рак» и «иная патология» по сравнению с базовой моделью.

Изменение тона	loss	val_	val_loss	val_	val_	val_	last_
		F1		mAP	Precision	Recall	epoch
0,01	0,016	0,796	0,092	0,792	0,861	0,743	80
0,02	0,026	0,809	0,104	0,814	0,853	0,776	65
0,03	0,018	0,783	0,103	0,771	0,873	0,712	82
0,1	0,020	0,793	0,091	0,799	0,852	0,750	66
0,2	0,019	0,802	0,092	0,790	0,891	0,729	65

Таблица 5.34 – Достигаемые метрики при изменении тона hsv_h

Таблица 5.35 – Достигаемые метрики при изменении насыщенности hsv_s

Изменение насыщенности	loss	val_ F1	val_loss	val_ mAP	val_ Precision	val_ Recall	last_ epoch
0,6	0,016	0,801	0,086	0,788	0,860	0,755	81
0,7	0,026	0,792	0,089	0,813	0,858	0,742	44
0,8	0,018	0,796	0,096	0,781	0,869	0,738	72

Таблица 5.36 – Достигаемые метрики при изменении яркости hsv_v

Изменение яркости	loss	val_ F1	val_loss	val_ mAP	val_ Precision	val_ Recall	last_ epoch
0,2	0,025	0,790	0,100	0,804	0,877	0,721	49
0,3	0,023	0,811	0,099	0,828	0,860	0,773	53
0,4	0,023	0,793	0,093	0,799	0,880	0,723	67

Аугментация в цветовых координатах	AP	АР	АР
HSV	(«рак»)	(«ранний рак»)	(«иная патология»)
hsv_h=0,02; hsv_s=0,7; hsv_v=0,3	0,906	0,888	0,691

Таблица 5.37 – Значения метрики АР для различных классов при использовании аугментации в цветовых координатах HSV

Исследования при геометрических трансформациях приведены в таблицах 5.38–5.41. Они показывают, что для наилучшие результаты достигаются при величине сдвига translate=0,3. При этом значении существенно увеличиваются значения метрики AP для всех трех классов. При масштабировании наилучшим выбором является значение параметра scale=0,8. При этом значении также видна положительная динамика метрики AP для всех трех классов.

Величина сдвига	loss	val_	val_loss	val_	val_	val_	last_
		F1		mAP	Precision	Recall	epoch
0,2	0,031	0,818	0,087	0,828	0,865	0,778	68
0,3	0,026	0,863	0,085	0,878	0,865	0,863	60
0,4	0,027	0,854	0,081	0,871	0,897	0,816	90

Таблица 5.38 – Достигаемые метрики при сдвиге (translate)

Таблица 5.39 – Значения метрики АР для различных классов при использовании сдвига (translate)

Величина сдвига	AP	AP	AP
	(«рак»)	(«ранний рак»)	(«иная патология»)
0,3	0,936	0,946	0,751

Величина	loss	val_	val_loss	val_	val_	val_	last_
масштабирования		F1		mAP	Precision	Recall	epoch
0,7	0,035	0,824	0,086	0,831	0,871	0,783	103
0,8	0,031	0,843	0,077	0,879	0,901	0,793	119
0,9	0,034	0,856	0,080	0,869	0,884	0,829	116

Таблица 5.40 – Достигаемые метрики при масштабировании (scale)

Таблица 5.41 – Значения метрики АР для различных классов при использовании масштабирования (scale)

Величина	АР	АР	AP
масштабирования	(«рак»)	(«ранний рак»)	(«иная патология»)
0,8	0,948	0,946	0,744

Исследования при горизонтальном отражении (fliplr=0,5) приведены в таблицах 5.42 и 5.43. Здесь также наблюдается рост значений метрики AP для всех трех классов, но менее значительный, чем в предыдущем случае.

Таблица 5.42 – Достигаемые метрики при использовании горизонтального отражения

Горизонтальное	loss	val_	val_loss	val_	val_	val_	last_
отражение		F1		mAP	Precision	Recall	epoch
fliplr=0,5	0,023	0,808	0,095	0,819	0,881	0,748	78

Таблица 5.43 – Значения метрики АР для различных классов при использовании горизонтального отражения

Горизонтальное	AP	AP	AP
отражение	(«рак»)	(«ранний рак»)	(«иная патология»)
fliplr=0,5	0,914	0,849	0,694

Исследования при мозаичном преобразовании и смешивании приведены в таблицах 5.44–5.46. Результаты показывают, что мозаичное преобразование существенно увеличивает значения метрик AP для всех трех рассматриваемых классов. Смешивание не оказывает существенное влияние на рост метрик качества.

Таблица 5.44 – Достигаемые метрики при использовании мозаичного преобразования

Тип аугментации	loss	val_ F1	val_loss	val_ mAP	val_ Precision	val_ Recall	last_ epoch
Мозаичное преобразование	0,029	0,873	0,071	0,886	0,92	0,83	68

Таблица 5.45 – Значения метрики АР для различных классов при использовании мозаичного преобразования

Тип аугментации	AP	AP	AP
	(«рак»)	(«ранний рак»)	(«иная патология»)
Мозаичное преобразование	0,956	0,959	0,745

Таблица 5.46 – Достигаемые метрики при использовании смешивания

Вероятность	loss	val_	val_loss	val_	val_	val_	last_
применения		F1		mAP	Precision	Recall	epoch
0,1	0,020	0,805	0,094	0,808	0,870	0,753	61
0,2	0,019	0,795	0,095	0,789	0,889	0,722	81
0,3	0,020	0,785	0,088	0,753	0,881	0,711	82

Описанные выше исследования позволили выбрать оптимизированные параметры аугментации. Их суммарное использование позволяет достигать значений метрик, приведенных в таблицах 5.47–5.48. Таким образом, процедура оптимизации процесса аугментации позволила увеличить метрику AP для класса «рак» на 0,057; для класса «ранний рак» на 0,122; для класса «иная патология» на 0,118, что является существенным увеличением производительности алгоритма детектирования на базе сети YOLOv7.

Таблица 5.47 – Достигаемые метрики дял алгоритма детектирования YOLOv7 с оптимизированными параметрами аугментации

Значения параметров	loss	val_ F1	val_loss	val_ mAP	val_ Precision	val_ Recall	last_ epoch
hsv_h=0,02; hsv_s=0,7; hsv_v=0,3 translate = 0,3 scale = 0,8 fliplr = 0,5	0,032	0,89	0,060	0,911	0,916	0,865	108
mosaic							

Таблица 5.48 – Значения метрики AP алгоритма детектирования YOLOv7 с оптимизированными параметрами аугментации

Значения параметров	AP («рак»)	АР («ранний рак»)	АР («иная патология»)
hsv_h=0,02;	0,963	0,979	0,789
hsv_s=0,7;			
hsv_v=0,3			
translate $= 0,3$			
scale $= 0,8$			
fliplr = 0,5			
mosaic			

Графики динамики метрик процесса обучения для выбора оптимизированных параметров аугментации приведены на рисунке 5.22. Они показывают сходимость

всех рассматриваемых метрик на более поздних эпохах обучения к существенно большим значениям.



Рисунок 5.22 – Динамика значений метрик от числа эпох обучения для нейронной сети YOLOv7 при использовании оптимизированных значений параметров аугментации



Рисунок 5.22 – Динамика значений метрик от числа эпох обучения для нейронной сети YOLOv7 при использовании оптимизированных значений параметров аугментации

Матрица ошибок для детектора на базе нейронной сети YOLOv7 при тестировании на базе изображений «С» приведена на рисунке 5.23



Рисунок 5.23 – Матрица ошибок детектора на базе нейронной сети YOLOv7 при тестировании на базе изображений «С»

5.9 Итоговое сравнение результатов работы нейросетевых детекторов

YOLOv7 – семейство моделей, а не одна конкретная архитектура. Выбор конкретной сети зависит от требований к точности, скорости и доступным вычислительным ресурсам. Официальными моделями из репозитория YOLOv7 являются следующие [113]:

- YOLOv7 базовая модель.
- YOLOv7-X увеличенная версия (extra large) с более высокой точностью, но большей вычислительной сложностью.
- YOLOv7-W6 версия, разработанная для входного разрешения 1280х1280 пикселей (wide). Используется как основа для каскадных моделей.
- YOLOv7-E6 основана на YOLOv7-W6, включает дополнительные улучшения.
- YOLOv7-D6 делает YOLOv7-Е6 глубже (Deeper).

- YOLOv7-E6E расширенная (Extended) и тщательно дообученная (Evolved) версия YOLOv7-E6.
- YOLOv7-tiny легкая и быстрая модель для ресурсоограниченных устройств (edge devices).

на Проведем тестирование перечисленных моделей базе «C» с оптимизированных параметров. Результаты исследования использованием В таблицах 5.49–5.50. Сравнение метрик показывает, приведены ЧТО на относительно небольшой базе изображений желудка существенно не проявляются различия между разными моделями из семейства YOLOv7. Небольшое преимущество показывает по большинству метрик сеть YOLOv7-W6, но выбор базовой сети для практической реализации СППВР следует считать обоснованным.

Нейронная сеть	loss	val_ F1	val_loss	val_ mAP	val_ Precision	val_ Recall	last_ epoch
YOLOv7 базовая с оптимизированными параметрами аугментации	0,032	0,890	0,060	0,912	0,916	0,865	108
YOLOv7-X	0,028	0,894	0,057	0,903	0,918	0,872	100
YOLOv7-W6	0,034	0,898	0,043	0,912	0,910	0,887	88
YOLOv7-E6	0,031	0,897	0,037	0,904	0,899	0,894	98
YOLOv7-D6	0,033	0,896	0,035	0,912	0,915	0,878	83
YOLOv7-E6E	0,027	0,899	0,037	0,906	0,928	0,874	96
YOLOv7-tiny	0,032	0,847	0,033	0,847	0,827	0,813	64

Таблица 5.49 – Анализ работы нейронных сетей из семейства YOLOv7 на базе «С»

Нейронная сеть	AP	AP	AP	
	(«рак»)	(«ранний рак»)	(«иная патология»)	
YOLOv7 базовая с оптимизированными параметрами аугментации	0,963	0,979	0,789	
YOLOv7-X	0,952	0,975	0,783	
YOLOv7-W6	0,942	0,997	0,798	
YOLOv7-E6	0,948	0,986	0,780	
YOLOv7-D6	0,960	0,975	0,741	
YOLOv7-E6E	0,947	0,989	0,783	
YOLOv7-tiny	0,927	0,975	0,742	

Таблица 5.50 – Значения метрики АР для сетей из семейства YOLOv7

Проведем итоговое сравнение трех подходов к детектированию объектов на изображениях желудка: базовой сети YOLOv7 с оптимизированными параметрами, алгоритма EfficientDet и сети трансформенного типа, рассмотренной выше на базе архитектуры D-FINE-L. Результаты исследования приведены в таблицах 5.51–5.52.

Таблица 5.51 – Анализ работы различных нейросетевых архитектур на базе «С»

Нейронная сеть	loss	val_ F1	val_ mAP	val_ Precision	val_ Recall	last_ epoch
YOLOv7 базовая	0,032	0,890	0,912	0,916	0,865	108
EfficientDet	0,218	0,872	0,953	0,898	0,848	24
D-FINE-L	0,032	0,943	0,897	0,924	0,973	80
Нейронная сеть	АР («рак»)	АР («ранний рак»)	АР («иная патология»)			
----------------	---------------	----------------------	--------------------------			
YOLOv7 базовая	0,963	0,979	0,789			
EfficientDet	0,889	0,911	0,902			
D-FINE-L	0,898	0,995	0,842			

Таблица 5.52 – Значения метрик АР для различных нейросетевых архитектур

Приведенные результаты показывают, что все три рассмотренных архитектурных подхода имеют высокие и сопоставимые между собой значения метрик. Выбор между ними для итоговой реализации СППВР должен учитывать другие параметры производительности.

5.10 Краткие выводы

1. Разработанные алгоритмы детектирования и классификации патологий желудка решают критически важную медицинскую задачу – ранней диагностики онкологических заболеваний. Автоматизация этого процесса способна повысить точность диагностики и расширить доступность процесса обнаружения новообразований.

2. Создана и размечена уникальная база эндоскопических изображений желудка (14 135 аннотированных снимков), включающая:

– Патологии в двух режимах съемки (с увеличением ×65 и без него).

– 12 классов патологий (рак, ранний рак, полипы, метаплазия и др.).

Были использованы следующие входные критерии качества (разрешение от 624×528 пикселей, отсутствие артефактов, контроль размытия и бликов и др.) и методика извлечения кадров из видеоданных.

3. Сравнение современных нейросетевых моделей выявило три ведущих архитектуры:

- Сеть EfficientDet + EfficientNetB0: максимальная точность (mAP = 0,953; F1=0,872), хорошая производительность.
- Сеть YOLOv7: баланс скорости и качества (mAP=0,89; F1=0,890) высокая точность предсказания классов «рак» и «ранний рак».
- Трансформерная архитектура (D-FINE-L) лидирует по метрике F1= 0,943, показала высокую точность для раннего рака (AP=0,995) и является явным лидером для класса «иные патологии» AP=0,902).

Выбор между ними зависит от этапа проектирования/реализации СППВР. Для начальных исследований хорошим выбором являются сети EfficientDet и YOLO. Для опытной эксплуатации предпочтительным является использование архитектуры YOLO, представляющей законченное промышленное решение. Для коммерческой версии целесообразным является выбор лицензионно открытой трансформерной модели D-FINE-L, которая также может использоваться в проектах по импортозамещению эндоскопического оборудования.

4. Была проведена оптимизация параметров процесса обучения нейронных сетей. Основные результаты:

- Оптимизатор: алгоритм Adam в большинстве случаев превзошел алгоритмы SGD/RMSProp.
- Аугментация: дает существенный прирост mAP для сети YOLOv7.
- Предобработка: деинтерлейсинг видео повысил mAP на 6,8% (особенно для режима с увеличением).

5. Выявлены следующие практические ограничения и проблемы:

- Низкое качество кадров из видео (размытие при перистальтике, блики).
- Малый объем данных для некоторых патологий (например, всего 29 карциноидов).
- Трудоемкость точной разметки демаркационных линий.

6. Разработанные нейросетевые алгоритмы соответствует требованиям реальной клинической практики:

- Скорость обработки кадров ≥20 FPS.
- Точность детекции раннего рака: АР ≥0,90 у лучших моделей.
- Подтверждена робастность нейросетевых алгоритмов при замене эндоскопа (OLYMPUS 160-190 серий).

ГЛАВА 6

РЕЗУЛЬТАТЫ ТЕСТИРОВАНИЯ АППАРАТНО-ПРОГРАММНЫХ КОМПЛЕКСОВ НА ЭНДОСКОПИЧЕСКИХ ВИДЕОИЗОБРАЖЕНИЯХ

Разработанные для эндоскопических аппаратно-программных комплексов алгоритмические решения прошли тестирование, включавшее оценку:

- функционала программного обеспечения при работе с нейросетевыми модулями детектирования, сегментации и классификации эндоскопических видеоизображений в режиме выбора различных источников видеоданных;
- возможность управления воспроизведением видеофайла и его записи на внутренний накопитель в режиме работающего нейросетевого модуля детектирования, сегментации и классификации патологий;
- возможность выгрузки численных характеристик и построения отчетов и уведомлений на основе выгруженных численных характеристик в медикоинформационную систему «ЭМИС».

Испытания проводились на высокопроизводительном компьютере следующей конфигурации: PowerCase 4U 850W, CPU Intel Core i9-10900K iB560/ 32Gb/SSDM2 2 x 1000Gb/ GPU – Nvidia RTX4080 стоимостью порядка 3 500 USD на момент написания настоящей работы. Также на этом компьютере запускалась медикоинформационная система. Для вычислительной оценки сложности ряда нейросетевых алгоритмов использовался мощный сервер (A-VSN-1) AMD Ryzen Threadripper PRO 3995WX с 64 ядрами и 128 потоками, 512 ГБ оперативной памяти, 4 видеокартами Nvidia Quadro RTX A5000 с 24 ГБ видеопамяти, предоставленный Центром ИИ ЯрГУ. Стоимость такого сервера в настоящий момент составляет порядка 50 000 USD.

Для испытаний использовались эндоскопические видеоданные, полученные с использованием эндоскопов GIF-HQ190 и GIF-HQ180, а также эндоскопической системы OLYMPUS EXERA III в режимах без оптического увеличения и с

оптическим увеличением. Результаты детектирования патологий разных классов выводились на монитор, а также загружались в ЭМИС. Рассмотрим результаты тестирования алгоритмических решений, разработанных в главах 3-5 настоящей работы.

6.1 Тестирование алгоритмов детектирования купола слепой кишки

Проведено тестирование рассмотренных в главе 3 алгоритмов детектирования купола слепой кишки на видеоизображениях. Для этого собран дополнительный набор из 117 видеопоследовательностей, из которых 100 содержали КСК, а 17 – его не содержали. Тестовая база имела следующие характеристики: средняя длительность видео – 8 мин 27 сек, количество кадров в секунду – 25, среднее нахождение области купола на записи – 9,5 сек. На рисунке 6.1 представлена иллюстрация работы разработанных алгоритмов (базовый АДК, с трекингом – АДКТ, с трекингом и оптическим потоком – АДКТ-ОП).

Исходя из представленных данных видно, что каждая последующая модернизация алгоритма позволяет сузить число предполагаемых точек с куполом слепой кишки на видеопоследовательности, а, следовательно, заключение системы о качестве проведения эндоскопического обследования становится наиболее точным и полным.

В таблице 6.1 представлены результаты тестирования разработанных версий алгоритма на независимых данных, не участвующих в обучении нейронных сетей. Для оценки точности предсказанного интервала не только с точки зрения покадровых метрик, но и со стороны корректности временного интервала использована метрика *IoU*, которая модифицирована для одномерного случая и вместо площадей ограничивающих объект рамок сравнивала временные метки и расстояния между ними [305].

Проведя анализ результатов, можно сделать вывод, что разработанный подход на основе детектора с алгоритмом трекинга и дополнительного классификатора направления оптического потока показывает наибольшую точность работы и,

несмотря на самую большую задержку на обработку одного кадра (в 4 раза выше, чем у базового алгоритма), его быстродействия достаточно для работы в режиме реального времени.



Рисунок 6.1 – Иллюстрация работы алгоритмов детектирования КСК: а) исходный интервал с куполом слепой кишки на тестовой записи;

б) пример работы базового АДК; в) пример работы АДКТ; г) пример работы АДКТ-ОП

Модификация алгоритма	F1-мера (покадровая)	IoU (по временному интервалу)	Время обработки одного кадра, (с)
АДК	0,41	0,33	0,008
АДКТ	0,75	0,56	0,025
АДКТ-ОП	0,89	0,83	0,033

Таблица 6.1 – Результат тестирования АДК

6.2 Тестирование алгоритмов детектирования колоректальных полипов

Проведено тестирование рассмотренных главе 4 алгоритмов В детектирования видеоизображениях колоректальных полипов на колоноскопических исследований. Для анализа алгоритма детектирования полипов непосредственно на видеопотоке с эндоскопической системы был собран специальный стенд, оборудованный платой видеозахвата (рисунок 6.2). С его помощью в Пироговском центре (ПРГВ) с сентября 2021 года по март 2022 года записаны 450 видеопоследовательностей колоноскопических исследований разрешением 1400×1080 пикселей. Из них 207 записей содержали полипы, временные интервалы которых отмечались врачами с помощью педалей на тестовом стенде.

Также совместно с врачами эндоскопического отделения ЯОКОБ собран набор видеоизображений для обучения и тестирования алгоритмов более низкого разрешения. Видеоданные в базе ЯОКОБ брались из архива, в котором они перекодировались с использованием видеокодека Н.264, а их разрешение варьировалось от 626×532 до 760×436 пикселей. Обучающий набор состоял из 50 видеозаписей, тестовый набор – из 9 видеозаписей [306, 307].



Рисунок 6.2 – Схема разработанного стенда с модулем анализа видеопотока

Для учета явления корреляции между соседними кадрами видеопоследовательности разработан и протестирован алгоритм постобработки на основе трекинга выделенных областей. Логика его работы аналогична алгоритму трекинга при детектировании КСК [308].

Результаты тестирования доработанного алгоритма на базе трансформера D-FINE-L приведены в таблице 6.2. Точность обнаружения полипов на обеих базах превышает по метрикам уровень 0,8, соответствующий по данным статистики среднему уровню обнаружения полипов врачом в скрининговой практике.

Метрика	ЯОКОБ без	ЯОКОБ с	ПРГВ без	ПРГВ с трекингом
	трекинга	трекингом	трекинга	пп п в с трекингом
AP@0,50	0,833	0,877	0,783	0,825
Р	0,764	0,844	0,735	0,827
R	0,810	0,882	0,757	0,855
$\overline{F1}$	0,796	0,863	0,741	0,846

Таблица 6.2 – Результаты тестирования алгоритма детектирования полипов с трекингом

Проведен анализ скорости работы для алгоритмов сегментации колоректальных полипов, рассмотренных в главе 4. Его результаты отображены в таблице 6.3.

Нейронная сеть	Количество параметров модели, млн	Размер батча, изображений	GFLOPS	Пропускная способность, изображений/сек	Используемая память GPU, МБ
SSformer-S	29,57	1	20,00	93	524
SSformer-S	29,57	16	320,05	390	1330
SSformer-L	66,22	1	34,61	36	736
SSformer-L	66,22	16	553,77	221	1556
U-Net	24,44	1	14,81	259	456
U-Net	24,44	16	237,04	835	1126
U-Net++	26,08	1	34,84	193	532
U-Net++	26,08	16	557,36	360	2106
DUCK-Net- 17	38,92	1	91,83	8,35	1768
DUCK-Net- 17	38,92	16	1469,24	48	3306
DUCK-Net- 34	155,41	1	315,35	6,32	5378
DUCK-Net- 34	155,41	16	5045,54	28,26	13618
MMP	28,77	1	13,88	25,03	694
MMP	28,77	16	222,00	149,74	1562

Таблица 6.3 – Результаты тестирования алгоритмов сегментации полипов по скорости работы

Для оценки вычислительной сложности и исключения фактора зависимости от версии библиотек и фреймворков все модели предварительно сконвертированы в универсальный формат ONNX (Open Neural Network Exchange). Далее работа обученной нейронной сети (инференс) осуществлялась в одной и той же программной среде с использованием библиотеки (фреймворка) onnxruntime на высокопроизводительном сервере A-VSN-1. Сравнительный анализ проводился при запуске на одном графическом ускорителе (GPU) для батчей из 1 и 16 изображений. Запуск на центральном процессоре (CPU) не оценивался. Разрешение входных RGB-изображений на входе нейронной сети составляло 352х352 пикселя. Для проведения тестирования в оптимальных температурных режимах работы GPU осуществлен предварительный «прогрев» видеокарты на протяжении 500 итераций. Оценивание пропускной способности осуществлялось на среднем значении из 5 000 итераций (1 итерация равняется одному обработанному батчу, время загрузки изображений в память GPU и постобработки отброшено).

Результаты исследований показывают, что предложенный по результатам тестов в главе 4 алгоритм сегментации ММР позволяет проводить обработку в режиме реального времени (для батча размером 16 изображений). При этом модель не занимает большой объем в памяти графического ускорителя по сравнению с аналогами. Размер модели в 28,77 млн. параметров также лишь незначительно превосходит традиционный нейросетевой алгоритм U-Net и ряд его широко используемых модификаций. Все перечисленное позволяет рекомендовать алгоритм ММР для использований в аппаратно-программных эндоскопических комплексах для сегментации полипов на видеоизображениях.

6.3 Тестирование алгоритмов детектирования и классификации патологий желудка

Проведен анализ скорости работы для **алгоритмов детектирования патологий желудка**, рассмотренных в главе 5 [309–311]. Его результаты отображены в таблице 6.4.

Тестирование производилось на сервере A-VSN-1 с использованием аналогичных настроек и параметров, рассмотренных в п. 6.2. Для YOLOv7 измерено влияние NMS (non-maximum suppression) на задержку: операция добавляет 0,5-2 мс к времени вывода, пропорционально числу bbox и размеру батча. Приведены сравнительные результаты с отключенным и активным режимом NMS. Принцип работы D-FINE-L данной постобработки не требует. Размер входных RGB изображений составлял 640х640 пикселя. Оценивание пропускной способности осуществлялось на среднем значении из 5 000 итераций (1 итерация равняется одному обработанному батчу, время загрузки изображений в память GPU отброшено).

Нейронная	Количество	Размер	NMS	GFLOPS	Пропускная	Используе-
сеть	параметров	батча,			способность,	мая память
	модели,	изобра-			изображе-	GPU, MБ
	МЛН	жений			ний/сек	
YOLOv7	36,49	1	-	51,65	141	843
YOLOv7	36,49	16	-	826,34	290	2 425
YOLOv7	36,49	1	Есть	51,65	129	843
YOLOv7	36,49	16	Есть	826,34	211	2 425
D-FINE-L	30,67	1	-	45,68	82	744
D-FINE_L	30,67	16	-	730,96	192	2 734

Таблица 6.4 – Результаты тестирования алгоритмов детектирования патологий желудка по скорости работы

Результаты замеров производительности демонстрируют, что при включении режима NMS в модели YOLOv7, трансформерные модели D-FINE-L сокращают исходное отставание по скорости: хотя их базовая задержка на этапе инференса на 15-25% выше из-за вычислительной сложности базовой реализации механизма внимания. Необходимость в постобработке рамок для детектирования (NMS добавляет YOLOv7 дополнительное время на обработку кадра) выводит архитектуры на сопоставимый уровень производительности.

С учетом лицензионных ограничений для модели YOLO7 (и более старших) при реализации итоговой СППВР с целью анализа видеоизображений желудка рекомендована модель D-FINE-L.

6.4 Разработка выходного интерфейса обмена данными между модулем нейросетевых алгоритмов и сопряженной МИС

Финальным этапом процесса анализа эндоскопических видеоданных является сохранение результатов детектирования и анализа патологий в устройство хранения, а также вывод этих результатов на монитор в понятной и легко

интерпретируемой форме. Для этого необходимо разработать выходной интерфейс обмена данными между модулем нейросетевых алгоритмов и сопряженной МИС.

При выводе информации о результатах детектирования объекта интереса (патологии, полип, купол) на экран осуществляется загрузка координат обрамляющей рамки, вычисленных модулем нейросетевых алгоритмов, а также захваченного кадра. Далее в соответствии с вычисленными координатами на текущий кадр наносится разметка, и изображение выводится на монитор.

Возможны два способа вывода изображения на монитор:

- Так как эндоскопические видеоданные обрабатываются на видеокарте NVIDIA, которая имеет свои HDMI-видеовыходы, то возможен способ подключения обыкновенного монитора через HDMI-кабель непосредственно к самой видеокарте.
- 2) Для вывода результатов алгоритма может использоваться одна из плат DeckLink, имеющая как SDI-вход (для захвата видеопотока), так и SDI-выход, который может быть подключен непосредственно к SDI-монитору. Мониторы, которые используются в эндоскопической стойке, также являются SDIмониторами. Таким образом, в стойке будут установлены два SDI-монитора, SDI-интерфейс будет один ИЗ которых через выводить поток видеоэндоскопа, а на второй будут выводиться результаты с видеокарты NVIDIA.

В ЯОКОБ в эндоскопическом отделении на всех рабочих станциях врачейэндоскопистов установлена система ЭМИС. Она является функциональным решением для врача-эндоскописта, позволяющим в процессе проведения эндоскопического исследования производить видеофиксацию с сохранением записи исследования на сервер, а также выполнять фотофиксацию обнаруженных патологий нажатием кнопки на эндоскопе. Также система позволяет легко создавать и настраивать протоколы проведения исследований, сохраняет информацию о пациентах и проведенных исследованиях в централизованную базу данных (рисунок 6.3).



Рисунок 6.3 – Интерфейс системы ЭМИС

В ЭМИС встроена функция определения момента окончания исследования. При получении системой ЭМИС сообщения об окончании исследования, она через REST API обращается к разработанному веб-серверу, развернутому на том же компьютере, с запросом о предоставлении результатов алгоритмов детектирования и классификации. Кроме обращения ЭМИС к разрабатываемой системе для получения результатов только что завершившегося исследования, возможен вариант анализа ранее полученных данных, хранящихся в ЭМИС.

Таким образом, разработанный модуль выходного интерфейса обмена данными между модулем нейросетевых алгоритмов и сопряженной медикоинформационной системой может функционировать в следующих режимах:

- вывод кадра эндоскопического исследования с нанесенной разметкой,
 выделяющей патологии, на экран во время проведения исследования;
- загрузка результатов анализа видеопотока в ЭМИС по завершению исследования;
- обработка отдельных кадров, хранящихся в историях болезни пациентов, с последующей выгрузкой результатов в ЭМИС.

6.5 Разработка опытного образца системы

Для создания опытного образца СППВР в эндоскопии разработанные ранее отдельные модули соединены в единую структуру в соответствии со схемой, приведенной на рисунке 6.4.



Рисунок 6.4 – Структурная схема опытного образца системы поддержки принятия врачебного решения в эндоскопии

Опытный образец состоит из следующих модулей с собственной законченной функциональностью [312, 313]:

- 1. Модуль видеозахвата осуществляет захват видеоизображений с высоким разрешением (до 4K) и относительно высокой частотой смены кадров (до 60 fps) со специализированной платы видеозахвата при прямом подключении эндоскопической системы к компьютеру, разбиение видеопотока на отдельные кадры, их проверку, помещение прошедшего проверку кадра в буфер памяти.
- 2. Модуль анализа изображений состоит из модулей анализа и классификации эндоскопических изображений, а также дополнительного алгоритма, осуществляющего выборку каждого n-го кадра для анализа и отмечающего полученный на выходе модуля классификации результат (координаты рамки и метку класса) на отобранных и пропущенных кадрах.
- 3. **Модуль АРІ** осуществляет взаимосвязь опытного образца с ЭМИС через специализированный протокол запросов/ответов.
- 4. **Модуль визуализации результатов** отображает найденные патологии на отдельном мониторе врача-эндоскописта.

При разработке опытного образца использовались следующие сторонние программные модули:

- Библиотека FFmpeg (поставляется в составе системы в собранном виде, лицензия на свободное использование LGPL-2). Правообладатель: публичная библиотека, общественное достояние. Репозиторий кода: https://github.com/FFmpeg/FFmpeg.
- 2. Библиотека OpenCV (поставляется в составе системы в собранном виде, лицензия на свободное использование LGPL-2). Правообладатель: публичная библиотека, общественное достояние. Репозиторий кода: https://github.com/opencv/opencv.
- Библиотека yaml-cpp (используется в составе системы в виде исходного кода, лицензия на свободное использование МІТ). Правообладатель: публичная библиотека, общественное достояние. Репозиторий кода: https://github.com/jbeder.

- 5. Библиотека rapidjson (поставляется в составе системы в виде исходного кода, лицензия на свободное использование МІТ). Правообладатель: публичная библиотека, общественное достояние. Репозиторий кода: https://github.com/Tencent/rapidjson.
- Библиотека QtFramework 5.12 (поставляется в составе дистрибутива Linux, лицензия на свободное использование LGPL-3). Правообладатель: публичная библиотека в составе Linux.
- 7. Библиотека SharedTools (поставляется в составе системы в виде исходного кода, лицензия на свободное использование МІТ). Правообладатель: публичная библиотека, общественное достояние. Репозиторий кода: https://github.com/hkarel/SharedTools.
- Библиотека PProtoCpp (поставляется в составе продукта в виде исходного кода, лицензия на свободное использование MIT). Правообладатель: публичная библиотека, общественное достояние. Репозиторий кода: https://github.com/hkarel/PProtoCpp.
- Библиотека CUDA-11.6. Проприетарная библиотека. Собственник NVIDIA. Бесплатных/публичных альтернатив нет. Поставляется корпорацией NVIDIA в составе дистрибутива Linux.
- Библиотека cudnn. Проприетарная библиотека. Собственник NVIDIA. Бесплатных/публичных альтернатив нет. Поставляется корпорацией NVIDIA в составе дистрибутива Linux.
- Библиотека tensorrt. Проприетарная библиотека. Собственник NVIDIA. Бесплатных/публичных альтернатив нет. Поставляется корпорацией NVIDIA в составе дистрибутива Linux.

Взаимодействие всех модулей осуществляется с помощью разработанных программных интерфейсов, обеспечивающих информационный обмен стандартным набором типов данных, поддерживаемых в языке программирования С++ и подключаемых библиотеках (STL, boost и др.) [314, 315].

При разработке опытного образца системы пройдены следующие процессы жизненного цикла:

- разработка;

- тестирование;

- опытная эксплуатация, обучение персонала;

- внедрение;

- эксплуатация.

В ходе эксплуатации система устанавливается на рабочем компьютере (сервере) эндоскопического центра. Обновления программного обеспечения заложены только в части дообучения нейронной сети. Рабочая операционная система (ОС), установленная при инсталляции, по умолчанию не меняется. В случае смены ОС приложения программы пересобираются под новую ОС, и программа направляется клиенту для переустановки.

Также, с установкой программы запускается специальный сервис, следящий за работой программы. Если она функционирует некорректно, то сервис перезапускает программу и может отправить уведомление разработчикам, которые дорабатывают программу в рамках технической поддержки.

Для работы системы необходимо организовать автоматизированное рабочее место со следующими характеристиками:

- Плата видеозахвата, поддерживающая запись несжатого 10-битного видео из источников с интерфейсами SDI и HDMI и обеспечивающая переключение между форматами видео SD и HD.
- Компьютер со следующими минимальными характеристиками: процессор архитектуры Intel/AMD, минимальное количество ядер: 8. Объем оперативной памяти: от 64 Гб. Видеокарта: GTX 4080Ti или более производительная, память не менее 12 Гб. Жесткий диск: доступное пространство минимум – 1 Тб, желательно 2 Тб и более. Сетевой адаптер для подключения к сети Ethernet/Internet. Монитор 22 дюйма или выше.

К необходимым программным требованиям для функционирования опытного образца системы относятся:

- OS Linux Ubuntu 20.04 LTS;
- GCC C++ 9.4 (идёт в комплекте с Ubuntu);

- Qt Framework 5.12.8 (идёт в комплекте с Ubuntu);
- Nvidia CUDA 11.6 (скачивается с https://developer.nvidia.com/cuda-11-6-0download-archive);
- Nvidia TensorRT 8 (скачивается с https://developer.nvidia.com/tensorrt-gettingstarted).

Для компиляции исходного текста в объектный код программного обеспечения используются механизм для сборки кода – компилятор GCC C++ 9.4.0 и механизм для построения приложений – Qt Framework 5.12.8, на котором написан интерфейс и базовые модули опытного образца. Данные механизмы работают на базе операционной системы Linux Ubuntu 20.04 LTS.

Для взаимодействия с GPU требуется использование следующих пакетов программного обеспечения: фреймворк для поддержки нейросетей – Nvidia CUDA 11.6; специализированный модуль Nvidia TensorRT 8.4, который позволяет использовать нейросетевую модель на конкретной GPU максимально эффективным образом, т.е. из открытого формата хранения весов нейронной сети ONNX перерабатывается в оптимальный бинарный код под конкретный GPU.

Входной информацией для системы являются:

- Файл конфигурации системы текстовый файл специального формата с описанием значений параметров, определяющих особенности ее эксплуатации.
- Видеосигнал, получаемый от эндоскопа цифровой компонентный некомпрессированный видеосигнал с канальным кодированием с модифицированным кодом без возвращения к нулю или набор изображений из медико-информационной системы в типовых растровых форматах изображений. Разрешение обрабатываемого видеоизображения должно быть не менее 720×576 и не более 1920×1080 пикселей. Частота видеоизображения 50 кадров в секунду с чересстрочной разверткой или 25 кадров в секунду с прогрессивной разверткой.

Выходные данные системы:

- Пространственные координаты областей, соответствующие возможным патологиям (для эндоскопических исследований с увеличением и без увеличения).
- Численная метрика от 0 до 1, характеризующая выполнение требований к изображениям в рамках эндоскопического исследования в рамках существующего регламента.

Система прошла тестирование с цифровым эндоскопическим оборудованием (специализированная стойка), позволяющим получать и сохранять качественные видеоизображения высокого разрешения (до 1920х1080 пикселей) с оптическим увеличением (до 100 крат) / без увеличения:

- 1. OLYMPUS EXERA II Q180;
- 2. OLYMPUS EXERA III HQ190;
- 3. OLYMPUS EXERA II Q160Z;
- 4. OLYMPUS EXERA II Q180Z;
- 5. OLYMPUS EXERA II Q180H;
- 6. OLYMPUS EXERA III HQ190;
- 7. OLYMPUS EXERA II Q160Z.

На рисунке 6.5 представлен пример интерфейса опытного образца системы. Он включает в себя поле с видеопотоком с камеры эндоскопа (1) и отмеченными на нем белыми рамками найденными полипами (2). В правой части экрана имеется информация о номере версии и коде ревизии программы для отслеживания актуальности, а также индикатор активации записи видео «Видеозапись активна» (3). Подсвечивающийся зелёным цветом индикатор означает, что видеозапись включена, выключенная подсветка – видеозапись выключена. В правом нижнем углу находятся кнопки включения полноэкранного режима и выхода из программы (4).



Рисунок 6.5 – Интерфейс опытного образца системы

В 2020-2024 гг. проведена опытная эксплуатация и доработка СППВР в эндоскопических отделениях ФГБУ «НМХЦ им. Н.И. Пирогова» Минздрава России (Пироговский центр, г. Москва), в ГБУЗ ЯО «Клиническая онкологическая больница» (г. Ярославль), в Медицинский центр диагностики и профилактики (г. Ярославль) (соответствующие документы 0 внедрении приведены В Приложении А). Система может быть использована как непосредственно во время эндоскопического исследования, так и для контроля качества исследования после его проведения и записи результатов в ЭМИС. Планируется дальнейшая доработка эксплуатация системы совместно с разработчиками ООО «А-Вижн» И (г. Ярославль), ООО «ЭМИС» (г. Ярославль).

6.6 Краткие выводы

Проведенные исследования позволяют сделать следующие основные выводы:

- 1. Протестирована точность алгоритмов детектирования купола слепой кишки. Разработанные алгоритмы (АДК, АДКТ, АДКТ-ОП) показали прогрессивный рост точности: F1-мера улучшилась с 0,41 до 0,89, метрика IoU для временных интервалов возросла с 0,33 до 0,83. Модернизация базового алгоритма с трекингом и оптическим потоком (АДКТ-ОП) обеспечила максимальную точность при задержке 0,033 с/кадр, что допустимо для сохранения режима реального времени.
- 2. Протестирована точность алгоритмов детектирования полипов. Интеграция методов трекинга в алгоритм детектирования D-FINE-L повысила метрику AP@0.50 на 4,4% (на базе ЯОКОБ) и 4,2% (на базе ПРГВ). Метрики Precision и Recall достигли значений 0,844–0,882, что несколько выше среднего уровня обнаружения полипов для начинающего врачаэндоскописта, проводящего скрининговую диагностику.
- 3. Оценка производительности для алгоритмов сегментации полипов. Разработанный алгоритм ММР продемонстрировал наилучший баланс скорости и ресурсов. Его пропускная способность 149,74 кадр/с (батч=16) при 28,77 млн параметров нейросетевой модели и потреблении 1562 МБ VRAM. Это в 1,8 раза быстрее традиционной нейросетевой модели U-Net++ при сопоставимой (и превосходящей) точности сегментации.
- 4. Сравнение скорости патологий для детекторов желудка. Рекомендуемая нейросетевая модель D-FINE-L, содержащая 30,67 млн. параметров, показала сопоставимую с моделью YOLOv7 производительность при включенном режиме NMS: 192 против 211 кадр/с (батч=16). Отсутствие необходимости в режиме NMS компенсирует 15-25% дополнительных вычислений И более высокую базовую задержку трансформерной архитектуры.

- 5. Разработка универсального интерфейса обработки и анализа видеоданных. Реализованы три режима интеграции со специализированной медико-информационной системой «ЭМИС»:
 - Вывод аннотированных кадров на монитор.
 - Автоматическая загрузка результатов по REST API.
 - Пакетная обработка архивных видеоданных.
- 6. Аппаратно-программные требования. Для функционирования опытного образца СППВР требуются:
 - GPU NVIDIA RTX 4080 (мин. 12 ГБ VRAM);
 - 64 ГБ ОЗУ;
 - Плата видеозахвата 4K/60 fps;
 - Стек программного обеспечения: Ubuntu 20.04, CUDA 11.6, TensorRT.
- 7. Клиническая валидация. Система успешно прошла тестирование и опытную эксплуатацию в Пироговском центре, ЯОКОБ и Медицинском центре диагностики и профилактики (2020–2024 гг.). Подтверждена работоспособность с эндоскопами OLYMPUS EXERA II/III. Партнерство с отечественными ИТ-компаниями ООО «А-Вижн» и ООО «ЭМИС» обеспечит дальнейшее развитие разработанной СППВР.

275 ЗАКЛЮЧЕНИЕ

Основные выводы и результаты диссертационной работы можно сформулировать в следующем виде.

- 1. Показано, что разработанные концепция, методы, системы, алгоритмы, методики, программное обеспечение, аппаратно-программные комплексы приводят к снижению доли пропущенных патологий, снижению влияния человеческого фактора, стандартизации заключений и контролю качества скрининговых и клинических исследований ЖКТ и тем самым способствуют повышению достоверности эндоскопических исследований желудочнокишечного тракта за счет применения методов глубокого машинного обучения для обработки и анализа видеоизображений. Таким образом, цель диссертационной работы достигнута.
- 2. Проведен критический анализ современного состояния в области систем автоматического анализа видеоизображений эндоскопических исследований ЖКТ. Он показал, что использование СППВР потенциально позволяет снизить временные и финансовые расходы на проведение исследований, повышает точность и эффективность диагностических процедур, облегчает и расширяет процесс обучения начинающих врачей-эндоскопистов.
- концепция СППВР ЖКТ. 3. Разработана построения В эндоскопии Врач, производящий визуальный анализ видеоизображений во время исследования, взаимодействует с первой системой, работающей в режиме реального времени (СППВР-РРВ). Поддержка в принятии решения для врача выводится на мониторе в виде обведенных (раскрашенных) областей, либо с помощью звуковых подсказок, фокусирующих его внимание. В процессе эндоскопического исследования сохраняется полная видеопоследовательность, а также набор статических изображений (кадров) по выбору врача, отражающих значимые области интереса. Данные загружаются в используемую МИС и взаимодействуют со второй системой в

режиме постобработки данных (СППВР-РПД). Она позволяет аннотировать проведенное исследование, проводить контроль качества медицинской процедуры в соответствии с ее регламентом, отбирать изображения с соответствующими областями интереса и др.

- 4. Усовершенствована архитектура системы ИИ, работающей в качестве программного ядра для СППВР в условиях практической ограниченности входного набора аннотированных эндоскопических видеоизображений. Особенностью системы являются дополнительные программно-аппаратные модули, позволяющие получать результат по точности обнаружения и локализации патологий на видеоданных, сопоставимый со средним результатом, показываемым врачом-эндоскопистом во время диагностики.
- 5. Созданы и зарегистрированы базы эндоскопических изображений с разметкой, аннотированием и врачебной валидацией областей появления патологий:
 - «IntestinesImages база изображений кишечника при проведении колоноскопических исследований».
 - «StomachImages база изображений желудка при проведении гастроскопических исследований».

Их использование в СППВР закладывает фундамент для дальнейшего совершенствования методов ИИ и алгоритмов анализа видеоизображений.

- 6. Разработана и апробирована методология физической аугментации и полуавтоматической разметки базы эндоскопических видеоданных для обучения нейросетевых моделей. Ее использование снижает трудозатраты при создании больших размеченных баз видеоизображений на 35-45%.
- 7. Созданы методы и алгоритмы предобработки, фильтрации, оценки качества и постобработки изображений, учитывающие специфику эндоскопических видеоданных. Их использование в ряде задач колоноскопических и гастроскопических исследований позволяет повысить метрики точности системы ИИ на 3-7%.

- 8. Разработан метод контроля качества скринингового колоноскопического исследования. Для ее решения разработан алгоритм детектирования купола слепой кишки на видеоизображениях, а также алгоритм определения некачественных изображений (размытых, с артефактами сжатия, неинформативных), сохраненных в МИС.
- 9. Разработаны и исследованы нейросетевые алгоритмы детектирования, сегментации классификации патологий видеоизображениях И на эндоскопических исследований ЖКТ. Они базируются на стандартных нейросетевых архитектурах, но содержат доработки в части процедуры постобработке данных, пред-И изображений, выборе аугментации оптимальных параметров для обучения нейронной сети. Оригинальные алгоритмы сегментации областей интереса на изображениях защищены двумя изобретение Российской Федерации. Разработанные патентами на нейросетевые алгоритмы обработки и анализа эндоскопического видеопотока интегрированы в следующие зарегистрированные программы для ЭВМ:
 - «EndoscopyVA.RT программа для анализа эндоскопических видеоизображений в реальном времени на основе методов искусственного интеллекта».
 - «EndoscopyDSS.Hub программа для визуализации результатов эндоскопических исследований с учетом анализа видеоданных методами искусственного интеллекта».
- Разработано специализированное программное обеспечение для контроля качества, визуализации и сопоставления результатов эндоскопической диагностики. Проведено его тестирование, опытная эксплуатация и внедрение в Медицинский центр диагностики и профилактики (г. Ярославль), в ООО «А-Вижн» (г. Ярославль), в ООО «ЭМИС» (г. Ярославль).
- 11. Разработан аппаратно-программный комплекс для захвата, разметки, обработки и хранения видеоизображений эндоскопических исследований. Проведено его тестирование, опытная эксплуатация и внедрение в ФГБУ «НМХЦ им. Н.И. Пирогова» Минздрава России (Пироговский центр,

г. Москва) и в ГБУЗ ЯО «Клиническая онкологическая больница» (г. Ярославль).

12. Осуществлена интеграция разработанных нейросетевых алгоритмов обработки эндоскопических видеоизображений с внешней специализированной эндоскопической МИС.

Задачи диссертационной работы успешно решены. Разработанные концепция, методы, системы, алгоритмы, методики, программное обеспечение, аппаратнопрограммные комплексы позволяют обнаруживать патологии на видеопотоке с эндоскопической системы в режиме реального времени. Они используются в СППВР для частичной автоматизации диагностики, а также для контроля качества, обучения и подготовки профильных врачей-эндоскопистов.

279

СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ

- АДК Алгоритм детектирования купола;
- АДКТ Алгоритм детектирования купола с использованием трекинга;
- ВОЗ Всемирная организация здравоохранения;
- ДНК Дезоксирибонуклеиновая кислота;
- ЖКТ Желудочно-кишечный тракт;
- ЗНО Злокачественные новообразования;
- ИИ Искусственный интеллект;
- ИТ Информационные технологии;
- КРР Колоректальный рак;
- КСК Купол слепой кишки;
- КТ Компьютерная томография;
- МИС Медицинская информационная система;
- МКБ-10 Международная классификация болезней;
- МО Машинное обучение;
- НС Нейронная сеть;
- ОП Оптический поток;
- ПО Программное обеспечение;

ПРГВ – Национальный медико-хирургический центр им. Н.И. Пирогова;

- РП Рабочие процессы;
- РРВ Режим реального времени;
- СНС Сверточная нейронная сеть;

СППВР – Система поддержки принятия врачебных решений;

- СППР Система поддержки принятия решений;
- СТБ Сверточно-трансформенный блок;
- УЗИ Ультразвуковое исследование;
- ЦОИ Цифровая обработка изображений;
- ЭМИС Эндоскопическая медицинская информационная система;
- ЭМК Электронная медицинская карта;

ЭЧ – Эффект чересстрочности;

ЯОКОБ – Ярославская областная клиническая онкологическая больница;

Adam – Adaptive Moment Estimation (оценка адаптивного момента) – алгоритм оптимизации, используемый для обучения нейронной сети;

AP – Average Precision (средняя точность);

API – Application Programming Interface (программный интерфейс приложения) – совокупность способов и правил, по которым разные программы взаимодействуют между собой и обмениваются данными;

BBox – Bounded Box – ограничивающие прямоугольники при детекции объектов на изображениях;

ВСЕ – Binary Cross-Entropy (бинарная перекрёстная энтропия) – функция потерь, используемая в задачах бинарной классификации;

ВМР – BitMaP – формат растровых изображений;

BRISQUE – Blind/Referenceless Image Spatial Quality Evaluator – Алгоритм оценки качества изображений без эталона, работающий в пространственной области;

CCD – Charge-Coupled Device (прибор с зарядовой связью) – русский эквивалент ПЗС;

CE Mark – Conformité Européenne – Европейский сертификат соответствия;

CIR – Cecal Intubation Rate – процент достижения купола слепой кишки при колоноскопических исследованиях;

CLAHE – Contrast Limited Adaptive Histogram Equalization – алгоритм адаптивного выравнивания гистограммы изображений;

СМОЅ – Complementary Metal-Oxide-Semiconductor (комплементарный металло-оксидный полупроводник) – русский эквивалент КМОП;

CNN – Convolutional Neural Network (сверточная нейронная сеть);

DICOM – Digital Imaging and Communications in Medicine – медицинский отраслевой стандарт создания, хранения, передачи и визуализации цифровых медицинских изображений и документов обследованных пациентов;

EHR – Electronic Health Records – электронные медицинские карты;

FDA – Food and Drug Administration – Управление по санитарному надзору за качеством пищевых продуктов и медикаментов – агентство Министерства здравоохранения и социальных служб США;

FIT – Fecal Immunochemical Test (иммунохимический анализ кала) – используются для скрининга колоректального рака;

FOBT – Fecal Occult Blood Test (анализы кала на скрытую кровь) – используются для скрининга колоректального рака;

GDPR – General Data Protection Regulation (общие правила защиты данных) – законодательный акт Европейского союза, который определяет порядок сбора, обработки, хранения и распространения персональных данных в странах Евросоюза и с участием его граждан;

HD – High Definition (высокое разрешение);

НІРАА – Health Insurance Portability and Accountability Act (закон о переносимости медицинского страхования и подотчетности) – закон США, принятый в 1996 году, устанавливающий стандарты защиты конфиденциальности и безопасности медицинской информации;

HITL – Human-in-the-Loop (человек в цикле) – обозначение класса полуавтоматических методов разметки изображений;

HSV – (Hue – цветовой тон, Saturation – насыщенность, Value – значение цвета) – цветовая модель;

HTTP – HyperText Transfer Protocol (протокол передачи гипертекста);

ICV – Ileocecal Valve (илеоцекальный клапан);

IoU – Intersection over Union (пересечение через объединение) – отношение площади пересечения областей к площади объединения;

JSON – JavaScript Object Notation – текстовый формат обмена данными, основанный на языке JavaScript;

LLM – Large Language Model (большая языковая модель);

LRG – Linear Regression (линейная регрессия);

mAP – Mean Average Precision (средняя точность) – средняя по всем классам величина AP;

MOS – Mean Opinion Score – средняя экспертная оценка качества цифрового изображения;

МРТ – Магнитно-резонансная томография;

MS COCO – Microsoft Common Objects in Context (общие объекты в контексте) – набор данных, разработанный компанией Microsoft;

MSE – Mean Square Error (среднеквадратичная ошибка);

NBI – Narrow Band Imaging (узкоспектральная визуализация);

NCCN – National Comprehensive Cancer Network (Национальная комплексная онкологическая сеть) в США;

NLP – Natural Language Processing (обработка естественного языка);

NoSQL – Not only SQL (не только SQL) – общее название технологий управления нереляционными базами данных;

NSS – Natural Scene Statistics (статистика естественных изображений);

PACS – Picture Archiving and Communication System (система хранения и передачи изображений);

PR – Perceptron (персептрон) – архитектура нейронной сети;

ResNet – Residual Network (остаточная сеть) – архитектура нейронных сетей, которая использует остаточные соединения для обучения очень глубоких моделей;

REST API – Representational State Transfer Application Programming Interface (интерфейс прикладного программирования для передачи репрезентативного состояния) – архитектурный стиль взаимодействия между клиентом и сервером через HTTP;

RIS – Radiology Information System (радиологическая информационная система);

ROI – Region of Interest (область интереса);

SNOMED CT – Systematized Nomenclature of Medicine – (Систематизированная медицинская номенклатура);

SQL – Structured Query Language (язык структурированных запросов);

SSD – Single Shot MultiBox Detector – однопроходная архитектура нейронной сети для детекции объектов;

TCI – Time to Cecal Intubation (время интубации слепой кишки) – время, необходимое для достижения купола слепой кишки при колоноскопических исследованиях;

VGG – Visual Geometry Group (группа визуальной геометрии) – архитектура нейронной сети для детекции объектов;

WHO – World Health Organization (всемирная организация здравоохранения);

WLI – White Light Imaging (визуализация в белом свете);

YOLO – You Only Look Once – однопроходная архитектура нейронной сети для детекции объектов.

ЛИТЕРАТУРА

Злокачественные новообразования в России в 2023 году (заболеваемость и смертность) / под ред. А.Д. Каприна [и др.]. – М.: МНИОИ им. П.А. Герцена – филиал ФГБУ «НМИЦ радиологии» Минздрава России, 2024. – 276 с.

 Bray, F. Global cancer statistics 2022: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries / F. Bray, M. Laversanne, H. Sung, J. Ferlay, R. L. Siegel, I. Soerjomataram, and A. Jemal // CA: A cancer journal for clinicians. – 2024. – Vol. 74, №. 3. – P. 229–263.

3. Никонов, Е.Л. Раннее выявление онкологических заболеваний желудочно–кишечного тракта: возможности развития эндоскопической службы в Российской Федерации / Никонов Е.Л., Куваев Р.О., Жарова М.Е., Куваева А.А., Кашин С.В. // Доказательная гастроэнтерология. – 2024. – №13 (4). – С. 69–80.

4. Маметьева, Ю.А. Колоректальный рак у лиц молодого возраста. Эпидемиологическая ситуация в Ярославской области / Маметьева Ю.А., Завьялов Д.В., Камкина Г.В., Кашин С.В., Нестеров П.В. // Доказательная гастроэнтерология. – 2019. – Т. 8, № 4. – С. 68–75.

5. Lambert, R. Mass screening programs in Japan: What can we learn in the west? / Lambert R. // Endoscopy. – 1998. – Vol. 30, №8. – P.721–723.

6. Triantafillidis, J.K. Early–onset gastrointestinal cancer: An epidemiological reality with great significance and implications / Triantafillidis J.K., Georgiou K., Konstadoulakis M.M., Papalois A.E. // World Journal of Gastrointestinal Oncology. – 2024. – Vol. 16, № 3. – P. 583–597. – doi: 10.4251/wjgo.v16.i3.583.

7. Xie, Y.Gastrointestinal cancers in China, the USA, and Europe / Xie Y., Shi L., He X., Luo Y. // Gastroenterology Report. – 2021. – Vol. 9, № 2. – P. 91–104. – doi: 10.1093/gastro/goab010.

 Shaukat, A. Current and future colorectal cancer screening strategies / Shaukat A., Levin T.R. // Nature Reviews Gastroenterology & Hepatology. – 2022. – Vol. 19, № 8. – P. 521–531. – doi: 10.1038/s41575–022–00612–y.

 Bénard, F. Systematic review of colorectal cancer screening guidelines for average–risk adults: Summarizing the current global recommendations / F. Bénard,
 A. N. Barkun, M. Martel [et al.] // World J Gastroenterol. – 2018. – №24. – P. 124–138. Драпкина, О.М. Скрининг: терминология, принципы и международный опыт / Драпкина О.М., Самородская И.В. // Профилактическая медицина. – 2019. – № 22(1). – С. 90–97.

11. Приказ Министерства здравоохранения Российской Федерации от 27.04.2021 г. №404н «Об утверждении Порядка проведения профилактического медицинского осмотра и диспансеризации определенных групп взрослого населения».

12. Десятов, Е. Н. Колоректальный скрининг: первые результаты и выводы /
Е. Н. Десятов, Ф. Ш. Алиев, Ю. Н. Кузнецов [и др.] // Медицинская наука и образование Урала. – 2016. – №4 (88). – С. 69–73.

 Десятов, Е.Н. Колоректальный скрининг. Первые результаты / Десятов Е.Н., Алиев Ф.Ш., Зуев В.Ю., Кирьянова А.С., Кузнецов В.Ю., Щепкин К.В. // Онкологическая колопроктология. – 2017. – № 7(4). – С. 31–37.

14. Ануфриева, Е.В. Проектный подход к организации скрининга колоректального рака как основа улучшения общественного здоровья населения / Ануфриева Е.В., Демидов Д.А., Чекасина Е.С. // Здоровье населения и среда обитания. – 2023. – № 31(11). – С. 7–14.

15. Десятов, Е.Н. Скрининг колоректального рака у работников предприятий РОСТЕХ Свердловской области. Пилотный проект / Десятов Е.Н., Алиев Ф.Ш., Цветков М.С., Верховодова И.С., Яковлева И.Л. // Колопроктология. – 2023. – № 22(2). – С. 64–69.

16. Сахибуллин, Р.Ф. Скрининг колоректального рака у сотрудников МВД России по Республике Татарстан. Первые результаты и перспективы развития / Сахибуллин Р.Ф., Амиров Н.Б., Фатыхов Р.Г., Тухватуллина Г.В. // Вестник современной клинической медицины. – 2021. – № 14(5). – С. 56–61.

17. Нестеров, П.В. Региональная клинико–экономическая модель скрининга колоректального рака / Нестеров П.В., Ухарский А.В., Кислов Н.В. // Исследования и практика в медицине. – 2020. – № 3, – С. 146–159.

18. Кашин, С.В. Принципы фотодокументации при выполнении эндоскопии верхних отделов пищеварительного тракта / Кашин С.В., Куваев Р.О., Никонов Е.Л., Филимендикова К.Ю. // Доказательная гастроэнтерология. – 2023. – № 12(2). – С. 76–86.

19. Кашин, С.В. Принципы фотодокументирования при выполнении колоноскопии / Кашин С.В., Завьялов Д.В. // Доказательная гастроэнтерология. – 2024. – № 13(1). – С. 86–93.

20. Приказ Департамента здравоохранения и фармации Ярославской области от 28.12.2017 г. №1568 «Регламент обмена данными эндоскопических исследований и стандартизированные протоколы фотодокументации».

21. Карпов, О.Э. Инновационные эндоскопические технологии в многопрофильном учреждении / О. Э. Карпов, П. С. Ветшев, А. С. Маады [и др.] // Хирургия. Журнал им. Н.И. Пирогова. – 2017. – № 5. – С. 52–59.

22. Кашин, С.В. Основные положения рекомендаций европейского общества гастроинтестинальной эндоскопии по подготовке толстой кишки к скрининговой колоноскопии / С.В. Кашин, Н.В. Нехайкова, Н.С. Видяева [и др.] // Доказательная гастроэнтерология. – 2017. – Т. 6, № 3. – С. 36–50.

23. Кашин, С.В. Стандарты качественной колоноскопии (пособие для врачей) / С.В. Кашин, Е.Л. Никонов, Н.В. Нехайкова [и др.] // Доказательная гастроэнтерология. – 2019. – №8(1). – С. 2003–2032. – doi: 10.17116/ dokgastro20198012003.

24. Hassan, C. The low impact of the guidelines: does it concern only colorectal polypectomy? / C. Hassan, R. Bisschops // Endosc. Int. Open. – 2020. – Vol. 8, № 3. – P. E456–E457.

25. Kwon, D. AI-generated images threaten science – here's how researchers hope to spot them / Kwon D. // Nature. – $2024. - N \ge 635. - P. 268-269. - doi: 10.1038/d41586-024-03542-8.$

26. Ali, S. Where do we stand in AI for endoscopic image analysis? Deciphering gaps and future directions / Ali, S. // npj Digital Medicine. – 2022. – №5. – P. 184. – doi: 10.1038/s41746-022-00733-3.

27. Song, Y. Polyp-Size: A Precise Endoscopic Dataset for AI-Driven Polyp Sizing / Song Y., Du S., Wang R. et.al. // Scientific Data. – 2025. – № 12. – P. 918. – doi: 10.1038/s41597-025-05251-x.

28. Parasa, S. Framework and metrics for the clinical use and implementation of artificial intelligence algorithms into endoscopy practice: recommendations from the American Society for Gastrointestinal Endoscopy Artificial Intelligence Task Force /

Parasa S., Repici A., Berzin T., Leggett C., Gross S.A., Sharma P. // Gastrointest Endosc. – 2023. – № 97(5). – P. 815-824. – doi: 10.1016/j.gie.2022.10.016.

29. Hann, A. Current status and limitations of artificial intelligence in colonoscopy
/ Hann A, Troya J, Fitting D. // United European Gastroenterology Journal. – 2021. –
№9(5). – P. 527-533. – doi: 10.1002/ueg2.12108.

30. Карпов, О.Э. Информационные технологии, вычислительные системы и искусственный интеллект в медицине / Карпов О.Э., Храмов А.Е. – М.: ДПК Пресс. 2022.

Терапевтическая эндоскопия желудочно-кишечного тракта: атлас / ред.
 Х.Дж.Чун, С.-К. Янг, М.-Г. Чой; пер. с англ. под ред. С.А. Палевской. –
 М.: ГЭОТАР–Медиа, 2017. – 480 с.

32. Палевская, С.А. Эндоскопия желудочно-кишечного тракта / Палевская С.А., Короткевич А.Г. – М.: ГЭОТАР-Медиа, 2020. – 752 с.

33. Белоусов, А.С. Диагностика, дифференциальная диагностика и лечение болезней органов пищеварения / Белоусов А.С., Водолагин В.Д., Жаков В.Л. – М.: Медицина, 2002. – 424 с.

34. Гончарик, И.И. Клиническая гастроэнтерология (болезни пищевода, желудка, кишечника): практ. пособие / Гончарик И.И. – Минск: Ураджай, 2002. – 335 с.

35. Cotton, P.B. Practical Gastrointestinal Endoscopy: The Fundamentals / Cotton P.B., Williams C.B. – 5th ed. – Oxford Blackwell Publishing, 2003.

36. Дворкович, В.П. Цифровые видеоинформационные системы (теория и практика) / Дворкович В.П., Дворкович А.В. – М.: Техносфера, 2012. – 1009 с.

37. Гонсалес, Р., Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. –
 М.: Техносфера, 2012.

38. Лукьяница, А.А. Цифровая обработка видеоизображений / Лукьяница А.А., Шишкин А.Г. – М.: Ай–Эс–Эс Пресс, 2009. – 518 с.

39. Клетте, Р. Компьютерное зрение. Теория и алгоритмы / Клетте Р. – М.: ДМК Пресс, 2019.

40. Форсайт, Д.А. Компьютерное зрение. Современный подход / Форсайт Д.А., Понс Д. – М.: «Вильямс», 2004. – 928 с.

41. Чобану, М.К. Многомерные многоскоростные системы обработки сигналов / Чобану М.К. – М.: Техносфера, 2009. – 480 с.

42. Rangayyan, R.M. Biomedical Image Analysis / Rangayyan R.M. – CRC Press.Boca Raton. 2005. – 1306 p.

43. Bovik, A. The Handbook of Image and Video Processing (Communication, Networking and Multimedia) / Bovik A. – Academic Press, 2005. – 891 p.

44. Безруков, В.Н. Системы цифрового вещательного и прикладного телевидения. Учебное пособие для вузов / Безруков В.Н., Балобанов В.Г. – М.: Горячая линия – Телеком, 2015. – 608 с.

45. Приоров, А.Л. Обработка и передача мультимедийной информации: учебное пособие / Приоров А.Л., Хрящев В.В., Топников А.И. – Ярославль: ЯрГУ. 2022. – 248 с.

46. Галушкин, А.И. Теория нейронных сетей / Галушкин А.И. // Сер. Нейрокомпьютеры и их применение. Кн.1. – М.: ИПРЖР, 2000.

47. Балухто, А.Н. Нейросетевые системы обработки информации и их применение в космической технике / Балухто А.Н. – М.: СИП РИА, 2000.

48. Горбань, А.Н. Обучение нейронных сетей / Горбань А.Н. – М.: СП Параграф, 1991.

49. Журавлев, Ю.И. Распознавание. Математические методы. Программная система. Практические применения / Журавлев Ю.И., Рязанов В.В., Сенько О.В. – М.: Фазис, 2005. – 159 с.

50. Вапник, В.Н. Теория распознавания образов (статистические проблемы обучения) / Вапник В.Н., Червоненкис А.Я. – М.: Наука, 1974. – 416 с.

51. Ланнэ, А.А. Нейронные цепи, тринадцатая проблема Гильберта и задачи обработки сигналов / Ланнэ А.А. // Вестник молодых учёных. Серия Технические науки. – 2001. – № 7. – С. 3-26.

52. Yin, L. A New Class of Nonlinear Filters — Neural Filters / Yin L., Astola J., Neuvo Y. // IEEE Trans. on Signal Processing. – 1999. – V. 41. № 3. – P. 1201-1222.

53. Хайкин, С. Нейронные сети: полный курс / Хайкин С. – 2-е издание. – М.: Вильямс, 2008. 1103 с.

54. Гудфеллоу, Я. Глубокое обучение / Гудфеллоу Я., Бенджио И., Курвилль А. – ДМК–Пресс, 2017. – 652 с.
55. Николенко, С.И. Глубокое обучение / Николенко С.И., Кадурин А.А., Архангельская Е.О. – Спб.: Питер, 2018. – 480 с.

56. Николенко, С.И. Машинное обучение: основы / Николенко С.И. – СПб.: Питер, 2025. – 608 с.

57. Бринк, Х. Машинное обучение / Бринк Х., Ричардс Д., Феверолф М. – Спб.: Питер, 2017. – 336 с.

58. Никонов, Е.Л. Международный опыт скрининга колоректального рака / Е.Л. Никонов, В.А. Аксенов, С.В. Кашин [и др.] // Доказательная гастроэнтерология. – 2017. – Т. 6, № 3. – С. 30–35.

59. Обухова, Н.А. Цветовые преобразования видеоэндоскопических изображений / Обухова Н.А., Мотыко А.А., Поздеев А.А., Савельева О.Р., Ян С. // В сборнике: GraphiCon 2024. Материалы 34–й Международной конференции по компьютерной графике и машинному зрению. Омск. – 2024. – С. 641–655.

60. Поздеев А.А. Методы анализа и обработки изображений видимого оптического диапазона в системах поддержки врачебных решений: дис. ... канд. техн. наук: 2.2.13 / Поздеев А.А. – Санкт–Петербург, 2023. – 169 с.

61. Pozdeev, A.A. Algorithms for Real–Time Endoscopy Image Processing Pipeline in Clinical Decision Support Systems / A.A. Pozdeev, N.A. Obukhova, A.A. Motyko // International Journal of Embedded and Real–Time Communication Systems (IJERTCS). – 2019. – Vol. 10, No. – P. 39–59.

62. Синь, Я. Методы обработки и визуализации мультиспектральных медицинских изображений: дис. ... канд. техн. наук: 2.2.13 / Синь Я. – Санкт– Петербург, 2025. – 181 с.

63. Early, D.S. Guidelines for sedation and anesthesia in GI endoscopy / Early D.S., Lightdale J.R., Vargo J.J., et al. for the ASGE Standarts of Practice Committee // Gastrointest Endosc. – 2018.

64. Баландина, Л.С. Оценка эффективности использования эндоскопического диагностического оборудования в субъектах Российской Федерации за 2022 год / Баландина Л.С., Базарова И.Н., Богомазова А.С. Социальные аспекты здоровья населения [сетевое издание]. – 2024. – № 70(S5):6. – Режим доступа: http://vestnik.mednet.ru/content/view/1673/30/lang,ru/. – doi: 10.21045/2071-5021-2024-70-S5-6.

65. Стандарты по направлению «Искусственный интеллект» [электронный ресурс]. – Режим доступа: https://www.rst.gov.ru/portal/gost/home/standarts/ aistandarts.

66. Зарегистрированные медицинские изделия, использующие технологии искусственного интеллекта [электронный pecypc]. – Режим доступа: https://webiomed.ru/blog/zaregistrirovannye-meditsinskie-izdeliia-ai/.

67. Каприн, А.Д. Технологии диагностики и скрининга в раннем выявлении злокачественных новообразований / А.Д. Каприн, Л.М. Александрова, В. В. Старинский [и др.] // Онкология. Журнал им. П.А. Герцена. – 2018. – Т. 7, № 1. – С. 34–40.

68. Евстогина, Ю.В. Наиболее эффективные стратегии скрининга колоректального рака / Ю. В. Евстогина, О. М. Драпкина // Профилактическая медицина. – 2019. – Т. 22. № 1. – С. 105–108.

69. Зорина, О.С. Скрининг на колоректальный рак в странах Европы / О.С. Зорина, Т.С. Сергалиев // Сборник статей по материалам XX международной научно–практической конференции «Научный форум: медицина, биология и химия». – 2019. – С. 19–28.

70. Леушина, Е.А. Скрининг колоректального рака у пациентов с заболеваниями желудочно–кишечного тракта / Е. А. Леушина // Колопроктология. – 2020. – Т. 19, № S1 (73). – С. 29.

71. Пузанов, Д.П. Скрининг колоректального рака. Обзор существующих методов и рекомендаций / Д. П. Пузанов, В. В. Половинкин, И. А. Пузанова // Инновационная медицина Кубани. – 2018. – № 1 (9). – С. 58–64.

72. Смирнова, Е.В. Опыт внедрения скрининга колоректального рака в рамках городской программы диспансеризации взрослого населения / Е. В. Смирнова, В. В. Дудка, А. В. Семенов [и др.] // Злокачественные опухоли. – 2017. – Т. 7, № 3–S1. – С. 205.

73. Смирнова, Е.В. Опыт внедрения скрининга колоректального рака в рамках городской программы диспансеризации взрослого населения / Е. В. Смирнова, В. В. Дудка, А. В. Семёнов [и др.] // Колопроктология. – 2017. – №S3 (61). – С. 64–65.

74. Факторы риска и методы скрининга колоректального рака. Учебнометодическое пособие / М. В. Кручинина, И. О. Светлова, А. А. Громов [и др.] // Новосибирск: ООО «Офсет-TM», 2019. – 160 с.

75. Ханевич, М.Д. Факторы риска и профилактика колоректального рака /
М. Д. Ханевич, А. В. Хазов, Г. Н. Хрыков [и др.] // Профилактическая медицина. –
2019. – Т. 22, № 3. – С. 107–111.

76. Хатьков, И.Е. Современные принципы скрининга, диагностики и терапии колоректального рака / И. Е. Хатьков, А. В. Каграманова, Н. Б. Захаржевская [и др.] // Терапевтический архив. – 2016. – Т. 88, № 2. – С. 90–96.

77. Maida, M. Quality measures improving endoscopic screening of colorectal cancer: a review of the literature / M. Maida, G. Morreale, E. Sinagra [et al.] // Expert Rev. Anticancer Ther. – 2019. – Vol. 19, No. – P. 223–235. – doi: 10.1080/14737140.2019.1565999.

78. Takeda, K. Accuracy of diagnosing invasive colorectal cancer using computer– aided endocytoscopy / K. Takeda, S. E. Kudo, Y. Mori [et al.] // Endoscopy. – 2017. – Vol. 49, № 8. – P. 798–802. – doi: 10.1055/s–0043–105486.

79. Saftoiu, A. Role of gastrointestinal endoscopy in the screening of digestive tract cancers in Europe: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement / A. Sáftoiu, C. Hassan, M. Areia [et al.] // Endoscopy. – 2020. – Vol. 52, № 4. – P. 293–304. – doi: 10.1055/a–1104–5245.

80. Засыпкина, Е.В. К вопросу о врачебных ошибках: методология и критерии определения / Засыпкина Е.В., Катрунов В.А., Кузнецова М.Н. // Бюл. медицинских Интернет–конференций. – 2016. – Т. 6, № 1. – С. 243–247.

81. Лудупова, Е.Ю. Врачебные ошибки. Литературный обзор / Лудупова Е.Ю. // Вестник Росздравнадзора. – 2016. – № 2. – С. 6–15.

82. Варданян, Г.Д. Врачебные ошибки: современное состояние проблемы / Варданян Г.Д., Аветисян Г.А., Джаноян Г.Дж. // Медицинская наука Армении НАН РА. – 2019. – № 4. – С. 105–119.

83. Сайт ООО «ПК Прометей» [электронный ресурс]. – Режим доступа: https://prom-endo.ru/production.

84. Khodashahri, N.G. Decision support system (DSS) / Khodashahri N.G.,
Sarabi M.M.H. // Singaporean Journal of Business Economics and Management Studies.
2013. – Vol. 1, № 6. – P. 95–102.

85. Dinevski, D. Dinevski D., Sarenac T., Bele U. Clinical Decision Support
Systems / Dinevski D., Sarenac T., Bele U. // Telemedicine Techniques and Applications.
2011. – № 1. – P. 185–210.

86. Кирсанова, А.В. Современное состояние и перспективы развития экспертных медицинских систем / Кирсанова А.В. // Новый университет. Серия «Технические науки». – 2015. – № 11. – С. 45–46.

87. Рамеев, О.А. Основы теории принятия решений в организационных системах управления / Рамеев О.А. – М. Горячая линия Телеком, 2023.

88. Малых, В.Л. Системы поддержки принятия решений в медицине /
В.Л. Малых // Программные системы: теория и приложения. – 2019. – Т. 10, № 2(41). – С. 155–184.

89. Dinevski, D. Clinical Decision Support Systems / Dinevski D., Sarenac T.,
Bele U. // Telemedicine Techniques and Applications. – 2011. – № 1. – Р. 185–210.

90. Фролова, М.С. Системы поддержки принятия решений для задач оснащения лечебных учреждений медицинской техникой / Фролова М.С., Фролов С.В., Толстухин И.А. // Университет им. В.И. Вернадского. Специальный выпуск. – 2014. – № 52. – С. 106–111.

91. Старков, Ю.Г. Эволюция диагностических технологий в эндоскопии и современные возможности выявления опухолей желудочно–кишечного тракта / Старков Ю.Г., Солодинина Е.Н., Шишин К.В. // Тихоокеанский медицинский журнал. – 2009. – Т. 2, № 36. – С. 35 – 39.

92. Bisschops, R. Performance measures for upper gastrointestinal endoscopy: a European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative / Bisschops R., Areia M., Coron E. // Endoscopy. – 2016. – Vol. 48, № 9. – P. 843–864.

93. Батухтин, Д.М. Анализ узкоспектральных эндоскопических изображений на внутренней поверхности пищевода / Батухтин Д.М., Пеганова Е.В., Митракова Н.Н. // Вестник Поволжского государственного технологического университета. Серия: радиотехнические и инфокоммуникационные системы. – 2014. – Vol. 23, № 4. – С. 45–57.

94. Обухова, Н. А. Исследование и разработка методов улучшения эндоскопических (медицинских) изображений / Н. А. Обухова, А. А. Мотыко, А. А. Поздеев // Известия высших учебных заведений России. Радиоэлектроника. – 2019. – Т. 22, № 2. – С. 22–30.

95. Степанова О.А. Алгоритмы распознавания гистологических и эндоскопических изображений для систем прикладного телевидения в медицине: дис. ... канд. техн. наук: 2.2.13 / Степанова Ольга Анатольевна. – Ярославль, 2019. – 122 с.

96. Поздеев, А.А. Повышение визуального качества эндоскопических изображений в системах поддержки врачебных решений / А.А. Поздеев // Вопросы радиоэлектроники. Сер. Техника телевидения. – 2022. – Вып. 1. – С. 3–10.

97. Поздеев, А.А. Метод повышения контраста медицинских видеоизображений с адаптивной глубиной коррекции для систем поддержки врачебных решений / А.А. Поздеев // Известия высших учебных заведений России. Радиоэлектроника. – 2022. – №5. – С. 91–103.

98. Технология NBI от OLYMPUS [Электронный ресурс]. – Режим доступа: https://www.olympus-russia.com/medical/ru/Продукция-и-решения/Медицинскиерешения/Узкоспектральная-визуализация-NBI/NBI-Гастроэнтерология.html (дата обращения: 04.04.2025).

99. Bashar, M. Automatic detection of informative frames from wireless capsule endoscopy images / Bashar M., Kitasaka T., Suenaga Y., Mekada Y., Mori K. // Med. Image Anal. – 2010. – № 14(3). – P. 449–470.

100. Karargyris, A. Wireless Capsule Endoscopy and Endoscopic Imaging: A Survey on Various Methodologies Presented / Karargyris A, Bourbakis N. // IEEE Eng. Med. Biol. Mag. $-2010. - N_{2} 29(1). - P. 72-83.$

101. Сайт ООО «ЭМИС» [электронный ресурс]. – Режим доступа: https://emisinfo.ru.

102. Rumelhart, D.E. Learning representations by back-propagating errors / Rumelhart D.E., Hinton G.E., Williams R.J.J. // Nature. – 1986. – № 323. – P. 533–536.

103. LeCun, Y. Convolutional networks for images, speech, and time-series / LeCun Y., Bengio Y. – The handbook of brain theory and neural networks. Cambridge, MA: MIT Press., 1995.

104. Krizhevsky, A. ImageNet classification with deep convolutional neural networks / Krizhevsky A., Sutskever I., Hinton G.E. // Advances in Neural Information Processing Systems. – 2012. – P. 1097–1105.

105. Shorten, C. A survey on image data augmentation for deep learning / Shorten C., Khoshgoftaar T. // J Big Data. $-2019. - N_{2}6. - P. 1-48.$

106. Blaivas, L. Are convolutional neural networks trained on imagenet images wearing rose-colored glasses? a quantitative comparison of imagenet, computed tomographic, magnetic resonance, chest x-ray, and point-of-care ultrasound images for quality [Электронный источник] / L. Blaivas, M. J. Blaivas // Ultrasound Med. – 2020. – Режим доступа: https://pubmed.ncbi.nlm.nih.gov/32757235/.

107. Girshick, R. Fast R–CNN / Girshick R. // Proceedings of the IEEE international conference on computer vision. – 2015. – P. 1440–1448.

108. Ren, S. Faster R–CNN: Towards real–time object detection with region proposal networks / Ren S., He K., Girshick R., Sun J. // Advances in neural information processing systems. – 2015. – P. 91–99.

109. Redmon, J. You only look once: Unified, real-time object detection / Redmon J., Divvala S., Girshick R., Farhadi A. // Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – P. 779–788.

110. Liu, W. SSD: Single Shot MultiBox Detector / Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C.–Y., Berg A. // European conference on computer vision. Springer. – 2016. – P. 21–37.

111. Ronneberger, O. U–Net: Convolutional Networks for Biomedical Image Segmentation / O. Ronneberger, P. Fischer, T. Brox // International Medical Image Computing and Computer–Assisted Intervention (MICCAI 2015), Munich, Germany, 2015.

112. Bochkovskiy, A. YOLOv4: Optimal Speed and Accuracy of Object Detection / Bochkovskiy A., Wang C.-Y., Liao H.-Y.M. // arXiv preprint arXiv:2004.10934, 2020.

113. Jocher, G. ultralytics/yolov5: v6.1. [Электронный ресурс] / Jocher G. et al. – Режим доступа: https://github.com/ultralytics/yolov5 (Дата обращения: 25.03.2022).

114. Tajbakhsh, N. Convolutional neural networks for medical image analysis: Full training or fine tuning? / Tajbakhsh N. et al. // IEEE transactions on medical imaging. – 2016. – Vol. 35(5). – P. 1299–1312.

115. Shin, H.–C. Deep Convolutional Neural Networks for Computer–Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning / Shin H.–C. et al. // IEEE Transactions on Medical Imaging. – 2016. – Vol. 35(5). – P. 1285–1298.

116. Goodfellow, I. Generative Adversarial Nets / Goodfellow I., Pouget–Abadie J., Mirza M., Xu B., Warde–Farley D., Ozair S., Courville A., Bengio Y. // Advances in Neural Information Processing Systems. – 2014. – P. 2672–2680.

117. Jha, D. DoubleU–Net: A Deep Convolutional Neural Network for Medical Image Segmentation / Jha D., Riegler M.A., Johansen D., Halvorsen P., Johansen H.D. // IEEE 33rd International Symposium on Computer–Based Medical Systems. – 2020. – P. 558–564.

118. Киселев, Н.С. Достаточный размер выборки: бутстрапирование правдоподобия / Киселев Н.С., Грабовой А.В. // Журнал вычислительной математики и математической физики. – 2025. – Т. 65., №2. – С. 235-242. – doi: 10.31857/S0044466925020094.

119. Vaswani, A. Attention Is All You Need [электронный ресурс] / Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A., Kaiser L., Polosukhin I. – Режим доступа: https://arxiv.org/abs/1706.03762.

120. Aksenov, Y. Linear Transformers with Learnable Kernel Functions are Better In-Context Models / Aksenov Y., Balagansky N., Vaina S., Shaposhnikov B., Gorbatovski A., Gavrilov D. // In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics, Bangkok, Thailand, 2024. – Vol. 1. – P. 9584–9597.

121. Li, Ruixiong. (2025). Target Detection Comparison Between YOLO and Transformer / Li, Ruixiong & He, Pengbo // 2025 IEEE 5th International Conference on Power, Electronics and Computer Applications. – P. 208-214.

122. Qiang, Chen. LW-DETR: A Transformer Replacement to YOLO for Real-Time Detection [электронный ресурс] / Qiang Chen, Xiangbo Su, Xinyu Zhang, Jian Wang, Jiahui Chen et al. – Режим доступа: https://arxiv.org/abs/2406.03459.

123. D-Fine: Redefine regression task in detrs as fine-grained distribution refinement [электронный ресурс]. – Режим доступа: https://arxiv.org/abs/2410.13842.

124. Федотов А.А. Основы цифровой обработки биомедицинских изображений: учеб. пособие / А.А. Федотов. – Самара: Изд–во СГАУ, 2013. – 108 с.

125. Бабаян, П.В. Обработка изображений в системах обнаружения и сопровождения объектов. Классические методы. Учебное пособие для вызов / Бабаян П.В., Муравьев В.С., Смирнов С.А., Стротов В.В. – М.: Горячая линия – Телеком, 2024. – 168 с.

126. Недзьведь, А.М. Анализ изображений для решения задач медицинской диагностики / А.М. Недзьведь, С.В. Абламейко. – Минск: ОИПИ НАН Беларуси, 2012. – 248 с.

127. Красильников, Н.Н. Цифровая обработка 2D- и 3D-изображений: учебное пособие / Н.Н. Красильников. – СПб.: БХВ-Петербург, 2011. – 608 с.

128. Обухова Н.А. Цифровая обработка изображений: учеб. пособие / Н. А. Обухова. – Санкт–Петербург: ЛЭТИ, 2024. – 117, [1] с.

129. Коротаев, В.В. Телевизионные измерительные системы: учеб. пособие / Коротаев В.В., Краснящих А.В. – СПб: СПбГУ ИТМО, 2008. – 108 с.

130. Коваленко, Д.А. Ассоциация сцен в эндоскопических видео / Коваленко Д.А., Гнатюк В.С. // GraphiCon 2017: Обработка и анализ биомедицинских изображений. Пермь. – 2017. – С. 269–274.

131. Авшаров, Е.М. Обработка медицинских изображений как необходимый инструментарий медицинского диагностического процесса / Авшаров Е.М., Абгарян М. Г., Сафарянц С.А. // Вестник рентгенологии и радиологии. – 2010. – № 3. – С. 54–61.

132. Münzer, B. (2018). Content–based processing and analysis of endoscopic images and videos: A survey / Münzer B. Schoeffmann K. Böszörmenyi L. // Multimedia Tools and Applications. – 2018. – Vol. 77. – P. 1323-1362. – doi: 10.1007/s11042–016–4219–z.

133. Хрящев, В.В. Детектирование и фильтрация импульсного шума со случайными значениями импульсов / Хрящев В.В., Приоров А.Л., Куйкин Д.К. // Цифровая обработка сигналов. – 2010. – № 1. – С. 18–22.

134. Хрящев, В. В. Улучшение качества ультразвуковых медицинских изображений / Хрящев В.В., Приоров А.Л., Сладков М.В. // Медицинская техника. – 2008. – № 4. – С. 11–13.

135. Хрящев, В.В. Применение переключающихся медианных фильтров для восстановления зашумленных изображений / Хрящев В.В., Приоров А.Л., Апальков И.В., Бухтояров С.С. // Вопросы радиоэлектроники. Серия общетехническая, выпуск 2. – 2006. – С. 137–147.

136. Хрящев, В.В. Удаление импульсного шума со случайными значениями импульсов из изображений / Хрящев В.В., Приоров А.Л., Голубев М.Н. // Радиотехника. – 2010. – № 5. – С. 72–79.

137. Khryashchev, V. Improvement of the Quality of Ultrasonic Medical Images / Khryashchev V., Priorov A., Sladkov M. // Biomedical Engineering. – 2008. – Vol. 42. No. 4. – P. 176–178.

138. Khryashchev, V. Image denoising using adaptive switching median filter / Khryashchev V., Apalkov I., Priorov A., Zvonarev P. // Proc. of the IEEE Int. Conf. on Image Processing (ICIP–2005). – Genoa, Italy, 2005. – P. I–117 – I–120.

139. Khryashchev, V. Vector median filter with directional detector for color image denoising / Khryashchev V., Kuykin D., Studenova A. // Proc. of Int. Conference of Signal and Image Engineering. – London. U.K., 6–8 July 2011. – P. 1699–1704.

140. Khryashchev, V. Random–valued impulse noise detection and removal in grayscale and color images / Khryashchev V., Shemyakov A., Gushchina O., Falcon D. // Proc. Int. MultiConference of Engineers and Computer Scientists 2012. – Hong Kong. 2012. – V. 1. – P. 692–697.

141. Khryashchev, V. An improved switching median filter for impulse noise removal / Khryashchev V., Kuykin D., Apalkov I. // Proc. of Int. IEEE Conference devoted to the 150–anniversary of Alexander S. Popov (EUROCON 2009). – Saint Petersburg. 2009. – P. 1315–1321.

142. Khryashchev, V. Modified progressive switched median filter for image enhancement / Khryashchev V., Kuykin D., Apalkov I. // Proc. of Int. conference on Computer Graphics and Vision (GraphiCon–2009). – Moscow. Russia. 2009. – P. 303–304.

143. Хрящев, В.В. Переключающийся медианный фильтр с блоком предварительного детектирования / Хрящев В.В. // Физический вестник ЯрГУ им. П.Г. Демидова: сб. науч. тр., Ярославль. 2006. – С. 90–97.

144. Khryashchev, V. JPEG2000 Ringing Artifact Reduction by Smart Bilateral Filter / Khryashchev V., Priorov A., Shmaglit L. // Proc. of the 2010 International Conference on Image Processing, Computer Vision and Pattern Recognition (IPCV 2010). – Las Vegas. USA. 2010. – Vol. 2. – P. 622–627.

145. Nikitin, A. Adaptive bilateral filter for JPEG 2000 deringing / Nikitin A.; Solovyev V.; Khryashchev V.; Priorov A. // Proc. of the 2011 International Conference on Image Processing, Computer Vision and Pattern Recognition (IPCV 2011). – Las Vegas. USA. 2010. – Vol. 2. – P. 227–232.

146. Khryashchev, V. A novel smart bilateral filter for ringing artifacts removal in JPEG2000 images / Khryashchev V., Apalkov I., Shmaglit L. // Proc. of the International Conference on Computer Graphics and Vision (GraphiCon–2010). – St. Petersburg. Russia. 2010. – P. 122–128.

147. Хрящев, В.В. Подавление артефактов «звона» при сжатии изображений / Хрящев В.В. // Радиоэлектронные средства получения, обработки и визуализации информации: докл. 4-й всероссийской конференции. – Нижний Новгород. 2014. – С. 279–281.

148. Yong, D. Visual Quality Assessment for Natural and Medical Image / Yong D.– Springer Berlin, Heidelberg, 2018.

149. Simoncelli, E. Natural image statistics and neural representation /E. Simoncelli,
B. Olshausen // Annual Review of Neuroscience. – 2001. – V. 24, № 1. – P. 1193–1216.

150. Wang, Z. A universal image quality index / Z. Wang, A Bovik // IEEE Signal Processing Letters. – 2002. – V. 9, № 3. – P. 81–84.

151. Wang, Z. Image quality assessment: from error visibility to structural similarity
/ Z. Wang, A.C. Bovik, H.R. Sheikh, E.P Simoncelli // IEEE Transactions on Image
Processing. – 2004. – Vol. 13, № 4. – P. 600–612.

152. Wang, Z. Mean squared error: Love it or leave it? – A new look at signal fidelity measures / Z. Wang, A. Bovik / IEEE Signal Processing Magazine. – 2009. – V. 26, № 1. – P. 98–117.

153. Хрящев, В.В. Модифицированный критерий оценки качества восстановленных изображений / Хрящев В.В., Арляпов С.А., Приоров А.Л. // Цифровая обработка сигналов. – 2006. – № 2. – С. 27–33.

154. Хрящев, В.В. Оценка качества цветных изображений при подавлении импульсного шума / Хрящев В.В., Приоров А.Л., Павлов Е.А., Герасимов Н.Б., Шемяков А.М. // Радиотехника. – 2013. – № 5. – С. 41–49.

155. Хрящев, В.В. Оценка качества JPEG2000 изображений / Хрящев В.В., Зараменский Д.А. // Сб. докл. 14-ой Всерос. конф. «Математические методы распознавания образов». – Суздаль. 2009. – С. 351–354.

156. Хрящев, В. В. Неэталонная оценка качества JPEG изображений / Хрящев В.В., Приоров А.Л., Саутов Е.Ю. // Цифровая обработка сигналов. – 2007. – № 3. – С. 15–19.

157. Хрящев, В.В. Оценка восстановленных изображений на основе универсального индекса качества / Хрящев В.В., Приоров А.Л., Балусов И.Л. // Радиотехника. – 2008. – № 12. – С. 23–28.

158. Хрящев, В.В. Определение типа искажения изображения в задаче неэталонной оценки качества / Хрящев В.В., Приоров А.Л., Соловьев В.Е., Шемяков А.М. // Нелинейный мир. – 2013. – Т. 11. № 1. – С. 32–35.

159. Хрящев, В.В. Анализ применения адаптивного дискретного косинусного преобразования в некоторых задачах цифровой обработки изображений / Хрящев В.В., Ганин А.Н., Гущина О.Н. // Успехи современной радиоэлектроники. – 2011. – № 2. – С. 72–80.

160. Sheikh, H. LIVE image quality assessment database release 2. 2006 / H. Sheikh et. all. [электронный ресурс]. – Режим доступа: http://live.ece.utexas.edu/research/ quality/subjective.htm.

161. Image database TID2013: Peculiarities, results, and perspectives / N. Ponomarenko et. all // Signal Processing: Image Communication. – 2015. – V. 30. – P. 57–77.

162. Стефаниди, А.Ф. Применение сверточных нейронных сетей в задаче мультимодальной идентификации / Стефаниди А.Ф., Приоров А.Л., Топников А.И., Хрящев В.В. // Цифровая обработка сигналов. – 2020. – № 2. – С. 52–58.

163. Стефаниди, А.Ф. Модификация VGG–архитектуры в задачах унимодальной и мультимодальной биометрии / Стефаниди А.Ф., Приоров А.Л., Топников А.И., Хрящев В.В. // Цифровая обработка сигналов. – 2020. – № 3. – С. 35–40.

164. Wojke, N. Simple online and realtime tracking with a deep association metric
/ Wojke N., Bewley A., Paulus D. // IEEE Int. Conference on Image Processing. – 2017.
– P. 3645–3649.

165. Bewley, A. Simple online and realtime tracking / Bewley A., Ge Z., Ott L., Ramos F., Upcroft B. // 2016 IEEE Int. Conference on Image Processing. – 2016. – P. 3464–3468.

166. Sabater, A. Robust and efficient post-processing for video object detection / Sabater A., Montesano L., Murillo A.C. // IEEE/RSJ Int. Conference on Intelligent Robots and Systems. – 2020. – P. 10536–10542.

167. Лебедев А.А. Исследование нейросетевых алгоритмов обнаружения объектов на видеоизображениях в медицинских системах прикладного телевидения: дис. ... канд. техн. наук: 2.2.13 / Лебедев А.А. – Ярославль, 2022. – 131 с.

168. Приоров, А.Л. Разработка и исследование алгоритма сегментации ядер клеток на гистологических изображениях / Приоров А.Л., Хрящев В.В., Степанова О.А., Среднякова А.С. // Биомедицинская радиоэлектроника. – 2018. – № 11. – С. 13–20.

169. Khryashchev, V. Histological Images Segmentation by Convolutional Neural Network with Morphological Post–filtration / Khryashchev V., Lebedev A., Stepanova O., Srednyakova A. // In: Favorskaya M., Jain L.C. (eds.) Computer Vision in Advanced Control Systems–5. Intelligent Systems Reference Library. – Springer International Publishing. Switzerland. – Vol. 175. – 2020. – P. 295–319.

170. Khryashchev, V. Using Convolutional Neural Networks in the Problem of Cell Nuclei Segmentation on Histological Images / Khryashchev V., Lebedev A., Stepanova O., Srednyakova A. // Recent Research in Control Engineering and Decision Making. ICIT 2019. Studies in Systems, Decision and Control. – vol 199. – Springer. Cham. – 2019. – P. 149-161.

171. Демидова Л.А., Титов С.Б. Гибридные алгоритмы анализа и обработки данных в задачах интеллектуальной поддержки принятия решений // М. Горячая линия – Телеком, 2017.

172. Хрящев, В.В. Концепция построения системы поддержки принятия врачебных решений в эндоскопии желудочно–кишечного тракта / Хрящев В.В. // Измерение. Мониторинг. Управление. Контроль. – 2025. – № 2. – С. 137–144.

173. Хрящев, В.В. Поддержка принятия врачебного решения в эндоскопии на основе методов искусственного интеллекта: задачи, проблемы, перспективы / Хрящев В.В. // Цифровая обработка сигналов и ее применение (DSPA–2024): докл. 26–й междунар. конф. – Москва, 2024. – С. 8.

174. Хрящев, В.В. Использование глубокого машинного обучения для обнаружения раннего рака при эндоскопическом исследовании желудочно–кишечного тракта / Хрящев В.В., Приоров А.Л. // Биотехнические, медицинские и экологические системы, измерительные устройства и робототехнические комплексы (БИОМЕДСИСТЕМЫ–2024): Рязань. 2024. – С. 8–12.

175. Свидетельство о государственной регистрации программы для ЭВМ № 2021660247. Piclab.EndoSolution – система поддержки принятия решений в эндоскопии на основе алгоритмов искусственного интеллекта / Хрящев В.В., Кашин С.В. – опубл. 23.06.2021.

176. Андержанова, А.С. Использование методов компьютерного зрения и искусственного интеллекта для анализа изображений в колоноскопии / Андержанова А.С., Котов Н.В., Хрящев В.В. // 78–я Научно–техническая конференция Санкт–Петербургского НТО РЭС им. А.С. Попова, посвященная Дню радио. – С.-Петербург, 2023. – С. 94–96.

177. Тихомиров, А.А. Исследование алгоритмов контроля качества эндоскопических изображений / Тихомиров А.А., Котов Н.В., Хрящев В.В. // Цифровая обработка сигналов и ее применение (DSPA–2023): докл. 25-й междунар. конф. – Москва, 2023. – С. 349–353.

178. Павлов, Н.А. Эталонные медицинские датасеты (MosMedData) для независимой внешней оценки алгоритмов на основе искусственного интеллекта в диагностике / Павлов Н.А., Андрейченко А.Е., Владзимирский А.В., Ревазян А.А., Кирпичев Ю.С., Морозов С.П. // Digital Diagnostics. – 2021. – №2(1). – Р. 49–65. – doi: 10.17816/DD60635.

179. Ахметвалеев, Р.Р. Методика разметки медицинских изображений с функцией кросс-проверки и интеллектуального сегментирования /

Ахметвалеев Р.Р., Шабанова К.И., Падукова А.А., Лакман И.А. // Оргздрав: новости, мнения, обучение. Вестник ВШОУЗ. – 2021. – Т. 7, № 3. – С. 62–69. – doi: 10.33029/2411- 8621-2021-7-3-62-69.

180. Ахметвалеев, Р.Р. Методика сегментации изображения для поддержки автоматической разметки объектов на эндоскопических изображениях / Ахметвалеев Р.Р., Лакман И.А., Попов Д.В., Красноперов М.В. // Информатизация и связь. – 2021. – № 2. – С. 146–152. – doi: https://doi.org/10.34219/ 2078-8320-2021-12-2-146-154.

181. Васильев, Ю.А. Обзор современных средств разметки цифровых диагностических изображений / Васильев Ю.А., Савкина Е.Ф., Владзимирский А.В., Омелянская О.В., Арзамасов К.М. // Казанский медицинский журнал. – 2023. – № 104(5). – С. 750–760. – doi: 10.17816/KMJ349060.

182. Сайт ООО «Точка зрения» [электронный ресурс]. – Режим доступа: www.tochka.ai.

183. Программа для разметки изображений [Электронный ресурс]. – Режим доступа: https://github.com/wkentaro/labelme (Дата обращения: 25.03.2022).

184. База об организации процесса разметки: команда, онбординг, метрики [электронный pecypc]. – Режим доступа: https://habr.com/ru/companies/alfa/ articles/883822/.

185. Rieke, N. The future of digital health with federated learning / Rieke, N., Hancox, J., Li, W. et al. // npj Digital Medicine. – 2020. – Vol. 3(1). – Article number: 119. – doi: 10.1038/s41746–020–00323–1.

186. Federated Learning [электронный ресурс]. – Режим доступа: https://www.ultralytics.com/ru/glossary/federated-learning.

187. Обучаем нейросети федеративным способом [электронный pecypc]. – Режим доступа: https://yandex.cloud/ru/blog/posts/2024/10/ml-models-and-federated-learning.

188. Xie, L. dFLMoE: decentralized federated learning via mixture of experts for medical data analysis [электронный ресурс] / Luan T., Cai W., Yan G., Chen Z., Xi N., Fang Y., Shen Q., Wu Z., Yuan J. – Режим доступа: https://arxiv.org/abs/2503.10412v1.

189. Kirchner, M. Federated EndoViT: Pretraining Vision Transformers via Federated Learning on Endoscopic Image Collections [электронный ресурс] / Kirchner M. et al. – Режим доступа: https://arxiv.org/html/2504.16612v1.

190. Kassem, H. Federated cycling (fedcy): Semi-supervised federated learning of surgical phases / H. Kassem, D. Alapatt, P. Mascagni, A. Karargyris, N. Padoy // IEEE Transactions on Medical Imaging. – 2023. – vol. 42. – P. 1920–1931.

191. Куваев, Р.О. Контроль качества эндоскопических исследований, перспективы автоматизированного анализа эндоскопических изображений / Куваев Р.О., Никонов Е.Л., Кашин С.В., Капранов В.А., Гвоздев А.А. // Кремлевская медицина. Клинический вестник. – 2013. – Т. 2. – С. 51–56.

192. Hoff, G. Cecum intubation rate as quality indicator in clinical versus screening colonoscopy / Hoff G., Holme O., Bretthauer M., Sandvei P., Darre-Næss O., Stallemo A., Wiig H., Hoie O., Noraberg G., Moritz V., de Lange T. // Endosc. Int. Open. – 2017. – №5(6). – P. E489-E495. – doi: 10.1055/s-0043-106180.

193. Rembacken, B. Quality in screening colonoscopy: position statement of the European Society of Gastrointestinal Endoscopy (ESGE) / Rembacken B., Hassan C., Riemann J.F. et al. // Endoscopy. – 2012. – № 44. – P. 957–968. – doi: 10.1055/s-0032-1325686.

194. Nagrath, N. Variation in cecal intubation rates between screening and symptomatic patient / Nagrath N., Phull P.S. // United European Gastroenterol J. – 2014. – N_{2} 2. – P. 295–300. – doi: 10.1177/2050640614536898.

195. Мтвралашвили, Д.А. Алгоритм на основе искусственного интеллекта для системы поддержки принятия врачебного решения при колоноскопии / Мтвралашвили Д.А., Шахматов Д.Г., Ликутов А.А., Запольский А.Г., Суслова Д.И., Бородинов А.А., Сушков О.И., Ачкасов С.И. // Колопроктология. – 2023. – Т. 22, № 2. – С. 92–102. – doi: 10.33878/2073–7556–2023–22–2–92–102.

196. Завьялов Д.В. Возможности ранней диагностики и лечения полипов и эпителиальных опухолей толстой кишки: дис. ... д-ра. мед. наук: 3.1.6, 3.1.9 / Завьялов Дмитрий Вячеславович. – Ярославль, 2021. – 300 с.

197. Королев, М.П. Принципы организации и критерии качества колоноскопии, выполняемой с целью скрининга колоректального рака. Электронный ресурс] / Королев М.П., Кашин С.В., Завьялов Д.В., Рассадина Е.А. и

др. – Режим доступа: http://www.endoscopy-yaroslavl.ru/content/100513_rekom (Дата обращения: 25.03.2022).

198. Вологжанина, Л. Г. Эффективная подготовка кишечника к скрининговой колоноскопии / Л. Г. Вологжанина, О. А. Игумнова, И. В. Петухова и др. // Колопроктология. – 2020. – Т. 19, № 3 (73). – С. 65–71.

199. Беслекоев, У.С. Колоноскопия – метод диагностики предраковых заболеваний и ранних форм рака толстой кишки / У.С.Беслекоев, Г.С.Доев, Б.А.Царукаев // В сборнике статей XIII Международного научно–практического конкурса. – 2018. – С. 261–265.

200. Веселов, В. В. Эндоскопическая диагностика плоских эпителиальных новообразований толстой кишки: автореф. дис. ... канд. мед. наук: 14.01.17 / Веселов Владимир Викторович. – М., 2018. – 24 с.

201. Галкова, З. В. Критерии качества проведения колоноскопии как элемент организации медицинской помощи при новообразованиях толстой кишки / З. В. Галкова // В сборнике: Материалы форума организаторов столичного здравоохранения. Клиники Москвы: практики устойчивого развития. – 2019. – С. 15–16.

202. Дуванский, В. А. Возможности виртуальной хромоскопии в эндоскопической диагностике новообразований толстой кишки (обзор литературы) / В. А. Дуванский, А. В. Белков // Лазерная медицина. – 2017. – Т. 21, № 1. – С. 45–48.

203. Дуванский, В. А. Современные эндоскопические технологии в диагностике неоплазий толстой кишки / В. А. Дуванский, Я. О. Чесалина // Экспериментальная и клиническая гастроэнтерология. – 2019. – № 4 (164). – С. 93–99.

204. Натансон, М. А. Факторы риска и скрининг рака ободочной кишки / М. А. Натансон // Справочник врача общей практики. – 2020. – № 4. – С. 13–22.

205. De Lange, T. Methodology to develop machine learning algorithms to improve performance in gastrointestinal endoscopy / De Lange T., Halvorsen P., Riegler M. // World journal of gastroenterology. – 2018. – Vol. 24(45). – P. 5057.

206. Rex, D.K. Quality indicators for colonoscopy / Rex D.K., Anderson J.C., Butterly L.F., et. al. // Gastrointestinal Endoscopy. – 2024. – Vol. 100(3). – P. 352–381. – doi: 10.1016/j.gie.2024.04.2905.

207. Areia, M. Cost–effectiveness of artificial intelligence for screening colonoscopy: a modelling study / Areia M., Mori Y., Correale L., et al. // Lancet Digit. Health. – 2022. –Vol. 4(6). – P. e436–e444. – doi: 10.1016/S2589–7500(22)00042–5.

208. Cho, M. Identification of cecum time-location in a colonoscopy video by deep learning analysis of colonoscope movement / Cho M., Kim J.H., Hong K.S., Kim J.S., Kong H.-J., Kim S. // PeerJ. – 2019. – \mathbb{N}_{2} 7. – P. e7256.

209. Обухова, Н. А. Автоматический анализ эндоскопических изображений для выявления и сегментации полипов / Н. А. Обухова, А. А. Мотыко, А. А. Поздеев // Статья в сборнике трудов 21-й Международной конференции «Цифровая обработка сигналов и её применение – DSPA–2019». – 2019. – С. 501–505.

210. Taghiakbari M. Automated Detection of Anatomical Landmarks During Colonoscopy Using a Deep Learning Model [Электронный источник] / M. Taghiakbari, S. H. Ghalehjegh, E. Jehanno et al. // Journal of the Canadian Association of Gastroenterology. – режим доступа: https://www.researchgate.net/publication/ 370486030_Automated_Detection_of_Anatomical_Landmarks_During_Colonoscopy_ Using_a_Deep_Learning_Model.

211. Lebedev, A. Automatic Identification of Appendiceal Orifice on Colonoscopy Images Using Deep Neural Network / Lebedev A., Khryashchev V., Kazina E., Zhuravleva A., Kashin S., Zavyalov D. // Proceedings of 2020 IEEE East–West Design and Test Symposium (EWDTS 2020), September 2020. – Varna, Bulgaria, 2020. – P. 367–371.

212. Guseynova, S. Deep Neural networks for confirming cecum achievement during screening colonoscopy / Guseynova S., Kashin S., Zavjalov D., Khryashchev V. // Journal of Gastroenterology and Hepatology. Asian Pacific digestive week 2024. – Bali, Indonesia. – 2024. – Vol. 39. – P. 223–224. – doi: 10.1111/jgh.16780.

213. Лебедев, А.А. Распознавание устья червеобразного отростка на эндоскопических изображениях прямой кишки на основе сверточной нейронной сети / Лебедев А.А., Хрящев В.В., Казина Е.М., Среднякова А.С., Журавлева А.С.

// Цифровая обработка сигналов и ее применение (DSPA-2020): докл. 22-й междунар. конф. – Москва, 2020. – С. 638–642.

214. Хрящев, В.В. Комбинированный алгоритм анализа изображений для контроля качества колоноскопического исследования / Хрящев В.В., Котов Н.В., Тихомиров А.А., Ненахов И.С. // Биомедицинская радиоэлектроника. – 2025. – Т. 28. № 1. – С. 64–74.

215. Хрящев, В.В. Использование методов глубокого машинного обучения в задаче обнаружения купола слепой кишки на видеоданных колоноскопического исследования / Хрящев В.В. // Модели, системы, сети в экономике, технике, природе и обществе. – 2023. – № 4. – С. 133–141.

216. Хрящев, В.В. Обнаружение купола слепой кишки на видеоэндоскопических изображениях / Хрящев В.В., Котов Н.В., Тихомиров А.А. // Цифровая обработка сигналов и ее применение (DSPA–2025): докл. 27-й междунар. конф. – Москва, 2024. – С. 119–122.

217. Pushkar, K. (2024). Object Tracking Using Computer Vision: A Review / Kadam, Pushkar & Fang, Gu & Zou, Ju. // Computers. № 13. P. 136. 10.3390/computers13060136.

218. Botao, Y. Joint Feature Learning and Relation Modeling for Tracking: A One-Stream Framework / Botao Y., Hong C., Bingpeng M., Shiguang S., Xilin C. // arXiv:2203.11991v4 [cs.CV] 20 Dec 2022.

219. Yifu, Z. ByteTrack: Multi-Object Tracking by Associating Every Detection Box / Yifu Z., Peize S., Yi J., Dongdong Y., Fucheng W., Zehuan Y., Ping L., Wenyu L., Xinggang W. // arXiv:2110.06864v3 [cs.CV] 7 Apr 2022.

220. Nir, A. BoT-SORT: Robust Associations Multi-Pedestrian Tracking / Nir A., Roy O., Ben-Zion B. // arXiv:2110.06864v3 [cs.CV] 7 Apr 2022.

221. Хрящев, В.В. Классификация эндоскопических изображений устья червеобразного отростка на основе методов глубокого машинного обучения / Хрящев В.В., Завьялов Д.В., Андержанова А.С. // Цифровая обработка сигналов. – 2023. – № 1. – С. 35–38.

222. Котов, Н.В. Разработка и тестирование алгоритма детектирования области купола слепой кишки на видеоизображениях колоноскопического исследования / Котов Н.В., Хрящев В.В. // Современные технологии обработки

сигналов: сборник докладов 4-й Всероссийской конференции. – Москва, 2023. – С. 200–204.

223. Horn, B. Determining optic flow / Horn B., Schunck B. // Artificial intelligence. - 1981. - Vol. 17. - P. 185-203.

224. Meinhardt-Llopis, E. Horn–Schunck Optical Flow with a Multi-Scale Strategy / Enric Meinhardt-Llopis, Javier Sanchez, Daniel Kondermann // Image Processing on Line. – 2013. – № 3. – P. 151–172. – doi: 10.5201/ipol.2013.20.

225. Седов, А.Г. Разработка и исследование алгоритмов оценки качества колоноскопических изображений / Седов А.Г., Хрящев В.В. // Цифровая обработка сигналов и ее применение (DSPA–2024): докл. 26–й междунар. конф. – Москва, 2024. – С. 244–247.

226. Седов, А.Г. Использование метрики PIQUE для оценки качества медицинских видеоизображений / Седов А.Г., Хрящев В.В., Приоров А.Л. // Цифровая обработка сигналов и ее применение (DSPA–2025): докл. 27-й междунар. конф. Москва, 2024. С. 189–191.

227. Nenakhov, I. No–Reference image quality assessment based on local binary patterns / Nenakhov I., Khryashchev V., Priorov A. // Proceedings of the IEEE East–West Design & Test Symposium (EWDTS). – 2016. – P. 529–532.

228. Mittal, A. Making a "completely blind" image quality analyzer / Mittal A., Soundararajan R., Bovik A. // IEEE Signal processing letters. – 2012. – № 20.3. – P. 209-212.

229. Хрящев, В.В. Неэталонный алгоритм оценки качества изображений с использованием дискретного косинусного преобразования / Хрящев В.В., Приоров А.Л., Павлов Е.А., Гущина О.Н. // Успехи современной радиоэлектроники. – 2013. – № 11, С. 3–13.

230. Герасимов, Н.Б. Оценка уровня блочности в изображениях, сжатых с использованием дискретного косинусного преобразования / Герасимов Н.Б., Приоров А.Л., Хрящев В.В. // Радиотехнические и телекоммуникационные системы. – 2014. – № 2(14). – С. 39–44.

231. Приоров, А.Л. Разработка и анализ алгоритма неэталонной оценки качества изображений на основе локальных бинарных шаблонов / Приоров А.Л.,

Ненахов И.С., Хрящев В.В. // Успехи современной радиоэлектроники. – 2016. – № 7. – С. 46–52.

232. Хрящев, В.В. К вопросу о неэталонной оценке качества телевизионных изображений / Хрящев В.В. // Телевидение: передача и обработка изображений: Материалы 14-й международной конференции. – С.-Петербург, 2017. – С. 21–24.

233. Chaofeng, Chen. TOPIQ: A Top–Down Approach from Semantics to Distortions for Image Quality Assessment / Chen C., Mo J., Hou J., Wu H., Liao L., Sun W, Yan Q., Lin W. // IEEE Transactions on Image Processing. – 2024. – Vol. 33. – P. 2404–2418.

234. Ying, Z. From Patches to Pictures (PaQ–2–PiQ): Mapping the Perceptual Space of Picture Quality / Ying Z., Niu H., Gupta P., Mahajan D., Ghadiyaram D., Bovik A. // arXiv:1912.10088 [cs.CV].

235. Chaofeng, C. IQA-PyTorch: PyTorch Toolbox for Image Quality Assessment [электронный источник] / Chaofeng C., Jiadi M. – режим доступа: https://github.com/chaofengc/IQA-PyTorch 2022.

236. Седов, А.Г. Оценка качества изображений на основе сравнения откликов сверточных слоев нейронной сети / Седов А.Г., Хрящев В.В., Приоров А.Л. // Телевидение: передача и обработка изображений: Материалы 15–й международной конференции. – С.-Петербург, 2018. – С. 232–234.

237. Хрящев, В.В. Неэталонная оценка качества изображений колоноскопических исследований / Хрящев В.В., Седов А.Г., Приоров А.Л. // Цифровая обработка сигналов. – 2024. – № 3. – С. 44–49.

238. Бакулин, И.Г. Применение искусственного интеллекта для анализа эндоскопических изображений при воспалительных заболеваниях кишечника / Бакулин И.Г., Расмагина И.А., Скалинская М.И., и др. // Терапия. – 2022. – № 7. – С. 7–17. – doi: 10.18565/therapy.2022.7.7–14.

239. Matsuda, T. Advances in image enhancement in colonoscopy for detection of adenomas / T. Matsuda, A. Ono, M. Sekiguchi [et al.] // Nat. Rev. Gastroenterol. Hepatol. – 2017. – Vol. 14, № 5. – P. 305–314. doi: 10.1038/nrgastro.2017.18.

240. McCarty, T. R. Role of image–enhanced endoscopy: how to improve colorectal polyp detection rates in the coming decade / T. R. McCarty, H. Alhara // Gastrointest. Endosc. – 2020. – Vol. 91, № 1. – P. 113–114. – doi: 10.1016/j.gie.2019.07.025.

241. The Paris endoscopic classification of superficial neoplastic lesions: esophagus, stomach and colon / Gastrointest. Endosc. – 2003. – Vol. 58, №6. Suppl. – P. S3–S4.

242. Завьялов, Д.В. Система САD ЕYE для детекции и дифференцировки новообразований толстой кишки в режиме реального времени / Завьялов Д.В., Кашин С.В., Гусейнова С.Р. // Доказательная гастроэнтерология. – 2024. – Т. 13 (2). – С. 50–54.

243. Mori, Y. Cost savings in colonoscopy with artificial intelligence–aided polyp diagnosis: an add–on analysis of a clinical trial (with video) / Mori Y., Kudo S., East J.E. et al. // American Society for Gastrointestinal Endoscopy. – 2020. – doi: 10.1016/j.gie.2020.03.3759.

244. Byrne, M.F. Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model / Byrne M.F., Chapados N., Soudan F. // Gut. – 2019. – Vol. 68. – P. 94–100.

245. Taghiakbari, M. Automated Detection of Anatomical Landmarks During Colonoscopy Using a Deep Learning Model / Taghiakbari M., Ghalehjegh S. Jehanno E. et. al. // Journal of the Canadian Association of Gastroenterology. – 2023. – Vol. 6, Issue 4. – P. 145-151. – doi: 6. 10.1093/jcag/gwad017.

246. Lui, T.K.L Is artificial intelligence the final answer to missed polyps in colonoscopy? / Lui T.K.L, Leung W.K. // World J Gastroenterol. – 2020. – Vol. 26 (35). – P. 5248–5255. – doi: 10.3748/wjg. v26.i35.5248.

247. Luo, Y Artificial Intelligence–Assisted Colonoscopy for Detection of Colon Polyps: A Prospective, Randomized Cohort Study / Luo Y, Zhang Y, Liu M, et al. // Journal of Gastrointestinal Surgery. – 2021. – Vol. 25(8). – P. 2011–2018. – doi: 10.1007/s11605–020–04802–4.

248. Komeda, Y. Computer–aided diagnosis based on convolutional neural network system for colorectal polyp classification: preliminary experience / Komeda Y., Handa H., Watanabe T. // Oncology. – 2017. – Vol. 93, N_{2} 1. – P. 30–34.

249. Wallace, M.B Impact of Artificial Intelligence on Miss Rate of Colorectal Neoplasia / Wallace M.B, Sharma P., Bhandari P., et al. // Gastroenterology. – 2022. – Vol. 163(1). – P. 295–304. – doi: 10.1053/j.gastro.2022.03.007.

250. Xu, H. Artificial Intelligence–Assisted Colonoscopy for Colorectal Cancer Screening: A Multicenter Randomized Controlled Trial / Xu H., Tang R.S.Y., Lam T.Y.T., et al. // Clin Gastroenterol Hepatol. – 2023. –Vol. 21(2). – P.337–346. – doi: 10.1016/j.cgh.2022.07.006.

251. Zhang, R. Automatic detection and classification of colorectal polyps by transferring low-level CNN features from nonmedical domain / Zhang R., Mak W., Zheng Y. // IEEE Journal of Biomedical and Health Informatics. – 2017. – Vol. 21, No 1. – P. 41–47.

252. Zhang, R. Polyp detection during colonoscopy using a regression-based convolutional neural network with a tracker / Zhang R., Zheng Y., Poon C.C.Y. // Pattern Recognition. – 2018. – Vol. 83, N_{2} 11. – P. 209–219.

253. Shin, Y. Abnormal colon polyp image synthesis using conditional adversarial networks for improved detection performance / Shin Y., Qadir H.A., Balasingham I. // IEEE Access. – 2018. – Vol. 6. – P. 56007–56017.

254. Jha, D. Real–Time Polyp Detection, Localization and Segmentation in Colonoscopy Using Deep Learning / Jha D., Ali S., Tomar N. K., Johansen H.D., Johansen D., Rittscher J., Riegler M.A., Halvorsen P. // Computer Vision and Pattern Recognition. – 2021. – P. 40496 – 40510.

255. Jha, D. Kvasir–SEG: A Segmented Polyp Dataset / Jha D. et al. // Proceedings of International Conference on Multimedia Modeling. – 2020. – P. 451–462.

256. Borgli, H. HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy [Электронный источник] / Н. Borgli, V. Thambawita, P. al.] H. Smedsrud [et // Sci Data. 2020. _ _ режим доступа: https://www.nature.com/articles/s41597-020-00622-y.

257. Smedsrud, P.H. Kvasir–Capsule, a video capsule endoscopy dataset / Smedsrud P.H., Thambawita V., Hicks S.A. et al. // Sci Data. – 2021. – Vol. 8. – P. 142.

258. Лебедев, А.А. Разработка алгоритма детектирования полипов на эндоскопических изображениях с использованием сверточных нейронных сетей / Лебедев А.А., Хрящев В.В., Среднякова А.С., Казина Е.М. // Цифровая обработка сигналов. – 2021. – № 2. – С. 55–60.

259. Лебедев, А.А. Разработка алгоритма детектирования полипов на видеопотоке с эндоскопической системы в режиме реального времени /

Лебедев А.А., Хрящев В.В. // СПбНТОРЭС: труды ежегодной НТК, 2022. – № 1(77). – С. 49–52.

260. Хрящев, В.В. Эффективный алгоритм обнаружения полипов на колоноскопических видеоданных / Хрящев В.В., Лебедев А.А., Приоров А.Л., Котов Н.В. // 78–я Всероссийская конференция «Радиоэлектронные устройства и системы для инфокоммуникационных технологий» (РЭУС–ИТ 2023). – Москва. – С. 232–237.

261. Хрящев, В.В. Сегментация изображений полипов при колоноскопическом исследовании с использованием нейронных сетей / Хрящев В.В. // Биомедицинская радиоэлектроника. – 2023. – Т. 26. № 4. – С. 66–72.

262. Хрящев, В.В. Использование алгоритмов цифровой обработки изображений в задаче попиксельного детектирования патологий в колоноскопии / Хрящев В.В., Приоров А.Л., Котов Н.В., Малыгин К.И. // Модели, системы, сети в экономике, технике, природе и обществе. – 2024. – № 4. – С. 86–97.

263. Свидетельство о государственной регистрации базы данных № 2024625493. IntestinesImages 1.0 – база изображений кишечника при проведении колоноскопических исследований / Хрящев В.В. – опубл. 26.11.2024.

264. Лебедев, А.А. Постобработка результатов работы нейросетевого детектора полипов на видеоизображениях колоноскопических исследований / Лебедев А.А., Приоров А.Л., Хрящев В.В., Среднякова А.С. // Цифровая обработка сигналов. – 2021. – № 4. – С. 67–71.

265. Хрящев, В.В. Исследование алгоритмов на базе нейросетевой архитектуры YOLO в задаче детектирования полипов на колоноскопических видеоданных / Хрящев В.В., Котов Н.В., Приоров А.Л. // 33–я Международная конференция по компьютерной графике и машинному зрению «ГрафиКон–2023». – Москва, ИПУ РАН. 2023. – С. 590–597.

266. Perone, C.S. Promises and limitations of deep learning for medical image segmentation / Perone C.S., Cohen–Adad J. // Journal of Medical Artificial Intelligence. – 2019. – Vol. 2. – doi: 10.21037/jmai.2019.01.01.

267. Wickstrom, K. Uncertainty and interpretability in convolutional neural networks for semantic segmentation of colorectal polyps / Wickstrom K., Kampffmeyer M., Jenssen R. // Medical Image Analysis. – 2020. – Vol. 60. – P. 1–18.

268. Пат. 2734058 С1 Российская Федерация. Система сегментации изображений зданий и сооружений / Островская А.А., Хрящев В.В. – № 2019141857; заявл. 17.12.2019; опубл. 12.10.2020, Бюл. № 29.

269. Пат. 2733823 С1 Российская Федерация. Система сегментации изображений участков недропользования открытого типа / Островская А.А., Хрящев В.В. – № 2019141854; заявл. 17.12.2019; опубл. 07.10.2020, Бюл. № 28.

270. Ahmed M., Seraj R., Islam S. M. S. The k-means algorithm: A comprehensive survey and performance evaluation // Electronics. – 2020. – T. 9. №. 8. – P. 1295.

271. Liu, D. Otsu method and K-means / Liu D., Yu J. // 2009 International conference on hybrid intelligent systems. IEEE. – 2009. – T. 1. – P. 344-349.

272. Ronneberger, O. U-Net: Convolutional Networks for Biomedical Image Segmentation / Ronneberger O., Fischer P., Brox T. // Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Lecture Notes in Computer Science. Springer. – 2015. – Vol. 9351. – P. 234-241.

273. Zhou, Z. UNet++: A Nested U-Net Architecture for Medical Image Segmentation [Электронный источник] / Zhou Z., Siddiquee M., Tajbakhsh N., Liang J. – режим доступа: https://arxiv.org/pdf/1807.10165

274. Jha, D. ResUNet++: An Advanced Architecture for Medical Image Segmentation / Jha D., Smedsrud P., Riegler M., Johansen D., Lange T. et. al. // in 21st IEEE International Symposium on Multimedia. – 2019.

275. Shi, W. SSformer: A Lightweight Transformer for Semantic Segmentation [Электронный источник] / Shi W., Xu J., Gao P. – режим доступа: https://arxiv.org/pdf/2208.02034.

276. Dumitru, R. Using DUCK-Net for polyp image segmentation / Dumitru R., Peteleaza D., Craciun C. // Scientific reports. – 2023. – T. 13. №. 1. – P. 9803.

277. Trinh, Q. Meta-Polyp: a baseline for efficient Polyp segmentation / Trinh Q. // 2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS). IEEE. – 2023. – P. 742-747.

278. Зырянов, Б.Н. Рак желудка / Зырянов Б.Н. – Томск, 1998. – 586 с.

279. Куваев, Р.О. Ранний рак желудка: методики скрининга, эндоскопической диагностики и малоинвазивного лечения / Куваев Р.О., Кашин С.В., Никонов Е.Л.,

Itoh T., Gotoda T., Gono K. // Доказательная гастроэнтерология, 2014, 3 (3), C. 44–51.

280. Куваев, Р.О. Helicobacter pylori–ассоциированный хронический гастрит: новые технологии эндоскопической диагностики / Куваев Р.О., Никонов Е.Л., Кашин С.В. // Доказательная гастроэнтерология, 2015, 4 (1), С. 19–24.

281. Куваев, Р.О. Современное эндоскопическое исследование желудка с использованием методик узкоспектральной и увеличительной эндоскопии: техника проведения и алгоритмы диагностики / Куваев Р.О., Кашин С.В. // Доказательная гастроэнтерология. – 2016. – Т. 2, № 5. – С. 3–13.

282. Itoh, T. Deep learning analyzes Helicobacter pylori infection by upper gastrointestinal endoscopy images / Itoh T., Kawahira H., Nakashima H. // Endoscopy International Open. 2018. Vol. 6, № 2. P. E139–E144.

283. Дунаева, О.А. Сегментация клинических эндоскопических изображений, основанная на классификации векторных топологических признаков / Дунаева О.А., Малкова Д.Б., Мячин М.Л., Эдельсбруннер Х. // Моделирование и анализ информационных систем, 20:6 (2013), С. 162–173.

284. Chan, A.T. Deep–learning system detects neoplasia in patients with barrett's esophagus with higher accuracy than endoscopists in a multistep training and validation study with benchmarking / Chan A.T., Williams C.S. // Gastroenterology. 2020. Vol. 158, N_{2} 4. P. 795–797.

285. Miyaki, R. A Computer System to Be Used With Laser–based Endoscopy for Quantitative Diagnosis of Early Gastric Cancer / Miyaki R., Tanaka S., Yoshida S. // Journal of clinical gastroenterology. 2014. Vol. 49, № 2. P. 1–14.

286. Li, L. Convolutional neural network for the diagnosis of early gastric cancer based on magnifying narrow band imaging / Li L., Shen Z., Chen Y., Zhang X. // Gastric Cancer. 2019. Vol. 23, № 2. P. 23–51.

287. Zhu, Y. Application of convolutional neural network in the diagnosis of the invasion depth of gastric cancer based on conventional endoscopy / Zhu Y., Wang Q., Xu M. // Gastrointest Endoscopy. 2019. Vol. 89, № 4. P. 806–815.

288. Ghatwary, N. Early esophageal adenocarcinoma detection using deep learning methods / Ghatwary N., Zolgharni M., Ye X. // International Journal of Computer Assisted Radiology and Surgery. 2019. Vol. 14, № 3. P. 1–11.

289. Khryashchev, V. Deep learning for gastric pathology detection in endoscopic images / Khryashchev V., Stepanova O., Lebedev A., Kashin S., Kuvaev R. // ACM International Conference Proceeding Series. 3rd International Conference on Graphics and Signal Processing. ICGSP 2019. – Hong Kong, 2019. – P. 90–94.

290. Lebedev, A. Convolutional neural network for early detection of gastric cancer by endoscopic video analysis / Lebedev A., Khryashchev V., Stefanidi A., Stepanova O., Kashin S., Kuvaev R. // Proc. SPIE 11433. Twelfth International Conference on Machine Vision (ICMV 2019).

291. Khryashchev, V. Analysis of Pathologies on Endoscopic Images of the Stomach Using SSD and RetinaNet Neural Network Architecture / Khryashchev V., Lebedev A., Stepanova O., Srednyakova A. // Proceedings of 2021 IEEE East–West Design and Test Symposium (EWDTS 2021), September 2021. – Batumi, Georgia, 2021. – P. 119–123.

292. Среднякова, А. С. Разработка системы поддержки принятия решений для врача-патоморфолога на основе алгоритмов машинного обучения / Среднякова А.С., Лебедев А.А., Хрящев В.В. // Нейрокомпьютеры и их применение: тезисы докладов 16-й Всероссийской научной конф. – Москва, МГППУ. 2018. – С. 226–227.

293. Лебедев, А. А. Разработка алгоритмов анализа изображений для классификации патологий слизистой оболочки желудка / Лебедев А.А., Степанова О.А., Юрченко Е.А., Хрящев В.В. // Цифровая обработка сигналов и ее применение (DSPA–2018): докл. 20–й междунар. конф. – Москва, 2018. – Т. 2. – С. 644–649.

294. Степанова, О.А. Использование глубокого обучения для анализа эндоскопического изображения желудка / Степанова О.А., Хрящев В.В. // Современные технологии обработки сигналов: Сб. докладов 1-й Всероссийской конференции. – Москва, 2018. – С. 248–252.

295. Степанова, О.А. Использование сверточной нейронной сети SSD для детектирования патологий при эндоскопии желудка / Степанова О.А., Лебедев А.А., Хрящев В.В., Приоров А.Л. // Цифровая обработка сигналов и ее применение (DSPA–2019): докл. 21–й междунар. конф. – Москва, 2019. – С. 533–537.

296. Лебедев, А.А. Распознавание патологий желудка по видеоизображениям эндоскопического исследования / Лебедев А.А., Хрящев В.В., Степанова О.А. // Новые информационные технологии и системы: сб. науч. ст. по материалам XVII Междунар. науч.–техн. конф. – Пенза, 2020. – С. 200–203.

297. Хрящев, В.В. Использование глубоких нейронных сетей для поиска патологий на эндоскопических изображениях желудка / Хрящев В.В., Среднякова А.С., Ганин А.Н., Кашин С.В. // Цифровая обработка сигналов и ее применение (DSPA–2021): докл. 23–й междунар. конф. – Москва, 2021. – С. 254–258.

298. Завьялов, Д.В. Алгоритм уточняющей диагностики и внутрипросветного эндоскопического удаления эпителиальных новообразований толстой кишки / Завьялов Д.В., Кашин С.В., Нестеров П.В., Нечипай А.М., Рыбачков В.В., Малашенко В.Н., Хрящев В.В. // Колопроктология. – Т. 20. № 1. – 2021. – С. 17–22.

299. Лебедев, А.А. Применение методов глубокого обучения для поддержки врачебного решения при эндоскопическом исследовании желудка / Лебедев А.А., Хрящев В.В., Кашин С.В., Среднякова А.С., Казина Е.М. // Модели, системы, сети в экономике, технике, природе и обществе. – 2021. – № 2(38). – С. 95–106.

300. Хрящев, В.В. Использование нейронной сети EfficientDet в задаче обнаружения патологий желудка на видеоизображениях эндоскопического исследования / Хрящев В.В., Приоров А.Л. // Модели, системы, сети в экономике, технике, природе и обществе. – 2023. – № 2. – С. 185–192.

301. Khryashchev, V.V. A medical decision support system using an artificial intelligence module for endoscopic examination of the stomach / Khryashchev V.V. // Biomedical Engineering. – 2024. – Vol. 57, P. 423–427.

302. Хрящев, В.В. Исследование нейросетевых алгоритмов для системы поддержки принятия врачебного решения при исследовании желудка в эндоскопии / Хрящев В.В. // Доклады XXV международной научно–технической конференции «Нейроинформатика – 2023». – Москва, НИЯУ МИФИ. 2023. – С. 380–390.

303. Хрящев, В.В. Исследование нейросетевых алгоритмов детектирования патологий на видеоданных гастроскопических исследований желудка / Хрящев В.В. // Проблемы информатики в образовании, управлении, экономике и

технике: сборник статей XXIV Международной научно-технической конференции. – Пенза: изд-во ПГУ. 2024. – С. 176–179.

304. Свидетельство о государственной регистрации базы данных № 2024625397. StomachImages 1.0 – база изображений желудка при проведении гастроскопических исследований / Хрящев В.В. – опубл. 22.11.2024.

305. Котов, Н.В. Исследование нейросетевой архитектуры YOLO для обнаружения изображения купола слепой кишки в колоноскопии / Котов Н.В., Хрящев В.В., Тихомиров А.А. // Цифровая обработка сигналов и ее применение (DSPA–2024): докл. 26–й междунар. конф. – Москва, 2024. – С. 151–155.

306. Хрящев, В.В. Анализ видеопотока с колоноскопических исследований на основе методов глубокого машинного обучения / Хрящев В.В. // Современные информационные технологии. Теория и практика: Материалы VI научно– практической конференции. – Череповец. ЧГУ. 2024. – Ч. 2. – С. 308–311.

307. Хрящев, В.В. Обнаружение полипов на колоноскопических изображениях при помощи алгоритмов на базе нейросетевой архитектуры YOLO / Хрящев В.В., Приоров А.Л., Лебедев А.А., Котов Н.А. // Цифровая обработка сигналов. – 2023. – № 3. – С. 27–31.

308. Khryashchev, V. Deep learning architecture for real-time polyps' detection on endoscopic video / Khryashchev V., Lebedev A., Priorov A., Kotov N., Zavyalov D. // Systems of Signal Synchronization, Generating and Processing in Telecommunications "SYNCHROINFO-2023": Conference Proceedings. – Pskov, Russia, 2023. – C. 1–5.

309. Хрящев, В.В. Тестирование нейросетевых алгоритмов детектирования патологий на эндоскопических видеоизображениях желудка / Хрящев В.В. // Сборник трудов XVII Международной отраслевой научно-технической конференция «Технологии информационного общества». – Москва, МТУСИ. – 2023. – С. 40–42.

310. Андержанова, А.С. Разработка нейросетевого алгоритма классификации патологий на эндоскопических видеоизображениях желудка / Андержанова А.С., Степанова О.А., Хрящев В.В., Волков Д.Б. // СПбНТОРЭС: труды ежегодной НТК, 2022. – № 1 (77). – С. 20–22.

311. Хрящев, В.В. Разработка и тестирование нейросетевого алгоритма сегментации эндоскопических изображений желудка / Хрящев В.В. //

78-я Всероссийская конференция "Радиоэлектронные устройства и системы для инфокоммуникационных технологий" (РЭУС–ИТ 2023). – Москва, 2023. – С. 245–249.

312. Хрящев, В.В. Разработка системы поддержки принятия врачебного решения для эндоскопических исследований желудка / Хрящев В.В. // Цифровая обработка сигналов и ее применение (DSPA–2023): докл. 25–й междунар. конф. – Москва, 2023. – С. 249–253.

313. Хрящев, В.В. Система поддержки принятия врачебного решения с использованием модуля искусственного интеллекта для эндоскопических исследований желудка / Хрящев В.В. // Медицинская техника. – 2023. – № 6. – С. 44–47.

314. Свидетельство о государственной регистрации программы для ЭВМ № 2025616549. EndoscopyVA.RT – программа для анализа эндоскопических видеоизображений в реальном времени на основе методов искусственного интеллекта / Хрящев В.В. – опубл. 18.03.2025.

315. Свидетельство о государственной регистрации программы для ЭВМ № 2025616548. EndoscopyDSS.Hub – программа для визуализации результатов эндоскопических исследований с учетом анализа видеоданных методами искусственного интеллекта / Хрящев В.В. – опубл. 18.03.2025.

ПРИЛОЖЕНИЕ А. ДОКУМЕНТЫ О ВНЕДРЕНИИ

- Акт о внедрении результатов диссертационного исследования в ФГБУ «НМХЦ им. Н.И. Пирогова» Минздрава России (Пироговский центр, г. Москва);
- Акт о внедрении результатов диссертационного исследования в ГБУЗ ЯО «Клиническая онкологическая больница» (г. Ярославль);
- Акт о внедрении результатов диссертационного исследования
 в Медицинский центр диагностики и профилактики (г. Ярославль);
- Акт о внедрении результатов диссертационного исследования в ООО «ЭМИС» (г. Ярославль);
- Акт о внедрении результатов диссертационного исследования в ООО «А-Вижн» (г. Ярославль);
- Справка об участии в выполнении НИР в ФГБОУ ВО «Ярославский государственный университет им. П.Г. Демидова».

МИНИСТЕРСТВО ЗДРАВООХРАНЕНИЯ «УТВЕРЖДАЮ» РОССИЙСКОЙ ФЕДЕРАЦИИ Заместитель генерального директора Федеральное государственное бюджетное учреждение Национальный по научной и образовательной медико-хирургический Центр деятельности имени Н.И. Пирогова (ФГБУ «НМХЦ им. Н.И. Пирогова» Минздрава России) ФГБУ «НМХЦ им. Н.И. Пирогова» 105203, Москва, ул. Нижняя Первомайская, дом 70 Минздрава России тел: +7 (499) 464-49-54, факс: +7 (495) 465-33-55, факс +7 (495) 465-09-52, e-mail: info@pirogov-center.ru З А.А. Пулин ОКПО 55246260 ОГРН 1027700077668 ИНН/КПП 7733108569/771901001 2025 г. 25.04.2025 No на №

АКТ

о внедрении результатов диссертационного исследования
 «Система поддержки принятия врачебных решений
 на основе анализа эндоскопических видеоизображений
 с применением методов искусственного интеллекта»

Настоящим актом подтверждается, что положения и результаты диссертационного исследования Хрящева Владимира Вячеславовича внедрены в практическую деятельность Института усовершенствования врачей ФГБУ «НМХЦ им. Н.И. Пирогова» Минздрава России.

Дата внедрения предложения: июнь 2022 года.

Предложено к внедрению: программно-аппаратный комплекс для системы поддержки принятия врачебного решения при колоноскопическом исследовании.

Используются следующие результаты диссертационной работы:

- нейросетевой алгоритм, позволяющий обнаруживать колоректальные полипы на видеопотоке с эндоскопической системы в режиме реального времени с точностью 80% и выше;
- методика подготовки базы видеоданных колоноскопических исследований с разметкой областей патологии для обучения нейросетевых моделей, позволяющая сократить соответствующие трудозатраты на 35-45%.

Использование предложенных результатов в медицинских видеоэндоскопических системах увеличивает информативность размеченных видеоматериалов, улучшает навигацию по ним, повышает достоверность эндоскопической диагностики, а также способствует контролю качества проведенных колоноскопических исследований.

Заведующий эндоскопическим отделением, профессор кафедры хирургии с курсом хирургической эндокринологии Института усовершенствования врачей ФГБУ «НМХЦ им. Н.И. Пирогова» Минздрава России, врач-эндоскопист, доктор медицинских наук, доцент

Maar got

Маады Аяс Сергеевич



АКТ

N⁰

OT

внедрения в ГБУЗ ЯО «Клиническая онкологическая больница» результатов диссертационного исследования Хрящева В.В., представленного на соискание ученой степени доктора технических наук

Настоящим актом комиссия в составе: Кашин Сергей Владимирович – заведующий эндоскопическим отделением, Гусейнова Сабина Руслановна – врач-эндоскопист, Куваев Роман Олегович – врач-гастроэнтеролог подтверждает факт того, что результаты диссертационного исследования Хрящева Владимира Вячеславовича «Система поддержки принятия врачебных решений на основе анализа эндоскопических видеоизображений с применением методов искусственного интеллекта» используются в работе эндоскопического отделения ГБУЗ ЯО «Клиническая онкологическая больница». Применяются следующие результаты диссертационной работы:

 программно-аппаратный комплекс, включающий алгоритмы для детектирования, сегментации и классификации патологий желудка, обрабатывающие эндоскопический видеопоток в режиме реального времени; нейросетевой алгоритм детектирования области купола слепой кишки на видеоданных для контроля качества скрининговых колоноскопических исследований.

Использование указанных результатов повышает эффективность эндоскопической диагностики желудочно-кишечного тракта на основе анализа видеоданных методами искусственного интеллекта.

Председатель: заведующий эндоскопическим отделением

UNIX

Кашин С.В.

Гусейнова С.Р.

Врач-эндоскопист

Врач-гастроэнтеролог высшей квалификационной категории

Куваев Р.О.

П.В. Нестеров

Главный врач

УТВЕРЖДАЮ» Генеральный директор ЭОО «Медицинский центр ностики и профилактики» Т.В. Крюкова мая 2025 г.

АКТ

о внедрении результатов диссертационной работы Хрящева Владимира Вячеславовича «Система поддержки принятия врачебных решений на основе анализа эндоскопических видеоизображений с применением методов искусственного интеллекта»

Мы, нижеподписавшиеся: Завьялов Дмитрий Вячеславович, заведующий отделением эндоскопии, Гусев Денис Владимирович – врач-эндоскопист, Гончаров Виталий Иванович – врач-эндоскопист удостоверяем факт того, что результаты диссертационного исследования Хрящева Владимира Вячеславовича, представленного на соискание ученой степени доктора технических наук по специальности 2.2.12 «Приборы, системы и изделия медицинского назначения» используются при проведении скрининговых колоноскопических исследований в ООО «Медицинский центр диагностики и профилактики».

Используются следующие результаты диссертационной работы:

- программно-аппаратный комплекс, позволяющий детектировать полипы в режиме реального времени;
- программный комплекс для локализации области купола слепой кишки на записанной видеопоследовательности колоноскопического исследования с целью контроля качества выполненной процедуры.

Предложенные в диссертационной работе алгоритмы анализа видеопотока с использованием методов искусственного интеллекта позволяют увеличить достоверность колоноскопий при скрининговом обследовании пациентов.

Заведующий отделением эндоскопии

Завьялов Д.В.

Гусев Д.В.

Врач-эндоскопист

Врач-эндоскопист

Гончаров В.И.

Общество с ограниченной ответственностью «Эндоскопические медицинские информационные системы» 150040, г. Ярославль пр. Октября 34/21, 123. ИНН 7606083512 / КПП 760601001 <u>info@emis-info.ru</u> www.emis-info.ru тел. (4852) 33-77-53

«УТВЕРЖДАЮ»

Генеральный директор ООО «ЭМИС» Р.Е. Плисс мялицинские 2025 г. информационные АКТ

о внедрении результатов диссертационной работы Хрящева Владимира Вячеславовича «Система поддержки принятия врачебных решений на основе анализа эндоскопических видеоизображений с применением методов искусственного интеллекта»

Настоящим актом Общество с ограниченной ответственностью «ЭМИС» (ООО «ЭМИС») подтверждает, что результаты диссертационного исследования Хрящева Владимира Вячеславовича, представленного на соискание ученой степени доктора технических наук по специальности 2.2.12 «Приборы, системы и изделия медицинского назначения» используются при разработке специализированной отечественной медицинскойинформационной системы «ЭМИС».

Предложенные в диссертационной работе алгоритмы анализа цифровых изображений И видеопоследовательностей С применением методов искусственного интеллекта используются ДЛЯ контроля качества эндоскопического исследования, что позволяет повысить эффективность данного метода медицинской диагностики при скрининговом обследовании пациентов.

Технический директор

phanparal

В.А. Капранов


внедрения результатов диссертационной работы Хрящева Владимира Вячеславовича, представленной на соискание ученой степени доктора технических наук

Настоящим актом подтверждается то, что результаты диссертационной работы Хрящева Владимира Вячеславовича, представленной на соискание ученой степени доктора технических наук по специальности 2.2.12 «Приборы, системы и изделия медицинского назначения» внедрены ООО «А-Вижн» при создании системы анализа видеопотока и контроля качества выполнения колоноскопического исследования на основе методов искусственного интеллекта.

Внедрены следующие результаты, полученные Хрящевым В.В. в рамках выполнения диссертационной работы:

- Система для контроля качества колоноскопического исследования, позволяющая детектировать купол слепой кишки и отличающаяся использованием алгоритмов сопровождения областей интереса и анализа оптического потока.
- Алгоритмы детектирования/сегментации полипов на колоноскопических видеоданных, работающие в условиях ограниченности входного набора аннотированных видеоизображений.



Конюхов А.С.

Куракин А.С.

Молчанов М.С.

Главный инженер

Заместитель начальника технического отдела

Ведущий программист



АКТ

об использовании результатов диссертационной работы Хрящева Владимира Вячеславовича

«Система поддержки принятия врачебных решений на основе анализа эндоскопических видеоизображений с применением методов искусственного интеллекта» на соискание ученой степени доктора технических наук по специальности 2.2.12. Приборы, системы и изделия медицинского назначения

Настоящим подтверждается, что результаты научных и практических работ, полученные Хрящевым В.В. при подготовке диссертационной работы «Система поддержки принятия врачебных решений на основе анализа эндоскопических видеоизображений с применением методов искусственного интеллекта» использовались при выполнении следующих проектов:

1. № 19-37-90153 «Исследование применимости технологий глубокого обучения и сверточных нейронных сетей для распознавания патологий желудка по видеоизображениям эндоскопического исследования», проводимого при поддержке Российского фонда фундаментальных исследований в 2019-2021 гг.,

2. № 2338-х/д, «Обучение и тестирование моделей нейронных сетей для сегментации, распознавания и классификации эндоскопических исследований», проводимого при поддержке ООО «СМАРТ АНАЛИЗ» в 2020-2021 гг.

Начальник управления научных исследований и инноваций канд. экон. наук, доцент

- Oynuf

Дугин А.Н.