

Федеральное государственное автономное образовательное учреждение
высшего образования

**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

На правах рукописи

Лаптев Владислав Витальевич

**СЕГМЕНТАЦИЯ ОБЪЕКТОВ С ОТСУТСТВИЕМ ЯВНЫХ
ХАРАКТЕРИСТИЧЕСКИХ ПРИЗНАКОВ НА МЕДИЦИНСКИХ
ИЗОБРАЖЕНИЯХ**

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

Диссертация на соискание ученой степени
Кандидат технических наук

Научный руководитель:
Доктор технических наук, профессор
Гергет Ольга Михайловна

Томск – 2025

Оглавление

ВВЕДЕНИЕ.....	5
Глава 1. Анализ предметной области и современных методов сегментации сложных объектов исследования.....	14
1.1 Представление изображений в медицине	14
1.2 Сегментация и локализация анатомических структур	17
1.3 Подготовка данных	22
1.3.1 Вычислительная сложность	23
1.3.2 Типы машинного обучения	25
1.3.3 Методы активации картирования классов.....	26
1.4 Архитектуры нейронной сети семантической сегментации	28
1.5 Математическая формулировка задачи семантической сегментации.....	31
1.6 Программные продукты для обработки медицинских данных.....	31
1.7 Выводы по первой главе.....	34
Глава 2. Метод и алгоритмы сегментации сложных объектов медицинских изображений с отсутствием явных характеристических признаков	36
2.1 Исходные данные и оборудование.....	36
2.2 Технология генерации сегментационной маски изображения.....	40
2.3 Метод формирования исходной маски изображения.....	43
2.3.1 Алгоритм формирования карты активации признаков	43
2.3.2 Алгоритм формирования трехканальной карты объекта исследования.....	51
2.4 Алгоритм аугментации данных	57

2.5	Метод обучения модели	59
2.6	Выводы по второй главе.....	66
Глава 3. Применение разработанных методов и алгоритмов в задачах подготовки данных и обучения моделей сегментации сложных объектов исследования.....		
3.1	Создание исходной маски объектов на основе карт признаков	68
3.1.1	Карты признаков	68
3.1.2	Результат улучшения маски с использования трехканального анализа.....	78
3.1.3	Выводы по работе метода создания сегментационных масок изображений.....	81
3.2	Результат обучения моделей в задаче сегментации нестабильных бляшек.....	83
3.2.1	Дизайн эксперимента.....	83
3.2.2	Тюнинг нейросетевых моделей	84
3.2.3	Обучение ансамбля нейросетевых моделей	86
3.3	Апробация предложенного метода в задаче сегментации гистологических аномалий	93
3.4	Выводы по третьей главе.....	104
ЗАКЛЮЧЕНИЕ		106
СПИСОК СОКРАЩЕНИЙ И ОПРЕДЕЛЕНИ		110
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ		112
Приложение А Оптимальные конфигурации сетей для исследуемых морфологических признаков бляшки.....		128

Приложение Б Показатели производительности исследуемых моделей на различных конфигурация набора данных.....	131
Приложение В Визуализация результатов работы модели сегментации атеросклеротической бляшки	135
Приложение Г Дополнительная информация к задаче сегментации гистологических аномалий.....	136
Приложение Д Акты внедрения результатов диссертационной работы.....	139
Приложение Е Свидетельства о государственной регистрации программы для ЭВМ	143

ВВЕДЕНИЕ

Актуальность темы исследования. В настоящее время алгоритмы анализа изображений с использованием средств машинного обучения стали неотъемлемой частью нашей жизни. С помощью глубокого обучения искусственный интеллект (ИИ) может выполнять такие задачи, как распознавание лиц [1-3], дорожных ситуаций [4-5], обработка и анализ снимков со спутника [6-8], медицинских изображений [9-11]. Медицинская диагностика стала одной из самых востребованных областей для применения ИИ. К популярным задачам относятся: локализация опухолей или других новообразований, оценка активности мозга, изучение плотности тканей и т.д. С помощью алгоритмов компьютерного зрения можно находить закономерности и аномалии на многомодальных снимках, таких как рентген, КТ, МРТ. Целью анализа медицинских данных является содействие специалисту для более быстрой и точной постановки диагноза, а также ассистирование при проведении процедуры интервенции.

Одной из основных задач обработки и анализа изображений является сегментация. Выделение на изображении области, согласно определенному критерию, позволяет не только обнаружить интересующий объект на изображении, но и получить значимые характеристики, такие как: границы объекта, форма, площадь и т.п. [12]. Существует большое множество алгоритмов сегментации, которые можно разделить на группы методов, основанные на:

- энтропии;
- кластеризации;
- локальных порогах;
- пиксельных корреляциях и атрибутах;
- искусственных нейронных сетях.

В настоящее время наибольший интерес представляют системы, основанные на алгоритмах, использующих искусственных нейронных сетей. Именно эта группа обладает такой характеристикой как обобщающая способность, что позволяет применять их на более широком спектре задач. Но стоит отметить, что для получения такой системы сегментации объектов высокой точности требуется большое количество качественных обучающих данных. Под обучающими данными подразумевается набор изображений и аннотаций к ним (описание анализируемого изображения). Как правило составлением аннотаций к данным занимаются эксперты в данной области, а под аннотацией сегментационной модели понимается маска изображения. Зачастую именно сбор данных и составление аннотаций занимает наибольшее количество времени при разработке систем компьютерного зрения. Причем составление сегментационной маски изображения является самым трудоемким из них [12].

В связи с высокими трудозатратами на создание качественных аннотаций особую актуальность приобретает анализ существующих подходов к автоматизации этого процесса. На сегодняшний день существует множество работ, посвященных сегментации объектов в общем контексте, а именно автомобилей, людей, дорог и других объектов [12-15]. В свою очередь, работы, направленные на сегментацию медицинских инструментов [16-17] или анатомических структур [18-19], представлено значительно меньшим числом, что объясняется низким качеством выходных данных, большой вариабельностью признаков, малым объемом выборок, высоким уровнем зашумленности и отсутствием глобального контекста и т.д.

Обнаружение и анализ сложных объектов исследования является актуальной задачей в настоящее время. Под сложным объектом исследования в данной работе понимается объект, не имеющий явных характеристических признаков. В частности, при диагностике заболеваний часто анализируют изображения, содержащие слабовыраженные объекты, полученные с помощью

специализированных методов визуализации, либо объекты микромира (например, клетки, ядра клеток). Следует отметить, что изображения такого типа часто имеют низкое качество. При этом визуализация различных органов, тканей, отдельных частей скелета человека чаще всего осуществляется с использованием специальных методов, таких как магнитно-резонансная томография, позитронно-эмиссионная томография, ультразвуковое исследование, а также методы световой и электронной микроскопии. Поскольку современные технологии позволяют получать все более крупные и сложные данные, для извлечения ценной информации из них необходимо использовать специализированные методы обнаружения и анализа сложных объектов. Обнаружение и анализ сложных объектов могут быть использованы в различных областях, таких как медицина, финансы, наука, технологии и другие. Например, в области медицины при диагностике заболеваний нередко приходится иметь дело с изображениями, на которых отображены едва уловимые объекты, полученные посредством специализированных исследований, либо объекты микромира (например, клетки, ядра клеток). При этом следует отметить, что зачастую такие изображения низкого качества. В качестве образа реальных объектов выступают изображения различных органов, тканей, отдельных частей скелета человека, которые получены при помощи специальных методов их визуализации:

- магниторезонансной томографии;
- позитронно-эмиссионной томографии;
- ультразвукового анализа;
- световой и электронной микроскопии.

Создание сегментационных масок изображений нестандартных мультимодальных данных является еще более сложной и трудоемкой задачей [21]. В связи с этим существует необходимость в разработке новых методов сегментации и визуализации объектов на входных данных низкого качества в режиме реального времени.

Целью диссертационной работы является разработка эффективной технологии подготовки сегментационных масок медицинских изображений, основанной на алгоритмах анализа карт признаков с последующей оптимизацией границ объектов с использованием FBA-Net модели, что позволяет повысить точность сегментации и сократить затраты на подготовку обучающих данных в условиях ограниченной аннотированной информации.

Для достижения поставленной цели в диссертации необходимо решить следующие исследовательские **задачи**:

- Провести комплексный анализ предметной области, включающий исследование специфики медицинской визуализации и задач сегментации объектов с отсутствием явных характеристических признаков.
- Выполнить анализ медицинских данных различных модальностей, включающий изучение их структурных особенностей, форматов хранения, специфических характеристик, методов их предварительной обработки, алгоритмов сегментации и классификации, а также современных подходов к визуализации медицинских изображений.
- Исследовать современные архитектуры сверточных нейронных сетей для задачи классификации.
- Провести анализ существующих методов и алгоритмов сегментации и локализации медицинских инструментов и анатомических структур.
- Разработать технологию автоматизированного формирования сегментационных масок медицинских изображений на основе анализа карт признаков классификационных моделей с последующей оптимизацией границ объектов.
- Предложить и исследовать метод автоматического выделения областей интереса на изображениях, а также разработать алгоритм формирования адаптивной трехканальной сегментационной карты.

- Реализовать алгоритм подбора оптимальных гиперпараметров для обучения глубоких нейронных сетей и провести вычислительный эксперимент с целью анализа эффективности предлагаемого метода сегментации в контексте медицинских задач.

Объектом исследования являются методы и алгоритмы сегментации и локализации объектов на медицинских изображениях.

Предметом исследования являются методы машинного и глубокого обучения, с помощью которых осуществляется сегментация и локализация области интереса.

Методы исследования. Выполненные в диссертационной работе исследования основаны на использовании методов цифровой обработки изображений, математического моделирования и алгоритмов машинного обучения.

Научная новизна диссертационной работы представлена следующими составляющими:

- Разработана технология автоматизированной сегментации изображений, базирующаяся на интерпретации признаков сверточных классификационных нейросетей с использованием модели FBA-Net, обеспечивающая точное выделение области интереса за счет последовательного потока данных между этапами, где каждый этап преобразует тензоры изображений с минимальными потерями информации. *Разработка специального математического и алгоритмического обеспечения систем анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта (пункт 5 паспорта специальности)*

- Предложен метод автоматического выделения областей интереса на изображениях по картам признаков классификационной модели посредством интерпретации пространственной информации, извлеченной из промежуточных слоев сверточной нейросети, что обеспечивает корректную сегментацию даже для

объектов с неявными характеристическими признаками. *Разработка методов и алгоритмов решения задач системного анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта (пункт 4 паспорта специальности)*

— Для эффективного применения модели FBA-Net разработан алгоритм автоматизированного формирования трехканальных карт на основе разделения изображения на зоны объекта, фона и области неопределенности, что обеспечивает детальный анализ объекта исследования и повышает точность сегментации анализируемого изображения. *Визуализация, трансформация и анализ информации на основе компьютерных методов обработки информации (пункт 12 паспорта специальности)*

Практическая ценность работы заключается в создании программного комплекса, позволяющего решать задачи подготовки большого обучающего набора данных и проводить обучение моделей семантической сегментации с использованием оптимального набора гиперпараметров. Кроме того, модели и методы, используемые для обработки медицинских данных, применяются при выполнении минимально инвазивных вмешательств в хирургической онкологии, нейрохирургии и сосудистой хирургии.

Реализация результатов. Результаты диссертационного исследования использованы при выполнении следующих проектов: национальный проект «Наука» № FSWW-2023-0007; национальный проект «Наука и университеты» для выполнения темы НИР НИИ комплексных проблем сердечно-сосудистых заболеваний № 0419-2024-0001 «Новые методы антиатеросклеротической терапии и решения на основе машинного обучения для автоматизированной диагностики и прогнозирования сердечно-сосудистых заболеваний»; грант РНФ № 23-75-10009 «Разработка автоматизированной системы выявления нестабильных бляшек на основе оптической когерентной томографии с использованием методов глубокого обучения».

Разработанные алгоритмы подготовки данных семантической сегментации внедрены в программные комплексы обучения моделей распознавания объектов на снимках в ООО «Fabula» (г. Томск) – при подготовке обучающего множества моделей семантической сегментации; в НИИ комплексных проблем сердечно-сосудистых заболеваний (г. Кемерово) для визуального ассистирования при проведении процедуры интервенции; в Сибирском государственном медицинском университете при визуальном анализе запатентованной системы «Офтальмик+».

Основные положения, выносимые на защиту:

- 1) Разработанная технология автоматизированной сегментации изображений позволяет сократить время подготовки обучающего множества более чем в 3 раза для выраженных структур по сравнению с ручным аннотированием.
- 2) Предложенный метод автоматического выделения областей интереса на изображениях обеспечивает сокращение трудозатрат на формирование сегментационных масок в среднем на 67%, что, в итоге, позволяет снизить трудоемкость подготовки данных для обучения нейросетевых моделей семантической сегментации.
- 3) Разработанный алгоритм построения трехканальных сегментационных карт обеспечивает повышение точности сегментации границ сложных анатомических структур в среднем на 20 % по сравнению с традиционными методами.

Достоверность результатов, представленных в диссертационном исследовании, подтверждается репрезентативной апробацией обученных моделей, демонстрирующей сопоставимость полученных метрик с эталонными показателями. Методологическая база исследования опирается на общепризнанные подходы к проектированию глубоких нейронных сетей для анализа визуальных данных, адаптированные с учетом специфики решаемых задач. Валидность предложенных алгоритмов и методов верифицирована в ходе серии

экспериментов, а их работоспособность подтверждается воспроизводимостью результатов при обработке разнородных графических данных.

Апробация результатов работы. Результаты диссертационной работы обсуждались и докладывались на следующих симпозиумах, семинарах и конференциях: 32-ая Международная научная конференция «GraphiCon 2022», г. Рязань, 2022 г.; 20-я международная научно-практическая конференция студентов, аспирантов и молодых ученых «Молодежь и современные информационные технологии», г. Томск, 2023 г. (доклад отмечен дипломом I степени); всероссийская конференция с международным участием «Молодые лидеры в медицинской науке», г. Томск, 2023 г.; третий всероссийском научно-образовательный форум с международным участием «Кардиология XXI века: альянсы и потенциал», г. Томск, 2023 г.; XXV Московский Международный Конгресс по рентгенэндоваскулярной хирургии, г. Москва, 2023 и 2024 г.; международная научно-практическая конференция «Электронные средства и системы управления», г. Томск, 2024 г. (доклад отмечен дипломом I степени).

Личный вклад соискателя. Основные научные результаты, выносимые на защиту получены автором лично. Данная работа выполнена в сотрудничестве с учеными и исследователями, оказавшими содействие при проведении настоящего исследования, а именно: постановка задач исследования по теме диссертации выполнена совместно с научным руководителем Гергет О.М. (д.т.н., профессор Томского политехнического университета); сбор, разметка и валидация данных оптической когерентной томографии выполнена совместно с Кочергиным Н.А. (д.м.н., НИИ комплексных сердечно-сосудистых заболеваний), обсуждение результатов и дальнейших разработок выполнена совместно с Даниловым В.В. (к.т.н., университет Помпеу Фабры, Барселона).

Публикации. Основные результаты исследований отражены в 10 публикациях: 7 статей в ведущих научных журналах рекомендованных ВАК, 5 из них проиндексированы в базах данных Scopus и (или) Web of Science (в том числе,

две статьи в журнале квартиля Q1); 3 статьи в сборниках трудов международных и российских конференций; 2 свидетельства о государственной регистрации программы для ЭВМ.

Структура и объем диссертации. Работа включает в себя введение, три главы, заключение, список литературы из 113 источников и приложения. Работа изложена на 145 страницах и включает в себя 31 рисунок и 11 таблиц.

Глава 1. Анализ предметной области и современных методов сегментации сложных объектов исследования

В данной главе описаны существующие подходы обработки медицинских данных с целью анализа сложных объектов исследования. Сложными объектами исследования являются структуры с отсутствием явных характеристических признаков. Основными особенностями таких данных являются:

- низкая различимость – схожесть визуальных характеристик с фоном и другими объектами исследования;
- отсутствие четких границ;
- малый размер;
- высокая вариативность формы и текстуры.

Такие структуры представляют значительную сложность для алгоритмов машинного обучения из-за неоднородности данных, шумов и ограниченности экспертной разметки. Особый интерес исследователей в таких отраслях, как: медицинская визуализация [21-22], промышленность [24] и экология [25] демонстрируют междисциплинарную значимость данного исследования.

В главе представлен детальный анализ современных архитектурных нейросетевых решений семантической сегментации, а также существующих методов и алгоритмов сегментации и локализации медицинских инструментов и анатомических структур. Приведены результаты исследования существующих программных продуктов, позволяющих автоматизировать процесс анализа медицинских данных.

1.1 Представление изображений в медицине

Современная практика хранения и передачи медицинских данных опирается на ряд универсальных стандартов, включая Analyze, Minc, NIfTI и DICOM [26-27].

Выбор формата существенно влияет на этапы и эффективность обработки информации, в особенности медицинских изображений. Некоторые из этих форматов имеют узкую специализацию и предназначены для изображений, полученных при конкретных методах исследования, тогда как другие, в частности DICOM, являются общепринятыми и поддерживают хранение практически любых типов медицинской визуализации. DICOM утвердился как наиболее распространенный стандарт, что обусловлено его повсеместной поддержкой в современном диагностическом оборудовании (МРТ, КТ, ПЭТ, ЭхоКГ) [28]. Широкое внедрение DICOM в России связано с импортом аппаратуры из стран, где он является де-факто стандартом, делая разработку национального формата нецелесообразной [29]. Данное распространение стимулировало создание множества программных продуктов для обработки и визуализации медицинских изображений в формате DICOM, которые могут поставляться вместе с оборудованием или функционировать как самостоятельные приложения [26], [28].

Важно отметить, что современные томографы и эхокардиографы генерируют не готовые изображения, а DICOM-файлы, содержащие как данные среза выбранной анатомической области, так и обширные метаданные. Последние включают дату исследования, информацию о пациенте, параметры изображения, данные для рендеринга и характеристики оборудования [26], [28]. Каждый файл кодирует количественные характеристики тканей в срезе (плотность для КТ или интенсивность сигнала для МРТ/ЭхоКГ), на основе которых строится итоговое изображение [26]. Физическая основа этих значений различается: рентгеновская плотность (КТ) коррелирует с физической плотностью ткани (высокая у костей, низкая у крови и мягких тканей), тогда как МРТ и ЭхоКГ регистрируют интенсивность обратного сигнала [29]. Данные о плотности, интенсивности хранятся в виде цифрового изображения (с разрешением съемки, размером пикселя/вокселя, методом представления), где значение пикселя соответствует измеренной характеристике ткани, а не цвету [28]. Визуализация, таким образом,

требует преобразования этих значений в цветовую карту посредством передаточной функции визуализатора. Часто в клинической практике используются готовые шаблоны (LUT - Look-Up Tables) для цветового кодирования различных диапазонов плотности/интенсивности [26], [28].

Подавляющее большинство методов 3D-реконструкции в медицине используют наборы двумерных срезов (слайсов), полученных с помощью различных модальностей (МРТ, ПЭТ, КТ и др.) [30]. Однако способы формирования этих изображений различаются: КТ-сканеры производят параллельные срезы с высоким контрастом, в то время как ультразвуковые сканеры и эхокардиографы могут создавать изображения с более низким контрастом как в параллельных, так и в веерообразных (под углом) геометриях [30]. Пример визуализации DICOM-данных представлен на рисунке 1.

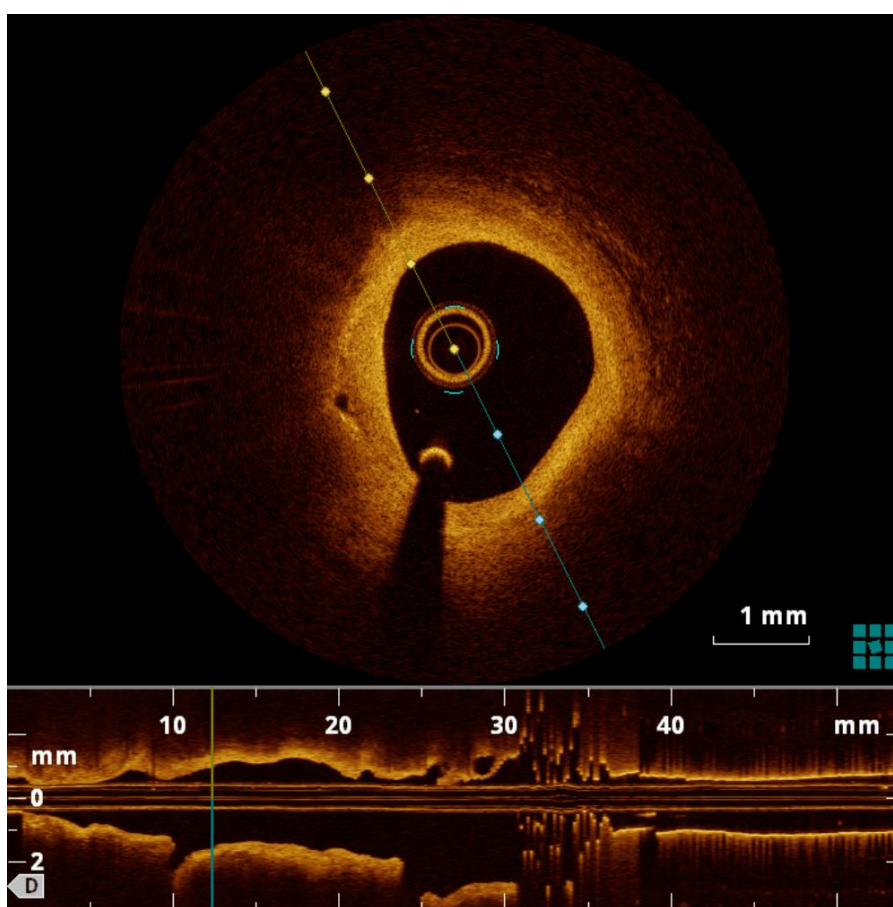


Рисунок 1 – Визуализации данных формата Dicom на примере ОСТ изображения (Кадр видеопоследовательности внутривенного перемещения катетера)

1.2 Сегментация и локализация анатомических структур

Выделение областей интереса на медицинских изображениях играет ключевую роль в клинической практике, повышая точность локализации органов и патологий и, как следствие, качество диагностики. Однако, существует проблема в отсутствии универсального метода сегментации. На конечный результат сильно оказывают влияние различные обстоятельства: исходное качество изображений, уровень и тип шумов, корректность предобработки. Это обуславливает необходимость множества различных методов сегментации и контурирования, каждый из которых эффективен для своего набора задач и позволяет минимизировать искажения и добиться качественного выделения структур [25-26]. Несмотря на то, что компьютерное зрение и обработка изображений – развивающиеся дисциплины, их методы за последние 20 лет стали неотъемлемой частью медицинских исследований. Важно подчеркнуть, что потенциал этой предметной области раскрыт научным сообществом еще не полностью.

Разработкой алгоритмов сегментации сложных объектов исследования, таких как обнаружение Covid-19 активно занимались Umut Özkaya, Şaban Öztürk и Mucahid Barstugan. В своей работе «Classification Using Deep Features Fusion and Ranking Technique [33] авторы добились качества F1score в 96.54% путем анализа CNN объединенных данных. Исследователи провели объединение 150 КТ снимков брюшной области до подачи в искусственную нейронную сеть (ИНС). В данной работе представлен детальный анализ КТ снимков ИНС для обнаружения Covid-19. Авторам удалось добиться точности обнаружения заболевания в 98% с моделью на базе ResNet50 и точностью оценки тяжести заболевания в 87,5%. Для достижения данного результата авторами было использовано 411 снимков КТ с присутствием заболевания Covid-19 и 100 аксиальных КТ-срезов от 60 пациентов с ручной сегментацией в виде изображений JPG. Анализ и выявление признаков заболевания по рентгеновским снимкам является весьма сложной задачей,

требующей участия высококвалифицированного специалиста, что не раз упоминается в подобных работах [29-31].

Зарубежные исследователи в работе «The Liver Tumor Segmentation Benchmark (LiTS)» [37] решают задачу автоматического контурирования печени и метастазов на КТ снимках. Применение метода глубокого обучения с адаптацией к вариациям контраста позволило достичь показателя точности по метрике DICE для печени 0.95, для опухолей 0.67. Предложенный в работе метод является эффективным при вариации яркости КТ-снимков. Также авторы провели сравнение с одним из классических методов сегментации – Region-growing [38]. Показатель точности, для которого на примере сегментации опухолей оказался менее 0.6.

Решения с использованием гибридных сетей также зарекомендовали себя при решении задачи сегментации сложных объектов исследования. Работа «Comparative analysis of modifications of U-Net neuronal network architectures in medical image segmentation» [39] демонстрирует итоговый показатель точности по метрике Dice – 0.92 для сегментации левого желудочка. В сравнении с атласными методами сегментации – atlas-based [40], которые демонстрируют значения лишь менее 0.85. В работе используется Гибридная 2D/3D CNN для анализа коротких и длинных осей сердца. Отметим, что использование мультимодальных данных, на примере 3D-контекста, дало возможность повысить точность сегментацию на 15%, как отмечают авторы работы.

Сегментацией сложных медицинских объектов исследования со стороны отечественных представителей активно занимаются: А.В. Работягов, В.В. Ляшенко, О.А. Кобылин [41]. В своей работе авторы представляют несколько различных подходов к сегментации медицинских изображений с целью выявления цитологических препаратов. Основной акцент в разработанном подходе авторы делают на использовании бионической модели. По результатам эксперимента исследователи сделали следующий вывод: проблема сегментации клеток на сложных изображениях цитологических препаратов существует, что в первую

очередь, связано с прозрачностью клеток по отношению к буферному раствору и наложением клеток друг на друга. К сожалению, решить данную проблему при помощи предложенной бионической модели пока не удалось.

Исследователям международного медицинский центр диагностики и лечения из г. Санкт-Петербург, напротив, удалось достичь высоких показателей эффективности при работе с мультимодальными данными. В своей работе «Нейросетевая модель детекции признаков поражения легких, ассоциированных с COVID-19, на аксиальных срезах нативной компьютерной томографии грудной клетки» [42] исследователи достигли конечного показателя точности по метрике качества IOU 0,82–0,97 с использованием решения на базе сверточной искусственной нейронной сети для сегментации вида энкодер-декодер с декодером Feature Pyramid Network (FPN) [43] и энкодером на основе классификационной нейронной сети EfficientNet-B5. В конечном результате авторы заявляют, что обученная модель может быть использована в качестве эффективного интеллектуального ассистента рентгенолога при работе с КТ-исследованиями пациентов с подозрением на коронавирусную инфекцию.

Важный вклад в работу со сложными объектами исследования в области медицины привнесли исследователи Самарского государственного медицинского университета. В своей работе «Автоматическая сегментация очагов демиелинизации при рассеянном склерозе» [44] авторы используют адаптированный для медицинских данных алгоритм YOLOv8 [45], обученный на размеченных вручную МРТ-снимках. Итоговая точность по метрике F1 зафиксирована показателем 0.61. Достигнутая точность является весомым показателем при работе с данными, сложного характера. Как отмечают авторы исследования: модель эффективно идентифицирует очаги демиелинизации различной локализации, типичной для рассеянного склероза. Но остается необходимость в повышении полноты для минимизации пропуска поражений.

Тестирование на независимых данных также подтвердило стабильность результатов модели.

Стоит отметить вклад исследователей Южного федерального университета, в своей работе «Futuristic Trends in Network and Communication Technology» [46] они представили самые перспективные методы анализа изображений на данный момент.

Примеры медицинских изображений с присутствием сложных объектов исследования представлены на рисунках 2-5.

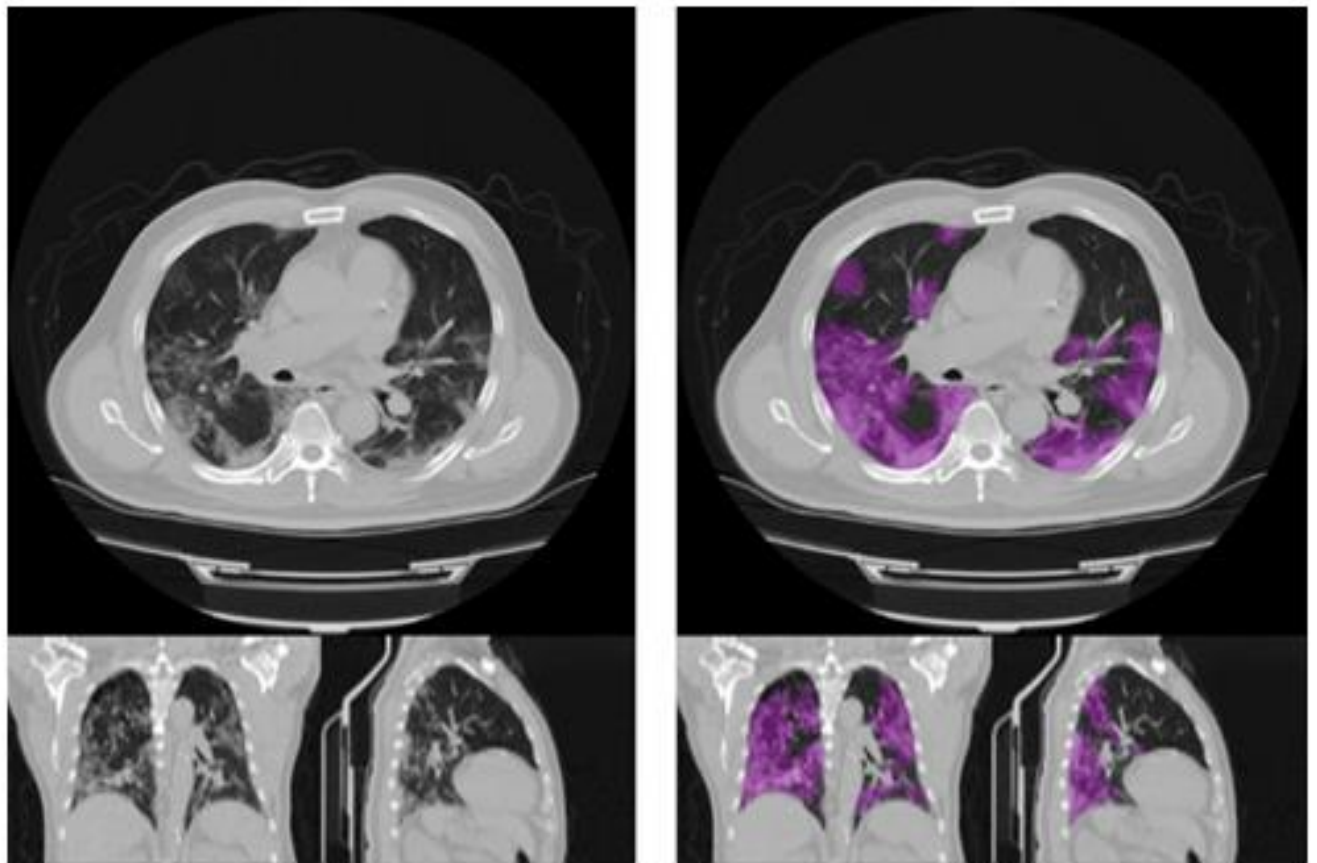


Рисунок 2 – Примеры исходных срезов КТ-изображения и соответствующие им аннотированные данные

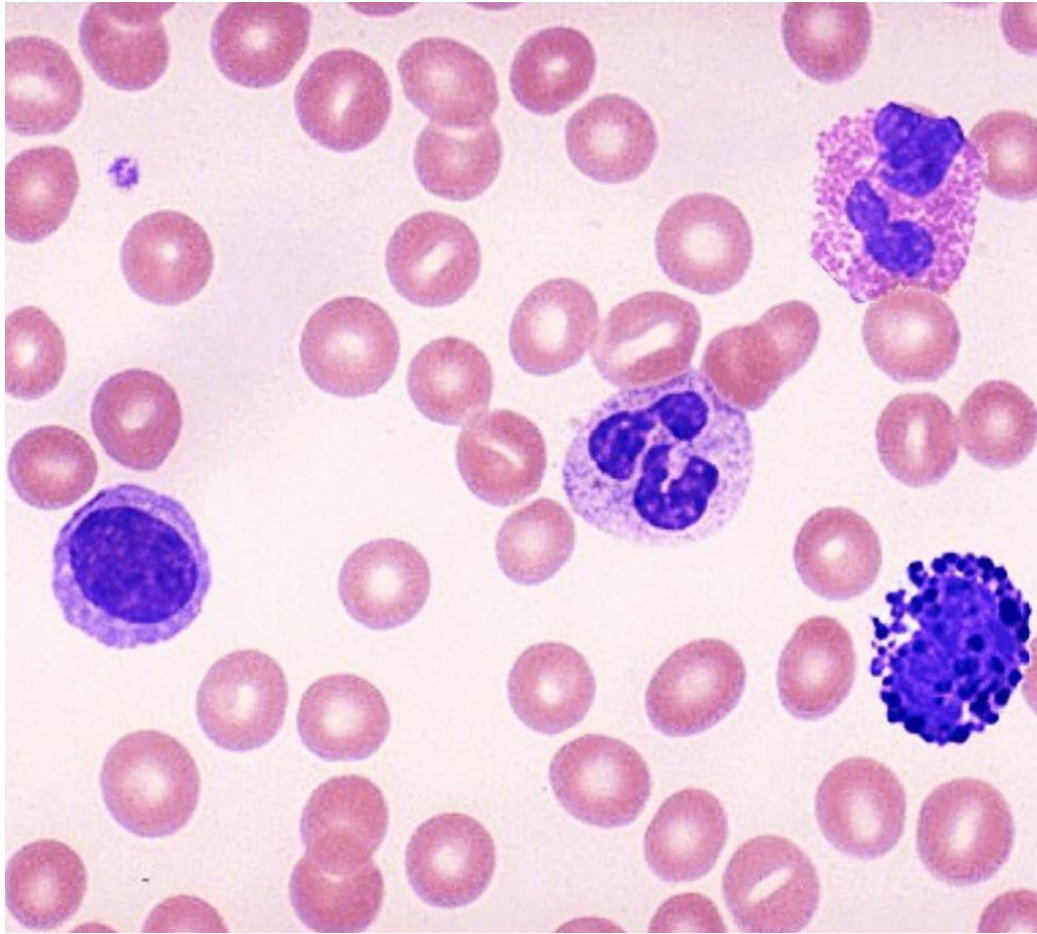


Рисунок 3 – Изображение цитологических препаратов

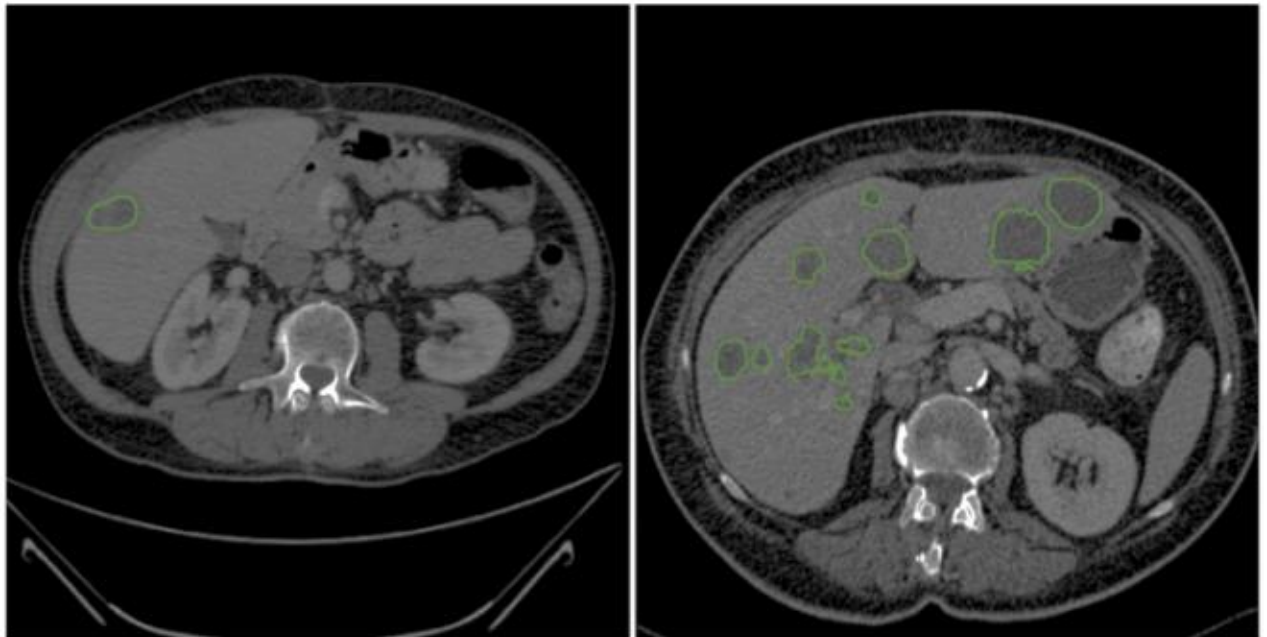


Рисунок 4 – Пример из набора данных LiTS, демонстрирующий разнообразие форм используемых КТ-снимков брюшной полости с контрастным усилением



Рисунок 5 – Гистологический срез ткани овцы

1.3 Подготовка данных

Как было отмечено в главе 1.2 для достижения высокого качества обнаружения сложных объектов исследования на медицинских данных, требуется большое количество обучающих данных. Справедливо отметить, что именно

разметка данных занимает большую часть времени проводимого исследования и разработки.

1.3.1 Вычислительная сложность

В рамках современных исследований в области компьютерного зрения ключевым аспектом, определяющим применимость алгоритмов в реальных условиях, является их вычислительная сложность.

Классификационные модели, такие как ResNet [47], EfficientNet [48] и Vision Transformer [49], ориентированы на присвоение входным данным меток из фиксированного набора классов. Их вычислительная сложность определяется архитектурой сети, объемом обучающих данных и стратегией оптимизации. Например, обучение ResNet-50 [50] на наборе данных ImageNet [51] требует около 90 эпох при использовании 8 GPU NVIDIA V100, что занимает приблизительно 12–15 часов. В работе «Lion: A Efficient Optimization Framework for Deep Learning» [52] демонстрируется, что использование оптимизатора Lion сокращает время обучения на 20% без потери точности. Однако ключевым ограничением остается зависимость от размера обучающего множества. Для достижения высокой обобщающей способности требуется не менее 1 млн размеченных изображений, что создает значительные трудозатраты на этапе подготовки данных.

Сегментационные модели, включая U-Net [53], Mask R-CNN [54] и DeepLabv3+ [55], характеризуются повышенной ресурсоемкостью из-за необходимости обработки пиксельных масок и генерации высокоточных границ. Например, обучение U-Net на наборе данных BraTS [56] занимает до 48 часов на 4 GPU NVIDIA A100. Стоит отметить, что сложность усугубляется при работе с мультимодальными данными: интеграция T1, T2 и FLAIR-режимов MPT увеличивает время обучения на 35–40%. В исследовании «nnU-Net: Self-adapting Framework for Medical Image Segmentation» [57] показано, что применение динамического размера обучающей партии и кэширования данных сокращает

длительность обучения на 25%. однако основным трудозатратным моментом остается подготовка экспертных масок. Ручная разметка одного 3D-скана мозга занимает у радиолога 4–6 часов, а для создания полного набора данных BraTS (1251 случай) требуется более 6000 человеко-часов.

Подготовка данных для обучения моделей компьютерного зрения включает этапы сбора, очистки, разметки и аугментации. Для классификационных задач основная нагрузка приходится на сбор репрезентативных данных и их категоризацию. Например, создание набора данных CIFAR-10 [58] объемом 60 тыс. изображений потребовало 500 человеко-часов на ручную разметку. В случае сегментации трудозатраты возрастают экспоненциально, так разметка одного изображения легких с COVID-19-поражениями занимает 15–20 минут для опытного врача, а для набора данных из 1000 изображений — до 300 часов [59].

Сравнение временных ресурсов для различных типов задач демонстрирует существенный дисбаланс. Например, подготовка данных для классификации гистологических срезов Camelyon16 [60] заняла 200 часов, тогда как сегментация аналогичного объема данных потребовала 1200 часов [61]. Эта разница обусловлена необходимостью пиксельной точности и коррекции артефактов, таких как размытые границы опухолей. В исследовании «Active Learning for Biomedical Image Segmentation» [62] показано, что применение активного обучения сокращает объем ручной разметки на 50% за счет итеративного выбора наиболее информативных образцов.

Процесс аннотирования данных для сегментации изображений является трудоемкой задачей, так как выполнять ее может только эксперт обладающей достаточной квалификацией, а выделять область на изображении приходится в ручном или полуавтоматическом режиме. Стоит отметить, что подготовка данных классификации изображений занимает много меньше времени и требует меньшего количества человеческих ресурсов. В связи с чем на данный момент активно развивается направление обучения искусственных нейронных сетей с

использованием подхода «обучение без учителя» и «использование данных автоматического аннотирования низкого качества».

1.3.2 Типы машинного обучения

Обучение без учителя (unsupervised learning) — это тип машинного обучения, при котором модель ИНС обучается на неразмеченных данных. Под неразмеченными данными понимается набор обучающих данных, которые не проходили процедуру аннотирования. Цель такого обучения состоит в том, чтобы выявить скрытые структуры, закономерности или распределения в данных, которые без явных указаний могли остаться незамеченными. Выделяют следующие основные характеристики обучения без учителя:

- **Отсутствие меток:** в обучении без учителя данные не имеют меток или категорий, что существенно отличает этот метод от обучения с учителем, где модель обучается на основе пар «вход-выход».
- **Обнаружение структуры:** модель анализирует данные и пытается выявить структуры, такие как кластеры или группировки, структурные связи, распределения и зависимости.
- **Анализ данных:** обучение без учителя часто используется для предварительного анализа с целью выявления потенциальных закономерностей или снижения размерности.

Использование данных автоматического аннотирования низкого качества. В основе предложенного в данной работе метода обучения положен алгоритм подобный обучению без учителя. Выбор данного метода обусловлен решением проблемы долгого и трудоемкого аннотирования данных. В данном методе будет использоваться парадигма обучения «вход-выход». Отметим, что принципиальным отличием является подготовка обучающего множества. Вместо подготовки экспертами сегментационных масок изображений будут использоваться маски «низкого качества», полученные в результате анализа

промежуточного результата работы модели классификации. Как отмечалось ранее подготовить данные для обучения классификационной модели не занимает значительное количество времени, для этого достаточно иметь разделение снимков пациентов на две категории: с присутствием объекта анализа и без объекта исследования. Проводя глубокий анализ промежуточного результата работы сети, можно выделить карту активации признаков присутствия класса в определенных областях изображения и использовать данный результат в качестве маски изображения.

В настоящее время существует большое множество алгоритмов, интерпретирующих результат работы моделей ИНС. Эти алгоритмы нацелены на объяснение внутренней работы сложных моделей и обеспечение большей прозрачности в принятии решений нейросетями. Они помогают понять, какие именно признаки данных оказывают наибольшее влияние на предсказания модели и как изменяется выход при варьировании входных данных.

1.3.3 Методы активации картирования классов

В диссертационной работе рассмотрены различные методы активационного картирования классов (CAM), включая GradCAM [63], HiResCAM [64], GradCAMElementWise [65], GradCAM++ [66], XGradCAM [67], EigenCAM [68], EigenGradCAM [69], ScoreCAM [70] и LayerCAM [71]. Перечисленные методы были специально направлены на анализ последнего слоя энкодера рассматриваемых моделей:

- GradCAM: градиентный метод, который определяет важные области изображения для классификации на основе градиентов активации. Данный метод визуализации помогает понять, какие части изображения способствуют принятию решения ИНС. Метод особенно полезен для глубоких нейронных сетей, применяемых в задачах компьютерного зрения.
- HiResCAM: версия GradCAM с высоким разрешением, предназначенная для более точного выделения областей активации. Данный

метод помогает лучше понимать, какие части изображения влияют на решение модели, с более высокой точностью и детализацией.

- GradCAMElementWise: разновидность GradCAM, учитывающая влияние каждого пикселя на итоговое предсказание. Этот метод позволяет более точно выделить значимые области, что способствует лучшему пониманию работы модели.
- GradCAM++: усовершенствованная версия GradCAM, учитывающая положительные и отрицательные градиенты активации, что позволяет получать более детализированные и точные карты активаций.
- XGradCAM: расширенная версия GradCAM, использующая дополнительные функции для повышения точности обнаружения важных областей, что помогает лучше понять, какие части изображения влияют на решение модели
- EigenCAM: метод, основанный на алгоритмах определения собственных значений для определения важных областей изображения. Eigen-CAM использует принципы линейной алгебры для выделения наиболее значимых признаков в активациях сверточных слоев, что позволяет получить более точные и детализированные карты важности.
- EigenGradCAM: разновидность EigenCAM, которая также учитывает градиенты активации для повышения точности определения важных областей. EigenGrad-CAM использует градиенты и анализ главных компонент для выделения наиболее значимых признаков, что позволяет улучшить понимание работы модели.
- LayerCAM: метод, анализирующий активации в различных слоях модели для определения важных особенностей изображения.
- ScoreCAM: метод, анализирующий оценки активаций вместо градиентов для построения карт важности. ScoreCAM устраняет зависимость от градиентов, что делает его более устойчивым.

Данные методы позволяют получить бесценные сведения о внутреннем устройстве нейросетевых моделей, осветив области интереса, выявленные в процессе классификации. Проводимый анализ объяснимости моделей подкрепляет уверенность в надежности и интерпретируемости результатов классификации, а также позволяет сформировать примерную сегментационную маску объекта исследования низкого качества.

1.4 Архитектуры нейронной сети семантической сегментации

На сегодняшний день существует большое множество архитектур искусственных нейронных сетей, для проведения экспериментов были выбраны девять популярных сегментационных нейронных сетей: U-Net [53], Unet++ [72], LinkNet [73], FPN [43], PSPNet [74], PAN [75], DeepLabV3 [76], DeepLabV3+ [25] и MA-Net [77]. Эти модели были выбраны на основе их доказанной эффективности при анализе сложных биомедицинских изображений.

Также помимо архитектуры нейросети на конечный результат и производительность напрямую влияет выбранный энкодер. Энкодер используется для преобразования входных данных в некоторый набор признаков меньшей размерности. В данной работе рассмотрены 9 популярных энкодеров различного объема: Resnet18, Efficient-b0, Timm-regnetx_002, Resnet50, Efficient-b3, Timm-regnetx_008, Resnet101, Efficient-b5, Timm-regnetx_064.

Сеть U-Net, широко применяемая в сегментации биомедицинских изображений, построена на архитектуре энкодера-декодера, который отлично справляется с захватом мелких деталей [53]. Ее способность выделять мелкие структуры, таких как кальцифицированные бляшки и фиброзные пробки, делает ее идеальной для сегментации ОКТ-изображений. Частое использование U-Net в смежных исследованиях еще больше укрепляет ее роль в качестве базовой модели [78], [79]. U-Net++ дополняет ее плотно связанными путями пропуска, улучшая распространение признаков и делает ее особенно эффективной для

сегментирования мелких, тонких структур, таких как микрососуды и тонкие фиброзные оболочки, что очень важно при ОКТ-изображении, где важна точность.

DeerLabV3 использует расширенные свертки для извлечения признаков при различных разрешениях без понижения дискретизации, сохраняя важные детали на ОКТ-изображениях [80]. Это очень важно для точного сегментирования сложных структур, таких как липидные ядра и кальцификаты. DeerLabV3+ улучшает эту модель с помощью усовершенствованного декодера, улучшая сегментацию границ объектов - что важно для выявления мелких особенностей, таких как фиброзные пробки [81]. В аналогичных исследованиях, DeerLabV3+ показала высокую производительность, достигнув чувствительности 85,8% и высокого коэффициента специфичности 0,837 при обнаружении липидных бляшек, а также среднего значения коэффициента Дайса $0,73 \pm 0,10$ для сегментации микрососудов [78].

Сеть пирамид характеристик (Feature Pyramid Network, FPN) отлично справляется с захватом многомасштабных характеристик, что подходит для сегментирования и количественной оценки атеросклеротических бляшек на ОКТ-изображениях [82]. Используя пирамидальные иерархии признаков, FPN улавливает как глобальную, так и мелкомасштабную информацию, повышая точность сегментации в кардиологической визуализации [83]. FPN также продемонстрировала быструю сходимость в аналогичных задачах сегментации, таких как обнаружение особенностей микрососудов в тканеинженерных сосудистых трансплантатах [84].

LinkNet предлагает оптимизированную архитектуру, которая обеспечивает баланс между вычислительной эффективностью и детальным извлечением признаков, что делает его пригодным для применения в режиме реального времени при анализе ОКТ-изображений. Его конструкция позволяет быстро обрабатывать изображения высокого разрешения, сохраняя при этом мелкие детали, необходимые для точного сегментирования сложных структур бляшек. В

конкретных задачах сегментации бляшек LinkNet-34 превзошел другие модели, такие как FPN и U-Net [85].

PSPNet представляет модуль объединения пирамид для агрегирования многомасштабной контекстуальной информации, что делает его высокоэффективным для сегментирования гетерогенных структур в ОКТ, таких как тонкие фиброзные оболочки и крупные кальцификаты [86]. PSPNet также продемонстрировал конкурентоспособную скорость сходимости в предыдущих исследованиях, занимая второе место после FPN [84].

Пирамидальная сеть внимания (PAN) сочетает в себе объединение пирамид с механизмами внимания, захватывая как локальный, так и глобальный контекст. Этот двойной механизм обеспечивает точную сегментацию мелких, критически важных объектов, таких как микрососуды и фиброзные капсулы, что очень важно при анализе ОКТ-изображений [87].

MA-Net (Multi-Attention Network) использует многомасштабные механизмы внимания, что позволяет ей улавливать мелкие детали, сохраняя при этом широкое контекстное понимание. Это делает MA-Net особенно эффективной для сегментирования сложных сердечно-сосудистых структур ОКТ изображений. Ее успешное применение для сегментации изображений IVUS еще раз подтверждает ее пригодность для решения этой задачи данной задачи [88].

Выбор этих девяти моделей представляет собой комплексный подход к решению специфических задач сегментации ОКТ-изображений. Каждая модель обладает определенными сильными сторонами в обработке локальных и глобальных особенностей, сохранении разрешения изображения и обработке сложных анатомических структур. Такое разнообразие архитектур обеспечивает прочную основу для проведения сравнения и оптимизации, обеспечивая выбор наиболее подходящей модели для будущей клинической интеграции.

1.5 Математическая формулировка задачи семантической сегментации

Процесс сегментации изображения состоит в детальной попиксельной (вокселевой) классификации и декомпозиции множества пикселей (вокселей) на специфические, связанные области. Как правило, сегментация осуществляется на основе яркостных характеристик для монохромных изображений и цветовых компонентов для полноцветных. В конечном итоге, цель сегментации состоит в том, чтобы разбить все поле N на области объектов $N_1, N_2, \dots, N_S$ и область фона N_B . В математическом виде задачу сегментации можно представить следующим образом:

$$\begin{cases} f[x, y, p(x, y)] \rightarrow f(n, p'^{(n)}), \\ f(N^k) \rightarrow N^l, \\ l < k \end{cases} \quad (1.1)$$

где:

- x, y – координаты исходного изображения;
- $p(x, y)$ – интенсивность пикселя;
- n – регион интереса;
- $p'^{(n)}$ – признак или вектор признаков сегментируемой области;
- k – число уникальных значений интенсивности в исходном изображении;
- l – количество сегментов после обработки.

1.6 Программные продукты для обработки медицинских данных

Актуальные решения для обработки медицинских данных охватывают разнообразные программные продукты, различающиеся по функциональным возможностям, специализации и степени комплексности. Можно выделить три ключевых стратегии реализации подобных программных систем:

Первая стратегия предполагает встраивание систем обработки медицинских изображений непосредственно в аппаратуру, с которой они впоследствии поставляются.

Вторая методология подразумевает разработку и внедрение таких систем в качестве автономного программного обеспечения.

Третья, более многообещающая и востребованная – системы, допускающие независимую от аппаратных средств установку, поскольку это обеспечивает возможность обработки изображений вне зависимости от применяемого оборудования. Стоит отметить, что подобные программные решения часто создаются для различных операционных платформ и устройств.

В настоящее время особое значение придается кроссплатформенности и портативности программного обеспечения. В открытом доступе имеется обширное количество программных продуктов, совместимых с операционными системами Windows, Linux и macOS. С целью обеспечения портируемости такие приложения часто разрабатываются под платформы Java или .NET. Из чего следует, что одним из критически важных аспектов является гарантирование совместимости систем. В свою очередь, диверсификация и постоянное совершенствование аппаратных средств медицинской диагностики принуждают производителей аппаратного и программного обеспечения использовать универсальные стандарты и протоколы информационного обмена. К популярным программным решениям по обработке и визуализации медицинских данных можно отнести:

- QLab от Philips (специализированный программный пакет для количественного анализа сердца и сосудов) [89].
- Syngo от Siemens (современный PACS и платформа для анализа медицинских изображений) [90].
- Innova EPVision 2.0 от General Electric (система трехмерной визуализации) [91].

- TRENION 3D HD от Zeiss (система трехмерной визуализации) [92].

Третий подход подразумевает создание системы удаленного доступа. Данный подход включает в себя все преимущества описанного ранее подхода, а также обладает рядом следующих преимуществ: все вычисления могут производиться централизованно на выделенном серверном оборудовании, клиентская часть, обладающая всеми свойствами кроссплатформенности, может быть запущена на множестве устройств одновременно, и не требует больших вычислительных ресурсов. Обновление вычислительной части также выполняется централизованно и не затрагивает клиентское оборудование. К популярным программным решениям по обработке и визуализации медицинских данных можно отнести:

- Teladoc Health (платформа для удаленных консультаций, интегрированная с системами диагностики) [93]
- Google Cloud Healthcare API (Облачная платформа для обработки медицинских данных, включая EHR, DICOM-изображения и геномные данные) [94]
- PACS (стандартная система для хранения, передачи и обработки медицинских изображений (КТ, МРТ, рентген)) [95]
- Flywheel Imaging Platform (Облачная платформа для управления и анализа медицинских изображений в исследованиях. Автоматическая сегментация МРТ-снимков с помощью предварительно обученных нейросетевых моделей) [96]

На основании изложенного материала становится ясно, что на сегодняшний день существует большое разнообразие техник для реализации сегментации анатомических компонентов и медицинских устройств. Однако большинство подобных методологий главным образом остаются в рамках научно-исследовательской деятельности и нечасто получают прикладное внедрение.

Важно также выделить многочисленные ограничения создаваемых подходов, среди которых их комплексность, требования к тщательной настройке и неполная валидация на реальных массивах данных. Несмотря на действующие разработки и программные системы, остается актуальной необходимость в прогрессе сферы обработки и визуализации медицинской информации с применением платформенно независимых и универсальных методов и концепций.

1.7 Выводы по первой главе

Исследование научных достижений в сфере решения проблем, связанных с результативной и высокоскоростной сегментацией целевых объектов в медицинской сфере, демонстрирует недостаточную проработанность данной тематики в российских и международных научных работах, несмотря на значительное количество публикаций. Ограниченный интерес научного сообщества к обработке и анализу медицинских визуальных данных обусловлен рядом факторов, включающих сложность типологии данных, зависимость файловой структуры от применяемого оборудования, значительные объемы информации, повышенные требования к сегментационной точности и временные издержки обработки. Особенно важно отметить, что такие диагностические техники, как эхокардиография, предполагают еще более жесткие критерии к распознавательным и аналитическим процедурам. Кроме того, медико-диагностические исследования часто сталкиваются с различными препятствиями, начиная с цифровых помех и произвольных движений пациентов (например, при дыхательных актах) и завершая нестандартными методиками, используемыми специалистами.

Анализ научной литературы в области сегментационных методов продемонстрировал, что стратегия, основанная на глубоком обучении, имеет наивысшую эффективность. К числу преимуществ нейросетевых архитектур относится их потенциал решения задач в условиях неопределенности взаимосвязи

между входными и выходными параметрами. Искусственные нейронные сети проявляют резистентность к шумовым воздействиям и обладают способностью автономной фильтрации менее существенных характеристик, концентрируясь на значимых признаках. В усложненных условиях точность нейросетевых моделей имеет незначительное снижение, что свидетельствует об их стабильности. К примеру, проблемы с получением данных возникают при нарушении связей с нейронами. Тем не менее, стоит отметить, что благодаря распределенному принципу информационного хранения и обширному количеству нейронных узлов, лишь существенная потеря соединений способна критически повлиять на конечную производительность. Также следует отметить, что параллельная обработка данных обеспечивает нейронным сетям сохранение высокой вычислительной производительности.

Анализ программно-аппаратных средств выявил, что большинство реконструктивных операций в настоящее время выполняются без применения навигации медицинских устройств. Они технологически сложны, дорогостоящи и имеют узкую специализацию. Также стоит отметить, что ряд производителей оборудования и средств медицинской диагностики прекратил поддержку программного обеспечения на территории Российской Федерации.

Глава 2. Метод и алгоритмы сегментации сложных объектов медицинских изображений с отсутствием явных характеристических признаков

В данной главе диссертационного исследования представлено решение задачи сегментации анатомических структур при проведении процедуры внутрисосудистой визуализации, осложненной ограниченной доступностью аннотированных данных и высокой сложностью ручной разметки. Предложена технология генерации сегментационных масок, основанная на интеграции метода анализа активационных карт признаков, полученных из классификационной модели, с последующей генерацией trimap-карт, выделяющих зоны неопределенности границ объектов. Для уточнения контуров использован механизм постобработки на основе FBA-Net модели, обеспечивающей адаптивное выравнивание границ. С целью оптимизации нейросетевой архитектуры представлена модификация алгоритма ВОНВ.

Целью диссертационной работы является разработка эффективной технологии подготовки сегментационных масок медицинских изображений, основанной на алгоритмах анализа карт признаков с последующей оптимизацией границ объектов с использованием FBA-Net модели, что позволяет повысить точность сегментации и сократить затраты на подготовку обучающих данных в условиях ограниченной аннотированной информации.

2.1 Исходные данные и оборудование

Исходные данные получены с двух устройств: от St. Jude Medical и LightLab Imaging. Данные были получены во время диагностических процедур в НИИ комплексных проблем сердечно-сосудистых заболеваний (Кемерово) и Тюменском кардиологическом научном центре (Тюмень). Общий объем данных составил 103 исследования оптической когерентной томографии (ОКТ): 77 снимок пациентов

мужского пола и 26 снимков пациентов женского пола. Каждое ОКТ исследование представляет собой трехканальный видеофайл с разрешением кадра 704x704 пикселя. Глубина обзора каждого ОКТ исследования варьируется от 215 до 270 кадров.

Для обеспечения клинической валидности исследования важно охарактеризовать не только технические параметры данных, но и состав исследуемой когорты. Когорта исследования состояла из 103 пациентов со стабильной ишемической болезнью сердца, средний возраст которых составил 69 лет (диапазон: 43-83 года). Значительная часть пациентов имела сопутствующие заболевания, в том числе 22 (21,4%) пациента с диабетом и перенесенным в анамнезе инфаркта миокарда (МИ). Клинические проявления были различными, при этом у большинства пациентов (94, 91,3%) наблюдалась стенокардия, классифицированная как функциональные классы 1-3, в то время как меньшая группа (9, 8,7%) демонстрировала незначительную ишемию. Эти разнообразные характеристики пациентов, представленные в таблице 1, повышают клиническую значимость набора данных и обеспечивают комплексную основу для изучения вариабельности атеросклеротических бляшек.

Таблица 1 – Исходные характеристики пациентов, включенных в исследование

Параметр	Значение
Пол:	
Мужской, n (%)	77 (74.7)
Женский, n (%)	26 (25.3)
Средний возраст, года [min – max]	69.0 [43.0–83.0]
Артериальная гипертензия, n (%)	92 (89.3)
Сахарный диабет, n (%)	22 (21.4)
Инфаркт миокарда, n (%)	22 (21.4)
Поливакулярные заболевания, n (%)	29 (28.2)
Стенокардия:	
Немая ишемия, n (%)	9 (8.7)
Функциональный класс 1, n (%)	24 (23.3)
Функциональный класс 2, n (%)	55 (53.4)
Функциональный класс 3, n (%)	15 (14.6)

Полученные 103 ОКТ-исследования, представленные 25 698 RGB-срезами, имеют четыре морфологических признака бляшки связанных с различными моделями развития атеросклеротической бляшки: сосудистый просвет, фиброзная капсула, липидное ядро и vasa vasorum [97]. Эти характеристики имеют решающее значение для понимания состава бляшки и ее клинических последствий, в частности для оценки риска сердечно-сосудистых событий. Более подробная информация об анализируемых морфологических характеристиках атеросклеротических бляшек представлена ниже [98]:

- Просвет (Lumen LM): Этот параметр просвета сосуда имеет решающее значение для оценки степени стеноза и кровотока, что крайне важно для

диагностики состояния пациента. Объект занимает центральное положение на изображении и имеет ярко выраженные очертания.

- Фиброзная капсула (Fibrous cap, FC): оценивается наличие и толщина фиброзной шапочки - участка бляшки с повышенной плотностью соединительной ткани. Толщина фиброзной шапочки связана со стабильностью бляшки, что делает ее анализ важным для прогнозирования риска разрыва. Данный объект не имеет ярко выраженных признаков и очертаний границ, при этом может достигать внушительных размеров, что также осложняет проводимый анализ.
- Липидное ядро (Lipid core, LC): Выявление и определение характеристик липидного ядра бляшки - области с высоким содержанием жировых отложений. Липидное ядро - ключевой показатель для оценки степени атеросклеротической активности и риска разрыва бляшки. Данный объект не имеет ярко выраженных признаков и очертаний границ подобно фиброзной шапочке. Размеры липидного ядра сравнительно невелики, а сам объект как правило имеет частичное перекрытие с фиброзной оболочкой.
- Vasa vasorum (Vasa vasorum, VV): Выявление кровеносных сосудов, пронизывающих стенку артерии и снабжающих ее питательными веществами. Наличие и распределение vasa vasorum связано с воспалительными процессами и может служить индикатором для оценки степени воспаления бляшки. Vasa vasorum встречается на анализируемых данных крайне редко при этом значения размера объекта принимает микроскопическое значение, что усложняет процесс обнаружения.

Все выявленные морфологических характеристиках атеросклеротических бляшек играют важную роль в понимании структуры и прогрессирования атеросклеротического заболевания. Каждая характеристика дает критическое представление о различных аспекты состава бляшки, влияя на клиническую оценку стабильности бляшки и риск сердечно-сосудистых событий. Сосудистый просвет

отражает степень стеноза, влияя на кровоток, в то время как фиброзная покрышка, липидное ядро и *vasa vasorum* являются важными маркерами для оценки уязвимости бляшки.

Предобработка и разметка данных осуществлялась на компьютере со следующими характеристиками: Intel Core i7-4820K CPU 3.7GHz, Nvidia GeForce 960 GT 2, 16 GB RAM и Windows 10. Обучение нейронных сетей осуществлялось на компьютере со следующими характеристиками: Intel Core i7-6850K CPU 3.6GHz, Nvidia TITAN X и Nvidia TITAN XP, 32 GB RAM и Linux 14.

2.2 Технология генерации сегментационной маски изображения

Технология представляет собой комплекс метода и алгоритмов, включающий анализ карт активации сверточных нейронных сетей, формирование вспомогательных трехканальных карт и уточнение границ объектов с использованием специализированной модели FBA-Net. Информационная передача между этапами, обеспечивает непрерывность обработки данных, минимизацию потерь информации и повышение точности сегментации в условиях ограниченных аннотированных данных. Разработанная технология представляет собой последовательный проход данных без обратных связей, где каждый этап преобразует входные данные в выходные, передавая их на следующий этап в форме тензоров или матриц изображений. Это обеспечивает минимизацию потерь информации, согласованность обработки и отсутствие циклических зависимостей, что упрощает реализацию и повышает воспроизводимость. Общий поток данных можно описать как цепочку преобразований: от исходного изображения к финальной сегментационной маске. Учитывая вышеизложенное, предложенная технология направлена на минимизацию ошибок сегментации в областях с неоднозначными границами, характерными для биомедицинских данных, таких как границы опухолей, сосудов или тканевых структур. Концепция предлагаемой технологии представлена в виде схемы на рисунке 1 и включает 3 основных этапа.

1. Анализ карт активации промежуточных слоев сверточных нейронных сетей с целью создания маски изображения:

- Формирование карты активации признаков на основе пространственных характеристик, полученных с последних сверточных слоев классификационной модели;
- Преобразование карты признаков в исходную маску изображения с выделением предполагаемой области объекта анализа.

Вход: исходное изображение

Выход: Нормализованная бинарная маска объекта.

Передача данных на второй этап осуществляется напрямую, в формате тензора без дополнительных преобразований. Преобразование – детерминированное, основанное на фиксированных весах модели ИНС.

2. Создание трехканальной сегментационной карты:

- Интерполяция исходной маски изображения до заданного разрешения;
- Фильтрация исходной маски объекта при помощи морфологических операций для устранения шумов;
- Формирования области объекта для выделения области с высокой достоверностью локализации целевого объекта выполняется морфологическая эрозия исходной маски.
- Формирование области принадлежности фона. Аналогично предыдущему этапу формируется область гарантированного фона путем применения операции дилатации к исходной маске с последующим инвертированием.
- Формирование зоны неопределенности. Зона неопределенности определяется как разность между исходной маской и надежными зонами.

Вход: бинарная маска объекта, полученная на 1-ом этапе.

Выход: Трехканальная карта объекта.

Передача данных на 3-ий этап прямая – ввод в модель FBA-Net без промежуточных изменений. Прямой, последовательный поток, преобразования – морфологические, с матрицами переходов на основе ядер, обеспечивающими локальные операции без глобальных обратных связей.

3. Уточнение границ сегментации:

- Анализ сформированной трехканальной маски изображения с использованием FBA-Net — специализированной архитектуры, ориентированной на детализированное восстановление контуров объектов и устранение ошибок на граничных участках.

Вход: трехканальная карта, полученная на 2-ом этапе.

Выход: финальная сегментационная маска объекта.

Характер потока прямой, без обратных связей между этапами. Выход 3-его этапа является конечным результатом работы технологии, готовым для использования в обучении сегментационных моделей.

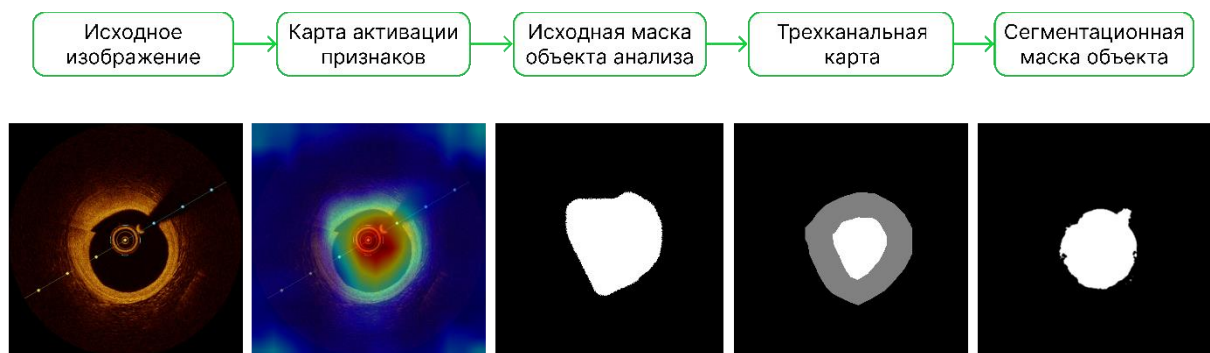


Рисунок 6 – Этапы технологии формирования сегментационной маски изображения

Предложенная технология обеспечивает автоматизированное формирование обучающих сегментационных масок изображений в условиях ограниченного объема аннотированных данных, сокращая трудоемкость подготовки обучающих выборок по сравнению с ручным аннотированием и повышая точность выделения границ сложных анатомических структур.

2.3 Метод формирования исходной маски изображения

Разработка алгоритмов точной сегментации объектов в медицинских изображениях требует интеграции методов глубокого обучения, адаптированных к особенностям низкоконтрастных и зашумленных данных. Предлагаемый метод направлен на автоматизированное формирование области присутствия объектов исследования, таких как атеросклеротические бляшки на снимках оптической когерентной томографии, и включает два взаимосвязанных алгоритма: алгоритм формирования карты активации признаков и алгоритм формирования бинарной маски объекта исследования. Данный метод реализован на начальном этапе разработанной технологии сегментации и обеспечивает первичную локализацию сложных анатомических структур с низкой контрастностью и отсутствием четких границ. Метод интегрирует информацию, извлеченную из промежуточных слоев сверточной нейронной сети, для создания маски, которая передается на последующие этапы технологии.

2.3.1 Алгоритм формирования карты активации признаков

Карта признаков. Классификационная задача представляет собой процесс получения категориального результата на основании совокупности признаков, к примеру, установление присутствия кота на фотографии или идентификация человеческого лица на изображении, диагностика онкологического заболевания у пациента и прочие случаи. Принципиальная схема функционирования классификационной модели изображений на примере CNN представлена на рисунке 7.

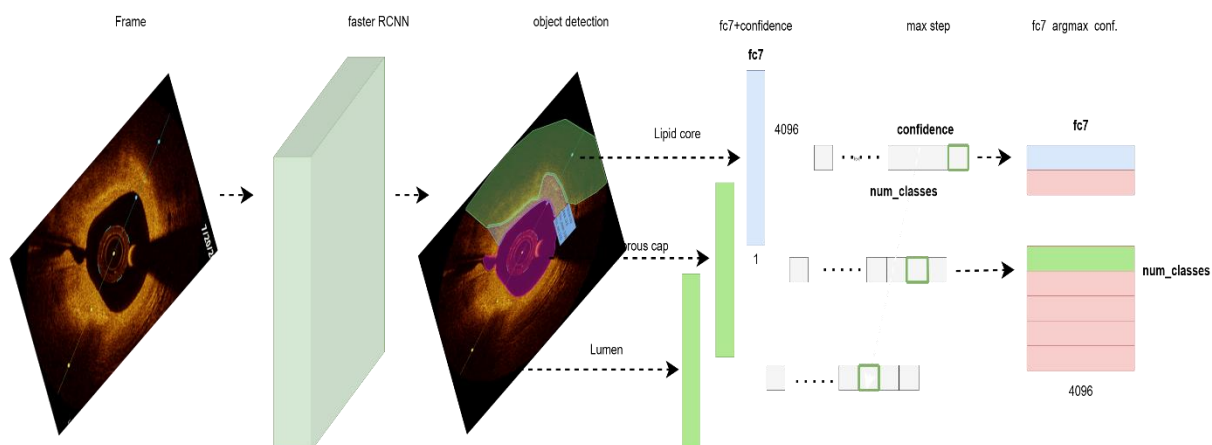


Рисунок 7 – Концептуальная схема работы нейросетевой модели классификации

Для получения категориального ответа на выходе, входное изображение проходит большой путь «сверток» с целью обобщения носимой информации. На каждом этапе свертки неинформативные характеристики изображения отбрасываются, а ключевые характеристики изображения (признаки) сохраняются. Непосредственное определение класса изображения происходит как правило на основе не более 1024 признаков.

В данной работе раскрывается концепция анализа тех самых ключевых признаков, отобранных моделью искусственной нейронной сети, на основе которых был произведен финальный ответ. Для демонстрации результат и выполнения расчета были выбраны изображения кадров видеопоследовательности фиксации прохождения хирургического катетера по артериальной вене человека.

Основная цель данного алгоритма — автоматизировать процесс первичной разметки сложных анатомических структур, обладающих низкой контрастностью и отсутствием выраженных границ. Предложенный алгоритм основан на интерпретации пространственной информации, извлеченной из промежуточных слоев сверточной искусственной нейронной сети и проекции значимых признаков на исходное изображение, где под значимыми признаками понимаются активации нейронов, соответствующие наиболее информативным областям изображения, определяемым на основе их вклада в итоговую классификацию, что дает

возможность получить предварительную маску области интереса без ручного выделения контура объекта. Алгоритм состоит из следующих этапов:

1. Проекция весов классификатора: веса выходного полносвязного слоя, соответствующие индексу предсказанного класса (определяемому через argmax), извлекаются и проецируются обратно на карту активаций последнего сверточного слоя, предшествующего Global Average Pooling. Реализация выполняется через матричное умножение:

$$M = A \times w_c, \quad (2.1)$$

$$A \in R^{(H_c \times W_c \times K)}, \quad (2.2)$$

где:

- A – трехмерный тензор карт активаций;
- H_c, W_c – высота и ширина карт признаков;
- K – количество каналов;
- w_c — веса целевого класса, что позволяет выделить каналы, наиболее значимые для классификации.

Графическая интерпретация данной матричной операции представлена на рисунке 8, где наглядно продемонстрирован процесс взвешенного суммирования карт активаций.

2. Реконструкция пространственной значимости: полученная двумерная карта M размерностью (H_c, W_c) , соответствующей разрешению сверточного слоя, подвергается билинейной интерполяции для масштабирования до исходного разрешения изображения (H, W) . Данная операция восстанавливает пространственное соответствие между абстрактными признаками и пиксельными координатами.

3. Синтез визуализации: нормализованная карта значимости $M_{\text{норм}} \in [0,1]$ накладывается на исходное изображение I с использованием альфа-смешивания:

$$R = \alpha \times I + (1 - \alpha) \times M_{\text{окраш}}, \quad (2.3)$$

где:

- $M_{\text{окраш}}$ — цветовое кодирование карты (например, Jet-палитра);
- α — коэффициент прозрачности, регулирующий баланс между визуализацией активаций и анатомическим контекстом.

Далее приведен псевдокод алгоритма формирования карты активации признаков.

Вход:

I – входное изображение, N – предварительно обученная классификационная CNN-модель, L – целевой сверточный слой для извлечения карт активаций.

Выход: $M_{\text{норм}}$ – нормализованная карта активации.

1. Загрузить изображение I и выполнить предварительную обработку (нормализация, масштабирование).
2. Передать I в модель N в режиме вычисления (инференс модели).
3. Определить индекс предсказанного класса C (индекс максимального значения).
4. Извлечь тензор карт активации $A \in R^{(H_c \times W_c \times K)}$ из слоя L .
5. Извлечь веса $w \in R^K$, соответствующие классу C , из полносвязного слоя.
6. Вычислить карту активации $M = \sum_k A_k \times w_k^c$, где $M \in R^{(H_c \times W_c)}$.
7. Выполнить билинейную интерполяцию карты M до разрешения (H, W) , получив $M_{\text{норм}}$.

8. Нормализовать $M_{\text{норм}}$ в диапазоне $[0,1]$.
9. Выполнить наложение $M_{\text{норм}}$ на исходное изображение I с использованием альфа-смешивания по формуле (2.3), получив визуализированную карту $I_{\text{виз}}$.

Алгоритм обеспечивает интерпретацию решений модели, визуализируя области изображения, максимизирующие активацию для предсказанного класса. Схематическое представление работы алгоритма приведено на рисунке 9. Пример практического применения алгоритма к реальным медицинским данным представлен на рисунке 10, где визуализирована карта активации для класса «сосудистый просвет» (Lumen) на кадре видеопоследовательности прохождения катетера. Интенсивность цвета (от синего к красному в Jet-палитре) отражает степень значимости пикселей для классификационного решения модели.

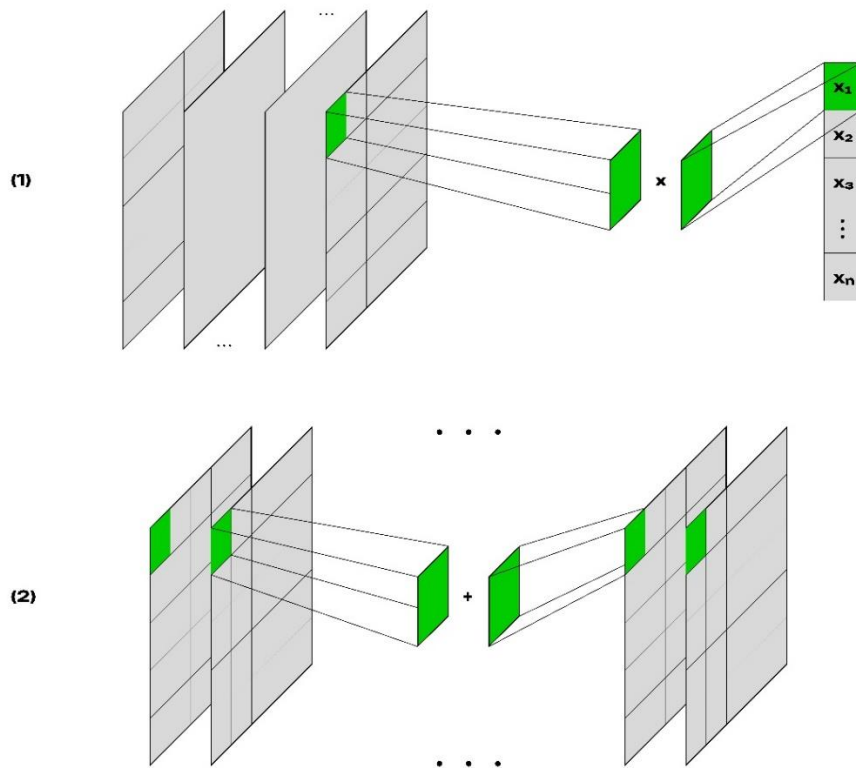


Рисунок 8 – Визуализация матричных операций

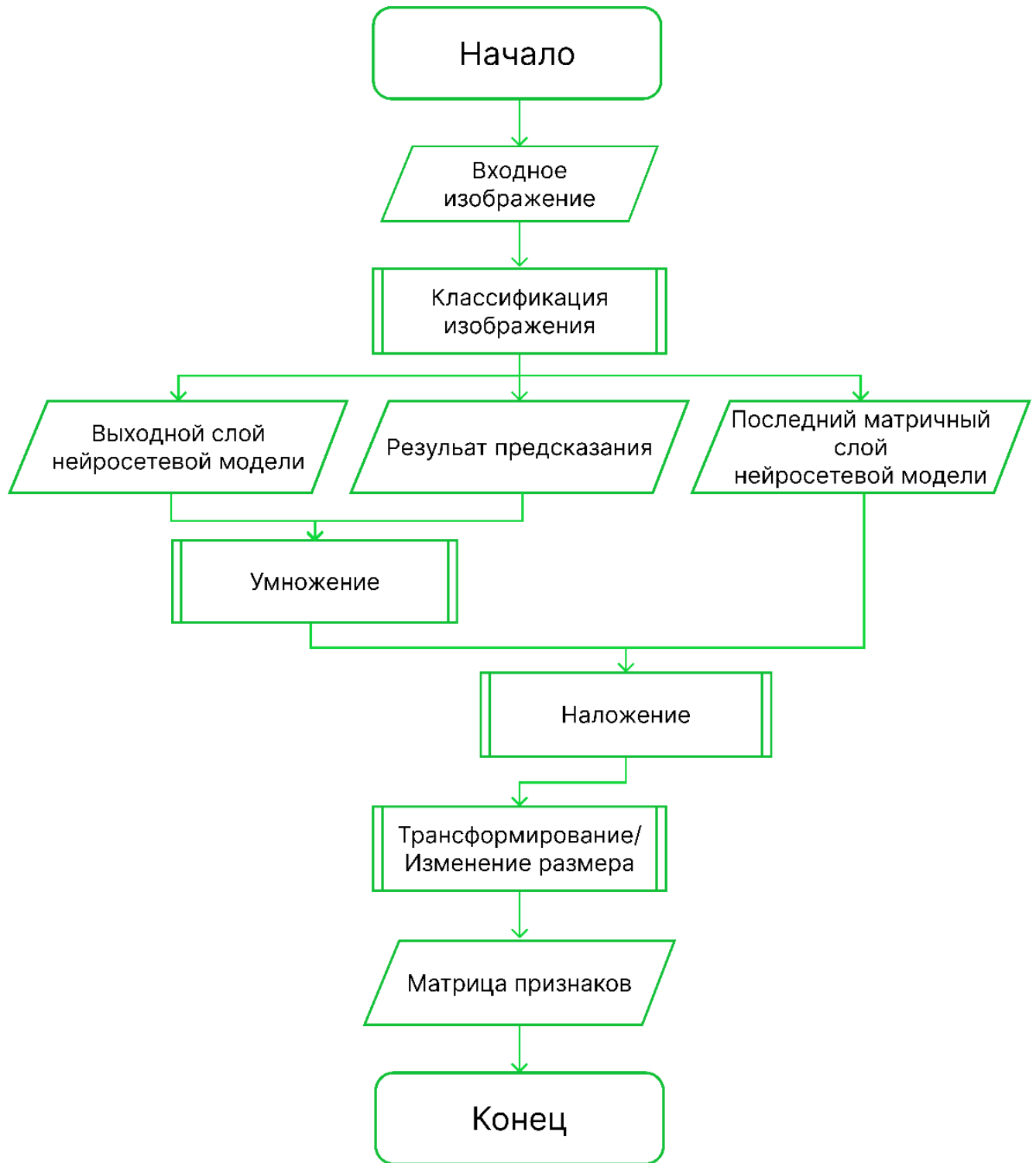


Рисунок 9 – Схематичное представление алгоритма формирования карты признаков



Рисунок 8 – Результат анализа карты признаков классификационной модели

Полученные результаты представляют собой вариант некой «тепловой карты следования за объектом», а также имеет схожее описание с сегментационной маской изображения. Данный результат позволяет выдвинуть гипотезу, что пространственное распределение активаций, взвешенных по значимости для целевого класса, можно интерпретировать как сегментационную маску изображения с обозначением позиции класса анализа.

Для решения поставленной задачи, а именно определение присутствия классов Lumen, Fibrous cap, Lipid core и Vasa vasorum на кадре видеопоследовательности было обучено 4 CNN модели (ResNet50, EfficientNetB5, Xception, ConvNeXtBase). Выбранные модели являются современными архитектурами глубокого обучения, доказавшими свою эффективность в задачах анализа медицинских изображений [99], [100]. Их ключевые преимущества включают способность работать с низкоконтрастными данными, малыми объектами и неоднородными текстурами, характерными для медицинских снимков. Данные модели покрывают широкий спектр медицинских задач: от анализа микроскопических объектов [101] до обработки 3D-данных [102]. Для достижения качественного результата за кратчайшие сроки мы использовали технологию трансферного обучения [103]. Используя модели, обученные на 350 миллионах изображений Imagenet и хранящие веса для определения 17000 классов,

в кратчайшие сроки были подготовлены модели классификации. Перед обучением к каждой из моделей был применен следующий набор преобразований:

1. Веса всей модели замораживаются.
2. Последние 30 слоев модели назначаются тренировочными
3. Последний слой CNN заменяется на GlobalAveragePooling2D
4. Сеть дополняется полносвязным слоем с количеством нейронов равным глубине карты признаков предыдущего слоя
5. Сеть дополняется полносвязным слоем с количеством нейронов равным двум и функцией активации Sigmoid.

Схема архитектуры ИНС представлена на рисунке ниже.

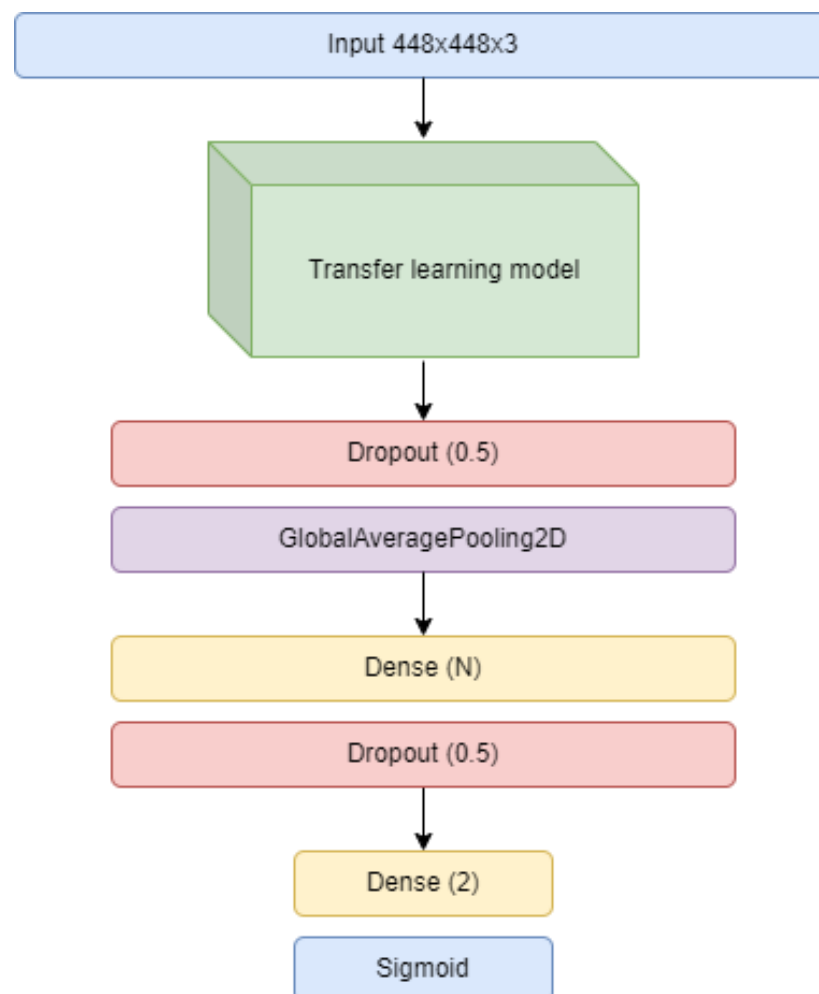


Рисунок 9 – Схема архитектуры ИНС модели классификации

После прохождения процедуры обучения к результирующей версии модели применяются следующие преобразование:

- Формирование второго выхода ИНС: в качестве вывода используется. Пространственные признаки высокого уровня, извлеченные из слоя, предшествующего операции Global Average Pooling, и вероятностное распределение классов, генерируемое финальным классификационным слоем. Параллельный вывод классификационных результатов позволяет установить корреляцию между пространственными активациями и итоговым решением модели, формируя основу для интерпретации вклада регионов изображения в предсказание. Такая архитектура обеспечивает декомпозицию процесса принятия решений, разделяя этапы формирования абстрактных признаков и их статистической интерпретации.
- Активация режима оценки модели (Eval, инициализация перехода нейросетевой модели в режим инференса (оценки), при котором деактивируются механизмы, специфичные для фазы обучения). Данная процедура выполняется единожды с целью эффективного использования ресурсов и времени анализа.

2.3.2 Алгоритм формирования трехканальной карты объекта исследования

Трехканальная карта объекта, или Trimap, представляет собой структурированную пространственную разметку изображения, предназначенную для решения задач сегментации и матирования объектов. В ее основе лежит разделение пикселей изображения на три взаимоисключающие категории, формирующие семантические зоны: детерминированный передний план (foreground), детерминированный фон (background) и неопределенную смешанную область (unknown region). Передний план соответствует пикселям, однозначно принадлежащим целевому объекту, фон – пикселям, не связанным с объектом, а смешанная зона включает пиксели, которые не могут быть однозначно отнесены

ни к одной из категорий из-за частичной прозрачности, сложных текстур или граничных артефактов.

Функционально Trimap выступает в качестве априорной маски, задающей контекст для алгоритмов матирования, которые вычисляют альфа-канал – полупрозрачную маску, количественно определяющую степень принадлежности каждого пикселя к объекту. Смешанная область, являясь буферной зоной, фокусирует вычислительные ресурсы на анализе переходных участков, где традиционные методы сегментации демонстрируют низкую точность из-за неоднозначности цветовых или текстурных характеристик.

Ключевая роль Trimap в компьютерном зрении обусловлена ее способностью снижать сложность задач разделения объектов и фона, особенно в условиях ограниченной контрастности, наличия шумов или динамического освещения. В современных подходах, таких как AlphaGAN [104] или методы, основанные на глубоком обучении, Trimap используется как входной сигнал для обучения нейросетей предсказывать альфа-каналы с субпиксельной точностью. Это позволяет достичь высокой детализации при выделении объектов с нежесткими границами, таких как волосы, дым или полупрозрачные материалы, что критически важно в приложениях видеомонтажа, дополненной реальности и медицинской визуализации.

Таким образом, трехканальная карта объекта служит не только инструментом разметки, но и математическим аппаратом, формализующим пространственные отношения между объектом и фоном, что обеспечивает воспроизводимость и интерпретируемость результатов в задачах анализа изображений.

Процесс составления Trimap включает несколько этапов:

1. Инициализация: первоначально пользователю или алгоритму необходимо задать начальные области для фона и объекта. Это можно сделать вручную, используя инструменты выделения в графических

редакторах, или автоматически, используя предварительные алгоритмы сегментации.

2. Определение границ: границы между объектом и фоном часто определяются с помощью алгоритмов, таких как градиентный анализ или методы машинного обучения. Пиксели, которые находятся рядом с границами, помещаются в неопределенную область.
3. Создание маски Trimap: на основе выделенных областей создается маска, где каждому пикселю присваивается одно из трех значений: 0 (фон), 128 (неопределенная область), 255 (объект).

В последние годы ведутся активные исследования по улучшению методов работы с Trimap:

- Улучшение алгоритмов глубокого обучения: развитие архитектур нейронных сетей, таких как U-Net, Deeplab и Mask R-CNN, значительно улучшило качество выделения объектов. Эти модели могут обучаться на больших наборах данных и учитывать множество параметров, что делает их более точными и универсальными.
- Интерактивные системы: создаются системы, которые позволяют пользователям интерактивно корректировать Trimap, обеспечивая быстрый и удобный процесс. Примеры таких систем включают инструменты в Adobe Photoshop и специализированные программы для видеообработки.
- Адаптивные методы: исследуются методы, которые могут адаптироваться к различным условиям съемки и типам изображений, обеспечивая высокую точность выделения независимо от конкретных особенностей сцены.

В рамках предлагаемой технологии подготовки аннотаций для сегментации изображений важным этапом является формирование трехканальной сегментационной карты — вспомогательной маски, которая разделяет исходное

изображение на три зоны: область с высокой уверенностью принадлежности к целевому объекту, фон и область неопределенности. Цель формирования трехканальной сегментационной маски – локализовать и уточнить границы сложных анатомических структур, в которых объекты часто имеют размытые или слабовыраженные контуры. Алгоритм формирования трехканальной карты состоит из следующих этапов:

Этап 1. Приведение к единому размеру. Полученная маска интерполируется до разрешения 1024×1024 пикселя.

Этап 2. Морфологическая очистка маски. Для устранения шумовых артефактов последовательно применяются морфологические операции открытия и закрытия.

$$B_{\text{фильтр}} = (B_{\text{строка}} \circ S) \bullet S, \quad (2.4)$$

где:

◦ – операция морфологического открытия:

$$B \circ S = (B \odot S) \oplus S, \quad (2.5)$$

• – операция морфологического закрытия:

$$B \bullet S = (B \oplus S) \odot S, \quad (2.6)$$

$\odot$ – морфологическая эрозия, $\oplus$ – морфологическая дилатация, S – структурный элемент.

Этап 3. Формирование области объекта. Выделения области, в которой уверенность присутствия объекта анализа максимальна. Для этого к исходной маске применяется морфологическая операция эрозии.

$$B \odot S = \{z \in E \mid S_z \subseteq B\}, \quad (2.7)$$

где:

- E – множество всех координат изображения;
- S_z – структурный элемент, сдвинутый в позицию z .

Этап 4. Формирование области принадлежности фона. Аналогично предыдущему этапу формируется область гарантированного фона путем применения операции дилатации к исходной маске с последующим инвертированием.

$$B \oplus S = \{x | (S + x) \cap B \neq \emptyset\} \quad (2.8)$$

Этап 5. Формирование зоны неопределенности. Зона неопределенности определяется как разность между исходной маской и надежными зонами. При этом:

$$T(x, y) = \left\{ \begin{array}{l} 1, \text{ если } F(x, y) = 1 (\text{объект}) \\ 0, \text{ если } B(x, y) = 1 (\text{фон}) \\ 0.5, \text{ в остальных случаях (зона неопределенности)} \end{array} \right\} \quad (2.9)$$

Пример формирования Trimar изображения представлен на рисунке ниже:

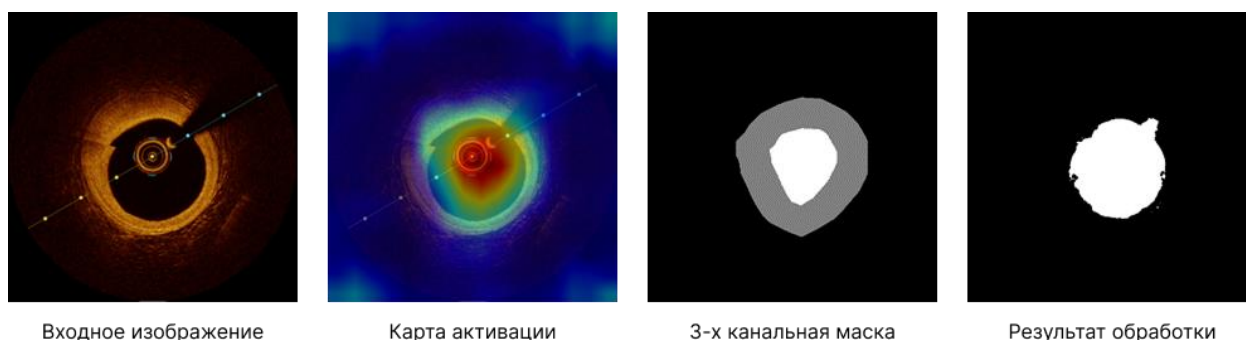


Рисунок 10 – Пример создания Trimar изображения

Полученное Trimar изображение анализируется FBA-Net моделью [106]. FBA-Net (Feature Boundary-Aware Network) – это глубокая нейронная сеть, разработанная для задач сегментации медицинских изображений с акцентом на точное определение границ объектов. Основная цель архитектуры – улучшить качество выделения сложных анатомических структур (опухолей, органов) за счет комбинации механизмов внимания к границам и семантическим признакам. Модель была предложена как решение проблем, связанных с неоднородностью интенсивности пикселей, шумами и размытыми границами в медицинских данных.

FBA-Net сочетает подходы U-Net и attention-механизмы, адаптированные для работы с граничными признаками.

FBA-Net демонстрирует высокую эффективность в задачах, где критична точность определения границ (онкология, кардиология). Ее архитектура, сочетающая анализ признаков и контуров, задает новый стандарт для медицинской сегментации.

Ключевые компоненты [107]:

1. Двухпоточный энкодер:

- Поток признаков: извлекает семантическую информацию (например, текстуру опухоли).
- Поток границ: детектирует границы объектов с помощью сверточных слоев, оптимизированных под задачи edge detection.

2. Boundary-Attention Module (BAM):

- Совмещает карты признаков и границ, усиливая релевантные области, используя формулу:

$$F_{out} = F_{features} * \sigma(F_{boundaries}) + F_{boundaries}, \quad (2.10)$$

где:

- σ – сигмоидная активация.

3. Ассиметричный декодер:

- Восстанавливает карту сегментации, комбинируя многоуровневые признаки энкодера.
- Использует skip-connections с взвешенным суммированием для минимизации потерь информации

Интеграция FBA-Net в методологию диссертационной работы обеспечивает существенное повышение точности автоматически формируемых сегментационных масок объектов. Способность модели выделять границы

объектов в условиях зашумленности и низкой контрастности напрямую коррелирует с целями исследования.

2.4 Алгоритм аугментации данных

Среди востребованных подходов при ограниченности данных выделяется методология аугментации. В рамках применения процесса аугментации данных осуществляется синтетическая генерация новых образцов на базе имеющегося материала. Путем применения разнообразных деформаций и преобразований к оригинальным изображениям исходный набор данных существенно увеличивается, что способствует значительному повышению эффективности модели. К наиболее распространенным видам искажений относятся модификации яркости и цветности, замещение фоновых областей, геометрические (проективные и аффинные) преобразования, добавление световых эффектов, шумовых компонентов, размытия и прочие. Следует подчеркнуть, что данная методология характеризуется относительной экономичностью и минимальными требованиями к человеческим ресурсам, поскольку исключает необходимость аннотирования и сбора дополнительных данных. Это объясняется тем, что весь процесс формирования изображений определяется выбранным алгоритмом и его настройками. Тем не менее, могут возникать затруднения при корректном определении параметров аугментации обучающего множества из исходных экземпляров. С одной стороны, требуется достаточное количество образцов для обучения нейросети распознаванию зашумленных примеров, с другой стороны, необходимо сохранять производительность на других категориях сложных изображений.

В данной работе планируется апробация представленных методов в задачах различного характера. С целью сохранения качества лабораторных условий было принято решение использовать единый унифицированный набор методов аугментации данных. Представленный ниже набор аугментационных

преобразований реализован с помощью библиотеки "*Albumentations*". Эти аугментации не только позволили расширить размер набора данных, но также служили как техника регуляризации, помогая смягчить переобучение во время обучения модели. Предложенный набор аугментаций включает следующие преобразования:

- Горизонтальное отражение с вероятностью 50%.
- Сдвиг, масштабирование и вращение с вероятностью 20%: позволяет производить случайные сдвиги, масштабирование и вращения в пределах заданных лимитов (лимит сдвига = 0.0625, лимит масштаба = 0.1 и лимит вращения = 15).
- Использование случайной меньшей области анализа с вероятностью 20%: применяется случайное обрезание с размерами, определенными как процент от размера входного изображения, в диапазоне от 0.8 до 0.9 от размера входа.
- Условное заполнение. Все изображения дополняются, чтобы обеспечить одинаковый размер для обработки.
- Гауссовский шум с вероятностью 20%: добавляет случайный шум к изображениям с переменным диапазоном интенсивности, где дисперсия варьируется от 3 до 10.
- Искажение перспективы с вероятностью 20%: применяет случайные перспективные преобразования к изображениям с масштабом от 0.05 до 0.1.
- Случайная коррекция яркости и контраста с вероятностью 20%: корректирует яркость и контраст изображений в пределах заданных лимитов (лимит яркости = 0.2, лимит контраста = 0.2).
- Настройка оттенка, насыщенности и значения с вероятностью 20%: сдвигает оттенок, насыщенность и значение изображений в пределах

заданных лимитов (лимит сдвига оттенка = 20, лимит сдвига насыщенности = 30, лимит сдвига значения = 20).

2.5 Метод обучения модели

Ключевыми характеристиками нейросетевой модели выступают ее способность к генерализации и прецизионность прогнозирования. Для достижения целевых значений указанных показателей требуется реализация комплекса специализированных процедур. К числу таких мероприятий относятся селекция оптимальной функции потерь, применение алгоритмов аугментации данных и настройка гиперпараметров. В связи с этим в данном разделе рассматриваются наиболее значимые операции для обеспечения требуемых показателей точности и обобщающей способности модели.

Тренировочный процесс ориентирован на оптимизацию сегментационных метрик с особым фокусом на коэффициенте сходства Дайса (DSC). Для реализации данной задачи применялась DSC-базированная функция потерь, расчет которой осуществляется по следующему принципу:

$$Loss = 1 - \frac{2 \sum (y_{true} \times y_{pred}) + \varepsilon}{\sum y_{true} + \sum y_{pred} + \varepsilon}, \quad (2.11)$$

где:

- y_{true} и y_{pred} представляют собой истинные и предсказанные значения соответственно, ε является константой (в данном исследовании равной 10^{-7}), которая используется для численной стабильности, чтобы избежать ошибок деления на ноль.

В качестве изменяемых параметров обучения в первую очередь выступают архитектура и энкодер искусственной нейронной сети. Также не маловажными параметрами являются метод оптимизации, скорость обучения, размер входного изображения, размер входной партии (батч), метод изменения скорости обучения.

В большинстве исследований для определения оптимальных гиперпараметров применяется метод перекрестной валидации. В настоящей работе все полученные результаты также прошли аналогичную верификацию. В процессе перекрестной валидации оценка моделей преимущественно базируется на метриках точности, в меньшей степени - на значениях функции потерь. Данная процедура предполагает анализ множества модельных вариантов с последующим выбором наиболее эффективного. Схема определения оптимальной модели посредством перекрестной валидации представлена на рисунке 13.

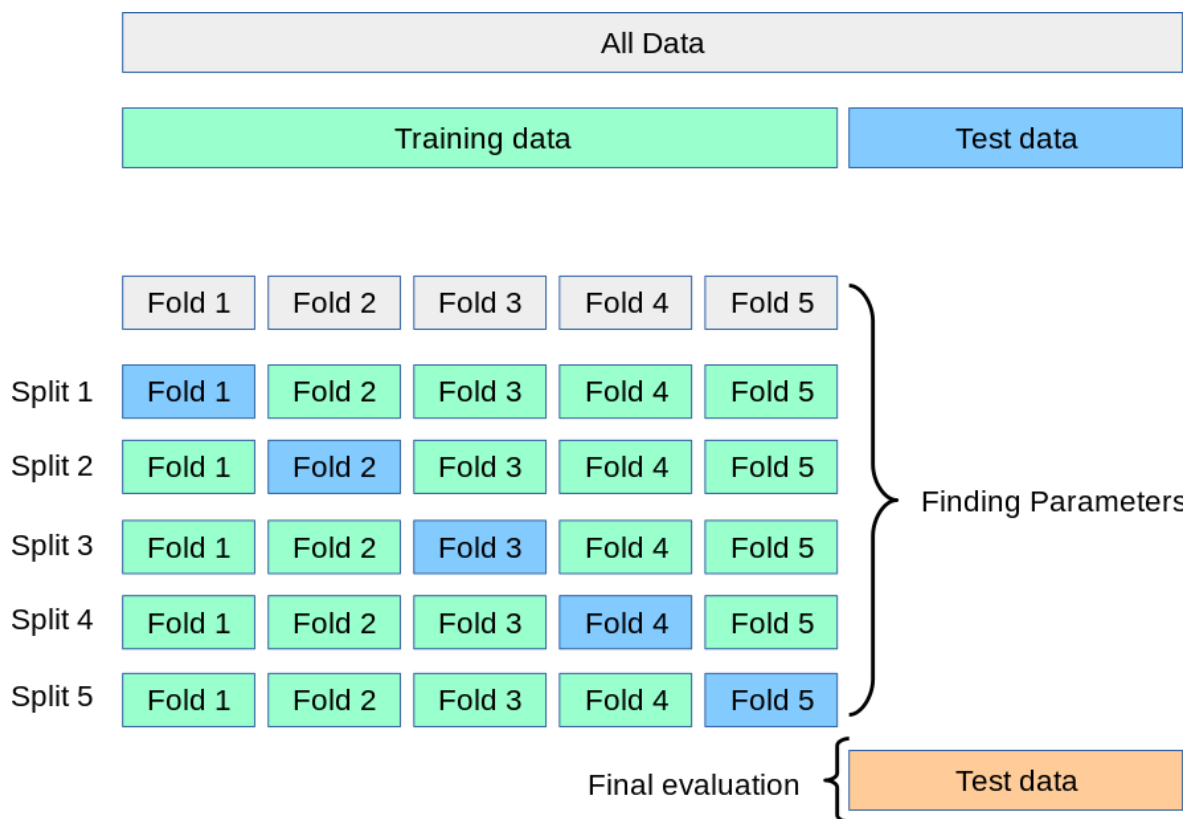


Рисунок 11 – Схема проведения перекрестной проверки

Стоит отметить, что процесс обучения нейросети для решения задачи сегментации может достигать до 14 ч и использовать процедуру полного перебора, а также кросс-валидацию для получения качественного сравнения не является эффективным.

В данной работе рассматривается использование модификации алгоритма гипертюнинга параметров нейросетевой модели на основе метода ВОНВ [108].

Кросс-валидация же в свою очередь будет применена только к тем моделям, что покажут наилучший результат, чтобы закрепить их показатели.

Стоит отметить, что настройка гиперпараметров проводилась на усеченном наборе данных, включающем 40 пациентов, а не на всей когорте из 103 пациентов. Такая корректировка обусловлена длительностью и вычислительной сложностью процесса. Однако исходя из предположения, что данные по 40 пациентам составляют репрезентативную выборку. Следовательно, ожидается, что оптимальные гипернастроенные сети, полученные из этого подмножества, будут иметь сопоставимую производительность при обучении на полном наборе данных, состоящем из 103 пациентов.

Поскольку приоритеты гиперпараметров меняются в процессе настройки, а некоторые гиперпараметры оказывают более значительное влияние на производительность сети, чем другие [109], то реализован подход сосредоточения усилий по настройке гиперпараметров на конкретных аспектах, вместо попыток оптимизировать каждый параметр. В частности, не осуществляется тонкая настройка таких гиперпараметров, как размер партии, тип нелинейности, параметры оптимизатора или размеры ядра свертки. Вместо этого внимание сосредоточено на тех гиперпараметрах, которые были показаны важными в проводимом исследовании [110], а именно: архитектура энкодера, размер входного изображения, выбор оптимизатора и скорость обучения. В таблице 2 приведено полное описание гиперпараметров, исследованных в процессе настройки, а также соответствующие используемые значения.

Таблица 2 – Гиперпараметры, используемые при оптимизации сетей

Гиперпараметр	Значение	Количество
Architecture	U-Net, U-Net++, DeepLabV3, DeepLabV3+, FPN, LinkNet, PSPNet, PAN, MA-Net	9
Encoder	ResNet-18, ResNet-50, ResNet-101, EfficientNet B0, EfficientNet B5, EfficientNet B7, RegNetX-200MF, RegNetX-6.4GF, RegNetY-12GF	9
Input size	512x512 to 896x896 with the step of 128x128 px	4
Optimizer	Adam, RAdam, RMSprop	3
Learning rate	10^{-3} , 10^{-4} , 10^{-5} , 10^{-6}	4

Стратегия поиска гиперпараметров, основывается на байесовском поиске, который, в отличие от случайного или сеточного поиска, принимает обоснованные решения. Байесовская оптимизация использует вероятностную модель, чтобы определить, какие значения использовать, через итерационный процесс тестирования значений на суррогатной функции перед оценкой функции цели. Кроме того, в работе применяется стратегию раннего завершения, HyperBand [111], для остановки плохо работающих конфигураций. В случае досрочного завершения HyperBand останавливает текущую конфигурацию, прежде чем приступить к работе с новым набором значений гиперпараметров. Сочетание байесовской оптимизации и досрочного завершения HyperBand образует так называемый «ВОНВ», подход, который обеспечивает более высокую вычислительную эффективность и устойчивость по сравнению с сеточным поиском, случайным поиском или стандартными реализациями Байеса или HyperBand.

В свою очередь байесовская оптимизация — это итерационный метод, позволяющий оценить оптимум функции, не дифференцируя ее. Кроме того, на каждой итерации метод указывает, в какой следующей точке мы с наибольшей вероятностью улучшим нашу текущую оценку оптимума. Это позволяет

значительно сократить количество вычислений функции, каждое из которых может быть довольно затратным по времени.

Подбор гиперпараметров тоже можно сформулировать в виде задачи, которая может решаться с помощью байесовской оптимизации. Пусть, например, наша функция — отображает зависимость значений валидационных метрик в зависимости от текущего сочетания гиперпараметров. Ее вычисление затратно по времени (нужно натренировать и провалидировать модель), но что является еще более важным, что мы не можем вычислить градиенты этой функции по ее переменным (нашим гиперпараметрам).

Таким образом, в работе предложен метод по улучшению алгоритма ВОНВ. Для наглядности сравнения, ему присвоено название ВОНВ+. Основная идея улучшения — динамическая адаптация на основе промежуточных результатов. В оригинальном ВОНВ исследуемая модель не обновляется в процессе работы алгоритма HyperBand. Это приводит к двум ключевым проблемам:

1. Неэффективное использование информации: результаты раннего завершения (отбракованные конфигурации) не учитываются при выборе новых гиперпараметров.
2. Смещение выборки: Исследуемая модель обучается на нерепрезентативной выборке, так как HyperBand удаляет "плохие" конфигурации досрочно, искажая распределение данных.

Пусть $D = \{(\theta_i, y_i)\}$ — набор данных, где θ_i — гиперпараметры, y_i — метрика. В классическом ВОНВ данные от ранних этапов HyperBand формируют выборку $D_{early} \in D$, которая не используется для обучения модели, что создает смещение:

$$p(\theta|D) \neq p(\theta|D \cup D_{early}) \quad (2.12)$$

где:

- $p(\theta|D)$ – апостериорное распределение модели. Кроме того, отбрасывание конфигураций нарушает репрезентативность D , что снижает качество итоговой модели.

Предлагаемая динамическая адаптация позволяет автоматически корректировать исследуемую модель на основе промежуточных результатов, улучшая исследование пространства гиперпараметров. Динамическая адаптация применяется после каждой итерации HyperBand. Процесс включает три этапа:

1. Кластеризация конфигураций – группировка конфигураций по схожим характеристикам и метрикам производительности, используя алгоритм $k - means$ с модификациями для работы с разнотипными данными. Использование модифицированного $k - means$ с метрикой, учитывающей разнотипные данные (категориальные, непрерывные гиперпараметры). Для смешанных данных применяется метрика Говера:

$$d(\theta_i, \theta_j) = \frac{1}{p} \sum_{k=1}^p \delta_k (\theta_i^{(k)}, \theta_j^{(k)}) \quad (2.13)$$

где:

- δ_k – метрика для k – го параметра.

2. Перерасчет весов моделей – учет вклада каждого кластера в итоговую метрику. Введение параметра: вес кластера $w_c \propto \max_{i \in c} y_i$, для фокусировки поиска на перспективных областях.

3. Перебалансировка выборки – корректировка распределение гиперпараметров для следующей итерации. Использование адаптивного семплирования, методом отжига для баланса между исследованием и эксплуатацией:

$$p(\theta) \leftarrow \alpha p_{model}(\theta) + (1 - \alpha) p_{uniform}(\theta) \quad (2.14)$$

Динамическая адаптация в ВОНВ+ позволяет интеллектуально перераспределять ресурсы между кластерами гиперпараметров, сокращая время оптимизации и повышая точность моделей. Метод особенно эффективен в задачах с большим пространством параметров, таких как настройка архитектур нейросетей для медицинской визуализации.

Предложенный алгоритм ВОНВ+ позиционируется в контексте современных методов оптимизации гиперпараметров, сочетающих стратегии ранней остановки и байесовского поиска. Классические подходы, такие как сеточный и случайный поиск, страдают от экспоненциального роста вычислительной сложности при увеличении размерности пространства параметров, что делает их непригодными для настройки глубоких нейросетей. Более продвинутые методы, включая байесовскую оптимизацию, демонстрируют эффективность в задачах с ограниченным бюджетом, но требуют точной настройки априорных распределений и чувствительны к зашумленности данных.

Гибридные методы, такие как ВОНВ, объединяют преимущества байесовской оптимизации и HyperBand, но сохраняют ключевое ограничение — статичность суррогатной модели, которая не адаптируется к данным, полученным в процессе «отбраковки» конфигураций. В отличие от ВОНВ, алгоритм FABOLAS [112] использует байесовскую оптимизацию с учетом оставшегося количества итераций, динамически корректируя распределение ресурсов между конфигурациями, однако не предусматривает механизмов кластеризации и перебалансировки, что снижает его применимость в задачах с высокой гетерогенностью данных, характерных для медицинской визуализации.

Главное преимущество ВОНВ+ перед аналогами заключается в динамической адаптации, которая устраняет информационные потери за счет интеграции данных от частично оцененных конфигураций. Это отличает его от ВОНВ, где суррогатная модель обучается только на завершенных испытаниях, игнорируя промежуточные результаты.

2.6 Выводы по второй главе

1. В связи с небольшим количеством примеров в исходном наборе данных и сложностью аннотирования была предложена технология синтеза сегментационных масок объектов, в основе которого лежит метод анализа активационных карт признаков, полученных из классификационной модели, с последующей генерацией trimap-карт, выделяющих зоны неопределенности границ объектов.

2. Представлен метод генерации исходной маски объекта исследования на основе анализа активационных карт признаков, полученных из классификационной модели, с целью формирования первичной сегментационной маски изображения.

3. Для повышения качества полученных сегментационных масок и сокращения времени валидации полученных результатов предложен и детально представлен алгоритм создания trimap изображений с дальнейшим применением FBA-Net модели для повышения конечной точности и обеспечивающей адаптивное выравнивание границ.

4. Предложена модификация алгоритма оптимизации гиперпараметров ВОНВ+ на основе алгоритма ВОНВ. Отличительной особенностью алгоритма ВОНВ+ и ВОНВ является использование данных, полученных в ходе «отбраковки» конфигураций гиперпараметров моделей. Предложенная модификация направлена на сокращение времени, затраченного на подбор оптимальной конфигурации ИНС. А также позволяет устранить недостаток искажения распределения данных. В результате «плохие» конфигурации не должны быть удалены досрочно, так как их влияние в частичном распределении может положительно отразиться на итоговом качестве выходной модели.

Глава 3. Применение разработанных методов и алгоритмов в задачах подготовки данных и обучения моделей сегментации сложных объектов исследования

В данной главе описываются результаты работы технологии синтеза сегментационных масок изображений. Представлено детальное описание алгоритма создания исходной маски объекта исследования на основе анализа карт признаков классификационной модели. Приведены результаты повышения конечной точности сегментирования с выравниванием границ, полученные в результате анализа trimap карты FBA-Net моделью. Апробация полученных результатов демонстрируется в задаче сегментации атеросклеротических бляшек.

Значимое место занимает результат настройки гиперпараметров нейросетевых моделей с использованием модифицированного алгоритма ВОНВ. В результате которого определено, что итоговым решением является нейросетевой ансамбль. Разработанный ансамбль нейронных сетей подвергался сравнительному анализу с набором широко используемых архитектур для мультиклассовой сегментации, включающих U-Net, Unet++, LinkNet, FPN, PSPNet, PAN, DeepLabV3, DeepLabV3+ и MA-Net.

С целью верификации эффективности и устойчивости разработанная технология и алгоритмические решения апробированы не только на задачах сегментации атеросклеротических бляшек, но и сегментации различных анатомических структур, таких как просвет артериолы, средняя оболочка артериолы, адвентиция артериолы, просвет вены, стенка вены, просвет капилляра, стенка капилляра, иммунные клетки и нервные стволы, что позволяет сделать вывод о универсальности разработанных метода и алгоритмов.

3.1 Создание исходной маски объектов на основе карт признаков

3.1.1 Карты признаков

Классификация изображений — это процесс автоматического разделения изображений на группы или классы по определенным критериям. Этот процесс является важным инструментом для упорядочивания и организации больших объемов данных. Для классификации изображений используются различные методы машинного обучения, поскольку они позволяют обрабатывать большие объемы данных и выделять наиболее значимые признаки, что дает возможность точно определять классы изображений. Подготовка данных для получения качественной модели классификации и ее обучение занимает сравнительно небольшое количество времени. Это преимущество было использовано для разработки нового метода генерации масок сегментации.

Для достижения качественного результата в кратчайшие сроки была применена технология трансферного обучения. Обучение модели классификации 4 классов атеросклеротических бляшек проводилось на предварительно обученных нейросетевых моделях на базе ResNet-50, EfficientNetB5, Xception, ConvNeXtBase обученные на 350 миллионах изображений Imagenet с сохранением весов для определения 17 000 классов. Перед обучением к модели был приведен ряд преобразований указанных в Главе 2. Данные модели были выбраны как архитектурные решения, хорошо зарекомендовавшие себя при решении сложных задач медицинской области.

Все модели обучались в равных условиях с заданным набором гиперпараметров, представленным в таблице 3. Обучение проводилось в течении 75 эпох с использованием предварительно размеченных 25 698 изображений. Результаты, полученные в ходе тестирования моделей приведены в таблице 4.

Таблица 3 – Гиперпараметры использованные для обучения моделей классификации

Гиперпараметр	Значение
Batch size	32
Input size	512x512
Optimizer	Adam
Learning rate	12^{-4}

Таблица 4 – Результаты работы моделей классификации на тестовом множестве

Наименование модели	Accuracy, %	Precision, %	Recall, %	Speed, s
ResNet-50	74.3	70.5	70.3	0,0018
EfficientNetB5	78.2	68.4	67.2	0,0025
Xception	97.8	91.7	91.7	0,0013
ConvNeXtBase	69.9	65.4	65.7	0,0021

Согласно полученным результатам, можно сделать вывод, что лидером среди моделей классификации атеросклеротических бляшек стала модель Xception с точностью классификации по метрике Accuracy 97,8% на тестовом множестве. Время обучения модели составило 5,5 часов, а время обработки одного изображения - 0,0013 с. Полученная эффективность модели позволяет использовать ее в дальнейших экспериментах создания сегментационной маски изображения.

Для оценки интерпретируемости моделей классификации и их ориентации на релевантные морфологические признаки бляшек были применены различные методы CAM, включая GradCAM, GradCAM++, HiResCAM, LayerCAM и другие. Представленные ниже сравнительные карты активации демонстрируют вариативность работы алгоритма CAM для признаков разного размера и

сложности, а также сравнение с результатами, полученными в ходе работы авторского алгоритма.

Ниже приведены результаты по каждой из характеристик бляшек, полученные в ходе анализа:

- Lumen:
 - LayerCAM, HiResCAM, GradCAMElementWise и EigenGradCAM постоянно демонстрировали превосходную производительность в определении границ просвета с высокой точностью. Эти алгоритмы эффективно подчеркивали контуры и минимизировали ложные активации за пределами просвета, что свидетельствует о надежной локализации признаков.
 - GradCAM, GradCAM++ и XGradCAM продемонстрировали умеренную производительность. Хотя они улавливали общую форму просвета, случайные активации на углах изображения или фоновые артефакты снижали их надежность.
 - EigenCAM показал низкие результаты, часто ошибочно классифицируя окружающие области как часть просвета, что снижало интерпретируемость.
- Fibrous cap:
 - LayerCAM, HiResCAM, GradCAMElementWise и EigenGradCAM отлично выделяли тонкий слой фиброзной капсулы, что является критической характеристикой для оценки уязвимости бляшки. Эти методы точно определяли границы фиброзной капсулы, даже в сложных случаях с диффузной текстурой.
 - GradCAM, GradCAM++ и XGradCAM показали средние результаты, при этом результаты иногда включали ложные активации на краях изображения.

- EigenCAM показал низкие результаты, часто ошибочно активируя области липидного ядра, а не на самой фиброзной структуре, что подрывает его клиническую значимость.
- Lipid core:
 - Сегментация липидного ядра оказалась сложной для всех протестированных алгоритмов. Слабые или неправильные активации наблюдались во всех методах CAM, включая GradCAM, HiResCAM, GradCAMElementWise, GradCAM++, XGradCAM, EigenCAM, EigenGradCAM и ScoreCAM, а также авторский алгоритм. Данное ограничение, вероятно, связано со сложной морфологией липидного ядра, нечеткими границами и встраиванием в окружающие ткани. Полученные данные подчеркивают необходимость совершенствования представленных алгоритмов, а также применение дальнейшей обработки с целью достижения качественных данных.
- Vasa Vasorum:
 - Большинство алгоритмов, за исключением EigenCAM, показали адекватные результаты при идентификации класса vasa vasorum. Исследуемые методы правильно выделили мелкие, менее отчетливые сосудистые структуры с незначительными отклонениями. EigenCAM демонстрировал значительные ошибки в классификации, часто активируясь в области просвета, где структуры vasa vasorum отсутствовали, что снижает его надежность при выявлении мелких анатомических особенностей

Полученные карты активации подтверждают ориентацию моделей на анатомически значимые области и обеспечивают механизм для выявления потенциальных областей неправильной классификации или двусмысленности. Примечательно, что, LayerCAM и HiResCAM, продемонстрировали стабильную производительность в создании точных визуализаций по всем признакам бляшек, превосходя другие методы благодаря своей способности агрегировать

многомасштабные активации и локализовать тонкие структуры (например, тонкие фиброзные капсулы) с минимальным количеством ложных срабатываний. Их градиентно-взвешенные механизмы внимания дополнительно подавляют артефакты изображения, присущие ОКТ, такие как спекл-шум, сохраняя при этом важнейшие морфологические детали. Применение этих методов позволяет в кратчайшие сроки получить сегментационные маски анатомических объектов удовлетворительной точности, что позволяет составить большой обучающий набор данных и повысить уверенность в клинической применимости будущей модели. Визуализация результата каждого рассмотренного метода приведена на рисунках 14–17.

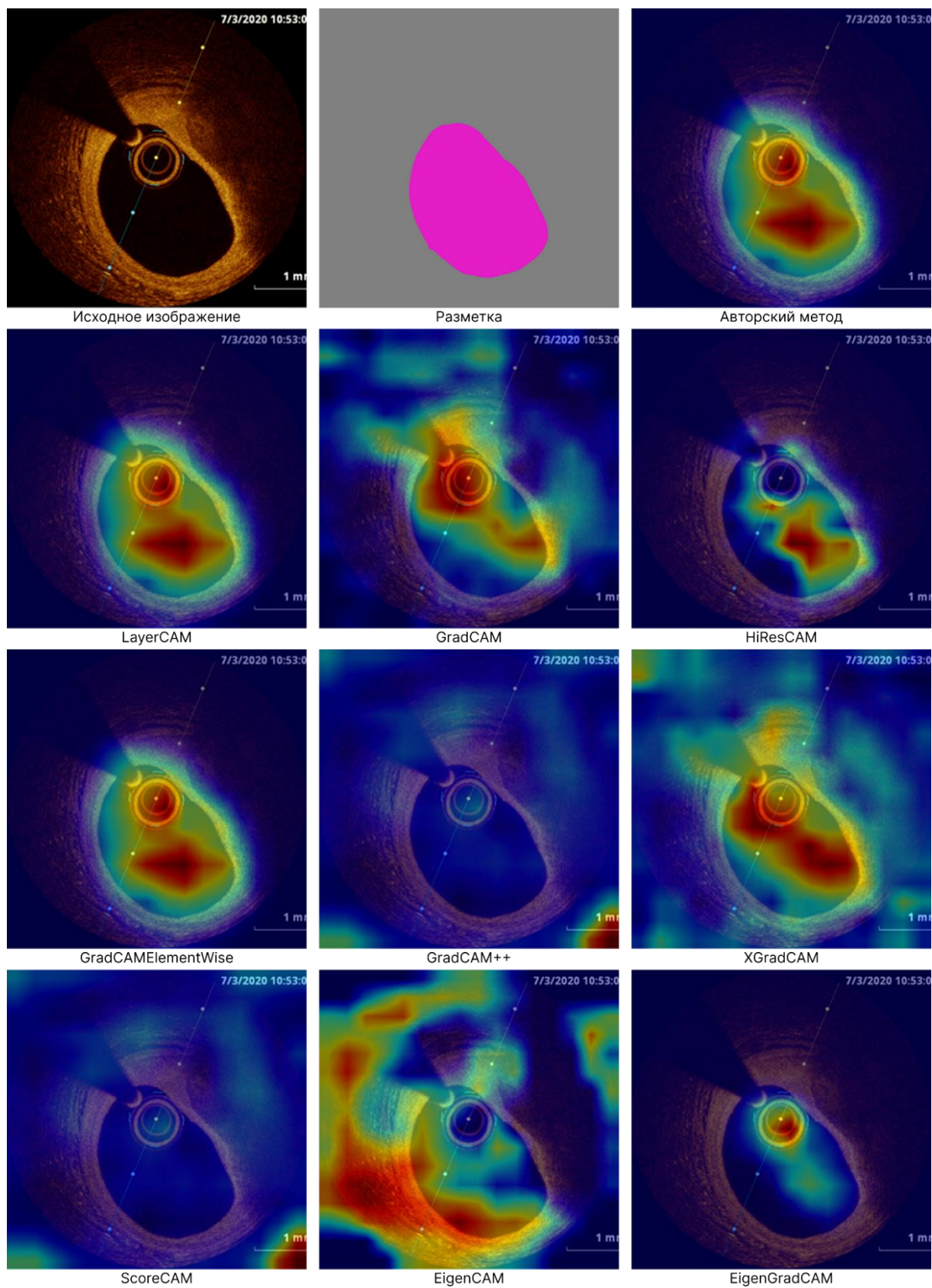


Рисунок 12 – Результат визуализации карты признаков. Lumen

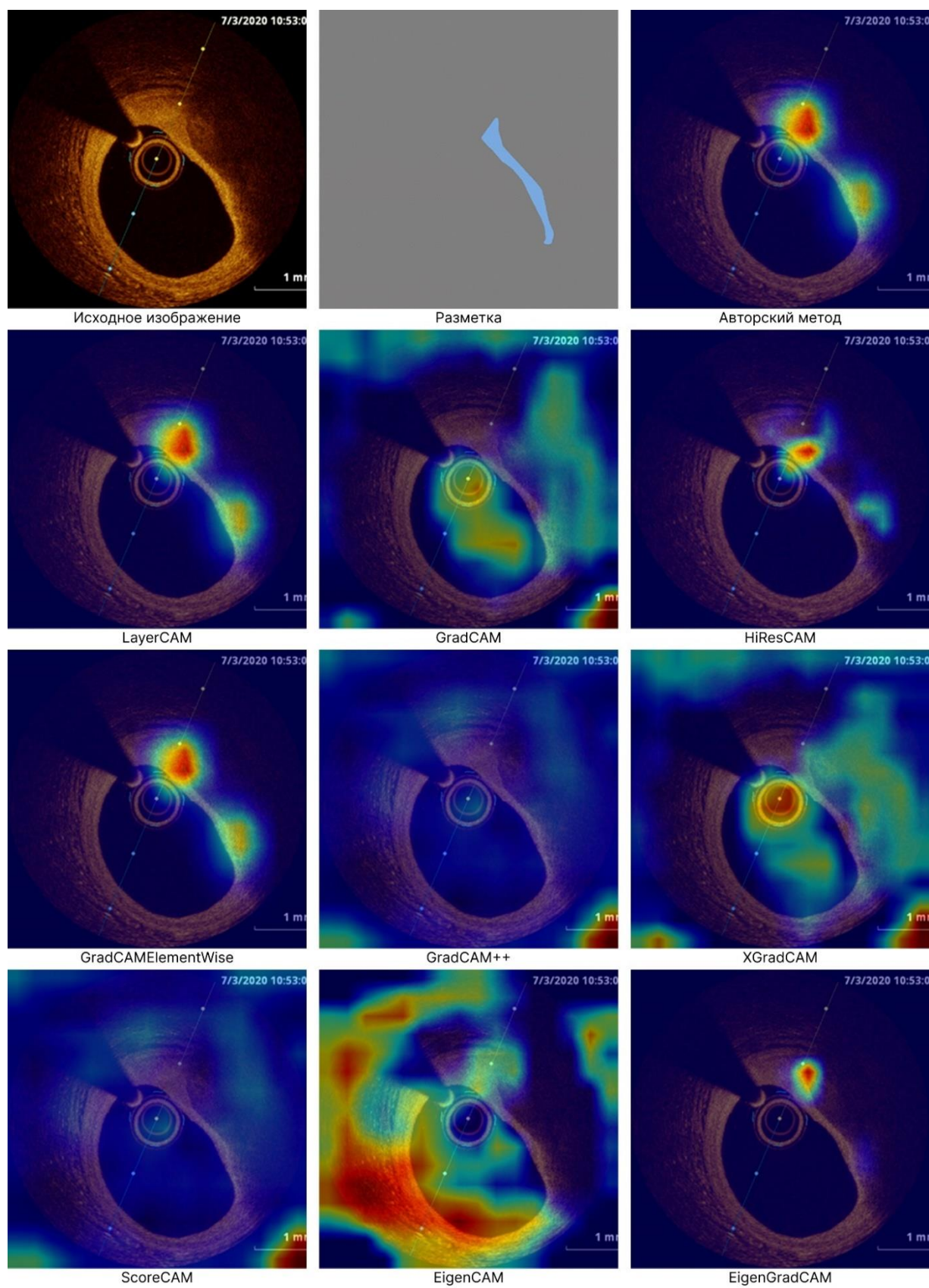


Рисунок 13 – Результат визуализации карты признаков. Fibrous cap

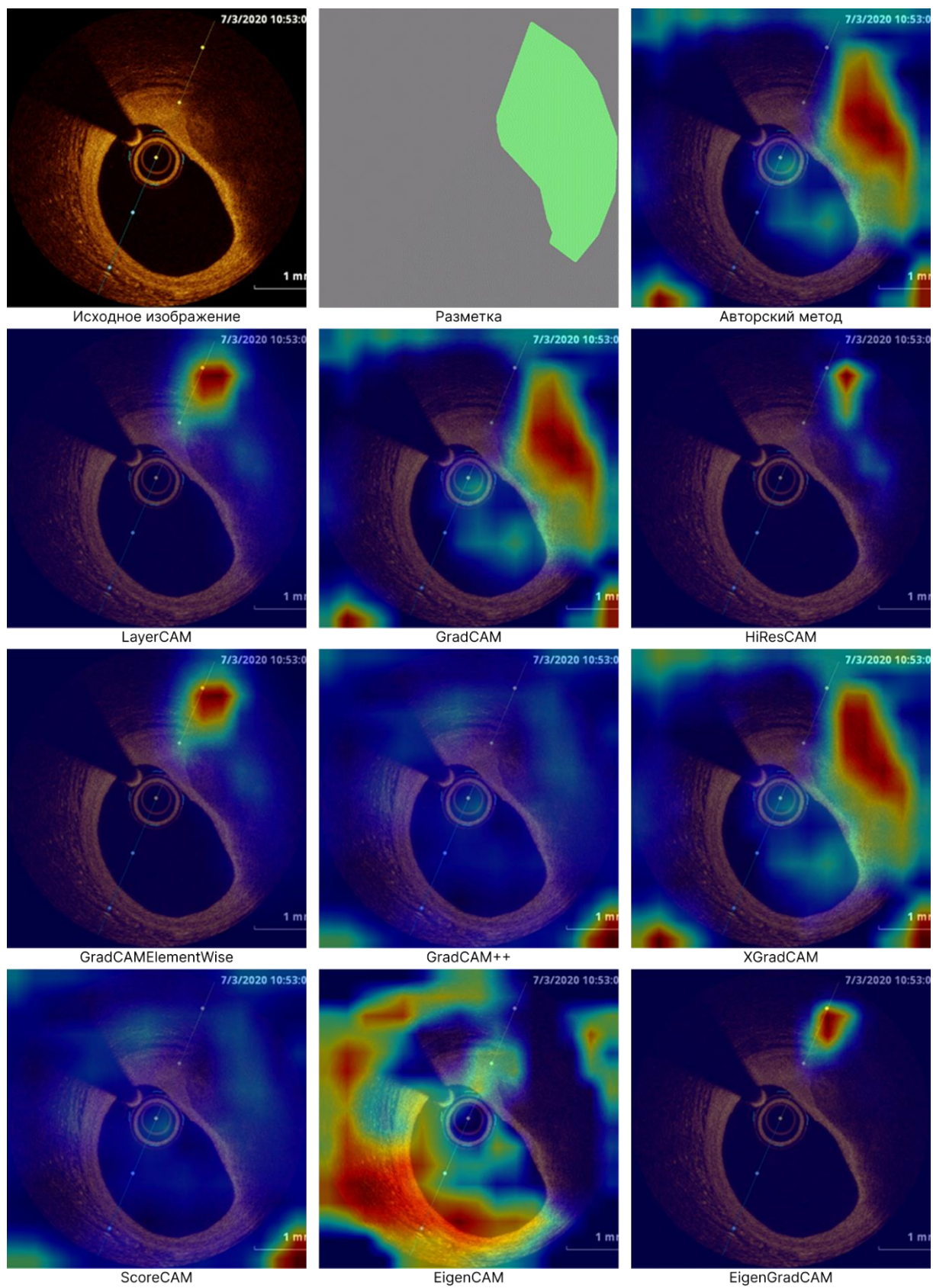


Рисунок 14 – Результат визуализации карты признаков. Lipid core

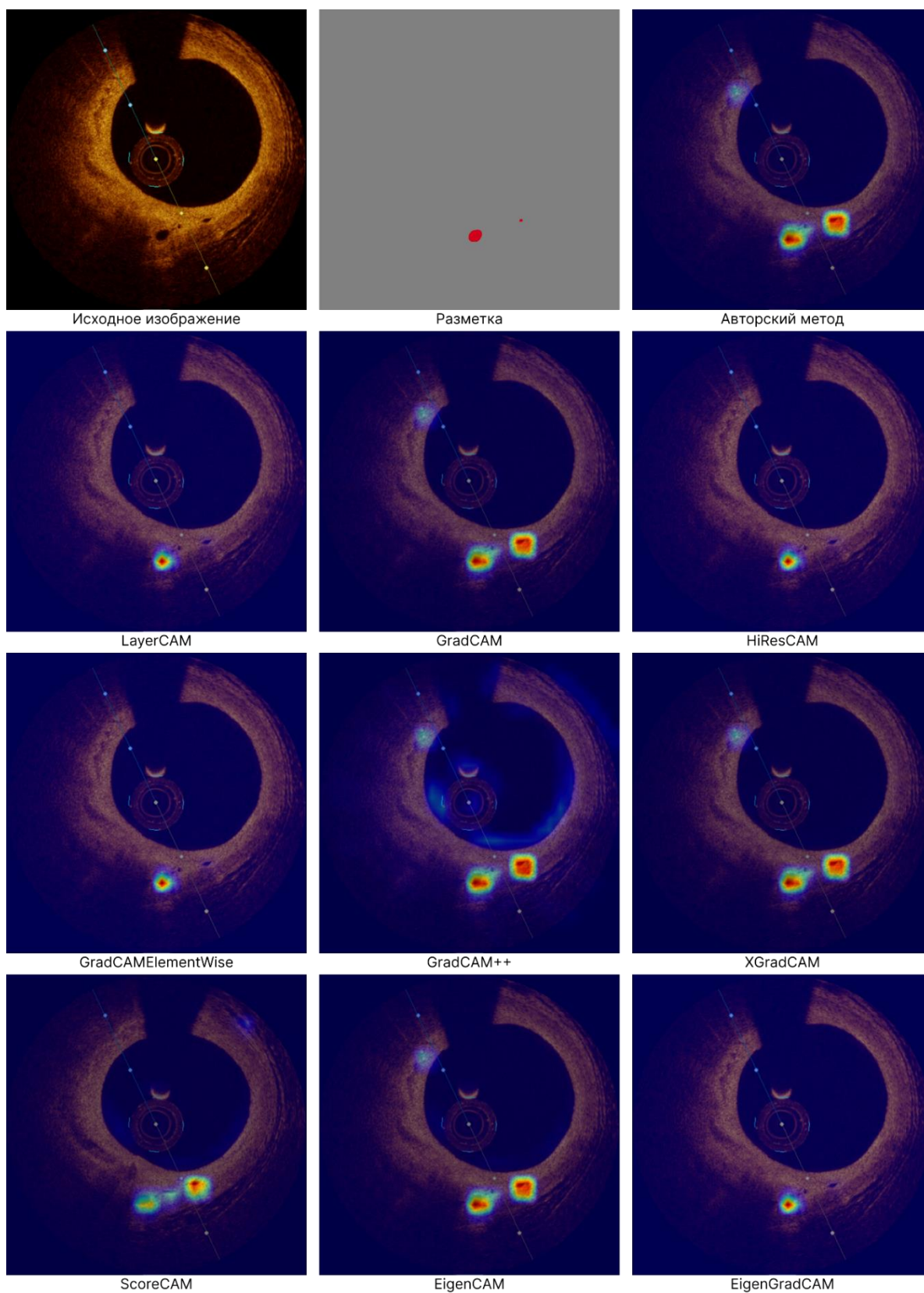


Рисунок 15 – Результат визуализации карты признаков. Vasa vasorum

Представленный авторский метод создания сегментационных масок изображений на основе карт признаков демонстрируем высокую производительность по сравнению с представленными в работе аналогами. В результате качественной оценки установлено, что представленный алгоритм не уступает аналогам, что конечно же связано со схожестью алгоритмов.

Метрика качества IOU (англ. Intersection over Union) является важным инструментом для оценки качества сегментации объектов на изображениях. Она также часто используется в задачах обнаружения объектов (object detection) и семантической сегментации (semantic segmentation). Идея IOU заключается в том, чтобы оценить степень перекрытия (overlap) между предсказанной границей объекта и истинной границей объекта на изображении. Метрика IOU вычисляется путем деления площади пересечения двух областей на площадь их объединения. Формула для расчета IOU выглядит следующим образом:

$$IoU = \text{Area of Overlap} / \text{Area of Union} \quad (3.1)$$

Чем выше значение IOU, тем лучше модель справляется с задачей сегментации или обнаружения объектов. Обычно IOU выражается в виде процентов (от 0 до 1) или в долях (от 0 до 100%).

При оценке моделей компьютерного зрения, часто используют пороговые значения IOU, например, $IOU \geq 0.5$ для считания детекции правильной. Также существуют вариации метрики IOU, например, средний IOU (mIOU), который усредняет значения IOU по всем категориям объектов на изображении.

Использование метрики IOU позволяет исследователям и инженерам оценивать и сравнивать качество различных моделей компьютерного зрения, что является важным шагом в разработке и улучшении алгоритмов сегментации и детекции объектов на изображениях. Ознакомиться с усредненными по классам количественными показателями в ходе анализа 100 изображений можно в таблице 5.

Таблица 5 – Количественные показатели анализа полученных масок изображений

Алгоритм	IOU, %	Speed, s
GradCAM	21.6	0.0068
HiResCAM	27.8	0.0072
GradCAMElementWise	16.3	0.007
GradCAM++	23.5	0.0071
XGradCAM	25.9	0.0071
ScoreCAM	22.4	0.0072
LayerCAM	28.1	0.0076
EigenGradCAM	9.2	0.008
Method	27.2	0.0024

В результате анализа полученных показателей в качестве итогового решения выбран авторский алгоритм. Данный подход не только продемонстрировал высокие показатели точности, уступающие модели фавориту на 0.9%, но и показал наивысший результат по характеристике – время анализа. Такие методы как LayerCAM и HiResCAM также имеют место для применения в представленном методе формирования маски сегментации изображения, за счет достижения наивысших показателей точности. Стоит отметить, что данные методы демонстрируют наименьшее проявление негативных артефактов во время анализа.

3.1.2 Результат улучшения маски с использования трехканального анализа

Заключительным этапом формирования сегментационной маски изображения является уточнение границ анатомической структуры. Представленный в работе алгоритм формирования трехканальной маски отлично подходит для решения поставленной задачи. На изображениях 18–21 демонстрируется результат работы данного алгоритма.

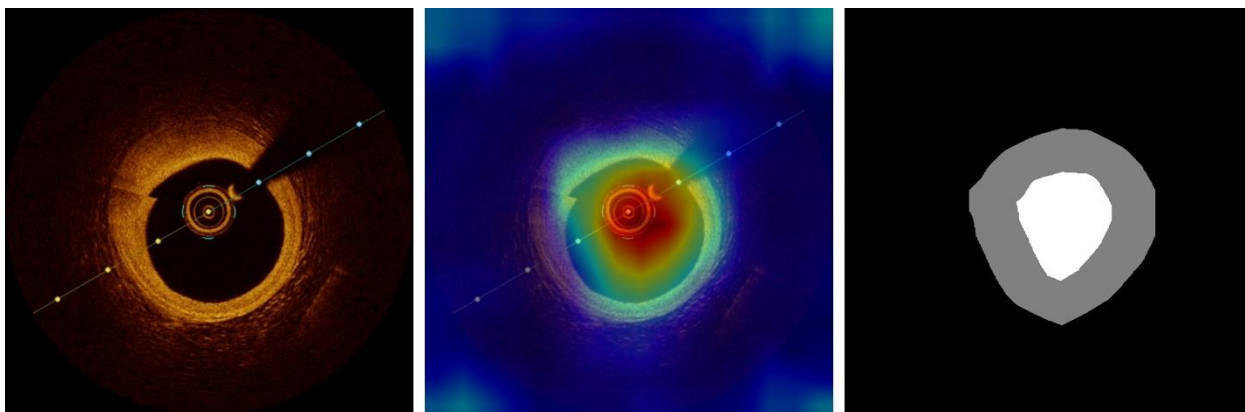


Рисунок 16 – Результат формирования трехканальном маски. Lumen

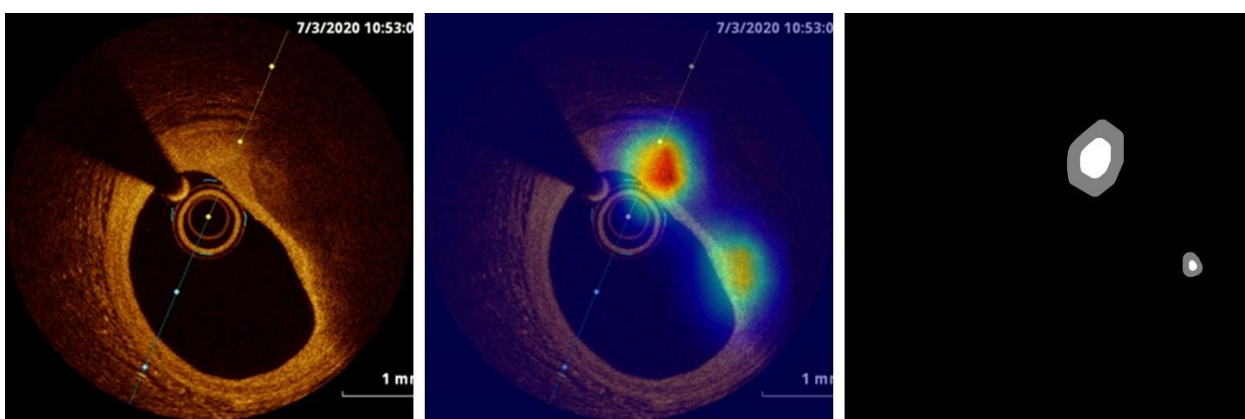


Рисунок 17 – Результат формирования трехканальном маски. Fibrous cap

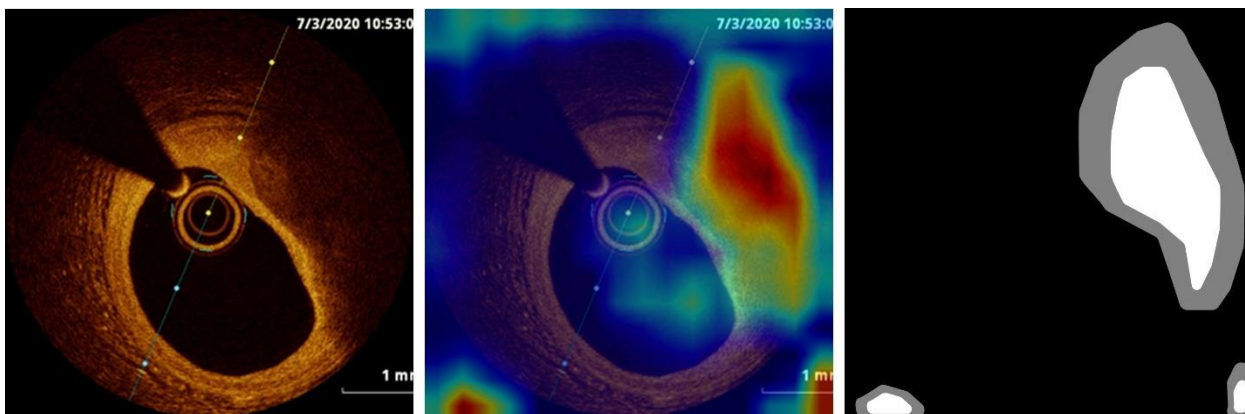


Рисунок 18 – Результат формирования трехканальном маски. Lipid core

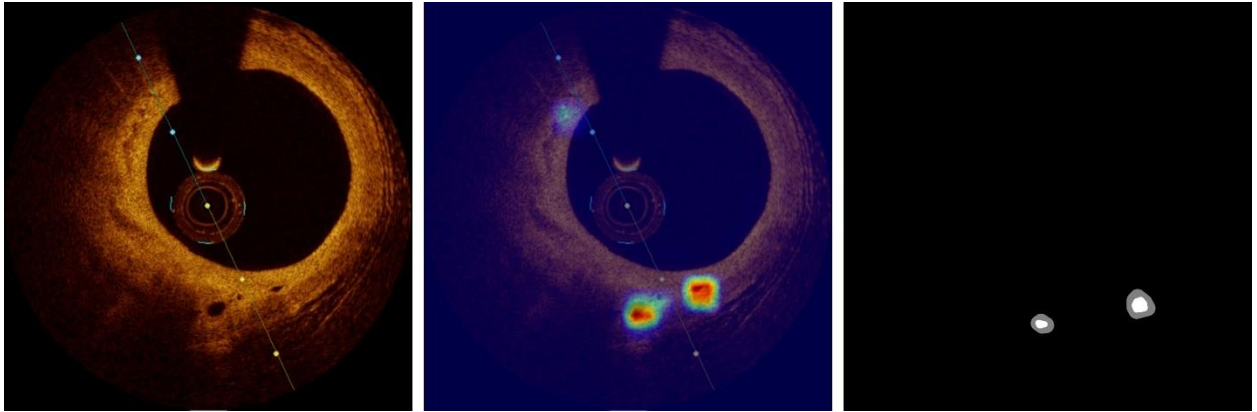


Рисунок 19 – Результат формирования трехканальной маски. Vasa vasorum

На представленных примерах отчетливо видно, что полученной на предыдущем этапе информации достаточно для формирования области объекта исследования, а также области неопределенности, полностью покрывающей область, где может располагаться объект анализа. Результаты анализа трехканальных масок FBA-Net моделью представлены на рисунке 22.

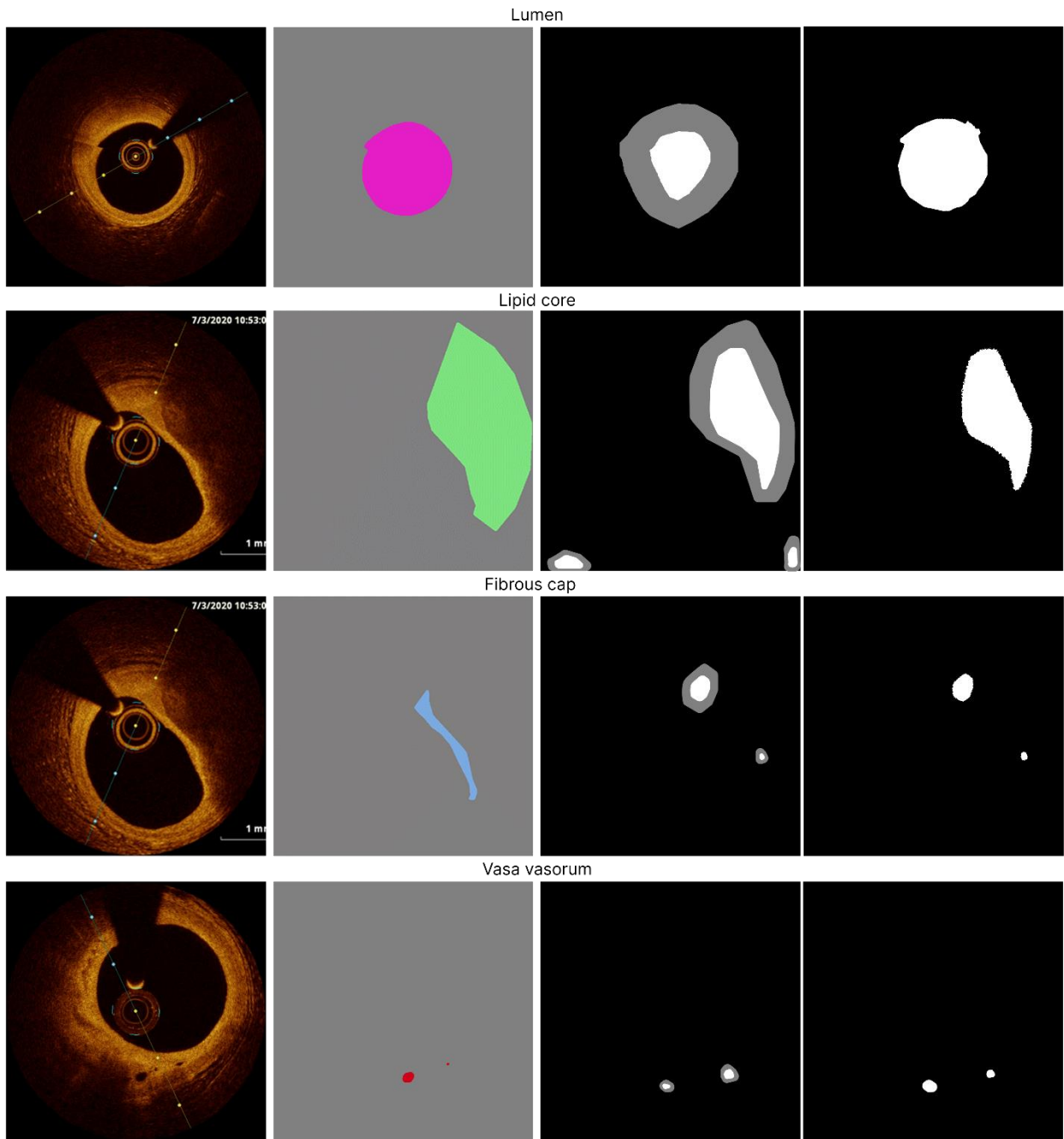


Рисунок 20 – Результат анализа трехканальной маски для классов: Lumen, Fibrous cap, Lipid core, Vasa vasorum

3.1.3 Выводы по работе метода создания сегментационных масок изображений

Из результатов, представленных на рисунке 22 видно, что представленный метод формирования сегментационных масок изображений подходит для тех объектов исследования, которые имеют выражение очертания, такие как: Lumen,

Vasa vasorum. Для объектов, которые не обладают данным признаком, такие как: Fibrous cap и Lipid core представленный метод формирования сегментационной маски не подходит. Основной проблемой для представленных классов выступает отсутствие явных границ объекта, в связи с чем FBA-Net модель не может провести детальное выделение области интереса, также стоит отметить, что именно для классов Fibrous cap и Lipid core создать исходную маску объекта смогло наименьшее число рассматриваемых алгоритмов синтеза карт внимания.

Несмотря на полученные результаты данный метод является эффективным инструментом формирования обучающего множества. Согласно показателям представленным в таблице 6 эффективность данного метода не может быть подвергнута сомнениям.

Таблица 6 – Показатели эффективности метода формирования сегментационных масок изображений, на примере класса Lumen и Vasa vasorum в расчете на 1 изображения

Algorithm	IOU, %	Precision	Recall	Время анализа, s
Ручная разметка	100%	100,0	100,0	176.3
Автоматическое создание маски объекта	88.6	89,7	92,8	0.14
Автоматическое создание маски объекта + экспертная доработка	98.4	98,2	99,0	54.5

Согласно результатам представленным в таблице 6 предложенный в работе метод формирования сегментационных масок изображений эффективнее ручного аннотирования более чем в 3 раза с учетом экспертной проверки. Также стоит отметить, что для формирования качественного обучающего множества как правило требуется не менее 3-х экспертов, для формирования объективного мнения. Важным моментом является ограниченность времени таких экспертов и загруженность прочей работой, так на разметку используемого в работе набора данных ушло 4 месяца. На обучение классификационной модели и формирования

сегментационных масок для классов Lumen и Vasa vasorum с учетом всех проводимых экспериментов было затрачено менее 1 месяца, что еще раз подчеркивает эффективность данного метода.

3.2 Результат обучения моделей в задаче сегментации нестабильных бляшек

3.2.1 Дизайн эксперимента

Разработка моделей сегментации обусловлена неравномерным распределением экземпляров исследуемых классов в наборе данных и отличительными характеристиками каждого класса. Как показано в таблице 7, наиболее распространенными являются объекты Lumen, включающие 21 808 экземпляров, за ними следуют Fibrous cap (7 226 экземпляров) и Lipid core (7 192 экземпляра). В отличие от этого, класс Vasa vasorum представлен значительно в меньшем количестве – всего 450 экземпляров. Такой дисбаланс создает уникальные проблемы для задач сегментации, поскольку производительность модели может быть смещена в сторону более распространенных классов, в то время как редкие классы могут быть недооценены. Для решения этих проблем в работе применяется гибридная стратегия сегментации, адаптированная к специфическому распределению и морфологической сложности каждого исследуемого класса:

- Lumen (LM): Опираясь на высокую представленность просвета в наборе данных, обучена одноклассовая модель.
- Fibrous cap (FC) and Lipid core (LC): Эти два класса, хотя и представлены умеренно, имеют общие морфологические черты, такие как сложные и перекрывающиеся границы. Двухклассовая модель была настроена и обучена для повышения точности сегментации для каждого, анализируемого признаков.
- Vasa vasorum (VV): Редкость этого класса обусловила необходимость создания специальной модели с одним классом, чтобы повысить эффективность работы, сосредоточившись исключительно на специфических характеристиках vasa vasorum.

Такой целенаправленный подход позволяет лучше справляться с дисбалансом классов, обеспечивая оптимальную настройку каждой сети сегментации, оптимально настроенной для соответствующего класса. Изолируя задачи сегментации и согласовывая дизайн модели с распределением классов, мы смягчаем проблемы, присущие несбалансированным наборам данных, и повышаем точность сегментации во всех классах.

Таблица 7 – Распределение морфологических признаков изображений по подмножествам

Множество	Подмножество	LM	FC	LC	VV	Всего объектов	Всего изображений
1	Train	17264	5610	5576	328	28778	16901
	Test	4544	1616	1616	122	7898	4492
2	Train	17554	5709	5690	237	29190	17207
	Test	4254	1517	1502	213	7486	4186
3	Train	17220	5600	5565	407	28792	16962
	Test	4588	1626	1627	43	7884	4431
4	Train	17813	5724	5686	416	29639	17473
	Test	3995	1502	1506	34	7037	3920
5	Train	17381	6261	6251	412	30405	17029
	Test	4427	965	941	38	6371	4364

3.2.2 Тюнинг нейросетевых моделей

В работе оценивалось 350 сетевых конфигураций для каждой модели (всего 1050 конфигураций). Основные результаты обобщены ниже, в Таблице 8, а также в Таблицах A1 - A3 приложения A:

Время настройки: Длительность процесса настройки значительно варьировалась между различными классами и конфигурациями. Сегментация FC и LC потребовала наибольшего времени – 712 часов, что, вероятно, отражает сложность этих конкретных классов. Для сравнения, модель сегментации LM завершила свою настройка была завершена за 309 часов. Сегментация Vasa vasorum, имея меньший набор данных, показала самое короткое время настройки - 103 часа.

Производительность моделей. Эффективность сегментации каждой модели оценивалась с помощью нескольких метрик, включая precision, recall, F1 score, IOU и коэффициент сходства Dice. Эти метрики вычислялись на усеченном наборе данных из 40 пациентов, использовавшемся исключительно для настройки модели, чтобы отразить производительность на этапе тонкой настройки. U-Net++ достигла наивысшего показателя DSC 0,989 для сегментации LM, продемонстрировав практически идеальное согласие между предсказанными и истинными сегментами. Для сегментации FC и LV, LinkNet достигла DSC 0,617, что свидетельствует об умеренной точности в среднем. Наконец, U-Net показала относительно высокие результаты относительно хорошо в сегментации Vasa vasorum, с DSC 0,784, отражающим хороший уровень точности для этого небольшого набора данных.

Сложность модели: Сложность модели оценивалась по количеству параметров и операций умножения-накопления (MAC). Модель U-Net++, используемая для сегментации LM, имела 68 миллионов параметров и требовала 249,8 GMACs вычислительной мощности. LinkNet, используемая для сегментации FC и LC, имел 64,4 миллиона параметров и продемонстрировал самую низкую вычислительной нагрузки, потребовав всего 9,0 GMACs. В отличие от этого, U-Net, применяемая для сегментации VV, имела наименьшее количество параметров - 31,9 млн. миллионов и потребовал 158 GMAC.

Подводя итог, можно засвидетельствовать, что полученные результаты подчеркивают вариативность производительности модели в различных задачах

сегментации. U-Net++ продемонстрировала превосходную производительность при сегментации Lumen, в то время как LinkNet и U-Net продемонстрировали устойчивость при сегментации Fibrous cap, Lipid core и сегментации Vasa vasorum. Эти результаты подчеркивают компромисс между точностью и сложностью модели точностью и сложностью модели, подчеркивая необходимость оптимизации под конкретную задачу при сегментации атеросклеротических бляшек.

Таблица 8 – Оптимальные гиперпараметры для изученных морфологических признаков бляшки

Гиперпараметр	Lumen	Fibrous cap & Lipid core	Vasa vasorum
Architecture	U-Net++	LinkNet	U-Net
Encoder	ResNet-101	EfficientNet B7	RegNetX-6.4GF
Input size	512x512	896x896	896x896
Optimizer	RMSprop	RMSprop	RAdam
Learning rate	10^{-5}	10^{-5}	10^{-3}
Parameters, M	68.0	64.4	31.9
MACs, G	249.8	9.0	158.0
Precision	0.991	0.700	0.814
Recall	0.993	0.977	1.000
F1	0.989	0.617	0.784
IoU	0.979	0.488	0.695
DSC	0.989	0.617	0.784

Проведение комплексной оценки динамики обучения для моделей позволило отобрать на этапе настройки гиперпараметров следующие фаворитные решения: U-Net++ для сегментации Lumen, LinkNet для сегментации Fibrous cap и Lipid core и U-Net для сегментации Vasa vasorum.

3.2.3 Обучение ансамбля нейросетевых моделей

Все представленные ранее модели были обучены в течение 125 эпох с использованием 5-кратной перекрестной проверки для обеспечения надежной

оценки их эффективности. Такая система перекрестной валидации позволила отследить тенденции в функции потерь и DSC в течение эпох для каждого анализируемого класса, что обеспечило детальный мониторинг поведения сходимости, что можно заметить на графике, представленном на рисунке 23.

В результате обучения отдельных моделей сегментации сделан следующий вывод. Модели сегментации Lumen, Fibrous cap и Lipid core демонстрируют четкое и последовательное уменьшение потерь, сопровождающееся прогрессивным улучшением DSC, что указывает на эффективное обучение и уточнения прогнозов сегментации. Модель сегментации vasa vasorum, напротив, демонстрировала более медленную скорость сходимости, сопровождающуюся значительными колебаниями в значениях потерь и DSC. Эти вариации подчеркивают присущую сложность и проблемы, связанные с сегментацией данной анатомической структуры.

Для систематической оценки сходимости были определены точки стабилизации для функции потерь и DSC, где инкрементные улучшения были минимальными. U-Net++ (сегментация Lumen) и LinkNet (сегментация Fibrous cap и Lipid core) достигли стабилизации быстро, требуя меньше эпох для достижения высоких значений DSC с минимальными потерями. В отличие от них, U-Net (сегментация Vasa vasorum) демонстрировала длительную конвергенцию и менее стабильные кривые обучения, что указывает на трудности в обучении тонких характеристик данного класса. Такие колебания отражают структурную сложность и изменчивость Vasa vasorum, что, вероятно, требует более тонкого выделения и агрегирования признаков.

Наблюдаемые различия в сходимости могут быть объяснены несколькими факторами, включая архитектурные различия между моделями, различия в инициализации параметров, и внутреннюю сложность сегментации каждого исследуемого признака. Например, модель Vasa vasorum столкнулась с проблемами в улавливании тонких морфологических особенностей, что

потребовало более длительного обучения и усовершенствованных архитектурных механизмов для эффективного анализа морфологических особенностей.

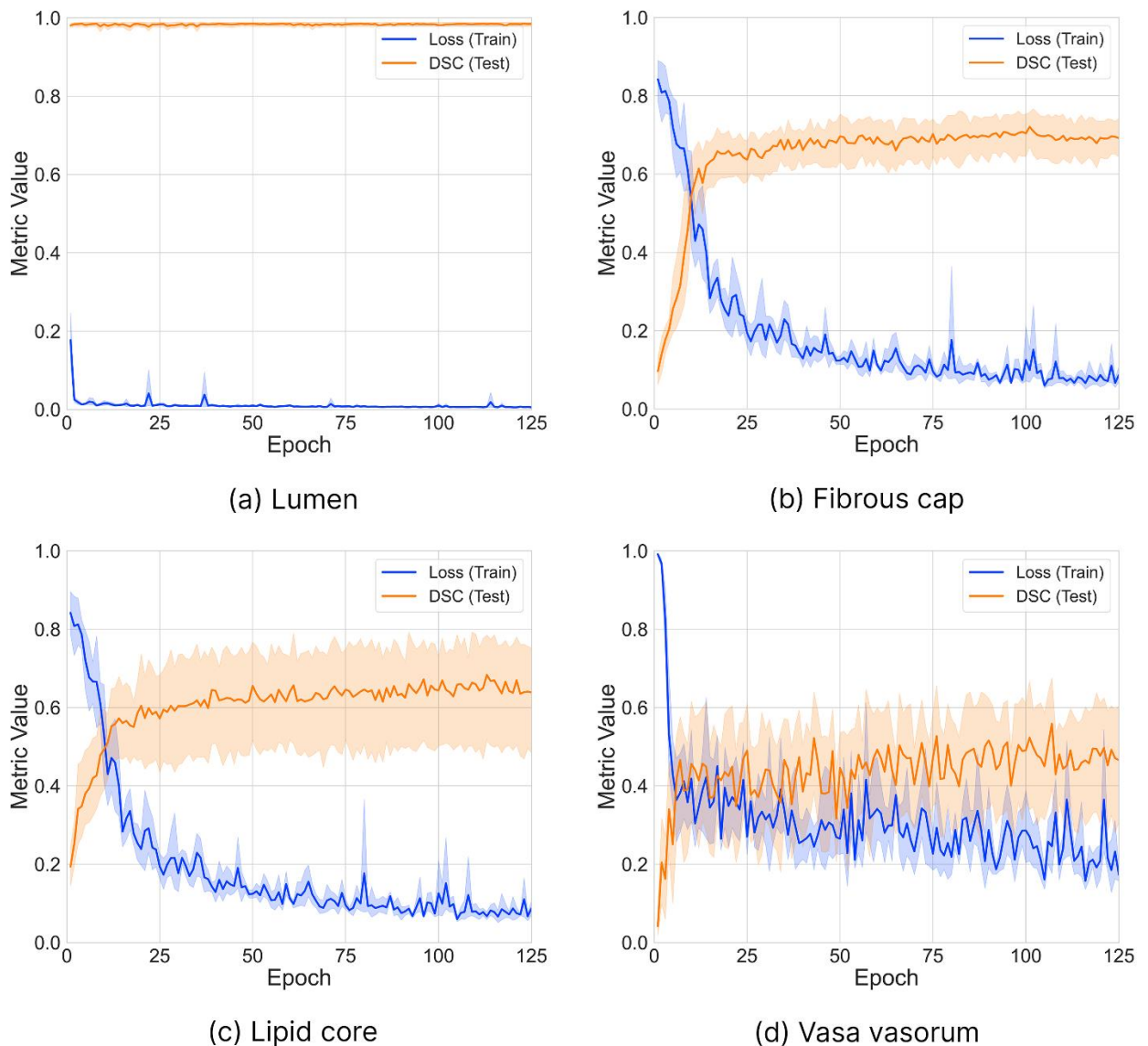


Рисунок 21 – Сравнительный анализ изменения функции потерь и DSC на этапах обучения и тестирования с помощью 5-кратной кросс-валидации с 95% доверительным интервалом

Оценка показателей эффективности сегментации выявила различные уровни точности для всех оцениваемых морфологических признаков бляшки, подробнее можно ознакомиться на рисунке 24 и таблице 9. Модель сегментации Lumen демонстрирует исключительную производительность, достигая DSC 0,987 и IOU 0,975. Данные показатели подчеркивают способность модели надежно и точно очерчивать структуры класса Lumen, подчеркивая ее надежность и пригодность

для клинических приложений, требующих точной сегментации. Модель Fibrous cap и Lipid core демонстрирует надежную работу с показателями DSC 0,736 и 0,751 соответственно. Данные показатели отражают эффективность модели в отображении сложных и запутанных границ этих структур, несмотря на трудности, связанные с их морфологией. Модель сегментации Vasa vasorum демонстрирует умеренную эффективность, получив DSC 0,610. Стоит отметить, что сегментация данного участка все еще остается сложной из-за тонких и менее четких границ. Результаты сегментации свидетельствуют о прогрессе в решении этих сложных задач. Интегрируя эти специфические для каждой задачи модели во взвешенный ансамбль, который учитывает распространенность классов и доверие к модели, объединенный ансамбль достиг взвешенного DSC 0,882 для всех исследуемых морфологических признаков бляшки. Данный подход представляет собой значительное улучшение по сравнению с отдельными моделями и ансамблем других комбинаций, демонстрируя способность гармонизировать различные архитектурные преимущества и смягчать проблемы дисбаланса классов, присущие гетерогенной характеристике бляшек. В качестве сравнения и демонстрации представленного преимущества можно обратиться к таблице Б1-Б3 приложение Б дополнительной информации, где представлены результаты исследования отличные от представленного.

Таблица 9 – Показатели эффективности сегментации для каждого морфологического признака бляшки, усредненные по 5 складкам во время перекрестной проверки

Класс	Precision	Recall	F1	IOU	DSC
Lumen	0.986	0.988	0.987	0.975	0.987
Fibrous cap	0.737	0.784	0.736	0.608	0.736
Lipid core	0.815	0.772	0.751	0.639	0.751
Vasa vasorum	0.664	0.630	0.610	0.511	0.610

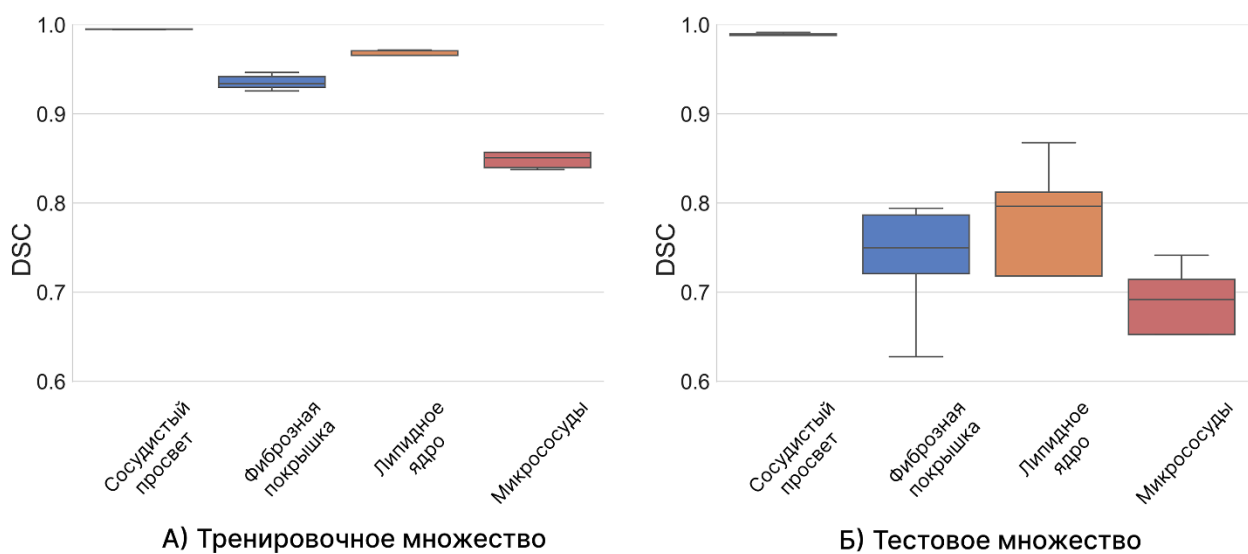


Рисунок 22 – Средняя производительность различных морфологических признаков бляшек

Для наглядной демонстрации эффективности сегментационного процесса в работе приведены три типичных примера, показывающих итоги идентификации морфологических особенностей атеросклеротической бляшки, представленных на рисунке 25. Представленные срезы отображают предсказания, полученные оптимальным ансамблем моделей: U-Net++ для сегментации Lumen, LinkNet для сегментации Fibrous cap и Lipid core, и U-Net для сегментации Vasa vasorum. Предсказания сопоставляются с соответствующими аннотациями, полученными в ходе экспертной разметки, что позволяет напрямую визуальную оценить эффективность модели.

Как показано на рисунке 25, сегментации Lumen и Vasa vasorum демонстрируют сильное совпадение с истинным положением дел, что подчеркивает эффективность ансамбля в выделении данных особенностей. Из чего следует вывод, что размер анализируемого объекта не оказывает негативного влияния на конечное решение. Данное утверждение подтверждается количественными метриками, полученными на тестовом множестве. Стоит отметить, что наблюдаются тонкие расхождения в сегментации Fibrous cap и Lipid core. Выявленные проблемы, вероятно, связаны с более сложной структурой и морфологией исследуемых элементов, которые значительно отличаются от Lumen и Vasa vasorum. Такие морфологии, как Fibrous cap и Lipid core являются компонентами ткани с диффузными и сложными границами, что затрудняет их точное определение. Отсутствие явной границы объекта не позволяет с высокой точностью определить истинность локализации исследуемого объекта.

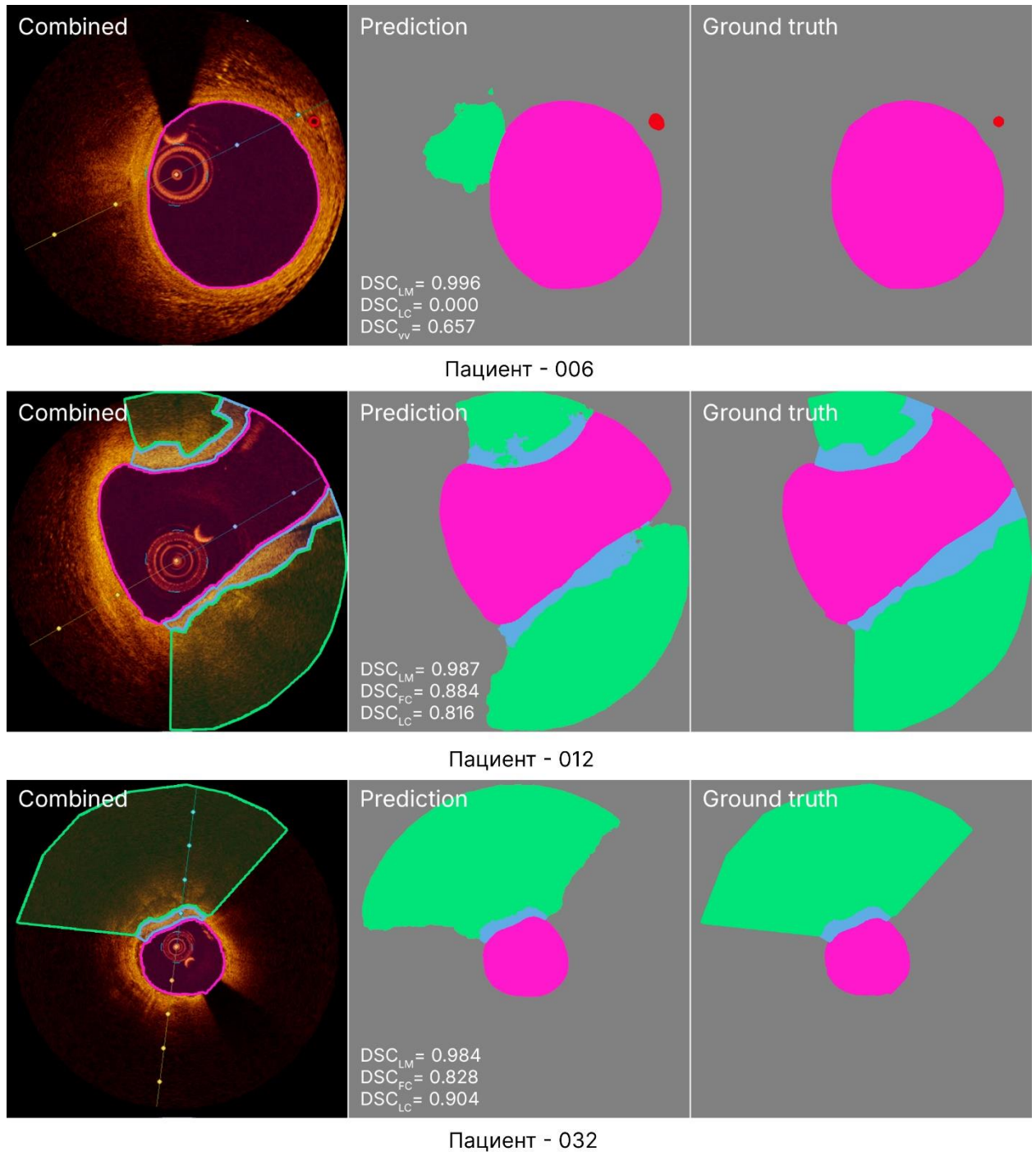


Рисунок 23 – Сравнение сегментации по истине и прогнозов ансамбля

В частности, Fibrous cap, представляет собой тонкий слой фиброзной ткани, покрывающий Lipid core, и его граница может значительно отличаться по текстуре и толщине. Между тем, Lipid core неправильная геометрия и частые вкрапления в окружающую ткань еще больше усложняют его сегментацию. Описанные внутренние структурные проблемы, вероятнее всего, способствуют наблюдаемым

различиям в точности сегментации. Представленные визуальные сравнения дополняют количественные показатели эффективности и позволяют понять, в каких именно областях необходимы улучшения. Дополнительная визуализация предсказаний моделей, подчеркивающая их производительность по различным морфологическим особенностям, представлена на рисунке В1 в приложении В.

3.3 Апробация предложенного метода в задаче сегментации гистологических аномалий

В данном разделе диссертационной работы представлено решение задачи автоматизации гистологической оценки тканеинженерных конструктов с использованием методов машинного обучения. В качестве исходных данных использован уникальный датасет из 104 цельных срезов (WSIs), полученных от 20 овец после шестимесячной имплантации сосудистых трансплантатов (TEVGs).

В исследовании использовались 104 WSI-изображения со средними размерами $135\,000 \times 123\,000$ пикселей, которые автоматически сегментировались на 99 831 фрагмент размерностью $3\,000 \times 3\,000$ пикселей. Полученные фрагменты обрабатывались посредством алгоритма Entropy Masker, применяющего специализированные энтропийные параметры [113] для эффективного маскирования гистопатологических WSI. Данная процедура селективной обработки обеспечивала сохранение исключительно тканесодержащих участков (рисунок 26), гарантируя отбор релевантного материала для дальнейших вычислительных процедур и аналитических операций.

Алгоритм Entropy Masker предназначен для разделения тканевых структур (передний план) и фоновых областей на цифровых гистологических препаратах полного среза (WSI), что особенно актуально при работе с образцами, отличающимися высокой степенью пористости и гетерогенности архитектуры. В основе данной технологии лежит использование энтропийных характеристик изображения — параметра, позволяющего численно характеризовать степень

хаотичности или вариативности пиксельного распределения в рамках определенной зоны или локального участка. Начальная стадия алгоритма включает гистерезисную цветовую сегментацию для устранения фоновых артефактов и неравномерностей. Впоследствии производится расчет локальной энтропии для каждого элемента изображения, что становится определяющим фактором для дифференциации между шумовыми компонентами фона и анатомическими структурами. Путем установки порогового значения определяется оптимальный диапазон энтропии, позволяющий различить фон и ткань. Каждому пикселю присваивается логическое значение (0 или 1) на основе сравнения его энтропии с порогом. Эти бинарные значения формируют маску сегментации ткани для каждого WSI. Метод обеспечивает точное выделение релевантных областей даже в условиях сложной морфологии тканей, что делает его особенно полезным для анализа гистологических изображений в биомедицинских исследованиях.

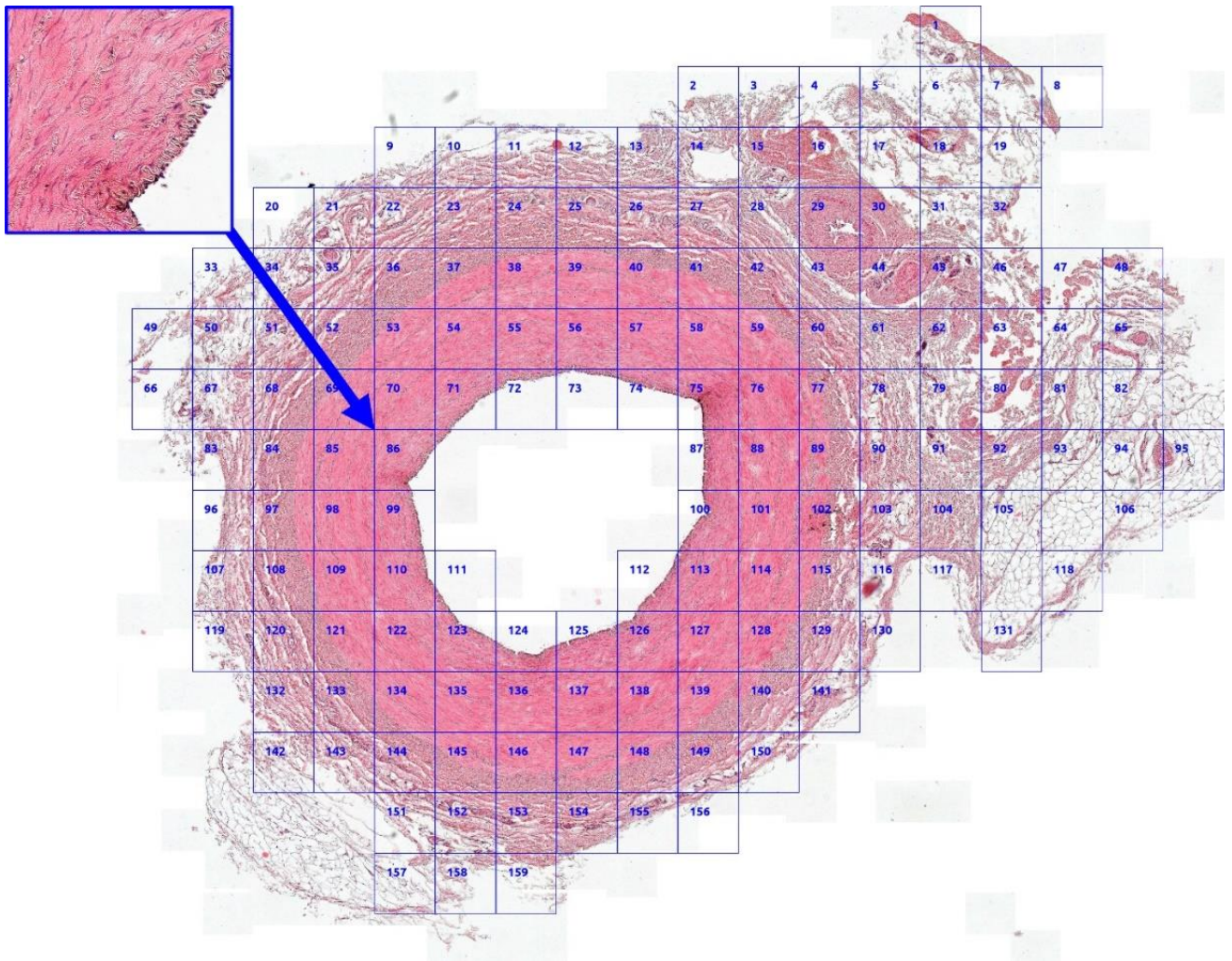


Рисунок 24 – Методика преобразования целого слайд-изображения в подмножество патчей

В конечном итоге был сформирован независимый набор из 1401 изображения. В результате тщательной аннотации было выявлено 9 гистологических признаков, связанных с различными моделями регенерации сосудистой ткани. Эти признаки включали просвет артериолы, медиа артериолы, адвентиция артериолы, просвет венулы, стенка венулы, капилляр, просвет капилляра, стенка капилляра, иммунные клетки и нервные стволы. Диаграмма распределения данных по классам представлена на рисунке 27.

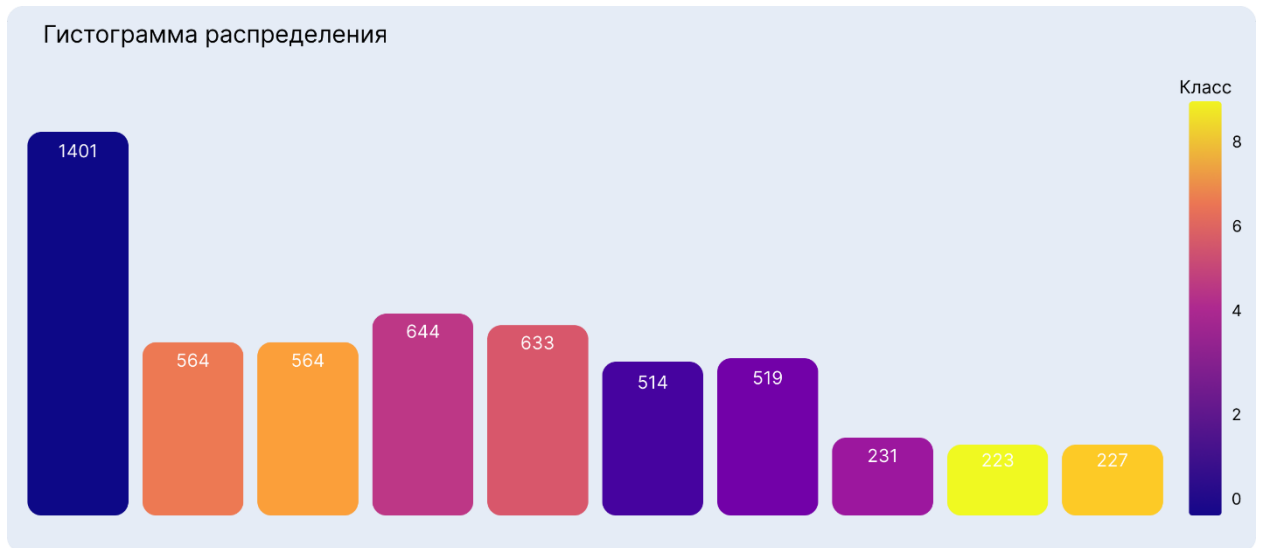


Рисунок 25 – Диаграмма распределения данных по классам

Классы, описанные при аннотации гистологических изображений, представляют различные гистологические структуры *vasa vasorum*, кластеров иммунных клеток и нервных стволов, которые совместно обеспечивают ремоделирование сосудистой ткани и адекватный ответ сосудов на вазоактивные стимулы. Краткое описание каждого класса приведено ниже:

- Просвет артериол (Arteriole Lumen, AL): Полый канал, по которому кровь течет через артериолы — мелкие сосуды, ответвляющиеся от артерий и транспортирующие насыщенную кислородом кровь от сердца к капиллярам.
- Медиа артериол (Arteriole Media, AM): Средний слой артериол, состоящий из гладкомышечной ткани и регулирующий артериальное давление.
- Адвентиция артериол (Arteriole Adventitia, AA): Наружный соединительнотканый слой артериол, обеспечивающий механическую поддержку и защиту.
- Просвет венул (Venule Lumen, VL): Центральное пространство венул — мелких сосудов, собирающих дезоксигенированную кровь из капилляров и возвращающих ее в вены.

- Стенка венул (Venule Wall, VW): Тканевый слой венул, включающий гладкие мышцы и соединительную ткань, регулирующий кровоток и поддерживающий целостность сосуда.
- Просвет капилляров (Capillary Lumen, CL): Полое пространство внутри капилляров — мельчайших сосудов, где происходит обмен кислородом и питательными веществами между кровью и тканями.
- Стенка капилляров (Capillary Wall, CW): Тонкий слой эндотелиальных клеток, формирующий капилляры и обеспечивающий обмен газами, нутриентами и продуктами метаболизма.
- Иммунные клетки (Immune Cells, IC): Различные типы клеток иммунной системы (макрофаги, лимфоциты, нейтрофилы), защищающие от патогенов и удаляющие клеточный дебрис.
- Нервные стволы (Nerve Trunks, NT): Пучки нервных волокон, окруженные соединительной тканью, передающие нервные импульсы к различным органам и обратно.

Все указанные структуры одинаково важны для корректной регенерации. Микрососуды различных типов снабжают регенерирующую сосудистую стенку кислородом и питательными веществами, удаляя метаболические отходы. Нервные стволы обеспечивают иннервацию и симпатическую регуляцию, адаптируя артериальное давление через вазоконстрикцию и вазорелаксацию. Кластеры иммунных клеток поддерживают локальный иммунитет и участвуют в деградации биodeградируемого полимера, из которого состоит сосудистый трансплантат.

Аналогично предыдущей задаче для формирования сегментационных карт изображений использована классификационная модель Xception. Конечная точность модели по метрике качества Ассигасу составила 89,93%. Результат работы метода создания сегментационных карт представлен на рисунке 28. Предлагаемый в работе метод позволил сформировать предварительную версию тренировочного множества семантической сегментации всего за 11 дней.

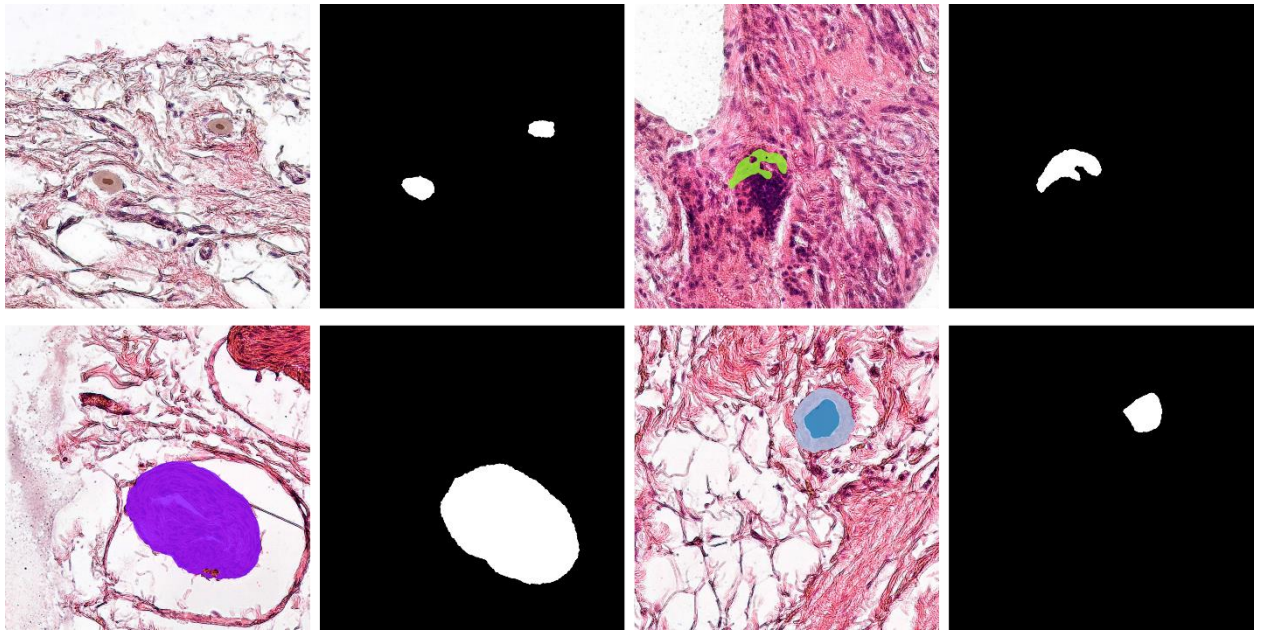


Рисунок 26 – Результат анализа трехканальной маски для классов: Arteriole Lumen, Arteriole Adventitia, Venule Lumen, Capillary Lumen, Immune Cells

Для сегментации девяти упомянутых гистологических структур мы тщательно отобрали и оптимизировали 6 различных сегментационных архитектур: U-Net, LinkNet, FPN, PSPNet, DeepLabV3, MA-Net. Важно подчеркнуть, что достижение окончательных конфигураций и параметров обучения для этих моделей потребовало тщательного процесса подбора гиперпараметров. Каждая архитектура прошла 200 испытаний различных конфигураций в ходе оптимизации гиперпараметров для обеспечения максимальной производительности.

Этап настройки гиперпараметров выявил, что U-Net и FPN достигли наивысших значений DSC, что делает их наиболее подходящими для сегментации гистологических данных. Также стоит отметить стабильность MA-Net и высокую вычислительную нагрузку DeepLabV3, подробнее ниже (детальная информация представлена в таблице 10):

- **Длительность:** Время настройки значительно варьировалось: FPN требовала больше всего времени (335 часов), а PSPNet — меньше всего (187 часов). Количество сбоев во время настройки было низким, при этом MA-Net

показала полную стабильность (0 сбоев). Сбои обычно возникали из-за конфигураций, требующих больше GPU-памяти, чем доступно.

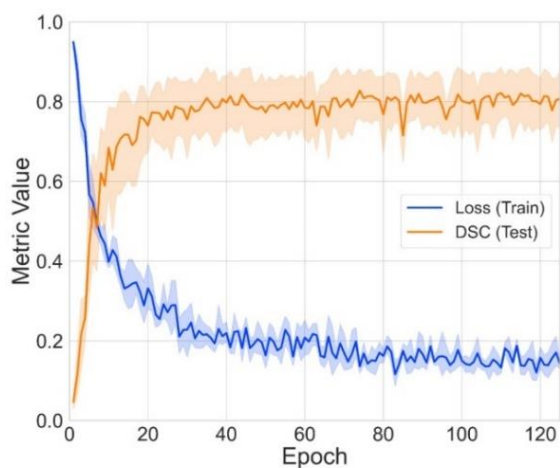
- **Энкодер:** наиболее часто использовался SE-ResNet-50. Однако для FPN и DeepLabV3/MA-Net применялись EfficientNet B7 и варианты RegNet соответственно.
- **Размер входных данных:** размеры входных данных адаптировались под каждую модель (от 512×512 до 896×896 пикселей), что отражает баланс между вычислительной эффективностью и детализацией, необходимой для точной сегментации.
- **Оптимизатор:** в большинстве случаев использовался RMSprop, за исключением U-Net (Adam) и PSPNet (RAdam).
- **Скорость обучения:** для всех моделей установлена единая скорость обучения 0.0001, что указывает на преимущество низких значений для сходимости задач сегментации.
- **Точность:** производительность оценивалась с помощью метрики DSC (коэффициент Dice) на валидационной выборке. DSC варьировался от 0.879 (PSPNet) до 0.906 (U-Net и FPN), что подтверждает наивысшую точность U-Net и FPN на этапе настройки.
- **Сложность:** количество параметров и FLOPs (операций с плавающей запятой в секунду) отражают вычислительную нагрузку. U-Net и LinkNet были схожи по этим показателям, тогда как MA-Net имела наибольшее число параметров (194.8 млн), а DeepLabV3 — наибольшие FLOPs (352.3 млрд), что делает его самой ресурсоемкой моделью.

Этап настройки гиперпараметров показал, что модели U-Net и FPN достигли наивысших значений DSC, что свидетельствует об их наибольшей пригодности для сегментации гистологических данных на валидационной выборке в рамках настройки. Кроме того, важными результатами стали стабильность MA-Net (отсутствие сбоев конфигураций) и высокая вычислительная нагрузка, характерная для DeepLabV3.

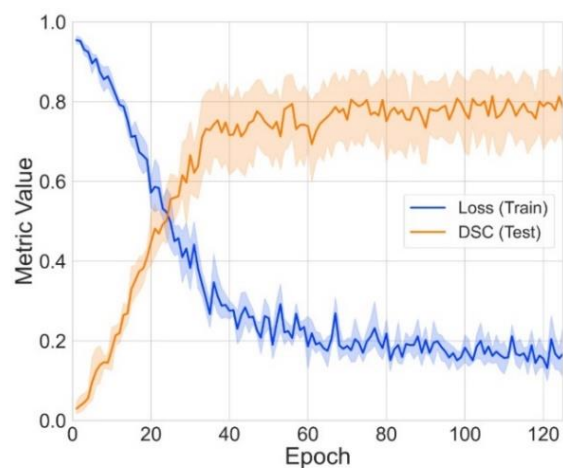
Таблица 10 – Оптимальные гиперпараметры для исследуемых сетей

Модель	Энкодер	Размер входного тензора	Алгоритм оптимизации	Скорость обучения
U-Net	SE-ResNet-50	896x896	Adam	0.0001
LinkNet	SE-ResNet-50	896x896	RMSprop	0.0001
FPN	EfficientNet B7	640x640	RMSprop	0.0001
PSPNet	SE-ResNet-50	640x640	RAdam	0.0001
DeepLabV3	RegNetX-6.4GF	768x768	RMSprop	0.0001
MA-Net	RegNetY-12GF	512x512	RMSprop	0.0001

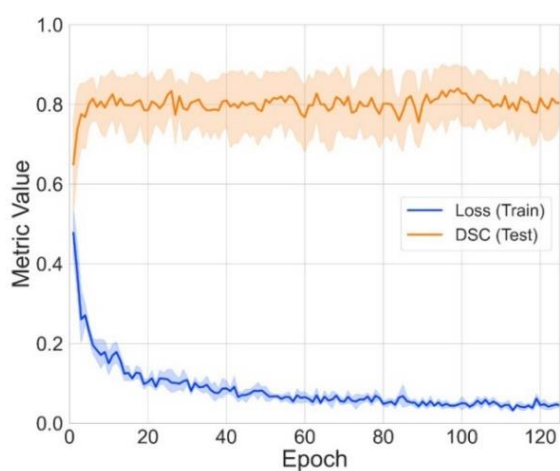
Последующий этап исследования включал детальное изучение эффективности и характеристик сходимости шести архитектур глубокого обучения: U-Net, LinkNet, FPN, PSPNet, DeepLabV3 и MA-Net. Комплексный анализ продолжительностью 125 эпох позволил проследить динамику потерь и коэффициента DSC для каждой нейросетевой модели в течение требуемого времени (см. рисунок 29). Полученные данные продемонстрировали консистентную тенденцию для всех исследуемых архитектур: прогрессивное уменьшение значений loss-функции и соответствующий рост показателя DSC в ходе тренировочного процесса. Указанные закономерности подтверждают способность моделей к адаптации и повышению точности сегментации в процессе обучения. Вместе с тем следует подчеркнуть наличие различий в темпах и устойчивости процесса конвергенции у различных архитектур.



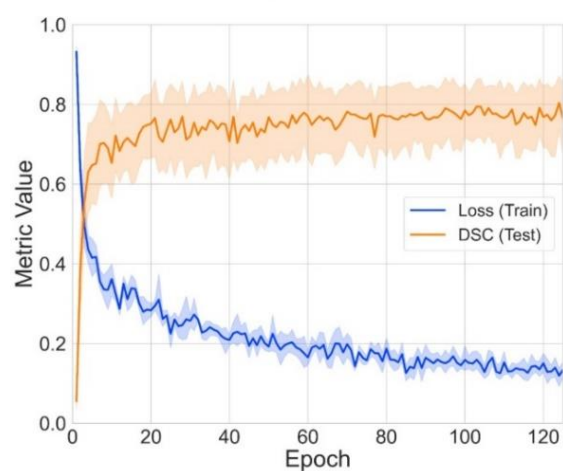
(a) U-Net



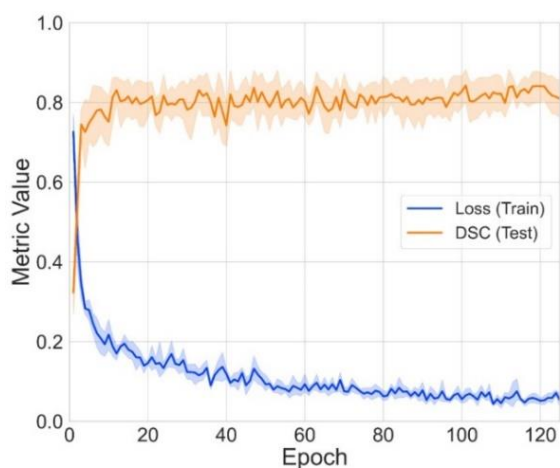
(b) LinkNet



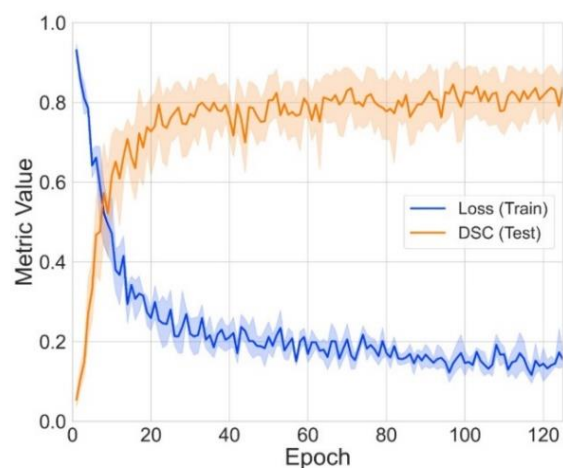
(c) FPN



(d) PSPNet



(e) DeepLabV3



(f) MA-Net

Рисунок 27 – Динамика изменения функции потерь и коэффициента DSC в процессе обучения и тестирования моделей

Результаты данного исследования предоставляют важные сведения о возможностях и ограничениях изученных моделей (см. таблицу 11, Рисунок 30 и приложение Г). Например, модель DeepLabV3 продемонстрировала стабильно высокую производительность для большинства анатомических структур, достигнув особенно высоких показателей в сегментации просвета и стенок венул, а также просвета и стенок капилляров. Это подчеркивает ее эффективность в точном определении венозных и капиллярных структур. Модель FPN, напротив, показала исключительную точность в сегментации иммунных клеток и нервных стволов, что указывает на ее способность анализировать сложные детали изображений. Архитектуры LinkNet и PSPNet показали неустойчивое качество работы. PSPNet характеризовалась наименьшим усредненным показателем DSC на уровне 0.823, демонстрируя затруднения при идентификации просвета венул и капилляров, а также их стеночных структур. Данные показатели указывают на существенные трудности в корректной локализации указанных анатомических образований.

Архитектура MA-Net оказалась лидером, показав наивысший средний DSC – 0.875. Она продемонстрировала высокую точность в сегментации просвета артериол и их адвентиции, что указывает на ее потенциал для задач сосудистой визуализации. Наконец, модель U-Net подтвердила свою универсальность, достигнув выдающихся результатов в сегментации медиа артериол.

Таблица 11 – Коэффициенты сходства по признакам и средние коэффициенты сходства по кубикам для исследуемых моделей

Модель	AL	AM	AA	VL	VW	CL	CW	IC	NT	Mean
U-Net	0.931	0.907	0.820	0.797	0.766	0.801	0.783	0.920	0.966	0.855
LinkNet	0.898	0.881	0.825	0.799	0.773	0.778	0.774	0.935	0.925	0.843
FPN	0.919	0.904	0.805	0.852	0.800	0.756	0.755	0.955	0.981	0.859
PSPNet	0.872	0.838	0.830	0.784	0.734	0.728	0.722	0.937	0.959	0.823
DeepLabV3	0.872	0.861	0.803	0.900	0.861	0.815	0.793	0.895	0.975	0.864
MA-Net	0.939	0.893	0.860	0.848	0.830	0.806	0.787	0.937	0.978	0.875

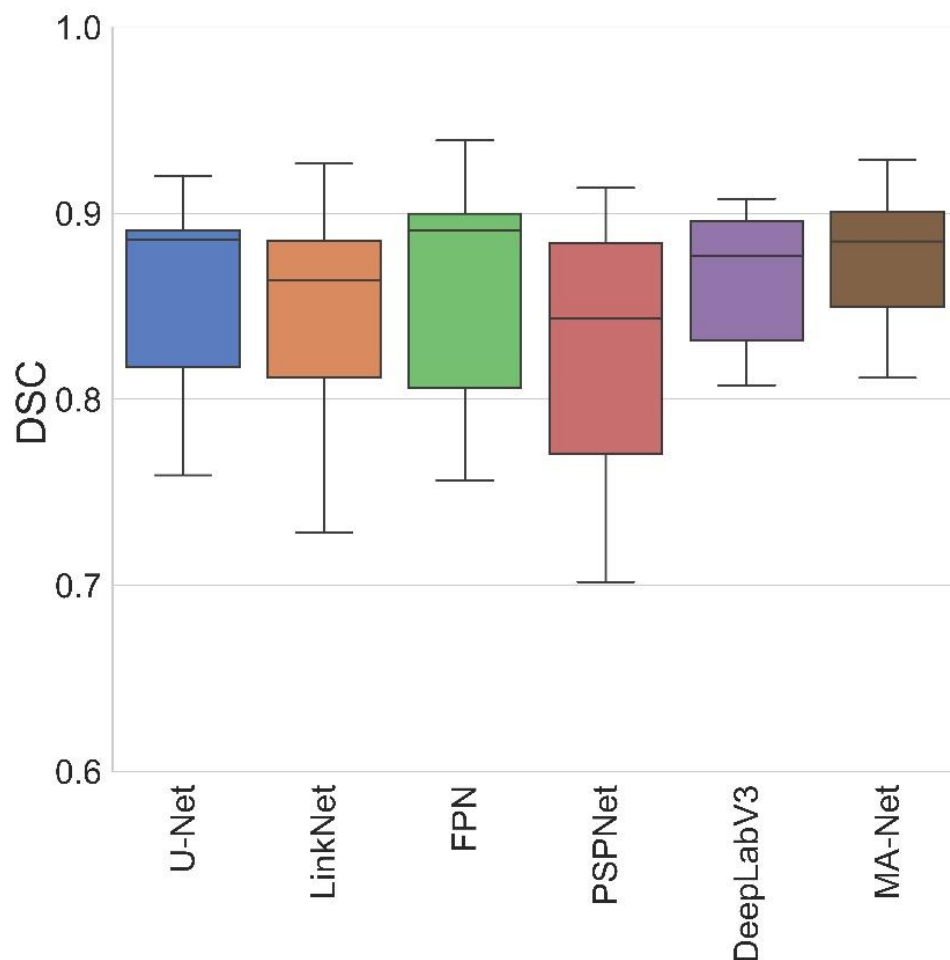


Рисунок 28 – Средняя производительность по функциям для различных моделей

Для визуализации предсказаний нейросетей на рисунке 31 представлены три фрагмента, демонстрирующих сегментацию изученных гистологических структур. Данные результаты получены с использованием оптимального решения — ансамбля из трех моделей (MA-Net, DeepLabV3 и FPN). Подробная визуализация всех моделей, использованных в исследовании, доступна на рисунках Г2 и Г3 Приложения Д.

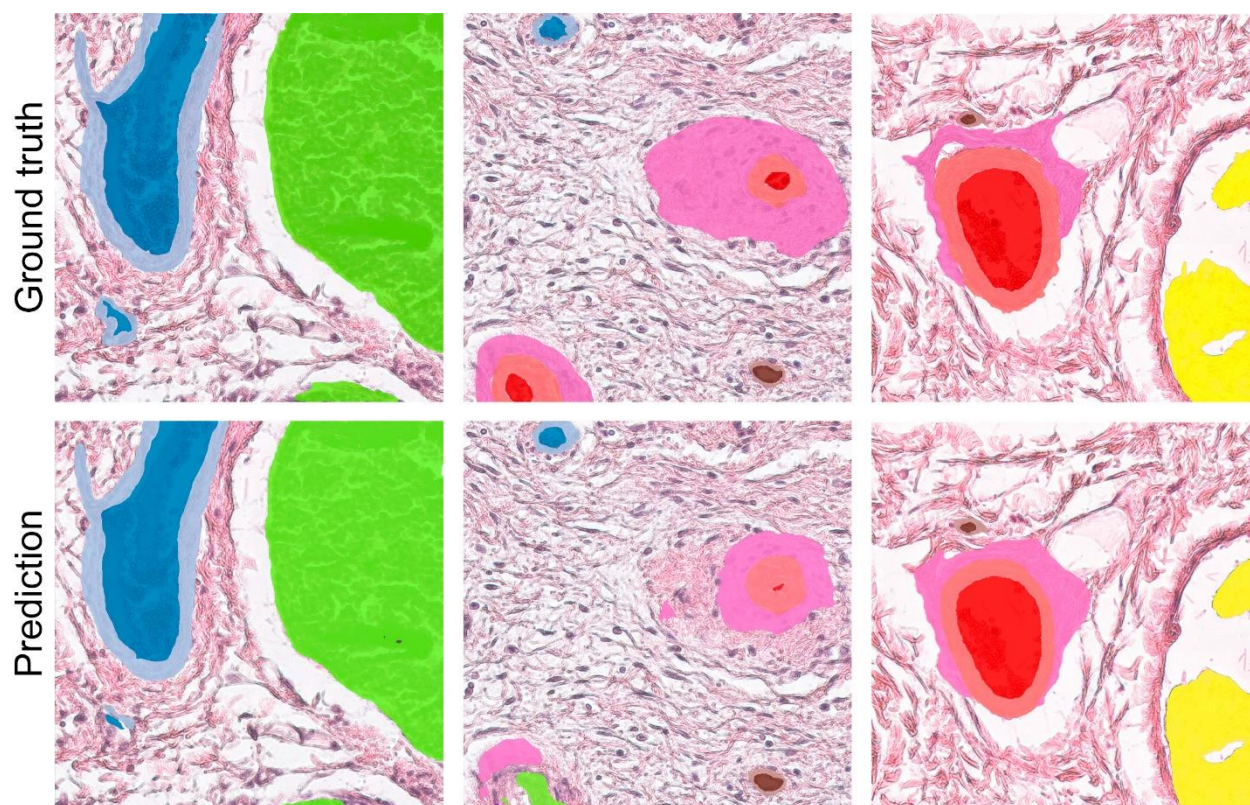


Рисунок 29 – Визуализация результатов работы модели в сравнение с экспертной разметкой

В результате зрительного и количественного анализа тестового множества видно, что все представленные в работе алгоритмы продемонстрировали свою эффективность и оказали положительное влияние на конечный результат.

3.4 Выводы по третьей главе

1. Предложенная технология формирования сегментационных масок изображений продемонстрировал положительные результаты: время подготовки

данных для обучения моделей семантической сегментации было сокращено более чем в 3 раза.

2. Предложенный метод автоматизированного выделения областей интереса на изображениях по картам признаков классификационной модели посредством интерпретации пространственной информации, извлеченной из промежуточных слоев сверточной нейросети, обеспечивает корректную сегментацию для объектов с неявными характеристическими признаками.

3. Предложенный алгоритм автоматизированного формирования адаптивной трехканальной сегментационной карты, основанный на разделении изображения на зоны объекта, фона и неопределенности с использованием морфологических операций, который обеспечивает высокую точность локализации границ сложных объектов исследования, отличается от аналогов повышенной производительностью и адаптивностью к размерам анализируемого изображения, а также повышением точности сегментации границ до 23,7% по сравнению с традиционными методами.

4. Предложенная модификация алгоритма оптимизации гиперпараметров позволила в рекордные сроки провести подбор оптимальных параметров нейросетевых архитектур для решения задачи семантической сегментации сложных объектов исследования на медицинских данных. Учет опыта прошлых запусков позволил с максимальной эффективностью провести подбор менее чем в 300 итерациях.

4. Апробация полученных результатов в задаче сегментации атеросклеротических бляшек и сегментации различных анатомических структур, таких как просвет артериолы, средняя оболочка артериолы, адвентиция артериолы, просвет венулы, стенка венулы, просвет капилляра, стенка капилляра, иммунные клетки и нервные стволы подтверждает результативность исследования. По каждой представленной задаче были достигнуты удовлетворительные показатели качества.

ЗАКЛЮЧЕНИЕ

Диссертационная работа посвящена разработке технологии автоматизированной подготовки данных семантической сегментации, включающей в себя: метод и алгоритмы сегментации объектов с отсутствием явных характеристических признаков на медицинских изображениях. Актуальность исследования в разработке эффективной технологии подготовки сегментационных масок медицинских изображений, основанной на методе анализа карты признаков классификационной модели для генерации исходной маски, алгоритме формирования трехканальной сегментационной карты с последующей оптимизацией границ объектов посредством FBA-Net модели, что обеспечивает повышение точности сегментации за счет минимизации артефактов и улучшения локализации слабовыраженных областей, а также оптимизирует процесс создания обучающих данных для построения высокоточной модели сегментации объектов с неявными характеристическими признаками в условиях ограниченной априорной информации обусловлена возрастающей потребностью в автоматизированных системах анализа медицинских данных, способных эффективно обрабатывать сложные изображения с низкой контрастностью, зашумленностью и отсутствием четких границ объектов. Решение этой задачи имеет значительное практическое значение для таких областей, как кардиология, онкология и гистология, где точность сегментации напрямую влияет на качество диагностики и лечения.

Основной целью работы является разработка эффективной технологии подготовки сегментационных масок медицинских изображений, основанной на анализе карт признаков классификационной модели с последующей оптимизацией границ объектов. Для достижения этой цели были решены следующие исследовательские задачи:

1. Проведен анализ предметной области и существующих методов сегментации, методов передачи и обработки информации, алгоритмов и

способов подготовки данных медицинских изображений. Систематизированы современные подходы к сегментации, включающие традиционные методы и архитектуры глубокого обучения. Выявлены ключевые ограничения: трудозатратность создания аннотаций, дисбаланс классов, нечеткие границы объектов. Обоснована необходимость разработки автоматизированных методов подготовки сегментационных масок для объектов с неявными характеристическими признаками.

2. Предложена технология предварительной подготовки аннотаций медицинских данных, сокращающая время подготовки тренировочного множества более чем в 3 раза. Предложенный новый метод создания сегментационных масок изображений, основанный на анализе карт признаков классификационной модели и последующей оптимизации границ с использованием FVA-модели, отличается от существующих аналогов высокой адаптивностью и точностью, особенно в условиях ограниченной априорной информации. Важной особенностью отмечается: разработка алгоритма формирования трехканальных сегментационных карт, который обеспечивает дополнительный анализ объекта локализации, что способствует повышению точности сегментации ~ на 23,7%.
3. Предложен метод автоматизированного выделения областей интереса на изображениях по картам признаков классификационной модели посредством интерпретации пространственной информации, извлеченной из промежуточных слоев сверточной нейросети, обеспечивающий корректную сегментацию для объектов с неявными характеристическими признаками.
4. Предложен алгоритм автоматизированного формирования адаптивной трехканальной сегментационной карты, основанный на разделении изображения на зоны объекта, фона и неопределенности с использованием

морфологических операций, который обеспечивает высокую точность локализации границ сложных объектов исследования, отличаясь от аналогов повышенной производительностью и адаптивностью к размерам анализируемого изображения, а также повышая точность сегментации границ до 23,7% по сравнению с традиционными методами.

5. Реализован алгоритм подбора оптимальных гиперпараметров для обучения нейронных сетей. Данный алгоритм основан на комбинации методов Hyperband и байесовской оптимизации, что позволило сократить время подбора оптимальных параметров нейросетевых моделей, за счет внедрения механизма учета накопленных знаний и достигнуть конечной точности модели семантической сегментации на 2,7% выше классического.
6. Реализован программный комплекс для автоматизированной подготовки сегментационных масок медицинских изображений, который интегрирует предложенный метод анализа карт признаков, алгоритм формирования трехканальных карт и оптимизации границ с использованием FBA-Net модели, обеспечивая сокращение времени подготовки обучающих данных более чем в 3 раза по сравнению с ручным аннотированием.

Практическая значимость работы подтверждается внедрением разработанного программного комплекса в реальные медицинские проекты, такие как автоматизированная диагностика сердечно-сосудистых заболеваний и выявление нестабильных бляшек на основе оптической когерентной томографии. Предложенные алгоритмы позволяют сократить время подготовки обучающих данных более чем в 3 раза и повысить точность сегментации ~ на 2,7%, что особенно важно для задач, требующих оперативного принятия решений, например, при хирургических вмешательствах. Результаты работы также могут быть использованы в образовательных программах для подготовки специалистов в области медицинской визуализации.

Экспериментальная часть работы подтверждается апробацией предложенного метода на двух типах медицинских данных: изображениях атеросклеротических бляшек и гистологических срезах. В ходе экспериментов было продемонстрировано, что метод обеспечивает подготовку данных для обучения моделей высокой точности сегментации для объектов с выраженными границами, таких как просвет сосудов (Lumen) и vasa vasorum. Для объектов с нечеткими границами, таких как фиброзная капсула (Fibrous cap) и липидное ядро (Lipid core), метод показал удовлетворительные результаты, однако требует дальнейшей доработки для повышения точности. Важным достижением является сокращение времени подготовки данных более чем в три раза по сравнению с ручной разметкой, что подтверждает эффективность предложенного подхода. Предложенные алгоритмы демонстрируют устойчивость к зашумленности и вариативности медицинских изображений, что делает их применимыми в реальных клинических условиях.

Перспективы дальнейших исследований связаны с совершенствованием метода для работы с объектами, не имеющими четких границ, а также с расширением области применения предложенных алгоритмов на другие типы медицинских данных. Результаты работы открывают новые возможности для автоматизации процессов диагностики и лечения, способствуя развитию персонализированной медицины и улучшению качества медицинской помощи.

СПИСОК СОКРАЩЕНИЙ И ОПРЕДЕЛЕНИ

Сегментация	– процесс анализа и разделения цифрового изображения на структурно или семантически однородные регионы (сегменты) путем присвоения каждому пикселю или группе пикселей уникальной категориальной метки
Сложные объекты исследования	– категория объектов, которые обладают рядом визуальных характеристик, существенно затрудняющих их обнаружение, локализацию и классификацию. Их сложность обусловлена низкой различимостью на фоне окружения, отсутствием четких границ, минимальным размером и/или неочевидными семантическими признаками
Гиперпараметры	– набор внешних конфигурационных параметров модели, которые определяются до начала процесса обучения и остаются неизменными в ходе оптимизации весов модели
Trimap	– трехканальная карта объекта
CNN	– Convolutional Neural Network (сверточная нейронная сеть)
IoU	– Intersection over Union. Метрика в обнаружении объектов, для оценки пересечения масок, которая измеряет, насколько хорошо предсказанное ограничительное поле перекрывается с истинным ограничительным полем
MRT	– магнитно-резонансная томография
КТ	– компьютерная томография

ОКТ	– оптическая когерентная томография
ИИ	– искусственный интеллект
ВОНВ	– ВОНВ: Robust and Efficient Hyperparameter Optimization at Scale. Передовой алгоритм оптимизации гиперпараметров
FBA-Net	– Feature Boundary-Aware Network – глубокая нейронная сеть, разработанная для задач сегментации медицинских изображений с акцентом на точное определение границ объектов
ИНС	– искусственная нейронная сеть
DSC	– the Dice similarity coefficient. Коэффициент сходства Дайса – безразмерный показатель сходства сравниваемых объектов

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- [1] E. Hjelmås and B. K. Low, “Face Detection: A Survey,” *Computer Vision and Image Understanding*, vol. 83, no. 3, pp. 236–274, Sep. 2001, doi: 10.1006/CVIU.2001.0921.
- [2] Б. Ю. С. ТИХОНОВА Т.С, “Основные подходы к отслеживанию и распознаванию лица.” Accessed: Apr. 12, 2025. [Online]. Available: <https://www.elibrary.ru/item.asp?id=26847529>
- [3] Д. А. Алексеевич, “Система поиска, выделения и распознавания лиц на изображениях,” *Известия Томского политехнического университета. Инжиниринг георесурсов*, vol. 318, no. 5, 2011, Accessed: Apr. 12, 2025. [Online]. Available: <https://cyberleninka.ru/article/n/sistema-poiska-vydeleniya-i-raspoznavaniya-lits-na-izobrazheniyah>
- [4] B. Coifman, D. Beymer, P. McLauchlan, and J. Malik, “A real-time computer vision system for vehicle tracking and traffic surveillance,” *Transp Res Part C Emerg Technol*, vol. 6, no. 4, pp. 271–288, Aug. 1998, doi: 10.1016/S0968-090X(98)00019-9.
- [5] В. В. Лаптев, Н. В. Лаптев, А. Х. Оздиев, and О. М. Гергет, “Анализ парковочного пространства посредством компьютерного зрения”.
- [6] A. Asokan, J. Anitha, M. Ciobanu, A. Gabor, A. Naaji, and D. J. Hemanth, “Image Processing Techniques for Analysis of Satellite Images for Historical Maps Classification—An Overview,” *Applied Sciences 2020, Vol. 10, Page 4207*, vol. 10, no. 12, p. 4207, Jun. 2020, doi: 10.3390/APP10124207.
- [7] С. Д.Н., “Методы и алгоритмы определения положения и ориентации беспилотного летательного аппарата с применением бортовых видеокамер,”

- Программные продукты и системы*, no. 1 (105), 2014, Accessed: Apr. 12, 2025. [Online]. Available: <https://cyberleninka.ru/article/n/metody-i-algoritmy-opredeleniya-polozheniya-i-orientatsii-bespilotnogo-letatel'nogo-apparata-s-primeneniem-bortovyh-videokamer>
- [8] K. V. Vik, A. A. Druki, D. S. Grigoriev, and V. G. Spitsyn, “Application of deep learning neural networks for solving the problem of forest fire segmentation on satellite images,” *Vestnik Tomskogo Gosudarstvennogo Universiteta - Upravlenie, Vychislitel'naya Tekhnika i Informatika*, no. 55, pp. 18–25, 2021, doi: 10.17223/19988605/55/3.
- [9] D. Shen, G. Wu, and H. Il Suk, “Deep Learning in Medical Image Analysis,” *Annu Rev Biomed Eng*, vol. 19, no. Volume 19, 2017, pp. 221–248, Jun. 2017, doi: 10.1146/ANNUREV-BIOENG-071516-044442/CITE/REFWORKS.
- [10] V. V. Danilov *et al.*, “Real-time coronary artery stenosis detection based on modern neural networks,” *Scientific Reports 2021 11:1*, vol. 11, no. 1, pp. 1–13, Apr. 2021, doi: 10.1038/s41598-021-87174-2.
- [11] V. V. Danilov *et al.*, “Automatic scoring of COVID-19 severity in X-ray imaging based on a novel deep learning workflow,” *Scientific Reports 2022 12:1*, vol. 12, no. 1, pp. 1–22, Jul. 2022, doi: 10.1038/s41598-022-15013-z.
- [12] В. В. Лаптев and О. М. Гергет, “Применение средств машинного обучения для формирования сегментационных масок,” *сборник трудов XX Международной научно-практической конференции студентов, аспирантов и молодых учёных*, pp. 40–41, Mar. 2023, Accessed: May 31, 2025. [Online]. Available: <http://earchive.tpu.ru/jspui/handle/11683/77941>
- [13] I. Ahmed, M. Ahmad, M. Nawaz, K. Haseeb, S. Khan, and G. Jeon, “Efficient topview person detector using point based transformation and lookup table,”

- Comput Commun*, vol. 147, pp. 188–197, Nov. 2019, doi: 10.1016/j.comcom.2019.08.015.
- [14] D. T. Nguyen, W. Li, and P. O. Ogunbona, “Human detection from images and videos: A survey,” *Pattern Recognit*, vol. 51, pp. 148–175, Mar. 2016, doi: 10.1016/J.PATCOG.2015.08.027.
- [15] C. Li, D. Song, R. Tong, and M. Tang, “Illumination-aware faster R-CNN for robust multispectral pedestrian detection,” *Pattern Recognit*, vol. 85, pp. 161–171, Jan. 2019, doi: 10.1016/j.patcog.2018.08.005.
- [16] B. Benjdira, T. Khursheed, A. Koubaa, A. Ammar, and K. Ouni, “Car Detection using Unmanned Aerial Vehicles: Comparison between Faster R-CNN and YOLOv3,” *2019 1st International Conference on Unmanned Vehicle Systems-Oman, UVS 2019*, Mar. 2019, doi: 10.1109/UVS.2019.8658300.
- [17] T. Ross *et al.*, “Robust Medical Instrument Segmentation Challenge 2019,” Mar. 2020, Accessed: Apr. 12, 2025. [Online]. Available: <https://arxiv.org/abs/2003.10299v2>
- [18] V. V. Danilov *et al.*, “Aortography keypoint tracking for transcatheter aortic valve implantation based on multi-task learning,” *Front Cardiovasc Med*, vol. 8, p. 697737, Jul. 2021, doi: 10.3389/FCVM.2021.697737/BIBTEX.
- [19] B. van Ginneken, M. B. Stegmann, and M. Loog, “Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database,” *Med Image Anal*, vol. 10, no. 1, pp. 19–40, Feb. 2006, doi: 10.1016/J.MEDIA.2005.02.002.
- [20] J. Liu, H. Su, and W. L. Nowinski, “A hybrid approach for segmentation of anatomic structures in medical images,” *Int J Comput Assist Radiol Surg*, vol. 3, no. 3–4, pp. 213–219, Jun. 2008, doi: 10.1007/S11548-008-0229-2/METRICS.

- [21] L. Pícek, K. Janouskova, M. Sulc, and J. Matas, “FungiTastic: A multi-modal dataset and benchmark for image categorization,” Aug. 2024, Accessed: Apr. 12, 2025. [Online]. Available: <https://arxiv.org/abs/2408.13632v2>
- [22] B. Lee and K. Paeng, “A robust and effective approach towards accurate metastasis detection and pN-stage classification in breast cancer,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11071 LNCS, pp. 841–850, 2018, doi: 10.1007/978-3-030-00934-2_93/TABLES/8.
- [23] A. F. Frangi, J. A. Schnabel, C. Davatzikos, C. Alberola-López, and G. Fichtinger, Eds., “Medical Image Computing and Computer Assisted Intervention – MICCAI 2018,” vol. 11071, 2018, doi: 10.1007/978-3-030-00934-2.
- [24] H. Liebl *et al.*, “A computed tomography vertebral segmentation dataset with anatomical variations and multi-vendor scanner data,” *Scientific Data 2021 8:1*, vol. 8, no. 1, pp. 1–7, Oct. 2021, doi: 10.1038/s41597-021-01060-0.
- [25] C. Wang, R. Zhang, and L. Chang, “A Study on the Dynamic Effects and Ecological Stress of Eco-Environment in the Headwaters of the Yangtze River Based on Improved DeepLab V3+ Network,” *Remote Sensing 2022, Vol. 14, Page 2225*, vol. 14, no. 9, p. 2225, May 2022, doi: 10.3390/RS14092225.
- [26] S. C. Horii, W. D. Bidgood, S. C. Horii, and W. D. Bidgood, “DICOM: a standard for medical imaging,” *SPIE*, vol. 1785, pp. 87–102, Jan. 1993, doi: 10.1117/12.139249.
- [27] D. A. . Clunie, “DICOM structured reporting,” p. 394, 2000.
- [28] O. S. . Panykh, “Digital imaging and communications in medicine (DICOM) : a practical introduction and survival guide,” p. 417, 2012.

- [29] W. D. Bidgood, S. C. Horii, F. W. Prior, and D. E. Van Syckle, “Understanding and Using DICOM, the Data Interchange Standard for Biomedical Imaging,” *Journal of the American Medical Informatics Association*, vol. 4, no. 3, pp. 199–212, 1997, doi: 10.1136/JAMIA.1997.0040199,.
- [30] S. Saxena and S. Paul, “High-Performance Medical Image Processing,” *High-Performance Medical Image Processing*, May 2022, doi: 10.1201/9781003190011.
- [31] P. A. Yushkevich *et al.*, “User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability,” *Neuroimage*, vol. 31, no. 3, pp. 1116–1128, Jul. 2006, doi: 10.1016/J.NEUROIMAGE.2006.01.015.
- [32] С. И. П. Г. О. М. ДАНИЛОВ В.В, “Метод поиска эффективной последовательности алгоритмов обработки медицинских данных при сегментации анатомических структур сердца.” Accessed: Apr. 12, 2025. [Online]. Available: <https://www.elibrary.ru/item.asp?id=35157500>
- [33] U. Özkaya, Ş. Öztürk, and M. Barstugan, “Coronavirus (COVID-19) Classification Using Deep Features Fusion and Ranking Technique,” *Studies in Big Data*, vol. 78, pp. 281–295, 2020, doi: 10.1007/978-3-030-55258-9_17.
- [34] A. Narin, C. Kaya, and Z. Pamuk, “Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks,” *Pattern Analysis and Applications*, vol. 24, no. 3, pp. 1207–1220, Aug. 2021, doi: 10.1007/s10044-021-00984-y.
- [35] Z. Tang *et al.*, “Severity Assessment of Coronavirus Disease 2019 (COVID-19) Using Quantitative Features from Chest CT Images,” Mar. 2020, Accessed: Apr. 12, 2025. [Online]. Available: <http://arxiv.org/abs/2003.11988>

- [36] R. Buizza, “Probabilistic prediction of COVID-19 infections for China and Italy, using an ensemble of stochastically-perturbed logistic curves,” Mar. 2020, Accessed: Apr. 12, 2025. [Online]. Available: <http://arxiv.org/abs/2003.06418>
- [37] P. Bilic *et al.*, “The Liver Tumor Segmentation Benchmark (LiTS),” *Med Image Anal*, vol. 84, p. 102680, Feb. 2023, doi: 10.1016/J.MEDIA.2022.102680.
- [38] A. Tremeau and N. Borel, “A region growing and merging algorithm to color segmentation,” *Pattern Recognit*, vol. 30, no. 7, pp. 1191–1203, Jul. 1997, doi: 10.1016/S0031-3203(96)00147-1.
- [39] A. M. Dostovalova *et al.*, “Сравнительный анализ модификаций нейросетевых архитектур U-Net в задаче сегментации медицинских изображений,” *Digital Diagnostics*, vol. 5, no. 4, pp. 833–853, Dec. 2024, doi: 10.17816/DD629866.
- [40] M. BachCuadra, V. Duay, and J. P. H. Thiran, “Atlas-based Segmentation,” *Handbook of Biomedical Imaging: Methodologies and Clinical Research*, pp. 221–244, Jan. 2015, doi: 10.1007/978-0-387-09749-7_12.
- [41] “ISSN 0485-8972 Радиотехника. 2016. Вып. 185 87 УДК 004.932:611.018.1 А.В. РАБОТЯГОВ, канд. те.” Accessed: Apr. 12, 2025. [Online]. Available: https://scholar.googleusercontent.com/scholar?q=cache:3TNxkWIsL30J:scholar.google.com/&hl=ru&as_sdt=0,5
- [42] П. В. Гаврилов *et al.*, “Нейросетевая модель детекции признаков поражения легких, ассоциированных с COVID-19, на аксиальных срезах нативной компьютерной томографии грудной клетки,” *Медицинский Альянс*, vol. 8, no. 2, Jul. 2020, doi: 10.36422/23076348-2020-8-2-6-13.
- [43] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, “Feature Pyramid Networks for Object Detection,” in *2017 IEEE Conference on Computer*

- Vision and Pattern Recognition (CVPR)*, IEEE, Jul. 2017, pp. 936–944. doi: 10.1109/CVPR.2017.106.
- [44] A. V. Zakharov *et al.*, “Автоматическая сегментация очагов демиелинизации при рассеянном склерозе,” *Наука и инновации в медицине*, vol. 9, no. 4, pp. 284–290, Dec. 2024, doi: 10.35693/SIM636947.
- [45] D. Reis, J. Kupec, J. Hong, and A. Daoudi, “Real-Time Flying Object Detection with YOLOv8,” May 2023, Accessed: Apr. 13, 2025. [Online]. Available: <https://arxiv.org/abs/2305.09972v2>
- [46] “Futuristic Trends in Network and Communication Technologies: Third ... - Google Книги.” Accessed: Apr. 29, 2025. [Online]. Available: https://books.google.ru/books?hl=ru&lr=&id=d7AmEAAAQBAJ&oi=fnd&pg=PR5&dq=Vladimir+Viktorovich+Kureychik+covid&ots=aqUPt9nMJ8&sig=ugkrbKrqeBA8otlyOiadlxRW9VI&redir_esc=y#v=onepage&q&f=false
- [47] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-December, pp. 770–778, Dec. 2015, doi: 10.1109/CVPR.2016.90.
- [48] M. Tan and Q. V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” *36th International Conference on Machine Learning, ICML 2019*, vol. 2019-June, pp. 10691–10700, May 2019, Accessed: Apr. 13, 2025. [Online]. Available: <https://arxiv.org/abs/1905.11946v5>
- [49] A. Dosovitskiy *et al.*, “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,” *ICLR 2021 - 9th International Conference on Learning Representations*, Oct. 2020, Accessed: Apr. 13, 2025. [Online]. Available: <https://arxiv.org/abs/2010.11929v2>

- [50] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-December, pp. 770–778, Dec. 2015, doi: 10.1109/CVPR.2016.90.
- [51] K. Yang, K. Qinami, L. Fei-Fei, J. Deng, and O. Russakovsky, “Towards Fairer Datasets: Filtering and Balancing the Distribution of the People Subtree in the ImageNet Hierarchy,” *FAT* 2020 - Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pp. 547–558, Dec. 2019, doi: 10.1145/3351095.3375709.
- [52] X. Chen *et al.*, “Symbolic Discovery of Optimization Algorithms,” Feb. 2023, Accessed: Apr. 13, 2025. [Online]. Available: <https://arxiv.org/abs/2302.06675v4>
- [53] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science*, Springer, 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.
- [54] K. He, G. Gkioxari, P. Dollar, and R. Girshick, “Mask R-CNN,” 2017.
- [55] L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, “Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11211 LNCS, pp. 833–851, Feb. 2018, doi: 10.1007/978-3-030-01234-2_49.
- [56] M. C. de Verdier *et al.*, “The 2024 Brain Tumor Segmentation (BraTS) Challenge: Glioma Segmentation on Post-treatment MRI,” *Eugenio Iglesias*, vol. 33, May 2024, Accessed: Apr. 13, 2025. [Online]. Available: <https://arxiv.org/abs/2405.18368v1>

- [57] F. Isensee *et al.*, “nnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation,” *Informatik aktuell*, p. 22, Sep. 2018, doi: 10.1007/978-3-658-25326-4_7.
- [58] B. Recht, R. Roelofs, L. Schmidt, and V. Shankar, “Do CIFAR-10 Classifiers Generalize to CIFAR-10?,” Jun. 2018, Accessed: Apr. 13, 2025. [Online]. Available: <https://arxiv.org/abs/1806.00451v1>
- [59] Р. Федерации and Е. Г. Камкин, “УТВЕРЖДАЮ Заместитель Министра здравоохранения”.
- [60] D. Wang, A. Khosla, R. Gargeya, H. Irshad, A. H. Beck, and B. Israel, “Deep Learning for Identifying Metastatic Breast Cancer”, Accessed: Apr. 13, 2025. [Online]. Available: <http://camelyon16.grand-challenge.org/>
- [61] Y. Liu *et al.*, “Detecting Cancer Metastases on Gigapixel Pathology Images,” Mar. 2017, Accessed: Apr. 13, 2025. [Online]. Available: <https://arxiv.org/abs/1703.02442v2>
- [62] L. Yang, Y. Zhang, J. Chen, S. Zhang, and D. Z. Chen, “Suggestive Annotation: A Deep Active Learning Framework for Biomedical Image Segmentation,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10435 LNCS, pp. 399–407, Jun. 2017, doi: 10.1007/978-3-319-66179-7_46.
- [63] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization,” in *2017 IEEE International Conference on Computer Vision (ICCV)*, IEEE, Oct. 2017, pp. 618–626. doi: 10.1109/ICCV.2017.74.

- [64] R. L. Draelos and L. Carin, “Use HiResCAM instead of Grad-CAM for faithful explanations of convolutional neural networks,” *ArXiv*, Nov. 2020, Accessed: May 04, 2024. [Online]. Available: <http://arxiv.org/abs/2011.08891>
- [65] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization,” *Int J Comput Vis*, vol. 128, no. 2, pp. 336–359, Oct. 2016, doi: 10.1007/s11263-019-01228-7.
- [66] A. Chattopadhyay, A. Sarkar, P. Howlader, and V. N. Balasubramanian, “Grad-CAM++: Generalized Gradient-Based Visual Explanations for Deep Convolutional Networks,” in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, IEEE, Mar. 2018, pp. 839–847. doi: 10.1109/WACV.2018.00097.
- [67] R. Fu, Q. Hu, X. Dong, Y. Guo, Y. Gao, and B. Li, “Axiom-based Grad-CAM: Towards Accurate Visualization and Explanation of CNNs,” *31st British Machine Vision Conference, BMVC 2020*, Aug. 2020, Accessed: May 04, 2024. [Online]. Available: <http://arxiv.org/abs/2008.02312>
- [68] M. B. Muhammad and M. Yeasin, “Eigen-CAM: Class Activation Map using Principal Components,” *Proceedings of the International Joint Conference on Neural Networks*, Aug. 2020, doi: 10.1109/IJCNN48605.2020.9206626.
- [69] A. Chaddad, Y. Hu, Y. Wu, B. Wen, and R. Kateb, “Generalizable and Explainable Deep Learning for Medical Image Computing: An Overview,” *Curr Opin Biomed Eng*, vol. 33, Mar. 2025, doi: 10.1016/j.cobme.2024.100567.
- [70] H. Wang *et al.*, “Score-CAM: Score-Weighted Visual Explanations for Convolutional Neural Networks,” *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, vol. 2020-June, pp. 111–119, Oct. 2019, doi: 10.1109/CVPRW50498.2020.00020.

- [71] J. Zheng *et al.*, “A Novel Defense Against Poisoning Attacks on Federated Learning: LayerCAM Augmented with Autoencoder,” Jun. 2024, Accessed: Apr. 13, 2025. [Online]. Available: <https://arxiv.org/abs/2406.02605v1>
- [72] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, “UNet++: A Nested U-Net Architecture for Medical Image Segmentation,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11045 LNCS, Springer Verlag, 2018, pp. 3–11. doi: 10.1007/978-3-030-00889-5_1.
- [73] A. Chaurasia and E. Culurciello, “LinkNet: Exploiting encoder representations for efficient semantic segmentation,” in *2017 IEEE Visual Communications and Image Processing (VCIP)*, IEEE, Dec. 2017, pp. 1–4. doi: 10.1109/VCIP.2017.8305148.
- [74] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, “Pyramid Scene Parsing Network,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jul. 2017, pp. 6230–6239. doi: 10.1109/CVPR.2017.660.
- [75] H. Li, P. Xiong, J. An, and L. Wang, “Pyramid Attention Network for Semantic Segmentation,” *ArXiv*, May 2018, [Online]. Available: <http://arxiv.org/abs/1805.10180>
- [76] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs,” *IEEE Trans Pattern Anal Mach Intell*, vol. 40, no. 4, pp. 834–848, Apr. 2018, doi: 10.1109/TPAMI.2017.2699184.
- [77] T. Fan, G. Wang, Y. Li, and H. Wang, “MA-Net: A Multi-Scale Attention Network for Liver and Tumor Segmentation,” *IEEE Access*, vol. 8, pp. 179656–179665, 2020, doi: 10.1109/ACCESS.2020.3025372.

- [78] J. Lee *et al.*, “Automated Segmentation of Microvessels in Intravascular OCT Images Using Deep Learning,” *Bioengineering*, vol. 9, no. 11, p. 648, Nov. 2022, doi: 10.3390/bioengineering9110648.
- [79] M. Chu *et al.*, “Artificial intelligence and optical coherence tomography for the automatic characterisation of human atherosclerotic plaques,” *EuroIntervention*, vol. 17, no. 1, pp. 41–50, May 2021, doi: 10.4244/EIJ-D-20-01355.
- [80] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs,” *IEEE Trans Pattern Anal Mach Intell*, vol. 40, no. 4, pp. 834–848, Apr. 2018, doi: 10.1109/TPAMI.2017.2699184.
- [81] J. Lee *et al.*, “Automated analysis of fibrous cap in intravascular optical coherence tomography images of coronary arteries,” *Sci Rep*, vol. 12, no. 1, p. 21454, Dec. 2022, doi: 10.1038/s41598-022-24884-1.
- [82] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, “Feature Pyramid Networks for Object Detection,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jul. 2017, pp. 936–944. doi: 10.1109/CVPR.2017.106.
- [83] B. Pu *et al.*, “MobileUNet-FPN: A Semantic Segmentation Model for Fetal Ultrasound Four-Chamber Segmentation in Edge Computing Environments,” *IEEE J Biomed Health Inform*, vol. 26, no. 11, pp. 5540–5550, Nov. 2022, doi: 10.1109/JBHI.2022.3182722.
- [84] V. V. Danilov *et al.*, “ML-driven segmentation of microvascular features during histological examination of tissue-engineered vascular grafts,” *Front Bioeng Biotechnol*, vol. 12, p. 1411680, Jun. 2024, doi: 10.3389/fbioe.2024.1411680.

- [85] A. Sulaiman *et al.*, “An intelligent LinkNet-34 model with EfficientNetB7 encoder for semantic segmentation of brain tumor,” *Sci Rep*, vol. 14, no. 1, p. 1345, Jan. 2024, doi: 10.1038/s41598-024-51472-2.
- [86] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, “Pyramid Scene Parsing Network,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jul. 2017, pp. 6230–6239. doi: 10.1109/CVPR.2017.660.
- [87] K. Li, Z. Wu, K.-C. Peng, J. Ernst, and Y. Fu, “Tell Me Where to Look: Guided Attention Inference Network,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2018, pp. 9215–9223. doi: 10.1109/CVPR.2018.00960.
- [88] L. Dong, W. Lu, X. Lu, X. Leng, J. Xiang, and C. Li, “Comparison of deep learning-based image segmentation methods for intravascular ultrasound on retrospective and large image cohort study,” *Biomed Eng Online*, vol. 22, no. 1, p. 111, Nov. 2023, doi: 10.1186/s12938-023-01171-2.
- [89] Z. Bulbul, F. Bitar, H. Tamim, Z. Issa, G. Siblini, and G. Di-Salvo, “Measurement of Left Ventricular Dimensions and Ejection Fraction in Neonates by Three-Dimensional Echocardiography: A Comparative Study Between Philips QLAB and TOMTEC Software—Are the Values Interchangeable?,” *Pediatr Cardiol*, vol. 42, no. 5, pp. 1111–1118, Jun. 2021, doi: 10.1007/S00246-021-02588-Z/METRICS.
- [90] Ó. Pera *et al.*, “Clinical Validation of Siemens’ Syngo.via Automatic Contouring System,” *Adv Radiat Oncol*, vol. 8, no. 3, p. 101177, May 2023, doi: 10.1016/J.ADRO.2023.101177.
- [91] G. Fang, X. Ma, M. Bi Mi, and X. Wang, “Isomorphic Pruning for Vision Models”.
- [92] D. Kotsougiani, C. A. Hundepool, L. F. Bulstra, D. M. Shin, A. Y. Shin, and A. T. Bishop, “The learning rate in three dimensional high definition video assisted

- microvascular anastomosis in a rat model,” *Journal of Plastic, Reconstructive & Aesthetic Surgery*, vol. 69, no. 11, pp. 1528–1536, Nov. 2016, doi: 10.1016/J.BJPS.2016.08.001.
- [93] L. Uscher-Pines and A. Mehrotra, “Analysis of teladoc use seems to indicate expanded access to care for patients without prior connection to a provider,” *Health Aff*, vol. 33, no. 2, pp. 258–264, Feb. 2014, doi: 10.1377/HLTHAFF.2013.0989.
- [94] “Cloud Healthcare API documentation | Google Cloud.” Accessed: Apr. 13, 2025. [Online]. Available: <https://cloud.google.com/healthcare-api/docs>
- [95] S. Yu, P. Wu, P. P. Liang, R. Salakhutdinov, and L. P. Morency, “PACS: A Dataset for Physical Audiovisual CommonSense Reasoning,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 13697 LNCS, pp. 292–309, Mar. 2022, doi: 10.1007/978-3-031-19836-6_17.
- [96] T. M. Tapera *et al.*, “FlywheelTools: Data Curation and Manipulation on the Flywheel Platform,” *Front Neuroinform*, vol. 15, p. 678403, Jun. 2021, doi: 10.3389/FNINF.2021.678403/FULL.
- [97] S. S. Virani *et al.*, “Heart Disease and Stroke Statistics - 2021 Update: A Report From the American Heart Association,” *Circulation*, vol. 143, no. 8, pp. E254–E743, Feb. 2021, doi: 10.1161/CIR.0000000000000950.
- [98] “Нестабильные атеросклеротические бляшки коронарных артерий у пациентов со стабильной ишемической болезнью сердца.” Accessed: Apr. 13, 2025. [Online]. Available: <https://cyberleninka.ru/article/n/nestabilnye-ateroskleroticheskie-blyashki-koronarnyh-arteriy-u-patsientov-so-stabilnoy-ishemicheskoy-boleznyu-serdtsa/viewer>

- [99] C. Chen *et al.*, “Improving protein-protein interactions prediction accuracy using XGBoost feature selection and stacked ensemble classifier,” *Comput Biol Med*, vol. 123, Aug. 2020, doi: 10.1016/j.combiomed.2020.103899.
- [100] D. Wang, H. He, and C. Wei, “Cellular and potential molecular mechanisms underlying transovarial transmission of the obligate symbiont *Sulcia* in cicadas,” *Environ Microbiol*, vol. 25, no. 4, pp. 836–852, Apr. 2023, doi: 10.1111/1462-2920.16310.
- [101] B.-E. Munkhsaikhan, “SOEs in China and AI Development: Legal Constraints for Foreign Investments towards China’s Digital Transformation,” *Open J Soc Sci*, vol. 12, no. 12, pp. 146–157, 2024, doi: 10.4236/JSS.2024.1212010.
- [102] A. Sekuboyina *et al.*, “VerSe: A Vertebrae labelling and segmentation benchmark for multi-detector CT images,” *Med Image Anal*, vol. 73, p. 102166, Oct. 2021, doi: 10.1016/J.MEDIA.2021.102166.
- [103] H.-Y. Zhou, A. Oliver, J. Wu, and Y. Zheng, “When Semi-Supervised Learning Meets Transfer Learning: Training Strategies, Models and Datasets,” *ArXiv*, Dec. 2018, Accessed: Mar. 13, 2024. [Online]. Available: <http://arxiv.org/abs/1812.05313>
- [104] S. Lutz, K. Amplianitis, and A. Smolic, “AlphaGAN: Generative adversarial networks for natural image matting,” *British Machine Vision Conference 2018, BMVC 2018*, Jul. 2018, Accessed: Apr. 13, 2025. [Online]. Available: <https://arxiv.org/abs/1807.10088v1>
- [105] P. Soille, “Opening and Closing,” *Morphological Image Analysis*, pp. 105–137, 2004, doi: 10.1007/978-3-662-05088-0_4.
- [106] Y. Chung, C. Lim, C. Huang, N. Marrouche, and J. Hamm, “FBA-Net: Foreground and Background Aware Contrastive Learning for Semi-Supervised Atrium

- Segmentation,” Jun. 2023, Accessed: Apr. 13, 2025. [Online]. Available: <http://arxiv.org/abs/2306.15189>
- [107] A. Hatamizadeh, D. Terzopoulos, and A. Myronenko, “End-to-End Boundary Aware Networks for Medical Image Segmentation,” Aug. 2019, Accessed: Apr. 13, 2025. [Online]. Available: <http://arxiv.org/abs/1908.08071>
- [108] S. Falkner, A. Klein, and F. Hutter, “BOHB: Robust and Efficient Hyperparameter Optimization at Scale,” *Proceedings of the 35th International Conference on Machine Learning*, vol. 4, pp. 1437–1446, Jul. 2018, Accessed: Mar. 03, 2024. [Online]. Available: <http://arxiv.org/abs/1807.01774>
- [109] J. Tobin, “Troubleshooting Deep Neural Networks.” [Online]. Available: <https://fullstackdeeplearning.com/spring2021/lecture-7/>
- [110] V. V. Danilov *et al.*, “Automatic scoring of COVID-19 severity in X-ray imaging based on a novel deep learning workflow,” *Sci Rep*, vol. 12, no. 1, p. 12791, Jul. 2022, doi: 10.1038/s41598-022-15013-z.
- [111] L. Li, K. Jamieson, G. DeSalvo, A. Rostamizadeh, and A. Talwalkar, “Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization,” *Journal of Machine Learning Research*, vol. 18, no. 1, pp. 6765–6816, Mar. 2016, Accessed: Mar. 23, 2024. [Online]. Available: <http://arxiv.org/abs/1603.06560>
- [112] “Fabulous Paper LLP - Manufacturer from Mayapuri, New Delhi, India | About Us.” Accessed: Apr. 13, 2025. [Online]. Available: <https://www.indiamart.com/fabulouspaper/profile.html>
- [113] Y. Song, L. Shen, L. Xing, and S. Ermon, “Solving Inverse Problems in Medical Imaging with Score-Based Generative Models,” *ICLR 2022 - 10th International Conference on Learning Representations*, Nov. 2021, Accessed: Apr. 30, 2025. [Online]. Available: <https://arxiv.org/pdf/2111.08005>

Приложение А

Оптимальные конфигурации сетей для исследуемых морфологических признаков бляшки

Таблица А.1 – Оптимальные конфигурации сети для сегментации Lumen

Architecture	Runtime, s	Encoder	Input size	Optimizer	LR	Precision	Recall	F1	IoU	DSC
U-Net++	2965	ResNet-101	512	RMSprop	10^{-5}	0.991	0.993	0.989	0.979	0.989
MA-Net	3552	RegNetY-12GF	512	RAdam	10^{-5}	0.990	0.994	0.989	0.978	0.989
U-Net	2876	ResNet-101	768	RAdam	10^{-5}	0.991	0.995	0.989	0.978	0.989
DeepLabV3	7004	ResNet-101	896	Adam	10^{-5}	0.990	0.991	0.989	0.978	0.989
DeepLabV3+	3559	ResNet-101	512	RMSprop	10^{-5}	0.991	0.989	0.989	0.978	0.989
FPN	1971	ResNet-101	512	RMSprop	10^{-5}	0.992	0.988	0.989	0.977	0.989
LinkNet	4797	EfficientNet B5	640	RAdam	10^{-4}	0.993	0.991	0.988	0.977	0.988
PAN	3593	EfficientNet B0	768	Adam	10^{-4}	0.992	0.990	0.988	0.977	0.988
PSPNet	3394	RegNetY-12GF	896	RAdam	10^{-4}	0.995	0.989	0.988	0.976	0.988

Таблица А.2 – Оптимальные конфигурации сети для сегментации Fibrous cap и Lipid core

Architecture	Runtime, s	Encoder	Input size	Optimizer	LR	Precision	Recall	F1	IoU	DSC
LinkNet	19942	EfficientNet B7	896	RMSprop	10^{-5}	0.700	0.977	0.617	0.488	0.617
U-Net++	13729	EfficientNet B7	640	RAdam	10^{-4}	0.659	0.925	0.602	0.455	0.602
MA-Net	4564	RegNetX-6.4GF	512	RAdam	10^{-4}	0.604	0.961	0.601	0.450	0.601
FPN	14894	EfficientNet B7	768	RMSprop	10^{-4}	0.647	0.751	0.592	0.438	0.592
DeepLabV3+	4381	ResNet-101	640	Adam	10^{-4}	0.626	0.807	0.591	0.449	0.591
U-Net	3322	ResNet-50	512	RAdam	10^{-4}	0.578	0.952	0.583	0.443	0.583
DeepLabV3	6031	RegNetX-6.4GF	512	RMSprop	10^{-4}	0.699	0.764	0.582	0.445	0.582
PAN	15580	EfficientNet B5	896	RAdam	10^{-3}	0.567	0.825	0.560	0.410	0.560
PSPNet	1822	ResNet-18	640	Adam	10^{-3}	0.656	0.862	0.535	0.385	0.535

Таблица А.3 – Оптимальные конфигурации сети для сегментации Vasa vasorum

Architecture	Runtime, s	Encoder	Input size	Optimizer	LR	Precision	Recall	F1	IoU	DSC
U-Net	1259	RegNetX-6.4GF	896	RAdam	10^{-3}	0.814	1.000	0.784	0.695	0.784
MA-Net	1252	ResNet-50	640	RAdam	10^{-3}	0.740	1.000	0.771	0.658	0.771
U-Net++	738	RegNetX-200MF	512	Adam	10^{-3}	0.845	0.972	0.759	0.671	0.759
LinkNet	1029	RegNetX-6.4GF	768	Adam	10^{-3}	0.902	0.929	0.758	0.643	0.758
DeepLabV3+	571	ResNet-50	512	RMSprop	10^{-3}	0.790	0.879	0.755	0.637	0.755
FPN	2100	EfficientNet B7	768	RMSprop	10^{-4}	0.874	0.821	0.751	0.651	0.751
PAN	1168	EfficientNet B5	512	Adam	10^{-3}	0.823	0.818	0.739	0.632	0.739
DeepLabV3	2731	ResNet-101	896	Adam	10^{-4}	0.841	0.908	0.719	0.617	0.719
PSPNet	1173	RegNetY-12GF	896	RMSprop	10^{-3}	0.833	0.848	0.706	0.587	0.706

Приложение Б

Показатели производительности исследуемых моделей на различных конфигурация набора данных

Таблица Б.1 – Показатели производительности исследуемых моделей на наборе данных с 4-я классами (Lumen, Fibrous cap, Lipid core, Vasa vasorum)

Model	Class	DSC	Precision	Recall	F1
DeepLabV3	Lumen	0.99	0.992	0.987	0.99
	Fibrous cap	0.639	0.613	0.753	0.639
	Lipid core	0.555	0.587	0.697	0.555
	Vasa vasorum	0.0089	0.0088	0.009	0.0089
	Mean	0.521	0.486	0.630	0.521
DeepLabV3+	Lumen	0.99	0.993	0.988	0.99
	Fibrous cap	0.592	0.592	0.681	0.592
	Lipid core	0.534	0.481	0.72	0.534
	Vasa vasorum	0.0088	0.0088	0.009	0.0088
	Mean	0.513	0.485	0.614	0.513
FPN	Lumen	0.989	0.99	0.988	0.989
	Fibrous cap	0.494	0.641	0.45	0.494
	Lipid core	0.53	0.559	0.568	0.53
	Vasa vasorum	0.0086	0.0086	0.0086	0.0086
	Mean	0.499	0.547	0.494	0.499
LinkNet	Lumen	0.99	0.989	0.991	0.99
	Fibrous cap	0.538	0.609	0.569	0.538
	Lipid core	0.48	0.442	0.654	0.48
	Vasa vasorum	0.0006	0.003	0.0096	0.0006
	Mean	0.49	0.524	0.518	0.49
MA-Net	Lumen	0.99	0.992	0.988	0.99
	Fibrous cap	0.684	0.639	0.786	0.684
	Lipid core	0.507	0.526	0.535	0.507
	Vasa vasorum	0.0091	0.0097	0.0085	0.0091
	Mean	0.529	0.516	0.596	0.529
PAN	Lumen	0.99	0.99	0.99	0.99
	Fibrous cap	0.571	0.563	0.654	0.571
	Lipid core	0.532	0.637	0.512	0.532
	Vasa vasorum	0.009	0.009	0.0089	0.009
	Mean	0.499	0.548	0.515	0.499

Продолжение таблицы Б.1 – Показатели производительности исследуемых моделей на наборе данных с 4-я классами (Lumen, Fibrous cap, Lipid core, Vasa vasorum)

Model	Class	DSC	Precision	Recall	F1
PSPNet	Lumen	0.985	0.988	0.982	0.985
	Fibrous cap	0.561	0.743	0.495	0.561
	Lipid core	0.593	0.586	0.722	0.593
	Vasa vasorum	0.0078	0.0071	0.0089	0.0078
	Mean	0.521	0.586	0.521	0.521
U-Net	Lumen	0.99	0.99	0.99	0.99
	Fibrous cap	0.617	0.612	0.696	0.617
	Lipid core	0.506	0.447	0.665	0.506
	Vasa vasorum	0.0092	0.0094	0.009	0.0092
	Mean	0.512	0.513	0.549	0.512
U-Net++	Lumen	0.99	0.99	0.99	0.99
	Fibrous cap	0.61	0.644	0.631	0.61
	Lipid core	0.521	0.53	0.626	0.521
	Vasa vasorum	0.0089	0.0089	0.0091	0.0089
	Mean	0.528	0.549	0.539	0.528

Таблица Б.2 – Показатели производительности исследуемых моделей на наборе данных с 3-я классами (Lumen, Fibrous cap, Lipid core)

Model	Class	DSC	Precision	Recall	F1
DeepLabV3	Lumen	0.984	0.977	0.990	0.984
	Fibrous cap	0.617	0.602	0.702	0.617
	Lipid core	0.497	0.415	0.748	0.497
	Mean	0.699	0.665	0.813	0.699
DeepLabV3+	Lumen	0.984	0.977	0.991	0.984
	Fibrous cap	0.664	0.638	0.778	0.664
	Lipid core	0.515	0.434	0.778	0.515
	Mean	0.721	0.683	0.849	0.721
FPN	Lumen	0.987	0.993	0.982	0.987
	Fibrous cap	0.480	0.750	0.398	0.480
	Lipid core	0.490	0.520	0.525	0.490
	Mean	0.652	0.754	0.635	0.652
LinkNet	Lumen	0.988	0.990	0.986	0.988
	Fibrous cap	0.572	0.595	0.611	0.572
	Lipid core	0.376	0.335	0.544	0.376
	Mean	0.645	0.640	0.714	0.645
MA-Net	Lumen	0.988	0.987	0.990	0.988
	Fibrous cap	0.612	0.654	0.627	0.612
	Lipid core	0.427	0.327	0.759	0.427
	Mean	0.676	0.656	0.792	0.676
PAN	Lumen	0.980	0.981	0.979	0.980
	Fibrous cap	0.625	0.695	0.618	0.625
	Lipid core	0.558	0.588	0.606	0.558
	Mean	0.721	0.755	0.734	0.721
PSPNet	Lumen	0.981	0.984	0.979	0.981
	Fibrous cap	0.480	0.722	0.395	0.480
	Lipid core	0.532	0.620	0.527	0.532
	Mean	0.664	0.776	0.634	0.664
U-Net	Lumen	0.988	0.987	0.990	0.988
	Fibrous cap	0.665	0.711	0.716	0.665
	Lipid core	0.449	0.404	0.726	0.449
	Mean	0.701	0.701	0.811	0.701
U-Net++	Lumen	0.986	0.986	0.986	0.986
	Fibrous cap	0.667	0.689	0.706	0.667
	Lipid core	0.489	0.519	0.624	0.489
	Mean	0.714	0.731	0.772	0.714

Таблица Б.3 – Показатели производительности исследуемых моделей на наборе данных с 1-м классом Vasa vasorum)

Model	Class	DSC	Precision	Recall	F1
DeepLabV3	Vasa vasorum	0.598	0.607	0.687	0.598
DeepLabV3+	Vasa vasorum	0.617	0.664	0.659	0.617
FPN	Vasa vasorum	0.66	0.637	0.761	0.66
LinkNet	Vasa vasorum	0.341	0.232	0.924	0.341
MA-Net	Vasa vasorum	0.721	0.676	0.833	0.721
PAN	Vasa vasorum	0.649	0.639	0.742	0.649
PSPNet	Vasa vasorum	0.527	0.515	0.647	0.527
U-Net	Vasa vasorum	0.616	0.533	0.844	0.616
U-Net++	Vasa vasorum	0.612	0.526	0.808	0.612

Приложение В

Визуализация результатов работы модели сегментации атеросклеротической бляшки

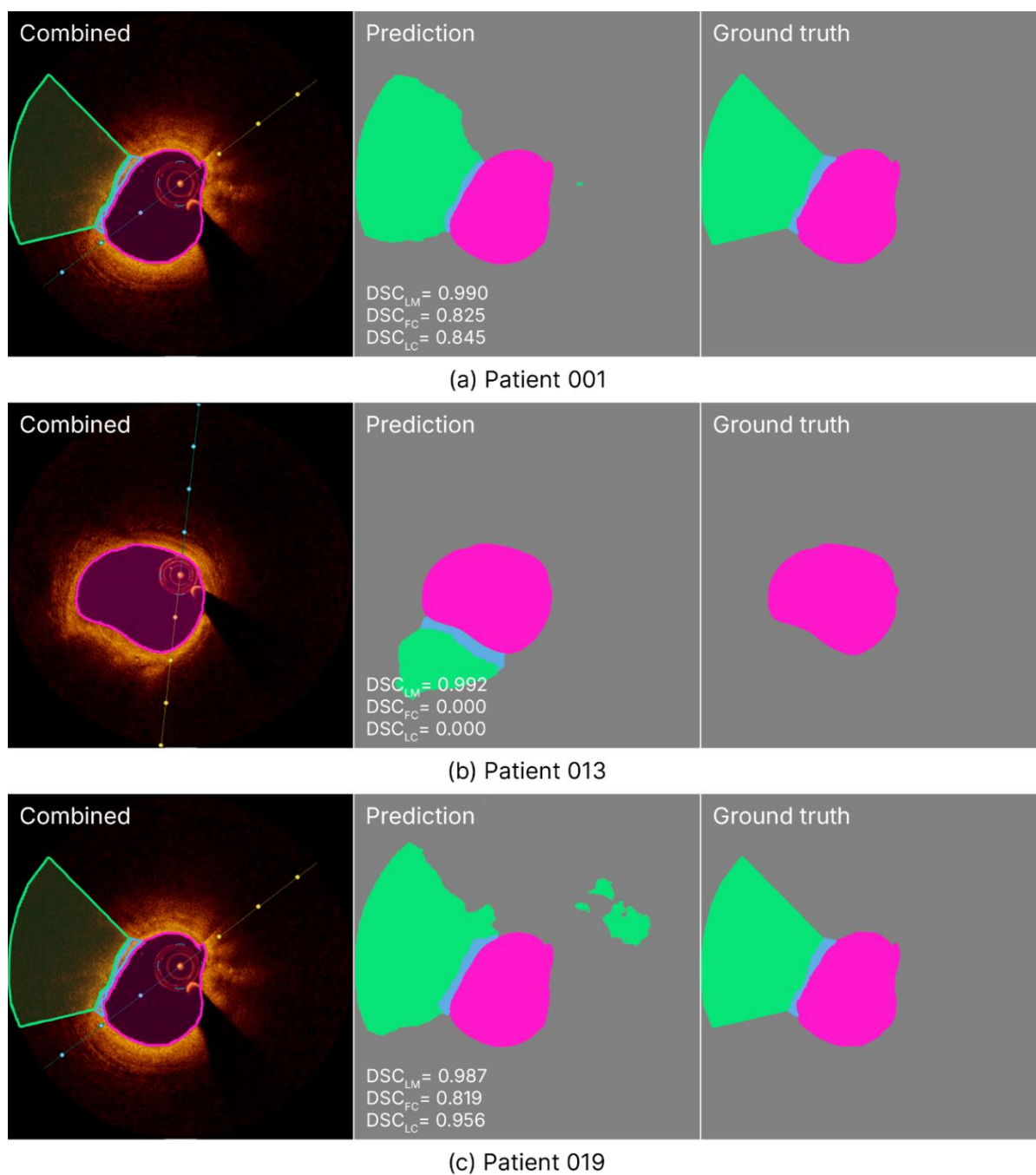


Рисунок В.1 – Сравнение работы сегментационного ансамбля нейросетевых моделей с результатами экспертной разметки

Приложение Г

Дополнительная информация к задаче сегментации гистологических аномалий

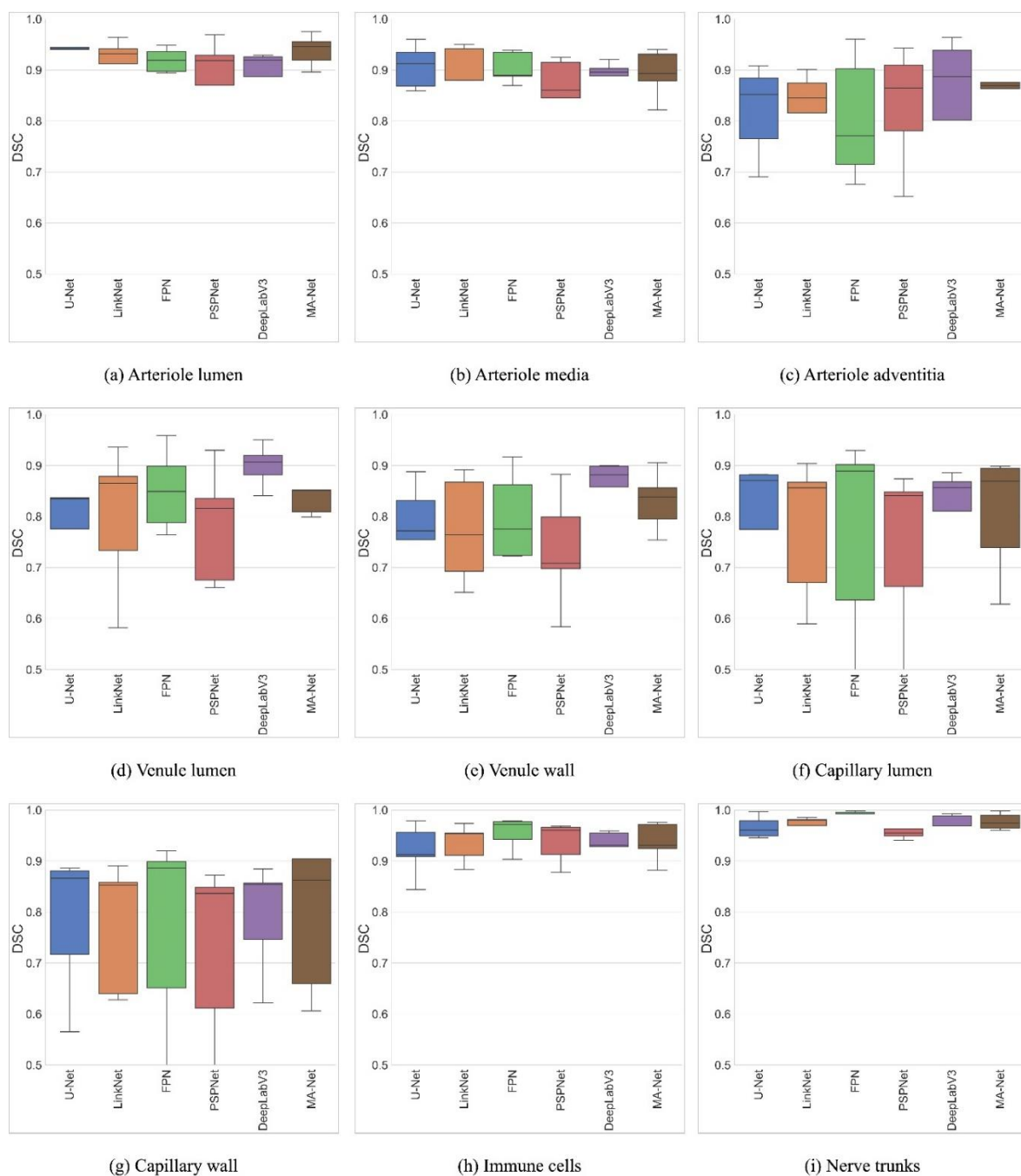


Рисунок Г.1 – Значения DSC на тестовом подмножестве по различным гистологическим признакам в оптимальную

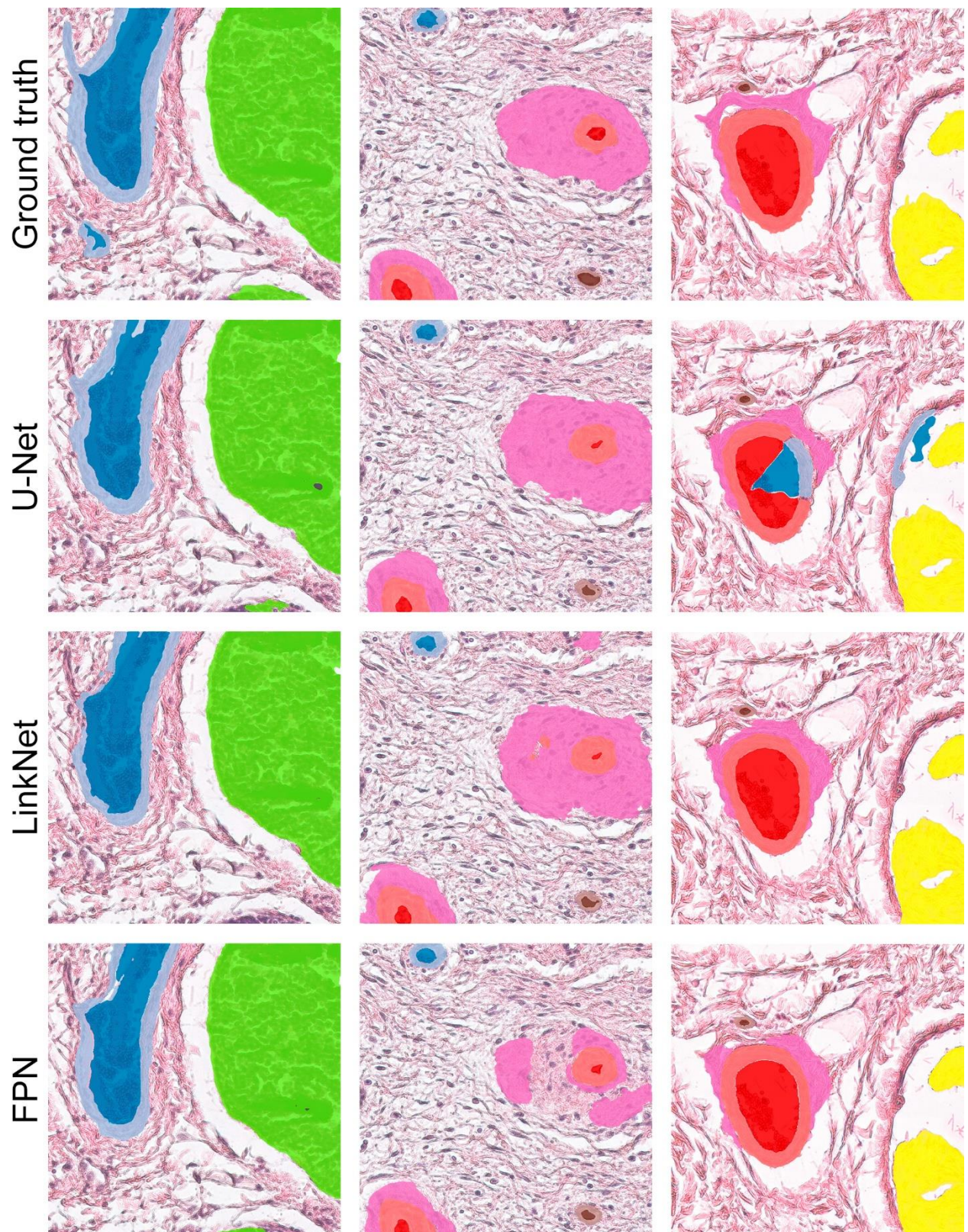


Рисунок Г.2 – Сравнение предсказаний U-Net, LinkNet и FPN с истинными данными

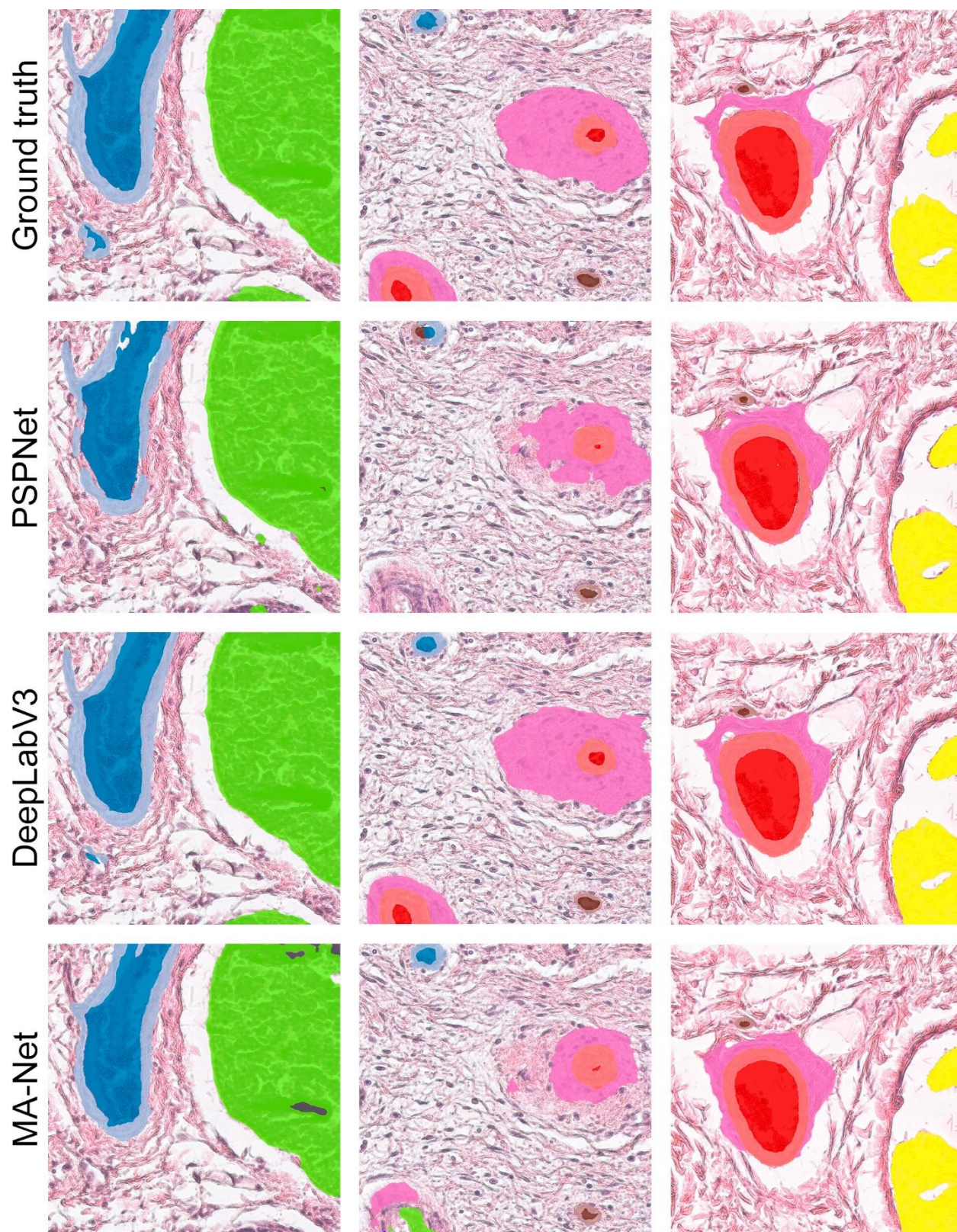


Рисунок Г.3 – Сравнение предсказаний PSPNet, DeepLabV3 и MA-Net с истинными данными

Приложение Д

Акты внедрения результатов диссертационной работы



ООО «Фабула»
634050, г. Томск, пр-т Ленина, д. 80а
fabula-ai.com, info@fabula-app.com

ИНН 7000005379
КПП 700001001
ОГРН 1237000002733

«Утверждаю»
Генеральный директор
ООО «Fabula»

А.Х. Оздиев
«16» декабря 2025

АКТ

О внедрении научных и практических результатов Лаптева Владислава Витальевича в ООО «Fabula»

Комиссия в составе: Оздиев А.Х., генеральный директор, Кадыров Р.И., директор по продукту, Гмырь В., технический директор — составила настоящий акт о том, что следующие результаты диссертационной работы Лаптева Владислава Витальевича используются при обучении и анализе моделей распознавания:

- Метод подготовки данных для обучения сегментационных моделей распознавания объектов на изображениях;
- Алгоритм формирования адаптивных трехканальных карт с целью дальнейшего анализа и уточнения границ объектов сегментации.

Генеральный директор

Директор по продукту

Технический директор

Оздиев А.Х.

Кадыров Р.И.

Гмырь В.

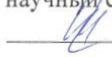


Минздрав России

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Сибирский государственный медицинский
университет» Министерства здравоохранения
Российской Федерации
(ФГБОУ ВО СибГМУ Минздрава России)
Московский тракт, д. 2, г. Томск, 634050
Телефон (3822) 53 04 23;
Факс (3822) 53 33 09
e-mail: office@ssmu.ru
http://www.ssmu.ru
ОКПО 01963539 ОГРН 1027000885251
ИНН 7018013613 КПП 701701001

УТВЕРЖДАЮ

Руководитель НТЦ «Цифровая
медицина и киберфизика», ведущий
научный сотрудник, к.м.н.

 И.В. Толмачев

14 мая 2025

АКТ внедрения результатов диссертационной работы

Внедрения результатов диссертационной работы Лаптева Владислава Витальевича на тему «Методы и алгоритмы сегментации объектов с отсутствием явных характеристических признаков на медицинских изображениях», представленной на соискание ученой степени кандидата технических наук по специальности 2.3.1 – Системный анализ, управление и обработка информации, статистика в запатентованный сервис «Офтальмик+» (Программа для ЭВМ / Офтальмик. Система хранения снимков глазного дна со встроенной системой поддержки принятия врачебных решений / 2023663680).

Мы, нижеподписавшиеся, комиссия в составе: председателя - руководителя отдела быстрого прототипирования медицинских изделий НТЦ «Цифровая медицина и киберфизика», научного сотрудника, к.т.н., доцент Солдатов А.А., и членов: эксперта отдела быстрого прототипирования медицинских изделий НТЦ «Цифровая медицина и киберфизика» Абраменко Н.С., научного сотрудника отдела разработки медицинского программного обеспечения НТЦ «Цифровая медицина и киберфизика» Кавериной И.С., заместителя руководителя НТЦ «Цифровая медицина и киберфизика» Желбуновой Е.К. удостоверяем, что результаты диссертационной работы Лаптева Владислава Витальевича внедрены в проект «Офтальмик+» с целью визуализации сегментационной маски выделения патологических очагов на изображении с использованием анализа карты признаков (тепловой карты).

Председатель

Члены комиссии



А.А. Солдатов

Н.С. Абраменко

И.С. Каверина

Е.К. Желбунова

Министерство науки и высшего образования
Российской Федерации



Федеральное государственное
бюджетное научное учреждение
«Научно-исследовательский институт
комплексных проблем
сердечно-сосудистых заболеваний»
(НИИ КПССЗ)

бульвар имени академика Л.С. Барбараша, стр. 6,
г. Кемерово, Кемеровская область – Кузбасс, 650002
тел. 8 (3842) 643-308, факс 8 (3842) 643-410
e-mail: reception@kemcardio.ru
www.kemcardio.ru
ОКПО 55608705; ОГРН 1034205024479;
ИНН/КПП 4205012290/420501001

№ _____
на № _____ от _____

Утверждаю
заведующий отделом Хирургии сердца
и сосудов
д.м.н. Ганюков Владимир Иванович

«23» июня 2025

АКТ

О внедрении научных и практических результатов

Разработанные Лаптевым Владиславом Витальевичем в рамках диссертационного исследования метод подготовки данных семантической сегментации, а также метод визуализации и интерпретации данных предсказания используются в работе НИИ КПССЗ, лаборатории тканевой инженерии и внутрисосудистой визуализации. Применение указанных методов позволило обучить модель для анализа ОКТ-снимков и автоматизировать процесс обнаружения атеросклеротических бляшек в практической деятельности лаборатории.

Заведующий отделением рентгенохирургических
методов диагностики лечения, к.м.н.

Торгунаков С.А.

Заведующий лабораторией Тканевой инженерии и
внутрисосудистой визуализации, к.м.н.

Кочергин Н.А.

Заведующий отделом Хирургии сердца и сосудов, д.м.н.



Приложение Е

Свидетельства о государственной регистрации программы для ЭВМ

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2022619805

Программа для ЭВМ «Многофункциональный мобильный комплекс мониторинга, ситуационного анализа и оповещения труднодоступных объектов с использованием беспилотных летательных аппаратов и ГИС-технологий (ММК МАО)»

Правообладатель: *Общество с ограниченной ответственностью «ИНКОМ» (RU)*

Авторы: *Сонькин Дмитрий Михайлович (RU), Плешков Алексей Геннадьевич (RU), Изумнов Артём Олегович (RU), Осипенко Денис Витальевич (RU), Слезенко Николай Сергеевич (RU), Погребной Вадим Юрьевич (RU), Волегова Любовь Владимировна (RU), Кравченко Андрей Александрович (RU), Лаптев Никита Витальевич (RU), Лаптев Владислав Витальевич (RU), Колпащников Дмитрий Юрьевич (RU), Гергет Ольга Михайловна (RU), Данилов Вячеслав Владимирович (RU)*

Заявка № **2022618392**

Дата поступления **05 мая 2022 г.**

Дата государственной регистрации

в Реестре программ для ЭВМ **26 мая 2022 г.**



Руководитель Федеральной службы
по интеллектуальной собственности

Ю.С. Зубов

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2025617776

**Автоматизированная система сегментации оптической
когерентной томографии с использованием методов
глубокого обучения**

Правообладатель: *Федеральное государственное бюджетное научное
учреждение «Научно-исследовательский институт комплексных
проблем сердечно-сосудистых заболеваний» (RU)*

Авторы: *Онищенко Павел Сергеевич (RU), Овчаренко Евгений
Андреевич (RU), Клышников Кирилл Юрьевич (RU), Кочергин
Никита Александрович (RU), Данилов Вячеслав Владимирович
(RU), Лаптев Владислав Витальевич (RU)*



Заявка № 2025616488

Дата поступления 24 марта 2025 г.

Дата государственной регистрации

в Реестре программ для ЭВМ 28 марта 2025 г.

Руководитель Федеральной службы
по интеллектуальной собственности

Ю.С. Зубов