

ТОМСКИЙ
ПОЛИТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ



На правах рукописи



Лаптев Владислав Витальевич

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

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

2.3.1 – Системный анализ, управление и обработка информации,
статистика

Томск – 2025

Работа выполнена в Национальном исследовательском Томском политехническом университете

Научный руководитель:

Гергет Ольга Михайловна

Доктор технических наук, профессор, ФГАОУ ВО «Национальный исследовательский Томский политехнический университет», профессор отделения информационных технологий

Официальные оппоненты:

Ходашинский Илья Александрович

Доктор технических наук, профессор, ФГАОУ ВО «Томский государственный университет систем управления и радиоэлектроники», профессор кафедры компьютерных систем в управлении и проектировании

Кистинев Юрий Владимирович

Доктор физико-математических наук, профессор, ФГАОУ ВО «Национальный исследовательский Томский государственный университет», профессор кафедры общей и экспериментальной физики

Защита состоится 24.12.2025 в ____ часов на заседании диссертационного совета ДС.ТПУ.15 Национального исследовательского Томского политехнического университета по адресу: Россия, 634034, г. Томск, ул. Советская, 84/3, ауд. 214.



С диссертацией можно ознакомиться в научно-технической библиотеке Томского политехнического университета и на сайте dis.tpu.ru при помощи QR-кода.

Автореферат разослан «__» _____ 2025 г.

Ученый секретарь
диссертационного совета ДС.ТПУ.15
Д.Т.Н.

A handwritten signature in blue ink, consisting of several fluid, overlapping strokes.

А.Я. Пак

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. В настоящее время медицинская диагностика является одной из наиболее перспективных и востребованных областей применения искусственного интеллекта (ИИ). К популярным задачам относятся: сегментация и локализация опухолей или других новообразований, оценка активности мозга, изучение плотности тканей и т.п. С помощью алгоритмов компьютерного зрения стало возможным находить закономерности и аномалии на многомодальных снимках, таких как рентген, КТ, МРТ. Целью автоматизированного анализа медицинских данных является содействие специалисту осуществлять быструю и точную постановку 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, но и получить значимые характеристики, такие как: границы объекта, форма, площадь и т.п. Существует множество алгоритмов сегментации изображений, которые можно классифицировать по следующим группам методов, на основе: энтропии, кластеризации, локальных порогов, пиксельных корреляциях, атрибутах, искусственных нейронных сетях.

В настоящее время наибольший интерес представляют системы, основанные на алгоритмах с использованием искусственных нейронных сетей. Именно эта группа обладает таким понятием как обобщающая способность, что позволяет применять их на более широком спектре задач. Отметим, 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, аннотирование данных выполняют высококвалифицированные эксперты, 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.

Задачи сегментации медицинских изображений активно исследуются в мировой и отечественной научной практике на протяжении последних десятилетий. Наибольшее развитие получили методы, основанные на классических алгоритмах обработки изображений, а также современные подходы с применением глубоких сверточных нейронных сетей (Convolutional neural network, CNN). Известные архитектуры, такие как U-Net, DeepLabV3+, PSPNet и их модификации, демонстрируют высокие результаты в сегментации стандартных объектов с четкими границами и выраженными характеристиками.

Отдельные исследования рассматривают применение карт активации для интерпретации работы нейронных сетей, однако их использование в качестве промежуточного источника формирования исходных сегментационных масок в задачах медицины ранее подробно не исследовалось. Следует отметить, что в существующих подходах отсутствует комплексная интеграция автоматического анализа карт признаков с последующей генерацией трехканальных масок изображений и оптимизацией границ объектов исследования с применением специализированных моделей, таких как FBA-Net (Feature Boundary-Aware Network).

Обнаружение и анализ сложных объектов исследования является актуальной научной задачей. 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.

Создание сегментационных масок изображений нестандартных мультимодальных данных является еще более сложной и трудоемкой задачей. 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 поставлены и решены следующие **задачи**:

1. Провести комплексный анализ предметной области, включающий исследование специфики медицинской визуализации и задач сегментации объектов с отсутствием явных характеристических признаков.
2. Выполнить анализ медицинских данных различных модальностей, включающий изучение их структурных особенностей, форматов хране-

ния, специфических характеристик, методов их предварительной обработки, алгоритмов сегментации и классификации, а также современных подходов к визуализации медицинских изображений.

3. Исследовать современные архитектуры сверточных нейронных сетей для задачи классификации.
4. Провести анализ существующих методов и алгоритмов сегментации и локализации медицинских инструментов и анатомических структур.
5. Разработать технологию автоматизированного формирования сегментационных масок медицинских изображений на основе анализа карт признаков классификационных моделей с последующей оптимизацией границ объектов.
6. Предложить и исследовать метод автоматического выделения областей интереса на изображениях, а также разработать алгоритм формирования адаптивной трехканальной сегментационной карты.
7. 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.

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

Научная новизна

1. Разработана технология автоматизированной сегментации изображений, базирующаяся на интерпретации признаков сверточных классификационных нейросетей с использованием модели FBA-Net, обеспечивающая точное выделение области интереса за счет организации последовательного потока данных между этапами, где на каждом этапе тензоры изображений вычисляются с минимальными потерями информации.
2. 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 для объектов с неявными характеристическими признаками.
3. Для эффективного применения модели FBA-Net разработан алгоритм автоматизированного формирования трехканальных карт на основе раз-

деления изображения на зоны объекта, фона и области неопределенности, что обеспечивает детальный анализ объекта исследования и повышает точность сегментации анализируемого изображения.

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. Достоверность предложенных алгоритмов и методов верифицирована в ходе серии экспериментов, а их работоспособность подтверждается воспроизводимостью результатов при обработке разнородных графических данных.

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, использующиеся для обработки медицинских данных, применяются при выполнении минимально инвазивных вмешательств в хирургической онкологии, нейрохирургии и сосудистой хирургии.

Положения, выносимые на защиту:

1. Разработанная технология автоматизированной сегментации изображений позволяет сократить время подготовки обучающего множества более чем в три 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.

Реализация и внедрение результатов работы. 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; 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+».

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).

Публикации. Основные результаты исследований отражены в 10 публикациях: 7 статей в ведущих научных журналах рекомендованных ВАК, 6 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, включая 31 рисунок и 11 таблиц.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обосновывается актуальность работы в данном научном направлении, формулируются цель и задачи исследования, определена

научная новизна, а также отражены практическая значимость и положения, выносимые на защиту.

В первой главе «Анализ предметной области и современных методов сегментации сложных объектов исследования» проведен комплексный обзор медицинской визуализации с точки зрения задач сегментации объектов с отсутствием явных характеристических признаков. Рассмотрены форматы и стандарты хранения данных (Analyze, Minc, NIfTI, DICOM), особенности изображений медицинских модальностей (МРТ, КТ, ОКТ, ПЭТ, ультразвуковые исследования), а также факторы, осложняющие сегментацию: низкая контрастность, отсутствие четких границ, высокий уровень шумов, малая величина целевых объектов. Рассмотрены основные понятия и методы семантической сегментации изображений в области медицинской визуализации. Приведен обзор классических и современных методов сегментации, таких как методы порогового выделения, регион–ориентированные подходы, алгоритмы на основе границ и современные методы сегментации с применением сверточных нейронных сетей (CNN). Особое внимание уделено существующим архитектурам, включая U-Net и ее модификации, а также рассмотрены методы генерации обучающих масок и подходы к полуавтоматической аннотации изображений. Отдельное внимание уделено оценке вычислительных затрат: показано, что сегментационные модели требуют существенно больших ресурсов и времени обучения по сравнению с классификационными, особенно при работе с мультимодальными данными и трехмерными наборами. Отмечено, что значительная часть трудозатрат приходится на ручную аннотацию данных, при этом точность разметки зависит от субъективного фактора эксперта. Обоснована актуальность использования промежуточных карт признаков классификационных моделей для автоматизированного выделения областей интереса и формулируются ограничения известных подходов, не позволяющие в полной мере снизить затраты на аннотирование данных. В результате анализа выявлены сильные стороны существующих решений и их ограничения, такие как неэффективность при слабовыраженных признаках, высокая стоимость подготовки обучающей выборки, снижение точности при анализе границ объектов. На основе полученных выводов сформулирована постановка задачи: разработка эффективной технологии автоматизированного формирования сегментационных масок изображений на основе карт признаков классификационной модели с последующим уточнением границ объектов. Концепция предлагаемой технологии представлена в виде схемы на рисунке 1 и включает следующие этапы: анализ активационных карт признаков, формирование трехканальных карт, уточнение границ сегментации, итоговое формирование обучающих масок.

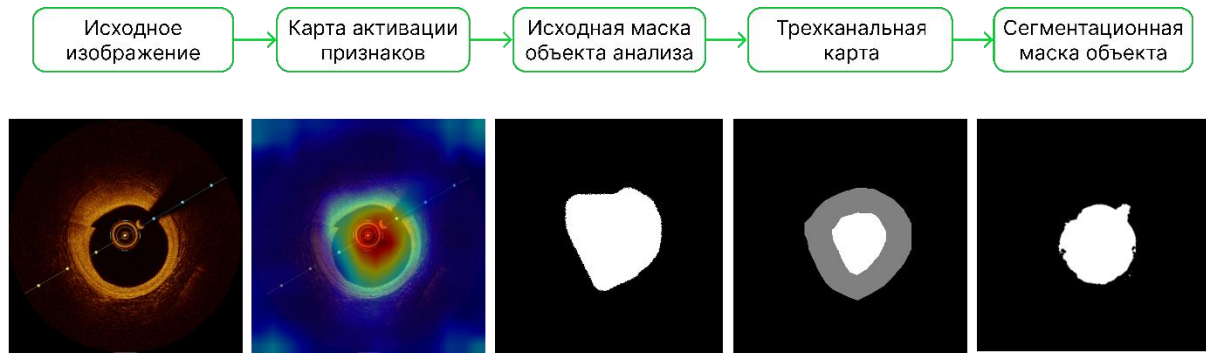


Рисунок 1 – Этапы технологии формирования сегментационной маски изображения

Во второй главе «Метод и алгоритмы сегментации сложных объектов медицинских изображений с отсутствием явных характеристических признаков» представлено подробное описание разработанной технологии, направленной на автоматизированную подготовку сегментационных данных с отсутствием явных характеристических признаков, таких как атеросклеротические бляшки на снимках оптической когерентной томографии (ОКТ). 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. Особое внимание уделено информационной передаче между этапами, обеспечивающей 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. Это обеспечивает минимизацию потерь информации, согласованность обработки и отсутствие циклических зависимостей, что упрощает реализацию и повышает воспроизводимость. Технология 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 слоёв классификационной модели;
 - Преобразование карты признаков в исходную маску изображения с выделением предполагаемой области объекта анализа.
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.

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.

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.

Метод формирования исходной маски изображения 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.

Алгоритм формирования карты активации признаков. 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 (например, Xception с

точностью классификации 97,8%, см. таблица 1), 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 на карту активаций слоя, предшествующего Global Average Pooling. Реализация выполняется через матричное умножение:

$$M = A \times w_c, \quad (1)$$

$$A \in R^{(H_c \times W_c \times K)}, \quad (2)$$

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.

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:

$$R = \alpha \times I + (1 - \alpha) \times M_{\text{окраш}}, \quad (3)$$

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 (3), получив визуализированную карту $I_{\text{виз}}$.

Алгоритм обеспечивает интерпретацию решений модели, визуализируя области изображения, максимизирующие активацию для предсказанного класса. Пример визуализации работы алгоритма для исследуемых классов представлен на рисунке 2.

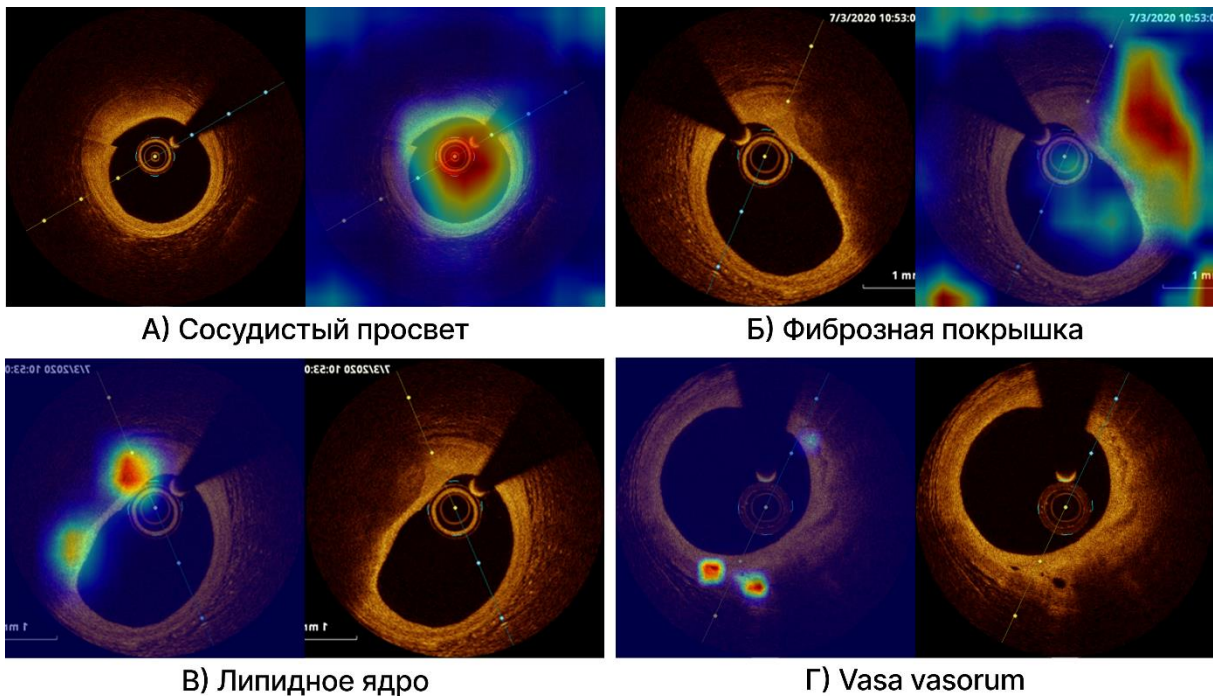


Рисунок 2 – Визуальное представление работы алгоритма формирования карты активации признаков

Рисунок иллюстрирует результаты визуализации работы предложенного алгоритма для различных морфологических признаков сосудистой бляшки: сосудистый просвет (А), фиброзная покрывка (Б), липидное ядро (В), vasa vasorum (Г). Для каждого признака представлены два изображения: исходное изображение и результат наложения карты активации признаков на исходное изображение, что обеспечивает удобство визуального анализа. Представленные результаты демонстрируют эффективность алгоритма, проявляющуюся в четком выделении всех исследуемых морфологических признаков на анализируемых изображениях.

Алгоритм формирования трехканальной карты. 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.

$$B_{\text{фильтр}} = (B_{\text{строк}} \circ S) \bullet S, \quad (4)$$

где $\circ$ — операция морфологического открытия:

$$B \circ S = (B \odot S) \oplus S, \quad (5)$$

$\bullet$ — операция морфологического закрытия:

$$B \bullet S = (B \oplus S) \odot S, \quad (6)$$

$\odot$ — 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.

$$B \odot S = \{z \in E \mid S_z \subseteq B\}, \quad (7)$$

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.

$$B \oplus S = \{x | (S + x) \cap B \neq \emptyset\} \quad (8)$$

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

$$T(x, y) = \left\{ \begin{array}{l} 1, \text{ если } F(x, y) = 1(\text{объект}) \\ 0, \text{ если } B(x, y) = 1(\text{фон}) \\ 0.5, \text{ в остальных случаях (зона неопределенности)} \end{array} \right\} \quad (9)$$

На данном этапе алгоритм позволяет получить трехканальную сегментационную карту, разделяющую изображение на три зоны: область объекта фон и область неопределенности, что обеспечивает автоматизированную локализацию сложных анатомических структур и упрощает последующую сегментацию.

В третьей главе «Применение разработанных методов и алгоритмов в задачах подготовки данных и обучения моделей сегментации сложных объектов исследования» представлены результаты тестирования разработанного метода и алгоритмов формирования исходной маски изображения, генерации трехканальных карт и последующего уточнения границ объектов с использованием FBA-Net модели. Проведена апробация предложенного метода формирования маски на реальных наборах медицинских изображений с низким контрастом и отсутствием явных характеристических признаков.

Исходные данные получены с двух устройств: от St. Jude Medical и LightLab Imaging при проведении диагностических процедур в НИИ комплексных проблем сердечно-сосудистых заболеваний (Кемерово) и Тюменском кардиологическом научном центре (Тюмень). Общий объем данных составил 103 ОКТ исследования, каждое исследование представляет собой трехканальный видеофайл с разрешением кадра 704×704 пикселя. продолжительностью от 215 до 270 кадров.

Разработанные методы и алгоритмы внедрены в программно-аппаратный комплекс, реализующий подготовку данных семантической сегментации.

Для тестирования метода создания сегментационных масок изображений обучены 4 модели сверточных нейронных сетей с применением технологии трансферного обучения на базе ResNet-50, EfficientNetB5, Xception, ConvNeXtBase. Данные модели выбраны как архитектурные решения, хорошо зарекомендовавшие себя при решении медицинских задач распозна-

вания сложных объектов. Обучение проводилось в течение 75 эпох с использованием, предварительно размеченных 25 698 изображений. Результаты тестирования моделей приведены в таблице 1.

Таблица 1 – Результаты работы моделей классификации на тестовом множестве

Наименование модели	Точность, %	Полнота, %	Количество параметров	Время вычисления, с
ResNet-50	70,5	70,3	25557032	0,0018
EfficientNetB5	68,4	67,2	30389784	0,0025
Xception	91,7	91,7	22855952	0,0013
ConvNeXtBase	65,4	65,7	88591464	0,0021

Согласно полученным результатам, сделан вывод, что лидером среди моделей классификации атеросклеротических бляшек стала модель Xception с точностью классификации по метрике качества Ассигасу 97,8% на тестовом наборе данных. 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 САМ, включая GradCAM, GradCAM++ и другие. Представленные ниже сравнительные карты активации демонстрируют вариативность работы алгоритма САМ для признаков разного размера и сложности, а также сравнение с результатами, полученными в ходе работы авторского алгоритма:

- Сосудистый просвет: LayerCAM, HiResCAM, GradCAMElementWise и EigenGradCAM постоянно демонстрировали превосходную точность 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.
- Фиброзная покрышка: 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 (объект исследования с отсутствием) случаях с диффузной текстурой.
- Липидное ядро: сегментация липидного ядра оказалась сложной для всех протестированных алгоритмов. Слабые или неправильные активации наблюдались во всех методах САМ, включая GradCAM, HiResCAM,

GradCAMElementWise, GradCAM++, XGradCAM, EigenCAM, EigenGradCAM и ScoreCAM, а также авторский алгоритм.

- 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.

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 стабильное качество работы 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. Применение данных 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 при анализе класса «сосудистый просвет» приведена на рисунке 3.

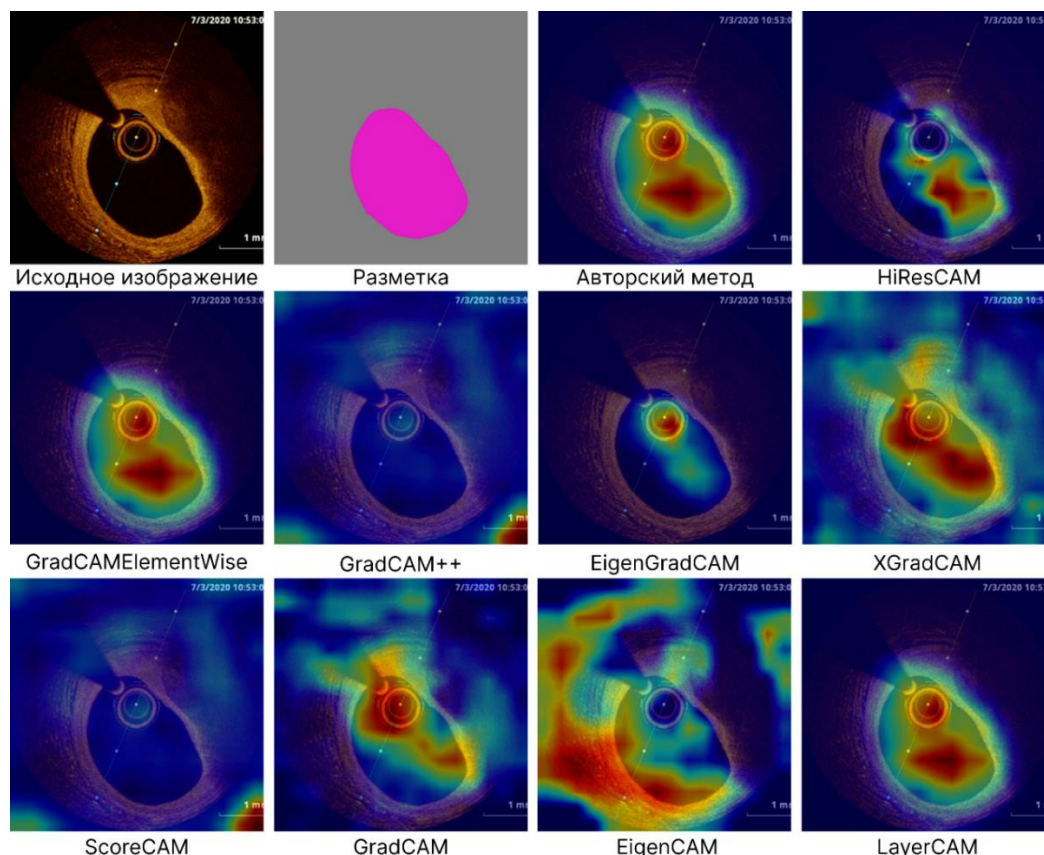


Рисунок 3 – Результат визуализации карты признаков класса сосудистый просвет

Как видно из рисунка 3 авторский метод формирования карт активации признаков для сегментации сложных анатомических структур в сравнении с исследуемыми аналогами. Для проведения визуальной оценки иллюстрация включает входное изображение, экспертную разметку и результаты работы каждого исследуемого метода.

Представленный авторский метод создания сегментационных масок изображений на основе карт признаков демонстрирует высокую точность по сравнению с представленными в работе аналогами. В результате качественной оценки установлено, что представленный алгоритм не уступает существующим аналогам по точности выделения морфологических признаков, что конечно же связано со схожестью рассмотренных алгоритмов. Однако предложенный алгоритм значительно сокращает время, затрачиваемое на аннотирование данных, благодаря автоматизации процесса формирования сегментационных карт. Усредненные по классам количественные показатели, полученные в ходе анализа 100 изображений, представлены в таблице 2.

Таблица 2 – Количественные показатели анализа полученных масок изображений

Алгоритм	IoU, %	Время вычисления, с	Алгоритм	IoU, %	Время вычисления, с
GradCAM	21,6	0,0068	LayerCAM	28,1	0,0076
HiResCAM	27,8	0,0072	EigenGradCAM	9,2	0,0080
GradCAM++	23,5	0,0071	EigenCAM	12,4	0,0088
XGradCAM	25,9	0,0071	GradCAMElementWise	16,3	0,0070
ScoreCAM	22,4	0,0072	Авторский метод	27,2	0,0024

В результате анализа полученных показателей в качестве итогового решения выбрана реализация предложенной технологии. Реализация метода и алгоритмов, лежащих в основе технологии не только продемонстрировала высокие показатели точности, уступающие модели фавориту на 0,9%, но и показала 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.

Заключительным этапом формирования сегментационной маски изображения является уточнение границ анатомической структуры с отсутствием явных характеристических признаков. Представленный в работе алгоритм формирования трехканальной маски отлично подходит для решения поставленной задачи. На рисунке 4 в столбце (в) демонстрируется результат работы данного алгоритма для каждого исследуемого класса.

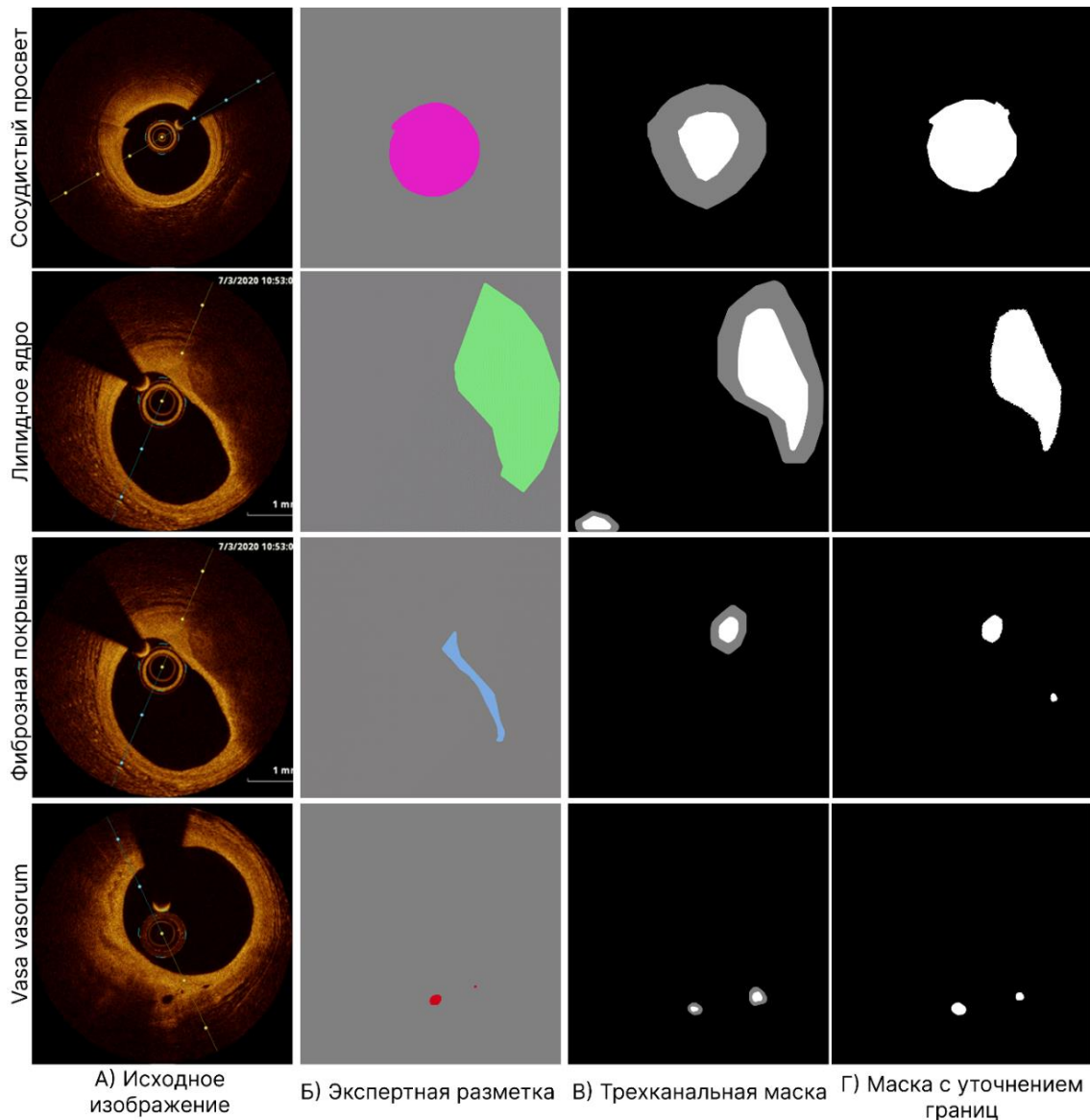


Рисунок 4 – Результат формирования маски для исследуемых классов

На представленных примерах отчетливо видно, что полученной на предыдущем этапе информации достаточно для формирования области объекта исследования, а также области неопределенности, практически полностью покрывающей область, где может располагаться объект анализа. Результаты анализа трехканальных масок FBA-Net моделью представлены на рисунке 4 в столбце (г).

Согласно результатам, представленным в таблице 3 и на рисунке 4 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, с целью 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 сосудистый просвет и *vasa vasorum* с учетом всех проводимых экспериментов было затрачено менее 1 месяца, что еще раз подчеркивает эффективность данного метода.

Таблица 3 – Показатели эффективности метода формирования сегментационных масок изображений, на примере класса сосудистый просвет и *vasa vasorum* в расчете 1 изображения

Алгоритм	IoU, %	Точность	Полнота	Время вычисления, с
Ручная разметка	100,0	100,0	100,0	176,30
Автоматическое создание маски объекта	88,6	89,7	92,8	0,14
Автоматическое создание маски объекта + экспертная доработка	98,4	98,2	99,0	54,50

Заключительным этапом работы является непосредственно обучение моделей семантической сегментации на сформированном множестве аннотированных данных. Учитывая неравномерное распределение экземпляров исследуемых классов в наборе данных и отличительные характеристики каждого класса в работе, применяется гибридная стратегия сегментации, адаптированная к специфическому распределению и морфологической сложности каждого исследуемого класса.

В результате обучения отдельных моделей сегментации сделан следующий вывод: модели сегментации сосудистого просвета, фиброзной оболочки и липидного ядра демонстрируют четкое и последовательное уменьшение потерь, сопровождающееся прогрессивным улучшением коэффициента Дайса (Dice Coefficient, 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. Представленные вариации подчеркивают присущую сложность и проблемы, связанные с сегментацией данной анатомической структуры подробнее с результатами обучения можно ознакомиться на рисунке 5 и таблице 4. Модель сегментации сосудистого просвета демонстрирует исключительную точность, достигая DSC 0,987 и коэффициент Жаккара (Intersection over Union, IoU) 0,975. Данные показатели подчеркивают способность модели надежно и точно очерчивать структуры класса, подчеркивая ее надежность

и пригодность для клинических приложений, требующих точной сегментации. Модель фиброзной покрышки и липидного ядра 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. Результаты сегментации свидетельствуют о прогрессе в решении этих сложных задач. Интегрируя 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.

Для визуальной оценки точности сегментации в работе представлено два репрезентативных среза, которые иллюстрируют результаты сегментации морфологических признаков бляшек, представленные на рисунке 6.

Таблица 4 – Показатели эффективности сегментации для каждого морфологического признака бляшки (перекрестная проверка)

Признак бляшки	Точность	Полнота	F-мера	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

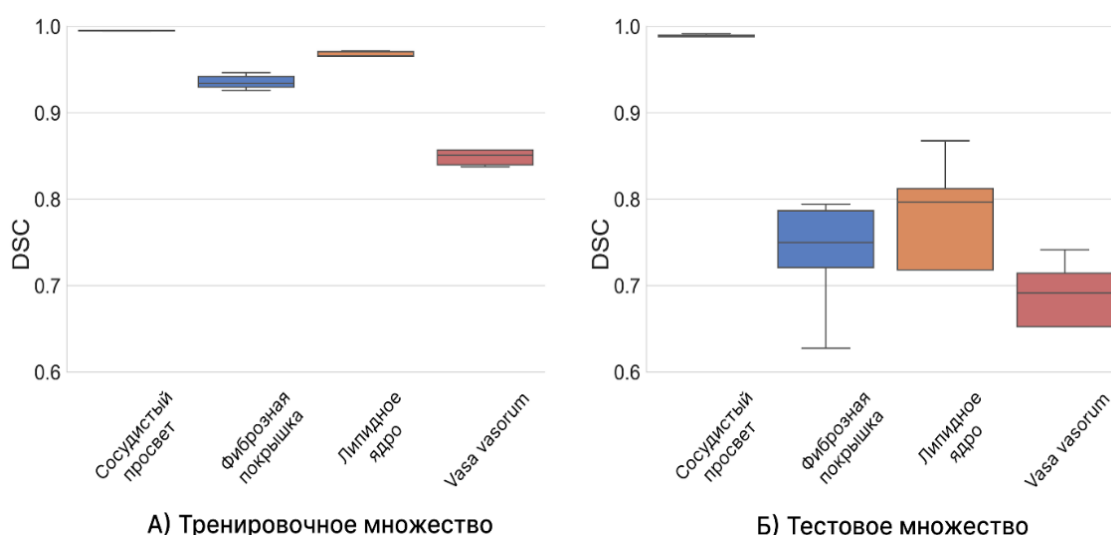


Рисунок 5 – Средняя точность исследуемых моделей при анализе различных морфологических признаков бляшек

Представленные срезы отображают предсказания, полученные оптимальным ансамблем моделей: U-Net++ для сегментации сосудистого просвета, LinkNet для сегментации фиброзной покрышки и липидного ядра и U-Net для сегментации vasa vasorum. Предсказания сопоставляются с соответствующими аннотациями, полученными в ходе экспертной разметки, что позволяет напрямую оценить эффективность модели.

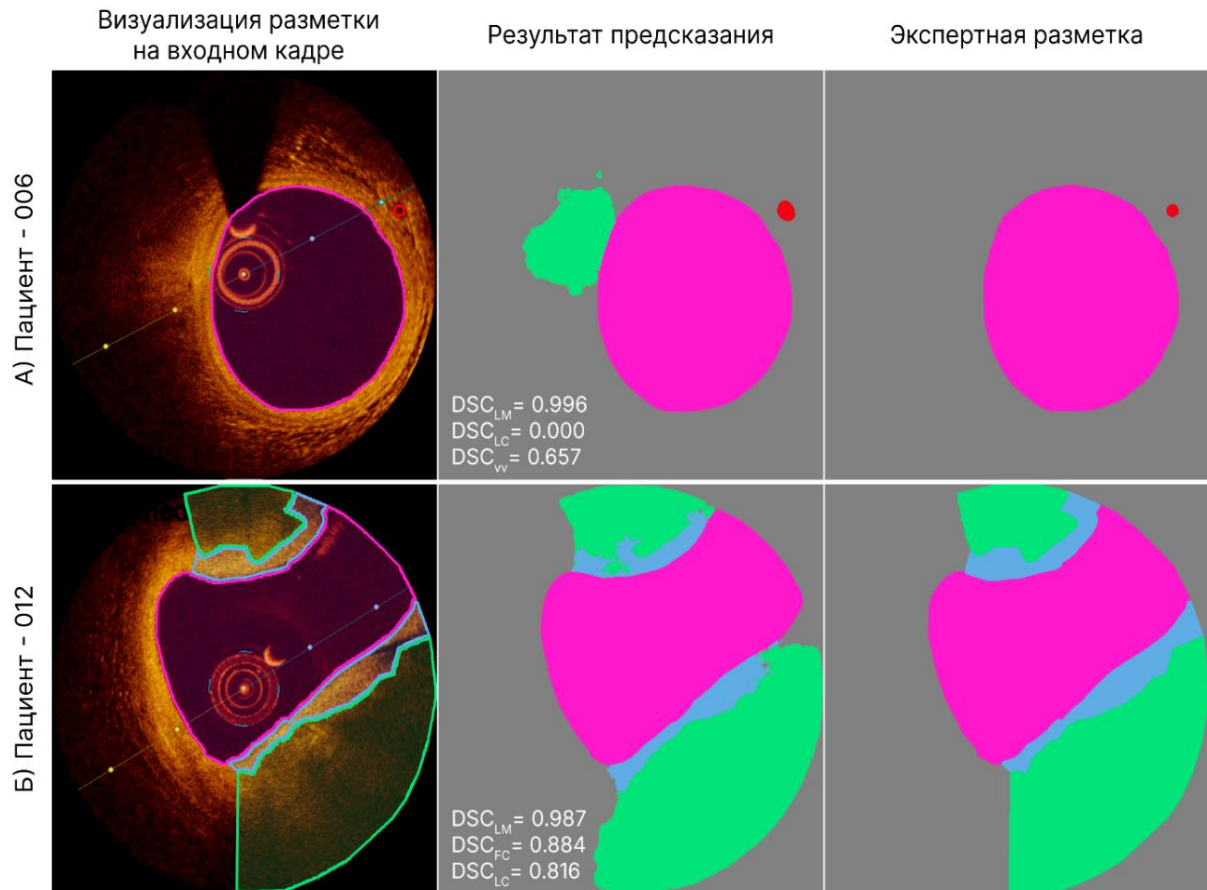


Рисунок 6 – Сравнение экспертной разметки и прогнозов ансамбля

ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ РАБОТЫ

1. Разработана технология автоматизированного формирования сегментационных масок на основе анализа карт признаков классификационной модели, обеспечивающая возможность эффективной подготовки масштабных обучающих наборов данных для сегментации сложных объектов исследования.
2. 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.

3. 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 точностью локализации и адаптивностью к размерам анализируемого изображения, а также повышая точность сегментации границ до 23,7% по сравнению с традиционными методами.
4. 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.
5. Результаты диссертационной работы были использованы при выполнении 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 РФФ № 23-75-10009 «Разработка автоматизированной системы выявления нестабильных бляшек на основе оптической когерентной томографии с использованием методов глубокого обучения». Экспериментальная апробация подтвердила эффективность предложенной технологии на реальных медицинских данных с достижением коэффициента DSC до 0,987 для четко выраженных структур.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в изданиях, рекомендованных ВАК

1. Лаптев В.В. Решение проблемы объединения прогнозов выделенных на изображении объектов дыма / Н.В. Лаптев, О.М. Гергет, А.А. Кравченко, Лаптев В.В., Д.Ю. Колпашиков // Информационные и математические технологии в науке и управлении. – 2022. – №2(26). – С.161-168

2. Лаптев В.В. Применение технологии object detection в задаче обнаружения ключевых точек аортографии / Лаптев В.В., Кочергин Н.А. // Комплексные проблемы сердечно-сосудистых заболеваний. – 2024. – Том 13. – № 1. – С. 144–151 (Scopus)
3. Laptev V.V. Application of Modern Object Tracking Technologies to the Task of Aortography Key Point Detection in Transcatheter Aortic Valve Implantation / V.V. Laptev, N.A. Kochergin // Scientific Visualization. – 2024. – Vol. 16. – № 2. – P.106–115 (Scopus).
4. Laptev V.V. ML-driven segmentation of microvascular features during histological examination of tissue-engineered vascular grafts / V. V. Danilov, V.V. Laptev, K. Yu. Klyshnikov [et al.] // Frontiers in Bioengineering and Biotechnology. – 2024. – Vol. 12 (Scopus, WoS)
5. Laptev V.V. A study of the effectiveness of convolutional neural network models in solving the problem of classifying images of the fundus in terms of the presence of signs of diabetic retinopathy // N.V. Laptev, V.V. Laptev, M.O. Pleshkov, Ya. N. Martusevich, E.G. Starikova, I.V. Tolmachev // Bio-medical Engineering. – 2025. – Vol. 58. – № 6. – P. 403–406 (Scopus, WoS)
6. Laptev V.V. Uncovering unstable plaques: deep learning segmentation in optical coherence tomography // Laptev V.V., Danilov V.V, Ovcharenko E.A, Klyshnikov K.Y, Kolesnikov A.Y, Arnt A.A, Bessonov I.S, Litvinyuk N.V, Kochergin N.A // Computer Optics. – 2025. – Vol. 49. – № 5. – P. 775–793 (Scopus, WoS)
7. Laptev V.V. Segmentation and quantification of atherosclerotic plaques in optical coherence tomography / V. V. Danilov, V.V. Laptev, K. Y. Klyshnikov, I. S. Bessonov [et al.] // Computers in Biology and Medicine. – 2025. – Vol. 197 (Scopus, WoS)

Свидетельства о регистрации программы для ЭВМ

8. Свидетельство о государственной регистрации программы для ЭВМ № 2022619805 (RU); заявка № 2022618392 от 05.05.2022, дата рег. 26.05.2022 // Сонькин Д.М., Плешков А.Г., Лаптев В.В. и др. Многофункциональный мобильный комплекс мониторинга, ситуационного анализа и оповещения труднодоступных объектов с использованием беспилотных летательных аппаратов и ГИС-технологий (ММК МАО).
9. Свидетельство о государственной регистрации программы для ЭВМ № 2025617776 (RU); заявка № 2025616488 от 24.03.2025, дата рег. 28.03.2025 // Онищенко П.С., Овчаренко Е.А., Лаптев В.В., и др. Автоматизированная система сегментации оптической когерентной томографии с использованием методов глубокого обучения.

Публикации в других изданиях

10. Лаптев В.В. Анализ парковочного пространства посредством компьютерного зрения / Лаптев В.В., Лаптев Н.В., Оздиев А.Х., Гергет О.М. // Труды международной конференции по компьютерной графике и зрению «Графикон», 19-21 сентября 2022 г., Рязань – Т. 32. – С. 508-517.
11. Лаптев, В. В. Применение средств машинного обучения для формирования сегментационных масок / Лаптев В. В., Гергет О. М. // Сборник трудов XX Международной научно-практической конференции студентов, аспирантов и молодых учёных «Молодежь и современные информационные технологии», 20-22 марта 2023 г., ТПУ. — Томск — С. 40-41.
12. Лаптев В.В. Визуальное ассистирование при транскатеторной имплантации аортального клапана, на основе технологии «pose estimation» / Лаптев В.В., О.М. Гергет // Труды XX Международной научно-практической конференции «Электронные средства и системы управления», 20–22 ноября 2024 г., ТУСУР. – Томск – Ч. 2. – С. 228-231