

## Методика полуавтоматической разметки данных: снижение трудоемкости формирования масок семантической сегментации

В. В. Лаптев, О. М. Гергет

Томский политехнический университет, Томск, Россия

**Аннотация.** В данной статье рассматривается методика полуавтоматической генерации сегментационных масок, основанная на использовании карт активации, извлекаемых из классификационных моделей глубокого обучения. Проведен анализ современных архитектур сверточных нейронных сетей (ResNet50, EfficientNetB5, Xception и ConvNeXtBase) и методов визуализации значимых признаков (GradCAM, HiResCAM, LayerCAM и др.) для выделения морфологических признаков на изображениях оптической когерентной томографии. Для повышения точности предложен этап постобработки с использованием модели FBA-Net, эффективно уточняющей границы объектов. Эксперименты показали, что предложенный подход позволяет достичь метрики IoU до 71,3 % на объектах с четкими структурными признаками, что существенно снижает трудоемкость разметки. Метод особенно актуален для медицинской визуализации, где объем ручной аннотации ограничен.

**Ключевые слова:** семантическая сегментация, маски изображений, карта внимания, карта признаков, сверточные нейронные сети, CAM, FBA-Net.

## Semi-Automatic generation of segmentation masks using deep learning activation maps

V. V. Laptev, O. M. Gerget

Tomsk Polytechnic University, Tomsk, Russia

**Abstract.** This article discusses a methodology for semi-automatic generation of segmentation masks based on the use of activation maps extracted from deep learning classification models. We conduct an analysis of modern convolutional neural network architectures (ResNet50, EfficientNetB5, Xception, and ConvNeXtBase) and salient feature visualization methods (GradCAM, HiResCAM, LayerCAM and others) for highlighting morphological features in optical coherence tomography images. To improve accuracy, a post-processing stage using the FBA-Net model is proposed, effectively refining object boundaries. Experiments demonstrated that the proposed approach achieves an Intersection over Union (IoU) metric of up to 71,3 % on objects with clear structural features, significantly reducing annotation workload. The method is particularly relevant for medical imaging, where manual annotation availability is limited.

**Keywords:** semantic segmentation, image masks, activation map, feature map, convolutional neural networks, CAM, FBA-Net.

### Введение

Сегментация изображений представляет собой одну из фундаментальных задач компьютерного зрения, направленную на разделение изображения на семантически значимые области. В медицинской визуализации эта задача приобретает критическую важность, так как точное выделение патологий, органов или анатомических структур напрямую влияет на качество диагностики и планирования лечения. Современные методы сегментации, основанные на глубоком обучении, демонстрируют высокую эффективность, однако их внедрение сталкивается с рядом ограничений, связанных с необходимостью масштабных аннотированных данных. Подготовка таких данных требует значительных временных и экспертных ресурсов, особенно в условиях пиксельной точности, необходимой для медицинских исследований. Например, разметка одного изображения с поражениями COVID-19 занимает до 20 минут, а для наборов из тысяч изображений трудозатраты достигают сотен часов [1]. Для сравнения: подготовка аналогичного объема данных для классификации требует на порядок меньше времени, что создает существенный дисбаланс в развитии методов анализа изображений. Эта проблема усугубляется в условиях ограниченности экспертных ресурсов и возрастающих требований к качеству медицинской диагностики, где ошибки сегментации напрямую влияют на достоверность клинических выводов.

Традиционные подходы к сегментации можно условно разделить на методы, основанные на анализе гистограмм, кластеризации, пороговой обработке и нейронных сетях. Последние, благодаря способности к обобщению и адаптации, становятся доминирующими, однако их эффективность напрямую зависит от качества и объема обучающих данных. В этом контексте актуальной остается

проблема разработки методов, минимизирующих зависимость от ручной разметки. Перспективным направлением является использование парадигм слабо контролируемого обучения, таких как обучение с частичными аннотациями или автоматическая генерация масок на основе промежуточных результатов классификационных моделей.

Актуальность работы определяется необходимостью минимизировать зависимость современных алгоритмов сегментации от ручной разметки без ущерба для их точности. Большинство существующих подходов, таких как U-Net [2] или DeepLab [3], демонстрируют высокую эффективность только при наличии полноценно размеченных данных. Однако в реальных сценариях, особенно в медицине, получение таких данных часто затруднено из-за этических, временных или финансовых ограничений. В качестве альтернативы предлагается использовать классификационные модели, обученные на слабо аннотированных данных (например, бинарных метках наличия/отсутствия объекта), для генерации первичных сегментационных масок. Ключевая гипотеза заключается в том, что карты активации признаков, извлекаемые из промежуточных слоев таких моделей, содержат достаточно информации для выделения областей интереса, а их последующее уточнение с помощью специализированных архитектур позволит достичь качества, сопоставимого с полностью контролируруемыми методами.

### Постановка задачи

В данной работе предложен инновационный подход, сочетающий преимущества классификационных моделей и методов постобработки для формирования сегментационных масок. Основная гипотеза исследования заключается в том, что карты активации признаков, извлекаемые из глубоких нейронных сетей, могут служить основой для генерации масок «низкого качества», которые затем могут быть уточнены с помощью специализированных архитектур. Для этого проведен сравнительный анализ методов активационного картирования (Class Activation Mapping, CAM), включая GradCAM [4], HiResCAM [5], GradCAM++ [6] и др., направленных на визуализацию значимых областей изображения. Эти методы, изначально разработанные для интерпретации решений нейросетей, адаптированы для формирования первичных масок, что позволяет избежать трудоемкого этапа ручной разметки. Ключевым этапом предложенного метода является использование трехканальной карты (Trimap), структурирующей изображение на три семантические зоны:

- детерминированный передний план;
- фон;
- неопределенную область.

Trimap формализует пространственные отношения между объектом и окружением, что особенно важно для медицинских данных с низкой контрастностью и зашумленностью. Для финального уточнения масок применена модель FBA-Net (Feature Boundary-Aware Network) [7], сочетающая механизмы внимания к границам и семантическим признакам. Архитектура FBA-Net, интегрирующая подходы U-Net и attention-механизмы, демонстрирует высокую эффективность в задачах, где критична точность определения контуров, таких как выделение опухолей или органов в условиях артефактов изображения.

Основная задача исследования формулируется следующим образом: разработка метода автоматической генерации сегментационных масок на основе карт активации классификационных моделей с их последующей оптимизацией с использованием FBA-Net.

### Теория

Исходные данные 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 исследований пациентов мужского пола и 26 исследований 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.

Полученные 103 ОКТ-исследования, представленные 25698 RGB-срезами, имеют четыре морфологических признака атеросклеротической бляшки: сосудистый просвет (Lumen, LM),

фиброзная капсула (Fibrous cap, FC), липидное ядро (Lipid core, LC) и vasa vasorum (VV) [8]. Представленные характеристики имеют решающее значение для понимания состава бляшки и ее клинических последствий, в частности для оценки риска сердечно-сосудистых событий.

Все выявленные морфологические характеристики атеросклеротических бляшек играют важную роль в понимании структуры и прогрессирования атеросклеротического заболевания. Каждая характеристика дает критическое представление о различных аспектах 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 покрывка, липидное ядро и vasa vasorum являются важными маркерами для оценки уязвимости бляшки.

Предлагаемая методика представляет собой полуавтоматический подход к генерации сегментационных масок на основе карт активации, извлекаемых из классификационных моделей глубокого обучения, с последующей постобработкой. Методика состоит из четырех основных этапов, показанных на рисунке 1.

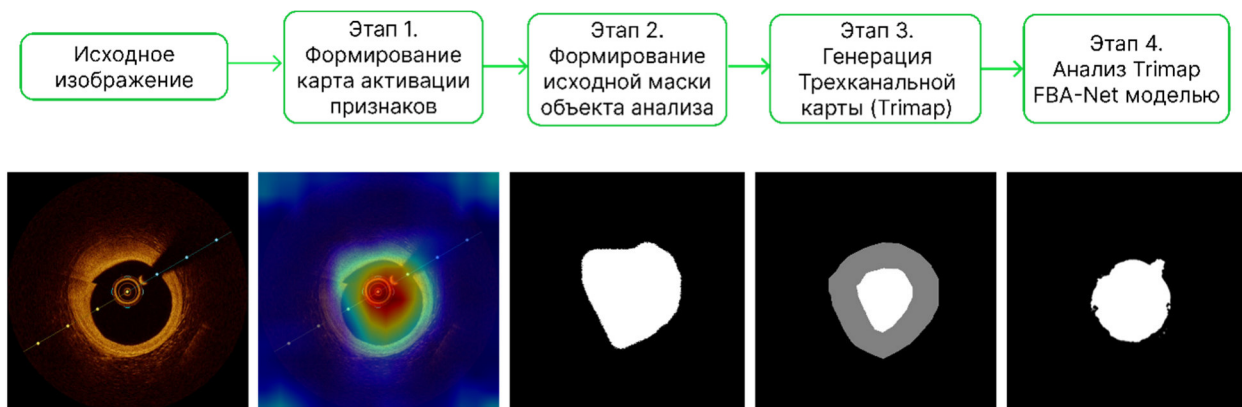


Рис. 1. Методика полуавтоматической генерации сегментационных масок

*Карта признаков.* Задача классификации – это получение категориального ответа на основе набора признаков. Концептуальная схема работы модели классификации изображения на примере сверточной нейронной сети приведена на рисунке 2.

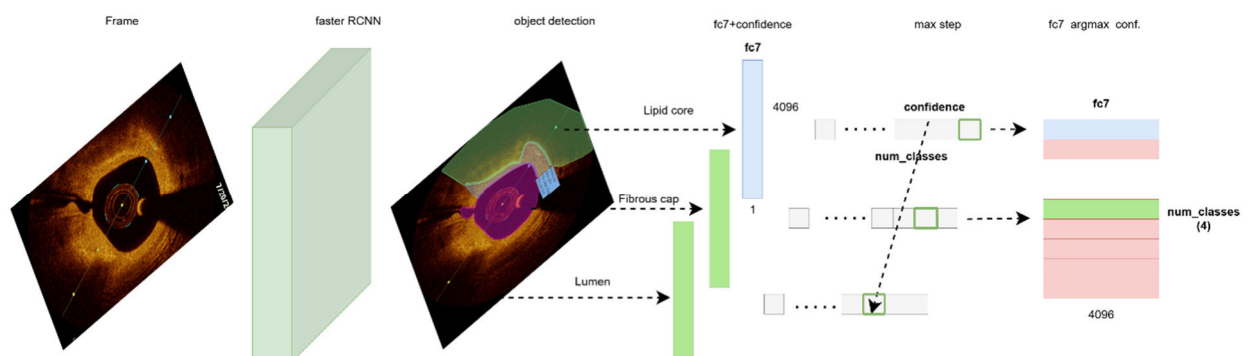


Рис. 2. Концептуальная схема работы нейросетевой модели классификации (на примере модели Faster RCNN)

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 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 генерации первичных сегментационных масок на основе классификационных моделей были

проанализированы и адаптированы методы активационного картирования (Class Activation Mapping, CAM), позволяющие визуализировать области изображения, наиболее значимые для предсказания модели. В исследовании рассмотрены 9 различных методов: GradCAM [4], HiResCAM [5], GradCAMElementWise [4], GradCAM++ [6], XGradCAM [9], EigenCAM [10], EigenGradCAM [11], ScoreCAM [12] и LayerCAM [13]. Перечисленные методы, основанные на анализе последнего сверточного слоя были выбраны из-за их способности локализовать морфологические признаки на ОКТ-изображениях, несмотря на низкую контрастность и наличие спекл-шума.

Полученные результаты (рис. 3) представляют собой вариант некой «тепловой карты следования за объектом», а также имеют схожее описание с сегментационной маской изображения. Данный результат позволяет выдвинуть гипотезу, что вектор признаков нейросетевой модели классификации можно интерпретировать как сегментационную маску изображения с обозначением позиции класса анализа. Полученная описанным способом маска изображения не имеет четких границ, а лишь примерно определяет местоположение анализируемого класса. С целью уточнения границ объекта в исследовании предложен подход, основанный на анализе Trimap карты FBA-Net моделью.

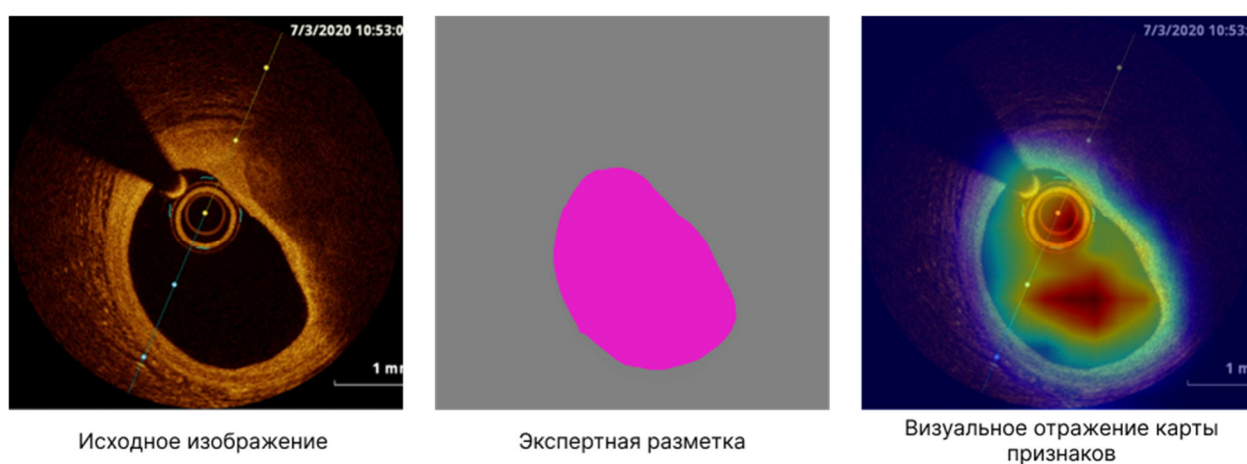


Рис. 3. Визуализация карты признаков (на примере анализа класса Lumen моделями Xception и LayerCAM)

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 [14] или методах, основанных на глубоком обучении, Trimap используется как входной сигнал для обучения нейросетей предсказанию альфа-каналов. Данный подход 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, применяя 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 неопределённую область.

3. Создание маски Trimap: на основе выделенных областей создаётся маска, где каждому пикселю присваивается одно из трёх значений: 0 (фон), 128 (неопределённая область), 255 (объект).

Для решения поставленной задачи Trimap будет использоваться в качестве валидатора. Полученные на предыдущем этапе маски изображений, сформированные из карты активации классификационной модели, подвергаются следующим преобразованиям:

1. Изображение приводится к размеру 1024x1024px.
2. Последовательное применение морфологических операций открытие и закрытие для устранения случайных шумов и незначительных пропусков.
3. Создается копия маски изображения.
4. Один из вариантов подвергается морфологической операции – дилатация. Данная маска в будущем будет представлять вариант Unknown.
5. Второй вариант подвергается морфологической операции – эрозия. Данная маска в будущем будет представлять вариант Foreground.

6. Заключительным этапом является наложение полученных масок с использованием представленной выше цветовой схемы.

Полученное Trimap изображение анализируется FBA-Net моделью [7]. 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 счёт 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.

Интеграция FBA-Net 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.

### Результаты эксперимента

Для решения поставленной задачи – определения 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 [15, 16]. 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 до обработки 3D-данных [17]. Для достижения качественного результата за кратчайшие сроки мы использовали технологию трансферного обучения [18]. С помощью моделей, обученных на 1,4 млн изображений ImageNet и хранящих веса для определения 17000 классов, в кратчайшие сроки были подготовлены модели классификации.

Все модели обучались в равных условиях с заданным набором гиперпараметров, представленным в таблице 1. Обучение проводилось в течение 75 эпох с использованием предварительно размеченных 25698 изображений. Результаты, полученные в ходе тестирования моделей, приведены в таблице 2.

Таблица 1. Гиперпараметры, использованные для обучения моделей классификации

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

Таблица 2. Результаты работы моделей классификации на тестовом множестве

Model name	Accuracy, %	Precision, %	Recall, %
ResNet-50	74,3	70,5	70,3
EfficientNetB5	78,2	68,4	67,2
Xception	97,8	91,7	91,7
ConvNeXtBase	69,9	65,4	65,7

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 ч, 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. Полученные сравнительные карты активации демонстрируют вариативность работы алгоритма CAM для признаков разного размера и сложности. Стоит отметить, что метрическая оценка не проводилась, так как полученные карты активации не позволили сформировать сегментационные маски с четкими очертаниями. В связи с этим было принято решение провести качественную оценку результатов экспертами, фокусируясь на способности методов локализовать морфологические признаки. 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. Данное ограничение, вероятно, связано со сложной морфологией липидного ядра, нечеткими границами и встраиванием в окружающие ткани.

- 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.



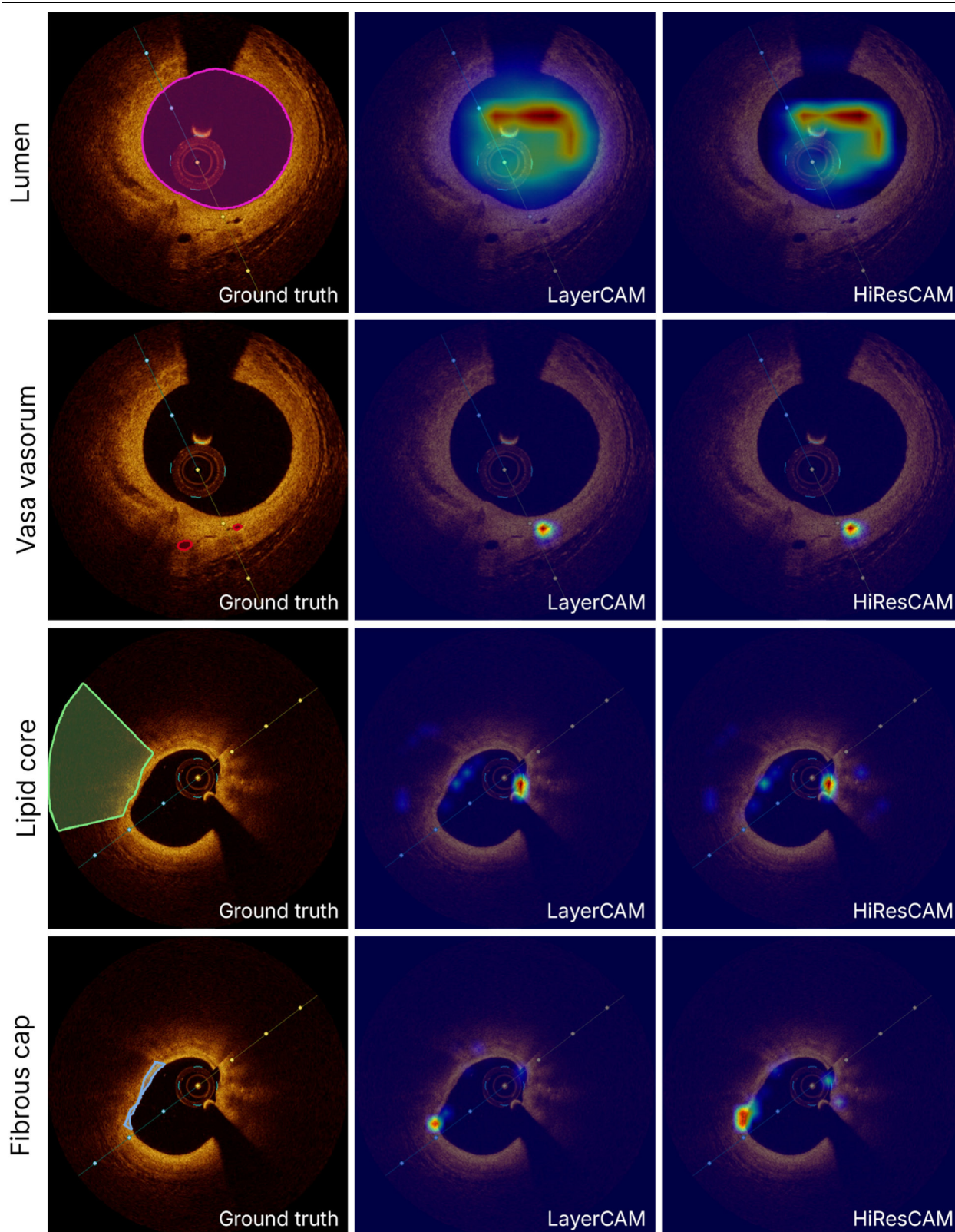


Рис. 4. Результат анализа карты признаков классификационной модели методами LayerCAM и HiResCAM

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 дает возможность составить большой обучающий набор данных и повысить уверенность в клинической применимости будущей модели. Визуализация результата для методов LayerCAM и HiResCAM приведена на рисунке 4.

Заключительным этапом формирования сегментационной маски изображения является уточнение границ анатомической структуры. Представленный в работе алгоритм формирования трехканальной маски отлично подходит для решения поставленной задачи. Визуализация этапа создания трехканальной маски и результата применения FBA-Net модели по результатам работы метода LayerCAM представлена на рисунке 5.

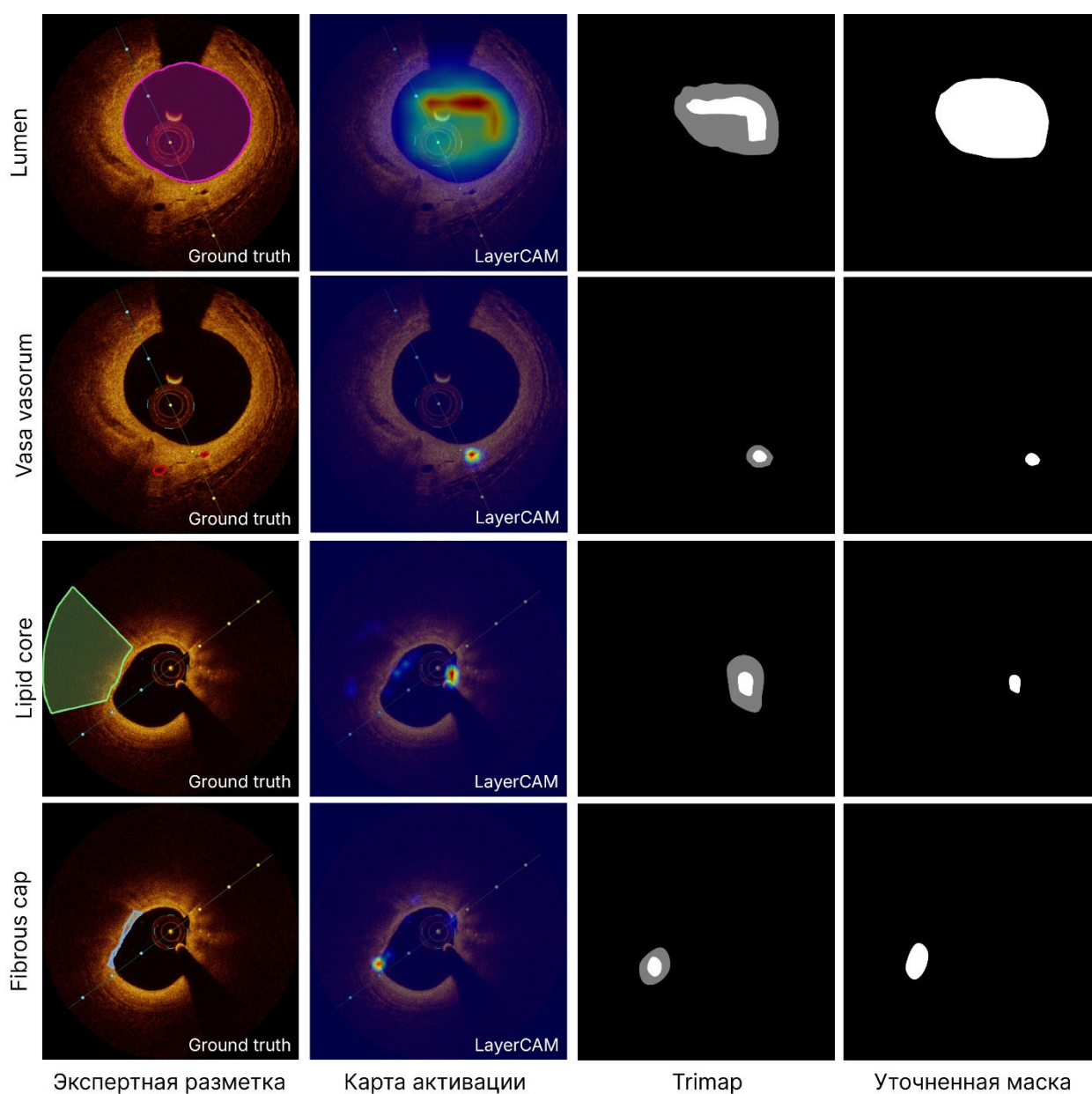


Рис. 5. Визуализация этапов работы представленного метода

Из результатов, представленных на рисунке 5, видно, что представленная методика формирования сегментационных масок изображений подходит для тех объектов исследования, которые имеют выраженные очертания, такие как Lumen и Vasa vasorum. Для объектов, что 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, представленные в таблице 3, данная методика является эффективным инструментом формирования обучающего множества.

Таблица 3. Показатели эффективности метода формирования сегментационных масок изображений на примере классов Lumen и Vasa vasorum в расчете на одно изображение

Algorithm	IoU, %	Speed, s
Ручная разметка	100,0	176,30
Автоматическое создание маски объекта	71,3	0,14
Автоматическое создание маски объекта + экспертная доработка	96,7	54,50

Согласно результатам, представленным в таблице 3, предложенная в работе методика формирования сегментационных масок изображений оказалась в 3 раза эффективнее ручного аннотирования. Также стоит отметить, что для формирования качественного обучающего множества, как правило, требуется не менее трех экспертов, что обеспечивает объективность 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.

### Обсуждение результатов

Экспериментальная часть работы подтверждает, что интеграция автоматически генерируемых масок и их последующее уточнение с помощью FBA-Net позволяет достичь метрики Intersection over Union (IoU) на уровне 87,3 % для классов объектов с выраженными структурными признаками. Это свидетельствует о потенциале комбинирования методов интерпретируемости глубоких нейронных сетей и продвинутых архитектур сегментации для снижения зависимости от ручной разметки. Полученные результаты открывают новые возможности для применения подхода в областях, где экспертные аннотации ограничены, включая онкологию, кардиологию и анализ гистологических срезов.

Работа вносит вклад в развитие методов слабо контролируемой сегментации, предлагая практический метод для минимизации трудозатрат на подготовку данных без ущерба для точности. Дальнейшие исследования могут быть направлены на адаптацию метода для работы с мультимодальными медицинскими данными и оптимизацию вычислительной эффективности архитектуры FBA-Net.

Научная новизна подхода заключается в комбинации интерпретируемости классификационных моделей и продвинутых методов постобработки. В отличие от традиционных методов слабого обучения, где маски генерируются эвристически, предложенный метод опирается на данные, извлеченные непосредственно из внутренних представлений нейросети, что повышает их семантическую согласованность. Использование FBA-Net, в свою очередь, позволяет компенсировать недостатки первичных масок за счет фокуса на граничных признаках и механизмах внимания.

### Заключение

В рамках исследования разработан метод полуавтоматической генерации масок семантической сегментации, сочетающий интерпретируемость классификационных моделей и постобработку с использованием FBA-Net. Данный подход позволяет формировать обучающие данные с приемлемым качеством при значительном снижении временных затрат на ручную разметку более чем в 3 раза. Это

особенно критично в области медицинской визуализации, где качество аннотаций напрямую влияет на клинические выводы.

Проведенные эксперименты продемонстрировали высокую точность на объектах с выраженными морфологическими границами (Lumen, Vasa vasorum), достигнув метрики IoU до 71,3 % в автоматическом режиме и до 96,7 % с экспертной доработкой. В то же время для объектов с нечеткими границами (Fibrous cap, Lipid core) методика демонстрирует качество ниже, чем традиционные подходы, что открывает перспективы для его дальнейшей оптимизации.

Предложенная методология обладает высоким потенциалом для применения в условиях ограниченных ресурсов аннотирования и может быть адаптирована к другим медицинским задачам и мультимодальным данным. В будущем планируется расширение метода для работы с 3D-данными и снижение вычислительной нагрузки архитектуры FBA-Net, что сделает его еще более применимым в клинической практике.

### Источник финансирования

Исследование выполнено при поддержке Министерства науки и высшего образования Российской Федерации, проект «Наука FSWW-2023-0007».

### Список литературы

1. Временные методические рекомендации «Профилактика, диагностика и лечение новой коронавирусной инфекции (COVID-19)». Версия 19 (27.05.2025). URL: [https://www.consultant.ru/document/cons\\_doc\\_LAW\\_347896](https://www.consultant.ru/document/cons_doc_LAW_347896).
2. 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.
3. 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.
4. 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.
5. 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. Available: <http://arxiv.org/abs/2011.08891>
6. 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.
7. 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. Available: <http://arxiv.org/abs/2306.15189>
8. 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.
9. 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. Available: <http://arxiv.org/abs/2008.02312>
10. 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.
11. 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.
12. 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.
13. 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>
14. 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. Available: <https://arxiv.org/abs/1807.10088v1>

- 
15. 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.
  16. 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.
  17. B.-E. Munkhsaikhan, “SOEs in China and AI Development: Legal Constraints for Foreign Investments towards China&#8217;s Digital Transformation,” *Open J Soc Sci*, vol. 12, no. 12, pp. 146–157, 2024, doi: 10.4236/JSS.2024.1212010.
  18. 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. Available: <http://arxiv.org/abs/1812.05313>