

Автоматическая сегментация гемангиомы печени на МРТ с использованием модифицированной архитектуры U-Net

Александр Кадан,

заведующий кафедрой системного программирования и компьютерной безопасности ГрГУ им. Янки Купалы, кандидат технических наук, доцент;
kadan@mf.grsu.by

Сергей Петров,

доцент кафедры системного программирования и компьютерной безопасности ГрГУ им. Янки Купалы, кандидат медицинских наук, магистр технических наук, доцент;
petrov@mf.grsu.by

Андрей Прокопович,

врач лучевой диагностики кабинета магнитно-резонансной томографии Гродненской университетской клиники, старший преподаватель кафедры лучевой диагностики и терапии Гродненского государственного медицинского университета;
andrey.doctor@gmail.com

Олег Зеньков,

заместитель главного врача Гродненской университетской клиники по организационно-методической работе

Артем Стурейко,

инженер электросвязи РУП Белтелеком, Гродненский филиал;
sturejko_aa_21@mf.grsu.by

Гемангиома печени – самая распространенная доброкачественная опухоль этого органа, встречающаяся, по разным данным, у 5–20% населения. Чаще всего эти образования небольшие, бессимптомные и обнаруживаются случайно при визуализации брюшной полости по другим показаниям. Однако гигантские гемангиомы (более 5 см) несут риск серьезных осложнений, включая болевой синдром, чувство переполнения в животе и жизнеугрожающие состояния: спонтанный разрыв с кровотечением, коагулопатию потребления (синдром Казабаха – Мерритт) или обструкцию желчных путей и крупных сосудов [1].

Аннотация. В работе представлен разработанный авторами метод автоматической сегментации гемангиомы печени на магнитно-резонансных снимках, основанный на модификациях сверточной архитектуры U-Net. Проведено экспериментальное сравнение базовой U-Net, U-Net++ и Attention U-Net. Подтверждено, что модель Attention U-Net с пространственно-канальным механизмом внимания (SCSE) и комбинированной функцией потерь (Dice + Focal Loss) достигает наилучших результатов, демонстрируя средний коэффициент Дайса (DSC) 84,65% на приватном наборе данных. Этот показатель сопоставим с результатами современных аналогов, особенно при работе с данными, характеризующимися значительным дисбалансом классов. Для апробации метода в условиях, приближенных к клиническим, разработан прототип программного обеспечения с графическим интерфейсом и возможностью интеграции в PACS-системы через протокол DICOMweb.

Ключевые слова: сегментация медицинских изображений, гемангиома печени, магнитно-резонансная томография, сверточные нейронные сети, глубокое обучение, МРТ, U-Net, Attention U-Net, DICOM, PACS.

Для цитирования: Кадан А., Петров С., Прокопович А., Зеньков О., Стурейко А. Автоматическая сегментация гемангиомы печени на МРТ с использованием модифицированной архитектуры U-Net // Наука и инновации. 2026. №4. С. 78–83.

<https://doi.org/10.29235/1818-9857-2026-04-78-83>

Точная оценка морфологических параметров опухоли – ее размеров, объема, локализации и взаимоотношения с окружающими структурами – критически важный этап диагностики, определяющий оптимальную тактику ведения, которая варьируется от динамического наблюдения до активного вмешательства, такого как трансартериальная химиоэмболизация (ТАХЭ) или хирургическая резекция [1].

Традиционно перечисленные параметры оцениваются врачами лучевой диагностики путем ручной или полуавтоматической сегментации опухоли. Однако у такого подхода есть существенные недостатки. Во-первых, он трудоемок (15–30 мин на исследование), что неприемлемо при большом потоке пациентов. Во-вторых, субъективен в случае ручной сегментации: меж- и внутриэкспертная вариабельность результатов достигает 15%, что снижает воспроизводимость и надежность оценок [2]. Наконец, метод плохо масштабируется, создавая затруднения в радиологическом процессе.

Магнитно-резонансная томография (МРТ) – золотой стандарт диагностики очаговых образований печени. Ее преимущества – высокая контрастность мягких тканей, отсутствие ионизирующего излучения и характерные для гемангиом сигнальные паттерны: выраженная гиперинтенсивность на T2-взвешенных изображениях и специфическая динамика накопления гадолиниевого контраста [3]. Эти особенности делают МРТ идеальной основой для разработки систем автоматизированного анализа.

Развитие глубокого обучения открыло новые возможности для автоматизации анализа медицинских изображений. Сверточ-

ные нейронные сети (например, U-Net) показали высокую эффективность в задачах семантической сегментации. Цель данной работы – разработка и валидация метода автоматической сегментации гемангиом печени на снимках МРТ на базе модифицированной архитектуры U-Net с оценкой его потенциала для интеграции в клиническую практику.

Архитектура U-Net: теоретические основы и модификации

Это уникальное решение, предложенное Олафом Роннебергером и соавторами в 2015 г., де-факто стало стандартом для многочисленных исследований [4]. Успех архитектуры обусловлен решениями, которые идеально подходят для специфики работы с медицинскими изображениями, где часто имеет место ограниченный набор данных, но требуется высокая точность локализации.

Популярность U-Net в биомедицинской сегментации объясняется синергией ее элементов. В основе лежит симметричная U-образная структура с энкодером (сжимающий путь) и декодером (расширяющий путь). Первый последовательно уменьшает разрешение, извлекая абстрактные контекстные признаки, а второй выполняет обратную задачу, восстанавливая разрешение для построения итоговой карты сегментации.

Ключевое нововведение U-Net – пропускные соединения (skip-connections). Они напрямую связывают слои энкодера и декодера, передавая низкоуровневые признаки высокого разрешения (точные границы, текстуры) в обход глубоких слоев сети. Этот короткий путь предотвращает

потерю деталей и обеспечивает высокую точность локализации, что является фундаментальным требованием для медицинской визуализации.

Наконец, еще два свойства закрепляют практическую ценность U-Net. Ее полносверточная природа делает ее гибкой к размерам входных данных, а небольшое количество параметров, по сравнению с другими классификационными сетями, обеспечивает высокую эффективность при обучении на малых данных. Эта способность избежать переобучения уже на сотнях, а не тысячах изображений, стала ключевым фактором ее успеха в медицинской сфере, где сбор больших размеченных датасетов затруднен.

На основе классической U-Net предложен ряд модификаций для решения ее ограничений и повышения производительности. В контексте сегментации изображений печени и ее поражений наиболее значимы представленные в *табл. 1* [5, 6].

Учитывая характеристики стоящей перед авторами задачи – сегментации на 2D МРТ-снимках с нечеткими границами и значительным влиянием окружающих тканей, для экспериментального сравнения были выбраны следующие архитектуры:

- **базовая U-Net** в качестве эталонной модели для оценки прироста производительности от модификаций;
- **U-Net++**: использует вложенные и плотные skip-connections для улучшения градиентного потока и агрегации признаков на разных масштабах. Такой подход полезен для сегментации гемангиом, размер которых может значительно варьироваться;

- **Attention U-Net (с модулем SCSE):** механизм пространственно-канального внимания (*Spatial and Channel Squeeze and Excitation*) позволяет модели адаптивно взвешивать важность отдельных областей изображения (пространственное внимание) и целых карт признаков (канальное внимание). Ожидается, что это позволяет эффективнее акцентировать внимание на признаках гемангиомы и игнорировать окружающий фон.

Материалы и методы

Для разработки и валидации моделей сегментации разработана методология, включающая подготовку набора данных, конвейер предобработки и аугментации, определение функций потерь, метрик и параметров экспериментальной среды. Исследование было основано на приватном наборе анонимизированных МРТ-данных брюшной полости 218 пациентов из архива Гродненской университетской клиники: серий аксиальных срезов в различных последовательностях: T1-взвешенные, T2-взвешенные и постконтрастные с динамическим усилением (DCE-MRI).

Эталонные аннотации (ground truth) созданы путем ручной попиксельной сегментации гемангиом. Разметку выполняли два независимых эксперта-рентгенолога с опытом более 5 лет, финальная маска формировалась после согласования спорных случаев. Исходные DICOM-изображения имели разрешение 512×512 пикселей. Для обучения преимущественно использовались T2-взвешенные варианты, так как на них

Модификация	Ключевые особенности	Преимущества для сегментации печени
3D U-Net	Использование 3D-сверток и 3D-пулинга для обработки объемных данных	Позволяет учитывать пространственный контекст между срезами, что улучшает когерентность сегментации и помогает отличать опухоли от сосудов
nnU-Net	Фреймворк для автоматической конфигурации и настройки гиперпараметров U-Net под конкретный датасет	Демонстрирует значительное превосходство над базовыми реализациями (например, DSC 84,7% vs 80,1% у базовой 3D U-Net на задаче сегментации опухолей печени), адаптируясь к особенностям данных
Attention U-Net	Интеграция весовых карт внимания (attention gates) в skip-connections	Модель учится «фокусироваться» на релевантных областях и подавлять признаки из нерелевантного фона, что особенно важно при низком контрасте между опухолью и здоровой паренхимой печени
ResU-Net	Внедрение остаточных блоков (residual blocks) из архитектуры ResNet	Облегчает обучение более глубоких сетей, предотвращая проблему затухания градиента и позволяя извлекать более сложные иерархические признаки

Таблица 1. Сравнение модификаций классической архитектуры U-Net

гемангиомы обладают характерным гиперинтенсивным сигналом, что облегчает их идентификацию. Набор данных отличается выраженным дисбалансом классов: пиксели гемангиом составляют менее 5% на срезах с патологией, а число срезов без нее значительно превышает число с ней. Эта особенность требует специальных подходов при обучении моделей.

Для стабилизации этого процесса и повышения обобщающей способности моделей реализован конвейер предобработки и аугментации данных. Для их унификации и устранения вариативности, связанной с настройками оборудования, все изображения проходили двухэтапную нормализацию. Сначала для устранения влияния экстремальных значений (артефактов или шума) значения интенсивности пикселей отсекались по 1-му и 99-му перцентилем их распределения. Затем проводилась Z-score стандартизация, при которой из каждого значения интенсивности вычиталось среднее и делилось на стандарт-

ное отклонение, рассчитанные по всему объему ткани.

Для расширения обучающей выборки и повышения устойчивости модели к искажениям применялись случайные преобразования в реальном времени. Геометрические аугментации включали повороты ($\pm 15^\circ$), аффинные сдвиги ($\pm 10\%$), масштабирование (0,9–1,1) и горизонтальные отражения; фотометрические – случайную коррекцию яркости и контрастности ($\pm 20\%$). Такой комплексный подход был ключевым для предотвращения переобучения модели на небольшом наборе данных.

Учитывая сильный дисбаланс классов, выбор стандартной функции потерь, такой как перекрестная энтропия, неэффективен. Поэтому в исследовании использовались комбинированные функции, специально разработанные для задач сегментации: сочетание Dice Loss + Focal Loss, включающее преимущества Dice Loss, напрямую оптимизирующей метрику перекрытия, и Focal Loss, придающей боль-

ший вес сложным для классификации пикселям (например, на границах опухоли) и снижающей влияние легко классифицируемого фона, и Dice Loss + Lovasz-Softmax Loss (последняя является суррогатной функцией для прямой оптимизации метрики Jaccard (IoU) и особенно эффективна для сегментации сложных, невыпуклых объектов).

Для количественного сравнения производительности моделей был задействован набор стандартных метрик, рекомендованных для задач семантической сегментации:

коэффициент Дайса (DSC):
 $DSC = 2 * |X \cap Y| / (|X| + |Y|);$

индекс Жаккарда (IoU):
 $IoU = |X \cap Y| / |X \cup Y|;$

точность (Precision):
 $Precision = TP / (TP + FP);$

полнота (Recall):
 $Recall = TP / (TP + FN),$

где X – предсказанная маска, Y – истинная маска, TP – истинно-положительные, FP – ложноположительные, FN – ложноотрицательные пиксели.

Все эксперименты проводились с применением фреймворка PyTorch. В качестве оптимизатора использовался Adam с начальной скоростью обучения (learning rate) 1e-3 и стандартными пара-

метрами. Обучение моделей производилось на вычислительной станции, оснащенной графическим ускорителем NVIDIA Tesla V100 с 32 Гб видеопамяти. Каждая модель обучалась на протяжении 200 эпох с задействованием механизма ранней остановки (early stopping): процесс прекращался, если значение функции потерь на валидационной выборке не улучшалось в течение 30 последовательных эпох.

Результаты и обсуждение

В ходе исследования была проведена количественная и качественная оценка трех архитектур: базовой U-Net, U-Net++ и Attention U-Net. Модели обучались и тестировались в идентичных условиях для обеспечения объективности сравнения.

Количественные результаты, усредненные по тестовой выборке, представлены в табл. 2. Они демонстрируют явное преимущество модифицированных архитектур над базовой версией U-Net.

Модель Attention U-Net с модулем SCSE показала наилучшие результаты по всем ключевым метрикам. Она достигла среднего значения DSC, равного 84,65%, что на 3,42 процентных пункта (п.п) выше, чем у базовой U-Net, и на 1,68 п.п. выше, чем у U-Net++. Важно отметить,

что улучшение наблюдается как по метрике Precision (снижение ложноположительных срабатываний), так и по метрике Recall (снижение пропусков опухолей). Это свидетельствует о том, что механизм внимания позволяет модели достичь лучшего баланса и более надежно идентифицировать целевые области.

Качественный анализ результатов сегментации подтверждает выводы, сделанные на основе количественных метрик. На рисунке представлены примеры работы модели Attention U-Net в сравнении с эталонной разметкой на сложных клинических случаях. Она демонстрирует более высокую робастность при работе с гемангиомами, имеющими неоднородную внутреннюю структуру или нечеткие, размытые границы. Благодаря механизму внимания модель научилась лучше отличать ткань опухоли от прилегающих сосудистых структур и паренхимы печени, что привело к снижению числа ложноположительных срабатываний на 18% по сравнению с базовой U-Net. В то же время улучшенная чувствительность позволила ей корректно идентифицировать небольшие участки опухоли, которые пропускались другими моделями.

Для позиционирования полученных результатов в контексте современных достижений было проведено их сравнение с данными из релевантных публикаций (табл. 3).

Полученный нами результат (DSC 84,7%) является высококонкурентоспособным. Он незначительно уступает методу [7] (86,9%), где использовалась более сложная двухэтапная 3D-архитектура, требующая значительно больших вычислительных ресурсов, но существенно превосходит показатель, полученный

Модель	DSC (коэф. Дайса)	IoU (индекс Жаккарда)	Precision (точность)	Recall (полнота)
Базовая U-Net	0,8123	0,6915	0,834	0,782
U-Net++	0,8297	0,7152	0,847	0,801
Attention U-Net (SCSE)	0,8465	0,7338	0,861	0,832

Таблица 2. Сравнительная оценка производительности моделей сегментации на тестовом наборе данных. Лучшие значения выделены жирным шрифтом

Hänsch et al. [9] (74,0%), что демонстрирует эффективность выбранного нами подхода к архитектуре и обучению. Важно подчеркнуть, что высокая производительность была достигнута на данных с выраженным классовым дисбалансом, что делает задачу объективно более сложной и приближенной к реальным условиям.

Все это свидетельствует о высокой эффективности предложенного подхода на основе архитектуры Attention U-Net для автоматической сегментации гемангиом печени, что обусловлено тремя факторами:

- **центральным элементом стал механизм внимания.** Интеграция модуля SCSE реализовала принцип адаптивной фокусировки [5, 6]: модель учится взвешивать значимость признаков, динамически усиливая харак-

терные для гемангиомы и ослабляя фоновые сигналы. Это объясняет значительное снижение ложноположительных срабатываний – главной проблемы базовых архитектур на низкоконтрастных изображениях;

- **стратегией оптимизации.** Выбор функции потерь является решающим для несбалансированных медицинских данных [9]. Примененная гибридная функция *Dice + Focal Loss* решает эту проблему на двух уровнях. *Dice Loss*, как метрика на уровне регионов, борется с дисбалансом, оптимизируя структурное сходство. *Focal Loss* же работает на уровне пикселей, концентрируя внимание модели на сложных примерах – пикселях

на границах опухоли, где неопределенность максимальна;

- **комплексной стратегией аугментации данных.** Обобщающая способность моделей зависит от разнообразия обучающей выборки. Применение широкого спектра геометрических и фотометрических преобразований позволило искусственно его создать. Этот подход, стандартный для медицинской визуализации [10], дал возможность обучить модель, эффективно работающую независимо от клинических вариаций: изменений положения пациента, артефактов и настроек МРТ-сканеров. Именно аугментация обеспечила робастность и генерализацию модели.

Однако у этого метода также есть ограничения, определяющие дальнейшее его применение, и главное из них – использование 2D-архитектуры. Обработывая каждый срез МРТ как независимое изображение, модель не анализирует трехмерный контекст и пространственную связь между срезами, что ведет к потере информации и может вызывать артефакты сегментации, например разрывы маски или неточности на полюсах образования.

Еще одна проблема – в генерализации модели. Обучение на приватном, одноцентровом наборе данных не гарантирует робастности на данных из других клиник с иными протоколами и оборудованием. Для подтверждения клинической применимости необходима валидация модели на внешних, мультицентрических наборах данных.

Вызовом остается и проблема «черного ящика». Разработанная модель, как и большинство сетей глубокого обучения, не объясняет свои решения. Такое отсутствие

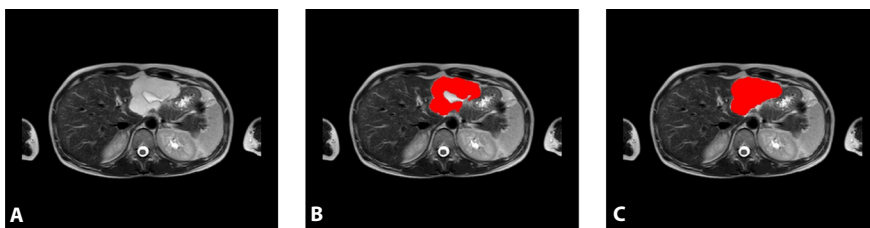


Рисунок. Примеры сегментации гемангиом моделью Attention U-Net. А – исходный T2-взвешенный МРТ-срез; В – эталонная сегментация, выполненная экспертом; С – сегментация, полученная с помощью модели Attention U-Net. Видно высокое совпадение контуров, включая сложные участки с неоднородной текстурой опухоли

Исследование	Метод/архитектура	DSC (%)	Характеристики датасета/задачи
Wu et al. x(x)(2023) [7]	Coarse-to-Fine Fusion (3D U-Net)	86,9	Сегментация малых опухолей, частный датасет
Hänsch et al. (2022) [8]	Анизотропная 3D aU-Net (мультимодель)	74,0	Сегментация опухолей на DCE-MRI, частный датасет
Наше исследование	2D Attention U-Net (SCSE)	84,7	Сегментация гемангиом, частный датасет со значительным дисбалансом

Таблица 3. Сравнение результатов данного исследования с опубликованными данными по сегментации опухолей печени на МРТ

интерпретируемости — серьезный барьер для доверия клиницистов, поскольку не позволяет понять причины ошибок и может скрывать систематические смещения алгоритма.

Для апробации разработанного метода в условиях, максимально приближенных к реальной клинической практике, был создан прототип программного обеспечения, реализующий следующий функционал:

- **графический пользовательский интерфейс (GUI)**, позволяющий врачу-рентгенологу загружать DICOM-серии МРТ-исследований;
 - **автоматическая сегментация**: после выбора исследования модель обрабатывает каждый срез и накладывает полученную маску на исходное изображение для визуального контроля;
 - **экспорт результатов**: пользователь может экспортировать сгенерированные маски в стандартные форматы (например, NIfTI или DICOM SEG) для дальнейшего использования в системах планирования лучевой терапии или для количественного анализа;
 - **интеграция с PACS**: прототип поддерживает базовую интеграцию с PACS-сервером Orthanc через протокол DICOMweb, что позволяет запрашивать и получать исследования напрямую из архива, эмулируя работу в едином медицинском информационном пространстве.
- Разработка данного прототипа – важный шаг, демонстрирующий, что созданный алгоритм представляет собой инструмент с реальным потенциалом для интеграции в радиологические рабочие процессы.

■ **Summary.** This paper presents a method for the automatic segmentation of liver hemangioma in magnetic resonance imaging (MRI) developed by the authors, based on modifications of the U-Net convolutional architecture. An experimental comparison was conducted between the baseline U-Net, U-Net++, and Attention U-Net. It was confirmed that the Attention U-Net model, utilizing a Spatial and Channel Squeeze-and-Excitation (scSE) attention mechanism and a combined loss function (Dice + Focal Loss), achieves the best performance, demonstrating a mean Dice Similarity Coefficient (DSC) of 84.65% on a private dataset. This result is comparable to state-of-the-art benchmarks, particularly when dealing with data characterized by significant class imbalance. To validate the method in conditions approximating clinical practice, a software prototype was developed featuring a graphical user interface and the capability for integration into PACS systems via the DICOMweb protocol.

■ **Keywords:** medical image segmentation, liver hemangioma, magnetic resonance imaging, convolutional neural networks, deep learning, MRI, U-Net, Attention U-Net, DICOM, PACS.

■ <https://doi.org/10.29235/1818-9857-2026-04-78-83>

Настоящее исследование продемонстрировало высокую эффективность модифицированной Attention U-Net в задаче автоматической сегментации гемангиом печени на T2-МРТ. Сочетание пространственно-канального механизма внимания и комбинированной функции потерь (Dice + Focal Loss) позволило достичь конкурентоспособных результатов (DSC 84,65%), значительно превзойдя базовые U-Net на сложном, несбалансированном наборе данных.

Апробированный прототип ПО подтвердил возможность интеграции метода в клиническую практику для автоматизации количественного анализа и снижения нагрузки на врачей-рентгенологов.

Несмотря на ограничения 2D-подхода, работа закладывает прочную основу для будущих исследований. Перспективы включают переход к 3D-архитектурам для учета объемной информации, внедрение мультимодального анализа для повышения диагностической точности и использование федеративного обучения для создания более робастных и обобщающих моделей. ■

Статья поступила в редакцию
11.08.2025 г.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Kacała A. et al, Evaluation of Predictive Factors for Transarterial Bleomycin–Lipiodol Embolization Success in Treating Giant Hepatic Hemangiomas // *Cancers*. 2025. №17(1). P. 42. <https://doi.org/10.3390/cancers17010042>.
2. A survey on deep learning in medical image analysis / G. Litjens [et al.] // *Medical Image Analysis*. 2017. V. 42. P. 60–88.
3. Albiin N. MRI of Focal Liver Lesions // *Curr Med Imaging Rev*. 2012. May, 8(2). P. 107–116. Doi: 10.2174/157340512800672216. PMID: 23049491; PMCID: PMC3462338.
4. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation // *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. 2015. P. 234–241.
5. Attention U-Net: Learning Where to Look for the Pancreas / O. Oktay [et al.] // <https://doi.org/10.48550/arXiv.1804.03999>.
6. Deep Residual Learning for Image Recognition / K. He [et al.] // *CVPR*. 2016. P. 770–778. Doi: 10.1109/CVPR.2016.90.
7. Coarse-to-Fine Liver Tumor Segmentation using 3D U-Net / H. Wu [et al.] // *Medical Image Analysis*. 2023. V. 85. P. 102734.
8. Multi-modal learning with an anisotropic 3D U-Net for the segmentation of liver tumors in DCE-MRI / A. Hänsch [et al.] // *Computerized Medical Imaging and Graphics*. 2022. V. 95. P. 102015.
9. Improving automatic liver tumor segmentation in late-phase MRI using multi-model training and 3D convolutional neural networks / Hänsch [et al.] // *Scientific Reports*. 2022. V.12. P. 12262. <https://doi.org/10.1038/s41598-022-16388-9>.
10. Explainable and Robust Deep Learning for Liver Segmentation Through U-Net Network / M. C. Brunese // *Diagnostics (Basel)*. 2025. Mar 31. V. 15(7). P. 878. Doi: 10.3390/diagnostics15070878. PMID: 40218228; PMCID: PMC11989174.