



Метод операторно-графовой релаксации для повышения качества автоматической сегментации анатомических структур сердца на медицинских изображениях

П.А. Пылов, ORCID: 0000-0003-3214-6592 <gedrosten@mail.ru>

Р.В. Майтак, ORCID: 0009-0009-4353-6925 <maytak.roman@mail.ru>

А.Г. Пимонов, ORCID: 0000-0002-9615-1895 <pag_vt@kuzstu.ru>

Кузбасский государственный технический университет имени Т.Ф. Горбачева,
Россия, 650000, г. Кемерово, ул. Весенняя, д. 28.

Аннотация. Предложен метод операторно-графовой релаксации для повышения качества автоматической сегментации желудочков сердца на двумерных срезах магнитно-резонансной томографии в трех ортогональных проекциях. В отличие от пиксельно-ориентированных подходов, осуществляющих независимую классификацию элементов изображения, предлагаемый метод рассматривает выход базовой модели U-Net как непрерывное вероятностное поле, подлежащее структурному согласованию на графе изображения. Томографический срез представляется в виде неориентированного графа, вершины которого соответствуют пикселям среза, ребра задаются отношением локального соседства, а веса ребер определяются пространственной близостью, перепадом интенсивности, выраженностью локального градиента и согласованностью вероятностных оценок соседних пикселей. Ключевым элементом метода является оператор переноса состояния, который в явном виде задается как линейное преобразование вектора вероятностей соседней вершины, зависящее от направления передачи информации, локального градиента интенсивности и матрицы анатомической совместимости классов. Параметры оператора задаются аналитически и не обучаются, что позволяет применять метод к вероятностному полю произвольной сегментационной модели без ее переобучения и без дополнительной разметки данных. Для итерационной процедуры релаксации доказано условие сжимаемости, обеспечивающее существование, единственность и геометрическую скорость сходимости к неподвижной точке, и сформулирован количественный критерий останова. Экспериментальная оценка выполнена на клиническом наборе данных с пятикратной кросс-валидацией на уровне пациентов и стратификацией по проекциям. Улучшение средних значений получено по всем десяти рассмотренным показателям – коэффициенту Дайса, HD95, Contour F1, Overreach и Underreach для обоих желудочков. На внешней выборке преимущество метода над базовой моделью сохранилось по девяти показателям из десяти при снижении абсолютных значений качества на 3,0–3,9 %.

Ключевые слова: сегментация медицинских изображений; магнитно-резонансная томография сердца; операторно-графовая релаксация; вероятностное поле сегментации; структурное согласование сегментации.

Для цитирования: Пылов П.А., Майтак Р.В., Пимонов А.Г. Метод операторно-графовой релаксации для повышения качества автоматической сегментации анатомических структур сердца на медицинских изображениях. Труды ИСП РАН, 2026, том 38, вып. 6, часть 1, стр. 233–352. DOI: 10.15514/ISPRAS-2026-38(6)-21.

Operator-graph relaxation for enhanced automatic segmentation of cardiac anatomical structures in medical images

P.A. Pylov, ORCID: 0000-0003-3214-6592 <gedrosten@mail.ru>

R.V. Maitak, ORCID: 0009-0009-4353-6925 <maytak.roman@mail.ru>

A.G. Pimonov, ORCID: 0000-0002-9615-1895 <pag_vt@kuzstu.ru>

T.F. Gorbachev Kuzbass State Technical University,
28, Vesennaya st., Kemerovo, 650000, Russia.

Abstract. This article presents a graph-based relaxation method that improves the quality of automatic segmentation of heart ventricles in two-dimensional magnetic resonance imaging (MRI) slices from three orthogonal projections. Unlike pixel-oriented approaches, which perform independent classification of image elements, the proposed method treats the output of the base U-Net model as a continuous probability field that is structurally aligned with the image graph. An MRI slice is represented as an undirected graph whose vertices correspond to the slice's pixels. The edges are defined by local neighborhood relations and the edge weights are determined by spatial proximity, intensity difference, magnitude of the local gradient, and consistency of probability estimates of neighboring pixels. A key element of the method is the state transfer operator, which is defined as a linear transformation of a neighboring vertex's probability vector depending on the direction of information transfer, local intensity gradient, and anatomical class compatibility matrix. The operator parameters are analytically specified and not learned. This allows the method to be applied to the probability field of an arbitrary segmentation model without retraining or additional data labeling. For the iterative relaxation procedure, a compressibility condition has been proven to ensure the existence and uniqueness of a fixed point and a geometric convergence rate, as well as a quantitative stopping criterion. An experimental evaluation was performed on a clinical dataset using five-fold cross-validation at the patient level and stratification by projections. The proposed method was compared with the baseline U-Net model and the nnU-Net and U-Net+ViT models. It was also compared with post-processing procedures, such as connected component selection, morphological closing, and a fully connected conditional random field, in terms of segmentation quality and computational cost. An ablation study evaluated the contribution of individual components of the method. External testing was additionally performed on an independent cohort of 100 studies acquired at a different medical institution located in another region, on a scanner from a different vendor; the algorithm version and all of its parameters were frozen before the reference annotations were accessed, and no fine-tuning was carried out. Improvements in mean values were observed for all ten considered metrics—Dice coefficient, HD95, Contour F1, Overreach, and Underreach for both ventricles. After adjusting for multiple comparisons, nine of the differences were statistically significant, and the reduction in right ventricular Underreach was interpreted as a trend. On the external cohort, the advantage of the method over the baseline model was retained for nine of the ten metrics, with absolute quality decreasing by 3.0–3.9%. Ventricular volumes and functional cardiac indices were not computed in this study; the reported results therefore characterize segmentation quality and are not direct evidence of improved accuracy of clinical measurements.

Keywords: medical image segmentation; magnetic resonance imaging of the heart; operator-graph relaxation; probabilistic segmentation field; structural consistency of segmentation.

For citation: Pylov P.A., Maitak R.V., Pimonov A.G. Operator-graph relaxation for enhanced automatic segmentation of cardiac anatomical structures in medical images. Trudy ISP RAN/Proc. ISP RAS, 2026, vol.38, issue 6, part 1, pp. 323–352 (in Russian). DOI: 10.15514/ISPRAS-2026-38(6)-21.

1. Введение

Заболевания сердечно-сосудистой системы стабильно занимают первое место среди причин смертности как в Российской Федерации, так и в мире. По данным Всемирной организации здравоохранения [1], заболевания данного типа ежегодно уносят до 19,8 миллионов жизней – это около трети от всех летальных исходов на планете. По данным Росстата, в 2022 году от болезней сердечно-сосудистой системы в России скончалось свыше 831 тысячи человек, что составило около 44 % от общего числа смертей в стране [2]. Необходимым условием снижения данных показателей является своевременная и точная диагностика структурных

патологий сердечной ткани, которая в значительной степени опирается на данные лучевой диагностики, прежде всего магнитно-резонансной томографии (МРТ) сердца.

МРТ сердца является золотым стандартом неинвазивной оценки анатомии и функции сердечных структур. Количественные параметры, извлекаемые из томографических данных, – объем полостей желудочков сердца, фракция выброса, масса миокарда, толщина стенок – играют ключевую роль в диагностике широкого спектра кардиологических нозологий: ишемической болезни сердца, кардиомиопатий, пороков сердца и хронической сердечной недостаточности [3]. Получение обозначенных показателей требует предварительной сегментации анатомических структур (левого и правого желудочков) на томографических изображениях. В клинической практике данная процедура традиционно выполняется вручную или полуавтоматически специалистами-кардиологами, что сопряжено со значительными временными затратами, зависимостью от квалификации оператора и низкой воспроизводимостью результатов между наблюдателями.

Точность автоматической сегментации важна не только для количественного анализа изображений, но и для построения пациент-специфических трехмерных реконструкций, применяемых при планировании вмешательств. Автоматизация способна снизить трудоемкость ручного контурирования и повысить воспроизводимость подготовки таких моделей. Вместе с тем в настоящей работе не оценивались ни точность трехмерной реконструкции, ни влияние сегментации на хирургическое планирование; предметом исследования является качество сегментации на двумерных срезах.

Целью данного исследования является разработка и сравнительная оценка метода операторно-графовой релаксации применительно к задаче автоматической сегментации анатомических структур сердца на томографических снимках. Вклад работы состоит в следующем:

- Предложен метод структурного согласования вероятностного поля сегментации, в котором перенос состояния между соседними вершинами графа изображения выполняется не операцией взвешенного усреднения, а линейным оператором, зависящим от направления передачи информации, локального градиента интенсивности и анатомической совместимости классов; тем самым оператор переноса является несимметричным относительно перестановки вершин ребра, что отличает предлагаемую схему от симметричных попарных потенциалов условных случайных полей.
- Для введенной итерационной процедуры получено достаточное условие сжимаемости отображения релаксации в равномерной норме, из которого следуют существование и единственность неподвижной точки, геометрическая скорость сходимости и явная оценка требуемого числа итераций; на этой основе сформулирован количественный критерий остановки, заменяющий качественное требование стационарности.
- Все параметры метода задаются аналитически и не подлежат обучению, вследствие чего метод применим к вероятностному полю произвольной сегментационной модели без ее переобучения, без изменения функции потерь и без дополнительной экспертной разметки.
- Выполнена внутренняя экспериментальная проверка на клиническом наборе данных с пятикратной кросс-валидацией на уровне пациентов, стратификацией выборок по проекциям, сопоставлением с моделями U-Net, nnU-Net и U-Net+ViT, а также с процедурами постобработки (отбор связанных компонент, морфологическое замыкание, полносвязное условное случайное поле) по качеству и вычислительным затратам и абляционным исследованием вклада отдельных компонентов метода.

- Выполнено внешнее тестирование на независимой выборке из 100 исследований, полученных в другом медицинском учреждении на томографе иного производителя, при фиксации версии алгоритма и всех его параметров до доступа к эталонной разметке; это позволяет оценить не только устойчивость эффекта внутри исходного набора, но и его сохранение при смене оборудования, протокола сканирования и практики экспертной разметки.

2. Обзор существующих подходов

2.1. Нейросетевые методы сегментации анатомических структур сердца

До широкого распространения моделей машинного обучения доминирующую роль в инструментарии кардиологов при выполнении сегментации занимали модельно-ориентированные подходы, в частности метод активных контуров (Active Contours, Snakes) [4]. Данный вариационный метод требовал значительного объема экспертных знаний для конструирования опорных признаков, на основе которых строится активный контур искомого объекта. Кроме того, построение контура по выделенным специалистом опорным признакам не отождествлялось с полноценным решением задачи сегментации структур сердца, поскольку патологические изменения анатомии сердечной ткани, которые часто и являлись причиной операционного вмешательства, существенно ограничивали применимость метода активных контуров.

Практическое применение моделей машинного обучения для анализа графической информации существенно повысило качество сегментации объектов в сравнении с классическими методами, в том числе и в кардиологии. Базовой архитектурой для сегментации медицинских изображений стала сверточная сеть U-Net с симметричной структурой «энкодер–декодер» и прямыми межуровневыми связями [5]. Ключевой отправной точкой развития машинного обучения для анализа кардиологических данных стал международный конкурс ACDC (Automated Cardiac Diagnosis Challenge), проведенный в рамках конференции MICCAI-2017 [6]. Набор данных, представленный для этого конкурса, включал МРТ-исследования 150 пациентов с экспертной разметкой левого и правого желудочков, а также миокарда. Сравнительные результаты участников продемонстрировали, что наиболее высокие значения коэффициента Дайса достигаются глубокими сверточными нейронными сетями, а обводка контура, максимально совпадающая с экспертной, была получена моделями семейства U-Net.

Дальнейшее развитие направления связано с автоматизацией выбора конфигурации модели. Комплекс nnU-Net [7] реализует самонастраивающийся конвейер, в котором предобработка, топология сети, стратегия обучения и постобработка выводятся из свойств набора данных; на широком множестве задач медицинской сегментации, включая ACDC, данный подход показал результаты на уровне лучших специализированных решений, что делает его естественным кандидатом на роль базовой модели в сравнительных исследованиях.

Для исследования кардиологических данных применяются также модели на основе архитектуры трансформера [8] и их комбинации со сверточными сетями. TransUNet [9] использует трансформерный энкодер для моделирования дальних пространственных зависимостей при сохранении декодирующей ветви, восстанавливающей детали границ. В более поздних работах гибридные CNN–ViT-схемы применяются не только как единая сеть, но и как взаимно согласуемые модели: в [10] объединены U-Net и Swin-UNet, а повышение качества достигается за счет согласованности на уровне данных и сети при ограниченном объеме разметки. Эти результаты подтверждают полезность глобального контекста, однако соответствующие решения требуют обучения существенно более ресурсоемкой архитектуры и не являются аналитической постобработкой произвольного вероятностного поля.

Отдельное направление образуют графовые нейронные сети, предсказывающие не пиксельную маску, а геометрически организованный контур. В работе [11], опирающейся на аппарат графовых сверточных сетей [12], графовая сеть использована для сегментации структур сердца на эхокардиографических данных: анатомически заданные кольца графа позволили уменьшить число грубо некорректных многокомпонентных результатов и использовать согласие с U-Net как индикатор качества входа и сегментации. В отличие от предлагаемого метода, такой подход содержит обучаемые веса, работает с заранее заданной топологией контура и требует отдельного обучения графовой модели.

2.2. Методы структурного согласования разметки

Ограничение пиксельно-независимой классификации состоит в том, что топологические ограничения структуры сердца не могут быть заданы внутри сверточной модели, вследствие чего возникают ложные изолированные области и разрывы контура анатомически цельного объекта [13]. Исторически данная проблема решается введением отдельного уровня согласования разметки, и предлагаемый в настоящей работе метод принадлежит именно к этому классу подходов, поэтому его контекст требует отдельного рассмотрения.

Первым классом таких методов являются процедуры релаксационной разметки, восходящие к работе [14], в которой согласование меток выполняется итерационным пересчетом вероятностей на основе коэффициентов совместимости соседних элементов сцены. Теоретические основания релаксационной разметки, включая условия сходимости и связь с задачами вариационного неравенства, были даны в [15]. Второй класс образуют марковские случайные поля и байесовские постановки восстановления изображений [16], в которых разметка определяется минимизацией энергии, объединяющей унарные и попарные члены. Практические алгоритмы минимизации таких энергий на графе изображения основаны на разрезах графа [17-18] и обеспечивают глобальный или приближенно-глобальный минимум для широкого класса потенциалов.

Третий класс составляют условные случайные поля, применяемые в качестве уровня уточнения выхода нейросетевой модели. Полносвязное условное случайное поле с гауссовыми потенциалами ребер [19] допускает эффективный приближенный вывод методом среднего поля и стало стандартным средством уточнения границ. В [20] показано, что итерации среднего поля могут быть представлены рекуррентной сетью и обучаться совместно с базовой моделью, а в [21] условное случайное поле использовано как этап уточнения выхода сети DeepLab. Для медицинских изображений эффективность такого уточнения подтверждена, в частности, при сегментации очаговых поражений головного мозга [22]. Отдельно отметим методы, встраивающие анатомические ограничения непосредственно в обучение: сети с анатомическими ограничениями на основе автокодировщика формы [23] и функции потерь, контролирующей топологию результата [24-25]. К этому же кругу относятся анизотропная диффузия [26-27], в которой сглаживание подавляется поперек границ, и простые процедуры постобработки – морфологическая фильтрация и отбор связных компонент, применяемые, в том числе, в самонастраивающихся конвейерах [7].

2.3. Позиционирование предлагаемого метода

Сопоставление перечисленных подходов с предлагаемым методом позволяет выделить его отличительные признаки. В условных случайных полях с гауссовыми потенциалами [19] взаимодействие соседних элементов описывается скалярным весом ребра, умноженным на функцию совместимости меток, симметричную относительно перестановки вершин; направление передачи информации в такой конструкции не различается. В предлагаемом методе перенос состояния задается матричным оператором, который зависит от градиента интенсивности в вершине-приемнике и от направления соответствующего ребра, вследствие

чего перенос из вершины (k, l) в вершину (i, j) и обратный перенос описываются различными операторами. Это позволяет подавлять передачу вероятностной массы через границу анатомической структуры избирательно, не блокируя согласование вдоль границы.

Прочие отличия касаются объекта и режима работы метода. Разрезы графа [17-18] оперируют бинарной разметкой; предлагаемый метод согласует непрерывное вероятностное поле и сохраняет информацию о неопределенности вблизи границ вплоть до финальной бинаризации. Графовые нейронные сети [11-12] требуют обучения, вмешательства в процедуру оптимизации базовой модели и дополнительной разметки – здесь же все параметры фиксированы аналитически. Анизотропная диффузия [26- 27] сглаживает поле интенсивностей, тогда как предлагаемая процедура действует на поле вероятностей и дополнительно учитывает анатомическую совместимость классов через матрицу взаимного подавления меток левого и правого желудочков. Морфологическая постобработка и отбор связных компонент удаляют области по геометрическому признаку; предлагаемый метод перераспределяет вероятностную массу согласно локальной структуре изображения, что позволяет не только подавлять изолированные ложноположительные включения, но и восстанавливать пропущенные фрагменты целевой структуры.

Таким образом, в настоящей работе предлагается метод операторно-графовой релаксации, направленный на уменьшение локальной несогласованности границ и нарушений связности в результатах нейросетевой сегментации. В предлагаемой схеме вероятностные карты рассматриваются не как окончательный результат, а как входное поле для структурного согласования на графе изображения.

3. Материалы и методы

3.1 Клинические данные и их предварительная обработка

Набор данных, используемый в настоящей работе, был предоставлен Федеральным государственным бюджетным научным учреждением «Научно-исследовательский институт комплексных проблем сердечно-сосудистых заболеваний» (НИИ КПССЗ, г. Кемерово). Данные являются обезличенными, получены в рамках стандартного клинического протокола обследования и представляют собой результаты МРТ-исследований сердца пациентов с различными кардиологическими нозологиями. Исследования основной выборки были выполнены на томографе Siemens MAGNETOM Sola с напряженностью магнитного поля 1,5 Т.

Независимо от основной выборки для внешнего тестирования сформирован второй набор данных, предоставленный многопрофильным медицинским центром ООО «Диалог Плюс» (г. Красноярск): 100 обезличенных МРТ-исследований сердца, выполненных на томографе Philips Ingenia с напряженностью магнитного поля 1,5 Т. Учреждения-источники расположены в разных субъектах Российской Федерации; пересечение по пациентам между выборками отсутствует. Внешняя выборка не использовалась ни при обучении базовых моделей, ни при подборе параметров ОГР; протокол ее обработки описан в разделе 3.9.

Экспертная разметка внешнего набора выполнена одним врачом-рентгенологом указанного центра – доктором медицинских наук со стажем работы по специальности не менее девяти лет. Использовалось то же правило присвоения меток, что и для основной выборки (0 – фон, 1 – полость правого желудочка, 2 – полость левого желудочка). Разметка выполнялась независимо от настоящего исследования и до применения к внешней выборке каких-либо моделей, то есть без доступа к их предсказаниям. Согласование разметки между учреждениями, совместный разбор расхождений и выработка единого консенсусного стандарта намеренно не проводились: целью внешнего тестирования являлась оценка работоспособности метода в условиях реального сдвига домена, одной из составляющих которого выступает различие локальных практик контурирования. Поскольку весь внешний

набор размечен одним специалистом, межнаблюдательная вариабельность эталона на этих данных не оценивалась; соответствующее ограничение рассматривается в разделе 5.1.

В обе выборки включались исследования, удовлетворяющие следующим условиям: наличие серии, полностью покрывающей полости обоих желудочков; техническое качество изображений, достаточное для контурирования границ эндокарда; наличие полной экспертной разметки обоих желудочков. Исключались исследования с выраженными артефактами движения или металлическими артефактами, препятствующими надежному выделению границ, неполные и прерванные серии, а также исследования с частичным покрытием желудочков.

Внешняя выборка сформирована ретроспективно из последовательных исследований, удовлетворявших перечисленным критериям и выполненным в период с февраля 2019 г. по март 2026 г.; дополнительный отбор по клиническим признакам не производился. Медиана возраста пациентов внешней выборки составила 41 год (диапазон от 22 до 67 лет), доля женщин – 48 %. Оба набора представлены в обезличенном виде; для внутренней выборки клинические метаданные, помимо факта наличия кардиологической патологии, не передавались, а нозологический состав обеих выборок не каталогизировался, поэтому формальное сопоставление клинического состава выборок не проводилось. Связанные с этим ограничения рассматриваются в разделе 5.1.

Оба набора данных обезличены и получены в рамках экспериментальной проверки метода. В соответствии с условиями их предоставления учреждениями-источниками наборы не подлежат публичному распространению и могут быть переданы только по обоснованному запросу при согласии правообладателей.

Исходный набор данных был получен в формате DICOM (*.dcm). Конвертирование серий в NIfTI выполнялось утилитой dcm2nii, запускаемой из воспроизводимого Python-конвейера; при преобразовании сохранялись пространственная ориентация, порядок срезов и метаданные трехмерного массива. Согласно экспертной разметке, каждый воксель имел одну из трех меток: 0 – фон, 1 – полость правого желудочка, 2 – полость левого желудочка. Итоговый набор включал 200 обезличенных томографических объемов, соответствующих 200 пациентам.

Хотя магнитно-резонансная томография сердца формирует трехмерный объем, в настоящей работе сознательно использована двумерная постановка, и ее выбор обусловлен четырьмя обстоятельствами. Во-первых, интерпретация томографических данных кардиологами и принятие диагностических решений традиционно выполняются по двумерным проекциям, а экспертное контурирование полостей левого и правого желудочков организовано посрезово, вследствие чего посрезовую структуру имеют и эталонная разметка, и исходный вычислительный конвейер. Во-вторых, предметом исследования является не архитектура сегментационной сети, а операторно-графовая релаксация как модуль структурного согласования; сопоставление методов на идентичных двумерных входах позволяет изолировать вклад этого модуля и не смешивать его с эффектом перехода к объемной архитектуре. В-третьих, модель операторно-графовой релаксации (ОГР), которая итерационно уточняет вероятности классов с учетом соседних пикселей, применяется к вероятностной карте отдельного среза и потому способна дополнять любую модель, выдающую такое поле, без ее переобучения; двумерная постановка сохраняет это свойство модульности. В-четвертых, посрезовая обработка требует существенно меньшего объема оперативной и видеопамати, что существенно при аппаратной конфигурации, указанной в табл. 1. Ограничения, вытекающие из отказа от межсрезового контекста, и требования к переходу в трехмерную постановку рассматриваются в разделе 5.2.

Каждый томографический объем был преобразован в набор двумерных срезов в трех вычислительных ортогональных проекциях (рис. 1), при этом для каждого среза формировались нормализованное полутоновое изображение, бинарная маска левого

желудочка, бинарная маска правого желудочка и RGB-визуализация с наложением масок на полутоновый снимок (левый желудочек выделен красным цветом, правый – синим). Полученные реконструкции не следует отождествлять со стандартными клиническими плоскостями магнитно-резонансной томографии сердца.

Извлечение среза с индексом i по оси axis из трехмерного объема D выполнялось операцией срезового индексирования: $D[i, :, :]$ при axis = 0 (сагиттальная проекция), $D[:, i, :]$ при axis = 1 (коронарная проекция), $D[:, :, i]$ при axis = 2 (аксиальная проекция). Нормализация интенсивностей каждого среза выполнялась робастным перцентильным методом (1), при помощи которого значения интенсивностей приводились к диапазону [0, 255] на основе 1-го и 99-го перцентилей распределения, что позволило устранить влияние выбросов (артефактов МРТ и слишком интенсивных включений) без потери клинически значимого динамического диапазона:

$$I_{norm} = clip\left(\frac{I - P_1}{P_{99} - P_1} \cdot 255, 0, 255\right). \tag{1}$$

В нормализации (1) P_{99} и P_1 – 99-й и 1-й перцентили интенсивностей среза соответственно; в случаях вырожденного диапазона ($P_{99} - P_1 < 10^{-6}$) срез заменялся нулевым массивом.

Табл. 1. Конфигурация базовой модели U-Net, режим ее обучения и вычислительная платформа эксперимента.
Table 1. Configuration of the baseline U-Net model, its training setup and the computing platform of the experiment.

Параметр	Значение
Размер входного изображения	256×256 пикселей
Число классов на выходе	2 (левый желудочек, правый желудочек); фон задается условием из раздела 3.4
Число обучаемых параметров	31 036 546
Весовой коэффициент αL в (2)	0,5
Весовой коэффициент βL в (2)	0,5
Оптимизатор	AdamW; $\beta_1 = 0,9$, $\beta_2 = 0,999$, weight decay = 10^{-4}
Начальная скорость обучения	1×10^{-3}
Расписание скорости обучения	CosineAnnealingLR, $\eta_{min} = 1 \times 10^{-6}$, $T_{max} = 150$
Размер пакета (batch size)	16
Максимальное число эпох	150
Критерий раннего останова	средний Dice на валидации; patience = 20, min $\Delta = 10^{-4}$
Вероятность SpatialDropout2d	0,20
Аугментация обучающей выборки	поворот $\pm 15^\circ$ ($p = 0,5$); сдвиг $\pm 10\%$ и масштаб $0,9 - 1,1$ ($p = 0,5$); эластическая деформация ($p = 0,2$); γ -коррекция $0,8 - 1,2$ ($p = 0,3$); гауссов шум $\sigma \leq 0,02$ ($p = 0,2$)
Взвешенное сэмплирование	по проекции и наличию LV/RV; веса обратно пропорциональны размеру страты
Инициализация весов	Kaiming normal для сверточных слоев; $\gamma = 1$, $\beta = 0$ для BatchNorm
Фиксация генераторов	seed = 42; deterministic algorithms = True
Программная платформа	Python 3.11; PyTorch 2.5.1; CUDA 12.4; cuDNN 9.1
Аппаратная конфигурация	CPU: Intel Core i7-11370H, 8 ядер; GPU: NVIDIA GeForce RTX 3050 Ti, 4 Гб VRAM; ОЗУ: 32 Гб.



Рис. 1. Пример разметки анатомических структур сердца на аксиальном срезе снимка магнитно-резонансной томографии: (а) исходный срез, (б) маска левого желудочка, (в) маска правого желудочка, (г) итоговая RGB-визуализация с нанесением масок на исходный срез.

Fig. 1. An example of the segmentation of anatomical structures of the heart in an axial MRI slice: (a) original slice, (b) left ventricular mask, (c) right ventricular mask, (d) final RGB visualization with the masks applied to the original slice.

Необходимо отметить, что в отличие от глобальной минимаксной нормализации, перцентильный подход позволяет обеспечить стабильный контраст мягких тканей при наличии радиочастотных артефактов и флуктуаций фонового сигнала, свойственных МРТ-данным.

После предварительной обработки 200 томографических объемов сформировано 26 260 двумерных срезов: 1 623 аксиальных, 12 140 коронарных и 12 497 сагиттальных. Соотношение 1 : 7,5 : 7,7 учитывалось при разбиении и сэмплировании. Внешняя выборка обрабатывалась тем же конвейером с теми же параметрами: из 100 объемов сформировано 12 940 срезов – 812 аксиальных, 6 018 коронарных и 6 110 сагиттальных. Интенсивностные изображения приводились к размеру 256×256 пикселей бикубической интерполяцией, а экспертные маски – исключительно методом ближайшего соседа, исключая появление промежуточных значений меток. После перцентильной нормализации значения интенсивности дополнительно масштабировались в диапазон [0; 1]; именно это представление используется в формулах (6) и (9).

3.2 Схема эксперимента, стратификация и кросс-валидация

Дисбаланс срезов учитывался на трех уровнях. Во-первых, все разбиения формировались на уровне пациентов, поэтому срезы одного исследования не попадали одновременно в обучающую и тестовую части. Во-вторых, обучающие пакеты формировались двухэтапным взвешенным сэмплером: сначала выбиралась проекция с вероятностью, обратно пропорциональной числу ее срезов, затем внутри проекции – страта по наличию целевой структуры (LV, RV или обеих), что предотвращало доминирование пустых и наиболее многочисленных классов срезов. В-третьих, качество оценивалось отдельно по проекциям; показатель «Все» рассчитывался по объединению предсказаний для всех тестовых пациентов, а не как невзвешенное среднее трех проекций.

Оценка выполнялась методом пятикратной кросс-валидации на уровне пациентов. Набор из 200 исследований разделялся на пять непересекающихся блоков (*fold*) по 40 пациентов. В каждой из пяти итераций 160 пациентов образовывали обучающую часть; из нее фиксированным стратифицированным отбором выделялись 24 пациента (15 %) для валидации и раннего останова, а 136 использовались для обновления параметров. Оставшиеся 40 пациентов служили тестовой частью. Все методы использовали одни и те же пять разбиений. В таблицах, включенных в следующие далее разделы, приводятся среднее и стандартное отклонение по итерациям; для статистического анализа каждому пациенту сопоставлялось одно значение каждой метрики, вычисленное в той итерации, где этот пациент входил в тестовую часть, что исключает псевдорепликацию коррелированных срезов.

3.3 Базовая модель U-Net и конфигурация обучения

В качестве базовой модели для сегментации желудочков сердца на проекциях трех анатомических плоскостей была выбрана полносверточная нейронная сеть U-Net с симметричной структурой «энкодер–декодер» и прямыми межуровневыми связями («skip connections»), соединяющими соответствующие уровни обеих ветвей сети [5]. Архитектура модели была адаптирована для решения задачи обработки двумерных срезов томографии сердца:

- основным ее элементом является модуль, включающий два последовательных блока, каждый из которых осуществляет четыре операции: свертку 3×3 без смещения, пакетную нормализацию, нелинейную активацию ReLU и опциональный пространственный Dropout; отказ от использования смещения в сверточных слоях связан с применением пакетной нормализации, позволяющей учитывать и компенсировать сдвиг распределения активаций;
- кодирующая ветвь модели (энкодер) состоит из четырех последовательных блоков EncoderBlock, каждый из которых применяет модуль DoubleConv к входному тензору, сохраняя результат в качестве карты признаков для реализации последующей прямой межуровневой связи;
- декодирующая часть состоит из четырех блоков DecoderBlock, в рамках которых осуществляется восстановление пространственного разрешения; для этого используется транспонированная свертка ConvTranspose2d с фиксированными параметрами (kernel = 2, stride = 2), после чего выполняется конкатенация с соответствующими картами признаков энкодера по каналному измерению;
- прямые межуровневые связи обеспечивают передачу критически значимой для точного формирования границ левого и правого желудочков высокочастотной пространственной информации, которая включает локальные градиенты и контуры анатомических объектов.

Обучение базовой модели осуществляется с использованием комбинированной функции потерь BCDiceLoss, представляющей собой линейную комбинацию бинарной кросс-энтропии по логитам и функции потерь Дайса:

$$L_{\text{total}} = \alpha_L L_{\text{BCE}} + \beta_L L_{\text{Dice}} \quad (2)$$

В комбинированной функции потерь (2) α_L и β_L являются весовыми коэффициентами; во избежание смещения обозначений, параметр релаксации, введенный ниже в (11), обозначен той же буквой α без индекса. Использование (2) обеспечивает численную устойчивость за счет объединения операций вычисления сигмоиды и логарифмической функции, предотвращая переполнение градиентов. Компонента бинарной кросс-энтропии L_{BCE} обеспечивает устойчивый градиентный сигнал и высокую чувствительность к локальным ошибкам классификации. Вторым слагаемым в (2) является компонента потерь Дайса L_{Dice} [28], которая, в отличие от кросс-энтропии, инвариантна к дисбалансу классов и ориентирована на согласование формы сегментируемых областей. Использование L_{Dice} принципиально важно при сегментации правого желудочка, характеризующегося относительно малой площадью и высокой вариабельностью формы.

Для nnU-Net использовалась конфигурация 2d nnU-Net v2: параметры предобработки, топология сети и расписание обучения определялись конвейером автоматически по свойствам набора данных, без ручной настройки. Разбиение на блоки задавалось извне и совпадало с разбиением пациентов, использованным для остальных моделей. Гибридная модель U-Net+ViT [8-9] использовала энкодер ViT-B/16, предварительно инициализированный на ImageNet-21k, AdamW с начальной скоростью 1×10^{-4} и размером

пакета 8; остальные правила разбиения, раннего останова, бинаризации и расчета метрик оставались идентичными.

3.4 Метод операторно-графовой релаксации

Предлагаемый метод операторно-графовой релаксации направлен на согласование вероятностного поля, формируемого базовой нейросетевой моделью. Цель метода – устранить несогласованности вероятностных оценок, приводящие к нарушению топологической и геометрической целостности формируемых анатомических структур. Метод применяется на этапе вывода к выходу базовой модели, не изменяет процедуру ее обучения и не содержит обучаемых параметров; все величины, входящие в его определение, вычисляются аналитически по текущему срезу и его вероятностному полю. В этом состоит принципиальное отличие от графовых нейронных сетей и обучаемых модулей уточнения, требующих совместной оптимизации с базовой моделью.

Пусть результат работы базовой модели на срезе I представлен вероятностным тензором $P \in [0,1]^{H \times W \times 2}$

где число 2 характеризует количество сегментируемых классов: $c = 1$ – левый желудочек, $c = 2$ – правый желудочек. Каждому пикселю с координатами (i, j) сопоставляется вектор состояния

$$s_{ij} = (p_{ij}^{(1)}, p_{ij}^{(2)})^T \in [0,1]^2 \quad (4)$$

что позволяет интерпретировать сегментацию как множество состояний, сохраняющее информацию о степени неопределенности на границах искоемых объектов. Фон не выделяется в отдельную компоненту состояния и соответствует условию $p_{ij}^{(1)} < \frac{1}{2}$ и $p_{ij}^{(2)} < \frac{1}{2}$.

Двумерная проекция томографического снимка представляется в виде неориентированного графа $G = (V, E)$, где V – множество вершин, соответствующих пикселям среза, а E – множество ребер, задающих локальные связи между пикселями. Каждой вершине $v_{ij} \in V$ сопоставляется состояние s_{ij} , а ее окрестность определяется отношением восьмисвязного соседства:

$$N(i, j) = \{(k, l): \max(|i - k|, |j - l|) = 1\} \quad (5)$$

Веса ребер отражают степень локальной согласованности соседних пикселей и определяются четырьмя величинами: евклидовым расстоянием между пикселями $d_{(ij),(kl)}$, перепадом интенсивности $\Delta I_{(ij),(kl)} = |I_{ij} - I_{kl}|$, средней выраженностью локального градиента $\bar{g}_{(ij),(kl)} = \frac{(g_{ij} + g_{kl})}{2}$, где $g_{ij} = \|\nabla_h I(v_{ij})\|_2$ – модуль дискретного градиента, вычисляемый оператором Собеля, и рассогласованием вероятностных оценок $\Delta p_{(ij),(kl)} = \|s_{ij} - s_{kl}\|_2$. Ненормированный вес ребра задается выражением

$$\tilde{\omega}_{ij,kl} = \exp\left(-\frac{d_{ij,kl}^2}{2\sigma_d^2} - \frac{\Delta I_{ij,kl}^2}{2\sigma_I^2} - \frac{\bar{g}_{ij,kl}^2}{2\sigma_g^2} - \frac{\Delta p_{ij,kl}^2}{2\sigma_p^2}\right) \quad (6)$$

где $\sigma_d, \sigma_I, \sigma_g, \sigma_p$ – параметры чувствительности веса к расстоянию, перепаду интенсивности, выраженности градиента и рассогласованию вероятностных оценок соответственно. Большие значения $\tilde{\omega}$ соответствуют близким, интенсивностно однородным, не разделенным границей пикселям с согласованными вероятностными оценками. Веса нормируются по окрестности вершины-приемника:

$$\omega_{ij,kl} = \frac{\tilde{\omega}_{ij,kl}}{\sum_{(k',l') \in N(i,j)} \tilde{\omega}_{ij,k'l'} + \varepsilon}, \quad \sum_{(k,l) \in N(i,j)} \omega_{ij,kl} \leq 1 \quad (7)$$

где $\varepsilon > 0$ – малая регуляризирующая константа. Условие нормировки (7) существенно: оно используется далее при доказательстве сжимаемости отображения релаксации.

Ключевым элементом метода является оператор переноса состояния $\Lambda_{(ij) \leftarrow (kl)}$, задающий преобразование состояния соседней вершины перед его учетом при обновлении. В отличие от схем взвешенного усреднения, оператор позволяет учитывать структурные особенности изображения: направление передачи информации, локальные градиентные характеристики и анатомическую совместимость классов. Оператор задается как линейное преобразование пространства состояний:

$$\Lambda_{ij \leftarrow kl}(s_{kl}) = \gamma_{ij,kl} C s_{kl} \quad (8)$$

в котором скалярный множитель $\gamma_{(ij),(kl)}$ описывает направленное затухание переноса, а матрица C – анатомическую совместимость классов. Множитель направленного затухания определяется проекцией градиента интенсивности в вершине-приемнике на направление ребра:

$$\gamma_{ij,kl} = \exp\left[-\frac{\max(0, \langle \nabla_h I(v_{ij}), e_{kl \rightarrow ij} \rangle)^2}{2\sigma_b^2}\right], \quad e_{kl \rightarrow ij} = \frac{x_{ij} - x_{kl}}{\|x_{ij} - x_{kl}\|_2} \quad (9)$$

где x_{ij} – координаты пикселя, $e_{kl \rightarrow ij}$ – единичный вектор направления переноса, а σ_b – параметр чувствительности к барьеру. Множитель удовлетворяет условию $0 < \gamma_{ij,kl} \leq 1$, близок к единице при переносе вдоль однородной области или вдоль границы и убывает при переносе через границу, соответствующую существенному нарастанию интенсивности в направлении переноса. Поскольку выражение (9) зависит от градиента в вершине-приемнике и направления ребра, в общем случае $\gamma_{ij,kl} \neq \gamma_{kl,ij}$; следовательно, оператор переноса несимметричен относительно перестановки вершин ребра. Это ограничивает проникновение вероятностной массы внутрь смежной анатомической структуры, не препятствуя согласованию состояний вдоль контура.

Матрица анатомической совместимости классов задается выражением

$$C = \frac{1}{1 + \kappa} \begin{pmatrix} 1 & -\kappa \\ -\kappa & 1 \end{pmatrix}, \quad 0 \leq \kappa < 1 \quad (10)$$

где параметр κ определяет силу взаимного подавления меток левого и правого желудочков. Внедиагональные элементы отрицательны, поэтому высокая вероятность принадлежности соседнего пикселя правому желудочку уменьшает переносимую в текущую вершину вероятность принадлежности левому желудочку, и наоборот. Тем самым в модель вводится анатомическое требование раздельности желудочков, которое не может быть выражено в схеме простого усреднения. Нормирующий множитель $\frac{1}{(1+\kappa)}$ обеспечивает выполнение условия $\|C\|_\infty = 1$, необходимого для сходимости процедуры релаксации. При $\kappa = 0$ матрица C вырождается в единичную, а оператор (8) – в скалярное направленное затухание; этот случай рассматривается далее как одна из конфигураций абляционного исследования. Согласование состояний осуществляется аффинной итерационной процедурой. При начальном условии $s_{ij}^{(0)} = (p_{ij}^{(1)}, p_{ij}^{(2)})^T$ обновление имеет вид

$$\tilde{s}_{ij}^{(t+1)} = (1 - \alpha) s_{ij}^{(0)} + \alpha \sum_{(k,l) \in N(i,j)} \omega_{ij,kl} \Lambda_{ij \leftarrow kl}(s_{kl}^{(t)}) \quad (11)$$

с последующим покомпонентным проецированием на допустимое множество состояний

$$s_{ij}^{(t+1)} = \Pi_{[0,1]^2}(\tilde{s}_{ij}^{(t+1)}), \quad \Pi_{[0,1]}(z) = \min(1, \max(0, z)) \quad (12)$$

В (11) параметр $\alpha \in (0; 1)$ задает долю структурно согласованного сообщения, а слагаемое $(1 - \alpha)s_{ij}^{(0)}$ сохраняет привязку к исходному вероятностному полю нейросети. Такая аффинная привязка устраняет вырождение итерации к нулевому полю и обеспечивает ненулевую неподвижную точку. Проецирование (12) необходимо, поскольку отрицательные внедиагональные элементы C могут вывести правую часть (11) за пределы $[0; 1]^2$. Покомпонентное проецирование на отрезок не увеличивает расстояний между состояниями,

то есть является нерастягивающим отображением, и потому не нарушает свойств сходимости, установленных в разделе 3.5. Обновление выполняется синхронно и не зависит от порядка обхода пикселей.

Итерации продолжают до выполнения количественного критерия остановки

$$\Delta^{(t)} = \max_{(i,j) \in V} \|s_{ij}^{(t+1)} - s_{ij}^{(t)}\|_{\infty} \leq \delta_s \quad \forall \quad t+1 = T_{\max} \quad (13)$$

где δ_s – допустимое изменение состояния между соседними итерациями, а $T_{\max}$ – ограничение на число итераций, исключающее неограниченный вычислительный процесс. Первое условие соответствует достижению стационарного состояния с заданной точностью и заменяет качественное требование стационарности количественной проверкой. Итоговая разметка формируется бинаризацией согласованного поля:

$$L_{ij} = \begin{cases} \arg \max_{c \in \{1,2\}} s_{ij}^{(T,c)}, & \max_c s_{ij}^{(T,c)} \geq \theta, \\ 0, & \text{иначе.} \end{cases} \quad (14)$$

где θ – порог бинаризации, T – номер итерации, на которой выполнен критерий (13). Никакой дополнительной морфологической фильтрации или отбора связанных компонент после (14) не выполняется; сравниваемые модели бинаризуются тем же порогом θ , что обеспечивает корректность сопоставления. Значения всех параметров метода приведены в табл. 2.

Табл. 2. Параметры метода операторно-графовой релаксации.

Table 2. Parameters of the operator-graph relaxation method.

Обозначение	Смысл параметра	Значение
α	доля структурного согласования относительно исходного поля, (11)	0,35
σ_d	чувствительность веса к расстоянию, (6)	1,0 пикс.
σ_I	чувствительность к перепаду нормированной интенсивности, (6)	0,10
σ_g	чувствительность к модулю нормированного градиента, (6)	0,15
σ_p	чувствительность к рассогласованию вероятностей, (6)	0,20
σ_b	чувствительность направленного затухания, (9)	0,12
κ	взаимное подавление меток LV/RV, (10)	0,15
ε	регуляризирующая константа, (7)	1×10^{-12}
δ_s	порог остановки, (13)	1×10^{-4}
$T_{\max}$	максимальное число итераций, (13)	20
T	фактическое число итераций, среднее $\pm$ СКО	$8,6 \pm 1,7$
θ	порог бинаризации, (14)	0,50

3.5 Сходимость и вычислительная сложность метода

Введенная аффинная процедура допускает строгое обоснование сходимости. Обозначим через S полное состояние поля, через $S^{(0)}$ – фиксированное исходное поле U-Net, а через F – отображение, задаваемое правой частью (11) с последующим проецированием (12).

Утверждение. Пусть выполнены условие нормировки (7), $\|C\|_{\infty} = 1$, $0 < \gamma_{ij,kl} \leq \gamma_{\max} \leq 1$ и $\alpha \in (0; 1)$. Тогда F является сжимающим отображением в равномерной норме с константой

$$q = \alpha \gamma_{\max} \quad (15)$$

Поскольку $q = \alpha \gamma_{\max} \leq \alpha < 1$, по принципу сжимающих отображений существует единственная неподвижная точка S^* , к которой последовательность $S^{(t+1)} = F(S^{(t)})$ сходится из любого начального состояния с геометрической скоростью

$$\|S^{(t)} - S^*\|_{\infty} \leq q^t \|S^{(0)} - S^*\|_{\infty} \quad (16)$$

Доказательство. Для произвольных состояний S и S' фиксированное слагаемое $(1 - \alpha)S^{(0)}$ сокращается. Из линейности оператора (8), нерастягиваемости проектора и неравенства треугольника для любой вершины (i, j) следует

$$\|F_{ij}(S) - F_{ij}(S')\|_{\infty} \leq \alpha \sum_{(k,l) \in N(i,j)} \omega_{ij,kl} \gamma_{ij,kl} \|C\|_{\infty} \|s_{kl} - s'_{kl}\|_{\infty} \quad (17)$$

Оценивая $\|s_{kl} - s'_{kl}\|_{\infty} \leq \|S - S'\|_{\infty}$, используя $\|C\|_{\infty} = 1$, $\gamma_{ij,kl} \leq \gamma_{\max}$ и нормировку (7), получаем

$$\|F(S) - F(S')\|_{\infty} \leq \alpha \gamma_{\max} \|S - S'\|_{\infty} \quad (18)$$

что совпадает с (15). Условие $\alpha \in (0; 1)$ обеспечивает $q < 1$ даже при $\gamma_{\max} = 1$; следовательно, дополнительное искусственное ограничение $\gamma_{\max} < 1$ не требуется. Утверждение доказано.

Аффинная привязка к $S^{(0)}$ имеет также содержательный смысл: графовая процедура уточняет исходный прогноз, но не заменяет его автономной диффузией. Из (16) следует явная верхняя оценка числа итераций, необходимых для достижения точности δ_s :

$$T_{\text{req}} \leq \left\lceil \frac{\ln \left(\frac{\delta_s}{\|S^{(0)} - S^*\|_{\infty}} \right)}{\ln q} \right\rceil \quad (19)$$

При $\alpha = 0,35$ и $\gamma_{\max} \leq 1$ теоретическая константа q не превышает 0,35. На тестовых предсказаниях критерий (13) выполнялся в среднем за $8,6 \pm 1,7$ итерации; распределение числа итераций скошено вправо, и ограничение $T_{\max} = 20$ достигалось менее чем в 0,3 % срезов – преимущественно на срезах с протяженной низкоконтрастной границей правого желудочка.

Вычислительная сложность одной итерации составляет $O(|V| \cdot |N| \cdot |C|^2)$, где $|V| = H \cdot W$, $|N| = 8$ и $|C| = 2$. Веса (6) и множители (9) вычисляются однократно для исходного изображения и поля $S^{(0)}$. Полная сложность равна $O(T \cdot H \cdot W)$, память – $O(H \cdot W)$. При значении batch = 1, после 100 прогревочных запусков и 1000 измерений с cudaDeviceSynchronize, прямой проход U-Net занимал $4,8 \pm 0,2$ мс на срез, этап ОГР – $1,7 \pm 0,1$ мс, суммарно $6,5 \pm 0,2$ мс. Накладные расходы относительно U-Net составили 35,4 %, а абсолютная добавка – 1,7 мс.

3.6 Метрики качества

Сравнение результатов работы моделей производилось по пяти метрикам, охватывающим два ключевых аспекта качества сегментации – степень перекрытия искомой области и точность воспроизведения границ объекта:

- коэффициент Дайса (Dice similarity coefficient, DSC) – основная метрика в задачах медицинской сегментации, характеризующая степень перекрытия предсказанной и эталонной масок [29]; значения метрики лежат в диапазоне [0, 1], где 1 соответствует идеальному совпадению;
- расстояние Хаусдорфа (HD95) – 95-й процентиль двустороннего расстояния Хаусдорфа между множествами граничных пикселей предсказанной и эталонной масок; в отличие от коэффициента Дайса, данная метрика чувствительна к локальным отклонениям границы и характеризует верхнюю часть распределения ошибок контурирования, причем меньшее значение соответствует лучшему результату [30]; применение 95-го перцентиля вместо полной метрики позволяет подавить влияние единичных артефактных пикселей разметки [31]. Величина HD95 потенциально релевантна последующему объемному анализу, однако объемные и функциональные показатели в настоящей работе не рассчитывались, поэтому изменение HD95 интерпретируется исключительно как изменение качества

сегментации, а не как свидетельство повышения точности клинических измерений [32];

- Contour F1 – среднее гармоническое между точностью (precision) и полнотой (recall), вычисляемое на уровне граничных пикселей; значения лежат в диапазоне [0, 1], где 1 соответствует наилучшему результату; метрика позволяет оценить качество воспроизведения контура вне зависимости от заполнения области и выявить случаи, когда модель корректно определяет область структуры, но систематически смещает ее границу; выбор данной метрики особенно важен для оценки качества сегментации правого желудочка, имеющего вытянутую серповидную форму с протяженными стенками;
- Overreach и Underreach – метрики, характеризующие род ошибки сегментации; Overreach – доля пикселей предсказанной маски, выходящих за пределы размеченной экспертами области (ложноположительные пиксели), значения метрики лежат в диапазоне [0, +∞), при этом меньшие значения соответствуют более точному контурированию границ объекта; Underreach – доля пикселей истинной маски, не покрытых предсказанием модели (ложноотрицательные пиксели), значения лежат в диапазоне [0, 1], где 0 соответствует полному покрытию искомого объекта.

Метрики вычислялись раздельно для левого и правого желудочков в каждой ортогональной проекции. Все метрики вычислялись на уровне отдельного среза; срезы, не содержащие целевой структуры ни в эталонной, ни в предсказанной маске, исключались из расчета граничных метрик.

3.7 Статистический анализ и абляционное исследование

Статистическая значимость различий оценивалась двусторонним критерием знаковых рангов Уилкоксона [33] при $\alpha_{stat} = 0,05$. Единицей наблюдения был пациент: для каждого из 200 пациентов по предсказаниям на тестовой части той итерации, в которой он участвовал, сначала агрегировалась соответствующая метрика по информативным срезам, после чего сравнивались парные значения для пациентов U-Net и ОГР. Такой порядок исключает искусственное увеличение объема выборки за счет коррелированных срезов. Величина эффекта рассчитывалась как $r = \frac{|Z|}{\sqrt{n}}$, $n = 200$; для десяти одновременно проверяемых показателей применялась поправка Холма–Бонферрони.

Для оценки вклада отдельных компонентов метода выполнено абляционное исследование, в рамках которого сопоставлены четыре конфигурации:

- M1 – базовая модель U-Net без согласования вероятностного поля;
- M2 – согласование взвешенным усреднением по окрестности, то есть замена оператора переноса (8) тождественным преобразованием при сохранении весов (6)–(7); данная конфигурация позволяет отделить вклад графового усреднения от вклада операторной структуры;
- M3 – операторно-графовая релаксация без учета анатомической совместимости классов ($\kappa = 0$), при которой оператор (8) сводится к направленному затуханию; конфигурация позволяет оценить вклад матрицы взаимного подавления меток;
- M4 – полная конфигурация предлагаемого метода.

Чувствительность исследовалась по сетке $\alpha \in \{0,15; 0,25; 0,35; 0,45; 0,55\}$ и $T_{max} \in \{4; 8; 12; 16; 20\}$. В каждой внешней итерации выбор выполнялся исключительно по ее внутренней валидационной подвыборке из 24 пациентов; соответствующая тестовая подвыборка из 40 пациентов при настройке не использовалась. Критерием служил средний коэффициент Дайса, усредненный по LV и RV. Во всех пяти внешних итерациях максимум либо плато в пределах разброса между итерациями достигались при $\alpha = 0,35$ и $T_{max} = 20$; эти значения

фиксируются до расчета метрик на тестовой части каждой итерации. В разделе 4.3 приведены не результаты на тестовых частях итераций, а сведения по усредненной по пяти внутренним валидационным подвыборкам поверхность качества.

3.8 Сравнительные методы постобработки и измерение вычислительных затрат

Помимо архитектурных базовых моделей, к тем же вероятностным картам U-Net применялись три близких по назначению варианта постобработки: (1) сохранение крупнейшей связной компоненты отдельно для LV и RV; (2) бинарное замыкание диском радиуса 2 пикселя с последующим сохранением крупнейшей компоненты; (3) DenseCRF с пятью итерациями среднего поля, пространственным стандартным отклонением 3 пикселя, параметрами билатерального ядра $\sigma_{xy} = 5$ пикселей и $\sigma_I = 0,10$, совместимостью Поттса, равной 3. Все параметры выбирались только по валидационным подвыборкам. После каждого метода применялся единый порог $\theta = 0,5$; ручная коррекция не выполнялась.

Число обучаемых параметров подсчитывалось средствами PyTorch. Время вывода измерялось на одном срезе 256×256 при batch = 1 в режиме model.eval() и torch.no_grad(). В таблицах указано полное время получения маски, включая постобработку.

3.9 Протокол внешнего тестирования

Внешнее тестирование выполнялось по протоколу, соответствующему рекомендациям CLAIM 2024 в части независимой оценки алгоритмов медицинской визуализации [34]. До получения доступа к эталонной разметке внешней выборки фиксировались веса базовой модели U-Net, обученной на внутренней выборке (использована модель той итерации, которая дала медианное значение валидационного коэффициента Дайса), полный набор параметров ОГР из табл. 2, порог бинаризации $\theta = 0,5$ и процедура предобработки, включая параметры перцентильной нормализации (1) и способ приведения срезов к размеру 256×256. Дообучение, тонкая настройка и подбор каких-либо параметров на внешних данных не выполнялись; внешняя выборка обрабатывалась однократно (табл. 3).

Первичными показателями переносимости считались парные разности коэффициента Дайса для левого и правого желудочков между ОГР и базовой моделью, вычисленные на уровне пациента. Заранее заданным критерием сохранения эффекта являлись положительный знак каждой из этих разностей и статистическая значимость после поправки на множественность сравнений. Вторичными показателями служили абсолютные значения всех десяти метрик раздела 3.6 и их относительное изменение по сравнению с внутренней оценкой. Единицей наблюдения, как и во внутреннем эксперименте, был пациент: метрики агрегировались по информативным срезам исследования, после чего парные значения для моделей U-Net и ОГР сравнивались двусторонним критерием знаковых рангов Уилкоксона [33] с поправкой Холма–Бонферрони на десять сравнений.

Для каждой метрики дополнительно вычислялась средняя парная разность $\Delta = \text{ОГР} - \text{U-Net}$ и ее 95-процентный доверительный интервал по нормальному приближению $\Delta \pm 1,96 \cdot s / \sqrt{n}$, где s – выборочное стандартное отклонение парных разностей, а $n = 100$ – число пациентов внешней выборки; правомерность нормального приближения для среднего парных разностей при таком объеме выборки обеспечивается центральной предельной теоремой и не требует нормальности распределения самих метрик. Доверительные интервалы приводятся как описательная мера величины эффекта, тогда как выводы о статистической значимости во всех случаях основаны на критерии Уилкоксона, не использующем предположений о виде распределения. Границы интервалов и результаты критерия согласованы между собой: интервал содержит ноль только для того показателя, по которому критерий не выявил значимого различия.

Табл. 3. Сопоставление характеристик внутренней и внешней выборок.
Table 3. Comparison of the internal and external cohorts.

Характеристика	Внутренняя выборка	Внешняя выборка	Значение для интерпретации
Учреждение	НИИ КПССЗ, г. Кемерово	ООО «Диалог Плюс», г. Красноярск	Разные города и субъекты РФ; пересечение по пациентам отсутствует
Число исследований / срезов	200 / 26 260	100 / 12 940	Единица наблюдения в обеих выборках – пациент
Оборудование	Siemens MAGNETOM Sola, 1,5 T	Philips Ingenia, 1,5 T	Смена производителя и протокола сканирования
Экспертная разметка	Специалисты НИИ КПССЗ	Один врач-рентгенолог, д.м.н., стаж ≥ 9 лет	Независимые разметчики; согласование намеренно не проводилось
Использование данных	Обучение, валидация, внутреннее тестирование	Только тестирование, однократный прогон	Дообучение и подбор параметров на внешних данных не выполнялись
Схема оценки	Пятикратная кросс-валидация на уровне пациентов	Зафиксированная версия алгоритма, примененная ко всей выборке	Разброс между итерациями против разброса между пациентами
Отчетные показатели	Табл. 4–8 (см. далее)	Табл. 10–11 (см. далее)	Единый набор метрик раздела 3.6

Существенное различие двух схем оценки состоит в следующем: внутренние показатели представляют собой среднее по пяти итерациям кросс-валидации, тогда как внешние получены однократным применением зафиксированного алгоритма ко всей выборке из 100 пациентов. Поэтому во внутренних таблицах разброс характеризует вариабельность между итерациями, а во внешней – между пациентами; прямое сопоставление величин стандартного отклонения между этими таблицами некорректно, и сравниваются только средние значения. Сопоставление характеристик выборок приведено в табл. 3.

4. Результаты

4.1 Сравнение с базовыми моделями и вычислительные затраты

В табл. 4 сопоставлены три нейросетевые архитектуры, традиционные процедуры постобработки и предлагаемый метод. Для каждой обучаемой архитектуры дополнительно показано применение ОГР к ее вероятностному полю; число параметров при этом не изменяется. Такое сравнение отделяет вклад структурного согласования от эффекта увеличения емкости нейросети. Результаты табл. 4 показывают, что простая фильтрация связанных компонент и морфологическое замыкание дают ограниченный прирост, а DenseCRF улучшает границы ценой почти четырехкратного увеличения времени относительно U-Net. nnU-Net и U-Net+ViT повышают исходное качество, однако требуют отдельного обучения; гибридная модель содержит в 3,39 раза больше параметров, чем U-Net. ОГР обеспечивает наибольший прирост для каждой из трех архитектур без добавления обучаемых параметров; относительный выигрыш закономерно уменьшается по мере усиления базовой модели. Дальнейший детальный анализ выполнен для U-Net как наименее ресурсоемкой и наиболее прозрачной базовой архитектуры.

Средние значения в табл. 5 улучшаются по всем десяти показателям, однако величина эффекта различается. Наиболее выражено снижение граничной ошибки: на аксиальных срезах HD95 уменьшился с 5,099 до 1,360 пикселя для LV и с 9,110 до 3,401 пикселя для RV; Contour F1 правого желудочка вырос на 0,127. Прирост Dice составил 0,055 для LV и 0,071 для RV. Термин «аксиальная» здесь обозначает одну из вычислительно реконструированных ортогональных проекций используемого набора, а не стандартную клиническую короткоосевую плоскость CMR. После поправки Холма-Бонферрони различия статистически значимы для метрик Dice, HD95, Contour F1 и Overreach обоих желудочков (pHolm < 0,001; r = 0,530–0,620; табл. 6), а также для Underreach L (p = 0,0033; pHolm = 0,0066; r = 0,208). Для метрики Underreach R различие не достигло уровня значимости после поправки (p = 0,0672; pHolm = 0,0672; r = 0,129), поэтому его снижение интерпретируется как тенденция. В пяти итерациях аксиальное расстояние HD95 L для модели U-Net составляло 5,430; 4,780; 5,210; 5,100; 4,975, а для модели ОГР – 1,520; 1,280; 1,410; 1,190; 1,400; HD95 R – 9,840; 8,670; 9,310; 8,910; 8,820 и 3,720; 3,180; 3,550; 3,290; 3,265 соответственно. Однонаправленный выигрыш во всех пяти итерациях подтверждает устойчивость эффекта к составу тестовой части.

Табл. 4. Сравнение базовых моделей по вычислительным затратам и качеству сегментации (среднее ± SKO по пяти итерациям).
Table 4. Comparison of the baseline models in terms of computational cost and segmentation quality (mean ± SD over five folds).

Модель	Параметры, млн	Время, мс	Dice L	Dice R	HD95 L	HD95 R
U-Net [5]	31,04	4,8 ± 0,2	0,8442 ± 0,0098	0,7381 ± 0,0147	8,6267 ± 0,5124	19,1842 ± 1,4386
U-Net + крупнейшая компонента	31,04	5,1 ± 0,2	0,8508 ± 0,0101	0,7465 ± 0,0139	8,0412 ± 0,5457	17,8664 ± 1,3074
U-Net + замыкание + компонента	31,04	5,4 ± 0,2	0,8559 ± 0,0088	0,7528 ± 0,0145	7,6126 ± 0,4589	17,1025 ± 1,3921
U-Net + DenseCRF [19]	31,04	18,7 ± 0,5	0,8674 ± 0,0094	0,7669 ± 0,0126	7,0183 ± 0,4918	16,2087 ± 1,1185
nnU-Net 2d [7]	30,74	6,6 ± 0,2	0,8732 ± 0,0079	0,7764 ± 0,0135	6,7845 ± 0,4026	15,6219 ± 1,2847
U-Net+ViT-B/16 [8, 9]	105,28	12,9 ± 0,4	0,8820 ± 0,0087	0,7827 ± 0,0118	6,4218 ± 0,4473	15,0386 ± 1,0632
U-Net + ОГР (рассматриваемый)	31,04	6,5 ± 0,2	0,8989 ± 0,0083	0,8094 ± 0,0109	5,5168 ± 0,4698	12,6256 ± 1,2147
nnU-Net + ОГР	30,74	8,3 ± 0,3	0,9111 ± 0,0071	0,8258 ± 0,0116	4,8841 ± 0,3274	11,4027 ± 1,0448
U-Net+ViT + ОГР	105,28	14,6 ± 0,4	0,9164 ± 0,0068	0,8312 ± 0,0097	4,6716 ± 0,3912	11,0865 ± 0,9035

Табл. 5. Сравнение U-Net и операторно-графовой релаксации (среднее ± СКО по пяти итерациям).
Table 5. Comparison of U-Net and operator-graph relaxation (mean ± SD over five folds).

Метрика	Проекция	Базовая модель U-Net	Операторно-графовый метод
Dice L	Все	0,8442 ± 0,0098	0,8989 ± 0,0083
	Аксиальная	0,8894 ± 0,0114	0,9387 ± 0,0080
	Коронарная	0,8326 ± 0,0107	0,8896 ± 0,0089
	Сагиттальная	0,8497 ± 0,0096	0,9028 ± 0,0074
Dice R	Все	0,7381 ± 0,0147	0,8094 ± 0,0109
	Аксиальная	0,7533 ± 0,0189	0,8261 ± 0,0146
	Коронарная	0,7350 ± 0,0156	0,8042 ± 0,0117
	Сагиттальная	0,7392 ± 0,0142	0,8122 ± 0,0102
HD95 L	Все	8,6267 ± 0,5124	5,5168 ± 0,4698
	Аксиальная	5,0990 ± 0,2444	1,3600 ± 0,1275
	Коронарная	8,7870 ± 0,5870	5,6370 ± 0,4762
	Сагиттальная	8,9300 ± 0,5510	5,9410 ± 0,4217
HD95 R	Все	19,1842 ± 1,4386	12,6256 ± 1,2147
	Аксиальная	9,1100 ± 0,4719	3,4010 ± 0,2256
	Коронарная	19,1330 ± 1,5170	13,2320 ± 1,3374
	Сагиттальная	20,5450 ± 1,4810	13,2370 ± 0,9846
Contour F1 L	Все	0,7896 ± 0,0152	0,8703 ± 0,0131
	Аксиальная	0,8758 ± 0,0136	0,9100 ± 0,0097
	Коронарная	0,7754 ± 0,0164	0,8947 ± 0,0143
	Сагиттальная	0,7922 ± 0,0148	0,8415 ± 0,0124
Contour F1 R	Все	0,6491 ± 0,0198	0,7536 ± 0,0169
	Аксиальная	0,7064 ± 0,0224	0,8334 ± 0,0172
	Коронарная	0,6378 ± 0,0207	0,7409 ± 0,0138
	Сагиттальная	0,6526 ± 0,0186	0,7555 ± 0,0165
Overreach L	Все	0,3786 ± 0,0275	0,2023 ± 0,0216
	Аксиальная	0,2058 ± 0,0221	0,1374 ± 0,0157
	Коронарная	0,3801 ± 0,0294	0,2033 ± 0,0235
	Сагиттальная	0,3997 ± 0,0267	0,2098 ± 0,0194
Overreach R	Все	0,3714 ± 0,0314	0,1979 ± 0,0254
	Аксиальная	0,3312 ± 0,0285	0,2389 ± 0,0231
	Коронарная	0,4198 ± 0,0342	0,2097 ± 0,0268
	Сагиттальная	0,3297 ± 0,0299	0,1812 ± 0,0217
Underreach L	Все	0,0849 ± 0,0087	0,0825 ± 0,0079
	Аксиальная	0,0835 ± 0,0101	0,0781 ± 0,0094
	Коронарная	0,0849 ± 0,0089	0,0809 ± 0,0087
	Сагиттальная	0,0852 ± 0,0082	0,0846 ± 0,0081
Underreach R	Все	0,1622 ± 0,0126	0,1556 ± 0,0119
	Аксиальная	0,1754 ± 0,0162	0,1050 ± 0,0128
	Коронарная	0,1469 ± 0,0118	0,1468 ± 0,0114
	Сагиттальная	0,1753 ± 0,0134	0,1707 ± 0,0121

Табл. 6. Парный критерий Уилкоксона по показателям пациентов (n = 200).
Table 6. Paired Wilcoxon test for patient-level out-of-fold metrics (n = 200).

Метрика	Z	p	pHolm	$r = \frac{ Z }{\sqrt{n}}$
Dice L	−8,61	< 0,001	< 0,001	0,609
Dice R	−8,13	< 0,001	< 0,001	0,575
HD95 L	−7,72	< 0,001	< 0,001	0,546
HD95 R	−7,49	< 0,001	< 0,001	0,530
Contour F1 L	−8,36	< 0,001	< 0,001	0,591
Contour F1 R	−8,02	< 0,001	< 0,001	0,567
Overreach L	−8,77	< 0,001	< 0,001	0,620
Overreach R	−8,51	< 0,001	< 0,001	0,602
Underreach L	−2,94	0,0033	0,0066	0,208
Underreach R	−1,83	0,0672	0,0672	0,129

4.3 Абляционное исследование

Результаты абляционного исследования, выполненного по протоколу раздела 3.7, приведены в табл. 7.

Табл. 7. Вклад отдельных компонентов метода (среднее ± СКО по пяти итерациям).
Table 7. Contribution of the individual components (mean ± SD over five folds).

Конфигурация	Оператор Λ	Матрица C	Dice L	Dice R	HD95 R	Contour F1 R
M1 – U-Net	не применяется	не применяется	0,8442 ± 0,0098	0,7381 ± 0,0147	19,1842 ± 1,4386	0,6491 ± 0,0198
M2 – взвешенное усреднение	Λ = I	не применяется	0,8687 ± 0,0092	0,7708 ± 0,0141	16,0314 ± 1,3578	0,6975 ± 0,0206
M3 – направленное затухание	Λ = γ · I	κ = 0	0,8906 ± 0,0080	0,7981 ± 0,0117	13,5572 ± 1,2435	0,7389 ± 0,0162
M4 – ОГР (полная)	Λ = γ · C	κ = 0,15	0,8989 ± 0,0083	0,8094 ± 0,0109	12,6256 ± 1,2147	0,7536 ± 0,0169

Переход от M1 к M2 увеличивает значение метрики Dice на 0,0245 для LV и на 0,0327 для RV, что характеризует вклад локального графового согласования. Добавление направленного затухания (переход от M2 к M3) дает еще 0,0219 и 0,0273 и снижает HD95 R на 2,474 пикселя. Матрица анатомической совместимости (переход от M3 к M4) обеспечивает дополнительный, но устойчивый прирост Dice 0,0083/0,0113 и уменьшение HD95 R на 0,932 пикселя. Таким образом, ни один компонент не объясняет итоговый результат в отдельности; наибольший вклад вносит направленный перенос, а матрица C усиливает разделение LV/RV. Чувствительность среднего коэффициента Дайса к параметру релаксации и числу итераций приведена в табл. 8. Максимум на внутренних валидационных подвыборках составил 0,8483 при α = 0,35 и T_{max} = 20; он отличается от среднего тестового значения 0,8542, полученного только после фиксации гиперпараметров (табл. 8).

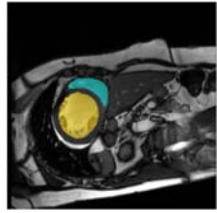
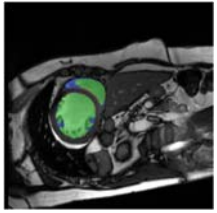
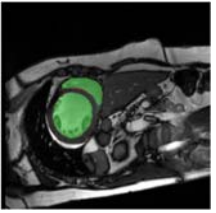
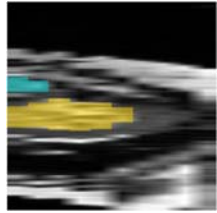
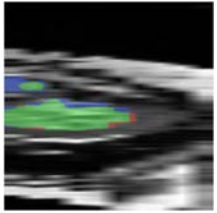
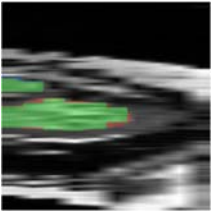
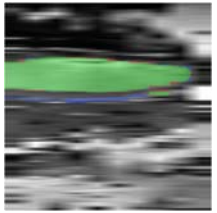
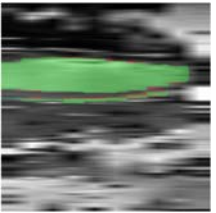
Табл. 8. Чувствительность среднего Dice (LV/RV) к α и T_{max} на валидационных данных.
Table 8. Sensitivity of mean Dice (LV/RV) to α and T_{max} on validation data.

$\alpha \setminus T_{max}$	4	8	12	16	20
0,15	0,8108	0,8179	0,8198	0,8201	0,8201
0,25	0,8226	0,8344	0,8381	0,8387	0,8388
0,35	0,8297	0,8436	0,8475	0,8481	0,8483
0,45	0,8271	0,8406	0,8442	0,8440	0,8436
0,55	0,8192	0,8328	0,8351	0,8339	0,8318

4.4 Визуальный анализ результатов

В табл. 9 сопоставлены результаты сегментации базовой модели и операторно-графового метода на репрезентативных срезах трех ортогональных проекций. Для каждого среза представлены экспертная разметка и предсказанные маски; зеленым обозначена истинно положительная область, синим – ложноотрицательная, красным – ложноположительная. Сравнительный визуальный анализ результатов (табл. 9) согласуется с количественными оценками (табл. 5), но служит только иллюстрацией и не заменяет статистического анализа. На аксиальных срезах обе модели локализируют левый желудочек, однако в маске U-Net заметны локальные пропуски. Для правого желудочка ОГР формирует более непрерывную границу, визуально ближе соответствующую экспертной разметке, тогда как базовая модель допускает фрагментацию в апикальной части.

Табл. 9. Сравнение экспертной разметки и предсказаний моделей.
Table 9. Comparison of expert annotations and model predictions.

Проекция	Экспертная разметка	Базовая модель U-Net	Операторно-графовый метод
Аксиальная			
Коронарная			
Сагиттальная			

На коронарной проекции экспертная разметка демонстрирует вытянутую форму левого желудочка и узкую область правого желудочка в верхней части среза. Базовая модель U-Net воспроизводит общий контур левого желудочка, но покрывает лишь часть размеченной области правого желудочка. После ОГР контур левого желудочка становится более гладким, а маска правого желудочка визуально лучше соответствует экспертной разметке. Эти наблюдения характеризуют выбранные примеры и не должны интерпретироваться как самостоятельное доказательство клинической применимости. На сагиттальной проекции различия наиболее заметны для правого желудочка, представленного узкой протяженной областью. Базовая модель выделяет лишь часть этой области, тогда как ОГР формирует более непрерывное покрытие вдоль ее протяженности. Поскольку межсрезовая связность и топологические показатели в работе не оценивались, данный пример свидетельствует только о локальном улучшении двумерной маски и не подтверждает трехмерную согласованность результата.

4.5 Результаты внешнего тестирования

Результаты применения зафиксированной версии алгоритма к внешней выборке ООО «Диалог Плюс» приведены в табл. 10. Для обеих моделей указаны среднее и стандартное отклонение по 100 пациентам, парная разность $\Delta = \text{ОГР} - \text{U-Net}$ с 95-процентным доверительным интервалом, а также относительное изменение среднего значения ОГР по сравнению с внутренней оценкой (табл. 5).

Табл. 10. Результаты внешнего тестирования на выборке ООО «Диалог Плюс» ($n = 100$ пациентов; среднее $\pm$ СКО по пациентам).
Table 10. External testing results on the Dialog Plus LLC cohort ($n = 100$ patients; mean $\pm$ SD across patients).

Метрика	U-Net	ОГР	Δ (ОГР – U-Net), 95 % ДИ	Изменение ОГР к внутренней оценке
Dice L	0,8021 $\pm$ 0,0647	0,8719 $\pm$ 0,0512	+0,0698 [0,0612; 0,0784]	–3,0 %
Dice R	0,6968 $\pm$ 0,0975	0,7795 $\pm$ 0,0834	+0,0827 [0,0709; 0,0945]	–3,7 %
HD95 L	9,3772 $\pm$ 2,6183	5,7154 $\pm$ 1,9436	–3,6618 [–4,1204; –3,2032]	+3,6 %
HD95 R	20,9875 $\pm$ 5,4128	13,1180 $\pm$ 4,2871	–7,8695 [–8,8712; –6,8678]	+3,9 %
Contour F1 L	0,7478 $\pm$ 0,0736	0,8425 $\pm$ 0,0604	+0,0947 [0,0836; 0,1058]	–3,2 %
Contour F1 R	0,6108 $\pm$ 0,1042	0,7250 $\pm$ 0,0917	+0,1142 [0,0995; 0,1289]	–3,8 %
Overreach L	0,4118 $\pm$ 0,0958	0,2092 $\pm$ 0,0721	–0,2026 [–0,2185; –0,1867]	+3,4 %
Overreach R	0,4059 $\pm$ 0,1074	0,2050 $\pm$ 0,0846	–0,2009 [–0,2183; –0,1835]	+3,6 %
Underreach L	0,0909 $\pm$ 0,0355	0,0851 $\pm$ 0,0318	–0,0058 [–0,0102; –0,0014]	+3,1 %
Underreach R	0,1749 $\pm$ 0,0537	0,1616 $\pm$ 0,0492	–0,0133 [–0,0284; +0,0018]	+3,9 %

Статистическая проверка парных различий на внешней выборке выполнена по протоколу раздела 3.9; результаты приведены в табл. 11. На внешней выборке абсолютные значения всех десяти метрик ухудшились по сравнению с внутренней оценкой, однако ухудшение оказалось умеренным и однородным: относительное изменение среднего значения лежит в диапазоне от 3,0 до 3,9 % и не превышает 4 % ни по одному показателю. Базовая модель при этом потеряла в качестве заметно больше: относительное ухудшение коэффициента Дайса составило 5,0 % для левого желудочка и 5,6 % для правого против 3,0 и 3,7 % у ОГР, а расстояние HD95 выросло на 8,7 и 9,4 % против 3,6 и 3,9 % соответственно. Различие имеет содержательное объяснение: параметры ОГР

задаются аналитически и не настраиваются на распределение обучающей выборки, поэтому смена сканера, протокола и практики разметки влияет на них лишь опосредованно – через качество входного вероятностного поля. Следствием является то, что абсолютная величина выигрыша от структурного согласования на внешних данных не только сохранилась, но и возросла: прирост коэффициента Дайса составил 0,0698 для левого желудочка и 0,0827 для правого против 0,0547 и 0,0713 во внутренней оценке.

Табл. 11. Парный критерий Уилкоксона на внешней выборке (n = 100).
Table 11. Paired Wilcoxon test on the external cohort (n = 100).

Метрика	Z	p	pHolm	$r = Z / \sqrt{n}$
Dice L	-7,12	< 0,001	< 0,001	0,712
Dice R	-6,84	< 0,001	< 0,001	0,684
HD95 L	-6,41	< 0,001	< 0,001	0,641
HD95 R	-6,27	< 0,001	< 0,001	0,627
Contour F1 L	-6,95	< 0,001	< 0,001	0,695
Contour F1 R	-6,73	< 0,001	< 0,001	0,673
Overreach L	-7,28	< 0,001	< 0,001	0,728
Overreach R	-7,05	< 0,001	< 0,001	0,705
Underreach L	-2,61	0,0090	0,0181	0,261
Underreach R	-1,74	0,0819	0,0819	0,174

Направление и структура эффекта на внешней выборке воспроизвели внутренний результат. После поправки Холма–Бонферрони различия статистически значимы для моделей Dice, HD95, Contour F1 и Overreach обоих желудочков (pHolm < 0,001; r = 0,627–0,728), а также для Underreach левого желудочка (p = 0,0090; pHolm = 0,0181; r = 0,261). Для Underreach правого желудочка различие вновь не достигло уровня значимости (p = 0,0819; r = 0,174), и 95-процентный доверительный интервал парной разности содержит ноль (табл. 10). Таким образом, преимущество ОГР статистически подтверждено по девяти показателям из десяти, а единственное исключение совпадает с исключением, полученным на внутренней выборке; это указывает скорее на систематическое ограничение метода в части восстановления пропущенных областей правого желудочка, чем на случайную вариацию или недостаток статистической мощности.

Полученный результат не следует трактовать как доказательство универсальной переносимости. Учреждения-источники расположены в разных субъектах Российской Федерации, однако обе выборки получены на аппаратуре с напряженностью магнитного поля 1,5 Т и в пределах одного макрорегиона, а объем внешней выборки (100 пациентов) позволяет оценить эффекты наблюдаемой величины, но недостаточен для надежной оценки редких режимов отказа. Ограниченная переносимость моделей глубокого обучения между центрами остается известной проблемой анализа медицинских изображений [35]; многоцентровые и многовендорные исследования кардиологической сегментации [36] подтверждают чувствительность качества к смене производителя оборудования и протокола сканирования. В настоящей работе варьировался производитель, но напряженность магнитного поля оставалась равной 1,5 Т. Расширение проверки на аппаратуру 3 Т, иные протоколы и данные других макрорегионов составляет предмет дальнейшей работы.

5. Обсуждение результатов

Полученные результаты показывают, что основной эффект операторно-графовой релаксации сосредоточен на границе: относительное улучшение HD95 и Contour F1 превышает прирост коэффициента Dice. Это согласуется с аффинным правилом (11): исходное поле U-Net сохраняется как унарная опора, а перераспределение вероятностей происходит прежде всего

в неопределенной приграничной зоне. Сравнение с DenseCRF и морфологическими процедурами подтверждает, что выигрыш не сводится к удалению малых компонент или симметричному сглаживанию.

Отдельного комментария требует сопоставление с морфологической постобработкой. Она удаляет изолированные ложноположительные включения, однако опирается лишь на геометрию бинарной маски и не использует ни исходное поле неопределенности, ни направленную информацию о границе, а потому неспособна восстановить пропущенный фрагмент целевой структуры. Предлагаемый метод действует в обе стороны, и величина второго эффекта поддается количественной оценке. Для Underreach левого желудочка снижение среднего значения с 0,0849 до 0,0825 значимо после поправки Холма–Бонферрони, но величина эффекта мала (r = 0,208); для Underreach правого желудочка среднее также снизилось, однако уровня значимости различие не достигло (pHolm = 0,0672). Надежно подтвержденными, таким образом, следует считать эффекты по Dice, HD95, Contour F1 и Overreach, тогда как вывод о восстановлении пропущенных областей правого желудочка носит предварительный характер: на внешней выборке различие по Underreach правого желудочка также не достигло уровня значимости (табл. 11), что указывает скорее на систематическое ограничение метода в этой части, чем на недостаток статистической мощности.

5.1 Внешняя обобщаемость и границы клинической интерпретации

Внешнее тестирование, описанное в разделах 3.9 и 4.5, устраняет наиболее существенное ограничение предыдущей версии исследования, но одновременно очерчивает границы допустимых выводов. Разбиение на уровне пациентов внутри основной выборки предотвращает утечку данных между обучающей и тестовой частями, однако само по себе не моделирует сдвиг распределения, обусловленный иным сканером, протоколом, составом пациентов и практикой разметки; именно этот сдвиг и воспроизводит внешняя выборка. Результаты показывают, что при переходе к другому учреждению и другому производителю оборудования эффект структурного согласования сохраняется по девяти показателям из десяти, а абсолютное качество снижается не более чем на 3,9 %. Это позволяет говорить о переносимости эффекта в пределах исследованных условий – два учреждения, две аппаратные платформы, напряженность поля 1,5 Т – но не об универсальной обобщаемости метода.

Три ограничения внешней проверки требуют отдельной оговорки. Во-первых, факторы сдвига домена в проведенном эксперименте изменяются совместно: одновременно различаются производитель и модель томографа, протокол сканирования, состав пациентов и практика контурирования, поэтому вклад каждого из них в наблюдаемое снижение качества по имеющимся данным разделить нельзя; полученные 3,0–3,9 % характеризуют совокупный эффект перехода между площадками, а не влияние отдельно взятого фактора. Нозологический состав выборок при этом не каталогизировался, вследствие чего оценка переносимости не стратифицирована по патологии и не позволяет выделить группы пациентов, для которых метод может работать хуже среднего; кроме того, внешняя выборка набиралась на протяжении семи лет, в течение которых могли измениться версия программного обеспечения томографа и детали протокола, что вносит дополнительную внутреннюю неоднородность. Во-вторых, внешний набор размечен одним специалистом, вследствие чего межнаблюдательная вариабельность внешнего эталона не оценивалась, и часть остаточного расхождения между предсказаниями и разметкой может объясняться индивидуальными особенностями контурирования, а не ошибками алгоритма. В-третьих, отказ от согласования разметки между учреждениями является одновременно достоинством и ограничением постановки: он сохраняет естественный сдвиг практик разметки, который и подлежал оценке, но исключает возможность отнести расхождения исключительно на счет качества сегментации. Устранение этих ограничений требует протокола с раздельным

варьированием факторов, со стратификацией по нозологиям и с независимой разметкой каждого исследования несколькими специалистами.

Клиническая интерпретация результатов ограничена выбранными конечными точками. В работе оценивались Dice, HD95, Contour F1, Overreach и Underreach, но не рассчитывались конечный диастолический и конечный систолический объемы, ударный объем, масса миокарда и фракция выброса; не изучались межнаблюдательная воспроизводимость, диагностические решения и исходы лечения. Расчет функциональных показателей магнитно-резонансной томографии сердца требует фазово согласованной сегментации по всему сердечному циклу и стандартизированной постобработки [32], тогда как использованные наборы данных содержат по одному статическому объему на исследование и не сопровождаются клиническим референсным стандартом. Улучшение сегментационных метрик поэтому следует рассматривать как технический результат: оно не является прямым доказательством повышения точности объемных или функциональных измерений и не обосновывает клиническое применение метода без отдельного исследования.

Осторожную количественную интерпретацию в этом направлении допускает лишь знак и величина систематической составляющей ошибки. Разность Overreach и Underreach характеризует преобладание ложноположительных пикселей над ложноотрицательными и, следовательно, направление возможного смещения при последующем вычислении площади и объема полости. Для левого желудочка эта разность снижается с +0,294 до +0,120 на внутренней выборке и с +0,321 до +0,124 на внешней; для правого желудочка – с +0,209 до +0,042 и с +0,231 до +0,043 соответственно. Операторно-графовая релаксация, таким образом, уменьшает систематическое завышение выделяемой площади приблизительно вдвое для левого желудочка и в четыре–пять раз для правого. Приведенные величины являются свойством метрик сегментации, а не результатом прямого измерения объема; вопрос о том, транслируется ли это свойство в снижение погрешности объемных показателей, требует отдельного исследования на фазово разрешенных данных с клиническим референсным стандартом.

5.2 Обоснование двумерной постановки и направление 3D-обобщения

Двумерная постановка выбрана по причинам, изложенным в разделе 3.1: она соответствует посрезовой организации экспертной разметки и клинической интерпретации, обеспечивает сопоставление методов на идентичных входах и сохраняет модульность ОГР как независимой от обучения процедуры постобработки. Согласование с 2D-конфигурацией базовых моделей позволяет не смешивать эффект операторно-графовой релаксации с преимуществами объемной архитектуры, обладающей иной емкостью и иным режимом обучения. Вместе с тем 2D-подход исключает межсрезовый контекст и не гарантирует согласованность масок в трехмерном объеме: обработка трех ортогональных реконструкций независимо друг от друга не эквивалентна объемному согласованию и в принципе допускает взаимно противоречивые решения на пересекающихся линиях. Межсрезовая связность и топологические характеристики трехмерного результата в настоящей работе не оценивались.

Полноценное 3D-обобщение требует замены восьмисвязной окрестности (5) на 26-связную, учета физического шага между срезами и анизотропии вокселя, а также сопоставления с объемными базовыми моделями, например V-Net и трехмерными конфигурациями nnU-Net [7, 28]. Принципиально важно, что доказательство сжимаемости раздела 3.5 не опирается на мощность окрестности: условие нормировки (7) и равенство $\|C\|_\infty = 1$ сохраняются при любом конечном числе соседей, поэтому оценка $q = \alpha \gamma_{\max} < 1$, существование и единственность неподвижной точки, геометрическая скорость сходимости и оценка (19) остаются в силе и для 26-связной окрестности. Переход к трехмерной постановке является, следовательно, вопросом вычислительных затрат, организации данных и сравнительного протокола, а не

пересмотра теоретического обоснования метода. Сопоставление постановок приведено в табл. 12.

Дополнительным ограничением остается упрощенная матрица анатомической совместимости (10): единственный параметр k описывает взаимное подавление меток левого и правого желудочков, но не различает типы переходов между смежными структурами – миокардом, предсердиями и магистральными сосудами. Расширение матрицы до полного набора анатомических пар и ее проверка на независимых данных должны предшествовать выводам о переносимости метода на иные задачи сегментации.

Табл. 12. Сопоставление 2D-постановки настоящей работы и потенциального 3D-обобщения.
Table 12. Comparison of the present 2D formulation with a potential 3D extension.

Критерий	2D-постановка настоящей работы	Потенциальная 3D-постановка	Следствие для интерпретации
Пространственный контекст	Один срез 256×256; восьмисвязная окрестность.	Полный объем; 26-связная окрестность.	2D-метод не моделирует истинную межсрезовую связность.
Сопоставимость	Одинаковые 2D-входы позволяют изолировать вклад ОГР.	Требуются объемные базовые модели и отдельное обучение.	Работа не является прямым сравнением качества 2D- и 3D-сетей.
Геометрия данных	Три вычислительные ортогональные реконструкции обрабатываются независимо.	Учитываются физический размер вокселя и возможная анизотропия.	Три проекции не заменяют объемную согласованность.
Ресурсы и модульность	Меньше памяти; применимость к готовому 2D-вероятностному полю.	Больше памяти; иные требования к данным и разметке.	2D выбрана для модульной постобработки, а не как заведомо более точная постановка.
Клинические конечные точки	Оценены сегментационные метрики отдельных срезов.	Возможна фазово согласованная объемная оценка.	EDV, ESV и EF в настоящей работе не оценены.

6. Заключение

Предложен метод операторно-графовой релаксации, уточняющий вероятностные карты сегментации левого и правого желудочков на двумерных срезах магнитно-резонансной томографии. Метод задан в замкнутом виде: определены граф изображения и восьмисвязная окрестность, нормированные веса ребер, несимметричный оператор переноса состояния с множителем направленного затухания и матрицей анатомической совместимости классов, аффинная итерация с привязкой к исходному вероятностному полю, проецирование на допустимое множество состояний и количественный критерий остановки. Для итерационного отображения доказана сжимаемость с константой $q = \alpha \gamma_{\max} < 1$, откуда следуют существование и единственность неподвижной точки, геометрическая скорость сходимости и явная оценка требуемого числа итераций. Метод не содержит обучаемых параметров и применяется на этапе вывода, не изменяя обучение базовой модели.

В рамках внутренней пятикратной кросс-валидации коэффициент Дайса вырос с 0,844 до 0,899 для левого желудочка и с 0,738 до 0,809 для правого, а расстояние HD95 для правого желудочка снизилось с 19,18 до 12,63 пикселя. После поправки Холма–Бонферрони значимыми оказались девять сравнений из десяти; для Underreach правого желудочка уровень значимости не достигнут. Метод превзошел отбор связанных компонент,

морфологическое замыкание и полносвязное условное случайное поле, а также улучшил выход моделей U-Net, nnU-Net и U-Net+ViT без добавления обучаемых параметров. Средняя добавка времени составила 1,7 мс на срез, или 35,4 % времени прямого прохода U-Net.

На независимой внешней выборке из 100 исследований, выполненных в другом медицинском учреждении на томографе иного производителя, зафиксированная до анализа версия алгоритма сохранила преимущество над базовой моделью по девяти показателям из десяти при том же единственном исключении – Underreach правого желудочка. Абсолютные значения метрик снизились на 3,0–3,9 % относительно внутренней оценки, тогда как базовая модель потеряла от 5,0 до 9,4 %, вследствие чего абсолютная величина прироста коэффициента Дайса на внешних данных оказалась выше внутренней: 0,070 и 0,083 против 0,055 и 0,071 для левого и правого желудочков соответственно.

В исследованных условиях операторно-графовую релаксацию можно рассматривать как воспроизводимый модуль структурного уточнения вероятностного поля произвольной сегментационной модели, сохраняющий эффективность при смене учреждения и производителя оборудования. Полученные результаты характеризуют качество сегментации; прямая оценка влияния метода на объемные и функциональные показатели сердца, равно как и проверка на аппаратуре с иной напряженностью магнитного поля и на данных других регионов, составляет предмет дальнейшей работы наряду с обобщением на трехмерную связность и автоматической настройкой параметров по свойствам конкретного исследования.

Список литературы / References

- [1]. Сердечно-сосудистые заболевания [Электронный ресурс] // Всемирная организация здравоохранения. – URL: <https://www.who.int/ru/news-room/fact-sheets/detail/cardiovascular-diseases-cvds> (дата обращения: 01.05.2026) // Cardiovascular diseases (CVDs). World Health Organization. Available at: <https://www.who.int/ru/news-room/fact-sheets/detail/cardiovascular-diseases-cvds>, accessed 01.05.2026 (in Russian).
- [2]. Федеральная служба государственной статистики (Росстат). Естественное движение населения Российской Федерации за 2022 год: статистический бюллетень. М.: Росстат, 2023, 75 с. // Federal State Statistics Service (Rosstat). Natural Population Change in the Russian Federation in 2022: Statistical Bulletin. Moscow: Rosstat, 2023, 75 pp. (in Russian).
- [3]. Chen C., Qin C., Qiu H., Tarroni G., Duan J., Bai W., Rueckert D. Deep learning for cardiac image segmentation: a review. *Frontiers in Cardiovascular Medicine*, 2020, vol. 7, p. 25. DOI: 10.3389/fcvm.2020.00025.
- [4]. Kass M., Witkin A., Terzopoulos D. Snakes: active contour models. *International Journal of Computer Vision*, 1988, vol. 1, no. 4, pp. 321-331. DOI: 10.1007/BF00133570.
- [5]. Ronneberger O., Fischer P., Brox T. U-Net: convolutional networks for biomedical image segmentation // *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015)*. Lecture Notes in Computer Science, 2015, vol. 9351, pp. 234-241. DOI: 10.1007/978-3-319-24574-4_28.
- [6]. Bernard O., Lalonde A., Zotti C. et al. Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved? *IEEE Transactions on Medical Imaging*, 2018, vol. 37, no. 11, pp. 2514-2525. DOI: 10.1109/TMI.2018.2837502.
- [7]. Isensee F., Jaeger P.F., Kohl S.A.A., Petersen J., Maier-Hein K.H. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 2021, vol. 18, no. 2, pp. 203-211. DOI: 10.1038/s41592-020-01008-z.
- [8]. Dosovitskiy A., Beyer L., Kolesnikov A. et al. An image is worth 16x16 words: transformers for image recognition at scale // *International Conference on Learning Representations (ICLR 2021)*, 2021. DOI: 10.48550/arXiv.2010.11929.
- [9]. Chen J., Lu Y., Yu Q. et al. TransUNet: transformers make strong encoders for medical image segmentation, 2021. DOI: 10.48550/arXiv.2102.04306.
- [10]. Guo S., Zhao X., Ren J. et al. Efficient cardiac MRI multi-structure segmentation for cardiovascular assessment with limited annotation by integrating data-level and network-level consistency. *npj Digital Medicine*, 2026, vol. 9, p. 328. DOI: 10.1038/s41746-026-02475-y.

- [11]. Van De Vyver G., Thomas S., Ben-Yosef G. et al. Toward robust cardiac segmentation using graph convolutional networks. *IEEE Access*, 2024, vol. 12, pp. 33876–33888. DOI: 10.1109/ACCESS.2024.3373046.
- [12]. Kipf T.N., Welling M. Semi-supervised classification with graph convolutional networks // *International Conference on Learning Representations (ICLR 2017)*, 2017. DOI: 10.48550/arXiv.1609.02907.
- [13]. Painchaud N., Skandarani Y., Judge T. et al. Cardiac segmentation with strong anatomical guarantees. *IEEE Transactions on Medical Imaging*, 2020, vol. 39, no. 11, pp. 3703-3713. DOI: 10.1109/TMI.2020.3003240.
- [14]. Rosenfeld A., Hummel R.A., Zucker S.W. Scene labeling by relaxation operations. *IEEE Transactions on Systems, Man and Cybernetics*, 1976, vol. SMC-6, no. 6, pp. 420-433. DOI: 10.1109/TSMC.1976.4309519.
- [15]. Hummel R.A., Zucker S.W. On the foundations of relaxation labeling processes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1983, vol. 5, no. 3, pp. 267-287. DOI: 10.1109/TPAMI.1983.4767390.
- [16]. Geman S., Geman D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1984, vol. 6, no. 6, pp. 721-741. DOI: 10.1109/TPAMI.1984.4767596.
- [17]. Boykov Y., Jolly M.-P. Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. *Proceedings of the 8th IEEE International Conference on Computer Vision (ICCV 2001)*, 2001, vol. 1, pp. 105-112. DOI: 10.1109/ICCV.2001.937505.
- [18]. Boykov Y., Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, vol. 26, no. 9, pp. 1124-1137. DOI: 10.1109/TPAMI.2004.60.
- [19]. Krähenbühl P., Koltun V. Efficient inference in fully connected CRFs with Gaussian edge potentials. *Advances in Neural Information Processing Systems (NIPS 2011)*, 2011, pp. 109-117.
- [20]. Zheng S., Jayasumana S., Romera-Paredes B. et al. Conditional random fields as recurrent neural networks. *Proceedings of the IEEE International Conference on Computer Vision (ICCV 2015)*, 2015, pp. 1529-1537. DOI: 10.1109/ICCV.2015.179.
- [21]. Chen L.-C., Papandreou G., Kokkinos I., Murphy K., Yuille A.L. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, vol. 40, no. 4, pp. 834-848. DOI: 10.1109/TPAMI.2017.2699184.
- [22]. Kamnitsas K., Ledig C., Newcombe V.F.J. et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis*, 2017, vol. 36, pp. 61-78. DOI: 10.1016/j.media.2016.10.004.
- [23]. Oktay O., Ferrante E., Kamnitsas K. et al. Anatomically constrained neural networks (ACNNs): application to cardiac image enhancement and segmentation. *IEEE Transactions on Medical Imaging*, 2018, vol. 37, no. 2, pp. 384-395. DOI: 10.1109/TMI.2017.2743464.
- [24]. Hu X., Li F., Samaras D., Chen C. Topology-preserving deep image segmentation. *Advances in Neural Information Processing Systems (NeurIPS 2019)*, 2019, pp. 5657-5668.
- [25]. Clough J.R., Byrne N., Oksuz I. et al. A topological loss function for deep-learning based image segmentation using persistent homology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, vol. 44, no. 12, pp. 8766-8778. DOI: 10.1109/TPAMI.2020.3013679.
- [26]. Perona P., Malik J. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, vol. 12, no. 7, pp. 629-639. DOI: 10.1109/34.56205.
- [27]. Weickert J. Anisotropic diffusion in image processing. Stuttgart: Teubner, 1998, 170 p.
- [28]. Milletari F., Navab N., Ahmadi S.-A. V-Net: fully convolutional neural networks for volumetric medical image segmentation. *Proceedings of the Fourth International Conference on 3D Vision (3DV)*, 2016, pp. 565-571. DOI: 10.1109/3DV.2016.79.
- [29]. Dice L.R. Measures of the amount of ecologic association between species. *Ecology*, 1945, vol. 26, no. 3, pp. 297-302. DOI: 10.2307/1932409.
- [30]. Huttenlocher D.P., Klanderman G.A., Rucklidge W.J. Comparing images using the Hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1993, vol. 15, no. 9, pp. 850-863. DOI: 10.1109/34.232073.
- [31]. Yeghiazaryan V., Voiculescu I. Family of boundary overlap metrics for the evaluation of medical image segmentation. *Journal of Medical Imaging*, 2018, vol. 5, no. 1, pp. 015006. DOI: 10.1117/1.JMI.5.1.015006.

- [32]. Schulz-Menger J., Bluemke D.A., Bremerich J. et al. Standardized image interpretation and post-processing in cardiovascular magnetic resonance – 2020 update. *Journal of Cardiovascular Magnetic Resonance*, 2020, vol. 22, article 19. DOI: 10.1186/s12968-020-00610-6.
- [33]. Wilcoxon F. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1945, vol. 1, no. 6, pp. 80-83. DOI: 10.2307/3001968.
- [34]. Tejani A.S., Klontzas M.E., Gatti A.A., Mongan J.T., Moy L., Park S.H., Kahn C.E. Jr.; CLAIM 2024 Update Panel. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): 2024 Update. *Radiology: Artificial Intelligence*, 2024, vol. 6, no. 4, e240300. DOI: 10.1148/ryai.240300.
- [35]. Litjens G., Kooi T., Bejnordi B.E. et al. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 2017, vol. 42, pp. 60-88. DOI: 10.1016/j.media.2017.07.005.
- [36]. Campello V.M., Gkontra P., Izquierdo C. et al. Multi-centre, multi-vendor and multi-disease cardiac segmentation: the M&Ms challenge. *IEEE Transactions on Medical Imaging*, 2021, vol. 40, no. 12, pp. 3543-3554. DOI: 10.1109/TMI.2021.3090082.

Информация об авторах / Information about authors

Петр Андреевич ПЫЛОВ – аспирант Кузбасского государственного технического университета имени Т.Ф. Горбачева. Сфера научных интересов: компьютерное зрение, обработка естественного языка, глубокое обучение, разработка интеллектуальных систем для автоматизации различных прикладных задач.

Petr Andreevich PYLOV – postgraduate student at the T.F. Gorbachev Kuzbass State Technical University. Research interests: computer vision, natural language processing, deep learning, development of intelligent systems for automation of various applied tasks.

Роман Вячеславович МАЙТАК – аспирант Кузбасского государственного технического университета имени Т.Ф. Горбачева. Научные интересы: обработка естественного языка, глубокое обучение, обработка текстовых и числовых данных моделями машинного обучения, автоматизация технологических задач.

Roman Viacheslavovich MAITAK – postgraduate student at the T.F. Gorbachev Kuzbass State Technical University. Research interests: natural language processing, deep learning, processing of text and numerical data with machine learning models, automation of technological tasks.

Александр Григорьевич ПИМОНОВ – доктор технических наук, профессор, заведующий кафедрой прикладных информационных технологий Кузбасского государственного технического университета имени Т.Ф. Горбачева. Сфера научных интересов: системы искусственного интеллекта, информационные технологии численного анализа и имитационного моделирования сложных систем.

Alexander Grigor'evich PIMONOV – Dr. Sci. (Tech.), Prof., Head of the Department of Applied Information Technologies at the T.F. Gorbachev Kuzbass State Technical University. Research interests: artificial intelligence systems, information technologies for numerical analysis and simulation of complex systems.