

Григорян Карен Сергеевич – Южный федеральный университет; e-mail: kgr@sfedu.ru; г. Таганрог, Россия; тел.: +79889490693; кафедра безопасности информационных технологий им. О.Б. Макаревича; аспирант.

Басан Елена Сергеевна – Южный федеральный университет; e-mail: ebasan@sfedu.com; г. Таганрог, Россия; тел.: +79880158697; кафедра безопасности информационных технологий им. О.Б. Макаревича; доцент.

Grigoryan Karen Sergeevich – Southern Federal University; e-mail: kgr@sfedu.ru; Taganrog, Russia; phone: +79889490693; the Department of Information Technology Security named after O.B. Makarevich; postgraduate student.

Basan Elena Sergeevna – Southern Federal University; e-mail: ebasan@sfedu.com; Taganrog, Russia; phone: +79880158697; the Department of Information Technology Security named after O.B. Makarevich; associate professor.

УДК 004.932

DOI 10.18522/2311-3103-2026-4-77-90

С.В. Ермоленко, И.Л. Каширина

УЛУЧШЕНИЕ СЕГМЕНТАЦИИ НА МУЛЬТИФАЗНЫХ КТ СНИМКАХ С ПОМОЩЬЮ ОБУЧАЕМОГО ФАЗОВОГО СОВМЕЩЕНИЯ

Автоматизированный анализ мультифазных КТ снимков, получаемых в разные моменты времени после введения контрастного вещества, крайне важен для точного выявления патологий, но связан с фундаментальной проблемой пространственного несовпадения изображений на разных фазах из-за дыхания и движения пациента, что приводит к значительному снижению точности автоматической сегментации. Существующие методы предварительной регистрации (совмещения фаз) требуют ручной настройки параметров и не интегрируются в обучаемые конвейеры, что препятствует полной автоматизации процесса сегментации. Цель данного исследования заключалась в разработке дифференцируемого метода совмещения мультифазных КТ-изображений на основе обучаемых линейных аффинных преобразований, полностью встроенного в конвейер обучения модели сегментации. В отличие от традиционных подходов, предложенный метод реализует дифференцируемый модуль аффинной регистрации (включающий перенос, поворот и масштабирование), параметры которого оптимизируются градиентным спуском без ручной настройки и встроены непосредственно в вычислительный граф модели nnU-Net. В работе на открытом датасете КТ органов брюшной полости WAW-TACE проведено сравнение предлагаемого метода с базовым подходом сегментации мультифазных изображений (основанном на простой конкатенации фаз без регистрации) и методом аффинной регистрации, реализованным в библиотеке ITK.Elastix. Относительно сегментации без регистрации достигнут существенный прирост метрик качества: коэффициента Dice на 43,19% и метрики ROC-AUC на 13,55%. По сравнению с Elastix улучшение составило 13,43% по Dice, 6,86% по ROC-AUC и 19,6% по точности определения количества патологий на КТ-снимке. Практическая значимость исследования заключается в разработке готового модуля на PyTorch для интеграции в конвейеры анализа КТ-изображений, обеспечивающего полную автоматизацию регистрации без необходимости в ручной настройке гиперпараметров и с высокой вычислительной эффективностью.

Мультифазная компьютерная томография; сегментация медицинских изображений; дифференцируемая регистрация; глубокое обучение; nnU-Net; аффинное преобразование.

S.V. Ermolenko, I.L. Kashirina

IMPROVING SEGMENTATION IN MULTIPHASE CT IMAGES USING TRAINABLE PHASE SUPERIMAGING

Joint Automated analysis of multiphase CT scans, acquired at different time points after contrast agent administration, is highly important for accurate pathology detection. However, it faces a fundamental problem of spatial misalignment between phases due to patient breathing and movement, which leads to a significant reduction in the accuracy of automatic segmentation. Existing pre-registration (phase alignment) methods require manual parameter tuning and are not integrated into trainable pipelines, hindering the full au-

tomation of the segmentation process. The aim of this study was to develop a differentiable method for aligning multiphase CT images based on trainable linear affine transformations, fully embedded into the segmentation model training pipeline. Unlike traditional approaches, the proposed method implements a differentiable affine registration module (including translation, rotation, and scaling), whose parameters are optimized via gradient descent without manual tuning and are directly integrated into the computational graph of the nnU-Net model. The study compared the proposed method with a baseline multiphase segmentation approach (based on simple phase concatenation without registration) and an affine registration method implemented using the ITK.Elastix library on the open abdominal CT dataset WAW-TACE. Compared to segmentation without registration, a substantial improvement in quality metrics was achieved: an increase in the Dice coefficient by 43.19% and in the ROC-AUC metric by 13.55%. Compared to Elastix, the improvements were 13.43% in Dice, 6.86% in ROC-AUC, and 19.6% in the accuracy of pathology count detection per CT scan. The practical significance of the research lies in the development of a ready-to-use PyTorch module for integration into CT image analysis pipelines. It enables fully automated registration without the need for manual hyperparameter tuning and offers high computational efficiency.

Multiphase computed tomography; medical image segmentation; differentiable registration; deep learning; nnU-Net; affine transformation.

Введение. Современные методы глубокого обучения доказали свою высокую эффективность в задачах обработки изображений, включая область медицинской визуализации. Компьютерная томография (КТ) служит одним из основных источников данных благодаря высокому пространственному разрешению. Нейросетевые модели, особенно для семантической сегментации и классификации патологий на КТ-снимках, активно внедряются в исследовательскую и клиническую практику. Основная проблема при автоматической сегментации КТ-изображений заключается в низком контрасте целевых объектов на фоне окружающих тканей и высокой вариативности их формы и размеров.

Для семантической сегментации медицинских изображений широко применяются свёрточные нейронные сети, в частности архитектуры семейства U-Net. Самоконфигурирующий фреймворк nnU-Net [1], построенный на основе U-Net, установил новый стандарт в области автоматической сегментации за счёт автоматизации конвейера обработки: он анализирует свойства входных данных (создает «отпечаток датасета»), определяет оптимальные параметры предобработки, выбирает архитектуру сети (2D или 3D конфигурацию), настраивает гиперпараметры обучения без ручного вмешательства, обеспечивает воспроизводимость и высокую точность в широком спектре задач сегментации [2].

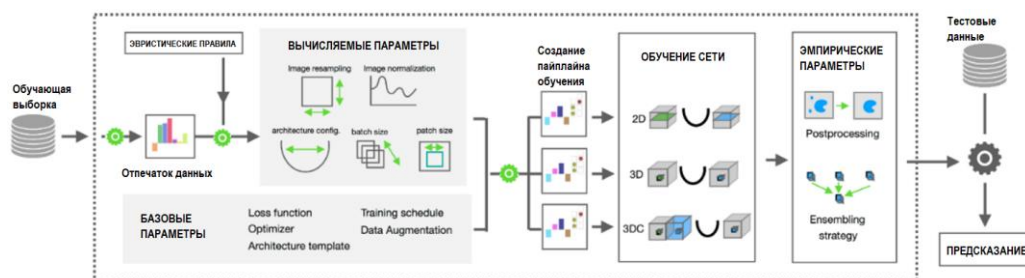


Рис. 1. Архитектура nnU-Net [3]

В научной литературе представлен обширный спектр модификаций U-Net, ориентированных на повышение точности и эффективности. Например, в работе [4] для повышения чёткости изображений перед сегментацией используется импульсно-связанная нейронная сеть (PCNN) в сочетании с U-Net. Архитектура UCR-Net [5] расширяет базовую U-Net структуру за счёт механизмов внимания, позволяя эффективнее захватывать многомасштабную контекстную информацию и высокоуровневые семантические признаки. В исследованиях [6, 7] применяются гибридные архитектуры, такие как Res-UNet и iRes-UNet, направленные на сохранение деталей в изображениях низкого качества и уточнение границ объектов. В работе [8] использован ансамблевый подход, когда для каждой фазы контрастирования обучается отдельная модель Res-UNet.

Для комплексного анализа динамических процессов, таких как распространение контрастного вещества, применяется мультифазная визуализация, подразумевающая получение серии снимков (фаз) в разные моменты времени. Совместный анализ этих фаз позволяет получать более полную информацию об объекте исследования, выявляя характеристики, недоступные при анализе отдельной фазы, что дает возможность для существенно более точной дифференциации патологий на КТ-снимках [9]. Однако фундаментальной проблемой при работе с такими данными является пространственное несоответствие между отдельными фазами, возникающее из-за движения объекта в промежутках между сканированиями. Решением этой проблемы является регистрация изображений – процесс пространственного преобразования, приводящий набор изображений к единой системе координат, что существенно повышает точность последующего анализа [10].

Стандартные методы аффинной регистрации, основанные на линейных преобразованиях, хотя и широко распространены, обладают рядом ограничений. Их эффективность часто снижается при значительных контрастных различиях между сравниваемыми изображениями. Кроме того, эти методы, как правило, не позволяют напрямую интегрировать этап регистрации в сквозной конвейер обучения нейронной сети для решения задачи сегментации. Гиперпараметры таких методов требуют ручного подбора, а итоговое качество сильно зависит от начальной инициализации, что снижает воспроизводимость и затрудняет масштабирование решения.

Современная научная литература предлагает обширный спектр подходов к решению задачи регистрации мультифазных изображений. Традиционные методы базируются на итеративной оптимизации параметров: в работе [11] оптимизируются параметр сдвига глобального аффинного преобразования для каждой пары объёмов, где одна из фаз выбирается в качестве эталонной. Альтернативный подход, основанный на глубоком обучении, использует модифицированные архитектуры нежесткой регистрации: VoxelMorph на основе U-Net формулирует задачу регистрации как функцию, отображающую пару входных изображений в поле деформации, что позволяет достичь скорости, сравнимой с традиционными методами при значительном ускорении вычислений [12, 13]. Применение адаптивного взвешивания неопределённости в функции потерь позволяет динамически балансировать вклад регистрации и сегментации, что обеспечивает улучшение обеих задач одновременно. Другие работы [14] применяют механизмы внимания, в частности механизмы кросс-модальной корреляции, для моделирования взаимосвязей как внутри одной фазы, так и между разными фазами. В работе [15] отказываются от явного этапа регистрации и вместо этого применяют модифицированные архитектуры сегментации: в работе используется SegNet, дополненная модулем пространственного пирамидального разложения (ASPP), который применяет параллельные диагональные свёртки с разными коэффициентами разрежения для эффективной захвата многомасштабного контекста и правильной интеграции информации из различных фаз без явного совмещения.

Однако существующие подходы по-прежнему недостаточно устойчивы к значительным вариациям интенсивности между фазами или требуют трудоемкой ручной настройки параметров. Кроме того, они не предназначены для интеграции в дифференцируемые вычислительные графы нейросетевых моделей, что делает регистрацию отдельным этапом конвейера, требующим дополнительных вычислительных ресурсов. Это ограничивает их применимость в современных ИИ-системах для автоматической обработки КТ-снимков.

Таким образом, актуальность настоящего исследования обусловлена необходимостью разработки эффективных методов обработки мультифазных изображений, где ключевой проблемой является компенсация пространственных рассогласований.

Предлагаемый в данном исследовании метод обучаемых линейных аффинных преобразований является целесообразным выбором для совмещения мультифазных КТ-изображений. Линейные преобразования эффективно устраняют глобальные смещения между фазами, вызванные движением и внешними возмущениями. В отличие от нелинейной регистрации, данный метод гарантирует топологическую согласованность структур, исключая риск нефизиологических искажений. Кроме того, он обеспечивает вычислительную эффективность, стабильность при вычислениях на GPU и позволяет создать быстрый и надёжный этап предварительной обработки.

Целью данного исследования является разработка метода дифференцируемого совмещения мультифазных КТ-изображений на основе обучаемых линейных аффинных преобразований для повышения точности последующей семантической сегментации с использованием фреймворка nnU-Net.

Для достижения поставленной цели решаются следующие задачи.

1. Реализация дифференцируемого модуля аффинной регистрации на основе фреймворка PyTorch, параметры которого оптимизируются с помощью градиентного спуска.
2. Интеграция данного модуля в единый каскадный конвейер обучения модели сегментации, включающий этап предварительной сегментации области интереса, фазового совмещения и финальной сегментации целевых объектов.
3. Проведение сравнительного анализа эффективности разработанного метода с традиционными подходами: прямая мультифазная сегментация без регистрации и сегментация с предварительной классической аффинной регистрацией с помощью инструментария Elastix/ITK.

Важным аспектом работы является создание целостного и адаптивного конвейера обработки данных, оптимизированного для выполнения на графических процессорах (GPU), что соответствует современным требованиям к вычислительным системам в области машинного обучения.

Описание решаемой задачи. В исследовании использован открытый датасет WAW-TACE [16], содержащий мультифазные трехмерные КТ-изображения брюшной полости от 233 пациентов с гепатоцеллюлярной карциномой в формате NIfTI. Выборка включает 185 мужчин и 48 женщин с медианой возраста 66 лет. Датасет охватывает четыре контрастные фазы: нативную (фаза 0, 200 исследований), артериальную (фаза 1, 230 исследований), портально-венозную (фаза 2, 231 исследование) и отсроченную (фаза 3, 193 исследования). Каждое КТ-исследование содержит переменное число аксиальных срезов в зависимости от анатомических особенностей пациента и использованного оборудования, при этом все срезы имеют пространственное разрешение $1,6 \times 1,6$ мм в плоскости, с толщиной среза, варьирующейся в диапазоне 1–3 мм. Общее количество аксиальных 2D срезов в датасете – 175 222, общее количество КТ-сканов – 854.

Трёхмерная природа данных, то есть, в среднем, $512 \times 512 \times 200 = 52\,428\,800$ вокселей (трехмерных пикселей) на одно исследование приводит к высокой вычислительной сложности их обработки, особенно при работе с мультифазными объёмами, где необходимо одновременно хранить и обрабатывать четыре полных 3D-сканирования на каждого пациента.

Датасет WAW-TACE включает 377 ручных сегментаций опухолей, размеченных опытными радиологами, и автоматические сегментации органов брюшной полости, полученные моделью TotalSegmentator [17]. Сегменты опухолей доступны только для одной из фаз (1, 2 или 3) в формате NRRD (Nearly Raw Raster Data) и требуют предварительной консолидации. Для каждого КТ-исследования имеется соответствующая сегментация печени для всех четырёх фаз.

Фундаментальной проблемой при работе с мультифазными данными является пространственное несовпадение анатомических структур между фазами. Несмотря на то, что все фазы получены для одного пациента в одном сеансе сканирования, дыхательные движения, физиологические смещения органов и движения пациента приводят к тому, что одна и та же анатомическая структура занимает разные пространственные координаты на разных фазах. Задача регистрации заключается в нахождении пространственного преобразования (сдвига, поворота, масштабирования и прочих деформаций), которое приводит все фазы в единую систему координат, обеспечивая повоксельное соответствие анатомических структур. Это выравнивание позволяет нейросетевым моделям сегментации корректно интегрировать информацию из разных фаз и выявлять диагностически значимые признаки, недоступные при анализе отдельной фазы.

Все фазы исследования в датасете WAW-TACE имеют различные пространственные размерности, что усложняет обучение нейросетевых моделей и требует предварительной обработки. После фильтрации данных (исключены пациенты с менее чем тремя

фазами) получена рабочая когорта из 215 пациентов (645 КТ-снимков): снимки 172 (80%) пациентов отнесены к обучающей выборке, 43 (20%) – к тестовой выборке со стратификацией по клиническим параметрам. Пример визуализации 2D-срезов всех фаз представлен на рис. 2.

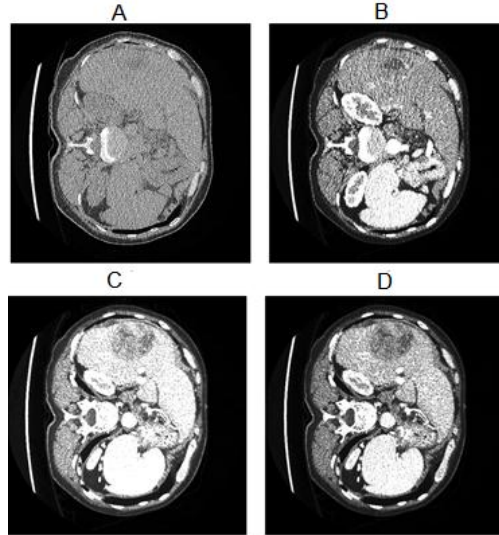


Рис. 2. Пример визуализации фаз контрастирования в датасете WAW-TACE: A – фаза 0; B – фаза 1; C – фаза 2; D – фаза 3

Методика исследования

1. Мультифазная сегментация без совмещения фаз (базовый подход)

В качестве базового уровня для последующего сравнения был реализован подход прямой мультимодальной сегментации с использованием стандартных средств фреймворка nnU-Net [1]. Данный эксперимент имел целью оценить эффективность сегментации без компенсации межфазовых пространственных рассогласований. В рамках данного подхода мультифазные КТ-объемы для каждого пациента напрямую конкатенировались по фазам, формируя многоканальный входной тензор. Архитектурно nnU-Net способна обрабатывать несколько каналов, однако она оперирует в предположении их пространственной согласованности. Ключевым ограничением этого подхода является то, что из-за отсутствия явного выравнивания один и тот же анатомический объект может занимать различные координаты в разных фазах. Поскольку операция свертки не обладает инвариантностью к таким смещениям, модель теряет способность к корректной агрегации мультифазной информации, что на практике часто приводит к игнорированию части входных каналов или существенным ошибкам в локализации объектов.

Для проведения эксперимента данные были подготовлены в соответствии с требованиями фреймворка. Модальности определялись как артериальная фаза (фаза 1), портально-венозная фаза (фаза 2) и дополнительная фаза (other), в качестве которой выбиралась нативная (фаза 0) или отсроченная (фаза 3), при наличии обеих предпочтение отдавалось фазе 3. При этом для каждого пациента явно определялась опорная фаза – та фаза, для которой в датасете присутствовала ручная сегментация опухоли, служащая эталоном для обучения и оценки. Целевые маски формировались путем объединения ручных сегментаций опухолей и автоматических масок печени, полученных моделью TotalSegmentator, с присвоением меток классов: 0 (фон), 1 (печень), 2 (опухоль). Для обеспечения идентичной размерности всех фаз для одного пациента выполнялось приведение неопорных фаз к размеру опорной фазы путем дополнения нулями. Далее был запущен стандартный для nnU-Net конвейер предобработки, включающий верификацию данных и автоматическое планирование эксперимента с помощью планировщика nnUNetPlannerResEncM [18] в конфигурации 3d_fullres, который учитывает остаточные

связи в энкодере. Обучение модели проводилось в течение 1000 эпох с применением стандартного оптимизатора и стратегии кросс-валидации по всем фолдам. Для контроля процесса и оценки качества использовался коэффициент Дайса, вычисляемый как

$$\text{Dice}(Y, \tilde{Y}) = \frac{2|Y \cap \tilde{Y}|}{|Y| + |\tilde{Y}|},$$

где Y – сегментация основной фазы, $\tilde{Y}$ – обработанная преобразованная сегментация другой фазы.

На рис. 3 отображена динамика функции потерь на обучающей (синим) и тестовой (красным) выборках и метрики качества сегментации Dice (зеленым) по эпохам. График показывают стабильную сходимость процесса обучения.

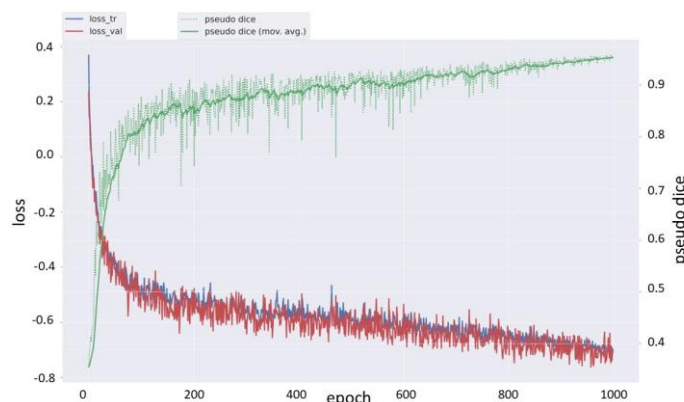


Рис. 3. Процесс обучения мультифазной сегментации без совмещения фаз

Итоговая инференс-оценка на тестовой выборке выполнялась командой `nnUNetv2_predict` с сохранением вероятностных карт.

2. Мультифазная сегментация с использованием аффинных преобразований (Elastix)

Для устранения пространственных несоответствий, выявленных при базовом подходе, во втором эксперименте был применён этап предварительной регистрации мультифазных изображений. Целью данного подхода была оценка эффективности классического метода аффинного совмещения для улучшения качества последующей сегментации. Для каждого пациента, аналогично предыдущей методике, три фазы КТ-изображений были разделены на опорную (фазу с доступной ручной разметкой опухоли) и две смещаемые (регистрируемые) относительно неё. Процесс совмещения выполнялся с использованием модуля `elastix` библиотеки ИТК [19]. Для повышения скорости сходимости и устойчивости процесса оптимизации параметров аффинного преобразования была реализована трёхуровневая стратегия на основе пирамиды изображений. На первом (грубом) уровне разрешение изображений уменьшалось в 4 раза по осям X и Y и в 2 раза по оси Z. На втором (среднем) уровне использовалось уменьшение в 2 раза по осям X и Y при полном разрешении по оси Z. Финальная, наиболее точная оптимизация параметров проводилась на третьем уровне с исходным разрешением данных.

В качестве метрики для оценки качества совмещения использовалась расширенная взаимная информация Маттеса с гистограммой из 128 бинов. Эта метрика, в отличие от методов, основанных на прямом сравнении интенсивности пикселей изображений, оценивает не прямое совпадение, а статистическую зависимость между распределениями яркости двух изображений. Она вычисляется на основе совместной гистограммы $p(i, j)$ интенсивностей опорного R и смещаемого F изображений:

$$I(R, F) = \sum_{i=1}^N \sum_{j=1}^M p(i, j) \log_2 \left(\frac{p(i, j)}{p_R(i) \cdot p_F(j)} \right),$$

где $p_R(i)$ и $p_F(j)$ – маргинальные (частные) распределения интенсивностей каждого изображения. Процесс оптимизации направлен на максимизацию этой взаимной информации, что соответствует поиску такого пространственного преобразования, при котором статистическая связь между изображениями становится наибольшей. Это делает метод устойчивым к нелинейным изменениям контраста между различными фазами КТ. Для эффективного вычисления метрики применялось адаптивное стохастическое семплирование: количество пространственных сэмплов увеличивалось с 1024 на первом уровне применяемой трёхуровневой стратегии до 4096 на последнем, при этом набор сэмплов обновлялся на каждой итерации. Поиск оптимальных параметров преобразования осуществлялся оптимизатором с ограничением в 1000 итераций на уровень.

В процессе регистрации были задействованы следующие механизмы: автоматическая инициализация начального приближения трансформации, автоматическая оценка масштабов параметров, многопоточные вычисления с использованием всех доступных ядер ЦПУ и распараллеливание вычисления метрики. Для применения найденного преобразования и получения зарегистрированного изображения использовалась кубическая В-сплайн интерполяция 3-го порядка. В отличие от линейной интерполяции, В-сплайн обеспечивает гладкую и дифференцируемую аппроксимацию интенсивности, используя взвешенную сумму значений соседних вокселей. Это минимизирует артефакты ступенчатости и размытия, сохраняя плавность анатомических границ, что важно для последующей точной сегментации. Данный набор параметров ранее успешно применялся для задач регистрации изображений брюшной полости [20]. После выполнения регистрации всех изображений обучающей выборки и формирования нового набора пространственно согласованных КТ-срезов была создана соответствующая структура проекта для обучения мультимодальной модели nnU-Net. Выбор модальностей, архитектуры и параметров обучения остался идентичным первому подходу. Процесс обучения завершился успешно. Перед применением обученной модели на тестовых данных для них были выполнены все те же этапы предобработки и регистрации.

Данный эксперимент позволил оценить вклад аффинной регистрации, реализованной с использованием модуля *elastix* библиотеки ИТК, как отдельного этапа предобработки. Его результаты служат важным ориентиром для сравнения с предлагаемым в работе методом дифференцируемого совмещения.

3. Мультифазная сегментация с обучаемым фазовым совмещением

Третьей, ключевой стадией исследования стало внедрение этапа дифференцируемого совмещения фаз, основанного на оптимизации градиентным спуском. В отличие от предыдущих подходов, где регистрация либо отсутствовала, либо выполнялась как отдельный шаг предобработки, данный метод интегрирует процесс пространственного выравнивания непосредственно в конвейер обучения нейронной сети. Это позволяет создавать единый вычислительный граф, в котором параметры преобразования и сегментации оптимизируются согласованно.

Основой предлагаемого метода является реализованный на PyTorch [21] модуль дифференцируемой аффинной регистрации. Его архитектура состоит из двух взаимосвязанных компонентов. Первый – это пользовательский класс, реализующий само трёхмерное аффинное преобразование. Второй – оптимизируемый блок, который управляет параметрами этого преобразования.

Для описания работы класса введем следующие обозначения.

1) Центр масс (центроид) сегментации. Для бинарной 3D маски печени V (где $V(x,y,z)=1$ для вокселей, принадлежащих органу, и 0 в противном случае) центр масс вычисляется как:

$$C = (C_x, C_y, C_z) = \frac{1}{N} \sum_{(x,y,z) \in \Omega} (x, y, z),$$

где $\Omega = \{(x, y, z) \in R^3 \mid V(x, y, z) = 1\}$ – множество вокселей маски, $N = |\Omega|$ – их общее количество.

2) Вектор переноса инициализируется как разность центров масс:

$$\Delta C = C_m - C_b,$$

где C_m – центр масс 3D сегментации фазы, которую необходимо сместить, C_b – центр масс 3D сегментации опорной фазы.

Вектор ΔC показывает смещение объекта в регистрируемой фазе относительно опорной по трём осям.

3) Матрица вращения R в 3D может быть представлен как композиция поворотов вокруг осей x, y, z на углы θ (углы Эйлера):

$$R = R_x(\theta_x) \times R_y(\theta_y) \times R_z(\theta_z),$$

где

$$R_x(\theta_x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_x) & -\sin(\theta_x) \\ 0 & \sin(\theta_x) & \cos(\theta_x) \end{bmatrix} \text{ – поворот вокруг оси } x,$$

$$R_y(\theta_y) = \begin{bmatrix} \cos(\theta_y) & 0 & \sin(\theta_y) \\ 0 & 1 & 0 \\ -\sin(\theta_y) & 0 & \cos(\theta_y) \end{bmatrix} \text{ – поворот вокруг оси } y,$$

$$R_z(\theta_z) = \begin{bmatrix} \cos(\theta_z) & -\sin(\theta_z) & 0 \\ \sin(\theta_z) & \cos(\theta_z) & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ – поворот вокруг оси } z,$$

Углы $\theta_x, \theta_y, \theta_z$ (в радианах) инициализируется нулевыми значениями и оптимизируются в процессе обучения.

4) Диагональная матрица масштабирования S задаёт коэффициенты изменения расстояний вдоль осей:

$$S = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & s_z \end{bmatrix},$$

где s_x, s_y, s_z – коэффициенты масштабирования вдоль соответствующих осей, инициализируется единицами и оптимизируются в процессе обучения.

5) Итоговая аффинная матрица преобразования A размером 3×4 объединяет вращение, масштабирование и перенос для преобразования однородных координат:

$$A = \begin{bmatrix} R \times S & t_x \\ & t_y \\ & t_z \end{bmatrix},$$

где $t = (t_x, t_y, t_z)$ – вектор переноса, являющийся обучаемым параметром, который инициализируется значением ΔC .

Функция `affine_grid` генерирует координатную сетку на основе матрицы AA , преобразуя координаты каждого вокселя согласно аффинному преобразованию. Затем функция `grid_sample` применяет эту сетку к исходному 3D-изображению, выполняя интерполяцию значений вокселей в новых координатах.

Ядром предлагаемого метода является модуль дифференцируемой аффинной регистрации. В отличие от классических подходов, параметры преобразования (вектор переноса $t = (t_x, t_y, t_z)$, углы Эйлера $\theta = (\theta_x, \theta_y, \theta_z)$, и коэффициенты масштабирования (s_x, s_y, s_z)) не вычисляются итеративно на основе метрики сходства изображений, а являются обучаемыми. Эти 9 параметров инициализируются разумными начальными значениями: вектор переноса t – разностью центров масс масок печени регистрируемой и опорной фаз, углы поворота – нулями, коэффициенты масштабирования – единицами.

Обучение модуля происходит на бинарных масках печени, что является ключевой особенностью предлагаемого подхода. В качестве функции потерь используется негативный коэффициент Dice:

$$L_{\text{reg}} = -\text{Dice}(S_{\text{mov}}, S_{\text{ref}}),$$

где S_{mov} – маска, преобразованная текущими параметрами, а S_{ref} – маска опорной фазы. Минимизация этой функции напрямую максимизирует пространственное совпадение анатомических структур. Благодаря полной дифференцируемости конвейера, градиенты от функции потерь распространяются через операцию интерполяции (`grid_sample`) и вычисления аффинной сетки (`affine_grid`) непосредственно к 9 параметрам преобразования, которые обновляются оптимизатором Adam [22].

После завершения обучения и достижения сходимости, оптимизированные параметры фиксируются. Итоговая аффинная матрица A применяется не к маскам, а к исходным полным 3D КТ-сканам соответствующей фазы. Это обеспечивает анатомически корректное выравнивание томографических данных в единое пространство опорной фазы, при этом само преобразование обучено на более простой и надёжной задаче совмещения бинарных масок.

Итоговый алгоритм сегментации с использованием обучаемого фазового совмещения имеет следующий вид.

Этап 1. Сегментация печени. На данном этапе была обучена модель nnU-Net в конфигурации `3d_lowres` для получения трёхмерных бинарных масок печени, необходимых для последующей регистрации. Для оптимизации использования памяти применялся планировщик `overwrite_target_spacing` с параметрами (2.0, 2.0, 3.0). Обученная модель продемонстрировала высокую точность сегментации печени: F1-мера составила 96.3%, а метрика Intersection over Union (IoU) – 91.2%.

Этап 2. Подготовка данных для регистрации. Для каждого пациента определялась опорная фаза (с доступной разметкой опухоли), вычислялся вектор смещения центров масс (ДС) между фазами, а все трёхмерные объёмы приводились к единой пространственной размерности. Для обучающей выборки использовались фактические маски печени, а для тестовой – маски, предсказанные моделью на Этапе 1.

Этап 3. Обучение и применение дифференцируемой регистрации. Для каждой пары «опорная-регистрируемая фаза» независимо обучался модуль дифференцируемой аффинной регистрации. После оптимизации полученное преобразование применялось к исходным трёхмерным КТ-сканам. Для последующей сегментации формировался многоканальный трёхмерный вход, где каждому каналу соответствовала определённая фаза: артериальная, портально-венозная и отсроченная фаза (если она отсутствовала, то нативная).

Этап 4. Обрезка данных. Все трёхмерные КТ-снимки и соответствующие маски были обрезаны по трёхмерной ограничивающей области маски печени. Эта операция позволила исключить из анализа значительный объём фоновых вокселей, что снизило дисбаланс классов и повысило вычислительную эффективность на следующих этапах.

Этап 5. Сегментация патологий. На подготовленных выровненных и обрезанных трёхмерных данных была обучена финальная модель nnU-Net для задачи бинарной сегментации «фон – патология». Динамика процесса обучения, представленная на рис. 4, подтвердила стабильную сходимость и достижение высокого качества модели

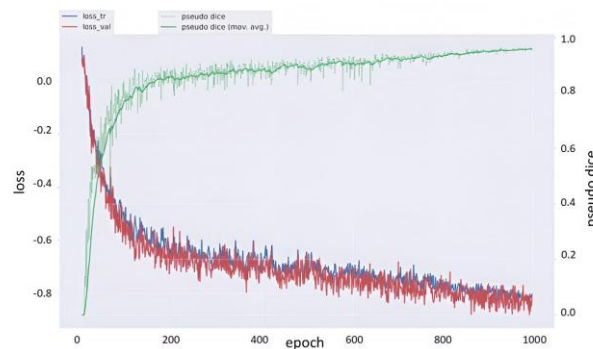


Рис. 4. Процесс обучения мультифазной сегментации с использованием совмещения фаз

Монотонный рост метрики Dice (зеленая кривая) из рис. 3 и последующая стабилизация на высоком значении (приближающемся к 1.0) количественно подтверждают, что модель не только минимизирует целевую функцию, но и достигает высокой практической точности в задаче сегментации, что является конечной целью обучения.

Вычислительные эксперименты проводились с использованием графического процессора NVIDIA GeForce RTX 3090 Ti с 24 ГБ видеопамати GDDR6X, что обеспечило значительное ускорение матричных операций и обработки трехмерных медицинских изображений. Поскольку реализуемый модуль совмещения фаз работал в пределах от 3 до 7 ГБ GPU, то для более эффективной обработки и преобразования данных были использованы параллельные вычисления с использованием multiprocessing.Pool с явным указанием метода инициализации spawn, обеспечивающего стабильную работу с CUDA для этапов 2-4.

Каждый процесс инициализируется с ограничением доли видеопамати GPU через torch.cuda.set_per_process_memory_fraction для предотвращения конфликтов ресурсов. Обработка отдельных КТ-снимков инкапсулирована в независимые задачи, выполняемые асинхронно через pool.imap в 4 параллельных потока, с последующей агрегацией результатов. Полный конвейер решения представлен на рис. 5.

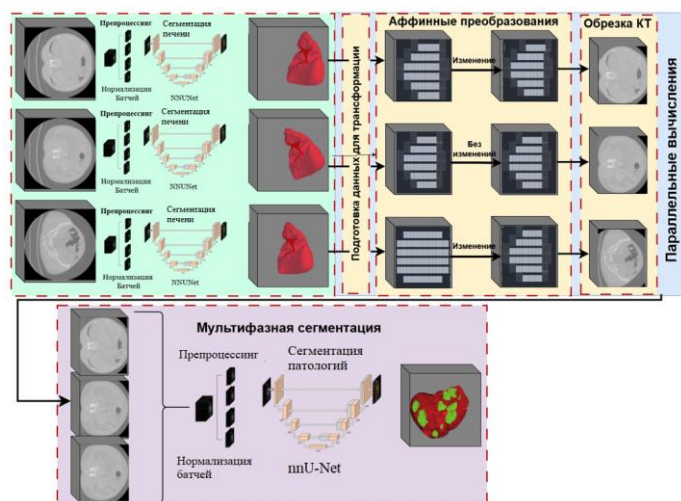


Рис. 5. Нейросетевой конвейер мультифазной сегментации с фазовым совмещением

Результаты и их обсуждение. Результаты первого эксперимента, в котором мультимодальный подход nnU-Net применялся без предварительного совмещения фаз, продемонстрировали его низкую точность и неэффективность для клинического применения. Это обосновало необходимость проведения последующих экспериментов с этапом регистрации: с использованием классических аффинных преобразований ИТК.Elastix и предложенного метода дифференцируемой регистрации на основе градиентного спуска.

Для комплексной оценки эффективности подходов, наряду со стандартными метриками, были введены две целевые пользовательские метрики, формализованные математически. Средняя вероятность в области патологии (mean probe) количественно оценивает уверенность модели внутри истинных патологических областей. Она вычисляется как среднее значение вероятностной карты предсказаний P только для тех вокселей, которые принадлежат истинной маске патологии M .

Формально, если $\Omega = \{r_i = (x, y, z) | M(r_i) = 1\}$ – множество вокселей патологии, $a = |\Omega|$ – их общее количество, то:

$$\text{mean probe} = \frac{1}{N} \sum_{r_i \in \Omega} P(r_i),$$

где $P(r_i) \in [0,1]$ –предсказанная моделью вероятность, что воксель r_i принадлежит к классу патологии. Высокое значение этой метрики указывает на то, что модель не только правильно локализует патологию на КТ-снимке, но и делает это с высокой уверенностью в её границах, что крайне важно для доверия к её результатам и их интерпретируемости.

Второй метрикой является доля случаев с совпадением количества патологических образований (cnt match). Для её вычисления количество патологий, предсказанных моделью на одном КТ-снимке (N_{pred}), сравнивается с их фактическим количеством (N_{fact}). Предсказанное количество определяется как число связанных компонент в предсказанной бинарной маске. Фактическое количество (N_{fact}) вычисляется аналогично из эталонной маски M . Итоговый бинарный индикатор для одного исследования определяется как:

$$\text{cnt match} = \begin{cases} 1, & \text{если } N_{pred} = N_{fact}, \\ 0, & \text{иначе} \end{cases}$$

В сводной таблице результатов (табл. 1) приводится усреднённое по всей тестовой выборке значение этой метрики, то есть доля исследований, в которых количество патологий было определено абсолютно точно. Эта метрика имеет непосредственное клиническое значение, так как корректное определение мультифокальности заболевания (количества отдельных опухолевых узлов) крайне важно для планирования дальнейшего лечения. Используемый в исследовании датасет WAW-TACE содержит ряд примеров с множественными очагами (до 10-15 в одном КТ-скане), поэтому важно, чтобы модель находила большую часть из них.

Таким образом, оценка подходов проводилась по набору базовых метрик, включающему коэффициент Дайса (Dice), точность (Precision), чувствительность, т.е. полноту (Sensitivity), долю правильных ответов (Accuracy), площадь под ROC-кривой (ROC-AUC), а также описанные выше пользовательские метрики mean probe и cnt match.

Основная метрика сегментации, коэффициент Дайса, показала значительный прирост при использовании алгоритма обучаемой регистрации: на 43% между первым и третьим подходами и на 13% между вторым и третьим. Это указывает на существенное превосходство метода с обучаемым совмещением. Высокие значения Accuracy, наблюдаемые во всех экспериментах, не являются показательными из-за сильного дисбаланса классов (незначительный объем опухоли относительно всего КТ-скана). Значимое улучшение Precision (на 48% относительно базового подхода) свидетельствует об эффективном подавлении ложноположительных срабатываний благодаря выравниванию фаз. Рост Sensitivity (на 36%) и ROC-AUC подтверждает улучшение в обнаружении истинных патологий. Пользовательские метрики также демонстрируют преимущество третьего подхода: средняя уверенность модели в области опухоли (mean probe) повысилась, а доля случаев с точным определением количества патологических образований (cnt match) увеличилась на 19.6% относительно метода Elastix.

Таблица 1

Сравнение разработанных подходов

Метрика	Подход 1, без регистрации	Подход 2, с регистрацией <i>Elastix</i>	Подход 3, разработанный
Dice	0,575	0,726	0,823
Accuracy	0,998	0,991	0,997
Precision	0,583	0,677	0,863
Sensitivity	0,600	0,705	0,817
ROC-AUC	0,799	0,850	0,908
Mean probe	0,598	0,755	0,811
Cnt match	0,56	0,42	0,67

Качественный анализ (рис. 6) подтверждает количественные результаты. Метод с обучаемым совмещением (В) корректно идентифицирует количество патологий, хотя и с небольшой погрешностью в размерах одной из них. Базовый подход без регистрации (с) допускает более серьезную ошибку в размерах и генерирует ложноположительное срабатывание.

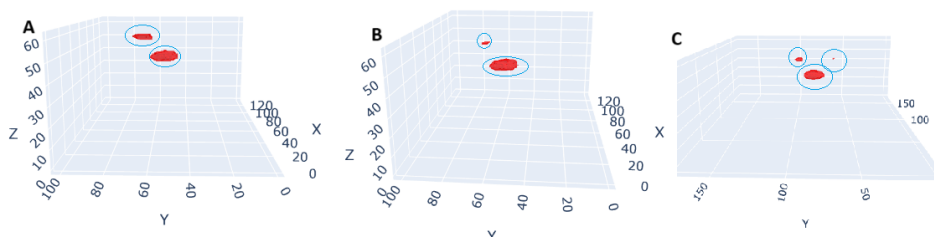


Рис. 6. Визуализация сегментов опухоли на конкретном пациенте:
а – фактическая; б – с совмещением подхода 3; в – без совмещения

Таким образом, метод мультифазной сегментации с обучаемым фазовым совмещением демонстрирует высокое качество, обеспечивая лучший баланс между точностью обнаружения и минимизацией ложных срабатываний. Его результаты приближаются к уровню, необходимому для клинического применения. К ограничениям подхода относятся: зависимость от качества предварительной сегментации печени, неспособность учесть нелинейные (нежесткие) деформации тканей из-за линейности аффинной модели, а также возросшие вычислительные затраты на обучение. Классический метод на основе Elastix также имеет недостатки, включая необходимость ручной настройки гиперпараметров, высокую ресурсоемкость и потенциальное размытие мелких структур при интерполяции, что в целом затрудняет его стабильное и масштабируемое применение в автоматизированных конвейерах.

Заключение. В работе была успешно решена задача повышения точности автоматической сегментации патологий на мультифазных КТ-изображениях печени. Ключевым результатом является разработка и валидация метода дифференцируемого фазового совмещения, основанного на обучаемых аффинных преобразованиях, которые интегрируются в единый конвейер глубокого обучения.

Проведённые эксперименты позволили сделать следующие основные выводы.

1. Прямая мультимодальная сегментация без учёта пространственных рассогласований между фазами показала свою принципиальную неэффективность, подтвердив необходимость этапа явной регистрации.

2. Предложенный метод дифференцируемого совмещения продемонстрировал значимое превосходство над классическим подходом на основе ITK.Elastix. Улучшение ключевых метрик (Dice и ROC-AUC) свидетельствует не только о лучшем пространственном соответствии, но и о более высокой дискриминативной способности модели при анализе выровненных данных.

3. Каскадная архитектура, в которой регистрация оптимизируется по анатомическим маскам печени, обеспечивает анатомически корректное выравнивание, устойчивое к вариациям контраста, и устраняет необходимость в ручной настройке гиперпараметров.

Главным практическим вкладом работы является создание готового к интеграции модуля, оптимизированного как по точности, так и по скорости работы. Реализация с поддержкой параллельных вычислений на GPU обеспечивает высокую производительность, что крайне важно для обработки больших объёмов трёхмерных медицинских данных в исследовательских и клинических конвейерах. Таким образом, метод позволяет превратить проблему мультифазной регистрации в эффективно решаемую задачу оптимизации средствами глубокого обучения.

Основными ограничениями метода являются его зависимость от качества первичной сегментации печени и применимость только к глобальным, линейным деформациям. Перспективы дальнейших исследований видятся в расширении предложенного подхода для учёта нелинейных деформаций, интеграции механизмов внимания для работы со сложными патологиями и адаптации архитектуры для задач сегментации других органов.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Isensee, F., Jaeger, P.F., Kohl, S. A., Petersen, J., & Maier-Hein, K.H.* nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation // *Nature methods*. – 2021. – 18 (2). – P. 203-211.
2. *Palazzo S., Zambetta G., Calbi R.* An overview of segmentation techniques for CT and MRI images: Clinical implications and future directions in medical diagnostics // *Medical Imaging Process & Technology*. – 2024.
3. *Куликов А.А., Каширина И.Л., Савкина Е.Ф.* Сегментация объемных образований печени на мультифазных КТ-изображениях с использованием фреймворка nnU-Net // *Моделирование, оптимизация и информационные технологии*. – 2025. – Т. 13, № 1 (48).
4. *Zang L. et al.* Research on liver cancer segmentation method based on PCNN image processing and SE-ResUnet // *Scientific reports*. – 2023. – Vol. 13, No. 1. – P. 12779.
5. *Sun Q. et al.* UCR-Net: U-shaped context residual network for medical image segmentation // *Computers in Biology and Medicine*. – 2022. – Vol. 151. – P. 106203.
6. *Liu S. et al.* Segmenting nailfold capillaries using an improved U-net network // *Microvascular Research*. – 2020. – Vol. 130. – P. 104011.
7. *Ly P. et al.* An improved residual U-Net with morphological-based loss function for automatic liver segmentation in computed tomography // *Math. Biosci. Eng.* – 2022. – Vol. 19, No. 2. – P. 1426-1447.
8. *Черемискин А.В., Каширина И.Л.* Сегментация мультифазных КТ-изображений с использованием ансамбля моделей RESUNET // *Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии*. – 2025. – № 3. – С. 140-152.
9. *Ермоленко С.В., Каширина И.Л., Старичкова Ю.В.* Постобработка результатов сегментации медицинских изображений // *Computational Nanotechnology*. – 2025. – Т. 12, № 2. – С. 109-118.
10. *Islam K.T., Wijewickrema S. & O'Leary S.* A deep learning based framework for the registration of three dimensional multi-modal medical images of the head // *Sci Rep*. – 2021. – 11, 1860. – <https://doi.org/10.1038/s41598-021-81044-7>.
11. *Uhm K.H. et al.* Deep learning for end-to-end kidney cancer diagnosis on multi-phase abdominal computed tomography // *NPJ precision oncology*. – 2021. – Vol. 5, No. 1. – P. 54.
12. *Pérez de Frutos J. et al.* Learning deep abdominal CT registration through adaptive loss weighting and synthetic data generation // *Plos one*. – 2023. – Vol. 18, No. 2. – P. e0282110.
13. *Balakrishnan G. et al.* Voxelmorph: a learning framework for deformable medical image registration // *IEEE transactions on medical imaging*. – 2019. – Vol. 38, No. 8. – P. 1788-1800.
14. *Wu L. et al.* Beyond radiologist-level liver lesion detection on multi-phase contrast-enhanced CT images by deep learning // *Iscience*. – 2023. – Vol. 26, No. 11.
15. *Nayantara P. V. et al.* Automatic liver segmentation from multiphase CT using modified SegNet and ASPP module // *SN Computer Science*. – 2024. – Vol. 5, No. 4. – P. 377.
16. *Bartnik K. et al.* WAW-TACE: a hepatocellular carcinoma multiphase CT dataset with segmentations, radiomics features, and clinical data // *Radiology: Artificial Intelligence*. – 2024. – Vol. 6, No. 6. – P. e240296.
17. *Wasserthal J. et al.* TotalSegmentator: robust segmentation of 104 anatomic structures in CT images // *Radiology: Artificial Intelligence*. – 2023. – Vol. 5, No. 5. – P. e230024.
18. *Isensee F. et al.* nnu-net revisited: A call for rigorous validation in 3d medical image segmentation // *International Conference on Medical Image Computing and Computer-Assisted Intervention*. – Cham: Springer Nature Switzerland, 2024. – P. 488-498.
19. *Klein S. et al.* Elastix: a toolbox for intensity-based medical image registration // *IEEE transactions on medical imaging*. – 2009. – Vol. 29, No. 1. – P. 196-205.
20. *Boussot V. et al.* IMPACT: A Generic Semantic Loss for Multimodal Medical Image Registration // *arXiv preprint arXiv:2503.24121*. – 2025.
21. *Paszke A. et al.* Pytorch: An imperative style, high-performance deep learning library // *Advances in neural information processing systems*. – 2019. – Vol. 32.
22. *Li Y., Zhang X.* Adaptive moment estimation optimization algorithm using projection gradient for deep learning // *Fifth International Conference on Applied Mathematics, Modelling, and Intelligent Computing (CAMMIC 2025)*. – SPIE, 2025. – Vol. 13644. – P. 501-511.

REFERENCES

1. *Isensee, F., Jaeger, P.F., Kohl, S. A., Petersen, J., & Maier-Hein, K.H.* nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation, *Nature methods*, 2021, 18 (2), pp. 203-211.
2. *Palazzo S., Zambetta G., Calbi R.* An overview of segmentation techniques for CT and MRI images: Clinical implications and future directions in medical diagnostics, *Medical Imaging Process & Technology*, 2024.
3. *Kulikov A.A., Kashirina I.L., Savkina E.F.* Segmentatsiya ob"emnykh obrazovaniy pecheni na mul'tifaznykh KT-izobrazheniyakh s ispol'zovaniem freymvorka nnU-Net [Segmentation of volumetric liver lesions on multiphase CT images using the nnU-Net framework], *Modelirovanie, optimizatsiya i informatsionnye tekhnologii* [Modeling, Optimization and Information Technology], 2025, Vol. 13, No. 1 (48).

4. Zang L. et al. Research on liver cancer segmentation method based on PCNN image processing and SE-ResUnet, *Scientific Reports*, 2023, Vol. 13, No. 1, pp. 12779.
5. Sun Q. et al. UCR-Net: U-shaped context residual network for medical image segmentation, *Computers in Biology and Medicine*, 2022, Vol. 151, pp. 106203.
6. Liu S. et al. Segmenting nailfold capillaries using an improved U-net network, *Microvascular Research*, 2020, Vol. 130, pp. 104011.
7. Lv P. et al. An improved residual U-Net with morphological-based loss function for automatic liver segmentation in computed tomography, *Math. Biosci. Eng.*, 2022, Vol. 19, No. 2, pp. 1426-1447.
8. Cheremiskin A.V., Kashirina I.L. Segmentatsiya mul'tifaznykh KT-izobrazheniy s ispol'zovaniem ansamblya modeley RESUNET [Segmentation of multiphase CT images using the RESUNET model ensemble], *Vestnik Voronezhskogo gosudarstvennogo universiteta. Seriya: Sistemnyy analiz i informatsionnye tekhnologii* [Bulletin of the Voronezh State University. Series: Systems Analysis and Information Technology], 2025, No. 3, pp. 140-152.
9. Ermolenko S.V., Kashirina I.L., Starichkova Yu.V. Postobrabotka rezul'tatov segmentatsii meditsinskikh izobrazheniy [Post-processing of medical image segmentation results], *Computational Nanotechnology*, 2025, Vol. 12, No. 2, pp. 109-118.
10. Islam K.T., Wijewickrema S. & O'Leary S. A deep learning based framework for the registration of three dimensional multi-modal medical images of the head, *Sci Rep.*, 2021, 11, 1860. Available at: <https://doi.org/10.1038/s41598-021-81044-7>.
11. Uhm K.H. et al. Deep learning for end-to-end kidney cancer diagnosis on multi-phase abdominal computed tomography, *NPJ precision oncology*, 2021, Vol. 5, No. 1, pp. 54.
12. Pérez de Frutos J. et al. Learning deep abdominal CT registration through adaptive loss weighting and synthetic data generation, *Plos one*, 2023, Vol. 18, No. 2, pp. e0282110.
13. Balakrishnan G. et al. Voxelmorph: a learning framework for deformable medical image registration, *IEEE transactions on medical imaging*, 2019, Vol. 38, No. 8, pp. 1788-1800.
14. Wu L. et al. Beyond radiologist-level liver lesion detection on multi-phase contrast-enhanced CT images by deep learning, *Iscience*, 2023, Vol. 26, No. 11.
15. Nayantara P.V. et al. Automatic liver segmentation from multiphase CT using modified SegNet and ASPP module, *SN Computer Science*, 2024, Vol. 5, No. 4, pp. 377.
16. Bartnik K. et al. WAW-TACE: a hepatocellular carcinoma multiphase CT dataset with segmentations, radiomics features, and clinical data, *Radiology: Artificial Intelligence*, 2024, Vol. 6, No. 6, pp. e240296.
17. Wasserthal J. et al. TotalSegmentator: robust segmentation of 104 anatomic structures in CT images, *Radiology: Artificial Intelligence*, 2023, Vol. 5, No. 5, pp. e230024.
18. Isensee F. et al. nnu-net revisited: A call for rigorous validation in 3d medical image segmentation, *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer Nature Switzerland, 2024, pp. 488-498.
19. Klein S. et al. Elastix: a toolbox for intensity-based medical image registration, *IEEE transactions on medical imaging*, 2009, Vol. 29, No. 1, pp. 196-205.
20. Boussoit V. et al. IMPACT: A Generic Semantic Loss for Multimodal Medical Image Registration, *arXiv preprint arXiv:2503.24121*, 2025.
21. Paszke A. et al. Pytorch: An imperative style, high-performance deep learning library, *Advances in neural information processing systems*, 2019, Vol. 32.
22. Li Y., Zhang X. Adaptive moment estimation optimization algorithm using projection gradient for deep learning, *Fifth International Conference on Applied Mathematics, Modelling, and Intelligent Computing (CAMMIC 2025)*. SPIE, 2025, Vol. 13644, pp. 501-511.

Ермоленко Сергей Витальевич – Воронежский государственный университет; e-mail: ermolenko@math.vsu.ru; г. Воронеж, Россия; тел.: +79601087190; аспирант.

Каширина Ирина Леонидовна – МИРЭА – Российский технологический университет; e-mail: kashirina@mirea.ru; г. Москва, Россия; тел.: +79036539293; д.т.н.; профессор кафедры технологий искусственного интеллекта.

Ermolenko Sergey Vitalievich – Voronezh State University; e-mail: ermolenko@math.vsu.ru; Voronezh, Russia; phone: +79601087190; graduate student.

Kashirina Irina Leonidovna – MIREA – Russian Technological University; e-mail: kash.irina@mail.ru; Moscow, Russia; phone: +79036539293; dr. of eng. sc.; professor of the Department of Artificial Intelligence Technologies.