

РУДЫКО М.О.,

ФГБОУ ВО «Курский государственный университет», Курск, Россия;
e-mail: tabeld@yandex.ru. ORCID: 0009-0001-3159-8741

ЯЦЕНКО Е.В.,

ФГБОУ ВО «Курский государственный университет», Курск, Россия;
e-mail: yatsenko1147@gmail.com. ORCID: 0009-0003-7144-0775

МАКАРОВ К.С.,

к.т.н., ФГБОУ ВО «Курский государственный университет», Курск, Россия;
e-mail: makarov_ks@kursksu.ru. ORCID: 0009-0008-8645-2847

МОДИФИЦИРОВАННАЯ МОДЕЛЬ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ БИНАРНОЙ СЕГМЕНТАЦИИ ПОЧЕЧНЫХ ПАТОЛОГИЙ

DOI: 10.25881/18110193_2026_3_90

Аннотация. Разработана модифицированная модель глубокого обучения для бинарной сегментации почечных патологий на компьютерной томографии. Актуальность исследования определяется потребностью в повышении точности и эффективности выявления аномалий на этапе первичной диагностики, особенно в условиях высокой нагрузки на медицинские учреждения.

Цель работы – создание универсальной системы автоматической детекции и локализации патологий почек без предварительной классификации их морфологического типа.

Методы. В основе модели лежит модифицированная архитектура U-Net с пятью уровнями энкодера и декодера. Внесены изменения: замена максимального пулинга на свёртку с шагом два, использование групповой нормализации и регуляризации методом исключений. Обучение проводилось на клиническом наборе данных из более чем 400 мультиспиральных компьютерных томографий с венозным контрастированием. Предварительная обработка включала преобразование интенсивностей в единицы Хаунсфилда, нормализацию без жёсткого ограничения диапазона и аугментацию данных. Для обучения применялась гибридная функция потерь, сочетающая коэффициент Дайса и бинарную кросс-энтропию.

Результаты. На независимой тестовой выборке достигнуто значение средневзвешенного коэффициента Дайса 0,85. Модель демонстрирует высокую точность сегментации при различных типах патологий – кистах (включая сложные по классификации Bosniak), солидных опухолях и смешанных структурах. Сохраняется устойчивость при анатомических аномалиях (подковообразная почка, дистопия) и при малых размерах очагов (менее 5 мм).

Выводы. Предложенная модель обеспечивает высокую универсальность и применимость в рутинной клинической практике. Возможность детекции широкого спектра патологий без предварительной типизации повышает её клиническую значимость и пригодность для скрининга.

Область применения. Результаты могут быть использованы в онкоурологии для первичного анализа компьютерной томографии, подготовки к хирургическому вмешательству и обучения медицинских специалистов. Перспективна интеграция модели в системы хранения и передачи медицинских изображений для автоматической предобработки исследований.

Ключевые слова: Опухоль почек, почечно-клеточный рак, нейронная сеть, искусственный интеллект, машинное обучение.

Для цитирования: Рудыко М.О., Яценко Е.В., Макаров К.С. Модифицированная модель глубокого обучения для бинарной сегментации почечных патологий. Врач и информационные технологии. 2026; 3: 90-103. DOI: 10.25881/18110193_2026_3_90.

RUDYKO M.O.,

Kursk State University, Kursk, Russia;
e-mail: tabeld@yandex.ru. ORCID: 0009-0001-3159-8741

YATSENKO E.V.,

Kursk State University, Kursk, Russia;
e-mail: yatsenko1147@gmail.com. ORCID: 0009-0003-7144-0775

MAKAROV K.S.,

PhD, Kursk State University, Kursk, Russia;
e-mail: makarov_ks@kursksu.ru. ORCID: 0009-0008-8645-2847

A MODIFIED DEEP LEARNING MODEL FOR BINARY SEGMENTATION OF RENAL PATHOLOGIES

DOI: 10.25881/18110193_2026_3_90

Abstract. A modified deep learning model for binary segmentation of renal pathologies in computed tomography (CT) scans has been developed. The relevance of this study stems from the need to improve the accuracy and efficiency of anomaly detection during primary diagnosis, particularly given the high workload faced by medical institutions.

The aim of the study is to create a universal system for the automatic detection and localization of renal pathologies without prior classification of their morphological type.

Methods. The model is based on a modified U-Net architecture featuring five levels in both the encoder and decoder. Key modifications included replacing max pooling with strided convolutions (stride of two), implementing group normalization, and incorporating group normalization and dropout regularization. The model was trained on a clinical dataset comprising over 400 contrast-enhanced multidetector computed tomography (MDCT) scans. Preprocessing included conversion of intensities into Hounsfield units, range-free normalization, and data augmentation. A hybrid loss function combining the Dice coefficient and binary cross-entropy was used during training.

Results. A weighted average Dice coefficient of 0.85 was achieved on an independent test set. The model demonstrates high segmentation accuracy across various types of pathologies, including cysts (including those classified as complex under the Bosniak system), solid tumors, and mixed structures. The model maintains robustness in the presence of anatomical anomalies (such as horseshoe kidney, renal ectopia) when dealing with small lesions (less than 5 mm in size).

Conclusions. The proposed model offers high versatility and applicability in routine clinical practice. Its ability to detect a wide range of renal pathologies without prior classification enhances its clinical significance and suitability for screening purposes.

Application Area. The results can be utilized in oncology for initial CT scan analysis, surgical planning, and training of medical professionals. Integrating the model into medical image archiving and communication systems for the automated preprocessing of imaging studies holds significant promise.

Keywords: Kidney tumor, renal cell carcinoma, neural network, artificial intelligence, machine learning.

For citation: Rudyko M.O., Yatsenko E.V., Makarov K.S. A modified deep learning model for binary segmentation of renal pathologies. Medical doctor and information technology. 2026; 3: 90-103. DOI: 10.25881/18110193_2026_3_90.

ВВЕДЕНИЕ

По данным Всемирной организации здравоохранения, рак почки по количеству заболевших на 2022 г. (434 840 новых случаев) занимает 14-е место в мире среди всех злокачественных образований [1]. В России рак почки занимает 10 место, составляя 3,9% от общего числа пациентов с онкологическими заболеваниями. Динамика показателей заболеваемости населения России злокачественными новообразованиями почек показывает прирост 10,81% в период 2013–2023 гг. [2].

Данная картина подчеркивает критическую важность эффективных методов ранней диагностики патологий почек, в первую очередь злокачественных новообразований. Визуализация, в частности компьютерная томография (КТ), является основным инструментом первичного выявления и оценки почечных образований [1]. Однако ручной анализ исследований, требующий значительного времени и высокой квалификации рентгенолога/уролога, подвержен риску субъективных ошибок и вариативности интерпретации, особенно в условиях высокой нагрузки и скрининговых программ [2]. Это создает предпосылки для внедрения систем искусственного интеллекта, способных автоматизировать и стандартизировать процесс обнаружения патологий.

На текущий момент в научной и инженерной литературе представлено множество подходов к задаче автоматической сегментации почек и почечных новообразований на снимках КТ. В фокусе исследований находятся как улучшение архитектур нейросетей, так и повышение устойчивости моделей в условиях гетерогенных клинических данных.

Одной из ключевых вех в развитии этой области стали международные соревнования KiTS19 и KiTS21 (Kidney Tumor Segmentation Challenge) [3], предложившие открытые размеченные наборы данных и унифицированные метрики (DSC, HD95), что обеспечило сравнимость различных методов. Можно выделить следующие наиболее заметные направления в существующих подходах к решению задачи сегментации почек, представленные в том числе в рамках KiTS19 и KiTS21.

Каскадные и multi-stage модели, например, Kid-Net или nnUNet-cascade, сначала локализируют

орган интереса (почку), а затем уточняют маску опухоли. Это повышает точность и снижает ложноположительные сегментации за счёт ограничения зоны внимания. Однако данные модели склонны к каскадному распространению ошибок.

Трансформерные архитектуры и гибридные решения, например, UNETR, Swin UNet и TransUNet, обеспечивают захват как локальных, так и глобальных зависимостей изображения [4]. Их применение особенно оправдано при работе с гетерогенными опухолями или в условиях слабовыраженных границ. При этом они требуют больших объёмов данных для стабильного обучения и вычислительных мощностей.

Таким образом, целью работы стало создание универсальной системы автоматической детекции и локализации патологий почек без предварительной классификации их морфологического типа.

МЕТОДЫ

Данные

Для обучения и валидации модели использован специально подготовленный клинический набор данных, собранный на базе ОБУЗ «Курского онкологического научно-клинического центра им. Г.Е. Островерхова» (далее – КОНКЦ им. Г.Е. Островерхова) и включающий в себя более 400 мультиспиральных КТ-исследований пациентов с подтверждёнными или предполагаемыми почечными патологиями. Совокупный объём данных составил более 30 000 аксиальных срезов.

В выборку были включены исследования пациентов с различными типами патологий: кисты (по классификации Bosniak I-IV), почечно-клеточный рак (ПКР), ангиомиолипомы, а также анатомические аномалии (подковообразная почка, дистопия почек).

Качество и структурированность обучающего набора данных имеют решающее значение при разработке систем на основе глубокого обучения в медицинской визуализации [5]. Для обеспечения достоверности результатов и воспроизводимости эксперимента была проведена систематическая работа по сбору, предобработке и аннотированию томографических данных.

Подготовка данных

В рамках подготовки данных для обучения особое внимание было уделено вопросам

обеспечения конфиденциальности и соблюдению нормативных требований в области обработки медицинской информации. Все КТ-исследования, использованные в работе, до загрузки данных в систему сегментации проходили процедуру анонимизации.

Удаление персонально идентифицируемой информации включало следующие этапы.

1. Удаление всех DICOM-тегов, содержащих потенциально идентифицируемые данные пациента, включая имя, дату рождения, номер полиса, идентификаторы учреждения, информацию о дате и времени сканирования, а также служебные комментарии, содержащие ФИО оператора и/или врача.
2. Проверку на отсутствие персональных данных в виде графической надписи или обводки персональных данных в заголовках или подписях к изображению (в ряде КТ-сканеров такие данные визуальным образом интегрированы в изображение).
3. Присвоение всем исследованиям псевдонимизированных идентификаторов, не содержащих сведений, позволяющих напрямую установить личность пациента.

Процедура анонимизации проводилась в соответствии с Постановлениями Правительства Российской Федерации № 1119 «Об утверждении Требований к защите персональных данных при обработке в информационных системах персональных данных» [6] и № 211 «О перечне мер, направленных на обеспечение выполнения обязанностей, предусмотренных Федеральным законом «О персональных данных»» [7], требованиями ГОСТ Р 55036-2012 [8] и Федеральным законом № 152-ФЗ «О персональных данных» [9].

В обучающий набор данных включались только те исследования, которые удовлетворяли следующим требованиям:

- 1) толщина среза не более 5 мм (для обеспечения высокой пространственной разрешающей способности);
- 2) обязательное наличие венозной фазы контрастирования;
- 3) отсутствие выраженных артефактов (в том числе от металлических имплантатов и движения пациента);
- 4) полная визуализация обеих почек во всех срезах (исключение: отсутствие органа

вследствие нефрэктомии);

- 5) стандартная ориентация томографических срезов (ось Z направлена краниально-каудально);

- 6) возраст пациентов старше 18 лет.

Исключались:

- 1) исследования без контрастирования;
- 2) неполные серии (например, локальные томограммы одного полюса);
- 3) нестандартные протоколы (например, КТ в положении лёжа на боку);
- 4) исследования детей и подростков.

Все томографические данные в рамках исследования были представлены в формате DICOM (Digital Imaging and Communications in Medicine) – стандартном формате хранения медицинских изображений, поддерживающем обширную метаинформацию, включая параметры съёмки, физические характеристики изображения, а также информацию об исследуемом объекте.

Для корректной работы модели и приведения данных к единообразному формату выполнялась предварительная обработка данных, состоящая в последовательном выполнении следующих этапов.

1. Преобразование значений пикселей в физически интерпретируемые единицы измерения в соответствии со шкалой Хаунсфилда (HU), отражающей плотность тканей относительно воды. Это реализовывалось с помощью встроенной функции `apply_modality_lut()` из библиотеки `pydicom.pixel_data_handlers.util`, применяющей таблицу преобразования (Modality LUT). В основе преобразования лежит формула:

$$HU = RawPixelValue \times RescaleSlope + RescaleIntercept \quad (1)$$

где *RawPixelValue* – исходное значение пикселя; *RescaleSlope* и *RescaleIntercept* – параметры, заданные в DICOM-заголовке, определяющие линейное преобразование интенсивностей.

Данный этап необходим, поскольку большинство диагностических алгоритмов ориентируются именно на значения по шкале Хаунсфилда, отражающие клинически значимую плотность тканей.

При построении шкалы условно принимается следующее: рентгеновская плотность воды равна 0 HU, воздуха – -1000 HU. Тогда, считая

рентгеновскую плотность материала линейно зависимой от его абсолютного коэффициента ослабления μ , получаем формулу для рентгеновской плотности в единицах шкалы Хаунсфилда:

$$HU = \frac{\mu - \mu_{\text{воды}}}{\mu_{\text{воды}} - \mu_{\text{воздуха}}} \times 1000 \quad (2)$$

где μ – линейный коэффициент ослабления рентгеновского излучения; $\mu_{\text{воды}}$ и $\mu_{\text{воздуха}}$ – коэффициент ослабления в воде и воздухе, соответственно (референсное значение).

Например, жировая ткань ~ -100 HU, вода ~ 0 HU, кость > 300 HU, кортикальный слой почек: от 25 до 30 HU и т.п. [10].

2. Для визуальной верификации корректности предобработки и экспертной разметки использовалась функция `apply_voi_lut()` из `rudicom`, преобразующая значения пикселей в 8-битный формат, пригодный для визуального анализа. Функция извлекает параметры Window Center (WC) и Window Width (WW) из DICOM-заголовка и выполняет линейное преобразование с ограничением результата в диапазоне [0; 255]. Преобразование выполняется по следующим формулам:

$$I_{\text{scaled}} = \left(\frac{I_{HU} - (WC - \frac{WW}{2})}{WW} \right) \times 255; \quad (3.1)$$

$$I_{\text{clipped}} = \min(\max(I_{\text{scaled}}, 0), 255) \quad (3.2)$$

где I_{HU} – исходное значение пикселя (в единицах Хаунсфилда), WC – центр окна, WW – ширина окна; I_{scaled} – промежуточное значение после масштабирования.

Применение `apply_voi_lut()` позволяет нивелировать фундаментальное несоответствие между глубиной данных (до 65 536 уровней интенсивности при 16-битной глубине). Этот механизм позволяет сфокусироваться на интересующем диапазоне HU, соответствующем тканям или патологическим изменениям.

Следует отметить, что функция `apply_voi_lut()` может использовать таблицу соответствия (VOI LUT), если она присутствует в DICOM-файле. Однако в используемом нами наборе данных тег VOI LUT Sequence отсутствовал, что обеспечивало строго линейное преобразование на основании параметров окна визуализации.

После преобразования в HU изображения нормализовались линейно в диапазон от 0 до 255 без жёсткого отсека значений по предопределённым порогам (например, [-1000, 400] или [-200, 300]). Такое решение позволяет сохранить максимальную диагностическую информацию о контексте изображения, включая возможные патологические отклонения за пределами «типичных» HU-интервалов. Используемый подход особенно важен при работе с редкими анатомическими вариантами и нетипичными опухолевыми тканями, характеристики которых могут выходить за рамки стандартных диапазонов.

Нормализация выполнялась по формуле:

$$I_{\text{norm}} = 255 \times \frac{I_{HU} - \min(I)}{\max(I) - \min(I)} \quad (4)$$

где I_{HU} – массив изображений в шкале Хаунсфилда, $\min(I)$, $\max(I)$ – глобальные минимумы и максимумы интенсивности в серии.

4. Все изображения были приведены к разрешению 512×512 пикселей. В случае отличий от исходного разрешения применялись:

- масштабирование с сохранением соотношения сторон;
- центрированное дополнение (padding) нулями или повторением краевых значений;
- линейная интерполяция.

Это обеспечивало унификацию входных данных и совместимость с архитектурой модели, ожидающей фиксированный входной размер.

Для повышения обобщающей способности модели и минимизации переобучения была реализована аугментация как на основе геометрических, так и фотометрических преобразований. В качестве геометрических были использованы случайные повороты (до $\pm 10^\circ$), отражения, масштабирование, *elastic deformation*, сдвиги; в качестве фотометрических – варьирование яркости, контрастности, гамма-коррекция. Все изменения были призваны расширить исходный набор данных, имитируя различные условия визуализации, анатомические вариации и технические особенности сканирования, встречающиеся в реальной клинической практике.

Особое внимание уделялось сохранению диагностической ценности изображений: например, при фотометрических преобразованиях диапазон изменения яркости и контраста ограничивался так, чтобы не затереть тонкие

патологические структуры, а геометрические искажения (включая elastic deformation) моделировали естественные деформации мягких тканей, не нарушая анатомической правдоподобности.

Таким образом, этап предварительной обработки обеспечил не только унификацию и стандартизацию входных изображений, но и сохранение клинически значимой информации, необходимой для обучения точной и устойчивой модели. Разнообразие анатомических особенностей и типов патологий в обучающем наборе данных повысило его диагностическую сложность и обеспечило базу для оценки универсальности модели.

РАЗРАБОТКА МОДЕЛИ

Для решения задачи сегментации патологии почек была разработана модель на основе

модифицированной архитектуры U-Net – одной из зарекомендовавших себя структур в области медицинской сегментации. За последние годы U-Net стал основой для множества модификаций, успешно применяемых к сегментации почек и опухолей на снимках КТ. Так, в рамках соревнования KiTS19 и KiTS21 часть топовых решений были основаны на вариантах U-Net [3].

Оригинальный U-Net (Ronneberger et al., 2015) [11] представляет собой симметричную энкодер-декодерную сеть с пропускными соединениями (skip connections), обеспечивающими сохранение пространственной информации при передаче признаков от сжимающих слоёв к восстанавливающим.

На рисунке 1 представлены последовательности свёрточных блоков, операции downsampling (красные стрелки) и upsampling

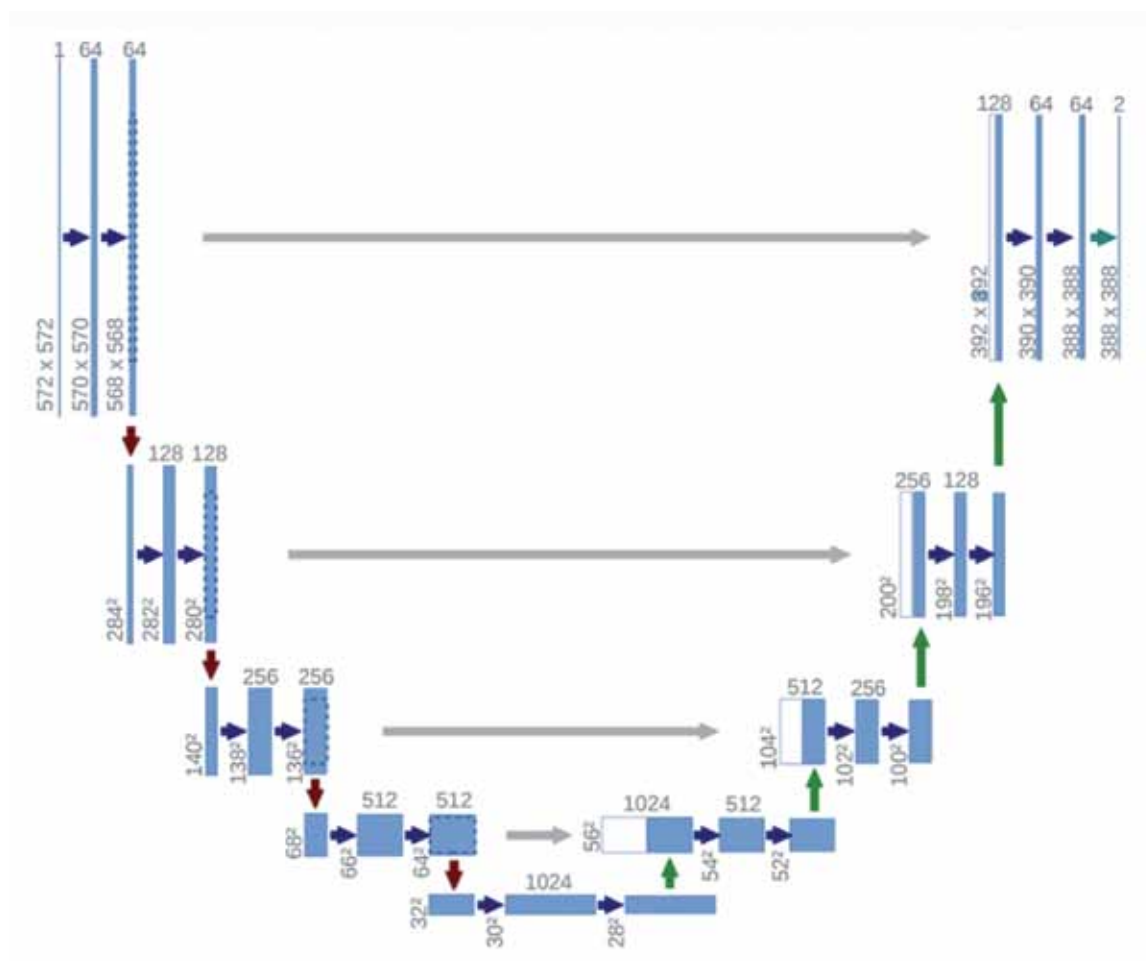


Рисунок 1 – Схематическая структура нейросетевой модели U-Net.

(зелёные стрелки), а также пропускные соединения (skip connections), обеспечивающие передачу пространственной информации между энкодером и декодером. Каждый блок помечен размерностью выходных признаков (высота $\times$ ширина $\times$ количество каналов). Модель принимает входное изображение размером $572 \times 572 \times 1$ и генерирует маску сегментации того же пространственного разрешения, где каждый пиксель содержит вероятность принадлежности к области интереса.

В нашей реализации были внесены следующие модификации, направленные на адаптацию архитектуры к особенностям КТ-изображений почек и расширение возможностей по извлечению контекстной информации.

Глубокая структура (Deeper U-Net): в модели используется пять уровней энкодера и соответствующих декодирующих блоков, что позволяет извлекать признаки на различных уровнях абстракции – от локальных краевых признаков до глобального контекста.

Свертка с шагом 2 вместо операции максимизирующего пулинга (MaxPooling): для уменьшения пространственного разрешения применяется свертка с шагом 2 (Strided Conv), что позволяет избежать потерь информации, характерных для пулинга.

MaxPooling применяется для агрегации информации за счёт выбора максимального значения в небольшой локальной области, что позволяет сократить размер данных, но может приводить к потере значимых признаков, особенно при слабо выраженных границах или мелких структурах. В отличие от этого, свёртка с шагом 2 выполняет пространственное уплотнение карт признаков, одновременно осуществляя фильтрацию и уменьшение размерности. Благодаря этому подходу достигается более плавное и информативное преобразование, поскольку все входные данные участвуют в вычислении выходных значений с учётом их пространственной взаимосвязи.

Такой подход обеспечивает сохранение важных признаков и улучшает устойчивость градиентного потока.

Вместо пакетной нормализации (Batch Normalization) используется групповая нормализация (Group Normalization), что критично при небольшом размере батча (batch size = 4), часто

используемом в задачах медицинской сегментации из-за ограничений по видеопамяти.

Метод исключений (Dropout) в блоках энкодера: для повышения обобщающей способности модели и предотвращения переобучения после каждой сверточной операции по уменьшению разрешения применяется Dropout, что обеспечивает умеренную регуляризацию без значительного замедления сходимости обучения [12].

Блок «бутылочного горлышка»: центральный слой модели (bottleneck) состоит из двух сверток 3×3 с последующей нормализацией, обеспечивает максимальную степень абстрагирования признаков перед переходом к фазе восстановления.

Декодер построен по принципу восстановления пространственного разрешения и состоит из последовательных блоков апсемплинга, реализованных с использованием билинейной интерполяции с масштабированием $\times 2$, за которыми следуют свёрточные блоки. Каждый уровень декодера получает skip-соединение (skip connection) от соответствующего уровня энкодера, что позволяет интегрировать высокоуровневые семантические признаки с детализированной пространственной информацией на более низких уровнях.

Такая архитектура способствует точному восстановлению границ и формы объекта интереса, под которым в данном исследовании понимается патологическое образование в почке, включая опухоли, кисты и смешанные структуры. Сохранение пространственной точности на этапе декодирования критически важно для корректной локализации и сегментации таких образований, особенно при малых размерах или слабой контрастности относительно окружающей паренхимы.

На финальном этапе применяется свертка 1×1 , уменьшающая количество каналов до числа классов (в данном случае – 1), после чего применяется сигмоида для получения вероятностной маски (каждый пиксель получает значение от 0 до 1, интерпретируемое как вероятность того, что данный пиксель принадлежит объекту интереса):

$$\hat{y} = \sigma(W * x + b), \quad \sigma(z) = \frac{1}{1 + e^{-z}} \quad (5)$$

где W и b – обучаемые параметры последнего слоя, $*$ – операция свертки, x – входной тензор

(изображение или признаки), z – логит, выход последнего слоя до активации, $\sigma(z)$ – сигмоидная функция, преобразующая логит z в значение от 0 до 1.

МЕТРИКИ

Задача сегментации в данном исследовании формализуется как задача бинарной пиксельной классификации: каждому пикселю входного изображения необходимо сопоставить метку принадлежности к патологической области (1) или к нормальной ткани (0). Для обучения модели использовалась функция потерь на основе коэффициента Дайса (Dice) и функции Binary Cross-Entropy (BCE). Такой подход обеспечивает устойчивое обучение в условиях сильного дисбаланса классов, обусловленного малой площадью патологических образований на фоне преобладающих здоровых тканей.

Классический коэффициент Дайса (D) определяется по формуле:

$$Dice(P, G) = \frac{2 \times |P \cap G|}{|P| + |G| + \epsilon} \quad (6)$$

где P – предсказанная маска (после применения сигмоиды), G – истинная маска, ϵ – малая положительная константа для избежания деления на ноль.

В качестве функции потерь используется коэффициент Дайса, ориентированный на перекрытие масок, особенно полезный при сильной несбалансированности классов:

$$L_{Dice} = 1 - Dice(P, G) \quad (7.1)$$

$$L_{Dice}(y, \hat{y}) = 1 - \frac{2 \sum y \times \hat{y} + \epsilon}{\sum y + \sum \hat{y} + \epsilon} \quad (7.2)$$

где ϵ – малая положительная константа для избежания деления на ноль, y – истинная бинарная маска, $\hat{y}$ – предсказанная вероятностная маска.

Для более устойчивого обучения также рассматривались гибридные функции потерь, комбинирующие коэффициент Дайса с бинарной кросс-энтропией (Binary Cross Entropy, BCE). BCE – классическая функция логистической регрессии, определяемая по формуле:

$$L_{BCE}(y, \hat{y}) = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)], \quad (8)$$

где $y_i \in \{0, 1\}$ – истинная маска, а $\hat{y}_i \in [0, 1]$ – предсказанная вероятная маска после сигмоиды.

Итоговая функция потерь:

$$L_{Total} = \alpha \times L_{BCE} + (1 - \alpha) \times L_{Dice}, \quad \alpha \in [0, 1] \quad (9)$$

Такой подход позволяет учитывать как локальные ошибки (достигается использованием BCE), так и глобальное перекрытие сегментируемых областей (с помощью Dice), что особенно эффективно при работе с трудноразличимыми или малыми объектами [13].

Для обучения данные были разделены на три подмножества:

- обучающее (train) – 70% исследований;
- валидационное (val) – 15%;
- тестовое (test) – 15%.

Разделение производилось по уникальным пациентам, исключая возможные утечки данных (data leakage) между подмножествами. Балансировка классов (патология/норма) была обеспечена на уровне исследования (60% пациентов с патологиями, 40% – норма).

Модель была реализована на фреймворке PyTorch 2.6.0 с использованием CUDA 12.6. Обучение производилось на одной видеокарте NVIDIA RTX A4000 (16 ГБ VRAM). Обучение модели занимало около 6 часов на одну эпоху, при объеме данных более 30 000 срезов.

На этапе валидации использовалось независимое подмножество данных, не встречавшееся модели в процессе обучения. Особое внимание уделялось качеству сегментации в условиях анатомической изменчивости (подковообразные почки, дистопия) и гетерогенности патологий.

Тестирование модели проводилось на отложенной выборке КТ-исследований, не использовавшихся на этапах обучения и валидации. Эта выборка включала случаи с разнообразными патологическими формами (простые кисты, сложные по классификации Bosniak, солидные образования, ангиомиолипомы), а также анатомические вариации (подковообразная почка, дистопия, гипоплазия).

РЕЗУЛЬТАТЫ

Участники тестирования

Оценка качества работы предложенной модифицированной архитектуры глубокого обучения проводилась на независимой тестовой выборке, не участвовавшей в обучении модели и полученной для тестирования дополнительно,

состоящей из 60 КТ-исследований (около 15 000 аксиальных срезов в венозной фазе), репрезентативно отражающих клиническое разнообразие почечных патологий и анатомических вариантов. Выборка была сформирована строго по принципу разделения по пациентам, что исключает утечку данных и обеспечивает объективность оценки обобщающей способности модели. Все исследования прошли процедуру анонимизации и были независимо размечены двумя сертифицированными рентгенологами с опытом работы в онкоурологии не менее 10 лет. В случае расхождений в аннотации применялось консенсусное решение с привлечением третьего эксперта.

Дополнительно проводился субгрупповой анализ по типам патологий и анатомическим особенностям, что позволило оценить устойчивость модели в сложных клинических сценариях.

Результаты тестирования

На независимой тестовой выборке предложенная модель достигла следующих показателей:

Средневзвешенный коэффициент Дайса: $0,85 \pm 0,06$

ИНТЕРПРЕТАЦИЯ РЕЗУЛЬТАТОВ

Анализ результатов по клиническим подгруппам (таблица 1) демонстрирует высокую и стабильную производительность модели на широком спектре патологий и анатомических конфигураций. Наибольшее значение коэффициента Дайса зафиксировано для простых кист (Bosniak I–II) – $0,88 \pm 0,04$, что обусловлено их чёткими границами, однородной внутренней структурой и высоким контрастом относительно окружающей паренхимы. Модель стабильно и точно воспроизводит форму и размер таких

образований, не включая прилегающие анатомические структуры.

Для сложных кист (Bosniak III–IV) средний DSC составил $0,87 \pm 0,07$, что лишь незначительно ниже результата для простых кист, несмотря на более выраженную неоднородность содержимого, наличие перегородок и зоны умеренного накопления контраста. Это свидетельствует о высокой способности модели к адаптации к сложным морфологическим паттернам. Это может быть объяснено тем, что сложные кисты, несмотря на неоднородность содержимого и наличие перегородок, обладают более выраженной капсулой и чётко определяемым объёмом, что способствует стабильному распознаванию. При этом модель адекватно сегментирует как гиподенсивные участки, так и зоны умеренного накопления контраста, что особенно важно для дифференциальной диагностики в рамках классификации Bosniak.

Наименьшее, но по-прежнему высокое значение DSC, зафиксировано для солидных опухолей, включая ПКР – $0,84 \pm 0,05$. Снижение метрики по сравнению с кистозными образованиями объясняется высокой гетерогенностью опухолевой ткани: сочетанием зон накопления контраста, некрозов, кальцинатов и инфильтративного роста. Тем не менее, модель стабильно выделяет основной объём образования, включая участки с низкой плотностью, что подтверждает её пригодность для задач локализации и объёмной оценки опухолей.

Важно отметить, что модель сохраняет высокую производительность и в условиях анатомических аномалий, где объём данных ограничен, но клиническая значимость высока. При подковообразной почке (4 случая) достигнут DSC $0,83 \pm 0,05$, что лишь на 0,01 ниже показателя для солидных опухолей. Это указывает на

Таблица 1 – Результаты сегментации по клиническим подгруппам

Подгруппа	Количество случаев	DSC
Простые кисты (Bosniak I–II)	22	$0,88 \pm 0,04$
Сложные кисты (Bosniak III–IV)	12	$0,87 \pm 0,07$
Солидные опухоли (включая ПКР)	20	$0,84 \pm 0,05$
Подковообразная почка	4	$0,83 \pm 0,05$
Дистопия почки	2	$0,82 \pm 0,04$

Примечание: DSC – коэффициент Дайса (Dice Similarity Coefficient); значения приведены как среднее $\pm$ стандартное отклонение.



Рисунок 2 – Простая киста (Bosniak I), левая почка.

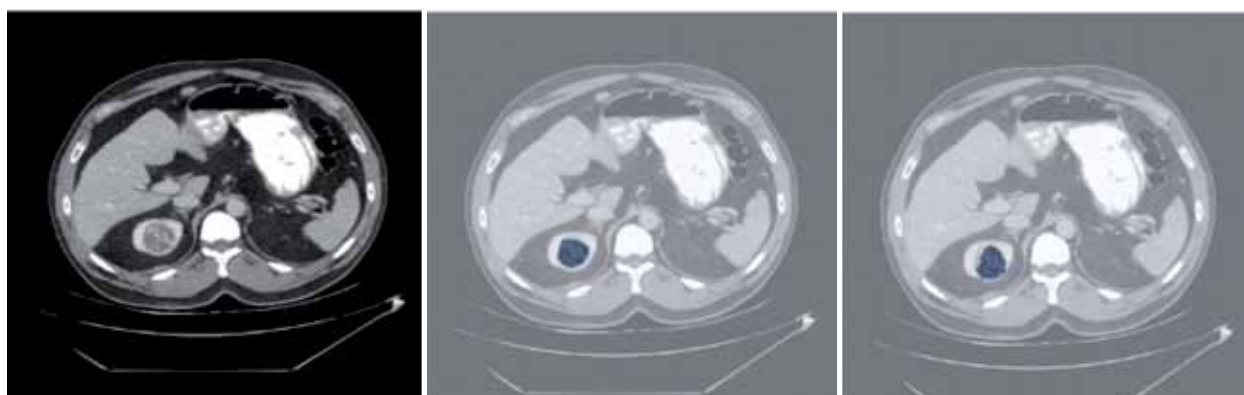


Рисунок 3 – Почечно-клеточный рак, солидное образование, правая почка.

устойчивость архитектуры к изменённой топографии и симметрии органа, включая точную сегментацию перешейка и полярных отделов. В двух случаях дистопии почки DSC составил $0,82 \pm 0,04$, что также находится в высоком диапазоне, несмотря на атипичное положение органа и возможные изменения в сосудистом рисунке.

Для наглядной демонстрации качества сегментации на рисунках 2–4 представлены примеры работы модели на различных клинических случаях. Каждый рисунок включает три компонента:

1. Исходное КТ-изображение.
2. Экспертная разметка (золотой стандарт), выполненная двумя независимыми рентгенологами.
3. Предсказанная моделью маска, наложенная на исходное изображение с полупрозрачностью для визуального сравнения.

Модель точно воспроизводит форму и размер

кистозного образования (рисунок 2). Наблюдается полное совпадение контуров с экспертной разметкой. Отсутствуют ложноположительные сегментации в области паренхимы или сосудов.

Модель демонстрирует нечёткое выделение контура опухоли (рисунок 3), однако в целом обеспечивает корректную сегментацию. Несмотря на выраженную морфологическую гетерогенность образования, неоднородное накопление контрастного препарата и наличие зон некротических изменений, достигнуто точное воспроизведение основного объёма опухоли, включая участки с пониженной плотностью. Отмечается отсутствие ложноположительной экстраполяции маски на прилежащие анатомические структуры. Приведённый случай подтверждает устойчивость алгоритма при обработке сложных радиологических паттернов, характерных для потенциально злокачественных новообразований, что имеет важное значение

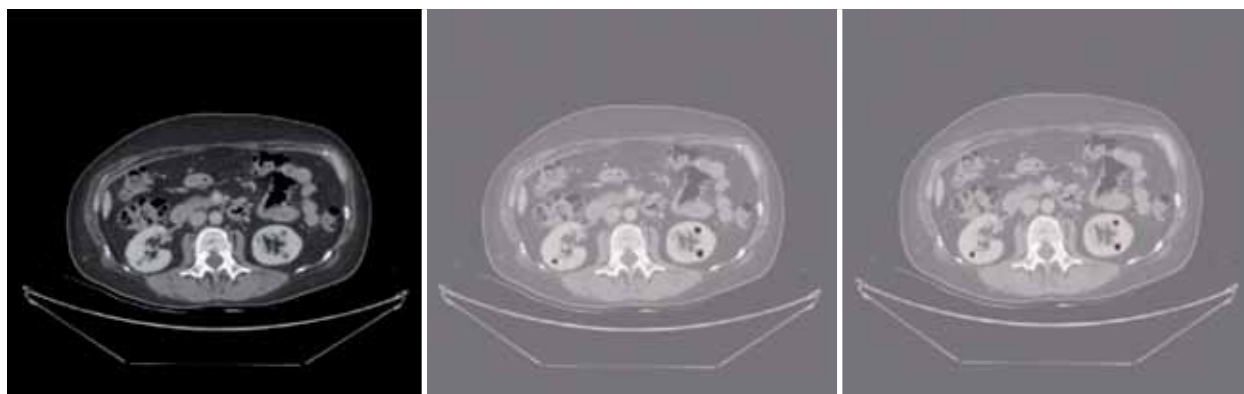


Рисунок 4 – Множественные почечные патологии при типичной анатомии.

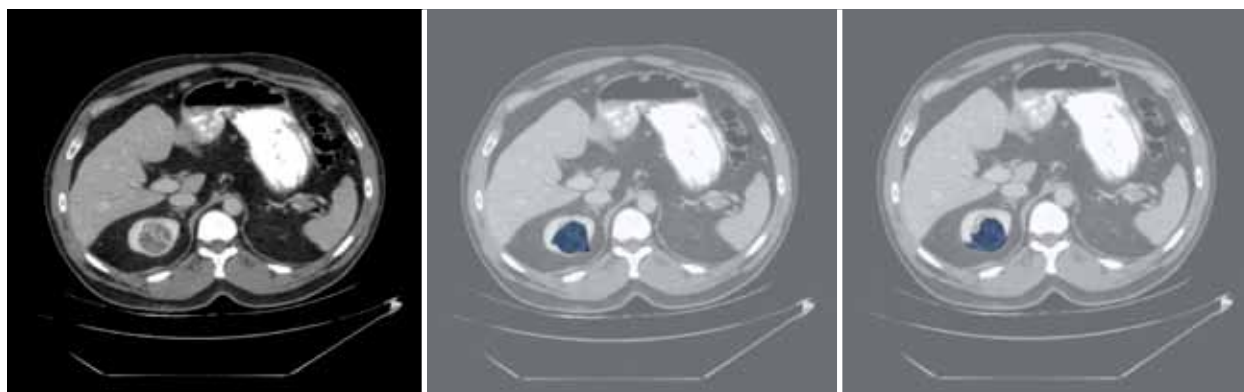


Рисунок 5 – Пример ошибки сегментации.

для ранней диагностики и предоперационного планирования.

Модель демонстрирует высокую точность сегментации в условиях множественных патологических очагов различной морфологии (рисунок 4). Все очаги, в том числе малые (<5 мм), корректно выделены без слияния или пропусков. Отсутствие ложноположительных находок и четкое воспроизведение границ подтверждают надёжность модели при высокой нагрузке патологии даже в условиях стандартной анатомии, что особенно важно для скрининговых и рутинных исследований.

ОГРАНИЧЕНИЯ

Визуальный анализ подтверждает интерпретируемость результатов – модель корректно выделяет кисты (в том числе многокамерные и перегородчатые, Bosniak III-IV), а также солидные опухоли. Ошибки сегментации наблюдаются

преимущественно в двух ситуациях: при наличии гиподенсивных фрагментов (рисунок 5), расположенных вблизи краев изображения, и в случаях, когда патологический очаг сливается с изменённой окружающей паренхимой.

Отдельное внимание заслуживает скорость работы предложенной модели, которая напрямую влияет на её применимость в клинической практике. В ходе тестирования были получены следующие временные характеристики:

- на устройстве с 4 ГБ видеопамяти обработка одного КТ-исследования занимает в среднем 6-8 минут (в зависимости от объёма данных и количества срезов);
- на системах с видеопамятью 8 ГБ и выше время анализа сокращается до 4 минут.

Эти показатели позволяют использовать модель как в условиях стационарных рабочих станций с ограниченными ресурсами, так и в высокопроизводительных средах, обеспечивая

баланс между скоростью и точностью без ущерба для качества сегментации.

Важным ограничением также является отсутствие мультицентральной выборки для тестирования, что не позволяет в полной мере оценить обобщающую способность модели на данных, полученных с использованием различных томографов, протоколов сканирования и в условиях иных клинических учреждений. Все исследования в данной работе были получены в рамках одного регионального онкологического центра, что может привести к неявному смещению (domain shift) и завышенной оценке производительности при внутренней валидации. Для подтверждения работоспособности и клинической применимости модели в реальных условиях требуется её внешняя валидация на независимых наборах данных из нескольких медицинских учреждений с гетерогенными техническими и популяционными характеристиками, что планируется в рамках последующих исследований.

ОБСУЖДЕНИЕ

Для объективной оценки эффективности предложенной модели её результаты были сопоставлены с данными победителей международных соревнований KiTS19 и KiTS21. Сравнение проведено по метрике DSC (коэффициент Дайса), при этом учтены различия в постановке задачи и составе набора данных (таблица 2).

Предложенная модель превосходит или соответствует современным аналогам по ключевым метрикам, несмотря на более широкую постановку задачи. В отличие от nnU-Net и других решений, ориентированных на сегментацию опухолей в рамках уже выделенной почки, наша модель решает обобщённую задачу – детекцию и сегментацию любой патологии, включая

доброкачественные кисты и анатомические аномалии. Это особенно важно для скрининга, где тип образования изначально неизвестен.

Достигнутые результаты подтверждают целесообразность применения специализированных архитектур, ориентированных на задачи медицинской визуализации. Предложенная модель сочетает глубокие свёрточные блоки с пространственным уменьшением разрешения (downsampling) и симметричный декодер с пропускными соединениями (skip connections), что обеспечивает высокую точность локализации границ патологических образований даже в условиях повышенного шума, вариативности режимов контрастирования и различий в технических параметрах компьютерных томографов. Использование линейной нормализации интенсивностей на этапе предобработки без жёсткого ограничения диапазона значений по шкале Хаунсфилда позволяет сохранить широкий спектр контрастных характеристик, включая мягкотканые компоненты и гиподенсивные включения.

Сравнение с литературными данными и аналогичными исследованиями свидетельствует о преимуществе предлагаемой архитектуры в условиях ограниченного обучающего набора за счёт ориентированного включения аномалий и редких патологий в обучающем наборе данных. Так, большинство открытых сегментационных данных (например, KiTS19, KiTS23) фокусируются преимущественно на ПКР и включают ограниченное число вариантов анатомического строения, тогда как в нашем исследовании представлен широкий спектр отклонений, встречающихся в практике онкоуролога.

Отдельного внимания заслуживает аспект потенциального внедрения в клиническую практику. Модель не требует дополнительного

Таблица 2 – Сравнение метрик качества сегментации с современными решениями

Исследование	DSC	Задача	Особенности данных
Isensee et al. (2021), KiTS21	0,84	Сегментация опухолей в пределах почки	Открытые данные KiTS21, ~300 исследований, преобладание ПКР, ограниченное число аномалий
Heller et al. (2019), KiTS19	0,81	Сегментация почек и опухолей	Открытые данные KiTS19 (300 КТ), разнообразные опухоли, включая сложные случаи
Предлагаемая модель	0,85	Бинарная сегментация всех патологий	Клинический набор данных (>400 КТ), включая кисты, опухоли, аномалии, сложные случаи

ввода от врача, не зависит от ручного выбора срезов и может быть интегрирована в PACS-систему для фоновой предобработки КТ-исследований, ускоряя процесс анализа и позволяя врачу сосредоточиться на интерпретации выявленных аномалий.

В будущем предполагается развитие модели в направлении мультиклассовой сегментации с дифференциацией по типу образования, что потребует расширения набора данных для обучения и привлечения мультидисциплинарной экспертной разметки. Кроме того, возможно усиление устойчивости модели за счёт self-supervised обучения и внедрения доменно-устойчивых подходов, что критично для масштабируемости в учреждениях с разной техникой и протоколами.

ВЫВОДЫ

Разработанная модифицированная модель глубокого обучения успешно решает поставленную цель, обеспечивая автоматическую детекцию и локализацию разнообразных почечных патологий без предварительной типизации морфологической структуры. Модель демонстрирует высокую точность сегментации (средневзвешенный коэффициент Дайса 0,85) на широком спектре патологий – от простых и сложных кист до солидных опухолей, включая анатомические аномалии и мелкие очаги (<5 мм).

С практической точки зрения, модель может быть полезна на нескольких этапах клинического маршрута пациента:

1. Первичный анализ КТ при подозрении на патологию.

Автоматическое выделение образований позволяет ускорить процесс интерпретации и снизить риск пропуска малых или атипичных очагов, особенно в скрининговых исследованиях или при перегрузке специалистов.

2. Подготовка к оперативному вмешательству.

Точная локализация солидных образований может быть использована в навигационных и 3D-моделирующих системах для планирования резекции или нефрэктомии.

3. Образование и подготовка молодых специалистов.

Визуализация автоматических контуров может использоваться как вспомогательное средство для обучения врачей-рентгенологов и урологов, помогая формировать клиническое мышление и оценку диагностических изображений.

В перспективе возможна интеграция модели в электронные истории болезни с автоматическим формированием отчётов, подсчётом объёма образований и извлечением ключевых диагностических признаков, что значительно упростит рутинную работу врача и повысит эффективность всей лечебно-диагностической цепи.

Источники финансирования

При поддержке Федерального государственного бюджетного учреждения «Фонд содействия развитию малых форм предприятий в научно-технической сфере» в рамках программы «Студенческий стартап» федерального проекта «Платформа университетского технологического предпринимательства».

ЛИТЕРАТУРА/REFERENCES

1. Российское общество онкоурологов и др. Клинические рекомендации. Рак паренхимы почки. Министерство здравоохранения Российской Федерации. – Москва, 2023. Доступно по: omnidocor.ru/upload/iblock/3f9/ucp3yrezax5ecu0d66dfuz1pvjl597yb.pdf. Дата обращения: 05.05.2024. [Rossijskoe obshchestvo onkourologov i dr. Klinicheskie rekomendacii. Rak parenhimy pochki. Ministerstvo zdravooohraneniya Rossijskoj Federacii. Moskva, 2023. Available at: omnidocor.ru/upload/iblock/3f9/ucp3yrezax5ecu0d66dfuz1pvjl597yb.pdf. Accessed 05.05.2024. (In Russ.)]
2. Злокачественные новообразования в России в 2023 году (заболеваемость и смертность) / под ред. А.Д. Каприна и др. – М.: МНИОИ им. П.А. Герцена – филиал ФГБУ «НМИЦ радиологии» Минздрава России, 2024. – 276 с. [Zlokachestvennyye novoobrazovaniya v Rossii v 2023 godu (zabolevaemost' i smertnost'). A.D. Kaprin i dr., editors. M.: MNIOI im. P.A. Gercena – filial FGBU «NMIC radiologii» Minzdrava Rossii, 2024. 276 s. (In Russ.)]

3. The 2023 Kidney and Kidney Tumor Segmentation Challenge (KiTS23). University of Minnesota Robotics Institute (MnRI); Helmholtz Imaging at DKFZ; Cleveland Clinic's Urologic Cancer Program. 2023. Available at: kits-challenge.org/kits23. Accessed 15.05.2024.
4. Hatamizadeh A, Tang Y, Nath V, Yang D, et al. UNETR: Transformers for 3D Medical Image Segmentation. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). 2022: 574-584.
5. Карпов О.Э., Субботин С.А., Шишканов Д.В. Использование медицинских данных для создания систем поддержки принятия врачебных решений // Системы поддержки принятия решений. – 2019. – №2. – С.16-23. [Karpov OE, Subbotin SA, SHishkanov DV. Ispol'zovanie medicinskih dannyh dlya sozdaniya sistem podderzhki prinyatiya vrachebnyh reshenij. Sistemy podderzhki prinyatiya reshenij. 2019; 2: 16-23. (In Russ.)]
6. Постановление Правительства Российской Федерации от 11 декабря 2012 г. №1119 «Об утверждении Требований к защите персональных данных при обработке в информационных системах персональных данных» // Собрание законодательства Российской Федерации. – 2012. – №50. – Ст.7347. [Postanovlenie Pravitel'stva Rossijskoj Federacii ot 11 dekabrya 2012 g. №1119 «Ob utverzhdenii Trebovanij k zashchite personal'nyh dannyh pri obrabotke v informacionnyh sistemah personal'nyh dannyh». Sobranie zakonodatel'stva Rossijskoj Federacii. 2012; 50. St.7347. (In Russ.)]
7. Постановление Правительства Российской Федерации от 21 марта 2012 г. №211 «О перечне мер, направленных на обеспечение выполнения обязанностей, предусмотренных Федеральным законом „О персональных данных“» (ред. от 06.09.2014) // Собрание законодательства Российской Федерации. – 2012. – №12. – Ст.1455. [Postanovlenie Pravitel'stva Rossijskoj Federacii ot 21 marta 2012 g. №211 «O perechne mer, napravlennyh na obespechenie vypolneniya obyazannostej, predusmotrennyh Federal'nym zakonom „O personal'nyh dannyh“» (red. ot 06.09.2014). Sobranie zakonodatel'stva Rossijskoj Federacii. 2012; 12. St. 1455. (In Russ.)]
8. ГОСТ Р 55036–2012. Информатизация здоровья. Псевдонимизация. Методы и рекомендации / ISO/TS 25237:2008. – Введ. 2013–07–01. – М.: Стандартиформ, 2012. – 16 с. [GOST R 55036–2012. Informatizaciya zdorov'ya. Psevdonimizaciya. Metody i rekomendacii / ISO/TS 25237:2008. Vved. 2013–07–01. M.: Standartinform, 2012. 16 p. (In Russ.)]
9. Федеральный закон от 27.07.2006 №152-ФЗ «О персональных данных» (с изм. и доп. от 08.08.2024) // Собрание законодательства Российской Федерации. – 2006. – №31. – Ст.3451. [Federal'nyj zakon ot 27.07.2006 № 152-FZ «O personal'nyh dannyh» (s izm. i dop. ot 08.08.2024). Sobranie zakonodatel'stva Rossijskoj Federacii. 2006; 31. St.3451. (In Russ.)]
10. Медло А.А., Кобцова Т.В., Башкирова Л.В., Логовой Ю.Н. Методы диагностики и оценки степени распространенности онкологических заболеваний // Практическая онкология. – 2017. – Т.18. – №1. – С.322-329. [Medlo AA, Kobcova TV, Bashkirova LV, Logovoj YUN. Metody diagnostiki i ocenki stepeni rasprostranennosti onkologicheskix zabolevanij. Prakticheskaya onkologiya. 2017; 18(1): 322-329. (In Russ.)]
11. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). 2015: P.234-241.
12. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research (JMLR). 2014; 15: 1929-1958.
13. Юрова А.С. Методы автоматизированной сегментации КТ-изображений брюшной полости: Дис. ... канд. физ.-мат. наук. – Москва, 2018. – 146 с. [YUrova AS. Metody avtomatizirovannoj segmentacii KT-izobrazhenij bryushnoj polosti. [dissertation] Moskva, 2018. 146 h. (In Russ.)]