

<https://doi.org/10.33878/2073-7556-2022-21-4-49-59>



Разработка и применение искусственного интеллекта для обнаружения метастазов в лимфатических узлах при колоректальном раке

Майновская О.А.¹, Ачкасов С.И.¹, Девяткин А.В.², Серых Е.В.², Рыбаков В.В.², Макамбаев Т.Х.², Суслова Д.И.³, Ряховская М.А.³

¹ФГБУ «НМИЦ колопроктологии имени А.Н. Рыжих» Минздрава России (ул. Саляма Адилы, д. 2, г. Москва, 123423, Россия)

²ООО «Вейв Аксес Сервис» (Каменноостровский просп., д. 11, корп. 2, г. Санкт-Петербург, 197046, Россия)

³АО НЦИ (ул. Усачева, д. 33, с.2, г. Москва, 121059, Россия)

РЕЗЮМЕ ЦЕЛЬ ИССЛЕДОВАНИЯ: создание размеченного набора данных (гистосканов лимфоузлов) для использования при разработке систем поддержки принятия врачебного решения (на основе машинного обучения) в патоморфологии, которые позволят определять наличие метастатического поражения лимфоузлов при колоректальном раке.

РЕЗУЛЬТАТЫ: датасет включил 432 файла с цифровыми изображениями и разметкой 1019 лимфоузлов, включавшие лимфоузлы с метастазами и без метастазов. На основе размеченных данных была обучена нейросетевая модель, определяющая вероятность метастатического поражения для каждого пикселя в области интереса — лимфоузле (Dice 0.863 для замещенной ткани, Dice таско 0.923). Кроме того, были реализованы методы пре- и постпроцессинга для представления входных данных в приемлемом для машинного обучения виде и представления ответа ИИ-модели в форме, удобной для восприятия пользователем. Дополнительно разработана нейросетевая модель, предсказывающая вероятность нахождения артефактов в цифровых изображениях лимфоузлов с возможностью формирования карты вероятностей артефактов (Dice таско 0.776; Dice — для артефактов 0.552; IoU таско 0.725 и IoU — для артефактов 0.451).

ЗАКЛЮЧЕНИЕ: разработанная модель представляет собой хорошую основу для реализации полноценного решения, на базе которого может быть разработана система помощи врачам в поиске и оценке замещения тканевых структур и определения метастатического поражения лимфоузлов, обнаружении артефактов и оценке качества цифровых изображений.

КЛЮЧЕВЫЕ СЛОВА: колоректальный рак, стадирование, лимфоузел, метастаз, компьютерная патология (CPATH), искусственный интеллект (ИИ), машинное обучение (ML)

КОНФЛИКТ ИНТЕРЕСОВ: авторы заявляют об отсутствии конфликта интересов

ДЛЯ ЦИТИРОВАНИЯ: Майновская О.А., Ачкасов С.И., Девяткин А.В., Серых Е.В., Рыбаков В.В., Макамбаев Т.Х., Суслова Д.И., Ряховская М.А. Разработка и применение искусственного интеллекта для обнаружения метастазов в лимфатических узлах при колоректальном раке. *Колопроктология*. 2022; т. 21, № 4, с. 49–59. <https://doi.org/10.33878/2073-7556-2022-21-4-49-59>

Development and application of artificial intelligence for the detection of the lymph nodes metastases in colorectal cancer

Olga A. Mainovskaya, Sergey I. Achkasov, Alexey V. Devyatkin, Elena V. Serykh, Vladimir V. Rybakov, Talgat Kh. Makambaev, Daria I. Suslova, Maria A. Ryakhovskaya

¹Ryzhikh National Medical Research Centre for Coloproctology (Salyama Adilya str., 2, Moscow, 123423, Russia)

²Wave Access Service (Kamennoostrovsky str., 11, bld. 2, St. Petersburg, 197046, Russia)

³JSC "National Center of Service integration" (Usacheva str., 33, bld. 2, Moscow, 121059, Russia)

ABSTRACT AIM: to create a marked data set (histoscans of lymph nodes) for use in the development of medical decision support systems (based on machine learning) in pathomorphology, which will allow determining the presence of metastatic lymph node lesions in CRC.

RESULTS: the dataset included 432 files with digital images and markings of 1019 lymph nodes, including lymph nodes with and without metastases. Based on the marked-up data, a neural network model was trained to determine the probability of metastatic lesion for each pixel in the area of interest — the lymph node (Dice 0.863 for the replaced tissue, Dice macro 0.923). In addition, pre- and postprocessing methods were implemented to represent input data in a form acceptable for machine learning and to represent the AI model's response in a form convenient for user perception. Additionally, a neural network model has been developed that predicts the probability of finding artifacts in digital images of lymph nodes with the possibility of forming an artifact probability map (Dice macro 0.776; Dice for artifacts 0.552; IoU macro 0.725 and IoU for artifacts 0.451).

CONCLUSION: the developed model is a good basis for the implementation of a full-fledged solution, on the basis of which a system can be developed to assist doctors in finding and evaluating the replacement of tissue structures and determining metastatic lymph node lesions, detecting artifacts and evaluating the quality of digital images.

KEYWORDS: Computational pathology (CPATH), artificial intelligence (AI), MachineLearning (ML), colorectal cancer (CRC), staging, lymph node metastasis (LNM)

CONFLICT OF INTEREST: the authors declare no conflict of interest

FOR CITATION: Mainovskaya O.A., Achkasov S.I., Devyatkin A.V., Serykh E.V., Rybakov V.V., Makambaev T.Kh., Suslova D.I., Ryakhovskaya M.A. Development and application of artificial intelligence for the detection of the lymph nodes metastases in colorectal cancer. *Koloproktologia*. 2022;21(4):49–59. (in Russ.). <https://doi.org/10.33878/2073-7556-2022-21-4-49-59>

АДРЕС ДЛЯ ПЕРЕПИСКИ: Майновская О.А., ФГБУ «НМИЦ колопроктологии имени А.Н. Рыжих» Минздрава России, ул. Саяма Адилы, д. 2, Москва, 123423, Россия; e-mail:info@gnck.ru

ADDRESS FOR CORRESPONDENCE: Mainovskaya O.A., Ryzhikh National Medical Research Centre for Coloproctology, Salyama Adilya str., 2, Moscow, 123423, Russia; e-mail:info@gnck.ru

Дата поступления — 24.08.2022
Received — 24.08.2022

После доработки — 20.09.2022
Revised — 20.09.2022

Принято к публикации — 09.11.2022
Accepted for publication — 09.11.2022

Точность диагностики и стадирования опухолей различной локализации определяет прогноз и выбор тактики лечения, во многом обеспечивая его успех. В диагностических дисциплинах, таких как лучевая диагностика и патология (патоморфология), диагностика опухолевого процесса основана на анализе изображений путем распознавания определенных паттернов и их интерпретации. Одной из главных проблем в диагностике является воспроизводимость диагностических признаков и паттернов среди врачей, что в условиях возрастающих требований прецизионной медицины может превысить возможности визуальной оценки человеком. Кроме того, проблема дефицита кадров в диагностических дисциплинах, особенно патоморфологии, приводит к увеличению нагрузки на врачей, которым приходится просматривать и анализировать большое количество изображений, что может отражаться на качестве диагностики. Развитие компьютерных технологий в патоморфологии, появление оборудования для сканирования гистологических препаратов с получением цифрового изображения всего гистологического среза/препарата с высоким разрешением — технология whole slide images (WSI) и возможность его исследования с помощью компьютерных программ анализа изображения значительно расширило возможность проведения морфометрических исследований и определения прогностических факторов при опухолях различных локализаций.

Применение компьютерных технологий для анализа цифровых изображений гистологических срезов

(Computational pathology, CPATH) с использованием алгоритмов машинного обучения (Machine learning, ML) и искусственного интеллекта (ИИ/Artificial intelligence) для анализа цифровых изображений гистологических препаратов раскрыло новые возможности CPATH для диагностики в патологии. За последние годы произошел переход к более продвинутым моделям ИИ с решением более сложных диагностических задач [1–4]. Примером перехода от анализа ограниченных областей к исследованию всего цифрового изображения препарата с целью диагностики могут служить исследования CAMELYON 16 и CAMELYON 17, в которых решалась задача обнаружения метастазов рака молочной железы в лимфатических узлах с использованием программ CPATH. Основной целью исследования являлось воспроизведение рутинной диагностики по обнаружению макро- и микрометастазов в лимфатических узлах посредством разработанных компьютерных алгоритмов. По результатам исследования лучшие тестируемые алгоритмы работали сопоставимо с лучшими патологами по определению макрометастазов, в то время как при обнаружении микрометастазов наиболее эффективные алгоритмы продемонстрировали лучший средний показатель AUC (0,885 [95%CI 0,812–0,997]) по сравнению с показателем AUC лучших патологов (0,808 [95%CI, 0,704–0,908]). CAMELYON16 явилось одним из первых исследований, продемонстрировавшим, что интерпретация цифровых изображений алгоритмами глубокого обучения может выполняться с точностью, которая может соперничать с человеческими

возможностями [6, 7]. Клиническая направленность CAMELYON16 явилась мощным стимулом, продвигающим разработку ИИ в данном направлении.

Следует отметить, что морфологическое исследование лимфатических узлов в удаленном препарате при опухоли любой локализации является трудоемким процессом, включающим обнаружение лимфоузлов в препарате и их дальнейшую оценку после гистологической обработки. В тоже время, именно количество исследованных лимфатических узлов отражает качество морфологического исследования удаленного препарата и позволяет максимально точно стадировать опухоль, определяя категорию pN [8, 9]. Таким образом, перед патологом стоит задача исследования большого количества лимфатических узлов посредством световой микроскопии от каждого пациента с целью обнаружения метастазов. Автоматизация повторяющихся и трудоемких задач в контексте клинической практики может оказать большое влияние на оптимизацию нагрузки патологов, что в условиях дефицита кадров позволит обеспечить необходимое качество гистологических исследований.

Колоректальный рак (КРР) является одним из наиболее распространенных онкологических заболеваний в мире, занимая второе место по уровню заболеваемости и смертности в России [10]. Улучшение результатов лечения и снижение смертности от колоректального рака является социально значимой задачей, одним из инструментов решения которой является повышение точности послеоперационного стадирования КРР путем оптимизации и повышения качества патоморфологического исследования операционного

препарата. Стадия КРР на основе результатов патоморфологического исследования в настоящее время является основным фактором прогноза, определяющим дальнейшую тактику лечения и прогноз. В действующих Клинических рекомендациях (2020), адъювантная химиотерапия рекомендуется всем больным колоректальным раком при стадии pT1-4N1-2M0. Статус лимфатических узлов, а именно их метастатическое поражение, является основным показателем, определяющим назначение адъювантной химиотерапии [11, 12].

В связи с этим, клинической задачей проведенного исследования явилось улучшение качества патоморфологической оценки лимфатических узлов и, следовательно, стадирования рака толстой кишки за счет создания перспективного информационного продукта, позволяющего автоматически распознавать на цифровых изображениях лимфатических узлов (гистосканах) метастатически пораженные лимфатические узлы с помощью технологий искусственного интеллекта (ИИ).

МАТЕРИАЛЫ И МЕТОДЫ

Исследование проводилось с использованием цифровых изображений лимфатических узлов, полученных с помощью сканирующего микроскопа Aperio T2 Leica (при увеличении $\times 20$) с применением технологии WSI в формате SVS с ручной разметкой нормальной лимфоидной ткани, метастазов и артефактов (с использованием пакета программ для анализа цифровых изображений Aperio Imagescope V12.4.0,

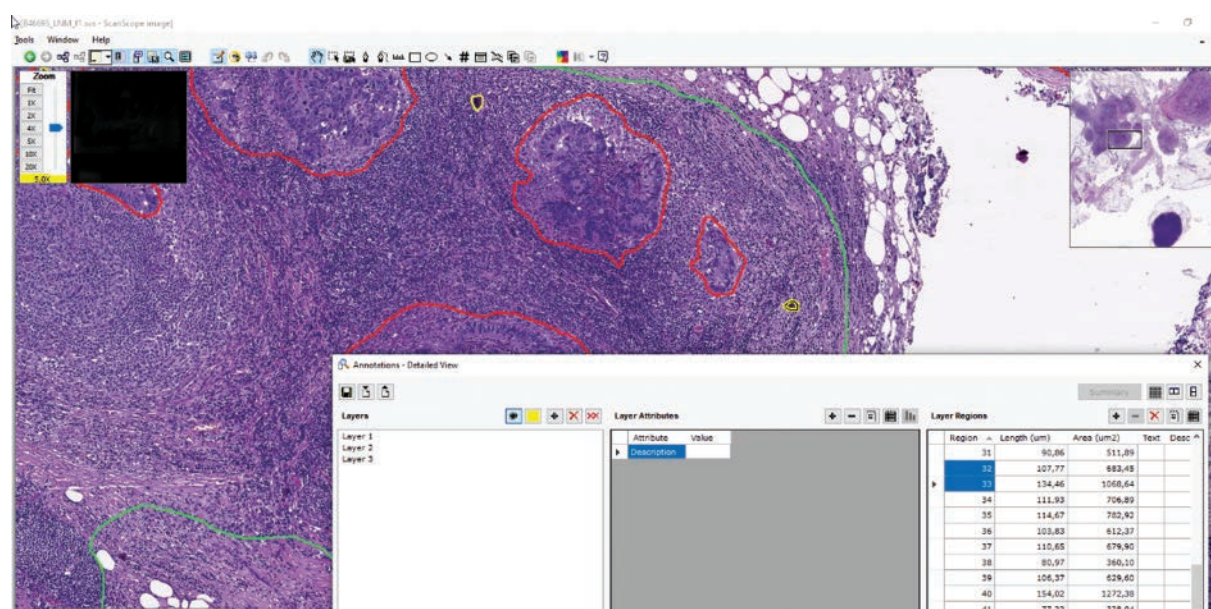


Рисунок 1. Пример отображения гистоскана и таблицы аннотаций в Aperio Image Scope

Figure 1. Example of histoscan and annotation table display in Aperio Image Scope

Aperio Technologies). Для создания базы данных был использован архивный материал отделения патологической анатомии ФГБУ «НМИЦ колопроктологии имени А.Н. Рыжих» Минздрава России.

В рамках исследования использовались снимки размером более 5000 пикселей по каждой стороне, хранящиеся в формате RGB8. Исходное разрешение находилось, в среднем, на уровне 0.5 мкм/пкс. Всего для работы было использовано 432 файла, что представляет порядка 148 Гб данных, содержащих цифровые изображения 1019 лимфоузлов (566 лимфоузлов с метастазами и 453 без метастазов). Разметка файлов представляла собой текстовые данные в формате XML.

При ручной разметке (аннотации) на цифровом изображении лимфоузлов выделялись зоны трех основных классов, требуемых для решения поставленной задачи: «Здоровая ткань» (неизменная лимфоидная ткань), «Замещенная ткань» (метастаз) и «Артефакты». Пример изображения разметки и управления описаниями представлен на рисунке 1. Поскольку распределение площади, количества артефактов и замещающей ткани существенно отличалось, была создана отдельная система определения артефактов для исключения дисбаланса при машинном обучении.

В процессе обучения глубоких нейросетей для задачи изображения на использованную модель использовался графический адаптер. Для того, чтобы избежать сжатия цифровых файлов в формате SVS и потери данных при адаптации большого разрешения использованных цифровых файлов к ограниченному

объему видеопамати на графическом адаптере, использовался метод разделения большого по размеру цифрового изображения по сетке на ячейки (crops — кропы), что позволило обрабатывать изображение в высоком разрешении, занимая при этом адекватное количество видеопамати для машинного обучения. Разделение изображения на кропы было реализовано на основе регулярной сетки с конкретным размером ячейки (за основу взята ячейка 1024 × 1024 пикселя) на конкретном разрешении (за основу взято разрешение 2 мкм/пкс). Схематичное изображение данного процесса представлено на рисунке 2.

В ходе машинного обучения проводилась аугментация данных (изменение данных в допустимых пределах для создания большей вариативности, с целью уменьшения шанса переобучения модели, а также повышения устойчивости модели к разного рода возмущениям).

В ходе обучения модели были использованы следующие аугментации:

- Конвертация в серое изображение (шанс = 30%)
- Отражение по вертикали (шанс = 50%)
- Отражение по горизонтали (шанс = 50%)
- Случайный поворот на 90 градусов (шанс = 50%)
- Один из (шанс = 40%)
- Гауссовый шум (шанс = 50%)
- Мультипликативный шум по каждому каналу (шанс = 30%)
- Сжатие изображение в пределах 90% качества (шанс = 50%)
- Один из (шанс = 20%)
- Размытие смещением (шанс = 20%)

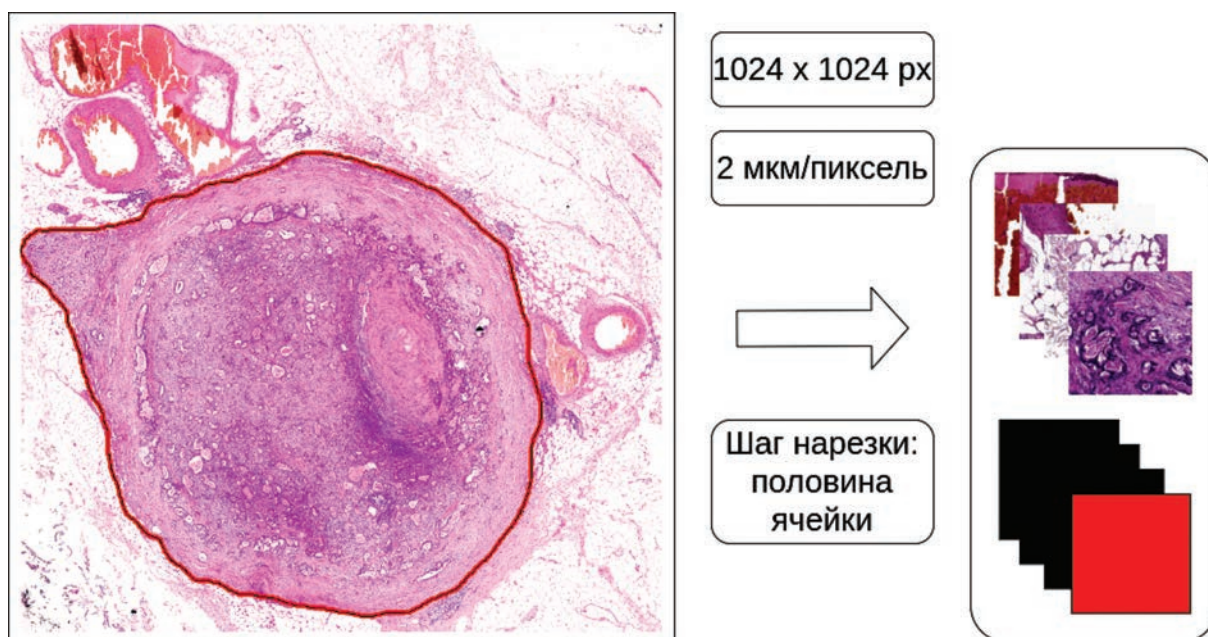


Рисунок 2. Схема методики разделения («нарезки») изображения
Figure 2. Diagram of the method of separation («slicing») of the image

- Медианное размытие (шанс = 20%)
- Гауссово размытие (шанс = 20%)
- Случайное размытие (шанс = 20%)
- Случайная нормализация/изменение яркости (шанс = 30%)
- Случайное изменение составляющих изображения в пространстве HSV (шанс = 30%)
- Случайное изменение масштаба (10%), смещения (10%), поворота (45 градусов) с черным/белым фоном (шанс = 50%)

Указанные аугментации позволяли изменять изображение в ходе обучения для уменьшения вероятности переобучения и создания искусственной разнообразности данных.

В качестве основы для обучения модели сегментации тканей была использована архитектура DeepLabV3 (<https://arxiv.org/abs/1706.05587>).

Для поиска артефактов за основу была взята архитектура Unet (<https://arxiv.org/abs/1505.04597>).

Обучение моделей производилось с использованием следующих параметров и алгоритмов:

- Оптимизатор Adam
- Начальный learning rate (LR) = $1e-4$
- LR диспетчер Reduce On Plateau
- Количество эпох ожидания = 10
- Коэффициент уменьшения LR = 0.5
- Минимальный LR = $1e-7$
- Количество эпох = 200
- Размер batch — в зависимости от размеров изображения и доступной видеопамяти от 8 до 40

Для загрузки пакета изображений с разрешением 1024×1024 использовалась видеокарта с объемом памяти 16 ГБ (Tesla T4). Выборки для обучения

и валидации формировались в виде набора пар изображение-маска в формате «.PNG». Изображения и маски представляли собой области, которые были вырезаны из конкретного положения изображения (Рис. 3).

Результаты обучения моделей

В качестве основного показателя для оценки результатов обучения моделей был использован показатель DICE (Мера Сёренсена), отображающий качество перекрытия предсказанной маски над истинной (ручная разметка изображения). Данная метрика в контексте задачи сегментации представляет собой отношение удвоенного пересечения истинной разметки и предсказанной к сумме их площадей. При идеальном наложении предсказанной маски на разметку DICE равняется единице, при полном несоответствии предсказания маске DICE равен 0. Альтернативным (более формальным) способом определения данной меры служит формула, где True Positive означает количество пикселей с предсказанием, совпадающим с разметкой, False Positive — пиксели, где была предсказана маска, но не было разметки и False Negative — пиксели, где была разметка, но модель не предсказала наличие маски (https://en.wikipedia.org/wiki/S%C3%B8rensen%E2%80%9393Dice_coefficient).

В таблицах с метриками указаны макро-метрики, которые рассчитываются через расчёт меры для каждого из классов, после чего берется их невзвешенное среднее, а также метрика для целевого класса по отдельности.

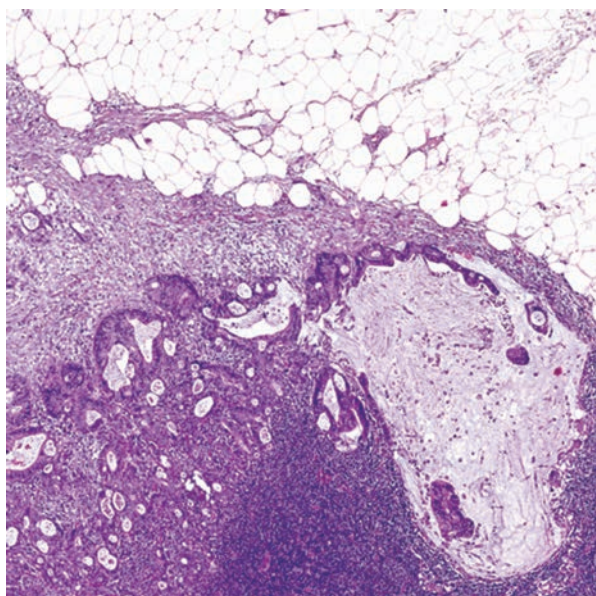


Рисунок 3. Пример пары кроп-маска

Figure 3. Example of a crop-mask pair

РЕЗУЛЬТАТЫ

Точность сегментирования ткани с помощью разработанной модели позволила сделать достаточно аккуратную оценку поражения лимфатических узлов. В результате обучения нейросети окончательные метрики сегментации тканей имели следующие показатели:

Показатель	Значение
Dice macro*	0,923
Dice замещенной ткани**	0,863

*Dice macro — общая усредненная оценка классификатора, применяющегося для нескольких одинаково значимых параметров (норма/патология, норма/артефакт); **Dice замещенной ткани — метастаз

Разработанный алгоритм определения лимфоузлов и классификации типа ткани (неизменная лимфоидная ткань/замещенная опухолью) работает на основе выделения крупных областей (зон интересов) контурами и обозначения их с помощью описывающих прямоугольников с дальнейшей оценкой процента замещенной ткани относительно всей найденной лимфоидной ткани в этой зоне (Рис. 4).

Диагностическая эффективность разработанной модели при определении неизменной лимфоидной ткани продемонстрировала чувствительность — 95,6% и специфичность — 98,9%; при выявлении метастазов в лимфоузлах чувствительность составила 87,3%, специфичность — 99,5%. При визуальной оценке работы нейросети получены весьма осмысленные результаты, показавшие высокую специфичность модели с выделением в ряде случаев большей зоны опухолевой ткани (метастаза), чем отмечено в разметке и способностью отмечать небольшие очаги опухоли в массиве лимфоидной ткани (Рис. 5).

Точность анализа цифрового изображения во многом зависит от его качества, которое, в свою очередь, определяется качеством исходного гистологического препарата. Если при просмотре патологом гистологического препарата или гистоскана любой артефакт (дефект гистологического препарата или его цифрового изображения, возникший в ходе гистологической обработки образца ткани или при сканировании) может быть легко определен и учтен при его оценке, то наличие артефакта при компьютерном анализе цифровых изображений может привести к неправильной оценке и значительному снижению точности анализа (Рис. 6). Поэтому параллельно с машинным обучением обнаружения метастазов в лимфатических узлах проводилось создание алгоритма распознавания артефактов в зонах интереса (ткань лимфоузлов). Предсказание артефактов происходило по аналогичной логике с сегментацией тканей — модель должна предсказывать вероятность нахождения артефактов в местах на изображении так, что по результату предсказаний можно сформировать карту вероятностей артефактов.

При решении данной задачи была добавлена метрика Intersection over Union (IoU). Она представляет собой простое отношение пересечения предсказанной и размеченной области к их объединению и во многом похожа на метрику DICE, однако в большей степени штрафует за отдельно взятые плохо предсказанные примеры, что важно учесть в контексте задачи определения артефактов.

https://en.wikipedia.org/wiki/Jaccard_index

Также как и для DICE, ниже приведены значения макро-метрики и метрики для целевого класса, которые были получены в результате обучения модели сегментации артефактов:

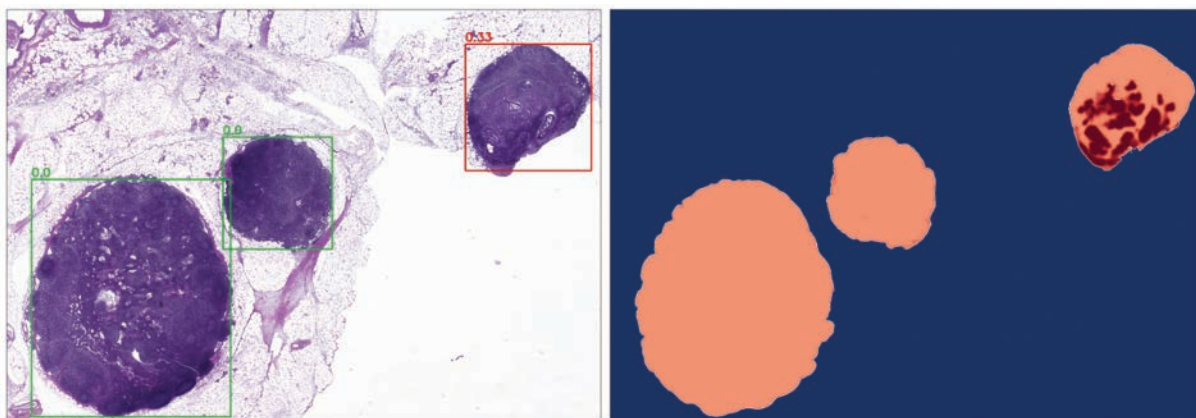


Рисунок 4. Пример использования метода определения неизменных лимфоузлов и лимфоузлов с метастазами/замещенной тканью (слева — разметка, справа — тепловая карта вероятностей предсказания)

Figure 4. An example of using the method for determining normal lymph nodes and lymph nodes metastases/replaced tissue (on the left is a markup, on the right is a heat map of prediction probabilities)

Показатель	Значение
Dice macro	0.776
Dice для артефактов	0.552
IoU macro	0.725
IoU для артефактов	0.451

**Dice macro/IoU macro — общая усредненная оценка классификатора, применяющегося для нескольких одинаково значимых параметров (норма/патология, норма/артефакт);*

При оценке работы модели сегментации артефактов отмечалась тенденция хорошего предсказания малых зон артефактов в ткани лимфоузлов (Рис. 7) и недостаточная точность в способности модели предсказывать большие артефакты, которые модель не отличала от фона или светлых зон в неизменной ткани лимфоузла (Рис. 8).

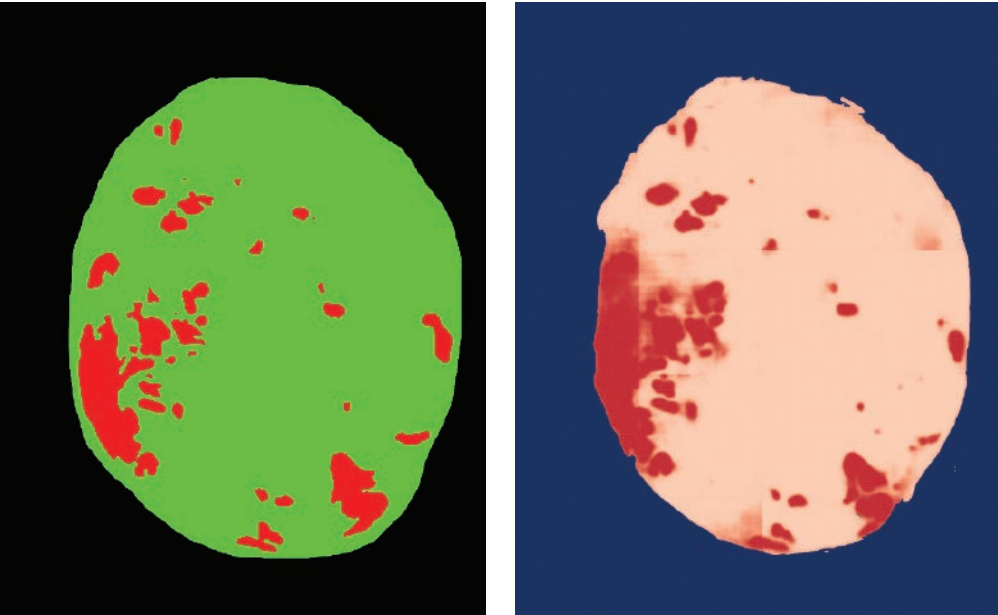


Рисунок 5. Пример специфичности модели (слева — разметка, справа — тепловая карта вероятностей предсказания)
Figure 5. An example of the specificity of the model (on the left — the markup, on the right — the heat map of prediction probabilities)

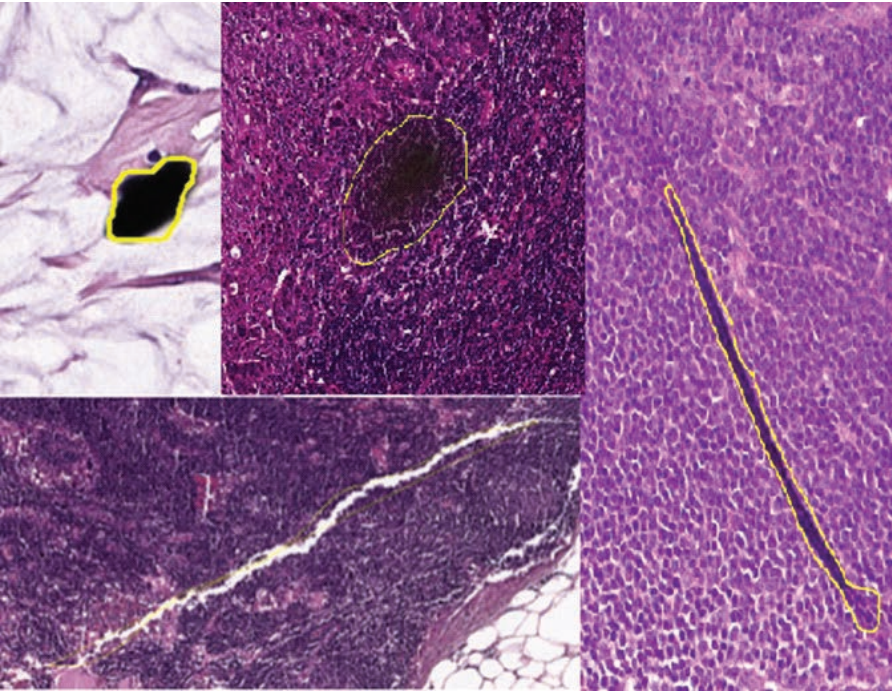


Рисунок 6. Примеры определяемых артефактов в гистосканах
Figure 6. Examples of detectable artifacts in histoscans

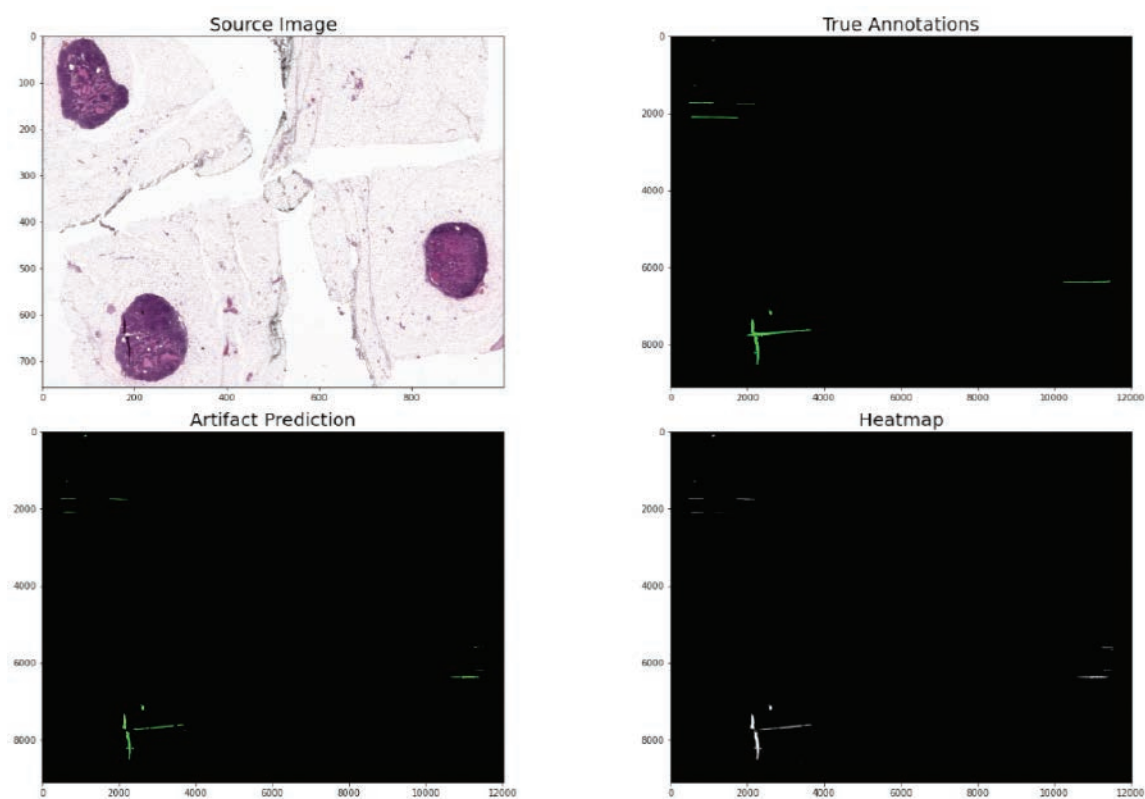


Рисунок 7. Пример хорошего предсказания небольших артефактов

Figure 7. A good prediction of small artifacts

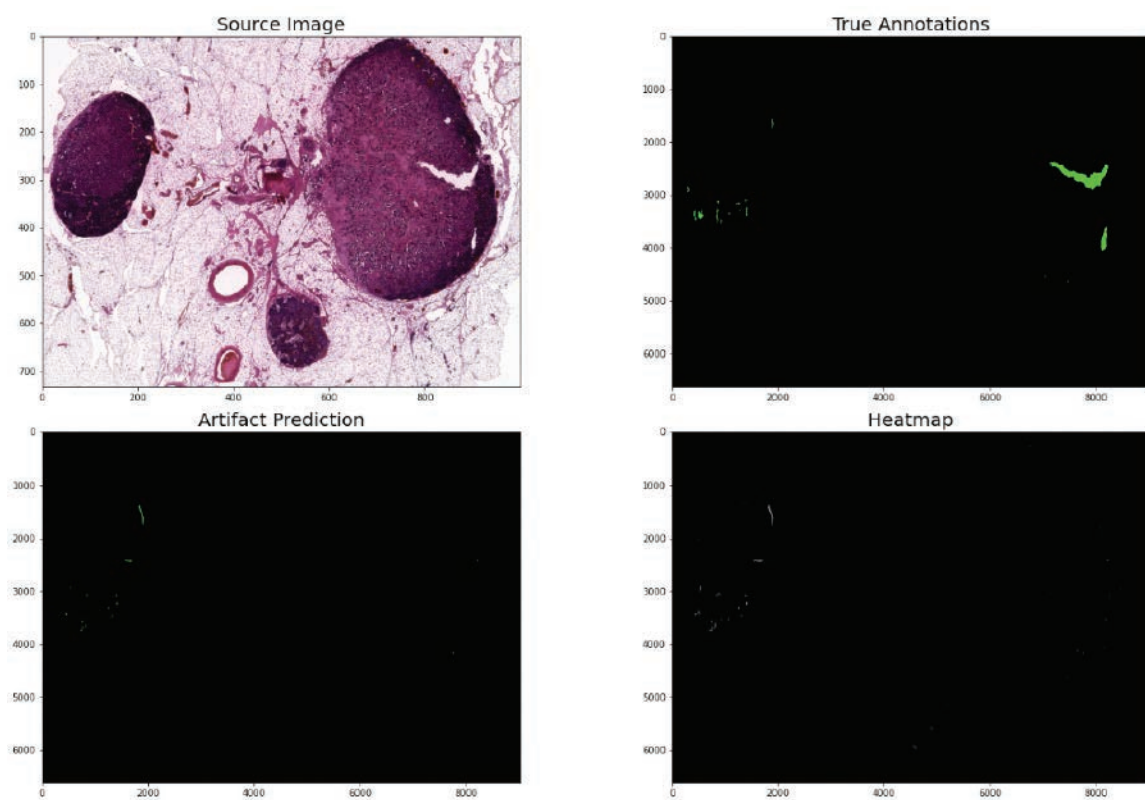


Рисунок 8. Пример ошибки в большом разрыве среза лимфоузла

Figure 8. An example of an error in a large rupture of a lymph node slice

ОБСУЖДЕНИЕ

Процесс разработки модели компьютерного анализа цифровых изображений в патологии обозначил ряд проблем, которые связаны с решением поставленной задачи. Прежде всего, это создание базы данных с большим количеством полноразмерных цифровых изображений (WSI), что необходимо в разработке алгоритмов CPATH. Выполнение данной задачи возможно в патологоанатомических лабораториях, оснащённых современными высокопроизводительными сканерами и обладающими достаточным количеством исходного гистологического материала. Объем исследований лимфатических узлов при КРР, проводимый в отделе патоморфологии ФГУБ «НМИЦ колопроктологии имени А.Н. Рыжих» Минздрава России, позволил создать значительную по объему базу данных для получения искомых результатов.

Сбор аннотаций WSI создавался путем ручной разметки исследуемых областей изображения (выделение различных типов ткани или расположения определенных типов клеток), что представляло собой очень трудоемкий процесс, требующий от патолога соответствующих знаний и квалификации. Однако для процесса машинного обучения (ML) с целью создания эффективных алгоритмов анализа большое количество WSI с ручной аннотацией является необходимым условием, чтобы модель CPATH научилась выполнять поставленную задачу. Это диктует необходимость разработки методов, облегчающих эффективное создание этих аннотаций.

В проведенном исследовании была создана база данных с ручной разметкой (аннотацией) порядка 432 файлов, содержащих WSI 1019 лимфоузлов, что позволило провести обучение модели ИИ для решения задачи сегментации как типа ткани, так и определения зон возможных артефактов на цифровом изображении.

Модель сегментации типа ткани, разработанная с использованием алгоритма DeepLab V3, была выбрана, исходя из экспериментов на основе вариации разрешения, типа модели и размера изображения. По представленным показателям модель сегментации типа ткани имеет высокую точность определения метастазов (замещений) в ткани лимфоузла со значением DICE около 0.9: (DICE макро — 0,923 и DICE замещения — 0,863), что позволило продемонстрировать разработанной модели ИИ хорошую диагностическую эффективность при определении неизменной лимфоидной ткани и выявлении метастазов в лимфоузлах с показателями чувствительности 95,6% и 87,3%, специфичности 98,9% и 99,5%, соответственно. Следует отметить, что полученные показатели позволяют разработанной модели ИИ

с высокой вероятностью избежать как ложноотрицательных, так и ложноположительных результатов, что является принципиально важным при оценке статуса лимфоузлов.

Модель сегментации артефактов показала значительно меньшую точность (метрики DICE и IoU составили для макро 0,776 и 0,725, соответственно; для замещения — 0,552 и 0,451, соответственно), что объясняется сложностью решения задачи из-за явно более разнообразного распределения по размерам целевых объектов и вариативности представлений артефактов. Результаты тестирования показали, что модель хорошо определяет малые артефакты в цифровых изображениях лимфоузлов, но при этом более крупные объекты, особенно разрывы, путает с зонами фона и светлыми включениями в лимфоидную ткань. Создание датасета с большим количеством возможных артефактов и обучение созданной модели в их распознавании позволит улучшить полученные результаты. Проведенное исследование показало, что решение задачи распознавания артефактов является актуальной проблемой, поскольку позволяет избежать ошибок при диагностике и, в тоже время, осуществлять контроль качества цифровых изображений, который является необходимой процедурой в цифровой патологии.

Следует отметить, что в ходе исследования для реализации обучения модели была разработана специальная предобработка цифровых изображений, позволяющая обрабатывать огромные WSI без крупных затрат как оперативной, так и видеопамати.

Тестирование созданной модели было проведено на группе отобранных WSI (тестовая выборка) из созданной нами базы данных. Для создания окончательного продукта необходимо тестирование полученной модели на цифровых изображениях из других источников, поскольку ранее проведенные исследования показали, что алгоритмы CPATH оптимально работают с данными из источников, на которых они были обучены, но хуже (иногда значительно) работают с данными из других датасетов [4]. Это позволит сделать алгоритмы CPATH как можно более надежными в ответ на вариации, которые могут встретиться в реальной практике. Дополнительным подходом к работе алгоритмов CPATH с различиями между цифровыми изображениями из разных источников данных является нормализация изображений по общему стандарту, что, в свою очередь, считается одной из основных проблем в цифровой патологии.

Тем не менее, результаты проведенных исследований с разработкой алгоритмов CPATH, которые показали эффективность, сравнимую с работой патологоанатомов, при выполнении ряда конкретных диагностических задач, как и результаты проведенного нами исследования можно рассматривать

как многообещающие. В настоящее время многие технические проблемы в разработке и применении ИИ в патологии уже преодолены, однако его роль в практической работе однозначно еще не определена [3,4,7]. Обсуждая вопрос о роли ИИ в патологии, важно понимать, что работа патолога заключается не только в исследовании фрагментов ткани под микроскопом. Она включает интеграцию диагностического процесса, клинических данных и конкретных обстоятельств пациента с последующим формулированием своего заключения о результатах морфологического исследования как для клиницистов, так и для пациента. Поэтому алгоритмы ИИ вряд ли в ближайшее время могут рассматриваться как самостоятельный инструмент для диагностики. Что может быть достигнуто в ближайшее время, так это создание алгоритмов ИИ, которые работают вместе с патологоанатомами, помогая сократить объем трудоемкой монотонной, но необходимой работы, в частности такой, как выявление метастазов в лимфатических узлах. Полученные результаты нашего исследования позволяют с определенной уверенностью говорить о возможности применения ИИ для решения подобного рода диагностических задач в патологии.

ЗАКЛЮЧЕНИЕ

Результаты, полученные в проведенном исследовании, демонстрируют, что разработанная модель представляет собой хорошую основу для реализации полноценного решения, на базе которого может быть разработана система помощи врачам в поиске и оценке замещения тканевых структур, обнаружении артефактов и оценке качества цифровых изображений. Помимо задачи сбора достаточно больших наборов аннотированных WSI, необходимо провести проспективные исследования, чтобы показать значение ИИ в патоморфологической диагностике.

ЛИТЕРАТУРА

1. Ciresan DC, Giusti A, Gambardella LM, Schmidhuber J. Mitosis detection in breast cancer histology images with deep neural networks. *Med Image Comput Assist Interv*. 2013;8150:411–418.
2. Cruz-Roa A, et al. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. *In Proc. SPIE Medical Imaging*. 2014;9041:904103.
3. Chetan L, Srinidhi et al. Deep neural network models for computational histopathology: A survey. *Medical Image Analysis*. 2020. www.elsevier.com/locate/media
4. Jeroen van der Laak et al. Deep learning in histopathology: the path to the clinic. *Nature Medicine*. May 2021;27:775–784.
5. Babak EB, Mitko V, Johannes van Diest P, et al. and the CAMELYON16 Consortium. *JAMA*. 2017;318(22):2199–2210.
6. Litjens G, et al. 1399 h&e-stained sentinel lymph node sections of breast cancer patients: the camelyon dataset. *Giga Science*. 2018;

УЧАСТИЕ АВТОРОВ

Концепция и дизайн исследования: *Девяткин А.В., Рыбаков В.В.*

Сбор и обработка материала: *Ряховская М.А., Макамбаев М.Т., Суслова Д.И., Майновская О.А.*

Статистическая обработка: *Серых Е.В.*

Написание текста: *Майновская О.А., Девяткин А.В.*

Редактирование: *Ачкасов С.И.*

AUTHORS CONTRIBUTION

Concept and design of the study: *Alexey V. Devyatkin, Vladimir V. Rybakov,*

Collection and processing of materials: *Maria A. Ryakhovskaya, Talgat Kh. Makambaev, Daria I. Suslova, Olga A. Mainovskaya*

Statistical processing: *Elena V. Serykh*

Text writing: *Olga A. Mainovskaya, Alexey V. Devyatkin*

Editing: *Sergey I. Achkasov*

СВЕДЕНИЯ ОБ АВТОРАХ (ORCID)

Майновская О.А. — 0000-0001-8189-3071

Ачкасов С.И. — 0000-0001-9294-5447

Девяткин А.В. — 0000-0002-5712-1626

Серых Е.В.

Рыбаков В.В.

Макамбаев Т.Х.

Суслова Д.И. — 0000-0003-1106-5486

Ряховская М.А. — 0000-0002-2243-1317

INFORMATION ABOUT THE AUTHORS (ORCID)

Olga A. Mainovskaya — 0000-0001-8189-3071

Sergey I. Achkasov — 0000-0001-9294-5447

Alexey V. Devyatkin — 0000-0002-5712-1626

Elena V. Serykh

Vladimir V. Rybakov

Talgat Kh. Makambaev

Daria I. Suslova — 0000-0003-1106-5486

Maria A. Ryakhovskaya — 0000-0002-2243-1317

7, gij065.

7. Bandi P, et al. From detection of individual metastases to classification of lymph node status at the patient level: the camelyon17 challenge. *IEEE Trans Med Imaging*. 2018;38:550–560.

8. Щаева С.Н. Анализ выживаемости при осложненном колоректальном раке. *Онкологическая колопроктология*. 2017;7(2):20–29. DOI: [10.17650/2220-3478-2017-7-2-20-29](https://doi.org/10.17650/2220-3478-2017-7-2-20-29).

9. Fortea-Sanchis C, Martínez-Ramos D, Escrig-Sos J. CUSUM charts in the quality control of colon cancer lymph node analysis: a population-registry study. *World J Surg Oncol*. 2018 Nov 30;16(1):230.

10. Злокачественные новообразования в России в 2019 году (заболеваемость и смертность). Под редакцией Каприна А.Д., Старинского В.В., Шахзадовой А.О./М., МНИОИ им. П.А. Герцена — филиал ФГБУ «НМИЦ радиологии» Минздрава России. 2020;252 с.

11. Клинические рекомендации «Злокачественные новообразова-

ния ободочной кишки и ректосигмоидного отдела». 2020. https://cr.minzdrav.gov.ru/schema/396_1

12. Клинические рекомендации «Рак прямой кишки». 2020. https://cr.minzdrav.gov.ru/schema/554_2

REFERENCES

1. Ciresan DC, Giusti A, Gambardella LM, Schmidhuber J. Mitosis detection in breast cancer histology images with deep neural networks. *Med Image Comput Comput Assist Interv*. 2013;8150:411–418.
2. Cruz-Roa A, et al. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. *In Proc. SPIE Medical Imaging*. 2014;9041:904103.
3. Chetan L. Srinidhi et al. Deep neural network models for computational histopathology: A survey. *Medical Image Analysis*. 2020. www.elsevier.com/locate/media
4. Jeroen van der Laak et al. Deep learning in histopathology: the path to the clinic. *Nature Medicine*. May 2021;27:775–784.
5. Babak EB, Mitko V, Johannes van Diest P, et al. and the CAMELYON16 Consortium. *JAMA*. 2017;318(22):2199–2210.
6. Litjens G, et al. 1399 h&e-stained sentinel lymph node sections of breast cancer patients: the camelyon dataset. *Giga Science*. 2018; 7, giy065.
7. Bandi P, et al. From detection of individual metastases to classification of lymph node status at the patient level: the camelyon17 challenge. *IEEE Trans Med Imaging*. 2018;38:550–560.
8. Shchaeva S.N. Survival analysis in complicated colorectal cancer. *Oncological coloproctology*. 2017;7(2):20–29. (in Russ.). DOI: [10.17650/2220-3478-2017-7-2-20-29](https://doi.org/10.17650/2220-3478-2017-7-2-20-29).
9. Fortea-Sanchis C, Martínez-Ramos D, Escrig-Sos J. CUSUM charts in the quality control of colon cancer lymph node analysis: a population-registry study. *World J Surg Oncol*. 2018 Nov 30;16(1):230.
10. Malignant neoplasms in Russia in 2019 (morbidity and mortality). Edited by Kaprin A.D., Starinsky V.V., Shakhzadova A.O./M., Herzen Moscow State Medical Research Institute — branch of the Federal State Budgetary Institution “NMIC of Radiology” of the Ministry of Health of Russia. 2020; 252 p. (in Russ.).
11. Clinical recommendations “Malignant neoplasms of the colon and rectosigmoid department”. 2020. https://cr.minzdrav.gov.ru/schema/396_1 (in Russ.).
12. Clinical recommendations “Rectal cancer”. 2020. https://cr.minzdrav.gov.ru/schema/554_2 (in Russ.).