

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ ПЕРВЫЙ МОСКОВСКИЙ
ГОСУДАРСТВЕННЫЙ МЕДИЦИНСКИЙ УНИВЕРСИТЕТ ИМЕНИ
И.М. СЕЧЕНОВА МИНИСТЕРСТВА ЗДРАВООХРАНЕНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ (СЕЧЕНОВСКИЙ УНИВЕРСИТЕТ)

Иванова Елена Ильинична

На правах рукописи


**Морфологическая градация рака почки с применением программных
решений на основе искусственного интеллекта**

1.5.22. Клеточная биология

3.3.2. Патологическая анатомия

Диссертация
на соискание учёной степени
кандидата медицинских наук

Научный руководитель:

PhD
Калабрезе Витторио

Научный руководитель:

кандидат медицинских наук, доцент
Файзуллин Алексей Леонидович

Москва – 2025

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	4
ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ	18
1.1 Машинное обучение в диагностике опухолей почки.....	18
1.1.1 Архитектура нейронных сетей для гистологического анализа изображений.....	18
1.1.2 Подготовка гистологических изображений.....	20
1.1.3 Отбор тканевых или клеточных морфологических признаков	23
1.1.4 Обучение моделей для анализа гистологических изображений	25
1.1.5 Автоматизированная детекция морфологических признаков	27
1.1.6 Направления прикладного применения моделей компьютерного зрения в гистологии.....	30
1.2 Применение компьютерного зрения для диагностики рака почки	33
1.2.1 Дифференциальная диагностика различных типов почечно-клеточного рака.....	33
1.2.2 Применение компьютерного зрения для диагностики светлоклеточного почечно-клеточного рака.....	37
1.2.3 Применение компьютерного зрения для диагностики папиллярного почечно-клеточного рака.....	44
1.2.4 Применение компьютерного зрения для диагностики хромофобного почечно-клеточного рака.....	46
1.3 Использование компьютерного зрения в диагностике отторжения биоптатов трансплантированных почек	47
ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ	55
2.1 Применение компьютерного зрения для сегментации почечных структур и диагностики отторжения трансплантата почки	55
2.1.1 Подготовка набора данных	55
2.1.2 Архитектура нейронной сети и ее обучение	58
2.1.3 Оценка эффективности сегментации	60
2.1.4 Подготовка скан-изображений	60

2.1.5 Морфологический анализ.....	62
2.1.6 Статистический анализ.....	62
2.2 Применение компьютерного зрения для анализа препаратов рака почки.....	63
2.2.1 Подготовка набора данных.....	63
2.2.2 Архитектура модели сегментации.....	65
2.2.3 Аннотирование и обучение модели классификации.....	66
2.2.4 Верификация.....	69
2.2.5 Прогнозирование выживаемости.....	70
ГЛАВА 3. ПРИМЕНЕНИЕ КОМПЬЮТЕРНОГО ЗРЕНИЯ ДЛЯ СЕГМЕНТАЦИИ ПОЧЕЧНЫХ СТРУКТУР И ДИАГНОСТИКИ ОТТОРЖЕНИЯ ТРАНСПЛАНТАТОВ ПОЧКИ.....	72
3.1 Эффективность многоклассовой сегментации.....	72
3.2 Фиброз и интерстициальная инфильтрация в биоптатах трансплантатов.....	73
3.3 Сравнение модели компьютерного зрения с экспертными оценками согласно Vanff-классификации.....	75
ГЛАВА 4. ПРИМЕНЕНИЕ КОМПЬЮТЕРНОГО ЗРЕНИЯ ДЛЯ ОЦЕНКИ СТЕПЕНИ ЗЛОКАЧЕСТВЕННОСТИ РАКА ПОЧКИ.....	83
4.1 Эффективность модели классификации.....	83
4.2 Верификация модели классификации.....	85
4.3 Обнаружение опухолевых клеток в образцах тканей.....	86
4.4 Анализ выживаемости.....	88
4.5 Морфологическое описание структуры опухоли.....	90
ГЛАВА 5. ЗАКЛЮЧЕНИЕ.....	94
ВЫВОДЫ.....	99
ПРАКТИЧЕСКИЕ РЕКОМЕНДАЦИИ.....	102
СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ.....	103
СПИСОК ЛИТЕРАТУРЫ.....	105

ВВЕДЕНИЕ

Актуальность темы исследования

Почечно-клеточный рак (ПКР) — это злокачественное новообразование, происходящее из эпителиальных клеток, выстилающих канальцы почки. Данный тип опухоли составляет примерно 3,8% всех злокачественных новообразований у взрослых, при этом в последние годы отмечается устойчивый рост как заболеваемости, так и смертности, связанной с ПКР [1,2]. В Российской Федерации рак почки занимает второе место по темпам увеличения заболеваемости у мужчин (рост на 38,8%) после рака предстательной железы (135,5%) и третье место у женщин (47,7%) после опухолей головного мозга (54,9%) и полости рта (49,2%) [1]. Хотя частично такой рост объясняется совершенствованием диагностических методов, сохраняющееся увеличение числа запущенных форм свидетельствует о реальном росте заболеваемости. Согласно глобальной статистике, в 2020 году зарегистрировано 271 249 новых случаев ПКР среди мужчин (6,1 на 100 000 населения) и 160 039 среди женщин (3,2 на 100 000 населения). Для сравнения: в 2012 году этот показатель составлял 144 тыс. случаев (91 тыс. у мужчин и 53 тыс. у женщин), а к 2018 году было зафиксировано дальнейшее увеличение, с прогнозируемым ростом на 2–3% в течение следующего десятилетия [1]. По данным Американского онкологического общества, к 2024 году прогнозировалось появление 81 610 новых случаев ПКР и 14 390 смертей от данного заболевания [2].

Рак почки является морфологически и молекулярно неоднородной группой опухолей, отличающихся строением, генетическими профилями, клиническим течением и ответом на терапию. Наиболее часто встречаются три гистологических типа: светлоклеточный (сПКР), папиллярный (пПКР) и хромофобный (хрПКР) [10].

Современные принципы классификации почечно-клеточного рака основаны на интеграции морфологических и молекулярных данных. Исторически типирование опухолей осуществлялось с учётом преобладающих клеточных и архитектурных особенностей, анатомического расположения, фоновой патологии почек, гистогенетического сходства с эмбриональными структурами (например, метанефросом), а также связи с наследственными синдромами [3-6].

Современные усовершенствованные классификационные подходы служат не только инструментом диагностики почечно-клеточного рака (ПКР), но и важным элементом планирования лечения и мониторинга его эффективности [4]. Морфологическая верификация как первичного опухолевого очага, так и метастатических поражений является обязательным этапом до начала терапии, поскольку позволяет точно определить гистологический тип опухоли. Послеоперационное стадирование также имеет решающее значение для оценки риска рецидива и прогноза заболевания. У примерно двух третей пациентов опухоль ограничена пределами почки, и в этой группе пятилетняя выживаемость достигает 93%. При инвазии опухоли в окружающие ткани, соседние органы или регионарные лимфатические узлы данный показатель снижается до 72% [5]. Классификация TNM является общепринятым стандартом для стадирования рака почки, учитывающим размеры опухоли, степень инвазии (включая распространение на капсулу почки, сосуды, чашечно-лоханочную систему и фасцию Героты) и наличие метастазов в лимфатических узлах или отдалённых органах [7].

Для опухолей I–II стадий, ограниченных почечной тканью и имеющих диаметр менее 7 см, как правило, выполняется органосохраняющая резекция. При III стадии наблюдается прорастание опухоли в крупные сосуды или метастазирование в регионарные лимфоузлы, что требует радикальной нефрэктомии. IV стадия характеризуется распространением опухоли за пределы почки с вовлечением других органов или отдалённых лимфатических узлов. В таких случаях тактика лечения определяется распространённостью процесса и общим состоянием пациента; возможны комбинированные схемы, включающие

хирургическое вмешательство, иммуно- и лучевую терапию [8]. Несмотря на развитие молекулярных и инструментальных методов диагностики, морфологическое исследование остаётся ключевым звеном в оценке опухолей почки. Определённые морфологические признаки, такие как саркоматоидная дифференцировка, служат индикаторами агрессивного течения и могут использоваться для выбора персонализированных схем терапии, включая таргетные и иммунотерапевтические подходы.

Патологогистологическое исследование с применением световой микроскопии остаётся основным методом определения гистологического типа и степени злокачественности опухолей почек. Однако традиционные классификационные подходы, включая систему Фурмана, имеют ряд ограничений. Эта система основана на визуальной оценке морфологии клеточных ядер, прежде всего формы и размера ядра, а также степени выраженности ядрышек, без учёта общей морфологической структуры клетки. Согласно классификации Фурман, выделяют четыре степени злокачественности, различающиеся по степени выраженности ядрышек при разном увеличении микроскопа [9,10].

В настоящее время наибольшее распространение получила классификация Всемирной организации здравоохранения и Международного общества урологической патологии (World Health Organization/International Society of Urological Pathology, WHO/ISUP). В соответствии с ней, для опухолей 1–3-й степени основным критерием является выраженность и видимость ядрышек в ядрах клеток, тогда как 4-я степень характеризуется выраженным ядерным полиморфизмом, гигантоклеточными элементами и признаками рабдоидной или саркоматоидной дифференцировки. В отличие от классификации Фурмана, применимой исключительно к светлоклеточному почечно-клеточному раку (сПКР), система WHO/ISUP используется также для папиллярного варианта (пПКР). Валидационные исследования для хромофобного ПКР (хрПКР) не выявили достоверной связи между степенью злокачественности по Фурману или WHO/ISUP и клиническим исходом, что послужило основанием для отказа от применения этих шкал при данной гистологической форме [11,12].

Одной из существенных трудностей практической морфологической диагностики остаётся высокая вариабельность интерпретации данных между разными патологоанатомами, а также значительная трудоёмкость визуальной оценки каждого ядра опухолевых клеток. В то же время, при морфологическом анализе необходимо учитывать целый комплекс прогностически значимых факторов: стадирование по системам TNM и WHO/ISUP, гистологический тип опухоли, наличие саркоматоидной дифференцировки, сосудистой инвазии и очагов некроза. Преодолеть указанные ограничения позволяет внедрение цифровых технологий и алгоритмов искусственного интеллекта. Автоматизированные системы анализа изображений и поддержки принятия решений способны выявлять тонкие морфологические различия между клиническими группами и значительно ускорять процесс диагностики.

Степень разработанности темы исследования

Хотя патогистологическое исследование по-прежнему остаётся основным инструментом для выбора оптимальной стратегии терапии, возможности ИИ в мультидисциплинарном анализе значительно превышают возможности отдельного специалиста. Технологии компьютерного зрения позволяют автоматизировать процесс классификации рака почки и выявлять специфические диагностические паттерны на клеточном и тканевом уровнях, включая области некроза или сосудистую инвазию опухоли [13–20]. С помощью ИИ можно также различать светлоклеточный почечно-клеточный рак и транслокационный почечно-клеточный рак, распознавая клетки с признаками хромосомных аномалий на гематоксилин-эозиновых препаратах, что позволяет существенно сократить затраты на сложные молекулярные и генетические исследования [21]. Автоматизированные системы классификации WHO/ISUP и Фурмана доказали свою эффективность в различении низкой (1–2 степень) и высокой (3–4 степень) злокачественности [22–24], однако их

прогностическая ценность ограничена из-за дискретной природы шкал, созданных с учётом практической применимости для патологоанатомов.

Развитие молекулярной биологии существенно расширило возможности традиционной световой микроскопии и способствовало появлению более комплексных систем классификации ПКР. Исследование молекулярных механизмов опухоли, в частности с применением секвенирования нового поколения (NGS — Next-Generation Sequencing), повышает точность диагностики и прогнозирования исходов. Анализ паттернов генной экспрессии помогает различать нормальную и опухолевую ткань, определять подтипы и стадию опухоли, выявлять ключевые биологические механизмы и потенциальные терапевтические мишени. Методы ИИ позволяют прогнозировать течение заболевания на основе выявленных генетических дефектов в опухолевой ткани, например, идентифицируя группы генов с коррелированной экспрессией и выбирая репрезентативный ген для каждой группы [25].

Таким образом, точная классификация и стадирование ПКР являются критически важными для эффективного ведения пациентов и планирования терапии. Интеграция молекулярных данных, таких как радиогеномика, с традиционным морфологическим анализом улучшает существующие системы классификации и повышает эффективность лечения пациентов с раком почки [26,27]. Среди наиболее перспективных направлений применения ИИ в медицине — прогнозная аналитика, диагностическая визуализация и поддержка клинических решений.

Цель и задачи исследования

Установить диагностическую и предиктивную значимость гистологических биомаркеров острого клеточного отторжения трансплантата почки и

светлоклеточного почечно-клеточного рака, рассчитанных на основе сегментации и классификации клеток и тканевых структур моделями компьютерного зрения.

Для достижения цели были поставлены следующие задачи:

- 1) Разработать оригинальные модели цифрового анализа изображений биоптатов трансплантированных почек и светлоклеточного почечно-клеточного рака и установить точность их работы.
- 2) Установить диагностические пороговые значения относительных площадей регионов фиброза и интерстициальной инфильтрации, полученных с помощью цифрового анализа изображений, для совершенствования клинικο-морфологической классификации острого клеточного отторжения трансплантата почки.
- 3) Установить количественные метрики относительного количества опухолевых клеток с видимыми ядрышками, полученные с помощью цифрового анализа изображений, для градации светлоклеточного почечно-клеточного рака почки на основе системы WHO/ISUP.
- 4) Установить прогностические группы и рассчитать среднюю ожидаемую продолжительность жизни пациентов со светлоклеточным почечно-клеточным раком почки на основе результатов цифрового анализа изображений и данных о выживаемости.
- 5) Провести сравнительный морфологический анализ и предложить описание тканевых паттернов прогностических групп пациентов со светлоклеточным почечно-клеточным раком, выделенных на основе анализа выживаемости, для совершенствования клинικο-морфологических классификаций злокачественных опухолей почки.

Научная новизна

- 1) Было предложено использование алгоритмов машинного обучения для анализа гистологических изображений почечной ткани, сочетающие сегментацию структур

с их классификацией, что позволяет подсчитывать точное количество структур в скан-изображении гистологического препарата почки.

2) Были валидированы разработанные методы автоматического определения ключевых морфологических признаков острого отторжения трансплантата почки (интерстициальной инфильтрации и фиброза), что способствует устранению вариабельности интерпретации результатов между экспертами.

3) Проведена валидация разработанной мультимодальной предиктивной модели, способная одновременно классифицировать опухолевые клетки в ткани светлоклеточного почечно-клеточного рака, предсказывать их степень злокачественности и выживаемость пациента на основании морфологических и клинических данных.

4) Было выделено четыре паттерна гистологического строения светлоклеточного рака, основанных на анализе моделью компьютерного зрения клеточного состава опухоли вкупе с данными об общей выживаемости пациентов после постановки диагноза.

Теоретическая и практическая значимость работы

Результаты данной работы подтверждают возможности применения моделей компьютерного зрения для анализа ткани почечных биоптатов при диагностике острого отторжения трансплантата почки, а также для оценки степени злокачественности светлоклеточного почечно-клеточного рака.

В этом исследовании представлены результаты работы модели компьютерного зрения, которая не только идентифицирует структуры тканей и предсказывает прогноз, но и предоставляет количественные альтернативы двум наиболее эффективным показателям — фиброзу и интерстициальной инфильтрации, — необходимыми для оценки в системе классификации Banff для оценки отторжения биоптатов почки. Эти морфологические данные имеют решающее значение для дифференциации между пограничными изменениями и

отторжением. В данном исследовании продемонстрировано, что автоматизация может решать проблему изменчивости между наблюдателями, что потенциально приведет к более персонализированному лечению и улучшению результатов лечения пациентов.

Используя компьютерное зрение для анализа тонких морфологических особенностей, в данной работе предложены более точные критерии классификации светлоклеточного почечно-клеточного рака с потенциалом для включения их при будущих пересмотрах в классификацию WHO/ISUP. Было исследовано, могут ли рассчитанные показатели служить гистологическим биомаркером, вычисляемым с помощью модели компьютерного зрения, и способным предсказывать экспертные оценки в рамках текущей классификации WHO/ISUP. Кроме того, была оценена прогностическая ценность биомаркера в отношении результатов выживаемости пациентов.

Данная работа устраняет разрыв между устоявшейся клинической системой и анализом на основе данных, демонстрируя необходимость таких моделей для преобразования текущей клинической практики.

Методология и методы исследования

Исследование выполнено с использованием современных методов компьютерного зрения, машинного обучения и цифровой патологии. Основные этапы работы включали сбор данных, их предобработку, валидацию разработанных алгоритмов анализа гистологических изображений, а также биоинформационный анализ для изучения роли морфологических и клинических параметров в диагностике и прогнозировании заболеваний.

Объектами исследования являлись гистологические изображения биоптатов почечных трансплантатов и светлоклеточного почечно-клеточного рака,

полученные из патологоанатомических отделений нескольких учреждений и открытых баз данных (TCGA).

Гистологическая обработка тканей осуществлялась стандартными методами, включая окрашивание гематоксилином и эозином. Далее гистологические препараты оцифровывались с использованием гистологических сканеров высокого разрешения.

Была оценена точность работы моделей глубокого обучения для автоматической сегментации и классификации структур почечной ткани, оценки площади фиброза и интерстициальной инфильтрации, входящих в классификацию Banff для оценки отторжения почки, а также клеточного состава опухоли при светлоклеточном почечно-клеточном раке. Выполнена интеграция морфологических и клинических данных для оценки выживаемости при светлоклеточном почечно-клеточном раке. Работа алгоритмов проверялась на независимых тестовых наборах данных, также сравнивались результаты автоматического и ручного анализа с использованием экспертной оценки.

Для проверки статистической значимости результатов использовались методы параметрического и непараметрического анализа, включая ANOVA, тесты Манна-Уитни и логистическую регрессию. Для анализа выживаемости был использован метод Каплана-Майера с логарифмическим критерием (тест Мантеля-Кокса). Для оценки точности работы моделей использовались метрики AUC-ROC, точность, полнота, F1-мера, коэффициент Дайса.

Личный вклад автора

Автор провел сбор исследовательского материала в архивах лабораторий и патологоанатомических отделений крупных стационаров, а также использовал открытые базы данных онкологических пациентов (TCGA). Он выполнял изучение гистологических препаратов при световой микроскопии и провел их детальный морфологический анализ. Автором были проаннотированы структуры на скан-

изображениях гистологических препаратов, проведена валидация моделей ИИ для сегментации и классификации, выполнен статистический анализ и визуализация данных, а также оценена предсказательная способность моделей. Он освоил применяемые в исследовании методы, сформулировал выводы, практические рекомендации и положения, выносимые на защиту. Результаты работы нашли отражение в научных публикациях и были представлены на научно-практических конференциях.

Положения, выносимые на защиту

- 1) Рассчитанные моделью компьютерного зрения площади участков интерстициальной инфильтрации и фиброза в трансплантированных почках позволяют с точностью более 97% отличить острое клеточное отторжение от пограничных изменений.
- 2) Рассчитанные моделью компьютерного зрения абсолютные и относительные метрики клеток с выраженными ядрышками в раке почки позволяют с точностью более 79% различать высокодифференцированные опухоли от низкодифференцированных.
- 3) Определен порог прогностической значимости клеток с выраженными ядрышками (11%) при раке почки, при превышении которого ожидаемая продолжительность жизни пациентов снижается с более чем 6 до 2,2 лет.

Соответствие диссертации паспорту научной специальности

Диссертационное исследование соответствует паспортам научных специальностей 1.5.22. Клеточная биология, 3.3.2. Патологическая анатомия; формулам специальности 1.5.22. Клеточная биология, гистология, цитология, клеточная биология – область науки, занимающаяся исследованием

происхождения, строения, развития, функционирования клеток и тканей, их взаимодействия в процессе жизнедеятельности организма как в норме, так и при различных патологических нарушениях. Области исследований: Изучение закономерностей цито- и гистогенеза, строения и функции клеток и тканей. Изучение закономерностей дифференцировки клеток и тканей, их физиологической регенерации и регуляции этих процессов, а также дифференцировки и жизнедеятельности недифференцированных клеток. Системный анализ взаимоотношений клеток, тканей и функциональных систем организмов – представителей всех царств. Сравнительно-эволюционное изучение тканевых элементов в связи с проблемой происхождения и филогенетического развития тканей. Исследование адаптации тканевых элементов к действию различных биологических, физических, химических и других факторов. Молекулярные, иммунологические и физиологические аспекты изучения клеток многоклеточных, малоклеточных и одноклеточных организмов в норме и патологии. Разработка экспериментальных моделей, методов цитологической диагностики, морфометрии, маркерной гисто- и цитохимии и др. 3.3.2. Патологическая анатомия – специальность, на современном этапе занимающаяся решением проблем теории (общая и частная патология человека и животных) и практики (диагноз, эффективность лечения, прогноз заболевания, ятрогении и др.) медицины. Объектом ее внимания являются клетки, ткани и органы, исследуемые прижизненно (биопсийный и операционный материал) и постмортально (аутопсийный материал). Методические приемы, используемые в современных патологоанатомических исследованиях: темнопольная, фазово-контрастная, поляризационная, люминесцентная и электронная микроскопия, гистохимия, гистоэнзимохимия, иммуноморфология, морфометрия, математический анализ и др. Значимость решения научных и технических проблем данной специальности для народного хозяйства состоит в совершенствовании диагностических мероприятий, создании теоретической и практической базы для разработки новых средств профилактики и терапии болезней.

Степень достоверности и апробация результатов

Достоверность полученных результатов определяется использованием достаточного количества образцов (334 скан-изображений препаратов, из которых 120 использовалось для первой части исследования, а 214 – для второй), полученных из 4-х независимых источников (открытых баз данных, лабораторий и патологоанатомических отделений стационаров), что является достаточным для получения достоверных результатов. Сканы гистологических препаратов были зарегистрированы в качестве баз данных (Тимашев П.С., Демура Т.А., Файзуллин А.Л., Шехтер А.Б., Бакулина О.А., Иванова Е.И., Валиева Я.М., Сержникова Н.Б. SECHENOV ATLAS OF CELL IMAGES – CLEAR CELL RENAL CELL CARCINOMA (SACI – CCRCC). Свидетельство о регистрации базы данных 2023622869, 29.08.2023. Заявка № 2023622635, от 11.08.2023; Тимашев П.С., Демура Т.А., Файзуллин А.Л., Шехтер А.Б., Бакулина О.А., Иванова Е.И., Валиева Я.М., Сержникова Н.Б. SECHENOV ATLAS OF TISSUE IMAGES – CLEAR CELL RENAL CELL CARCINOMA (SATI – CCRCC). Свидетельство о регистрации базы данных 2023622870, 21.08.2023. Заявка № 2023622651, от 11.08.2023). Применяемые методы исследования полностью соответствуют поставленным задачам, а изложенные положения, выводы и рекомендации являются обоснованными. Достоверность результатов подтверждается применением известных морфологических методов исследования с использованием сертифицированного научного оборудования и строгим соблюдением методик, современных методов количественной статистики и сертифицированных программных пакетов. Результаты проведенного исследования использованы в учебной работе в Институте Регенеративной Медицины ФГАОУ ВО Первый МГМУ имени И.М. Сеченова Минздрава России (Сеченовский Университет).

Апробация диссертации состоялась на заседании Института регенеративной медицины Научно-технологического парка биомедицины ФГАОУ ВО Первый

МГМУ имени И.М. Сеченова (Сеченовский Университет) протокол № 13 от 21.10.2025 г.

Основные положения и результаты диссертационной работы докладывались и обсуждались на следующих конференциях:

- 1) Применение искусственного интеллекта в анализе скан-изображений гистологических препаратов почки / Вычислительная биология и искусственный интеллект в персонализированной медицине (г. Москва, 9–11 августа 2024 г.).
- 2) Применение искусственного интеллекта для диагностики отторжения трансплантата почки / VII Сеченовский международный биомедицинский саммит: мегатренды в биомедицине (SIBS-2023) (г. Москва, 8–9 ноября 2023 г.).
- 3) Применение программных решений на основе искусственного интеллекта для морфологической градации рака почки / Цифровое морфологическое пространство. Сегодня и завтра (г. Москва, 18 апреля 2024 г.).
- 4) Инновационная классификация светлоклеточного рака почки: мультимодальный анализ характеристик опухоли с помощью искусственного интеллекта / Вычислительная биология и искусственный интеллект в персонализированной медицине, (г. Москва, 7-9 августа 2024 г.).
- 5) Новый подход к морфологической градации рака почки при помощи программных решений на основе искусственного интеллекта / IV Российский Диагностический Саммит (г. Москва, 2-4 октября 2024 г.).
- 6) Применение программных решений на основе искусственного интеллекта для морфологической градации рака почки / Финал конкурса НИР «IX Всероссийский Молодежный Научный Форум Наука Будущего — Наука Молодых», (г. Самара, 28 октября – 1 ноября 2024 г.).
- 7) Новый подход к морфологической градации рака почки при помощи программных решений на основе искусственного интеллекта / VIII Сеченовский Международный биомедицинский саммит (SIBS-2024), (г. Москва, 6-7 ноября 2024 г.)
- 8) Морфологическая градация рака почки с применением программных решений на основе искусственного интеллекта / V Международный саммит по цифровой

медицине и информационным технологиям в здравоохранении «Цифроайтимед» (г. Москва, 1 октября 2025 г.).

Публикации по теме диссертации

По результатам исследования автором опубликовано 13 работ, в том числе 1 научная статья в журнале, включенном в Перечень рецензируемых научных изданий Сеченовского Университета / Перечень ВАК при Минобрнауки России, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук; 2 статьи в изданиях, индексируемых в международных базах Scopus, PubMed, 4 иные публикации по результатам исследования, 6 публикаций в сборниках материалов международных и всероссийских научных конференций.

Структура и объем диссертации

Диссертация выполнена в формате монографии, занимает 116 страниц машинописного текста, включает 5 таблиц и 19 рисунков. Структурно работа состоит из введения, четырёх глав, заключения, выводов и списка литературы, включающего 9 работ отечественных и 83 работы зарубежных авторов.

ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ

1.1 Машинное обучение в диагностике опухолей почки

1.1.1 Архитектура нейронных сетей для гистологического анализа изображений

Методы машинного обучения традиционно подразделяются на контролируемые и неконтролируемые, или, иначе говоря, обучение с учителем и без учителя. В первом случае алгоритмы используют размеченные данные — то есть заранее известные правильные ответы — для извлечения признаков и построения моделей, решающих задачи классификации или регрессии. Классификация направлена на отнесение входных данных к определённым категориям, в то время как регрессия используется для прогнозирования непрерывных величин, например, показателей выживаемости. В противоположность этому, обучение без учителя — это тип машинного обучения, при котором алгоритм работает с неразмеченными данными, то есть без заранее известных правильных ответов. Целью является выявление скрытых закономерностей, кластеров или структур в данных. Примеры таких методов — кластеризация (k-means, DBSCAN) и снижение размерности (PCA, t-SNE) [28].

Искусственные нейронные сети (НС) представляют собой один из наиболее известных и широко применяемых инструментов искусственного интеллекта. НС — это модель машинного обучения, вдохновлённая принципами работы человеческого мозга. Она состоит из множества взаимосвязанных узлов — искусственных нейронов, организованных в слои (входной, скрытые и выходной). Каждый нейрон получает данные, преобразует их с помощью весов и функции активации, а затем передаёт результат дальше. Благодаря многослойной структуре нейронные сети способны распознавать сложные закономерности и извлекать признаки из данных, что делает их особенно эффективными при анализе

изображений. Если сеть содержит множество последовательных слоёв, её относят к категории глубоких нейронных сетей. Для обучения таких моделей данные, как правило, разделяют на несколько подмножеств: обучающее, валидационное и тестовое, что позволяет оптимизировать параметры сети и оценить её способность к обобщению [28].

Алгоритм компьютерного анализа гистологических изображений почечно-клеточного рака (ПКР) с использованием нейронных сетей включает несколько последовательных этапов: отбор цифровых скан-изображений, выделение областей интереса, сегментацию и аннотацию клеточных структур, извлечение морфологических признаков, применение классификаторов и последующее прогнозирование. Подробное описание этих этапов представлено далее (Рисунок 1).

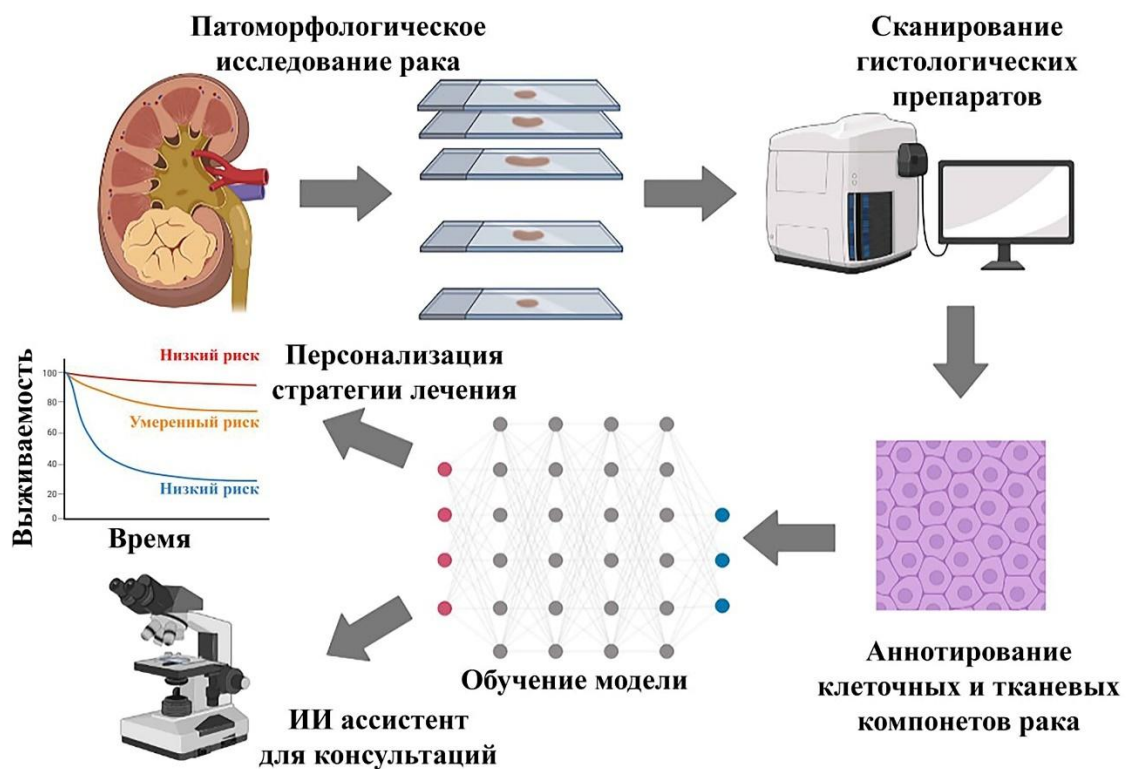


Рисунок 1 – Типичный рутинный процесс гистологического анализа изображений рака почки с помощью технологий компьютерного зрения

1.1.2 Подготовка гистологических изображений

Для обучения моделей искусственного интеллекта (ИИ) применяются цифровые изображения гистологических препаратов, или WSI (Whole Slide Images) которые можно увеличивать, перемещать и аннотировать на экране компьютера так же, как при работе с традиционным микроскопом. Этот подход, известный как виртуальная микроскопия, подразумевает полное сканирование предметного стекла и получение цифрового файла с высоким разрешением, который воспроизводит микроскопическую картину ткани. Репозитории содержат изображения, сопровождаемые клиническими, морфологическими и молекулярными аннотациями, что позволяет использовать их для задач классификации, сегментации и прогнозирования. Среди наиболее известных открытых источников — The Cancer Genome Atlas (TCGA), The Cancer Imaging Archive (TCIA), Human Protein Atlas и PAIP Challenge Dataset. Эти ресурсы предоставляют доступ к тысячам оцифрованных препаратов различных типов опухолей, включая ПКР, и служат основой для обучения и валидации моделей машинного обучения [22].

После получения WSI проводится выбор характерных участков ткани — областей интереса (ROI, Region of Interest), которые служат для последующего анализа. Обычно выделение ROI выполняется несколькими опытными патологоанатомами независимо друг от друга, что повышает точность и воспроизводимость данных. Размер выбранных участков, как правило, варьирует от 1000 до 8000 мкм в периметре в зависимости от исследуемой задачи. На следующем этапе из ROI извлекаются изображения клеток, которые аннотируются специалистами с указанием их принадлежности к определённым морфологическим классам. Выделение ROI позволяет сократить объём данных для обработки и сфокусировать вычислительные ресурсы на наиболее значимых фрагментах скана, что повышает точность и эффективность анализа при разработке и обучении нейронных сетей [23].

Далее проводится сегментация аннотированных структур — это процесс разделения цифрового изображения на отдельные области или объекты, обладающие схожими визуальными характеристиками (например, цветом, текстурой или формой). В цифровой патологии сегментация применяется для автоматического выделения клеток, ядер, тканей или структур опухоли на гистологических изображениях. Она может выполняться на разных уровнях — от сегментации отдельных ядер и клеток до разграничения опухолевой и нормальной ткани. Методы сегментации подразделяются на классические алгоритмы извлечения признаков и современные подходы, основанные на глубоких сверточных НС (CNN, Convolutional Neural Networks). К классическим методам относится пороговая сегментация (thresholding) – это простой подход, при котором изображение разделяется на объекты и фон на основании значения яркости или цвета пикселей. Этот метод эффективен при хорошем контрасте между тканями, но чувствителен к шумам и неравномерному окрашиванию. Методы на основе границ (edge-based segmentation) основаны на поиске резких переходов интенсивности (границ), которые соответствуют контурам структур. Метод хорошо выделяет очертания клеток, но может давать разрывы контуров при слабовыраженных границах. Региональные методы (region-based segmentation) включают подходы типа «распространения области» (region growing), где сегментация начинается с точки — «затравки», и затем область постепенно расширяется, пока соседние пиксели удовлетворяют заданным критериям однородности. Такие методы часто применяются для выделения опухолевых зон. Современные методики основаны на использовании сверточных НС, среди которых наиболее популярны U-Net, Mask R-CNN, DeepLab и их модификации. Они обучаются на размеченных данных и способны автоматически выделять сложные структуры, включая клеточные ядра и ткани, с высокой точностью [29].

В контексте анализа изображений почечно-клеточного рака сегментация часто реализуется поэтапно. На первом этапе может выполняться адаптивное пороговое выделение ядер в цветовой модели HSV (Hue, Saturation, Value), позволяющее отделить клетки от фона. Затем применяется метод маркер-

контролируемой сегментации WaterShed, обеспечивающий корректное разделение соприкасающихся и перекрывающихся ядер [22]. Для повышения точности выделения границ клеток может использоваться вейвлет-преобразование (wavelet transform), которое уменьшает шум и улучшает четкость контуров. Вейвлет-преобразование — это математический метод анализа сигналов и изображений, позволяющий разложить их на компоненты, содержащие информацию о частотах и локализации этих частот во времени или пространстве. Изображение (или сигнал) представляется в виде суммы базовых функций — вейвлетов, каждая из которых характеризуется масштабом (размером) и положением. Таким образом, можно выделить области с различной степенью детализации. Вейвлет-преобразование часто используется для извлечения текстурных признаков ткани (например, различия между опухолевыми и нормальными структурами), сжатия изображений без значительной потери качества, фильтрации шумов и улучшения контрастности микроскопических изображений, предварительной обработки данных перед анализом с помощью методов машинного обучения [29].

Перед сегментацией часто выполняют предварительную обработку изображений, включающую масштабирование и нормализацию яркости, реконструкцию изображения, устранение артефактов и шумов, коррекцию цвета, контрастирование и выравнивание освещённости, деконволюцию цвета. Деконволюция позволяет разделить вклад каждого красителя — например, изолировать канал гематоксилина, отражающий ядра клеток, и канал эозина, характеризующий цитоплазму и стромальные компоненты. После завершения сегментации изображения делятся на небольшие фрагменты — так называемые «патчи», содержащие отдельные клеточные структуры, которые используются для обучения и валидации ИИ-моделей [30].

1.1.3 Отбор тканевых или клеточных морфологических признаков

Сегментированные морфологические структуры могут быть описаны с помощью количественных параметров, называемых дескрипторами. Дескрипторы — это цифровое представление формы, структуры, цвета и текстуры биологических элементов, которое позволяет алгоритмам машинного обучения анализировать изображения так же, как это делает патологоанатом, но в числовом виде. Для извлечения максимально возможного числа информативных дескрипторов из изображений применяются различные методы анализа признаков. Выделяют несколько основных групп дескрипторов: геометрические (формальные), цветовые, текстурные, статистические и структурные показатели. В контексте машинного обучения, направленного на определение степени злокачественности опухоли, ключевыми характеристиками ядер клеток выступают их размер, форма, текстура и цвет.

Цветовые параметры могут анализироваться в различных цветовых пространствах — RGB (red, green, blue), HSV (hue, saturation, value), LAB color space, а также в канале гематоксилина при деконволюции красителей. К числу количественных характеристик, описывающих цвет, относятся такие статистические показатели, как среднее значение, стандартное отклонение, медиана, асимметрия и эксцесс [22,31]. Для оценки геометрии ядер используют показатели площади, длины большой и малой осей, периметра, выпуклой площади и среднего диаметра [29]. Эти признаки позволяют количественно описать вариации формы клеточных структур, характерные для различных степеней дифференцировки опухоли. Тектурные особенности ядер описываются с помощью целого ряда статистических и структурных показателей, среди которых — энергия, корреляция Харалика, суммарная дисперсия, момент обратной разницы, энтропия, инерция, суммарная энтропия, контраст и показатели корреляционной информации [22,29]. Анализ текстуры основывается на изучении пространственного распределения уровней серого, отражающего микроструктуру

ткани и локализацию клеточных маркеров в ядрах и цитоплазме. Для этого используются специализированные матричные методы: матрица совпадений уровней серого (GLCM - Gray-Level Co-Occurrence Matrix) описывает взаиморасположение пикселей с определённой интенсивностью, а матрица длины пробега уровней серого (GLRLM - Grey-Level Run Length Matrix) характеризует последовательности одинаковых интенсивностей в заданном направлении, что позволяет выявить структурные особенности и неоднородность текстуры (Рисунок 2) [29].

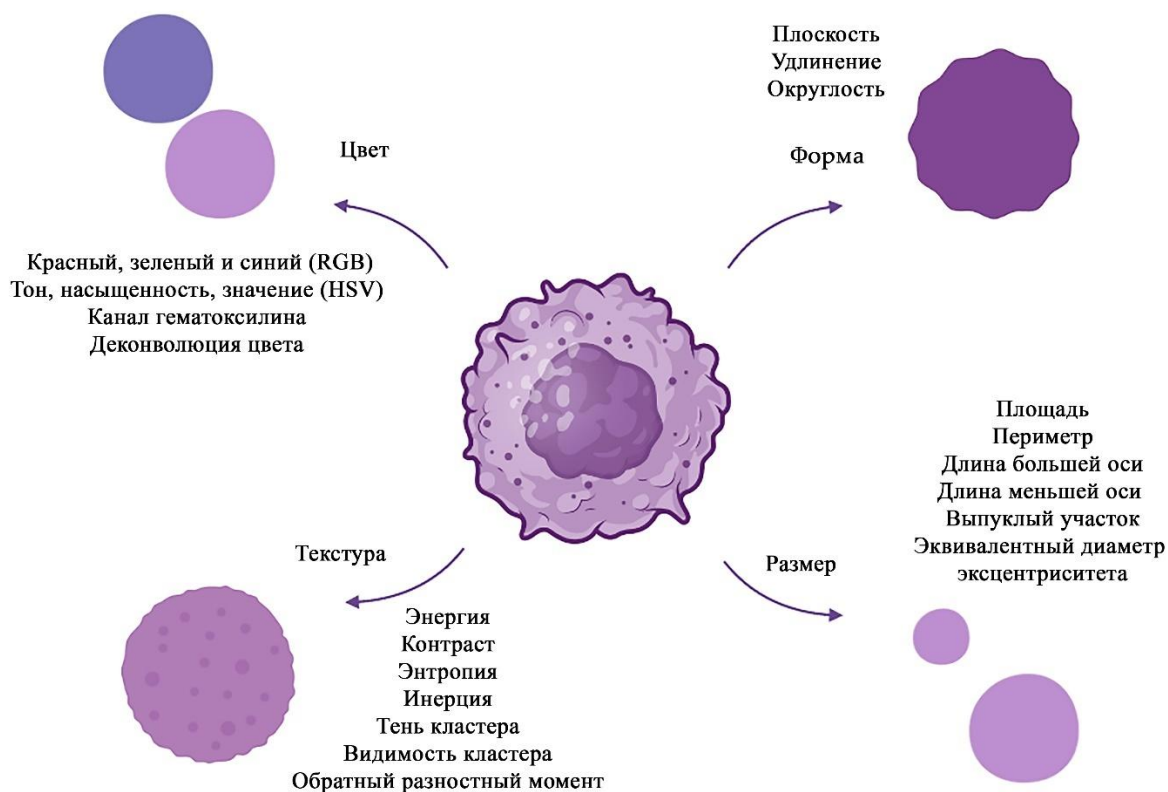


Рисунок 2 – Дескрипторы ядра опухолевой клетки

Далее количественные показатели морфологических признаков подвергаются обработке с использованием технологий искусственного интеллекта. На начальном этапе число дескрипторов может достигать от нескольких десятков до нескольких тысяч. После этого выполняется сокращение их до ограниченного

набора (обычно 5–10) оптимальных морфологических признаков, которые затем применяются в качестве входных данных для классификаторов. Для выбора наиболее значимых дескрипторов применяются различные методы, включая: анализ главных компонент, оценку корреляции между признаками; корреляцию между признаками и классами; ранжирование признаков с использованием методов линейных или нелинейных опорных векторов (SVM – Support Vector Machine); вычисление средних значений и дисперсий признаков для различных классов с последующим объединением в общий показатель качества; генетические алгоритмы; оптимизацию роя бинарных частиц; метод случайного леса (RF – Random Forest) и другие. Каждый из подходов характеризуется определенным набором анализируемых параметров. Чтобы получить более полную и комплексную оценку морфологических характеристик, обычно используют несколько алгоритмов, объединенных в ансамбль [29].

1.1.4 Обучение моделей для анализа гистологических изображений

Для решения задач классификации чаще всего применяются методы опорных векторов и случайного леса. Метод случайного леса позволяет оценивать значимость признаков при классификации классов. Важность признака определяется путем перестановки его значений и последующего измерения увеличения ошибки классификации по сравнению с исходными данными. Метод случайного леса (Random Forest, RF) — это ансамблевый алгоритм машинного обучения, основанный на построении множества решающих деревьев для задач классификации и регрессии. Принцип работы заключается в нескольких шагах: для каждого дерева выбирается случайная подвыборка данных из исходного обучающего набора (с повторениями — бутстрэп). Далее при каждом разбиении узла дерева выбирается случайный поднабор признаков, и выбирается оптимальный признак для разделения данных и затем каждое дерево делает свой прогноз. Для классификации результат определяется голосованием большинства

деревьев, для регрессии — усреднением предсказаний [29]. Метод опорных векторов (SVM, Support Vector Machine) — это алгоритм машинного обучения, используемый для задач классификации и регрессии. Его основная цель — найти оптимальную разделяющую гиперплоскость между классами данных в многомерном пространстве признаков. Каждое наблюдение представляется в виде вектора признаков, далее алгоритм ищет гиперплоскость, которая максимизирует зазор (margin) между ближайшими точками разных классов — эти точки называются опорными векторами. В случае, если классы не линейно разделимы, данные могут быть преобразованы в пространство большей размерности с помощью ядерных функций (например, RBF, полиномиальное ядро), чтобы гиперплоскость могла разделять классы. После обучения модели проходят тестирование, а их эффективность оценивается с использованием 10-кратной перекрестной проверки (10-кратного CV). Критериями оценки служат диагностическая точность, чувствительность, специфичность, а также положительная и отрицательная прогностические ценности [29].

Сверточные нейронные сети (CNN) применяются для анализа гистологических изображений с целью распознавания паттернов и являются частью технологий глубокого обучения. Для более точной визуализации могут использоваться гибридные архитектуры, объединяющие CNN с полносвязными нейронными сетями или сетями выходного уровня [32]. При работе с новым изображением CNN сначала применяет свертку с использованием различных фильтров, которые систематически сканируют изображение для выявления характерных паттернов, таких как края или кривые. Для каждого фильтра формируется карта признаков, получаемая как скалярное произведение между фильтром и исходным изображением. На следующем этапе проводится операция подвыборки (pooling), уменьшающая соседние группы пикселей до одного пикселя с максимальным значением. Подвыборка служит для уменьшения размера изображения, снижения вычислительной нагрузки и создания усредненных представлений деталей. В итоге многомерное изображение преобразуется в вектор, пригодный для дальнейшей обработки. В гистологическом анализе подвыборка

помогает выделять ключевые морфологические особенности клеток и тканей, сохраняя важную информацию при уменьшении объёма данных для обработки

Во время обучения модель корректирует параметры, минимизируя функцию потерь, измеряемую как перекрестная энтропия между оценкой патологоанатома и прогнозом модели. Обучение нейронной сети представляет собой процесс настройки её весов и параметров для минимизации ошибки между предсказаниями модели и известными правильными ответами. Данные проходят через сеть в прямом направлении, формируя прогноз, после чего вычисляется функция потерь, отражающая расхождение между предсказанным и реальным результатом. Затем с помощью метода обратного распространения ошибки и алгоритмов оптимизации, таких как градиентный спуск, веса сети корректируются. Один полный проход всех обучающих данных через сеть называется эпохой обучения, и процесс обычно повторяется многократно, с периодической проверкой на валидационном наборе, чтобы контролировать способность модели к обобщению и предотвращать переобучение. Обучение проводится многократно на всех данных, каждая итерация называется эпохой. Если после нескольких эпох ошибок не выявляется, модель начинает формировать прогнозы на основе обучающих данных. Валидация обеспечивает дополнительную настройку параметров и улучшение показателей работы модели. Обучение завершается, когда производительность модели на валидации перестает улучшаться [33].

1.1.5 Автоматизированная детекция морфологических признаков

Нейронные сети могут применяться для автоматизации извлечения ROI с опухолевой тканью из каждого WSI. Автоматизация ROI на гистологических изображениях предполагает использование алгоритмов искусственного интеллекта для предварительного обнаружения участков ткани, наиболее информативных для анализа. После автоматического отбора ROI патологоанатом проверяет и при необходимости корректирует результаты, удаляя случайно включённые фрагменты

или артефакты. Такая автоматизация позволяет ускорить процесс анализа больших наборов WSI, снижает субъективность оценки и обеспечивает более последовательное выделение областей с опухолевой тканью, зонами некроза или иммунной инфильтрацией, которые важны для диагностики и прогноза ПКР. При обучении ИИ-модели, направленной на различение нормальной и опухолевой ткани при диагностике ПКР, необходимо учитывать как ткань самого образования, так и различные типы окружающих тканей, включая интерстиций почки и паренхиматозные структуры. Обучающие наборы данных в таких случаях обычно включают ROI одинакового разрешения и увеличения с зонами иммунной инфильтрации, полями некроза, нормальной и дистрофической почечной паренхимой и другими структурами [34].

Для машинного обучения, направленного на оценку топографии ПКР, требуется обучение модели распознавать каждый тип ткани. Разработка набора алгоритмов для такой классификации является сложной задачей, поскольку микроокружение опухоли содержит множество различных клеточных типов, каждый из которых требует специфического набора признаков для распознавания. В связи с этим исследователи нередко используют неконтролируемое обучение, например автоэнкодеры, для классификации клеток на основе их морфологических характеристик. Автоэнкодеры состоят из двух основных частей: кодировщика, который сжимает входные данные в компактное представление (латентное пространство), и декодировщика, который восстанавливает исходные данные из этого представления. В контексте анализа гистологических изображений автоэнкодеры позволяют выявлять скрытые паттерны и особенности клеток, классифицировать различные типы тканей и уменьшать размерность данных, что облегчает последующую обработку и обучение моделей машинного обучения. Многоуровневый разреженный автоэнкодер (SSAE – stacked sparse autoencoder) представляет собой нейронную сеть, состоящую из нескольких слоев разреженных автоэнкодеров (SAE), где выход каждого слоя служит входом для последующего [30].

Факторы, такие как васкуляризация, наличие некроза и прорастание опухоли в соседние ткани (почечная капсула, фасция Героты), могут влиять на распространение опухоли и прогноз заболевания. Современные инструменты машинного обучения открывают новые возможности для количественной оценки клеточного состава и пространственной организации опухолевой ткани. Изучение гетерогенности опухоли с помощью ИИ предполагает количественную оценку разнообразия клеточных типов и структурных особенностей внутри опухолевой ткани. Современные модели ИИ способны анализировать микроскопические изображения и выявлять различия в морфологии, васкуляризации, наличии очагов некроза и составе опухолевого микроокружения, включая иммунные клетки, фибробласты и эндотелиоциты. Такой подход позволяет оценивать пространственное распределение клеток, степень инфильтрации иммунными элементами и архитектурную неоднородность ткани, что важно для прогнозирования агрессивности опухоли и ответа на терапию. Использование ИИ обеспечивает более точную, воспроизводимую и масштабируемую характеристику опухолевой гетерогенности по сравнению с традиционной визуальной оценкой. При этом гетерогенность микроокружения может способствовать резистентности к системной терапии [35].

Анализ иммуногистохимических препаратов с помощью ИИ позволяет количественно и качественно оценивать присутствие и распределение различных клеточных популяций и маркеров внутри опухолевой ткани. Модели ИИ могут автоматически распознавать положительно окрашенные клетки, различать типы иммунных клеток (например, $CD3^+$ Т-лимфоциты, $CD8^+$ цитотоксические Т-клетки, $CD20^+$ В-лимфоциты, макрофаги $CD68^+/CD163^+$) и измерять их плотность и пространственное расположение. Для более точного выявления опухолевых сосудов ИИ-модель может обучаться на скан-изображениях иммуногистохимических препаратов с антителами к сосудистому эндотелию ($CD31$, $CD34$) [16,17]. Кроме того, ИИ позволяет оценивать другие маркеры, включая PD-L1 (иммуносупрессия), Ki-67 (пролиферация), PanCK (эпителиальный маркер) и виментин (мезенхимальный маркер). Использование ИИ в этом контексте

обеспечивает высокую точность, воспроизводимость и масштабируемость анализа, снижая субъективность визуальной оценки и позволяя интегрировать данные в прогностические модели и планирование персонализированного лечения [36].

1.1.6 Направления прикладного применения моделей компьютерного зрения в гистологии

После этапов обучения и валидации модели машинного обучения способны прогнозировать выживаемость пациентов с учетом клинических и эпидемиологических данных, таких как пол, возраст и стадия опухоли по системе TNM. Для выбора наиболее информативных признаков часто применяется регрессия LASSO (Least Absolute Shrinkage and Selection Operator), которая сочетает отбор объясняющих переменных и регуляризацию для снижения дисперсии. Этот подход сочетает линейную регрессию с регуляризацией L1, которая штрафует величину коэффициентов признаков, приводя к обнулению наименее информативных параметров. В результате LASSO автоматически выделяет набор признаков, оказывающих наибольшее влияние на прогнозируемую переменную, например, выживаемость пациента или риск рецидива опухоли. Метод особенно полезен при анализе больших наборов данных с высокой размерностью, позволяя создавать более интерпретируемые и стабильные модели, а также снижать шум и избыточную информацию в исходных признаках. Прогнозирование общей выживаемости, риска рецидива и других исходов имеет важное значение для разработки персонализированных планов лечения и организации дальнейшего наблюдения за пациентом [19].

Еще одной перспективной задачей является использование предварительно обученных нейронных сетей для диагностики различных типов опухолей. Эти модели уже обучены выявлять общие паттерны, такие как текстура ткани, форма ядер и структурная организация клеток, что позволяет переносить их на новые

задачи без необходимости обучения с нуля. Возможность переносить ранее полученные знания на новые случаи представляет собой важный инструмент для поддержки клинических решений. Поскольку биологическое поведение и морфологические признаки злокачественности схожи для разных карцином, патологоанатомы могут применять общие знания о распространенных патологиях для предварительной оценки редких и атипичных поражений. В контексте гистологического анализа это особенно полезно для классификации редких или атипичных опухолей, где количество размеченных данных ограничено. Достижения глубокого обучения позволили сверточным нейронным сетям решать сложные задачи классификации изображений. Однако диагностические нейронные сети редко используются вне контекста обучения. В связи с этим исследователи предлагают применять предварительно обученные модели к гистоморфологическому анализу, используя их как универсальные цифровые инструменты для аннотирования тканей, классификации и прогнозирования течения заболевания [34,37].

Анализ генетических дефектов при ПКР представляет собой важное направление в изучении молекулярных основ заболевания, позволяя выявлять ключевые изменения, влияющие на биологическое поведение опухоли. Генетические мутации и хромосомные аномалии могут определять скорость роста опухоли, её инвазивность, способность к метастазированию и ответ на различные терапевтические подходы. Современные методы искусственного интеллекта позволяют интегрировать гистологические данные с информацией о генетических и протеомных нарушениях, что даёт возможность выявлять скрытые закономерности между морфологией клеток и молекулярными профилями опухоли. Для этого используются большие базы данных, такие как CPTAC (Clinical Proteomic Tumor Analysis Consortium) и TCGA, содержащие как секвенированные образцы тканей, так и соответствующие им гистологические изображения. Такие интегративные подходы позволяют не только определять влияние конкретных генетических повреждений на прогноз заболевания, но и выявлять потенциальные биомаркеры, включая микроРНК и белковые продукты генов, которые могут

использоваться для предсказания агрессивности опухоли и выбора индивидуальной терапии. Благодаря этим технологиям становится возможным прогнозирование выживаемости пациентов, стратификация риска и разработка персонализированных схем лечения, что особенно важно для подтипов ПКР с высокой гетерогенностью и различной чувствительностью к стандартным методам терапии [38]. Кроме того, модели ИИ, в частности на основе случайного леса, применяются для выдвижения гипотез о влиянии различных генетических повреждений на прогноз и выживаемость пациентов [33].

Важным аспектом является точная предоперационная диагностика с использованием неинвазивных методов, например радиомики. Радиомика включает количественный анализ медицинских изображений, таких как КТ, МРТ и ПЭТ, с целью выявления скрытых паттернов и признаков, которые могут быть недоступны визуальной оценке. Объединение радиомических данных с цифровой патологией открывает возможность корреляции морфологических признаков на уровне клеток и тканевых структур с визуальными характеристиками опухоли на радиологических изображениях. Ряд исследований показал, что анализ рентгенологических признаков на венозно-фазовой компьютерной томографии с применением методов машинного обучения позволяет достаточно точно различать подтипы опухолей почек [39,40]. В отдельных работах гистопатологические характеристики опухоли (метастатический и неметастатический ПКР) оценивались с использованием ИИ в сочетании с клиническими данными, такими как МРТ и КТ [41]. Другие исследователи изучали связь между степенью опухоли по WHO/ISUP и радиологическими признаками у 406 пациентов, применяя SVM вместе с тремя методами отбора признаков: регрессией LASSO, рекурсивным устранением признаков (RFE – Recursive Feature Elimination) и алгоритмом Relief [42]. Такой интегративный подход способствует более точной стратификации пациентов, прогнозированию агрессивности опухоли и выживаемости, а также помогает в разработке персонализированных стратегий лечения, учитывающих как микроархитектуру ткани, так и макроскопические особенности опухолевого очага.

1.2 Применение компьютерного зрения для диагностики рака почки

1.2.1 Дифференциальная диагностика различных типов почечно-клеточного рака

Для точного определения гистологического типа опухоли при различных формах ПКР, а также для эффективного наблюдения за пациентами, важна надежная морфологическая диагностика. Цифровая патология предоставляет возможности для решения этой задачи. Внедрение технологий ИИ в повседневную гистопатологическую практику открывает доступ к дополнительным аналитическим инструментам, позволяющим точно определять тип рака и ускорять процесс постановки диагноза (Рисунок 3).

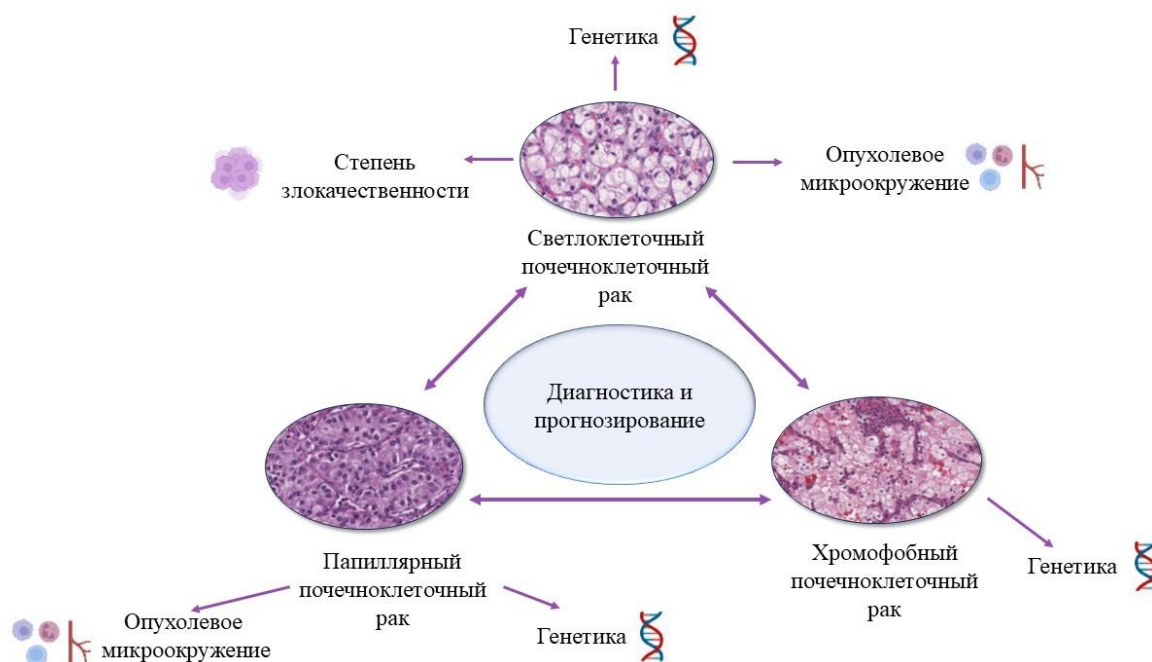


Рисунок 3 – Дифференциальная диагностика между основными типами почечно-клеточного рака

Помимо точного определения гистологического типа опухоли, важной задачей является различение злокачественных и доброкачественных

новообразований, что особенно актуально для онкоцитомы и хПКР, а также метанефральной аденомы и пПКР. В одном исследовании было проанализировано 48 оцифрованных гистологических изображений (по 12 для каждого типа — сПКР, пПКР, хПКР и онкоцитомы). Скан-изображения преобразовывались в четырехуровневые оттенки серого, затем выполнялась сегментация, после чего из ПКР извлекались морфологические признаки для классификации подтипов опухолей. Использование многоклассового байесовского классификатора с многомерными гауссовскими распределениями позволило достигнуть точности свыше 90% [43].

В другом исследовании показано применение методов глубокого обучения для автоматической классификации подтипов ПКР и прогнозирования исходов на основе цифровых гистопатологических изображений. Сверточные нейронные сети, обученные на полных сканах препаратов, отличали сПКР и хПКР от нормальной ткани с точностью 93,39% и 87,34% соответственно. Модель, различающая сПКР, хПКР и пПКР, достигла точности 94,07%. Для решения задачи многоклассовой классификации исследователи использовали метод опорных векторов, что позволило разбить задачу на несколько бинарных подзадач, повысив производительность и компенсировав дисбаланс данных. Из областей опухоли с высокой вероятностью, выявленных сверточной сетью, извлекались морфологические признаки, на основе которых формировался индекс риска, достоверно коррелирующий с выживаемостью пациентов с сПКР [37].

В другой работе авторы обошлись без сегментации и применили метод Харриса для выявления ключевых точек на 48 скан-изображениях четырех типов почечных опухолей. В этом случае байесовский классификатор показал точность 83% [44]. Мэнгдан Чжу и соавторы разработали глубокую нейронную сеть для классификации оцифрованных гистологических препаратов после резекции и биопсии на пять категорий: сПКР, пПКР, хПКР, онкоцитомы и нормальная ткань. Конвейер глубокого обучения позволял классифицировать препарат в целом, а также визуализировать ключевые области на уровне участков для проверки достоверности модели. Средняя площадь под ROC-кривой (AUC) составила 0,98

(95% ДИ: 0,97–1,00) для резекционных образцов, 0,98 (95% ДИ: 0,96–1,00) для биопсийных данных Дартмут-Хичкок и 0,97 (95% ДИ: 0,96–0,98) для TCGA [45].

Кроме точного определения гистологического типа опухоли, важным аспектом является различие злокачественных и доброкачественных новообразований. В одном исследовании был сформирован набор гистологических сканов, включающий сПКР, пПКР, хПКР, светлоклеточный папиллярный ПКР (спПКР), онкоцитому и метанефральную аденому. Для обучения классификатора на основе ИИ было использовано 298 071 «патч» размером 350×350 пикселей. Применение коммерческого инструмента Google AutoML Vision показало среднюю точность 76% (auPRC) после одного часа обучения. После восьми часов обучения финальная модель достигла 93% auPRC и корректно идентифицировала 47 из 55 случаев (сПКР 11/13, пПКР 14/15, хПКР 10/11, спПКР 2/4, онкоцитомы 8/9, метанефральная аденома 2/3) [46].

В другом исследовании использовался ансамбль из трех сверточных нейронных сетей, каждая из которых анализировала «патчи» разного размера (малые 250×250 , средние 350×350 и большие 450×450). На 64 отсканированных препаратах определялись четыре типа ткани: жировая, паренхима почки, сПКР и светлоклеточный папиллярный ПКР. Конечный результат работы ансамбля представлял собой карту изображения, где каждый пиксель классифицировался по классу, а для согласованности карта обрабатывалась с помощью сглаживания случайных полей Гаусса–Маркова. Ансамбль продемонстрировал высокую точность идентификации классов пикселей (0,89 через ResNet18) с ожидаемой точностью 0,92 [47].

Различия в экспрессии и локализации иммуногистохимических маркеров в подтипах рака почки имеют ключевое значение для прогноза и реакции на терапию. Отдельные антитела позволяют прогнозировать исход заболевания (например, PD-L1, Ki-67), тогда как другие служат маркерами васкуляризации (CD31) и иммунной инфильтрации (CD3, CD8, CD20, CD68). В одном исследовании анализировалось микроокружение 83 опухолей различных типов (пПКР, карцинома собирательных протоков и высококодифференцированная уротелиальная карцинома) с

использованием 10 иммуногистохимических маркеров. Классификатор на основе случайного леса HALO определял экспрессию CD68, CD163, CD3, CD8, CD20 и виментина, а для CD31, Ki-67 и PD-L1 применялся алгоритм колокализации объектов, рассчитывая количество положительно окрашенных клеток на 1 мм². Статистический анализ выявил различия между гистологическими типами: CD3+ и CD20+ клетки преобладали в карциноме собирательных трубочек, CD68+ — в пПКР типа 1, а PD-L1 был выше в метастатических опухолях. Ki-67 был ниже в пПКР типа 1 по сравнению с типом 2 (IRR = 0,47, 95% ДИ = 0,21–0,92, p = 0,017) [36].

Для масштабирования и повышения точности диагностики применяются предварительно обученные нейронные сети. Кевин Фауст и коллеги использовали CNN, изначально обученную на морфологии опухолей головного мозга, для анализа 550 цифровых изображений двух типов ПКР (396 сПКР и 154 пПКР). Из извлеченных векторов признаков проводилась неконтролируемая кластеризация, что позволило исследовать клиническую и биологическую значимость подгрупп пациентов. Всего было выделено 512 морфологических признаков для анализа внутри- и межпациентных различий. Использование предварительно обученной сети позволяет расширять модели на новые патологии, минимизируя различия в морфологии ткани без дополнительной настройки [34].

Обученные модели ИИ на основе гистологических данных могут использоваться не только для диагностики, но и для прогнозирования течения заболевания. Так, в одном исследовании сверточная НС (CNN) была обучена на гистологических скан-изображениях из базы TCGA для различения опухолевой и нормальной ткани, а также определения гистологического типа ПКР (сПКР, пПКР, хПКР) с точностью 87,69%. С помощью CNN выделялись зоны опухоли с высокой вероятностью риска, формируя вероятностные тепловые карты. Для каждого типа ПКР из опухолевой ткани извлекались морфологические характеристики клеток и ядер, включая площадь и периметр. Индекс риска рассчитывался с применением регуляризованной модели Кокса-LASSO для каждого признака и подтверждался двухэтапной перекрестной проверкой. В результате было выявлено, что 13

клеточных и 6 ядерных характеристик статистически значимо связаны с выживаемостью пациентов ($p < 0,05$) [37].

Нейронные сети способны выполнять несколько задач одновременно: идентифицировать опухолевую и нормальную ткань, определять гистологический тип и степень злокачественности, а также прогнозировать исход заболевания. Эндрю Чемпион и соавторы классифицировали гистологические скан-изображения ПКР по четырём степеням злокачественности по Фурману, используя показатели цвета и текстуры ядер, формы ядер и топологические признаки. Когорта включала 47% пациентов с сПКР, 33% — с пПКР и 20% — с хПКР; всего было проанализировано 160 оцифрованных гистологических препаратов. Из каждого изображения извлекалось 1316 признаков цвета, текстуры и топологии с использованием бинарной сегментации. Модель достигла точности 90,4% при классификации на четыре степени злокачественности по Фурману [31].

В другом исследовании сверточная нейронная сеть одновременно различала опухолевую и нормальную ткань, определяла тип ПКР и его степень злокачественности по Фурману. Модель показала общую точность 99,1% при различении нормальной паренхимы и ПКР (чувствительность 100%, специфичность 97,1%), точность классификации сПКР, пПКР и хПКР составила 97,5%, а определение степени злокачественности — 98,4% [33]. Однако использование системы Фурман для хПКР остаётся спорным, поэтому многие патологи предпочитают классификацию Панера [48,49].

1.2.2 Применение компьютерного зрения для диагностики светлоклеточного почечно-клеточного рака

Светлоклеточный почечно-клеточный рак (сПКР) представляет собой морфологически разнообразную злокачественную опухоль, происходящую из эпителиальных клеток почечных канальцев. Опухолевые клетки преимущественно имеют прозрачную цитоплазму, реже — эозинофильную, и окружены плотной

сетью кровеносных сосудов. СПКР является самым распространенным типом ПКР у взрослых, составляя 60–75% всех случаев. Чаще всего заболевание возникает спонтанно, без выявленных предрасполагающих факторов, и нередко обнаруживается случайно, например при обследованиях, не связанных с патологией почек. Метастазы наблюдаются у 25% пациентов, при этом средняя продолжительность жизни составляет 6–12 месяцев. У 70–90% больных выявляются изменения в сегменте 3p длинного плеча третьей хромосомы, включая инактивацию гена фон Гиппеля–Линдау вследствие миссенс-мутаций или гиперметилирования промотора [4].

Основным прогностическим критерием при сПКР является степень злокачественности опухоли, оцениваемая по морфологии ядер. Ранее для этого использовалась четырехступенчатая классификация Фурман, основанная на форме ядер и визуализации ядрышек, а современная система WHO/ISUP также включает четыре степени злокачественности. Для упрощения анализа применяют двухступенчатые подходы, объединяющие степени в высокодифференцированные (1–2) и низкодифференцированные (3–4). В одном исследовании была разработана нейронная сеть для распознавания высоко- и низкодифференцированных сПКР по классификации Фурман с прогнозированием выживаемости пациентов. Из 72 количественных морфологических и текстурных признаков LASSO-модель, построенная на данных 42 пациентов, включала 26 признаков и продемонстрировала чувствительность 84,6% и специфичность 81,3%. Модель была протестирована на 160 случаях с учетом возраста, пола, степени злокачественности, стадии TNM и лечения, а оценка пропорционального риска Кокса подтвердила прогностическую значимость предсказанных степеней [22].

Высокую и низкую степень злокачественности также возможно определять автоматически по размерам ядер. Для этого использовался автоматизированный анализ 39 оцифрованных гистологических препаратов сПКР, окрашенных гематоксилином и эозином. Регрессия ядра позволяла оценивать пространственное распределение размеров по всему препарату. Результаты коррелировали с ядерными градациями по Фурману, установленными патологоанатомами. На

основе распределения размеров ядер создавались тепловые карты с характерными диаметрами: для высокодифференцированного сПКР — 6 мкм, для низкодифференцированного — 9 мкм. Для классификации применялись алгоритм распознавания цвета WS-Recognizer и метод опорных векторов с использованием RGB-каналов. Статистический анализ показал значимые различия максимального размера ядра между низко- и высокодифференцированными опухолями ($p < 0,001$). Анализ рабочих характеристик приёмника продемонстрировал отличное разделение: частота ложноположительных результатов — 0,2, частота истинно положительных — 1,0, площадь под кривой составила 0,97 [23].

Крук и соавторы продемонстрировали возможность автоматического определения степени злокачественности сПКР по классификации Фурман с общей точностью 96,7%, при этом чувствительность и специфичность по отдельным степеням колебались от 87,3 до 99,3%. Анализ изображений включал сегментацию с применением вейвлет-преобразования на этапе предварительной обработки, последующее использование нескольких методов отбора признаков и формирование ансамбля классификаторов на финальном этапе распознавания. Диагностические признаки ядер, основанные на текстуре, геометрии, цвете и гистограмме, отбирались с помощью нескольких подходов, каждый из которых применял уникальный механизм отбора. Полученные наборы признаков позволяли создавать ансамбли классификаторов на основе машин опорных векторов и случайного леса [29].

В других исследованиях создавались автоматизированные конвейеры для разделения сПКР на две группы по степени злокачественности, с использованием методов машинного обучения и извлечения признаков на основе интенсивности пикселей для анализа морфологии ядер и ядрышек. Конвейеры выделяли области изображения с ядрышками при помощи двух детекторов и количественно оценивали плеоморфные паттерны ядер, объединяя признаки из различных участков изображения [24].

Помимо стандартных степеней Фурмана или WHO/ISUP, в ряде исследований вводится «нулевая» степень, соответствующая нормальной почечной ткани.

Qingyuan Zheng и коллеги разработали многоклассовую модель глубокого обучения SSL-CLAM для оценки степени злокачественности по классификации Фурман на скан-изображениях почечно-клеточного рака. Авторы также использовали подход слияния человек–машина для проверки дополнительной ценности модели. SSL-CLAM успешно различала пять уровней (0–4) сПКР, достигая AUC 0,917 и 0,887 на внутреннем и внешнем наборах данных соответственно, и демонстрировала высокую эффективность при двухуровневой оценке (степень 0, степень 1/2 и степень 3/4) [50].

В другом исследовании была предложена полностью автоматизированная нейронная сеть RCCGNet для анализа гистопатологических изображений почек. Два набора данных (КМС и BreakHis) включали скан-изображения, классифицированные как нормальные или опухолевые (степени 0–4). Эффективность RCCGNet оценивалась с использованием стандартных метрик: для КМС точность достигла 90,14%, F1-мера — 89,06%, для BreakHis — 90,09% и 88,90% соответственно [51].

Методы машинного обучения также применяются для прогнозирования течения сПКР и оптимизации клинического ведения. Например, разработана мультимодальная модель глубокого обучения (MMDLM), объединяющая индивидуальную 18-слойную резидуальную сеть (ResNet18) для каждого типа данных: гистологических изображений, КТ и МРТ-снимков, а также геномных данных. Обучение проводилось поэтапно: сначала на гистологических препаратах, затем с добавлением визуальных и геномных данных. В исследовании участвовали данные 113 пациентов, точность модели составила $83,43\% \pm 11,62\%$, при максимуме 100% по результатам 12-кратной перекрестной проверки. Авторы показали, что прогноз MMDLM является независимым прогностическим фактором и превосходит другие клинические показатели [41].

В другом исследовании показано, что оценки, полученные с помощью разработанного конвейера, коррелируют ($R = 0,59$) с мультигенной системой прогностической оценки, важной для пациентов с сПКР. Для этого был создан конвейер на основе машинного обучения и методов извлечения признаков на основе

интенсивности пикселей для анализа ядер [24]. Авторы продемонстрировали, что морфологическая сигнатура, извлеченная с помощью машинного обучения, может служить новым прогностическим маркером. Сегментация и обнаружение ядер в опухолевой ткани выполнялись с использованием программного обеспечения QuPath на трех когортных выборках, где определялись морфологические характеристики клеток, ядер и цитоплазмы. Для отбора наиболее значимых признаков применялся пакет Glimnet, а коэффициенты признаков использовались для построения модели. Среднее время наблюдения в когортных выборках составило $26,4 \pm 27,8$ и $43,2 \pm 30,6$ месяцев. В анализе методом LASSO были выделены следующие прогностические параметры: округлость ядра, минимальный калибр ядра, средняя оптическая плотность гематоксилина ядра, минимальная оптическая плотность гематоксилина и стандартное отклонение. Эти показатели легли в основу многослойной системы перцептрона (MLPS), которая достоверно различала пациентов с высоким риском выживаемости с отношением рисков 15,05, 4,49 и 1,65 в трех независимых когортах. Регрессионный анализ Кокса подтвердил, что MLPS может служить независимым прогностическим фактором. Интегрированная номограмма, объединяющая MLPS, систему стадирования и степень злокачественности, повысила точность прогнозирования выживаемости: AUC ROC составила 89,5%, 90,0%, 88,5% и 85,9% для оценки безрецидивной выживаемости через 1, 3, 5 и 10 лет после постановки диагноза [52].

Прогнозирование течения сПКР также может основываться на анализе васкуляризации, площади некрозов и других гистоархитектурных характеристик. Наличие участков некроза в опухоли является неблагоприятным прогностическим признаком, а плотность сосудов коррелирует с инвазивностью опухоли и вероятностью метастазирования. Иммуногистохимическое окрашивание эндотелиальных клеток и оцифровка препаратов облегчают выявление сосудов. В одном исследовании разработан алгоритм для анализа взаимосвязи между экспрессией генов и морфологией сосудов. Были созданы два инструмента машинного обучения: первый — для идентификации эндотелиальных клеток, второй — для определения границ сосудов и периваскулярных областей. Для

повышения точности использовались дублированные срезы, окрашенные гематоксилином и эозином, с последующей иммуногистохимической обработкой, что позволило создать маску сосудистых областей. Девять характеристик сосудов оказались предикторами выживаемости ($n = 64$, $RR = 2,3$). С помощью двух обобщенных линейных моделей с 14 генами (14VF и 14GT) пациентов разделили на группы с высокой и низкой выживаемостью ($RR_{14VF} = 2,4$, $RR_{14GT} = 3,33$), что подтвердило эффективность цифрового анализа изображений и машинного обучения для разработки морфологически ориентированных прогностических сигнатур экспрессии генов [15].

Кроме того, данные иммуногистохимического анализа инфильтрирующих опухоль лимфоцитов (TIL) применяются для прогнозирования злокачественного потенциала сПКР. В исследовании оценивалась экспрессия CD3+ Т-клеток, CD8+ цитотоксических Т-лимфоцитов, регуляторных Т-клеток, В-клеток, плазматических клеток, макрофагов, гранулоцитов, PD-1 и PD-L1 у 756 пациентов с сПКР. Оцифровка и количественный анализ проводились с помощью платформы HALO, а цитонуклеарный модуль (v1.4e1.6) с классификатором тканей использовался для различения опухолевой и нормальной ткани и оценки числа положительно окрашенных клеток и плотности клеток в области анализа. Одномерный анализ выживаемости показал, что высокая инфильтрация В- и Т-клетками и PD-1-позитивными клетками ассоциировалась с благоприятным прогнозом, тогда как высокий уровень гранулоцитов, макрофагов, цитотоксических Т-клеток и экспрессии PD-L1 — с неблагоприятным исходом. Высокая плотность цитотоксических лимфоцитов или В-клеток и выраженность PD-L1 рассматривались как независимые прогностические маркеры, а также учитывались при оценке прогноза после терапии ниволумабом [53].

Применение ИИ для анализа нескольких молекулярно-генетических маркеров в образцах обладает значительным прогностическим потенциалом и может быть объединено с морфологическими характеристиками для улучшения стратификации риска и персонализированной терапии. Несмотря на успехи ингибиторов иммунных контрольных точек, у примерно трети пациентов со сПКР

продолжают возникать рецидивы. Современные прогностические алгоритмы, например шкала Лейбовича, базируются на ручной оценке морфологических признаков и подвержены субъективной вариабельности. Более того, такие подходы не учитывают гетерогенную молекулярную среду микроокружения опухоли, которая также имеет прогностическое значение.

В одном исследовании был разработан полуавтоматический метод анализа 62 маркеров и их комбинаций в 150 первичных сПКР с использованием мультиплексной иммунофлуоресценции, цифрового пространственного профилирования NanoString GeoMx и ИИ-анализа изображений. Изображения обрабатывались в Definiens Tissue Studio и Indica Labs Halo AI. Сегментация стромы выполнялась только в опухолевой ткани, а ядерная сегментация основывалась на интенсивности окрашивания Hoechst. Было выявлено, что коэкспрессия маркеров раковых стволовых клеток и эпителиально-мезенхимального перехода (OCT4, ZEB1) ассоциировалась с неблагоприятным исходом. Кроме того, маркеры OCT4, CD8, CD34 и CD163 позволяли эффективно стратифицировать риск у пациентов на промежуточной фазе лечения [54].

Гистологические характеристики сПКР тесно связаны с генетическими и эпигенетическими данными, которые также могут выступать прогностическими маркерами. В одном исследовании использовались методы машинного обучения, включая глубокие нейронные сети, для анализа связи между протеомикой и гистологией. На первом этапе строилась диагностическая модель на основе протеомных данных с применением классификатора случайного леса (RF) на 997 белковых продуктах, которая различала нормальные и опухолевые образцы с точностью 0,98 по результатам 10-кратной перекрестной проверки (чувствительность 0,97, специфичность 0,99). На втором этапе создавалась сверточная нейронная сеть (CNN) для различения сПКР и нормальной ткани почки с точностью 0,95, чувствительностью 1 и специфичностью 0,93 на тестовом наборе. Далее модели совместно тестировались на 24 образцах и обучались выявлять корреляции между протеомными данными, соответствующими генами и прогнозом, основанным на гистологических признаках [24].

ИИ также демонстрирует перспективы в диагностике редких подтипов с генетическими нарушениями. Транслокационная почечно-клеточная карцинома TFE3 Xp11.2 (TFE3-RCC) отличается более агрессивным течением по сравнению с другими подтипами ПКР и трудна для выявления при стандартной световой микроскопии. Для решения этой задачи был создан автоматизированный вычислительный конвейер для извлечения признаков из гистологических изображений. Выявлено 52 морфологических признака, способных различать TFE3-RCC и светлоклеточный ПКР. Тестирование на внешнем наборе данных показало высокую точность с AUC 0,842–0,894, что позволяет фиксировать даже минимальные морфологические различия и существенно повышает эффективность диагностики редких вариантов ПКР [21].

1.2.3 Применение компьютерного зрения для диагностики папиллярного почечно-клеточного рака

Папиллярный почечно-клеточный рак (пПКР) развивается из эпителия дистальных извитых канальцев и составляет около 10–15% всех случаев почечно-клеточного рака. В сравнении с светлоклеточным вариантом, пПКР имеет более благоприятный прогноз, с пятилетней выживаемостью на уровне 80–85%. Это гетерогенное заболевание, представленное двумя основными гистологическими подтипами: тип 1 включает сосочковые и тубулярные структуры, выстланные мелкими клетками с базофильной цитоплазмой и небольшими овальными ядрами, а тип 2 состоит из сосочков, покрытых крупными клетками с эозинофильной цитоплазмой и большими сферическими ядрами с отчетливо выраженными ядрышками. В отдельных случаях пПКР может проявлять мультифокальное или индифферентное течение, в то время как одиночные опухоли нередко демонстрируют агрессивный летальный фенотип [4].

Микроокружение опухоли оказывает существенное влияние на прогноз заболевания и позволяет выявлять гистологические биомаркеры, характерные для

конкретных подтипов. В одном исследовании был разработан новый автоматизированный конвейер для топологического анализа клеточных структур микроокружения, протестированный на 190 отсканированных гистологических препаратах пациентов с пПКР из базы TCGA. Далее был проведён однофакторный анализ выживаемости с использованием 50 морфологических признаков и двух патологогистологических критериев — стадии TNM и подтипа пПКР. Стадии II и III TNM были объединены для сравнения со стадией I. Пациенты с пПКР типа 2 продемонстрировали худший прогноз по сравнению с типом 1. Результаты показали, что топологические признаки эффективно позволяют стратифицировать пациентов по выживаемости на ранней и промежуточной стадиях. Значения AUC для предсказания бинарного 5-летнего исхода по стадии и подтипу составили 0,63 и 0,66, соответственно, а индекс риска достиг AUC 0,78. Такой подход не только углубляет понимание топологической организации опухолевых тканей, но и может быть объединён с геномными данными для разработки комплексных биомаркеров [22].

Совмещение морфологических и генетических данных представляет собой дополнительный инструмент для прогнозирования исхода пПКР. В одном из исследований была выполнена всесторонняя характеристика иммунной микросреды опухолей с использованием генетического анализа и мультиомного биоинформатического подхода. Результаты показали, что пПКР демонстрирует признаки «горячей» опухоли, однако наличие CD8⁺ Т-клеток не препятствовало прогрессированию заболевания. Это указывает на потенциальную клиническую ценность терапии, направленной на увеличение числа CD8⁺ Т-клеток. Более того, экспрессия CCL5 и FASLG может способствовать формированию иммуносупрессивной микросреды, что открывает новые возможности для исследований персонализированной иммунотерапии у пациентов с пПКР [55].

Ле Ли и коллеги провели биоинформатический анализ для выявления потенциальных биомаркеров, связанных с повреждением ДНК и окислительным стрессом в пПКР. Данные РНК-секвенирования были получены из базы TCGA, а дифференциально экспрессируемые гены (DEG) определялись с использованием

различных алгоритмов кластеризации и классификации. В результате были выделены гены BDKRB1, NMUR2, PMCH и SAA1, которые могут служить перспективными предиктивными биомаркерами и потенциальными терапевтическими мишенями для пПКР [56].

1.2.4 Применение компьютерного зрения для диагностики хромофобного почечно-клеточного рака

Хромофобный почечно-клеточный рак (хПКР) формируется из кортикальных собирательных трубочек и встречается примерно в 3–5% случаев всех ПКР. По прогнозу данная опухоль значительно благоприятнее светлоклеточного и папиллярного вариантов, при этом пятилетняя выживаемость превышает 90%. Клеточные ядра хПКР имеют характерный морщинистый, «изюмовидный» вид, а наличие двуядерных клеток наблюдается достаточно часто. Высокая морфологическая вариабельность затрудняет постановку диагноза, особенно для неопытных патологоанатомов, и может приводить к ошибочной идентификации сПКР. Несмотря на существующие предложения по классификации хПКР, ни одна из них не получила широкого клинического применения, а классификация WHO/ISUP пока не прошла валидацию для этого подтипа. В литературе обсуждаются двух-, трёх- и четырёхступенчатые схемы оценки злокачественности, из которых наиболее известна трёхступенчатая система Панера. При этом ни стадирование по TNM, ни наличие саркоматоидной дифференцировки не оказывают существенного влияния на прогноз [4].

Согласно рекомендациям ВОЗ, хПКР делится на классический и эозинофильный подтипы. Классический вариант характеризуется крупными клетками с ретикулярной цитоплазмой и отчетливо видимыми клеточными мембранами (бледные клетки). В одном исследовании были проанализированы три когорты: 42 пациента из отделения патологии и молекулярной патологии университетской больницы Цюриха, 199 пациентов из японских медицинских

учреждений и 66 пациентов из TCGA. В обеих когортах выживаемость не различалась между эозинофильным и классическим подтипами. Для изучения взаимосвязи генотипа и фенотипа был проведён полногеномный анализ CNV на платформе Affymetrix OncoScan CNV в 33 образцах хПКР. В объединённой выборке из Цюриха и TCGA потери хромосом 1, 2, 6, 10, 13 и 17 встречались значительно чаще в классическом варианте ($p < 0,05$), что указывает на более высокую хромосомную нестабильность и позволяет отличать два подтипа хПКР [57].

Особую сложность представляет дифференциация хПКР и доброкачественной почечной онкоцитомы, так как эти опухоли обладают схожей морфологией, но требуют различных подходов к ведению пациента. В литературе описаны многочисленные исследования, где для разделения онкоцитомы и хПКР применялось машинное обучение на данных компьютерной томографии и других радиологических методов [58–60]. В одном исследовании Кевин Бреннан и коллеги изучали метилирование ДНК в свежзамороженных образцах онкоцитомы и хПКР с использованием методов машинного обучения для выявления дифференциально метилированных сайтов CpG. Эти маркеры оказались надежными для различения опухолей. Примечательно, что онкоцитома демонстрировала более выраженные аномалии метилирования по сравнению с хПКР. Всего было идентифицировано 79 CpG с различиями между онкоцитомой и хПКР. Разработанная диагностическая модель позволяла различать онкоцитому и хПКР с AUC = 0,96 (95% ДИ 0,88–1,00) при 10-кратной перекрестной проверке. Профиль CpG также позволял дифференцировать онкоцитому от других подтипов ПКР и нормальной ткани, что делает его перспективным диагностическим биомаркером [61].

1.3 Использование компьютерного зрения в диагностике отторжения биоптатов трансплантированных почек

Трансплантация почки считается наиболее оптимальной стратегией лечения хронической почечной недостаточности, значительно продлевающей жизнь и

улучшающей ее качество для пациентов [62]. Первая успешная операция по трансплантации почки была проведена в 1954 году американским хирургом Джозефом Мюрреем [63]. Количество операций по пересадке почки достигло рекордного уровня в 22 393 в США в 2018 году, увеличившись на 6,5 % по сравнению с 2017 годом [64]. Продолжительность жизни трансплантированного органа зависит от многих факторов, которые можно достоверно оценить с помощью морфологического исследования. Потеря функции трансплантата может быть вызвана как отторжением, так и рецидивом основного заболевания. Отторжение остается преобладающей причиной неудачных трансплантаций, несмотря на прогресс в иммуносупрессивной терапии. Отторжение считается иммунологической реакцией реципиента на донорские антигены, что приводит к разрушению пересаженной ткани или органа. Оно делится на две основные подгруппы на основе механизма патогенеза: отторжение, опосредованное Т-клетками (клеточное), и отторжение, опосредованное антителами (гуморальное). В некоторых случаях отторжение имеет смешанную природу. Отторжение также может быть острым или хроническим, в зависимости от течения заболевания [65]. Острое отторжение развивается у 7,0 % взрослых реципиентов в течение первого года после трансплантации с частотой 9,1 % у реципиентов в возрасте 18–34 % и 6,1 % у реципиентов старше 65 лет [66]. Гистологическое исследование биоптатов трансплантатов является «золотым стандартом» диагностики отторжения в клинике.

Было разработано несколько морфологических классификаций, в первую очередь для клеточного отторжения. В частности, классификация Banff является наиболее широко используемой патологами. Эта классификация была впервые предложена Ким Солез и Лоррейн К. Ракузен в 1991 году и с тех пор несколько раз пересматривалась [67]. В издании Banff-классификации 2019 года указаны следующие прогностические признаки: фиброз [c], интерстициальная инфильтрация [i], тубулит [t], артериит [v] для клеточного отторжения; гломерулит [g], перитубулярный капиллярит [ptc] и отложение C4d [C4d] для гуморального

отторжения [67]. Каждому признаку присваивается оценка (от 0 до 3 баллов), которая используется патологом для диагностики состояния трансплантата.

Несмотря на свою полноту, классификация Banff имеет некоторые ограничения. Наиболее важным из них является вариабельность между наблюдателями: оценка тяжести отторжения в биопсии может различаться у врачей в зависимости от их квалификации и интерпретации морфологических данных. Banff-классификация имеет довольно большое количество критериев, оценка которых основана на полуколичественном методе. Этот метод может быть субъективным, поскольку патологоанатомы должны мысленно рассчитывать относительные площади и относительные количественные показатели, что приводит к потенциальным различиям в оценках. Пограничные изменения являются особенно сложным диагностическим признаком, поскольку пациенты с невыраженными изменениями в ткани трансплантата получают иное лечение, чем пациенты с отторжением [68]. Диагностические критерии для этой группы были дважды пересмотрены в классификации Banff [68].

Эти проблемы можно решить с помощью использования машинного обучения. ИИ может помочь стандартизировать исследования, устранить различия между наблюдателями и повысить объективность за счет точной количественной оценки морфологических характеристик. Последний признак важен для четкого разграничения отторжения трансплантата от пограничных изменений. Кроме того, ИИ может помочь разработать четкие критерии, основанные на алгоритмических расчетах компьютерных программ, улучшая воспроизводимость классификации Banff в клинической практике. Нефропатология является одним из передовых направлений для применения цифровой патологии. ИИ интенсивно использовался для сегментации структур в биопсиях почек и их количественной оценки. Сверточные НС применялись для дифференциации нормальных и патологических тканей на скан-изображениях препаратов почки [69,70]. Впоследствии технологии ИИ стали более сложными и применялись для оценки тяжести патологических изменений [67,69]. Некоторые исследователи используют ИИ для оценки степени фиброза в интерстиции и атрофии почечных канальцев при хроническом

заболевании почек. Автоматизированная оценка васкуляризации тканей и инфильтрации иммунными клетками использовались для оценки воспаления в паренхиме почки [69,71–73]. Интересной задачей является способность ИИ точно вычислять наиболее важные прогностические морфологические паттерны, например, фиброз и интерстициальную инфильтрацию, чтобы предсказывать экспертные оценки и дифференцировать пограничные изменения от острого отторжения.

Исследования ИИ в области нефропатологии выделяются среди всех исследований цифровой патологии благодаря сосредоточению исследований на неопухолевых тканях. Повторные пункционные биопсии трансплантатов почек и удаление фрагментов или целых почек по поводу опухолевых процессов позволили накопить большое количество гистологических препаратов, необходимых для изучения нормального строения паренхимы почки. В статье Бутельи и соавторов сверточные НС были обучены для сегментации здоровой почечной паренхимы крыс. Далее технологии машинного обучения применялись для оценки образцов с заболеваниями (болезнь минимальных изменений, острое повреждение канальцев почки и гломерулонефрит). На основе примерно 72 700 аннотаций были идентифицированы шесть основных структур почек: клубочки, пучки капилляров клубочка, канальцы, вены, артерии и их просветы. Модель также использовалась для оценки расширения интерстиция, тубулярной дилатации и атрофии и вариабельности размера клубочков [74]. В более ранней статье НС была обучена для сегментации 11 классов структур тканей в биопсиях почечных трансплантатов. Расчеты использовались для прогнозирования нескольких баллов системы Banff (ci, pts, i), и была обнаружена корреляция между экспертной оценкой и сверточной НС [73]. Несмотря на то, что выявление структур и аномалий в почечной ткани является несомненно важным, однако расстановка акцентов на оценке фиброза и/или интерстициальной инфильтрации позволяет проводить более ценные прогностические оценки острого отторжения почечного трансплантата.

Относительно недавно ИИ стал использоваться для дифференциации нормальной ткани и патологии. Методы обработки изображений, основанные на

машинном обучении, могут определять нормальные и склерозированные клубочки с точностью до 0,98 [70]. Значительным достижением является применение ИИ для диагностики различных заболеваний почек, а в частности для диагностики состояния трансплантата почки [75]. Оценка фиброзных изменений в биопсиях почечного трансплантата важна, поскольку эта метрика коррелирует с потерей функции почки. Дискутабельным является метод окрашивания ткани почки, который должен стать стандартом в цифровой патологии. В исследовании Фарриса и соавторов интерстициальный фиброз оценивался с помощью иммуногистохимической реакции с антителом к коллагену III типа, а также при окрашивании трихромом по Массону. Авторы считают, что использование дополнительных методов окрашивания гистологических препаратов (включая иммуногистохимию), а не рутинно используемых гематоксилина и эозина, может способствовать стандартизации процесса. Однако следует отметить, что как иммуно-, так и гистохимические методы окрашивания должны быть валидированы в каждой лаборатории для обеспечения одинаковых условий оцифровки препаратов. В этой же публикации авторы использовали алгоритм Aperio ImageScope Positive Pixel Count для расчета процента атрофии канальцев и инфильтрации в корковом веществе [76].

Стоит отметить, что несколько описанных алгоритмов ИИ были обучены на скан-изображениях препаратов, окрашенных специализированными красителями (трихром по Массону, иммуногистохимические реакции с антителами к коллагенам, CD34, CD3), которые используются с различными вариациями в реагентах и протоколах в разных лабораториях или редко используются в повседневной практике патолога [69,71–73]. С другой стороны, иммуногистохимическое исследование в сочетании с сегментацией стромы необходимо для оценки с помощью ИИ отложений C4d в перитубулярных капиллярах, которая успешно применялось для обнаружения гуморального отторжения [77]. В одной из статей в качестве системы поддержки принятия врачебных решений, авторы применили ИИ для определения выраженности воспалительной инфильтрации в интерстициальной ткани биоптатов почек после

трансплантации. Иммунные клетки различных типов (Т- киллеры, Т-хелперы, В-лимфоциты, макрофаги) оценивались на основе таких параметров, как ядерно-цитоплазматическое соотношение, плотность клеток и микроокружение. Однако исследование проводилось на иммуногистохимических препаратах, что значительно усложняет и замедляет внедрение данного подхода в клиническую практику [78].

В другой публикации интерстициальная инфильтрация оценивалась с использованием встроенных методов в программное обеспечение сканера Aperio, а также настраиваемого алгоритма с открытым исходным кодом и применялась к образцам, окрашенным антителами к CD3. Также рассчитывалась клеточная плотность на 1 мм². Этот метод можно считать широко воспроизводимым в различных лабораториях, оснащенных необходимыми инструментами, но эти результаты транслируются не для всего ряда существующих гистологических сканеров [72]. Поэтому важной задачей для ученых является разработка ИИ, которые применяются непосредственно к скан-изображениям препарата, а не в операционной системе гистологического сканера, как шаг в направлении многомодульной патоморфологической экосистемы. С другой стороны, во многих публикациях сообщалось о ценных результатах, которые были получены при помощи коммерческих решений с морфометрическим программным обеспечением производителя без какой-либо тканевой специфичности, например, алгоритме подсчета положительных пикселей Aperio (Leica Biosystems) или алгоритме плотности микрососудов Aperio (Leica Biosystems) [69,71–73].

В одной из публикаций исследователи создали конвейер для прогнозирования ранней и долгосрочной выживаемости пациентов с трансплантатами почек. Он основан на оценке патологических изменений в первоначальной биопсии и биопсии трансплантата почки год спустя. Аномалии в интерстиции и канальцах, а также инфильтрированные иммунными клетками регионы оценивались в биоптатах в соответствии с классификацией Vanff 2007 года. Этот метод представляет собой новый подход к стратификации риска аллотрансплантатов и посттрансплантационному мониторингу в клинической практике [79].

Машинное обучение постепенно внедряется в рутинную морфологическую оценку образцов почек. В одной из статей авторы использовали сверточную НС для оценки препаратов при иммунофлуоресцентном исследовании биопсий почек и идентифицировали различные типы иммуноглобулинов, каппа- и лямбда-цепи, а также компоненты системы комплемента, отложенные в клубочковых капиллярах. Исследователи отметили, что использование ИИ в качестве системы поддержки принятия решений для патологоанатомов ускоряет процесс оценки биопсии в десятки раз [80]. Следующим направлением развития для сегментации структур почки может являться использование 3D-гистологических моделей на основе серийных гистологических препаратов [81]. Можно представить, как ценность морфологического анализа возрастет с добавлением третьего измерения, поскольку этот переход сравним с переходом от обычной рентгенографии к рентгеновской компьютерной томографии. Будущими интересными направлениями работы для многих исследователей в нефропатологии является создание автоматизированной морфометрической оценки клубочков и канальцев, которая может конкурировать с очень сложными методами, такими как электронная микроскопия.

Одним из ограничений работ программ на основе ИИ является проведение оценки только «вручную» выбранной корковой части биоптата почки. Интеграция программ в медицинскую практику потребовала бы проведения семантической сегментации структур почечной ткани. Однако, есть основания полагать, что эту проблему можно решить посредством использования трансформеров, однако, это требует значительных вычислительных ресурсов [82,83].

Непростой задачей для исследователей является детекция иммунных клеток внутри канальцев и артерий, что могло бы помочь в оценке тубулита и артериита. Автоматизированный количественный анализ тубулита будет ценным инструментом для анализа разнообразного спектра патологий, включая тубулоинтерстициальный нефрит, различные типы отторжения и вирусные поражения трансплантата. Однако, похоже, что автоматизированная оценка тубулита остается нерешенной проблемой, в первую очередь связанной с классификацией изображений между нормальными канальцами, дистрофическими

канальцами и канальцами с тубулитом. Эта проблема выходит за рамки автоматизации классификации Vanff и требует принципиально иного подхода.

Интеграция искусственного интеллекта в нефропатологию продемонстрировала значительный потенциал для диагностики, прогнозирования и лечения отторжения почечного трансплантата и других заболеваний почек. Использование моделей на основе ИИ для оценки инфильтрации иммунными клетками, фиброзных изменений и различных патологических процессов показало многообещающие результаты с точки зрения точности и эффективности, помогая врачам в процессе принятия решений [84].

ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ

2.1 Применение компьютерного зрения для сегментации почечных структур и диагностики отторжения трансплантата почки

2.1.1 Подготовка набора данных

Для первой части исследования было отобрано 120 биоптатов почечных трансплантатов, а также сопроводительные клинические данные из архива патологоанатомического отделения Российского Научного Центра Хирургии им. Б.В. Петровского (Москва, Россия). Изучаемая когорта состояла из 120 пациентов, в том числе 74 мужчин (61,66 %) и 46 женщин (38,34 %). Средний возраст пациентов составил $39 \pm 10,28$ года и варьировался от 18 до 74 лет. Все пациенты проходили лечение в отделении Трансплантации почки в Российском научном центре хирургии им. Б.В. Петровского с диагнозами, связанными с осложнениями после трансплантации почки (МКБ-10: Т86.1). Критериями включения являлись: (1) возраст пациента старше 18 лет; (2) подтвержденный диагноз Т-клеточного отторжения в анамнезе. Критериями исключения были: (1) наличие нефросклероза в образцах; (2) клинические и/или морфологические признаки гуморального отторжения; (3) наличие артефактов на изображении. Нефросклероз считался критерием исключения, т.к. при нем наблюдалось замещение почечной паренхимы соединительной тканью, что делало невозможным достоверное выявление отторжения. Механизмы развития нефросклероза существенно отличаются от отторжения трансплантата и включают рецидив первичного заболевания, гипертонический ангиосклероз и такие воспалительные состояния как пиелонефрит или вирусные инфекции. Гистологические препараты были изготовлены из пункционных биоптатов почки, которые фиксировались в формалине, а затем подвергались стандартной гистологической обработке, включающей заливку в парафиновые блоки, микротомию и окрашивание

гематоксилином и эозином. Поиск образцов почечных трансплантатов проводился с помощью лабораторной информационной системы, что позволило получить сбалансированную выборку пациентов, основанную на оценке интерстициальной инфильтрации (35 образцов $i = 0$; 36 образцов с $i = 1$; 39 образцов с $i = 2-3$). Препараты биоптатов были изготовлены в период с 2004 по 2022 год в соответствии с данными информационной системы (Таблица 1).

Таблица 1 – Характеристика набора данных почечных биоптатов

	Количество скан-изображений	
	фиброз (сi)	Интерстициальная инфильтрация (i)
Увеличение x20 (размер пикселя = 0.5056 мкм, размер фрагмента = 250x250)		
значение = 0	18	13
значение = 1	8	12
значение = 2	3	3
значение = 3	0	1
Итого	29	29
Увеличение x40 (размер пикселя = 0.2519 мкм, размер фрагмента = 500x500)		
значение = 0	30	22
значение = 1	33	24
значение = 2	14	32
значение = 3	4	3
Итого	81	81
Всего	110	110

Случаи с нефросклерозом были исключены из исследования. Было отсканировано 29 образцов с помощью Leica Aperio AT2 при увеличении в 20 раз и 81 образец – при увеличении в 40 раз. Решение сканировать при разных увеличениях было принято для того, чтобы гарантировать точность модели в больницах, где приоритет отдается скорости рутинного анализа. Все скан-изображения биоптатов были анонимными и не содержали ссылок на клинические случаи. Идентификаторы пациентов были разделены случайным образом на обучающий и валидационный наборы, чтобы предотвратить утечку данных. Этот метод гарантировал, что скан-изображения в обучающем наборе были полностью

отличны от скан-изображений в наборе для валидации, тем самым обеспечивая надежную оценку эффективности модели и возможность обобщения на невидимые данные. Скан-изображения почечных биоптатов были расклассифицированы согласно Vanff-классификации пересмотра 2019 года, являющейся актуальной на сегодняшний день и не имеющей критических изменений с момента последнего пересмотра в 2022 году. Повторная оценка проводилась вслепую с использованием скан-изображений в формате .svs в программе CaseViewer (3DHistech, Венгрия).

В 10 скан-изображениях, содержащих по 2-3 срезов биоптатов, были проаннотированы все структуры с помощью программного обеспечения QuPath v.0.4.3. 120 наиболее отличающихся друг от друга изображений были обрезаны на квадраты размером 512×512 пикселей или 256×256 микрометров и использовались для обучения и тестирования модели. В 110 скан-изображениях из валидационного набора данных были выбраны области интересов для автоматической сегментации в QuPath v.0.4.3. Области включали корковый слой почки, капсула и мозговой слой исключались из анализа.

Коэффициент Дайса использовался для оценки сходства между предсказанной НС маской сегментации и истинной. Коэффициент Дайса варьировался от 0 (отсутствие перекрытия) до 1 (идеальное перекрытие). Коэффициент Дайса вычислялся по следующей формуле:

$$Dice\ Score = \frac{2*TP}{2*TP+FN+FP},$$

где TP — количество истинно положительных результатов, TN — количество истинно отрицательных результатов, FP — количество ложноположительных результатов, а FN — количество ложноотрицательных результатов. Ожидаемые диапазоны коэффициента Дайса для хорошей, приемлемой и плохой оценок могут быть $> 0,8$, $0,6-0,8$ и $< 0,6$ соответственно. Для статистического анализа была рассчитана t-статистика. Дизайн исследования представлен на Рисунке 4.

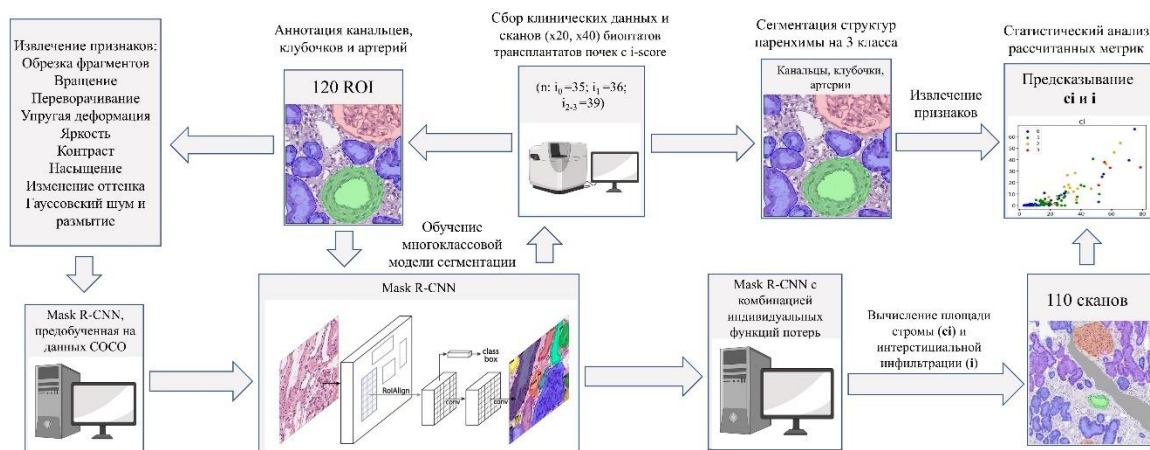


Рисунок 4 - Дизайн изучения моделей компьютерного зрения для сегментации структур почки и диагностики отторжения почечных трансплантатов

2.1.2 Архитектура нейронной сети и ее обучение

Модель глубокого обучения была натренирована на наборе данных, состоящем из 52 скан-изображений, с применением методов аугментации данных для расширения эффективного размера набора данных. Чтобы смягчить переобучение и обеспечить надежную производительность модели, была использована 20-кратная перекрестная проверка. Работа окончательной модели была оценена на отдельном тестовом наборе из 68 изображений.

Для сегментации использовались следующие классы: «клубочки», «канальцы», «артерии», «стромы» и «инфильтрация». Для сегментации «клубочков», «канальцев» и «артерий» была выбрана архитектура Mask R-CNN из-за ее доказанной эффективности в задачах сегментации структур особенно в анализе медицинских изображений. Также использовался фреймворк Detectron2 для имплементации и обучения модели [85].

Модель Mask R-CNN была изначально предварительно обучена на наборе данных COCO (Common Objects in Context). Этот этап предварительной подготовки имел важное значение, поскольку он позволял модели изучать общие характеристики из большого и разнообразного набора данных, который затем

настраивался для конкретной задачи гистопатологической сегментации. Набор данных COCO был выбран, поскольку он предоставляет богатый набор аннотированных изображений для обнаружения и сегментации объектов, что хорошо согласуется с задачей идентификации и сегментации определенных структур на скан-изображениях. Предварительно обученная модель была настроена с помощью исходной многозадачной функции потерь:

$$L = L_{cls} + L_{box} + L_{mask},$$

где L_{cls} — потеря классификации, L_{box} — потеря регрессии ограничивающего прямоугольника, а L_{mask} — потеря прогнозирования маски. Эта многозадачная функция потерь позволяет модели одновременно обучаться классификации объектов, уточнять их ограничивающие прямоугольники и генерировать точные маски сегментации. Для оптимизации был использован алгоритм стохастического градиентного спуска (SGD - Stochastic gradient descent) со скоростью обучения 0,003. SGD был выбран из-за его эффективности в обучении глубоких НС и его способности избегать локальных минимумов посредством стохастических обновлений. Конкретная скорость обучения 0,003 была определена путем эмпирической настройки, балансируя между скоростью сходимости и стабильностью. Эта относительно низкая скорость обучения помогает предотвратить перерегулирование во время оптимизации и позволяет более детально обновлять параметры модели.

На каждой из 20 перекрестных проверок модель Mask R-CNN была обучена в течение 300 эпох с размером батча 16. Размер батча был выбран для баланса вычислительной эффективности и производительности модели с учетом доступной памяти графического процессора и сложности задачи сегментации.

Чтобы улучшить возможности генерализации модели и учесть изменения в морфологии и окрашивании тканей, использовались обширные методы аугментации данных, которые включали геометрические преобразования (вращение, переворачивание, упругую деформацию, масштабирование), корректировки цвета и интенсивности (яркость, контрастность, насыщенность,

смещение оттенка), а также шум и размытие (гауссовский шум и гауссовское размытие).

Эти методы аугментации помогают модели научиться инвариантности к различным преобразованиям изображения и артефактам, которые могут присутствовать в скан-изображениях препаратов, тем самым повышая ее надежность и производительность на невидимых данных.

2.1.3 Оценка эффективности сегментации

Среднее значение точности (mAP – mean average precision) обычно используется для анализа производительности систем сегментации. MAP рассчитывается путем нахождения среднего значения точности (AP) для каждого класса сегментации, а затем усреднения по ряду классов.

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i,$$

где параметр N — это количество классов. Для расчета AP была построена кривая точности-полноты и вычислена площадь под ней.

2.1.4 Подготовка скан-изображений

Первичная сегментация ткани биоптатов была выполнена путем применения порогового значения интенсивности света к изображениям в оттенках серого. Для анализа всего скан-изображения использовалось скользящее окно, поскольку НС принимала в качестве входных данных небольшие фрагменты со стороны размером 50 микрометров (99 пикселей для скан-изображениях с увеличением x20). Следует отметить, что небольшие экземпляры на границе скользящего окна не учитывались, поскольку их классы были определены как нестабильные. Когда сканирование было завершено, экземпляры одного класса объединялись с сильным

пересечением. Выходным результатом модели была маска сегментации для трех классов («клубочки», «канальцы», «артерии»).

Все ткани, которые не были отнесены к этим классам, были определены как строма. В нормальных образцах почечной ткани строма имела небольшое пространство и окружала структуры в ткани почки на расстоянии примерно 15–20 микрометров (что эквивалентно $30 \times 30 - 40 \times 40$ пикселей на скан-изображениях при 20-кратном увеличении). Эти области были визуализированы как светло-голубые в маске сегментации и были исключены при расчете показателей фиброза и интерстициальной инфильтрации. В случаях отека пространство вокруг структур ткани превышало вычисленную относительную площадь стромы. Наоборот, в областях фиброза соединительная ткань замещала почечные структуры, уменьшая предположительно нормальную площадь стромы вокруг структур и приводя к высокой относительной площади фиброза. Поиск ядер и лимфоцитов проводился в пространстве стромы. Для идентификации лимфоцитов использовались алгоритм порогового разделения и разделения на основе водораздела. Круглые клетки отбирались путем расчета эллиптичности (эллипс рассчитывался для каждого объекта). Далее найденные объекты классифицировались как ядра или лимфоциты в зависимости от размера и цвета. Соотношение полуосей для лимфоцитов было меньше 2. Параметр цвета измерялся на серых изображениях по шкале от 0 до 255. Порог 95 применялся для средней яркости лимфоцитов. Далее, области, содержащие большое количество ядер и лимфоцитов, считались областями интерстициальной инфильтрации. Зонами интерстициальной инфильтрации считалась область размером 50×50 микрометров (99×99 пикселей для скан-изображений с увеличением $\times 20$), содержащая 2 лимфоцита и 20 ядер и более.

2.1.5 Морфологический анализ

Было проведено морфологическое исследование скан-изображений с наложенными на них масками сегментации в валидационном наборе данных ($n = 110$). Сегментация областей стромы анализировалась на предмет точных прогнозов и ошибок модели. Случаи с расхождением между экспертными оценками и автоматическими расчетами были изучены и описаны в разделе результатов. Для оценки обобщаемости модели использовали 50 скан-изображений из открытой базы данных СРТАС-ССРСС в качестве независимой внешней когорты [86]. На каждом скан-изображении была выбрана одна квадратная область интереса с неопухолевой почечной тканью с периметром 1024 микрометра для аннотации. В этих регионах интереса были вручную проаннотированы структуры «канальцы», «клубочки» и «артерии», как описано ранее. Результаты сегментации сравнивались с реальными данными, была оценена точность прогнозирования алгоритма с помощью коэффициента Дайса.

2.1.6 Статистический анализ

Статистический анализ количественных данных был выполнен с использованием GraphPad Prism версии 8.00 для Windows (GraphPad Software, Inc., Сан-Диего, Калифорния, США) и R v4.1.2 с библиотеками betareg v3.1–4, car v3.0–13 и emmeans v1.7.3. Для оценки различий между группами использовались модели бета-регрессии, а затем однофакторный дисперсионный анализ и апостериорный анализ Тьюки. Предположения модели проверялись с использованием библиотеки performance v0.9.0. Для сравнения характеристик был выполнен ROC-анализ, и был рассчитан J-индекс Юдена. Двумерная визуализация рассчитанных значений была выполнена с использованием библиотек Python v.3.10 и NumPy v.1.21.1 и pandas

v.1.4.1. Результаты статистического анализа были представлены в виде диаграмм типа violin. Статистическая значимость была установлена на уровне $p < 0,05$.

2.2 Применение компьютерного зрения для анализа препаратов рака почки

2.2.1 Подготовка набора данных

Для второй части исследования были использованы гистологические изображения и клинические данные пациентов из трех независимых источников: Атласа генома рака (TCGA), патоморфологической цифровой лаборатории UNIM (Москва, Россия) и Института клинической морфологии и цифровой патологии Сеченовского университета (Москва, Россия) [87]. Было отобрано только по одному гистологическому отсканированному препарату из каждого клинического случая. Внутренние наборы данных включали 49 образцов сПКР из UNIM и 83 образца из Института клинической морфологии и цифровой патологии. Образцы прошли стандартную гистологическую пробоподготовку: ткань опухоли была зафиксирована в 10%-ом нейтральном формалине, залита в парафин и нарезана при помощи микротомы на срезы толщиной 2-3 мкм. Далее срезы были окрашены гематоксилином и эозином и оцифрованы на увеличении $\times 40$ с помощью сканера Leica Aperio AT2 (Leica Microsystems, Вецлар, Германия) для получения скан-изображений (WSI). Образцы были анонимизированы и не содержали никакой клинической информации.

Кроме того, в качестве внешнего набора данных, было использовано 105 случаев сПКР из TCGA, из которых 70 случаев - для оценки прогностической ценности расчетов модели. Были загружены исходные скан-изображения препаратов с сПКР и клиническая информация для этих случаев с помощью программного обеспечения GDC Data Transfer Tool. В общей сложности для

анализа было отобрано 167 скан-изображений сПКР со степенью злокачественности от 1 до 4 по системе Фурман. Эти скан-изображения были пересмотрены и реклассифицированы в соответствии с системой WHO/ISUP. Изображения просматривались вслепую в формате файла .svs с использованием программного обеспечения Case Viewer (3DHistech, Венгрия). После применения критериев включения и исключения из исследования, окончательный набор данных включал 144 пациента с сПКР различной степени злокачественности. Критериями включения были: 1) Возраст пациента старше 18 лет; и 2) подтвержденный диагноз сПКР. Критериями исключения были: 1) наличие других диагнозов, обозначенных буквой «С» в Международной классификации болезней 10-го пересмотра (МКБ-10); 2) наличие артефактов на изображениях; и 3) низкое качество подготовки образцов. Дизайн исследования представлен на Рисунке 5.

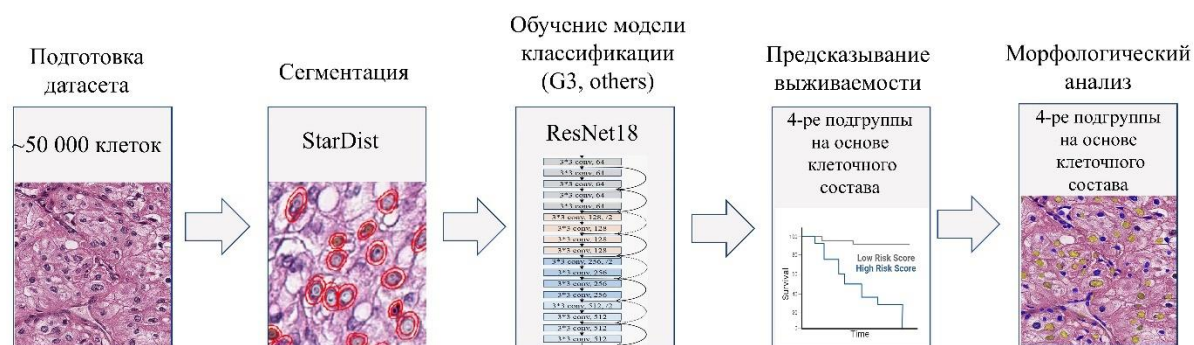


Рисунок 5 – Дизайн изучения моделей компьютерного зрения для определения степени злокачественности сПКР

Распределение скан-изображений по классам представлено в таблице 2.

Таблица 2 - Характеристика набора данных

Степень злокачественности	TCGA (внешний набор данных)	UNIM (внутренний набор данных)	Институт клинической морфологии и цифровой патологии (внутренний набор данных)
Обучающий набор			
1	1	-	5
2	7	-	7
3	1	-	4
4	3	-	1
Итого	12	-	17
Предсказывание степени злокачественности (валидационный набор)			
1	1	6	20
2	12	13	28
3	2	14	10
4	6	1	2
Итого	21	34	60
Предсказывание выживаемости (верификационный набор)			
1	-	-	-
2	46	-	-
3	7	-	-
4	17	-	-
Итого	70	-	-

2.2.2 Архитектура модели сегментации

Модель Stardist была использована для конвейера сегментации-классификации из-за ее способности идентифицировать границы ядер на слайд-изображениях, окрашенных гематоксилином и эозином [88,89]. Эта модель предсказывает расстояние (r) до границы объекта по фиксированному набору лучей и вероятность (d) появления каждого объекта, создавая полный набор возможных полигонов-кандидатов для данного изображения. Для достижения конечного результата сегментации к этим кандидатам применяется немаксимальное

подавление. Плагин Stardist был установлен в программное обеспечение QuPath версии 0.4.3 для помощи экспертам при аннотировании скан-изображений. Затем сегментированные ядра и аннотации экспертов были использованы для обучения классификационной модели ResNet18 [88,89].

2.2.3 Аннотирование и обучение модели классификации

Для обучения модели классификации было использовано 12 сканов из когорты TCGA и 17 скан-изображений из Института клинической морфологии и цифровой патологии. Области интереса были выбраны с помощью программы QuPath v.0.4.3 и экспортированы в виде файлов .json. Эти области включали опухолевую ткань, содержащую как опухолевые, так и стромальные клетки, кроме областей диффузного некроза и областей с плохой фокусировкой во время сканирования. Отдельные «размытые» клетки, часто вызванные артефактами микротомии и неоптимальным фокусом в гистологических сканерах, обычно наблюдались в скан-изображениях, полученных при помощи различных сканеров, и считались нормальным явлением. В каждом скан-изображении было выделено 10-15 квадратов со стороной 500 мкм (областей интереса) для дальнейшей аннотации в них. Области интереса случайным образом были разделены на обучающие, валидационные и тестовые наборы, в результате чего было получено 277, 51 и 32 уникальных областей интереса соответственно. Используя алгоритм Stardist были определены границы полигонов клеток, было извлечено каждое изображение клетки и использовано совместно со связанными с ними метки для обучения модели классификации ResNet18. Для обучения и валидации в 328 областях интереса клетки в опухоли были отнесены к одному из девяти классов: G1+, G1-, G2+, G2-, G3+, G3-, G4+, G4- и стромальные клетки. К классу G1+ относились клетки с небольшими, четко выраженными ядрами без видимых ядрышек, а клетки со слегка размытыми, округлыми ядрами, где ядрышки не были видны, к G1-. Клетки с эозинофильным, округлым ядрышком, заметным при

увеличении от $\times 400$ до $\times 100$ в центре ядра, были классифицированы как G2+. Клетки с размытым, базофильным ядрышком, видимым при том же увеличении, были отнесены в класс G2-. Клетки G3+ имели крупные базофильные ядра с четкими или слегка неровными круглыми темно-фиолетовыми или светлыми эозинофильными ядрышками, заметными при увеличении $\times 100$. Если ядрышки имели звездчатую форму или размытые границы, эти клетки были классифицированы как G3-. Гигантские многоядерные опухолевые клетки и клетки с рабдоидными или саркоматоидными признаками были помечены как G4+ или G4- на основе их формы и фокуса. Все клетки внутри областей интереса были проаннотированы с помощью программного обеспечения QuPath v.0.4.3.

В результате был подготовлен набор данных на основе классификации WHO/ISUP и адаптирован для включения широкого спектра клеток, которые наблюдались в срезе ткани опухоли. Клетки класса G4+ и G4- были исключены из обучения из-за их низкого содержания, в среднем, менее 10 клеток в одной области интереса в ткани опухоли со степенью злокачественности 4. Ошибки классификации в основном возникали между соседними классами, такими как G2-/G3-, G1-/G2- и G2+/G3-.

Для тестового набора была применена упрощенная аннотация, состоящая из четырех классов (G1+, G2+, G3+ и стромальные клетки) в 32 областях интереса, чтобы оценить как модель, обученная исключительно на «идеальных» изображениях клеток (исключая клетки G1-, G2-, G3-, G4+ и G4-), выполняла предсказание и идентификацию клеток класса G3+. Ключевой задачей в этой оценке была классификация между клетками G3+ и другими (others), которые включали все стромальные и опухолевые клетки без выраженных ядрышек. В таблице 3 демонстрируется распределение клеток разных классов в обучающем, валидационном и тестовом наборах изображений клеток.

Таблица 3 – Распределение клеток по классам в обучающем, валидационном и тестовом наборах

Класс	Обучающий набор	Валидационный набор	Тестовый набор
	Количество клеток, %	Количество клеток, %	Количество клеток, %
G1+	5420; 11,82	816; 12,55	1101; 14,72
G1-	1701; 3,71	201; 3,09	-
G2+	6295; 13,73	764; 11,75	4452; 3,97
G2-	23122; 50,42	3016; 46,40	-
G3+	1075; 2,34	127; 1,95	271; 65,19
G3-	3728; 8,13	487; 7,49	-
Стромальные клетки	4514; 9,84	1089; 16,75	1005; 16,12

Во время обучения каждой клетке назначался класс, а результаты сегментации и классификации были в дальнейшем оценены для обеспечения точности алгоритма. В окончательном алгоритме использовалось 41 341 клетка для обучения, 5 411 клеток для валидации и 5 824 клетки для тестирования, включая 271 клетку типа Grade 3. Процесс обучения включал использование предварительно обученной модели ResNet18 для классификации клеток по двум категориям: Grade 3 и Others. Все параметры модели использовались во время обучения. Обучение длилось 100 эпох с ранней остановкой, если качество проверки не улучшалось в течение 10 последовательных эпох. Выбиралась модель из эпохи с наивысшим качеством проверки. Для расчета точности, специфичности и полноты обученных моделей имелось два набора опухолевых клеток. Метриками оценки для классификации были F1-мера и абсолютная разница между точностью и полнотой. Был использован оптимизатор Adam с параметрами: скорость обучения = 0,002, снижение веса = $1e-4$ и размером батча = 128. Аугментация изображений, такая как изменение цвета и размытие, использовалась для улучшения результатов, но не во время тестирования.

Подводя итог, данная обученная модель использовала области интереса на скан-изображениях сПКР в качестве входных данных, сегментировала ядра клеток, использовала маски для извлечения клеток из исходного изображения и классифицировала каждую клетку. Выходные данные включали список координат, классов и вероятностей для каждой клетки, которые можно было импортировать в программное обеспечение для анализа изображений, например QuPath.

2.2.4 Верификация

Для верификации было использовано 115 скан-изображений из трех наборов данных: TCGA ($n = 21$), UNIM ($n = 34$) и Института клинической морфологии и цифровой патологии ($n = 60$). На каждом слайде была выделена квадратная область интереса с периметром 4000 мкм и содержащая от 400 до 1200 клеток. Данные области были проанализированы с наложенными на них масками сегментации, сгенерированными моделью глубокого обучения. Далее в этих зонах опухолевые клетки были классифицированы на два класса: G3 и others.

Специфичность и чувствительность прогнозов модели глубокого обучения оценивались с использованием кривой рабочих характеристик приемника (ROC - Receiver operating characteristic) и площади под ROC-кривой (AUC ROC - Area Under the ROC Curve).

Статистический анализ экспериментальных количественных данных проводили с использованием программного обеспечения GraphPad Prism версии 10.2.3.403 для Windows (GraphPad Software, Inc., Сан-Диего, Калифорния, США) и R версии 4.1.2, с библиотеками betareg версии 3.1-4, car версии 3.0-13 и emmeans версии 1.7.3. Для сравнения характеристик был проведен ROC-анализ и рассчитан J-индекс Юдена (Youden). Количество клеток G3 в 1 мм² и процент клеток G3 анализировались методом однофакторного дисперсионного анализа с последующим множественным сравнительным тестом Тьюки. Результаты статистического анализа были представлены в виде диаграмм типа boxplot («ящик

с усами»). Двумерная визуализация вычисленных значений была выполнена с использованием библиотек Python версии 3.10, NumPy версии 1.21.1 и pandas версии 1.4.1. Статистическая значимость была установлена на уровне $p < 0,05$.

2.2.5 Прогнозирование выживаемости

Для оценки выживаемости было использовано 70 скан-изображений с сопроводительной клинической информацией из открытой базы данных онкологических пациентов TCGA. В выборку вошли 20 женщин и 50 мужчин, средний возраст составил $61,14 \pm 12,54$ года, стадии заболевания по TNM распределились следующим образом: 1-я наблюдалась у 27 пациентов (38,57%), 2-я – у 6 пациентов (8,57%), 3-я – у 17 пациентов (24,28%), 4-я – у 18 пациентов (28,58%). Образцы были разделены на две группы: 35 пациентов, которые умерли в течение 5-ти лет после постановки диагноза, и 35 пациентов, которые были живы более 5-ти лет. В набор данных были включены пациенты с сПКР 2-й, 3-й и 4-й степенями злокачественности. Скан-изображения были реклассифицированы в соответствии с системой WHO/ISUP. Далее они были проанализированы при помощи обученной модели для разделения клеток на два класса (G3 и others).

Из каждого слайда была выделена квадратная область интересов с периметром 4000 мкм, содержащая от 400 до 1200 клеток. Чтобы оценить взаимосвязь между клеточным составом опухоли и выживаемостью пациентов, данные были разделены на подгруппы, используя два пороговых значения для переменных G3 и others, определенных с помощью J-индекса Юдена. Это позволило сравнивать выживаемость пациентов с разным клеточным составом в опухоли. Пороговые значения составили 533 и 4133 для классов G3+ и others соответственно. На основе этих пороговых значений было сформировано четыре подгруппы, соответствующие четырем квадрантам: (1) G3+/others-; (2) G3+/others+; (3) G3-/others-; (4) G3-/others+. В каждом квадранте выживаемость оценивалась с использованием метода Каплана-Мейера с логарифмическим критерием (тест

Мантела-Кокса). Распределение выборок пациентов по полученным квадрантам в соответствии со степенью злокачественности оценивались с помощью точного критерия Фишера.

ГЛАВА 3. ПРИМЕНЕНИЕ КОМПЬЮТЕРНОГО ЗРЕНИЯ ДЛЯ СЕГМЕНТАЦИИ ПОЧЕЧНЫХ СТРУКТУР И ДИАГНОСТИКИ ОТТОРЖЕНИЯ ТРАНСПЛАНТАТОВ ПОЧКИ

3.1 Эффективность многоклассовой сегментации

В паренхиме биоптатов почки структуры были проаннотированы с присвоением им следующих классов: «канальцы», «клубочки» и «артерии». Оставшаяся ткань относилась в 4-й класс «стромы». Коэффициент Дайса для предсказания экспертных аннотаций без присвоения конкретного класса составил 0,9065 (Q2, Min-Max: 0,8291 – 0,9297). Для визуализации производительности модели была использована матрица ошибок. Каждый столбец матрицы представлял образцы в фактическом (истинном) классе, в то время как каждое необработанное поле представляло образцы в предсказанном классе (Рисунок 6).

		GROUND TRUTH			
		Канальцы	Клубочки	Артерии	Строма
ПРЕДСКАЗЫВАНИЕ	Канальцы	0.815000	0.000000	0.003000	0.180000
	Клубочки	0.000000	0.962000	0.000000	0.037000
	Артерии	0.176000	0.000000	0.676000	0.147000

Рисунок 6 – Матрица ошибок при сегментации

Объемы классов были приблизительно соотнесены как 40:3:2 («канальцы», «клубочки» и «артерии»). Небольшая ошибка 1-го типа 0,18 наблюдалась в классе

«канальцы» в пользу «стромы», а также ошибки 1-го типа 0,176 и 0,147 в классе «артерии» в пользу «канальцы» «и «строма». Лучшим сегментированным классом были «клубочки» (точность 0,962). Производительность модели позволила получить высококачественный прогноз. Обученная нейронная сеть продемонстрировала значение средней точности $m [AP] ([0,5:0,95]) = 0,74$ при сегментации на три класса («канальцы», «клубочки» и «артерии»). Коэффициент Дайса измерялся для классов, которые были сегментированы моделью ИИ: канальцы - 0,8310, клубочки - 0,9242, артерии - 0,7087, строма - 0,7919.

3.2 Фиброз и интерстициальная инфильтрация в биоптатах трансплантатов

Полупрозрачные маски сегментации накладывались на исходные сканирования биоптатов, что позволило оценить точность прогнозов класса тканей с морфологической точки зрения (Рисунок 7).

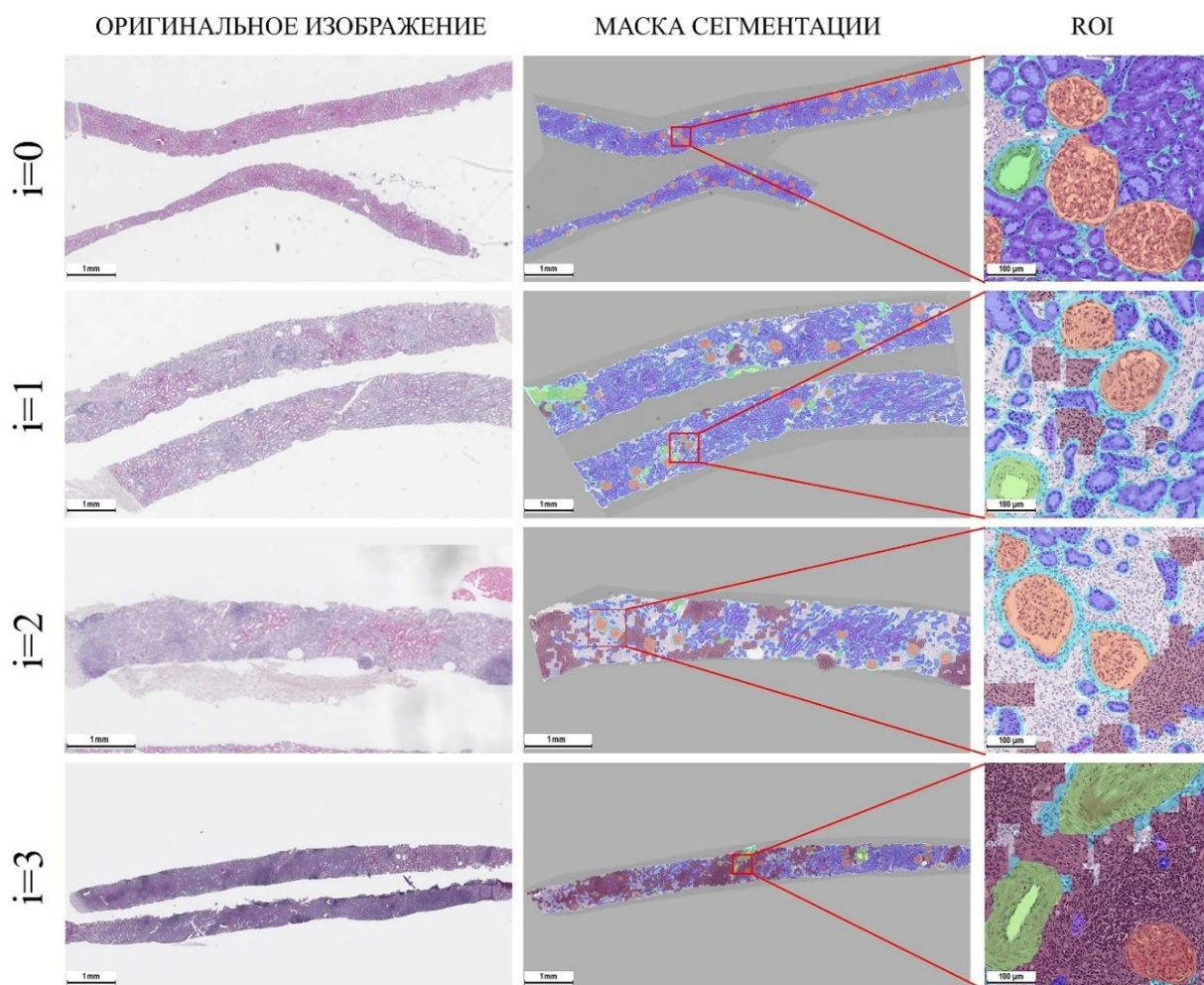


Рисунок 7 - Сегментация структур почечной ткани на скан-изображениях биоптата и в выделенных областях интереса

Стоит отметить, что не было существенных различий в качестве масок сегментации между скан-изображениями, сделанными при разном увеличении. Клубочки (окрашенные на маске скан-изображения оранжевым цветом) были успешно сегментированы в их истинных границах, которые включали как сосудистую сеть, так и капсулу Боумена, и представляли собой овальные структуры с приблизительным диаметром 100–150 мкм. Интересно, что модель ИИ сегментировала даже некоторые клубочки незначительного размера.

Почечные канальцы (синего цвета) представляли собой плотно упакованные пучки полигональных структур. Они сливались в более крупные элементы маски сегментации в некоторых случаях, но чаще имели приблизительный размер 30–50

мкм. Канальцы обнаруживались с довольно хорошим результатом, даже когда они были расположены в областях фиброза и интерстициальной инфильтрации.

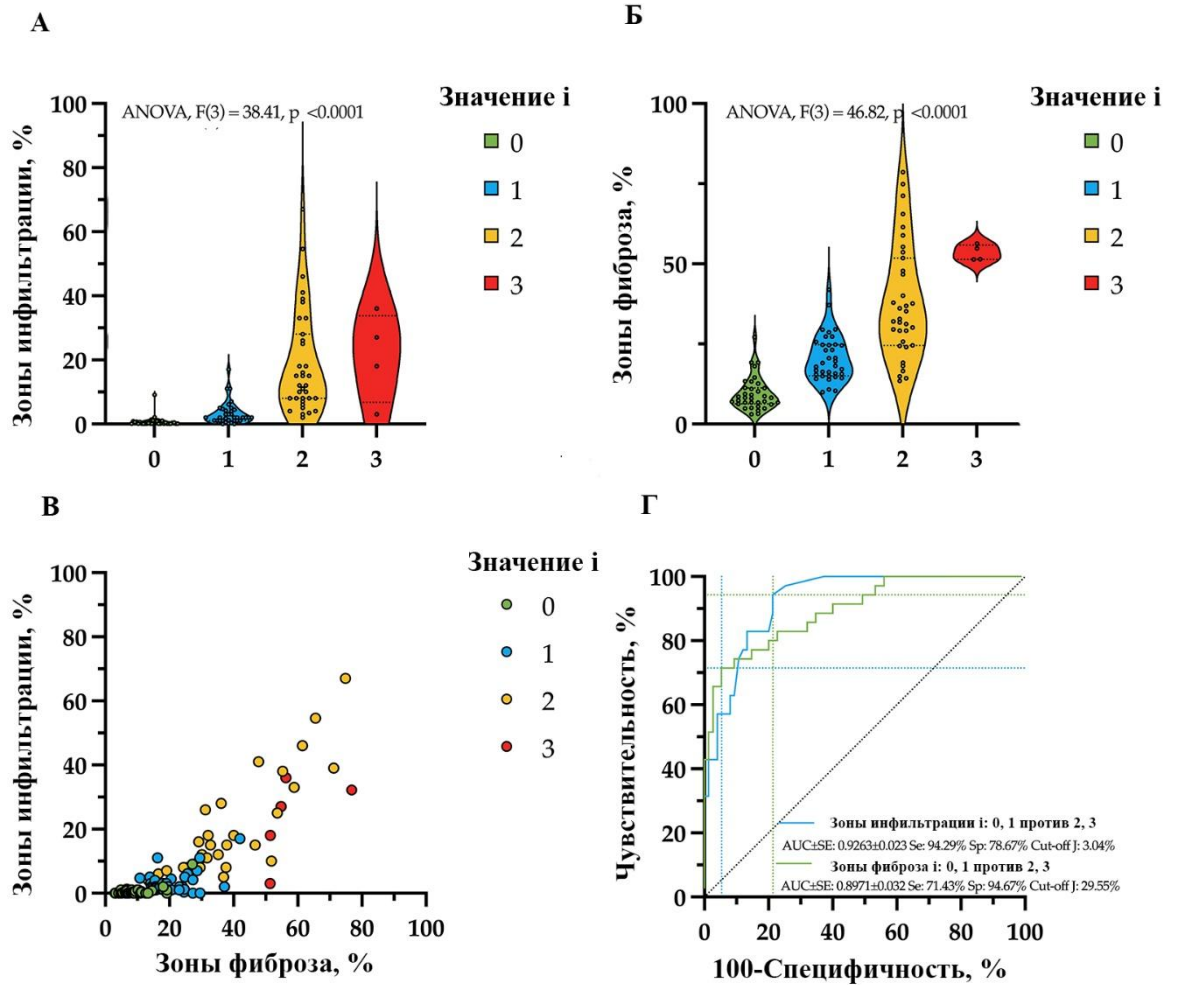
Артерии (окрашенные зеленым цветом) обнаруживались в их внешних границах вместе с просветом. Программа детектировала как мелкие (размером около 100 мкм), так и более крупные артерии, некоторые из которых были только частично представлены в биоптатах.

Строма (серого цвета) определялась как пространство ткани, расположенное между всеми структурами, упомянутыми выше. Важно отметить, что существует нормальный объем стромальной ткани, состоящий из тонких пучков коллагеновых волокон и мелких капилляров, которые образуют пространство вокруг структур паренхимы. Это пространство не учитывалось при оценке фиброза и было окрашено на маске сегментации светло-голубым цветом.

Интерстициальная инфильтрация (коричневого цвета) являлась частью стромы, которая удовлетворяла критериям как морфологии лимфоидных клеток, так и их плотности в регионе. В группе $i = 0$ в этих зонах лимфоциты практически отсутствовали и постепенно увеличивались на относительной площади в более клинически неблагоприятных группах.

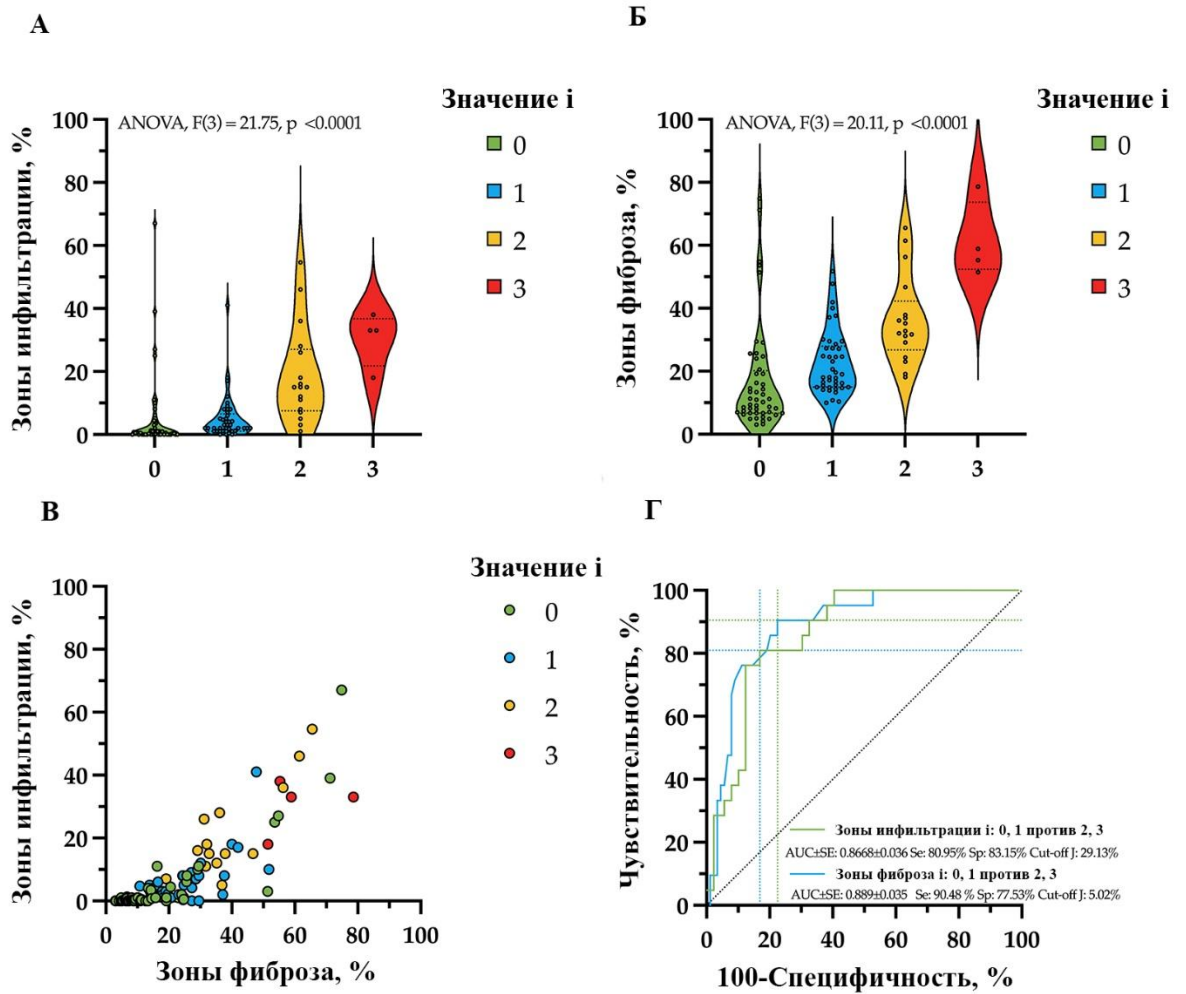
3.3 Сравнение модели компьютерного зрения с экспертными оценками согласно Vanff-классификации

Была обнаружена сильная положительная нелинейная корреляция между измеренными областями инфильтрации и фиброзом (коэффициент Спирмена $r = 0,8212$, $p < 0,0001$, рисунок 8, 9).



А, Б - Сравнительный анализ рассчитанных измерений фиброза и интерстициальной инфильтрации, модель бета-регрессии, дисперсионный анализ (ANOVA), график типа violin с точечными медианами и 25, 75 % процентилем. В - Распределение рассчитанных параметров по баллам. Г - Анализ рабочей характеристики приемника (ROC) рассчитанных параметров и оценка отсечки для баллов

Рисунок 8 - Статистический анализ рассчитанных относительных площадей фиброза и интерстициальной инфильтрации для экспертных оценок интерстициальной инфильтрации (i)



А, Б - Сравнительный анализ рассчитанных измерений фиброза и интерстициальной инфильтрации, модель бета-регрессии, дисперсионный анализ (ANOVA), график типа violin с точечными медианами и 25, 75 % процентилем. В - Распределение рассчитанных параметров по баллам. Г - Анализ рабочей характеристики приемника (ROC) рассчитанных параметров и оценка отсечки для баллов

Рисунок 9 - Статистический анализ рассчитанных относительных площадей фиброза и интерстициальной инфильтрации для экспертных оценок фиброза (s_i)

Также обнаруживалась значительная связь между показателем s_i и площадью инфильтрации (ANOVA, $df = 3$, $F = 21,75$, $p < 0,0001$, таблицы 4 и 5, рисунок 8), а также между показателем s_i и площадью фиброза (ANOVA, $df = 3$, $F = 20,11$, $p < 0,0001$, таблицы 4 и 5, рисунок 8).

Таблица 4 - Описательная статистика по измеренным площадям интерстициальной инфильтрации и фиброза в соответствии с баллами по Vanff-классификации

Значение Измеряемая площадь	0	1	2	3	Значение p (ANOVA)
сi	n=48	n=41	n=17	n=4	Всего 110
инфильтрация	3.82 (2.43– 5.21)	6.46 (4.27– 8.64)	18.83 (12.61– 25.05)	32.64 (16.43– 48.85)	<0.0001
Фиброз	16.54 (13.51– 19.57)	24.19 (20.32– 28.06)	36.86 (29.97– 43.75)	60.67 (46.27– 75.08)	<0.0001
i	n=35	n=36	n=35	n=4	Всего 110
инфильтрация	2.09 (1.24– 2.95)	4.44 (2.87– 6.01)	17.58 (13.91– 21.25)	19.57 (8.48– 30.67)	<0.0001
фиброз	10.98 (8.61– 13.35)	21.55 (18.33– 24.78)	37.52 (33.6– 41.45)	53.26 (41.27– 65.26)	<0.0001
Модель бета-регрессии со значениями μ (95% ДИ); жирным шрифтом выделены значения p, значимые при $p < 0,05$					

Таблица 5 – Сравнение средних различий в площадях инфильтрации и фиброза по баллам Vanff-классификации (95% ДИ)

Значение/ Измеряемая площадь	Значение 1 в сравнении с 0	Значение р	Значение 2 в сравнении с 0	Значение р	Значение 2 в сравнении с 1	Значение р	Значение 3 в сравнении с 0	Значение р	Значение 3 в сравнении с 1	Значение р	Значение 3 в сравнении с 2	Значение р
сi												
инфильтрация	2.63 (0.36 – 4.33)	0.066 4	15.01 (10.73– 20.24)	<0.001	12.38 (8.19– 18.08)	0.007	28.81 (2.96– 32.01)	0.008	26.18 (0.54– 29.72)	0.009	13.81 (- 13.21– 17.2)	0.398
фиброз	7.65 (1.35 – 13.95)	0.008	20.32 (10.51– 30.14)	<0.001	12.67 (2.35– 23)	0.007	44.14 (24.82– 63.46)	<0.001	36.49 (16.92– 56.06)	<0.001	23.81 (2.87– 44.75)	0.0183

Продолжение Таблицы 5

i												
инфильтрация	2.35 (0.36– 4.33)	0.012	15.49 (10.73– 20.24)	<0.01	13.14 (8.19– 18.08)	<0.01	17.48 (2.96– 32.01)	0.010	15.13 (0.54– 29.72)	0.038	2 (- 13.21– 17.2)	0.9 86 8
фиброз	10.58 (5.42– 15.73)	<0.001	26.55 (20.58– 32.52)	<0.001	15.97 (9.34– 22.6)	<0.001	42.29 (26.26– 58.31)	<0.001	31.71 (15.43– 47.99)	<0.001	15.74 (- 0.8– 32.28)	0.0 69 0
Апостериорный критерий Тьюки; значения представляют собой среднюю разницу (95% ДИ); жирным шрифтом выделены значения p, значимые при p<0,05												

Было показано, что увеличение показателя s_i было достоверно связано с увеличением площади инфильтрации (таблицы 3 и 4, рисунок 7): показатель $s_i 0$ – 3,82 % (95 % ДИ, 2,43–5,21 %); показатель $s_i 1$ – 6,46 % (95 % ДИ, 4,27–8,64 %, значение p [$s_i 1$ в сравнении $s_i 0$] = 0,0664); показатель $s_i 2$ – 18,83 % (95 % ДИ, 12,61–25,05 %, значение p [$s_i 2$ в сравнении $s_i 1$] = 0,0007); показатель $s_i 3$ – 32,64 % (95 % ДИ, 16,43–48,85 %, значение p [$s_i 3$ в сравнении $s_i 2$] = 0,3988). Увеличение показателя s_i также было связано с большей площадью фиброза (таблицы 4 и 5, рисунок 7): показатель $s_i 0$ – 16,54 % (95 % ДИ, 13,51–19,57 %); показатель $s_i 1$ – 24,19 % (95 % ДИ, 20,32–28,06 %, значение p [s_i в сравнении с $s_i 0$] = 0,0098); показатель $s_i 2$ – 36,86 % (95 % ДИ, 29,97–43,75 %, значение p [$s_i 2$ в сравнении с $s_i 1$] = 0,0087); показатель $s_i 3$ – 60,67 % (95 % ДИ, 46,27–75,08 %, значение p [$s_i 3$ в сравнении с $s_i 2$] = 0,0183).

При ROC-анализе (Рисунок 8) была продемонстрирована высокая способность прогнозирования оценки $s_i 0$ и 1 по сравнению с 2 и 3 на основе измеренных параметров площади инфильтрации (AUC = 0,889, Чувствительность = 90,48 %, Специфичность = 77,53 %, J-индекс Юдена = 5,02 %) и площади фиброза (AUC = 0,8669, Чувствительность = 80,95 %, Специфичность = 83,15 %, J-индекс Юдена = 29,19 %).

Существовала значимая связь между показателем i и площадью инфильтрации (ANOVA, $df=3$, $F=38,41$, $p < 0,0001$, таблицы 4 и 5, рисунок 5), а

также между показателем i и площадью фиброза (ANOVA, $df=3$, $F=46,82$, $p < 0,0001$, таблицы 2 и 3, рисунок 8B).

Увеличение показателя i было достоверно связано с увеличением площади инфильтрации (таблицы 4 и 5, рисунок 9): показатель i_0 – 2,09 % (95 % ДИ, 1,24–2,95 %); показатель i_1 – 4,44 % (95 % ДИ, 2,87–6,01 %, значение $p [i_1 \text{ в сравнении } i_0] = 0,0127$); показатель i_2 – 17,58 % (95 % ДИ, 13,91–21,25 %, значение $p [i_2 \text{ в сравнении } i_1] < 0,0001$); показатель i_3 – 19,57 % (95 % ДИ, 8,48–30,67 %, значение $p [i_3 \text{ в сравнении } i_2] = 0,9868$). Увеличение показателя i также было связано с большей площадью фиброза (таблицы 4 и 5, рисунок 9): показатель i_0 – 10,98 % (95 % ДИ, 8,61–13,35 %); показатель i_1 – 21,55 % (95 % ДИ, 18,33–24,78 %, значение $p [i_1 \text{ в сравнении } i_0] < 0,0001$); показатель i_2 – 37,52 % (95 % ДИ, 33,6–41,45 %, значение $p [i_2 \text{ в сравнении } i_1] < 0,0001$); показатель i_3 – 53,26 % (95 % ДИ, 41,27–65,26 %, значение $p [i_3 \text{ в сравнении } i_2] = 0,0690$).

При ROC-анализе (Рисунок 9) была продемонстрирована высокая способность прогнозирования значения i_0 и i_1 против i_2 и i_3 на основе измеренных параметров области инфильтрации (AUC = 0,9263, Чувствительность = 94,29 %, Специфичность = 78,67 %, J-индекс Юдена = 3,04 %) и области фиброза (AUC = 0,8971, Чувствительность = 71,43 %, Специфичность = 94,67 %, J-индекс Юдена = 29,55 %).

Чтобы оценить производительность модели, ее точность сравнивалась на скан-изображениях биоптатов с экспертными аннотациями из внешнего набора данных. Коэффициенты Дайса для различных классов аннотаций были следующими: «канальцы»: 0,8612, «клубочки»: 0,9567, «артерии»: 0,7036 и фон («строма»): 0,7033. Сравнивая коэффициенты Дайса между внутренним и внешним наборами данных, становится очевидным, что модель стабильно работает в разных наборах данных, с немного более высокими коэффициентами Дайса для аннотаций структур паренхимы почки во внутреннем наборе данных. Рисунок 10 демонстрирует эффективность модели в точной сегментации структур в почечной ткани.

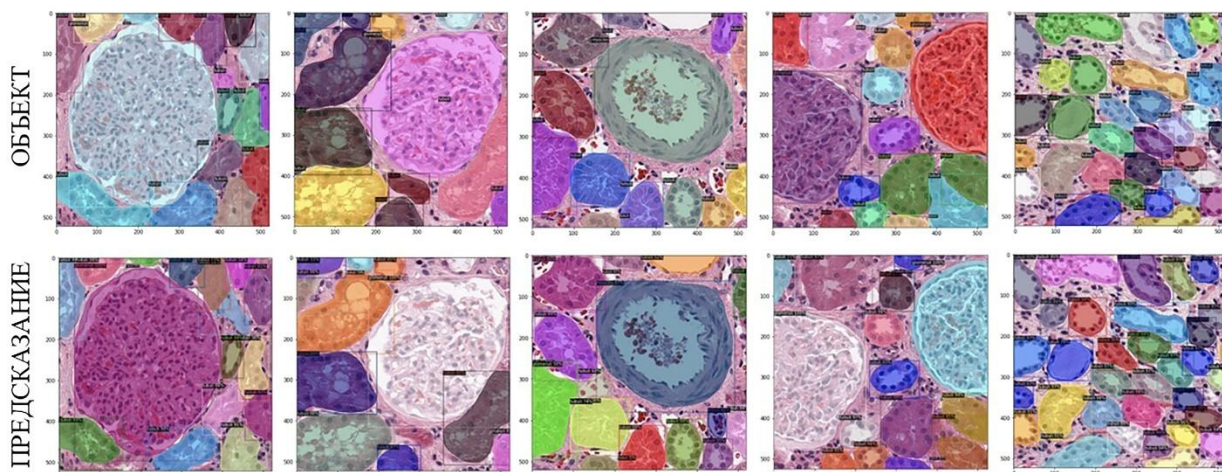


Рисунок 10 – Эффективность работы модели сегментации структур почечной паренхимы

Сравнение между истинными данными и прогнозом имело высокую степень точности, при этом модель практически полностью соответствовала экспертным аннотациям. Несмотря на присутствие незначительных расхождений, не было отмечено значительного снижения общей производительности алгоритма. Примечательно, что алгоритм точно идентифицировал измененные структуры, включая каналы с признаками дистрофии и сосуды с небольшим гиалинозом. Визуальная проверка подтверждает количественные показатели, например, коэффициент Дайса, при оценке эффективности модели. Кроме того, матрица ошибок на Рисунке 11 иллюстрирует производительность модели при различении классов.

МАТРИЦА ОШИБОК

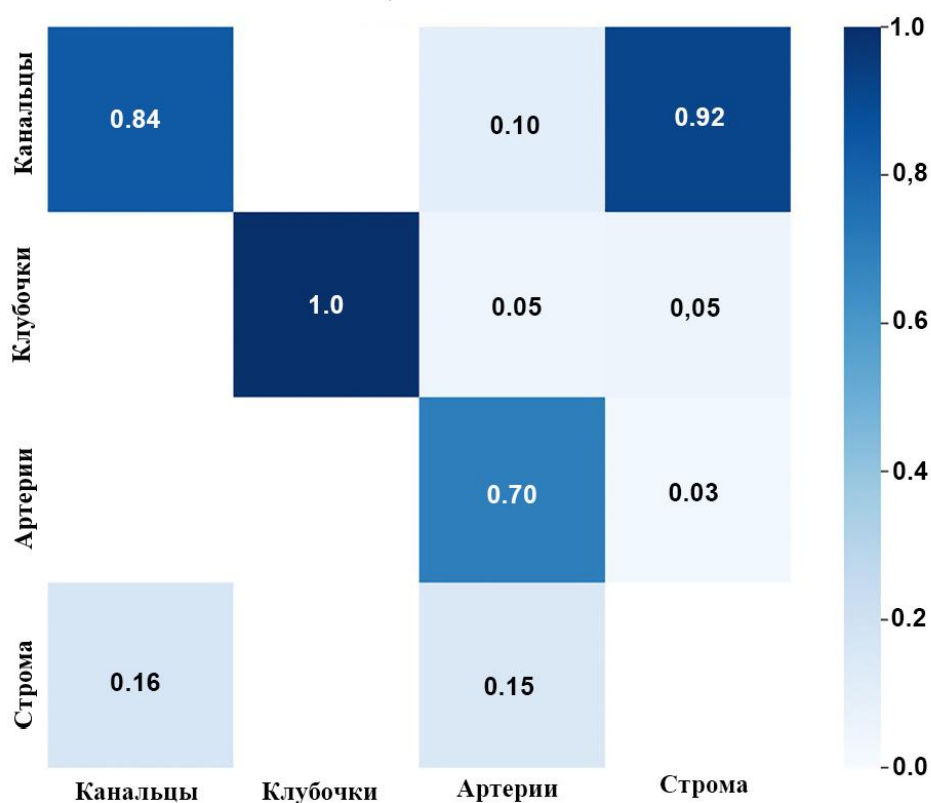


Рисунок 11 – матрица ошибок при сегментировании структур паренхимы почки

Сравнение с матрицей ошибок внутреннего набора данных демонстрирует работу алгоритма в различных наборах данных. Обнаружение класса «клубочки» довольно высоким, в то время как другие классы имели умеренную вариабельность.

ГЛАВА 4. ПРИМЕНЕНИЕ КОМПЬЮТЕРНОГО ЗРЕНИЯ ДЛЯ ОЦЕНКИ СТЕПЕНИ ЗЛОКАЧЕСТВЕННОСТИ РАКА ПОЧКИ

4.1 Эффективность модели классификации

На скан-изображениях рака почки различных степеней злокачественности была проверена производительность модели, а также проводилось сравнение ее точности с экспертными оценками на двух валидационных и тестовых наборах. Статистические метрики для валидационного набора: точность (accuracy): 0,88; точность (precision): 0,60; полнота: 0,51; F1-мера: 0,56. Для четырехклассового тестового набора были получены следующих метрики: точность (accuracy): 0,96; точность (precision): 0,87; полнота: 0,87; F1-мера: 0,87 (Рисунок 12).



Рисунок 12 - Матрица ошибок валидации классификации клеток G3

Алгоритм неправильно классифицировал некоторые опухолевые клетки как Grade 3. Однако, когда эти клетки позже были проанализированы при изучении препаратов, их отнесли преимущественно в класс G3 (Рисунок 13).

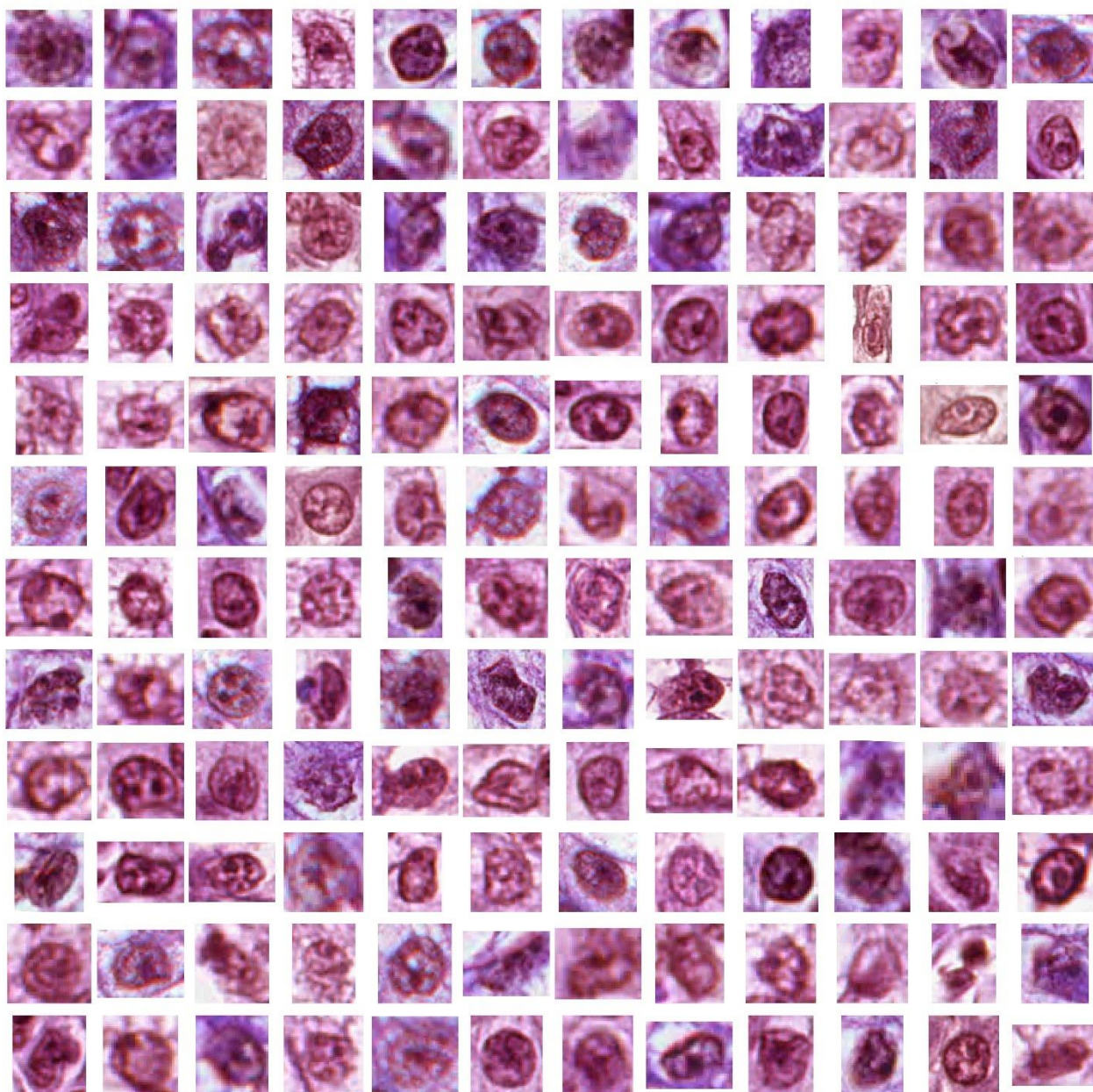
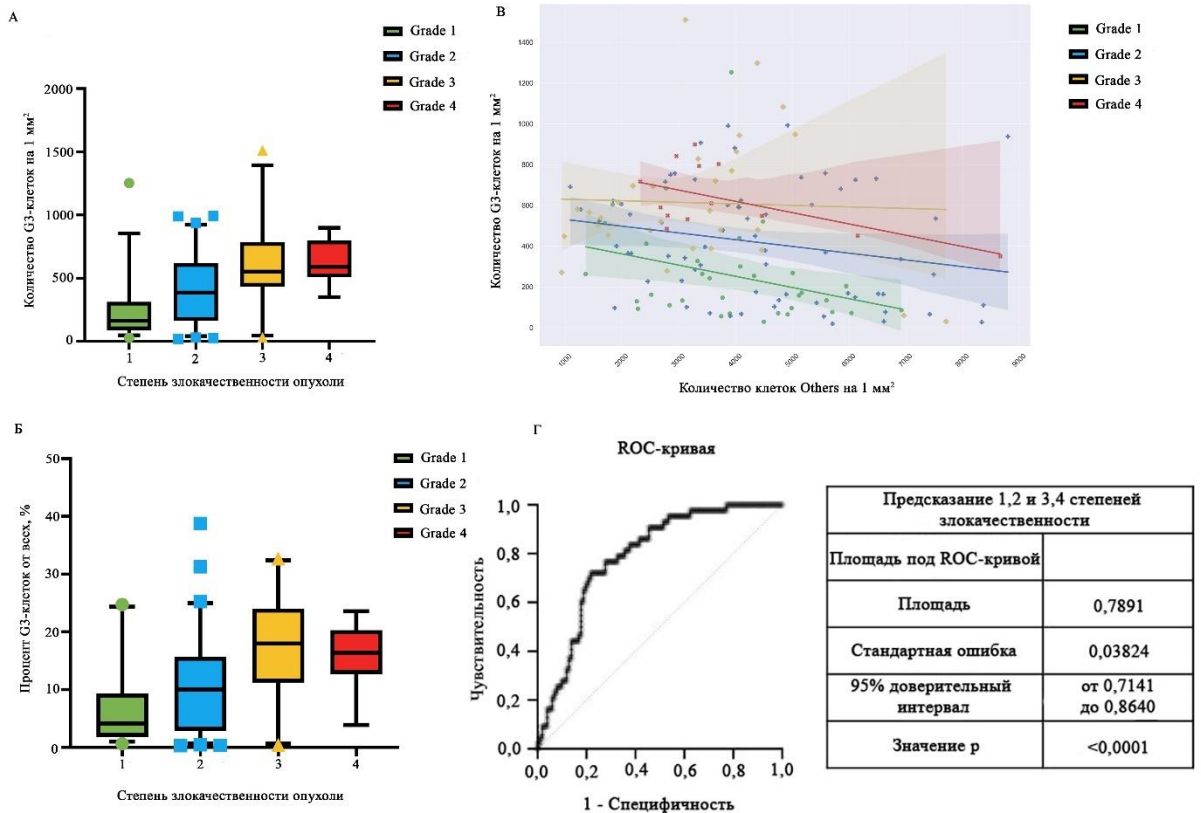


Рисунок 13 - Изображения клеток, которые модель ошибочно отнесла к G3

При помощи рассчитанных метрик было показано, что модель способна надежно идентифицировать клетки с выраженными ядрышками, что явилось причиной для исследования прогностического значения этих метрик.

4.2 Верификация модели классификации

Была исследована корреляция между отношением клеток G3 к общему числу клеток, а также экспертной общей степенью злокачественности опухоли. Анализ общего количества клеток и процента клеток G3 от него выявил статистически значимые различия (общее число клеток G3: 1-я степень против 2-й, 3-й и 4-й степени, 2-я степень против 3-й степени; процент клеток G3: 1-я степень против 3-й и 4-й степени, 2-я степень против 3-й степени) или тенденции (общее число клеток G3: 2-я степень против 4-й степени, $p=0,06$; процент клеток G3: 1-я степень против 2-й степени, $p=0,18$, 2-я степень против 4-й степени, $p=0,12$) между всеми группами, кроме 3-й и 4-й степени (рисунок 13). В частности, наблюдалось увеличение содержания клеток G3 в группах от 1-й степени до 3-й степени. Были определены пороговые значения для клеток типа G3 на 1 мм^2 в ткани светлоклеточного почечно-клеточного рака от 1 до 4 степеней злокачественности: $254,76 \pm 246,96$ ($7,08 \pm 0,07 \%$); $418,83 \pm 247,22$ ($10,43 \pm 0,081 \%$); $612,9 \pm 323,90$ ($17,93 \pm 0,082 \%$) и $628,46 \pm 167,71$ ($15,69 \pm 0,057 \%$). Стоит отметить, что в опухолях 4-й степени доля клеток с видимыми ядрышками увеличивалась незначительно. Эта тенденция связана с тем, что по мере прогрессирования опухолей до 4 степени злокачественности происходит сдвиг в клеточной морфологии. В сПКР со степенью злокачественности 4 часто встречаются такие особенности, как причудливое ядро, деформация ядра опухолевой клетки, саркоидная или рабдоидная дифференцировка и образование гигантских клеток. Эти морфологические изменения усложняют процесс классификации, поскольку эти атипичные клетки не вписываются в категории, используемые для опухолей более низкой степени. Наличие таких клеток в опухолях 4 степени злокачественности может приводить к снижению доли клеток, которые можно напрямую классифицировать как степень злокачественности 3.



А - Плотность клеток G3 на 1 мм² в опухолевой ткани. Б - Процентное соотношение клеток 3-й степени в опухолевой ткани. В - Корреляция между соотношением клеток 3-й степени и степенью опухоли. Г - ROC-анализ для прогнозирования степени злокачественности 1-2 по сравнению со степенями 3-4.

Рисунок 14 - Статистический анализ клеточного профиля и предсказывание экспертных оценок

Модель продемонстрировала высокую производительность в бинарной классификации между случаями низкой (1-2) и высокой (3-4) степенями злокачественности, достигнув AUC $0,79 \pm 0,04$. Чувствительность модели составила 0,89, а ее специфичность — 0,59.

4.3 Обнаружение опухолевых клеток в образцах тканей

Полупрозрачные маски накладывались на исходные изображения сПКР, позволяющие оценить сегментацию клеток G3 в опухолях. На масках клетки с видимыми ядрышками (G3) были выделены желтым цветом, в то время как

стромальные клетки и клетки 1, 2 и 4-й степени (others) были окрашены в синий цвет. Применение масок показало, что клетки G3 присутствовали в опухолях всех степеней злокачественности по WHO/ISUP, однако их количество варьировалось (Рисунок 15). Хотя, в отличие от системы классификации WHO/ISUP, все сканирования с сПКР 1-й и 2-й степеней злокачественности включали минимальное или относительно небольшое количество клеток с видимыми ядрышками в ядрах.

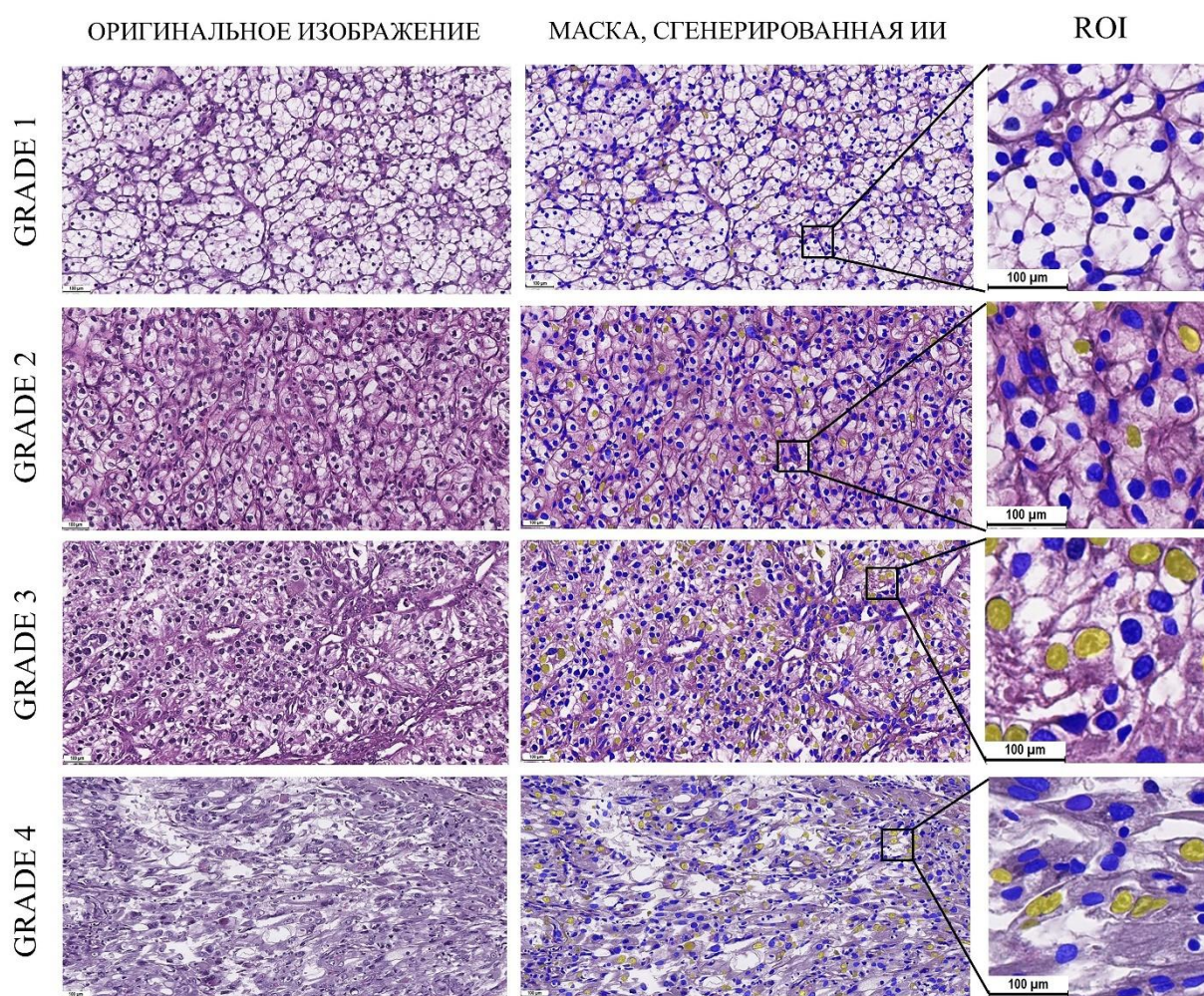


Рисунок 15 - Маска, сгенерированная моделью компьютерного зрения для сканирования сПКР с различными степенями злокачественности

Наиболее интересным морфологическим открытием было то, что в опухолях разных пациентов с 3-й степенью злокачественности наблюдалась значительная неоднородность в содержании клеток с видимыми ядрышками. Относительное количество клеток G3 варьировалось от 0,39% до 32,62%. Если рассматривать эти

клетки как показатель метастатического потенциала опухолевой ткани, то такие различия в клеточном профиле могут свидетельствовать против использования дискретной классификации в пользу определения постоянного прогностического показателя (Рисунок 16).

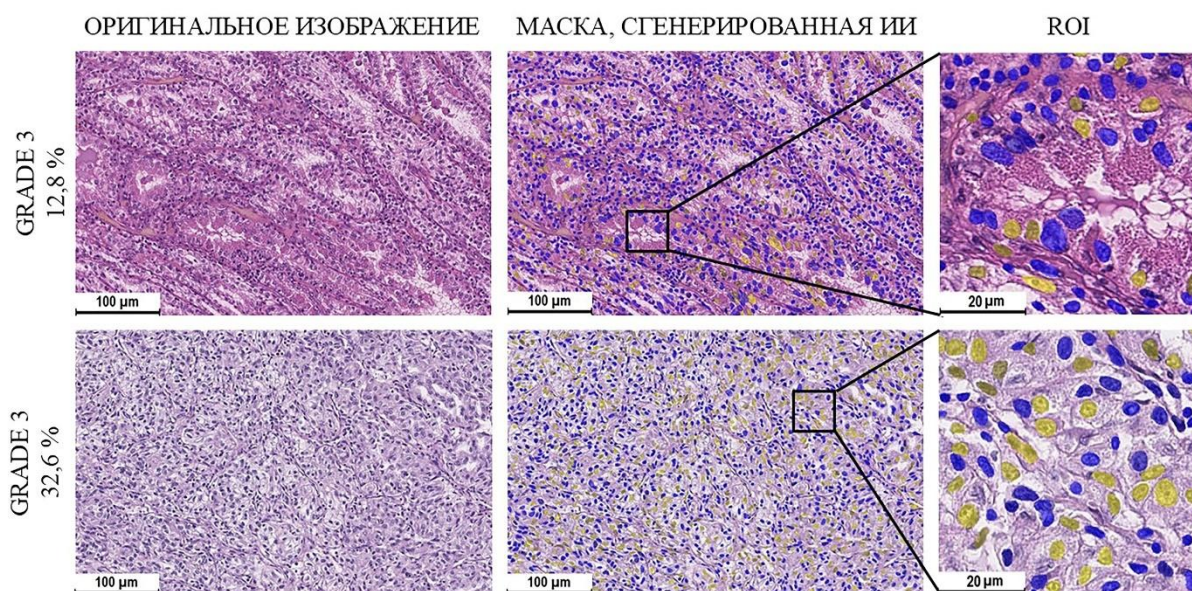


Рисунок 16 - Изменение содержания клеток G3 в скан-изображениях сПКР 3 степени злокачественности

4.4 Анализ выживаемости

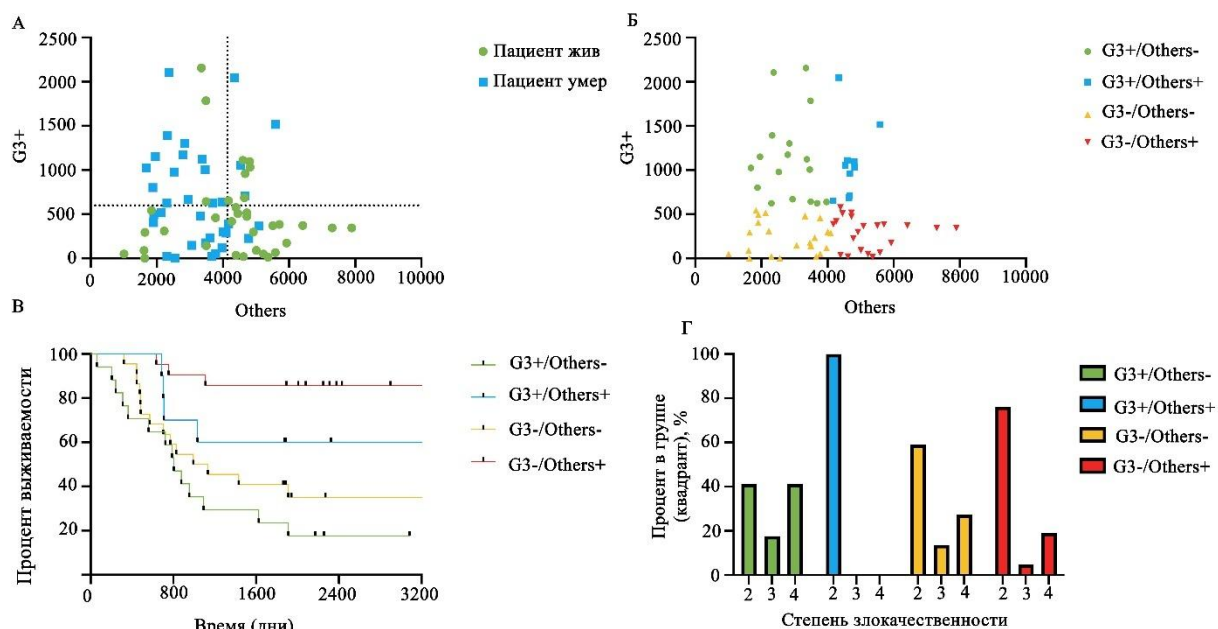
Показатели распределения клеток G3 и others, рассчитанные при помощи модели, были использованы чтобы выяснить, имеют ли они независимую прогностическую ценность. На основе прогнозируемых значений G3 и others на мм² образца ткани из набора данных, содержащего информацию о выживаемости пациентов, были определены два пороговых значения для опухолевых тканей: 533 для G3 и 4133 для G3+ и others клеток на мм². Используя эти пороговые значения, набор данных был разделен на четыре квадранта:

1. Grade 3+/Others-: >533 G3 клеток и \leq 4133 Others клеток/ мм²
2. Grade 3+/Others+: >533 G3 клеток и >4133 Others клеток/ мм²

3. Grade 3-/Others-: ≤ 533 G3 клеток и < 4133 Others клеток/ мм^2

4. Grade 3-/Others+: ≤ 533 G3 клеток и ≥ 4133 Others клеток/ мм^2

Среднее время выживания для всего набора данных составило 1912 дней. Две группы с низким количеством others имели среднюю продолжительность жизни менее 3 лет. У пациентов, относящихся к квадранту Grade 3+/Others-, медиана выживаемости составила 804 дня, в то время как у пациентов, относящихся к квадранту Grade 3-/Others-, медиана выживаемости составила 1062,5 дня. Было предсказано, что две группы с большим количеством others проживут более 6 лет. Однако точное среднее время выживания для них не было определено из-за недостатка данных (Рисунок 17).



А - Распределение типов клеток G3 и others среди пациентов в зависимости от результата (цветовая кодировка). Пунктирные линии представляют собой пороговые значения, используемые для разделения набора данных на подгруппы выживших и умерших пациентов. Б - Распределение квадрантов на основе пороговых значений. В - Анализ выживаемости Каплана-Майера, показывающий прогнозы выживаемости для каждого квадранта. Г - Распределение квадрантов по степени злокачественности

Рисунок 17 - Статистический анализ выживаемости пациентов со сПКР

Кривые выживаемости для четырех квадрантов продемонстрировали статистически значимые различия (p -значение = 0,0005, логарифмический

критерий), что указывает на различные результаты выживаемости, связанные с принадлежностью к квадранту. Кроме того, распределение оценок было в значительной степени связано с принадлежностью к определенному квадранту, о чем свидетельствует статистическая тенденция (p -значение = 0,06349), которая была подтверждена точным критерием Фишера.

4.5 Морфологическое описание структуры опухоли

Выявленная связь между клеточным составом опухоли и выживаемостью пациентов с сПКР стала причиной повторного пересмотра скан-изображений сПКР различной степени злокачественности. Было выявлено 4-ре различных морфологических паттерна, основанных на количестве клеток с видимыми ядрышками и общем количестве клеток в опухолевой ткани (Рисунок 18).

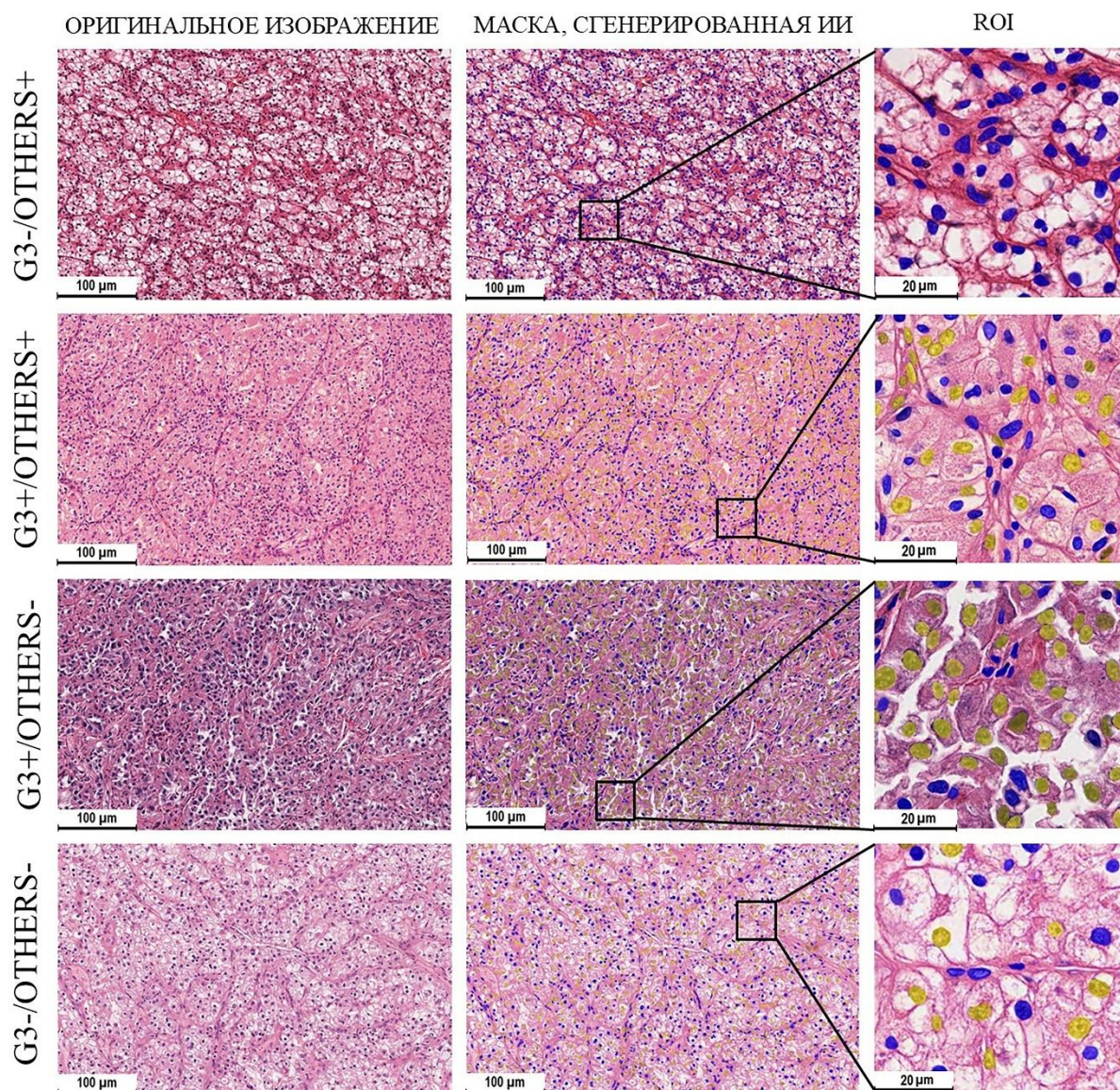


Рисунок 18 - Гистологический анализ скан-изображений сПКР с 4 клеточными паттернами

В первом («мономорфном») определялось небольшое количество клеток G3 ($5,24\% \pm 3,55\%$), в то время как количество клеток others было высоким. Такое соотношение клеток чаще всего наблюдалось в опухолях 2-й степени злокачественности у живых пациентов в течение длительного времени после постановки диагноза. Этот клеточный паттерн был характерен для высокодифференцированного сПКР. При гистологическом анализе эти опухоли демонстрировали гнездные или альвеолярные паттерны роста. Клетки имели оптически пустую цитоплазму и в основном однородные, небольшие, базофильные

ядра с отсутствующими или едва заметными ядрышками. Эта морфология в первую очередь соответствовала 1-й и 2-й степеням по классификации WHO/ISUP, где ядерная атипия минимальна, а видимость ядрышка не является определяющим признаком.

Во втором («плюралистическом») паттерне наблюдалось большое количество как клеток G3, так и клеток others. Этот клеточный состав присутствовал как у живых, так и у умерших пациентов, при этом время выживания превышало шесть лет. Хотя эксперты определяли степень злокачественности этих опухолей как 2, в них присутствовала заметная доля клеток G3 ($18,45\% \pm 5,57\%$). Паттерны роста опухоли были альвеолярными, гнездными или тубулярными с редкими кистами и кровоизлияниями. Морфологически клетки были круглыми с прозрачной цитоплазмой и имели умеренную вариабельность в размере и форме. Ядра имели слегка неровные границы, овальную форму, светло-фиолетовый цвет и ядрышки, видимые при увеличении $\times 400$. Интересно, что, хотя клетки G3 были идентифицированы моделью, они не были доминирующей характеристикой опухолевой ткани. Наличие клеток G3 в ткани опухоли предполагает ее более агрессивное поведение, тем не менее эксперты классифицировали их как степень злокачественности 2, возможно, из-за их низкого процента. Несмотря на общую более низкую распространенность ядерной атипии, строгое следование системе ВОЗ/ISUP требует классифицировать все эти случаи как степень злокачественности 3.

В третьем («ядрышковом») паттерне отмечалось преобладание клеток G3 ($28,29\% \pm 10,04\%$) и над клетками others. Пациенты с этим паттерном имели самое короткое ожидаемое время выживания — 2,2 года. Интересно, что эти случаи чаще имели экспертную оценку степени злокачественности как 2 или 4, чем 3. Это подчеркивает вариабельность между наблюдателями, что является распространенной проблемой при классификации сПКР по степени злокачественности. При морфологическом исследовании образцов определялись тубулярные, гнездные или альвеолярные паттерны роста опухоли, а также наличие очагов некроза, кровоизлияний и единичных кист. Опухоль состояла из большого

количества клеток с прозрачной цитоплазмой и пурпурными крупными ядрами с неровными границами. Эозинофильные ядрышки были видны при небольшом увеличении (x100). Некоторые клетки демонстрировали признаки анаплазии, ядерного полиморфизма с выраженной вариабельностью в размере и форме. Этот паттерн характерен для сПКР 3 степени злокачественности согласно классификации WHO/ISUP, где заметные ядрышки являются ключевым признаком.

Последний («дистрофический») паттерн характеризовался низким количеством клеток обоих типов (менее 5000 в области интереса площадью 1 мм²). Пациенты с этим клеточным составом имели плохой прогноз - 2,7 лет. Паттерны роста опухоли были альвеолярными или тубулярными, с редкими кровоизлияниями, обширными областями некроза и выраженным отеком. Клетки с заметными ядрышками были крупными и редкими, имели светлую эозинофильную цитоплазму и большие ядра с неровными границами. Ядрышки были заметны при увеличении x100 (8,56% ± 7,06%). При изучении препаратов от умерших пациентов отмечалось, что клетки G3 были очень редкими, что указывало на выраженные дистрофические процессы в опухолевой ткани, преобладающие над пролиферацией раковых клеток. В 4 из 21 случаев дистрофического паттерна наблюдалось низкое относительное количество клеток G3 (от 0,39% до 1,35%), что приводило к плотности менее 50 клеток G3 на 1 мкм². Низкая плотность клеток указывала на обширные дистрофические процессы и обширные зоны некроза опухолевой ткани, в частности, на наличие гомогенных скоплений погибших клеток. Этот паттерн коррелировал со степенями злокачественности 3 и 4, где ядерный полиморфизм является выраженным, а некроз обширным. Данные результаты согласуются с предложениями включить некроз в систему градации WHO/ISUP для повышения ее прогностической точности [13-15].

ГЛАВА 5. ЗАКЛЮЧЕНИЕ

В данном исследовании представлены результаты валидации при обработке гистологических скан-изображениях биоптатов почки моделью компьютерного зрения, которая не только идентифицирует структуры тканей и предсказывает прогноз, но и предоставляет количественные альтернативы двум наиболее эффективным показателям — фиброзу и интерстициальной инфильтрации, — необходимым для диагностики в рамках системы классификации Banff. Эти морфологические данные имеют решающее значение для дифференциации пограничных изменений и острого отторжения трансплантата почки. В данной работе продемонстрировано, что модель компьютерного зрения может стать полезным инструментом в рутинной патологоанатомической практике благодаря способности обнаруживать области интерстициальной инфильтрации в биоптате и предсказывать соответствующую клиническую оценку (i) с высокой точностью. Это утверждение основано на прогнозировании экспертных оценок при ROC-анализе, а также на оценке чувствительности и специфичности. При валидации данного алгоритма было показано, что он может точно предсказывать значение i (интерстициальной инфильтрации) в баллах, а не приблизительно, как при традиционной световой микроскопии. Также модель позволила определить четкие границы классов по значению i , которые можно использовать для дальнейшего уточнения классификации Banff. Стоит отметить, что алгоритм может быть полезен практикующим патологоанатомам в рутинной практике, когда есть сомнения в различении $i = 1$ и $i = 2$ по классификации Banff. Эта незначительная разница в оценках определяет терапевтическую стратегию и может стать вопросом жизни и смерти для пациента (Рисунок 19).

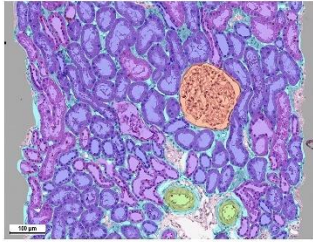
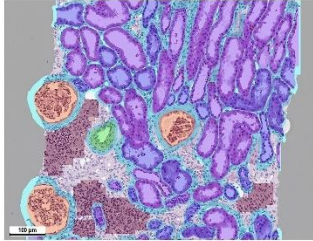
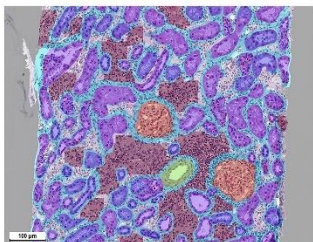
ЗНАЧЕНИЕ ПО КЛАССИФИКАЦИИ BANFF	СЕГМЕНТАЦИЯ	МОРФОЛОГИЯ	КЛИНИЧЕСКИЕ ИСХОДЫ
i=0 (S < 10%) отсутствие интерстициальной инфильтрации		Отсутствие или незначительные скопления лимфоцитов, указывающие на неспецифическую фоновую реакцию	Адекватное лечение против отторжения трансплантата
i=1 (S = 10-25%) слабая интерстициальная инфильтрация		Пограничные изменения: в строме трансплантата имеется несколько очагов скопления лимфоцитов	Мониторинг; незначительная корректировка или усиление текущего режима иммуносупрессии
i=2 (S = 25-50%) умеренная интерстициальная инфильтрация		T-клетки инфильтрируют трансплантат, что приводит к повреждению тканей и изменению морфологии, характеризующемуся воспалением, гибелью клеток и фиброзом	При i=2+ диагностируется острое T-клеточное отторжение; необходима альтернативная стратегия иммуносупрессии (высокие дозы стероидов, антитимоцитарный глобулин).

Рисунок 19 – Связь морфологических и клинических изменений с тактикой лечения при разных оценках по классификации Banff

Применение технологий ИИ в рутинном патологоанатомическом анализе биоптатов пересаженных почек повышает объективность и эффективность диагностики острого отторжения, а также потенциально приведет к более персонализированной терапии и улучшению результатов лечения пациентов. Кроме того, данная работа иллюстрирует, что сочетание НС с традиционными моделями машинного обучения может обеспечить комплексный анализ как тканевых, так и клеточных компонентов биологических тканей, а также устраняет разрыв между устоявшейся клинической системой и анализом на основе данных, демонстрируя необходимость использования ИИ-моделей для преобразования текущей клинической практики.

Применение технологий компьютерного зрения может масштабироваться и на онкологические процессы в различных органах, в том числе и почки. Например,

все больше внимания уделяется использованию ИИ для повышения точности диагностики, охватывающей от подробного клеточного анализа для постановки диагноза и определения степени злокачественности опухоли до углубленных оценок микросреды опухоли и идентификации прогностических гистологических биомаркеров — все в рамках существующих клинических протоколов. Медицинские системы принятия решений могут стать незаменимыми инструментами, помогающими онкопатологам в их повседневной работе. Многие исследовательские группы и IT-компании по всему миру сосредоточены на автоматизации трудоемких задач, используя алгоритмы для выполнения крупномасштабного анализа, например, классификации каждой отдельной клетки в образцах тканей, что не актуально в ежедневной практике патологоанатома.

В данной работе была измерена плотность клеток с видимыми ядрышками в опухолевой ткани сПКР различных степеней злокачественности, а также оценивалось ее влияние на исход заболевания. Была отмечена значительная вариабельность в клеточном составе опухолей, классифицированных как Grade 3 по классификации WHO/ISUP. Эта изменчивость подчеркивает необходимость более тонкой системы классификации, которая могла бы привести к более персонализированным стратегиям лечения пациентов с сПКР. Кроме того, данное исследование проливает свет на ключевую роль клеточного состава опухоли в прогрессировании заболевания и продемонстрировать его роль как прогностического биомаркера. С помощью детального морфологического и статистического анализа было выявлено четыре различных паттерна, основанных на соотношении клеток типа G3 и других клеток.

В данном исследовании был представлен основанный на анализе данных подход к классификации рака почки, который решает ограничения текущих систем путем оценивания клеточного состава в опухоли. Выявив четыре паттерна опухолевой ткани, было продемонстрировано, что как относительное количество клеток с выраженными ядрышками, так и наличие дистрофических изменений, таких как очаговый некроз и отек, имеют независимую прогностическую ценность [90]. Эти результаты указывают на то, что включение дополнительных

морфометрических параметров в будущем пересмотры классификации WHO/ISUP может значительно повысить ее прогностическую точность, позволяя делать более персонализированные прогнозы результатов лечения пациентов. Этот подход предлагает многообещающий путь для повышения биологической значимости классификации рака почки, что потенциально приведет к более индивидуальным и эффективным стратегиям лечения.

Основные тенденции в вычислительной патологии, которые предоставят новые знания о гистологии, включают мультимодальный анализ и трансферное обучение. Клинические гистологические и протеономные методы можно объединить для повышения точности прогнозирования рецидива рака почки. Также ИИ может играть важную роль в диагностике редких подтипов заболеваний, основанных на молекулярно-генетических особенностях [23]. Прогрессирование сПКР достигается не только за счет неограниченного роста опухолевых клеток, но также поддерживается, стимулируется и питается за счет опухолевого микроокружения [18,19]. Наконец, комбинированное использование моделей ИИ со специализированными функциями (например, классификаторами клеток) и большими языковыми моделями имеет большой потенциал для выявления скрытых корреляций и обеспечения более глубокого понимания биологии опухолей.

Ожидается, что в обозримом будущем внедрение ИИ в патологию опередит его применение во многих других областях. Во многом это объясняется тем, что эксперты могут быстро оценить полезность продукта, интегрированного с ИИ. В ближайшие годы решения на основе ИИ позволят улучшить рутинную работу патологоанатомов и впоследствии внедрить инновации в области мультимодального анализа, включая морфологию, компьютерную томографию и генетику, а также биофотонику. Однако существует необходимость в тщательных исследованиях программ на основе ИИ, чтобы подкрепить полученные результаты эмпирическими данными. В настоящее время несколько центров располагают обширными базами данных и проводят рандомизированные исследования с использованием различных моделей сканеров, методов окрашивания и протоколов иммуногистохимических исследований. Эти ресурсы могут способствовать

проведению клинических испытаний, в которых традиционные методы диагностики дополняются инновационными, прокладывая путь к научно обоснованному и дополненному ИИ будущему в патологии. Дальнейшая эволюция извлечения признаков из скан-изображений препаратов и разработка эффективных вычислительных методов для гистологического анализа являются будущим для диагностики, основанной на данных [13,20,90-92].

ВЫВОДЫ

- 1) Были разработаны оригинальные модели цифрового анализа изображений биоптатов трансплантированных почек и светлоклеточного почечно-клеточного рака и установить точность их работы. Точность работы алгоритма цифрового анализа изображений для биоптатов почечных трансплантатов была оценена при расчете Dice-коэффициента и составляла для почечных канальцев - 0,8310, почечных клубочков - 0,9242, артерий - 0,7087, стромы - 0,7919. 3) При оценке аннотированных клеток в ткани светлоклеточного почечно-клеточного рака были достигнуты высокие показатели: значение F1 равнялось 0,87. Была установлена высокая точность: количество ложнонегативных клеток составило всего 9,16 %, что свидетельствует о высокой точности определения клеток с видимыми ядрышками.
- 2) Полученные результаты позволили точно количественно оценивать патологические изменения в строме почки. Точность предсказания экспертных оценок фиброза и интерстициальной инфильтрации по системе диагностики острого отторжения трансплантата почки Banff достигала 97% при использовании методов компьютерной морфометрии и цифрового анализа. Пороговыми значениями площади фиброза от 0 до 1 степени острого клеточного отторжения трансплантата почки составило 16,54 %, от 1 до 2 - 24,19 %, от 2 до 3 - 36,86 % и более 3 - 60,67 %. Также были установлены пороговые значения площади интерстициальной инфильтрации при различных степенях острого клеточного отторжения биоптата трансплантированной почки: от 0 до 1 степени - 2,09 %, от 1 до 2 - 4,44 %, от 2 до 3 - 17,58 %, более 3 - 19,57 %.
- 3) В ходе исследования была оценена работа модели цифрового анализа изображений для сегментации и классификации опухолевых клеток в ткани светлоклеточного почечно-клеточного рака, которая продемонстрировала высокую точность в определении степени злокачественности опухоли, предсказывая ядрышковую градацию (высоко- и низкодифференцированные по системе WHO/ISUP) с AUC ROC 0,79. Были установлены количественные метрики для клеток с видимыми ядрышками на 1 мм² в ткани светлоклеточного почечно-

клеточного рака от 1 до 4 степеней злокачественности: $254,76 \pm 246,96$ ($7,08 \pm 0,07$ %); $418,83 \pm 247,22$ ($10,43 \pm 0,081$ %); $612,9 \pm 323,90$ ($17,93 \pm 0,082$ %) и $628,46 \pm 167,71$ ($15,69 \pm 0,057$ %).

4) Были установлены 4 подгруппы пациентов (гистологических паттернов) на основе относительного соотношения клеток с видимыми ядрышками к остальным клеткам. Пороговые значения составили 533 для клеток с видимыми ядрышками и 4133 для остальных клеток на 1 мм. Была установлена средняя выживаемость пациентов в каждой подгруппе, длительная продолжительность жизни пациентов отмечалась в двух прогностически благоприятных подгруппах и составила более 6 лет. Также было определено два наиболее неблагоприятных паттерна (более 533 клеток с видимыми ядрышками и менее 4133 остальных клеток на 1 мм² и более 533 клеток с видимыми ядрышками и более 4133 остальных клеток на 1 мм²), в которых выживаемость равнялась 2,2 и 2,7 лет соответственно. Было показано, что важно не только присутствие клеток с видимыми ядрышками в ткани опухоли, а также их относительное количество: в опухолях с умеренной клеточной плотностью процент клеток с видимыми ядрышками должен превышать 11%, но и низкая клеточная плотность является прогностически неблагоприятным фактором, что свидетельствует о некротических и дистрофических процессах в опухоли и подтверждает необходимость учета морфометрических характеристик при стратификации риска и выборе тактики лечения.

5) Для выделенных по результатам анализа выживаемости прогностических групп были предложены следующие названия: мономорфный, плюралистический, ядрышковый, дистрофический. В мономорфном паттерне наблюдалось малое количество клеток с видимыми ядрышками ($5,24\% \pm 3,55\%$), что свидетельствует о низкой пролиферативной активности. В плюралистическом паттерне соотношение клеток обоих типов было равным ($18,45\% \pm 5,57\%$). В ядрышковом паттерне отмечалось наибольшее количество клеток с видимыми ядрышками ($28,29\% \pm 10,04\%$), что свидетельствует о высоком риске метастазирования опухоли. В дистрофическом паттерне наблюдались обширные поля некрозов, малое количество клеток с видимыми ядрышками ($8,56\% \pm 7,06\%$) и остальных, а также

их выраженные дистрофические изменения. Полученные данные могут быть применены для совершенствования клинико-морфологических классификаций злокачественных опухолей почки, в частности указывают на важность учитывания некротических изменений, который на настоящий момент не включены в классификацию WHO/ISUP. Пороговые значения для клеток с выраженными ядрышками могут применяться для градации рака почки с помощью программ на основе ИИ для подсчета клеток или же с помощью ручного подсчета при анализе нескольких полей зрения на увеличении микроскопа x400.

ПРАКТИЧЕСКИЕ РЕКОМЕНДАЦИИ

- 1) Рекомендуется внедрять модели на основе компьютерного зрения для оценки биоптатов от пациентов после трансплантации почки, что позволит своевременно выявлять признаки острого отторжения в ткани трансплантата и корректировать терапию на ранних стадиях.
- 2) Рекомендуется внедрение моделей компьютерного зрения для автоматизированного анализа гистологических изображений сПКР. Эти модели могут повысить точность степени злокачественности опухоли и снизить субъективность при оценке большого массива опухолевых клеток, что значительно улучшит результаты диагностики и прогнозирования исхода заболевания у пациентов.

СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ

ПКР – почечно-клеточный рак

сПКР – светлоклеточный почечно-клеточный рак

пПКР – папиллярный почечноклеточный рак

хПКР – хромофобный почечноклеточный рак

TNM (Tumor, Nodus, Metastasis) – опухоль, узел, метастаз

WHO/ISUP (World Health Organization/International Society of Urological Pathology)

- Всемирная Организация Здравоохранения/Международное Общество Урологической Патологии

ИИ – искусственный интеллект

NGS (Next-Generation Sequencing) – секвенирование нового поколения

ИГХ – иммуногистохимия

WSI (Whole Slide Image) – целноslideовое изображение

SVM (Support Vector Machine) – метод опорных векторов

НС – нейронная сеть

TCGA (The Cancer Genome Atlas) - Атлас генома рака

CPTAC (Clinical Proteomic Tumor Analysis Consortium) – Консорциум клинического протеомного анализа опухолей

ROI (Region of interest) – область интереса

CNN (Convolutional Neural Network) – сверточная нейронная сеть

HSV (hue, saturation and value) - тон, насыщенность, значение

SSAE (stacked sparse autoencoder) - многоуровневый разреженный автоэнкодер

RF (random forest) – метод случайного леса

RGB (red, green, blue) – красный, зеленый, синий

LASSO (Least Absolute Shrinkage and Selection Operator) - метод оценивания коэффициентов линейной регрессионной модели

RFE (Recursive Feature Elimination) - рекурсивное устранение признаков

спПКР – светлоклеточный папиллярный почечно-клеточный рак

КТ – компьютерная томография

МРТ – магнитно-резонансная томография

MMDLM (multimodal deep learning model) – мультимодальная модель глубокого обучения

MLPS (multilayer perceptron system) - многослойная система перцептрона

ROC (Receiver operating characteristic) - кривая рабочих характеристик приемника

AUC ROC (Area Under the ROC Curve) – площадь под кривой рабочих характеристик приемника

TIL (tumor-infiltrating lymphocytes) - инфильтрирующие опухоль лимфоциты

TFE3- RCC - транслокационная почечно-клеточная карцинома TFE3 Xp11.2

DEG (differentially expressed genes) - дифференциально экспрессируемые гены

CNV (Copy number variation) – вариация числа копий

CPG - цитозинфосфатгуанин

COCO (Common Objects in Context) – общие объекты в контексте

SGD (Stochastic gradient descent) - стохастический градиентный спуск

mAP (mean average precision) - среднее значение точности

ДИ – доверительный интервал

СПИСОК ЛИТЕРАТУРЫ

1. Алексеева, Г. Н. Рак почки в регионе Сибири и Дальнего Востока. Эпидемиологические особенности, факторы риска, ранняя диагностика и эффективность лечения : специальность 14.01.12 «Онкология» : автореферат диссертации на соискание степени доктора медицинских наук / Алексеева Галина Николаевна ; ФГБНУ Томский национальный исследовательский медицинский центр Российской академии наук. – Томск, 2019. – 47 с.
2. American Cancer Society: official website. Key Statistics About Kidney Cancer. [Electronic resource] – URL: <https://www.cancer.org/cancer/types/kidney-cancer/about/key-statistics.html> (дата обращения: 12.01.2024).
3. International variations and trends in renal cell carcinoma incidence and mortality / A. Znaor, J. Lortet-Tieulent, M. Laversanne, [et al.] // European Urology. – 2015. – Vol. 67. – № 3. – P. 519–530.
4. WHO Classification of Tumours Editorial Board. Urinary and Male Genital Tumours. – Lyon : International Agency for Research on Cancer, 2022. — 356 p. – ISBN 978-92-832-4512-4. – Текст : непосредственный.
5. Множественные гепатоцеллюлярные аденомы и почечноклеточный рак, ассоциированные с приёмом бодибилдером анаболических андрогенных стероидов / М. М. Морозова, Е. И. Иванова, Н. К. Чардаров, О. В. Должанский, Г. А. Шатверян, Ю. Р. Камалов // Хирургия. Журнал им. Н. И. Пирогова. — 2021. — № 10. — С. 105–109.
6. Иванова, Е. И. Сочетание светлоклеточного почечно-клеточного рака и высокодифференцированной нейроэндокринной опухоли почки / Е. И. Иванова, С. Е. Соловьева, О. В. Должанский // Клиническая и экспериментальная морфология. — 2022. — Т. 11. — № 3. — С. 67–72.
7. American Cancer Society: official website. Kidney Cancer Early Detection, Diagnosis, and Staging, American Cancer Society. [Electronic resource] – URL:

<https://www.cancer.org/cancer/types/kidney-cancer/detection-diagnosis-staging.html>

(дата обращения: 12.01.2024).

8. International Union against Cancer (UICC) : TNM Classification of Malignant Tumours. 8th ed. / J.D. Brierley, M.K. Gospodarowicz, C. Wittekind ; под редакцией J.D. Brierley, M.K. Gospodarowicz, C. Wittekind – Oxford : Wiley, 2017. – 253 p. – ISBN 9781119263579 – Текст : непосредственный.

9. American Cancer Society: official website. Treatment of Kidney Cancer by Stage, American Cancer Society. [Electronic resource] – URL: <https://www.cancer.org/cancer/types/kidney-cancer/treating/by-stage.html> (дата обращения: 12.01.2024).

10. Fuhrman, S.A. Prognostic significance of morphologic parameters in renal cell carcinoma / S. A. Fuhrman, L. C. Lasky, C. Limas // American Journal of Surgical Pathology. – 1982. – Vol. 6. – № 7. – P. 655–663.

11. Validation of the World Health Organization/International Society of Urological Pathology grading for Chinese patients with clear cell renal cell carcinoma / Q. Xiao, X. Yi, X. Guan [et al.] // Translational Andrology and Urology. – 2020. – Vol. 9. – № 6. – P. 2665–2674.

12. Warren, A.Y. WHO/ISUP classification, grading and pathological staging of renal cell carcinoma: standards and controversies / A. Y. Warren, D. Harrison // World Journal of Urology. – 2018. – Vol. 36. – № 12. – P. 1913–1926.

13. Artificial intelligence assists in the detection of blood vessels in whole slide images: practical benefits for oncological pathology / A. Timakova, V. Ananov, A. Fayzullin, V. Makarov, E. Ivanova, A. Shekhter, P. Timashev // Biomolecules. — 2023. — Vol. 13. — e1327. — [Electronic resource] – URL: <https://www.mdpi.com/2218-273X/13/9/1327>.

14. Периимплантационное таргетирование с помощью модифицированных аденовирусных векторов / Н. Б. Сережникова, А. Л. Файзуллин, А. А. Антошин, А. А. Тимакова, Е. И. Иванова, Б. П. Ершов, Н. М. Файзуллина, И. В. Уласов, А. С. Малоголовкин, С. Л. Котова, П. С. Тимашев // Журнал анатомии и гистопатологии. — 2024. — Т. 13. — № 4. — С. 61–74.

15. A novel machine learning approach reveals latent vascular phenotypes predictive of renal cancer outcome / N. Ing, F. Huang, A. Conley [et al.] // *Scientific Reports*. – 2017. – Vol. 7. – № 1. – P. 13190. – [Electronic resource]. – URL: <https://www.nature.com/articles/s41598-017-13196-4>.
16. Development and Validation of a Histological Method to Measure Microvessel Density in Whole-Slide Images of Cancer Tissue / K. M. Marien, V. Croons, Y. Waumans [et al.] // *PLoS One*. – 2016. – Vol. 11. – № 9. – [Electronic resource]. – P. e0161496. – URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0161496>.
17. The potential of artificial intelligence-based applications in kidney pathology / R. D. Büllow, J. N. Marsh, S. J. Swamidass [et al.] // *Current opinion in nephrology and hypertension*. – 2022. – Vol. 31. – № 3. – P. 251–257.
18. Artificial intelligence and machine learning in nephropathology / J. U. Becker, D. Mayerich, M. Padmanabhan [et al.] // *Kidney International*. – 2020. – Vol. 98. – № 1. – P. 65–75.
19. Prognostic factors and predictive models in renal cell carcinoma: a contemporary review / M. Sun, S. F. Shariat, C. Cheng [et al.] // *European Urology*. – 2011. – Vol. 60. – № 4. – P. 644–661.
20. Empowering renal cancer management with AI and digital pathology: pathology, diagnostics and prognosis / E. Ivanova, A. Fayzullin, V. Grinin, D. Ermilov, A. Arutyunyan, P. Timashev, A. Shekhter // *Biomedicines*. — 2023. — Vol. 11. — e2875. — [Electronic resource] – URL: <https://www.mdpi.com/2227-9059/11/11/2875>.
21. Computational analysis of pathological images enables a better diagnosis of TFE3 Xp11.2 translocation renal cell carcinoma / J. Cheng, Z. Han, R. Mehra [et al.] // *Nature Communications*. – 2020. – Vol. 11. – № 1. – P. 1778. – [Electronic resource]. – URL: <https://www.nature.com/articles/s41467-020-15671-5>.
22. Automated clear cell renal carcinoma grade classification with prognostic significance / K. Tian, C. A. Rubadue, D. I. Lin [et al.] // *PLoS One*. – 2019. – Vol. 14. – № 10. – P. e0222641. – [Electronic resource]. – URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0222641>.

23. Automated grading of renal cell carcinoma using whole slide imaging / F. C. Yeh, A. V. Parwani, L. Pantanowitz [et al.] // *Journal of Pathology Informatics*. – 2014. – Vol. 5. – № 1. – P. 23. – [Electronic resource]. – URL: linkinghub.elsevier.com/retrieve/pii/S2153353922002929.
24. Automated renal cancer grading using nuclear pleomorphic patterns / D. A. Holdbrook, M. Singh, Y. Choudhury [et al.] // *JCO Clinical Cancer Informatics*. – 2018. – Vol. 2. – P. 1–12.
25. A review of machine learning methods of feature selection and classification for autism spectrum disorder / M. Rahman, O. L. Usman, R. C. Muniyandi [et al.] // *Brain Sciences*. – 2020. – Vol. 10. – № 12. – P. 949. – [Electronic resource]. – URL: <https://www.mdpi.com/2076-3425/10/12/949>.
26. Radiogenomics in renal cancer management—current evidence and future prospects / M. Ferro, G. Musi, M. Marchioni [et al.] // *International Journal of Molecular Sciences*. – 2023. – Vol. 24. – № 9. – P. 4615. – [Electronic resource]. – URL: <https://scispace.com/papers/radiogenomics-in-renal-cancer-management-current-evidence-14iucujl>.
27. Artificial intelligence and radiomics in evaluation of kidney lesions: a comprehensive literature review / M. Ferro, F. Crocetto, B. Barone [et al.] // *Therapeutic Advances in Urology*. – 2023. – Vol. 15. – P. 17562872231164803. – [Electronic resource]. – URL: <https://journals.sagepub.com/doi/10.1177/17562872231164803>.
28. The role of artificial intelligence in the diagnosis and prognosis of renal cell tumors / B. M. Naghavi-Behzad, C. B. Petersen, M. Vogsen [et al.] // *Diagnostics*. – 2021. – Vol. 11. – № 11. – P. 206. – [Electronic resource]. – URL: <https://www.mdpi.com/2075-4418/11/2/206>
29. Ensemble of classifiers and wavelet transformation for improved recognition of Fuhrman grading in clear-cell renal carcinoma / M. Kruk, J. Kurek, S. Osowski [et al.] // *Biocybernetics and Biomedical Engineering*. – 2017. – Vol. 37. – № 3. – P. 357–364.
30. Identification of topological features in renal tumor microenvironment associated with patient survival / J. Cheng, X. Mo, X. Wang [et al.] // *Bioinformatics*. – 2018. – Vol. 34. – № 6. – P. 1024–1030.

31. Semantic interpretation of robust imaging features for Fuhrman grading of renal carcinoma/A. Champion, G. Lu, M. Walker [et al.]. — Текст: непосредственный // Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society / IEEE. — Chicago: IEEE, 2014. — P. 6446–6449.
32. Connecting Histopathology Imaging and Proteomics in Kidney Cancer through Machine Learning / F. Azuaje, S.-Y. Kim, D. P. Hernandez [et al.] // Journal of Clinical Medicine. — 2019. — Vol. 8. — №10. — P. 1535. — [Electronic resource]. — URL: <https://www.mdpi.com/2077-0383/8/9/1535>.
33. Development and Validation of a Deep-learning Model to Assist with Renal Cell Carcinoma Histopathologic Interpretation / M. Fenstermaker, S. A. Tomlins, K. Singh [et al.] // Urology. — 2020. — Vol. 144. — P. 152–157.
34. Unsupervised Resolution of Histomorphologic Heterogeneity in Renal Cell Carcinoma Using a Brain Tumor-Educated Neural Network / K. Faust, A. Roohi, A. J. Leon [et al.] // JCO Clinical Cancer Informatics. — 2020. — Vol. 4. — P. 811–821.
35. Cellular and Molecular Players in the Tumor Microenvironment of Renal Cell Carcinoma / F. Lasorsa, M. Rutigliano, M. Milella [et al.] // Journal of Clinical Medicine. — 2023. — Vol. 12. — № 12. — P. 3888. — [Electronic resource]. — URL: <https://www.mdpi.com/2077-0383/12/12/3888>.
36. Characterizing the tumor microenvironment in rare renal cancer histological types / N. C. Synnott, M. L. Poeta, M. Costantini [et al.] // Journal of Pathology and Clinical Research. — 2022. — Vol. 8. — № 1. — P. 88–98.
37. Tabibu, S. Pan-Renal Cell Carcinoma classification and survival prediction from histopathology images using deep learning / S. Tabibu, P. K. Vinod, C. V. Jawahar // Scientific Reports. — 2019. — Vol. 9. — № 1.— P. 10509. — [Electronic resource]. — URL: <https://www.nature.com/articles/s41598-019-46718-3>.
38. Urinary MicroRNAs as Biomarkers of Urological Cancers: A Systematic Review / A. Aveta, S. Cilio, R. Contieri [et al.] // International Journal of Molecular Sciences. — 2023. — Vol. 24. — № 13. — P. 10846. — [Electronic resource].— URL: <https://www.mdpi.com/1422-0067/24/13/10846>.

39. Radiomic Features and Machine Learning for the Discrimination of Renal Tumor Histological Subtypes: A Pragmatic Study Using Clinical-Routine Computed Tomography / J. Uhlig, A. Leha, L. M. Delonge [et al.] // *Cancers*. – 2020. – Vol. 12. – № 10. – P. 3010. – [Electronic resource]. – URL: <https://www.mdpi.com/2072-6694/12/10/3010>.
40. A comprehensive computer-assisted diagnosis system for early assessment of renal cancer tumors / M. Shehata, A. Alksas, R. T. Abouelkheir [et al.] // *Sensors*. – 2021. – Vol. 21. – [Electronic resource]. – P. 4928. – URL: <https://www.mdpi.com/1424-8220/21/14/4928>.
41. Multimodal Deep Learning for Prognosis Prediction in Renal Cancer / S. Schulz, A.-C. Woerl, F. Jungmoann [et al.] // *Frontiers in Oncology*. – 2021. – Vol. 11. – P. 788740. – [Electronic resource]. – URL: <https://www.frontiersin.org/journals/oncology/articles/10.3389/fonc.2021.788740/full>.
42. Machine learning-based CT radiomics approach for predicting WHO/ISUP nuclear grade of clear cell renal cell carcinoma: An exploratory and comparative study / Y. Xu, F. Lv, H. Guo [et al.] // *Insights into Imaging*. – 2021. – Vol. 12. – №1. – P. 170. – [Electronic resource]. – URL: <https://insightsimaging.springeropen.com/articles/10.1186/s13244-021-01107-1>.
43. Automated Renal Cell Carcinoma Subtype Classification Using Morphological, Textural and Wavelets Based Features / Q. Chaudry, S.H. Raza, A.N. Young [et al.] // *Journal of Signal Processing Systems*. – 2009. – Vol. 55. – P. 15–23.
44. Automated classification of renal cell carcinoma subtypes using bag-of-features / S.H. Raza, R.M. Parry, Y. Sharma [et al.]. — Текст: непосредственный // *Proceedings of the 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, IEEE*. –Buenos Aires: IEEE, 2010. – P. 6749–6752.
45. Development and evaluation of a deep neural network for histologic classification of renal cell carcinoma on biopsy and surgical resection slides / M. Zhu, B. Ren, R. Richards [et al.] // *Scientific Reports*. – 2021. – Vol. 11. – № 1. – P. 7080. – [Electronic resource]. – URL: <https://www.nature.com/articles/s41598-021-86540-4>.

46. Artificial intelligence-based multi-class histopathologic classification of kidney neoplasms / D. D. Gondim, K. I. Al-Obaidy, M. T. Idrees [et al.] // *Journal of Pathology Informatics*. – 2023. – Vol. 14. – P. 100299. – [Electronic resource]. –URL: <https://www.sciencedirect.com/science/article/pii/S215335392300113X?via%3Dihub>.
47. A pyramidal deep learning pipeline for kidney whole-slide histology images classification / H. Abdeltawab, F. Khalifa, M. Ghazal [et al.] // *Scientific Reports*. – 2021. – Vol. 11. – № 1. – P. 21867. – [Electronic resource]. – URL: <https://www.nature.com/articles/s41598-021-99735-6>.
48. A Novel Tumor Grading Scheme for Chromophobe Renal Cell Carcinoma: Prognostic Utility and Comparison With Fuhrman Nuclear Grade / G. P. Paner, M. B. Amin, I. Alvarado-Cabrero [et al.] // *The American Journal of Surgical Pathology*. – 2010. – Vol. 34. – № 9. – P. 1233–1240.
49. Chromophobe Renal Cell Carcinoma: The Impact of Tumor Grade on Outcome / J. C. Cheville, C. M. Lohse, W. R. Sukov [et al.] // *The American Journal of Surgical Pathology*. – 2012. – Vol. 36. – № 6. – P. 851–856.
50. A Weakly Supervised Deep Learning Model and Human–Machine Fusion for Accurate Grading of Renal Cell Carcinoma from Histopathology Slides / Q. Zheng, R. Yang, H. Xu [et al.] // *Cancers*. – 2023. – Vol. 15. – № 12. – P. 3198. – [Electronic resource]. – URL: <https://www.mdpi.com/2072-6694/15/12/3198>.
51. A novel dataset and efficient deep learning framework for automated grading of renal cell carcinoma from kidney histopathology images / A. K. Chanchal, S. Lal, R. Kumar [et al.] // *Scientific Reports*. – 2023. – Vol. 13. – № 1. – P. 5728. – [Electronic resource]. –URL: <https://www.nature.com/articles/s41598-023-31275-7>.
52. Machine learning-based pathomics signature could act as a novel prognostic marker for patients with clear cell renal cell carcinoma / S. Chen, L. Jiang, F. Gao [et al.] // *British Journal of Cancer*. – 2022. – Vol. 126. – № 5. – P. 771–777.
53. Prognostic and predictive value of tumor-infiltrating leukocytes and of immune checkpoint molecules PD1 and PDL1 in clear cell renal cell carcinoma / P. J. Stenzel, M. Schindeldecker, K. E. Tagscherer [et al.] // *Translational Oncology*. – 2020. – Vol. 13. – № 2. – P. 336–345.

54. Use of high-plex data reveals novel insights into the tumour microenvironment of clear cell renal cell carcinoma / R. De Filippis, G. Wölflein, I. H. Um, P. D. Caie, S. Warren, A. White [et al.] // *Cancers*. – 2022. – Vol. 14. – № 21. – P. 5387. – [Electronic resource]. – URL: <https://www.mdpi.com/2072-6694/14/21/5387>.
55. Multidimensional analyses of tumor immune microenvironment reveal the possible rationality of immunotherapy and identify high immunotherapy response subtypes for renal papillary cell carcinoma / B. Wei, M. Yu, J. Yao, M. Jiang, J. An, J. Yang [et al.] // *Frontiers in Immunology*. – 2021. – Vol. 12. – P. 657951. – [Electronic resource]. – URL: <https://www.frontiersin.org/journals/immunology/articles/10.3389/fimmu.2021.657951/full>.
56. Identification of prognostic markers of DNA damage and oxidative stress in diagnosing papillary renal cell carcinoma based on high-throughput bioinformatics screening / L. Li, X. Liu, Y. Wen, P. Liu, T. Sun // *Journal of Oncology*. – 2023. — P. 4640563. – [Electronic resource]. – URL: <https://onlinelibrary.wiley.com/doi/10.1155/2023/4640563>.
57. Classic chromophobe renal cell carcinoma incur a larger number of chromosomal losses than seen in the eosinophilic subtype / R. Ohashi, P. Schraml, S. Angori [et al.] // *Cancers*. – 2019. – Vol. 11. – № 10. – P. 1492. – [Electronic resource]. – URL: <https://www.mdpi.com/2072-6694/11/10/1492>.
58. Comparative analysis for the distinction of chromophobe renal cell carcinoma from renal oncocytoma in computed tomography imaging using machine learning radiomics analysis / A. J. Alhussaini, J. D. Steele, G. Nabi // *Cancers*. – 2022. – Vol. 14. – № 15. – P. 3609. – [Electronic resource]. – URL: <https://www.mdpi.com/2072-6694/14/15/3609>.
59. Automated differentiation of benign renal oncocytoma and chromophobe renal cell carcinoma on computed tomography using deep learning / A. Baghdadi, N. A. Aldhaam, A. S. Elsayed, A. A. Hussein, L. A. Cavuoto, E. Kauffman, K. A. Guru // *BJU International*. – 2020. – Vol. 125. – № 4. – P. 553–560.
60. Value of radiomics in differential diagnosis of chromophobe renal cell carcinoma and renal oncocytoma / Y. Li, X. Huang, Y. Xia, L. Long // *Abdominal Radiology*. – 2020. – Vol. 45. – № 10. – P. 3193–3201.

61. Development of a DNA Methylation-Based Diagnostic Signature to Distinguish Benign Oncocytoma From Renal Cell Carcinoma / K. Brennan, T. J. Metzner, C. S. Kao [et al.] // *JCO Precision Oncology*. – 2020. – Vol. 4. – P. PO.20.00015. – [Electronic resource]. – URL: <https://ascopubs.org/doi/10.1200/PO.20.00015>.
62. Comparison of mortality in all patients on dialysis, patients on dialysis awaiting transplantation, and recipients of a first cadaveric transplant / R. A. Wolfe, V. B. Ashby, E. L. Milford [et al.] // *The New England Journal of Medicine*. – 1999. – Vol. 341. – № 23. – P. 1725–1730.
63. Die Geschichte der Nierentransplantation / M. Hatzinger, M. Stastny, P. Grützmacher [et al.] // *Urologe*. – 2016. – Vol. 55. – P. 1353–1359.
64. US Renal Data System 2018 Annual Data Report: Epidemiology of Kidney Disease in the United States / R. Saran, B. Robinson, K. C. Abbott [et al.] // *American Journal of Kidney Diseases*. – 2019. – Vol. 73, Suppl. 1. – P. A7–A8.
65. Jeong, H.J. Diagnosis of renal transplant rejection: Banff classification and beyond / H. J. Jeong // *Kidney Research and Clinical Practice*. – 2020. – Vol. 39. – № 1. – P. 17–31.
66. OPTN/SRTR 2019 Annual Data Report: Kidney / A. Hart, K. L. Lentine, J. M. Smith [et al.] // *American Journal of Transplantation*. – 2021. – Vol. 21, Suppl. 2. – P. 21–137.
67. A 2018 Reference Guide to the Banff Classification of Renal Allograft Pathology / C. Roufosse, N. Simmonds, M. Clahsen-van Groningen [et al.] // *Transplantation*. – 2018. – Vol. 102. – № 11. – P. 1795–1814.
68. Loupy, A. Thirty years of the International Banff Classification for Allograft Pathology: the past, present, and future of kidney transplant diagnostics / A. Loupy, M. Mengel, M. Haas // *Kidney International*. – 2022. – Vol. 101. – № 4. – P. 678–691.
69. Renal Medullary and Cortical Correlates in Fibrosis, Epithelial Mass, Microvasculature, and Microanatomy Using Whole Slide Image Analysis Morphometry / A. B. Farris, C. L. Ellis, T. E. Rogers [et al.] // *PLoS One*. – 2016. – Vol. 11. – № 8. – P. e0161019. – [Electronic resource]. – URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0161019>.

70. A neural network for glomerulus classification based on histological images of kidney biopsy / G. D. Cascarano, F. S. Debitonto, R. Lemma [et al.] // *BMC Medical Informatics and Decision Making*. – 2021. – Vol. 21. – Suppl. 1. – P. 300. – [Electronic resource]. – URL: <https://bmcmmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-021-01650-3>.
71. Computerized histomorphometric assessment of protocol renal transplant biopsy specimens for surrogate markers of chronic rejection / M. L. Nicholson, E. Bailey, S. Williams [et al.] // *Transplantation*. – 1999. – Vol. 68. – № 2. – P. 236–241.
72. Development of CD3 cell quantitation algorithms for renal allograft biopsy rejection assessment utilizing open source image analysis software / A. Moon, G. H. Smith, J. Kong [et al.] // *Virchows Archiv*. – 2018. – Vol. 472. – № 2. – P. 259–269.
73. Deep learning-based histopathologic assessment of kidney tissue / M. Hermsen, T. de Bel, M. den Boer [et al.] // *Journal of the American Society of Nephrology*. – 2019. – Vol. 30. – № 10. – P. 1968–1979.
74. Deep learning-based segmentation and quantification in experimental kidney histopathology / N. Bouteldja, B. M. Klinkhammer, R. D. Bülow [et al.] // *Journal of the American Society of Nephrology*. – 2021. – Vol. 32. – № 1. – P. 52–68.
75. Deep learning-based classification of kidney transplant pathology: a retrospective, multicentre, proof-of-concept study / J. Kers, R. D. Bülow, B. M. Klinkhammer [et al.] // *The Lancet Digital Health*. – 2022. – Vol. 4. – № 1. – P. e18–e26.
76. Banff fibrosis study: multicenter visual assessment and computerized analysis of interstitial fibrosis in kidney biopsies / A. B. Farris, S. Chan, J. Climenhaga [et al.] // *American Journal of Transplantation*. – 2014. – Vol. 14. – № 4. – P. 897–907.
77. A fully automated system using a convolutional neural network to predict renal allograft rejection: extra-validation with giga-pixel immunostained slides / Y. G. Kim, G. Choi, H. Go [et al.] // *Scientific Reports*. – 2019. – Vol. 9. – № 1. – P. 5123. – [Electronic resource]. – URL: www.nature.com/articles/s41598-019-41479-5.
78. Automatic evaluation of histological prognostic factors using two consecutive convolutional neural networks on kidney samples / E. Marechal, A. Jaugey, G. Tarris [et

- al.] // *Clinical Journal of the American Society of Nephrology*. – 2022. – Vol. 17. – № 2. – P. 260–270.
79. Deep learning identified pathological abnormalities predictive of graft loss in kidney transplant biopsies / Z. Yi, Y. Li, C. Yang [et al.] // *Kidney International*. – 2022. – Vol. 101. – № 2. – P. 288–298.
80. Evaluation of the Classification Accuracy of the Kidney Biopsy Direct Immunofluorescence through Convolutional Neural Networks / G. Ligabue, F. Pollastri, F. Fontana [et al.] // *Clinical Journal of the American Society of Nephrology*. – 2020. – Vol. 15. – № 10. – P. 1445–1454.
81. CODA: quantitative 3D reconstruction of large tissues at cellular resolution / A. L. Kiemen, A. M. Braxton, M. P. Grahn [et al.] // *Nature Methods*. – 2022. – Vol. 19. – № 11. – P. 1490–1499.
82. Generative Adversarial Networks for Facilitating Stain-Independent Supervised and Unsupervised Segmentation: A Study on Kidney Histology / M. Gadermayr, L. Gupta, V. Appel [et al.] // *IEEE Transactions on Medical Imaging*. – 2019. – Vol. 38. – № 10. – P. 2293–2302.
83. Colour adaptive generative networks for stain normalisation of histopathology images / C. Cong, S. Liu, A. Di Ieva [et al.] // *Medical Image Analysis*. – 2022. – Vol. 82. – P. 102580. – [Electronic resource]. – URL: <https://www.sciencedirect.com/science/article/abs/pii/S1361841522002195?via%3Dihub>.
84. Towards accurate and efficient diagnoses in nephropathology: an AI-based approach for assessing kidney transplant rejection / A. Fayzullin, E. Ivanova, V. Grinin, D. Ermilov, S. Solovyeva, M. Balyasin, A. Bakulina, P. Nikitin, Y. Valieva, A. Kalinichenko, A. Arutyunyan, A. Lychagin, P. Timashev // *Computational and Structural Biotechnology Journal*. — 2024. — Vol. 24. — P. 571–582.
85. GitHub Repository : official website. Detectron2 [Electronic resource] – URL: <https://github.com/facebookresearch/detectron2> (дата обращения: 14.05.2025).
86. National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC). The Clinical Proteomic Tumor Analysis Consortium Clear Cell Renal Cell Carcinoma

- Collection (CPTAC-CCRCC), 2018. США : база данных. URL: <https://www.cancerimagingarchive.net/collections/cptac-ccrcc/> (дата обращения: 15.04.2025). Режим доступа: для всех пользователей.
87. The Cancer Genome Atlas Kidney Renal Clear Cell Carcinoma Collection (TCGA-KIRC), Version 3, 2016. США : база данных. URL: <https://www.cancerimagingarchive.net/collections/kirc/> (дата обращения: 15.04.2025). Режим доступа: для всех пользователей.
88. Cell detection with star-convex polygons/U. Schmidt, M. Weigert, C. Broaddus [et al.] — Текст: непосредственный // Medical Image Computing and Computer Assisted Intervention – MICCAI, 2018. – Singapore: 2018. – P. 265–273.
89. Weiger, M., Nuclei instance segmentation and classification in histopathology images with Stardist/M. Weiger, U. Schmidt. — Текст: непосредственный // IEEE International Symposium on Biomedical Imaging Challenges (ISBIC), IEEE. – Kolkata: IEEE, 2022. – P. 1–4.
90. Predicting Grade and Patient Survival in Renal Cancer Using Machine Learning Analysis of Nucleolar Prominence / E. Ivanova, A. Fayzullin, V. Grinin, D. Zhavoronkov, D. Ermilov, M. Balyasin, A. Timakova, A. Bakulina, Y. Osmanov, E. Rudenko, A. Arutyunyan, R. Parchiev, N. Shved, M. Astaeva, A. Lychagin, T. Demura, P. Timashev // Cancer Medicine. — 2025. — Vol. 14. — № 17. – e71196. – [Electronic resource] – URL: <https://onlinelibrary.wiley.com/doi/10.1002/cam4.71196>.
91. Next-Generation Morphometry for pathomics-data mining in histopathology / D. L. Hölscher, N. Bouteldja, M. Joodaki [et al.] // Nature Communications. – 2023. – Vol. 14. – № 1. – P. 470. – [Electronic resource]. – URL: <https://www.nature.com/articles/s41467-023-36173-0>.
92. Хабр : сайт. – URL: <https://habr.com> (дата обращения: 10.11.2025).