

Федеральное государственное автономное образовательное учреждение
высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»
Физтех-школа Биологической и Медицинской Физики
Кафедра молекулярной и клеточной биологии

Направление подготовки / специальность: 03.03.01 Прикладные математика и физика
(бакалавриат)

Направленность (профиль) подготовки: Молекулярная биофизика и физика живых систем

АНАЛИЗ МИКРООКРУЖЕНИЯ ТКАНИ ОПУХОЛИ МЕТОДОМ МУЛЬТИПЛЕКСНОЙ ИММУНОФЛУОРЕСЦЕНЦИИ

(бакалаврская работа)

Студент:

Бруттан Мария Вячеславовна

(подпись студента)

Научный руководитель:

Атауллаханов Равшан Иноятович,
д-р мед. наук, проф.

(подпись научного руководителя)

Консультант (при наличии):

(подпись консультанта)

Москва 2020

Оглавление

Список используемых сокращений	3
Введение	4
Обзор литературы	6
О раке предстательной железы	6
О микроокружении ткани опухоли	10
Метод циклической иммунофлуоресценции CyCIF	13
Другие методы комплексного исследования тканей	19
Подходы к автоматической сегментации клеток	22
Применение конволюционных нейронных сетей в решении задачи сегментации	23
Материалы и методы	25
Результаты	34
Обсуждение	49
Выводы	50
Список литературы	51

Список используемых сокращений

- MxIF – multiplexed immunofluorescence
- CycIF – cyclic immunofluorescence
- FFPE – formalin-fixed, paraffin-embedded
- FF – fresh frozen
- CNN – convolutional neural network
- mpMRI, MPT – multiparametric magnetic resonance imaging
- H&E – Hematoxylin and Eosin staining
- ИHC – immunohistochemistry
- SOM – self-organizing map
- ISUP – International Society of Urological Pathology
- ВОЗ – Всемирная организация здравоохранения
- GS – Gleason score
- PSA, ПСА – простатспецифический антиген
- MIBI – Multiplexed Ion Beam Imaging
- CODEX – codetection by indexing
- MALDI-TOF – matrix-assisted laser desorption/ionization with time of flight mass-analysator

Введение

Онкологические заболевания – одна из наиболее частых причин смерти в мире. По данным Всемирной организации здравоохранения, рак занимает вторую позицию в списке лидирующих причин смертности, уступая только сердечно-сосудистым заболеваниям.¹ Каждая шестая смерть в мире происходит в следствие опухолевого процесса. При этом разные виды рака отличаются по частоте выявления и опасности для жизни человека. Согласно информации из базы данных GLOBOCAN, на 2018 год, исследуемый в данной работе рак простаты является вторым по частоте диагностирования у мужчин среди всех онкологических заболеваний (13,5% случаев) после рака лёгких (14,5 % случаев).²

Необходимо отметить, что существует тенденция ежегодного увеличения диагностируемых случаев онкологических заболеваний.³⁻⁵ Таким образом, в настоящее время перед обществом особенно остро стоит проблема их выявления и лечения.

Существует большое количество методов исследования тканей опухоли и способов лечения этих заболеваний. Новые методы лечения, такие как таргетная и иммунная терапия, являются следствием значительного научного прогресса в понимании молекулярных механизмов канцерогенеза и разнообразных клеточных взаимодействий в области опухолевого микроокружения. Для определения клеточного состава ткани и внутриклеточной экспрессии различных молекулярных компонент, в современной диагностической практике одним из наиболее популярных является метод проточной цитометрии⁶. Однако, данный метод, как и многие другие лабораторные методы, в которых происходит разрушение структуры самой ткани, позволяет оценить клеточный состав лишь количественно, не давая информации об архитектуре опухоли и расположении клеток относительно друг друга. Возможность достаточно быстро и качественно получать такую информацию из образца открывает большое поле для научных исследований и потенциально может сыграть важную роль в развитии медицины. Принятый в медицинской практике метод иммуногистохимии (ИНС) используется в основном для визуализации небольшого набора маркеров на слайде. Сложность задачи визуализации достаточного набора маркеров для определения местоположения и типов клеток в образце связана с техническими ограничениями для одновременной визуализации разных маркеров, длительным процессом окрашивания, риском разрушения образца в процессе работы, а также необходимостью в разработке и применении алгоритмов машинного обучения для быстрой и качественной обработки полученных изображений и идентификации клеточных контуров.

В последние годы появляются новые методы иммуногистохимических исследований, позволяющие визуализировать до 60 маркеров на одном слайде. В данной

работе будут исследованы возможности одного из методов MxIF (Multiplexed Immunofluorescence) – циклической иммунофлуоресценции (CyCIF)⁷⁻⁹, на примере 15 образцов аденокарциномы предстательной железы и представлен способ обработки и анализа флуоресцентных изображений ткани при помощи машинного обучения.

При диагностике рака простаты и наблюдении за течением этого заболевания используется метод mpMRI^{10,11}, который безопаснее регулярного проведения биопсии, так как было показано, что частый забор биологического материала может способствовать развитию воспаления у больного. Такой побочный эффект наблюдался у 1-4% пациентов.¹² Важно отметить, что MRI не всегда позволяет выявить злокачественное разрастание ткани простаты. Так, в работе 2019 года¹³ было показано, что 23 из 89 пациентов получило ложно-отрицательное заключение на основании МРТ-исследования, при этом 17.4% из них имели средний и высокий риск в соответствии с Decipher Genomic Classifier.¹⁴ Это же подтверждается рядом других исследований, которые своей целью ставили определение факторов, влияющих на видимость злокачественного образования при mpMRI.^{15,16}

Целями данной работы являлись изучение возможностей метода MxIF при исследовании образцов аденокарциномы предстательной железы и поиск критериев, позволяющих определить различие в образцах, определяемых (MRI-visible) и не определяемых (MRI-invisible) как опухолевые по результатам mpMRI.

Задачи для реализации поставленной цели включали в себя:

- решить задачу сегментации клеток на основе изображений, полученных методом мультиплексной иммунофлуоресценции;
- определить клеточные типы на основании классификации по экспрессии имеющихся маркеров;
- проанализировать взаимное расположение клеток на срезах;
- реализовать метрики сравнения опухолевого и не злокачественного эпителия, а также стромального микроокружения опухоли в тканях MRI-visible и MRI-invisible пациентов.

Обзор литературы

О раке предстательной железы

Среди всего многообразия видов онкологических заболеваний наиболее часто диагностируемыми и широко представленными являются карциномы, или рак, происходящий из эпителиальных клеток различных органов. Аденокарциномами называют злокачественные новообразования, которые формируются из секретирующих эпителиальных клеток.¹⁷ По данным Американского Института Исследования Рака, среди всех диагностированных случаев онкологии у мужчин в 2018 году, рак простаты располагается на втором месте после рака лёгких по частоте встречаемости.⁵¹ Также необходимо отметить, что приблизительно в 95% случаев рак предстательной железы является аденокарциномой.^{18,19}

Предстательная железа является экзокринной железой мужского организма у млекопитающих, располагающейся в области малого таза. Простата состоит из нескольких трубно-альвеолярных желез (30-50), которые выделяют секрет непосредственно в мочеиспускательный канал, железы окружены стромальным компонентом с большой долей гладкой мускулатуры в своем составе. При рассмотрении строения простаты в ней выделяют 4 железистые зоны: центральную (ЦЗ, CZ), которой соответствуют примерно 25% железистой ткани, периферическую (ПЗ, PZ), занимающую 70%, и две переходные (транзиторные) зоны (ТЗ, TZ), которым принадлежит 5% всей железистой ткани.²⁰ Основание центральной зоны находится у шейки мочевого пузыря, а вершина располагается у мочеиспускательного канала, окружая часть проксимального отдела мочеиспускательного канала; ее пересекают эякуляторные протоки. Ацинусы крупные, с неровными контурами, а разветвления очень сложные, с однородной текстурой. Периферическая зона занимает середину и вершину простаты, поэтому является наиболее доступной областью для цифрового ректального исследования. Переходная зона расположена вокруг проксимального отдела мочеиспускательного канала и имеет небольшие, простые, округлые железы, окруженные стромой, подобной той, что находится в центральной зоне.

Злокачественные изменения в каждой из зон происходят неравномерно, наибольшее число случаев рака предстательной железы (70%) связаны с видоизменениями ПЗ, тогда как опухоли в ЦЗ составляют 5-10% от всех случаев. Простата не содержит капсульной структуры как таковой, существует только конденсация железистой стромы, называемая “капсулой”. Эта “капсула” содержит фибромышечную строму, образующую тонкую пластинку вокруг железы. Верхушка предстательной железы лишена этой

стромальной пластинки, что приводит к возникновению области анатомической слабости, называемой трапециевидной областью. Эта область предстательной железы играет важную роль в распространении опухоли, возникшей в районе верхушки железы, в другие ее части.

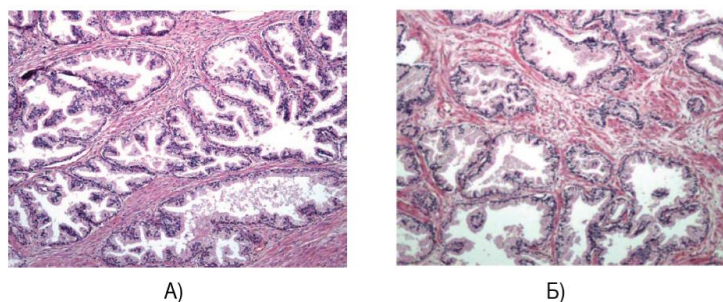


Рисунок 1 – Н&Е изображение здоровой ткани предстательной железы А) переходной зоны и Б) периферической зоны. Источник: Joan C. Vilanova et al. “Atlas of Multiparametric Prostate MRI” \ Springer International Publishing AG – 2018 – С.48

Как уже говорилось выше, рак предстательной железы – это железистая злокачественная неоплазия, преимущественно секреторного фенотипа, и лишь в небольшом проценте случаев она имеет другую гистологическую морфологию. Данный тип рака относится к сравнительно медленно прогрессирующим заболеваниям. Однако при длительном развитии и при отсутствии лечения ткань может разрастаться и сдавливать близлежащие органы, а также на поздних стадиях может врастать в органы малого таза, опухолевые клетки с током крови могут распространяться в другие органы и таким образом провоцировать развитие отдаленного метастазирования.

Основным методом диагностики данного заболевания является анализ уровня простатспецифического антигена (PSA, ПСА), однако необходимо учитывать тот факт, что «ПСА – это органоспецифичный маркер, который в норме секретируется эпителиальными клетками предстательной железы. Сывороточный уровень ПСА может повышаться не только при наличии опухоли, но и при наличии доброкачественной гиперплазии, а также при хроническом простатите». Граничным значением специалисты обычно называют 4 нг/мл. Однако показано, что уровень ПСА в крови от 3 до 4 нг/мл ассоциируется с риском рака простаты порядка 27% и риском опухоли с GS >7 в районе 6.7%. То есть нет такого значения ПСА, при содержании фермента ниже которого человек может быть уверен, что у него нет риска развития рака предстательной железы.²¹

Типичными морфологическими признаками опухоли предстательной железы являются микрогландулярная монослойная пролиферация с отсутствием базальных клеток и нерегулярными ядрами. Для оценки рака предстательной железы ВОЗ рекомендует использовать шкалу Глисона в соответствии с критериями и группами оценок

Международного общества урологической патологии (ISUP) 2014 года.²² В основе оценки региона по системе Глисона лежат архитектурные изменения ткани.

Аденокарциномы классифицируют как:

- мелкоацинарная;
- крупноацинарная;
- солидно-трабекулярная;
- эндометриоидная;
- железисто-кистозная;
- слизеобразующая;
- крибозная;
- папиллярная.

Так как ткань опухоли предстательной железы в большинстве случаев имеет неоднородную морфологическую структуру, гистопатологами выделяется наиболее распространённая гистологическая градация, которая определяет первичный балл по шкале Глисона, далее выбирается следующий по частоте встречаемости паттерн дифференцировки, таким образом, выделяется вторичный балл. Сумма Глисона (от 2 до 10 баллов) является результатом сложения первичной и вторичной оценок.²³

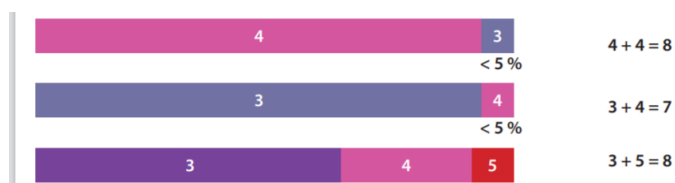


Рисунок 2 – Примеры определения суммы Глисона. Источник: М.В. Ковылина, Е.А. Прилепская, Д.Ю. Пушкарь “ Патоморфологическая диагностика рака предстательной железы, рака мочевого пузыря и рака почки” \ \ ООО «Издательский дом «АБВ-пресс» – 2017 – С.10

При этом гистологи выделяют 4 паттерна видоизменений ткани в соответствии с ISUP:

1. Слитые железы состоят из группы желез, не разделенных стромой;
2. Крибриформные железы представляют собой разросшийся железистый эпителий с множественными забитыми просветами;
3. Плохо определяемые железы означают плохо сформированные или отсутствующие железистые просветы. Классификация по этому подтипу требует наличия скопления таких желез, чтобы исключить возможность желез, относящихся к паттерну Глисона 3;

4. Гломерулоидные железы определяются как расширенные железы, содержащие внутрипросветные кривбриформные структуры с одной точкой прикрепления, что придает схожесть с почечным клубочком.

На основании наличия или отсутствия данных структур и их комбинаций выделяются 5 групп опухоли. Описание каждой группой и ее значение по шкале Глисона (GS) указаны в Таблице 1.^{20,22}

Группа	GS	Описание
1	≤ 6	Присутствуют только отдельные хорошо сформированные ацинусы
2	$3+4=7$	В основном хорошо сформированные ацинусы с небольшим числом компонент, относящихся к паттернам 1-3
3	$4+3=7$	В основном железы, относящиеся к паттернам 1-3, при этом присутствуют нормально сформированные ацинусы
4	8	Присутствуют только железы, относящиеся к паттернам 1-3, или в основном сформированные ацинусы и небольшой участок, лишенный желез, или в ацинусы в основном отсутствуют и есть небольшой участок с нормально сформированными ацинусами
5	9-10	Отсутствие желез (возможно наличие некроза), выделение паттернов 1-3 не является обязательным

Таблица 1 – Описание характеристик опухоли для разных значений суммы Глисона

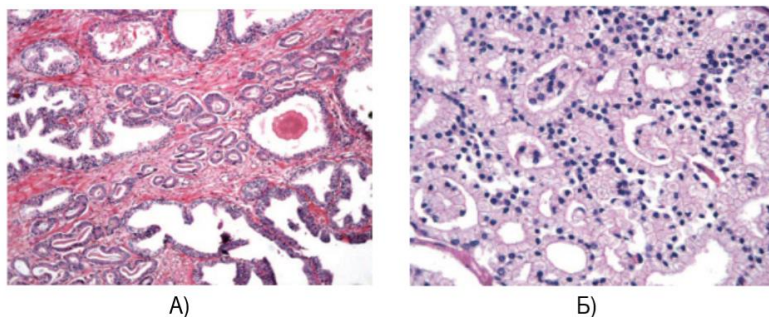


Рисунок 3 – H&E изображение опухоли предстательной железы А) GS = 3 и Б) GS = 4 со слитыми железами. Источник: Joan C. Vilanova et al. “Atlas of Multiparametric Prostate MRI” \ Springer International Publishing AG – 2018 – С.49

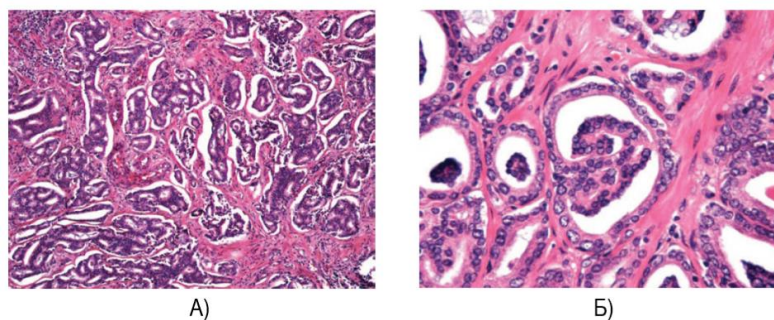


Рисунок 4 – Н&Е изображение опухоли предстательной железы А) GS = 4 с крибриформным паттерном и Б) GS = 4 с гломерулоидными железами. Источник: Joan C. Vilanova et al. “Atlas of Multiparametric Prostate MRI” \ Springer International Publishing AG – 2018 – С.49-50

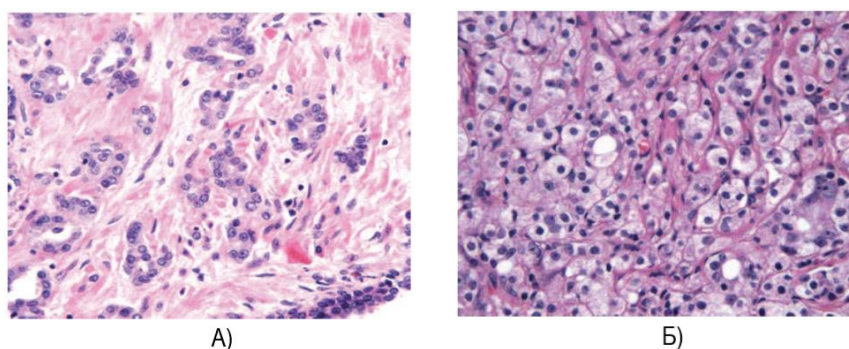


Рисунок 5 – Н&Е изображение опухоли предстательной железы А) GS = 4 с паттерном плохо сформированных желез и Б) GS = 5. Источник: Joan C. Vilanova et al. “Atlas of Multiparametric Prostate MRI” \ Springer International Publishing AG – 2018 – С.50

При классификации опухоли к одной из групп необходимо также учитывать, ткань какой из зон предстательной железы рассматривается, так как разные части отличаются по своей морфологии, что сказывается на результатах исследования, в том числе при проведении mpMRI. Так периферическая зона характеризуется рыхлым стромальным компартментом, который распределяется разнонаправленно, формируя сеть, отличается высоким содержанием воды. В центральной части простаты строма компактна и имеет толстые мышечные пучки с меньшим содержанием воды, чем в периферической зоне.

О микроокружении ткани опухоли

Микроокружение опухоли включает в себя здоровые клетки организма, внеклеточный матрикс и секретируемые клетками вещества, кровеносные сосуды, которые окружают и/или питают опухоль, которая в свою очередь также может влиять на микроокружение. Оно в свою очередь также оказывает влияние на злокачественное образование.^{17,24,25} В ряде исследований было показано, что взаимодействие между

эпителиальными клетками и стромальными важно для регуляции пролиферации, инвазии, ангиогенеза и метастазирования опухоли.²⁶⁻²⁸ Принимая во внимание этот факт, в последнее время при типировании опухолей используются не только фенотипические и генотипические изменения в раковых клетках, но и учитывается их окружение. Этот фактор также играет важную роль в прогнозировании течения заболевания и его лечении.

Микроокружение может как поддерживать опухоль, так и сдерживать ее развитие. Впервые это было показано в 1984 году на куриных эмбрионах,²⁹ в которых при введении сильного онкогена не наблюдалось развития злокачественных новообразований. Авторы исследования предположили, что в эмбрионах не хватает факторов, необходимых для поддержания опухоли, иными словами, исследователи наблюдали влияние микроокружения на её развитие.

Также было показано, что стромальные клетки могут создавать окружение, благоприятно влияющее на развитие и метастазирование опухоли. Например, клетки костного мозга могут колонизироваться в местах отдаленных метастазов опухоли и таким образом создавать некий каркас, который в дальнейшем будет способствовать благоприятному развитию рака на новом месте. Аналогичную роль играют и процессы, нарушающие окружение тканей в результате хронического воспаления или мутагенеза.^{24,27}

Сказанное выше подтверждает гипотезу, что нацеливание на микроокружение может стать важным терапевтическим подходом для лечения онкологических заболеваний. Преимуществом данного подхода можно назвать тот факт, что здоровые клетки, составляющие микроокружение, являются генетически стабильными и, таким образом, менее пластичны, то есть менее склонны приобретать устойчивость к лекарственным препаратам.²⁴ Принимая во внимание данное наблюдение, в настоящее время ведется разработка и начинается применение таких методов лечения рака, как ингибиторы контрольных точек³⁰ и ряд других. В то же время было показано, что при удалении макрофагов, ассоциированных с карциномой в мышинных моделях метастатического рака молочной железы и немелкоклеточного рака легких, наблюдается снижение ангиогенеза, роста и метастазирования опухоли. При таргетировании опухоль-ассоциированных фибробластов в мультирезистентных карциномах толстой кишки и молочной железы³¹ у мышей подавляется первичный рост опухоли и увеличивается поглощение химиотерапевтических препаратов.

mpMRI исследование ткани опухоли предстательной железы

Как уже было упомянуто во введении, одним из методов подтверждения диагноза у пациентов с подозрением на рак простаты является метод mpMRI,^{10,11} который можно считать безопаснее биопсии. Основанием для данного утверждения является тот факт, что в среднем от 1 до 4% пациентов^{12,32,33} могут иметь осложнения после биопсии, выражающееся в развитии воспаления. Однако при mpMRI некоторые области не идентифицируются как опухолевые, при этом некоторые из этих не видимых на МРТ образцов классифицируются как области повышенного риска разрастания и возможного метастазирования. Например, в статье 2019 года¹³ было показано, что из 89 пациентов, принимавших участие в исследовании, 23 образца не были детектированы при проведении mpMRI как опухолевые, но 17.4% из них имели средний и высокий риск в соответствии с the Decipher Genomic Classifier.¹⁴ Также в другом исследовании на большем количестве образцов (1213 региона на 588 пациентов) было показано, что при данном исследовании у 34% от всех пациентов не учитывается в среднем один клинически значимый регион.³⁴ И было показано, что структуру опухоли (мультифокальная или одиночная) и ее размер можно считать лимитирующими факторами для данного типа исследований. Однако точные критерии, которые отличают одни образцы от других и влияют на видимость опухоли на МРТ-диаграмме, остаются неясными. Ряд исследований был посвящен данному вопросу. Например, при исследовании 10 пациентов и 26 опухолевых регионов было обнаружено, что 14 областей были не видны на МРТ, что составляет 54% от общего числа исследуемых областей.¹⁶ Учитывая малый размер выборки, сложно делать выводы о различии двух групп регионов, но авторы создали модель, которая на разнице в экспрессии 9 генов предсказала с достаточно высокой точностью результат МРТ-диагностики на тестовой когорте из 375 пациентов. Данные гены были связаны с клеточной и стромальной организацией. При этом не было замечено статистически значимых связей видимости на МРТ со способностью к метастазированию и влиянию на выживаемость. Данный результат показывает, что регионы, которые не видны при проведении mpMRI, являются опасными для пациента. Исследования на модельных мышях также доказывают тот факт, что экспрессия генов, связанных с ростом клеток и клеточной организацией, коррелирует с видимостью на МРТ, то есть пониженная плотность клеток и их пролиферативный потенциал могут быть факторами, определяющими невидимость опухоли при mpMRI.¹⁵ При этом отмечается, что несмотря на в целом низкую точность диагностики, данный метод позволяет выявлять клинически

значимые регионы и улучшает точность биопсии, которая уступает в чувствительности диагностики радикальной простатэктомии, но является более предпочтительной для пациентов.^{35,36}

Как говорилось в части, посвященной строению предстательной железы и характеристикам ее злокачественных видоизменений, разные зоны простаты отличаются по своей морфологии, в том числе по содержанию воды в них, что сказывается на результатах МРТ. При этом существует некоторая корреляция между описанными ранее группами и возможностями МРТ-диагностики. Карциномы предстательной железы 1-й группы ($GS \leq 3+3$) образованы железами, похожими друг на друга, с регулярной стромой, которая располагается между нормальными железами, так что количество воды может варьироваться и напоминать периферическую зону нормальной ткани. Опухоли с $GS = 4$ могут иметь разнородные области. Для подтипов паттерна 4 со слитыми и плохо сформированными железами характерны плотно прилегающие друг к другу клетки с небольшим количеством межклеточной воды и поэтому с малой подвижностью молекул. С другой стороны, кривриформный и гломерулоидный подтипы характеризуются рыхлыми железами с обильным количеством воды и большей подвижностью молекул, напоминающих нормальные ткани периферической зоны. При этом для участков с $GS = 5$ структура ткани может представлять собой как компактные скопления неопластических клеток с небольшим количеством межклеточной жидкости, так и клеточные скопления между участками толстой соединительной ткани или твердые образования с центральным некрозом. Поскольку *mpMRI*-изображение зависит от клеточного распределения и подвижности молекул воды, легко понять, что некоторые неопластические патологии могут быть ошибочно приняты за опухолевое поражение при условии, что они характеризуются достаточно большими скоплениями изолированных клеток в жидкой среде или некрозом, как это происходит при воспалительных процессах.

Метод циклической иммуофлуоресценции CyCIF

В последнее время при диагностике и лечении злокачественных заболеваний всё больше внимания уделяется микроокружению ткани опухоли. Ввиду этого возрастает необходимость в разработке и использовании методик, позволяющих визуализировать большое количество биомаркеров в рамках одного образца. На сегодняшний день уже существует некоторое количество методов, позволяющих осуществлять комплексное

исследование FFPE-образцов. Все они являются относительно новыми, каждый имеет ряд своих достоинств и недостатков. При одновременной окраске несколькими антителами с флуоресцентными метками, необходимо учитывать ширину спектра излучения используемых флуорофоров для предотвращения перекрытия сигналов.⁷

До появления методов, позволяющих проводить одновременную визуализацию многих белков, присутствующих в образце, основными способами исследования гетерогенности опухоли, ее состава в эпителиальной и стромальной части, иммунного инфильтрата и возможных взаимодействий между клетками являлись метод проточной цитометрии и секвенирование РНК единичных клеток.^{37,38} Данные методы показывают, что опухоли содержат множество клеточных состояний с различными транскрипционными программами. Таким образом, более глубокое понимание спектра и динамики клеточных состояний при опухолевых заболеваниях может иметь решающее значение для создания достоверных моделей и продвижения терапевтических стратегий, направленных на решение сложных проблем этого вида заболеваний. Однако главным недостатком этих методов является необходимость разрушения образца, при котором теряется информация о взаимном расположении клеток на слайде.

Указанного выше недостатка лишены методы иммуногистохимии. Методика ИНС исследования заключается в обработке образца антителами, конъюгированными с флуоресцентным красителем или ферментом, для идентификации местоположения выбранных эпитопов.

В данной работе рассматриваются возможности метода мультиплексной иммунофлуоресценции, который также называется методом циклической иммунофлуоресценции, позволяющего одновременно визуализировать большое количество молекулярных маркеров (точное количество зависит от типа ткани и требований специалистов).^{7,9} Таким образом, данный метод позволяет не только узнать фенотипический портрет клеток в виде экспрессии различных маркеров, но и показать сложную структуру новообразования. Одним из преимуществ данного метода является его доступность: открытые протоколы проведения эксперимента и использование классических реагентов.

Преимущества метода CyCIF:

- Позволяет получить информацию о пространственном распределении маркеров на плоских срезах;
- Дает возможность изображения приблизительно до 60 маркеров для каждого образца в широком диапазоне типов тканей и опухолей³⁹;

- Позволяет исследователям использовать интересующие их наборы антител для анализа конкретных тканей и решения определенных задач;
- Не требует обязательного использования специального оборудования и реагентов;
- Решает проблему сбора информации о больших образцах;
- Генерирует данные высокого разрешения с помощью оптической микроскопии;
- Позволяет работать как с FF-, так и с FFPE-препаратами.

При использовании данного метода окрашивание антителами с флуоресцентными метками производится циклически. На первом этапе для уменьшения фоновой флуоресценции и неспецифического связывания антител проводится этап предварительного окрашивания. Данный этап включает инкубацию образца с вторичными антителами с последующим окислением флуорофора в растворе перекиси водорода с высоким рН в присутствии света (этот шаг также называется «fluorophore bleaching»). Следующие циклы состоят из 4 шагов:

1. Окрашивание антителами к 3-4 маркерам;
2. Окрашивание ДНК, как правило, с помощью DAPI (как в случае данных от коллабораторов BostonGene) или Hoechst 33342 (как описано в статье, посвященной исследованию данного метода⁹);
3. Снятие многоканального изображения при высоком и низком увеличении;
4. Отделение антител с флуорофором от образца и их последующая смывка в буферном растворе.

После завершения всех циклов производится H&E-окрашивание.

Визуализация образцов площадью несколько квадратных сантиметров с разрешением ~1 мкм обычно выполняется на микроскопах, специализированных для сканирования флуоресцентных слайдов.

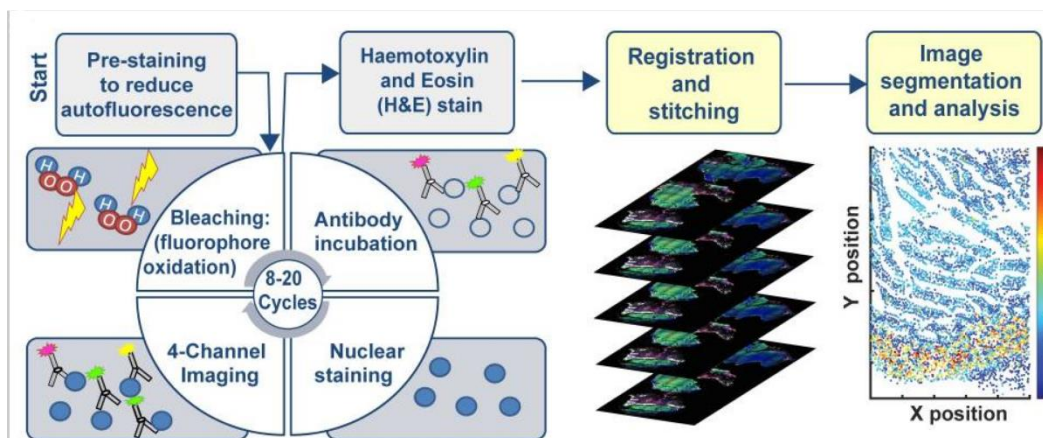


Рисунок 6 – Схема применения метода CyCIF. Источник: Lin et al. “Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes” \ eLife – 2018

Необходимость в окрашивании ДНК на каждом цикле объясняется необходимостью контроля качества изображений и облегчения сопоставления изображений, полученных на разных циклах.

Показано, что соотношение сигнал/шум увеличивается с каждым последующим этапом, как результат проведения процедур предобработки образца и последующей смывки флуорофоров на каждом из циклов. При этом необходимо заметить, что эффективность инактивации флуорофора в условиях высоких значений pH изменяется в зависимости от флуорофора, но практически не зависит от антитела, с которым флуорофор связан.

В качестве проверки влияния процедуры на сохранность препарата для 48 образцов биопсии 16 различных здоровых и опухолевых тканей проводилось 10 циклов окрашивания. Было замечено, что скорость возникновения повреждений зависит от типа ткани. Так, для образцов рака простаты к 10 циклу сохранялось до 70% от исходного числа клеток, в то время как, например, для образцов ткани легких и лимфатических узлов этот показатель составлял 90%. Немаловажным является тот факт, что неравномерность сохранности клеток наблюдается и в пределах одного типа ткани.

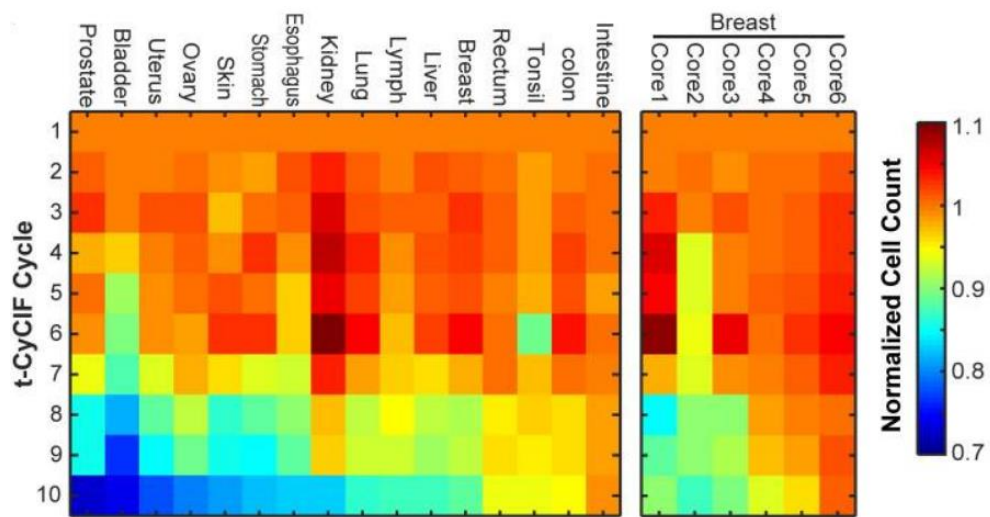


Рисунок 7 – Иллюстрация сохранности тканей на разных циклах CyCIF в зависимости от типа ткани и образца. Источник: Lin et al. “Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes” \ eLife – 2018

Так как образцы обрабатываются антителами к разным белкам циклически, необходимо принимать во внимание проблему воспроизводимости результатов в зависимости от очередности маркеров и окрашиваемых регионов. Для окрашивания одним и тем же набором маркеров на разных циклах для одного образца был определен коэффициент корреляции, значения которого составляли от 0.8 до 0.95 в зависимости от исследуемого маркера. При этом после масштабирования изображений для выравнивания диапазона интенсивности (путем выравнивания гистограммы), паттерны окрашивания были неразличимы во всех циклах. При сравнении окрашивание между разными слайдами для одних и тех же антител и номеров циклов, перекрытие в распределениях интенсивности было высоким (>0.85), что демонстрирует хорошую воспроизводимость от образца к образцу. В целом, было обнаружено, что повторяемость результатов окрашивания была самой высокой, когда антитела использовались в одном цикле на обоих образцах, более низкой при использовании антител в разных циклах на образце, и самой низкой, когда и порядок, и образец были разными. Таким образом, для минимизации различий в окрашивании рекомендуется проводить исследование, обрабатывая все FFPE-образцы в одном порядке.

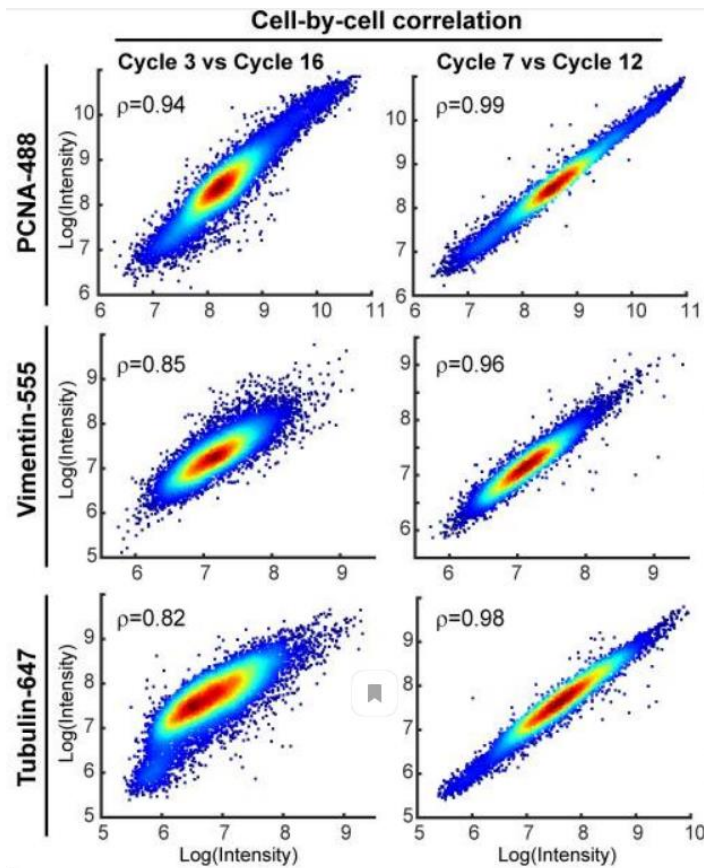


Рисунок 8 – Корреляция окрашивания трех маркеров на разных циклах в рамках одного образца. Источник: Lin et al. “Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes” || eLife – 2018

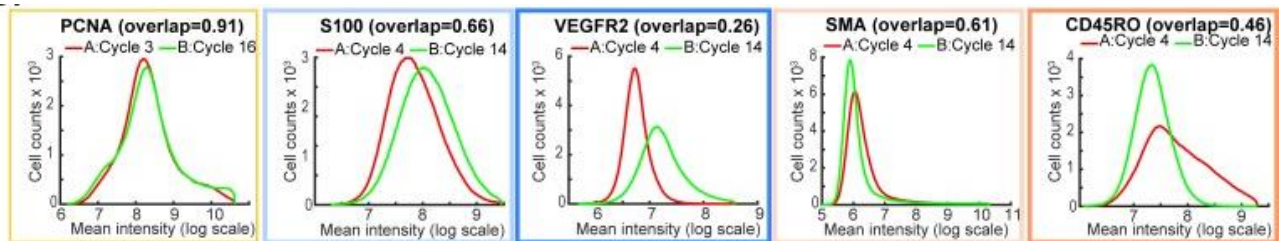


Рисунок 9 – Перекрытие спектров флуоресценции для 5 маркеров на 4 и 14 циклах окрашивания CyCIF. Источник: Lin et al. “Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes” || eLife – 2018

Effect of cycle number	Overlap Integral	Antibodies
Little or no change	>0.8	PCNA, E2F1, gH2ax, cJUN, NGFR, pERK, HLA, MITF, pS6, Keratin, beta-Catenin, E-Cadherin, FoxP3
Minor Increase with Cycle Number	0.6 to 0.8	S100, Ki67, pH3, Vimentin, pRB
Major increase with Cycle number	<0.6	VEGFR2, EGFR, CD8a, Tubulin
Minor Decrease with Cycle Number	0.6 to 0.8	CD11b, HES1, SMA
Major Decrease with Cycle Number	<0.6	CD45RO, PD1, CD3

Рисунок 10 – Сводная таблица повторяемости спектров флуоресценции в зависимости от цикла окрашивания. Источник: Lin et al. “Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes” || eLife – 2018

Учитывая всё, сказанное выше, можно утверждать, что изображения сусIF позволяют наглядно показать особенности в масштабе всей опухоли, отдельных опухолевых клеток и субклеточных структур.

Другие методы комплексного исследования тканей

Помимо упомянутого выше метода, существует ряд других, позволяющих проводить многопараметрическое исследование препаратов различных тканей и получать изображения различных маркеров с высоким разрешением. В качестве примеров таких методов будут рассмотрены Multiplexed Ion Beam Imaging (MIBI™) Technology и codetection by indexing (CODEX). Данные методы отличаются тем, что позволяют проводить анализ без многократных процедур окрашивания и смывки, а также избежать проблемы перекрывания сигналов различных мишеней(в случае MIBI), или же снижают влияние неспецифического окрашивания и разрушения образца за счёт особой технологии присоединения флуорофора.

Первым рассмотрим метод получения мультиплексных изображений методом направленного ионного потока (MIBI) в сочетании с MALDI-TOF масс-спектрометрией, позволяющий проводить *in situ* анализ содержания нескольких десятков белков. При первом описании данного метода говорилось, что он дает возможность визуализировать до ста мишеней на одном FFPE-образце.⁴⁰ В данном методе к антителам прикреплены изотопы металлов группы лантаноидов, которые не свойственны биологическим объектам. После чего проводится облучение образца пучком первичных ионов методом “мягкой” ионизации MALDI, с последующим анализом для каждой точки образца её изотопного состава, для чего используется времяпролетный масс-анализатор. В зависимости от интересующего элемента, чувствительность MIBI может достигать $1/10^6$ с динамическим диапазоном 10^5 и разрешением, сравнимым с высоким увеличением световой микроскопии.^{40,41} При этом могут использоваться как FFPE-ткани, так и суспензии клеток, в том и другом случаях образец необходимо зафиксировать на проводящей подложке для облучения ее ионным пучком. Также было показано, что прикрепление изотопа металла не влияет на способность антител узнавать свои эпитопы, демонстрируя при этом сопоставимый уровень интенсивности окрашивания с использованием немодифицированных антител.

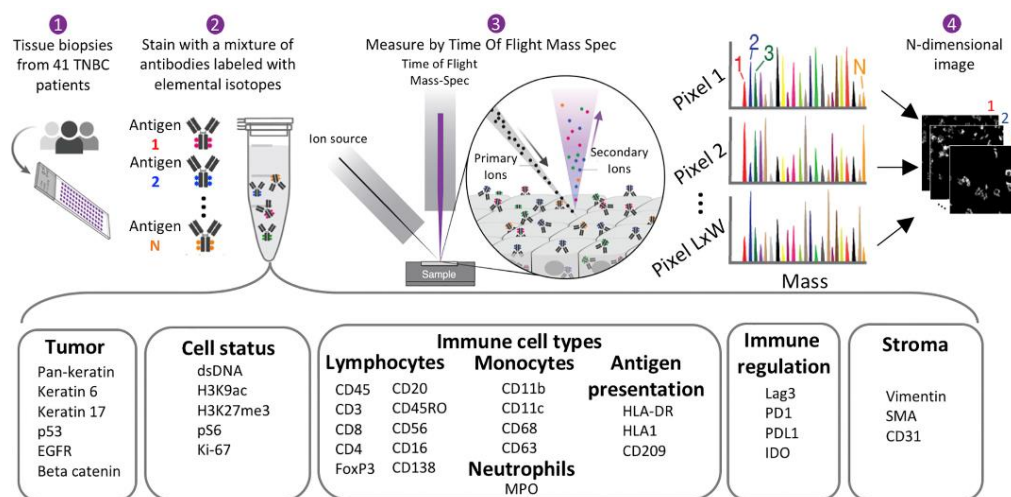


Рисунок 11 – Схема проведения исследования методом MIBI-TOF. Источник: Keren et al. “A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging” \ Cell – 2018

Преимуществами данного метода является отсутствие фонового сигнала, обусловленного аутофлуоресценцией, при этом динамический диапазон составляет уже 10^5 , превышая этот показатель для иммунофлуоресценции и ферментативного ИНС в 100 и 1000 раз соответственно. Поскольку массовая точность может разрешать меньше доли Дальтона даже при самом низком разрешении, то никакого спектрального перекрытия между соседними по массе элементарными репортерами не наблюдается. Ко всему прочему, поскольку массовые метки не разрушаются, образцы остаются стабильными в течение долгого времени.

Тем не менее, данный метод не решает проблему неспецифического связывания антител и невозможности исследовать все интересные эпитопы посредством антител.

В то же время информация о большом числе молекул, присутствующих в образце позволяет изучать клеточный состав, в том числе иммунный инфильтрат. При этом было показано, что изучение совместного присутствия различных типов иммунных клеток позволяет определить скоординированность иммунного ответа на присутствие злокачественного новообразования. Еще одним немаловажным фактором является степень васкуляризации, которая играет двоякую роль, с одной стороны доставляя питание тканям опухоли, с другой стороны являясь источником клеток иммунитета.

Так как этот метод позволяет сохранить информацию о пространственной организации FFPE-ткани, можно выделять области, где присутствует иммунный инфильтрат или где клетки иммунной системы располагаются на периферии злокачественного образования. А также имеется возможность оценить индивидуальную эффективность иммунотерапии путем оценки не только расположения иммунных клеток

относительно опухоли, но и визуализацией экспрессии регуляторных белков, таких как PD-1, PD-L1, IDO.

Другим методом получения мультиплексных изображений является CODEX. Преимуществом данного метода перед описанным выше MIBI является компактность и простота прибора, который автоматизирует подачу реагентов. Изначально данный метод был создан для замороженных образцов биопсии,⁷ но сейчас уже существуют оптимизации CODEX для использования на FFPE-данных.⁴² Основным отличием от описанных выше методов CyCIF и MIBI является то, что антитела, которыми обрабатывается исследуемая ткань, имеют уникальную олигонуклеотидную последовательность, которая является субстратом для отжигания на ней соответствующей последовательности с флуоресцентной меткой. Таким образом, возможно одновременное закрепление до 40 антител со специфичными полинуклеотидными метками, после чего проводится поэтапное окрашивание групп эпитопов по 2-5 флуорофоров, соединенных с олигонуклеотидами.

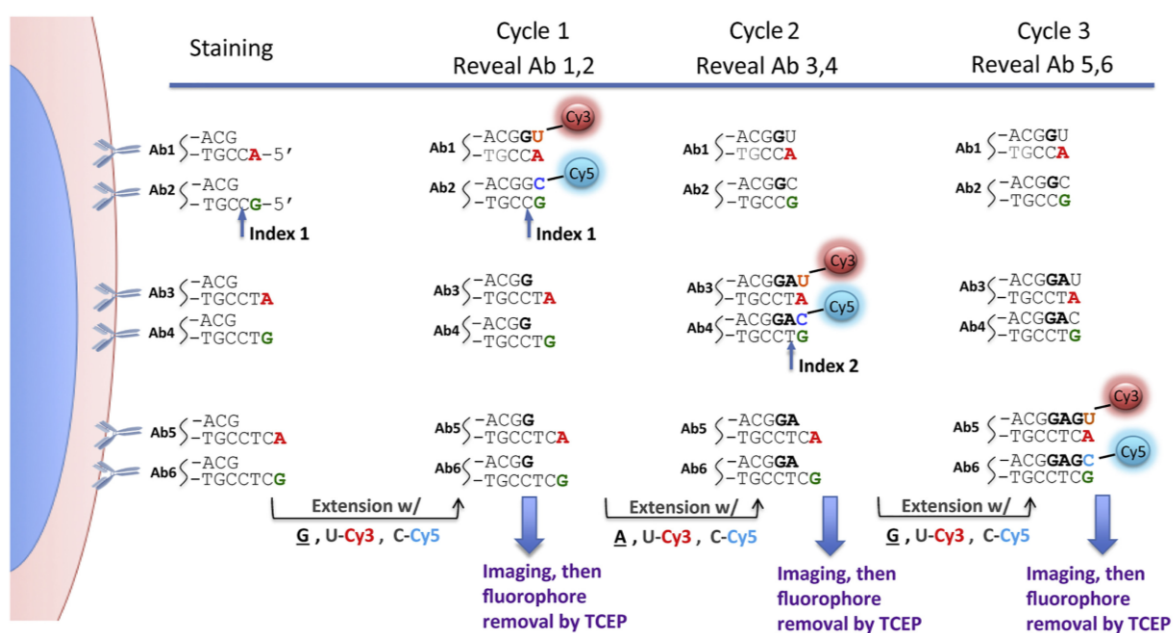


Рисунок 12 – Схема работы метода CODEX. Источник: Goltsev et al. “Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging” || Cell – 2018

Визуализация циклов окрашивания CODEX показала ровные цикловые сигналы с низким фоном (среднее отношение сигнал/шум = 85/1), эффективным (~98%) высвобождением флуорофора путем расщепления с помощью TCEP (Трис(2-карбоксиил)фосфина гидрохлорид) и отсутствием переноса сигнала между циклами. Линейный регрессионный анализ выявил низкое ухудшение сигнала (~0,79% за цикл) и приемлемый фон (начиная с ~1,1% и увеличиваясь на 0,06% за цикл).

Данный метод, как и описанные выше, позволяет визуализировать клеточные контакты, оценивать иммунный инфильтрат, а также выделять так называемые i-niche, определяемые как клеточное окружение конкретной клетки, таким образом иллюстрируя отличия в микроокружении различных образцов опухоли.

Все описанные выше методы позволяют проводить исследование образцов FFPE без повреждения их архитектуры, при этом получая информацию о локализации нескольких десятков белков в рамках одного участка, что дает возможность проводить комплексное изучение опухолей различной природы.

Подходы к автоматической сегментации клеток

Для проведения количественных оценок изображений тканей опухолей необходимо провести выделение отдельных клеток образца. Классическим примером программы, позволяющей осуществлять автоматическое выделение клеточных контуров, является CellProfiler (<https://cellprofiler.org/>), который после первого описания в 2006 году за 4 года своего существования был процитирован в более чем 250 статьях.⁴³ Данная программа позволяет решать различные биологические вопросы количественно, включая стандартные анализы (например, количество клеток и их приблизительный размер, содержание белка на основании флуоресцентного сигнала в каждой клетке) и в некоторых случаях сложные морфологические анализы (например, форма клеток/органелл или субклеточных структур).⁴⁴ CellProfiler доступен в виде приложения для Macintosh OS X и Microsoft Windows и может быть скомпилирован для Linux.⁴³

Первые объекты, идентифицируемые на изображении (называемые первичными объектами), часто являются ядрами, определяются они по окрашенным ДНК изображениям, хотя первичными объектами также могут быть и целые клетки. Как правило, ядра клеток обычно легче поддаются идентификации благодаря их более однородной морфологии, высокому контрасту относительно фона при окрашивании и неперекрыванию между соседними ядрами. В CellProfiler встроено несколько простых алгоритмов для случаев, когда первичные объекты хорошо разделяются, не сливаются и ярки относительно фона. Для определения ядер, как правило, используется алгоритм Otsu,⁴⁵ в основе которого лежит поиск порогового значения для отсечки невалидных значений путем максимизации межклассовой дисперсии. При этом для эффективной

идентификации сливающихся объектов CellProfiler используется модульная трехступенчатая стратегия. Во-первых, трудно определяемые объекты распознаются и разделяются; во-вторых, обнаруживаются разделительные линии между объектами; и в-третьих, некоторые из полученных объектов либо удаляются, либо объединяются вместе на основе измерений, например, их размера или формы. После того, как первичные объекты идентифицированы, края вторичных объектов, которые окружают каждый первичный объект (как правило, клеточные мембраны), могут быть определены, данный этап считается проще идентификации первичных. Другие субклеточные компартменты также могут быть обозначены на слайде, включая цитоплазму (как часть каждой клетки, исключая ядро) и клеточную или ядерную мембрану (край клетки или ядра соответственно). Далее программа позволяет определить большое количество признаков для каждой идентифицированной клетки или субклеточного компартмента, включая площадь, форму и интенсивность.

Достоинствами данного программного обеспечения является возможность автоматизации процесса сегментации и вычисления параметров объектов на изображении, а также работа с наборами слайдов, относящихся к разным регионам одного пациента и/или разных пациентов. Учитывая, что типичный анализ изображений занимает примерно две минуты, один процессор может обрабатывать 30 изображений в час, а кластер из 100 процессоров обрабатывает 3000 изображений в час.⁴⁴ Изображения для анализа могут содержаться в одной директории или в наборе директорий. В то время как сотни изображений могут быть проанализированы на одном компьютере, такого вычислительного решения недостаточно для тысяч или миллионов изображений, характерных для крупномасштабных экранов. В последнем случае анализ CellProfiler может быть запущен на вычислительном кластере, используя преимущества аппаратной инфраструктуры для параллельной обработки любого количества изображений.

Применение конволюционных нейронных сетей в решении задачи сегментации

Нетрудно заметить, что описанная выше программа CellProfiler имеет ряд недостатков, например, трудности с определением клеток 3D структур, где ядра могут накладываться друг на друга; определение клеточных границ как расширение на заданное расстояние относительно клеточной мембраны не всегда является эффективным. Ручное

выделение клеток при большом объеме данных не представляется возможным, поэтому в последнее время разрабатываются архитектуры нейронных сетей для решения задачи клеточной сегментации.⁴⁶ В 2015 году была представлена конволюционная нейронная сеть UNet, представляющая собой автоэнкодер, которая позволяет проводить сегментацию при наличии небольшого количества тренировочных данных.⁴⁷ Положительной чертой данной архитектуры является сохранение информации для слоев декодера о соответствующих слоях энкодера. Таким образом, повышается разрешение выходной маски. Этой же цели служит выделение большого числа характерных признаков на каждой итерации при работе с входными данными. При этом входное изображение сегментируется не одновременно, а поэтапно, то есть для предсказания объектов используется только та часть изображения, для которой имеется полный контекст на входных данных. При этом для краевых областей информация восстанавливается путем отражения исходного изображения.

При работе с биологическими объектами в контексте машинного обучения основную сложность представляет получение необходимого для качественного обучения объема данных. Для увеличения тренировочной выборки применяется процедура аугментации, которая заключается в искажении исходных изображений путем отражений относительно вертикальной и горизонтальной осей, повороты, изменения яркости и контрастности и т.д. Такого рода деформации являются естественными для биологических изображений и позволяют увеличить размер выборки.

Нейронная сеть состоит из сжимающей (энкодер) и восстанавливающей (декодер) изображение частей. Энкодер представляет собой классическую CNN, включающую в себя применение двух конволюций с ядром 3×3 , каждая сопровождается применением активационной функции ReLU к результату, после чего проводится сжатие слоя применением max pooling с ядром 2×2 . На каждой итерации число признаков увеличивается вдвое. На этапе декодера применяется деконволюция с ядром 2×2 , после чего число признаков уменьшается в два раза, далее изображения слоя конкатенируется с соответствующим слоем энкодера, после этого дважды применяется конволюция с ядром 3×3 , каждый из которых сопровождается применением функции ReLU. Стоит заметить, что при каждом применении конволюции изображение несколько уменьшается, поскольку теряется информация о краевых областях. Последний слой – конволюция 1×1 для определения по имеющимся 64 признакам принадлежности каждого пикселя к определенному классу. Таким образом, в сумме имеется 23 конволюционных слоя.

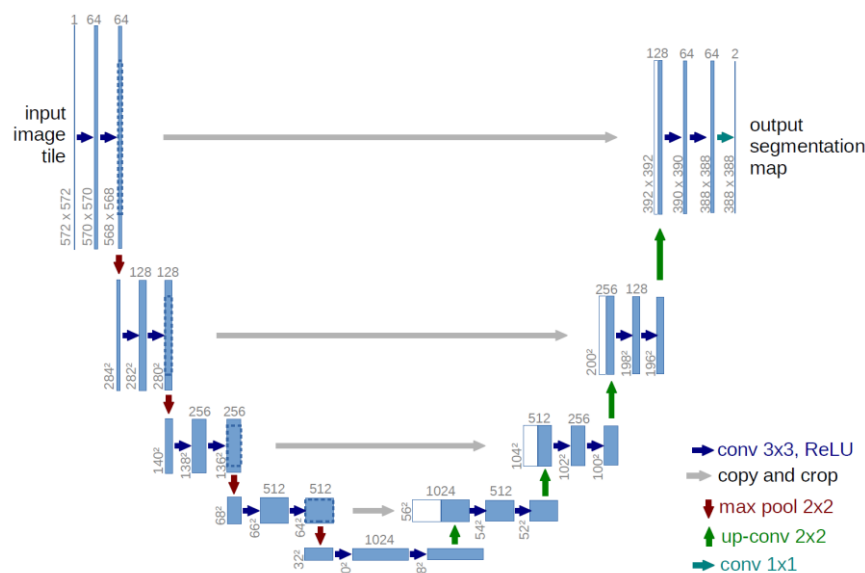


Рисунок 13 – Архитектура нейронной сети UNet. Источник: Olaf Ronneberger, Philipp Fischer, Thomas Brox “U-Net: Convolutional Networks for Biomedical Image Segmentation” \ arXiv.org – 2015

Обучение сети проводится методом стохастического градиентного спуска,⁴⁸ отличие которого от классического градиентного спуска заключается в том, что градиент вычисляется на некотором подмножестве данных, а не на всех имеющихся значениях. Функцией потерь, к которой применяется оптимизатор, является кросс-энтропия, при этом данные нормализуются с помощью функции softmax.

UNet показала себя на соревнованиях the ISBI cell tracking challenge 2014 как перспективный метод для проведения сегментации со значением IoU = 92% для изображений глиобластомы, полученных методом фазового контраста, и 77% для клеток HeLa, обследованных с помощью дифференциальной интерференционно-контрастной микроскопии, что превышает этот показатель для других алгоритмов.⁴⁷

Материалы и методы

● Когорта PRAD

В работе изучались изображения, полученные методом СуСIF, для 15 пациентов с диагнозом “аденокарцинома предстательной железы”. При этом 7 пациентов были отнесены к группе, диагноз которой не подтверждается результатами МРТ, 8 пациентов – к группе с диагнозом, подтвержденным МРТ-диагностикой. Для каждого пациента имелись изображения 14 маркеров, которые указаны в Таблице 2.

CD4	FOXP3
NaKATPase	PD1
PCK26	PD-L1
CD45	Ki67
CD31	CD8
cCasp3	CD68
S6	DAPI

Таблица 2 – Клеточные маркеры когорты PRAD

● Препроцессинг изображений

Изображения для каждого пациента когорты являлись 16-битными и имели разные площади. На начальном этапе обработки изображения разбивались на регионы размером 2048*2048 пикселей.

● Клеточная сегментация

Процедура сегментации проводилась с помощью конволюционной нейронной сети UNet,⁴⁷ после чего полученные маски корректировались алгоритмом **watershed**, реализованным с помощью библиотеки OpenCV, для дополнительного разделения объектов. Логика работы нейронной сети включала в себя три этапа:

- 1) Область считается клеткой только при наличии в ней ядра (маркер DAPI);
- 2) Замкнутая граница свечения маркера NaKATPase считается клеточной мембраной;
- 3) При отсутствии значимой экспрессии NaKATPase вблизи ядра, клеточная мембрана определялась как замкнутый контур на расстоянии ~15 пикселей от ядра, при условии отсутствия других ядер в засекаемом пространстве.

Для сегментации клеток на вход нейронной сети подавались изображения маркеров DAPI(ядра), NaKATPase(мембрана), S6(цитоплазма).

В обучающую выборку входили клеточные маски с ручной разметкой, дополненные изображениями, полученными с помощью программы CellProfiler (Carpenter Lab is based at the Broad Institute of MIT and Harvard in Cambridge, Massachusetts, USA, <https://personal.broadinstitute.org/anne/>). Более подробные сведения о данных для обучения CNN представлены в разделе “Создание обучающей выборки”.

В процессе обучения алгоритм учится предсказывать как можно больше пикселей, соответствующих клеточной маске, при этом неизбежно теряя информацию о тонких границах между клетками в местах плотных скоплений. Эта проблема не всегда может быть решена с помощью watershed, поэтому сеть была обучена предсказывать маски границ между клетками, сгенерированные автоматически путем одновременного расширения клеточных контуров в масках обучающей выборки и сохранения информации об их пересечениях. На выходе сети получали 2 класса бинарных изображений – маску клеток и маску границ между ними. Дополнительное вычитание второго изображения из первого даёт более точную сегментацию с лучшим разделением клеточных контуров. Таким образом, обучающие данные включали в себя три компонента: флуоресцентные изображения, объединенные в трехканальный массив; клеточные маски; маски границ между клетками.

Каждое изображение было разделено на квадраты размером 224*224 пикселя. В процессе обучения, для увеличения вариативности обучающей выборки, к ним применялись аугментации: случайный поворот от 0 до 270 градусов с отражением относительно вертикальной и/или горизонтальной оси. Стандартизация изображений производилась поканально путем вычитания средних значений яркости изображений маркера и последующим делением на среднеквадратичное отклонение, вычисленных на тренировочных данных.

Выбранная архитектура UNet включала в себя энкодер и декодер, состоявшие из 4 слоёв каждый. Один слой – это дважды примененная к изображению последовательность из конволюции с ядром 3*3, padding = 1, stride = 1 и функции активации ReLU. Между слоями выполнялось сжатие max pooling с ядром 2*2 на этапе энкодера и билинейная экстраполяция между слоями декодера.

Для проведения машинного обучения использовалась библиотека PyTorch на языке python. При обучении нейронной сети для решения задачи семантической сегментации данные делились на две части: обучающая выборка (85%) и тестовая (15%). Само

обучение проводилось на видеокарте NVidia Tesla P100, размер батча составлял 25 изображений, обучение длилось 100 эпох. Использовался оптимизатор Adam, параметр learning rate = 0.01, функция потерь – бинарная кросс-энтропия в сочетании с dice loss.

$$LOSS = BCE * weight + DICE * (1 - weight);$$

$$BCE = -\sum[y_i \ln p_i + (1 - y_i) \ln(1 - p_i)];$$

$$DICE = 1 - 2 \frac{\sum y_i p_i}{\sum y_i + \sum p_i},$$

где LOSS – loss-функция, BCE – бинарная кросс-энтропия, DICE – dice-loss, y_i – истинное значение, p_i – предсказанное значение пикселя, weight – вес функции.

Обученная CNN была использована для проведения сегментации регионов всех пациентов когорты. Изображения нормализовались аналогично с процедурой препроцессинга для обучающей выборки. Маски, полученные на выходе нейронной сети, были обработаны с помощью функции *watershed*.

После завершения процедуры сегментации для каждой клетки были сняты параметры, такие как ее локальные координаты на слайде и средняя интенсивность для каждого из 14 маркеров. На выходе для каждого пациента когорты был получен датафрейм с параметрами, необходимыми для дальнейшей работы с изображениями.

Использованные библиотеки: `alumentations`, `numpy`, `opencv`, `sklearn.model_selection`, `torch`, `torchvision`, `torchsummary`, локальные библиотеки BostonGene Corporation.

● Создание обучающей выборки

Обучающая выборка состояла из двух типов клеточных масок: полученных с помощью программы CellProfiler и масок ручной разметки.

Ручная разметка проводилась с помощью программы GIMP (<https://www.gimp.org/>) и графического планшета Wacom One. Разметка учитывала расположение ядер на основании изображений маркера DAPI, клеточные мембраны выделялись по свечению NaKATPase, при этом принималось во внимание наличие внутри клетки сигнала маркера S6. В среде GIMP были созданы следующие слои в указанном порядке:

1. Черно-белая маска (области клеток соответствует значение 255, фону – 0), прозрачность 8%
2. Композитное изображение (NaKATPase + S6 + DAPI), полученное с помощью программы ImageJ (<https://imagej.net/ImageJ/>)
3. NaKATPase

4. DAPI

При автоматической сегментации в программе CellProfiler использовались следующие инструменты:

→ *White Top-hat Transformation*

Фильтрация по типу “белого цилиндра” возвращает объекты, которые меньше заданного объекта (выбирается в разделе structuring element) и ярче, чем их окружение. Таким образом, можно уменьшить влияние свечения соседних объектов друг на друга.

→ *IdentifyPrimaryObjects*

Определяются границы ядер по маркеру dapi. Задача – отделить округлые области ядер от цитоплазмы и фона, для этого применяется threshold (на основе алгоритма Оцу, описанного в обзоре литературы), разделяющий область значений интенсивности на три класса: фон, промежуточный (как правило, цитоплазма) и ядро. Пренебрегая средним классом, очерчиваются границы ядра.

→ *IdentifySecondaryObjects*

Областью цитоплазмы считается расширение от ядра до ближайших клеток не более чем на заданное расстояние.

● Процедура клеточной классификации

Для определения принадлежности клеток к определенным клеточным типам была проведена предобработка данных о средней интенсивности маркеров с применением гиперболического арксинуса для корректировки диапазона данных.

Для упрощения процесса типирования клеток был применен алгоритм FlowSOM,⁴⁹ в основе которого лежит нейронная сеть с обучением без учителя – самоорганизующаяся карта Кохонена,⁵⁰ на основе результатов которой возможно построение минимального остовного дерева, вершинами которого являются выделенные кластеры. Чем дальше друг от друга находятся вершины, тем меньше между ними общего. В работе была использована реализация на языке R фреймворка BioConductor *flowCore*.

Самоорганизующаяся карта Кохонена строится в пространстве, размерность которого d равна количеству используемых маркеров (11 в данном случае), на котором на первом шаге в произвольном порядке k точек из данных соотносятся с k узлами, где k – предопределенное количество кластеров. Окрестность точки определяется как расстояние

Чебышева в двумерной сетке узлов. Самоорганизующаяся карта обучается путем многократного выбора точки из набора данных, нахождения ближайшего к ней узла и обновления всех узлов в окрестности этого узла. В процессе работы алгоритма размер окрестности и коэффициент обучения уменьшаются. В итоге каждая точка из набора данных принадлежит тому узлу, который наиболее похож на нее, что приводит к окончательной кластеризации.

Для корректной работы алгоритма и уменьшения погрешности кластеризации разработчиками FlowSOM рекомендуется количество кластеров назначать больше, чем предполагаемое количество клеточных типов. При ожидаемом выделении 9 клеточных типов, было решено выделять 20 кластеров для каждого из пациентов.

После того, как с помощью FlowSOM была определена принадлежность к определенному кластеру для каждой клетки пациента, типирование проводилось на основании диаграмм распределения интенсивностей используемых маркеров в каждом кластере. Для проверки качества типирования использовались двумерные tSNE-проекции клеток. Ожидаемые клеточные типы, с учетом исходного набора маркеров, указаны в Таблице 3.

Сигнатура	Клеточный тип
CD8+CD45+	Т-киллеры
CD4+CD45+	Т-хелперы
CD4+FoxP3+	Регуляторные Т-клетки
CD68+	Макрофаги
CD68+PD-L1+	Макрофаги PD-L1
PCK26+	Эпителиальные клетки
PCK26+Ki67+	Пролиферирующие эпителиальные клетки
PCK26+Ki67+PD-L1+	Клетки опухоли с экспрессией PD-L1
CD31+	Эндотелий сосудов

Таблица 3 – Предполагаемые клеточные типы и их сигнатуры для пациентов когорты PRAD

Использованные библиотеки: pandas, numpy, matplotlib, MulticoreTSNE, gry2.robjects, packages("BiocManager"), локальные библиотеки BostonGene Corporation.

- **Бинарные маски эндотелия, эпителия, стромы**

Для создания масок использовались методы библиотеки OpenCV python.

Основной функцией является созданная специально для данного проекта функция *make_thresholded_image*:

```
def make_thresholded_image(img: np.ndarray, thresh_lim: int, num_blur: int) ->
np.ndarray:
    """
    :param img: input image, numpy.ndarray, 8-bit
    :param thresh_lim: parameter for cv2.threshold(img, thresh_lim, 255,
cv2.THRESH_BINARY)
    :param num_blur: the number of circles of blurring
    :return: thresholded image
    """
    if not np.issubdtype(img.dtype, np.uint8):
        raise TypeError("Only 8-bit images are available, but you use " + str(img.dtype))

    img_gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY,
cv2.IMREAD_UNCHANGED)

    img_blurred = img_gray
    for x in range(num_blur):
        img_blurred = cv2.blur(img_gray, ksize=(3, 3))
    img_gray = None

    _, thresholded = cv2.threshold(img_blurred, thresh_lim, 255, cv2.THRESH_BINARY)
    img_blurred = None

    return thresholded
```

Маски эпителия формировались на основании изображений маркера PCK26, изображение которого передавалось в функцию *make_thresholded_image*, для масок эндотелия – CD31. После чего производилась полуавтоматическая коррекция полученных масок.

Для создания масок стромы учитывались маркеры PCK26 и S6. На первом этапе получали маску эпителия, инвертировали ее, таким образом, получая неэпителиальный компартмент каждого слайда. Затем выполнялась коррекция на основе изображения маркера S6 для выявления областей, не содержащих клетки (в том числе просветов ацинусов). Полученные маски подвергались полуавтоматической корректировке.

Основные использованные библиотеки: `opencv`, `numpy`, локальные библиотеки BostonGene Corporation.

- **Метрики клеточных составов**

Для оценки гетерогенности образцов по клеточному составу выполнялись графическая визуализация процентного соотношения клеток, плотности их распределения в разных областях опухоли, а также построение `box-plot`'ов для этих плотностей. Расчет статистической значимости проводился с применением U-критерия Манна–Уитни с двумя концами.

Изображения процентного отношения клеточных составов строились с применением методов библиотеки `pandas` и локальной библиотеки BostonGene Corporation в различных областях опухоли: эпителий с видимыми отличиями от нормы и без таковых, стромы, соответствующая указанным типам эпителия.

Плотность клеток определялась с использованием методов библиотек `pandas`, `numpy`, `opencv`, локальной библиотеки BostonGene Corporation. Данные представляли собой отношение числа клеток в рассматриваемой области к ее площади.

- **Сравнение стромальных распределений**

Для определения различий между стромальным распределением в незлокачественных и опухолевых зонах проводилась квадратная тесселяция каждой зоны, то есть разделение на небольшие квадраты одинакового размера, с последующим вычислением процентного содержания стромальной маски в каждом квадрате. Размер квадратов был выбран для поддержания максимальной вариабельности значений, чтобы отличить участки с крупными и плотными стромальными элементами от разреженных участков, позволяя сформировать равномерное распределение на слайде. Далее гистограммы двух распределений для одного пациента сравнивались в разных зонах (нормы/опухоли) на основании величины расстояния Вассерштейна между ними. Для расчетов использовался написанный специально для данной задачи код, а также методы библиотек Python: `OpenCV`, `SciPy` и ряд других.

- **Определение групп клеточных соседств**

Для расчета клеточных соседей использовался алгоритм `sKDTree` библиотеки `scipy`. Ближайшие соседи считались как количество клеток каждого типа на расстоянии не

более 80 пикселей, дальние соседи – на расстоянии до 300 пикселей. При этом средний радиус клетки составлял 25 пикселей.

Полученная таблица далее использовалась для определения групп клеточных соседств (Communities). Для этого данные о количестве клеток нормировались с помощью функции `scale` библиотеки `sklearn.preprocessing`, после чего полученная таблица передавалась в алгоритм кластеризации `phenograph`, выделяющий кластеры методом полного перебора. Выделенные кластеры типировались с помощью веб-приложения Phantanus (<https://artyomovlab.wustl.edu/phantanus/>).

Результаты

Первым этапом обработки изображений, полученных методом мультиплексной иммунофлуоресценции, является проведение сегментации с целью выделения контуров отдельных клеток. Возможно решение данной задачи с применением специальных программ, таких как CellProfiler, которая является одной из рекомендованных для проведения автоматической сегментации биологических объектов. Но в работе алгоритмов данной программы был обнаружен ряд недостатков. Во-первых, сегментация проводится на основании только изображений ядер, клеточные границы определяются как расширение от ядра на заданное расстояние (15 пикселей в данной работе), то есть не учитываются маркеры клеточных мембран. Во-вторых, из-за неоднородности флуоресцентного сигнала возможна пересегментация ядер, соответственно, мы получаем неверное представление о количестве клеток, что скажется на последующих этапах работы. Сказанное выше объясняет необходимость искать другой подход к автоматической сегментации. Решение было найдено в применении конволюционной нейронной сети UNet, созданной с использованием фреймворка PyTorch, и создании масок ручной разметки, которые легли в основу обучающей выборки нейронной сети. Для увеличения разнообразия данных применялась процедура аугментации, для чего использовалась библиотека albumentations (Python). Подробно архитектура нейронной сети, а также процедура создания обучающей выборки описаны в разделе «Материалы и методы». Примеры флуоресцентных изображений, подаваемых на вход UNet, представлены на Рисунке 14.

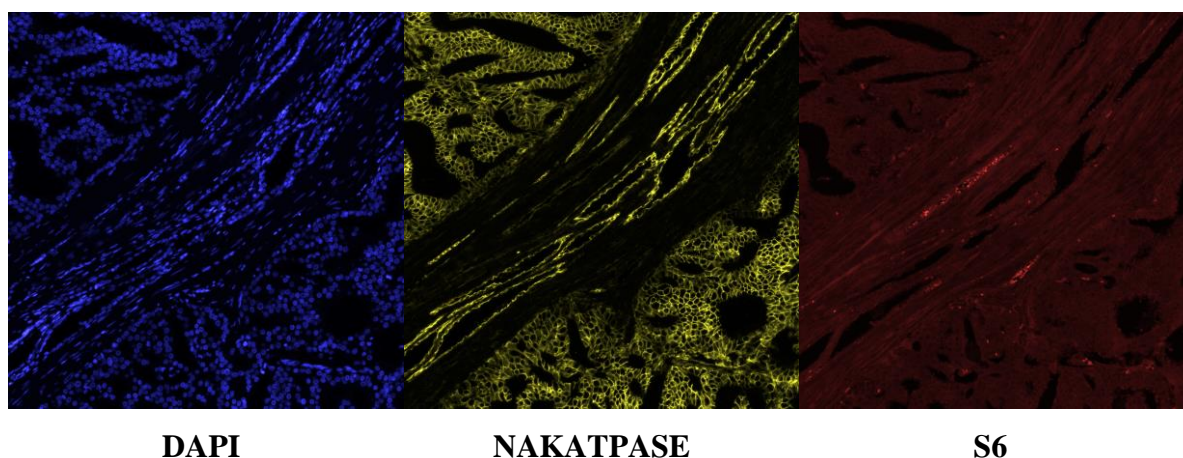


Рисунок 14 – Примеры флуоресцентных изображений одного региона маркеров, использовавшихся при обучении нейронной сети для решения задачи клеточной сегментации. DAPI – маркер ядер, NAKATPASE – маркер клеточных мембран, S6 – маркер цитоплазмы.

Обученная модель применялась для всех изображений когорты PRAD, состоящей из 15 пациентов, для каждого из которых имелись флуоресцентные изображения 14 маркеров на срезах FFPE-препаратов предстательной железы. У пациентов была диагностирована аденокарцинома предстательной железы, после чего была проведена радикальная простатэктомия. Пример результата применения обученной нейронной сети представлен на Рисунке 15.

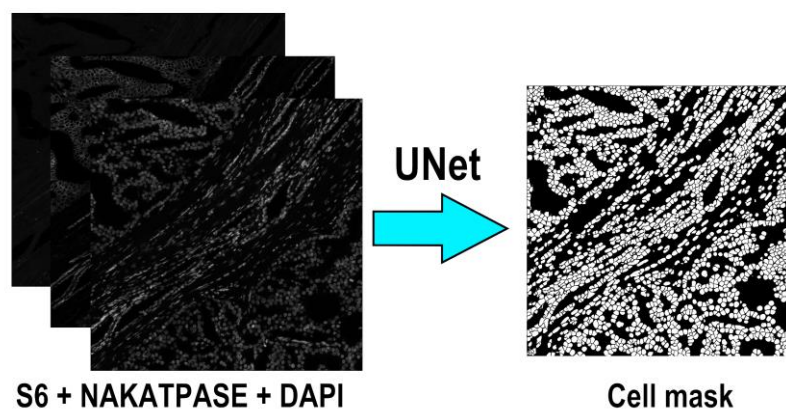


Рисунок 15 – Схема решения задачи сегментации с помощью UNet. На входе – флуоресцентные изображения маркеров, на выходе – Cell mask: бинарное изображение.

После выделения клеточных контуров для всех образцов, были написаны функции для снятия метрик, таких как: координаты клетки в регионе (изображения маркеров были предварительно разделены на квадраты 2048*2048 пикселей, каждый из которых имел аннотацию гистопатолога о наличии или отсутствии там злокачественных изменений ткани предстательной железы) и средней интенсивности свечения каждого из 14 маркеров внутри клеточного контура.

Предварительный анализ флуоресцентных изображений выявил неспецифичное окрашивание для некоторых маркеров, не позволяющее аналитически определить области значимой экспрессии. В дальнейшем анализе не использовались следующие маркеры: cCasp3, PD-1, PD-L1.

Следующим этапом являлось клеточное типирование. Для этого первым шагом провели кластеризацию с помощью FlowSOM, который активно применяется в анализе результатов проточной цитометрии, выходные данные которой похожи на таблицы метрик, получаемых после сегментации MxIF изображений. Схема работы данного алгоритма представлена на Рисунке 16. После чего каждый кластер был отнесен к определенному клеточному типу на основании гистограмм распределения интенсивности

свечения маркеров в клетках, принадлежащих данному кластеру. Данные об интенсивностях свечения маркеров представлены на Рисунке 18. В результате было выделено 9 клеточных типов, указанных в Таблице 4.

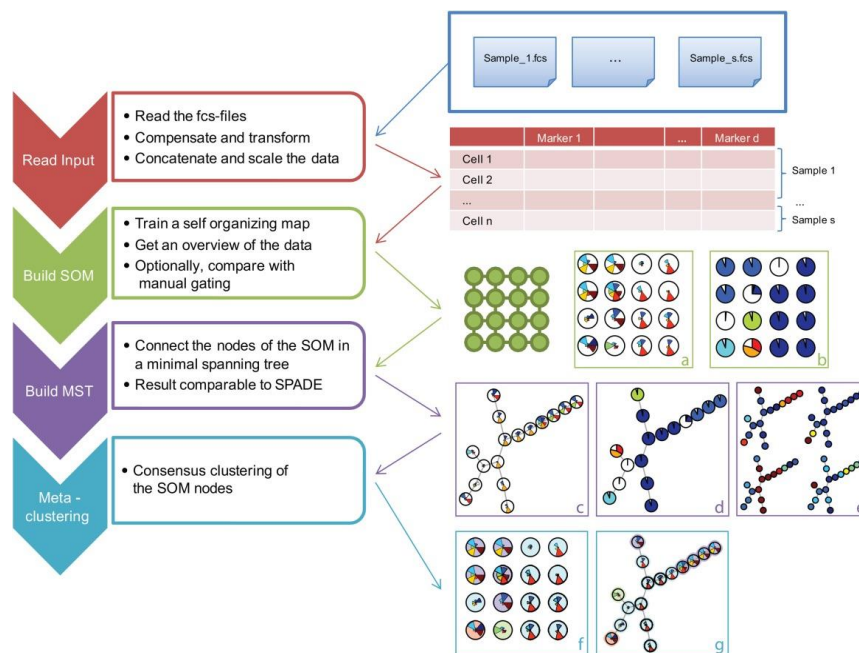


Рисунок 16 – Архитектура нейронной сети UNet. Источник: Sofie Van Gassen et al. “ FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data” || Cytometry – 2015

Сигнатура	Клеточный тип
CD8+CD45+	Т-киллеры (CD8 T-cells)
CD4+CD45+	Т-хелперы (CD4 T-cells)
CD4+FoxP3+	Регуляторные Т-клетки (T-reg)
CD68+	Макрофаги (Macrophages)
CD31+	Эндотелий сосудов (Endothelium)
PCK26+ (и данные гистопатолога)	Опухолевый эпителий (Malignant epithelium)
PCK26+ (и данные гистопатолога)	Эпителиальные клетки (Non-malignant epithelium)
PCK26+Ki67+	Пролиферирующие эпителиальные клетки (Tumor KI67+)
Недостаточно имеющих маркеров	Unclassified

Таблица 4 – Клеточные типы, определенные для пациентов когорты PRAD, и их сигнатуры

Эпителий (клетки с высокой интенсивностью маркера PCK26) был разделен на здоровый эпителий и опухолевый на основании аннотации гистопатолога с использованием шкалы Глисона. Клеткам, которые не удалось отнести ни к одному клеточному типу ввиду ограниченности имеющегося набора маркеров, был присвоен тип Unclassified. Пример визуализации клеток разных типов на гистологическом слайде одного из пациентов изображен на Рисунке 19.

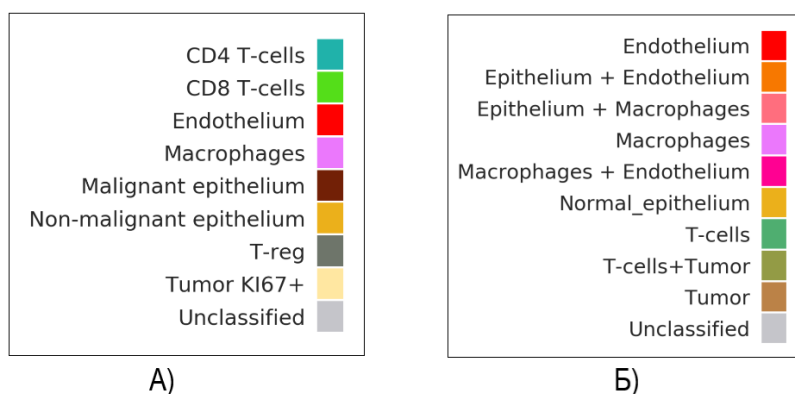


Рисунок 17 – Цветовые палитры А) для клеточных типов Б) для выделенных групп клеточных соседств

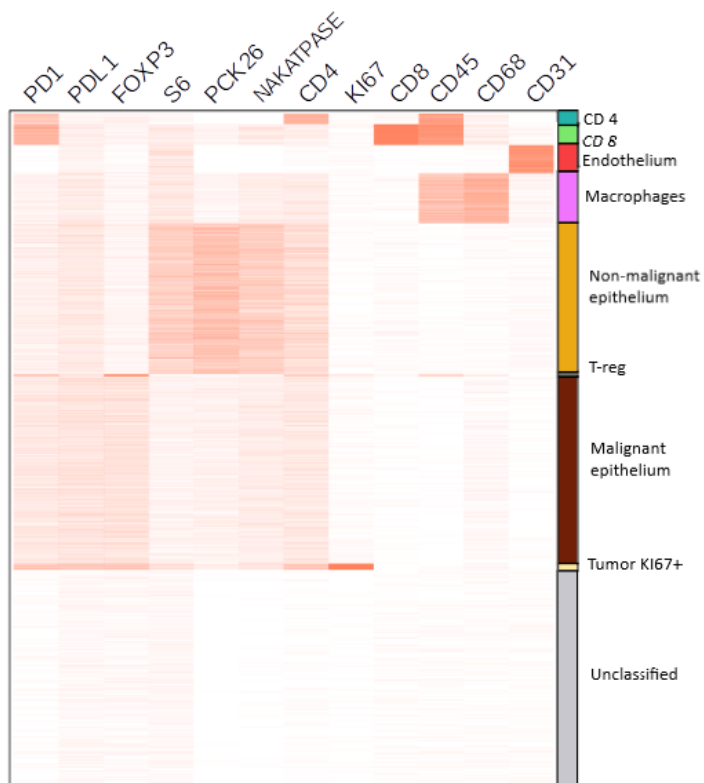


Рисунок 18 – Интенсивности свечения маркеров, которые использовались при выделении клеточных типов. Цветовая палитра соответствует указанной на Рисунке 17

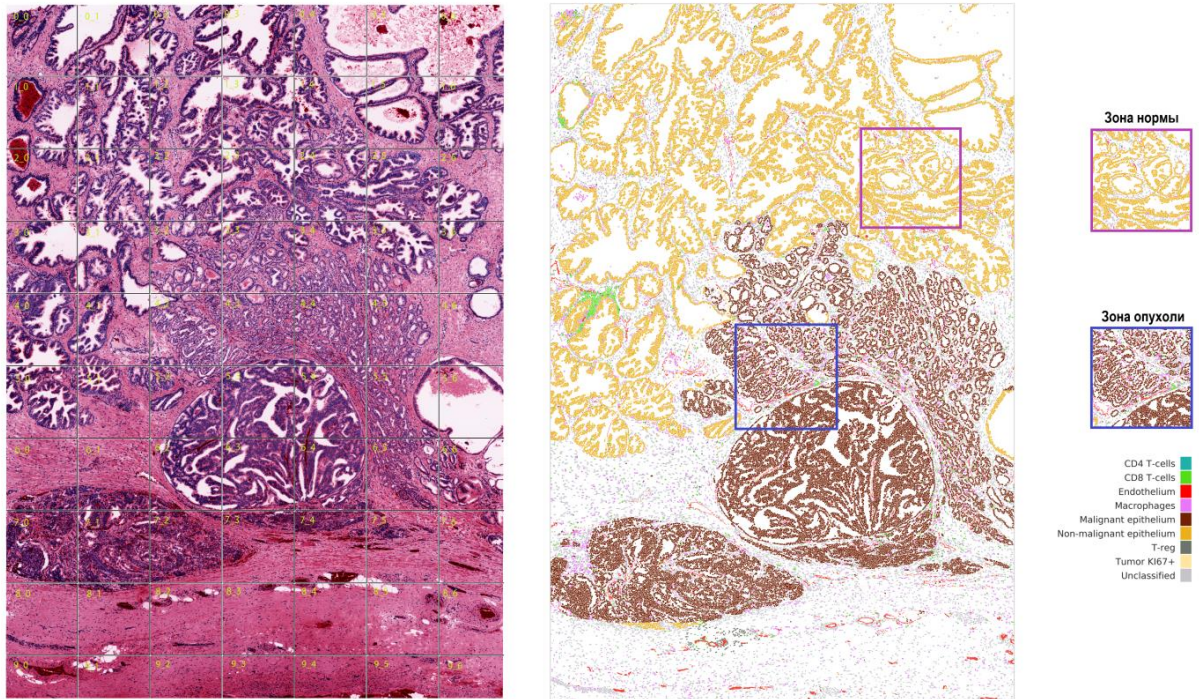


Рисунок 19 – Н&Е изображение и визуализация клеточных типов на всем слайде и в выбранных регионах на примере одного пациента

Для дальнейшего анализа полученных данных было принято решение создать маски эпителия и стромального компартмента. Для этих целей были реализованы функции, преобразующие флуоресцентные изображения PCK26 и PCK26+S6 в бинарные изображения, представляющие собой маски эпителиальной ткани и стромы соответственно. После чего полученные маски были алгоритмически разделены на области нормы и опухоли на основании аннотации гистолога. Созданные маски для одного из пациентов когорты представлены на Рисунках 20-21. На Рисунке 22 изображена композитная маска, отражающая стромальный и эпителиальный компартменты опухоли одного из пациентов.

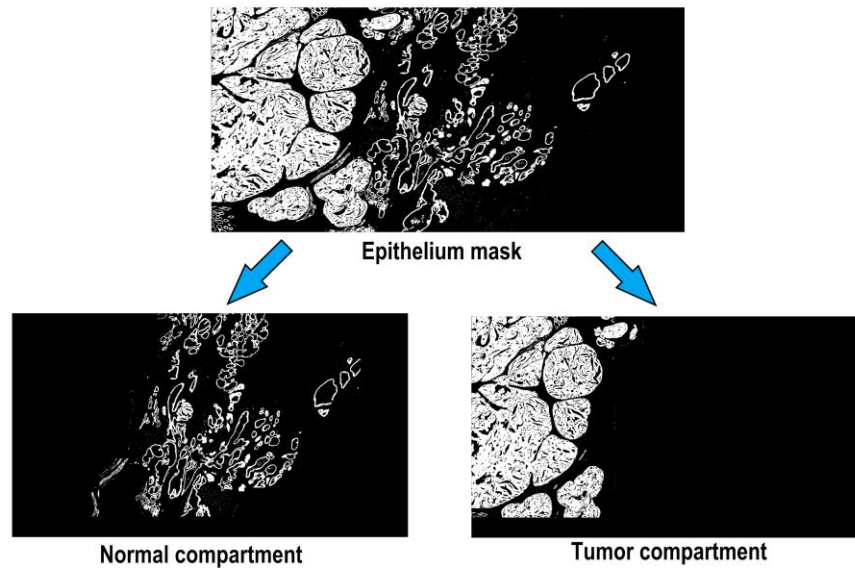


Рисунок 20 – Пример созданной эпителиальной маски и ее разделение на нормальный эпителий (Normal compartment) и опухолевый эпителий (Tumor compartment)

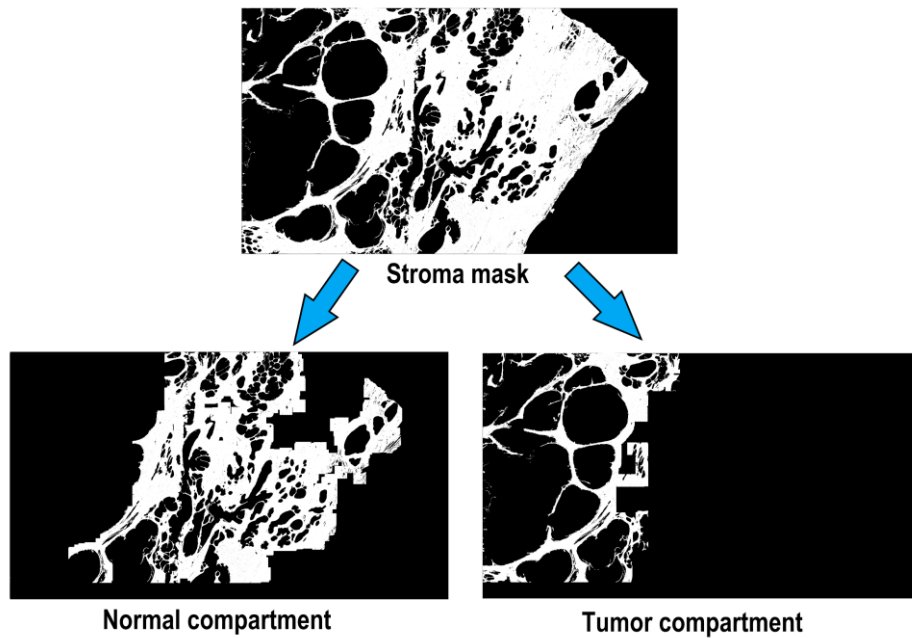


Рисунок 21 – Пример созданной маски стромы и ее разделение на строму в зоне нормального эпителия (Normal compartment) и опухолевого эпителия (Tumor compartment)

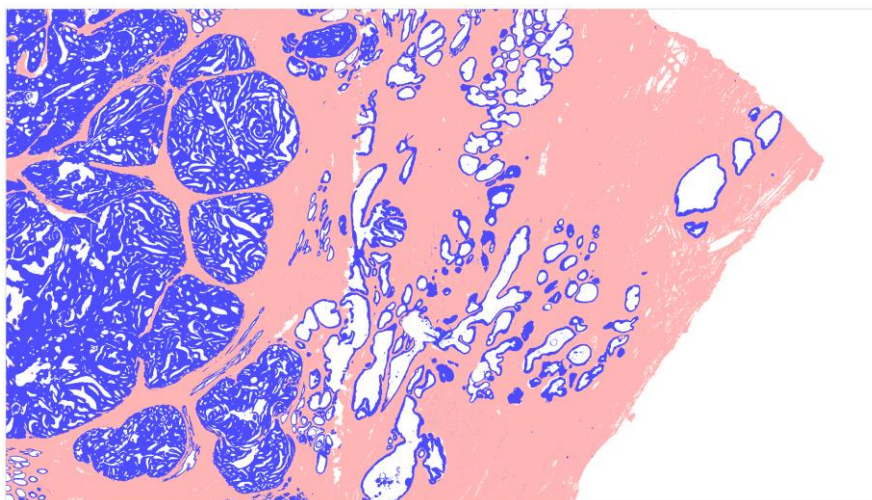


Рисунок 22 – Пример композитного изображения масок компартов опухоли на примере одного из пациентов. Синим – область эпителия, розовым – область стромы.

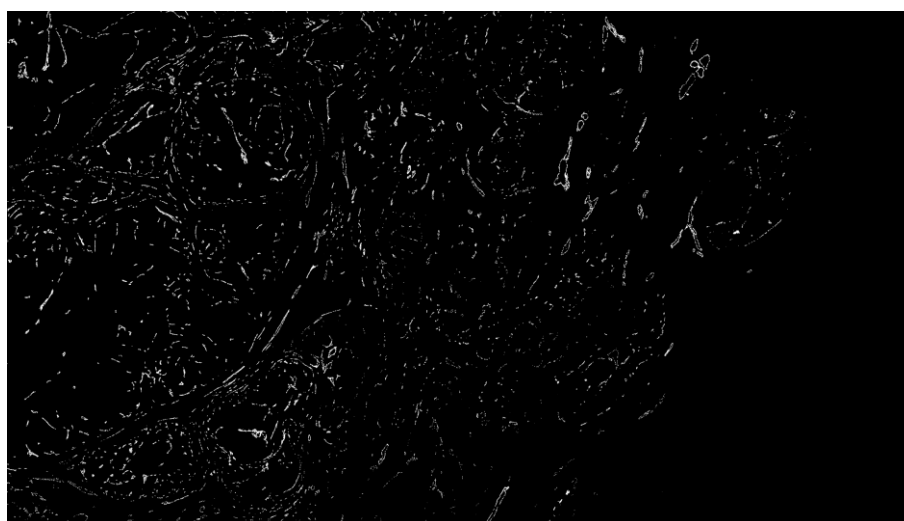


Рисунок 23 – Пример маски эндотелия сосудов для одного пациента

Также для всех пациентов были созданы маски эндотелия сосудов (Рисунок 23) на основе изображений свечения маркера CD31. При анализе сосуды изображались с использованием полученных масок, например, на Рисунке 19.

На основании проведенной классификации были построены метрики клеточного состава для каждого пациента. Для этого использовались методы библиотек `pandas`, `numpy` и локальных библиотек BostonGene Corporation, подробное описание создания метрик содержится в разделе «Материалы и методы».

Были построены диаграммы как процентного отношения содержания клеток всех типов, так и плотностей распределения клеток в областях злокачественного и здорового эпителия, а также в стромальных компартаментах соответствующих областей. Результаты анализа представлены на Рисунках 24–26.

Зона нормального эпителия

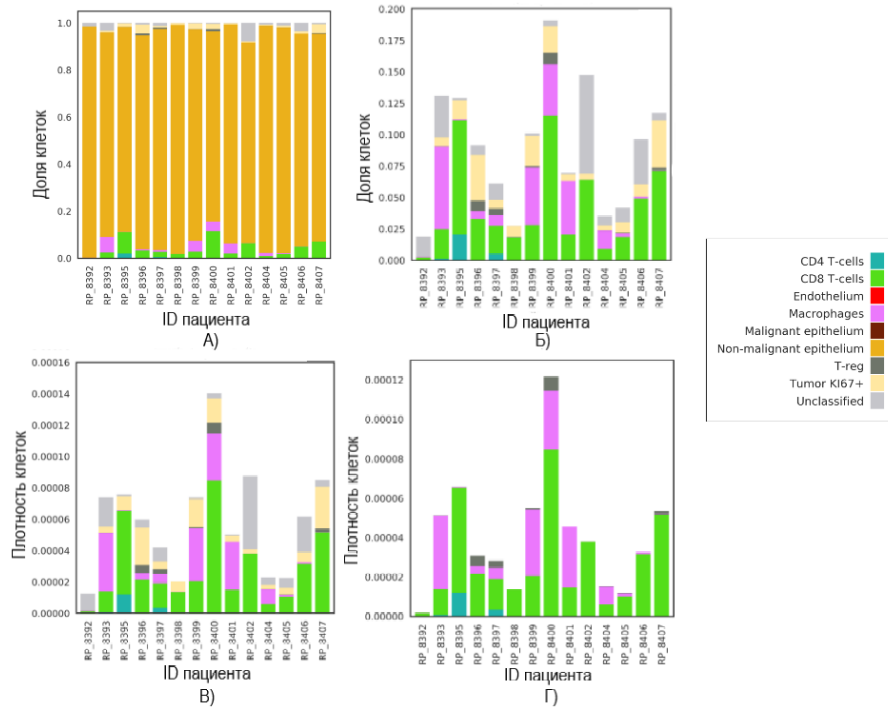


Рисунок 24 – Диаграммы клеточного состава когорты PRAD в области нормального эпителия. А) Доля клеток разных типов от общего числа клеток пациента в данном регионе. Б) Доля клеток без учета эпителиальных относительно всех клеток области. В) Плотность клеток как число клеток по отношению к площади области. Г) Плотность иммунных клеток как отношение их числа к площади области. Цветовая палитра соответствует указанной на Рисунке 17

Зона опухолевого эпителия

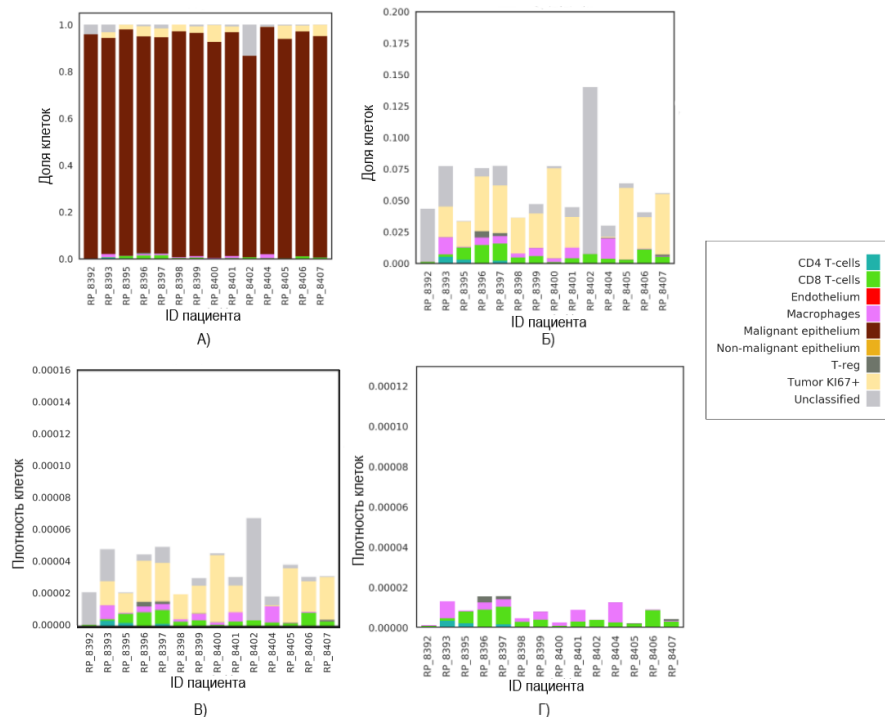


Рисунок 25 – Диаграммы клеточного состава когорты PRAD в области опухолевого эпителия. А) Доля клеток разных типов от общего числа клеток пациента в данном регионе. Б) Доля клеток без учета эпителиальных относительно всех клеток области. В) Плотность клеток как число клеток по отношению к площади области. Г) Плотность иммунных клеток как отношение их числа к площади области. Цветовая палитра соответствует указанной на Рисунке 17

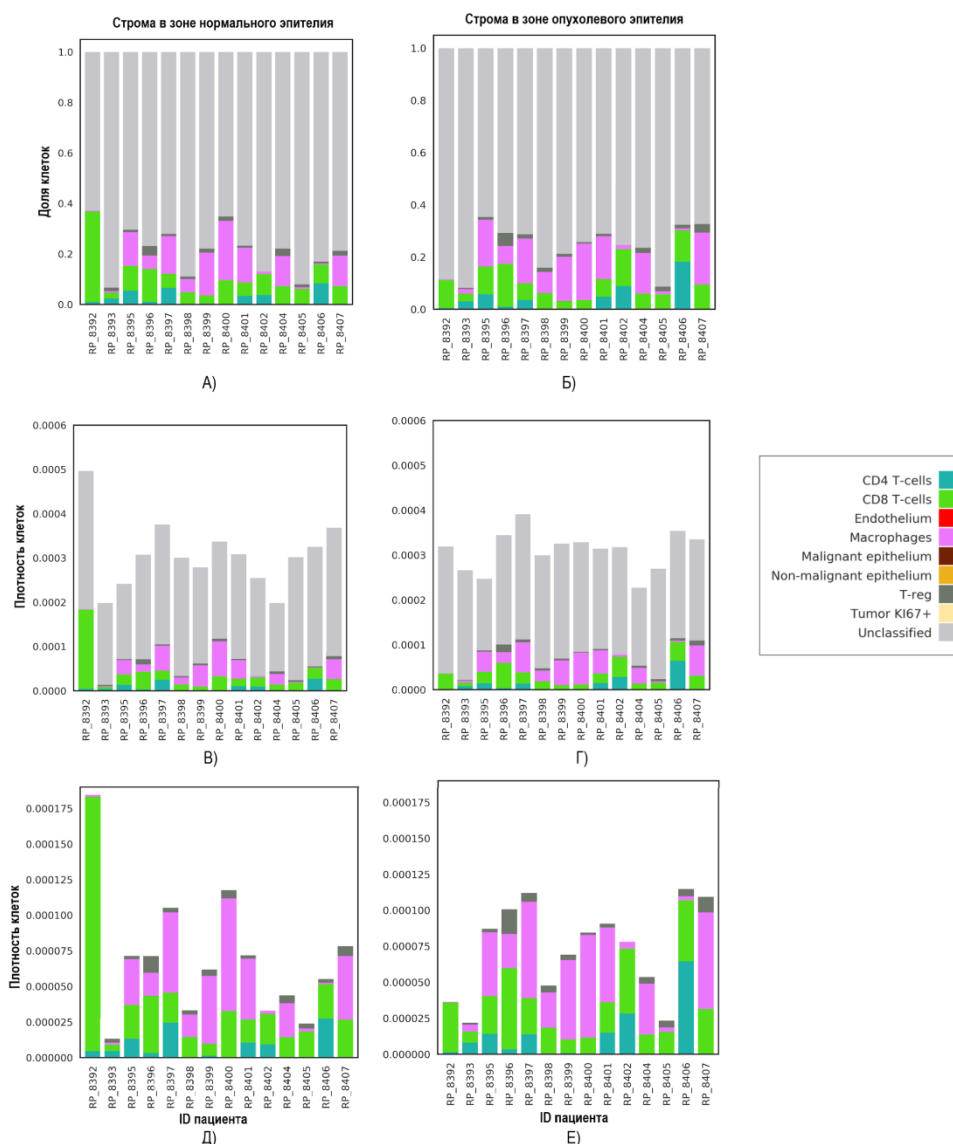


Рисунок 26 – Диаграммы клеточного состава когорты PRAD в стромальной области. А, Б) Доля клеток разных типов от общего числа клеток пациента в данном регионе. В, Г) Плотность клеток как число клеток по отношению к площади области. Д, Е) Плотность иммунных клеток как отношение их числа к площади области. Цветовая палитра соответствует указанной на Рисунке 17

Пациенты когорты PRAD были разделены на две группы, на основании заключения врача по результатам МРТ-диагностики. Предыдущие исследования показали, что mrMRI выявляет не все области образования опухоли, таким образом, невозможно гарантировать отсутствие заболевания на основании данного исследования. Одной из целей данной работы был поиск возможных метрик, позволяющих предположить признаки, по которым пациенты одной группы отличаются от пациентов другой.

Для получения количественных результатов сравнения клеточного состава пациентов двух групп были построены box-plot'ы для плотности распределения клеток разных типов, а также рассчитано p-value. Результаты данного анализа представлены ниже на Рисунках 27–30.

В зоне опухолевого эпителия

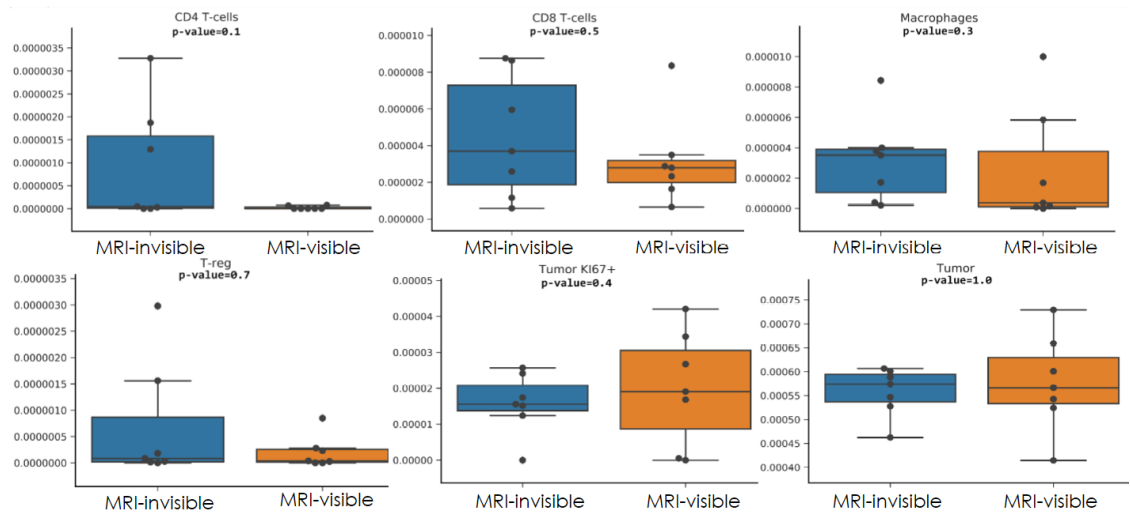


Рисунок 27 – Вох-plot’ы по клеточной плотности в области опухолевого эпителия для пациентов когорты PRAD, разделенных на две группы: MRI-invisible – опухоль не видна на МРТ, MRI-visible – видна

В зоне нормального эпителия

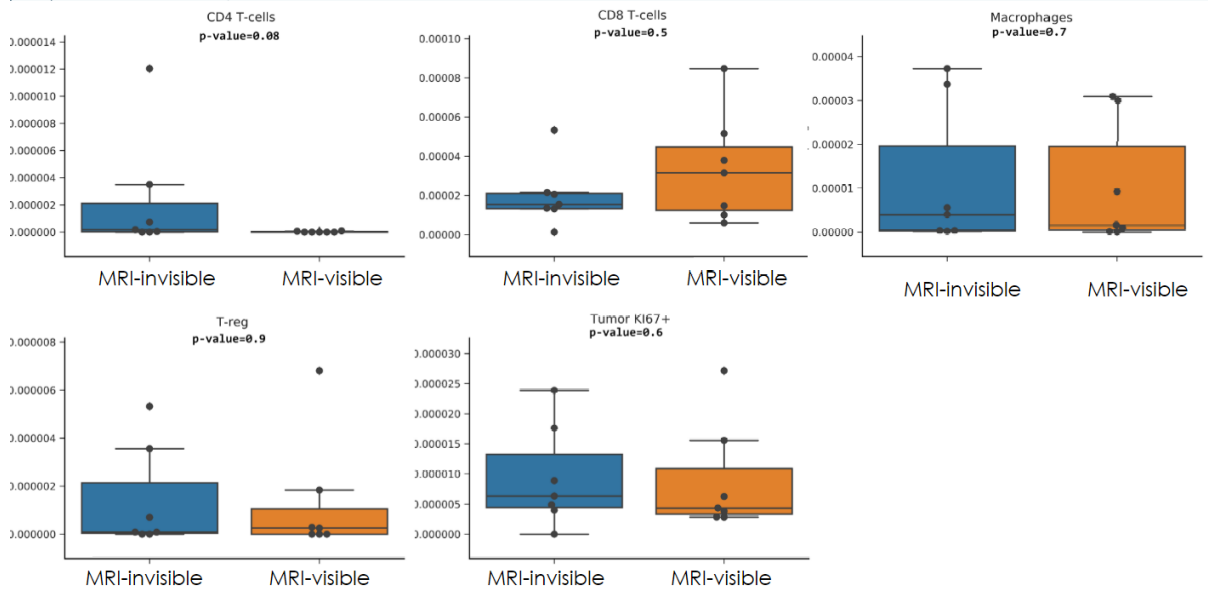


Рисунок 28 – Вох-plot’ы по клеточной плотности в области нормального эпителия для пациентов когорты PRAD, разделенных на две группы: MRI-invisible – опухоль не видна на МРТ, MRI-visible – видна

Строма в зоне опухолевого эпителия

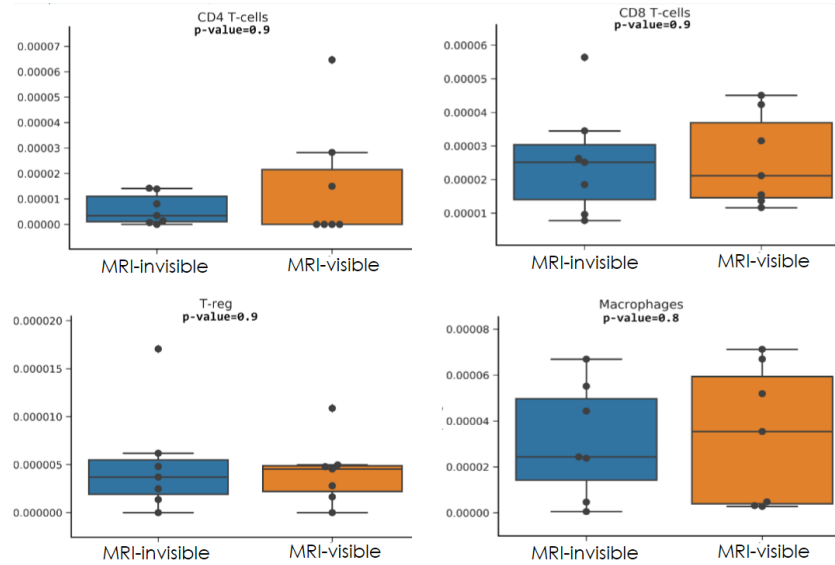


Рисунок 29 – Box-plot’ы по клеточной плотности в стромальном компартменте в области опухолевого эпителия для пациентов когорты PRAD, разделенных на две группы: MRI-invisible – опухоль не видна на МРТ, MRI-visible – видна

Строма в зоне нормального эпителия

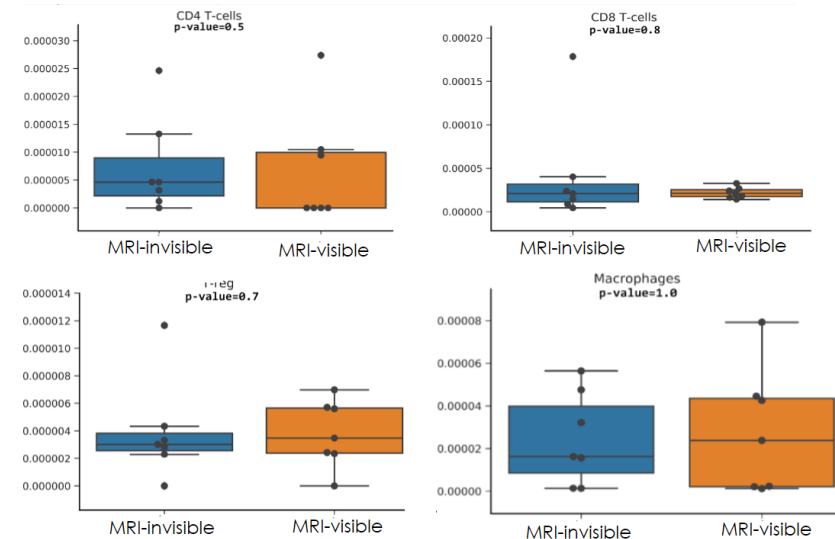


Рисунок 30 – Box-plot’ы по клеточной плотности в стромальном компартменте в области нормального эпителия для пациентов когорты PRAD, разделенных на две группы: MRI-invisible – опухоль не видна на МРТ, MRI-visible – видна

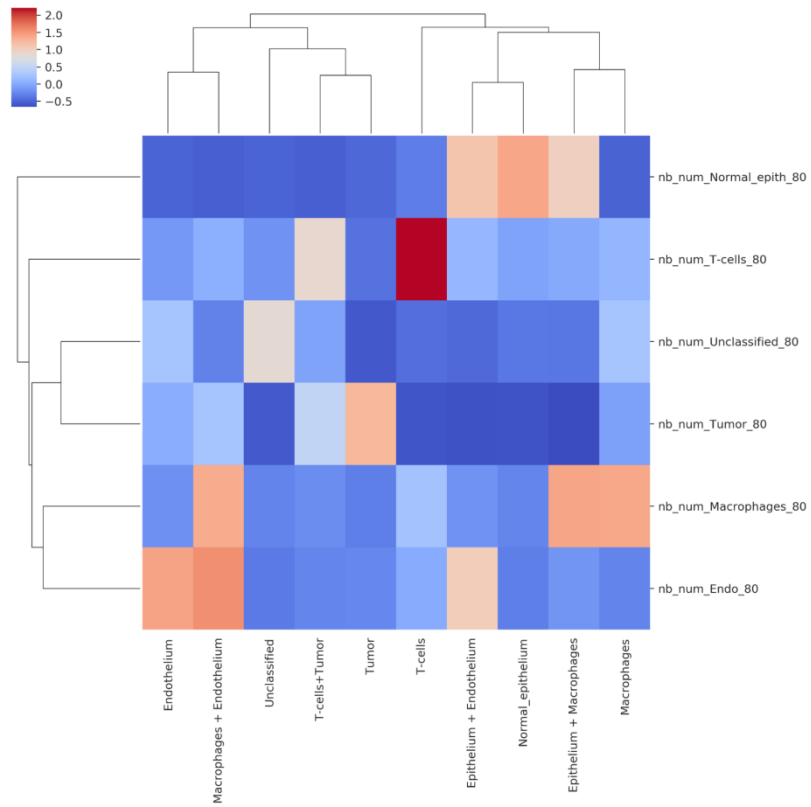
Дополнительно был проведен анализ клеточных соседств. Для этого для каждой выделенной клетки были рассчитаны ее ближние (на расстоянии не более 80 пикселей) и дальние (на расстоянии менее 300 пикселей) соседи каждого типа, после чего клетки были отнесены к группе клеточных соседств на основании наиболее представленного в их окружении клеточного типа. Результаты анализа представлены на Рисунках 31–33. Для удобства типирования Т-киллеры, Т-хелперы и регуляторные Т-клетки были отнесены к

клеточному типу T-cells; клетки, относящиеся к клеточным типам Tumor и Tumor KI67+, – к типу Tumor.

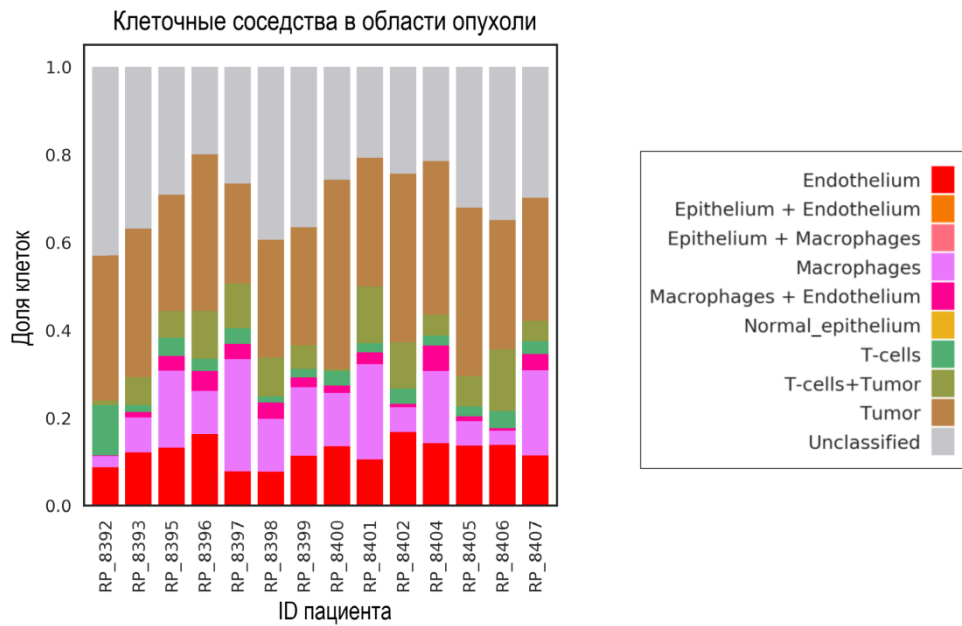
В результате было выделено 10 различных групп клеточных соседств, которые представлены в Таблице 5.

Группа клеточного соседства	Типы клеток, наиболее представленные в радиусе ближайшего соседства
Endothelium	Эндотелий сосудов (Endo)
Epithelium + Endothelium	Эпителиальные клетки без видимых изменений морфологии относительно нормы (Normal_epith) и эндотелий сосудов
Epithelium + Macrophages	Эпителиальные клетки без видимых изменений морфологии относительно нормы + макрофаги
Macrophages	Макрофаги
Macrophages + Endothelium	Макрофаги и эндотелий сосудов
Normal_epithelium	Эпителиальные клетки без видимых изменений морфологии относительно нормы
T-cells	T-клетки (может содержать CD4, CD8, регуляторные T-клетки)
T-cells + Tumor	T-клетки и опухолевые эпителиальные клетки
Tumor	Опухолевые эпителиальные клетки
Unclassified	Клетки с неопределенным клеточным типом

Таблица 5 – Описание выделенных групп клеточных соседств



A)



Б)

Рисунок 31 – Выделение групп клеточных соседств А) heatmap, где интенсивность цвета отражает среднее значение стандартизированного количества ближних соседей каждого типа для каждой группы клеточных соседств Б) доля клеток каждой группы в зоне опухоли как отношение числа клеток каждой группы к их общему числу

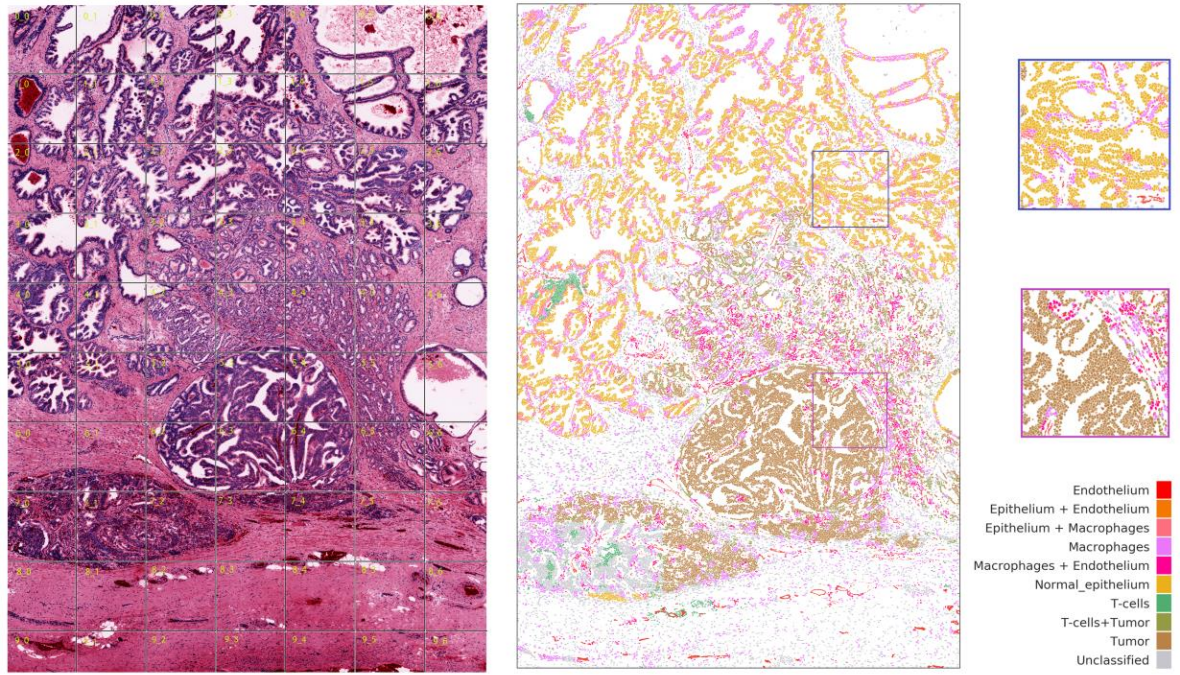


Рисунок 32 – Н&Е изображение и визуализация групп клеточных соседств на всем слайде и в выбранных регионах на примере одного пациента

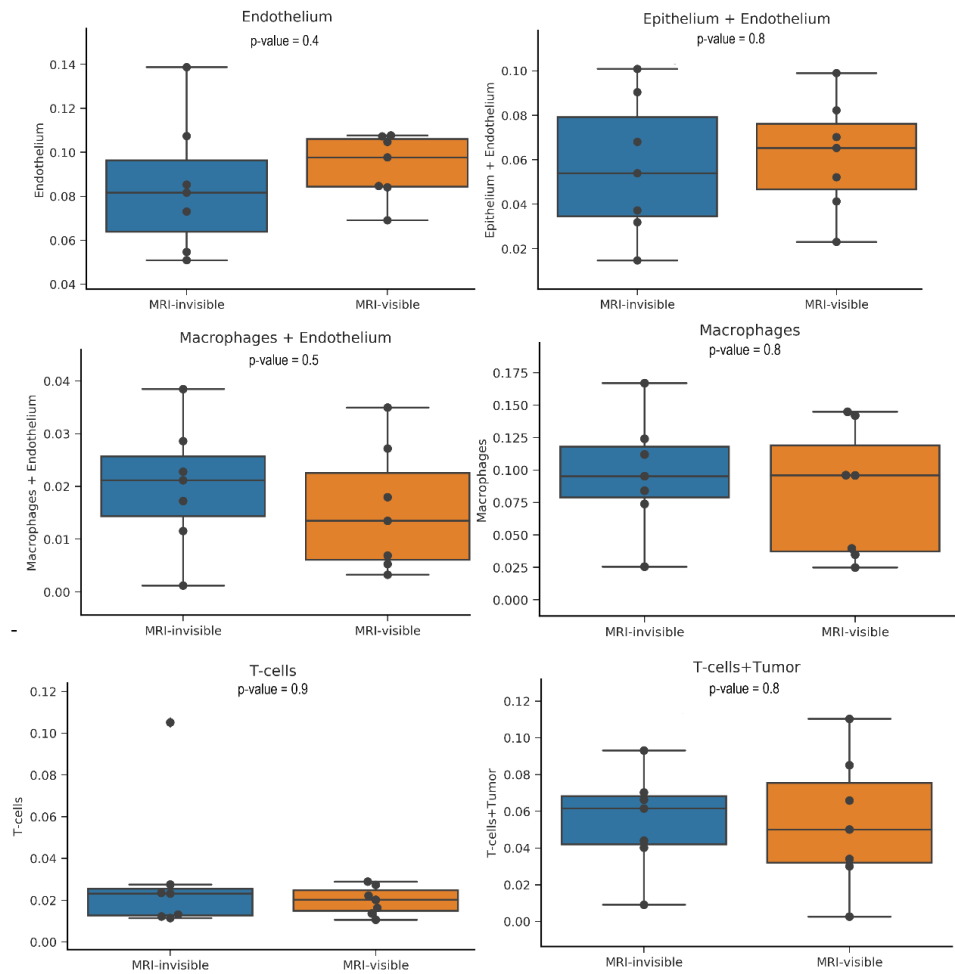


Рисунок 33 – Box-plot'ы для доли клеток, относящихся к разным группам клеточных соседств, для пациентов MRI-invisible – опухоль не видна на МРТ и MRI-visible – видна

На основании имеющихся для когорты N&E изображений было замечено, что пациенты из группы MRI-visible отличаются более ярко выраженным изменением архитектуры стромального компартмента в зоне малигнизации по отношению к не видоизмененным зонам вблизи опухоли. Для получения количественной оценки данного параметра была проведена квадратная тесселяция маски стромального компартмента в зонах опухоли и близлежащих, построены гистограммы распределения плотности стромы в данных областях и рассчитано расстояние Вассерштайна для каждого пациента. После чего на основе этих данных были построены box-plot'ы и рассчитано p-value. Результат анализа показан на Рисунке 34 (А).

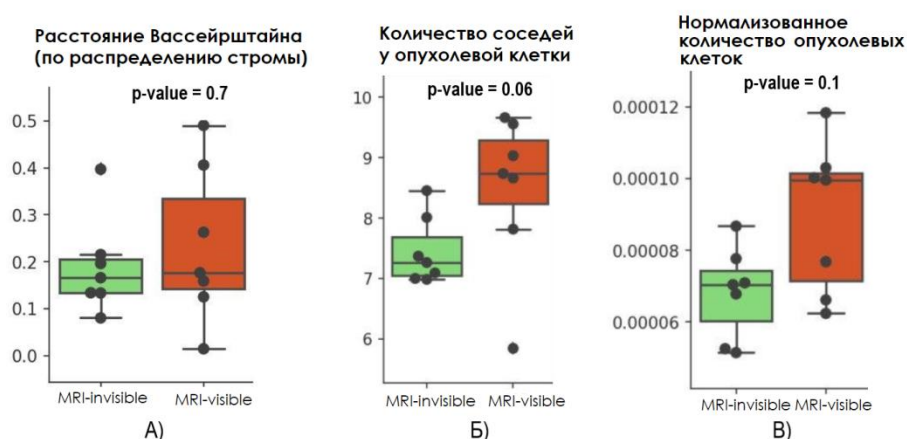


Рисунок 34 – Box-plot'ы для пациентов когорты PRAD, разделенных на две группы: MRI-invisible – опухоль не детектируется на МРТ, MRI-visible – детектируется. А) Расстояние Вассерштайна для распределений стромы в зонах опухолевого и неопухолевого эпителия Б) количество опухолевых клеток в соседстве у опухолевых клеток в ацинусах В) плотность опухолевых клеток как их количество, деленное на площадь опухолевой зоны

Другой предположительной метрикой для определения различий между двумя группами пациентов является увеличение количества клеток в злокачественном эпителии. Для получения данной метрики были созданы функции, позволяющие рассчитать клеточное соседство, то есть количество клеток каждого типа в заданном радиусе выбранной клетки для всех клеток образца. Далее были выделены опухолевые клетки и оценено количество соседствующих с ними клеток этого же типа, другими словами, было рассчитано количество соседей у эпителиальных клеток в ацинусах, подверженных злокачественным изменениям. Таким образом, были получены метрики, показывающие нормализованное количество опухолевых клеток для двух групп и различия в количестве соседей у опухолевой клетки для них, что отражено на Рисунке 34 (Б, В).

Обсуждение

Проведенный анализ флуоресцентных изображений показал, что использование нейронной сети, обученной на данных с включением масок, полученных путем проведения ручной разметки слайдов гистологических препаратов, улучшает качество сегментации по сравнению с применением программы CellProfiler. Оценка проводилась при консультации с гистопатологом и путём сравнения данных, полученных после обработки нейронной сетью и CellProfiler одних и тех же слайдов.

Анализ клеточного состава показал, что образцы опухоли разных пациентов, несмотря на одинаковый диагноз – аденокарцинома предстательной железы, – показывают гетерогенность как между препаратами разных пациентов, так и между разными регионами одного пациента. Что подтверждает важность комплексного исследования тканей злокачественных образований, чтобы учесть индивидуальные особенности опухоли конкретного пациента.

При количественной оценке различий в клеточных составах пациентов, принадлежащих к разным группам на основании видимости опухоли на их МРТ-изображениях, не было получено статистически значимых отличий ни для одного типа клеток.

Помимо описанного выше анализа, была исследована пространственная архитектура в области активной малигнизации и проведено сравнение с зонами без видимых злокачественных изменений. На основе метрики Вассерштайна было показано, что для когорты PRAD наблюдаются различия между распределениями стромы в указанных областях, но статистически значимых отличий для двух групп MRI обнаружено не было. А также для пациентов рассматриваемой когорты было показано, что у образцов, относящихся к группе ложно-отрицательных по МРТ, нормализованное количество злокачественных клеток и среднее количество опухолевых соседей на одну опухолевую клетку эпителия ниже, чем соответствующие параметры для образцов, новообразования которых видны при МРТ-исследовании. Исходя из чего можно сделать вывод, что для рассмотренных 15 пациентов предполагается большая схожесть в строении эпителия и стромы для опухолевых и здоровых регионов в случае ложно-положительного результата mpMRI. Однако, ввиду малого размера выборки, невозможно гарантировать сохранения данной тенденции для большего количества образцов.

Выводы

В данной работе был проведен анализ флуоресцентных изображений, полученных методом циклической иммуофлуоресценции для изучения возможностей исследования тканей без нарушения их архитектуры:

1. Обучена нейронная сеть для выделения клеточных контуров на изображениях;
2. Выделены 9 клеточных типов на основании интенсивности свечения имеющихся маркеров;
3. Проанализировано взаимное расположение клеток на слайдах гистологических препаратов аденокарциномы предстательной железы;
4. Проведено сравнение плотностей распределения клеток в разных регионах опухоли для двух групп пациентов, сформированных на основании результатов МРТ-диагностики. Статистически значимых отличий получено не было;
5. Проведен анализ групп клеточных соседств, выделено 10 групп. Статистически значимых отличий по пропорциям этих групп для пациентов с определяемым и не определяемым по МРТ диагнозом не обнаружено;
6. Проведено вычисление различий между распределениями стромы в зонах опухоли и зонах без видимых злокачественных изменений с помощью метрики Вассерштайна для каждого пациента когорты. Для двух исследуемых групп пациентов получено $p\text{-value} = 0.7$ при сравнении по этому параметру;
7. Существенное различие между двумя группами получено по критерию среднего количества опухолевых соседей у опухолевой клетки эпителия в ацинусах с видимыми изменениями морфологии относительно неопухолевой ткани ($p\text{-value} = 0.06$).

Список литературы

1. Cancer. <https://www.who.int/news-room/fact-sheets/detail/cancer>.
2. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **68**, 394–424 (2018).
3. Three measures of cancer mortality. <https://ourworldindata.org/grapher/cancer-deaths-rate-and-age-standardized-rate-index>.
4. Number of people with cancer. <https://ourworldindata.org/grapher/number-of-people-with-cancer>.
5. Cancer incidence. <https://ourworldindata.org/grapher/cancer-incidence>.
6. McKinnon, K. M. Flow Cytometry: An Overview. *Curr. Protoc. Immunol.* **120**,.
7. *Biomarkers for Immunotherapy of Cancer: Methods and Protocols*. vol. 2055 (Springer New York, 2020).
8. Lin, J.-R., Fallahi-Sichani, M. & Sorger, P. K. Highly multiplexed imaging of single cells using a high-throughput cyclic immunofluorescence method. *Nat. Commun.* **6**, 1–7 (2015).
9. Lin, J.-R. *et al.* Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes. *Elife* **7**,.
10. Martorana, E. *et al.* Prostate MRI and transperineal TRUS/MRI fusion biopsy for prostate cancer detection: clinical practice updates. *Turkish Journal of Urology* **45**, 237–244.
11. Drost, F. H. *et al.* Prostate MRI, with or without MRI-targeted biopsy, and systematic biopsy for detecting prostate cancer. *Cochrane Database Syst. Rev.* **2019**, (2019).
12. Wagenlehner, F. M. E., Pilatz, A., Waliszewski, P., Weidner, W. & Johansen, T. E. B. Reducing infection rates after prostate biopsy. *Nat. Rev. Urol.* **11**, 80–86.
13. Correlation between MRI phenotypes and a genomic classifier of prostate cancer: preliminary findings. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6684343/>.
14. Dalela, D., Löppenberg, B., Sood, A., Sammon, J. & Abdollah, F. Contemporary Role of the Decipher® Test in Prostate Cancer Management: Current Practice and Future Perspectives. *Rev. Urol.* **18**, 1–9 (2016).
15. Li, P. *et al.* Genes involved in prostate cancer progression determine MRI visibility. *Theranostics* **8**, 1752–1765 (2018).
16. Salami, S. S. *et al.* Biologic Significance of Magnetic Resonance Imaging Invisibility in Localized Prostate Cancer. *JCO Precision Oncology* 1–12.
17. About Cancer. *National Cancer Institute* <https://www.cancer.gov/about-cancer>.
18. Alizadeh, M. & Alizadeh, S. Survey of clinical and pathological characteristics and

- outcomes of patients with prostate cancer. *Glob. J. Health Sci.* **6**, 49–57 (2014).
19. Histology of prostate cancer. <http://oncolex.org/Prostate-cancer/Background/Histology>.
 20. Vilanova, J. C., Catalá, V., Algaba, F. & Laucirica, O. *Atlas of Multiparametric Prostate MRI: With PI-RADS Approach and Anatomic-MRI-Pathological Correlation*. (Springer, 2017).
 21. Thompson, I. M. *et al.* Prevalence of prostate cancer among men with a prostate-specific antigen level $< \text{or} = 4.0$ ng per milliliter. *N. Engl. J. Med.* **350**, 2239–2246 (2004).
 22. Epstein, J. I. *et al.* The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: Definition of Grading Patterns and Proposal for a New Grading System. *Am. J. Surg. Pathol.* **40**, 244–252 (2016).
 23. Федерации, М. З. Р. *Клинические рекомендации 'Рак предстательной железы'*. vol. 71 (2018).
 24. Hu, M. & Polyak, K. Microenvironmental regulation of cancer development. *Curr. Opin. Genet. Dev.* **18**, 27–34.
 25. Baghban, R. *et al.* Tumor microenvironment complexity and therapeutic implications at a glance. *Cell Commun. Signal.* **18**, (2020).
 26. Elenbaas, B. & Weinberg, R. A. Heterotypic signaling between epithelial tumor cells and fibroblasts in carcinoma formation. *Exp. Cell Res.* **264**, 169–184 (2001).
 27. Joyce, J. A. & Pollard, J. W. Microenvironmental regulation of metastasis. *Nat. Rev. Cancer* **9**, 239–252.
 28. Maman, S. & Witz, I. P. A history of exploring cancer in context. *Nat. Rev. Cancer* **18**, 359–376.
 29. Dolberg, D. S. & Bissell, M. J. Inability of Rous sarcoma virus to cause sarcomas in the avian embryo. *Nature* **309**, 552–556 (1984).
 30. Wei, S. C., Duffy, C. R. & Allison, J. P. Fundamental Mechanisms of Immune Checkpoint Blockade Therapy. *Cancer Discov.* **8**, 1069–1086 (2018).
 31. Luo, Y. *et al.* Targeting tumor-associated macrophages as a novel strategy against breast cancer. *J. Clin. Invest.* **116**, 2132–2141.
 32. Lee, S.-J. Infection after transrectal ultrasound-guided prostate biopsy. *Korean J. Urol.* **56**, 346–350.
 33. Wu, Y.-P. *et al.* Risk factors for infectious complications following transrectal ultrasound-guided prostate biopsy. *Infect. Drug Resist.* **11**, 1491–1497 (2018).
 34. Johnson, D. C. *et al.* Detection of Individual Prostate Cancer Foci via Multiparametric Magnetic Resonance Imaging. *Eur. Urol.* **75**, 712–720.
 35. Chang, S. S. & Cookson, M. S. *Prostate Cancer: Clinical Case Scenarios*. (Springer, 2018).

36. Pessoa, R. R. *et al.* Value of 3-Tesla multiparametric magnetic resonance imaging and targeted biopsy for improved risk stratification in patients considered for active surveillance. *BJU Int.* **119**, 535–542.
37. Patel, A. P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396–1401 (2014).
38. Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).
39. McCarthy, M. E. & Birtwistle, M. R. Highly Multiplexed, Quantitative Tissue Imaging at Cellular Resolution. *Curr. Pathobiol. Rep.* **7**, 109–118 (2019).
40. Angelo, M. *et al.* Multiplexed ion beam imaging of human breast tumors. *Nat. Med.* **20**, 436–442 (2014).
41. Keren, L. *et al.* A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging. *Cell* **174**, 1373–1387.e19 (2018).
42. Goltsev, Y. *et al.* Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* **174**, 968–981.e15 (2018).
43. Kamentsky, L. *et al.* Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software. *Bioinformatics* **27**, 1179–1180 (2011).
44. Carpenter, A. E. *et al.* CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, R100 (2006).
45. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *Plan. Perspect.* **285**, 296 (1975).
46. Kromp, F. *et al.* Deep Learning architectures for generalized immunofluorescence based nuclear image segmentation.
47. Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv [cs.CV]* (2015).
48. Sra, S., Nowozin, S. & Wright, S. J. *Optimization for Machine Learning*. (MIT Press, 2012).
49. Van Gassen, S. *et al.* FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry A* **87**, 636–645 (2015).
50. Kohonen, T. The self-organizing map. *Proc. IEEE* **78**, 1464–1480 (1990).
51. Worldwide cancer data. *World Cancer Research Fund International*
<https://www.wcrf.org/dietandcancer/cancer-trends/worldwide-cancer-data>