

Федеральное государственное автономное образовательное
учреждение высшего образования
«Национальный исследовательский университет
«Высшая школа экономики»

На правах рукописи

Янив Коэн

**ИК-спектрография и томография тканей человека
и их анализ методами машинного обучения**

РЕЗЮМЕ ДИССЕРТАЦИИ
на соискание ученой степени
кандидата технических наук

Научный руководитель:
доктор технических наук, профессор
Крук Евгений Аврамович

Москва – 2022

Актуальность темы исследования

Диагностика рака является особенно активной областью исследований; появляется большое количество публикаций, особенно в отношении рака шейки матки, легких, предстательной железы, толстой кишки, пищевода, желудка, головного мозга и кожи и т.д. В течение последних десятилетий научное общество заявило о нескольких распространенных врожденных раковых заболеваниях, связанных со сквозными отверстиями, и современных методах диагностики.

В настоящее время золотым стандартом диагностики является гистопатология, основанная на иссечении и окрашивании биологических тканей. Однако, главная проблема этого метода заключается в том, что для получения окончательных результатов исследования может понадобиться от нескольких дней до нескольких недель, в связи с затратами времени на подготовку тканей и многоэтапной процедурой обработки. Кроме того, биопсия требует инвазивного удаления тканей, что может быть болезненно и дорого; за один сеанс может быть проведена биопсия очень ограниченного количества тканей. И последнее, но не менее важное, взятые образцы биопсии должны быть сохранены и перевезены в лабораторию для проведения экспертного анализа; это увеличивает стоимость, повышает вероятность неправильной обработки, уничтожения или потери образцов, а также вызывает значительную задержку в получении результатов. Наконец, интерпретация результатов биопсии обычно проводится с помощью микроскопического анализа, что довольно субъективно и плохо подходит для систематического толкования.

Поэтому в медицинской диагностике большой интерес проявляется к безопасным неоперативным технологиям обнаружения с улучшенной чув-

ствительностью, способным выявлять внутренние раковые заболевания на более ранней стадии, а также к усовершенствованным методам идентификации для дифференциации рака, доброкачественных состояний и других патологий внутренних тканей.

Важной областью исследований в этой области в настоящее время является поиск возможностей для как можно более раннего и неинвазивного выявления участков поврежденных тканей для их оперативного удаления. В недавних исследованиях использовались различные методы, включающие оптические, акустические, магнитные и рентгеновские устройства, а также соответствующие методы. В работах Минского, был разработан метод конфокальной сканирующей микроскопии, то есть оптический метод визуализации для повышения оптического разрешения и контрастности. В работах Хаунсфилда представлен метод «СТ Scan» рентгеновской компьютерной томографии, то есть методика медицинской визуализации для использования в радиологии для неинвазивного получения подробных изображений тела для диагностических целей. В работах Мэнсфилда и Лотербура, разработан метод магниторезонансной томографии (МРТ) для медицинской визуализации в радиологии для формирования изображений анатомии и физиологических процессов в организме.

В качестве примера, при болезни Барретта и язвенном колите, признаком канцерогенеза является плоская дисплазия. По действующим стандартам, в данном случае берутся множественные случайные образцы биопсии в поисках пресловутой иголки в стоге сена. Однако получение большого количества случайных образцов для биопсии утомительно, дорого, создает потенциальные осложнения и приводит к потенциальным ложноотрицательным результатам (поскольку обычно берется менее 5% образцов слизистой оболочки).

В качестве другого примера, методы для улучшения обнаружения аде-номы обладают недостатками: либо эндоскопическими слепыми пятнами (за складками, изгибами), либо визуализированные тонкие повреждения в поле зрения (например, плоские и вдавленные повреждения). Большая часть усилий на сегодняшний день была направлена на усовершенствование эндоскопа (высокая четкость, узкополосная визуализация, автофлуоресценция и т. д.) или контрастных веществ (молекулярная диагностика, хромозэндоскопия). Упомянутые попытки улучшают способность обнаружения, но не устраняют проблему полностью.

В ряде исследований использовались способы, предполагающие инфракрасную спектроскопию и томографию с преобразованием Фурье в сочетании с нарушенным полным внутренним отражением (ATR-FTIR) для анализа слюнных экзосом от рака полости рта или для выявления стволовых клеток в исследованиях рака и к характеристике клеточного цикла вариаций. ИК-спектры могут быть использованы в качестве чувствительного маркера структурных изменений клеток и перестройки, происходящей в клетках, и большинство биомолекул приводят к ИК поглощению в полосе между 1800 и 700 см^{-1} , которые известны как характерные зоны, или зоны первичного поглощения.

FTIR-спектроскопия и визуализирующая спектроскопия криосекций являются широко исследованными методами диагностики злокачественных новообразований, которые дают результаты, сопоставимые с обычным исследованием патологии замороженных срезов (FS). Эти методы получили все большее признание в последние годы, но их применение в клинической практике все еще ограничено. Лишь относительно недавно значительные улучшения в технологии позволили получить чувствительность, необходимую для изучения биологических молекул, и только в течение последнего десятилетия стала возможной FTIR-томография. Неко-

торые из наиболее важных технологических разработок включают интерферометр, высокочувствительные детекторы и матричные детекторы, мощные источники света и технологию полного ослабленного отражения (ATR). Развитие вычислительной мощности также позволило быстро обрабатывать большие наборы данных. Таким образом, разработки реализованы в Медицинской инфракрасной оптической системе (IROS).

Несмотря на достижения в этой области, остается сложной задача разработки быстрого и надежного метода интраоперационной диагностики рака, который может быть адаптирован и внедрен в клиническую практику. Существующие методы, как правило, сложны и отнимают много времени, и поэтому их трудно перевести в клинические условия.

Подход к процессу диагностики, включающий ATR-FTIR, кажется очень многообещающим. Однако, нерешенными остаются следующие вопросы: а) методы идентификации спектральных паттернов, отличающих раковые ткани от здоровых; б) методы анализа наборов данных; в) простые и стабильные приборы.

Цели и задачи диссертации

Таким образом, **цель** данной работы состоит в том, чтобы разработать неинвазивную, быструю, компактную, удаленную, портативную и высокочувствительную методику интраоперационной диагностики и соответствующий инструмент. Для решения этой проблемы мы планируем использовать измерения контраста разности температур на фоне с использованием ATR-FTIR.

Для достижения этой цели необходимо решить следующие **задачи**:

1. разработать воспроизводимые методы ИК-томографии и спектроскопии, которые можно превратить в клинические инструменты для более ранней неинвазивной диагностики аномалий внутренних тканей человека: опухолей и заболеваний;
2. выбрать наиболее подходящие из существующих в настоящее время диагностических, аналитических и статистических методов, чтобы учесть и ослабить некоторые недостатки, а также предложить более оптимальные методы анализа исходных данных, наблюдаемых экспериментально;
3. разработать диагностические приборы на основе инфракрасной камеры и FTIR-спектрометра, которые были бы достаточно простыми и стабильными инструментами в клинической практике;
4. разработать методы машинного обучения, которые были бы достаточно простыми и стабильными для использования в клинической практике;
5. выбрать наилучшую платформу машинного обучения для спектрального анализа тканей человека с целью ранней неинвазивной диагностики и дифференциации раковых и нераковых аномалий кожи.

Научная новизна работы

Данная работа вносит свой вклад в решение проблемы ранней неинвазивной диагностики аномалий тканей человека: опухолей и заболеваний – на основании теоретических основ ИК-томографии и спектроскопии, более эффективного использования инструментария для этих целей, аналитических и статистических методов анализа исходных данных, а также элементов искусственных нейронных сетей и алгоритмов машинного обучения. Таким образом:

1. использована облачная база данных с методами машинного обучения в области идентификации и классификации рака кожи, где это ранее не применялось, что позволяет повысить способность системы правильно идентифицировать и классифицировать биохимический состав тканей и помочь врачу в качестве системы поддержки принятия решений реального времени на месте проведения операции;
2. экспериментально установлены тепловые сигнатуры различных тканей организма человека, соответствующие нормальному и абнормальному состояниям;
3. предложен новый метод обнаружения и идентификации рака желудка, колоректального рака и рака шейки матки, отличающийся от других методов в том, что обычные методы основаны на субъективном осмотре в видимом спектральном диапазоне, тогда как разработанный метод основан на характеристиках перегрева раковых тканей и обнаружении этих аномалий с использованием инфракрасной камеры ближнего или среднего диапазона;
4. предложен новый метод предварительной классификации аномалий, который отличается от других методов тем, что он выполняется с использованием ИК-оптического диапазона в диапазоне 3–12 мкм;
5. предложен новый способ ранней диагностики или удаления рака на основании ИК-перегрева, основанный на пассивной визуализации изменения тепловой подписи в опухоли из-за быстрого изменения окружающей температуры вокруг этих новообразований с применением методов нагрева и охлаждения тканей. Этот метод позволяет отбирать и характеризовать эти опухоли более точно на ранних стадиях их пространственно-временной эволюции, где они не могут быть обнаружены с помощью обычного человеческого или машин-

ного зрения. Значение этих новых методов нагрева и охлаждения в том, что они показывают, что удаление рака можно проводить неинвазивно, чтобы обеспечить полное удаление рака внутри/снаружи человеческого тела.

Практическая значимость

Практическое применение предложенного подхода, основанного на объединенной платформе междисциплинарных методов, теоретических и экспериментальных, позволило создать Медицинскую инфракрасную оптическую систему IROS, которая представляет собой настольное устройство для диагностики тканей в режиме реального времени и использует FTIR-спектроскопию и принцип ATR для точной диагностики ткани;

на основании полученных результатов был разработан новый эндоскоп, содержащий массив инфракрасных оптических микродетекторов, работающих в диапазоне длины волны 3–12 мкм, работающих как в интегральном, так и в спектральном режимах, основанных на знаке и амплитуде контраста злокачественных аномалий по сравнению с таковыми для нормальных живых тканей;

была создана обученная искусственная нейронная сеть, служащая для прогнозирования рака и других патологий на основе измерений с помощью устройства FTIR-ATR. Надежность предлагаемого метода нейронной сети была проверена на данных, собранных с помощью медицинского устройства IROS и полученных с помощью биопсии;

на основе комбинации ИК-тепловизионной визуализации и спектрального анализа средствами машинного обучения, была разработана система поддержки принятия решений для раннего выявления и идентификации рака в режиме реального времени на месте проведения операции;

медицинская IROS вместе с системой поддержки принятия решений служат основой для выполнения оптической биопсии или спектральной гистопатологии тканей в реальном времени на месте в операционной. Это мощный инструмент для оперативного интраоперационного экономически эффективного простого в использовании неразрушающего автоматизированного раннего неинвазивного обнаружения и идентификации различных видов внутренних раковых образований тканей человека.

Основные положения, выносимые на защиту

- 1) предложен новый подход к ИК-томографии и спектрографии с соответствующей теоретической и концептуальной основой;
- 2) предложен усовершенствованный метод раннего обнаружения с помощью теплового ИК-подхода, основанный на пассивной визуализации изменений тепловой сигнатуры в опухолях из-за незначительных изменений температуры окружающей среды вокруг этих опухолей после применения методов нагрева и охлаждения ткани, что позволяет обнаруживать и более точно характеризовать эти опухоли на более ранних стадиях их пространственно-временной эволюции;
- 3) предложены новые экспериментальные методы, основанные на использовании спектроскопии ATR-FTIR в качестве диагностического инструмента для быстрой интраоперационной идентификации злокачественных новообразований;
- 4) предложена адаптация структуры и алгоритмов искусственной нейронной сети к исследованию аномалий тканей: опухолей, раковых заболеваний и метастазов, для экспресс-анализа аномалий такого рода.
- 5) предложена единая платформа, сочетающая элементы теоретической базы, ATR-FTIR, нейронную структуру, алгоритмы и элементы машинного обучения, позволяющая медицинскому специалисту полу-

чить простой, но недорогой мощный инструмент для раннего бесконтактного обнаружения и идентификации опухолей, а также возможность мониторинга удаления опухоли во время операции с помощью чувствительной камеры методом нагрева и охлаждения.

Методология диссертационного исследования

Использование облачной базы данных совместно с методами машинного обучения позволяет в рамках разрабатываемой системы повысить скорость идентификации/классификации различного биохимического состава тканей и помогает врачу в режиме реального времени в качестве системы поддержки принятия решений. В этой работе мы исследовали возможность использования ATR-FTIR- спектроскопии необработанных, свежесрезанных образцов тканей GI в качестве диагностического инструмента для быстрой интраоперационной идентификации злокачественных новообразований.

Эндоскопия является наиболее важным методом визуализации в области гастроэнтерологии. Это помогает врачу принимать более обоснованные решения, более точно ставить диагноз и, конечно, является мощным терапевтическим инструментом. Прогнозирование и окончательная выживаемость рака желудочно-кишечного тракта зависят от стадии заболевания. Необходимы методы обнаружения желудочно-кишечных поражений на ранней стадии, что трудно сделать с помощью обычной эндоскопии с белым светом оптической визуализации, поскольку в таких образцах изменения в морфологических характеристиках невелики. С момента появления видеоэндоскопов в 1990 году произошел большой прогресс в качестве изображения. Эндоскопы высокой четкости и экраны в настоящее время становятся стандартом. Параллельно с разработкой эндоскопов мы продолжали искать другие методы усовершенствования оптической части системы. Рассматривались следующие технологии: хромоэндоскопия, NBI,

FICE автофлуоресценция и конфокальная лазерная эндоскопия. В наших исследованиях мы учитываем некоторые недостатки этих технологий, с которыми мы встречались в ходе анализа, такие как длительность, цена и необходимость подготовки операторов. Учитывая, что человеческое тело излучает в ИК-диапазоне, и каждая ткань – либо нормальная, либо не нормальная – имеет свою тепловую подпись, мы в нашей диссертации предложили улучшить раннее выявление на основе теплового ИК-подхода, основанного на пассивной визуализации. Учет изменений тепловой подписи в опухолях из-за мельчайших изменений температуры окружающей среды после применения методов нагрева и охлаждения ткани, оказался полезным для более точной характеристики этих опухолей на ранних стадиях их пространственно-временной эволюции.

Сочетание ИК тепловизионных изображений и машинного обучения с использованием спектрального анализа позволяет создать новую экономически оправданную платформу для раннего обнаружения рака, которая может использоваться не в лабораторных условиях и в режиме реального времени. Таким образом, в рамках настоящей работы разрабатывается система поддержки принятия врачебного решения по диагностике рака с помощью оптической биопсии и спектральной гистопатологии ткани.

Степень достоверности результатов

Достоверность полученных результатов в диссертационном исследовании подтверждается:

- соответствием экспериментальных данных, полученных по предложенным методикам и по стандартным методикам;
- соответствием полученных результатов фундаментальным физическим и биологическим принципам;
- численным моделированием на современных пакетах расчета.

Личный вклад автора

- разработка и внедрение математических моделей и алгоритмов;
- доказательство теорем, разработка алгоритмов и их тестирование;
- сбор данных, проведение вычислительных, пре-клинических и клинических экспериментов.

Апробация результатов исследований

Работа прошла апробацию на следующих международных конференциях:

- 1) “Method of Infrared Thermography for Earlier Diagnostics of Gastric Colorectal and Cervical Cancer.” Innovation in Medicine and Healthcare 2016. InMed 2016. Smart Innovation, Systems and Technologies, Tenerife, Spain, 15-17 June 2016.
- 2) "Real-Time, On-Site, Machine Learning Identification Methodology of Intrinsic Human Cancers Based on Infra- Red Spectral Analysis – Clinical Results”. 11th International Multi-Conference on Complexity, Informatics and Cybernetics (IMCIC 2020) Orlando, USA, March, 2020. (Chosen as “best 25%-30% of the papers, which authors registered for IMCIC 2020”)
- 3) “Artificial Neural Network in Predicting Cancer Based on Infrared Spectroscopy”. Intelligent Decision Technologies. IDT 2020. Smart Innovation, Systems and Technologies, Split City, CROATIA, 17-19 June 2020. KES Virtual Conference center.

Список опубликованных статей, в которых отражены основные научные результаты диссертации

- *Список статей, опубликованных в журналах, включенных в международную систему цитирования Scopus:*

- 1) B.-Z. Dekel, A. Zilberman, N. Blaunstein, Y. Cohen, M.B. Sergeev, L.L. Varlamova and G.S. Polishchuk (2016), “Method of Infrared Thermography for Earlier Diagnostics of Gastric Colorectal and Cervical Cancer” (*Метод инфракрасной термографии для ранней диагностики колоректального рака желудка и рака шейки матки*). In: Chen YW., Tanaka S., Howlett R., Jain L. (eds) Innovation in Medicine and Healthcare 2016. InMed 2016. Smart Innovation, Systems and Technologies, vol 60. Springer, Cham. Pages 83-92 (DOI: https://doi.org/10.1007/978-3-319-39687-3_8)
- 2) Y. Cohen, A. Zilberman, B.-Z. Dekel, E. Krouk, “Real-Time, On-Site, Machine Learning Identification Methodology of Intrinsic Human Cancers Based on Infra- Red Spectral Analysis – Clinical Results,” (*Методология идентификации внутренних раковых заболеваний человека на месте в режиме реального времени с использованием машинного обучения на основе инфракрасного спектрального анализа - клинические результаты*) Proceedings of The 11th International Multi-Conference on Complexity, Informatics and Cybernetics (IMCIC 2020) Orlando, USA, March, 2020, pages 72 – 77;.
- 3) Cohen Y., Zilberman A., Dekel B.Z., Krouk E. (2020) “Artificial Neural Network in Predicting Cancer Based on Infrared Spectroscopy” (*Методология идентификации внутренних раковых заболеваний человека на месте в режиме реального времени с использованием машинного обучения на основе инфракрасного спектрального анализа - клинические результаты*). In: Czarnowski I., Howlett R., Jain L. (eds) Intelligent Decision Technologies. IDT 2020. Smart Innovation, Systems and Technologies, vol 193, pp 141-153. Springer, Singapore (DOI: https://doi.org/10.1007/978-981-15-5925-9_12)
- 4) Y. Cohen, A. Zilberman, B.-Z. Dekel, E. Krouk, "Real-Time, On-Site, Machine Learning Identification Methodology of Intrinsic Human Can-

cers Based on Infra- Red Spectral Analysis – Clinical Results” (*Методология идентификации внутренних раковых заболеваний человека на месте в режиме реального времени с использованием машинного обучения на основе инфракрасного спектрального анализа - клинические результаты*). Journal of Systemics, Cybernetics and Informatics (JSCI), vol. 18, No. 2, 2020.

- *Список статей, опубликованных в медицинских журналах:*

5) Y. Cohen, B.-Z. Dekel, E. Krouk, and N. Blaunstein “Method of Infrared Thermography for Earlier Diagnostics of Gastric Colorectal and Cervical Cancer” (*Метод инфракрасной термографии для ранней диагностики колоректального рака желудка и рака шейки матки*). EC Gastroenterology and Digestive System, Vol. 6.10 , pp. 916-933., 2019.

Структура диссертации

Во введении приведены: актуальность исследования, литературный анализ, цели и задачи, научная новизна, практическая значимость, методология исследования.

В главе 1 показана способность обнаруживать на мышах небольшие опухоли и метастазы (которые не видны невооруженным глазом или с помощью визуальной камеры), с использованием тепловой камеры с нагреванием или охлаждением, что также может быть использовано для мониторинга удаления опухоли. Кроме того, в ходе тепловизионной лапароскопической процедуры для удаления опухолей был снят процесс удаления, чтобы показать разницу в температуре между нормальными тканями и опухолью. В главе представлена также математическая модель исследуемого процесса.

Глава 2 посвящена разработке комбинированного устройства и метода для анализа тканей и биологических клеток с помощью Медицинской системы IROS.

В главе обсуждаются принципы и устройство для обнаружения патологий с помощью FTIR-ATR, обсуждается реализация основных элементов разработанного устройства, описываются потоки информации и дается полная схема системы передачи данных от каждого пациента различных больниц в Центр для сбора информации и для отправления решения, принятого медицинским персоналом после анализа результатов машинного обучения.

В главе 3 выбран и выстроен классификатор для различения нормальных, раковых тканей и тканей с другими патологическими изменениями по измеренным данным спектроскопии FTIR. Описывается подготовка данных для предварительной обработки и подход к их классификации на основе машинного обучения. Представленные методы классификации используют:

- 1) регрессия методом дробных наименьших квадратов (PLSR);
- 2) регрессия основных компонентом (PCR);
- 3) классификатор линейного дискриминационного анализа (LDA);
- 4) наивный байесовский классификатор (NBC).

Указанные методы позволяют выбрать биомаркеры для распознавания нормальных и злокачественных клеток и получить характеристики злокачественности в спектральном регионе Mid-IR.

Основной целью **главы 4** является представление результатов оценки искусственной нейронной сети (ANN) в прогнозировании рака и других патологий на основе измерений с помощью устройства FTIR-ATR. Для

классификации заболевания была использована нейронная сеть с контролируемым обучением (реальные результаты известны заранее) для классификации заболевания: рак/не рак или рак-полип-нормальный. Предварительные результаты основаны на 76 образцах влажной ткани из клинических испытаний. Проведено обсуждение концепции искусственной нейронной сети (ANN), биологических и искусственных нейронных сетей, разработки классификаторов Feed-forward ANN и алгоритмов обучения сети.

В главе 5 представлено заключение по результатам исследований, представленных в предыдущих главах. Исследование показало, что результаты были достигнуты путем применения машинного обучения (ML) на основе методов классификации 76 влажных образцов: "Рак" или "не рак": правильная классификация была получена во всех 100% испытаний; неправильная классификация - 0,0%.,

В главе 6 показано возможное практическое применение алгоритмов, методов и методик, предлагаемых для визуализации и идентификации аномалий тканей, включая различные опухоли и внутренние раковые заболевания, на основе нашего опыта и результатов, полученных в ходе 4-летних исследований.

Наконец, **в приложении** мы даем обзор исследований, основанных на таких аспектах, как спектр электромагнитных волн; инфракрасной спектроскопии с Фурье-преобразованием (FTIR); применение FTIR для диагностики в медицине; колебательной спектроскопии и режимы ИК-спектроскопического отбора, включая режимы FTIR, наиболее часто используемые для проверки клеточного материала.

Заключение

В организме человека иногда развиваются опухоли и другие внутренние поражения, которые нарушают работоспособность пораженных органов и в конечном итоге приводят к смерти. Их раннее выявление и идентификация имеет решающее значение для здоровья и выживания пациентов.

- 1) В этой работе мы представили платформу на основе ИК для обнаружения опухолей (на основе метода томографии) и их идентификации (на основе метода спектрографии); предложен новый подход к ИК-томографии и спектрографии с соответствующей теоретической и концептуальной основой;
- 2) предложен усовершенствованный метод раннего обнаружения с помощью теплового ИК-подхода, основанный на пассивной визуализации изменений тепловой сигнатуры в опухолях из-за незначительных изменений температуры окружающей среды вокруг этих опухолей после применения методов нагрева и охлаждения ткани, что позволяет обнаруживать и более точно характеризовать эти опухоли на более ранних стадиях их пространственно-временной эволюции;
- 3) предложены новые экспериментальные методы, основанные на использовании спектроскопии ATR-FTIR в качестве диагностического инструмента для быстрой интраоперационной идентификации злокачественных новообразований;
- 4) предложена адаптация структуры и алгоритмов искусственной нейронной сети к исследованию аномалий тканей: опухолей, раковых заболеваний и метастазов, для экспресс-анализа аномалий такого рода.
- 5) предложена единая платформа, сочетающая элементы теоретической базы, ATR-FTIR, нейронную структуру, алгоритмы и элементы ма-

шинного обучения, позволяющая медицинскому специалисту получить простой, но недорогой мощный инструмент для раннего бесконтактного обнаружения и идентификации опухолей, а также возможность мониторинга удаления опухоли во время операции с помощью чувствительной камеры методом нагрева и охлаждения.