

На правах рукописи

Фролов Олег Олегович

**РАЗРАБОТКА СПОСОБА АНАЛИЗА СПЕКТРОВ
КОМБИНАЦИОННОГО РАССЕЯНИЯ ДЛЯ ПРИМЕНЕНИЯ В
СТОМАТОЛОГИИ**

2.2.12. Приборы, системы и изделия медицинского назначения

АВТОРЕФЕРАТ

диссертации на соискание учёной степени

кандидата технических наук

Самара – 2025

Работа выполнена в федеральном государственном автономном образовательном учреждении высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева» (Самарский университет) на кафедре лазерных и биотехнических систем.

Научный руководитель:

Тимченко Павел Евгеньевич, кандидат физико-математических наук, доцент, федеральное государственное автономное образовательное учреждение высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева», кафедра лазерных и биотехнических систем, доцент кафедры.

Официальные оппоненты:

Генина Элина Алексеевна, доктор физико-математических наук, доцент, федеральное государственное бюджетное образовательное учреждение высшего образования «Саратовский национальный исследовательский государственный университет имени Н.Г. Чернышевского», кафедра оптики и биофотоники, профессор кафедры;

Зайцев Кирилл Игоревич, доктор физико-математических наук, федеральное государственное бюджетное учреждение науки Федеральный исследовательский центр «Институт общей физики им. А.М. Прохорова Российской академии наук» (ИОФ РАН), лаборатория широкополосной диэлектрической спектроскопии отдела субмиллиметровой спектроскопии, ведущий научный сотрудник.

Ведущая организация:

Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский Томский государственный университет», г. Томск.

Защита состоится 19 марта 2026 г. в 10:00 часов на заседании диссертационного совета 24.2.379.11, созданного на базе федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева», по адресу: 443086, г. Самара, Московское шоссе, д. 34.

С диссертацией можно ознакомиться в библиотеке и на официальном сайте федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева» по электронному адресу: https://ssau.ru/storage/pages/6858/file_692006fec6c605.49356687.pdf.

Автореферат разослан «___» _____ 2026 г.

Ученый секретарь
диссертационного совета 24.2.379.11
д.т.н., доцент

В.А. Зеленский

Общая характеристика работы

Актуальность темы.

Высокая распространенность заболеваний пародонта, достигающая, по данным некоторых авторов, 98 % среди взрослого населения, подразумевает актуальность исследований, направленных на разработку методов диагностики пародонтита. На сегодняшний момент в большинстве случаев заболевание выявляют лишь на стадии активных клинических проявлений. Часто диагностика осложняется тем, что к пародонтиту приводит множество факторов. Нередко он является следствием других заболеваний. Как и при проблемах любой нозологии, ранняя диагностика пародонтита, вероятно, позволила бы купировать активный процесс и значительно улучшить прогноз заболевания. На сегодняшний день золотым стандартом диагностики пародонтита являются рентгеновские методы исследования, но они не всегда способны выявить ранние признаки воспалительного процесса, поскольку костные изменения наблюдаются лишь на более поздних стадиях заболевания. Актуальной задачей является развитие средств и методов диагностики стоматологических заболеваний, позволяющих врачу произвести своевременную и точную диагностику.

Степень разработанности темы.

В современном состоянии развития диагностики пародонтита применяются различные методы, каждый из которых обладает своими сильными и слабыми сторонами. Так, рентгеновские методы, являющиеся золотым стандартом в диагностике, демонстрируют высокую объективность, контрастность и детализацию костных структур, что позволяет достигать точности порядка 0,91 (0,83 – 0,96, 95 % доверительный интервал (ДИ)) с чувствительностью до 0,90 (0,82 – 0,96, 95 % ДИ) (J.Jarupat). Однако, они не всегда выявляют ранние признаки воспалительного процесса, используют ионизирующее излучение и зависят от опыта оператора.

Флюоресцентная спектроскопия характеризуется высокой чувствительностью к изменениям минерального состава и доступностью оборудования, однако ее молекулярная специфичность ограничена из-за широкого спектра перекрывающихся флюоресцентных сигналов (Y.Park).

При этом ИК-спектроскопия, применяемая для анализа белков, липидов и углеводов, демонстрирует точность классификации до 0,95 в сочетании с методами машинного обучения (S.Fujii, S.Silva), но ограничена из-за чувствительности к присутствию воды и недостаточной пространственной разрешающей способности.

Спектроскопия комбинационного рассеяния (КР) ранее использовалась с целью различения трёх основных пародонтопатогенов Авторами (J.Zhang и др.) заявлена точность классификации на уровне ROC AUC 0,94.

Особое внимание уделено методам анализа дентинных материалов при деминерализации. Исследования в этой области активно используют оптическую когерентную томографию (ОКТ), в том числе swept-source ОКТ, как показано в работе Matsuzaki (2021), а также поляризационно-чувствительную оптическую когерентную томографию (ОКТ) (PS-ОКТ), продемонстрированную в исследовании Manesh (2009) для неразрушающей оценки кариозных изменений в дентине. Дополнительно, исследование Koga (2016) показало, что частичная деминерализация дентина способствует более выраженной регенерации костной ткани, что подтверждает важность выбора оптимальных методов оценки. Кроме того, для анализа процессов деминерализации используются различные спектроскопические методы, позволяющие фиксировать изменения в химическом составе дентина, что дополняется вспомогательными методами, такими как ручная деконволюция спектров и анализ ультраструктурных характеристик с использованием РЭМ. Также применяется и ультразвуковая диагностика (Denis и др. 2018) для классификации уровней деминерализации костной ткани. Максимальная точность и специфичность классификации достигала 0,80 и 0,83.

Учитывая как возможности методов, используемых для диагностики пародонтита, так и особенности методик анализа дентинных материалов при деминерализации, спектроскопия комбинационного рассеяния представляется перспективным инструментом, способным объединить преимущества различных подходов. Спектроскопия КР неинвазивна, обеспечивает высокую молекулярную детализацию, позволяет проводить исследования *in vivo* без разрушительной подготовки образцов и эффективно работает в водной среде, что делает её незаменимым элементом в динамическом мониторинге состояния зубных тканей. При этом метод сталкивается с проблемами влияния флюоресцентного фона, сложностями алгоритмической обработки и высокой требовательностью к калибровке оборудования.

Именно поэтому данная диссертационная работа направлена на разработку и оптимизацию способа анализа спектров комбинационного рассеяния, что позволит решить указанные проблемы и объединить достоинства различных способов для повышения точности и надежности диагностики пародонтита.

Цель диссертационной работы.

Целью диссертационного исследования является разработка и экспериментальная валидация способа анализа спектров комбинационного рассеяния биоматериалов, позволяющего проводить неинвазивный контроль состояния твердых тканей зубов и костной ткани при пародонтите, а также оценку биоматериалов.

Задачи диссертационного исследования:

1. Разработка способа анализа спектров комбинационного рассеяния биоматериалов, включающего алгоритмы предобработки и декомпозиции.
2. Определение возможности оценки состояния дентинных материалов по спектральным критериям, для оценки их деминерализации на основе выделения ключевых амплитудных признаков, с помощью спектроскопии комбинационного рассеяния.
3. Апробация способа экспресс-диагностики пародонтита на основе анализа спектров комбинационного рассеяния эмали и костной ткани с целью оценки его диагностической эффективности.

Научная новизна работы:

1. Разработан итеративный алгоритм декомпозиции спектров комбинационного рассеяния биологических материалов, основанный на автоматической оптимизации числа линий, их амплитуд, положений и ширин. Алгоритм при взаимном перекрытии полос до 75 % обеспечивает точность определения позиций линий не хуже $0,1 \text{ см}^{-1}$, относительную погрешность амплитуд и ширин $\leq 0,3 \%$ при детерминации модели $R^2 \geq 0,995$.

2. Предложен способ оценки деминерализации дентинных материалов, основанный на извлечении информативных признаков из декомпозированных спектральных линий комбинационного рассеяния в диапазоне $630\text{--}1450 \text{ см}^{-1}$. Полученные признаки используются в модели логистической регрессии для вероятностной классификации образцов по наличию признаков деминерализации с точностью ROC AUC 1,0 (0,99–1,0, 95% ДИ) для определения факта деминерализации.

3. Предложен способ экспресс-диагностики пародонтита, основанный на выявлении структурных признаков изменения ткани зуба при комплексной обработке декомпозированных спектральных данных комбинационного рассеяния с применением логистической регрессии. Способ позволяет диагностировать признаки пародонтита при анализе спектров комбинационного рассеяния костной ткани на уровне ROC AUC 0,88 (0,78–0,96, 0,95 ДИ), а для спектров эмали ROC AUC составляет 0,94 (0,89–0,98, 0,95 ДИ) при чувствительности и специфичности классификации пародонтита на уровне 0,86 (0,76–0,94, 0,95 ДИ) и 0,86 (0,76–0,94, 0,95 ДИ).

Соответствие паспорту специальности 2.2.12 Приборы, системы и изделия медицинского назначения.

Область исследования соответствует направлениям исследований паспорта научной специальности 2.2.12. Приборы, системы и изделия медицинского назначения в части:

п.14. Методы, модели и алгоритмы, включая распознавание образов, для медицинских информационных и интеллектуальных систем обеспечивающих повышение эффективности медико-биологических исследований и врачебных решений;

п.21. Методы и средства искусственного интеллекта для медико-биологических исследований;

п.22. Специальное программное обеспечение, предназначенное для профилактики, диагностики, лечения и медицинской реабилитации заболеваний, мониторинга состояния организма человека и проведения медицинских исследований.

Положения, выносимые на защиту:

1. Применение итеративного алгоритма декомпозиции спектров комбинационного рассеяния биологических материалов позволяет обеспечить точность определения положения центров спектральных линий не хуже $0,1 \text{ см}^{-1}$ с относительной погрешностью определения амплитуды и ширины линии ниже 0,3 % при взаимном перекрытии полос до 75 % и детерминации модели R^2 не ниже 0,995. (Соответствие пункту 22 паспорта специальности).

2. Использование предложенного способа оценки деминерализации дентинных материалов по спектрам комбинационного рассеяния, с применением логистической регрессии для вероятностной классификации образцов по информативным признакам декомпозированных спектральных линий, позволяет проводить диагностику с точностью на уровне ROC AUC 1,0 (0,99-1,0, 0,95 ДИ). (Соответствие пункту 14, 21 паспорта специальности).

3. Использование предложенного способа экспресс-диагностики пародонтита на основе спектров комбинационного рассеяния эмали и костной ткани, обладающего чувствительностью к ранним изменениям минерального состава, с применением логистической регрессии, позволяет диагностировать признаки пародонтита на уровне ROC AUC 0,94 (0,89-0,98, 0,95 ДИ) для эмали и 0,88 (0,78-0,96, 0,95 ДИ) для костной ткани. (Соответствие пункту 14, 21 паспорта специальности).

Практическая значимость.

Практическая значимость заключается в разработке программного обеспечения, реализующего разработанный способ анализа спектров комбинационного рассеяния для диагностики пародонтита и оценки качества биоимплантов в процессе их производства.

Способ позволяет выявлять признаки пародонтита, в том числе на ранних стадиях развития заболевания, оценивать деминерализацию дентинных материалов с промежуточными значениями.

Результаты диссертационного исследования внедрены в научно-исследовательском институте биотехнологий Самарского государственного медицинского университета (НИИ БиоТех СамГМУ) (ПРИЛОЖЕНИЕ А) (акт о внедрении результатов кандидатской диссертационной работы № 1230 / 40-23-3345 от 10.07.2025) и в стоматологической клинике «Центр восстановительной стоматологии» (ПРИЛОЖЕНИЕ Б) (акт о внедрении результатов кандидатской диссертационной работы от 15.01.2025).

Результаты, вошедшие в диссертационное исследование, были использованы при выполнении работ по теме: «Разработка программного обеспечения для идентификации и аттестации компонентов клеточно-тканевых имплантатов с помощью оптического метода при их производстве» (договор № 14483ГУ/2019 от 17.07.2019 г. с Фондом содействия инновациям). А также были использованы при выполнении работ по проекту 18-315-20017 мол_а_вед «Новые фундаментальные подходы в комплексной оценке структуры костной ткани в норме, при экспериментальном остеопорозе и его профилактики с помощью продуктов биотехнологий».

Достоверность полученных результатов обеспечивается строгостью применённых методов обработки данных и машинного обучения, а также воспроизводимостью экспериментальных измерений. Корректность интерпретации спектров подтверждается статистическим анализом (ROC-AUC, p -значение) и соответствием полученных данных

результатам инструментальных и клинических методов оценки. Основные результаты апробированы в профильных учреждениях и опубликованы в рецензируемых научных изданиях, входящих в базы данных Scopus, Web of Science и ВАК.

Методы исследования.

В работе применялись методы экспериментальной спектроскопии комбинационного рассеяния, а также методы цифровой обработки сигналов, математического моделирования, статистического анализа и машинного обучения. Спектральные данные получены с использованием рамановского спектрометра с лазерным источником возбуждения на длине волны 785 нм.

Первичная предобработка спектров включала: интерполяцию, преобразование длины волны в рамановский сдвиг [см^{-1}], нормализацию (по площади под кривой), а также сглаживание сигналов с использованием фильтра максимального правдоподобия MLE-SG на основе алгоритма Савицкого-Голея. Коррекция флуоресцентной составляющей спектров и базовой линии осуществлялась с использованием алгоритм Improved Modified Multi-Polynomial Fitting (улучшенный модифицированный алгоритм многополиномиального аппроксимирования (I-ModPoly)).

Отличительной особенностью предложенного подхода является использование оригинального итеративного алгоритма декомпозиции спектров, разработанного автором, который учитывает наложение спектральных линий с коэффициентом перекрытия до 75 %.

Для построения моделей классификации использовались алгоритмы логистической регрессии, а также методы машинного обучения с контролем переобучения. Качество моделей оценивалось по метрикам точности, чувствительности, специфичности и площади под ROC-кривой (AUC).

Вся математическая обработка и визуализация данных проводилась с использованием языка программирования Python в средах разработки Jupyter Notebook и PyCharm. Использовались библиотеки NumPy, SciPy, Pandas, Matplotlib, scikit-learn и другие специализированные модули для обработки спектроскопических данных.

Личный вклад автора.

Личный вклад автора заключается в самостоятельной разработке способа обработки и анализа спектров комбинационного рассеяния, включая создание оригинального алгоритма декомпозиции спектров с высоким коэффициентом перекрытия линий. Автором лично выполнено программное обеспечение, реализующее предложенные алгоритмы, включая автоматизированную обработку спектральных данных и построение моделей машинного обучения для диагностики стоматологических заболеваний. Также автором проведена апробация разработанных способов на реальных биологических образцах, статистический анализ, визуализация и интерпретация полученных результатов.

Все оригинальные результаты, представленные в диссертации, получены автором лично либо при его непосредственном участии. Постановка задач и обсуждение результатов осуществлялись совместно с научным руководителем.

Публикации.

По теме диссертации опубликовано 7 статей в научных журналах и изданиях, рекомендованных Высшей аттестационной комиссией и/или рецензируемых базой данных Scopus и WoS. Получено 1 свидетельство о регистрации РИД.

Апробация результатов диссертации.

Основные положения и результаты исследований докладывались и обсуждались на следующих конференциях: X международной конференции по информационным технологиям и нанотехнологиям (ITNT-2024), г. Самара (2024 г.); XXI Всероссийском молодежном Самарском конкурсе-конференция по оптике, лазерной физике и физике плазмы, г. Самара (2022, 2023 гг.); Международной конференции для молодых ученых и студентов Saratov Fall Meeting – International School for Junior Scientists and Students on Optics, Laser Physics & Biophotonics, г. Саратов (2022, 2023 гг.); Международной конференция ALT'23 (2023 г.), г. Самара; Всероссийской научной конференции с международным

участием «Енисейская Фотоника» (2022 г.), г. Красноярск; Всероссийской научно-практической конференции с международным участием «Аспирантские чтения» (2022, 2023 гг.), г. Самара; Международной конференции «Лазерно-информационные технологии» (2021, 2022 гг.), г. Новороссийск; Международной конференции «Fundamentals of laser-assisted micro- and nanotechnologies» (FLAMN-22) (2022 г.), г. Санкт-Петербург; Международной конференции «Физика.СПб» (2021 г.), г. Санкт-Петербург; Международной конференции «EEhPolytech» (2021 г.), г. Санкт-Петербург; Международной конференции «European Conferences on Biomedical Optics Conference» (2021 г.), г. Мюнхен.

Структура и объем диссертации.

Диссертация состоит из введения, пяти глав, заключения, списка сокращений и условных обозначений, списка литературы, включающего 149 наименований, и двух приложений. Работа изложена на 187 страницах машинописного текста, содержит 76 рисунков, 16 таблиц.

Содержание работы

Во введении обоснована актуальность выбранной темы, приведён обзор существующих работ, сформулированы цель и задачи исследования, изложены основные результаты работы, их научная новизна, приведены положения, выносимые на защиту, описана структура диссертации.

В первой главе диссертации «Оптические методы анализа биоматериалов в стоматологии» обобщаются современные неразрушающие подходы к исследованию твёрдых тканей зуба и костей челюстно-лицевой области, выделяются их возможности и ограничения.

Рассмотрены хирургические показания к использованию костнопластических материалов, в частности дентинных трансплантатов из деминерализованного дентина. Показано, что для оценки степени деминерализации активно применяются оптическая когерентная томография (SS-OCT и PS-OCT), растровая электронная микроскопия, гиперспектральная NIR-съёмка и ультразвук.

Описаны источники и особенности обработки аллогенных брэфоматрикс (деминерализованного костного матрикса) и ювенильного дентина из молочных зубов.

Рентгенография и КТ остаются «золотым стандартом» для оценки потери костной ткани (точность до 0,91, чувствительность до 0,91) но мало информативны на ранних этапах воспаления. РЭМ даёт ультраструктурные данные, но требует деструкции образца. Флюоресцентные методы (QLF) и FTIR-спектроскопия обеспечивают высокую чувствительность к изменениям минерального и органического состава (ROC AUC до 0,94) при невысокой молекулярной специфичности и чувствительности к воде. Слюнные тесты (ферменты, BANA) и опросники удобны для скрининга, но менее чувствительны к ранней стадии заболевания.

Рассмотрены преимущества спектроскопии КР: неинвазивность, высокая молекулярная детализация, работа в водной среде и возможность портативной реализации.

Перечисленные методы дополняют друг друга: традиционные рентгенографические и электронно-микроскопические подходы дают структурные и морфологические данные, а спектроскопические (флюоресценция, FTIR, КР) — химические маркеры. Среди них спектроскопия КР обладает оптимальным сочетанием чувствительности, молекулярной специфичности и удобства неинвазивного применения, что обосновывает выбор её в качестве основы разработанного в диссертации способа анализа деминерализации и диагностики пародонтита.

Во второй главе подробно рассмотрены существующие подходы к обработке спектров комбинационного рассеяния (КР) в сложных многокомпонентных средах и выделены три ключевых блока предобработки:

I. Предобработка спектров:

1. Удаление спайков (despiking) – итеративные алгоритмы подавления узких высокоинтенсивных артефактов, «cosmic spikes»;
2. Конвертация в см^{-1} – переход от нанометров к волновому числу;
3. Нормализация, например, по площади под кривой (trapezoidal rule);
4. Сглаживание – фильтр Савицкого–Голея (MLE-SG).

II. Коррекция базовой линии (удаление флюоресцентного фона):

1. ModPoly (модифицированный полиномиальный подбор) – итеративное удаление пиков из фона и аппроксимация сглаженного полинома; эффективно, но может вносить артефакты при взаимном перекрытии линии более 50 %;
2. I-ModPoly (улучшенный ModPoly) – учёт статистики шума и автоматическое удаление участков с пиками на каждом шаге, что позволяет минимизировать субъективность выбора опорных точек и избежать «искусственных» пиков;
3. Прочие методы.

III. Алгоритмы декомпозиции спектров на отдельные линии:

1. Слепая и полуслепая деконволюция: используют априорную инструментальную функцию и регуляризаторы для восстановления исходного сигнала при слабом перекрытии полос;
2. Жёсткое моделирование пиков:
 - 1) аппроксимация Voigt-функциями (сведение Лоренцевых и Гауссовых профилей) для учёта инструментальных уширений;
 - 2) гибридные алгоритмы (генетические и градиентные оптимизаторы) для одновременного подбора числа линий, их амплитуд, положений и ширины;
 - 3) Вейвлет- и Гильберт-Хуан разложения (HVD) для более точного отделения перекрывающихся пиков из шума.

Выявлены ограничения: большинство классических алгоритмов рассчитано на простые неорганические спектры с взаимным перекрытием линии не более 25 %, при этом параметры (число пиков, позиции, ширины) фиксируются заранее и плохо адаптируются к биоматериалам с широким спектром вариаций

Полученные выводы обосновывают необходимость разработки нового алгоритма декомпозиции, способного на основе единой схемы предобработки, автоматической коррекции базовой линии и итеративного поиска состава компонент адекватно описывать спектры сложных биоматериалов. Эта задача и решается в третьей главе диссертации.

В третьей главе детально описан новый алгоритм декомпозиции спектров КР биоматериалов, включающий следующие особенности:

1. Общая архитектура алгоритма.

В качестве входных данных используются спектры после удаления фона (I-ModPoly), нормализации и сглаживания. Алгоритм состоит из двух компонент: итеративного подбора состава компонент (для каждого спектра) и построения обобщённой модели (агрегация результатов по всем спектрам).

2. Итеративный поиск линий КР.

По алгоритму «добавь-и-проверяй» последовательно вводятся новые элементарные линии, каждая итерация завершается нелинейной аппроксимацией (Least-Squares, Trust Region Reflective) и проверкой информационных критериев (AIC/BIC). Процесс повторяется до тех пор, пока добавление новой линии не улучшит модель по AIC/BIC и по назначенному порогу амплитуды линии. При наличии заданного шаблона заранее известных линий, верифицируется этот шаблон и используется в качестве итоговой модели.

3. Агрегация и кластеризация параметров.

Собранные по всем спектрам координаты центров линий подвергаются кластеризации (по оригинальному алгоритму с использованием DBSCAN и бутстрепа), что выделяет

устойчивые группы компонент. Для каждого кластера вычисляются средняя позиция x_0 и диапазон разброса амплитуд и полуширин, формируя окончательную модель.

4. Оценка устойчивости к шуму и искажениям.

Проведено моделирование спектров с разными SNR (10–50 дБ) и добавлением искусственного фона. Ошибка восстановления положения пиков не превышает $0,1 \text{ см}^{-1}$, амплитуд и ширины — 0,3 % при реалистичном SNR = 1000.

5. Анализ неопределённостей и их влияние на классификацию.

На основе доверительных интервалов параметров сформированы случайные версии вектора признаков. Продемонстрировано, что качество последующей логистической регрессии остаётся неизменным — модель держит 1,0 accuracy и ROC AUC даже при максимальных флуктуациях.

Разработанный алгоритм декомпозиции не требует жёсткого задания числа компонент и автоматически адаптируется к спектральным особенностям каждого образца, обеспечивает точность на уровне $R^2 = 0,995$ и робастность при взаимном перекрытии линий до 75 % и шуме. Лёгкая модификация позволяет переносить алгоритм на другие биоматериалы.

В четвёртой главе диссертации описано, как разработанный анализ спектров КР был применён к оценке дентинных материалов в разных условиях.

В разделе 4.1 проводилась оценка процесса деминерализации дентина. Были изготовлены 36 образцов из дентина, полученного при удалении зубов по ортодонтическим показаниям, с последующей обработкой в растворах HCl (1,2 Н и 1,8 Н). По готовым образцам зарегистрировано 344 спектра комбинационного рассеяния: 194 спектра минерализованного и 150 спектров полностью деминерализованного дентина.

На рисунке 1 представлены усреднённые спектры двух групп. В деминерализованных образцах заметно снижение интенсивности линий КР ($429, 579, 956$ и 1067 см^{-1}) и одновременно относительное усиление амидных и протеиновых линий КР ($850\text{--}875, 1244, 1446, 1664 \text{ см}^{-1}$).

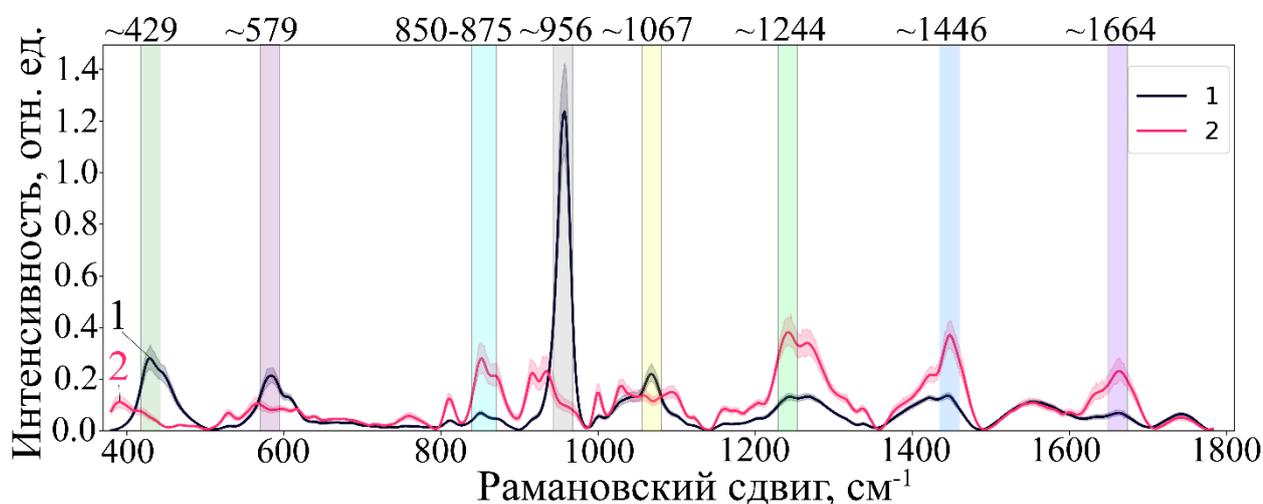


Рисунок 1 – Усредненные спектры КР дентинного материала: 1 – минерализованный; 2 – деминерализованный материал дентина

Для построения классификатора из 167 декомпозированных линий методом ANOVA отобрано 83 признака, затем с помощью рекурсивного исключения (RFE) – 37, и, наконец, по критериям корреляции – 5 наиболее информативных амплитуд ($870, 956, 1244, 1446$ и 1664 см^{-1}).

Перед сравнением групп проверено, что распределения всех пяти признаков значительно отличаются от нормального (критерий Шапиро–Уилка, $p < 0,01$), поэтому применялись непараметрический пермутационный тест ($p \approx 10^{-5}$) и критерий Манна–Уитни, который показан $p \approx 10^{-5}$ для каждого признака. Гистограммы распределений показали чёткое смещение медиан и отсутствия перекрытия между классами.

В качестве базовой модели машинного обучения выбрана логистическая регрессия, так как исходя из статистического анализа распределений значений признаков можно выдвинуть гипотезу, что существует линейная разделяющая граница. Также объем выборки ограничен, модель хорошо интерпретируема и позволяет получать вероятности отнесения к классу, не требует большого количества данных для обучения.

При выборе модели логистической регрессии также учитывалось, что применение других методов: SVM с нелинейным ядром, решающее дерево, случайный лес, CatBoost существенного прироста качества модели не помогло достигнуть.

Классификатор построен на логистической регрессии в конвейере с QuantileTransformer и калибровкой сигмодой, с оптимизацией гиперпараметров (l2-регуляризация, C=0,28, solver=newton-cholesky, max_iter=5000, class_weight='balanced'). Данные были разделены на обучающую (258 спектров) и тестовую (86 спектров) выборки в отношении 75:25.

Как видно из таблицы 1, модель достигает точности 1,0, ROC AUC = 1,0, precision, recall и F1 = 1,0 как на обучении, так и на тесте. Распределение «decision score» показывает чёткое разделение групп, а PCA-проекция, матрица решений и ROC-кривая (рисунок 2) подтверждают полную линейную делимость без переобучения.

Таблица 1. Значения метрик модели логистической регрессии, доверительные интервалы рассчитаны для уровня вероятности не менее 95%

Метрика	Обучающая выборка (258 спектров)	Тестовая выборка (86 спектров)
Точность (accuracy)	1,0 (1,0-1,0)	1,0 (1,0-1,0)
ROC AUC	1,0 (0,99-1,0)	1,0 (0,99-1,0)
Точность (precision)	1,0 (1,0-1,0)	1,0 (1,0-1,0)
Чувствительность (recall)	1,0 (1,0-1,0)	1,0 (1,0-1,0)
Специфичность	1,0 (1,0-1,0)	1,0 (1,0-1,0)
Оценка F1 (F-мера)	1,0 (1,0-1,0)	1,0 (1,0-1,0)
Логарифмическая потеря (Logloss)	0,097 (0,088-0,11)	0,080 (0,067-0,094)

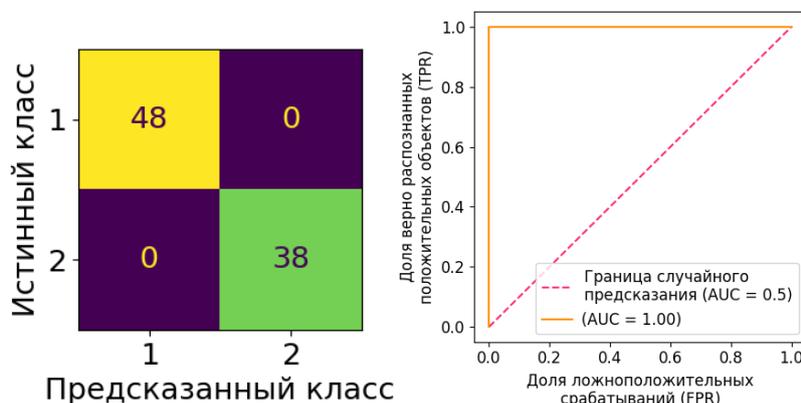


Рисунок 2 – Матрица решений (слева) и Кривая ошибок первого и второго рода (ROC AUC curve) (справа)

Таким образом, в разделе 4.1 доказано, что предложенный способ позволяет определять факт деминерализации дентина по пяти амплитудным признакам и интегральному «decision score».

В разделе 4.2 сравнивался деминерализованный ювенильный дентин (47 образцов) и брэфоматрикссы из кортикальной и губчатой кости (по 36 образцов каждого типа). На рисунке 3 приведены усреднённые нормированные спектры трёх групп. Статистический анализ подтвердил значимые отличия между каждой парой групп ($p < 0,01$), причём кортикальная кость по спектральному профилю ближе к деминерализованному дентину

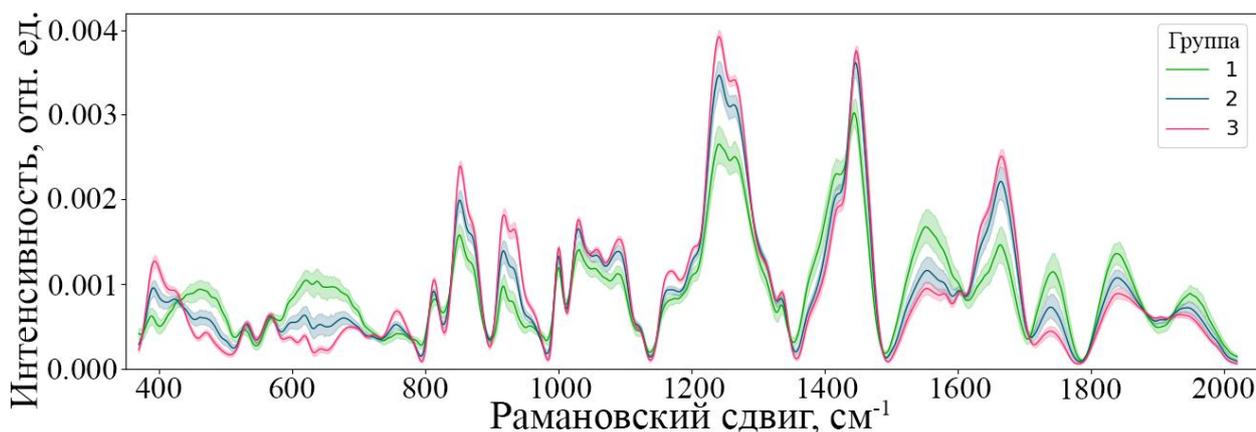


Рисунок 3 – Усредненные спектры КР исследуемых образцов: 1 — брeфоматрикс из губчатой кости, 2 — брeфоматрикс из кортикальной кости, 3 — биоматериалы из деминерализованных ювенильных зубов

Глава 5 посвящена изучению возможности диагностики пародонтита и оценки его лечения с помощью спектроскопии комбинационного рассеяния и разработанного способа обработки спектров КР. В рамках этой главы:

1. Приведены результаты оптической диагностики изменений эмали при пародонтите (раздел 5.1);
2. Оценен состав костной ткани в очагах пародонтита (раздел 5.2);
3. Проанализировано влияние профессиональной гигиены и кюретажа на спектральные свойства твердых тканей зубов (раздел 5.3).

Раздел 5.1. В исследование включены 42 зуба (моляры, премоляры, клыки) пациентов 35–70 лет, удалённых по поводу тяжёлого хронического пародонтита (глубина карманов ≥ 6 мм, подвижность III–IV степени, убыль костной ткани $\geq \frac{1}{2}$ длины корня) и по ортодонтическим показаниям (контроль). В каждой области на эмали регистрировалось по 3 спектра в 3–5 точках поверхности. Кроме того, *in vivo* измерения спектров эмали проведены на 22 зубах одного добровольца с локальным пародонтитом. В итоге проанализировано 193 спектра КР из группы пародонтита и 222 — контрольной.

Экспериментальный стенд включал рамановский пробник RPB-785, полупроводниковый лазер 785 нм/400 мВт, спектрограф Shamrock SR-303i с камерой ANDOR DV-420A-OE (охлаждение -60°C , разрешение 0,15 нм) и компьютер. Регистрация спектров велась в диапазоне $320\text{--}2200\text{ см}^{-1}$ при расстоянии зонда 7 мм.

Анализ спектров здоровых тканей зубов и тканей зубов при пародонтите показывает, что основные спектральные различия тканей зубов при пародонтите проявляются в области спектра $630 - 690$, ~ 850 , $1230 - 1280$, ~ 1319 , $1420 - 1460$, ~ 1552 , $1610 - 1670$ и $\sim 1745\text{ см}^{-1}$ (рисунок 4). Декомпозиция спектров выполнялась предложенным алгоритмом в главе 3. Предобработка включала интерполяцию в волновые числа, нормализацию по площади, сглаживание MLE-Savitzky-Golay ($\sigma=4$) и коррекцию базовой линии I-ModPoly.

Многоэтапный статистический анализ (U-тест, пермутационный тест, рекурсивный отбор признаков) выявил четыре ключевых признака: амплитуды линий ~ 649 , ~ 850 и $\sim 1618\text{ см}^{-1}$ и позицию максимума линии $\sim 957\text{ см}^{-1}$.

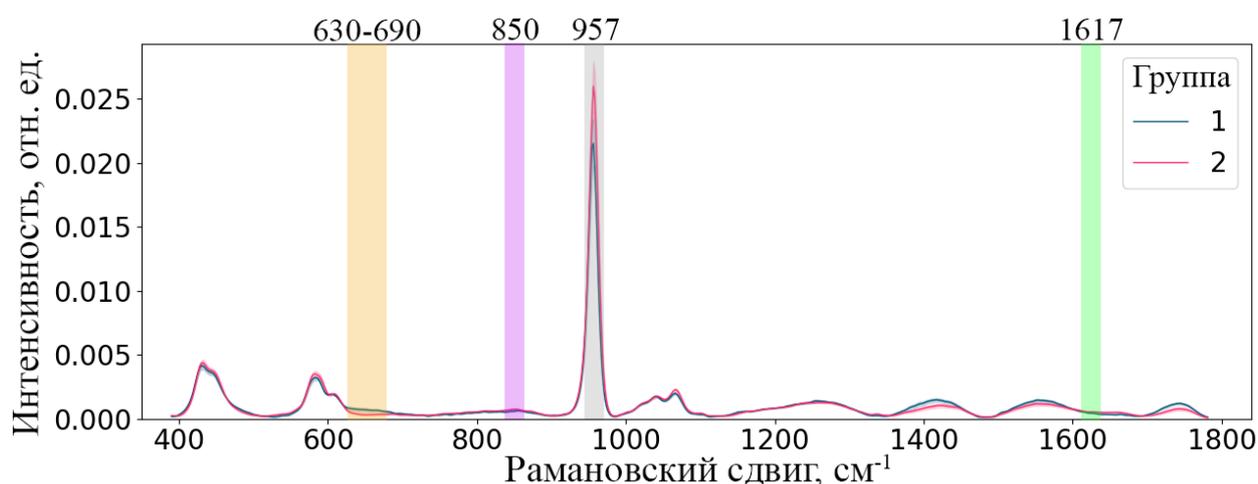


Рисунок 4 – Усредненные спектры КР эмали зубов: 1 – контрольная группа (синий), 2 – с диагнозом пародонтит (красный)

Пермутационный тест показал $p < 0,001$ для всех отобранных признаков, мощность теста $> 0,95$, что подтверждает высокую статистическую значимость наблюдаемых изменений эмали.

Для классификации была использована логистическая регрессия с расширением обучающей выборки через генерацию синтетических данных CTGAN (увеличение каждой группы в 2 раза). Итоговый набор включал 222 контрольных и 193 патологических спектра; в качестве признаков использованы 4 декомпозированных линии. Конвейер обработки данных включал QuantileTransformer и калиброванную логистическую регрессию, гиперпараметры оптимизировались с помощью библиотеки Optuna. Модель обучена на 311 спектрах. Тестовая выборка составила 104 спектра. На тестовой выборке достигнуты: точность (accuracy) 0,88, ROC AUC 0,94, специфичность 0,86, F1-score $\approx 0,89$ (таблица 2). Анализ SHAP-значений подтвердил ключевую роль линии в предсказаниях.

Таблица 2. Значения метрик модели логистической регрессии, доверительные интервалы рассчитаны для уровня вероятности не менее 95%

Метрика	Обучающая выборка (311 спектров)	Тестовая выборка (104 спектра)
Точность (accuracy)	0,90 (0,86-0,93)	0,88 (0,82-0,94)
ROC AUC	0,95 (0,93-0,97)	0,94 (0,89-0,98)
Точность (precision)	0,93 (0,84-0,98)	0,90 (0,85-0,94)
Чувствительность (recall)	0,91 (0,86-0,95)	0,86 (0,76-0,94)
Специфичность	0,91 (0,86-0,95)	0,86 (0,76-0,94)
Оценка F1 (F-мера)	0,90 (0,87-0,94)	0,89 (0,82-0,94)
Логарифмическая потеря (Logloss)	0,31 (0,26-0,35)	0,33 (0,26-0,42)

Кривая ошибок первого и второго рода (ROC AUC кривая) показана на рисунке 5, площадь под кривой (AUC) составляет 0,94, что говорит о приемлемой классифицирующей способности модели.

Сравнительная модель, построенная только на амплитудах линий КР ~ 649 , ~ 850 , ~ 957 и ~ 1618 cm^{-1} , показала заметное ухудшение ROC AUC (0,81 на тесте) и точности при том же наборе спектров. U-тест выявил значимое снижение информативности ряда признаков без декомпозиции.

Разработанный подход к предобработке, декомпозиции и анализу спектров КР эмали позволяет достичь высокой диагностической эффективности экспресс-диагностики пародонтита (ROC AUC = 0,94), превосходя подход без декомпозиции.

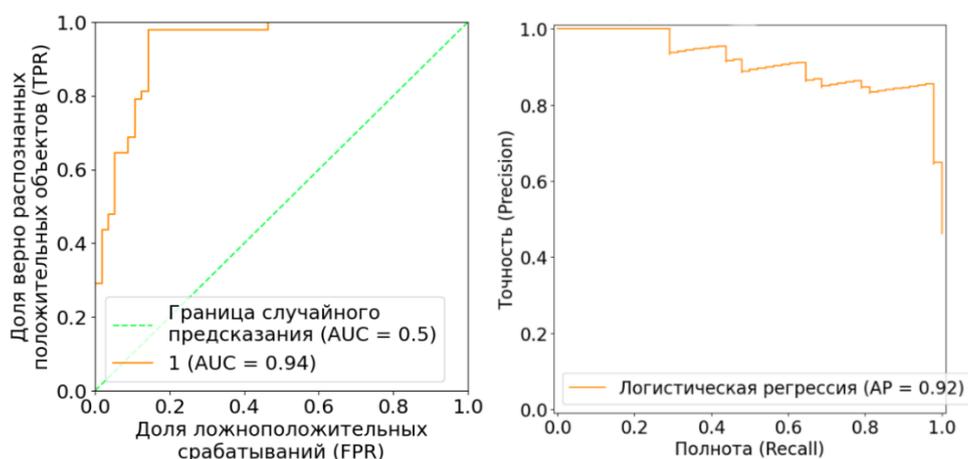


Рисунок 5 – Кривая ошибок первого и второго рода (ROC AUC кривая) (слева) и кривая точности – полноты (Precision – Recall кривая) (справа)

В итоге, признаки после квантильного преобразования к нормальному распределению подаются на вход логистической регрессии, которая выдаёт откалиброванную вероятность p , по которой образцы с $p \leq 0,3$ считаются «здоровыми», $0,3 < p \leq 0,7$ – «пограничными», а $p > 0,7$ – с признаками пародонтита (для бинарного решения в экспериментах использован порог 0,5).

Проведена *in vivo*-валидация разработанной модели на 40 спектрах эмали добровольцев (диагноз при измерениях был неизвестен): при пороге $p > 0,5$ модель классифицировала 35 из них (88 %) как «пародонтит», причём 28 спектров (70 %) получили очень высокую вероятность ($p > 0,8$), 7 спектров (17,5 %) — пограничные оценки ($0,3 < p \leq 0,7$) и лишь 2 спектра (5 %) — здоровые ($p \leq 0,3$).

В разделе 5.2 изучались удалённые зубы (моляры, премоляры, клыки) пациентов 35–70 лет с тяжёлым хроническим пародонтитом и контрольные образцы (ортодонтические удаления). В каждой из трёх микроскопических областей (эмаль, дентин, цемент) снимали спектры в нескольких точках поверхности; всего проанализировано 72 спектра: 46 — пародонтит, 26 — норма.

Нормированные усреднённые спектры КР кости показали при пародонтите относительный рост интенсивности линий КР $\sim 1318 \text{ см}^{-1}$ и $\sim 1744 \text{ см}^{-1}$ (ассоциированные с органическим матриксом и бактериальной ДНК) и снижение амплитуды линий КР: $\sim 852 \text{ см}^{-1}$, $\sim 1039 \text{ см}^{-1}$, $\sim 1067 \text{ см}^{-1}$, $\sim 1244 \text{ см}^{-1}$, $\sim 1444 \text{ см}^{-1}$, $\sim 1663 \text{ см}^{-1}$ по сравнению с контрольной группой (рисунок 6). Декомпозиция выполнялась итеративным алгоритмом (глава 3) с Trust-Region Reflective оптимизацией, $\text{HWHM} \leq 12 \text{ см}^{-1}$, что обеспечило точность $R^2 = 99,82 \%$, $\chi^2 = 2 \cdot 10^{-6}$, $\text{AIC} = -16262$, $\text{BIC} = -14267$.

Из 141 декомпозированной линии методом ANOVA отобраны 70 признаков; затем CatBoost-RFE сократил их до 25; при корреляционном фильтре и экспертном анализе оставлено 5 взаимно слабо коррелированных и статистически значимых амплитуд линий: ~ 852 , ~ 1199 , ~ 1318 , ~ 1663 и $\sim 1744 \text{ см}^{-1}$.

Непараметрический пермутационный тест подтвердил достоверность отличий по всем пяти признакам ($p = 0,00081$; мощность = 0,99 при минимуме 8 спектров на группу)

В качестве классификатора использована логистическая регрессия, обученная на исходных и синтетически дополненных данных CTGAN (общий объём 216 спектров), с оптимизацией гиперпараметров по ROC AUC. На тестовой выборке (25 % образцов) модель показала следующие метрики качества (таблица 3).

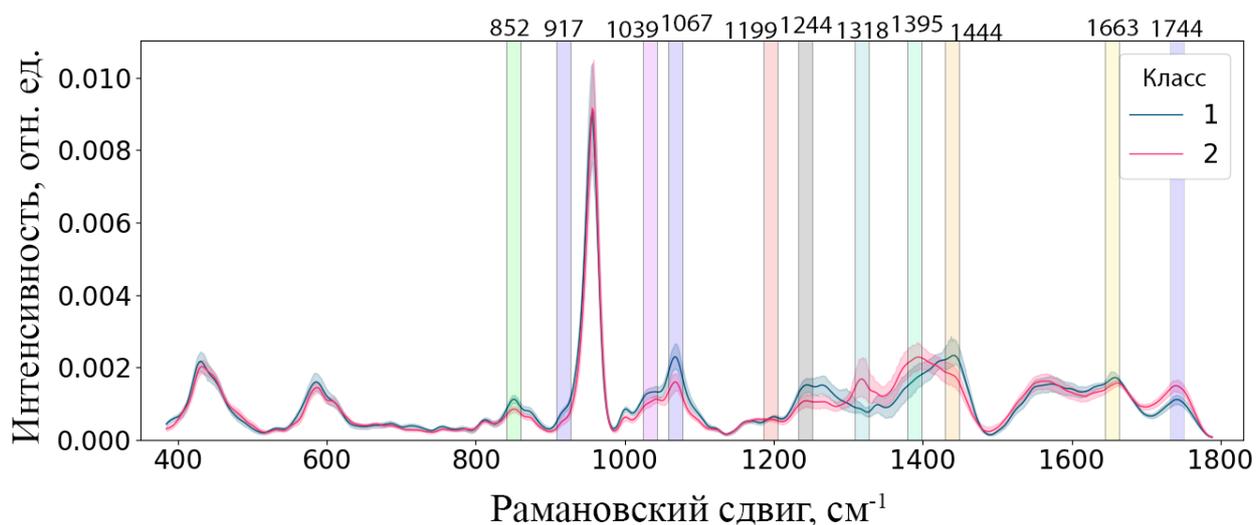


Рисунок 6 – Усредненные спектры КР исследуемых образцов: 1 – костная ткань при пародонтите, 2 – костная ткань в норме

Таблица 3. Значения метрик модели логистической регрессии, доверительные интервалы рассчитаны для уровня вероятности не менее 95 %

Метрика	Обучающая выборка (162 спектра)	Тестовая выборка (54 спектра)
Точность (accuracy)	0,85 (0,80-0,90)	0,80 (0,69-0,91)
ROC AUC	0,92 (0,87-0,96)	0,88 (0,78-0,96)
Точность (precision)	0,86 (0,75-0,95)	0,68 (0,47-0,87)
Чувствительность (recall)	0,79 (0,59-0,95)	0,71 (0,59-0,82)
Специфичность	0,79 (0,59-0,95)	0,71 (0,59-0,82)
Оценка F1 (F-мера)	0,77 (0,68-0,86)	0,72 (0,55-0,86)
Логарифмическая потеря (Logloss)	0,38 (0,33-0,44)	0,43 (0,35-0,52)

Кривая ошибок первого и второго рода (ROC AUC curve) показана на рисунке 7, площадь под кривой (AUC) составляет 0,91, что говорит о достаточной классифицирующей способности модели.

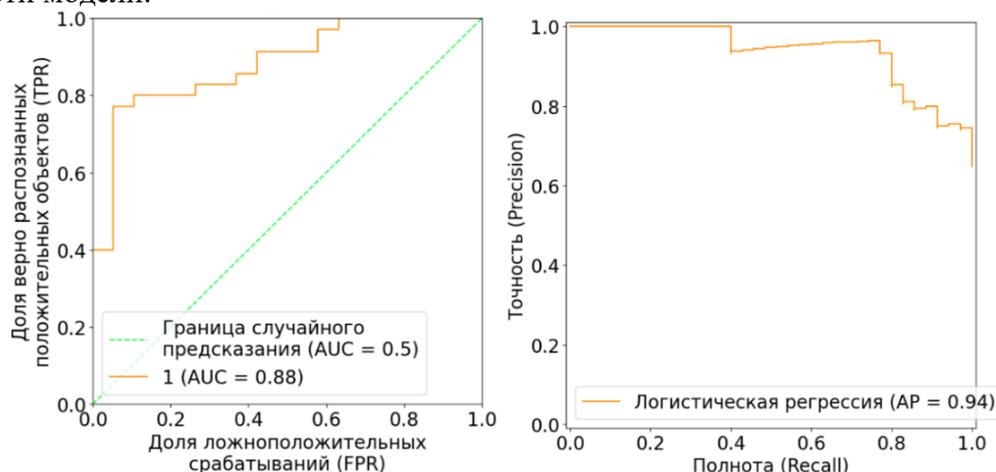


Рисунок 7 – Кривая ошибок первого и второго рода (ROC AUC curve) (слева) и кривая точности – полноты (Precision – Recall curve) (справа)

Разработанный способ позволяет выделить в спектрах костной ткани при пародонтите устойчивые биомаркеры: повышенные амплитуды линий $\sim 1318 \text{ см}^{-1}$ и $\sim 1744 \text{ см}^{-1}$ и снижение амплитуды линий КР, соответствующих минеральным компонентам. Классификатор на основе декомпозированных признаков демонстрирует приемлемую точность

(ROC AUC 0,88), что обосновывает его применение для экспресс-диагностики патологических изменений в костной ткани при пародонтите.

Модель оценивает состояние костной ткани при пародонтите по пяти ключевым спектральным признакам (амплитуды декомпозированных линий КР $\sim 852 \text{ см}^{-1}$, $\sim 1199 \text{ см}^{-1}$, $\sim 1318 \text{ см}^{-1}$, $\sim 1663 \text{ см}^{-1}$ и $\sim 1744 \text{ см}^{-1}$), которые после квантильного преобразования к нормальному распределению подаются на вход логистической регрессии, которая выдаёт откалиброванную вероятность p наличия патологических изменений, причём для бинарного решения обычно используется порог $p > 0,5$: образцы с $p \leq 0,5$ считаются нормой, а с $p > 0,5$ — признаком пародонтита.

В разделе 5.3 исследовано влияние закрытого кюретажа и профессиональной гигиены (через 6 недель после вмешательства) на спектральные характеристики эмали, дентина и цемента зубов пациентов с хроническим пародонтитом. Проанализированы 12 удалённых зубов (моляры и премоляры), в трёх зонах каждого: эмали, дентина и цемента. Всего получено 86 спектров эмали, 44 спектра дентина и 66 спектров цемента.

Сравнительный анализ усреднённых спектров КР двух групп эмали показал минимальные, но статистически значимые изменения в амплитудах фосфатных линий $\nu_1 \sim 957 \text{ см}^{-1}$ после лечения. В спектрах КР дентина незначительные сдвиги амплитуд линий β -карбоната ($\sim 1070 \text{ см}^{-1}$) и ν_1 ($\sim 957 \text{ см}^{-1}$).

Коэффициенты корреляции Пирсона между группами «до» и «после» составили 0,998 (эмаль), 0,998 (дентин) и 0,999 (цемент), что указывает на общее сходство спектров при локальных различиях.

Перестановочный тест показал, что все три ткани имеют значимые различия до и после лечения (эмаль: $p = 0,016$; дентин: $p = 0,039$; цемент: $p = 0,012$; мощность $\geq 0,99$ при минимуме 8–11 спектров), при $\alpha = 0,05$. Но выявленные изменения не достигли статистической значимости $\alpha = 0,01$.

Заключение

В настоящей диссертационной работе предложен и экспериментально подтверждён способ анализа спектров комбинационного рассеяния для диагностики признаков пародонтита и оценки качества дентинных материалов. Доказана возможность неинвазивного контроля состояния твёрдых тканей зубов и костной ткани.

1. Разработан способ анализа спектров комбинационного рассеяния биоматериалов, включающий этапы интерполяции, нормализации, коррекции базовой линии и сглаживания спектральных данных, а также оригинальный итеративный алгоритм декомпозиции спектров. Данный подход позволяет выделять элементарные спектральные компоненты даже при высоком уровне взаимного перекрытия линий до 75 % и шумовых помех, что существенно снижает ошибку интерпретации исходных данных.

Алгоритм декомпозиции отличается от других известных применением кластеризации и генерацией итоговой модели на основе множества спектров. Алгоритм позволяет учитывать сдвиг положения линий КР биологических материалов и других параметров, работу с асимметричными линиями. Позволяет корректно выявлять линии за счет определения итоговой модели на основе композиции множества отдельных моделей.

Алгоритм обеспечивает точность определения позиций линий не хуже $0,1 \text{ см}^{-1}$, относительную погрешность амплитуд и ширин $\leq 0,3 \%$ и детерминацию модели $R^2 \geq 0,995$.

2. Предложен способ оценки деминерализации дентинных материалов по спектрам комбинационного рассеяния, основанный на извлечении информативных признаков из декомпозированных спектральных линий в диапазоне $630\text{--}1450 \text{ см}^{-1}$ с применением логистической регрессии для вероятностной классификации образцов, позволяющей с точностью ROC AUC 1,0 (0,99–1,0, 95% ДИ) определять факт деминерализации.

3. Предложен способ экспресс-диагностики признаков пародонтита по спектрам эмали и костной ткани с применением логистической регрессии. Способ позволяет проводить диагностику на уровне ROC AUC 0,94 (0,89–0,98, 0,95 ДИ) для эмали и 0,88 (0,78–

0,96, 0,95 ДИ) для костной ткани. Способ демонстрирует чувствительность и специфичность на уровне 0,86 (0,76–0,94, 0,95 ДИ) и 0,86 (0,76–0,94, 0,95 ДИ).

Способ рекомендован к применению для диагностики пародонтита у пациентов возрастной группы 35-70 лет.

Перспективы дальнейшей разработки по теме диссертации предполагают интеграцию способа в портативные диагностические инструменты, работающие в режиме реального времени, что повысит их клиническую применимость для стоматологической и медицинской диагностики. Существует значительный потенциал для внедрения алгоритмов машинного обучения для повышения точности и эффективности спектрального анализа, что позволяет более точно выявлять патологические изменения в тканях.

Список основных работ, опубликованных автором по теме диссертации

Публикации в изданиях, рекомендуемых ВАК, индексируемых Scopus/Web of Science

1. Decomposition Method for Raman Spectra of Dentine / **О. О. Frolov** et al. // J of Biomedical Photonics & Eng. – 2024. – Vol. 10, No. 3. – 030303-1 – DOI 10.18287/JBRE24.10.030303.
2. Применение метода спектроскопии комбинационного рассеяния для оценки дентинных материалов в процессе их изготовления / П. Е. Тимченко, Е. В. Тимченко, Л. Т. Волова, **О. О. Фролов** // Оптический журнал. – 2021. – Т. 88, № 9. – С. 3-8. – DOI 10.17586/1023-5086-2021-88-09-03-08.
3. Хемометрический анализ биоимплантатов из твердой мозговой оболочки при их изготовлении / П. Е. Тимченко, Е. В. Тимченко, Л. Т. Волова, **О.О. Фролов** // Оптический журнал. – 2019. – Т. 86, № 1. – С. 13-20. – DOI 10.17586/1023-5086-2019-86-01-13-20.
4. Экспериментальные исследования гидроксиапатита методом спектроскопии комбинационного рассеяния / П. Е. Тимченко, Е. В. Тимченко, Е. В. Писарева, М.Ю. Власов, Л.Т. Волова, **О.О. Фролов**, А.Р. Калимуллина // Оптический журнал. – 2018. – Т. 85, № 3. – С. 12-18.
5. Optical Assessment of In-Office Teeth Whitening Influence on Tooth Enamel Spectral Composition In Vivo / Е. V. Timchenko, **О. О. Frolov**, О. А. Magsumova [et al.] // Optical Memory and Neural Networks. – 2022. – Vol. 31, No. 4. – P. 413-417. – DOI 10.3103/s1060992x22040099.
6. Raman spectroscopy for assessment of hard dental tissues in periodontitis treatment / Е. V. Timchenko, **О. О. Frolov**, Р. Е. Timchenko [et al.] // Diagnostics. – 2021. – Vol. 11, No. 9. – DOI 10.3390/diagnostics11091595.
7. Timchenko E., Timchenko P., Volova L., **Frolov O.**, Zibin M., Bazhutova I. Raman spectroscopy of changes in the tissues of teeth with periodontitis. Diagnostics. 2020. Т. 10. № 11.

Зарегистрированные результаты интеллектуальной деятельности

1. Патент № 2752035 С1 Российская Федерация, МПК А61К 35/32, А61К 6/80, А61L 27/02. Способ одновременного получения деминерализованного дентина и минерально-органического компонента из зубов: № 2021100652: заявл. 13.01.2021: опубл. 22.07.2021 / Л. Т. Волова, **О. О. Фролов**, П. Е. Тимченко [и др.].