

На правах рукописи



Вражнов Денис Александрович

**ДИАГНОСТИКА ГЛИОМЫ ГОЛОВНОГО МОЗГА
НА ОСНОВЕ ТЕРАГЕРЦОВОЙ СПЕКТРОСКОПИИ КРОВИ
И МЕТОДОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

1.3.21. Медицинская физика

Автореферат

диссертации на соискание учёной степени
кандидата физико-математических наук

Томск – 2025

Работа выполнена в федеральном государственном автономном образовательном учреждении высшего образования «Национальный исследовательский Томский государственный университет».

Научный руководитель: доктор физико-математических наук, профессор
Кистенев Юрий Владимирович

Официальные оппоненты:

Зайцев Владимир Юрьевич, член-корреспондент РАН, доктор физико-математических наук, старший научный сотрудник, Федеральное государственное бюджетное научное учреждение «Федеральный исследовательский центр Институт прикладной физики им. А.В. Гапонова-Грехова Российской академии наук», лаборатория волновых методов исследования структурно-неоднородных сред, заведующий лабораторией

Дунаев Андрей Валерьевич, доктор технических наук, доцент, федеральное государственное бюджетное образовательное учреждение высшего образования «Орловский государственный университет имени И.С. Тургенева», научно-технологический центр биомедицинской фотоники, ведущий научный сотрудник

Ежов Дмитрий Михайлович, кандидат физико-математических наук, федеральное государственное бюджетное образовательное учреждение высшего образования «Национальный исследовательский Томский государственный университет», кафедра оптики и спектроскопии, доцент

Защита состоится 23 октября 2025 г. в 14⁰⁰ часов на заседании диссертационного совета «НИ ТГУ.1.3.03», созданного на базе физического факультета федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский Томский государственный университет», по адресу: 634050, г. Томск, пр. Ленина 36 (учебный корпус № 13 ТГУ (НИИ ББ), аудитория 108).

С диссертацией можно ознакомиться в Научной библиотеке и на официальном сайте федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский Томский государственный университет» www.tsu.ru.

Материалы по защите диссертации размещены на официальном сайте ТГУ: <https://dissertations.tsu.ru/PublicApplications/Details/2e0aec84-ad1c-4160-9d49-009dd6b5cfc7>

Автореферат разослан « ____ » сентября 2025 г.

Ученый секретарь
диссертационного совета
кандидат физико-математических наук



Князькова Анастасия Игоревна

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность исследования

Всемирная организация здравоохранения (ВОЗ) считает борьбу с раком одной из приоритетных задач [1, 2]. Наиболее часто встречающимися злокачественными опухолями центральной нервной системы являются глиомы, и, в частности, мультиформная глиобластома (МГБ), на долю которой приходится 14,3% всех опухолей и 49,1% злокачественных опухолей [3]. МГБ является наиболее агрессивным, инвазивным и недифференцированным типом опухолей (IV класс по классификации ВОЗ [4, 5]).

Одной из причин высокой вероятности неблагоприятного исхода МГБ является поздняя диагностика, поскольку большинство существующих методов медицинской визуализации, таких как магнитно-резонансная томография [6, 7, 8] и компьютерная томография [9], неэффективны для диагностики опухолей небольшого размера. Альтернативой являются диагностические методы, основанные на детекции молекулярных маркеров глиом – продуктов метаболизма опухолей.

Оптические методы, такие как флуоресцентная визуализация [10], многофотонная микроскопия [11], фотоакустическая визуализация [12], спектроскопия комбинационного рассеяния [13, 14], абсорбционная [15, 16] и терагерцовая (ТГц) [17, 18, 19] спектроскопия широко используются для обнаружения уникальных спектральных характеристик биологических тканей, включая молекулярные биомаркеры новообразований, при хирургической резекции пораженной ткани. Такие методы чувствительны к химическому составу ткани, но имеют малую глубину проникновения для неинвазивного обнаружения глиомы *in vivo*.

Применение оптических методов для анализа биологических жидкостей, таких как кровь, моча, спинномозговая жидкость, слюна и др. характеризуется неинвазивностью/малоинвазивностью и меньшей сложностью манипуляции по сравнению со стандартной биопсией [20, 20]. Было показано, что кровь, как и другие жидкости организма, содержит циркулирующие опухолевые клетки и молекулярные маркеры – нуклеиновые кислоты, внеклеточные везикулы, белки и метаболиты, анализ которых позволяет получить информацию о развитии рака [22]. Таким образом контроль молекулярных маркеров глиомы в крови может стать эквивалентным стандартной биопсии. Важно, что подобные молекулярные маркеры могут появляться в биологических жидкостях на начальных этапах онкогенеза, открывая путь к ранней диагностике глиомы [23, 24].

Терагерцовая (ТГц) спектроскопия жидкостей является относительно новой областью по отношению к спектроскопии в инфракрасном, ультрафиолетовом и видимом диапазонах. В настоящее время наиболее распространена ТГц спектроскопия во временной области (ТГц СВО), основанная на генерации зондирующих ТГц импульсов при помощи GaAs антенны с накачкой сериями фемтосекундных лазерных импульсов и синхронной регистрацией отраженного или поглощенного сигнала второй GaAs антенной. Для анализа данных ТГц СВО

применяется переход в частотную область при помощи преобразования Фурье, где они принимают вид зависимости интенсивности сигнала (либо таких характеристик, как коэффициент поглощения, показатель преломления) от частоты. Такие спектры имеют высокую размерность (сотни частот) и корреляцию в широких диапазонах частот, что, в свою очередь, приводит к высокой схожести спектров при оценках на основе евклидовой метрики [25]. Это существенно затрудняет или делает невозможным их анализ с помощью методов математической статистики. Альтернативой является паттерный подход, в котором анализируются выборочные данные на совокупности частот – спектральные сигнатуры.

Машинное обучение (МО) является мощным инструментом для построения предиктивных моделей и обработки спектральных сигнатур, включая дифференциацию наборов данных с высокой схожестью [26] за счет генерации метрик, учитывающих особенности конкретных наборов данных. Современные методы МО, например, методов глубокого обучения (ГО) создают нелинейные модели данных с миллиардами настраиваемых гиперпараметров, позволяющие обнаружить латентные связи между входными и выходными данными [27]. Недостатком таких моделей данных является слабая интерпретируемость – фактически, они представляют собой «черный ящик», не позволяющий оценить причину, по которой было сделано то или иное предсказание [28].

В области биомедицины особый интерес представляют интерпретируемые модели данных, поскольку они позволяют показать взаимосвязь метаболических путей, метаболитов опухолей и течение болезни [29]. Таким образом, разработка методов МО, генерирующих интерпретируемые предиктивные модели на основе данных ТГц оптической биопсии является актуальной задачей.

Степень разработанности темы исследования

На момент формулировки цели диссертации в области методов диагностики глиомы головного мозга на основе терагерцовой спектроскопии крови и методов искусственного интеллекта состояние проблемы выглядело следующим образом.

В литературе были представлены работы по основным метаболитам МГБ в тканях человека и мелких лабораторных животных (МЛЖ), полученные методами магнитной резонансной спектроскопии (МРС), масс-спектрометрии и хроматографии. В работе [30] были изучены инфракрасные (ИК) спектры поглощения высушенных капель плазмы крови у пациентов с опухолями головного мозга и обнаружены наиболее информативные области 1200–950 см⁻¹, содержащие линии поглощения связей Р–О, С–О, простых и сложных эфирных связей С–ОО и Р–ОО, присутствующих в липидах, фосфорилированных белках, углеводах, креатинфосфокиназах и других соединениях.

В области развития методов МО для анализа ТГц спектров был представлен ряд работ посвященных анализу твердых образцов [31] и их двумерной визуализации. Отсутствовали работы по анализу проб сыворотки и плазмы крови больных МГБ методами ТГц спектроскопии. Слабо были представлены методы машинного и глубокого обучения создания предиктивных моделей, на основе анализа данных ТГц спектроскопии образцов биологических жидкостей. Также

отсутствовали сведения о наиболее характерных ТГц диапазонах для метаболитов МГБ в крови.

Особенностью данной диссертационной работы заключается в использовании методов МО для построения предиктивных диагностических моделей МГБ на основе анализа спектральных сигнатур плазмы и сыворотки крови в ТГц диапазоне спектра. Такой подход требует решения следующих проблем: создание предиктивных моделей на основе различных конвейеров машинного обучения и выбор оптимальных с точки зрения метрик качества, выделение информативных диапазонов или отдельных частот, верификация полученных результатов.

Цель и задачи диссертационной работы

Целью диссертационной работы является разработка и верификация методов диагностики глиомы головного мозга мелких лабораторных животных и человека на основе анализа данных ТГц спектроскопии образцов сыворотки и плазмы крови с использованием методов искусственного интеллекта.

В соответствии с поставленной целью решались следующие задачи:

1. Подготовить базы данных ТГц спектроскопии образцов плазмы крови здоровых МЛЖ, МЛЖ с ксенографтовой моделью глиомы на основе клеток U87, МЛЖ с травматическим поражением мозга.
2. Подготовить базы данных ТГц спектроскопии образцов сыворотки крови пациентов с глиомой, пациентов после краниэктомии, здоровых доноров.
3. Провести разведывательный анализ на подготовленных базах данных для предварительной оценки делимости данных.
4. Разработать базовые конвейеры машинного обучения для построения и верификации предиктивных моделей с возможностью селекции информативных признаков на подготовленных базах данных.
5. Подобрать оптимальные конфигурации конвейеров машинного обучения и соответствующие параметры алгоритмов.
6. Получить оценки метрик качества для построенных предиктивных моделей.

Объектами исследования в данной диссертационной работе являлись образцы сыворотки крови следующих групп МЛЖ: здоровых, с травматическими поражениями головного мозга и с ксенографтовой моделью глиомы на основе клеток U87; образцы плазмы крови следующих групп добровольцев: здоровых, пациентов после краниэктомии и с глиомой головного мозга.

Научная новизна исследования

Впервые были предложены предиктивные модели для оценки динамики глиомы и дифференциации образцов сыворотки крови МЛЖ с ксенографтовой моделью глиомы на основе клеток U87 и контрольной группы на основе данных ТГц спектроскопии плазмы крови МЛЖ. Получены значения соответствующих информативных ТГц частот в динамике развития МГБ.

Впервые разработаны предиктивные модели для попарной дифференциации групп с глиомой, здоровых доноров и пациентов после краниэктомии на основе данных ТГц спектроскопии сыворотки крови. Также получены данные об информативных ТГц частотах для дифференции соответствующих групп.

Впервые разработаны предиктивные модели для дифференциации образцов сыворотки крови МЛЖ с ксенографтовой моделью глиомы на основе клеток U87 и группы с травматическим поражением головного мозга. Впервые выявлены информативные частоты ТГц для соответствующих образцов плазмы крови МЛЖ.

Научная и практическая значимость работы

Результаты работы позволяют получить общие представления об изменениях оптических характеристик сыворотки и плазмы крови пациентов в динамике развития глиомы головного мозга и при травматических поражениях головного мозга. Это важно при применении оптических методов для малоинвазивной и интраоперационной диагностики глиомы и травматических поражений мозга.

Методология и методы исследования

Спектральные характеристики (коэффициенты поглощения и показатели преломления при прохождении ТГц излучения через образец) определялись стандартными методами ТГц спектроскопии с временным разрешением. Обработка и анализ получаемых экспериментальных данных осуществлялись с применением интерпретируемых методов машинного обучения.

Положения, выносимые на защиту

1. Предиктивная модель, построенная при помощи конвейера методов машинного обучения, предназначенного для анализа данных ТГц спектроскопии сыворотки крови мелких лабораторных животных в диапазоне 0,2-1,7 ТГц, включающего предварительную обработку данных фильтром Савицкого-Голея и методом изолированного леса, уменьшение размерности методом главных компонент, классификацию методом опорных векторов с линейным ядром позволяет дифференцировать образцы сыворотки крови мелких лабораторных животных с ксенографтовой моделью глиомы на основе клеток U87 и контрольной группы на третьей неделе эксперимента со средними значениями чувствительности, специфичности, точности и площади под ROC-кривой не менее 90% на использованной экспериментальной выборке.

2. Предиктивная модель, построенная при помощи конвейера методов машинного обучения с подобранными оптимальными значениями гиперпараметров, включающей предварительную обработку данных фильтром Савицкого-Голея с шириной окна 5 и степенью полинома 2, классификацию методом стохастического градиентного бустинга с глубиной деревьев 2, автоматической балансировкой классов, логарифмической функцией потерь, позволяет попарно дифференцировать группы с глиомой, здоровых доноров и пациентов после краниэктомии со средним значением площади под ROC-кривой не менее 90% на использованной экспериментальной выборке.

3. Предиктивная модель, построенная при помощи конвейера методов машинного обучения, включающего фильтр с Савицкого-Голея с шириной окна 25 и степенью полинома 2, нормализацию данных путем вычитания среднего и деления на дисперсию, ансамблевый классификатор на основе комбинации метода опорных векторов с линейным ядром, стохастического градиентного бустинга и случайного леса с подобранными оптимальными гиперпараметрами, позволяет дифференцировать группу мелких лабораторных животных с травматическим

повреждением головного мозга от группы мелких лабораторных животных с ксенографтовой моделью глиомы на основе клеток U87, со средними значениями чувствительности, специфичности и точности равными 95%, 100% и 97%, соответственно, на использованной экспериментальной выборке.

Таблица 1 – Оптимальные параметры методов

| Оптимальные параметры метода случайный лес | Оптимальные параметры метода стохастический градиентный бустинг |
|---|---|
| Функция качества: критерий Джини; Количество деревьев в лесу: 100; Минимальное количество образцов, необходимое для разбиения внутреннего узла: 6; Минимальное количество образцов, необходимое для нахождения в узле листа: 3; Использовалась балансировка подвыборки и бутстрап | Функция качества: логистическая функция потерь; Максимальное количество деревьев: 300, Глубина деревьев: 2, Скорость обучения: 0.03, Количество итераций при вычислении узла листа: 10, Максимальное количество листов: 16; Использовалась балансировка классов |

4. Частоты 0,37, 0,40, 0,55, 0,60, 0,70 и 0,90 ТГц являются наиболее информативными для дифференцирования группы мелких лабораторных животных с травматическим повреждением головного мозга и группы мелких лабораторных животных с ксенографтовой моделью глиомы на основе клеток U87.

Личный вклад автора диссертации

Автор лично участвовал в обработке полученных данных, анализе и обсуждении полученных результатов, в написании научных статей и апробации результатов исследований на конференциях, симпозиумах. Определение основного направления диссертационной работы, формулировка темы, постановка задач, обсуждение результатов, обсуждение текста диссертационной работы, ее основных положений и выводов, осуществлялась совместно с научным руководителем диссертационной работы: заведующим лабораторией лазерного молекулярного имиджинга и машинного обучения федерального государственного бюджетного образовательного учреждения высшего образования «Национальный исследовательский Томский государственный университет», профессором, доктором физико-математических наук, Ю.В. Кистеневым.

Степень достоверности полученных результатов

Достоверность полученных результатов диссертационной работы подтверждается использованием апробированных экспериментальных моделей и методов данных, измерений, согласованностью с результатами независимых исследований других авторов.

Апробация результатов исследования

Основные положения диссертации докладывались и обсуждались на следующих конференциях: международная конференция Advanced Laser Technologies ALT`21, (Москва, 06-10 октября 2021), международная конференция 46 International Conference on Infrared, Millimeter and Terahertz Waves «IRMMW-THz

2021» (29 августа – 3 сентября 3, 2021, Ченгду, Китай), V Международная конференция «Terahertz and Microwave Radiation: Generation, Detection and Applications – TERA 2020» (Томск, 24-26 августа 2020), Saratov fall meeting 2020 (Саратов, 28 сентября - 1 октября 2020), международная конференция International Conference Laser Optics 2020 «ICLO 2020» (Санкт Петербург, 2-6 ноября 2020), XV Международная конференция по импульсным лазерам и применениям лазеров, AMPL-2021(Томск, 12-17 сентября 2021), Saratov Fall Meeting SFM21, 9th International Symposium Optics and Biophotonics (Саратов, 27 сентября -1 октября 2021), Saratov Fall Meeting SFM22, 10th International Symposium Optics and Biophotonics (Саратов, 26-30 сентября 2022), 5-я Конференция с международным участием «Терагерцевое и микроволновое излучение: генерация, детектирование и приложения» - ТЕРА-2023 (Москва, 27 февраля - 2 марта 2023), XX Symposium on High Resolution Molecular Spectroscopy HighRus-2023 (Иркутск, 03 июля 2023 – 07 июля 2023).

Публикации

Основные результаты диссертационной работы опубликованы в 13 работах, в том числе в 6 статьях в журналах, включенных в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук (все статьи опубликованы в зарубежных научных журналах, входящих в Web of Science и / или Scopus), 6 публикациях в сборниках материалов конференций, представленных в изданиях, входящих в Web of Science; 1 монографии.

Конкурсная поддержка работы

Исследования были поддержаны грантом РФФИ (№19-52-55004 «Ранняя диагностика глиомы головного мозга на основе терагерцовой и инфракрасной спектроскопии, спектроскопии комбинационного рассеяния света»), грантом Министерства науки и высшего образования Российской Федерации (№ 075-15-2024-557 от 25.04.2024 г. «Развитие фундаментальных основ новых технологий обеспечения безопасности жизнедеятельности на основе интеграции мультимодальной радиоволновой и оптической дистанционной сенсорики, и искусственного интеллекта»), оказана государственная поддержка в рамках программы 5-100 в целях повышения их конкурентной способности среди ведущих мировых научно-образовательных центров, в рамках программы развития Томского государственного университета (Приоритет-2030) и в рамках конкурса компании Хуавэй для обучающихся Томского государственного университета.

Структура и объем диссертации

Диссертация состоит из введения, пяти глав, заключения и списка использованных источников и литературы из 198 наименований. Работа изложена на 168 страницах, включает 69 рисунков и 14 таблиц.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении представлена общая характеристика работы, актуальность темы исследования, цели и задачи настоящей работы; описаны объекты исследования;

сформулированы защищаемые положения; приведена структура и краткое содержание диссертации.

В первой главе представлено описание структуры и строения головного мозга. Приведена информация о злокачественных новообразованиях головного мозга, классификации глиом. Рассмотрены молекулярные маркеры глиом, существующие биологические модели глиом и инструментальные методы их диагностики. Изложена информация о ТГц спектроскопии крови. Представлены методы и подходы к анализу биомедицинских данных на основе искусственного интеллекта. В частности, приведен базовый конвейер машинного обучения для построения предиктивных моделей в задаче медицинской диагностики и рассмотрены его этапы.

Во второй главе приведено обоснование выбора модели глиомы МЛЖ, образцов крови в качестве объекта исследования, метода ТГц СВЧ в качестве инструментального, а методов искусственного интеллекта в качестве аналитических методов анализа.

В данной работе исследование проводилось на свободных от специфических патогенов мышцах-самцах линии SCID (SHO-PrkdcscidHrhr) в возрасте 6-7 недель. Использовалась ортотопическая ксенографтовая модель глиомы на основе линии клеток U87. Для оценки размеров опухоли мышам под наркозом проводили прижизненную МРТ, перед операцией и на 7, 14 и 21-й дни после инъекции.

В динамике развития МГБ объем опухоли составлял $2,6 \pm 0,4$, $10,6 \pm 1,8$ и $89,6 \pm 11,5$ мм³ на 7, 14 и 21 день после введения опухолевых клеток, соответственно. На 7-й, 14-й и 21-й день после инъекции животным проводили декапитацию, после чего собирали образцы крови и центрифугировали в течение 15 мин при 1000 g. Далее, сыворотку помещали в отдельные пробирки и замораживали при -80° С. Объем сыворотки составил около 200 мкл. Мозг каждой мыши помещали в индивидуальную пробирку целиком, после этого быстро замораживали жидким азотом. Перед измерениями ТГц спектров образцы сыворотки предварительно размораживали.

Номера групп и животных приведены в Таблице 2. Количество животных в каждой группе указано в скобках.

Таблица 2 – Номера групп и животных

| № | Номер группы неделя (количество) | Номер животных |
|---|----------------------------------|--------------------|
| 1 | U87-1 неделя (5) | 2, 6, 8, 9, 10 |
| 2 | Контроль к U87-1 неделя (5) | 11, 12, 16, 17, 18 |
| 3 | U87-2 неделя (10) | 21-30 |
| 4 | Контроль к U87-2 неделя (10) | 31-40 |
| 5 | U87-3 неделя (7) | 42-48 |
| 6 | Контроль к U87-3 неделя (10) | 51-60 |
| 7 | U87-4 неделя (4) | 61-64 |
| 8 | Контроль к U87-4 неделя (5) | 72, 75, 77, 78, 80 |

В рамках диссертационной работы были исследованы образцы плазмы крови пациентов, представленные Е. В. Ступак, врачом-нейрохирургом, к.м.н.,

Новосибирского научно–исследовательского института травматологии и ортопедии им. Я. Л. Цивьяна.

Образцы периферической крови собирали в вакутейнеры с ЭДТА (Sarstedt AG&Co. KG, Нюмбрехт, Германия). Плазму отделяли центрифугированием при 2800 g в течение 15 минут при + 4 С. Образцы плазмы крови замораживали и хранили при –80 С. Перед измерениями сыворотки предварительно размораживали. Характеристика и количество образцов представлены в таблице 3.

Таблица 3 – Характеристика исследованных образцов плазмы крови добровольцев

| Образец | Количество | Характеристика |
|-------------------------|------------|---|
| Плазма крови (группа 1) | 14 | Гистологически верифицирован диагноз глиомы 2-4 стадии согласно классификации ВОЗ |
| Плазма крови (группа 2) | 13 | Пациенты с дефектами черепа после краниэктомии, без онкологии |
| Плазма крови (группа 3) | 11 | Доноры, практически здоровые лица |

В данной работе для снятия спектров использовался коммерческий ТГц СВО спектрометр T-SPEC (EKSPLA, Вильнюс, Литва, <https://www.teravil.lt/t-spec.php>) со следующими техническими характеристиками:

- Спектральное разрешение до 10 ГГц.
- Две линии задержки, быстрая и медленная. Первая позволяет получать спектры в реальном времени со скоростью до 10 спектров/сек.
- Динамический диапазон превышает 90 дБ при 0.4 ТГц.

Все измерения проводились при комнатной температуре ($21 \pm 1^\circ\text{C}$).

В третьей главе описаны результаты анализа данных ТГц спектроскопии сыворотки крови мышей методами машинного обучения для построения предиктивных моделей для диагностики МГБ. Приведено описание групп и анализируемых данных. Представлен модифицированный конвейер машинного обучения на основе метода опорных векторов с линейным ядром и метода главных компонент, позволяющий выделить информативные частоты для дифференции исследуемых групп в динамике развития глиомы. На основе проведенного анализа дана оценка связи ТГц-поглощения сыворотки крови и размера опухоли, а также способов повышения метрик качества полученных предиктивных моделей.

Описание исследуемых групп и данных

ТГц спектры пропускания образцов сыворотки крови мышей измерялись в диапазоне от 0,2 до 1,7 ТГц с помощью спектрометра T-SPEC (EKSPLA, Вильнюс, Литва). В качестве референсного использовался ТГц-сигнал, зарегистрированный от пустой кюветы.

Конвейер машинного обучения для диагностики U87 глиобластомы у мышей

Для построения предиктивных моделей использовался следующий конвейер МО (рисунок 1). Конвейер был реализован на языке Python версии 3.8.

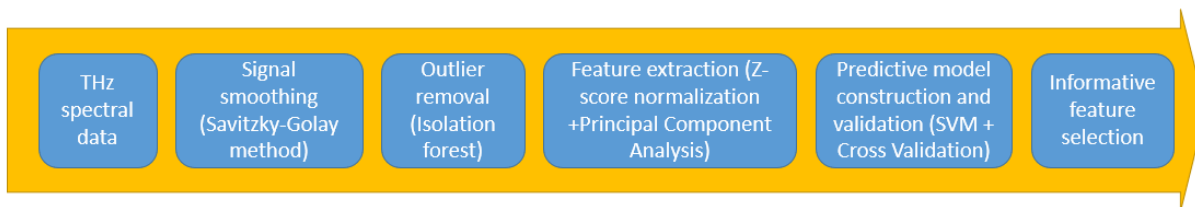


Рисунок 1 – Реализованный конвейер МО для анализа спектральных данных ТГц сыворотки крови мышей

На начальном этапе ТГц спектры были сглажены при помощи фильтра Савицкого-Голея со следующими параметрами: ширина окна фильтрации 15, степень сглаживающего полинома 2. Для удаления выбросов использовался метод изолированных лесов (ИЛ) с автоматическим подбором параметров. Перед применением метода главных компонент (МГК) данные были стандартизованы. При уменьшении размерности были оставлены 10 главных компонент, чтобы покрыть более 95 % объясненной дисперсии в данных. Для выбора информативных признаков были проанализированы строки матрицы нагрузок. Разделимость данных анализировалась с помощью метода опорных векторов (МОВ) с линейным ядром. Валидация предиктивных моделей осуществлялась при помощи k-кратной перекрестной проверки (ПП). В качестве метрик качества работы классификаторы были использован анализ кривой работы приемника (ROC) и площади под кривой (AUC).

На последнем этапе был разработан алгоритм последовательного выделения признаков, объединяющий МГК и линейный МОВ (рисунок 2).

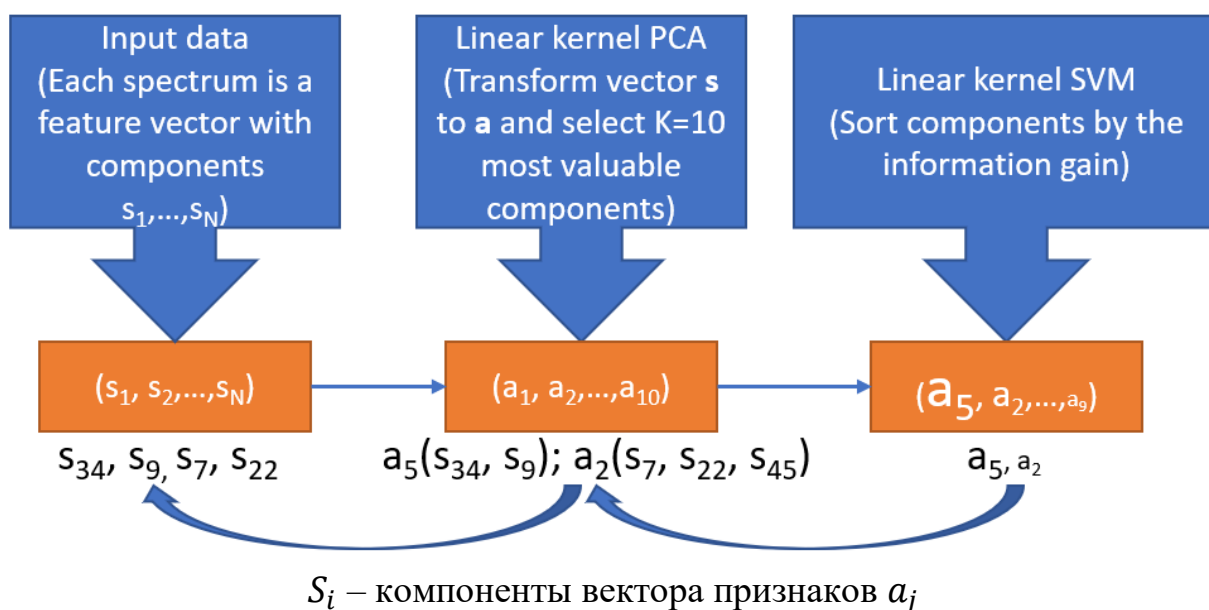


Рисунок 2 – Алгоритм последовательного выделения признаков на основе МГК и МОВ с линейным ядром

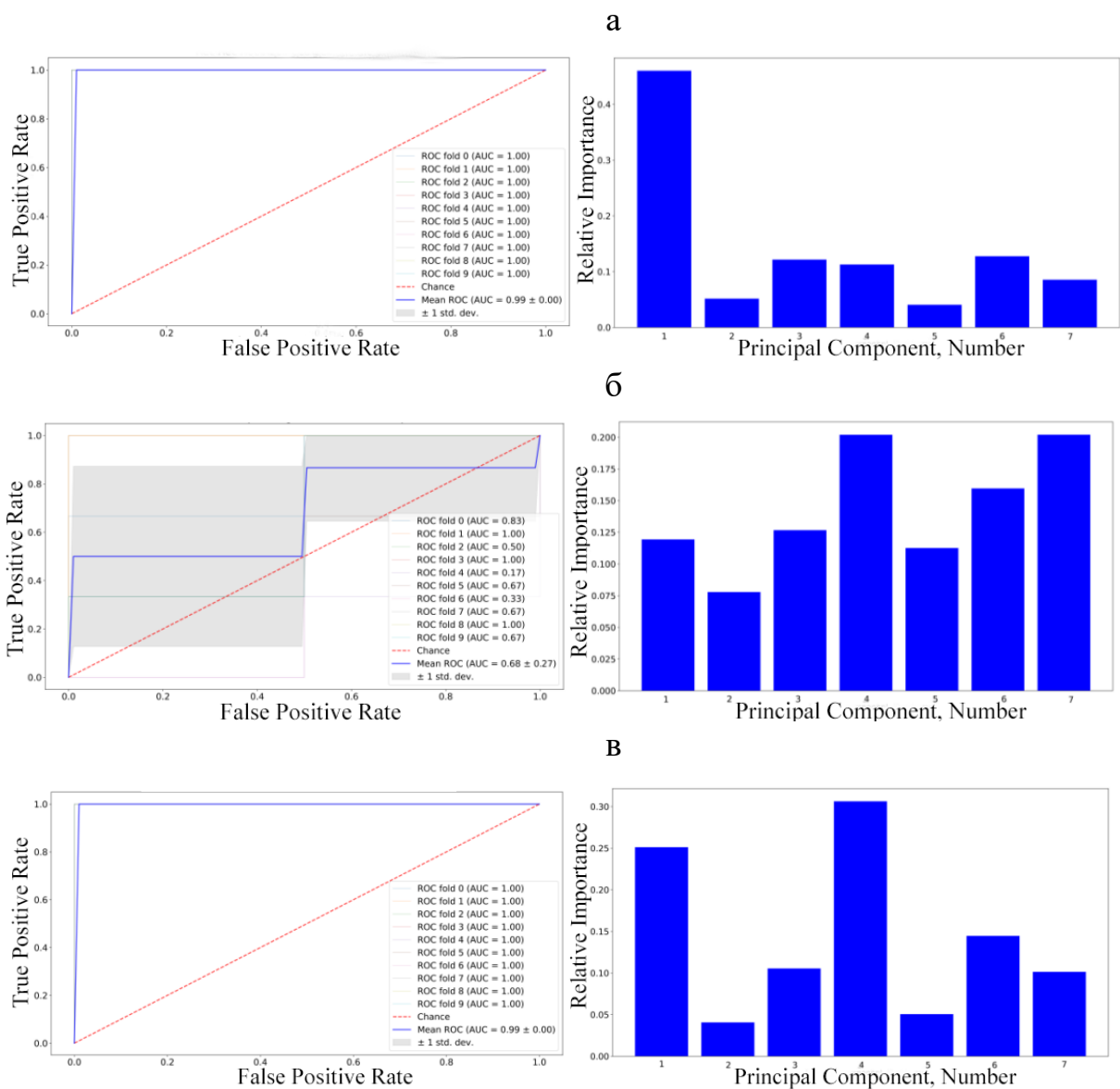
Результаты анализа ТГц спектров сыворотки крови мышей с U87 глиобластомой методами машинного обучения

Значительный рост опухоли наблюдался с 14 по 21 день после введения клеток U87 в подкорковые структуры мозга. Рост опухоли вызывает существенные изменение состава крови [23]. Показатели качества работы МОВ для различных

групп приведены в таблице 4 и на рисунке 4. Также, на рисунке 4 даны наиболее информативные главные компоненты (ГК).

Таблица 4 – Усредненные показатели качества работы классификатора МОВ с линейным ядром для контрольной и опухолевой групп для 1-й, 2-й и 3-й недель эксперимента с удалением и без удаления выбросов

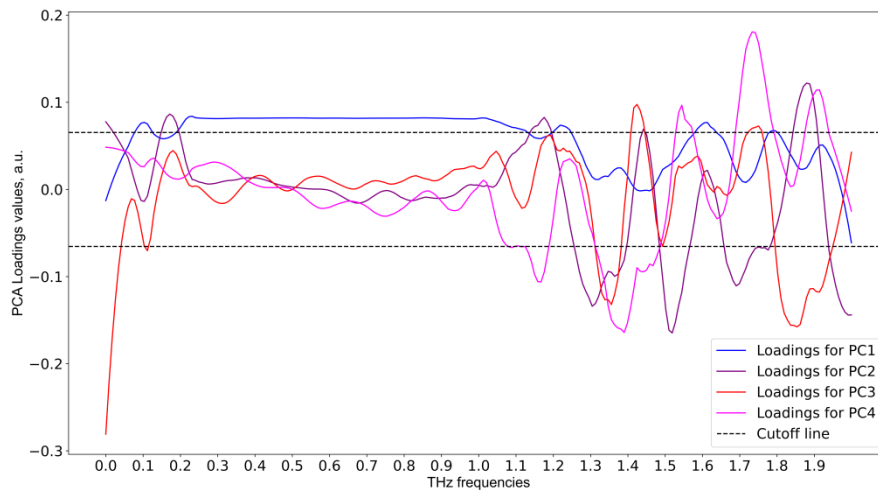
| Неделя | Выбросы удалены | Чувствительность | Специфичность | Точность |
|--------|-----------------|------------------|---------------|----------|
| 1 | Нет | 0,43 | 0,69 | 0,57 |
| 1 | Да | 1 | 1 | 1 |
| 2 | Нет | 0,43 | 0,57 | 0,50 |
| 2 | Да | 0,65 | 0,83 | 0,76 |
| 3 | Нет | 0,7 | 0,88 | 0,82 |
| 3 | Да | 0,85 | 1 | 0,93 |



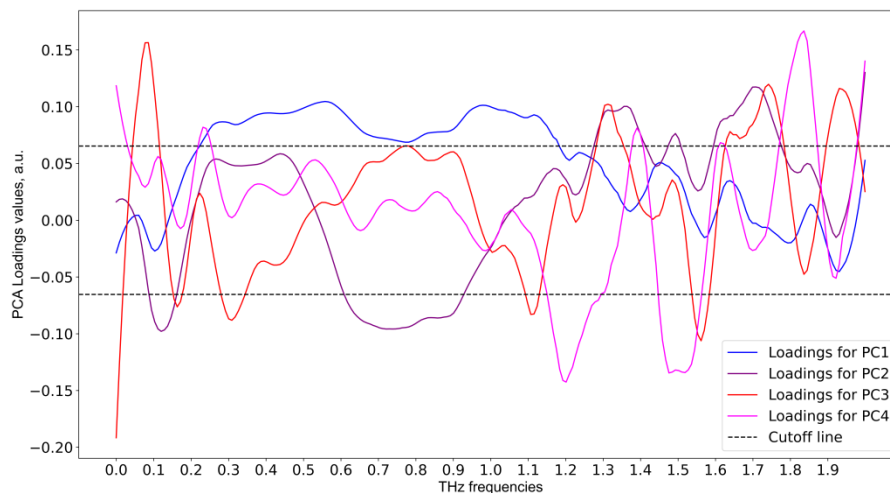
а – 1-я неделя эксперимента, б – 2-я неделя, в – 3-я неделя эксперимента
Рисунок 3 – ROC-AUC анализ и информативные ГК, найденные при помощи метода МОВ с линейным ядром

Из рисунка 3 видно, что по данным классификатора МОВ информативной главной компонентой для 1-й недели является ГК1, для 2-й недели - ГК4 и ГК7, для 3-й недели - ГК4 и ГК1. Следует отметить, что точность МОВ и значения ROC-AUC для 2-й недели низкие, поэтому данные модели были исключены из анализа. Далее, был проведен анализ матрицы нагрузок (рисунок 4). Прерывистые линии отсечения на рисунке 5 указывают на пороговые значения: пики выше положительной линии отсечения и ниже отрицательной линии отсечения считаются информативными.

а



б



а – 1-я неделя, б – 3-я неделя эксперимента

Рисунок 4 – Анализ матрицы нагрузок в методе МГК для выделения информативных признаков по данным ТГц спектроскопии сыворотки крови

Как видно из рисунка 4, для первой недели информативные частоты составляют 0,11; 0,22; 1,2; 1,6 ТГц для РС1. Для третьей недели информативные частоты составляют: 0,56; 1,0 ТГц для РС1, 0,22; 1,2, 1,48, 1,52; 1,82 ТГц для РС4.

В четвертой главе представлены результаты анализа данных ТГц спектроскопии плазмы крови методами машинного обучения для диагностики МГБ у добровольцев. Дано описание групп и анализируемых данных. Приведена оценка

однородности групп на основе методов кластеризации. Представлен модифицированный конвейер машинного обучения, указаны оптимальные значения гиперпараметров для его этапов и метрики качества предиктивных моделей для попарной дифференциации групп с глиомой, здоровых добровольцев и пациентов с дефектами черепа после краниэктомии. Также приведены наиболее информативные ТГц частоты для построенных предиктивных моделей.

Описание исследуемых групп и данных

В рамках поставленной задачи было проведено два эксперимента: 1 – оценивалась разделимость групп 1 и 3 (плазмы крови пациентов с МГБ и доноров) и 2 – разделимость групп 1 и 2 (плазмы крови пациентов с МГБ и пациентов с дефектами черепа после краниэктомии).

Конвейер машинного обучения для диагностики глиобластомы у пациентов

Анализ экспериментальных ТГц-данных проводился в соответствии с конвейером МО, представленным на рисунке 5. На первом этапе данные были разделены на две группы: пациенты с глиомой против здоровых доноров (эксперимент 1) и пациенты с глиомой против пациентов без глиомы (эксперимент 2). Конвейер вычислительной обработки включал сглаживание ТГц-спектров фильтром Савицкого-Голея, снижение размерности и визуализацию методами МГК и t-SNE, проверку разделимости данных с помощью МОВ с линейным ядром, случайного леса (СЛ), стохастического градиентного бустинга (СГБ). Выбор последних был ограничен требованием интерпретируемости результатов за счет информативного выделения признаков. Экспериментальным путем были определены оптимальные значения перплексии и выбрана косинусная метрика. Для валидации предиктивных моделей применялась k-кратная ПП, ROC-AUC анализ, а также значения чувствительности и специфичности.

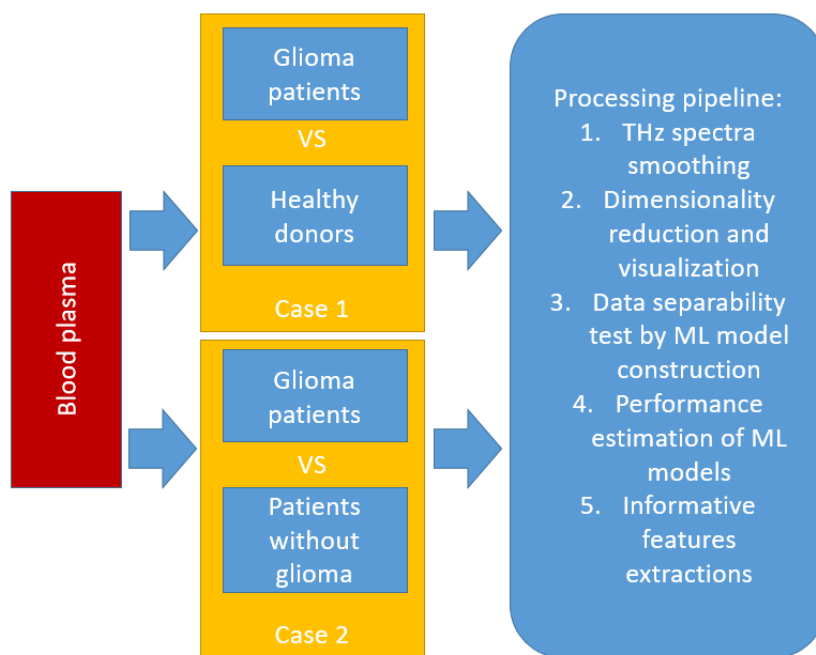


Рисунок 5 – МО конвейер для обработки экспериментальных ТГц данных добровольцев

Результаты анализа ТГц спектров сыворотки крови добровольцев методами машинного обучения

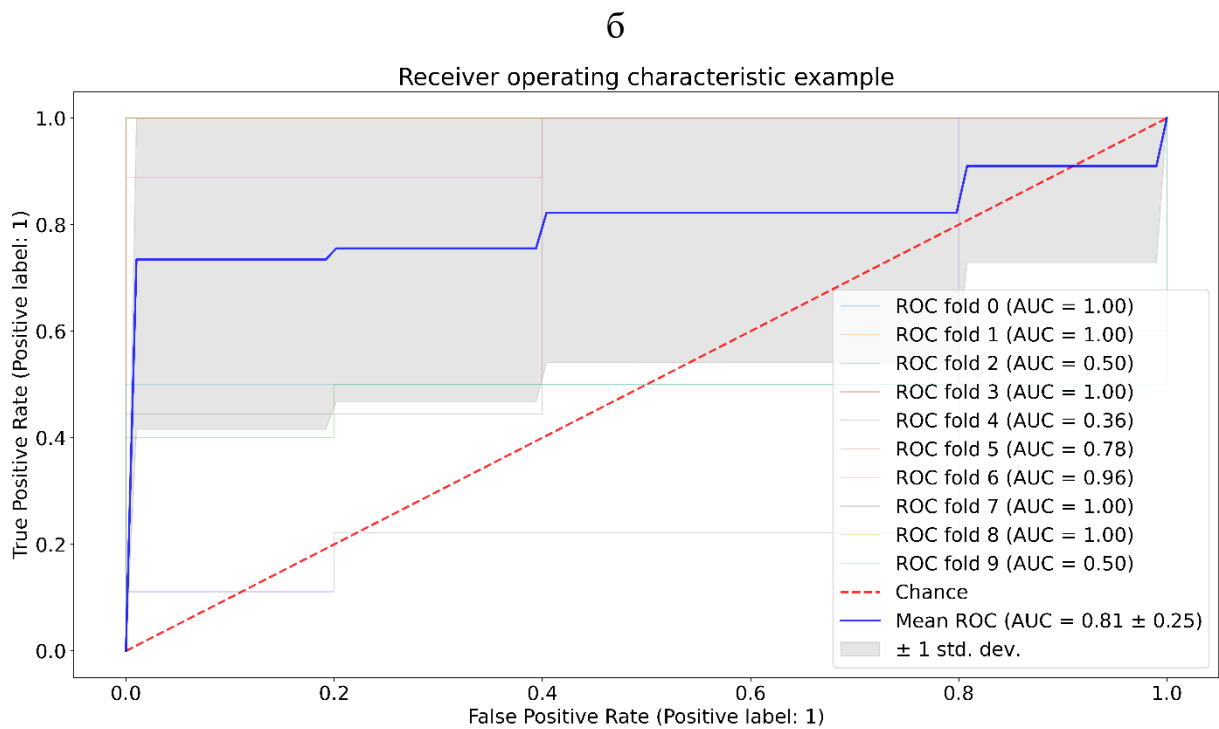
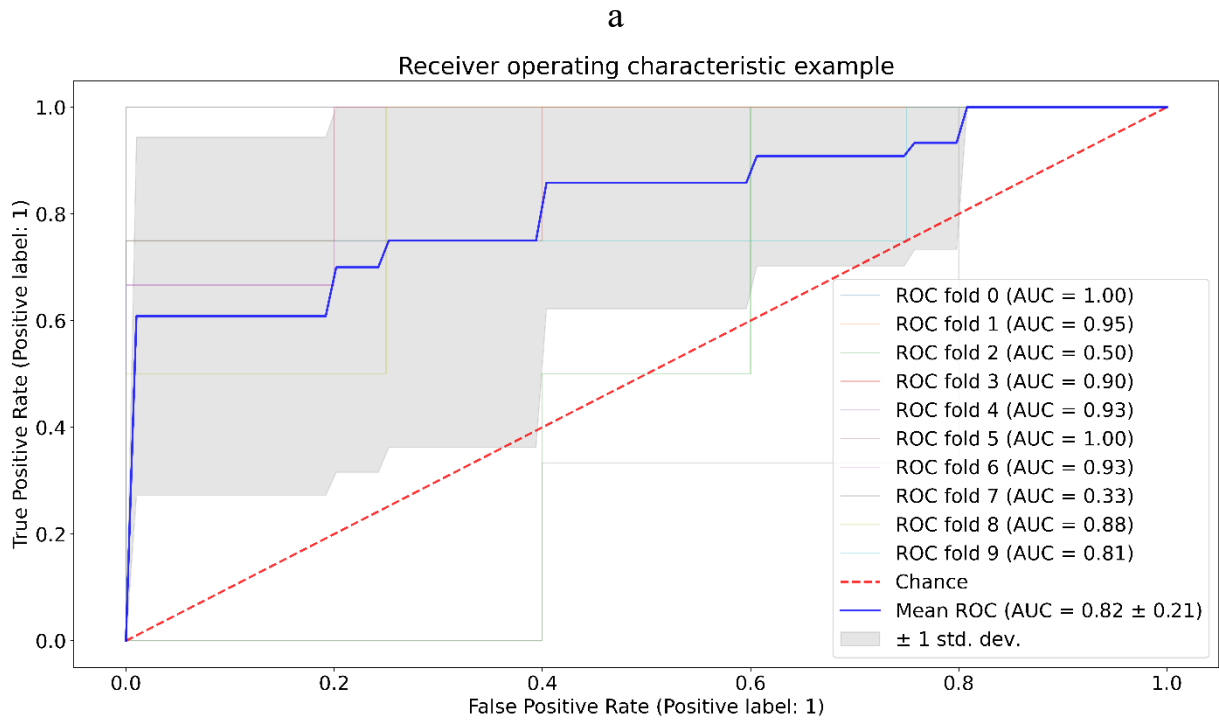
На первом этапе была проведена оценка оптимального значения размера окна для фильтра Савицкого-Голея по средним значениям чувствительности (SE), специфичности (SP), точности (ACC), прецизионности (PR) при 5-кратной ПП методов МОВ с линейным ядром, СЛ и СГБ. Порядок полинома выбран равным 2, так как 3-й порядок слишком сильно сглаживает спектр интенсивности убирая информативные пики, а 1-й порядок не оказывает существенного влияния. Согласно таблице 5, оптимальным значением размера окна является 5, так как большинство метрик выше, чем при других значениях.

Таблица 5 – Оценка оптимальных значений параметра window_length для фильтра Савицкого-Голея

| Значение Window_length Value | 5 | 15 | 31 | Без сглаживания |
|------------------------------|------------------------|----------------|-----------------------|-----------------------|
| Среднее SE для МОВ/СЛ/СГБt | 0,60/ 0,62/0,68 | 0,60/0,57/0,58 | 0,70/0,52/0,68 | 0,60/0,57/0,58 |
| Среднее SP для МОВ/СЛ/СГБt | 0,47/0,50/0,50 | 0,39/0,47/0,44 | 0,28/0,42/0,42 | 0,60/0,57/0,58 |
| Среднее ACC для МОВ/СЛ/СГБt | 0,55/ 0,57/0,61 | 0,52/0,53/0,53 | 0,54/0,48/0,58 | 0,60/0,57/0,58 |
| Среднее PR для МОВ/СЛ/СГБt | 0,65/0,67/0,69 | 0,62/0,64/0,64 | 0,62/0,60/0,66 | 0,60/0,57/0,58 |

На втором этапе было выполнено снижение размерности отфильтрованных данных и последующая визуализация с помощью методов МГК и t-SNE. Были проведены тесты применения МГК со стандартизацией данных и без нее и установлено, что без стандартизации на экспериментальных данных всех групп он лучше разделяет данные.

На третьем этапе проводился анализ разделимости данных с помощью предиктивной модели на основе МОВ с линейным ядром, СЛ, СГБ. Качество работы классификатора оценивалось с помощью ROC-кривой и анализа площади под кривой (AUC). Наилучшие показатели качества работы продемонстрировал метод СГБ (рисунок 6). Следовательно, СГБ лучше всего отделяют группу пациентов с МГБ от здоровых доноров и пациентов без МГБ (с дефектами черепа при краниэктомии).



а – СГБ эксперимент 1; б – СГБ эксперимент 2
 Рисунок 6 – ROC-анализ AUC для эксперимента 1 и эксперимента 2

На заключительном этапе анализа данных были выделены информативные частоты. Результаты представлены на рисунке 7.

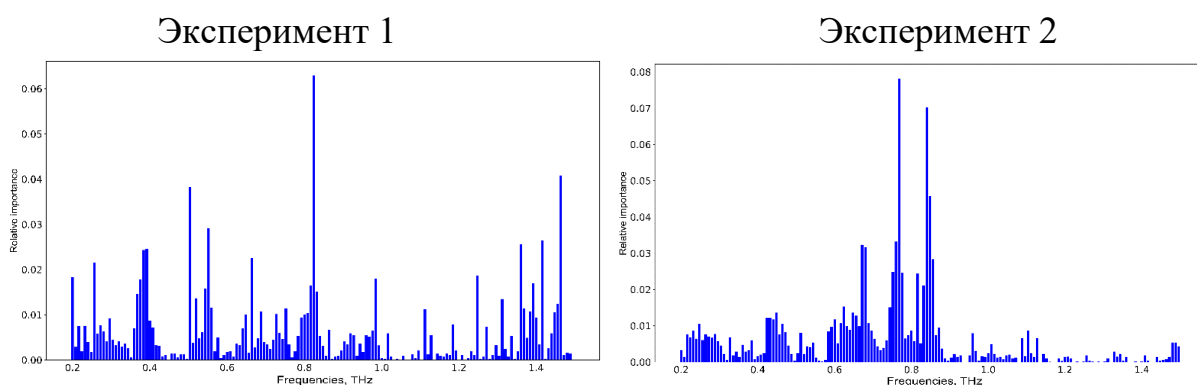


Рисунок 7 – Информативные частоты для эксперимента 1 и 2 для классификатора СГБ

Среди пациентов с МГБ было 5 пар пациентов, у которых образцы плазмы крови были взяты до и через неделю после хирургического вмешательства по удалению опухоли. Для установления различий в ТГц-спектрах до и после операции были применены методы МО. Было предположено, что некоторые особенности конкретного пациента нивелируются и различия в спектрах могут быть связаны с наличием или отсутствием опухоли до и после операции. Для дифференциации групп пациентов до и после операции использовались методы МОВ, СЛ, СГБ. Лучшие показатели точности были получены для метода СГБ (рисунок 8).

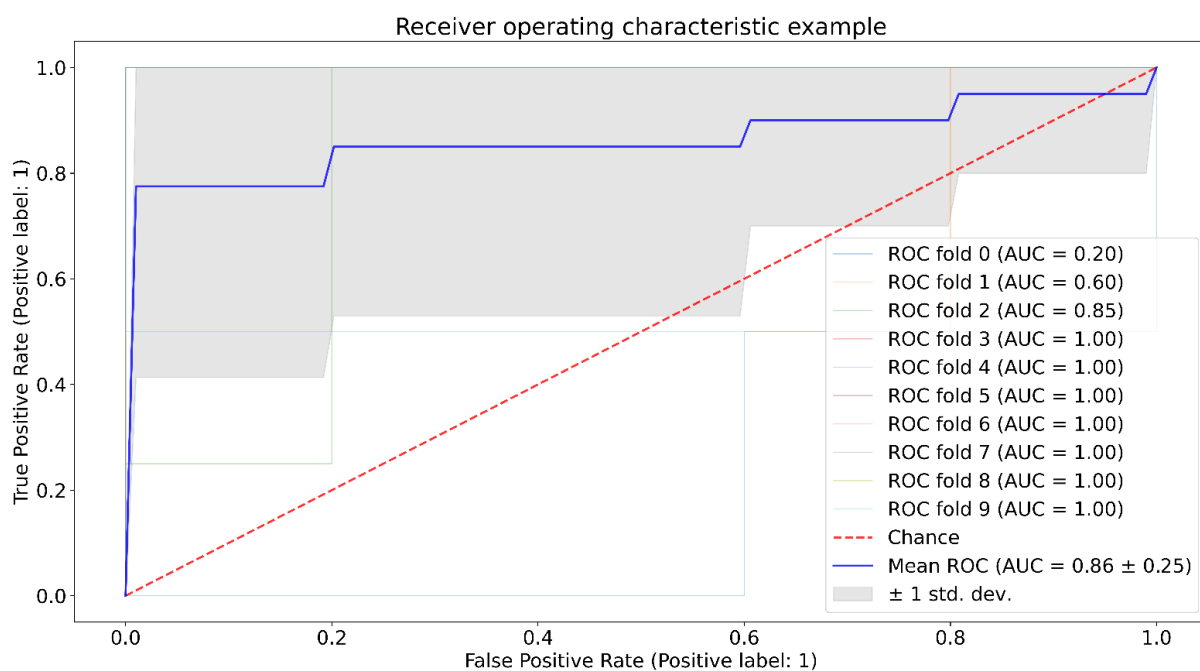


Рисунок 8 – График ROC-AUC кривой для метода СГБ

Как видно из рисунка 8, метод СГБ позволил различить плазму крови пациентов до и после операции с хорошей дискриминацией ($AUC=0,86-0,92$). Наиболее информативными частотами для полученных предиктивных моделей являются 0,56 и 1,2 ТГц.

В пятой главе представлены результаты анализа данных ТГц спектроскопии сыворотки крови методами машинного обучения для дифференциации МГБ и

травматического поражения головного мозга на образцах МЛЖ. На основе проведенного разведывательного анализа методами машинного обучения без учителя разработана и реализована схема эксперимента. Приведены описания групп, конвейера машинного обучения с соответствующими оптимальными значениям гиперпараметров и метрики качества и наиболее информативные признаки для построенных предиктивных моделей.

Описание исследуемых групп и данных

В данной главе на модели МЛЖ изучена возможность дифференциации глиобластомы от травматического повреждения головного мозга (ТПМ) с помощью анализа сыворотки крови методом терагерцовой спектроскопии во временной области и машинного обучения. Образцы культуральной среды и суспензии клеток глиобластомы человека U87 в культуральной среде были введены в подкорковые структуры головного мозга групп мышей, названных соответственно группами введения культуральной среды (ВКС) и группами глиобластомы (МГБ). Через первую, вторую и третью недели после инъекции были взяты образцы сыворотки крови и измерены спектры их терагерцового излучения. В течение первой недели инъекция вызывала острое воспаление в мозге, поэтому группа ВКС на первой неделе эксперимента соответствовала состоянию травматического повреждения мозга (ТПМ). На третьей неделе эксперимента острое воспаление практически исчезло в группах ВКС. В то же время группа МГБ, в которую вводили клетки глиобластомы человека U87, имела самый большой размер опухоли.

Согласно выдвинутой гипотезе, основной причиной этого стали воспалительные процессы в тканях головного мозга, возникшие на фоне ТПМ в месте инъекции после введения культуральной среды или клеток U87 в первую неделю эксперимента. Для проверки этой гипотезы был предложен следующий анализ экспериментальных данных. Во-первых, проверялась однородность групп на 7-й, 14-й и 21-й день после инъекции. Во-вторых, если между этими группами имелась разница, то третья группа ВКС рассматривалась как мыши со здоровым мозгом, а первая и вторая группы ВКС соответствовали ТПМ. Их сравнение с третьей группой ВКС позволило выделить информативные признаки ТПМ. Затем, сравнивая первую группу ВКС (связанную с ТПМ) и третью группу МГБ (21-й день после введения клеток глиобластомы человека U87), можно было найти разницу между травматическим повреждением мозга и глиобластомой. Список соответствующих моделей МО приведен на рисунке 9. Модели 1-3 использовались для проверки однородности групп ВКС, а модель 4 - для проверки разделения между группами ТПМ и МГБ.

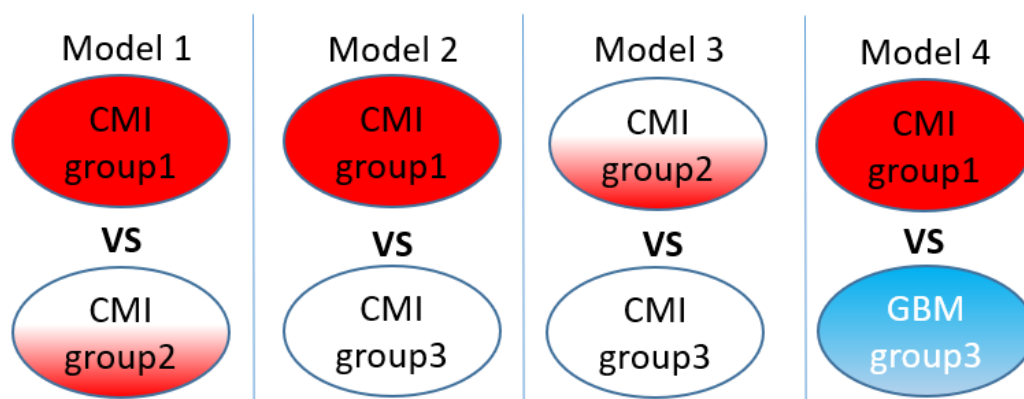


Рисунок 9 – Список исследуемых бинарных моделей МО

Конвейер машинного обучения для дифференциации МГБ и травматического поражения головного мозга

Разведывательный визуальный анализ ТГц-спектров проводился с помощью двух алгоритмов снижения размерности (МГК и t-SNE) и трех алгоритмов классификации (МОВ, СЛ и СГБ).

Построенные модели прогнозных данных были проверены с помощью 10-кратной ПП, кривой операционных характеристик приемника и соответствующего анализа площади под кривой. Предложенный конвейер МО (рисунок 10) позволил отличить группу травматических повреждений головного мозга от группы глиобластомы с чувствительностью 95 %, специфичностью 100 % и точностью 97 % с помощью метода СГБ. Наиболее информативными признаками для различения этих групп оказались 0,37, 0,40, 0,55, 0,60, 0,70 и 0,90 ТГц. Это позволяет сделать вывод, что анализ сыворотки крови мышей с помощью терагерцовой спектроскопии во временной области и машинного обучения позволяет дифференцировать глиобластому от травматического повреждения мозга.

Для проверки однородности групп ВКС использовались методы МО без учителя (МГК и t-SNE).

Для выбора оптимальных гиперпараметров для каждой модели данных использовался алгоритм GridSearchCV из Python библиотеки sklearn.model.selection. Используя ПП на заданной сетке поиска параметров, алгоритм находил предиктивную модель данных с наивысшими показателями точности для заданного алгоритма МО. Этот шаг был выполнен для того, чтобы избежать нестабильности алгоритмов выбора признаков на моделях с параметрами по умолчанию. Кроме того, оценивалось влияние стандартизации данных (вычитания среднего и деления на дисперсию) на методы снижения размерности. Для корректного сравнения ТГц-данные нормировались по максимальному значению интенсивности.

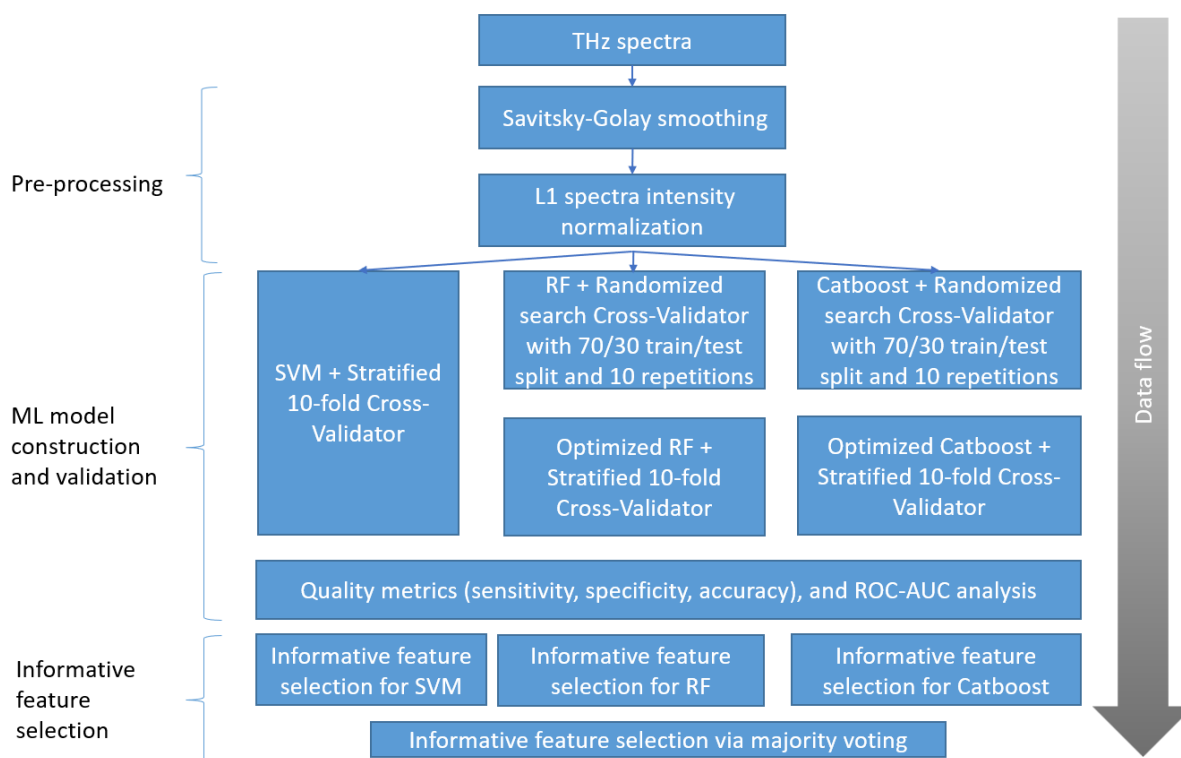


Рисунок 10 – Реализованный конвейер МО

Результаты анализа ТГц спектров сыворотки крови мышей для дифференциации МГБ и травматического поражения головного мозга методами машинного обучения

Представление групп ВКС в пространстве признаков с помощью метода t-SNE показано на рисунке 11. Хорошо видно, что группа ВКС 1 (за исключением шести красных точек вблизи области с координатами (-2, -2), соответствующих спектрам одного образца) отделена от групп ВКС 2 и 3. Точки из групп ВКС 2 и 3 сильно разбросаны и перемешаны между собой. Это объясняется тем, что побочные эффекты от инъекций еще не прошли. Этот результат согласуется с выводами, сделанными на основе анализа 1H-MRS, о том, что нейровоспаление проходит к третьей неделе.

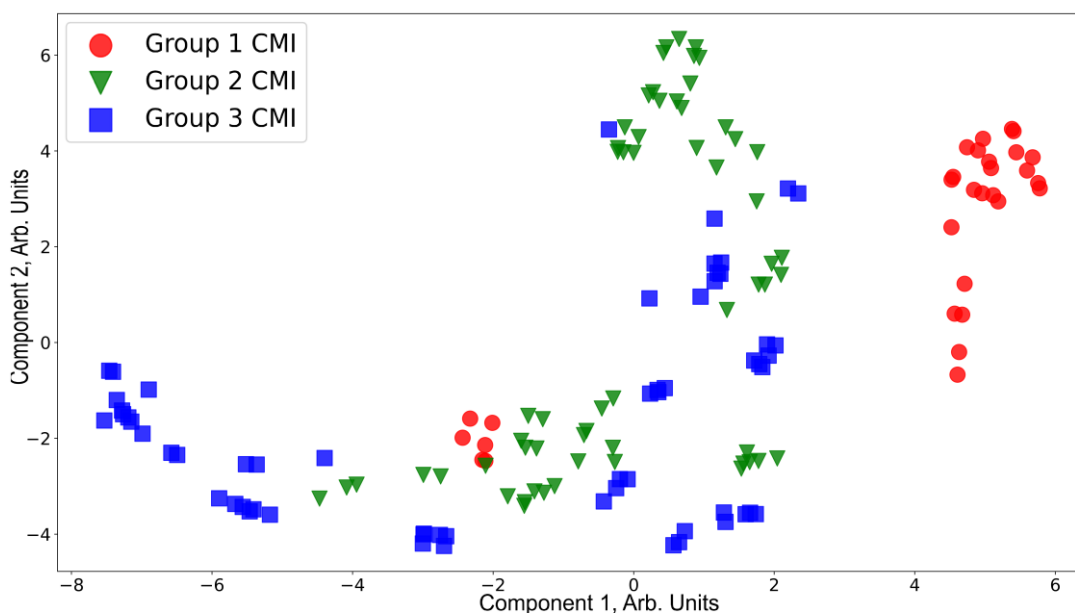


Рисунок 11 – t-SNE визуализация групп ВКС

Подбор оптимальных гиперпараметров для СЛ реализован при помощи RandomizedSearchCV при кратности равной 10. Параметры СГБ (библиотека Python Catboost) были оптимизированы процедурой `catboost.randomized_search`. Оптимальные значения параметров СЛ и СГБ приведены в таблице 6.

Таблица 6 – Оптимальные значения параметров для предлагаемого конвейера МО

| Модель | Оптимальные параметры СЛ согласно именованию в модуле <code>sklearn.ensemble.RandomForestClassifier</code> | Оптимальные параметры СГБ согласно именованию в модуле <code>catboost.CatBoostClassifier</code> | Размер окна в методе СГ, точки |
|------------------------|---|---|--------------------------------|
| 1 (ВКС 1 против ВКС 2) | 'n_estimators': 100, 'min_samples_split': 2, 'min_samples_leaf': 1, 'max_depth': 20, 'class_weight': 'balanced_subsample', 'bootstrap': True | 'eval_metric': 'Logloss', 'iterations': 300, 'depth': 6, 'auto_class_weights': 'Balanced', 'loss_function': 'Logloss', 'learning_rate': 0.1, 'leaf_estimation_iterations': 10, 'max_leaves': 64 | 25 |
| 2 (ВКС 1 против ВКС 3) | 'n_estimators': 50, 'min_samples_split': 2, 'min_samples_leaf': 1, 'max_depth': 60, 'class_weight': 'balanced_subsample', 'bootstrap': True | 'eval_metric': 'Logloss', 'iterations': 300, 'depth': 4, 'auto_class_weights': 'Balanced', 'loss_function': 'Logloss', 'learning_rate': 0.1, 'leaf_estimation_iterations': 10, 'max_leaves': 16 | 25 |

Окончание таблицы 6

| | | | |
|------------------------|---|--|----|
| 3 (ВКС 2 против ВКС 3) | 'n_estimators': 100, 'min_samples_split': 2, 'min_samples_leaf': 1, 'max_depth': 10, 'class_weight': 'balanced_subsample', 'bootstrap': True | 'eval_metric': 'Logloss', 'iterations': 300, 'depth': 2, 'auto_class_weights': 'Balanced', 'loss_function': 'Logloss', 'learning_rate': 0.1, 'leaf_estimation_iterations': 10, 'max_leaves': 4 | 25 |
| 4 (ВКС 1 против МГБ 3) | 'n_estimators': 100, 'min_samples_split': 6, 'min_samples_leaf': 3, 'max_depth': 110, 'class_weight': 'balanced_subsample', 'bootstrap': True | 'eval_metric': 'Logloss', 'iterations': 300, 'depth': 2, 'auto_class_weights': 'Balanced', 'loss_function': 'Logloss', 'learning_rate': 0.03, 'leaf_estimation_iterations': 10, 'max_leaves': 16 | 25 |

Было проведено попарное сравнение групп ВКС по метрикам качества и информативным признакам: 1-я неделя против 2-й, 1-я неделя против 3-й, 2-я неделя против 3-й. Наилучшие показатели ROC-AUC были достигнуты классификаторами МОВ и СГБ (табл. 7). Идеальная разделяемость была между группами ВКС 1 и 3, что подтверждает предыдущий вывод о периоде нейровоспаления.

Таблица 7 – Усредненные показатели ROC-AUC в группах ВКС для различных классификаторов

| Классификатор | ВКС группа 1 против 2 | ВКС группа 1 против 3 | ВКС группа 2 против 3 |
|---------------|-----------------------|-----------------------|-----------------------|
| МОВ | 0,99 | 0,99 | 0,74 |
| СЛ | 0,98 | 0,99 | 0,88 |
| СГБ | 0,98 | 0,99 | 0,92 |

Далее, было проведено сравнение группы ВКС 1 и группы МГБ 3, чтобы дифференцировать ТПМ и МГБ. Результаты, представленные в таблице 8, показывают, что все классификаторы продемонстрировали высокие показатели качества, при этом специфичность для предиктивных моделей на основе МОВ и СЛ имеет большую дисперсию.

Таблица 8 – Усредненные показатели ROC-AUC, чувствительности, специфичности и точности \pm дисперсия дифференциации группы 1 СМІ и группы 3 GBM с помощью различных классификаторов

| Классификатор | AUC, отн. ед. | Чувствительность, % | Специфичность, % | Точность, % |
|---------------|-----------------|---------------------|------------------|-----------------|
| МОВ | 0,98 \pm 0,05 | 100,0 \pm 0,0 | 80,0 \pm 4,0 | 92,0 \pm 17,0 |
| СЛ | 0,95 \pm 0,15 | 95,0 \pm 15,0 | 80,0 \pm 4,0 | 89,0 \pm 23,0 |
| СГБ | 0,95 \pm 0,12 | 95,0 \pm 15,0 | 100,0 \pm 0,0 | 97,0 \pm 8,0 |

Информативные признаки, отличающие ТПМ и МГБ при помощи классификатора МОВ, соответствуют 0,375, 0,4, 0,55, 0,6, 0,7, 0,9 ТГц.

ЗАКЛЮЧЕНИЕ

В диссертационной работе проведено исследование образцов сыворотки крови мелких лабораторных животных: здоровых, с ортотопической ксенографтовой моделью глиомы на основе линии клеток U87, с травматическими поражениями головного мозга, а также образцов плазмы крови больных глиомой, здоровых добровольцев и пациентов после краниэктомии на основе анализа данных ТГц спектроскопии с использованием методов искусственного интеллекта.

Разработана предиктивная модель для дифференцирования образцов сыворотки крови мелких лабораторных животных с ксенографтовой моделью глиомы на основе клеток U87 и группы контроля на основе анализа данных ТГц спектроскопии в диапазоне 0,2-1,7 ТГц. Конвейер машинного обучения, использованный для построения данной предиктивной модели, включал предварительную обработку данных фильтром Савицкого-Голея и удаление выбросов методом изолированного леса, уменьшение размерности методом главных компонент, классификацию методом опорных векторов с линейным ядром. Использование данной комбинации методов позволило обнаружить информативные ТГц диапазоны для делимости изучаемых групп. Средние значения чувствительности, специфичности, точности и площади под ROC-кривой для разработанных предиктивных моделей составили не менее 90% на использованной экспериментальной выборке. Предложен последовательный алгоритм классификации с выделением информативных признаков на основе МГК и МОВ с линейным ядром.

Разработаны предиктивные модели для попарной дифференциации групп с глиомой, здоровых доноров и пациентов после краниэктомии на основе данных ТГц спектроскопии сыворотки крови. Соответствующий конвейер МО с подобранными оптимальными значениями гиперпараметров, включал предварительную обработку данных фильтром Савицкого-Голея с шириной окна 5 и степенью полинома 2, классификацию методом стохастического градиентного бустинга с глубиной деревьев 2, автоматической балансировкой классов, логарифмической функцией потерь. Среднее значение показателя площади под ROC-кривой для моделей на основе СЛ и СГБ на использованной экспериментальной выборке составило 0,92 и 0,86 соответственно. Также получены данные об информативных ТГц частотах для дифференциации соответствующих групп – 0,56 и 1,2 ТГц.

Построены предиктивные модели для дифференциации образцов сыворотки крови МЛЖ с ксенографтовой моделью глиомы на основе клеток U87 и группы с травматическим поражением головного мозга. Предиктивная модель, построенная при помощи конвейера методов машинного обучения, включающего фильтр с Савицкого-Голея с шириной окна 25 и степенью полинома 2, нормализацию данных путем вычитания среднего и деления на дисперсию, ансамблевый классификатор на основе комбинации метода опорных векторов с линейным ядром, стохастического градиентного бустинга и случайного леса с подобранными оптимальными

гиперпараметрами, позволила дифференцировать группу мелких лабораторных животных с травматическим повреждением головного мозга от группы мелких лабораторных животных с ксенографтовой моделью глиомы на основе клеток U87, со средними значениями чувствительности, специфичности и точности равными 95%, 100% и 97%, соответственно, на использованной экспериментальной выборке. Обнаружены наиболее информативные частоты 0,37, 0,40, 0,55, 0,60, 0,70 и 0,90 ТГц для дифференциации соответствующих образцов плазмы крови МЛЖ с травматическим повреждением головного мозга и группы с моделью глиомы.

Результаты работы дают общие представления об модификациях оптических характеристик сыворотки и плазмы крови больных в динамике развития глиомы головного мозга и при травматических поражениях головного мозга. Эта информация важна при использовании методов оптической биопсии для малоинвазивной и интраоперационной диагностики глиомы и травматических поражений мозга.

Благодарности

Автор выражает благодарность своему научному руководителю доктору физико-математических наук, профессору физического факультета Кистеневу Юрию Владимировичу за постоянное внимание к работе, неизменную поддержку, заинтересованное обсуждение и доброжелательную критику получаемых результатов. Автор выражает благодарность доктору физико-математических наук, профессору Шаповалову Александру Васильевичу за моральную поддержку и внимание к диссертационной работе, а также всем сотрудникам лаборатории лазерного молекулярного имиджинга и машинного обучения и лаборатории биофотоники за доброжелательное отношение, поддержку во время работы и при подготовке к защите диссертации. Автор благодарит доктора медицинских наук, профессора Черкасову Ольгу Павловну и кандидата физико-математических наук Князькову Анастасию Игоревну за неоценимую помощь в подготовке материалов диссертации.

СПИСОК ИСТОЧНИКОВ, ЦИТИРУЕМЫХ В АВТОРЕФЕРАТЕ

1. Agnihotri S. Glioblastoma, a brief review of history, molecular genetics, animal models and novel therapeutic strategies / S. Agnihotri, K. E. Burrell, A. Wolf, S. Jalali, C. Hawkins, J. T. Rutka, G. Zadeh // *Archivum immunologiae et therapiae experimentalis*. – 2013. – Vol. 61. – P. 25–41.
2. Wesseling P. WHO 2016 Classification of gliomas / P. Wesseling, D. Capper // *Neuropathology and applied neurobiology*. – 2018. – Vol. 44, №. 2. – P. 139–150.
3. Ostrom Q. T. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2014–2018 / Q. T. Ostrom, G. Cioffi, K. Waite, C. Kruchko, J. S. Barnholtz-Sloan // *Neuro-oncology*. – 2021. – Vol. 23, is. Supplement 3. – P. iii1–iii105.
4. Torp S. H. The WHO 2021 Classification of Central Nervous System tumours: a practical update on what neurosurgeons need to know—a minireview / S. H. Torp, O. Solheim, A. J. Skjelsvik // *Acta Neurochirurgica*. – 2022. – Vol. 164, is. 9. – P. 2453–2464.
5. Hanif F. Glioblastoma multiforme: a review of its epidemiology and pathogenesis through clinical presentation and treatment / F. Hanif, K. Muzaffar, K. Perveen, S. M. Malhi, S. U. Simjee // *Asian Pacific journal of cancer prevention: APJCP*. – 2017. – Vol. 18, №. 1. – P. 3.
6. Hammoud M. A. Prognostic significance of preoperative MRI scans in glioblastoma multiforme / M. A. Hammoud, R. Sawaya, W. Shi, P. F. Thall, N. E. Leeds // *Journal of neuro-oncology*. – 1996. – Vol. 27. – P. 65–73.
7. Wangaryattawanich P. Multicenter imaging outcomes study of The Cancer Genome Atlas glioblastoma patient cohort: imaging predictors of overall and progression-free survival / P. Wangaryattawanich, M. Hatami, J. Wang, G. Thomas, A. Flanders, J. Kirby, M. Wintermark, E. S. Huang, A. S. Bakhtiari, M. M. Luedi, S. S. Hashmi, D. L. Rubin, J. Y. Chen, S. N. Hwang, J. Freymann, C. A. Holder, P. O. Zinn, R. R. Colen // *Neuro-oncology*. – 2015. – Vol. 17, №. 11. – P. 1525–1537.
8. Bulik M. Potential of MR spectroscopy for assessment of glioma grading // M. Bulik, R. Jancalek, J. Vanicek, A. Skoch, M. Mechl // *Clinical neurology and neurosurgery*. – 2013. – Vol. 115, is. 2. – P. 146–153.
9. Leijenaar R. T. H. External validation of a prognostic CT-based radiomic signature in oropharyngeal squamous cell carcinoma / R. T. H. Leijenaar, S. Carvalho, F. J. P. Hoebbers, H. J. W. L. Aerts, W. J. C. van Elmpt, Sh. H. Huang, B. Chan, J. N. Waldron, B. O'sullivan, Ph. Lambin // *Acta oncologica*. – 2015. – Vol. 54, №. 9. – P. 1423–1429.
10. Honda N. Determination of optical properties of human brain tumor tissues from 350 to 1000 nm to investigate the cause of false negatives in fluorescence-guided resection with 5-aminolevulinic acid / N. Honda, K. Ishii, Y. Kajimoto, T. Kuroiwa, K. Awazu // *Journal of biomedical optics*. – 2018. – Vol. 23, №. 7. – P. 075006.
11. Vasefi F. Review of the potential of optical technologies for cancer diagnosis in neurosurgery: a step toward intraoperative neurophotonics / F. Vasefi, N. MacKinnon, D. L. Farkas, B. Kateb // *Neurophotonics*. – 2017. – Vol. 4, №. 1. – P. 011010–011010.

12. Liang B. Impacts of the murine skull on high-frequency transcranial photoacoustic brain imaging / B. Liang, W. Liu, Q. Zhan, M. Li, M. Zhuang, Q. H. Liu, J. Yao // *Journal of biophotonics*. – 2019. – Vol. 12, №. 7. – P. e201800466.
13. Jermyn M. Raman spectroscopy detects distant invasive brain cancer cells centimeters beyond MRI capability in humans / M. Jermyn, J. Desroches, J. Mercier, K. St-Arnaud, M. C. Guiot, F. Leblond, K. Petrecca // *Biomedical optics express*. – 2016. – Vol. 7, №. 12. – P. 5129–5137.MLA
14. Kong K. Raman spectroscopy for medical diagnostics—From in-vitro biofluid assays to in-vivo cancer detection / K. Kong, C. Kendall, N. Stone, I. Notingher // *Advanced drug delivery reviews*. – 2015. – Vol. 89. – P. 121–134.
15. Genina E. A. Optical properties of brain tissues at the different stages of glioma development in rats: pilot study / E. A. Genina, A. N. Bashkatov, D. K. Tuchina, P. A. Dyachenko (Timoshina), N. Navolokin, A. Shirokov, A. Khorovodov, A. Terskov, M. Klimova, A. Mamedova, I. Blokhina, I. Agranovich, E. Zinchenko, O. V. Semyachkina-Glushkovskaya, V. V. Tuchin // *Biomedical optics express*. – 2019. – Vol. 10, №. 10. – P. 5182–5197.
16. Cameron J. M. Clinical validation of a spectroscopic liquid biopsy for earlier detection of brain cancer / J. M. Cameron, P. M. Brennan, G. Antoniou, H. J. Butler, L. Christie, J. J. A. Conn, T. Curran, E. Gray, M. G. Hegarty, M. Jenkinson, D. Orringer, D. S. Palmer, A. Sala, B. R. Smith, M. J. Baker // *Neuro-oncology advances*. – 2022. – Vol. 4, №. 1. – P. vdac024.
17. Zaytsev K. I. The progress and perspectives of terahertz technology for diagnosis of neoplasms: a review / K. I. Zaytsev, I. N. Dolganova, N. V. Chernomyrdin, G. M. Katyba, A. A. Gavdush, O. P. Cherkasova, G. A. Komandin, M. A. Shchedrina, A. N. Khodan, D. S. Ponomarev, I. V. Reshetov, V. E. Karasik, M. Skorobogatiy, V. N. Kurlov, V. V. Tuchin // *Journal of Optics*. – 2019. – Vol. 22, №. 1. – P. 013001.
18. Gavdush A. A. Terahertz spectroscopy of gelatin-embedded human brain gliomas of different grades: a road toward intraoperative THz diagnosis / A. A. Gavdush, N. V. Chernomyrdin, K. M. Malakhov, S.-I. T. Beshplav, I. N. Dolganova, A. V. Kosyrkova, P. V. Nikitin, G. R. Musina, G. M. Katyba, I. V. Reshetov, O. P. Cherkasova, G. A. Komandin, V. E. Karasik, A. A. Potapov, V. V. Tuchin, K. I. Zaytsev // *Journal of biomedical optics*. – 2019. – Vol. 24, №. 2. – P. 027001-027001.
19. Gavdush A. A. Terahertz dielectric spectroscopy of human brain gliomas and intact tissues ex vivo: double-Debye and double-overdamped-oscillator models of dielectric response / A. A. Gavdush, N. V. Chernomyrdin, G. A. Komandin, I. N. Dolganova, P. V. Nikitin, G. R. Musina, G. M. Katyba, A. S. Kucheryavenko, I. V. Reshetov, A. A. Potapov, V. V. Tuchin, K. I. Zaytsev // *Biomedical Optics Express*. – 2020. – Vol. 12, №. 1. – P. 69–83.
20. Lone S. N. Liquid biopsy: a step closer to transform diagnosis, prognosis and future of cancer treatments / S. N. Lone, S. Nisar, T. Masoodi, M. Singh, A. Rizwan, S. Hashem, W. El-Rifai, D. Bedognetti, S. K. Batra, M. Haris, A. A. Bhat, M. A. Macha // *Molecular cancer*. – 2022. – Vol. 21, №. 1. – P. 79.

21. Ronvaux L. Liquid biopsy in glioblastoma / L. Ronvaux, M. Riva, A. Coosemans, M. Herzog, G. Rommelaere, N. Donis, L. D'Hondt, J. Douxfils // *Cancers*. – 2022. – Vol. 14, №. 14. – P. 3394.
22. Ali H. Blood-based biomarkers for glioma in the context of gliomagenesis: a systematic review / H. Ali, R. Harting, R. de Vries, M. Ali, T. Wurdinger, M. G. Best // *Frontiers in Oncology*. – 2021. – Vol. 11. – P. 665235.
23. Figueroa J. M. Detection of glioblastoma in biofluids / J. M. Figueroa, B. S. Carter // *Journal of Neurosurgery*. – 2017. – Vol. 129, №. 2. – P. 334–340.
24. Jelski W. Molecular and circulating biomarkers of brain tumors / W. Jelski, B. Mroczko // *International journal of molecular sciences*. – 2021. – Vol. 22, №. 13. – P. 7039.
25. Smolyanskaya O. A. Terahertz biophotonics as a tool for studies of dielectric and spectral properties of biological tissues and liquids / O. A. Smolyanskaya, N. V. Chernomyrdin, A. A. Konovko, K. I. Zaytsev, I. A. Ozheredov, O. P. Cherkasova, M. M. Nazarov, J.-P. Guillet, S. A. Kozlov, Yu. V. Kistenev, J.-L. Coutaz, P. Mounaix, V. L. Vaks, J.-H. Son, H. Cheon, V. P. Wallace, Yu. Feldman, I. Popov, A. N. Yaroslavsky, A. P. Shkurinov, V. V. Tuchin // *Progress in Quantum Electronics*. – 2018. – Vol. 62. – P. 1–77.
26. Kistenev Y. V. Diagnosis of oral lichen planus from analysis of saliva samples using terahertz time-domain spectroscopy and chemometrics / Y. V. Kistenev, A. V. Borisov, M. A. Titarenko, O. D. Baydik, A. V. Shapovalov // *Journal of Biomedical Optics*. – 2018. – Vol. 23, №. 4. – P. 045001–045001.
27. Ahmed S. F. Deep learning modelling techniques: current progress, applications, advantages, and challenges / S. F. Ahmed, M. S. B. Alam, M. Hassan, M. R. Rozbu, T. Ishtiak, N. Raza, M. Mofijur, A. B. M. S. Ali, A. H. Gandomi // *Artificial Intelligence Review*. – 2023. – Vol. 56, №. 11. – P. 13521–13617.
28. Angelov P. P. Explainable artificial intelligence: an analytical review / P. P. Angelov, E. A. Soares, R. Jiang, N. I. Arnold, P. M. Atkinson // *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. – 2021. – Vol. 11, №. 5. – P. e1424.
29. Vrazhnov D. Discovering glioma tissue through its biomarkers' detection in blood by Raman spectroscopy and machine learning / D. Vrazhnov, A. Mankova, E. Stupak, Y. Kistenev, A. Shkurinov, O. Cherkasova // *Pharmaceutics*. – 2023. – Vol. 15, №. 1. – P. 203.
30. Maslova G. T. The use of morphostructural analysis and IR spectrometry for diagnosis of the brain tumors / G. T. Maslova, J. I. Bulochik, A. P. Zajogin, A. S. Mavrichev, L. A. Derzhavets, A. S. Trubetskaya, A. V. Titova // *Doklady BGUIR*. – 2018. – №. 7. – P. 25-29.MLA
31. Park H. Machine learning techniques for THz imaging and time-domain spectroscopy / H. Park, J. H. Son // *Sensors*. – 2021. – Vol. 21, №. 4. – P. 1186.
32. Shevelev O. B. In vivo MRS study of long-term effects of traumatic intracranial injection of a culture medium in mice / O. B. Shevelev, O. P. Cherkasova, I. A. Razumov, E. L. Zavjalov // *Vavilov Journal of Genetics and Breeding*. – 2023. – Vol. 27, №. 6. – P. 633.

33. Knyazkova A. I. Paraffin-embedded prostate cancer tissue grading using terahertz spectroscopy and machine learning / A. I. Knyazkova, A. V. Borisov, L. V. Spirina, Y. V. Kistenev // *Journal of Infrared, Millimeter, and Terahertz Waves.* – 2020. – Vol. 41, №. 9. – P. 1089–1104.

Список работ, опубликованных по теме диссертации

Статьи в журналах, включенных в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук:

1*. Tretyakov A. Qualitative and Quantitative Analysis of Volatile Molecular Biomarkers in Breath Using THz-IR Spectroscopy and Machine Learning / A. Tretyakov, **D. Vrazhnov**, A. Shkurinov, V. Zasedatel, Y. Kistenev // *Applied Sciences.* – 2024. – Vol. 14, № 24. – Article number 11521. – 17 p. – URL: <https://www.mdpi.com/2076-3417/14/24/11521> (access date: 10.04.2025). – DOI: 10.3390/app14241152. – 1,47 / 0,29 а.л. (*Web of Science*).

2*. **Vrazhnov D. A.** Terahertz time-domain spectroscopy of blood serum for differentiation of glioblastoma and traumatic brain injury / D. A. Vrazhnov, D. A. Ovchinnikova, T. V. Kabanova, A. G. Paulish, Yu. V. Kistenev, N. A. Nikolaev, O. P. Cherkasova // *Applied Sciences.* – 2024. – Vol. 14, is. 7. – Article number 2872. – 21 p. – URL: <https://www.mdpi.com/2076-3417/14/7/2872> (access date: 10.04.2025). – DOI: 10.3390/app14072872. – 1,74 / 0,25 а.л. (*Web of Science*).

3*. Cherkasova O. P. Terahertz Spectroscopy of Mouse Blood Serum in the Dynamics of Experimental Glioblastom / O. P. Cherkasova, M. R. Konnikova, M. M. Nazarov, **D. A. Vrazhnov**, Yu. V. Kistenev, A. P. Shkurinov // *Journal of Biomedical Photonics & Engineering.* – 2023. – Vol. 9, № 3. – Article number 030308. – 7 p. – URL: <https://jbpe.ssau.ru/index.php/JBPE/article/view/8990>. (access date: 10.04.2025). – DOI: 10.18287/JBPE23.09.030308. – 0,75 / 0,12 а.л. (*Scopus*).

4*. Cherkasova O. Terahertz Time-Domain Spectroscopy of Glioma Patient Blood Plasma: Diagnosis and Treatment / O. Cherkasova, **D. Vrazhnov**, A. Knyazkova, M. Konnikova, E. Stupak, V. Glotov, V. Stupak, N. Nikolaev, A. Paulish, Y. Peng, Yu.V. Kistenev, A. Shkurinov // *Applied Sciences.* – 2023. – Vol. 13, is. 9. – Article number 5434. – 17 p. – URL: <https://www.mdpi.com/2076-3417/13/9/5434> (access date: 10.04.2025). – DOI: 10.3390/app13095434. – 1,22 / 0,1 а.л. (*Web of Science*).

5*. **Vrazhnov D.** Analysis of Mouse Blood Serum in the Dynamics of U87 Glioblastoma by Terahertz Spectroscopy and Machine Learning / D. Vrazhnov, A. Knyazkova, M. Konnikova, O. Shevelev, I. Razumov, E. Zavjalov, Y. Kistenev, A. Shkurinov, O. Cherkasova // *Applied Sciences.* – 2022. – Vol. 12, is. 20. – Article number 10533. – 18 p. – URL: <https://www.mdpi.com/2076-3417/12/20/10533> (access date: 10.04.2025). – DOI: 10.3390/app122010533. – 1,02 / 0,11 а.л. (*Web of Science*).

6*. Cherkasova O. Diagnosis of glioma molecular markers by terahertz technologies / O. Cherkasova, Y. Peng, M. Konnikova, Y. Kistenev, C. Shi, **D. A. Vrazhnov**, O. Shevelev, E. Zavjalov, S. Kuznetsov, A. Shkurinov // *Photonics.* –

2021. – Vol. 8, is. 1. – Article number 22. – URL: <https://www.mdpi.com/2304-6732/8/1/22> (access date: 10.04.2025). – 30 p. – DOI: 10.3390/photonics8010022. – 2,66 / 0,27 а.л. (*Web of Science*).

Публикации в сборниках материалов конференции, представленных в изданиях, входящих в Web of Science:

7*. Kistenev Y. V. Applications of THz laser spectroscopy and machine learning for medical diagnostics / Y. V. Kistenev, A. V. Borisov, A. I. Knyazkova, E. A. Sandykova, V. V. Nikolaev, **D. A. Vrazhnov** // EPJ Web of Conferences. – 2018. – Vol. 195 : 3rd International Conference “Terahertz and Microwave Radiation: Generation, Detection and Applications” (TERA–2018). Nizhny Novgorod, Russia, October 22–25, 2018. – Article number 10006. – 2 p. – URL: https://www.epj-conferences.org/articles/epjconf/abs/2018/30/epjconf_tera2018_10006/epjconf_tera2018_10006.html (access date: 10.04.2025). – 0,12 / 0,02 а.л.

8*. Kistenev Y. V. IR and THz imaging of paraffin embedded cancer tissues / Y. V. Kistenev, A. V. Borisov, V. V. Nikolaev, **D. A. Vrazhnov**, A. I. Knyazkova // Proceedings of IEEE : 44rd International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz – 2019). Paris, France, September 01–06, 2019. – New York, 2019. – Article number 8874212. – 2 p. – URL: <https://ieeexplore.ieee.org/document/8874212> (access date: 10.04.2025). – 0,12 / 0,02 а.л.

9*. Kistenev Y. V. Medical diagnosis using NIR and THz tissue imaging and machine learning methods / Y. V. Kistenev, V. V. Tuchin, A. V. Borisov, E. N. Lazareva, V. V. Nikolaev, D. K. Tuchina, **D. A. Vrazhnov**, I. Yu. Yanina // Proceedings of SPIE. – 2019. – Vol. 10877 : Dynamics and Fluctuations in Biomedical Photonics XVI. San Francisco, California, United States, February 02–03, 2019. – Article number 10877J. – 11 p. – URL: <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10877/2508166/Medical-diagnosis-using-NIR-and-THz-tissue-imaging-and-machine/10.1117/12.2508166.short> (access date: 10.04.2025). – DOI: 10.1117/12.2508166. – 1,09 / 0,14 а.л.

10*. Cherkasova O. The possibilities of optical methods in the early diagnosis of gliomas / O. Cherkasova, A. Mankova, M. Konnikova, P. Solyankin, **D. Vrazhnov**, Yu. Kistenev, A. Sinko, Y. Peng, E.L. Zavjalov // Proceedings of IEEE : International Conference Laser Optics 2020 (ICLO – 2020). St. Petersburg, Russia, November 02–06, 2020. – 2020. – Article number 9285845. – 1 p. – URL: <https://ieeexplore.ieee.org/document/9285845> (access date: 10.04.2025). – 0,13 / 0,015 а.л.

11*. Lykina A. Study of dry pellets of blood plasma using THz spectroscopy / A. Lykina, M. Konnikova, M. Cherniaeva, P. Gavrilova, I. Mustafin, E. Domracheva, V. Anfertev, **D. Vrazhnov**, V. Prishchepa, Y. Toropova, D. Korolev, A. Shkurinov // Proceedings of IEEE : International Conference Laser Optics 2020 (ICLO 2020). St. Petersburg, Russia, November 02–06, 2020. – 2020. – Article number 9285903. – 1 p. – URL: <https://ieeexplore.ieee.org/document/92859030> (access date: 10.04.2025). – 0,12 / 0,01 а.л.

12*. Cherkasova O. Terahertz and infrared absorption spectra of mouse blood plasma in glioma development / O. Cherkasova, M. Konnikova, T. Heinz, M. Nazarov, A. Kuryanova, N. Aksenova, A. Mankova, **D. Vrazhnov**, Y. Kistenev, Y. Peng,

A. Shkurinov // Proceedings of IEEE : 2021 46th International Conference on Infrared, Millimeter and Terahertz Waves (IRMMW-THz – 2021). Chengdu, China, August 30 – September 03, 2021. – 2021. – Article number 9567163. – 2 p. – URL: <https://ieeexplore.ieee.org/document/9567163> (access date: 10.04.2025). – 0,25 / 0,02 а.л.

Монография:

13*. Kistenev Y. V. Medical applications of laser molecular imaging and machine learning / Y. V. Kistenev, A. V. Borisov, **D. A. Vrazhnov**. – Bellingham : Society of Photo-Optical Instrumentation Engineers, 2021. – 246 p. – 11,03 / 3,68 а.л.

Издание подготовлено в авторской редакции.

Отпечатано на участке цифровой печати

Издательства Томского государственного университета

Заказ № 7775 от «04» июля 2025 г. Тираж 100 экз.

г. Томск, Московский тр. 8, тел. (3822) 53-15-28

publish.tsu.ru