СОВРЕМЕННЫЕ ПОДХОДЫ ПРОГНОЗИРОВАНИЯ ДЛЯ ДИАГНОСТИКИ РАКА МОЛОЧНОЙ ЖЕЛЕЗЫ НА ОСНОВЕ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ

Мухамедиева Дилдора Кабиловна[1], Дошанова Малика Юлдашовна[1], Шаазизова Мадина Элдаровна[1]

[1]ТУИТ имени Мухаммада ал-Хоразмий

K E Y W O R D S

Машинное зрение, классификация, сортировка, зерно, анализ и обработка изображений, реальное время

A B S T R A C T

В статье рассматриваются современные подходы прогнозирования и классификации при диагностике рака молочной железы на основе интеллектуального анализа данных. Описаны наиболее используемые датасеты и модальности, этапы предобработки и извлечения признаков, классические и глубокие модели, методы валидации и оценки, вопросы интерпретируемости, клинической валидации, регуляторики и этические аспекты. Приведены рекомендации по практическому внедрению и направления дальнейших исследований.

Введение

Рак молочной железы – одна из ведущих причин заболеваемости и смертности среди женщин во многих странах. Ранняя и точная диагностика существенно повышает шансы на успешное лечение. За последние 10 лет наблюдается быстрый рост исследований, использующих машинное обучение (ML) и глубокое обучение (DL) для автоматизации и улучшения диагностики BC на данных изображений и клинико-лабораторных данных. Обоснование: современные методы дают возможность извлечь тонкие паттерны, недоступные визуальной оценке, а также строить персонализированные риск-прогнозы.

Источники данных и используемые модальности в исследованиях рака молочной железы отличаются разнообразием и дополняют друг друга, обеспечивая комплексный подход к диагностике и прогнозированию заболевания.

Наиболее распространённой модальностью является маммография, применяемая в рамках скрининговых программ. Существуют популярные открытые базы данных, такие как DDSM, INbreast и MIAS, которые широко используются для обучения и тестирования алгоритмов компьютерного анализа. Помимо этого, нередко применяются собственные («in-house») наборы изображений, а также объединённые датасеты, включая агрегированные коллекции с платформы Kaggle. Для задач классификации опухолей по табличным признакам также используются базы наподобие UCI Breast Cancer Wisconsin, содержащие числовые характеристики, полученные, например, из данных тонкоигольной аспирационной биопсии (FNA features).

Ультразвуковая диагностика (УЗИ) применяется как дополнение к маммографии, особенно в случаях высокой плотности тканей, когда рентгеновские изображения менее информативны.

Магнитно-резонансная томография (МРТ) и позитронно-эмиссионная томография, совмещённая с компьютерной томографией (ПЭТ/КТ), предоставляют дополнительную функциональную и морфологическую информацию. Эти модальности активно используются в исследованиях по радиомике для прогнозирования чувствительности опухоли к терапии и оценки прогноза заболевания.

(cc) BY

Отдельную категорию составляют гистопатологические изображения, которые применяются для анализа тканей на клеточном уровне. Такие изображения обладают высоким разрешением и требуют тщательной сегментации для выделения патологических участков.

Кроме изображений, важную роль играют клинические и молекулярные данные, включая возраст, медицинский анамнез, результаты генетических тестов и другие параметры. В современных подходах эти данные часто комбинируются с визуальными признаками, формируя мультимодальные модели, способные обеспечивать более точную и информативную диагностику.

Современные исследования в области применения искусственного интеллекта

Современная литература и эмпирические исследования свидетельствуют о стремительном развитии методов искусственного интеллекта в области диагностики рака молочной железы, охватывая широкий спектр подходов – от радиомики и классического машинного обучения до глубоких нейронных сетей и клинических испытаний.

Современные исследования в области применения искусственного интеллекта (ИИ) и радиомики для диагностики рака молочной железы демонстрируют значительный прогресс – как в разработке алгоритмов, так и в создании клинически применимых решений. Ниже представлены ключевые работы, оказавшие существенное влияние на развитие этого направления.

В работе L.Wang и др. [1] представлен комплексный анализ применения методов глубокого обучения (Deep Learning, DL) для задач автоматизированного распознавания патологий на маммографических изображениях. Автор демонстрирует высокую эффективность сверточных нейронных сетей в выявлении злокачественных образований и обсуждает возможности интеграции таких моделей в клинические рабочие процессы, подчёркивая необходимость мультицентричной валидации и интерпретируемости алгоритмов.

Исследование Y.J.Qi и др. [2] обобщает современные достижения в области радиомики, рассматривая количественное извлечение признаков из изображений как ключевой элемент прецизионной онкологии. Работа акцентирует внимание на потенциале радиомики для прогнозирования ответов на терапию и формировании персонализированных стратегий лечения, а также на проблемах стандартизации и интеграции радиомических данных с геномными и клиническими источниками.

В обзоре C.R.Taylor и др. [3] рассмотрены актуальные направления применения искусственного интеллекта в маммографическом скрининге, включая методы сегментации, классификации и риск-моделирования. Автор отмечает значительный рост интереса к гибридным и мультимодальным архитектурам, а также подчёркивает важность доказательной базы для их внедрения в клиническую практику.

К числу фундаментальных открытых ресурсов относится UCI Machine Learning Repository – Breast Cancer Wisconsin (Diagnostic) [4], который широко используется в качестве эталонного набора данных для обучения и тестирования моделей машинного обучения. Этот датасет представляет собой табличную форму признаков, извлечённых из микроскопических изображений клеток, и служит важной платформой для исследований в области классификации опухолей молочной железы.

Одним из первых масштабных подтверждений эффективности глубокого обучения в онкопатологии стало исследование B.E.Bejnordi и др. [5], посвящённое диагностической оценке алгоритмов для детекции метастазов в лимфатических узлах у женщин с раком молочной железы. Используя сверточные нейронные сети (CNN) и данные соревнования CAMELYON16, авторы продемонстрировали, что алгоритмы глубокого обучения могут достигать, а в отдельных случаях – превосходить точность патологов при обнаружении метастазов на

цифровых гистологических слайдах. Это исследование стало важной вехой, подтвердившей практическую жизнеспособность ИИ в поддержке решений патологоанатомов.

Само соревнование CAMELYON16 (2016), организованное как международный бенчмарк по детекции метастазов на цельнослайдовых изображениях (Whole-Slide Images, WSI), заложило основу для систематического сравнения алгоритмов. Привлекая десятки исследовательских команд, оно продемонстрировало потенциал машинного обучения для автоматизированного анализа цифровой патологии и сформировало эталон для последующих исследований.

Крупное мультицентровое исследование McKinney и др. [5] стало одним из первых международных примеров масштабной валидации систем искусственного интеллекта для маммографического скрининга. Обученная на сотнях тысяч маммограмм из разных стран, сверточная нейронная сеть показала сопоставимую или лучшую эффективность по сравнению с врачами-рентгенологами, одновременно снижая количество ложноположительных и ложноотрицательных заключений. Работа подчеркнула необходимость дальнейших клинических испытаний и стандартизации процедур валидации.

В исследовании Rodríguez-Ruiz и др. [7] было показано, что использование ИИ-помощника в процессе анализа маммограмм способствует повышению чувствительности радиологов при выявлении опухолей, не увеличивая при этом существенно время чтения. Этот результат подтвердил практическую пользу ИИ в поддержке клинического принятия решений.

Работа A. Yala и др. [8] представила мультимодельный подход, объединяющий признаки, извлечённые CNN из маммографических изображений, с традиционными клиническими факторами для прогнозирования риска развития рака. Модель показала превосходство над классическими риск-шкалами, что продемонстрировало эффективность интеграции изображений и клинических данных в единую систему прогнозирования.

Исследование R.S.Lee и др. [9], представившее CBIS-DDSM, стало важным шагом в развитии открытых данных для обучения систем компьютерной диагностики. Курированный и размеченный набор маммограмм обеспечил стандартизированную базу для разработки и сравнения алгоритмов, стимулировав развитие CAD (Computer-Aided Detection) и DL-моделей для маммографии.

Работа Campanella и др. [10] продемонстрировала возможность обучения клинически применимых моделей патологии с использованием слабонаблюдаемого обучения (weakly supervised / multiple-instance learning). Модели обучались на сотнях тысяч цельнослайдовых изображений без пиксельной разметки, используя лишь диагноз как метку, что позволило существенно сократить трудозатраты на аннотацию и расширить масштабы цифровой патологии, включая рак молочной железы.

Методологический вклад в развитие радиомики внесли P.Lambin и др. [11], представившие концепцию радиомического подхода как моста между медицинской визуализацией и персонализированной медициной. Авторы систематизировали основные этапы радиомического пайплайна – от предобработки изображений до построения прогностических моделей, подчеркнув значение стандартизации и валидации.

В последующем обзоре Conti и др. [12] была обобщена информация о применении радиомики для диагностики, прогнозирования и мониторинга рака молочной железы на основе различных модальностей (маммография, МРТ, УЗИ, ПЭТ). Авторы систематизировали доказательства эффективности радиомики, отметив при этом существующие ограничения – небольшие выборки и отсутствие единых протоколов.

Более современный обзор Qi Y.J. и др. [2] представили анализ достижений радиомики за 2020–2024 годы, подчеркнув развитие мультимодальных и радиогеномных подходов,

а также тенденцию к мультицентровой стандартизации и клинической интеграции.

Исторически важным источником для тестирования алгоритмов машинного обучения стал датасет Breast Cancer Wisconsin (Diagnostic), опубликованный Wolberg и др. [13] и включённый в UCI Machine Learning Repository. Несмотря на относительную простоту и табличный формат данных (характеристики ядер клеток из тонкоигольных аспирационных образцов), этот набор послужил классическим бенчмарком для ранних исследований и стал отправной точкой в развитии алгоритмов диагностики на основе структурированных признаков.

Наконец, в отчётах и новостях последних лет отражены значимые регуляторные и клинические события. В частности, в 2025 году сообщается о крупном клиническом испытании NHS (National Health Service, Великобритания), направленном на оценку эффективности систем искусственного интеллекта в реальных условиях диагностики рака молочной железы. Кроме того, FDA зарегистрировало ряд новых AI-инструментов для оценки риска по маммограммам, что свидетельствует о постепенном переходе технологий от исследовательской стадии к клиническому применению.

Совокупность этих источников подтверждает, что развитие искусственного интеллекта в маммографии находится на этапе активной трансляции научных достижений в клиническую практику, сопровождающейся ростом доказательной базы, регуляторных инициатив и международного сотрудничества. В совокупности эти исследования формируют научную и практическую основу современного направления ИИ в маммографической диагностике и цифровой патологии – от фундаментальных радиомических принципов до клинически валидированных решений, готовых к интеграции в медицинскую практику.

Этапы обработки данных и извлечения признаков

Этапы обработки данных и извлечения признаков включают несколько последовательных шагов, направленных на подготовку изображений и получение информативных характеристик для последующего анализа.

На первом этапе проводится предобработка изображений, которая включает нормализацию интенсивности, подавление шума, выравнивание и регистрацию, а также коррекцию артефактов. Для маммограмм дополнительно выполняется удаление меток и посторонних элементов, маскирование областей, не относящихся к исследуемой зоне (например, уха или рамки), а также приведение изображений к единому разрешению.

Далее выполняется сегментация, то есть выделение области интереса (ROI, region of interest). Этот процесс может быть ручным, полуавтоматическим или полностью автоматическим. В последнее время широко применяются нейросетевые методы, такие как архитектура U-Net и её модификации, позволяющие точно отделять опухолевые области от здоровых тканей.

Следующий этап – радиомика (radiomics), которая представляет собой извлечение большого числа количественных признаков из изображений. Эти признаки могут описывать форму, текстуру, распределение интенсивности (гистограммы) и различные фильтровые характеристики. Радиомика позволяет преобразовать изображение в табличный набор числовых признаков, пригодный для применения классических алгоритмов машинного обучения.

Помимо этого, широко используются глубокие признаки, извлекаемые с помощью современных архитектур глубоких нейронных сетей, таких как CNN и Transformer. Эти модели автоматически формируют информативные представления изображений, зачастую превосходящие традиционные радиомические признаки. Для повышения эффективности нередко применяется подход transfer learning – использование моделей, предварительно обученных на больших наборах данных (например, ImageNet), с последующим дообучением на

специализированных медицинских изображениях.

Модели прогнозирования и классификации

Модели прогнозирования и классификации, применяемые в медицинской диагностике, охватывают широкий спектр методов машинного обучения – от классических статистических подходов до современных архитектур глубокого обучения.

К числу классических методов машинного обучения (ML) относятся логистическая регрессия, метод опорных векторов (SVM), случайный лес (Random Forest) и алгоритмы градиентного бустинга, такие как XGBoost и LightGBM. Эти методы характеризуются высокой интерпретируемостью, устойчивостью к переобучению на ограниченных выборках и способностью эффективно работать при наличии качественно извлечённых признаков. Их преимущество заключается в возможности анализа структурированных данных и построения прозрачных моделей, что особенно важно в медицинских приложениях, где объяснимость решений имеет ключевое значение.

Методы глубокого обучения (Deep Learning, DL), в частности сверточные нейронные сети (CNN) – такие как ResNet, VGG, DenseNet, – обеспечивают автоматическое извлечение признаков из изображений и демонстрируют высокую эффективность при анализе больших объемов данных. В последние годы получили развитие гибридные архитектуры, сочетающие CNN и Transformer, что позволяет учитывать как локальные, так и глобальные зависимости в изображениях. Для задач сегментации и локализации патологических областей широко применяются специализированные архитектуры, включая U-Net и Mask R-CNN, обеспечивающие точное выделение контуров опухолевых структур.

Особое внимание уделяется разработке мультимодальных и ансамблевых моделей, которые объединяют различные типы данных: клинические параметры, радиомические признаки и представления, извлечённые нейросетевыми моделями. Интеграция разнородных источников информации позволяет значительно повысить точность прогнозирования, улучшить стратификацию пациентов по рисковым группам и повысить достоверность предсказаний клинических исходов. Такой подход способствует созданию более устойчивых и обобщающих моделей, способных адаптироваться к сложным биомедицинским сценариям.

Валидация моделей и метрики

Процесс валидации моделей и оценки их эффективности является неотъемлемой частью разработки и внедрения систем искусственного интеллекта в медицинскую диагностику. Корректная оценка качества предсказаний обеспечивает достоверность результатов и позволяет определить пригодность модели для практического применения.

Для количественной оценки производительности прогнозных моделей используются различные метрики, отражающие как общую точность, так и качество отдельных аспектов классификации. Наиболее распространёнными являются площадь под кривой ошибок (AUC-ROC), площадь под кривой Precision-Recall (AUPRC), точность (accuracy), чувствительность (recall, sensitivity), специфичность (specificity) и F1-мера (F1-score). В клинических приложениях особое значение придаётся чувствительности – показателю способности модели выявлять положительные случаи (например, опухоли), так как минимизация ложных отрицательных результатов напрямую связана с безопасностью пациента. Не менее важна и специфичность, характеризующая способность системы избегать ложноположительных классификаций, что позволяет снизить количество ненужных обследований и эмоциональную нагрузку на пациентов.

Этап валидации моделей включает ряд процедур, направленных на проверку устойчивости и обобщающей способности алгоритма. Наиболее часто применяются кросс-валидация (cross-validation) и внешняя

(cc) BY

валидация на независимых тестовых наборах данных (external test set). В медицинской сфере особенно важно осуществлять разделение данных на уровне пациентов (patient-level split), а не отдельных изображений, чтобы избежать утечки информации между обучающей и тестовой выборками. Для получения достоверной оценки клинической применимости рекомендуется проводить мультицентричную валидацию, включающую данные из различных медицинских учреждений и сканеров.

Дополнительным аспектом оценки является калибровка прогноза, направленная на проверку соответствия предсказанных вероятностей фактическим исходам. Для этого используются калибровочные кривые (calibration curves) и Brier score, позволяющие оценить, насколько модель корректно оценивает риск наступления события. Хорошо откалиброванные модели обеспечивают не только высокую точность классификации, но и надёжность вероятностных прогнозов, что особенно важно при принятии клинических решений.

Интерпретируемость и объяснимость

Одним из ключевых требований к применению методов искусственного интеллекта в медицине является обеспечение интерпретируемости и объяснимости моделей. В условиях клинической практики важно не только получать точные предсказания, но и понимать, почему алгоритм принял то или иное решение. Это необходимо для повышения доверия со стороны врачей, выявления возможных источников ошибок и обеспечения соответствия медицинским стандартам принятия решений.

Современные методы объяснимого искусственного интеллекта (Explainable AI, XAI) предоставляют различные подходы к визуализации и интерпретации внутренней логики моделей. Для нейронных сетей, применяемых при анализе медицинских изображений, широко используются методы Grad-CAM и saliency maps, позволяющие визуализировать, какие области изображения внесли наибольший вклад в решение модели. Такие карты важности помогают клиницистам убедиться, что внимание модели действительно сосредоточено на патологических участках, а не на посторонних элементах снимка.

Для моделей, основанных на табличных или радиомических признаках, применяются алгоритмы LIME (Local Interpretable Model-Agnostic Explanations) и SHAP (SHapley Additive exPlanations). Эти методы позволяют количественно оценить вклад каждого признака в формирование конкретного предсказания, обеспечивая локальную и глобальную интерпретацию поведения модели.

Дополнительно важную роль играет анализ значимости радиомических признаков, который позволяет выявить наиболее информативные характеристики, связанные с биологическими и морфологическими особенностями опухоли. Такой подход способствует не только повышению прозрачности алгоритмов, но и обнаружению потенциальных биомедицинских артефактов, например, ложных корреляций с техническими элементами изображения (метками, рамками и т.п.).

Таким образом, развитие инструментов интерпретируемости и объяснимости способствует интеграции искусственного интеллекта в клиническую практику, укрепляет доверие к автоматизированным системам и делает их использование более безопасным и обоснованным с точки зрения доказательной медицины.

Клиническая интеграция и регуляторика

Вопрос клинической интеграции и регуляторного одобрения систем искусственного интеллекта в области маммографии приобретает всё более важное значение по мере развития технологий компьютерной диагностики. За последние годы наблюдается существенный рост числа программных решений, прошедших официальные регуляторные процедуры и получивших одобрение со стороны регулирующих органов, в частности FDA (Food and Drug Administration, США). Это

отражает повышение уровня зрелости технологий и их приближение к реальному медицинскому применению.

Тем не менее, процесс внедрения искусственного интеллекта в клиническую практику остаётся сравнительно медленным. Основными барьерами являются необходимость строгого подтверждения доказательной эффективности моделей, проблемы обобщаемости (generalizability) алгоритмов при переходе на данные, полученные в других медицинских учреждениях, а также вопросы юридической и этической ответственности при ошибках автоматизированной диагностики. Для широкого внедрения требуется не только демонстрация высокой точности на ретроспективных наборах данных, но и подтверждение стабильной работы модели в условиях реальной клинической нагрузки.

С этой целью в последние годы активно проводятся проспективные исследования и клинические испытания, направленные на оценку влияния AI-инструментов на качество диагностики, скорость анализа и клинические исходы. Одним из заметных примеров являются масштабные испытания в рамках Национальной службы здравоохранения Великобритании (NHS), где оценивается эффективность систем компьютерного анализа маммограмм в условиях реальных рабочих процессов радиологов.

Таким образом, клиническая интеграция технологий искусственного интеллекта в маммографии требует комплексного подхода, включающего не только техническое совершенствование моделей, но и разработку нормативных, этических и организационных механизмов, обеспечивающих надёжность, безопасность и юридическую прозрачность их применения в медицинской практике.

Практические рекомендации для разработки и внедрения

Практические рекомендации для разработки и внедрения систем искусственного интеллекта в маммографии охватывают широкий спектр аспектов, связанных с обеспечением качества данных, корректностью построения моделей, а также их надёжной эксплуатацией в клинических условиях.

Одним из ключевых факторов успеха является качество данных и разметки. Создание достоверных обучающих выборок требует значительных инвестиций в экспертную аннотацию изображений, выполняемую несколькими независимыми специалистами – рентгенологами или патологами. Такая многократная разметка обеспечивает консенсусное суждение и снижает риск систематических ошибок. Кроме того, важно фиксировать степень уверенности аннотаторов и контекст интерпретации, что повышает надёжность используемых данных.

Важным методологическим требованием является корректное разделение данных при обучении и тестировании моделей. Следует исключать пересечение данных одного пациента между обучающей и тестовой выборками (patient-level split), чтобы предотвратить утечку информации и переоценку качества модели. Для объективной оценки обобщающей способности рекомендуется проводить мультицентричную валидацию, проверяя работу алгоритма на данных из различных клиник, полученных с использованием разных аппаратов и протоколов визуализации.

Особое внимание должно уделяться учёту популяционной неоднородности. Различия в плотности молочной ткани, этнических особенностях, параметрах оборудования и условиях съёмки могут существенно влиять на результаты работы моделей. Поэтому требуется проведение стратифицированного анализа, позволяющего оценить устойчивость и справедливость алгоритма для разных подгрупп пациентов.

Наиболее перспективным направлением считается применение мультимодальных подходов, сочетающих данные медицинской визуализации с клиническими и лабораторными параметрами. Такие модели, объединяющие изображения, радиомические признаки и клиническую информацию, демонстрируют более высокую точность и

(cc) BY

клиническую значимость по сравнению с чисто изображенческими системами.

Помимо статистических показателей (AUC, F1-score и др.), необходимо проводить оценку клинической пользы алгоритмов. Это включает анализ влияния внедрения системы на рабочий поток радиологов, время постановки диагноза, количество проводимых биопсий, а также конечные клинические исходы.

Не менее важным аспектом является интерпретируемость и удобство пользовательского интерфейса (UI). Для повышения доверия врачей следует интегрировать визуальные объяснения – карты внимания (Grad-CAM), тепловые карты вероятностей и отображение значимости признаков. Это позволяет специалистам верифицировать выводы модели и использовать их в контексте клинического суждения.

Наконец, на этапе эксплуатации системы необходимо обеспечить регулярное переобучение и мониторинг качества в производственной среде. Это включает отслеживание дрейфа данных (изменений статистических свойств входных данных) и деградации производительности модели с течением времени. Постоянная оценка и адаптация алгоритмов являются обязательным условием поддержания их клинической надёжности и соответствия текущим стандартам медицинской практики.

Текущие достижения и примеры реального внедрения

В последние годы наблюдается существенный прогресс в области применения методов глубокого обучения (Deep Learning, DL) для анализа маммографических изображений. Современные нейросетевые системы демонстрируют высокие показатели точности при решении задач детекции и классификации патологий, что подтверждается результатами многочисленных исследований и метаобзоров, опубликованных в период 2023–2025 годов. Эти работы отмечают стабильное улучшение качества алгоритмов, их способности к обобщению и повышению воспроизводимости результатов при использовании различных наборов данных и условий визуализации.

Особое значение имеют регуляторные достижения последних лет, отражающие постепенное продвижение технологий искусственного интеллекта от лабораторных прототипов к реальной клинической практике. В 2024–2025 годах зафиксированы случаи авторизации и рассмотрения регуляторными органами (включая FDA) ряда AI-инструментов, предназначенных для оценки риска развития рака молочной железы на основе маммограмм. Эти системы прошли этапы клинической валидации и начали внедряться в ограниченных медицинских учреждениях, что свидетельствует о росте доверия к искусственному интеллекту со стороны профессионального медицинского сообщества.

В то же время, несмотря на достигнутый прогресс, сохраняется необходимость в проведении доказательных клинических испытаний, направленных на подтверждение устойчивости и клинической значимости моделей в реальных условиях. Только при наличии убедительной эмпирической базы и прозрачных данных о влиянии алгоритмов на клинические исходы можно обеспечить ответственное и безопасное внедрение AI-систем в практику скрининга и диагностики.

Таким образом, текущие достижения в области глубокого обучения для маммографии указывают на переход от этапа технологического развития к фазе клинической интеграции, требующей не только совершенствования моделей, но и строгого соблюдения принципов доказательной медицины, регуляторных стандартов и этических норм.

Ограничения и вызовы

Несмотря на значительный прогресс в области применения искусственного интеллекта для анализа маммографических изображений, существует ряд ограничений и вызовов, препятствующих широкому и безопасному внедрению таких технологий в клиническую практику.

(cc) BY

Одной из ключевых проблем остаётся недостаток крупных размеченных наборов данных, охватывающих различные медицинские учреждения, оборудование и популяции пациентов. Отсутствие мультицентричных данных с клиническим наблюдением (follow-up) существенно ограничивает возможности для построения устойчивых и обобщающих моделей. Большинство существующих наборов формируются в рамках отдельных исследовательских центров и не отражают реального разнообразия клинических случаев, что повышает риск переобучения и снижает надёжность результатов при переносе моделей на новые данные.

Вторым важным вызовом является проблема биаса (смещения) и несправедливости в моделях искусственного интеллекта. Алгоритмы, обученные на несбалансированных или нерепрезентативных выборках, могут демонстрировать пониженную точность при анализе изображений пациентов из определённых подгрупп – например, по возрасту, расовой принадлежности, плотности тканей молочной железы или техническим характеристикам оборудования. Такие различия ставят под угрозу клиническую справедливость и требуют разработки механизмов адаптивного обучения и стратифицированной оценки производительности.

Не менее существенной проблемой остаются ограниченная интерпретируемость и недостаток доверия к алгоритмам со стороны клиницистов. Многие современные модели функционируют как «чёрные коробки», предоставляя высокоточные прогнозы без ясного объяснения своих решений. Это вызывает обоснованные сомнения в медицинском сообществе и препятствует принятию технологий в повседневную практику, особенно в критически значимых сценариях диагностики.

Дополнительные сложности связаны с регуляторными и юридическими аспектами, включая вопросы защиты персональных данных и медицинской конфиденциальности. Международные стандарты, такие как GDPR (General Data Protection Regulation) в Европе и HIPAA (Health Insurance Portability and Accountability Act) в США, требуют строгого соблюдения правил обработки, хранения и передачи медицинской информации. Эти нормы создают дополнительные организационные и технические барьеры при сборе и обмене клиническими данными для обучения моделей.

Наконец, важным вызовом остаётся реалистичная оценка клинической пользы. Высокие значения метрик, таких как AUC или F1-score, полученные в ретроспективных исследованиях, не всегда коррелируют с улучшением реальных исходов пациентов. Для подтверждения клинической значимости требуется проведение проспективных испытаний, в которых оценивается влияние внедрения AI-инструментов на процессы принятия решений, скорость диагностики, частоту ложных биопсий и конечные результаты лечения.

Таким образом, преодоление указанных ограничений требует системного подхода, включающего развитие инфраструктуры медицинских данных, повышение прозрачности моделей, совершенствование регуляторной базы и постоянное взаимодействие между исследователями, клиницистами и регуляторными органами.

Перспективы исследований

Современные направления развития искусственного интеллекта в маммографической диагностике определяются стремлением к созданию более интегрированных, интерпретируемых и клинически значимых систем. В ближайшие годы ключевые перспективы исследований будут связаны с развитием мультимодальных архитектур, интеграцией молекулярных данных, совершенствованием методов обучения и стандартизацией оценки эффективности.

Одним из наиболее перспективных направлений является разработка мультимодальных и трансформерных архитектур, способных агрегировать и обрабатывать информацию из различных источников – изображений, клинических

данных, гистологических и молекулярных профилей. Использование трансформеров (Transformer-based models) позволяет учитывать сложные пространственные и контекстуальные зависимости, обеспечивая более точную и устойчивую интерпретацию диагностических данных.

Особое внимание уделяется развитию радиогеномики (radiogenomics) – междисциплинарного подхода, объединяющего радиомические признаки, извлечённые из медицинских изображений, с геномными и молекулярными данными. Такой синтез открывает путь к персонализированной медицине, где прогнозы и решения по лечению адаптируются к индивидуальному биологическому профилю пациента.

Перспективным направлением также является внедрение методов онлайн-обучения (online learning) и федеративного обучения (federated learning), позволяющих обучать модели на распределённых данных из разных медицинских учреждений без необходимости их централизованного хранения. Это решение не только повышает обобщающую способность моделей, но и обеспечивает соответствие требованиям конфиденциальности и регуляторным нормам, таким как GDPR и HIPAA.

Важной задачей становится оценка системного влияния AI-технологий на здравоохранение, включая анализ их стоимостной эффективности (cost-effectiveness), влияние на клинические рабочие процессы (workflow impact trials) и качество обслуживания пациентов. Такие исследования позволяют не только определить технологическую пользу алгоритмов, но и оценить их реальный вклад в повышение эффективности здравоохранения.

Наконец, ключевое значение для дальнейшего прогресса имеет стандартизация методик и создание открытых бенчмарков. Разработка согласованных протоколов оценки, единых наборов тестовых данных и прозрачных критериев сравнения моделей позволит обеспечить честную и воспроизводимую оценку алгоритмов, способствуя объективному развитию области и повышению доверия со стороны клинического сообщества.

Таким образом, будущее исследований в области AI для маммографии связано с формированием устойчивой, интероперабельной и клинически ориентированной экосистемы, объединяющей достижения в области машинного обучения, биомедицины и цифрового здравоохранения.

ЗАКЛЮЧЕНИЕ

Интеллектуальный анализ данных кардинально расширил инструментарий диагностики рака молочной железы: от повышения точности интерпретации маммограмм до персонализированных прогнозов риска и ответа на терапию. В то же время переход от исследовательских результатов к широкому клиническому применению требует строгой валидации, внимания к справедливости и интерпретируемости, а также соответствия нормативно-правовым требованиям. Сбалансированное сочетание radiomics, глубоких нейросетей и клинических данных, вместе с мультицентровыми исследованиями – наиболее перспективный путь для безопасного и эффективного внедрения AI-инструментов в онкологическую практику.

ЛИТЕРАТУРА

1. Wang L. Mammography with deep learning for breast cancer detection. Front. Oncol. 14:1281922. 2024. 01-16. doi: 10.3389/fonc.2024.1281922
2. Qi Y.J., Su G.H., You C., et al. Radiomics in breast cancer: current advances and future directions. Cell Rep Med.2024; 5(9):101719. DOI: 10.1016/j.xcrm.2024.101719
3. Taylor CR, Monga N, Johnson C, Hawley JR, Patel M. Artificial intelligence applications in breast imaging: current status and future directions. Diagnostics (Basel) 2023;13:2041.
4. UCI Machine Learning Repository – Breast Cancer Wisconsin (Diagnostic).
5. Bejnordi, B. E., Veta, M., Van Diest, P. J., Van Ginneken, B., Karssemeijer, N., Litjens, G., ... & CAMELYON16 Consortium.

(cc) BY

(2017). Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. Jama, 318(22), 2199-2210.
6. McKinney SM, et al. International evaluation of an AI system for breast cancer screening. Nature. 2020; 577:89–94
7. Rodriguez-Ruiz A, Lang K, Gubern-Merida A et al (2019) Stand-alone artificial intelligence for breast cancer detection in mammography: comparison with 101 radiologists. J Natl Cancer Inst 111:916–922
8. Yala, A. · Lehman, C. · Schuster, T. et al. A deep learning mammography-based model for improved breast cancer risk prediction Radiology. 2019; 292:60-66.
9. Rebecca Sawyer Lee, Francisco Gimenez, Assaf Hoogi, Kanae Kawai Miyake, Mia Gorovoy, and Daniel L Rubin. A curated mammography data set for use in computeraided detection and diagnosis research. Scientific Data, 4:170177, 2017. ISSN 2052-4463. doi: 10.1038/sdata.2017.177. URL https://doi.org/10.1038/sdata.2017.177.
10. Campanella, G., Hanna, M.G., Geneslaw, L. et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. Nat Med 25, 1301–1309 (2019). https://doi.org/10.1038/s41591-019-0508-1
11. Lambin, P., et al. (2017). Radiomics: the bridge between medical imaging and personalized medicine. Nature Reviews Clinical Oncology, 14(12), 749-762. https://doi.org/10.1038/nrclinonc.2017.141
12. Conti A, Duggento A, Indovina I, Guerrisi M, Toschi N. 2021. Radiomics in breast cancer classification and prediction. Semin. Cancer Biol. 72:238–50
13. Wolberg,William, Mangasarian, Olvi, Street,Nick, and Street,W.. (1995). Breast Cancer Wisconsin (Diagnostic). UCI Machine Learning Repository. https://doi.org/10.24432/C5DW2B.
14. Reports on regulatory/clinical trials and news: NHS trial of AI for breast cancer diagnosis (2025); FDA authorizations and recent AI risk tools (2025).