Наборы данных для маммографии и искусственный интеллект для выявления рака молочной железы: проблемы, смещение и будущие направления

Источник: Frontiers in AI — Medicine

Оригинал: https://www.frontiersin.org/articles/10.3389/frai.2026.1933770

диагностика | машинное обучение | наборы данных | онкология | радиология | справедливость ИИ

Введение

Рак молочной железы по-прежнему остается одной из основных причин смертности от онкологических заболеваний среди женщин во всем мире, и раннее выявление имеет решающее значение. Маммография является основным методом визуализации для скрининга рака молочной железы, а недавние достижения в области искусственного интеллекта (ИИ) значительно улучшили автоматизированное выявление рака молочной железы. В данном обзоре рассматривается подход, ориентированный на данные, анализируются ключевые характеристики общедоступных наборов данных, вклад данных маммографии в разработку ИИ для выявления рака молочной железы, а также последствия для клинической обобщаемости.

Несмотря на то, что ИИ играет важную роль, этот обзор фокусируется на данных, критически анализируя характеристики общедоступных наборов данных маммографии и роль данных маммографии в разработке, производительности и клинической обобщаемости моделей ИИ для выявления рака молочной железы. Наборы данных сравниваются по

следующим аспектам: размер набора данных, качество аннотаций, метод визуализации, клинические метаданные, доступность набора данных и демографическое разнообразие.

Анализ показывает, что качество, разнообразие и репрезентативность используемых данных имеют критическое значение для производительности и обобщаемости моделей ИИ, охватывая период от традиционного машинного обучения до современной эпохи глубокого обучения, систем на основе трансформеров, мультимодального обучения, объяснимого ИИ и обучения с сохранением конфиденциальности.

Заключение

Результаты показывают, что помимо разработки архитектур моделей ИИ, наличие стандартизированных, разнообразных и клинически репрезентативных наборов данных является необходимостью для дальнейшего прогресса в области ИИ-ассистированной маммографии. В обзоре подчеркивается важность создания всеобъемлющего эталонного набора данных для маммографии, который был бы большим и разнообразным, чтобы повысить воспроизводимость результатов, обеспечить справедливую оценку моделей и разработать надежные, клинически жизнеспособные системы ИИ для выявления рака молочной железы.

Перевод выполнен: 31.10.2026 | ai4med.ru

Машинный перевод. Рекомендуем сверять с оригиналом при клиническом использовании.