Диагностика редких генетических заболеваний по изображениям лица с использованием визуального промптинга и мультимодальных больших языковых моделей: исследование диагностической точности

Источник: JMIR Medical Informatics

Дата публикации: 2024-11-04

Оригинал: https://medinform.jmir.org/2026/1/e105860

генетика | диагностика | компьютерное зрение | мультимодальные LLM | редкие заболевания

Предпосылки: Редкие генетические заболевания часто характеризуются узнаваемыми фенотипическими особенностями лица. Неясно, могут ли мультимодальные большие языковые модели (MLLM) приоритизировать диагностику редких генетических заболеваний на основе изображений лица, а также улучшают ли этот процесс повышение разрешения изображений или визуальные подсказки. **Цель:** Данное исследование было направлено на оценку диагностической точности MLLM для редких генетических заболеваний с использованием изображений лица, а также на оценку того, улучшает ли сверхразрешение (super-resolution) или визуальные подсказки в виде красных кругов производительность моделей. **Методы:** В данном диагностическом исследовании точности использовались 4746 изображений лица из базы данных GestaltMatcher (GMDB), включая 882 изображения в подмножестве GMDB-Frequent (заболевания, имеющие не менее 6 изображений) и 3864 изображения в подмножестве GMDB-Rare (заболевания, имеющие менее 6 изображений). Три локально развернутые открытые модели MLLM с большим количеством параметров — Gemma-4-31B,

Qwen3.6-35B-A3B и Qwen3.8-27B — генерировали список из 50 ранжированных кандидатов на диагноз в условиях использования исходных изображений, изображений с улучшением через CodeFormer, изображений с визуальной подсказкой в виде красных кругов, сгенерированных с помощью Qwen-Image-Edit-2511, изображений с несоответствующими красными кругами, а также в рамках двухэтапного процесса преобразования изображения в фенотип. Кандидаты на диагноз были сопоставлены с идентификаторами онтологии заболеваний MONDO (Mondo Disease Ontology). Кроме того, мы оценивали случайный базовый уровень, качество визуальных подсказок, области, аннотированные экспертами, стабильность при повторных выводах, уверенность, сообщаемая моделями, перплексивность списка диагнозов, пересечение предсказаний и состав заболеваний. **Результаты:** Точность в топ-50 варьировалась от 13,95% до 18,25% в подмножестве GMDB-Frequent и от 0,57% до 1,35% в подмножестве GMDB-Rare. Визуальные подсказки в виде красных кругов привели к численному увеличению точности в 5 из 6 сравнений моделей и подмножеств. Модель Qwen3.8-27B показала значительное улучшение в подмножестве GMDB-Rare, при этом точность увеличилась с 0,57% до 1,27% (скорректированный P=.009). В подгруппе с качественными визуальными подсказками в GMDB-Frequent модель Gemma-4-31B показала значительные улучшения по сравнению с условием сверхразрешения в обоих анализах, оцененных врачами, с увеличением точности на 5,76 процентных пункта (скорректированный по методу Хольма P=.005) и на 4,46 процентных пункта (скорректированный по методу Хольма P=.048) соответственно. Вводные данные с сопоставленными красными кругами обеспечивали более высокую точность по сравнению с несопоставленными вводами во всех сравнениях моделей и подмножеств, при этом скорректированный P>.05. Рабочий процесс преобразования изображения в фенотип улучшил точность для Gemma-4-31B в GMDB-Frequent и для Qwen3.8-27B в GMDB-Rare. В подмножестве GMDB-Rare случайный базовый уровень достиг точности 8,15% и превзошел все 3 модели в условиях исходных изображений и визуальных подсказок в виде красных кругов. Наборы предсказаний оставались существенно перекрывающимися после замены изображений пациентов, а предсказания в топ-1 концентрировались на небольшом числе заболеваний. **Выводы:** В рамках оцененных экспериментальных настроек три MLLM показали ограниченные способности к диагностике редких генетических заболеваний на основе изображений лица, особенно заболеваний из подмножества GMDB-Rare. Визуальные подсказки и диагностика на основе преобразования изображения в фенотип принесли пользу в конкретных моделях и подгруппах. Анализ пересечения предсказаний, несопоставленных изображений и состава заболеваний указывает на то, что выходные данные

моделей слабо ограничивались информацией о лице конкретного пациента. Дальнейшие исследования должны оценить сравнения с врачами, проспективные данные из нескольких центров и мультимодальные диагностические рабочие процессы, основанные на фенотипах.

Перевод выполнен: 04.11.2026 | ai4med.ru

Машинный перевод. Рекомендуем сверять с оригиналом при клиническом использовании.