

Классификация педиатрических стоматологических заболеваний по панорамным рентгенограммам с использованием трансформера естественного языка и моделей глубокого обучения

Источник: Frontiers in AI — Medicine

Оригинал: <https://www.frontiersin.org/articles/10.3389/frai.2026.1754498>

NLP

глубокое обучение

диагностика

компьютерное зрение

радиология

стоматология

Введение

Точная классификация педиатрических стоматологических заболеваний по панорамным рентгенограммам имеет решающее значение для ранней диагностики и эффективного планирования лечения. Хотя модели глубокого обучения традиционно работают непосредственно с изображениями, текстовые представления, сгенерированные на основе рентгенограмм, могут представлять альтернативную стратегию для классификации заболеваний.

Методы

В данном исследовании был предложен текстово-ориентированный подход, в котором трансформатор естественного языка использовался для генерации структурированных текстовых описаний из панорамных рентгенограмм. Эти описания впоследствии классифицировались для бинарной классификации заболеваний с использованием трех архитектур глубокого обучения: одномерной сверточной нейронной сети (1D-CNN), сети долгой краткосрочной памяти (LSTM) и предобученной модели двунаправленных

энкодерных представлений из трансформера (BERT). Производительность моделей оценивалась и сравнивалась с тремя предобученными сверточными нейронными сетями, обученными непосредственно на рентгенографических изображениях.

Результаты

1D-CNN показала наивысшую производительность с точностью 84%, продемонстрировав сбалансированную классификацию по категориям заболеваний. Модель BERT достигла точности 77%, показав сильную производительность в выявлении периапикальных инфекций, но сравнительно более низкую чувствительность при идентификации кариеса. Модель LSTM показала значительно худшие результаты, достигнув точности 57%. Оба текстовых подхода на основе 1D-CNN и BERT превзошли три предобученные модели CNN на основе изображений.

Обсуждение

Эти результаты указывают на то, что текстовая классификация панорамных рентгенограмм является потенциальной альтернативой традиционным методам глубокого обучения на основе изображений. Языково-ориентированные модели демонстрируют перспективы для интерпретации рентгенограмм; однако остаются проблемы с достижением последовательной обобщаемости по типам заболеваний. Будущие исследования должны быть сосредоточены на улучшении качества генерации описаний «рентгенограмма-текст», разработке гибридных архитектур, интегрирующих текстовые и визуальные признаки, а также на валидации производительности на более крупных и разнообразных наборах данных для укрепления клинической применимости.