

Поперечная стратификация уязвимости к падениям при болезни Паркинсона через мультимодальную интеграцию речи, походки, мимики и движений верхней конечности

Источник: Frontiers in Digital Health

Оригинал: <https://www.frontiersin.org/articles/10.3389/fdgth.2026.1937343>

болезнь Паркинсона

глубокое обучение

диагностика

мобильность

мультимодальный ИИ

принципы работы моделей

цифровые биомаркеры

Падения являются значимой причиной заболеваемости при болезни Паркинсона (БП), однако текущие подходы к поперечной оценке уязвимости, связанной с падениями, часто опираются на униформодальные показатели мобильности с ограниченной способностью характеризовать многомерные проявления этого состояния. В данном исследовании представлена архитектура мультимодального глубокого обучения (DL), которая включает мимику, голос, динамику походки и движения верхних конечностей для классификации поперечной уязвимости, связанной с падениями, при БП, используя текущий статус постуральной нестабильности и историю падений. Данные были собраны от 147 пациентов, и модальность-специфичные модели были систематически прунированы (оптимизированы путем удаления избыточных параметров) для уменьшения размера модели, задержки вывода (инференса) и вычислительных требований при сохранении классификационной производительности, что обеспечило первоначальную оценку вычислительной эффективности в экспериментальной среде. Были оценены различные стратегии слияния данных, включая позднее, промежуточное, основанное на внимании (attention-based), основанное на трансформерах (transformer-based) и с помощью больших языковых моделей (LLM) методы, с использованием кросс-валидации с исключением одного

субъекта (leave-one-subject-out). Анализ с помощью больших языковых моделей был исследовательским и основывался на одном зафиксированном запуске API для каждой конфигурации «участник-модальность». Прунинг привел к значительному улучшению эффективности без заметного снижения точности, при этом походка продемонстрировала наивысшую устойчивость. Производительность слияния варьировалась в зависимости от методов и комбинаций модальностей. Слияние на основе внимания (attention-based fusion) с использованием данных о лице и походке достигло наблюдаемого макро-F1 показателя (macro-f1) на уровне 91,67%. Исследовательская конфигурация слияния «лицо-верхние конечности» с помощью большой языковой модели показала наблюдаемый макро-F1 показатель на уровне 93,75% в одном зафиксированном запуске; поскольку стабильность при повторных запусках и точная воспроизводимость с использованием конкретной версии модели не были установлены, этот результат не следует интерпретировать как доказательство воспроизводимого превосходства. Слияние на основе трансформеров (Transformer fusion) показало наблюдаемый макро-F1 показатель на уровне 85,02%, когда были включены все четыре модальности. Значимо то, что речь и, в особенности, выразительность лица обеспечивали дополнительные сигналы к традиционным оценкам походки и верхних конечностей, подчеркивая потенциал недостаточно изученных аксиальных характеристик в качестве дополнительных цифровых маркеров, связанных с уязвимостью к падениям. Эти выводы поддерживают дальнейшее изучение мультимодального искусственного интеллекта для поперечной классификации уязвимости, связанной с падениями, при БП. Прогностическое предсказание падений, внешняя обобщаемость и производительность на целевых устройствах остаются задачами для оценки в независимых будущих исследованиях.