

Мультимодальная классификация кожных поражений на основе легкой сети гиперслияния с затвором низкого ранга

Источник: Frontiers in Digital Health

Оригинал: <https://www.frontiersin.org/articles/10.3389/fdgth.2026.1947161>

дерматология

диагностика

интерпретируемость

мультимодальный ИИ

онкология

сверточные нейронные сети

Рак кожи создает значительную глобальную нагрузку на общественное здравоохранение, а ранняя, неинвазивная и точная классификация поражений является критически важным условием для улучшения долгосрочных показателей выживаемости пациентов. В рутинной дерматологической практике широко применяется мультимодальная интеграция дерматоскопических изображений и клинических метаданных (включая возраст пациента, пол и анатомическую локализацию поражения). Тем не менее, современные передовые рамки мультимодального слияния все еще сталкиваются с рядом нерешенных критических узких мест при развертывании в реальной клинической практике. Эти неустраняемые ограничения в первую очередь включают негативный перенос между модальностями, переобучение в условиях ограниченного количества аннотированных выборок, недостаточное взаимодействие признаков между модальностями и «черный ящик» природы принятия решений моделью; все эти факторы в совокупности снижают операционную надежность и реализуемость развертывания существующих мультимодальных диагностических рамок в реальных условиях.

Для системного решения этих нерешенных задач в данной работе представлена сеть гиперслияния с затвором низкого ранга (Gated Low-Rank Hyper-Fusion Network, GLHF-Net) для автоматической классификации поражений кожи. Предложенная архитектура использует легковесный вариант Swin Transformer в качестве основы для извлечения визуальных признаков, построенную на базе двухэтапного конвейера динамического мультимодального слияния. Во-первых, механизм канального затвора, управляемый метаданными (Metadata-Driven Channel-wise Gating Mechanism, Meta-Gate), адаптивно подавляет шум на уровне каналов, не имеющий отношения к задаче, в высокоразмерных картах визуальных признаков, смягчая расхождение распределений между табличными клиническими метаданными и латентными вложениями изображений. Во-вторых, встроенный модуль низкорангового гипердинамического слияния (Low-Rank Hyper-Dynamic Fusion, LR-HDF) выполняет ограниченную реконструкцию глубоких признаков при явном регуляризаторе низкого ранга, эффективно снижая риск эмпирического переобучения на небольших аннотированных дерматологических наборах данных.

Обширные эксперименты по валидации, проведенные на общедоступном эталонном наборе данных HAM10000, демонстрируют, что предложенная модель GLHF-Net достигает надежной диагностической точности с площадью под ROC-кривой (AUC) 0,9200, площадью под частичной ROC-кривой (pAUC) 0,8405, чувствительностью (Sensitivity) 0,8935 и отрицательной прогностической ценностью (NPV) 0,9682, а также демонстрирует высокую конкурентоспособность по сравнению как с методами обычного конкатенирования признаков, так и с методами, основанными на сложных механизмах внимания. Результаты визуализации Grad-CAM подтверждают, что предложенная модель может точно локализовать основные области поражений, эффективно подавляя помехи, такие как закрывающие волосы, что предоставляет важнейшие интерпретируемые доказательства клинической надежности модели. Обладая сильной способностью противостоять переобучению и высокой скоростью вывода (121,1 кадра в секунду на GPU и 20,35 кадра в секунду на CPU), эта рамка предоставляет надежное легковесное решение для развертывания интеллектуального скрининга рака кожи в клинических условиях.

Перевод выполнен: 08.10.2026 | ai4med.ru

Машинный перевод. Рекомендуем сверять с оригиналом при клиническом использовании.