

Легкая гибридная архитектура глубокого обучения для обнаружения нескольких таблеток, распознавания их атрибутов, анализа оттисков на основе OCR и извлечения метаданных

Источник: Frontiers in AI — Medicine

Оригинал: <https://www.frontiersin.org/articles/10.3389/frai.2026.1890296>

глубокое обучение

диагностика

компьютерное зрение

медицинские технологии

фармаконадзор

Введение

Нежелательные лекарственные реакции (ADEs) остаются основной причиной предотвратимых осложнений в здравоохранении из-за неправильной идентификации таблеток, ошибок в дозировке и путаницы между визуально идентичными препаратами, особенно среди пожилых людей, лиц с нарушениями зрения и людей с ограниченной медицинской грамотностью. Последние достижения в области искусственного интеллекта и компьютерного зрения позволили создать системы автоматического распознавания таблеток. Однако многие существующие методы рассматривают обнаружение, классификацию и распознавание надписей как изолированные задачи, не предоставляя при этом облегченную единую структуру, подходящую для развертывания в сфере здравоохранения в режиме реального времени.

Методы

В данном исследовании представлена облегченная гибридная структура для многообъектного распознавания таблеток и их множественных атрибутов с целью поддержки **фармаконадзора** и систем поддержки принятия клинических решений на базе искусственного интеллекта. Предложенная структура объединяет обнаружение таблеток с использованием **YOLOv26n**, классификацию на основе **MobileNetV4-ConvSmall**, распознавание надписей на основе **EasyOCR** и извлечение метаданных в единый конвейер прогнозирования. Для классификации использовался специализированный эталонный набор данных, содержащий 400 классов таблеток, по 50 изображений для каждого класса, что в общей сложности составляет 20 000 изображений. Каждое изображение содержит две таблетки, что составляет 40 000 вручную размеченных экземпляров таблеток, использованных для обнаружения. Для оценки производительности, надежности и возможности развертывания структуры были проведены сравнительные эксперименты, функциональные сравнения с существующими системами распознавания таблеток и качественные оценки на различных устройствах.

Результаты

Модель **YOLOv26n** достигла точности (precision) 0,963, полноты (recall) 0,982, **mAP@50** 0,989 и **mAP@50-95** 0,981, в то время как **MobileNetV4-ConvSmall** достигла точности классификации классов 99,60% с более высокими показателями точности определения формы и цвета таблеток. Сравнительный анализ оптического распознавания символов (**OCR**) показывает, что **EasyOCR** обеспечил более надежную работу по распознаванию надписей, чем **TesseractOCR** и **TrOCR**. Сравнительная оценка структуры продемонстрировала, что комбинация **YOLOv26n** с **MobileNetV4-ConvSmall** обеспечила высокую общую производительность распознавания, достигнув точности **Top-1** 89,83% и точности **Top-5** 99,49%. Качественная оценка на различных устройствах с использованием камер смартфонов и прямой загрузки файлов на ноутбук продемонстрировала стабильную работу в условиях различного уровня шума на изображениях.

Обсуждение

Предложенная структура демонстрирует потенциал облегченных архитектур глубокого обучения для масштабируемой идентификации лекарственных средств и приложений фармаконадзора с поддержкой ИИ. Благодаря интеграции обнаружения, классификации, распознавания надписей и

извлечения метаданных в единую структуру, предлагаемая система обладает потенциалом для поддержки идентификации лекарств и улучшения рабочих процессов вспомогательной медицинской помощи. Окончательное клиническое решение и проверка лекарственного препарата должны оставаться под контролем медицинского работника.

Перевод выполнен: 20.07.2026 | ai4med.ru

Машинный перевод. Рекомендуем сверять с оригиналом при клиническом использовании.