

АРХИТЕКТУРЫ ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ АНАЛИЗА МЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ

Ходжиметов Э. Ш., Захарьев В. А.

Кафедра интеллектуальных информационных технологий, Кафедра систем управления
Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {khodjimetovelyor05012003pr}@gmail.com

Представлен подход к анализу медицинских изображений на основе моделей глубокого обучения и современных архитектур компьютерного зрения. Использование сверточных нейронных сетей и трансформерных моделей позволяет повысить точность диагностики патологий на рентгеновских, МРТ и КТ-снимках. Описаны программные средства обработки изображений и результаты тестирования на открытых медицинских датасетах.

ВВЕДЕНИЕ

Анализ медицинских изображений является одной из ключевых задач современной медицины, так как позволяет своевременно выявлять заболевания и контролировать их развитие. Традиционные методы диагностики, основанные на визуальном анализе снимков врачом-специалистом, подвержены субъективности и ограничены физическими возможностями человека. Развитие технологий машинного обучения и компьютерного зрения позволило автоматизировать процесс обработки изображений и повысить точность диагностических решений. Использование нейронных сетей обеспечивает выделение скрытых информативных признаков, недоступных для классических алгоритмов фильтрации и сегментации [1].

I. АРХИТЕКТУРЫ И МЕТОДЫ АНАЛИЗА

Современные методы автоматизированной интерпретации медицинских изображений сочетают достижения компьютерного зрения, глубокого обучения и медицинской визуализации. За последнее десятилетие произошёл переход от классических методов с ручным проектированием признаков к полностью обучаемым моделям на основе глубоких нейронных сетей – прежде всего свёрточных (CNN) и трансформеров (ViT).

CNN остаются основой благодаря эффективному извлечению локальных пространственно-инвариантных признаков, что идеально для обнаружения локальных аномалий (очагов, затемнений, кальцификатов). Ключевые архитектуры:

- ResNet – за счёт остаточных связей позволяют обучать очень глубокие сети, необходимые для детекции малоконтрастных (subtle) патологий;
- DenseNet – плотное соединение слоёв улучшает поток градиента и переиспользование признаков, повышая чувствительность к мелким паттернам (например, ранним интерстициальным изменениям на CXR);
- EfficientNet – compound scaling сбалансированно увеличивает глубину, ширину и раз-

решение, что критично для работы с высокодетализированными рентгенами, КТ и МРТ.

Трансформерные модели (ViT, Swin-Transformer) дополняют CNN, обеспечивая глобальный анализ изображения и лучшее распознавание сложных структур, таких как опухоли на МРТ [2].

Обработка изображений включает следующие этапы:

- нормализация и предобработка снимков;
- выделение признаков с помощью предобученных моделей;
- классификация или сегментация патологических областей.

В данной работе использована предобученная модель EfficientNet с адаптацией на медицинский датасет CheXpert, содержащий рентгеновские снимки органов грудной клетки. Наиболее перспективным направлением дальнейшего развития представляется использование ансамбля специализированных моделей, где каждая решает свою подзадачу: одна отвечает за высокоточную локализацию грубых аномалий, другая – за обнаружение малоконтрастных изменений, третья – за глобальный контекст и дифференциальную диагностику. Такое разделение ответственности позволяет повысить общую надёжность и интерпретируемость системы, а также эффективно бороться с проблемой domain shift. Принцип взаимодействия компонентов внутри общего ансамбля и последовательность этапов обработки представлены на блок-схеме алгоритма (рисунк 1).

II. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

Оценка эффективности предлагаемых методов была проведена на двух широко используемых в научном сообществе наборах данных грудных рентгенограмм – CheXpert и ChestX-ray14. Набор CheXpert включает пять ключевых целевых классов: ателектаз, кардиомегалия, консолидация, отёк лёгких и плевральный выпот. Набор ChestX-ray14, в свою очередь, содержит более ши-

рокий спектр диагнозов, среди которых для целей сопоставления в настоящей работе выделяются пневмоторакс и пневмония – патологии, традиционно представляющие особый интерес для задач первичного автоматизированного триажа.

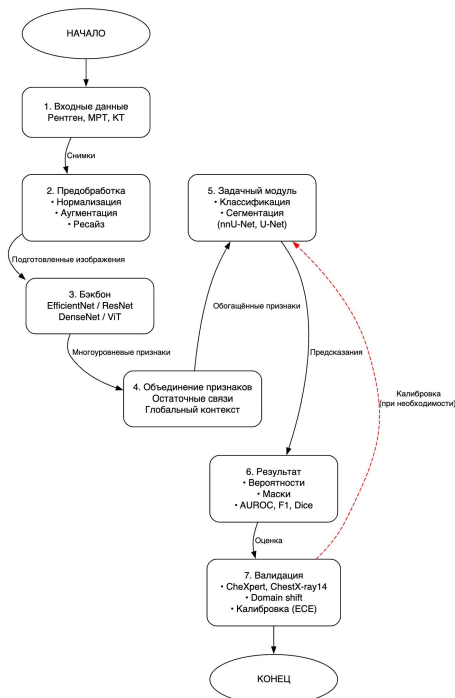


Рис. 1 – Алгоритм анализа медицинских изображений на основе ансамбля моделей

Основной метрикой оценки выбран AUROC – наиболее устойчивый показатель при дисбалансе классов и разных порогах. Дополнительно используется масго-F1, учитывающий чувствительность и точность по всем классам без весов. Это позволяет одновременно оценивать дискриминационную способность модели и её устойчивость к перекосу данных.

Согласно опубликованным результатам модели CheXzero (zero-shot/self-supervised) на тестовом наборе CheXpert, значения AUROC составляют: плевральный выпот – 0.958 (95 ДИ: 0.953–0.963), отёк – 0.961 (0.946–0.974), консолидация – 0.871 (0.851–0.888), кардиомегалия – 0.898 (0.894–0.903), ателектаз – 0.798 (0.780–0.817). Среднее значение AUROC по этим пяти классам – 0.897.

По метрике масго-F1 различия между моделью и радиологами статистически незначимы: 0.606 (95 ДИ: 0.571–0.638) против 0.619 (0.585–0.642).

Для набора ChestX-ray14 приведены результаты базовой модели ResNet-50 из оригинальной работы: пневмоторакс – 0.7891, пневмония – 0.6333. Среднее значение AUROC по этим двум классам – 0.711.

Полученные результаты показывают, что для визуально выраженных патологий (например, выпот, отёк) AUROC достигает ≈ 0.96 , тогда

как для менее однозначных находок (ателектаз, пневмония) значения заметно ниже. При внешней валидации на данных из других медицинских учреждений (domain shift) возможное снижение показателей требует проведения внешних тестов и оценки калибровки (например, по ECE или с использованием диаграмм надёжности)[6].

III. Выводы

Модели глубокого обучения продемонстрировали высокую эффективность в анализе медицинских изображений. Однако их клиническая применимость определяется не только точностью на внутренних наборах данных, но и устойчивостью к сдвигу домена, корректной калибровкой и полнотой отчётности.

Модели глубокого обучения показывают высокую эффективность в анализе медицинских изображений, но их клиническая применимость зависит не только от точности на внутренних датасетах, а от устойчивости к domain shift, правильной калибровки и полноты отчётности. Для сегментации надёжным стандартом остаются U-архитектуры и автонастраиваемый nnU-Net; перспективу межцентровой переносимости дают foundation-модели (MedSAM и аналоги).

Практическое внедрение требует внешней валидации, соблюдения TRIPOD+AI и EU AI Act (для высокорисковых систем). Рекомендуется использовать nnU-Net/MONAI как воспроизводимую основу, интегрировать трансформерные и foundation-модели, оценивать не только AUROC/Dice, но и калибровку (ECE, диаграммы надёжности), проводить внешние тесты, внедрять мониторинг drift и повторную калибровку, а также обеспечивать полную документацию и отчётность.

СПИСОК ЛИТЕРАТУРЫ

1. Ганичев П. А., Тихомирова А. А., Дохов М. А. Перспективы использования искусственного интеллекта в радиологии: краткий обзор // Визуализация в медицине. – 2023. – Т. 4, № 4. – С. 7-14.
2. Куликов А. А., Каширина И. Л., Савкина Е. Ф. Сегментация объёмных образований печени на мультифазных КТ-изображениях с использованием фреймворка nnU-Net // Моделирование, оптимизация и информационные технологии. – 2025. – Т. 13, № 1. – DOI: 10.26102/2310-6018/2025.48.1.040.
3. Лейхтер С. В. Обзор методов регистрации изображений на основе нежёстких преобразований // Математические структуры и моделирование. – 2024. – № 1(69). – С. 88–93. – DOI: 10.24147/2222-8772.2024.1.88-93.
4. Семантическая сегментация изображений с использованием методов глубокого обучения: учебные материалы. – ННГУ им. Н. И. Лобачевского, 2020. – 46 с.
5. Кочина Е. В., Швайко П. М., Лихолап О. Б. Создание и печать 3D-модели фрагмента лицевого скелета при помощи программы 3D Slicer // (материалы, описание методики; русскоязычная публикация).