VLM для диагностического анализа рентгенограмм легких

Руководитель
Крылов Андрей Серджевич

Закончил ВМК МГУ в 1978, защитил кандидатскую диссертацию в 1983, докторскую диссертацию в 2010, доктор физико-математических наук, профессор по специальности № 1.2.2. Математическое моделирование, численные методы и комплексы программ, профессор кафедры математической физики ВМК МГУ, заведующий лабораторией математических методов обработки изображений ВМК МГУ. Руководитель магистерской программы магистерской программы “Компьютерное зрение, графика и обработка изображений“ факультета ВМК МГУ и студенческого спецсеминара “Обработка изображений и компьютерное моделирование“ кафедры математической физики. Читает ряд курсов по обработке и анализу изображений, включая ежегодно обновляемый курс “Использование математических моделей в глубоком обучении”. Под его научным руководством защищены 86 дипломных работ и 14 кандидатских диссертаций. Имеет более 250 научных публикаций, включая 16 публикаций в журналах Q1 WoS. Является главным редактором журнала издательства Шпрингер “Computational Mathematics and Modeling” (Q4), членом редколлегии журнала “Machine Intelligence Research” (Q1 в 8 категориях, включая Artificial Intellegence и Computer Vision and Pattern Recognition). Входил в программные комитеты более чем 50 международных конференций и конгрессов, являясь руководителем программного комитета ряда из них. В 2024 году являлся руководителем программного комитета 9th International Conference on Biomedical Imaging, Signal Processing (ICBSP 2024), Hong Kong. Руководил большим числом грантов РНФ, РФФИ, и федеральных целевых программ, включая грант БРИКС и проектами с фирмами Intel, Samsung, Huawei.

https://istina.msu.ru/workers/396274/
https://scholar.google.com/citations?user=22qstzIAAAAJ&hl=ru&oi=ao
https://dblp.uni-trier.de/pid/24/7547.html

Зачем делаем?
Согласно «Правилам проведения рентгенологических исследований» Минздрава РФ при лучевой диагностике необходимо проводить двойное чтение рентгенограмм врачами с составлением подробного описания снимка и заключения с указанием рентгенологических синдромов, ретроспективное изучение рентгенограмм, обеспечение взаимосвязи медорганизаций при проведении рентгенологических исследований, рационализация использования ресурсов диагностических подразделений, в том числе с использованием автоматизированных систем. Применение алгоритмов анализа рентгенограмм на основе искусственного интеллекта, способных обрабатывать текстовые запросы и генерировать текстовые описания изображений, в качестве вспомогательного инструмента врача позволят снизить нагрузку на специалистов, выявлять заболевания на более ранних стадиях за счёт дополнительного этапа анализа снимков, проводить поиск изображений в базах медорганизаций по заданным признакам для более глубокого анализа и диагностики.

Что делаем?
Разрабатываем VLM-модели и алгоритмы на их основе путём создания собственных и модификации существующих кодировщиков изображений и медицинских текстов, адаптации существующих моделей для обработки рентгенограмм грудной клетки различного разрешения и российских и русскоязычных данным. Разрабатываем систему поиска изображений грудной клетки с заданными свойствами на основании текстовых запросов и похожих изображений.

Как достигнем цель?

1. Анализ существующих методов и решений

- Изучение принципов работы ведущих моделей для совместной обработки и анализа медицинскинских изображений (в особенности рентгенограмм грудной клетки) и текстов, таких как LLaVA-Med, BiomedGPT, SERPENT-VLM, XrayGPT, RadVLM, Maira-2, а также агентных систем (CheXagent, MedRAX).
- Изучение принципов работы ведущих моделей-кодировщиков изображений и текста, поддерживающих работу с изображениями разных разрешений и русским языком, каких как SigLIP
- Поиск точек для улучшения существующих методов.

2. Разработка собственных методов

- Использование архитектур на основе Transformer как основы для кодировщика рентгенограмм грудной клетки.
- Дообучение ведущих языковых моделей-кодировщиков и декодировщиков для адаптации к домену русскоязычных заключений рентгенологов. Внедрение механизмов контроля качества выходного текста, таких как Self-Refine.
- Разработка метода сцепки двух моделей для совместного кодирования визуальной и текстовой информации и обучение полученной VLM на диагностические задачи.
- Разработка метода поиска рентгенограмм и медицинскинских заключений по изображению или текстовому запросу.

3. Тестирование и валидация

Оцениваем эффективность разработанных подходов на открытых наборах данных и бенчмарках (CheXbench, ChestAgentBench и др.) и на собственных данных [Robustness analysis of chest X-ray computer tuberculosis diagnosis / Y. Pchelintsev, A. Khvostikov, O. Buchatskaia, N. Nikiforova, L. Shepeleva, E. Prokopev, L. Parolina, A. Krylov // Computational Mathematics and Modeling. — 2023. — T. 33, № 4. — C. 472—486].

Как измерить достижение результата?

- Для диагностических алгоритмов оцениваем качество сгенерированного текста с точки зрения логичности высказываний, правильности языка, точности анализа рентгенограмм с учётом Рекомендаций Минздрава РФ к покупке изделий на основе ИИ (определённый перечень рентгенологических синдромов; качество их обнаружения в терминах точности, чувствительности и специфичности).
- Для алгоритмов поиска изображений и заключений оцениваем точность поиска метриками ранжирования такими как Recall@k, Precision@k, NDCG, MAP, MRR.