Marking medical images with Natural Language Processing

1. Liver reporting directions
2. Demo visual expectations

Q1.1 Liver Dataset Categories

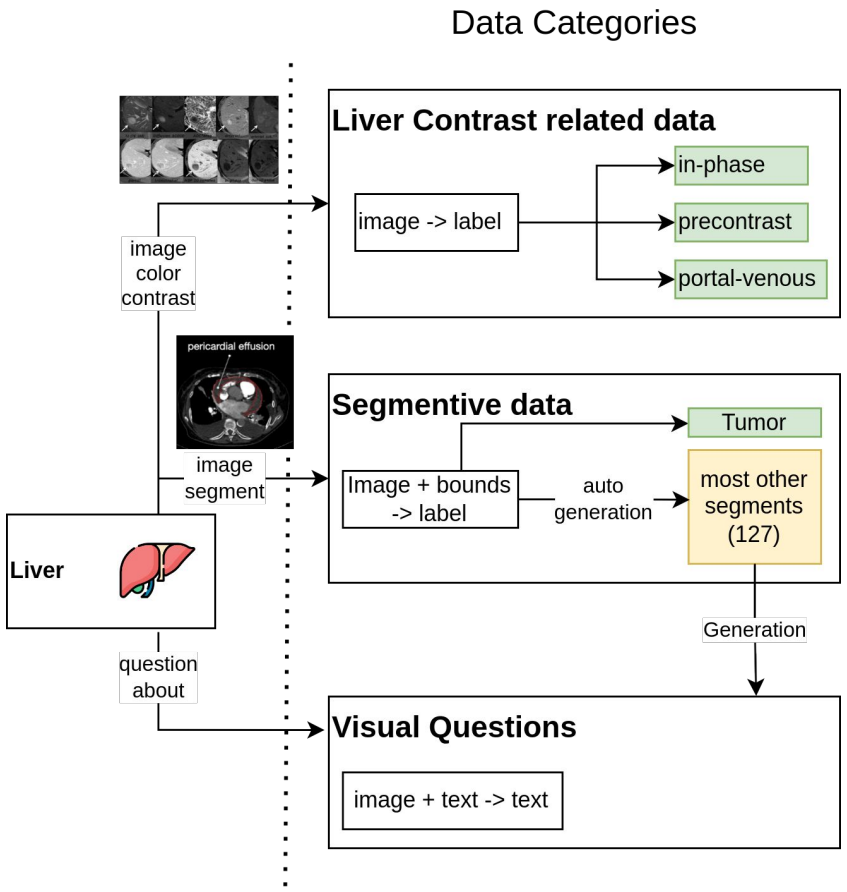
Q: How the anomalies could be revealed:

A: Three categories to be study them

1. Liver Contrast
2. Image Segments
3. Visual Matching (Questions, Reports)

Cope with missed data:

1. **Auto total segmentation [video] [tool]**
2. Questionnaire training organization (PCM-VQA)
3. Chain of thought on these questions



Q1.2 Existed Reporting Schemas

Li-RADS (2021) — provides two templates for reporting liver observations:

- in a single continuous paragraph
- in a structured format with keywords and imaging findings.

Concept of completing the report:

1. Align questions with the related report forms.
2. Use separated models for reporting, those that **aware of the reporting scheme** (for example ChatGPT-3.5 aware of Li-RADS)

Q2.1 Existed Open Models Capabilities in MMI

Showcase the comparison on how generalized open models that could be enhanced for the task could perform with image captioning out-of-the-box

Model 1. Moondream (1B) — tiniest model for experiments at small scale.

Model 2. LLaMA-1.5B

Findings:

- Existence of brief awareness of medical background from the point of view of professionals
- At present models can understand the layout of the image (data clearance)

Q2.1 Existed Proprietary Models Capabilities in MMI

Med-Gemini [advanced-paper]

HIGHLIGHT THE MAIN FEATURES

JohnSnowLab: Visual NLP, Healthcare NLP [list-of-notebooks]

These models are helpful for annotating texts due to the specific NER mentions / relations for further seeking options in external medical documents (citations)

Findings:

- **Inability to modify**
- **Domain oriented**
- **Utilized resources:**

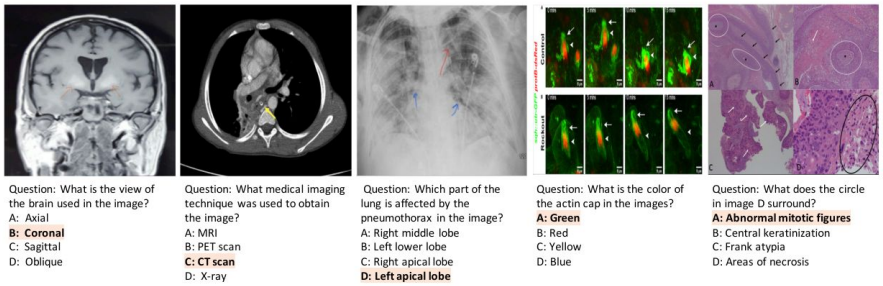
Q2.2 Prospects on combining in demo

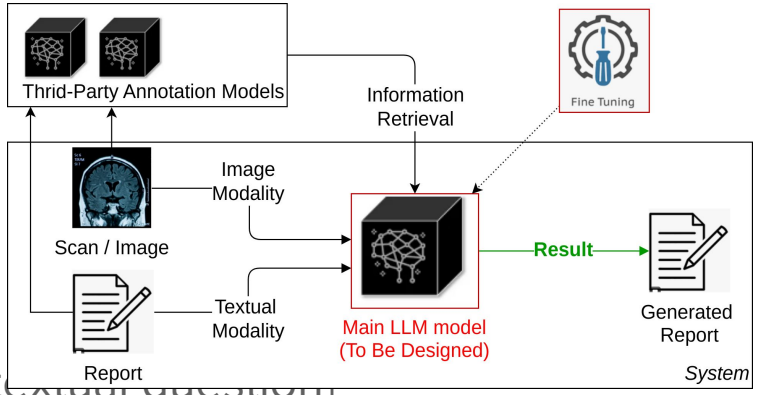
Diagram of the whole architecture.

1. (Maybe add RAG concepts for segmenta
their findings in IR and then using the most

2. Mention **TotalSegmentation** here.

3. Raise the question on the external **databases with medical articles (citation feature)**





Supported interactions:

- **Prompting** [Image + textual question]
- **Chatting** [conversation history, reasoning]

Questions

- Maybe something not covered with data concepts?

- With the Liver, can we go further in data rather than questionnaire based on image + text + segments + list of answers? *[Novelty Contribution]*

Med-GEMINI data details

Information from the following [advanced-paper]

Table 1 | Overview of the training datasets. More than 7 million data samples from 3.7 million medical images and cases is used for fine-tuning and further instruction-tuning of Gemini for medical applications in Med-Gemini. This includes diverse set of modalities including 2D and 3D radiology images, pathology, ophthalmology, and genomic data. These datasets includes mostly free text paired with medical data, which eliminates the need for expensive expert labeling of the training data.

Modality	Dataset	No. examples	No. Images	Description
Radiology (2D)	Slake-VQA	4,919	450	Radiology images & QA pairs
	MIMIC-CXR	2,142,892	231,483	Radiology images & free-form reports
	Digital Knee X-ray	1,469	1,469	Knee X-ray images & labels
	CXR-US2	132,680	132,680	Radiology images & free-form reports
	NLST	2,199	2,199	2D CT slices & free-form reports
	CT-US1	3,207	3,207	2D CT slices & free-form reports
Radiology (3D)	CT-US1	657,719	657,719	3D CT images & free-form reports
Pathology	PathVQA	19,654	2,599	Pathology images & QA pairs
	Histopathology	1,550,976	207,603	Histopathology images, captions, & QA pairs
Dermatology	PAD-UFES-20	2,047	2,047	Skin lesion images & labels
Ophthalmology	EyePACS	14,406	14,406	Fundus images & labels
Medical VQA	PMC	2,246,656	2,246,656	PubMed Central images & caption pairs
	MedVQA	12,664	3,168	Medical images & QA pairs
Genomics	UK Biobank	259,225	259,225	Genomic data & disease outcomes

PMC-VQA Paper training details

Takeaway: Authors experiment only with the Radiology Images.

Table 1: Comparison of existing medical VQA datasets with PMC-VQA, demonstrating the significant increase in size and diversity achieved by our dataset.

Dataset	Modality	Source	Images	QA pairs
VQA-RAD [18]	Radiology	MedPix® database	0.3k	3.5k
PathVQA [12]	Pathology	PEIR Digital Library [14]	5k	32.8k
SLAKE [23]	Radiology	MSD [3], ChestX-ray8 [36], CHAOS [15]	0.7k	14k
VQA-Med-2021 [5]	Radiology	MedPix® database	5k	5k
PMC-VQA	Mixture*	PubMed Central®	149k	227k

* Mixture: Radiology, Pathology, Microscopy, Signals, Generic biomedical illlustrations, *etc*.

Evaluation

Public benchmarks:

- **VQA-RAD**
- **SLAKE**

Domain-oriented open systems

Non-officially fine-tuned LLaMA and CLIP:

- **PMC-LLaMA & PMC-CLIP (MedViNT tuning)**

Why we have to have our data

Initially it is better to go with the public benchmarks (better work recognition).

Report assessment

- Existed studies are mainly focused on VQA which is subpart of the REPORT.
- Report-level evaluation (liver for example)
- Decompose reports into series of more complex VQA questions ->
- **CoT:** Propose more complicated training data generation process rather than in PMC-VQA, which involves report decomposition into VQA questions (chain-of-questions) that are necessary for reporting.