

## 1. CancerLLM: Al for Oncology

This language model, with 7 billion parameters, is designed to assist with cancer-related tasks.

Trained on 2.67 million clinical records and 515,500 pathology reports across 17 types of cancer, CancerLLM outperforms existing models by 7.61% in classification accuracy (F1 score).

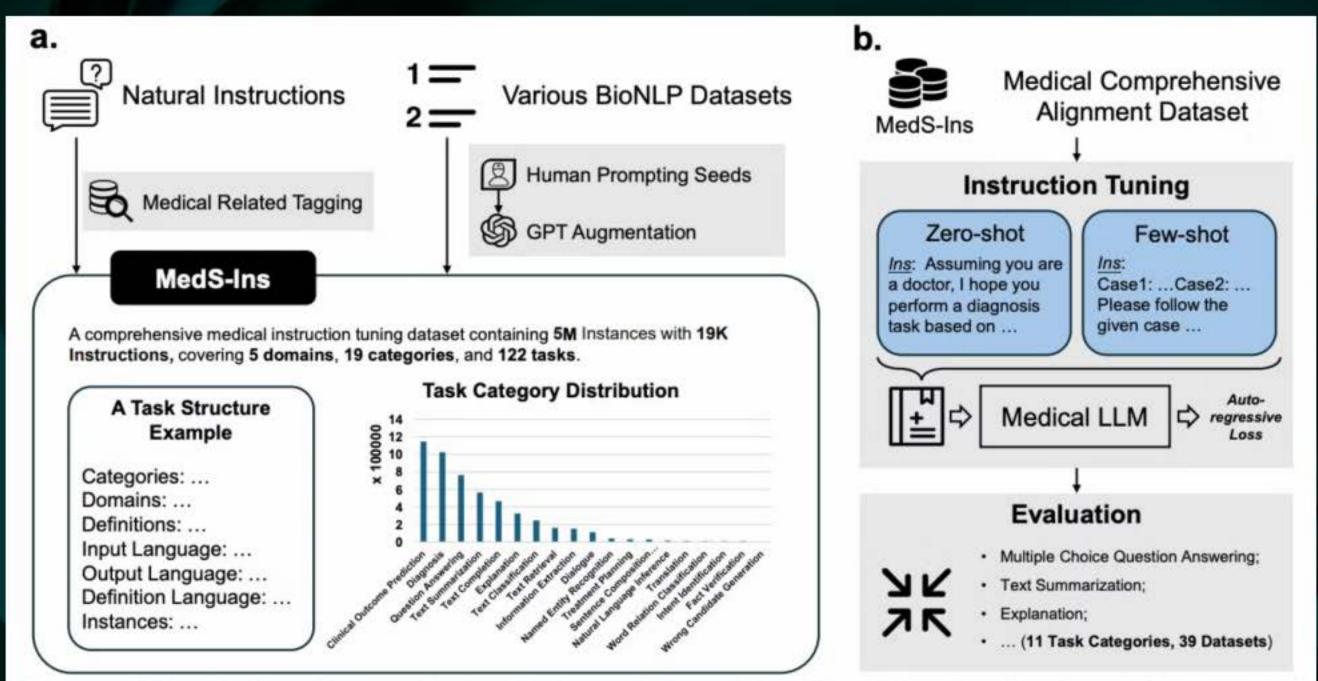


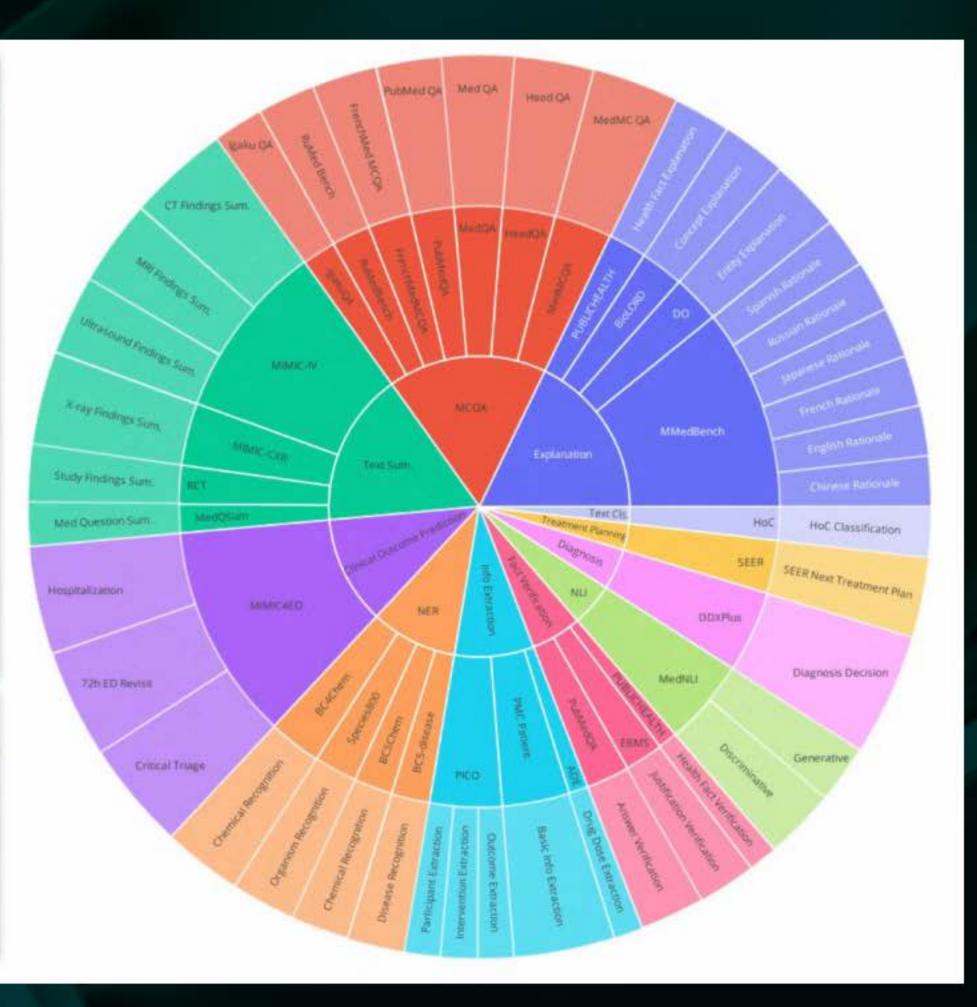
Figure 3 | The pipeline of our method. a The data collection pipeline. We mainly collect data through filtering natural instructions and prompting well-organized BioNLP datasets. b The training and evaluation pipeline for our model leveraging the collected MedS-Ins. We leverage the instruction tuning training method to combine different datasets and evaluate the final model on multiple benchmarks comprehensively.

## 2. MedUnA: Al for Medical Image Processing

Medical Unsupervised Adaptation (MedUnA) is a two-step method. First, class labels generated by an LLM are processed through a text encoder (BioBERT), aligning textual embeddings with class labels. Then, a trained adapter is integrated with the MedCLIP visual encoder to improve image-text alignment.

Category	Number		
Diagnosis	134529		
Text Summarization	41238		
NER	37585		
Information Extraction	16856		
MCQA	13701		
Clinical Outcome Prediction	4500		
Treatment Planning	3535		
Explanation	4486		
NLI	2844		
Fact Verification	2142		
Text Classification	158		

Item	Number		
Categories	11		
Source Dataset	28		
Tasks	52		
Samples	0.3M		



## 3. MedQA Benchmarking with LLM

This study explores how LLMs can automate the evaluation of medical Q&A systems, mimicking human expert analysis with a margin of error as low as 0.62 on a scale from 0 to 3.

#### Case Study: Asthma Treatment Evaluation

Question: "Has the patient been prescribed treatment for asthma?"

Ground Truth: Yes, albuterol sulfate HFA 90 mcg/actuation aerosol inhaler, Stiolto Respirat 2.5 mcg-

2.5 mcg/actuation, ProAir HFA 90 mcg/actuation aerosol inhaler, Ventolin HFA 90 mcg/actuation aerosol,

Trelegy Ellipta 100 mcg-62.5 mcg-25 mcg, methylprednisolone 4 mg tablets in a dose pack, Spiriva Respirat 2.5 mcg/actuation solution for inhalation, Medrol (Pak) 4 mg tablets in a dose pack.

Q&A System Response: Yes, the patient has been pre-

scribed albuterol sulfate 2.5 mg/3 ml (0.083%) solution for nebulization,

albuterol sulfate hfa 90 mcg/actuation aerosol inhaler, trelegy ellipta 100 mcg-62.5 mcg-25 mcg powder

for inhalation, benzonatate 200 mg capsule, and ventolin hfa 90 mcg/actuation aerosol inhaler.

		Precision	Succinctness	Correctness	Hallucination	Completeness	Coherence
<b>Evaluations:</b>	Medical Team	3	3	3	3	3	3
	LLM	3	2	2	2	2	3

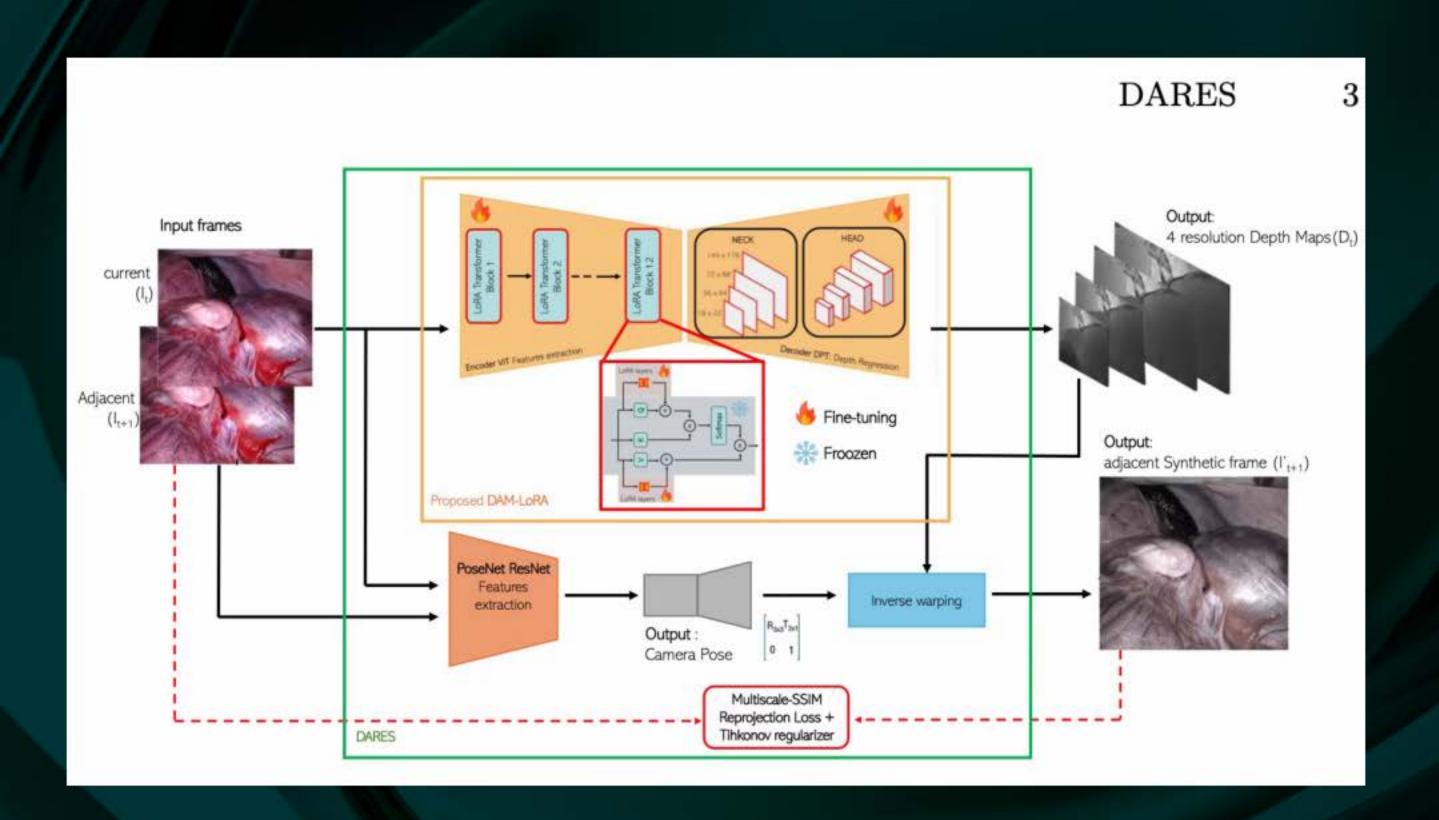
LLM Explanation: The response is highly relevant and succinct, providing a list of medications that are indeed used for treating asthma. However, it is not completely medically correct as it includes benzonatate,

which is a non-narcotic oral medication used to suppress coughs and not typically used in asthma treatment. This affects the medical correctness, introduces a hallucination, and impacts the completeness of the response.

Figure 2: Case Study: Comparison of actual and LLM-recommended asthma treatments. Correct recommendations are highlighted in yellow, incorrect in red. The medical team gave perfect scores, while the LLM identified specific issues in medical correctness, hallucination, and completeness, showcasing its potential in medical recommendations.

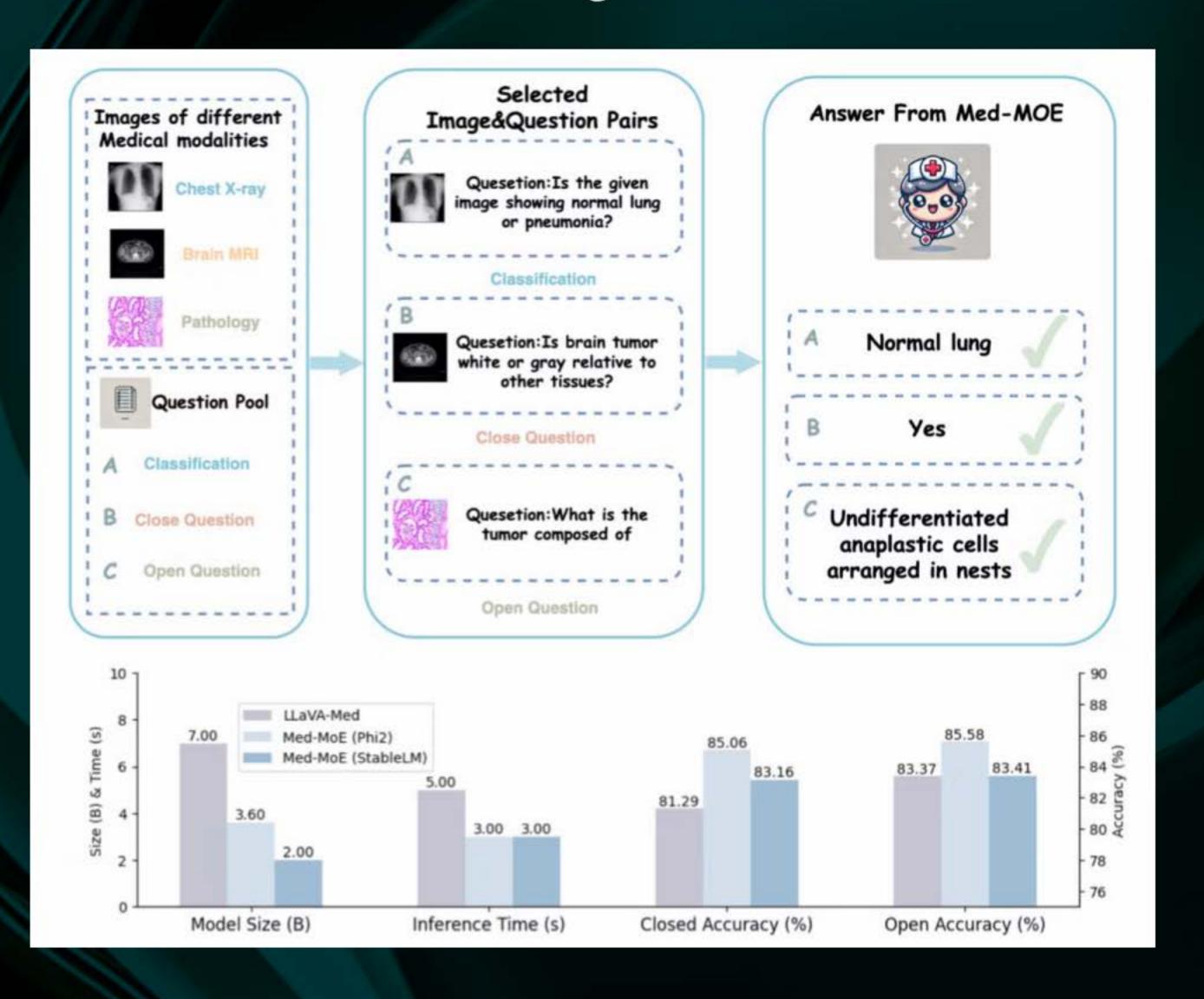
## 4. DARES: Al in Robotic Surgery

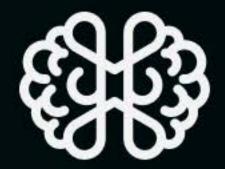
This base model assists in robotic endoscopic surgery by performing self-supervised monocular depth estimation. The method, code, and model are available for use in various robotic surgery tasks.



## 5. Med-MoE: Mixture-of-Experts for Medical VLM

Med-MoE is a lightweight framework for handling multimodal medical tasks. It aligns medical images with LLM tokens, selects experts for specific instructions, and fine-tunes the models for targeted tasks.





# HEALING WITH AI

