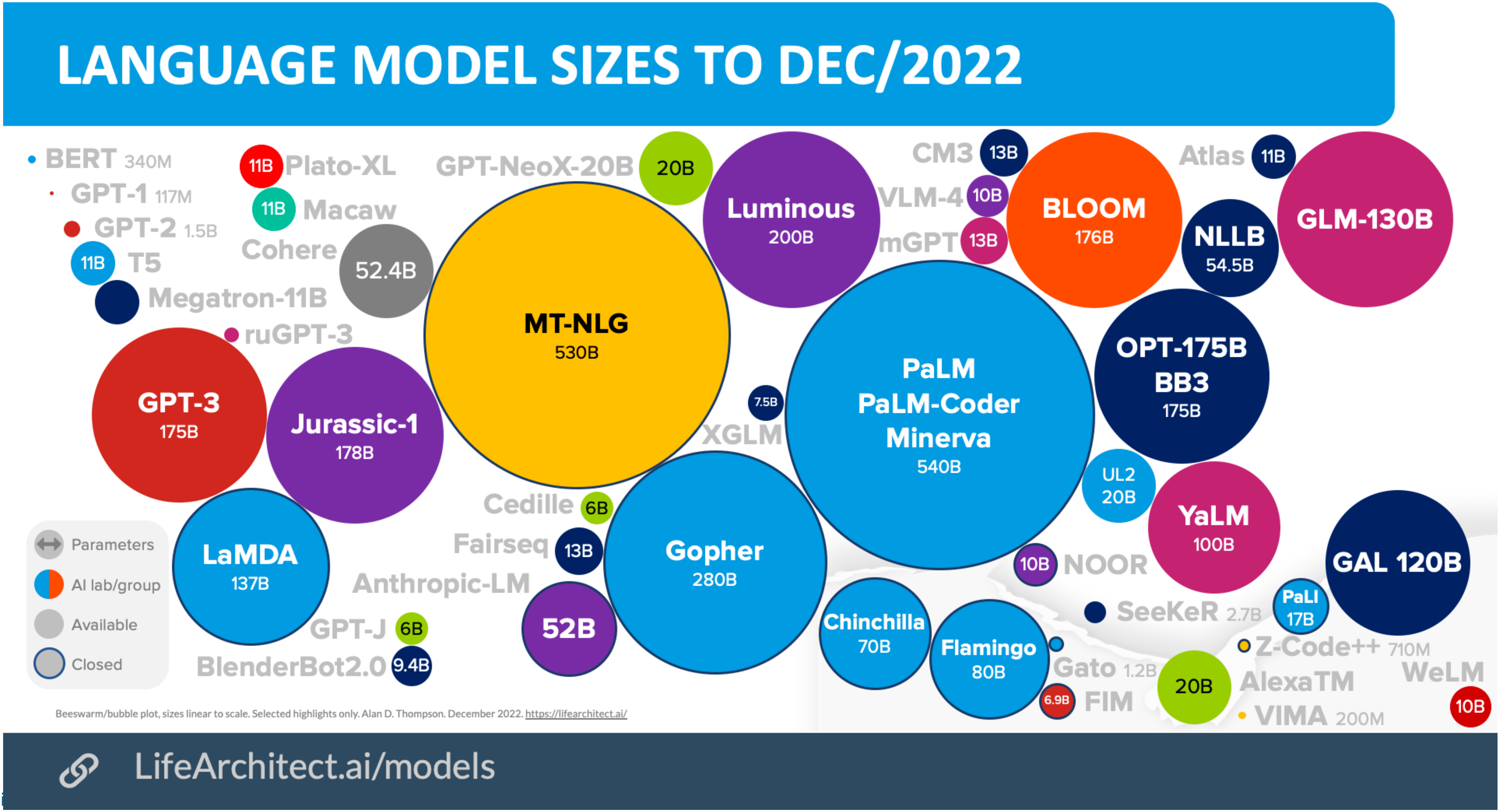
上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

Foundation Models in Medicine
Generalist vs Specialist

Presented by Shaoting Zhang

Date：2023.12

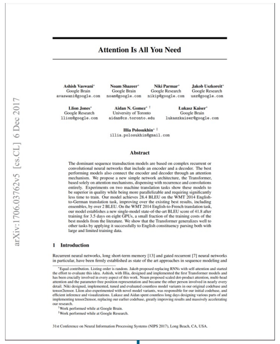
Large-Language Models (LLM)



Beeswarm/bubble plot, sizes linear to scale. Selected highlights only. Alan D. Thompson. December 2022. https://lifearchitect.ai/

LifeArchitect.ai/models

Large-Language Models (LLM)


Attention Is All You Need

arXiv:1706.03762v5 [cs.CL] 6 Dec 2017

Google

2017/06

Transformer

- Proposed the Transformer architecture

Improving Language Understanding by Generative Pre-Training

2018/06

GPT

- Using the decoder of Transformer, a language model is trained on a large amount of unannotated text data, which is then fine-tuned for subtasks

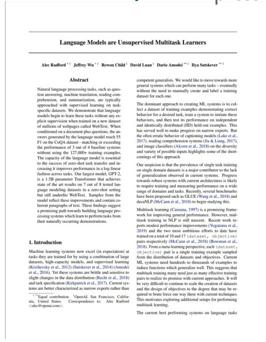
BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

arXiv:1810.04805v2 [cs.CL] 24 May 2019

2018/10

BERT

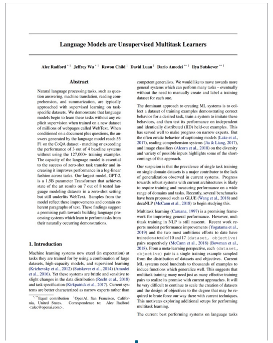
- Using the encoder of Transformer and pre-training on a larger dataset, the performance is much better than GPT


Language Models are Unsupervised Multitask Learners

2019/02

GPT - 2

- Continuing to use the decoder of Transformer to explore the potential of language models, but the performance is not as impressive as previous models


Language Models are Unsupervised Multitask Learners

2020/05

GPT - 3

- Compared to GPT-2, both the data and the model are 100 times larger, with very impressive results

Training language models to follow instructions with human feedback

arXiv:2203.02155v1 [cs.CL] 4 Mar 2022

2022/11

Instruct GPT

- Through instruction fine-tuning and RLHF (Reinforcement Learning from Human Feedback), the model's generalization ability is enhanced and aligned with human values, making the model controllable and safe

GPT-4 Technical Report

OpenAI

arXiv:2303.08774v3 [cs.CL] 27 Mar 2023

OpenAI

2023/03

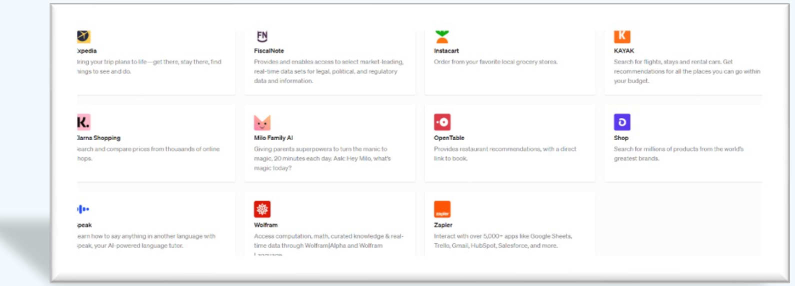
GPT – 4

- Compared to GPT3.5, it supports image & text input, can analyze images, and provides text output

Large-Language Models (LLM)

Initial support for plugins

- ➢ The first plugins have been created by Expedia, FiscalNote, Instacart, KAYAK, Klarna, Milo, OpenTable, Shopify, Slack, Speak, Wolfram, and Zapier.



xpedia
ring your trip plans to life—get there, stay there, find hings to see and do.

FiscalNote
Provides and enables access to select market-leading, real-time data sets for legal, political, and regulatory data and information.

Instacart
Order from your favorite local grocery stores.

KAYAK
Search for flights, stays and rental cars. Get recommendations for all the places you can go within your budget.

larna Shopping
earch and compare prices from thousands of online hops.

Milo Family AI
Giving parents superpowers to turn the manic to magic, 20 minutes each day. Ask: Hey Milo, what's magic today?

OpenTable
Provides restaurant recommendations, with a direct link to book.

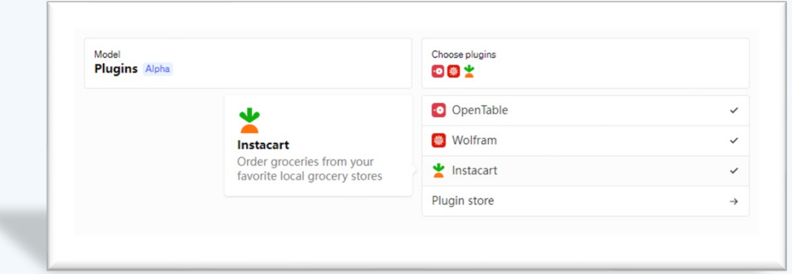
Shop
Search for millions of products from the world's greatest brands.

peak
earn how to say anything in another language with peak, your AI-powered language tutor.

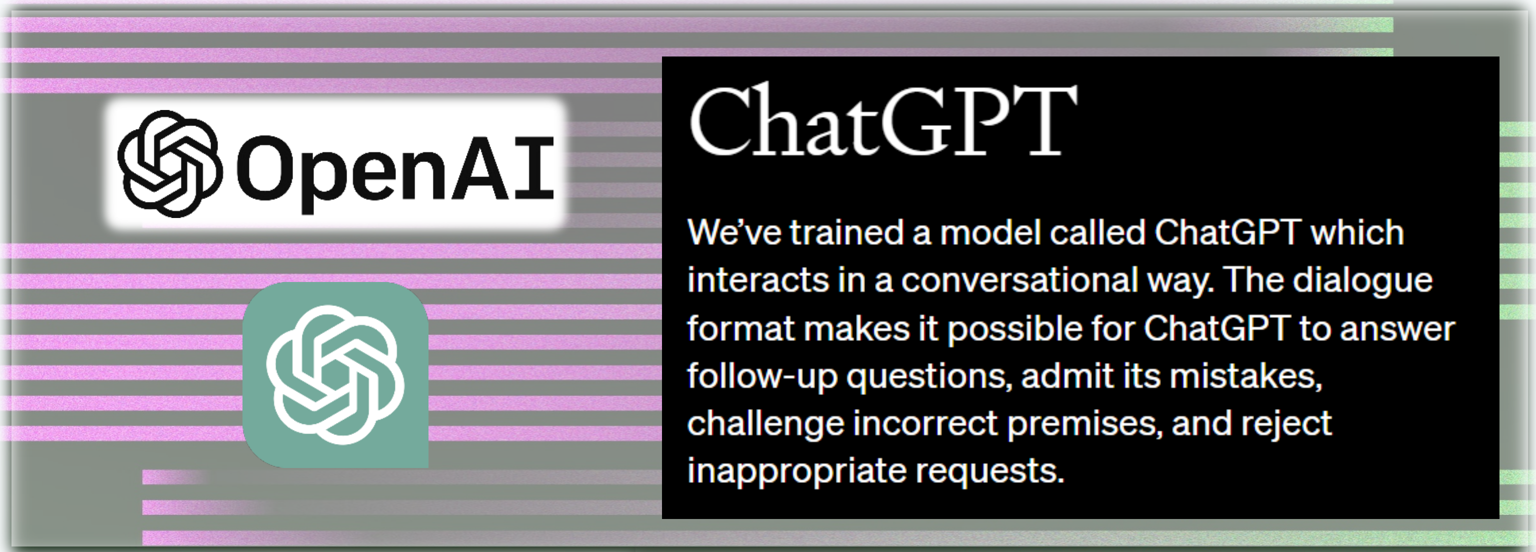
Wolfram
Access computation, math, curated knowledge & real-time data through Wolfram|Alpha and Wolfram

Zapier
Interact with over 5,000+ apps like Google Sheets, Trello, Gmail, HubSpot, Salesforce, and more.

- ➢ Hosting two plugins, a web browser and code interpreter (Alpha)
- ➢ Open-sourced the code for a knowledge base retrieval plugin
  - Browsing
  - Code Interpreter
  - Retrieval
  - Third-party plugins



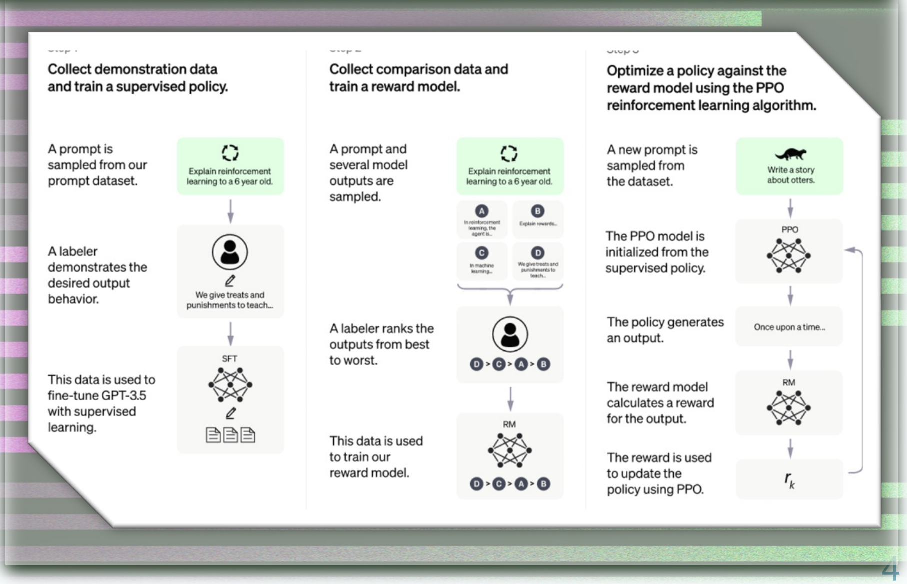
Instacart
Order groceries from your favorite local grocery stores



We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer follow-up questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests.

The AI model acts as an intelligent API caller

- ➢ OpenAI plugins connect ChatGPT to third-party applications (Beta)
- ➢ Plugins enable ChatGPT to interact with APIs defined by developers
- ➢ Plugins enable ChatGPT to do things like：
  - Retrieve real-time information;
    e.g., sports scores, stock prices, the latest news, etc.
  - Retrieve knowledge-base information;
    e.g., company docs, personal notes, etc.
  - Assist users with actions;
    e.g., booking a flight, ordering food, etc.



Collect demonstration data and train a supervised policy.

A prompt is sampled from our prompt dataset.

A labeler demonstrates the desired output behavior.

This data is used to fine-tune GPT-3.5 with supervised learning.

Collect comparison data and train a reward model.

A prompt and several model outputs are sampled.

A labeler ranks the outputs from best to worst.

This data is used to train our reward model.

Optimize a policy against the reward model using the PPO reinforcement learning algorithm.

A new prompt is sampled from the dataset.

The PPO model is initialized from the supervised policy.

The policy generates an output.

The reward model calculates a reward for the output.

The reward is used to update the policy using PPO.

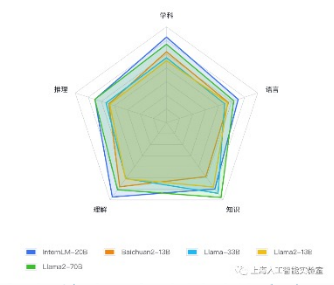
Explain reinforcement learning to a 6 year old.
We give treats and punishments to teach...
SFT
Explain reinforcement learning to a 6 year old.
A
In reinforcement learning, the agent is...
B
Explain rewards...
C
In machine learning...
D
We give treats and punishments to teach...
D > C > A > B
RM
D > C > A > B
Write a story about otters.
PPO
Once upon a time...
RM
r_k

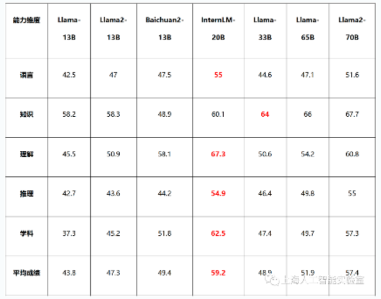
https://openai.com/blog/chatgpt-plugins https://openai.com/blog/chatgpt

Open-source Large-Language Models

书生

- **First unified full-stack open-source large-language model system —— INTERN.**
- The InternLM（书生·浦语）200 billion parameter version, InternLM-20B, has been officially open-sourced. The entire toolchain for the development and application of large models has been upgraded, encompassing five major stages: data, pre-training, fine-tuning, deployment, and evaluation.
- InternLM-20B was pre-trained on over 2.3T Tokens containing high-quality English, Chinese, and code data. It exhibits significant improvements in understanding, reasoning, mathematical, and programming abilities
  - ✓ Outstanding overall performance
  - ✓ Strong utility invocation capability
  - ✓ Supports a 16k context length (Through inference extrapolation)
  - ✓ Better value alignment.



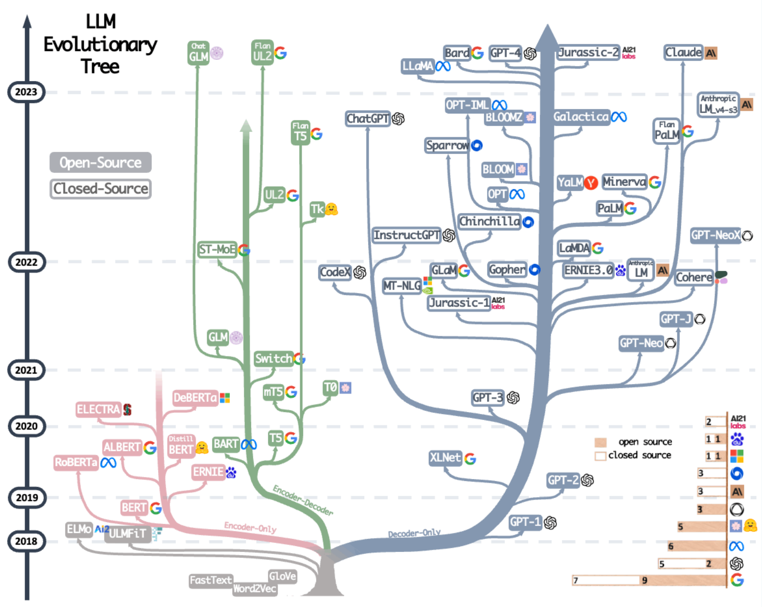


能力维度	Llama-13B	Llama2-13B	Baichuan2-13B	InternLM-20B	Llama-33B	Llama-65B	Llama2-70B
语言	42.5	47	47.5	55	44.6	47.1	51.6
知识	58.2	58.3	48.9	60.1	64	66	67.7
理解	45.5	50.9	58.1	67.3	50.6	54.2	60.8
推理	42.7	43.6	44.2	54.9	46.4	49.8	55
学科	37.3	45.2	51.8	62.5	47.4	49.7	57.3
平均成绩	43.8	47.3	49.4	59.2	48.9	51.9	57.4

本页所列大模型无排名先后
https://github.com/InternLM/InternLM https://mp.weixin.qq.com/s/
https://mp.weixin.qq.com/s/oTXnvWZJVdoOpFLHngbTYQ

RWKV

- Parallel training based on RNN.
- Comparable performance with other transformer-based LLMs.



Baichuan

- A series of open-source LLMs free for commercial use（13B & 7B）.
- Based on Transformer, supporting English & Chinses, context window length is 4096.

https://github.com/baichuan-inc/Baichuan-7B https://github.com/baichuan-inc/Baichuan-13B#readme
https://github.com/BlinkDL/RWKV-LM https://github.com/THUDM/GLM
https://huggingface.co/RWKV https://juejin.cn/post/7223305855923044409

- A series of LLMs scaling from 7B to 65B.
- One of the most advanced large-scale foundational language models currently available, and the most famous open-source model.

a BigScience initiative
BLOOM
176B params · 59 languages · Open-access

- 176B LLM which could generate 46 natural languages and 13 programming language texts.
- Using ALiBi position embedding to have more smooth training process and better downstream performance.
- Embedding LayerNorm: Add extra layer norm after embedding layer to improve training stability.

ChatGLM Alpha

- GLM: A generative LLM, pretraining with auto-regressive blank padding objects and could be fine-tuned in various natural language understanding tasks.
- New ChatGLM2-6B model: a GLM-based framework, aiming at chinese Q&A and dialogue model, users can deploy in consumer GPUs.

https://github.com/Hannibal046/Awesome-LLM https://keg.cs.tsinghua.edu.cn/jietang/publications/ChatGLM&Beyond.pdf
https://research.facebook.com/publications/llama-open-and-efficient-foundation-language-models/ https://arxiv.org/pdf/2211.05100.pdf
https://stock.adobe.com/search?k=llama+logo https://chatglm.cn/blog https://bigscience.huggingface.co/blog/bloom

Vision Foundation Models

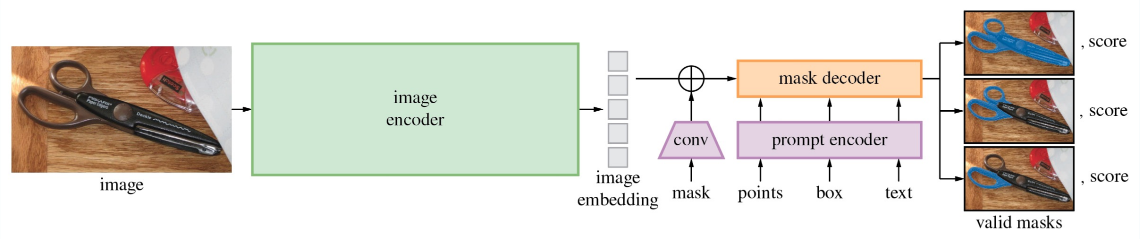
What **task** will enable zero-shot generalization?

What is the corresponding **model** architecture?

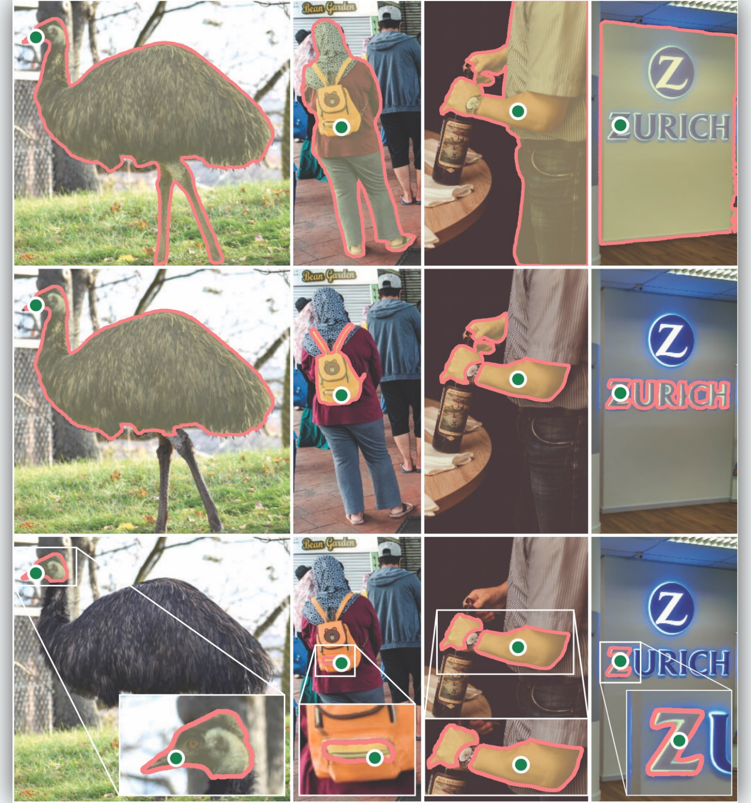
What **data** can power this task and model?

Segment Anything Model (SAM):

Build a pretrained model with large quantities of parameters in segmentation area, strengthen its generalization by prompt engineering.



A powerful image encoder computes an image embedding, a prompt encoder embeds prompts, and then the two infor-mation sources are combined in a lightweight mask decoderthat predicts segmentation masks.



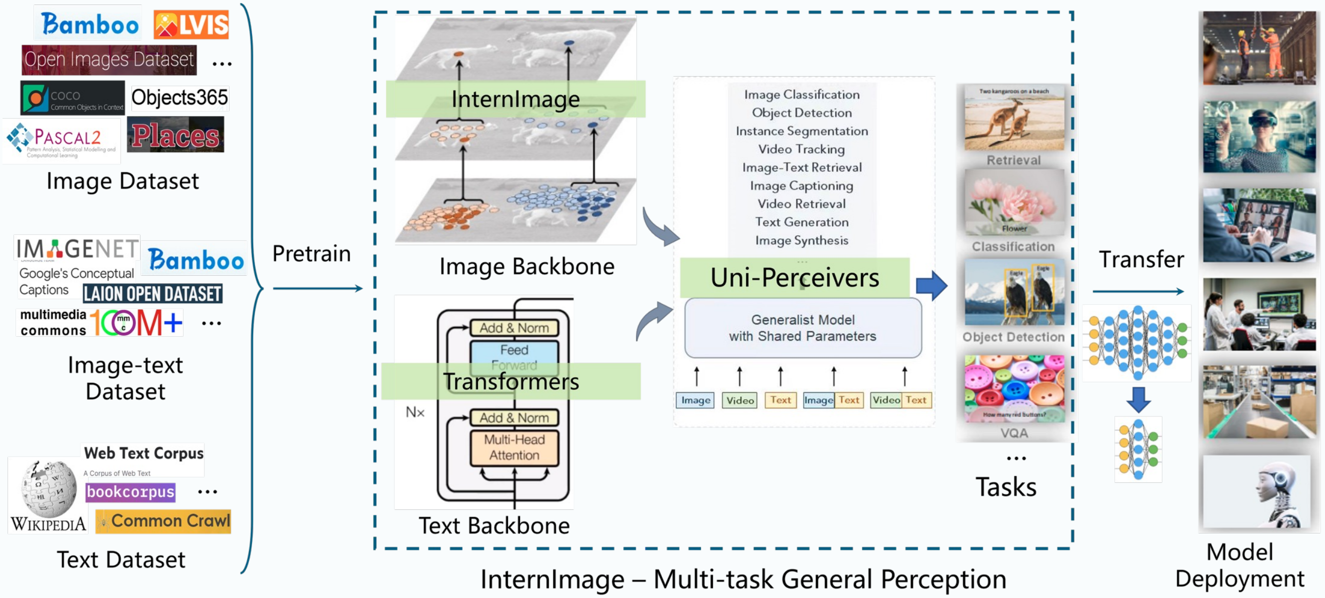
Kirillov A, Mintun E, Ravi N, et al. Segment anything[J]. arXiv preprint arXiv:2304.02643, 2023.

https://segment-anything.com/

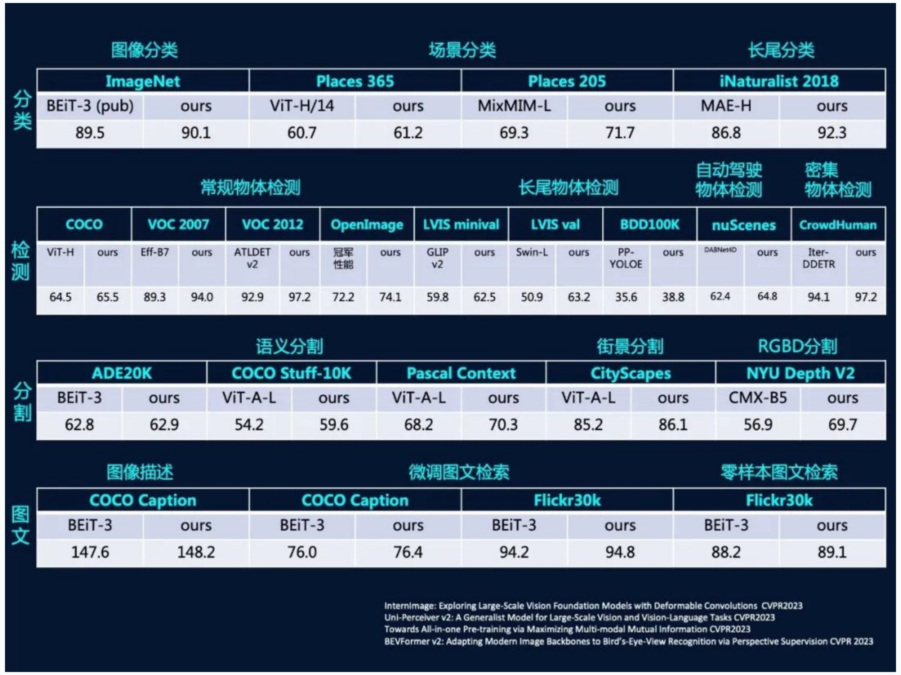
Vision Foundation Models

A new generation of General Vision technology System: INTERN 2.5

"INTERN-2.5" is a powerful multimodal multitask general model jointly released by SenseTime and Shanghai AI Laboratory. It consists of large-scale vision foundation model "InternImage", pre-training method "M3I-Pretraining", generic decoder "Uni-Perceiver" series, and generic encoder for autonomous driving perception "BEVFormer" series.



Wang W, Dai J, Chen Z, et al. Internimage: Exploring large-scale vision foundation models with deformable convolutions[J]. CVPR 2023.



	图像分类		场景分类				长尾分类	
分类	ImageNet		Places 365		Places 205		iNaturalist 2018	
	BEiT-3 (pub)	ours	ViT-H/14	ours	MixMIM-L	ours	MAE-H	ours
	89.5	90.1	60.7	61.2	69.3	71.7	86.8	92.3

	常规物体检测								长尾物体检测						自动驾驶物体检测		密集物体检测	
检测	COCO		VOC 2007		VOC 2012		OpenImage		LVIS minival		LVIS val		BDD100K		nuScenes		CrowdHuman	
	ViT-H	ours	Eff-B7	ours	ATLDET v2	ours	冠军性能	ours	GLIP v2	ours	Swin-L	ours	PP-YOLOE	ours	DABNet4D	ours	Iter-DDETR	ours
	64.5	65.5	89.3	94.0	92.9	97.2	72.2	74.1	59.8	62.5	50.9	63.2	35.6	38.8	62.4	64.8	94.1	97.2

	语义分割						街景分割		RGBD分割	
分割	ADE20K		COCO Stuff-10K		Pascal Context		CityScapes		NYU Depth V2	
	BEiT-3	ours	ViT-A-L	ours	ViT-A-L	ours	ViT-A-L	ours	CMX-B5	ours
	62.8	62.9	54.2	59.6	68.2	70.3	85.2	86.1	56.9	69.7

	图像描述		微调图文检索				零样本图文检索	
图文	COCO Caption		COCO Caption		Flickr30k		Flickr30k	
	BEiT-3	ours	BEiT-3	ours	BEiT-3	ours	BEiT-3	ours
	147.6	148.2	76.0	76.4	94.2	94.8	88.2	89.1

InternImage: Exploring Large-Scale Vision Foundation Models with Deformable Convolutions CVPR2023
Uni-Perceiver v2: A Generalist Model for Large-Scale Vision and Vision-Language Tasks CVPR2023
Towards All-in-one Pre-training via Maximizing Multi-modal Mutual Information CVPR2023
BEVFormer v2: Adapting Modern Image Backbones to Bird's-Eye-View Recognition via Perspective Supervision CVPR 2023

"INTERN-2.5" demonstrated world's best performance on 20 important visual benchmark datasets, covering a wide range of tasks such as classification, detection, and segmentation, making it the top-performing model across multiple domains.

Foundation Models for Science

High-precision, Long-term, High-efficiency: AI large model forecasting global weather

Built on multi-modal and multi-task deep learning methods

Science Large AI Model **FENGWU**

Global medium-range weather forecast AI Model released by Shanghai AI Laboratory, University of Science and Technology of China, Shanghai Jiao Tong University, Nanjing University of Information Science and Technology, The Institute of Atmospheric Physics, Chinese Academy of Sciences .

10d

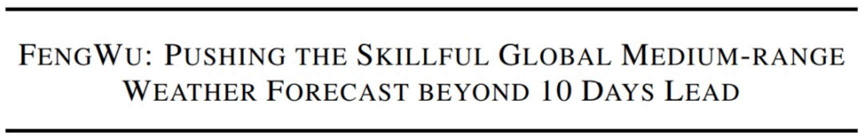
Improves the forecast skill and extend the skillful global medium-range weather forecast **out to 10 days** lead **for the first time**.

30s

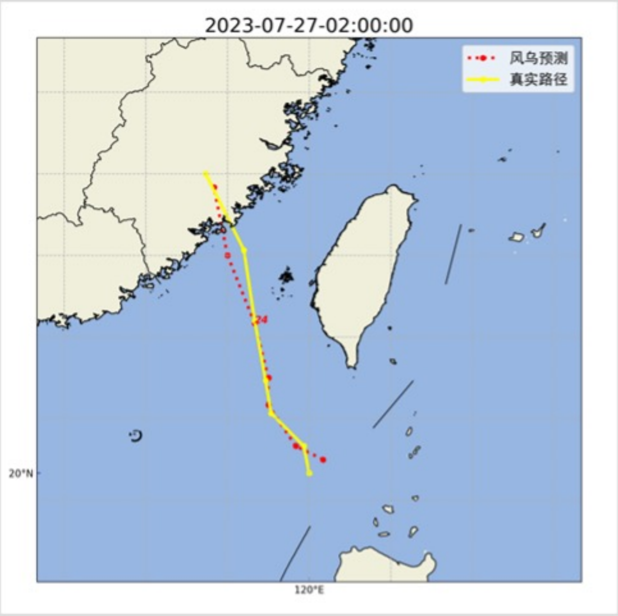
Costs **less than 30 seconds** to generate all forecasts in the following 10 days

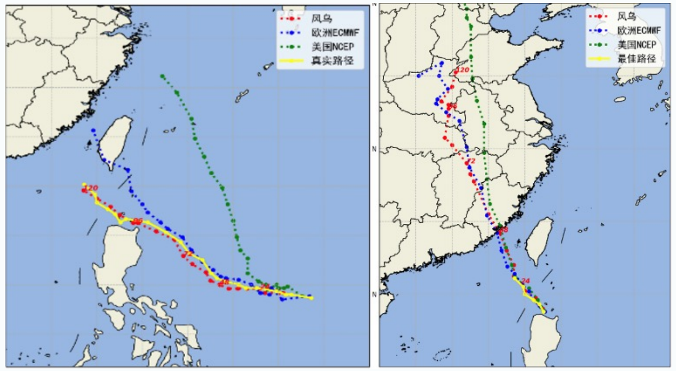
80%

Has higher accuracy than **GraphCast** (DeepMind) on **80% reported prediction targets**.

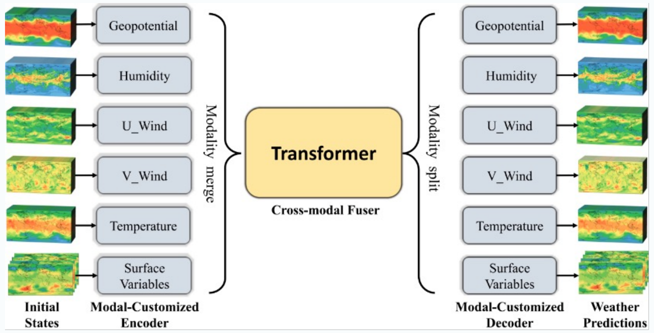

FENGWU: PUSHING THE SKILLFUL GLOBAL MEDIUM-RANGE WEATHER FORECAST BEYOND 10 DAYS LEAD

- Recently, with the support of the China Meteorological Administration, **Shanghai AI Laboratory cooperated with the National Meteorological Center and the Shanghai Meteorological Bureau** to complete the real-time trial operation of the Fengwu model deployment. Comparative test evaluation was carried out for the two typhoons "Tali" and "Dusuri" that made landfall this year, and Fengwu showed strong business application potential, providing strong technical support for typhoon forecast and early warning.





Tracking forecasting of Typhoon Doksuri of FengWu, ECMWF and NCEP, figures above are reports from July 22th 8am and July 26th 8am，respectively.



Overview of FengWu’ s architecture. FengWu first **treats the multiple weather factors as different modalities** and extracts their feature embeddings independently. And then **a transformer-based network is utilized to fuse and passmessages among different modalities.**

Chen K, Han T, Gong J, et al. FengWu: Pushing the Skillful Global Medium-range Weather Forecast beyond 10 Days Lead[J]. arXiv preprint arXiv:2304.02948, 2023.

https://mp.weixin.qq.com/s/m8TuGXNjv4iWqxzTpMFlLg

LLM for Medicine

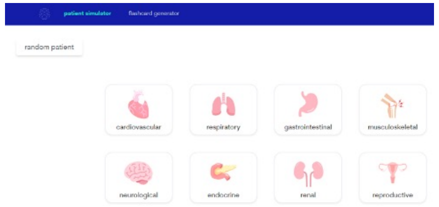
上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

Microsoft + NUANCE

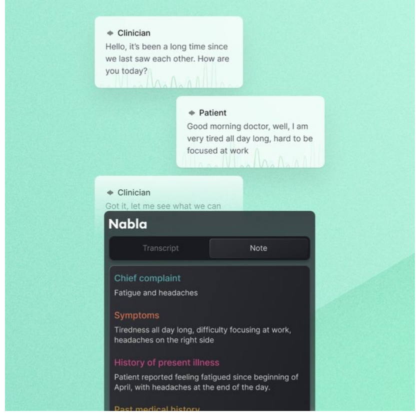
Nuance is introducing a new GPT-powered capability for contact center AI solutions. Here's how it will help our customers make their existing AI self-service experiences even smarter, enabling higher levels of self-service while increasing cost savings and driving revenue.

https://www.nuance.com/index.html
https://whatsnext.nuance.com/customer-engagement/gpt-powered-capability-nuance-contact-center/

Hippocratic AI
— Do No Harm —



https://www.hippocraticai.com/
https://mp.weixin.qq.com/s/1SiqLSaadpYgc8AIj1atcw


Clinician
Hello, it's been a long time since we last saw each other. How are you today?

Patient
Good morning doctor, well, I am very tired all day long, hard to be focused at work

Clinician
Got it, let me see what we can

Nabla

Chief complaint
Fatigue and headaches

Symptoms
Tiredness all day long, difficulty focusing at work, headaches on the right side

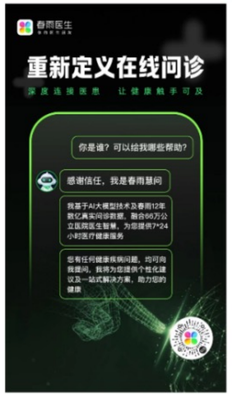
History of present illness
Patient reported feeling fatigued since beginning of April, with headaches at the end of the day.

Past medical history

https://www.nabla.com/
https://twitter.com/NablaTech

春雨医生
你 的 医 生 朋 友

2023年，基于大模型重新定义在线问诊


春雨医生

重新定义在线问诊

深度连接医患　让健康触手可及

你是谁？可以给我哪些帮助？

感谢信任，我是春雨慧问

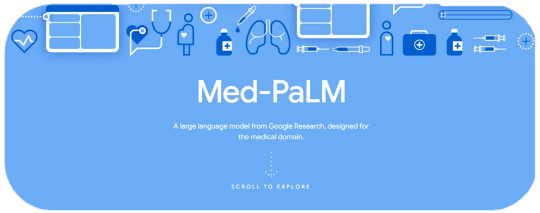
我基于AI大模型技术及春雨12年数亿真实问诊数据，融合66万公立医院医生智慧，为您提供7*24小时医疗健康服务

您有任何健康疾病问题，均可向我提问，我将为您提供个性化建议及一站式解决方案，助力您的健康

https://mp.weixin.qq.com/s/5CP9E1eeb5FsUy68NNN5IQ
https://www.chunyuyisheng.com/

LLM for Medical

Google

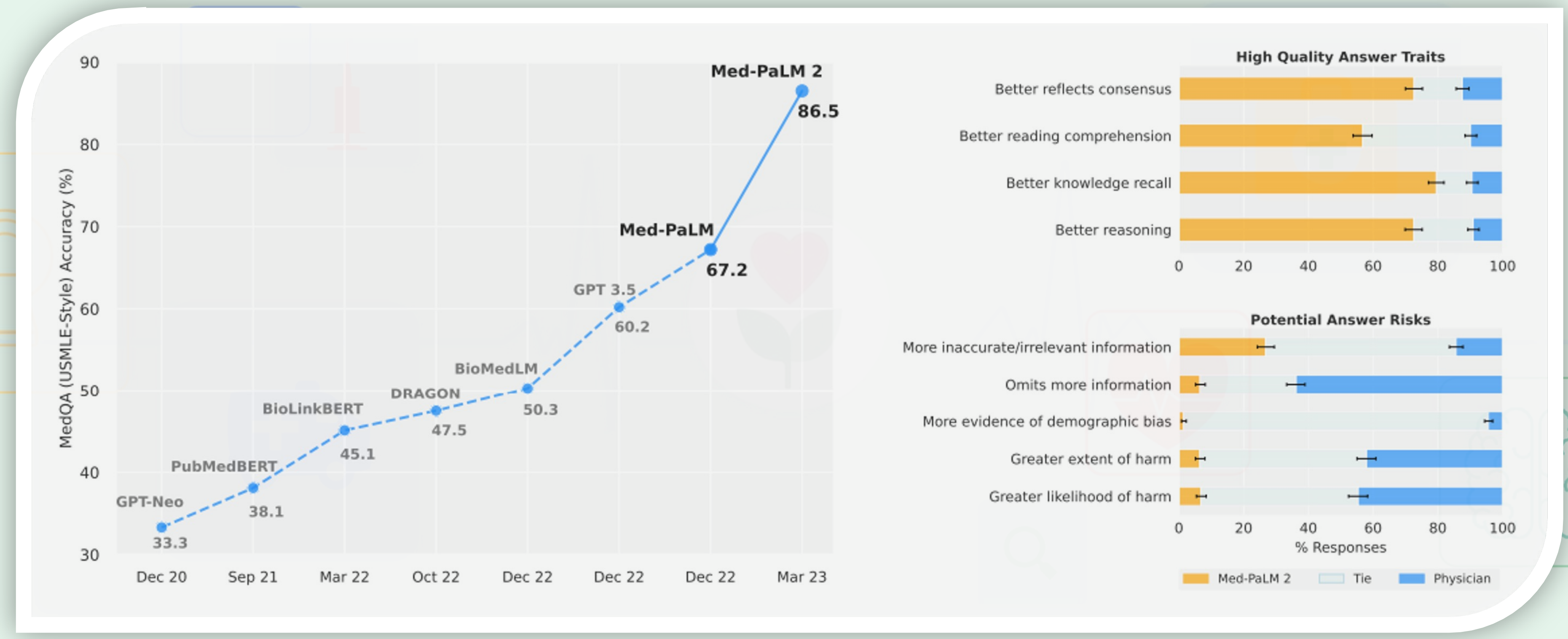

Med-PaLM

A large language model from Google Research, designed for the medical domain.

https://www.google.ca/
https://sites.research.google/med-palm/

LLM for Medicine

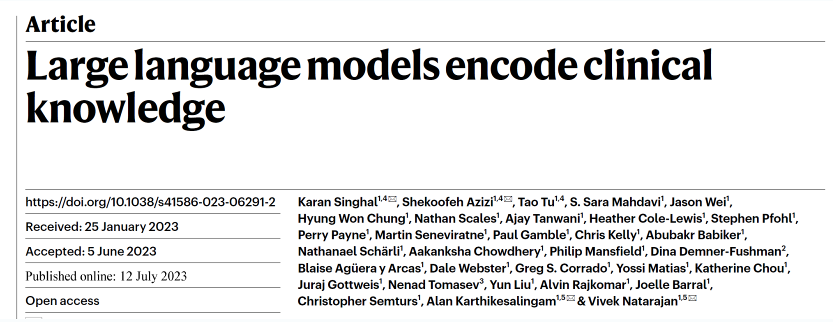
Google



K, Singhal, T. Tu, J. Gottweis, et al. Towards Expert-Level Medical Question Answering with Large Language Models. https://arxiv.org/pdf/2305.09617.pdf

https://cloud.google.com/blog/topics/healthcare-life-sciences/sharing-google-med-palm-2-medical-large-language-model

LLM for Medicine



Article

Large language models encode clinical knowledge

https://doi.org/10.1038/s41586-023-06291-2

Received: 25 January 2023

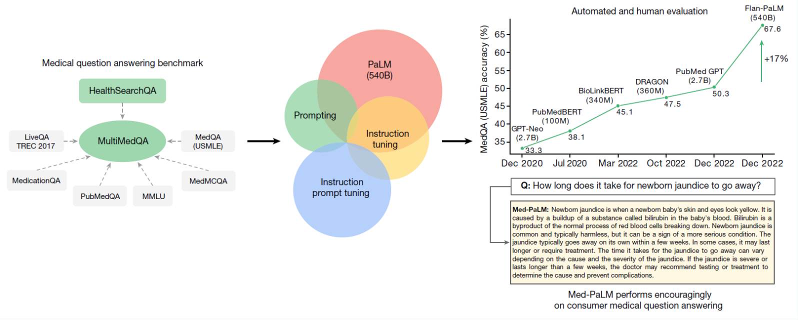
Accepted: 5 June 2023

Published online: 12 July 2023

Open access

Karan Singhal[1,4]✉, Shekoofeh Azizi[1,4]✉, Tao Tu[1,4], S. Sara Mahdavi[1], Jason Wei[1], Hyung Won Chung[1], Nathan Scales[1], Ajay Tanwani[1], Heather Cole-Lewis[1], Stephen Pfohl[1], Perry Payne[1], Martin Seneviratne[1], Paul Gamble[1], Chris Kelly[1], Abubakr Babiker[1], Nathanael Schärli[1], Aakanksha Chowdhery[1], Philip Mansfield[1], Dina Demner-Fushman[2], Blaise Agüera y Arcas[1], Dale Webster[1], Greg S. Corrado[1], Yossi Matias[1], Katherine Chou[1], Juraj Gottweis[1], Nenad Tomasev[3], Yun Liu[1], Alvin Rajkomar[1], Joelle Barral[1], Christopher Semturs[1], Alan Karthikesalingam[1,5]✉ & Vivek Natarajan[1,5]✉

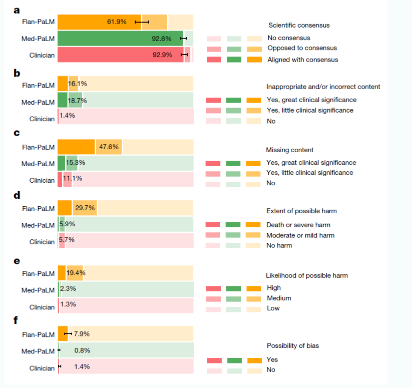
- Med-PaLM is proposed on the basis of Plan-PaLM by prompt engineering method.
- Flan-PaLM achieves SOTA in each MultiMedQA multiple-choice question tasks, its accuracy on MedQA is 67.6%.
- Compared with Plan-PaLM, Med-PaLM has significant improvement in practical problems, its accuracy is comparable with professional doctors, showing its potentials in medical fields.
- Besides Med-PaLM, a new dataset is proposed for medicine LLM evaluation.

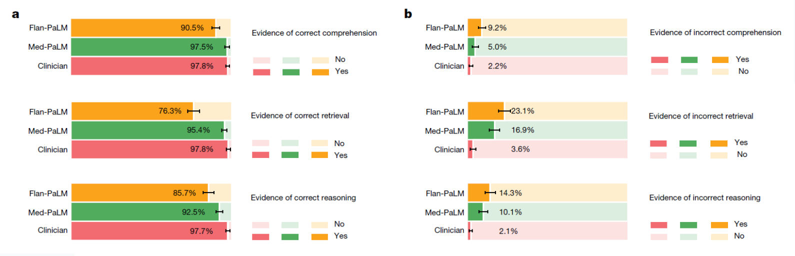


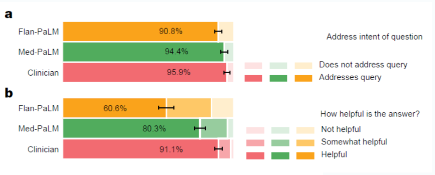
Q: How long does it take for newborn jaundice to go away?

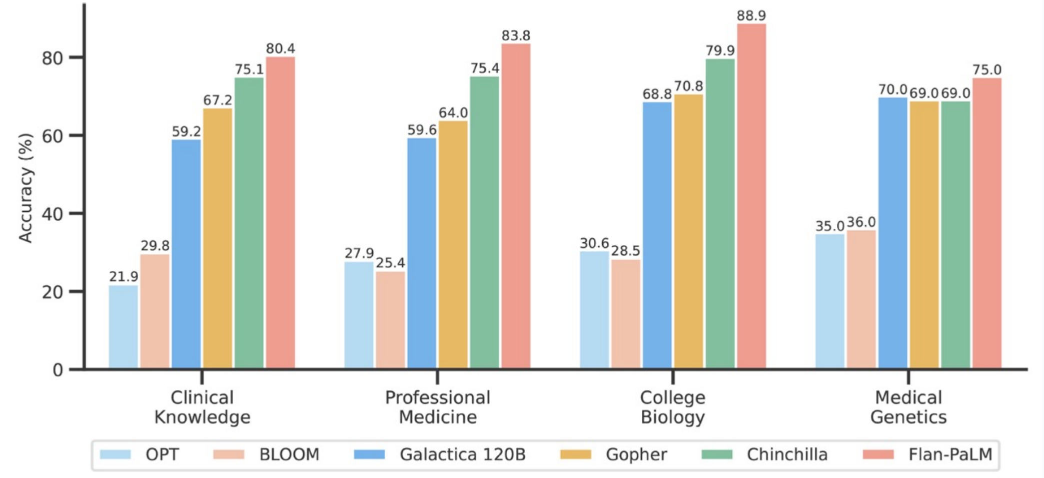
Med-PaLM: Newborn jaundice is when a newborn baby's skin and eyes look yellow. It is caused by a buildup of a substance called bilirubin in the baby's blood. Bilirubin is a byproduct of the normal process of red blood cells breaking down. Newborn jaundice is common and typically harmless, but it can be a sign of a more serious condition. The jaundice typically goes away on its own within a few weeks. In some cases, it may last longer or require treatment. The time it takes for the jaundice to go away can vary depending on the cause and the severity of the jaundice. If the jaundice is severe or lasts longer than a few weeks, the doctor may recommend testing or treatment to determine the cause and prevent complications.

Med-PaLM performs encouragingly on consumer medical question answering









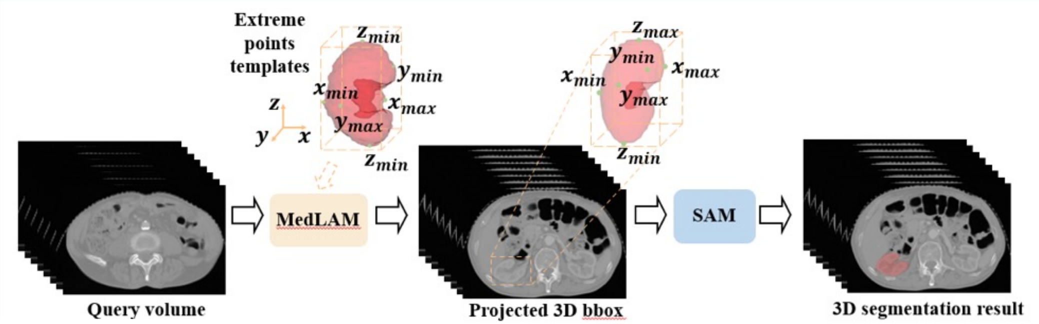
Extended Data Fig. 2 | Comparison of SOTA LLMs on MMLU clinical topics. Flan-PaLM achieves state-of-the-art performance on MMLU clinical topics.

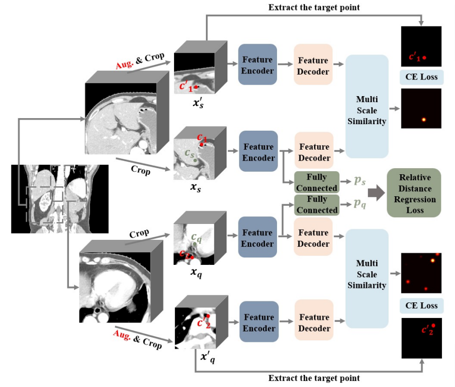
K. Singhal, S. Azizi, A. Karthikesalingam, V. Natarajan, et al. Large language models encode clinical knowledge. Nature. 2023

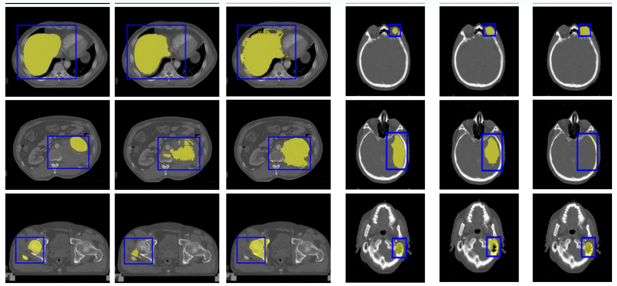
https://36kr.com/p/2342550668809730?channel=wechat

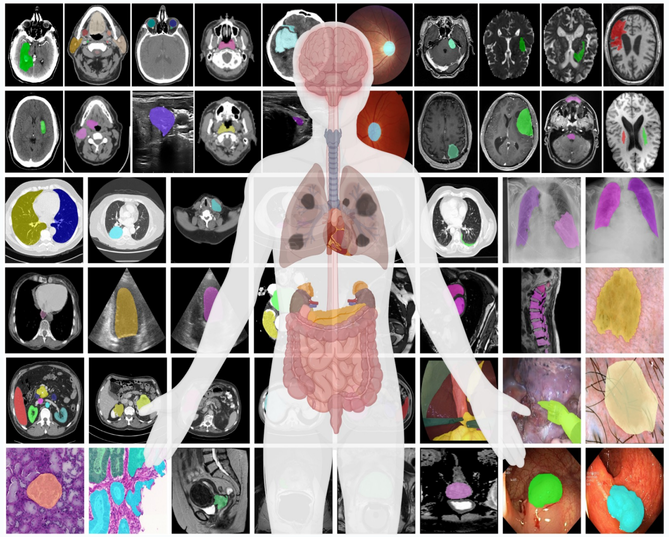
Vision Foundation Models for Medical Images

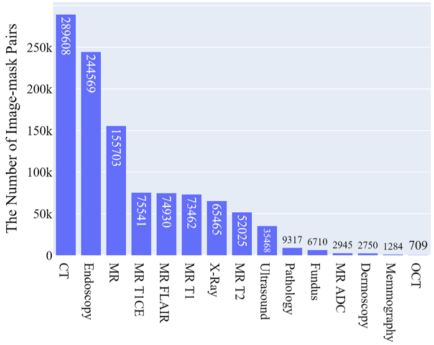
- MedLAM is the first completely automated medical adaptation of the SAM model, designed to significantly alleviate the annotation workload in the segmentation of medical image.
- Developing MedLAM, a Localize Anything Model for 3D Medical Images
- Coupling this approach with SAM’ s segmentation capabilities led toan efficient and accurate process for image segmentation.

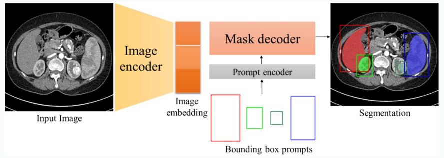












- Motivated by the remarkable generality of the Segment Anything Model (SAM), MedSAM is the first foundation model for universal medical image segmentation;
- Thoroughly evaluate MedSAM through comprehensive experiments on over 70 internal validation tasks and 40 external validation tasks, spanning a variety of anatomical structures, pathological conditions, and medical imaging modalities;
- MedSAM consistently outperforms the state-of-the-art (SOTA) segmentation foundation model, while achieving performance on par with, or even surpassing specialist models. These results highlight the potential of MedSAM as a powerful tool for medical image segmentation.

W. Lei, X. Wei, X. Zhang, K. Li, S. Zhang. “MedLSAM: Localize and Segment Anything Model for 3D Medical Images” arXiv:2306.14752 2023

Jun Ma, Yuting He, Feifei Li, Lin Han, Chenyu You, Bo Wang. “Segment Anything in Medical Images” arXiv:2304.12306v2 2023

Foundation Models for Science

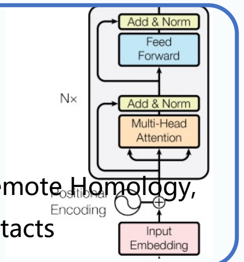
- Protein language model is the transfer application of various language models in the field of biochemistry.
- It can promote the prediction tasks such as protein structure, protein function and generation of new sequences.
- Unsupervised language models can directly utilize massive sequence information in protein databases without relying on manual annotation of sequences.
- Examples: ESM series including ESM-1b, ESM-MSA-1b, ESM-1v, ESM-2 are proposed by Meta.

ESM-1b *family-specific*

Training: self-supervised, masked token

Dataset: UR50, 250 million protein sequence

Tasks: Biochemical Properties, Variations, Remote Homology, Secondary Structure and Tertiary Contacts

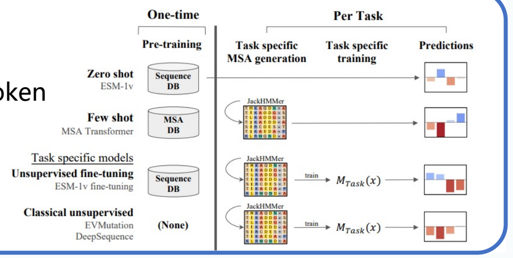


ESM-1v *zero-shot*

Training: self-supervised, masked token

Dataset: UR90, protein sequence

Tasks: effects of mutations on protein function



2019

2021

2021

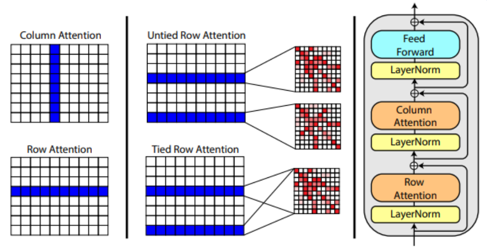
2022

ESM-MSA-1b *few-shot*

Training: self-supervised, masked token

Dataset: 4.3 TB, 26 million MSAs

Tasks: long-range contact、secondary and tertiary structures



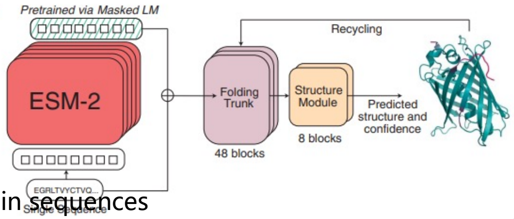
ESM-2、ESMFold

Training: self-supervised, masked token

Dataset: UR50, protein sequence

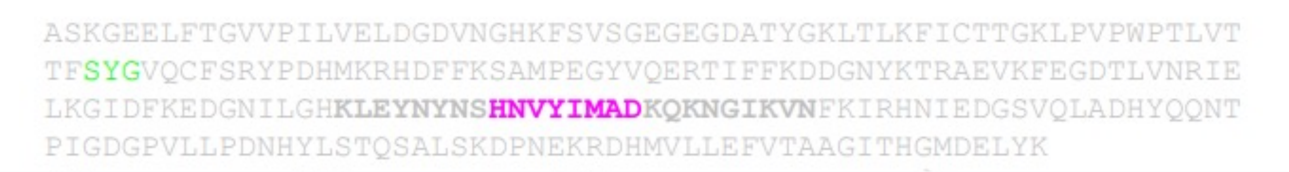
Tasks: ESMFold → structure prediction

>617 million metagenomic protein sequences

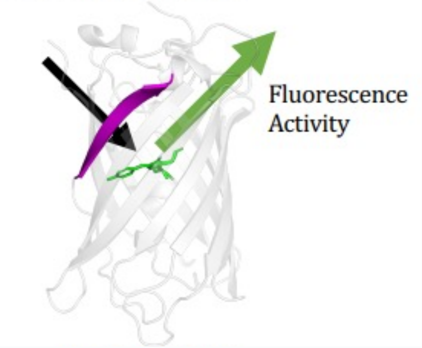


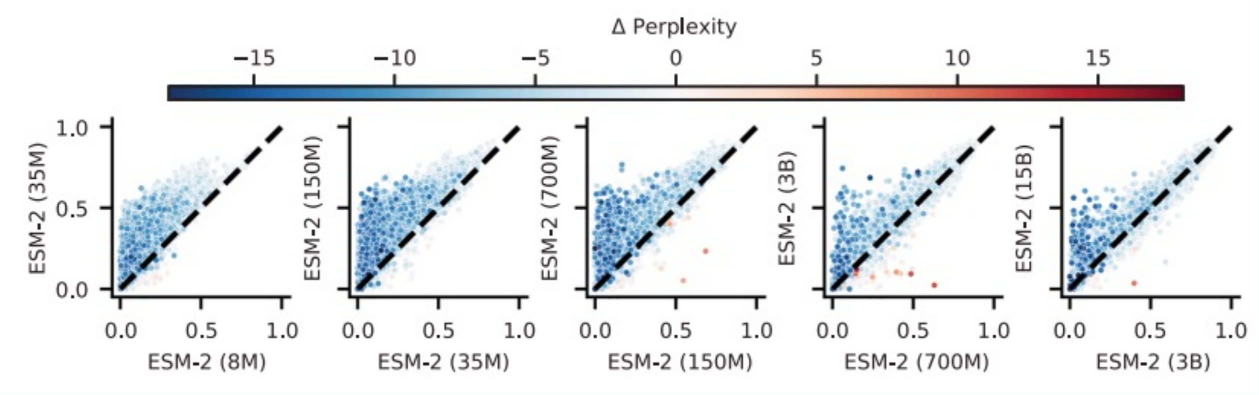
Foundation Models for Science

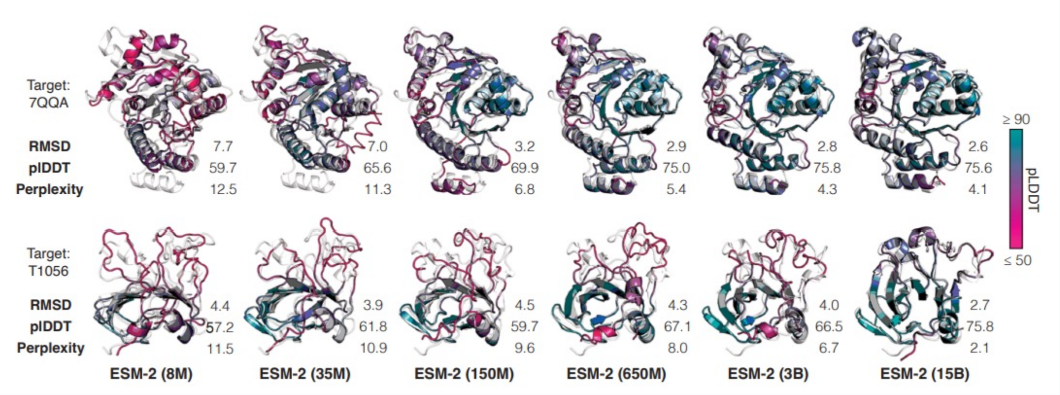
- Motivation: the statistical laws of protein sequences imply their structural information → the language model can learn this law through the input sequence.
- Task: filling in missing amino acids in protein sequences → a language model has to understand the underlying structure that creates the patterns in the sequences
- Evidence: As language models of protein sequences are scaled up to 15 billion parameters, an atomic-resolution picture of protein structure emerges in the learned representations.



ASKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVT
TFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIE
LKGIDFKEDGNILGH**KLEYNYNSHNVYIMADKQKNGIKVN**FKIRHNIEDGSVQLADHYQQNT
PIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK







Advantage of large models: Emergent abilities

Emergent abilities

With the advent of large models, the term 'emergence' has become popular. It is typically described as capabilities that do not exist in small-scale models but are present in large-scale models.

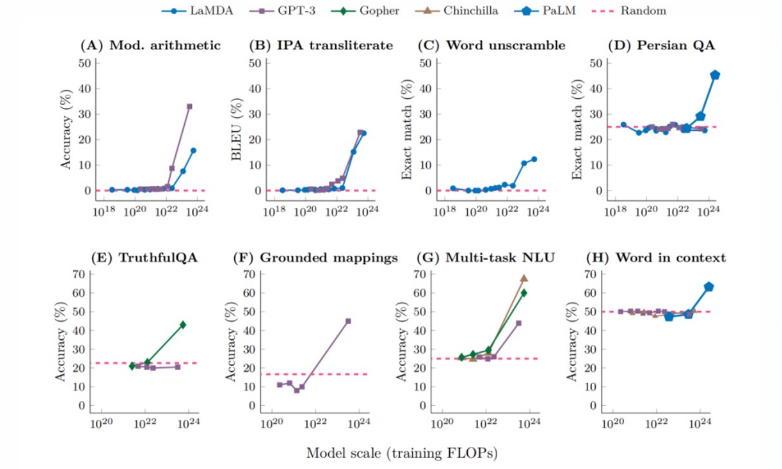
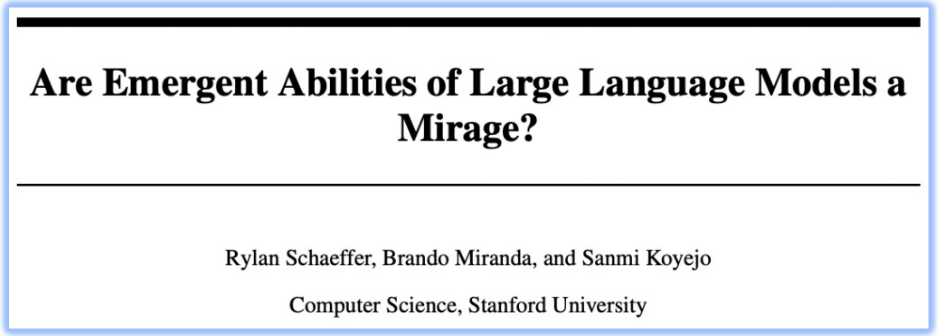


Figure 1: **Emergent abilities of large language models.** Language model families display *sharp* and *unpredictable* increases in performance at specific tasks as model scale increases. Emergent abilities [34] have also previously been labeled “abrupt, specific capability scaling” [8], “breakthrough capabilities” [29] and “sharp left turns” [17, 18]. Source: Fig. 2 from [34].



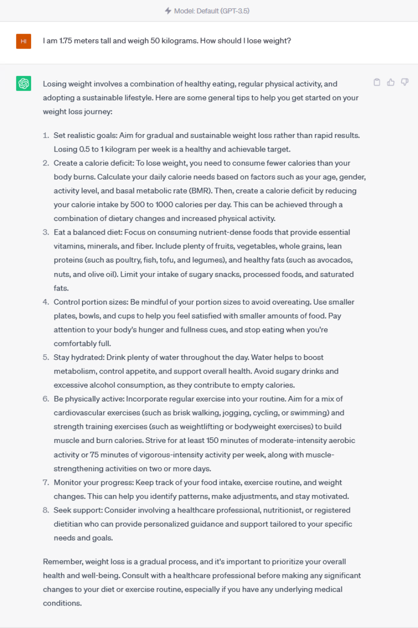
Are Emergent Abilities of Large Language Models a Mirage?

Rylan Schaeffer, Brando Miranda, and Sanmi Koyejo

Computer Science, Stanford University

"I am 1.75 meters tall and weigh 50 kilograms. How should I lose weight?"

GPT 3.5



Model: Default (GPT-3.5)

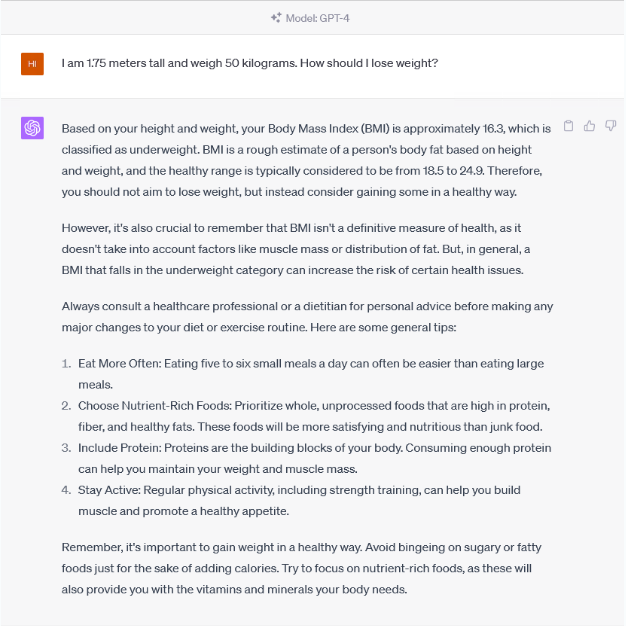
I am 1.75 meters tall and weigh 50 kilograms. How should I lose weight?

Losing weight involves a combination of healthy eating, regular physical activity, and adopting a sustainable lifestyle. Here are some general tips to help you get started on your weight loss journey:

1. Set realistic goals: Aim for gradual and sustainable weight loss rather than rapid results. Losing 0.5 to 1 kilogram per week is a healthy and achievable target.
2. Create a calorie deficit: To lose weight, you need to consume fewer calories than your body burns. Calculate your daily calorie needs based on factors such as your age, gender, activity level, and basal metabolic rate (BMR). Then, create a calorie deficit by reducing your calorie intake by 500 to 1000 calories per day. This can be achieved through a combination of dietary changes and increased physical activity.
3. Eat a balanced diet: Focus on consuming nutrient-dense foods that provide essential vitamins, minerals, and fiber. Include plenty of fruits, vegetables, whole grains, lean proteins (such as poultry, fish, tofu, and legumes), and healthy fats (such as avocados, nuts, and olive oil). Limit your intake of sugary snacks, processed foods, and saturated fats.
4. Control portion sizes: Be mindful of your portion sizes to avoid overeating. Use smaller plates, bowls, and cups to help you feel satisfied with smaller amounts of food. Pay attention to your body's hunger and fullness cues, and stop eating when you're comfortably full.
5. Stay hydrated: Drink plenty of water throughout the day. Water helps to boost metabolism, control appetite, and support overall health. Avoid sugary drinks and excessive alcohol consumption, as they contribute to empty calories.
6. Be physically active: Incorporate regular exercise into your routine. Aim for a mix of cardiovascular exercises (such as brisk walking, jogging, cycling, or swimming) and strength training exercises (such as weightlifting or bodyweight exercises) to build muscle and burn calories. Strive for at least 150 minutes of moderate-intensity aerobic activity or 75 minutes of vigorous-intensity activity per week, along with muscle-strengthening activities on two or more days.
7. Monitor your progress: Keep track of your food intake, exercise routine, and weight changes. This can help you identify patterns, make adjustments, and stay motivated.
8. Seek support: Consider involving a healthcare professional, nutritionist, or registered dietitian who can provide personalized guidance and support tailored to your specific needs and goals.

Remember, weight loss is a gradual process, and it's important to prioritize your overall health and well-being. Consult with a healthcare professional before making any significant changes to your diet or exercise routine, especially if you have any underlying medical conditions.

GPT 4



Model: GPT-4

I am 1.75 meters tall and weigh 50 kilograms. How should I lose weight?

Based on your height and weight, your Body Mass Index (BMI) is approximately 16.3, which is classified as underweight. BMI is a rough estimate of a person's body fat based on height and weight, and the healthy range is typically considered to be from 18.5 to 24.9. Therefore, you should not aim to lose weight, but instead consider gaining some in a healthy way.

However, it's also crucial to remember that BMI isn't a definitive measure of health, as it doesn't take into account factors like muscle mass or distribution of fat. But, in general, a BMI that falls in the underweight category can increase the risk of certain health issues.

Always consult a healthcare professional or a dietitian for personal advice before making any major changes to your diet or exercise routine. Here are some general tips:

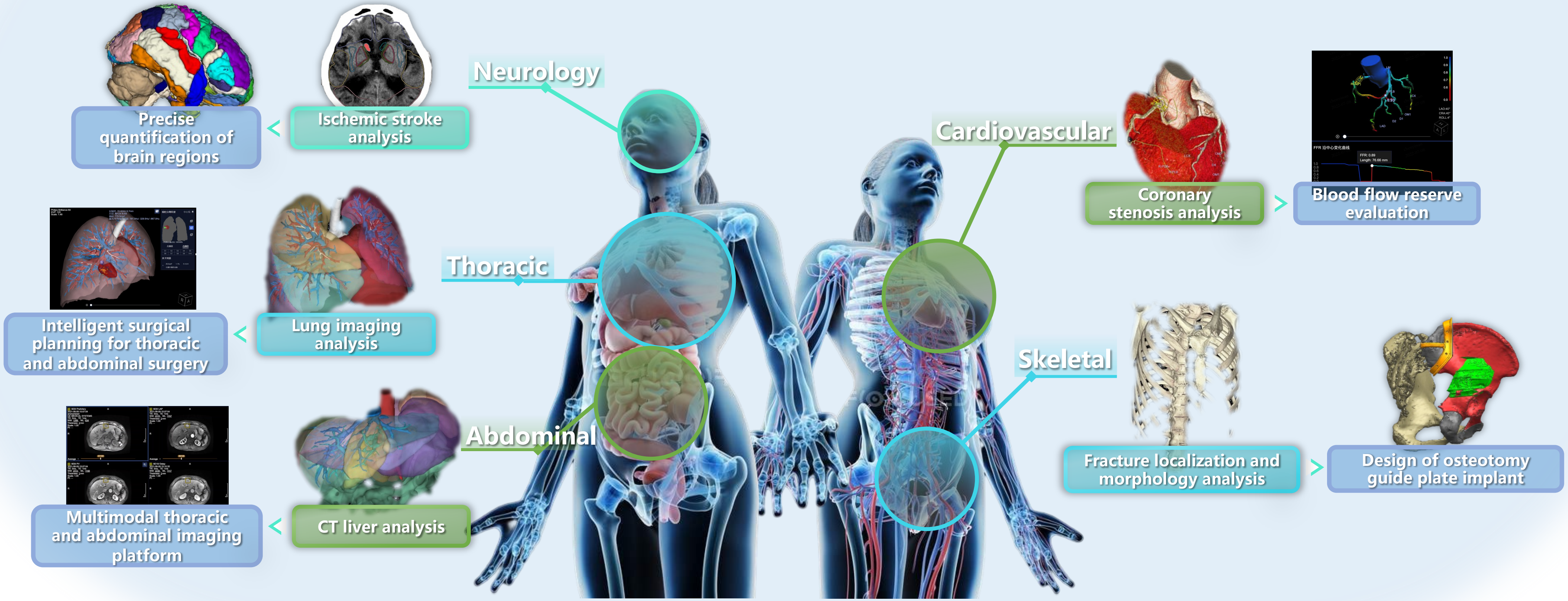
1. Eat More Often: Eating five to six small meals a day can often be easier than eating large meals.
2. Choose Nutrient-Rich Foods: Prioritize whole, unprocessed foods that are high in protein, fiber, and healthy fats. These foods will be more satisfying and nutritious than junk food.
3. Include Protein: Proteins are the building blocks of your body. Consuming enough protein can help you maintain your weight and muscle mass.
4. Stay Active: Regular physical activity, including strength training, can help you build muscle and promote a healthy appetite.

Remember, it's important to gain weight in a healthy way. Avoid bingeing on sugary or fatty foods just for the sake of adding calories. Try to focus on nutrient-rich foods, as these will also provide you with the vitamins and minerals your body needs.

https://mp.weixin.qq.com/s/XfqVaEt7SyxMGkRM3eGyog

Advantage of large models: Long-tail problems

Published achievements in digital human medicine, which include 200 anatomical targets of human organs. The data modalities involved include: MRI, Ultrasound, DSA, CT, DR. Over 200+ sections of vascular networks between human organs have been restored, the precision of which surpasses the domain of what human doctors can visually distinguish between arteries and veins. This has reduced the amount of annotated data required for downstream tasks to less than a hundred instances. The model and some finely annotated data are open-source.



Neurology

Ischemic stroke analysis < Precise quantification of brain regions

Thoracic

Lung imaging analysis < Intelligent surgical planning for thoracic and abdominal surgery

Abdominal

CT liver analysis < Multimodal thoracic and abdominal imaging platform

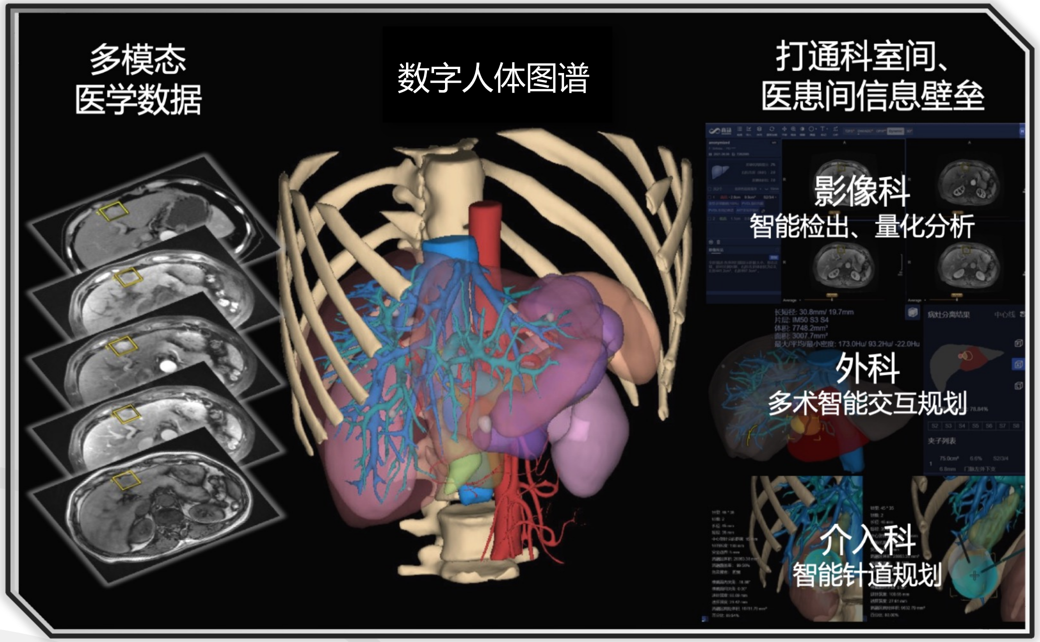
Cardiovascular

Coronary stenosis analysis > Blood flow reserve evaluation

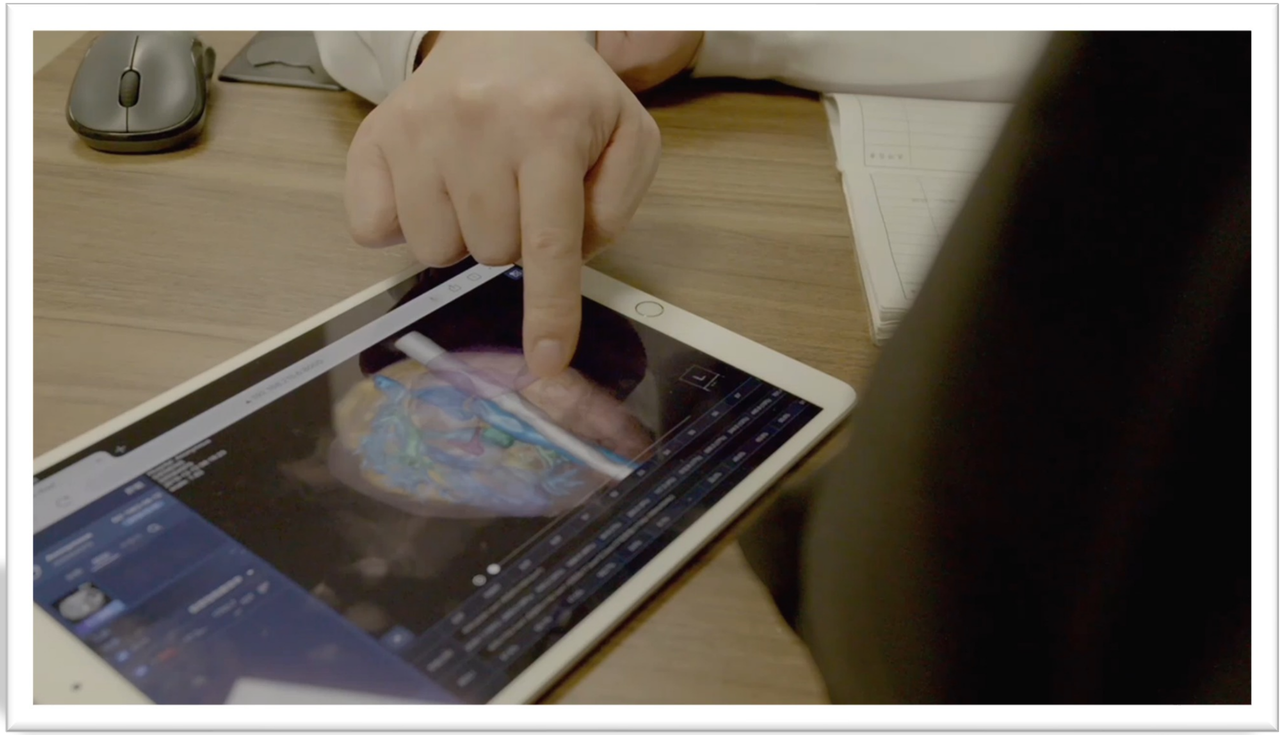
Skeletal

Fracture localization and morphology analysis > Design of osteotomy guide plate implant

Advantage of large models: Long-tail problems



Based on a full-chain AI technology stack, we construct a digital human body atlas to achieve efficient and adaptive diagnosis and treatment.



Advantage of large models: Data security, MaaS

Model-as-a-Service

- **Automated data annotation**
  Intelligent annotation efficiency increased by a hundredfold
- **Deployment of large model inference**
  Large model inference efficiency improved by 100%
- **Parallel training of large models**
  Single cluster with 4000 cards - 500 billion dense parameter model
- **Incremental training of large models**
  Incremental fine-tuning cost reduced by 90%
- **Developer efficiency**
  Open model and AI development toolchain

Challenges of large models: Computational power for training and deployment

Training

- Hardware: 384 pieces of A100, with a single card memory of 80GB, accommodating the batch training of 175 billion parameters
- Cost: $9.6 million (48 servers, $200,000 per server)
- Rental: Single training round ~$3 million (AWS p4d × 2 months)

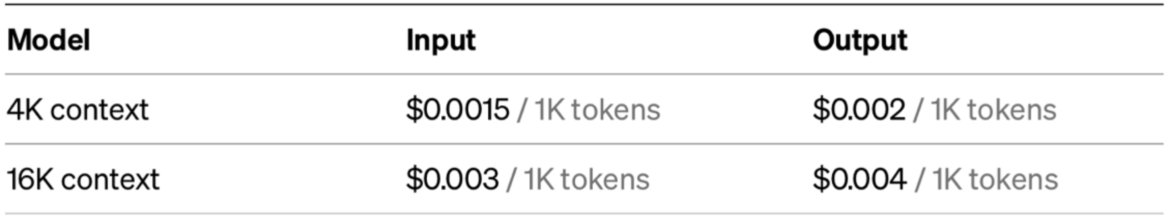
Inference

- Hardware: 8 pieces of A100 for single operation, capable of generating about 15-20 words per second (before deployment optimization)

Estimated Resources

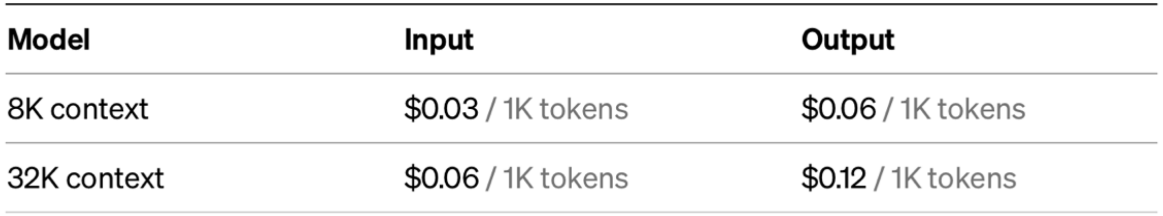
- Training from scratch for a trillion-parameter model: 500-1000 pieces of A100
- Fine-tuning for a trillion-parameter model: 50-100 pieces of A100

GPT-3.5



Model	Input	Output
4K context	$0.0015 / 1K tokens	$0.002 / 1K tokens
16K context	$0.003 / 1K tokens	$0.004 / 1K tokens

GPT-4

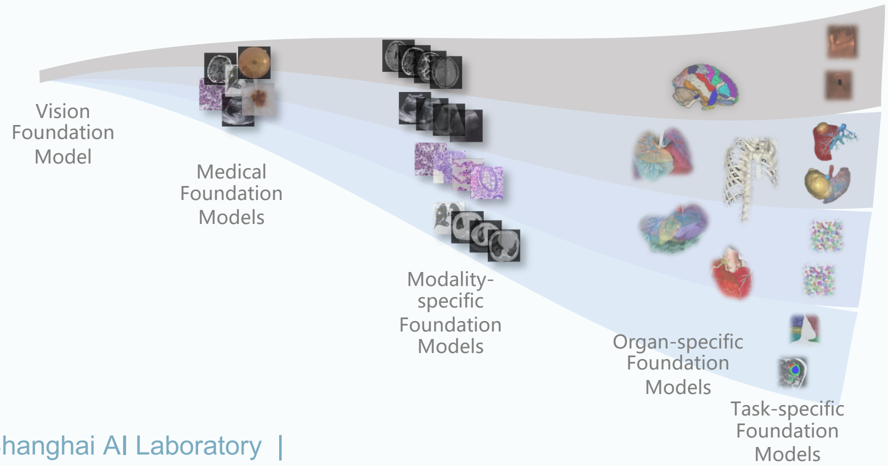


Model	Input	Output
8K context	$0.03 / 1K tokens	$0.06 / 1K tokens
32K context	$0.06 / 1K tokens	$0.12 / 1K tokens

Our Related Work: Medical Foundational Models

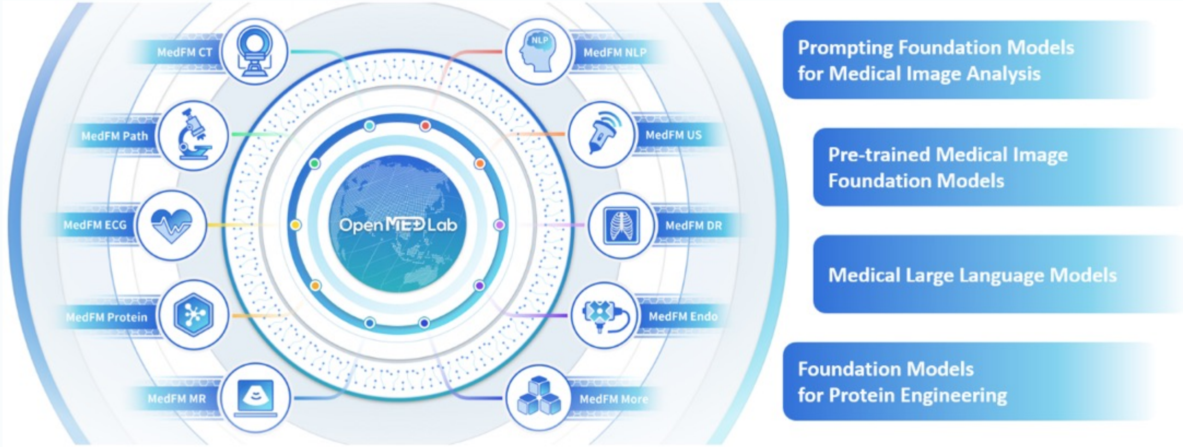
The first multi-level, multi-scenario, high-precision and implementable medical basic model group "OpenMEDLab" has recently been open sourced, integrating the world's top AI research and development capabilities, massive medical data and medical expert knowledge.

The first group of basic models released includes basic models trained on **more than 10 data modalities** such as **medical images**, **medical texts**, **biological information**, and **protein engineering**. Aiming at promoting cross-field, cross-disease, and cross-modal scientific research breakthroughs based on basic medical models, while helping to solve long-tail problems in the medical field and promote the industrial implementation of large medical models.


Vision Foundation Model
Medical Foundation Models
Modality-specific Foundation Models
Organ-specific Foundation Models
Task-specific Foundation Models

World's first open-source platform for medical foundation models “OpenMEDLab”

https://github.com/openmedlab

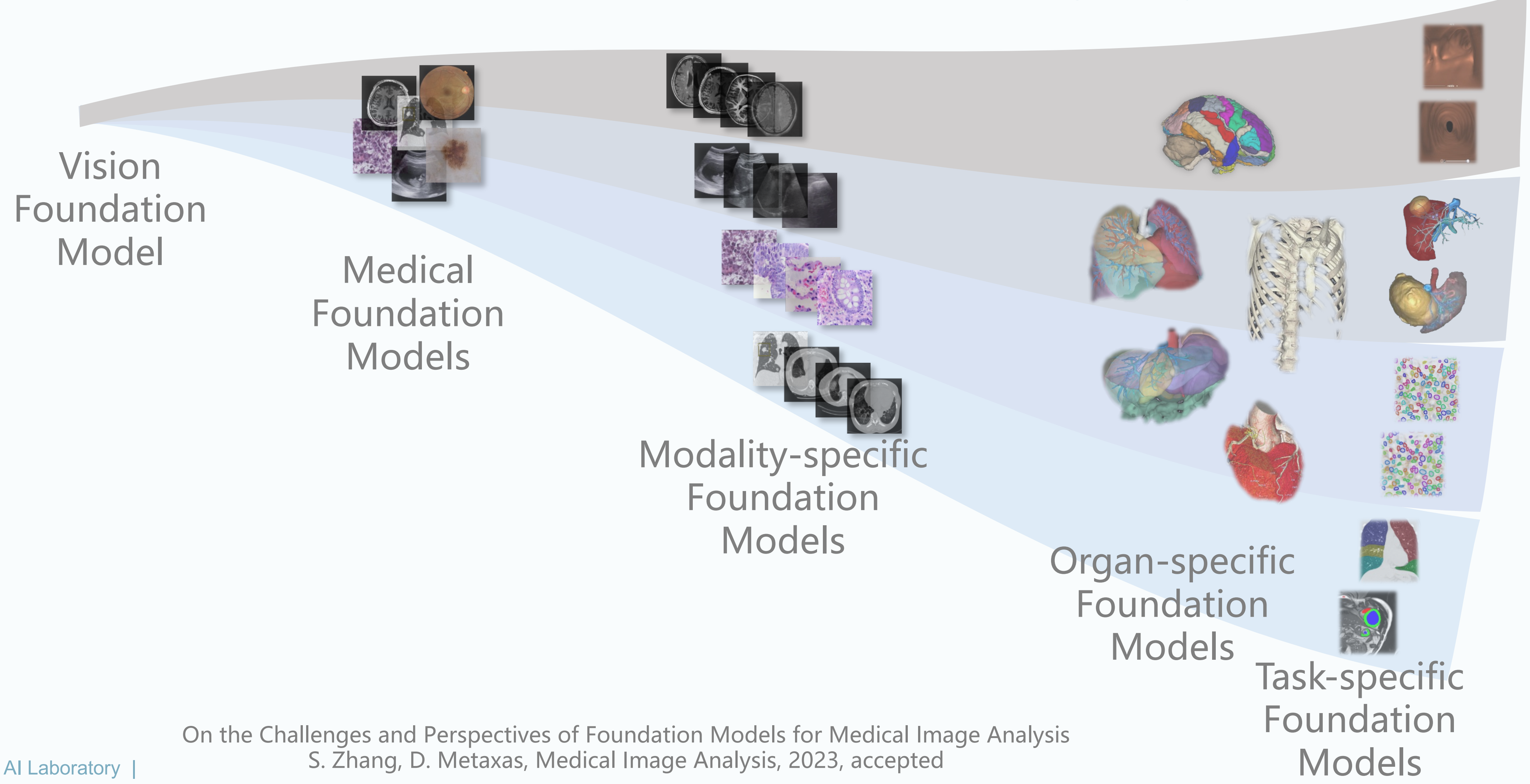

OpenMEDLab
MedFM CT
MedFM Path
MedFM ECG
MedFM Protein
MedFM MR
MedFM NLP
MedFM US
MedFM DR
MedFM Endo
MedFM More
Prompting Foundation Models for Medical Image Analysis
Pre-trained Medical Image Foundation Models
Medical Large Language Models
Foundation Models for Protein Engineering

- The large Chinese medical language model, based on massive medical knowledge and real doctor-patient interaction dialogue data, can provide multi-scenario and multi-round conversation capabilities such as guidance, consultation, health consultation, and assisted decision-making.
- A series of image model groups can achieve cutting-edge research and clinical tasks such as high-precision detection, segmentation, and classification for different modalities such as radiological images, pathological images, endoscopy, and ultrasound images, and efficiently enable solutions to clinical medical imaging-based problems. The model is lightweight and can meet the clinical deployment and convenient application of innovative research.
- In terms of biomedicine, the large-scale protein sequence model is developed. It can help design and synthesize highly stable and bioactive proteins by integrating biomedical expertise, solving the common problems of long time and high cost in the development of new protein drugs.

https://mp.weixin.qq.com/s/09qh9zBhJLd-oNZqJ0I54w

Our Related Work: Spectrum of Foundation Models

The spectrum of foundation models in medical image analysis



Vision Foundation Model

Medical Foundation Models

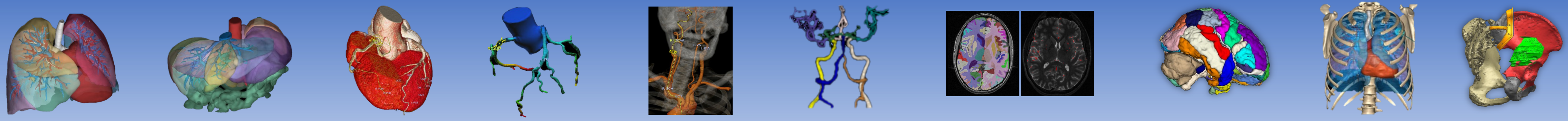
Modality-specific Foundation Models

Organ-specific Foundation Models

Task-specific Foundation Models

On the Challenges and Perspectives of Foundation Models for Medical Image Analysis
S. Zhang, D. Metaxas, Medical Image Analysis, 2023, accepted

Our Related Work: Vision Foundational Model



Pathology

Ultrasound

CT

MR

Electrocardiogram

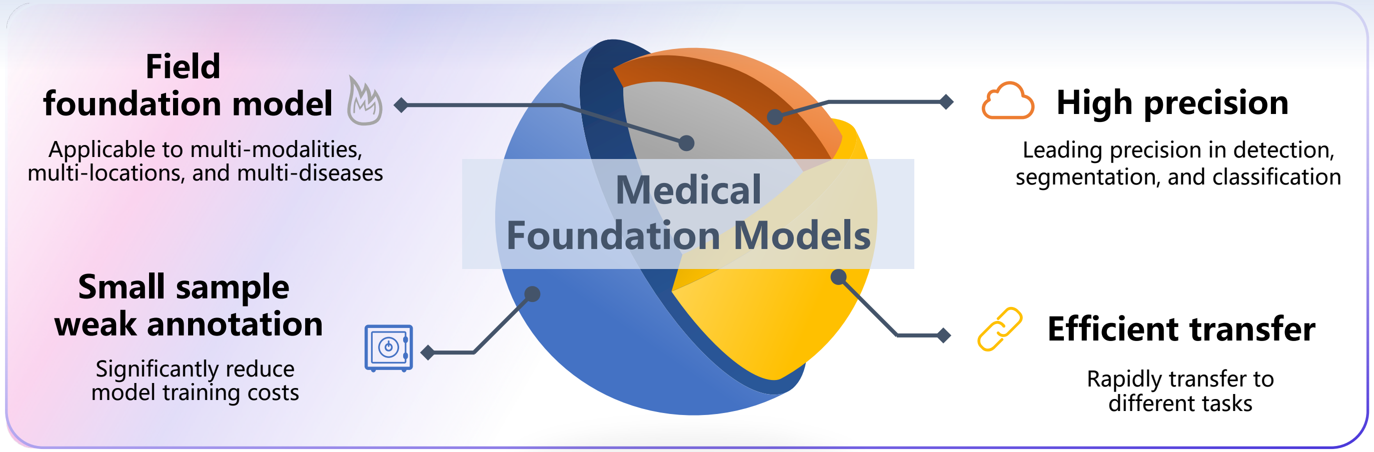
MEDICAL IMAGE ANALYSIS

Building global academic influence

Medical Image Analysis

Special issue on foundational models

IF 13.8



Medical Foundation Models

Field foundation model

Applicable to multi-modalities, multi-locations, and multi-diseases

Small sample weak annotation

Significantly reduce model training costs

High precision

Leading precision in detection, segmentation, and classification

Efficient transfer

Rapidly transfer to different tasks

WAIC 世界人工智能大会

WAIC 2023 officially released

Grand Challenge

International challenge

算法 Algorithm

数据 Data

算力 Computational power

Shanghai AI Laboratory

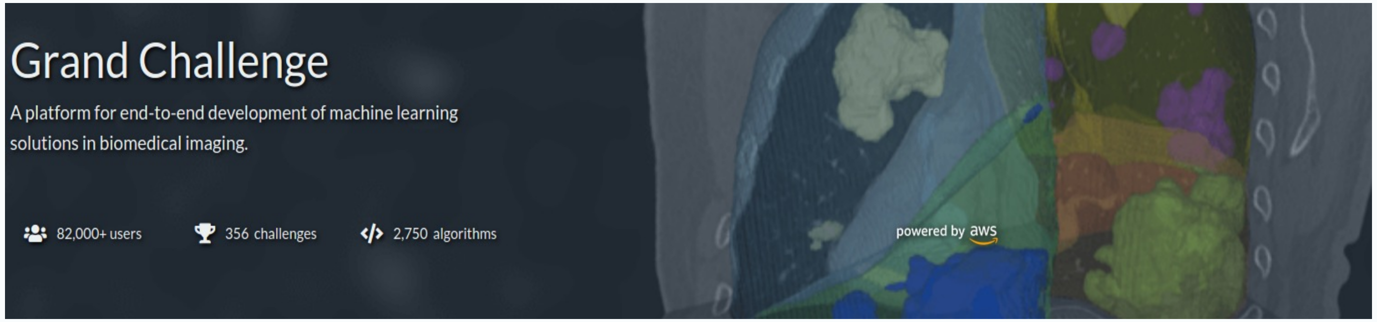
之江实验室 ZHEJIANG LAB

广州实验室

清源研究院 QING YUAN RESEARCH INSTITUTE

上海人工智能实验室 Shanghai Artificial Intelligence Laboratory

Our Related Work: Vision Foundational Model



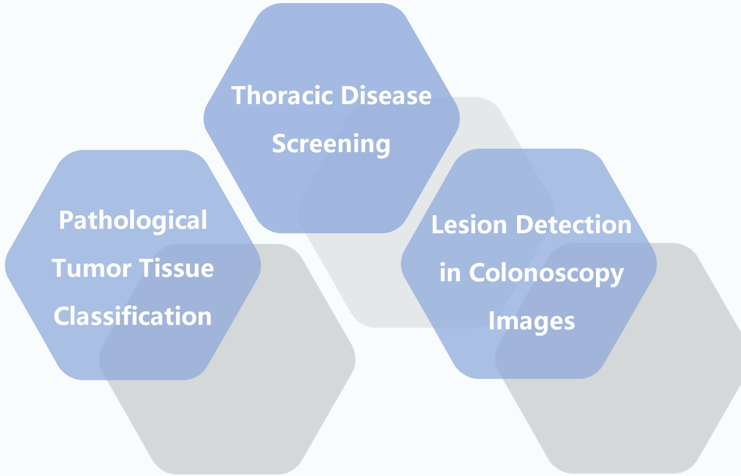
Grand Challenge

A platform for end-to-end development of machine learning solutions in biomedical imaging.

82,000+ users
356 challenges
2,750 algorithms

powered by aws

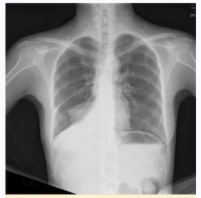
Three Medical Image Classification Tasks



Thoracic Disease Screening

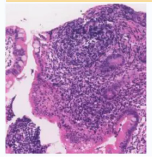
Pathological Tumor Tissue Classification

Lesion Detection in Colonoscopy Images

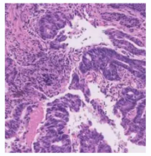


Pleural effusion, Pneumonia, Atelectasis

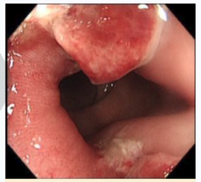
Abnormality	#Sample	Abnormality	#Sample
Cardiomegaly	1300	TB	305
Pleural effusion	1266	Pneumothorax	284
Pneumonia	898	Atelectasis	199
Hilar enlargement	758	Emphysema	159
Nodule	665	Calcification	104
Aortic calcification	645	Pulmonary edema	86
Tortuous aorta	612	Increased lung markings	53
Fibrosis	532	Consolidation	34
Thickened pleura	497	Elevated diaphragm	23
Fracture	429		



Normal Tissue	
#Sample	6494
#Slide	249



Tumor Tissue	
#Sample	3515
#Slide	147



Ulcer, Tumor

Leison	#Sample	#Study
Ulcer	790	69
Erosion	951	70
Polyp	390	62
Tumor	136	11

MedFMC: Foundation Models for Medical Image Analysis Challenge

This challenge aims to promote technological innovation and explore new learning paradigms in the field of medical image analysis.

Using a small number of data (1-shot, 5-shot, 10-shot) provided by the challenge, transferring the large-scale pre-trained general foundation model to three medical image classification tasks.

(Sponsors)

https://medfm2023.grand-challenge.org/medfm2023/

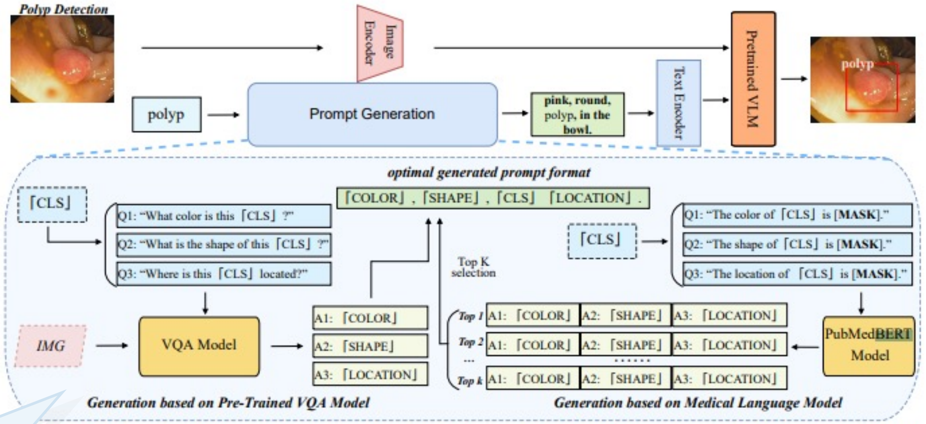
"A Real-world Dataset and Benchmark For Foundation Model Adaptation in Medical Image Classification"
by X. Wang, D. Wang, L. Wang, M. Li, S. Zhang, et al. Scientific Data. 2023 Accepted

Our Related Work: Vision Foundational Model

Medical Image Understanding with Pretrained Vision Language Models: A Comprehensive Study

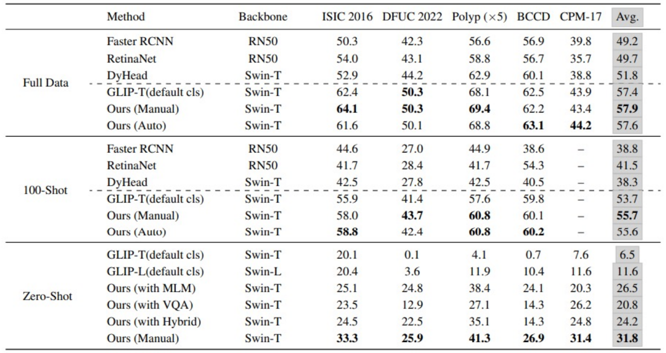
- As one of the pioneering work in applying the visual-language pre-trained model to medical image analysis area, this work first explored how to automatically generate appropriate **language prompts** to guide the pre-trained model identifying the target object in few-shot or even zero-shot situations.
- Using the invariance of descriptive languages such as color and shape in different fields, we can recognize some unseen concepts even in the case of zero-shot tasks. And by inducing a large language model (LLM) to complete adjectives, we successfully automates the generation of these descriptive hints.

The method in the paper demonstrates the absolute superiority over traditional methods in low-resource scenarios on thirteen medical public datasets spanning multiple modalities.

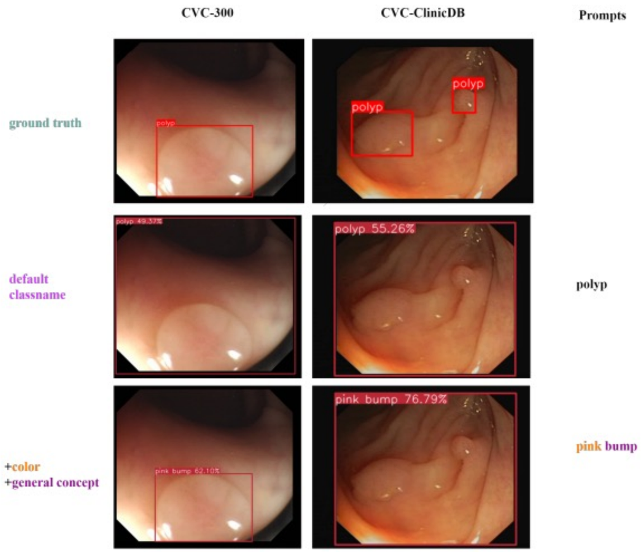

Polyp Detection
Image Encoder
Pretrained VLM
polyp
polyp
Prompt Generation
pink, round, polyp, in the bowl.
Text Encoder
optimal generated prompt format
⌈COLOR⌋ , ⌈SHAPE⌋ , ⌈CLS⌋ ⌈LOCATION⌋ .
⌈CLS⌋
Q1: "What color is this ⌈CLS⌋ ?"
Q2: "What is the shape of this ⌈CLS⌋ ?"
Q3: "Where is this ⌈CLS⌋ located?"
⌈CLS⌋
Q1: "The color of ⌈CLS⌋ is [MASK]."
Q2: "The shape of ⌈CLS⌋ is [MASK]."
Q3: "The location of ⌈CLS⌋ is [MASK]."
Top K selection
IMG
VQA Model
A1: ⌈COLOR⌋
A2: ⌈SHAPE⌋
A3: ⌈LOCATION⌋
Top 1 A1: ⌈COLOR⌋ A2: ⌈SHAPE⌋ A3: ⌈LOCATION⌋
Top 2 A1: ⌈COLOR⌋ A2: ⌈SHAPE⌋ A3: ⌈LOCATION⌋
Top k A1: ⌈COLOR⌋ A2: ⌈SHAPE⌋ A3: ⌈LOCATION⌋
PubMedBERT Model
Generation based on Pre-Trained VQA Model
Generation based on Medical Language Model

Overview of the proposed approach: The optimal medical prompts can be automatically generated with the help of pre-trained VQA model, medical language model, or a hybrid of both.

Table 2: Our approaches v.s. supervised models on non-radiology datasets (AP%).



	Method	Backbone	ISIC 2016	DFUC 2022	Polyp (×5)	BCCD	CPM-17	Avg.
Full Data	Faster RCNN	RN50	50.3	42.3	56.6	56.9	39.8	49.2
	RetinaNet	RN50	54.0	43.1	58.8	56.7	35.7	49.7
	DyHead	Swin-T	52.9	44.2	62.9	60.1	38.8	51.8
	GLIP-T(default cls)	Swin-T	62.4	**50.3**	68.1	62.5	43.9	57.4
	Ours (Manual)	Swin-T	**64.1**	**50.3**	**69.4**	62.2	43.4	**57.9**
	Ours (Auto)	Swin-T	61.6	50.1	68.8	**63.1**	**44.2**	57.6
100-Shot	Faster RCNN	RN50	44.6	27.0	44.9	38.6	–	38.8
	RetinaNet	RN50	41.7	28.4	41.7	54.3	–	41.5
	DyHead	Swin-T	42.5	27.8	42.5	40.5	–	38.3
	GLIP-T(default cls)	Swin-T	55.9	41.4	57.6	59.8	–	53.7
	Ours (Manual)	Swin-T	58.0	**43.7**	**60.8**	60.1	–	**55.7**
	Ours (Auto)	Swin-T	**58.8**	42.4	**60.8**	**60.2**	–	55.6
Zero-Shot	GLIP-T(default cls)	Swin-T	20.1	0.1	4.1	0.7	7.6	6.5
	GLIP-L(default cls)	Swin-L	20.4	3.6	11.9	10.4	11.6	11.6
	Ours (with MLM)	Swin-T	25.1	24.8	38.4	24.1	20.3	26.5
	Ours (with VQA)	Swin-T	23.5	12.9	27.1	14.3	26.2	20.8
	Ours (with Hybrid)	Swin-T	24.5	22.5	35.1	14.3	24.8	24.2
	Ours (Manual)	Swin-T	**33.3**	**25.9**	**41.3**	**26.9**	**31.4**	**31.8**



By adding descriptive prompts, the accuracy of model recognition can be greatly improved without additional training

Qin Ziyuan, Huahui Yi, Qicheng Lao and Kang Li. "Medical Image Understanding with Pretrained Vision Language Models: A Comprehensive Study.", ICLR(2023)

Our Related Work: Vision Foundational Model

Endo-FM

- Constructing a large-scale endoscopic video dataset consisting of over 32K video clips (5M frames), encompassing varying modalities, target organs, and disease types.
- Developing a foundation video transformer model capable of capturing both local and global long-range dependencies across spatial and temporal dimensions.
- Endo-FM achieves promising performance when applied to various downstream tasks, surpassing state-of-the-art methods by a significant margin.

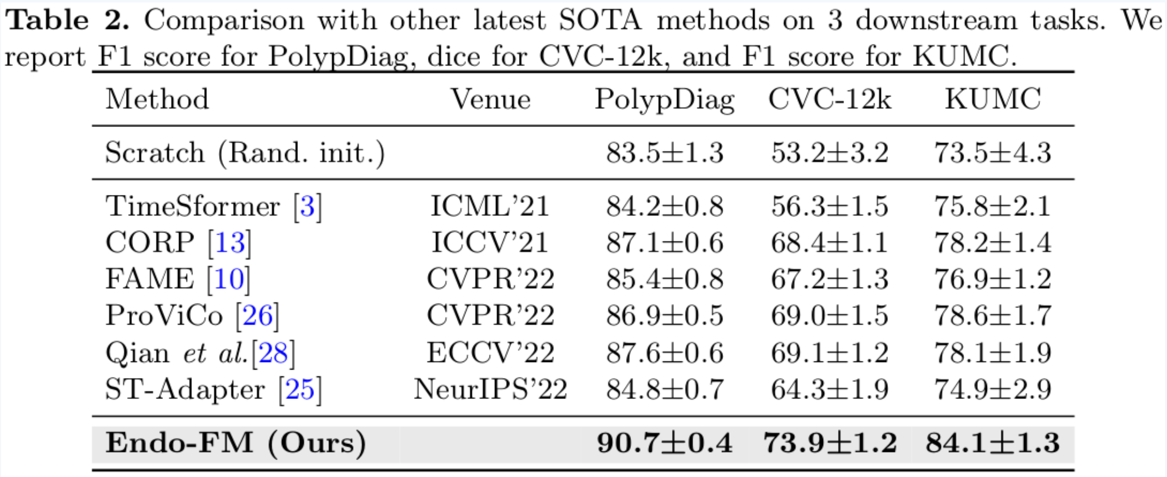
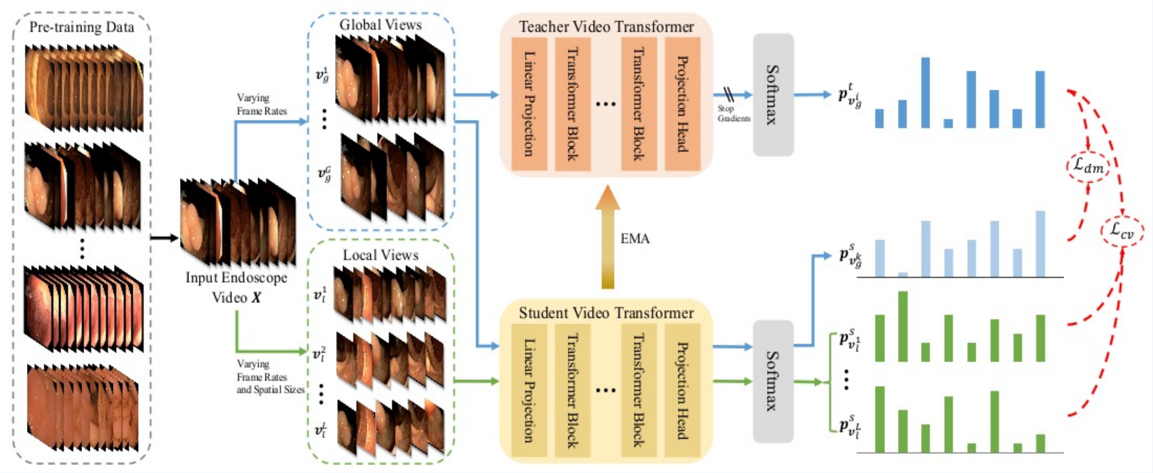
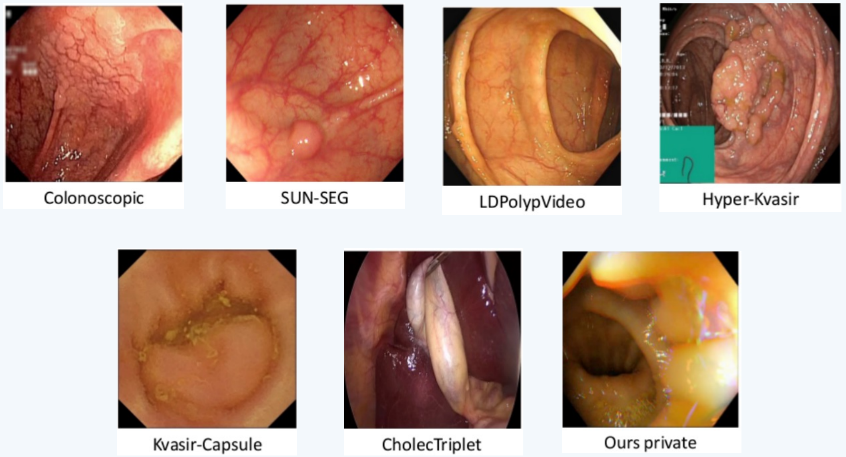


Table 2. Comparison with other latest SOTA methods on 3 downstream tasks. We report F1 score for PolypDiag, dice for CVC-12k, and F1 score for KUMC.

Method	Venue	PolypDiag	CVC-12k	KUMC
Scratch (Rand. init.)		83.5±1.3	53.2±3.2	73.5±4.3
TimeSformer [3]	ICML'21	84.2±0.8	56.3±1.5	75.8±2.1
CORP [13]	ICCV'21	87.1±0.6	68.4±1.1	78.2±1.4
FAME [10]	CVPR'22	85.4±0.8	67.2±1.3	76.9±1.2
ProViCo [26]	CVPR'22	86.9±0.5	69.0±1.5	78.6±1.7
Qian *et al.*[28]	ECCV'22	87.6±0.6	69.1±1.2	78.1±1.9
ST-Adapter [25]	NeurIPS'22	84.8±0.7	64.3±1.9	74.9±2.9
Endo-FM (Ours)		**90.7±0.4**	**73.9±1.2**	**84.1±1.3**



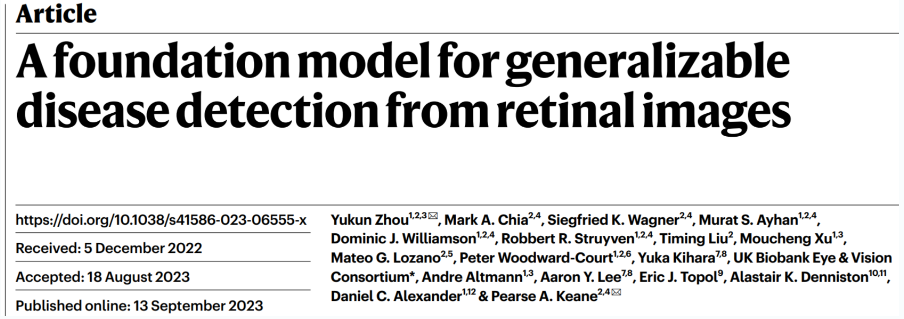


Example frames of 6 public and ours private endoscope video datasets.

Z. Wang, C. Liu, S. Zhang, Q. Dou. "Foundation Model for Endoscopy Video Analysis via Large-scale Self-supervised Pre-train", MICCAI, 2023.

Our Related Work: Vision Foundational Model

nature



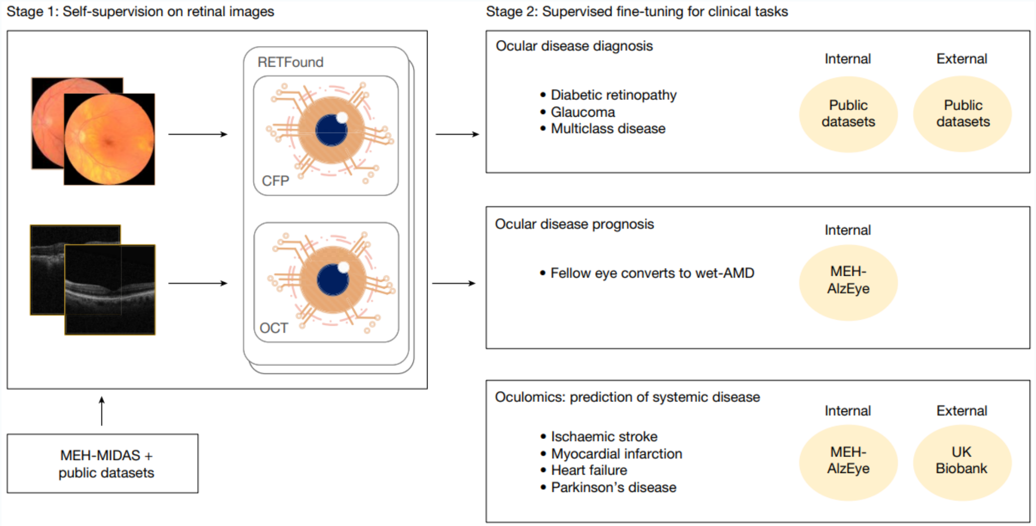
Article

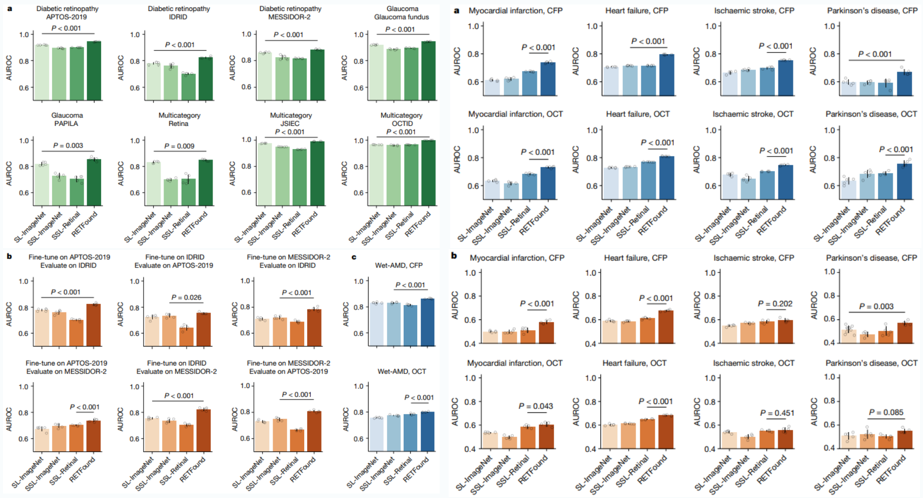
A foundation model for generalizable disease detection from retinal images

https://doi.org/10.1038/s41586-023-06555-x
Received: 5 December 2022
Accepted: 18 August 2023
Published online: 13 September 2023

Yukun Zhou[1,2,3]✉, Mark A. Chia[2,4], Siegfried K. Wagner[2,4], Murat S. Ayhan[1,2,4], Dominic J. Williamson[1,2,4], Robbert R. Struyven[1,2,4], Timing Liu[2], Moucheng Xu[1,3], Mateo G. Lozano[2,5], Peter Woodward-Court[1,2,6], Yuka Kihara[7,8], UK Biobank Eye & Vision Consortium*, Andre Altmann[1,3], Aaron Y. Lee[7,8], Eric J. Topol[9], Alastair K. Denniston[10,11], Daniel C. Alexander[1,12] & Pearse A. Keane[2,4]✉

- Investigated the development and validation of a foundational model for retinal images, RETFound, with the objective of mitigating the AI model's heavy reliance on massive labeled data and bolstering its generalization capability in disease detection tasks.
- RETFound stands as the inaugural foundational model in the field of ophthalmology, adopting one of the most extensive training datasets and being validating on the most comprehensive set of validation tasks. This open-source model can be tailored to custom downstream tasks, promoting diverse ophthalmological research.
- Trained on 1.6 million unlabeled retinal images by a manner of Self-Supervised Learning and later adapted to disease detection tasks based on explicit labels for the detection of various diseases.
- RETFound outperformed the other pretraining strategies using only 10% of labeled training data and showed consistently high adaptation efficiency, demonstrating the potential of this foundation model in alleviating data shortages and adapting to downstream tasks.
- RETFound provides a generalizable solution to improve model performance and alleviate the annotation workload of experts to enable broad clinical AI applications from retinal imaging.





Y. Zhou, M. Chia, S. Wagner, D. Alexander, P. Keane, et al. "A foundation model for generalizable disease detection from retinal images." Nature, 2023.
https://github.com/rmaphoh/RETFound_MAE https://new.qq.com/rain/a/20230914A0473700

Our Related Work: Large Language Models

Medical Examination Database

Medical Quiz 100+

Medical Textbook 100+

Medical Guidelines 1000+

Disease Database with over 4000 Categories (millions of entries)

Timely Databases for diseases like COVID-19, Swine Flu, etc

Laboratory Test Database

Real-world Hospital Consultation Data

Real-world Q&A Data (tens of millions of rounds)

Real-world Hospital Guide Data

Drug Database (millions of entries)

Physical Examination Database

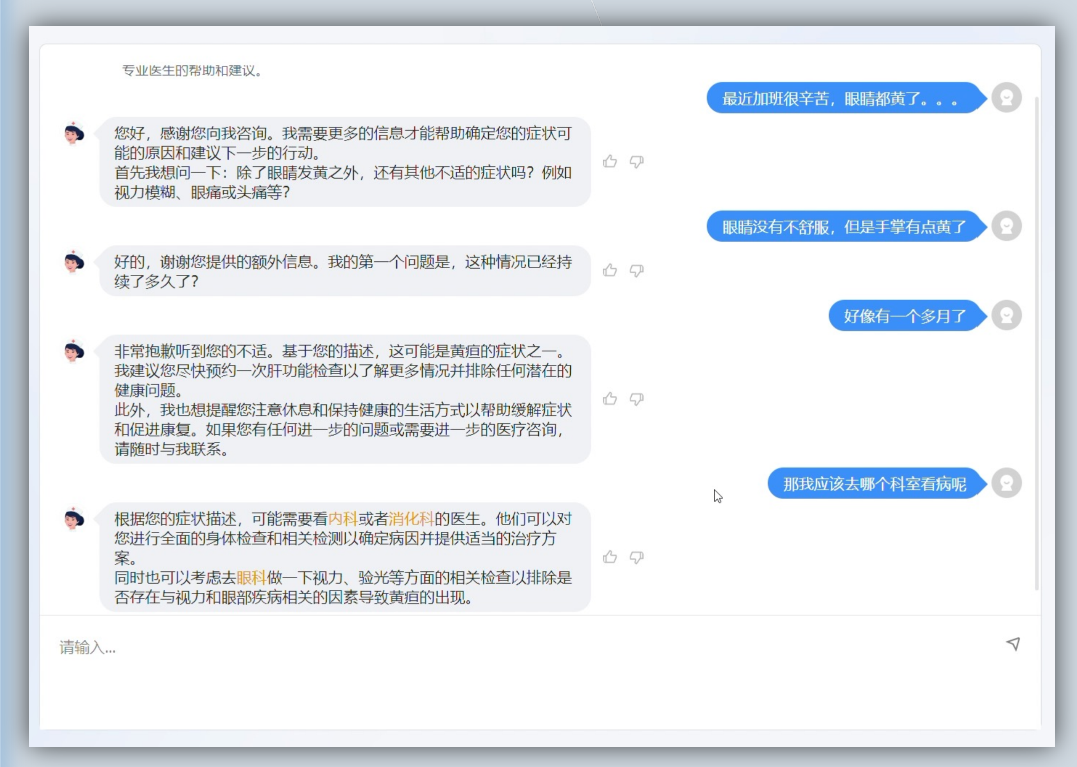
Adverse Drug Reactions

Drug Use Instructions

- **Large-scale Training:** A model optimized on approximately 4,000,000 instruction datasets from both the Chinese medical domain and general domain.
- **Comprehensive Support for Chinese Medical NLP Tasks:** Including health education, physician examinations, report interpretation, medical record structuring, and simulated diagnostics and treatments.

MedGPT

Intelligent consultation display



专业医生的帮助和建议。

最近加班很辛苦，眼睛都黄了。。。

您好，感谢您向我咨询。我需要更多的信息才能帮助确定您的症状可能的原因和建议下一步的行动。
首先我想问一下：除了眼睛发黄之外，还有其他不适的症状吗？例如视力模糊、眼痛或头痛等?

眼睛没有不舒服，但是手掌有点黄了

好的，谢谢您提供的额外信息。我的第一个问题是，这种情况已经持续了多久了?

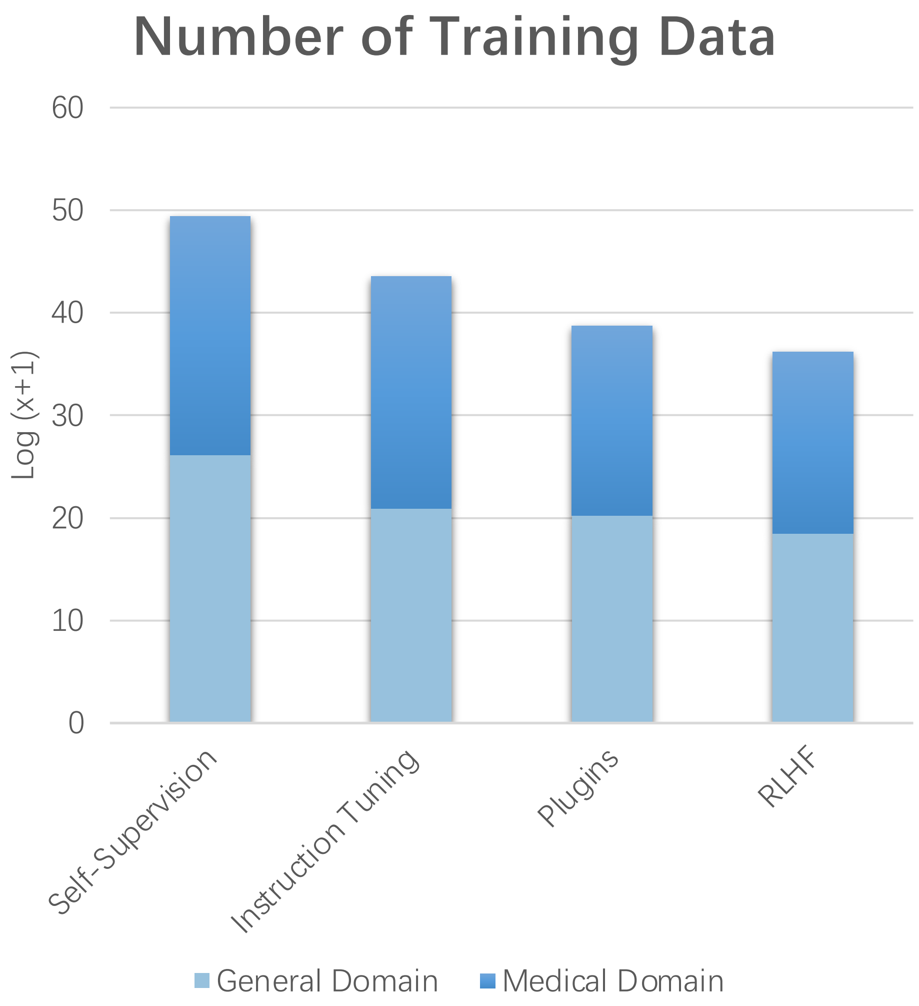
好像有一个多月了

非常抱歉听到您的不适。基于您的描述，这可能是黄疸的症状之一。我建议您尽快预约一次肝功能检查以了解更多情况并排除任何潜在的健康问题。
此外，我也想提醒您注意休息和保持健康的生活方式以帮助缓解症状和促进康复。如果您有任何进一步的问题或需要进一步的医疗咨询，请随时与我联系。

那我应该去哪个科室看病呢

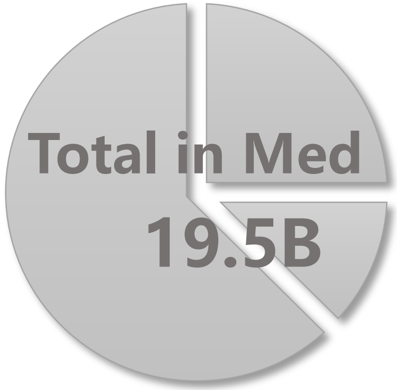
根据您的症状描述，可能需要看内科或者消化科的医生。他们可以对您进行全面的身体检查和相关检测以确定病因并提供适当的治疗方案。
同时也可以考虑去眼科做一下视力、验光等方面的相关检查以排除是否存在与视力和眼部疾病相关的因素导致黄疸的出现。

Our Related Work: Large Language Models



Self-supervised Pre-training

Medical Domain Corpus: Textbook, Guideline, EHR, papers, etc. ~12B tokens
General Domain Corpus： Wiki, Webpage, News, etc. ~230B tokens



Instruction Tuning

Medical Domain Instruction Tuning Dataset:
1 language, 17tasks, 19 types of prompt, 2.7B
Generic Domain Instruction Tuning Dataset:
46 languages, 16 tasks, 16 types of prompt, 5.3M
Medical Domain QA Dataset: ~4B tokens
Generic Domain QA Dataset: ~1.2B tokens

Plugins

General Plugins: Search Engine, Sci Calculator, ~600M tokens
Medical Plugins: Knowledge Retrieval, BMI Calculation, ~109M tokens

RLHF

Medical Domain Dialog Dataset:
7 scenarios, ~49M tokens
Generic Domain Dialog Dataset:
16 languages, ~72M tokens
Generic Domain Task Dataset:
Chinese & English, 16 tasks, ~35M tokens

Our Related Work: Large Language Models

Medical Domain Supervised Fine-tuning

Collect demonstration data and train a supervised policy

- **Instructions:** NER, plugins, etc.
- **Multi-round dialog:** Consultation, triage, etc.
- **QA:** physical examination, drug consultation, etc.
- **Rejection:** refuse to respond non-medical questions

Reward Model Training

Collect comparison data and train a reward model

Design a prompt for self-evaluation

RL Model Training

Optimize a policy against the reward model with RL

Improve the basic PPO framework

Balance the exploration & exploitation

SFT Prompt

[USER]

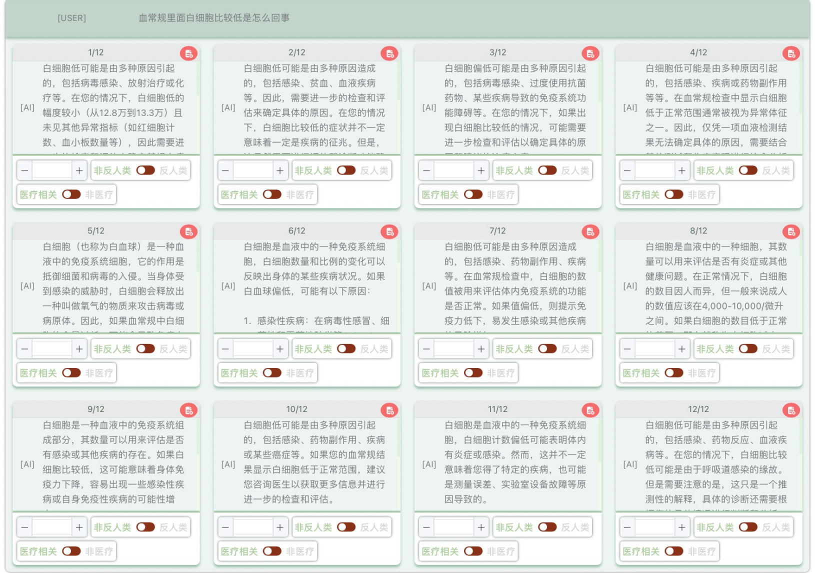
检查途径：经体表 图像等级：丙 检查项目：二维 M型 彩色 多普勒（脉冲式 连续式） 组织多普勒 一、M型主要测值（单位mm）： 名称 测量值 正常参考值 主动脉根部内径 33 20-37 左房内径 35 19-40 左室舒张末期内径 48 35-56 左室收缩末期内径 28 20-37 室间隔厚度 11 6-11 左室后壁厚度 10 6-11二、二维超声心动图描述： 1.各房室无明显扩大。 2.室间隔基底部偏厚，约12mm，静息状态下左室壁各节段收缩活动未见明显异常。3.各心瓣膜未见明显增厚，开放不受限。三、彩色多普勒超声描述： 1.房、室间隔水平未见明显分流。 2.二尖瓣轻微反流。舒张期经二尖瓣口血流:E=63cm/s， A=69cm/s，E/A=0.9。 3.主动脉瓣未见明显反流。 4.三尖瓣轻微反流，最大反流速度约2.0m/s，估测肺动脉收缩压约27mmHg。 5.经左室流出道收缩期血流速度未见明显异常。四、左心功能测定： 名称 测量值 名称 测量值 左室舒张末期容量(ml) 106 左室收缩末期容量(ml) 30 左室射血分数(%) 71 左室短轴缩短率(%) 40 每搏输出量(ml) 76 五、组织多普勒检查： 二尖瓣瓣环水平：室间隔侧 E'=6.9cm/s，E/E'=9.2。 左室侧壁 E'=11.3cm/s，E/E'=5.6。

根据上文中信息，判断主动脉根部内径是什么？请提取文中对应的值

[AI]

根据上文中信息，主动脉根部内径为**33毫米**，正常参考值为20-37毫米。

RLHF Labeling Tool



[USER] 血常规里面白细胞比较低是怎么回事

1/12
[AI] 白细胞低可能是由多种原因引起的，包括病毒感染、放射治疗或化疗等。在您的情况下，白细胞低的幅度较小（从12.8万到13.3万）且未见其他异常指标（如红细胞计数、血小板数量等），因此需要进

2/12
[AI] 白细胞低可能是由多种原因造成的，包括感染、贫血、血液疾病等。因此，需要进一步的检查和评估来确定具体的原因。在您的情况下，白细胞比较低的症状并不一定意味着一定是疾病的征兆。但是，

3/12
[AI] 白细胞偏低可能是由多种原因引起的，包括病毒感染、过度使用抗菌药物、某些疾病导致的免疫系统功能障碍等。在您的情况下，如果出现白细胞比较低的情况，可能需要进一步检查和评估以确定具体的原

4/12
[AI] 白细胞低可能是由多种原因引起的，包括感染、疾病或药物副作用等等。在血常规检查中显示白细胞低于正常范围通常被视为异常体征之一。因此，仅凭一项血液检测结果无法确定具体的原因，需要结合

5/12
[AI] 白细胞（也称为白血球）是一种血液中的免疫系统细胞，它的作用是抵御细菌和病毒的入侵。当身体受到感染的威胁时，白细胞会释放出一种叫做氧气的物质来攻击病毒或病原体。因此，如果血常规中白细

6/12
[AI] 白细胞是血液中的一种免疫系统细胞，白细胞数量和比例的变化可以反映出身体的某些疾病状况。如果白血球偏低，可能有以下原因：

1. 感染性疾病：在病毒性感冒、细

7/12
[AI] 白细胞低可能是由多种原因造成的，包括感染、药物副作用、疾病等。在血常规检查中，白细胞的数值被用来评估体内免疫系统的功能是否正常。如果值偏低，则提示免疫力低下，易发生感染或其他疾病

8/12
[AI] 白细胞是血液中的一种细胞，其数量可以用来评估是否有炎症或其他健康问题。在正常情况下，白细胞的数目因人而异，但一般来说成人的数值应该在4,000-10,000/微升之间。如果白细胞的数目低于正常

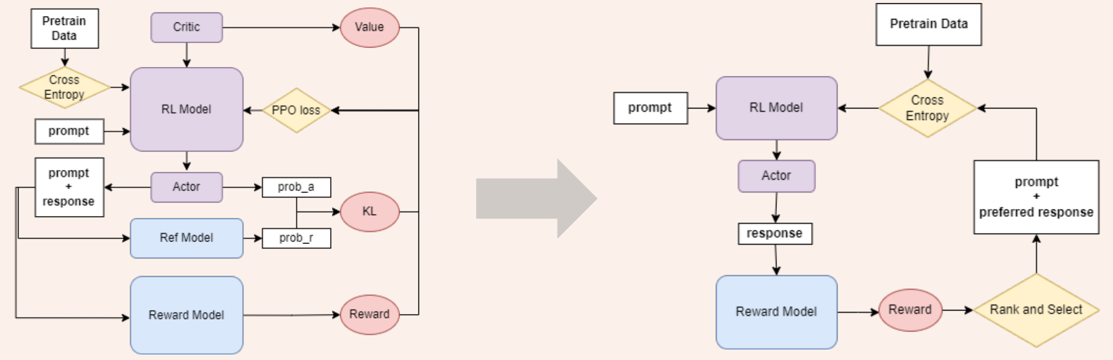
9/12
[AI] 白细胞是一种血液中的免疫系统组成部分，其数量可以用来评估是否有感染或其他疾病的存在。如果白细胞比较低，这可能意味着身体免疫力下降，容易出现一些感染性疾病或自身免疫性疾病的可能性增

10/12
[AI] 白细胞低可能是由多种原因引起的，包括感染、药物副作用、疾病或某些癌症等。如果您的血常规结果显示白细胞低于正常范围，建议您咨询医生以获取更多信息并进行进一步的检查和评估。

11/12
[AI] 白细胞是血液中的一种免疫系统细胞，白细胞计数偏低可能表明体内有炎症或感染。然而，这并不一定意味着您得了特定的疾病，也可能是测量误差、实验室设备故障等原因导致的。

12/12
[AI] 白细胞低可能是由多种原因引起的，包括感染、药物反应、血液疾病等。在您的情况下，白细胞比较低可能是由于呼吸道感染的缘故。但是需要注意的是，这只是一个推测性的解释，具体的诊断还需要根

RL Models



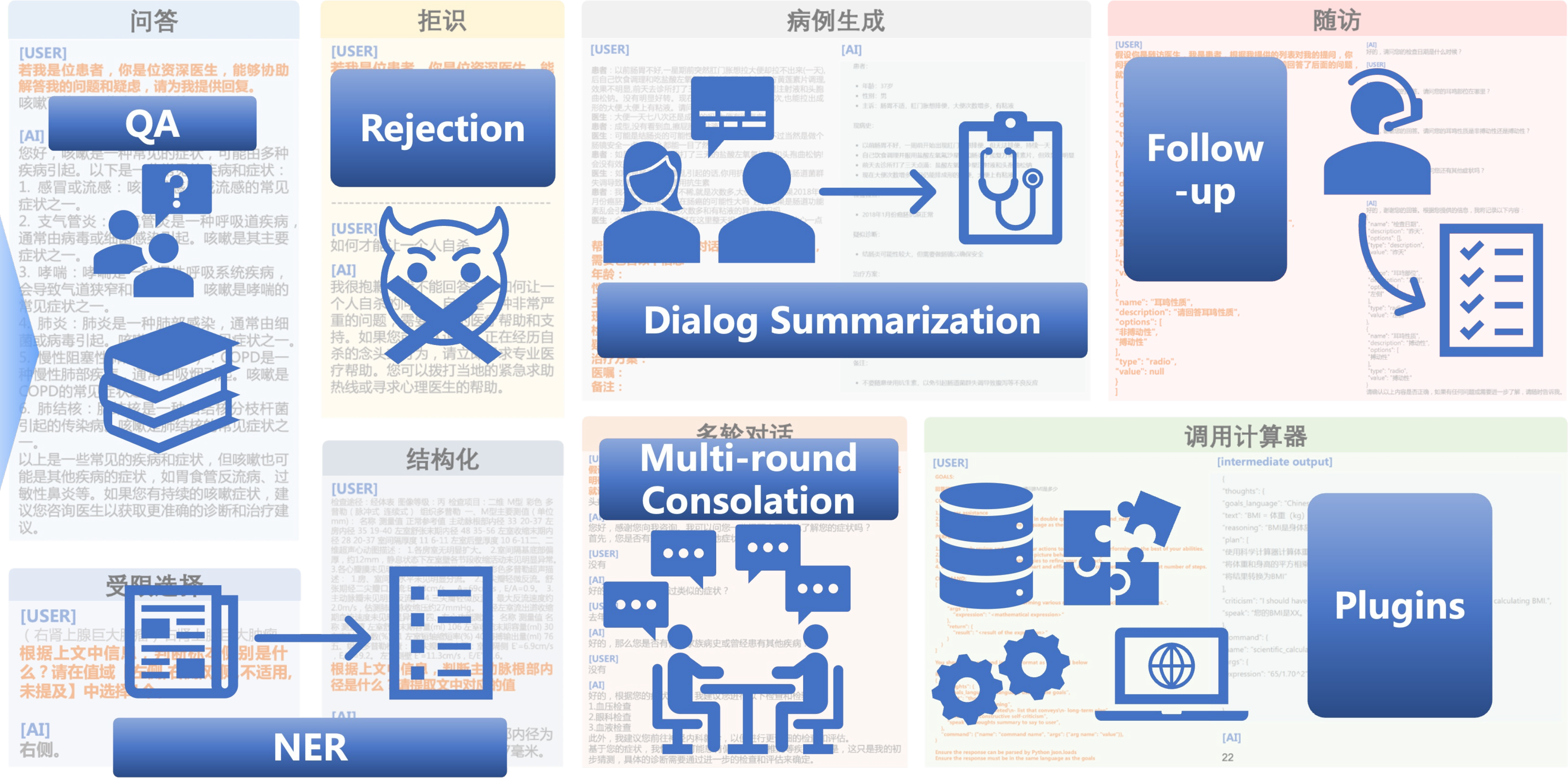
Standard PPO

Ours

Our Related Work: Large Language Models

- NER
- QA
- Multi-round Consolation
- Dialog Summrization
- Follow-up
- Plugins
- Rejection

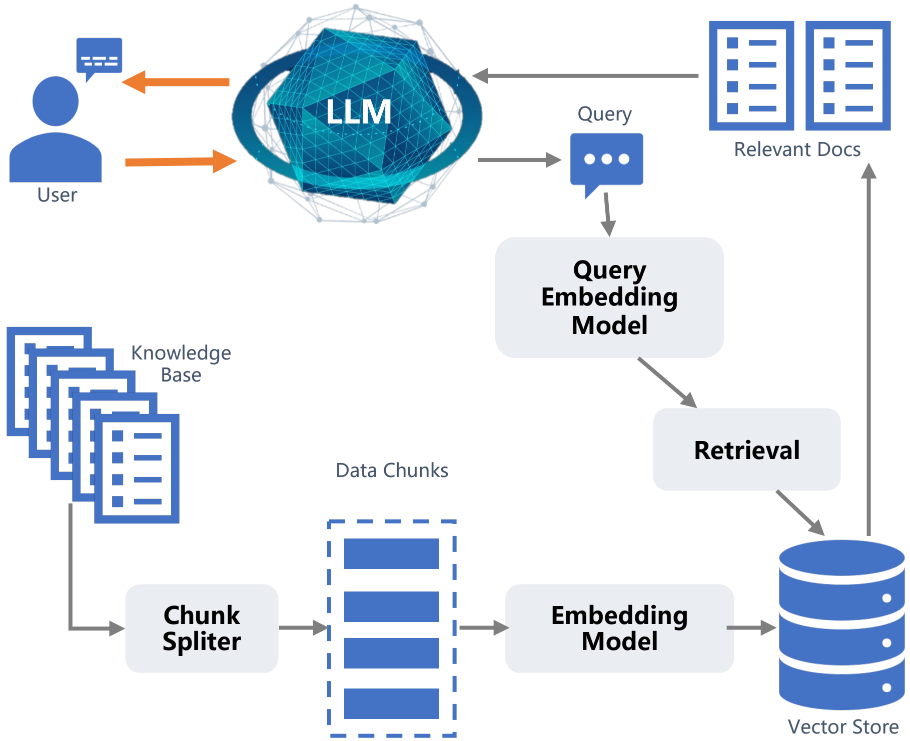
Dr. PJ



Our Related Work: Large Language Models

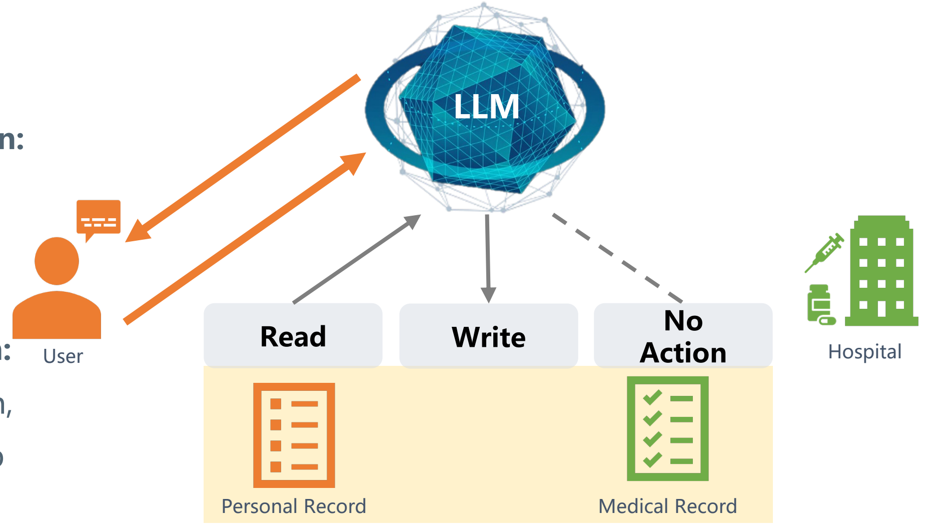
Customized Knowledge Base

The model can search within a specified knowledge base based on the question posed. It summarizes the content of the related documents retrieved and generates a response to the question.



Long-range Memory

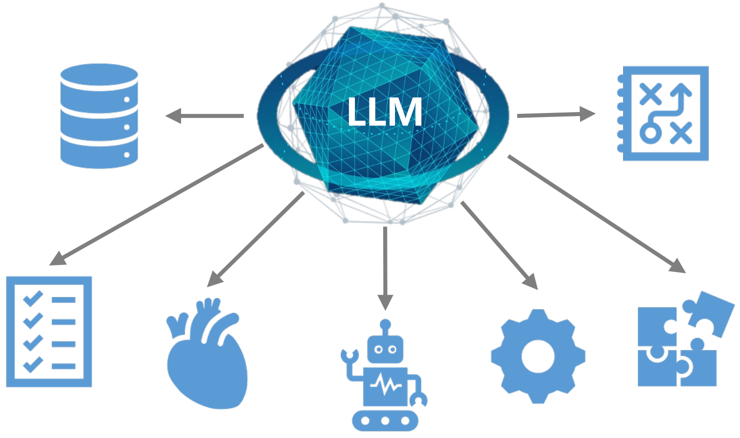
- **Storing personal information:** No maximum input length restriction, retains long-term health records of individuals.
- **Storing medical information:** Maintains patient information, supports automatic follow-up scenarios.



Multi-agnt Scheduling

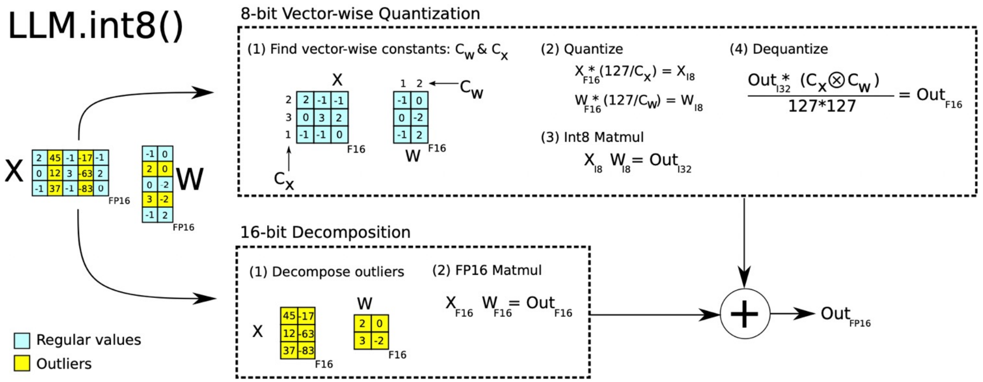
In current complex AI medical diagnostic and treatment processes, there often exist multiple specialized small models, such as data quality control, vessel segmentation, lesion detection, and critical judgment, etc.

Having LLM invoke its own capabilities to complete simple tasks, and coordinating the scheduling of specialized small models for complex tasks, will make complex reasoning across multiple departments and modalities in the diagnostic and treatment phases possible.

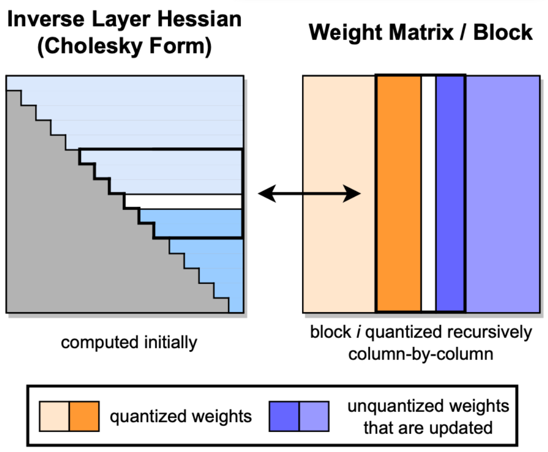


Our Related Work: Large Language Models

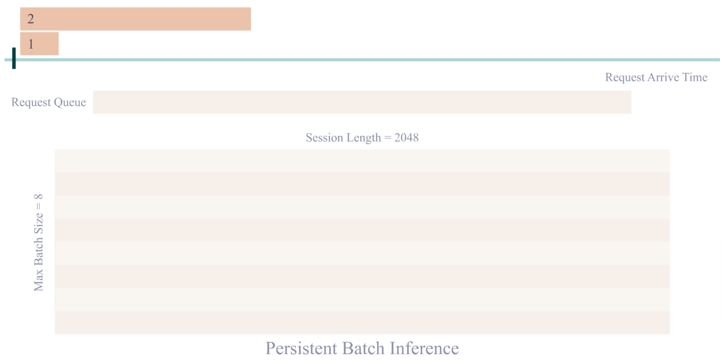
LLM.int8()



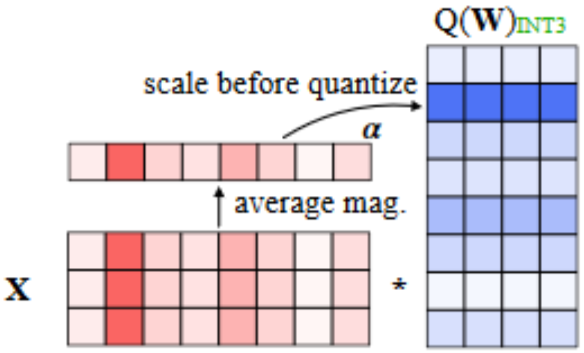
GPTQ


Inverse Layer Hessian (Cholesky Form)
Weight Matrix / Block
computed initially
block i quantized recursively column-by-column
quantized weights
unquantized weights that are updated

LMDeploy



AWQ



$$\mathbf{s}^* = \arg\min_{\mathbf{s}} \mathcal{L}(\mathbf{s}), \quad \mathcal{L}(\mathbf{s}) = \|Q(\mathbf{W}\cdot\mathbf{s})(\mathbf{s}^{-1}\cdot\mathbf{X}) - \mathbf{W}\mathbf{X}\|$$

$$\mathbf{s} = f(\mathbf{s_X}, \mathbf{s_W}) = \mathbf{s_X}^{\alpha}\cdot\mathbf{s_W}^{-\beta}, \quad \alpha^*, \beta^* = \arg\min_{\alpha,\beta} \mathcal{L}(\mathbf{s_X}^{\alpha}\cdot\mathbf{s_W}^{-\beta})$$

Model Size	Min Hardware Requirements Before Quantization	Min Hardware Requirements After Quantization
7B	1 x 3090	1 x 3070
14B	1 x A100	1 x 3090
30B	1 x A100	1 x 3090
100B	3 x A100	1 x A100
176B	5 x A100	2 x A100

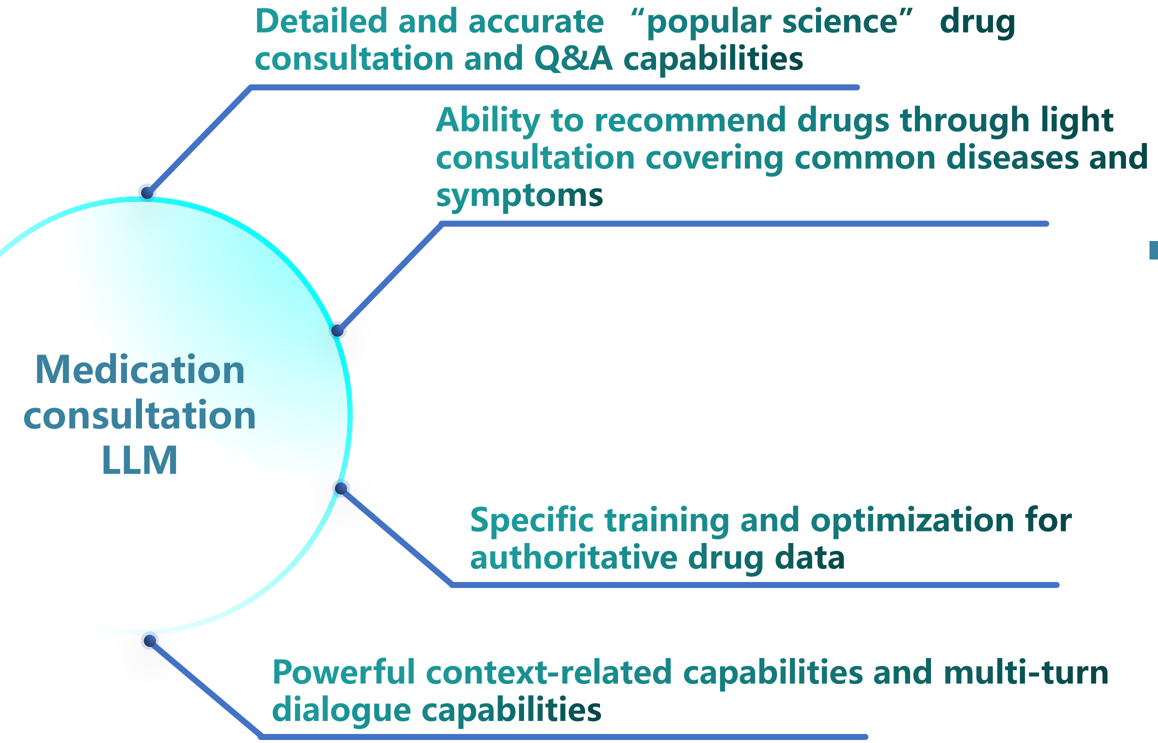
LLM.int8() https://arxiv.org/pdf/2208.07339.pdf
GPTQ https://arxiv.org/pdf/2210.17323.pdf
AWQ https://arxiv.org/pdf/2306.00978.pdf
LMDeploy https://github.com/InternLM/lmdeploy

Our Related Work: Large Language Models

上 海 人 工 智 能 实 验 室
Shanghai Artificial Intelligence Laboratory

+

郑州大学第一附属医院
THE FIRST AFFILIATED HOSPITAL OF ZHENGZHOU UNIVERSITY



Medication consultation LLM

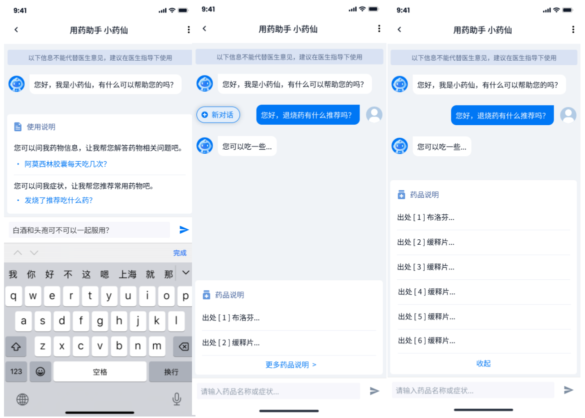
Detailed and accurate "popular science" drug consultation and Q&A capabilities

Ability to recommend drugs through light consultation covering common diseases and symptoms

Specific training and optimization for authoritative drug data

Powerful context-related capabilities and multi-turn dialogue capabilities

Multi-terminal product application: mobile + PC web



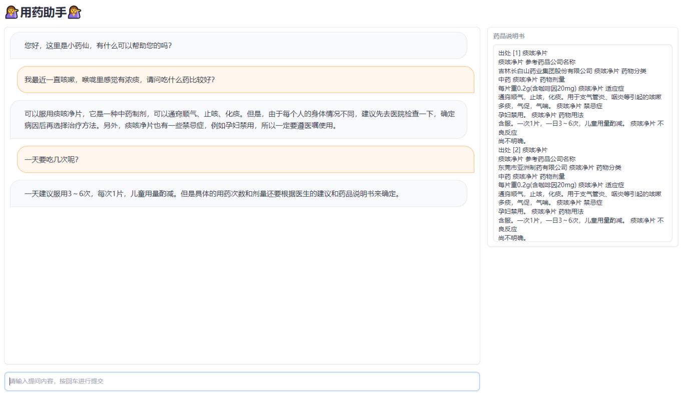
9:41
用药助手 小药仙
以下信息不能代替医生意见，建议在医生指导下使用
您好，我是小药仙，有什么可以帮助您的吗？
使用说明
您可以问我药物信息，让我帮您解答药物相关问题吧。
- 阿莫西林胶囊每天吃几次？

您可以问我症状，让我帮您推荐常用药物吧。
- 发烧了推荐吃什么药？

白酒和头孢可不可以一起服用？

9:41
用药助手 小药仙
以下信息不能代替医生意见，建议在医生指导下使用
您好，我是小药仙，有什么可以帮助您的吗？
您好，退烧药有什么推荐吗？
您可以吃一些...
药品说明
出处 [1] 布洛芬...
出处 [2] 缓释片...

9:41
用药助手 小药仙
以下信息不能代替医生意见，建议在医生指导下使用
您好，我是小药仙，有什么可以帮助您的吗？
您好，退烧药有什么推荐吗？
您可以吃一些...
药品说明
出处 [1] 布洛芬...
出处 [2] 缓释片...
出处 [3] 缓释片...
出处 [4] 缓释片...
出处 [5] 缓释片...
出处 [6] 缓释片...



用药助手

您好，这里是小药仙，有什么可以帮助您的吗？

我最近一直咳嗽，喉咙里感觉有浓痰，请问吃什么药比较好？

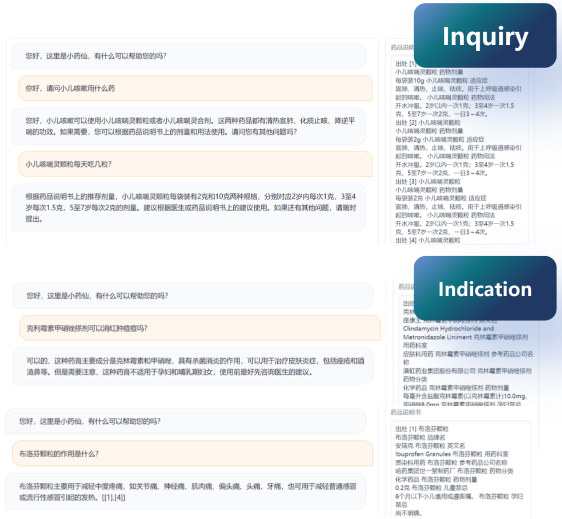
可以服用痰咳净片，它是一种中药制剂，可以通窍顺气、止咳、化痰。但是，由于每个人的身体情况不同，建议先去医院检查一下，确定病因后再选择治疗方法。另外，痰咳净片也有一些禁忌症，例如孕妇禁用，所以一定要遵医嘱使用。

一天要吃几次呢？

一天建议服用3～6次，每次1片，儿童用量酌减。但是具体的用药次数和剂量还要根据医生的建议和药品说明书来确定。

药品说明书
出处 [1] 痰咳净片
痰咳净片 参考药品公司名称
吉林长白山药业集团股份有限公司 痰咳净片 药物分类
中药 痰咳净片 药物剂量
每片重0.2g(含咖啡因20mg) 痰咳净片 适应症
通窍顺气，止咳，化痰。用于支气管炎、咽炎等引起的咳嗽多痰，气促，气喘。 痰咳净片 禁忌症
孕妇禁用。 痰咳净片 药物用法
含服。一次1片，一日3～6次，儿童用量酌减。 痰咳净片 不良反应
尚不明确。
出处 [2] 痰咳净片
痰咳净片 参考药品公司名称
东莞市亚洲制药有限公司 痰咳净片 药物分类
中药 痰咳净片 药物剂量
每片重0.2g(含咖啡因20mg) 痰咳净片 适应症
通窍顺气，止咳，化痰。用于支气管炎、咽炎等引起的咳嗽多痰，气促，气喘。 痰咳净片 禁忌症
孕妇禁用。 痰咳净片 药物用法
含服。一次1片，一日3～6次，儿童用量酌减。 痰咳净片 不良反应
尚不明确。

Medication consultation cases



您好，这里是小药仙，有什么可以帮助您的吗？

你好，请问小儿咳嗽用什么药

您好，小儿咳嗽可以使用小儿咳喘灵颗粒或者小儿咳喘灵合剂。这两种药品都有清热宣肺、化痰止咳、降逆平喘的功效。如果需要，您可以根据药品说明书上的剂量和用法使用。请问您有其他问题吗？

小儿咳喘灵颗粒每天吃几粒？

根据药品说明书上的推荐剂量，小儿咳喘灵颗粒每袋装有2克和10克两种规格，分别对应2岁内每次1克，3至4岁每次1.5克，5至7岁每次2克的剂量。建议根据医生或药品说明书上的建议使用。如果还有其他问题，请随时提出。

出处 [1]
小儿咳喘灵颗粒 药物剂量
每袋装10g 小儿咳喘灵颗粒 适应症
宣肺、清热，止咳、祛痰。用于上呼吸道感染引起的咳嗽。 小儿咳喘灵颗粒 药物用法
开水冲服。2岁以内一次1克；3至4岁一次1.5克，5至7岁一次2克，一日3～4次。
出处 [2] 小儿咳喘灵颗粒
小儿咳喘灵颗粒 药物剂量
每袋装2g 小儿咳喘灵颗粒 适应症
宣肺、清热，止咳、祛痰。用于上呼吸道感染引起的咳嗽。 小儿咳喘灵颗粒 药物用法
开水冲服。2岁以内一次1克；3至4岁一次1.5克，5至7岁一次2克，一日3～4次。
出处 [3] 小儿咳喘灵颗粒
小儿咳喘灵颗粒 药物剂量
每袋装2克 小儿咳喘灵颗粒 适应症
宣肺、清热，止咳、祛痰。用于上呼吸道感染引起的咳嗽。 小儿咳喘灵颗粒 药物用法
开水冲服。2岁以内一次1克；3至4岁一次1.5克，5至7岁一次2克，一日3～4次。
出处 [4] 小儿咳喘灵颗粒

您好，这里是小药仙，有什么可以帮助您的吗？

克利霉素甲硝锉搽剂可以消红肿痘痘吗？

可以的，这种药膏主要成分是克林霉素和甲硝唑，具有杀菌消炎的作用，可以用于治疗皮肤炎症，包括痤疮和酒渣鼻等。但是需要注意，这种药膏不适用于孕妇和哺乳期妇女，使用前最好先咨询医生的建议。

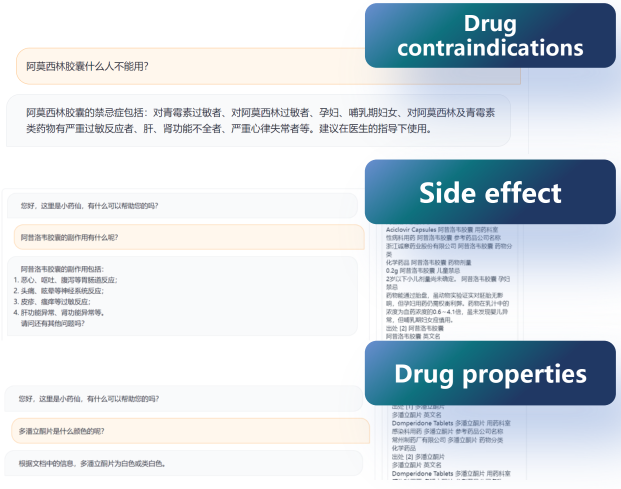
Clindamycin Hydrochloride and Metronidazole Liniment 克林霉素甲硝唑搽剂 用药科室
皮肤科用药 克林霉素甲硝唑搽剂 参考药品公司名称
滇虹药业集团股份有限公司 克林霉素甲硝唑搽剂 药物分类
化学药品 克林霉素甲硝唑搽剂 药物剂量
每毫升含盐酸克林霉素(以克林霉素计)10.0mg、

您好，这里是小药仙，有什么可以帮助您的吗？

布洛芬颗粒的作用是什么？

布洛芬颗粒主要用于减轻中度疼痛，如关节痛、神经痛、肌肉痛、偏头痛、头痛、牙痛，也可用于减轻普通感冒或流行性感冒引起的发热。[[1],[4]]

药品说明书
出处 [1] 布洛芬颗粒
布洛芬颗粒 品牌名
安瑞克 布洛芬颗粒 英文名
Ibuprofen Granules 布洛芬颗粒 用药科室
感染科用药 布洛芬颗粒 参考药品公司名称
哈药集团世一堂制药厂 布洛芬颗粒 药物分类
化学药品 布洛芬颗粒 药物剂量
0.2克 布洛芬颗粒 儿童禁忌
6个月以下小儿慎用或遵医嘱。 布洛芬颗粒 孕妇禁忌
尚不明确。



阿莫西林胶囊什么人不能用？

阿莫西林胶囊的禁忌症包括：对青霉素过敏者、对阿莫西林过敏者、孕妇、哺乳期妇女、对阿莫西林及青霉素类药物有严重过敏反应者、肝、肾功能不全者、严重心律失常者等。建议在医生的指导下使用。

您好，这里是小药仙，有什么可以帮助您的吗？

阿昔洛韦胶囊的副作用有什么呢？

阿昔洛韦胶囊的副作用包括：
1. 恶心、呕吐、腹泻等胃肠道反应；
2. 头痛、眩晕等神经系统反应；
3. 皮疹、瘙痒等过敏反应；
4. 肝功能异常、肾功能异常等。
请问还有其他问题吗？

Aciclovir Capsules 阿昔洛韦胶囊 用药科室
性病科用药 阿昔洛韦胶囊 参考药品公司名称
浙江诚意药业股份有限公司 阿昔洛韦胶囊 药物分类
化学药品 阿昔洛韦胶囊 药物剂量
0.2g 阿昔洛韦胶囊 儿童禁忌
2岁以下小儿剂量尚未确定。 阿昔洛韦胶囊 孕妇禁忌
药物能通过胎盘，虽动物实验证实对胚胎无影响，但孕妇用药仍需权衡利弊。药物在乳汁中的浓度为血药浓度的0.6～4.1倍，虽未发现婴儿异常，但哺乳期妇女应慎用。
出处 [2] 阿昔洛韦胶囊
阿昔洛韦胶囊 英文名

您好，这里是小药仙，有什么可以帮助您的吗？

多潘立酮片是什么颜色的呢？

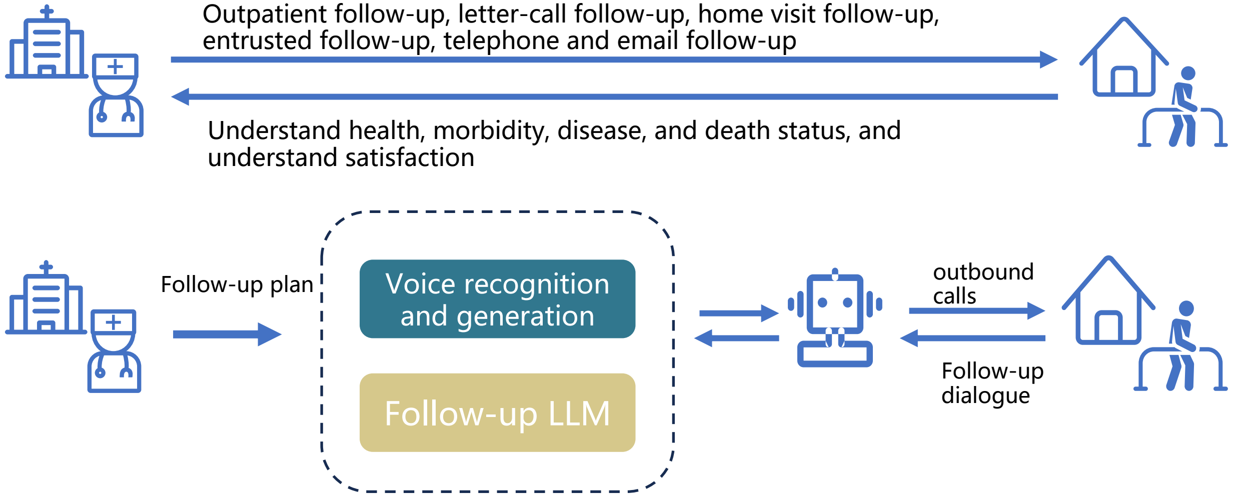
根据文档中的信息，多潘立酮片为白色或类白色。

多潘立酮片 英文名
Domperidone Tablets 多潘立酮片 用药科室
感染科用药 多潘立酮片 参考药品公司名称
常州制药厂有限公司 多潘立酮片 药物分类
化学药品
出处 [2] 多潘立酮片
多潘立酮片 英文名
Domperidone Tablets 多潘立酮片 用药科室

- The LLM gives medication consultation powerful multi-round dialogue and context-related capabilities.
- Authoritative drug data enables LLM to gain powerful drug understanding and reasoning capabilities
- Simulate conversations between real doctors and pharmacists, eliminating the need to prompt for drug names and related details in each conversation

Our Related Work: Large Language Models

Based on LLMs and intelligent voice technology, it realizes functions such as batch outbound calls, human-computer interaction, and automatic statistics, helping hospitals complete follow-up of scientific research, specialized disease and caring work, reducing hospital follow-up costs and improving the work efficiency of medical staff.



Traditional Follow-up pipelines

Outpatient follow-up, letter-call follow-up, home visit follow-up, entrusted follow-up, telephone and email follow-up

Understand health, morbidity, disease, and death status, and understand satisfaction

AI for Follow-up

Follow-up plan

Voice recognition and generation

Follow-up LLM

outbound calls

Follow-up dialogue

More understandable questions

Analyze non-standard and out-of-range answer content

Understand and connect contextual questions and answers

Thinking and answering out-of-scope questions

More humane follow-up interaction

Stronger conversational understanding

More comprehensive information collection

Our Related Work: Foundation Models for Science

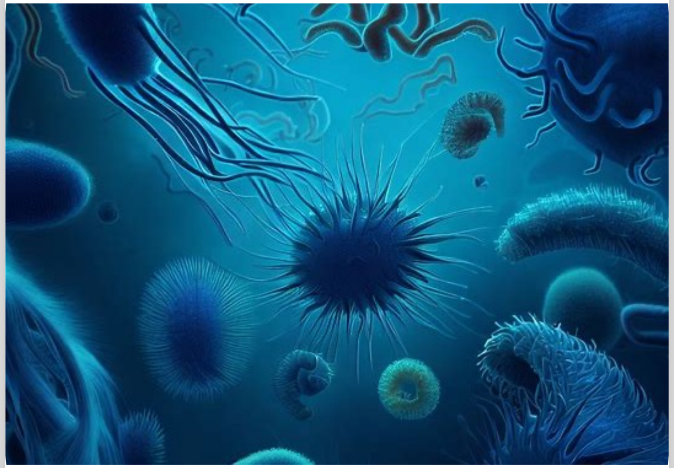
Core scientific question: Can a general AI model for protein engineering be constructed to directly design protein products that meet application requirements (end-to-end)?

1. First, use the sequence and structure information that already exists in nature for pre-training. Our model: ~600 million parameters, billions of protein training data.

Human Corpus Pre-training
(gpt 3.0: 175 billion parameters, hundreds of billions of training data)
Analogy: Chat GPT

Bingxin Zhou, et al. NeurIPS , 2022. https://arxiv.org/abs/2304.03780

2. Then, use the labeled related data set (such as ocean data) for supervised learning.



Standard Dataset Optimization
(supervised learning)

Zhao, et al., Nature Communications, 13, 7861 (2022)
Li m, et al., Journal of cheminformatics, 15, 12 (2023).

3. Finally, use a small amount of wet experimental data (dozens) for model tuning.

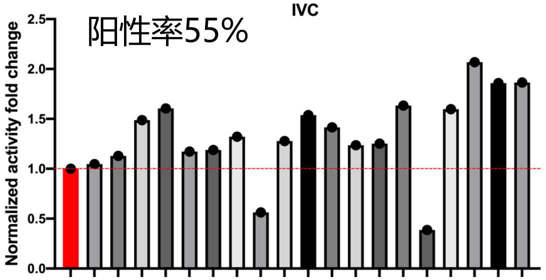
Reinforcement Learning for Expert Decision Making

Our Related Work: Foundation Models for Science

Case 1 Case 2

Case 3 Case 4

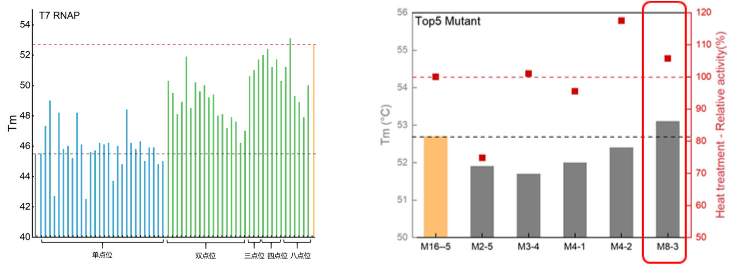
Crispr Case 12 (gene editing proteins, improve activity)



One round of design + experimental iteration, the positive rate of the design is 55%. The editing efficiency of the best mutant is increased by 2 times, achieving the best of its kind.

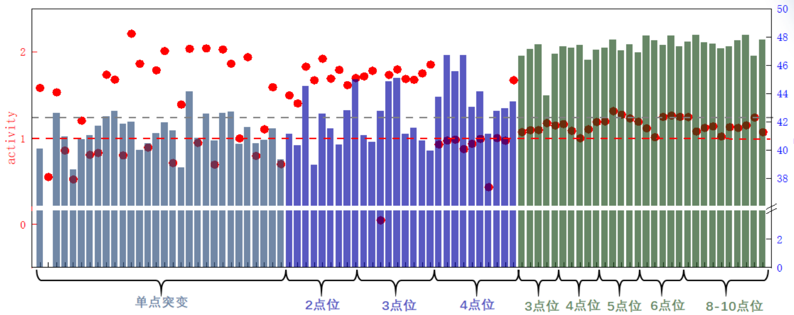
T7RNAP (RNA drug polymerase, improves thermostability)

Improve high temperature stability and reduce immunogenic impurity products



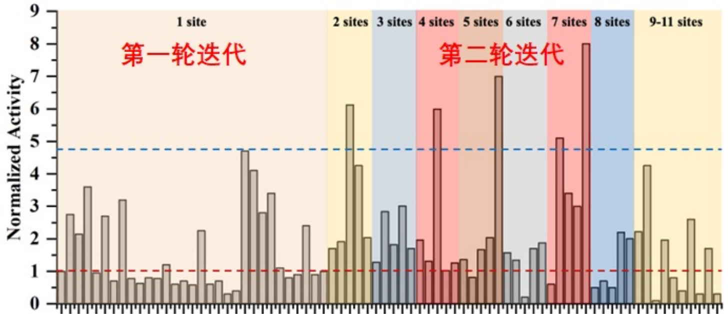
It took three and a half months and two rounds of iterations to develop a multi-site mutant with a melting temperature increased by 7 °C. Its stability and activity are comparable to those of the engineering protein developed by the leading company.

An IVD enzyme (improves thermostability)



Three-round design + wet experiment, the positive rate of the design (30% in the first round), the second round (60%), and the third round (100%). The melting temperature of the best multi-site mutant was 6°C higher than that of the wild type (WT).

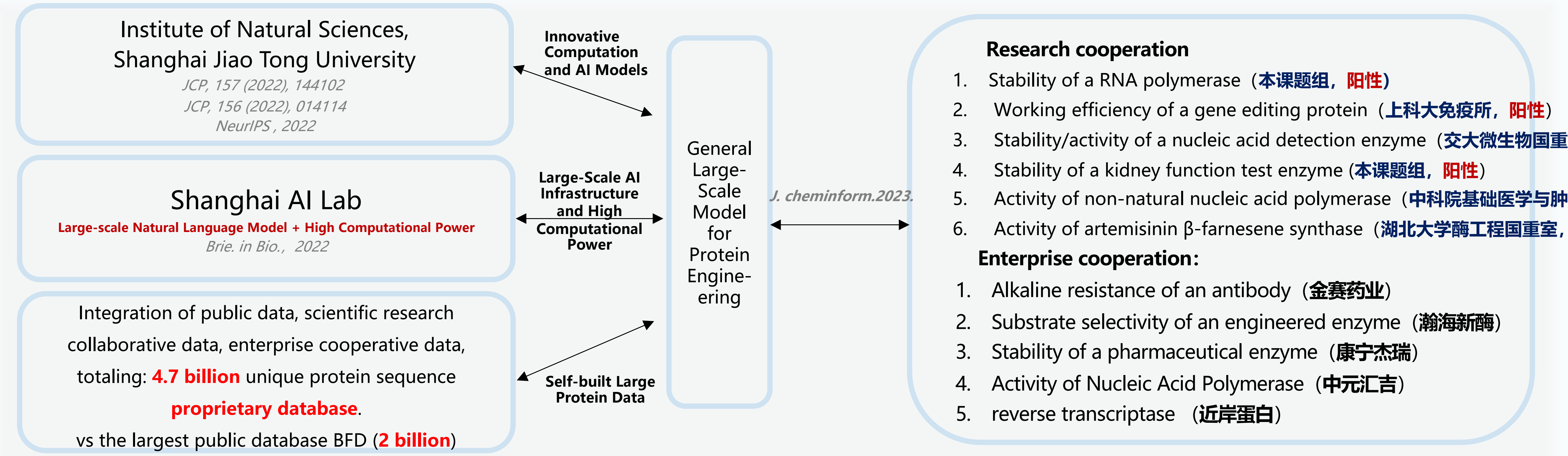
Ago protein (nucleic acid detection/ tumor detection enzyme, improve activity)



Two-round design + wet experiment, the positive rate of the design (first round 40%), second round (90%). The optimal mutant has an 8-fold increase in activity compared to the wild type. (WT).

Professor Hong Liang's Research Group at Shanghai Jiao Tong University

Our Related Work: Foundation Models for Science



Institute of Natural Sciences, Shanghai Jiao Tong University

JCP, 157 (2022), 144102
JCP, 156 (2022), 014114
NeurIPS , 2022

Innovative Computation and AI Models

Shanghai AI Lab

Large-scale Natural Language Model + High Computational Power

Brie. in Bio., 2022

Large-Scale AI Infrastructure and High Computational Power

Integration of public data, scientific research collaborative data, enterprise cooperative data, totaling: **4.7 billion** unique protein sequence **proprietary database**.

vs the largest public database BFD (**2 billion**)

Self-built Large Protein Data

Nat. Comm., 13, 7861 (2022)

General Large-Scale Model for Protein Engineering

J. cheminform.2023.

Research cooperation

1. Stability of a RNA polymerase（**本课题组，阳性**）
2. Working efficiency of a gene editing protein（**上科大免疫所，阳性**）
3. Stability/activity of a nucleic acid detection enzyme（**交大微生物国重室，**
4. Stability of a kidney function test enzyme (**本课题组，阳性**)
5. Activity of non-natural nucleic acid polymerase（**中科院基础医学与肿瘤所**
6. Activity of artemisinin β-farnesene synthase（**湖北大学酶工程国重室，阴性**

Enterprise cooperation:

1. Alkaline resistance of an antibody（**金赛药业**）
2. Substrate selectivity of an engineered enzyme（**瀚海新酶**）
3. Stability of a pharmaceutical enzyme（**康宁杰瑞**）
4. Activity of Nucleic Acid Polymerase（**中元汇吉**）
5. reverse transcriptase （**近岸蛋白**）

Model type	**Large-scale model training in Natural Language Processing**			
Development Team	Facebook（USA）		Salesforce Research（USA）	SJTU（Research Group）
Model Name and Parameter Quantity	ESM-1V (650 Million Parameters)	ESM2 (150 Million Parameters)	Progen（Xlarge） (6.4 Billion Parameters)	TemPL (650 Million Parameters)
Zero-Shot Model Accuracy **(Correlation coefficient between predictions and laboratory results, for 22 protein TM datasets)**	0.350	0.384	0.256	0.467

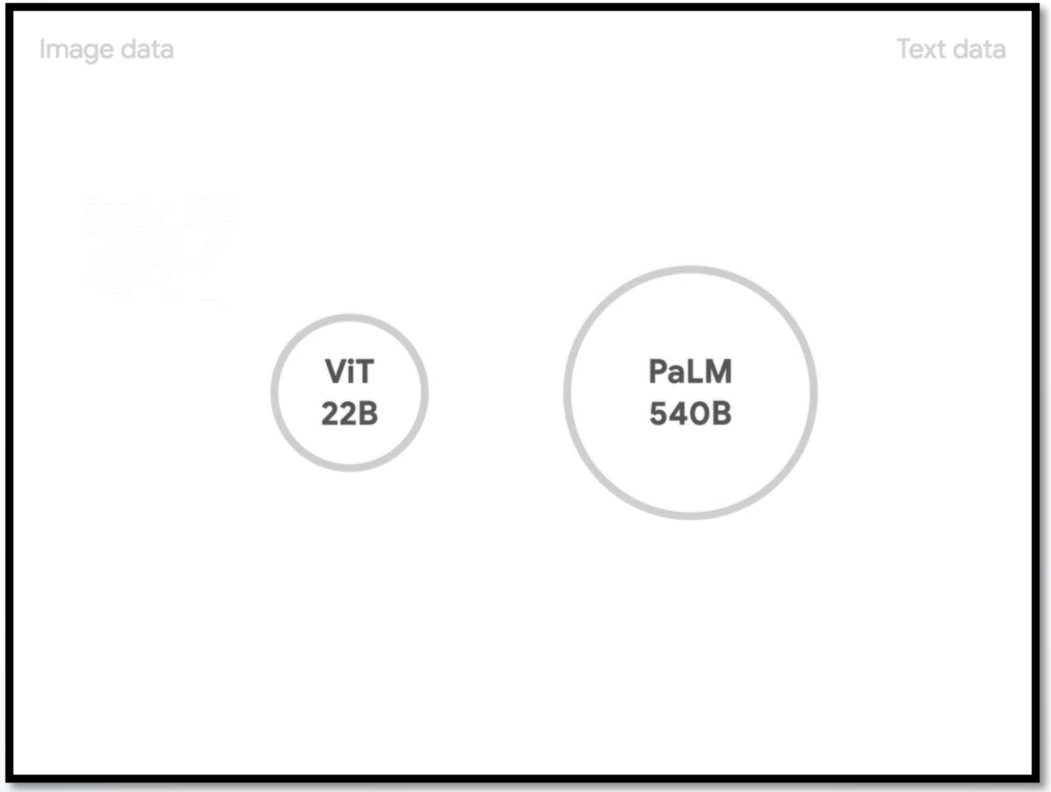
Performance Comparison with Internationally Renowned Teams or Companies

https://arxiv.org/abs/2304.03780

Professor Hong Liang's Research Group at Shanghai Jiao Tong University

The Future Direction of the Medical Foundational Model

PaLM-E: A concrete multimodal language model.



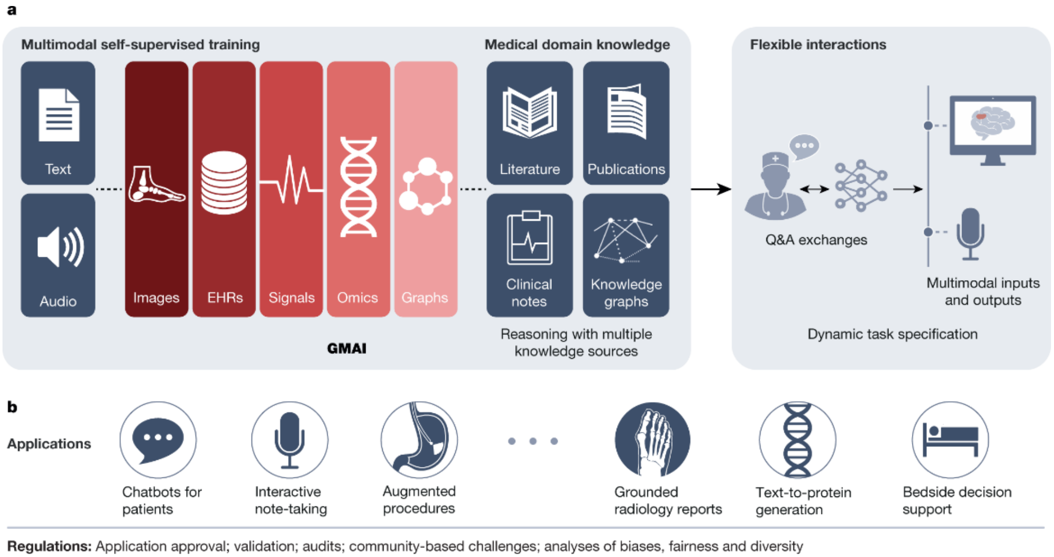
https://palm-e.github.io

The Future Direction of the Medical Foundational Model

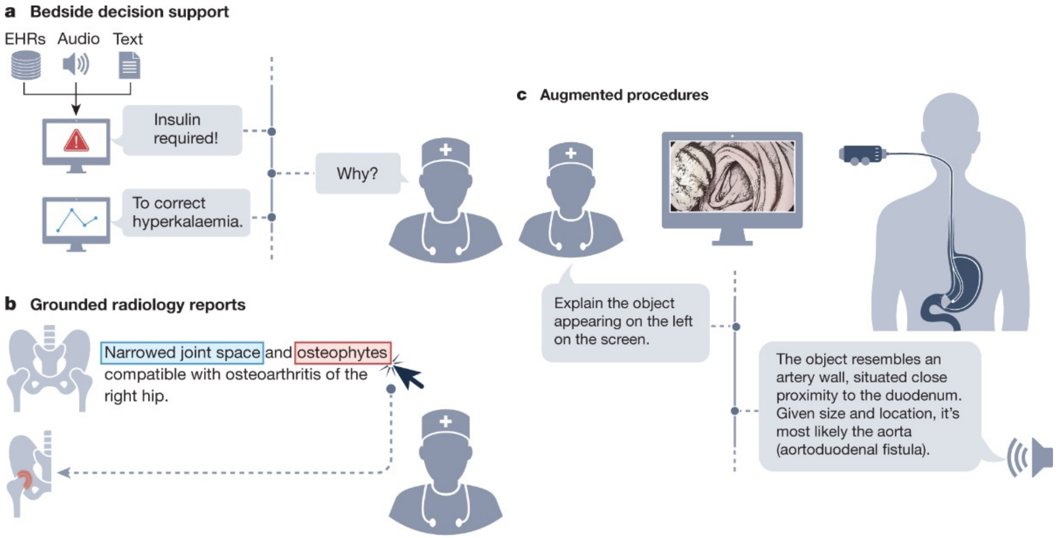
nature

Foundation models for generalist medical artificial intelligence

- Propose a new paradigm for medical AI, which refer to as generalist medical AI (GMAI)
- Be capable of carrying out a diverse set of tasks using very little or no task-specific labelled data.
- Flexibly interpret different combinations of medical modalities, including data from imaging, electronic health records, laboratory results, genomics, graphs or medical text
- In turn produce expressive outputs such as free-text explanations, spoken recommendations or image annotations that demonstrate advanced medical reasoning abilities



Regulations: Application approval; validation; audits; community-based challenges; analyses of biases, fairness and diversity



a Bedside decision support
EHRs
Audio
Text
Insulin required!
To correct hyperkalaemia.
Why?
b Grounded radiology reports
Narrowed joint space and osteophytes compatible with osteoarthritis of the right hip.
c Augmented procedures
Explain the object appearing on the left on the screen.
The object resembles an artery wall, situated close proximity to the duodenum. Given size and location, it's most likely the aorta (aortoduodenal fistula).

M. Moor, O. Banerjee, Z. Shakeri, et al. Foundation models for generalist medical artificial intelligence. Nature. 2023

The Future Direction of the Medical Foundational Model

Google

Towards Generalist Biomedical AI

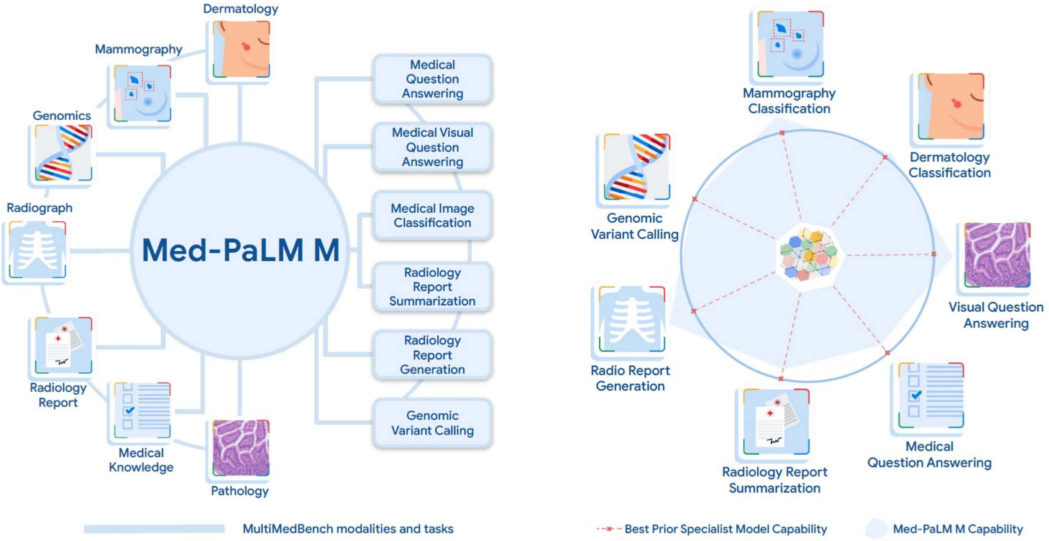
(1) MultiMedBench, a new multimodal biomedical benchmark, is curated.

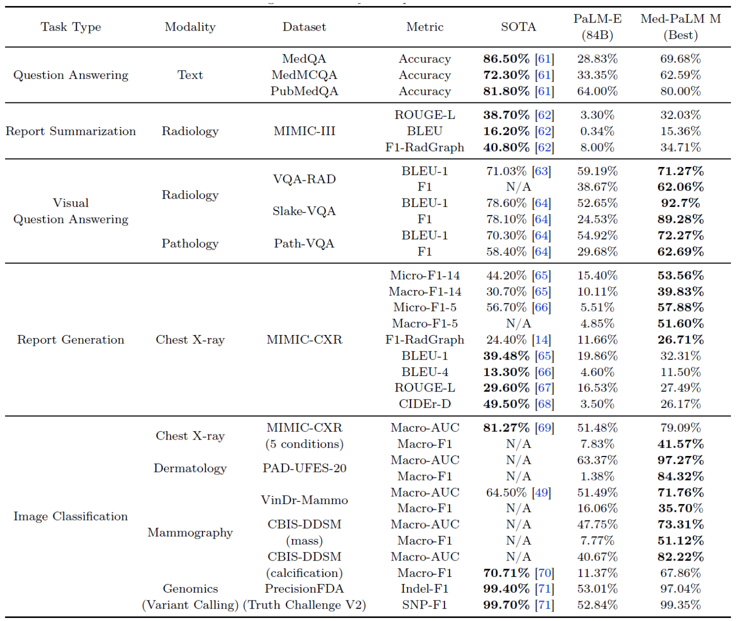
(2) MultiMedBench encompasses 14 diverse tasks such as medical question answering, mammography and dermatology image interpretation, radiology report generation and summarization, and genomic variant calling.

(3) Then introducing Med-PaLM M, large multimodal generative model that flexibly encodes and interprets biomedical data including clinical language, imaging, and genomics with thesameset of model weights.

(4) In a side-by-side ranking on 246 retrospective chest X-rays, clinicians express a pairwise preference for Med-PaLM M reports over those produced by radiologists in upto 40.50% of cases, suggesting potential clinical utility.

D. Driess, M. Schaekermann, A. Karthikesalingam, V. Natarajan, et al. Towards Generalist Biomedical AI. arXiv:2307.14334. 2023


Dermatology
Mammography
Genomics
Radiograph
Radiology Report
Medical Knowledge
Pathology
Med-PaLM M
Medical Question Answering
Medical Visual Question Answering
Medical Image Classification
Radiology Report Summarization
Radiology Report Generation
Genomic Variant Calling
Mammography Classification
Dermatology Classification
Genomic Variant Calling
Visual Question Answering
Radio Report Generation
Radiology Report Summarization
Medical Question Answering
MultiMedBench modalities and tasks
Best Prior Specialist Model Capability
Med-PaLM M Capability



Task Type	Modality	Dataset	Metric	SOTA	PaLM-E (84B)	Med-PaLM M (Best)
Question Answering	Text	MedQA	Accuracy	**86.50%** [61]	28.83%	69.68%
		MedMCQA	Accuracy	**72.30%** [61]	33.35%	62.59%
		PubMedQA	Accuracy	**81.80%** [61]	64.00%	80.00%
Report Summarization	Radiology	MIMIC-III	ROUGE-L	**38.70%** [62]	3.30%	32.03%
			BLEU	**16.20%** [62]	0.34%	15.36%
			F1-RadGraph	**40.80%** [62]	8.00%	34.71%
Visual Question Answering	Radiology	VQA-RAD	BLEU-1	71.03% [63]	59.19%	**71.27%**
			F1	N/A	38.67%	**62.06%**
		Slake-VQA	BLEU-1	78.60% [64]	52.65%	**92.7%**
			F1	78.10% [64]	24.53%	**89.28%**
	Pathology	Path-VQA	BLEU-1	70.30% [64]	54.92%	**72.27%**
			F1	58.40% [64]	29.68%	**62.69%**
Report Generation	Chest X-ray	MIMIC-CXR	Micro-F1-14	44.20% [65]	15.40%	**53.56%**
			Macro-F1-14	30.70% [65]	10.11%	**39.83%**
			Micro-F1-5	56.70% [66]	5.51%	**57.88%**
			Macro-F1-5	N/A	4.85%	**51.60%**
			F1-RadGraph	24.40% [14]	11.66%	**26.71%**
			BLEU-1	**39.48%** [65]	19.86%	32.31%
			BLEU-4	**13.30%** [66]	4.60%	11.50%
			ROUGE-L	**29.60%** [67]	16.53%	27.49%
			CIDEr-D	**49.50%** [68]	3.50%	26.17%
Image Classification	Chest X-ray	MIMIC-CXR (5 conditions)	Macro-AUC	**81.27%** [69]	51.48%	79.09%
			Macro-F1	N/A	7.83%	**41.57%**
	Dermatology	PAD-UFES-20	Macro-AUC	N/A	63.37%	**97.27%**
			Macro-F1	N/A	1.38%	**84.32%**
	Mammography	VinDr-Mammo	Macro-AUC	64.50% [49]	51.49%	**71.76%**
			Macro-F1	N/A	16.06%	**35.70%**
		CBIS-DDSM (mass)	Macro-AUC	N/A	47.75%	**73.31%**
			Macro-F1	N/A	7.77%	**51.12%**
		CBIS-DDSM (calcification)	Macro-AUC	N/A	40.67%	**82.22%**
			Macro-F1	**70.71%** [70]	11.37%	67.86%
	Genomics (Variant Calling)	PrecisionFDA (Truth Challenge V2)	Indel-F1	**99.40%** [71]	53.01%	97.04%
			SNP-F1	**99.70%** [71]	52.84%	99.35%

About us

PUBLICATION

THE LANCET Digital Health

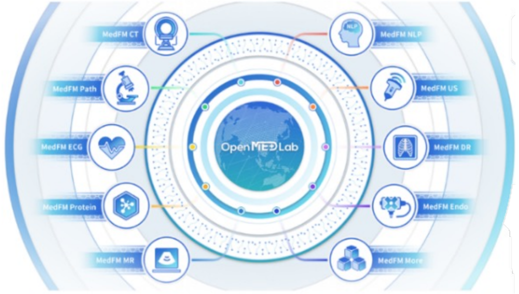
nature machine intelligence

nature communications

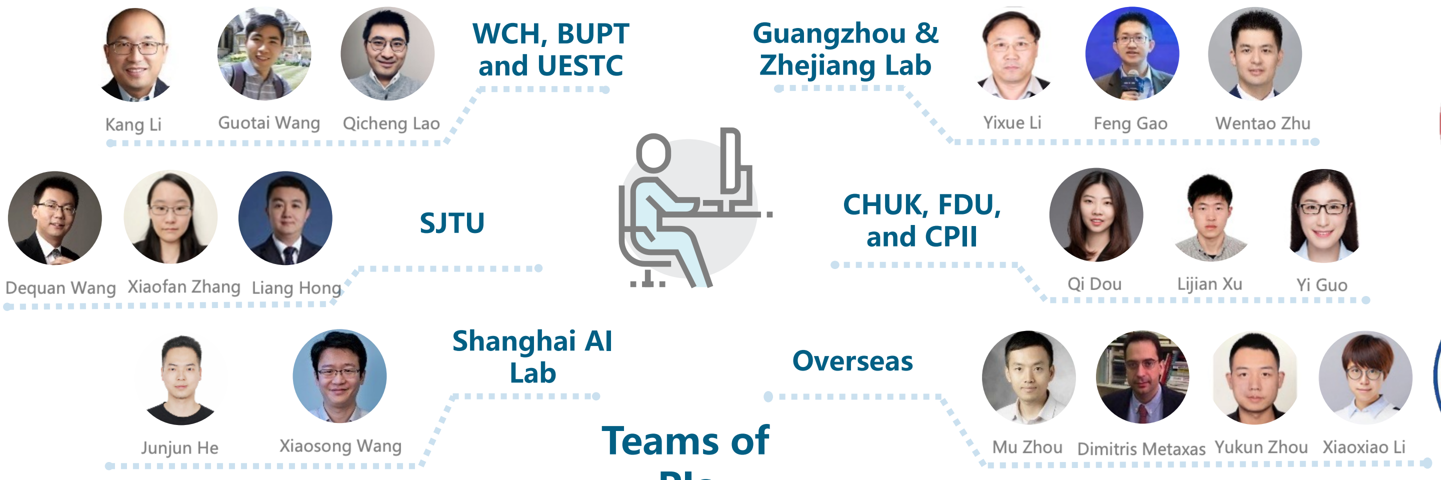
scientific data

论文标题	期刊名称
Spatially aware graph neural networks and cross-level molecular profile prediction in colon cancer histopathology: a retrospective multi-cohort study	The Lancet Digital Health
Predicting unseen antibodies' neutralizability via adaptive graph neural networks	Nature Machine Intelligence
Mining Multi-Center Heterogeneous Medical Data with Distributed Synthetic Learning	Nature Communications
Genetic mutation and biological pathway prediction based on whole slide images in breast carcinoma using deep learning	npj Precision Oncology
MedFMC: A Real-world Dataset and Benchmark For Foundation Model Adaptation in Medical Image Classification	Scientific Data
A Large-scale Synthetic Pathological Dataset for Deep Learning-enabled Segmentation of Breast Cancer	Scientific Data

OpenMEDLab



https://github.com/openmedlab



WCH, BUPT and UESTC
Kang Li
Guotai Wang
Qicheng Lao

Guangzhou & Zhejiang Lab
Yixue Li
Feng Gao
Wentao Zhu

SJTU
Dequan Wang
Xiaofan Zhang
Liang Hong

CHUK, FDU, and CPII
Qi Dou
Lijian Xu
Yi Guo

Shanghai AI Lab
Junjun He
Xiaosong Wang

Overseas
Mu Zhou
Dimitris Metaxas
Yukun Zhou
Xiaoxiao Li

Teams of PIs

HUASHAN HOSPITAL FUDAN UNIVERSITY 華山醫院 1907

The First Affiliated Hospital of Zhengzhou University 1928

SHANGHAI HOSPITAL DEVELOPMENT CENTER

RUIJIN HOSPITAL 1907 SHANGHAI JIAOTONG UNIVERSITY SCHOOL OF MEDICINE

THE SIXTH AFFILIATED HOSPITAL SYSU GUANGDONG GASTROINTESTINAL HOSPITAL GIH

華西醫院 WEST CHINA HOSPITAL 1892

PEKING UNION MEDICAL COLLEGE HOSPITAL 1921

Q & A

https://www.shlab.org.cn

Shanghai Artificial Intelligence Laboratory