中国科学院国家天文台
NATIONAL ASTRONOMICAL OBSERVATORIES, CAS



JW-Flare：基于微调大模型的耀斑预报方法及其智能体开发

邵明福、王慧、佟立越、李瑀旸等

指导教师：林佳本　　国家天文台怀柔观测基地

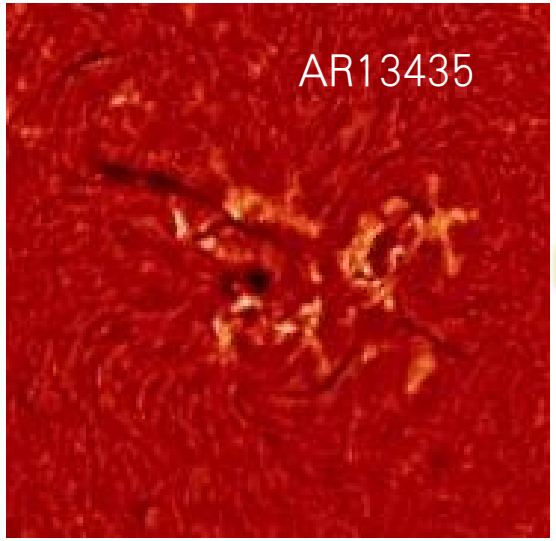
2025-11-28 @海口

NAOC
国家天文台
NATIONAL ASTRONOMICAL OBSERVATORIES

第一部分：研究背景及现状



研究背景及现状:



怀柔太阳观测基地
全日面Hα望远镜

- 耀斑预报
  - 数据驱动方法
    - 统计学方法
    - 机器学习方法
      - SVM
      - KNN
      - MLP
    - 深度学习方法
      - 卷积神经网络(CNN)
      - 长短时记忆网络(LSTM)
      - 生成对抗网络(GAN)
      - Transformer
  - 物理模型

Continue ?

研究背景及现状:

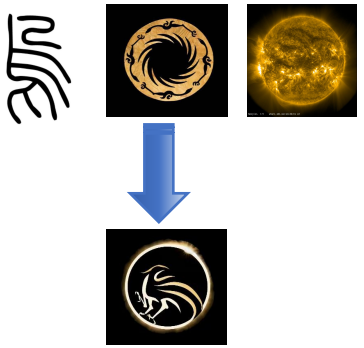
Transformer → ChatGPT(LLM)

模型结构: 多层 Transformer 解码器

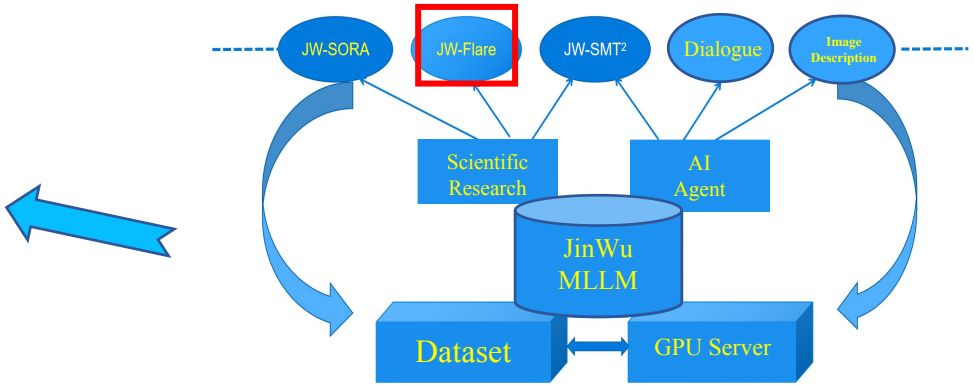
训练数据: 大规模, 多样化互联网数据

训练方法: 自监督预训练、有监督微调等

Jin Wu MLLM



多模态 LLM



JW-Flare

第一个基于LLM的耀斑预报模型

第二部分：数据处理



公开可用的数据集:

Solar active region magnetogram image dataset for studies of space weatherBoucheron
(Laura E. Boucheron)

数据来源:SDO/HMI、201005-201812、LoS Magnetogram

标签: GOES/events.txt

1) Fits: 1,372,004 HMI 磁图, 1655 活动区 (AR11064-AR12731), 600 × 600图像大小, 537GB.

2) Fits: 950,047 HMI 磁图, 1570 活动区 (经纬度范围 ± 60 °), 物理参数, 耀斑标签, 600 × 600 大小, 375GB.

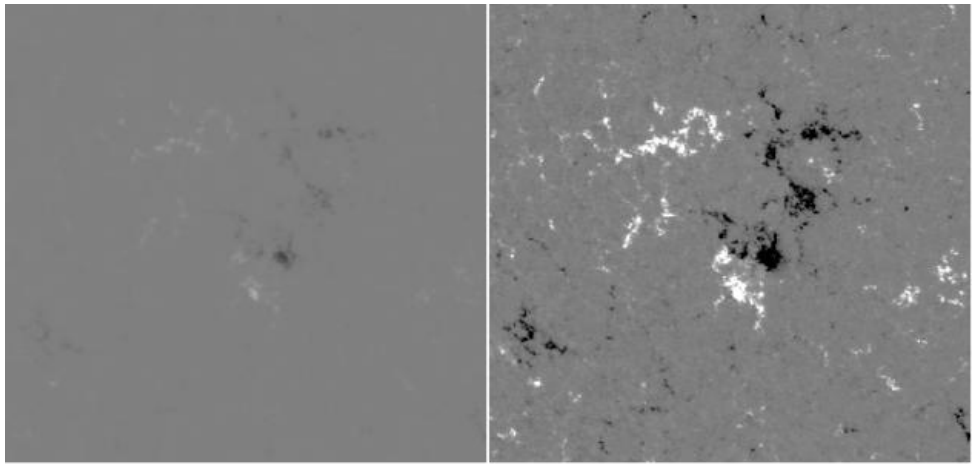
3) PNG: 物理参数, 224 × 224 , 15GB.

耀斑等级:
(e.g., "C1.0", "M1.0", "X1.0")
无耀斑: "0"

预测时间 :
24 hours

数据预处理：

1) 阈值: [-200,200]



(a) (b)

$$I = \left[MinMax(0,400, I + 200) * \frac{255}{400} \right]$$

2) 类别不平衡问题:

Boucheron 数据集:
950,047 磁图, Train:Val:Test=15:2:2

训练集:
(≥C class*2) : (None*1/2) = 1:1
(≥M5.0 class*14) : (None*1/10) = 1:1
(≥X class*30) : (None*1/10) = 1:1

JW-Flare微调数据集:

Dataset for JW-Flare:

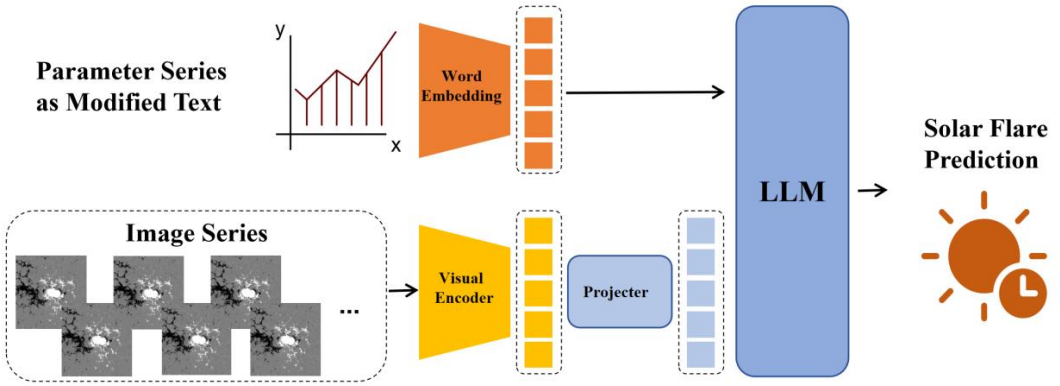
Model	Level	Train	Test	Val
JW-Flare (single)	>C	603328	94757	95933
	>M5.0	153976	94757	95933
	>X	143321	94757	95933
JW-Flare (series)	>C	121128	18582	18822
	>M5.0	28726	18582	18822
	>X	30219	18949	19184

Model	JSONL
JW-Flare (series)	{"**query**": "Given a set of 15 magnetograms of the solar active region, Please follow the steps below to conduct a step-by-step analysis of the sequence of 15 solar images and predict whether a flare is likely to erupt in the next 24 hours: Step 1: Analyze key features of each solar active region magnetogram. Step 2: Compare the sequence of images to identify dynamic changes (e.g., magnetic neutral line length, total unsigned magnetic flux) and their potential impact on flare likelihood. Step 3: Relate these changes to historical data or physical laws to infer if they align with typical flare eruption characteristics. Step 4: Summarize your conclusions and assess the likelihood of a flare eruption.<image>The physical parameters: Magnetic Neutral Line(NL) length: 1 4 6 2 . 0 0 0 pixel, Total unsigned flux: 1 6 8 7 5 1 1 . 0 0 0 Wb<image>...<image>...<image>...<image>...<image>...<image>...<image>...<image>...<image>...<image>...<image>...<image>... . Return the answer as one of the following options: 'A: Flare', 'B: None'. Only return the Option Letters, not the Description.", "**response**": "**A**", "**history**": [], "**images**": ["/***.png", "/***.png", "/***.png", "/***.png", "/***.png", "/***.png", "/***.png", "/***.png", "/***.png", "/***.png", "/***.png", "/***.png", "/***.png", "/***.png", "/***.png"]} ....
JW-Flare (single)	{"**query**": "<image>Given a magnetogram of the solar active region and The corresponding physical parameters: Magnetic Neutral Line(NL) length: 7 3 . 0 0 0 pixel, Total unsigned flux: 9 1 7 5 5 . 0 0 0 Wb. Please follow the steps below to conduct a step-by-step analysis of the solar images and predict whether a flare is likely to erupt in the next 24 hours: Step 1: Analyze key features of solar active region magnetogram(e.g., magnetic neutral line length, total unsigned magnetic flux) and their potential impact on flare likelihood. Step 2: Relate the features to physical laws to infer if it align with typical flare eruption characteristics. Step 3: Summarize your conclusions and assess the likelihood of a flare eruption. Return the answer as one of the following options: 'A: Flare', 'B: None'. Only return the Option Letters, not the Description.", "**response**": "**B**", "**history**": [], "**images**": ["/***.png"]} ....

第三部分：模型结构及训练



JW-Flare模型特点：



1) 第一个基于LLM的耀斑预报模型

2) 关键TSS指标: M5.0耀斑： **0.88** ； X级耀斑： 0.95, 超出现有方法约 **10%**.

3) 真实的泛化能力表现: 怀柔35cm SMFT 望远镜, ASOS/FMG, SDO/HMI

4) 可解释性实验和概率输出: 引导 MLLM和LLM可视化工具

5) 消融实验与Scaling_law:

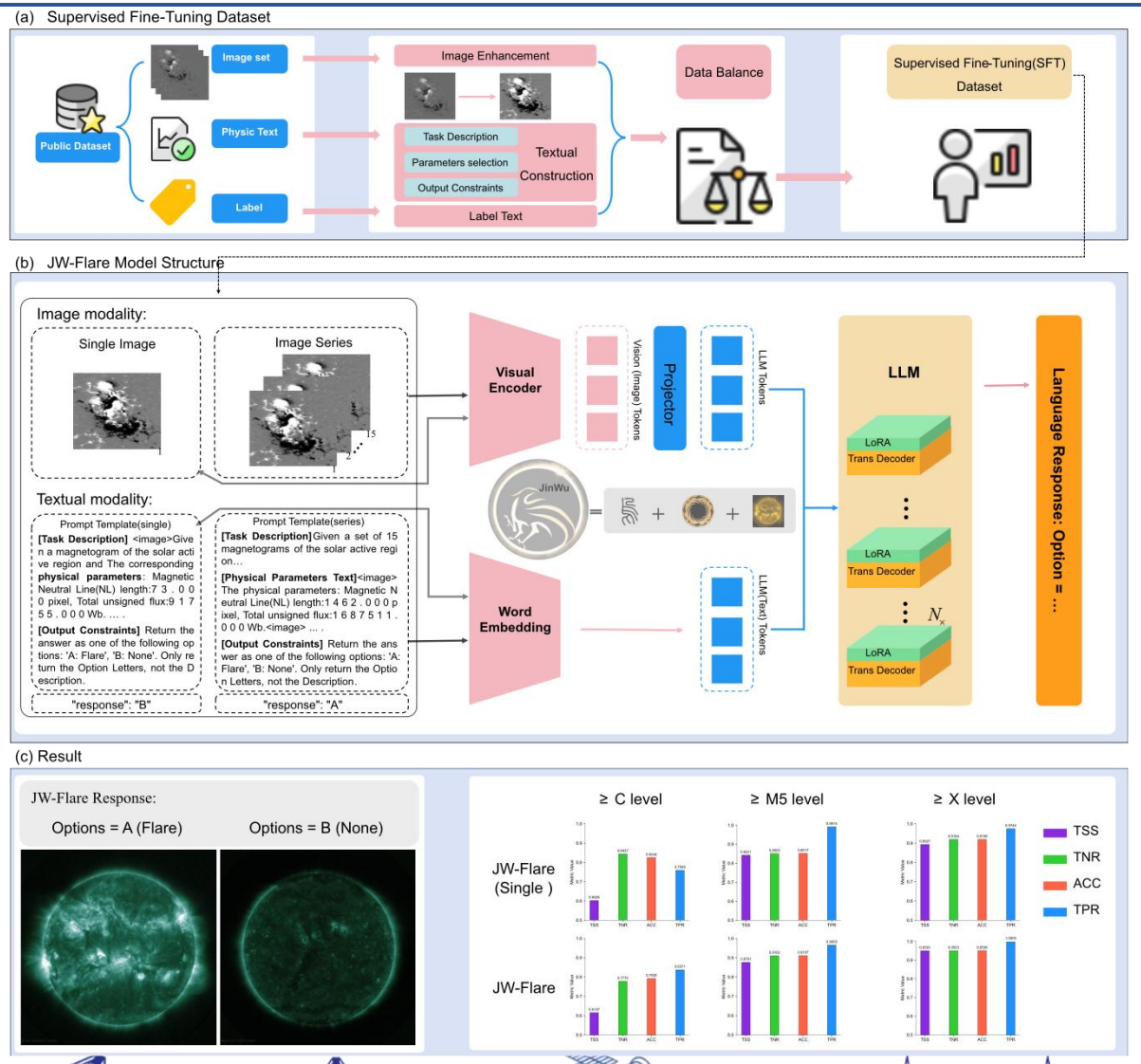
JW-Flare模型结构：

输入/输出：

文本:
自然语言 +物理参数

图像:
单张磁图 or
序列磁图(15 images/3h ,image/12 min)

Option=A(Flare)
Option=B(None)



JW-Flare模型训练:

实验环境: 四块GPUs: NVIDIA RTX 6000 Ada GPU(48GB)

Pytorch and Modelscope Swift(Scalable lightWeight Infrastructure for Fine-Tuning)

模型参数: 训练时间(单卡 GPU): ≥ C ,52 hours; ≥ M5/X ,15 hours

Model	Model_type	Sft_type	Lora_target_modules	Deepspeed	Batch_size	Epochs	Lora_rank	Lora_alpha	Lora_dropout
JW-Flare	Qwen2-vl-7b-instruct	lora	ALL	default-zero2	1	1	32	64	0.1

NAOC
国家天文台
NATIONAL ASTRONOMICAL OBSERVATORIES

第四部分：实验结果



JW-Flare模型：

实验结果：

Model	Level	TP	TN	FP	FN	TPR	TNR	HSS	Accuarcy	Percision	F1	**TSS**
JW-Flare (single)	≥ C	16345	61772	11443	5197	0.7588	0.8437	0.5465	0.8244	0.5882	0.6627	**0.6025**
JW-Flare (series)	≥ C	3767	10950	3132	733	0.8371	0.7776	0.5203	0.7920	0.5460	0.6609	**0.6147**
JW-Flare (single)	≥ M5	826	79879	14045	7	0.9916	0.8505	0.0900	0.8517	0.0555	0.1052	**0.8421**
JW-Flare (series)	≥ M5	170	16753	1653	6	0.9659	0.9102	0.1555	0.9107	0.0933	0.1701	**0.8761**
JW-Flare (single)	≥ X	342	86700	7706	9	0.9744	0.9184	0.0749	0.9186	0.0425	0.0814	**0.8927**
JW-Flare (series)	≥ X	79	17933	937	0	1.0000	0.9503	0.1376	0.9506	0.0778	0.1443	**0.9503**

Boucheron数据集评测结果:

公开数据集的评测结果:

Model	Level	TP	TN	FP	FN	TPR	TNR	HSS	Accuracy	Precision	F1	**TSS**
SVM	$\geq C$	16123	57042	16173	5419	0.7484	0.7791	0.4485	0.7721	0.4992	0.5989	0.5275
VGG	$\geq C$	14977	59611	13604	6565	0.6952	0.8142	0.4567	0.7872	0.5240	0.5976	0.5094
Swin_transformer	$\geq C$	16186	58927	14288	5356	0.7514	0.8048	0.4852	0.7927	0.5311	0.6223	0.5562
JW-Flare (single)	$\geq$ C	16345	61772	11443	5197	0.7588	**0.8437**	**0.5465**	**0.8244**	**0.5882**	**0.6627**	0.6025
JW-Flare (series)	$\geq$ C	3767	10950	3132	733	**0.8371**	0.7776	0.5203	0.7920	0.5460	0.6609	**0.6147**

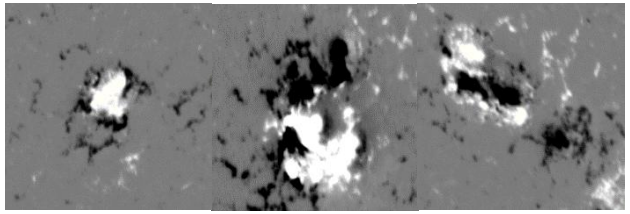
JW-Flare与领域内方法对比结果:

Method	Model	Type	Input	Level	Time	Table	TPR	TNR	ACC	HSS	TSS
Huang et al. (2018)	CNN	AR.C 2	Single	C/M/X	24h	T4	0.73/0.85/0.87	0.76/0.81/0.85	0.76*/0.81*/0.85*	0.34/0.14/0.03	0.49/0.66/0.71
Nishizuka et al. (2018)	DNN	AR.F 1*	Single	C+/M+	24h	T3	0.81/0.95	0.82*/0.85*	0.82/0.86	0.53/0.26	0.63/0.80
Park et al. (2018)	CNN	FD 2	Single	C+*	24h*	T4	0.80*	0.89*	0.84	0.69	0.69
Li et al. (2020)	CNN	AR.C 2	Single	C+/M+	24h	T2	0.889/0.817	0.79*/0.932*	0.861/0.891	0.671/0.759	0.679/0.749
Nishizuka et al. (2021)	DNN,t=0.5	AR.F 1*	Sing					0.99*/0.99*	0.99/0.99	0.64/0.06	0.70/0.24
Zheng et al. (2021)	CNN	AR.C 2	Single	C+/M+	24h	T3	0.895/0.818	0.808*/0.921*	0.869/0.887	0.692/0.746	0.703/0.739
Deng et al. (2021)	GAN+CNN	AR.C 2	Single	C/M/X	24h	T8	0.734/0.746/0.830	0.912*/0.907*/0.932*	0.865/0.864/0.909	0.644/0.655/0.770	0.646/0.653/0.762
Sun et al. (2022a)	3D CNN	AR.C 3	Series	C+/M+	24h	T5	0.862/0.925	0.893*/0.901*	0.879/0.904	0.756/0.667	0.756/0.826
Abduallah et al. (2023)	Transformer	AR.C 1	Series	C+/M+/M 5+	24h	T1	0.891/0.842/0.853	0.944*/0.997/0.965*	0.915/0.928/0.964	---	**0.835**/0.839/0.818
JW-Flare(single)	MLLM	AR.C 2	Single	C+/M5+/X	24h	T3	0.759/0.992/0.974	0.844/0.850/0.918	0.824/0.852/0.918	0.547/0.090/0.075	0.603/0.842/0.893
JW-Flare(series)	MLLM	AR.C 3	Series	C+/M5+/X	24h	T3	0.837/0.966/1.000	0.778/0.910/0.950	0.792/0.911/0.950	0.520/0.156/0.137	0.615/**0.876**/**0.950**

SOTA: 超出约 10%

JW-Flare真实泛化性能测试:

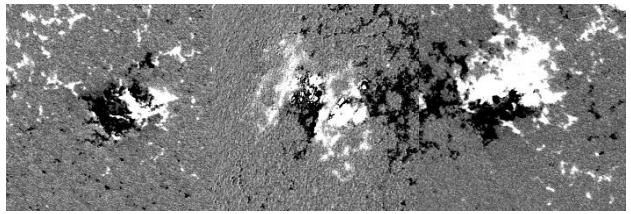
HSOS:



M5.0 X9.3 M5.0

AR13078 AR12673 AR13296

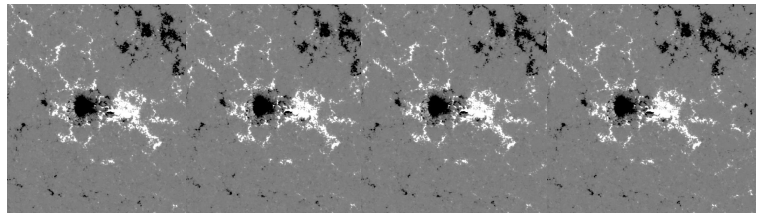
FMG:



X1.6 X1.1 M5.7

AR13663 AR13697 AR13691

HMI:



None

AR13613

JW-Flare(single): ≥M5.0

JW-Flare(series): ≥ X

Data	Input	TP	TN	FP	FN	TPR	TNR	HSS	accuracy	percision	F1	**TSS**
HSOS	Single	400	0	0	83	---	---	---	0.8282	---	---	---
FMG	Single	289	0	0	2	---	---	---	0.9931	---	---	---
HMI	Series	236	21852	1732	95	0.7130	0.9266	0.1860	0.9236	0.1199	0.2053	**0.6396**

JW-Flare可解释性实验： CoT LLM

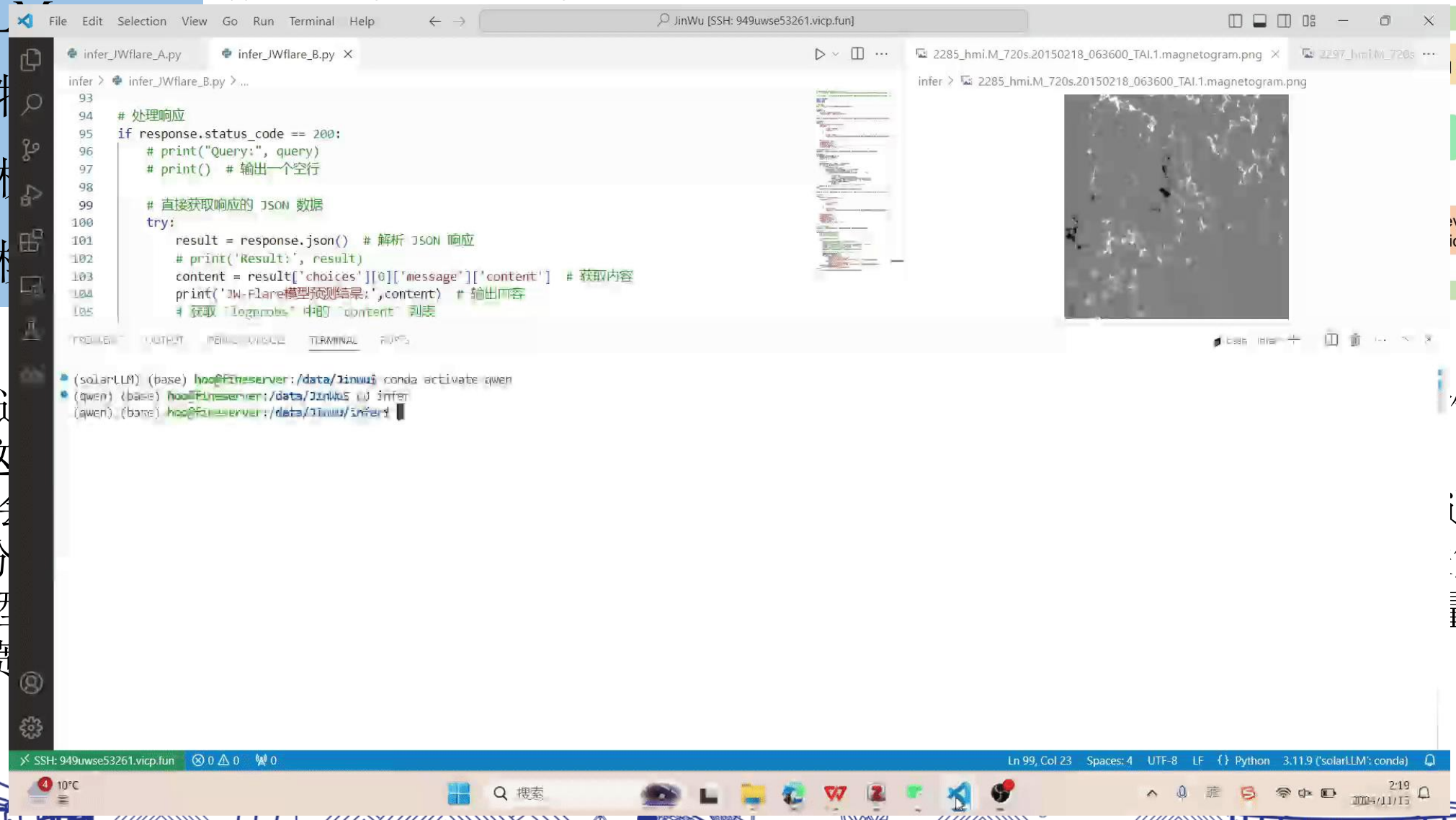
CoT LL

1)数据

2)物理

3)综合

(a) The reasoning and decision-making process in CoT LLM



1） 模型

式——这

2） 模型

数协同分

3） 该模

以得出喷

勺前兆模

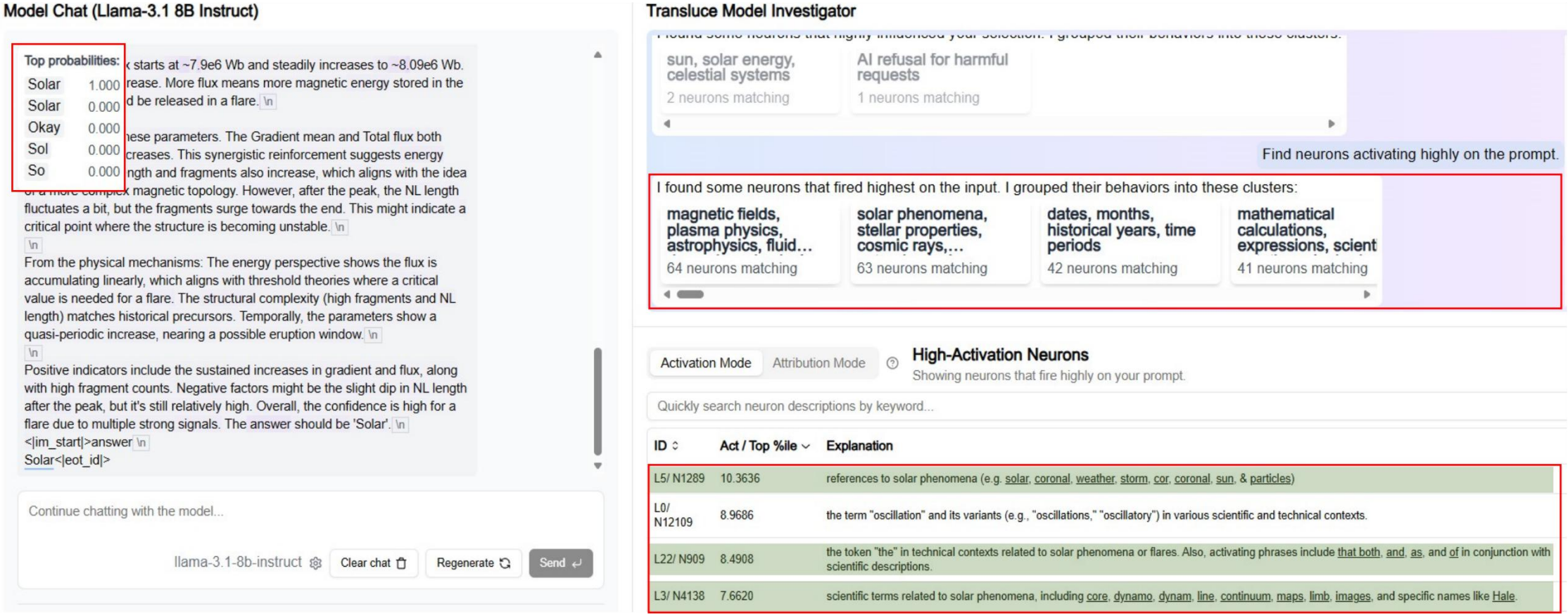
通过多参

触发效应。

量评估，

JW-Flare可解释性实验： Monitor 工具

(b) Monitor Project Interface



消融实验、Scaling law与基座模型:

基座MLLM:

≥ C Flare

Deepseek-vl-7b-chat

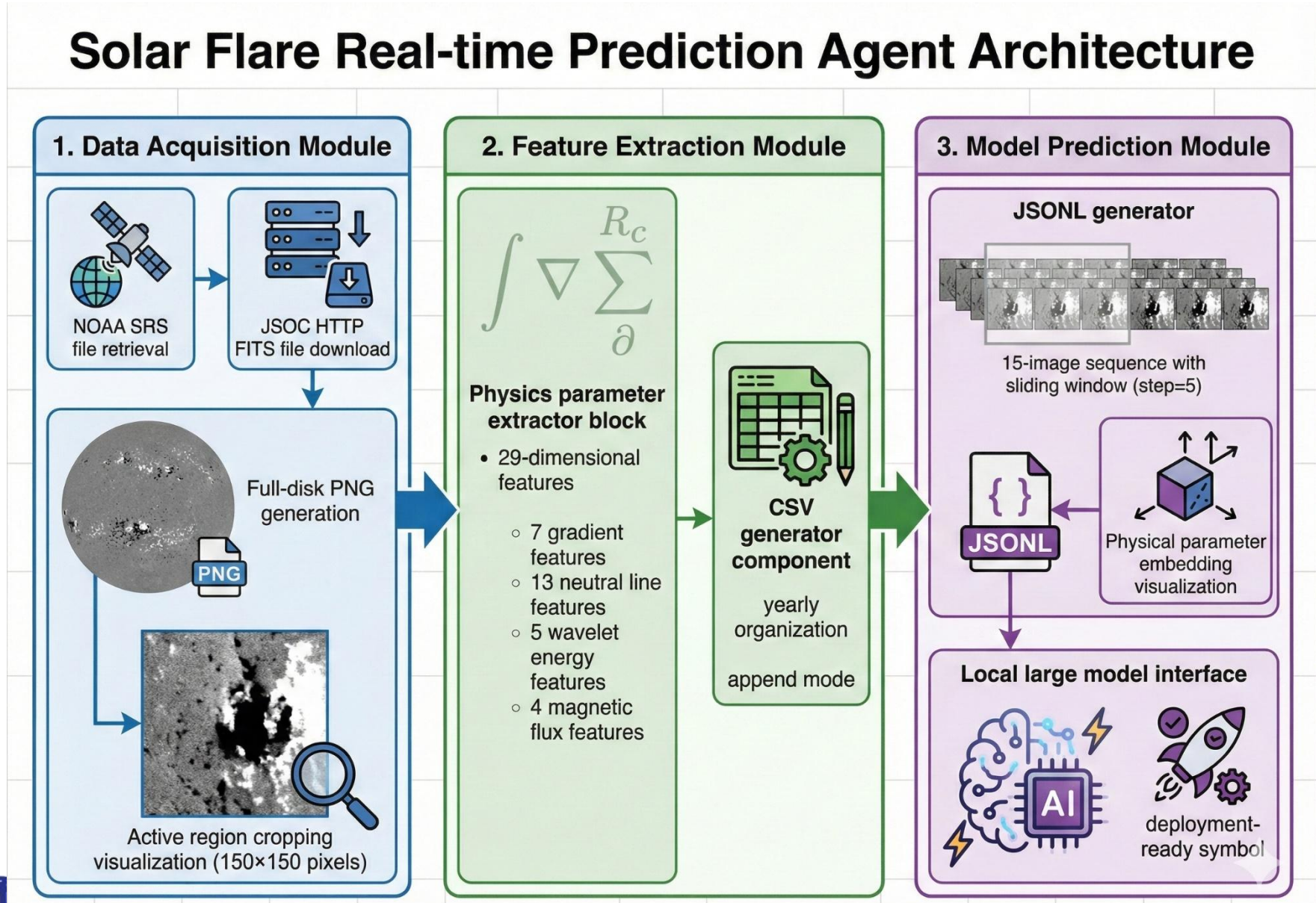
Qwen2-vl-7b-instruct

Model	TP	TN	FP	FN	TPR	TNR	HSS	accuarcy	percision	F1	TSS
Deepseek-single	13697	65024	8191	7845	0.6358	**0.8881**	0.5210	**0.8308**	**0.6258**	0.6308	0.5240
Deepseek-series	3231	12087	2553	1078	0.7498	0.8256	0.5134	0.8084	0.5586	0.6402	*0.5754*
Qwen2-single	18146	56429	16786	3396	**0.8424**	0.7707	0.5028	0.7870	0.5195	0.6426	0.6131
Qwen2-series	3460	11865	2217	1040	0.7889	0.8426	**0.5763**	0.8296	0.6156	**0.6915**	***0.6315***

第五部分：智能体应用开发



JW-Flare智能体工具:



第六部分：总结与展望



总结与展望

https://arxiv.org/abs/2511.08970

总结:

1)卓越性能：在预测最剧烈的X级耀斑时，实现了100%的真正例率，性能远超传统方法。

2)创新方法：成功利用了大语言模型固有的两大优势:

知识基础：模型预训练阶段获得 的太阳物理知识为耀斑预测奠定 了基础。

核心技术：通过构建多模态监督微调数据集（结合物理参数与磁场图）并应用领域适配的提示微调技术，实现了高精度预测。

3)有效验证：通过可解释性实验、缩放定律分析和消融研究，证实了方法的有效性及各技术组件的协同作用。

展望:

1)数据维度优化:

时间尺度：纳入年代尺度的长期太阳活动数据，以捕捉更完整的演化模式。

预测范围：在多小时至数天的不同时间窗口上进行训练验证，提升模型在多时序任务中的适应性。

空间层次：开发分层融合策略，整合太阳大气不同区域（如光球层、色球层、日冕层）的观测数据，提升系统分析的完整性。

2)技术维度演进:

将持续追踪并集成如DeepSeek-R1、Qwen等最先进的开源大模型技术，确保JW-Flare核心引擎始终保持领先。

中国科学院国家天文台
NATIONAL ASTRONOMICAL OBSERVATORIES, CAS



感谢各位老师批评指正!