United Nations/Germany workshop on the International Space Weather Initiative 2024

(*Preparing for the Solar Maximum*)

AI large model for solar activity forecasting

LONG XU

Ningbo University

10 – 14 June 2024 Neustrelitz, Germany

Outline

1、 AI and Deep learning

2、 AI modeling for solar astronomy

3、 AI large model for solar astronomy

4、 Summary

1. AI and deep Learning

Intelligence?

- The ability to learn and solve problems;
- The ability to solve new problems, act rationally, and act like a human;
- There are various levels of intelligence among humans, many animals, and some machines;

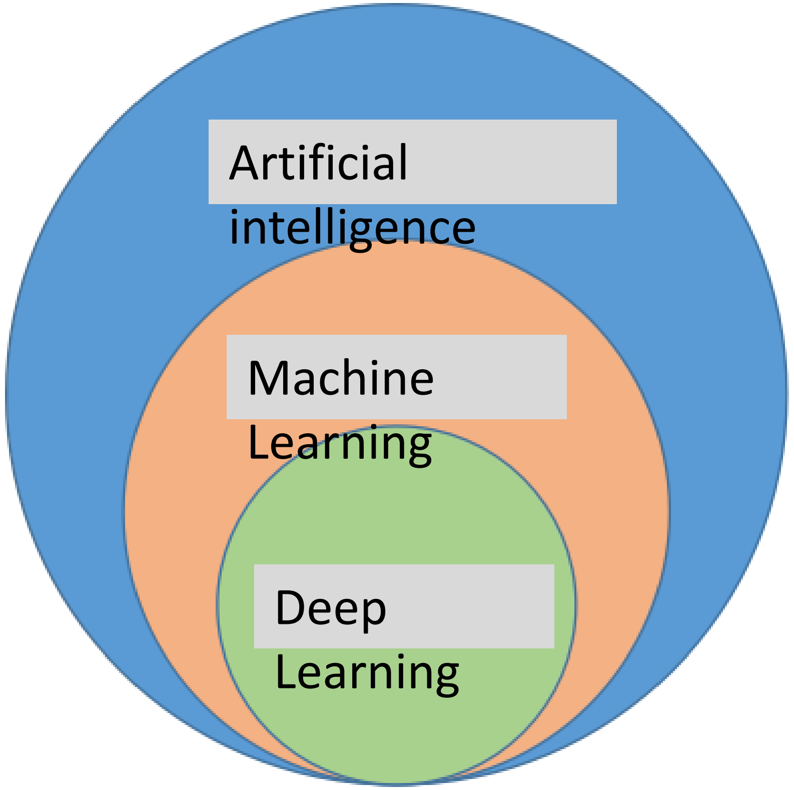
Natural intelligence?

- It refers to the intelligence of various organisms created by natural evolution;
- Human intelligence is the most advanced, developed, and representative natural intelligence on Earth;

Artificial intelligence?

- It refers to artificial intelligence/machine intelligence, which is inspired by the working mechanism of natural intelligence and implemented on machines;
- Artificial intelligence technology is the use of various artificial methods to enable machines to have certain intelligence.

1. AI and deep Learning



Artificial Intelligence (AI): It involves the research and development of theories, methods, technologies, and application systems used to **simulate, extend, and expand human intelligence**.

Machine Learning (ML): It specializes in the study of **using computers to simulate/implement human learning** behaviors in order to acquire new knowledge or skills and reorganize existing knowledge structures to continuously improve its own performance.

Deep Learning (DL): It is a branch of machine learning, focusing on **multi-layer (deep)** neural networks and **feature learning**.

1. AI and deep Learning

History of AI

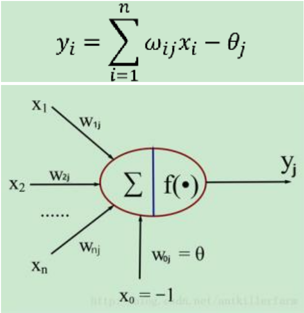
1940s

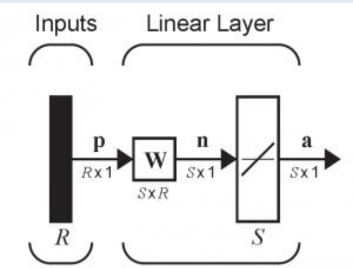
Start-up

MP Model

Summation+Threshold

Hebb Learning





AI founders(1956, Dartmouth)

peak

trough

1960s First Wave

Perception（1957）

peak

1980s Second Wave

Hopfield Network（1982）

Restricted Boltzmann Machine （1986，2006）

BP（1974，85）

trough

peak

2000s

Third Wave

Deep NN：

DBN

DBM

Deep CNN

RNN

ResNet

DenseNet

LSTM

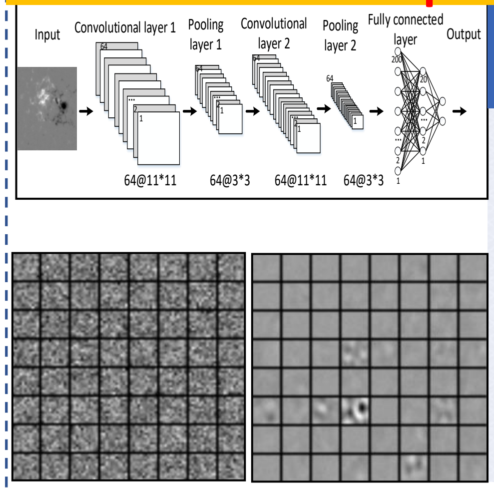
Transformer

2018 Turing Award Recipients

2. AI modeling for solar astronomy

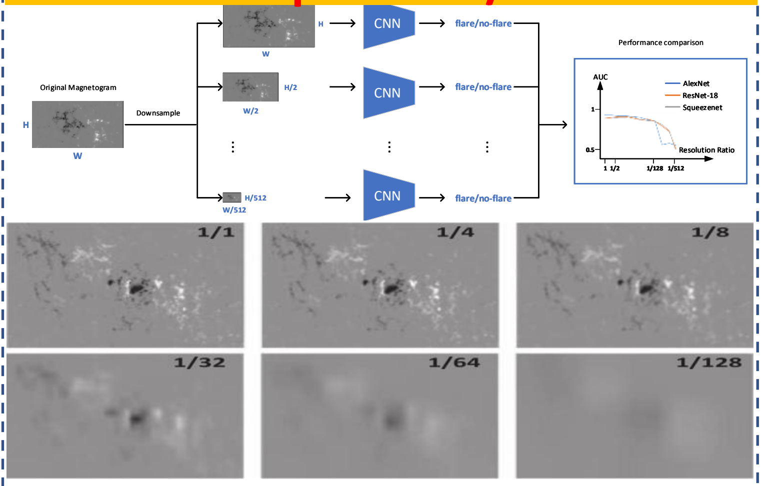
2.1 AI for solar flare forecast

International pioneering model for solar flare prediction



HMI Science Nuggets

Magnetogram can be down-sampled to 1/64



HMI Science Nuggets of SDO
2021 "Chinese Highly-cited paper award"

Deep learning model is unsensitive to resolution of magnetogram;

- We proposed a deep cnn for solar flare forecast, extracting image features from magnetogram directly without prioknoledge of solar flare.
- A dataset was established and published.
- This algorithm has become a baseline of AI in solar flare forecast.
- It was included in the HMI Science Nuggets, and awarded the Chinese Highly-cited Paper in 2021.

X. Huang, H. Wang, L. Xu, ApJ 2018, vol. 856, no. 1, id. 7(IF:5.52, TOP)

S. Liu, L. Xu, et al., ApJS 2022, vol. 941, no. 1(IF： 9.2, TOP)

2. AI modeling for solar astronomy

2.1 AI for solar flare forecast

Validating solar flare forecast using Multiband solar observation

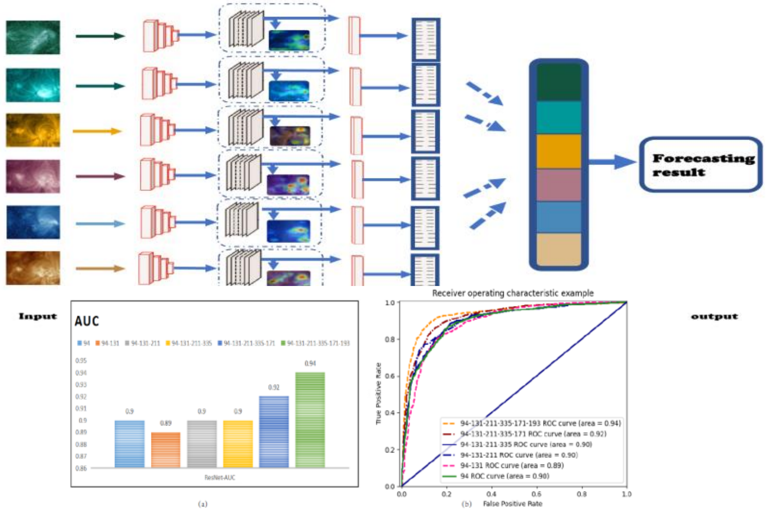
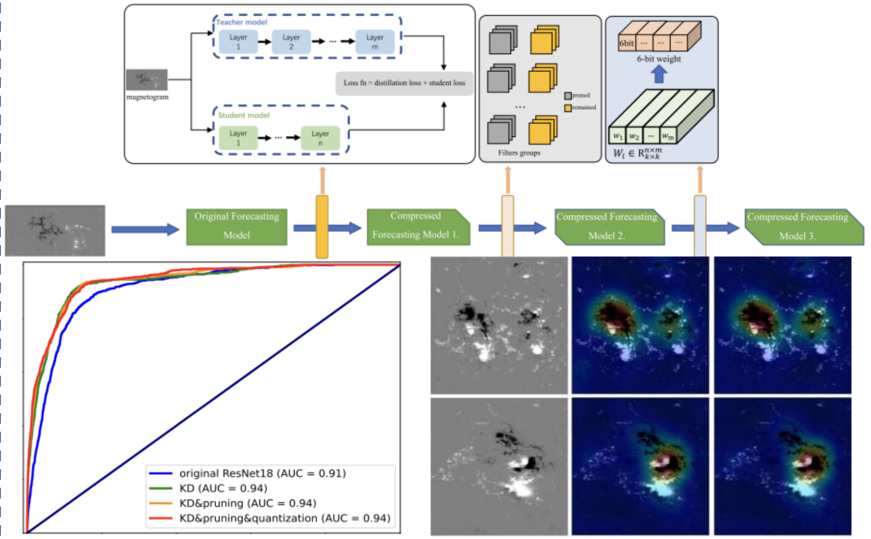


Figure 11. AUC (a) and ROC (b) charts of the combined forecasting model with multi-wavelength results on Resnet-18.

Validate the feasibility for flare forecast from UV/EUV observations using deep learning models

*D. Sun, X. Huang, Z. Zhao and **L. Xu**, ApJS 2023（IF：9.2, TOP）, vol. 266, no. 1, id. 8*

Lightweight of deep learning models



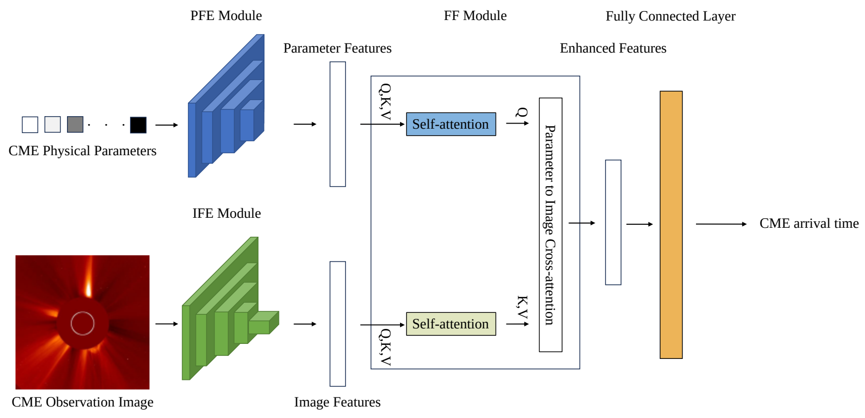
A framework of integrating Knowledge distillation, pruning, quantification；Compress the deep learning models to 1.67% of its original size without obvious performance drop.

*K. Feng, **Long Xu***, ApJS 2022（IF：9.2, TOP）, vol. 268, no. 59*

2. AI modeling for solar astronomy

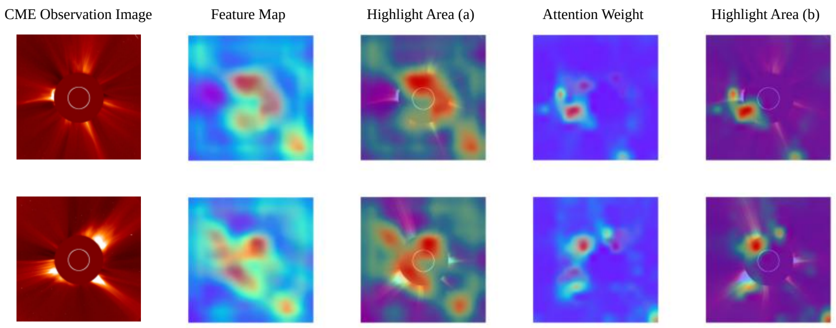
2.2 AI for CME arrival time

Multi-modal CME Forecast Model



International pioneering model for CME using both CEM image and CME parameters, and attention module to fuse two modalities.

Visualization analysis



The region of interest in regular CNNs is larger than that of the algorithm proposed in this paper, which only covers the main area of CME. This indicates that the algorithm in this paper focuses on crucial information while ignoring irrelevant details.

Statistics

Methods	MAE (hours)	RMSE (hours)
Sudar et al. (2015)	11.56	-
Liu et al. (2018)	11.58 ± 1.21	15.52 ± 1.58
Ours (PFE module)	9.63 ± 0.93	12.04 ± 1.12

Methods	MAE (hours)	RMSE (hours)
Wang et al. (2019)	12.42	-
VGG (Simonyan & Zisserman 2014)	12.79 ± 1.25	16.23 ± 1.56
DenseNet (Huang et al. 2017)	11.63 ± 1.34	15.57 ± 1.67
MobileNet (Howard et al. 2017)	10.50 ± 1.27	14.02 ± 1.33
Ours (IFE module)	9.51 ± 1.03	12.19 ± 1.21

Methods	MAE (hours)	RMSE (hours)
Ours (PFE module)	9.63 ± 0.93	12.04 ± 1.12
Ours (IFE module)	9.51 ± 1.03	12.19 ± 1.21
Ours (PFE + IFE)	9.06 ± 0.91	11.70 ± 1.23
[illegible]	[illegible]	[illegible]

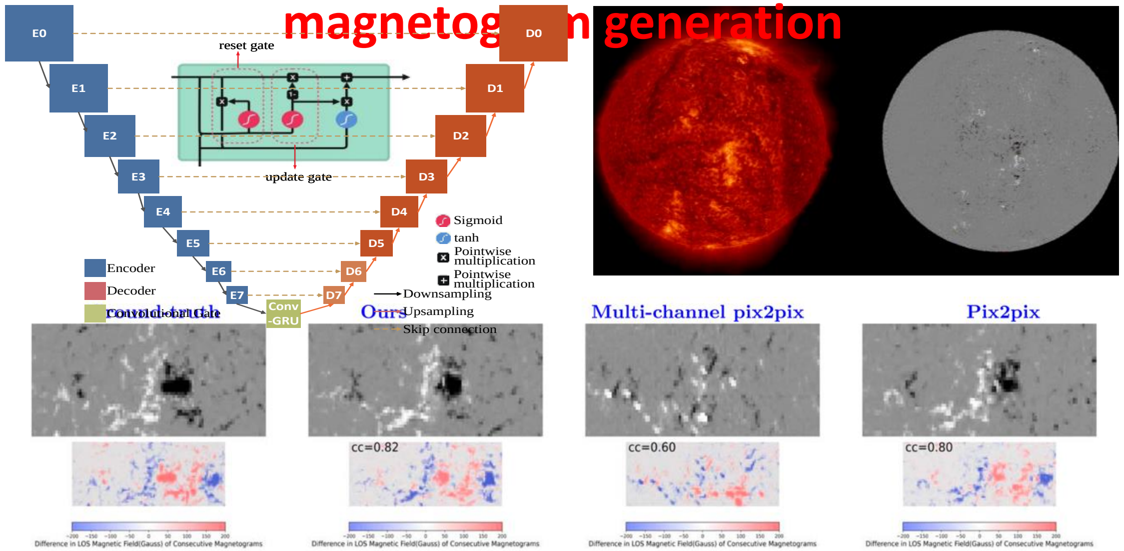
The experimental results show that the algorithm proposed in this paper has a prediction error of 9.63 hours for the arrival time of CMEs, which outperforms the existing best algorithms.

Yufeng Zhong, Dong Zhao, Xin Huang, ***Long Xu*****, ApJS 2024 (IF: 9.2), vol. 271, no. 2, id. 31, 2024*

2. AI modeling for solar astronomy

2.3 AI for magnetogram generation

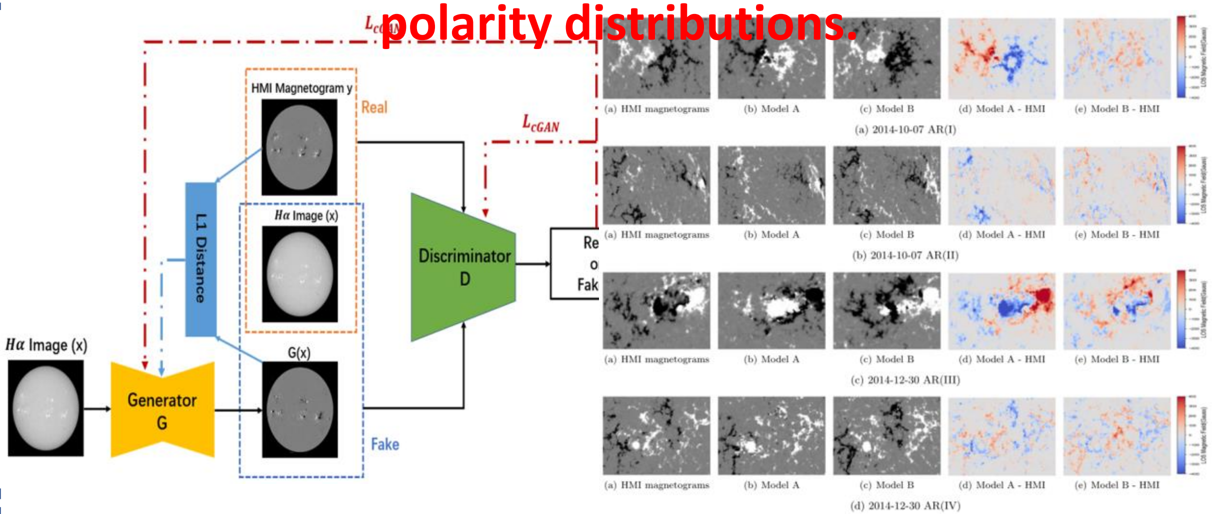
The first dynamic neural network for magnetogram generation



A dynamic neural network model is proposed for the first time to generate magnetograms from EUV images, which alleviates the magnetic pole fluctuation of static models in generating magnetogram sequences.

*W. Sun, **L. Xu***, et.al., ApJS 2022（IF：9.2）, vol 262, no. 2, id. 45*

Time series data can guide the generation of better polarity distributions.



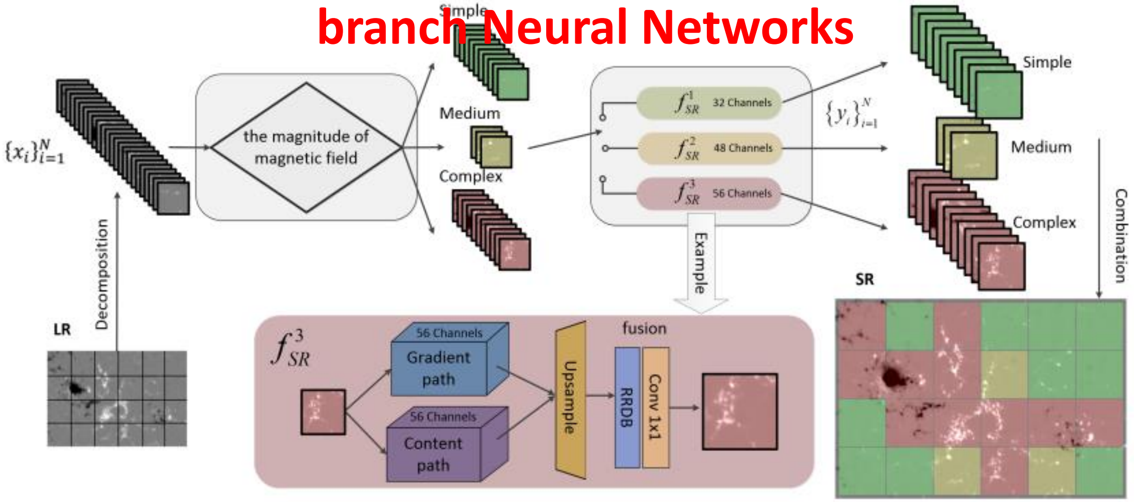
A GAN framework was proposed to generate SDO/HMI magnetograms from Hα images. By investigating two data arrangements during model training, random shuffle and chronological, we found that the time series information benefits better magnetic field polarity.

*F. Gao, T. Liu, W. Sun, **Long Xu***, ApJS 2023（IF：9.2）, vol. 266, no. 2, id. 19*

2. AI modeling for solar astronomy

2.3 AI for magnetic field extrapolation

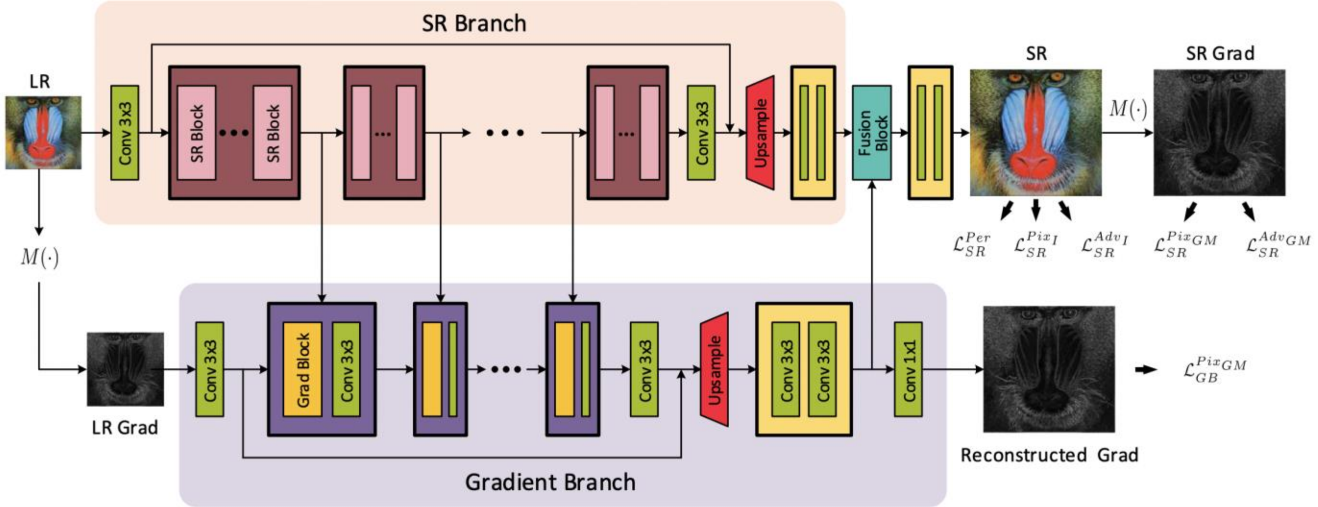
Super-resolution of Magnetogram using Multi-branch Neural Networks



	PS/N	SSIM	CC	RMSE
Bicubic	31.3669	0.8487	0.9109	27.4155
HighRes-Net	32.8036	0.8633	0.9322	23.1134
SPSR-GAN	31.1899	0.8041	0.8942	27.7780
MBSR (Shannon entropy)	33.4007	0.8763	0.9347	21.5153
MBSR (TUMF)	33.4418	0.8765	0.9350	21.4112

*Fengping Dou, **Long Xu***, Zhixiang Ren, Dong Zhao, ApJS 2024 （IF：9.2）, vol. 271, no. 1, id. 9, 2024*

Three branches are for simple, medium and complex patches respectively;
Simple branch has the least channels.



Splitting of the Data Set for Training, Validation, and Testing

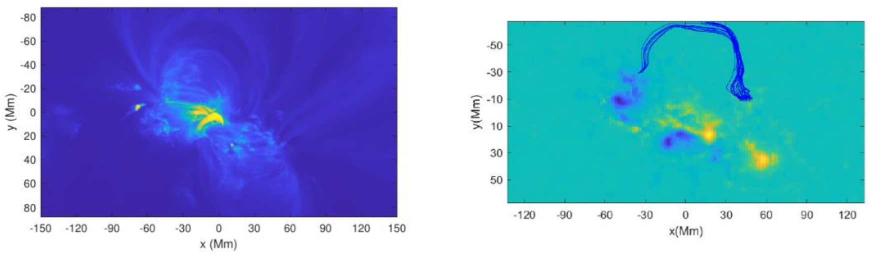
	Training Set	Validation Set	Test Set
Time interval	2010-5–2010-11-15	2010-11-15–2010-11-30	2010-12
Number of samples	6282	170	290
Number of patches	141,349	3825	6830

2. AI modeling for solar astronomy

2.4 AI for magnetic field extrapolation

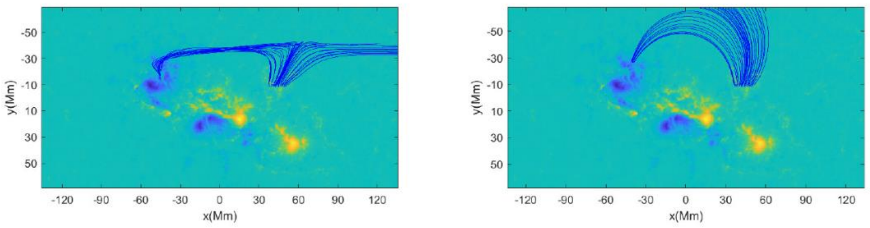
Magneto-hydro-static state /non-force-free field	Force-free field
$\boldsymbol{j}\times\boldsymbol{B}-\nabla p+\rho \boldsymbol{g}=0$	$\boldsymbol{j}\times\boldsymbol{B}=0$ (No Lorentz force)
$\nabla\times\boldsymbol{B}=\mu_0\boldsymbol{j}$	$\nabla\times\boldsymbol{B}=\mu_0\boldsymbol{j}$
$\nabla\cdot\boldsymbol{B}=0$	$\nabla\cdot\boldsymbol{B}=0$

Computing result on AR 11158

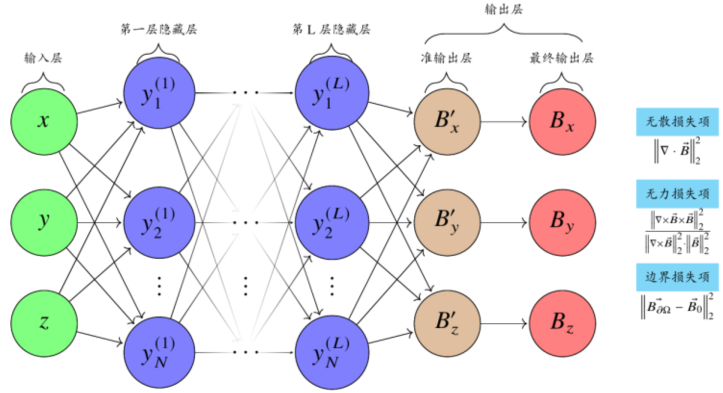


(a) 193 Å波段观测成像[96]

(b) 无力场神经网络外推法



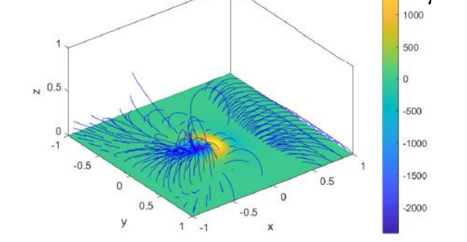
Physics-informed neural network (PINN) for



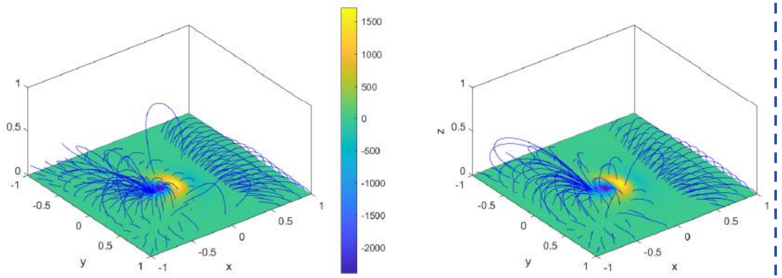
Grid-free: alternative numerical computation schemes such as finite difference, finite volume,

Comparison among reference、our model and potential field

$$(1-\mu^2)\frac{d^2F}{d\mu^2}+n(n+1)F+a^2\frac{1+n}{n}F^{1+\frac{2}{n}}=0, F(\pm1)=0, \frac{dF}{d\mu}(\mu=-1)=10$$



(a) 参考磁场



- Deep neural network became one of the numerical algorithms of partial differential equation, exhibiting good ability;
- Differential equations are widely applied in various engineering fields and natural science fields to describe scientific phenomena and engineering systems, such as in magnetohydrodynamic systems;
- In magnetic field extrapolation, traditional methods require the grid of space, its precision is related to grid while

2. AI modeling for solar astronomy

2.4 AI for magnetic field extrapolation

Data Time Span: 2010-2022

Data Computing Time: 2021.12.10-2022.03.19

Data Volume >280T

Total amount of data >73000

Website of dataset:

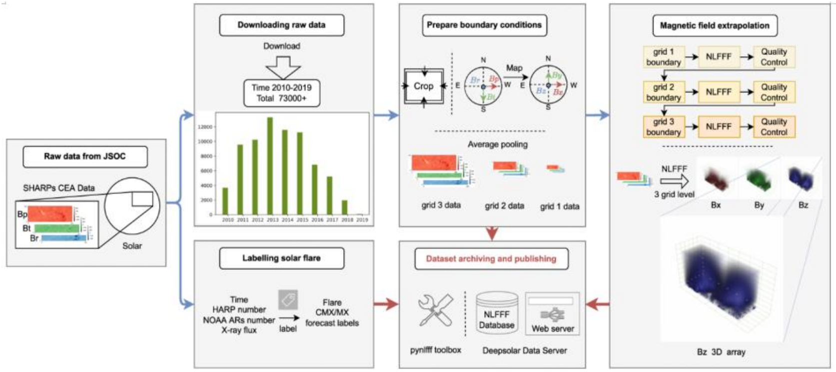
https://nlfff.dataset.deepsolar.space/zh/index.html

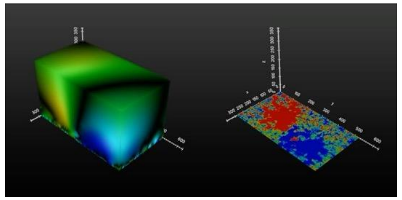
Publish *Nature Scientific Data* paper:

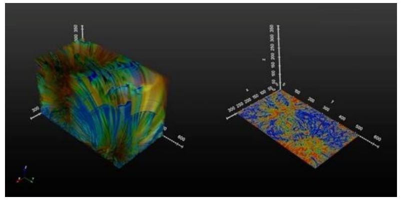
https://www.nature.com/sdata/

Publish the largest dataset of 3D magnetogram (the largest spatio-temporal resolution, the biggest time span)

Large-scale nonlinear force-free magnetogram extrapolation







Zhengrui Zhao, Long Xu*, et.al., Nature Scientific data （IF:8.5）, vol. 10, id. 178, 2023.

computing power and excellent hardware and software platform support provided by the "Pengcheng Cloud Brain II", we have successfully completed the magnetic field extrapolation for all active regions from 2010 to the present. We have improved the CPU parallelization and GPU acceleration of magnetic field extrapolation, resulting in a speed-up of 1000 times. The database

2. AI modeling for solar astronomy

2.4 AI for magnetic field extrapolation

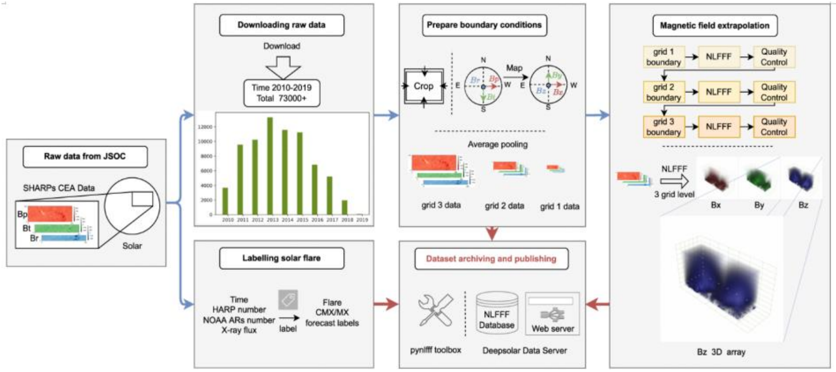
Data Time Span:	2010-2022
Data Computing Time:	2021.12.10-2022.03.19
Data Volume	>280T
Total amount of data	~73000

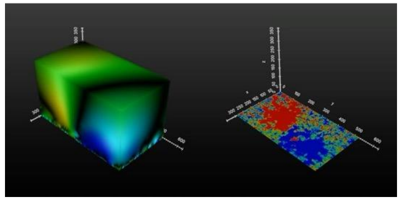
Website of dataset:
https://nlfff.dataset.deepsolar.space/zh/index.html

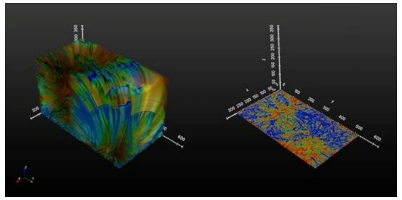
Publish *Nature Scientific Data* paper:
https://www.nature.com/sdata/

Publish the largest dataset of 3D magnetogram (the largest spatio-temporal resolution, the biggest time span)

Large-scale nonlinear force-free magnetogram extrapolation







Zhengrui Zhao, Long Xu, et.al., Nature Scientific data* (IF:8.5), vol. 10, id. 178, 2023.

Challenges:

1. GPU memory optimization and computation scheduling for high-precision scientific computing;
2. Ultra-large file reading and writing on the cloud;
3. Complex process scheduling control for magnetic field extrapolation algorithms;
4. Program deployment in hybrid software environments.

2. AI modeling for solar astronomy

2.4 AI for magnetic field extrapolation

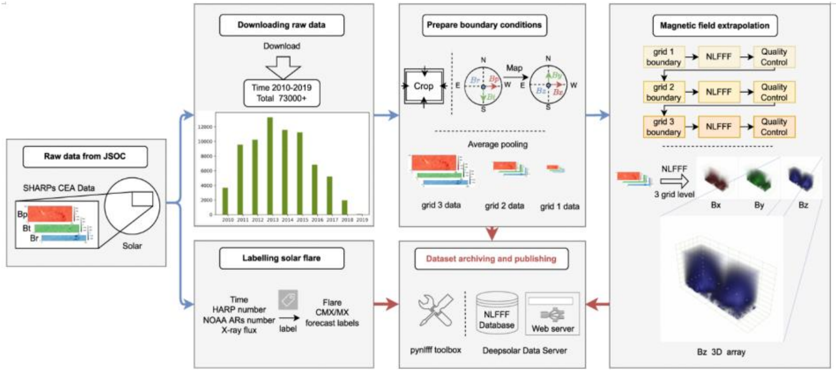
Data Time Span:	**2010-2022**
Data Computing Time:	**2021.12.10-2022.03.19**
Data Volume	**>280T**
Total amount of data	**~73000**

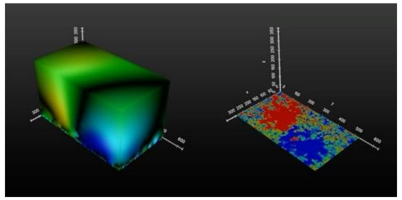
Website of dataset:
https://nlfff.dataset.deepsolar.space/zh/index.html

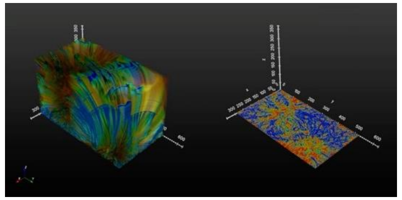
Publish *Nature Scientific Data* paper:
https://www.nature.com/sdata/

Publish the largest dataset of 3D magnetogram (the largest spatio-temporal resolution, the biggest time span)

Large-scale nonlinear force-free magnetogram extrapolation







Zhengrui Zhao, Long Xu, et.al., Nature Scientific data （IF:8.5）, vol. 10, id. 178, 2023.*

Eight large-scale databases for training AI models have been constructed and are currently available (http://www.deepsolar.space/). The corresponding algorithms and data resources have also been open-sourced via (https://openi.pcl.ac.cn).

3. AI large model for solar astronomy

Large-scale Model?

- Large models refer to AI models with a significant number of parameters, often in the billions or trillions.
- They can process vast amounts of data, enabling complex tasks like language understanding, image recognition, and more.

Developments

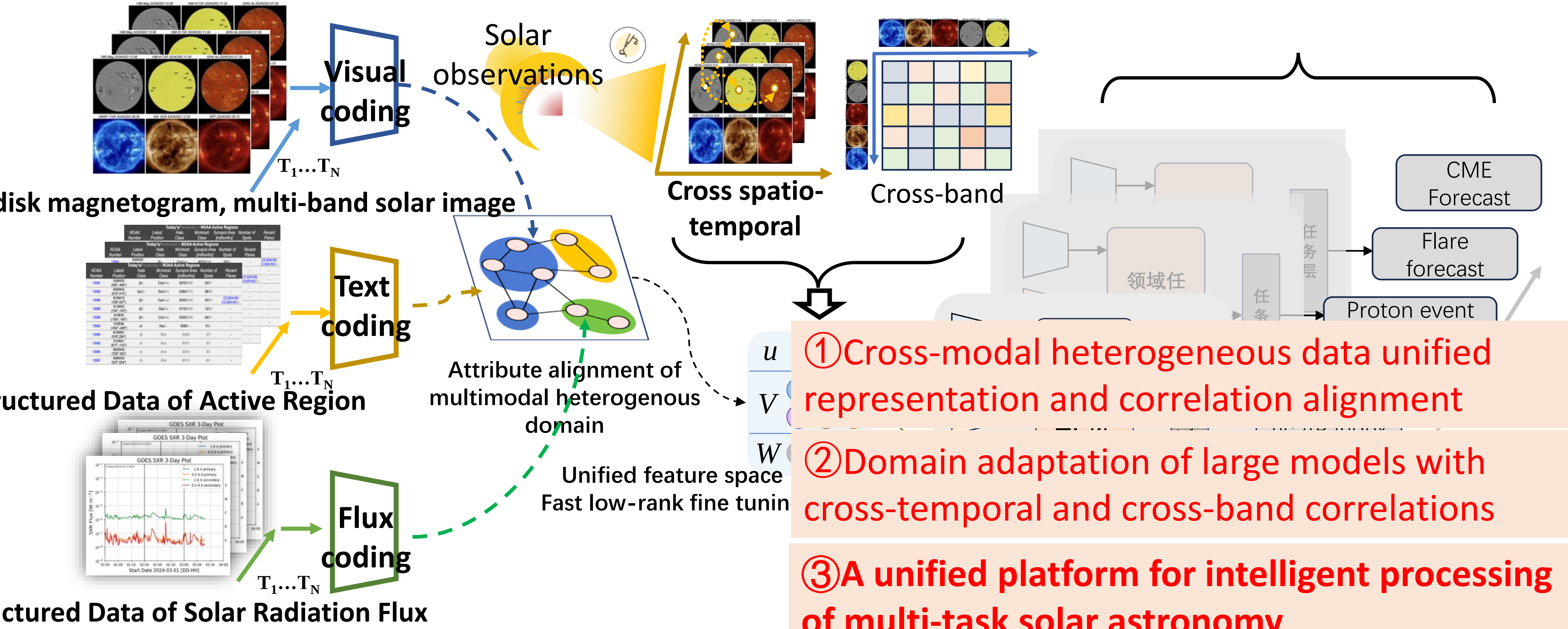
- The introduction of the Transformer in 2017 revolutionized NLP.
- The release of GPT-2 and GPT-3 showed the capabilities of LLMs.
- GPT-4o is astonishing in multimodality, seamlessly integrating multiple modalities; closer to human-like interaction; human-like emotions and sentiments.
- Image Large Models: CLIP, SAM, etc.

Applications

- Language translation, NLP, sentiment analysis, Image recognition, object detection, and computer vision;
- Medical diagnosis, drug discovery, healthcare, autonomous vehicles and robotics, financial analysis and stock prediction.

3. AI large model for solar astronomy

An AI large model for solar astronomy



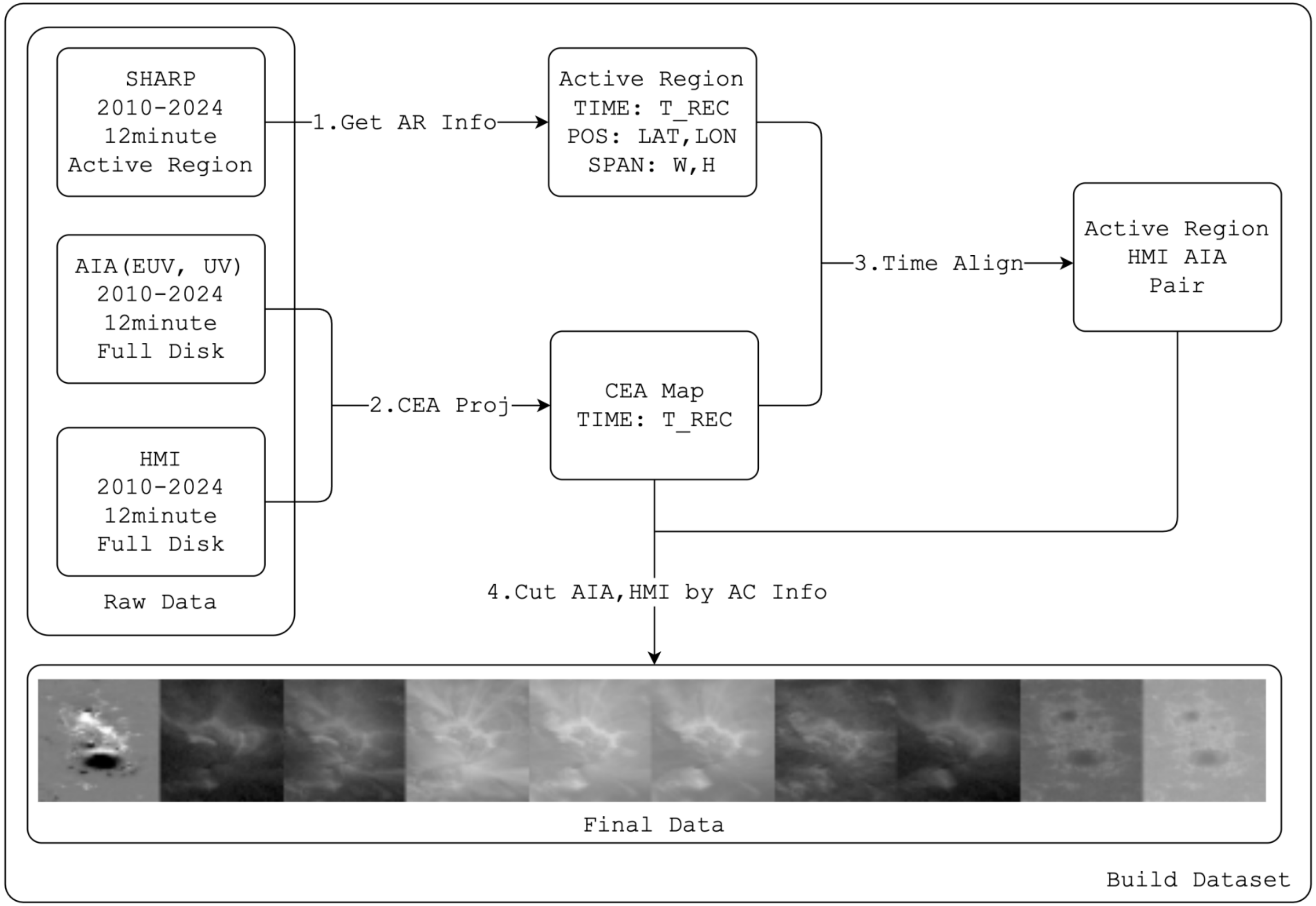
①Cross-modal heterogeneous data unified representation and correlation alignment

②Domain adaptation of large models with cross-temporal and cross-band correlations

③A unified platform for intelligent processing of multi-task solar astronomy

3. AI large model for solar astronomy

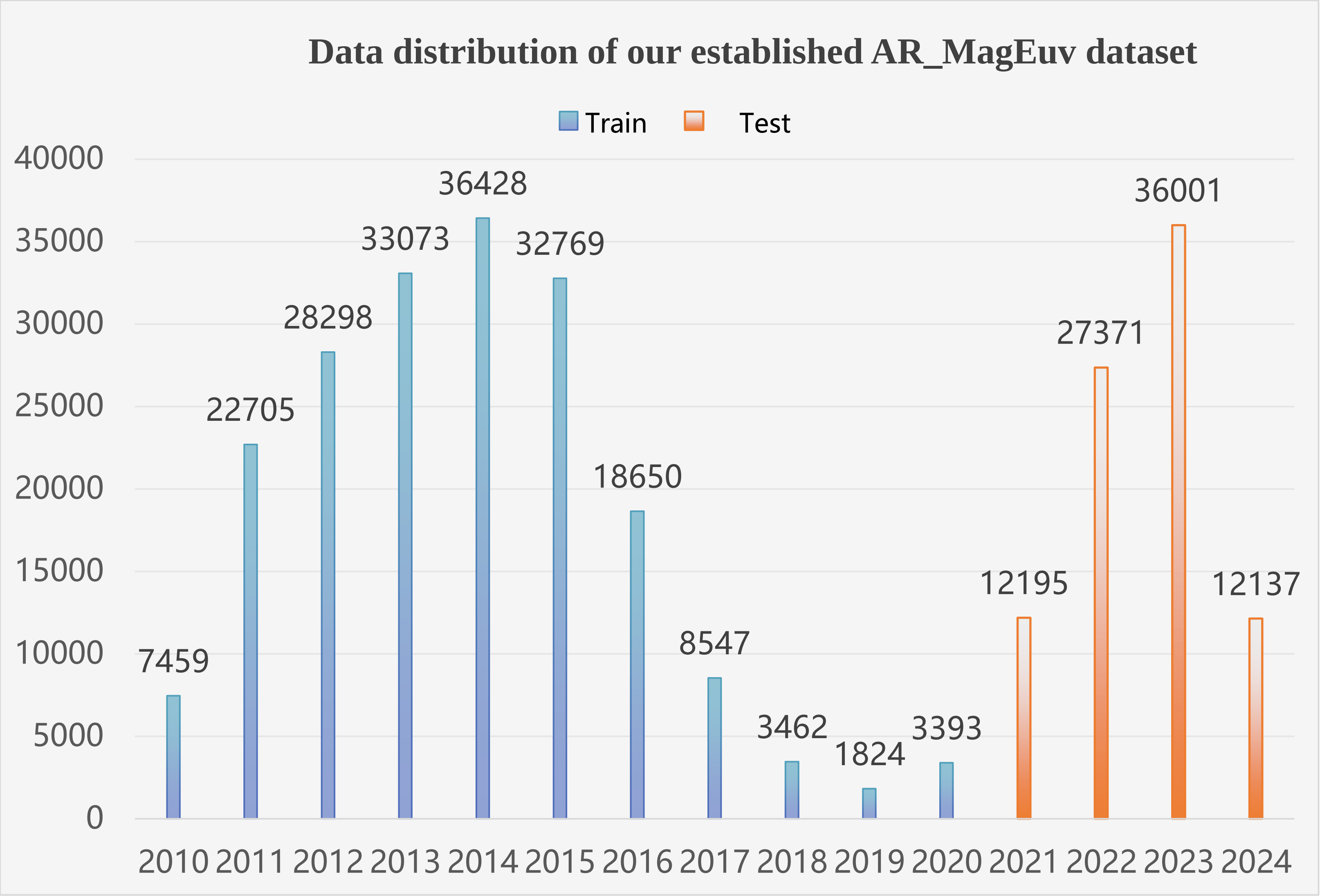
Building a dataset of active region(magnetogram and multi-band UV/EUV images)



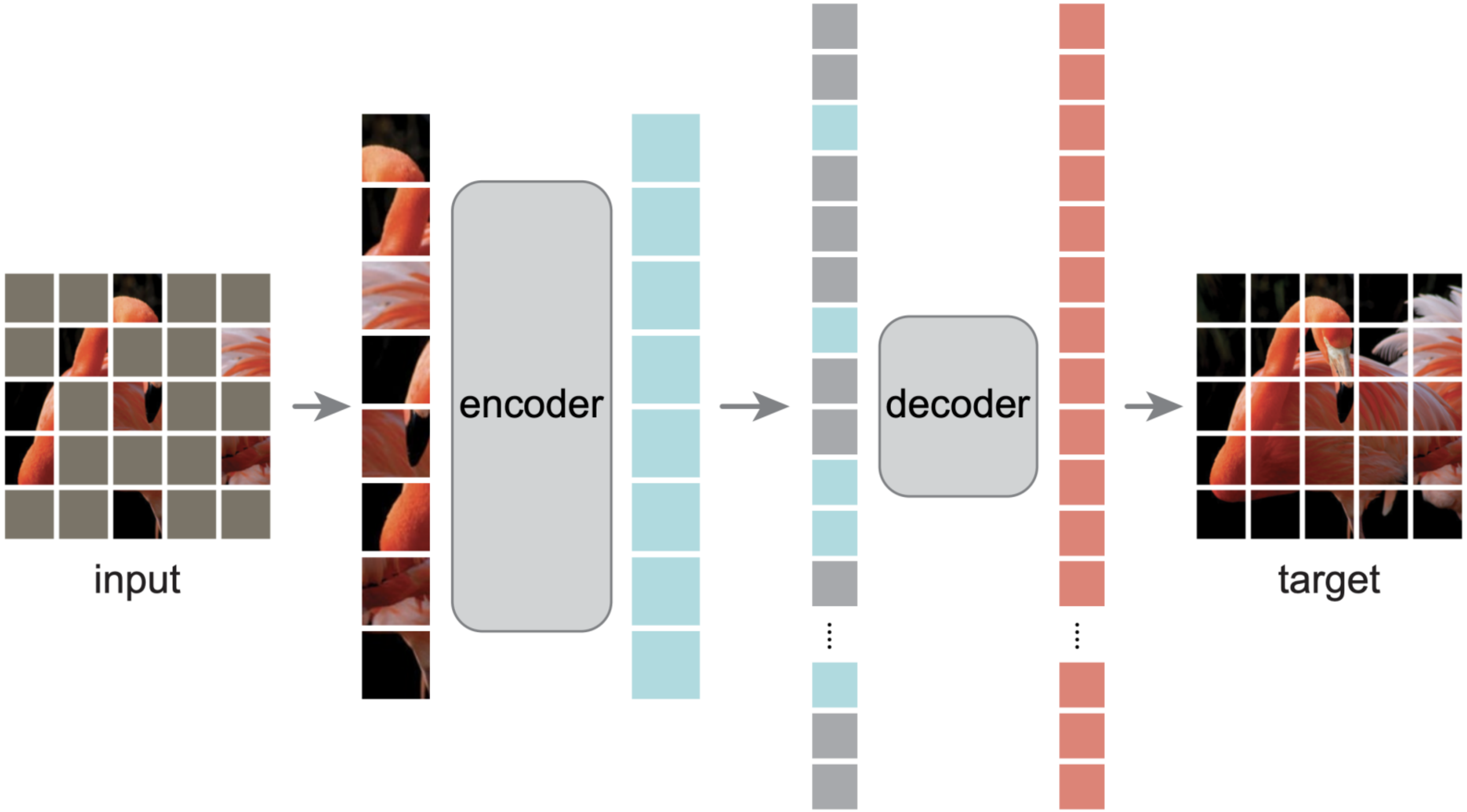
3. AI large model for solar astronomy

Chronological Data Distribution Chart

Train	196608
Test	87704



The Baseline of AI Large Model: Mask-AutoEncoder (MAE)



3. AI large model for solar astronomy

3. AI large model for solar astronomy

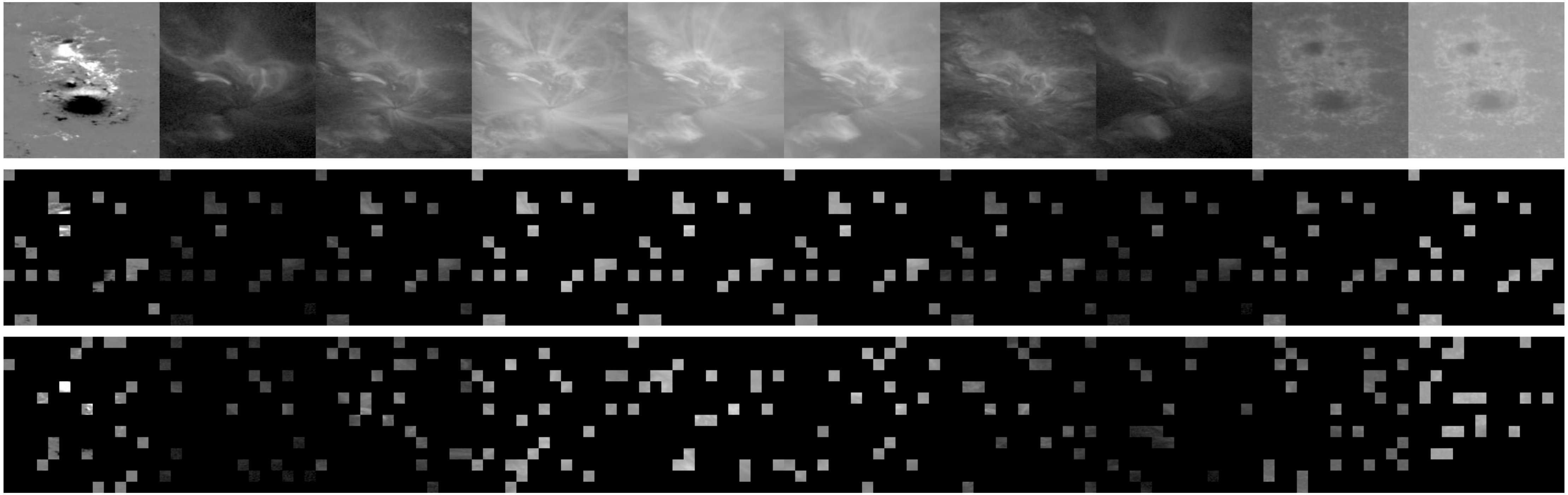
Configuration of Pre-Trained Model

Configurations	
Input Size	224x224x10
Optimizer	adamw
Epochs	300
Warmup Epochs	400

Volume		
Model	Parameters(M)	FLOPS(G)
Tiny	1.27	0.25
Small	18.85	3.67
Base	72.36	14.18
Large	250.11	49.01

3. AI large model for solar astronomy

Different masks among channels: ensure cross-band reference between channels



3. AI large model for solar astronomy

The performance of MAE： quality of generated images

Model	Mask ratio	PSNR
Tiny	0.9	**32.21**
	0.75	**34.26**
	0.5	**34.62**
Small	0.9	**33.55**
	0.75	**36.90**
	0.5	**36.82**
Base	0.9	**33.78**
	0.75	**36.45**
	0.5	**36.76**
Large	0.9	**33.48**
	0.75	**35.67**
	0.5	**37.22**

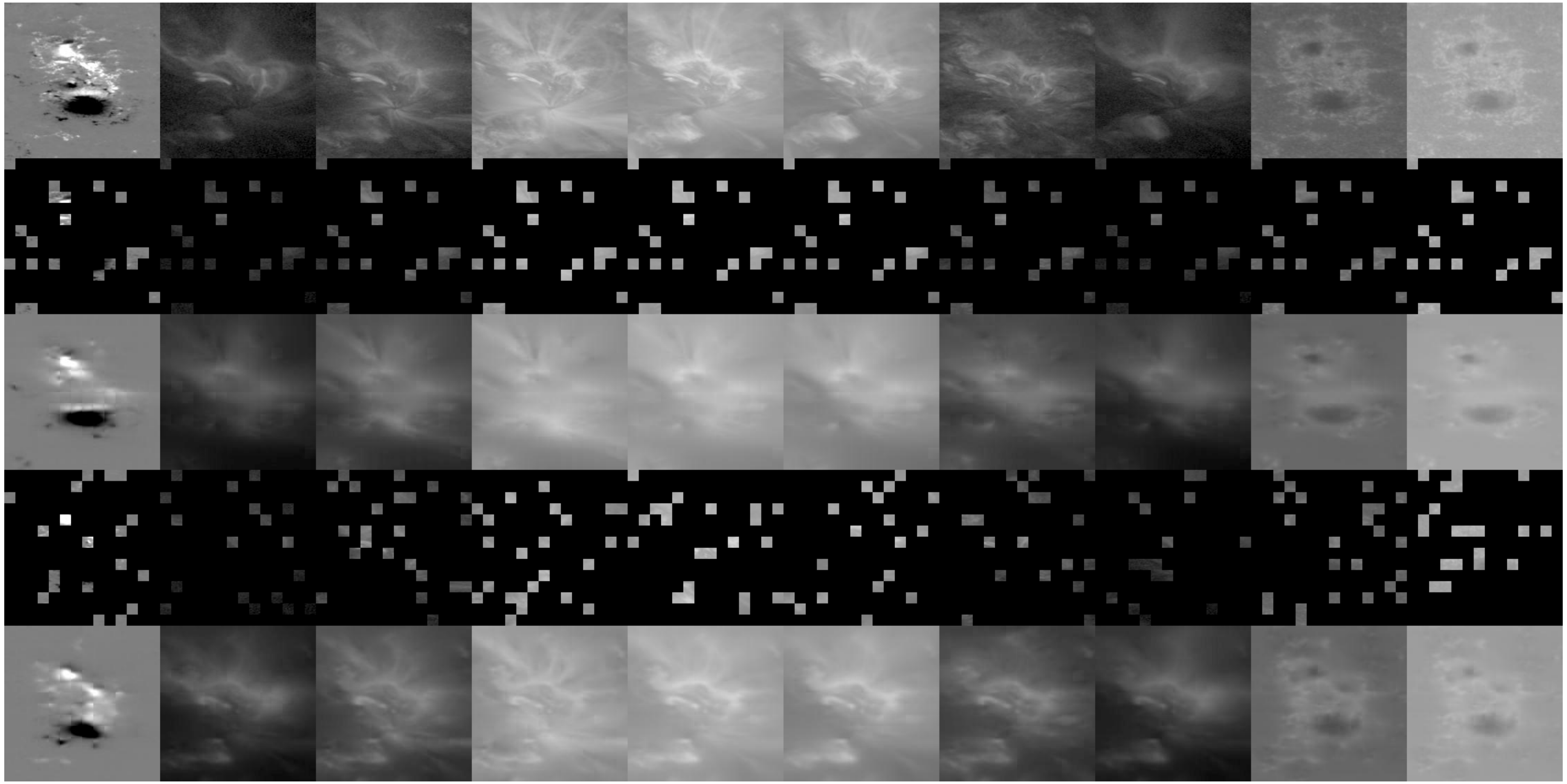
3. AI large model for solar astronomy

The comparison between the same mask and random mask

Small model;
mask_ratio=0.9

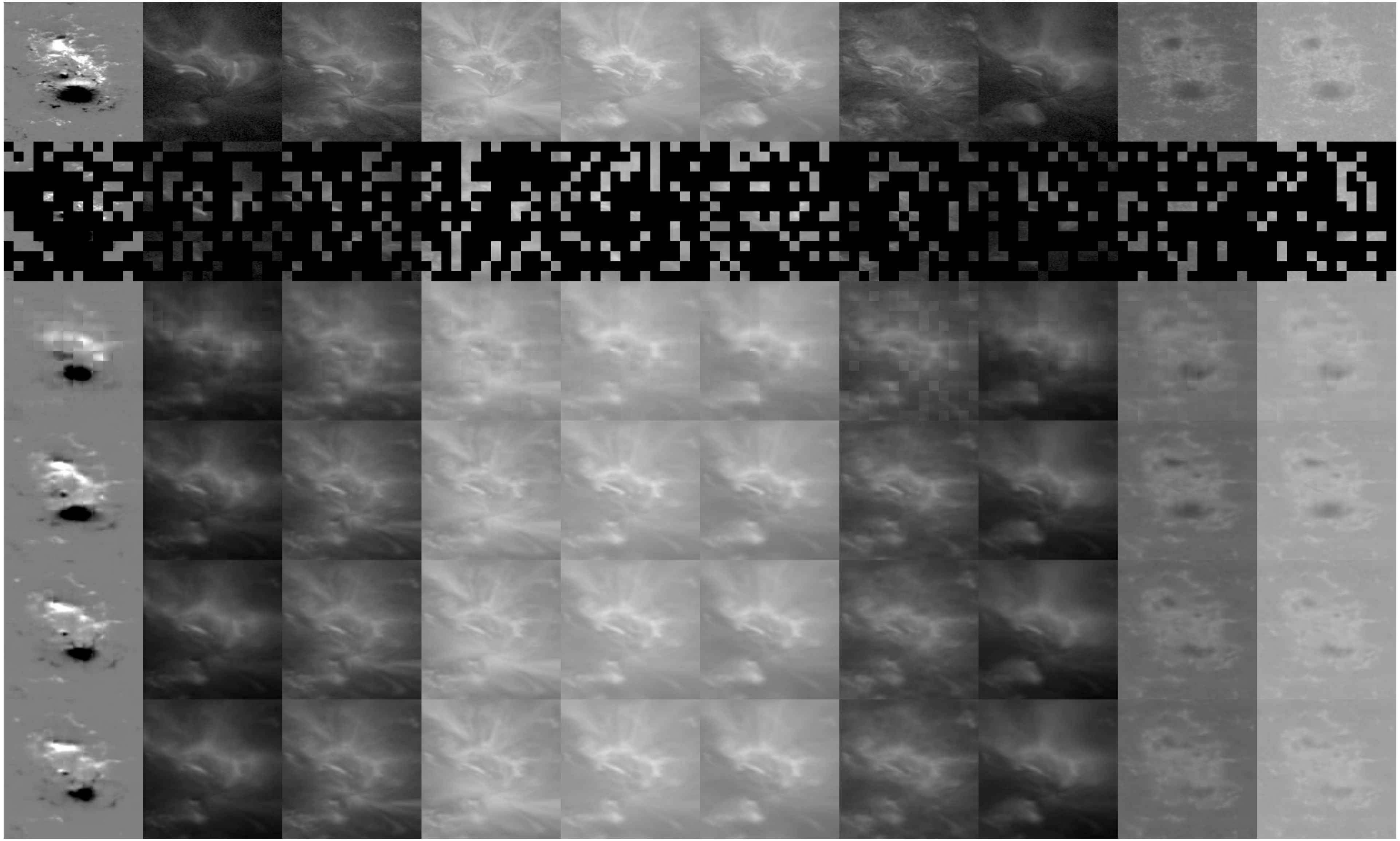
The same masks for all channels

Random masks between channels



3. AI large model for solar astronomy

Comparison among different models with random mask



mask_type=random
mask_ratio=0.75

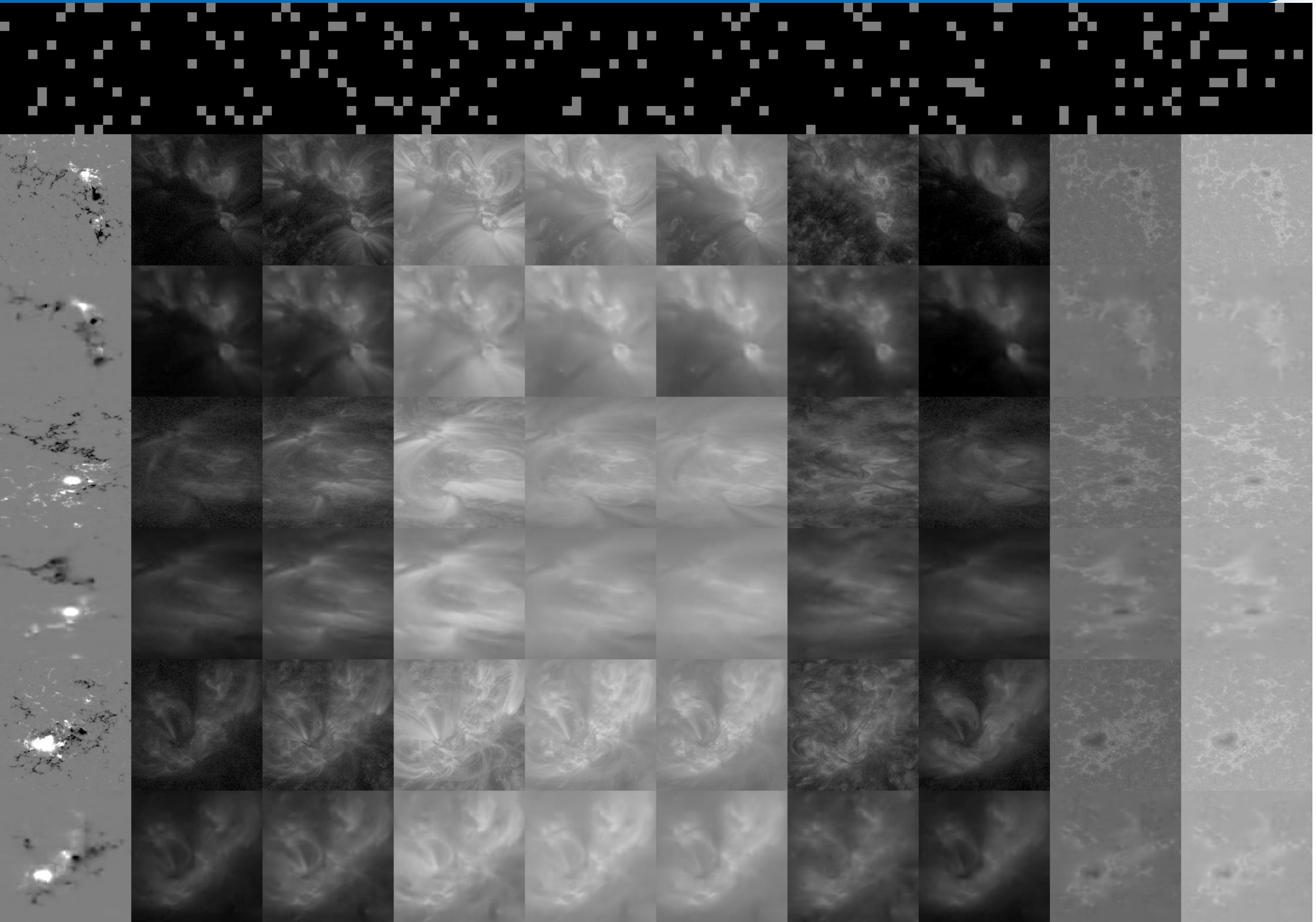
Tiny

Small

Base

Large

3. AI large model for solar astronomy

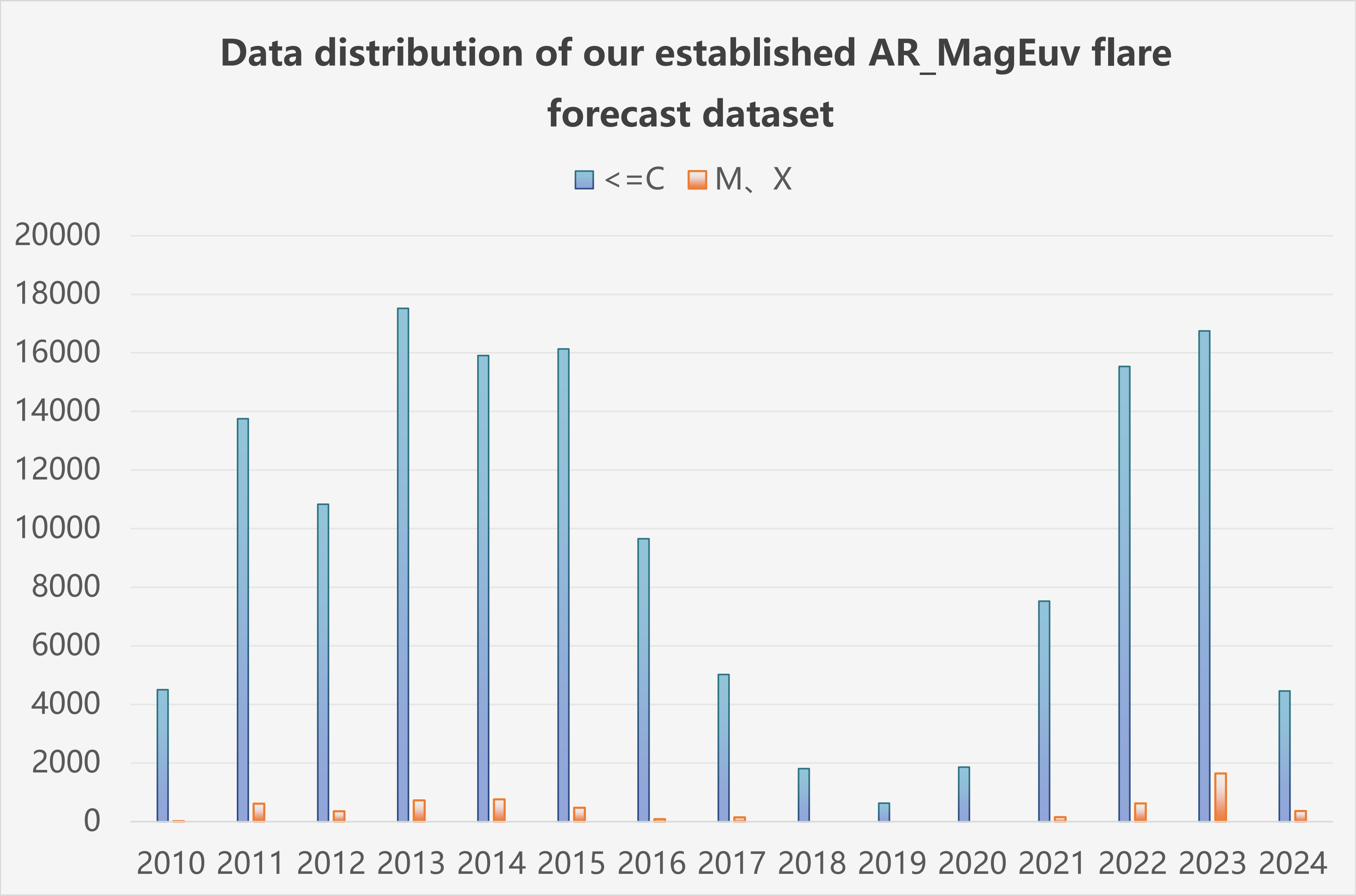


base model
mask_type=random
mask_ratio=0.9

3. AI large model for solar astronomy

24-h flare forecast (M&X)

	<C	M、X
Train	97623	3181
Test	44265	2768



3. AI large model for solar astronomy

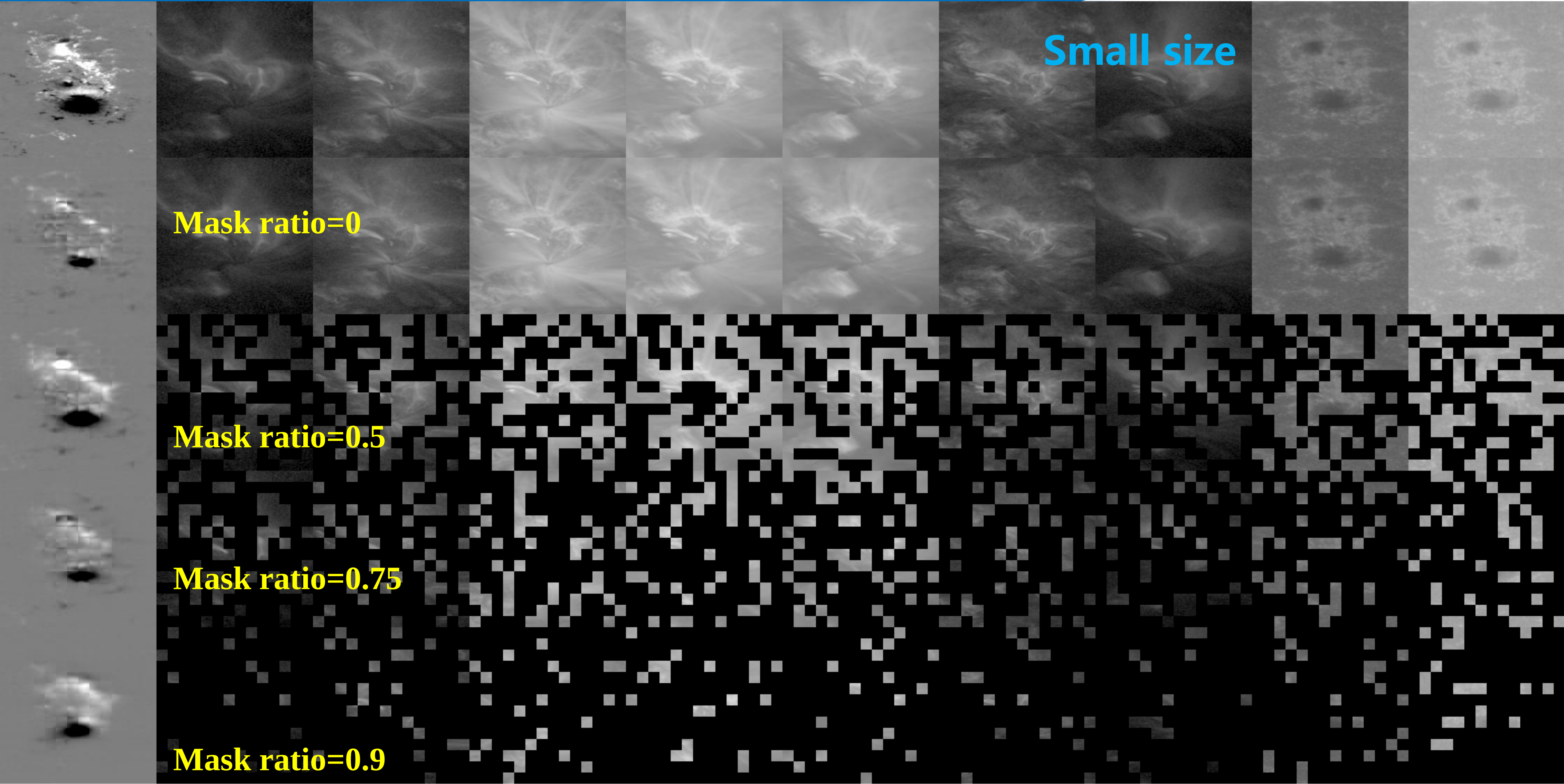
Performance of flare forecast

	Actual True	Actual False
Predict Pos	1895	5098
Predict Neg	873	39167

Auc	0.8943
Acc	0.8730
Precision	0.2710
Recall	0.6846

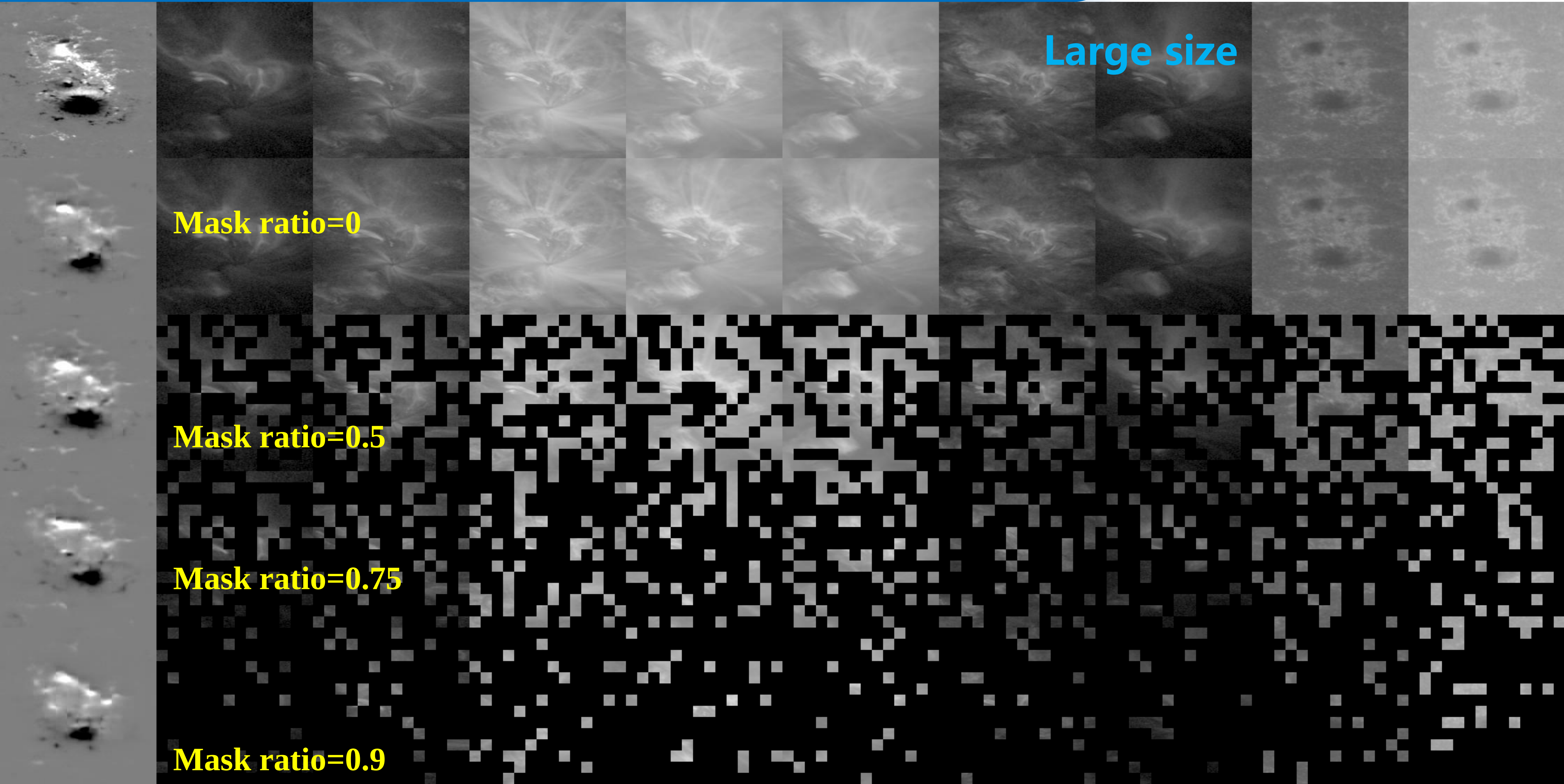
3. AI large model for solar astronomy

Magnetogram generation



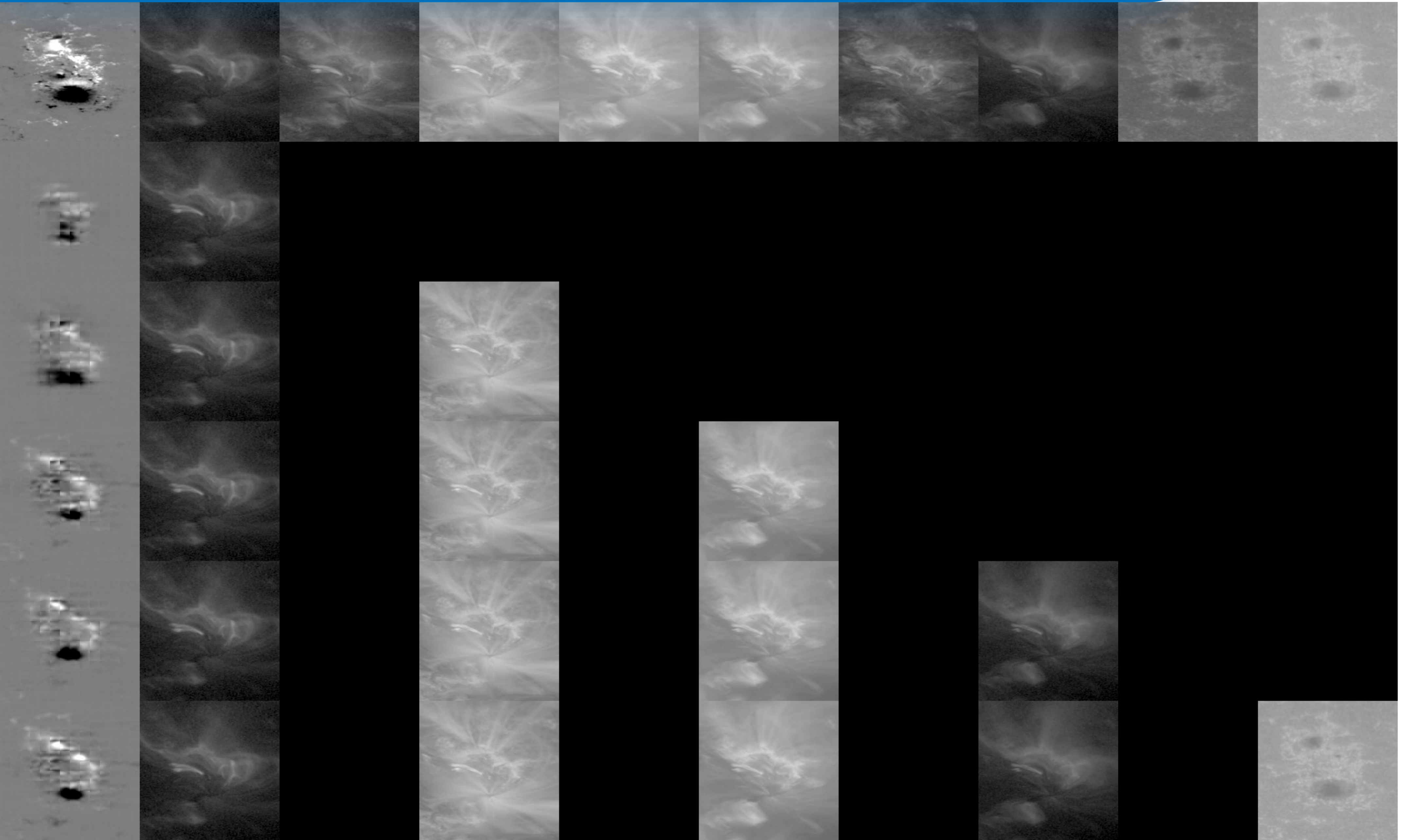
3. AI large model for solar astronomy

Magnetogram generation



3. AI large model for solar astronomy

Magnetogram generation



One channel

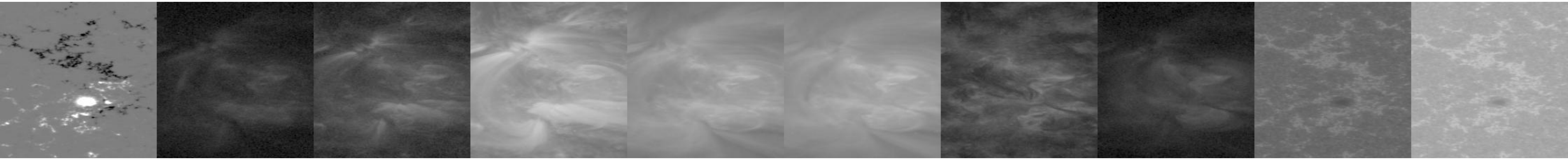
Two channels

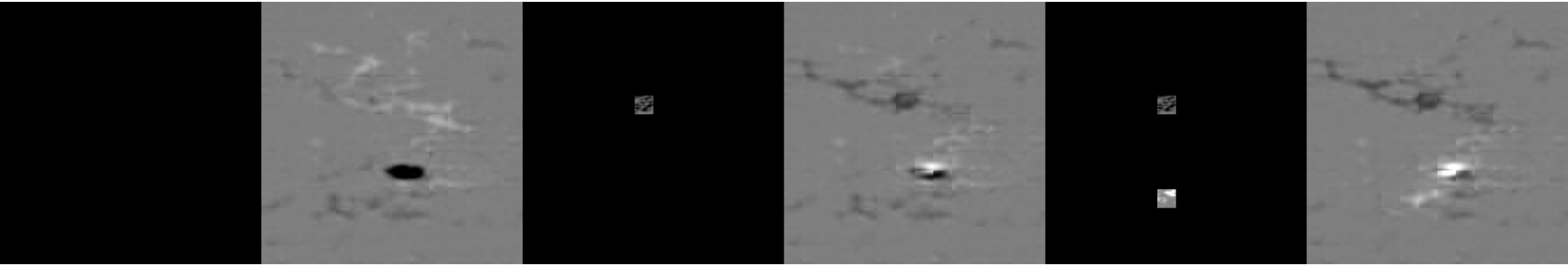
Three channels

Four channels

Five channels

Magnetogram's Polarity

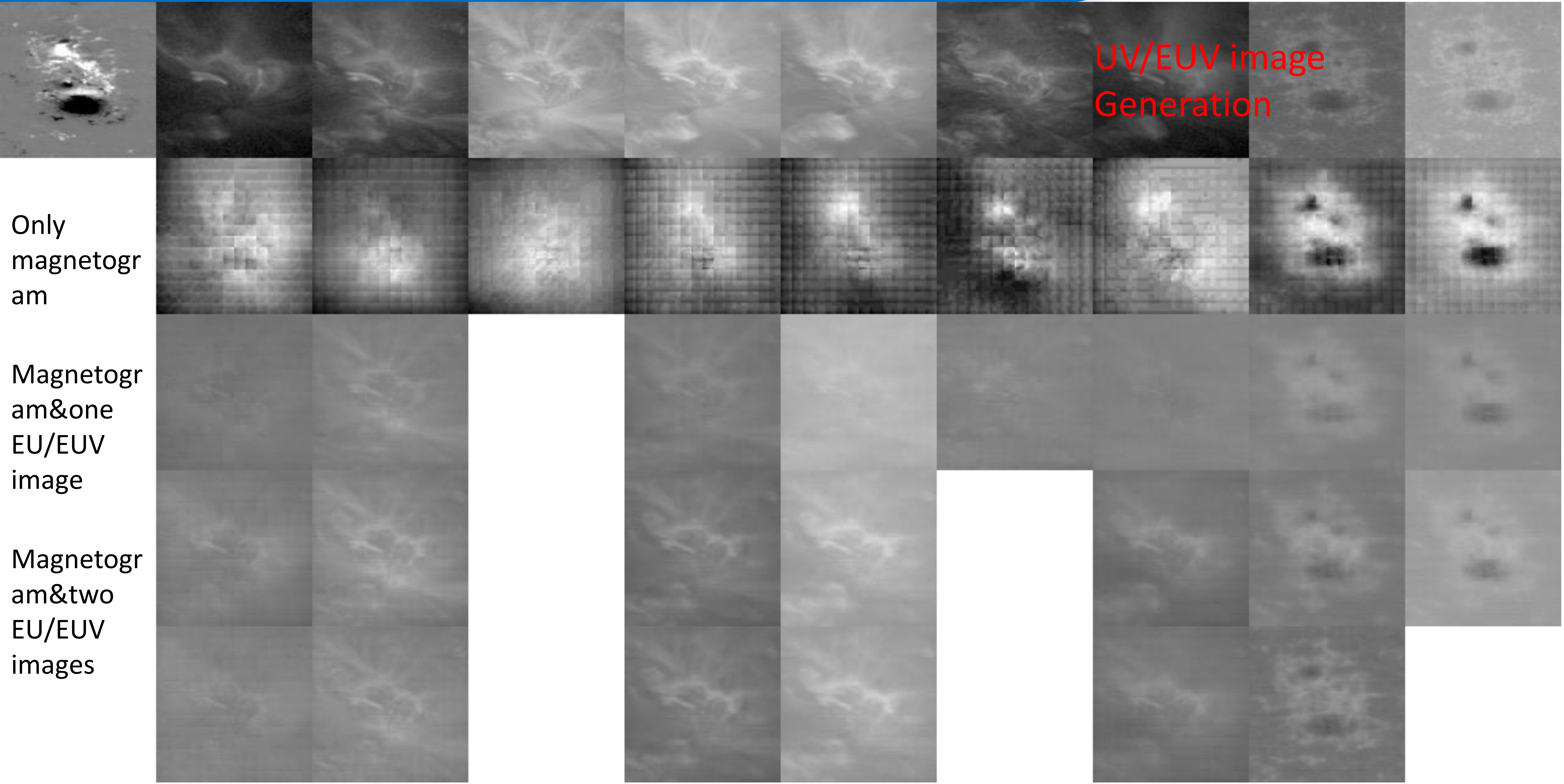




It is interesting that the prompt of magnetogram’s polarity (very small patch of magnetogram) help the model to correct the polarity of generated magnetogram.
In practice, the prompt can be any information, including hand-drawn image, text (current model can not accept text).

3. AI large model for solar astronomy

model_size: small
mask_type: random



4. Summary

- AI has been widely used in various intelligence processing of solar astronomy, achieving big success;
- AI has been witnessed its advantages in various single task of solar astronomy (image processing, generation and prediction);
- It is the beginning that the AI large model is developed for solar astronomy.
- The AI large model is an unified model integrated multiple tasks in an AI model to process a multi-task integrated multiple intelligence tasks; in addition, it is trained over a very huge database, usually with a unsupervised manner.
- The better AI large model should be multiple modality, especially benefits from Language Large Model (LLM).

Thank you for your attention