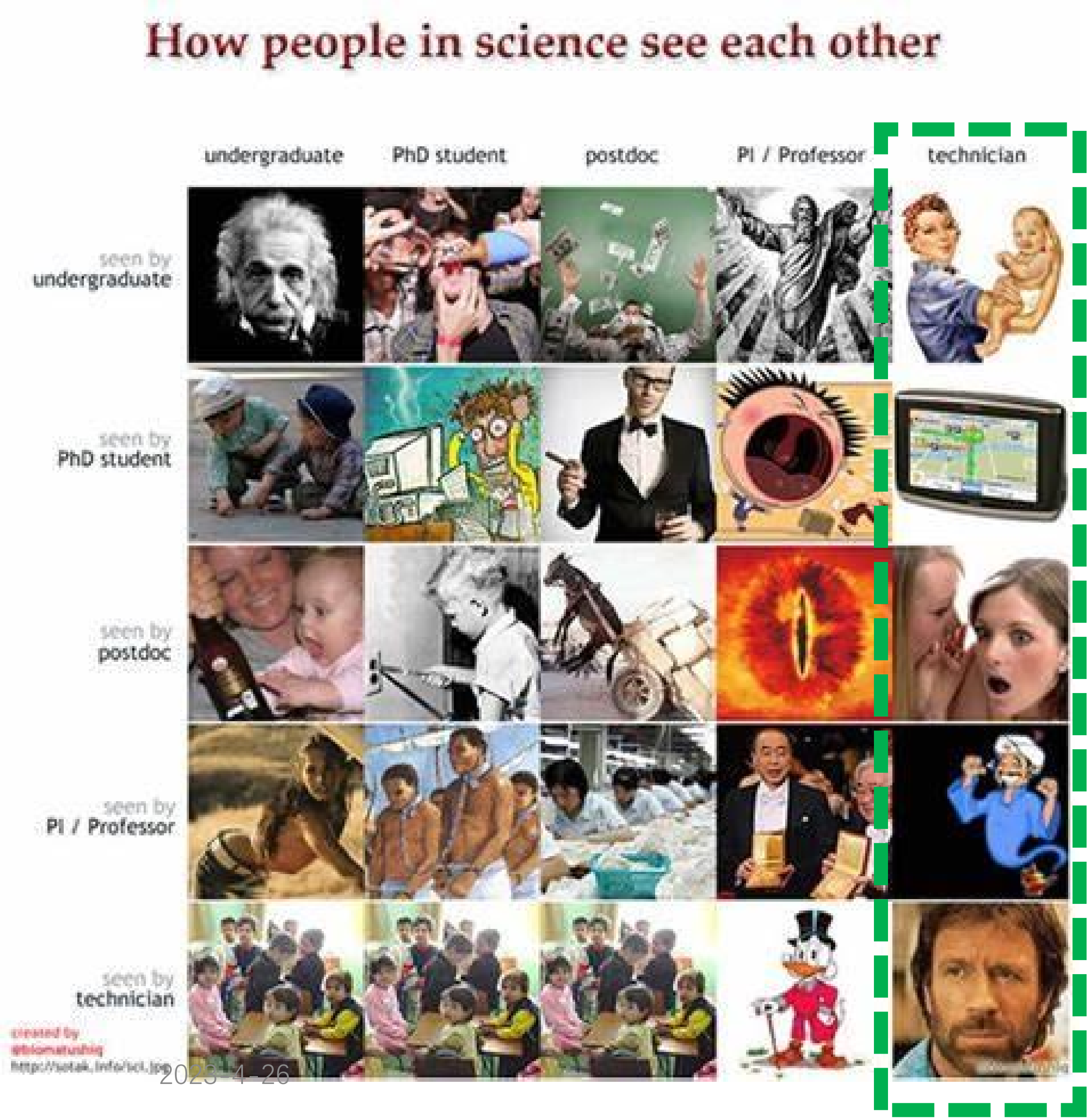
# 利用AI发现新天文现象

王锋 (GZU/PCL), 崔辰州 (NAOC), 陕欢源 (SHAO), 李楠 (NAOC), 季凯帆(YNAO), 刘元 (NAOC), 李江涛 (PMO), 李正阳(NIAOT), 刘慧根(NJU), 潘海武 (NAOC), 李珊珊(NAOC), 刘文波(TYUT), 孙瑞琪(TYUT), 吕佳蒙(TYUT) &More

**贾鹏 (TYUT)**

# 两个问题




天文学家对于AI的希望?

解决更多问题

更快

更容易用

更便宜…


# 三个层次

Level 0：将某些成熟算法套用到天文观测/分析/数据处理问题上

Level 1：针对天文数据，解决人力繁杂/难以解析描述的问题

Level 2：利用AI发现新现象和新问题（**需要冲刺的方向**）

# Outline

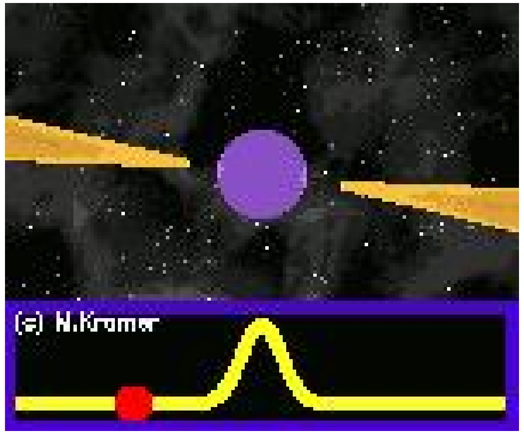
- **不确定、不常见和未知的天文现象**

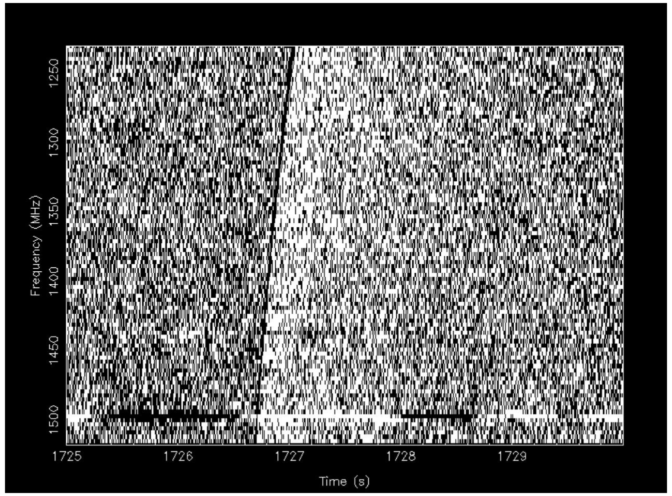
- 发现不常见天文现象需要的框架

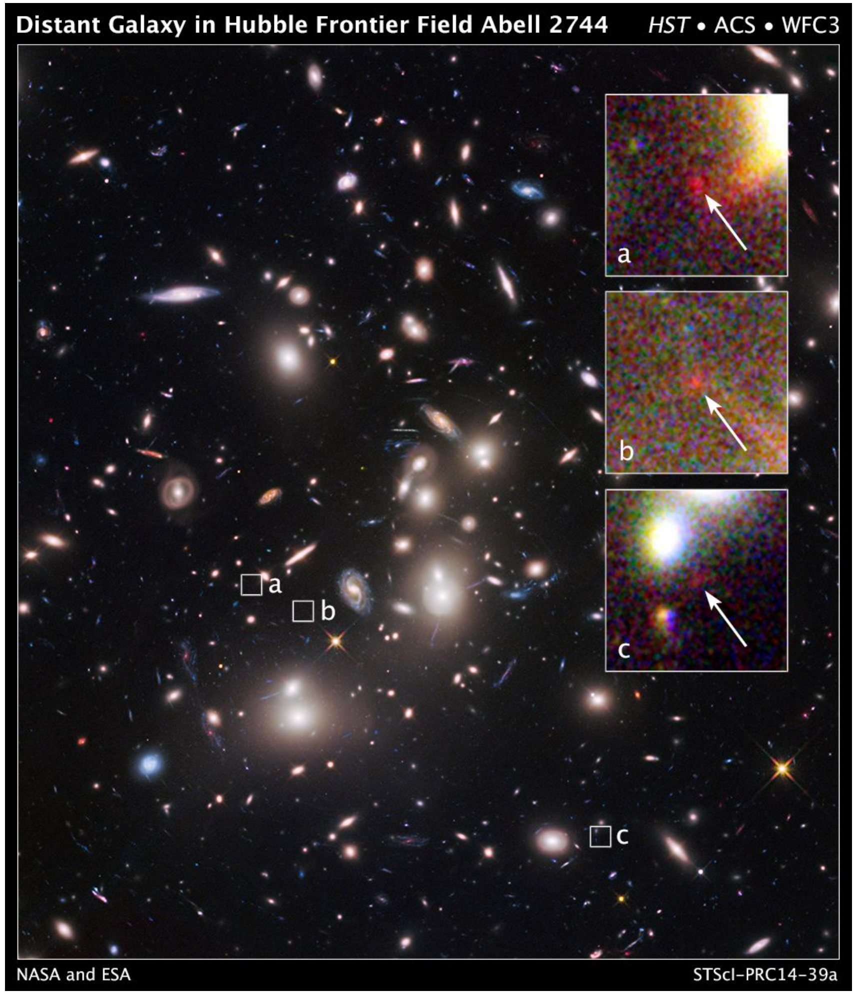
- 初步结果

# 不确定Uncertain、不常见Unusual 、未知Unknown

## 科学观测所能带来的:

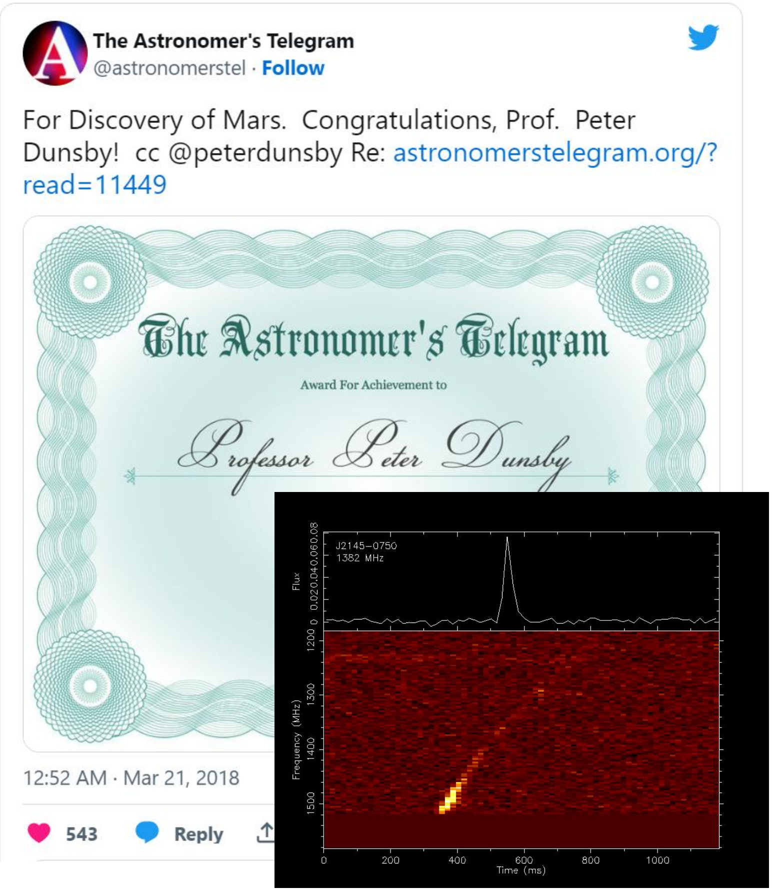



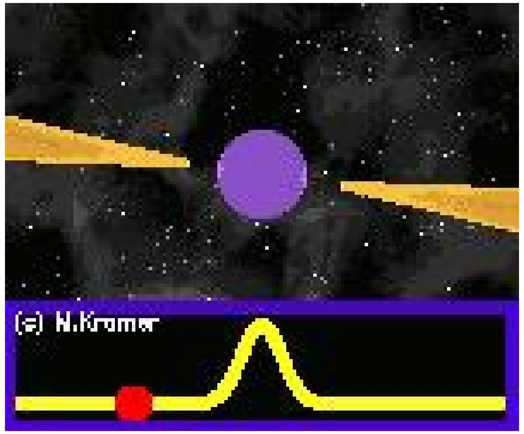

Uncertainty(Mars/Peryton) Unknown(Pulsar/Fast Radio Burst) Unusual (High-z galaxy/exoplanet/TDEs)

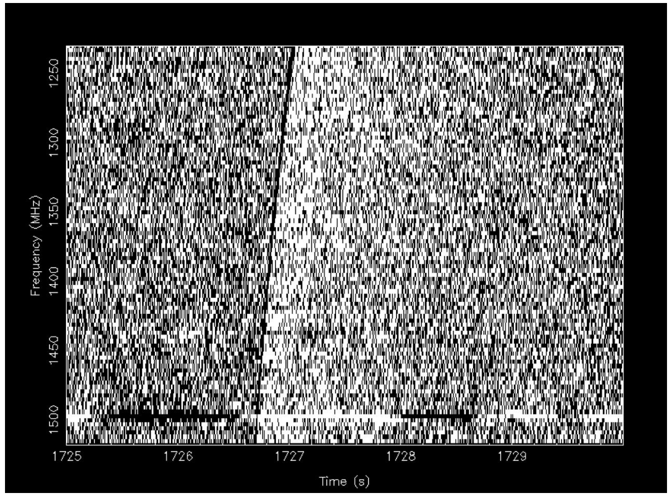
# 不确定Uncertain、不常见Unusual 、未知Unknown

## 科学观测所能带来的:

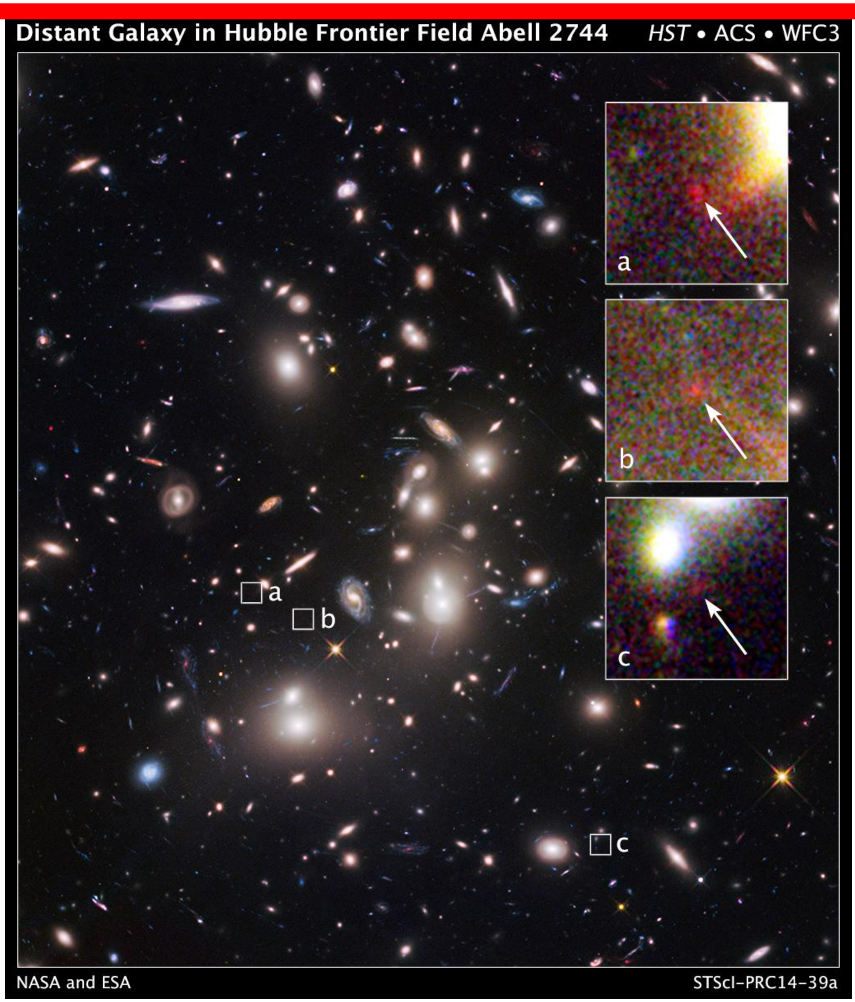



Uncertainty(Mars/Peryton)



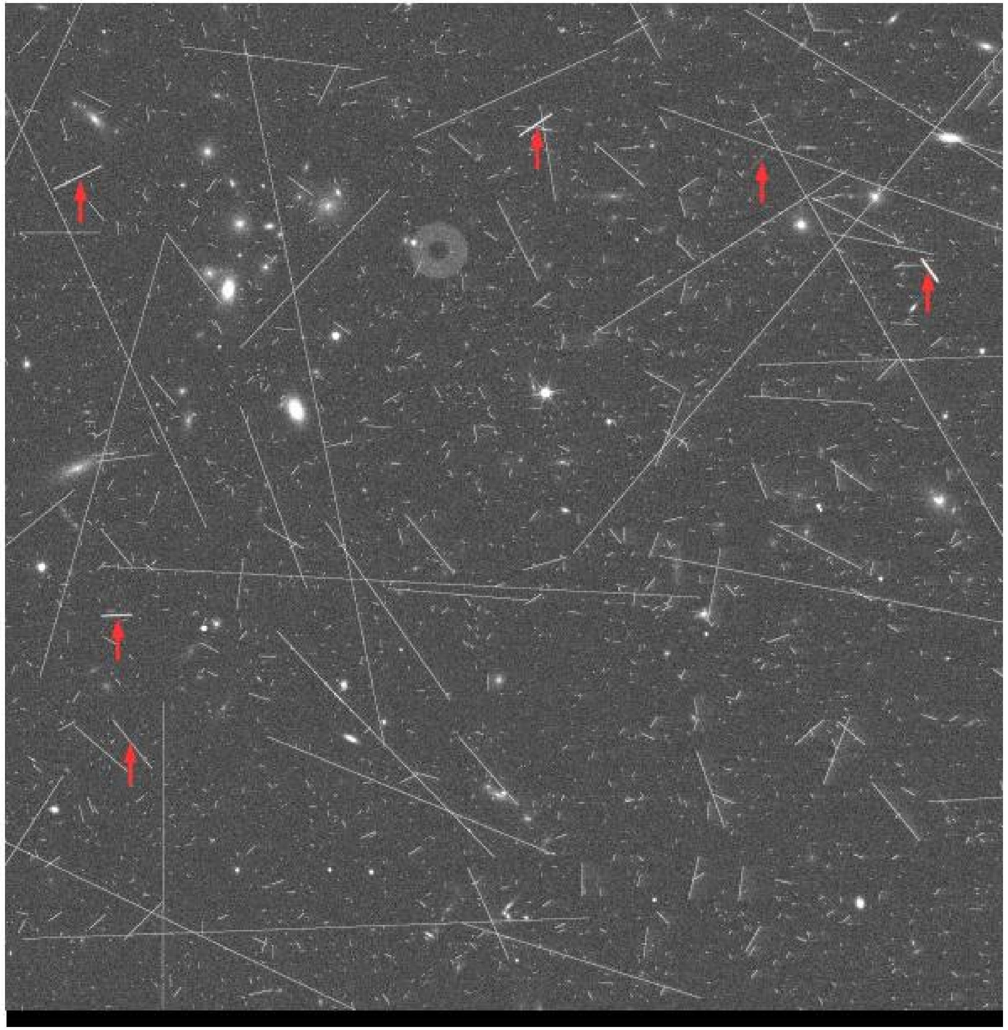

Unknown(Pulsar/Fast Radio Burst)




Unusual (High-z galaxy/exoplanet/TDEs)

# 不确定Uncertain、不常见Unusual 、未知Unknown

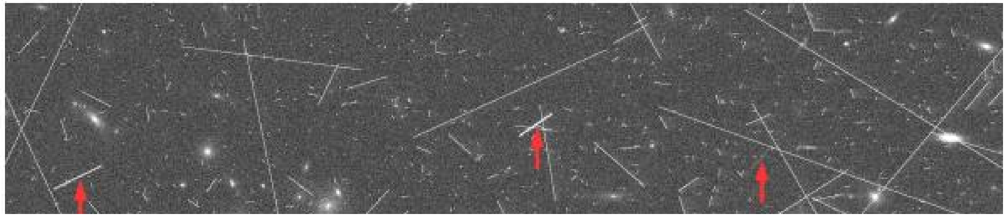
**Distinguishing between unknown, uncertain and unusual is HARD.**



Cosmic Ray
Asteroids
From: A&A 644, A35 (2020)

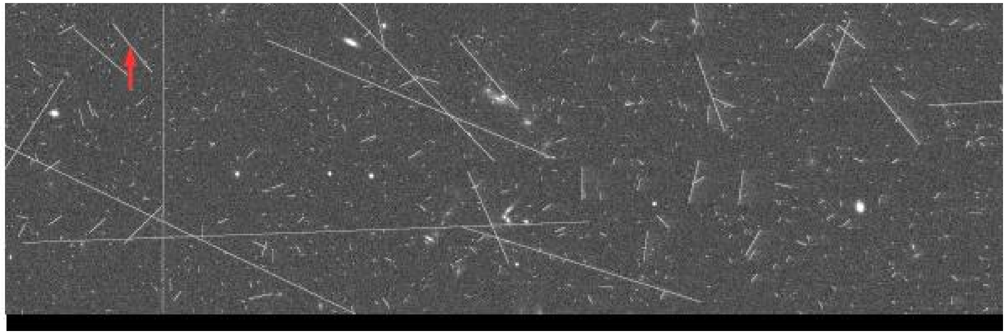
# 不确定Uncertain、不常见Unusual 、未知Unknown

**Distinguishing between unknown, uncertain and unusual is HARD.**

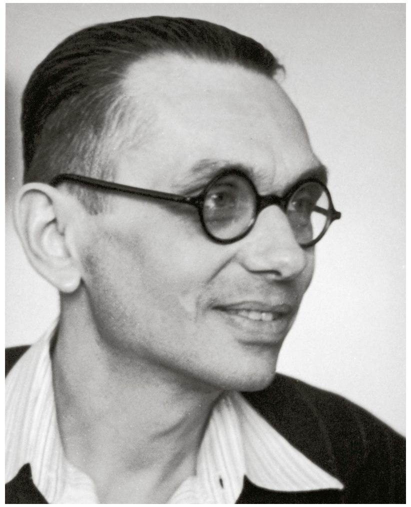


Could AI distinguish between uncertainty and unusual?

Probably not…



Cosmic Ray
Asteroids
From: A&A 644, A35 (2020)

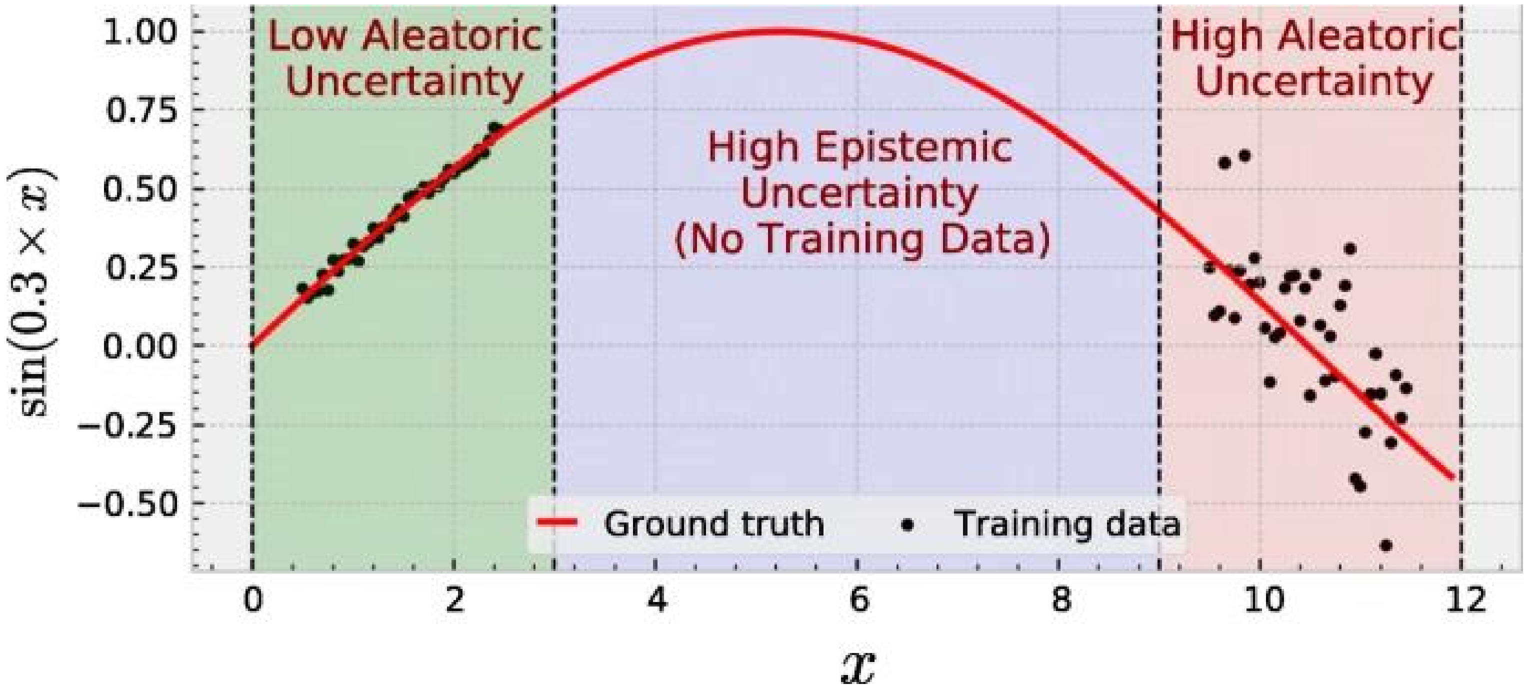


Gödel's second incompleteness theorem shows that, under general assumptions, this **canonical consistency statement Cons(F) will not be provable in F**.

# 不确定Uncertain、不常见Unusual 、未知Unknown

**Let’s talk about the philosophy:**

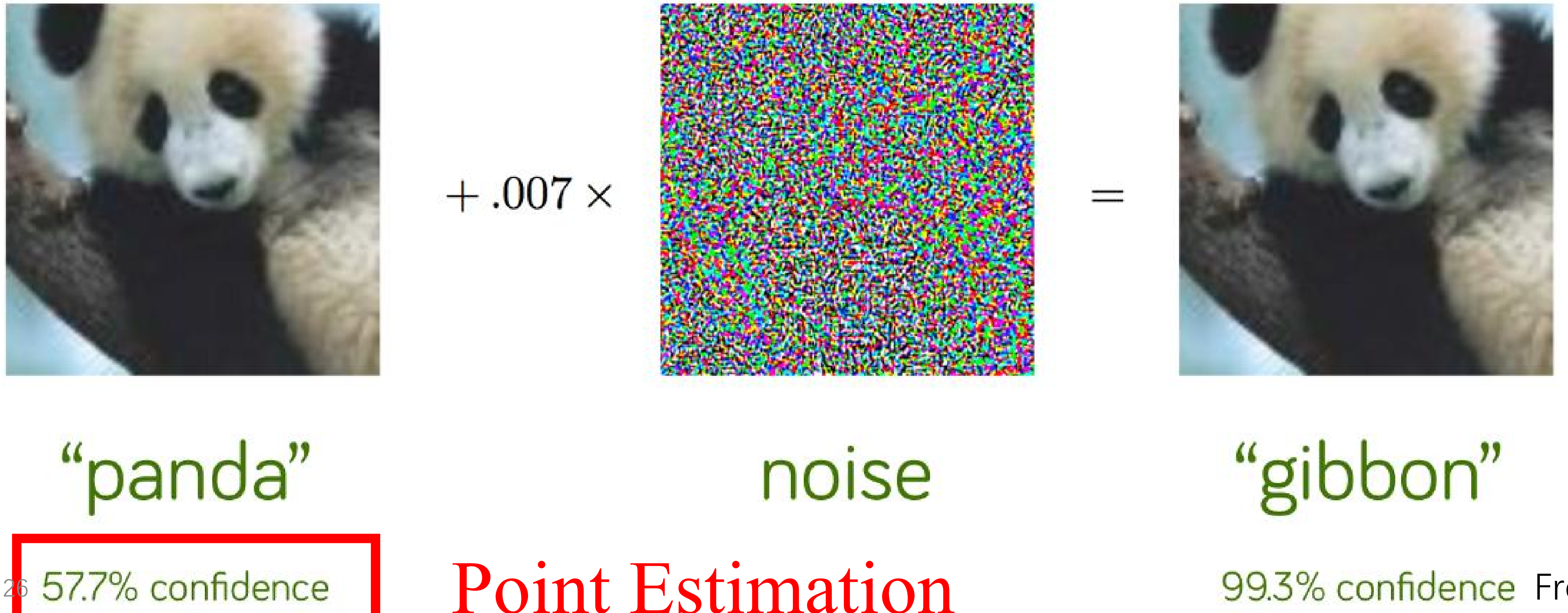
**the world is full of uncertainties and no one would be omnipotent…**




Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods
https://link.springer.com/article/10.1007/s10994-021-05946-3
Exploiting epistemic uncertainty of the deep learning models to generate adversarial samples
https://link.springer.com/article/10.1007/s11042-022-12132-7

# 不确定Uncertain、不常见Unusual 、未知Unknown

Devil exists in the TRAINING SET …
PSF/Noises/Temporal Noises…




Point Estimation

From NVIDIA

# Outlines

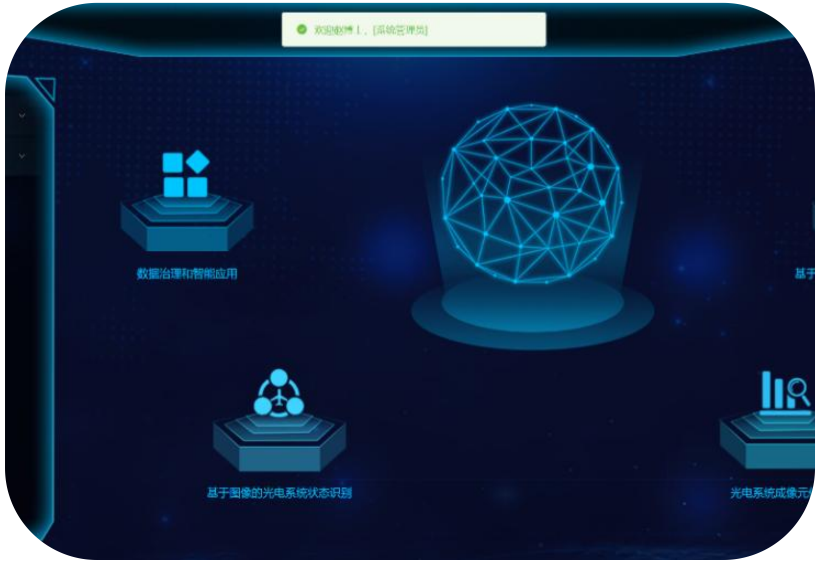
太原理工大學
TAIYUAN UNIVERSITY OF TECHNOLOGY

- Uncertain, Unusual and Unknown
- **Pipeline for Discovery of Unusual Astronomical Targets**
- Initial Results

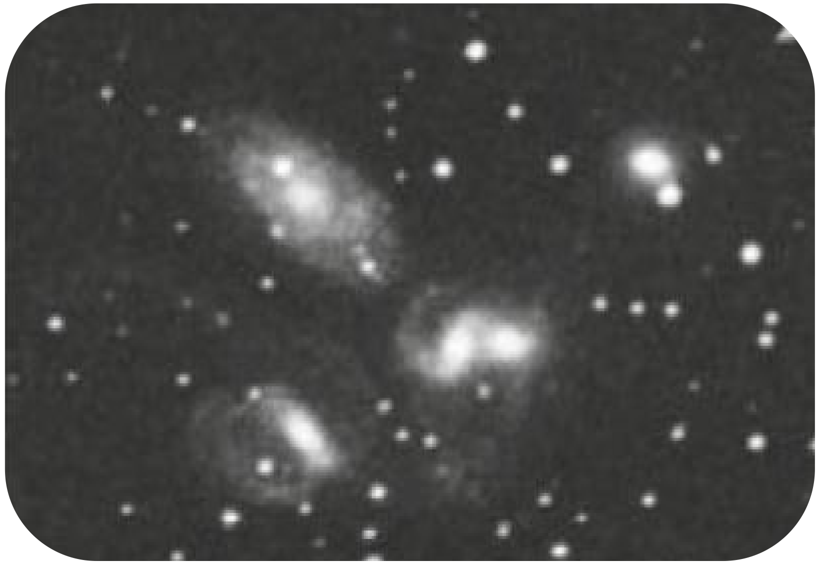
# Pipeline for Discovery of Unusual Astronomical Targets

TAIYUAN UNIVERSITY OF TECHNOLOGY

## A LARGE Bayes Model for Astronomical Communities



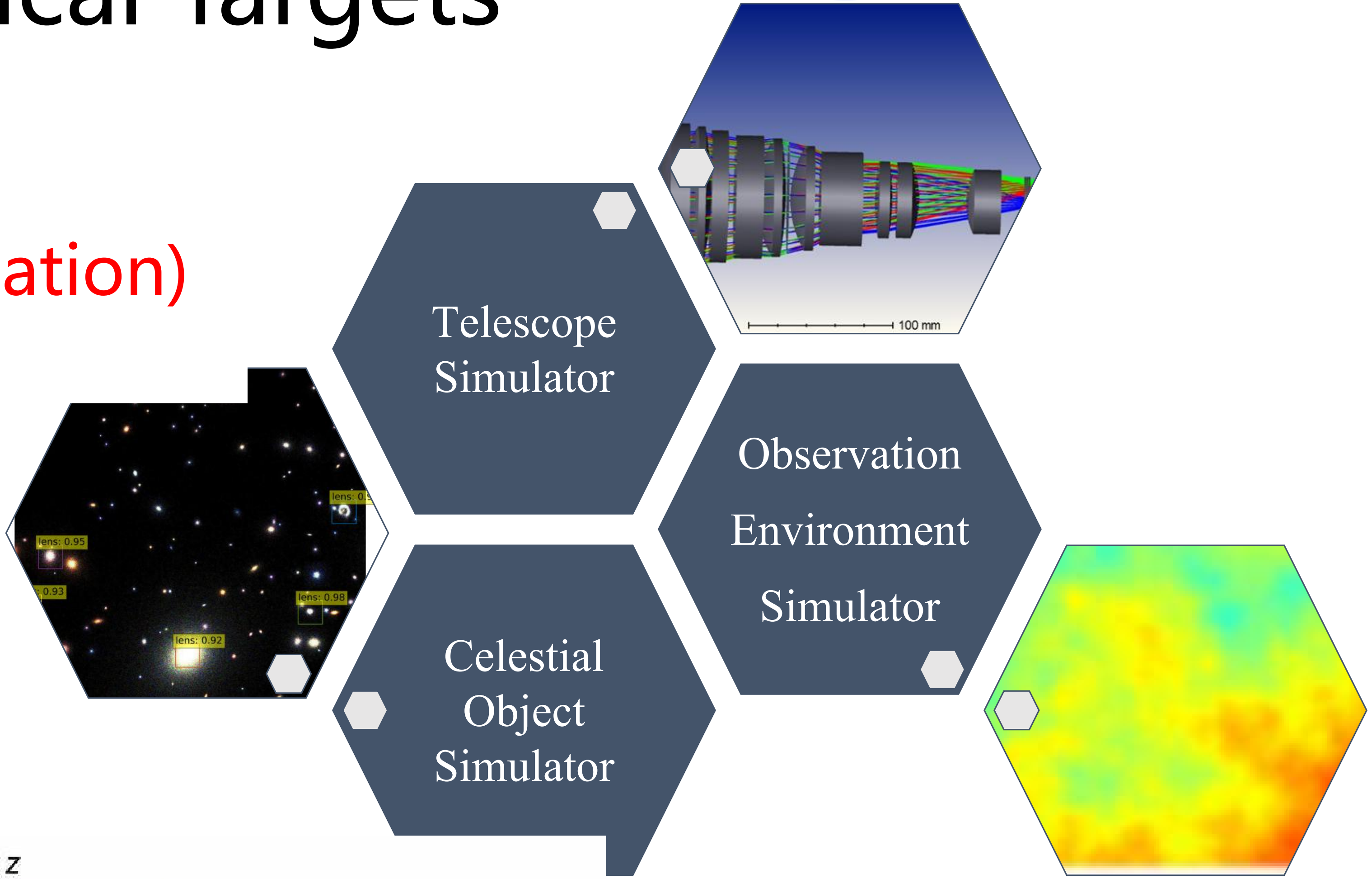

**Digital Twin of Observations**



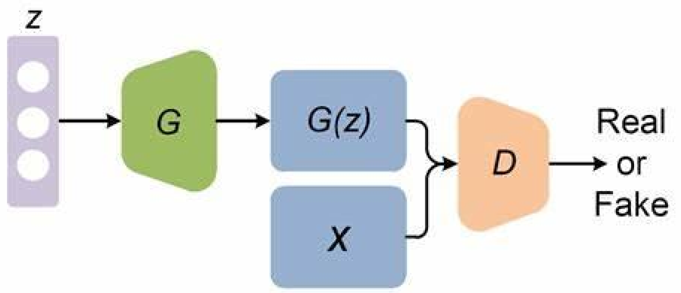
**Prior Information from Large Model**

# Pipeline for Discovery of Unusual Astronomical Targets

Embed Prior Information
(Widest idea)
(Theoretical & Instrumentation)
into simulation data


**Digital Twin of Observations**



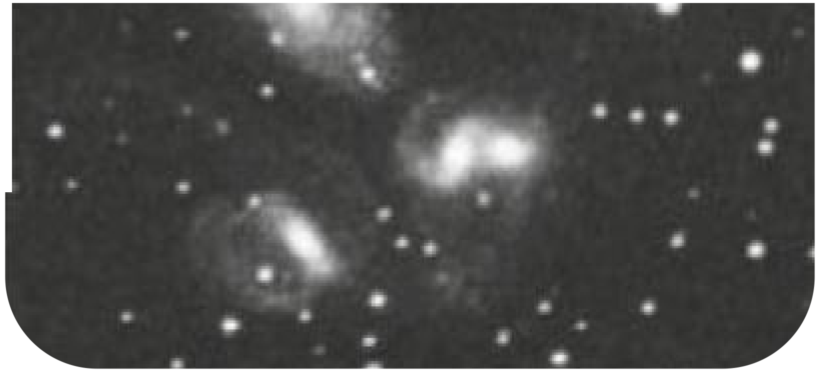
# Pipeline for Discovery of Unusual Astronomical Targets


Data from Citizen Scientists

Prior Observation Results

Data without Label


Embed Prior Information (Observation Data) into pre-trained models

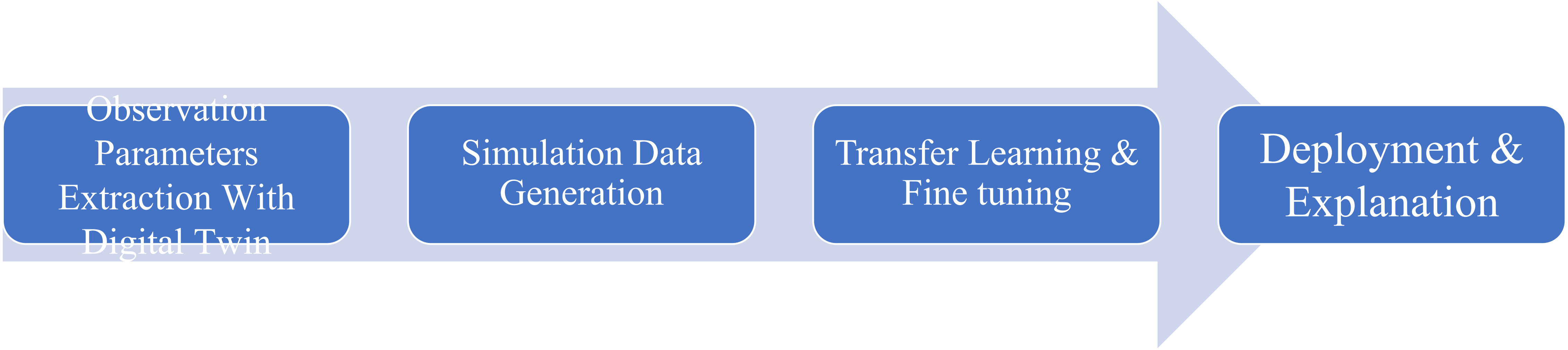


**Prior Information from LARGE Model**

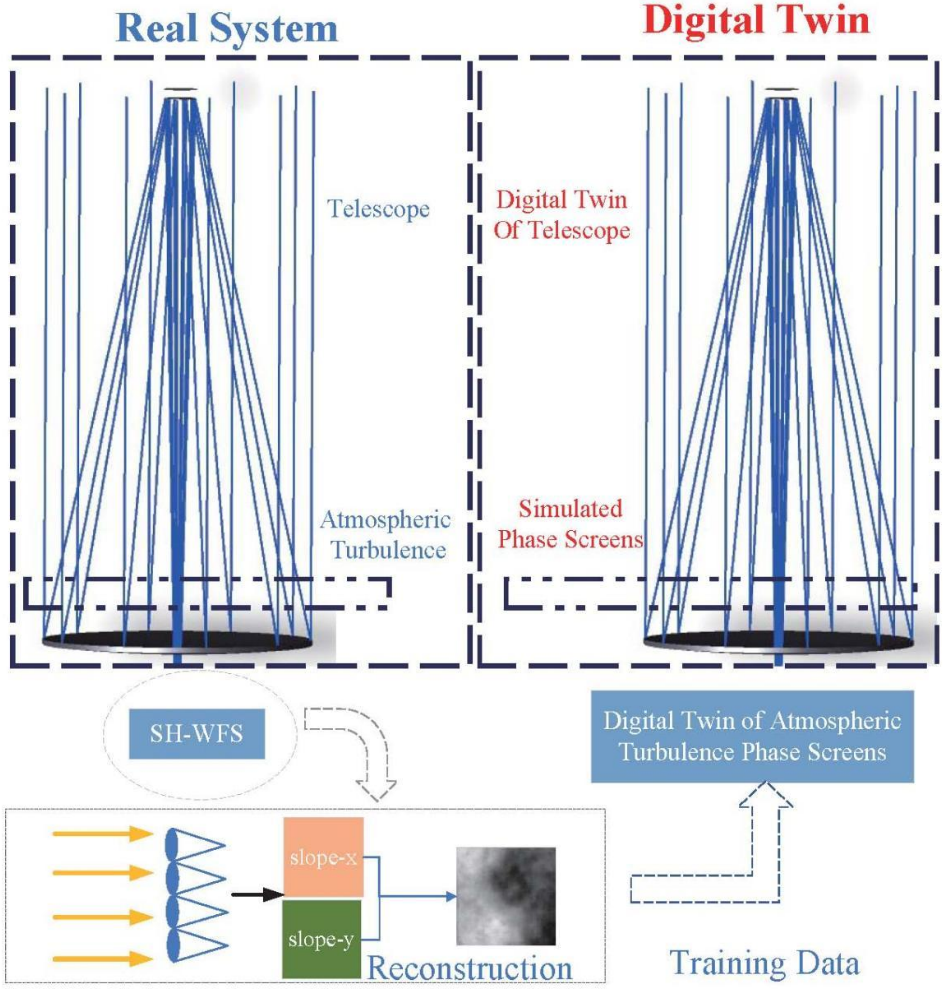
# Pipeline for Discovery of Unusual Astronomical Targets

TAIYUAN UNIVERSITY OF TECHNOLOGY

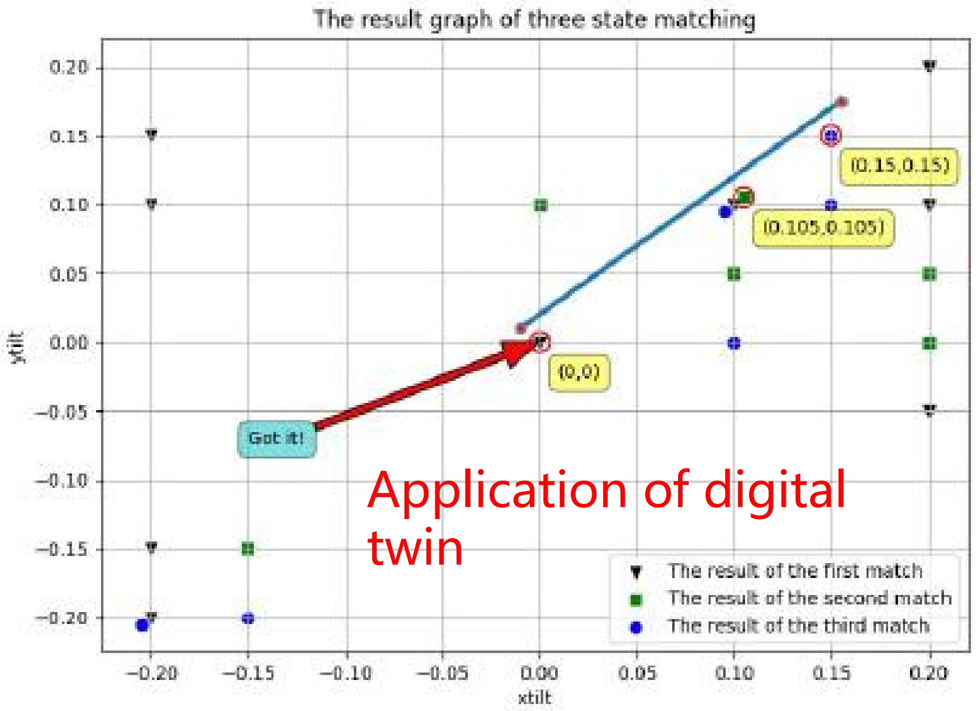
## A LARGE Bayes Model for Astronomical Communities



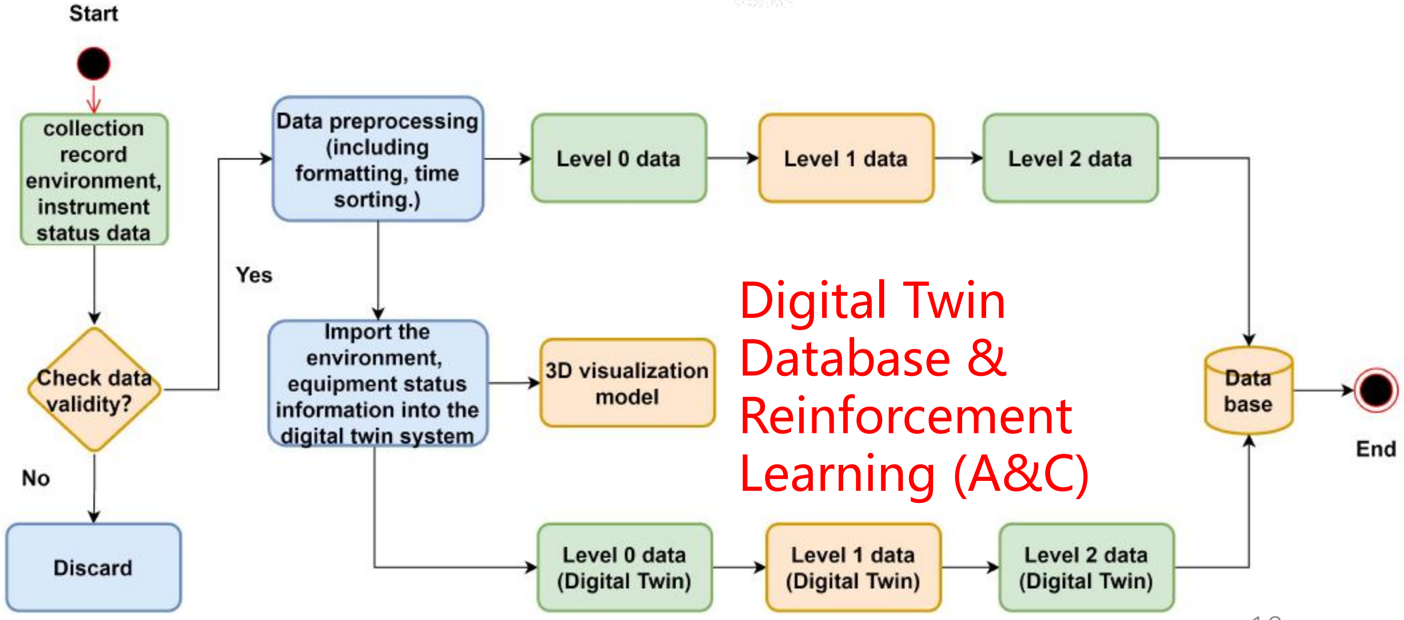

Observation Parameters Extraction With Digital Twin:
Integrate data from telemetry data from telescopes, instruments and all




Principle of digital twin




Application of digital twin




Digital Twin Database & Reinforcement Learning (A&C)

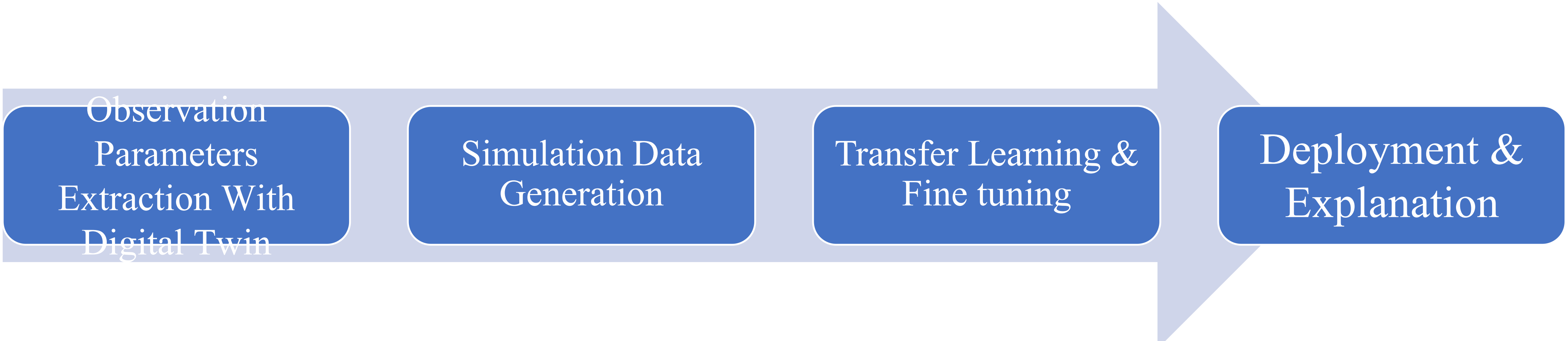
From Peng et.al. OE 2022

From Zhan et.al. SPIE 2022

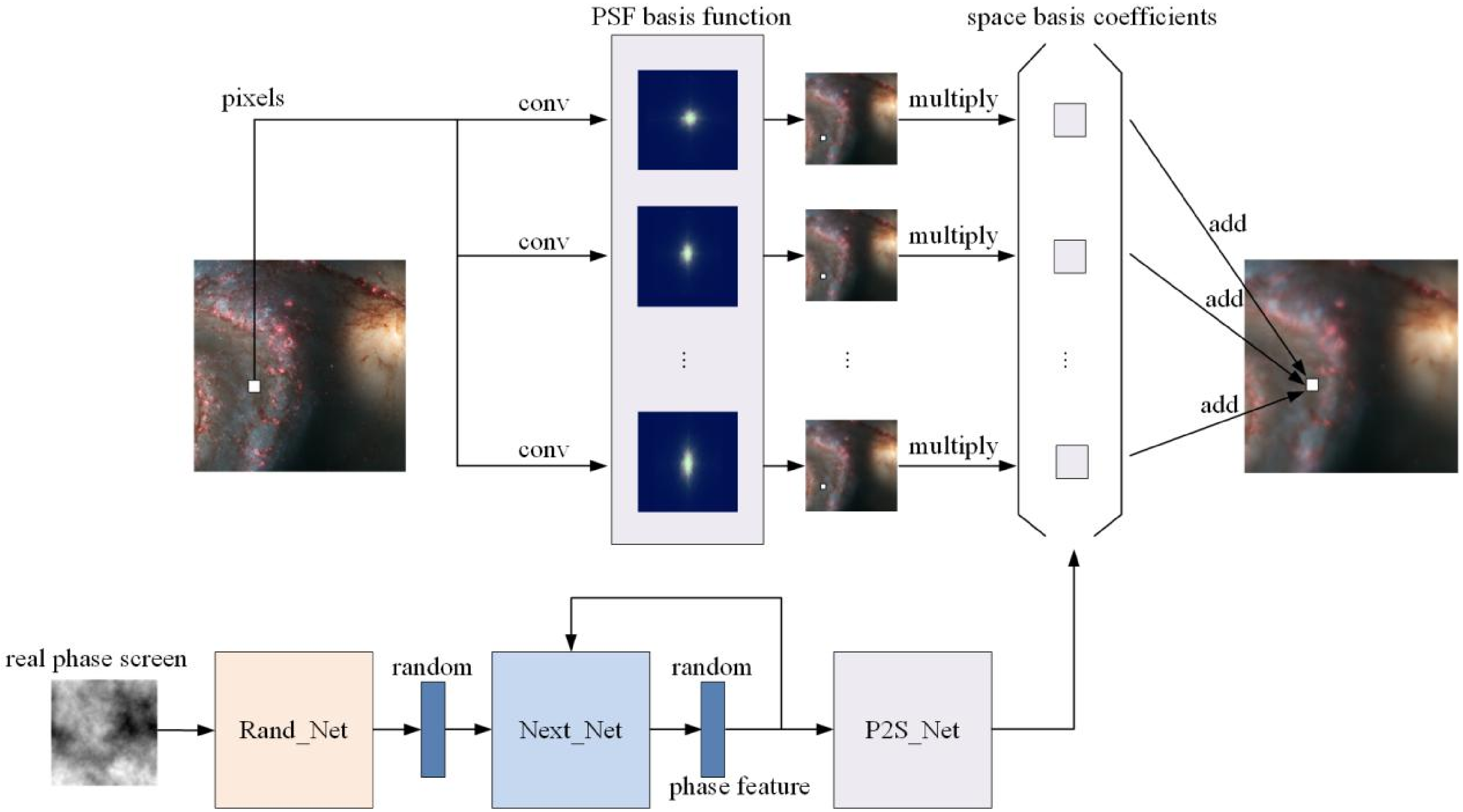
# Pipeline for Discovery of Unusual Astronomical Targets

太原理工大学
TAIYUAN UNIVERSITY OF TECHNOLOGY

## A LARGE Bayes Model for Astronomical Communities




Simulation Data Generation:
Integrate prior information of astronomical targets twin (light curve, morphology of galaxies, energy distribution) with digital twin (PSF, environment)




Realistic
Speed is important (1000 times faster than ordinary phase-> PSF -> Convolution method)

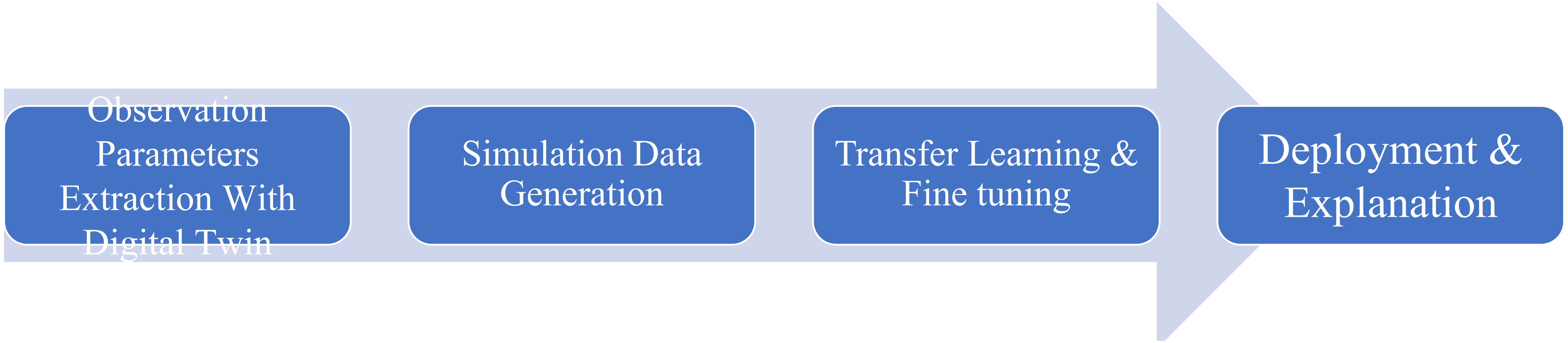


From Zhang et.al. SPIE 2022

# Pipeline for Discovery of Unusual Astronomical Targets

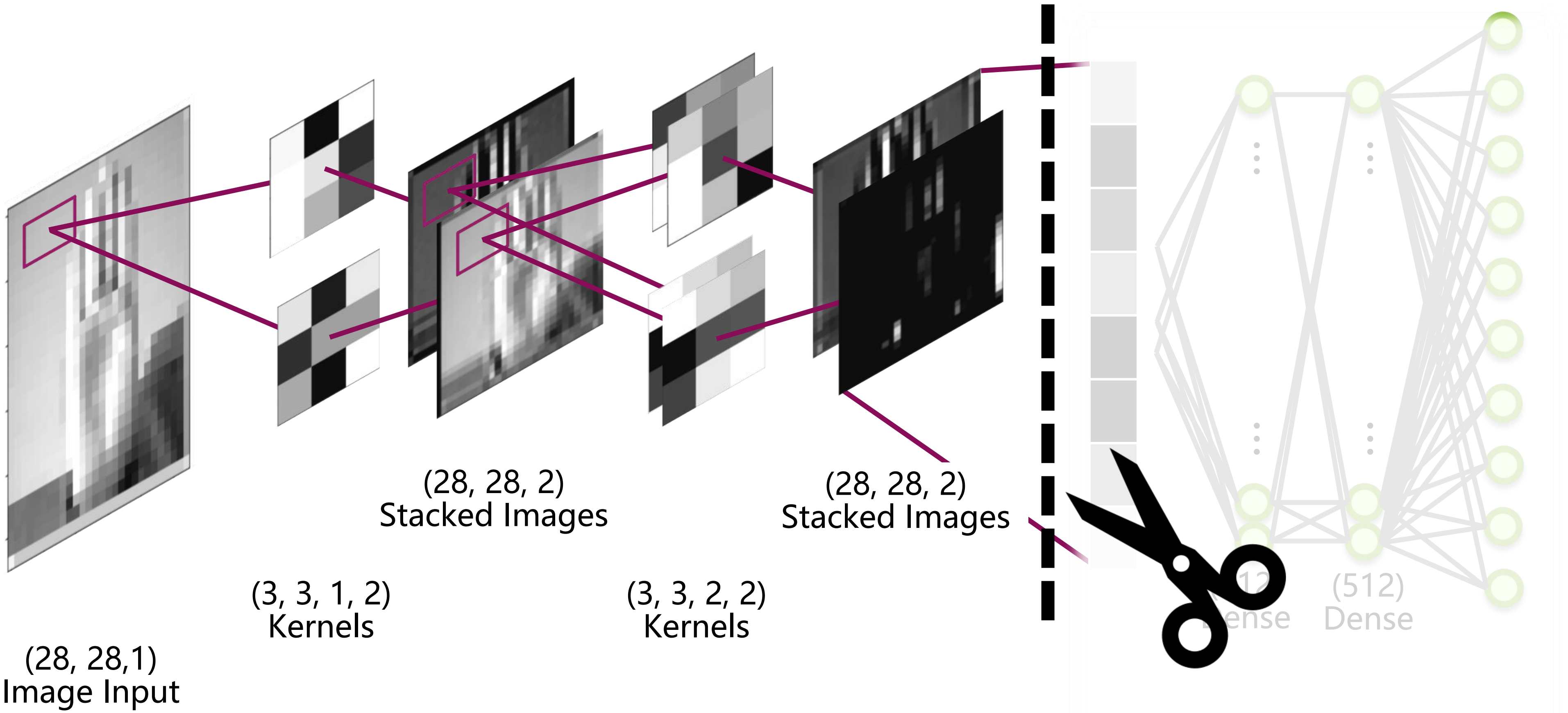
太原理工大学
TAIYUAN UNIVERSITY OF TECHNOLOGY

## A LARGE Bayes Model for Astronomical Communities




Transfer Learning and fine tuning:
Load big model pre-trained with real observation data
Then transfer learning/fine tuning to real observation data.

# TRANSFER LEARNING




Could be Transformer/CNN/RNN and trained with supervised learning
/unsupervised learning (masking etc...)
TRAINABLE/UNTRAINABLE

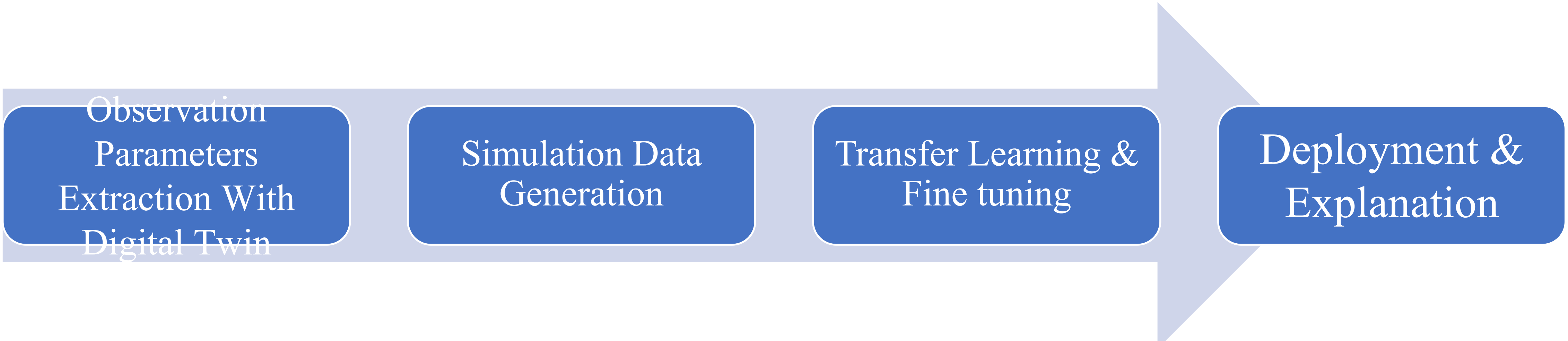
Bayesian Neural Network

From NVIDIA

# Pipeline for Discovery of Unusual Astronomical Targets

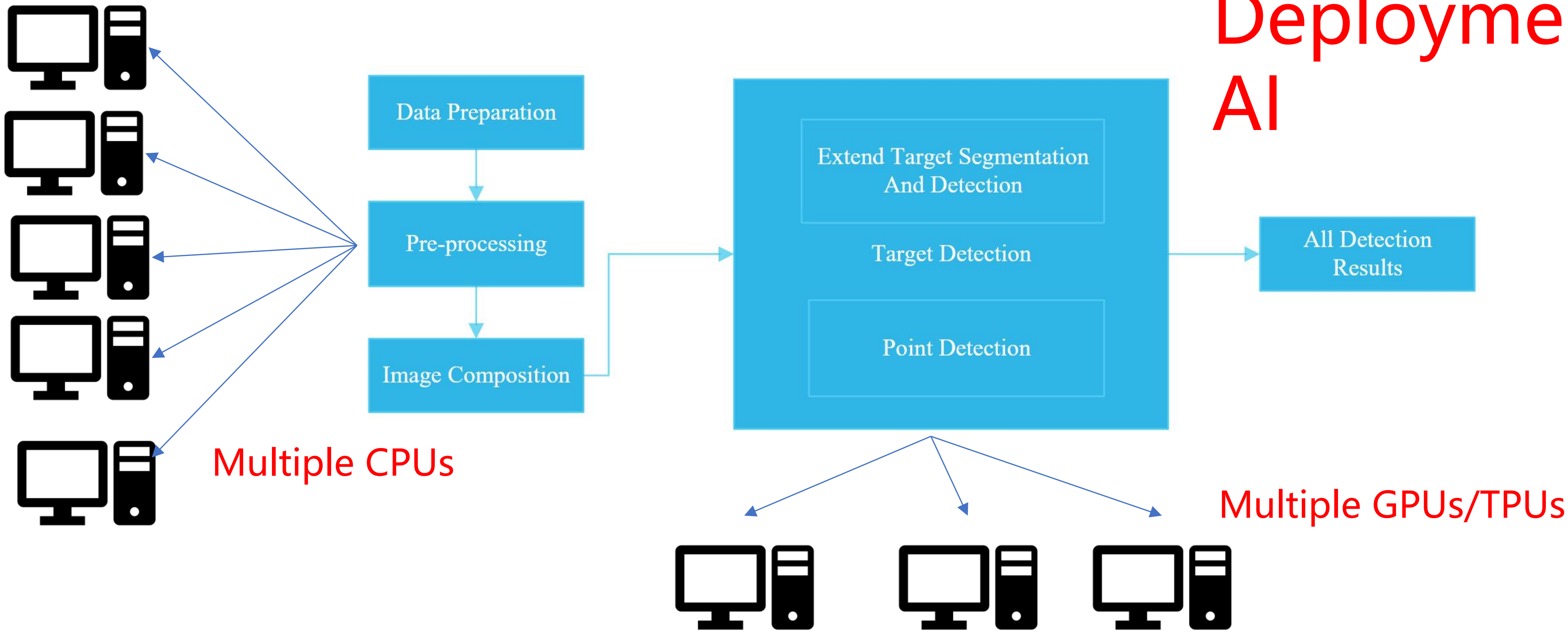
太原理工大学
TAIYUAN UNIVERSITY OF TECHNOLOGY

## A LARGE Bayes Model for Astronomical Communities




Deployment & Explanation:
Deployment of pipeline is important (Not a single toy model)
Explanation could help us to better investigate performance and analyze errors

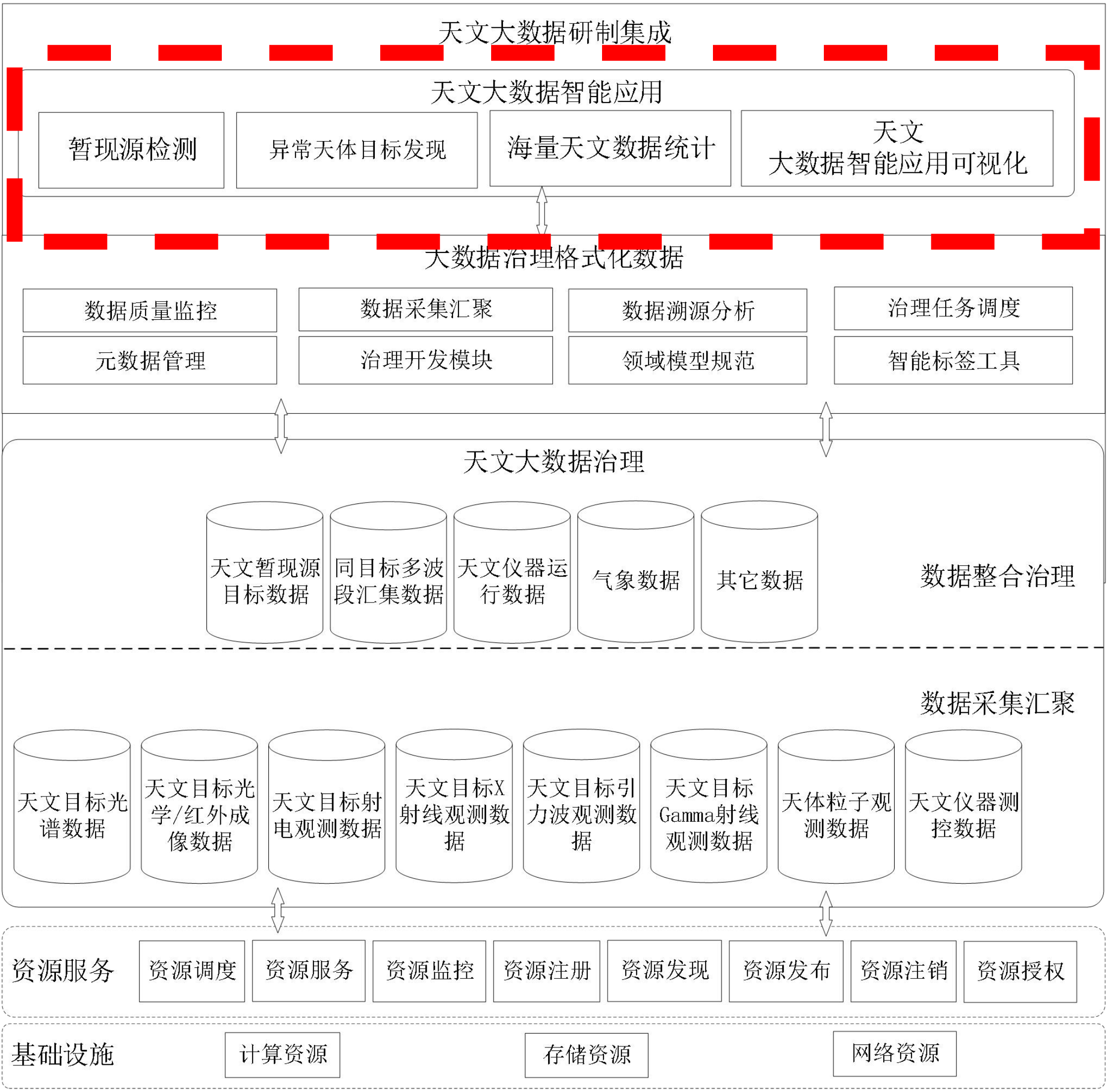
# Deployment of AI



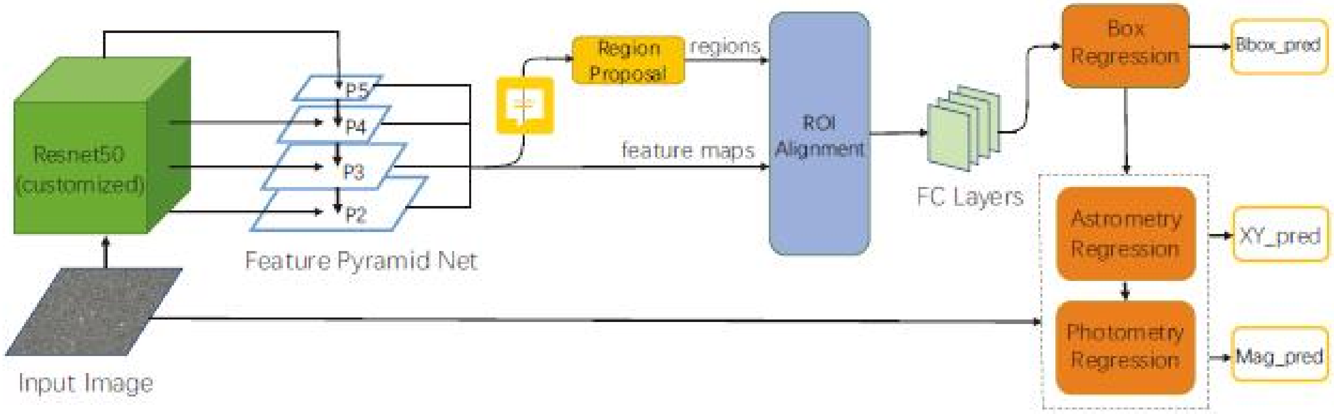

**Multiple CPUs/Clusters/GPUs + AUTOML**
**Computation workload control**
Technology Stack with:
Spark/Hadoop/Redis/Kafka/Openmpi/Docker/Kubernetes
Software containers (NGC, TesorFlow hub, PyTorch hub)/Triton....

**Speed & Money are always important**
Technology Stack with:
(Pruning/Quantization/Distillation) TensorRT/Neural Network Architecture Search/Work flow Control
Detection -> Light Curve Classification
Spectrum Cross-Check
Segmentation

# Deployment of AI




Architecture Design of Data System




Architecture Design of Neural Network
(Separate Design for Different Parts)

AND

Explanation Architecture for Different Parts
(Features/Separate Explainable Abilities)

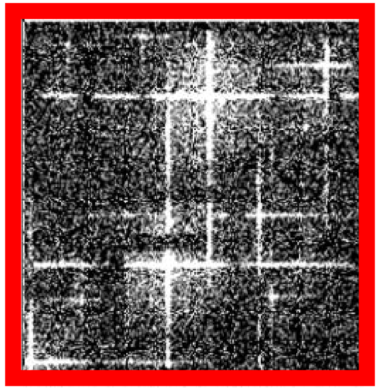
From Jia et.al. 2020

# Outlines

- Uncertain, Unusual and Unknown
- Pipeline for Discovery of Unusual Astronomical Targets
- **Initial Results**

# Initial Applications





X-ray: EP Detection

Energy

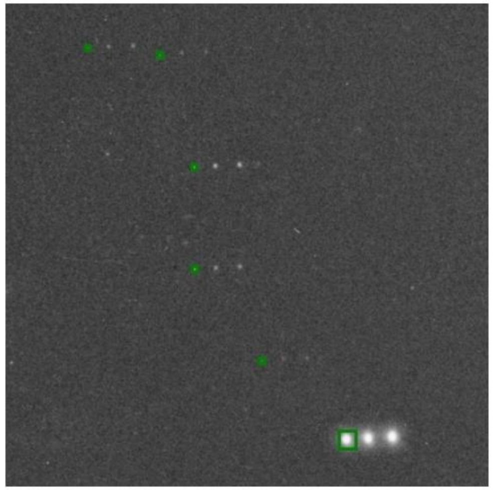
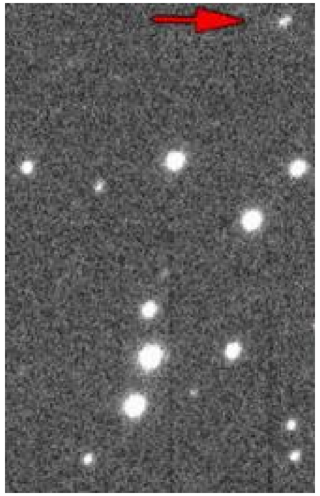
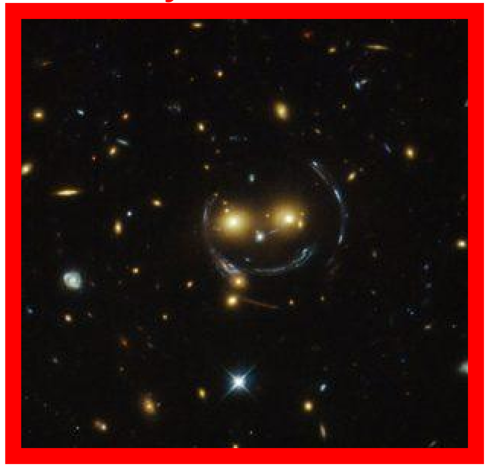


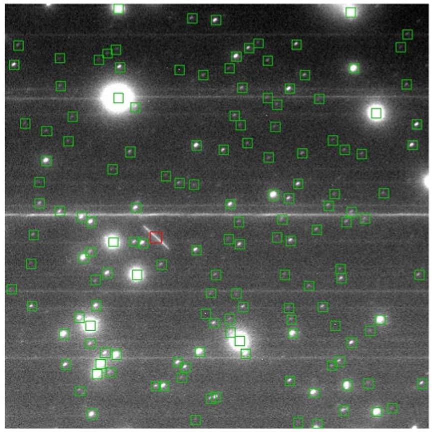
Photo-Plates ~ 100 yrs



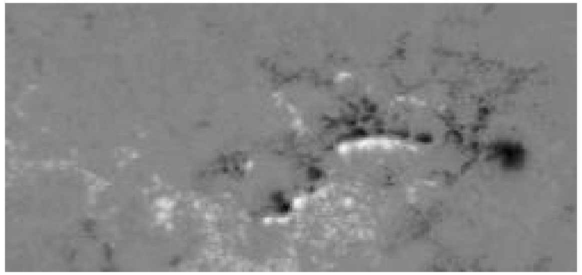
NEOs/Exo-Planets ~ Days



Optical: GC Strong Lensing

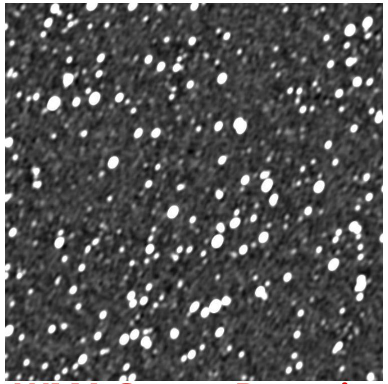


Space Debris ~ Hours



Solar Flare ~ minutes

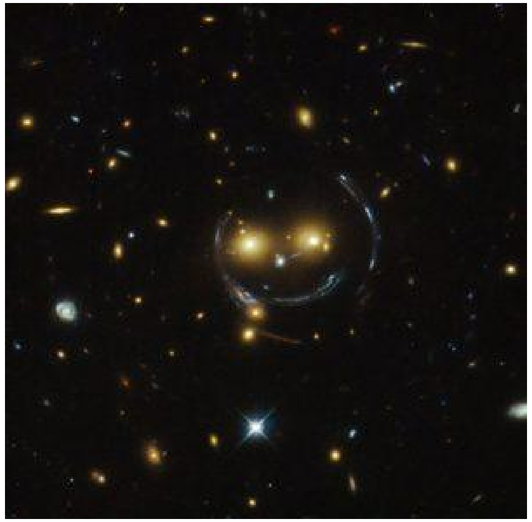
Time

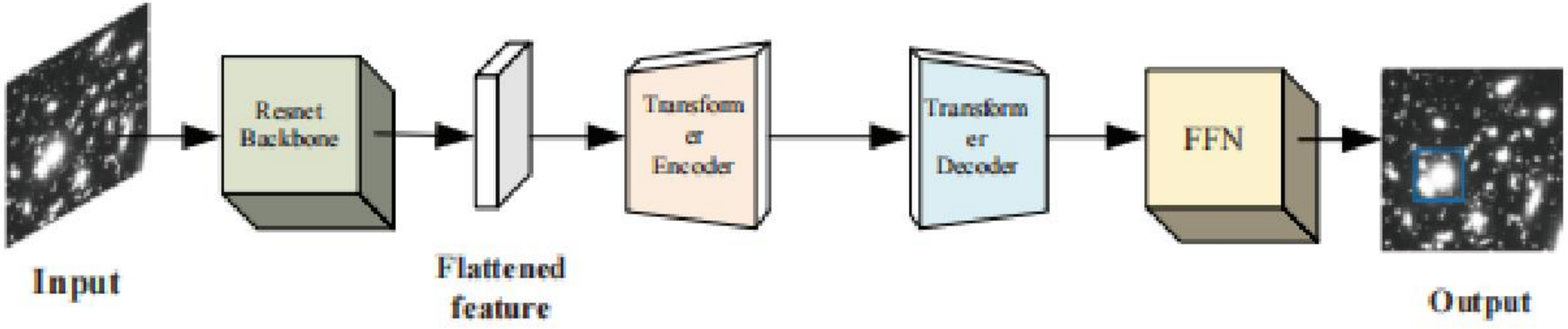


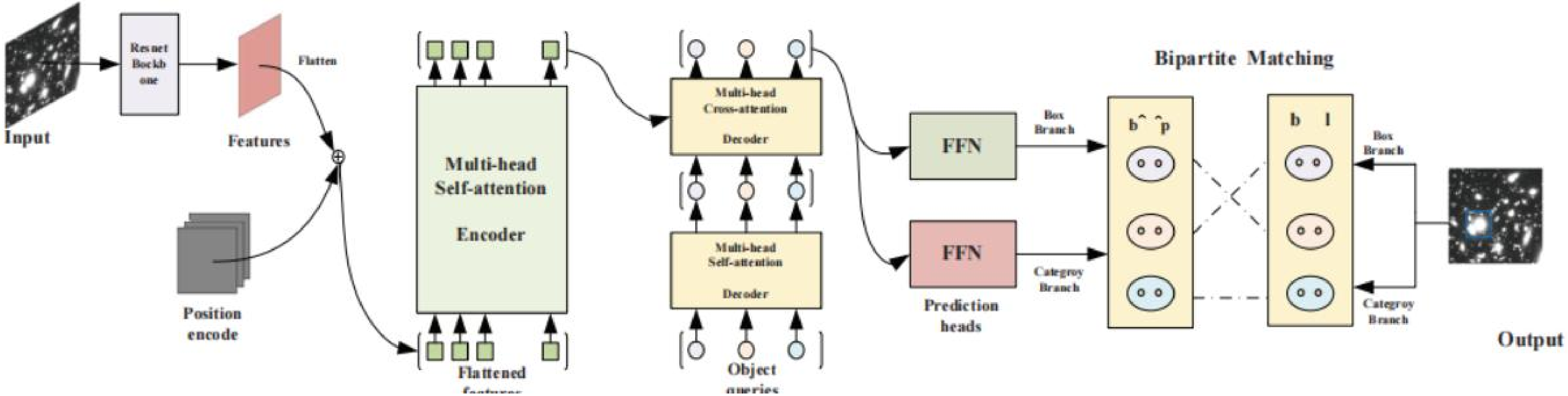
Radio: WMA Source Detection

# Initial Results 1 – Optical Images

**Trained with Simulated Data (Realistic Strong Lensing System)**







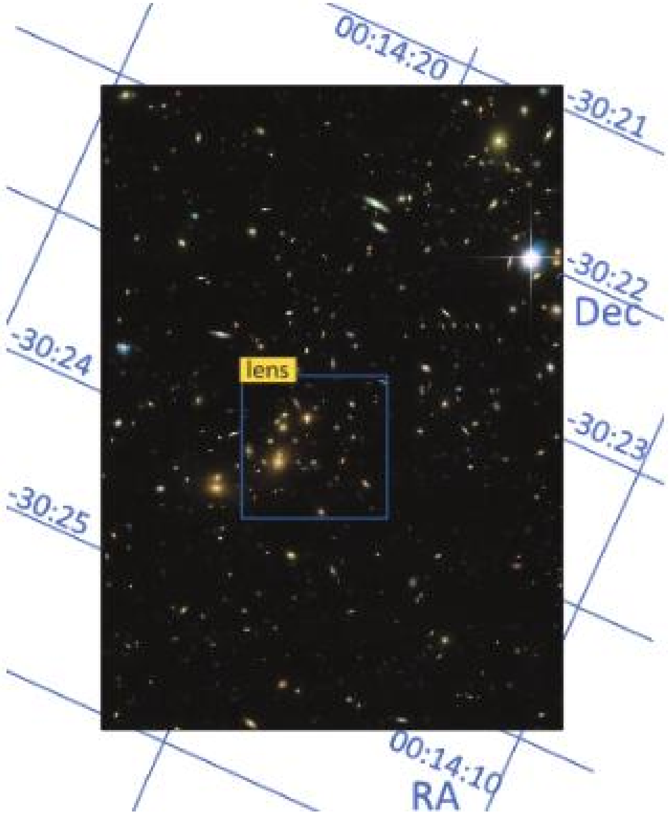


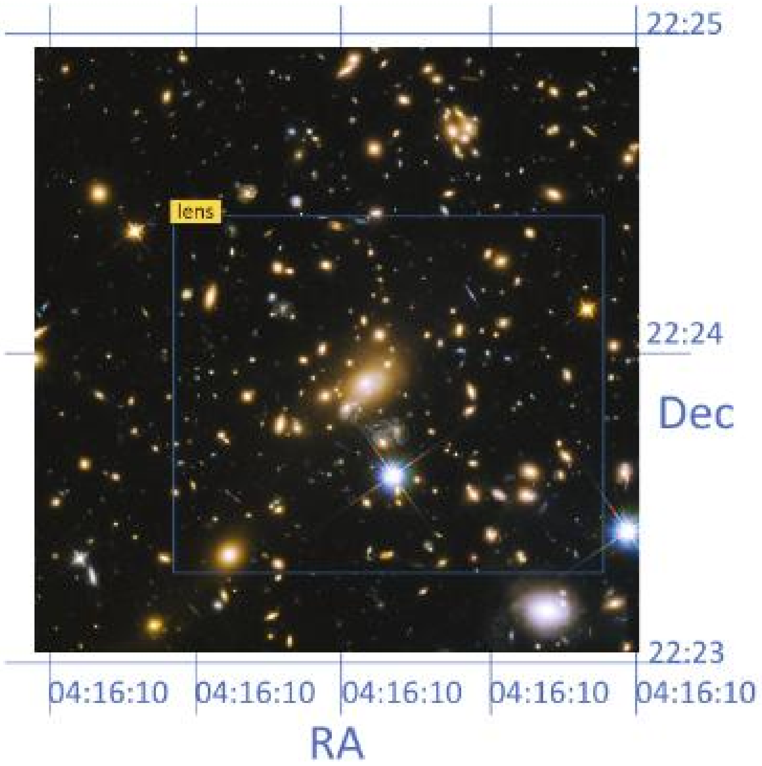

From Jia et.al. 2022

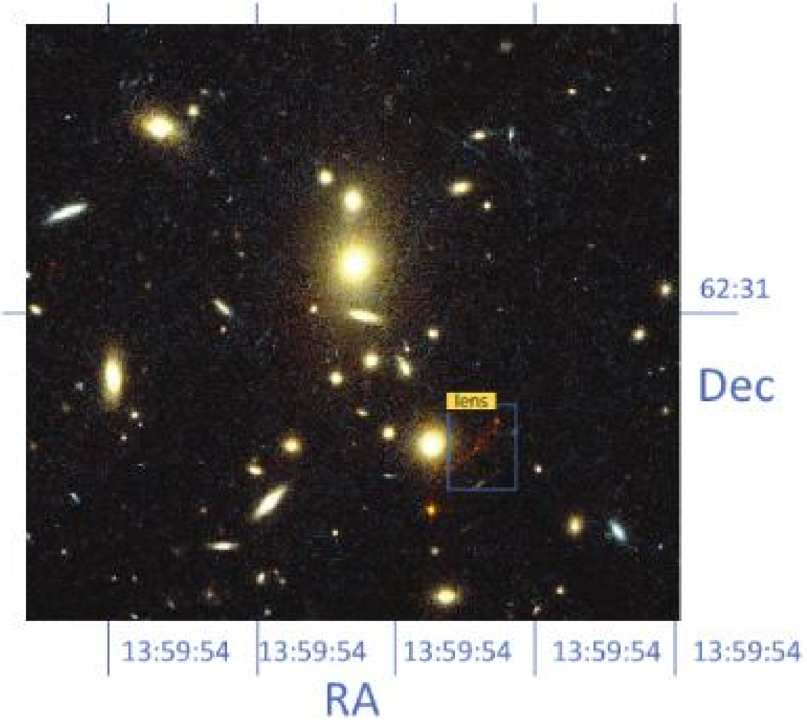
# Initial Results 1 – Optical Images



**Direct Application to Real Observation Data**



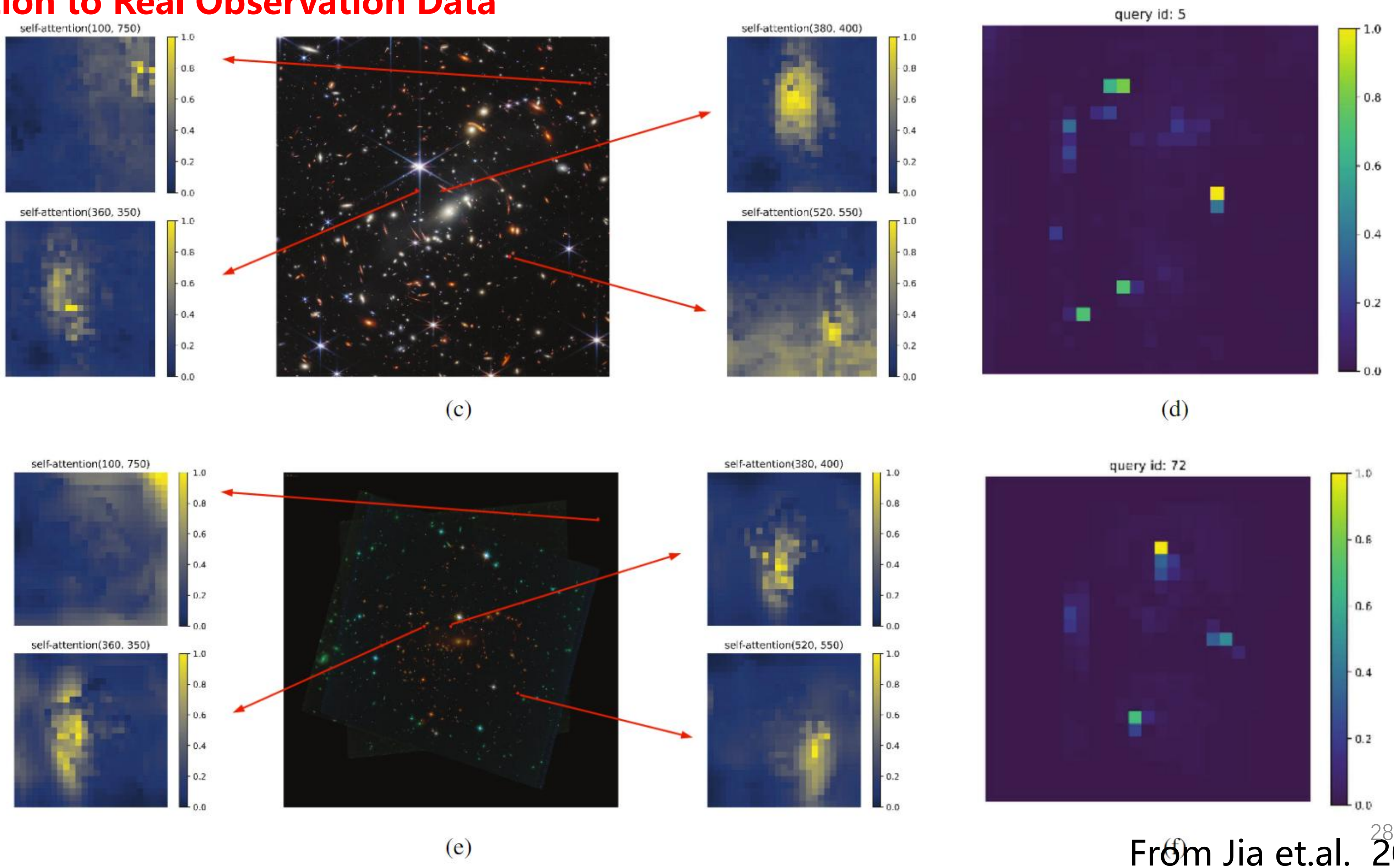




From Jia et.al. 2022

# Initial Results 1 – Optical Images

太原理工大學
TAIYUAN UNIVERSITY OF TECHNOLOGY

**Direct Application to Real Observation Data**

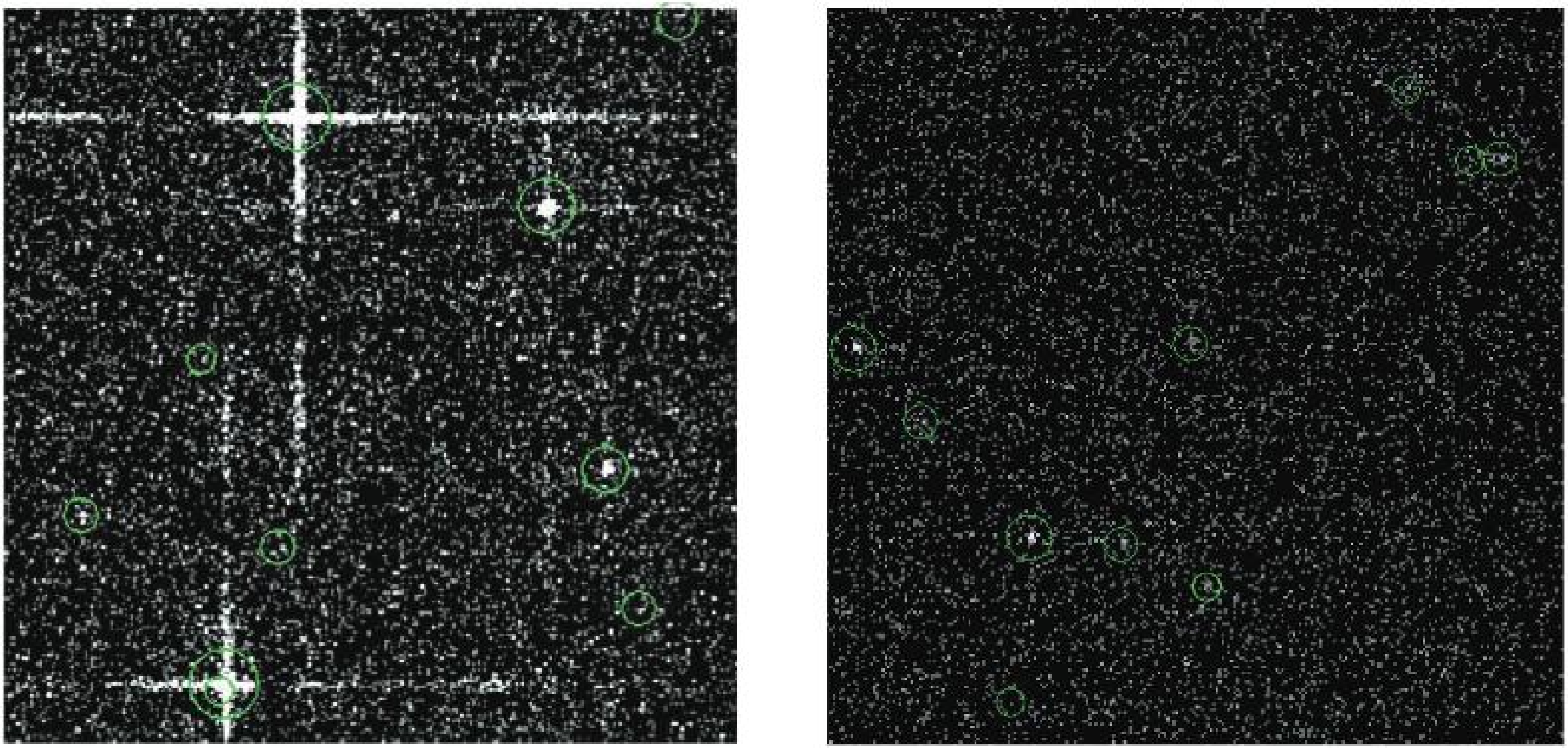



From Jia et.al. 2022

# Initial Results 2 – X-Ray Images

TAIYUAN UNIVERSITY OF TECHNOLOGY

**Trained with Simulated Data (X-Ray Lobster Eye Data)**

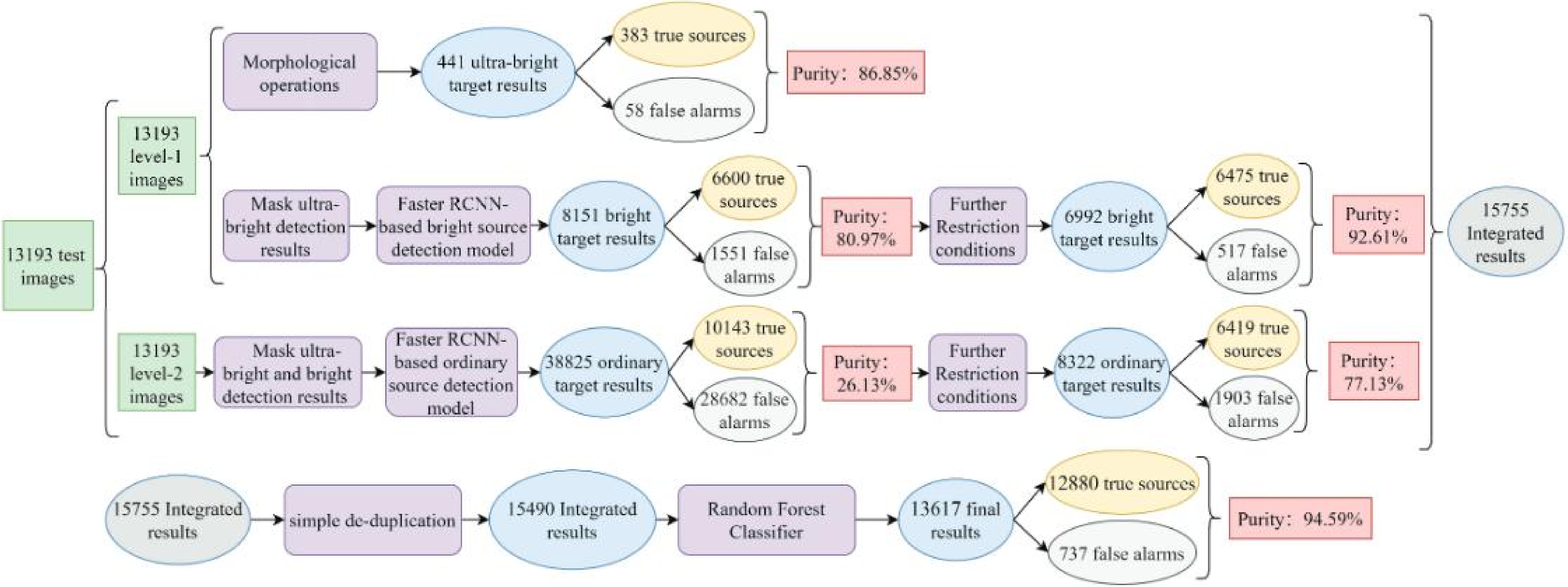


From Jia et.al. 2023

# Initial Results 2 – X-Ray Images

太原理工大学 TAIYUAN UNIVERSITY OF TECHNOLOGY

**Trained with Simulated Data (X-Ray Lobster Eye Data)**





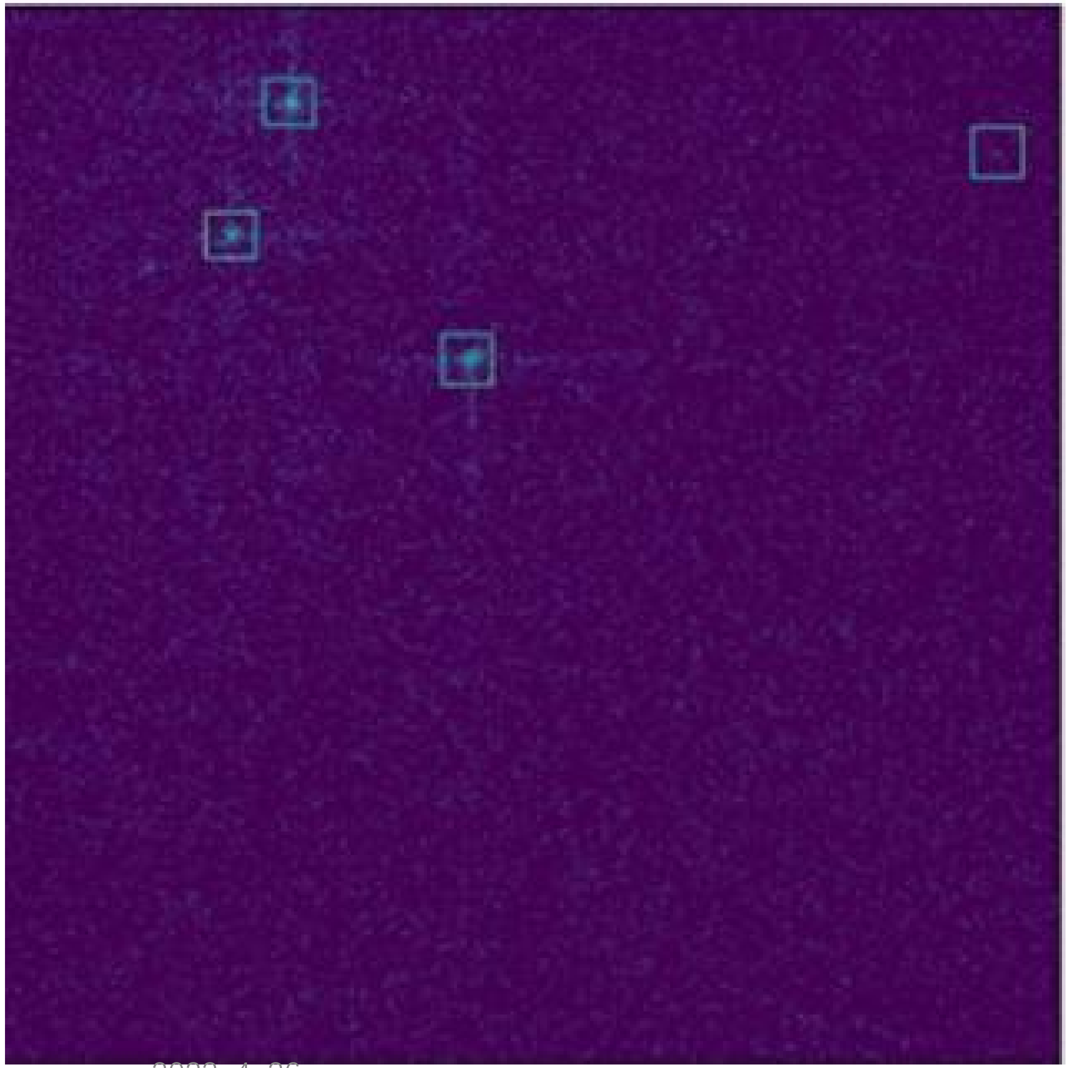


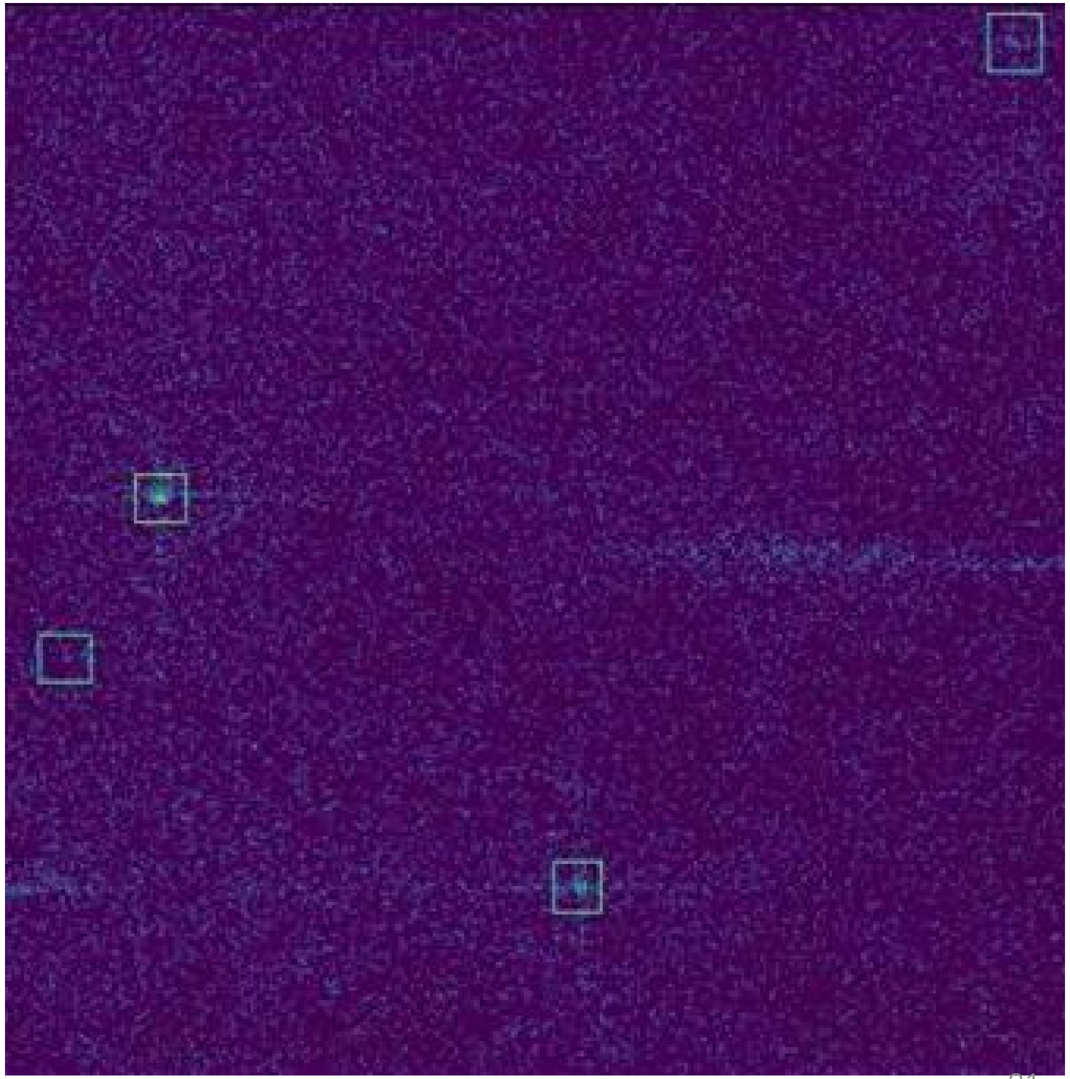

From Jia et.al. 2023

# Initial Results 2 – X-Ray Images

太原理工大学
TAIYUAN UNIVERSITY OF TECHNOLOGY

**Applied to Real Observation Data (LEIA)**





From Jia et.al. 2023

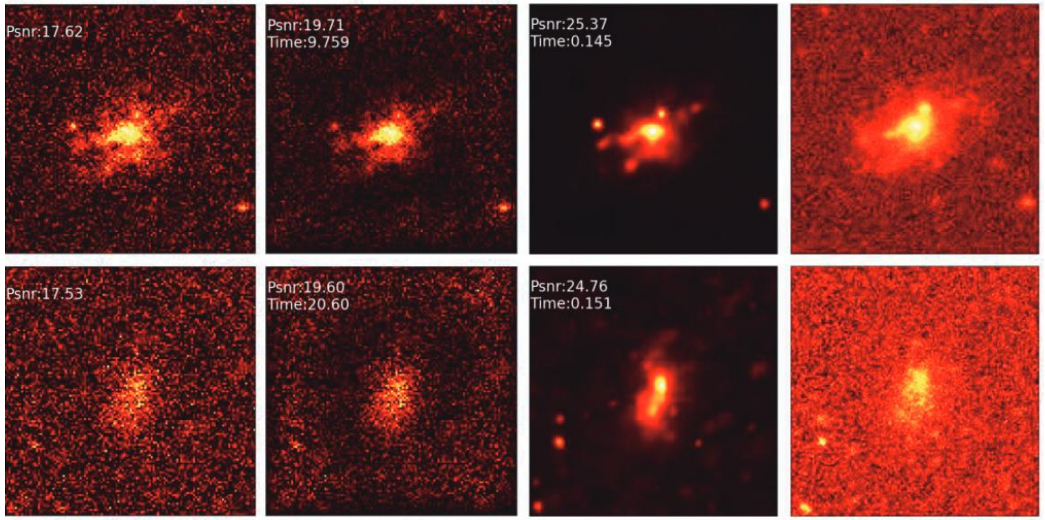
# And more ....

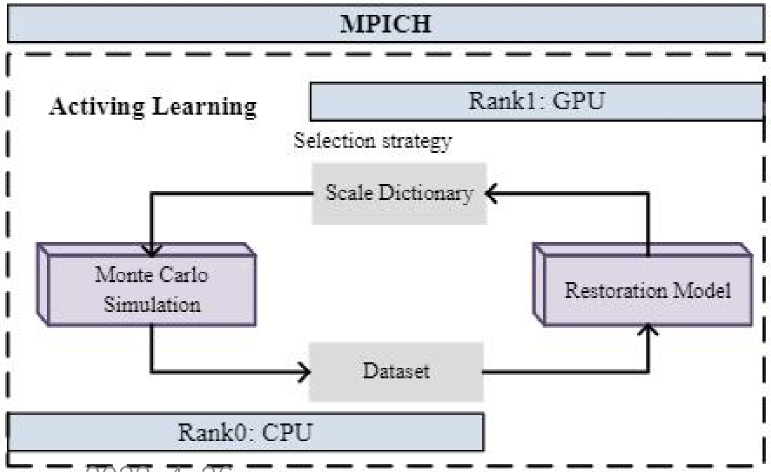
Lessons Learnt:

·It would be necessary to build pipeline to discover unusual astronomical targets.

· Connect digital twin with large model would be a possible way.

· Data augmentation and pre-processing are important but too many to discuss...



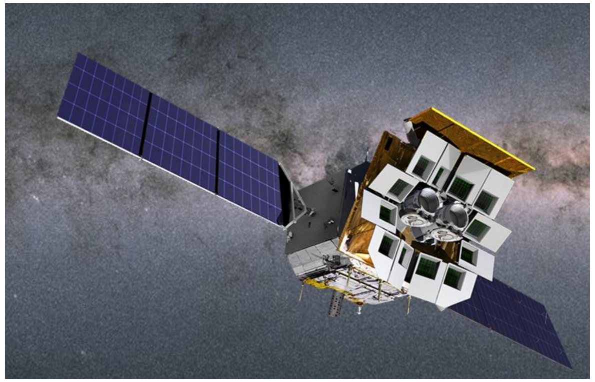




From Jia et.al. 2023

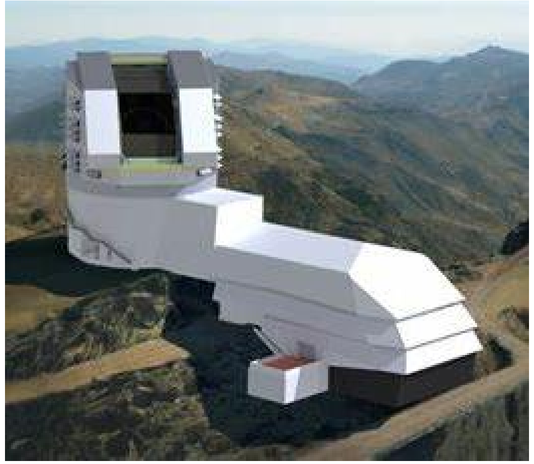
# Future

Waiting for data and working for new discoveries
(CSST, EP, Sitian, LOT, SKA, LSST and all other scientific projects)

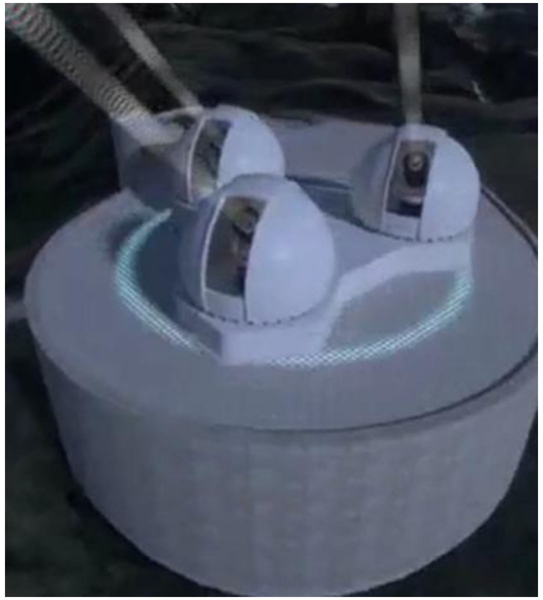
From https://english.cas.cn/newsroom/cas_media/202205/t20220507_305162.shtml



From https://ep.bao.ac.cn/ep/



From https://www.lsst.org/about/tel-site



Sitian Project