AI + Astronomy: Models, Data, Discovery

PI-AstroDeconv: Physics-Informed Deep Learning for Telescope Beam Deconvolution

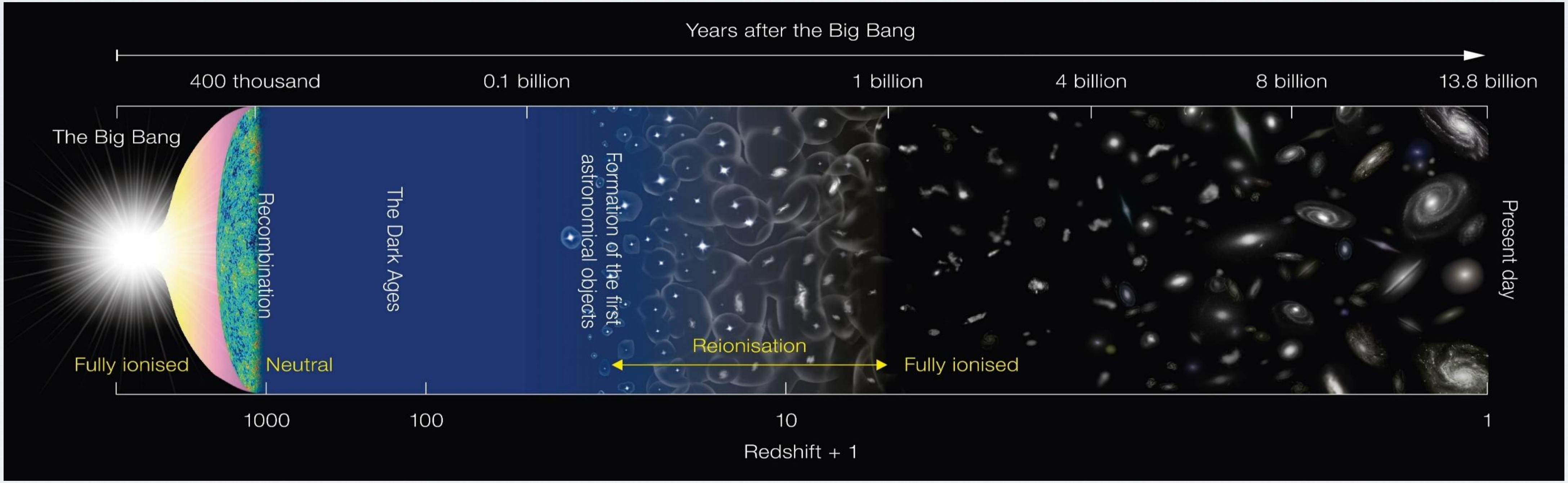
Shulei Ni • Zhejiang Lab

2026. 03. 31 • EPFL

Content

Content

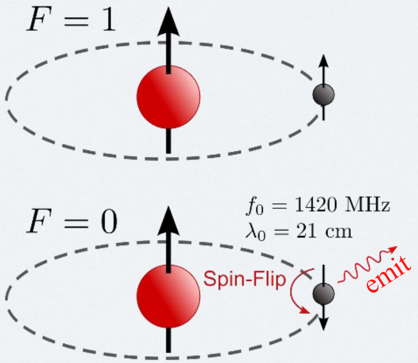
1.1 HI Intensity Mapping



Epoch of Universe:

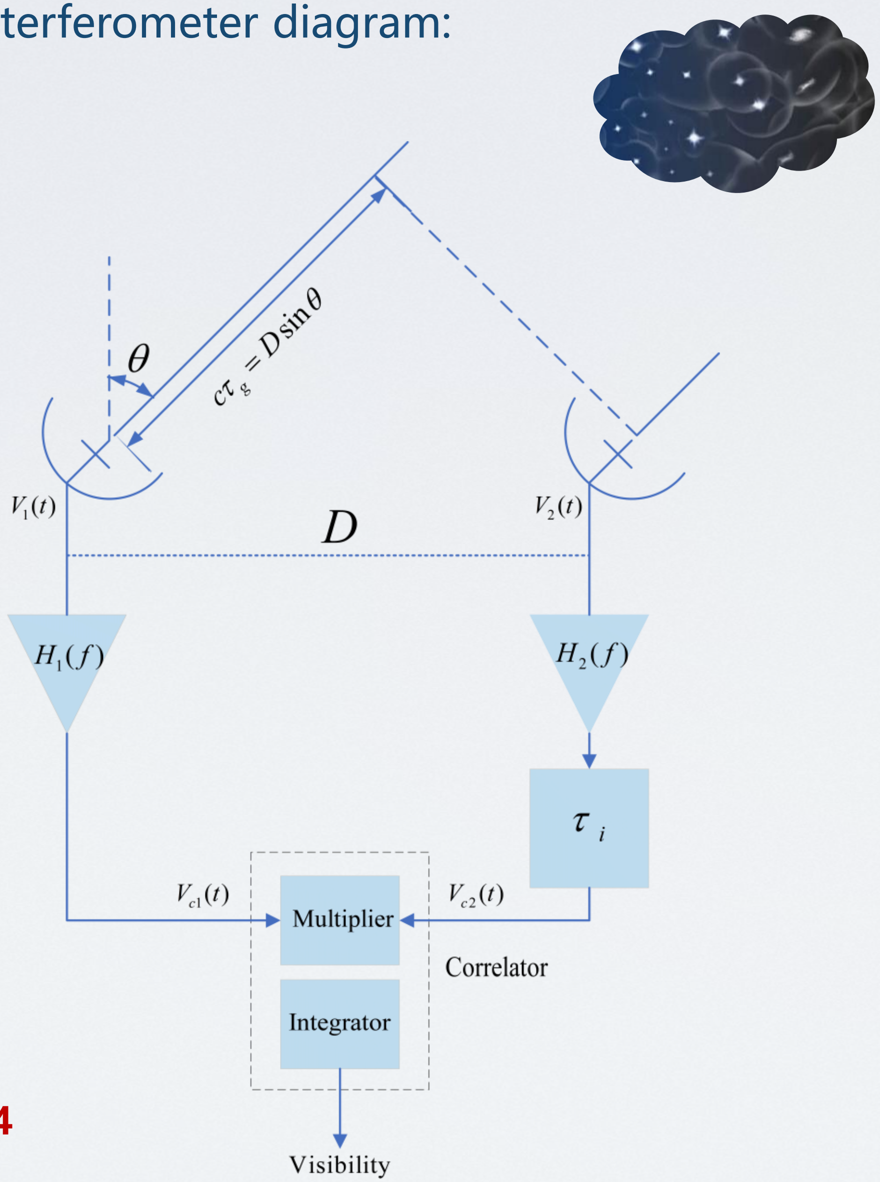
1. Cosmic Dark Ages (30 < z < 1100, 0.38-101 MY)
2. Cosmic Dawn (15 < z < 30, 101-272 MY)
3. Epoch of Reionization (6 < z < 15, 272 – 941 MY)

Hyperfine Structure：



1.2 HI Observation Pipeline

Interferometer diagram:



Visibility(V):

$$V(\mu,\nu) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A(x,y)I(x,y)e^{i2\pi(\mu x+\nu y)}dxdy$$

FFT:

$$I'(x,y) = A(x,y)I(x,y) = \int_{-\infty}^{\infty} V(\mu,\nu)e^{-i2\pi(\mu x+\nu y)}d\mu d\nu$$

Gridded Visibility Function and DFT, Dirty Image:

$$I_D(x,y) = \sum_k g(\mu_k,\nu_k)V(\mu_k,\nu_k)e^{-i2\pi(\mu_k x+\nu_k y)}$$

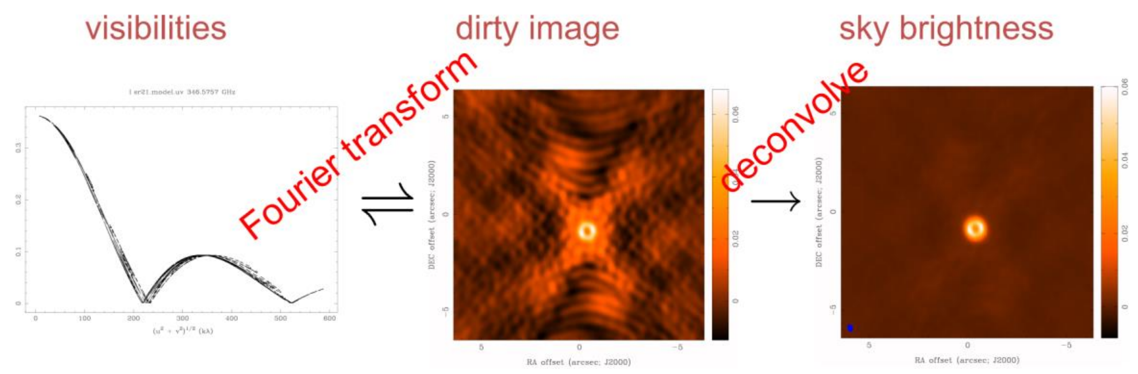
FFT ⇋ Convolution :

$$\boldsymbol{I_D(x,y) = P_D(x,y) \otimes I'(x,y)}$$

I' is real image, P_D is the dirty beam (PSF):

1.2 HI Observation Pipeline

Formula(traditional):



$$\begin{tikzcd}
V(\mu,\nu) \arrow[r, "\text{FFT}"] & \text{Dirty Image} = & \text{Dirty Beam} \arrow[d, leftrightarrow] & \otimes & \text{Real Image} \arrow[d, leftrightarrow] \\
& \text{Clean Image} = & \text{Clean Beam} & \otimes & \text{Model}
\end{tikzcd}$$

W **WSClean**

Project ID: 19986999

```
wsclean -reorder -use-wgridder -parallel-gridding 10 -weight natural -oversampling
4095 -kernel-size 15 -nwlayers 1000 -grid-mode kb -taper-edge 100 -padding 2 -
name out_name -size 2048 2048 -scale 16asec -niter 0 -pol xx -make-psf in_ms
```

CASA

```
tclean(vis='xx.ms', imagename='SNR.MultiScale', deconvolver='multiscale',
scales=[0,6,10,30,60], smallscalebias=0.9, imsize=4095, cell='16arcsec', pblimit=-0.01,
niter=1000,weighting='briggs', stokes='I', robust=0.0, interactive=False,
threshold='0.12mJy', savemodel='modelcolumn')
```

Challenges:

Computing Power :

1. Need to test different niter values;
2. Large niter → high CPU, memory, and time cost.

Noise:

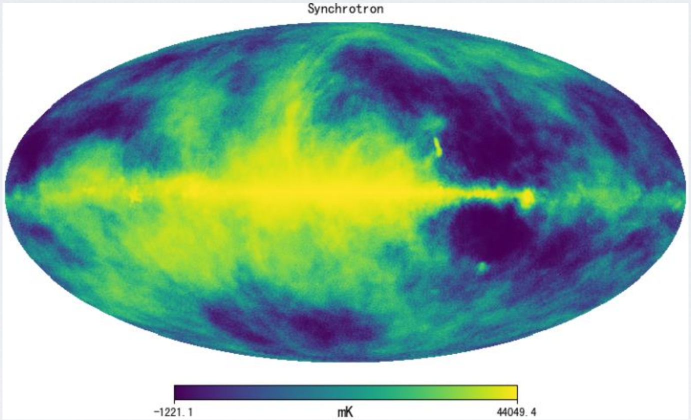
1. Foregrounds, $10^5 - 10^6$;
2. Beam Effect.

Deep Learning?

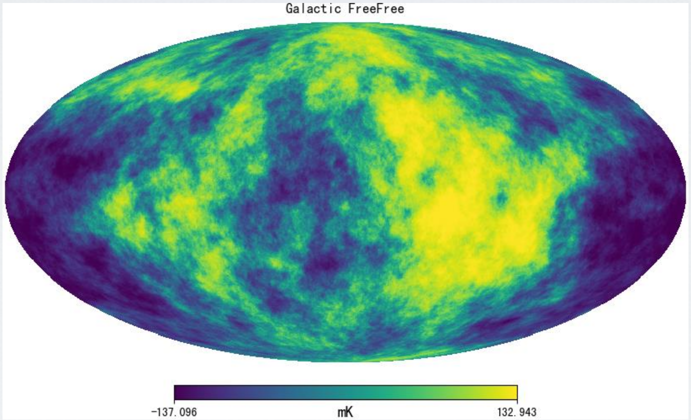
Content

2.1 Datasets

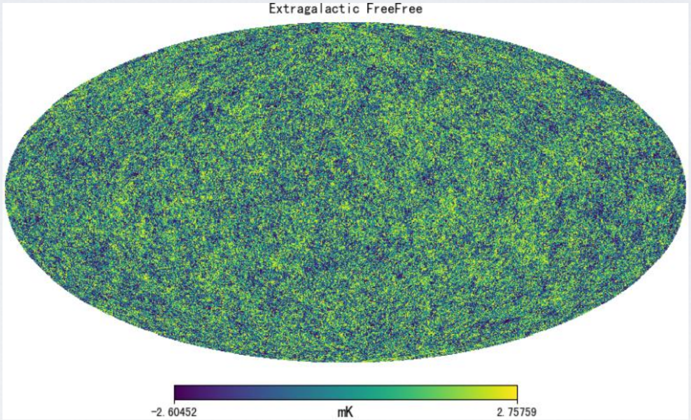
DATA: simulate with CRIME

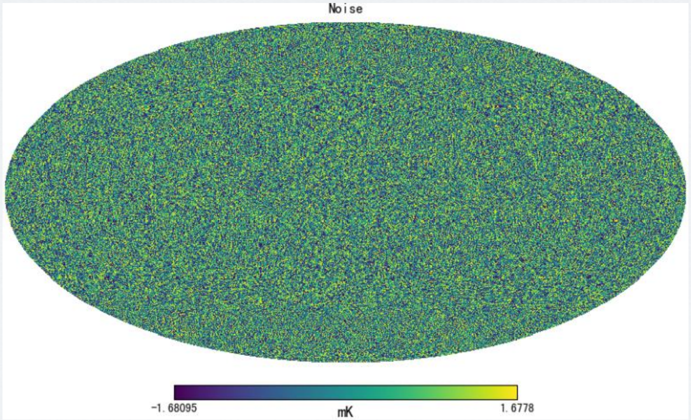


+

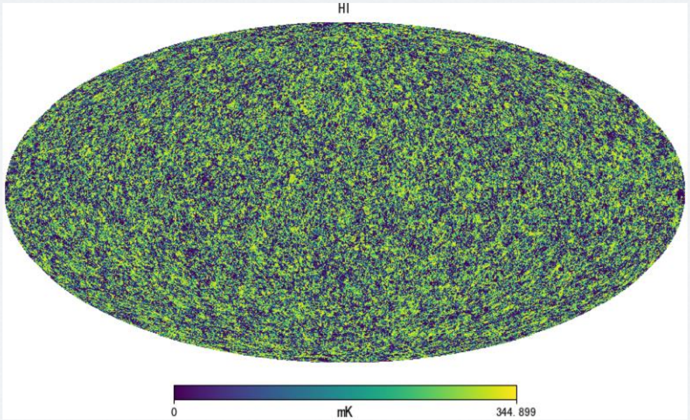


+

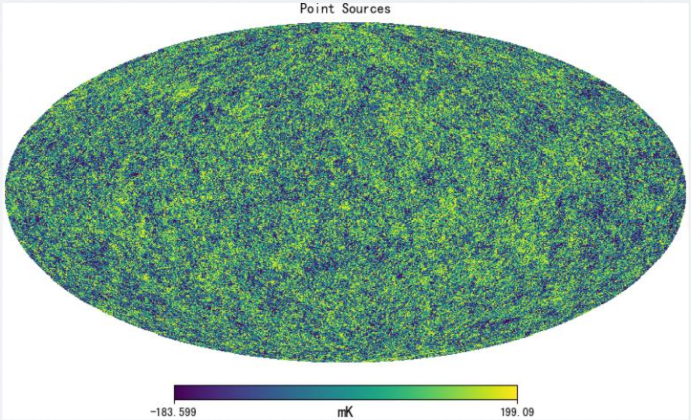




+



+



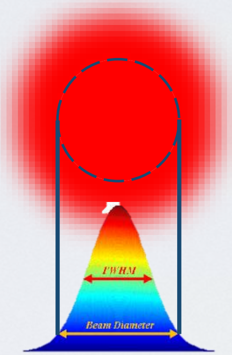
2.1 Datasets

Beam: simulation

Gaussian Beam:

$$B_g(\nu,\theta) = e^{-4ln2(\theta/\Delta\theta_g)^2}$$

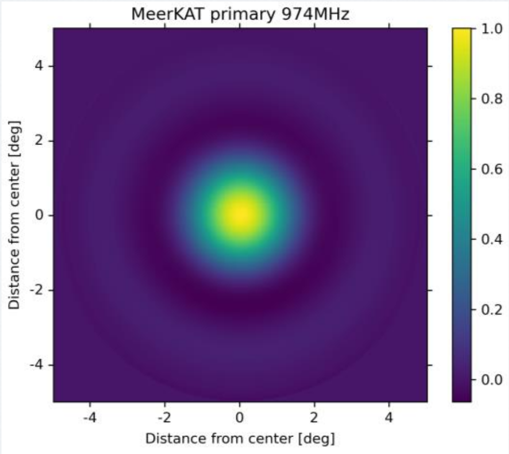
$$\Delta\theta_g = 1.16\frac{\lambda}{D}$$

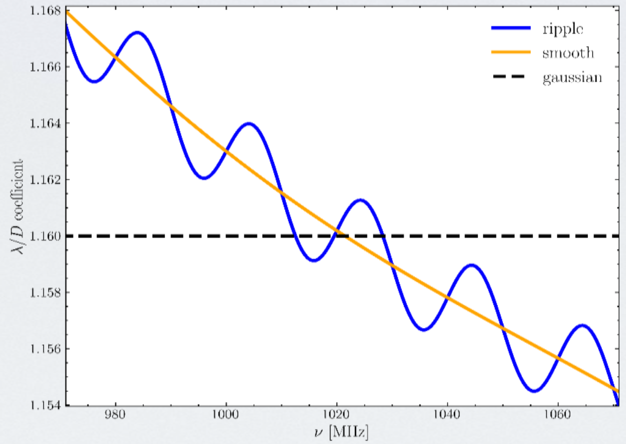


MeerKAT Beam:

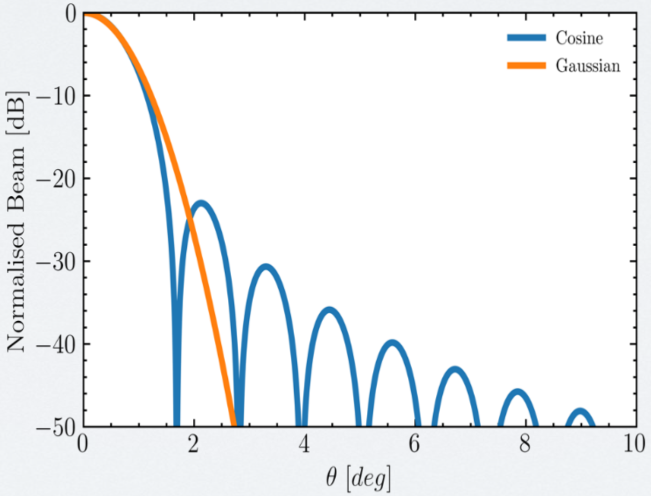
$$B_c(\nu,\theta) = \left[\frac{cos(1.189\theta\pi/\Delta\theta_c)}{1-4(1.189\theta/\Delta\theta_c)^2}\right]^2$$

$$\Delta\theta_c = \frac{\lambda}{D}\left(\sum_{d=0}^{8} a_d\nu^d + Asin\left(\frac{2\pi\nu}{T}\right)\right)$$

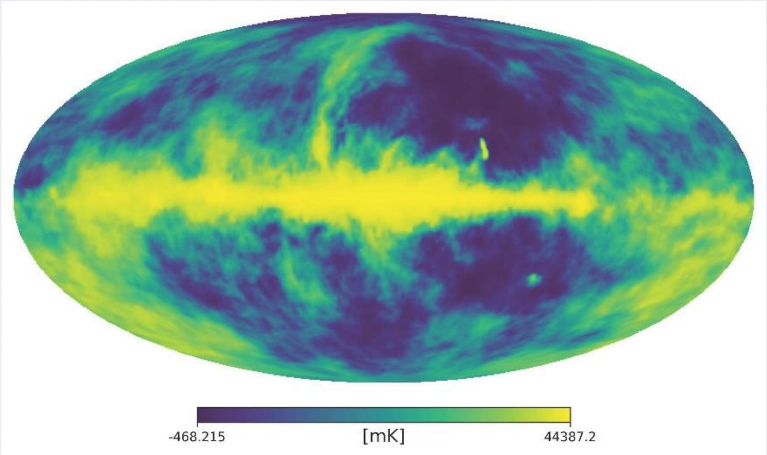




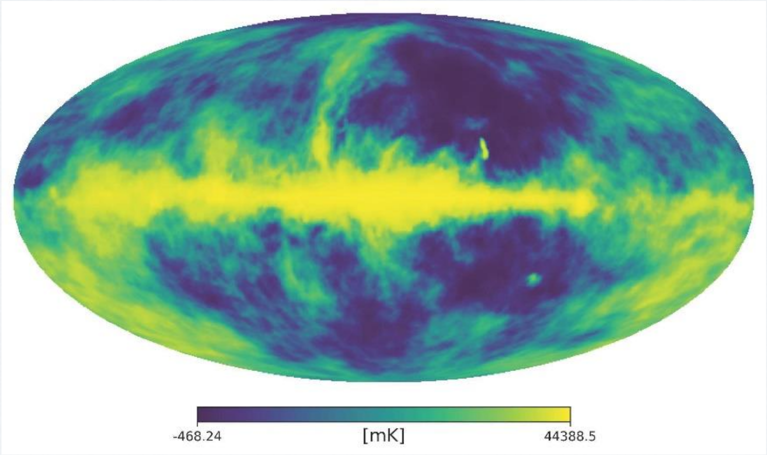
FWHM vs Freq



Beam Power



Gaussian Beam Smoothing



Cosine Beam Smoothing

2.1 Datasets

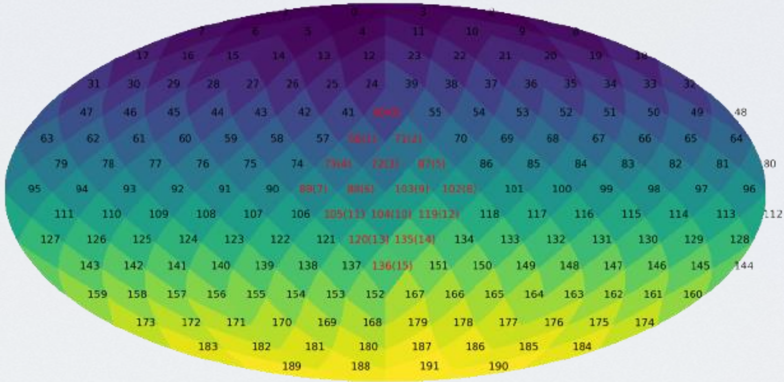
Nside=1:

12 pixels

Upgrade

1→4

Nside=4(based map):

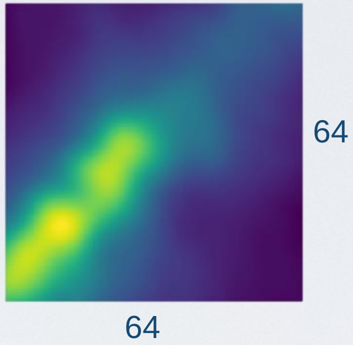


192 pixels

Upgrade

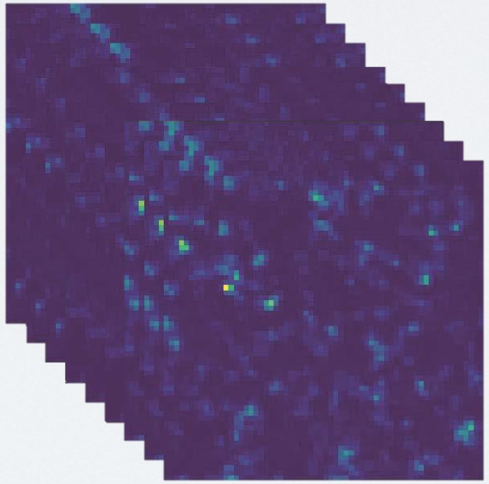
4→256

Nside=256(one patch):

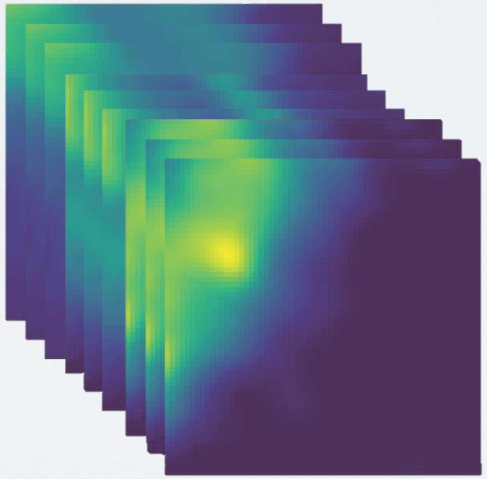


192*(64*64) pixels

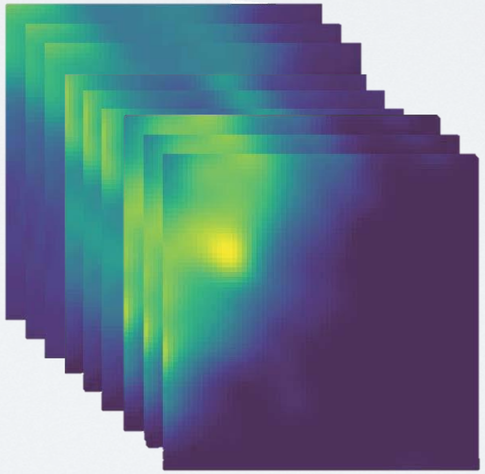
Datasests:



Label(HI)



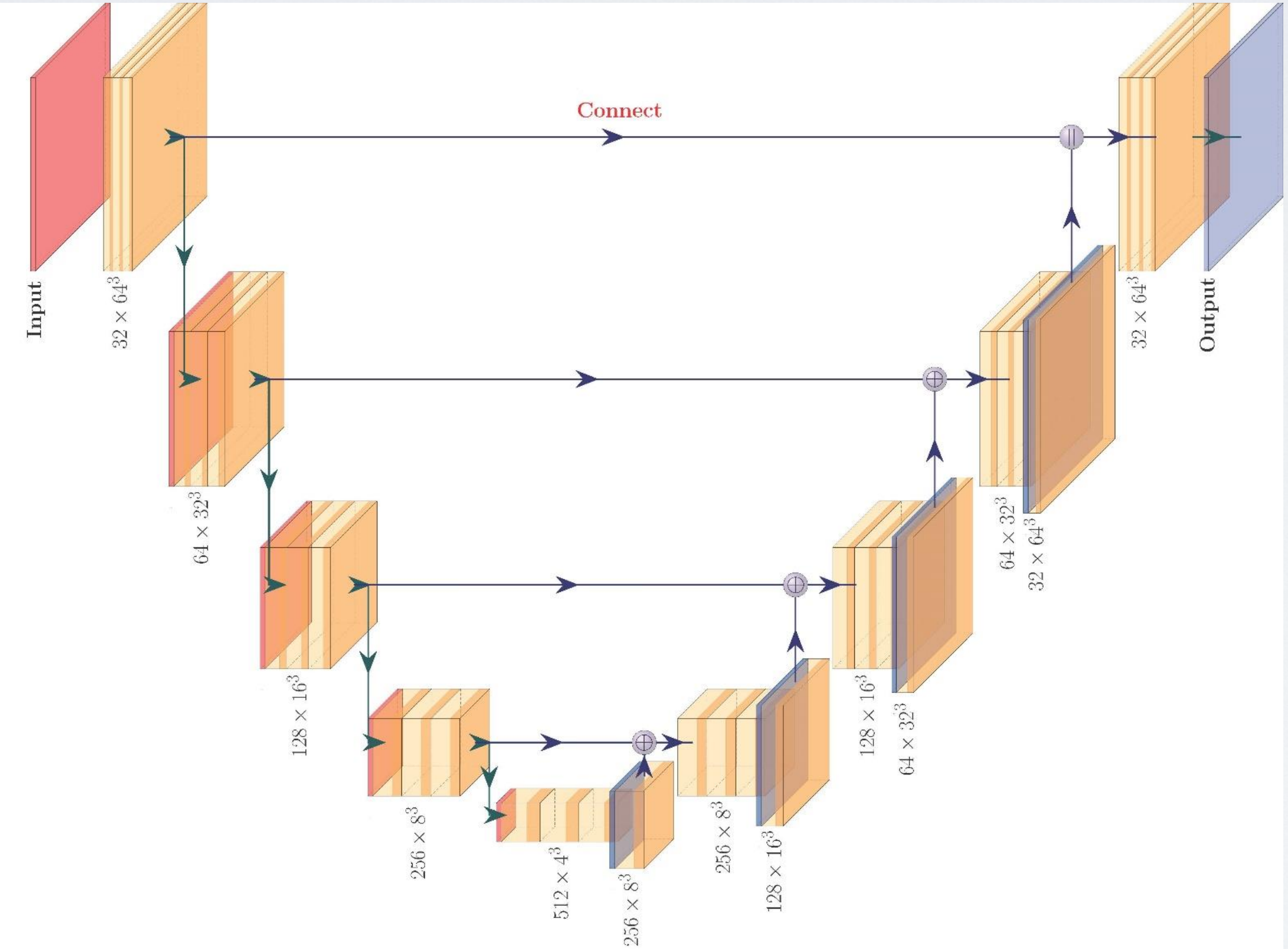
Gaussian Smoothing(HI+FG)



Cosine Smoothing(HI+FG)

2.2 Unet-3D

Unet-3D Architecture:



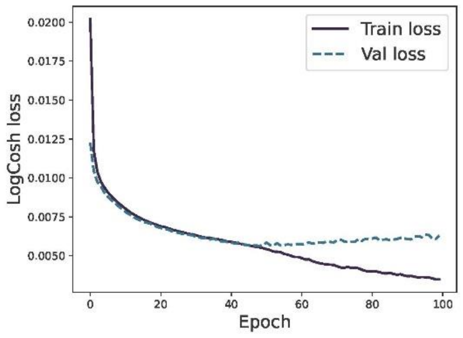
Loss Function:

$$Loss(p,t) = \sum_i \log\cosh(p_i - t_i)$$

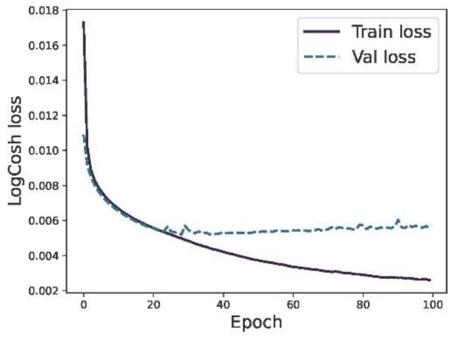
Hyperparameters:

lr	$n_{filters}$	***batch size***	***Optimizer***
10^{-4}	32	16	$NAdam$

Loss vs Epoch:



Gaussian Beam



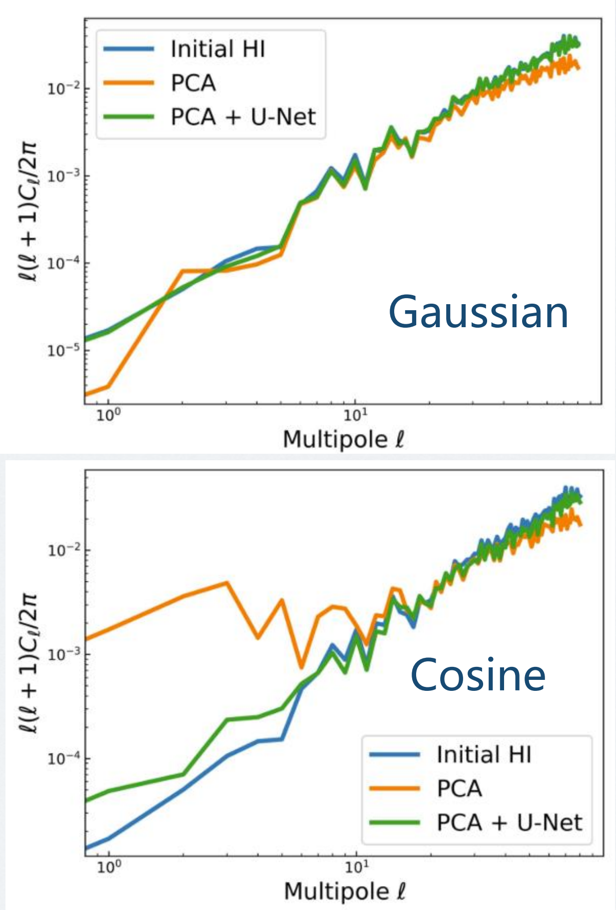
Cosine Beam

2.2 Unet-3D

Train: $\mathrm{PCA}_{res}[\mathrm{Smooth}(FG + HI)]$

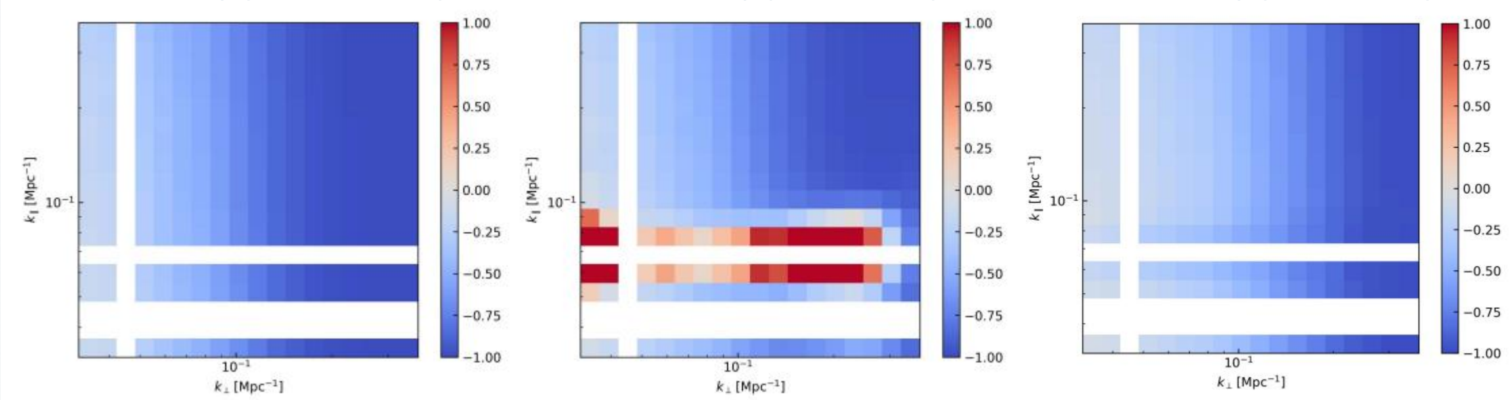
Label: HI

Angular Power Spectrum：



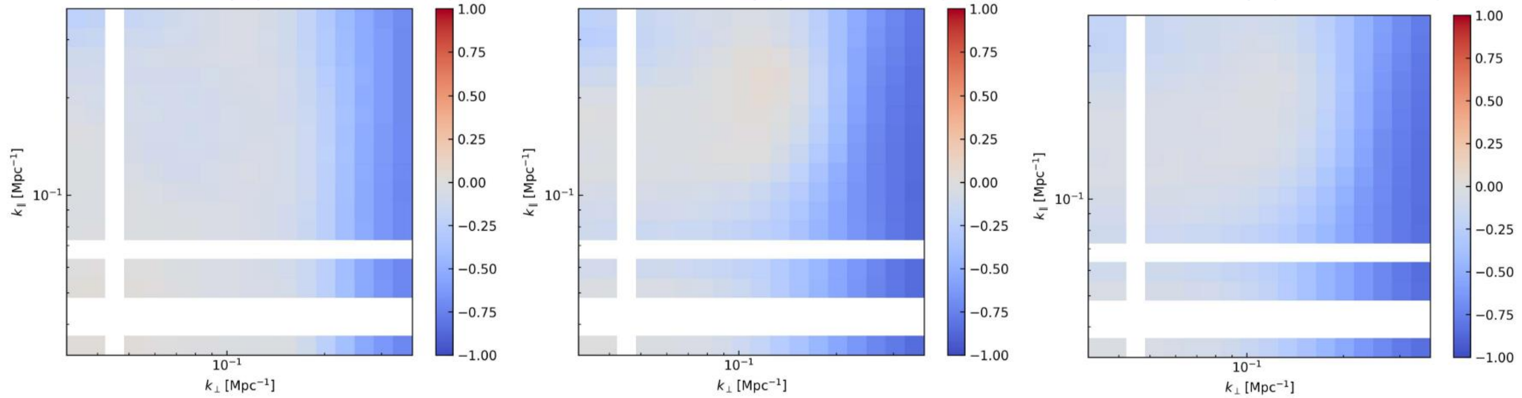
2D auto-correlation power spectrum：

PCA cleaned map (Gaussian Beam) PCA cleaned map (Cosine Beam) U-Net cleaned map (Cosine Beam)



2D cross-correlation power spectrum：

PCA cleaned map (Gaussian Beam) PCA cleaned map (Cosine Beam) U-Net cleaned map (Cosine Beam)

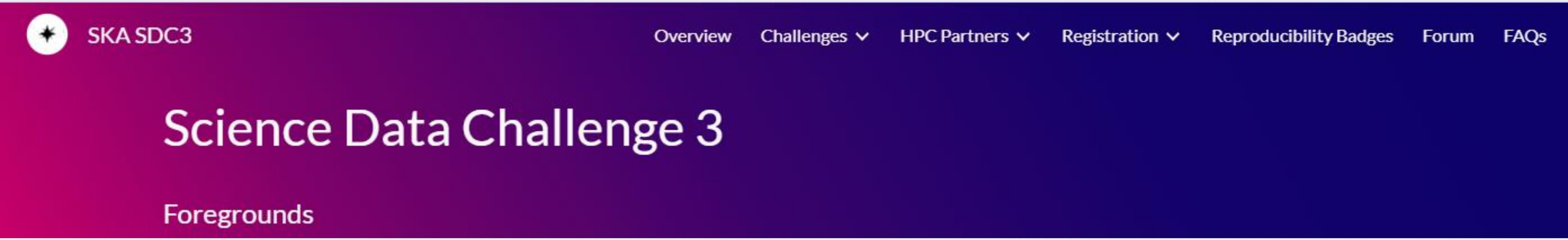


Supervised Model

APJ 934.1 (2022): 83

Content

3.1 SDC3a



Science Data Challenge 3

Foregrounds

SDC3a:

Objective: Detect the faint Hydrogen-21cm signal from the Epoch of Reionisation (EoR).

Datasets(image and visibility):

1. Volume: 7.5TB
2. Freq: 106-196(z=[6, 15])
3. Resolution: 16x16 arcsec, 100kHz
4. Pixels Number(RA/Dec): 2048x2048; (cube: 900x2048x2048)
5. Synthesised Beam and Thermal noise

Task: Perform Foreground Removal to isolate the EoR signal from the noisy data.

3.2 PI-AstroDeconv

Results of Foreground Removal (based on the Unet-3D job):

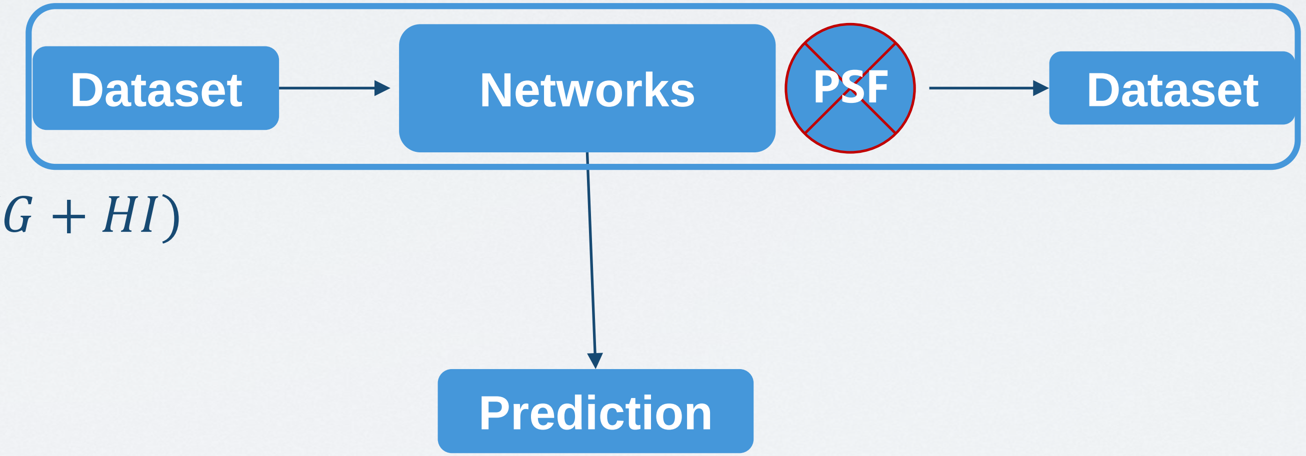
Success (√) : simple beam model (gaussian beam), PCA can remove the foreground;

Failure (×) : complex beam model (MeerKAT beam, cosine beam), PCA can not work.

Formula (PCA, Smooth Beam, U-Net as an operator**):**

$$\mathrm{UNet}\{\mathrm{PCA}_{res}[\mathrm{Smooth}_{beam}\,(FG+HI)]\}\approx HI \quad \Longleftrightarrow \quad \mathrm{PCA}_{res}\{\mathrm{UNet}[\mathrm{Smooth}_{beam}\,(FG+HI)]\}\approx HI$$

PI–AstroDeconv:



$$(\mathrm{UNet}+\mathrm{Conv}_{Beam})[Beam(FG+HI)] = Beam(FG+HI)$$

Unet: foreground removal

Conv_{Beam}: beam effect

Dataset: Train = Label : Beam$(FG+HI)$

3.2 PI-AstroDeconv

PI-AstroDeconv Architecture :

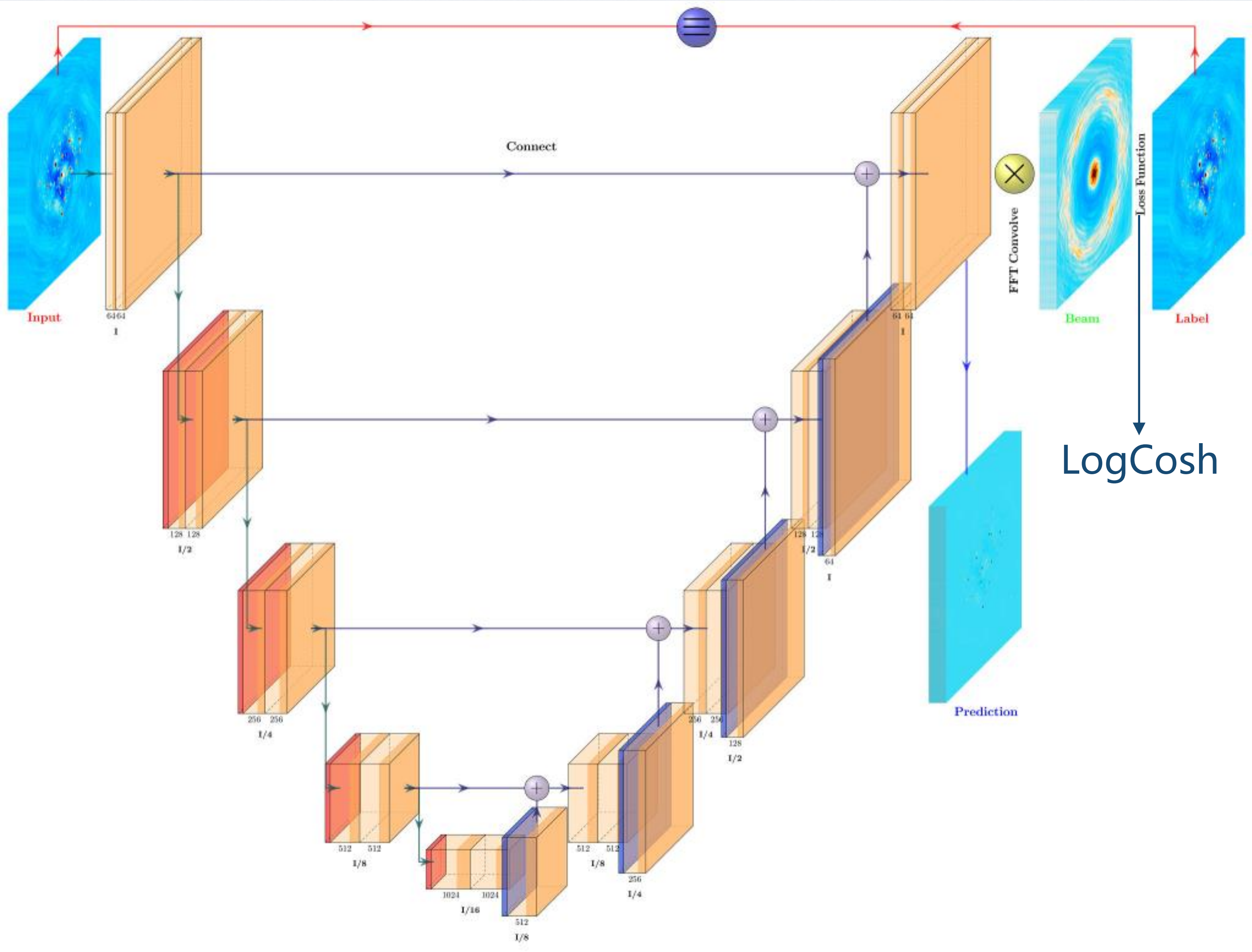
Pink Box: U-Net network

$\text{Input} = \text{Label(Output)} = \text{Image}_{\text{smoothed}}$

$\text{Prediction} = \text{Image}$

$\text{Beam} = \text{Dirty Beam}$

$\text{Dirty Image} = \text{Dirty Beam} \otimes \text{Real Image}$



LogCosh

3.2 PI-AstroDeconv

SDC3a:

Inputs: Large size, (2048*2048)

Beam : Large size, (2048*2048)

FFT Convolution :

- Standard Convolution(w=h=n):

$$O(n \times n) = O(n^2)$$

Time consumption (scipy conv):

$$\Delta t_{\mathrm{conv}} = 37797.2303 \text{ sec}$$

- FFT Convolution(w=h=n):

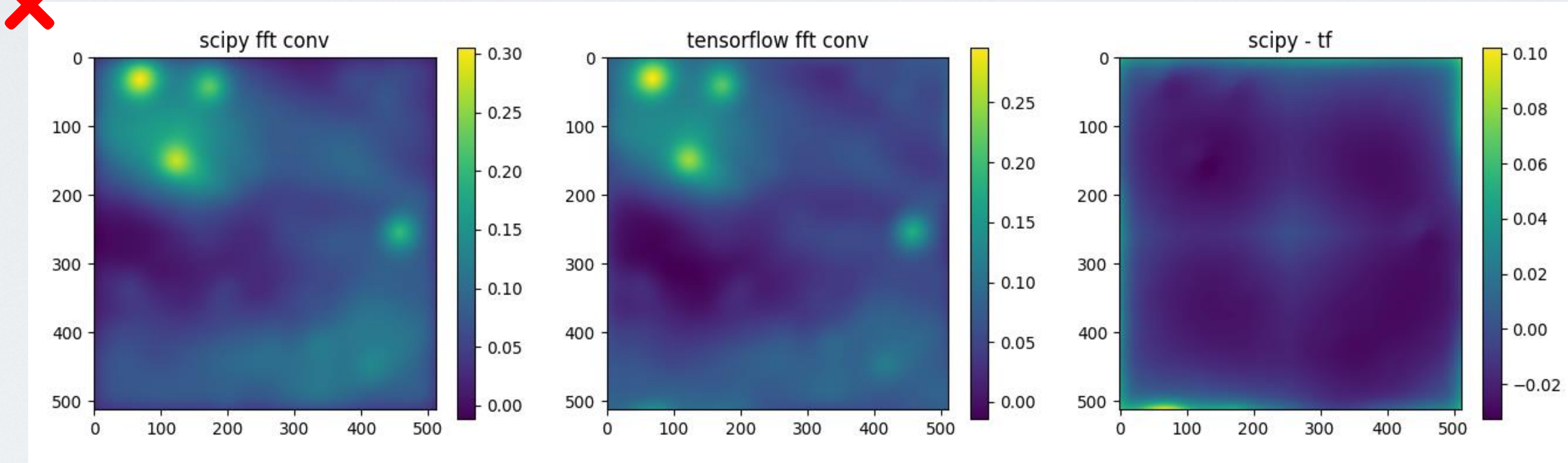
$$O(n \times n) = O[n \times \log(n)]$$

Time consumption (scipy fft conv):

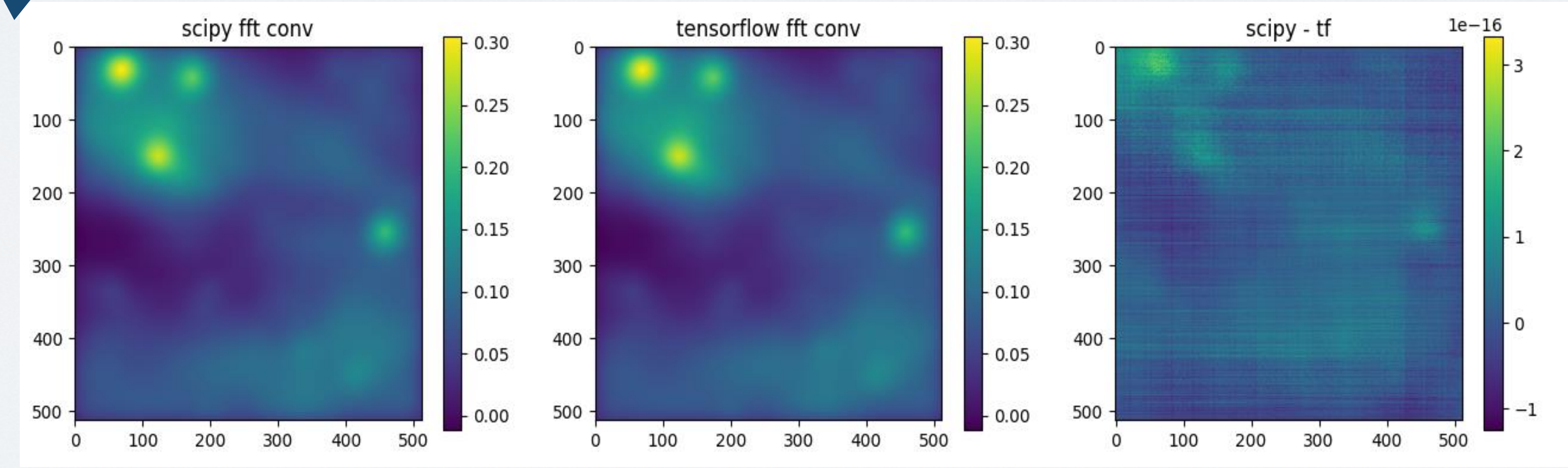
$$\Delta t_{\mathrm{fft}} = 0.5391 \text{ sec}$$

$$\frac{\Delta t_{\mathrm{conv}}}{\Delta t_{\mathrm{fft}}} = \mathbf{70111.7238}$$

✗



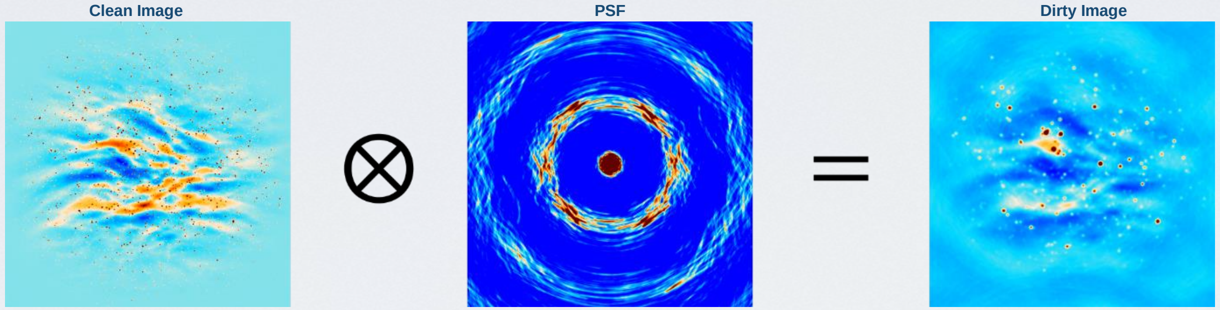
✓

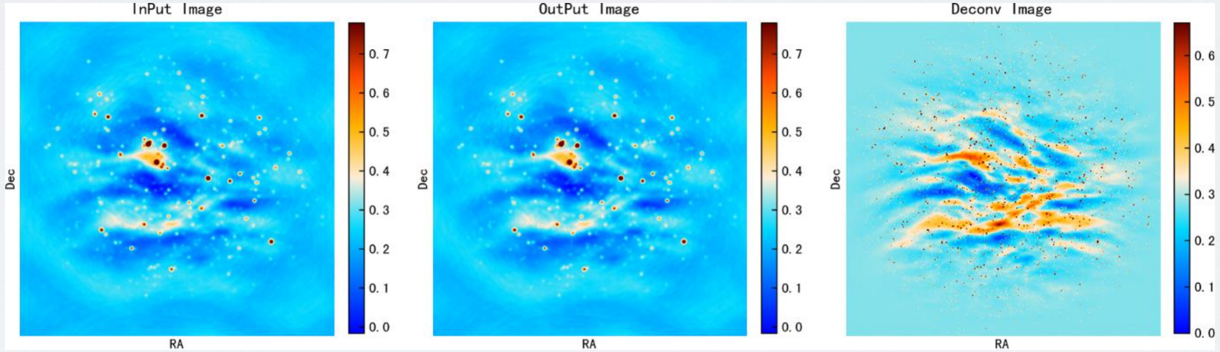


3.2 PI-AstroDeconv

SDC3a Dataset:

Slice : Freq = 171.7MHz



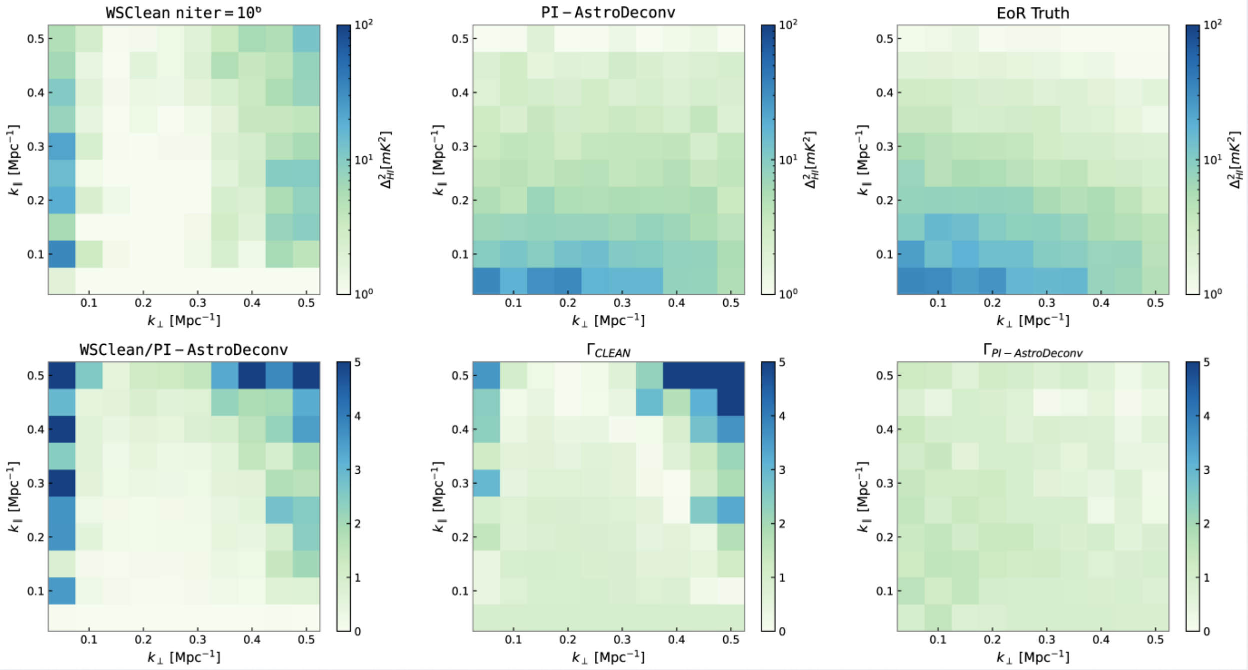


$$R(k_{\parallel}, k_{\perp}) = \left|\frac{P_{pred}(k_{\parallel}, k_{\perp})}{P_{truth}(k_{\parallel}, k_{\perp})} - 1\right|$$

$$\Gamma = Mean\{R(k_{\parallel}, k_{\perp})_X\}$$

SDC3a Results

Freq band: 166MHz – 181MHz



SDC3a:

Top 3% (9/324, score) and 1% (3/324, chi-square) in SKA Data Challenge 3a (SDC3a, Mar–Oct 2023), among 324 global teams (77 from China).

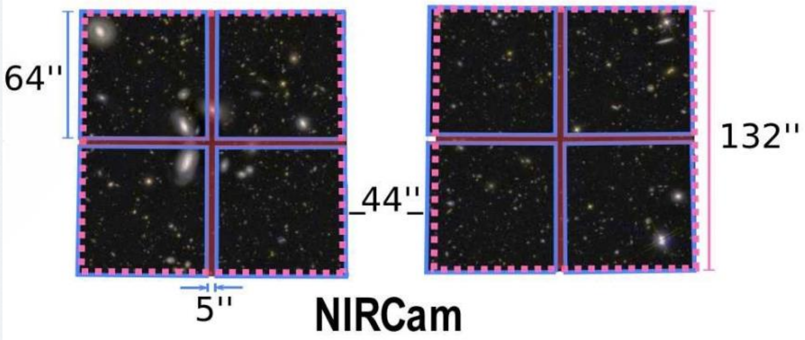
MN, 2025, 543(2): 1092-1119; APJ, 2025, 990(2): 122

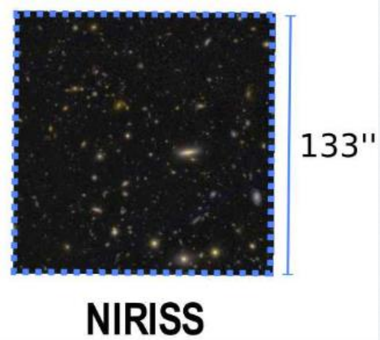
Content

4.1 JWST

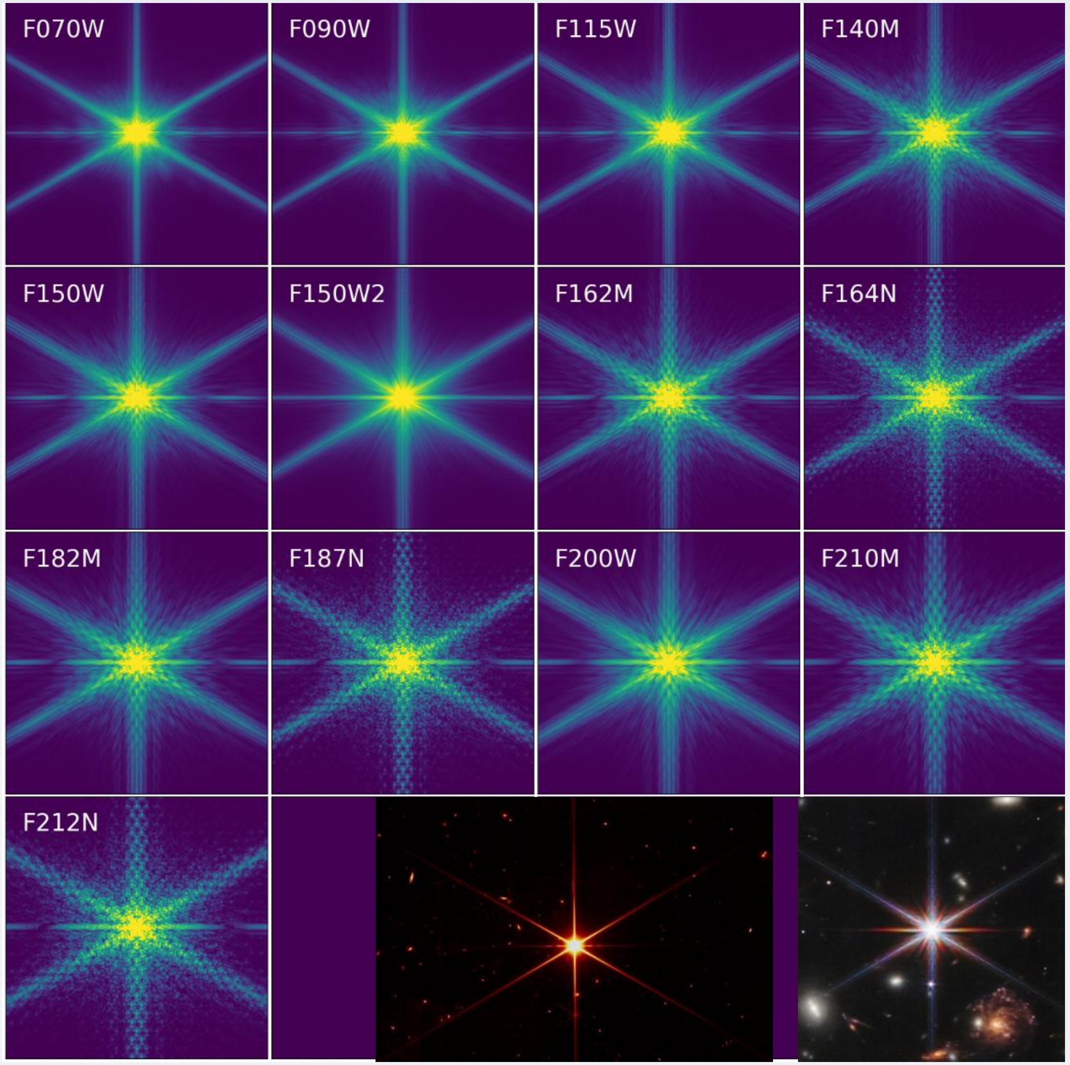
JWST Datasets

NIRCam	SW (Short Wavelength)	0.6–2.3 µm
	Filters	F070W, F090W, F115W, F150W, F200W
	LW (Long Wavelength)	2.4–5.0 µm
	Filters	F277W、F356W、F444W
NIRISS	Wavelength	0.6–5.0 µm
	Broadband Filters	F090W, F115W, F150W, F200W, F277W, F356W, F444W
	Medium-band filters	F140M, F158M, F380M, F430M, F480M



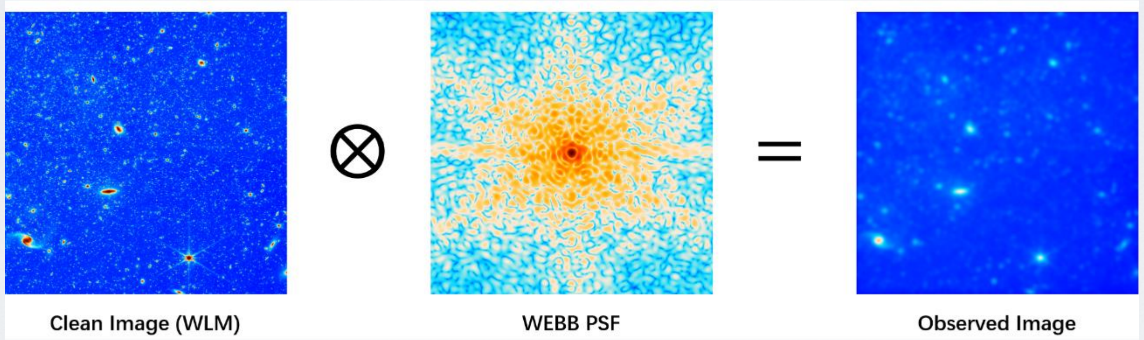


JWST PSF:



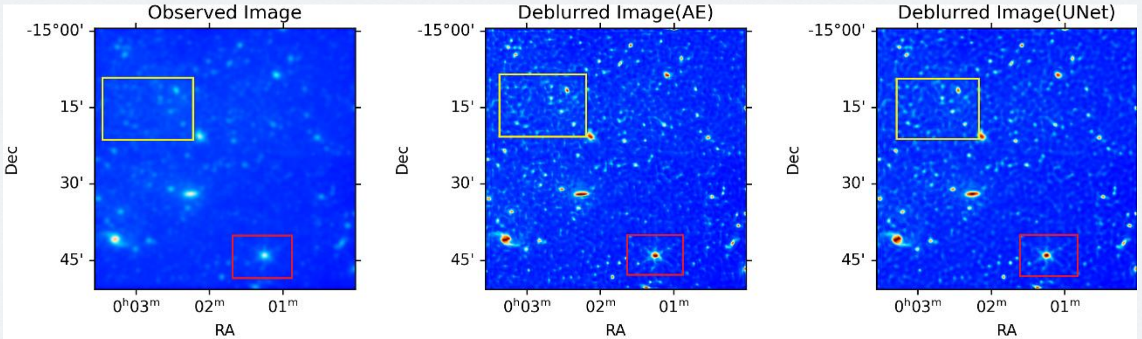
4.1 JWST

WEBB Test Datasets:



Clean Image (WLM) WEBB PSF Observed Image

Results(Image):



Results(SSIM&PSNR):

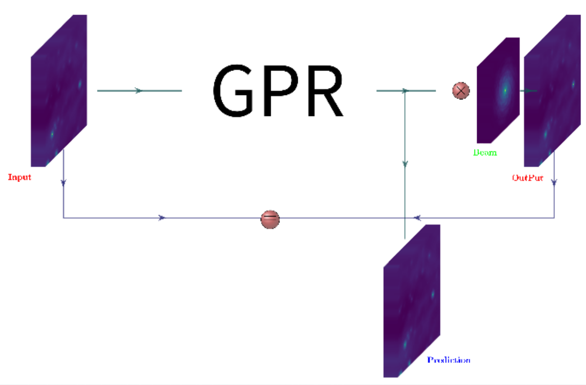
Methods	Image Quality Metrics	Images No. 1	No. 2	No. 3	No. 4	No. 5
Treitel [1969]	SSIM	0.6688	0.5452	0.6683	0.6611	0.5903
	PSNR(dB)	23.16	21.13	25.97	20.85	21.67
Bai et al. [2019]	SSIM	0.7305	0.3581	0.7298	0.7033	0.6720
	PSNR(dB)	19.13	20.05	24.90	19.71	22.04
Chen et al. [2021]	SSIM	0.7842	0.5967	0.6463	0.8085	0.7674
	PSNR(dB)	25.24	23.78	25.69	25.80	27.31
Nan and Ji [2020]	SSIM	0.7752	0.5686	0.6129	0.7127	0.6406
	PSNR(dB)	24.30	22.41	24.75	22.35	23.54
Ren et al. [2020]	SSIM	0.6683	0.7688	0.7173	0.5368	0.5837
	PSNR(dB)	24.59	25.99	25.94	23.21	25.21
PI-AstroDeconv (AutoEncoder)	SSIM	0.7973	0.8116	0.8156	0.8072	0.7960
	PSNR(dB)	29.28	32.81	29.06	33.94	31.60
PI-AstroDeconv (U-Net)	SSIM	**0.8186**	**0.8668**	**0.8391**	**0.8970**	**0.8706**
	PSNR(dB)	**32.27**	**37.30**	**35.81**	**39.61**	**38.02**

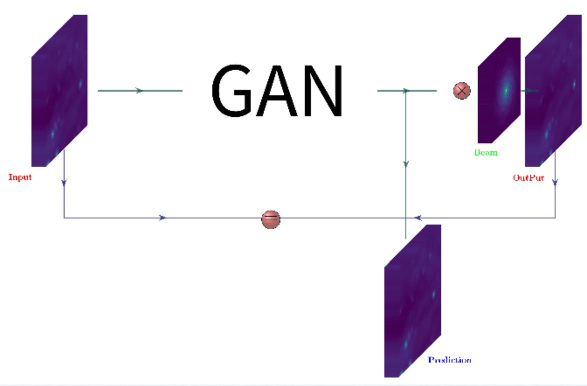
arXiv:2403.01692

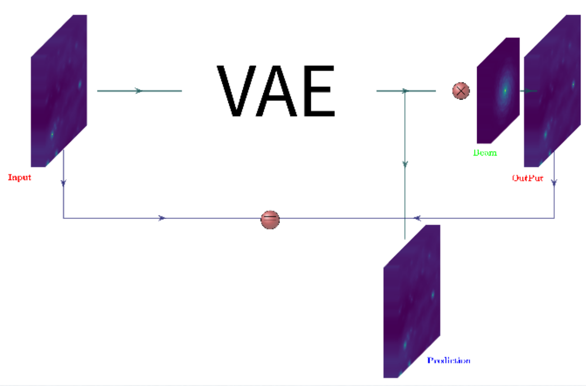
4.2 Network Extension

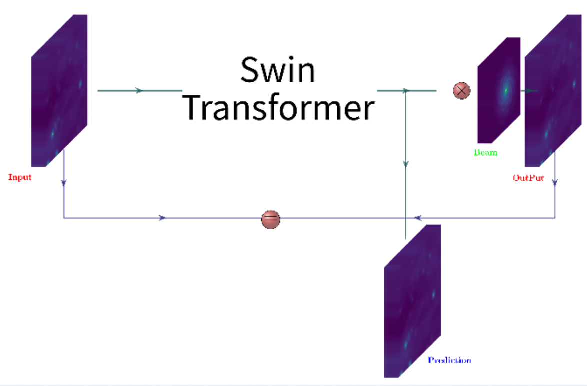
Different Networks:

Replaceable with any regression model.



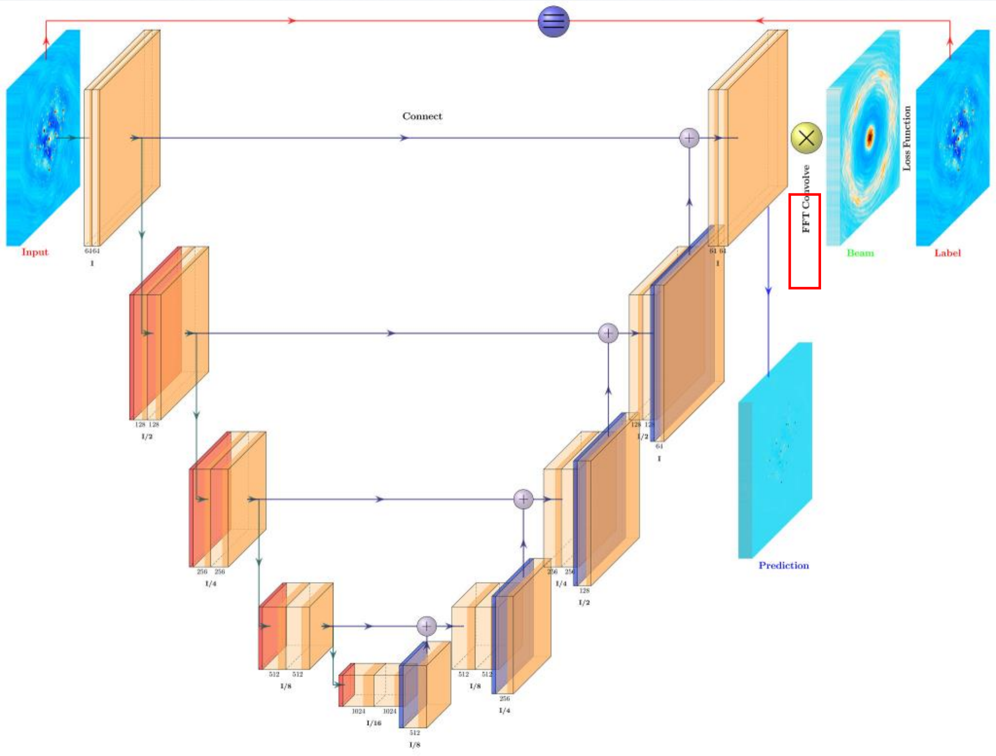






Different Beams:

Works with partial telescope beams:



EHT, ALMA (CCNU), SDSS(Paris Observatory),

VLA, HST…

5. Conclusions

- Build Unet-3D, for foreground removal
- Built PI-AstroDeconv, a smart tool that mixes physics informations with AI to clean telescope images.
- It succeeds not just for radio telescopes (SKA) but also for other telescopes like JWST.

Thanks!