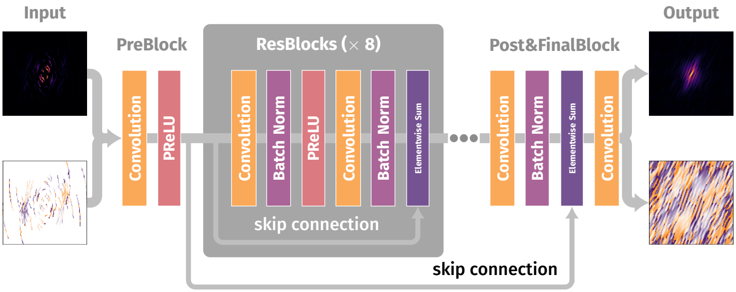
technische universität dortmund

e5 experimentelle physik 5
astroteilchenphysik



Deep learning-based cleaning in radio interferometry

Kevin Schmidt

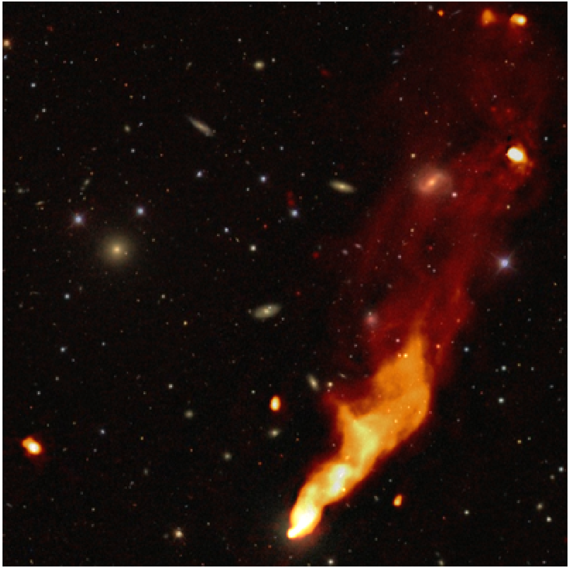
November 11th 2022

Astroparticle Physics / WG Elsässer
Department of Physics, TU Dortmund University

Motivation

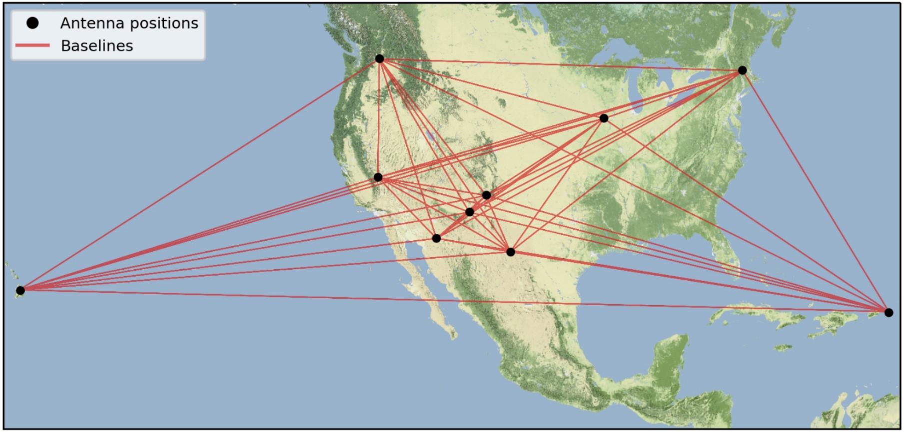
Motivation

- High-resolution images
- Straight forward cleaning
- Utilize machine learning



Motivation

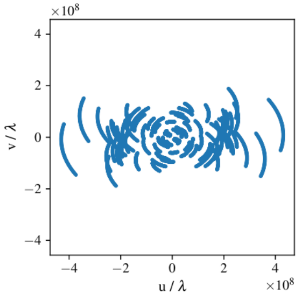
Rayleigh criterion: $\Theta \approx 1.22 \cdot \frac{\lambda}{D}$



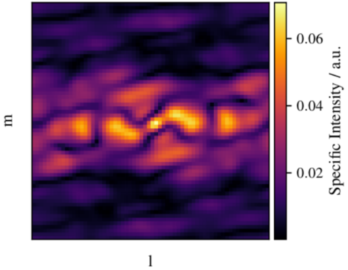
Motivation

Problem:

- Incomplete data samples
- Measurements in Fourier space
- Noise corruption
- Hidden source features



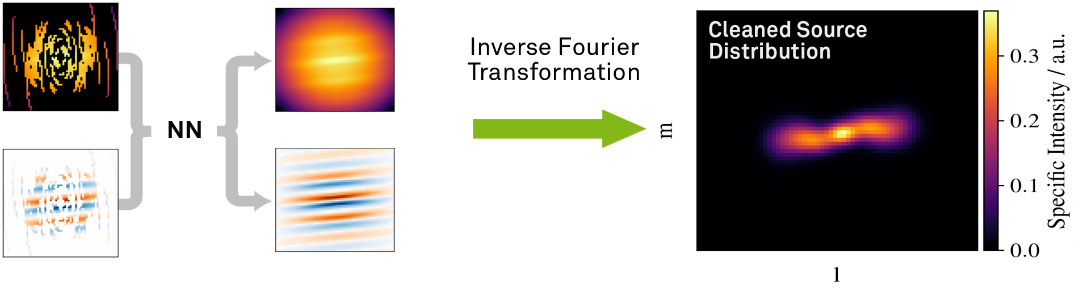
FFT



Motivation

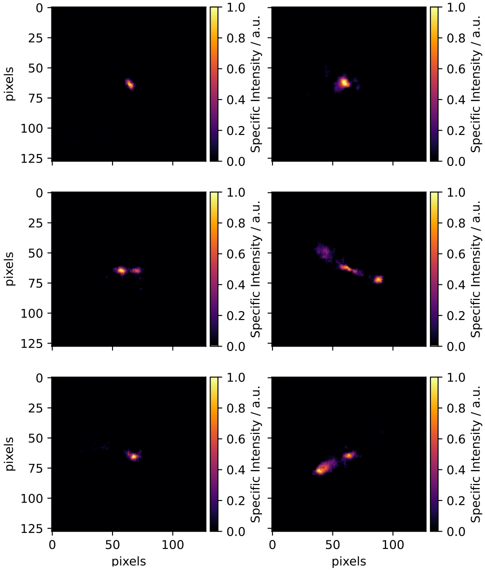
Solution:

- Use neural networks for data reconstruction
- Simulation of radio galaxy
- Simulation of radio interferometer responses
- Creation of data sets

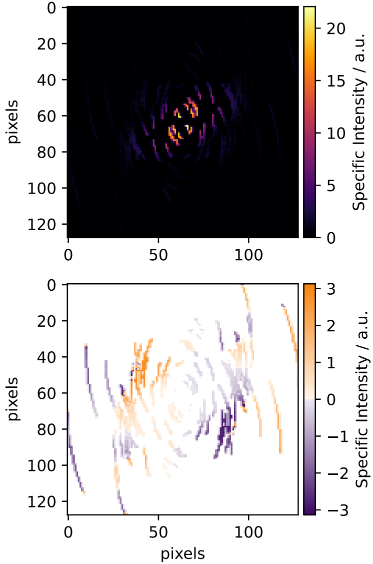


Radio galaxy simulations

- Realistic source shapes
- GAN trained on data from FIRST catalog
- Different source types
- Focus on FR I and FR II sources



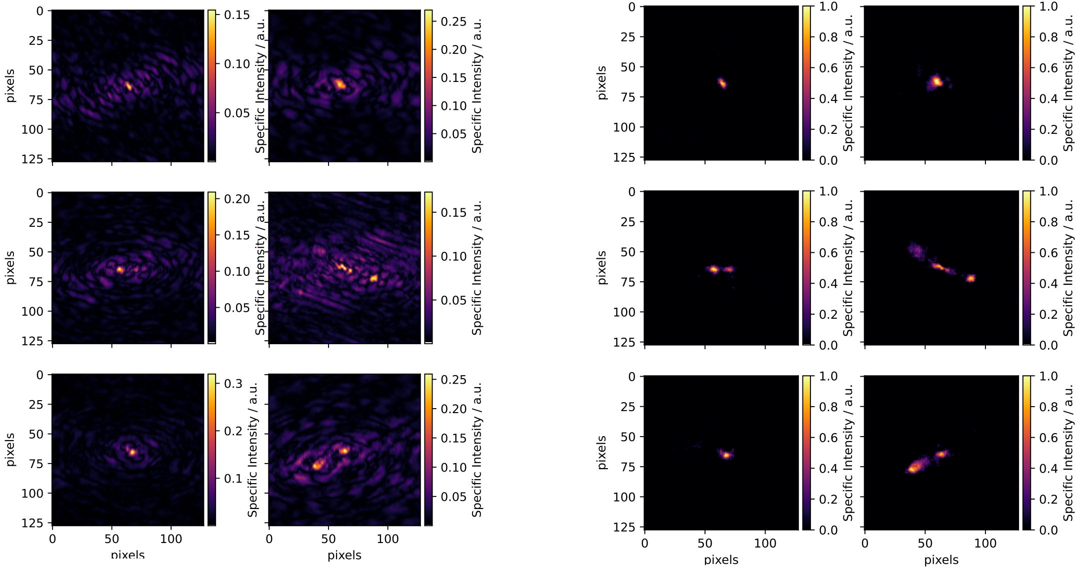
Radio interferometer observation simulations



Sampling masks

- Sparse data coverage
- Dependent on array layout and source position
- Different observation lengths

Dirty images



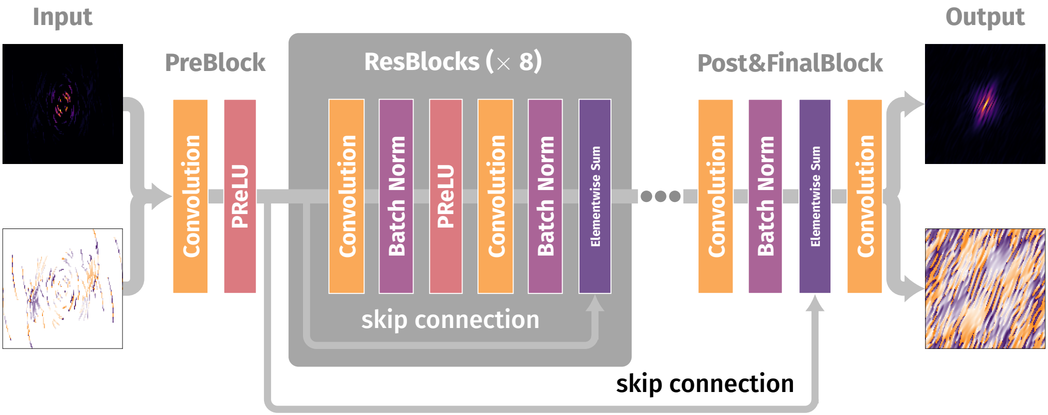
Problem definition

- **Goal:**
  Inpainting of missing information
- **Missing information:**
  Maximal corrupted pixels
- **Compareable problem:**
  Upscaling in super-resolution
- **Common architecture:**
  SRResNet

Bicubic

SRResNet

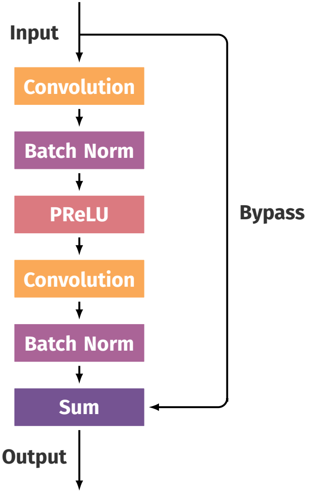
Architecture: Overview



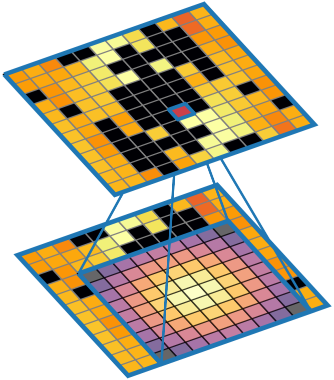
Architecture: Residual learning

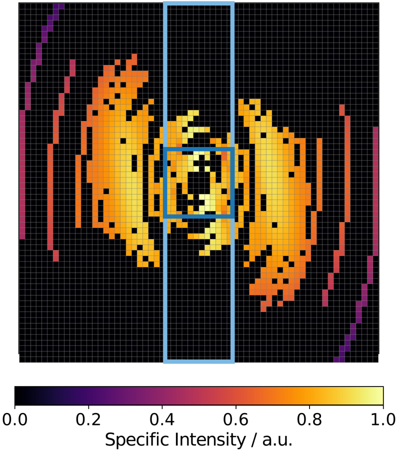
$F(x) = y - x$

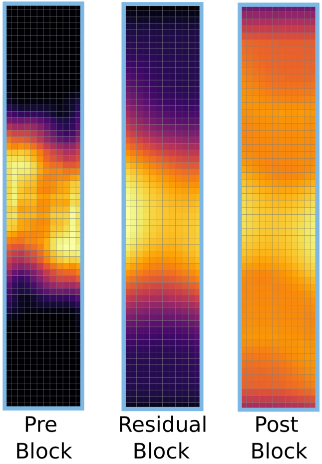
x: Input
y: Output



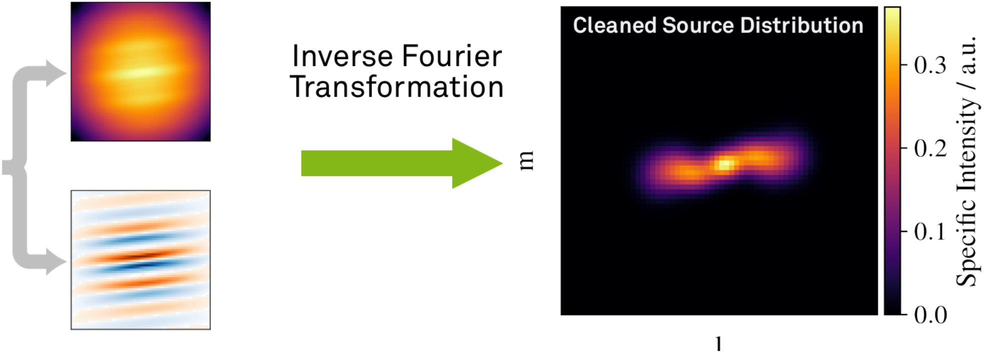
Architecture: Insights







Architecture: Predictions



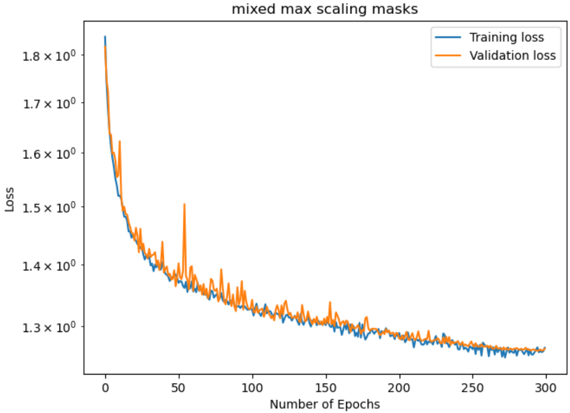
Model training

- 300 epochs
- ≈14 hours
- Modified L1-loss:

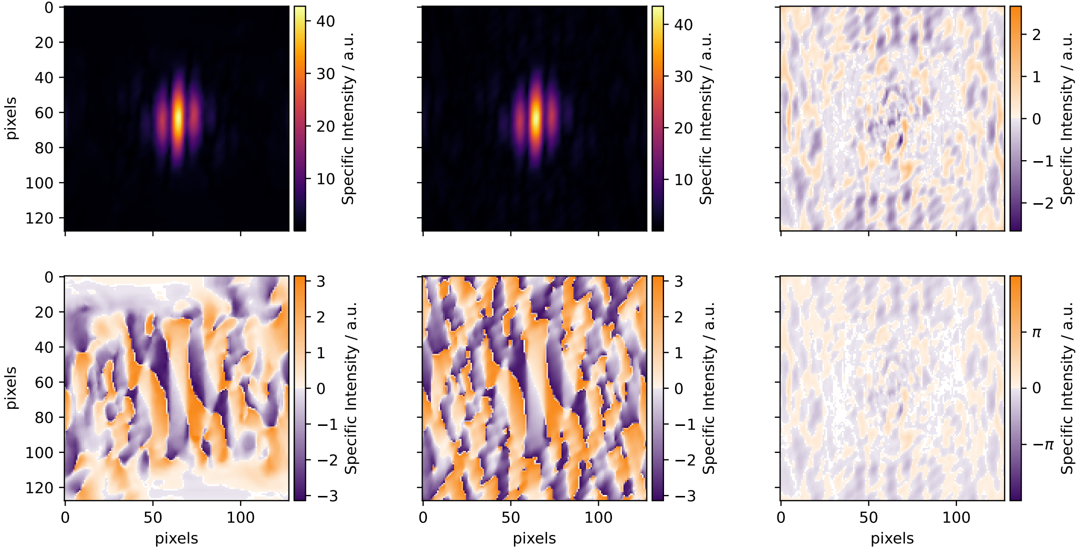
$$\text{Loss} = \text{L1}\left(x_{\text{amp}}, y_{\text{amp}}\right) + \text{L1}\left(\text{HardTanh}(x_{\text{phase}}), y_{\text{phase}}\right),$$

$$\text{with} \;\; \text{L1}(x, y) = |\, x - y \,|,$$

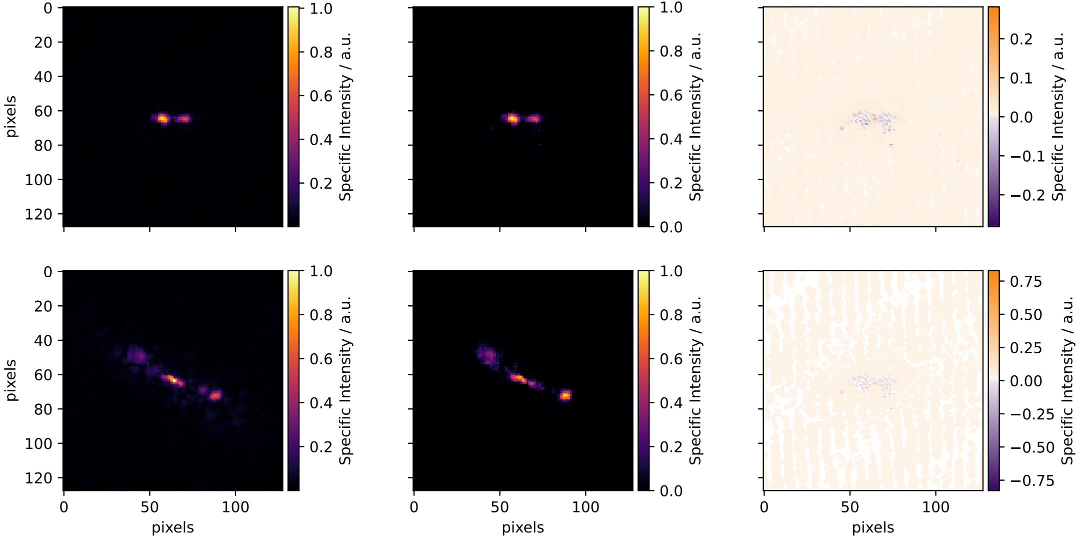
$$\text{and} \;\; \text{HardTanh}(x) = \begin{cases} \pi, & \text{if x} > \pi \\ -\pi, & \text{if x} < -\pi \\ 0, & \text{otherwise} \end{cases},$$



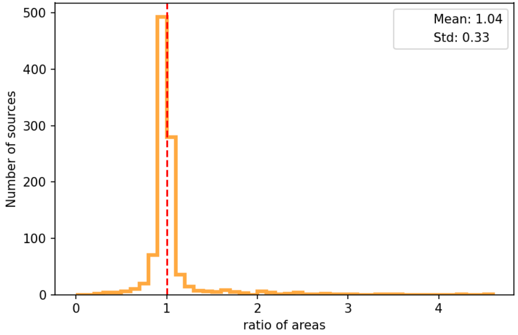
Reconstructions

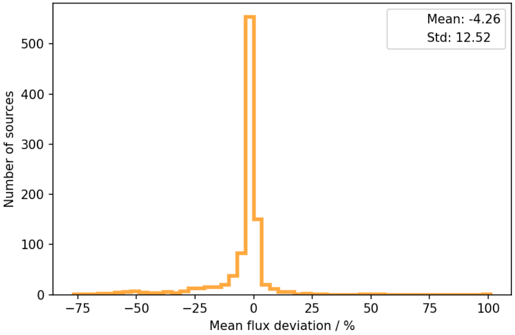


Source reconstructions



Evaluation


Mean: 1.04
Std: 0.33
Number of sources
ratio of areas
0
100
200
300
400
500
0
1
2
3
4

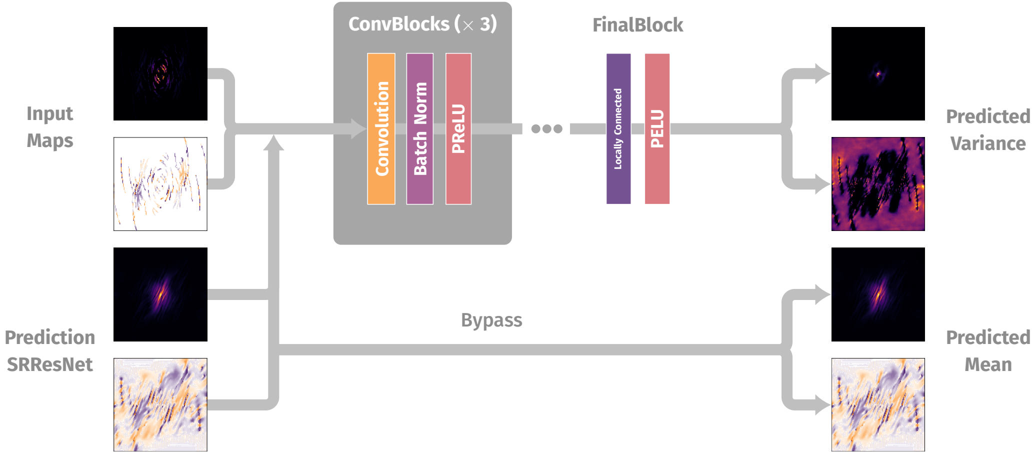

Mean: -4.26
Std: 12.52
Number of sources
Mean flux deviation / %
0
100
200
300
400
500
−75
−50
−25
0
25
50
75
100

Uncertainty estimates

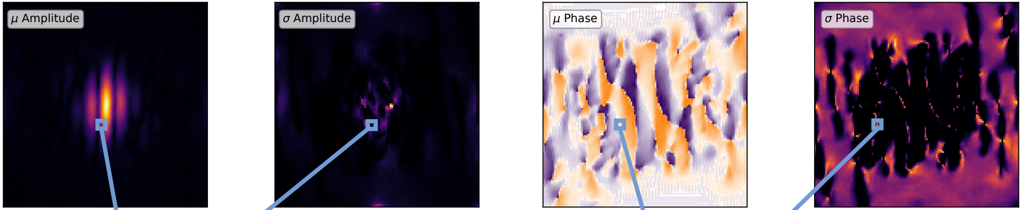
$$L_{\text{NLL}}(\mu, \sigma, y) = \text{stop}(\sigma^2)^{\beta} \cdot \left(\frac{\log(\sigma^2)}{2} + \frac{(\mu - y)^2}{2\sigma^2} \right), \quad \text{with} \;\; \beta = 0.5$$

- Adjusted loss function
- Predict μ and σ^2
- New uncertainty architecture

Uncertainty architecture

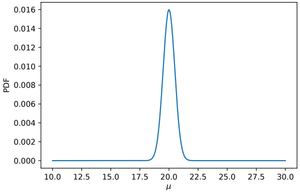


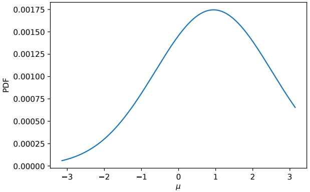
Sampling



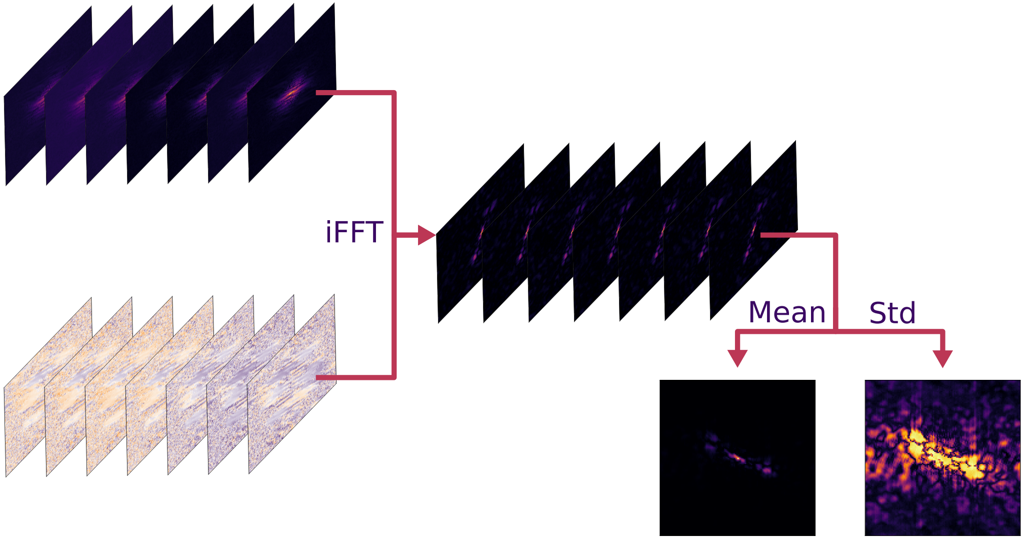
$$g(x|\mu = 20, \sigma = 0.5) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

$$g(x|\mu = 0.3\pi, \sigma = 0.5\pi) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

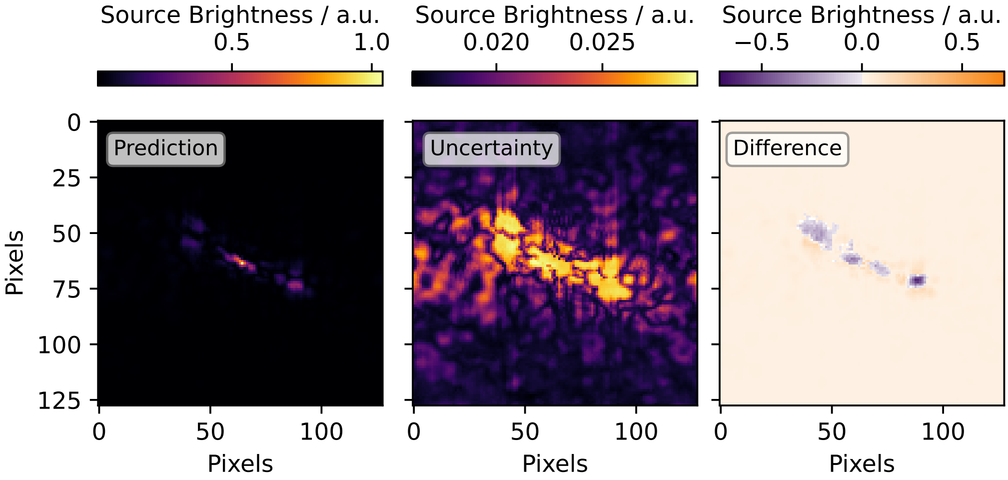




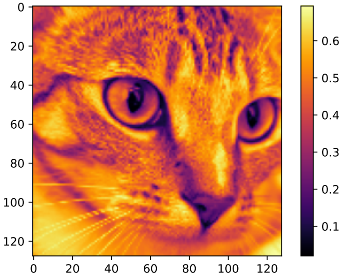
Sampling

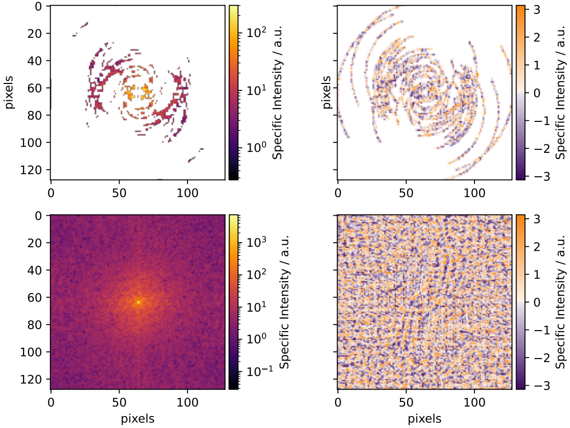


Prediction

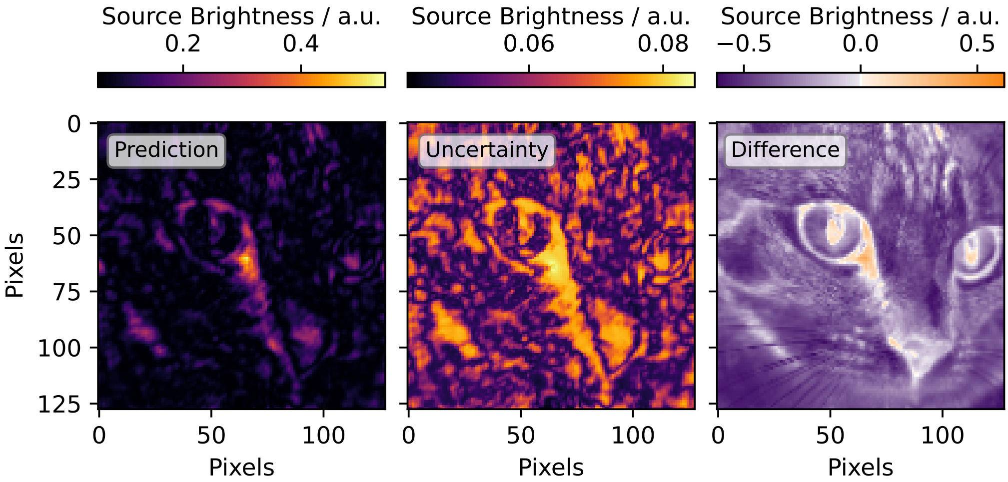

Source Brightness / a.u.
0.5
1.0
Prediction
Source Brightness / a.u.
0.020
0.025
Uncertainty
Source Brightness / a.u.
−0.5
0.0
0.5
Difference
Pixels
0
25
50
75
100
125
0
50
100
Pixels

Untrained source shapes



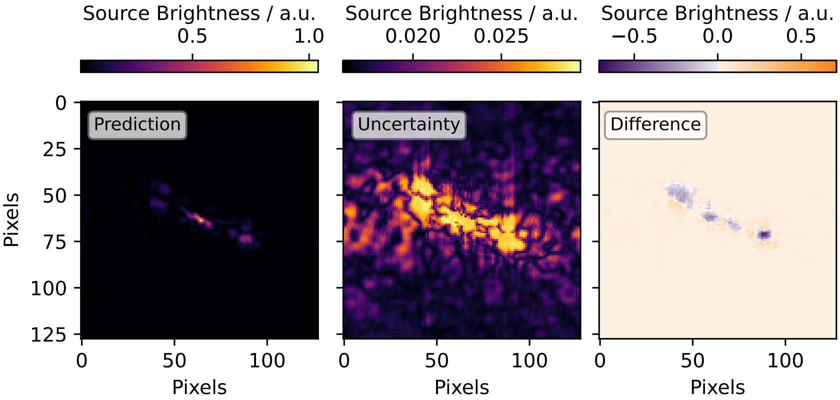


Prediction cat



Summary

- Easy integration into framework
- Training works well without major adaptations
- Bright pixel in the center of the prediction
- Good performance when used with sampling masks



Outlook

- Enhance evaluation methods
- Updated GAN model
- Improved simulations with `RIME`