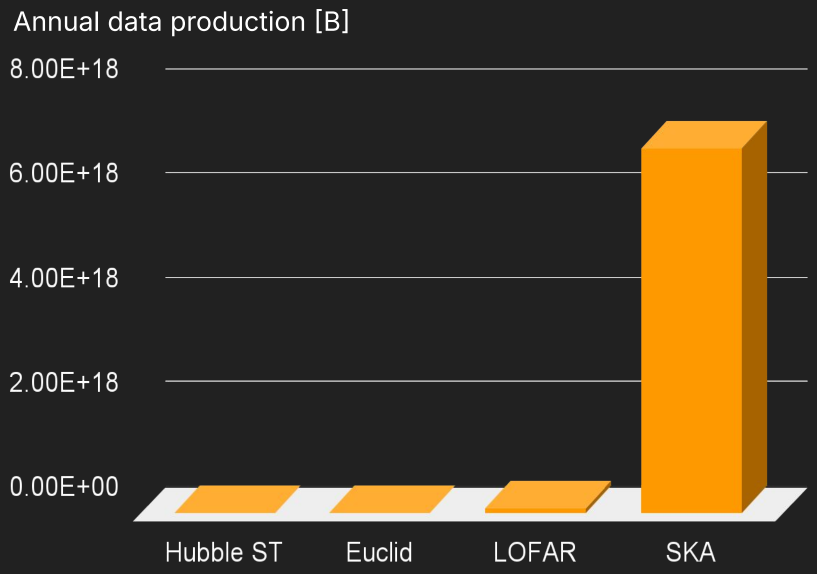
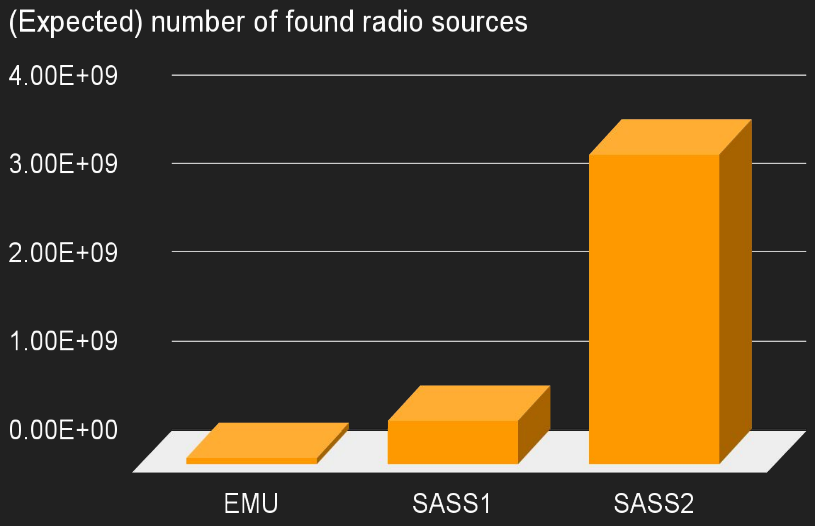
Radio galaxy classification with scattering-transform-based generative augmentation

Markus Bredberg
markus.bredberg@epfl.ch
PhD student, LASTRO, EPFL,
Supervisor: Emma Tolley
27th of January 2025

With great data comes great responsibilities



Credit: NASA, ESA, IRA, SKAO



Credit: Norris et al., 2011; Norris et al., 2014;

In the context of my PhD

1 — 2 — 3

Radio galaxy generation and classification

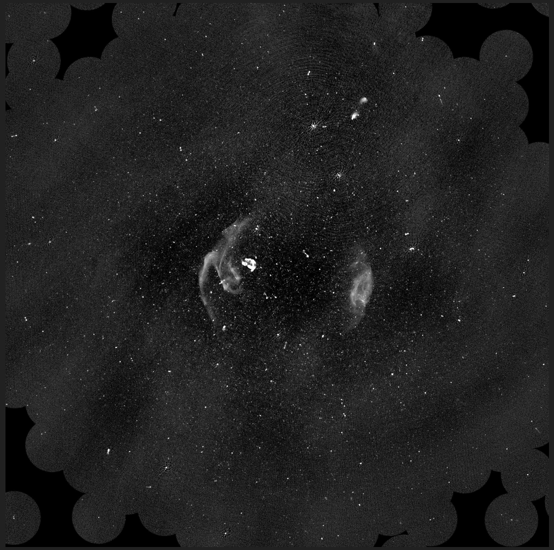
- FIRST Data

Credit: FIRST dataset, Rustige+23

Diffuse cluster radio emission generation and classification

- MGCLS and MERGHERS data
- X-ray, optical, polarisation, spectral index

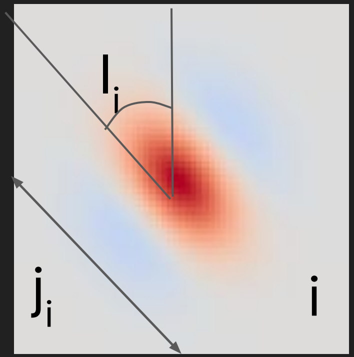
Early-stage analysis



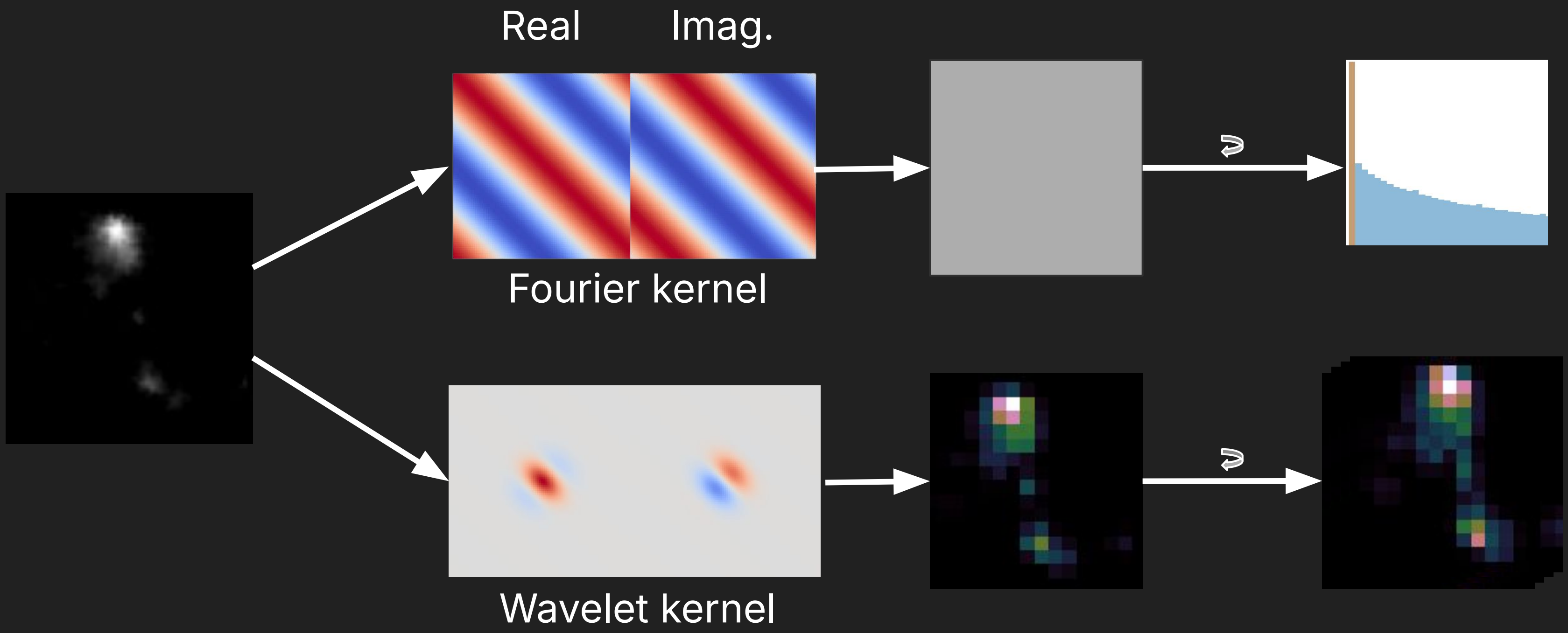
Courtesy: Konstantinos Kolokythas

Morlet wavelets are sinusoids with Gaussian envelopes

$$\psi_{j,\ell}(u) = 2^{-2j}\psi(2^{-j} r_\theta u) \quad \text{for} \quad 0 \leq \ell < L,\ \theta = \ell\pi/L,\ 0 \leq j < J$$



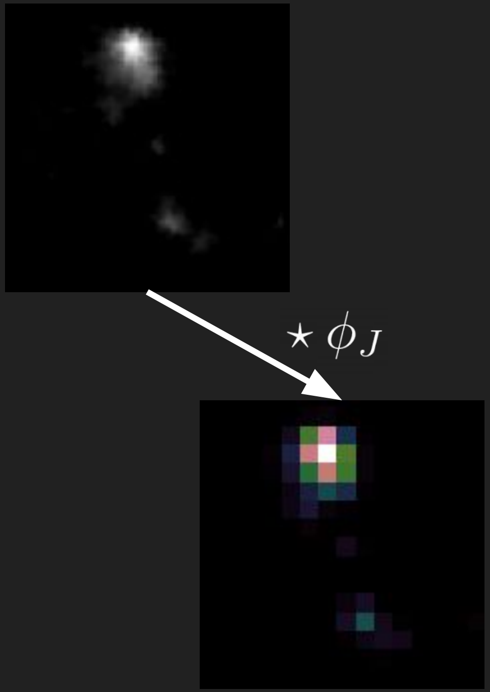
Wavelet kernels are localised and extract features



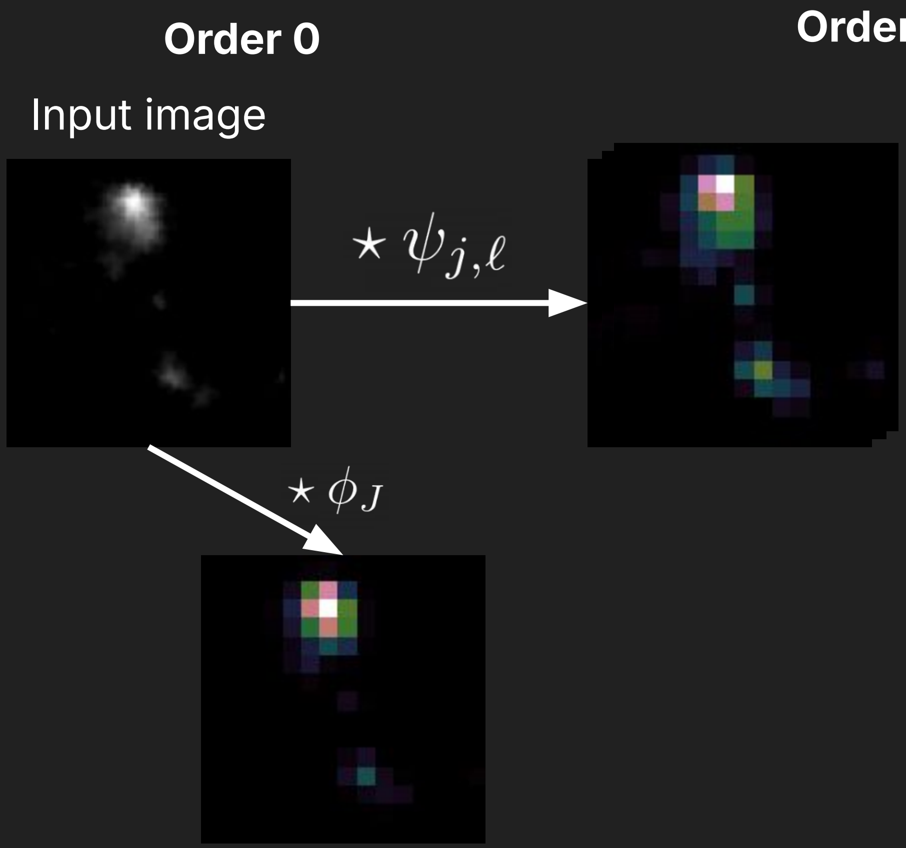
Scattering transform is a cascade of wavelet transforms

Order 0

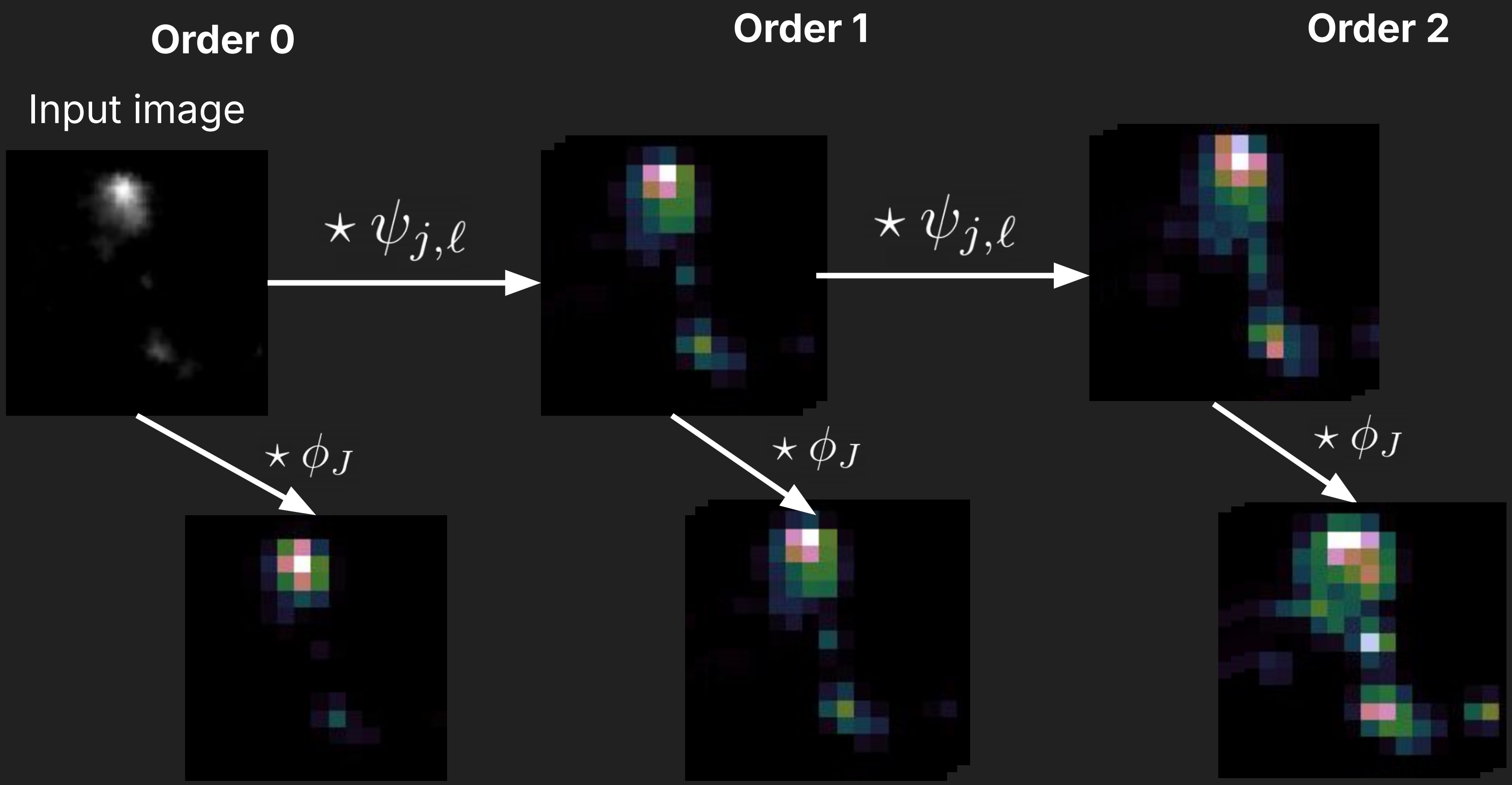
Input image



Scattering transform is a cascade of wavelet transforms



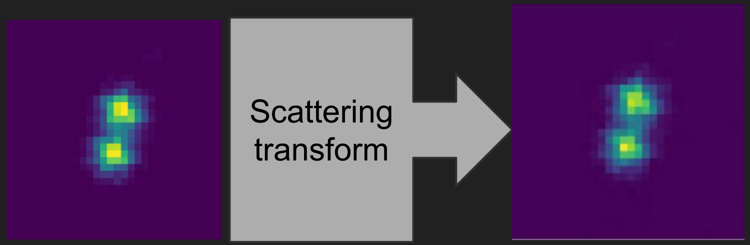
Scattering transform is a cascade of wavelet transforms



Multidimensional dependency require variable mixing

Tensor fitting

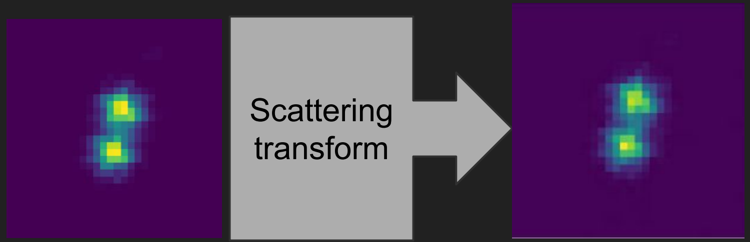
Scattering coefficients are dependent



Multidimensional dependency require variable mixing

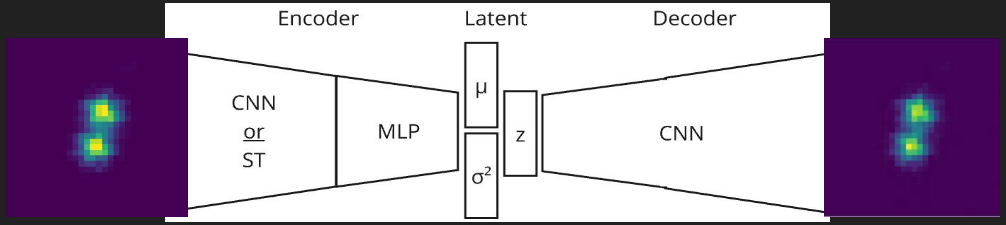
Tensor fitting

Scattering coefficients are dependent



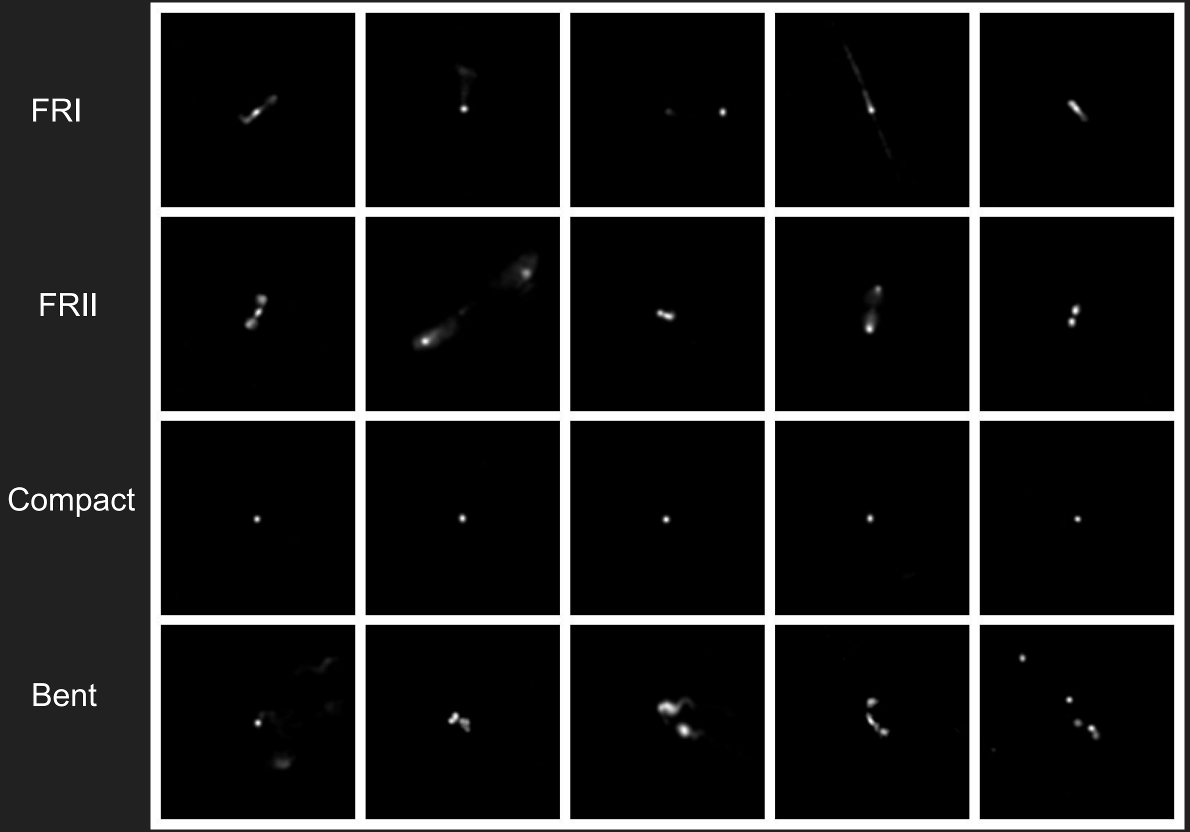
Variational Autoencoder

Latent variables are independent



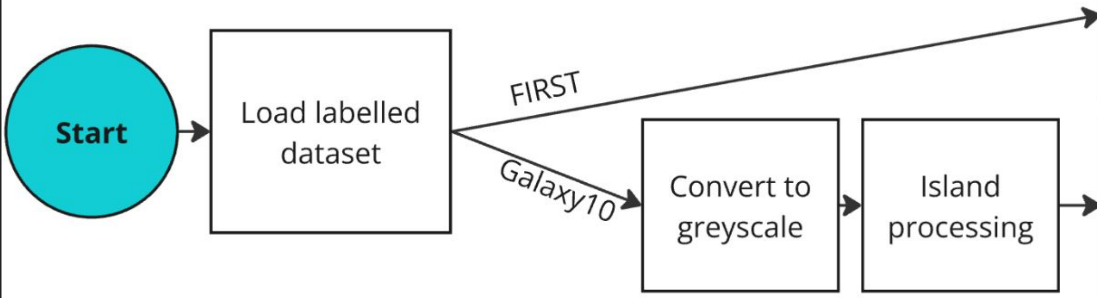
FIRST dataset is a popular radio galaxy benchmark

- 2158 radio galaxies
  - FRI: 495
  - FRII: 924
  - Compact: 391
  - Bent: 348
- 300x300 pixels (greyscale)
- No noise or artefacts

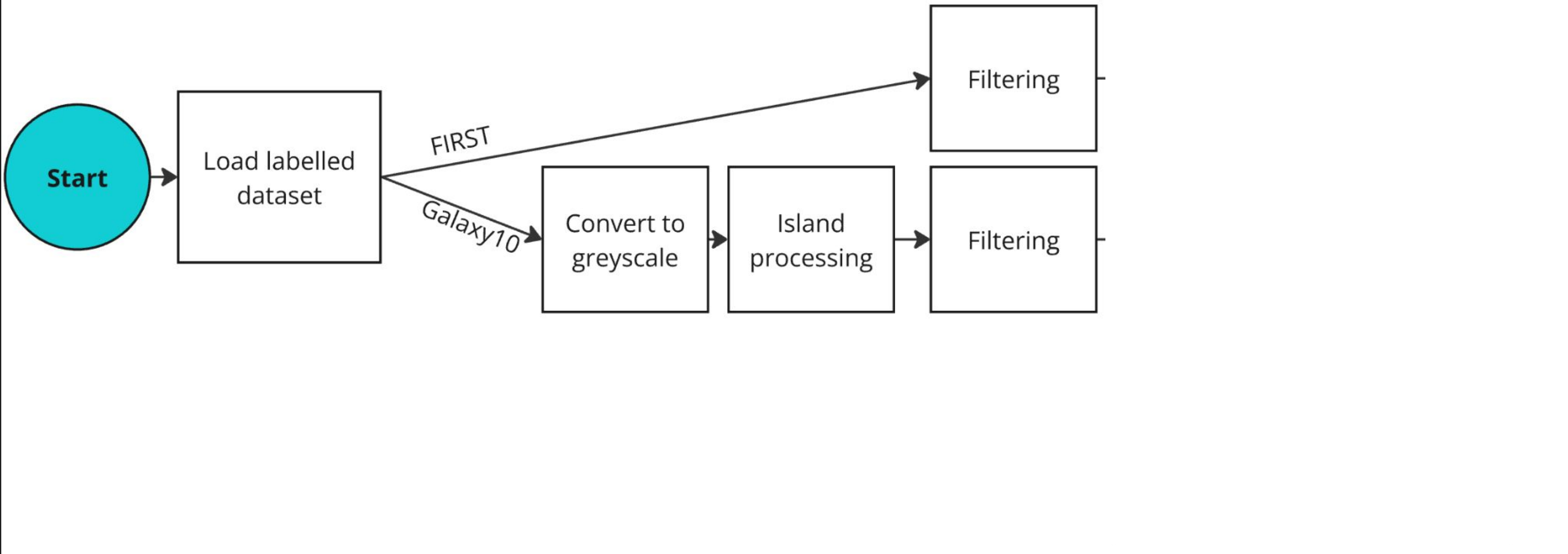


Credit: FIRST dataset, Griese et al. 2022

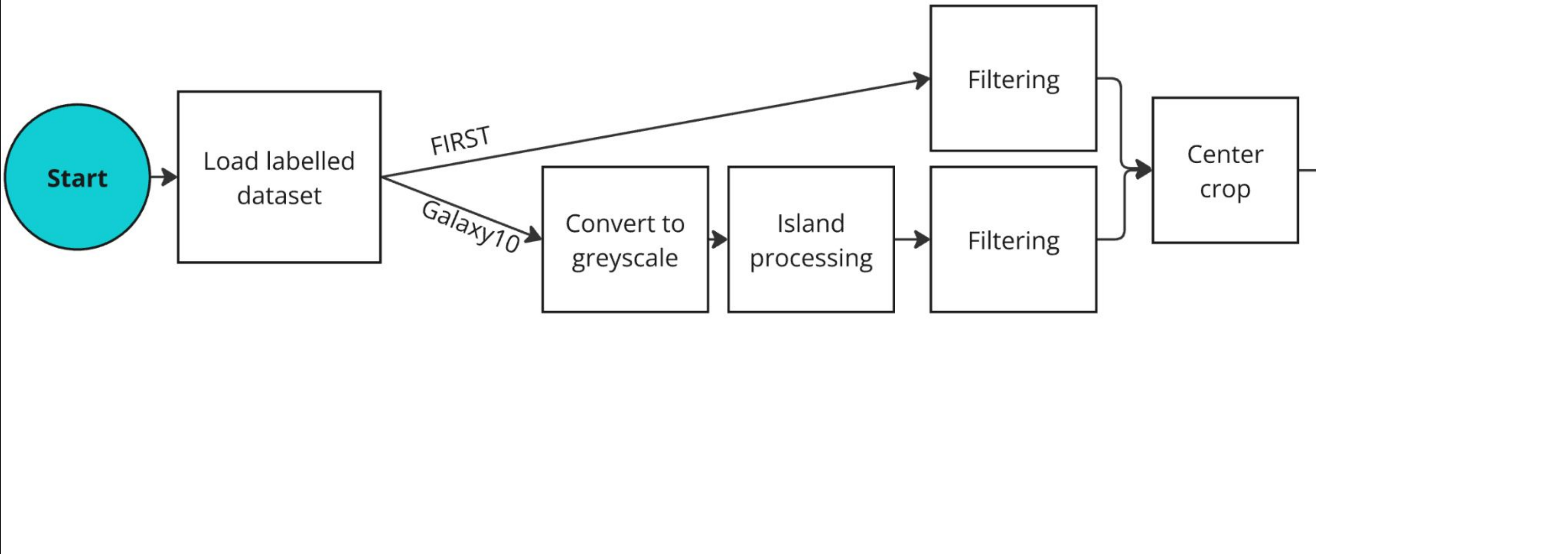
Workflow pipeline for generative modelling



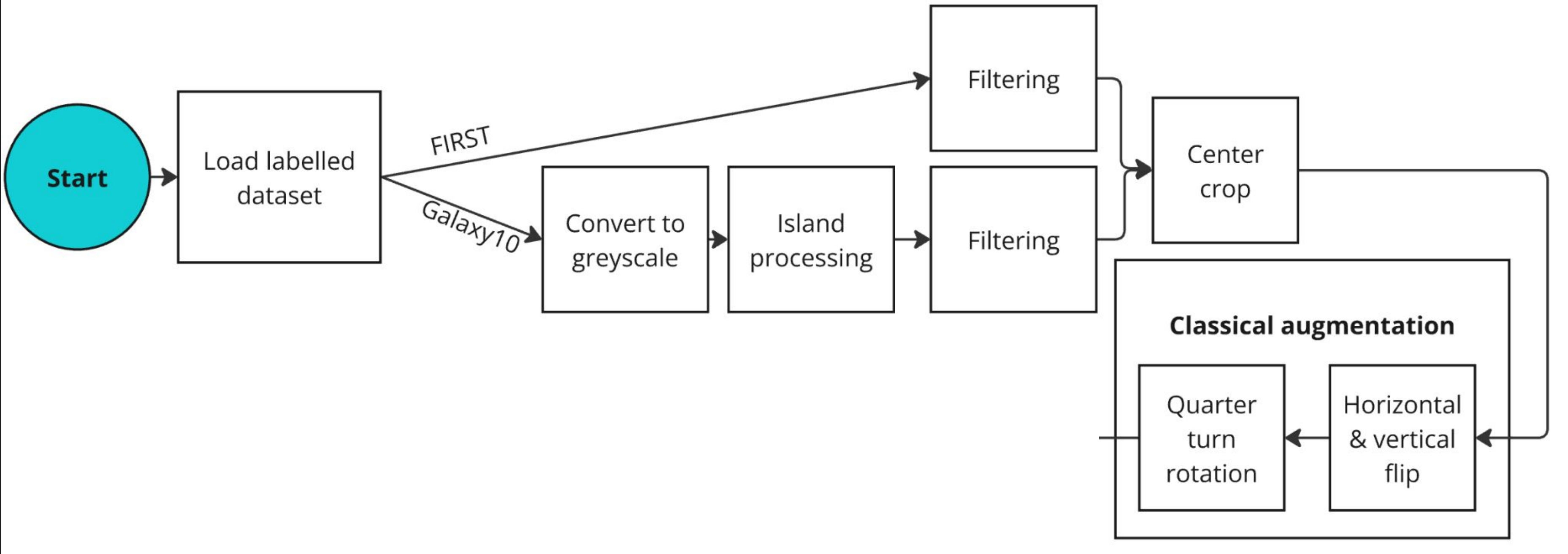
Workflow pipeline for generative modelling



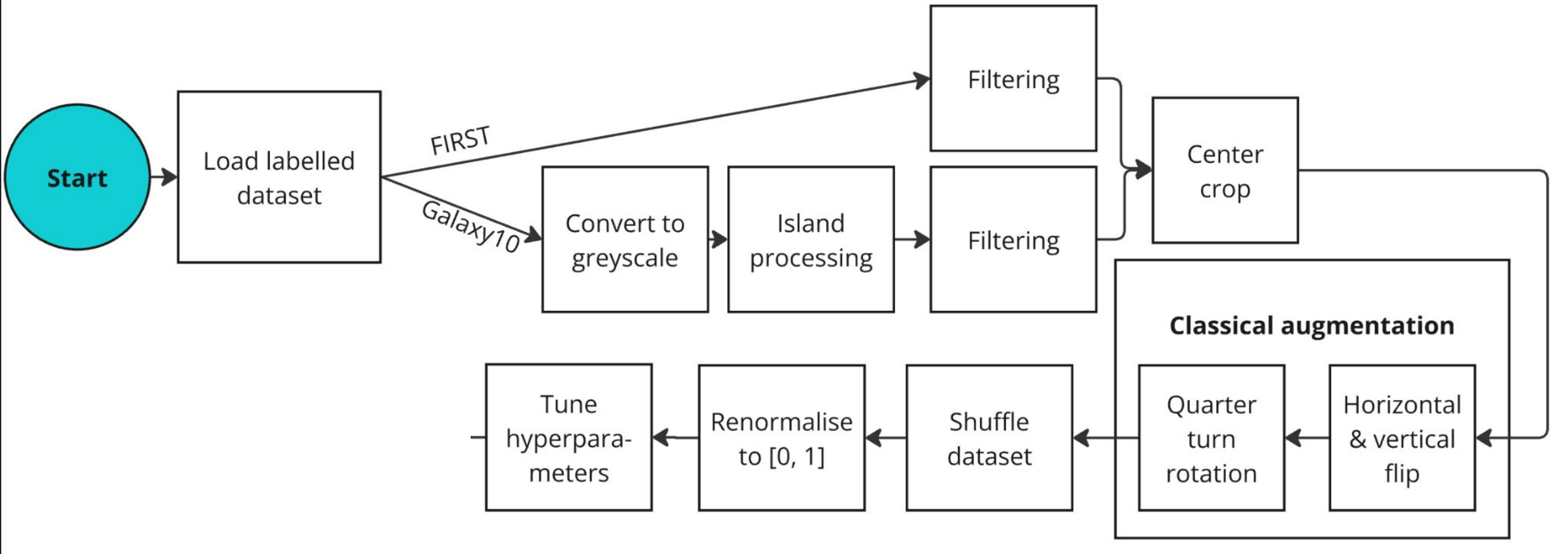
Workflow pipeline for generative modelling



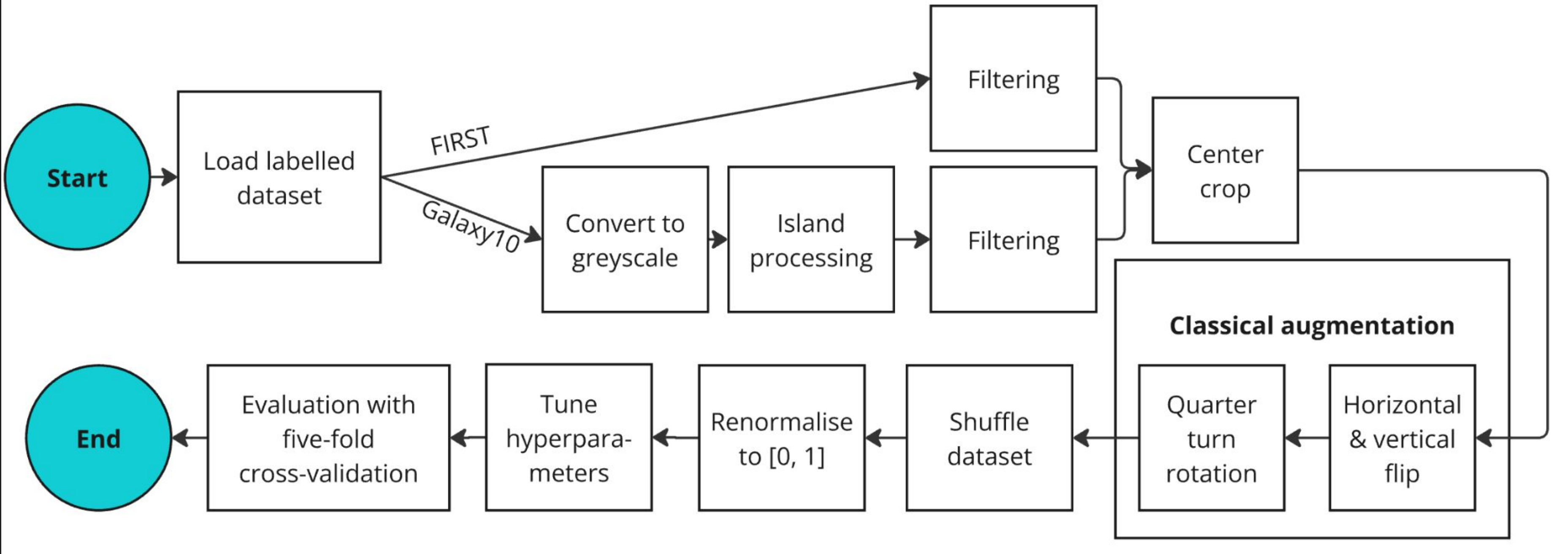
Workflow pipeline for generative modelling



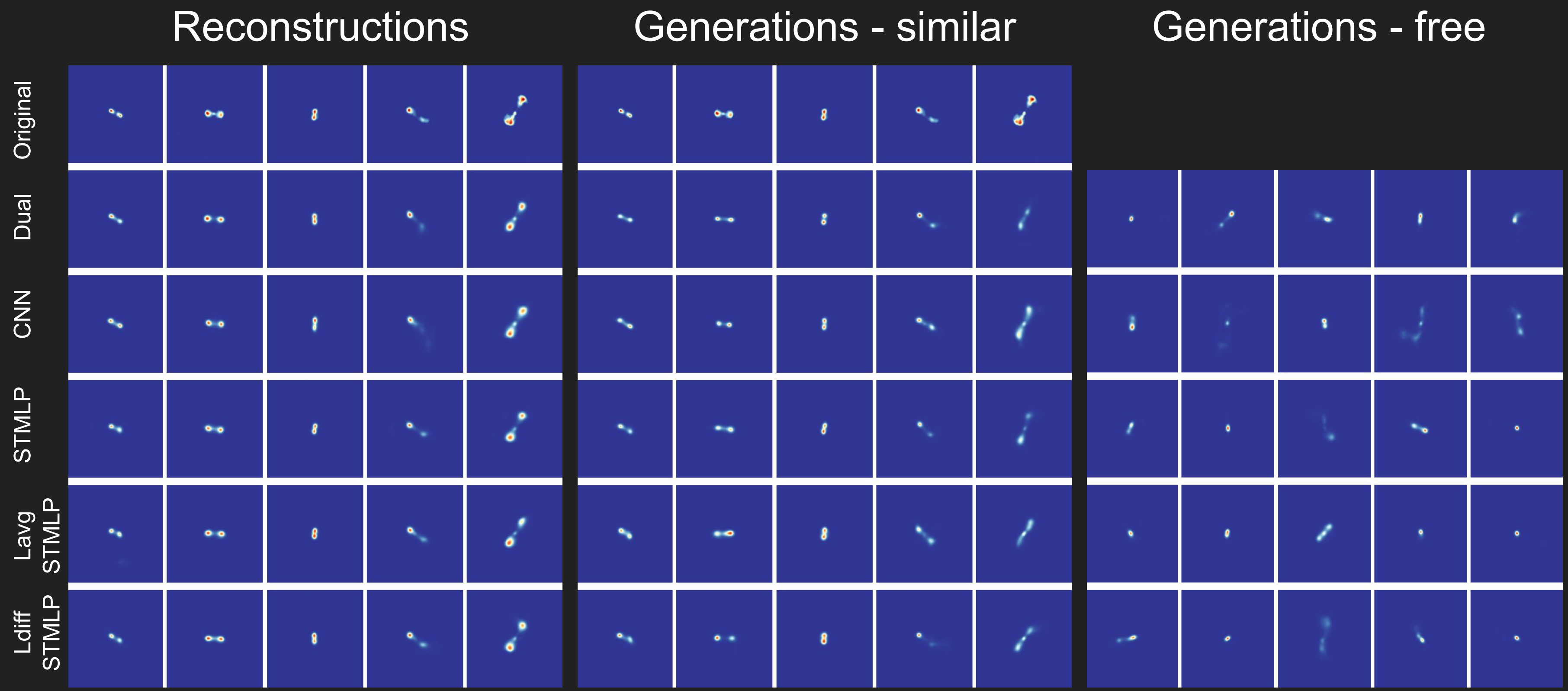
Workflow pipeline for generative modelling



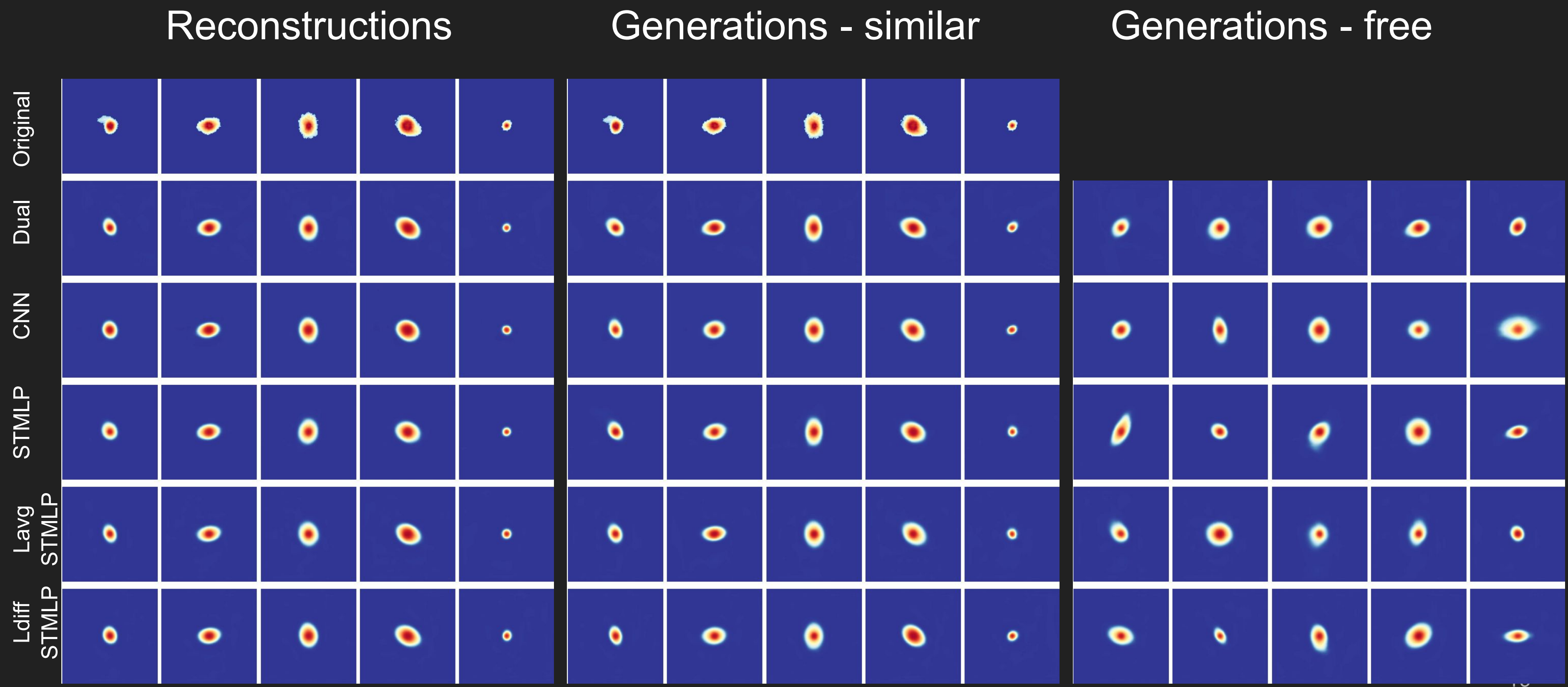
Workflow pipeline for generative modelling



Models capture general features of FRII galaxies

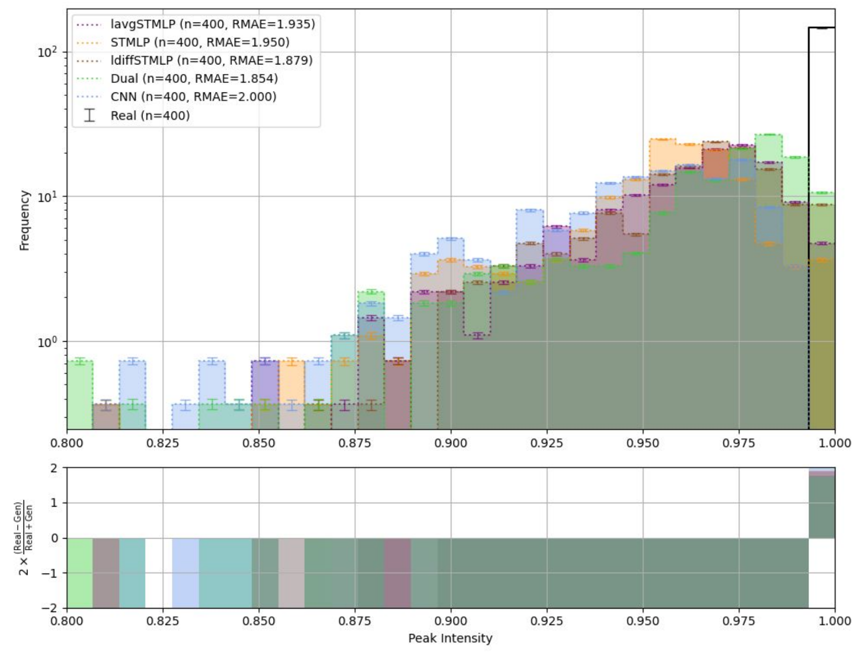


Galaxy10 provides less sparse sources and is easier



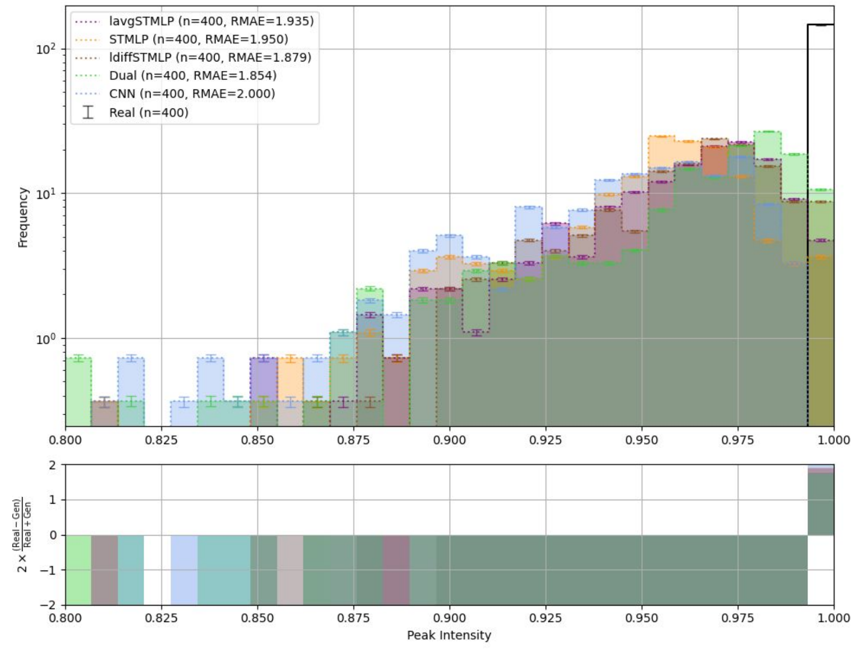
The VAEs capture Galaxy10 peak pixels well

Reconstructions of Galaxy10

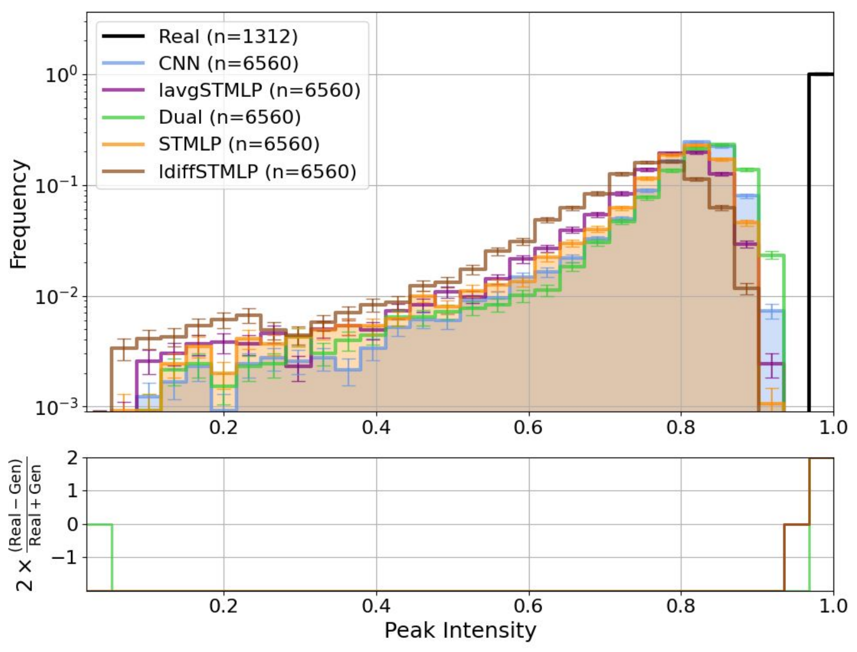


The VAEs capture Galaxy10 peak pixels well

Reconstructions of Galaxy10

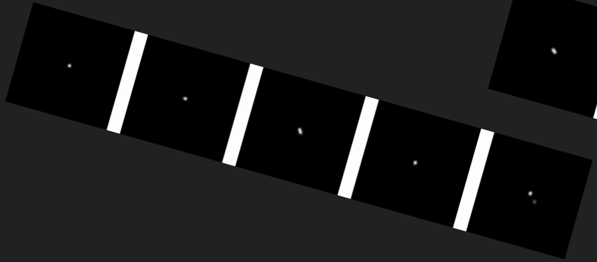


Reconstructions of FIRST

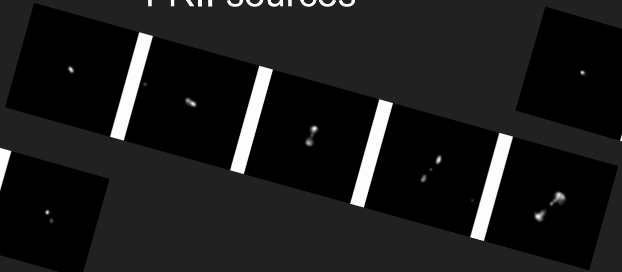


Peak intensity is worse for irregular sources

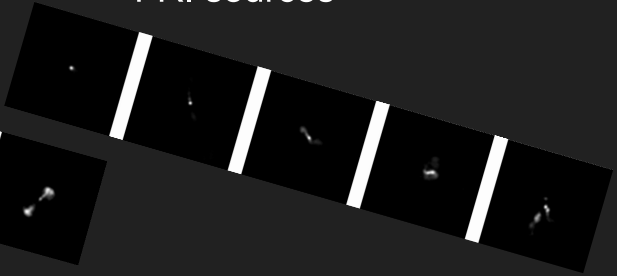
Compact sources



FRII sources



FRI sources



Peak intensity is worse for irregular sources

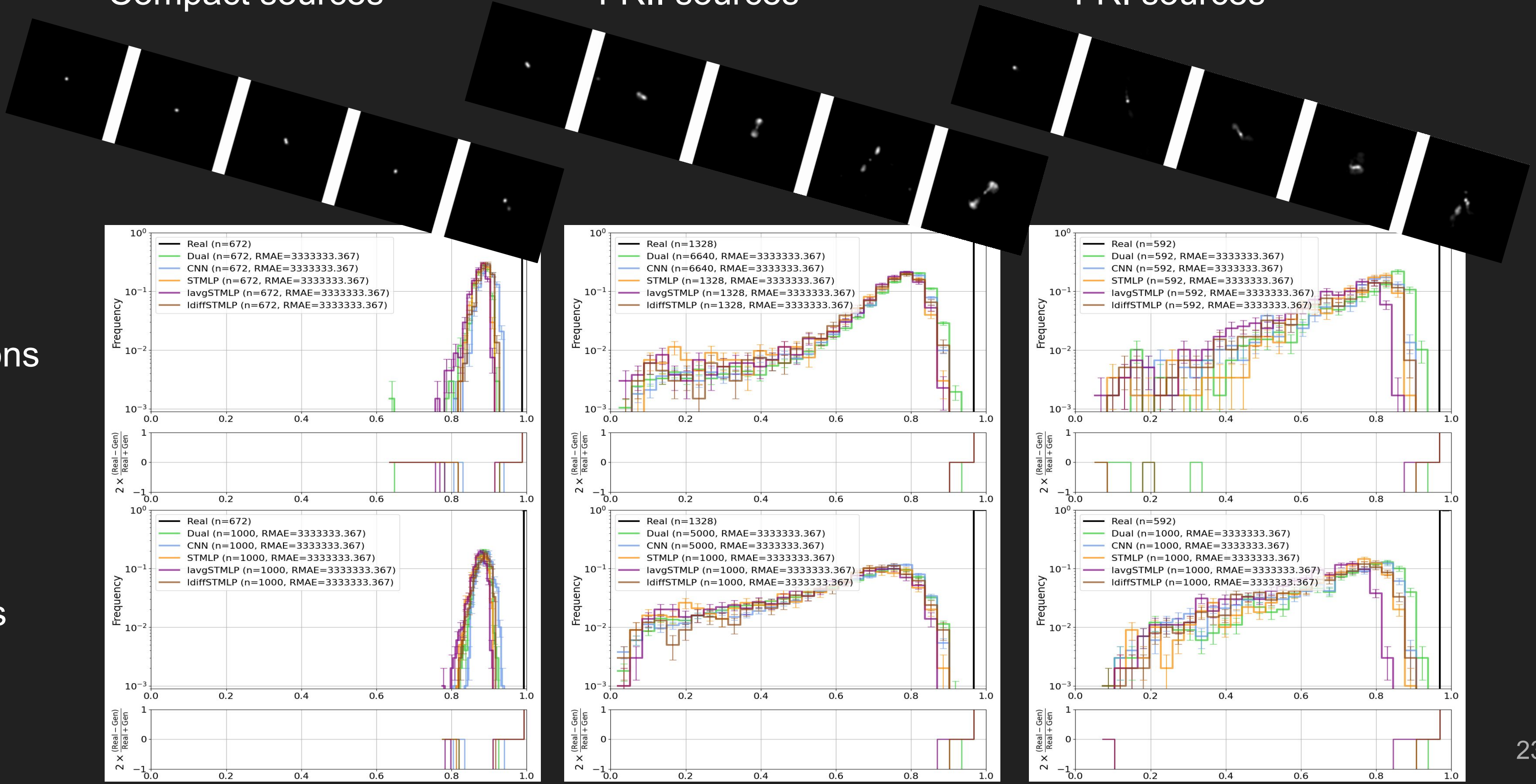
Compact sources

FRII sources

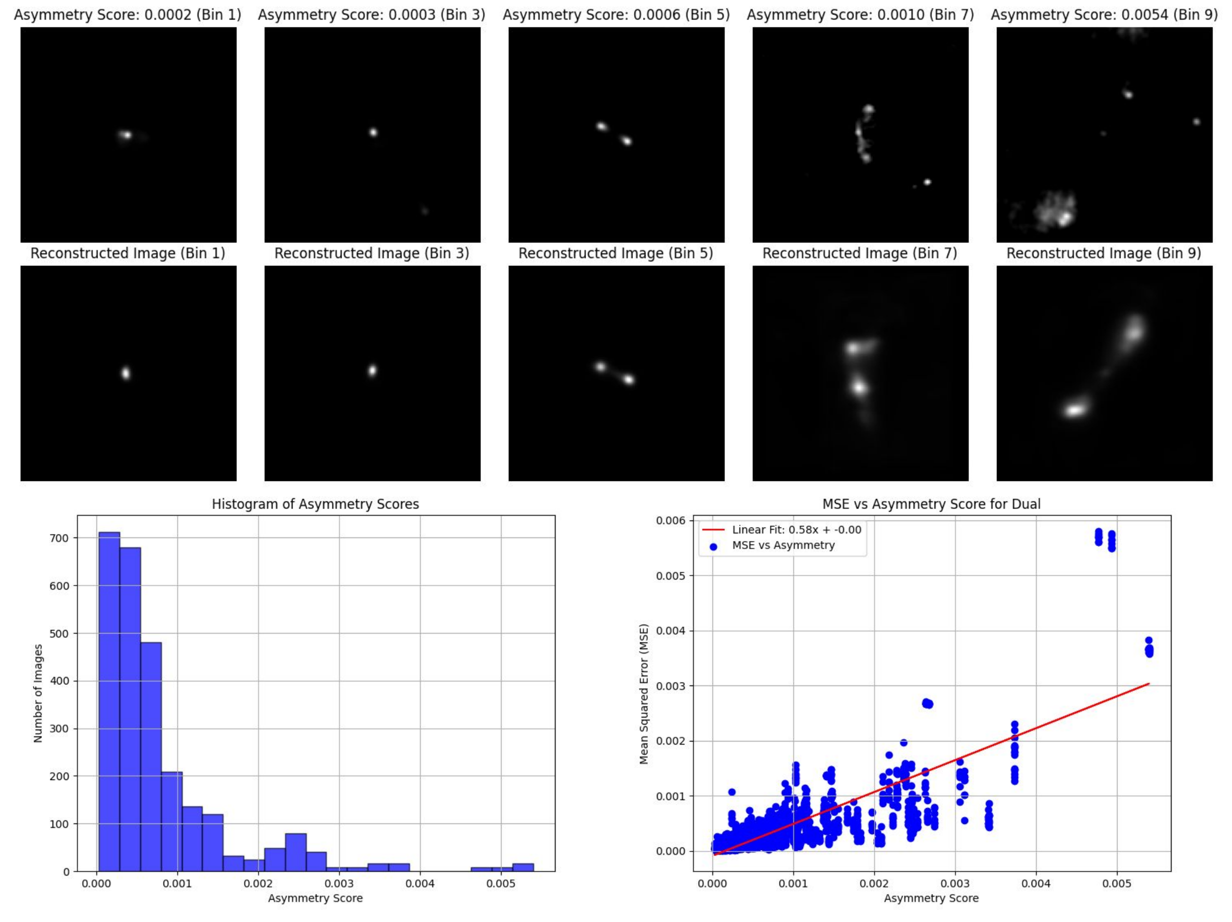
FRI sources

Reconstructions

Generations



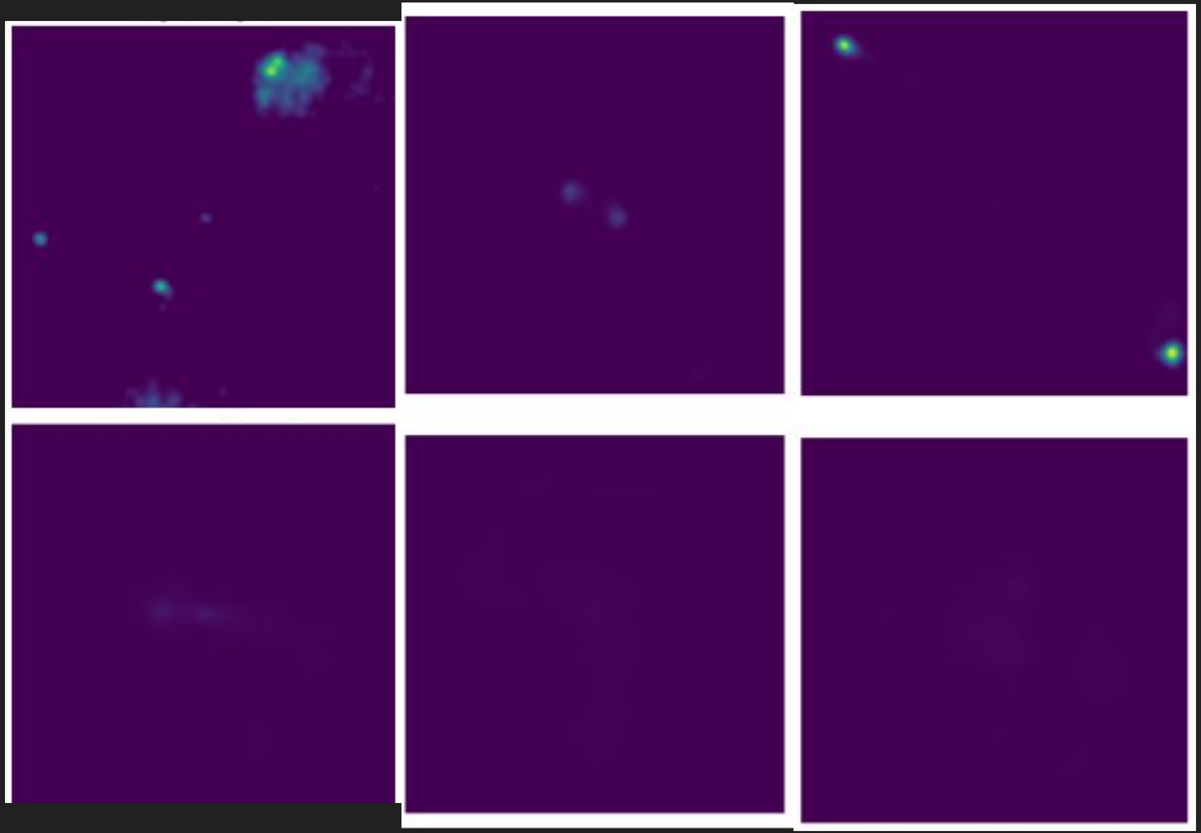
MSE scales with asymmetry



Difficult reconstructions are asymmetric, faint and spread out

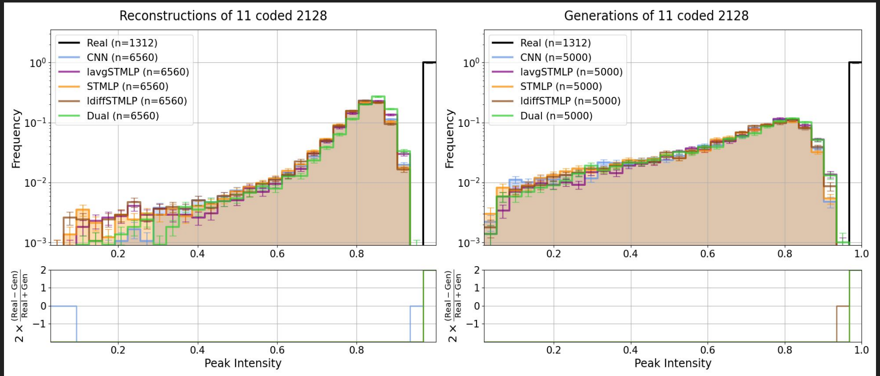
Originals

Reconstructions

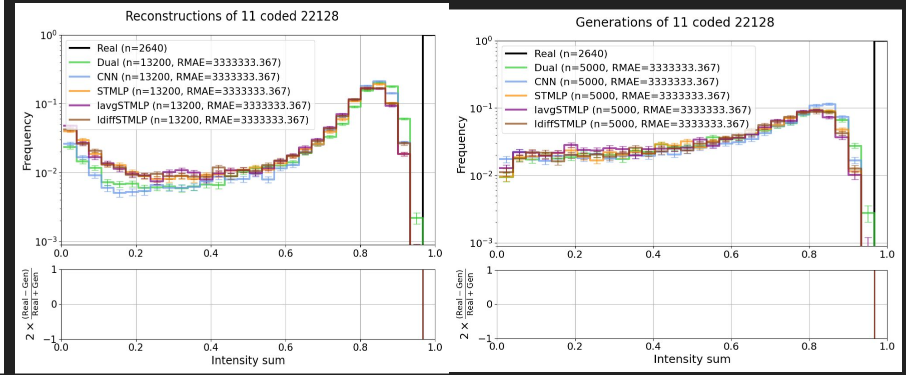


Filtering reduces risk of faint constructions

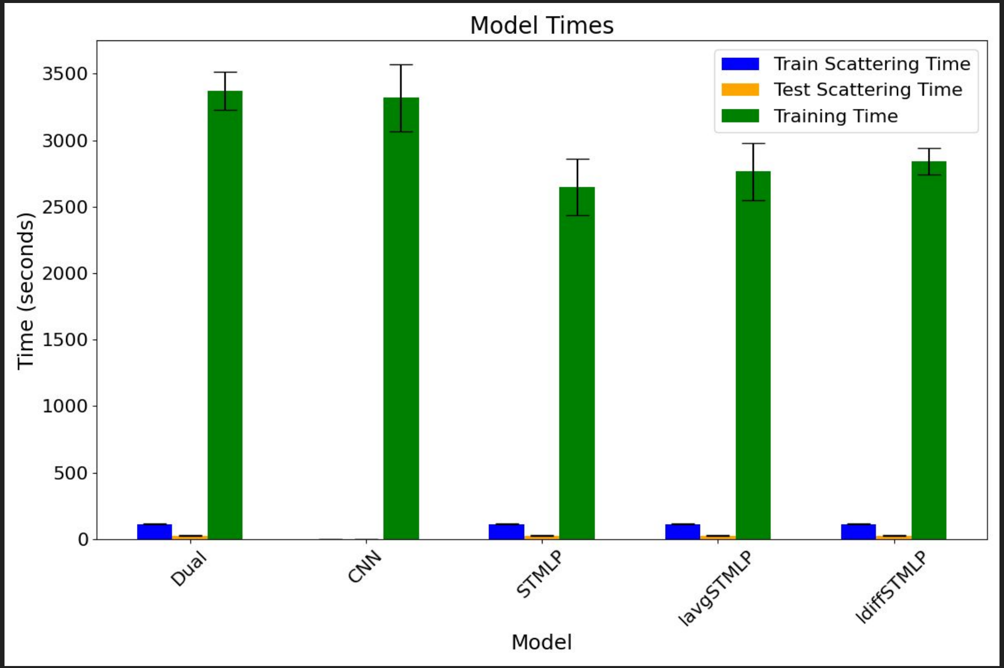
Filtered



Not filtered

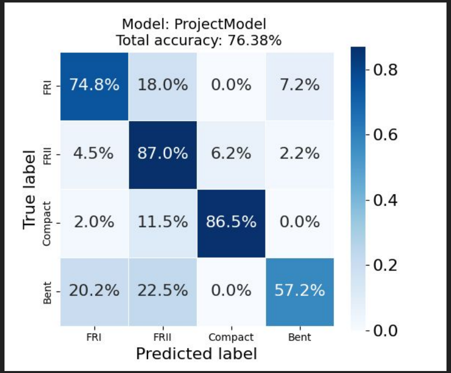


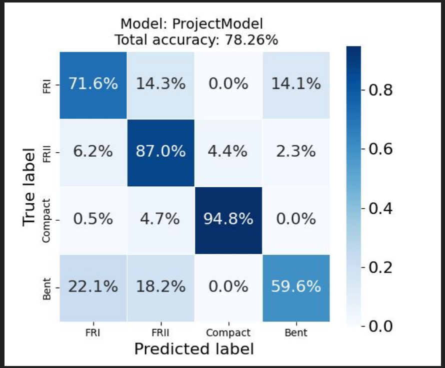
Scattering transform saves training time

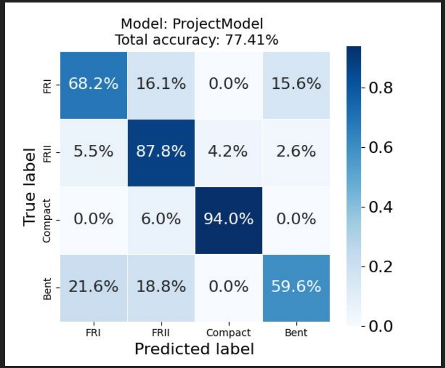


Using Dual VAE no classification improvement is made

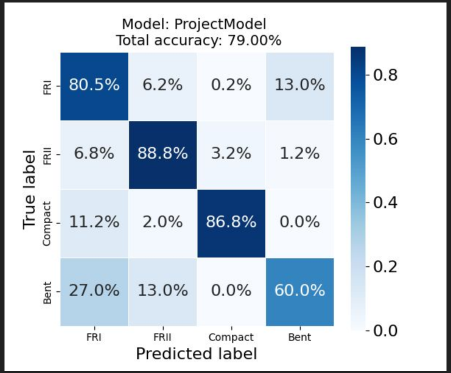
λ=0

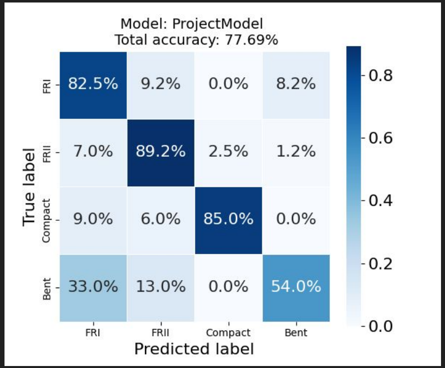


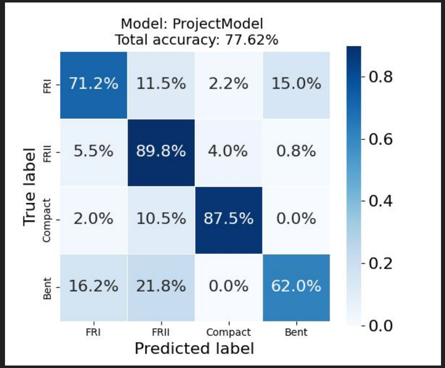




λ=1

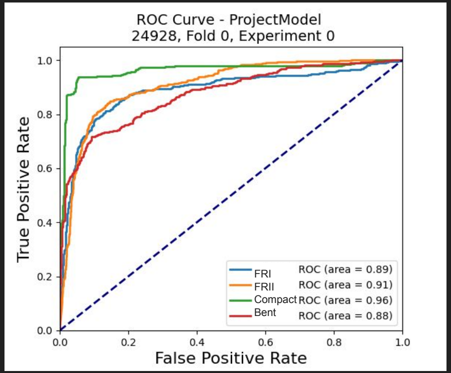


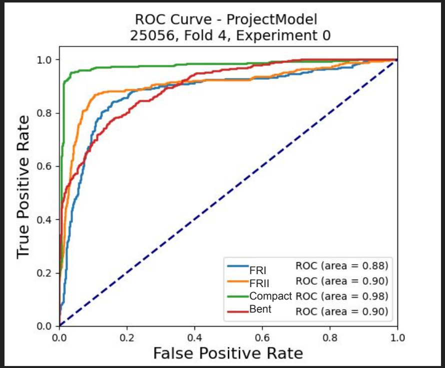


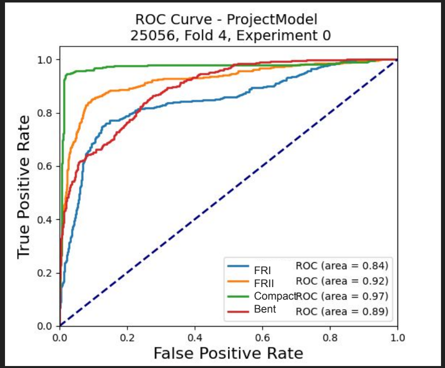


Generated images facilitate classification of FRII and bent sources

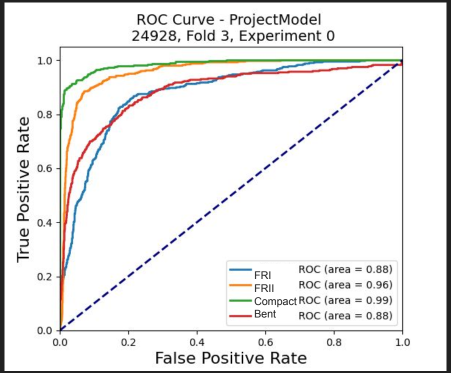
λ=0

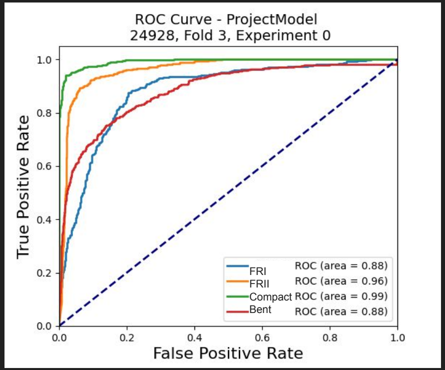


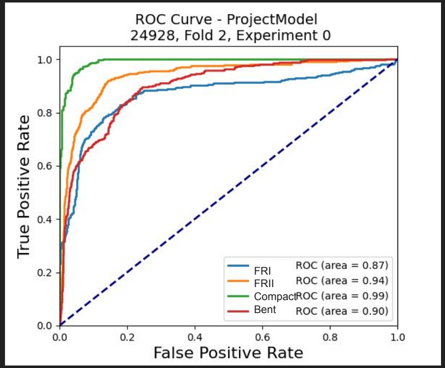




λ=1







Conclusions

- VAE-generated images are smooth, faint, but overall realistic
- Classification performance seems to be only marginally improved by artificial augmentation
- More tests with different VAEs and levels of artificial generation will be made
- This method will soon be applied on diffuse cluster radio emission

Thank you!
Questions?

Markus Bredberg

markus.bredberg@epfl.ch

Backup slides

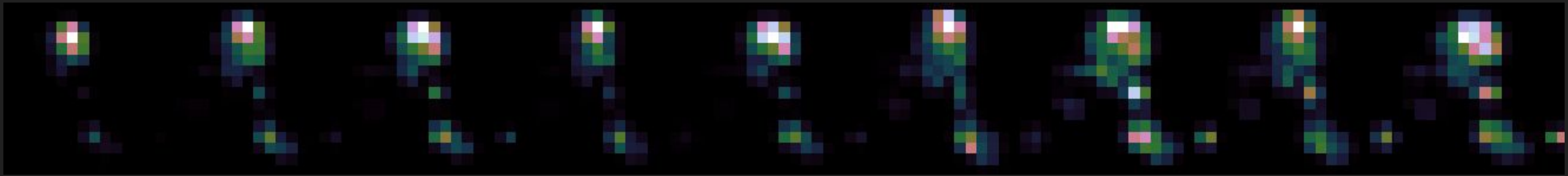
Scattering transforms are iterative wavelet transforms

$$\Psi_{J,L}(x) = \left[X, \left|X \star \psi_{j,\ell}\right|, \left|\left|X \star \psi_{j,\ell}\right| \star \psi_{j',\ell'}\right|\right]_{1 \le \ell, \ell' \le L,\ 1 \le j < j' \le J}$$

$$\phi_J(u) = 2^{-2J} \phi(2^{-J} u)$$

$$\begin{aligned} S_{J,L}(X) &= \Psi_{J,L}(X) \star \phi_J \\ &= \left[X \star \phi_J, \left|X \star \psi_{j,\ell}\right| \star \phi_J, \left|\left|X \star \psi_{j,\ell}\right| \star \psi_{j',\ell'}\right| \star \phi_J\right]_{1 \le j < j' \le J, 1 \le \ell, \ell' \le L} \end{aligned}$$

$X \star \phi_2$ $\quad |X \star \psi_{0,0}| \star \phi_2$ $\quad |X \star \psi_{0,1}| \star \phi_2$ $\quad |X \star \psi_{1,0}| \star \phi_2$ $\quad |X \star \psi_{1,1}| \star \phi_2$ $\quad ||X \star \psi_{0,0}| \star \psi_{1,0}| \star \phi_2$ $\quad ||X \star \psi_{0,0}| \star \psi_{1,1}| \star \phi_2$ $\quad ||X \star \psi_{0,1}| \star \psi_{1,0}| \star \phi_2$ $\quad ||X \star \psi_{0,1}| \star \psi_{1,1}| \star \phi_2$



Low pass filter for scattering transform

Littlewood-Paley equality $$\sum_{j=-\infty}^{+\infty} |\hat{\psi}(2^j\omega)|^2 = 1 \ , \ \forall \omega > 0$$

Conservation of energy

Capture lower frequencies of signal $[-2^{-J}\pi, 2^{-J}\pi]$

With low-pass filter $\hat{\Phi}_J(\omega) = \left(\sum_{j=J+1}^{+\infty} |\hat{\psi}(2^j\omega)|^2 \right)^{1/2}$ **satisfying** $\int \Phi_J(t)dt = 1$

$\rightarrow$ **preserves norm and is therefore invertible**

The scattering transform is an iterative wavelet transform

Wavelet transform $WX = X \star \psi_{j,l}(t)$

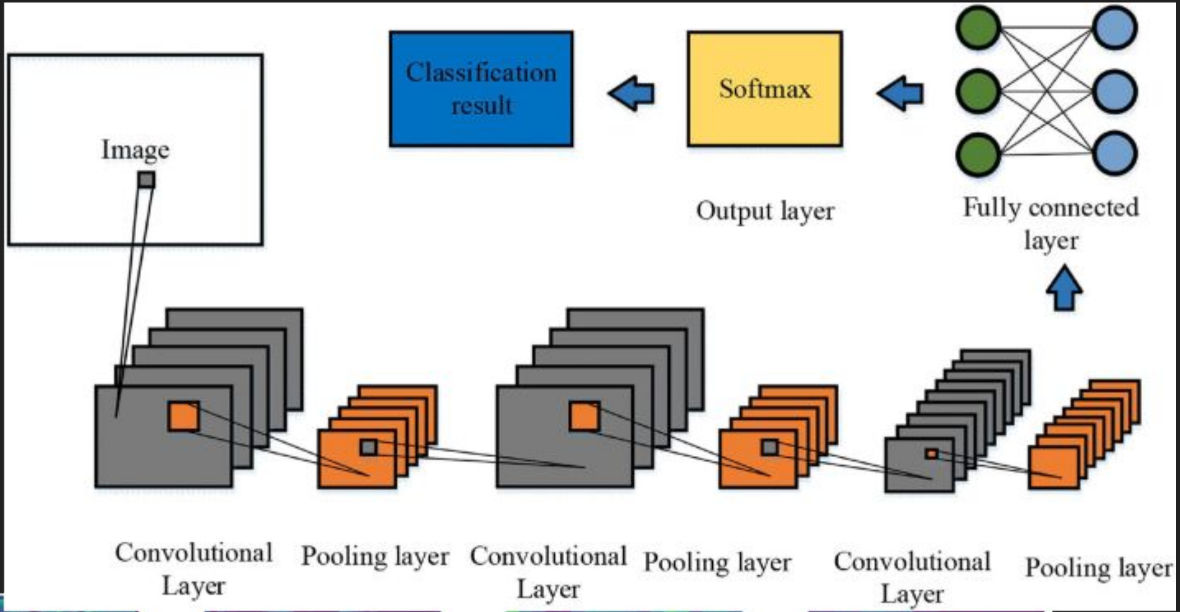
Scattering transform $SX = W|WX|$

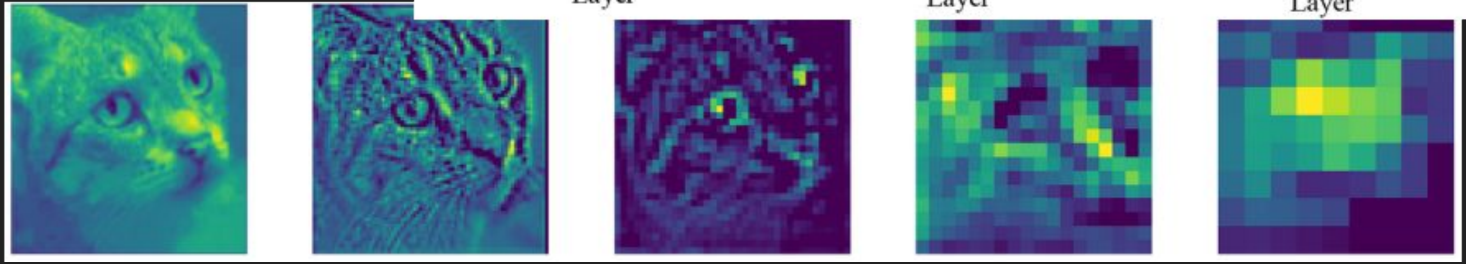
$$S_0 = X \star \Phi$$

$$S_1(j_1, l_1) = \langle |X \star \psi_{j_1,l_1}| \star \Phi \rangle$$

$$S_2(j_1, j_2, l_1, l_2) = \langle ||X \star \psi_{j_1,l_1}| \star \psi_{j_2,l_2}|\Phi \rangle$$

Convolutional neural networks encodes features



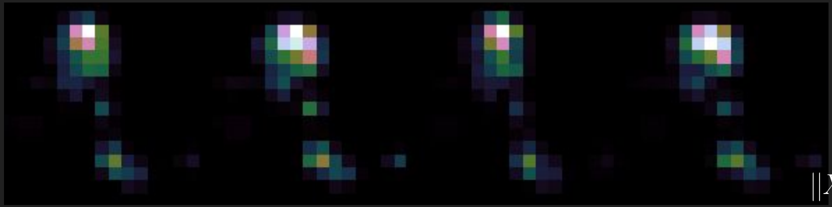


Scattering transform is an iterative wavelet transform

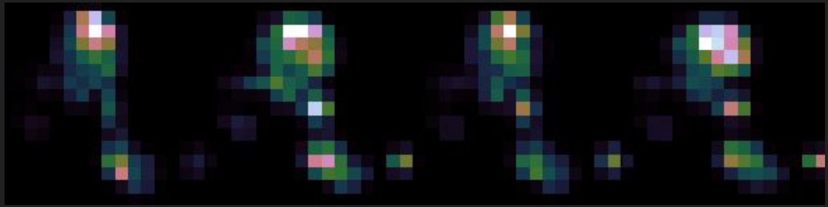
$$
\begin{aligned}
S_{J,L}(X) &= \Psi_{J,L}(X) \star \phi_J \\
&= [\underbrace{X \star \phi_J}_{\text{Order 0}}, \underbrace{|X \star \psi_{j,\ell}| \star \phi_J}_{\text{Order 1}}, \underbrace{||X \star \psi_{j,\ell}| \star \psi_{j',\ell'}| \star \phi_J}_{\text{Order 2}}]_{1 \le j < j' \le J, 1 \le \ell, \ell' \le L}
\end{aligned}
$$

$X \star \phi_2$

$|X \star \psi_{0,0}| \star \phi_2$ $|X \star \psi_{0,1}| \star \phi_2$ $|X \star \psi_{1,0}| \star \phi_2$ $|X \star \psi_{1,1}| \star \phi_2$



$||X \star \psi_{0,0}| \star \psi_{1,0}| \star \phi_2$ $||X \star \psi_{0,0}| \star \psi_{1,1}| \star \phi_2$ $||X \star \psi_{0,1}| \star \psi_{1,0}| \star \phi_2$ $||X \star \psi_{0,1}| \star \psi_{1,1}| \star \phi_2$

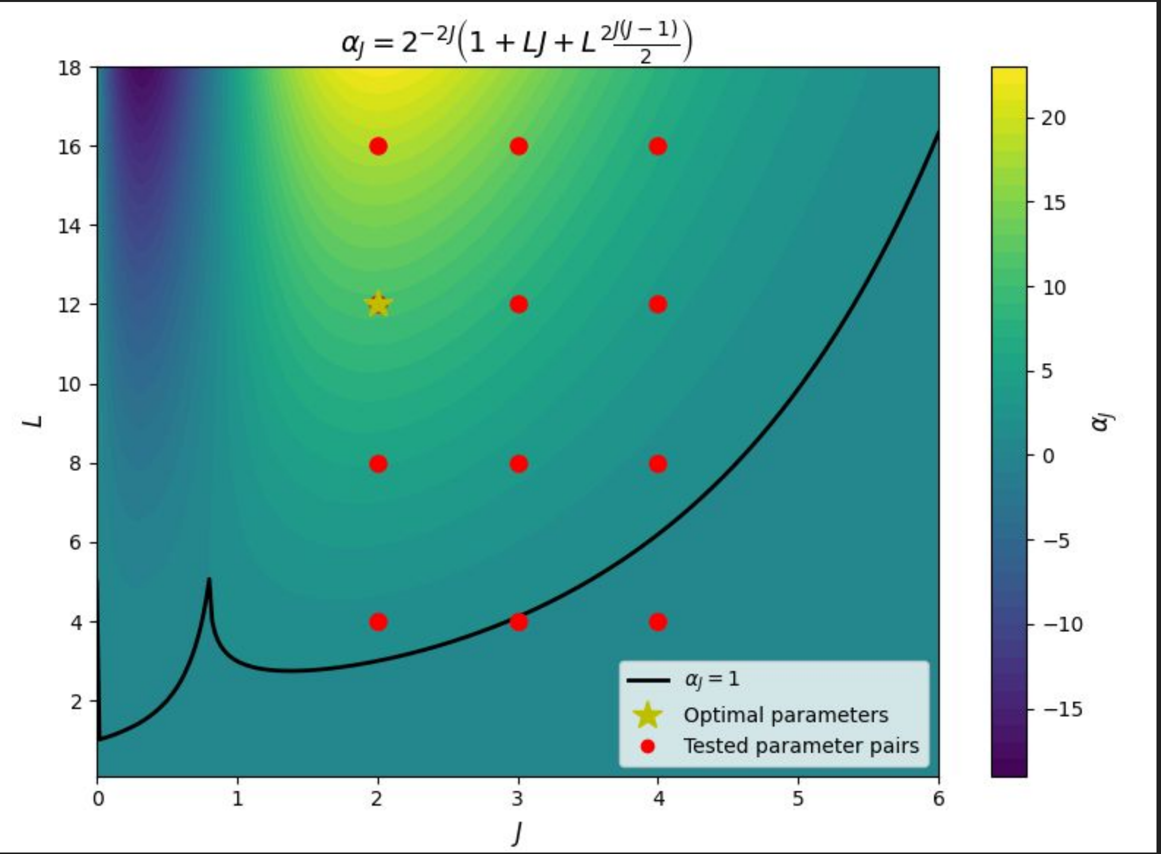


$\star$ = Convolution

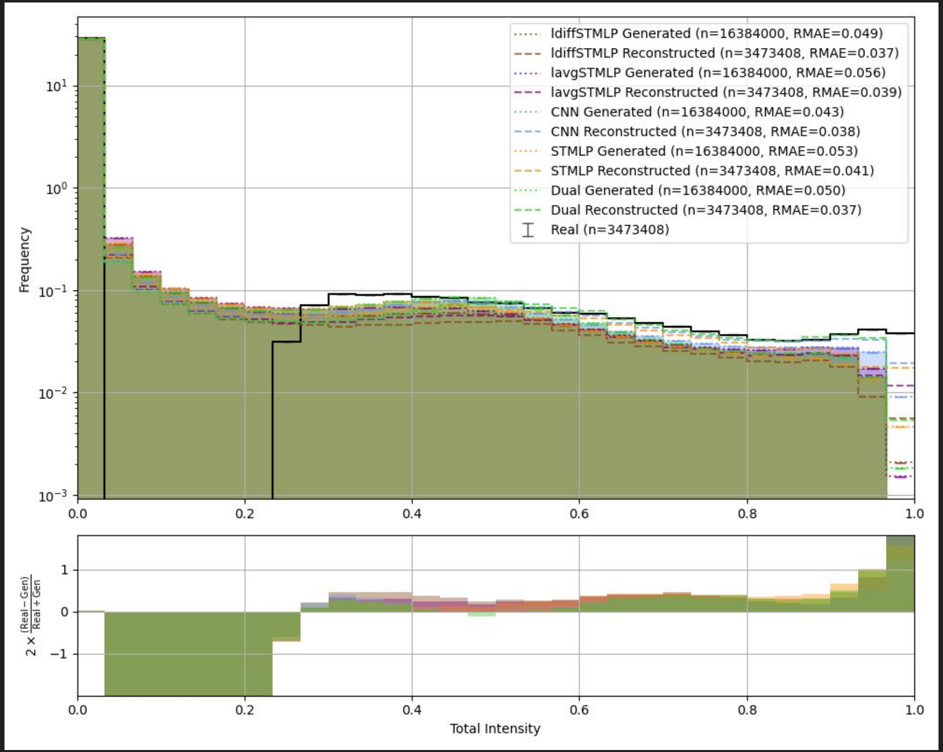
ϕ_J = Low-pass filter

$\psi_{j,l}$ = Wavelet

Performance improves with number of scattering coefficients

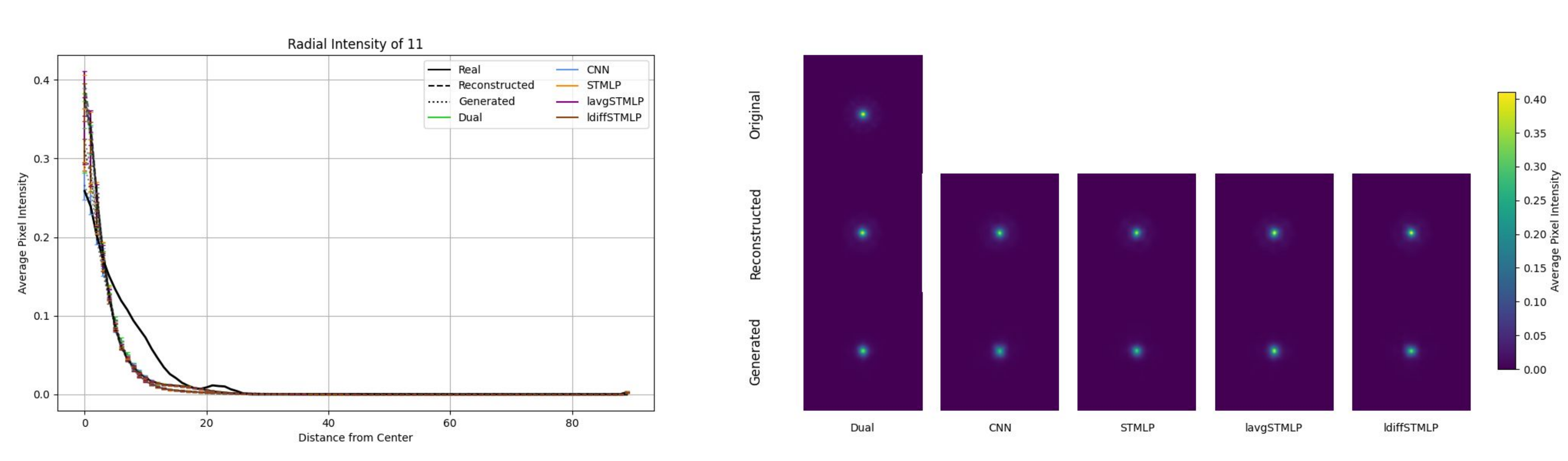


VAEs learns total pixel distribution well for galaxy10

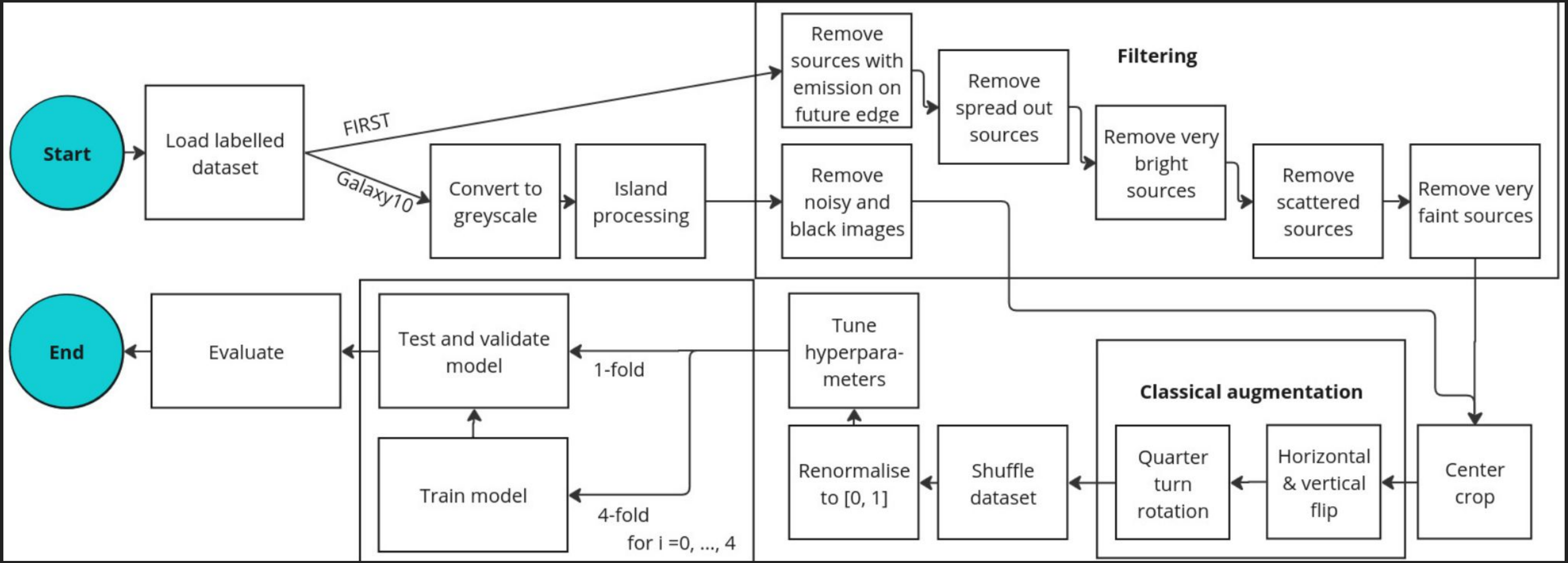


Radial profiles are smoother and steeper than originals

Two plots from
te left one.
Logarithmic
x-axis.



Filtering was an important part in the generative modelling

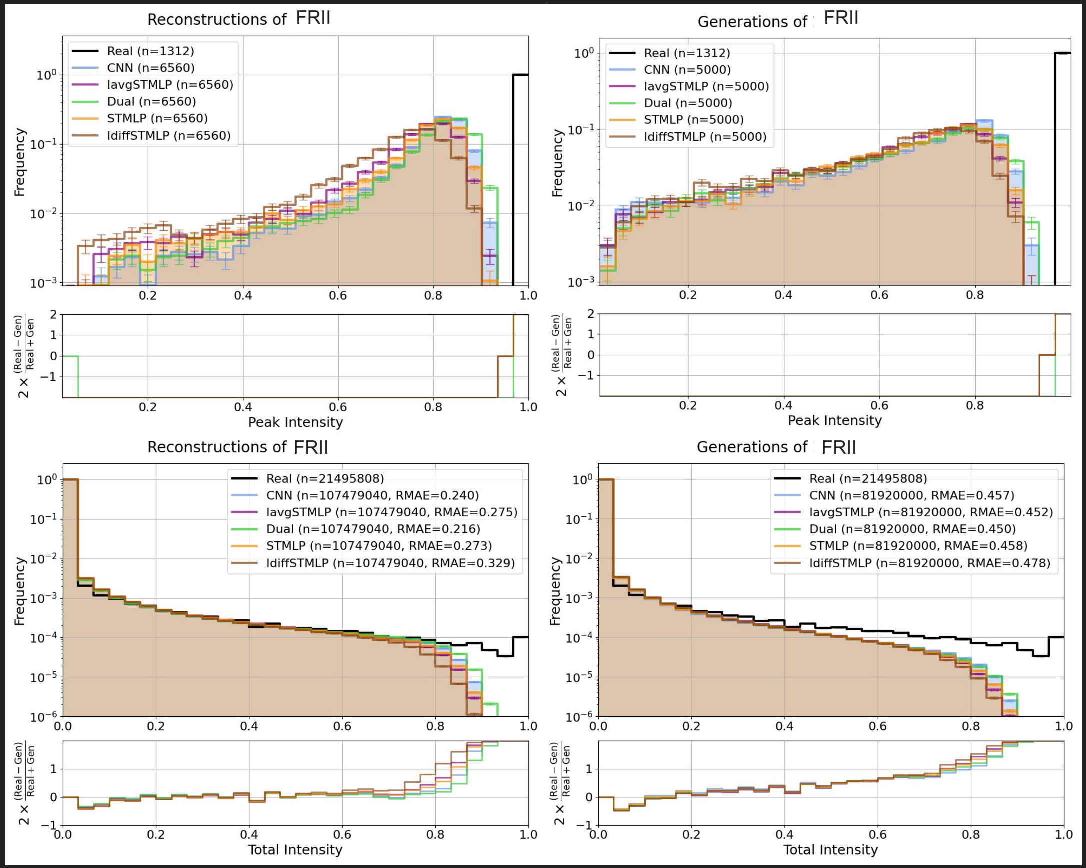


Larger maximum scale J leads to fainter generations

Value of brightest pixel

J=2

Values of all pixels

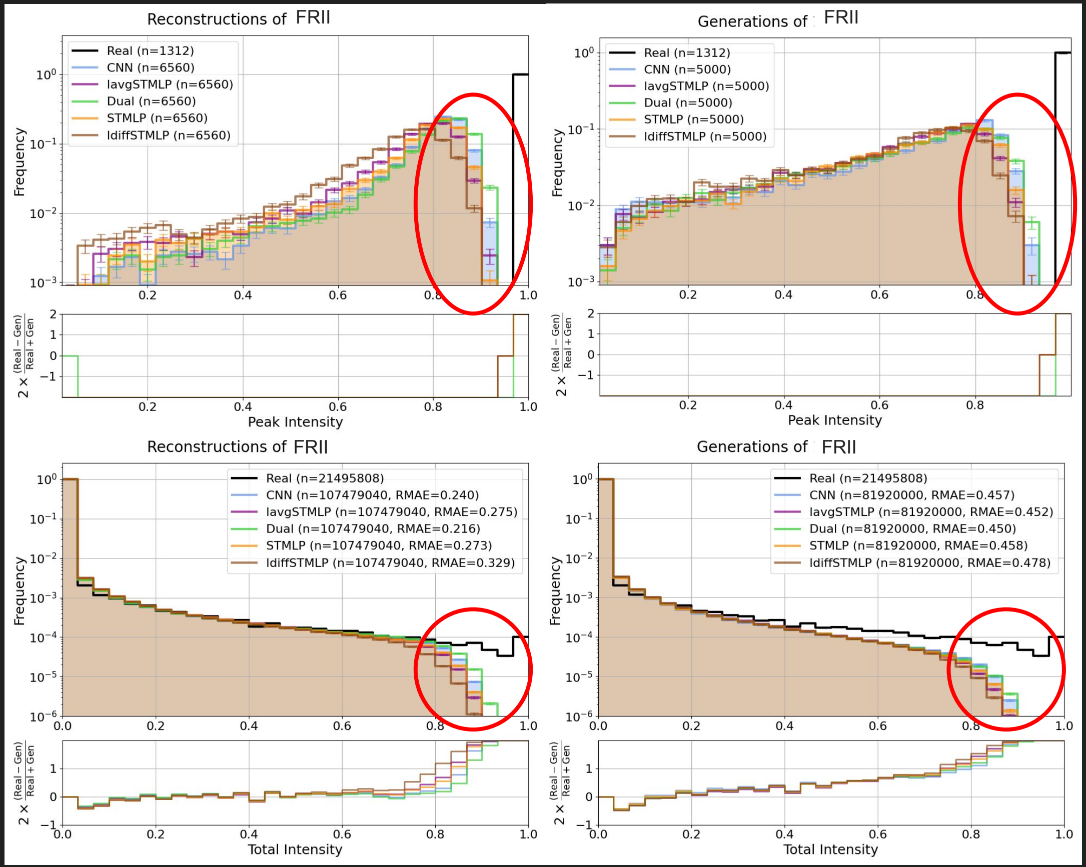


Larger maximum scale J leads to fainter generations

Value of brightest pixel

J=2

Values of all pixels

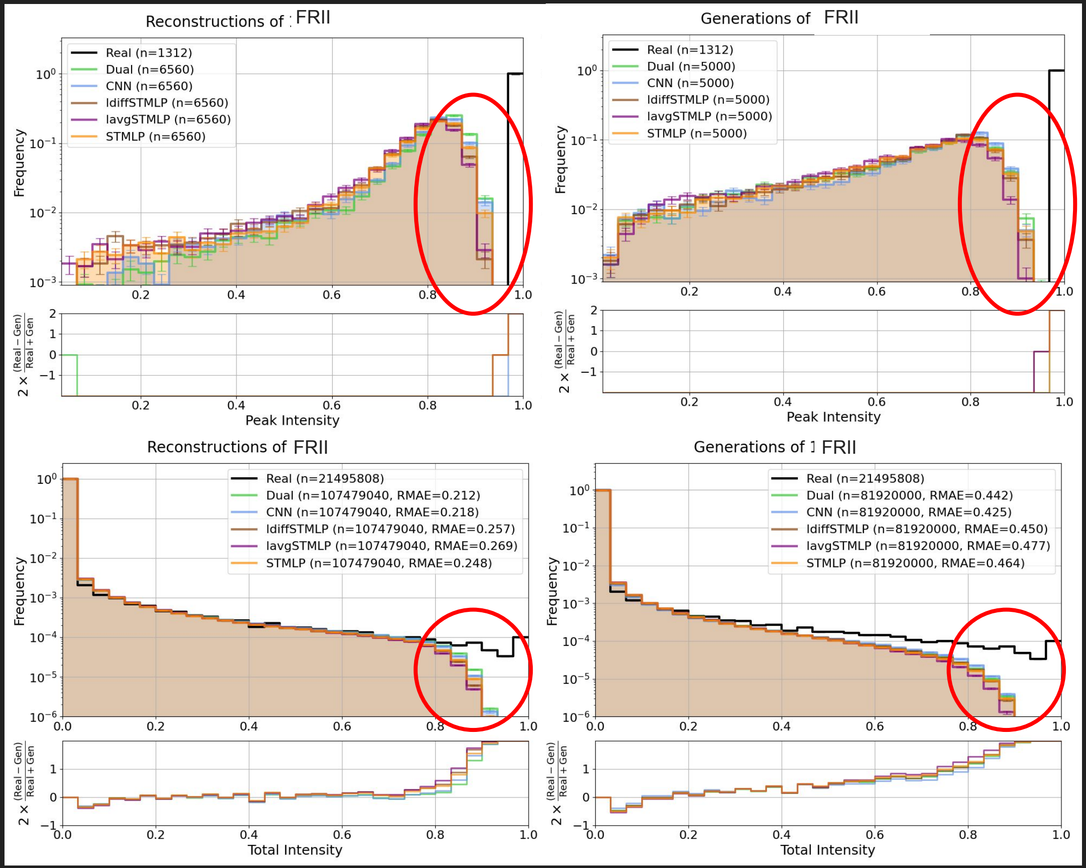


Larger maximum scale J leads to fainter generations

J=3

Value of brightest pixel

Values of all pixels

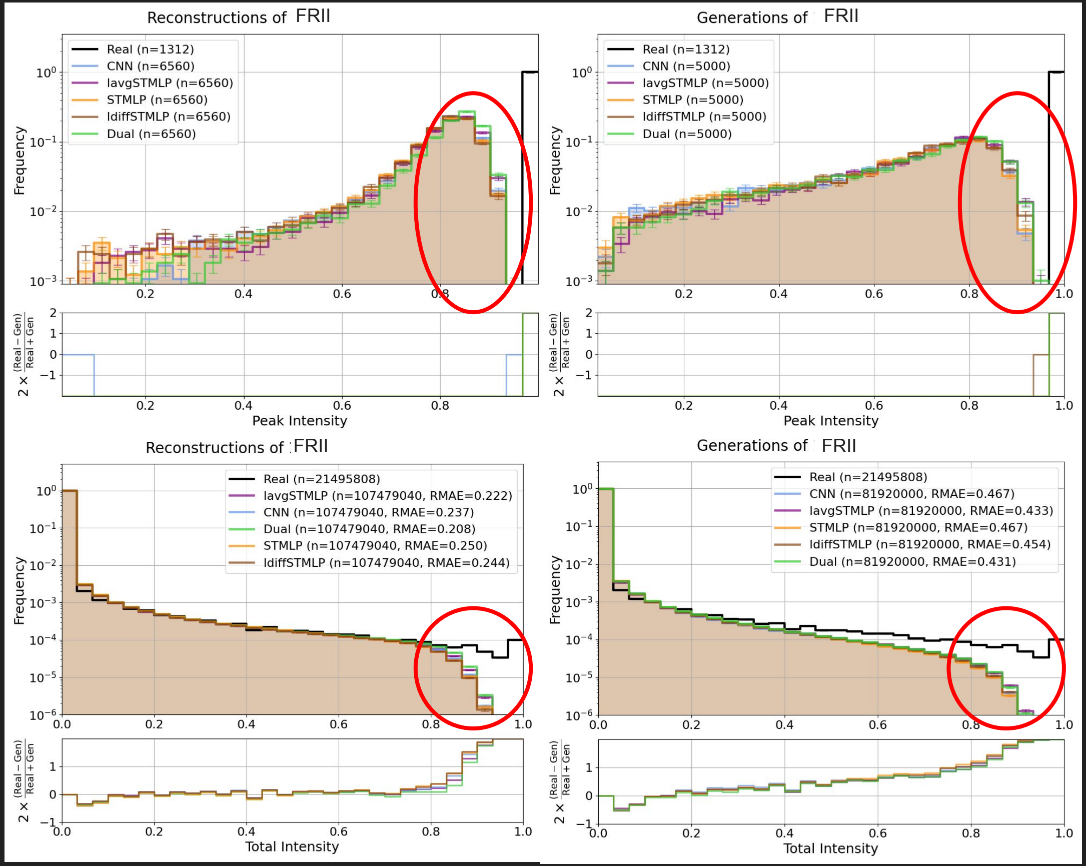


Larger maximum scale J leads to fainter generations

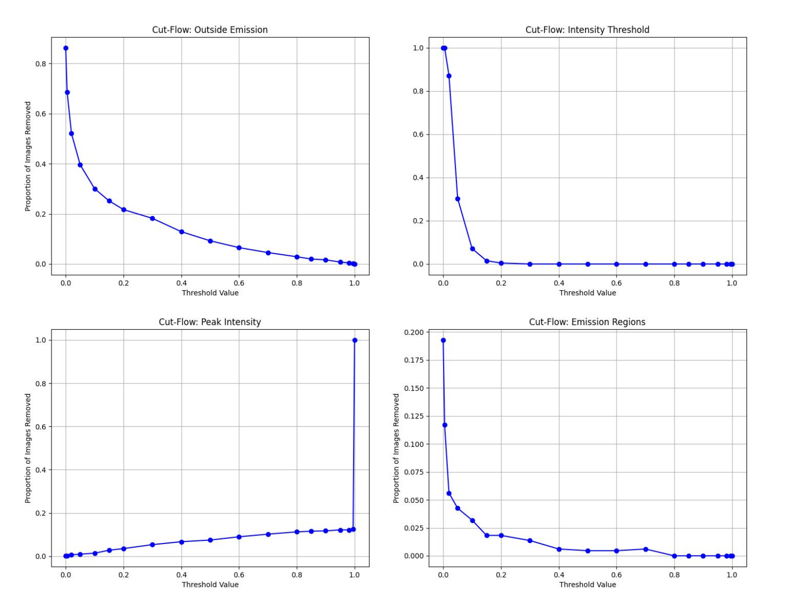
J=4

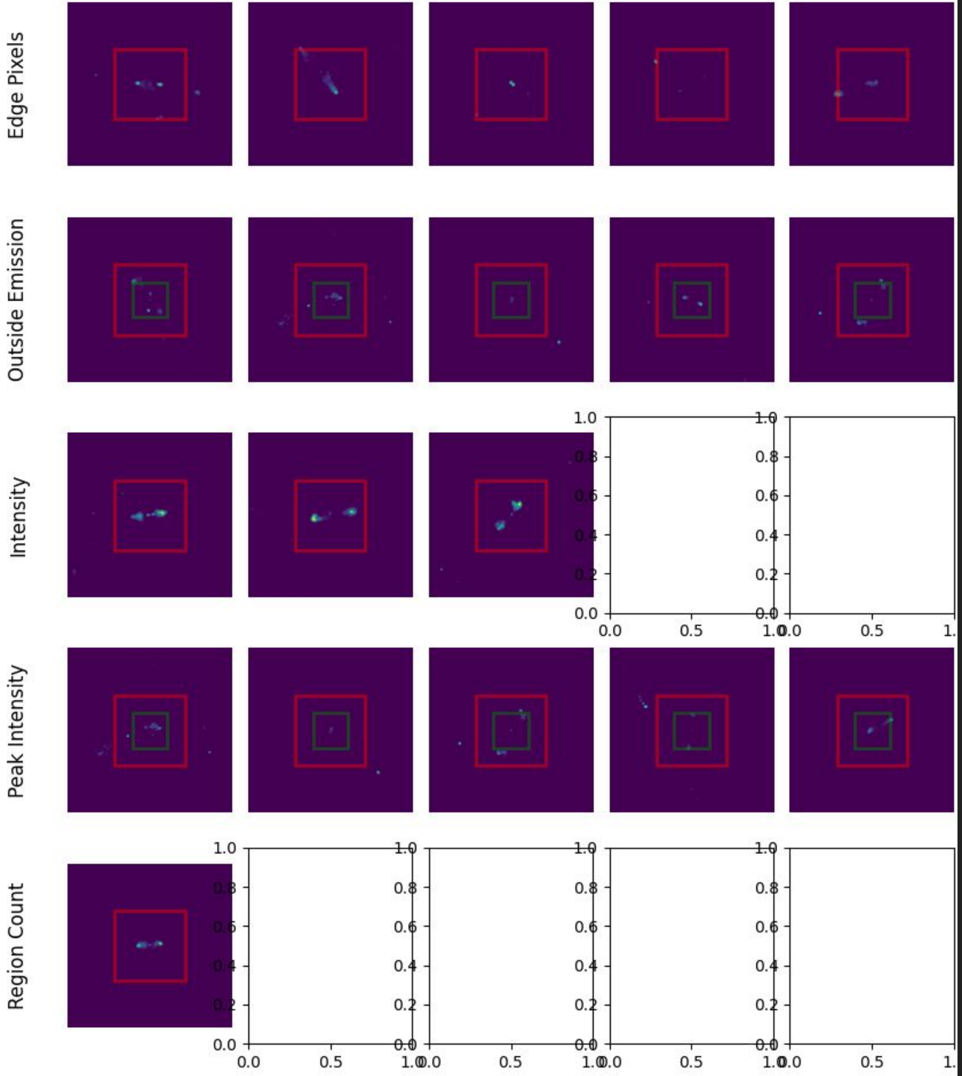
Value of brightest pixel

Values of all pixels



Filtering helps but is difficult

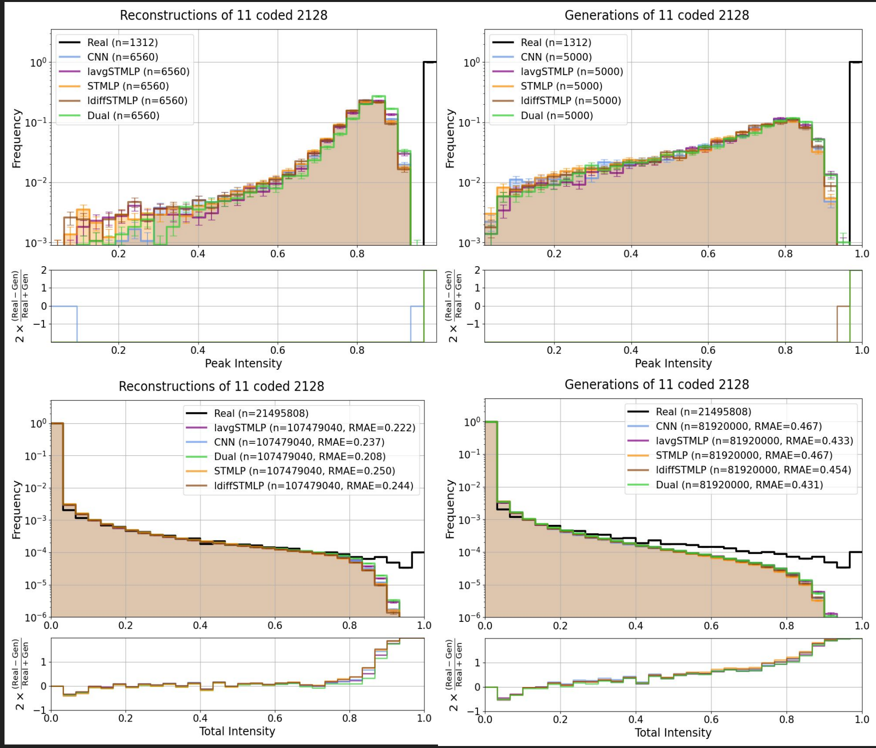




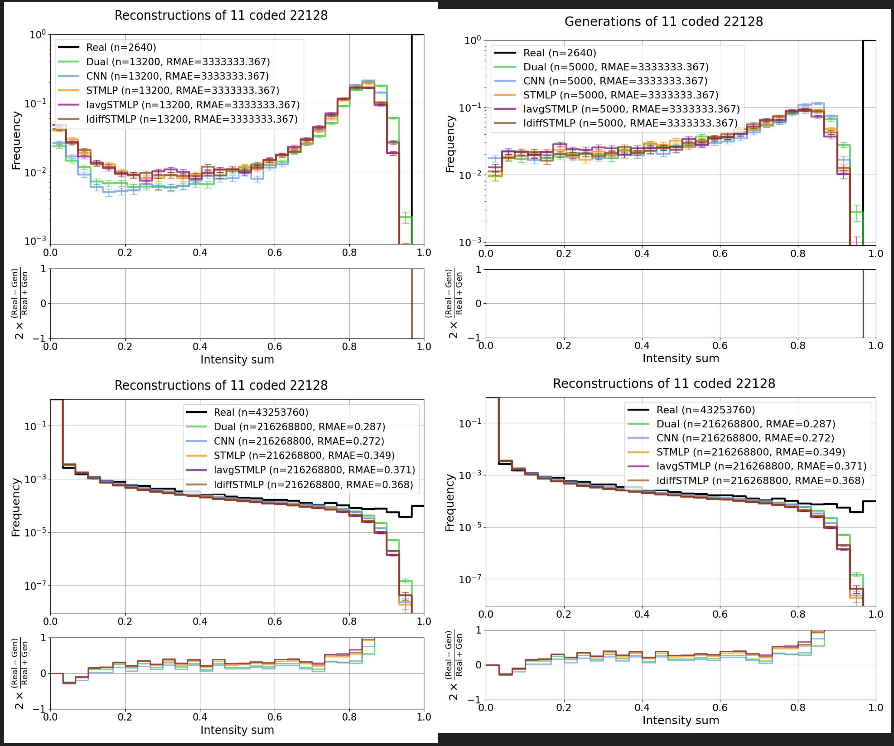
Thresholds and filtering

Non-filtered vs filtered

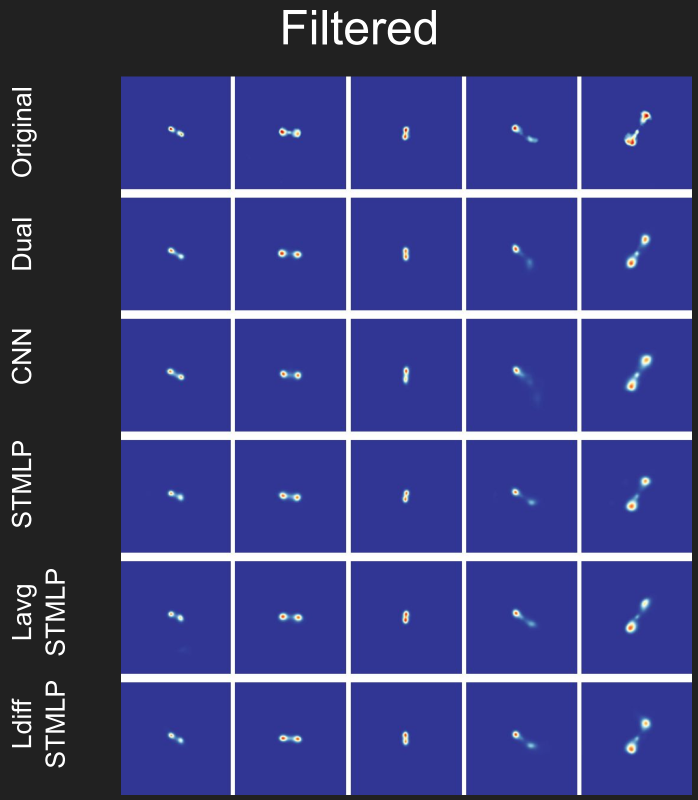
Filtered

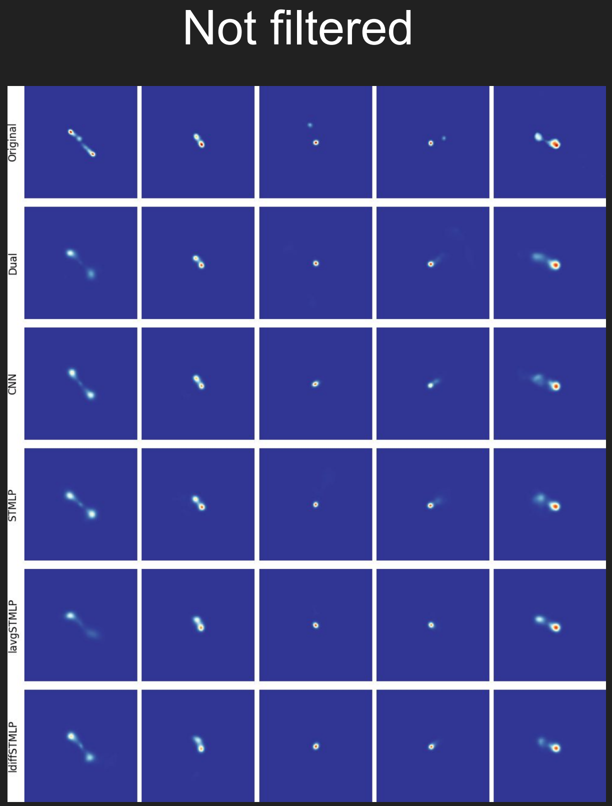


Not filtered

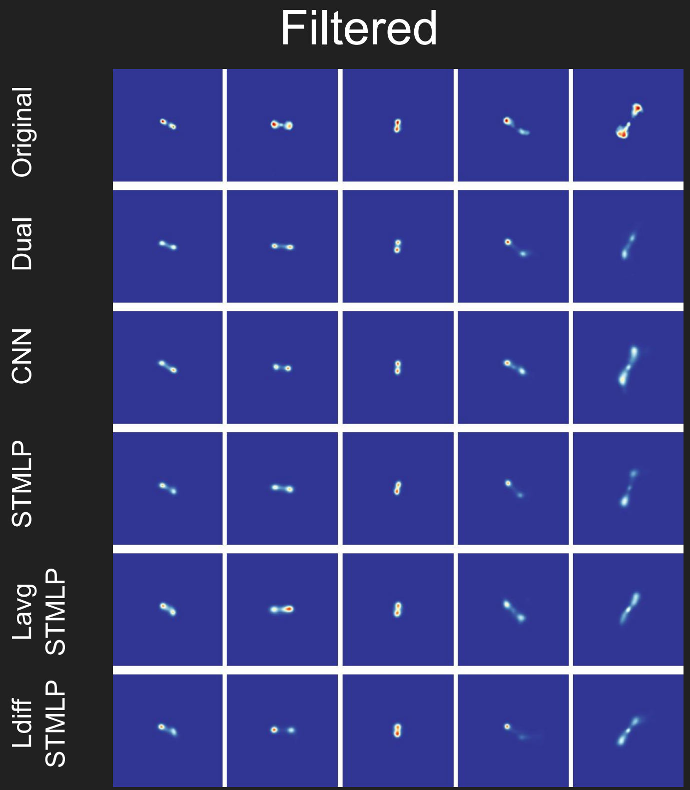


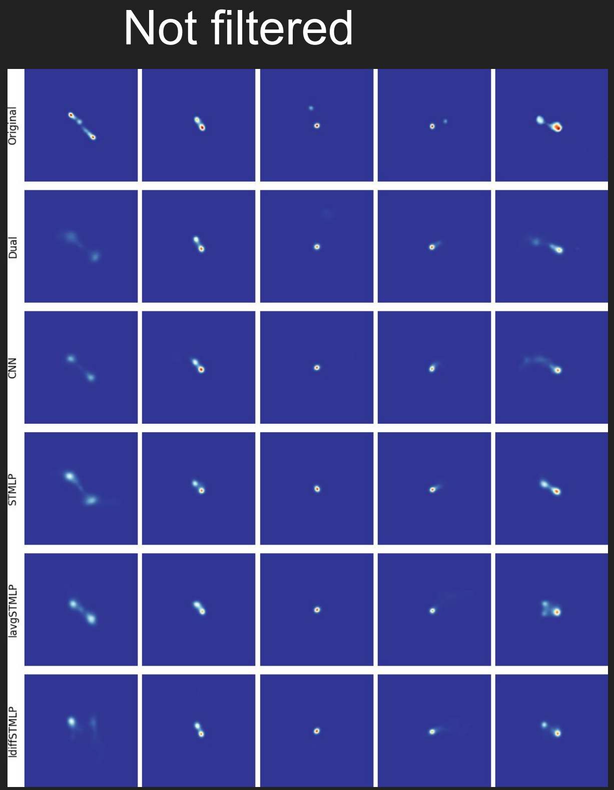
Reconstructed FRII sources





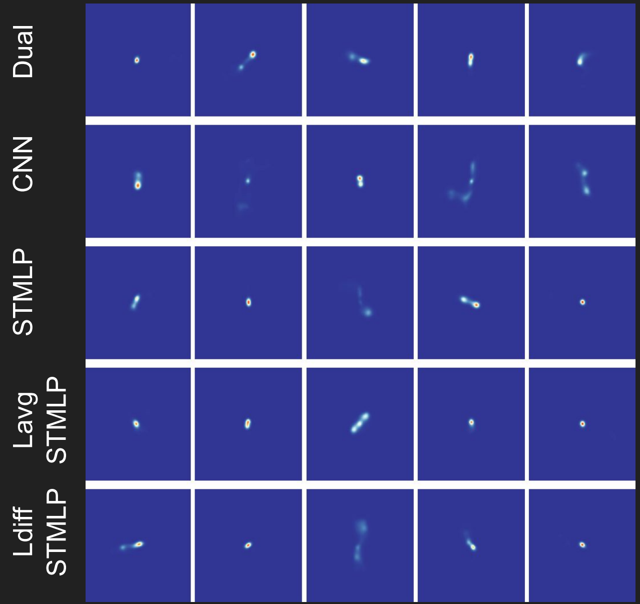
Generated FRII sources





Generated FRII sources

Filtered



Not filtered