TOSCA

Weak lensing for optical and radio surveys

Radio Shape Measurement

Joint ARGOS-TITAN-TOSCA workshop, Crete

Priyamvad Tripathi

6 June, 2024

Observatoire de la Côte d'Azur

Table of contents

Introduction

Overview

- Galaxies sampled as discrete points in the Fourier domain

Overview

- Galaxies sampled as discrete points in the Fourier domain
- $$V(u,v,w) = \int \frac{1}{n} I(l,m) \exp[-2\pi i(ul + vm + w(n-1))] \, \mathrm{d}l\mathrm{d}m \tag{1}$$

  u,v are Fourier domain coordinates, l,m are image domain coordinates, w corresponds to the non-coplanarity of antenna pairs, and $n = \sqrt{1 - l^2 - m^2}$

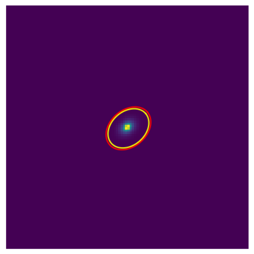
Overview

- Galaxies sampled as discrete points in the Fourier domain
- $$V(u,v,w) = \int \frac{1}{n} I(l,m) \exp[-2\pi i(ul + vm + w(n-1))] \, \mathrm{d}l \mathrm{d}m \tag{1}$$
  u,v are Fourier domain coordinates, l,m are image domain coordinates, w corresponds to the non-coplanarity of antenna pairs, and $n = \sqrt{1 - l^2 - m^2}$
- General radio-interferometric approach for image reconstruction is an iterative process: Minor and Major cycles

Overview

- Galaxies sampled as discrete points in the Fourier domain
-
$$V(u, v, w) = \int \frac{1}{n} I(l, m) \exp[-2\pi i(ul + vm + w(n-1))] \, dl dm \tag{1}$$
u,v are Fourier domain coordinates, l,m are image domain coordinates, w corresponds to the non-coplanarity of antenna pairs, and $n = \sqrt{1 - l^2 - m^2}$
- General radio-interferometric approach for image reconstruction is an iterative process: Minor and Major cycles
- Such a process is computationally expensive and can introduce non-linear biases [Patel et al., 2014]

Example

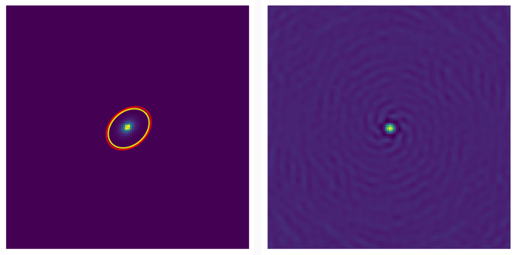


(a) x

$\epsilon = [0.022, -0.154]$

Figure 1: Example of a simulated galaxy with ellipticity $\epsilon = [0.024, -0.152]$.

Example

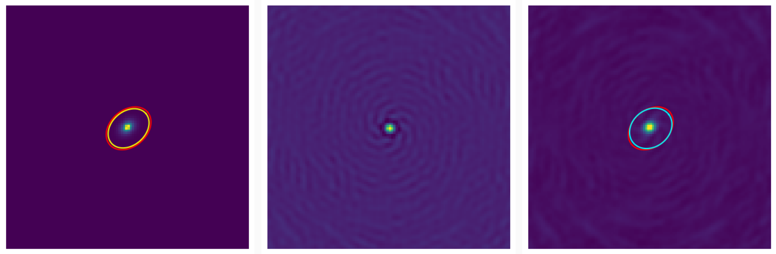


(a) x

$\epsilon = [0.022, -0.154]$

(b) h

Figure 1: Example of a simulated galaxy with ellipticity $\epsilon = [0.024, -0.152]$.

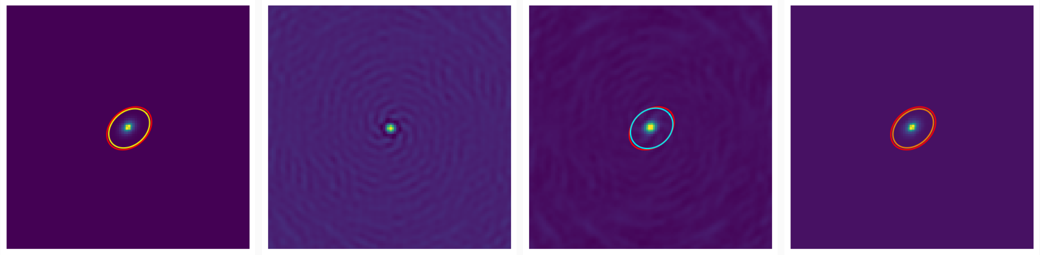


(a) x
$\epsilon = [0.022, -0.154]$

(b) h

(c) x^D
$\epsilon = [0.026, -0.099]$

Figure 1: Example of a simulated galaxy with ellipticity $\epsilon = [0.024, -0.152]$.



(a) x $\epsilon = [0.022, -0.154]$

(b) h

(c) x^D $\epsilon = [0.026, -0.099]$

(d) $\hat{x}$ $\epsilon = [0.021, -0.144]$

Figure 1: Example of a simulated galaxy with ellipticity $\epsilon = [0.024, -0.152]$. The images plotted from left to right are the true image x, PSF h, dirty image x^D, and reconstructed image $\hat{x}$.

Methodology

Image feature extraction

- We used an E(2) equivariant convolutional neural network (CNN) [Weiler and Cesa, 2019] for our feature extraction layer

- We used an E(2) equivariant convolutional neural network (CNN) [Weiler and Cesa, 2019] for our feature extraction layer
- Equivariance is enforced in the structure by using convolution kernels expressed in a steerable basis of the $E(2)$ group:

$$k(\mathbf{x}|\mathbf{w}) = \sum_{\ell=1}^{8} w_\ell(r) Y_\ell(\alpha) \tag{2}$$

where $\mathbf{x} = (r, \alpha)$, $Y_\ell(\alpha) = e^{i\ell\alpha}$ are the basis vectors and the kernel weights $w_\ell(r)$ have a radial symmetry.

- We used an E(2) equivariant convolutional neural network (CNN) [Weiler and Cesa, 2019] for our feature extraction layer
- Equivariance is enforced in the structure by using convolution kernels expressed in a steerable basis of the $E(2)$ group:

$$k(\boldsymbol{x}|\boldsymbol{w}) = \sum_{\ell=1}^{8} w_\ell(r) Y_\ell(\alpha) \tag{2}$$

where $\boldsymbol{x} = (r, \alpha)$, $Y_\ell(\alpha) = e^{i\ell\alpha}$ are the basis vectors and the kernel weights $w_\ell(r)$ have a radial symmetry.

- Produces a vector feature map that is equivariant to the actions of the $E(2)$ group:

$$\mathrm{C}_{E(2)}[G(\hat{x})] = G[\mathrm{C}_{E(2)}(\hat{x})] \tag{3}$$

- **Population Model:** Star-Forming Galaxies (SFGs) catalogue from the Tiered-Radio Extragalactic Continuum Simulation (T- RECS) [Bonaldi et al., 2018]
- Isolated galaxies with galaxy center at antenna pointing: 128×128 pixels

Simulating galaxy images

- **Population Model:** Star-Forming Galaxies (SFGs) catalogue from the Tiered-Radio Extragalactic Continuum Simulation (T- RECS) [Bonaldi et al., 2018]
- Isolated galaxies with galaxy center at antenna pointing: 128×128 pixels
- Visibilities based on SKA-MID configuration: 197 antennas, 936144 unique visibility positions
- Perturb the visibilities with noise sampled from $\mathcal{N}(0, (\sigma/20)^2)$

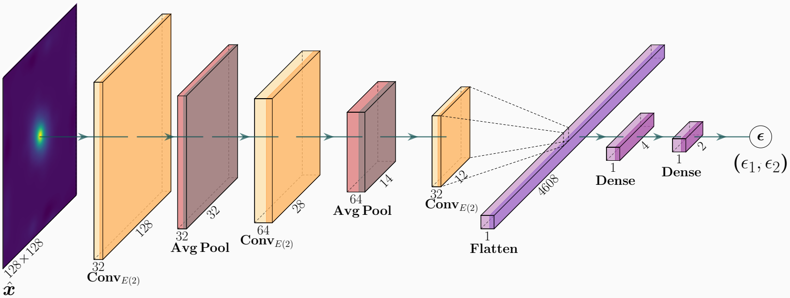
Simulating galaxy images

- **Population Model:** Star-Forming Galaxies (SFGs) catalogue from the Tiered-Radio Extragalactic Continuum Simulation (T- RECS) [Bonaldi et al., 2018]
- Isolated galaxies with galaxy center at antenna pointing: 128×128 pixels
- Visibilities based on SKA-MID configuration: 197 antennas, 936144 unique visibility positions
- Perturb the visibilities with noise sampled from $\mathcal{N}(0, (\sigma/20)^2)$
- Visibilities are gridded followed by Inverse FFT to get dirty image x^D

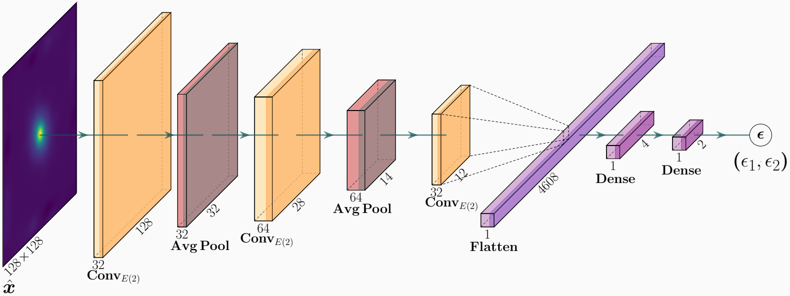
Simulating galaxy images

- **Population Model:** Star-Forming Galaxies (SFGs) catalogue from the Tiered-Radio Extragalactic Continuum Simulation (T- RECS) [Bonaldi et al., 2018]
- Isolated galaxies with galaxy center at antenna pointing: 128×128 pixels
- Visibilities based on SKA-MID configuration: 197 antennas, 936144 unique visibility positions
- Perturb the visibilities with noise sampled from $\mathcal{N}(0, (\sigma/20)^2)$
- Visibilities are gridded followed by Inverse FFT to get dirty image x^D
- PSF h is then deconvolved from the dirty image using MS-Clean to get reconstructed image $\hat{x}$

Fixed PSF

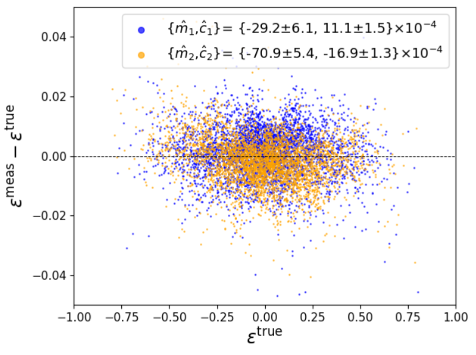


128×128
$\hat{x}$
32
128
$\mathbf{Conv}_{E(2)}$
32
32
Avg Pool
64
28
$\mathbf{Conv}_{E(2)}$
64
14
Avg Pool
32
12
$\mathbf{Conv}_{E(2)}$
1
4608
Flatten
1
4
Dense
1
2
Dense
ϵ
(ϵ_1, ϵ_2)



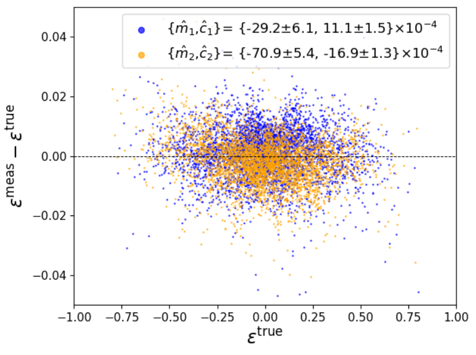
- 20,000 galaxies with varying size and intrinsic ellipticity

$$\hat{\boldsymbol{\theta}}_{\text{eq}} = \underset{\boldsymbol{\theta}_{\text{eq}}}{\operatorname{argmin}} \, \mathbb{E}_{(\hat{x}, \boldsymbol{\epsilon}^{\text{true}})}[\|N_{\boldsymbol{\theta}_{\text{eq}}}(\hat{x}) - \boldsymbol{\epsilon}^{\text{true}}\|^2] \tag{4}$$

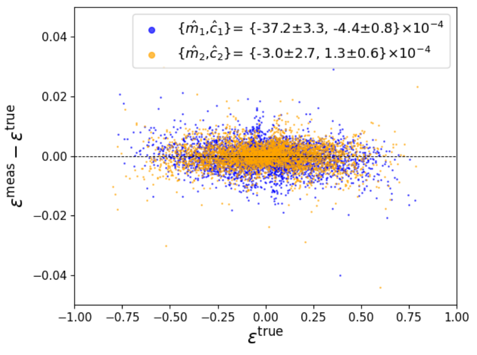


(a) Traditional Network

Figure 2: Comparison of $\Delta\boldsymbol{\epsilon}$ vs $\boldsymbol{\epsilon}^{\text{true}}$. The two components $\Delta\epsilon_1$ and $\Delta\epsilon_2$ are plotted in blue and orange respectively. The legend indicates the linear bias in ellipticity measurement: $\Delta\epsilon_i = m_i\epsilon_i^{\text{true}} + c_i$



(a) Traditional Network



(b) Equivariant Network

Figure 2: Comparison of $\Delta\boldsymbol{\epsilon}$ vs $\boldsymbol{\epsilon}^{\text{true}}$. The two components $\Delta\epsilon_1$ and $\Delta\epsilon_2$ are plotted in blue and orange respectively. The legend indicates the linear bias in ellipticity measurement: $\Delta\epsilon_i = m_i\epsilon_i^{\text{true}} + c_i$

Variable PSF

PSF Encoding

- PSF can be computed using the uv sampling pattern and galaxy position

PSF Encoding

- PSF can be computed using the uv sampling pattern and galaxy position
- The reconstructed images contain artifacts from the corresponding PSFs

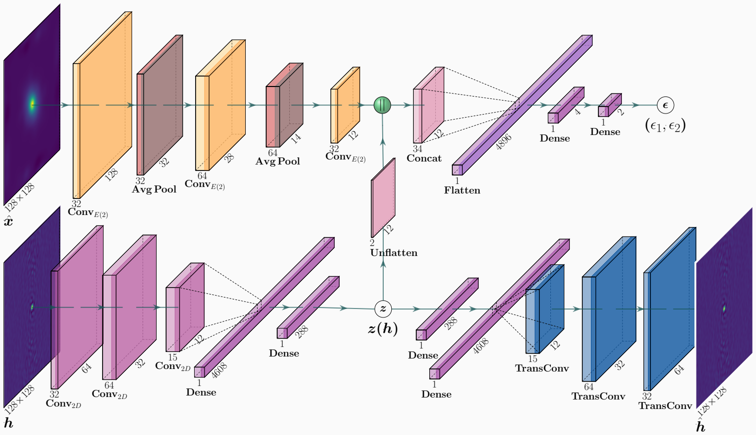
PSF Encoding

- PSF can be computed using the uv sampling pattern and galaxy position
- The reconstructed images contain artifacts from the corresponding PSFs
- We first train an autoencoder to encode the PSFs to a latent space

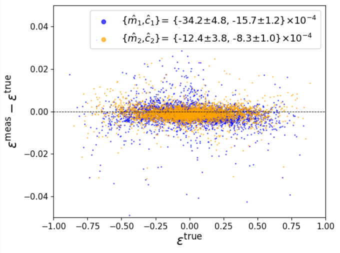
- PSF can be computed using the uv sampling pattern and galaxy position
- The reconstructed images contain artifacts from the corresponding PSFs
- We first train an autoencoder to encode the PSFs to a latent space
-

$$\{\hat{\boldsymbol{\theta}}_{\mathrm{E}}, \hat{\boldsymbol{\theta}}_{\mathrm{D}}\} = \underset{\{\boldsymbol{\theta}_{\mathrm{E}},\boldsymbol{\theta}_{\mathrm{D}}\}}{\operatorname{argmin}} \mathbb{E}_h[\|h - \hat{h}_{\boldsymbol{\theta}_{\mathrm{D}}}[\boldsymbol{z}_{\boldsymbol{\theta}_{\mathrm{E}}}(h)]\|^2] \tag{5}$$

where $\hat{h}_{\boldsymbol{\theta}_{\mathrm{D}}}$ is the output from the decoder, $\boldsymbol{z}_{\boldsymbol{\theta}_{\mathrm{E}}}$ is the output from the encoder and $\{\boldsymbol{\theta}_{\mathrm{E}},\boldsymbol{\theta}_{\mathrm{D}}\}$ the encoder-decoder architecture parameters.

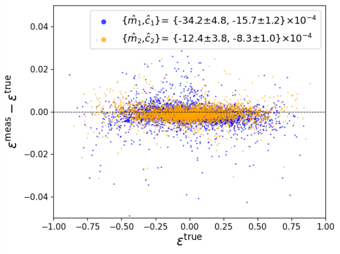


20,000 galaxies/PSF pairs with varying size and intrinsic ellipticity

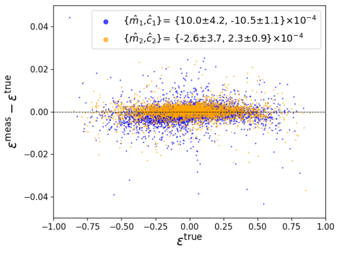


(a) Reconstructed Images: 1000 MS-Clean cycles

Figure 3

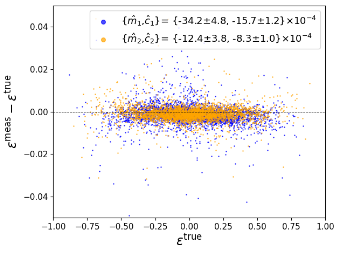


(a) Reconstructed Images: 1000 MS-Clean cycles

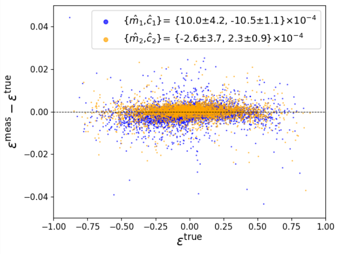


(b) Reconstructed Images: 500 MS-Clean cycles

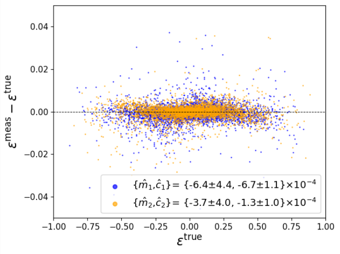
Figure 3



(a) Reconstructed Images: 1000 MS-Clean cycles



(b) Reconstructed Images: 500 MS-Clean cycles

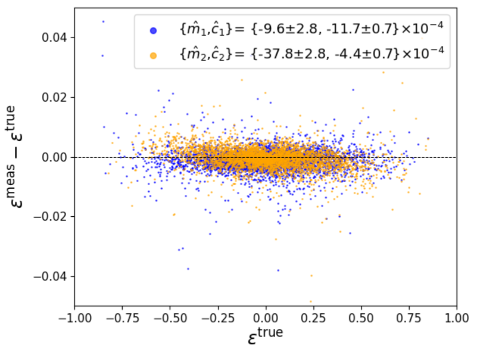


(c) Dirty Images

Figure 3

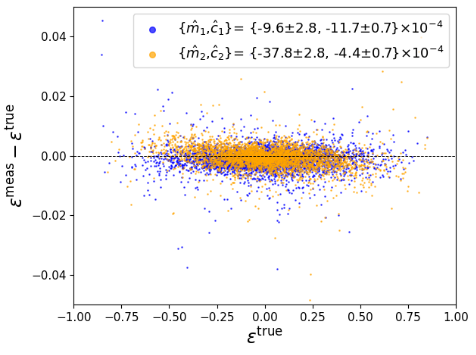
Figure 4: Galaxies following a Sérsic brightness profile: $I(r) = I_0 \exp[-(\frac{r}{r_0})^{\frac{1}{n}}]$ with index n drawn from $\mathcal{U}(1, 4)$

Model Bias

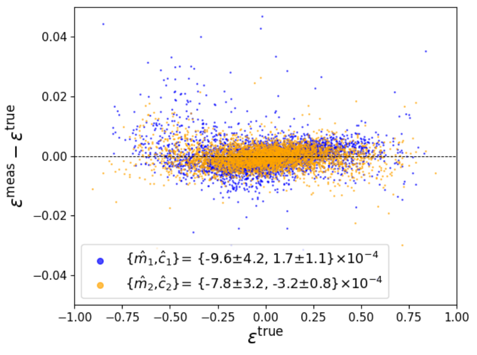


(a) Reconstructed Images

Figure 4: Galaxies following a Sérsic brightness profile: $I(r) = I_0 \exp[-(\frac{r}{r_0})^{\frac{1}{n}}]$ with index n drawn from $\mathcal{U}(1, 4)$



(a) Reconstructed Images



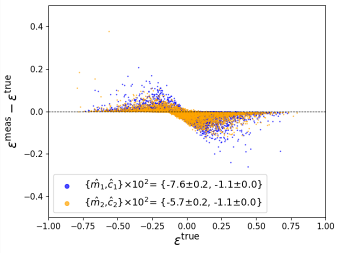
(b) Dirty Images

Figure 4: Galaxies following a Sérsic brightness profile: $I(r) = I_0 \exp[-(\frac{r}{r_0})^{\frac{1}{n}}]$ with index n drawn from $\mathcal{U}(1, 4)$

Comparison with Other Works

ShapeNet Deconvolution [Nammour et al., 2022]

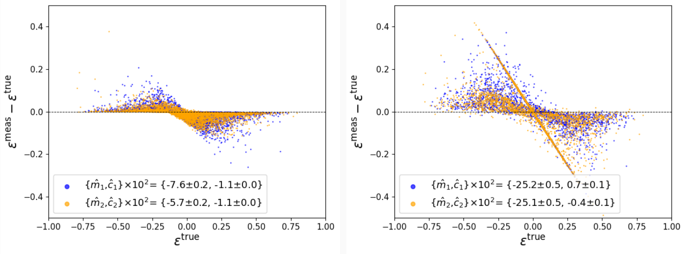
- Tikhonov solution: $\tilde{x} = (H^T H + \lambda \Gamma^T \Gamma)^{-1} H^T x^D$ where H corresponds to the PSF operator, Γ corresponds to Tikhonov linear filter and λ is the regularisation weight.
- A UNET architecture is then trained to learn the mapping b/w the Tikhonov output and the true image.
- The network is trained to minimize the following loss function: $l(\hat{x}) = ||\hat{x} - x||^2 + \gamma M(\hat{x})$
- $M(\hat{x}) = \sum_{i=1}^{6} \omega_i \langle \hat{x} - x, u_i \rangle$ is a shape constraint with $\{\omega_i\}$ and $\{u_i\}$ are constant scalar weights and images respectively



(a) $x^D = x * h + n$

Figure 5

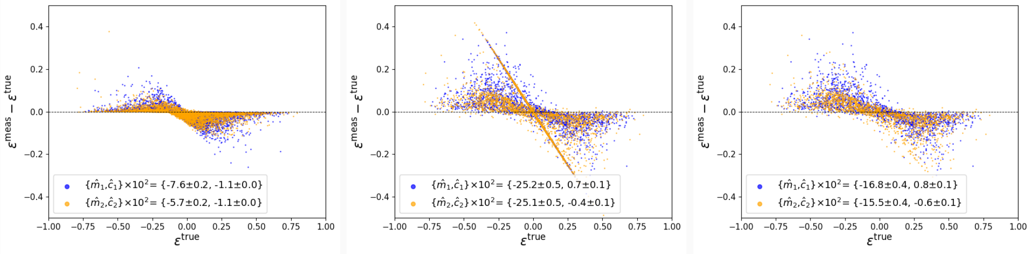
Comparison with ShapeNet



(a) $x^D = x * h + n$

(b) $X^D = f(x, h, w, D/G, \text{vis})$

Figure 5



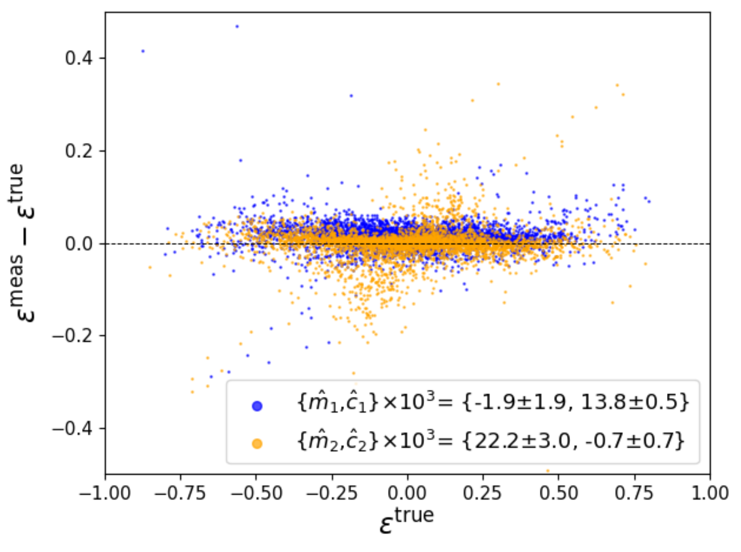
(a) $x^D = x * h + n$

(b) $X^D = f(x, h, w, D/G, \mathrm{vis})$

(c) Same as b with $\sim 25\%$ objects removed

Figure 5

- Reconstruct image by deconvolving the PSF from the dirty image and estimate ellipticity $\boldsymbol{\epsilon}_k^{\text{calc}}$
- In the residual image, inject model sources with the same size and flux properties, but known ellipticity $\epsilon_i^{\text{inp}} = \{0, \ \pm 0.2375, \ \pm 0.475, \ \pm 0.7125, \ \pm 0.95\}$
- Simulate visibilities $\Rightarrow$ Dirty Image $\Rightarrow$ Reconstructed Image $\Rightarrow$ Measure ellipticity $\boldsymbol{\epsilon}^{\text{obs}}$
- Fit second order 2D polynomial $b_k(\epsilon_1^{\text{inp}}, \epsilon_2^{\text{inp}}) = \epsilon_1^{\text{obs}} - \epsilon_1^{\text{inp}}$
- Calibrate observed ellipticities using $\epsilon_{1,k}^{\text{SC}} = \epsilon_{1,k}^{\text{calc}} - b_k(\epsilon_{1,k}^{\text{calc}}, \epsilon_{2,k}^{\text{calc}})$
- Repeat for ϵ_2


0.4
0.2
0.0
−0.2
−0.4
$\epsilon^{\mathrm{meas}} - \epsilon^{\mathrm{true}}$
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
ϵ^{true}
$\{\hat{m}_1, \hat{c}_1\} \times 10^3 = \{-1.9 \pm 1.9, 13.8 \pm 0.5\}$
$\{\hat{m}_2, \hat{c}_2\} \times 10^3 = \{22.2 \pm 3.0, -0.7 \pm 0.7\}$

Summary

Table 1: Linear Bias in Ellipticity estimates (at the order of 10^{-3})

		m_1	c_1	m_2	c_2
Fixed PSF	Trad Net	$\mathbf{2.9 \pm 0.6}$	1.1 ± 0.2	7.1 ± 0.5	-1.7 ± 0.1
	Eq Net	3.7 ± 0.3	$\mathbf{-0.4 \pm 0.1}$	$\mathbf{-0.3 \pm 0.2}$	$\mathbf{0.1 \pm 0.1}$
Variable PSF	Recon 500	1.0 ± 0.4	-1.1 ± 0.1	$\mathbf{-0.3 \pm 0.4}$	0.2 ± 0.1
	Recon 1000	-3.4 ± 0.5	-1.6 ± 0.1	-1.2 ± 0.4	-0.8 ± 0.1
	Dirty	$\mathbf{-0.6 \pm 0.4}$	$\mathbf{-0.7 \pm 0.1}$	-0.4 ± 0.4	$\mathbf{-0.1 \pm 0.1}$
	Shapenet Decon	76.1 ± 2.0	-11.3 ± 0.1	57.3 ± 2.1	-11.1 ± 0.0
	SuperCLASS Calib	1.9 ± 1.9	13.8 ± 0.5	22.2 ± 3.0	-0.7 ± 0.7
Sersic Gal	Recon	$\mathbf{1.0 \pm 0.3}$	-1.2 ± 0.1	-3.8 ± 0.3	-0.4 ± 0.1
	Dirty	-1.0 ± 0.4	$\mathbf{0.2 \pm 0.1}$	$\mathbf{-0.8 \pm 0.3}$	$\mathbf{-0.3 \pm 0.1}$

Conclusion

- Equivariant convolutions are better than traditional convolutions for our problem

Conclusion

- Equivariant convolutions are better than traditional convolutions for our problem
- PSF can be encoded into the network

Conclusion

- Equivariant convolutions are better than traditional convolutions for our problem
- PSF can be encoded into the network
- The specifics of the image reconstruction process are not that important and the network can be trained directly with the dirty images

Conclusion

- Equivariant convolutions are better than traditional convolutions for our problem
- PSF can be encoded into the network
- The specifics of the image reconstruction process are not that important and the network can be trained directly with the dirty images
- The network can recover ellipticities with similar/better linear biases as other popular methods

Conclusion

- Equivariant convolutions are better than traditional convolutions for our problem
- PSF can be encoded into the network
- The specifics of the image reconstruction process are not that important and the network can be trained directly with the dirty images
- The network can recover ellipticities with similar/better linear biases as other popular methods
- Can work with galaxies with different intensity profiles

Thank you for your time

Reported Statistics

Table 2: Comparison of MAE: $\frac{1}{N_{\text{obj}}} \sum_{n=1}^{N_{\text{obj}}} |\Delta \epsilon_i^n|$

	MAE ϵ_1 (ϵ_2)	N_{obj}
ShapeNet Paper	$7.34\ (7.70) \times 10^{-2}$	3000
Case I	$2.37\ (2.69) \times 10^{-2}$	3810
Case II	$7.11\ (7.02) \times 10^{-2}$	3247
Fiducial Network (Recon Images)	$3.82\ (2.74) \times 10^{-3}$	3993
Fiducial Network (Dirty Images)	$4.16\ (3.59) \times 10^{-3}$	3993

RadioLensfit [Rivi et al., 2016]

- Works using visibilities
- Galaxy brightness profile: $I(r) = I_0 \exp(-r/\alpha)$,
- Transformation matrix $\mathbf{A}$ with ellipticity parameters $\mathbf{e} = (e_1, e_2)$ such that:

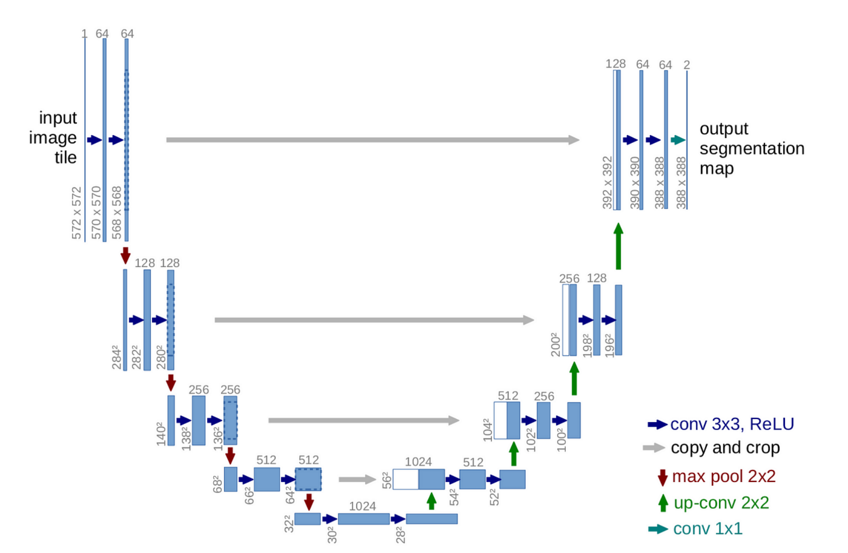
$$\begin{pmatrix} l_r \\ m_r \end{pmatrix} = \mathbf{A}x = \begin{pmatrix} 1-e_1 & -e_2 \\ -e_2 & 1+e_1 \end{pmatrix} \times \begin{pmatrix} l \\ m \end{pmatrix}$$

- Observed visibility due to a galaxy at point $\mathbf{k} = (u, v)$ can be given by:

$$V_s(u, v) = \frac{2\pi\alpha^2 I_0}{|\mathbf{A}|(1 + 4\pi^2\alpha^2|\mathbf{A}^{-\intercal}k|)^{3/2}} \times \exp 2\pi i \mathbf{k}^\intercal \mathbf{x_0} \qquad (6)$$

- Perform a Bayesian marginalization of the likelihood over I_0, α and source centroid position $\mathbf{x_0} = (l_0, m_0) \Rightarrow P(\mathbf{A}|D)$

UNET ARCHITECTURE



Bonaldi, A., Bonato, M., Galluzzi, V., Harrison, I., Massardi, M., Kay, S., De Zotti, G., and Brown, M. L. (2018).
The tiered radio extragalactic continuum simulation (t-recs).
MNRAS, 482(1):2–19.

Harrison, I., Brown, M. L., Tunbridge, B., Thomas, D. B., Hillier, T., Thomson, A. P., Whittaker, L., Abdalla, F. B., Battye, R. A., Bonaldi, A., Camera, S., Casey, C. M., Demetroullas, C., Hales, C. A., Jackson, N. J., Kay, S. T., Manning, S. M., Peters, A., Riseley, C. J., and Watson, R. A. (2020).
Superclass – iii. weak lensing from radio and optical observations in data release 1.
MNRAS, 495(2):1737–1759.

References II

- Nammour, F., Akhaury, U., Girard, J. N., Lanusse, F., Sureau, F., Ben Ali, C., and Starck, J. L. (2022).
  ShapeNet: Shape constraint for galaxy image deconvolution.
  A&A, 663:A69.
- Patel, P., Abdalla, F. B., Bacon, D. J., Rowe, B., Smirnov, O. M., and Beswick, R. J. (2014).
  Weak lensing measurements in simulations of radio images.
  MNRAS, 444(3):2893–2909.
- Rivi, M., Miller, L., Makhathini, S., and Abdalla, F. B. (2016).
  Radio weak lensing shear measurement in the visibility domain – i. methodology.
  MNRAS, 463(2):1881–1890.
- Weiler, M. and Cesa, G. (2019).
  General e (2)-equivariant steerable cnns.
  Advances in neural information processing systems, 32.