Rencontres du Vietnam 2025

Inference of Weak Lensing Parameters from Blended Galaxies using Generative Neural Networks

Nguyen Thi Yen Binh

Supervisors: **Wassim Kabalan, Cecile Roucelle, Alexandre Boucaud, Eric Aubourg**

August 13, 2025

- Introduction
- Variational Autoencoder
- Pipeline Overview
- Results
- Conclusion

Weak gravitational lensing

1 Introduction

Gravitational Lensing: the gravitational deflection of light from distant galaxies by intervening matter.

- Galaxy images distorted in *shape* and *size*.
- Statistical measurement of *cosmic shear*.
- Map the 3-D *matter distribution*

→ Probe the evolution of dark energy

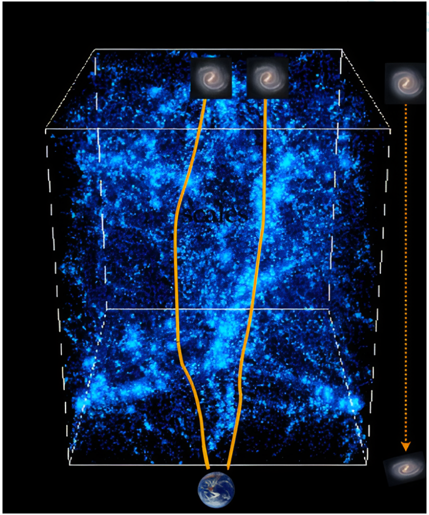


Figure 1: Weak gravitational Lensing


LSST Vera C. Rubin Observatory

Galaxy Blending

1 Introduction

Figure 2: Blended scenes in LSST

- In preparation for LSST Vera C. Rubin data.

Galaxy Blending

Figure 2: Blended scenes in LSST

- In preparation for LSST Vera C. Rubin data.
- About 62% of the detected objects will be predicted as blended

Galaxy Blending

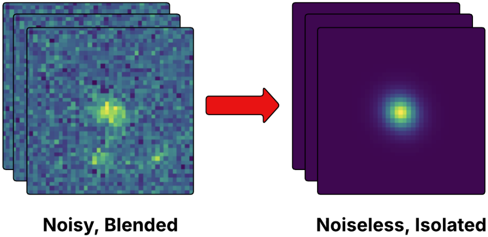
Figure 2: Blended scenes in LSST

- In preparation for LSST Vera C. Rubin data.
- About 62% of the detected objects will be predicted as blended
  → Systematic effects on shear measurement

Galaxy Blending

Figure 2: Blended scenes in LSST

- In preparation for LSST Vera C. Rubin data.
- About 62% of the detected objects will be predicted as blended
  → Systematic effects on shear measurement
  → Data-driven deblending algorithm



Goals

1 Introduction

1. **Denoise + Deblend galaxy images**
   — Redistribute flux in each pixel
   — Combine multi-band information
   — Reconstruct complex morphologies

Goals

1. **Denoise + Deblend galaxy images**
   - Redistribute flux in each pixel
   - Combine multi-band information
   - Reconstruct complex morphologies

→ Variational Autoencoder (Kingma & Welling 2019)

Goals

1. **Denoise + Deblend galaxy images**
   - Redistribute flux in each pixel
   - Combine multi-band information
   - Reconstruct complex morphologies

   → Variational Autoencoder (Kingma & Welling 2019)

2. **Extract weak lensing parameters**
   - Ellipticities (e_1, e_2)
   - Redshifts z

Goals

1. **Denoise + Deblend galaxy images**
   — Redistribute flux in each pixel
   — Combine multi-band information
   — Reconstruct complex morphologies

   → Variational Autoencoder (Kingma & Welling 2019)
2. **Extract weak lensing parameters**
   — Ellipticities (e_1, e_2)
   — Redshifts z

   → Regressor

Goals

1. **Denoise + Deblend galaxy images**
   - Redistribute flux in each pixel
   - Combine multi-band information
   - Reconstruct complex morphologies

   → Variational Autoencoder (Kingma & Welling 2019)

2. **Extract weak lensing parameters**
   - Ellipticities (e_1, e_2)
   - Redshifts z

   → Regressor

3. **Test on real data**

Goals

1. **Denoise + Deblend galaxy images**
   — Redistribute flux in each pixel
   — Combine multi-band information
   — Reconstruct complex morphologies

   → Variational Autoencoder (Kingma & Welling 2019)

2. **Extract weak lensing parameters**
   — Ellipticities (e_1, e_2)
   — Redshifts z

   → Regressor

3. **Test on real data**
   → Dark Energy Survey (DES)

Table of Contents

- Introduction
- Variational Autoencoder
- Pipeline Overview
- Results
- Conclusion

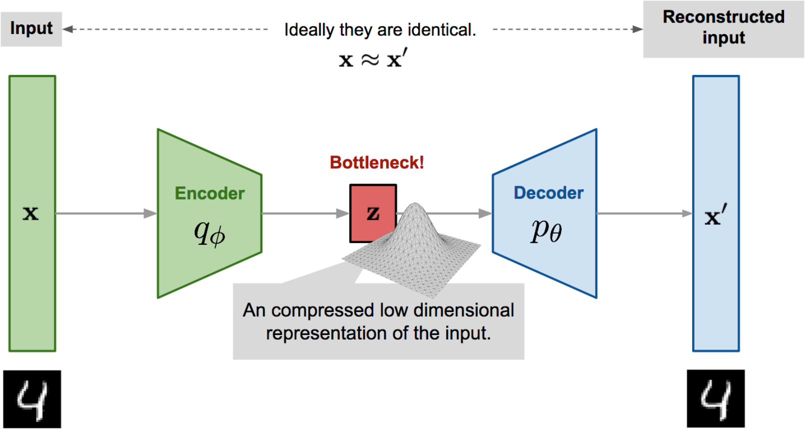
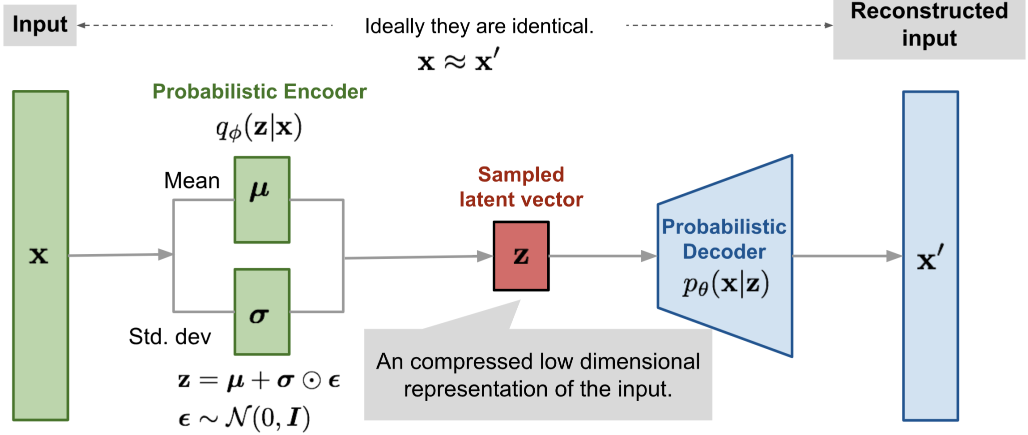


Figure 3: VAE (source: Lilian Weng)

Loss function:

$$\mathcal{L}_{\mathrm{VAE}}(\mathbf{x}) = -\underbrace{\mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}\left[\log p_\theta(\mathbf{x} \mid \mathbf{z})\right]}_{\text{reconstruction}} + \underbrace{D_{\mathrm{KL}}\left(q_\phi(\mathbf{z} \mid \mathbf{x}) \| p(\mathbf{z})\right)}_{\text{regularization}} \tag{1}$$

Reparameterization trick



Loss function:

$$\mathcal{L}_{\text{VAE}} = \underbrace{\frac{1}{2}||\mathbf{x}' - \mathbf{x}||^2}_{\text{reconstruction}} - \underbrace{\frac{1}{2}\left[\log \boldsymbol{\sigma}^2 - \boldsymbol{\mu}^2 - \boldsymbol{\sigma}^2 + 1\right]}_{\text{regularization}} \tag{2}$$

Balancing between the reconstruction term and regularization term:

$$\mathcal{L}_{\beta-\mathrm{VAE}}(\mathbf{x}) = \mathcal{L}_{\mathrm{recon}} + \beta \mathcal{L}_{\mathrm{KL}} \quad (\beta > 0) \tag{3}$$

- $\beta > 1$: disentangled representations $\nearrow$, reconstruction quality $\searrow$
- $\beta < 1$: disentangled representations $\searrow$, reconstruction quality $\nearrow$

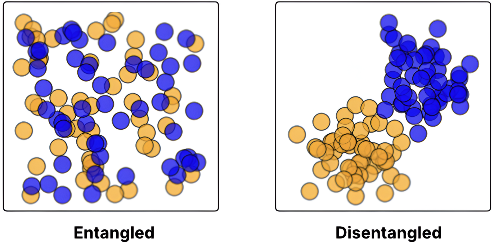
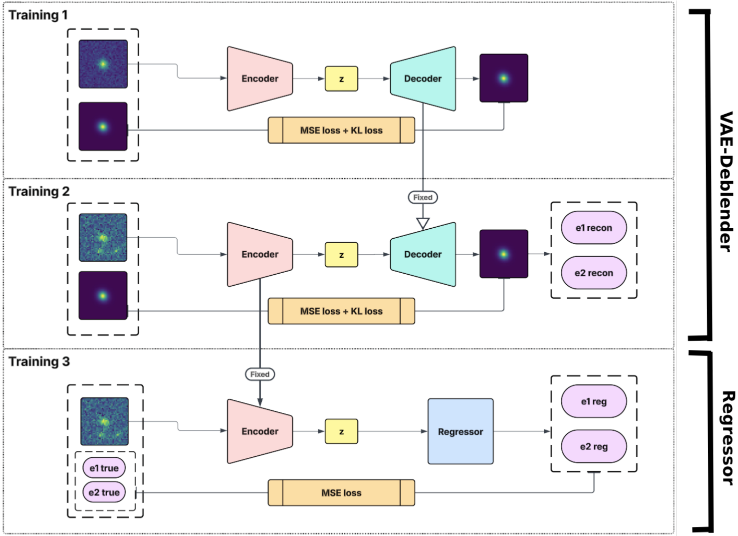


Table of Contents

- Introduction
- Variational Autoencoder
- Pipeline Overview
- Results
- Conclusion



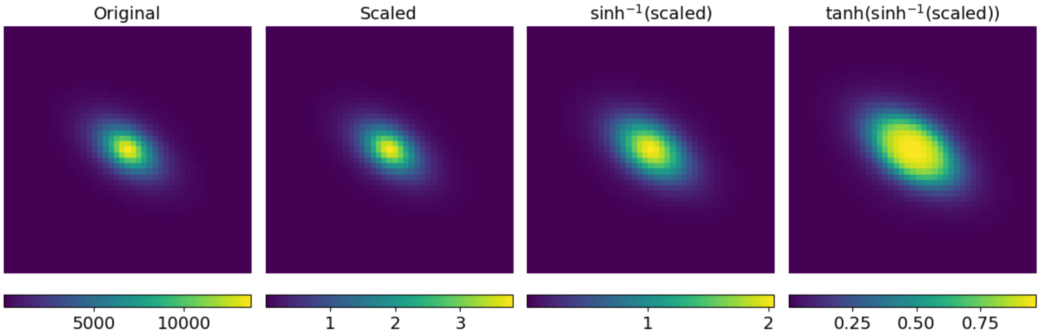
1. **Training 1**: Learning the representation of isolated galaxies.
2. **Training 2**: Deblend and reconstruct isolated galaxies from blended scenes.
3. **Training 3**: Extract weak lensing parameters from latent space.

Dataset Preparation: Normalization

Normalization function (Arcelin et al. 2020):

$$x_b = \tanh\left(\sinh^{-1}\left(\beta \frac{x_{\mathrm{raw},b}}{\langle \max\left(x_{\mathrm{raw},b}\right)\rangle_b}\right)\right) \tag{4}$$

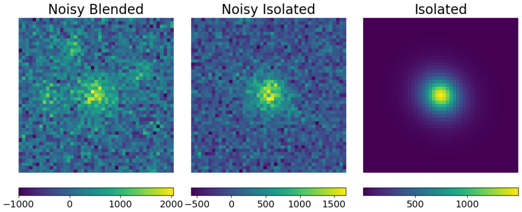
where $\langle \max\left(x_{\mathrm{raw},b}\right)\rangle_b$ is the mean of the distribution of the maximum pixel values in the b band of input images and $\beta = 2.5$.

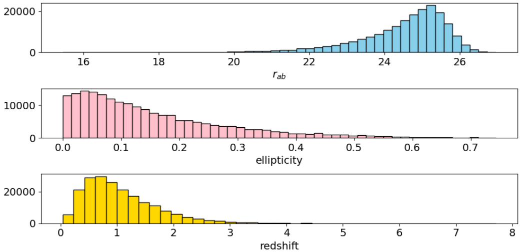


Training Dataset: simulated LSST data

Simulations for a 10-year LSST with `Blending ToolKits` (Mendoza et al. 2025)

- Catalog: `Catsim`
- Pixel size: 45×45
- 6 bands `ugrizY`
- 200,000 images
- Noisy isolated images for *Training 1*
- Blended images for *Training 2*
- Clean isolated images for optimization of *Loss function*





Testing Dataset: real DES data

- **Assumption:** Ellipticities from `SVA1 im3shape` are assumed to be ground truth.
- **Selection function:**
  - Valid ellipticity components (e_1, e_2).
  - SNR > 20.
  - Matching with DES "gold" catalog to obtain the position `(RA, Dec)`.
- **Image retrieval:**
  - `Image Cutout Tool` from NOIRLab Astro Data Lab
  - Brightest galaxy in the center.
  - 5 bands `grizY`

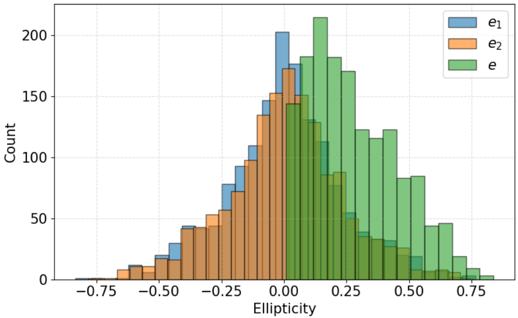
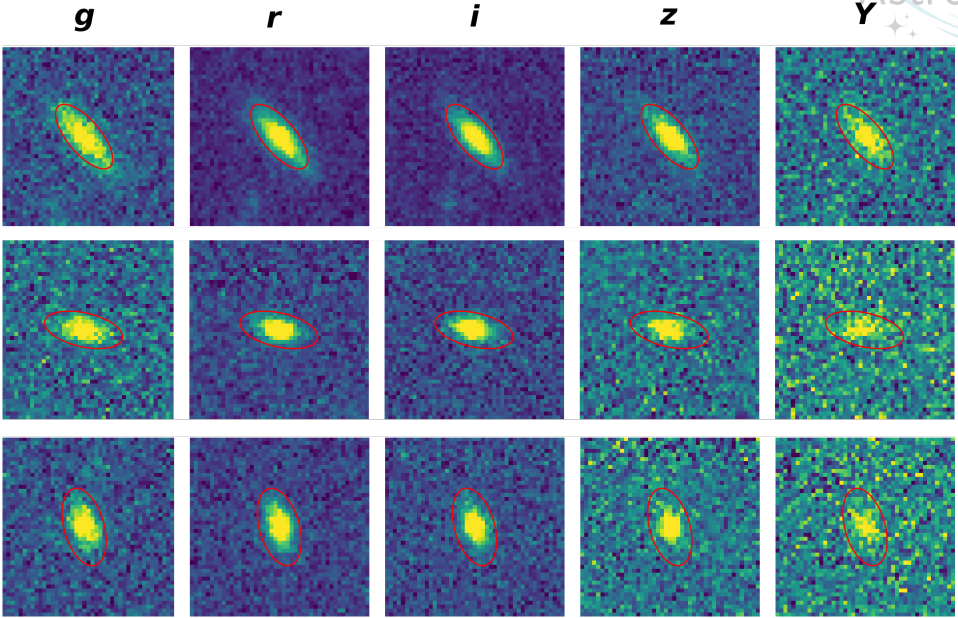


Figure 5: Distribution of ellipticities selected from `im3shape` catalog

real DES data



Dynamic Range

Dynamic range: the range of pixel intensities

- To apply the pipeline from LSST to DES, the dynamic ranges have to be matched for each band.

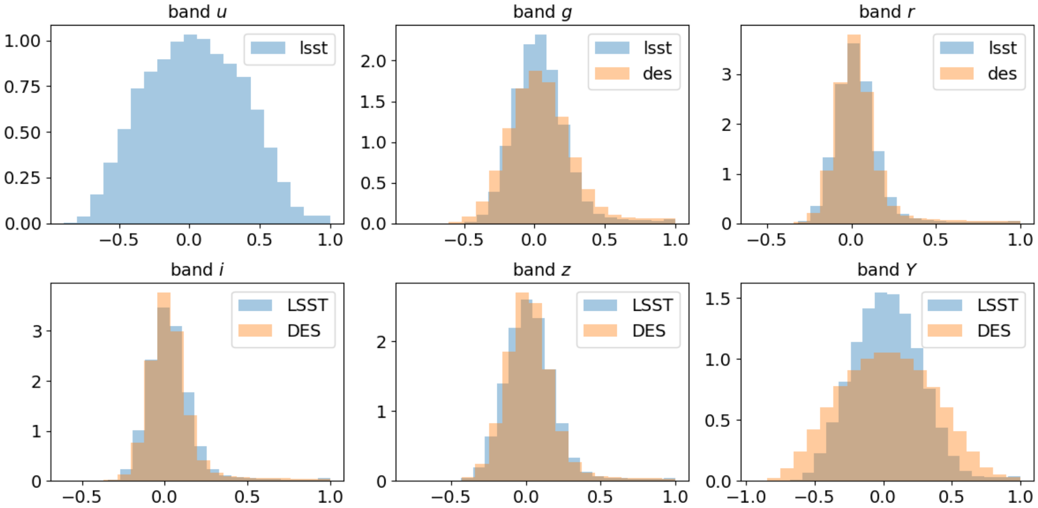
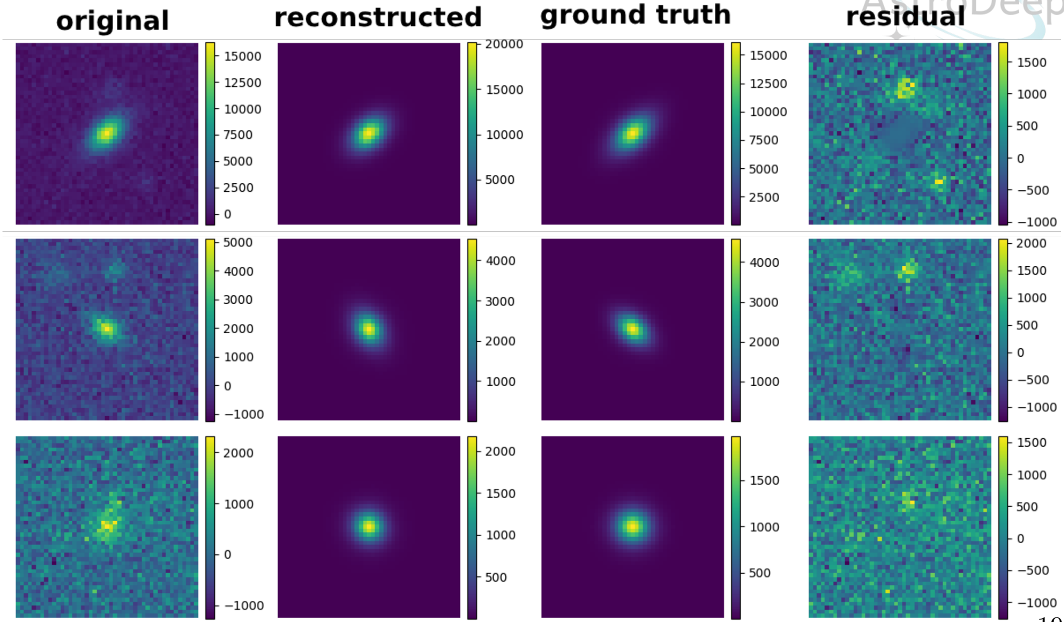


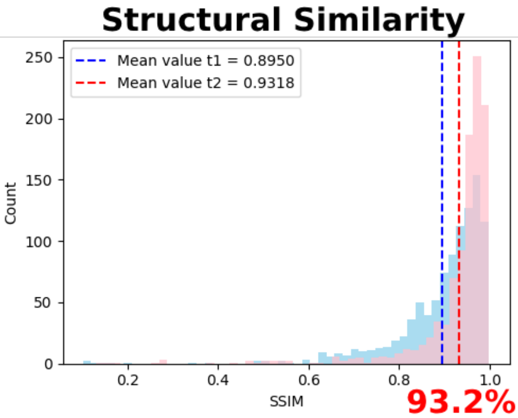
Table of Contents

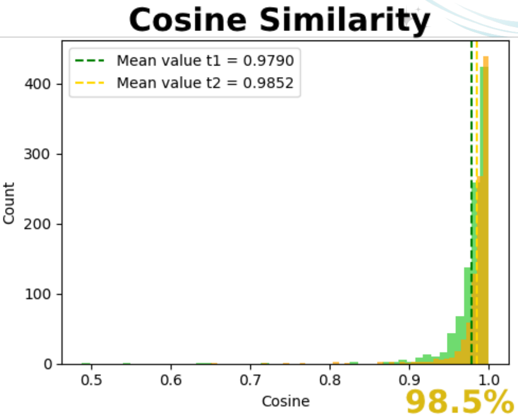
- Introduction
- Variational Autoencoder
- Pipeline Overview
- Results
- Conclusion

Deblender trainings: VAE reconstruction



Morphology & shape of reconstructed images

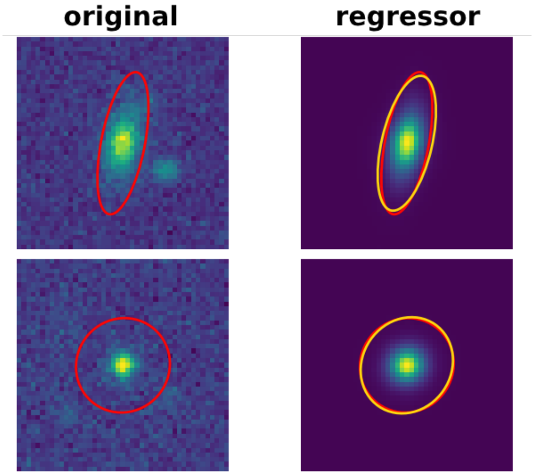




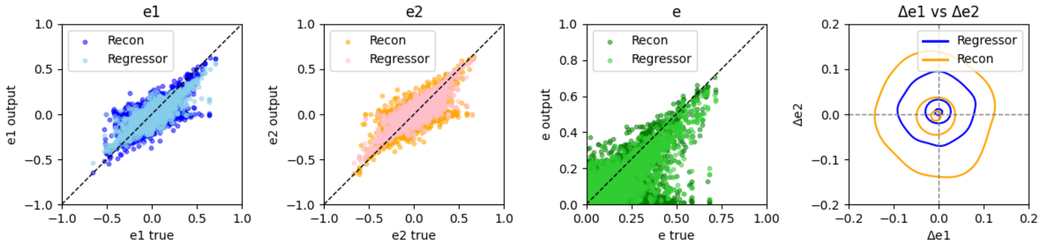
$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{\left(\mu_x^2 + \mu_y^2 + c_1\right)\left(\sigma_x^2 + \sigma_y^2 + c_2\right)}$$

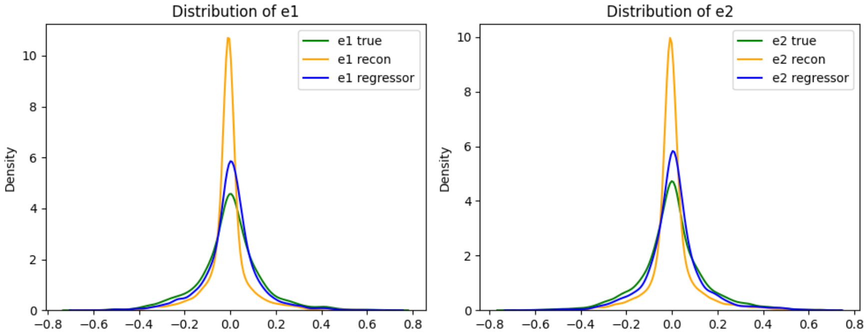
$$S_C(A,B) = \frac{\mathbf{A}\cdot\mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|}$$

Deblender trainings: Regressor

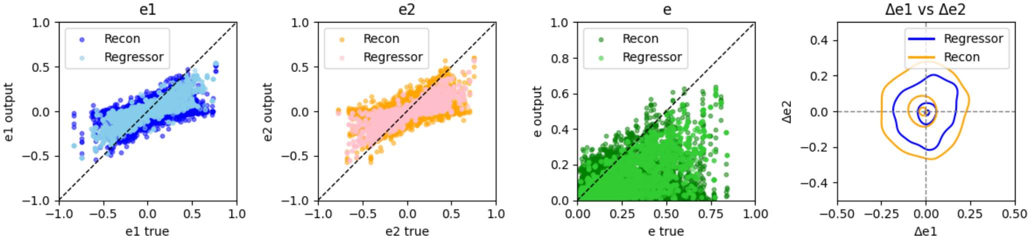


Ellipticity: simulated LSST





Ellipticity: real DES



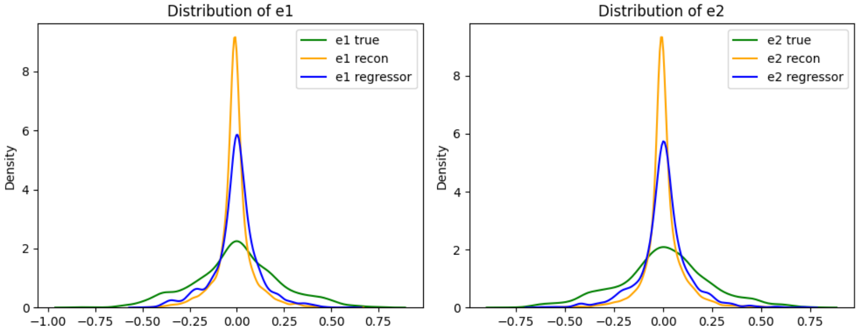


Table of Contents

- Introduction
- Variational Autoencoder
- Pipeline Overview
- Results
- Conclusion

Conclusion

Key Points

- **Data:** Simulated LSST-like images + real DES data.
- **Method:** VAE-based deblender + ellipticity regressor.
- **Result:** Regressor outperforms VAE in ellipticity accuracy.
- **Limits:** Roundness bias; weaker performance on real DES images.

Conclusion

Key Points

- **Data:** Simulated LSST-like images + real DES data.
- **Method:** VAE-based deblender + ellipticity regressor.
- **Result:** Regressor outperforms VAE in ellipticity accuracy.
- **Limits:** Roundness bias; weaker performance on real DES images.

Next Steps

- Normalizing flows → reduce roundness bias.
- Better generalization to real survey data.
- Extract redshift from latent space.