Large Synoptic Survey Telescope

université PARIS DIDEROT PARIS 7

Deblending galaxies with Variational Autoencoder: a multi-bands, multi-instruments analysis

(*Arcelin, Doux, Roucelle, Aubourg*)

Bastien Arcelin
03/04/2019

Workshop GPU - CC IN2P3

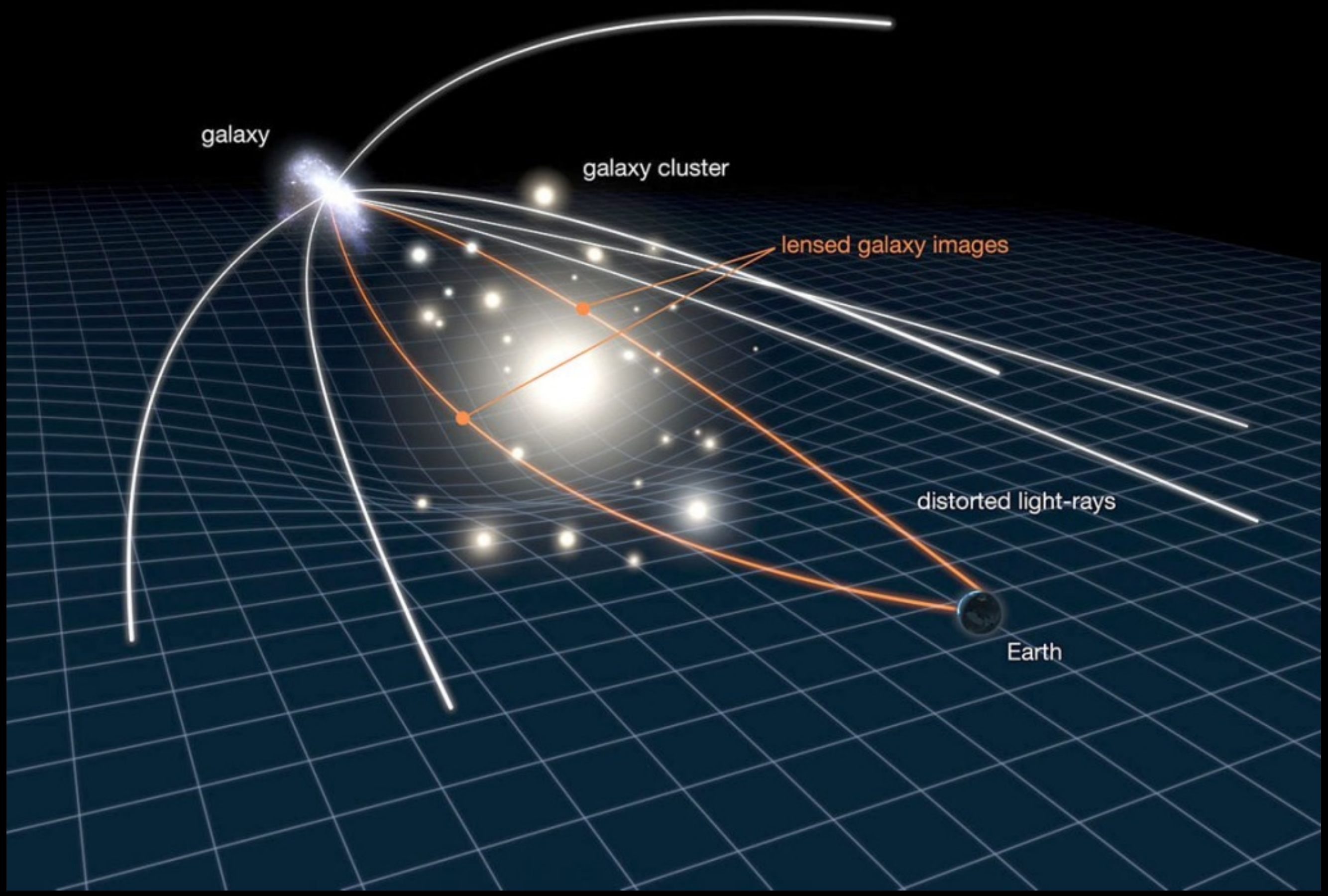
LSST: need for machine learning ?

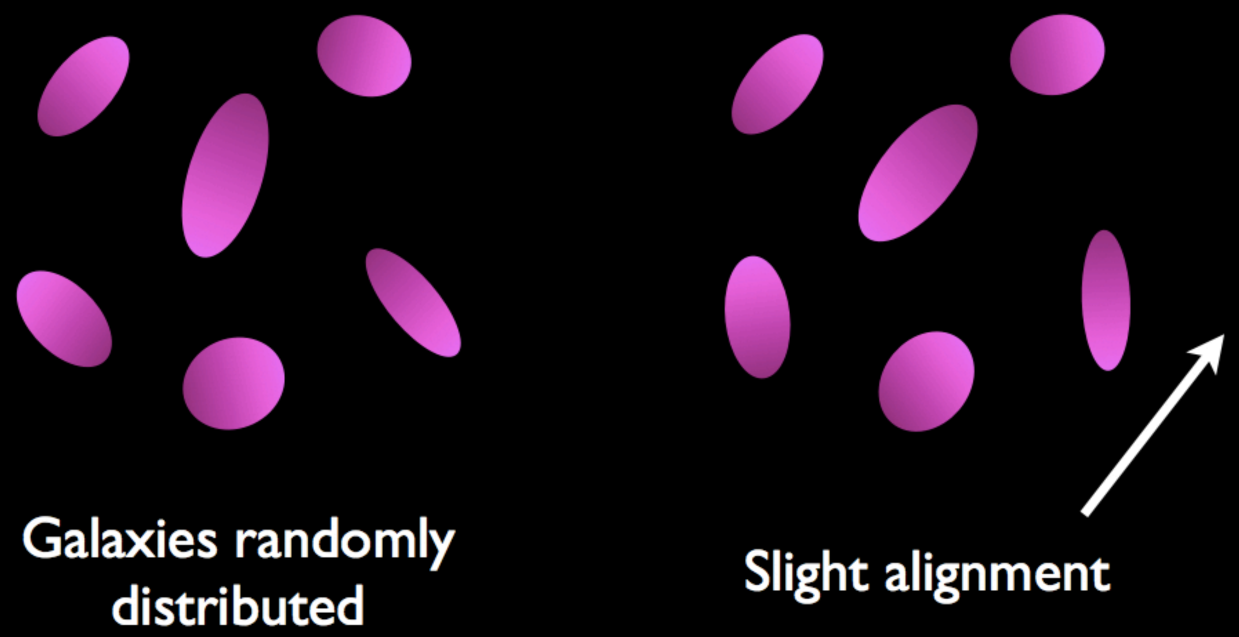
- LSST : Large Synoptic Survey Telescope
- Being built north Chile
- **20 terabytes of data per night**
- 10 years of operation : **around 60 petabytes** of data

Need fast analysis methods —> Machine learning

Weak lensing to probe dark energy

- **Lensing** due to mass: **deforms** the images of the galaxies
- Correlations between orientations and shapes of neighbour galaxies: **cosmic shear**





LSST Data

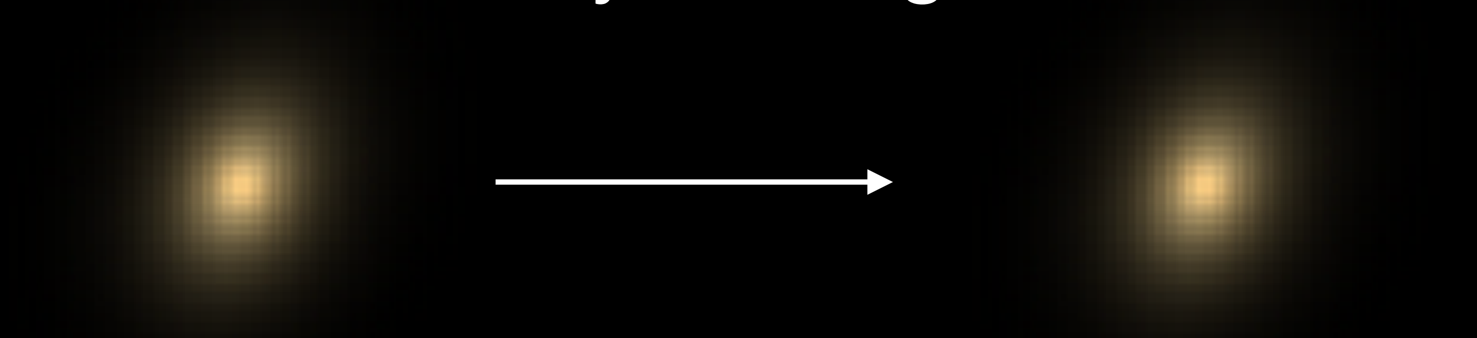
- Will look a lot like HSC data (hsc-release.mtk.nao.ac.jp/hscMap2/)
- For HSC 58% of the detected objects are identified as blended*
- Systematic in shear measurement

From Cyrille Doux
(LSST Europe, 14 Juin 2018)

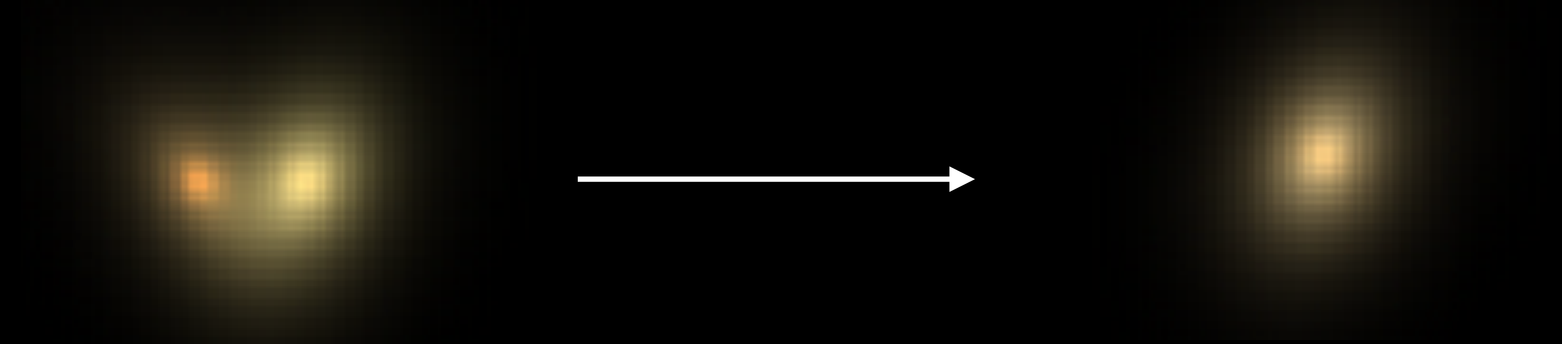
* Bosch et al. (2017)

Machine learning for deblending

- Two neural networks with each a dedicated function:
  - the variational autoencoder (VAE): reproduce precisely the shape and luminosity of the galaxies



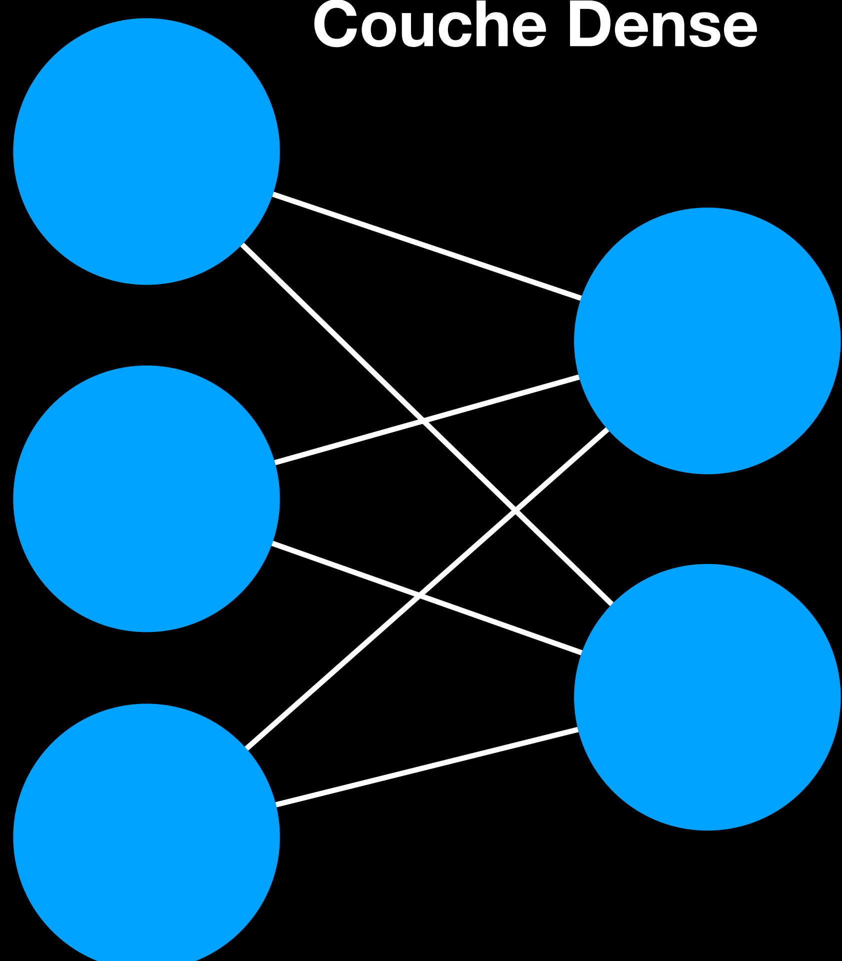
  - the deblender : separate the blended galaxies and reproduce accurately the shape and luminosity of the centered galaxy



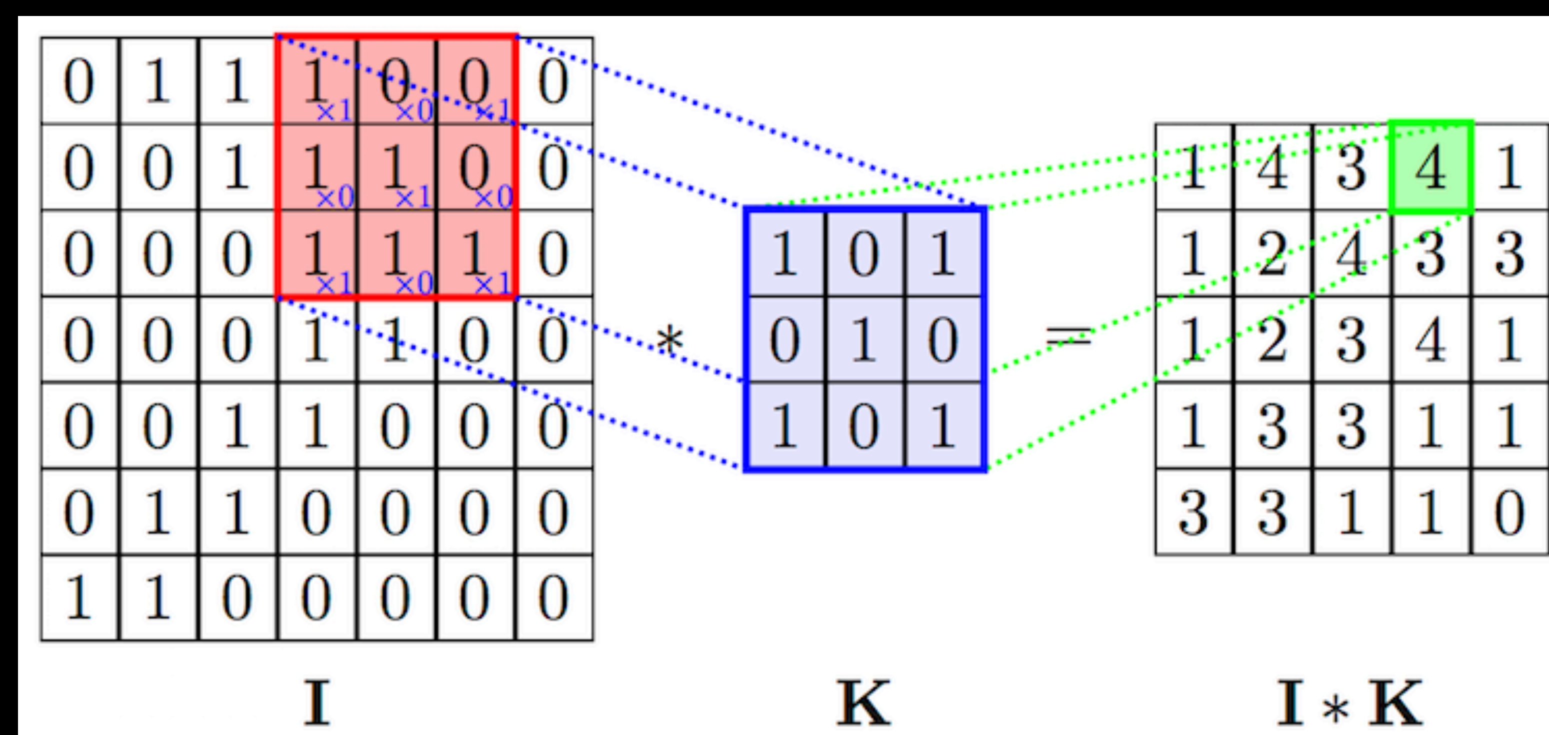
Architecture de réseaux de neurones

- Différentes architectures sont possibles:

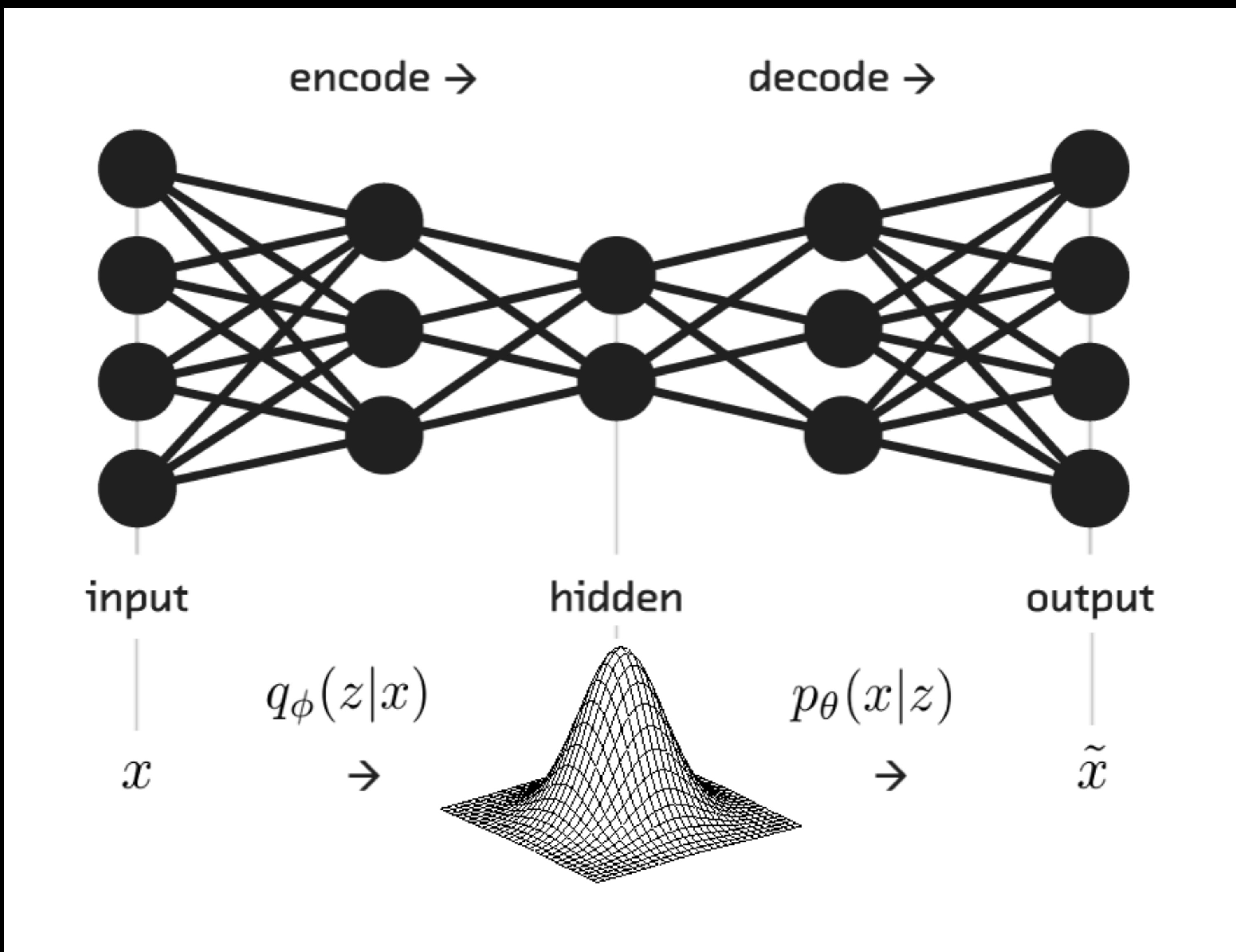
Couche Dense



Couche Convolutionnelle



Variational autoencoder



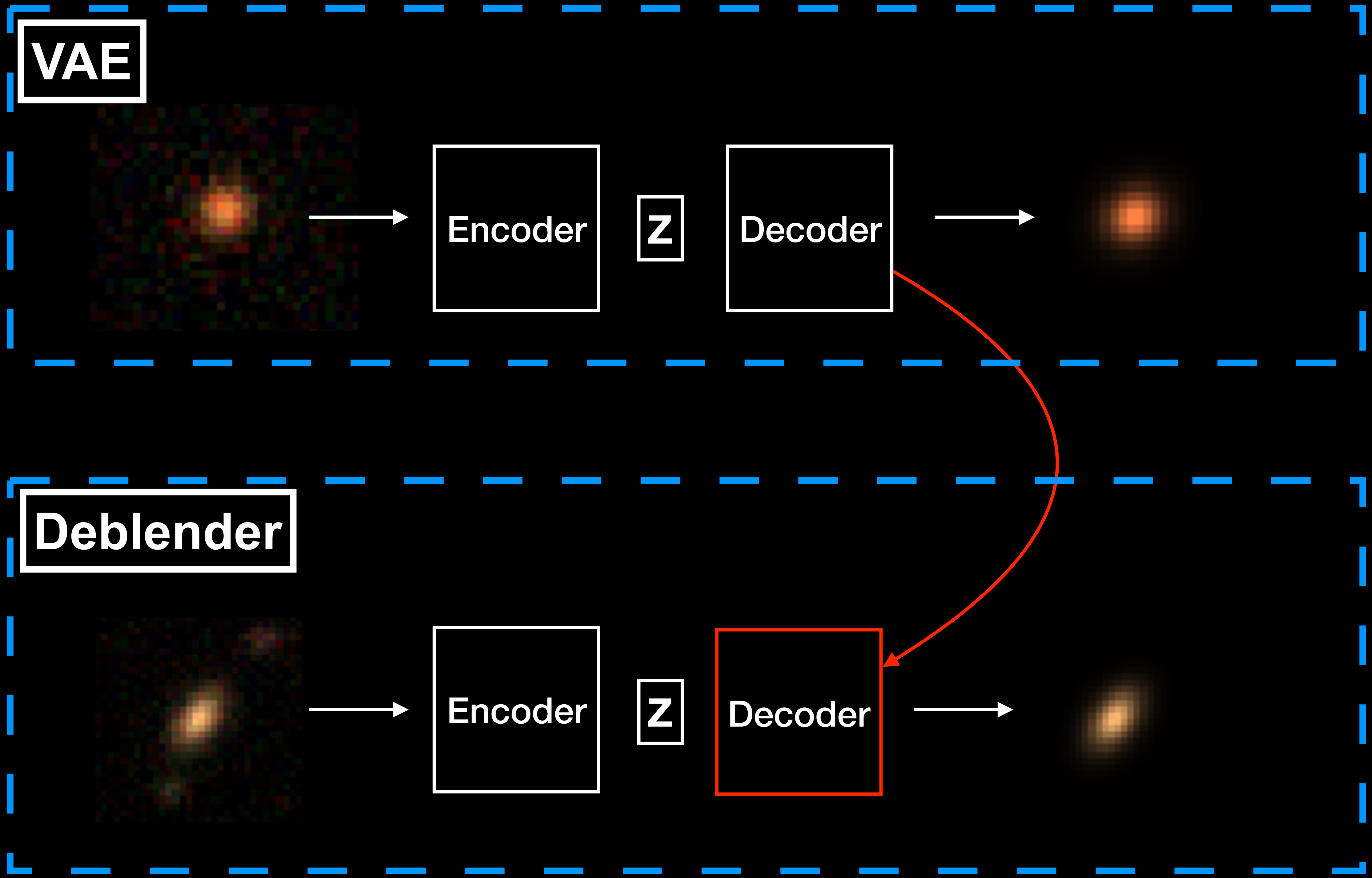
- http://blog.fastforwardlabs.com

Loss function :

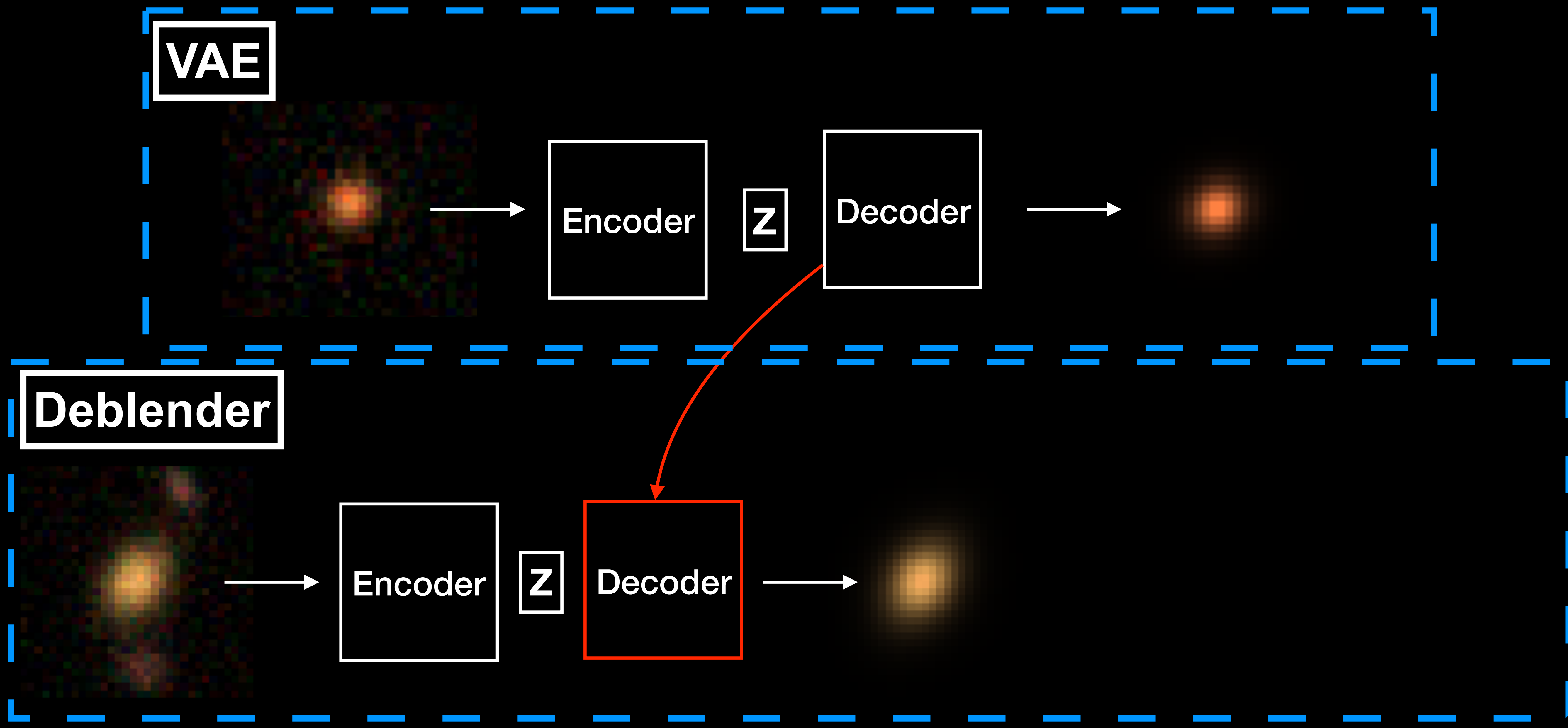
$$ELBO = -D_{KL}\left(q_\phi(z|x) \parallel p(z)\right) + \mathbb{E}(\log p_\theta(x|z))$$

Encoder **Decoder**

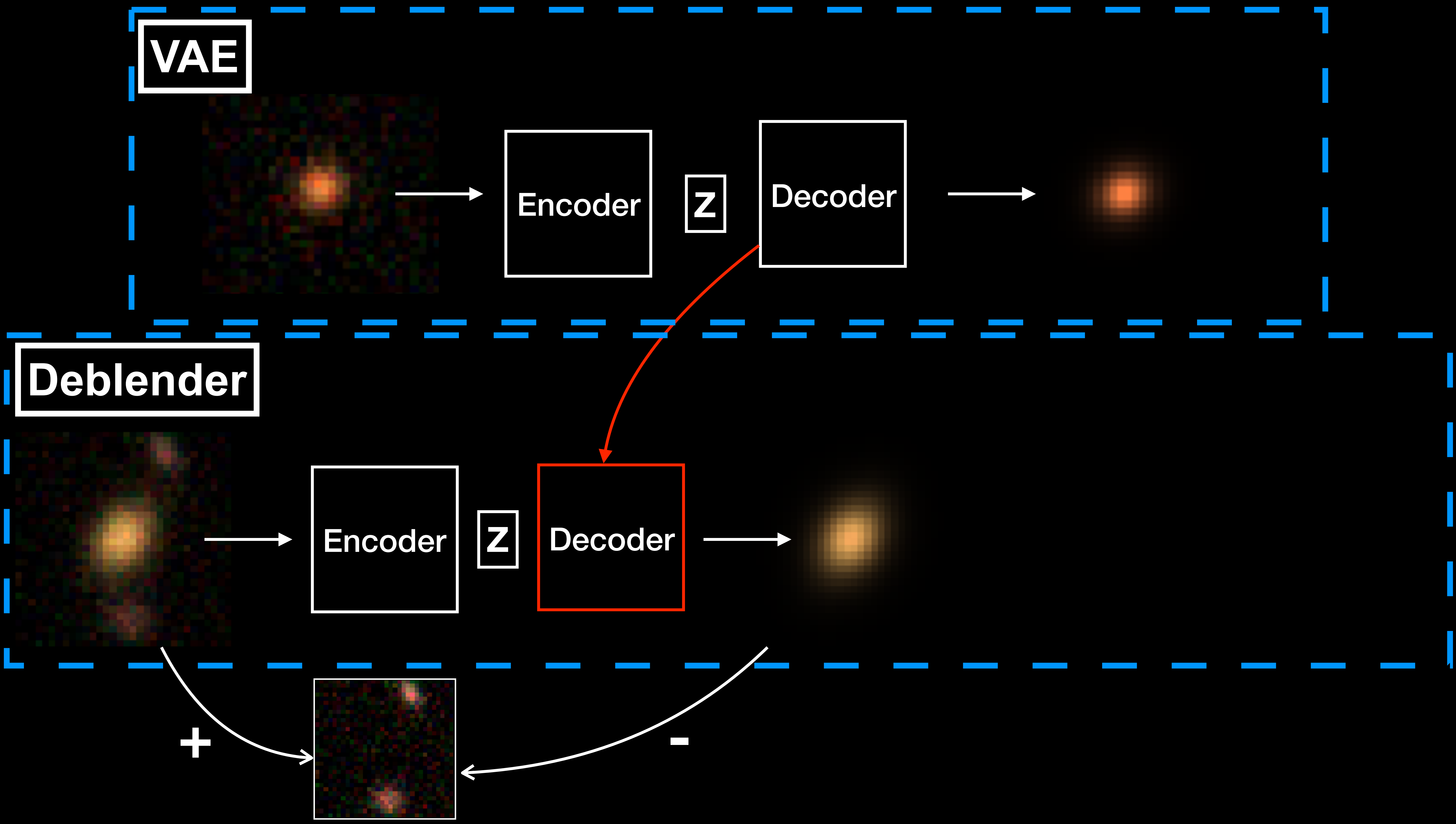
Pipeline



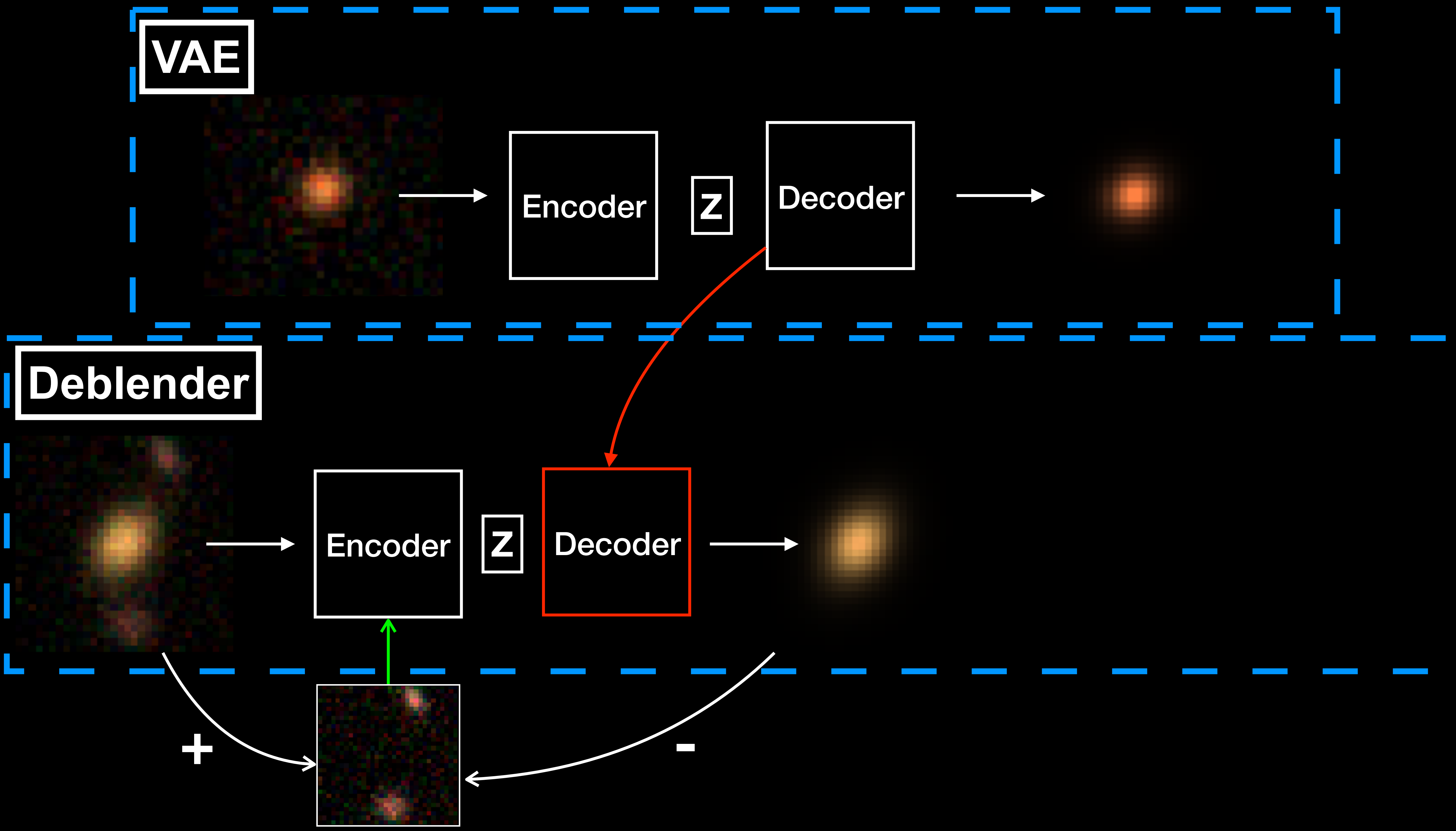
Pipeline for iterative deblender



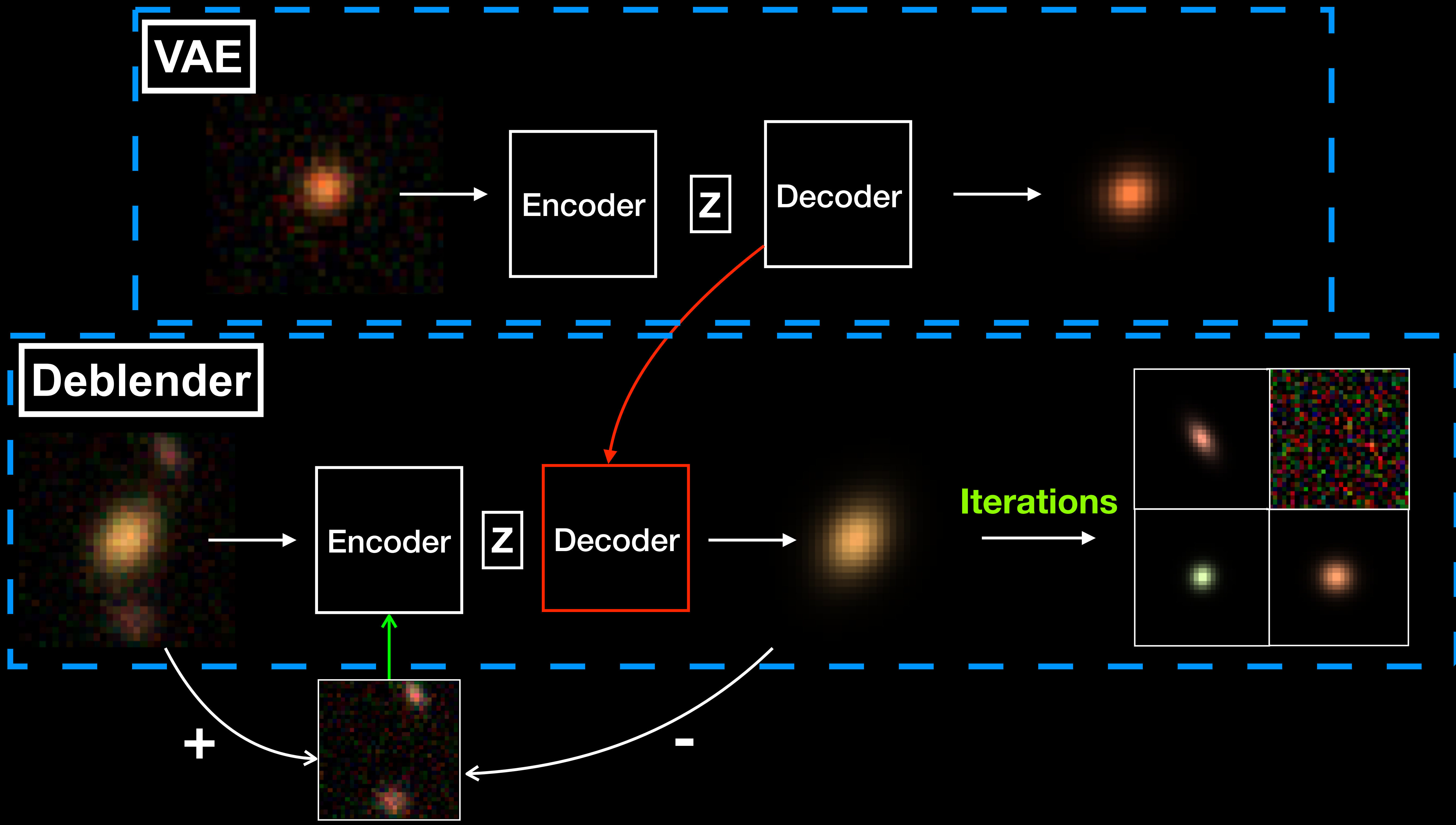
Pipeline for iterative deblender



Pipeline for iterative deblender

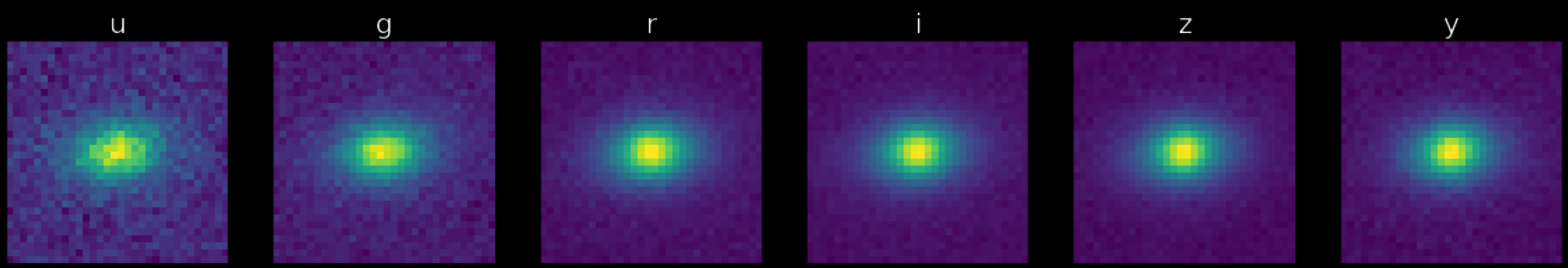


Pipeline for iterative deblender



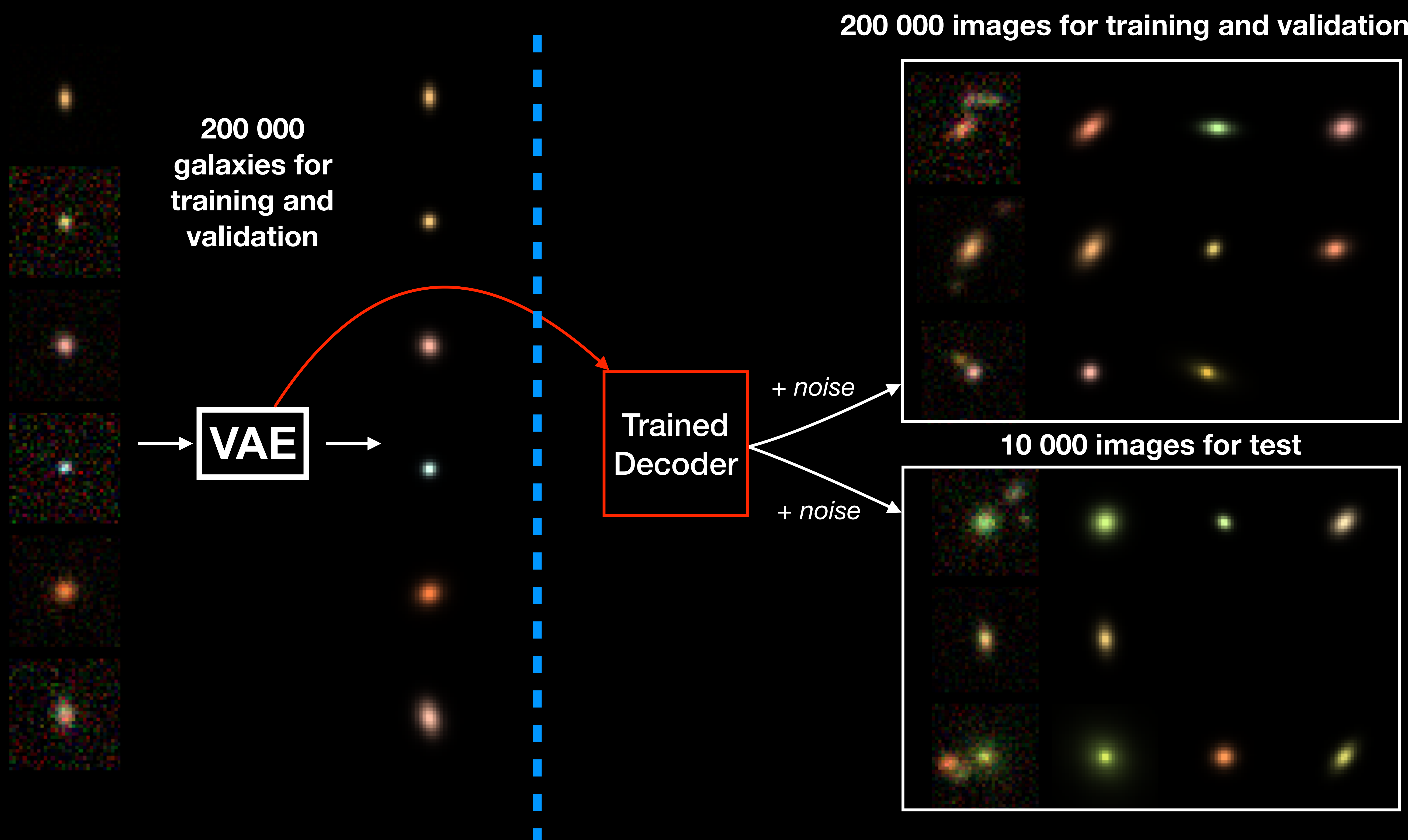
Training and validation data

- 200 000 galaxies simulated with GalSim from COSMOS catalog (Hubble Space Telescope) : 6 LSST pass-band filters

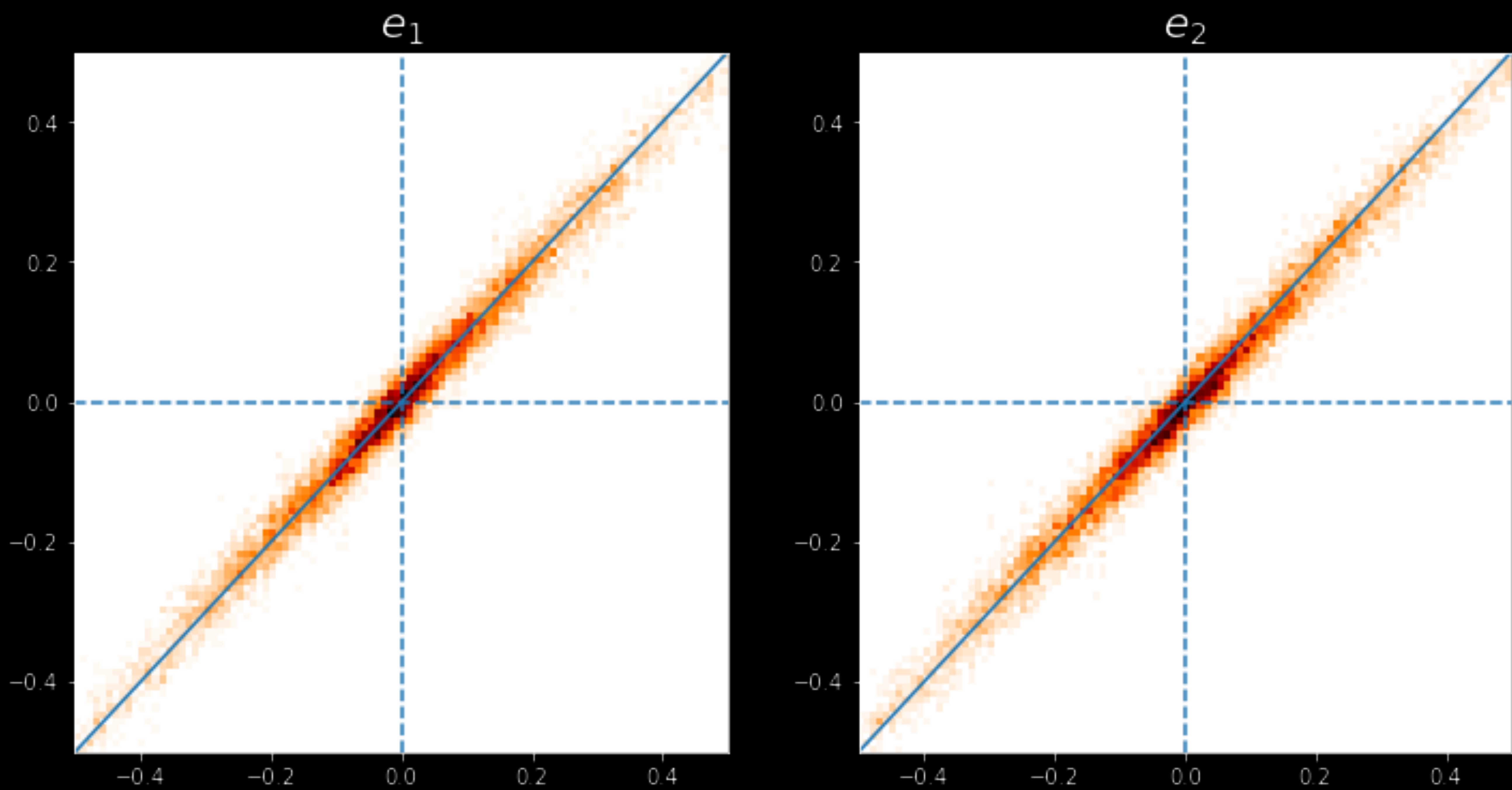


- Use the trained decoder from the VAE to create 200 000 images of blended galaxies (around 60 time faster than GalSim)

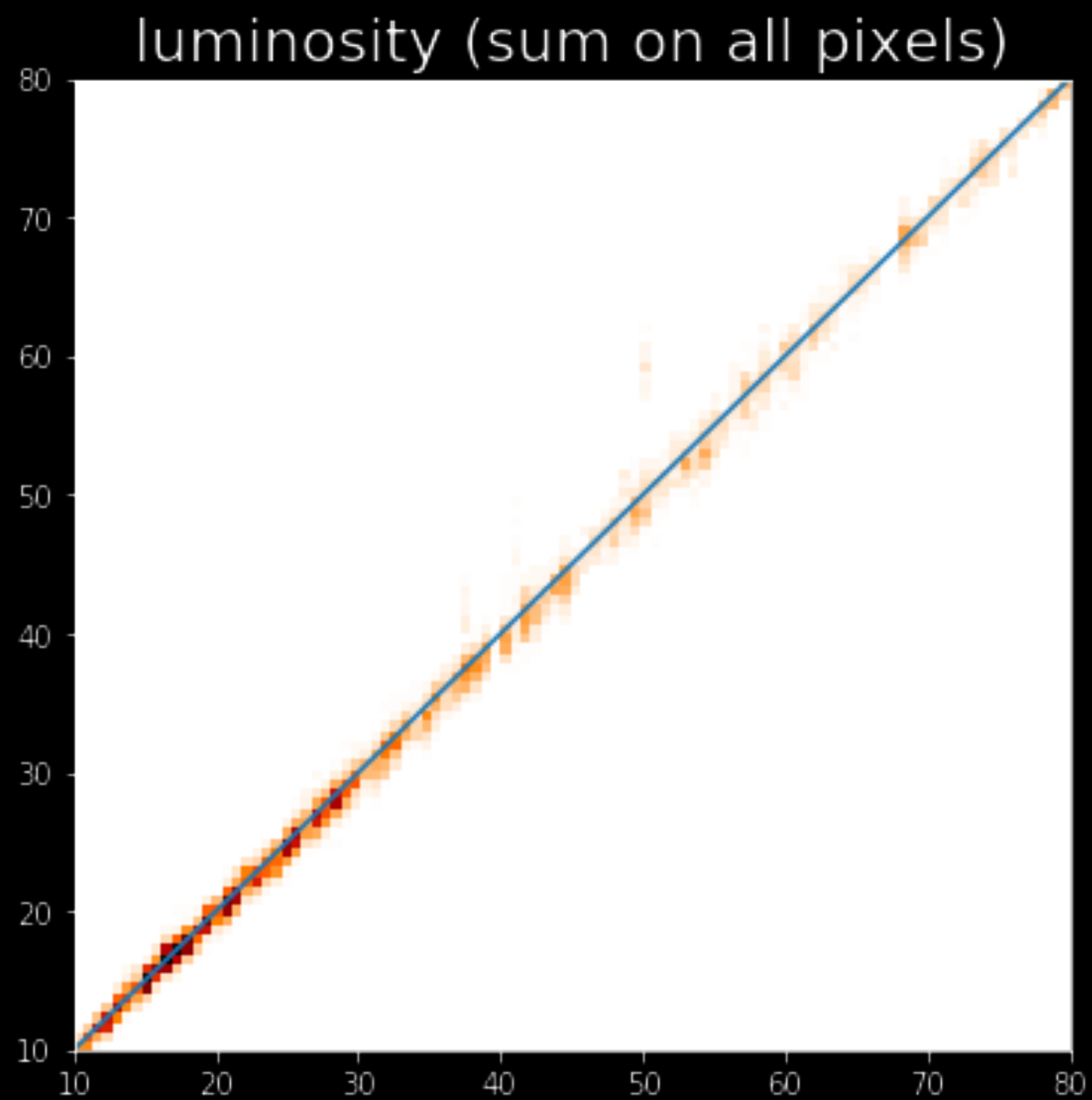
Training and validation data



VAE: with LSST R band-pass filter



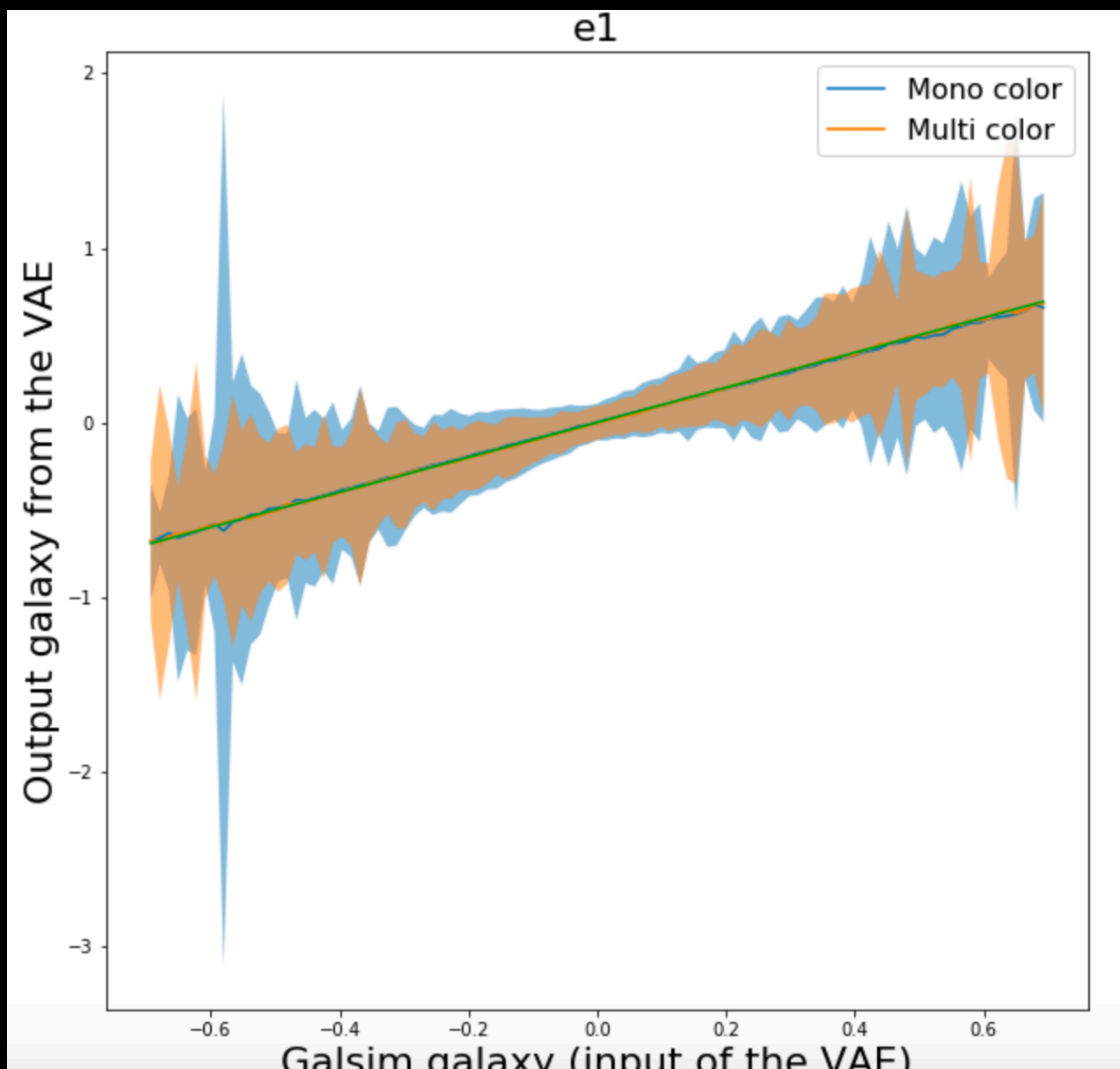
VAE: with LSST R band-pass filter

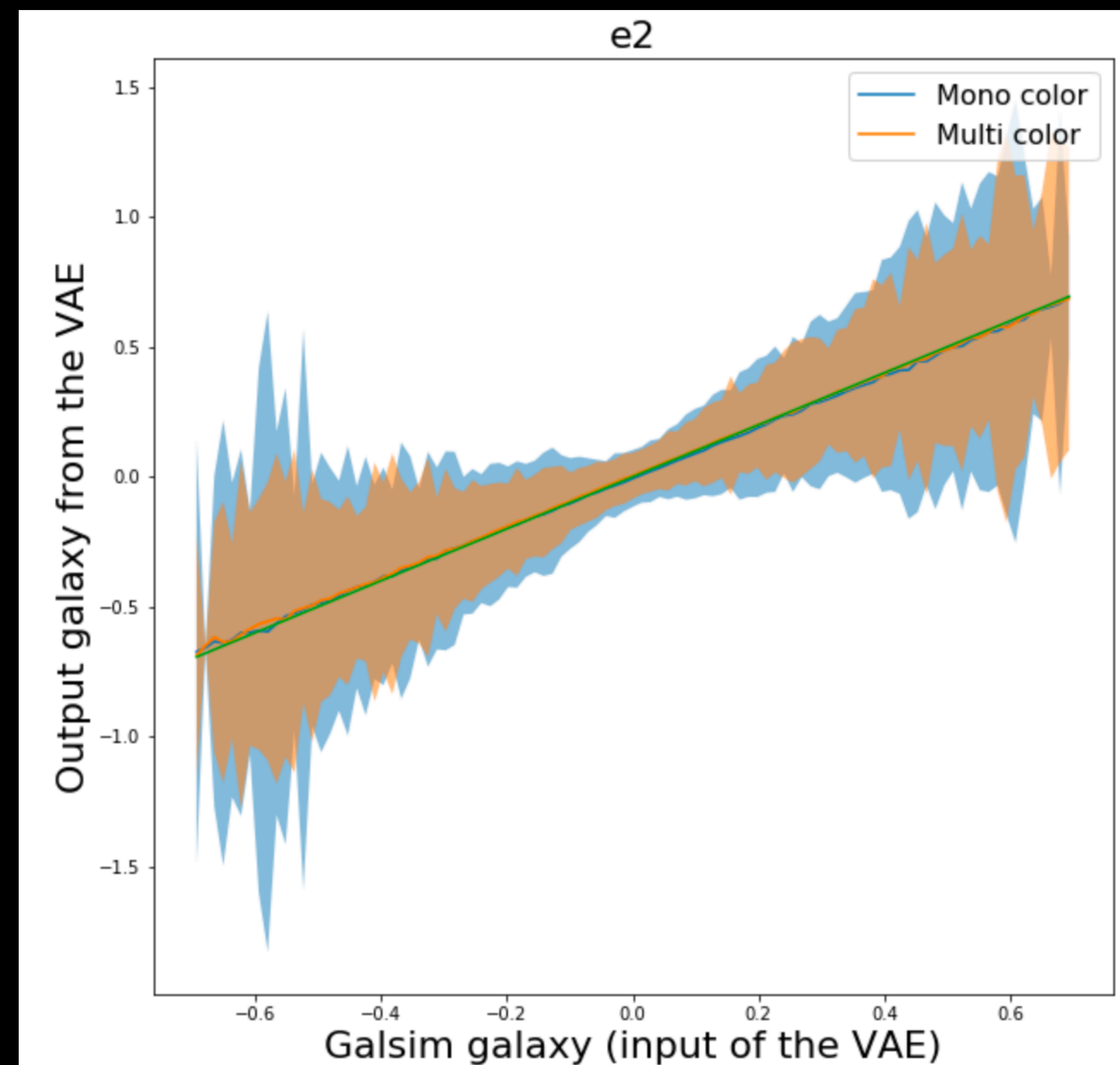


VAE

mono-wavelength/multi-wavelengths comparison

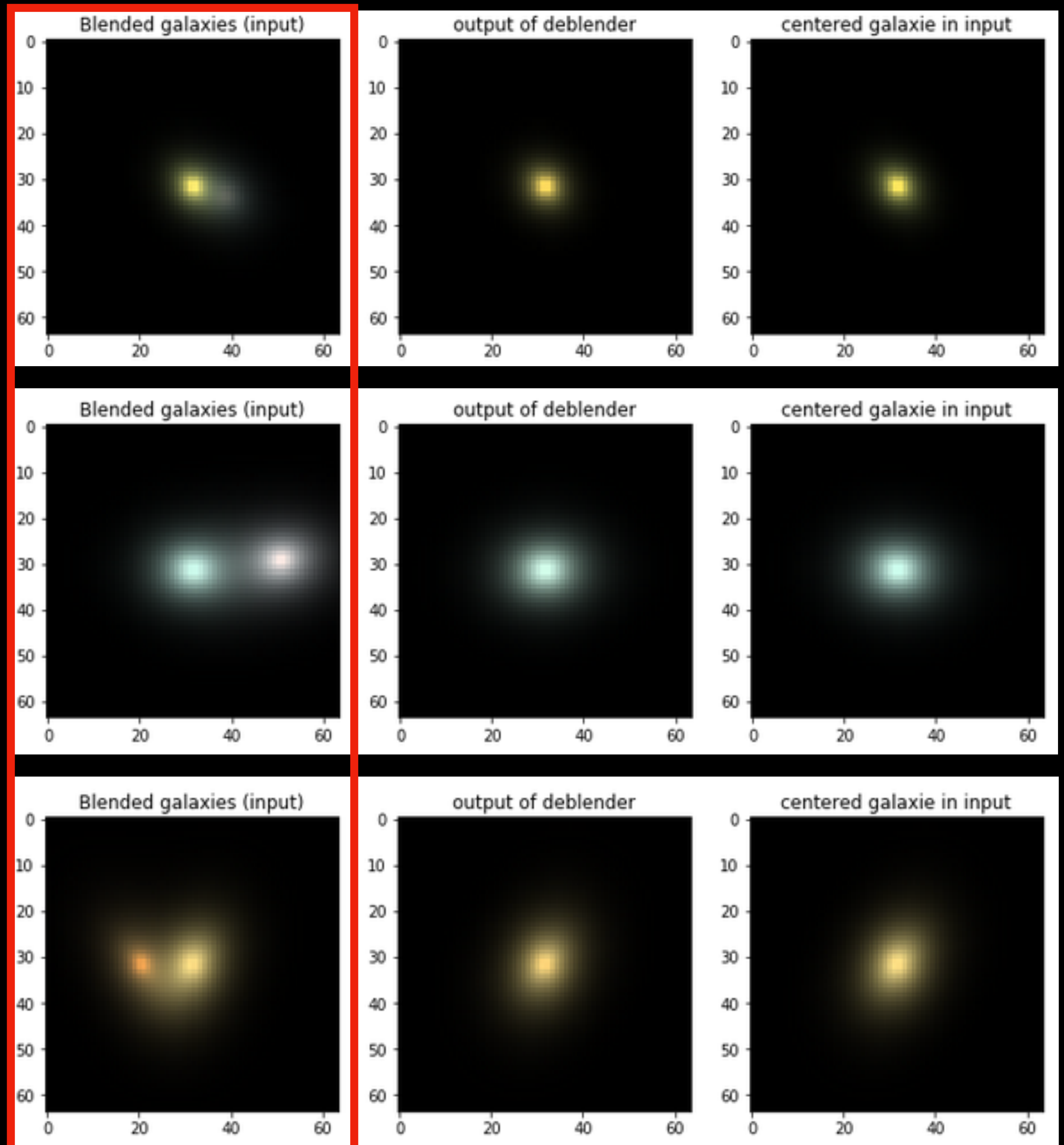
- 6 passbands of LSST (comparison done on r band)



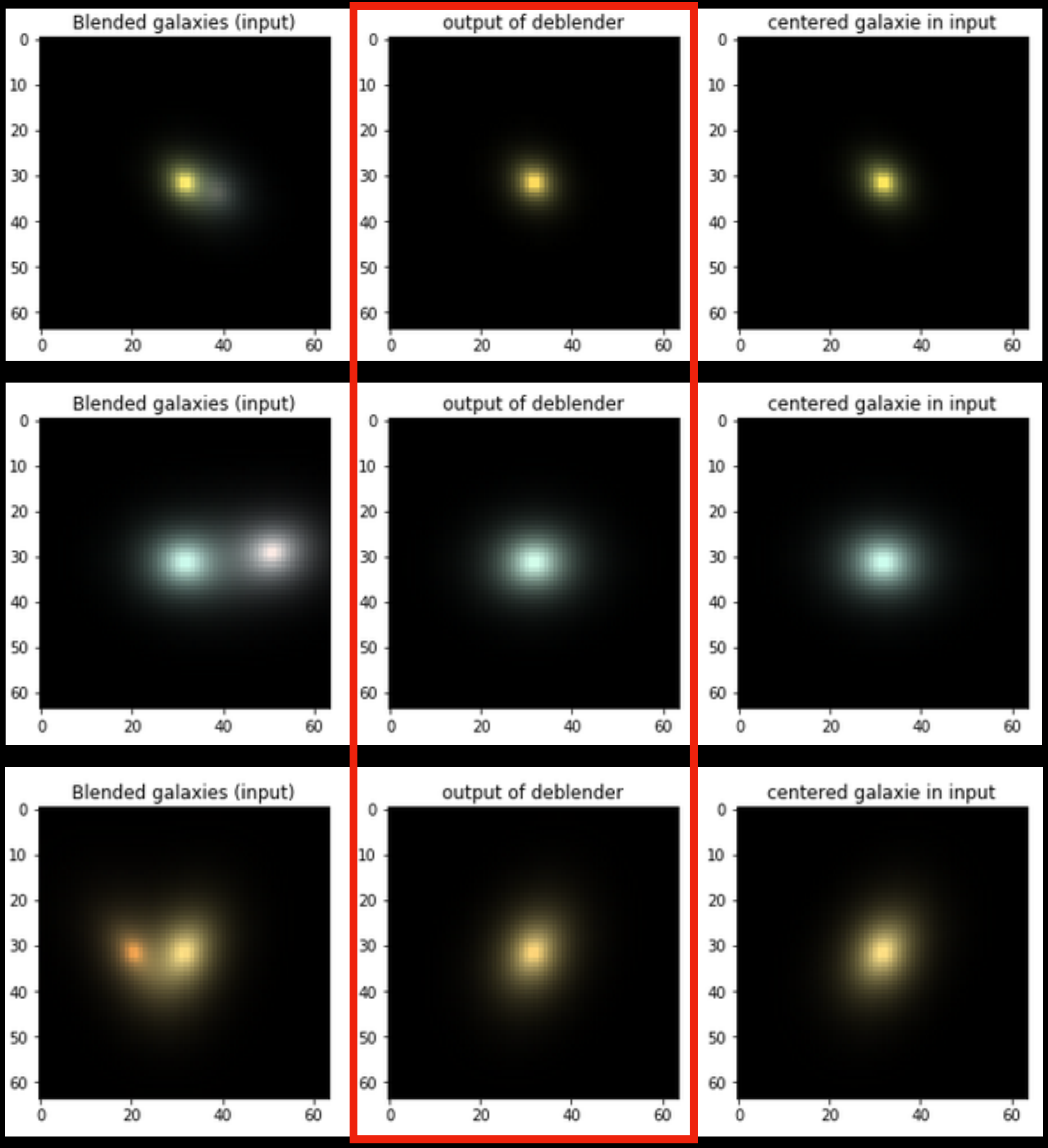


➡ VAE gives better results using 6 bands rather than one

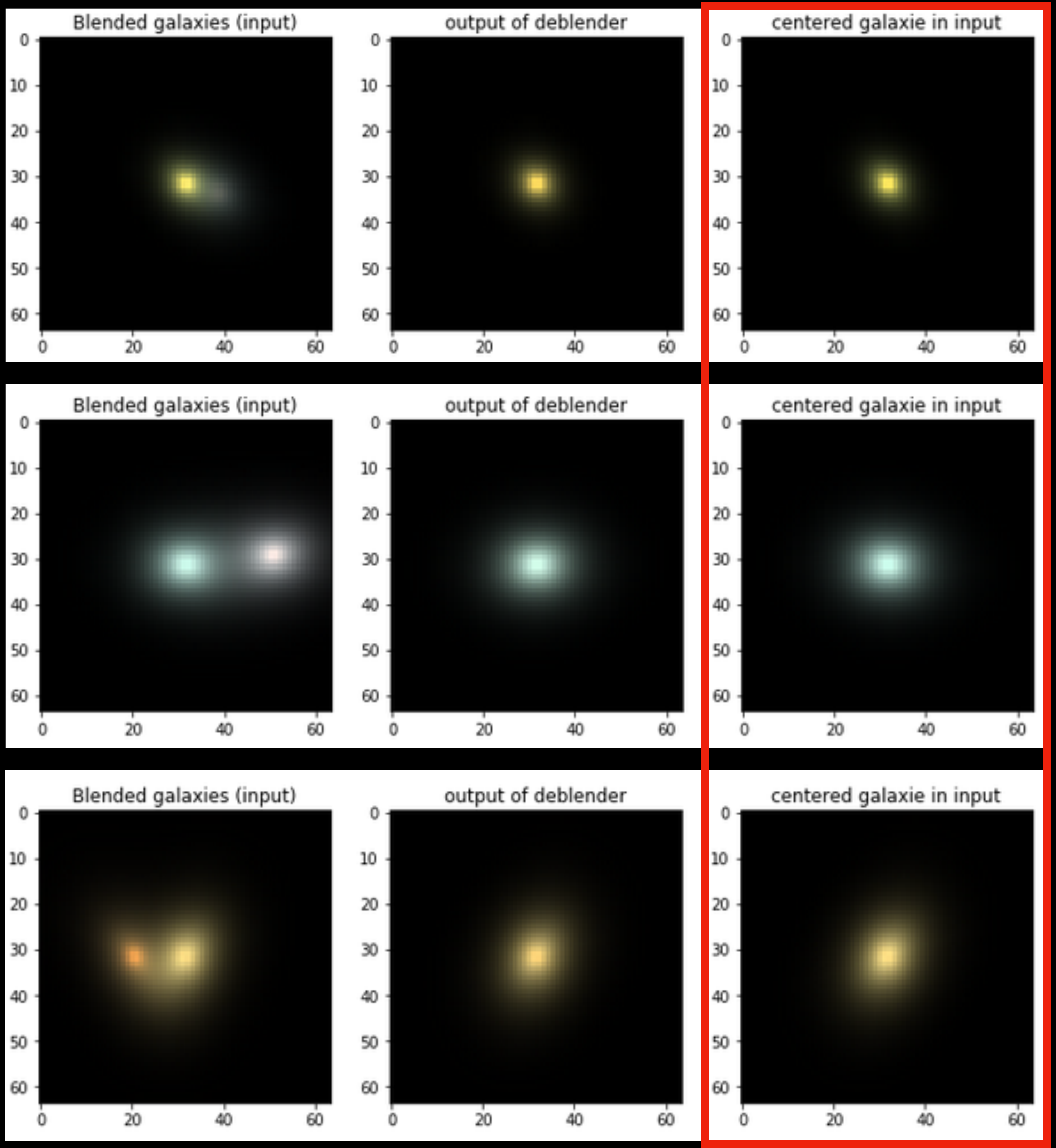
Deblender : results



Deblender : results

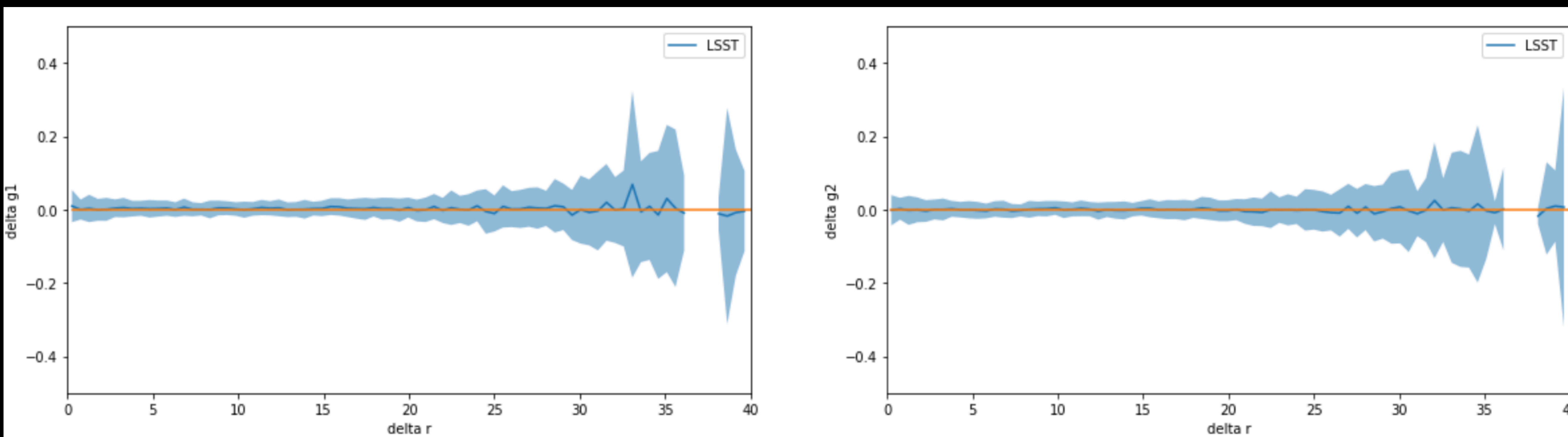


Deblender : results



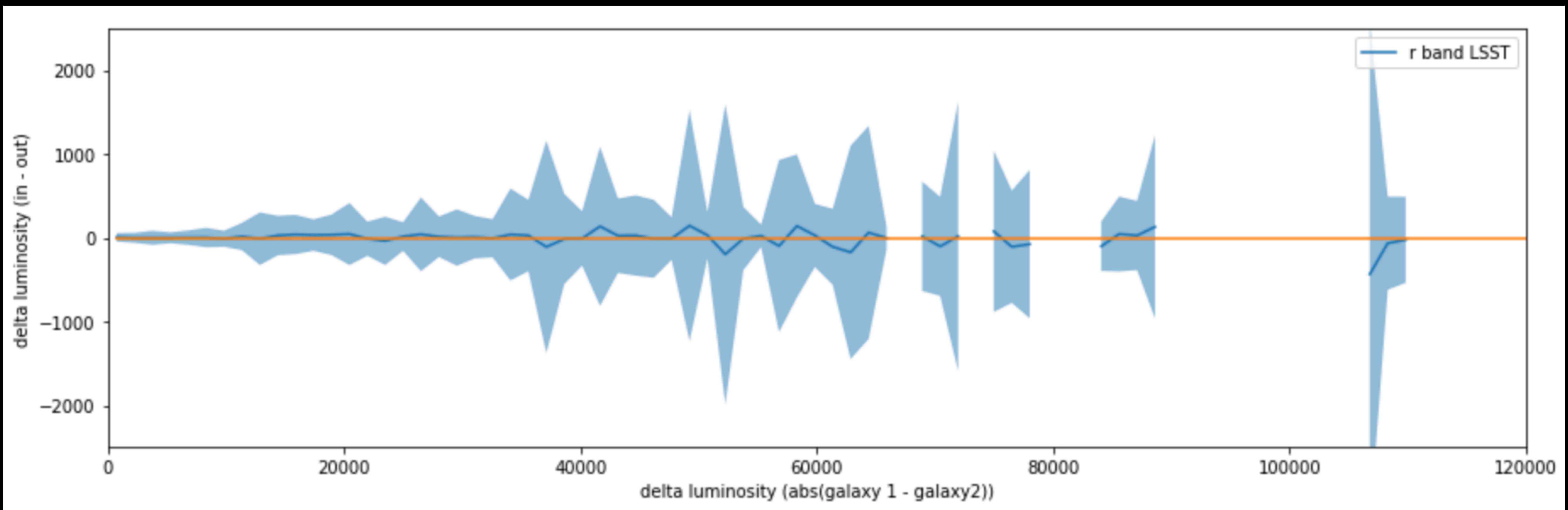
Deblender : results

shape reproduction



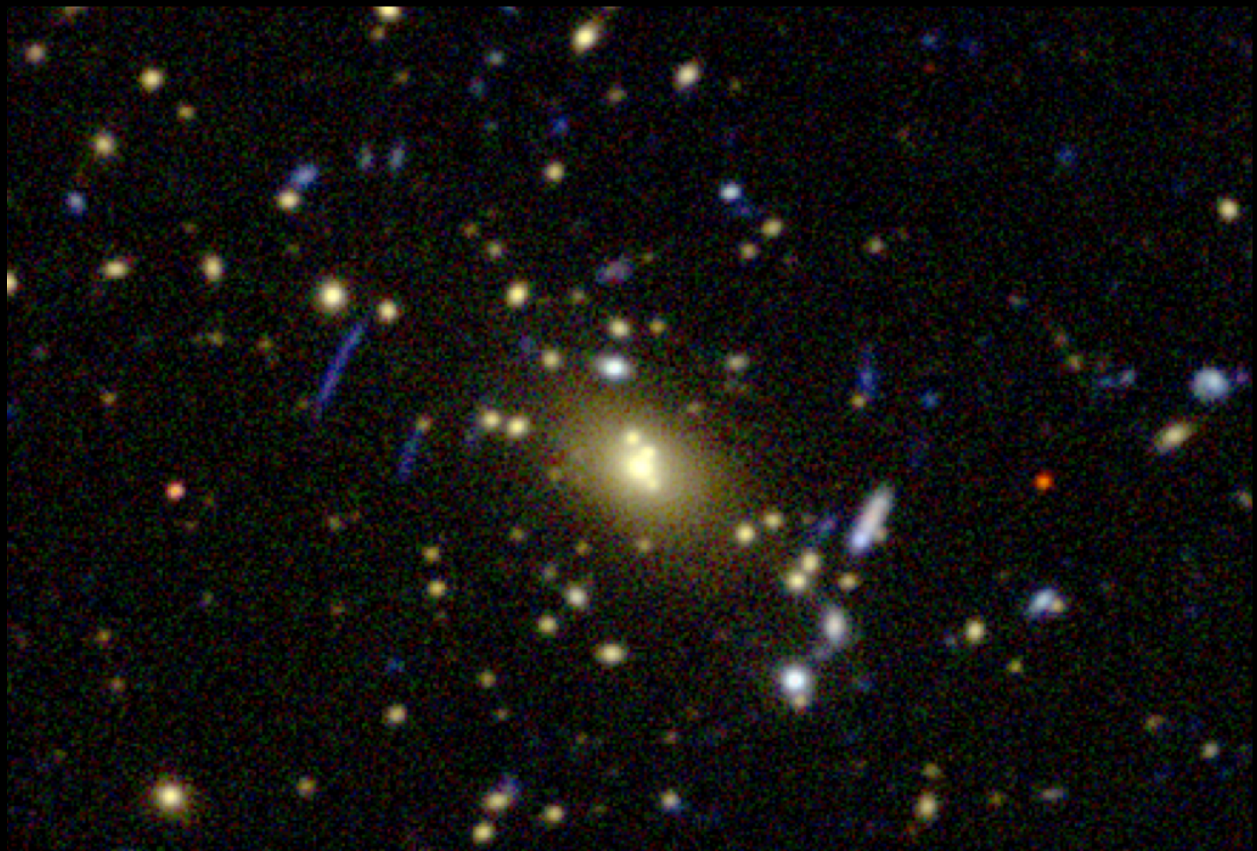
Deblender : results

luminosity reproduction



LSST+Euclid Data

- Why using Euclid data:
  - ✓ Adding infrared bands (x3)
  - ✓ Adding a visible band
  - ✓ Different PSF

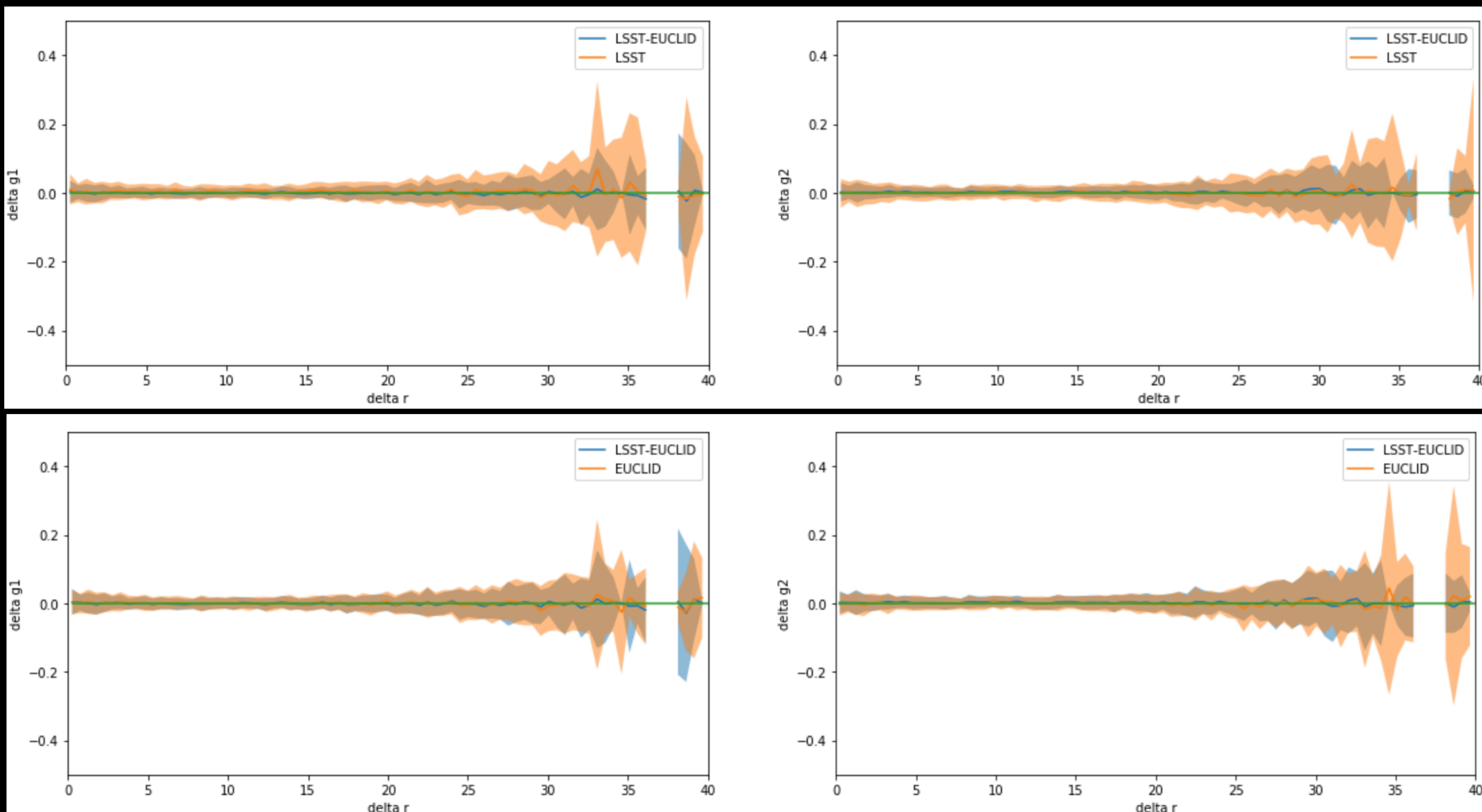


DES data - Peter Melchior's slides

CLASH WFC3/IR data - Peter Melchior's slides

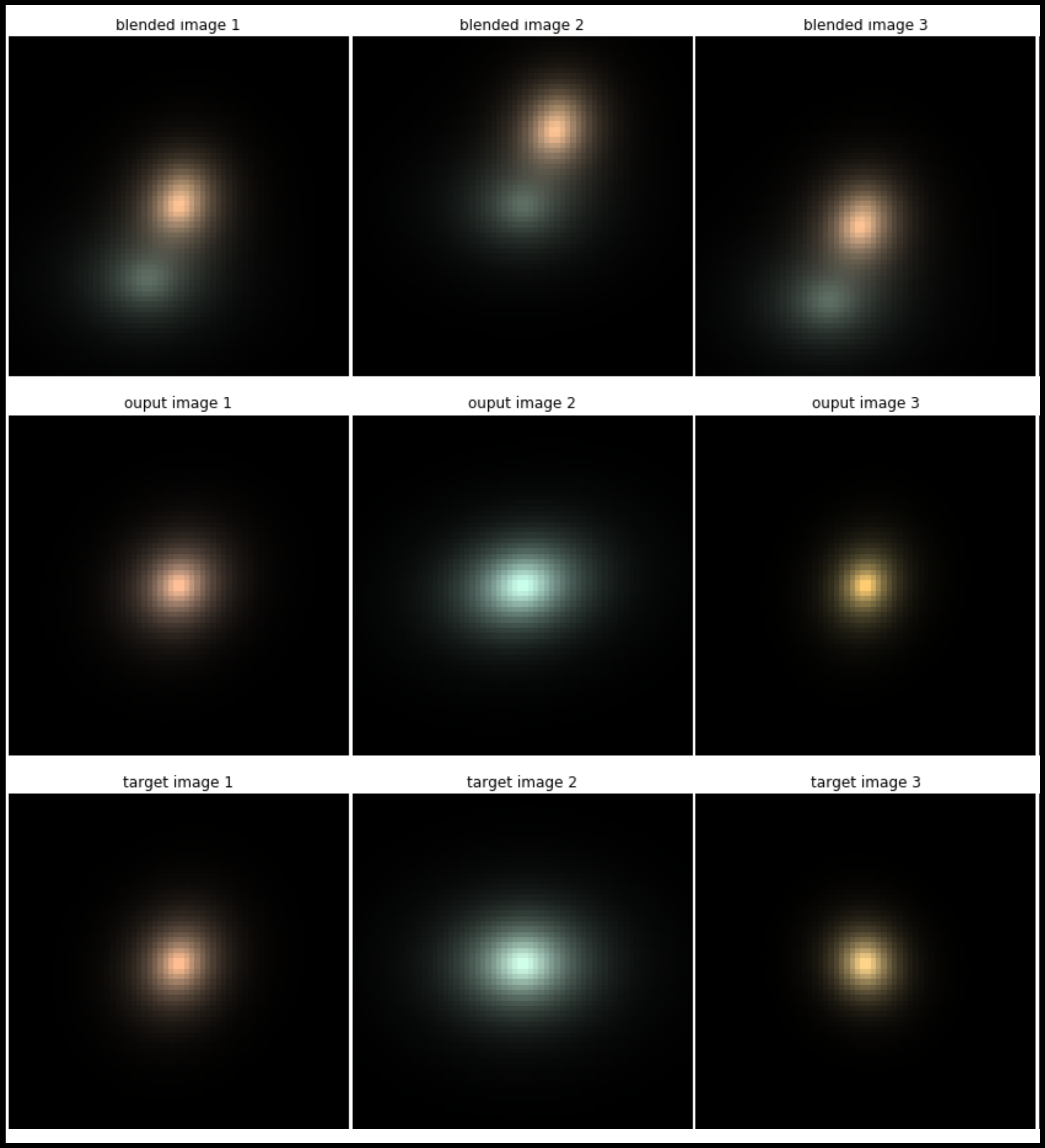
LSST + Euclid : results

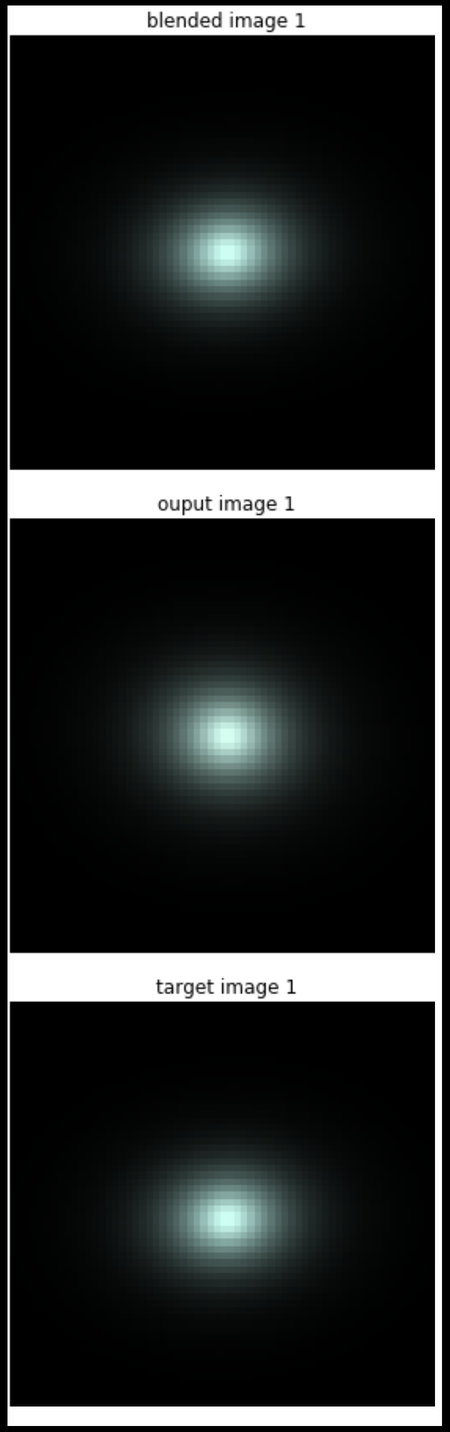
Figure of merit: shape reproduction



Deblender for N galaxies

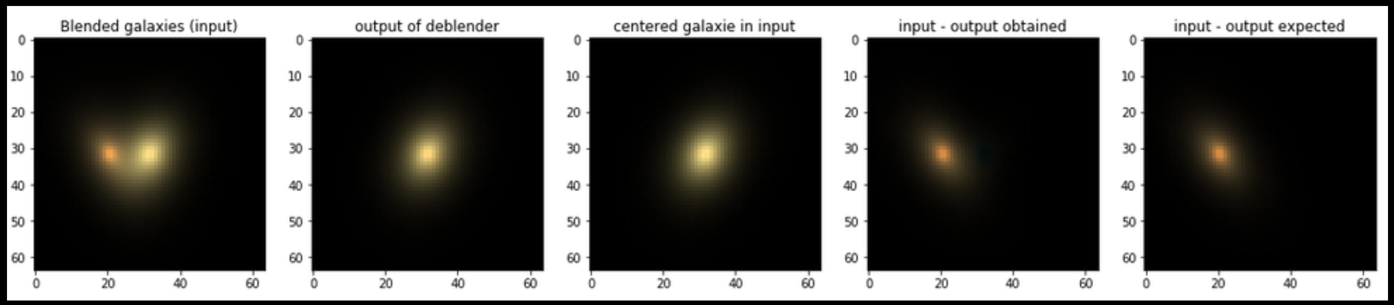
(N = 1, 2, 3 ou 4)





Conclusion

- VAE :
  - Can be used as galaxy generator : (around 60 time) faster than with classical tools.
  - Can be used on real data



- Deblender :
  - Reproduction of the shape and of the luminosity.

Conclusion

- **Using several passbands —> improvement of the precision**

- **Using several instruments (LSST+Euclid - different PSF) —> improvement of the precision for the shape reproduction**

Wait for the paper !

Arcelin et al. (2019). Deblending galaxies with variational Autoencoder: a multi-bands, multi-instruments analysis. *in prep*.

What's next ?

- Improve networks architecture
  - 4 convolutional layers for now
  - Iterations by substracting the output of the deblender to the blended image
- Implement the detection part and the patch construction part
- Ask p(z) and p(shear) and not the reconstruction of the centered galaxy