DEBLENDING GALAXIES WITH GENERATIVE ADVERSARIAL NETWORKS

SHOUBANEH HEMMATI (CALTECH/ IPAC)

+ ERIC HUFF, HOOSHANG NAYYERI, AGN`ES FERT´E, PETER MELCHIOR, BAHRAM MOBASHER, JASON RHODES, ABTIN SHAHIDI, HARRY TEPLITZ

MOTIVATION

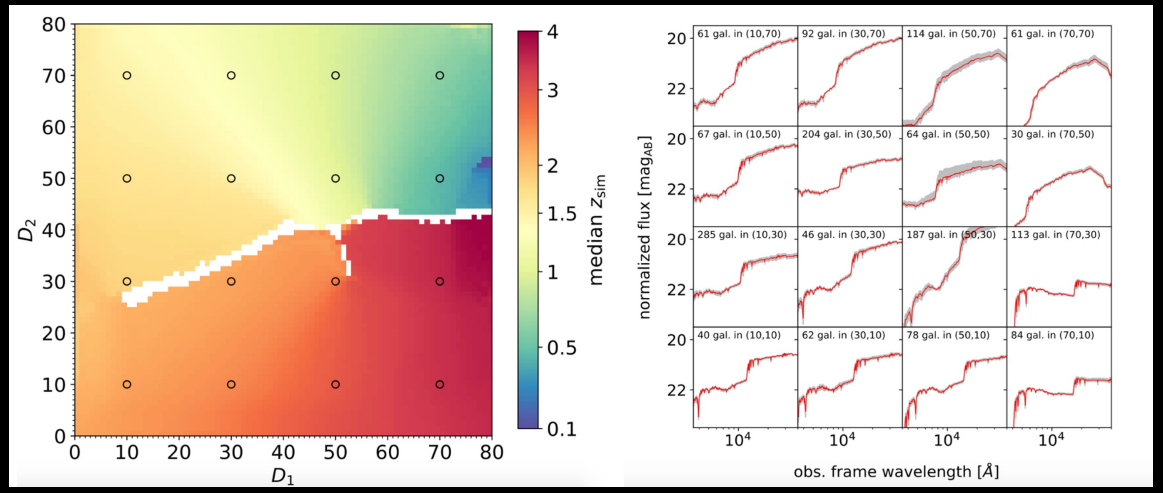
Galaxy evolution

- KPC-scale properties of galaxies,
- Quenching, clumps, environment
- z~0 CII Luminosity functions
- high z massive and evolved
- … we know a lot about galaxies.

Cosmology?

- many systematics come from galaxies, redshifts, blends, shapes, …
- Do we know a lot about galaxies?

GALAXY DISTRIBUTIONS



Davidzon et al. 2019

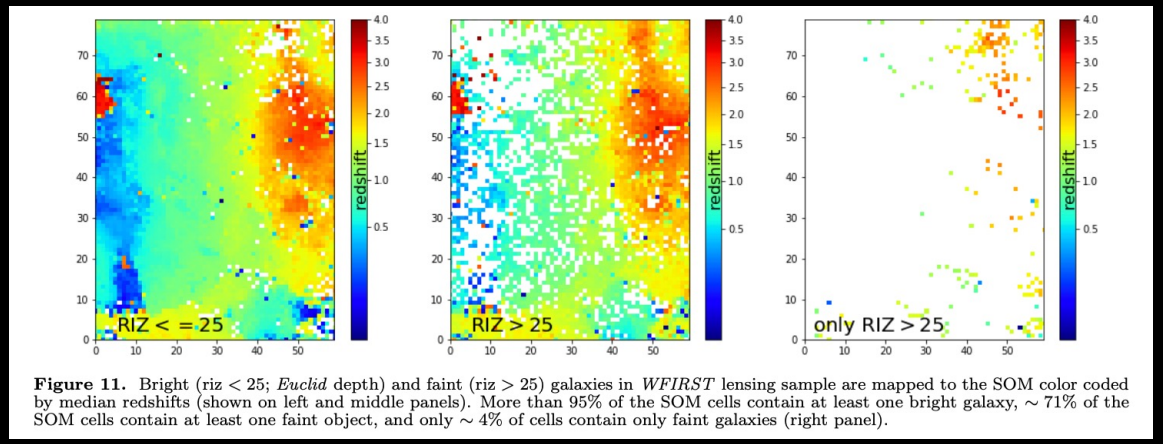
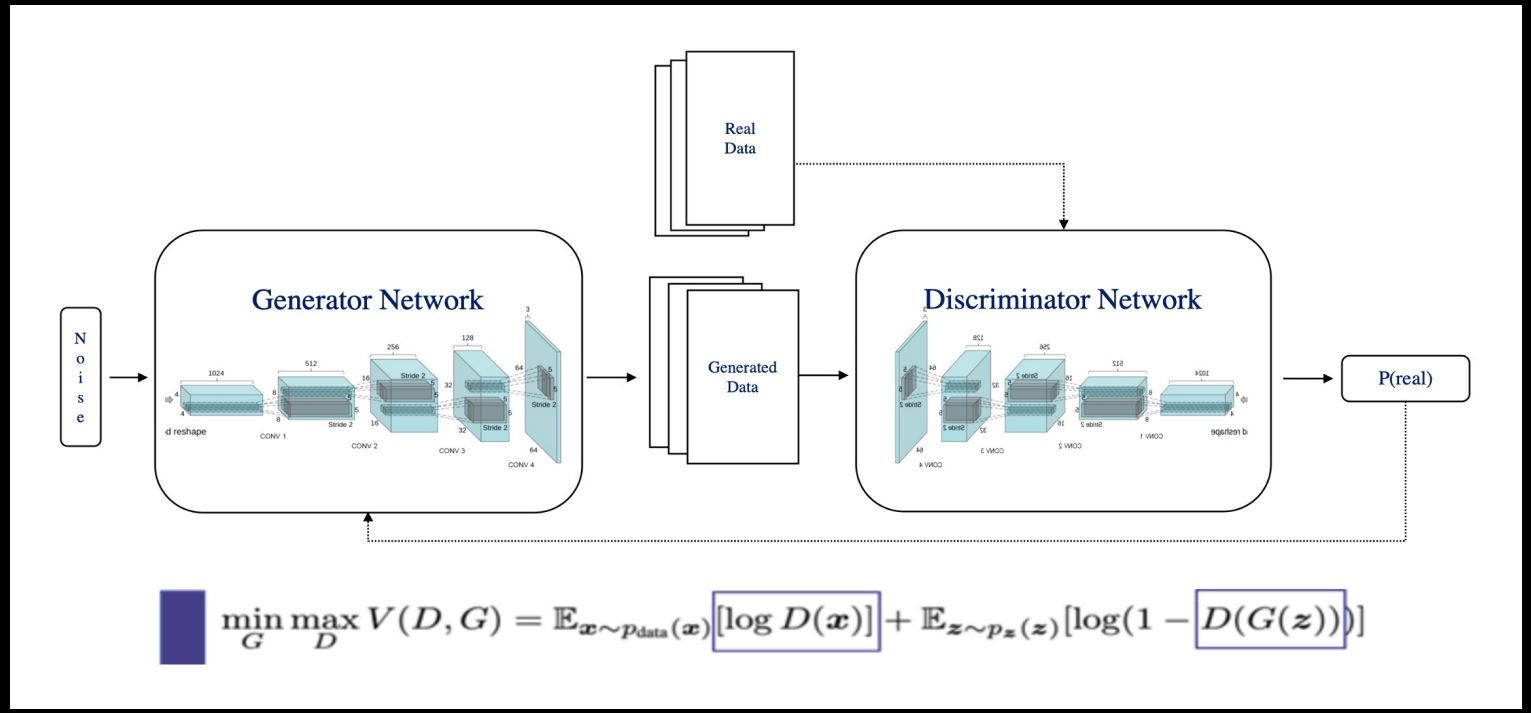


Figure 11. Bright (riz < 25; *Euclid* depth) and faint (riz > 25) galaxies in *WFIRST* lensing sample are mapped to the SOM color coded by median redshifts (shown on left and middle panels). More than 95% of the SOM cells contain at least one bright galaxy, $\sim 71\%$ of the SOM cells contain at least one faint object, and only $\sim 4\%$ of cells contain only faint galaxies (right panel).

Hemmati et al. 2019a

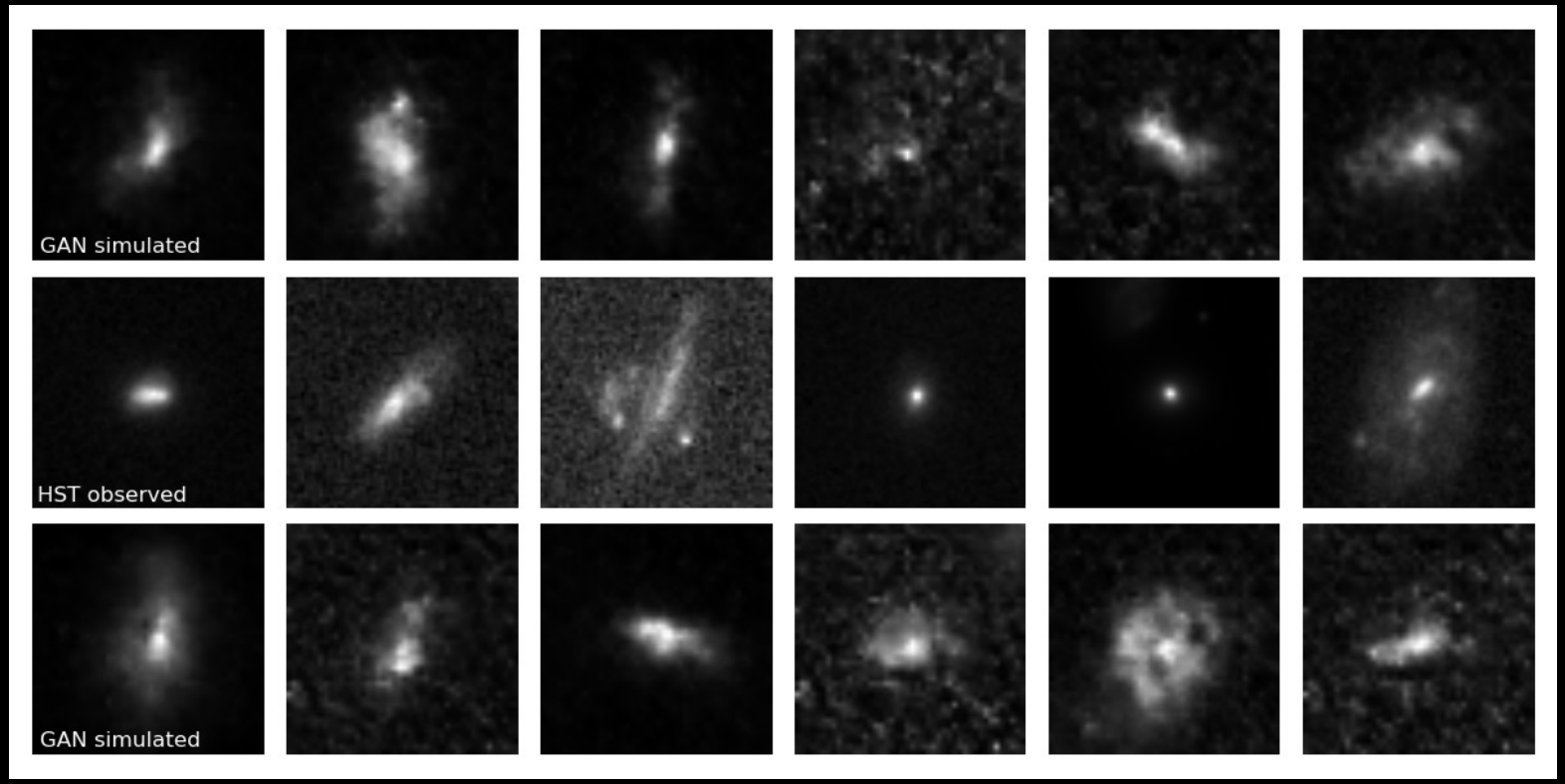
GENERATIVE ADVERSARIAL NETWORKS

GENERATIVE ADVERSARIAL NETWORKS



CANDELS HST CUTOUT SIMULATION

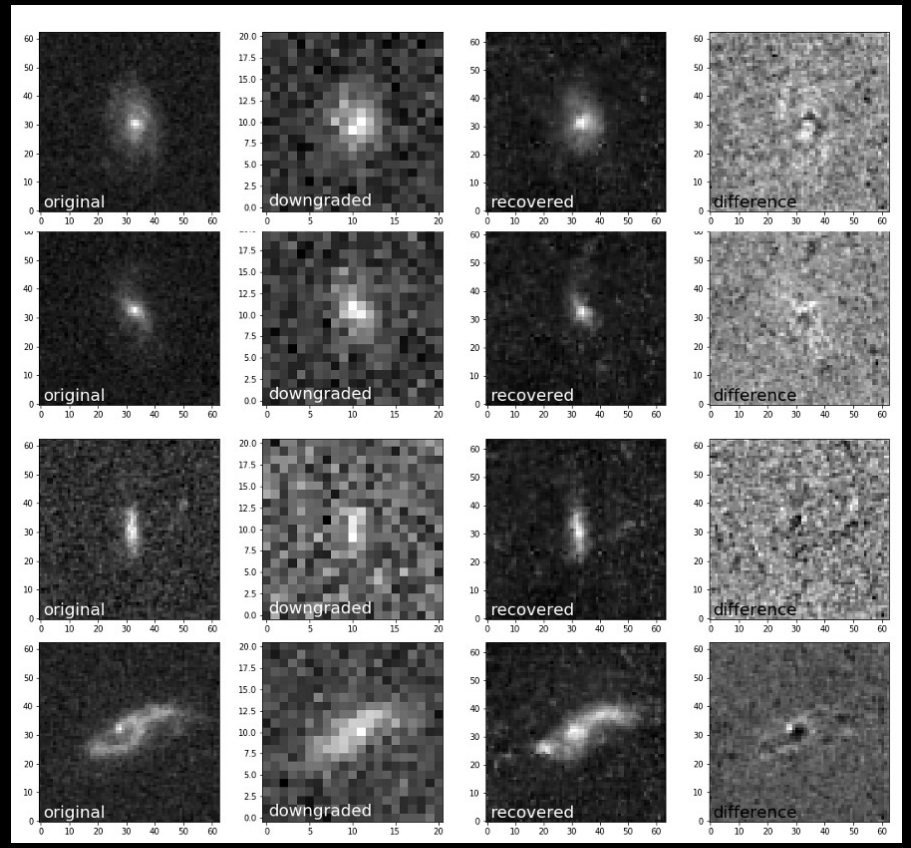
- CANDELS has some of the deepest optical, NIR data HST data
- HST F775W cutouts of z<2 galaxies, normalized, used for training
- Realistic looking HST cutouts simulated with a simple GAN, fast, clumps, …
- Scientifically reliable? Distributions correct?



CONDITIONAL GAN

- Image enhancement
  - super resolution
  - in-filling
  - denoising

pixel scale, PSF, noise
HSC --> HST

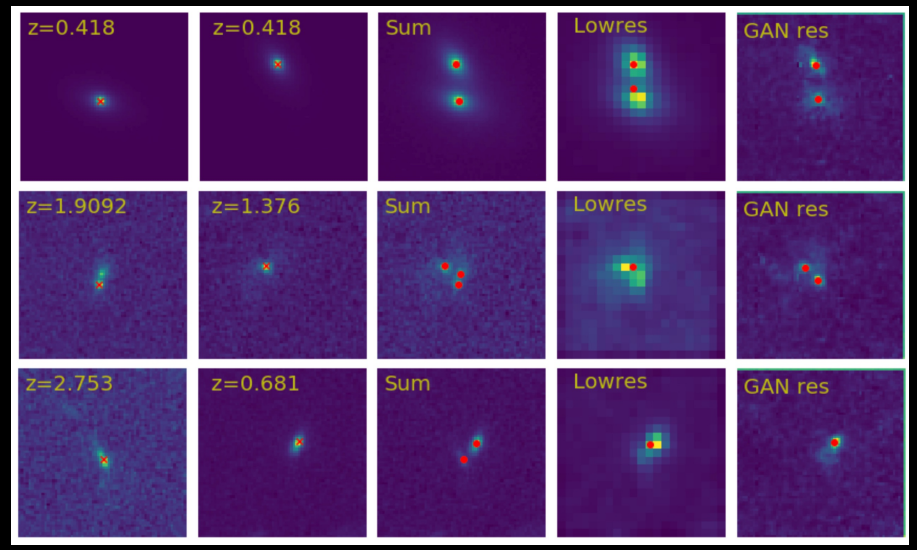


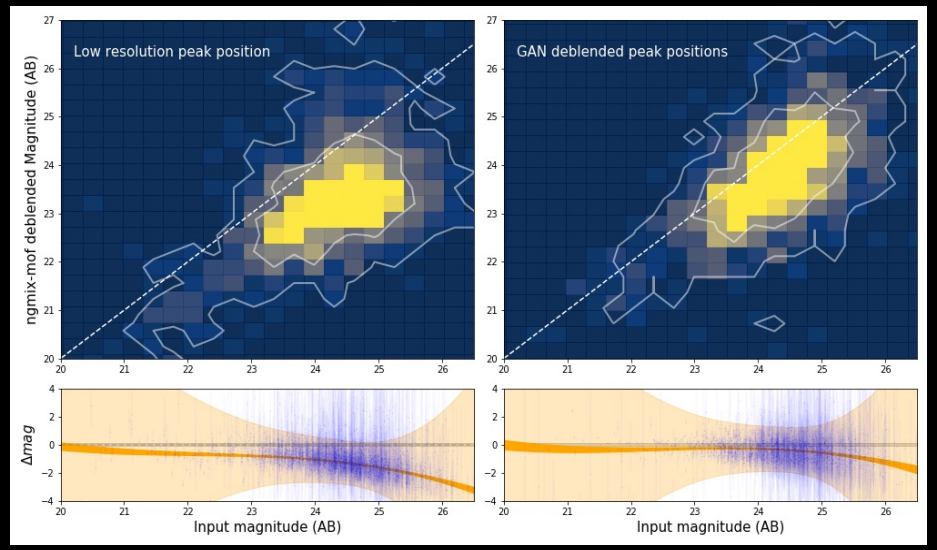
THE ARCHITECTURE I USED

- On pytorch
- Three NVIDIA GPUs (each >3000 cuda cores)
- ~4 hours to train with couple thousand galaxy images
- Not that deep, 4-5 convolution layers in both generator and discriminator
- Largest challenge: hyper parameter tunning (padding, stride, kernel size, number of layers, …)
- Activation ReLU mostly, TanH/sigmoid last layer, which adds non-linearity to the network
- Loss function: minimax cross entropy enough for this work
- ADAM optimizer (efficient stochastic gradient decent with low memory requirement)
- Codes on my github (but super unorganized) https://github.com/xoubish/disks

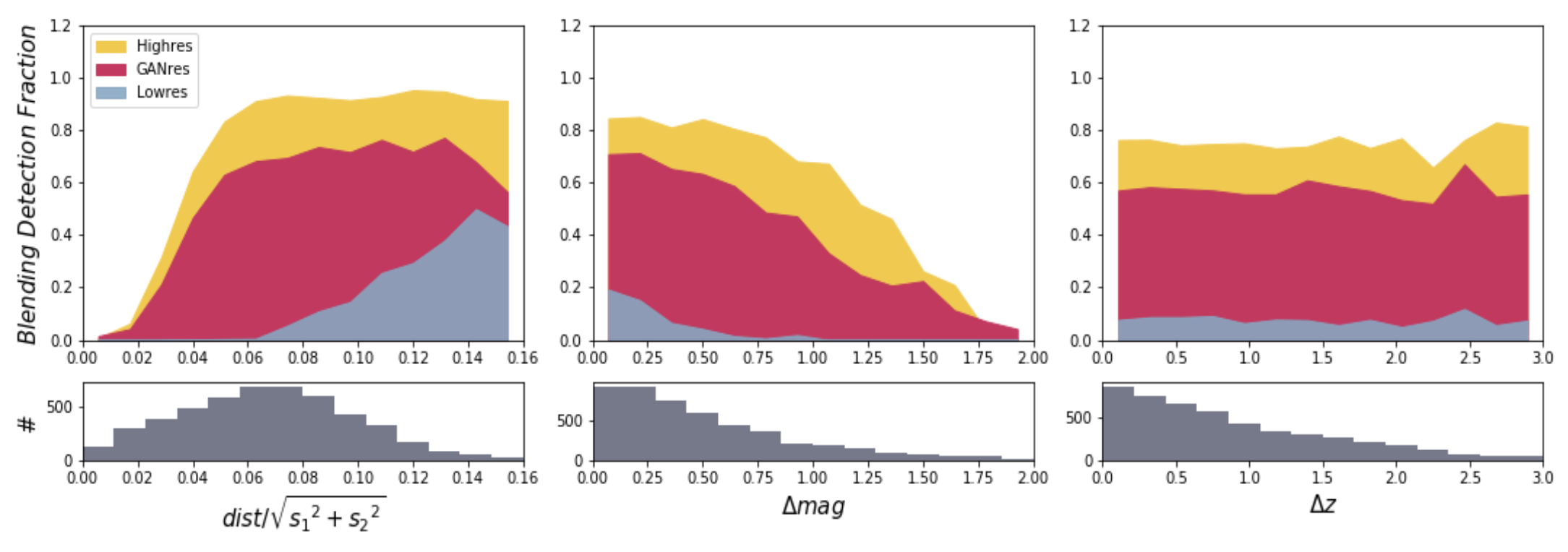
SCIENCE EXAMPLE: DEBLENDING

- A large fraction of galaxies (~ 58%) are blended in HSC or LSST depths (Bosch et al. 2018)
- simulated a blended sample with HST, various distances, orientations, brightness, redshifts, …
- For detection used skimage Laplacian detection and code by E. Sheldon, similar result
- Photometry with NGMIX: implements Gaussian mixture models for 2D images. Both the PSF profile and the galaxy are modeled using mixtures of Gaussians. Convolutions are thus performed analytically, resulting in fast model generation as compared to methods that perform the convolution in Fourier space. For the galaxy model, NGMIX supports exponential disks and de Vaucouleurs and Sérsic profiles; these are implemented approximately as a sum of Gaussians using the fits from Hogg & Lang (2013)

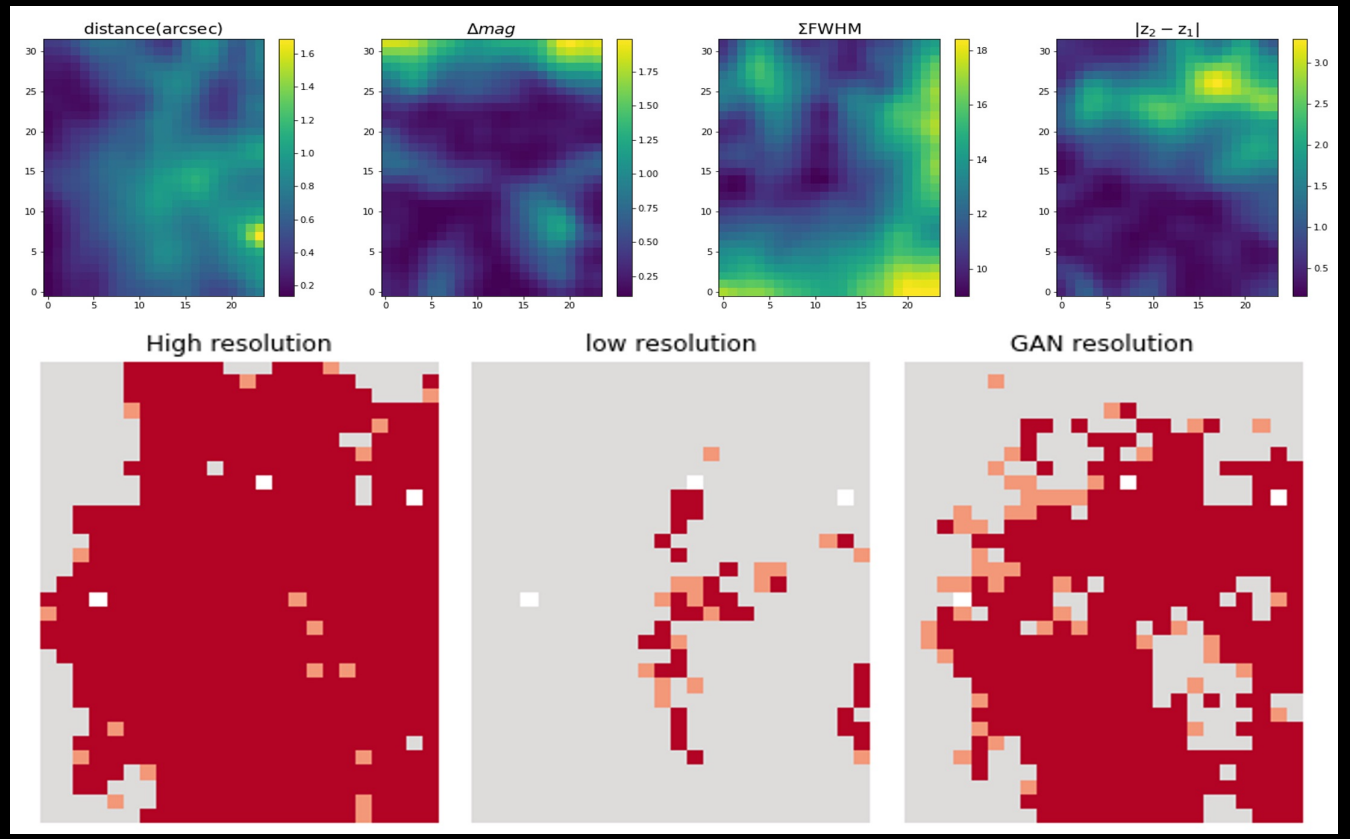




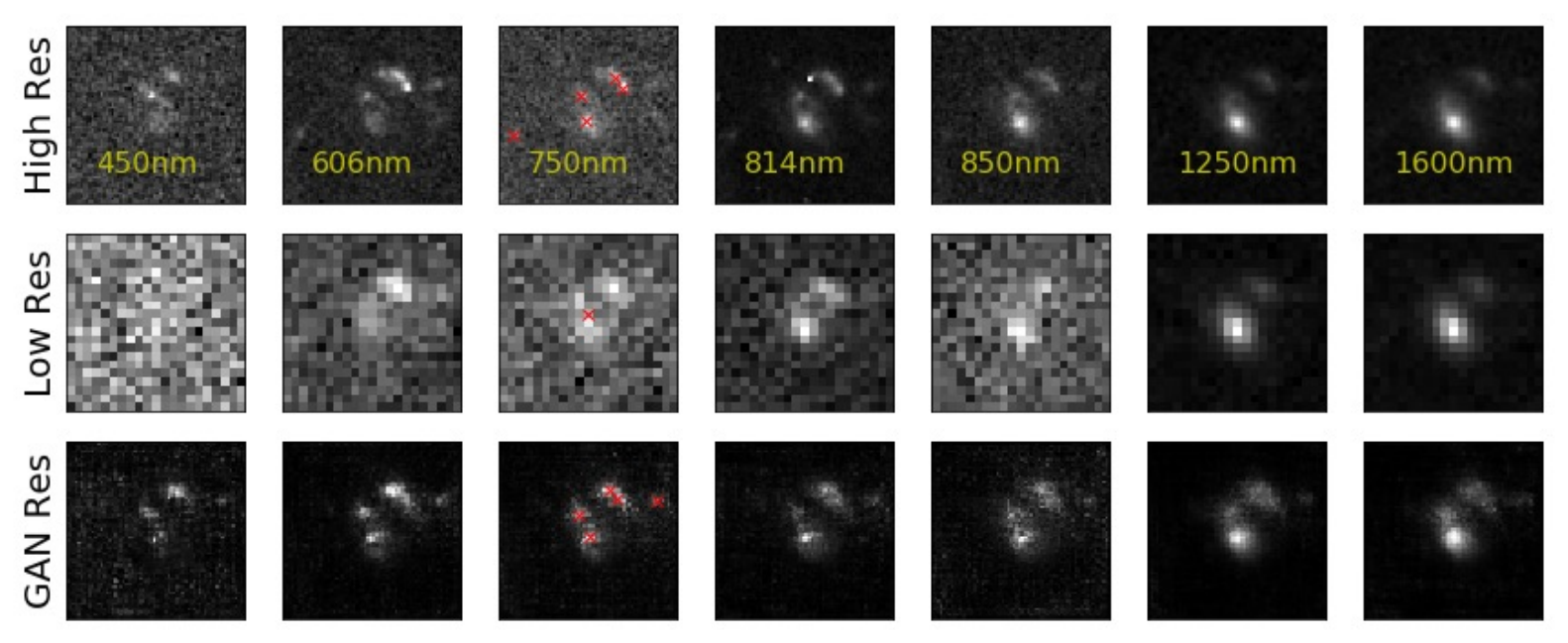
WHERE DOES IT HELPS MORE?

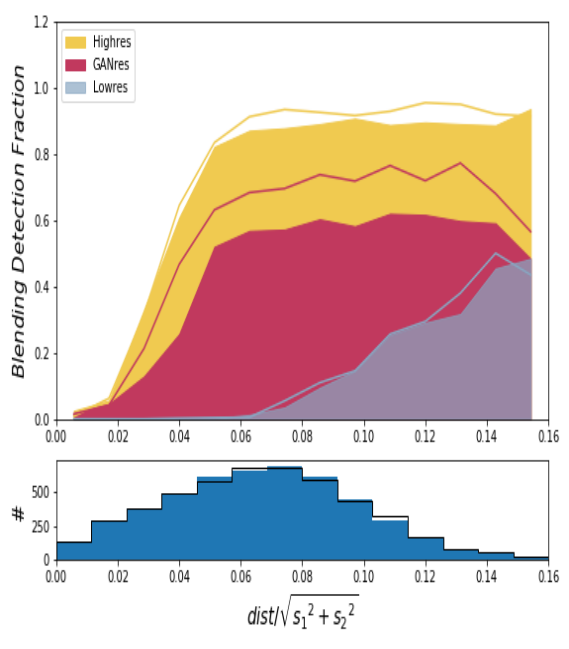


WHERE DOES IT HELPS MORE?



MULTI-WAVEBAND GAN





DISCUSSION

- Paper submitted six months ago …
- Do we know the galaxy distribution well enough to draw and generate from them?
  - How are new observations changing the known distribution
    - Extensions in wavelength
    - Resolution, spatial and spectral
- How to best add uncertainty to enhanced data products? (reason I signed up for this conference)
- The best networks to enhance data? Very eager to try diffusion models now …