



ML accelerators for Nifty

David Smolinski

Table of content

- Motivation
- Radio interferometry
- Modeling the UV coverage
- Sky model
- Architectures
- Next steps

Motivation

- Nifty and resolve implement very fast forward models
- Sampling mock data from priors is fast
- The backpropagation, i.e. the inference is computationally costly
- Use Machine Learning
- Con: We have training phases
- Pro: after training the evaluation is cheap
- → Use forward model to produce pretrained networks

Radio interferometry

- Motivation: Angular resolution $\Theta = \frac{\lambda}{D}$
- Interferometry with long baselines increases resolution
- No free lunch
- Low data coverages (sparse UV plane)
- Images need postprocessing
- Challenge in the upcoming years: High data rates



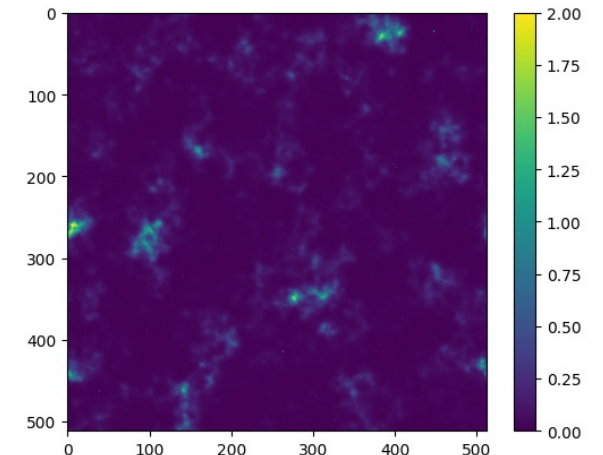
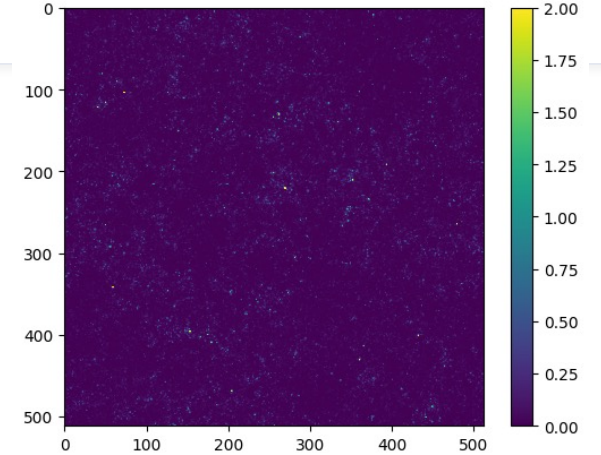
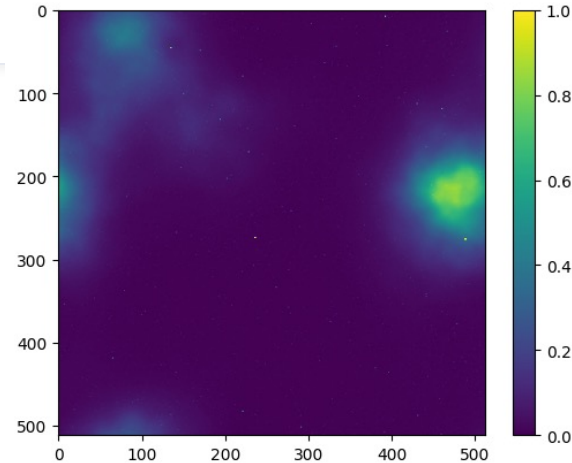
Picture: SARAO

Modeling the UV coverage

- Calculating UV coordinates for a given observation
- Pyvisgen (Kevin Schmidt, Felix Geyer @ TU Dortmund)
- RIME formalism to calculate visibilities
- Using the UV coverages of pyvisgen
- Nifty + Resolve to simulate a mock sky

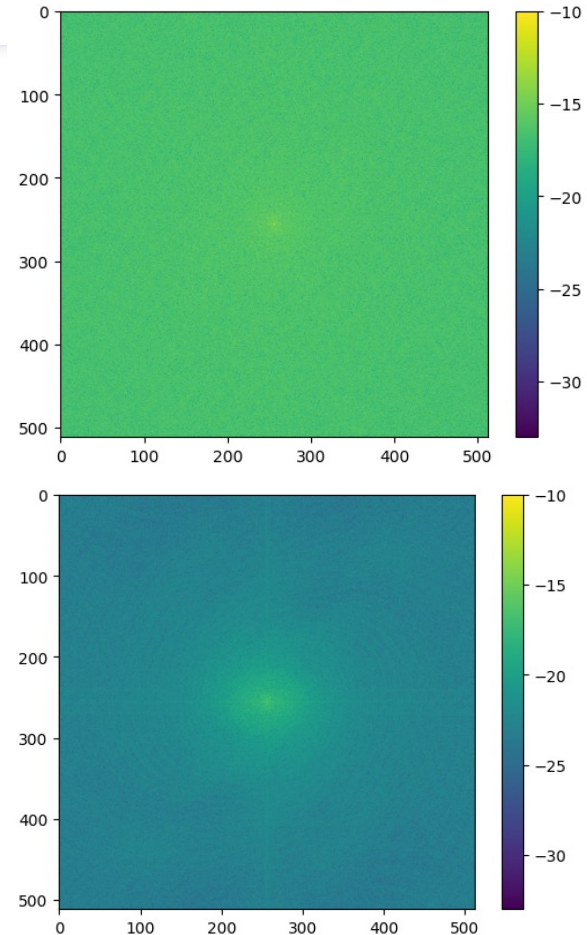
Sky model

- Diffuse emissions = correlated field model
- Point sources = Inverse gamma operator
- Apply to random gaussian field



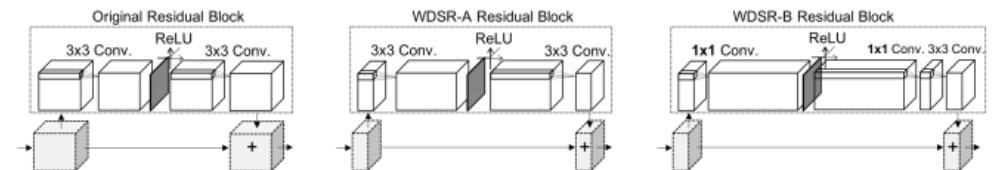
Sky model

- Full UV coverage \rightarrow Fourier transform
- Sparse UV coverage \rightarrow
 - Use UV pointings
 - Create observation in resolve
 - Apply radio-response to sky model
 - Fourier transformation



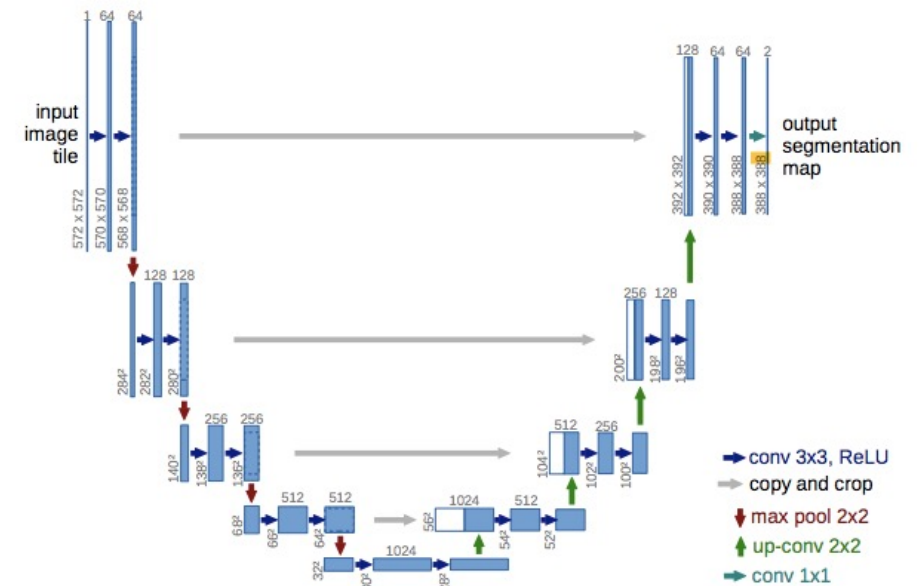
Architectures

- WDSR (= Wide Activation for Efficient and Accurate Image Super-Resolution)
- Head: Convolution-layer
- Body: Combination of resblocks
- Resblock: B three convolutional layers, skip connection, weight normalization
- Tail: Convolution and Pixel shuffle



Architectures

- UNet
- Encoder blocks: convolution blocks + Batch norms+ MaxPooling
- Decoder: convolution transpose blocks + Batch norms



Challenges

- In the moment: Overfitting
- Trainings losses are good
- Validation loss is bad
- → The model is not able to apply to new, unseen data

- Batch normalization?

Next steps

- Get rid of overfitting
- Comparison of UNet, WDSR, radionets, resolve, CLEAN
- Generalize to different Nifty/ resolve models
- Use structure of Nifty to design models
- Pretrained models?
- Graph neural networks?



Questions and Discussion