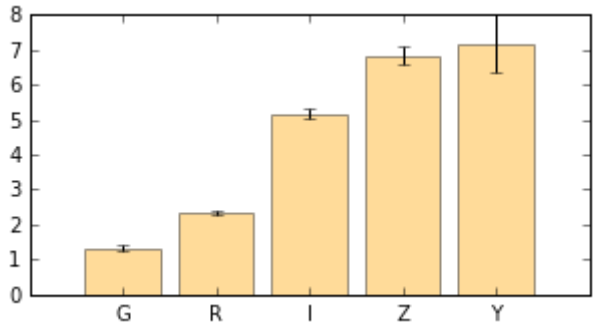
Kanazawa Institute of Technology

Recent Results from Deep Learning

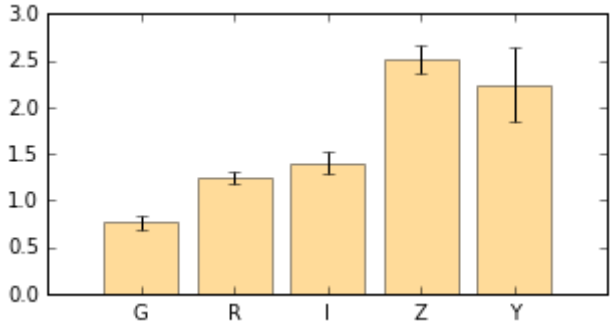
May 18, 2017

Jun Nakano
Kanazawa Institute of Technology
http://nakanolab.net/

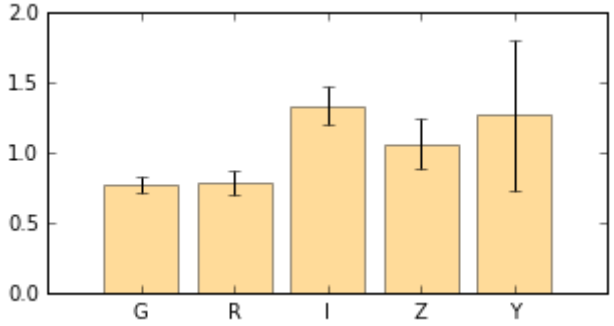
Predicting z from 5-Band **Flux** of Subaru HSC



z =
→ 0.73



→ 1.81

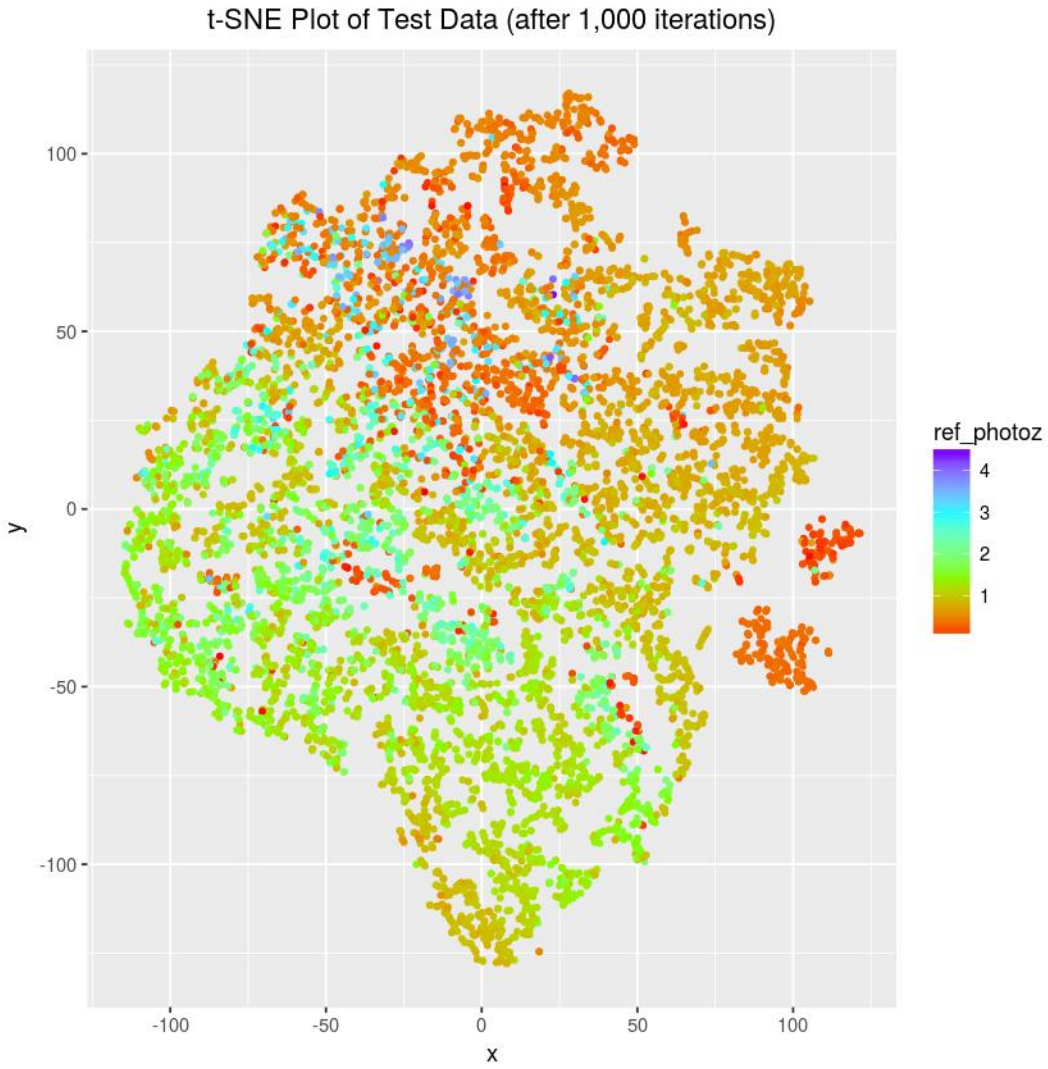


→ 2.93

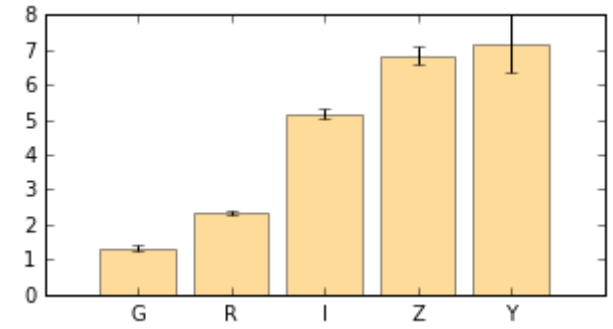
Degeneracy makes prediction hard!

K.I.T.

- t-SNE (t-Distributed Stochastic Neighbor Embedding) plot of (G, R, I, Z, Y) data points

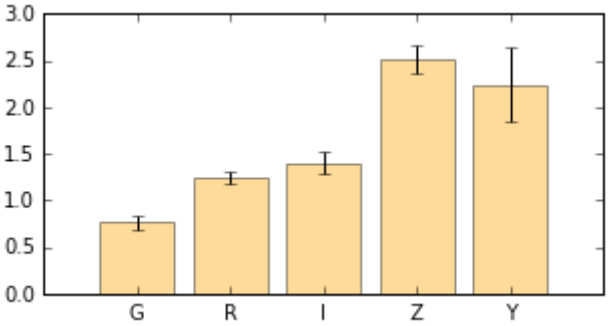


Predicting z from 5-Band **Flux** of Subaru HSC



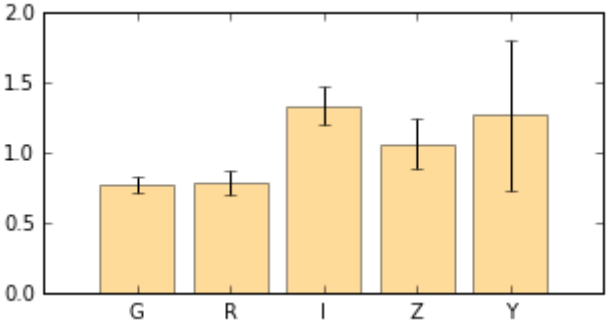
+ other features
and/or physics

z =
→ 0.73



+ other features
and/or physics

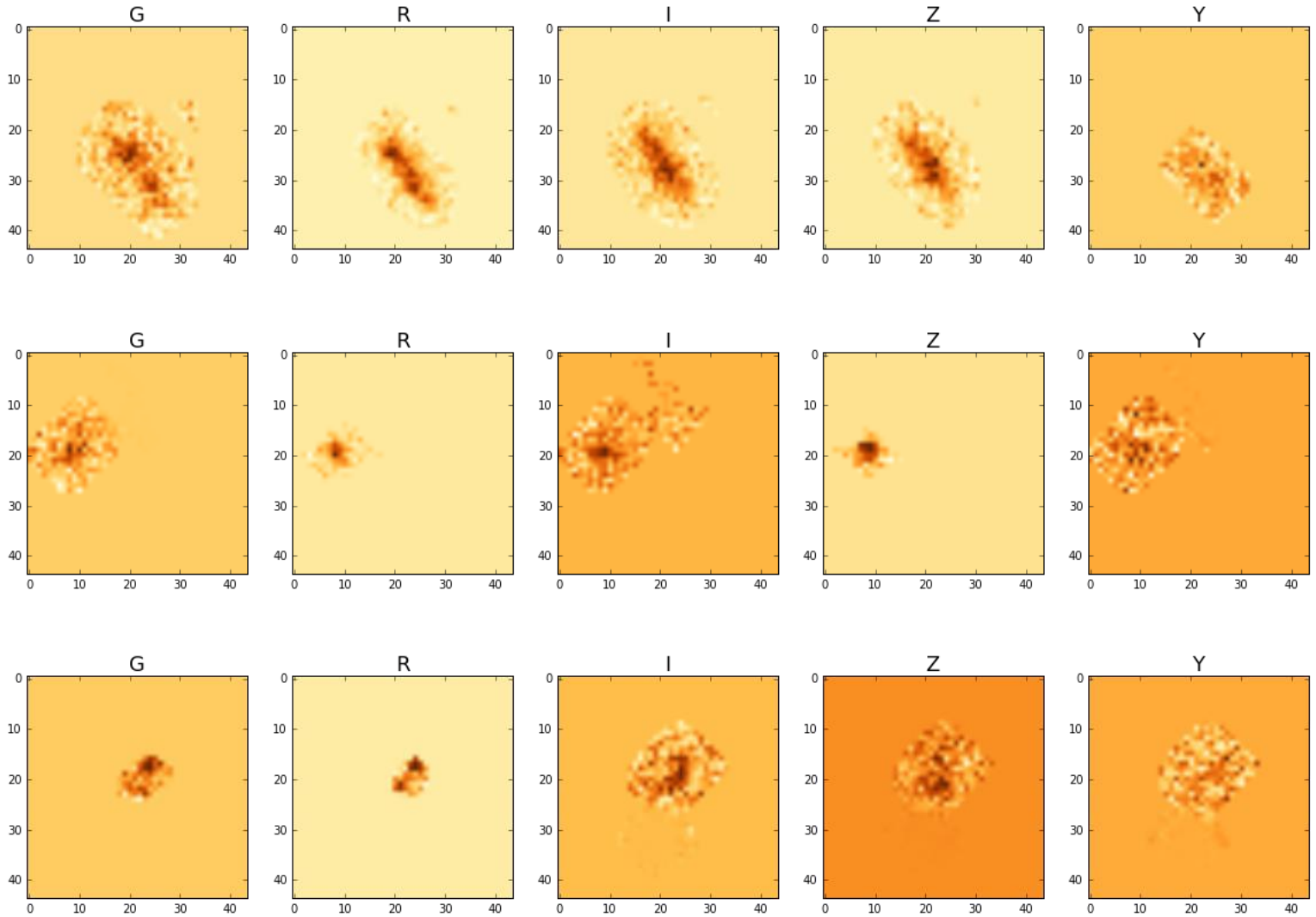
→ 1.81



+ other features
and/or physics

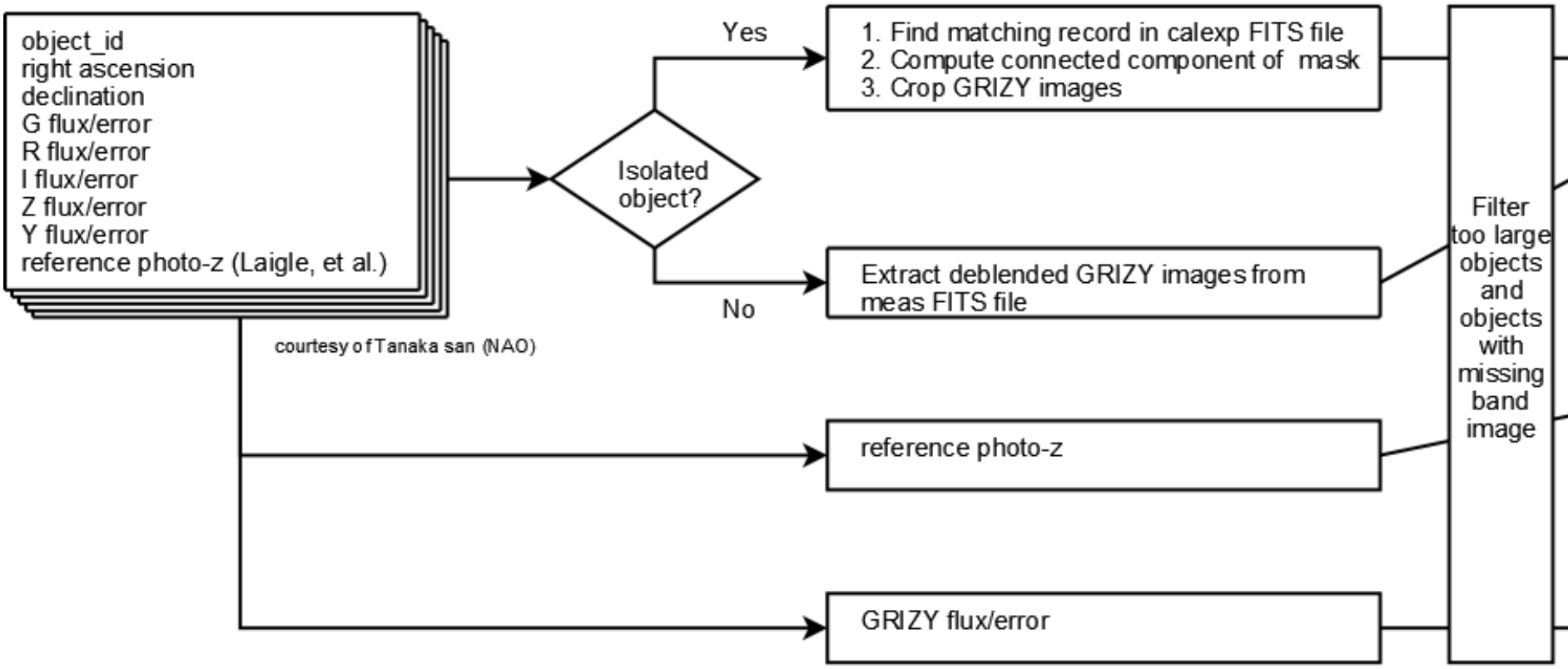
→ 2.93

Predicting z from 5-Band Images of HSC



Data processing pipeline (1)

K.I.T.

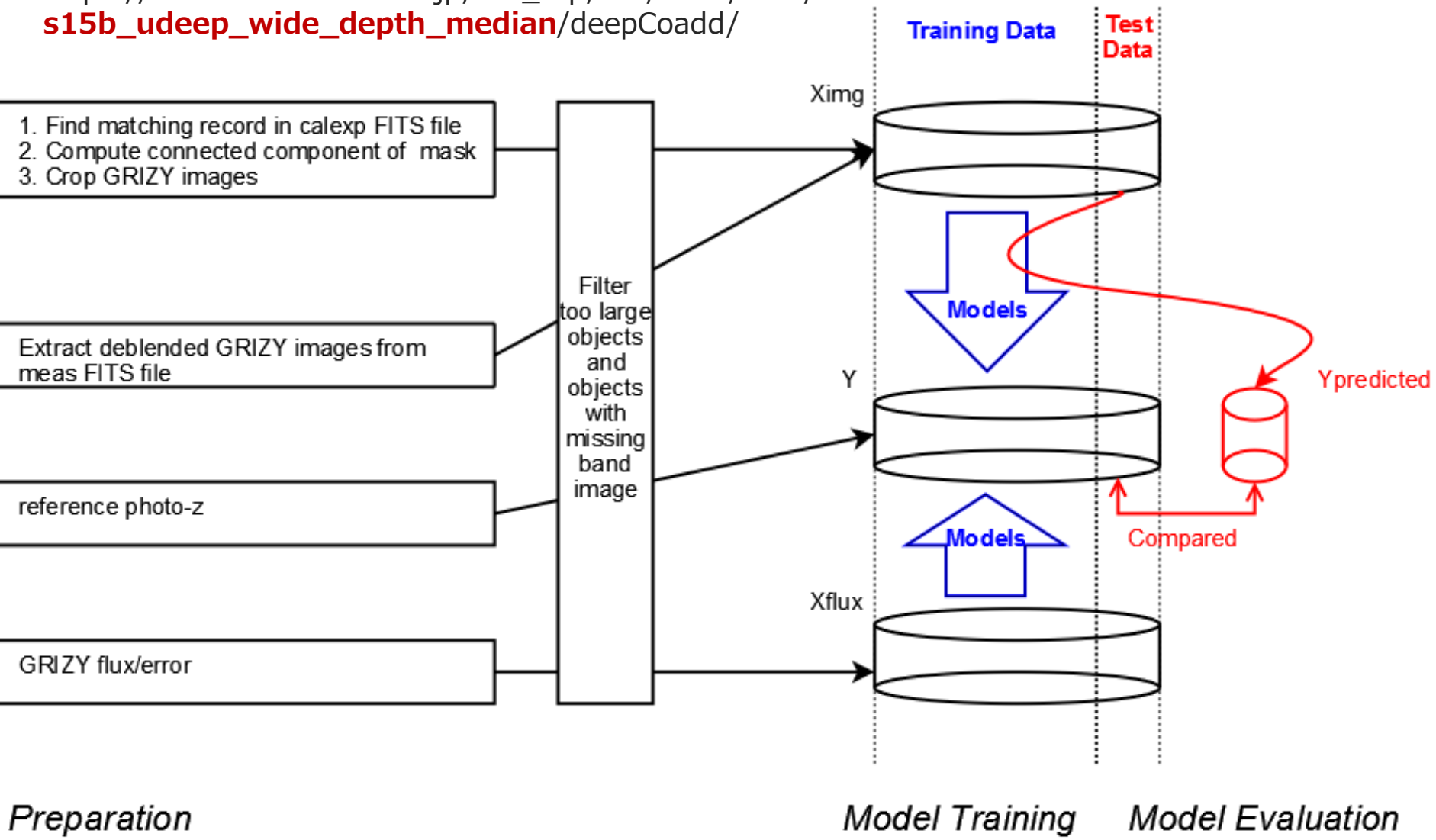


Data Preparation

Data processing pipeline (2)

K.I.T.

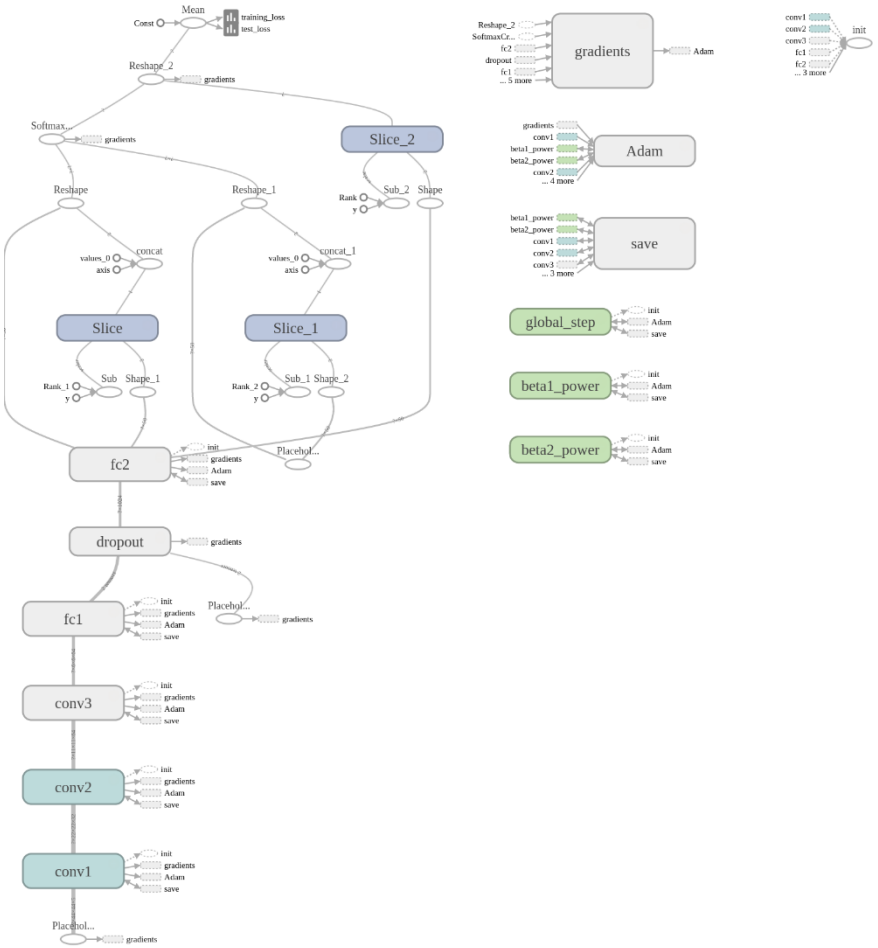
https://hscdata.mtk.nao.ac.jp/hsc_ssp/dr1/s15b/data/
s15b_udeep_wide_depth_median/deepCoadd/



Neural net architecture

K.I.T.

- Convolutional Neural Network (CNN)

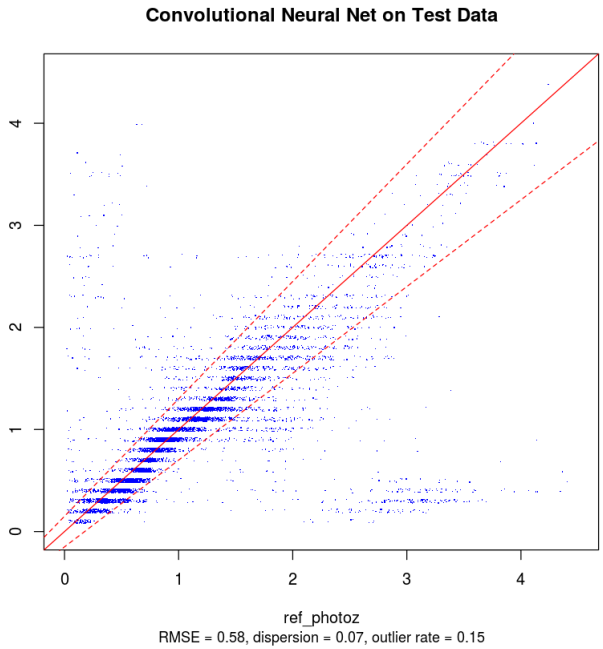


Random forest vs. CNN

K.I.T.

Aggregated Flux

Raw Image

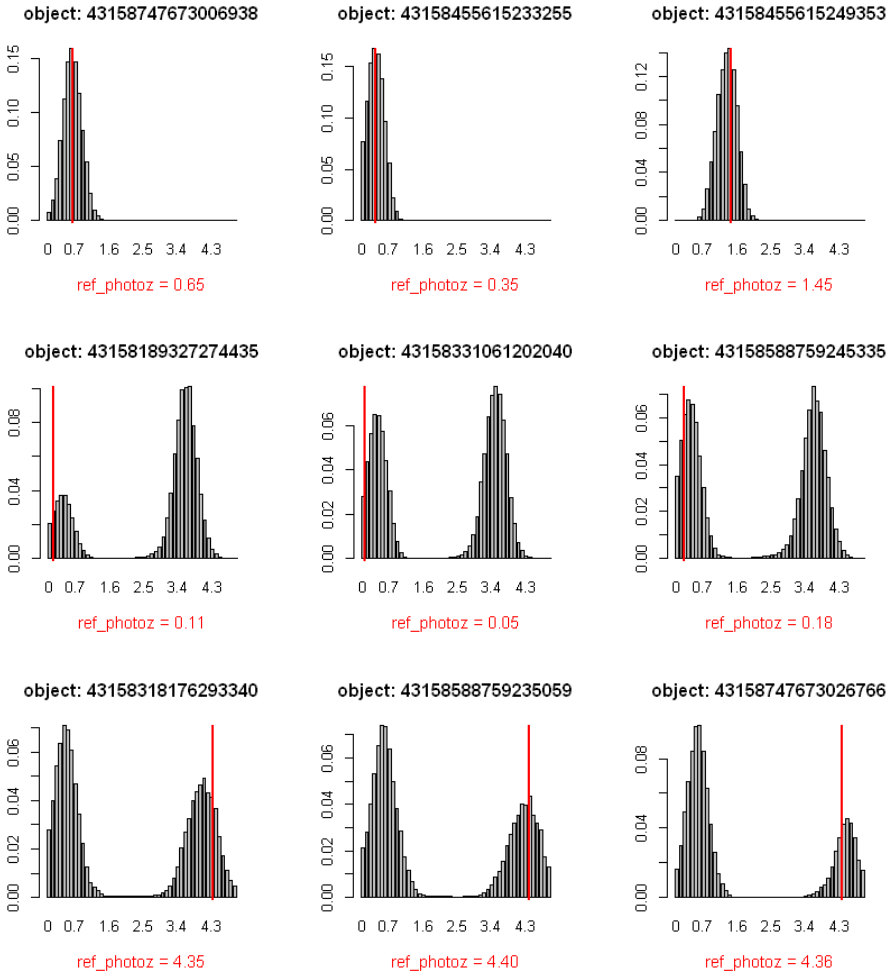


	Random Forest	CNN	
RMSE	0.52	0.58	+12%
Dispersion	0.11	**0.07**	-36%
Outlier Rate	0.29	**0.15**	-48%

Raw outputs from CNN

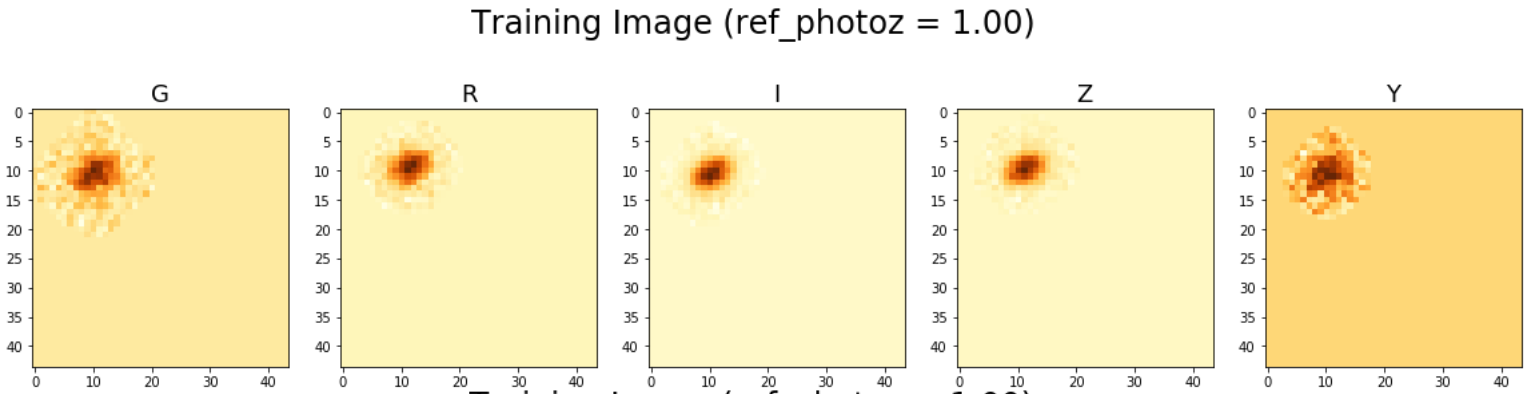
K.I.T.

- Our CNN model outputs "probability distribution" of z
  - Almost right
  - Hugely overestimated
  - Hugely underestimated

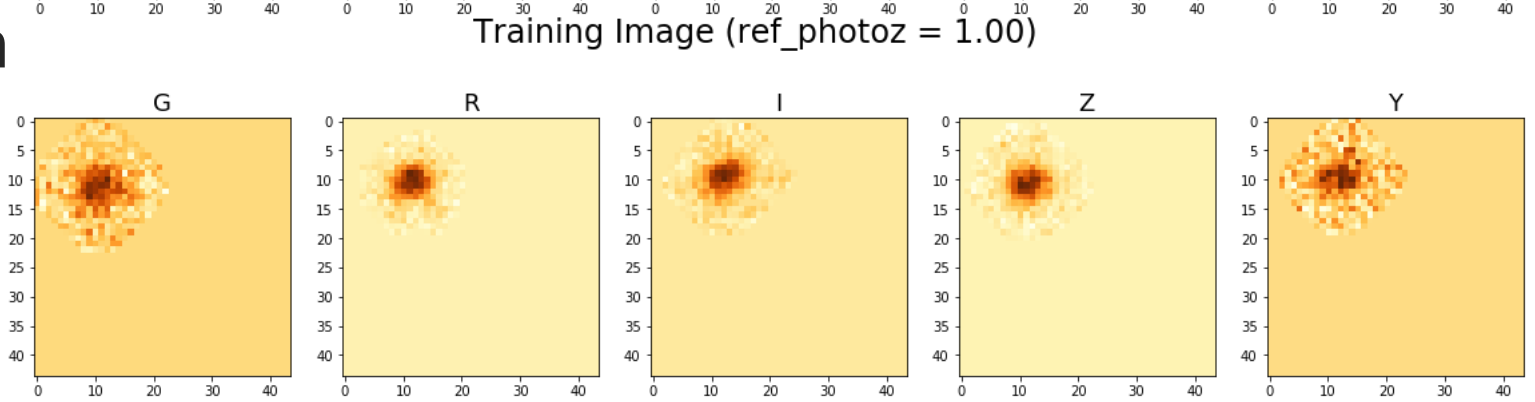


How does it perform on different seeings?

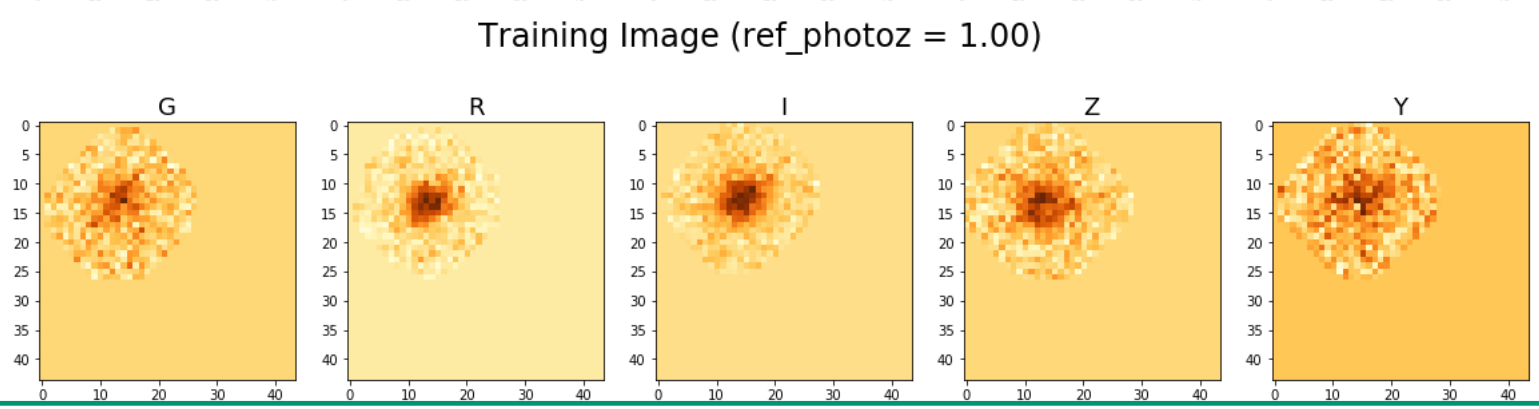
- Best



- Median



- Worst

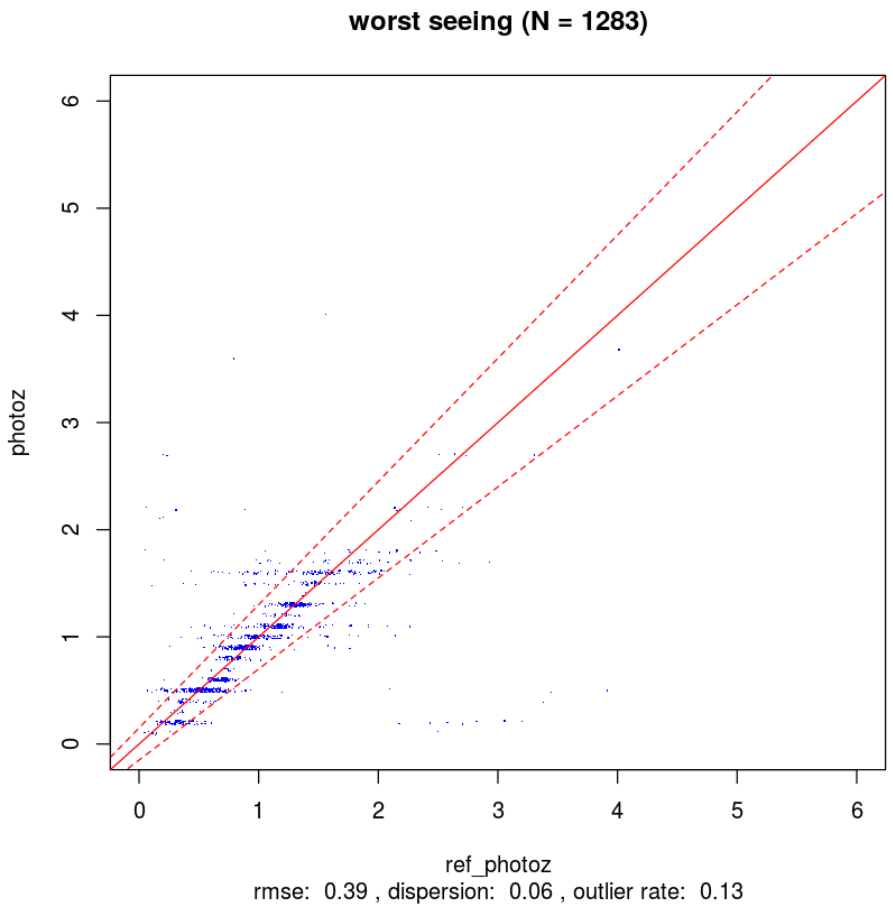


Images from different seeings

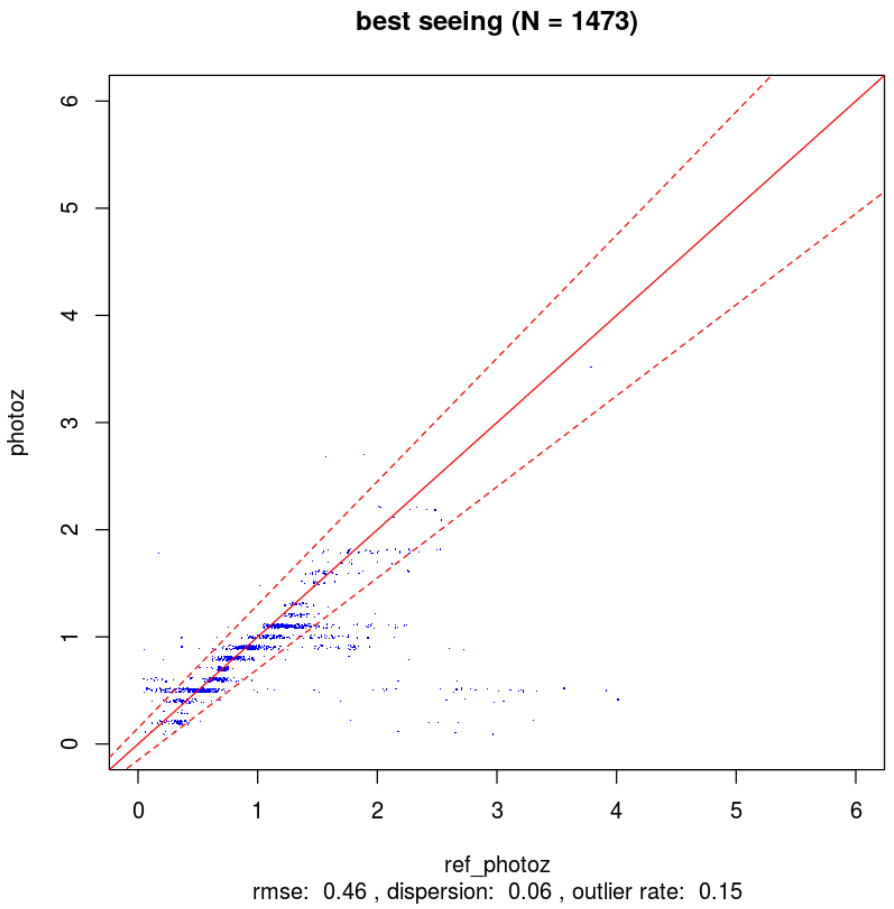
K.I.T.

- Model trained on median seeing

Worst seeing (13% outliers)



Best seeing (15% outliers)



Summary

K.I.T.

- Convolutional Neural Network: Powerful tool for predicting photo-z (especially for reducing outlier rates)
- Raw output of CNN can be considered as a probability distribution of photo-z and may be used for identifying could-be outliers (e.g., double peaks)
- Future Plan
  - More extensive hyper-parameter search in design space (# of layers, size of filters, activation functions, etc.)
  - Stacking with predictors based on physics