KIT | Kanazawa Institute of Technology

Predicting Photo-z using Deep Learning

February 10th, 2018

Jun Nakano

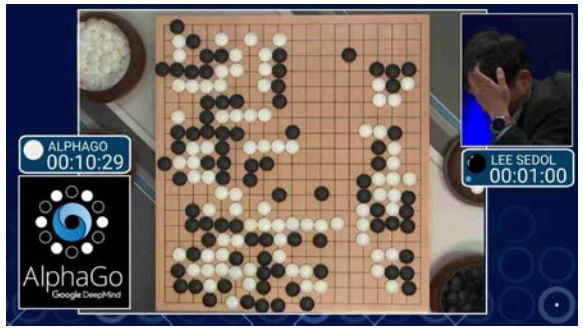
Department of Information and Computer Engineering

Kanazawa Institute of Technology

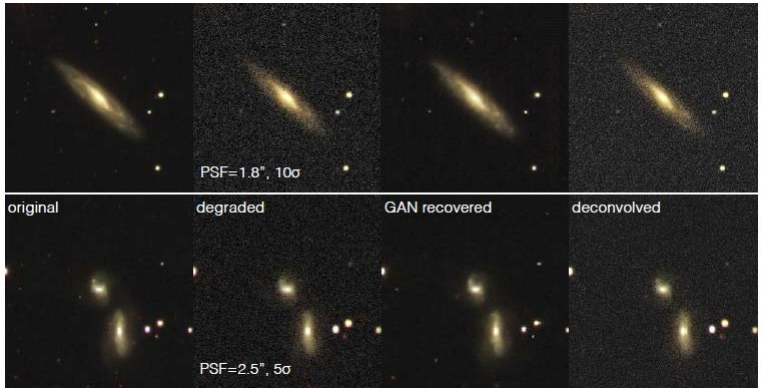
https://www.nakanolab.net/

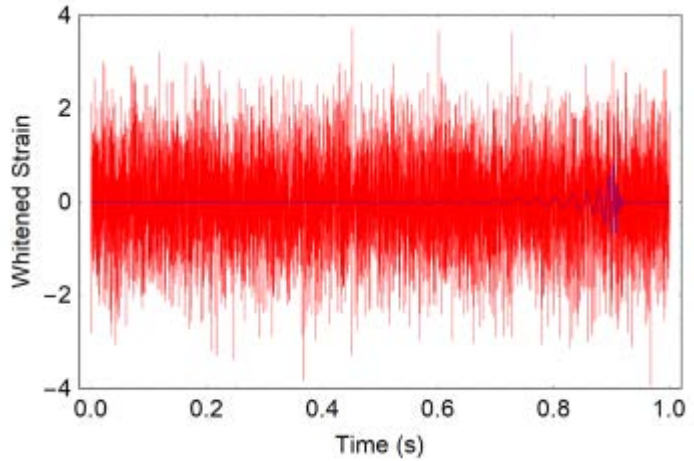
Deep Learning is Everywhere

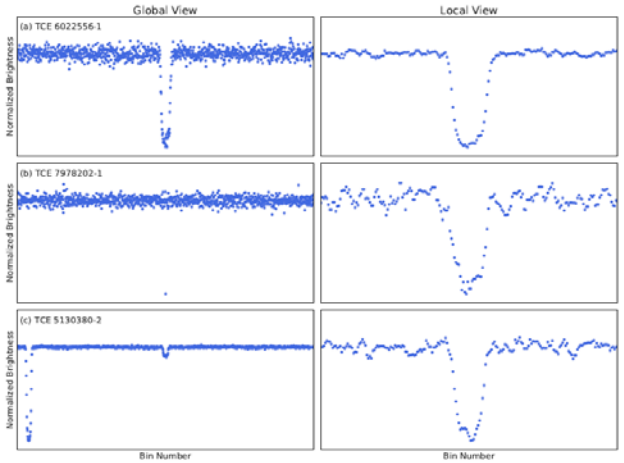
- In image recognitions, language translations, playing intelligent games, etc.



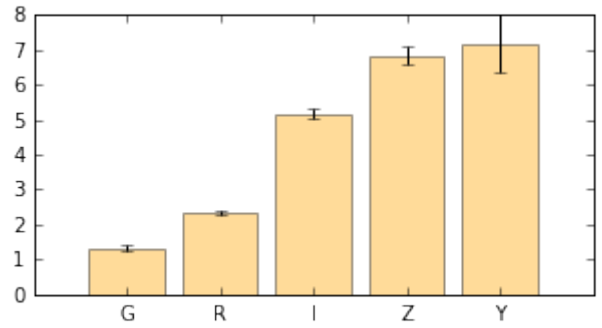
- And in astrophysics, too!



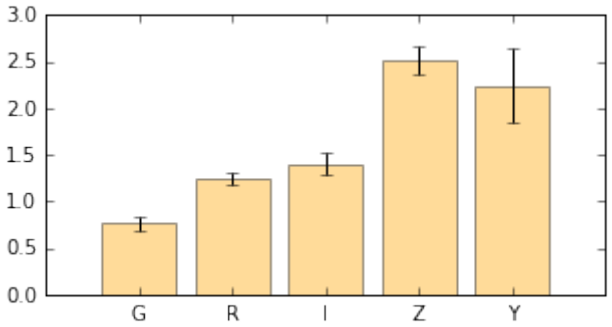




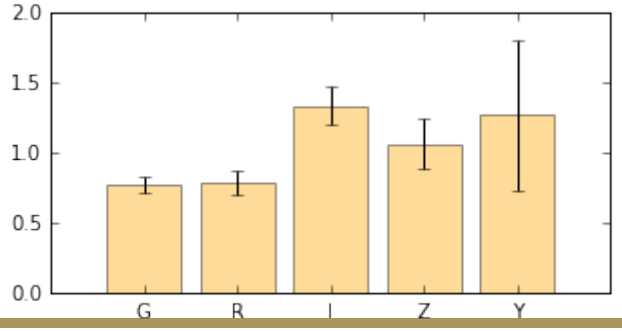
Predicting z from 5-Band Flux of Subaru HSC



z =
→ 0.73



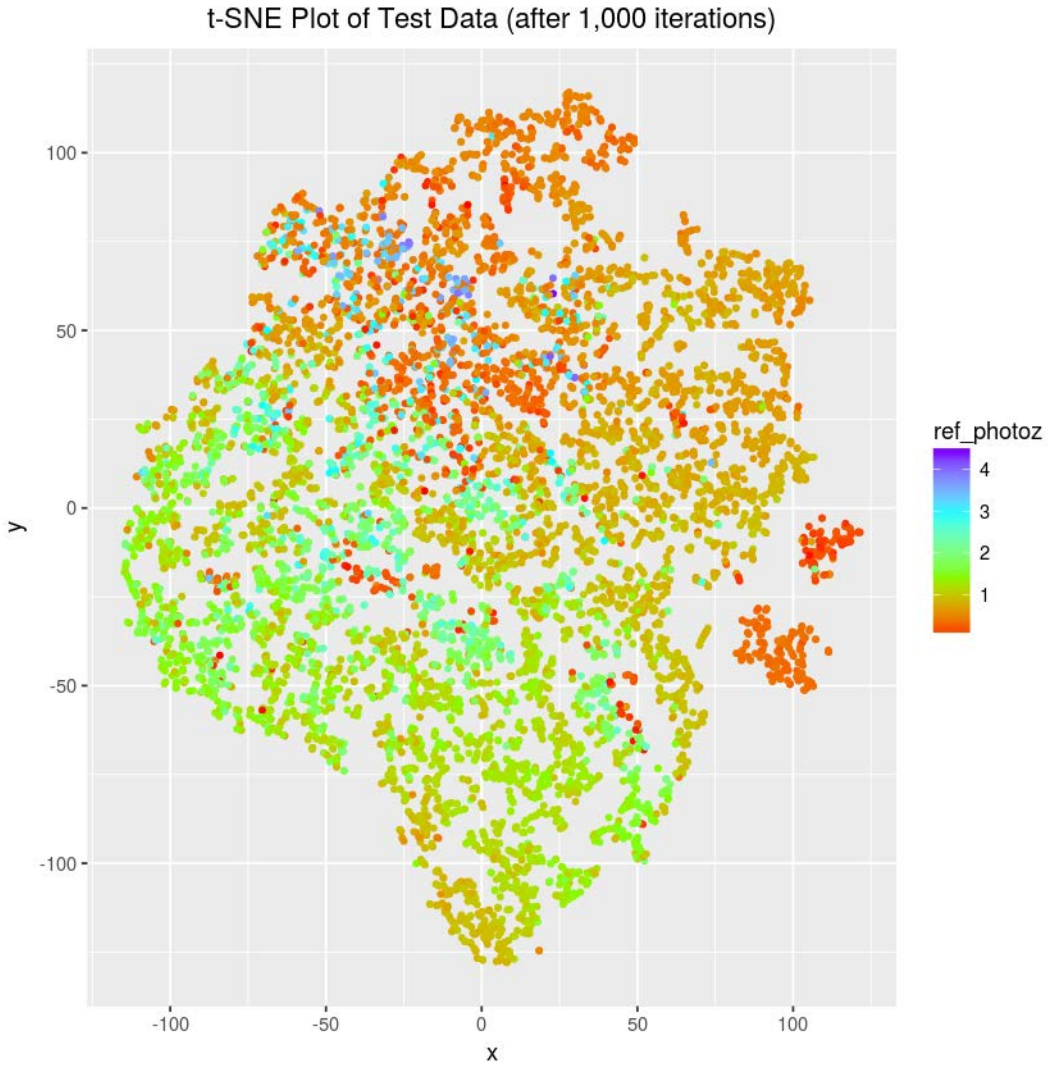
→ 1.81



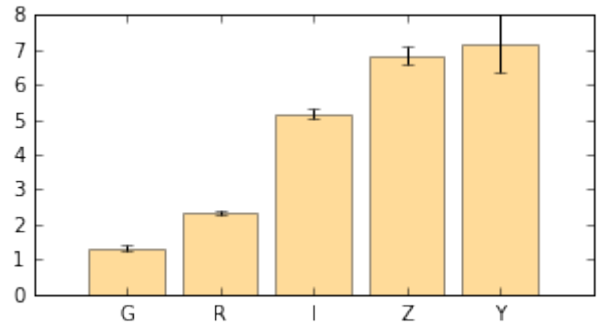
→ 2.93

Degeneracy makes prediction hard!

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

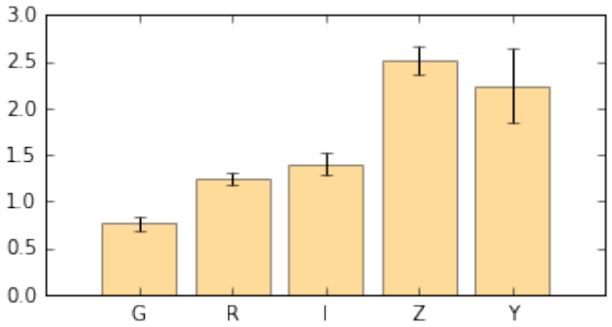


Predicting z from 5-Band Flux of Subaru HSC



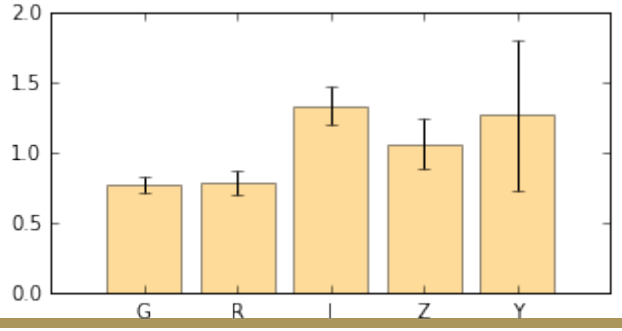
+ other features and/or physics

z = → 0.73



+ other features and/or physics

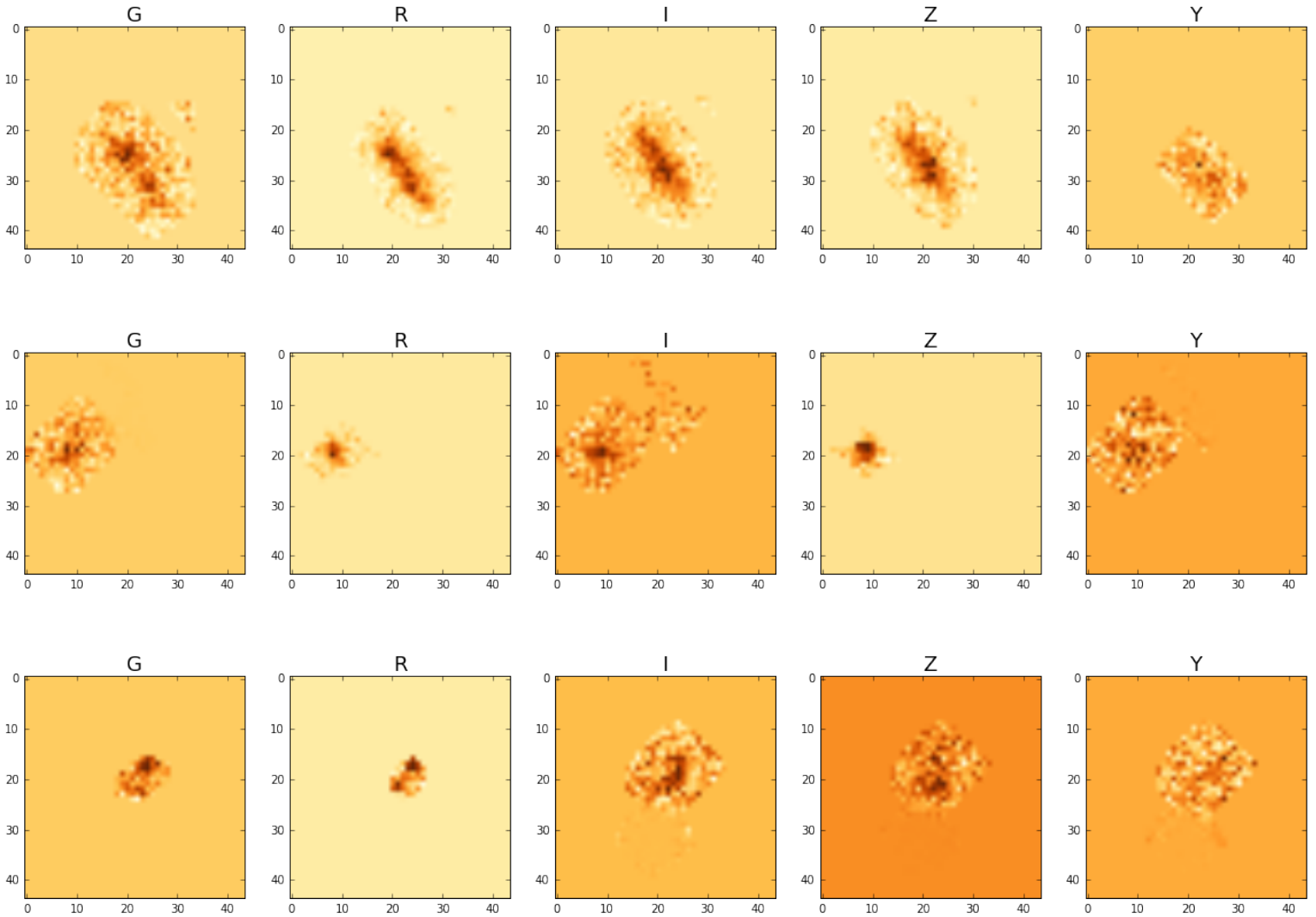
→ 1.81



+ other features and/or physics

→ 2.93

Predicting z from 5-Band Images of HSC



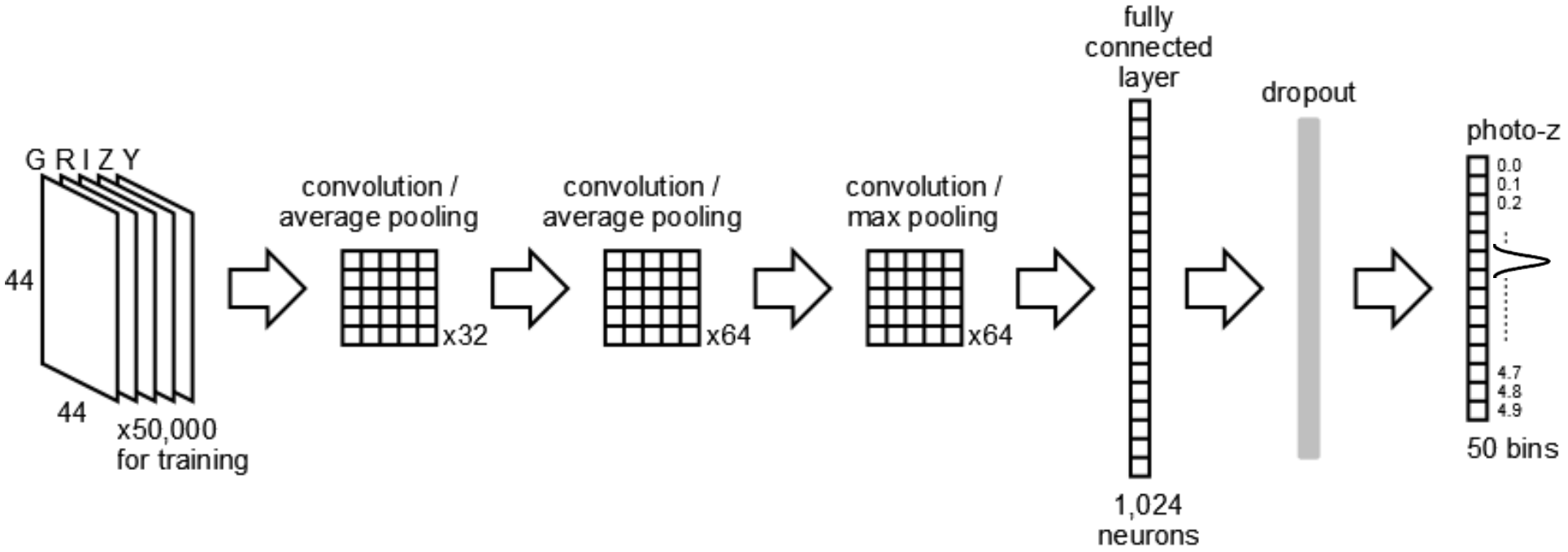
Dataset

- s l 5b_udeep_wide_depth_median with high precision photo-z by Laigle et al.

- 40,000 training data
  - for computing edge weights in NN
- 10,000 validation data
  - for choosing the best hyper-parameter configuration
- 8,629 test data
  - for getting the performance of the chosen model

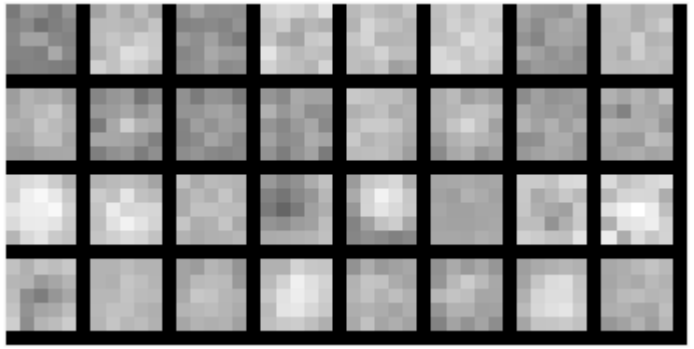
Neural network architectures (I)

- Convolutional Neural Network (CNN)
  - ~3 million parameters

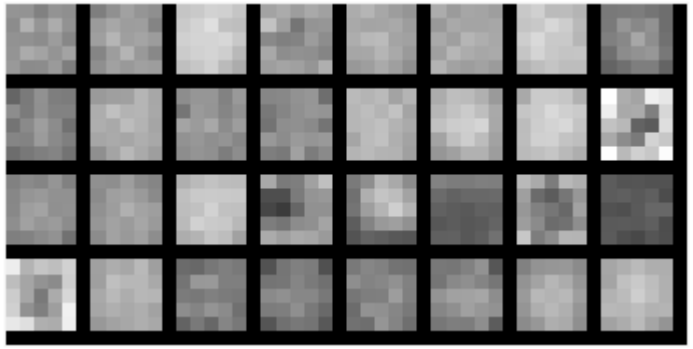


Examples of filters after training

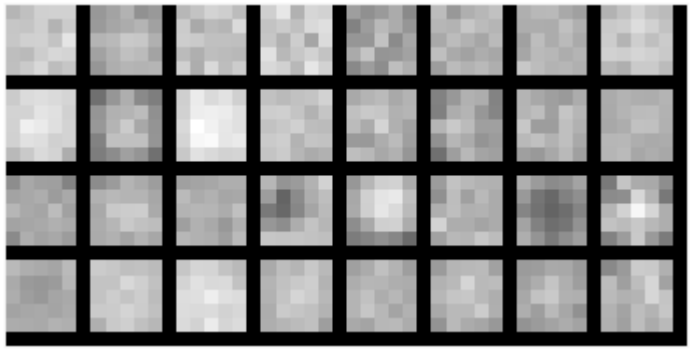
- conv1 /weights/I



- conv1 /weights/R



- conv1 /weights/Z



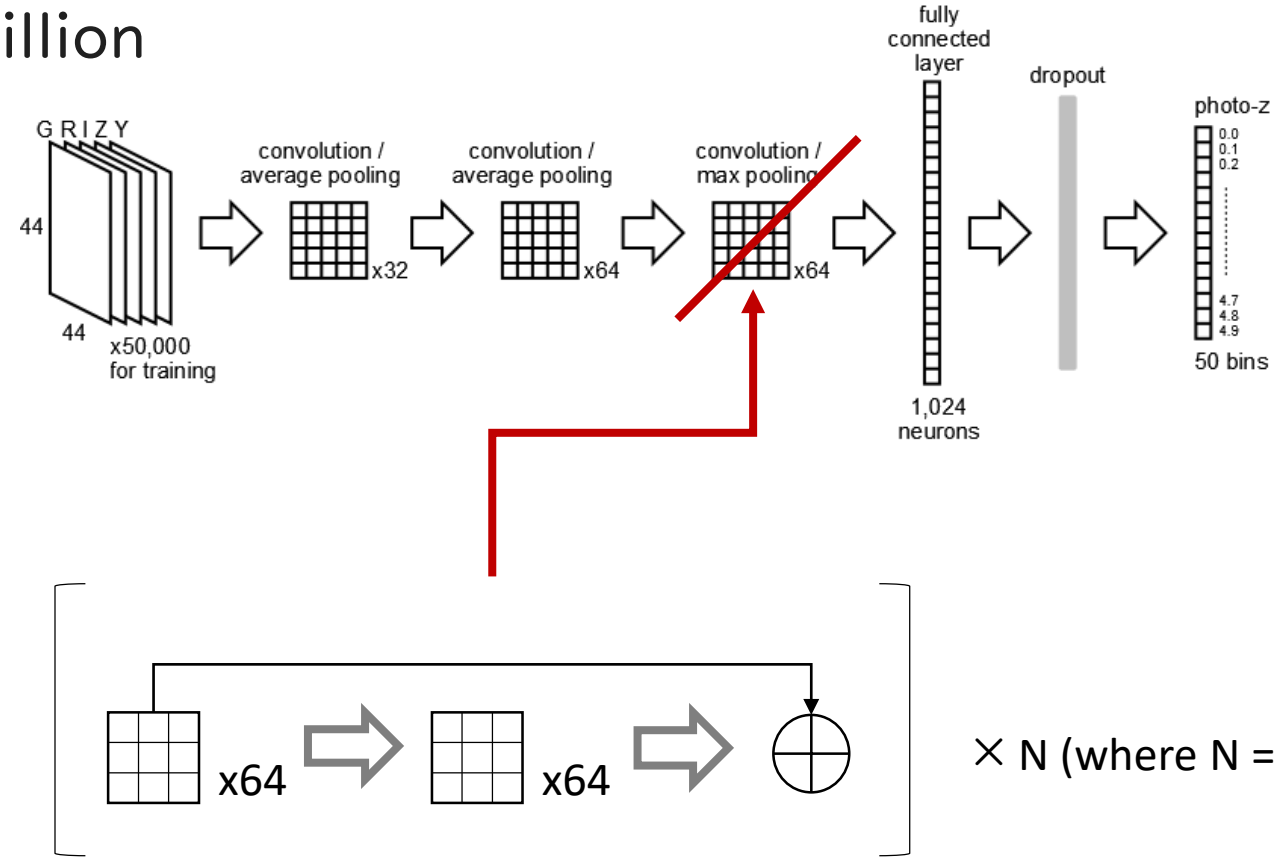
Challenges of deep learning

- Overfitting
  - Need complex model to fit data
  - But too much complexity tends to overfit training data and does not generalize well (for unseen data)

  - Mitigations
    - Regularization
    - Data augmentation
    - Dropout (NN only)
    - Batch normalization (NN only)
    - Shortcut connections for learning residuals (ResNet only)

Neural network architectures (2)

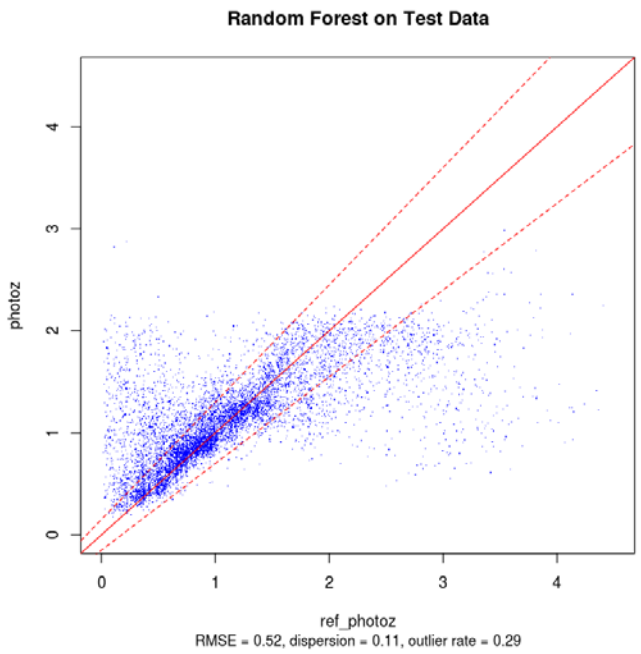
- ResNet (Residual Network)
  - ~(1 + 0.1N) million



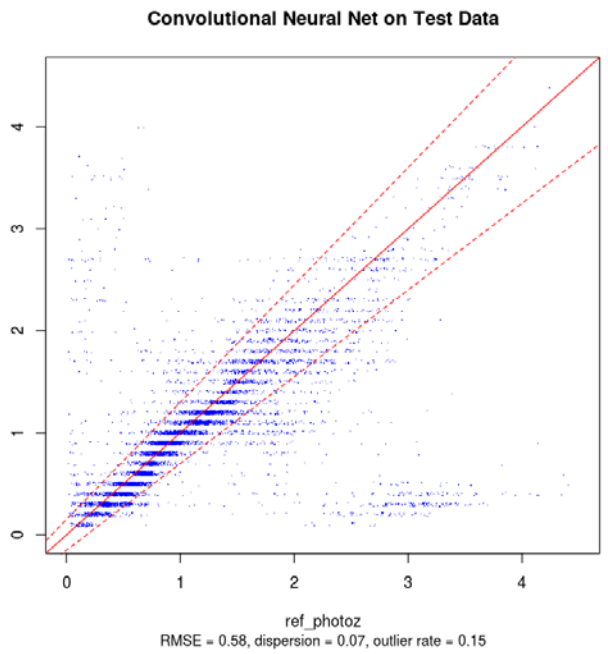
× N (where N = 20, 50, 70, 100)

Random forest vs. neural networks

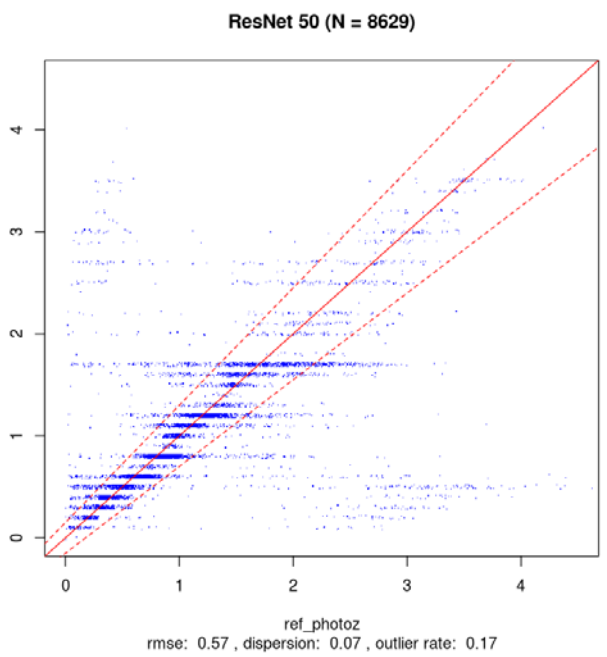
Random Forest



CNN



ResNet 100

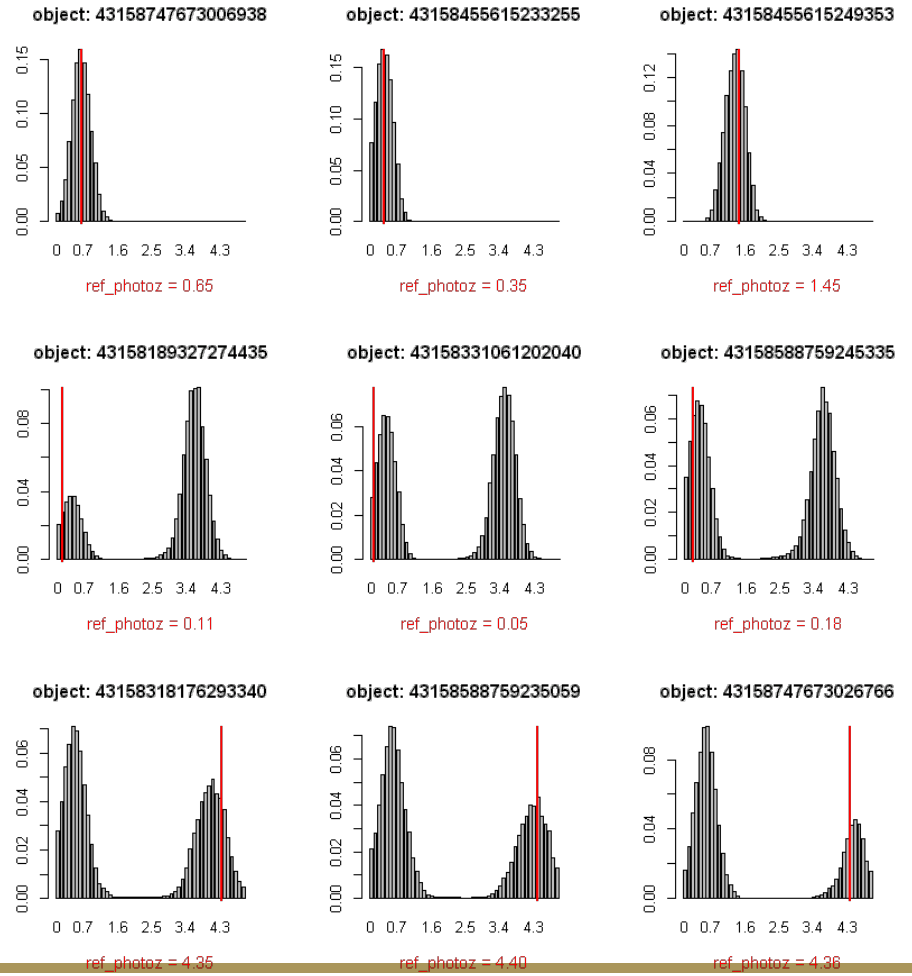


	Random Forest	CNN	ResNet 100
RMSE	0.52	0.58	0.57
Dispersion	0.11	**0.07**	**0.07**
Outlier Rate	0.29	**0.17**	**0.17**

Raw outputs from CNN

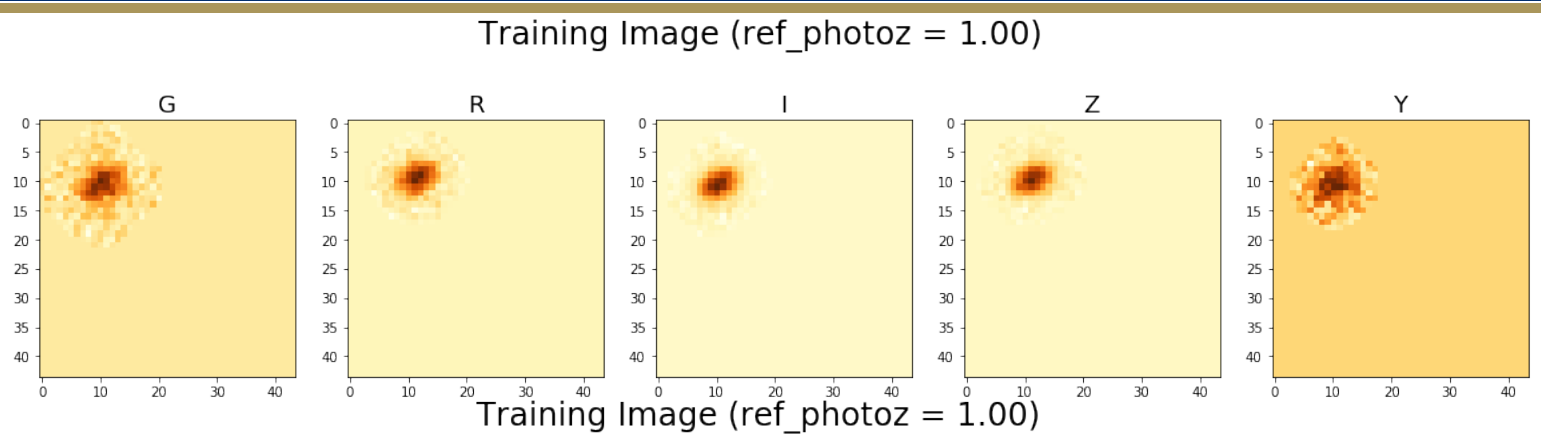
- Our CNN model outputs "probability distribution" of z

  - Almost right
  - Hugely overestimated
  - Hugely underestimated

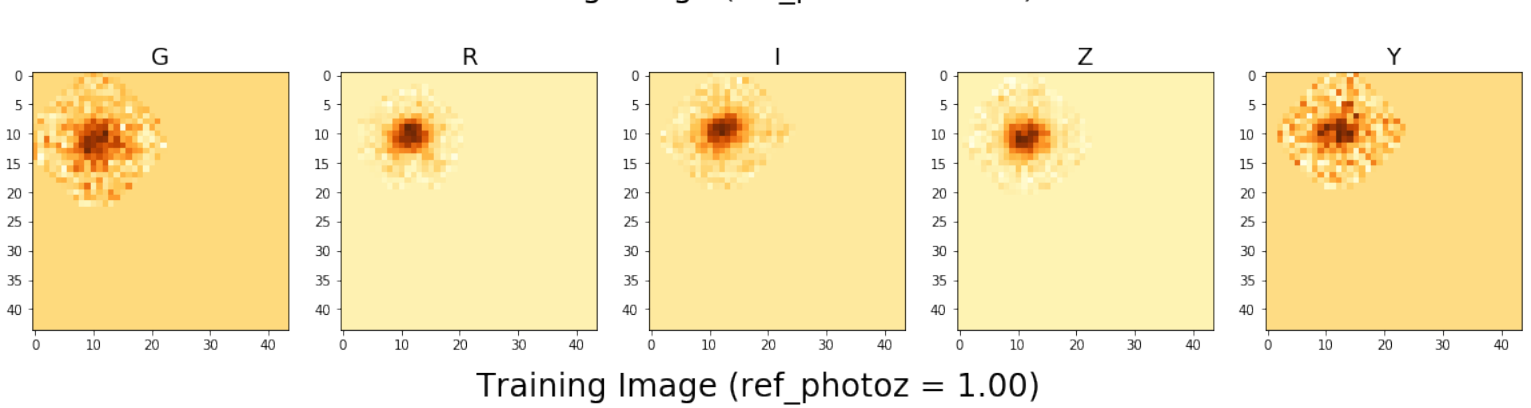


How does it perform on different seeings?

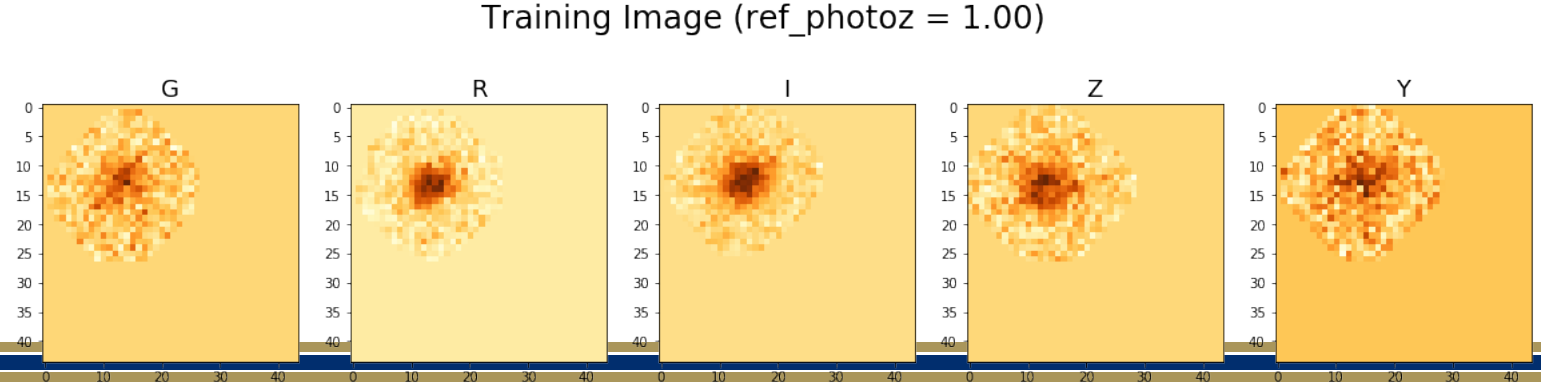
- Best



- Median



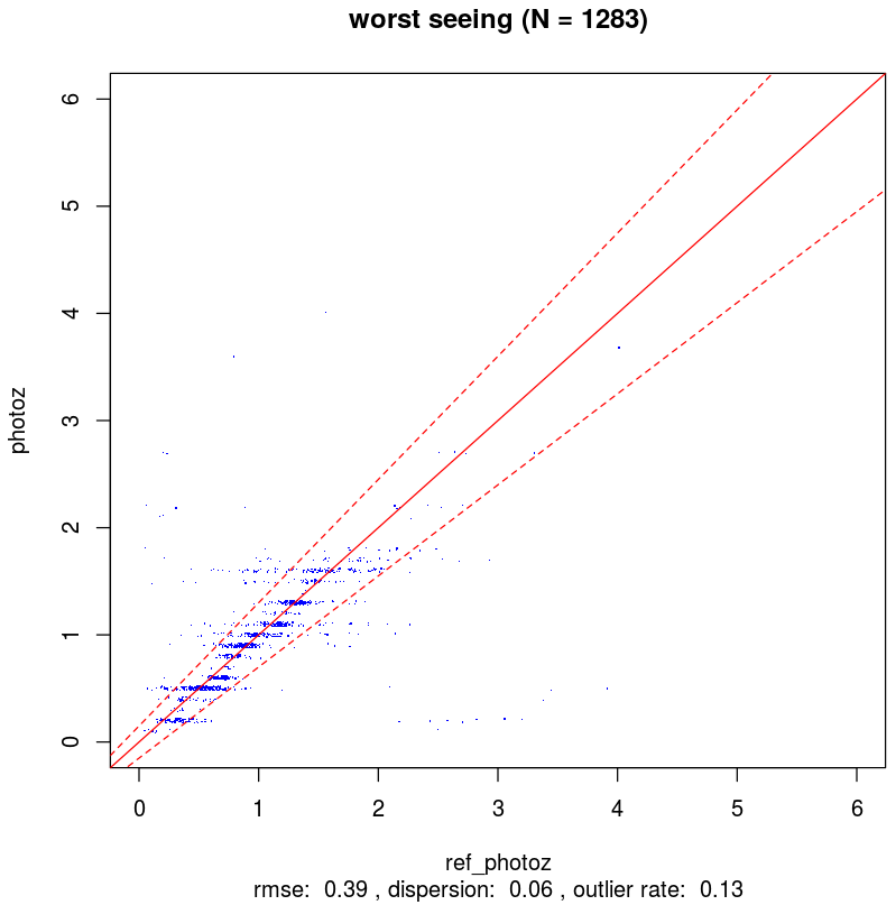
- Worst



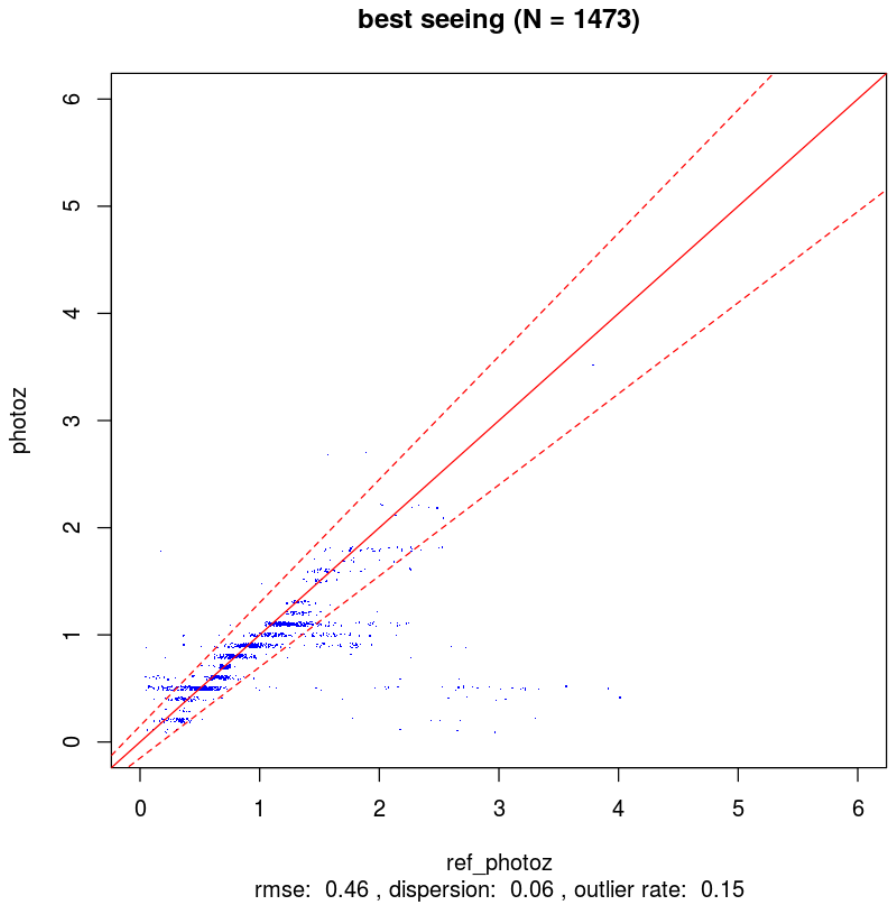
Images from different seeings

- Model trained on median seeing

Worst seeing (13% outliers)



Best seeing (15% outliers)



Summary

- Deep Neural Network: Powerful tool for predicting photo-z (especially for reducing outlier rates)

- Raw output of CNN/ResNet can be considered as a probability distribution of photo-z and may be used for identifying could-be outliers (e.g., double peaks)