Using Deep Learning for spectral classification, redshift prediction and anomaly detection

Speaker: Fucheng Zhong

School of Physics and Astronomy, Sun Yat-Sen University
School of Physics, Napoli Federico II University

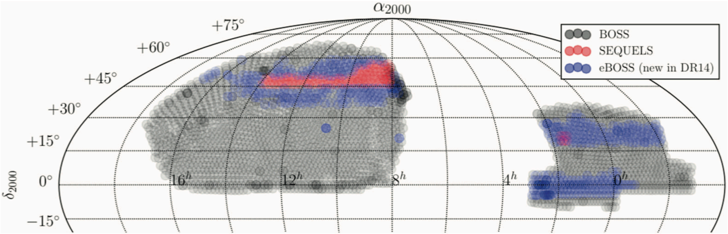
7th July 2024

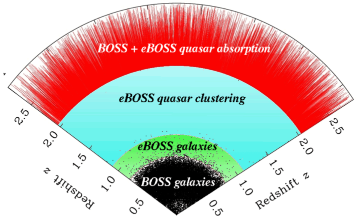
Contents

Motivations

- Build an ML pipeline for next-generation sky survey (4MOST).
- Pure data drive results can be complementary to the classical pipeline.
- The statistical feature of ML is used for uncertainty estimation.
- Fast and efficient feature of ML to face challenges of the enormous data volume of next-generation sky surveys.
- The transient objects require a fast pipeline.
- ...

SDSS overview



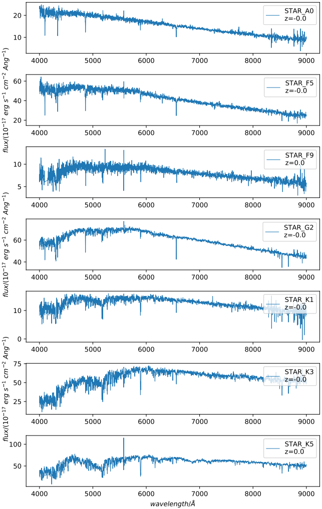


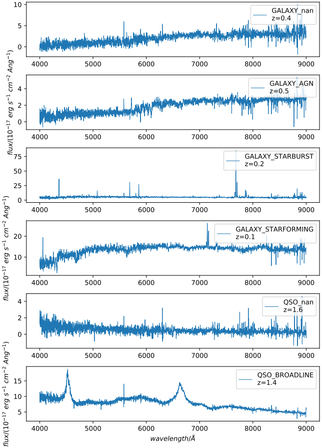
sdss4.org

ss4-DR14

- $>$ One million
- Cover $> 10,000$ deg2 (BOSS)
- 1,000-fiber spectrograph, resolution $R \sim 2000$
- Wavelength: 360-1000 nm

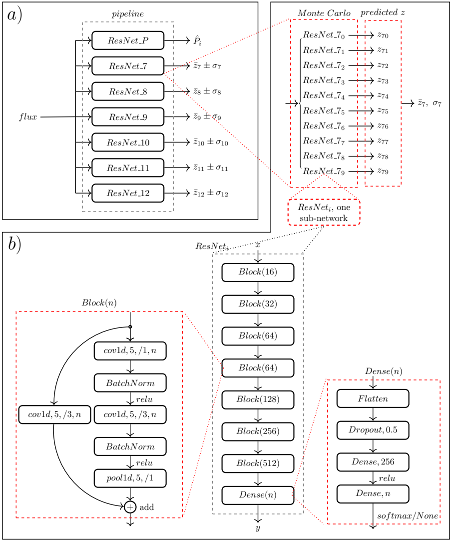
Training data





SDSS spectra, 13-subclass, 20k spectra used for each subclass.

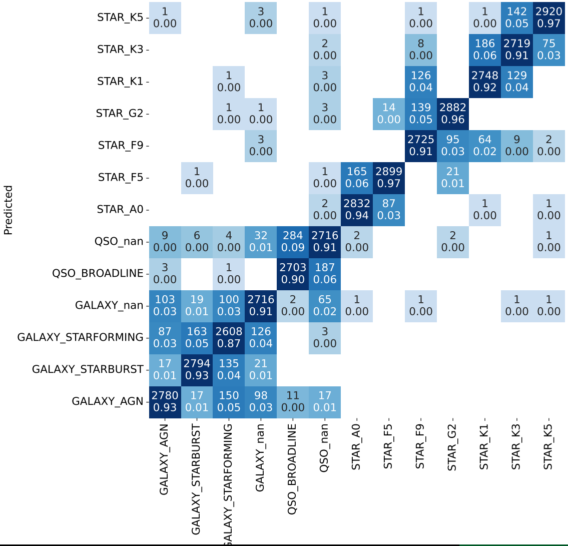
Pipeline



Architecture of pipeline.

- P_i gives probability.
- $\bar{z}$ gives expectation.
- σ give error.

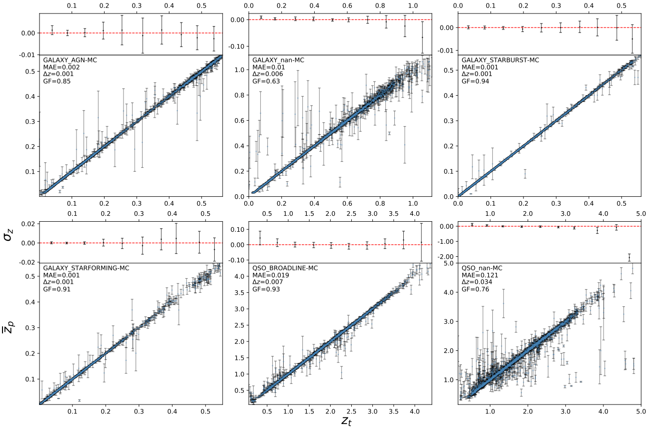
SDSS confusion matrix



SDSS classification

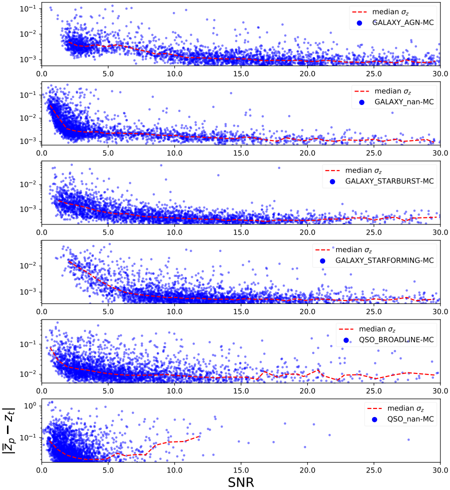
- Most subclass accuracy $> 90\%$.
- Average accuracy $\approx 92\%$.

SDSS redshift:



- $GF \approx 0.9$ for High SNR.
- $\Delta_z \sim \mathcal{O}(10^{-3})$ for galaxy.
- $\Delta_z \sim \mathcal{O}(10^{-2})$ for QSO.

SNR and uncertainty



Uncertainty is strong correlated with SNR when $SNR < 1$.

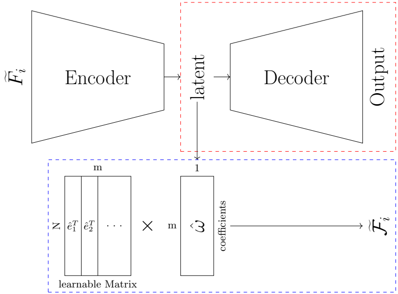
arxiv:2311.04146

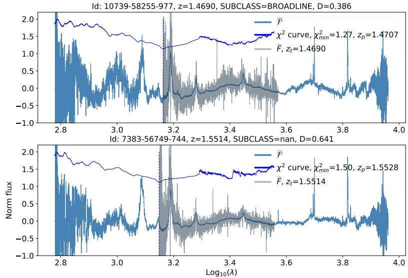
Improvement

Ideas of improve the z accuracy:

- More labeled data
  $\rightarrow$ Each pixel points as labels to learned – reconstruction. ("Cheap" or free lunch)
- z-shift translation symmetry in logarithmic scale
  $\rightarrow$ $ln(\lambda_o) - \ln(\lambda_e) = ln(1 + z)$ (prior)
- Statistically, the spectra are rough similar.
  $\rightarrow$ z causes we observed the different rest frame wavelength
- Reconstruction/modeling help detect anomaly and obtain residual.
  $\rightarrow$ A more quantify and visualized result.

Model: reconstruction & contrasting

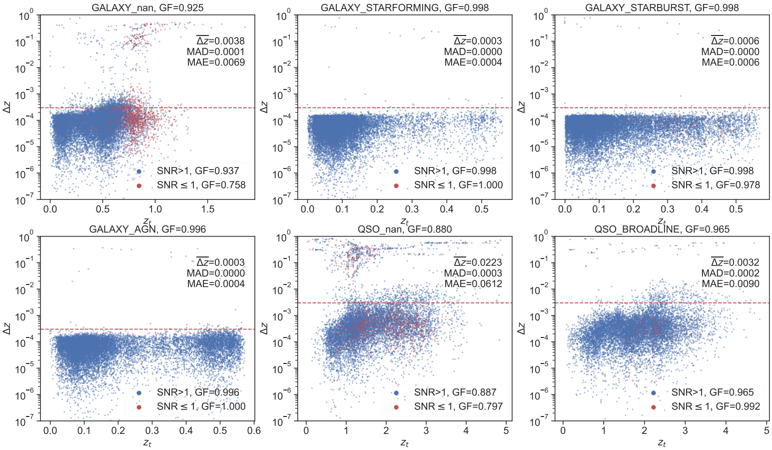




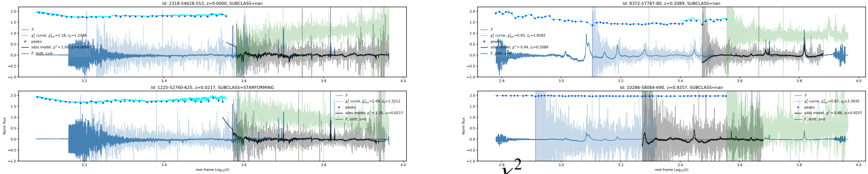
Reconstructing the spectrum and comparing the original and reconstructed.

Redshift accuracy

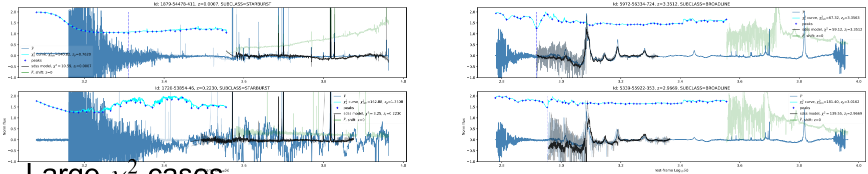
4MOST requirement $> 99\%$ for high SNR galaxy



Anomaly cases: degeneracy & large χ^2

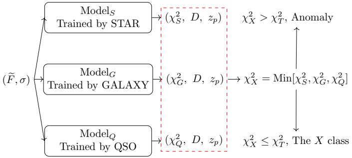


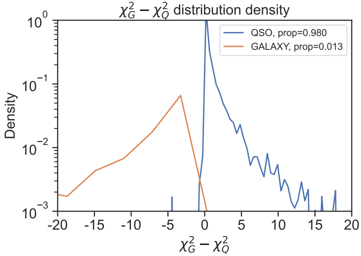
Redshift degeneracy cases, def $D = \frac{\chi^2_{2nd}}{\chi^2_{min}} - 1$, to remov outliers.

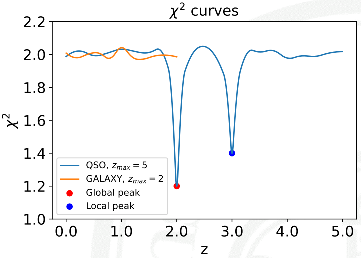


Large χ^2 cases

Pipeline

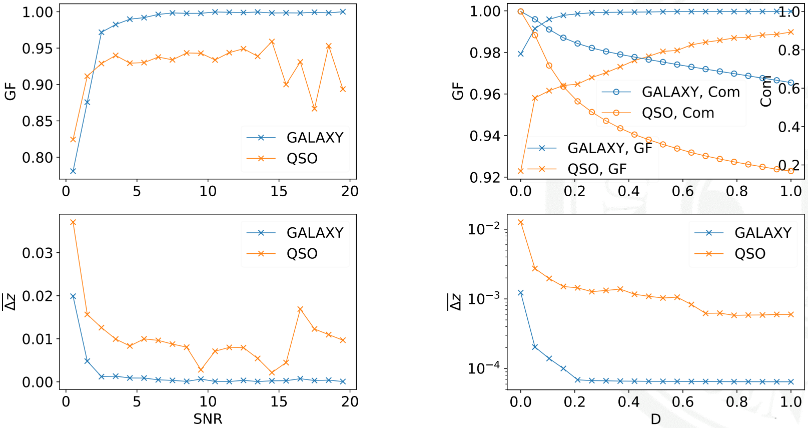




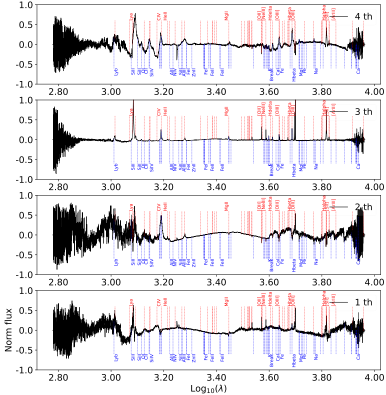


Scanning the whole redshift and class space with the pixel size interval, mapping the flux and error into a χ^2 curve. Classification and redshift estimate just finding the global minimum.

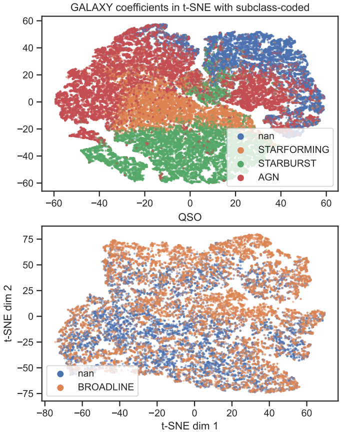
SNR, Degeneracy, GF and Completeness:



Visualization



Eigenspectra learned by network



tSNE projection of coefficients

Comparison and conclusion

Table 2. The comparison of different methods

Method	MAD GALAXY	GF GALAXY	MAD QSO	GF QSO	Time sec/spec/core
QXP++	0.00037	0.980	1.15	0.055	0.4
Redmonster	0.000027	0.999	0.000452	0.977	<0.4
pPXF	0.07723	0.854	-	-	0.03
GaSNet-III	0.000040	0.993	0.000282	0.919	0.001

Table A1. The performance of different training data.

Train num	GF	$\bar{\Delta}z$	MAD	MAE
5k QSO	91.65%	0.01438	0.00026	0.04117
5k GALAXY	96.79%	0.00216	3.75×10^{-5}	0.00376
10k QSO	92.28%	0.01288	0.00024	0.03480
10k GALAXY	97.50%	0.00148	3.71×10^{-5}	0.00255
20k QSO	92.29%	0.01263	0.00023	0.03481
20k GALAXY	97.93%	0.00123	3.66×10^{-5}	0.00208

Table C1. The performance using different number of eigenspectra.

eigenspectra num	GF	$\bar{\Delta}z$	MAD	MAE
3 QSO eigenspectra	91.51%	0.01517	0.00024	0.03952
3 GALAXY eigenspectra	98.06%	0.00145	3.66×10^{-5}	0.00217
5 QSO eigenspectra	90.97%	0.01309	0.00024	0.03496
5 GALAXY eigenspectra	97.79%	0.00187	3.70×10^{-5}	0.00291
10 QSO eigenspectra	92.29%	0.01263	0.00023	0.03481
10 GALAXY eigenspectra	97.93%	0.00123	3.66×10^{-5}	0.00208

Reaching the same accuracy as the classical method with faster speed, requiring a low training sample compared with former deep learning (GaSNet-II), using few learned eigenspectra to restore the spectra, it agrees with the conclusion before.

Thank you for your attention!