The quasar luminosity function at z~5 improved by artificial neural network and Bayesian information criterion

Suhyun Shin (SNU)

Advisor : Myungshin Im
Co-worker : Yongjung kim

GSM2022

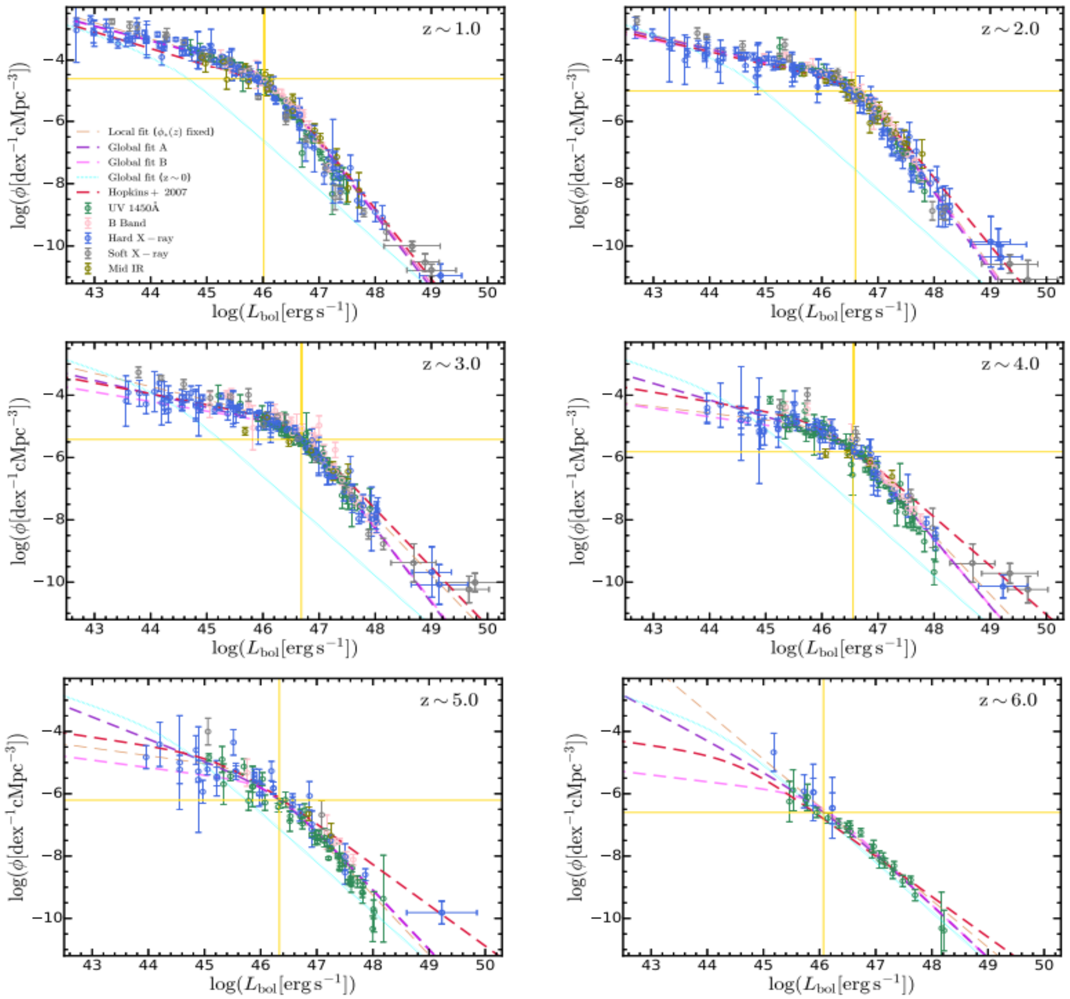
Quasar Luminosity Function (LF)

Shen+2020

- Representing **the demography depen
  dent on the luminosity**

- A key to investigate **the evolution of q
  uasars (SMBH and host galaxies)** acr
  oss the cosmic time

- Can be used for **estimating the contri
  bution of quasar to keeping the IGM
  ionized**

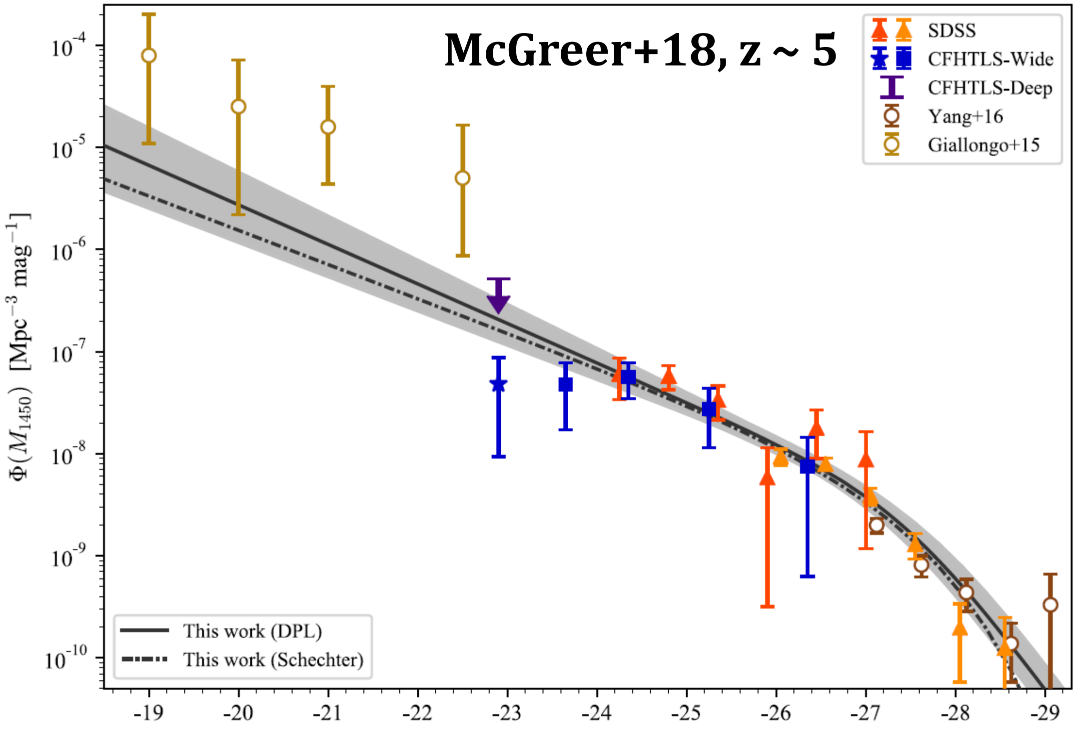


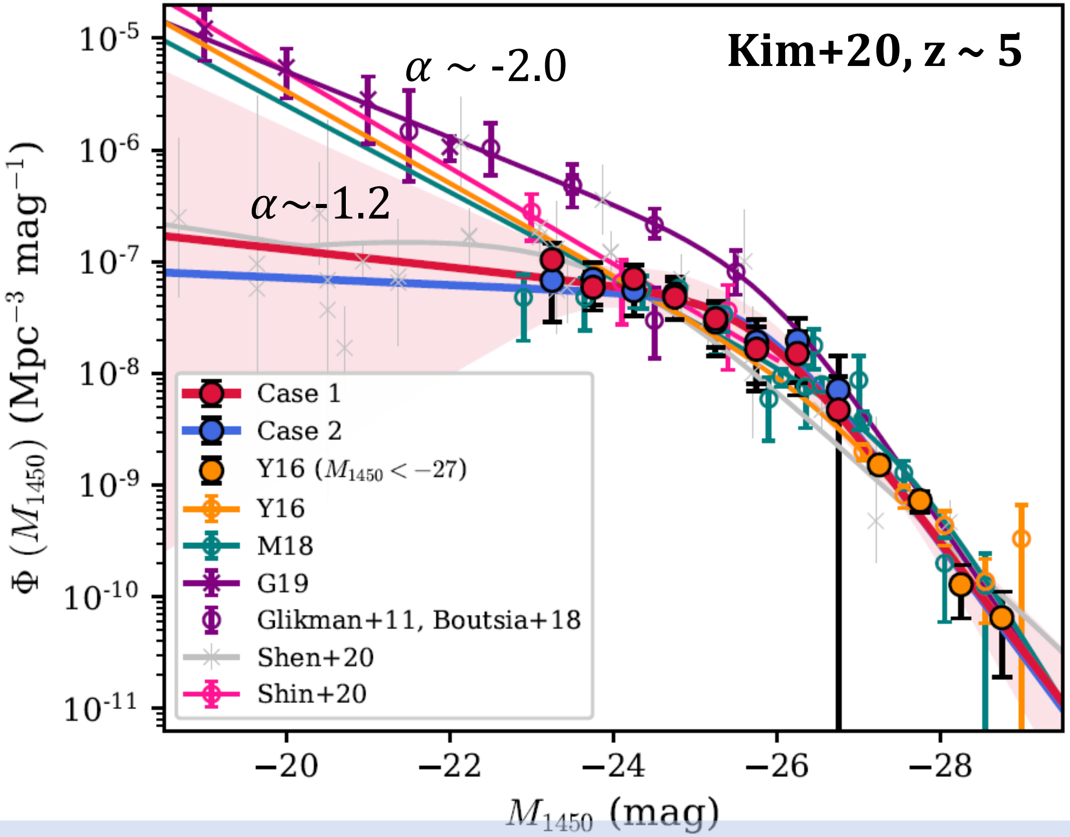
→ A lot of studies give efforts to construct quasar LF at various redshifts

Quasar Luminosity Function (LF)

$$\Phi_{\rm model}(M_{1450}) = \frac{\Phi^*(z=5.2)}{10^{0.4(\alpha+1)(M_{1450}-M^*_{1450})} + 10^{0.4(\beta+1)(M_{1450}-M^*_{1450})}}.$$

Faint-end slope



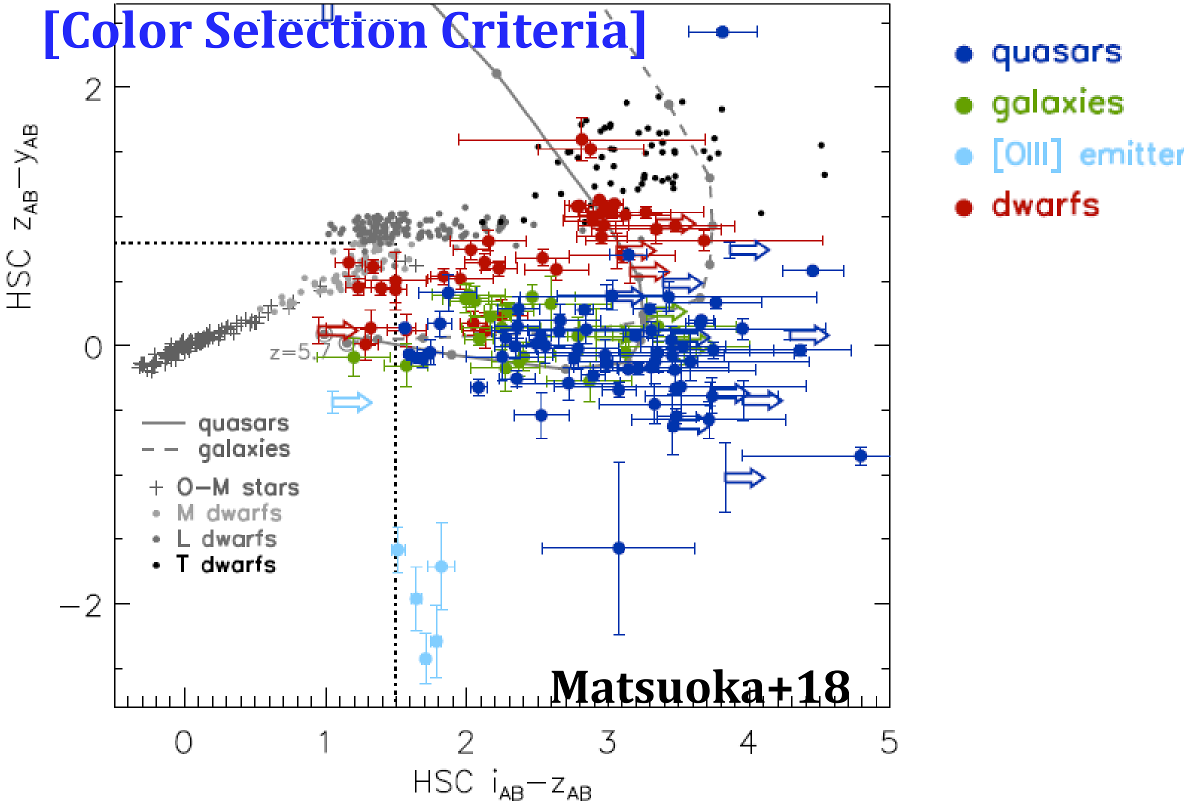


→ The discrepancies between the quasar LFs at high redshift exist in the faint regime

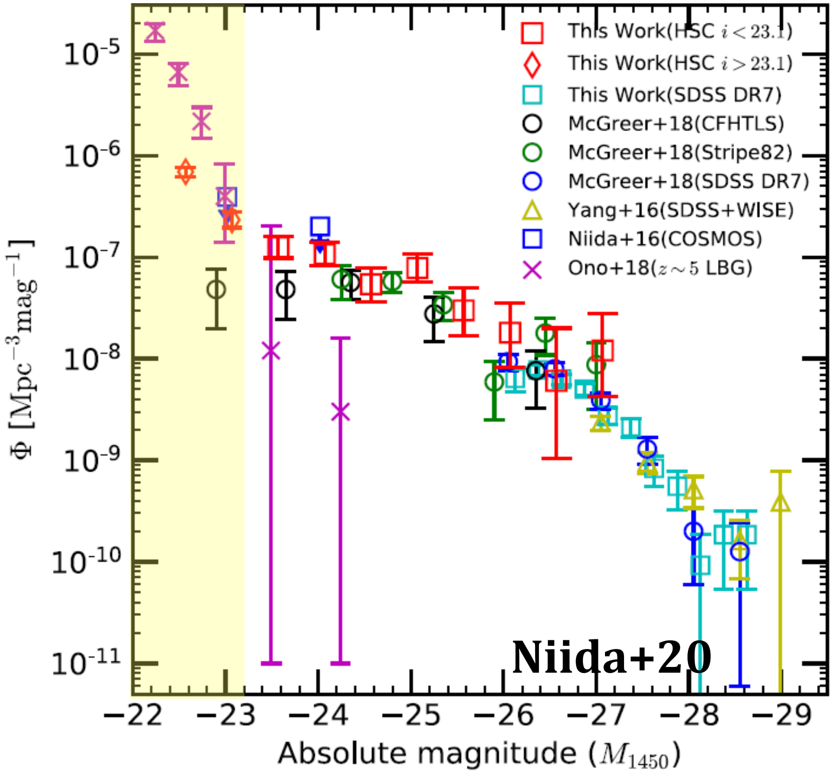
Contamination sources of quasar survey

- **The contamination of M-dwarf stars and high-z galaxies** makes it difficult to select promising quasar candidates

[Color Selection Criteria]



Matsuoka+18



Niida+20

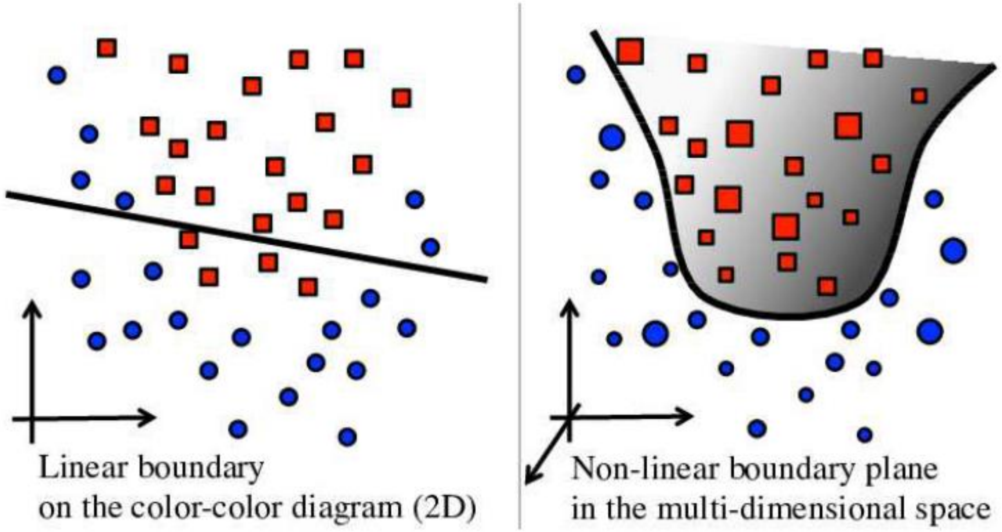
→ The need for new attempt to separate high-z quasar from other contamination sources

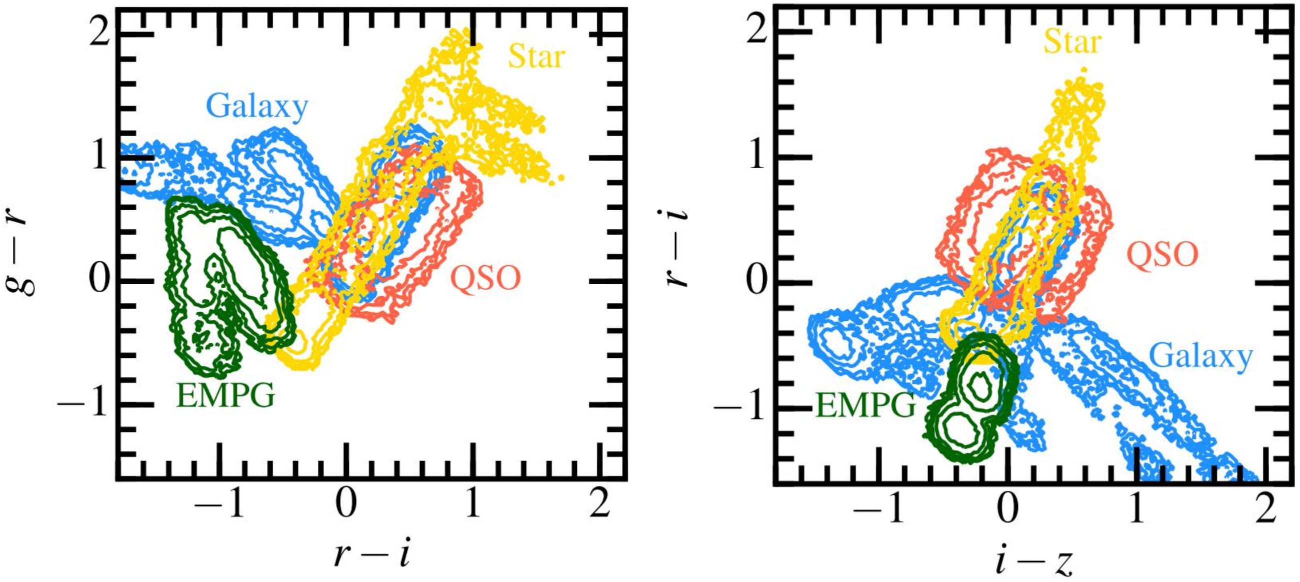
→ **Deep learning and ΔBIC calculation**

Merits of Deep Learning

- known as **great solvers to classification problem**
- Use criteria in **multi-dimensional space**
- Enable us to separate objects based on **non-linear boundaries**
- **Optimize these boundaries** by decreasing a loss
- **Computing time** for DNN selection is much shorter than SED fitting

Kojima+20



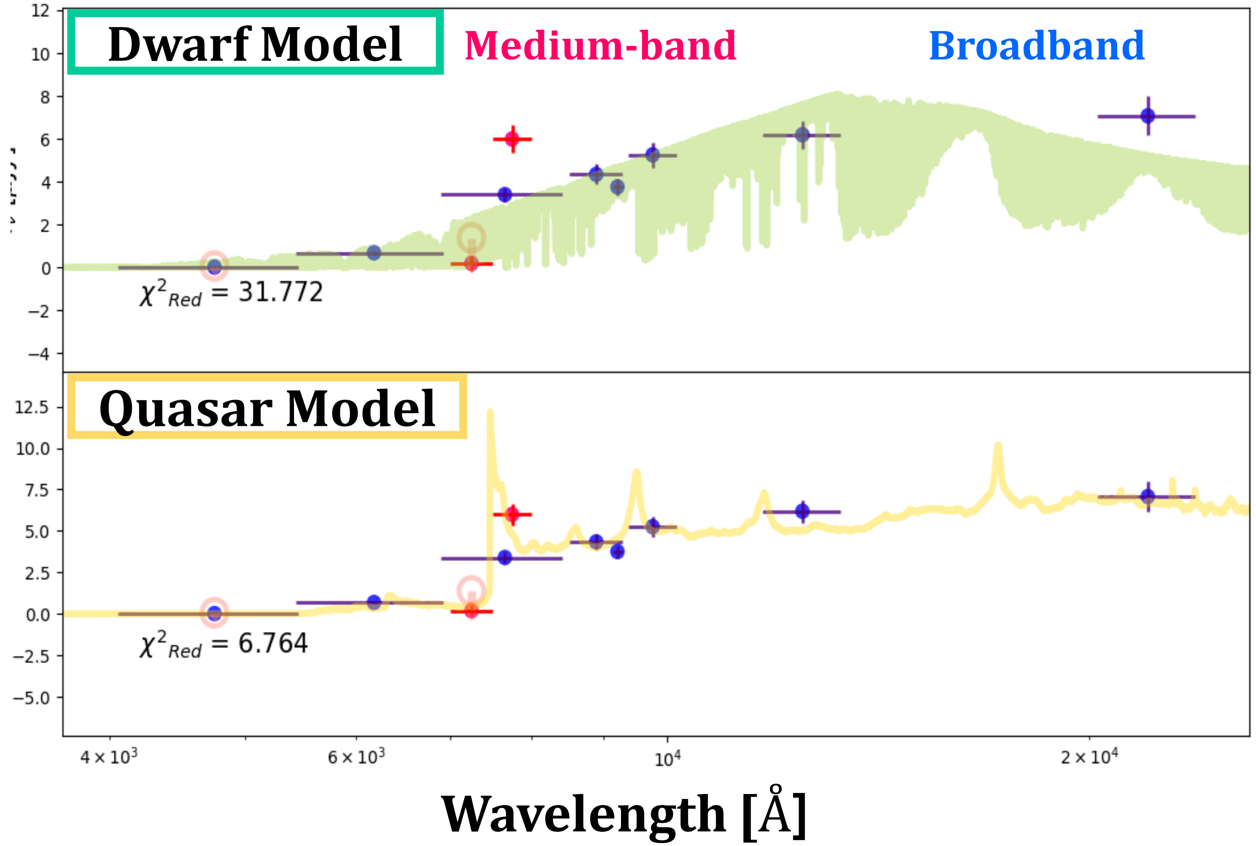


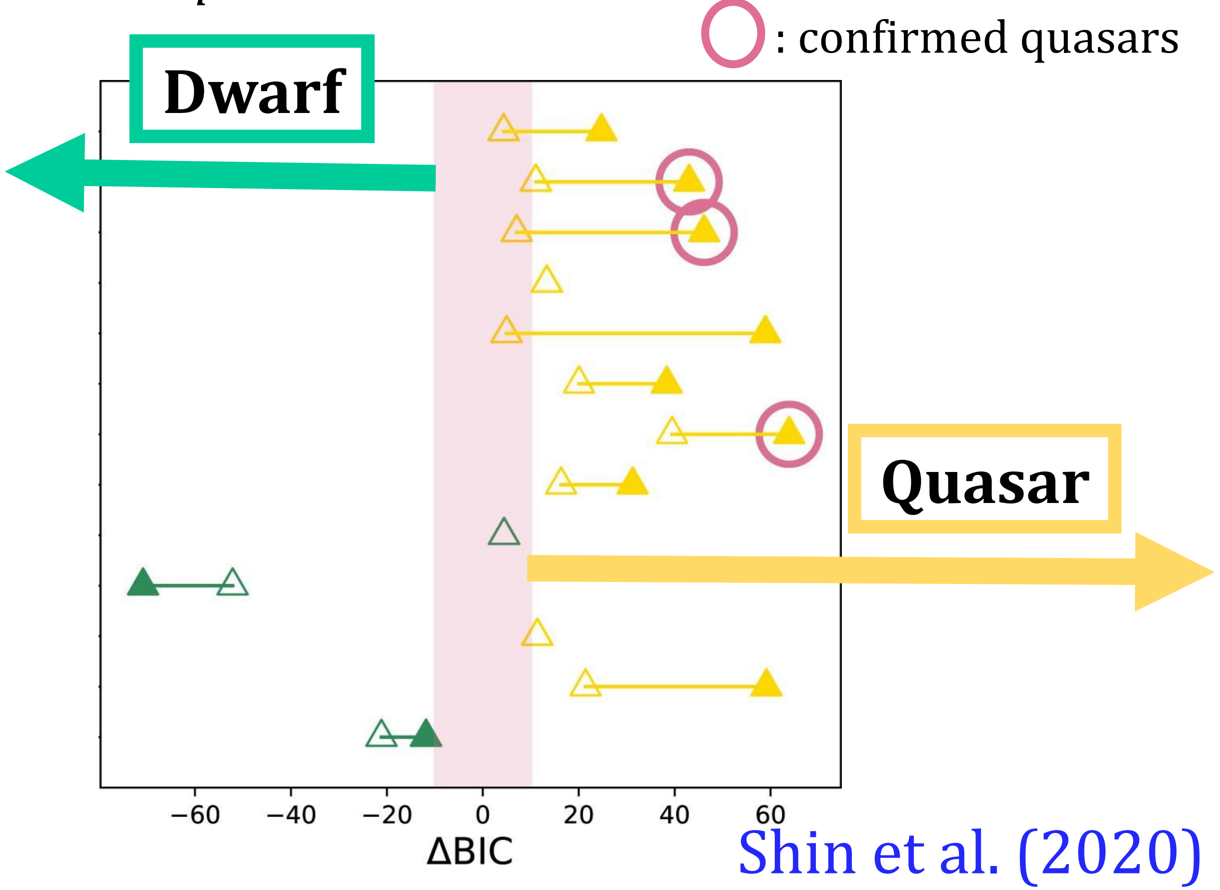
Merits of ΔBIC calculation

L : likelihood
k : the number of parameters
n : the number of data

$$\text{Bayesian information criterion(BIC)} = -2\ln L + k\ln n$$

$$\Delta\text{BIC} = [-2\ln L + k\ln n]_{dwarf} - [-2\ln L + k\ln n]_{quasar}$$ **(> 10 : very strong)**

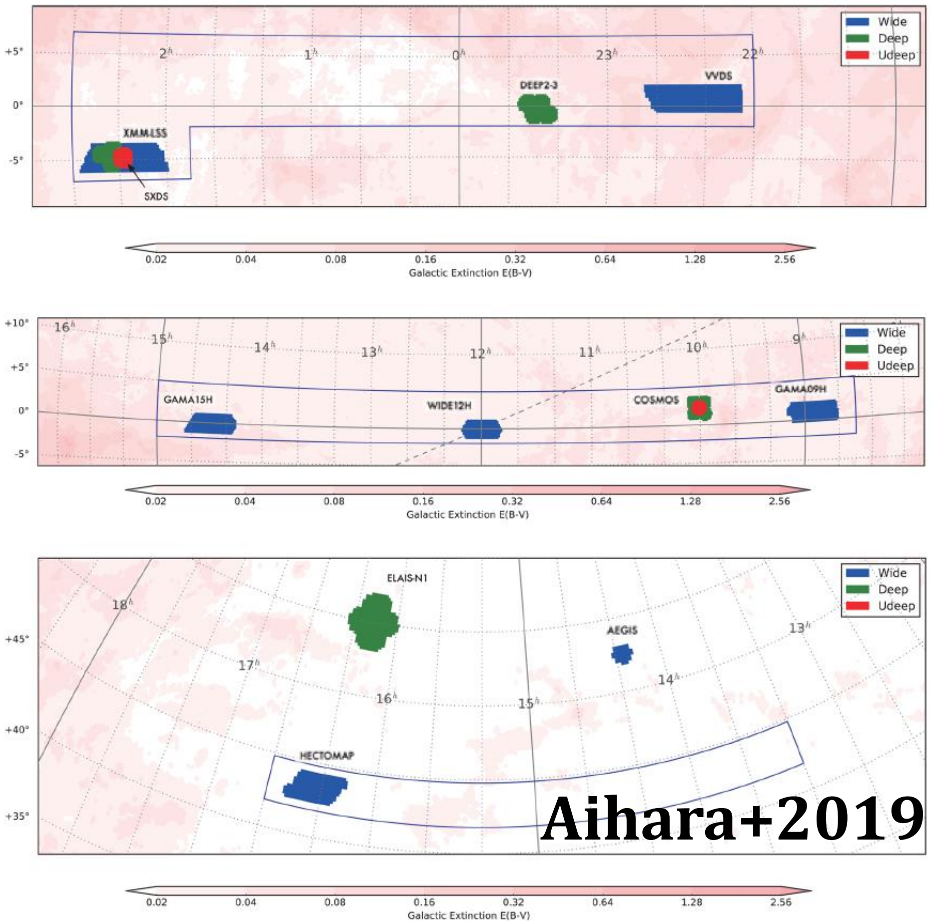




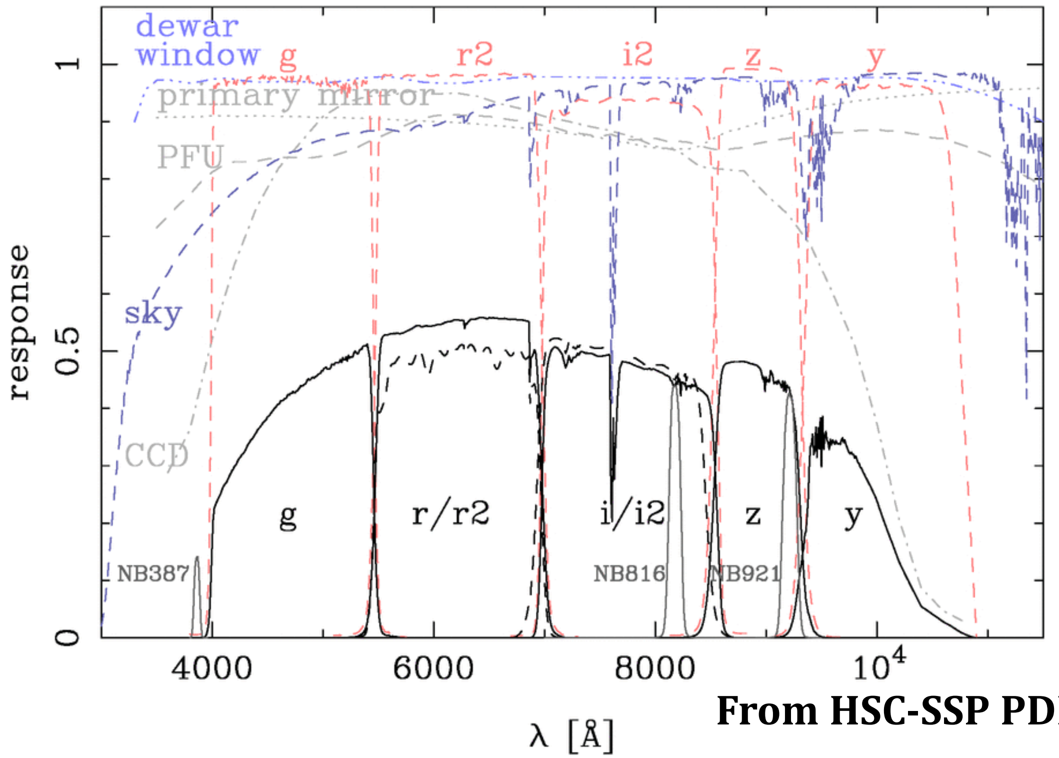
Shin et al. (2020)

→ Efficient to discern quasar-like candidates from M dwarf contamination

HSC-SSP (PDR2)



Aihara+2019



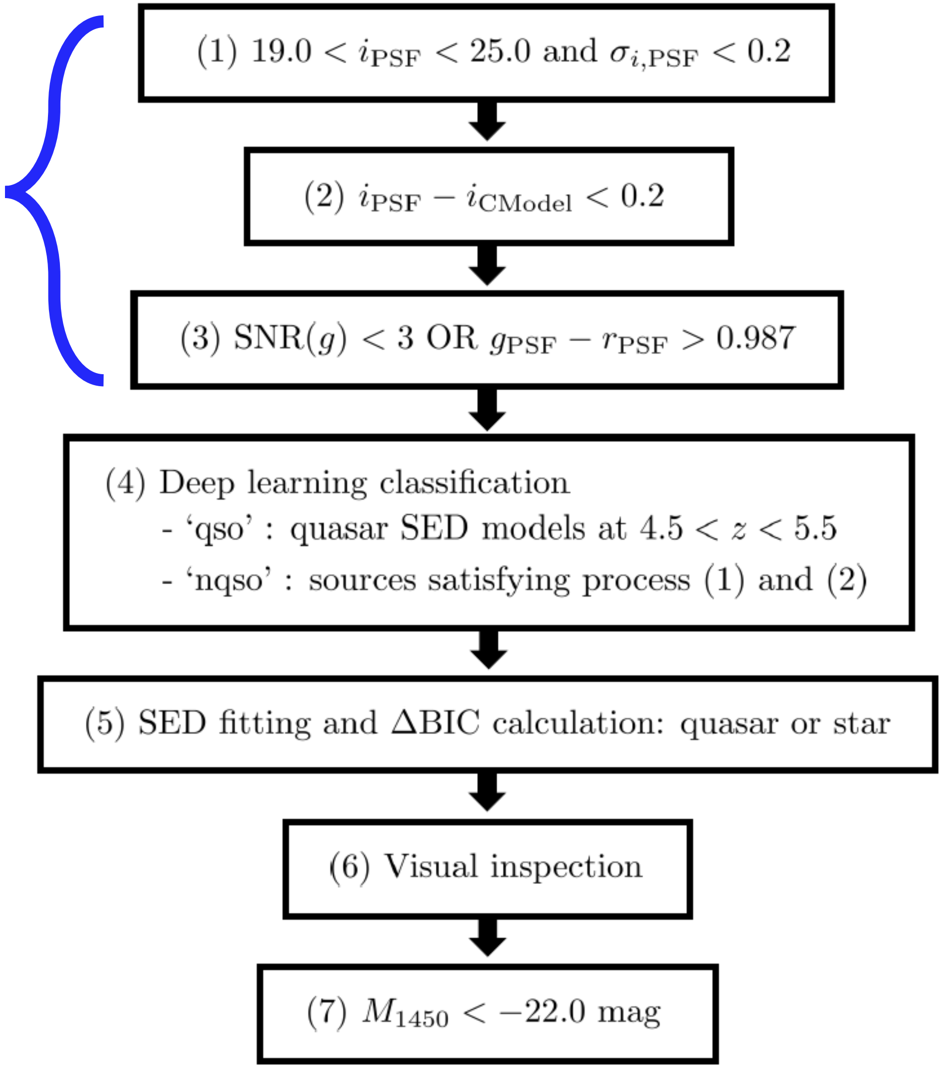
From HSC-SSP PDR2 homepage

Layer	Depth	Area	Filter
Wide	i_{AB} ~ 26.2	**1400**	g,r,i,z,y
Deep	i_{AB} ~ 26.7	**26**	g,r,i,z,y + ~~NB316~~,NB816, NB921
Ultra-deep	i_{AB} ~ 28.0	**3.6**	g,r,i,z,y + NB316,NB816, NB921

→ **Using the Deep layer of HSC-SSP (PDR2) :** i_{AB} **~ 26.7** (5-σ depth), Area ~ 26 deg^2

→ make our quasar LF **~ 1.0 mag deeper** than previous quasar LFs

Quasar candidates selection



Pre-selection

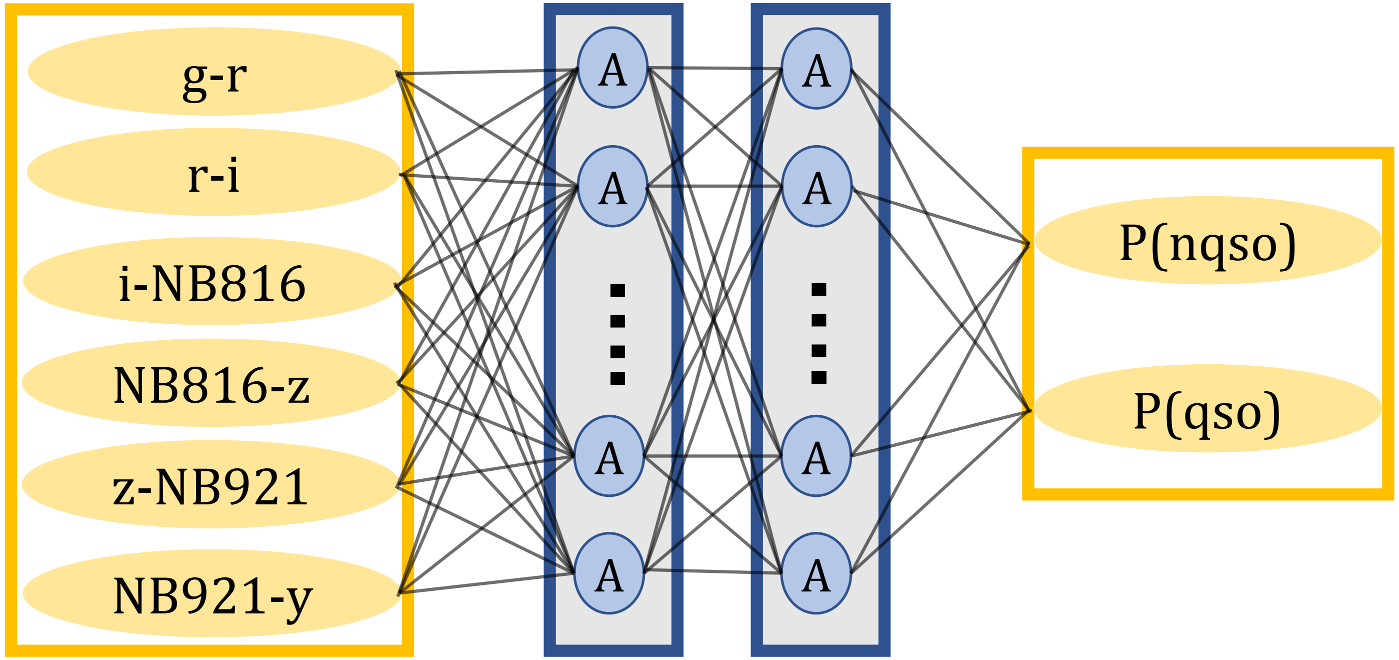
→ Reliable photometry

→ **Point source selection-** exclude extended local galaxies

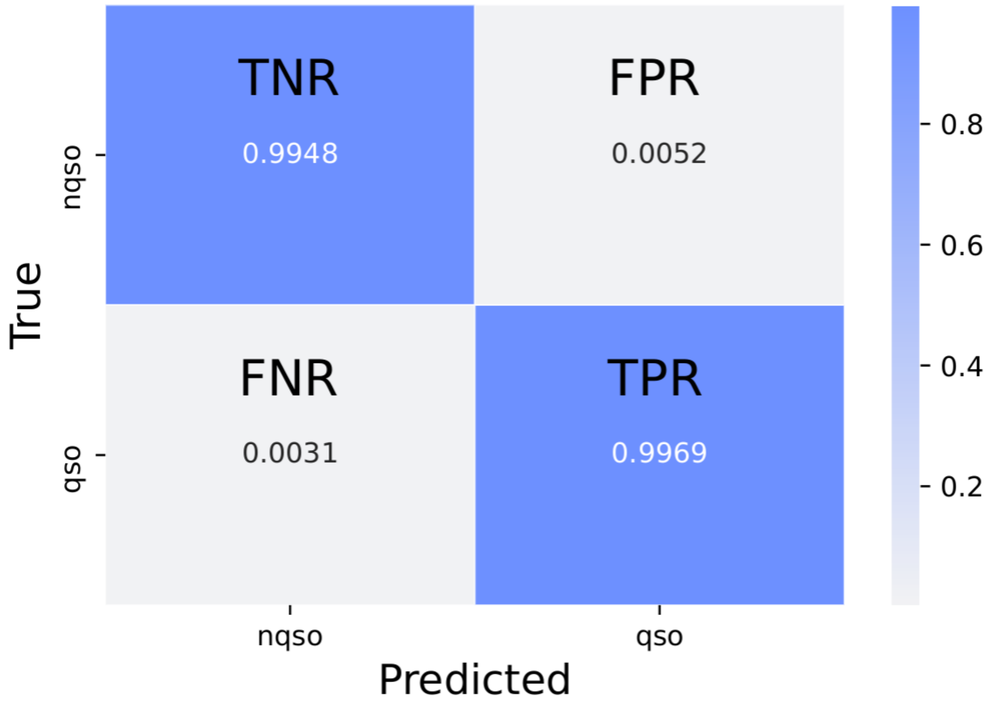
→ **Red objects**

Deep learning (DL)

Input **Hidden** **Output**



Confusion matrix



A : Activation function, ReLU

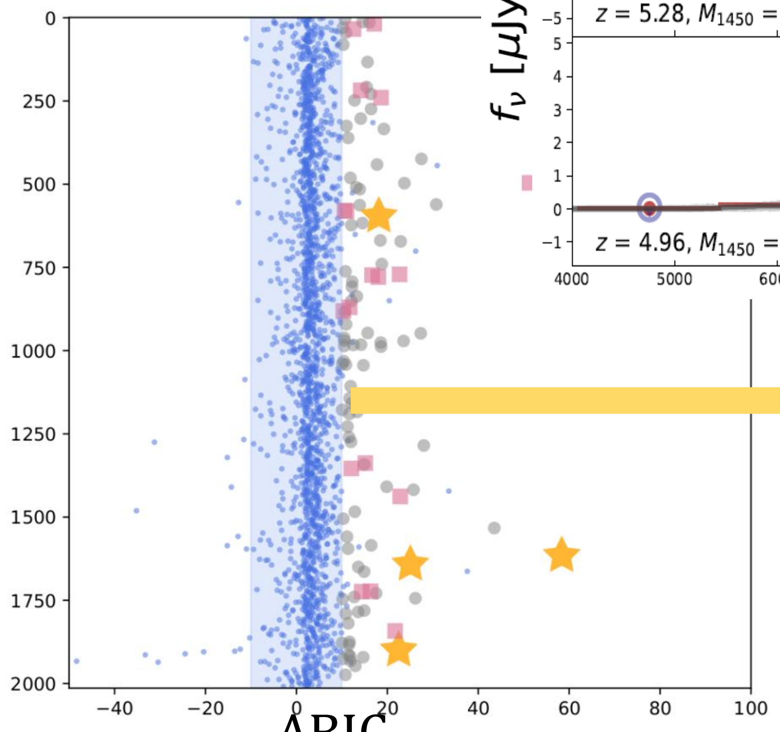
→ **Classification: i-band-detected Point sources** (nqso) vs **quasars at 4.5 < z < 5.5** (qso)

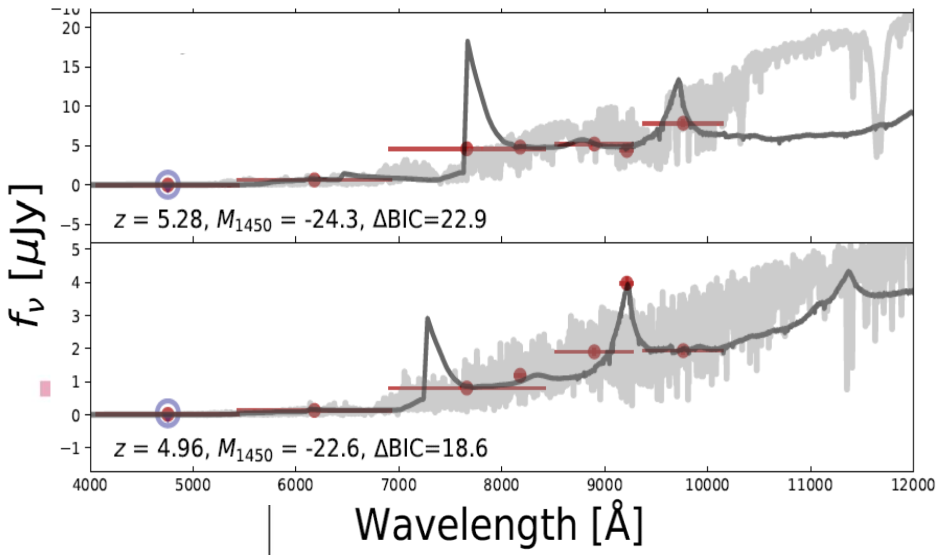
→ DL can make criteria in **multi-dimensional space** (g-r, r-i, i-NB816, NB816-z, z-NB921, NB921-y)

ΔBIC calculation

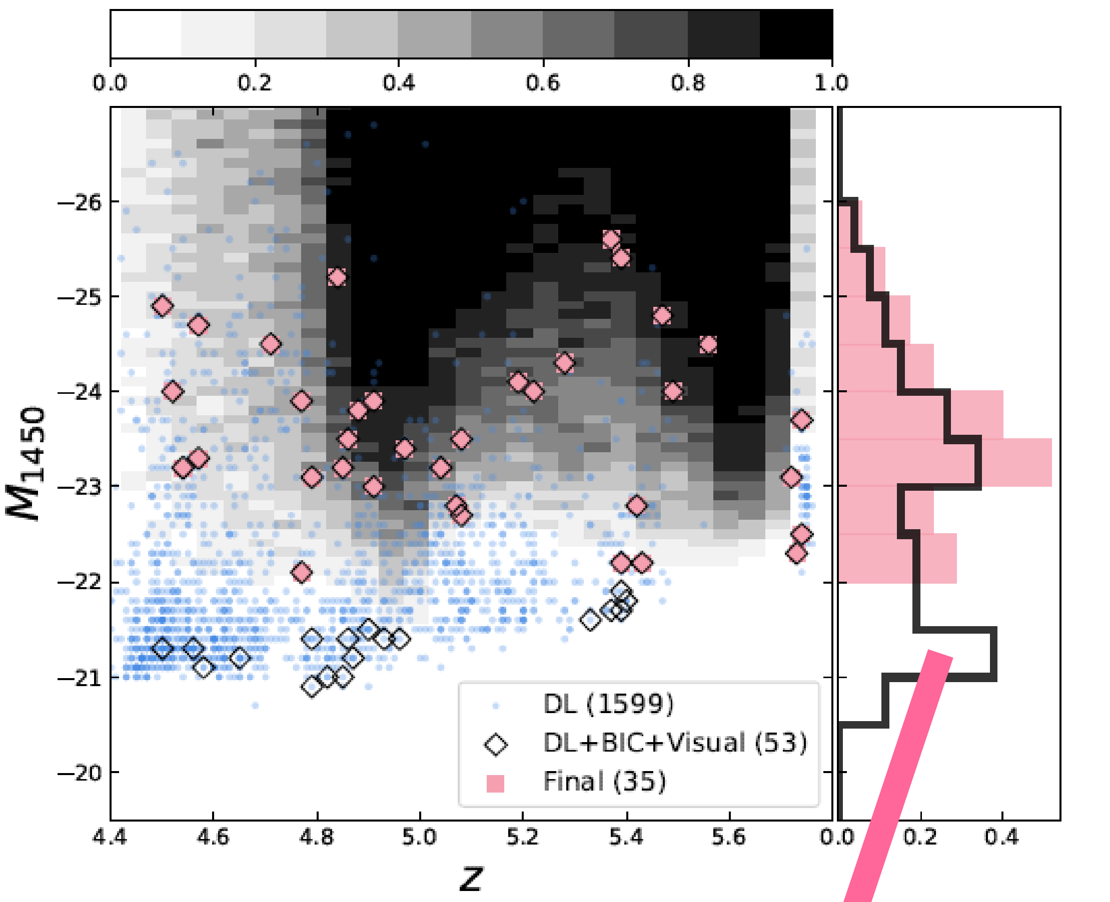
SED fitting using **quasar** and **star** models

|ΔBIC| < 10





Qso Candidates



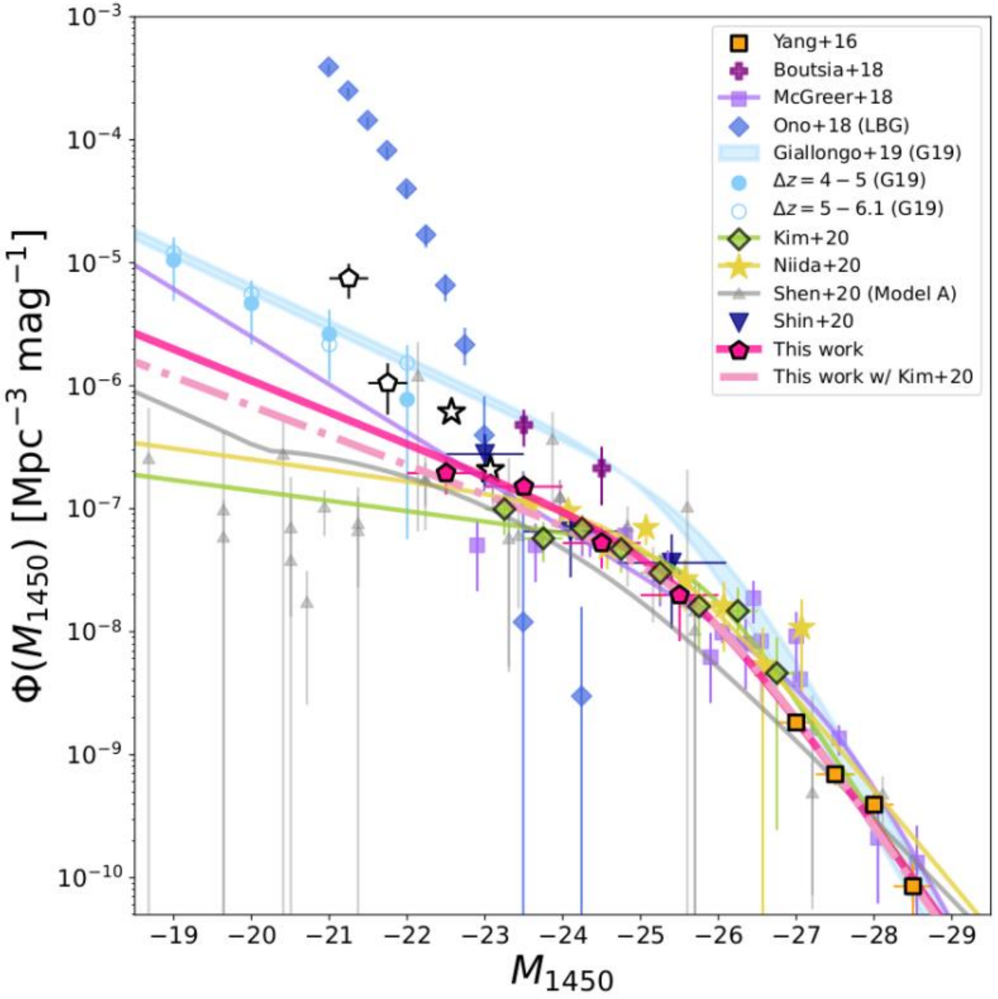
Bimodal distribution

→ Possible contamination from LBG

→ Consider quasar candidates with M_{1450}<-22.0 mag

→ Using ΔBIC calculation, **exclude star-like candidates**

The quasar LF



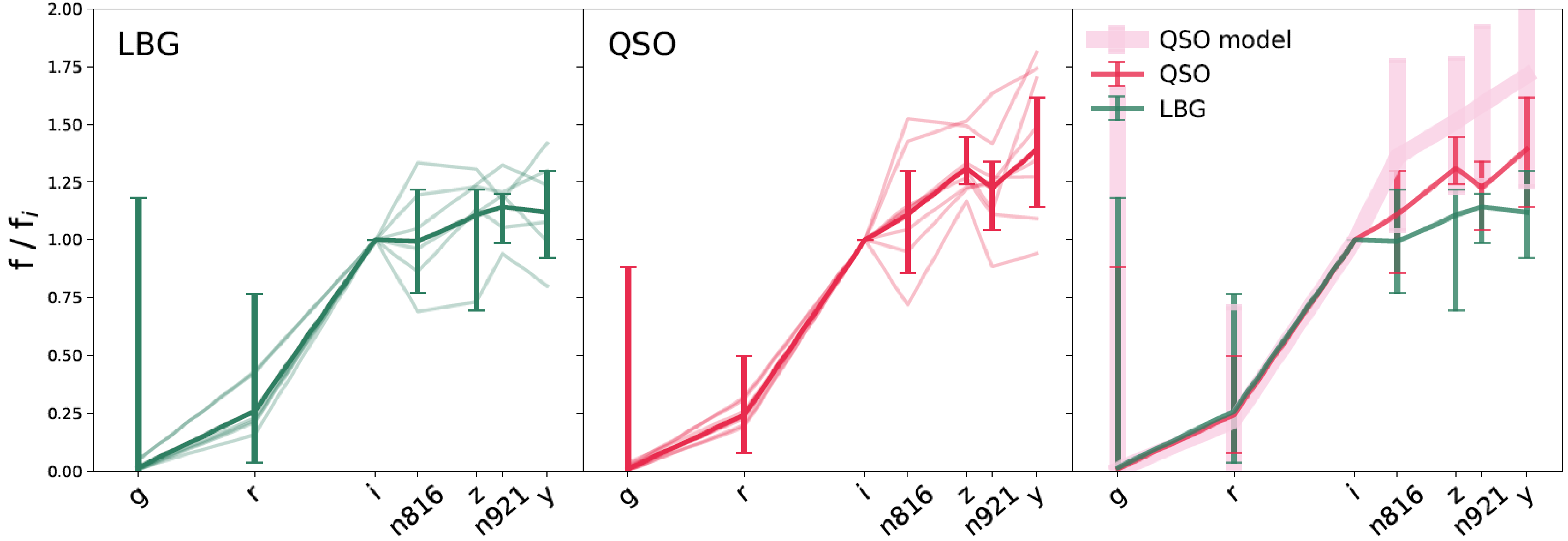
- **Built QLF reaching M_{1450} = -22.0 mag** without multiwavelength data
- **Derived $\alpha = -1.61^{+0.21}_{-0.19}$** is consistent with the best-fit α in Niida+20 and Kim+20 within 1-σ level.
- Quasars did **not significantly contribute the ionizing background.**
- **Clear difference of the number density at $M_{1450} > -23.5$** between Ono+18 (LBGs at z~5) and this work**, implying lower contamination rate of high-z galaxies.**

→**How to check** the efficiency of our quasar selection process ? ***Contamination & Recovery***

Galaxy contamination

Using spectroscopically confirmed galaxies and AGNs at z = 4-7 from HSC-SSP (Ono+18)

All observed fluxes of QSOs/LBGs are normalized to their i-band fluxes

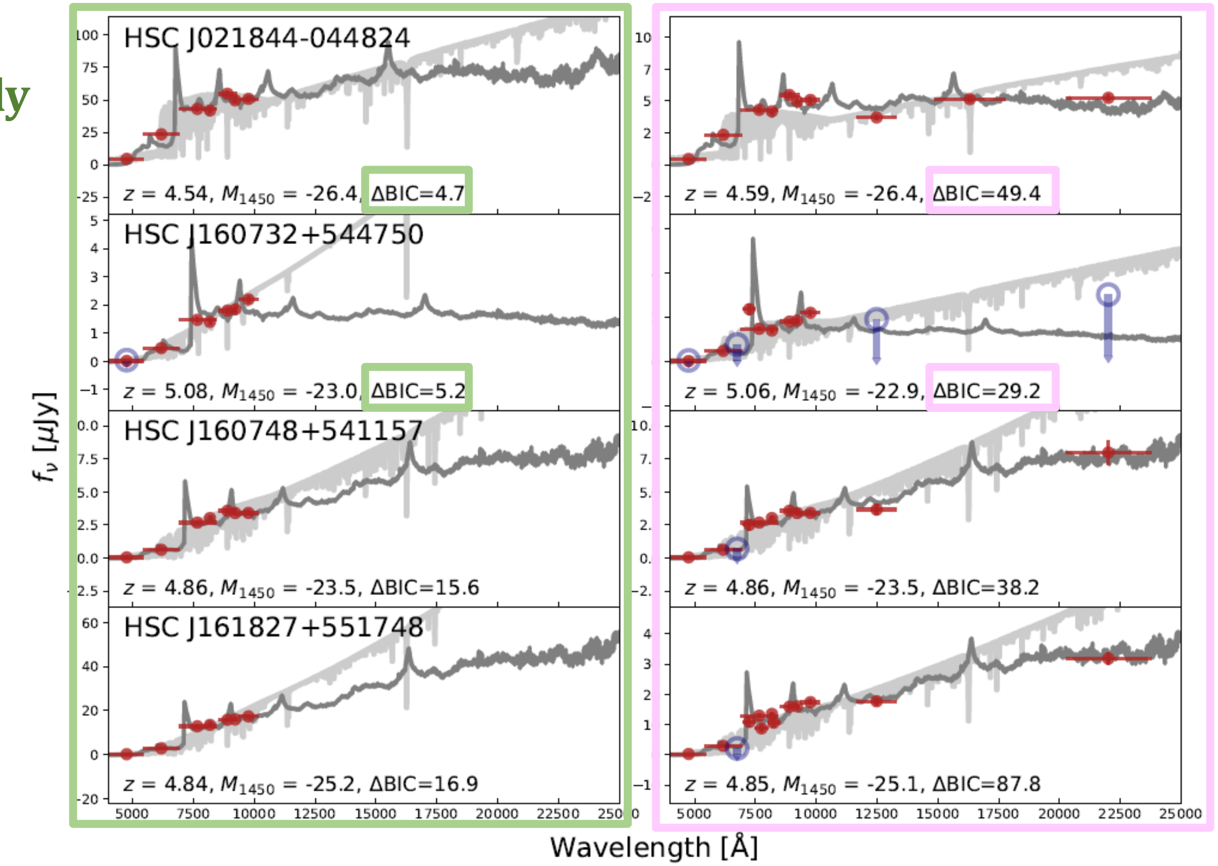


- Niida+20 Color selection → 3/8 (4.0 < z < 6.0)
- **Deep learning & Bayesian statistics → 1/8** (4.0 < z < 6.0)

Quasar recovery of our selection

One quasar and three promising candidates in Shin et al. (2020)

HSC only



HSC
+NIR(IMS or DXS)
+Medium-bands

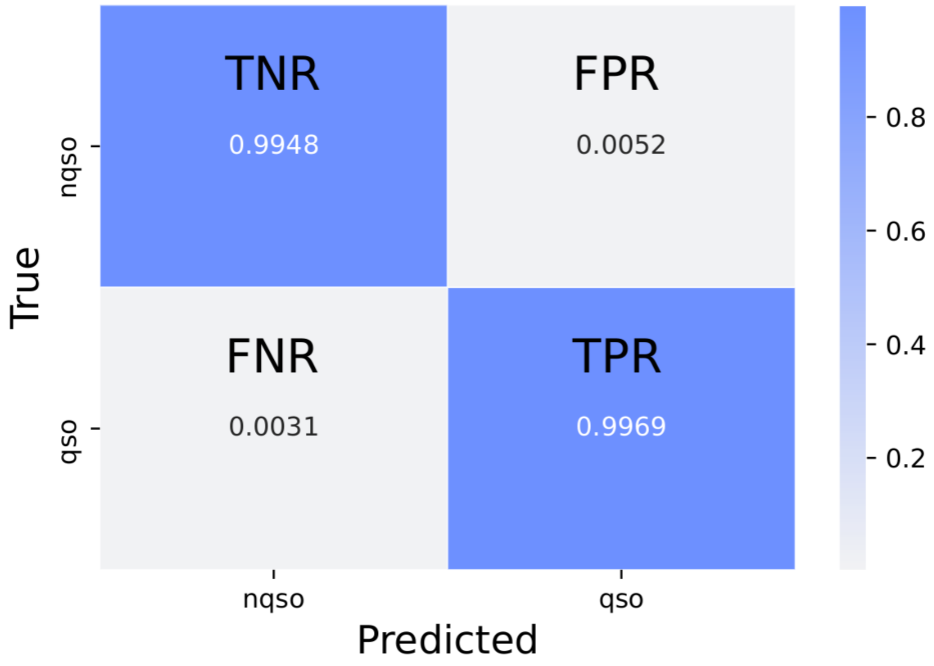
→ **Recover known quasars (5/6) and promising candidates (2/3)** in the survey

→ Without multiwavelength data, we could miss quasars (2/9 ~ 22%)

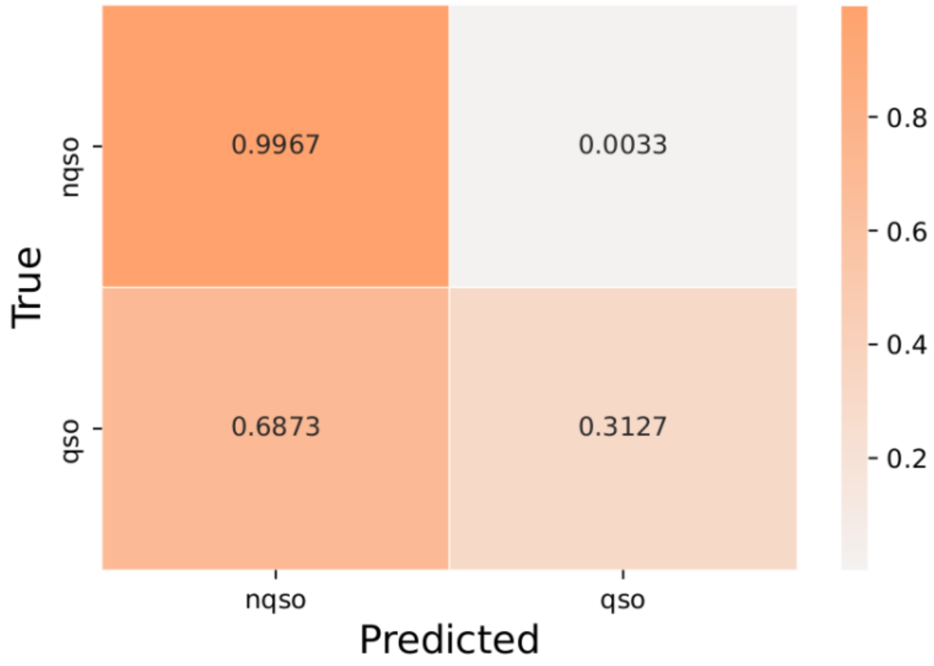
Confusion matrices for two methods

Using quasar models at z=4.5 – 5.5 as a test set

Deep learning



Color selection (Niida+20)



→ Color selection is only effective for finding quasars at z=4.7 – 5.1

→ **Low FPR (~ 0.5%) & Higher TPR (~ 99.7% vs. 31.3 %) of Deep learning**

Summary

Shin, Im & Kim, submitted

- **Searching for faint quasars** with $-26 \lesssim M_{1450} \lesssim \mathbf{-22.0}$ (i ~ 24.0 mag) at z ~ 5 over an area of 16 deg^2 using the **optical imaging data only**
- **Adopting deep learning technique** to efficiently select quasar candidates from non-quasar objects
- **Performing SED fitting** using quasar and star SED models
- **Comparing the fitting results via Bayesian information criterion (BIC)** calculation to select quasar-like objects
- **Building the quasar LF** with 5 confirmed quasars and 30 promising candidates and obtaining $\boldsymbol{\alpha = -1.61^{+0.21}_{-0.19}}$
- Our selection can **minimize galaxy contamination** and also **maximize the quasar recovery compared to conventional color-selection.**