EXPLORING MULTI-BAND IMAGING FOR IDENTIFYING Z > 6.5 QUASARS

A Contrastive Learning Approach Using HSC Data

Laura N. Martínez-Ramírez[1,2,3*], Julien Wolf[3], Silvia Belladitta[3], Raphael Hviding[3], Franz Bauer[1,2], Eduardo Bañados[3]

[1] Pontificia Universidad Católica de Chile
[2] Millennium Institute of Astrophysics
[3] Max-Planck-Institut für Astronomie
*lmartinez@mpia.de

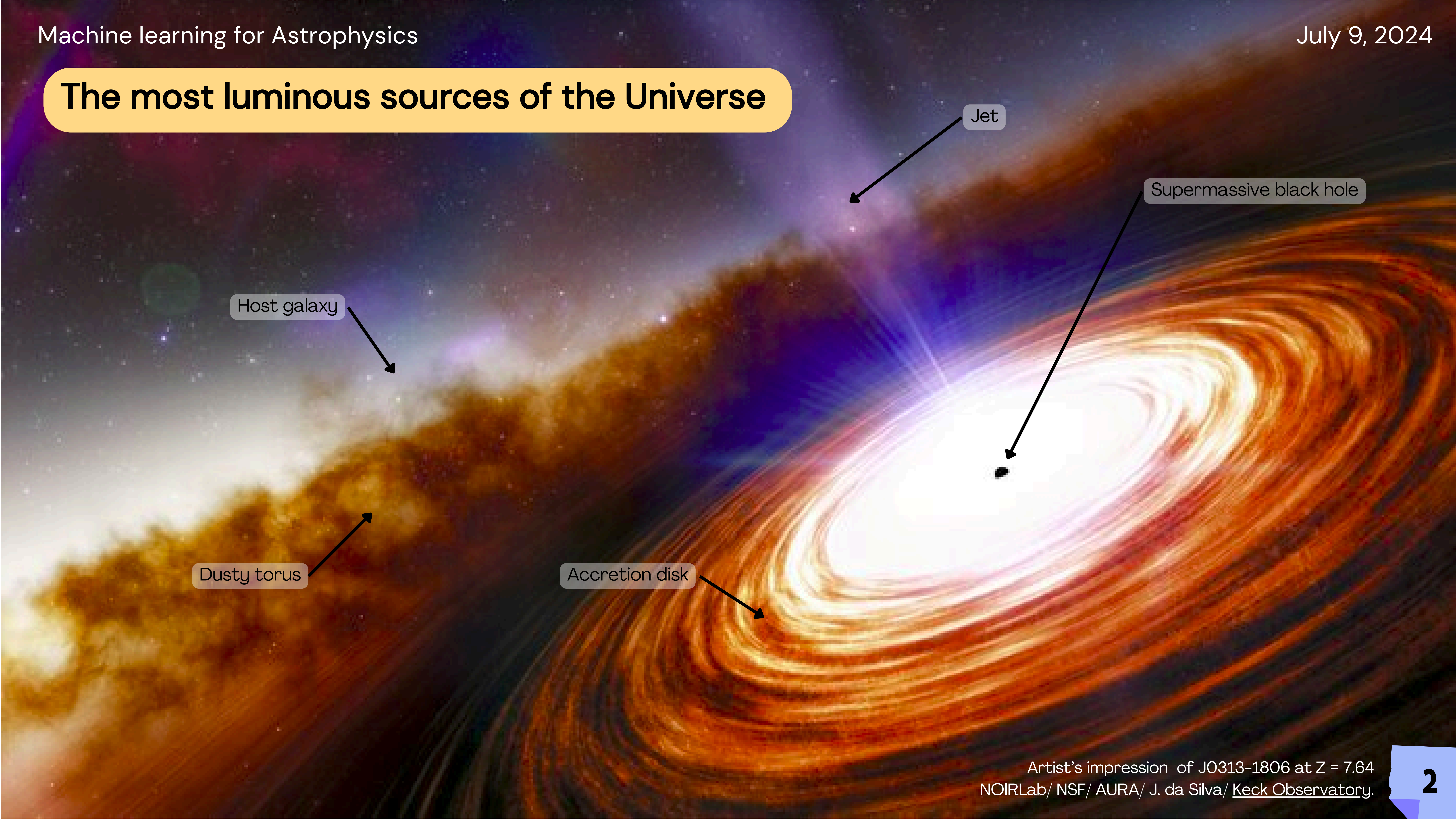
INSTITUTO DE ASTROFÍSICA
FACULTAD DE FÍSICA

INSTITUTO MILENIO DE ASTROFÍSICA

UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386

MACHINE LEARNING FOR ASTROPHYSICS
2ND EDITION CATANIA, 8-12 JULY, 2024

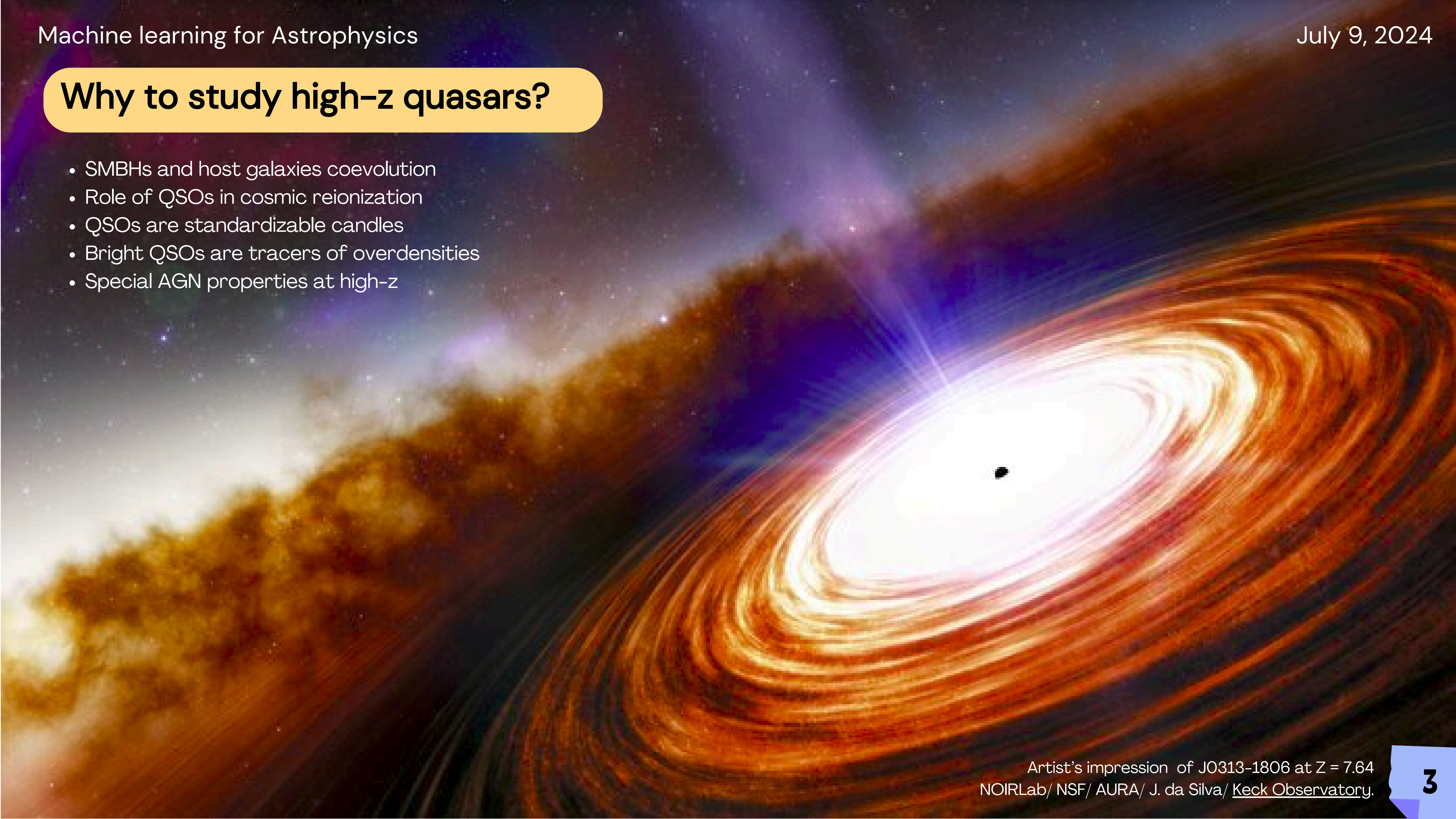
The most luminous sources of the Universe



Artist's impression of J0313-1806 at Z = 7.64
NOIRLab/ NSF/ AURA/ J. da Silva/ Keck Observatory.

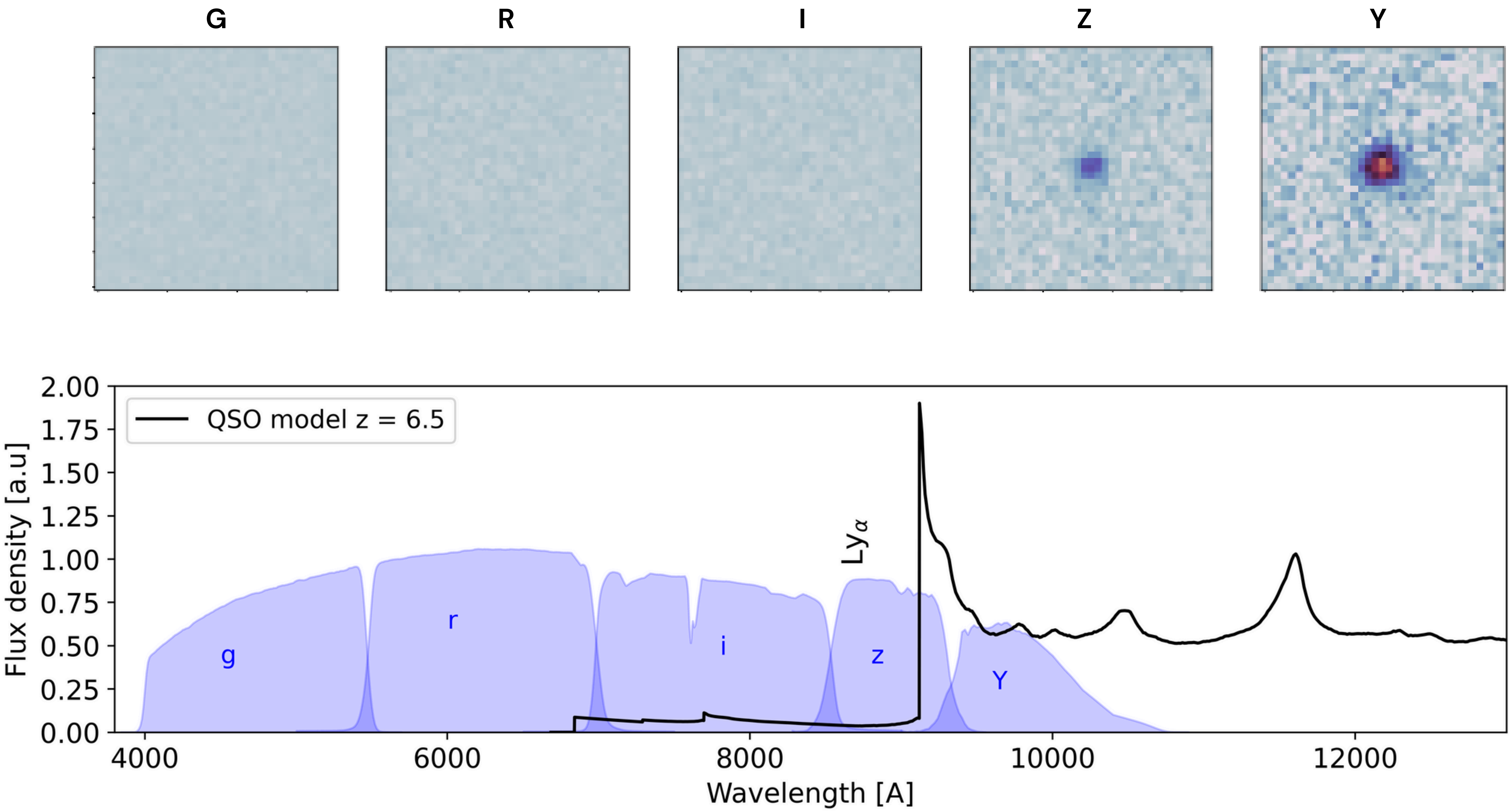
Why to study high-z quasars?

- SMBHs and host galaxies coevolution
- Role of QSOs in cosmic reionization
- QSOs are standardizable candles
- Bright QSOs are tracers of overdensities
- Special AGN properties at high-z

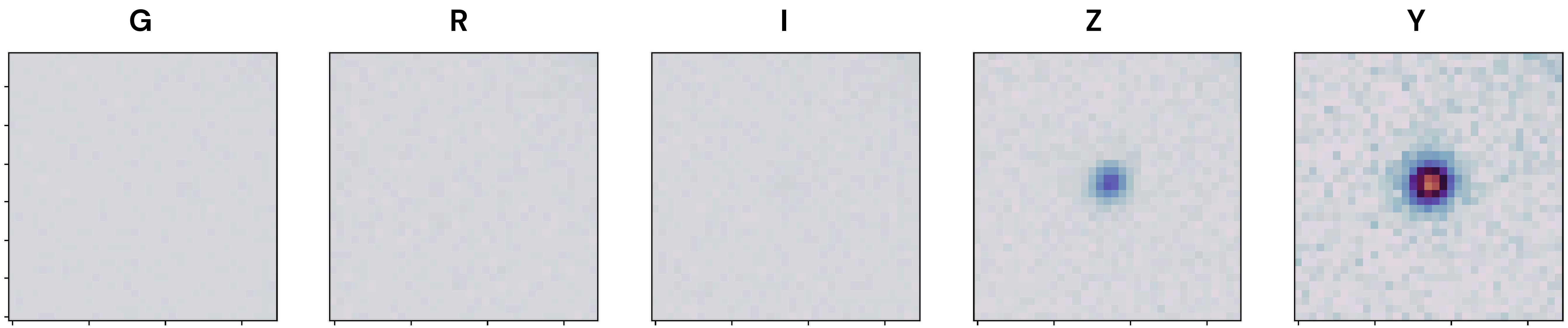


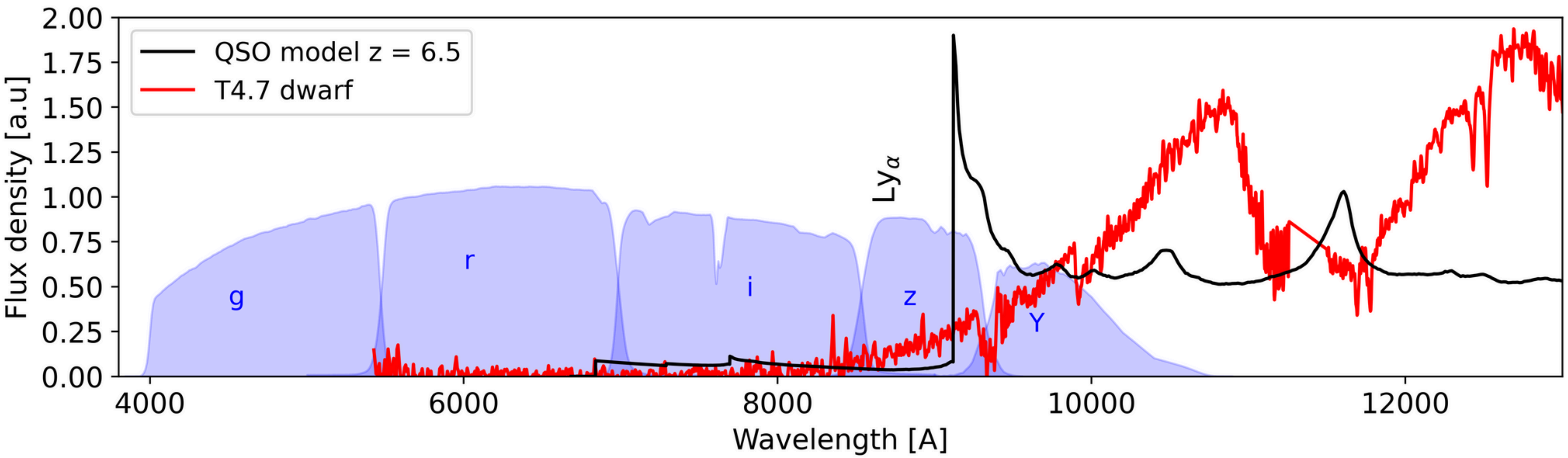
Artist's impression of J0313-1806 at Z = 7.64
NOIRLab/ NSF/ AURA/ J. da Silva/ Keck Observatory.

How to find them? Band-dropouts

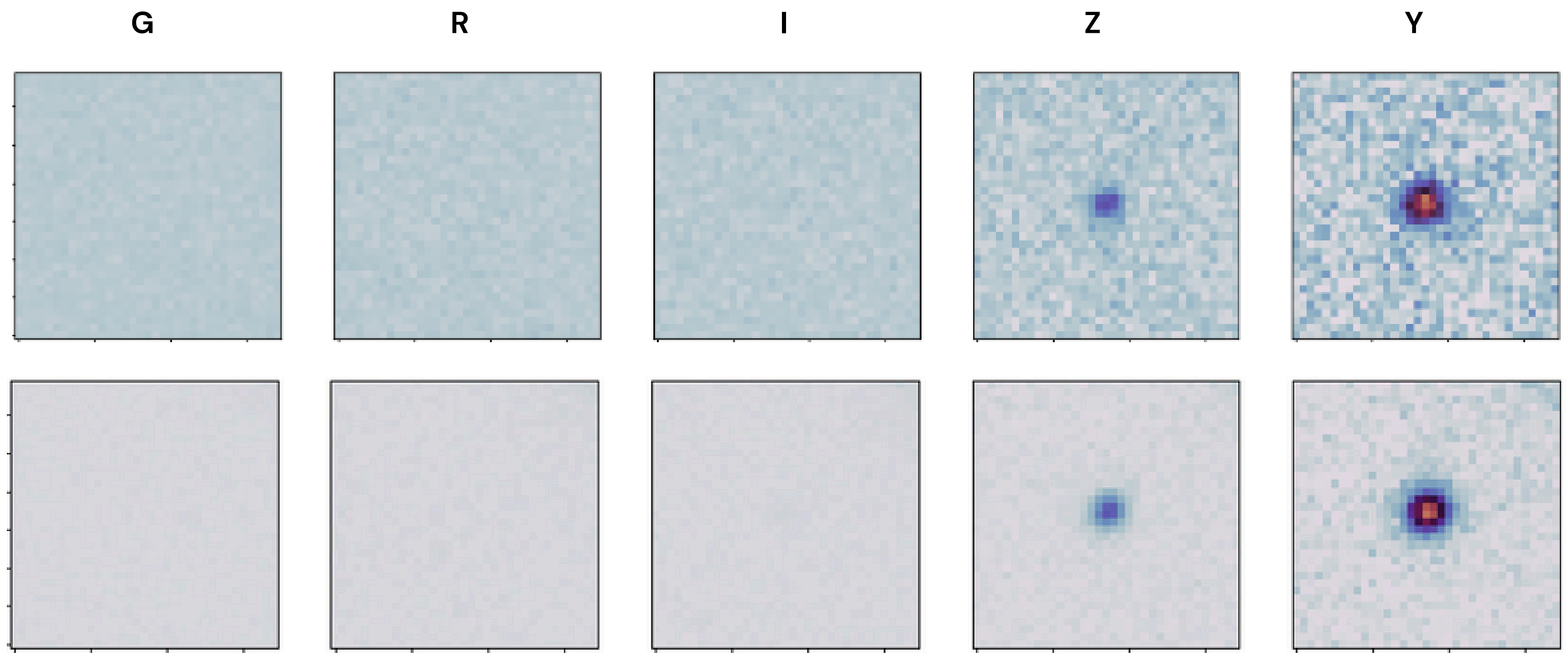


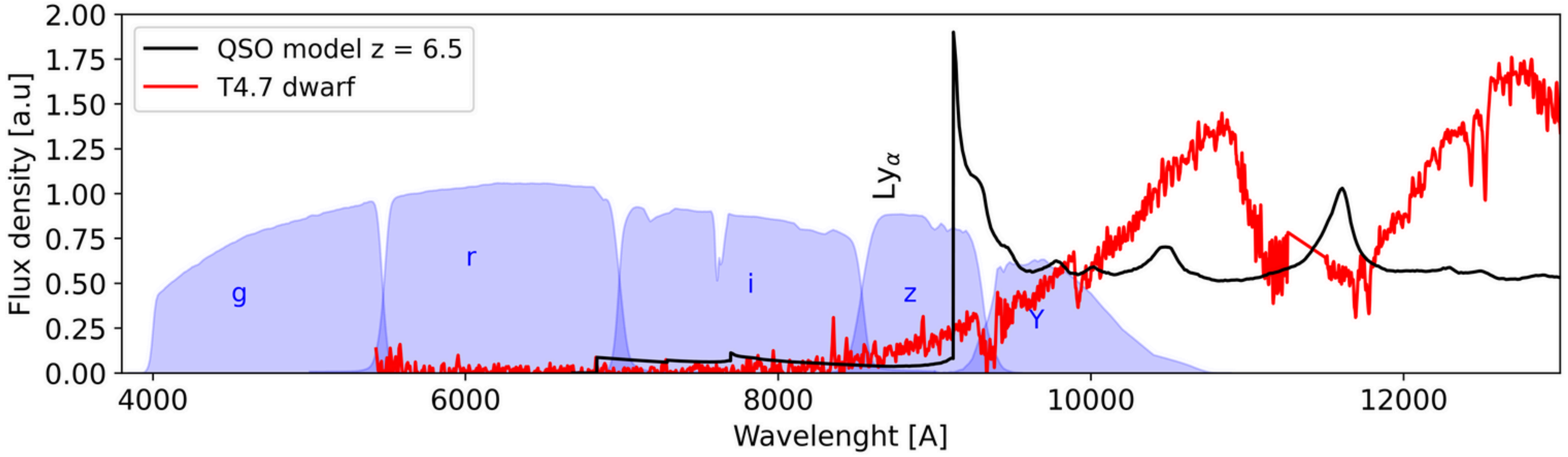
How to find them? Band-dropouts





Finding needles in a haystack

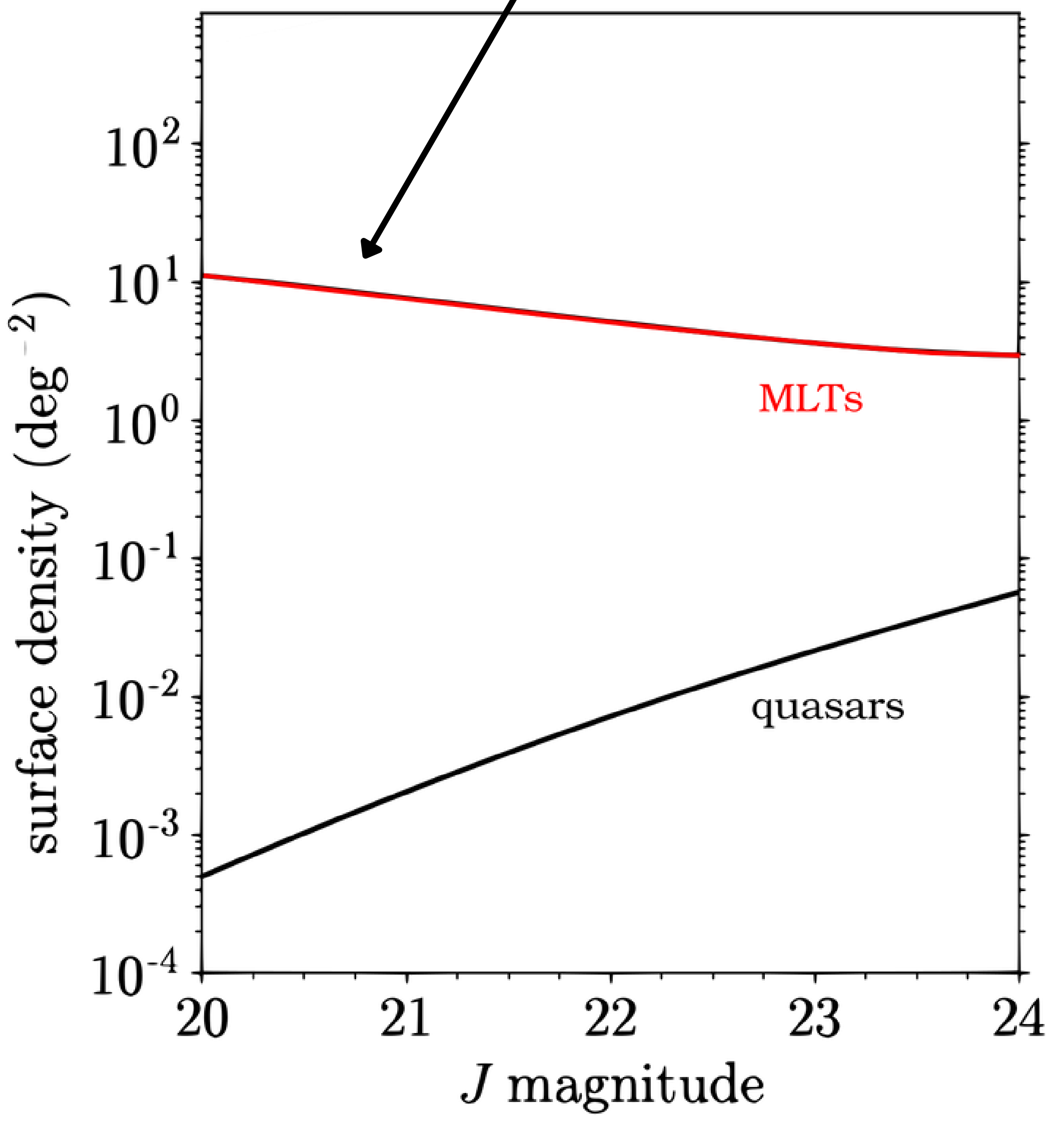




< 1 QSOs per Gpc^3 at z > 6

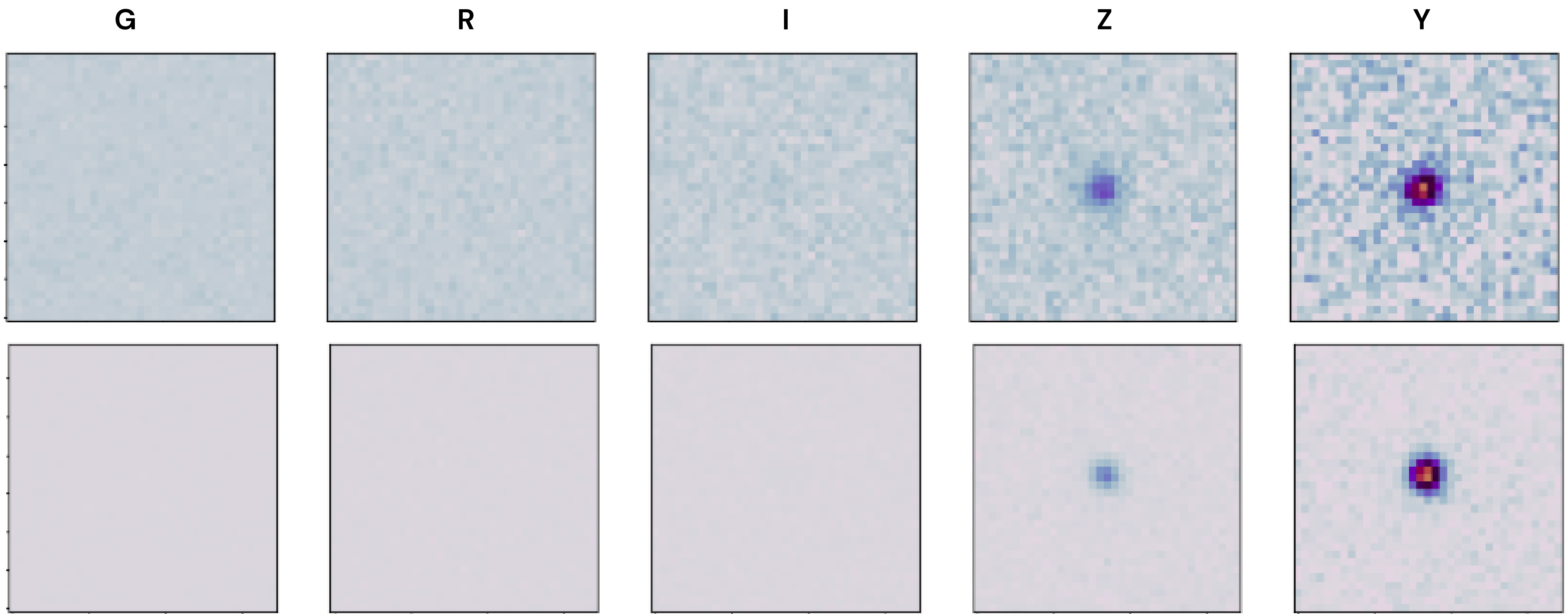
3 QSOs known at z > 7.5

Over 2 orders of magnitude more numerous



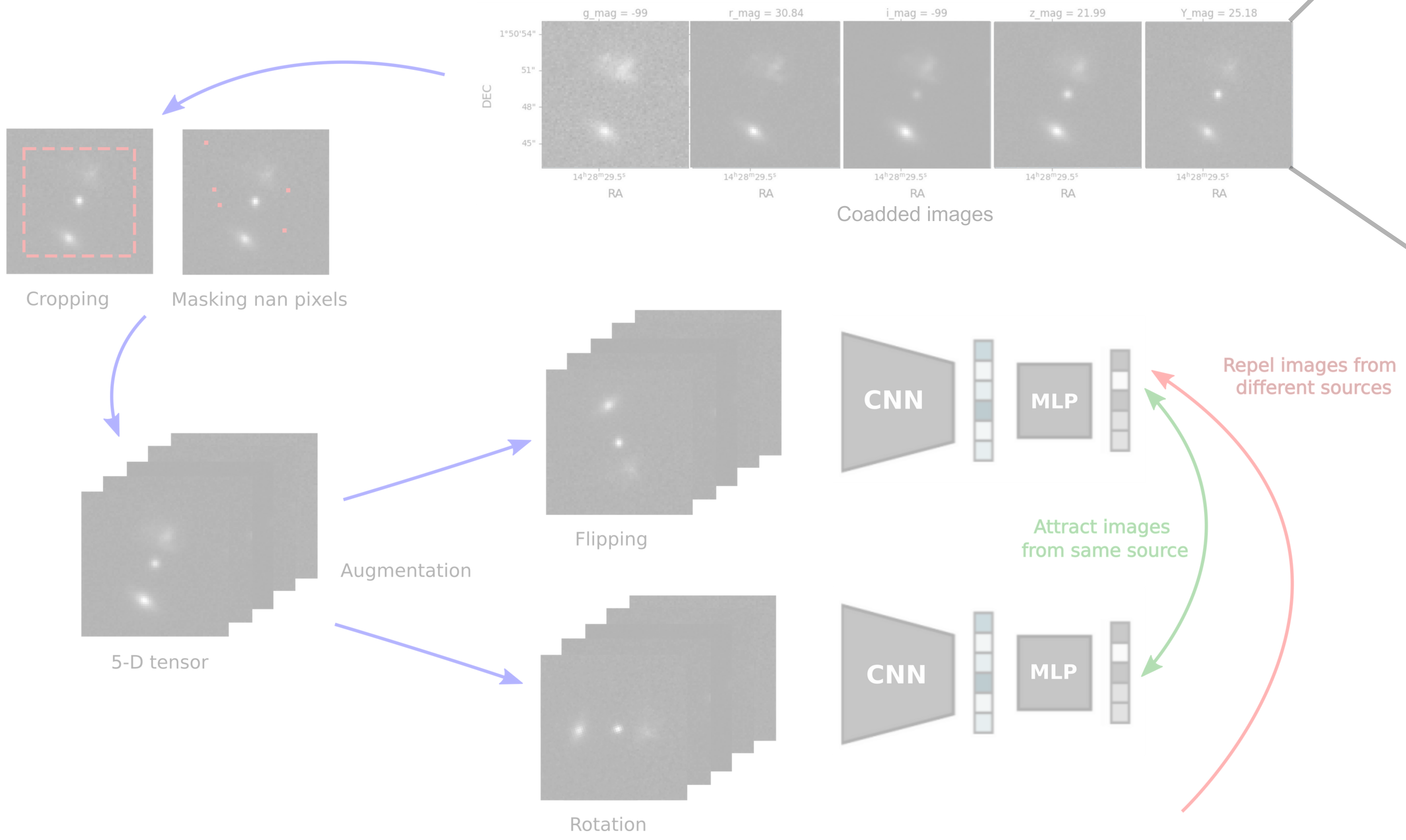
Barnett et al. (2016)

Which one is a z=6.4 quasar?



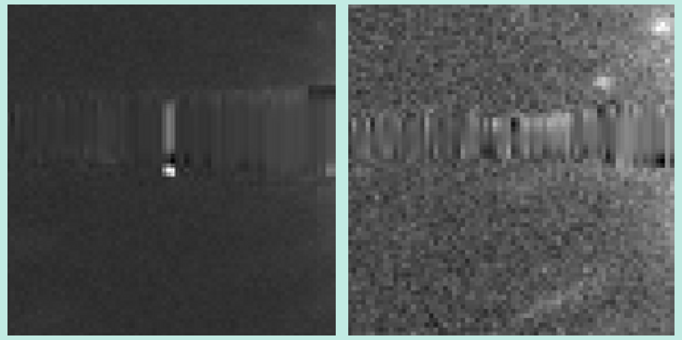
Despite they look similar, we know they are substantially different astrophysical objects.

Contrastive learning with optical data



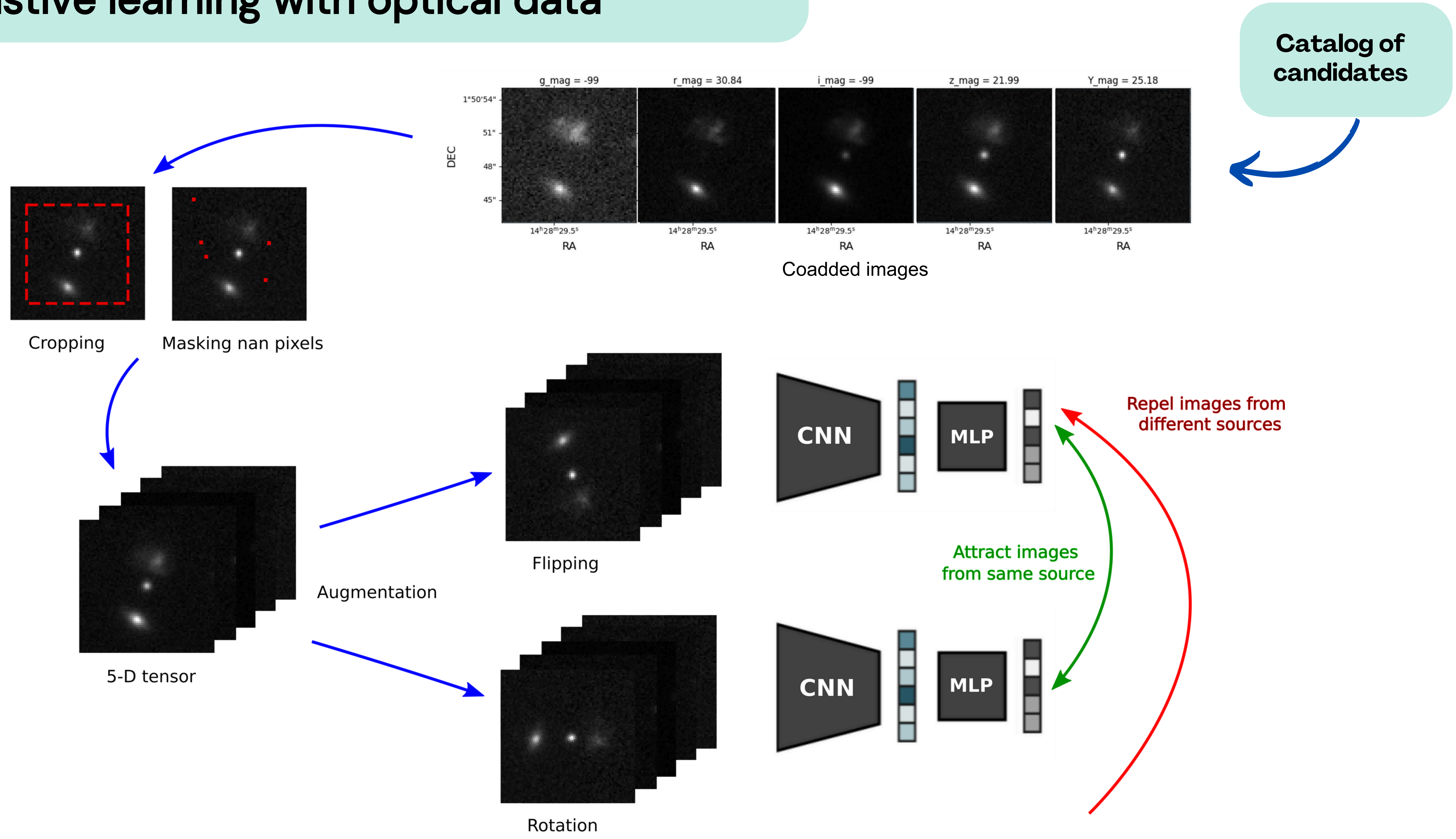
Catalog of candidates

- Dropout in I and z bands (accounting for Ly-α)
- Morphology
- Artifacts removal



Examples of CR and edge flags

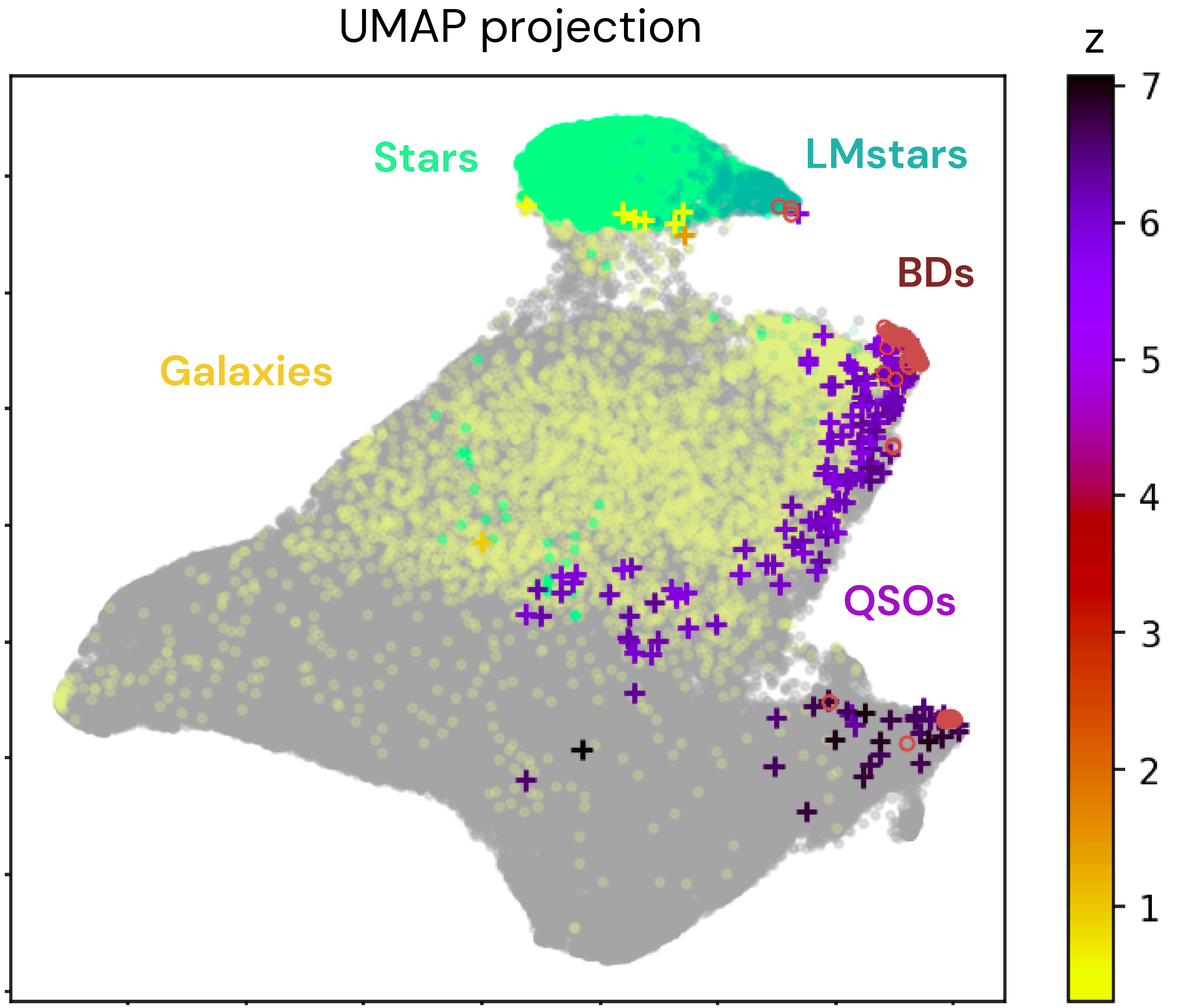
Contrastive learning with optical data



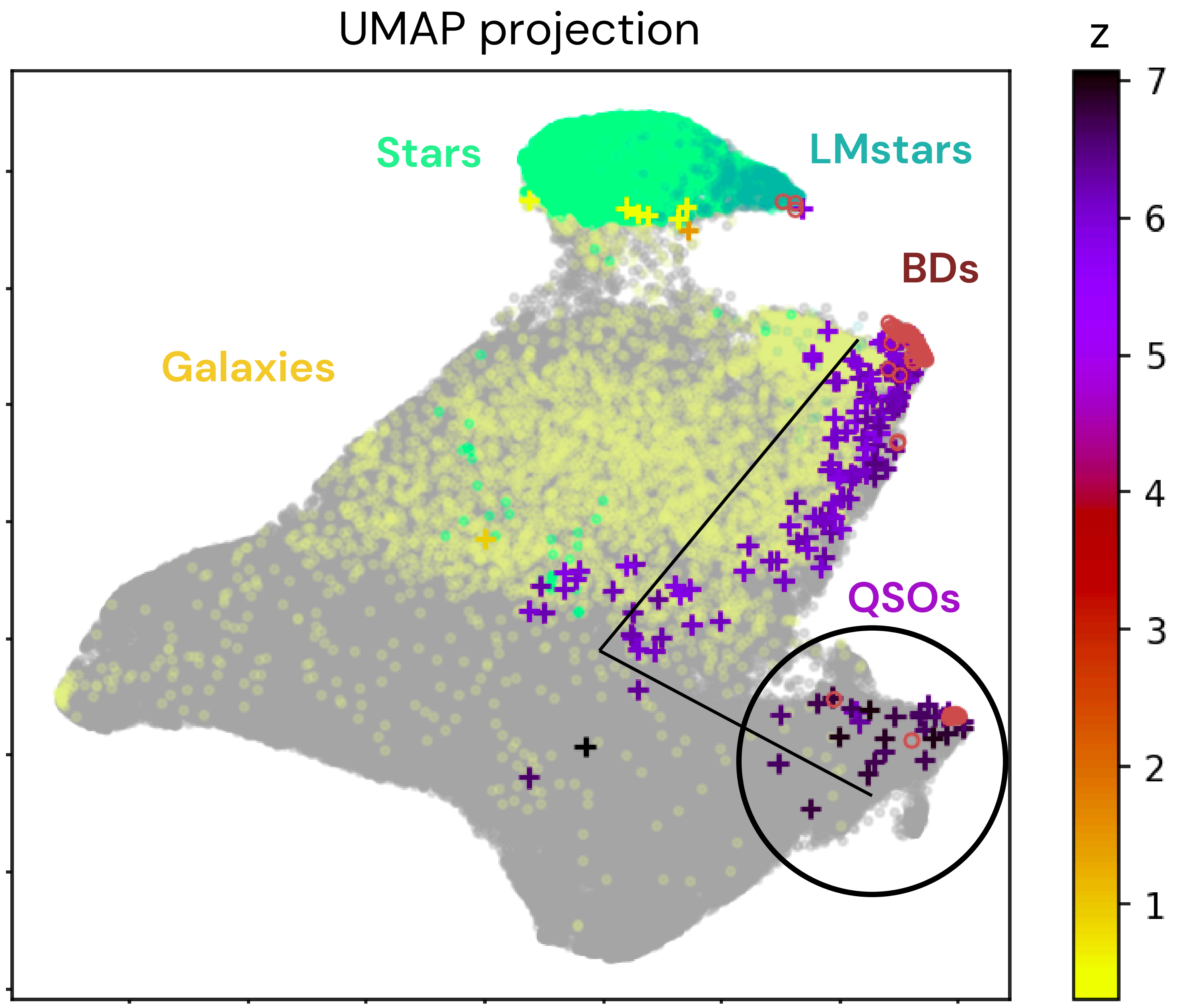
Preliminary results: quasar evolutionary track

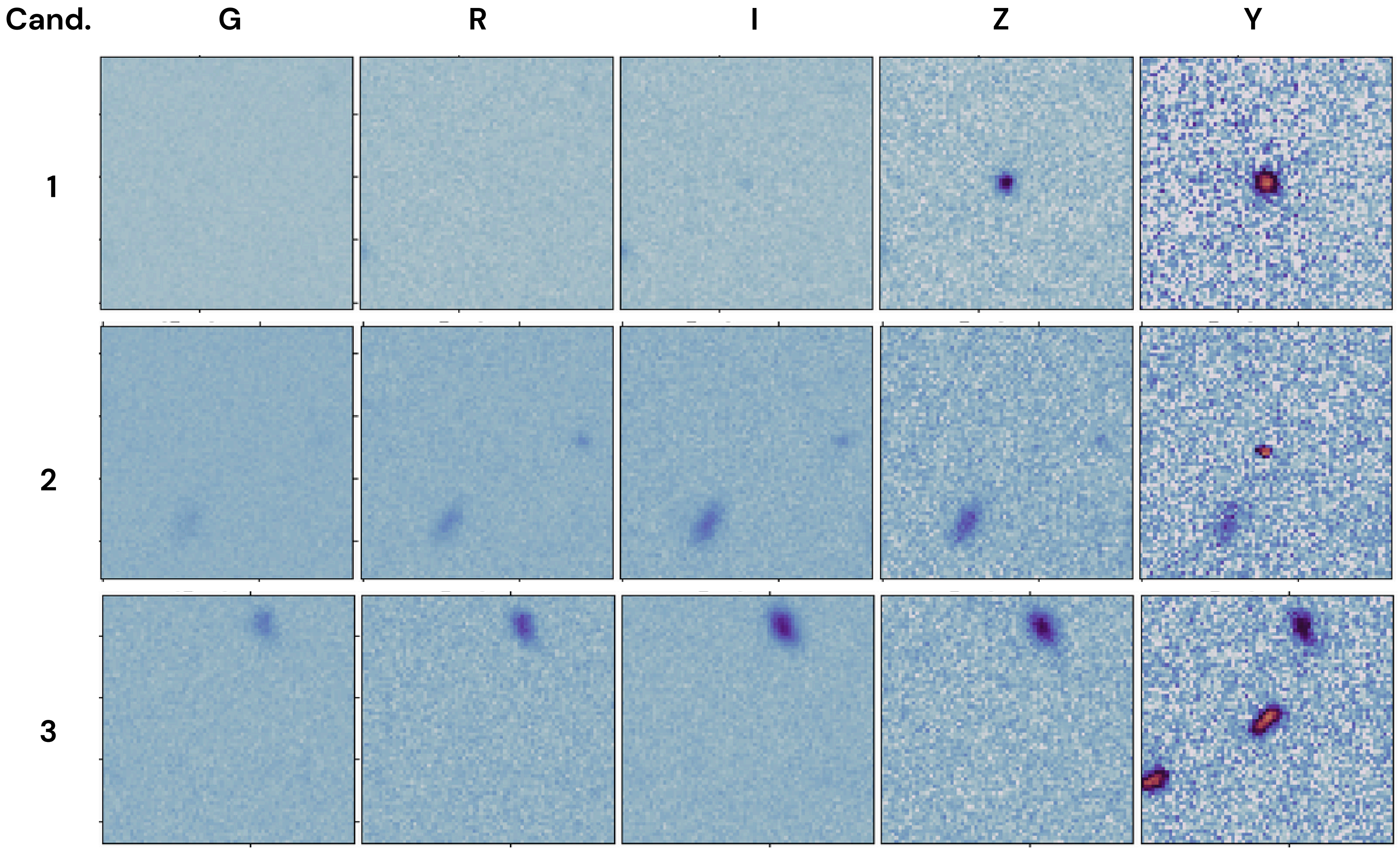
UMAP projection

Preliminary results: quasar evolutionary track



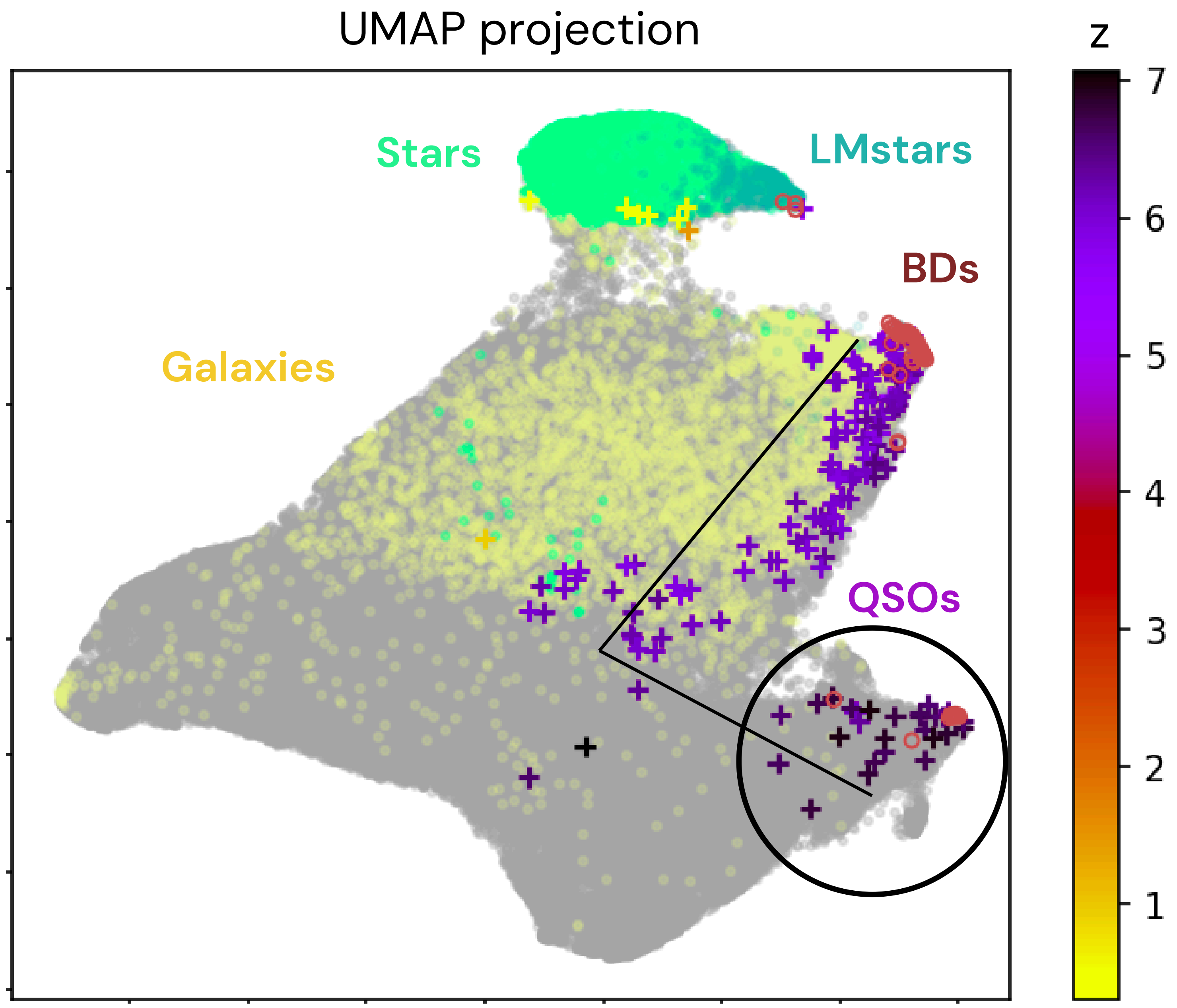
Preliminary results: quasar evolutionary track

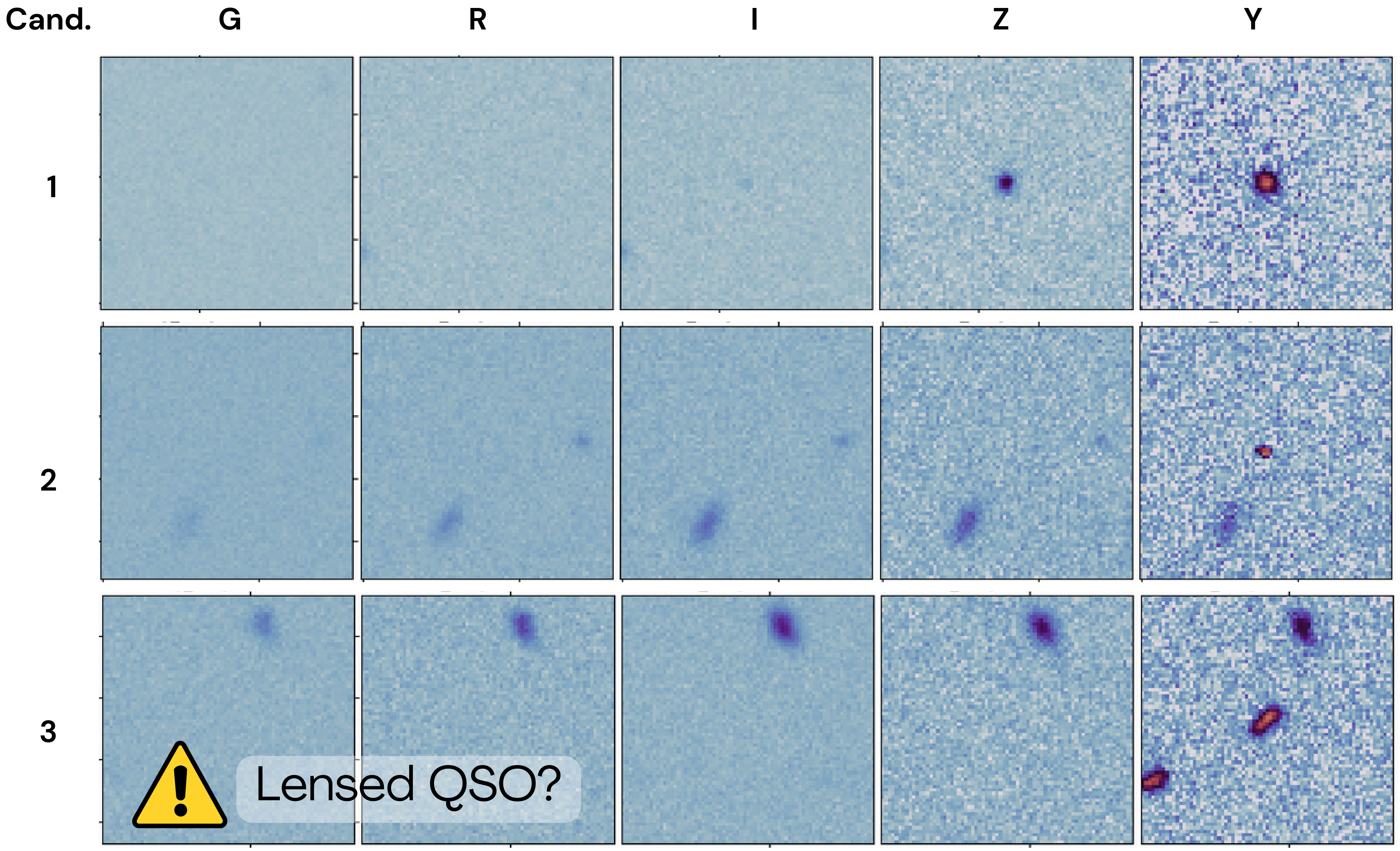




Follow-up observations:
We got telescope time (CNTAC/MAGELLAN and MPIA/LBT)

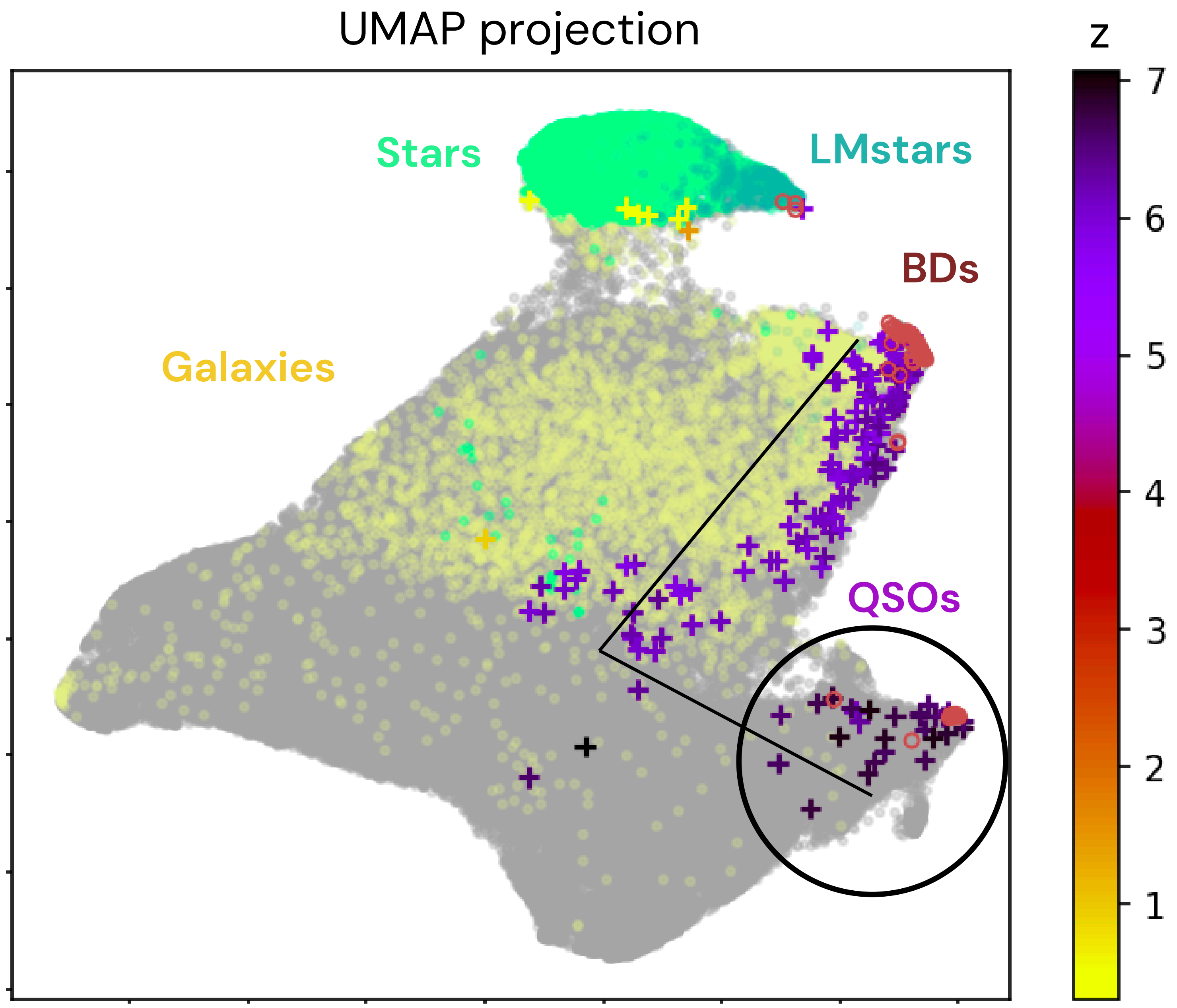
Preliminary results: quasar evolutionary track

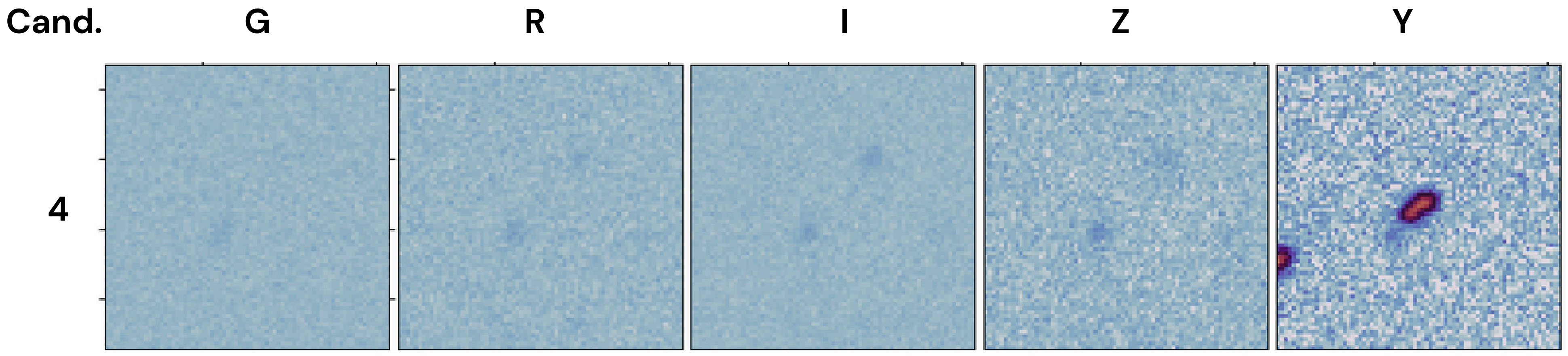




Follow-up observations:
We got telescope time (CNTAC/MAGELLAN and MPIA/LBT)

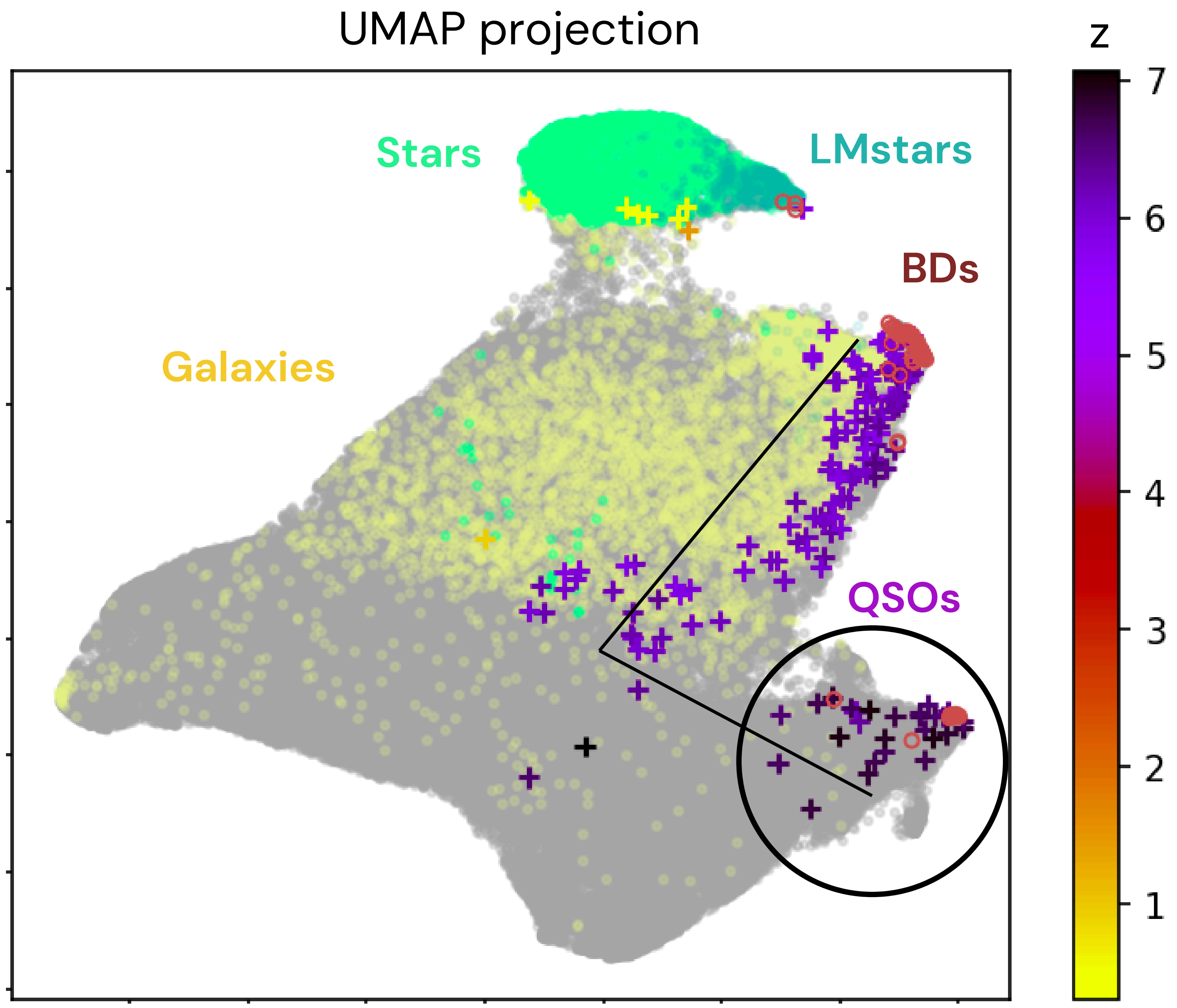
Preliminary results: quasar evolutionary track

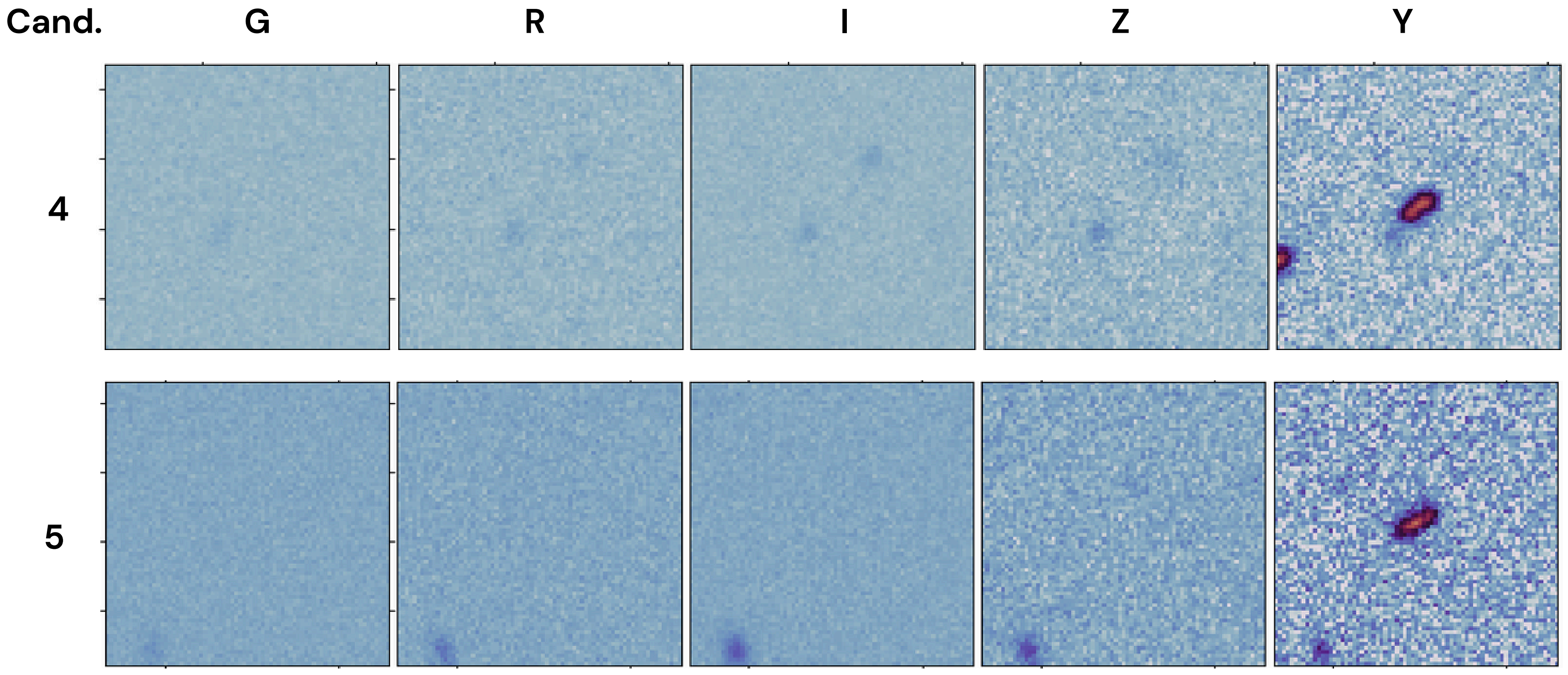




Follow-up observations:
We got telescope time (CNTAC/MAGELLAN and MPIA/LBT)

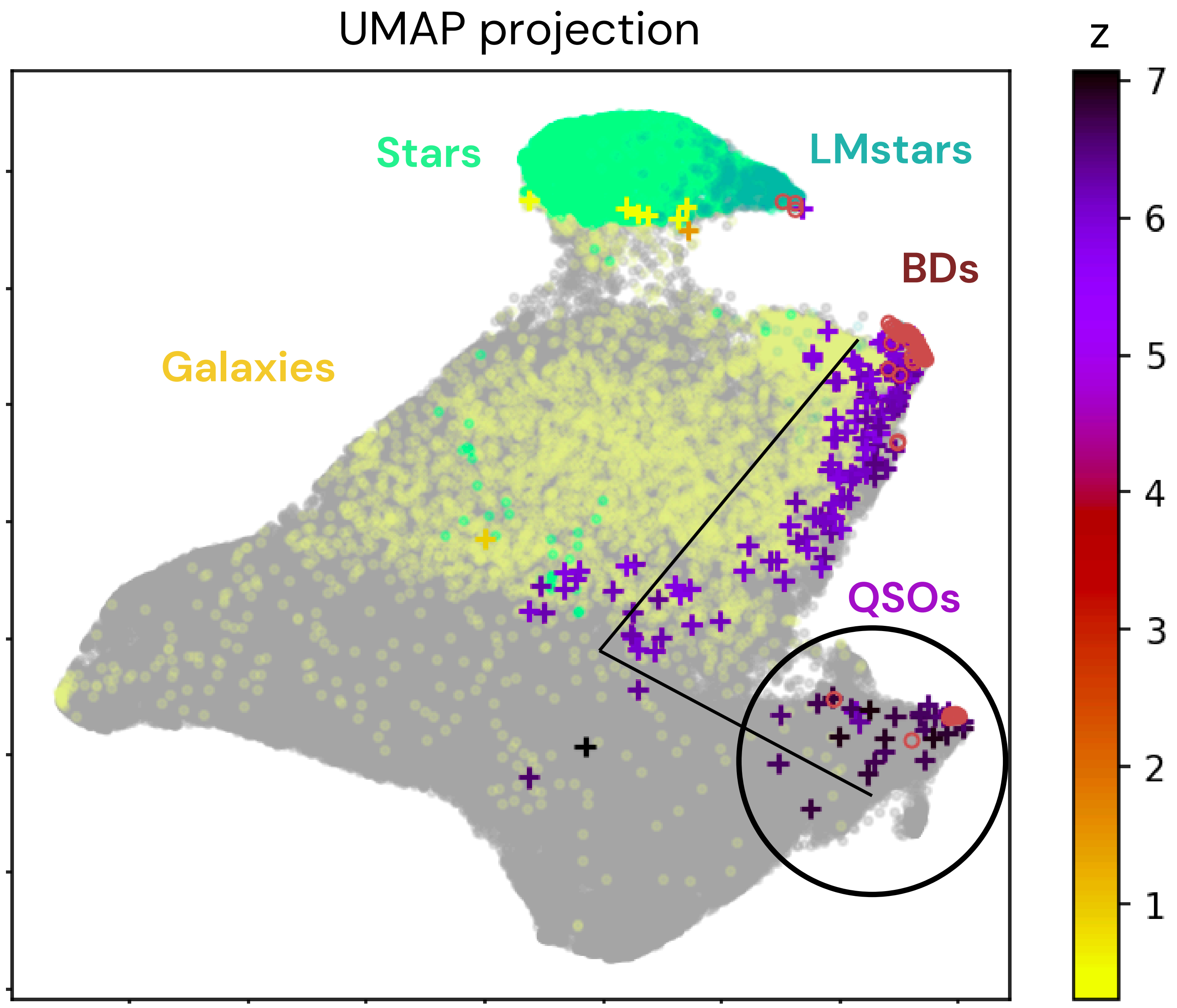
Preliminary results: quasar evolutionary track

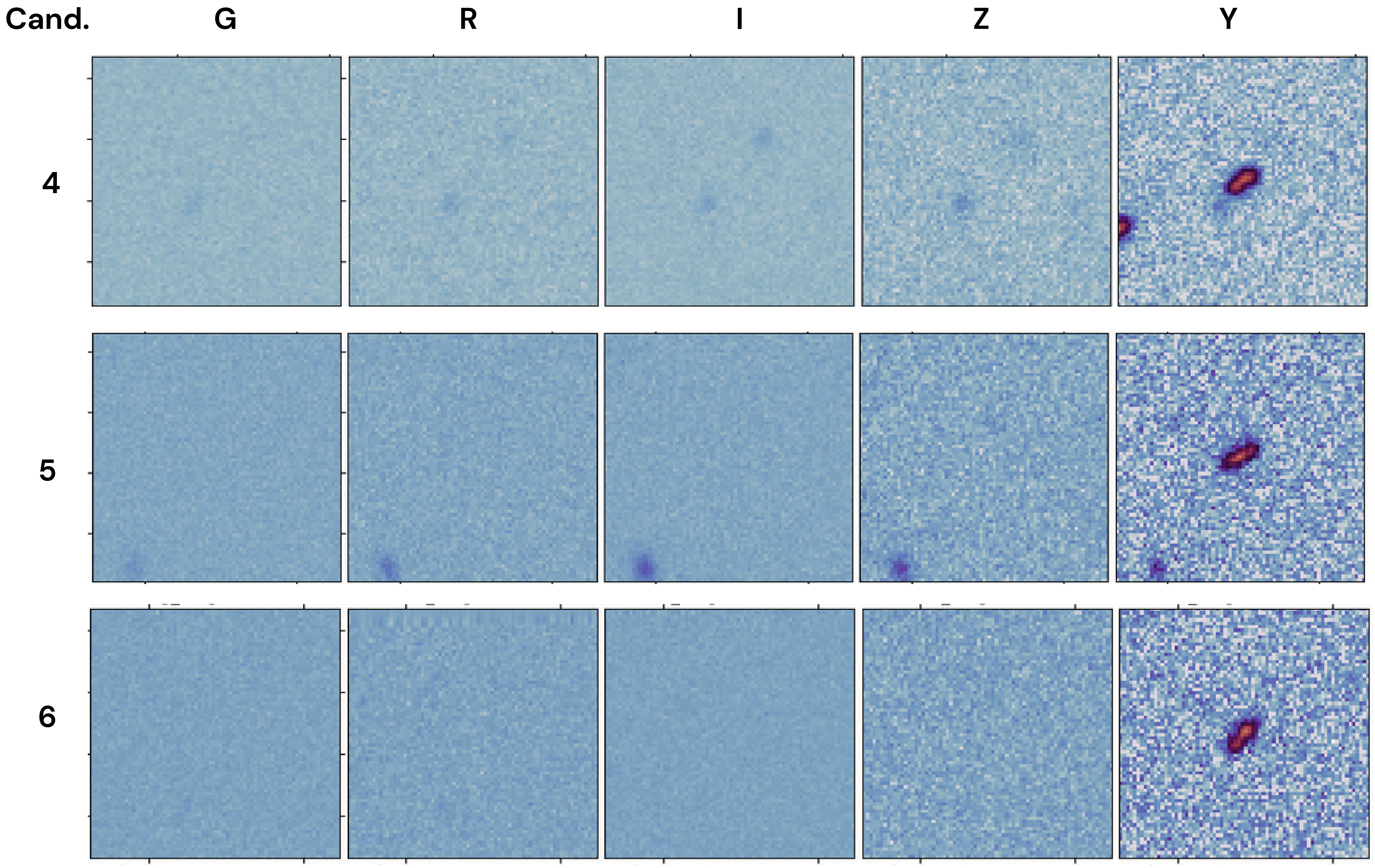




Follow-up observations:
We got telescope time (CNTAC/MAGELLAN and MPIA/LBT)

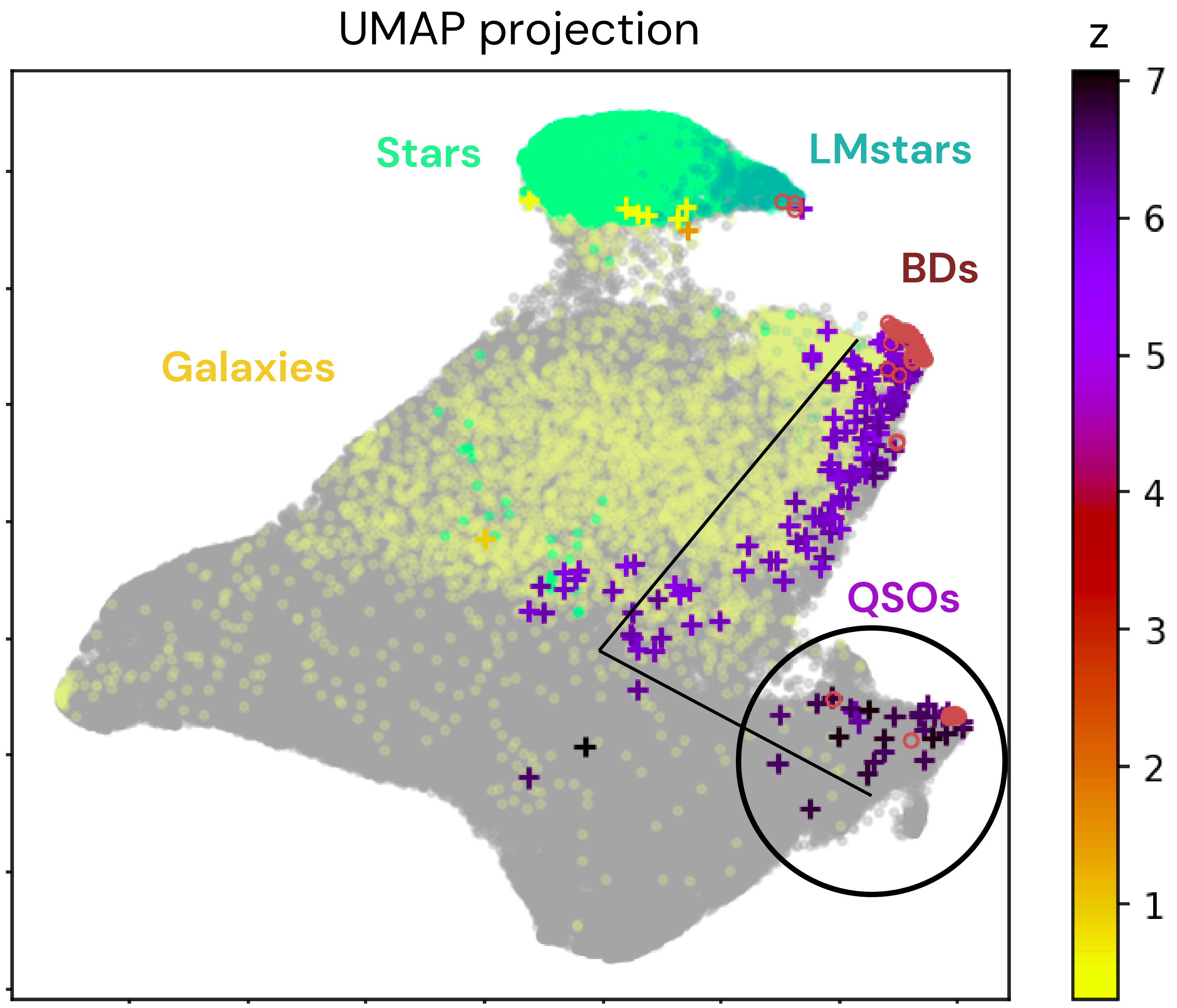
Preliminary results: quasar evolutionary track

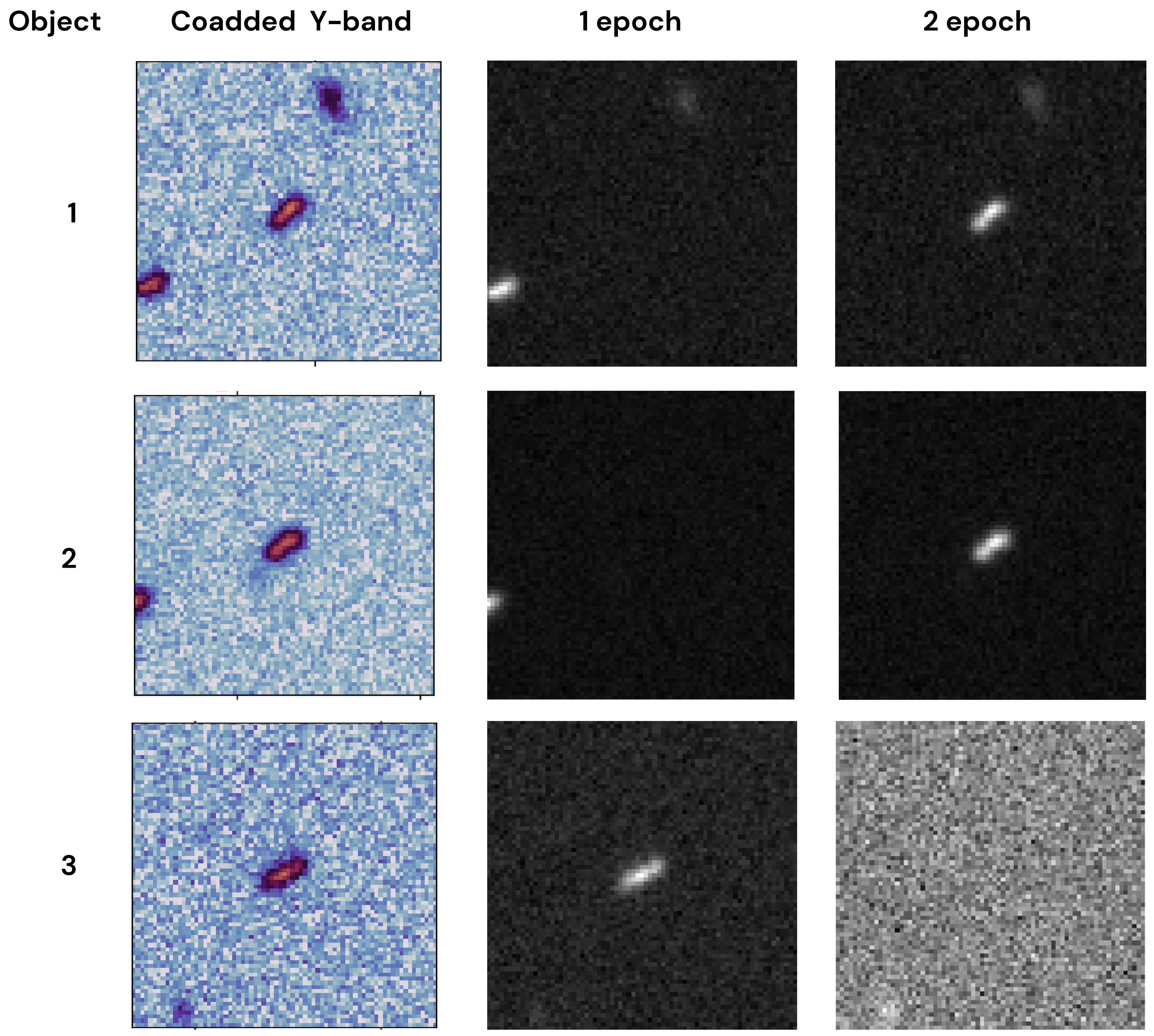




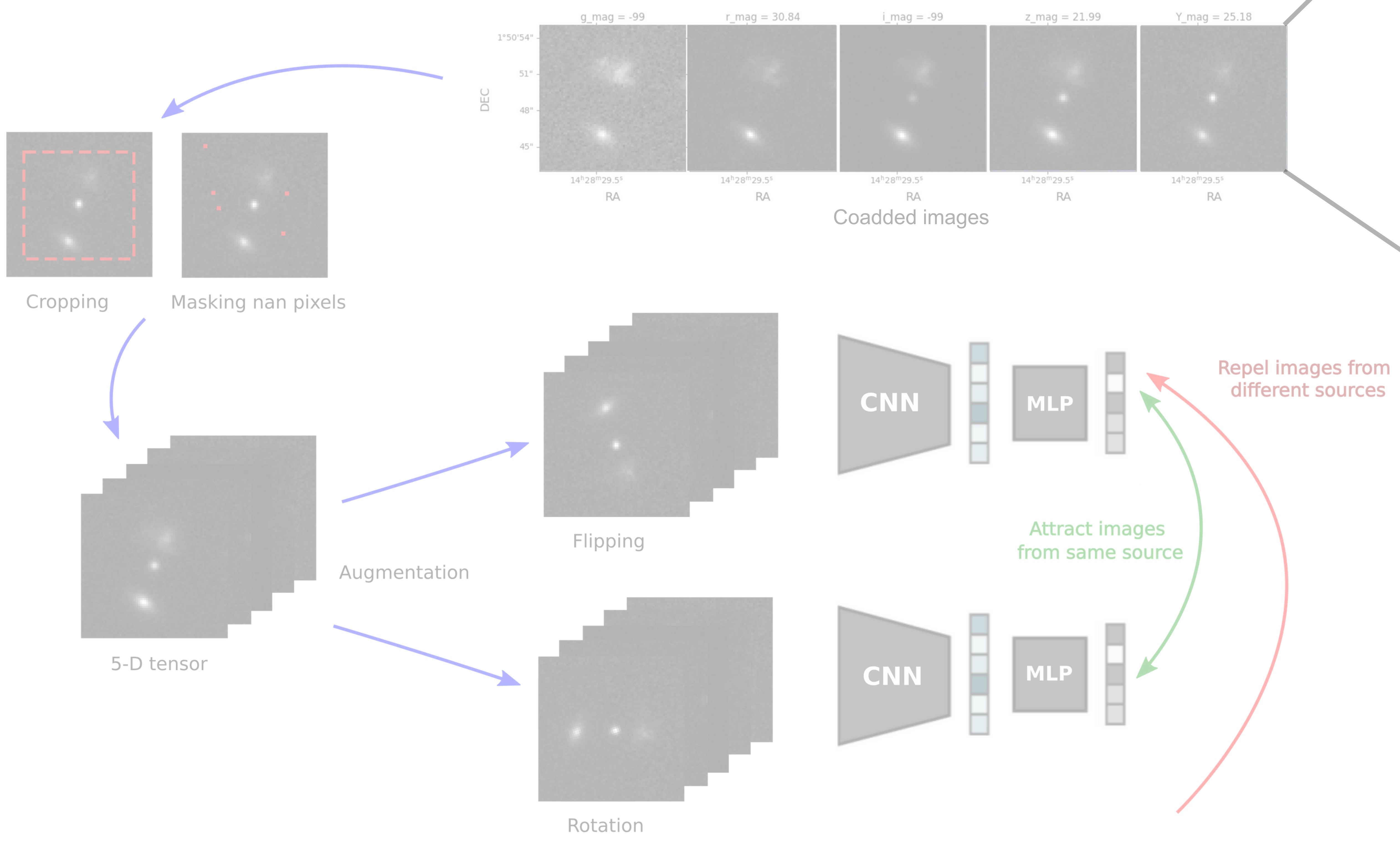
Follow-up observations:
We got telescope time (CNTAC/MAGELLAN and MPIA/LBT)

Preliminary results: moving objects!



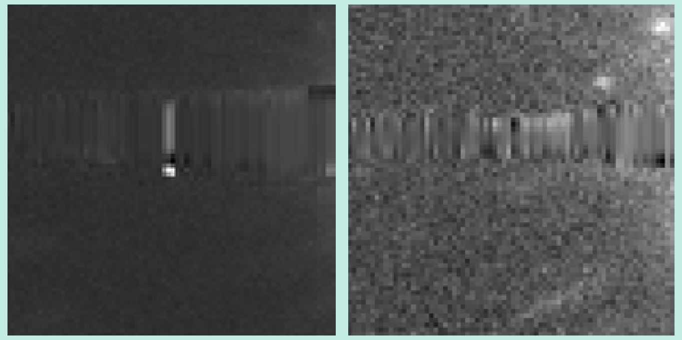


Preliminary results: removing faint sources



Catalog of candidates

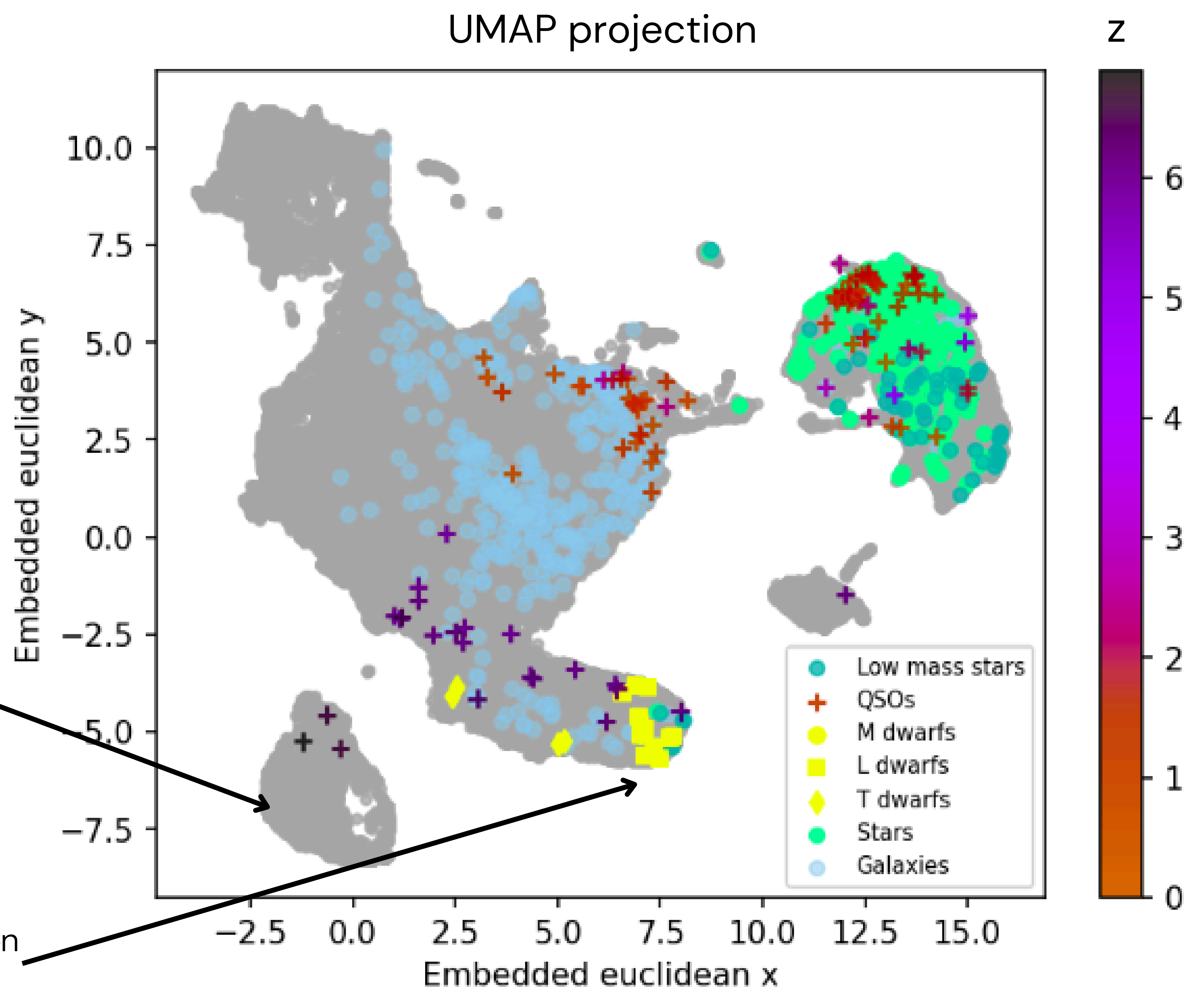
- Dropout in I and z bands (accounting for Ly-α)
- Morphology
- Artifacts removal



Examples of CR and edge flags

Magnitude cut

Preliminary results: removing faint sources

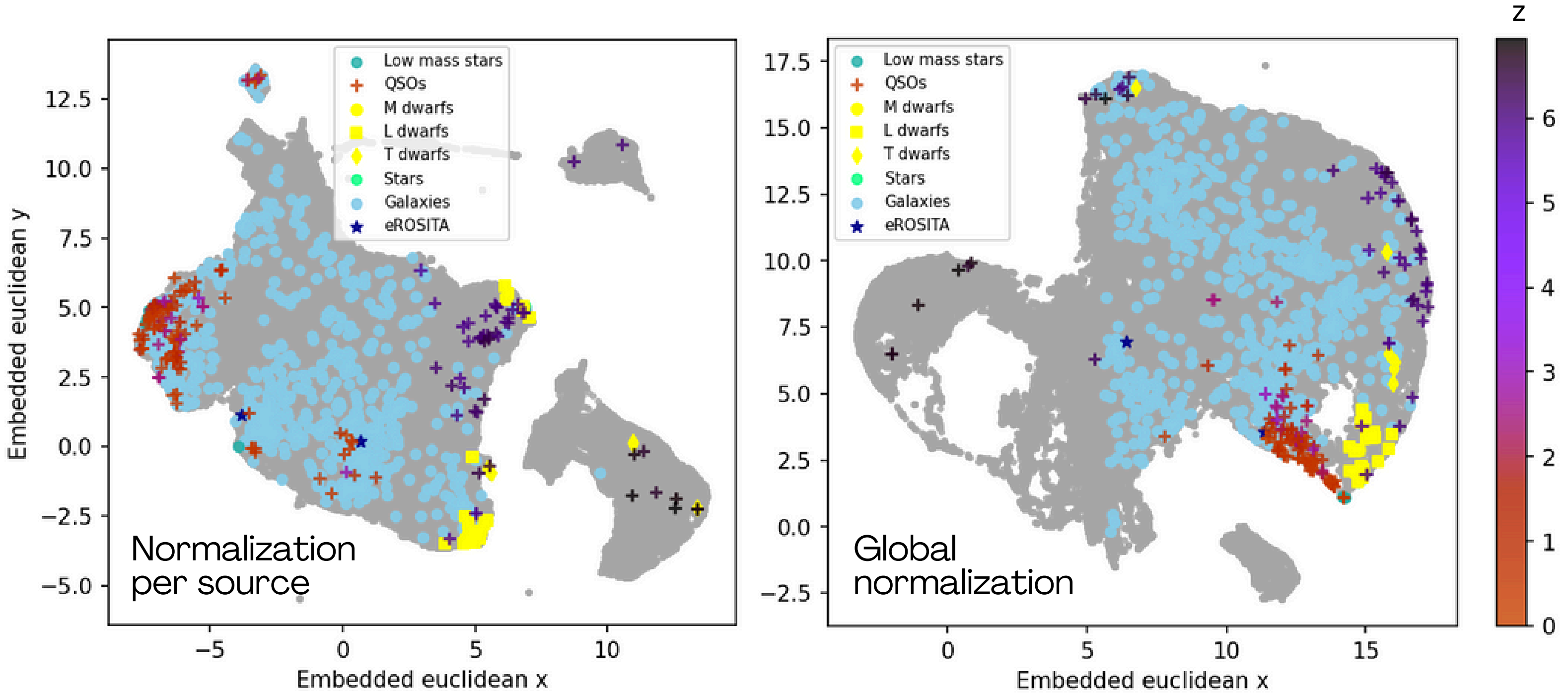


Island of z-dropouts

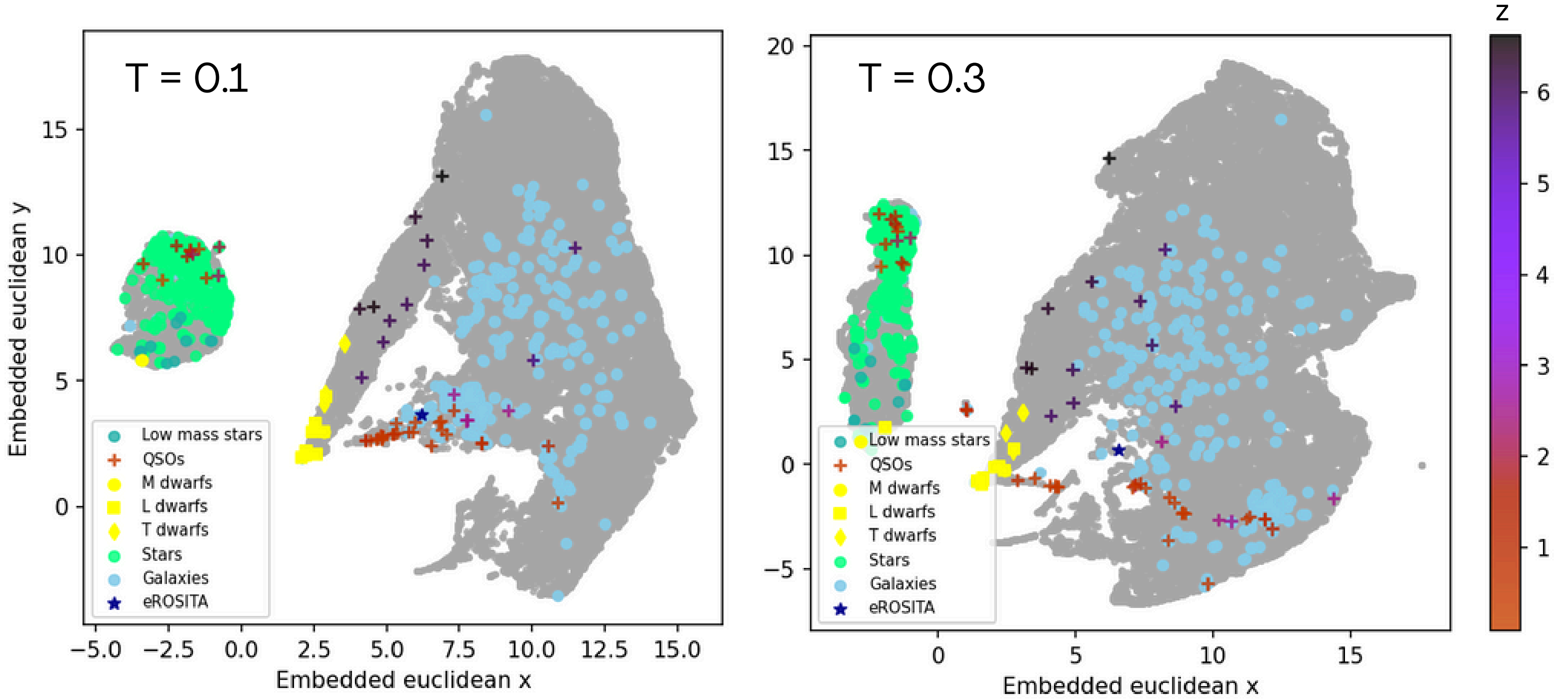
Better isolation of dwarfs

Preliminary results: Parameters fine tunning

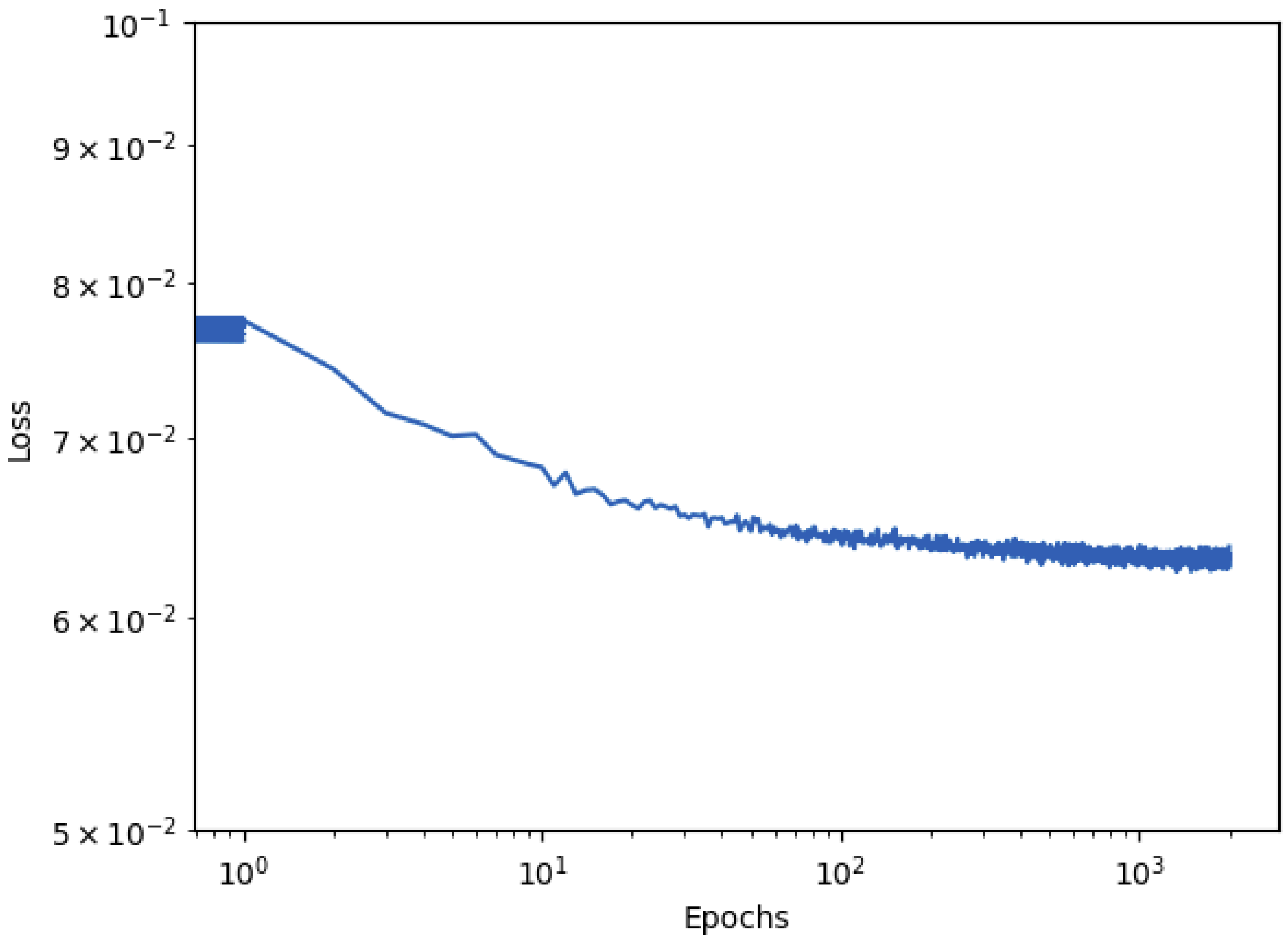
- **Tensor normalization**



- Temperature in contrastive loss function

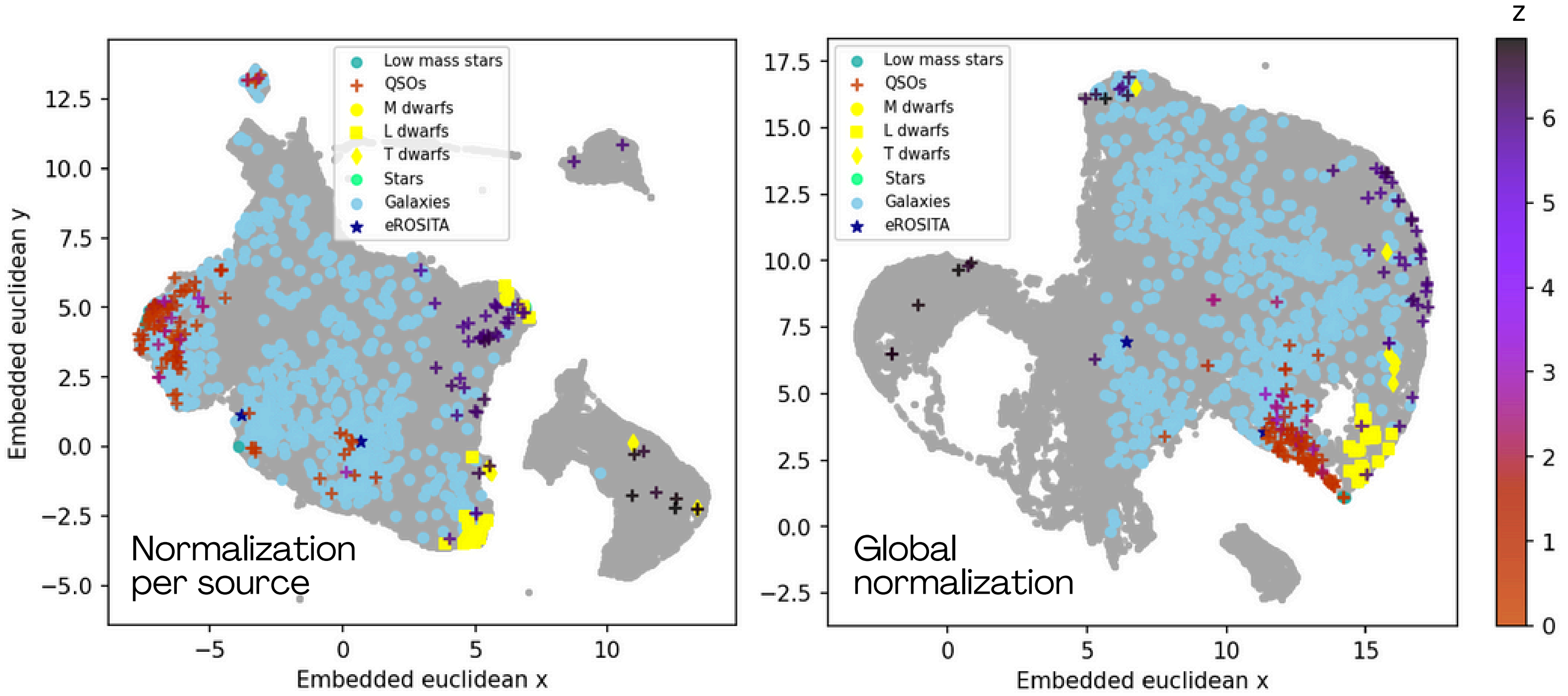


- **Training set**
- Image scales
- **Number of neighbors**
- **Minimum distance**
- Metrics
- Number of epochs

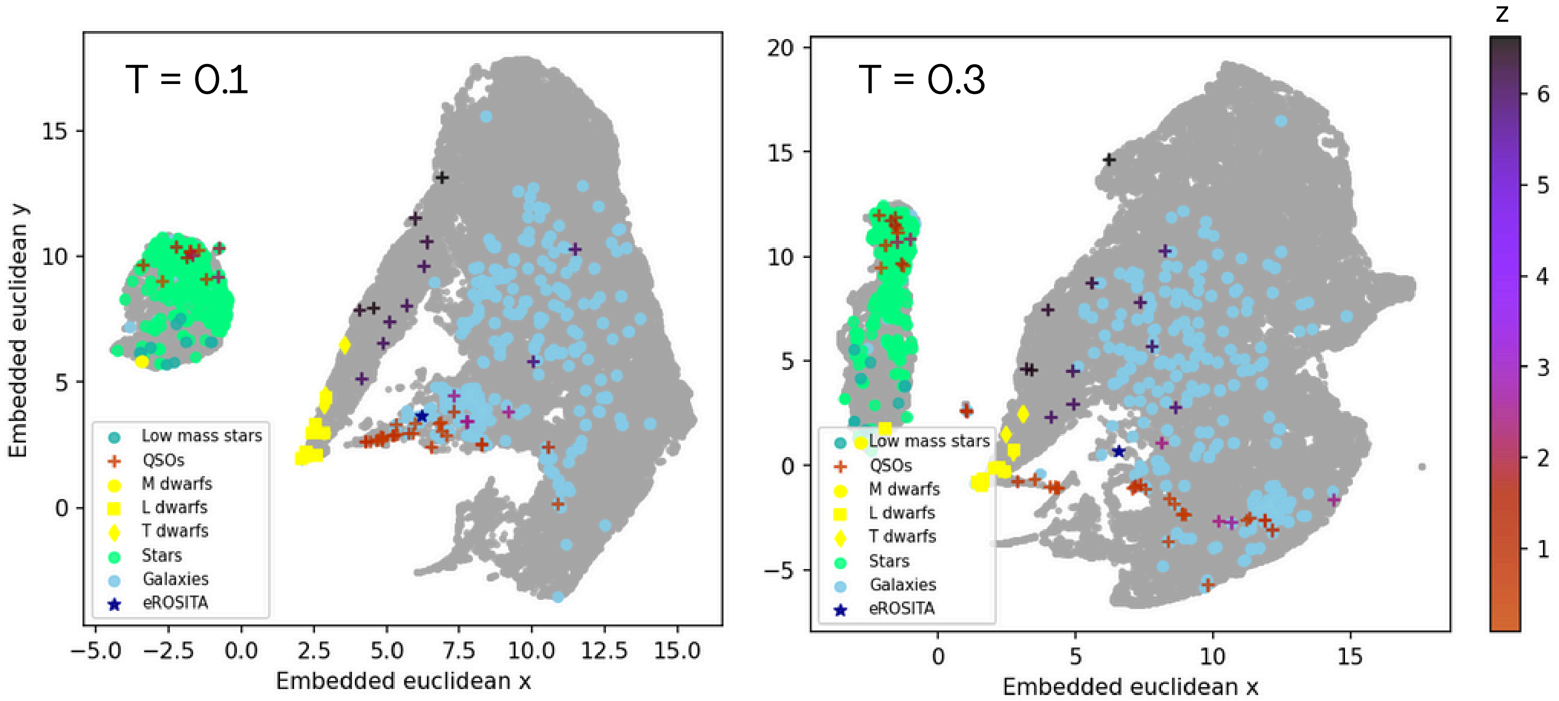


Preliminary results: Parameters fine tunning

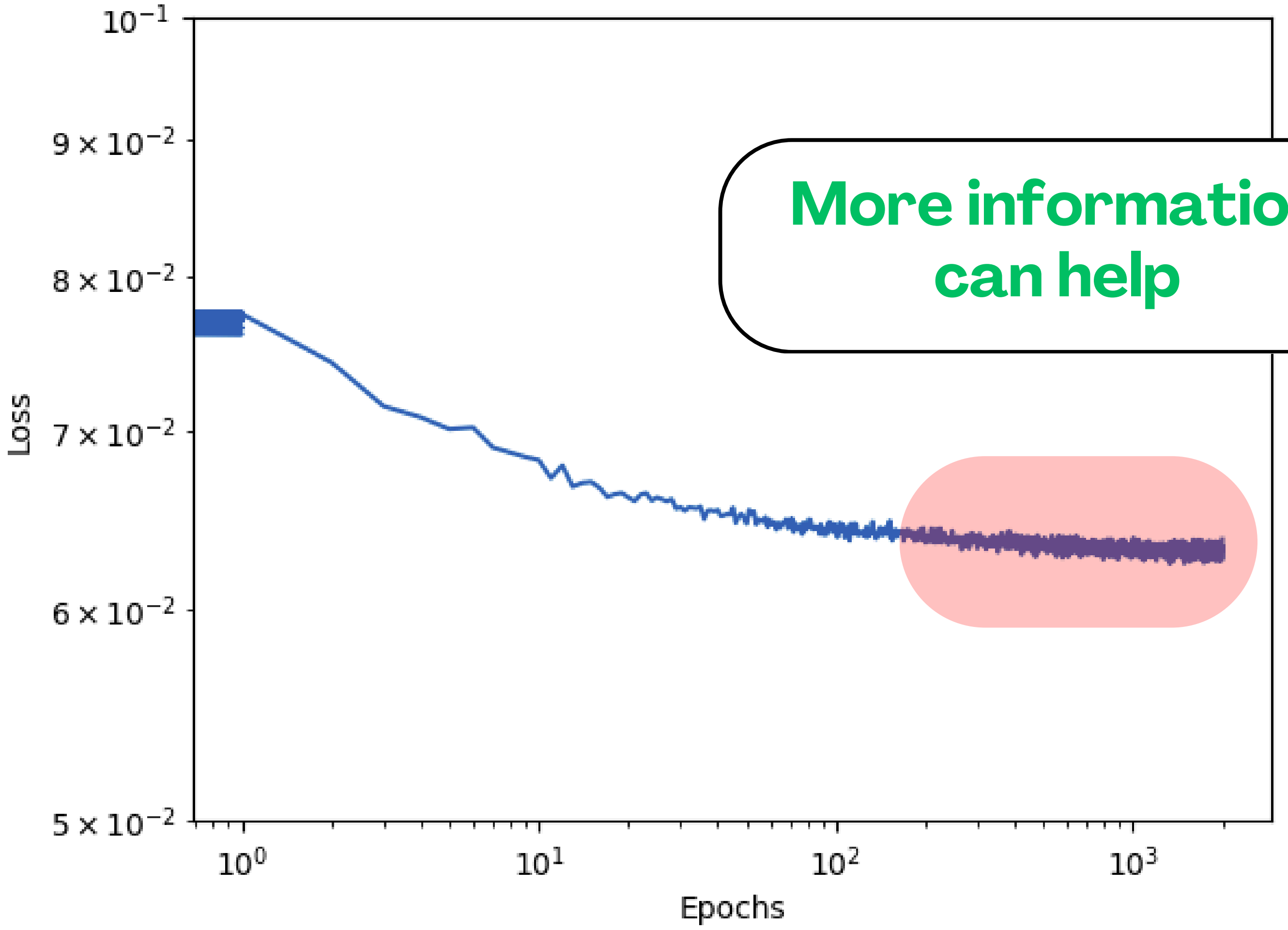
- **Tensor normalization**



- Temperature in contrastive loss function



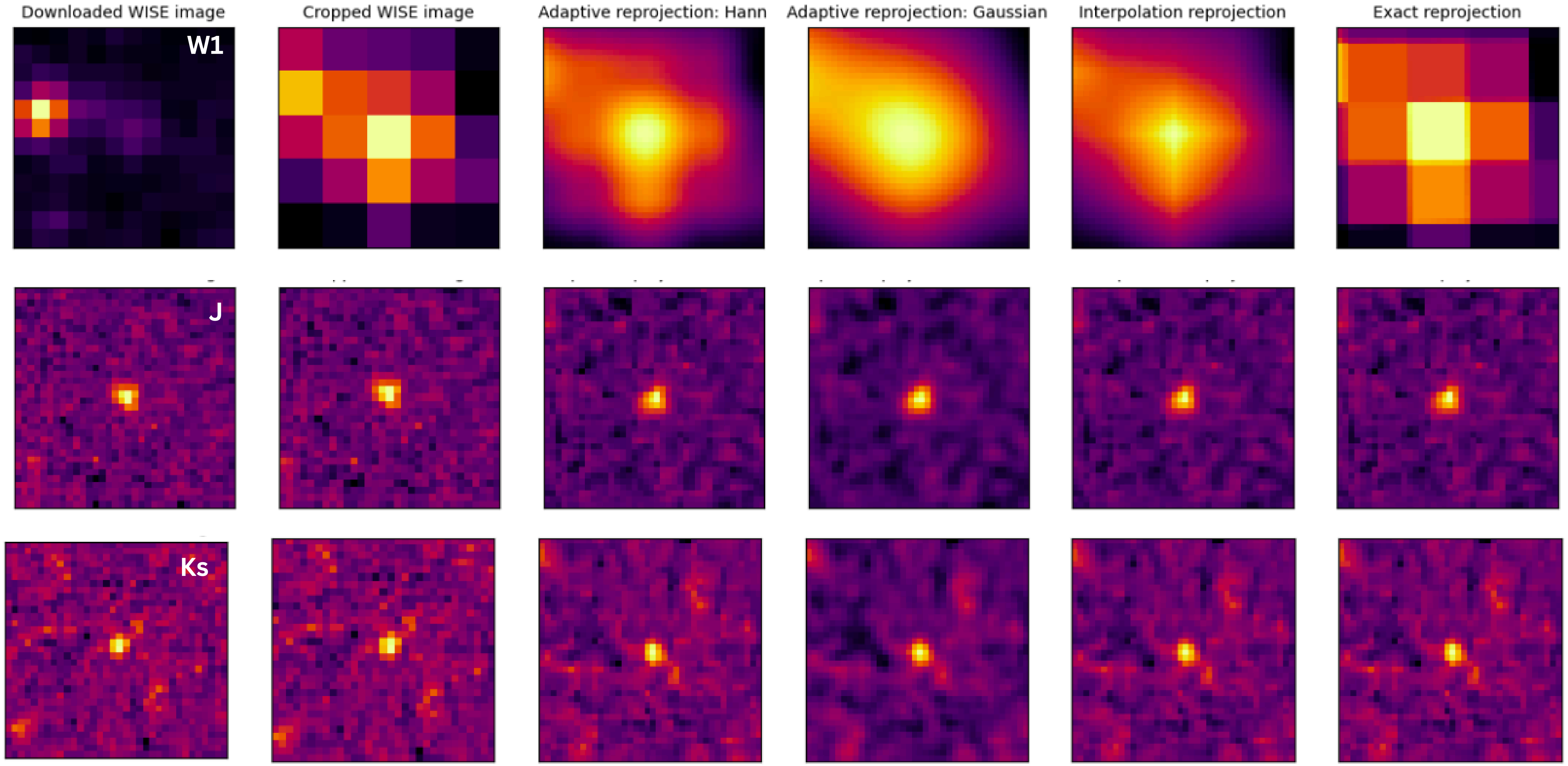
- **Training set**
- Image scales
- **Number of neighbors**
- **Minimum distance**
- Metrics
- Number of epochs



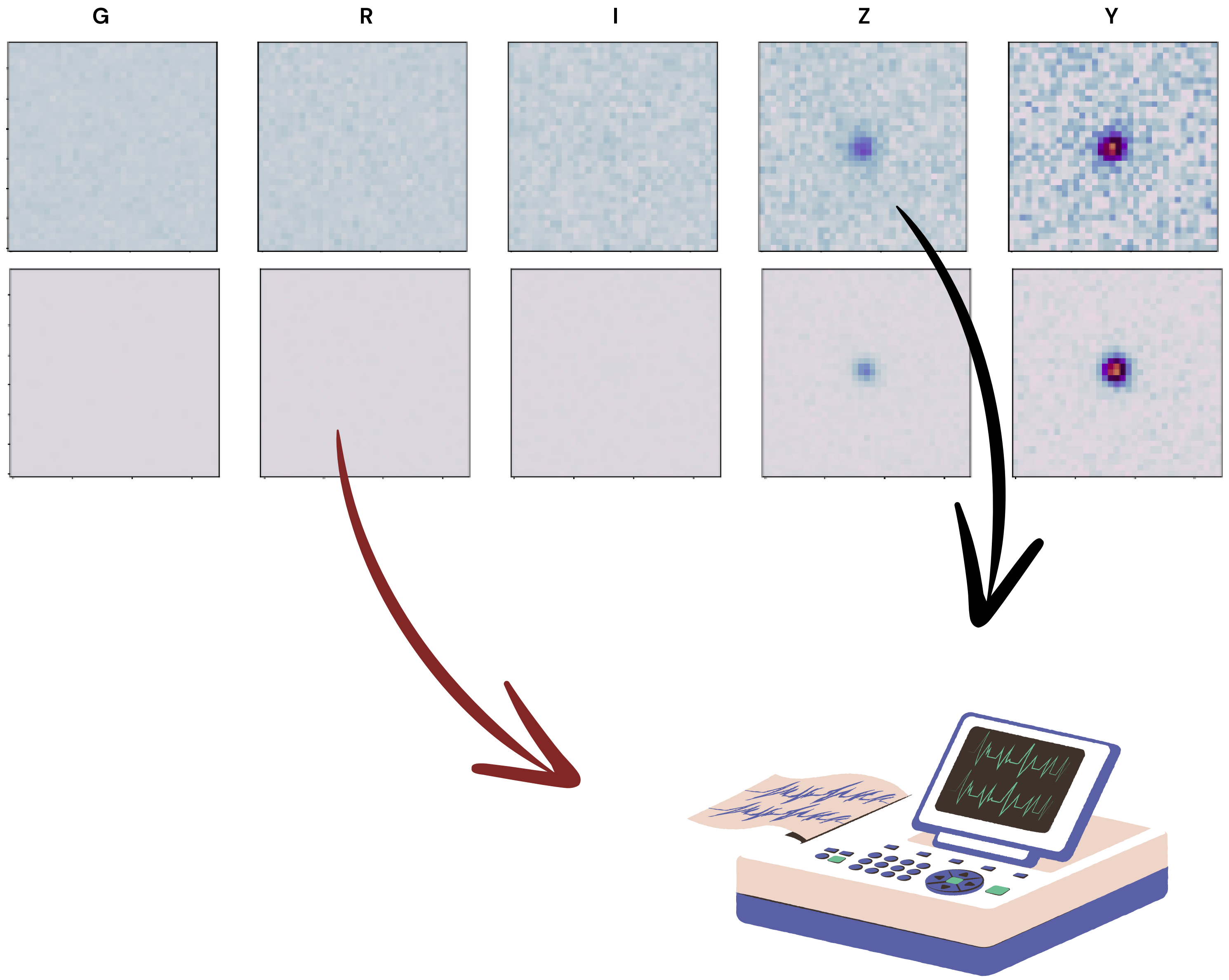
More information can help

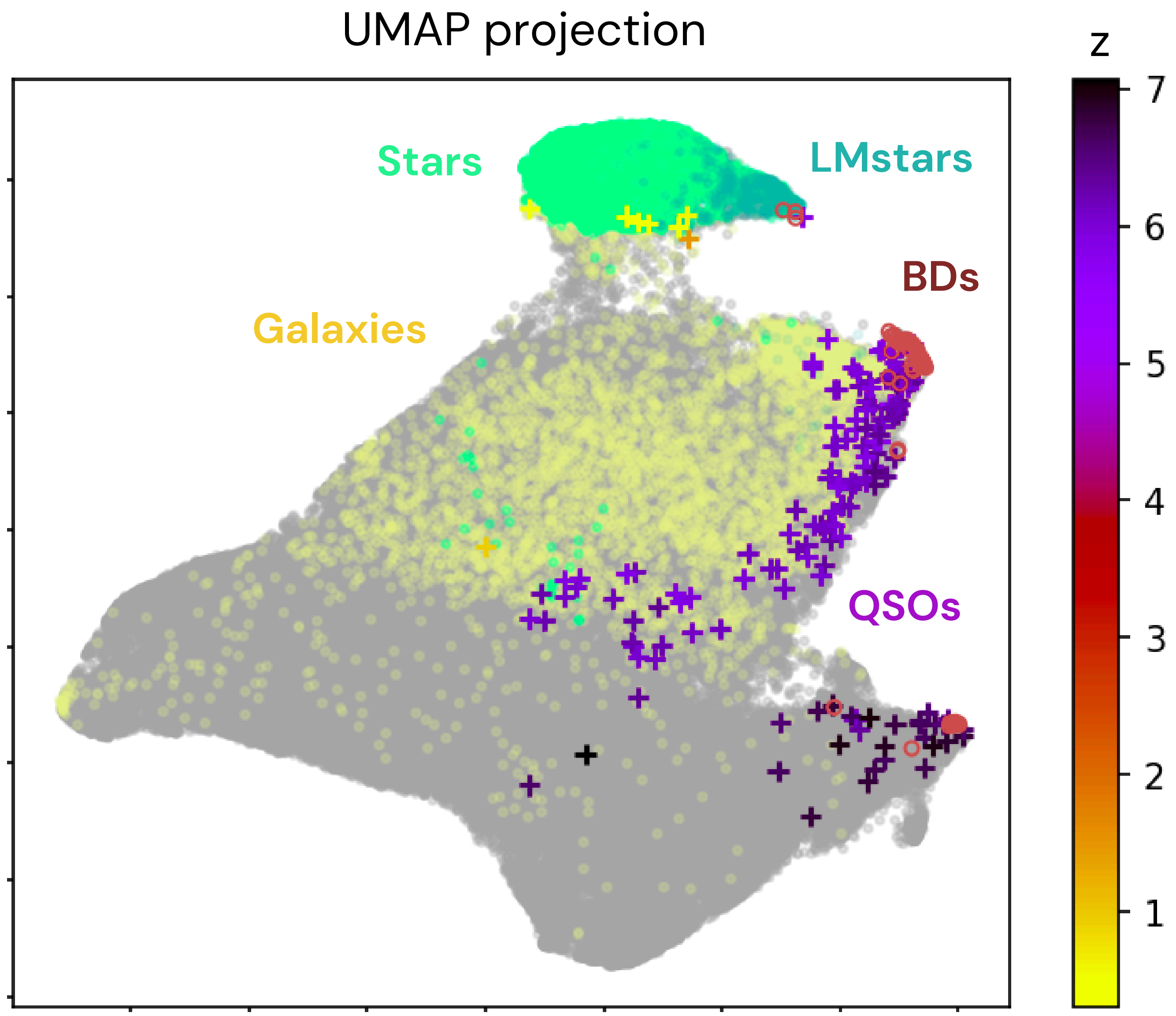
Work in progress: adding IR images

The challenge: different resolutions which means non-homogeneoús image pixel-sizes

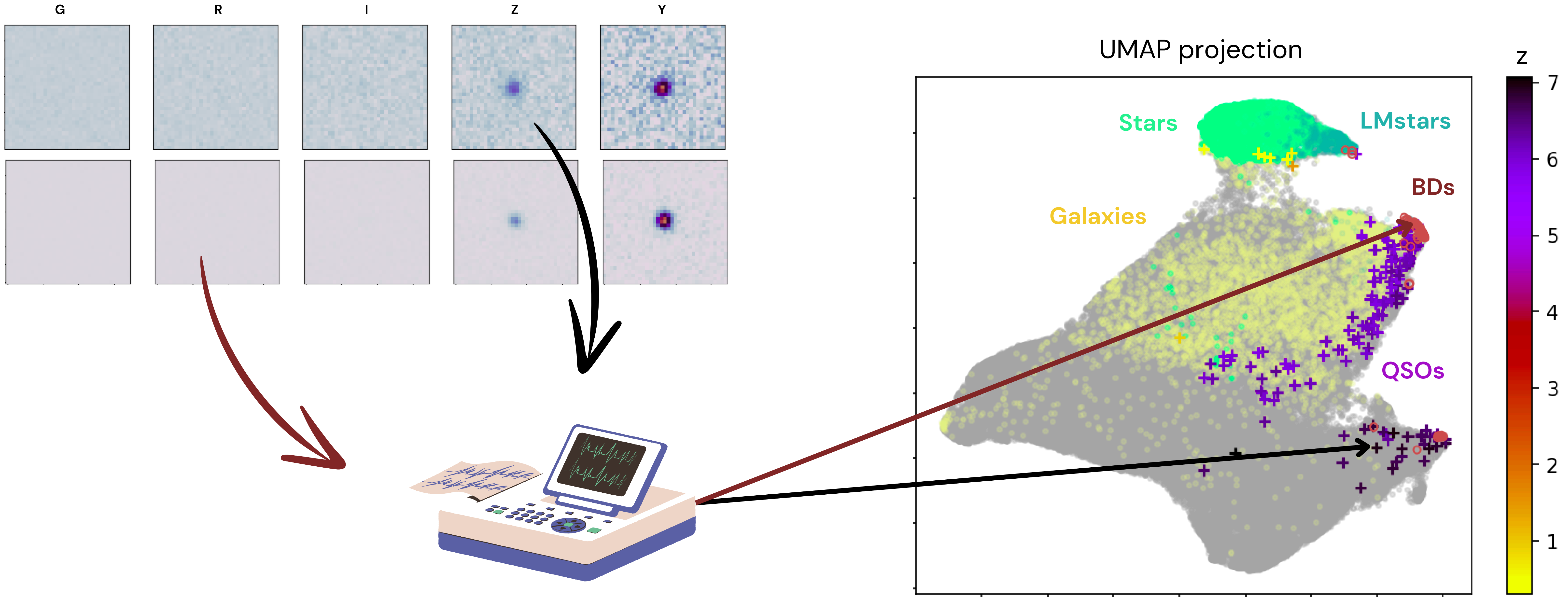


Which one is a z=6.4 quasar?

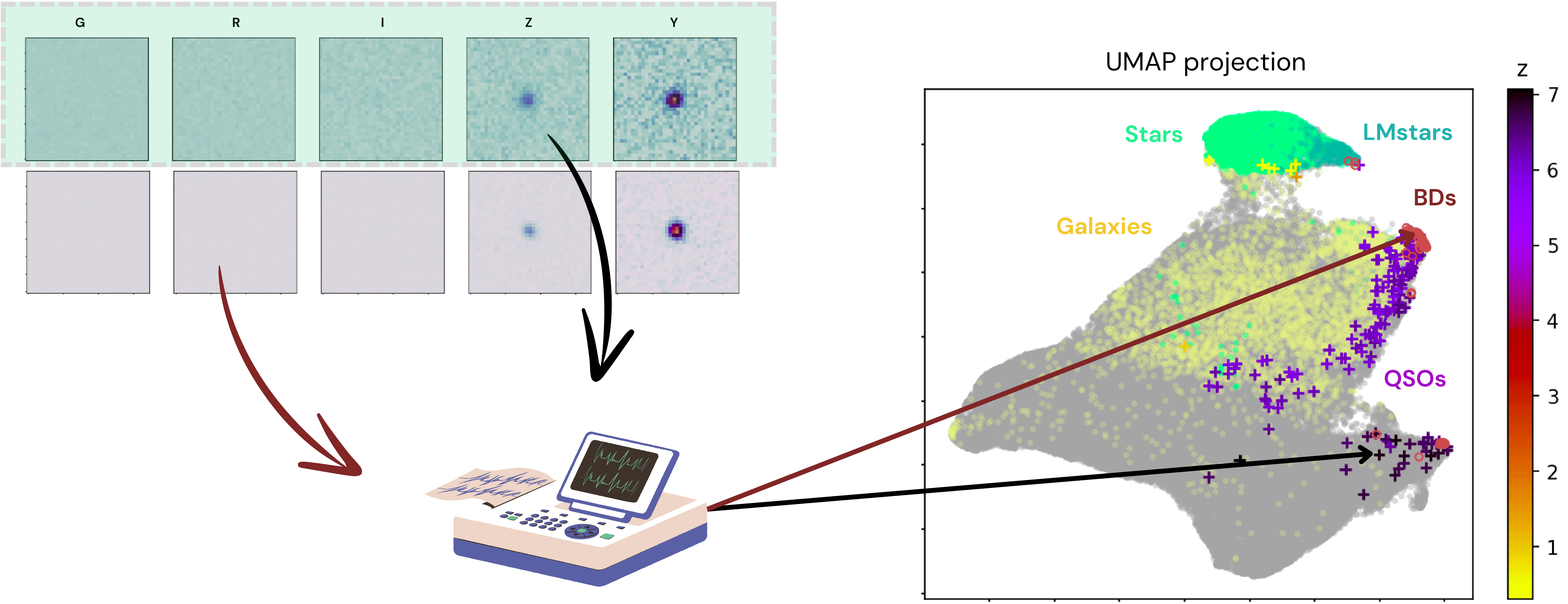




Which one is a z=6.4 quasar?



Which one is a z=6.4 quasar?



Take-away message

- **High-z QSOs are key to understand evolution of the Universe.**
- Representations on the low-dimensional embedded space by contrastive learning found: a **QSO evolutionary track**, a **z ~7 QSO island** and **brown dwarfs peninsula**.
- Future work:
  - Observing runs to confirm the selection and characterize the contaminants.
  - Hyperparameters fine tunning, IR data addition and augmentation function exploration to **improve performance.**

A lot of work to do! :)

THANK YOU!

BACK UP

Deep learning techniques in Astronomy

Supervised Unsupervised

Application		Model	CNNs	Enc.	Gene.	BNN	RNN	Trans.	GNN
1. Computer Vision	Classification	Morphology	✓	✓					
		Strong Lenses	✓*	✓*					
		Transients					✓*)	✓*)	
	Segmentation			✓*	✓*				
2. Galaxy Properties		Photoz	✓			✓			
		Structure	✓*)						
		Stellar Populations	✓*						
		Lensing	✓*			✓*			
		Physical Processes	✓*						
		Dark Matter	✓*			✓*			✓*
3. Discovery		Visualization	✓	✓	✓				
		Outliers	✓	✓	✓		✓		
		Laws							✓*
4. Cosmology		Emulation	✓*	✓*	✓*	✓*			
		Cosmological inference	✓*		✓*	✓*			

Huertas-Company & Lanusse (2023)

Deep learning techniques in Astronomy

Supervised Unsupervised

Application \ Model			CNNs	Enc.	Gene.	BNN	RNN	Trans.	GNN
1. Computer Vision	Classification	Morphology	✓	✓					
		Strong Lenses	✓*	✓*					
		Transients					✓*)	✓*)	
	Segmentation			✓*	✓*				

Challenge 1 Small (and biased) labelled datasets

Solution 1.A Transfer Learning	Domínguez Sánchez et al. (2019), Samudre et al. (2022), Lukic et al. (2019)
Solution 1.B Simulated dataset	Jacobs et al. (2017), Vega-Ferrero et al. (2021)
Solution 1.C Self-supervised learning	Hayat et al. (2021)
Solution 1.D Active Learning and similar	Walmsley et al. (2020)

Application \ Model			CNNs	Enc.	Gene.	BNN	RNN	Trans.	GNN
		Dark Matter	✓*			✓*			✓*
3. Discovery		Visualization	✓	✓	✓				
		Outliers	✓	✓	✓		✓		
		Laws							✓*
4. Cosmology		Emulation	✓*	✓*	✓*	✓*			
		Cosmological inference	✓*		✓*	✓*			

Huertas-Company & Lanusse (2023)

Why to study high-z quasars?

- SMBHs and host galaxies coevolution

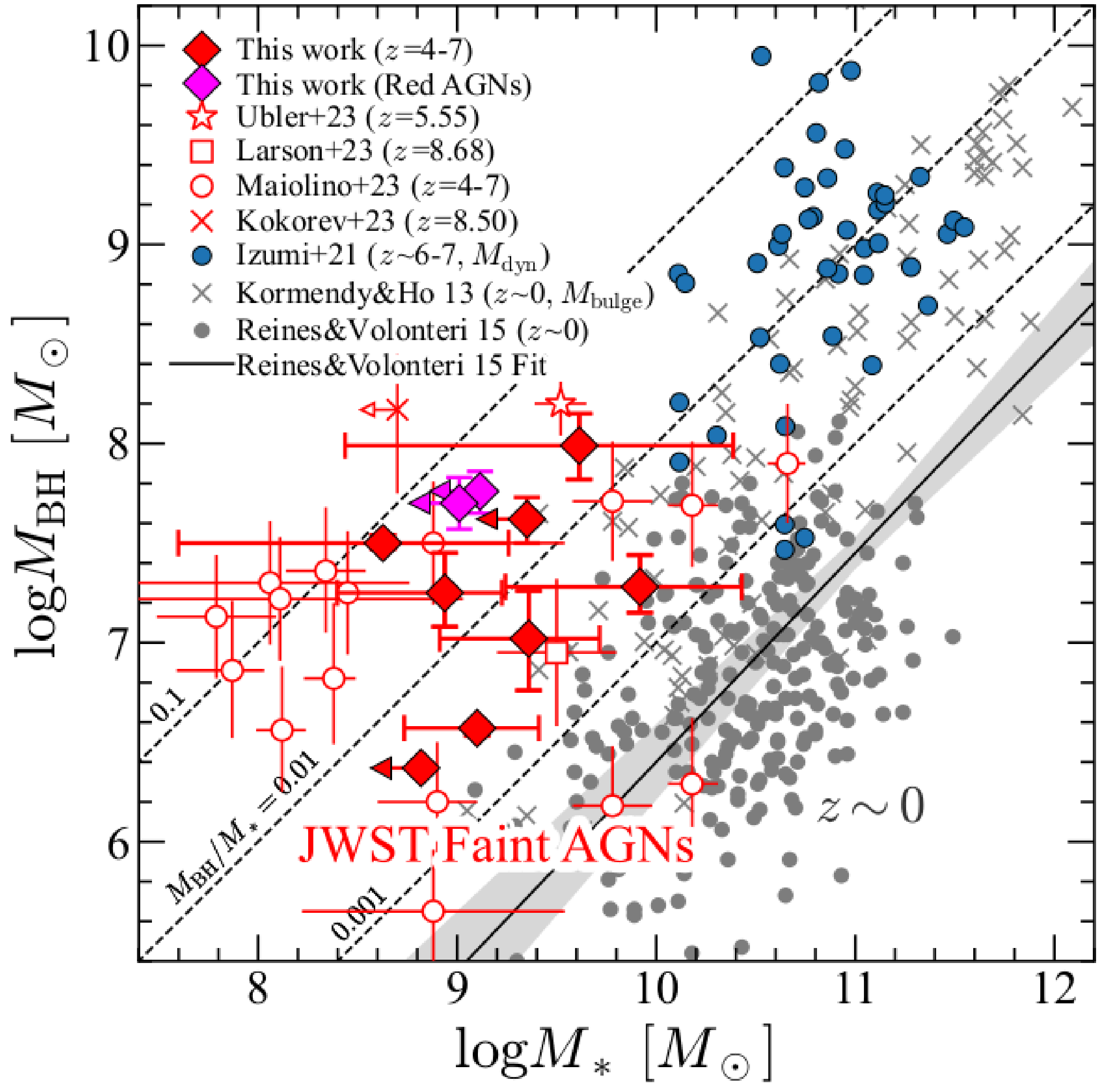


Figure from Harikane et al. (2023)

- Role of QSOs in cosmic reionization

- QSOs are standardizable candles

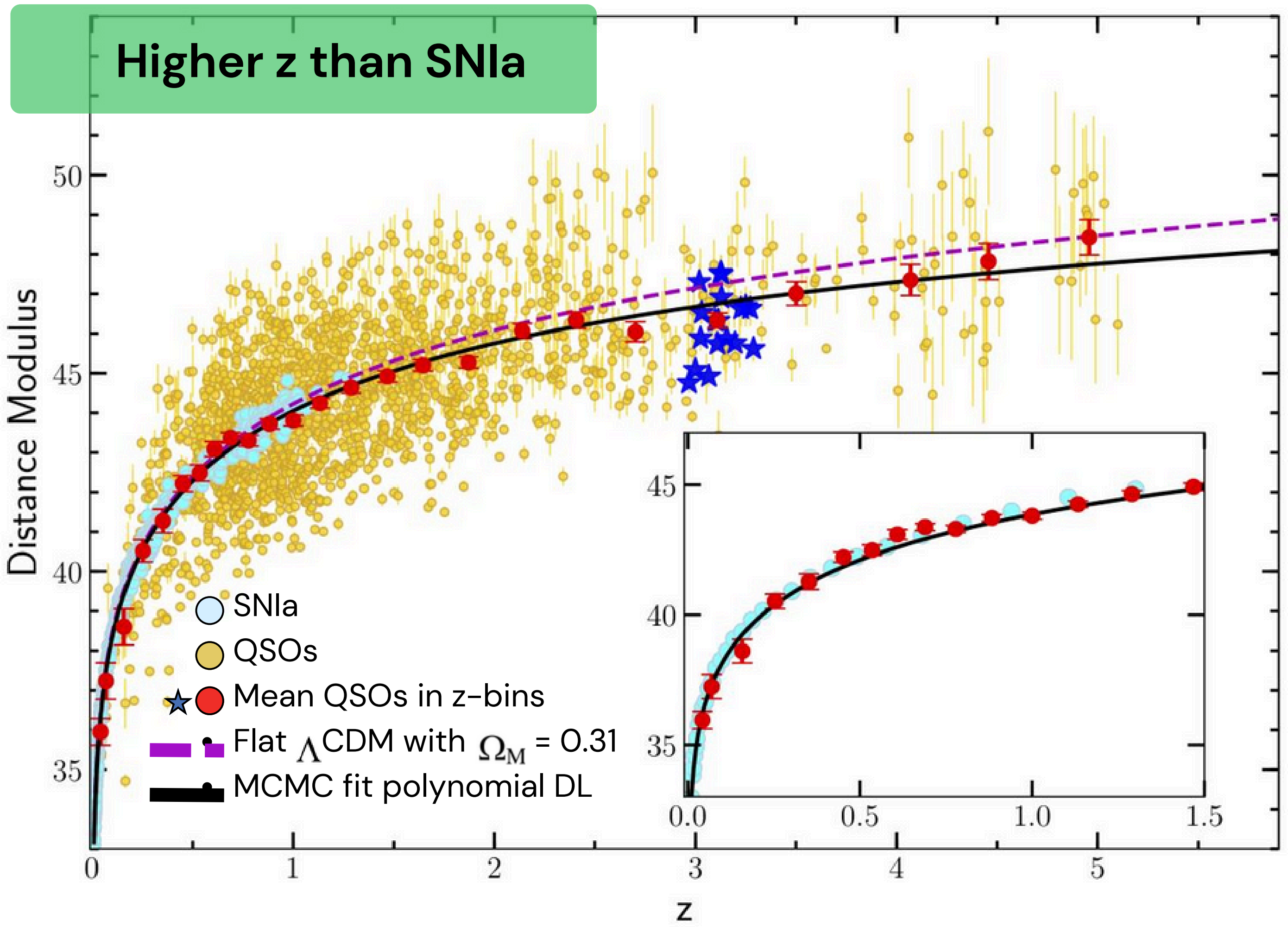


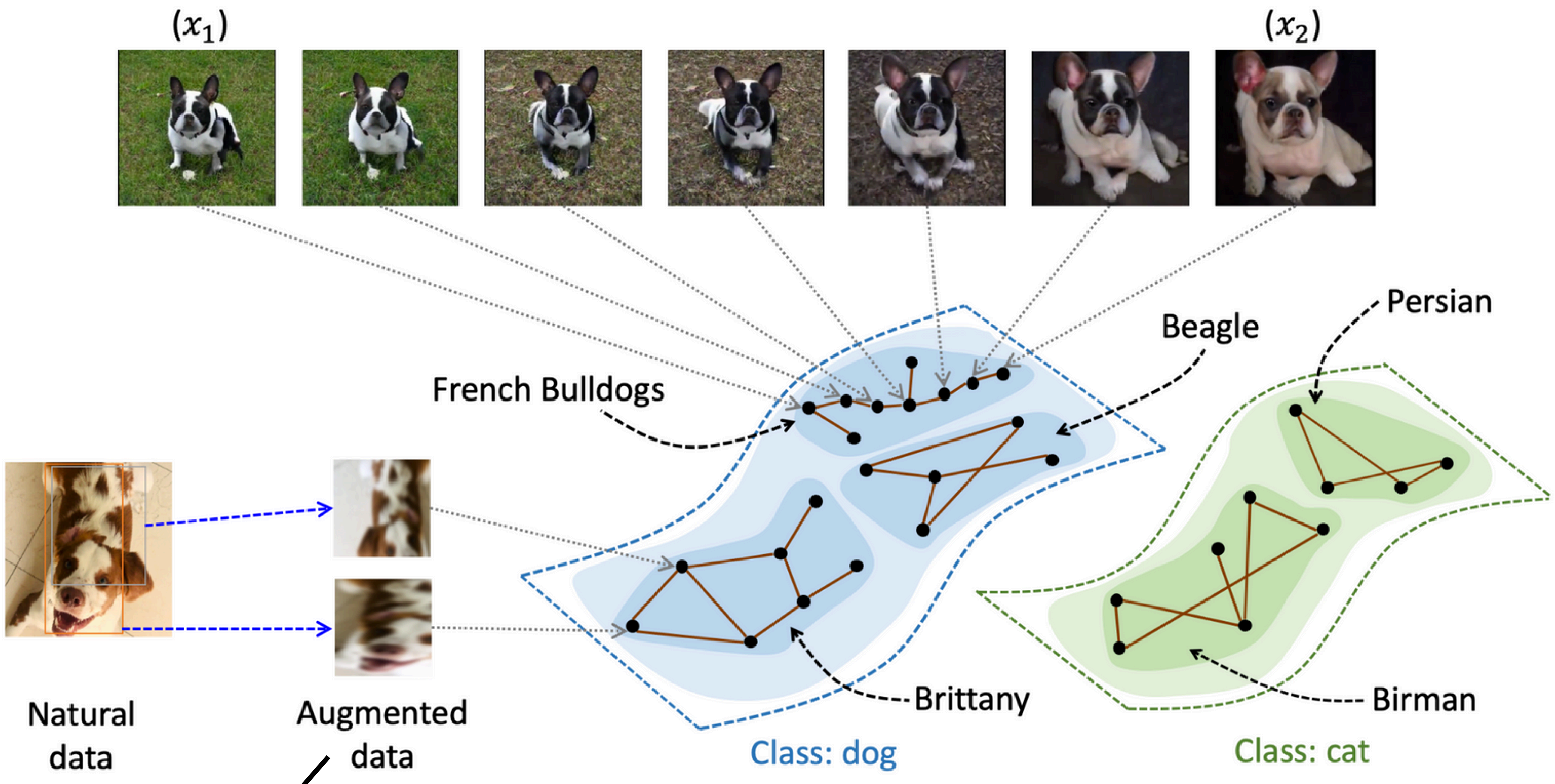
Figure from Risaliti & Lusso (2017)

- Bright QSOs are tracers of overdensities

- Special AGN properties at high-z

Contrastive learning

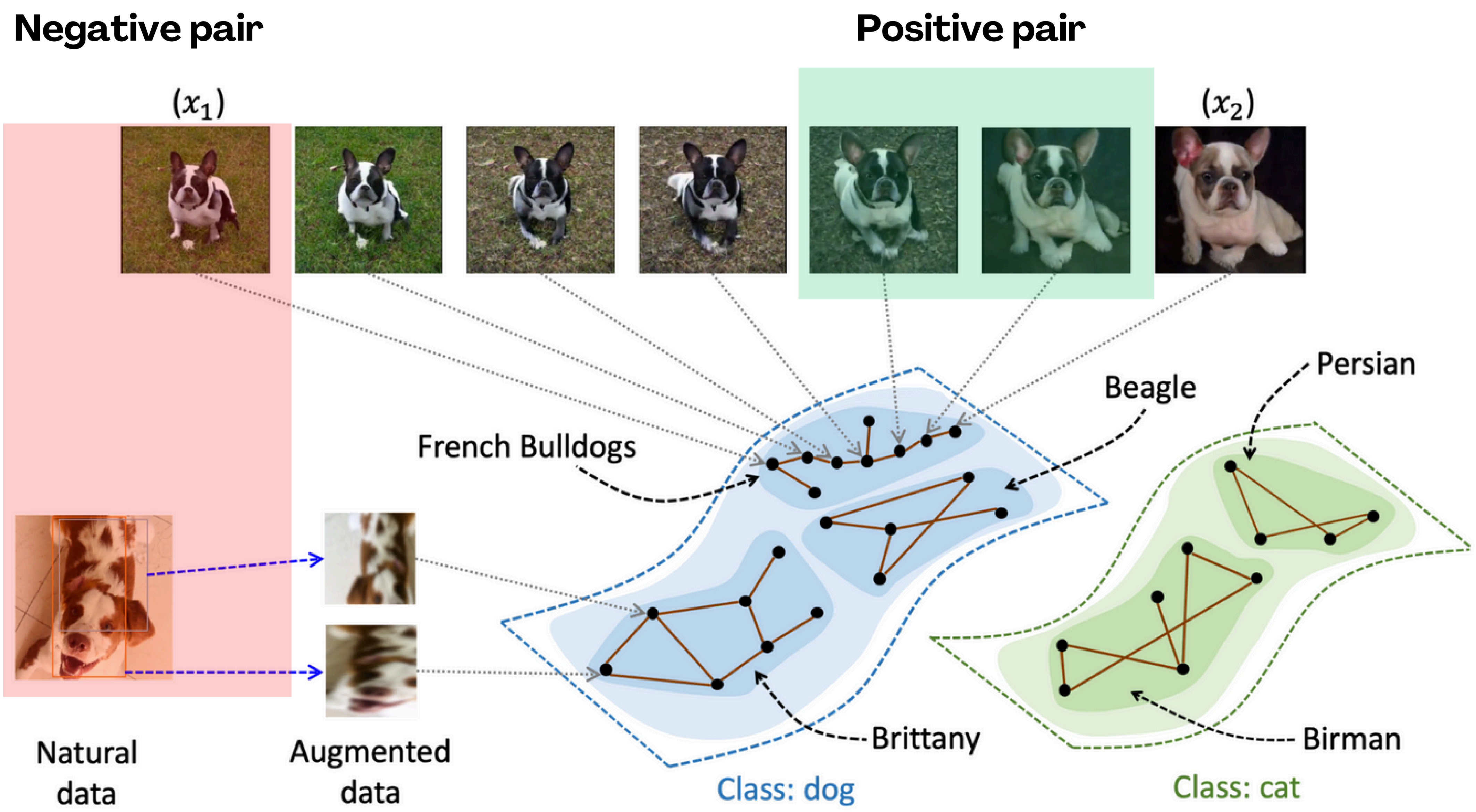
It is a self-supervised technique that trains a neural network (NN) to **project** input data onto a **low-dimensional embedding space** while minimizing the distance between similar objects.



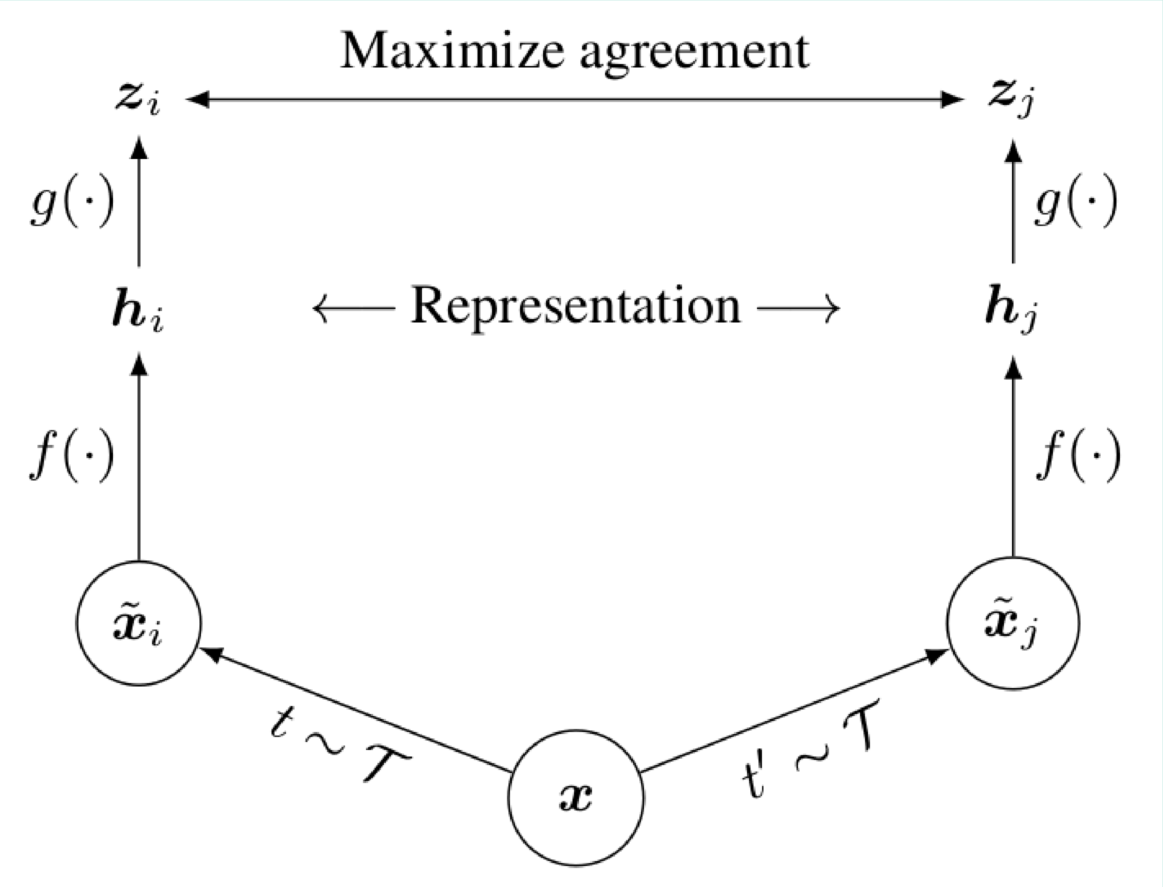
to make the representation invariant to different views of the same objects

HaoChen, Wie & Ma (2022)

Contrastive learning



HaoChen, Wie & Ma (2022)

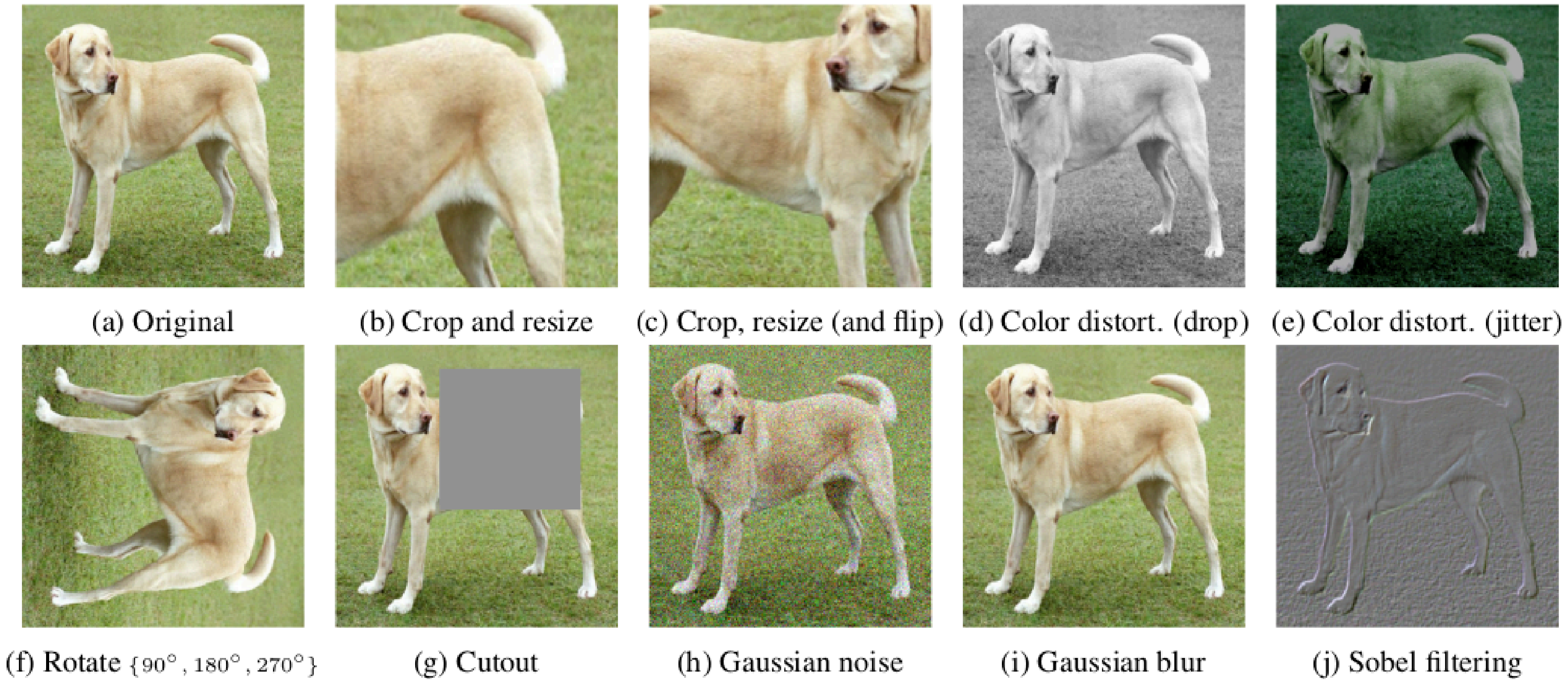


Chen et al. (2020)

Main components:

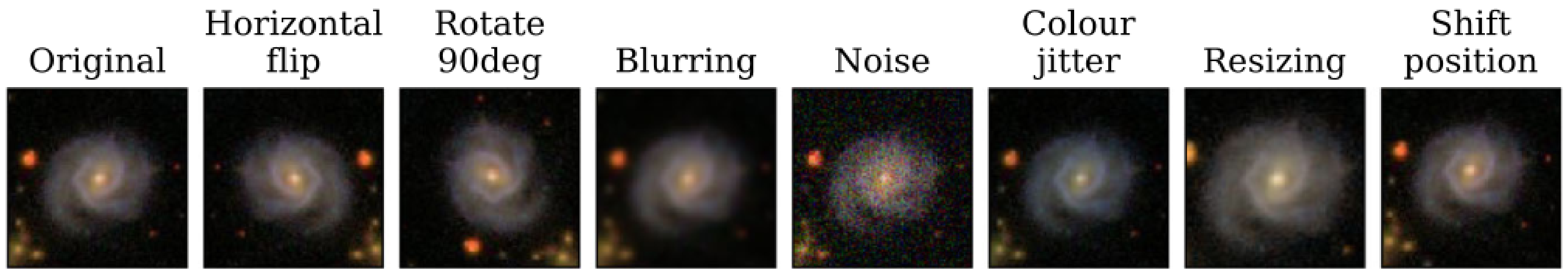
- A stochastic data augmentation
- A NN base encoder f(.)
- A small NN projection head g(.)
- A contrastive loss function

Contrastive learning

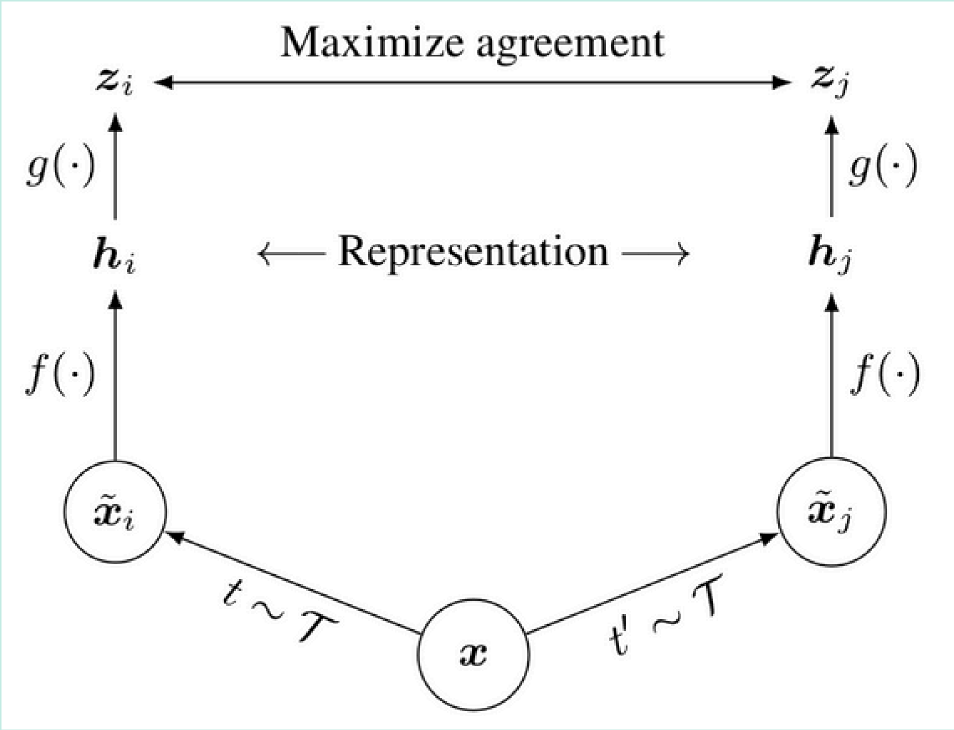


Chen et al. (2020)

Perturbations can be tuned for a science case, for example to make the representations independent to instrumental and/or selection biases.



Huertas-Company, Sarmiento and Knapen (2023)



Chen et al. (2020)

Main components:

- A stochastic **data augmentation**
- A NN base encoder f(.)
- A small NN projection head g(.)
- A contrastive loss function

Contrastive learning

- The **encoder** extracts representation vectors from augmented data during the training, and natural data for downstream tasks.
- The **projection head** maps the representations to the space where contrastive loss is computed. It's used only during the training phase.

Normalized temperature-scaled cross entropy **loss function**

$$\mathrm{sim}(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u}^\top \boldsymbol{v} / \|\boldsymbol{u}\| \|\boldsymbol{v}\|$$

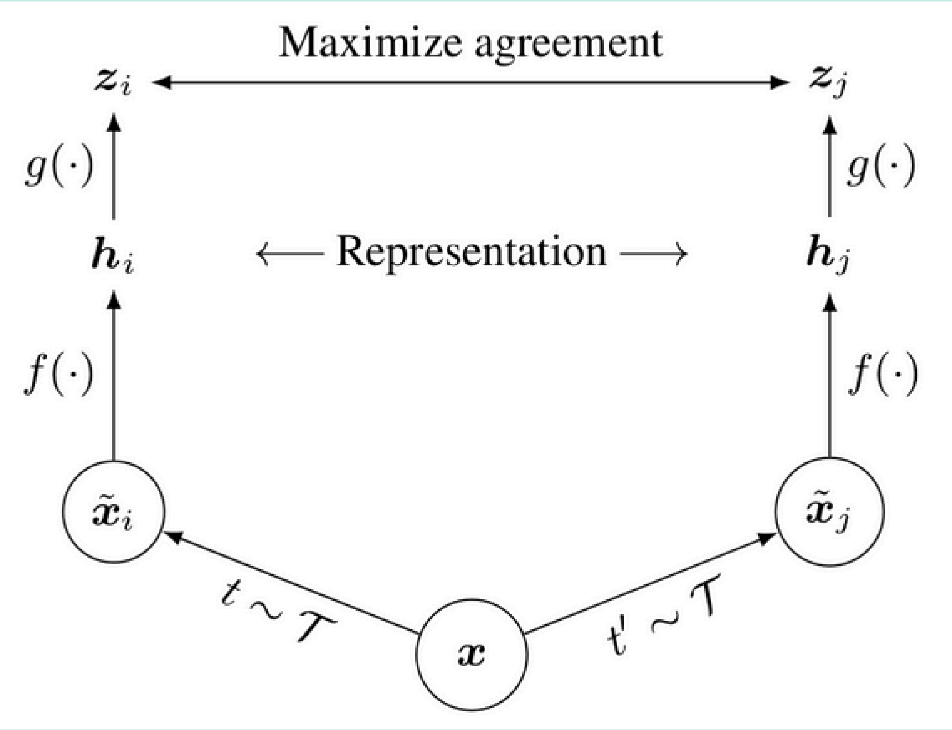
Cosine similarity

$$\ell_{i,j} = -\log \frac{\exp(\mathrm{sim}(\boldsymbol{z}_i, \boldsymbol{z}_j)/\tau)}{\sum_{k=1}^{2N} \mathbb{1}_{[k \neq i]} \exp(\mathrm{sim}(\boldsymbol{z}_i, \boldsymbol{z}_k)/\tau)}$$

Temperature parameter

Name	Negative loss function
NT-Xent	$\boldsymbol{u}^T \boldsymbol{v}^+ / \tau - \log \sum_{\boldsymbol{v} \in \{\boldsymbol{v}^+, \boldsymbol{v}^-\}} \exp(\boldsymbol{u}^T \boldsymbol{v} / \tau)$
NT-Logistic	$\log \sigma(\boldsymbol{u}^T \boldsymbol{v}^+ / \tau) + \log \sigma(-\boldsymbol{u}^T \boldsymbol{v}^- / \tau)$
Margin Triplet	$-\max(\boldsymbol{u}^T \boldsymbol{v}^- - \boldsymbol{u}^T \boldsymbol{v}^+ + m, 0)$

Chen et al. (2020)



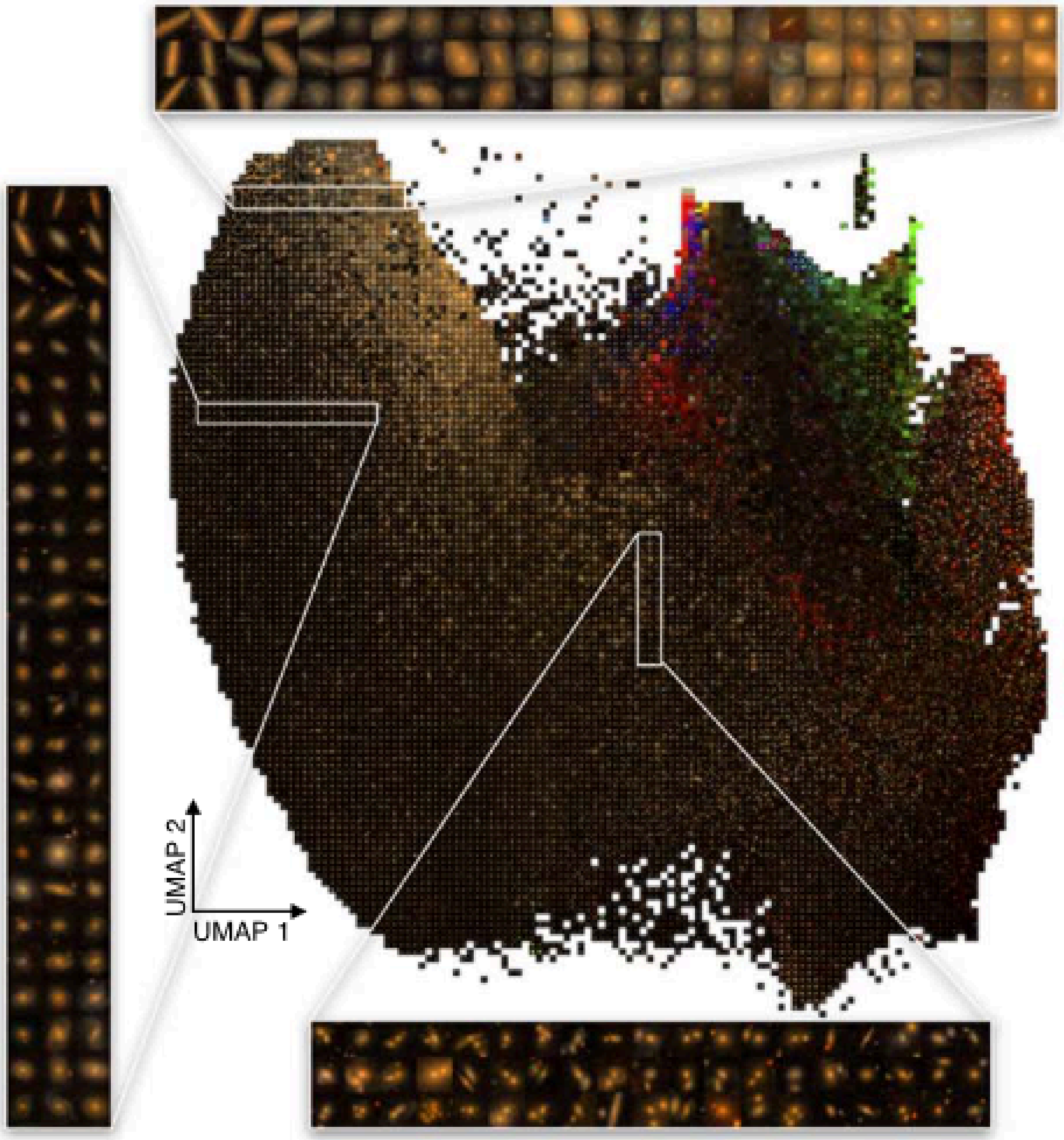
Chen et al. (2020)

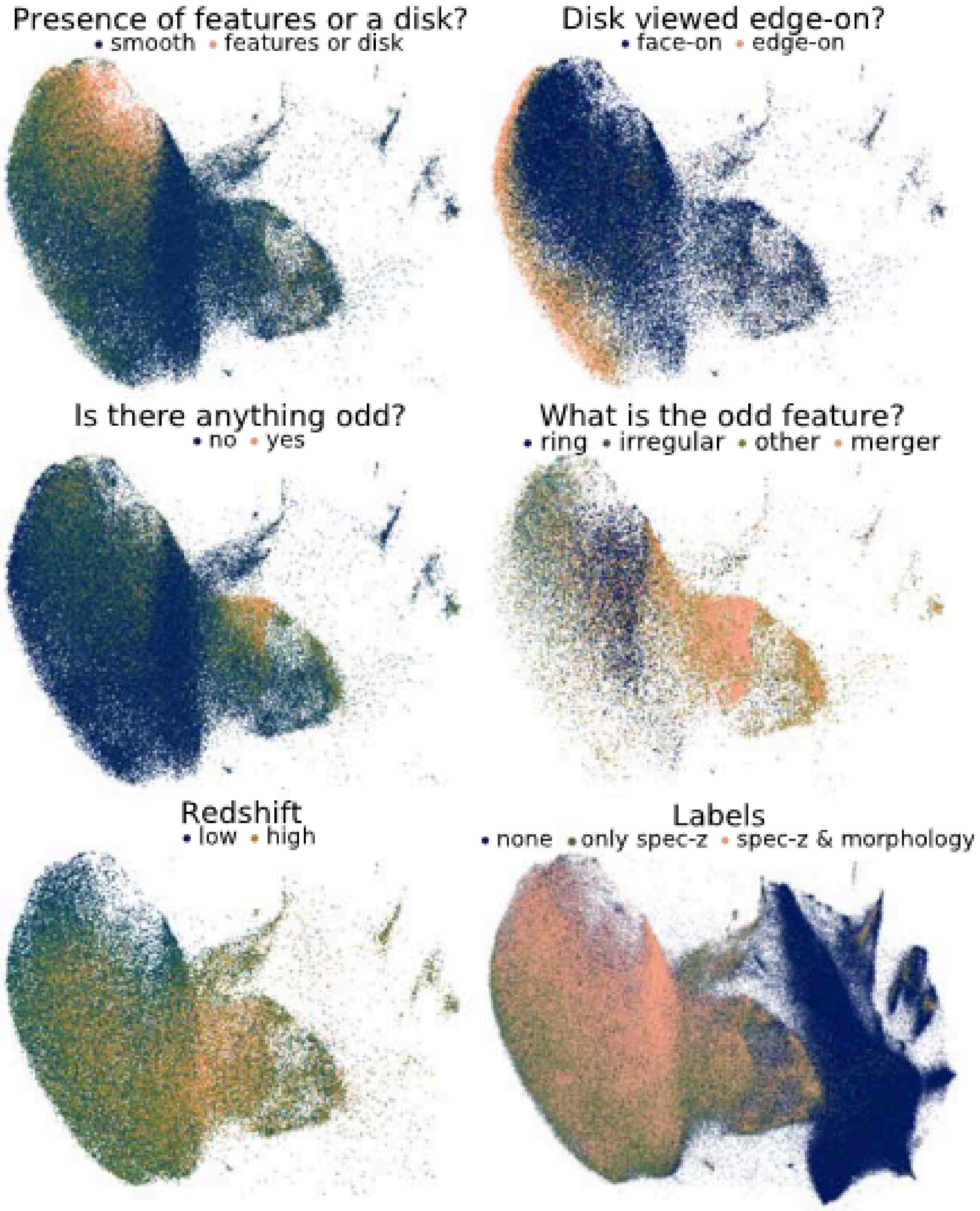
Main components:

- A stochastic data augmentation
- A NN base **encoder f(.)**
- A small NN **projection head g(.)**
- A **contrastive loss function**

Some examples

Self-supervised learning on SDSS images with perturbations such as rotation, cropping and extinction.



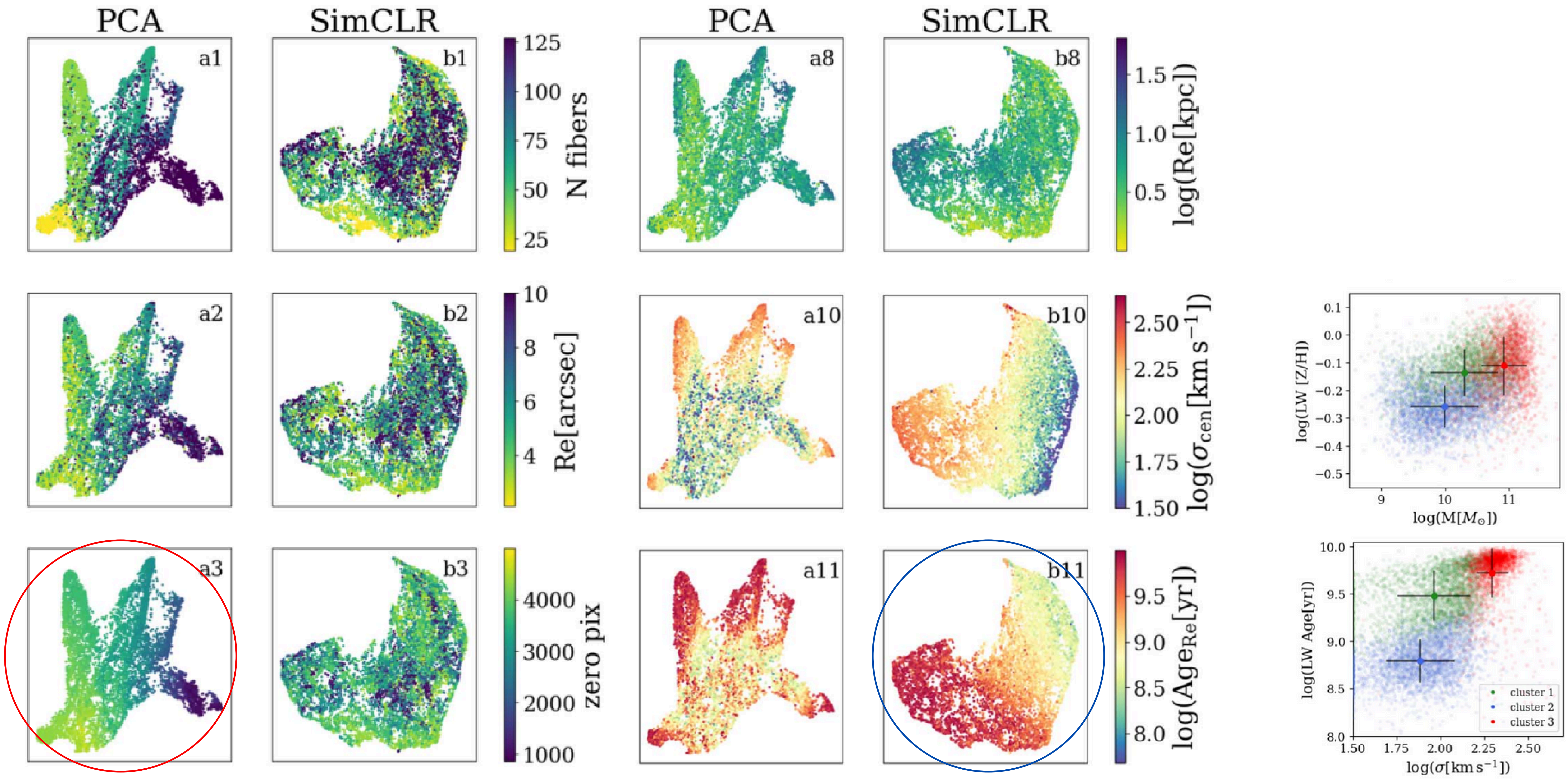


Hayat et al. (2021)

Some examples

Application to nearby galaxies from Manga survey by using maps of stellar population properties and kinematic maps.

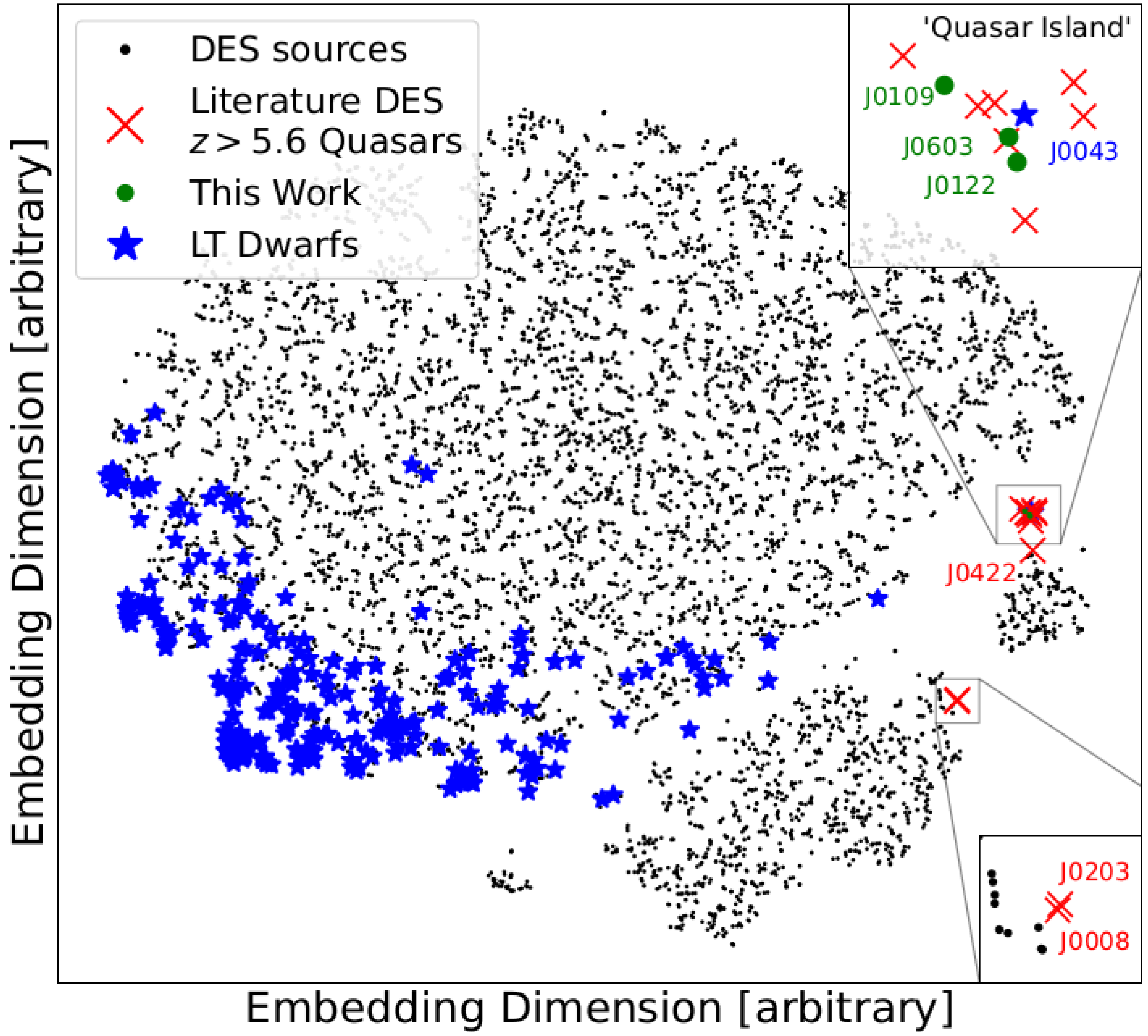
High dependence
on non-physical
parameters



Sarmiento et al. (2021)

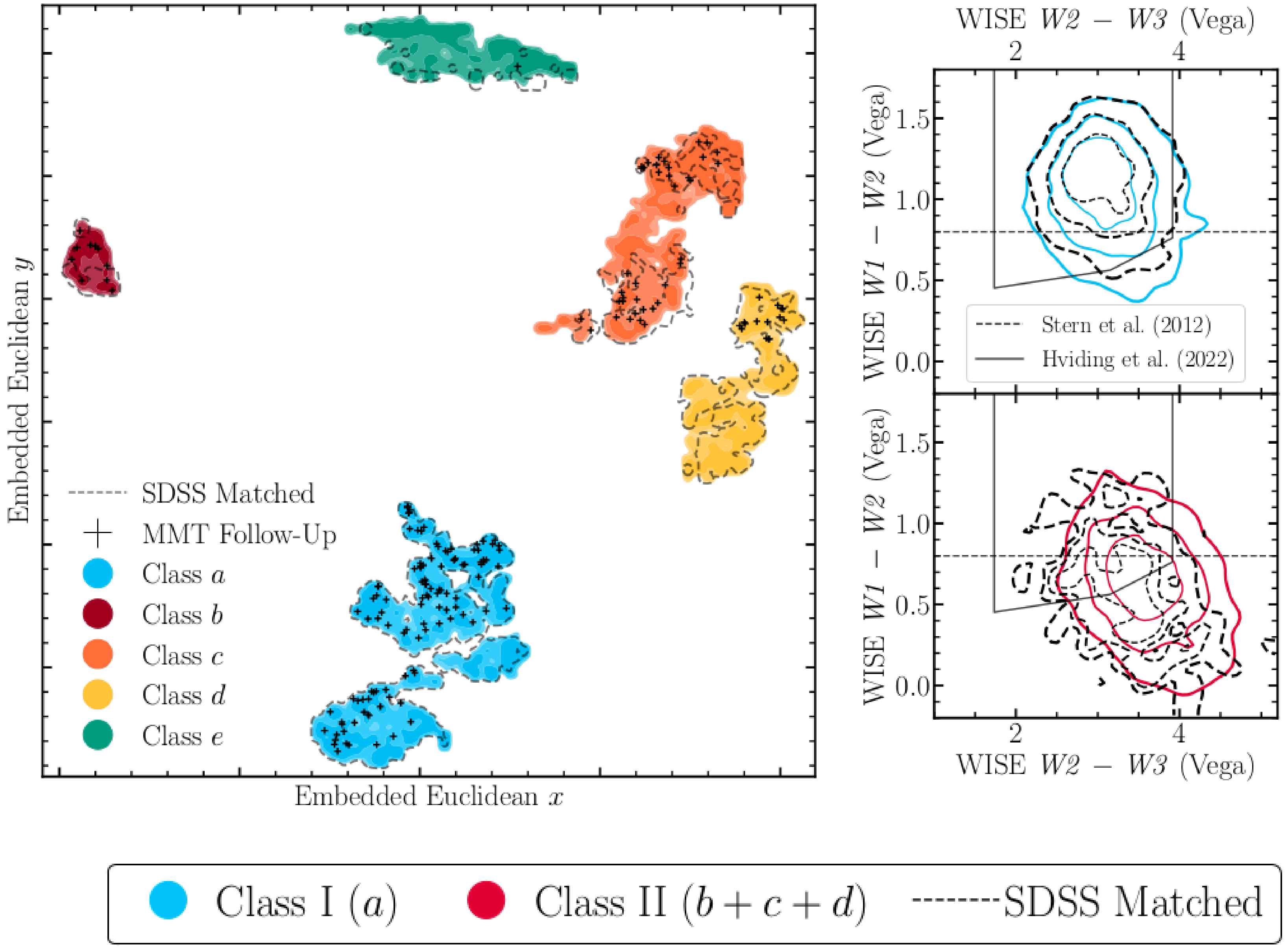
Some examples

Search for z > 6 QSOs with DES DR2 images



Byrne et al. (2024)

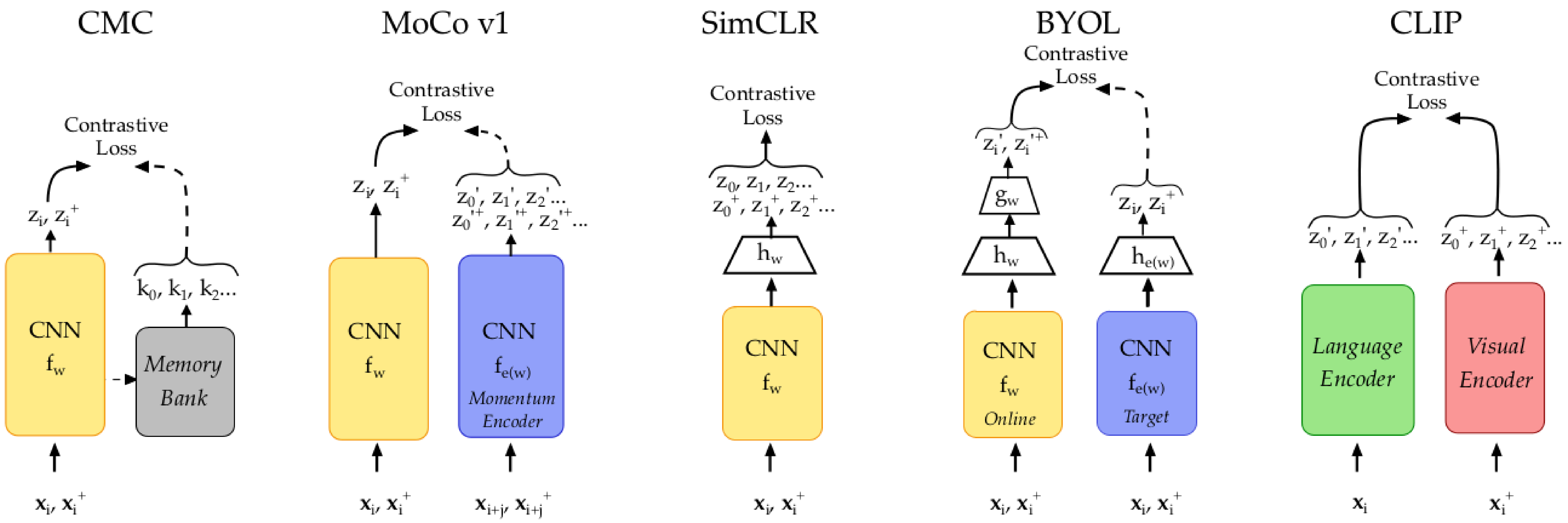
Search for obscured AGNs with color catalogs based on HSC SPP, allWISE and unWISE



Hviding et al. (2024)

CL framework zoo

Contrastive multiview coding | Momentum contrastive | Simple contrastive learning | Bootstrap your own latent | Contrastive language-image pre-training



For previous representations of negative pairs

Calculates updated representations by including the immediately previous batch

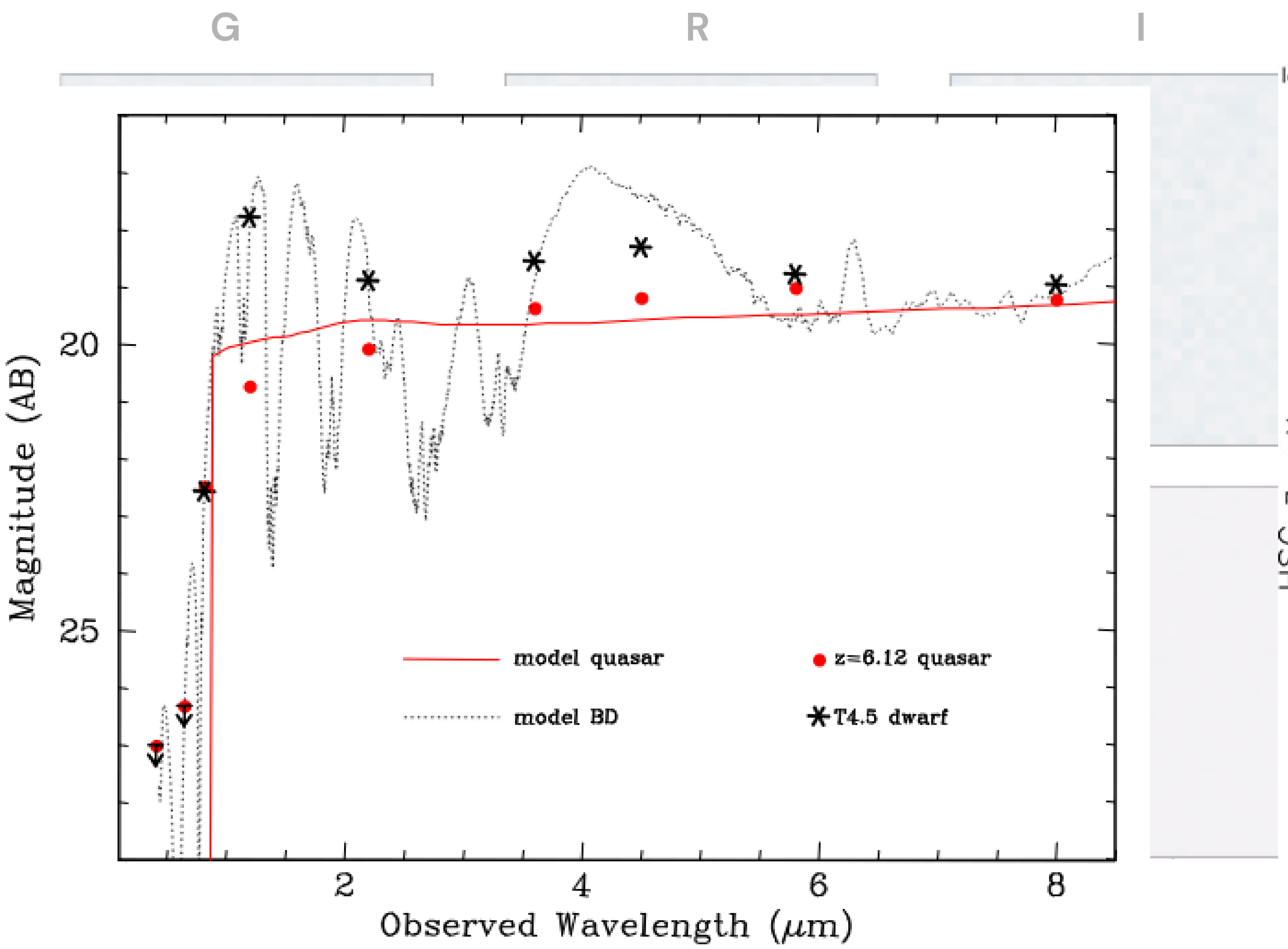
Negative pairs limited by batch size

Additional projection function

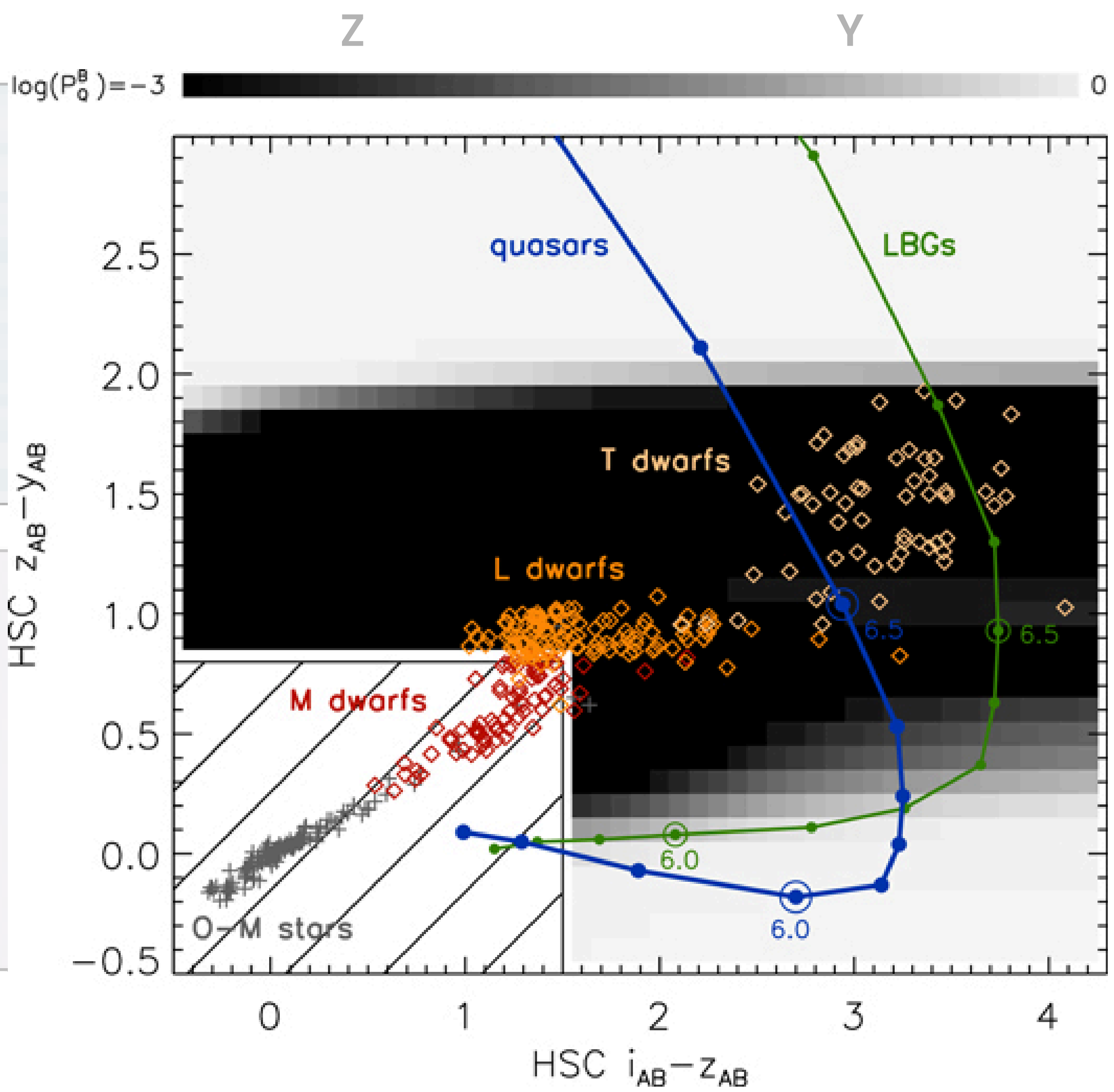
Combining different data types

Huertas-Company, Sarmiento and Knapen (2023)

Which one is a z=6.4 quasar?

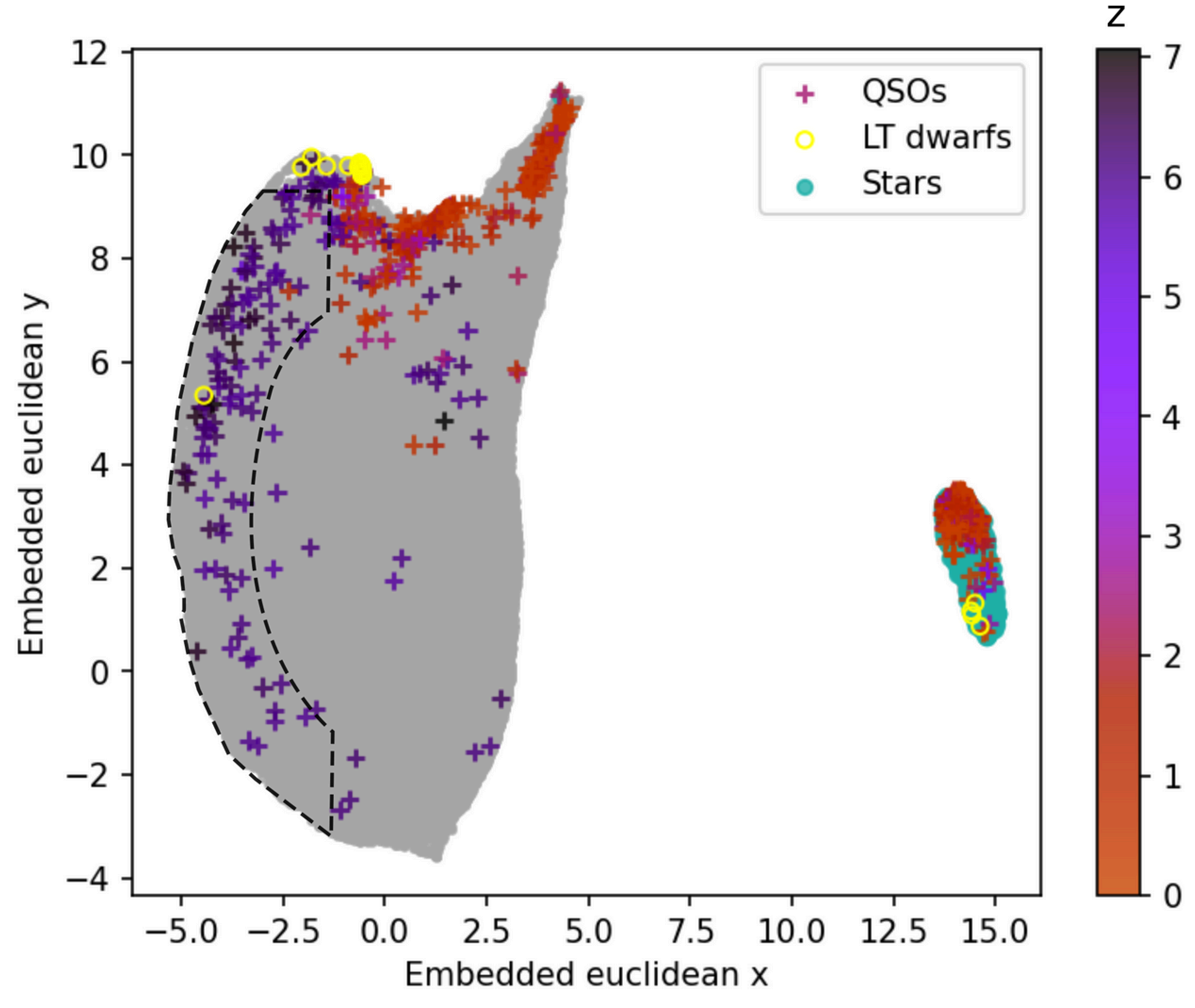


Stern et al. (2008)



Matsuoka et al. (2016)

Self-supervised contrastive learning for LBT proposal

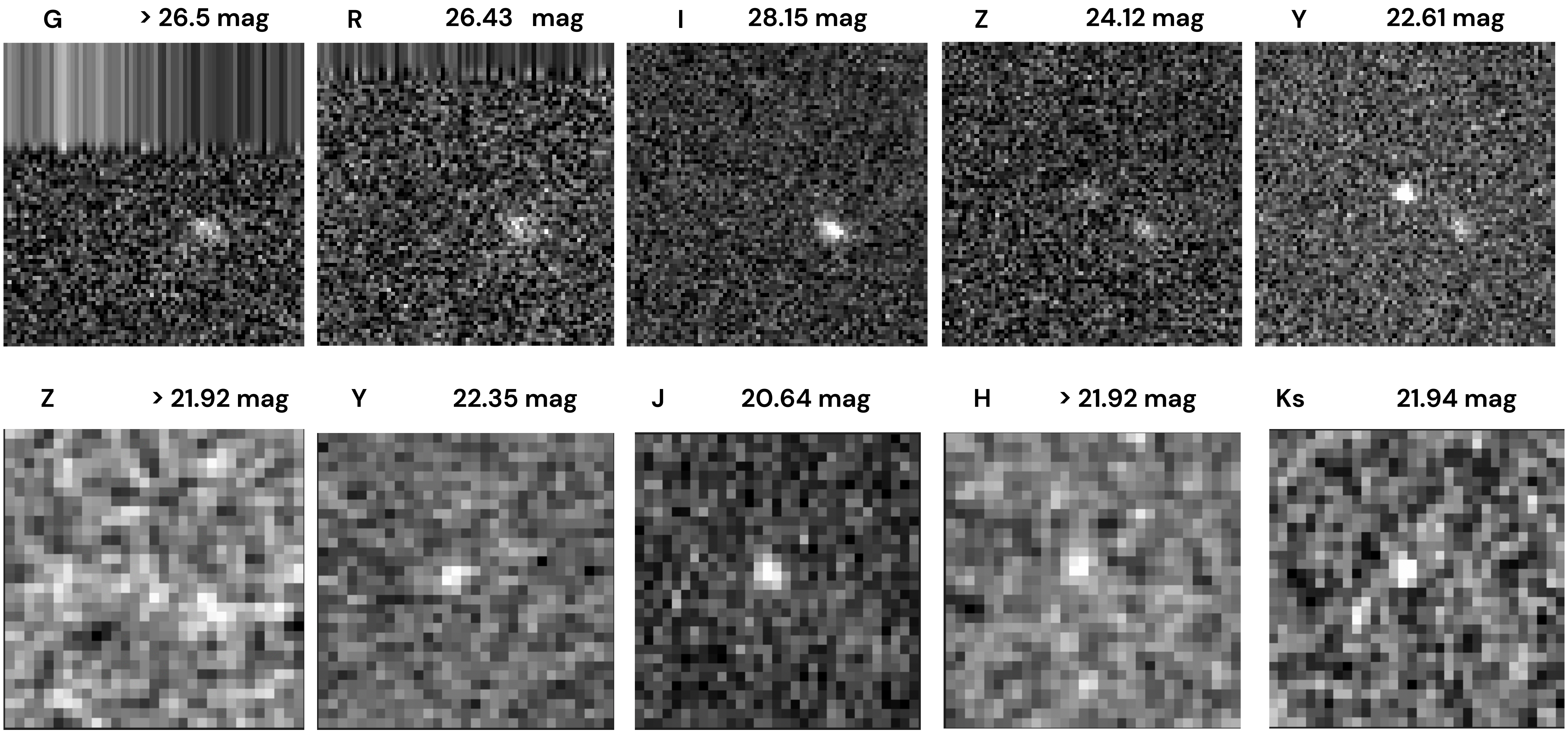


1 normalization value for all the tensor
We keep track of brighter vs faint sources

Main improvements:

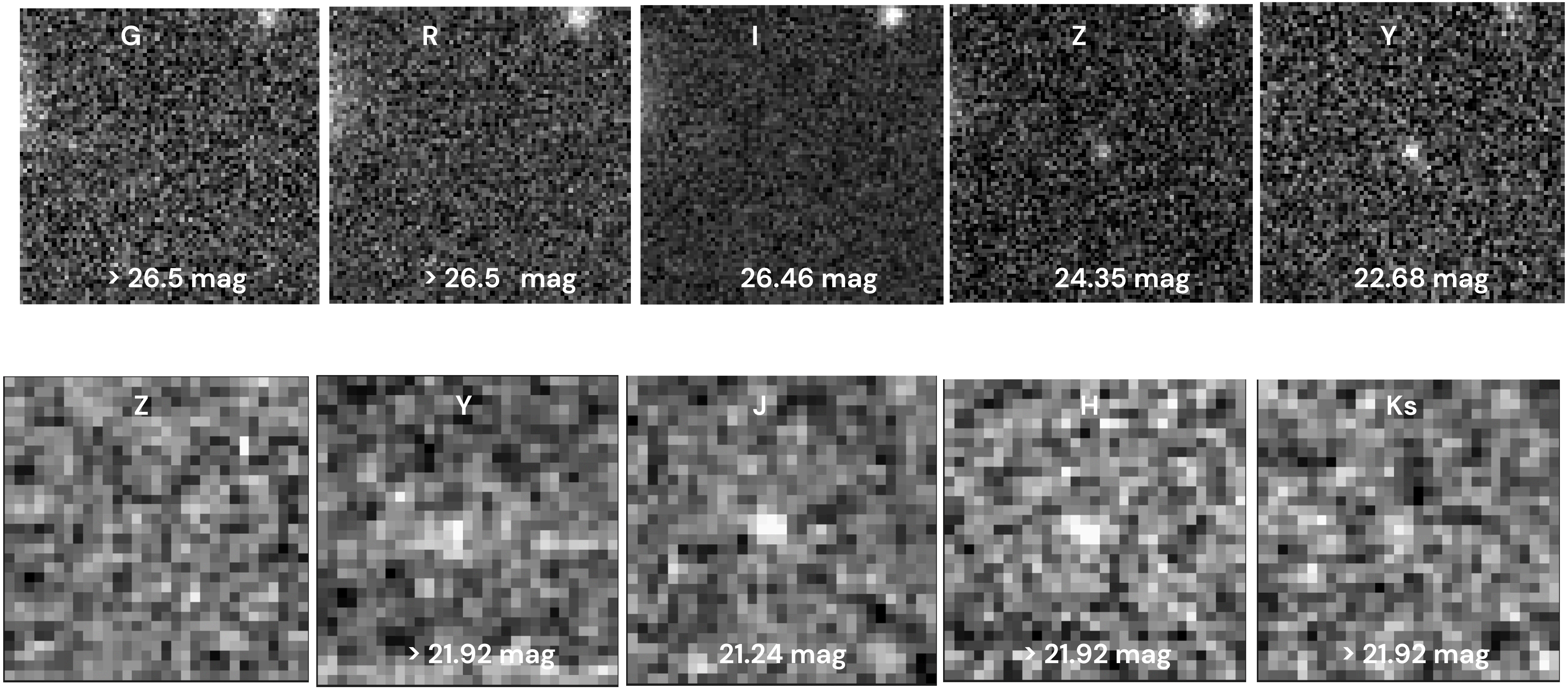
- **No CR contamination**
- **Inclusion of low-z QSOs labels**
- **Trained with more spatially distributed data**

Preliminary results: VHS_DR4 constraint

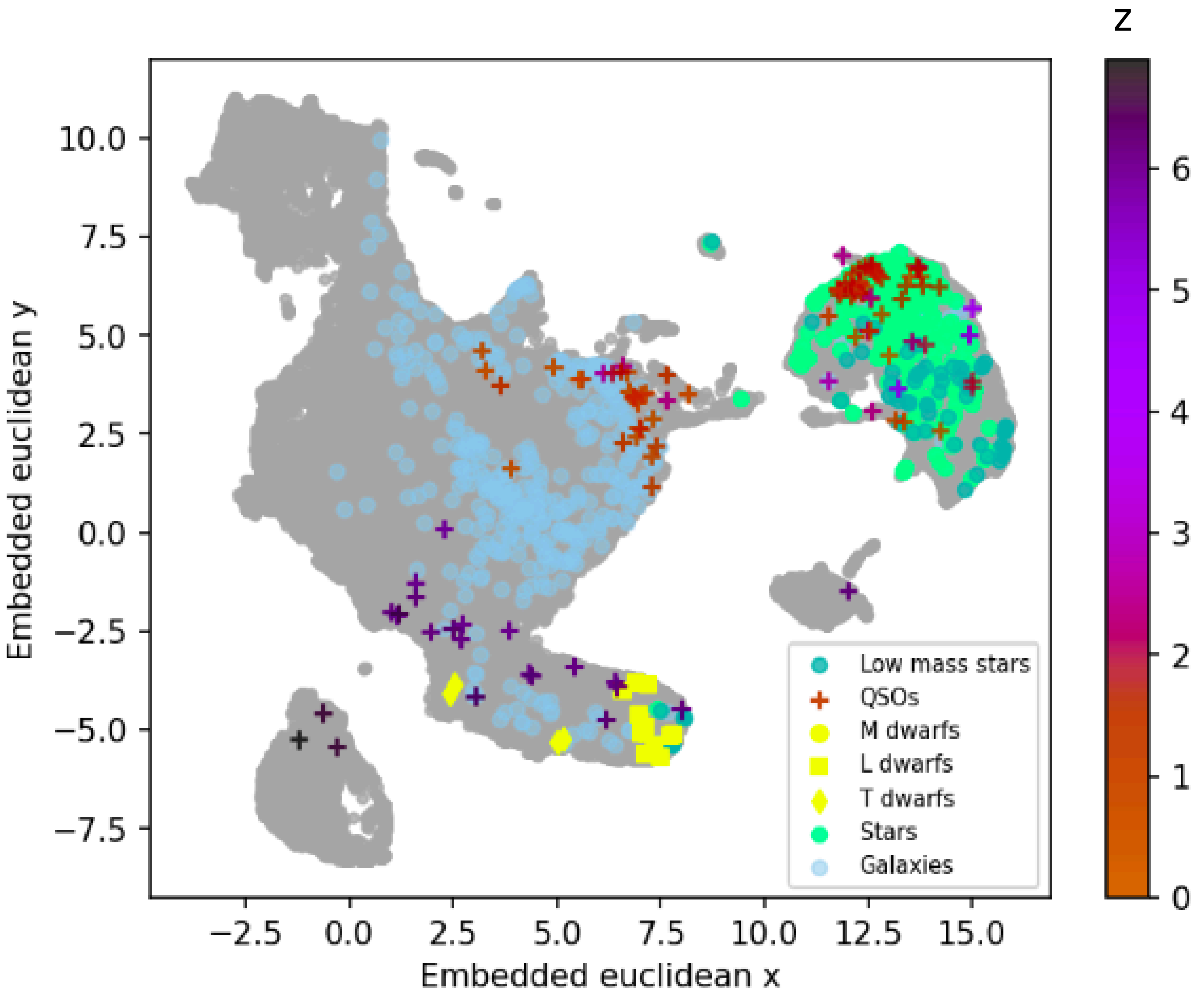


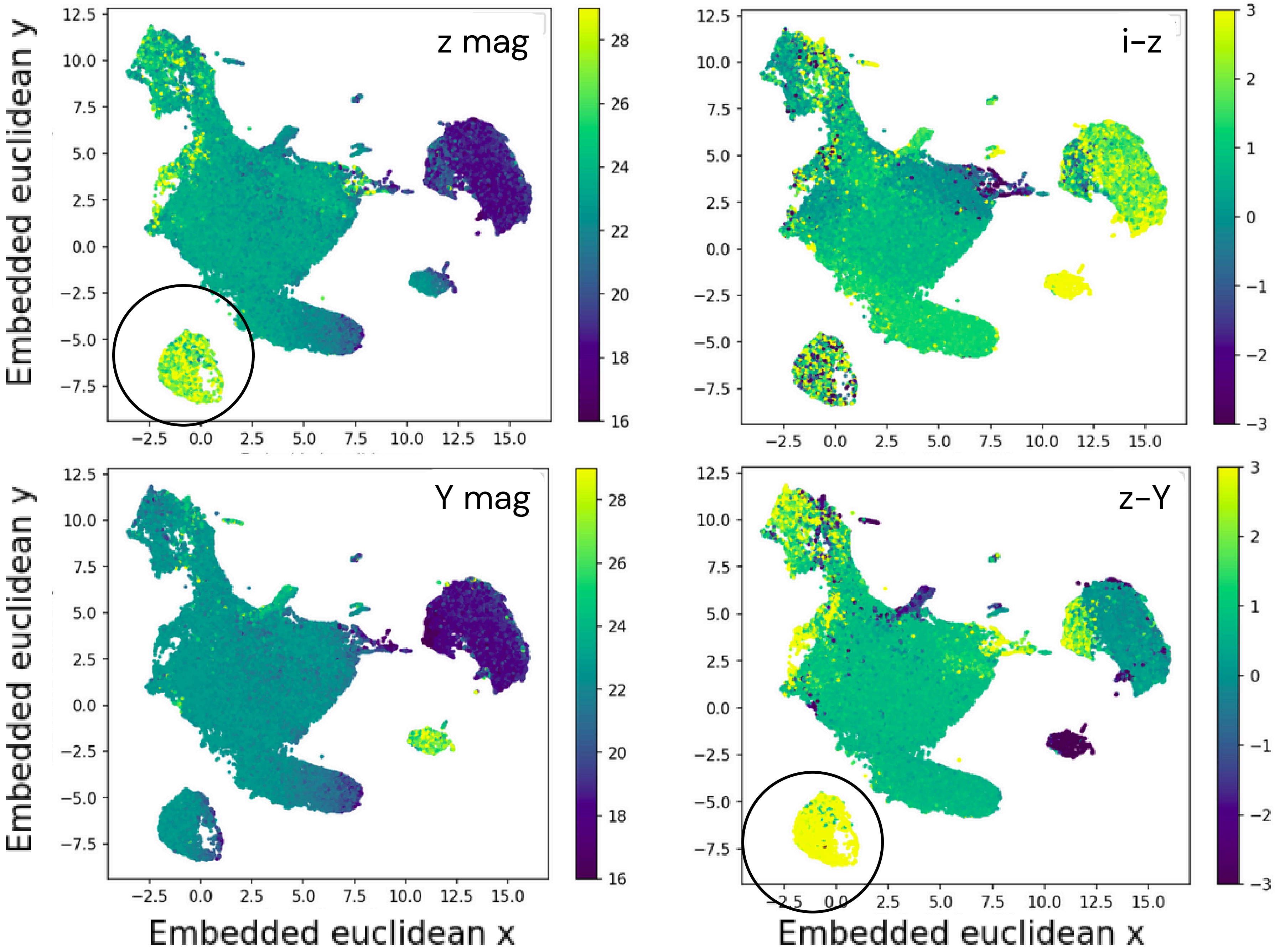
Self-supervised contrastive learning

216.53808851852585 -1.5044673363370582



Preliminary results: VHS_DR4 constraint

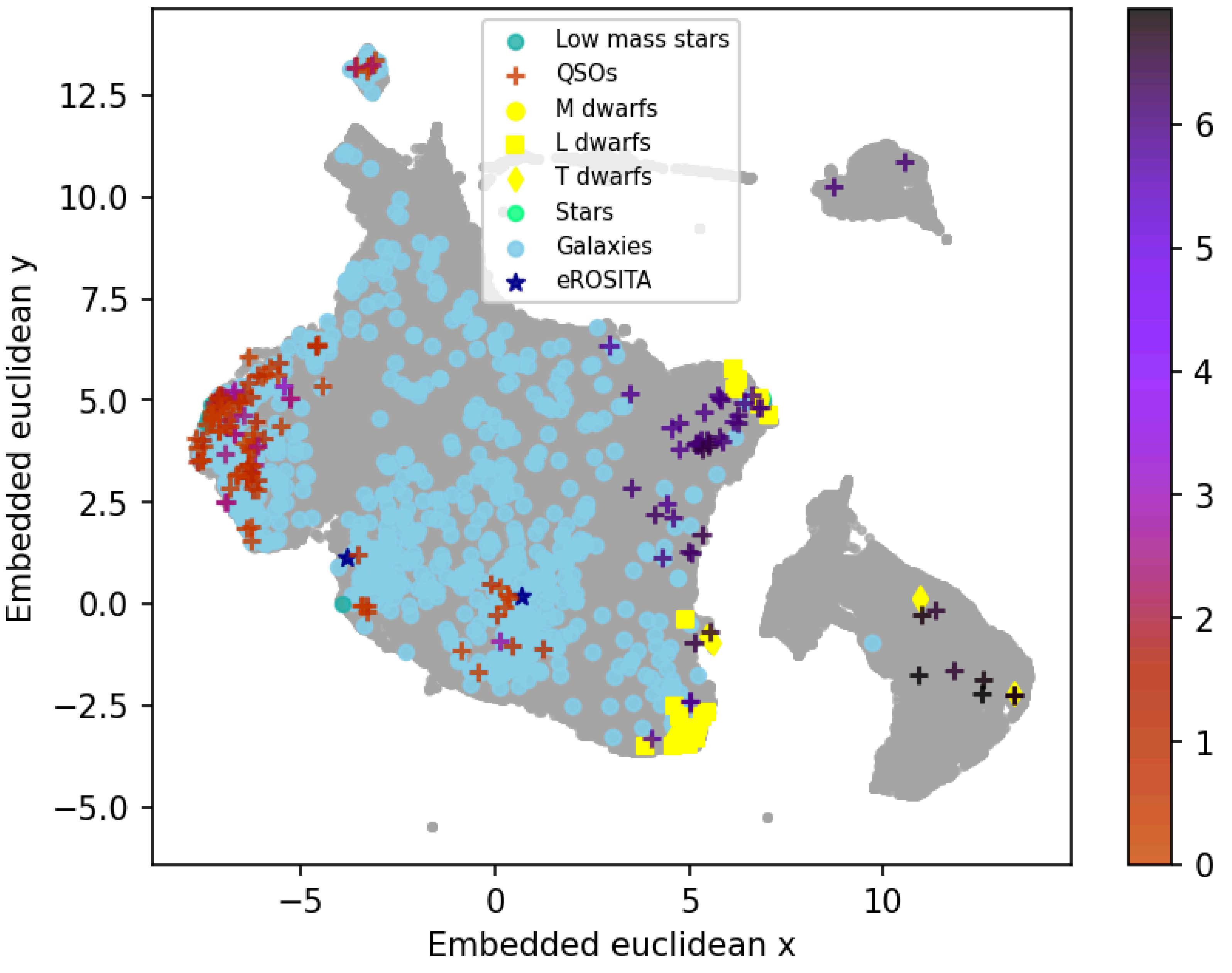


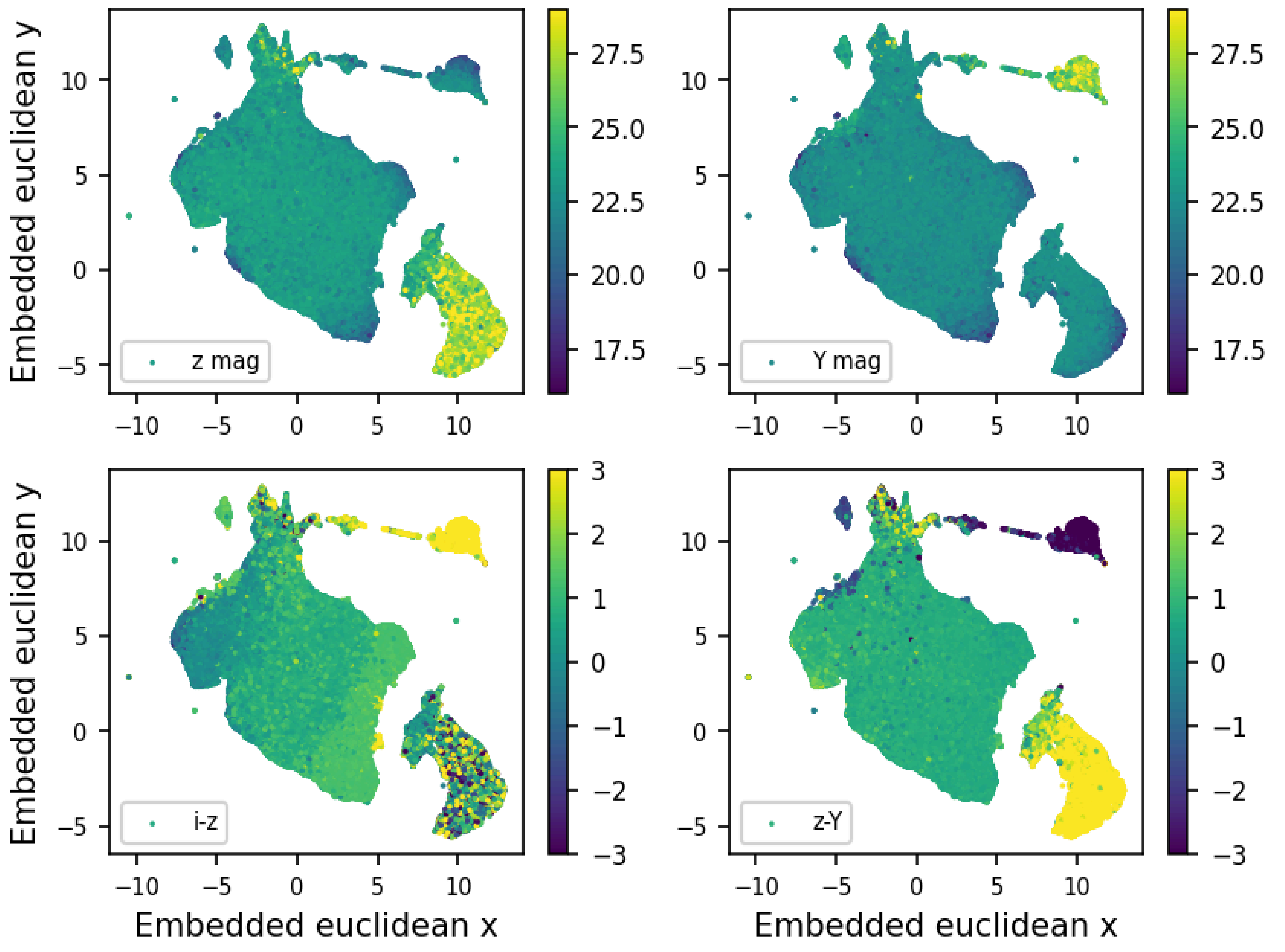


Island of z-dropouts

Self-supervised contrastive learning

mag cut catalog, 6x6 arcsec, individual norm





Query in DECaLS