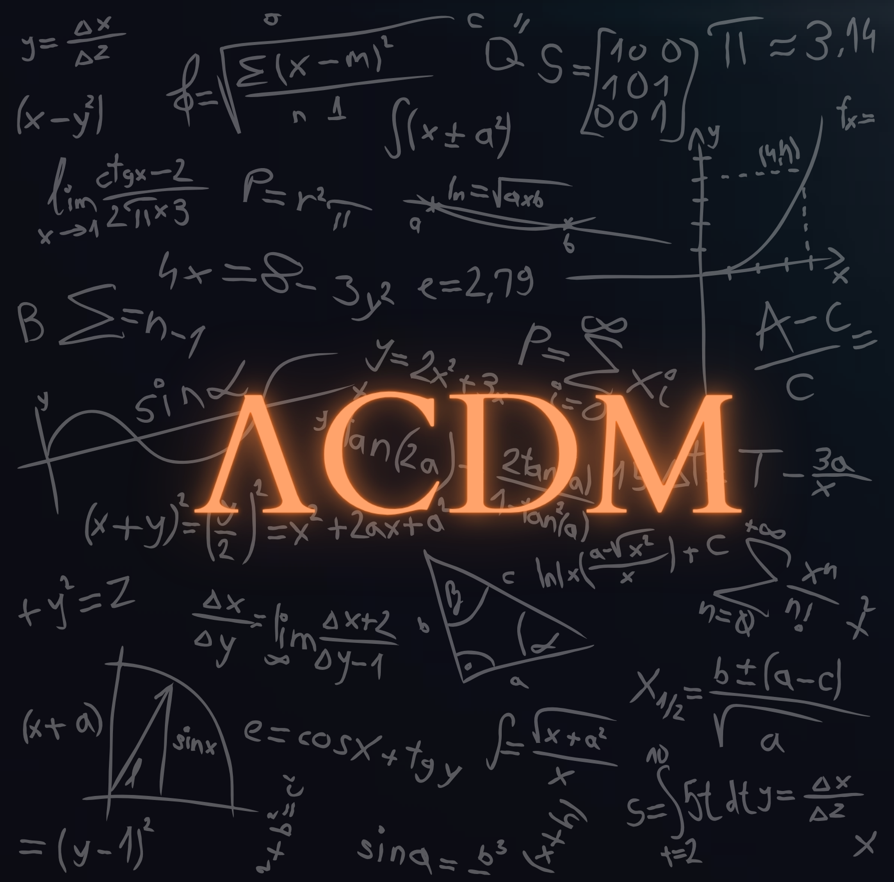
Galaxy redsfhit estimation from multi-band images with Deep Learning

Strengths and challenges

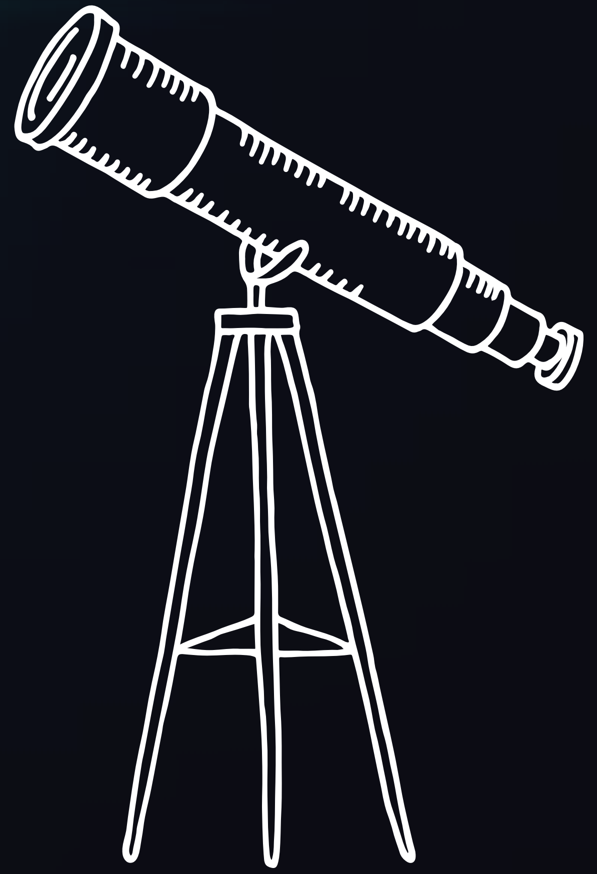
Reda Ait Ouahmed
Stéphane Arnouts, Jérôme Pasquet, Marie Treyer

Observing and understanding the Universe

To understand the Universe



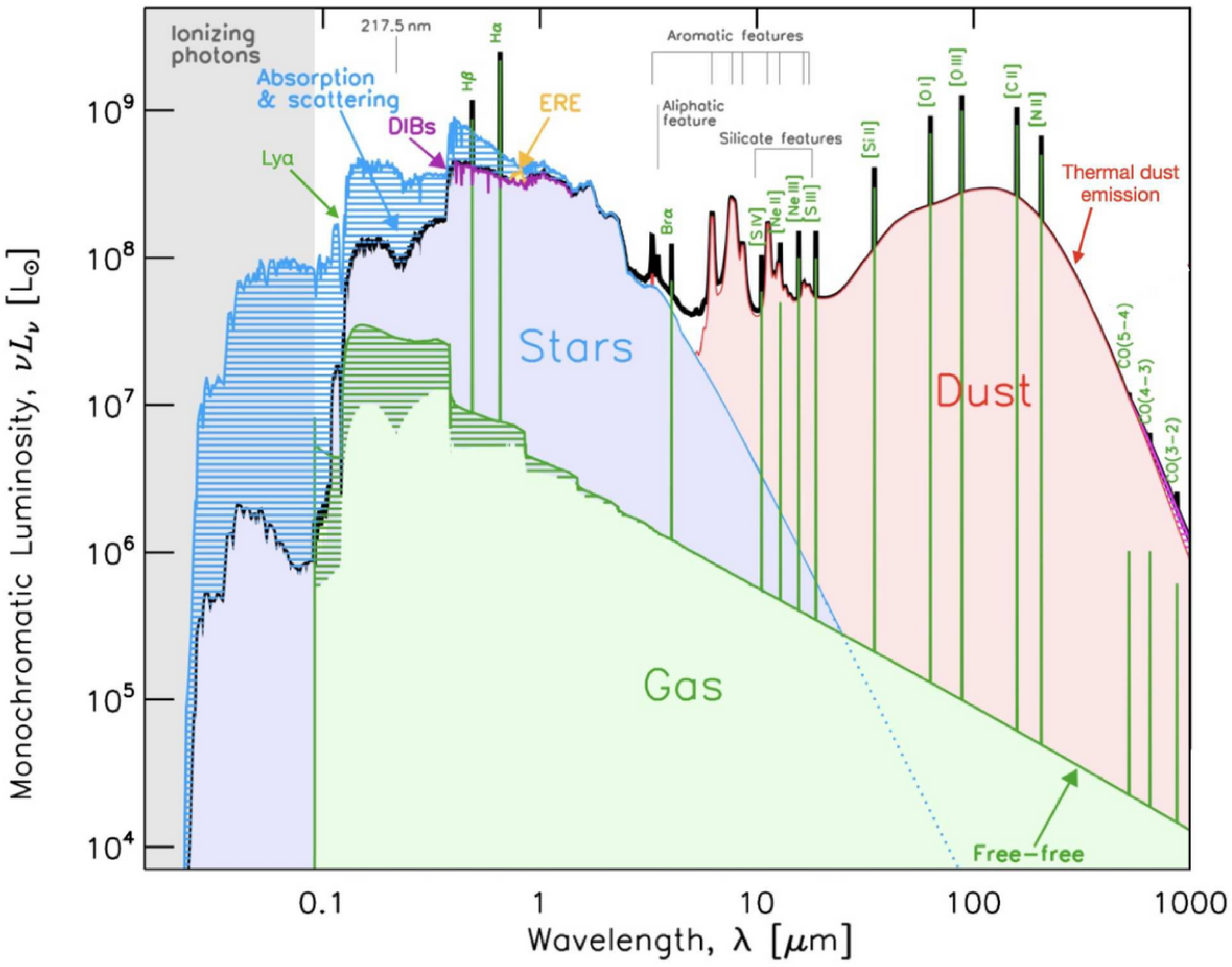
Theoritical models **constrained** by observations



FROM 2D TO 3D

How can we measure galaxy distances ?

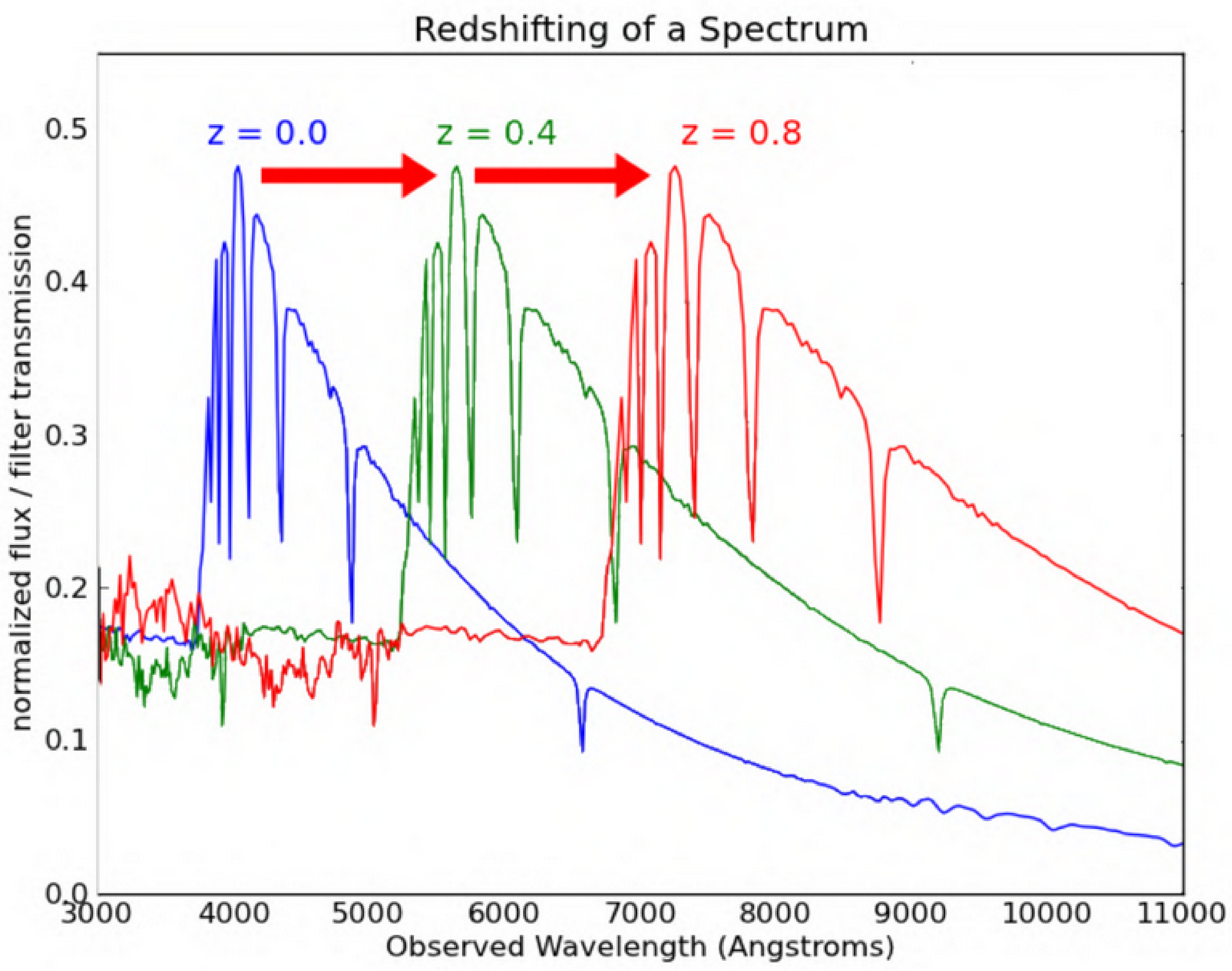
Galaxy Spectrum



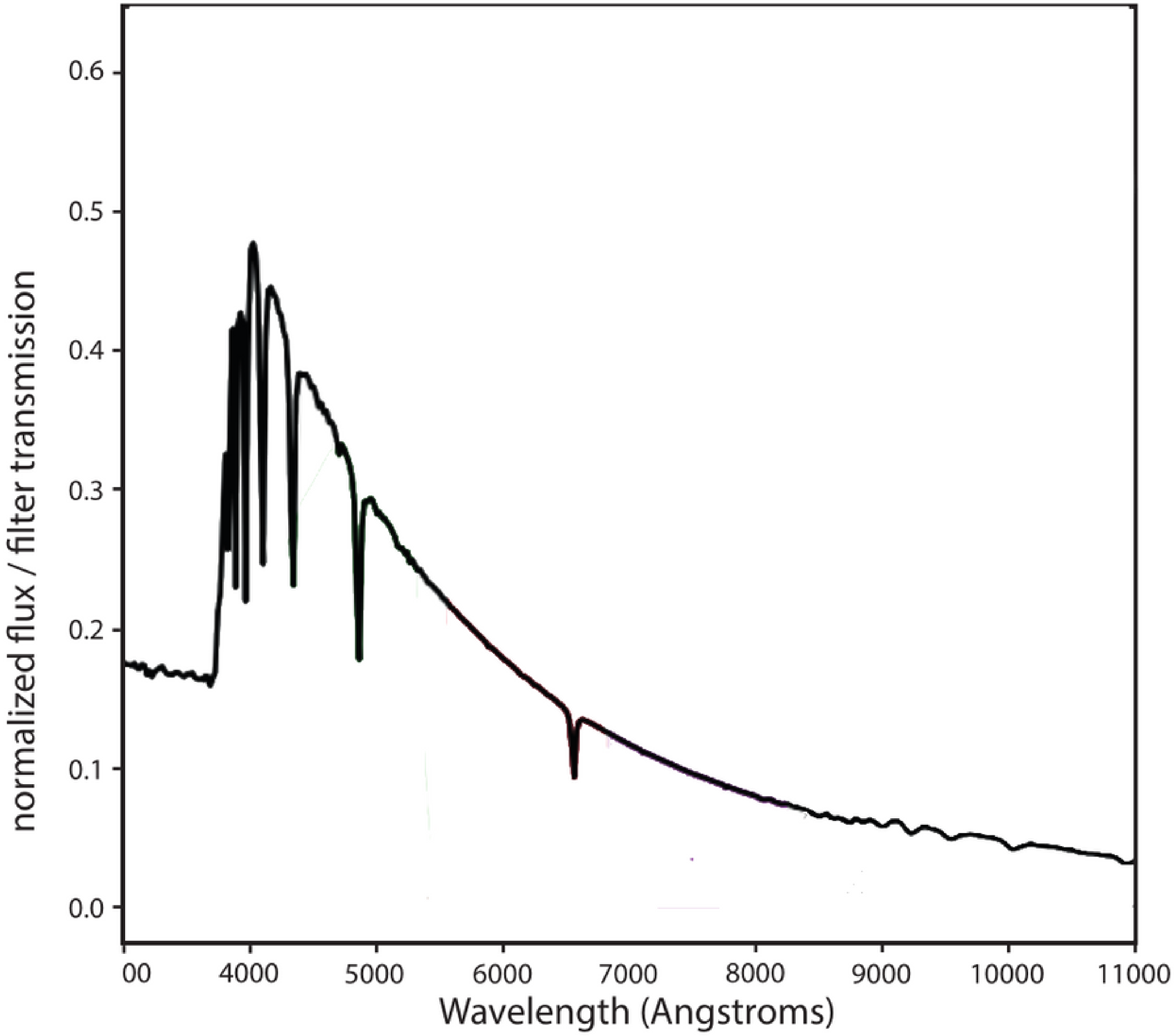
Adapted from Galliano et al. (2018)

Galaxy Spectrum Redshifted

$$z = (\lambda_{obs} - \lambda_{rest}) / \lambda_{rest}$$

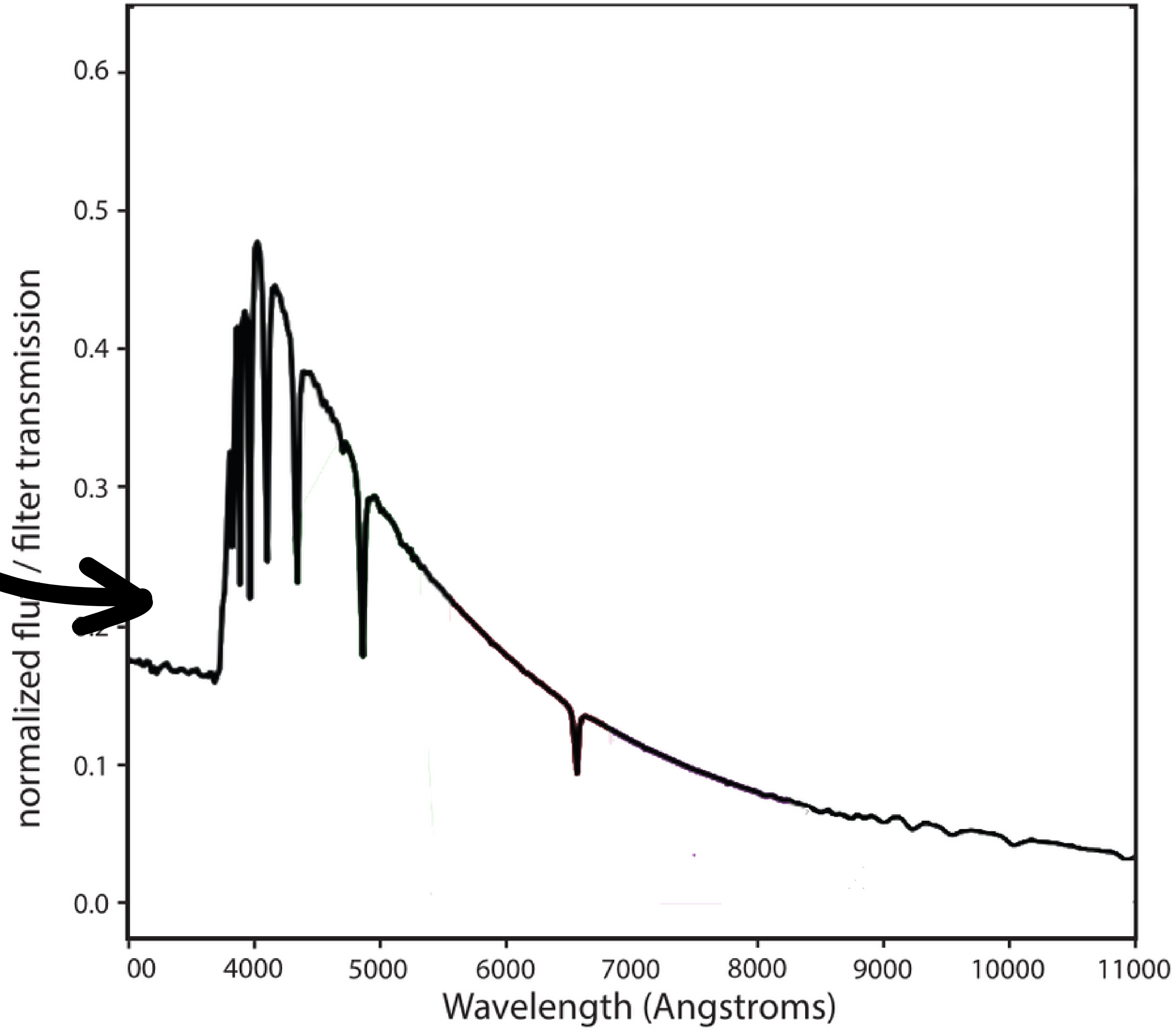


Spectroscopy



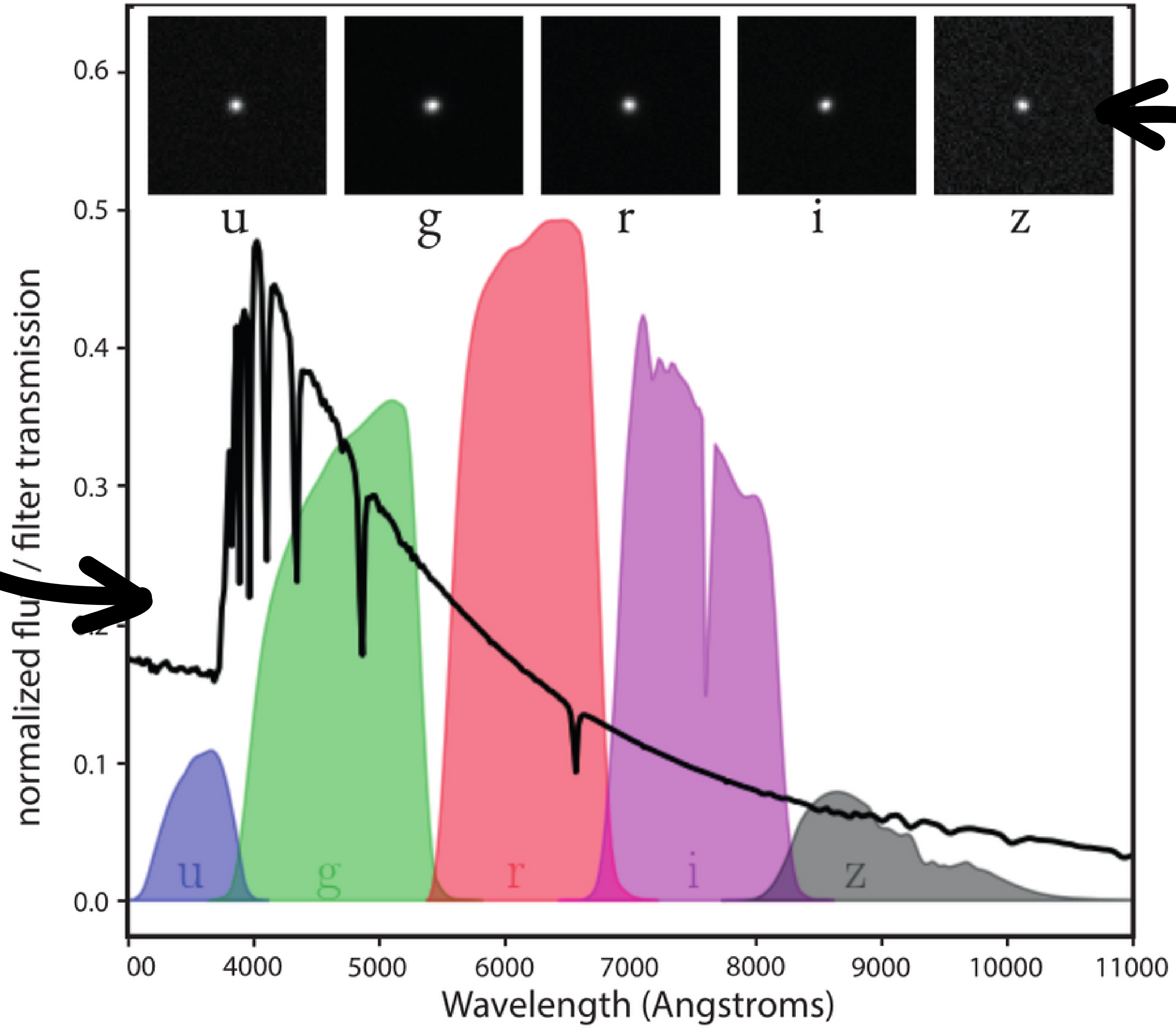
Spectroscopy

- High resolution view
- Precise redshift estimation
- Too slow



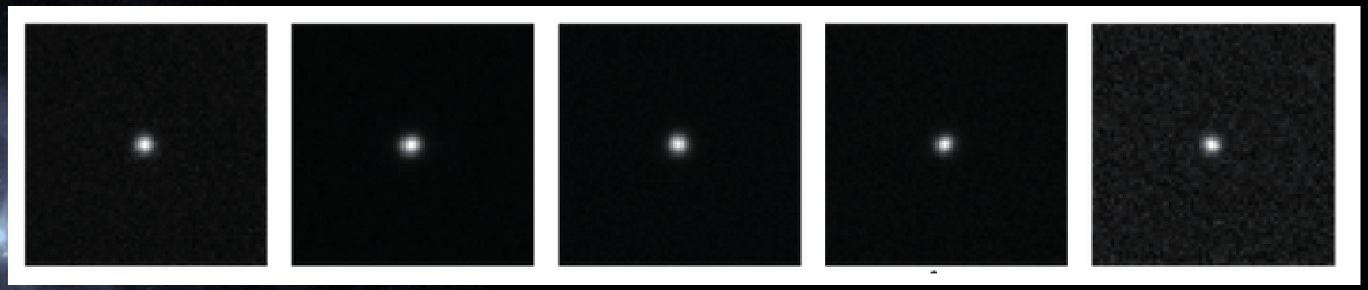
Spectroscopy and photometry

- High resolution view
- Precise redshift estimation
- Too slow



- Low resolution view
- Lower precision of redshift estimation
- Fast

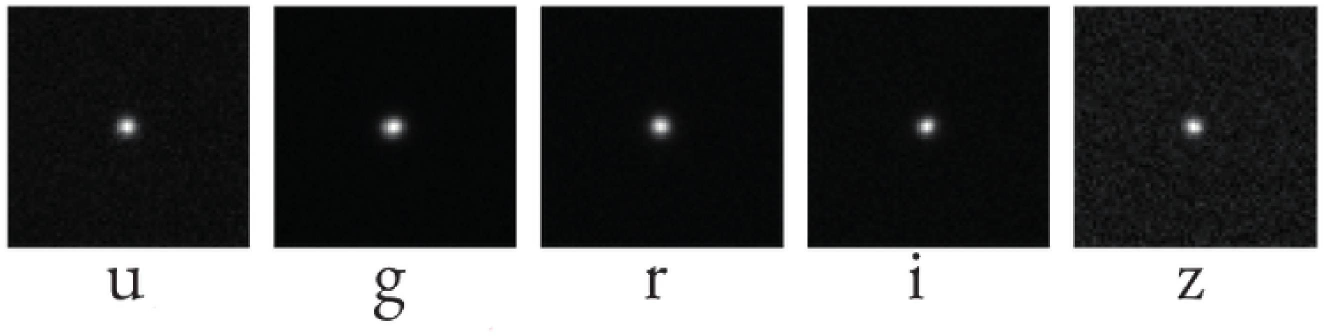
Redshift estimation from multiband photometric images



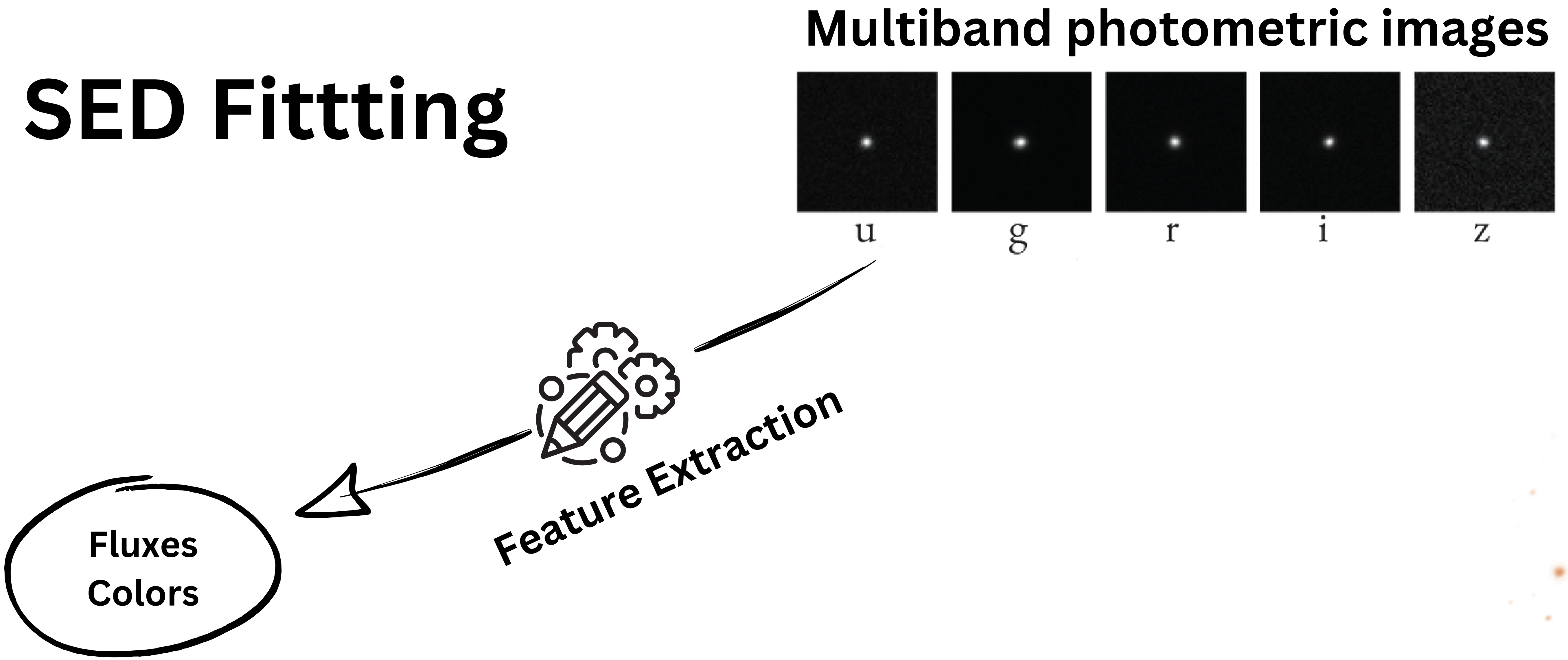
- SED Fittting
- Machine Learning
- Deep Learning

SED Fittting

Multiband photometric images

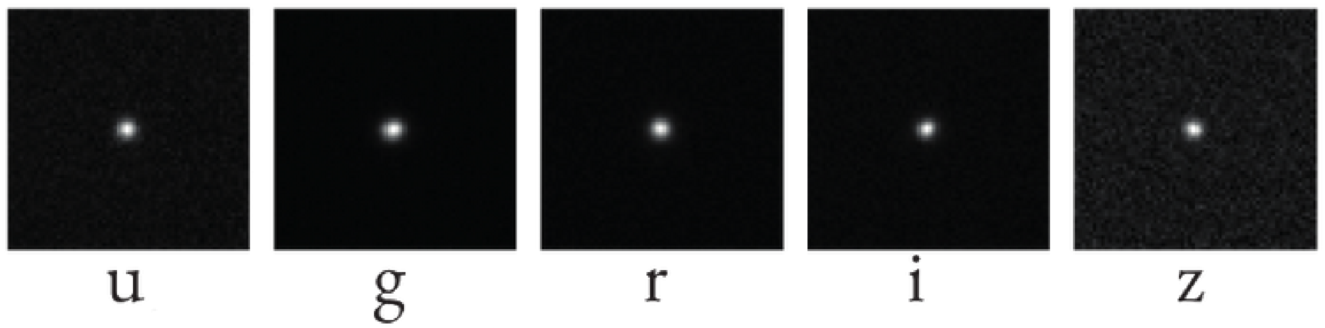


SED Fitttting



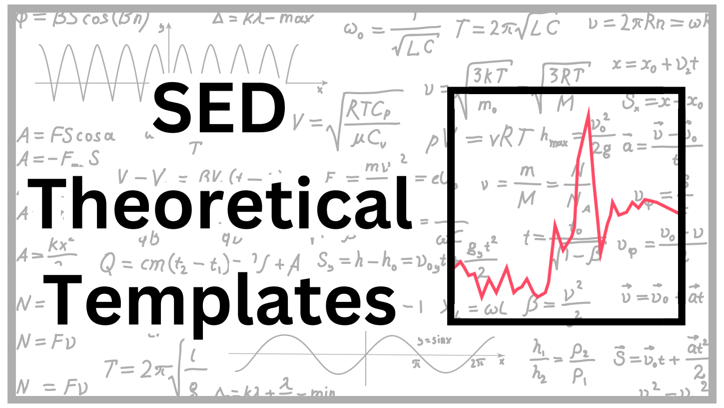
SED Fittting

Multiband photometric images



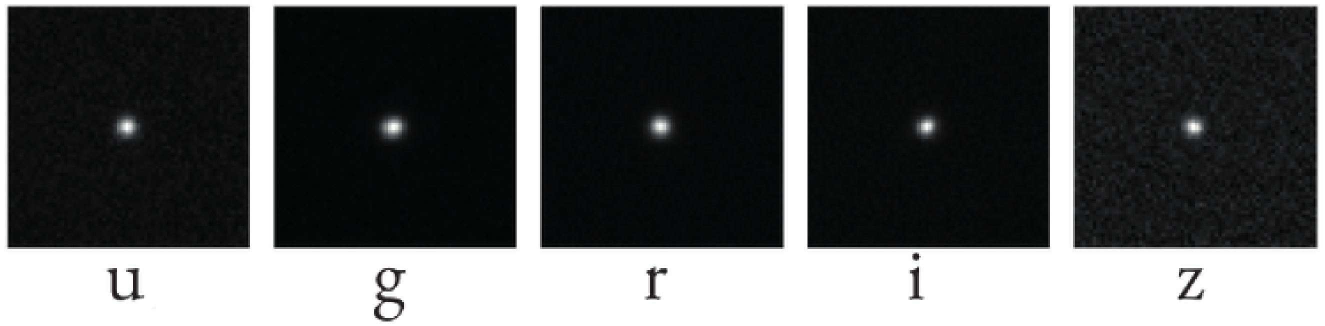
Feature Extraction

Fluxes
Colors



SED Fittting

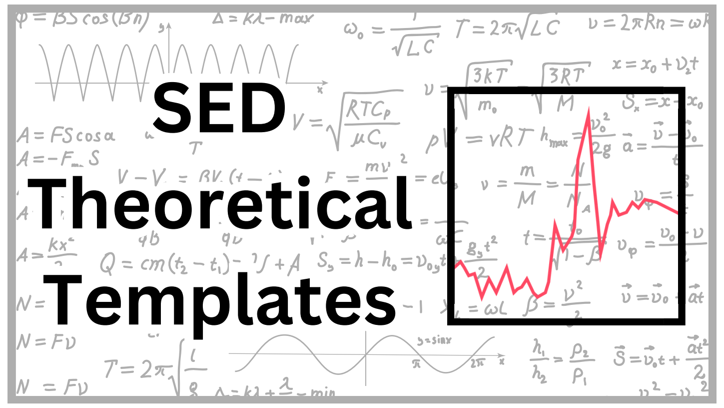
Multiband photometric images



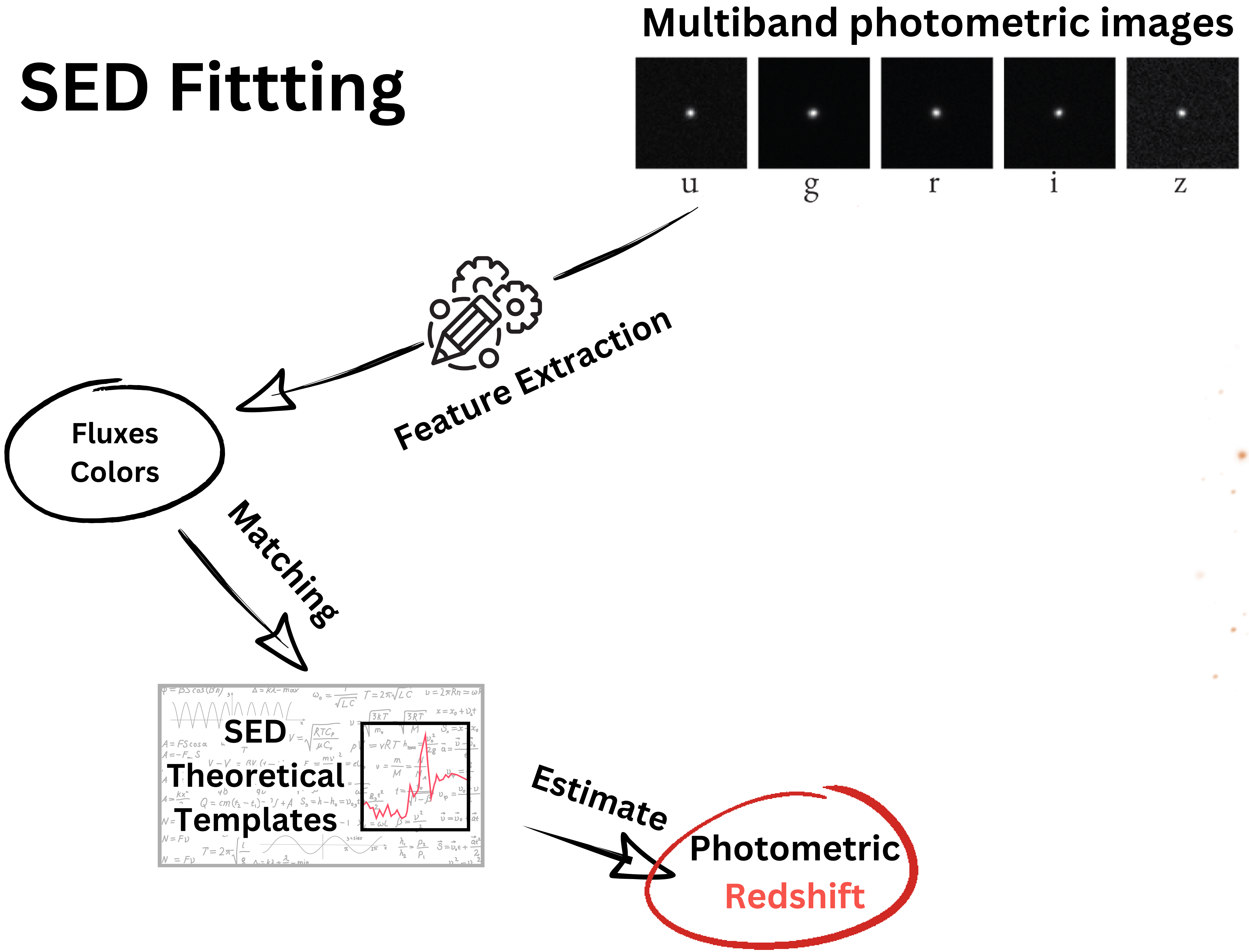
Feature Extraction

Fluxes
Colors

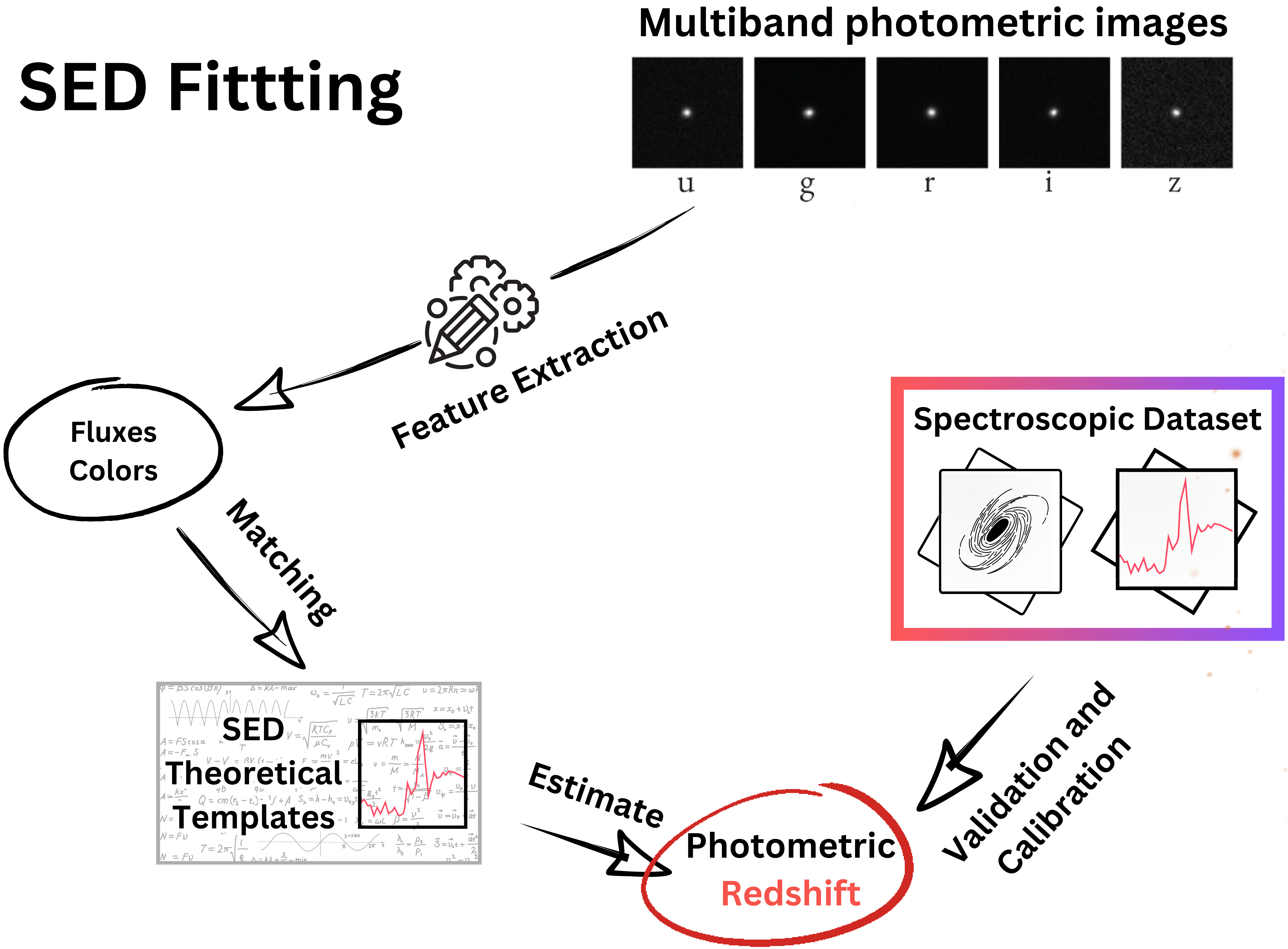
Matching



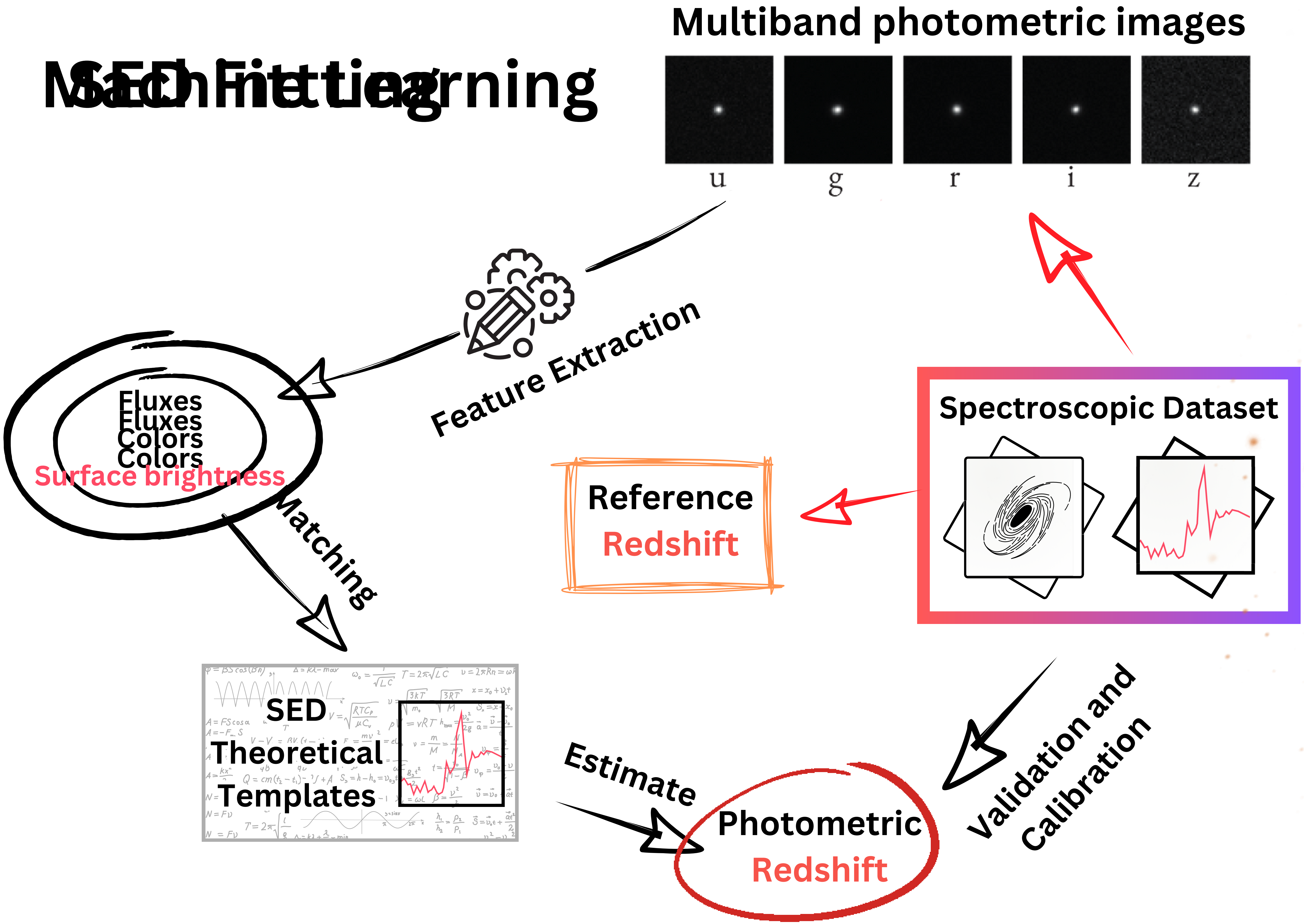
SED Fitttting



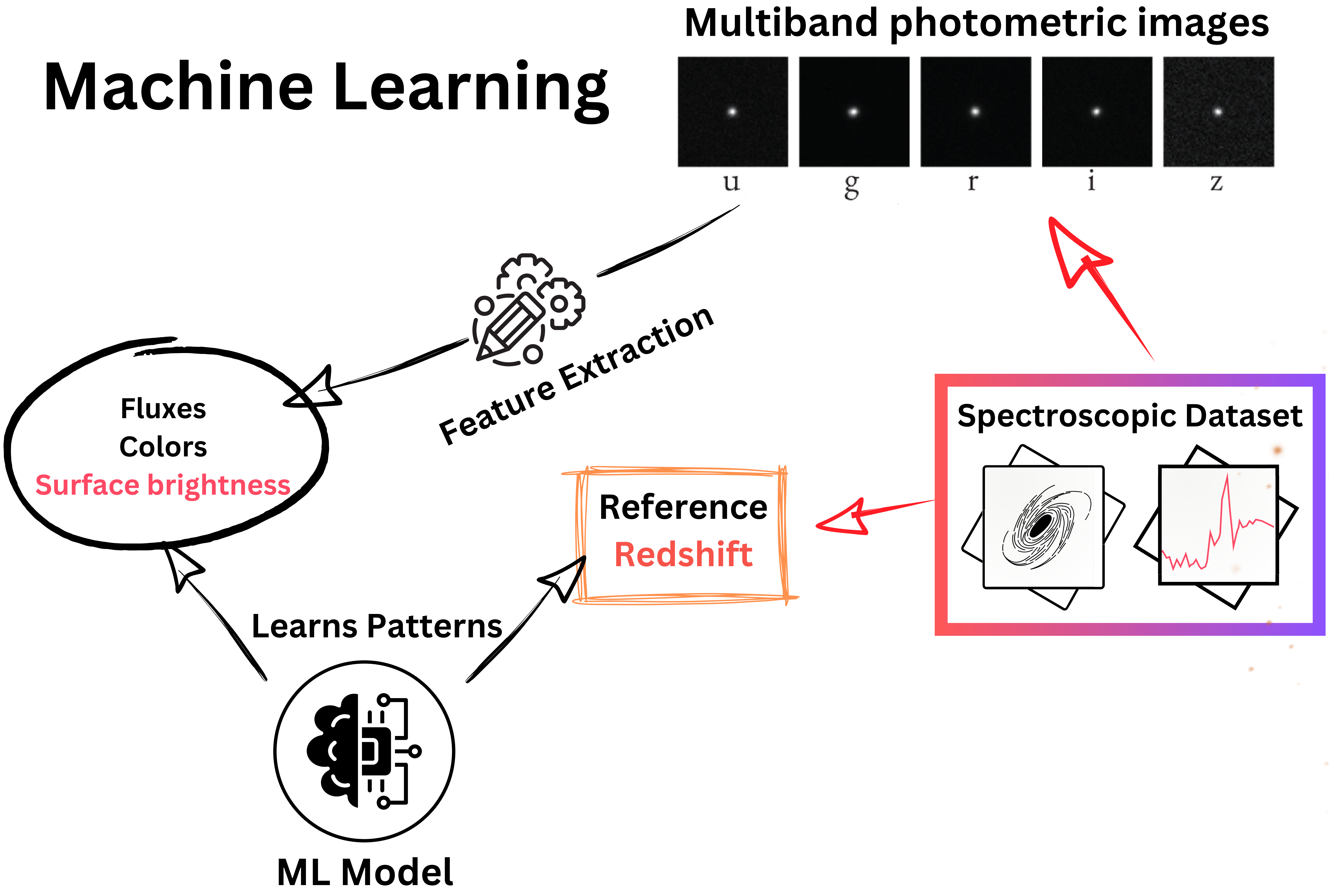
SED Fitttting



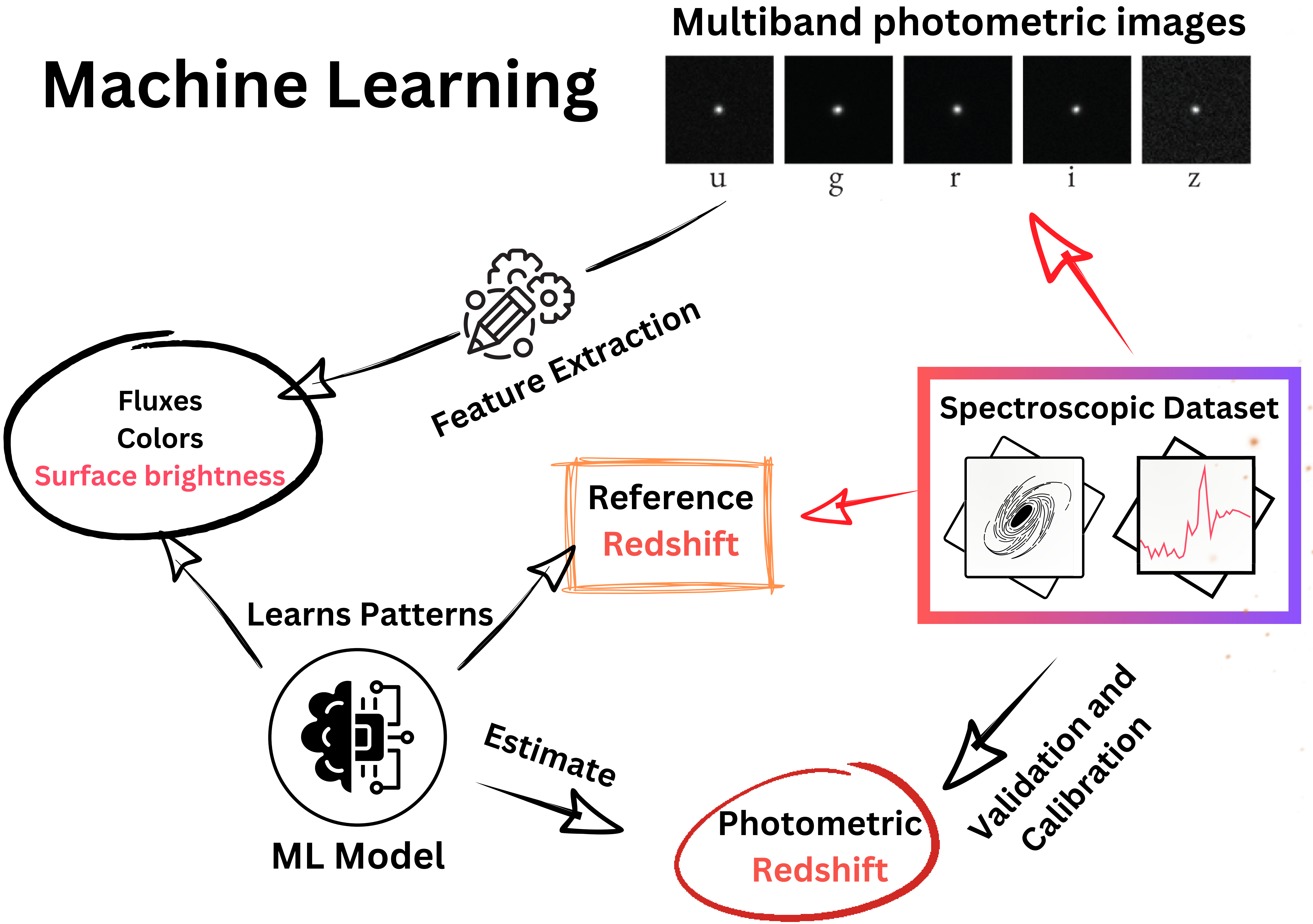
SED Fitting
Machine Learning



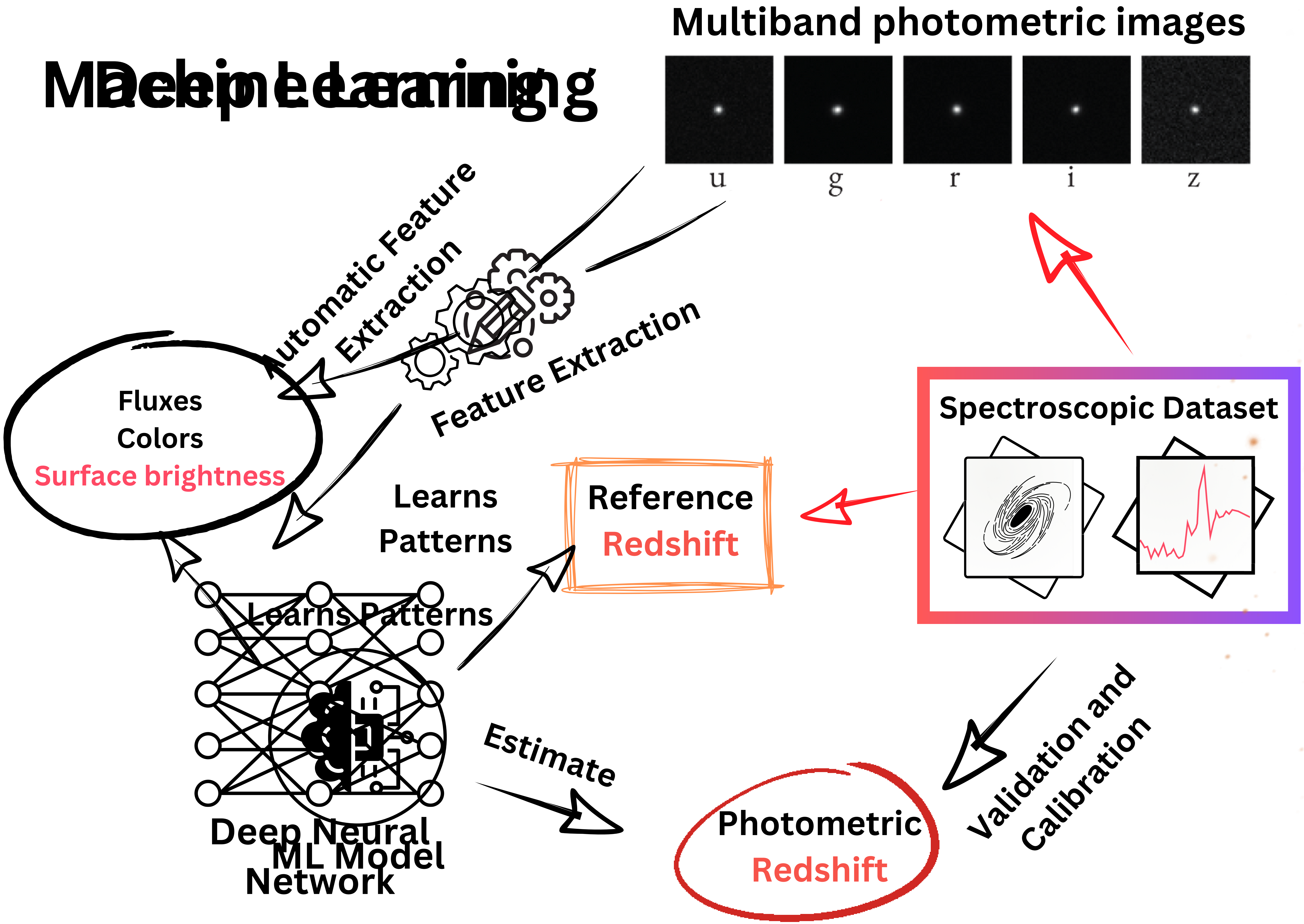
Machine Learning



Machine Learning



Machine Learning / Deep Learning



Understanding the Universe

Theortical Models

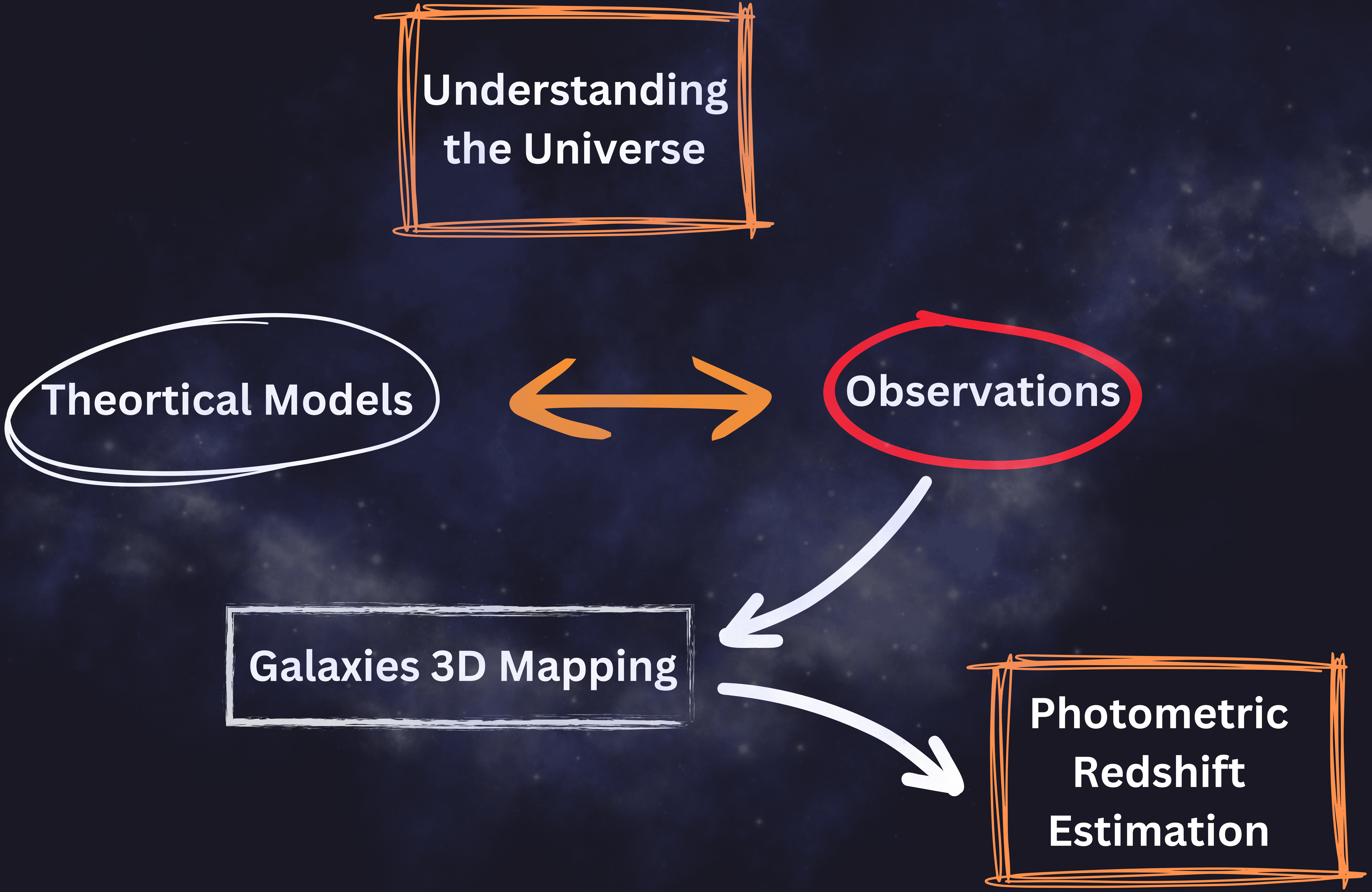
Observations

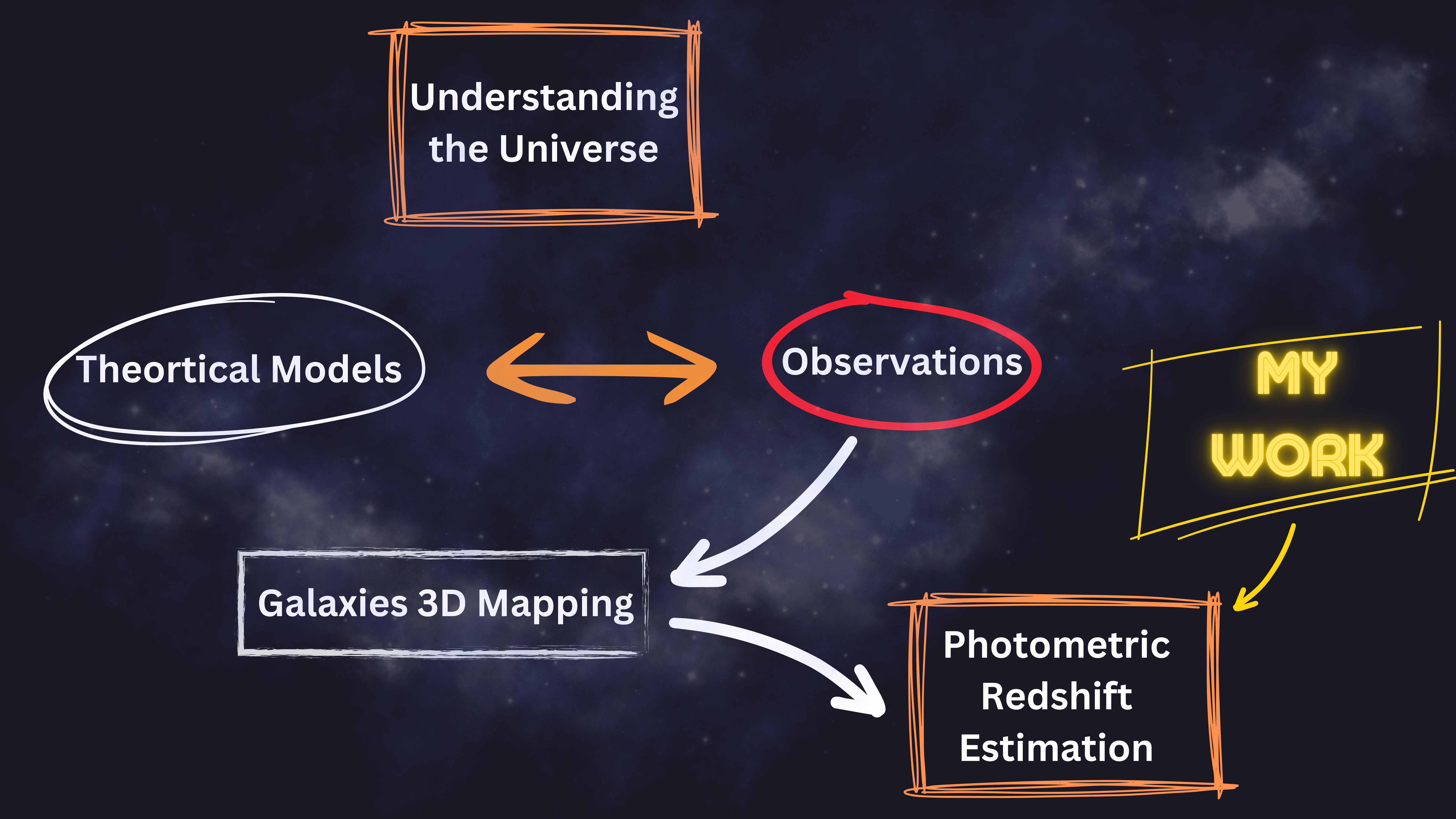
Understanding the Universe

Theortical Models

Observations

Galaxies 3D Mapping


Understanding the Universe
Theortical Models
Observations
Galaxies 3D Mapping
Photometric Redshift Estimation

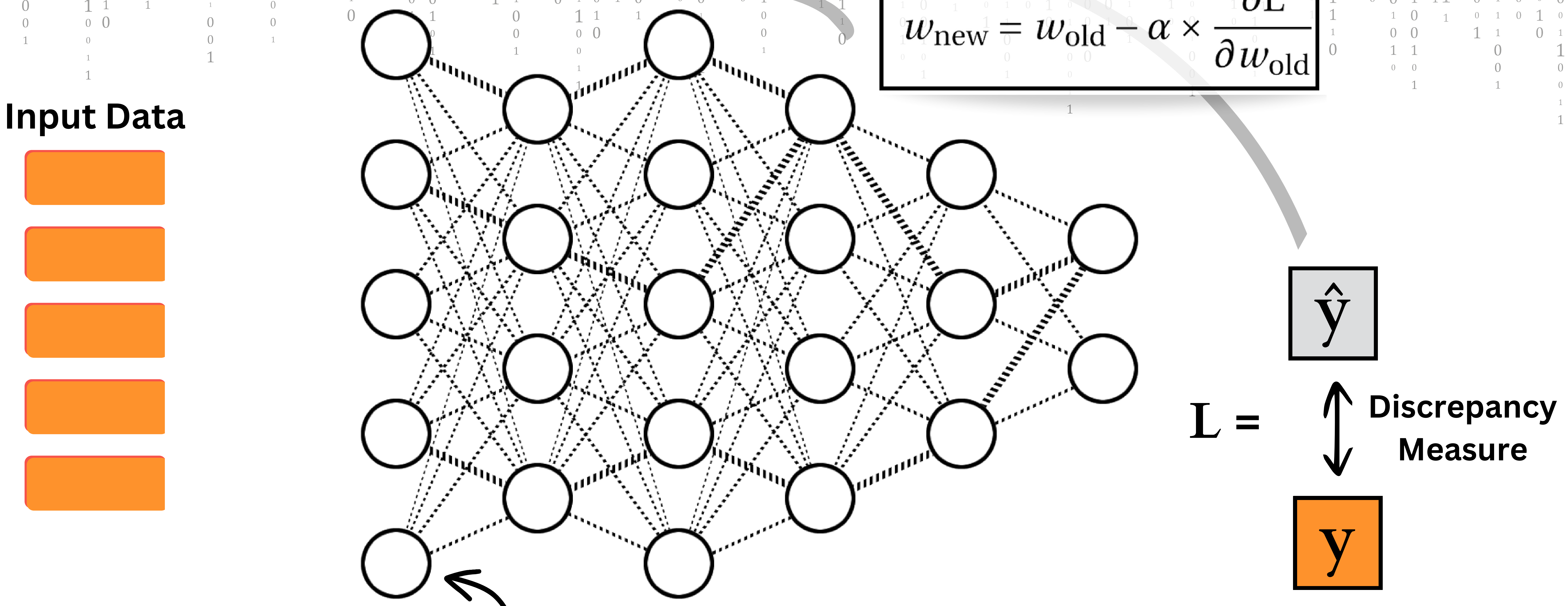

Understanding the Universe
Theortical Models
Observations
MY WORK
Galaxies 3D Mapping
Photometric Redshift Estimation

Presentation Plan

- Context Of This Work : Deep Learning And Photometric Redshifts
- First Contribution : Multimodality For Improved Photometric Redshifts
- Second Contribution : Application To The HSC Deep Survey

Deep Neural Networks

$$w_{\text{new}} = w_{\text{old}} - \alpha \times \frac{\partial \text{L}}{\partial w_{\text{old}}}$$

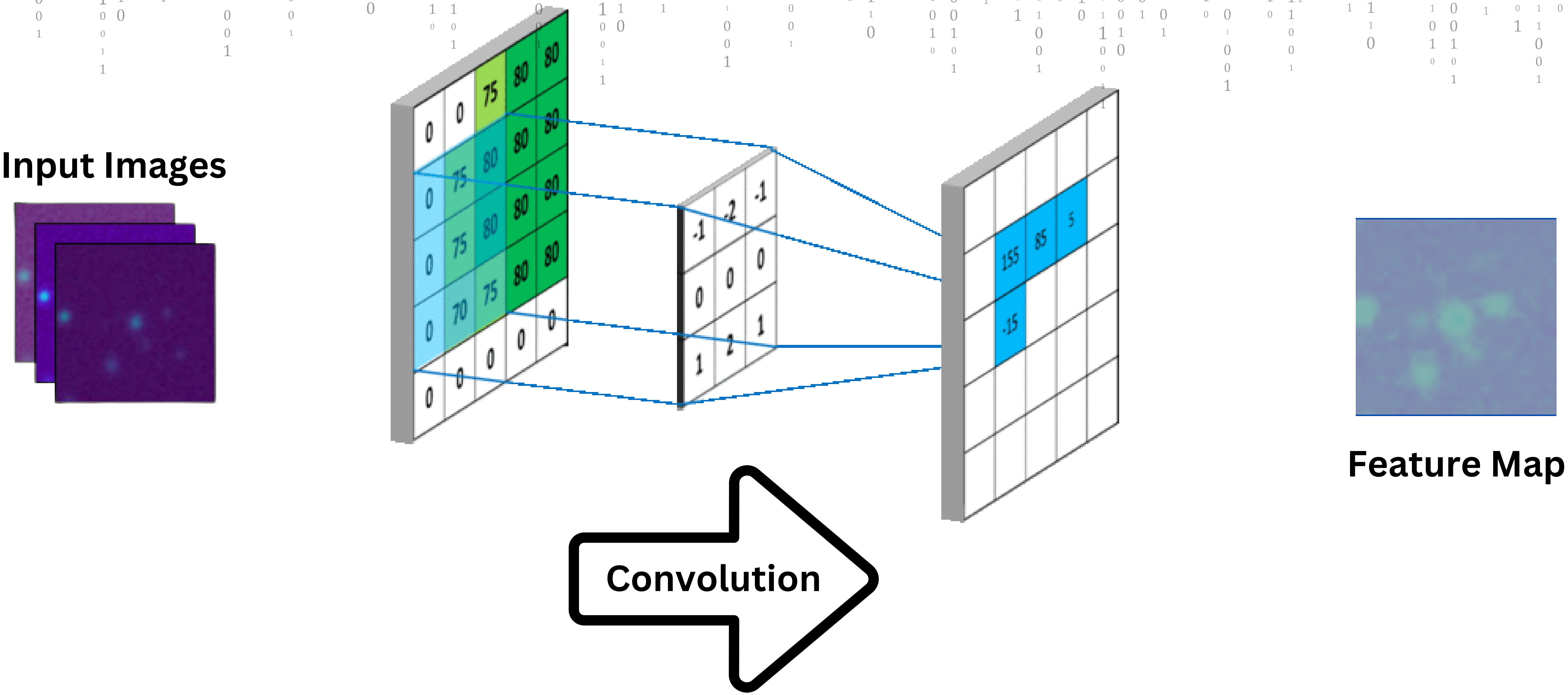


$$\hat{y} = \phi(w_1 \times x_1 + w_2 \times x_2 + \ldots + w_n \times x_n + b)$$

Context Of This Work

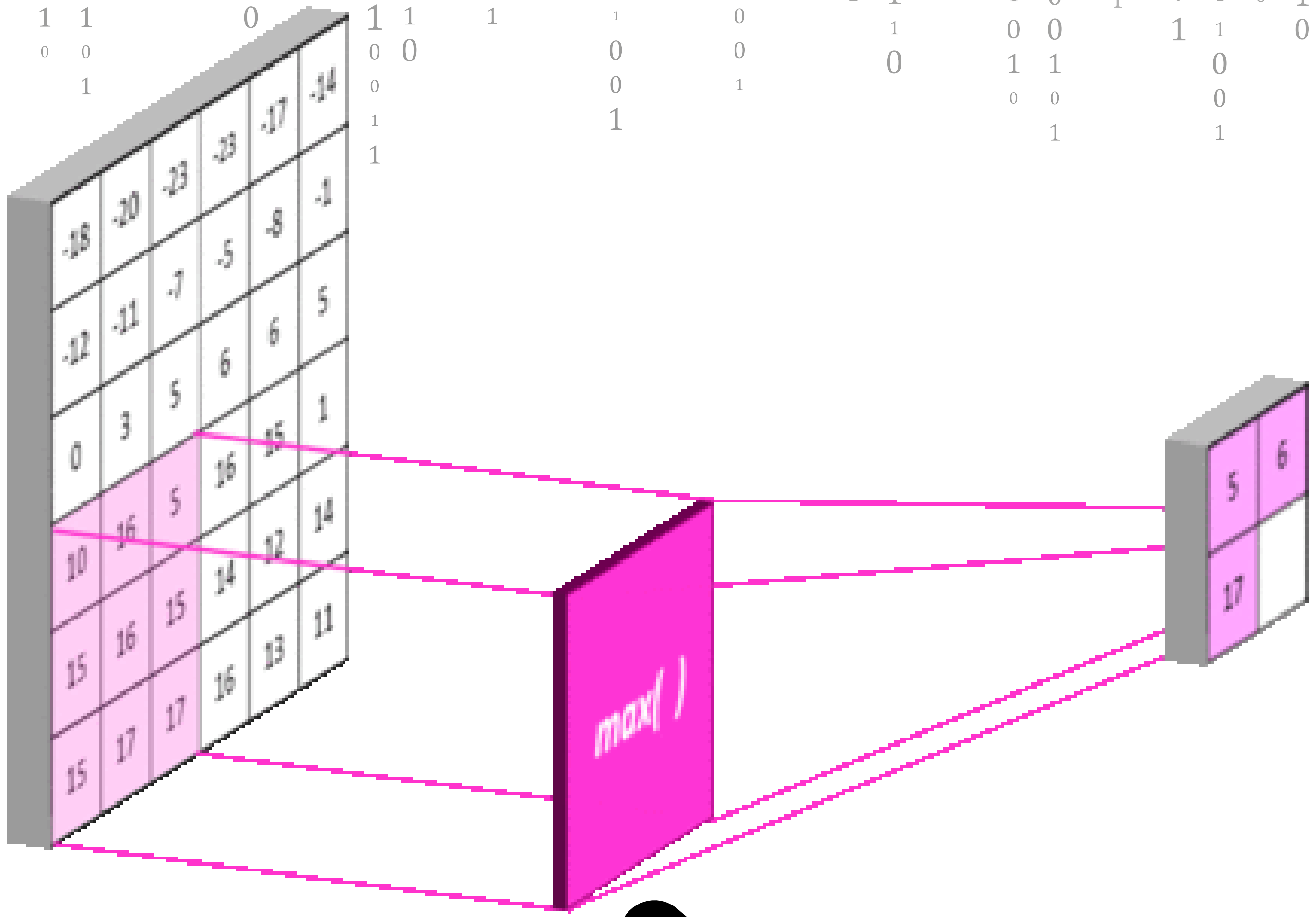
Convolutional Neural Networks

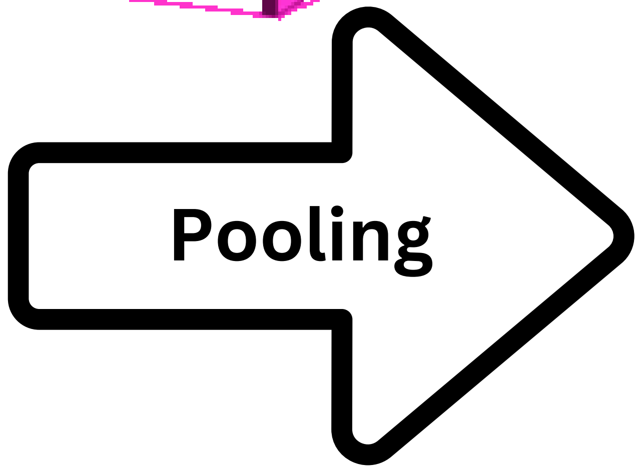
Context Of This Work



Convolutional Neural Networks

Feature Map

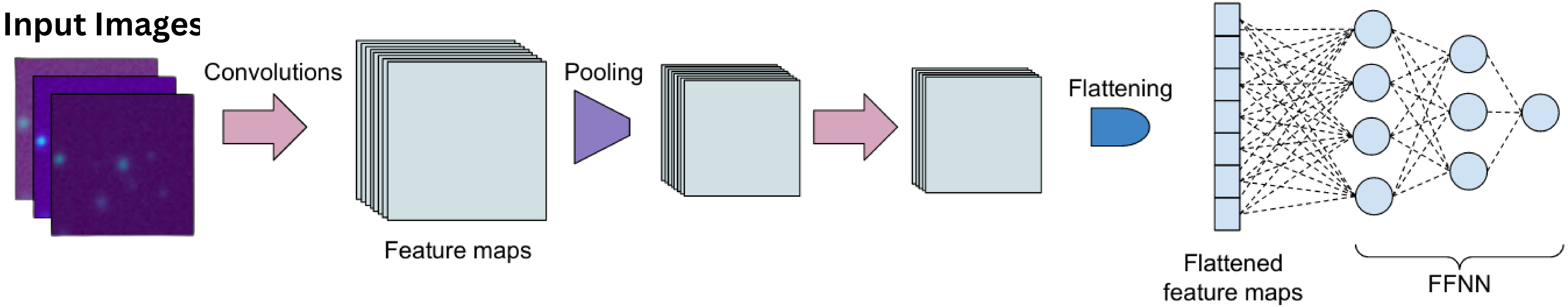




Reduced Feature Map

Context Of This Work

Convolutional Neural Networks



Context Of This Work

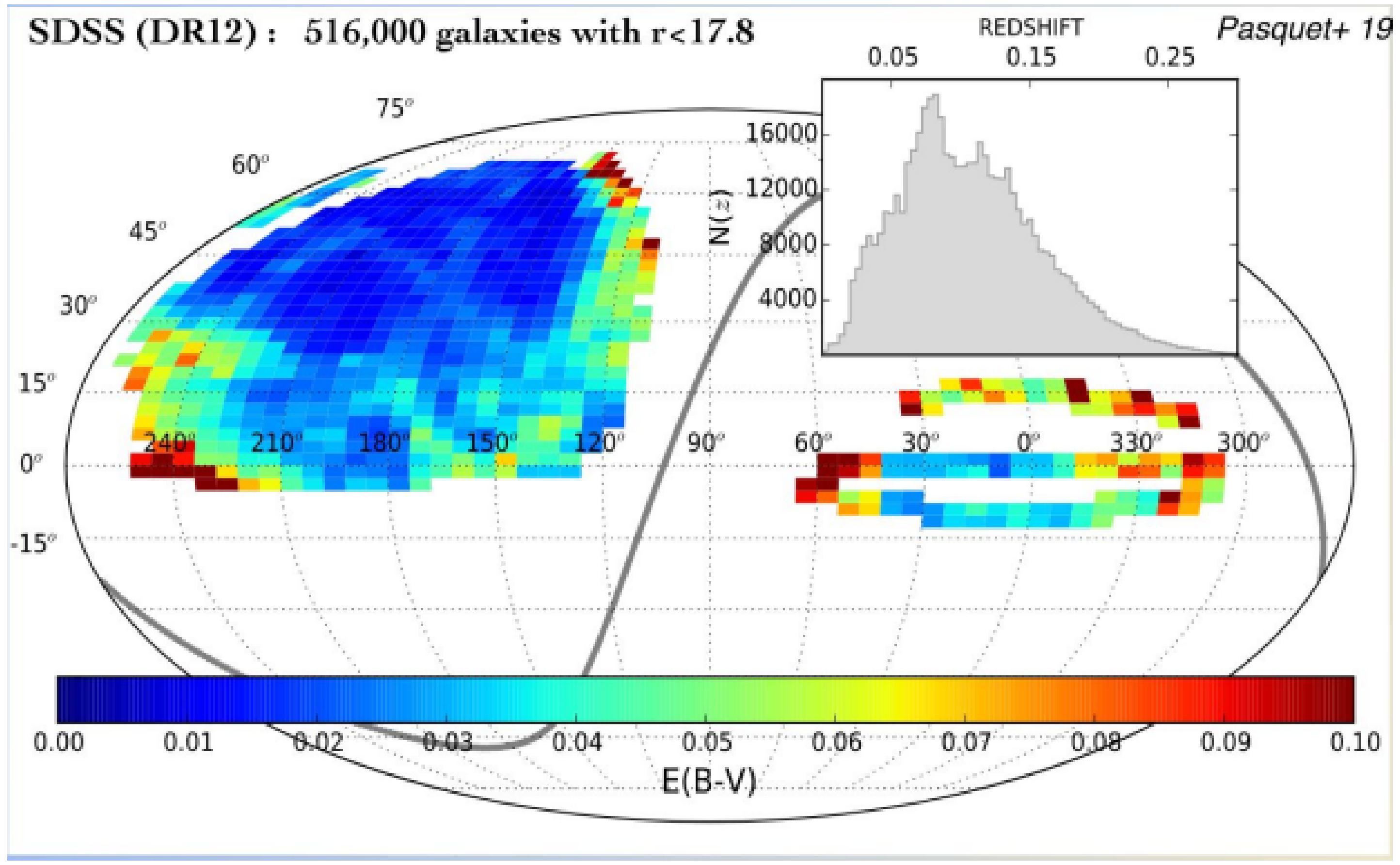
Metrics of Photometric Redshifts

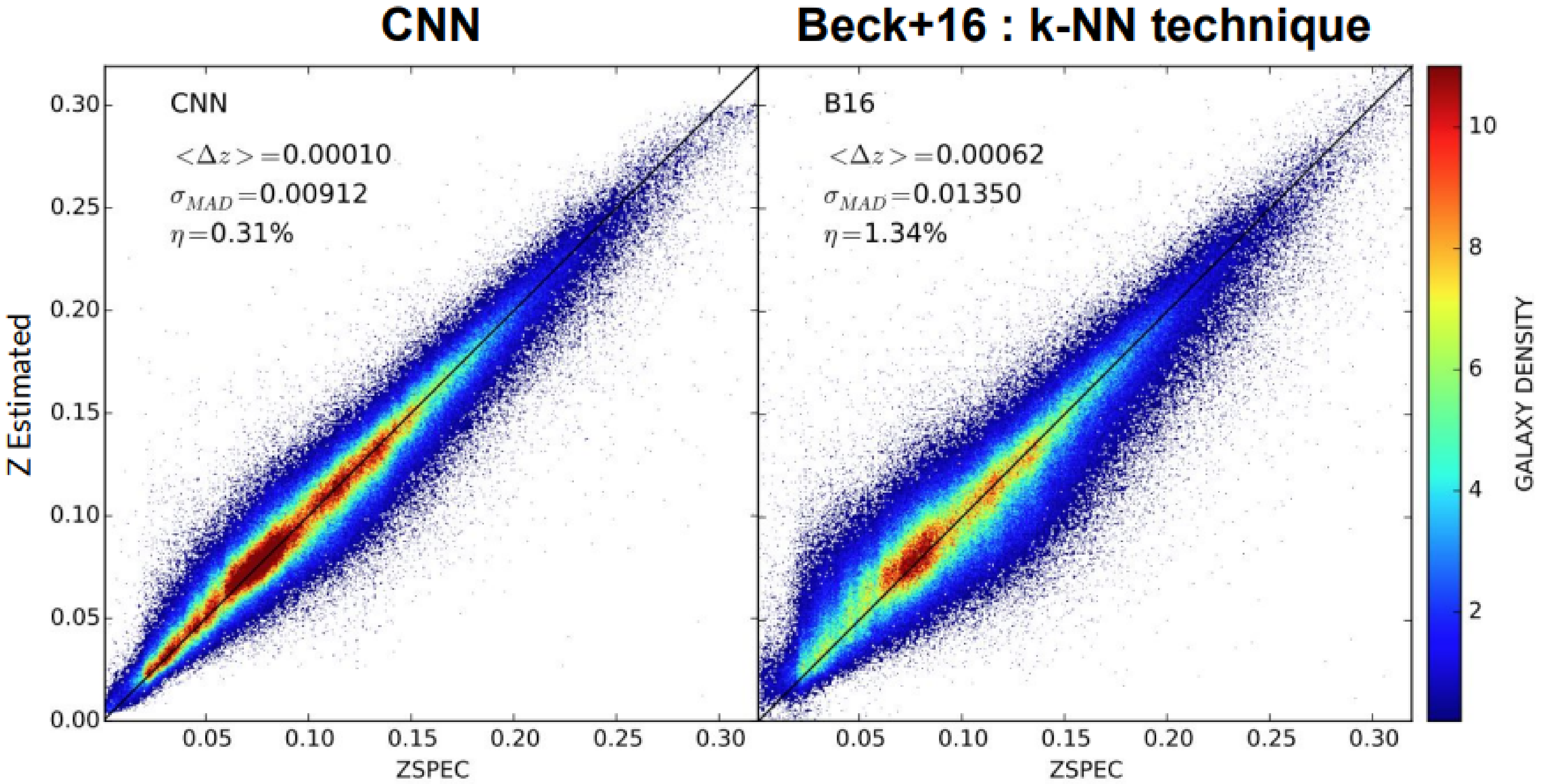
- Residuals : $\Delta z = (z_{\text{phot}} - z_{\text{spec}})/(1 + z_{\text{spec}})$
- Normalized Mad (Median absolute deviation) : $\text{Mad} = 1.48 * \text{Median}(|\Delta z - \text{Median}(\Delta z)|)$
- Outliers fraction: fraction of objects with $|\Delta z| \geq 0.15$ (or $|\Delta z| \geq 0.05$ for the SDSS)
- Bias: $\text{Bias} = \text{Mean}(\Delta z)$

Deep Leaning And Photometric Redshifts

Photometric redshifts from SDSS images using a Convolutional Neural Network

Johanna Pasquet[1], E. Bertin[2], M. Treyer[3], S. Arnouts[3] and D. Fouchez[1]





Redshift Estimation With Deep Learning

Capsule Networks

Self-supervised Learning

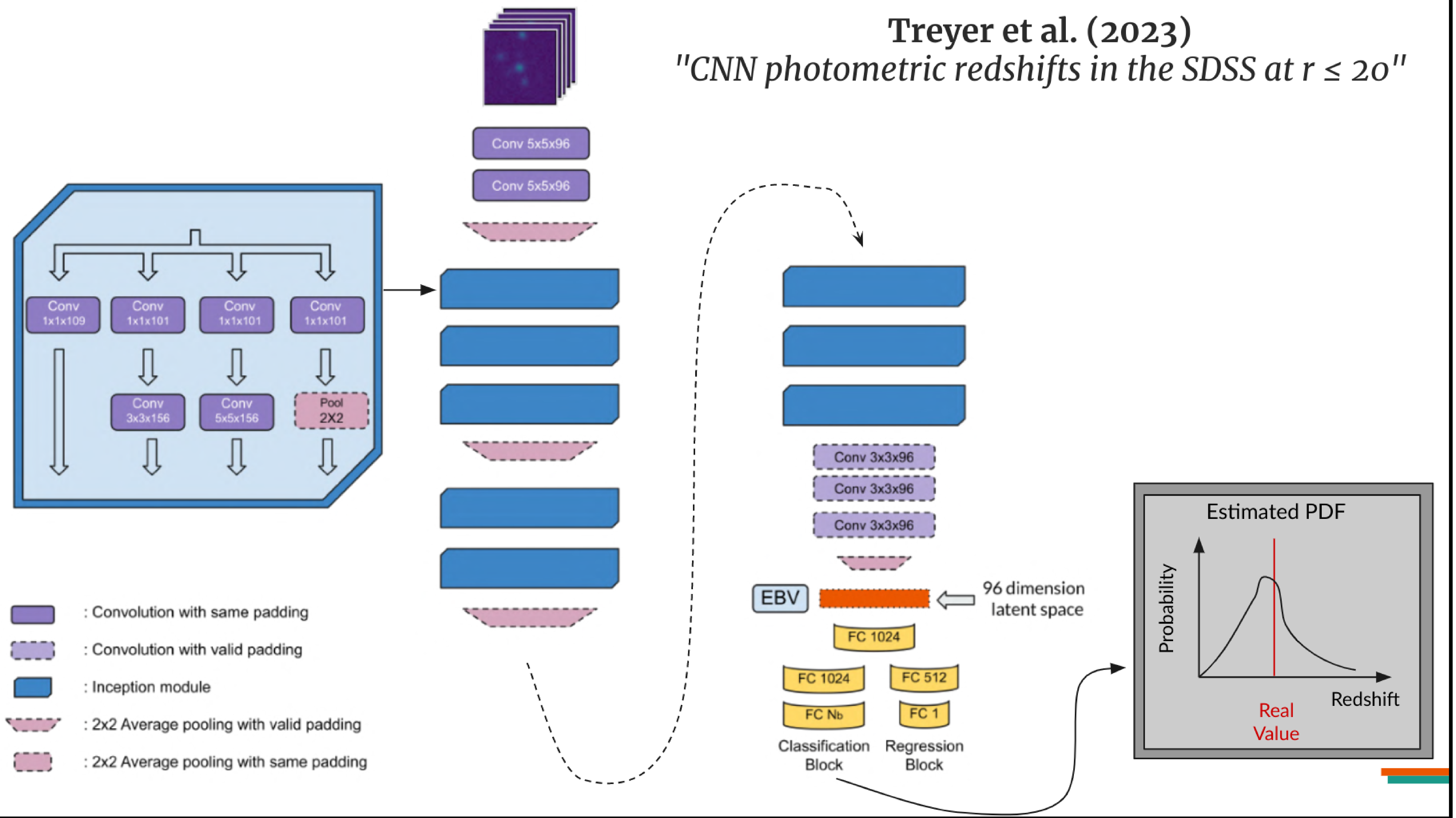
Improved Version of P19

Experiences	σ 10^{-3}	η %	$<\Delta z>$ 10^{-3}
SDSS r < 17.8			
P19	9.08	0.31	**0.04**
Dey et al. (2021)	8.98	0.19	0.07
Hayat et al. (2021)	8.25	0.21	0.1
Treyer et al. (2023)	8.00	0.18	-0.31

Baseline Network

Treyer et al. (2023)

"CNN photometric redshifts in the SDSS at r ≤ 20"



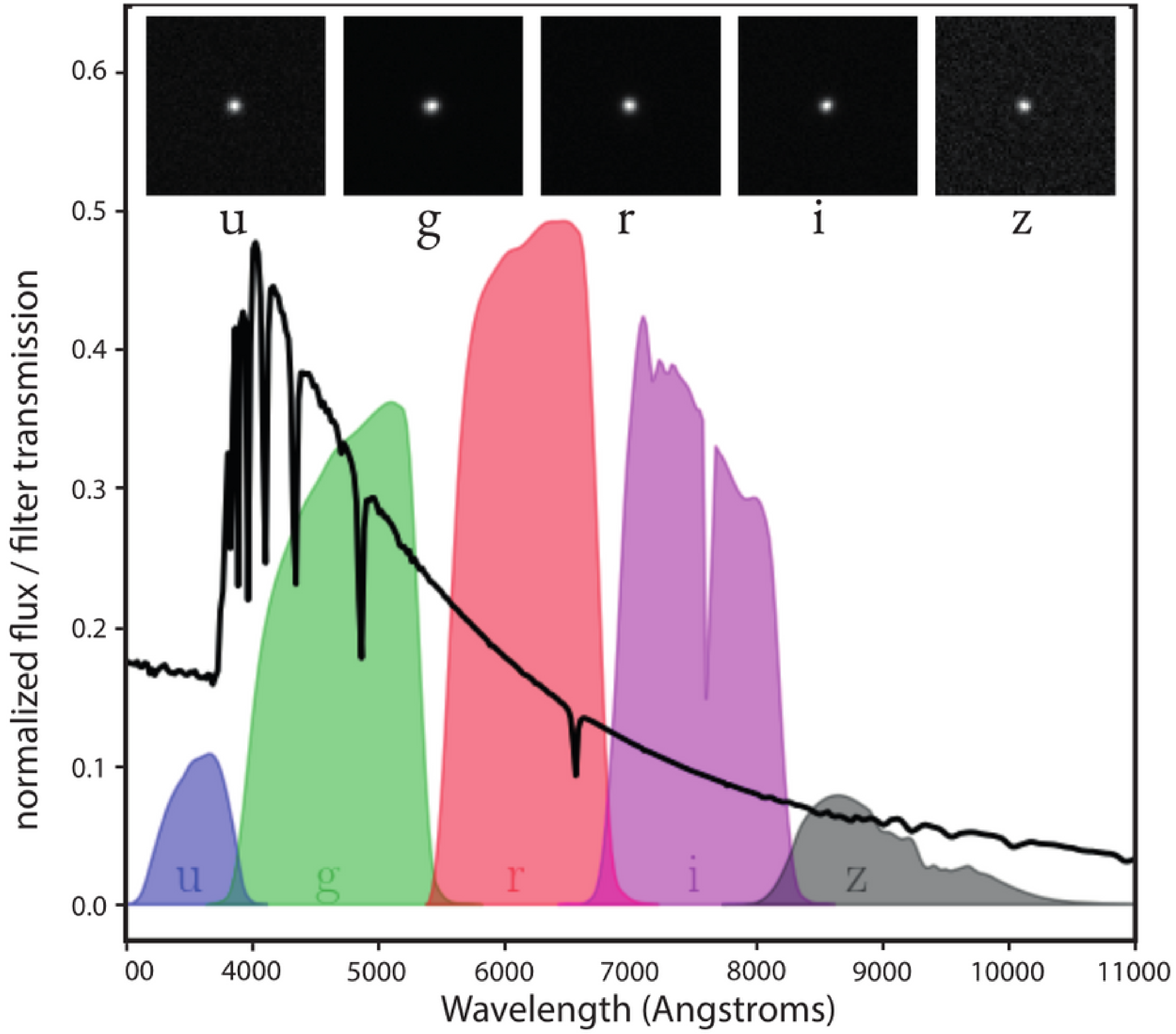
First Contribution

Multimodality For Improved Photometric Redshifts

Can we improve the network architecture to further improve the redshift estimation quality ?

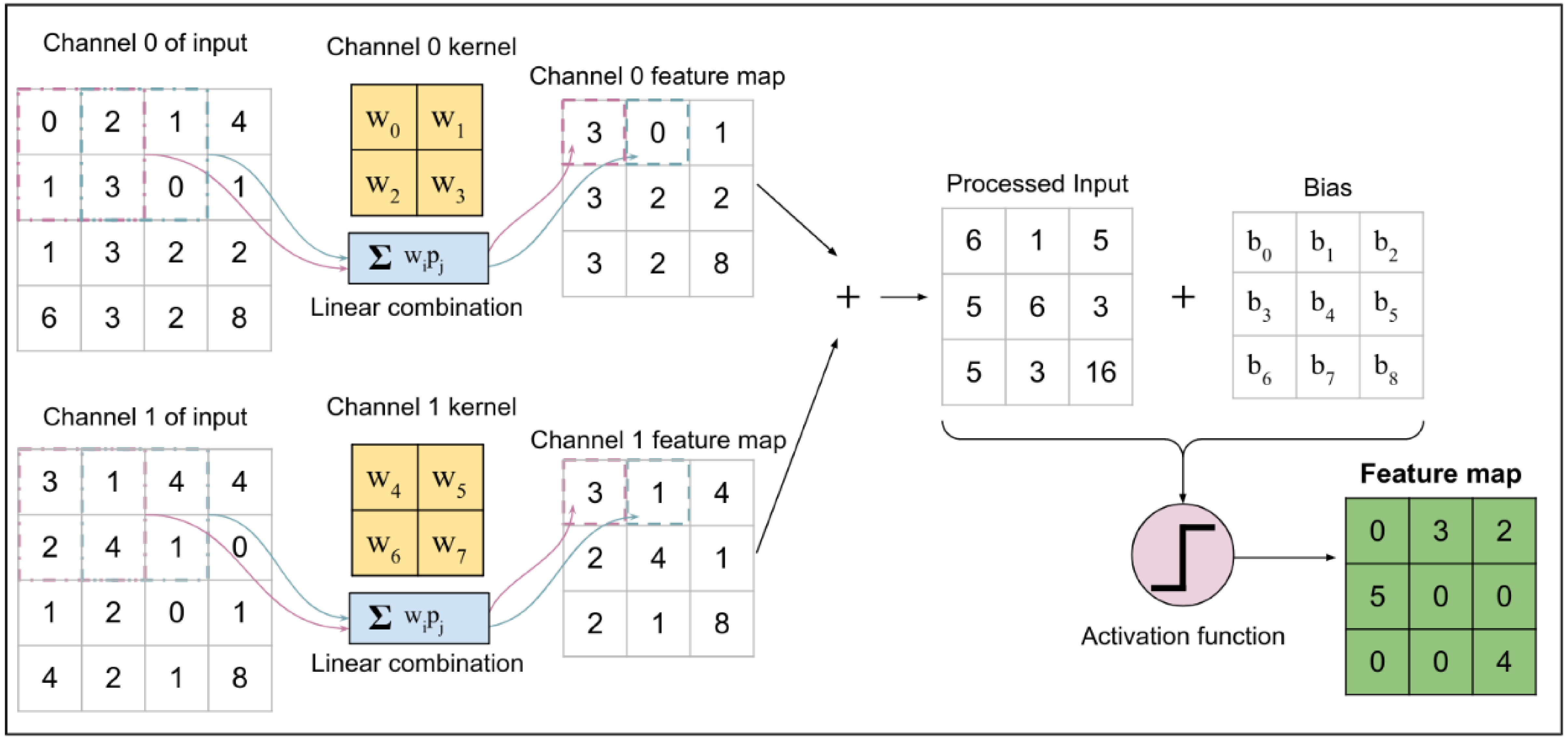
Suboptimal Input Processing ?

Correlation between the bands is indicative of the galaxy SED

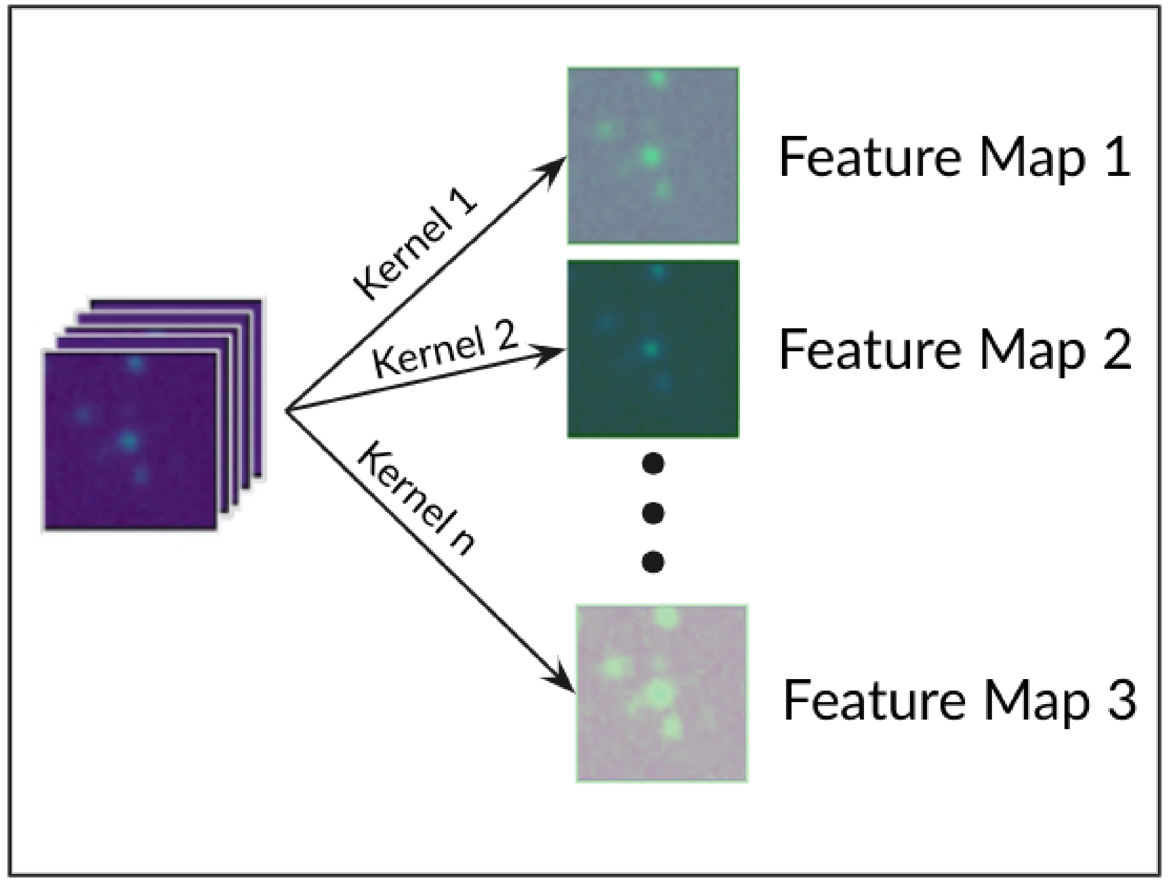


First Contribution

Suboptimal Input Processing ?



Convolution Operation

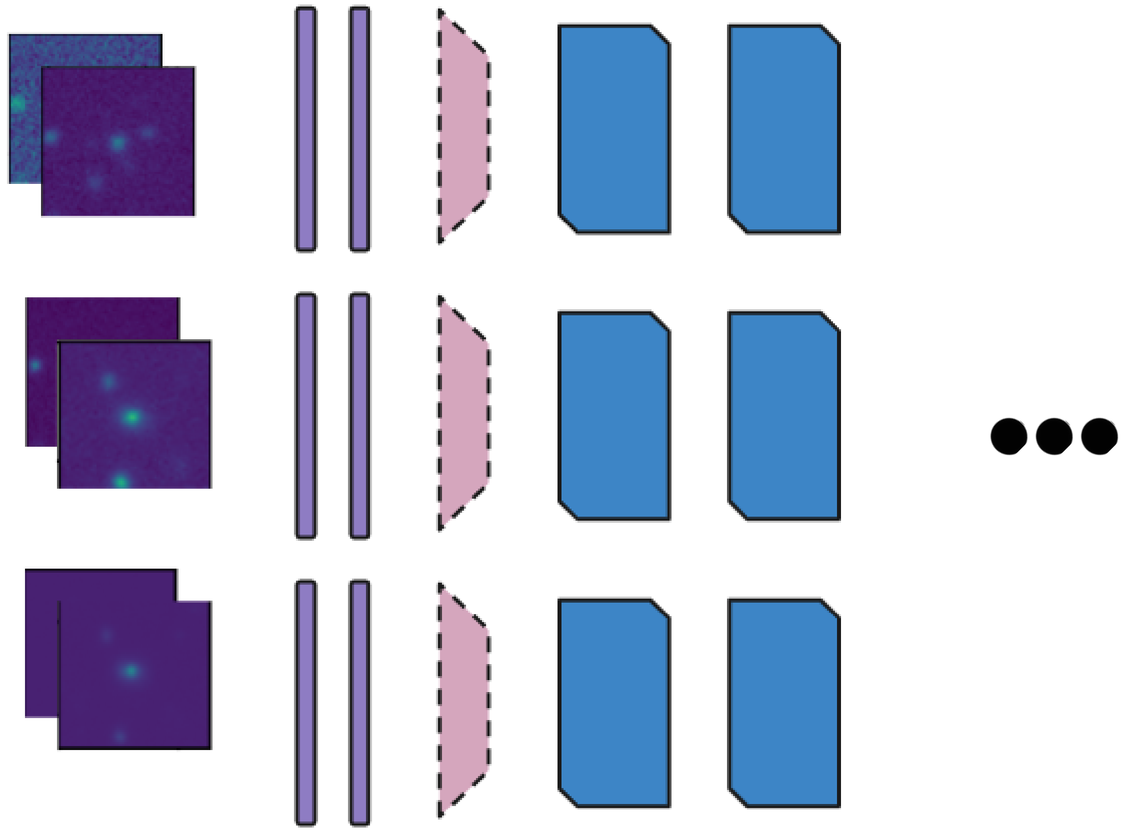


First Convolution Layer

First Contribution

Suboptimal Input Processing ?

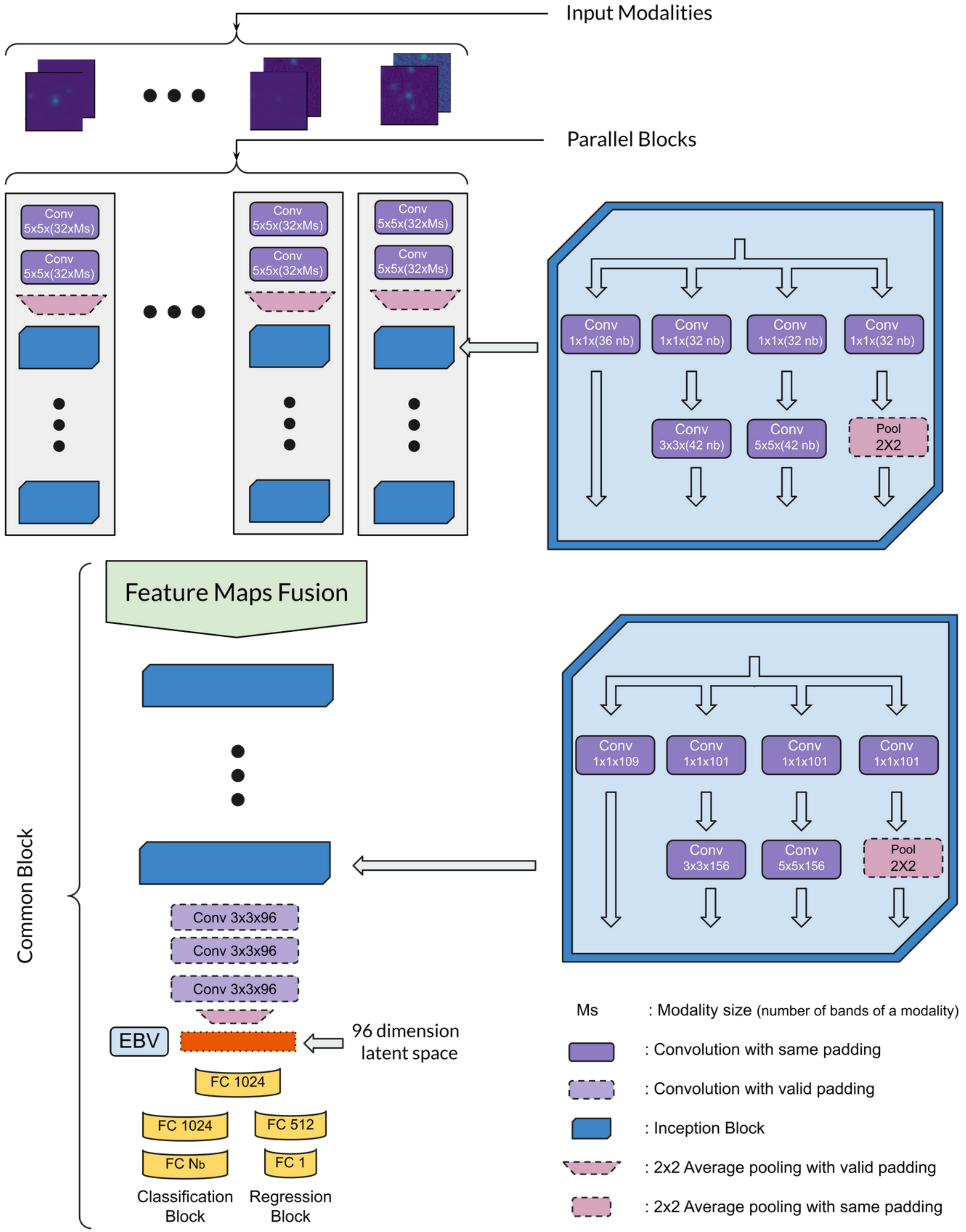
Proposed Solution : Parallel processing of small sets of bands (Multimodality)



First Contribution

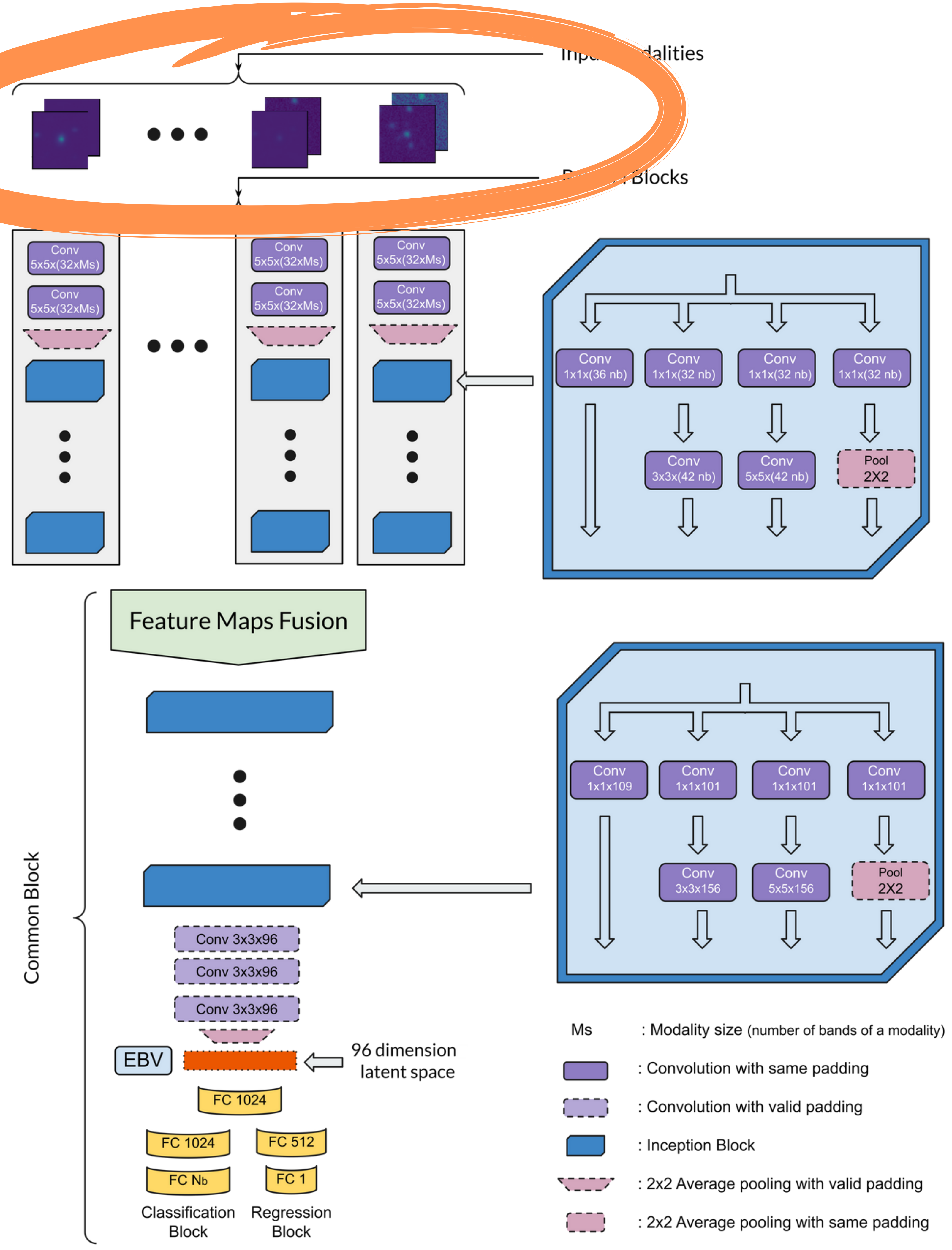
Generic Architecture

First Contribution



Generic Architecture

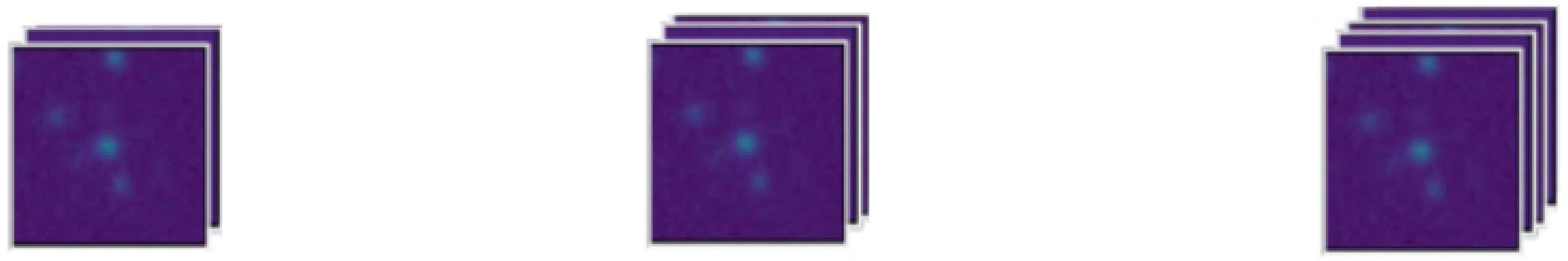
- How many bands in each modality ? **=> Modality Size**
- How to group bands into a modality ? **=> Modality Order**



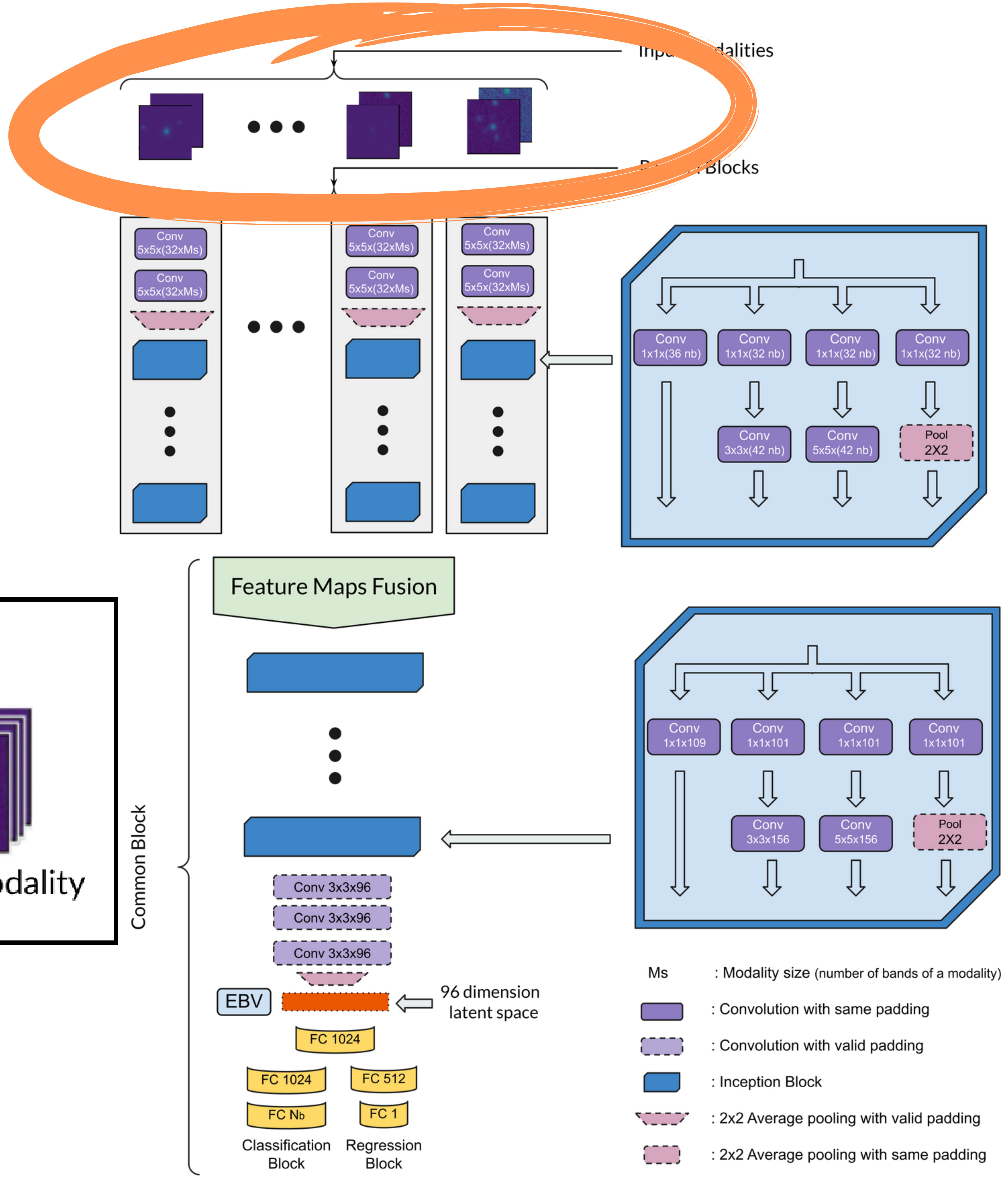
First Contribution

Generic Architecture

Modality Size

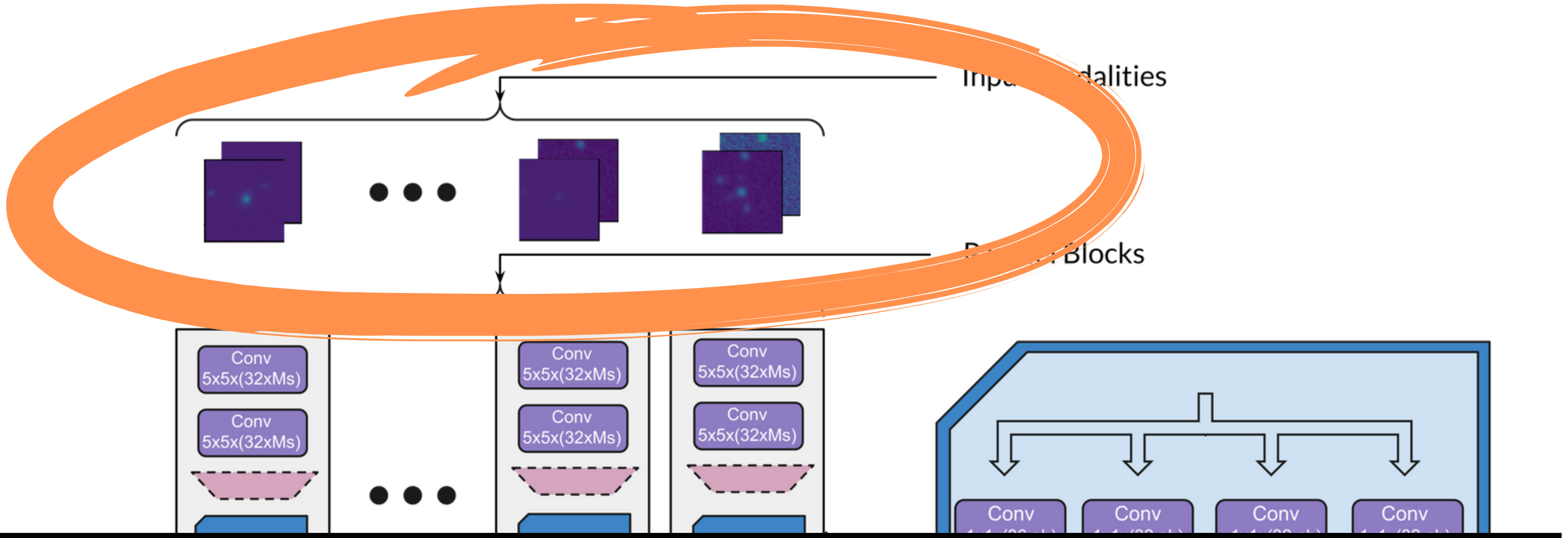


2 bands modality | 3 bands modality | 4 bands modality



First Contribution

Generic Architecture



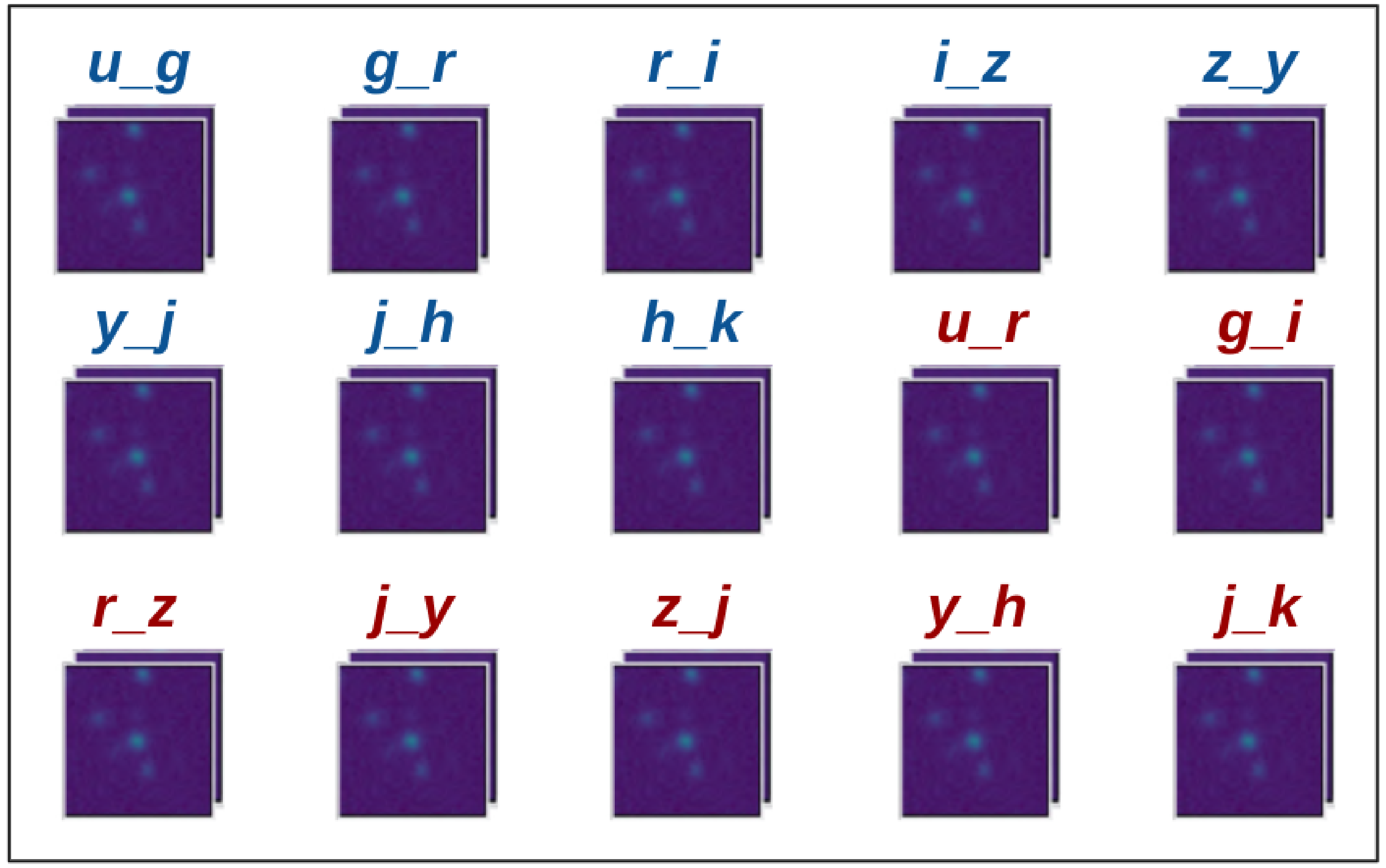
Modality Order

Taking the ***ugrizyjhk*** bands as an example :

Different order and size modalities composition

	2 bands	3 bands	4 bands
1st order	u_g, g_r, r_i, i_z, z_y, y_j, j_h, h_k	u_g_r, g_r_i, r_i_z, i_z_y, z_y_j, y_j_h, j_h_k	u_g_r_i, g_r_i_z, r_i_z_y, i_z_y_j, z_y_j_h, y_j_h_k
2nd order	u_r, g_i, r_z, i_y, z_j, y_h, j_k	u_r_z, g_i_y, r_z_j, i_y_h, z_j_k	u_r_z_j, g_i_y_h, r_z_j_k
3rd order	u_i, g_z, r_y, i_j, z_h, y_k	u_i_j, g_z_h, r_y_k	

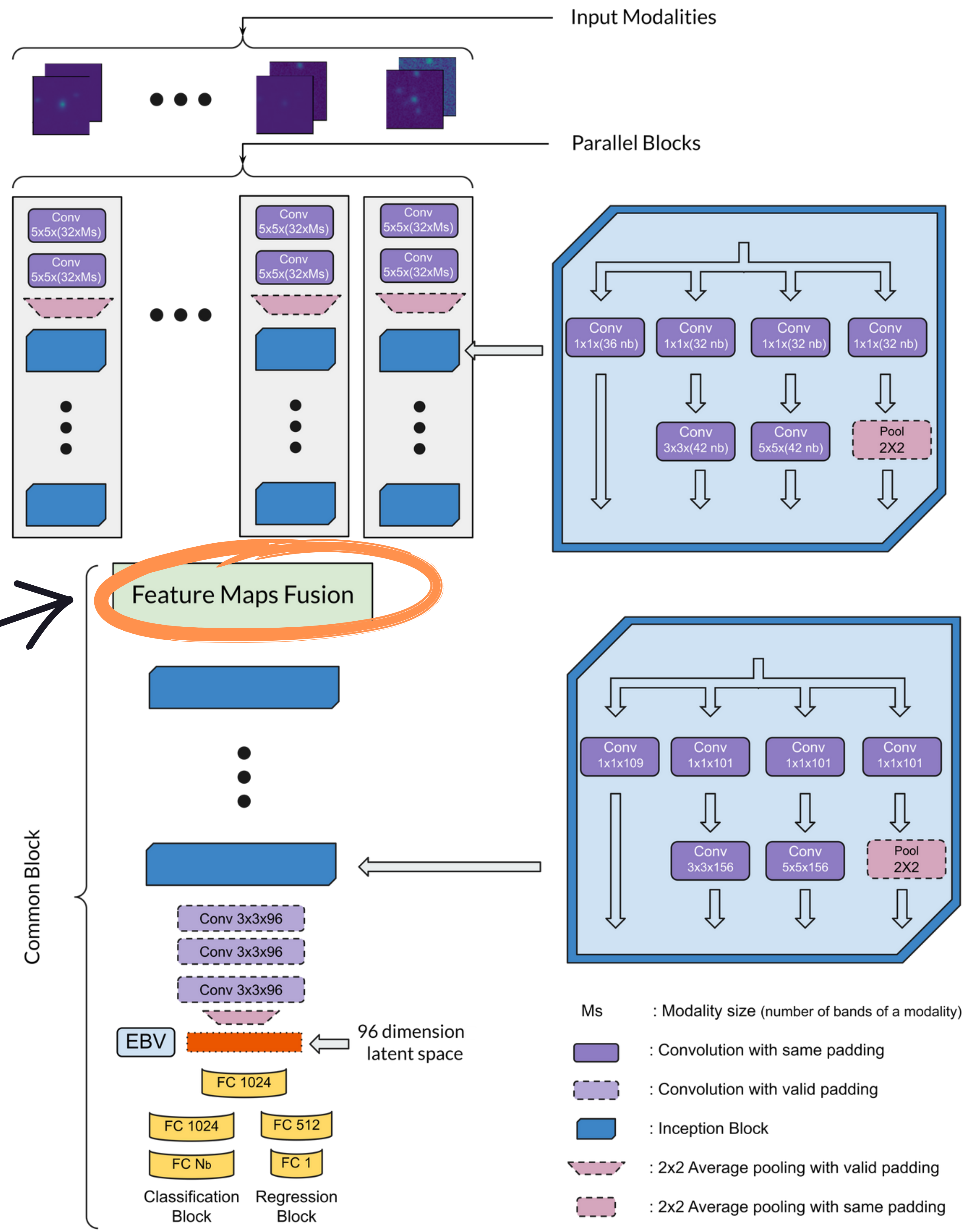
Ex : 2 band first and second order modalities



First Contribution

Generic Architecture

- Where in the network should the processed modalities be fused ? **=> Fusion Stage**



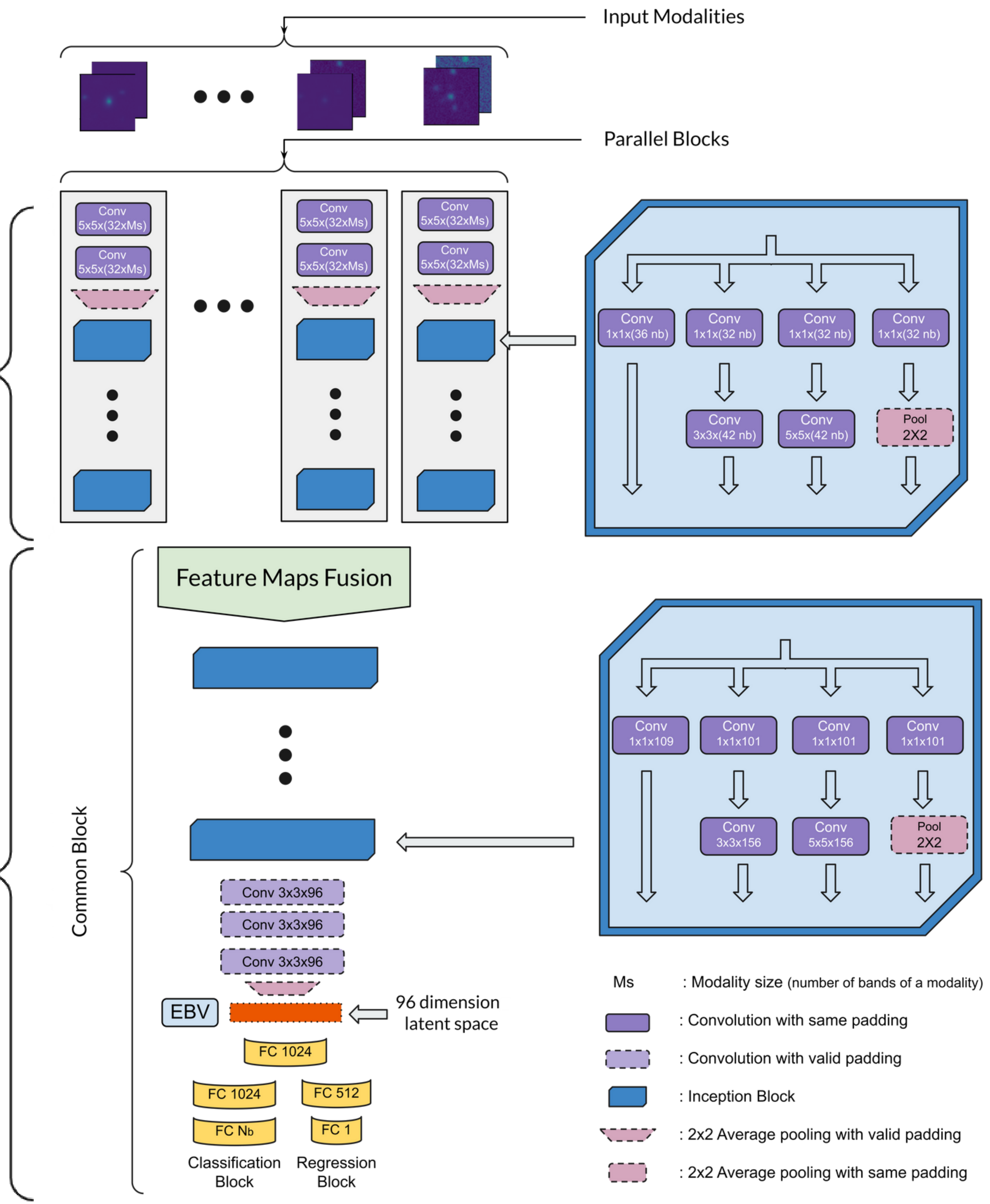
First Contribution

Generic Architecture

Fusion Stage

Depth of Parallel Block ?

Depth of Common Block ?



First Contribution

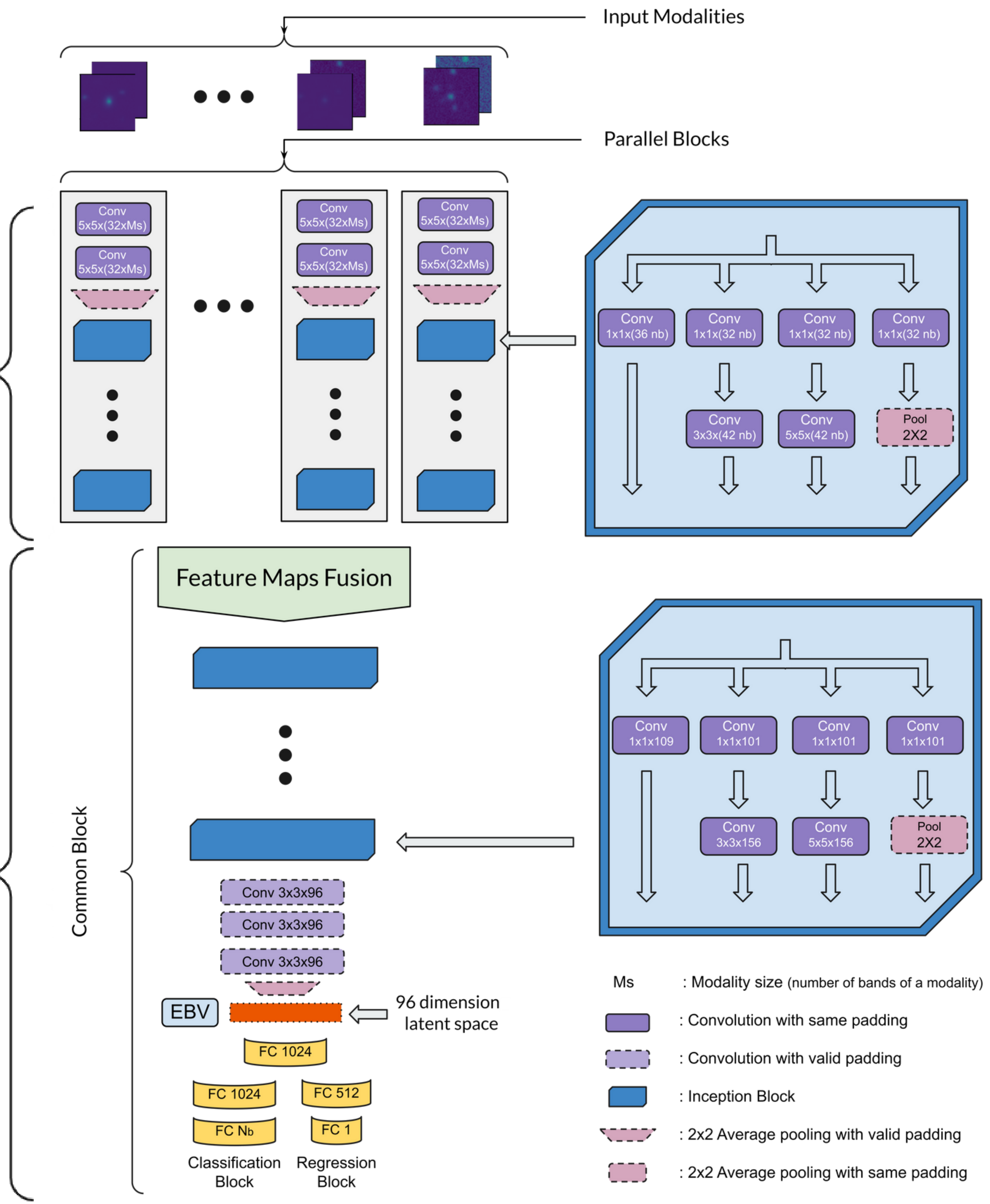
Generic Architecture

Fusion Stage

Depth of Parallel Block ?

- Early Fusion : 25 % Parallel, 75 % Common
- Middle Fusion : 50 % Parallel, 50 % Common
- Late Fusion : 75 % Parallel, 25 % Common

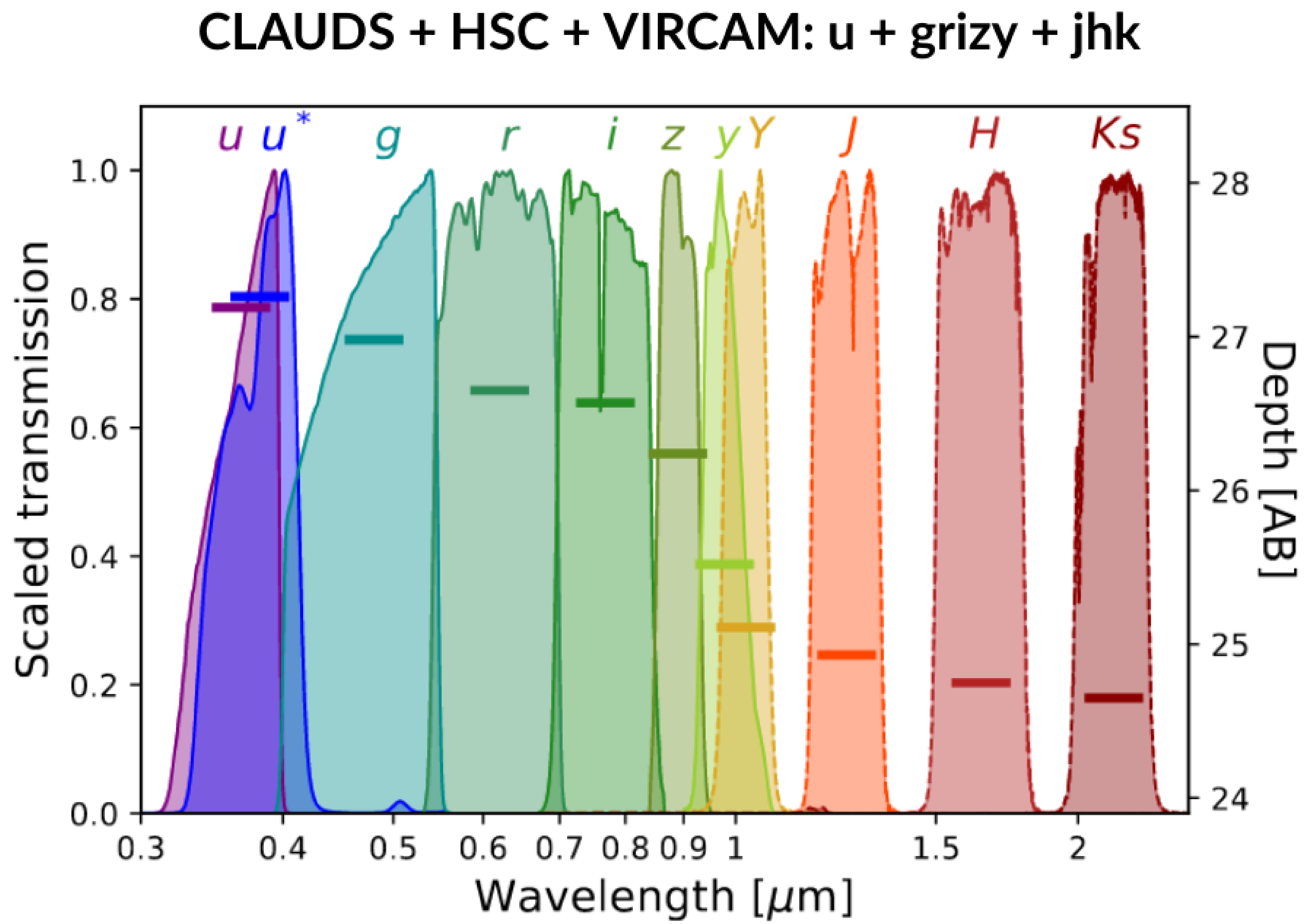
Depth of Common Block ?

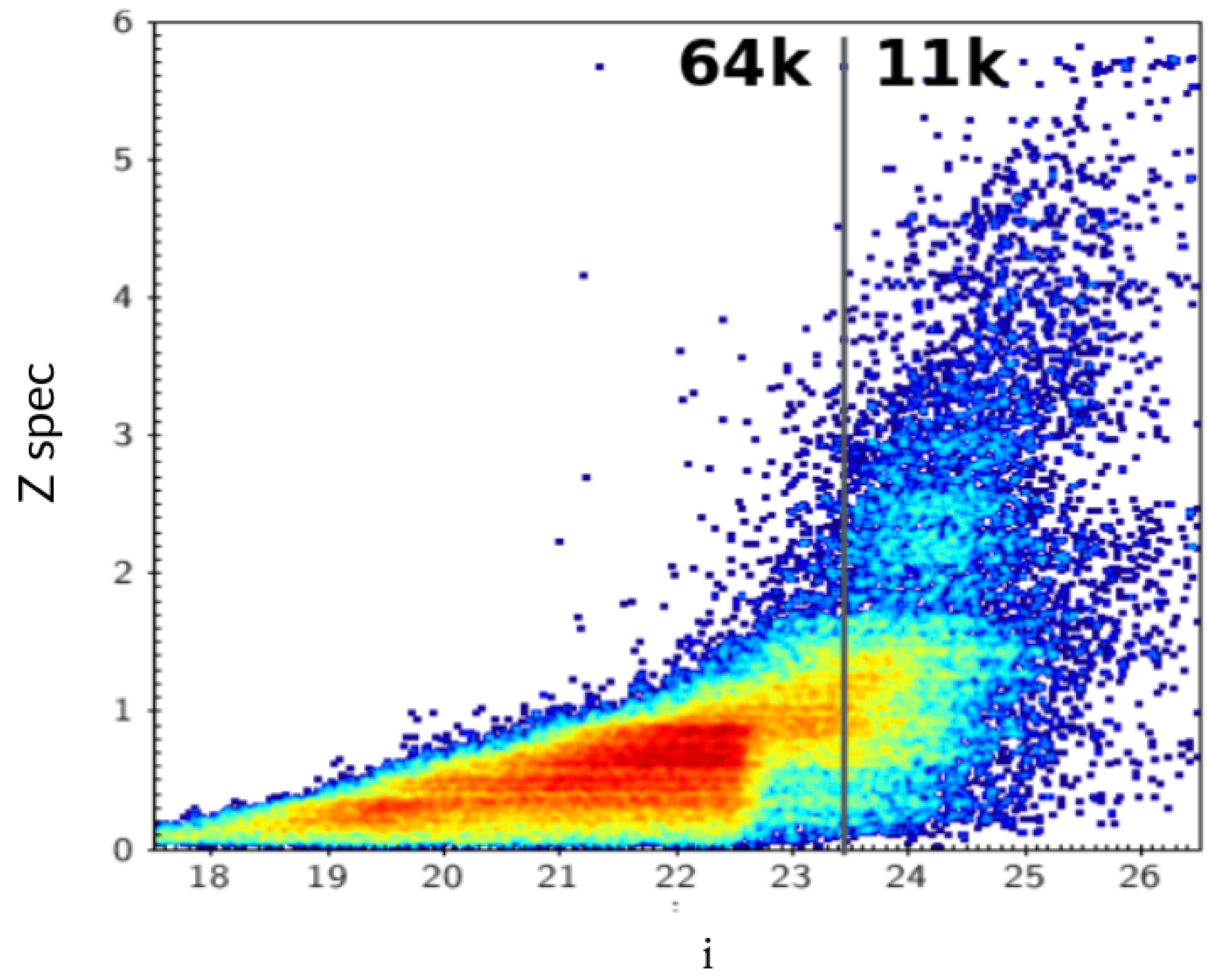


First Contribution

Exploring Multimodality

High Redshift Spectroscopic Sample : HSC - CLAUDS

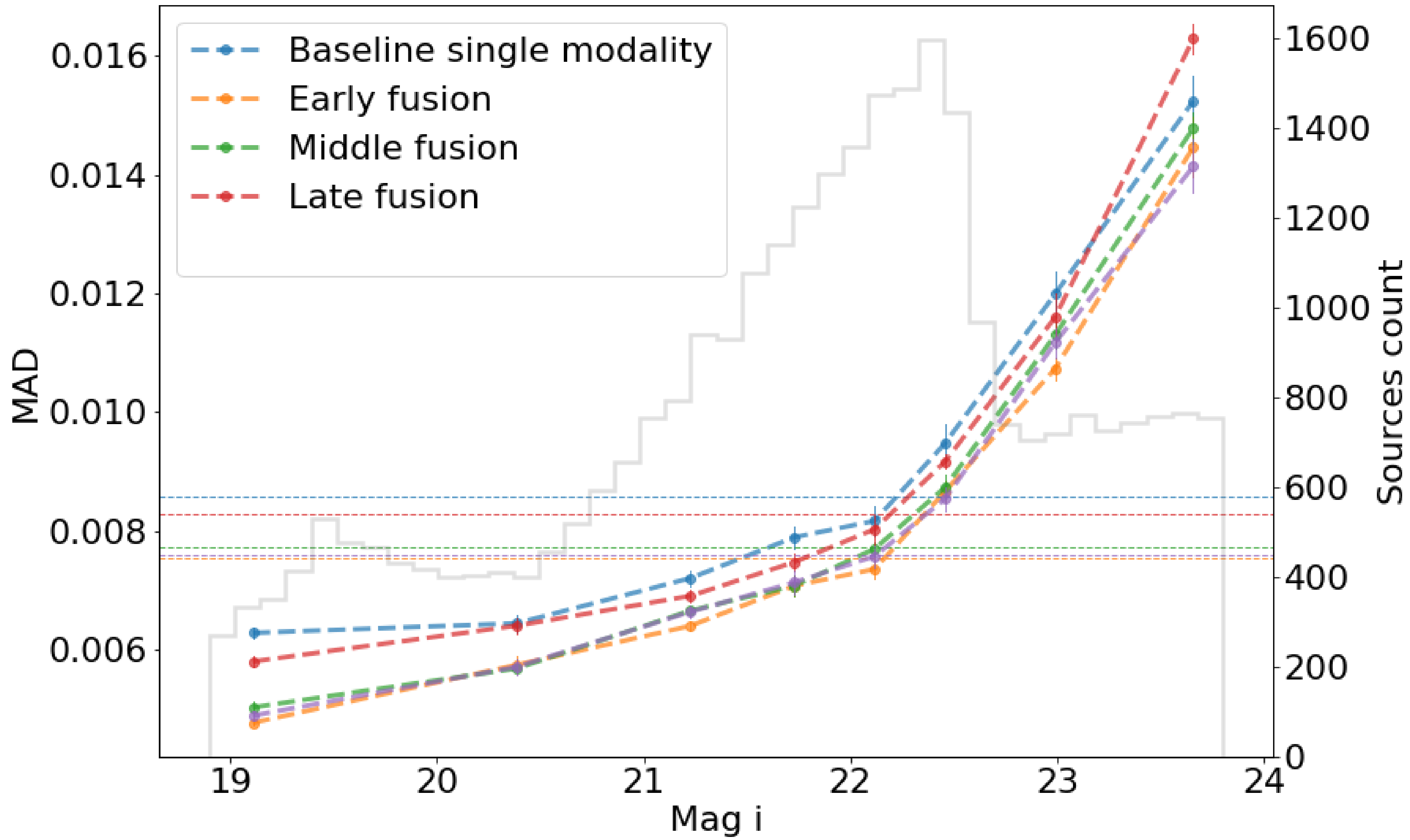




First Contribution

Exploring Multimodality

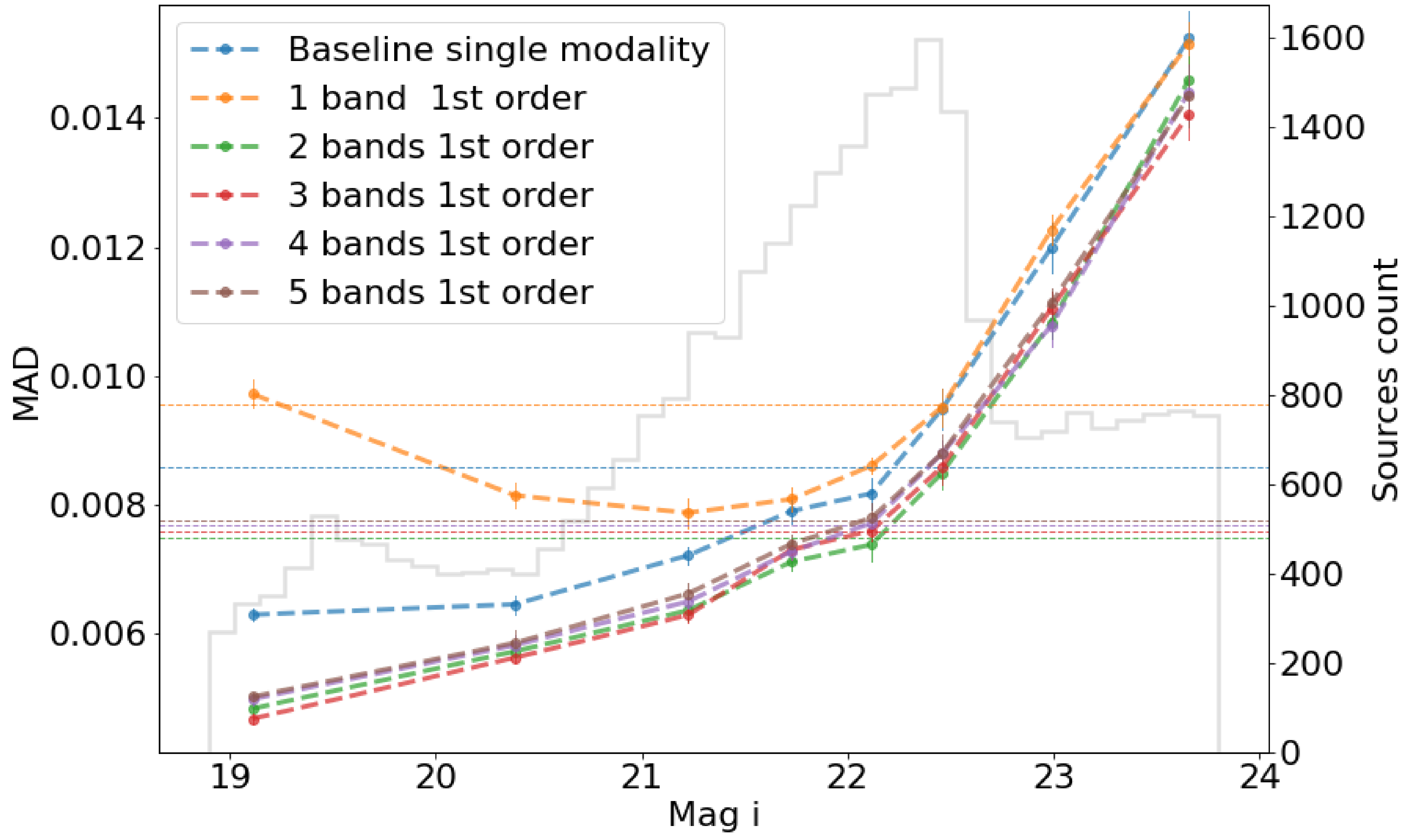
Using 2 band first and second order modalities, **Early Fusion yields best results**.



First Contribution

Exploring Multimodality

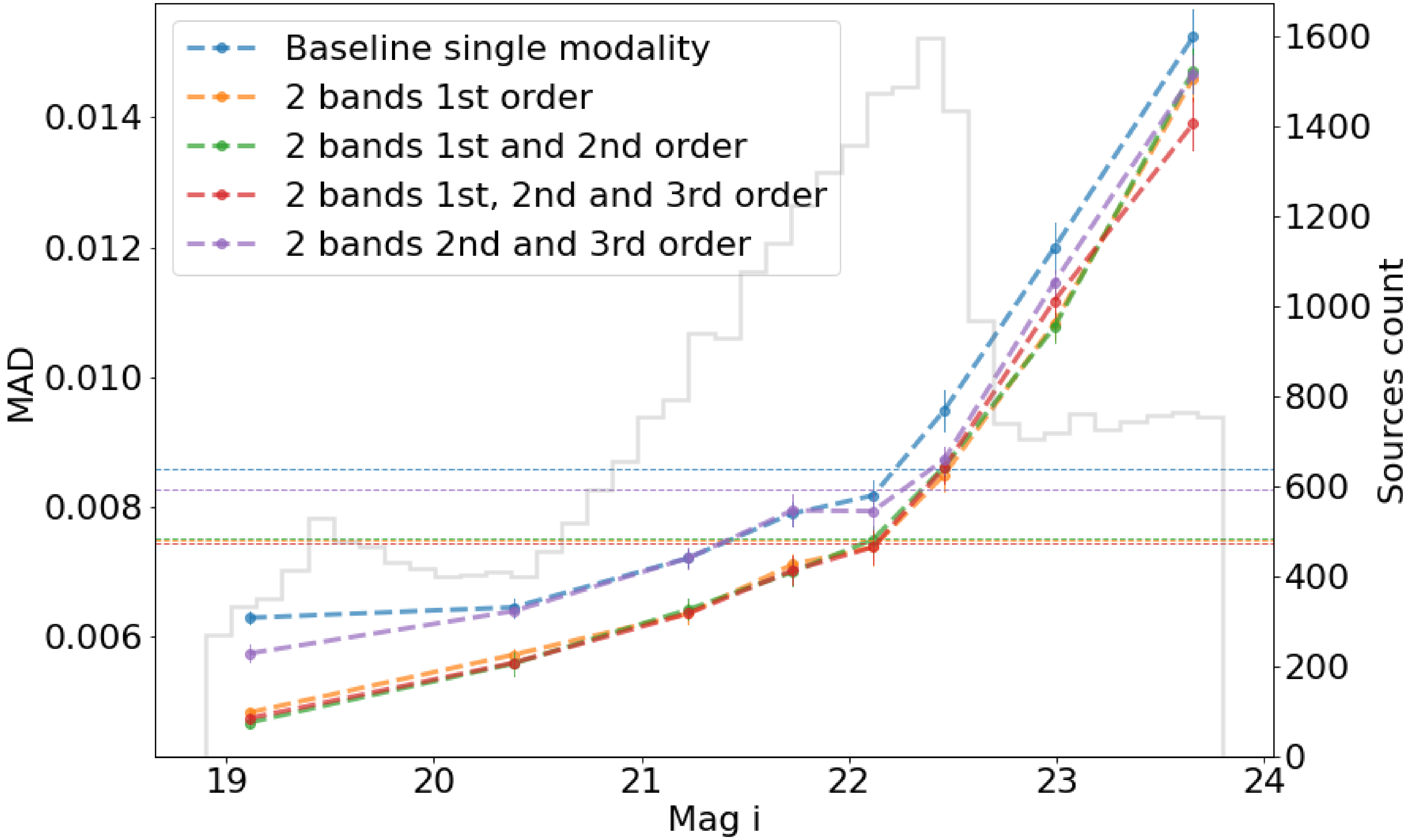
Using first order modalities and early fusion, **modalities of size two and more produce optimal results**



First Contribution

Exploring Multimodality

Using 2 band modalities and early fusion, we see that **First Order modalities are the most important**



First Contribution

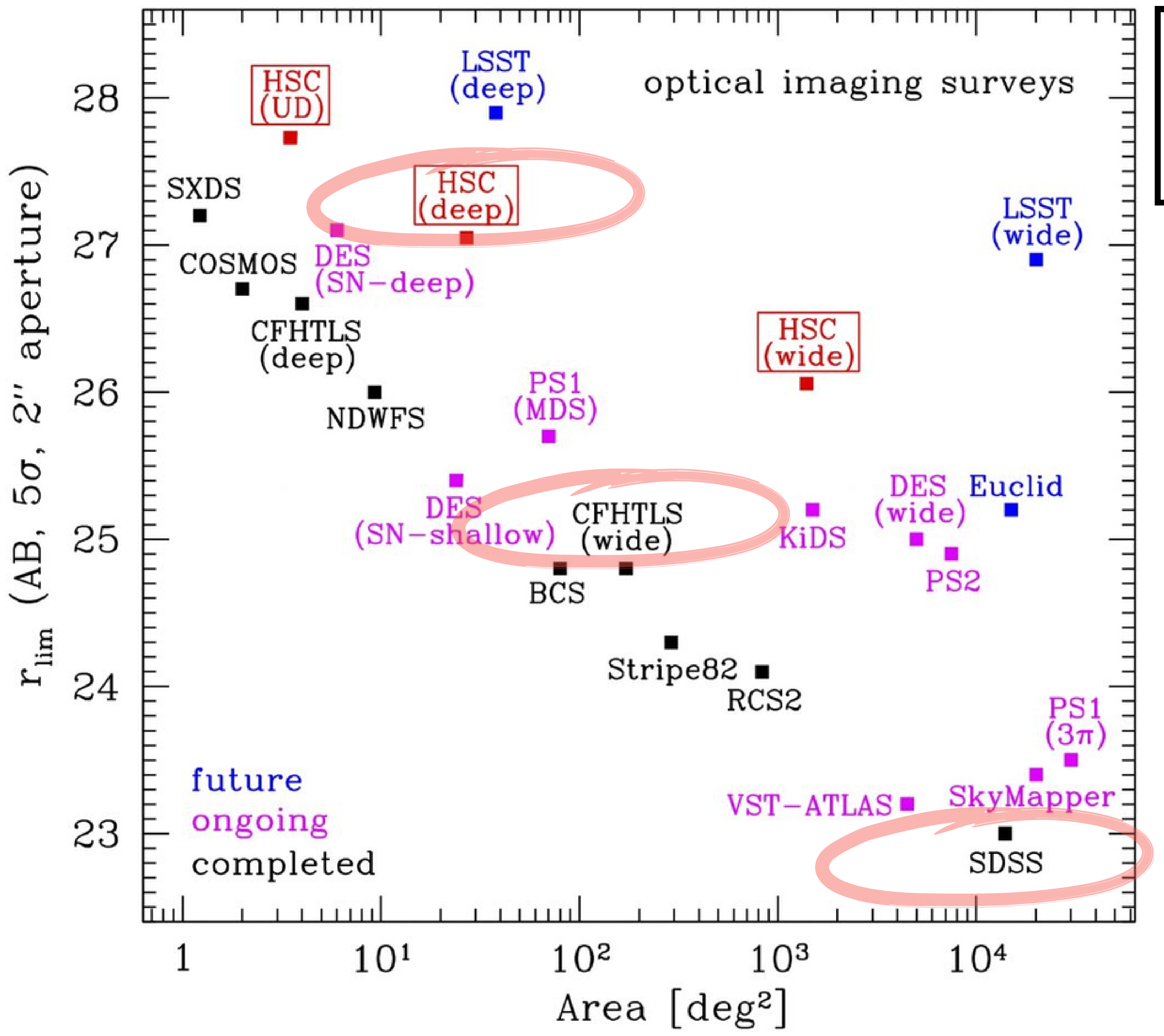
Exploring Multimodality

Most optimal and simplest configuration

- Early Fusion
- 2 band modalities
- First order modalities

Exploring Multimodality

Improved metrics under various conditions



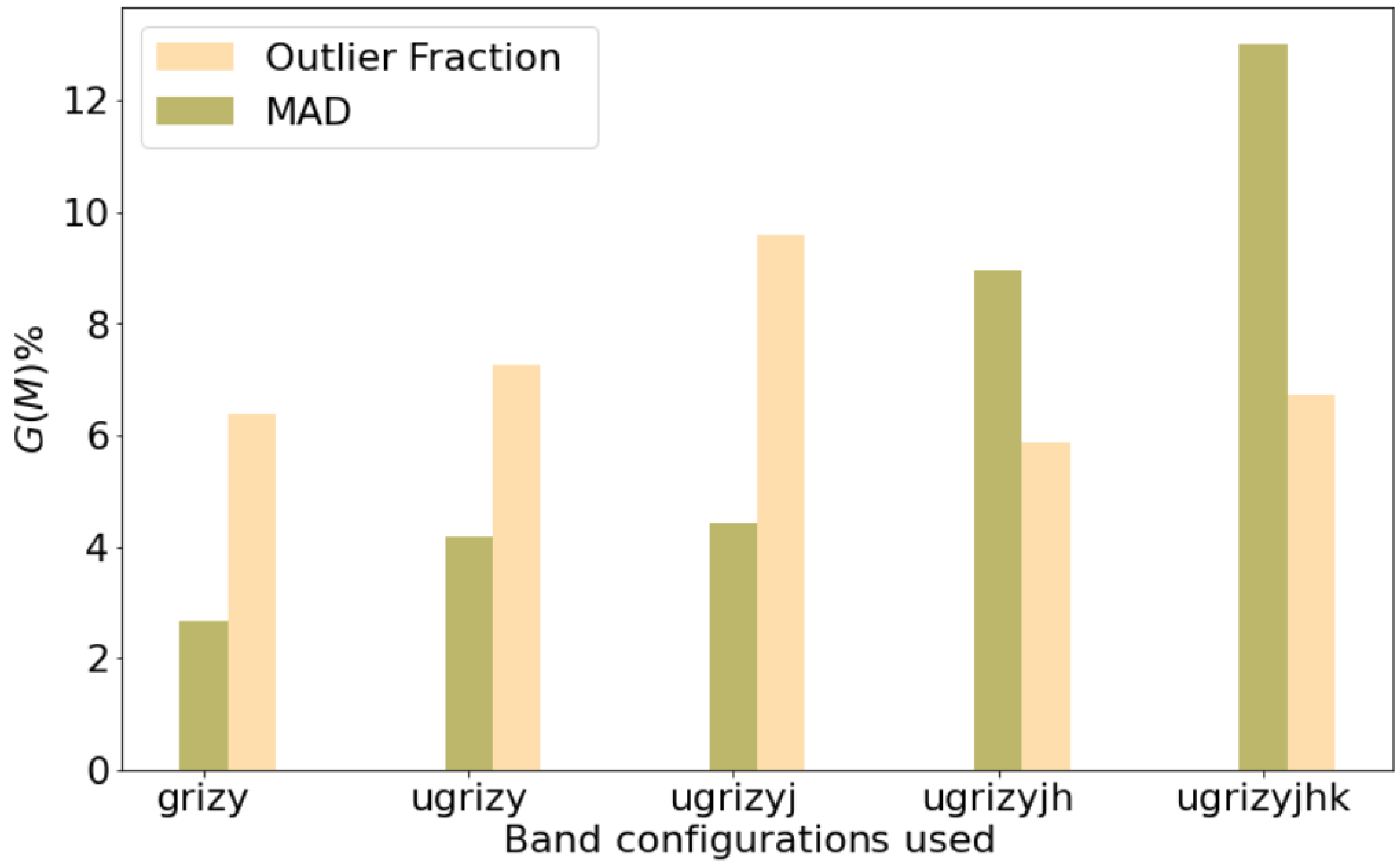
$$G(M) = \frac{|M_B| - |M_M|}{|M_B|}$$

Relative gain of multimodality compared to the baseline on a metric

Experiences	σ 10^{-3}	η %	$<\Delta z>$ 10^{-3}	Count 10^3
SDSS				
Baseline	07.99	0.18	0.34	516.5
Multimodal	**07.85**	**0.16**	**0.31**	516.5
$G(M)$	1.74%	10.88%	6.28%	-
p_{value}	0.0	0.0	0.0	-
CFHTLS				
Baseline	16.01	0.85	**0.22**	108.5
Multimodal	**15.35**	**0.79**	0.29	108.5
$G(M)$	4.13%	7.22%	-24.05%	-
p_{value}	0.0	0.0002	0.15	-
HSC-6b				
Baseline	09.14	1.25	1.97	46.8
Multimodal	**08.87**	**1.20**	**1.63**	46.8
$G(M)$	2.96%	3.94%	17.33%	-
p_{value}	0.0	0.0575	0.04	-
HSC-9b				
Baseline	08.41	1.24	**1.58**	33.1
Multimodal	**07.60**	**1.19**	1.64	33.1
$G(M)$	10.1%	3.67%	-3.1%	-
p_{value}	0.0	0.11	0.40	-
HSC-9b with 3DHST redshifts				
Baseline	14.44	2.46	13.28	2.2
Multimodal	**13.88**	**2.37**	**10.6**	2.2
$G(M)$	3.93%	3.71%	20.19%	-
p_{value}	0.069	0.27	0.10	-
HSC-9b with PRIMUS redshifts				
Baseline	12.34	2.66	11.84	15
Multimodal	**11.38**	**1.85**	**09.23**	15
$G(M)$	7.74%	30.4%	22.01%	-
p_{value}	0.0	0.0	0.0	-
HSC-9b with COSMOS2020 photometric redshifts				
Baseline	12.01	1.01	8.74	43.7
Multimodal	**11.46**	**0.83**	**6.82**	43.7
$G(M)$	4.57%	17.08%	21.97%	-
p_{value}	0.0	0.0	0.0001	-

Exploring Multimodality

Relative gain based on available photometric bands

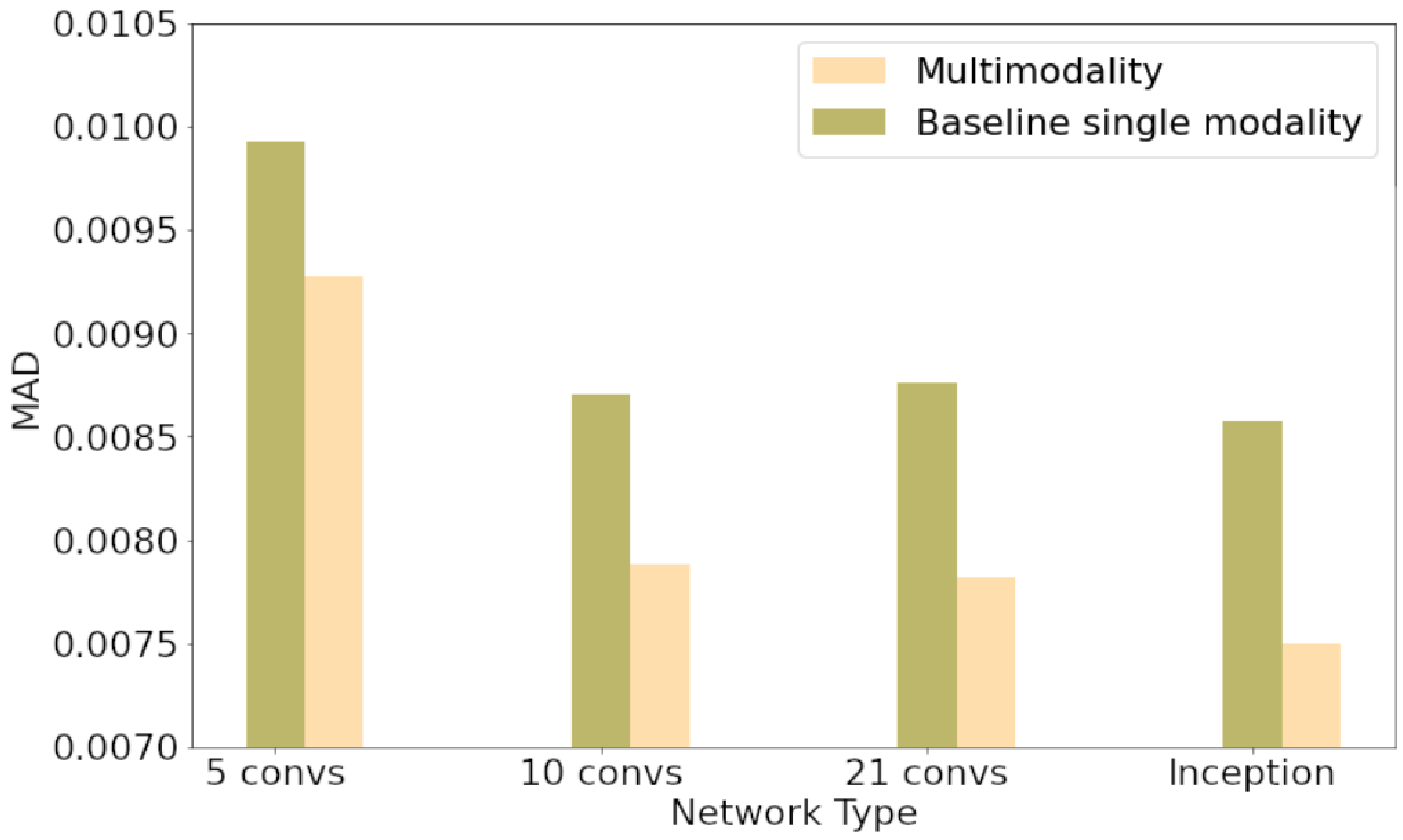


First Contribution

Exploring Multimodality

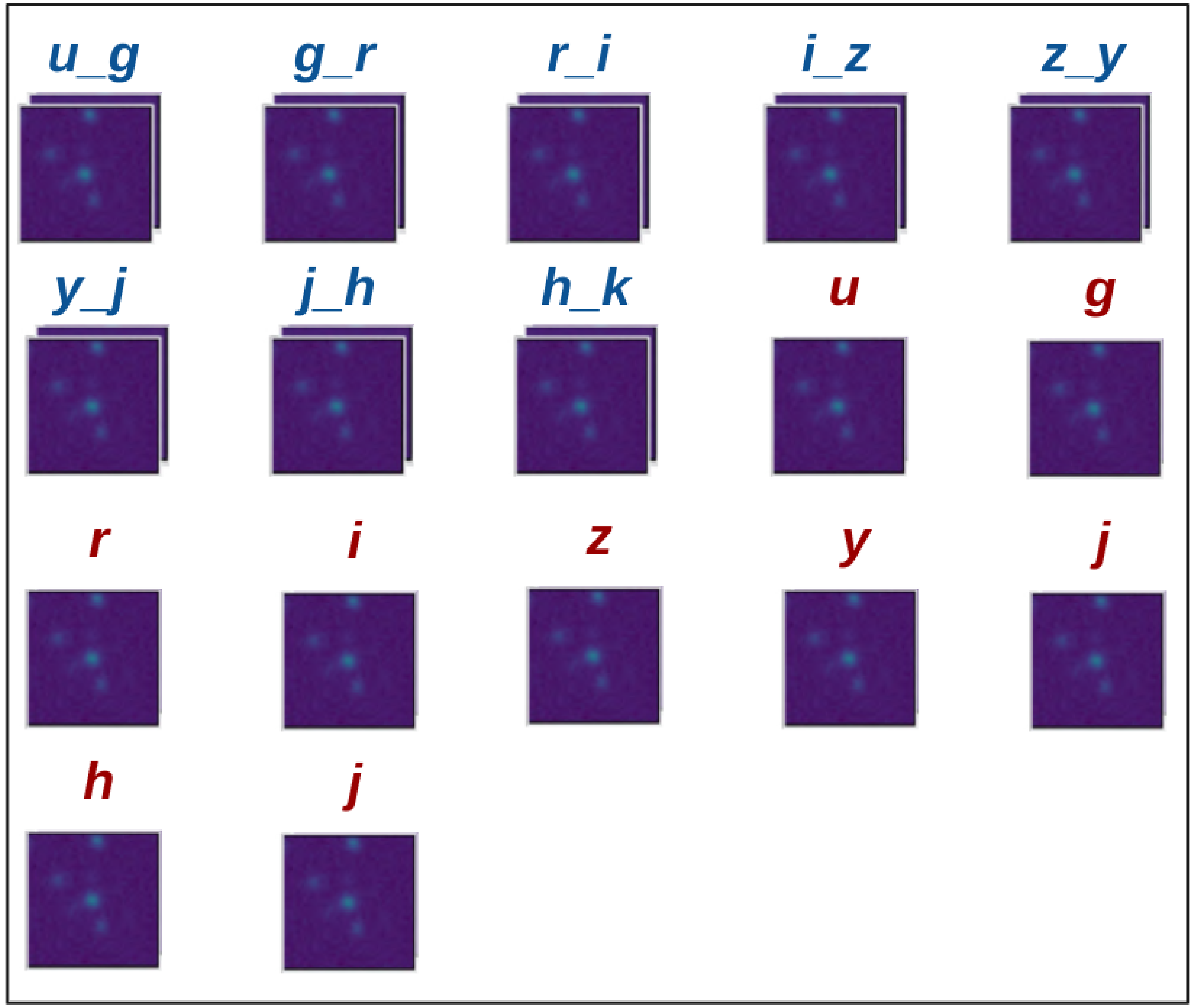
First Contribution

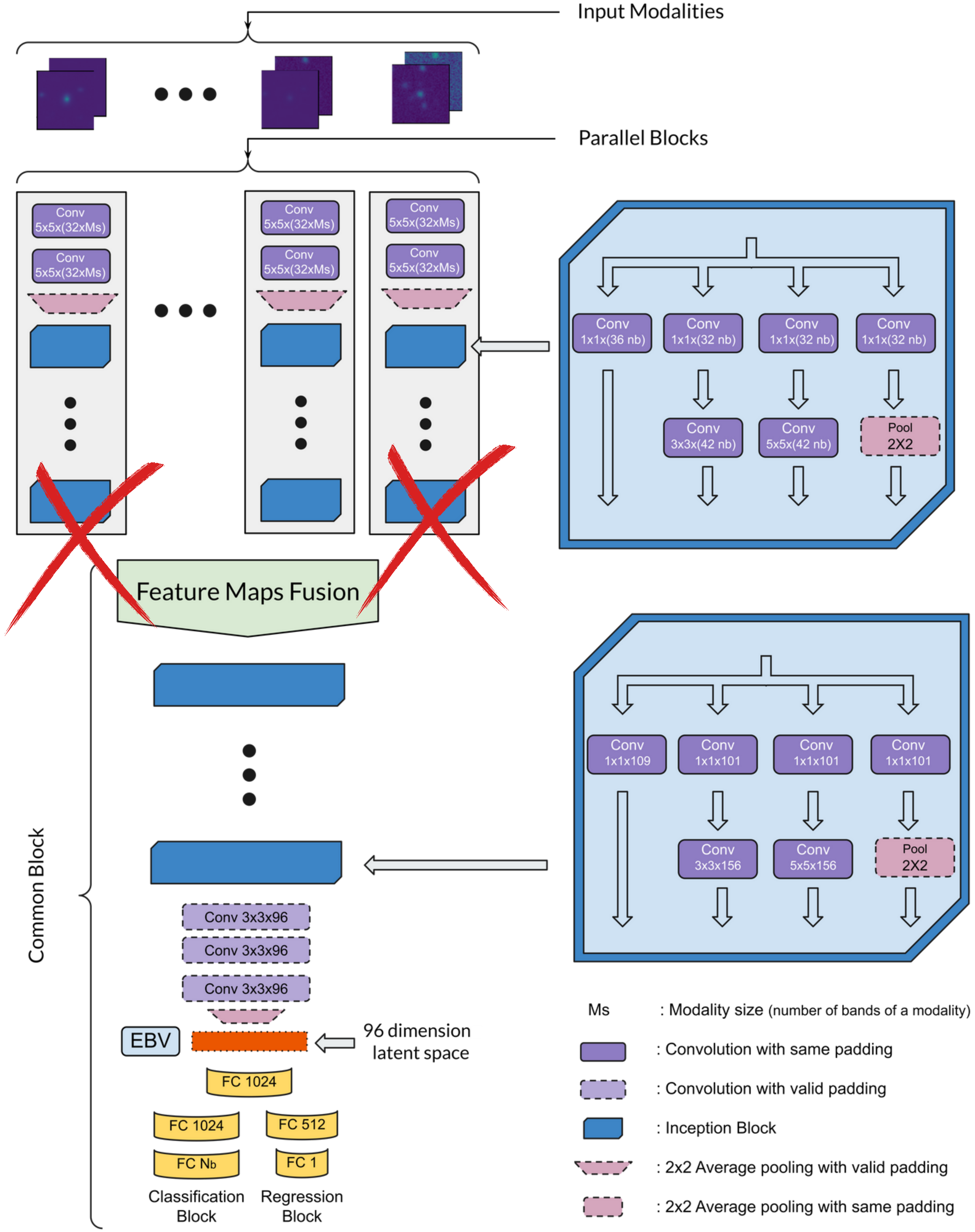
Performance using different network depths



Exploring Multimodality

Multimodality Dropout



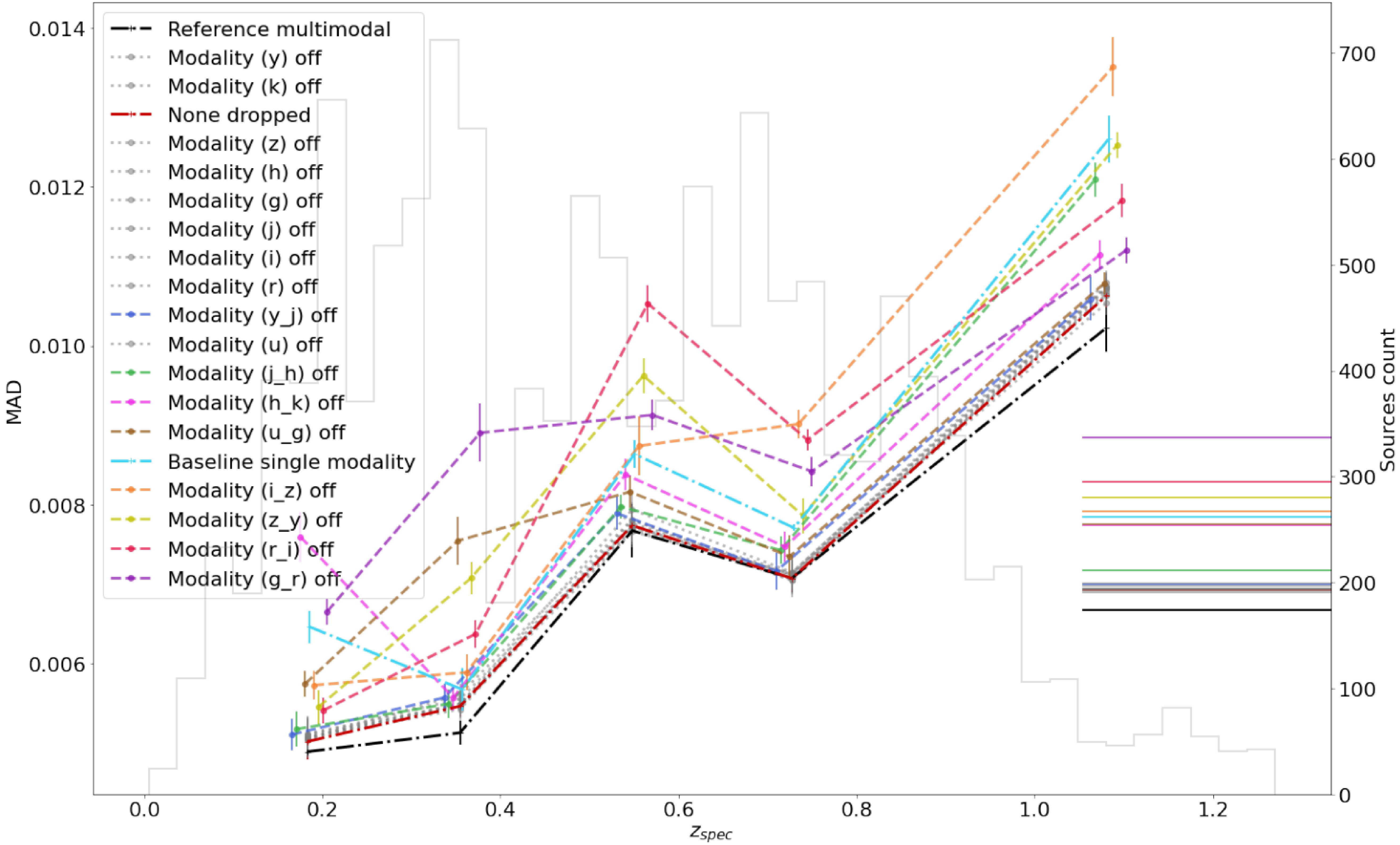


First Contribution

Exploring Multimodality

Multimodality Dropout

one modality dropped at test time



First Contribution

Contribution Published **: Ait-Ouahmed et al. 2023. A&A**

- **Introduction of simple yet efficient method to optimize CNN redshift estimations**
- **Multimodality improves redshift estimation precision independently of the dataset and the CNN depth**
- **Multimodality achieves new state of the art redshift precision on the SDSS MGS.**
- **Multimodality dropout allows to isolate the effect of bands correlation and study it.**

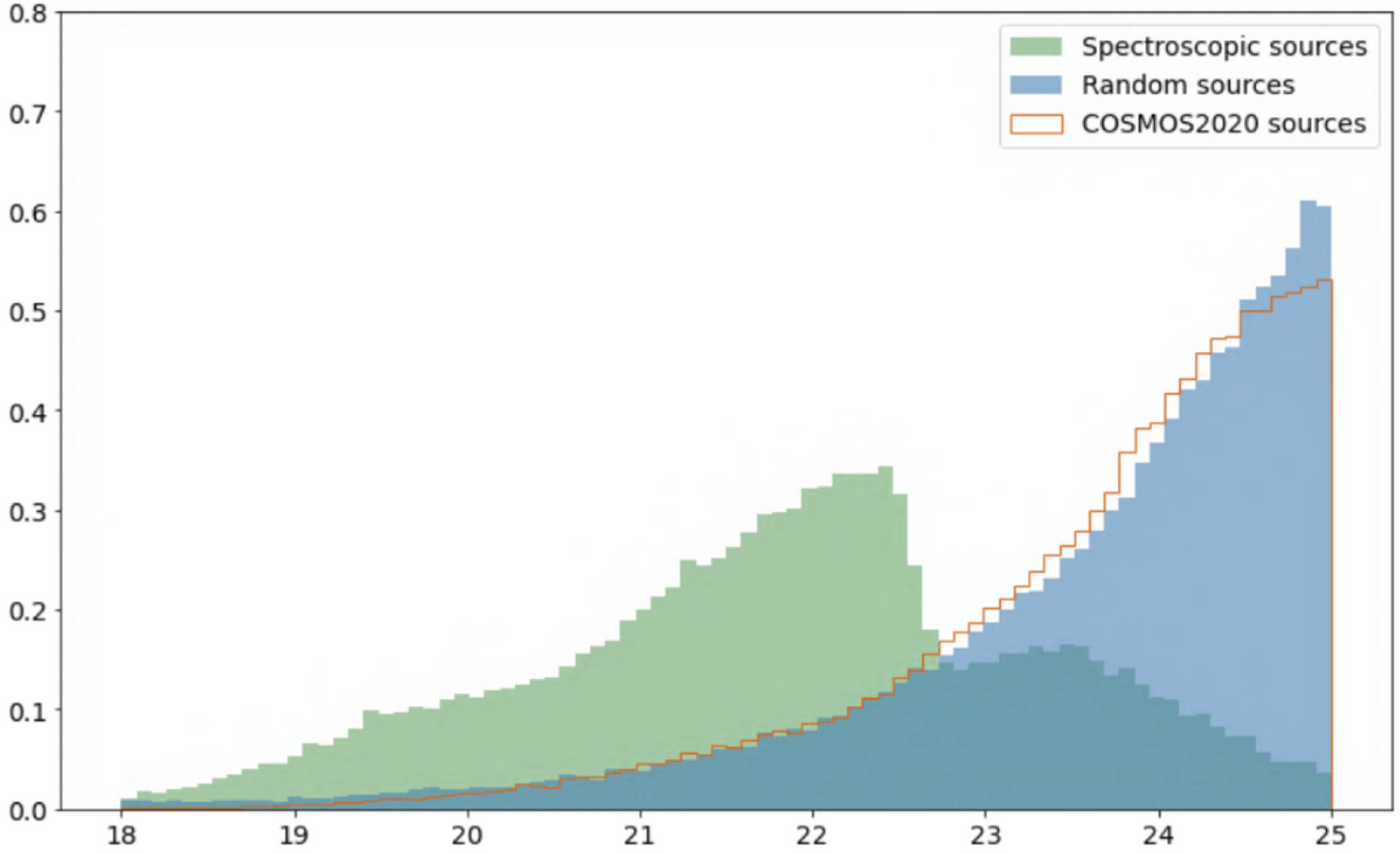
Second Contribution

Application To The HSC Deep Survey

What challanges for a realstic application tackling the needs of current and futur deep surveys ?

Building A Realstic Dataset

Previous application was not realistic



Second Contribution

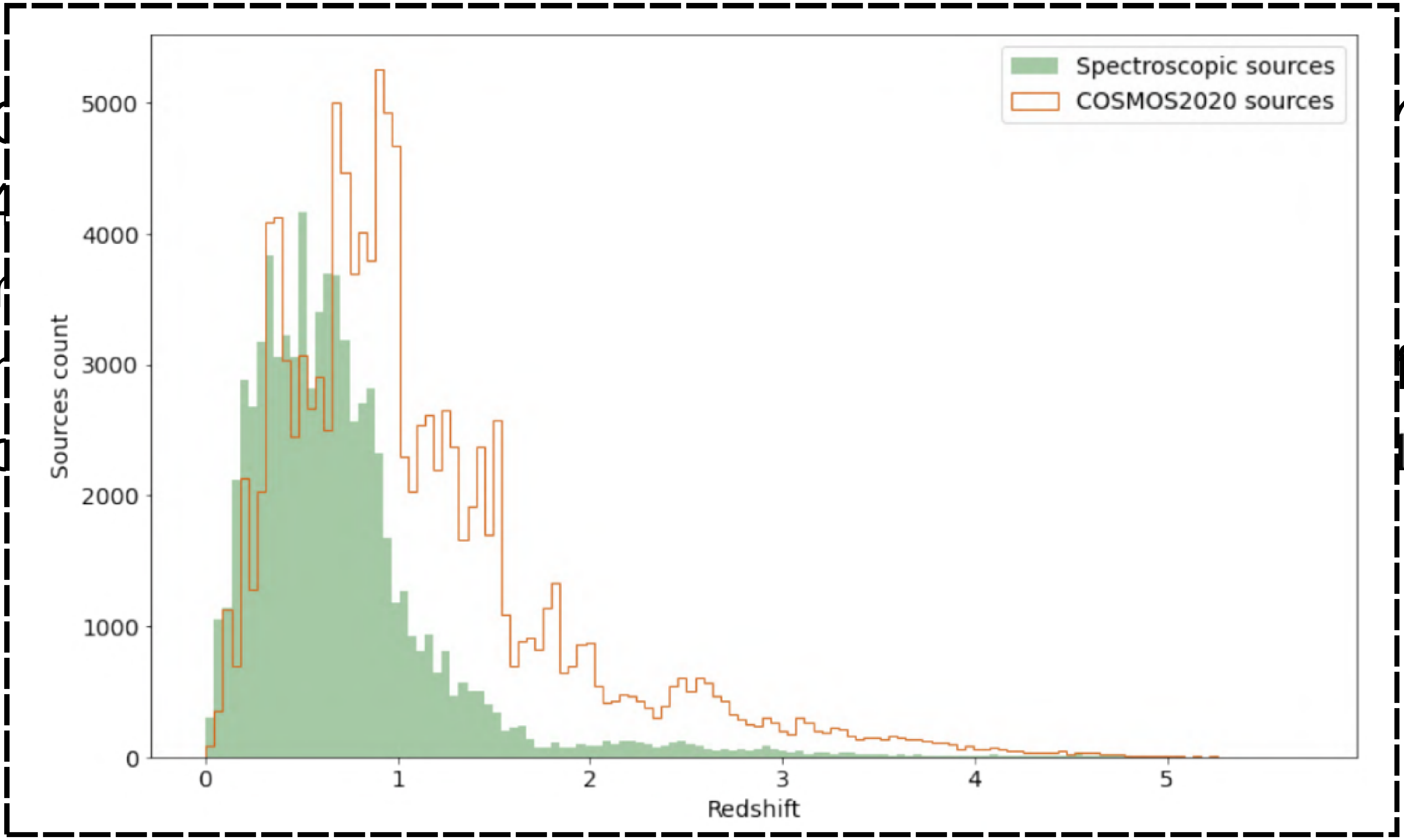
Building A Realstic Dataset

COSMOS2020 For faint sources

- 30 band photometric redshifts from *Weaver et al. (2022)*
- 4 different photometric redshifts were estimated based on different SED Fitting methods
- The mean and standard deviation of these 4 redshifts are computed, we retain the ones satisfying : $\sigma(z) \leq 0.1(1+\bar{z})$

Building A Realstic Dataset

COSMOS2020 For faint sources



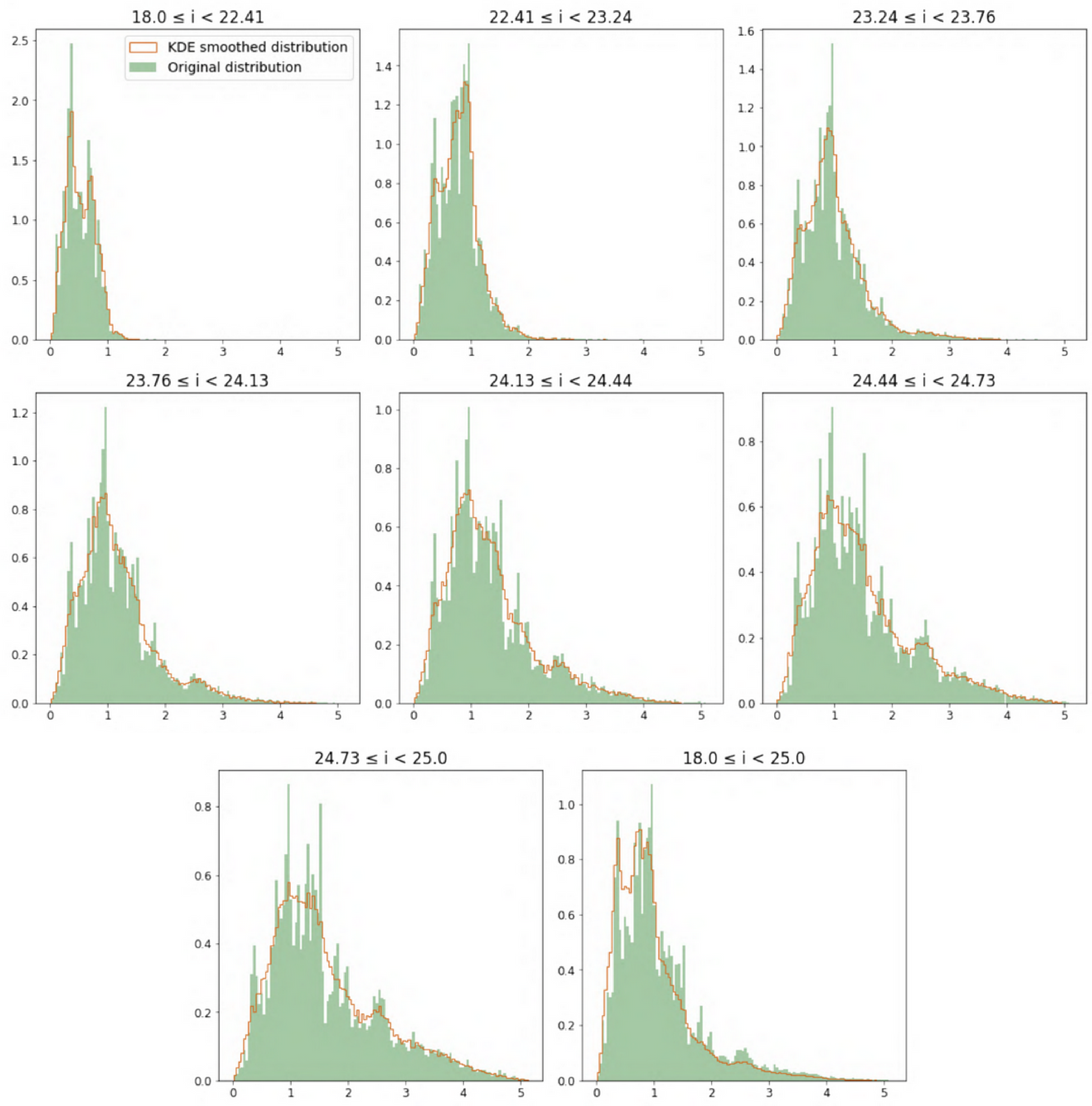
- 3(... *(2022)*
- 4 ... based
  or ...
- Th ... fts are
  co ... $1 + \bar{z})$

Second Contribution

Building A Realstic Dataset

Merging and smoothing

- Using Self Orgnizing Maps, spectroscopic sources are merged with cosmos2020 optimizing representativity and label quality
- Smoothing using the Kernel density estimation technique



Second Contribution

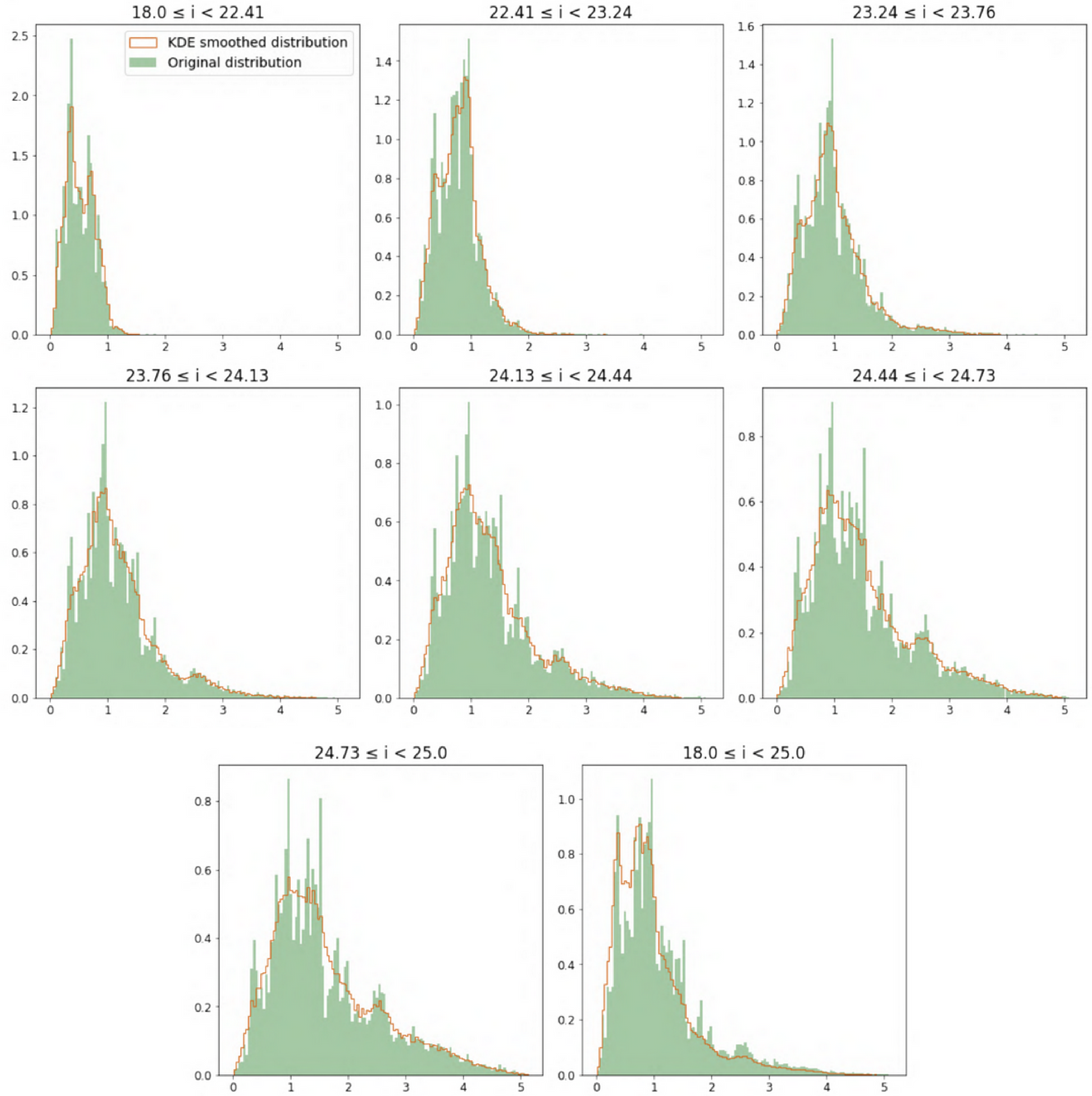
Building A Realstic Dataset

Merging and smoothing

- Using Self Orgnizing Maps, spectroscopic sources are merged with cosmos2020 optimizing representativity and label quality
- Smoothing using the Kernel density estimation technique

Result

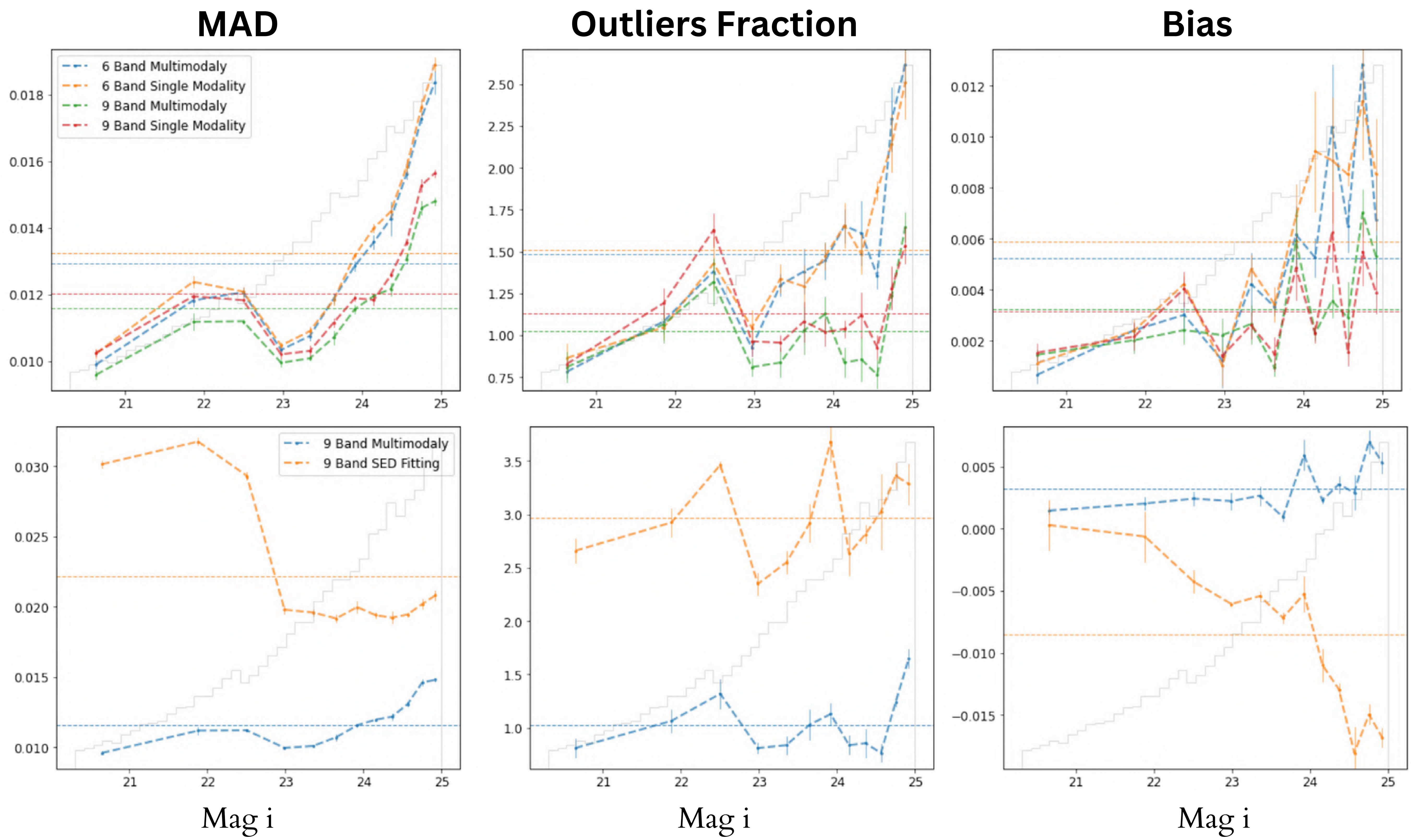
Realistic representative dataset with best redshift labels available



Second Contribution

Initial Results

Good Overall Cross Validated Performance

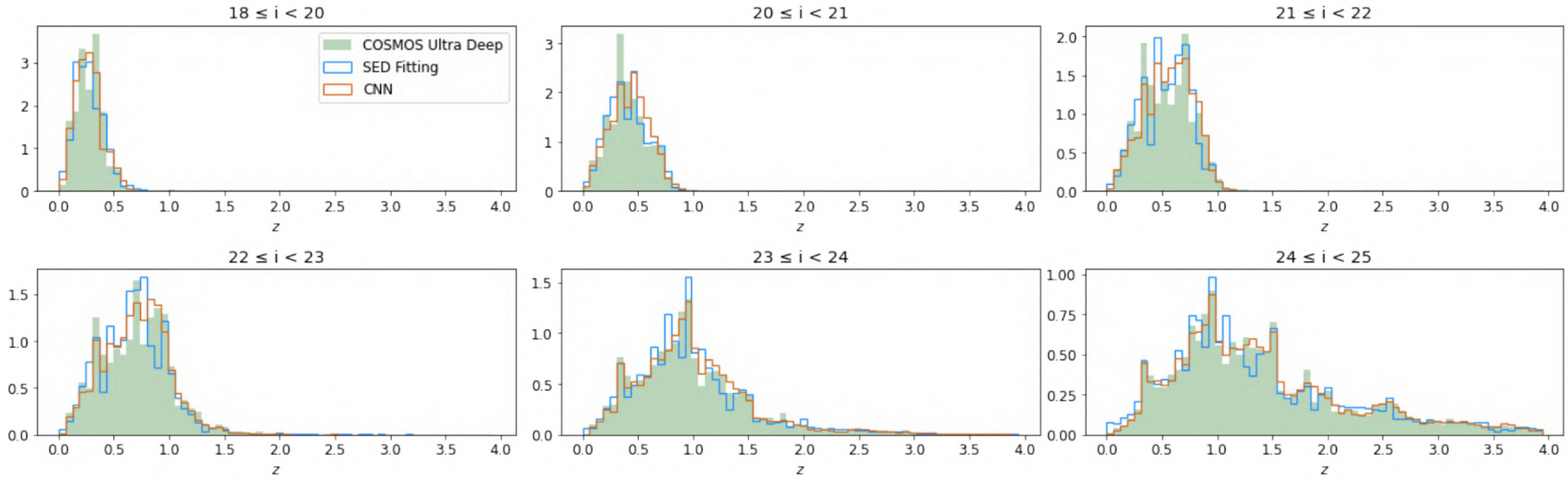


Second Contribution

Initial Results

N(z) Retrieval Performance

Cross Validation Results

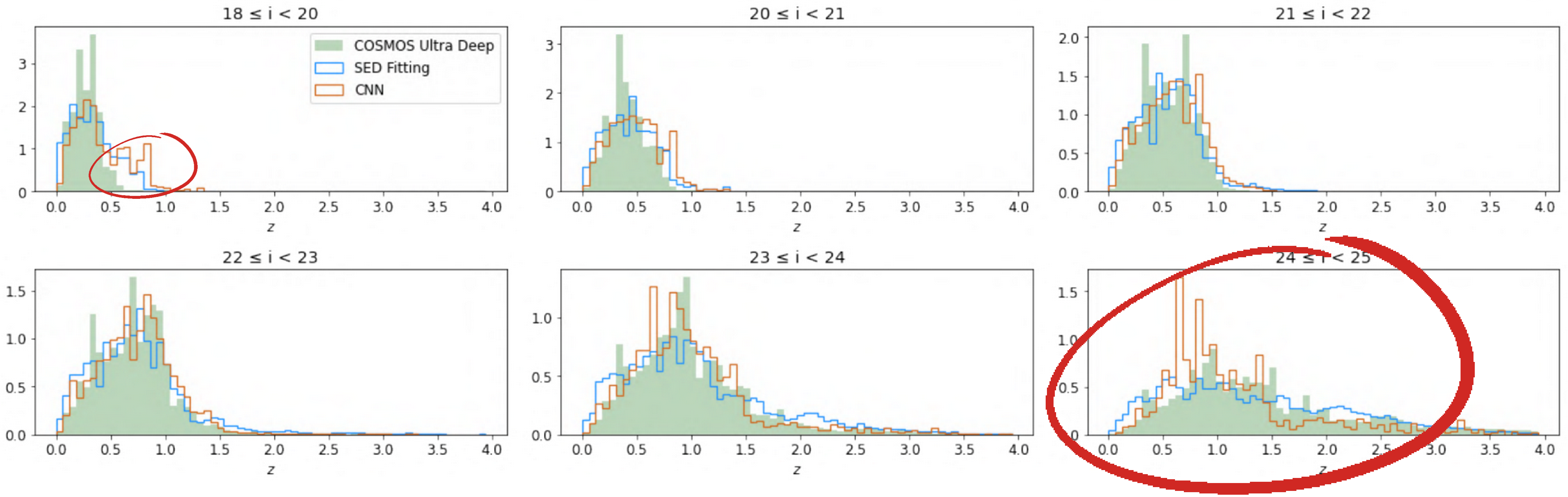


Second Contribution

Initial Results

N(z) Retrieval Performance

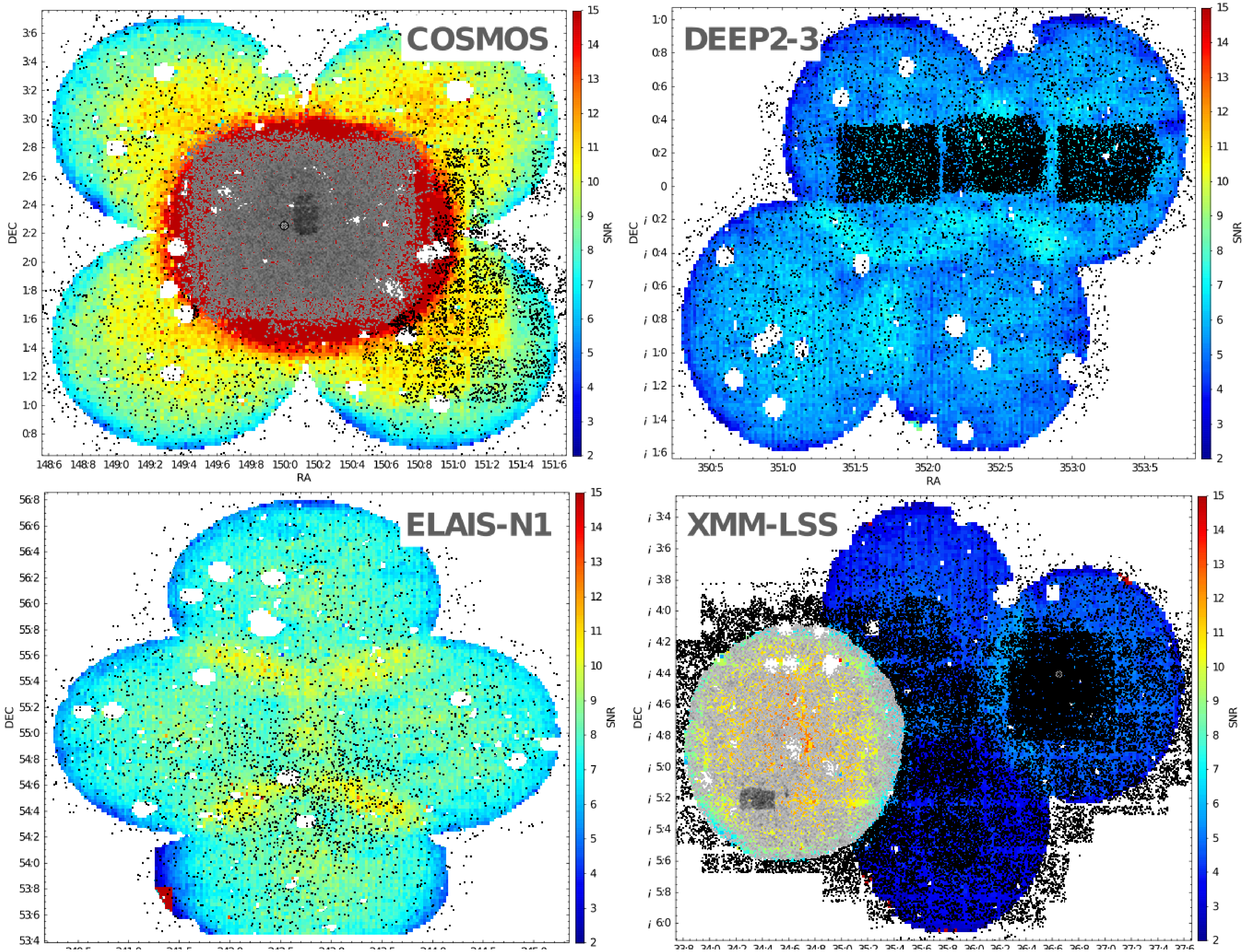
Inference on unlabeled set Results



Second Contribution

Domain Mismatch Problem

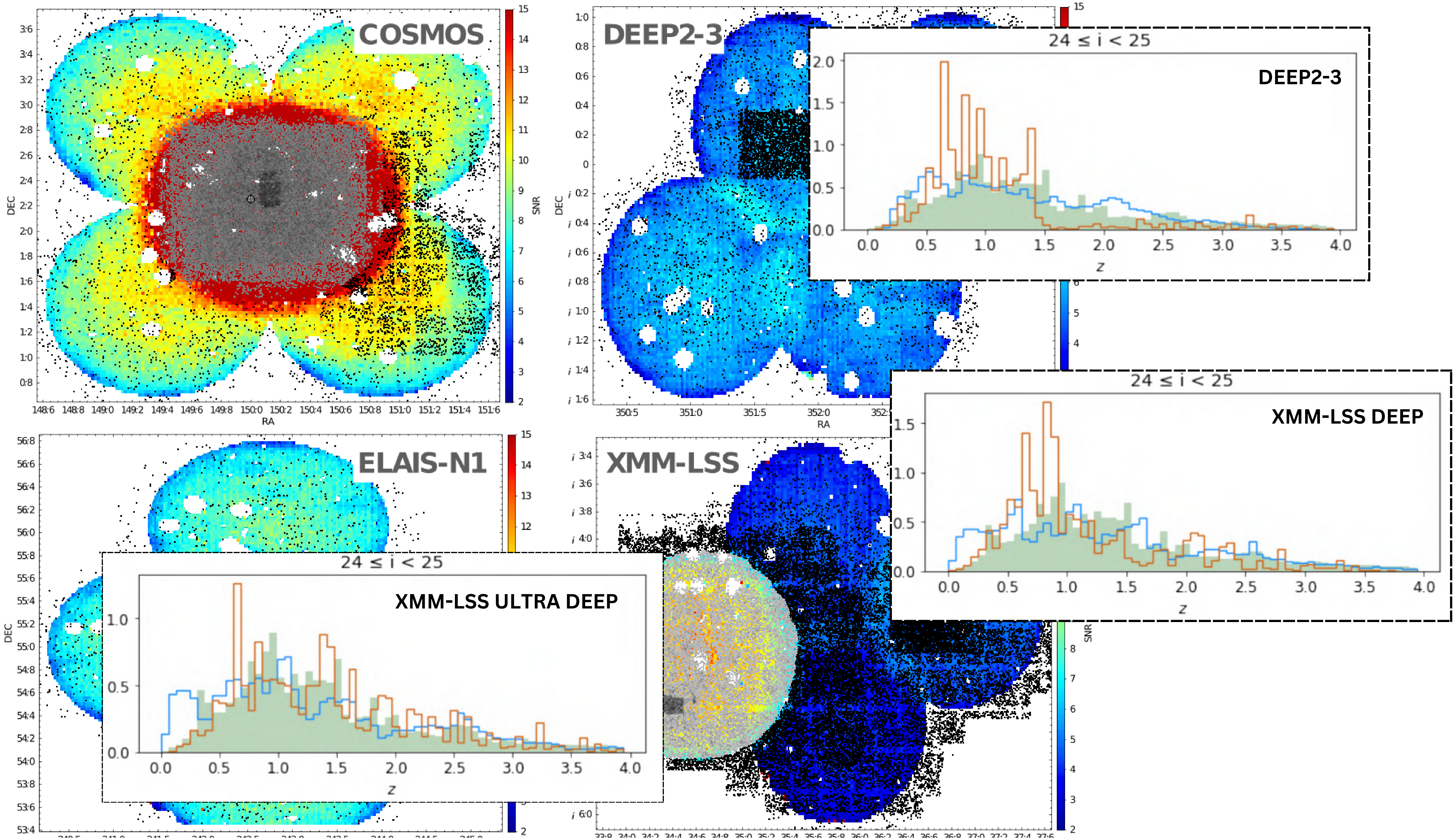
Different image acquisition conditions



Second Contribution

Domain Mismatch Problem

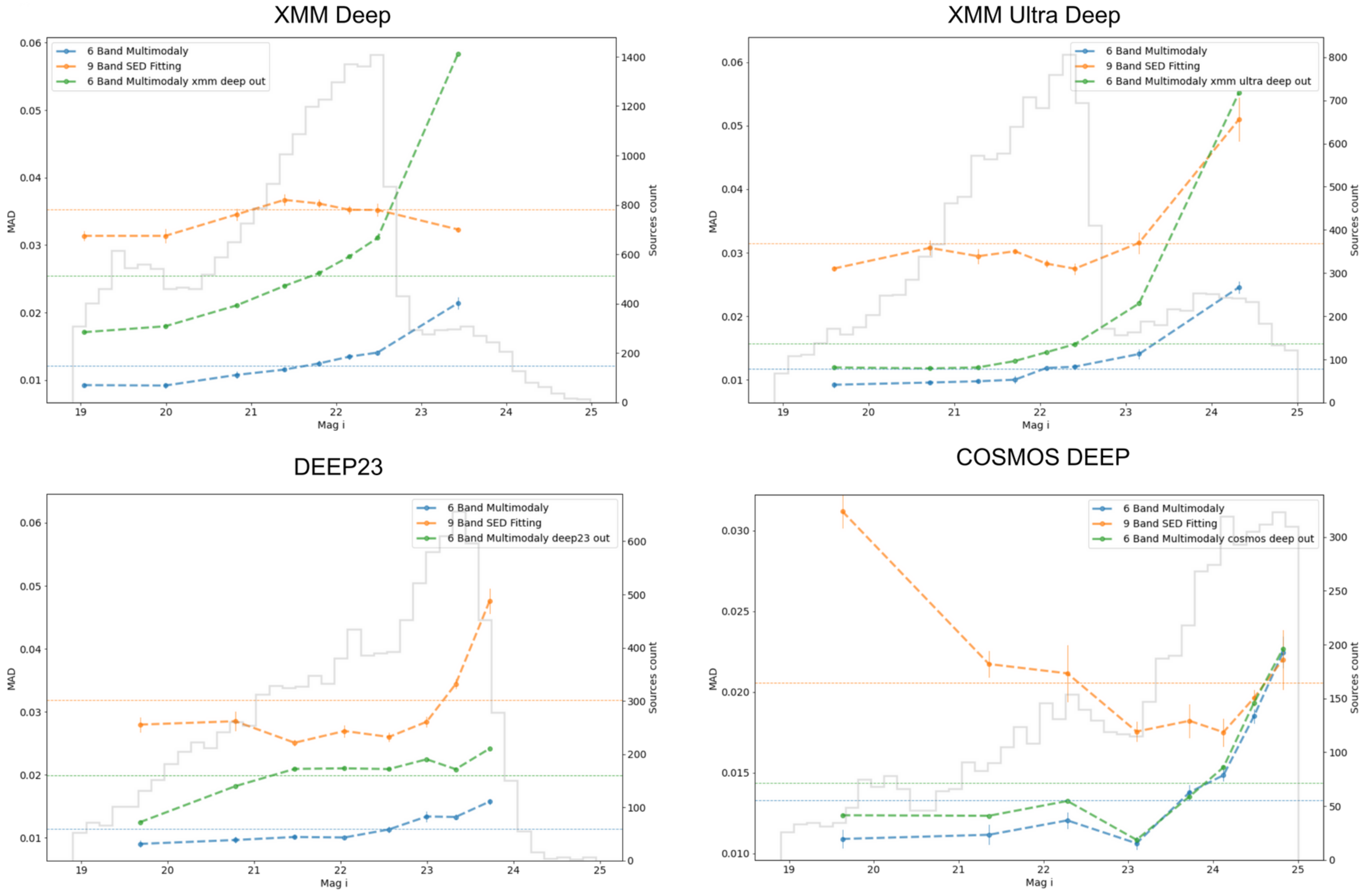
Different image acquisition conditions



Second Contribution

Domain Mismatch Problem

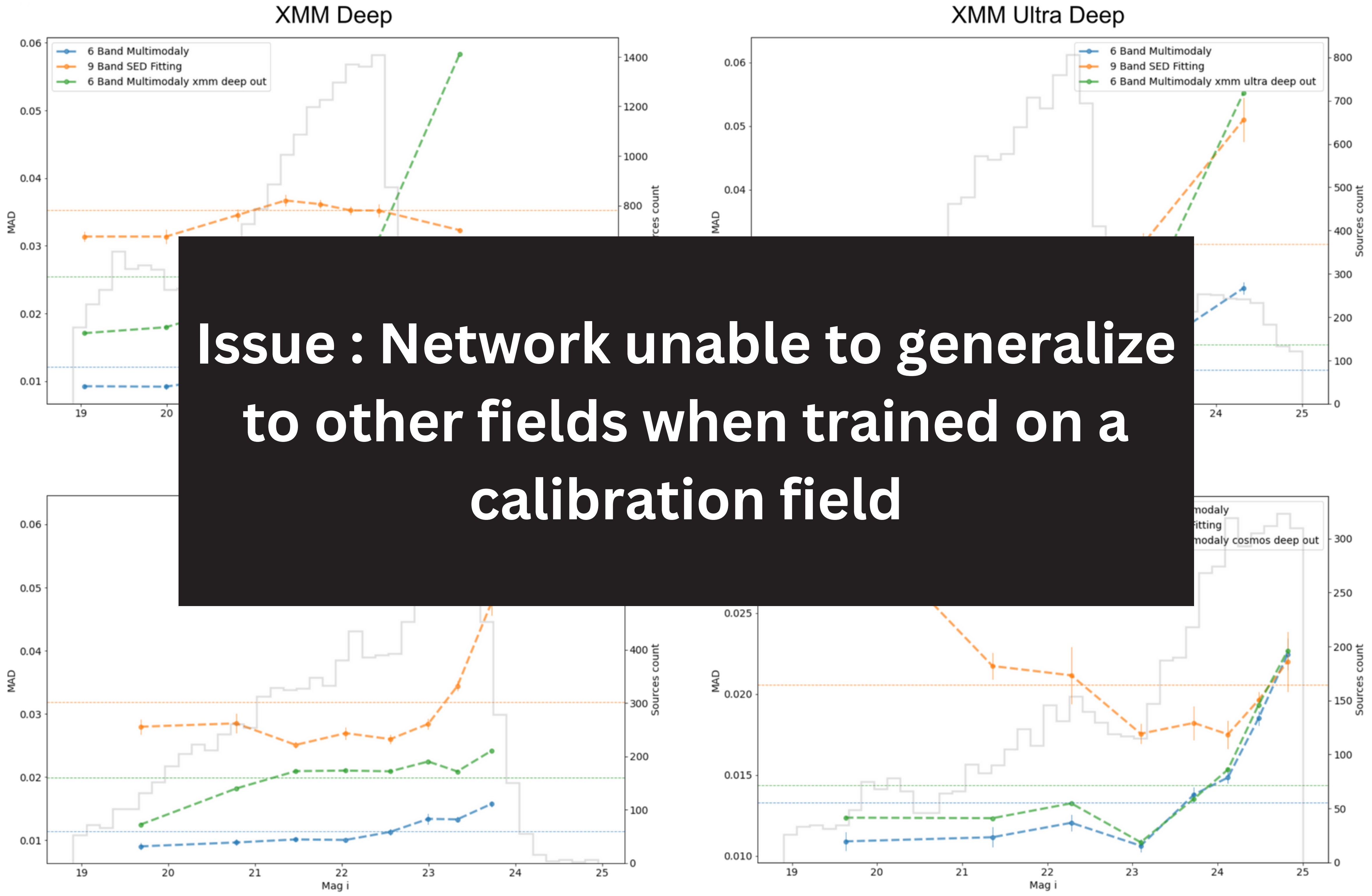
Train on one field and infer on others (COSMOS Ultra Deep)



Second Contribution

Domain Mismatch Problem

Train on one field and infer on others (COSMOS Ultra Deep)



Issue : Network unable to generalize to other fields when trained on a calibration field

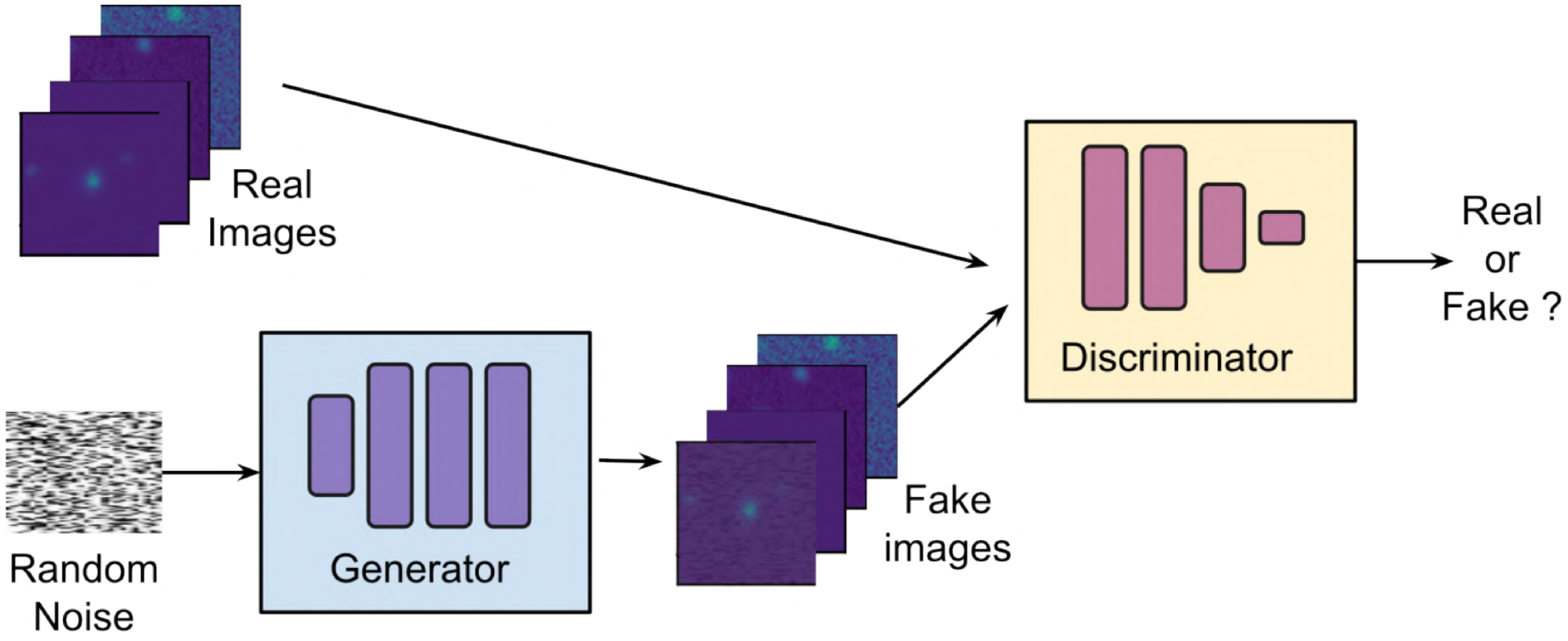
Second Contribution

Proposed Solution

Adversarial Domain Adaptation

Origins : Generative Adversarial Networks (GANs)

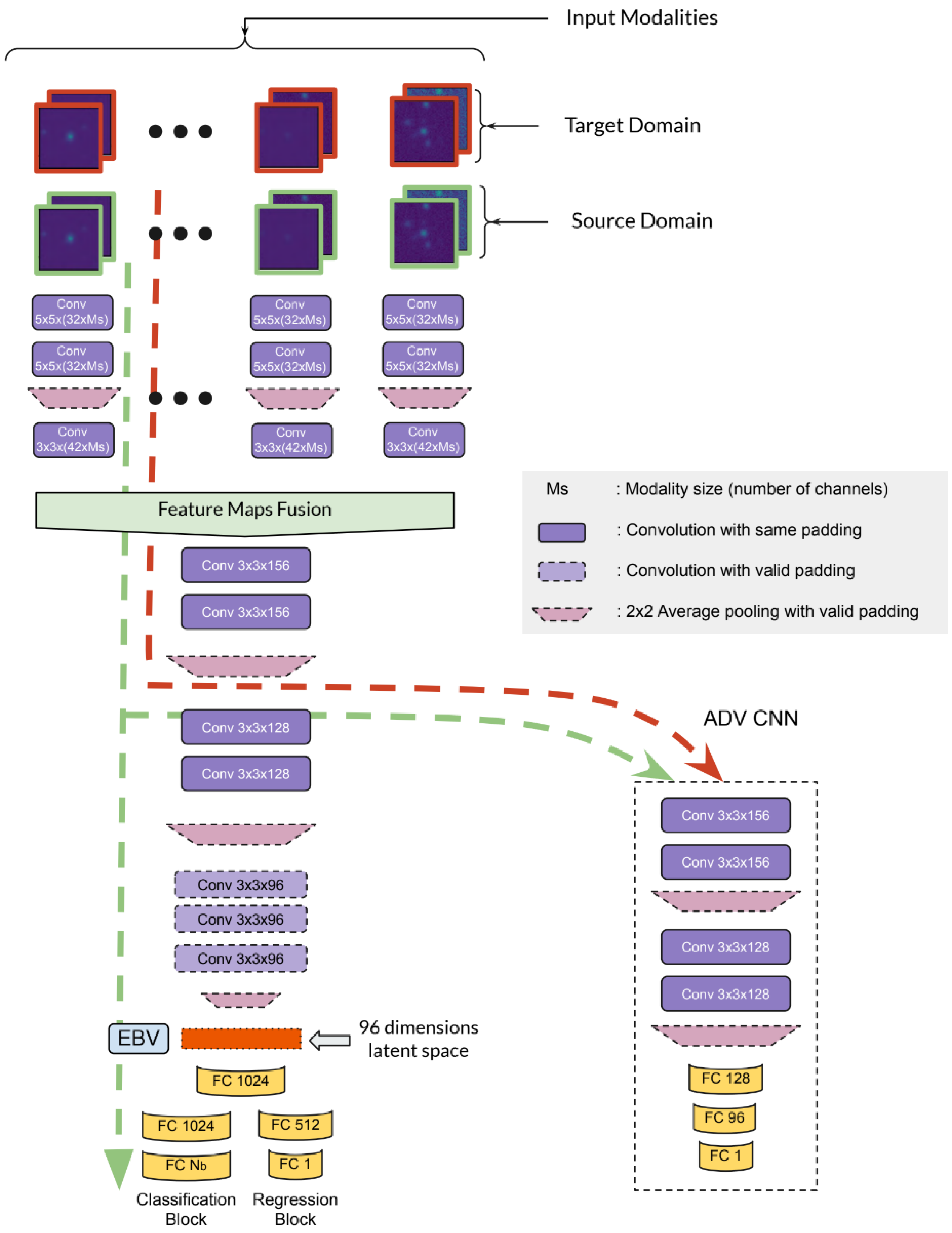
Goodfellow et al. (2014)



Second Contribution

Proposed Solution

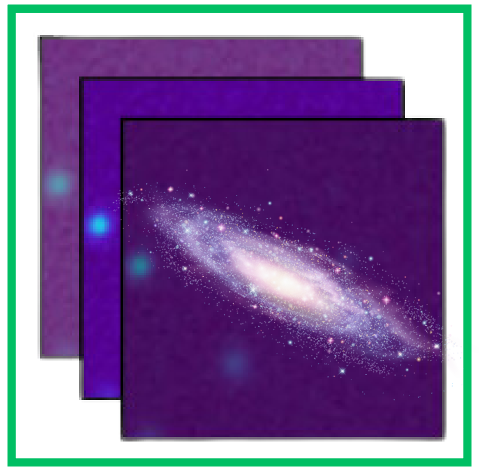
Adversarial Domain Adaptation, The Architecture



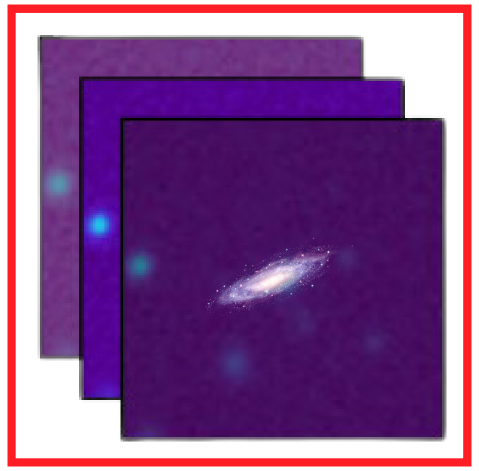
Second Contribution

Proposed Solution

Potential issue : Negative Domain Transfer

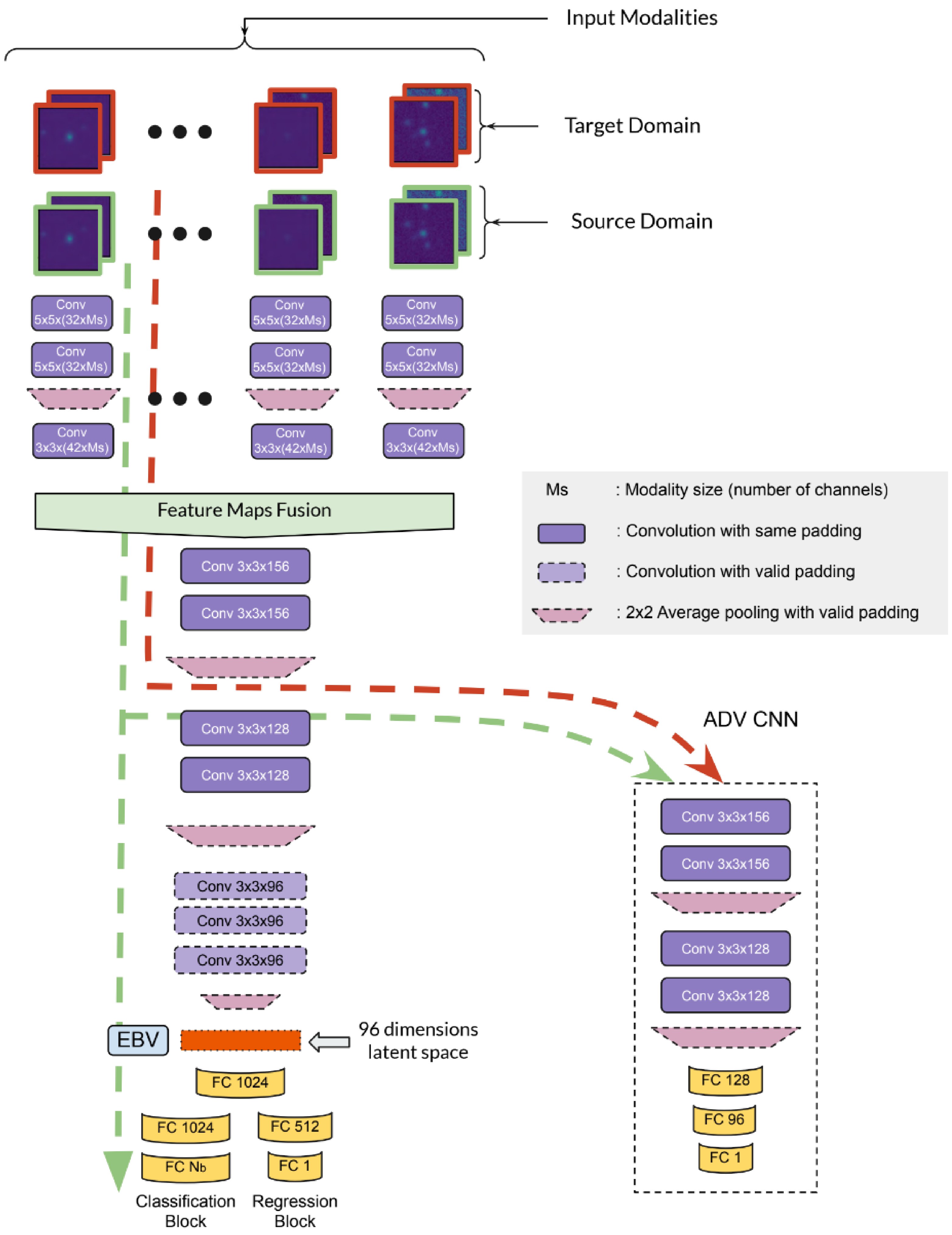


Source Domain



Target Domain

Network encouraged to produce indistinguishable representations

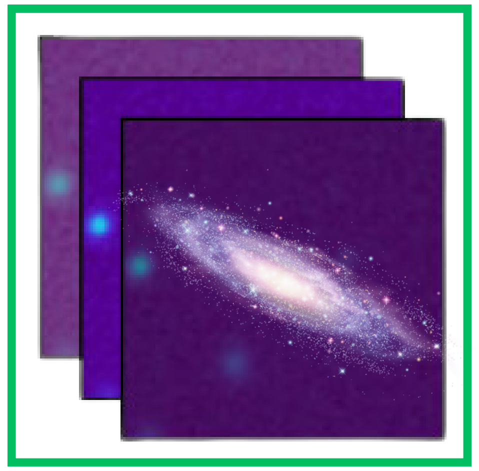


Second Contribution

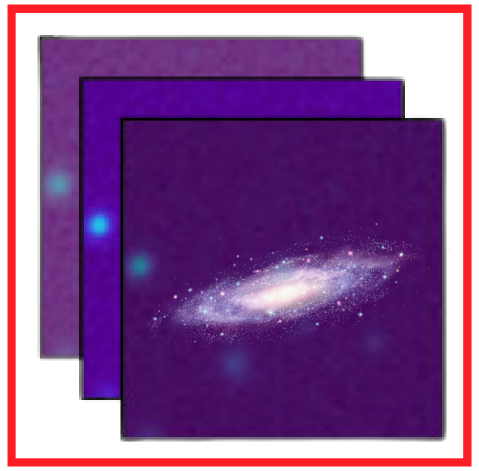
Proposed Solution

Two Steps Fix :

1 - Guided Batch Selection

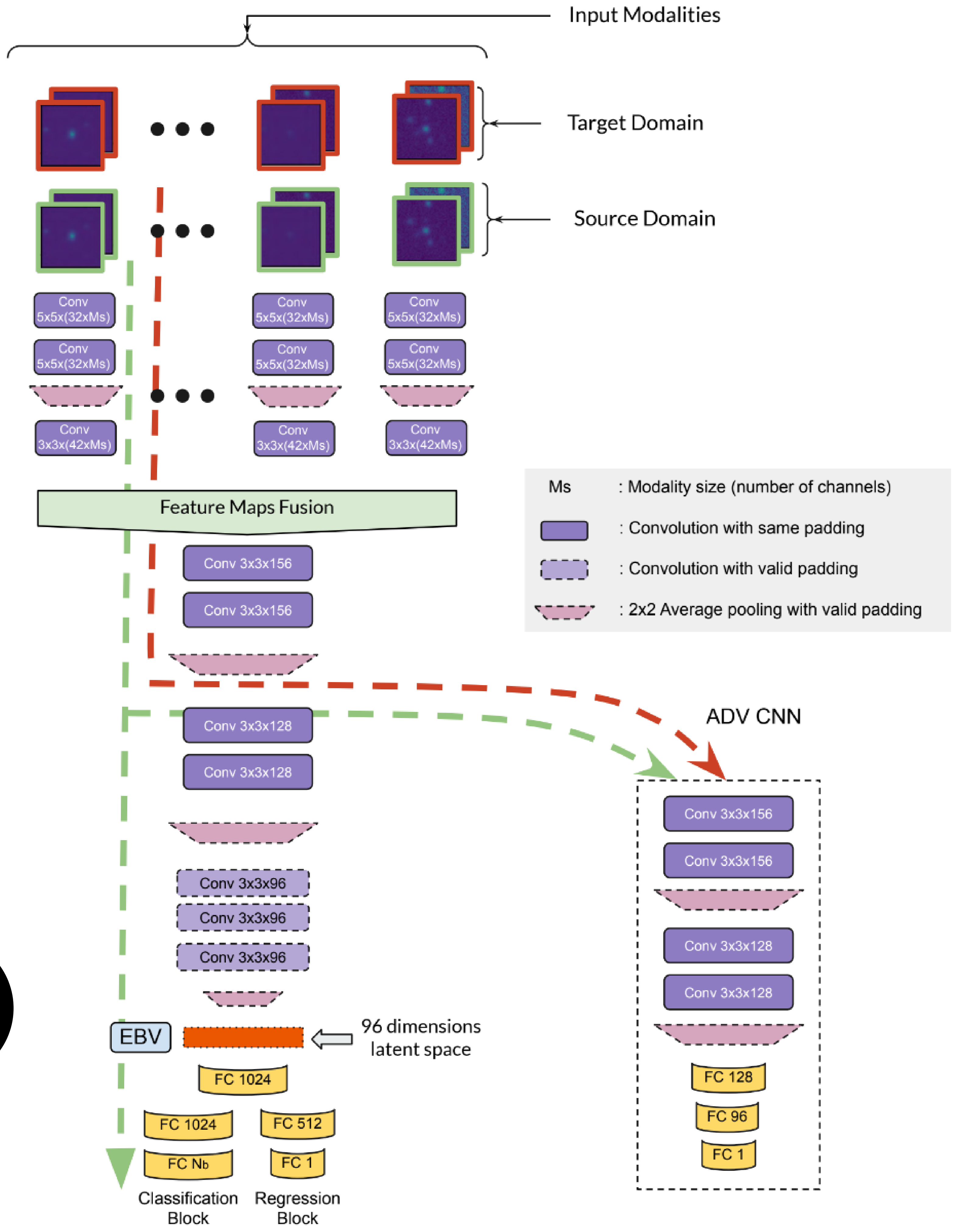


Source Domain



Target Domain

For each source image, a corresponding target image is selected based on photometric magnitude



Second Contribution

Proposed Solution

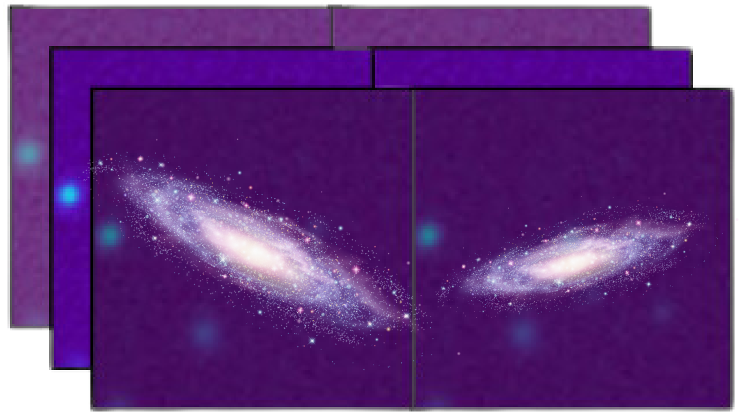
Second Contribution

Two Steps Fix :

1 - Guided Batch Selection

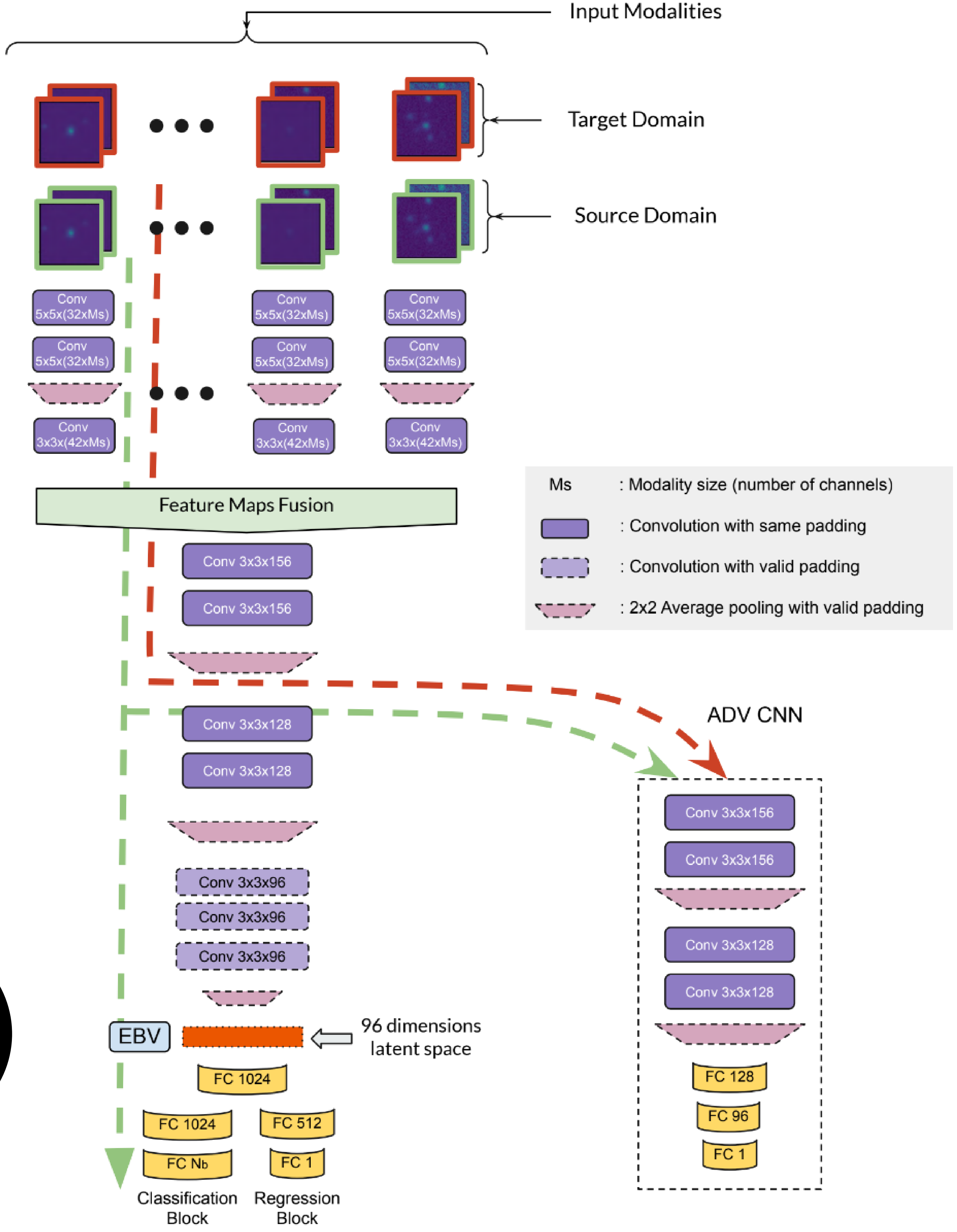
2 - Pairing The Selections

Discriminator Input



Source Domain Target Domain

The selection are paired at the input of the discriminator, which has to estimate if a pair comes from the sam field or not

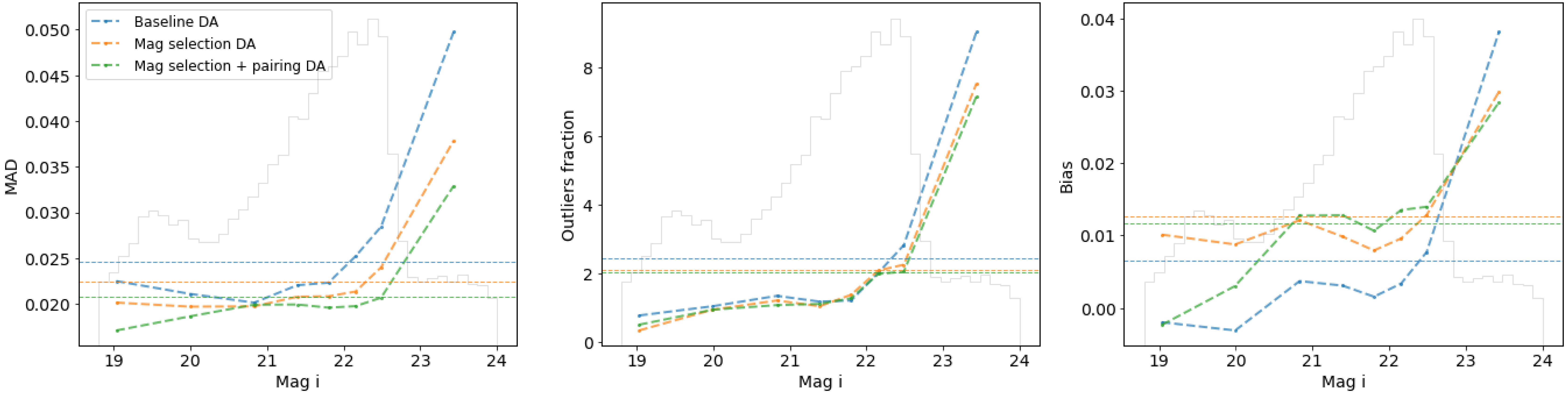


Proposed Solution

XMM DEEP as a study case

Negative Transfer Solution

Training on COSMOS Ultra DEEP, Infering on XMM DEEP



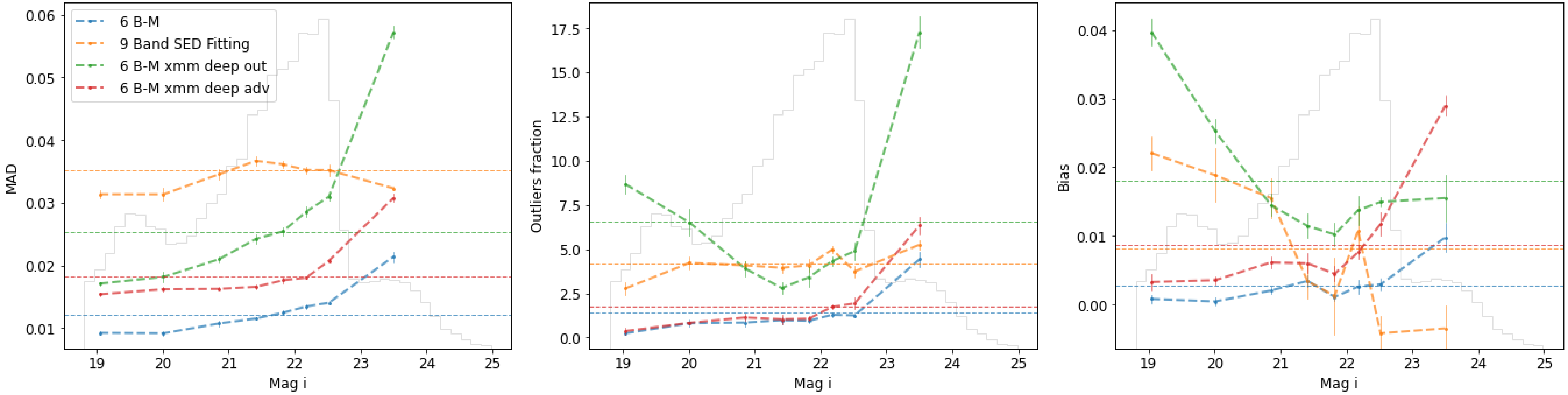
Second Contribution

Proposed Solution

XMM DEEP as a study case

SED Fitting

Classical Cross Validation Results
(Not relaiable at faint magnitudes)



Training On COSMOS Ultra Deep With No DA

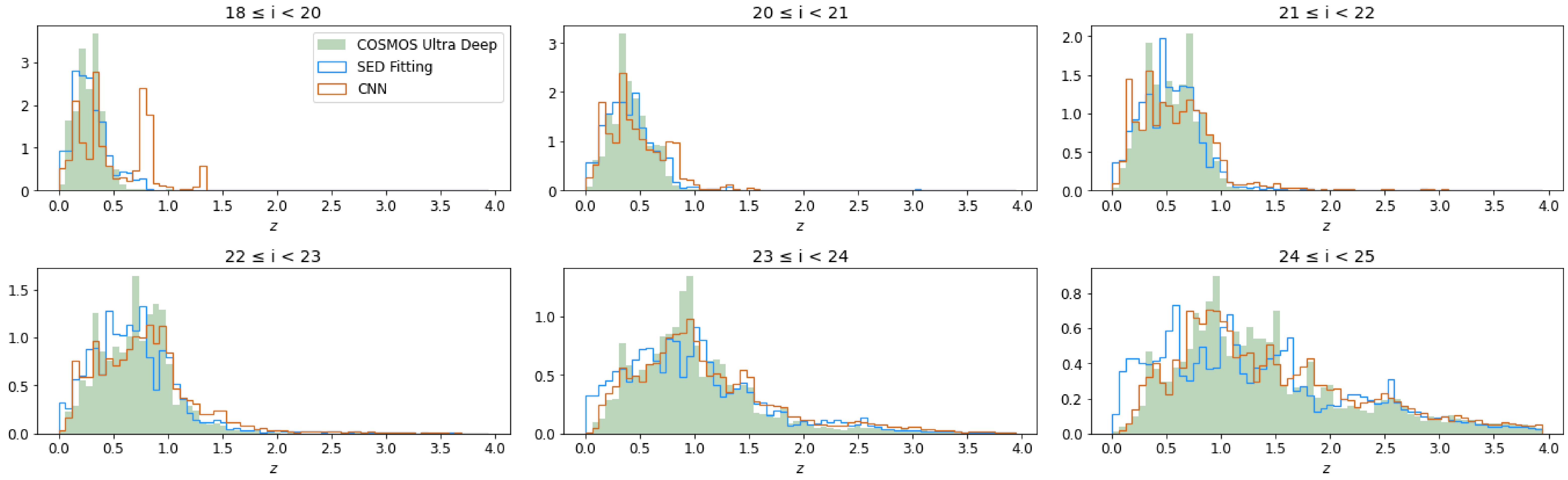
Training On COSMOS Ultra Deep With DA

Second Contribution

Proposed Solution

XMM DEEP as a study case

Good N(z) Retrieval Performance With DA

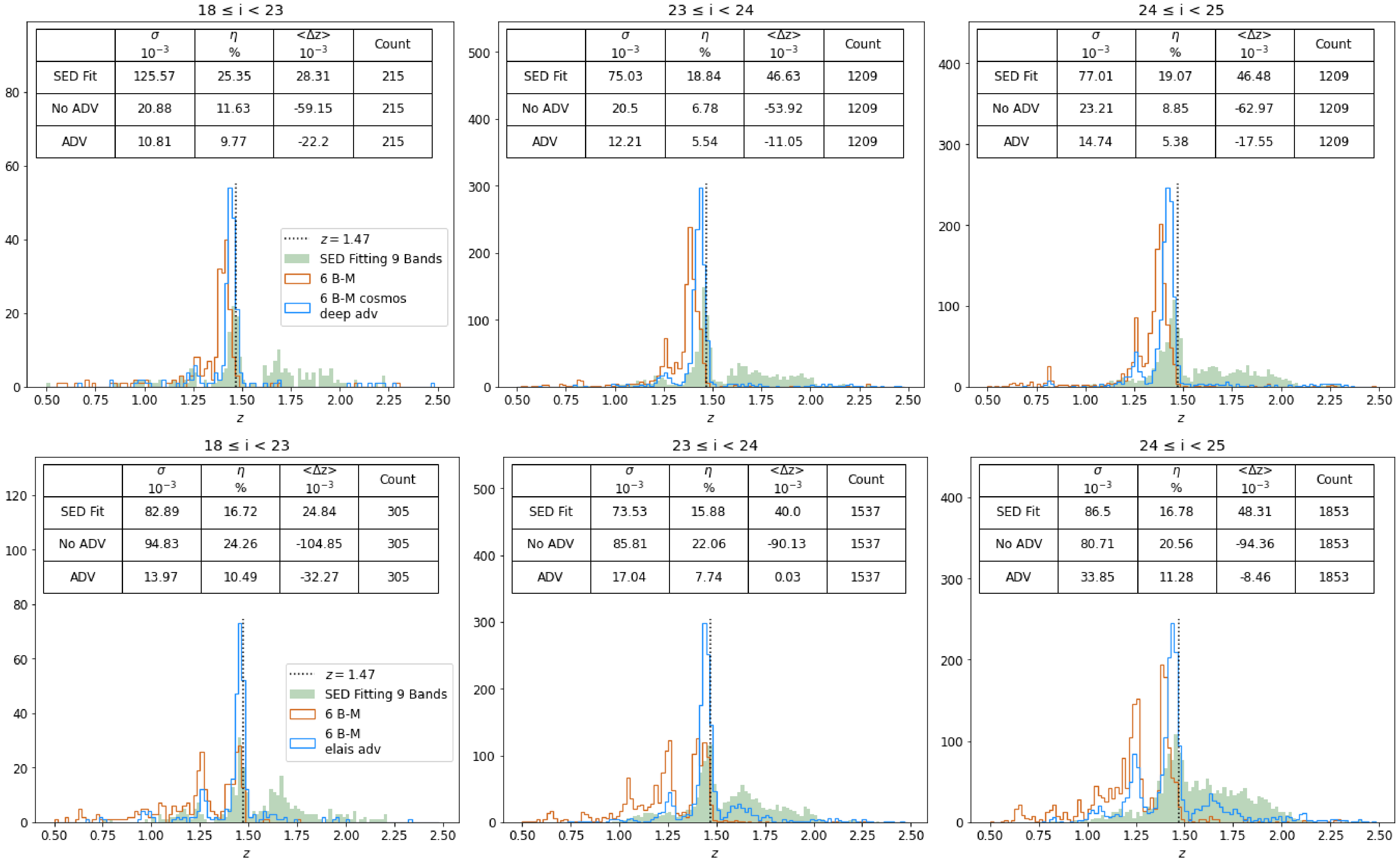


Second Contribution

Proposed Solution

Independent Performance Test

[OII] Emission line galaxies selected from narrowband observations at redshiftz=1.47



COSMOS DEEP FIELD

18 ≤ i < 23

	σ 10^{-3}	η %	<Δz> 10^{-3}	Count
SED Fit	125.57	25.35	28.31	215
No ADV	20.88	11.63	-59.15	215
ADV	10.81	9.77	-22.2	215

23 ≤ i < 24

	σ 10^{-3}	η %	<Δz> 10^{-3}	Count
SED Fit	75.03	18.84	46.63	1209
No ADV	20.5	6.78	-53.92	1209
ADV	12.21	5.54	-11.05	1209

24 ≤ i < 25

	σ 10^{-3}	η %	<Δz> 10^{-3}	Count
SED Fit	77.01	19.07	46.48	1209
No ADV	23.21	8.85	-62.97	1209
ADV	14.74	5.38	-17.55	1209

ELAIS FIELD

18 ≤ i < 23

	σ 10^{-3}	η %	<Δz> 10^{-3}	Count
SED Fit	82.89	16.72	24.84	305
No ADV	94.83	24.26	-104.85	305
ADV	13.97	10.49	-32.27	305

23 ≤ i < 24

	σ 10^{-3}	η %	<Δz> 10^{-3}	Count
SED Fit	73.53	15.88	40.0	1537
No ADV	85.81	22.06	-90.13	1537
ADV	17.04	7.74	0.03	1537

24 ≤ i < 25

	σ 10^{-3}	η %	<Δz> 10^{-3}	Count
SED Fit	86.5	16.78	48.31	1853
No ADV	80.71	20.56	-94.36	1853
ADV	33.85	11.28	-8.46	1853

Second Contribution

Contribution In Writing For Publication

- **Charecterization of the Domain Mismatch Problem For CNNs In Deep Surveys**
- **Adapted Solution using Adversarial Domain Adaptation For Training on A Calibration Field And Genralizing to other fields**

Thank You