HELENA DOMÍNGUEZ SÁNCHEZ

DEEP LEARNING FOR GALAXY MORPHOLOGY

With the contribution of:

M. Huertas-Company
J. Vega Ferrero
R. Sarmiento
I. Damjanov
M. Bernardi
F. Buitrago
C. Bottrell
G.Martin
J. Tous

Centro de Estudios de Física del Cosmos de Aragón

IAP COLLOQUIUM - 1ST DECEMBER 2023

Ayuda ICTS-MRR-2021-03-CEFCA, financiada por:
Financiado por la Unión Europea
NextGenerationEU

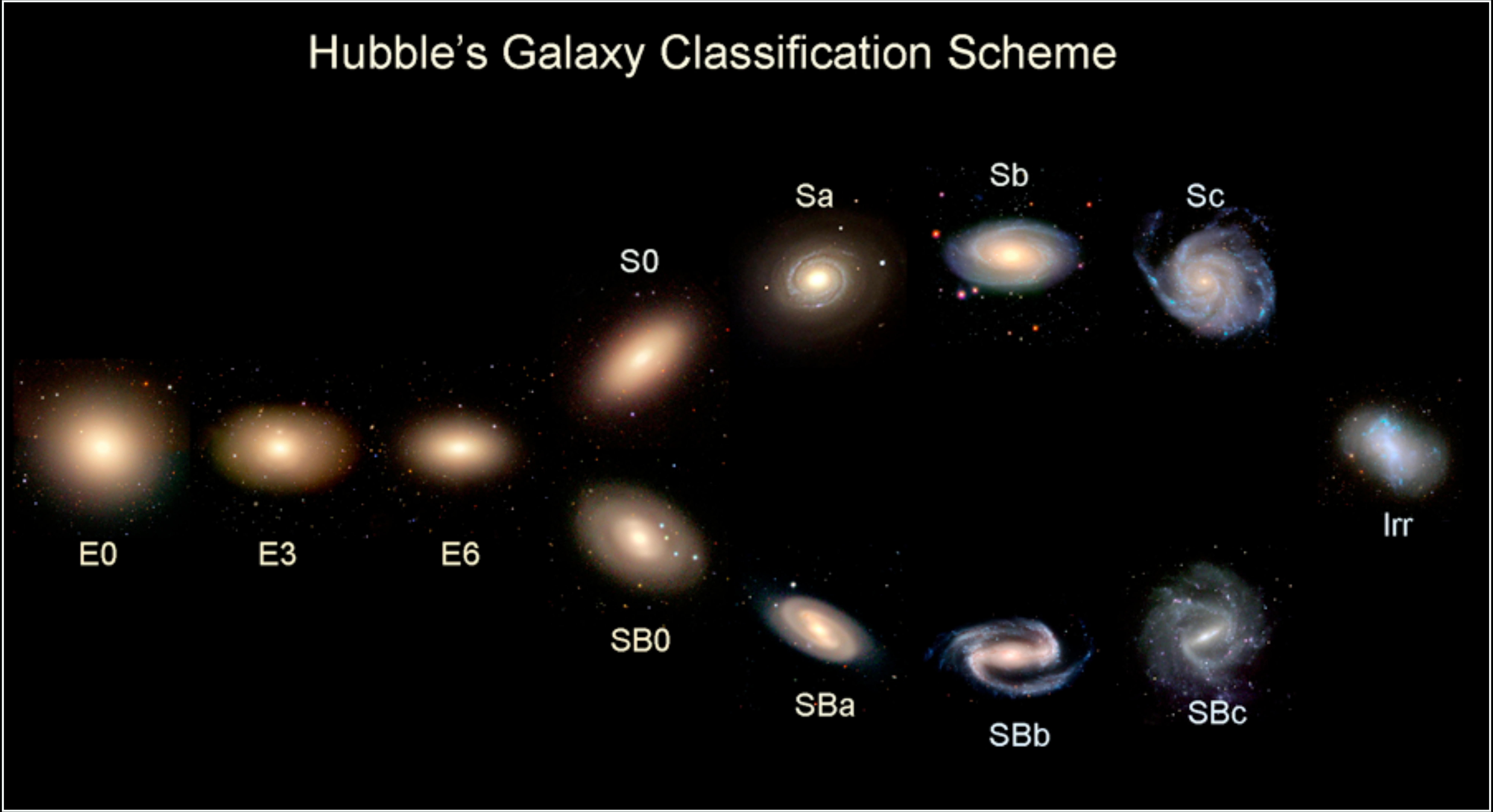
OUTLINE

- Why care about galaxy morphology?
- Why Deep Learning?
- Supervised learning: CNNs
- What if there are no labels?
  - Transfer Learning
  - ‘Emulations’
  - Unsupervised learning
- Going deeper

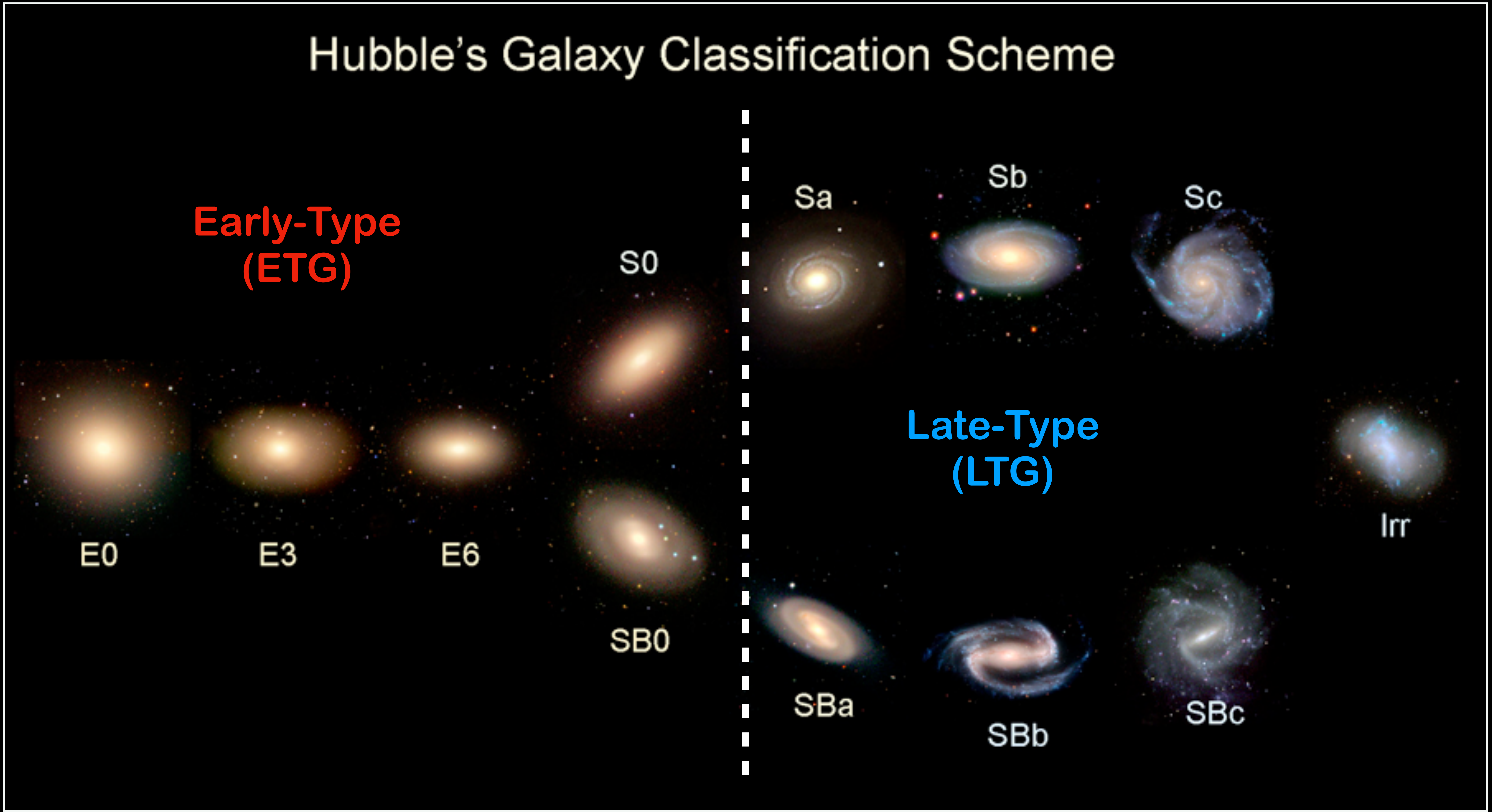
OUTLINE

- Why care about galaxy morphology?
- Why Deep Learning?
- Supervised learning: CNNs
- What if there are no labels?
  - Transfer Learning
  - ‘Emulations’
  - Unsupervised learning
- Going deeper

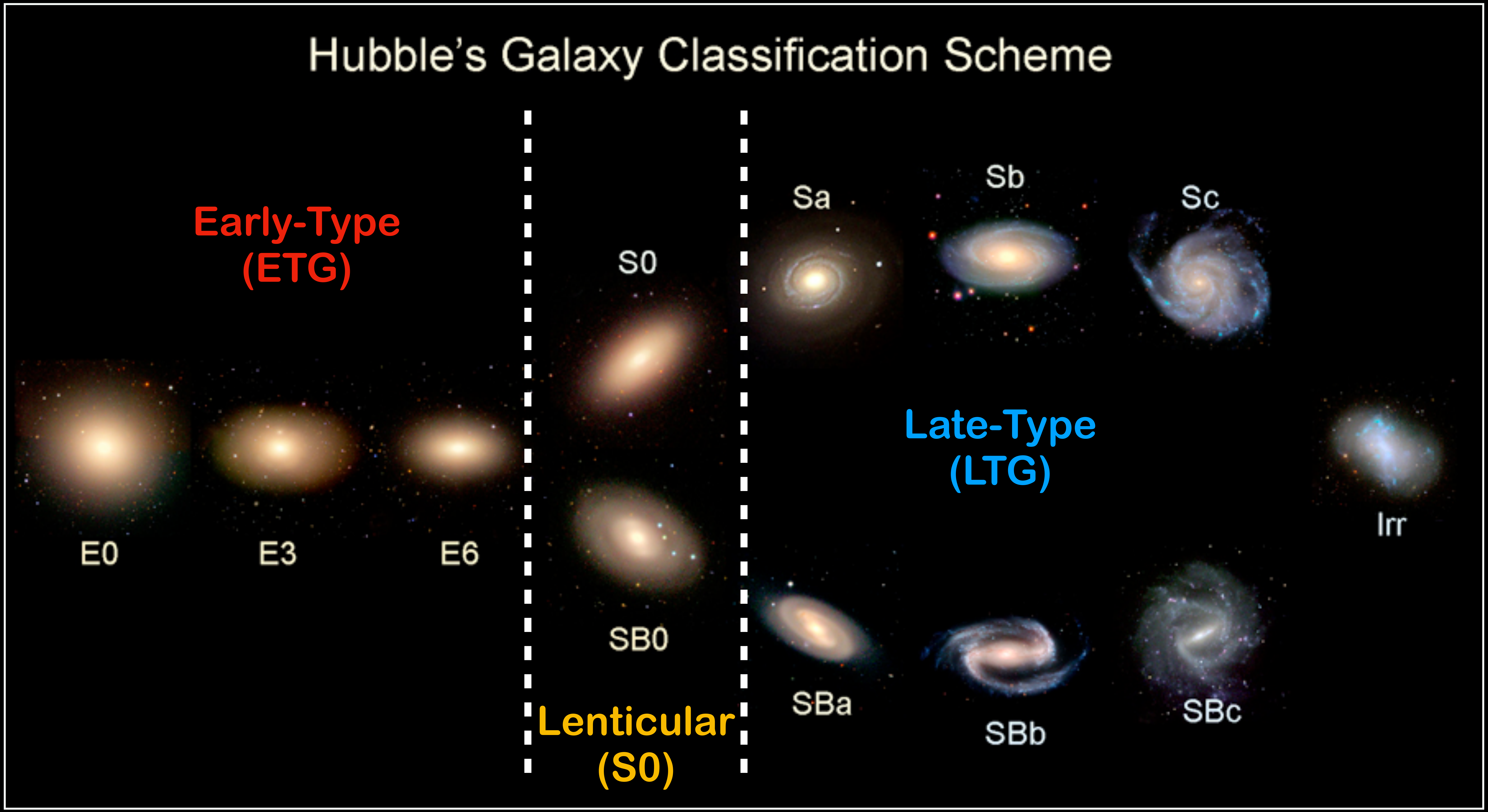
Hubble+1926,1936



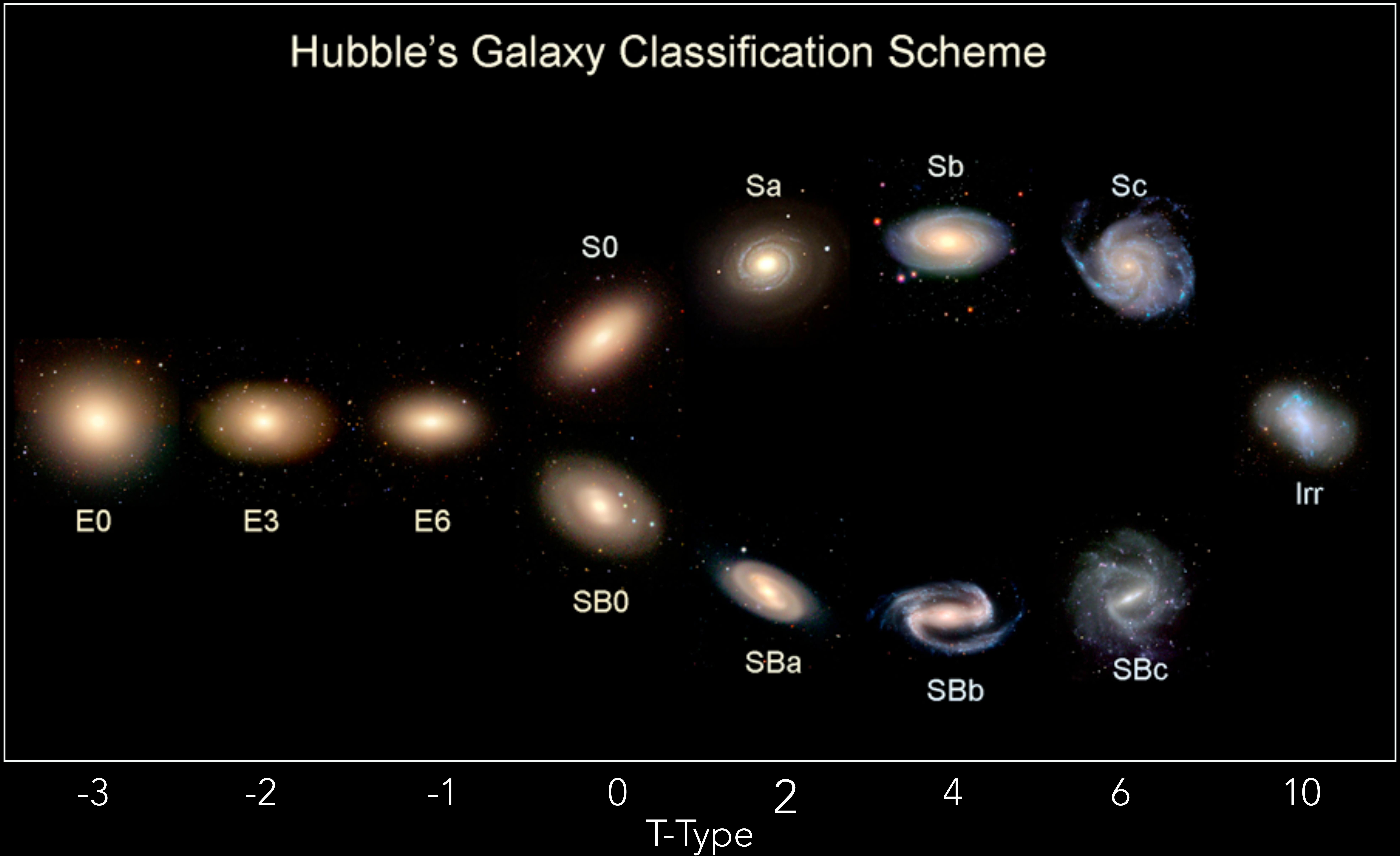
Hubble+1926,1936



Hubble+1926,1936



de Vaucouleurs+1963



SPIRAL
(LTG)

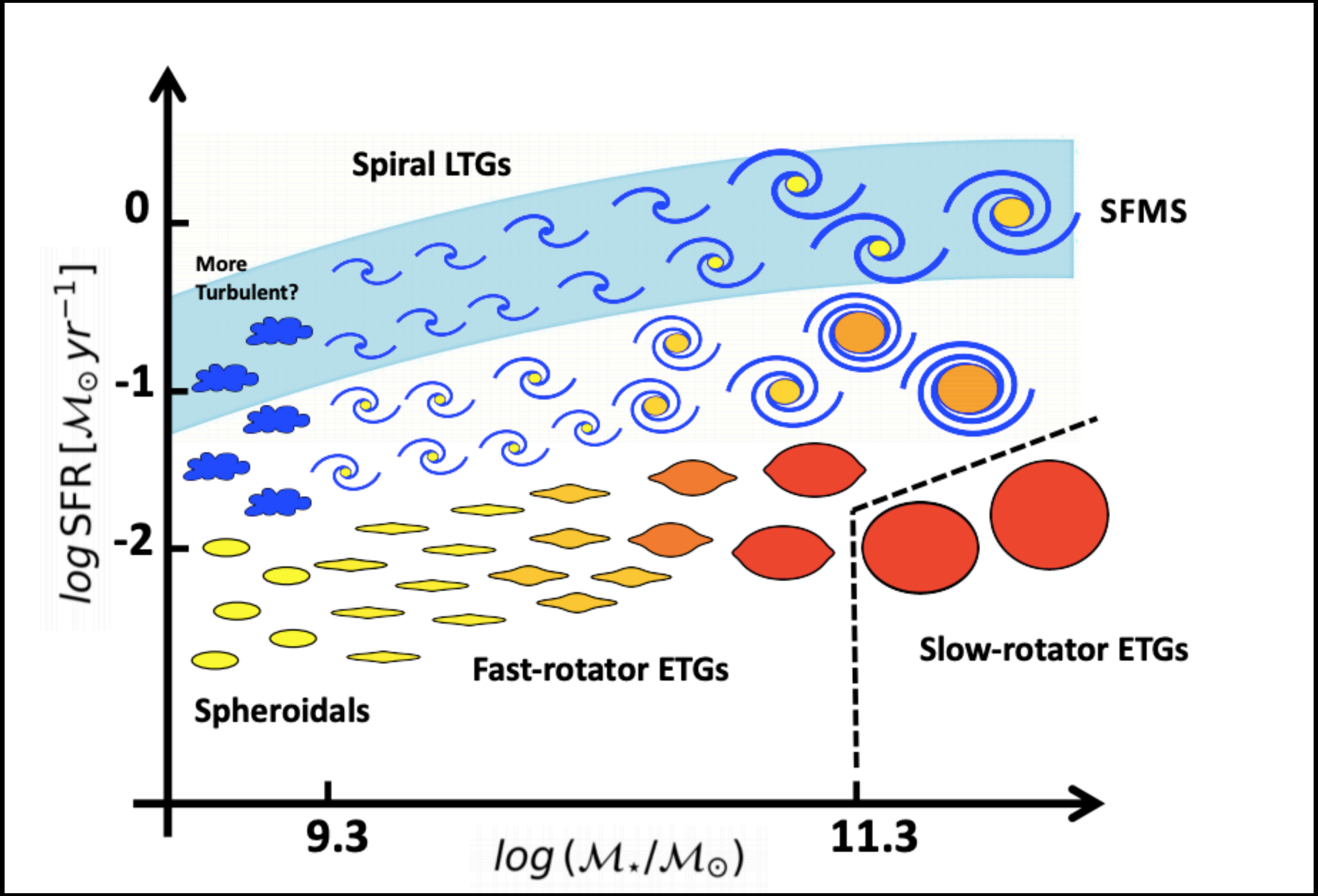
- Intermediate/low mass galaxies
- Young stellar populations
- On-going star formation
- Rotationally supported

ELLIPTICAL
(ETG)

- Most massive galaxies
- Old stellar populations (red colors)
- Kinematics dominated by velocity dispersion

Morphology related to fundamental galaxy properties

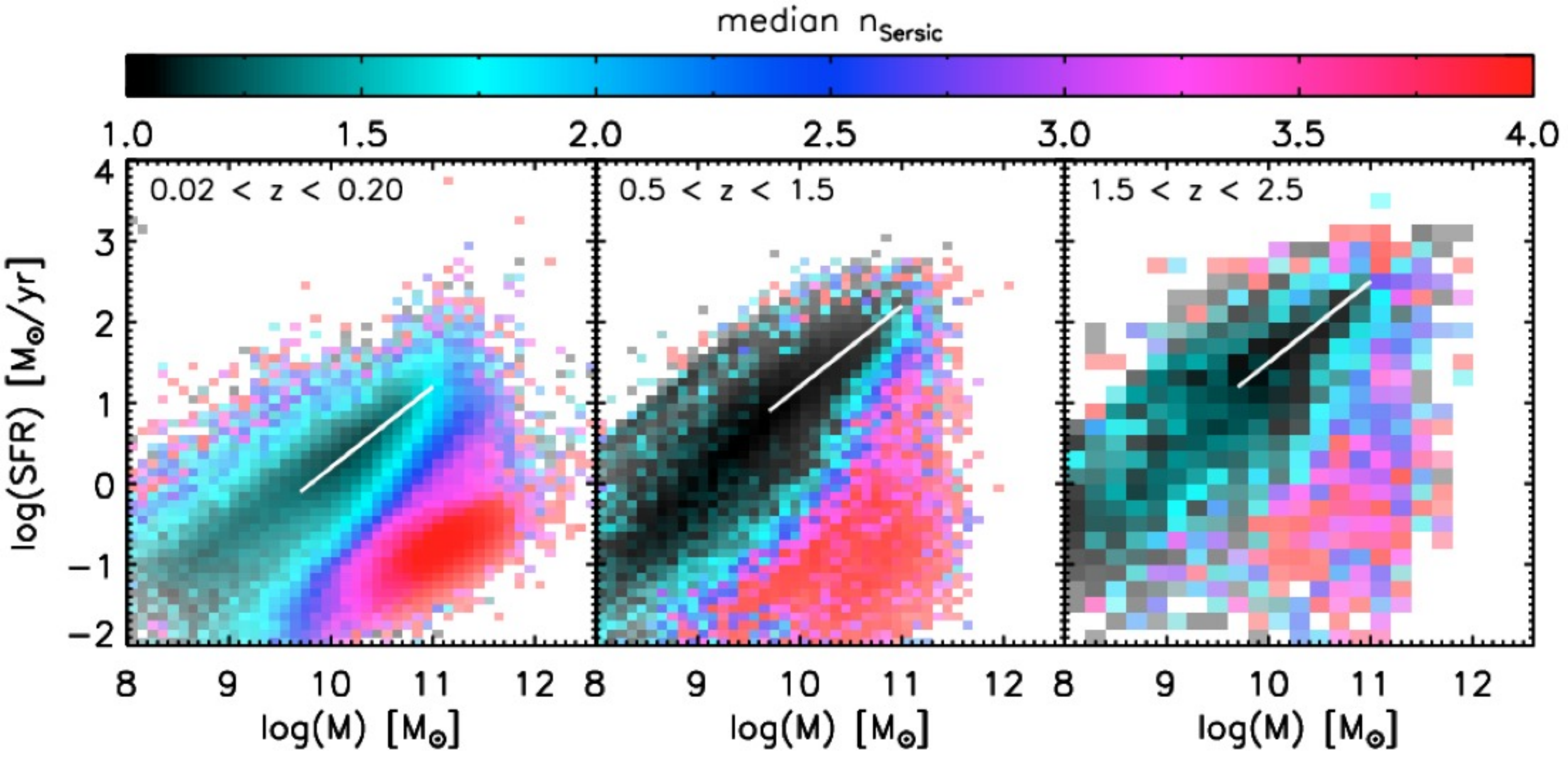
Number of stars formed per year



Galaxy Mass (in units of solar masses)

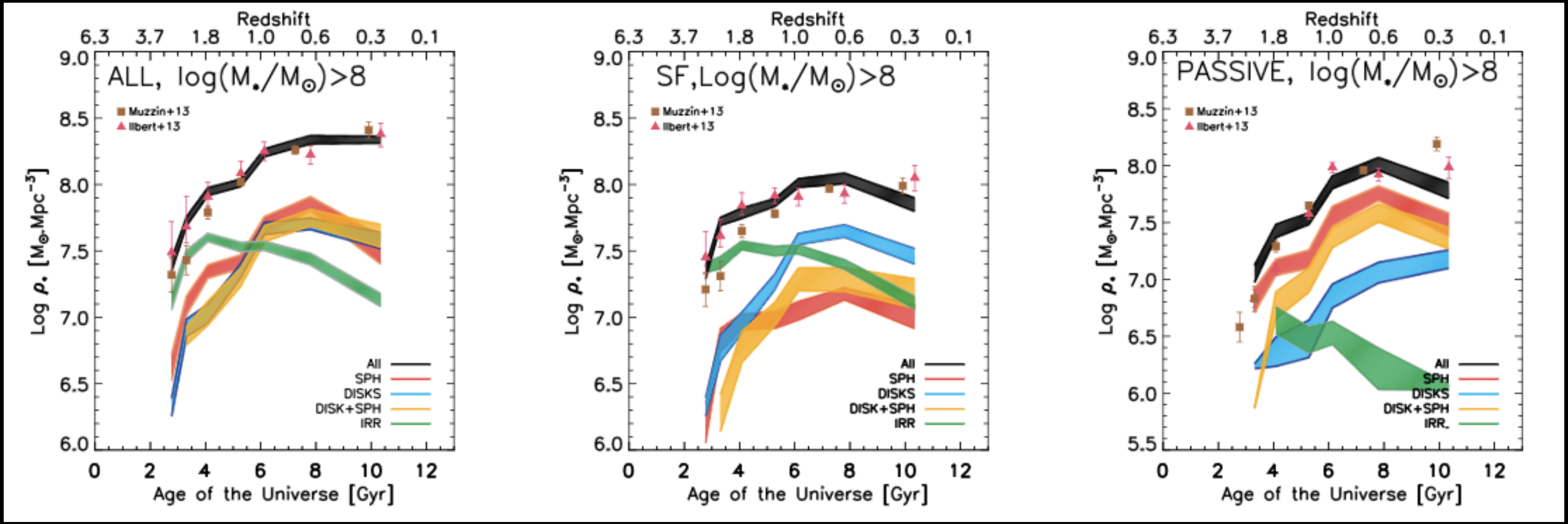
Wang+2020

Morphology related to fundamental galaxy properties



Wuyts+2011

Morphology-SFR evolution



ALL • SPH • DISKS • DISKS+SPH • IRR

Huertas-Company+2016

Understanding how morphology relates to galaxy properties and in which way they affect galaxy assembly and evolution is one of the major challenges of present day astronomy.

It is crucial to have accurate morphological classifications for large samples of galaxies

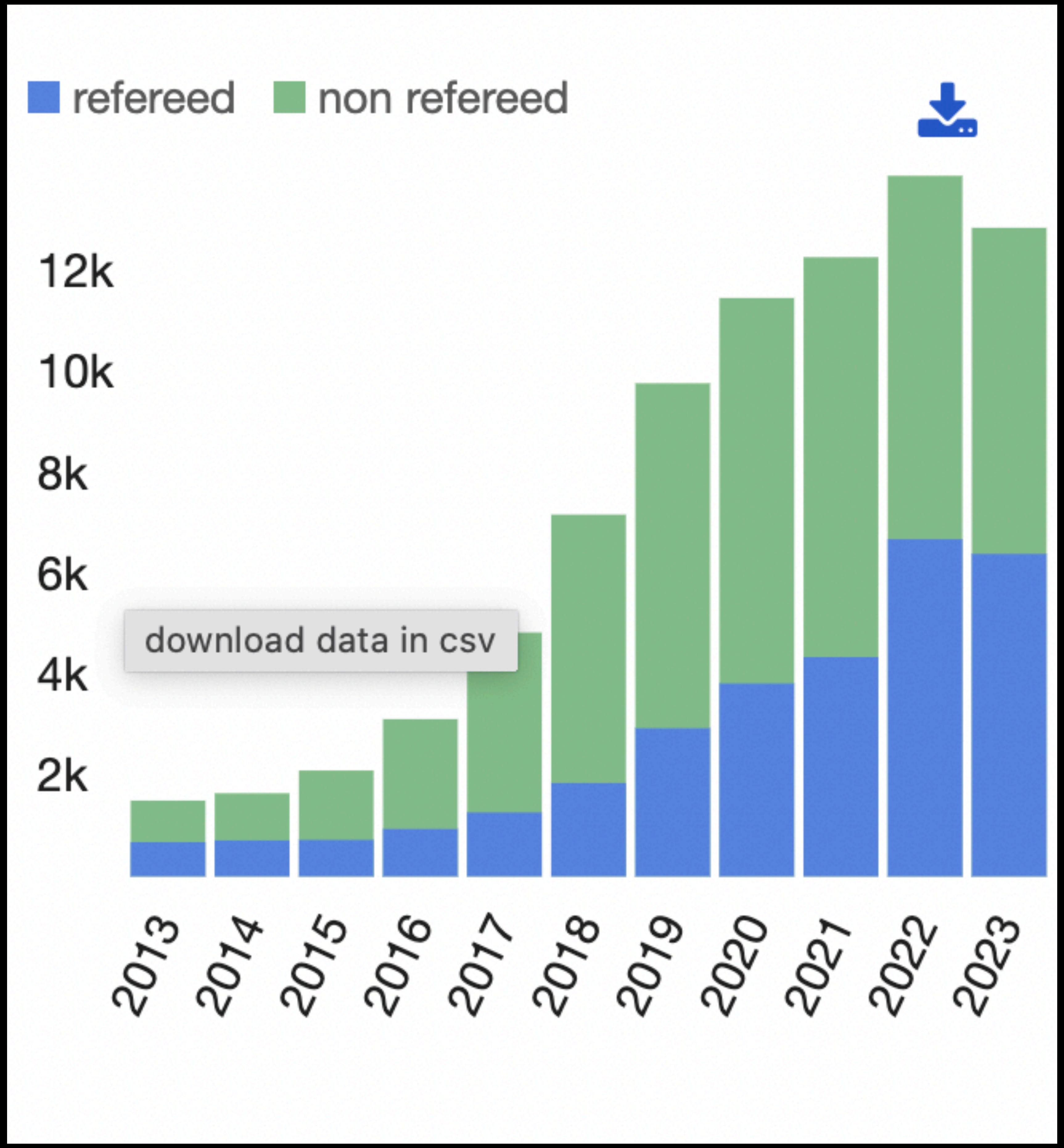
– (SUPERVISED) DEEP LEARNING –

OUTLINE

- Why care about galaxy morphology?
- Why Deep Learning?
- Supervised learning: CNNs
- What if there are no labels?
  - Transfer Learning
  - 'Emulations'
  - Unsupervised learning
- Going deeper

WHY DEEP LEARNING?

- Astronomy is already in the Big Data Era (Euclid, LSST, etc.)
- DL is optimized for image analysis (e.g., MINST)
- DL has been successfully applied to many astronomical tasks (classification, regression, clustering).
- DL has matured a lot and is starting to be normalised as a methodology tool.



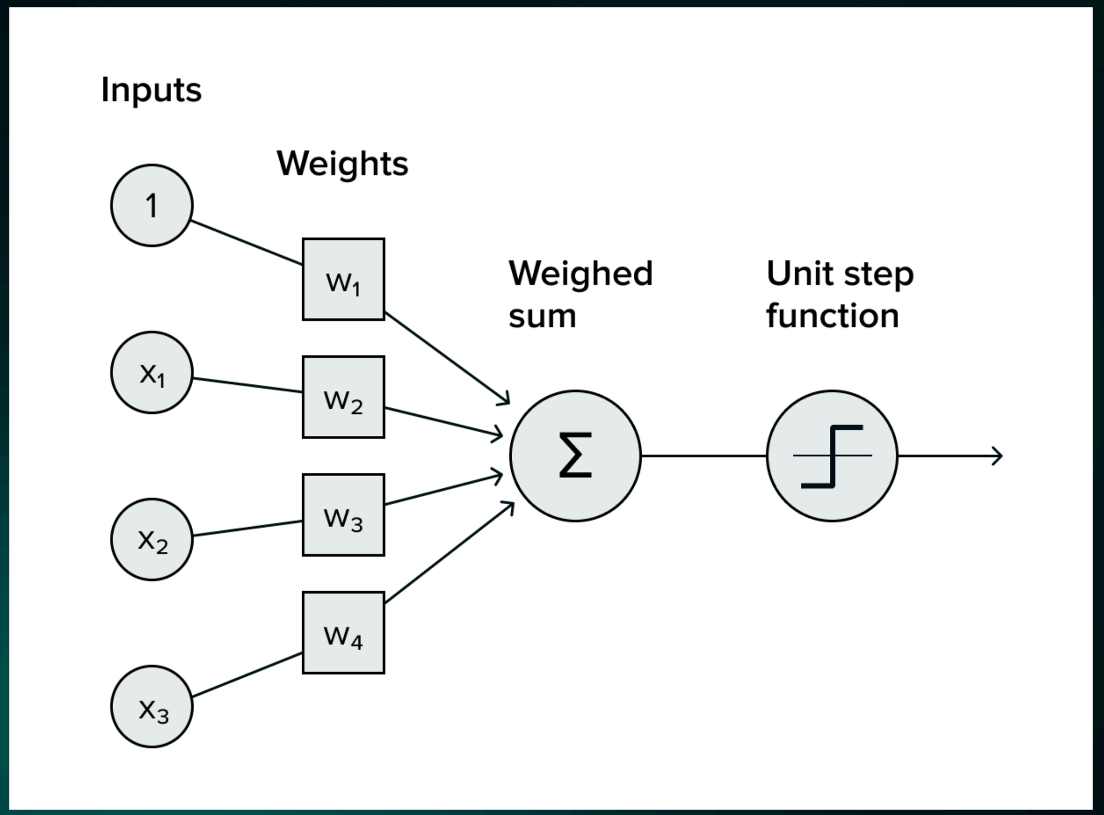
Papers with ‘convolutional’ in their abstract

Source: NASA ADS

WHAT IS DEEP LEARNING?

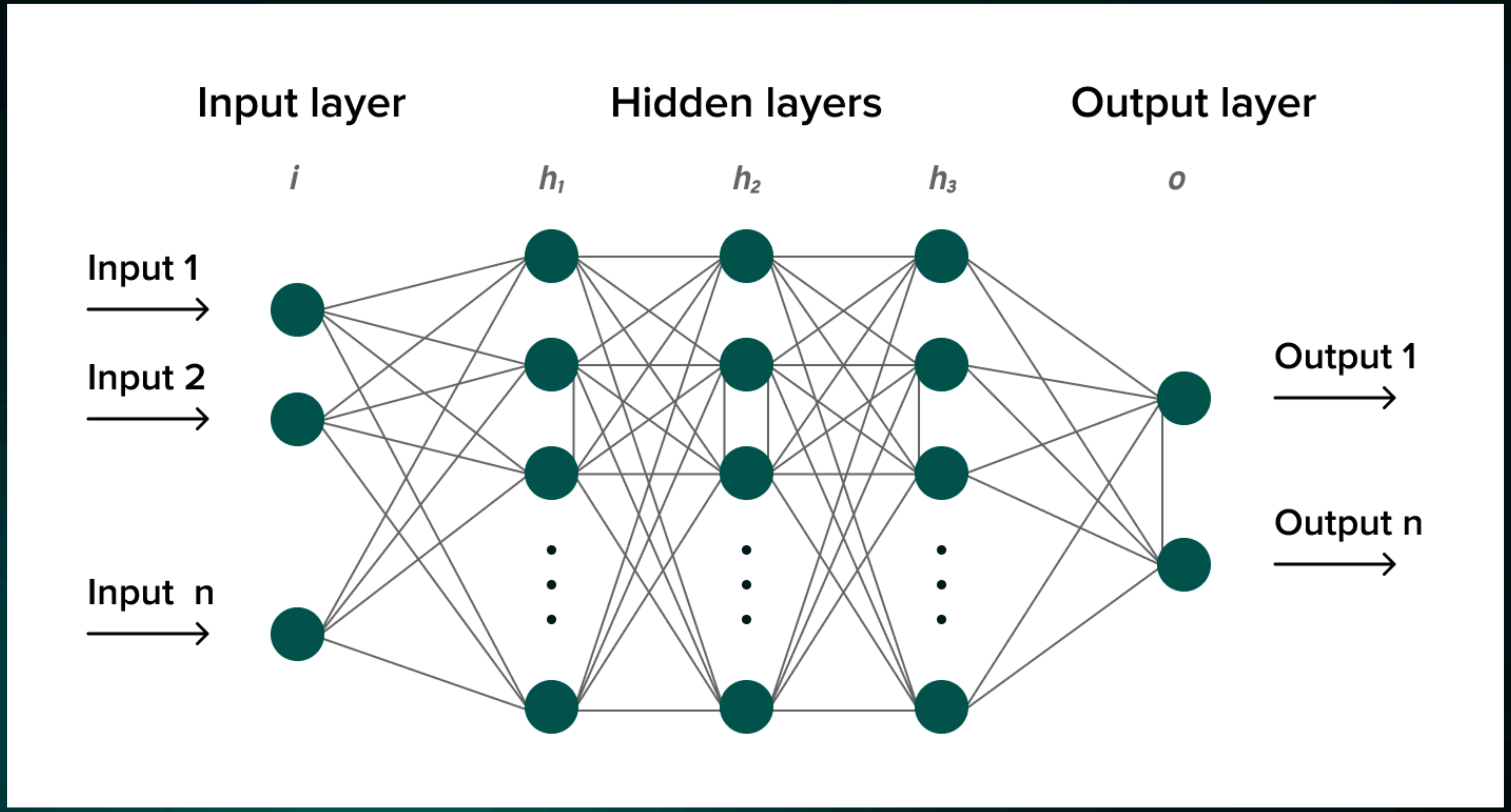
y = input values

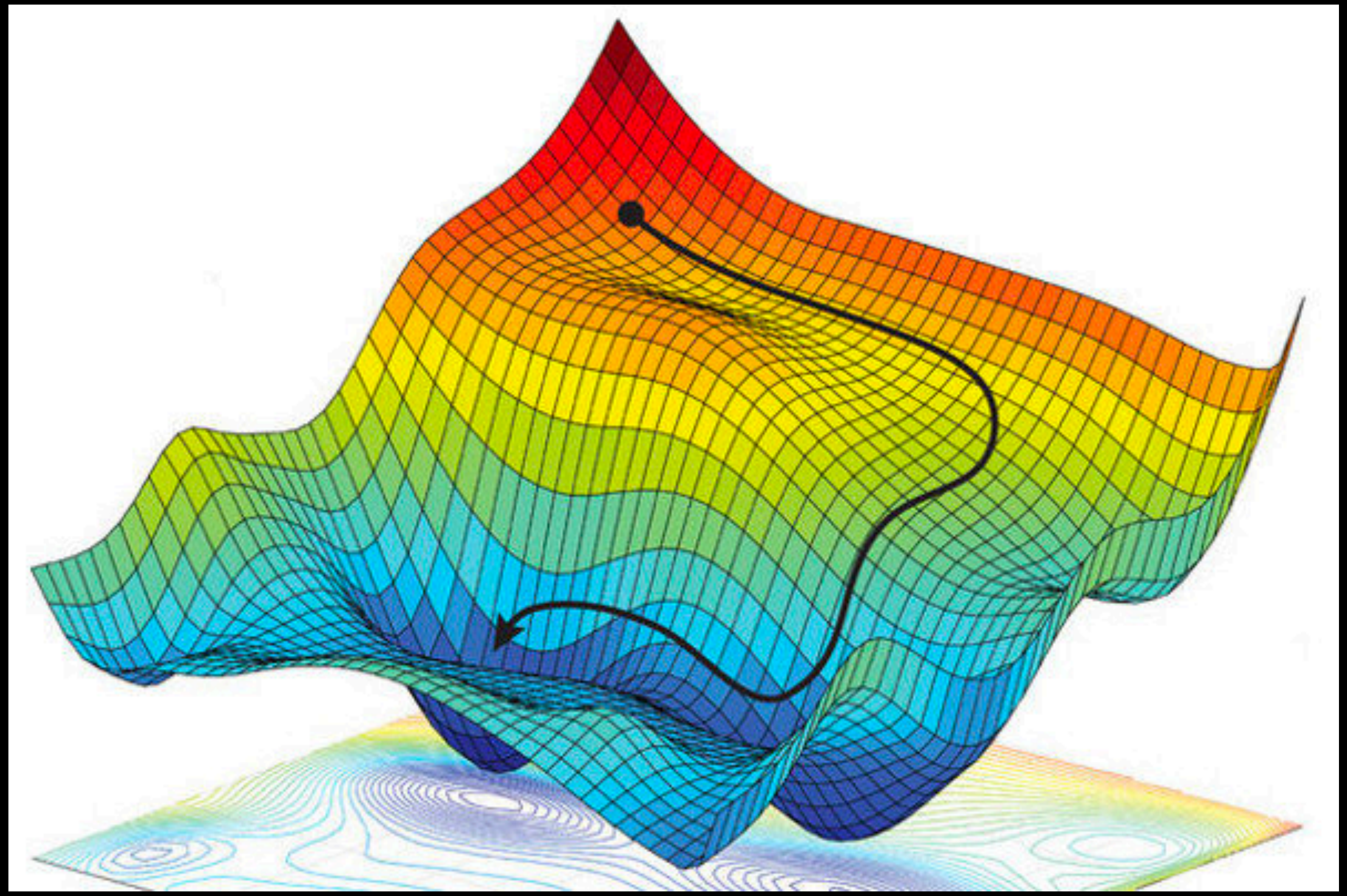
p(y) = output values



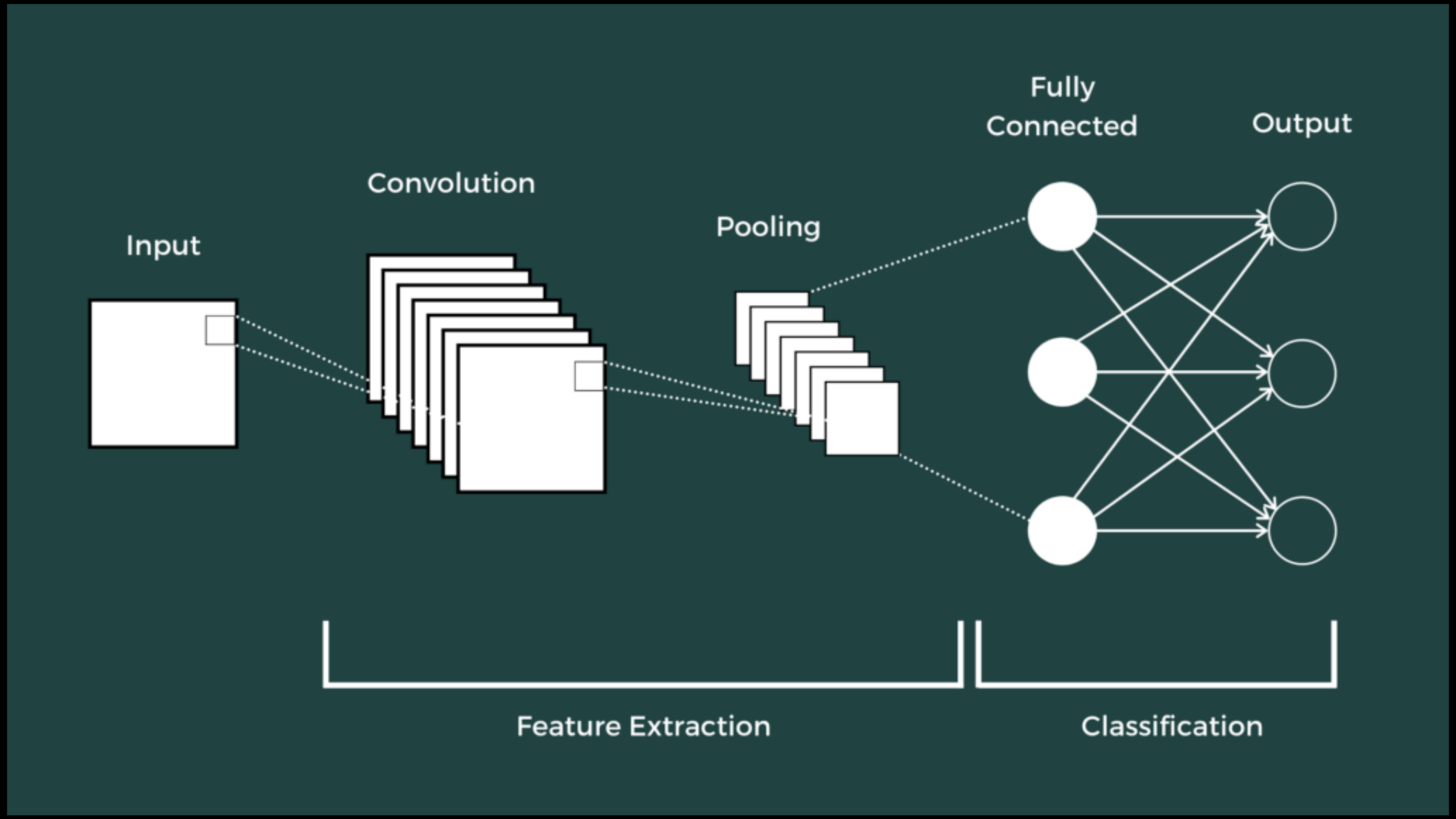
$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot log(p(y_i)) + (1 - y_i) \cdot log(1 - p(y_i))$$

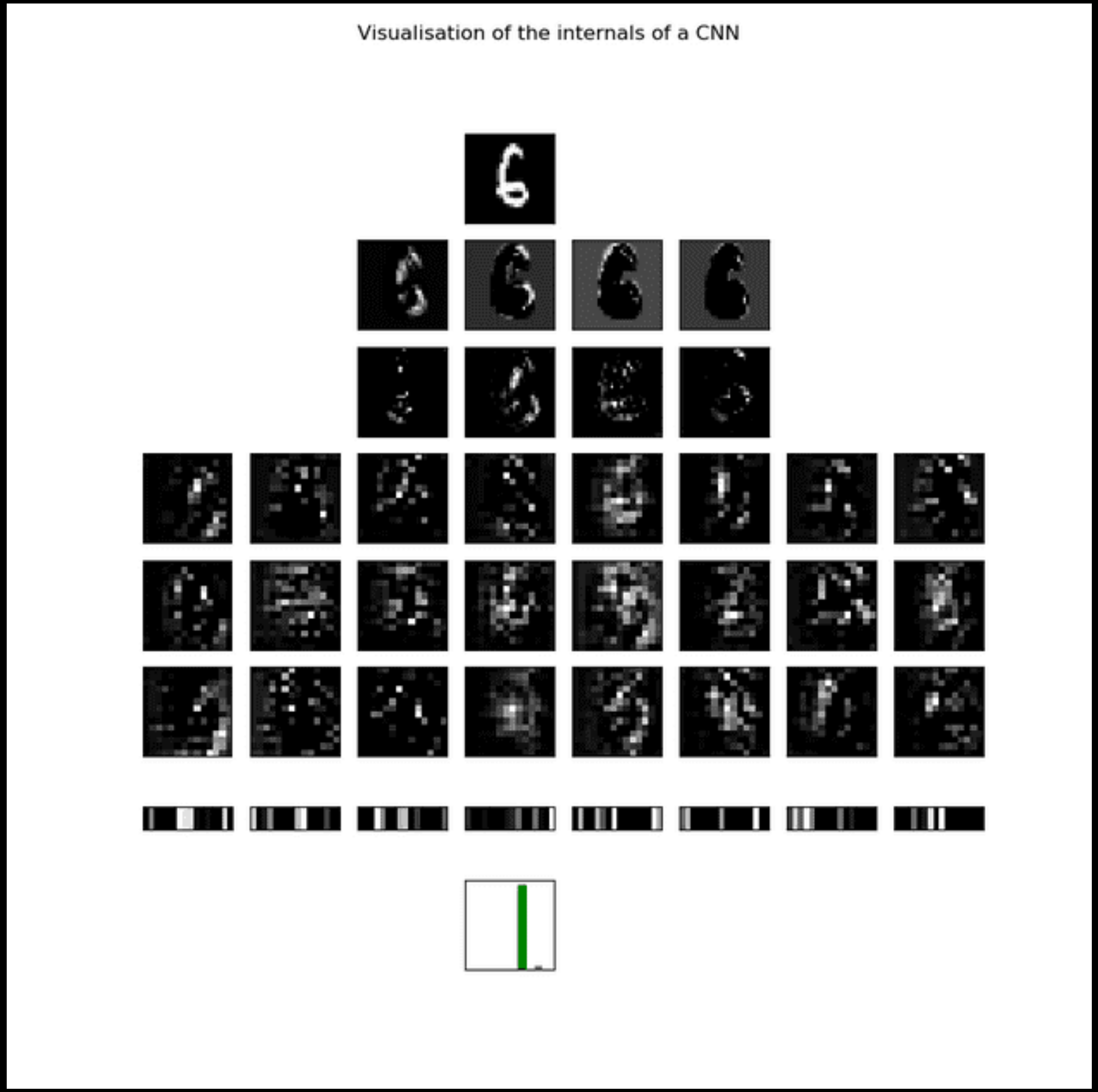
Binary Cross-Entropy / Log Loss





WHAT IS DEEP LEARNING?





WHY DEEP LEARNING?

Dieleman+15, Huertas-Company+15, Aniyan+17, Charnock+17, Gieseke+17, Jacobs+17, Petrillo+17, Schawinski+17, Alhassan+18, Dominguez-Sanchez+18, George+18, Hinners+18, Lukic+18, Moss+18, Razzano+18, Schaefer+18, Allen+19, Burke+19, Carrasco-Davis+19, Chatterjee+19, Davies+19, Dominguez-Sanchez+19, Fusse+19, Glaser+19, Ishida+19, Jacobs+19, Katebi+19, Lanusse+19Liu+19, Lukic+19, Metcalf+19, Muthukrishna+19, Petrillo+19, Reiman+19, Boucaud+20, Chiani+20, Ghosh+20, Gomez+20, Hausen+20, Hlozek+20, Hosseinzadeh+20. Huang+20, Li+20, Moller+20, Paillassa+20, Tadaki+20, Vargas dos Santos+20, Walmsley+20, Wei+20, Allam+21, Arcelin+21, Becker+21, Bretonniere+21, Burhanudin+21, Davison+21, Donoso-Oliva+21, Jia+21, Lauritsen+21, Ono+21, Ruan+21, Sadegho+21, Tang+21, Tanoglidis+21, Vojtekova+21, Dhar+22Hausen+22, Orwat+22, Pimentel+22, Rezaei+22, Samudre+22, Shen+22, Walmsley+22

WHY DEEP LEARNING?

Dieleman+15, Huertas-Company+15, Aniyan+17, Charnock+17, Gieseke+17, Jacobs+17, Petrillo+17, Schawinski+17, Alhassan+18, Dominguez-Sanchez+18, George+18, Hinners+18, Lukic+18, Moss+18, Razzano+18, Schaefer+18, Allen+19, Burke+19, Carrasco-Davis+19, Chatterjee+19, Davies+19, Dominguez-Sanchez+19, Fusse+19, Glaser+19, Ishida+19, Jacobs+19, Katebi+19, Lanusse+19Liu+19, Lukic+19, Metcalf+19, Muthukrishna+19, Petrillo+19, Reiman+19, Boucaud+20, Chiani+20, Ghosh+20, Gomez+20, Hausen+20, Hlozek+20, Hosseinzadeh+20. Huang+20, Li+20, Moller+20, Paillassa+20, Tadaki+20, Vargas dos Santos+20, Walmsley+20, Wei+20, Allam+21, Arcelin+21, Becker+21, Bretonniere+21, Burhanudin+21, Davison+21, Donoso-Oliva+21, Jia+21, Lauritsen+21, Ono+21, Ruan+21, Sadegho+21, Tang+21, Tanoglidis+21, Vojtekova+21, Dhar+22Hausen+22, Orwat+22, Pimentel+22, Rezaei+22, Samudre+22, Shen+22, Walmsley+22

OUTLINE

- Why care about galaxy morphology?
- Why Deep Learning?
- Supervised learning: CNNs
- Transfer Learning
- Going fainter
- Going deeper

CATALOGUES & DATA USED FOR TRAINING

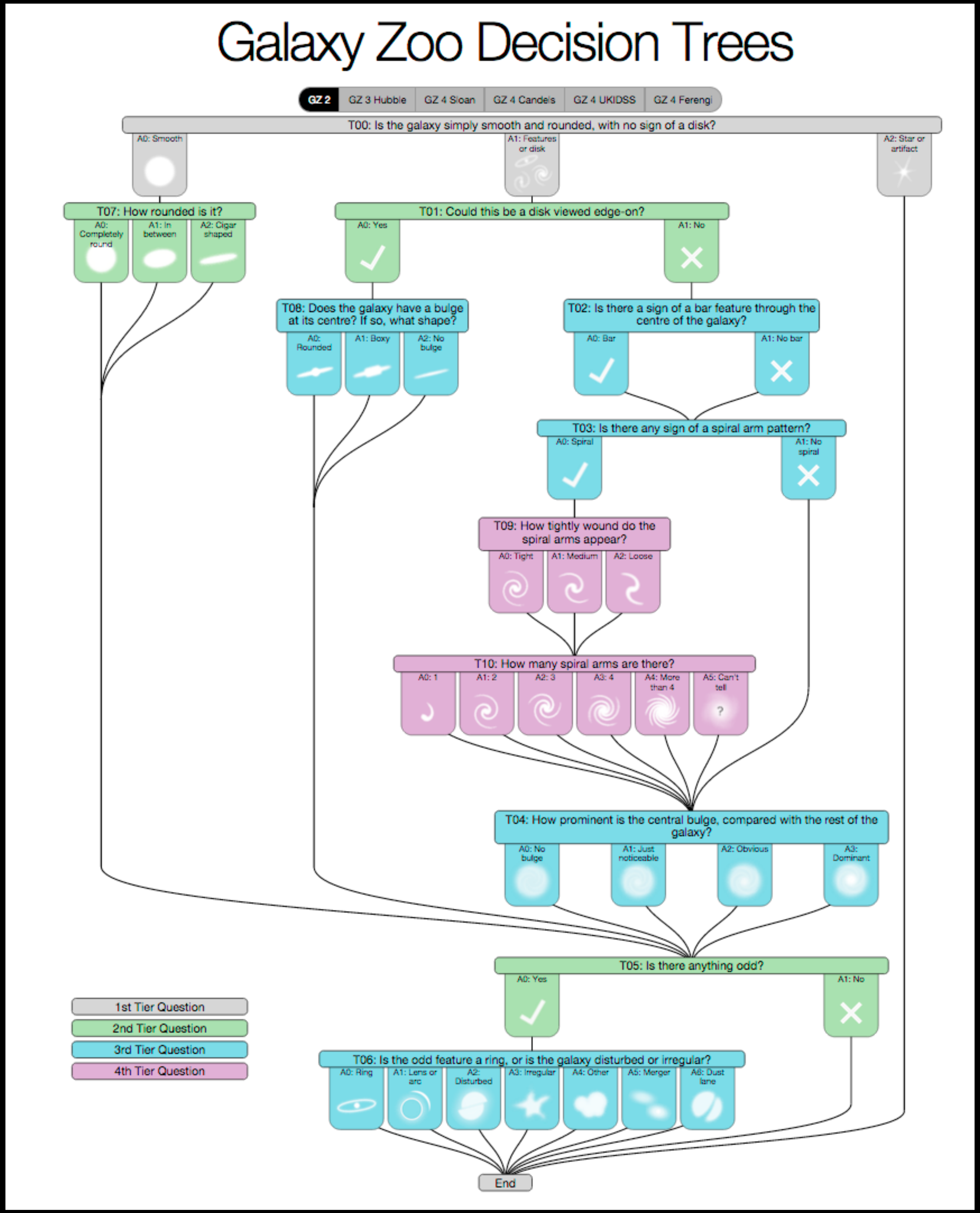
- **Sloan Digital Sky Survey**
  - 1/3 of the Skye, 0.5 billion objects
  - 2.5m telescope (Apache Point Observatory)
- **Galaxy Zoo** (Willet+2013 catalogue)
  - 240,000 galaxies
  - Science citizen project
  - Decision tree
- **Nair & Abraham** (N&A+2010 catalogue)
  - 14,000 galaxies
  - professional astronomers
  - Detailed classification (T-Type)

CATALOGUES & DATA USED FOR TRAINING

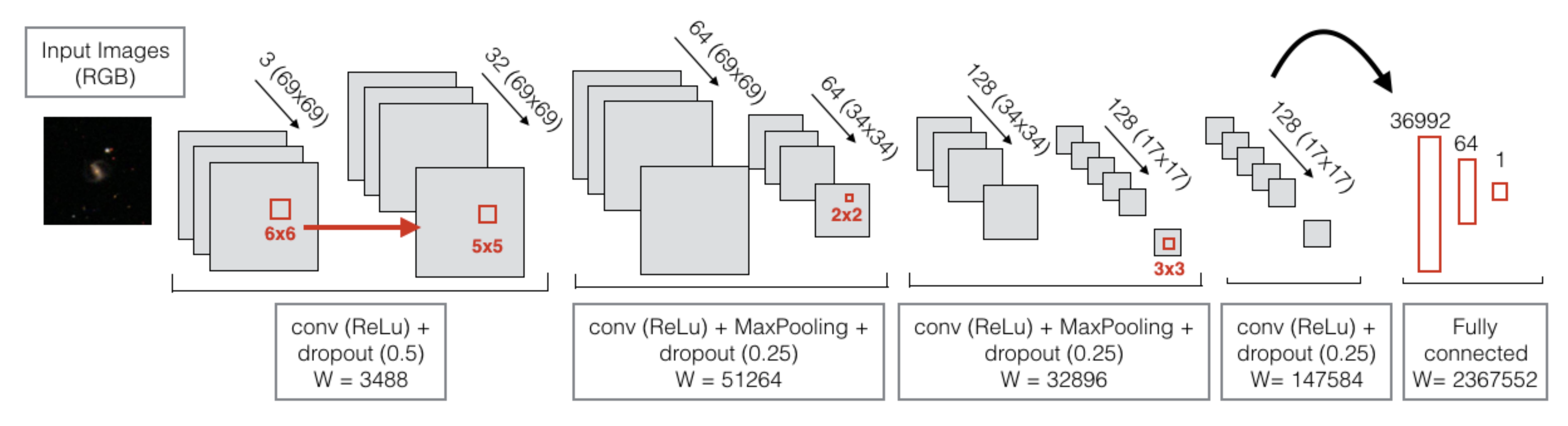
- **Sloan Digital Sky Survey**
  - 1/3 of the Skye, 0.5 billion objects
  - 2.5m telescope (Apache Point Observatory)

- **Galaxy Zoo** (Willet+2013 catalogue)
  - 240,000 galaxies
  - Science citizen project
  - Decision tree

- **Nair & Abraham** (N&A+2010 catalogue)
  - 14,000 galaxies
  - professional astronomers
  - Detailed classification (T-Type)



(VANILLA) CONVOLUTIONAL NEURAL NETWORK



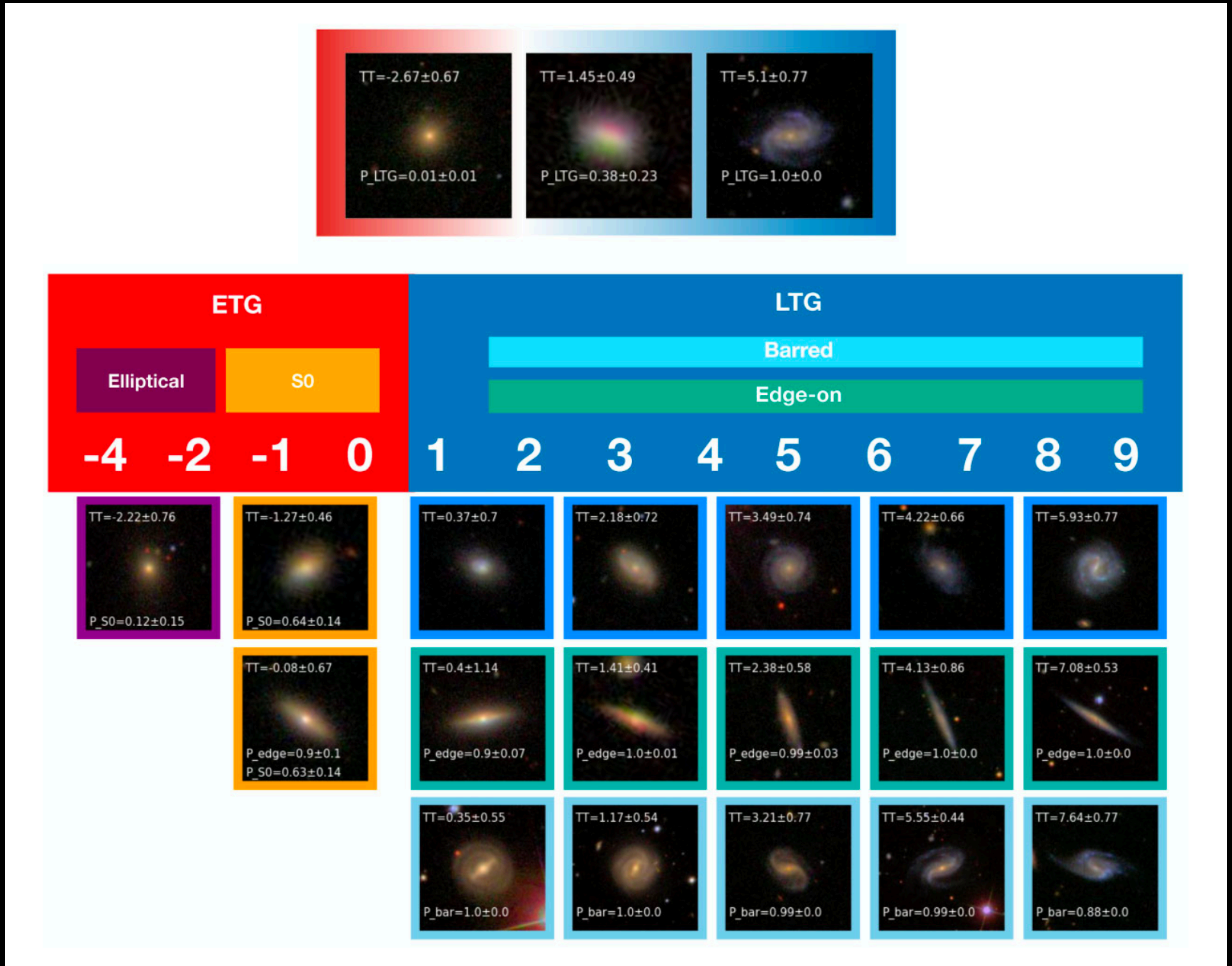
Input= flux per pixel in 3 bands

Cutout size proportional to galaxy size Re

Domínguez Sánchez + 2018

(Following Dieleman+2015, Huertas-Company+2015)

CLASSIFICATION SCHEME

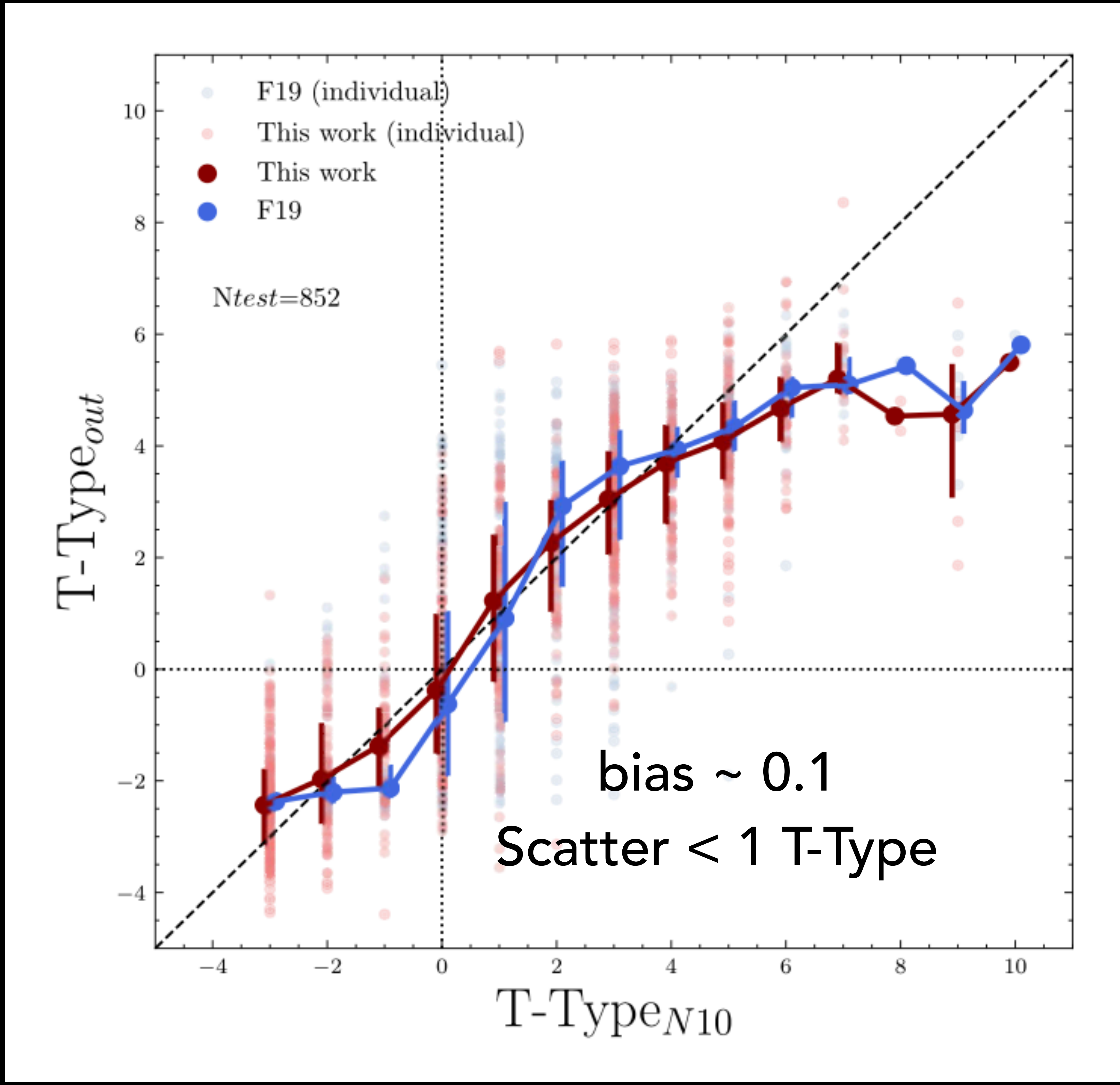


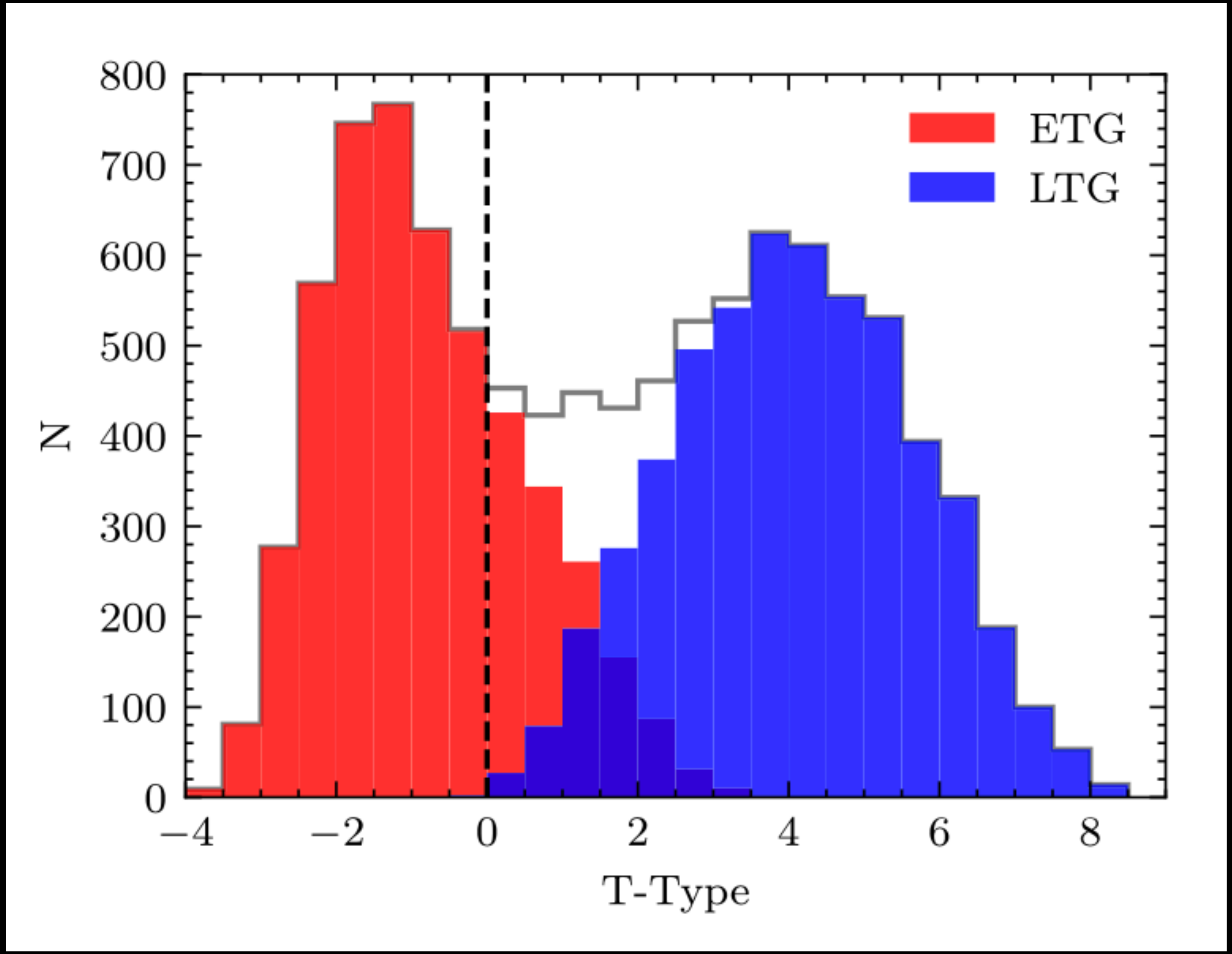
Improvements:

- T-Type better determined
- New ETG/LTG model
- Uncertainty estimation
- Visual classification

Domínguez Sánchez+2022

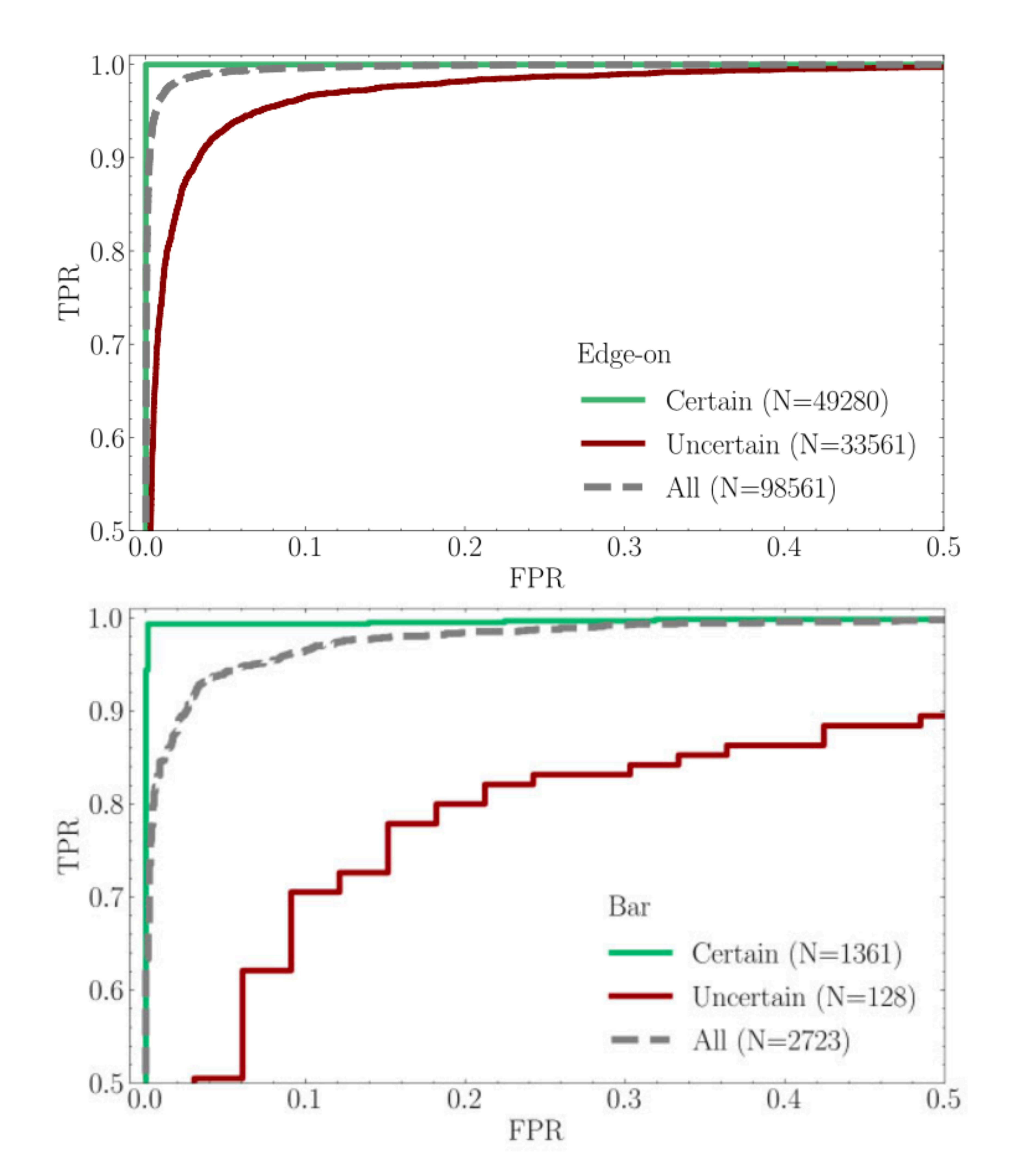
T-TYPE VS ETG/LTG MODELS



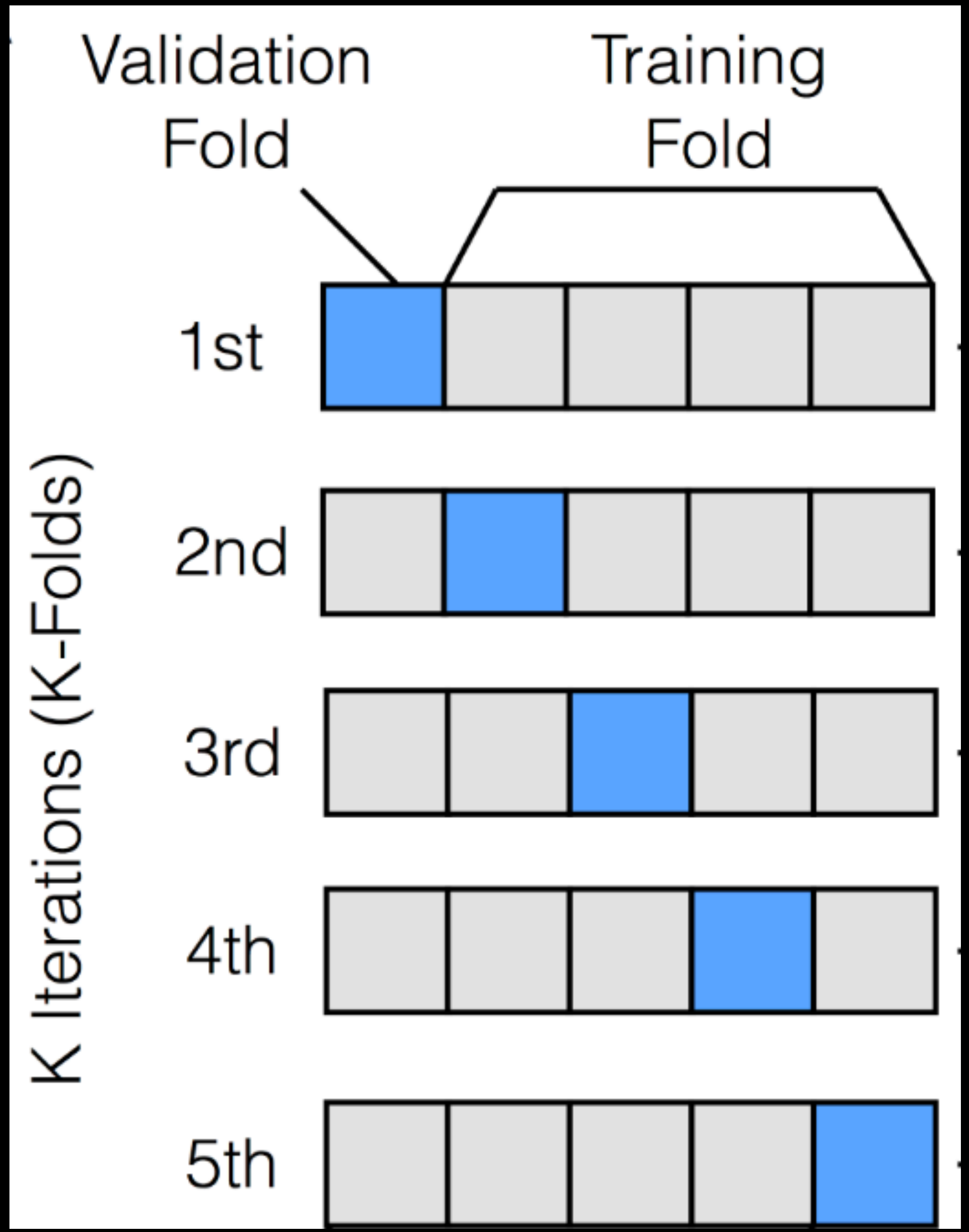


Domínguez Sánchez+2022

BINARY MODELS PERFORMANCE & UNCERTAINTIES



Uncertainties are derived as the standard deviation of the output of 10 k-folded models

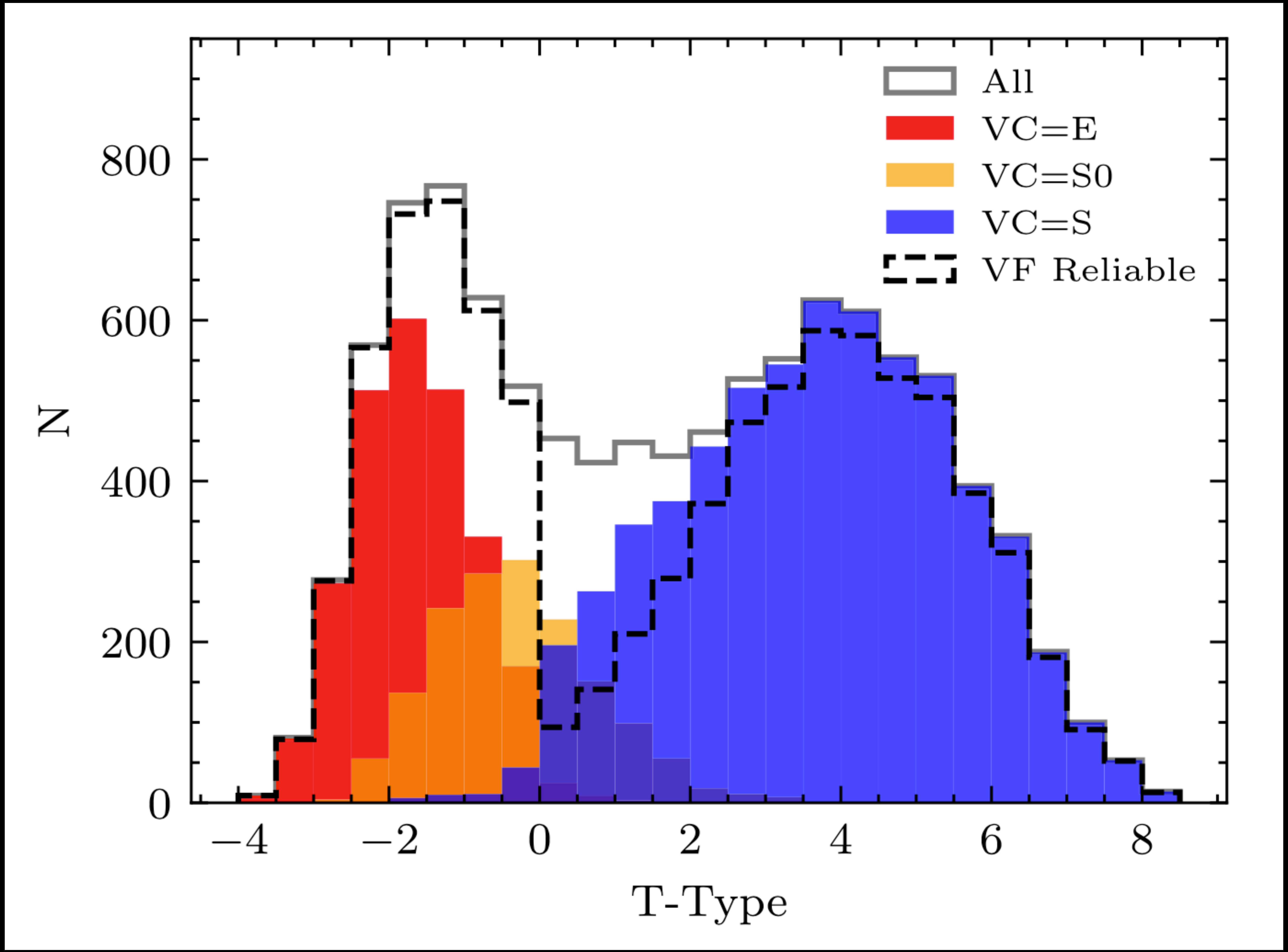


Model	N_{test}	% Positives	Accuracy	Precision	Recall	F1
$P_{\text{edge-on}}$	98 561	14	0.98	0.87	0.98	0.93
P_{bar}	2723	50	0.93	0.92	0.90	0.93

Domínguez Sánchez+2022

T-TYPE VS ETG/LTG MODELS

VISUAL CLASSIFICATION

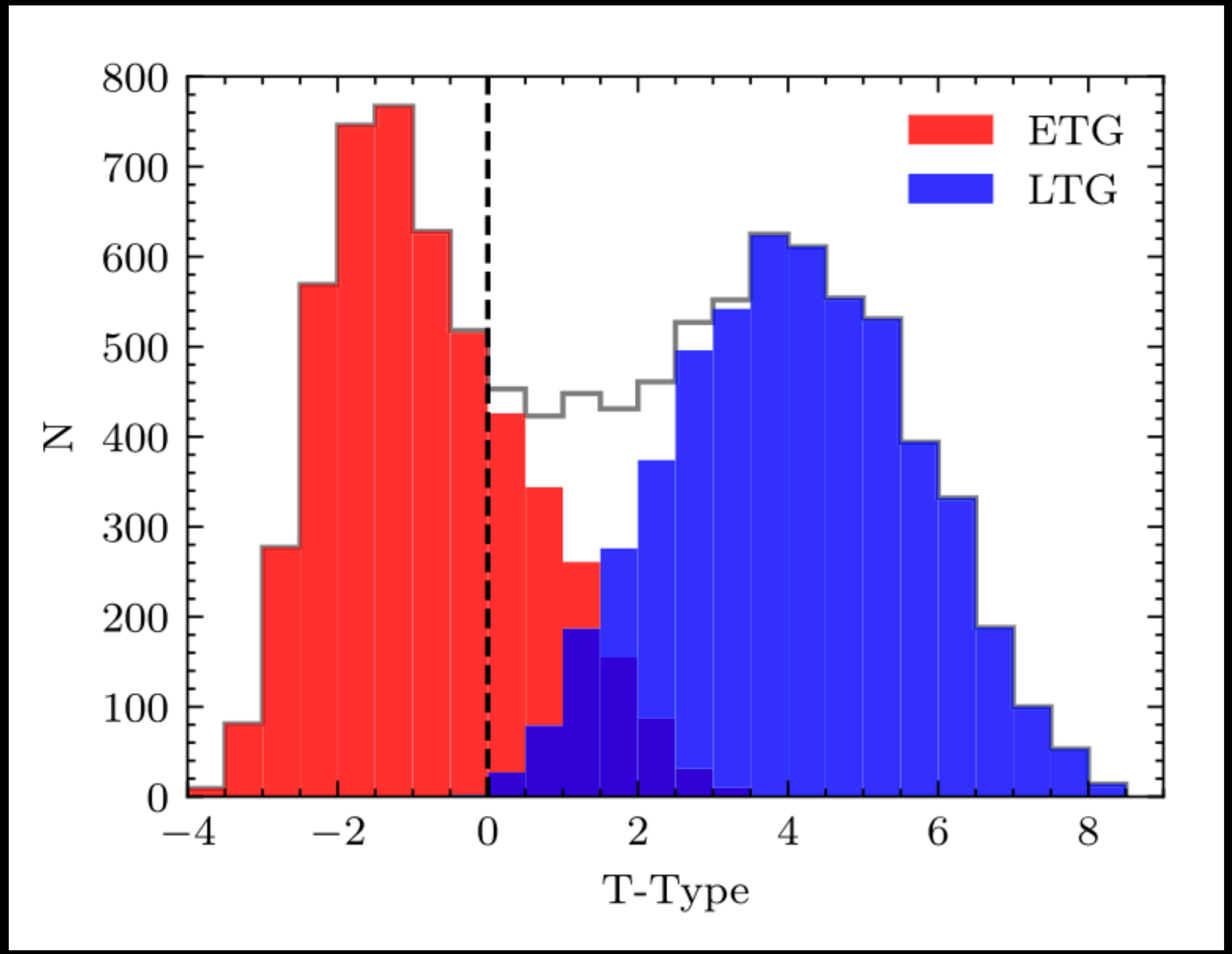


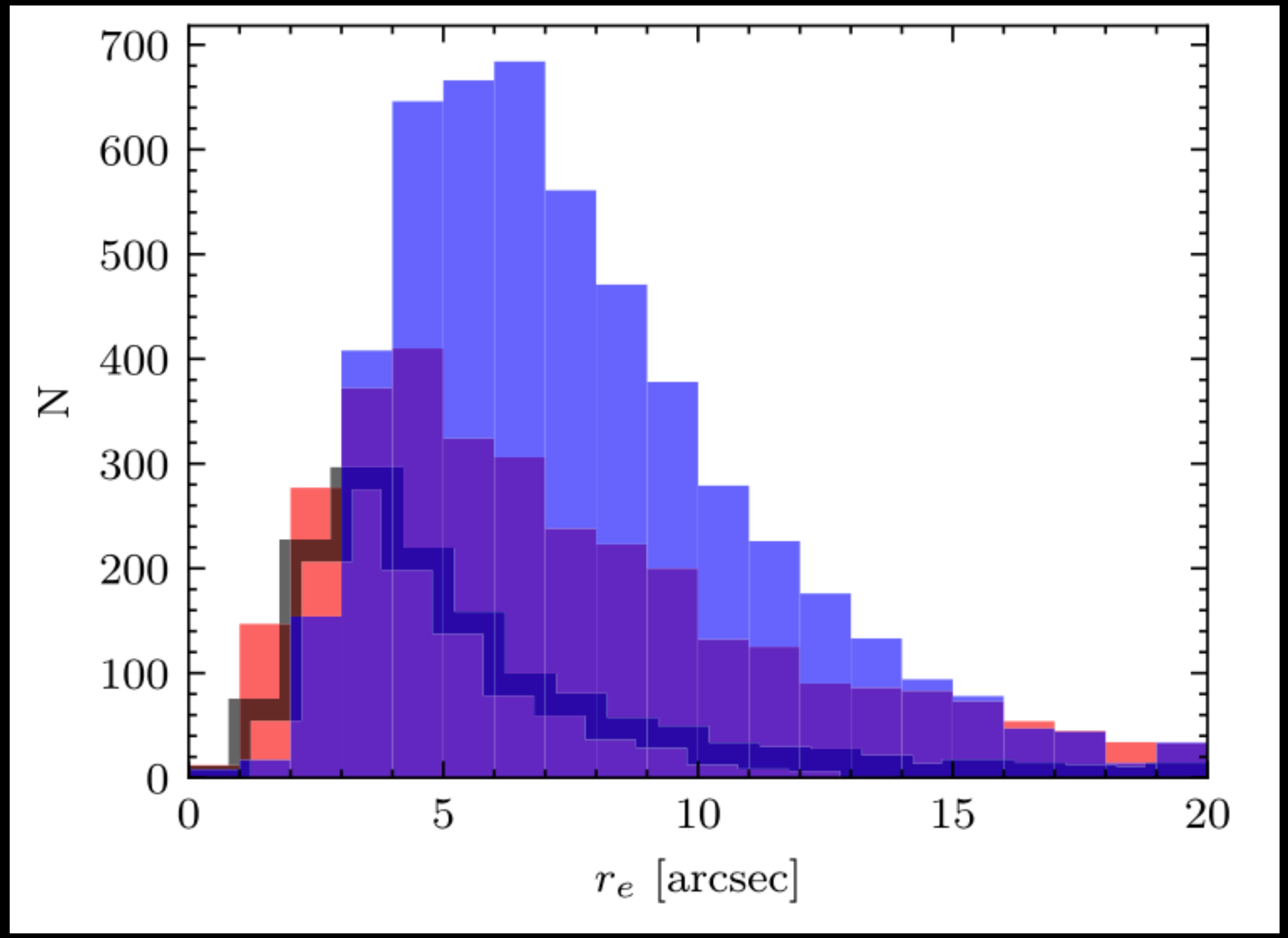
Domínguez Sánchez+2022

T-TYPE VS ETG/LTG MODELS

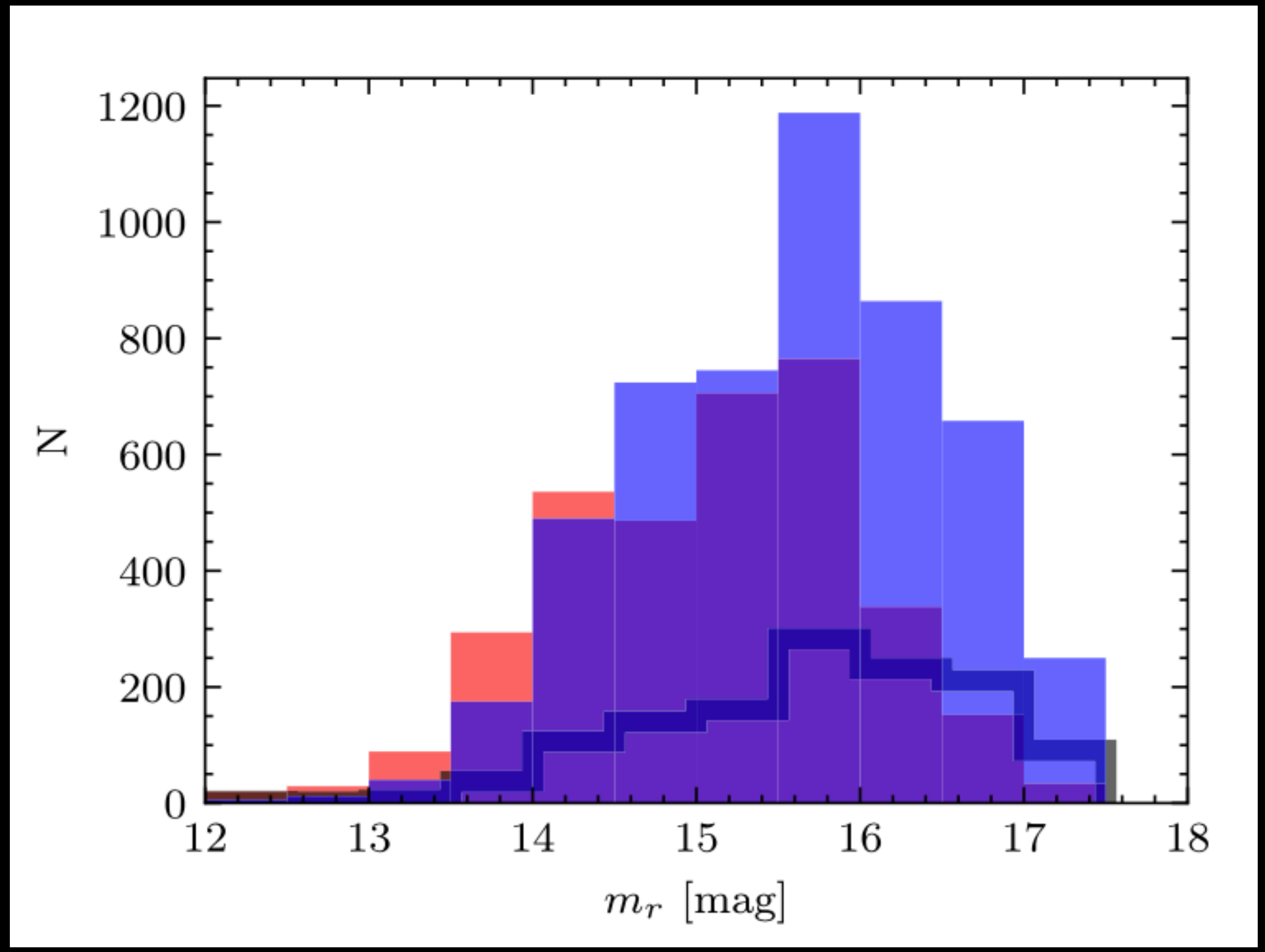
Galaxies classified as ETG with T-Type > 0 tend to be small, faint and with low central velocity dispersions.

Consistent with disk-like galaxies with no clear spiral structure.

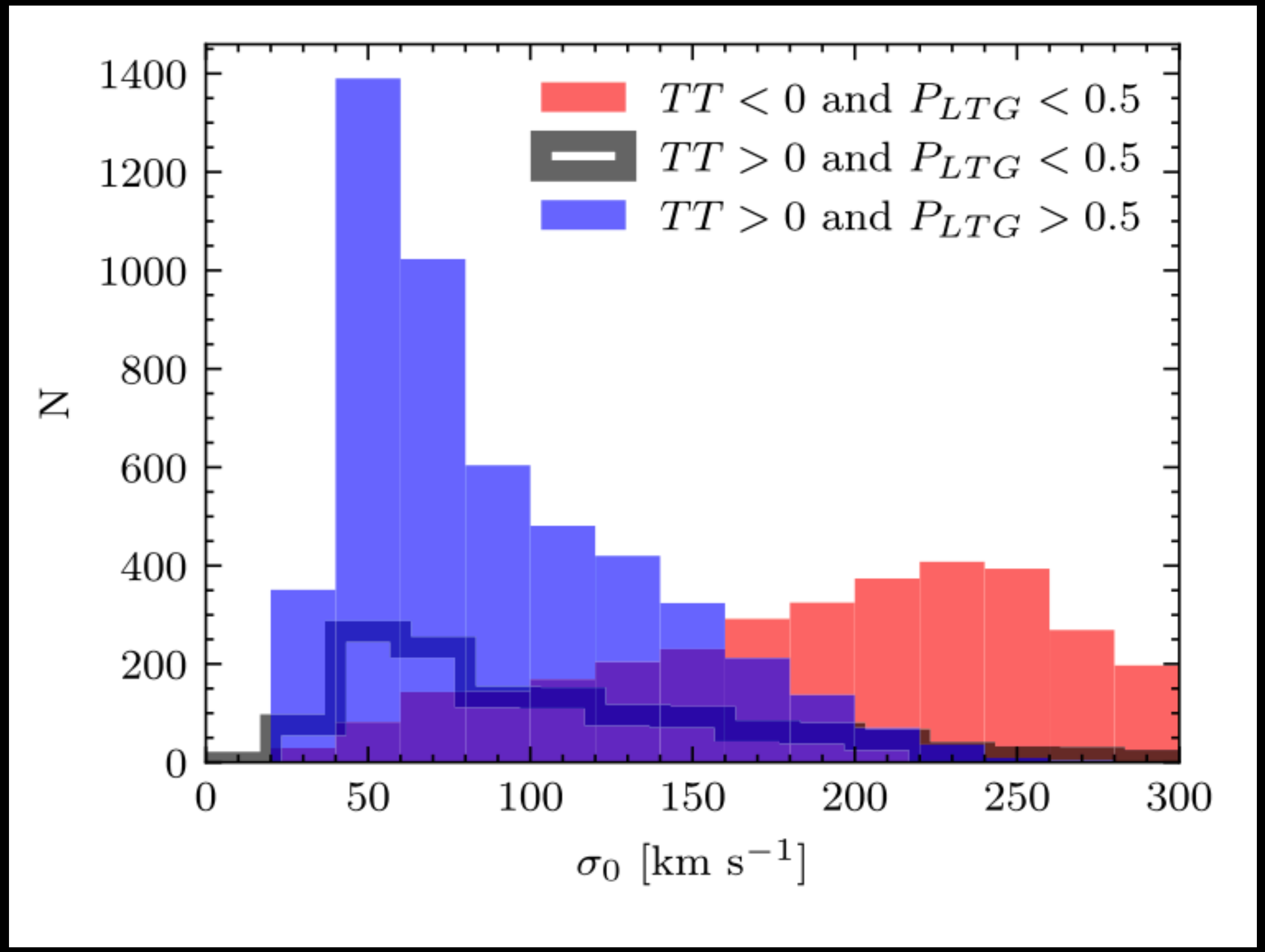




SIZE

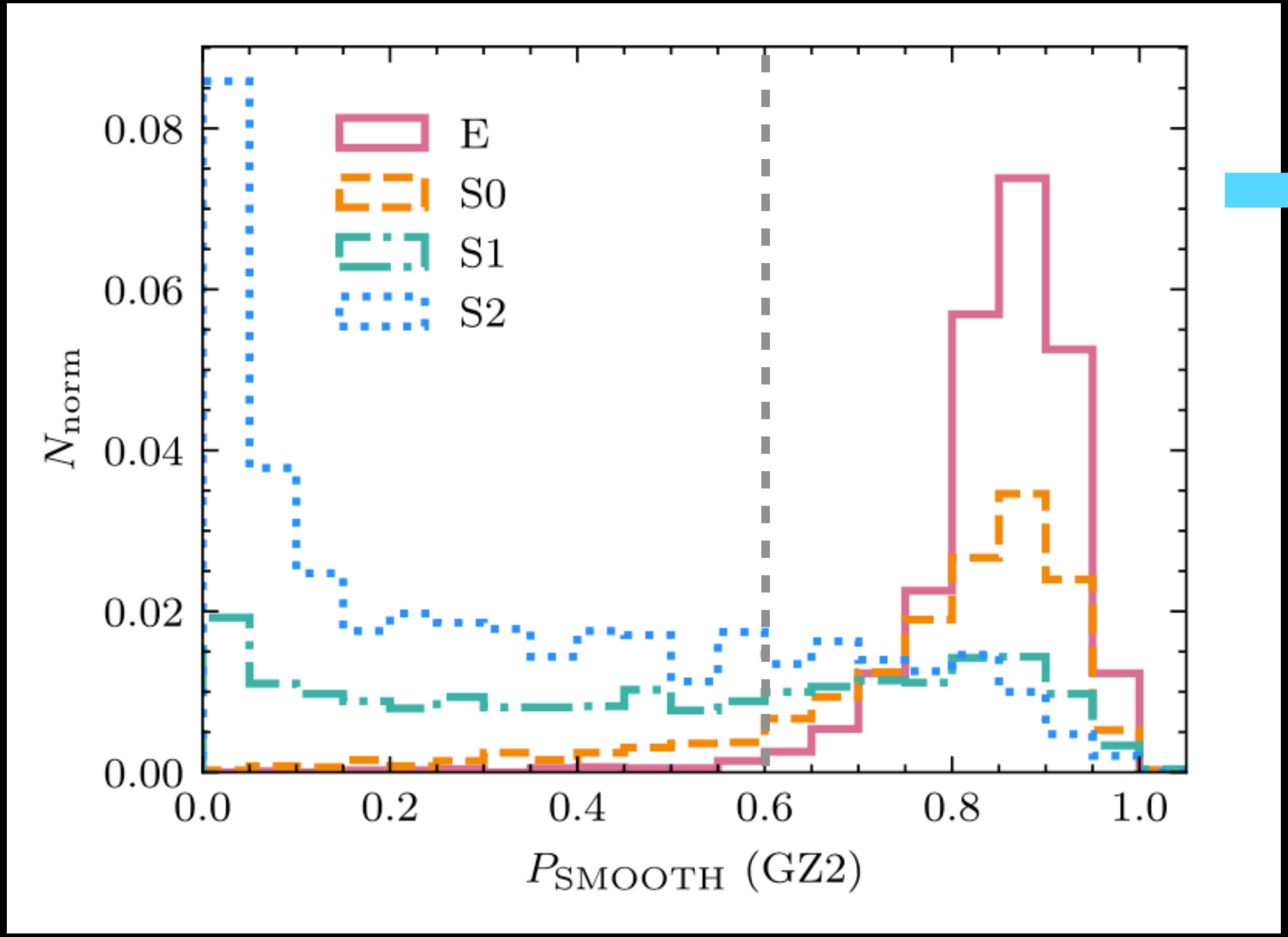


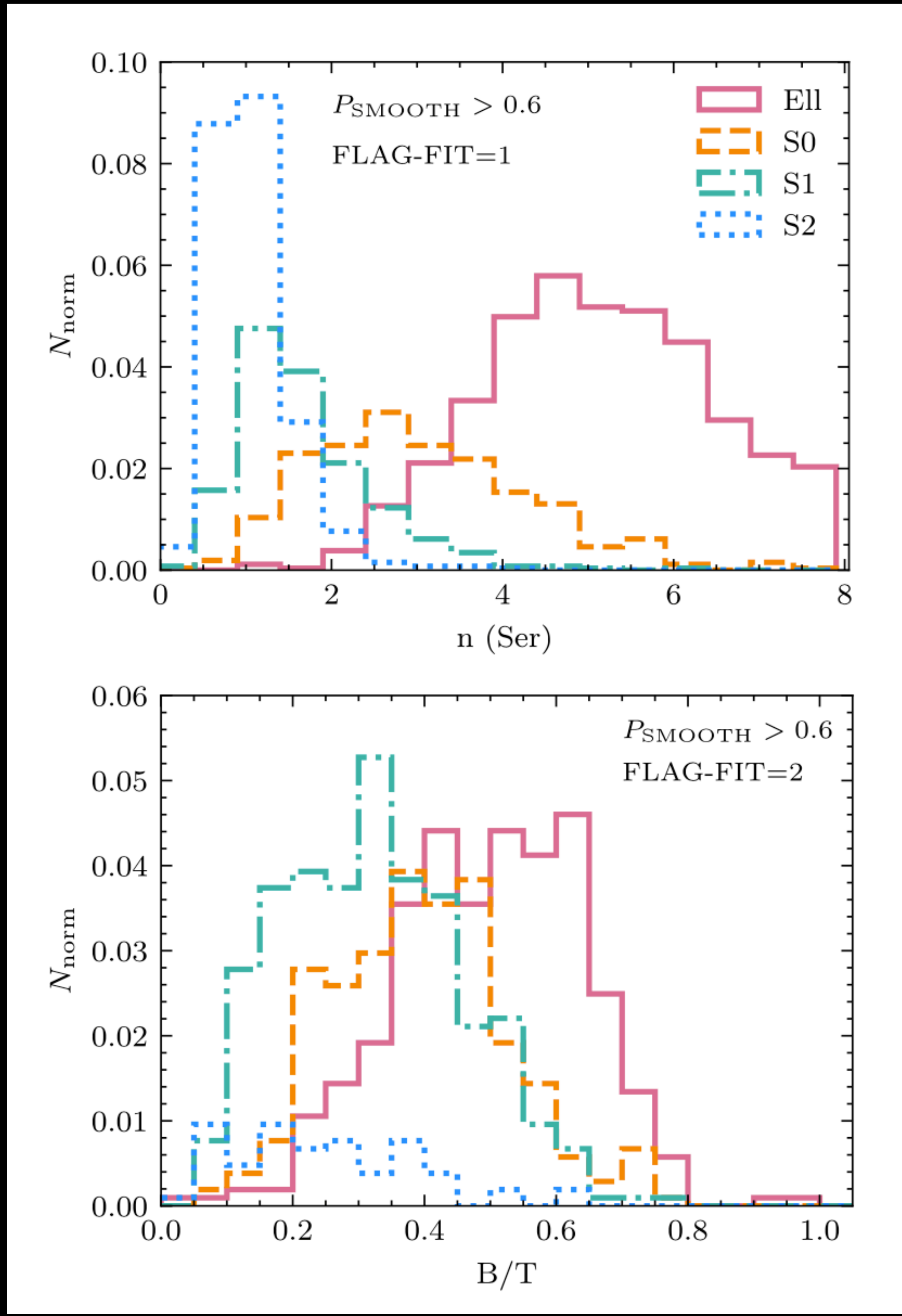
BRIGHTNESS



ROTATION

DS+22 VS GALAXY ZOO CLASSIFICATION





- A significant fraction of galaxies with P_smooth > 0.6 are spirals (according to DS22)
- They have low Sèrsic indices and B/T values
- Being 'smooth' is not equivalent to being elliptical!

Domínguez Sánchez+2022, Fischer+2019

PUBLIC CATALOGUES

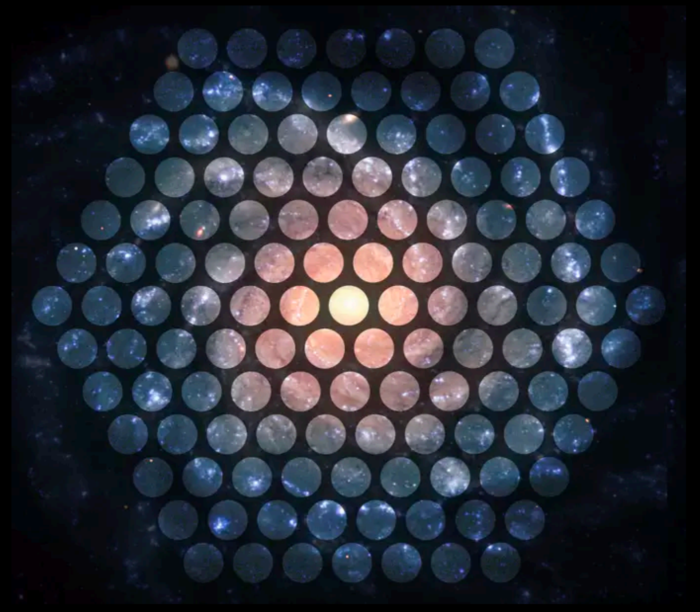
- **SDSS:** 600,000 galaxies; DS+18, DS+23

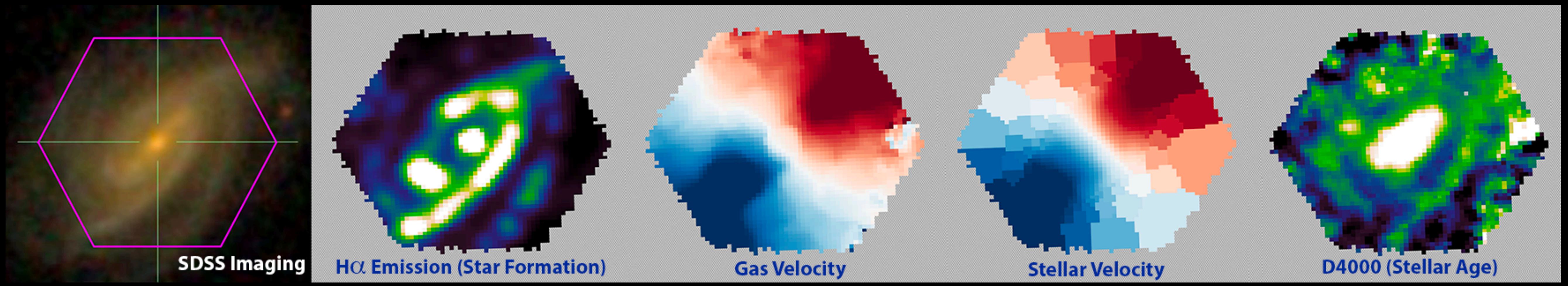
  https://archive.cefca.es/ancillary_data/sdss_morphological_catalogues/sdss_morphological_catalogues.tar.gz

- **MaNGA:** 10,000 galaxies; Fischer+19, DS+22
  - *MaNGA Morphology Deep Learning DR17 catalog*

    https://www.sdss4.org/dr17/data_access/value-added-catalogs/?vac_id=manga-morphology-deep-learning-dr17-catalog
  - *MaNGA Pymorph DR17 photometric catalog*

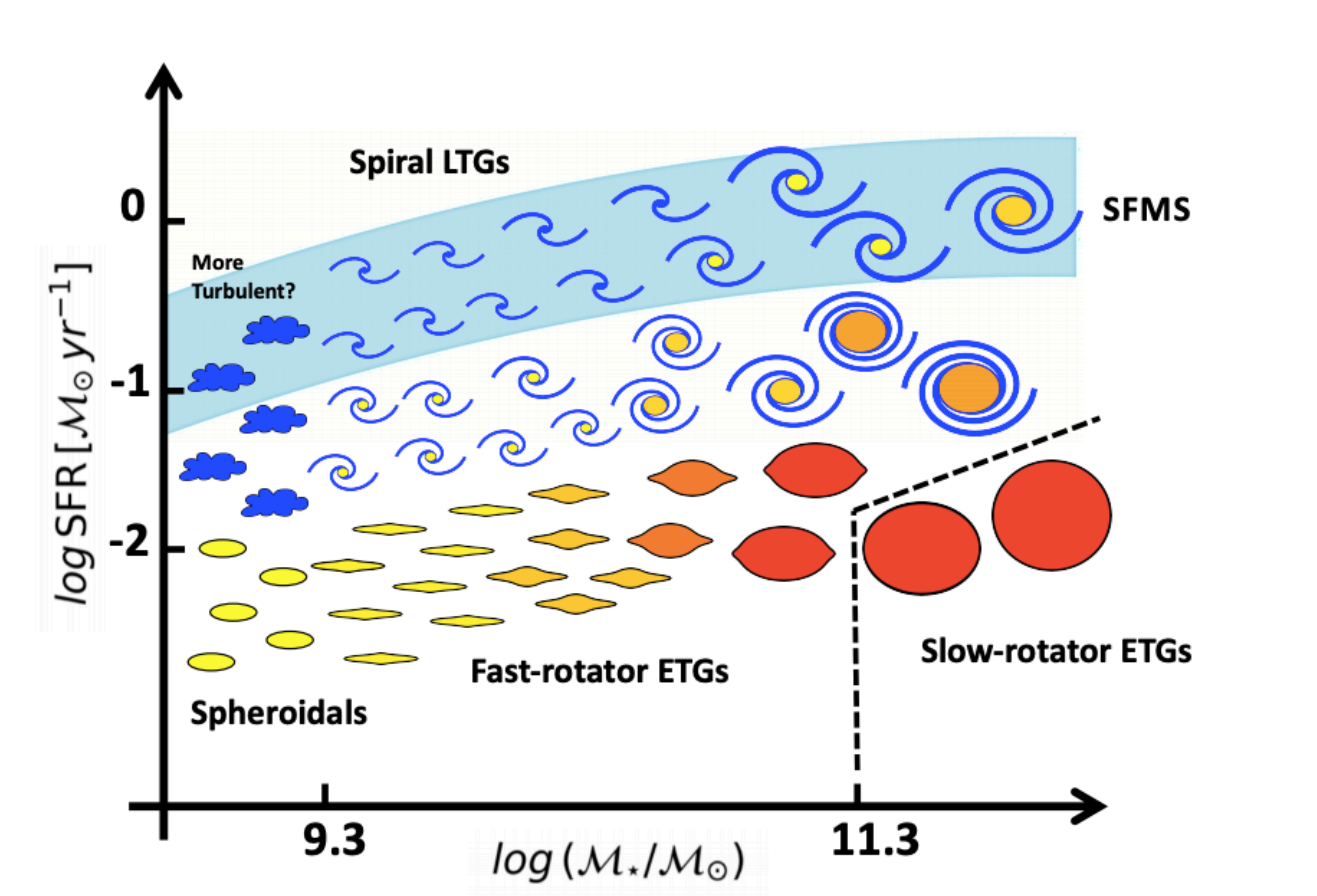
    https://www.sdss4.org/dr17/data_access/value-added-catalogs/?vac_id=manga-pymorph-dr17-photometric-catalog





Morphology related to fundamental galaxy properties

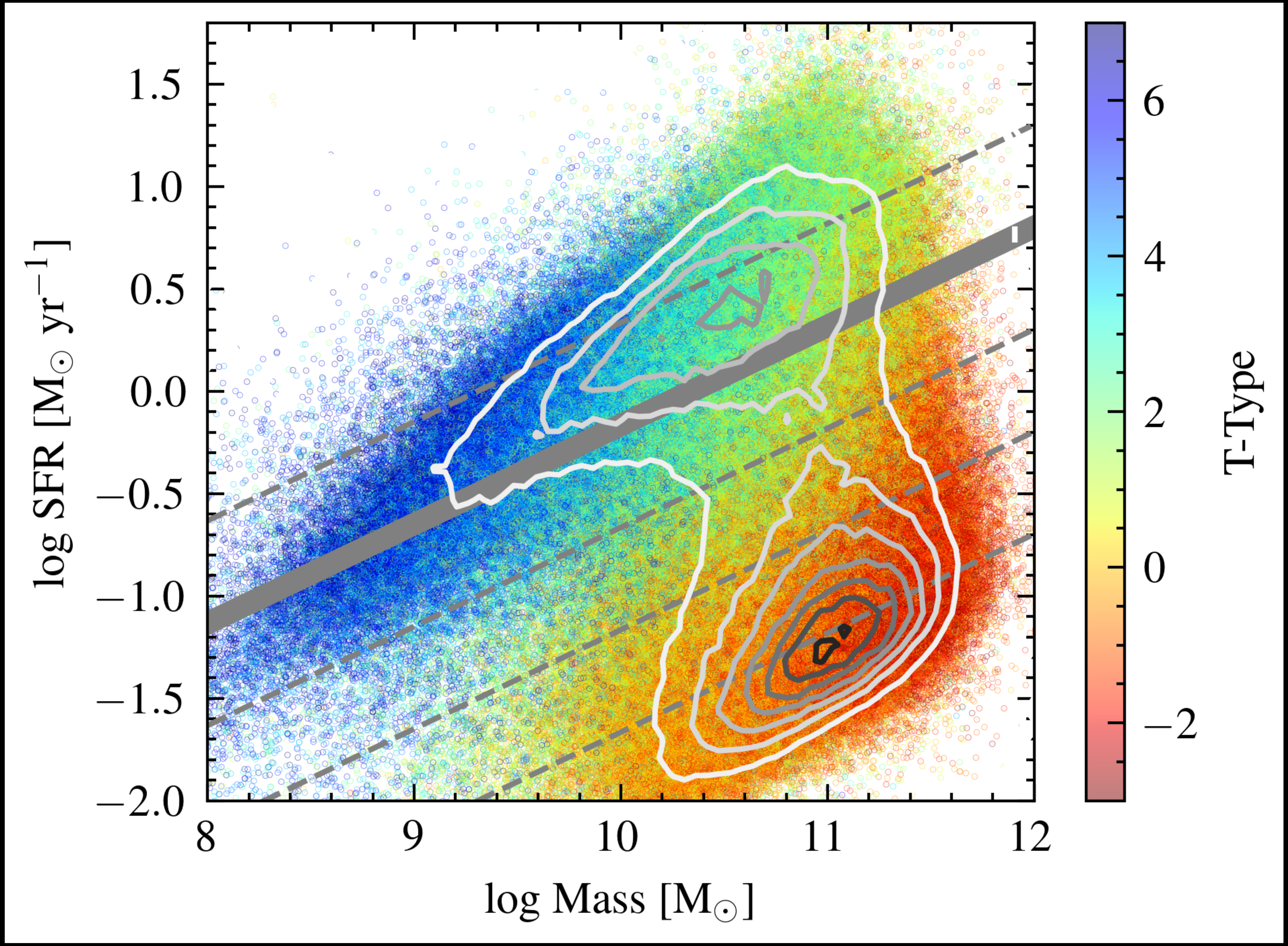
Number of stars formed per year



Galaxy Mass (in units of solar masses)

Morphology related to fundamental galaxy properties

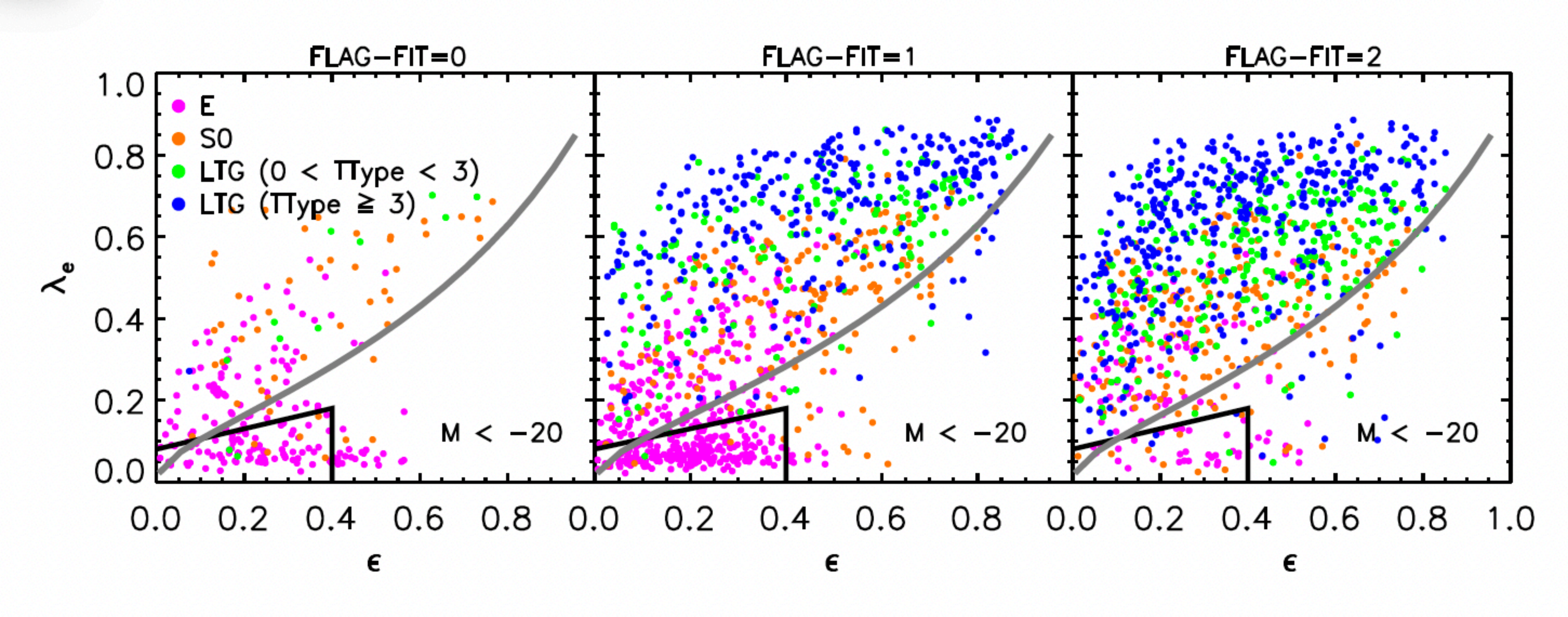
Number of stars formed per year



Galaxy Mass (in units of solar masses)

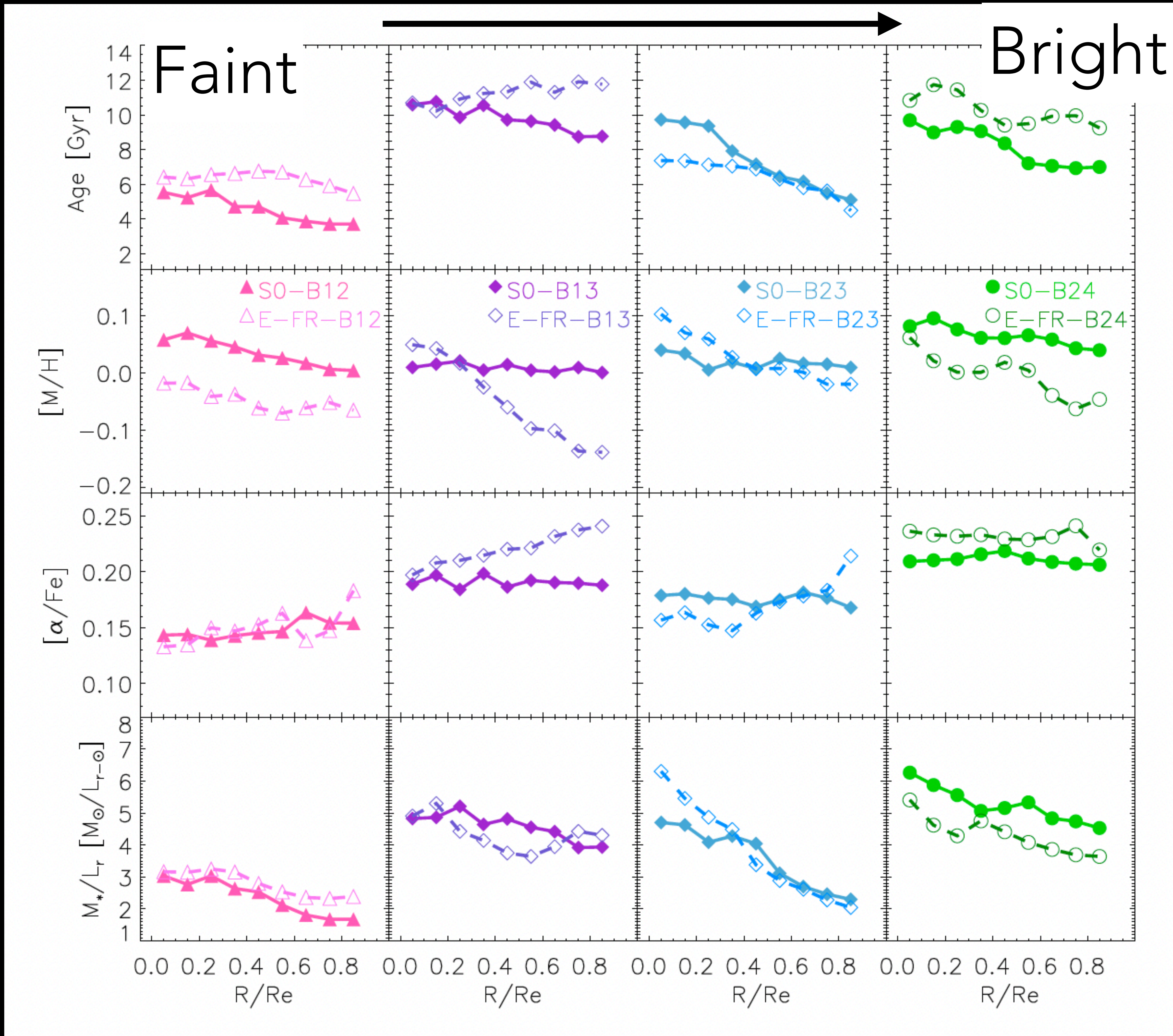
Domínguez Sánchez+2023

Morphology related to fundamental galaxy properties



Fischer, Domínguez Sánchez+2019

Morphology related to fundamental galaxy properties



We were able to study the stellar population gradients of E and S0 separated according to their kinematics.

E-FR and S0 have different properties (not the same objects)

Domínguez Sánchez+2019b
Bernardi, DS+2019
Domínguez Sánchez+2020
Bernardi, DS+2023

OUTLINE

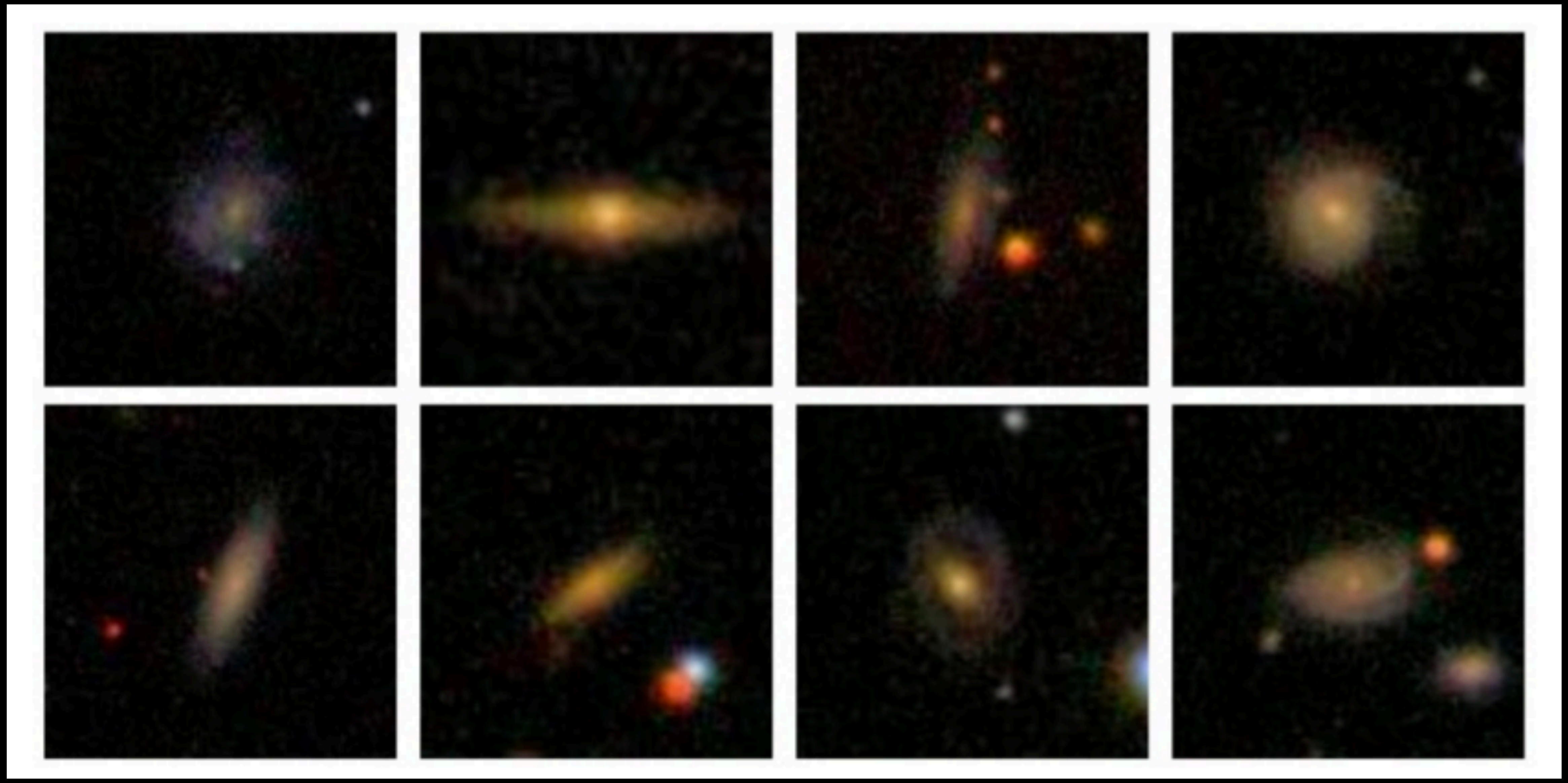
- Why care about galaxy morphology?
- Why Deep Learning?
- Supervised learning: CNNs
- What if there are no labels?
  - Transfer Learning
  - 'Emulations'
  - Unsupervised learning
- Going deeper

TRANSFER LEARNING

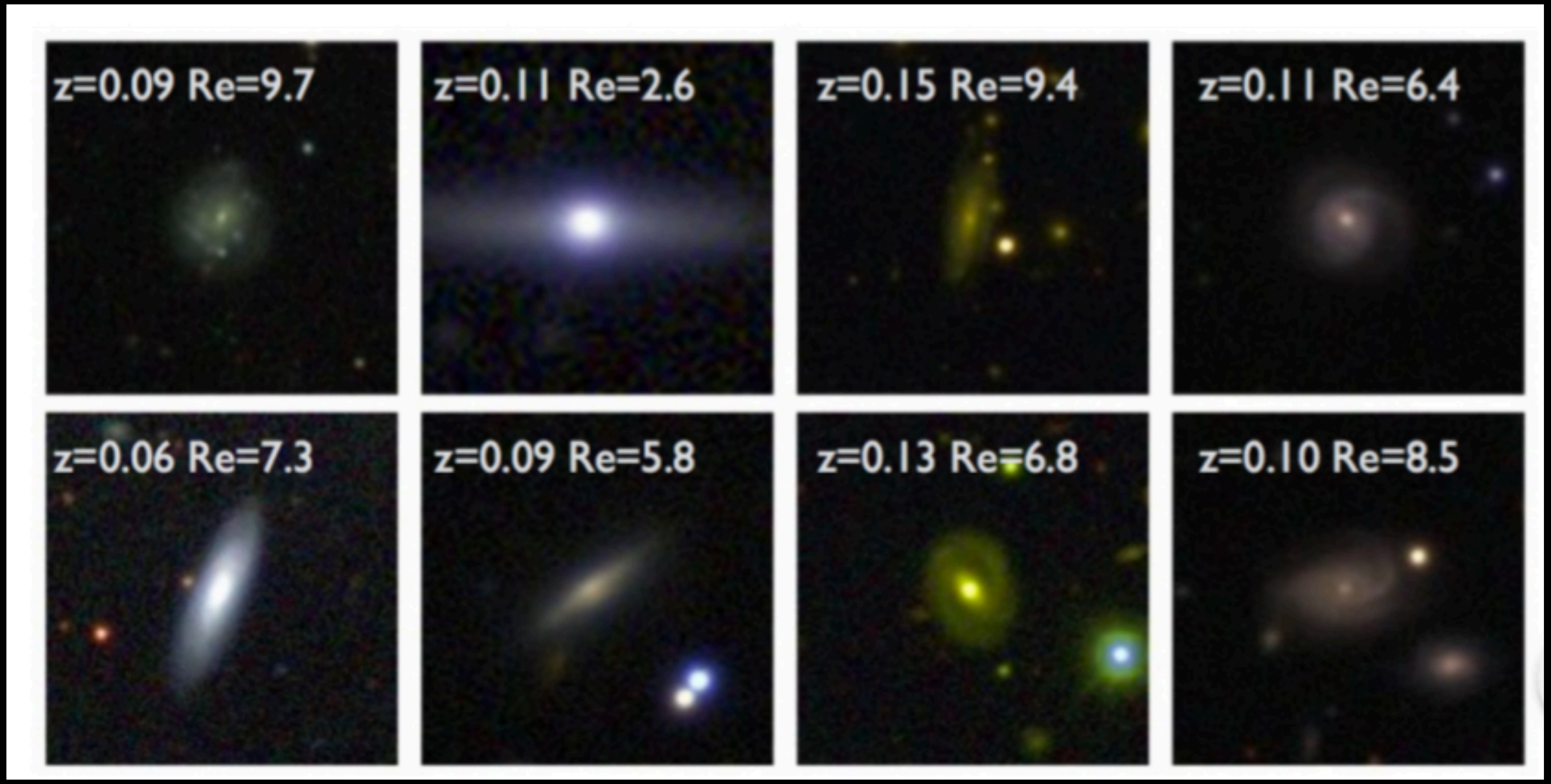
Supervised deep learning algorithms rely on large training samples (5000-10000 labeled galaxies). A key question in view of using them in future Big Data surveys such as LSST or Euclid is…

How much of the knowledge acquired from an existing survey can be exported to a new dataset?

SDSS



DES

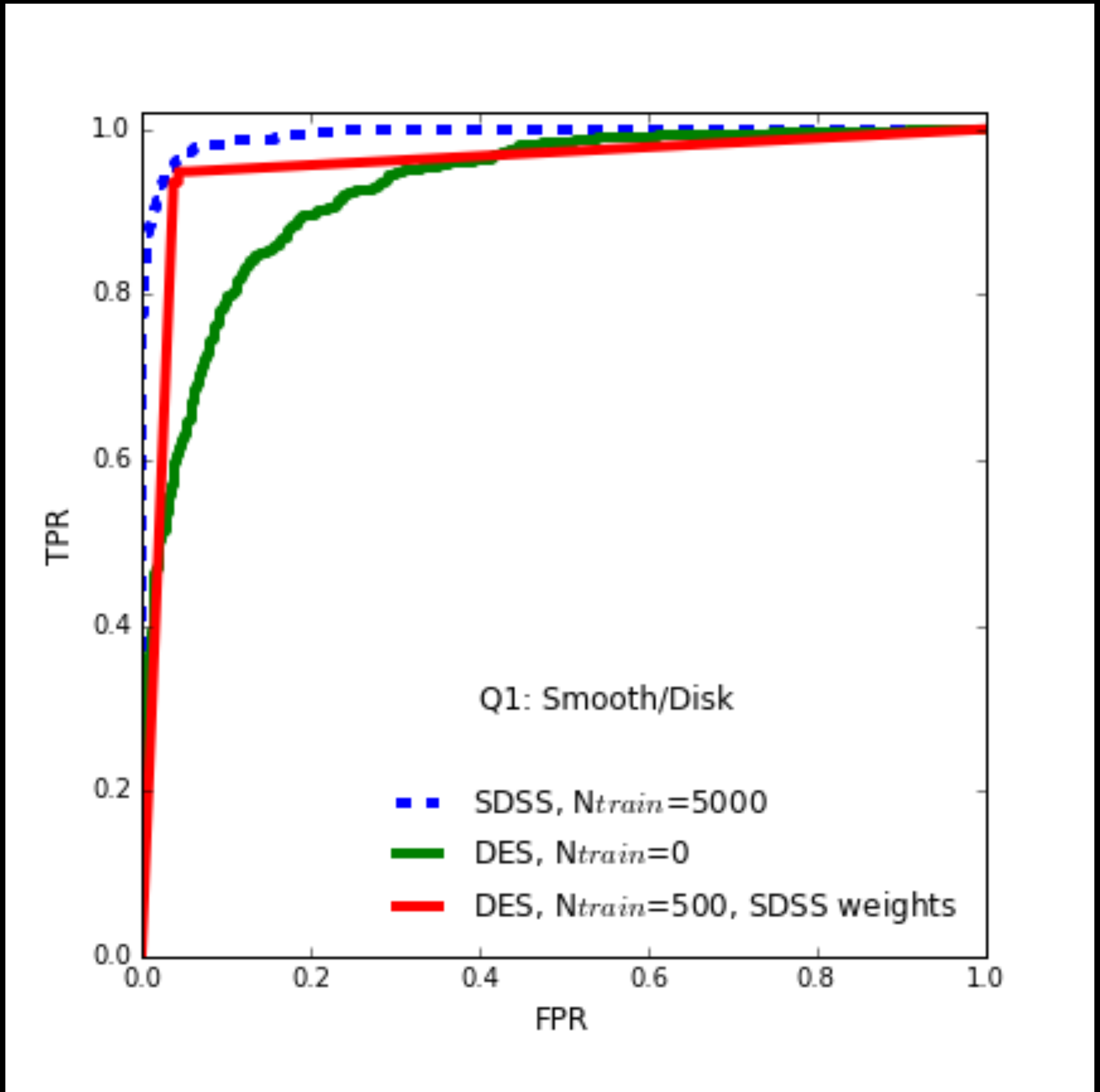


TRANSFER LEARNING

(1) Apply SDSS-based models directly to Dark Energy Survey (DES)images.

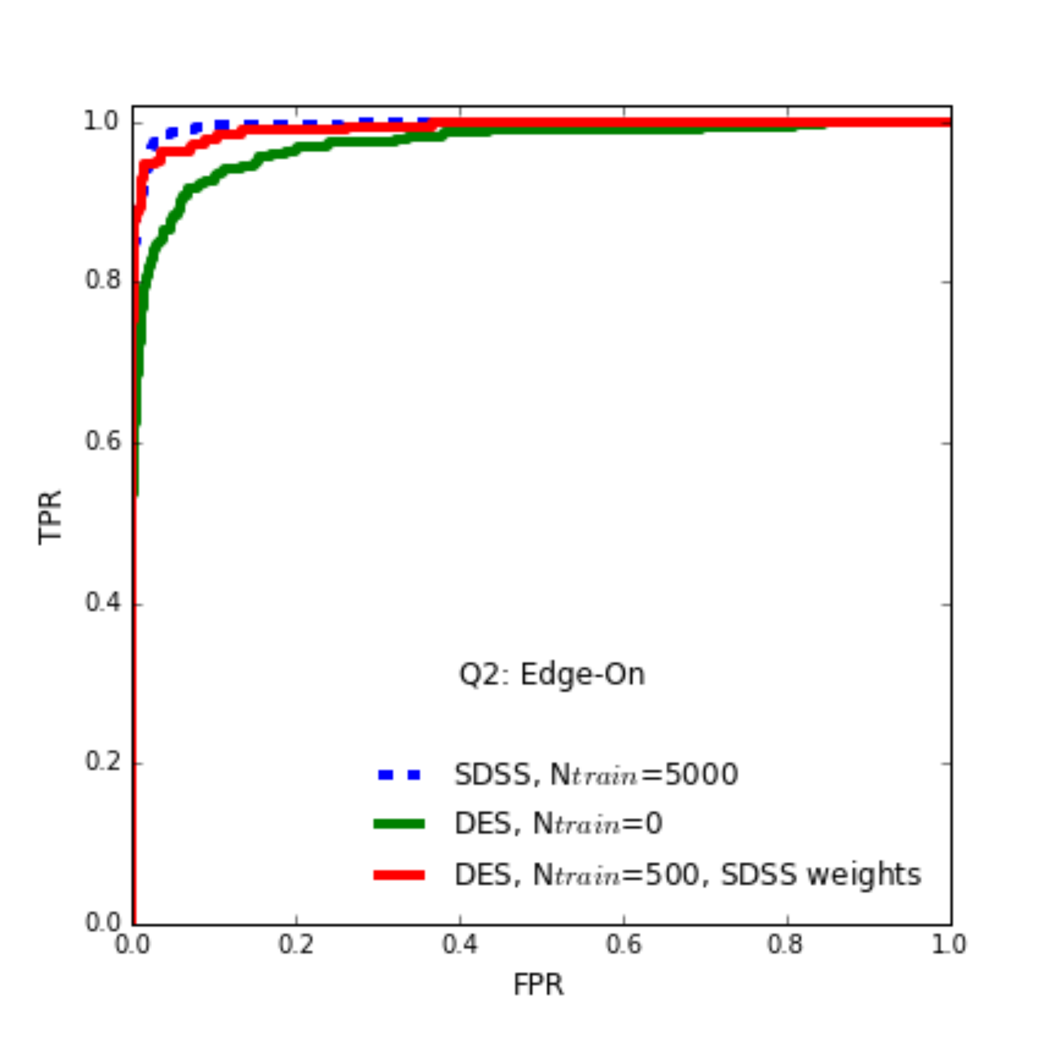
(2) Fast domain adaptation step: train with N~500 DES galaxies using as a starting point the weights from SDSS-based models.

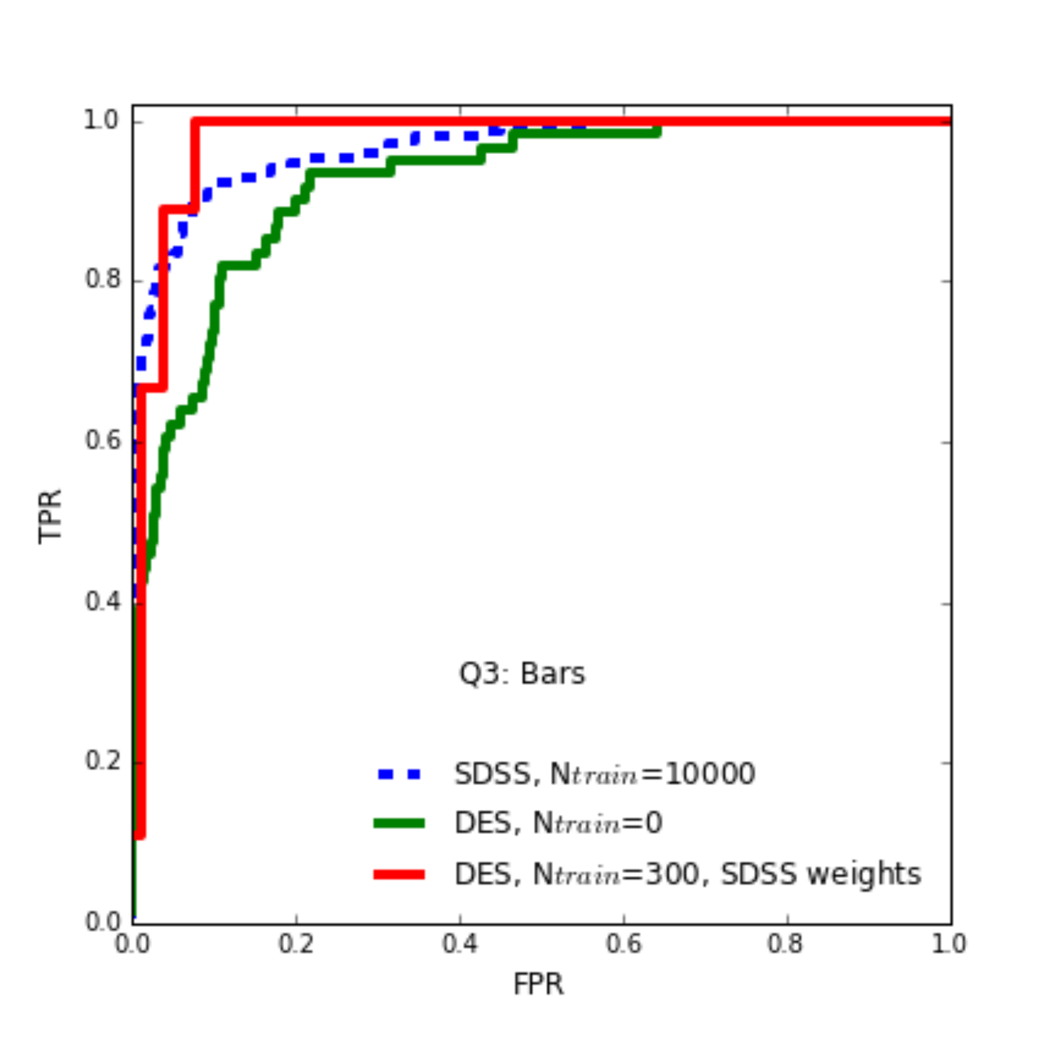
(3) Test DES model on sample not used for training.



Domínguez Sánchez+2019

TRANSFER LEARNING



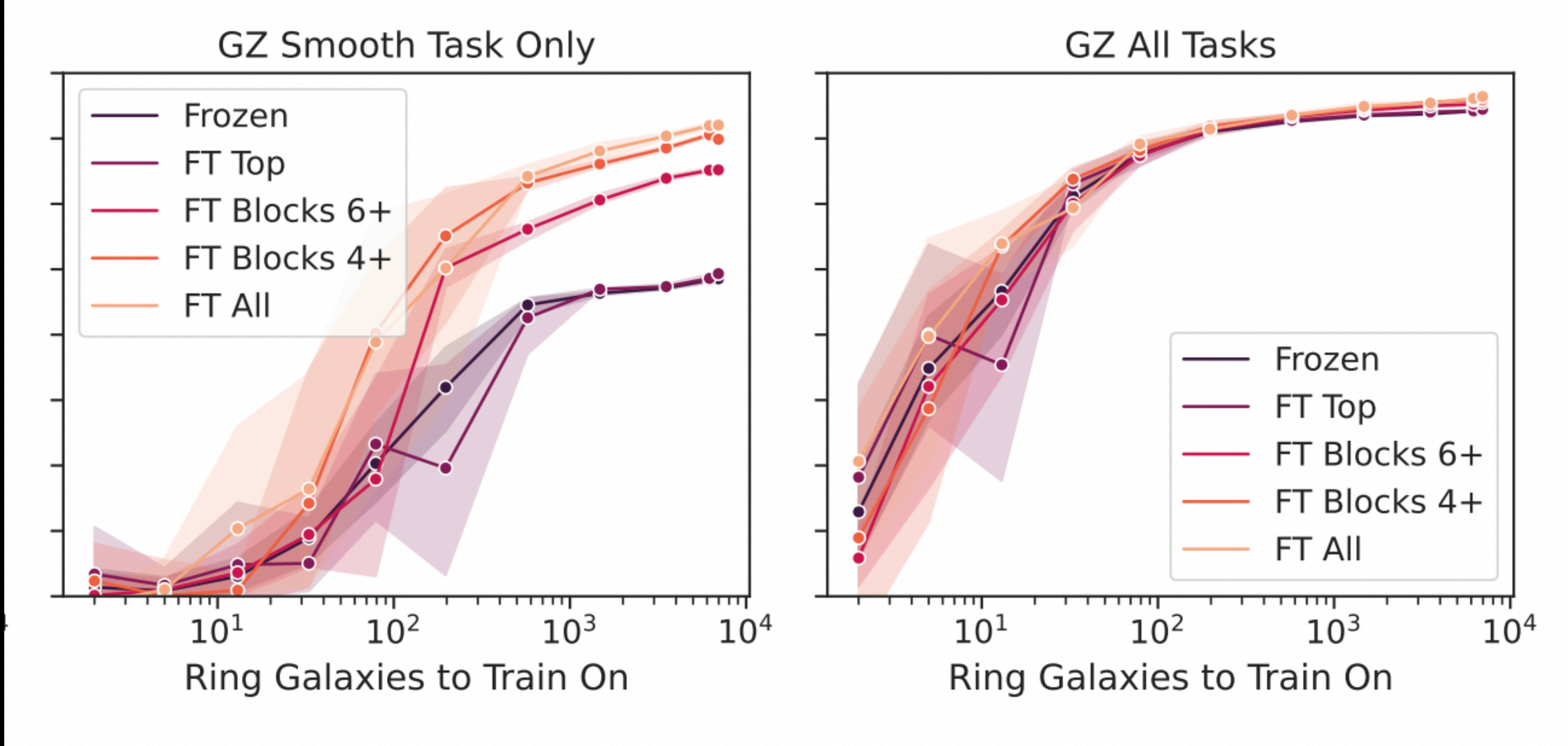


"Recycling" features/weights learned from a different sample helps reducing the training sample by one order of magnitude.

Question	Survey	$Ntrain$	$Ntest$	$Npos$	TPR	Prec.	Acc.
	SDSS	5000	3370	674	0.93	0.91	0.97
Q1	DES	0	2409	797	0.48	0.92	0.81
Smooth/Disk	DES	500	238	78	0.95	0.91	0.95
	SDSS	5000	2687	396	0.98	0.80	0.96
Q2	DES	0	2851	536	0.91	0.73	0.96
Edge-on	DES	500	738	187	0.96	0.86	0.95
	SDSS	10000	1806	169	0.76	0.79	0.96
Q3	DES	0	1768	61	0.57	0.35	0.95
Bar sign	DES	300	86	9	0.89	0.73	0.95

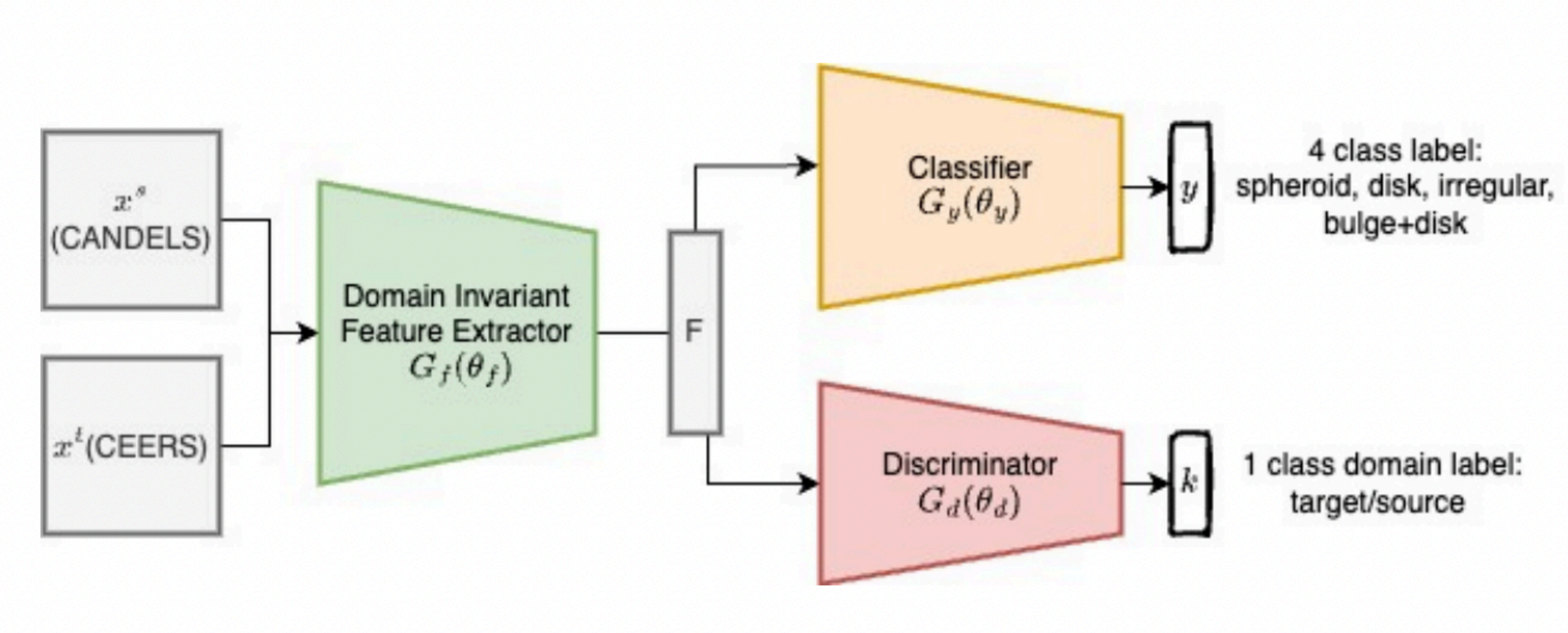
Domínguez Sánchez+2019

TRANSFER LEARNING

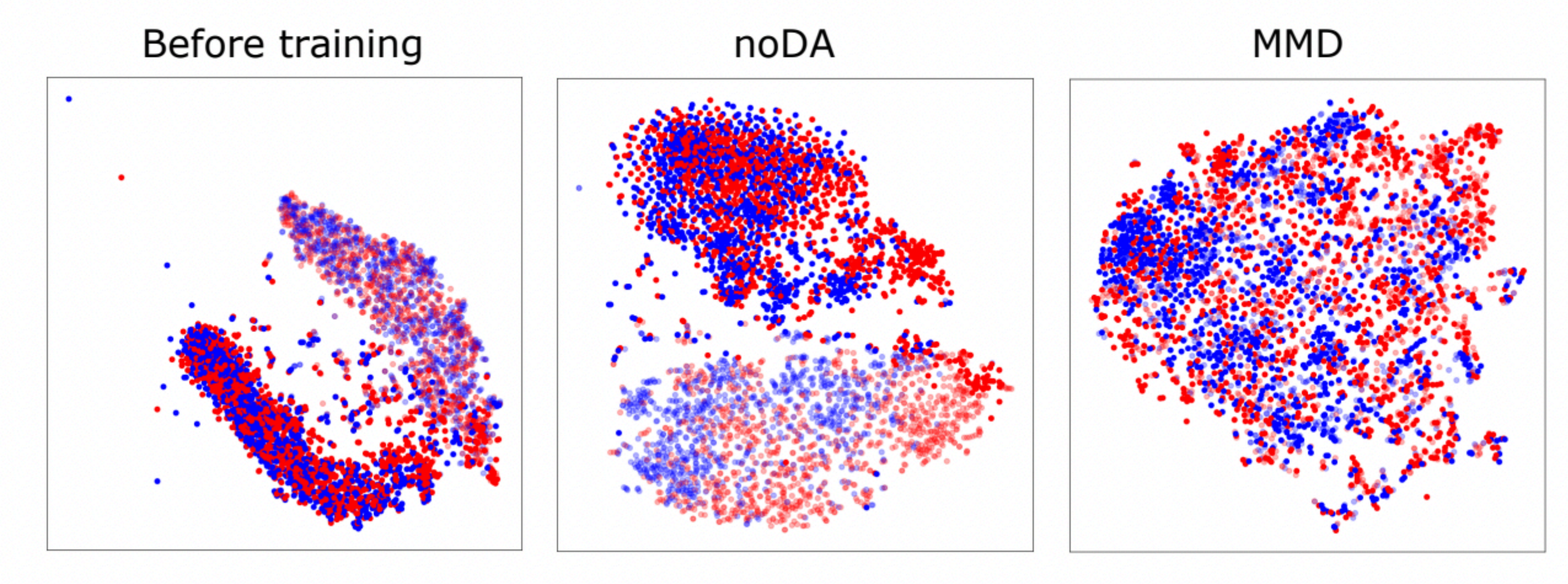


Walmsley+2022

DOMAIN ADAPTATION



Huertas Company+2023



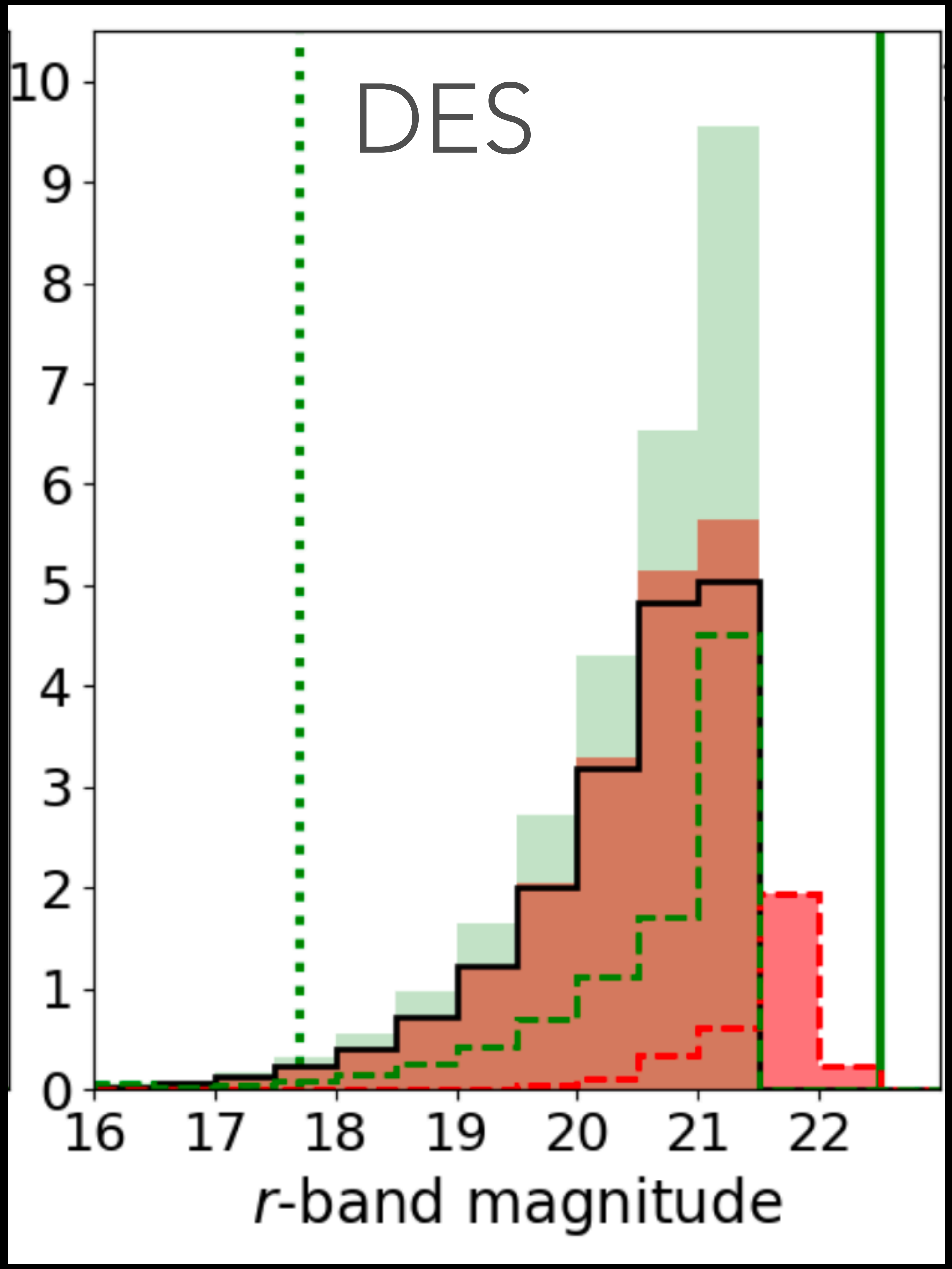
Ćiprijanović+2020, 2021, 2023

See also Izbicki+2016; Xu+23

HOW TO CLASSIFY DISTANT (FAINT) GALAXIES?

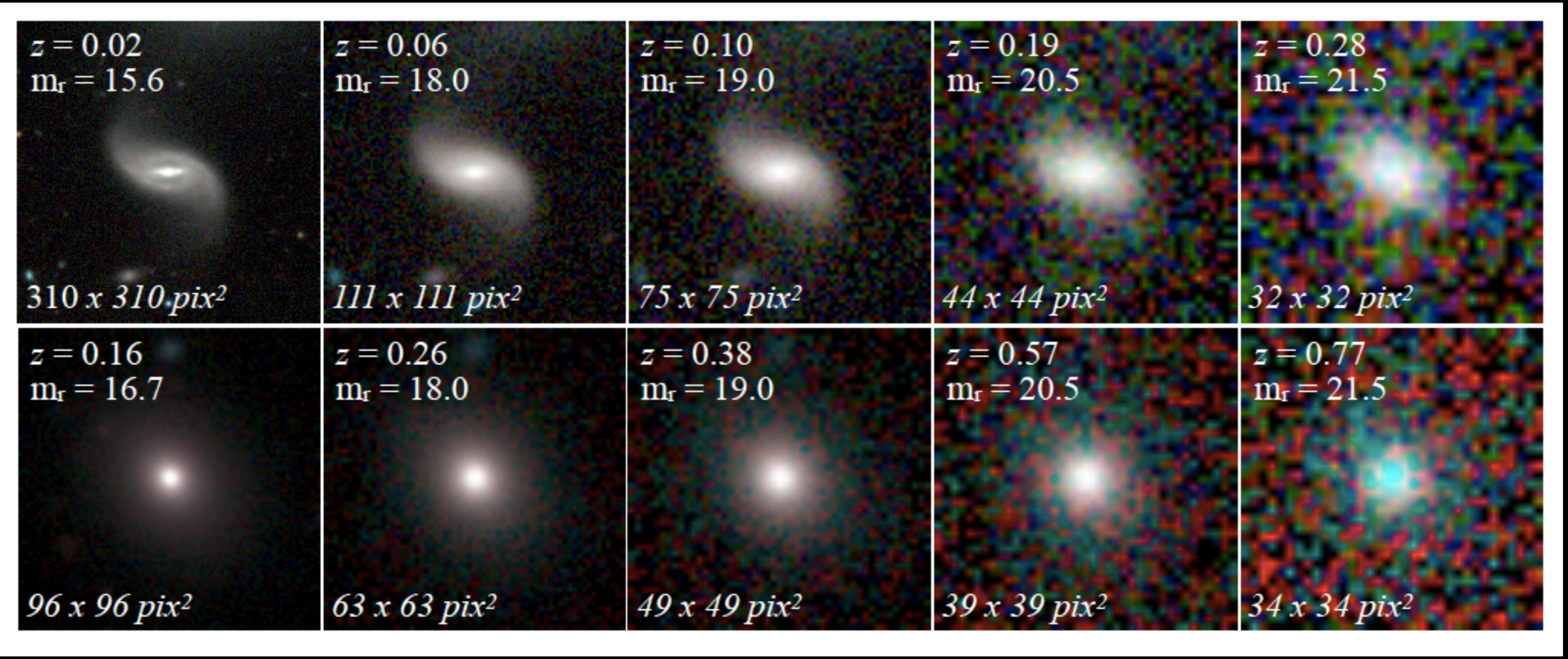
- To understand the impact of morphology in galaxy evolution, we need to know how morphology changes with cosmic time (z=redshift)
- Distant galaxies look fainter (cosmological dimming)
- Catalogues used for labelling training the sample are limited to bright galaxies (m < 17 mag)
- Need labelled samples of faint galaxies… but they are difficult to classify by eye!

HOW TO CLASSIFY DISTANT (FAINT) GALAXIES?



- To understand the impact of morphology in galaxy evolution, we need to know how morphology changes with cosmic time (z=redshift)
- Distant galaxies look fainter (cosmological dimming)
- Catalogues used for labelling training the sample are limited to bright galaxies (m < 17 mag)
- Need labelled samples of faint galaxies… but they are difficult to classify by eye!

HOW TO CLASSIFY DISTANT (FAINT) GALAXIES?



'Emulate' real DES galaxies as if they were further away from us (higher z)

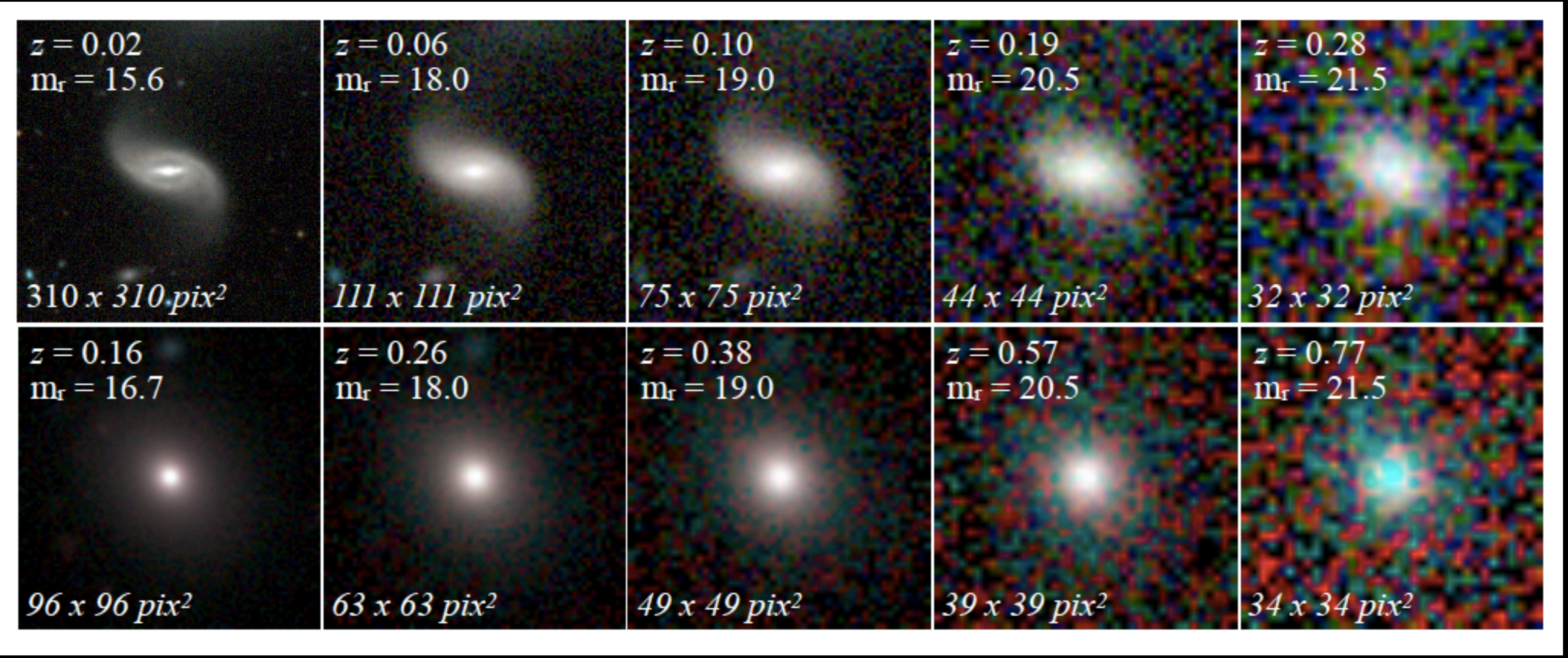
✓Cosmological dimming: flux and size (N pixels)

✓ k-correction + evolution

✓ PSF +noise

Vega-Ferrero, DS+2021

HOW TO CLASSIFY DISTANT (FAINT) GALAXIES?



SPIRAL
(LTG)

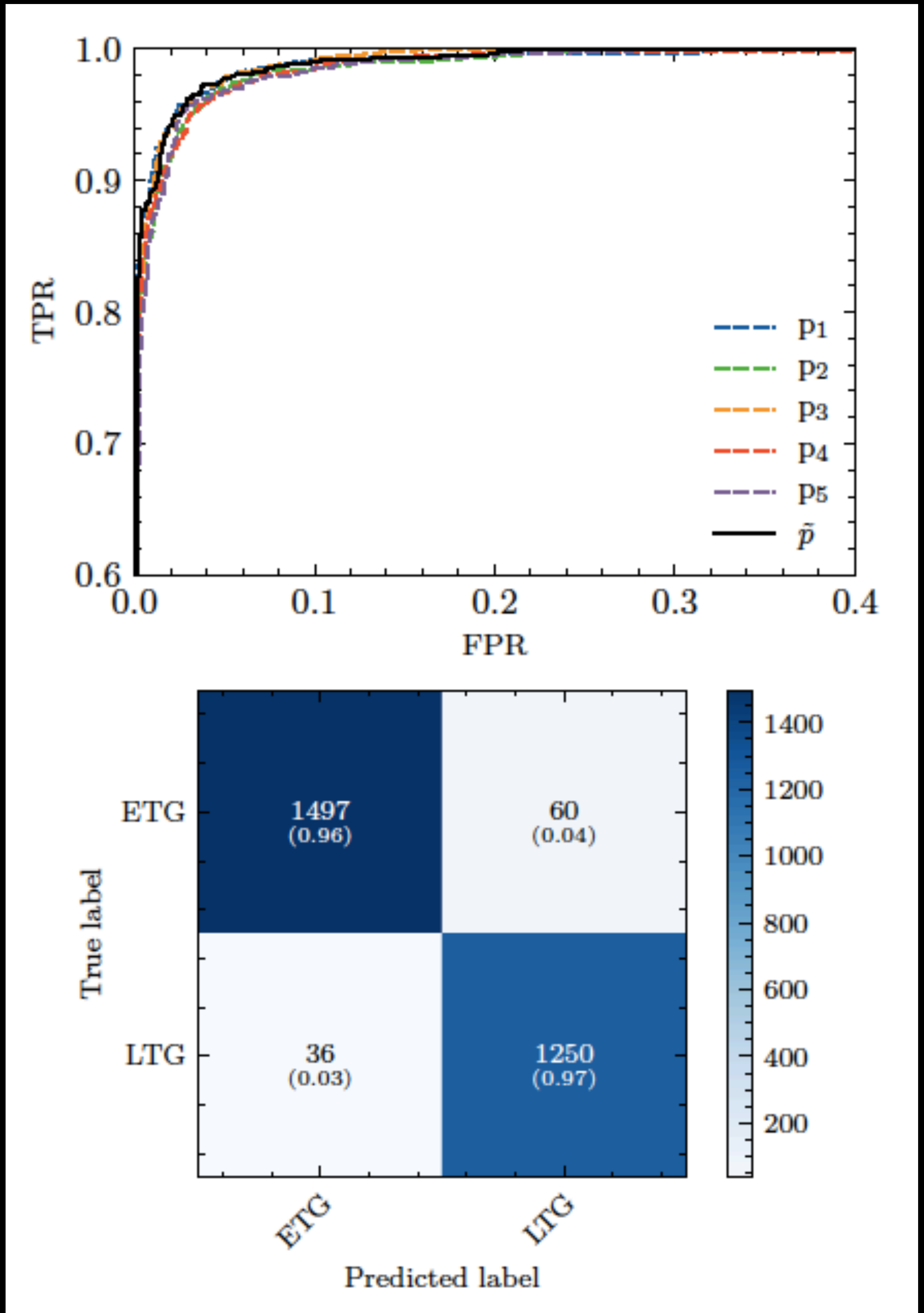
ELLIPTICAL
(ETG)

Trick: Use original labels for training

Can machines recover features hidden to the human eye?

Vega-Ferrero, DS+2021

HOW TO CLASSIFY DISTANT (FAINT) GALAXIES?



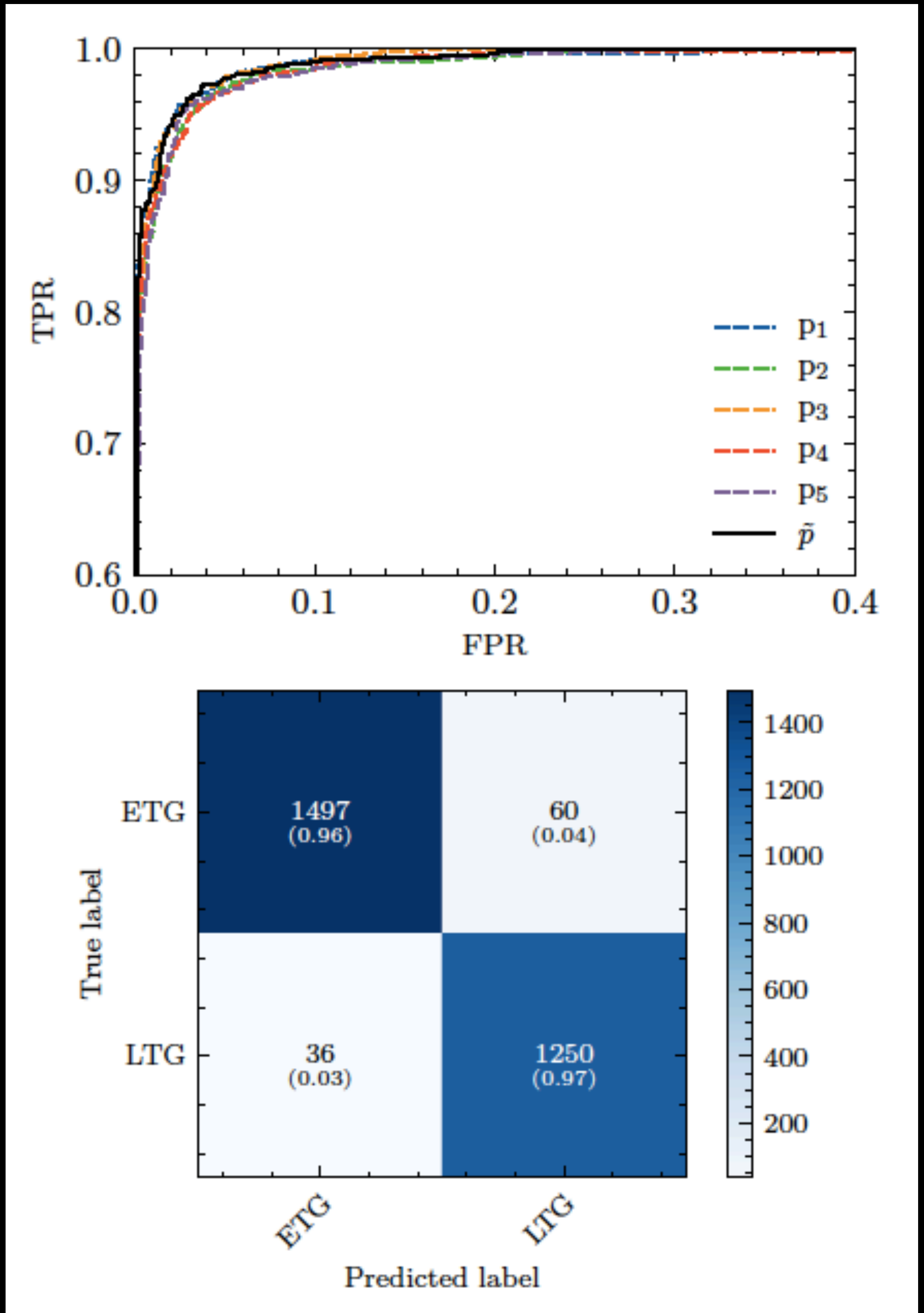
ELLIPTICAL VS SPIRAL

K-folding with 5 different training sets

Model	p_{thr}	ROC AUC	Prec	R	Acc
p_1	0.49	0.99	0.97	0.96	0.97
p_2	0.46	0.99	0.95	0.96	0.96
p_3	0.39	0.99	0.96	0.97	0.97
p_4	0.35	0.99	0.95	0.97	0.96
p_5	0.54	0.99	0.96	0.96	0.96
$\tilde{p}$	0.40	0.99	0.95	0.97	0.97

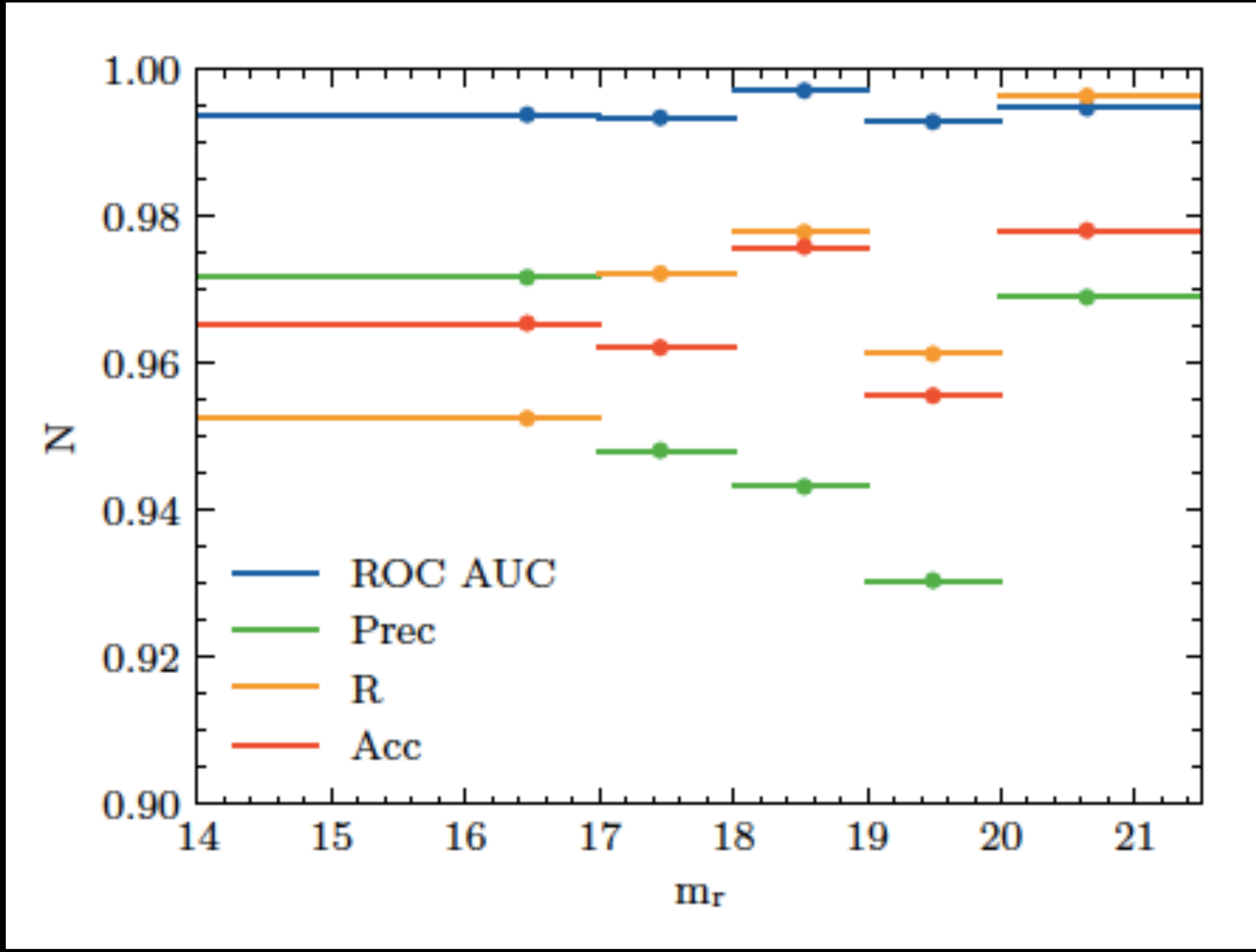
Vega-Ferrero, DS+2021

HOW TO CLASSIFY DISTANT (FAINT) GALAXIES?



Are the results affected by observed magnitude (faintness of the galaxy)?

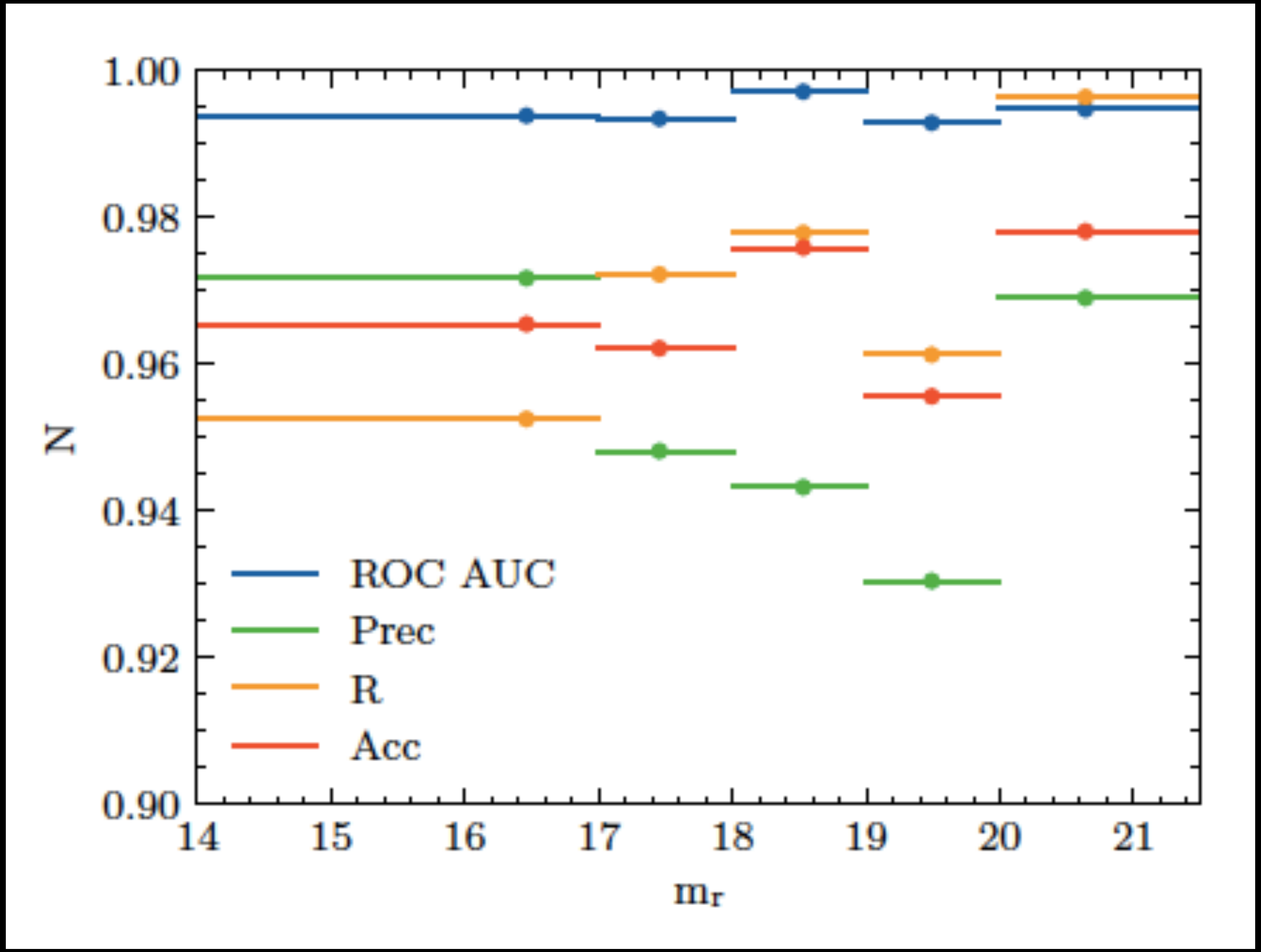
NO!



Vega-Ferrero, DS+2021

HOW TO CLASSIFY DISTANT (FAINT) GALAXIES?

Are the results affected by observed magnitude (faintness of the galaxy)?

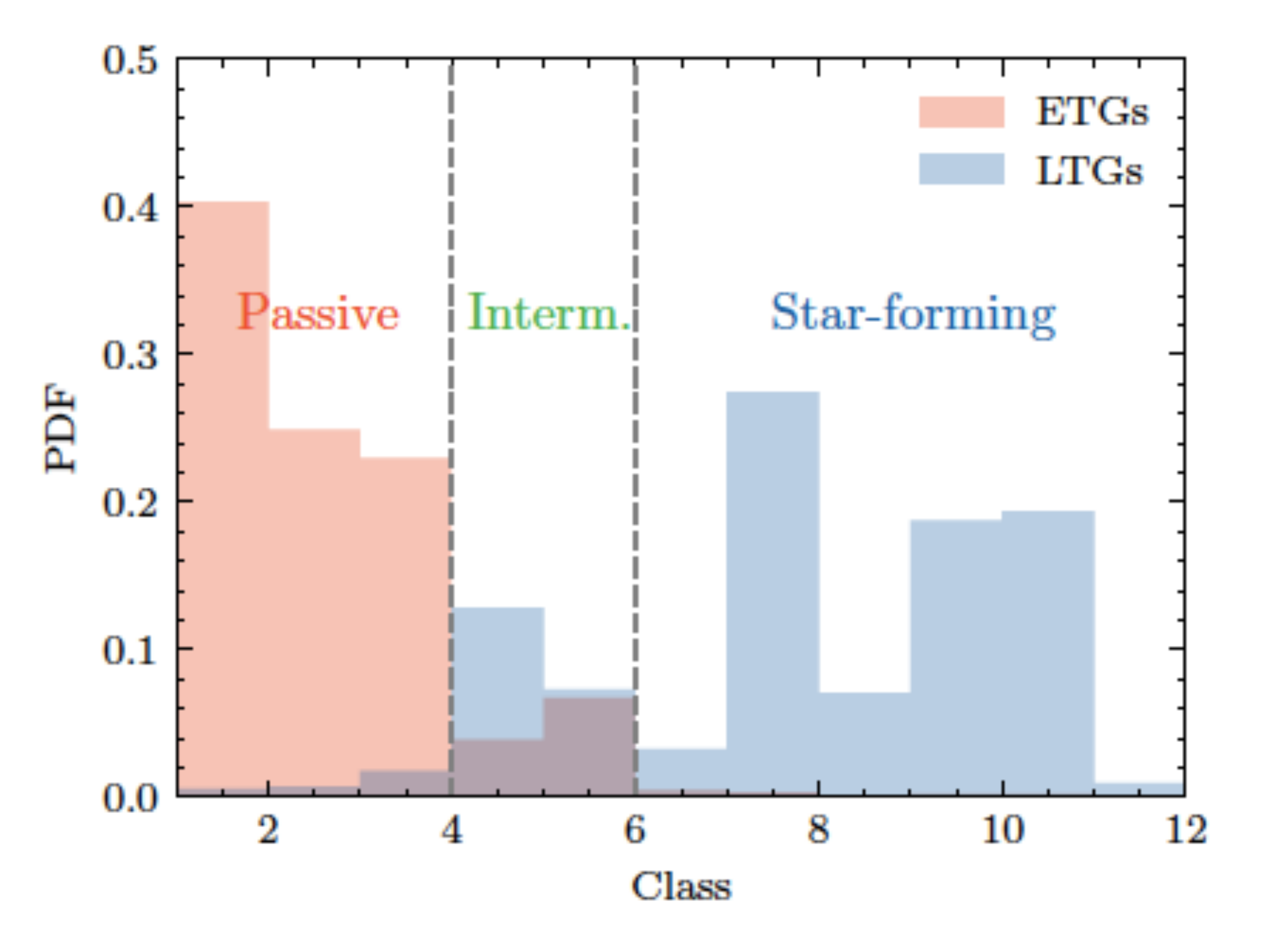


Can machines recover features hidden to the human eye?

YES!

Vega-Ferrero, DS+2021

TEST ON REAL DES DISTANT GALAXIES



- VIPERS spectral classification (accurate data, related to stellar populations)

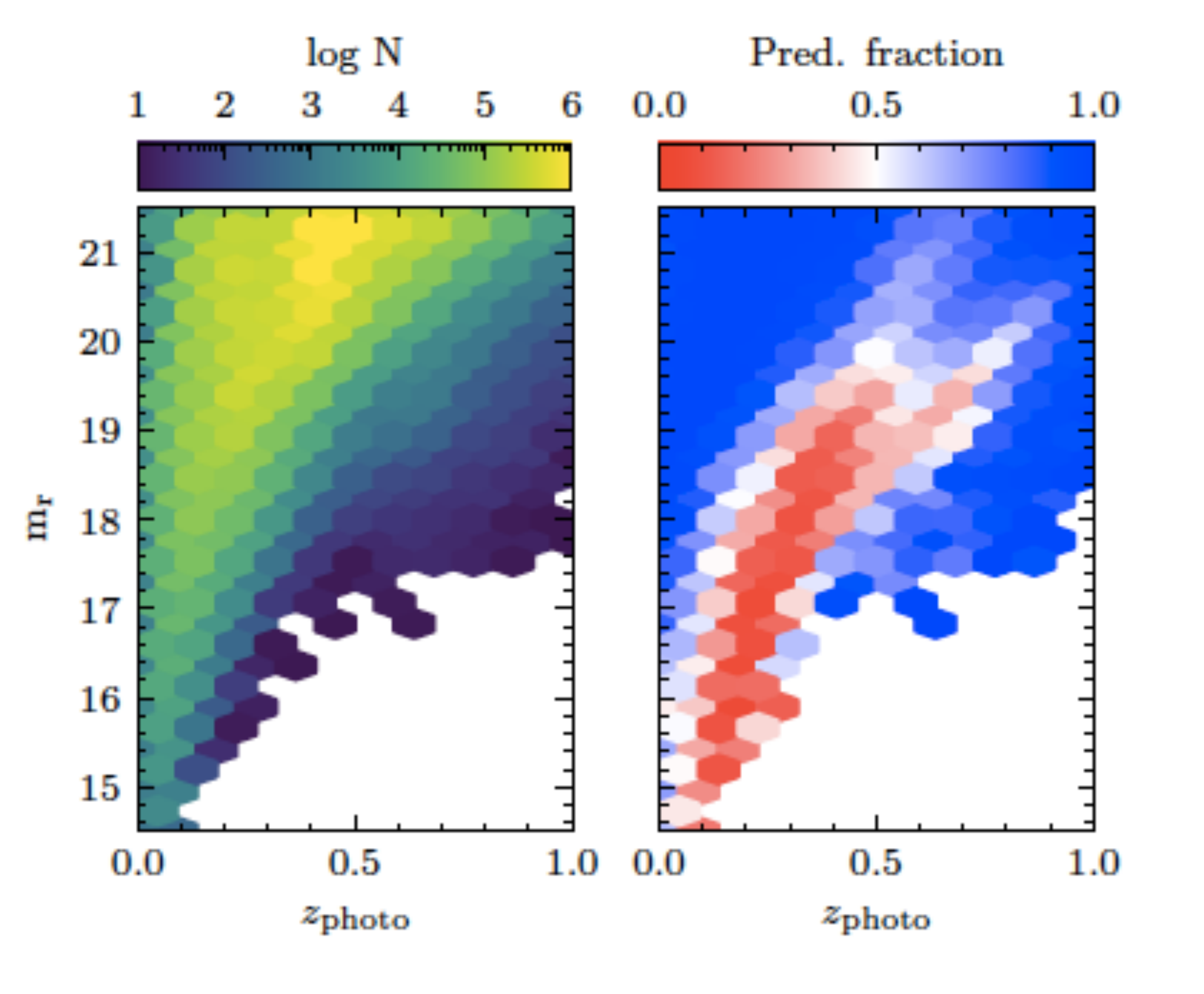
- Morphological classification clearly correlate with spectral class:

✓ 97 % TP (LTG with class > 4)
✓ 89% TN (ETGs with class < 4)

Vega-Ferrero, DS+2021

DES MORPHOLOGICAL CATALOGUE

https://des.ncsa.illinois.edu/releases/y3a2/gal-morphology

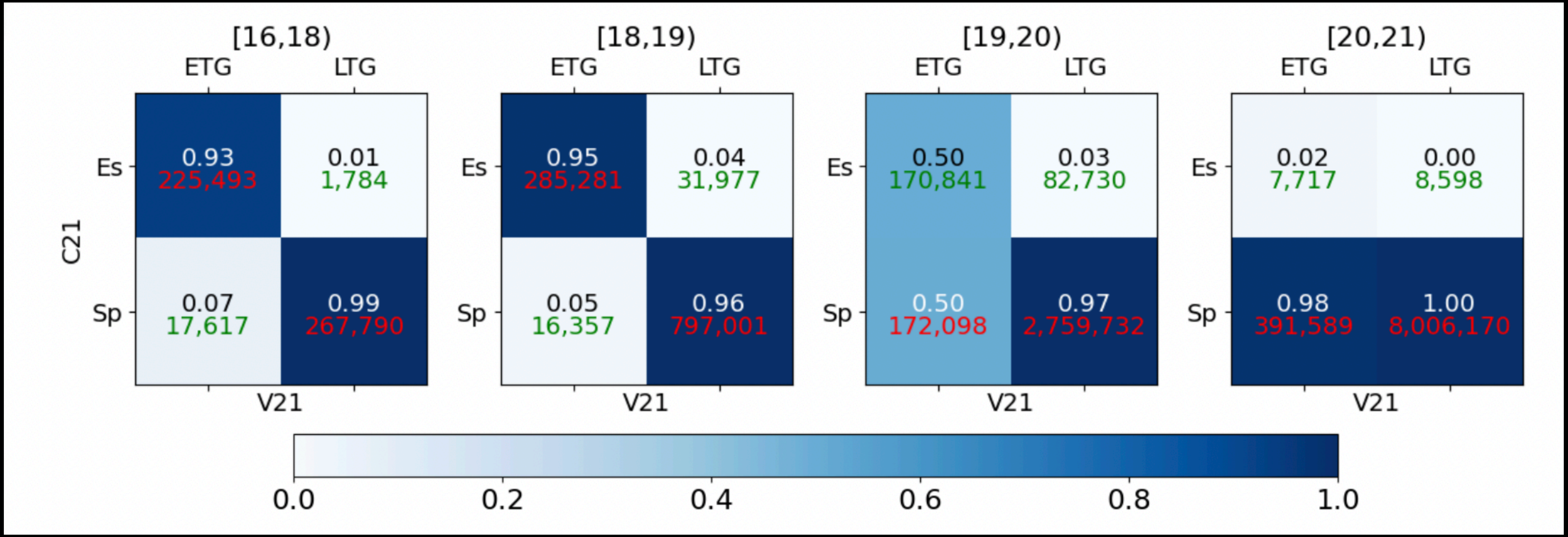


LARGEST MORPHOLOGICAL CATALOG UP TO DATE!

- ✓27 million galaxies
- ✓Up to 21.5 mag
- ✓12% ETGs, 88% LTGs
- ✓Released with paper

Vega-Ferrero, DS+2021

WHAT HAPPENS WHEN YOUR TRAINING SAMPLE IS NOT IN THE SAME DOMAIN AS YOUR TARGET SAMPLE?



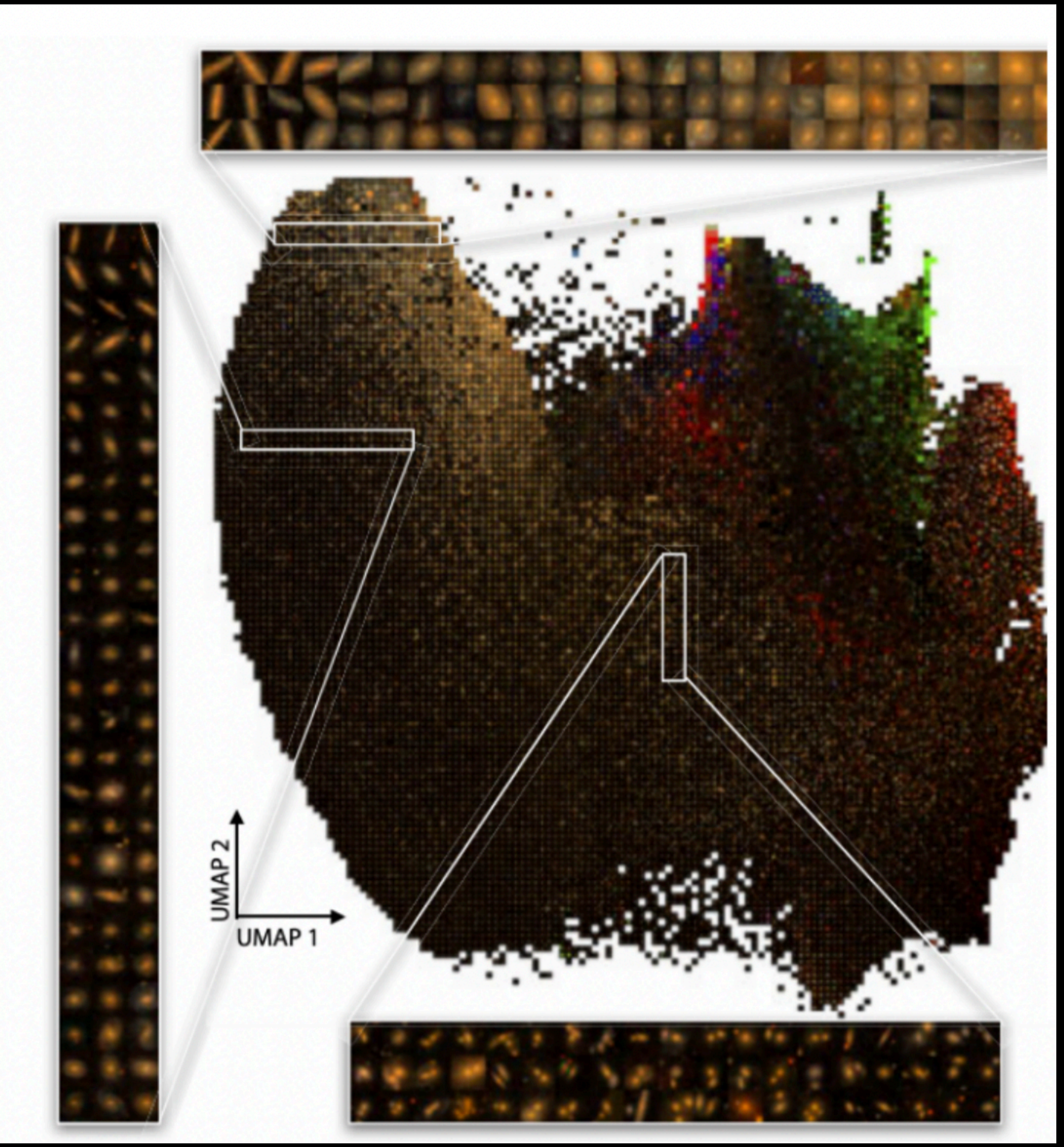
✓ Comparison of VF+21 classification (trained with emulations) and Cheng+21 catalogue (trained with mag < 18)

Cheng, DS, JVF +2023

UNSUPERVISED LEARNING

- No need for labels (only for interpreting the results)
- Let the algorithms find structure on the data
- Excellent for anomaly detection

  - Clustering (e.g., k-means)
  - Dimensionality reduction (t-SNE, UMAP, PCA)
  - Contrastive learning



Cheng+21, Spindler+21, Zhou+21, Stein+21, Walmsley+22b, Slipcevic+22, Shen+22, Doorembos+22, Lamdouar+22, Guo+22, Mercea+23, Lanusse+23

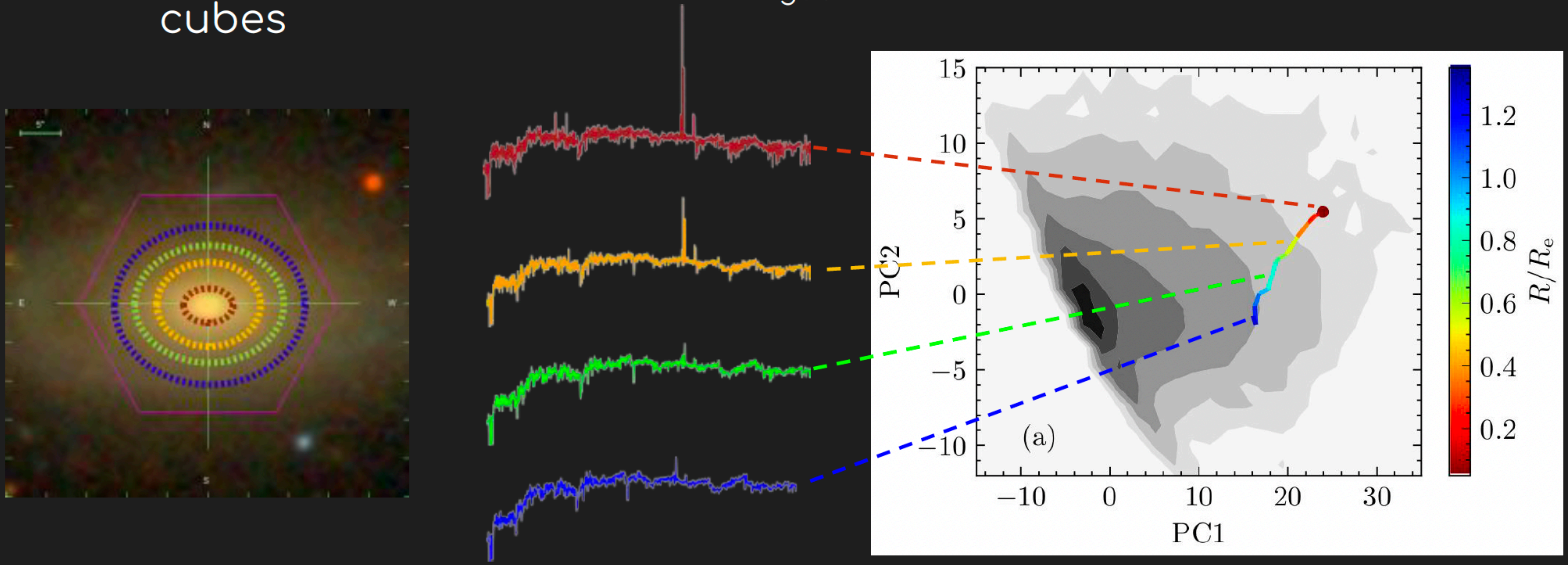
PCA

ACTIVITY GRADIENTS IN MANGA S0 2D SPECTRA

Radial binning of ~1000 MaNGA data cubes

Mean flux as a function of $R_{gal.}$

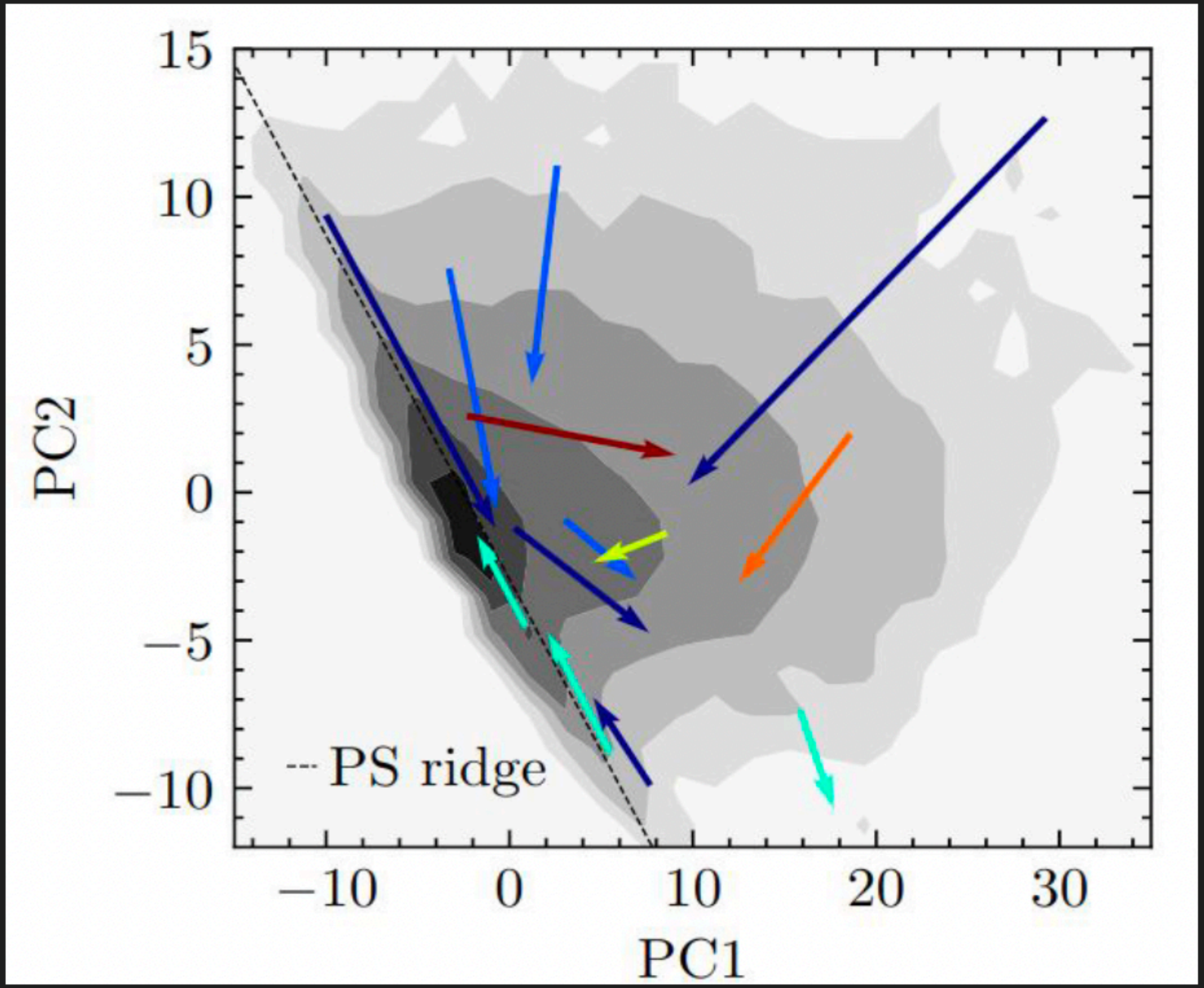
Projection into the first 2 principal components



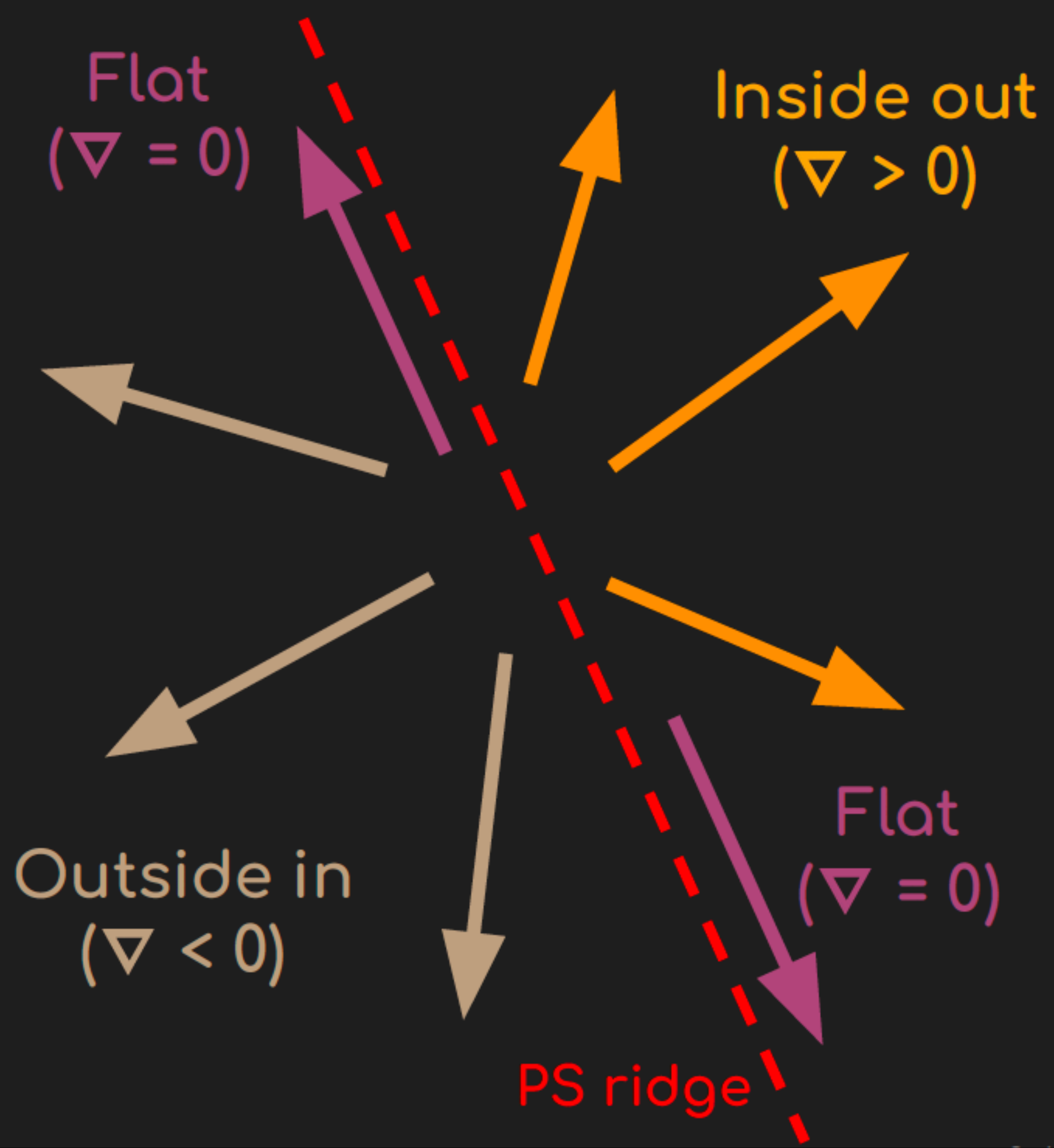
Tous+2021, 23,

PCA
ACTIVITY GRADIENTS IN MANGA S0 2D SPECTRA

The PC1 – PC2 profiles are converted into vectors

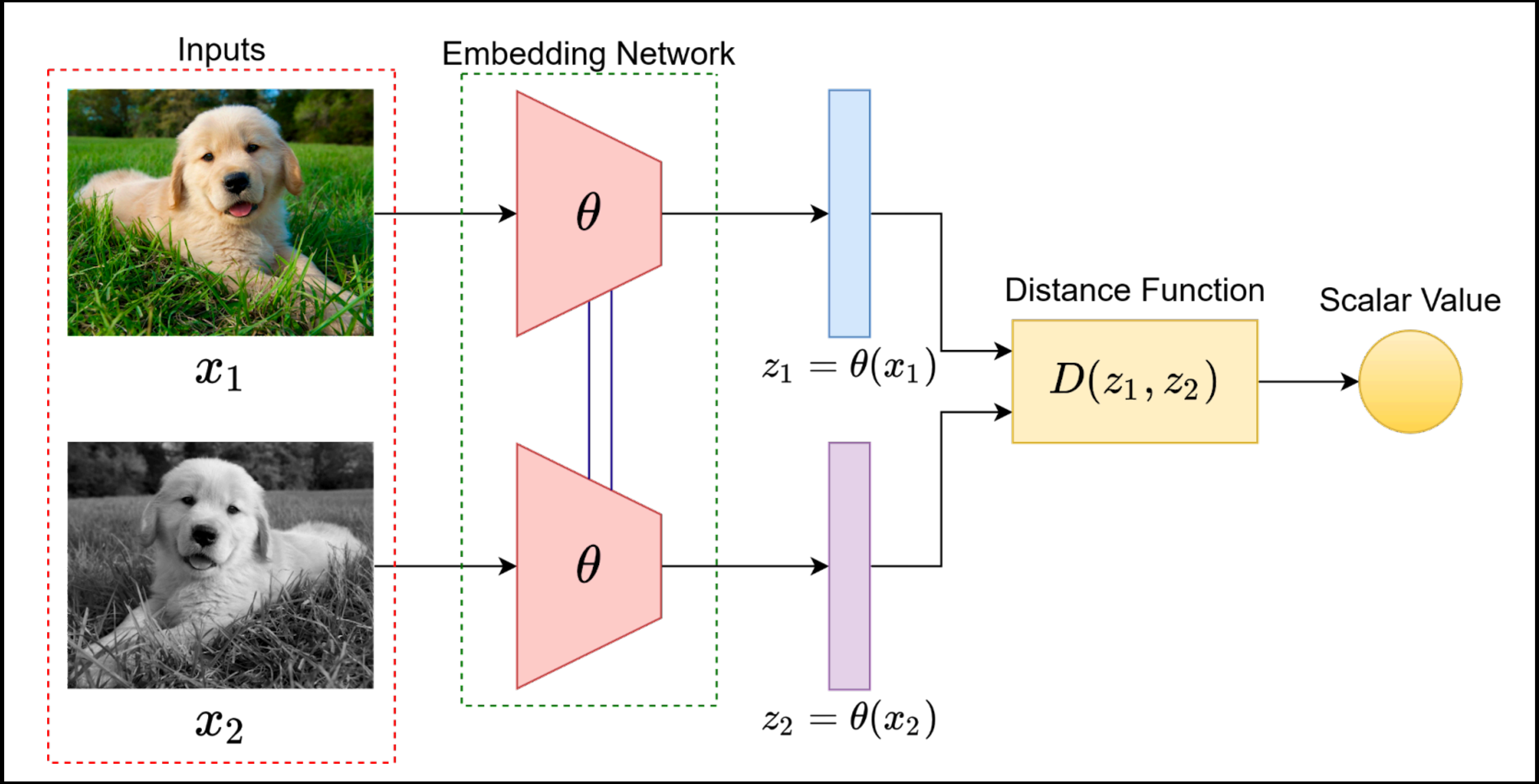


Orientation of vectors
→ sign of activity gradient



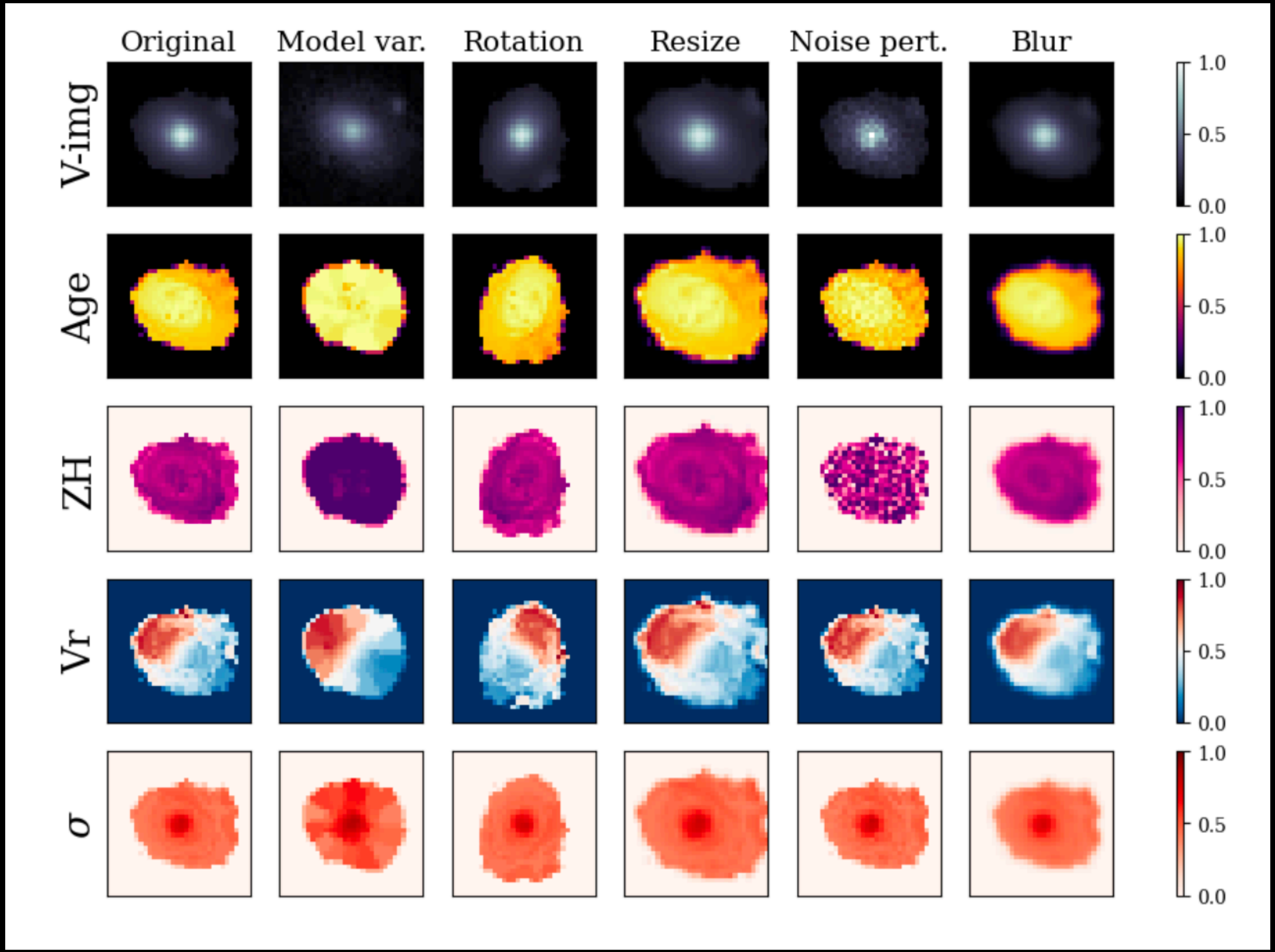
Tous+2021, 23,

SELF-SUPERVISED MACHINE LEARNING



SELF-SUPERVISED LEARNING

MANGA S0 2D SPECTRA

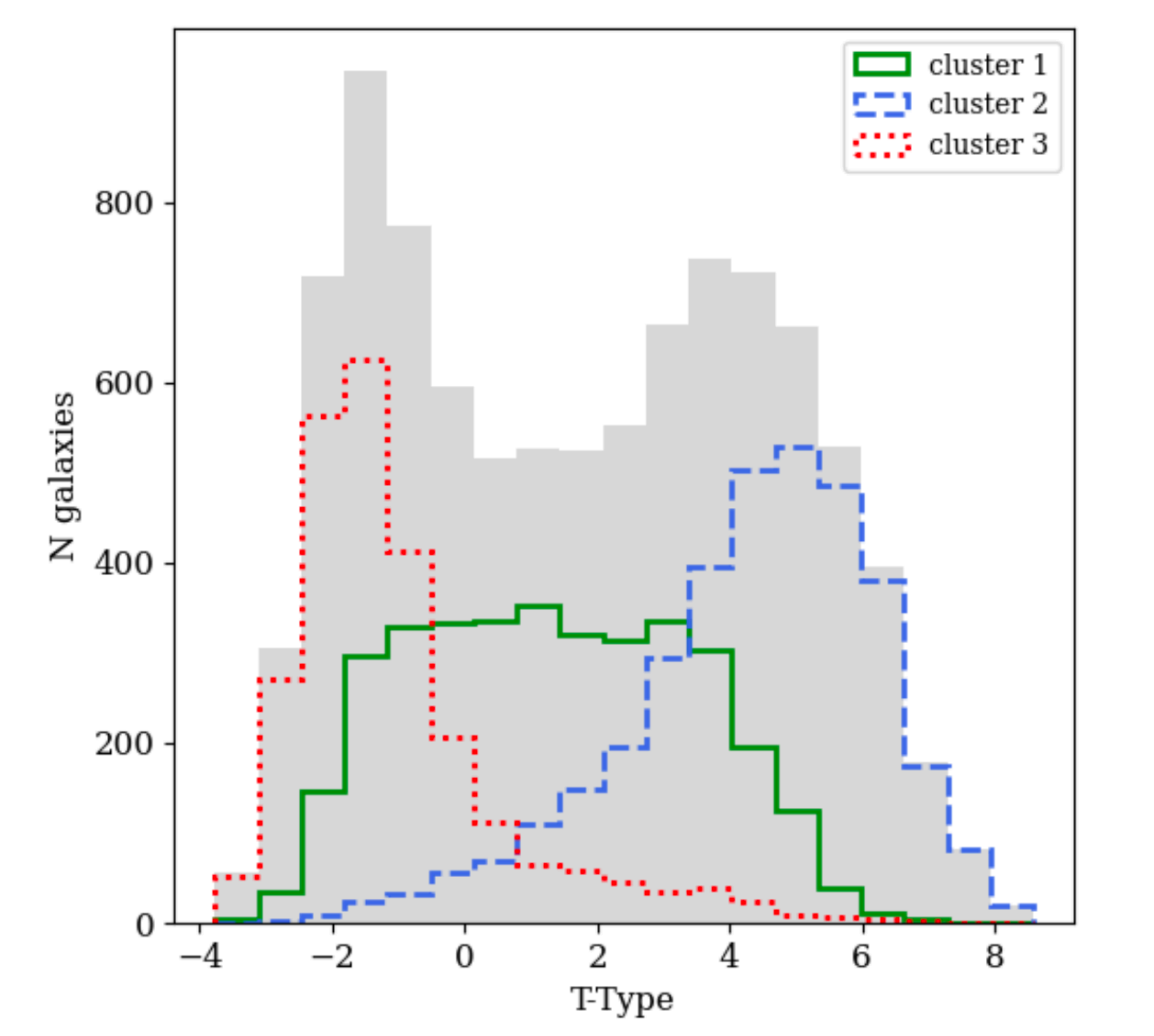


- Input: MaNGA maps of galaxy parameters (Flux, age, metallicity, velocity)
- Augmentations
- Dimensionality reduction: UMAP
- Clustering: k-means

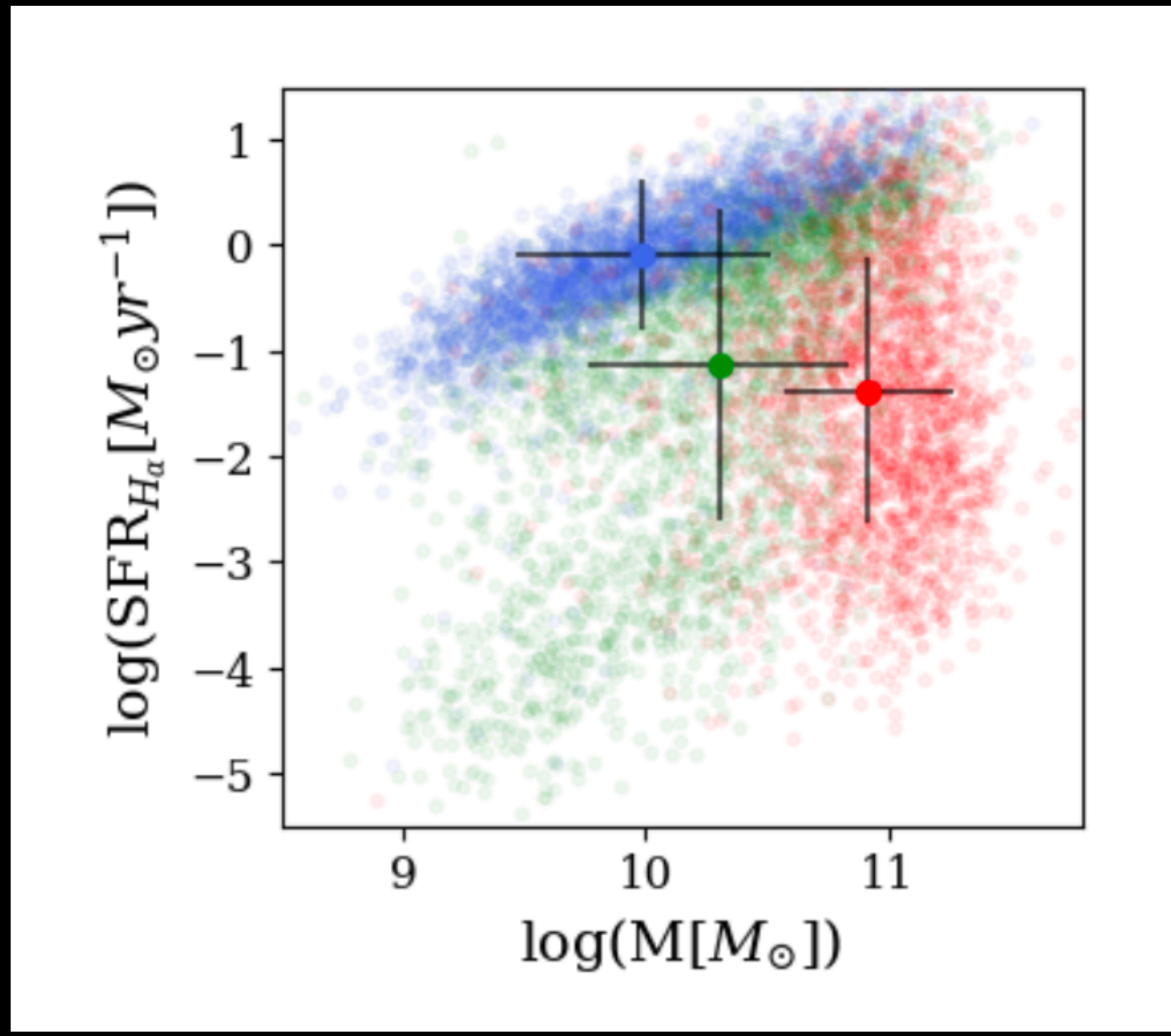
Sarmiento+2021

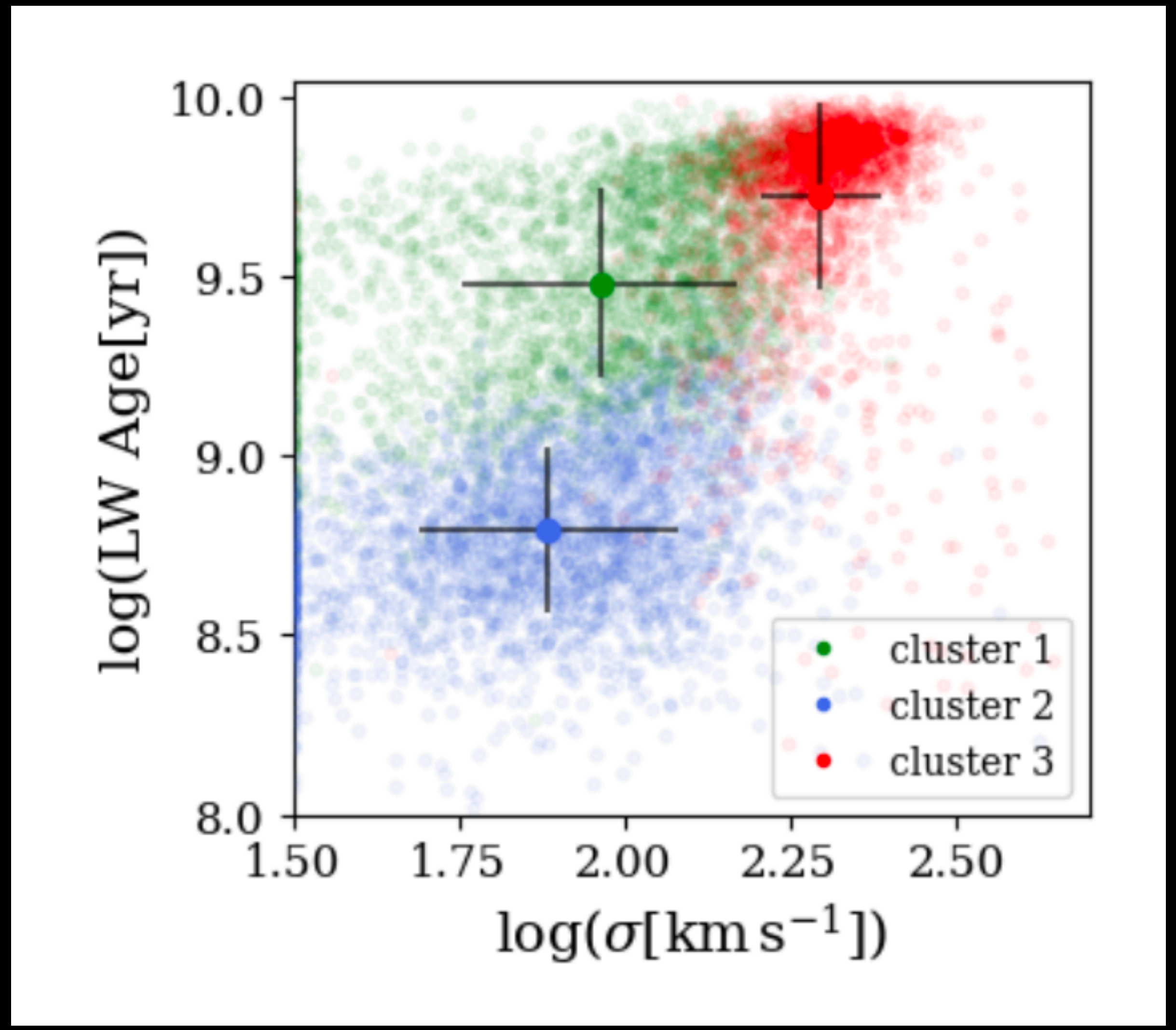
SELF-SUPERVISED LEARNING

MANGA S0 2D SPECTRA



Clusters show good correlation with 'traditional' morphology and physical properties.

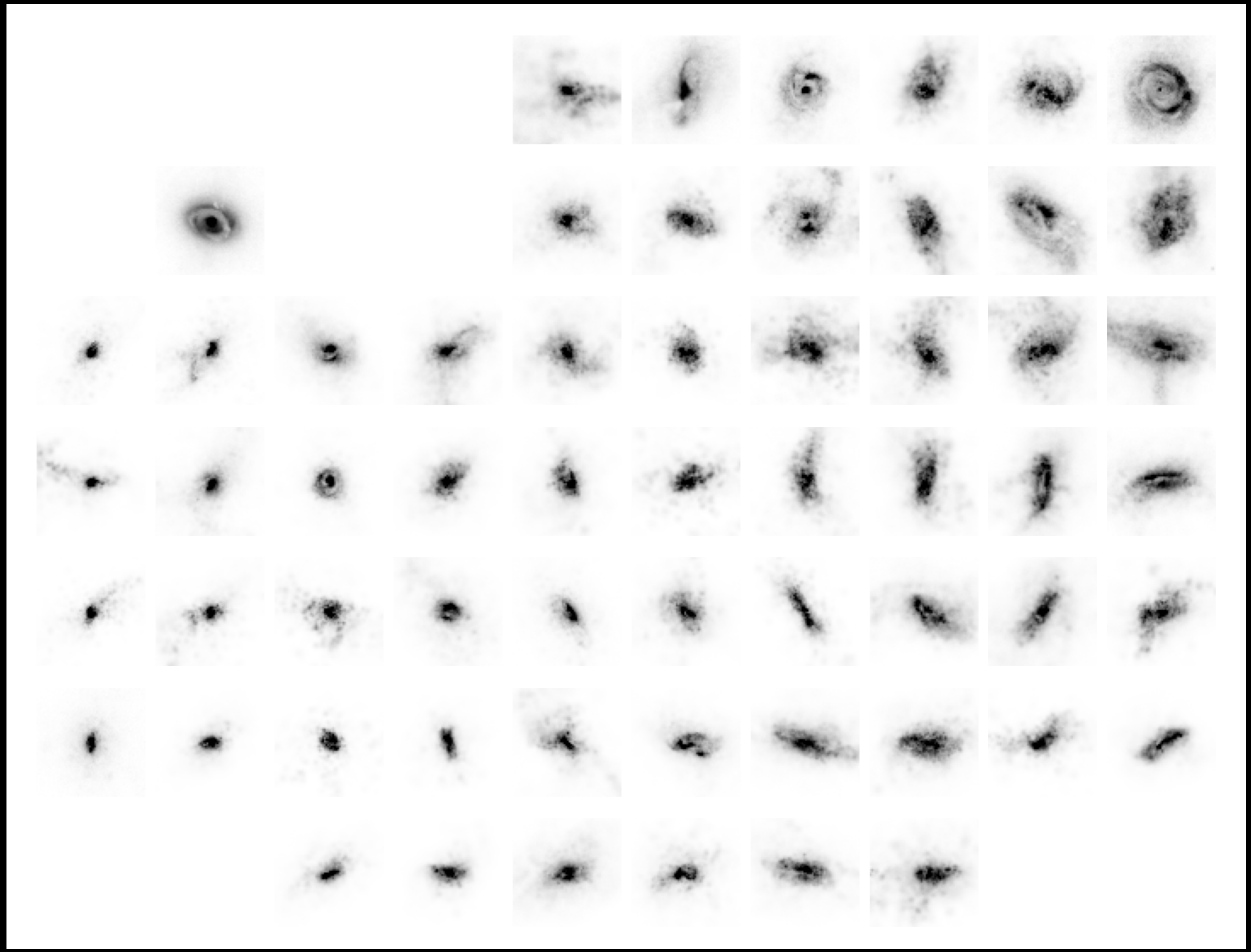


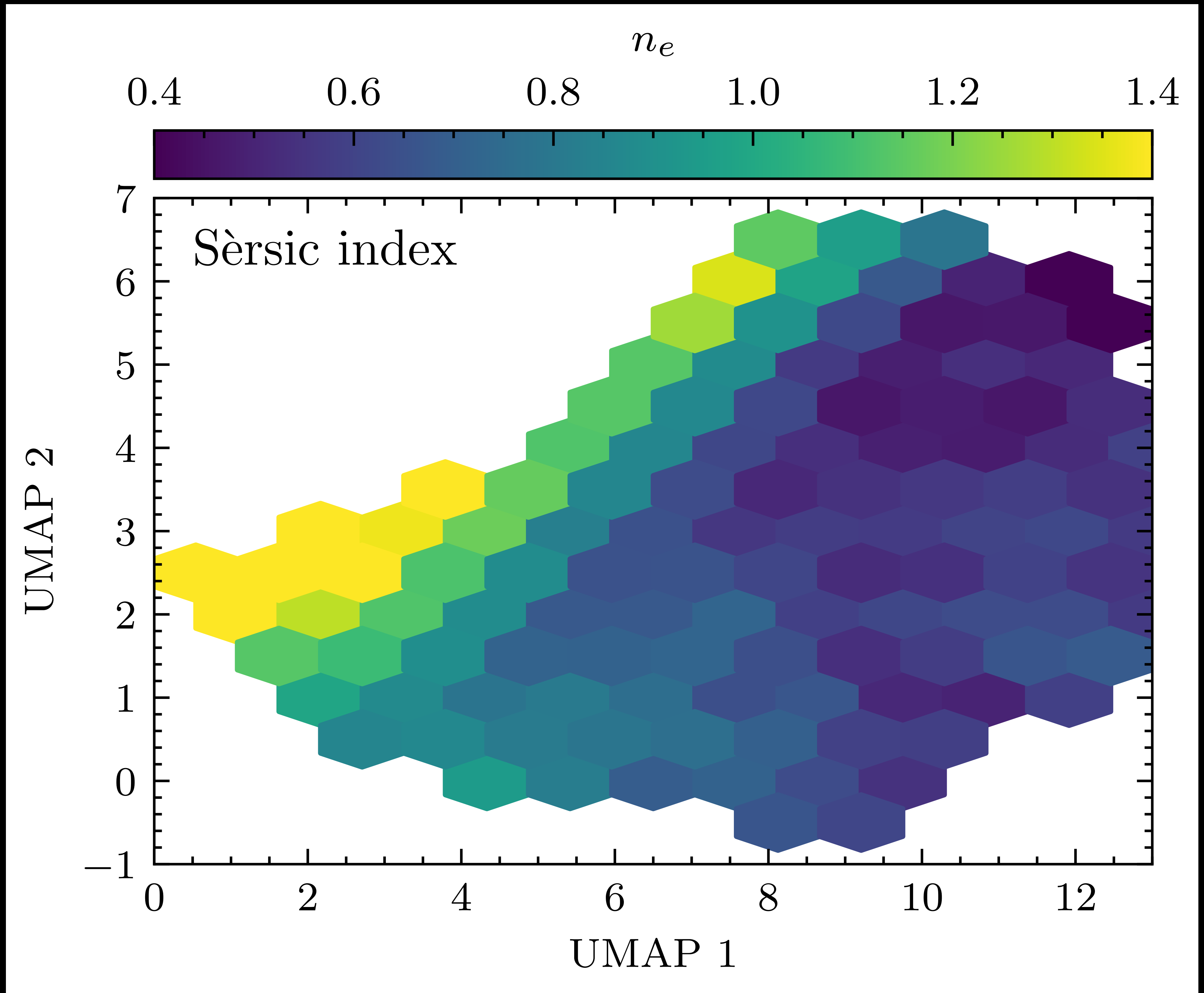


Sarmiento+2021

SELF-SUPERVISED LEARNING

JWST/CEERS IMAGES



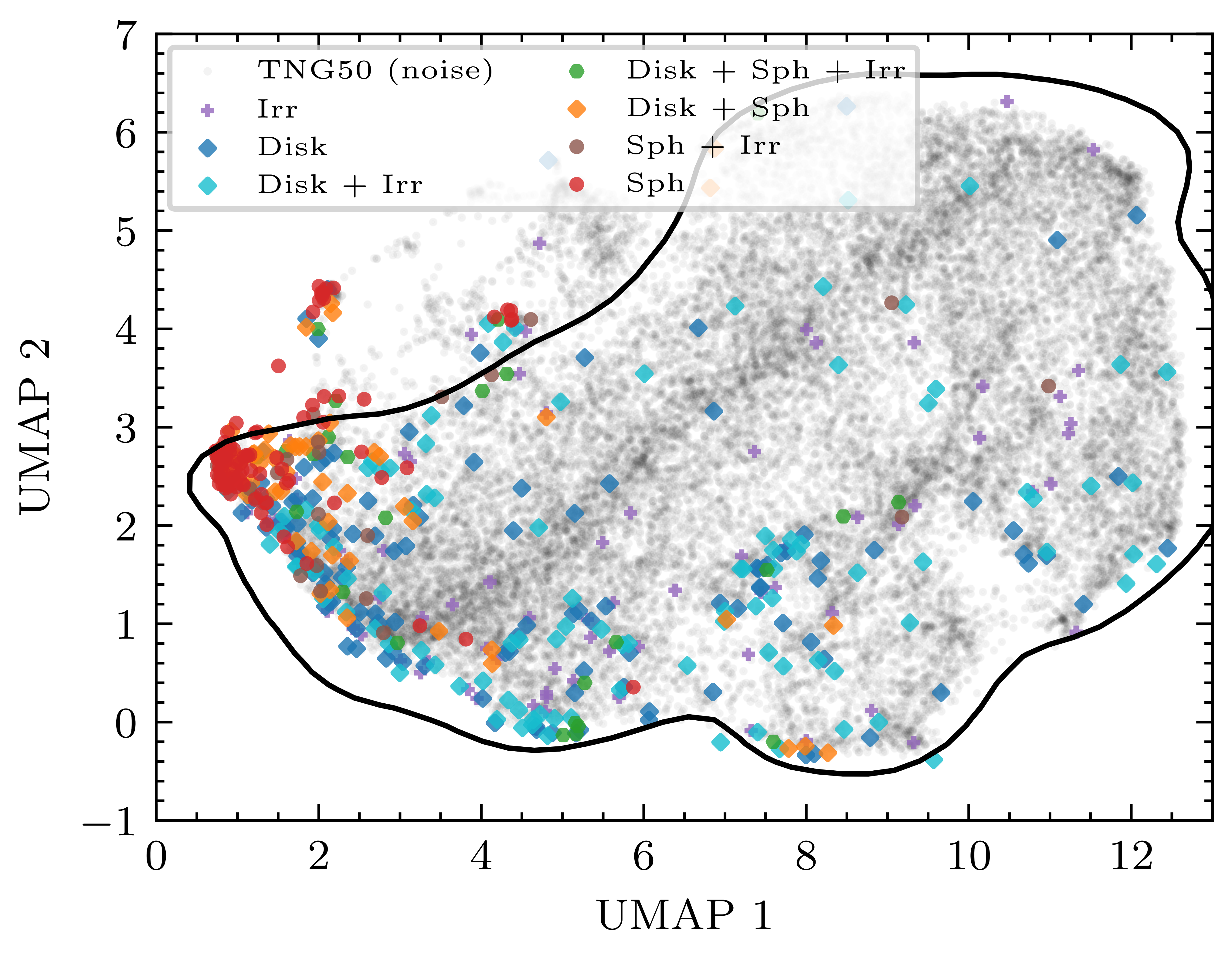


Vega-Ferrero+2023

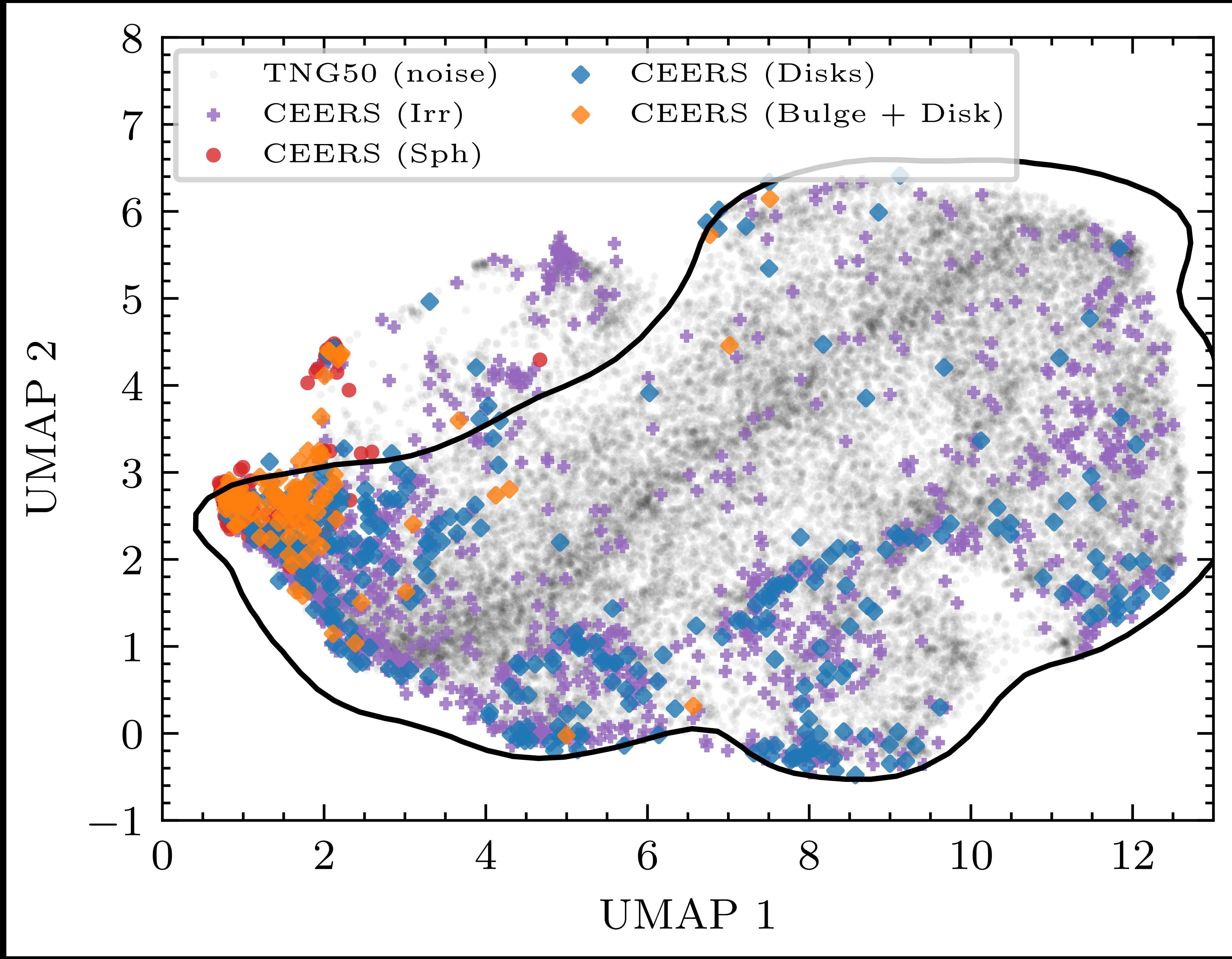
SELF-SUPERVISED LEARNING

JWST/CEERS IMAGES

MHC+2023b
CNN-based
Domain adaptation from CANDELS labels



Kartaltepe+2022
Visual classifications



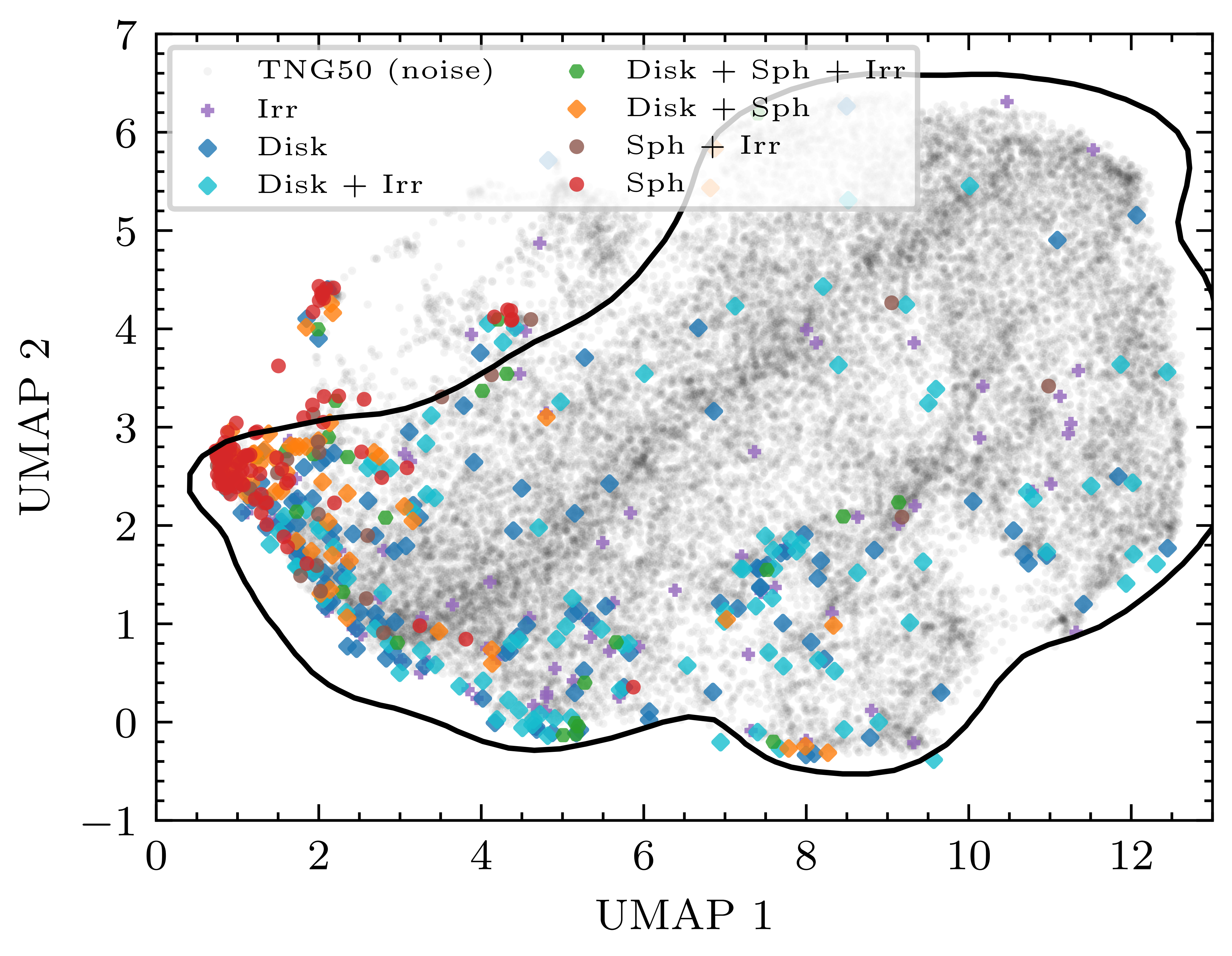
Vega-Ferrero+2023

SELF-SUPERVISED LEARNING

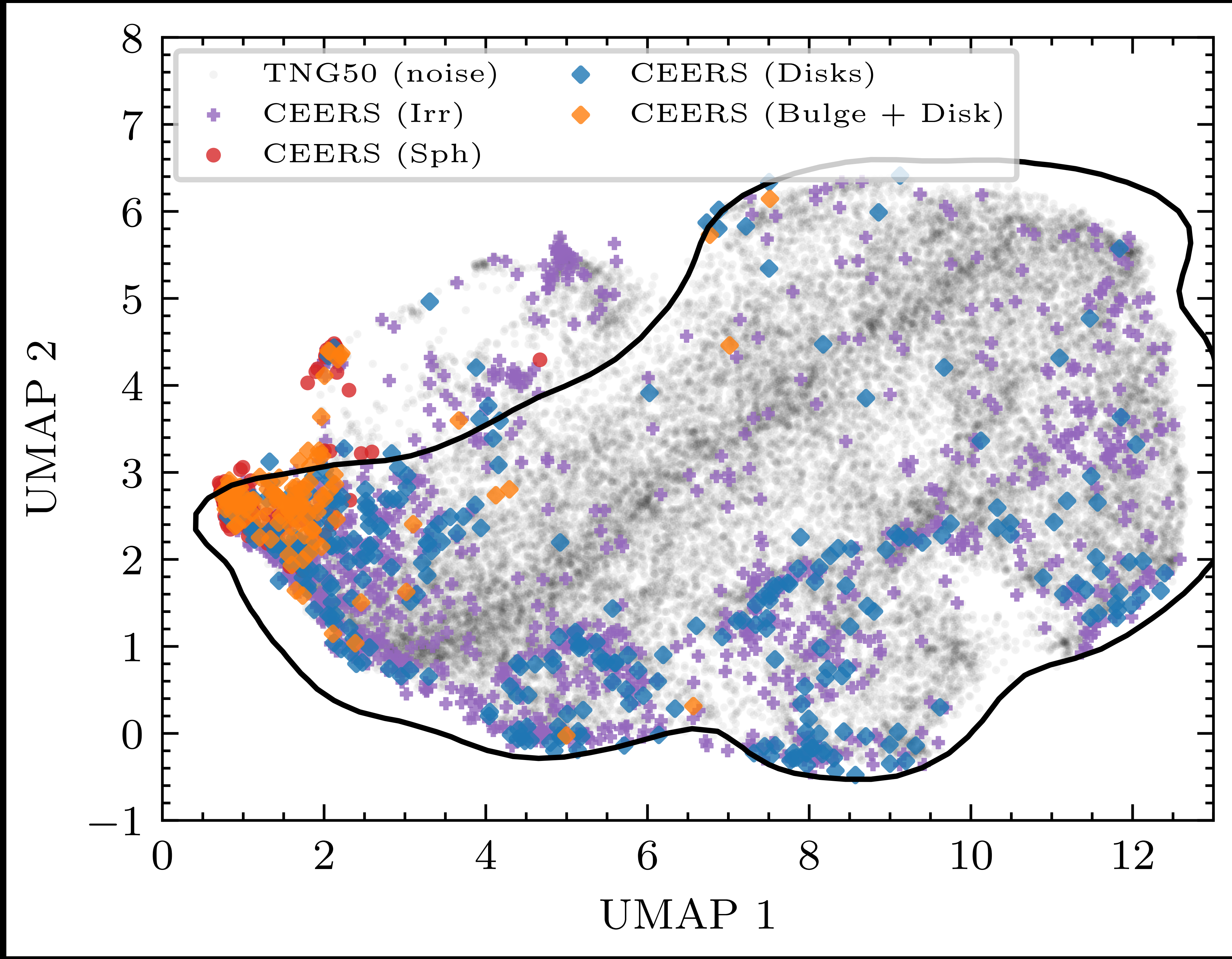
JWST/CEERS IMAGES

SEE TALK AT 15:00

MHC+2023b
CNN-based
Domain adaptation from CANDELS labels



Kartaltepe+2022
Visual classifications



Vega-Ferrero+2023

OUTLINE

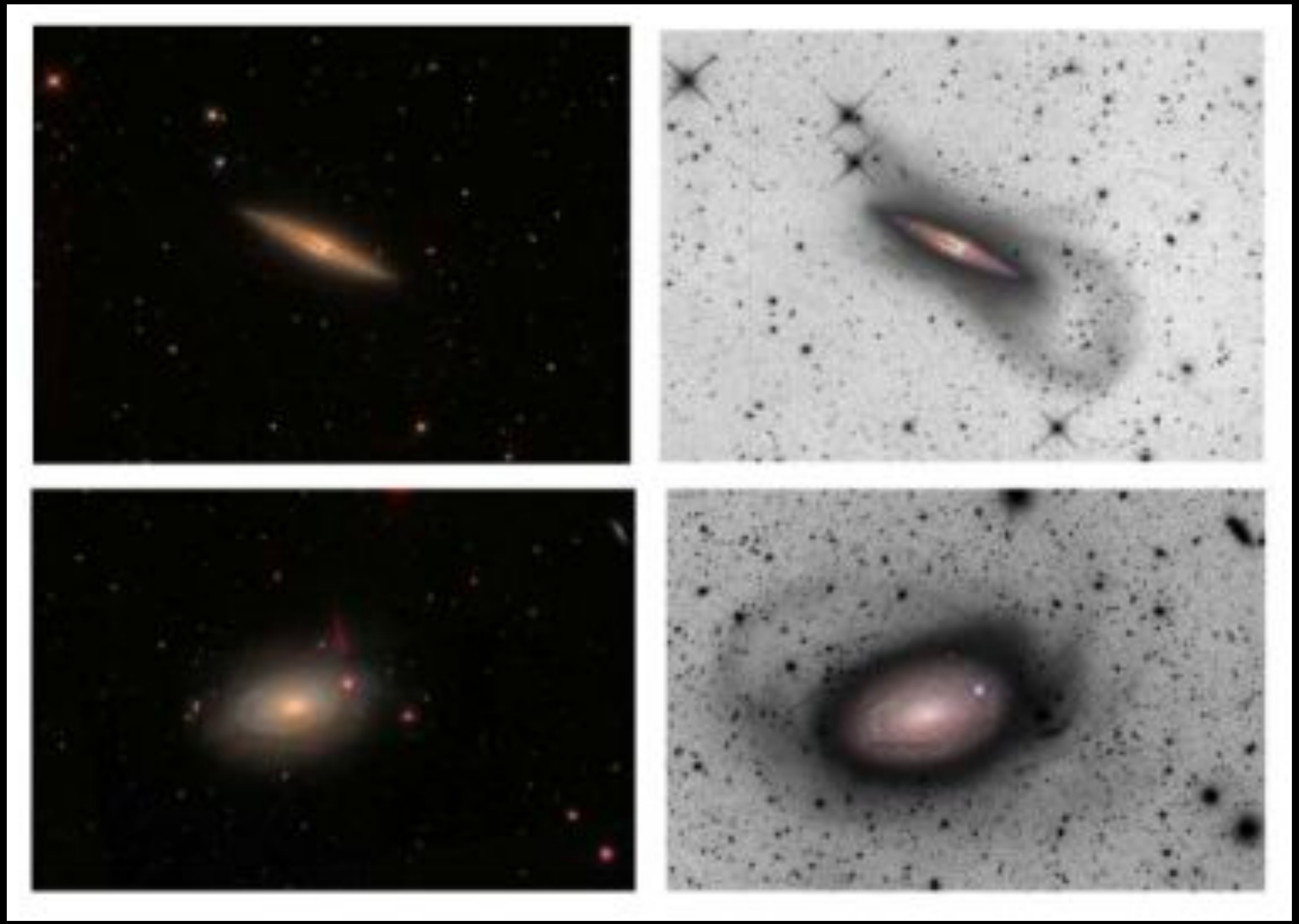
- Why care about galaxy morphology?
- Why Deep Learning?
- Supervised learning: CNNs
- What if there are no labels?
  - Transfer Learning
  - 'Emulations'
  - Unsupervised learning
- Going deeper

DILEMA DEEP IMAGING AND DEEP LEARNING FOR GALAXY MASS ASSEMBLY

- The frequency and characteristics of low surface brightness features can be used to disentangle the different formation channels (in situ star formation versus accreted stars).
- Classical 'cat/dog' problem: optimal for CNN but…
- Tidal streams are difficult to detect: short lived (only a few dynamical periods) and faint (imprints in the outskirts of galaxies)
- Need deep imaging and large areas!

SDSS

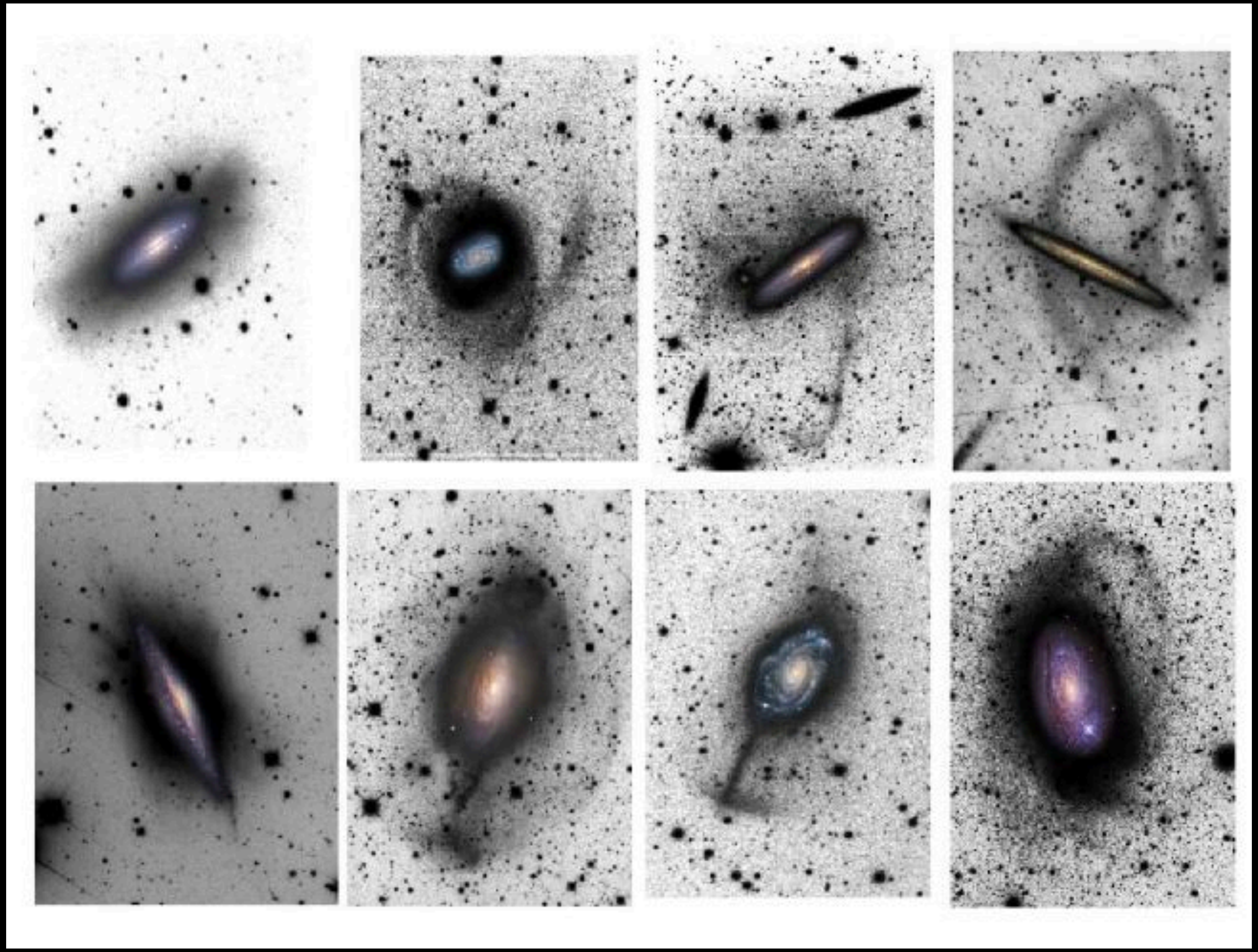
Martínez Delgado+2009



Chonis+2011

DILEMA DEEP IMAGING AND DEEP LEARNING FOR GALAXY MASS ASSEMBLY

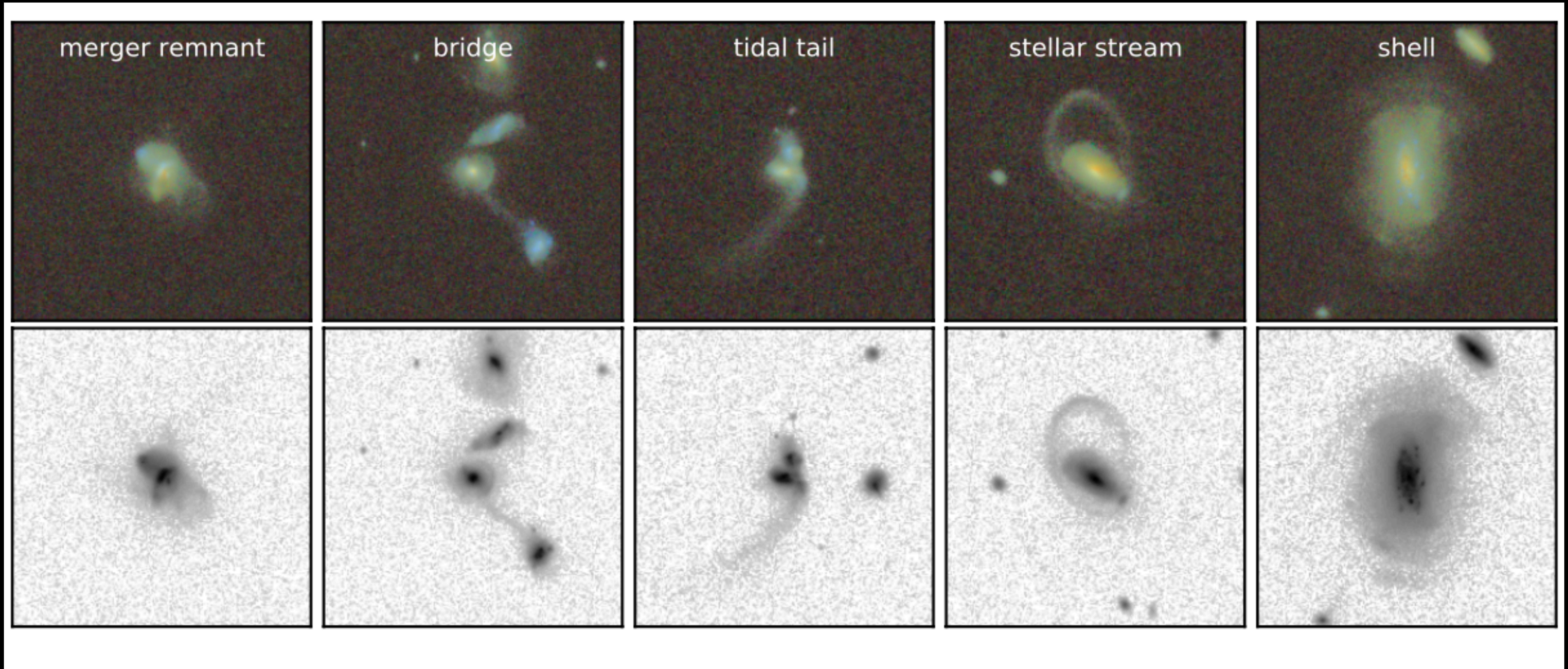
Martínez Delgado+2008, 10



- Need deep imaging and large areas!
- Small training samples available
- On-going efforts (e.g. Sola+22)
- Pioneer attempt by Walmsley+19
  - CFHTLS-Wide Survey
  - 1316 galaxies, 305 labelled as tidal streams
  - 76% accuracy

DILEMA DEEP IMAGING AND DEEP LEARNING FOR GALAXY MASS ASSEMBLY

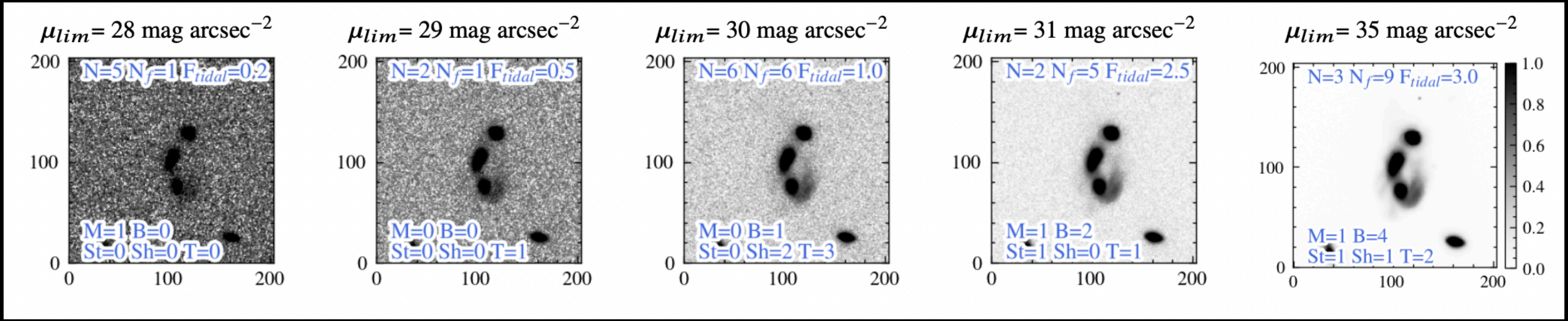
Images Classified by professional astronomers into different types of features



DILEMA DEEP IMAGING AND DEEP LEARNING FOR GALAXY MASS ASSEMBLY

Labelling data:
Binary classification
$F_{Tidal}=N_{features}/N_{classifiers}$

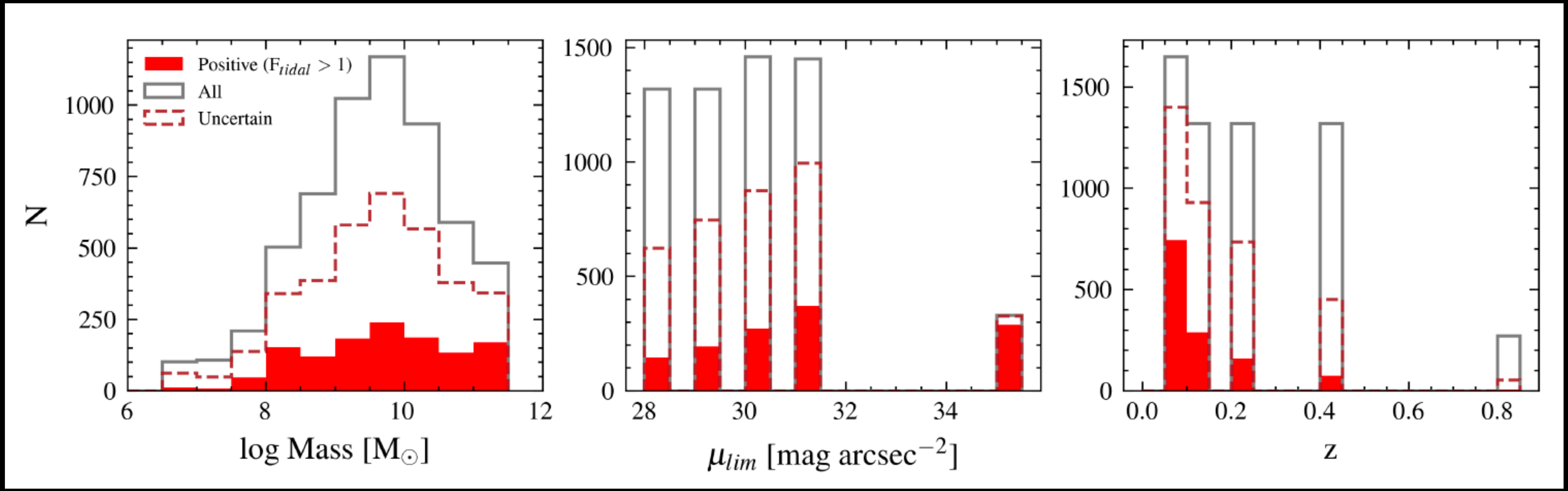
Negatives: $F_{Tidal} = 0$
Positives: $F_{Tidal} >= 1$
Uncertain: $F_{Tidal} = (0, 1)$

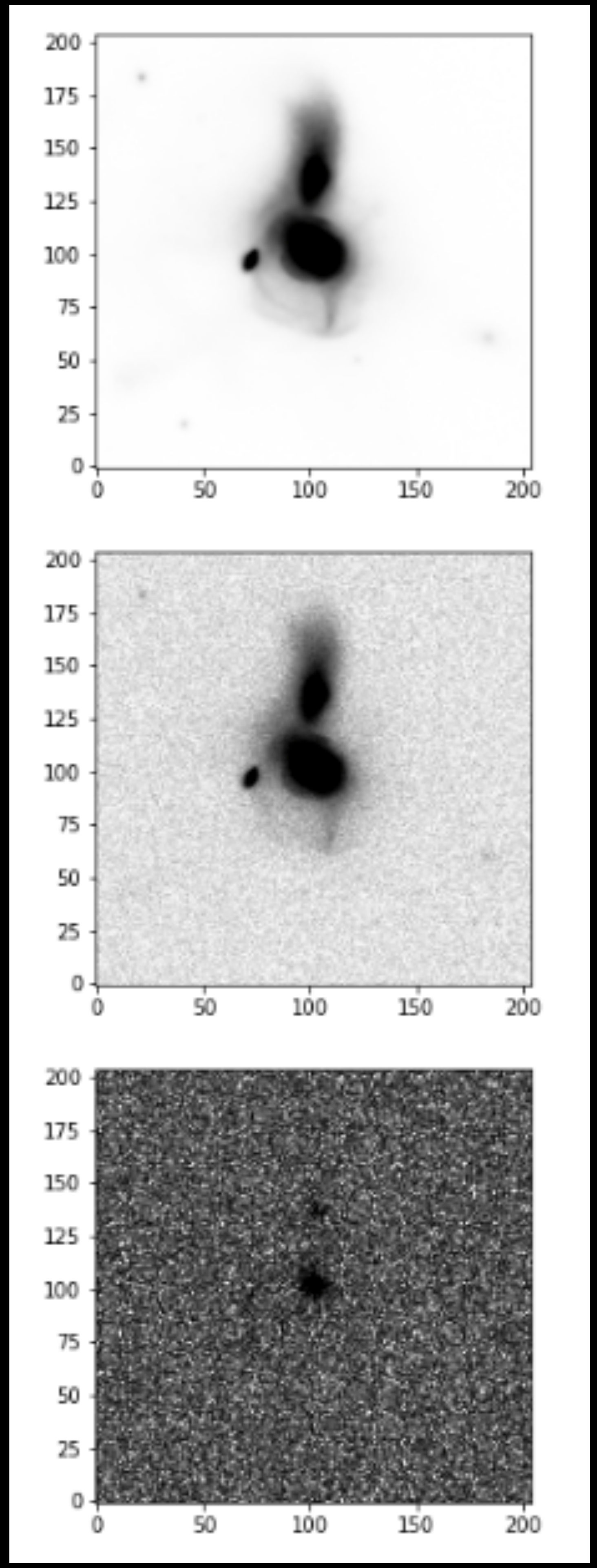


Domínguez Sánchez+2023

DILEMA DEEP IMAGING AND DEEP LEARNING FOR GALAXY MASS ASSEMBLY

Strong dependence of detections with image properties, specially surface brightness and redshift (stamp size)

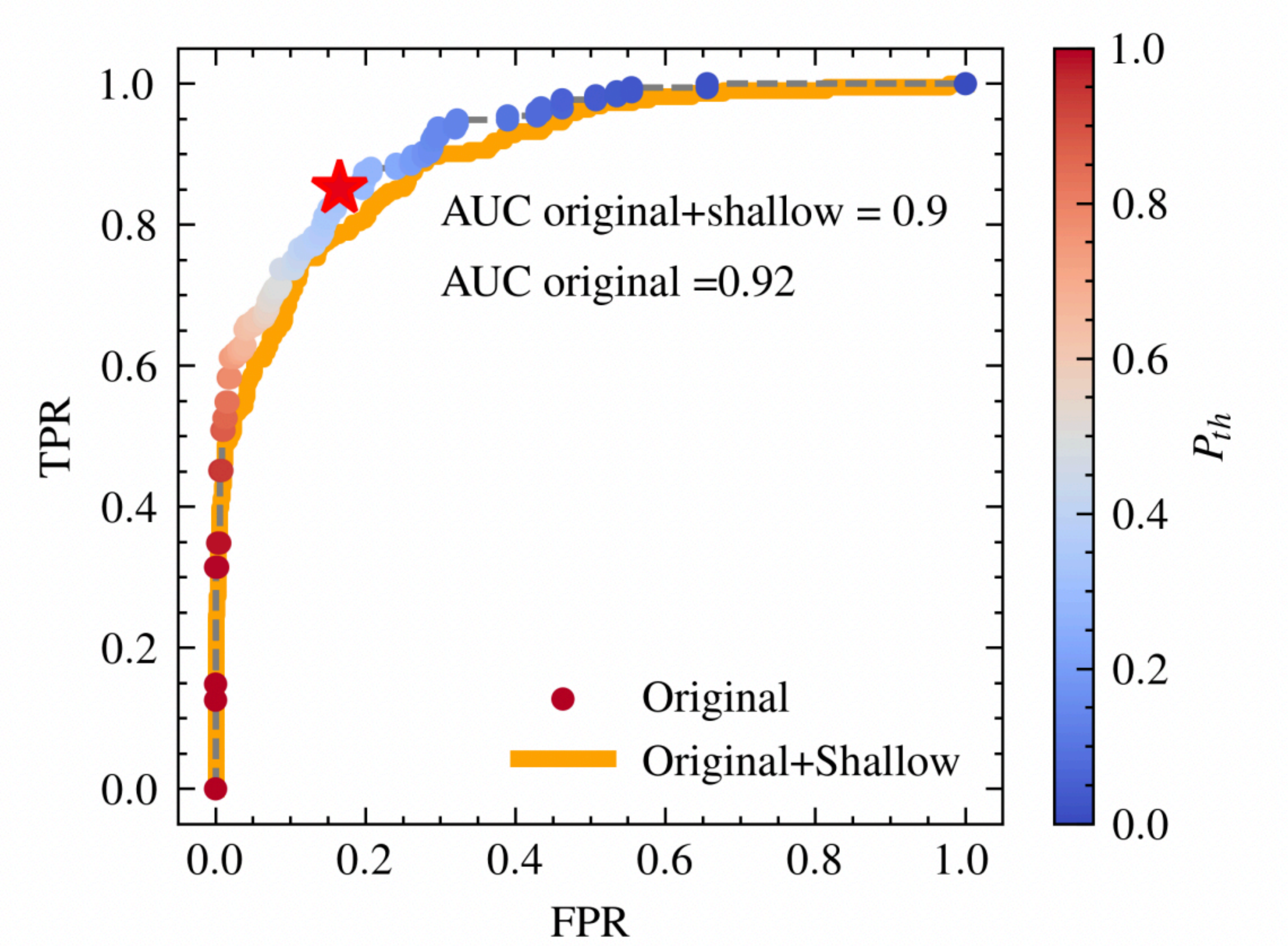


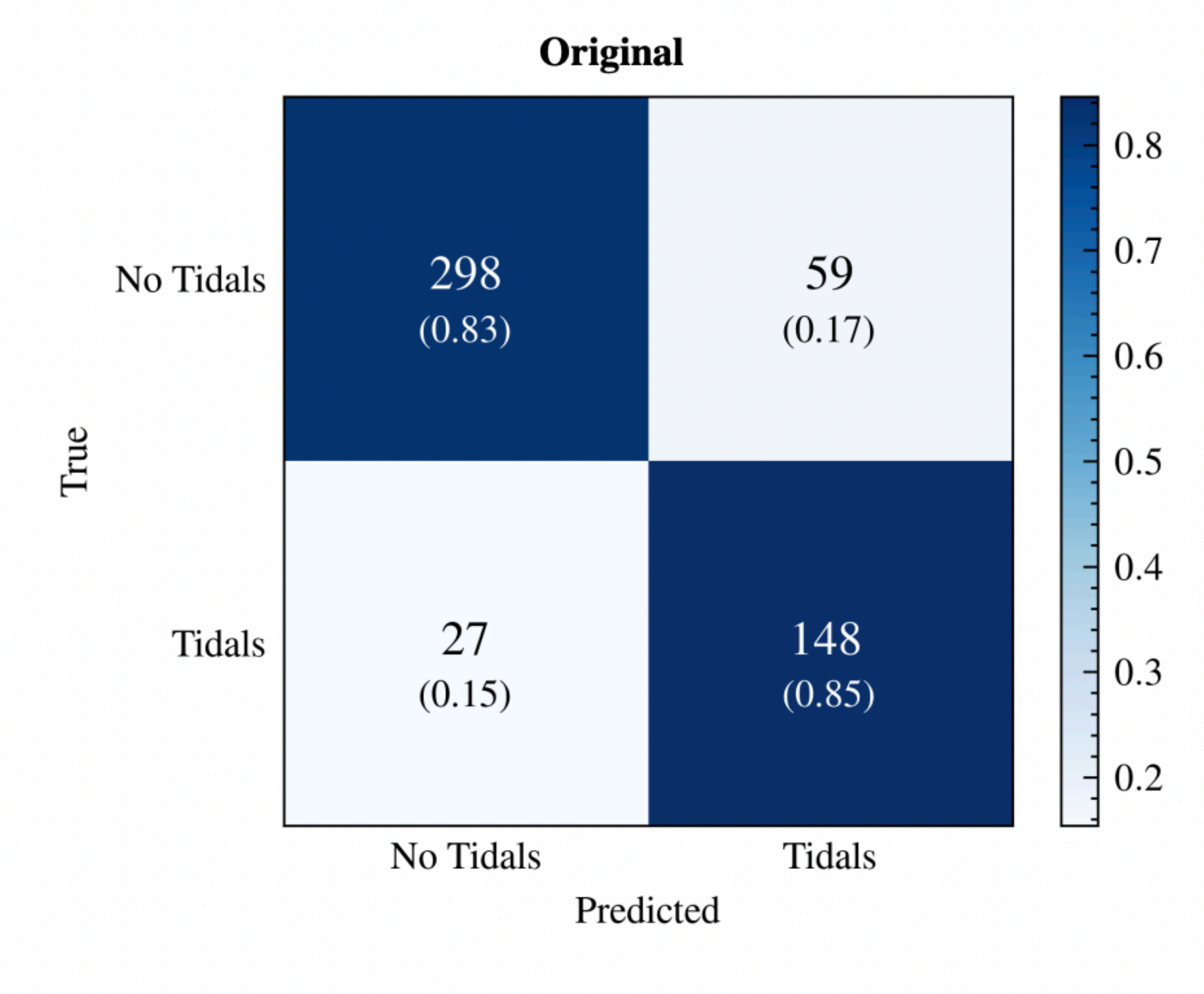


DS+2023

DILEMA DEEP IMAGING AND DEEP LEARNING FOR GALAXY MASS ASSEMBLY

Test sample	N_{test}	% Positives	P_{th}	Accuracy	Precision	Recall	F1
Original	532	33	0.32	0.84	0.72	0.85	0.78
Original+shallow	820	27	0.31	0.85	0.71	0.75	0.73

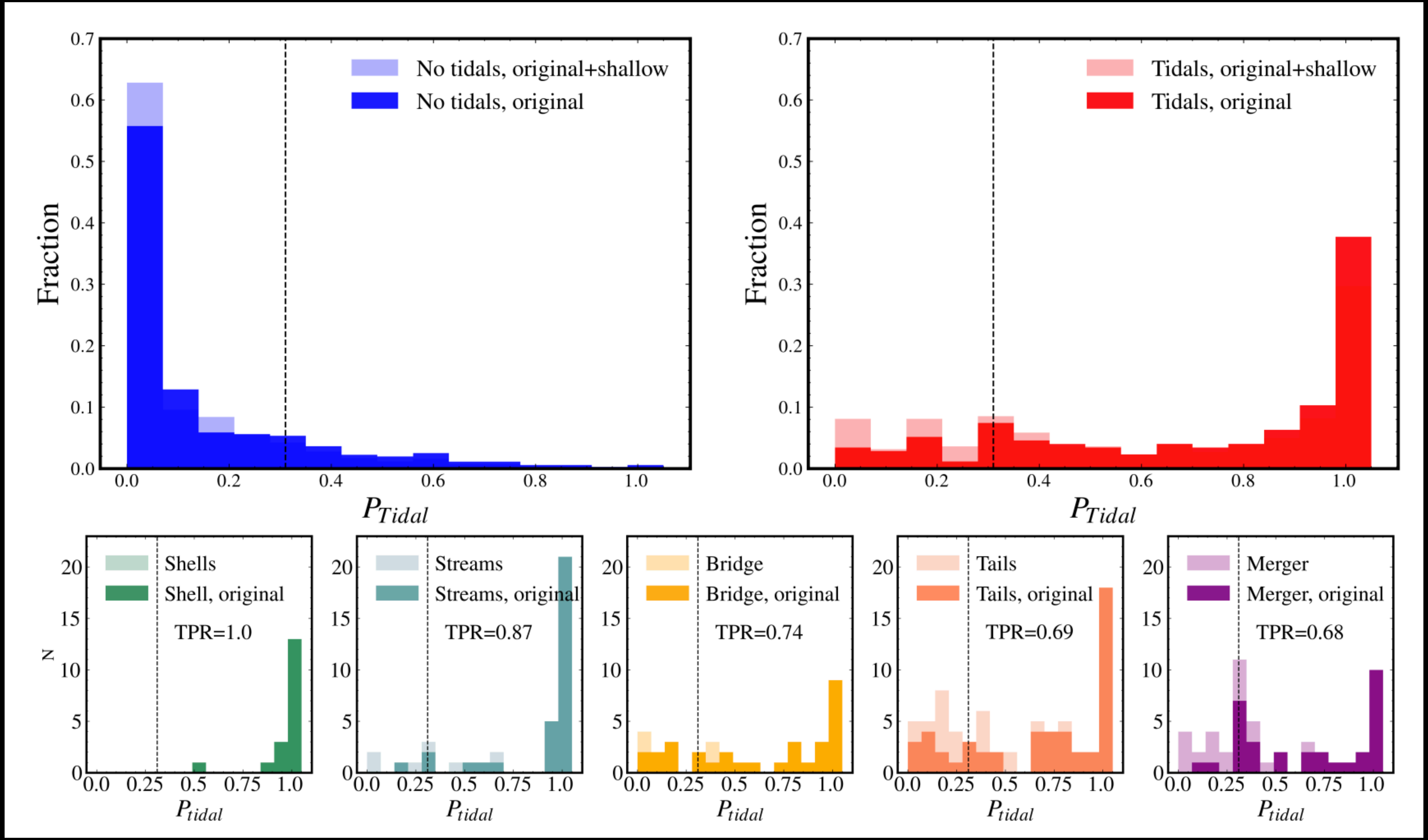




DS+2023

DILEMA DEEP IMAGING AND DEEP LEARNING FOR GALAXY MASS ASSEMBLY

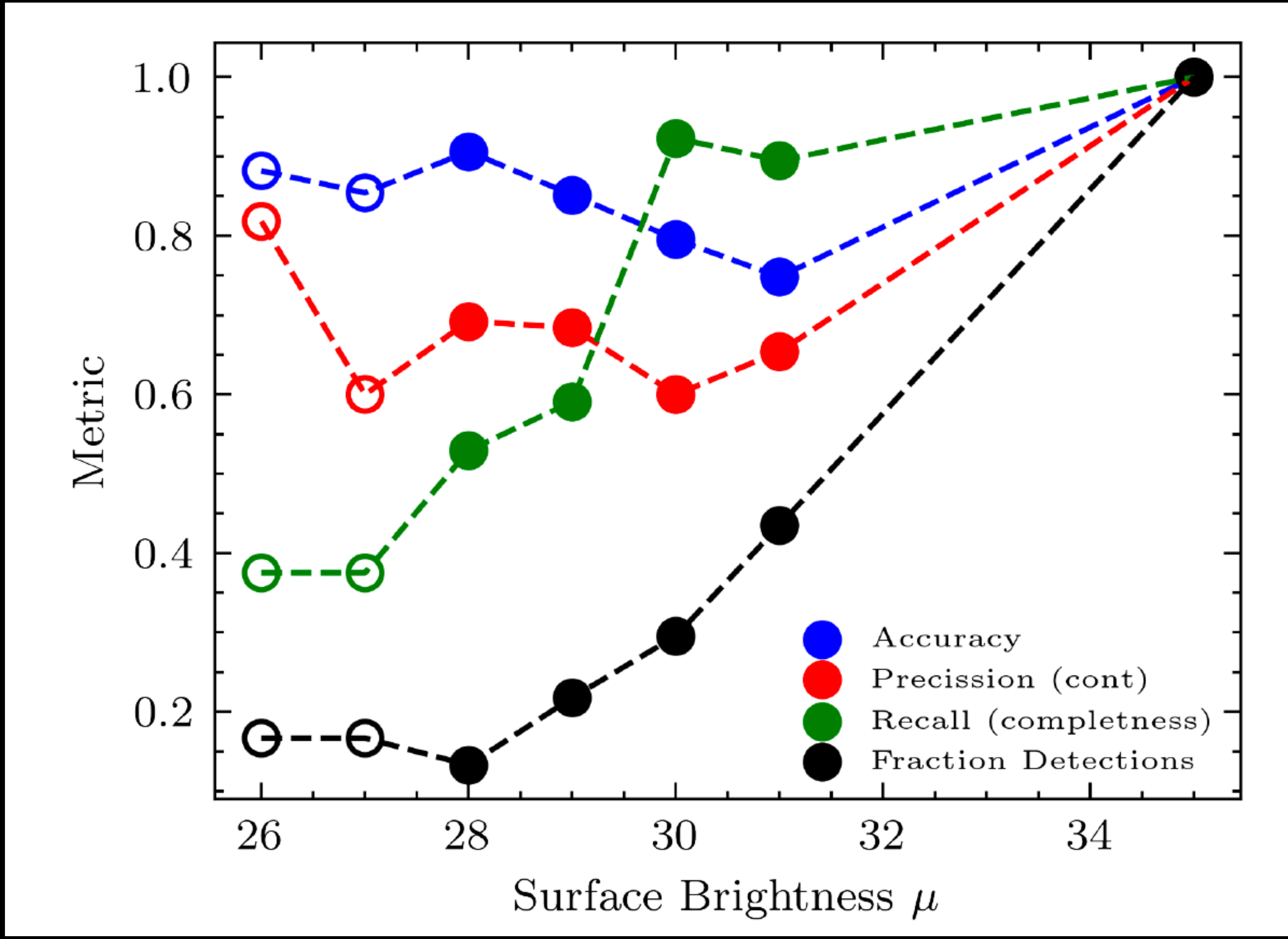
Strong dependence of the performance with different types of tidal features



DS+2023

DILEMA **D**EEP **I**MAGING AND DEEP **LE**ARNING FOR GALAXY **M**ASS **A**SSEMBLY

- Completeness strongly affected by LSB
- Purity and accuracy more stable



DS+2023

WHAT ABOUT REAL DATA?

Performance on real HSC data is worst (AUC=0.64), even with transfer learning techniques

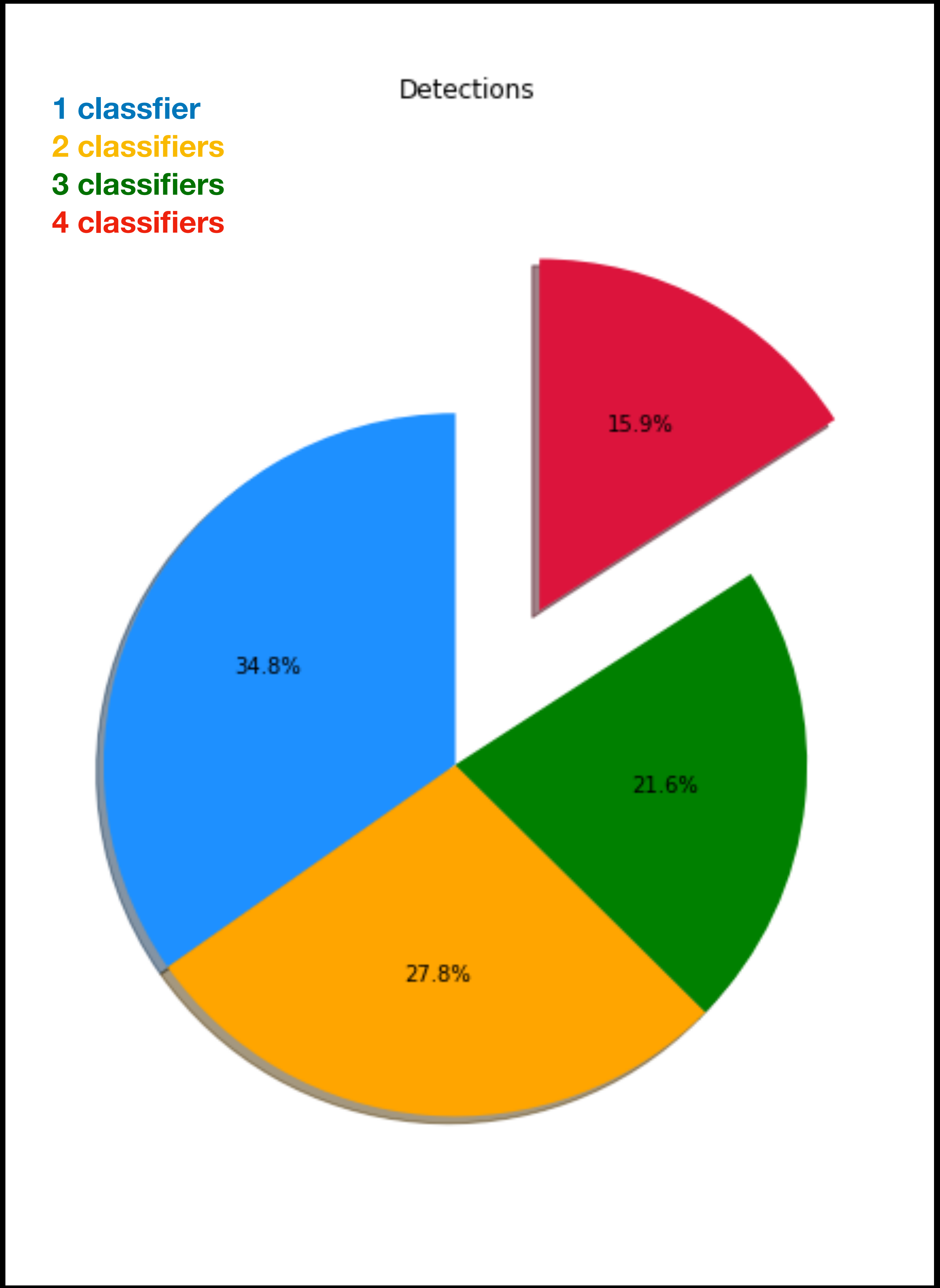
Why?

- Classifications are very uncertain
- Data is much shallower (μ=26 mag arcsec^{-2})
- Simulations are not completely 'realistic':

  Angular resolution (1″ vs 0.167″)

  No background sources

  No artefacts or sky subtraction residuals



ARRAKIHS

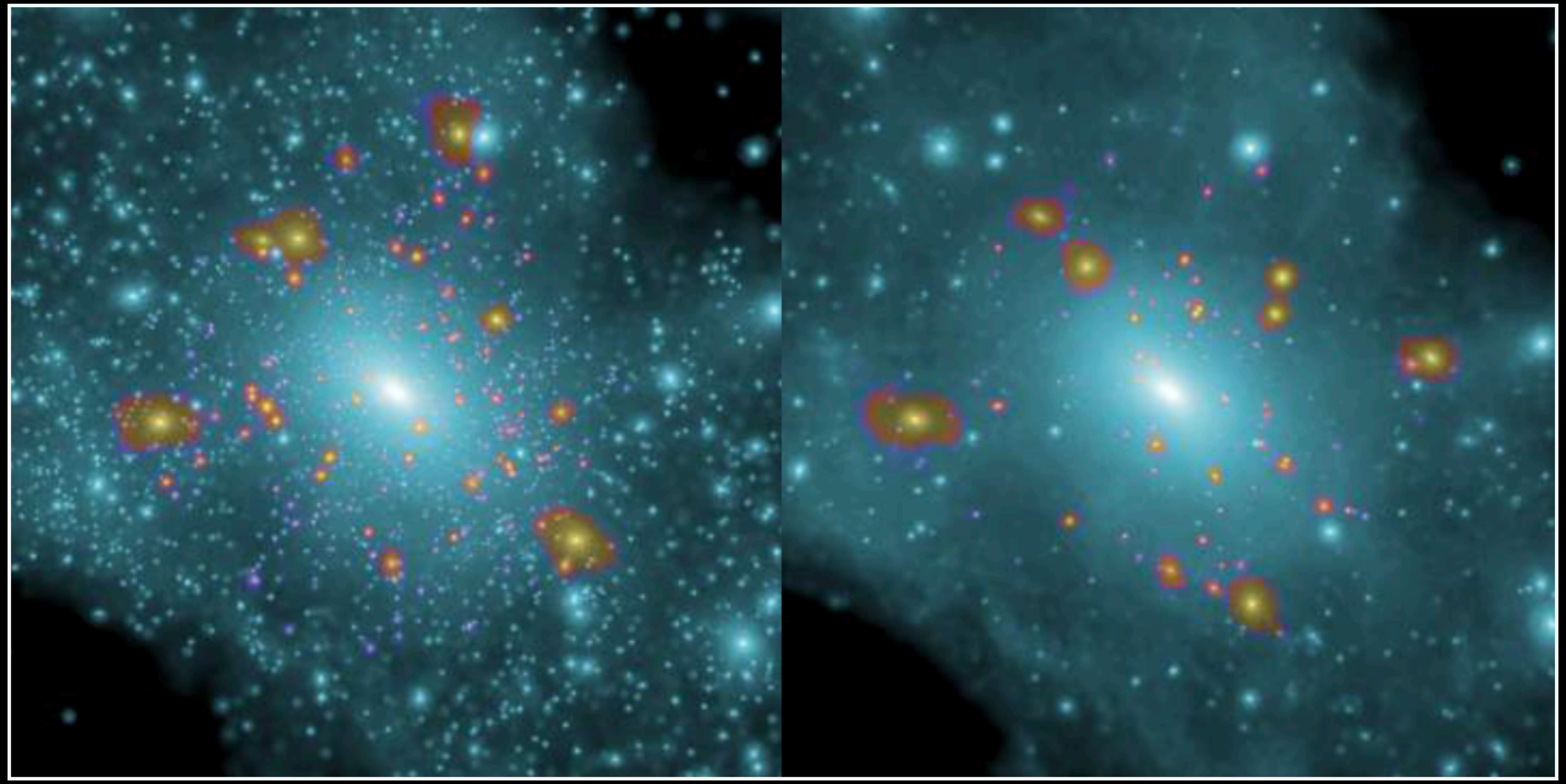
ANALYSIS OF RESOLVED REMNANTS OF ACCRETED GALAXIES AS A KEY INSTRUMENT FOR HALOS SURVEYS

ESA F-Mission (approved Dec 2022)
International consortia lead by Spain (P.I. R. Guzmán)

Ultra deep images (μ=31 mag arcsec^{-2}) arcsec of ~100 Milky Way-like galaxies in 2 visible and 2 IR bands with FWHM=0.8 to study:

- the nature of Dark Matter and ΛCDM
- satellite galaxies beyond the local group
- stellar streams and diffuse light

CDM WDM



ARRAKIHS

ARRAKIHS

ANALYSIS OF RESOLVED REMNANTS OF ACCRETED GALAXIES AS A KEY INSTRUMENT FOR HALOS SURVEYS

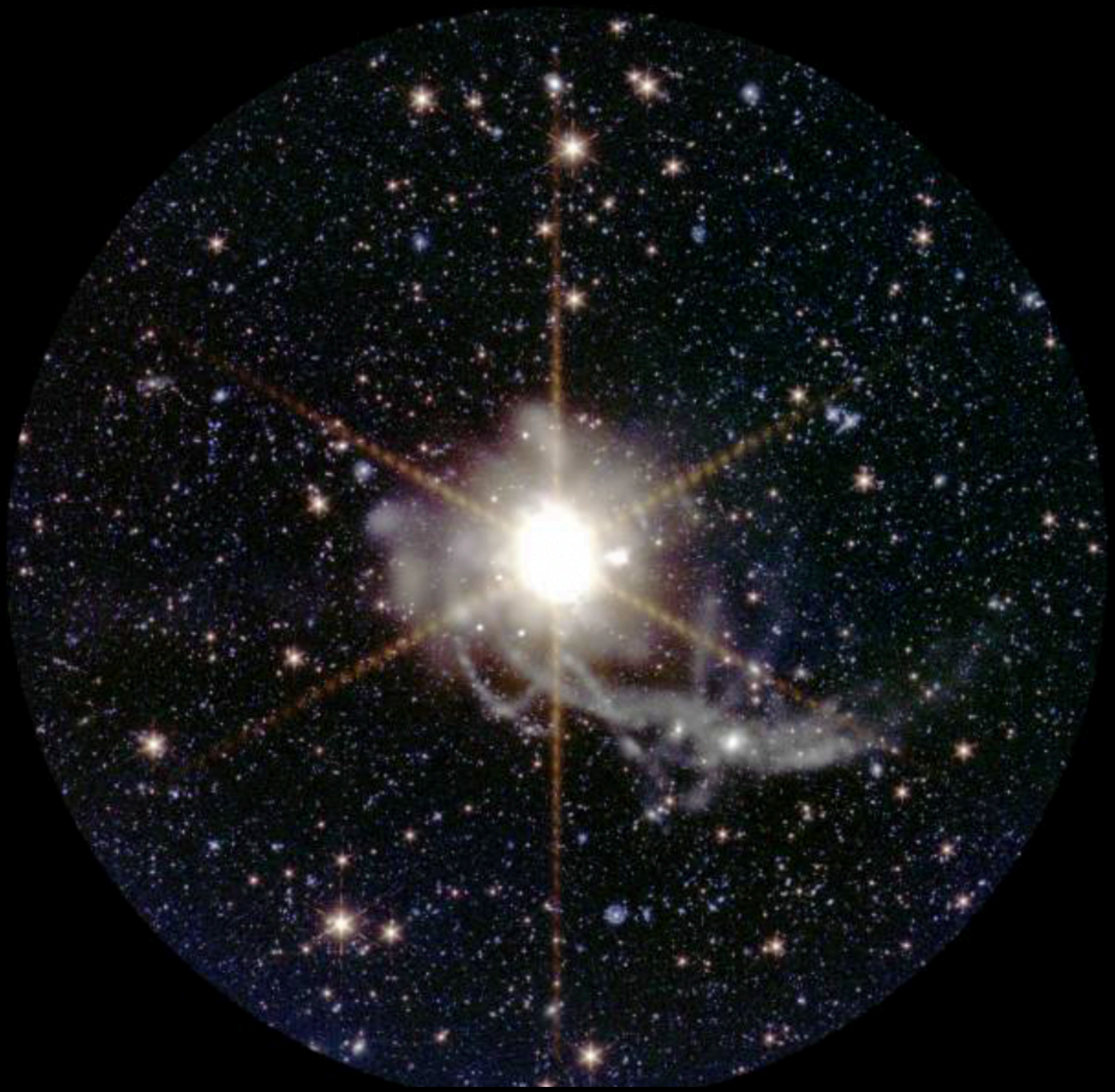


Figure 8: *The ARRAKIHS instrument with two iSIM170 binocular telescopes*

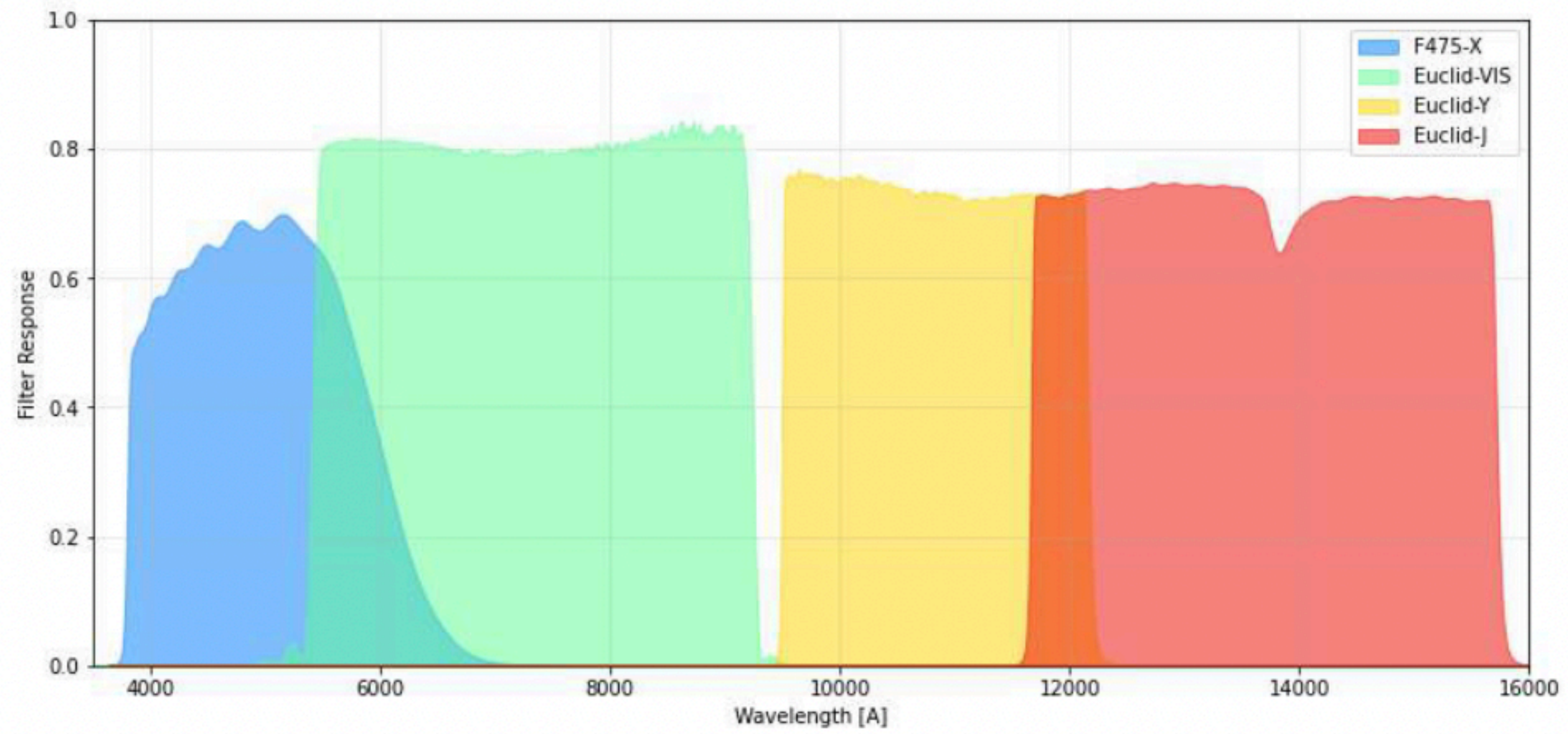


Figure 11: *ARRAKIHS filter system.*

ARRAKIHS

ANALYSIS OF RESOLVED REMNANTS OF ACCRETED GALAXIES AS A KEY INSTRUMENT FOR HALOS SURVEYS

I am the coordinator of WP S-10 Structure Analysis, in particular on ML techniques for the detection and characterisation of satellite galaxies and tidal features.

We have two open positions at CEFCA:

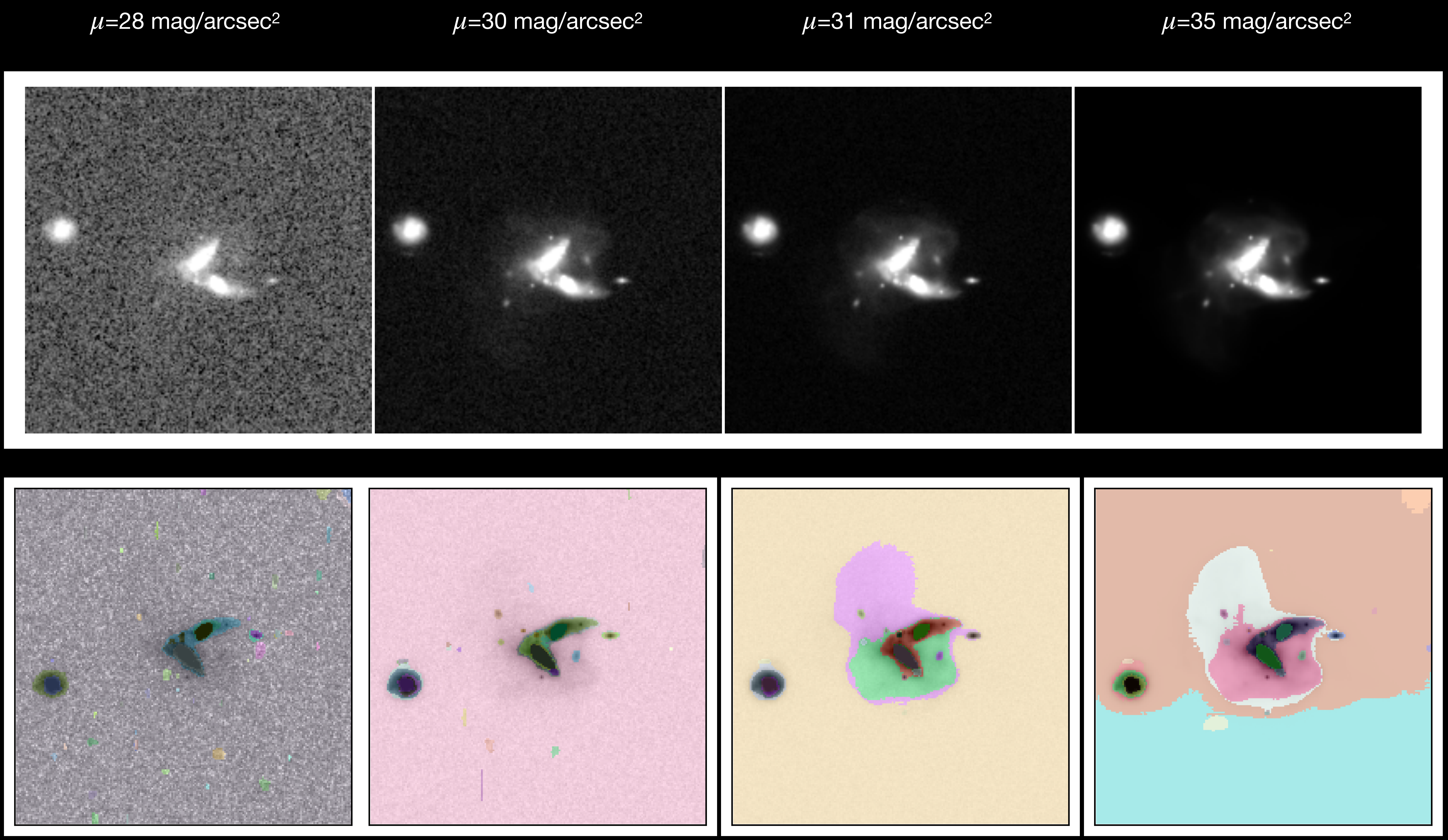
- Post-doc for ARRAKIHS
  https://www.cefca.es/cefca_en/reference_0118

- Scientific Software Engineer for ARRAKIHS
  https://www.cefca.es/cefca_en/reference_0119

SAM: Segment **Anything** Model

https://segment-anything.com/

SAM: Segment Anything Model



TAKE HOME MESSAGES

DL extremely successful for classifying galaxy images BUT:

- Need large training samples of labeled galaxies from the same domain (magnitude, instrument, etc.).
- Detecting faint/subtle features (tidal interactions) is still challenging.
- Transfer Learning reduces the necessary training sample by one order of magnitude (but still need some labels).
- 'Emulated' faint galaxies are a clever approach to have labelled samples at high-z.
- Unsupervised learning (self-supervised or PCA) unveils useful properties of galaxies w/o the need of labels.
- Do not blindly apply your models to data from different domain as the training sample !!