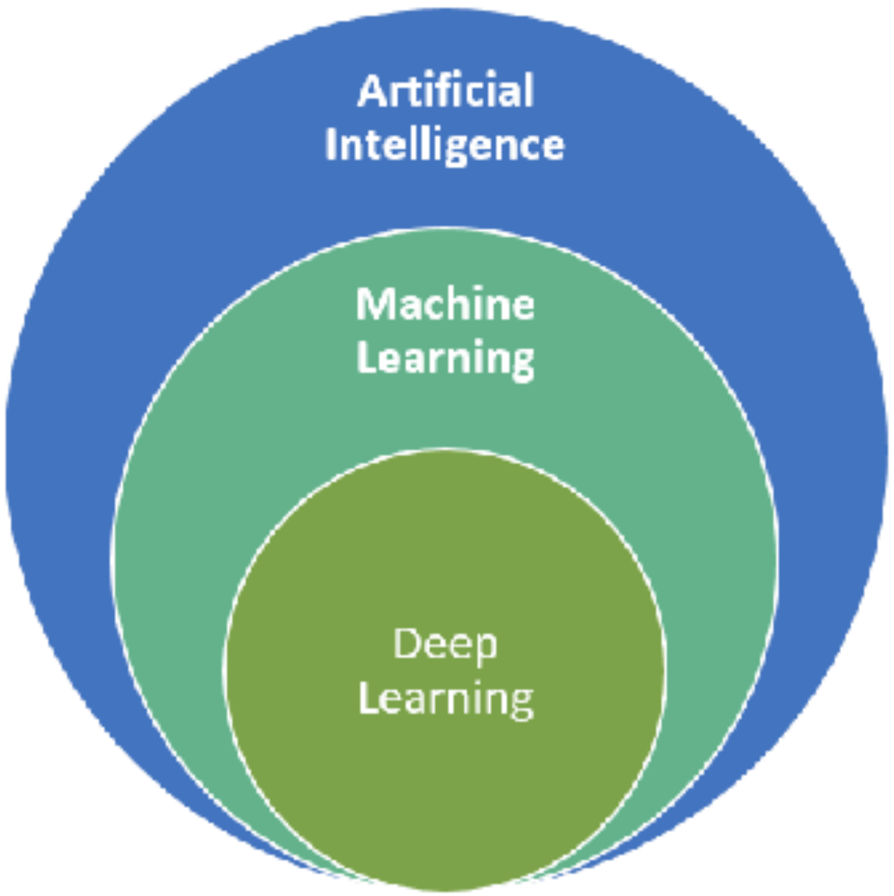
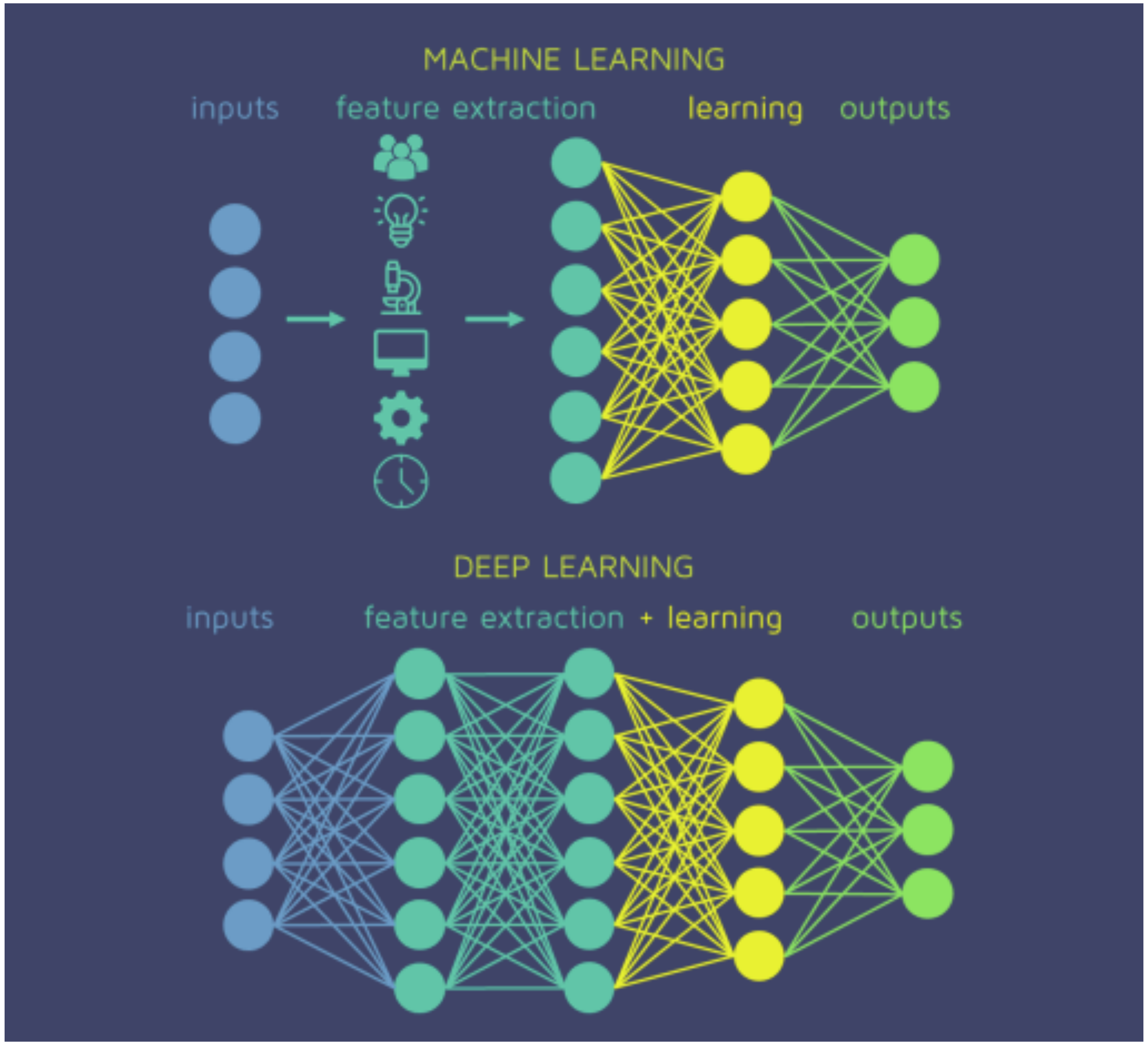
Astro Machine Learning (galaxy evolution biased)

Helena Domínguez Sánchez
Marc Huertas Company

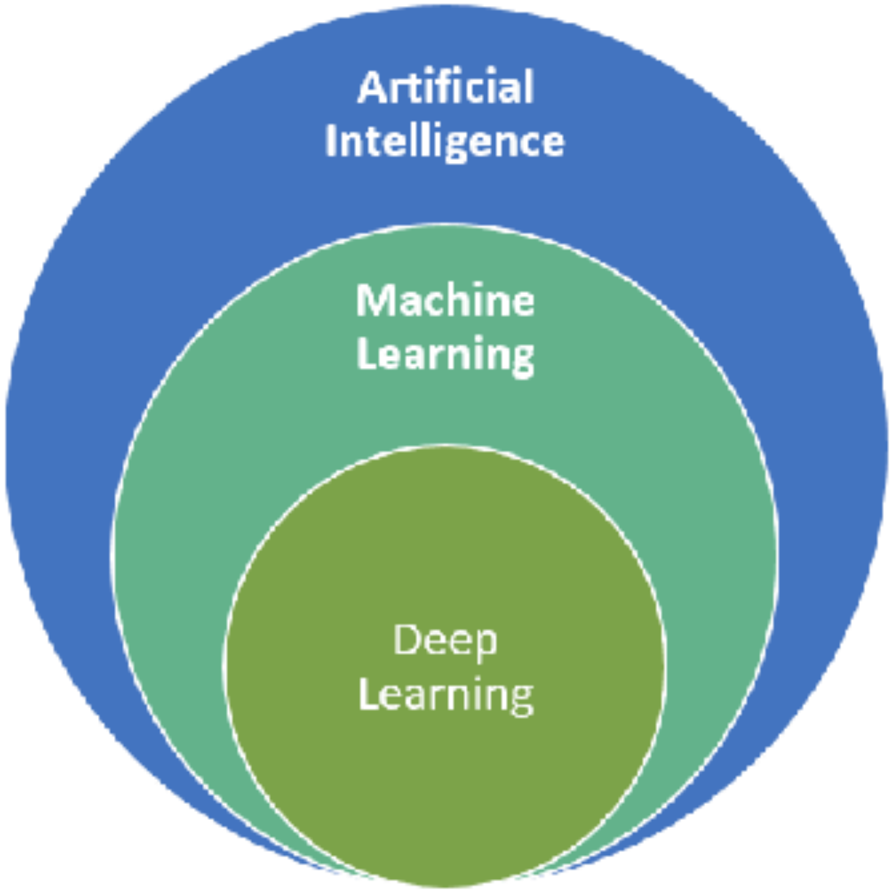
SEA 2020 - REUNIÓN CIENTÍFICA VIRTUAL

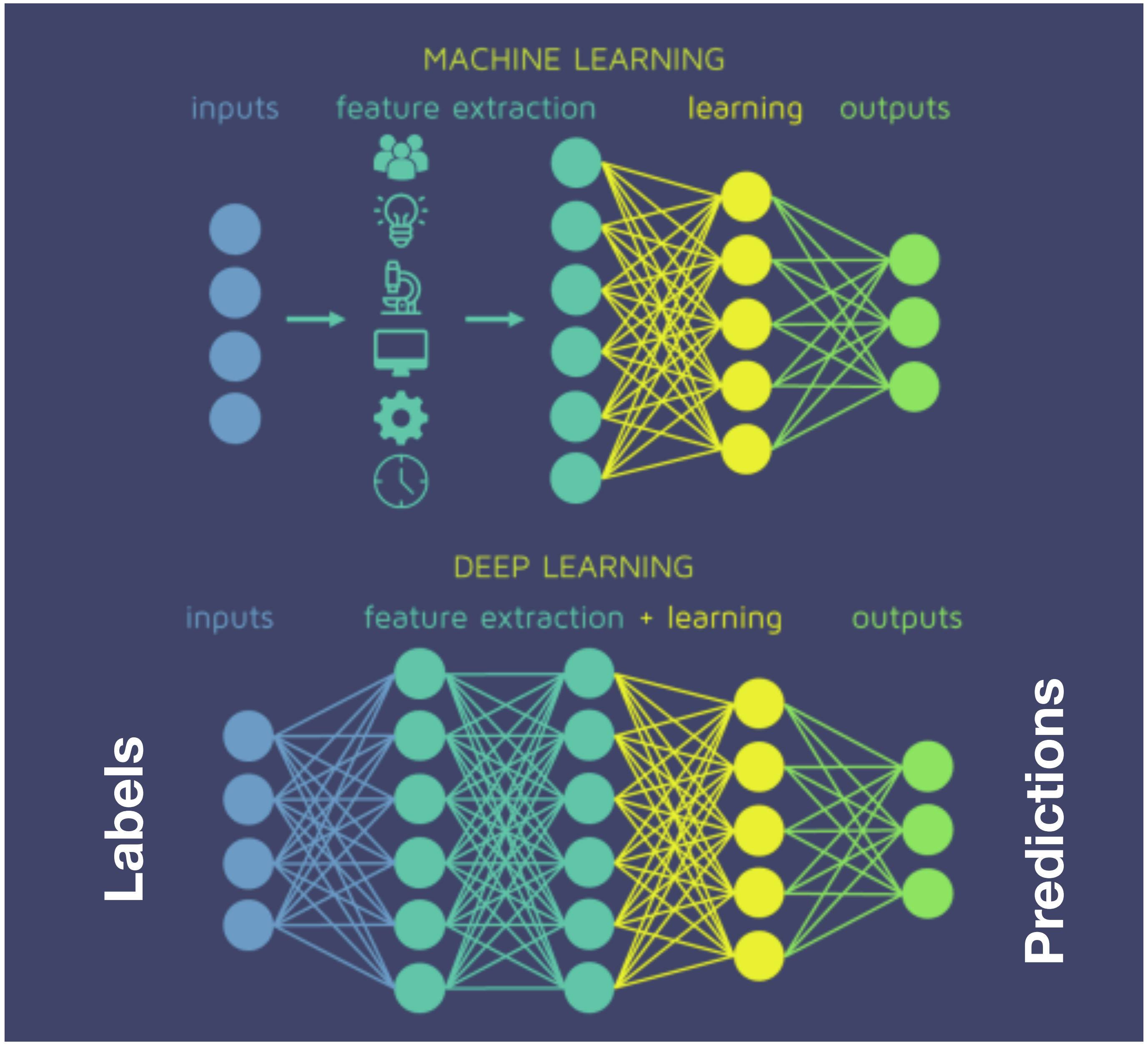
What is Deep Learning?



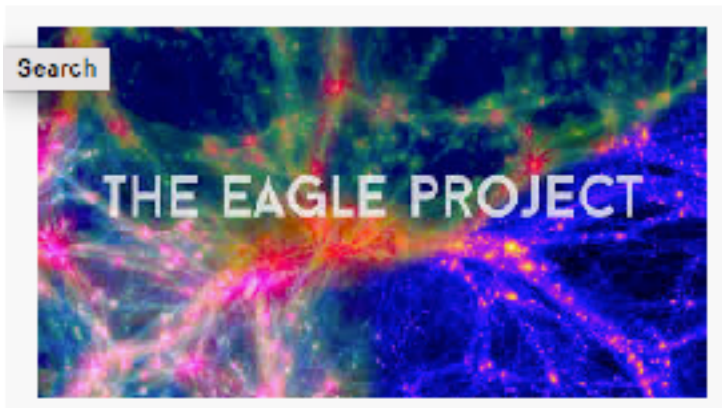


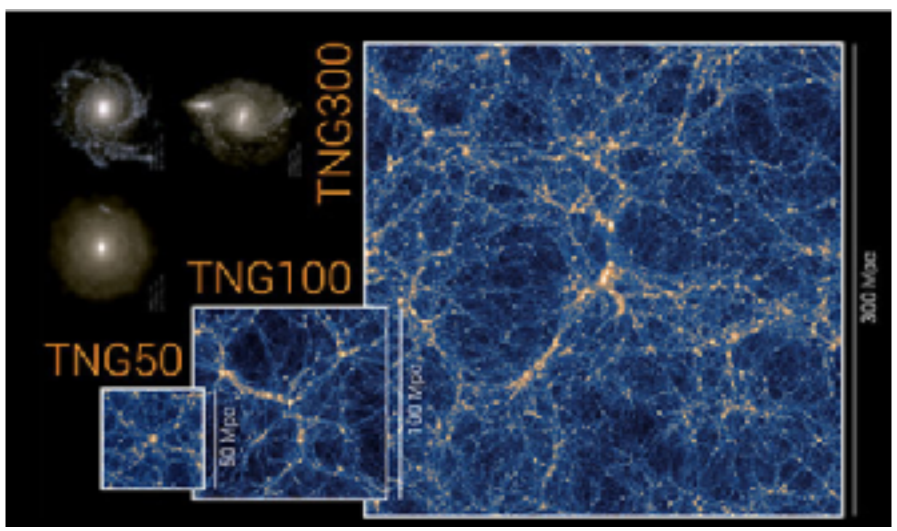
What is Supervised Deep Learning?





Why DL in astronomy?



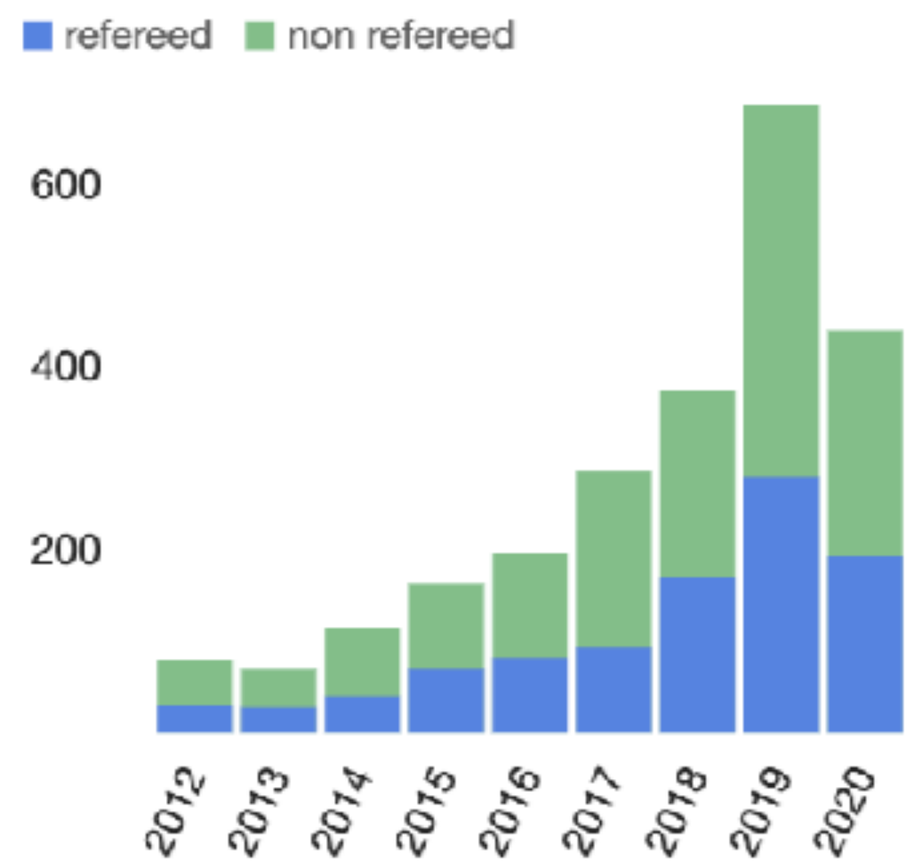


The Horizon Simulation

We don’t have the capacity to “look at” future astronomical datasets (BIG DATA)!

DL and astronomy

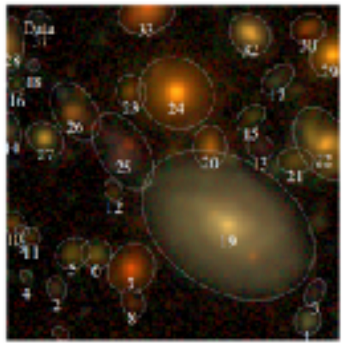
- DL has been successfully applied to myriad of astronomical task.
- DL has matured a lot since first papers (~2015).
- DL is starting to be normalised as a methodology tool.

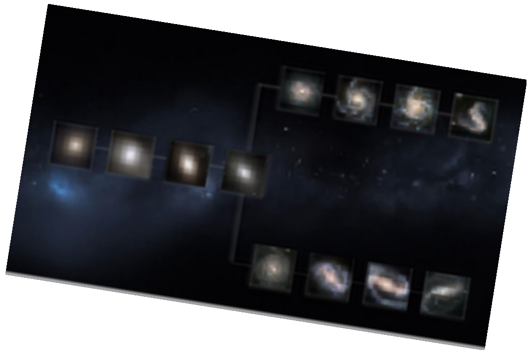


Applications in astronomy

✓**Object detection and segmentation**

- Deblending (e.g, Bocaud+19, Arcelin+2020)
- Source extraction (e.g., Hausen+19)



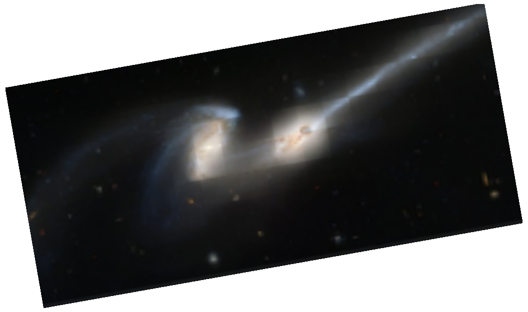


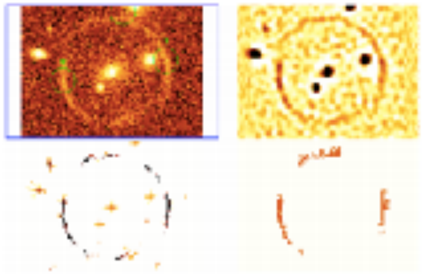
✓**Classification**

- Morphology (e.g., Dieleman+15, Huertas-Company+15, Domínguez Sánchez+18, Cheng+2020)
- Tidal streams (e.g., Walmsley+18), mergers (e.g., Bottrell+19, Snyder+19, Pearson+19), gravitational lenses (e.g., Petrillo+19, Metcalf+19, Jacobs+19, Cheng+20)

✓**Regression**

- Photo-z (e.g., Pasquet+18, Campagne+2020)
- Cluster Masses (e.g., Ho+20, Yan+20, Su+20)
- Galaxy morphometry (e.g., Tuccillo+18)





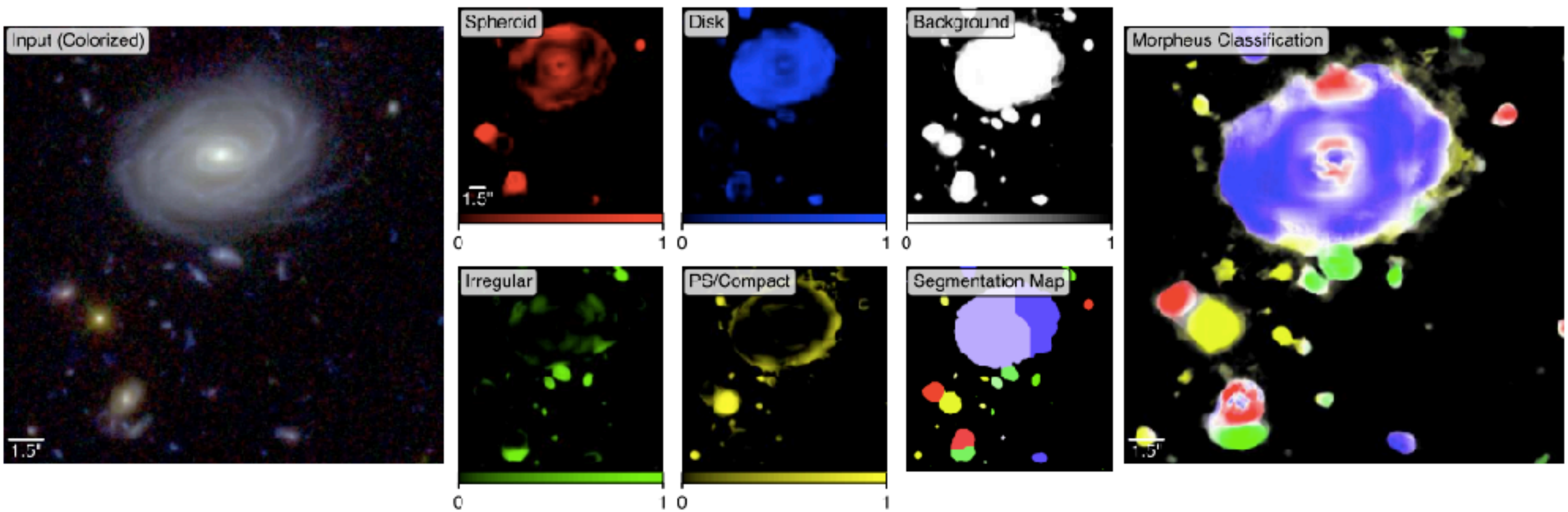
Galaxy Morphologies

Morpheus: A Deep Learning Framework For Pixel-Level Analysis of Astronomical Image Data

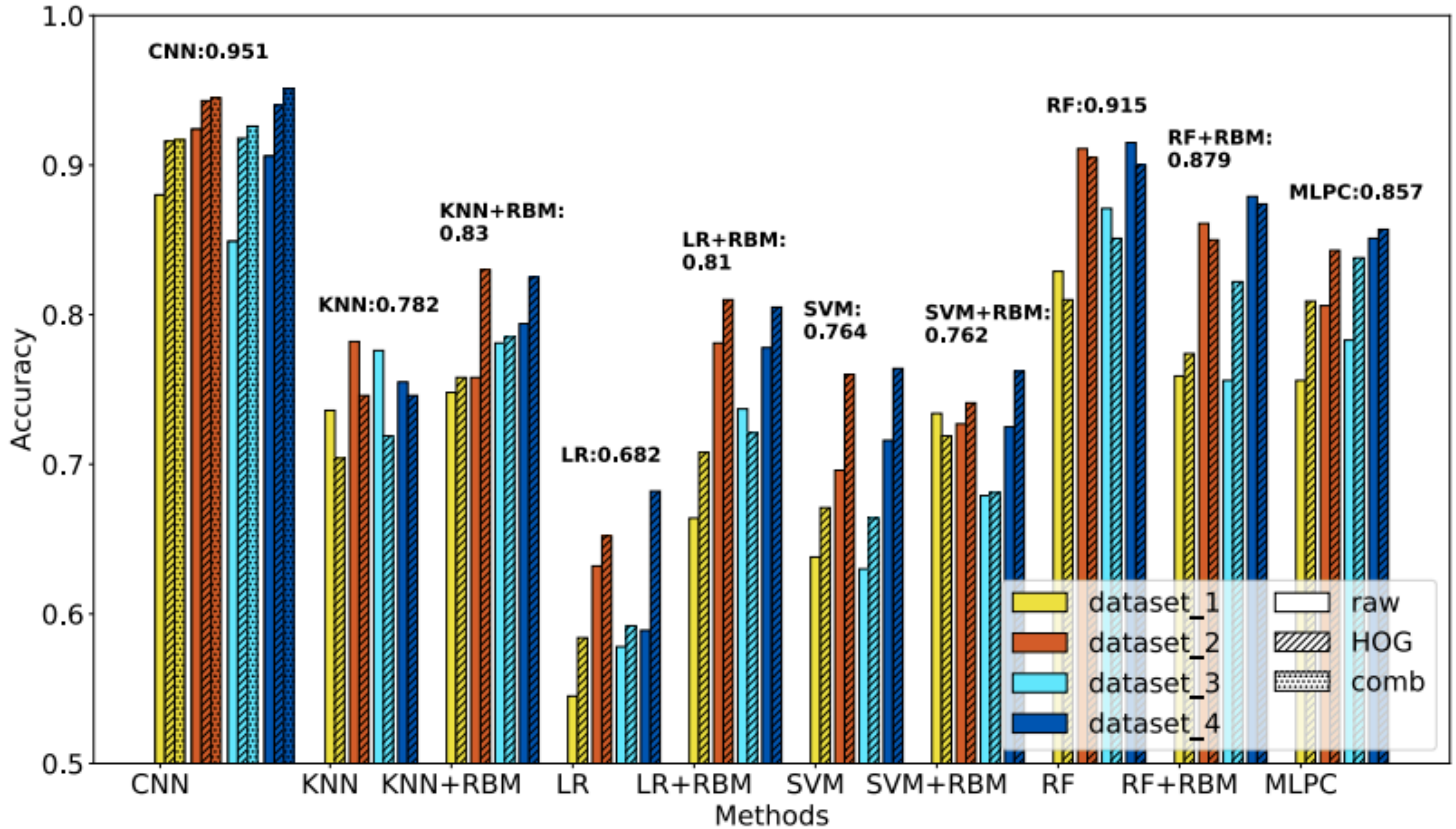
RYAN HAUSEN[1] AND BRANT E. ROBERTSON[2,3]

[1]Department of Computer Science and Engineering, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064 USA
[2]Department of Astronomy and Astrophysics, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064 USA
[3]Institute for Advanced Study, 1 Einstein Drive, Princeton, NJ 08540 USA

Object detection + segmentation

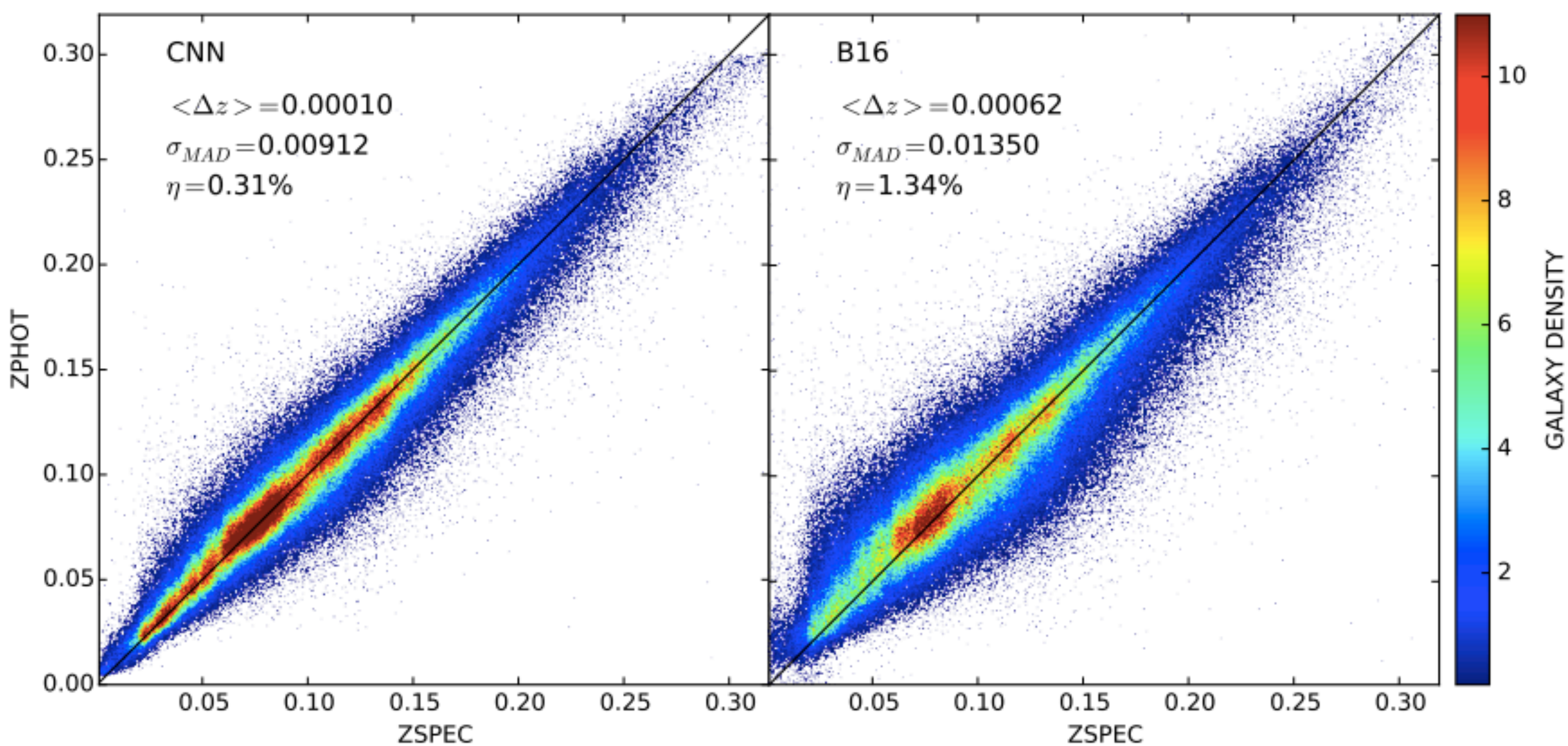


CNN are the most accurate



Morphological classification of galaxies (Cheng+2020)

CNN are the most accurate



Photometric Redshifts (Pasquet+18)

BUT…

- One of the main challenges to apply supervised DL is to have on hand a **large enough** set of pre-labelled images to be used as a **training sample** coming from the same domain (e.g., instrument, resolution, magnitude, etc.).

- How can we overcome this issue?

   A. Transfer learning

   B. Simulations

   C. Generative Networks

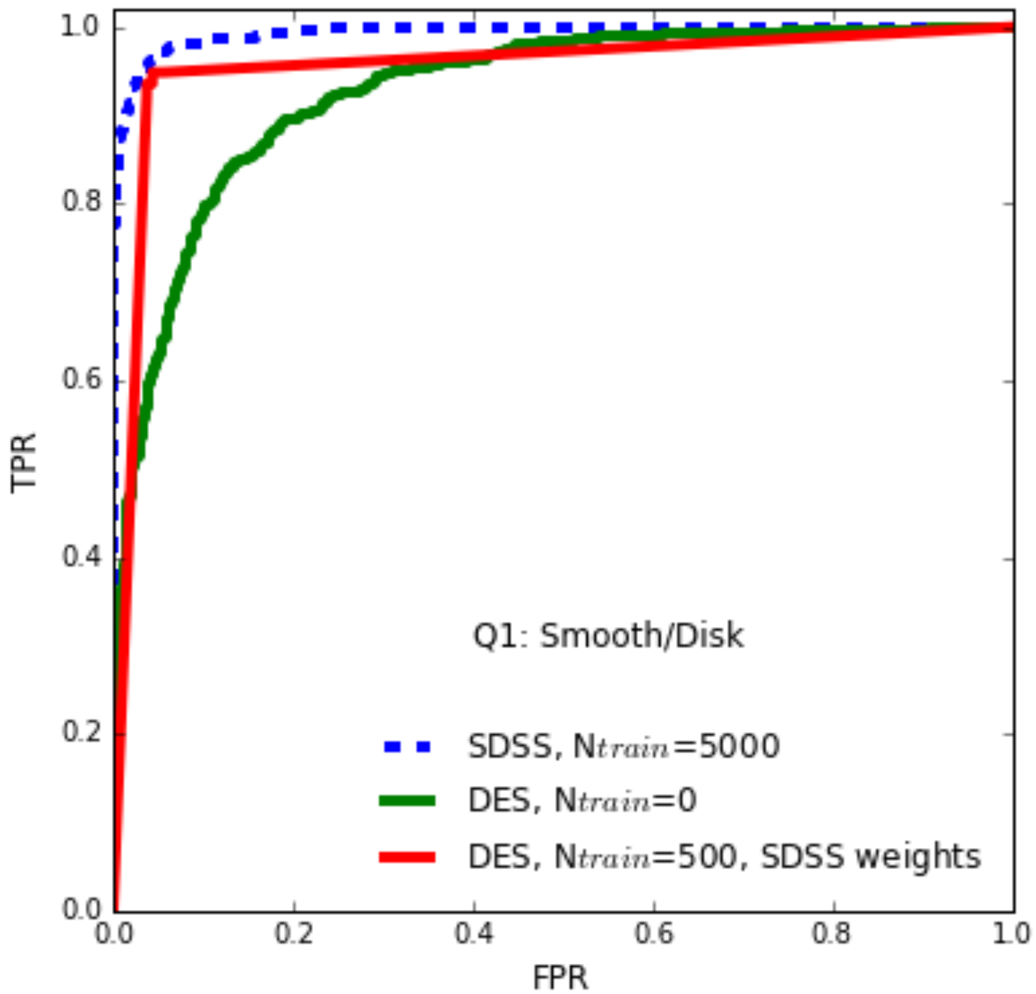
Transfer Learning

- Instrument characteristics (PSF, spatial resolution, depth) affect the result of DL models trained with images from different surveys.

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

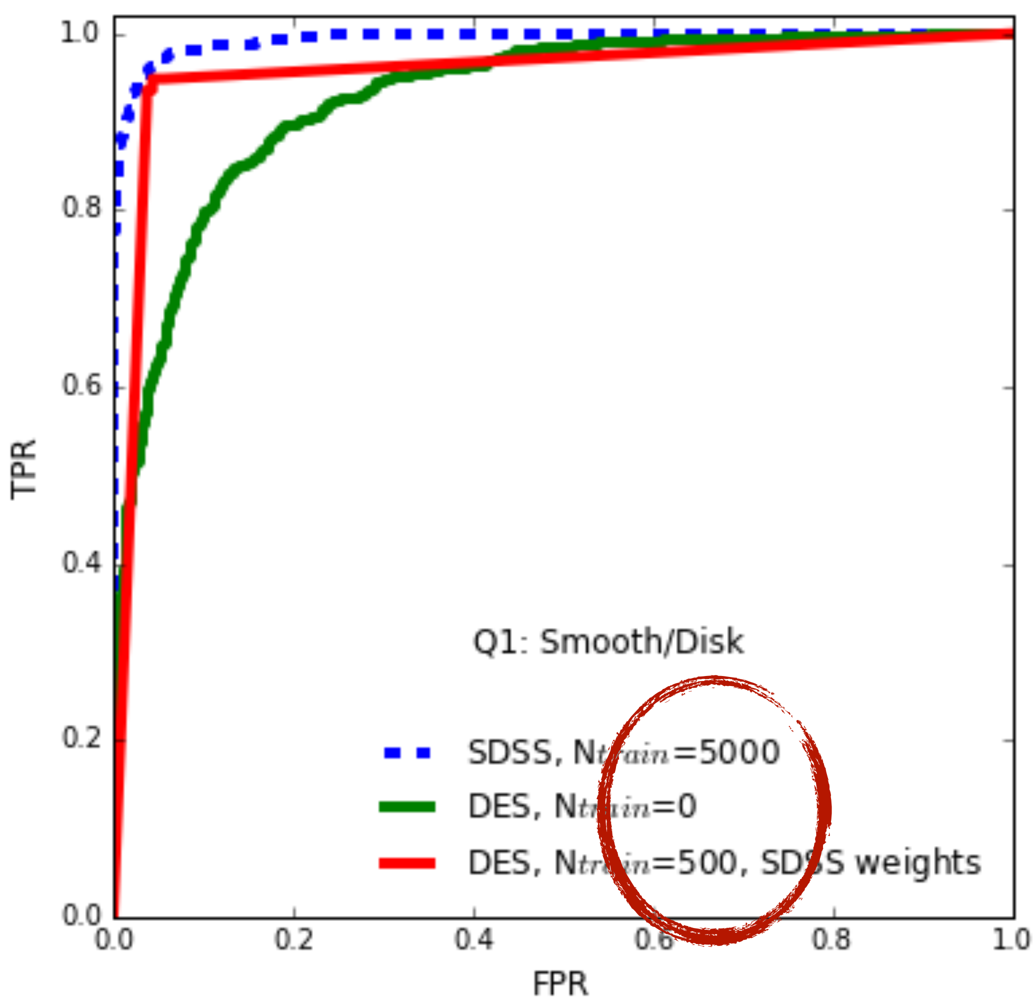
- Can Convolutional Neural Networks transfer knowledge between surveys?

Transfer learning: SDSS to DES



Domínguez Sánchez et al. (2019)

Transfer learning: SDSS to DES



Domínguez Sánchez et al. (2019)

- Can Convolutional Neural Networks transfer knowledge between surveys?

YES!

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

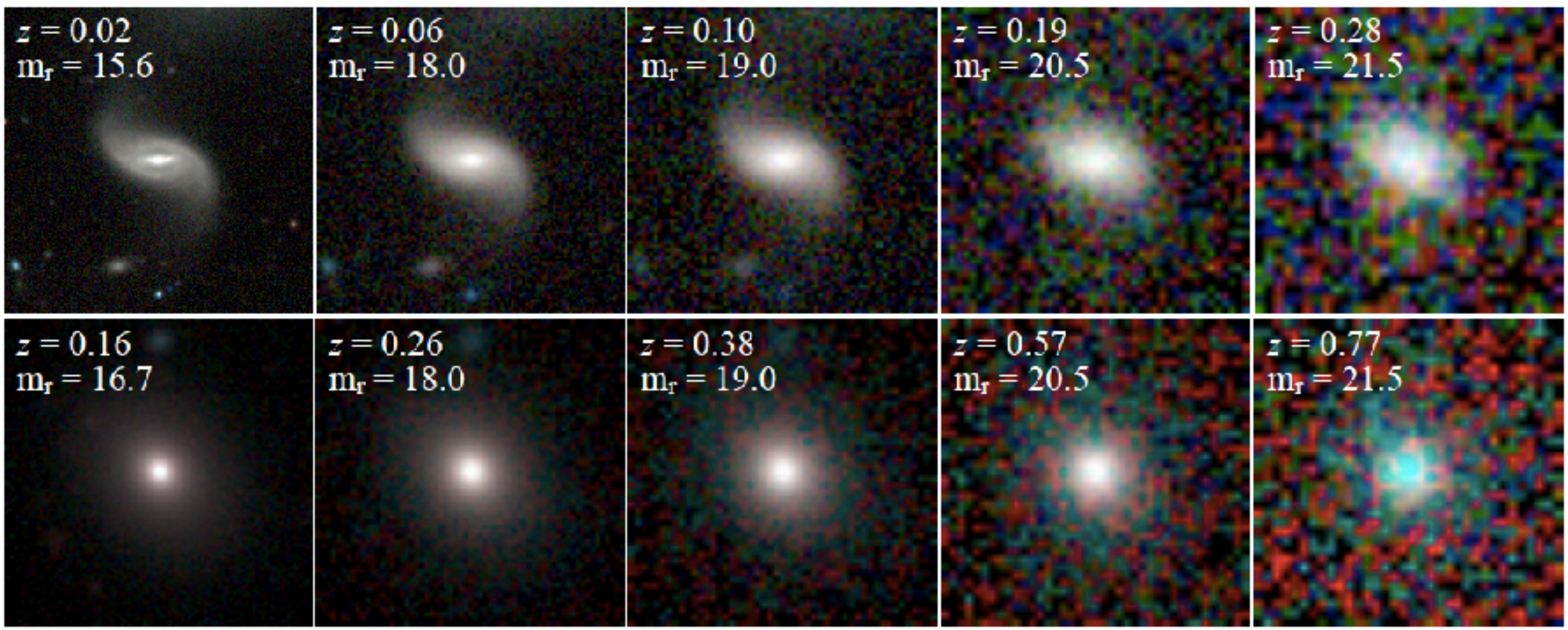
BUT…

- Current morphological catalogues are limited to very **bright observed** magnitudes (mr < 18; e.g. Nair& Abraham+10, GZOO, DS+18).

- Use galaxies with known morphological classifications and simulate them at higher redshifts.

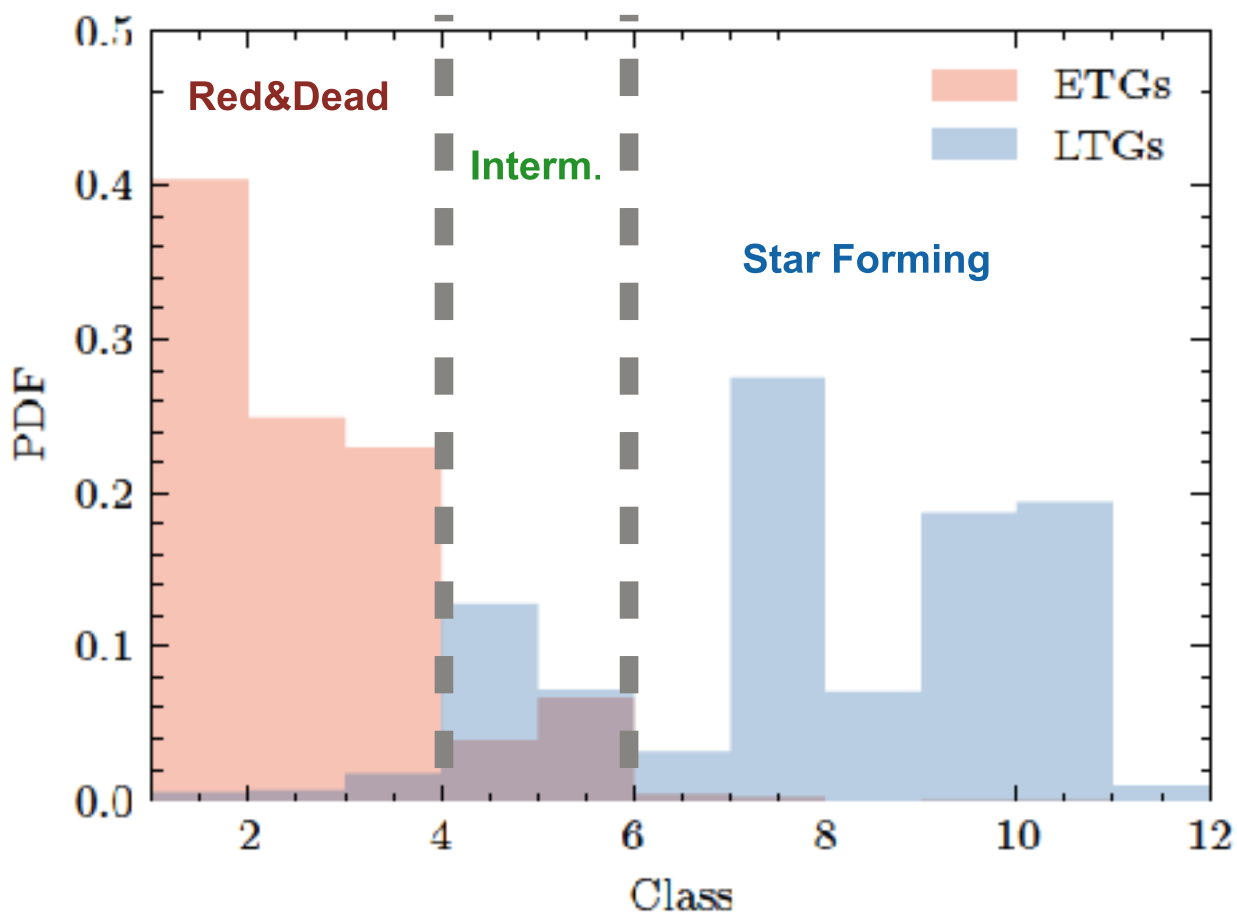
Simulations: Redshifting DES galaxies

- Can machines recover features hidden to human eye?



Vega-Ferrero, Domínguez Sánchez, Huertas-Company et al. (in prep.)

TEST on real DES galaxies

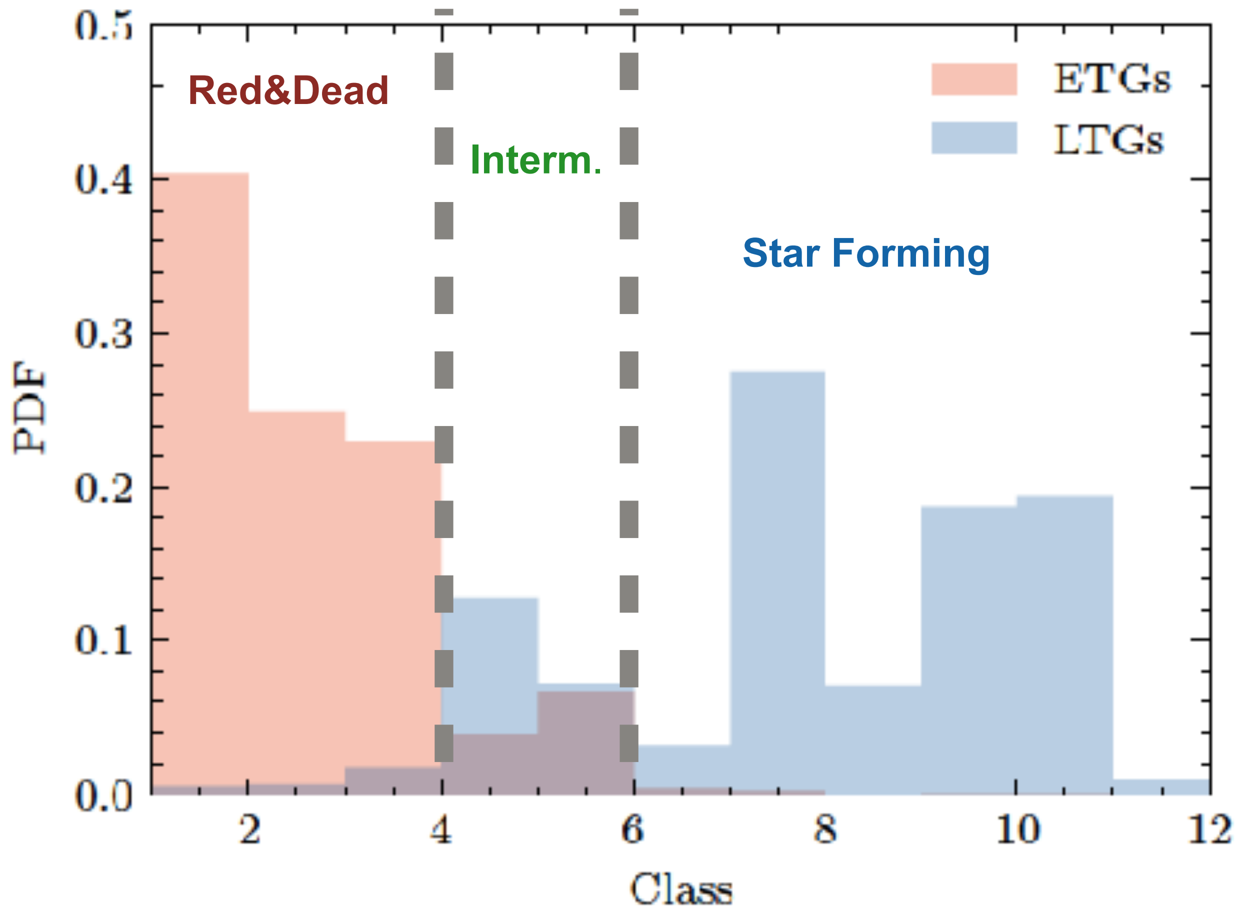


Model	p_{thr}	ROC AUC	Prec	R	Acc
p_1	0.47	0.99	0.96	0.97	0.97
p_2	0.46	0.99	0.95	0.96	0.96
p_3	0.39	0.99	0.96	0.97	0.97
p_4	0.41	0.99	0.96	0.96	0.96
p_5	0.54	0.99	0.96	0.96	0.96
$\tilde{p}$	0.41	0.99	0.96	0.97	0.97
$p_{edge-on}$	0.25	1.00	0.81	0.99	0.97

Comparison with VIPERS spectral classification (Siudek+2018):
✓97 % LTGs have Class > 4 and 89 % ETGs have Class < 4

- Can machines recover features hidden to human eye? **YES!**

TEST on real DES galaxies



Largest morphological catalogue up to date!

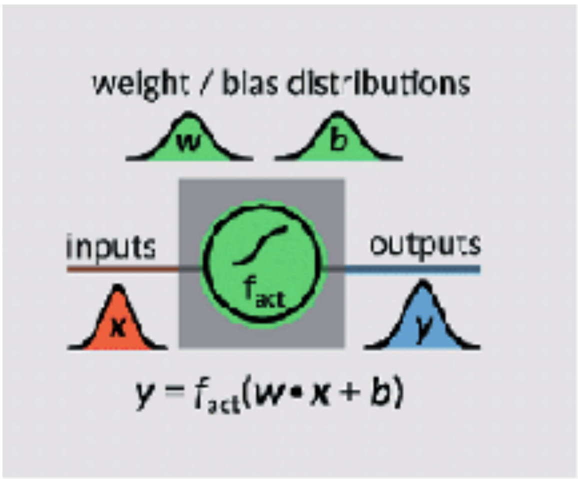
✓ 27 million galaxies up to $m_r < 21.5$

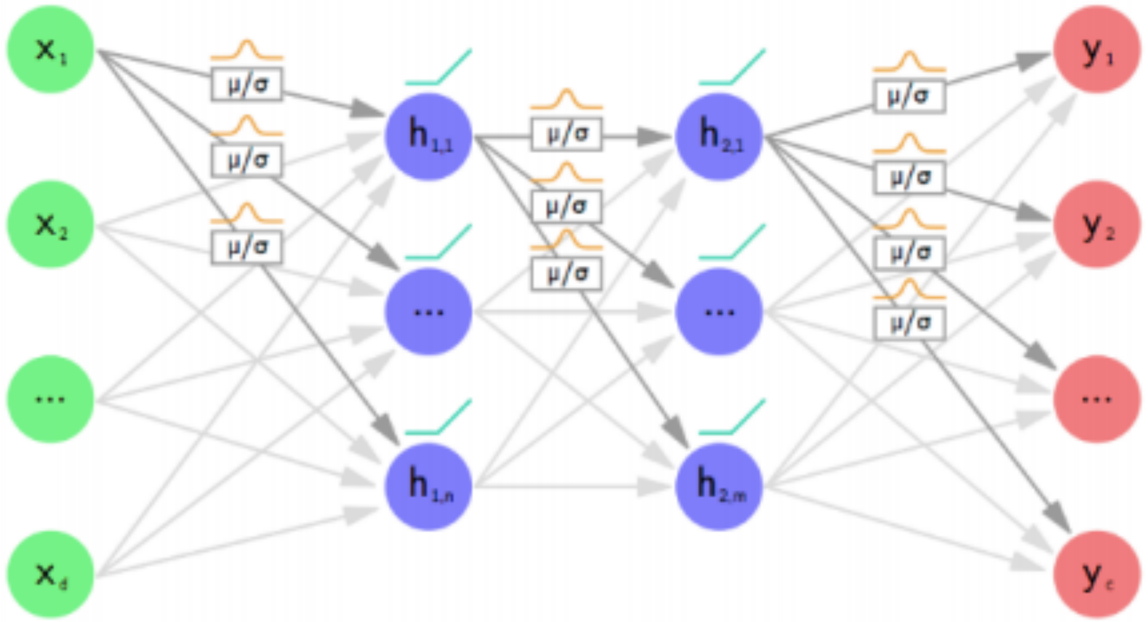
Vega-Ferrero, Domínguez Sánchez, Huertas-Company et al. (in prep.)

What about errors?

- Reliability and confidence estimates of CNN are important for astronomy.

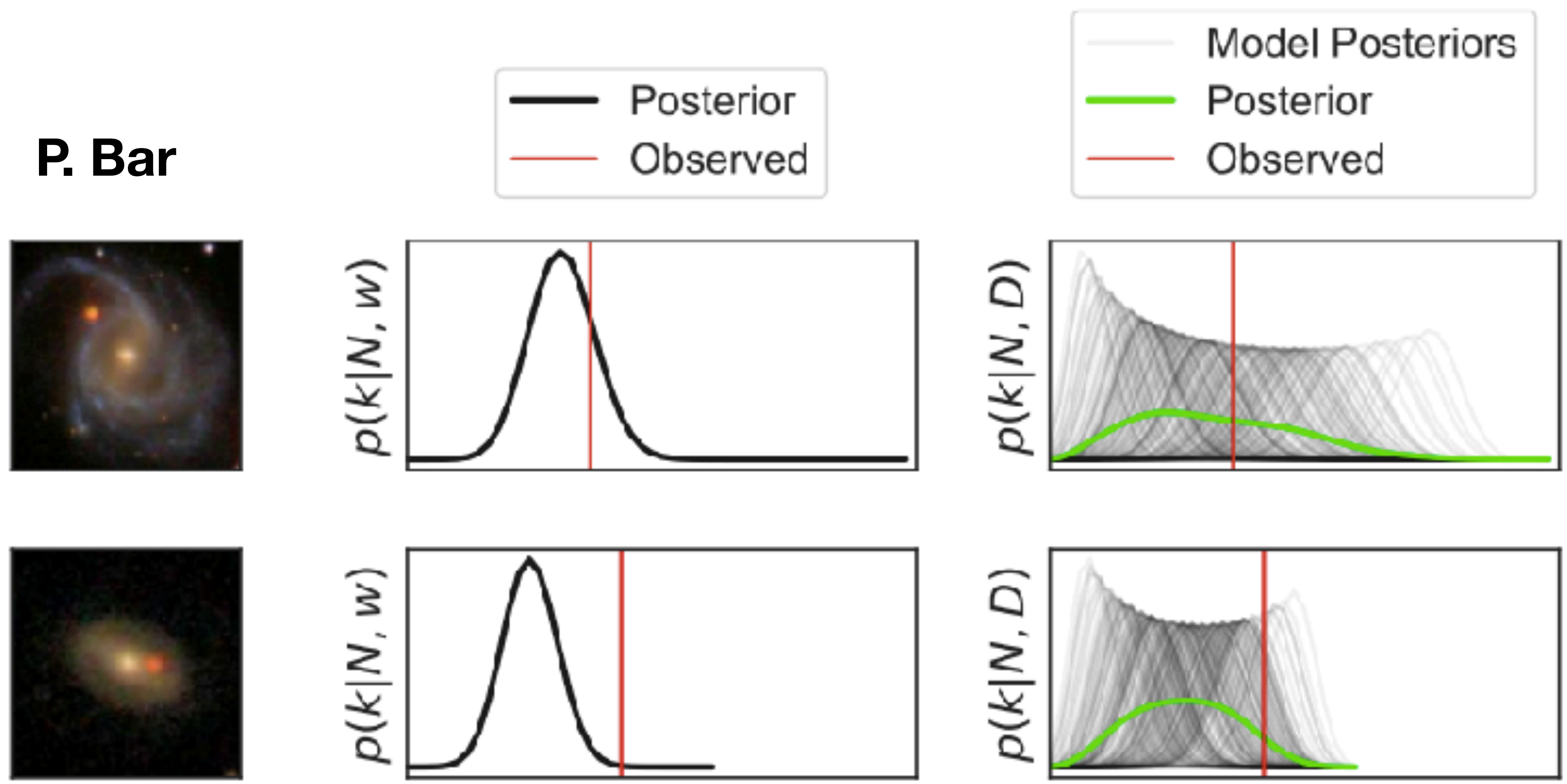
- **Bayesian Neural Networks** directly model the uncertainty of the estimated network weights (e.g., Perreault Levasseur+17, Lin+20, Ho+20).





What about errors?

P. Bar



Bayesian Neural Networks + MonteCarlo Dropout (Walsmley+2019)

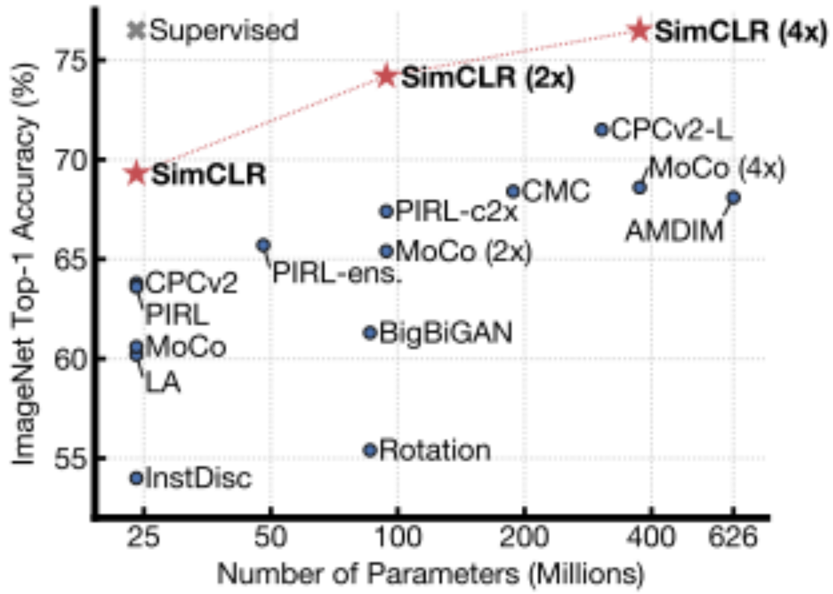
- “Pure” Reinforcement Learning (cherry)
  - The machine predicts a scalar reward given once in a while.
  - A few bits for some samples

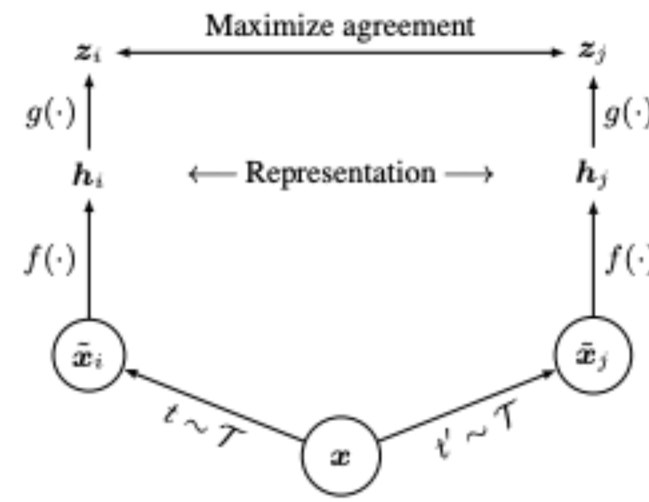
- Supervised Learning (icing)
  - The machine predicts a category or a few numbers for each input
  - Predicting human-supplied data
  - 10→10,000 bits per sample

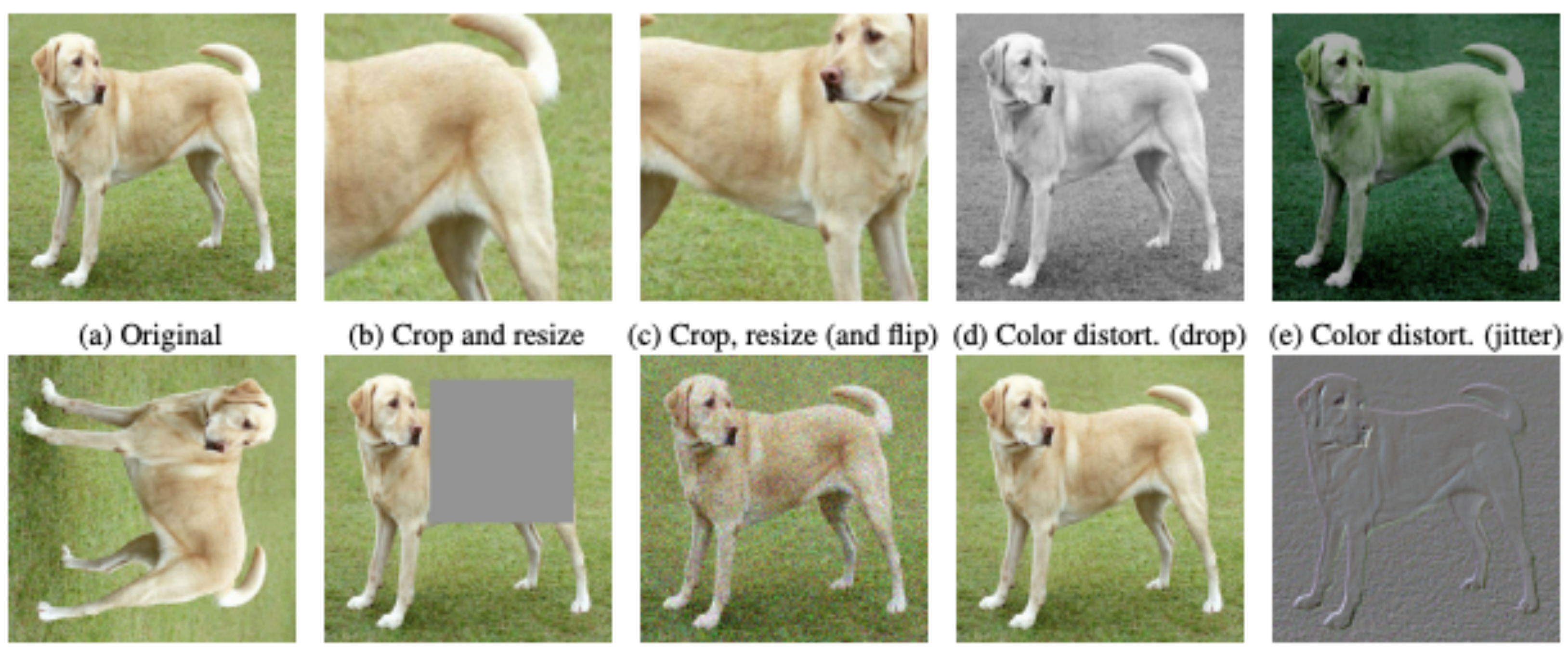
- Self-Supervised Learning (cake génoise)
  - The machine predicts any part of its input for any observed part.
  - Predicts future frames in videos
  - Millions of bits per sample

@ Y. LECUN

CLASSIFICATION WITOUT LABELS: Self Supervised Learning







(a) Original (b) Crop and resize (c) Crop, resize (and flip) (d) Color distort. (drop) (e) Color distort. (jitter)

(f) Rotate $\{90°, 180°, 270°\}$ (g) Cutout (h) Gaussian noise (i) Gaussian blur (j) Sobel filtering

Chen+2020

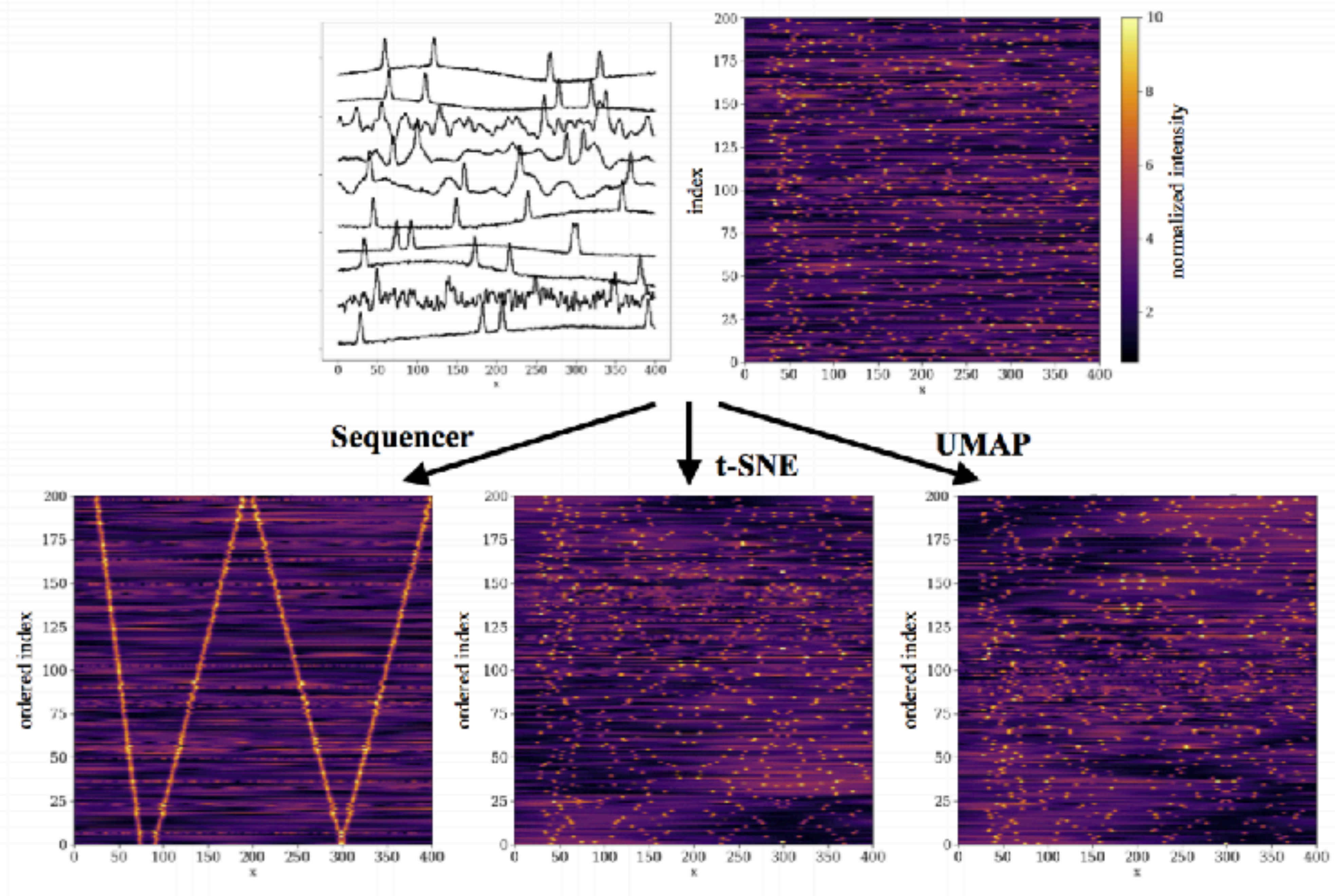
TWO MAIN CHALLENGES FOR THE NEAR FUTURE...

1. CAN ML HELP IN MAKING DISCOVERIES IN THE BIG-DATA ERA?

2. CAN ML HELP IN LEARNING PHYSICS?

HOW DO WE MAKE DISCOVERIES?

DATA VISUALIZATION IS KEY



Baron+20

DETECTING OUTLIERS …

$$p(X)$$

you data follows some probability distribution p

Then an object will be anomalous if:

$$p(x_i) < \epsilon$$

UNKNOWN UNKNOWNS IS WHERE INTERESTING (NEW) SCIENCE WILL BE FOUND

GENERATIVE MODELS DO PRECISELY THAT:

DATA:

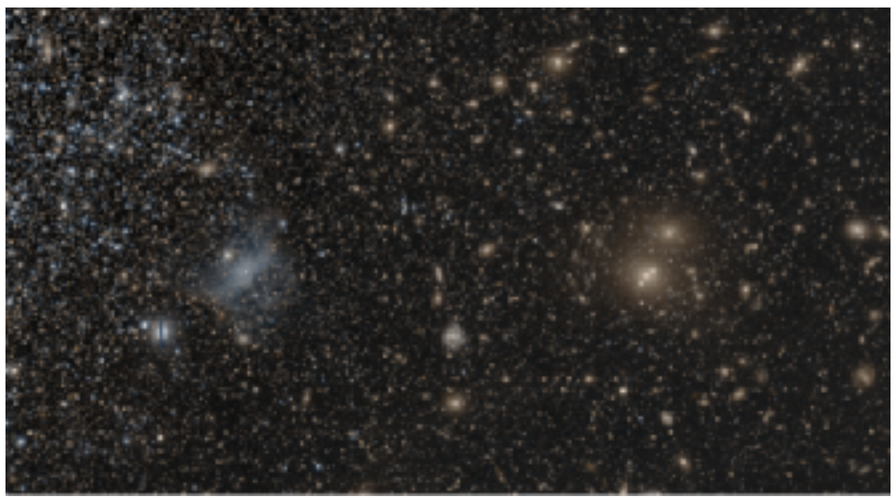
X

DEEP GENERATIVE MODEL

PROBABILITY DENSITY FUNCTION:

$P(X)$

Hyper Suprime Cam Survey (HSC) (Subaru telesocpe 1400 sq. deg, r~26)



WGAN

$P_{HSC}(X)$

x_i

(individual detection)

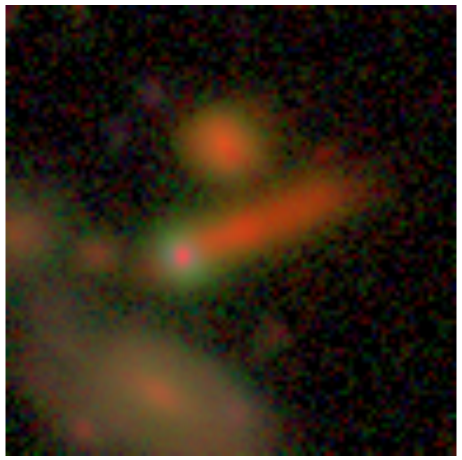
$P_{HSC}(x_i)$

identify most "unlikely" objects

Storey-Fisher, MHC, Leauthaud+ (in prep)

SOME EXAMPLES OF HSC "ANOMALIES"

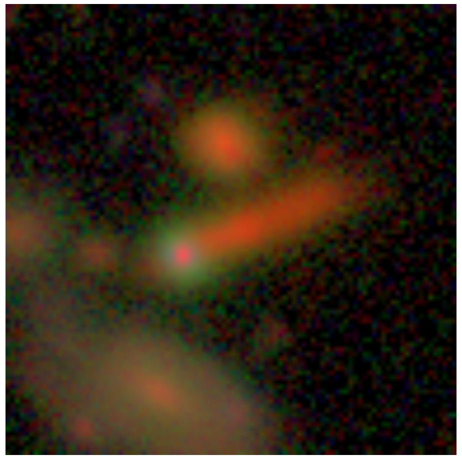
HSC



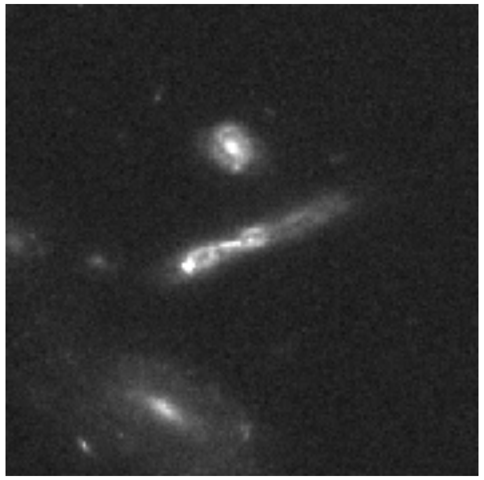
Sotrey-Fisher, MHC, Leauthaud+ (in prep)

SOME EXAMPLES OF HSC "ANOMALIES"

HSC



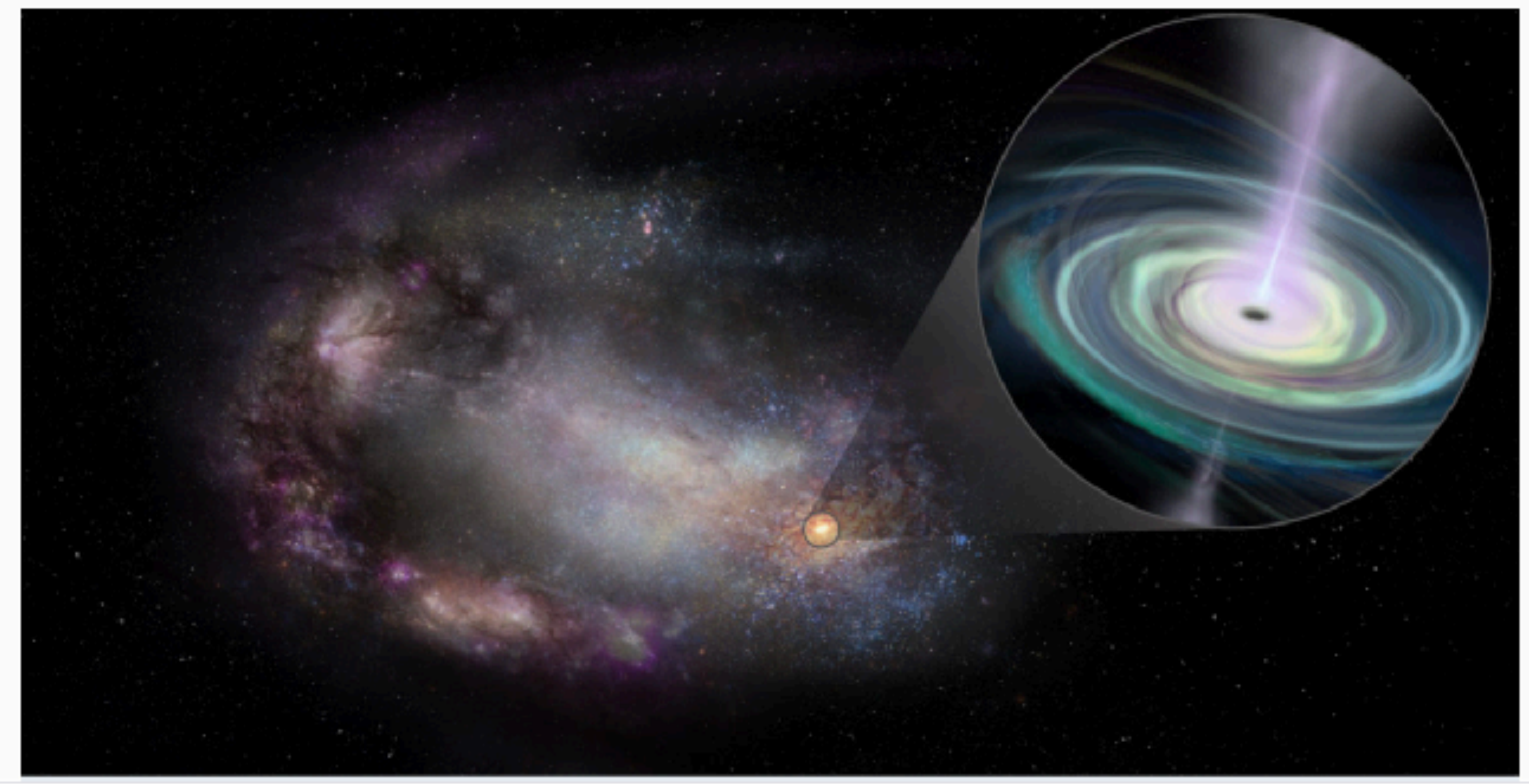
HST



Followed-up with Keck

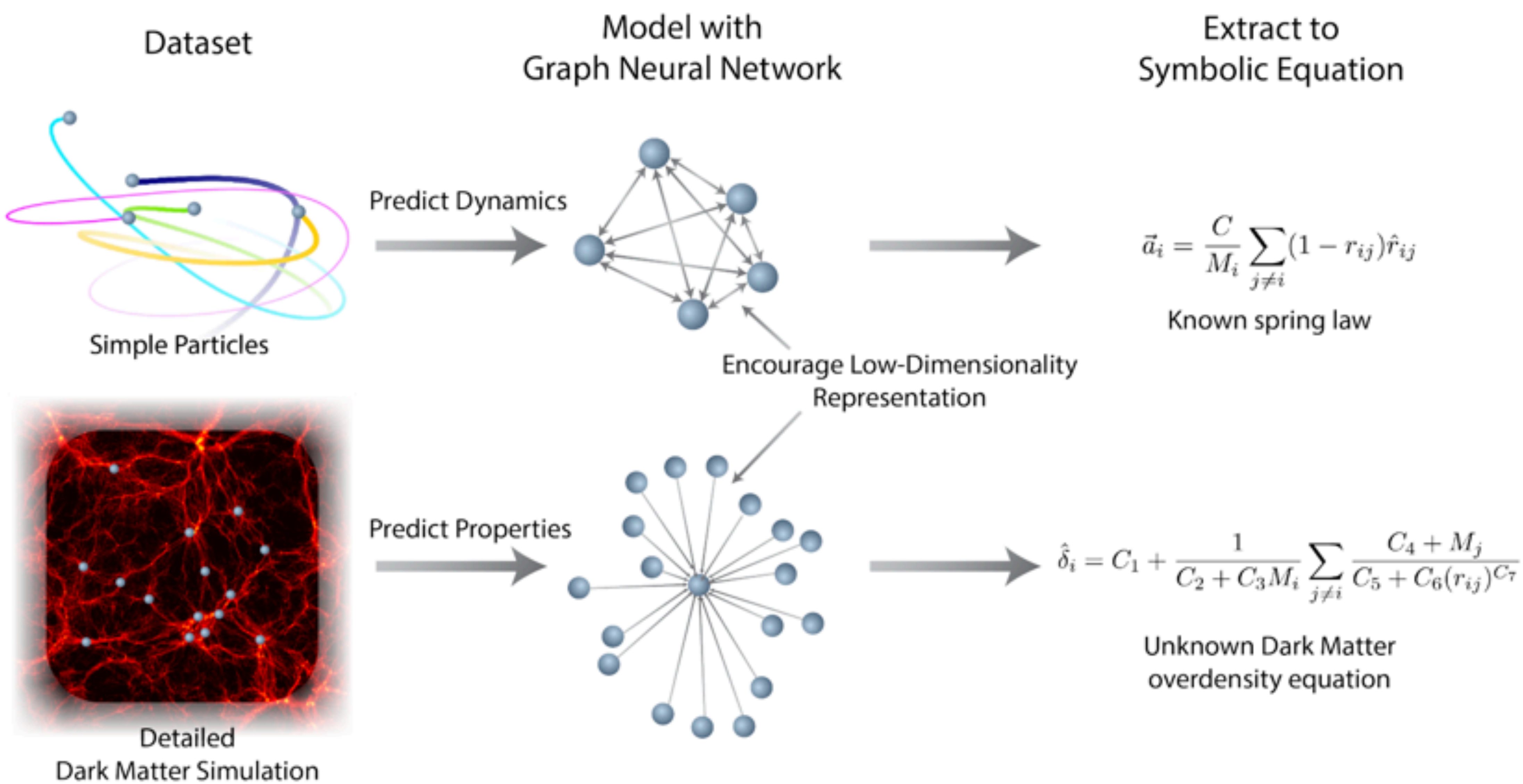
wandering black hole?

https://public.nrao.edu/news/wandering-black-holes-dwarf-galaxies/



Sotrey-Fisher, MHC, Leauthaud+ (in prep)

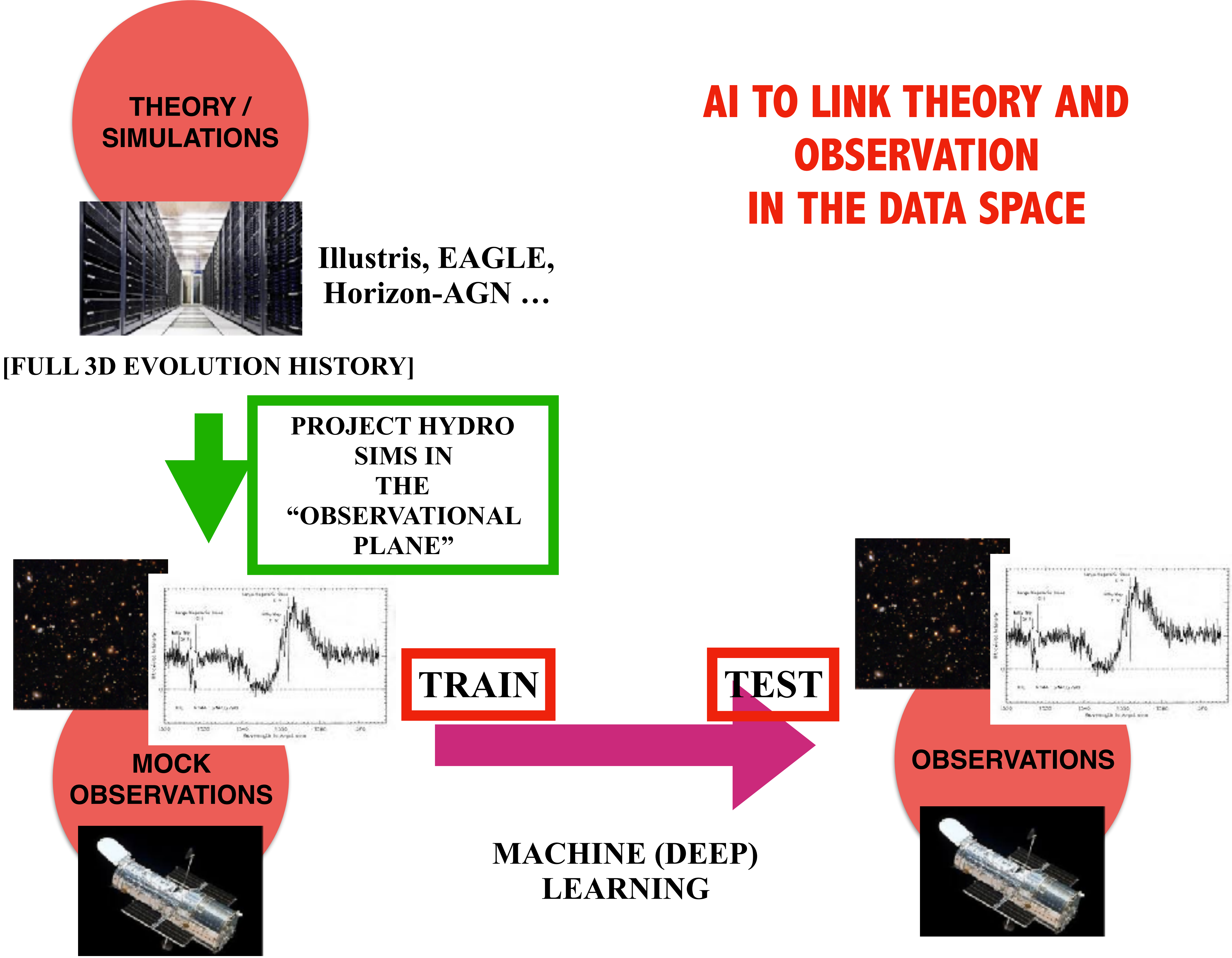
HOW DO WE LEARN PHYSICS?


Dataset
Model with
Graph Neural Network
Extract to
Symbolic Equation
Simple Particles
Predict Dynamics
$\vec{a}_i = \frac{C}{M_i}\sum_{j\neq i}(1-r_{ij})\hat{r}_{ij}$
Known spring law
Encourage Low-Dimensionality
Representation
Predict Properties
$\hat{\delta}_i = C_1 + \frac{1}{C_2 + C_3 M_i}\sum_{j\neq i}\frac{C_4 + M_j}{C_5 + C_6 (r_{ij})^{C_7}}$
Unknown Dark Matter
overdensity equation
Detailed
Dark Matter Simulation

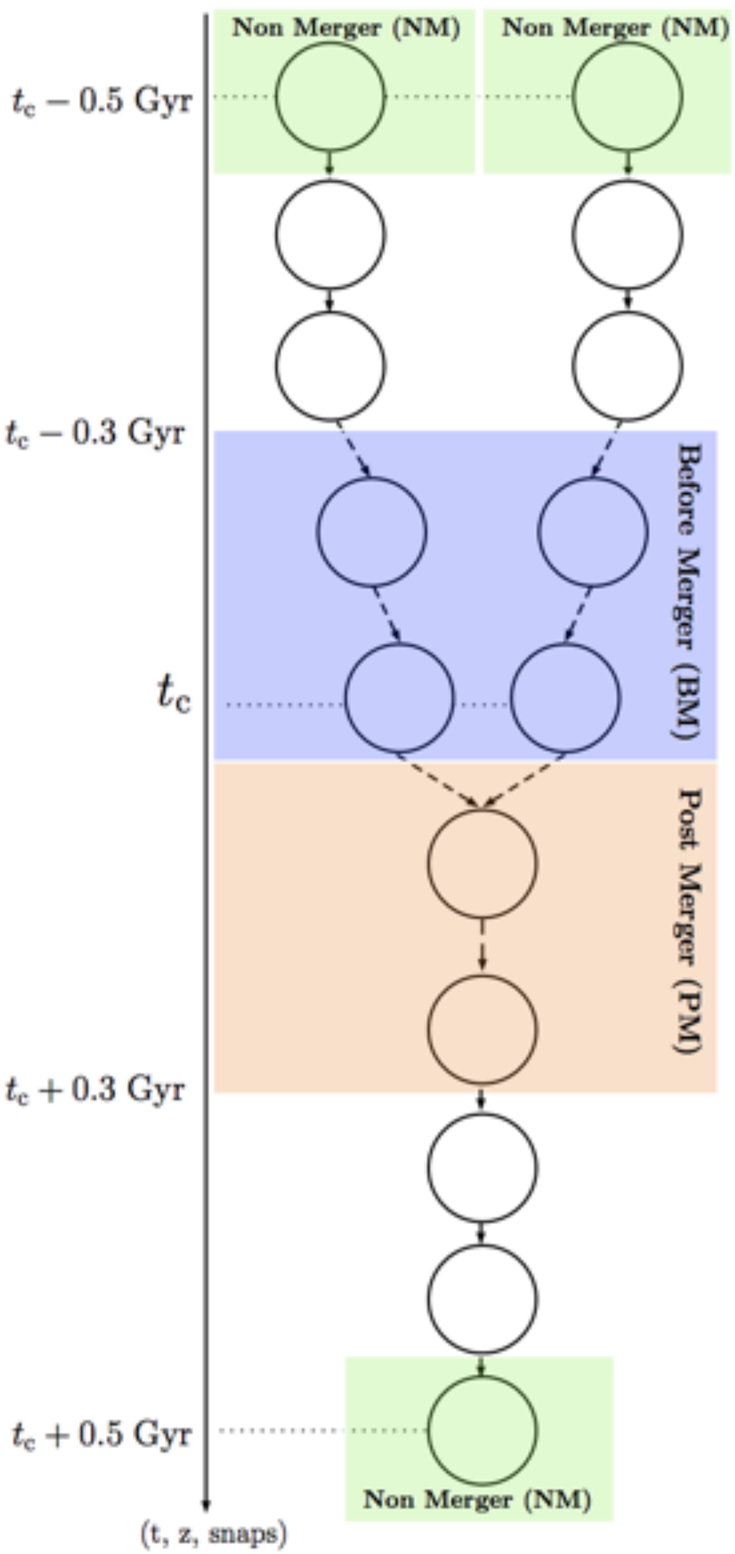
Cranmer+20

ADVANCED COSMOLOGICAL SIMULATIONS NOW CAPTURE THE TIME EVOLUTION OF GALAXIES

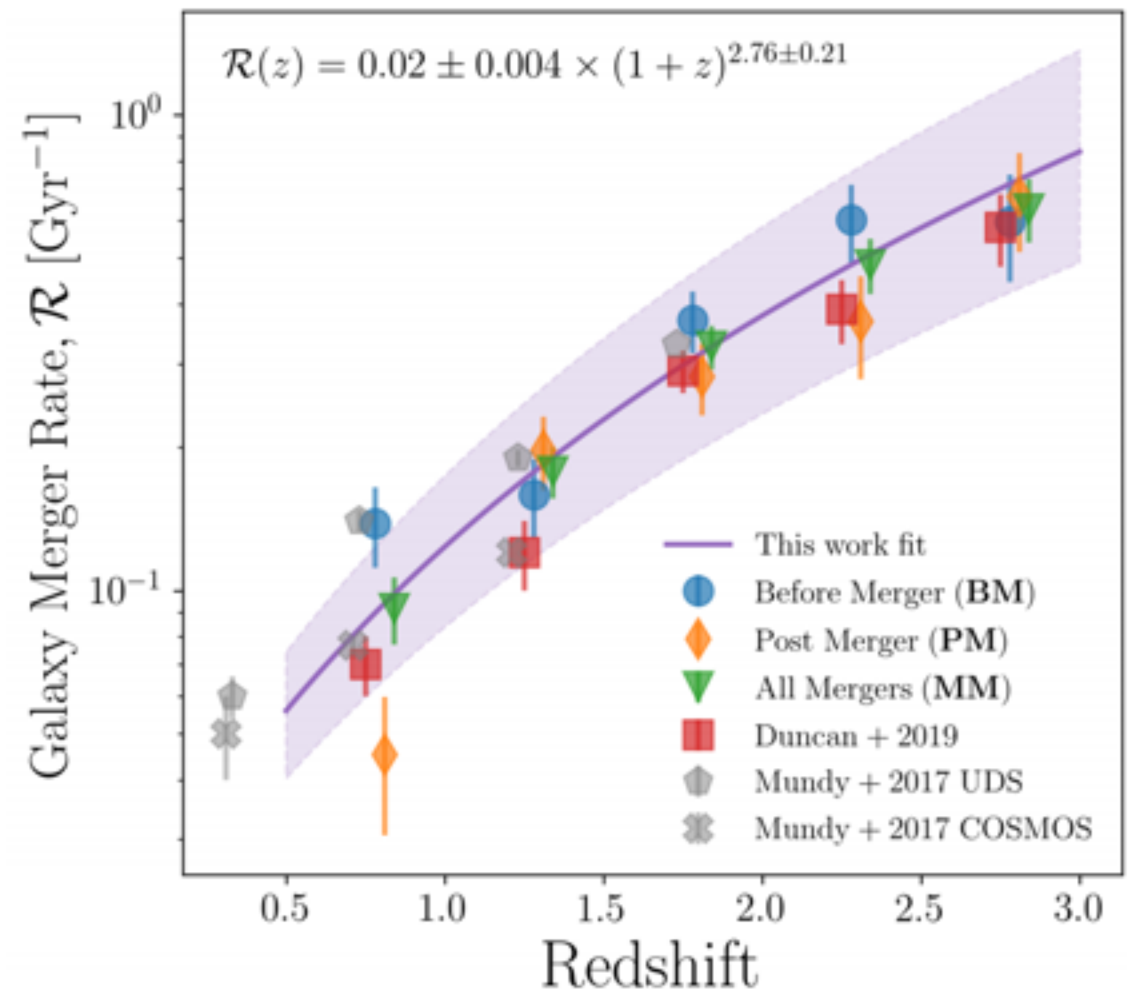
TNG SIMULATION
(https://www.tng-project.org/)



AI TO LINK THEORY AND
OBSERVATION
IN THE DATA SPACE
THEORY /
SIMULATIONS
Illustris, EAGLE,
Horizon-AGN …
[FULL 3D EVOLUTION HISTORY]
PROJECT HYDRO
SIMS IN
THE
“OBSERVATIONAL
PLANE”
TRAIN
TEST
MOCK
OBSERVATIONS
MACHINE (DEEP)
LEARNING
OBSERVATIONS



Non Merger (NM)
Non Merger (NM)
$t_c - 0.5$ Gyr
$t_c - 0.3$ Gyr
t_c
Before Merger (BM)
Post Merger (PM)
$t_c + 0.3$ Gyr
$t_c + 0.5$ Gyr
Non Merger (NM)
(t, z, snaps)

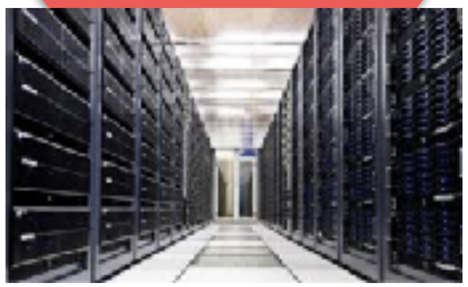


$\mathcal{R}(z) = 0.02 \pm 0.004 \times (1+z)^{2.76\pm0.21}$
Galaxy Merger Rate, $\mathcal{R}$ [Gyr^{-1}]
Redshift
This work fit
Before Merger (BM)
Post Merger (PM)
All Mergers (MM)
Duncan + 2019
Mundy + 2017 UDS
Mundy + 2017 COSMOS

Ferreira+20

AI TO CHECK IF THE PHYSICS IS CORRECT....

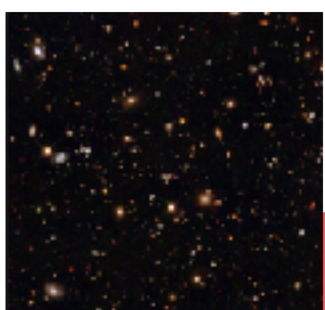
THEORY / SIMULATIONS



Illustris, EAGLE, Horizon-AGN ...

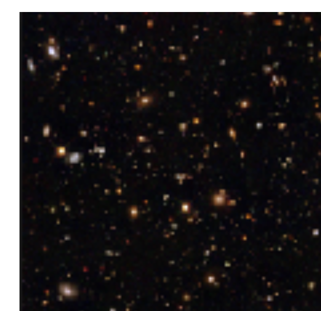
[FULL 3D EVOLUTION HISTORY]

FORWARD MODEL HYDRO SIMS IN THE "OBSERVATIONAL PLANE"



MOCK OBSERVATIONS

MACHINE (DEEP) LEARNING



OBSERVATIONS

MATCHING OBSERVED AND SIMULATED PROBABILITY DISTRIBUTIONS

$$p(x_i|\theta_{SDSS}) \sim p(x_i^s|\theta_{SDSS})$$

OBSERVED GALAXY

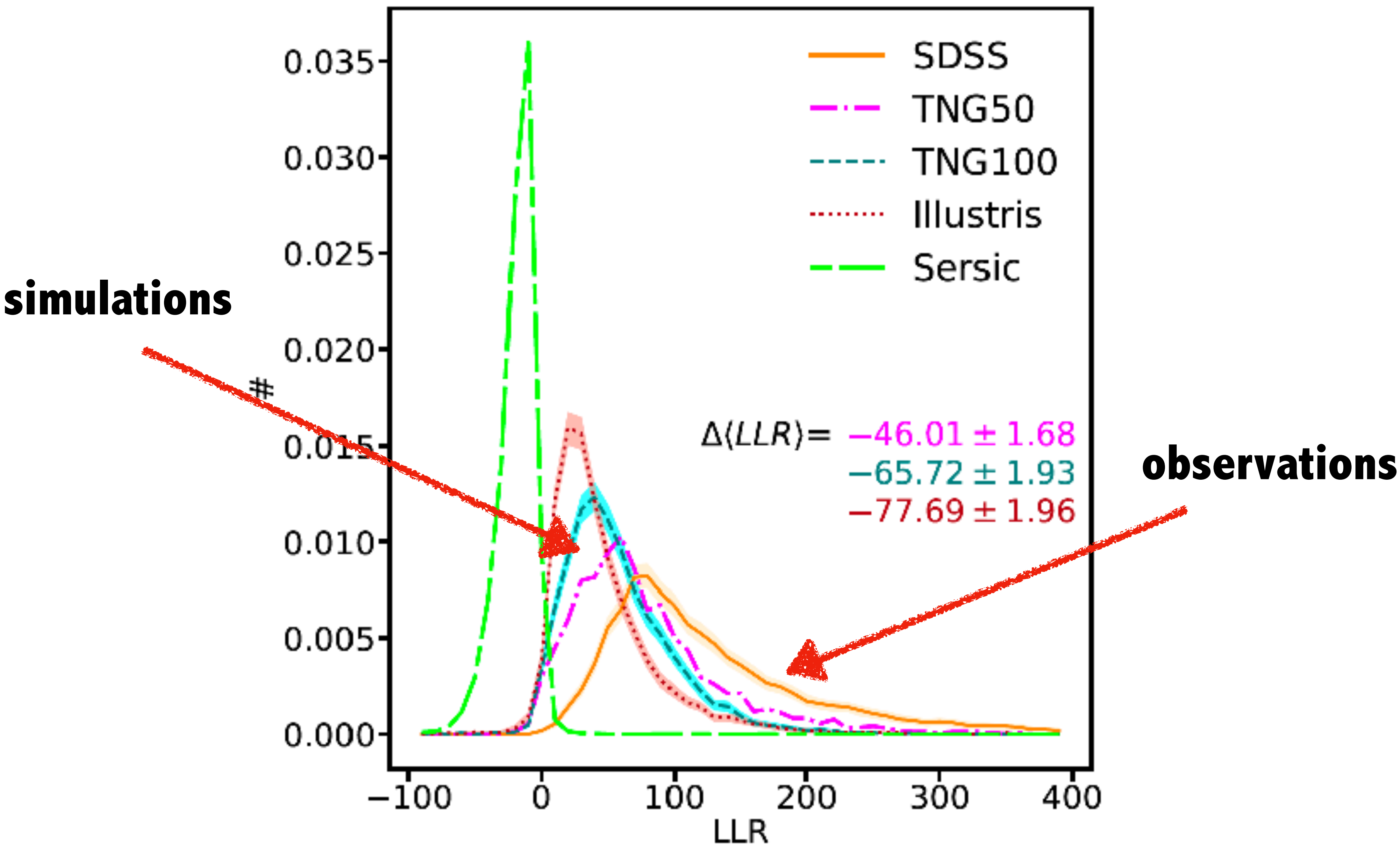
SIMULATED GALAXY

OBSERVATIONAL PROPERTIES

OBSERVATIONAL PROPERTIES

SEE ALSO VERY NICE PAPER FROM TODAY: https://arxiv.org/abs/2007.05535

SIMULATIONS IMPROVE OVER TIME BUT STILL HAVE LOWER LIKLEIHOODS THAN THE DATA



Zanisi, MHC, Lanusse+20

SUMMARY

- Deep supervised learning is still an infant but on his way to maturity
  - It will most probably be integrated at different levels in many future pipelines for object classification, segmentation and other regressive tasks.
- <u>**There is many more - and moving very fast! :**</u>
  - <u>blind search of interesting objects</u> in large surveys
  - comparing simulations with observations and eventually constraining physics …

SUMMARY

- Deep supervised learning is still an infant but on his way to maturity
  - It will most probably be integrated at different levels in many future pipelines for object classification, segmentation and other regressive tasks.
- **There is many more - and moving very fast! :**
  - blind search of interesting objects in large surveys
  - comparing simulations with observations and eventually constraining physics …

JOIN OUR Bi-WEEKLY ML-PHYSICS DISCUSSIONS: **mlclub.net**