Investigating galaxy evolution with deep learning

Marc Huertas-Company

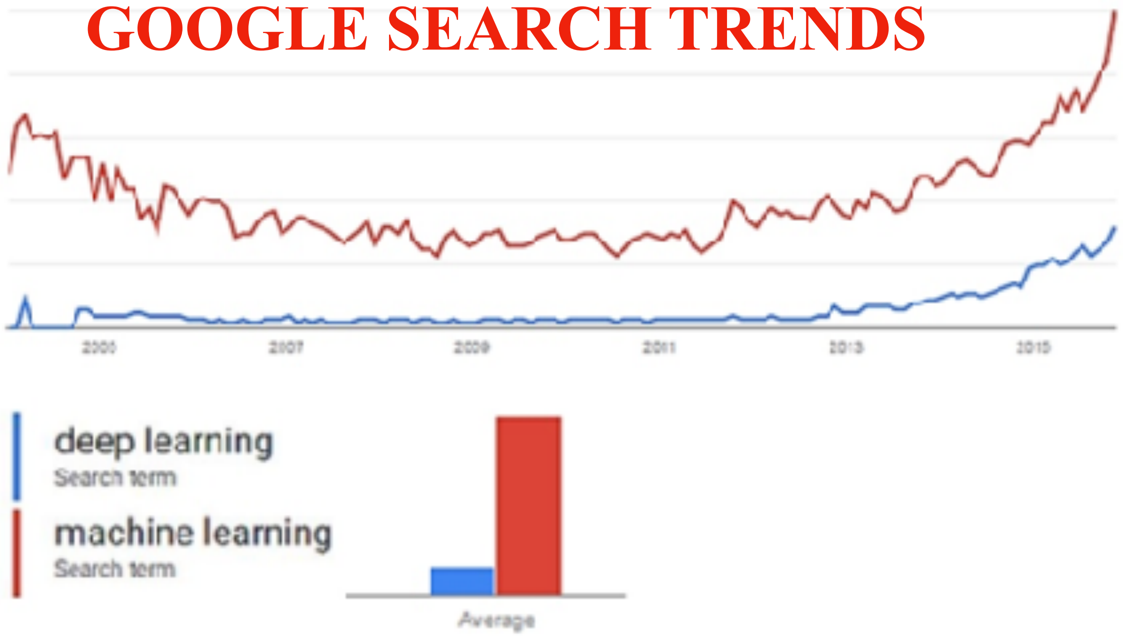
Instituto de Astrofísica de Canarias
Observatoire de Paris
Université Paris Diderot
Institut Universitaire de France

Berta Margalef-Bentabol, Fernando Caro, Christoph Lee, Alexandre Boucaud, Avishai Dekel, Joel Primack, Tom Charnock, Annalisa Pillepich, Vicente Rodriguez (+ TNG team), Emille Ishida (+ COIN initiative), Helena Dominguez-Sanchez …

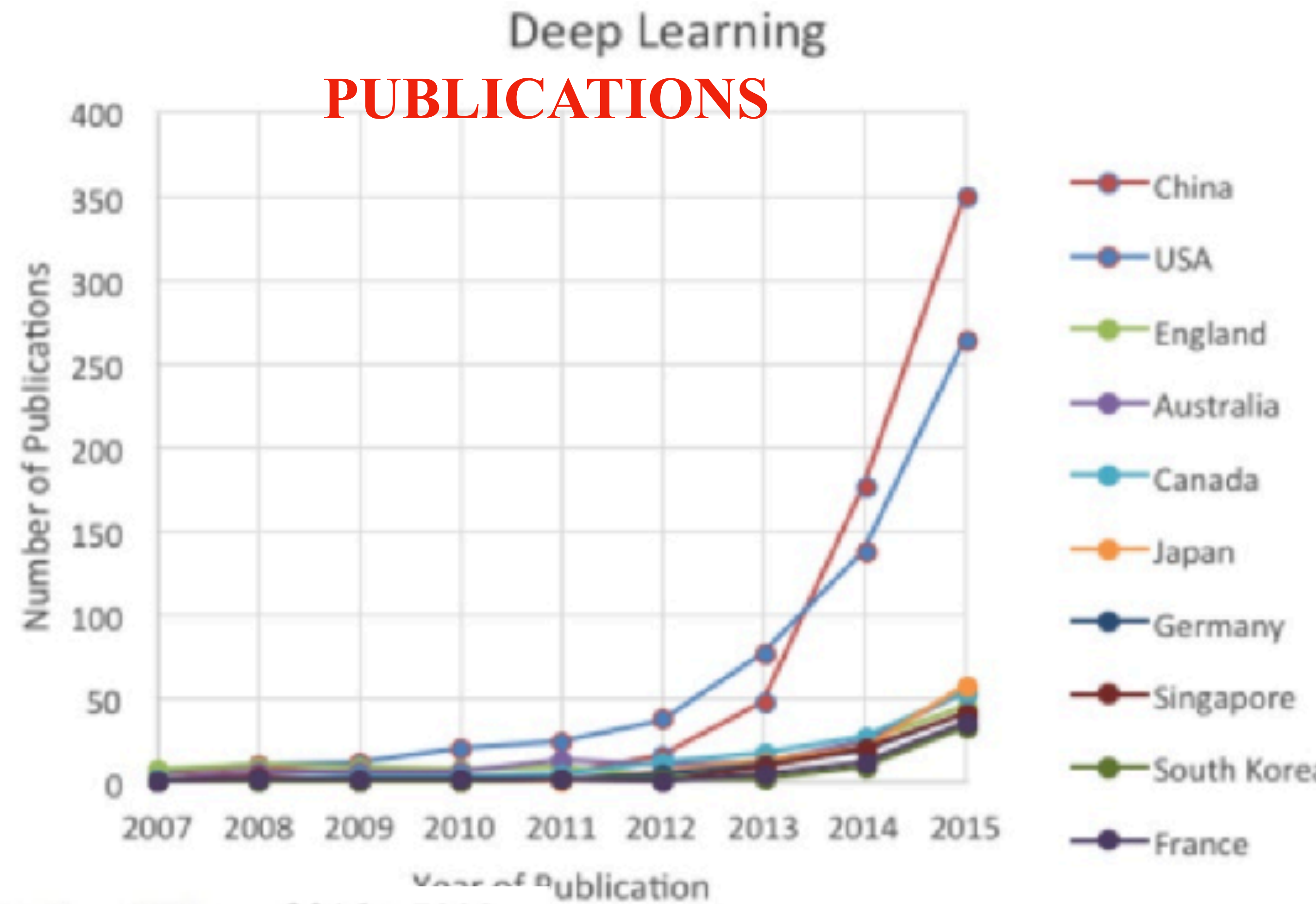
AAS 233rd meeting, Machine Learning in Astronomical Data Analysis, Seattle 2019

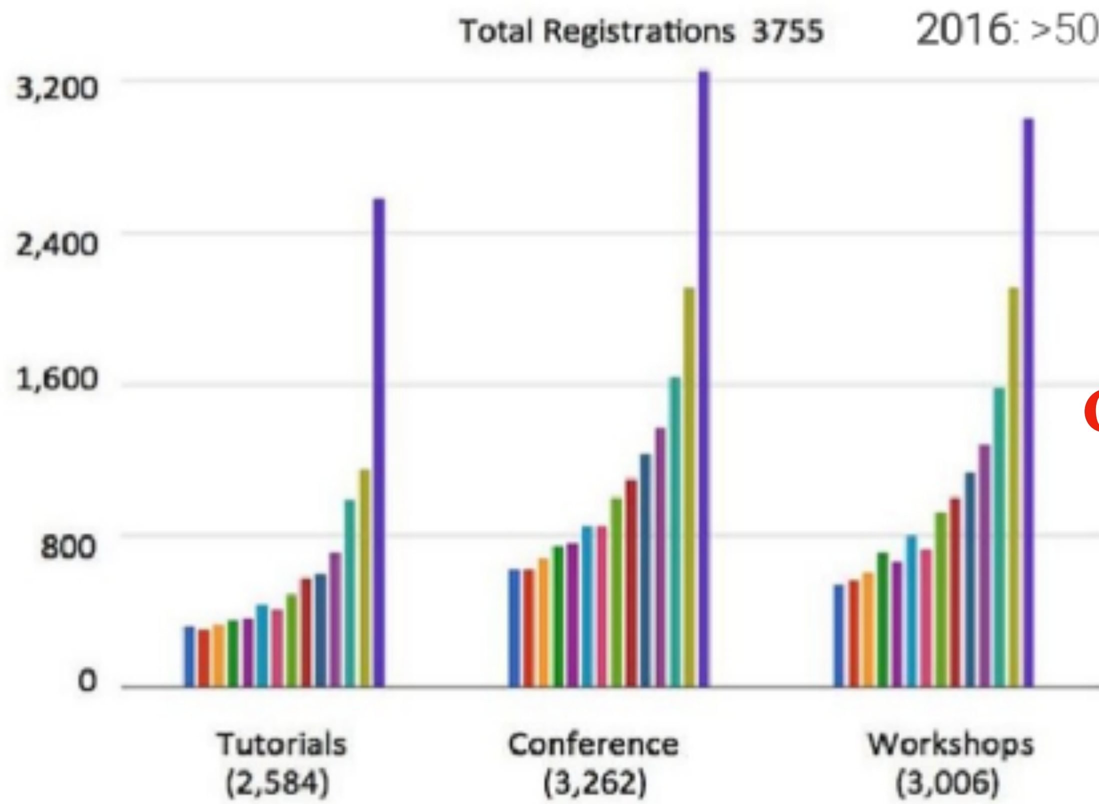
AI FEVER?

GOOGLE SEARCH TRENDS



PUBLICATIONS

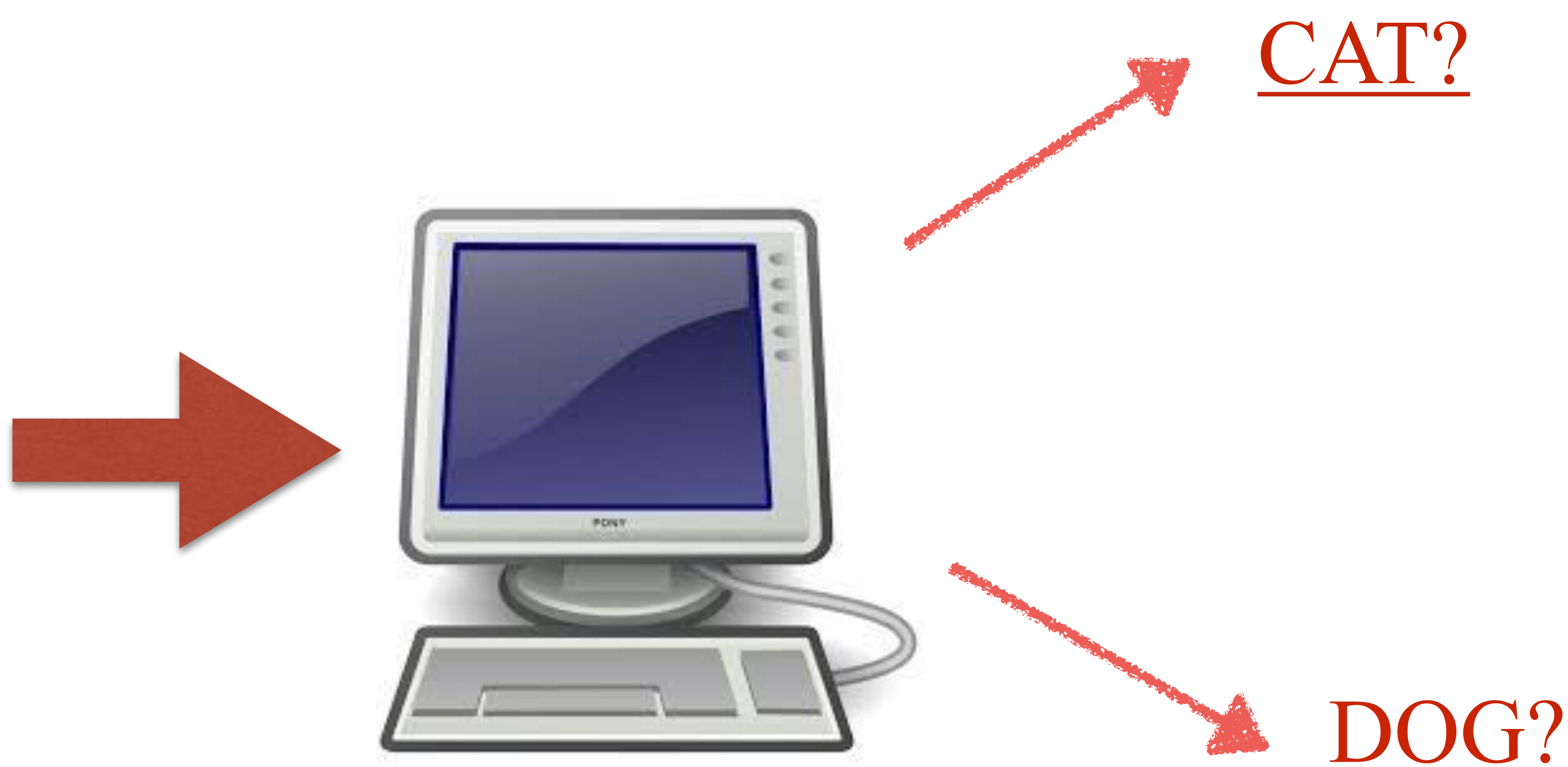




CONFERENCES

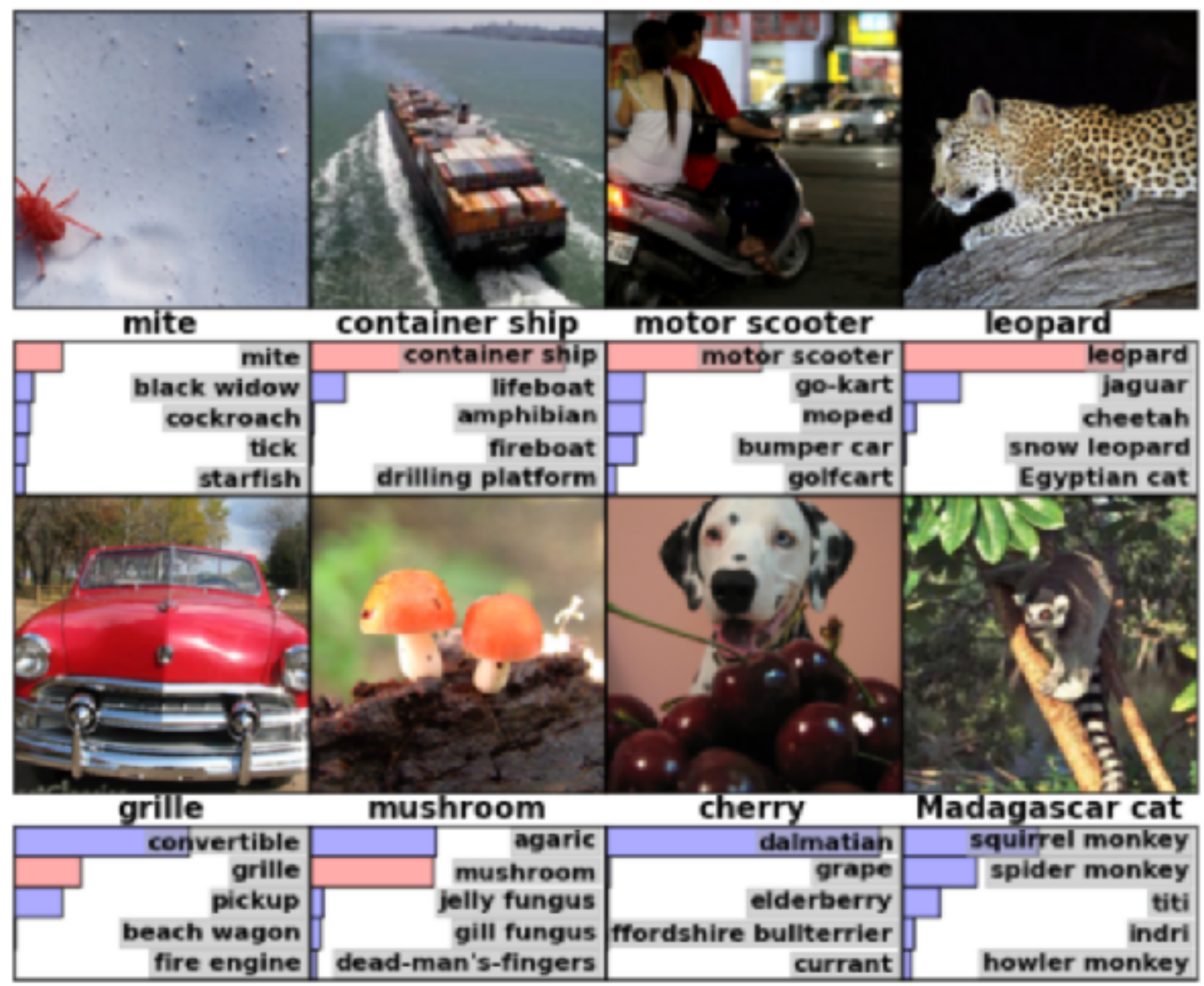
Source

BEFORE 2012….



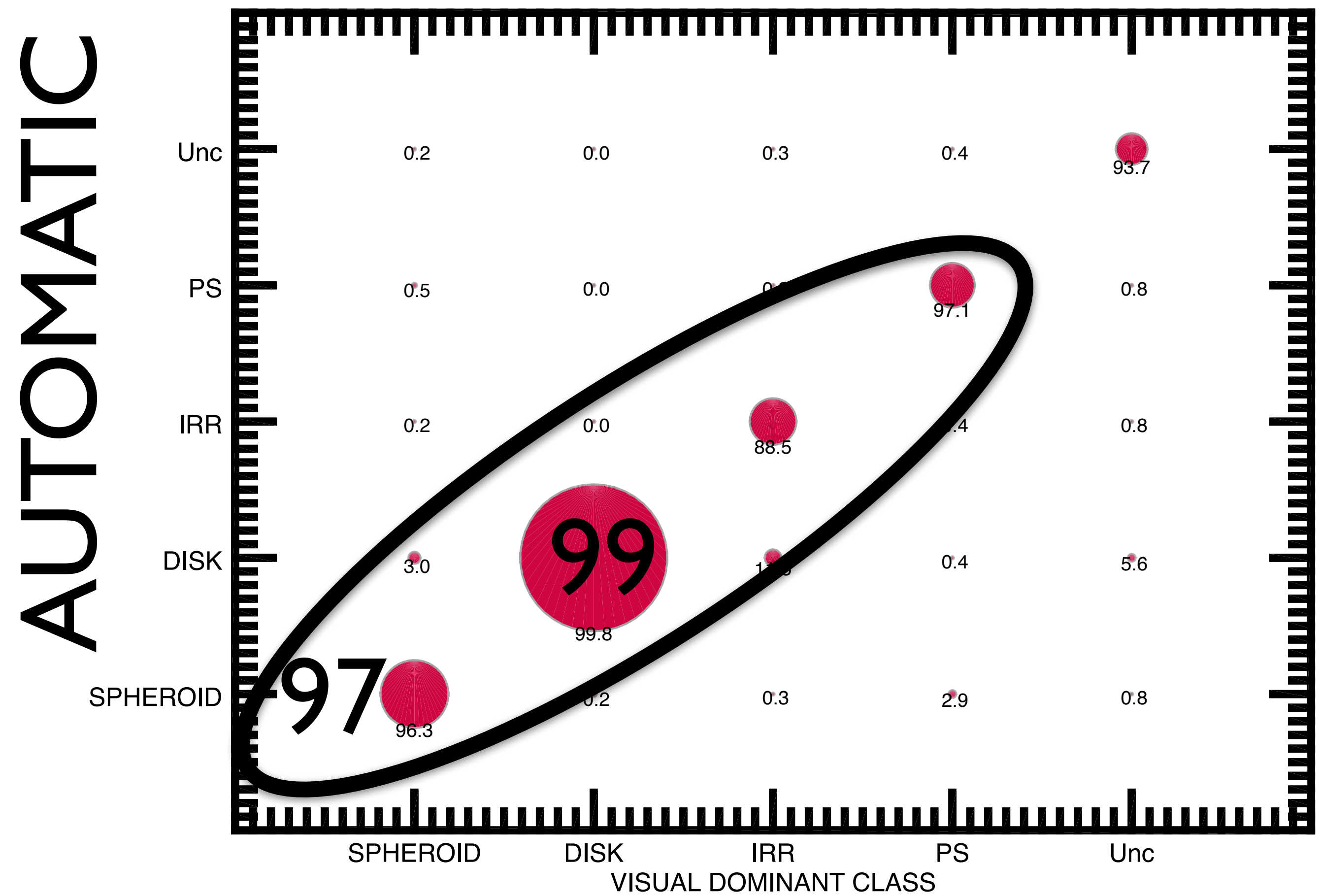
TRIVIAL HUMAN TASKS REMAINED CHALLENGING FOR COMPUTERS

AFTER 2012



IT HAS BECOME TRIVIAL….

“OUR CATS AND DOGS”: GALAXY MORPHOLOGY



VISUAL

CNNs

DEEP LEARNING SOLVES THE PROBLEM OF GALAXY MORPHOLOGICAL CLASSIFICATION?

MHC+15b

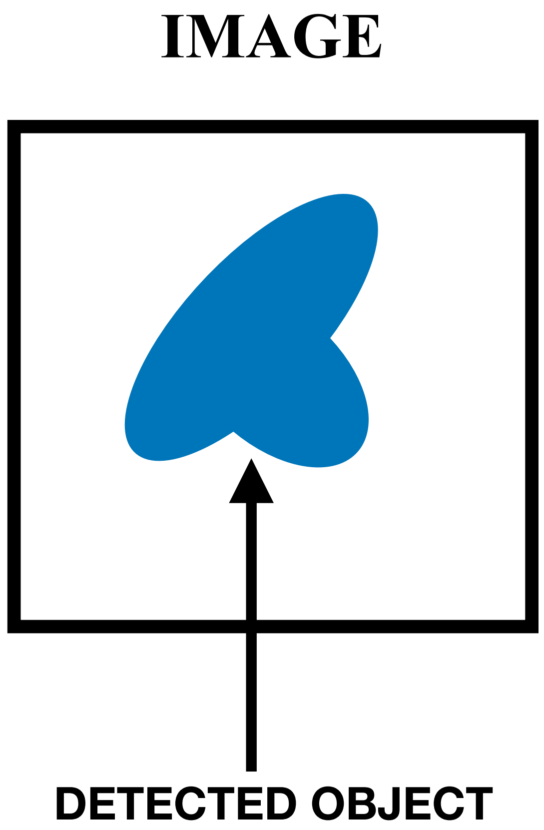
TALK ON WEDNESDAY AFTERNOON -

SPECIAL SESSION
ON "MODERN" GALAXY MORPHOLOGIES

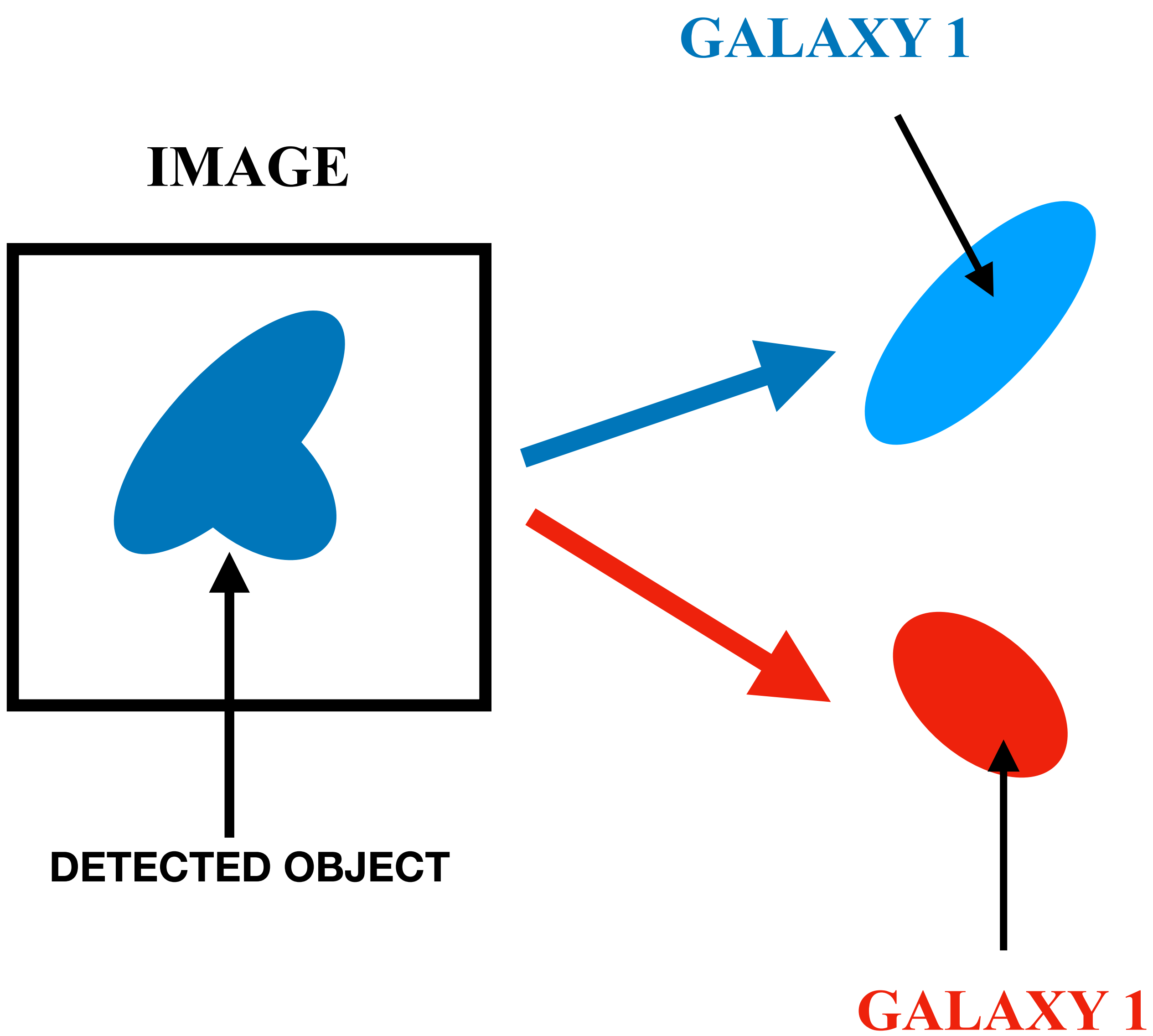
2 TAKE HOME MESSAGES FOR TODAY

1. MOST OF THE PROCESSING WE DO ON IMAGES CAN BE DONE WITH AI - POSSIBLY MORE EFFICIENTLY AND MORE ACCURATELY

2. WE CAN LEARN SOME PHYSICS BY USING AI TO LINK SIMULATIONS AND OBSERVATIONS



Boucaud+19



Boucaud+19

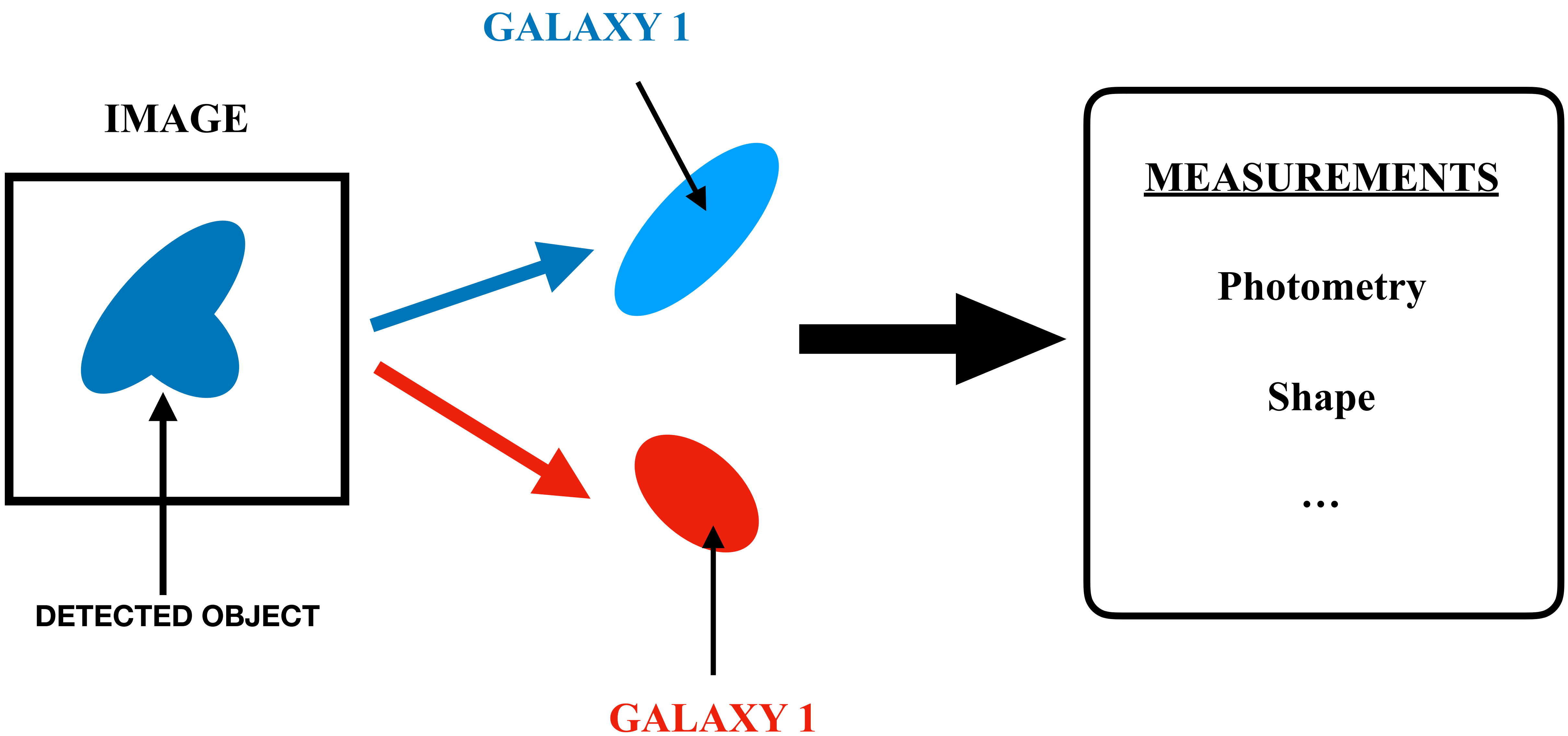
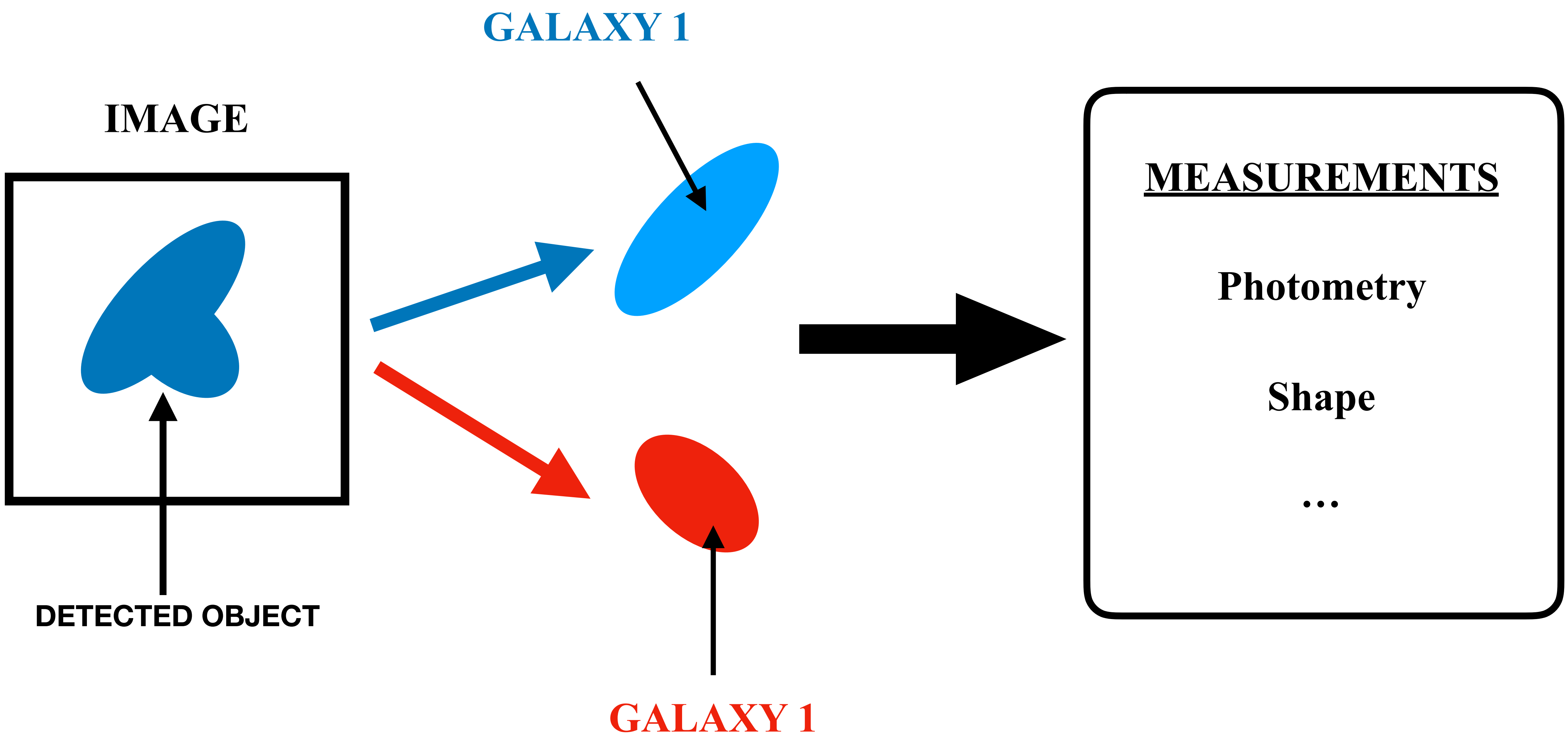
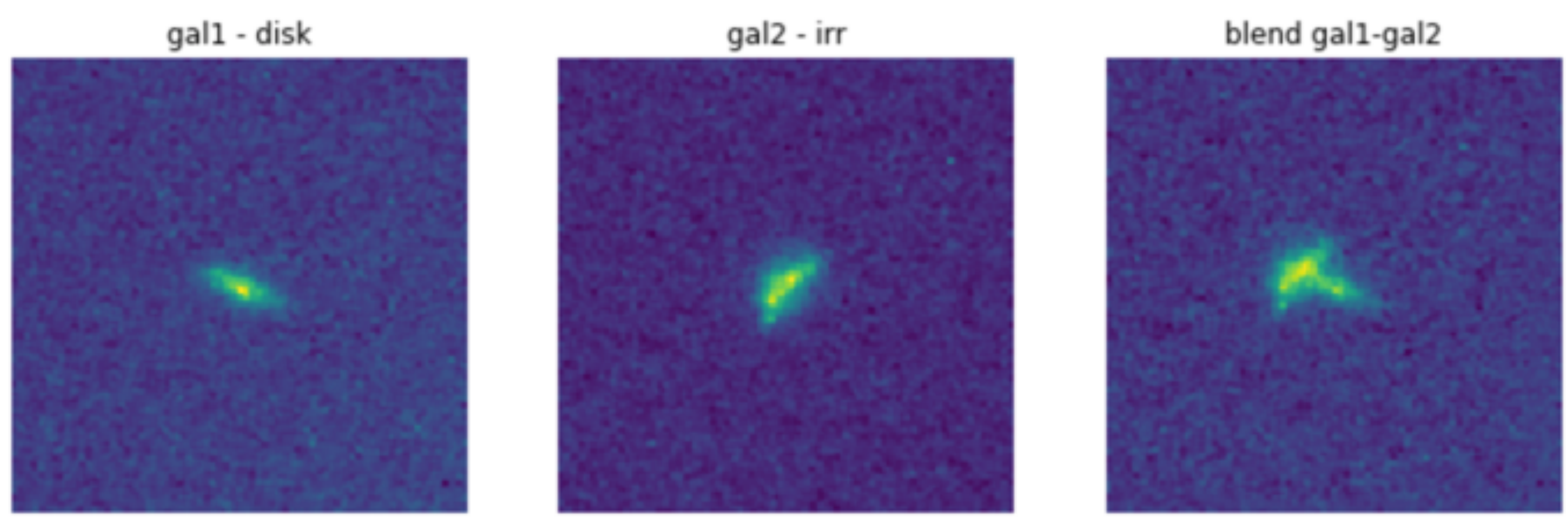


IMAGE
DETECTED OBJECT
GALAXY 1
GALAXY 1
MEASUREMENTS
Photometry
Shape
...

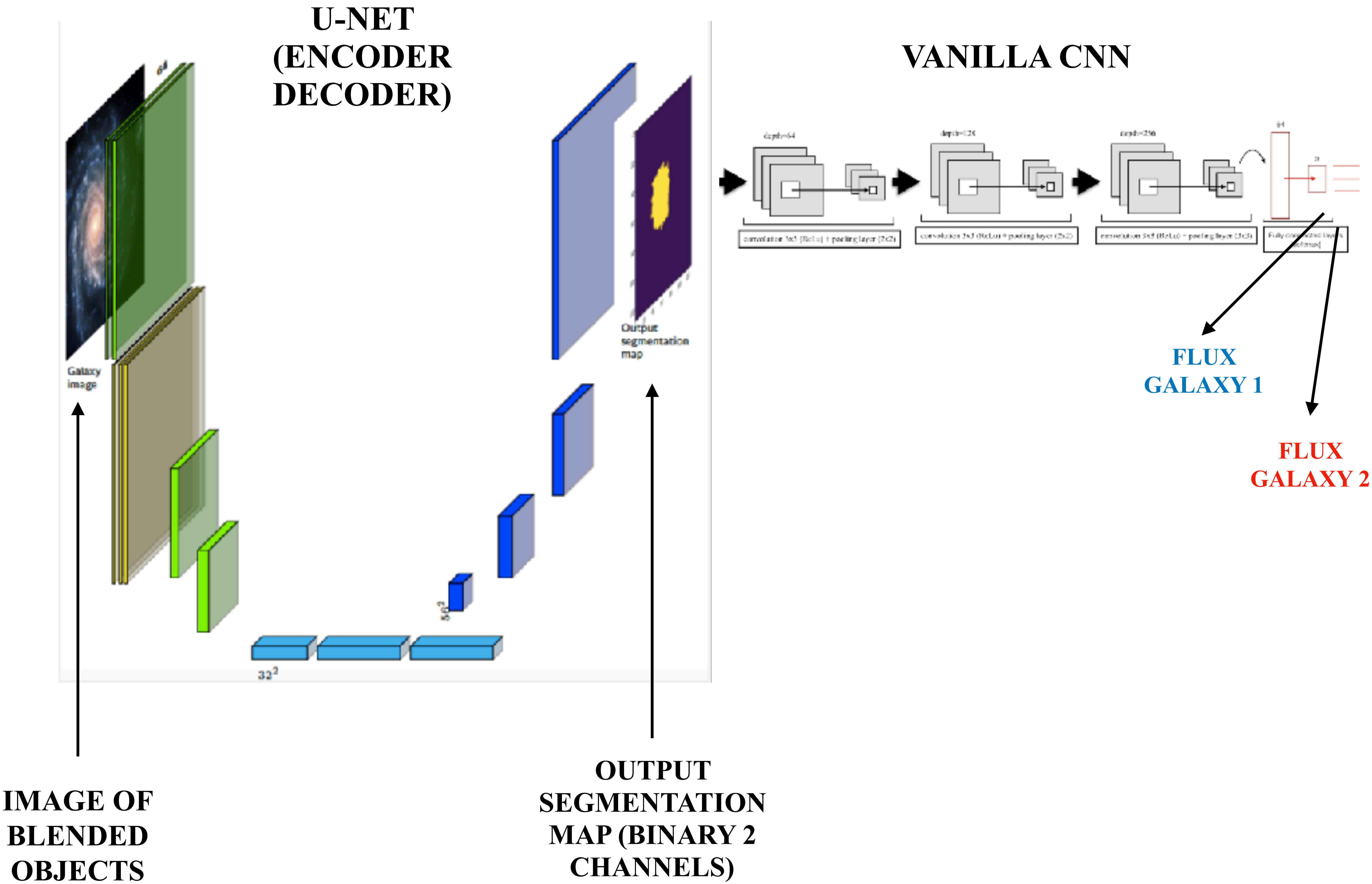
Boucaud+19



ISOLATED GALAXIES IN CANDELS ARTIFICIALLY BLENDED



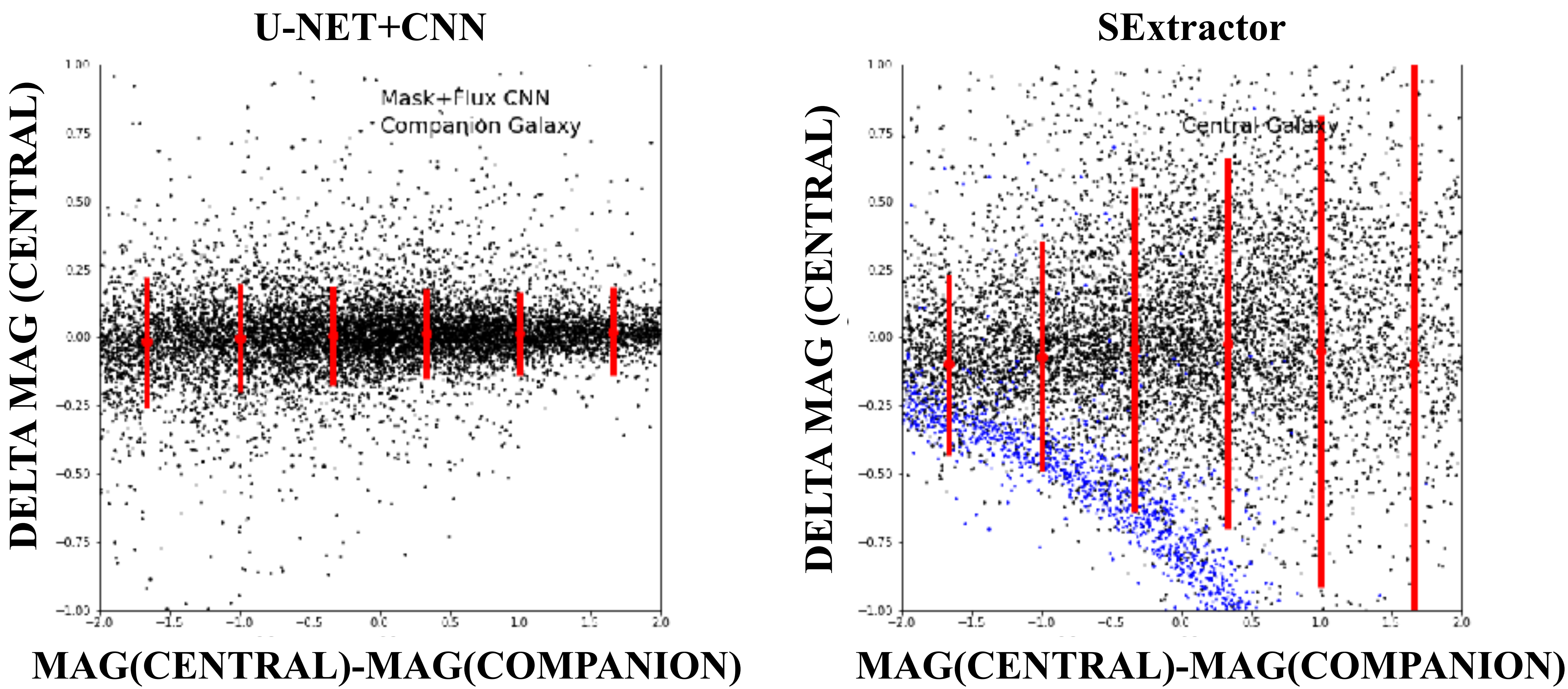
Boucaud+19



U-NET
(ENCODER
DECODER)
VANILLA CNN
Galaxy image
Output segmentation map
32²
FLUX GALAXY 1
FLUX GALAXY 2
IMAGE OF BLENDED OBJECTS
OUTPUT SEGMENTATION MAP (BINARY 2 CHANNELS)

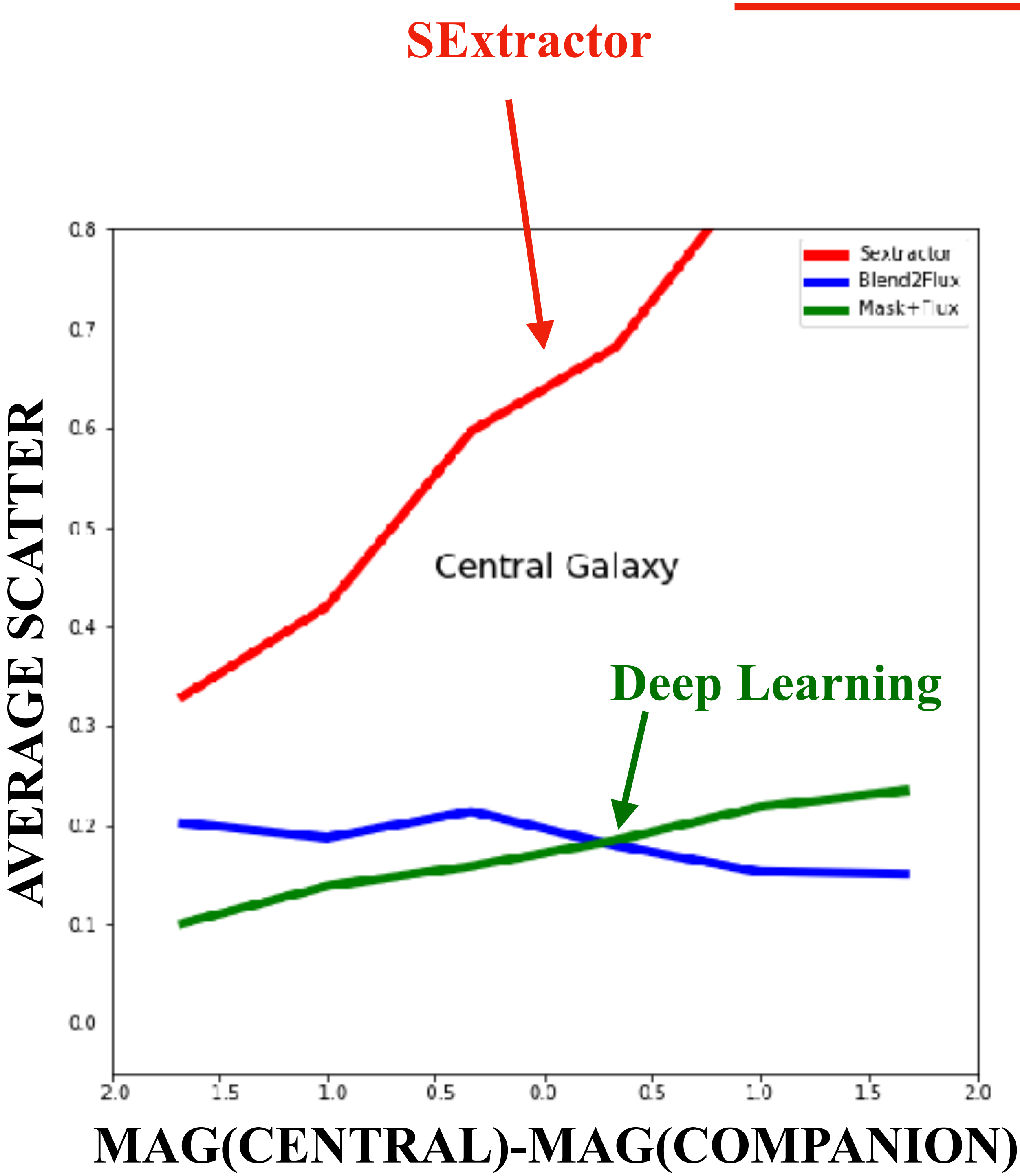
Boucaud+19

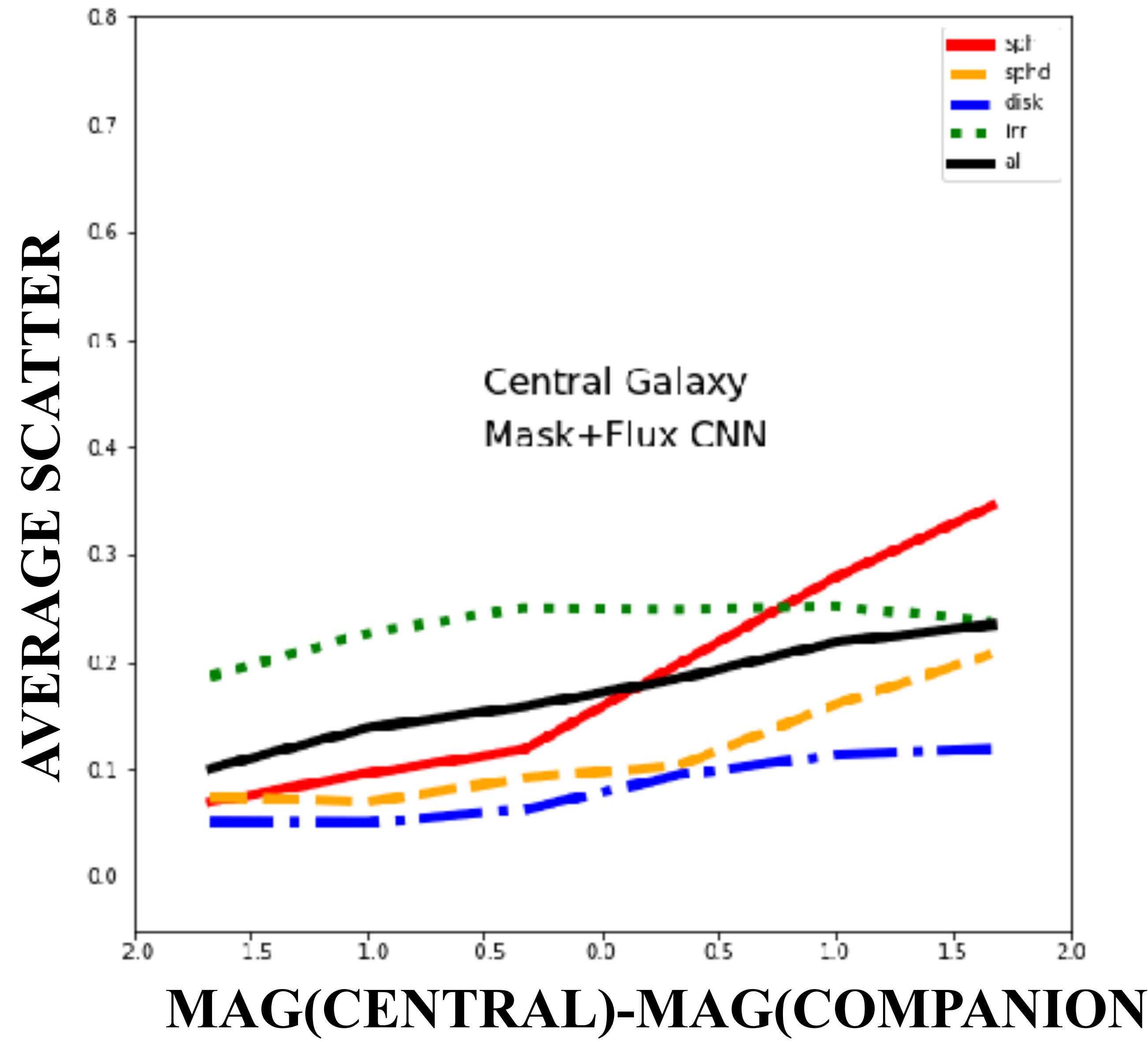
PHOTOMETRY OF BLENDED SOURCES



Boucaud, MHC+19

PHOTOMETRY OF BLENDED SOURCES





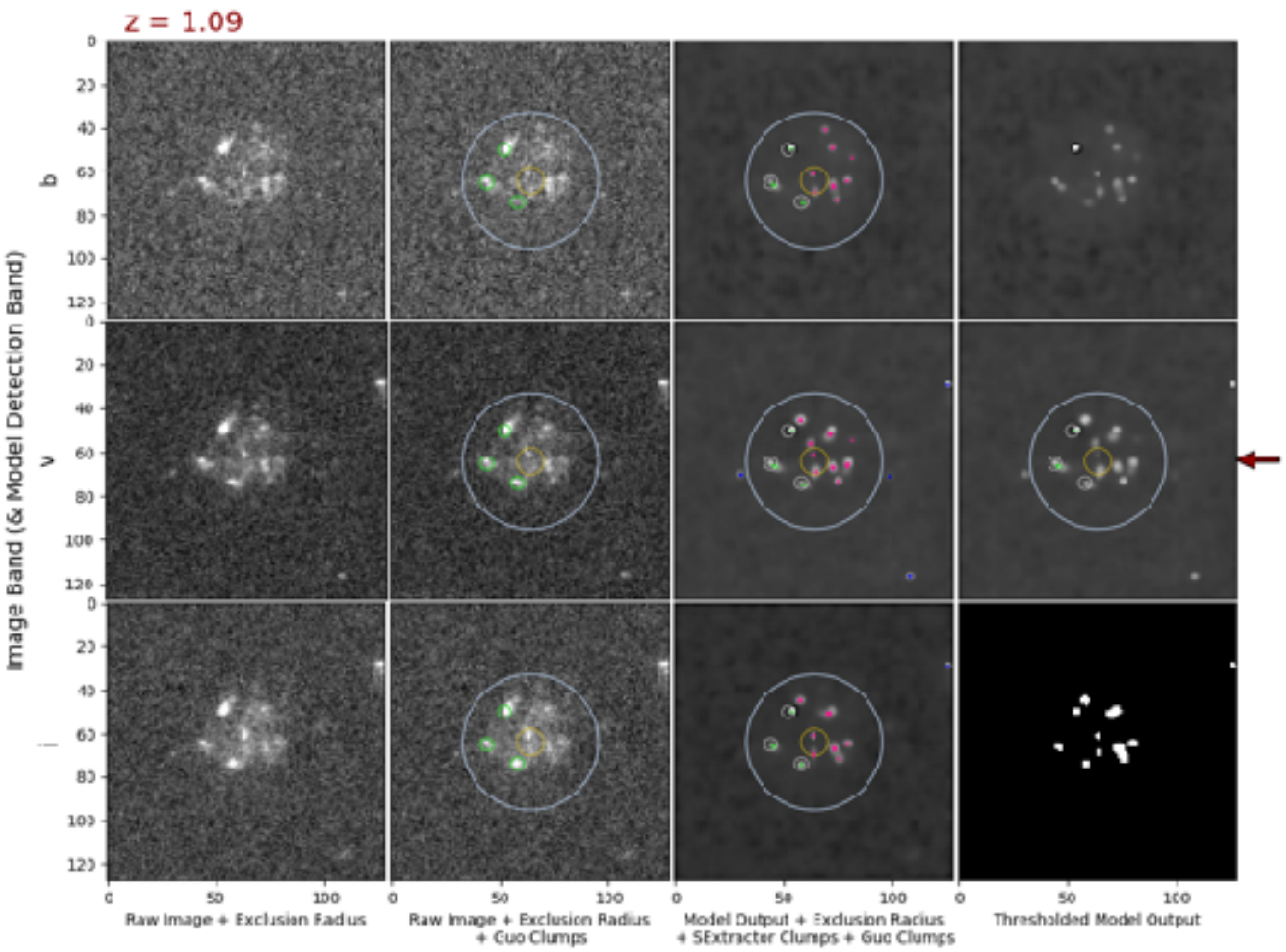
CONSTANT <0.2 SCATTER. A FACTOR 3-4 BETTER THAN SExtractor

REASONABLE BEHAVIOR FOR DIFFERENT MORPHOLOGIES

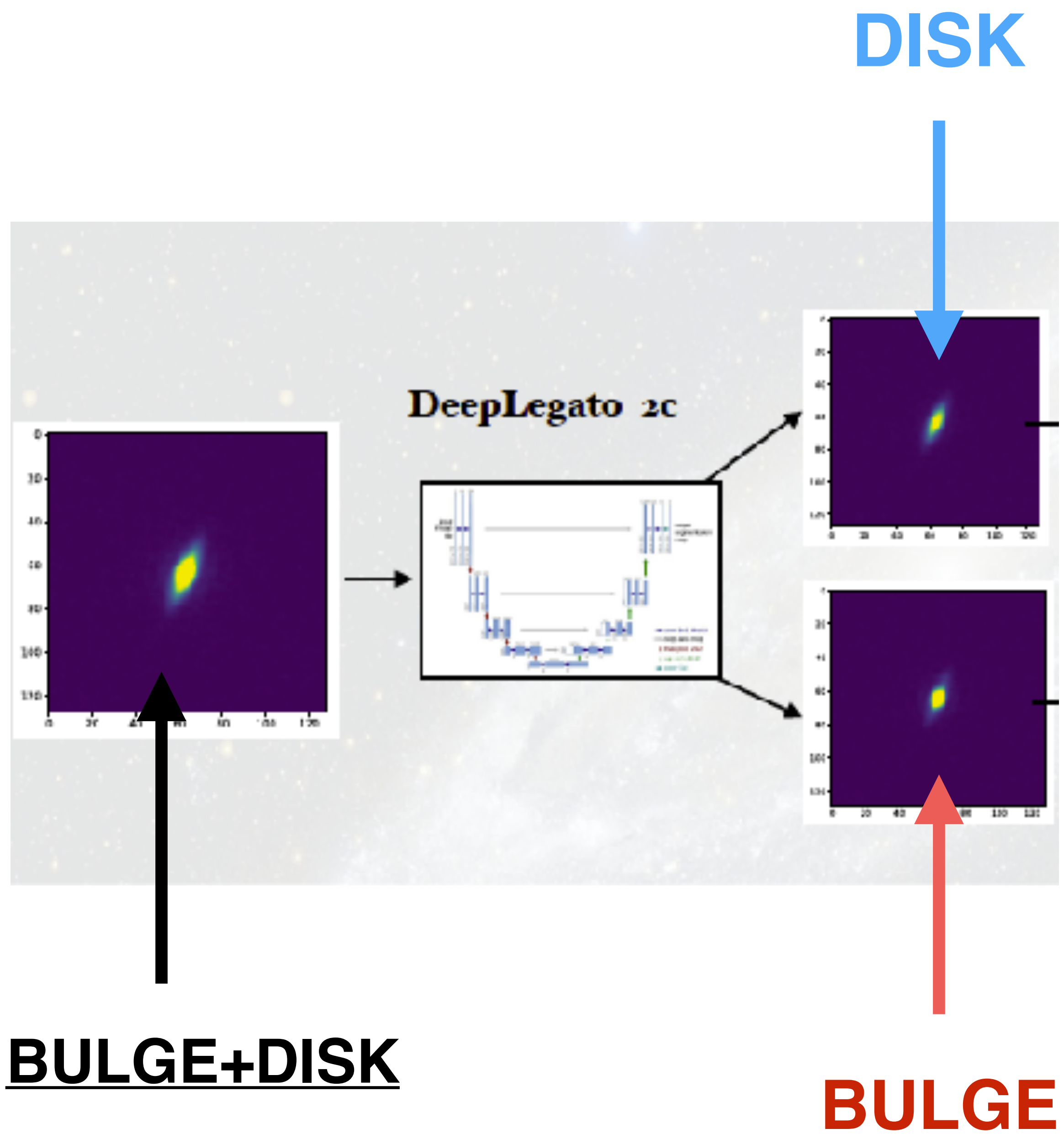
Boucaud, MHC+19

SIMILAR APPROACHES CAN BE USED FOR…

DETECTION OF CLUMPS IN HIGH REDSHIFT GALAXIES



BULGE-DISK DECOMPOSITIONS



Lee, MHC+19

Tuccillo, MHC+19

THEORY / SIMULATIONS

Illustris, EAGLE, Horizon-AGN …

[FULL 3D EVOLUTION HISTORY]

AI TO LINK THEORY AND OBSERVATION IN THE DATA SPACE

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”

ASSUMPTIONS OF MASS TO LIGHT CONVERSION
+ DUST
+PSF
+ NOISE

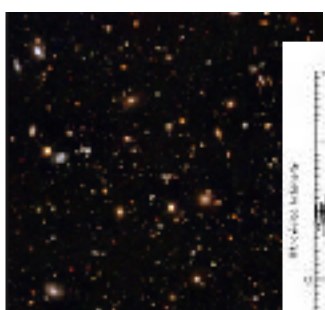
AI TO LINK THEORY AND OBSERVATION IN THE DATA SPACE

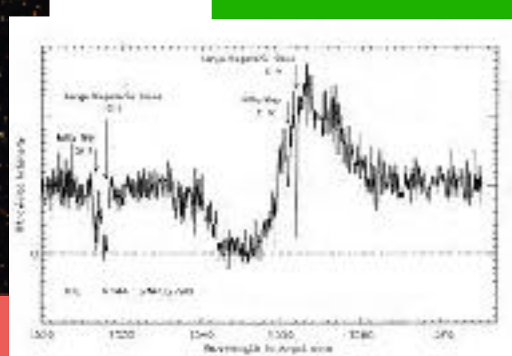
THEORY / SIMULATIONS

Illustris, EAGLE, Horizon-AGN …

[FULL 3D EVOLUTION HISTORY]

PROJECT HYDRO SIMS IN THE “OBSERVATIONAL PLANE”





MOCK OBSERVATIONS

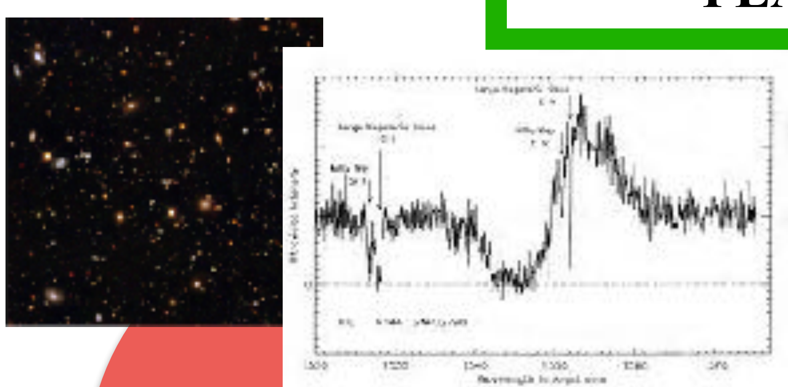
AI TO LINK THEORY AND OBSERVATION IN THE DATA SPACE

THEORY / SIMULATIONS

Illustris, EAGLE, Horizon-AGN …

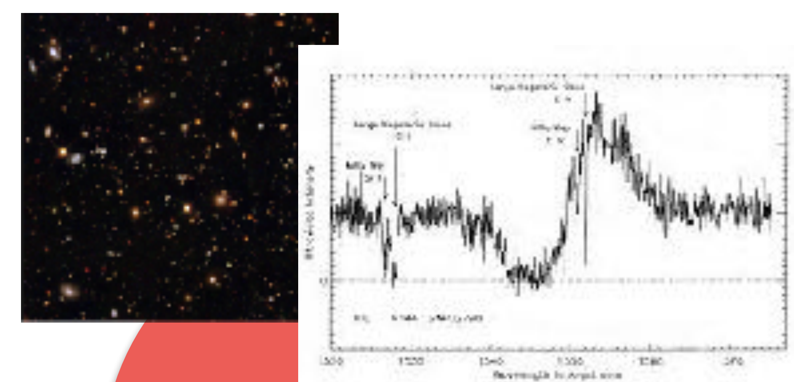
[FULL 3D EVOLUTION HISTORY]

PROJECT HYDRO SIMS IN THE “OBSERVATIONAL PLANE”



MOCK OBSERVATIONS

MACHINE (DEEP) LEARNING



OBSERVATIONS

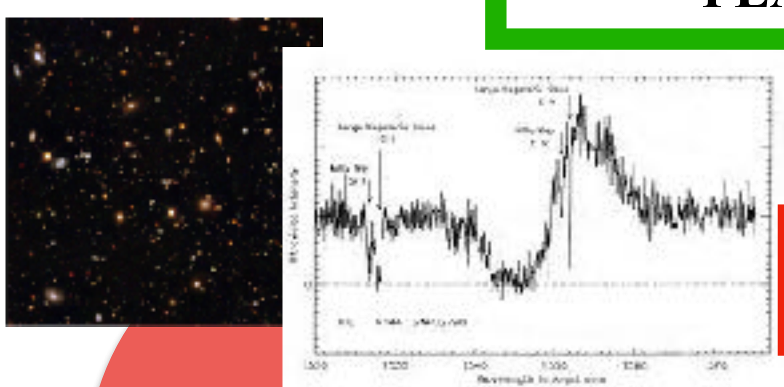
AI TO LINK THEORY AND OBSERVATION IN THE DATA SPACE

THEORY / SIMULATIONS

Illustris, EAGLE, Horizon-AGN …

[FULL 3D EVOLUTION HISTORY]

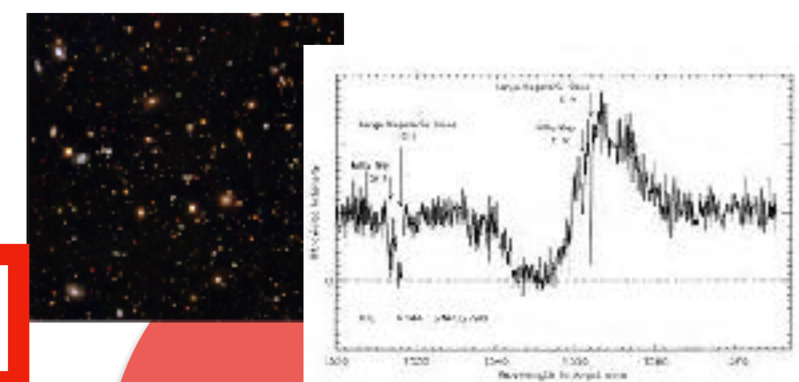
PROJECT HYDRO SIMS IN THE "OBSERVATIONAL PLANE"



TEST

MOCK OBSERVATIONS

TRAIN



OBSERVATIONS

MACHINE (DEEP) LEARNING

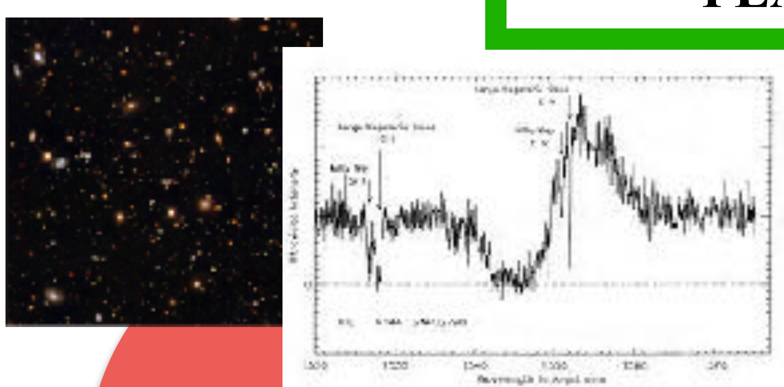
AI TO LINK THEORY AND OBSERVATION IN THE DATA SPACE

THEORY / SIMULATIONS

Illustris, EAGLE, Horizon-AGN …

[FULL 3D EVOLUTION HISTORY]

PROJECT HYDRO SIMS IN THE "OBSERVATIONAL PLANE"

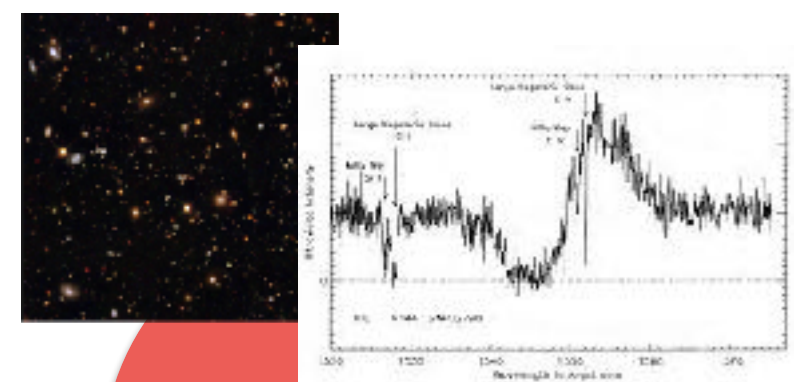


MOCK OBSERVATIONS

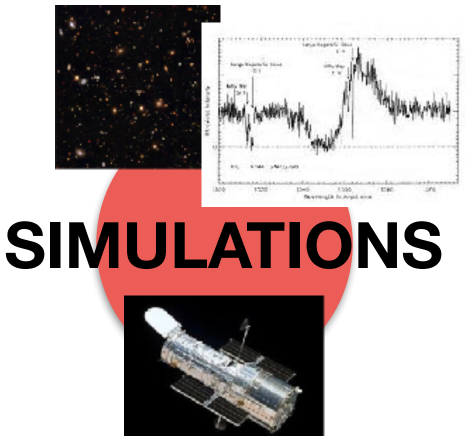
TRAIN

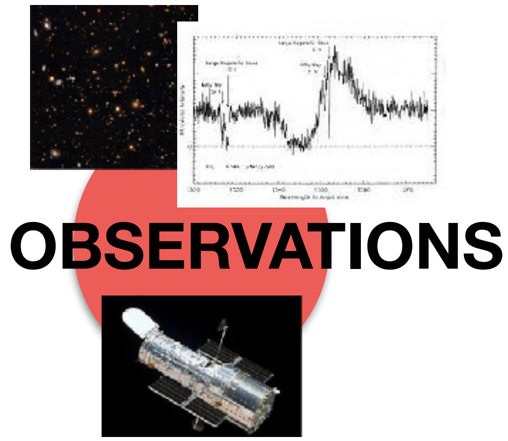
TEST

MACHINE (DEEP) LEARNING



OBSERVATIONS





HOW REALISTIC ARE ILLUSTRIS TNG MORPHOLOGIES?

SDSS DR7

TRAIN

Nair&Abraham10
GZOO

CNN MODEL
TO ESTIMATE GALAXY "VISUAL" MORPHOLOGY

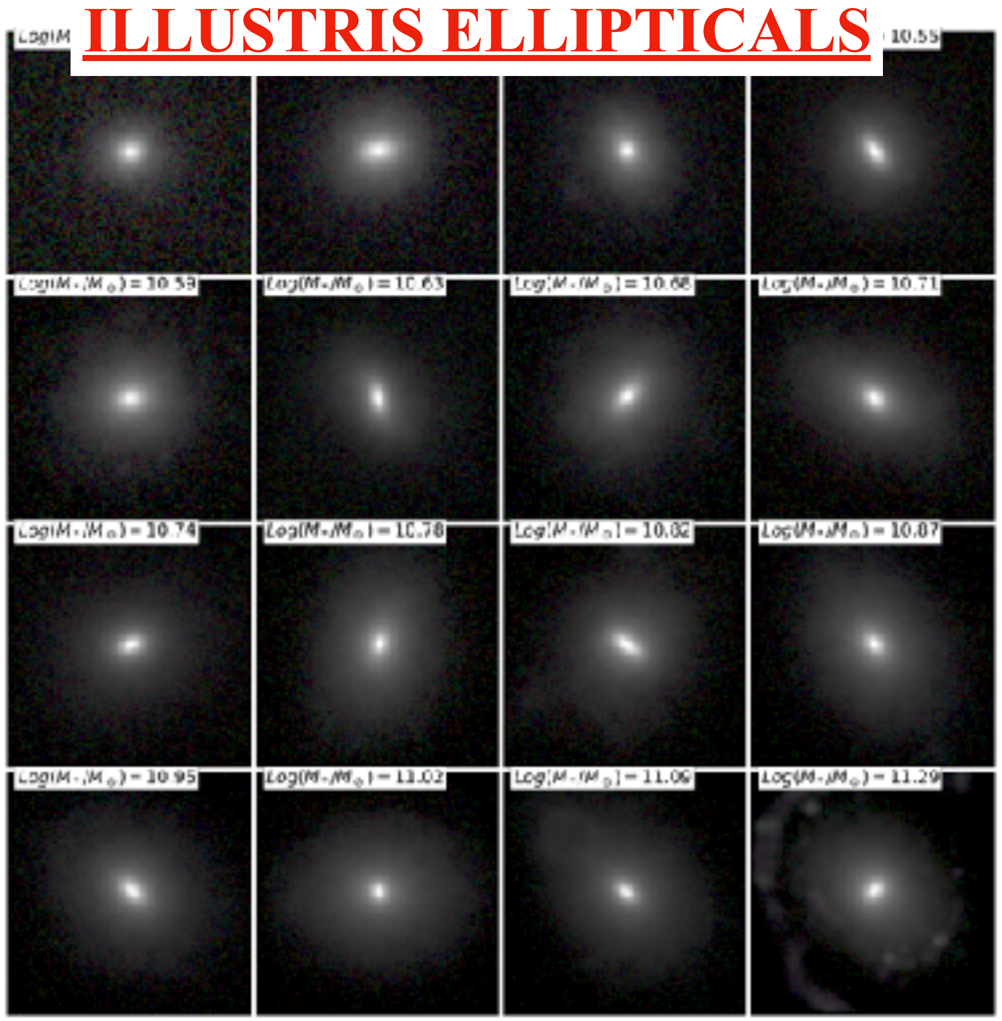
Dominguez-Sanchez+18

CLASSIFY

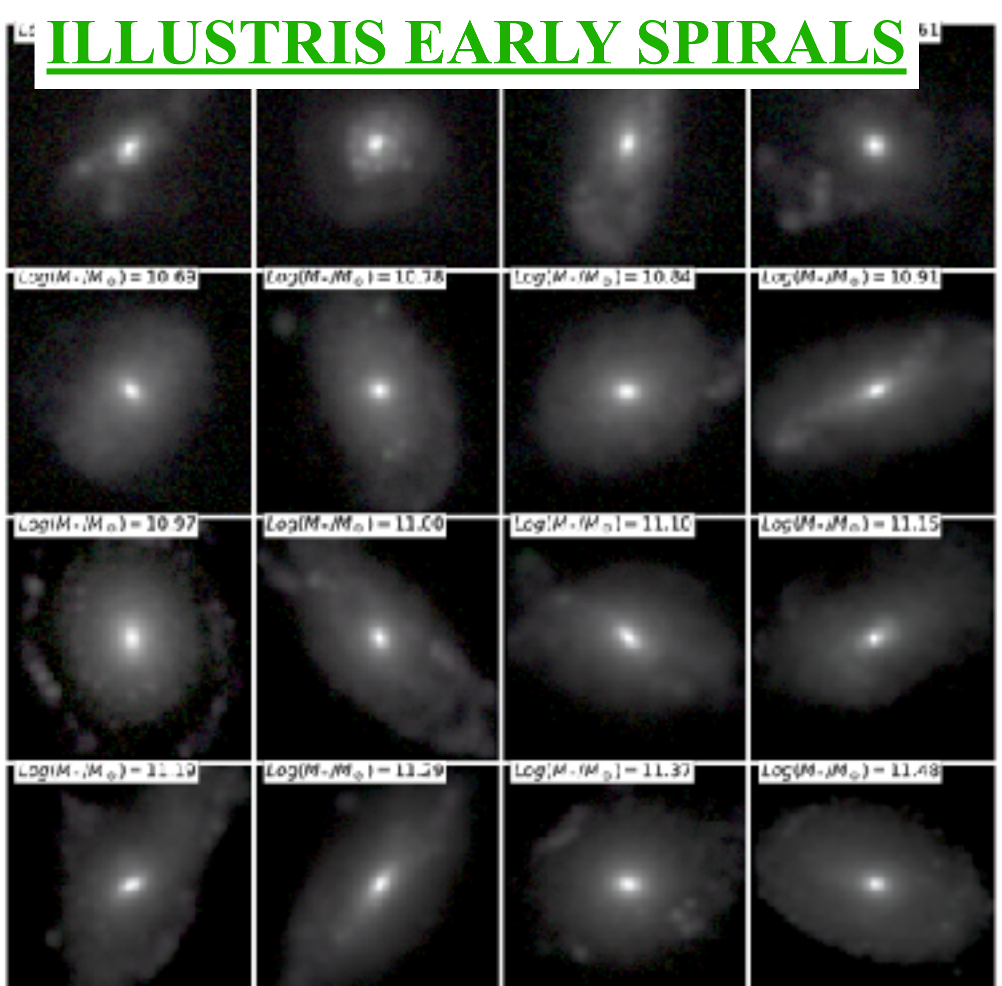
ILLUTRIS TNG

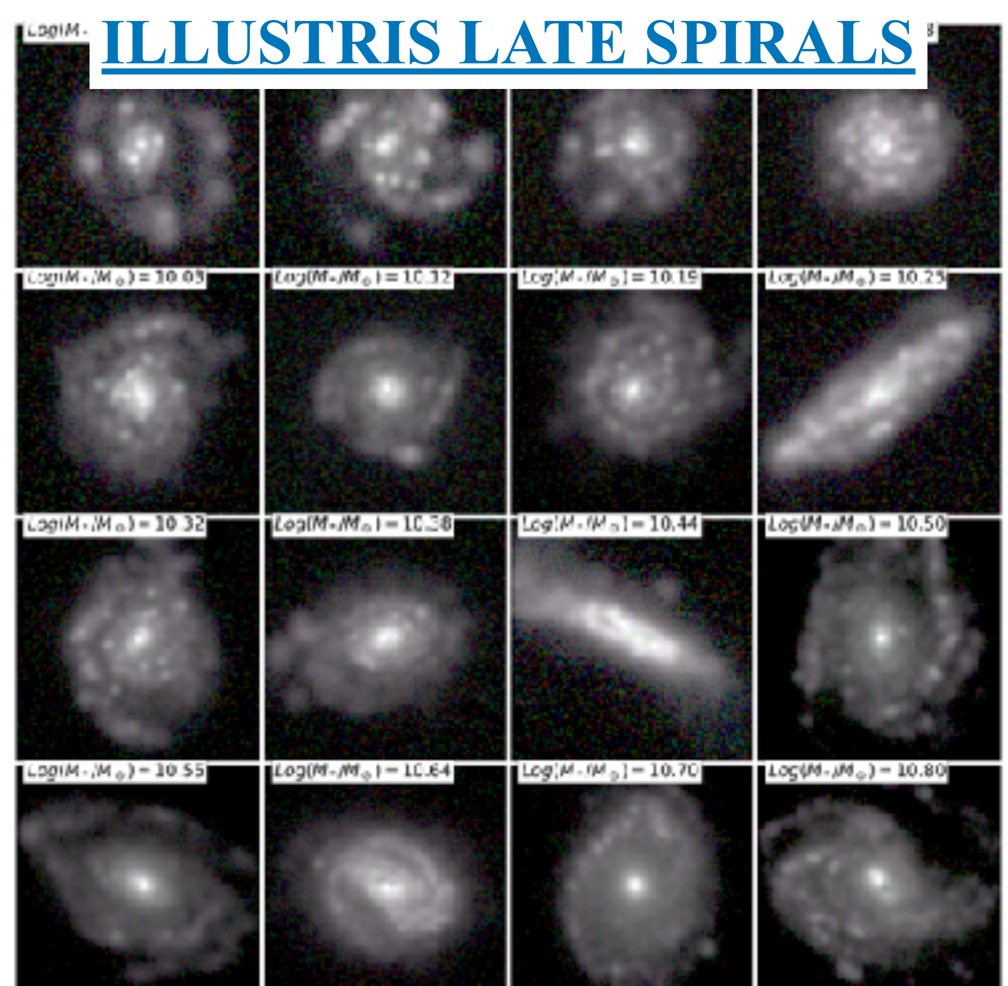
MOCK SDSS GALAXIES

MHC+19

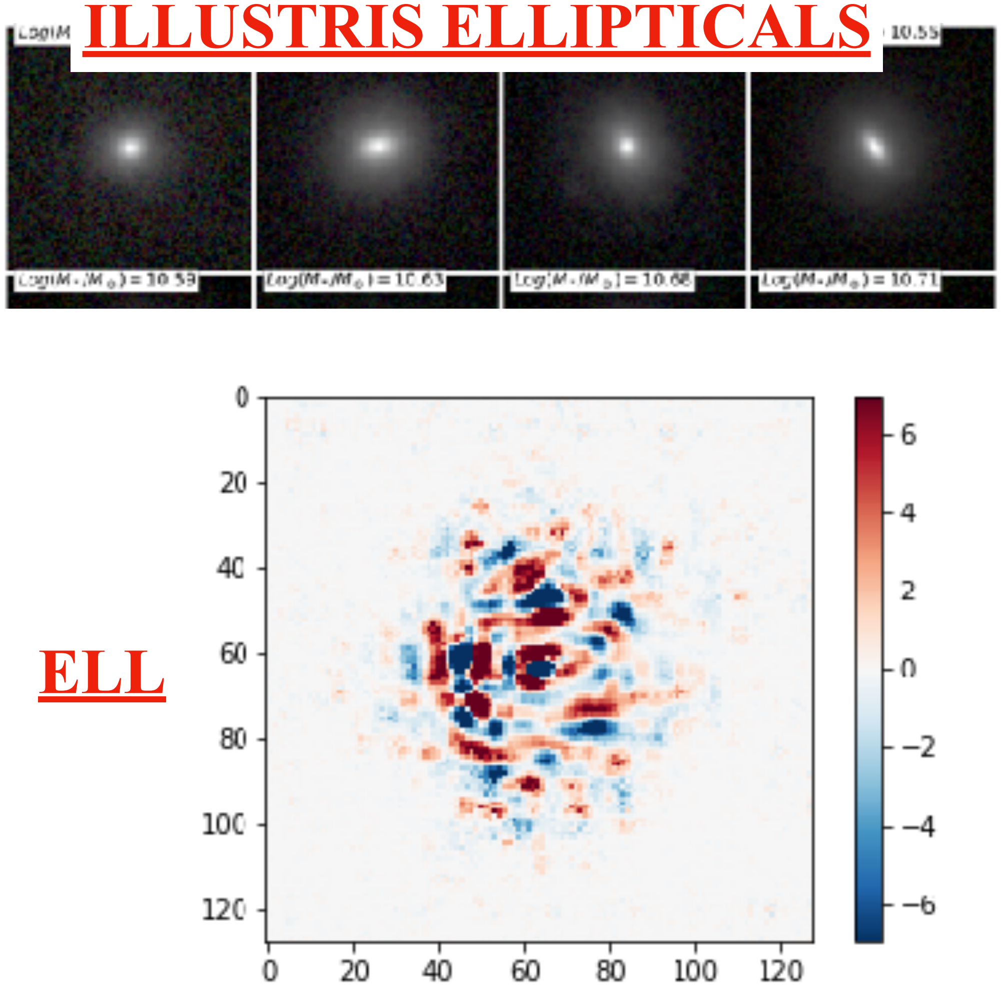

ILLUSTRIS ELLIPTICALS


ILLUSTRIS S0s

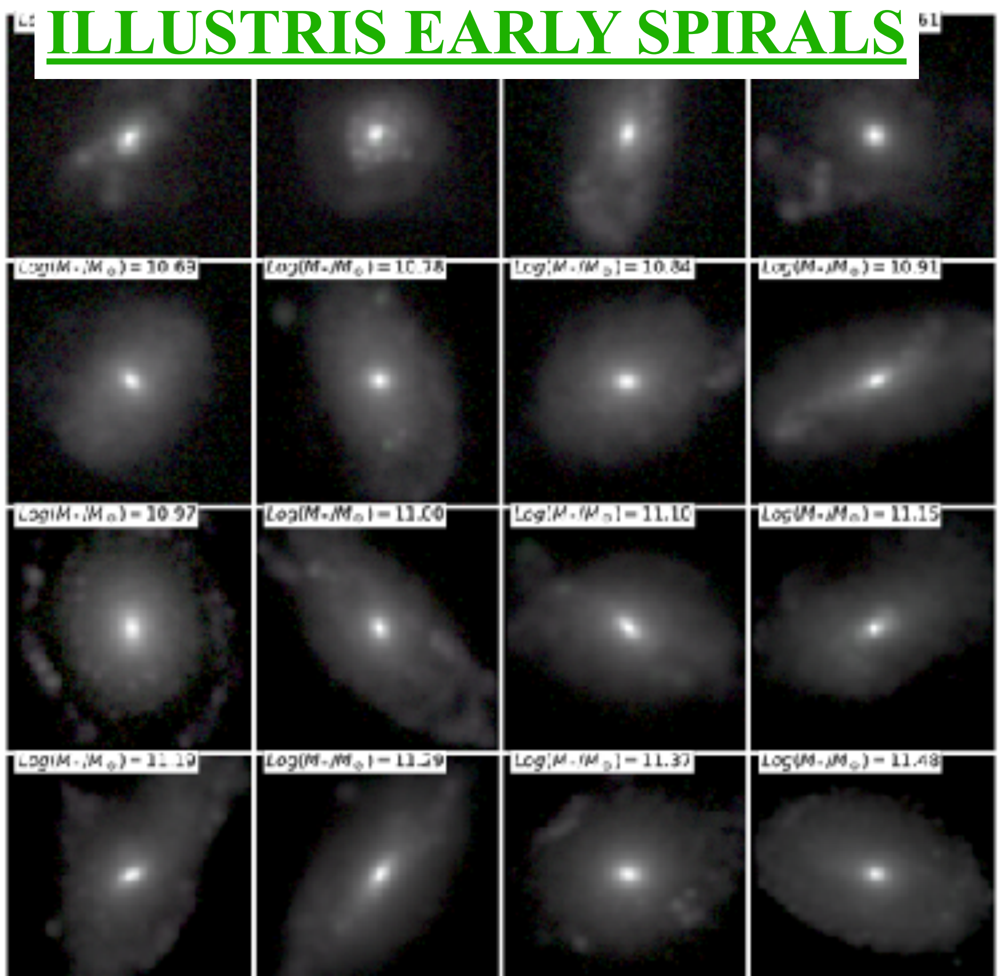

ILLUSTRIS EARLY SPIRALS

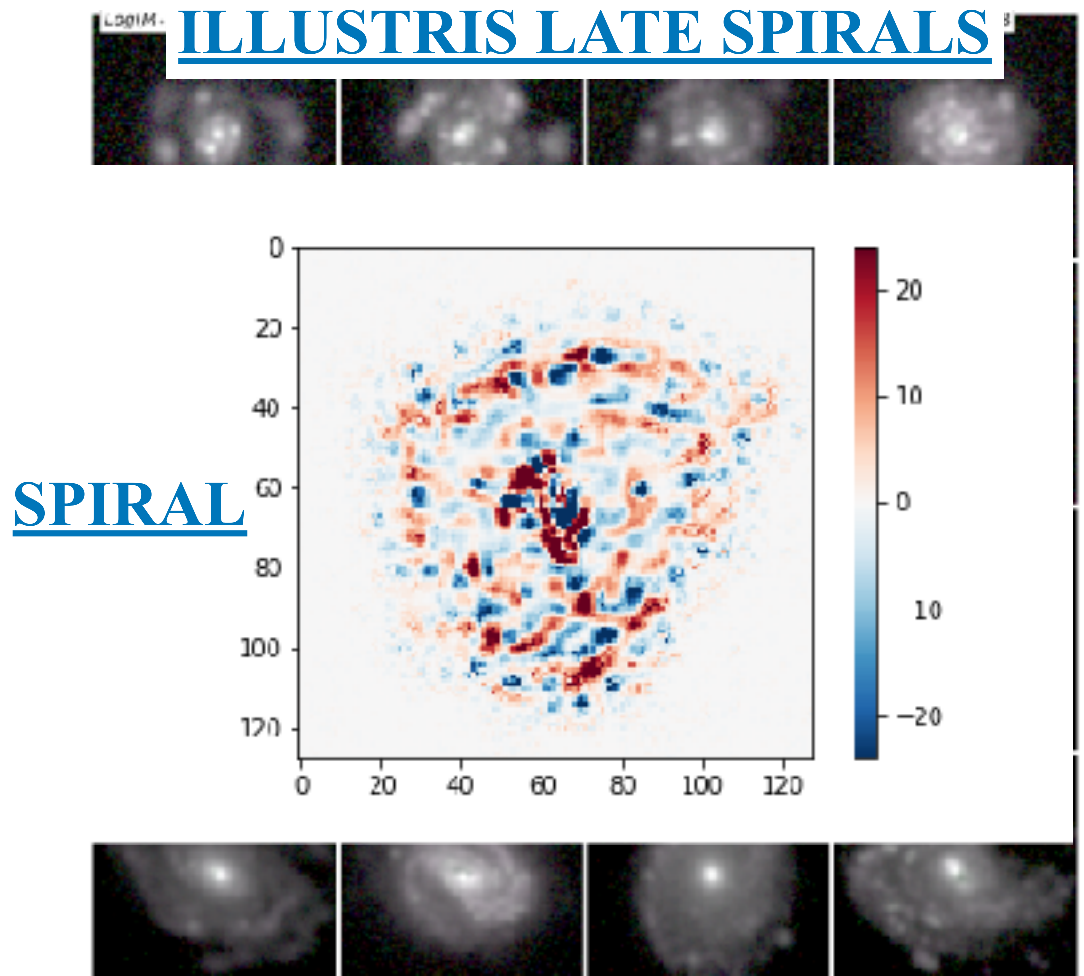

ILLUSTRIS LATE SPIRALS

MHC+19


ILLUSTRIS ELLIPTICALS
ELL

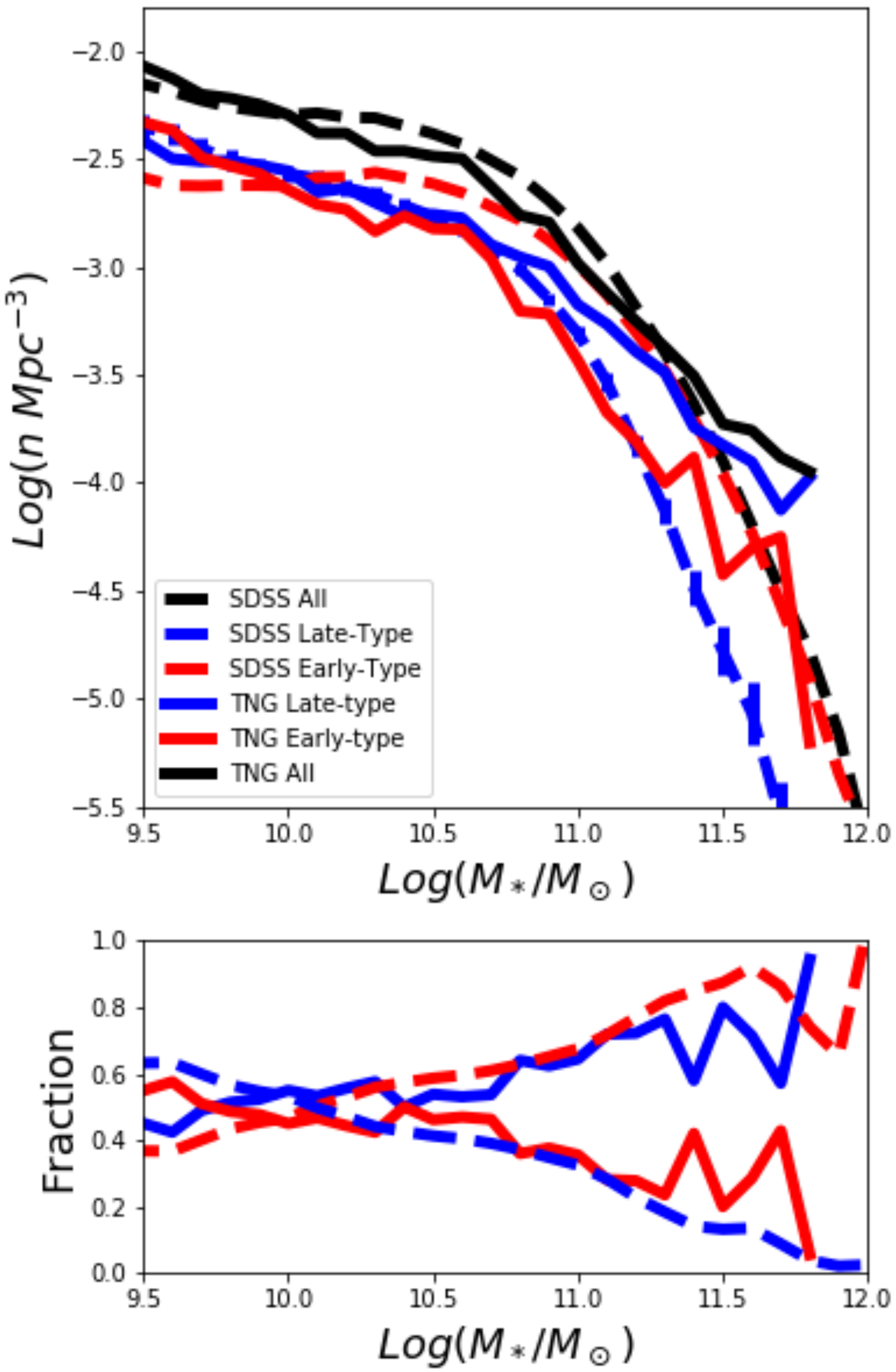

ILLUSTRIS S0s
ATTRIBUTION TECHNIQUES PROVIDE SOME INFORMATION BUT STILL RATHER LIMITED


ILLUSTRIS EARLY SPIRALS


ILLUSTRIS LATE SPIRALS
SPIRAL

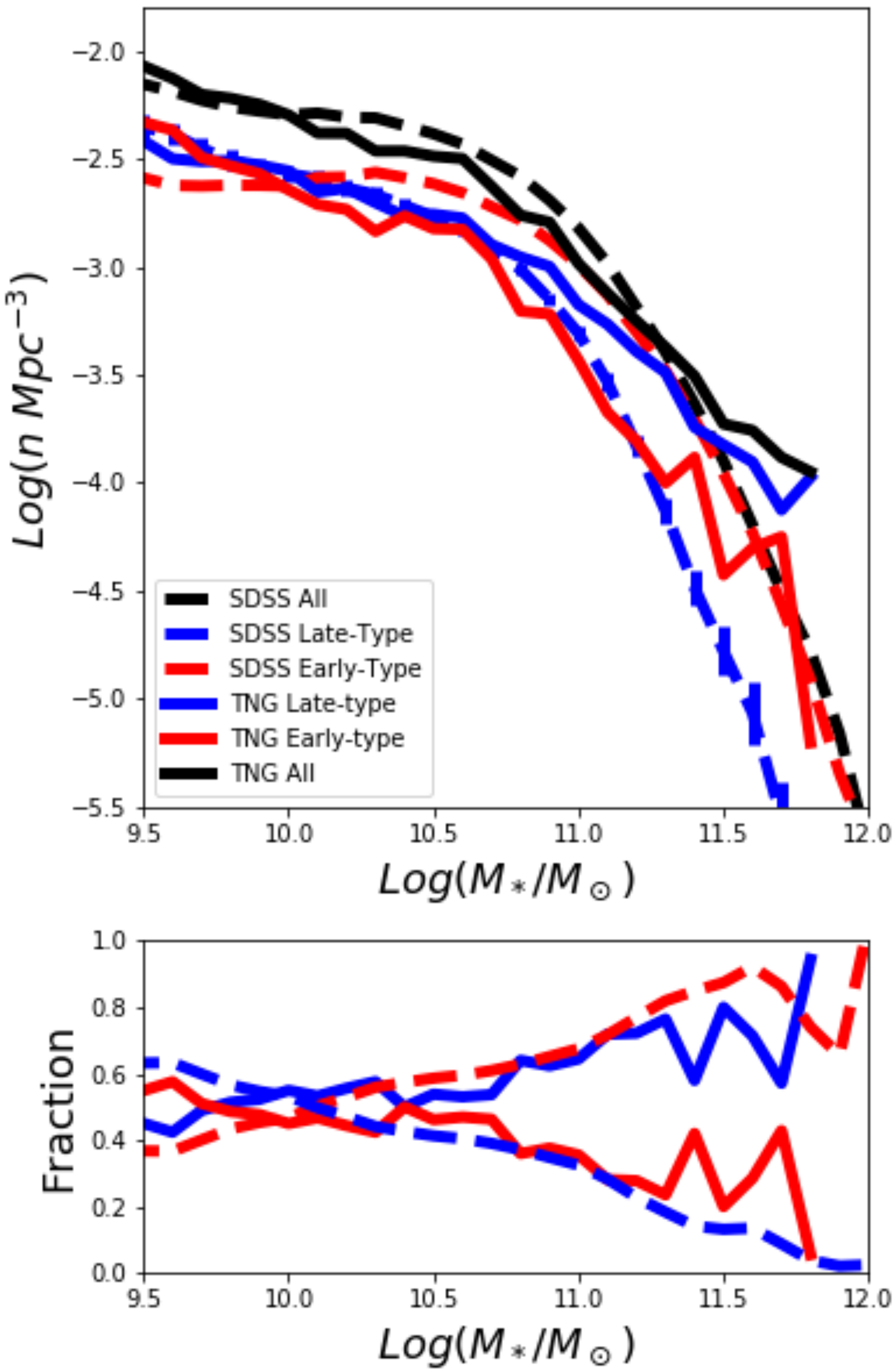
MHC+19

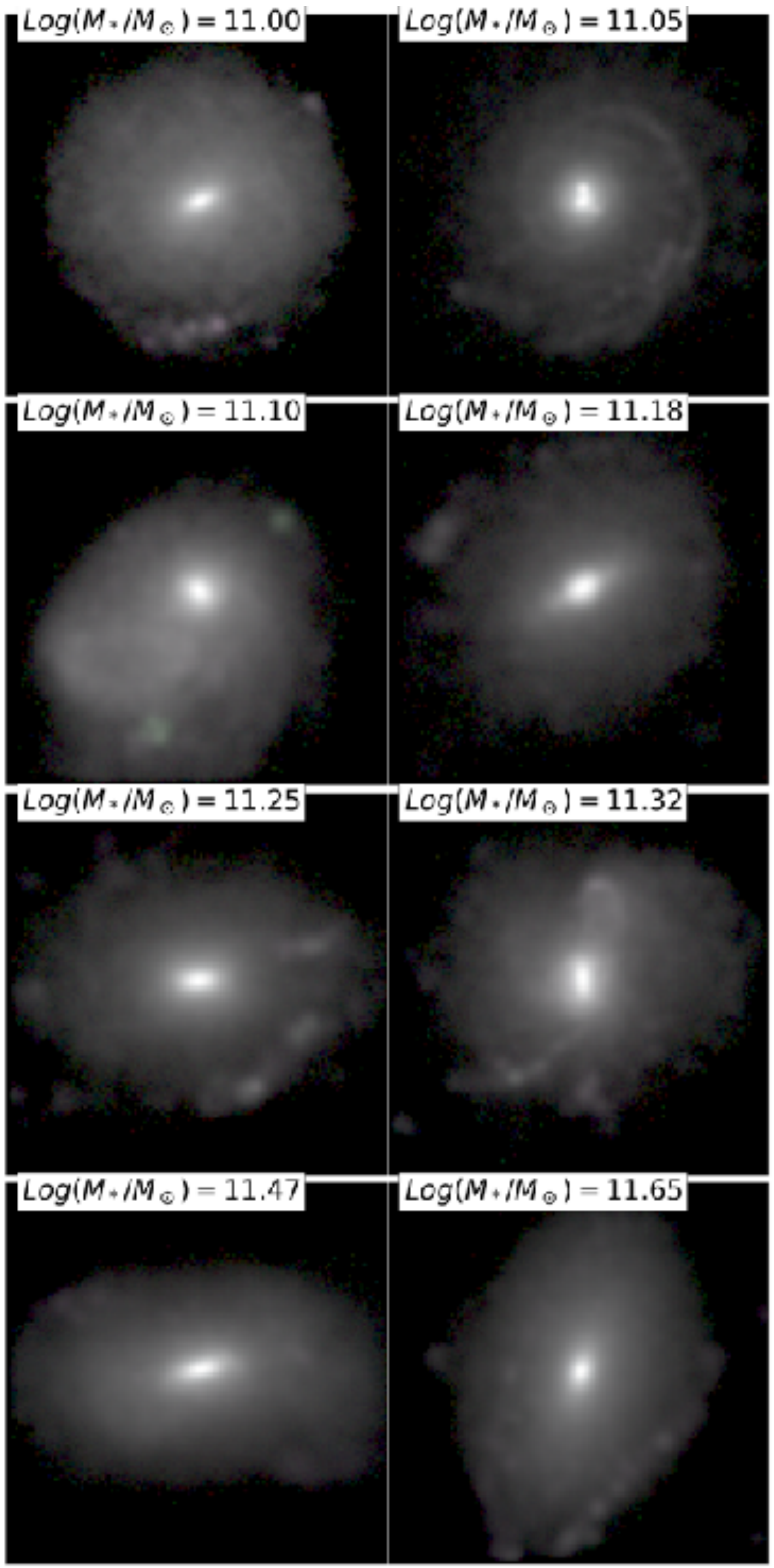
THE HIGH MASS END OF THE STELLAR MASS FUNCTION OF ILLUSTRIS TNG IS DOMINATED BY DISKS



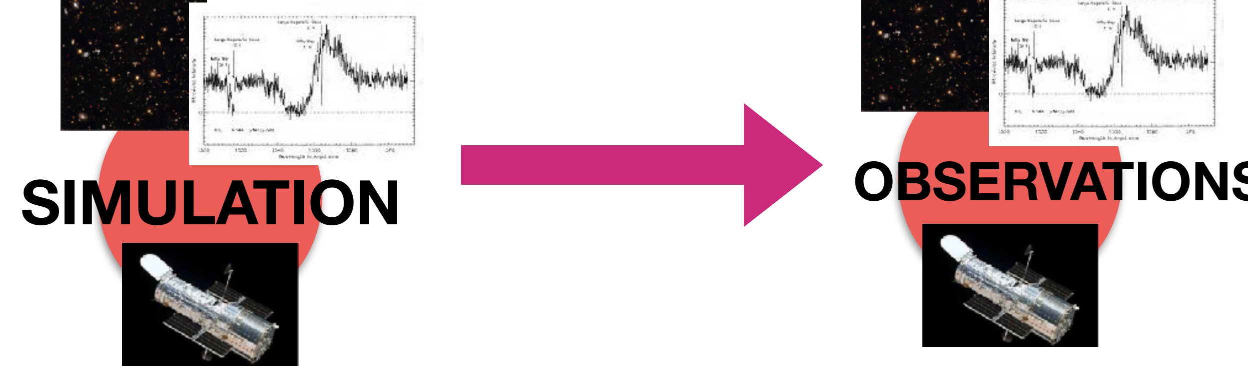
MHC+19

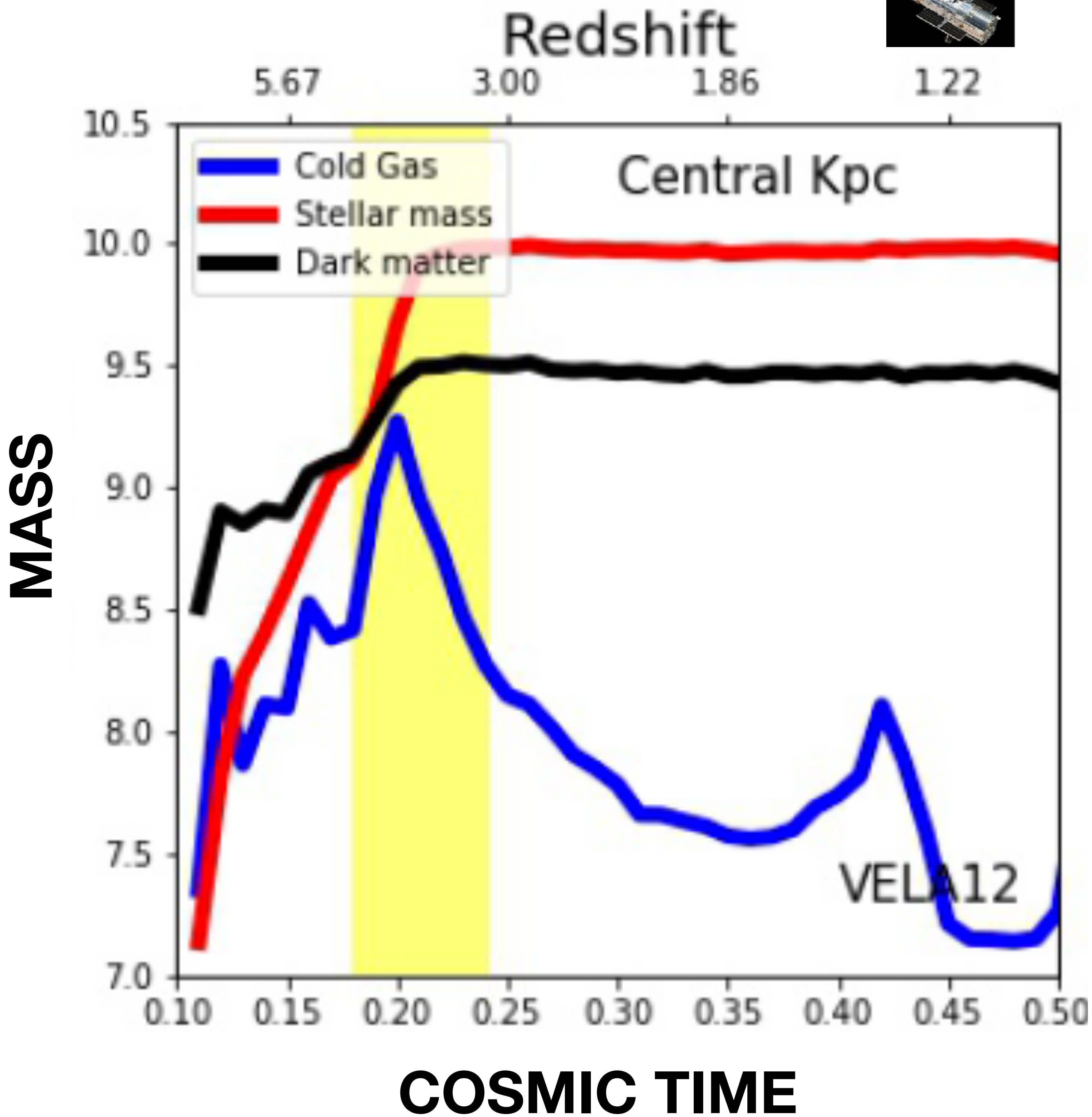
THE HIGH MASS END OF THE STELLAR MASS FUNCTION OF ILLUSTRIS TNG IS DOMINATED BY DISKS





IDENTIFYING COMPACTION IN THE VELA ZOOM-IN OF SIMULATIONS

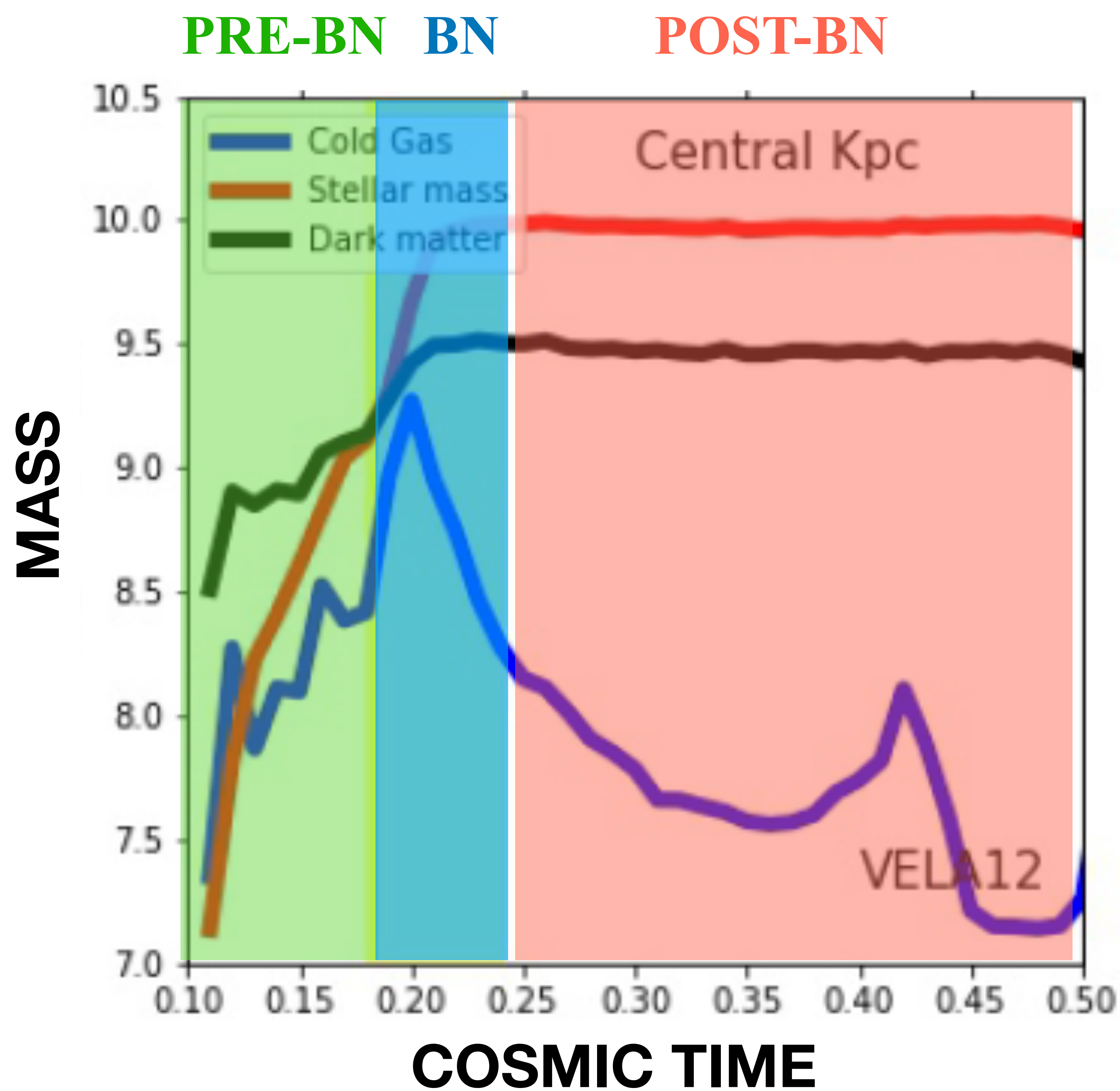




VELA suite of hydro simulations [Ceverino, Dekel+]

Dekel&Burkert+14, Ceverino+15, Zolotov+15,Tacchella+16a,b, Dekel+18

IDENTIFYING COMPACTION IN THE VELA ZOOM-IN OF SIMULATIONS

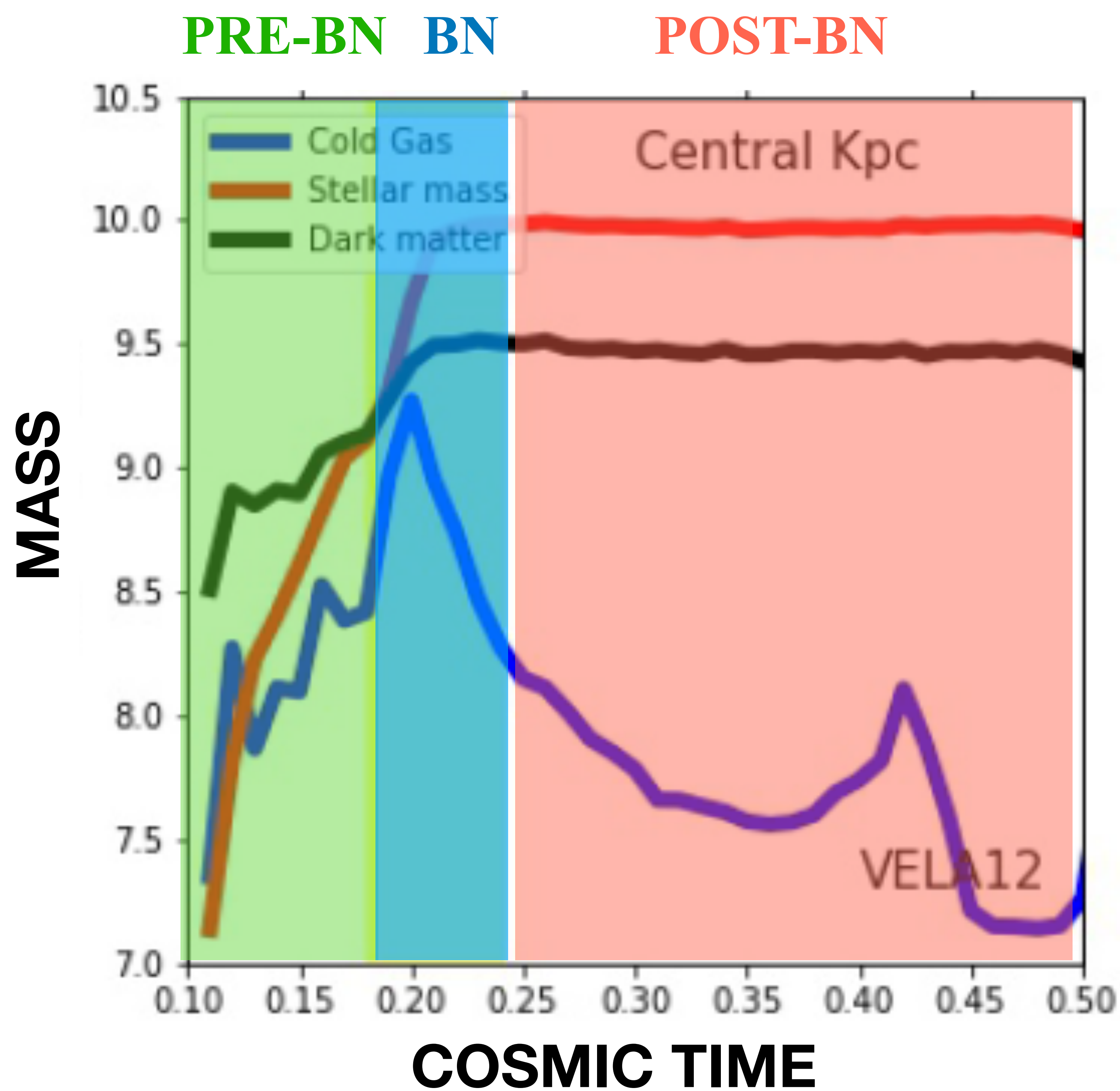


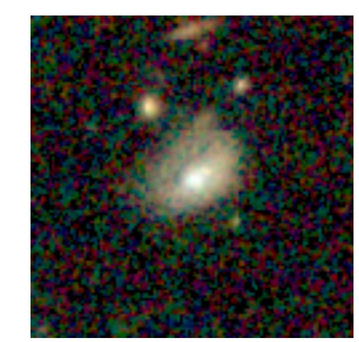
LABELLING EXCLUSIVELY BASED ON THE FULL HISTORY OF THE GALAXY FROM ADVANCED NUMERICAL SIMULATIONS BASED ON PHYSICAL PRINCIPLES

Dekel&Burkert+14, Ceverino+15, Zolotov+15,Tacchella+16a,b, Dekel+18

MHC+18

IN THE OBSERVATIONS WE ONLY HAVE ONE SNAPSHOT OF A GALAXY AT A GIVEN TIME

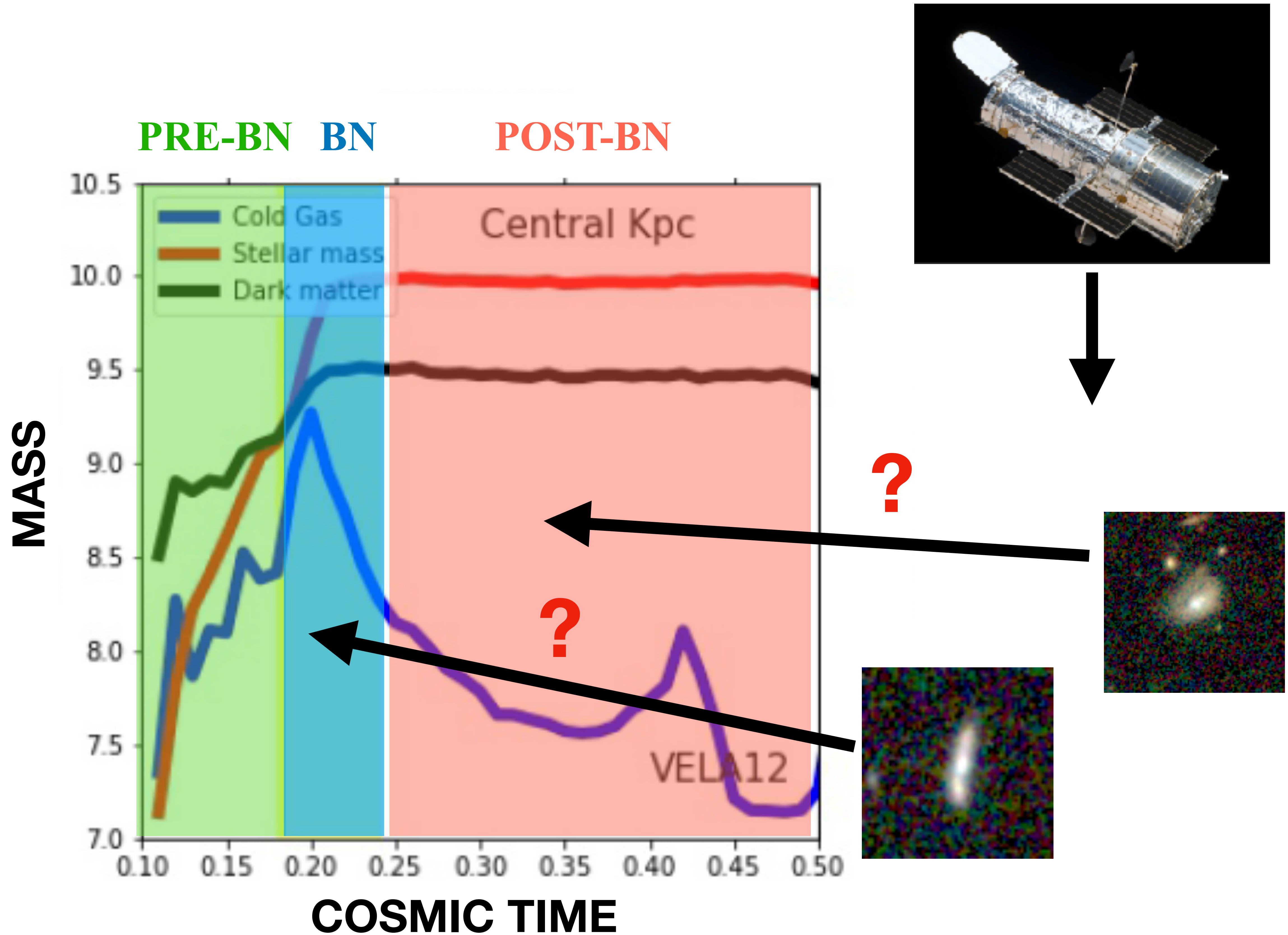




Dekel&Burkert+14, Ceverino+15, Zolotov+15,Tacchella+16a,b, Dekel+18

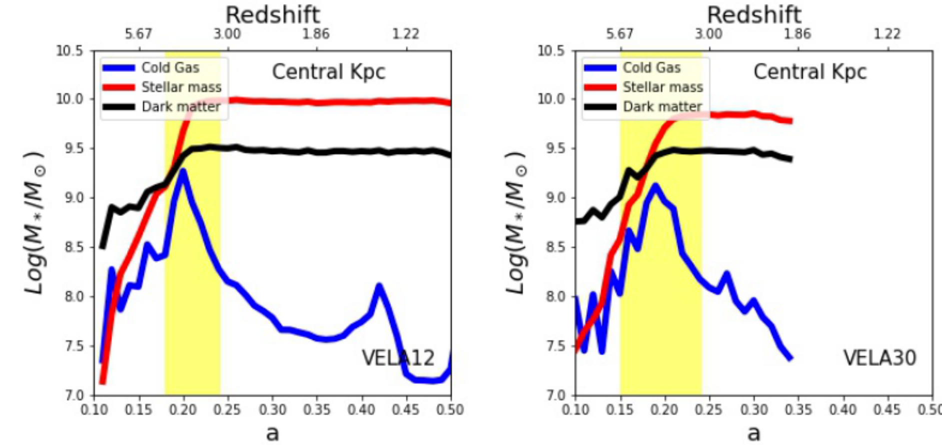
MHC+18

HOW CAN WE ESTIMATE THE PHASE FROM A UNIQUE IMAGE?

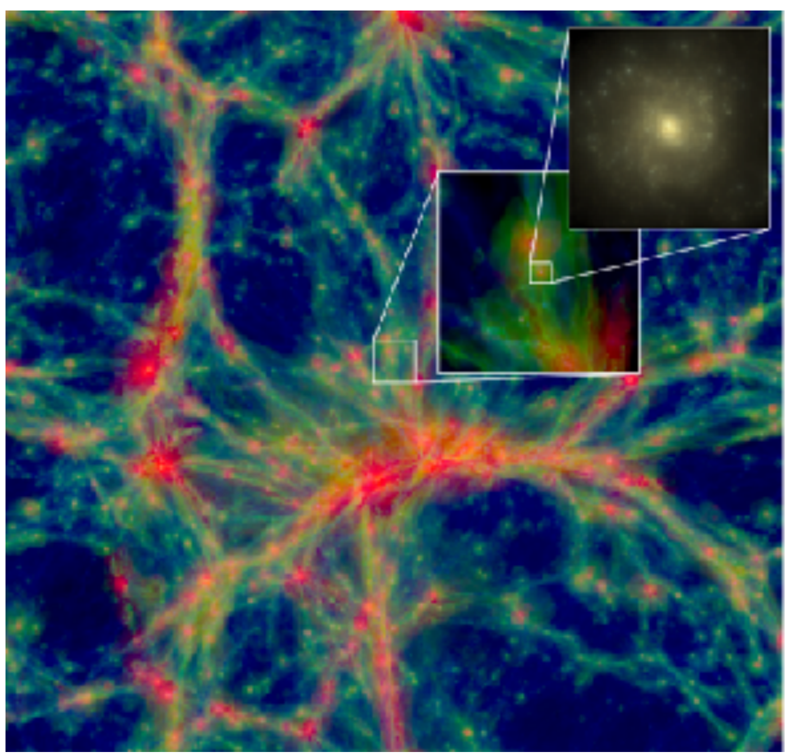


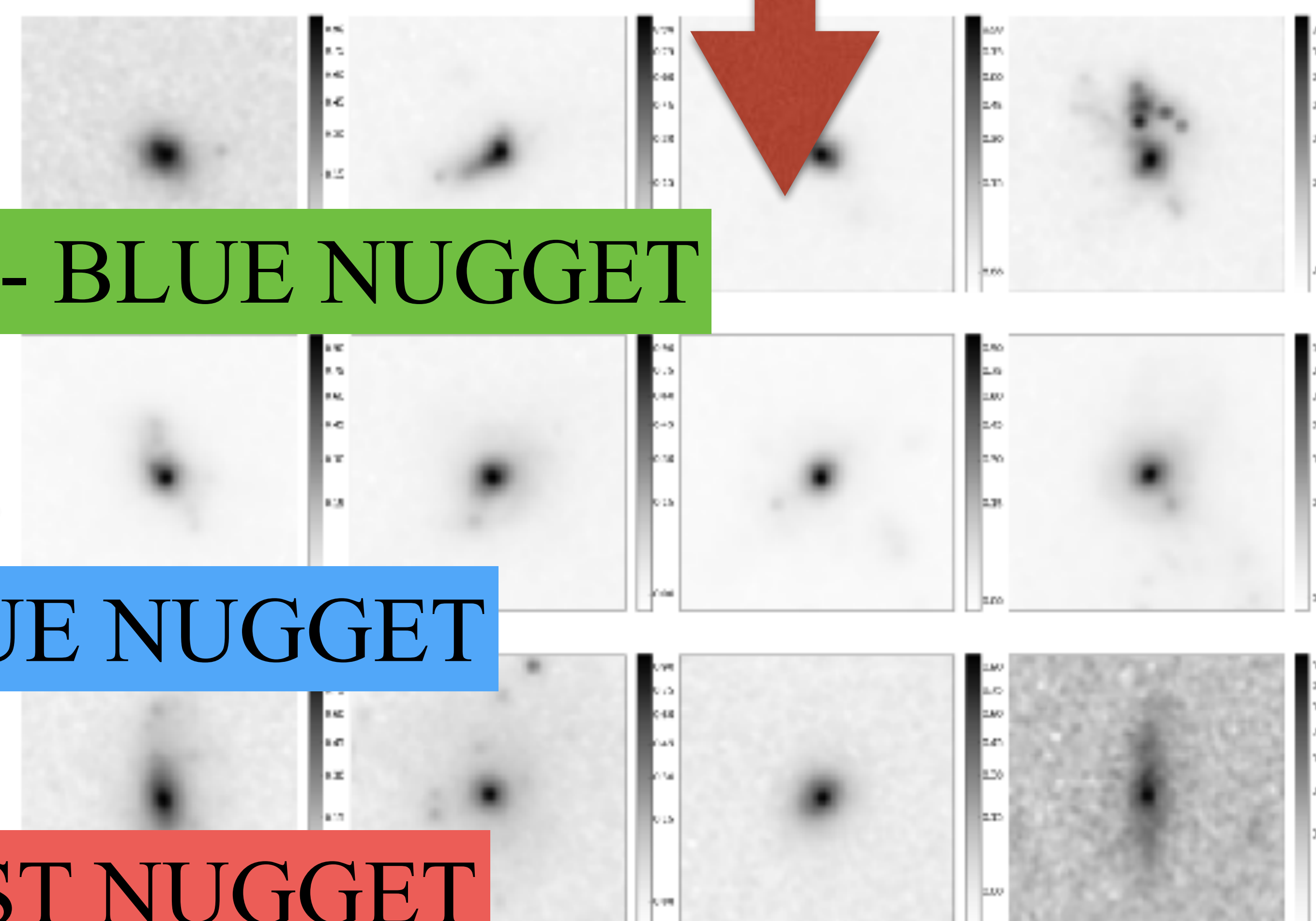
Dekel&Burkert+14, Ceverino+15, Zolotov+15,Tacchella+16a,b, Dekel+18

MHC+18



USE THE FORMATION HISTORY OF EACH GALAXY
TO LABEL IMAGES …





PRE- BLUE NUGGET

BLUE NUGGET

POST NUGGET

[COSMOLOGICAL SIMULATION]

PROJECT THE 3D OBJECT TO GENERATE “MOCK” IMAGES
AS OBSERVED BY HUBBLE

Pre-Blue-Nugget-Stage | Blue-Nugget-Stage | Post-Blue-Nugget-Stage

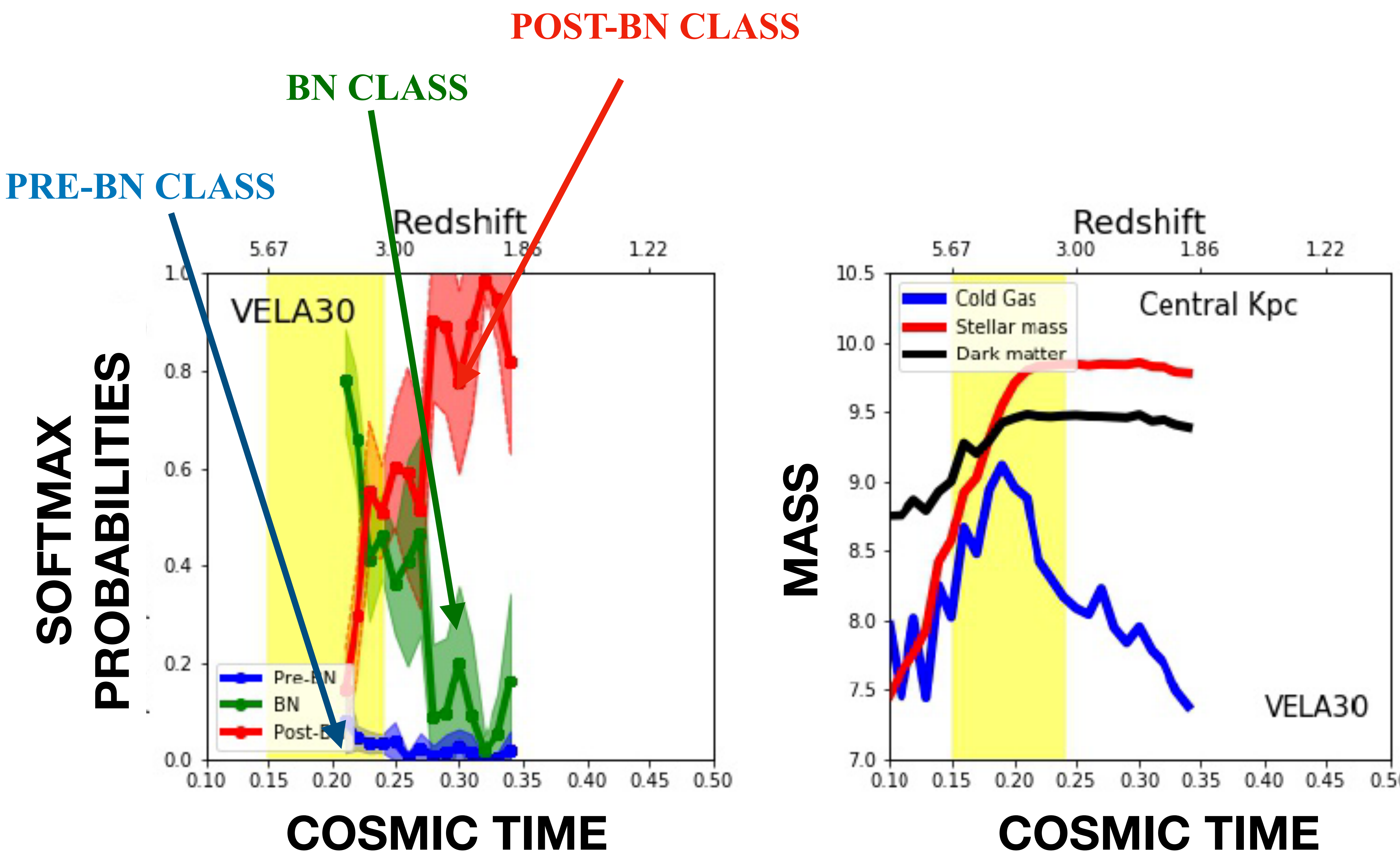
VELA HIGH RESOLUTION

VELA HST RESOLUTION

CANDELS

MHC+18

THE OUPUT SOFTMAX PROBABILITY TRACKS THE GAS MASS

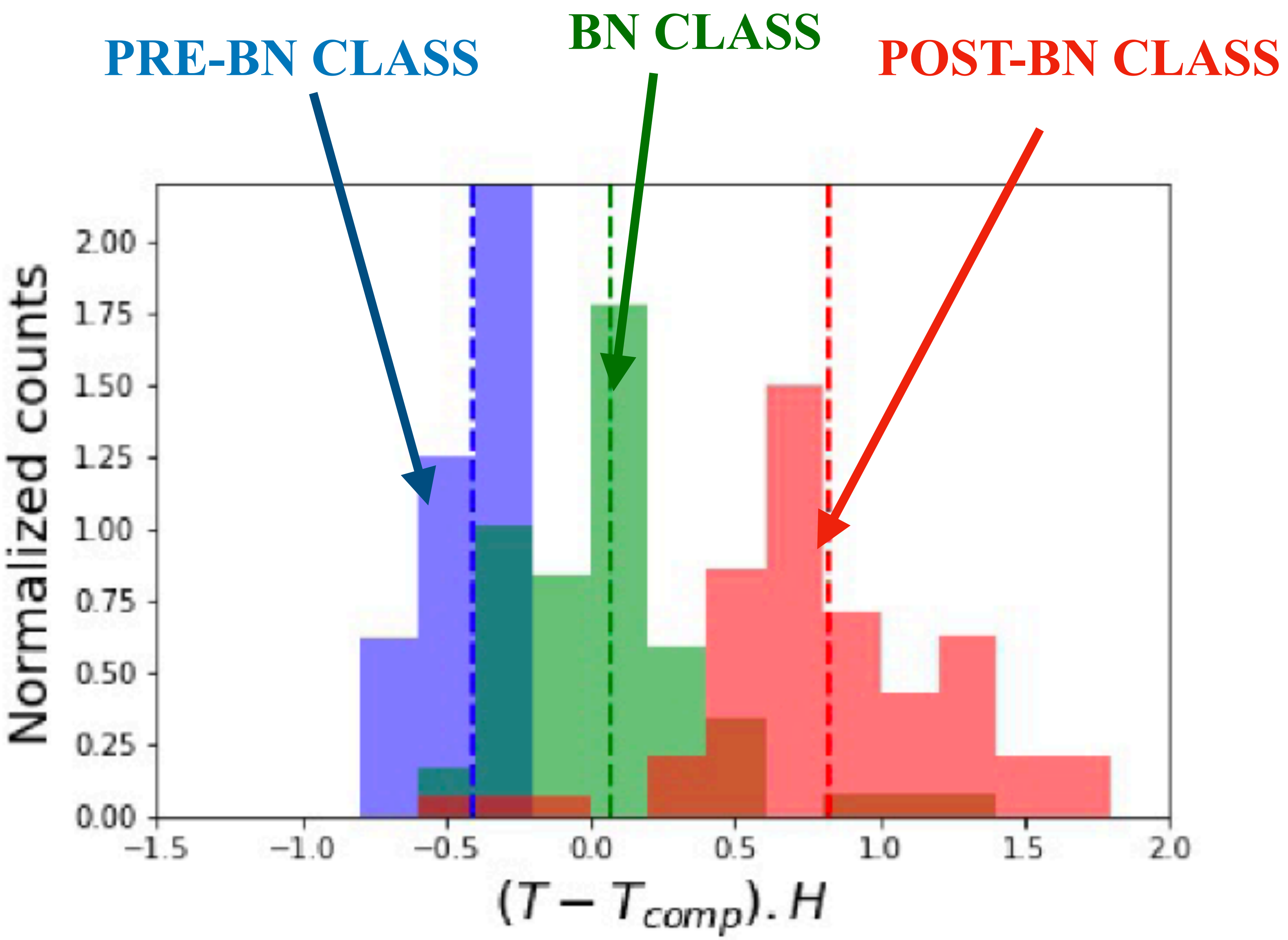


<u>DEEP LEARNING</u>

<u>SIMULATION</u>

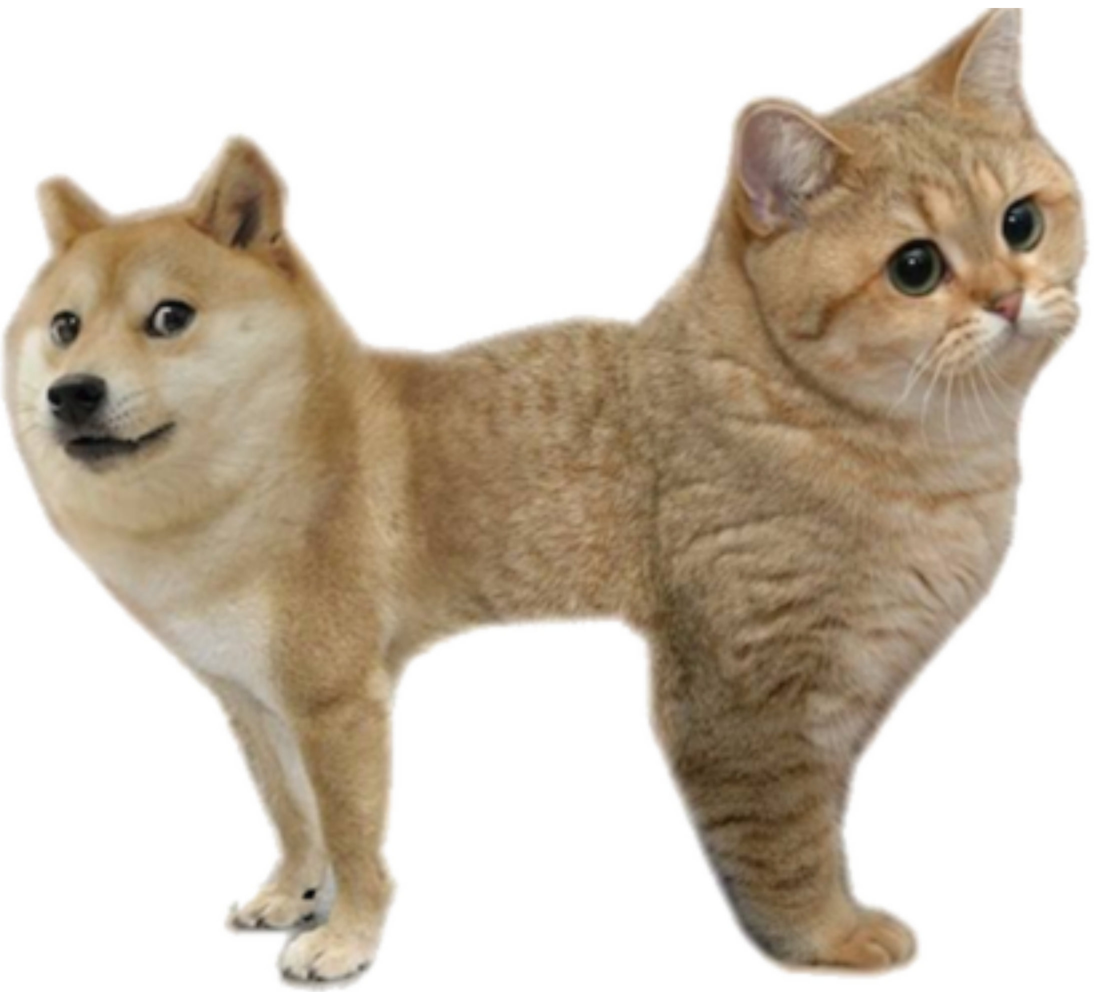
MHC+18

CONSTRAINTS ON THE OBSERVABILITY TIMESCALE



MHC+18

CAPTURING THE MODEL UNCERTAINTY



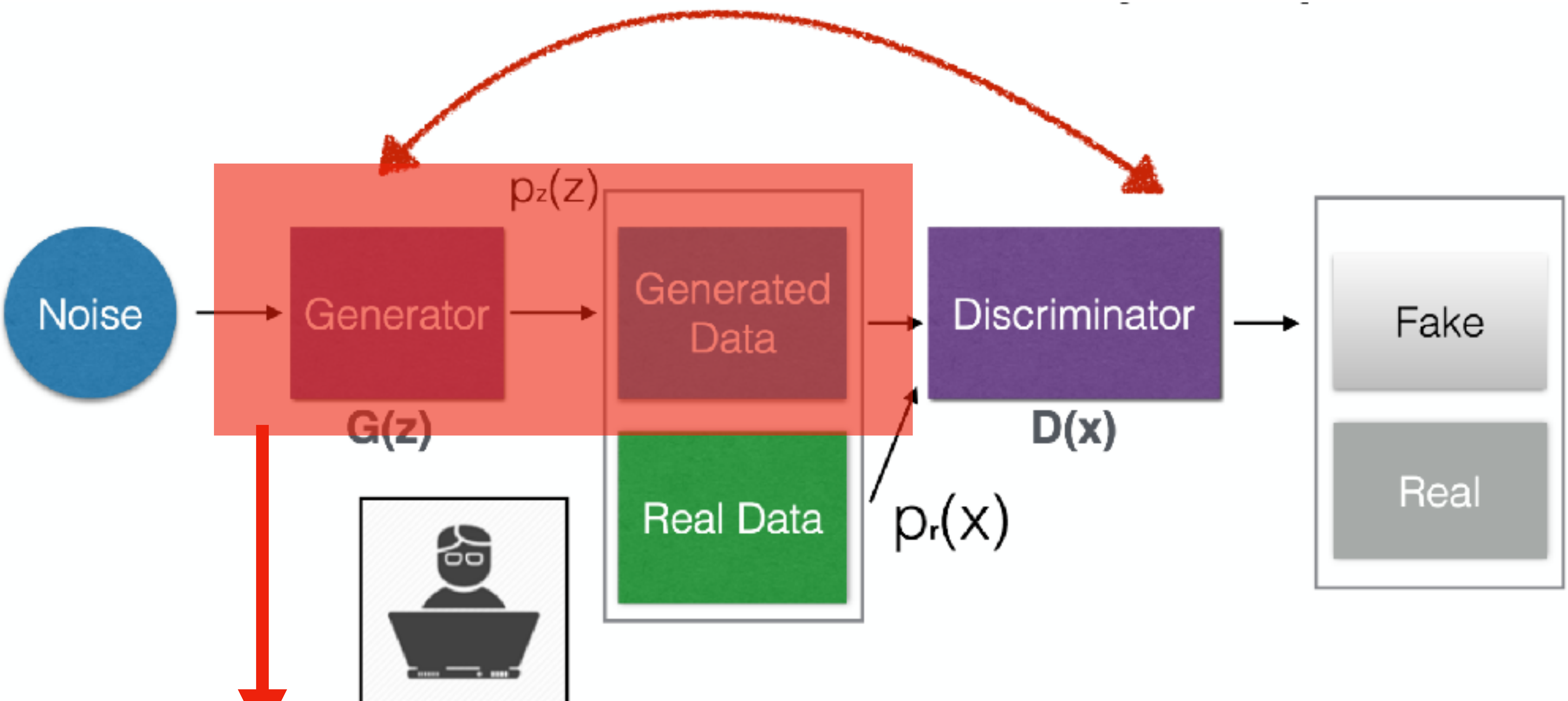
SEVERAL OPTIONS EXIST IN THE LITTERATURE…

NO TIME TO TALK ABOUT BNNs

GENERATIVE ADVERSARIAL NETWORKS

(Goodfellow+)

TWO COMPETING NETWORKS

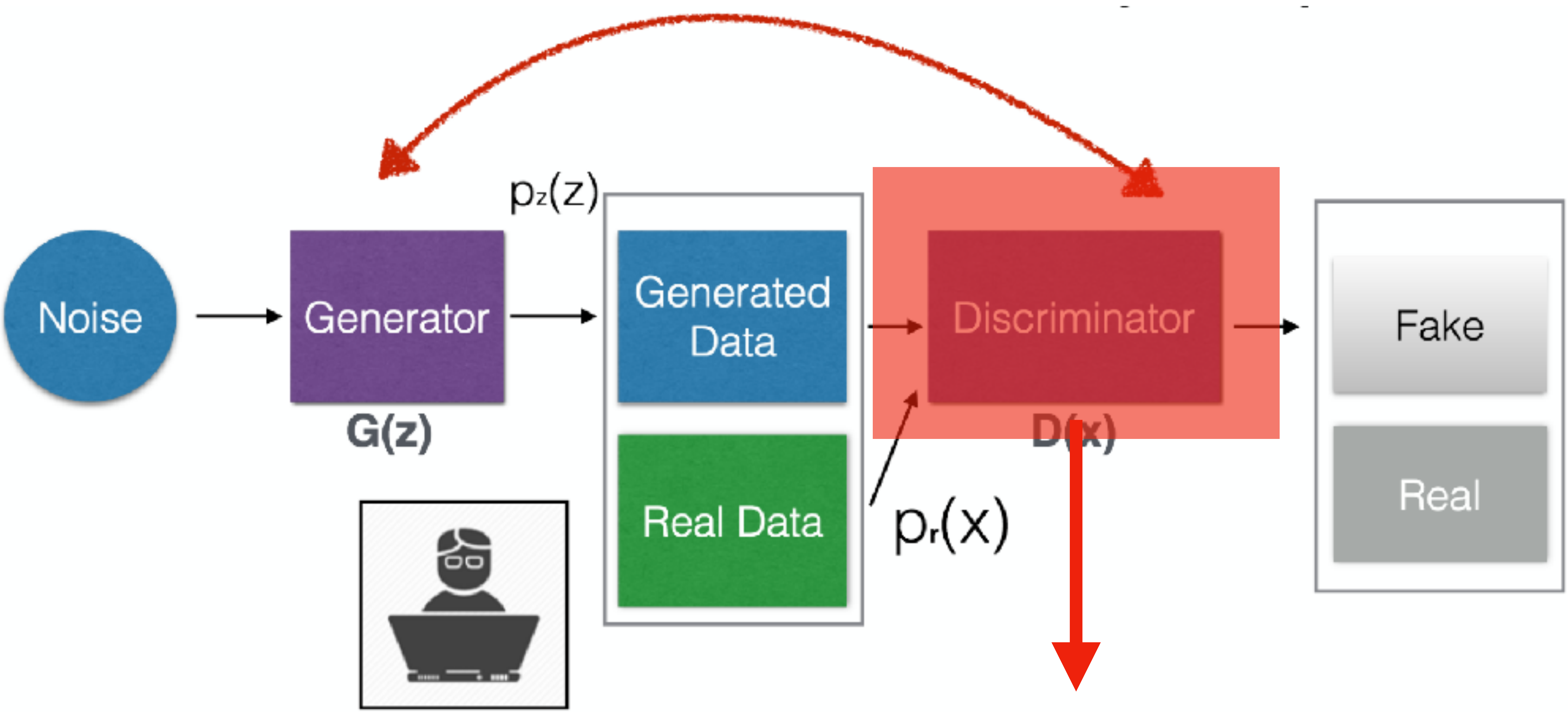


Every 2 iterations the generator is trained to force the discriminator to classify as real

GENERATIVE ADVERSARIAL NETWORKS

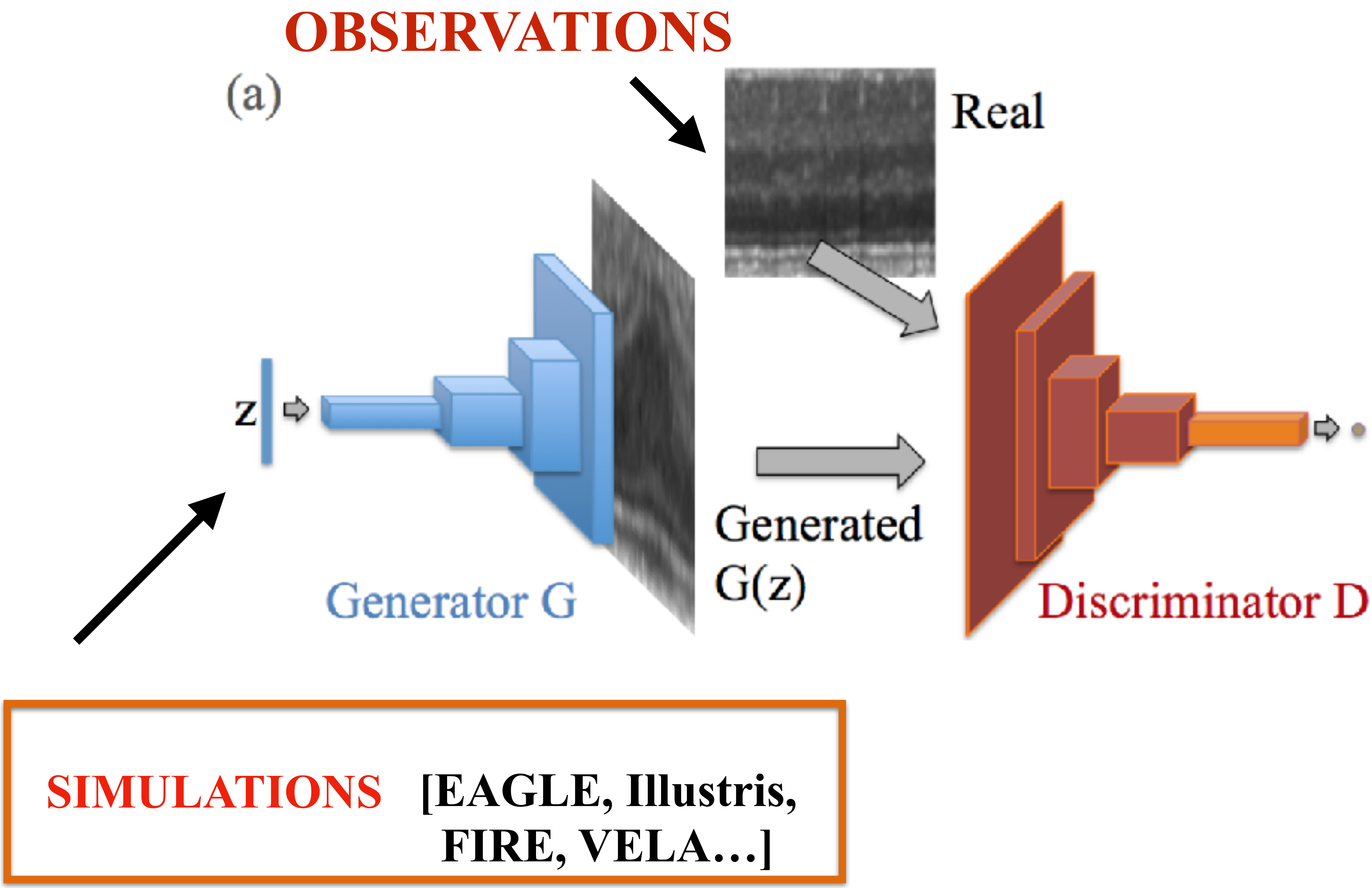
(Goodfellow+)

TWO COMPETING NETWORKS

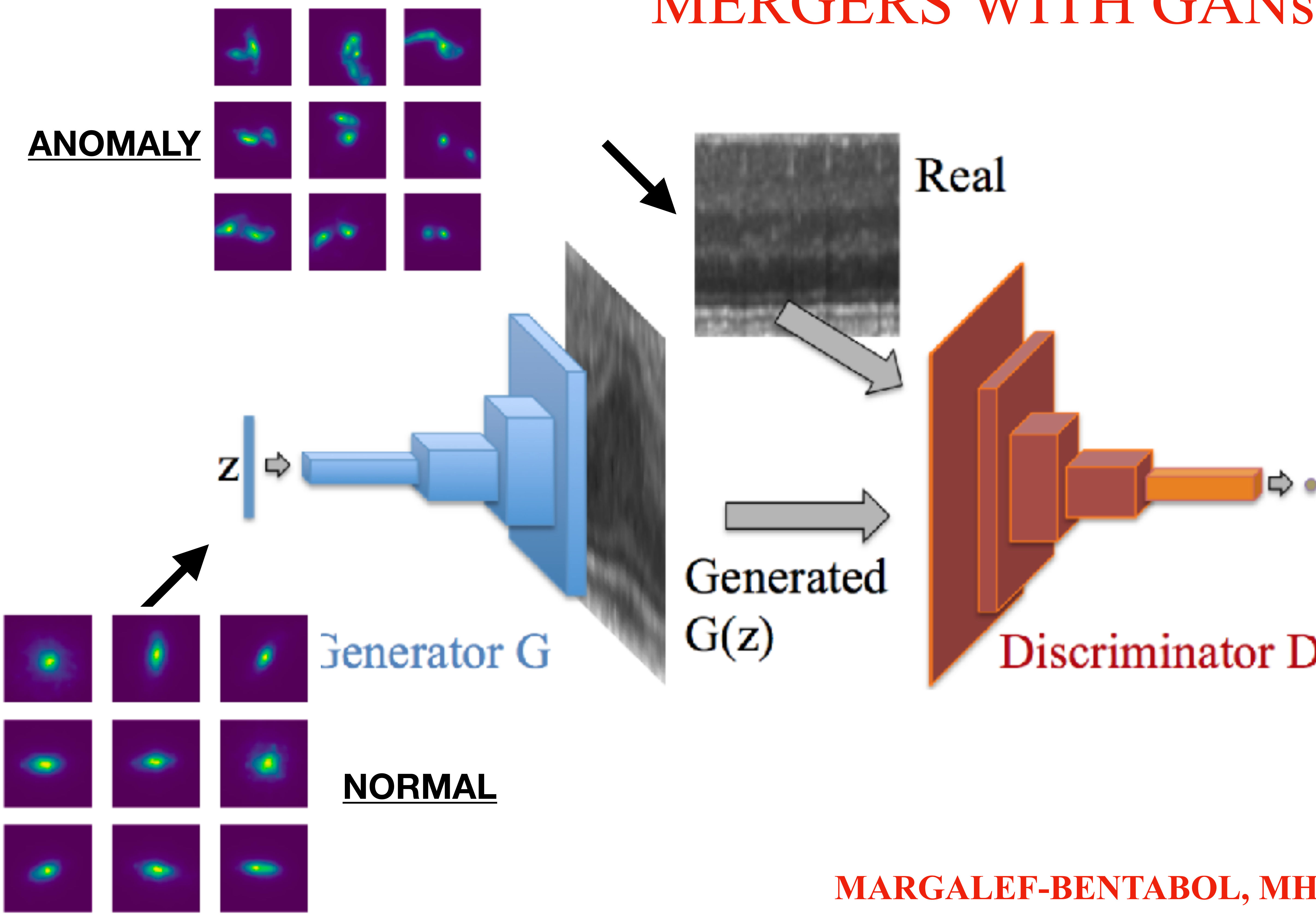


Every 2 iterations the discriminator is trained to force to distinguish between real and fake

ANOMALY DETECTION WITH GANs

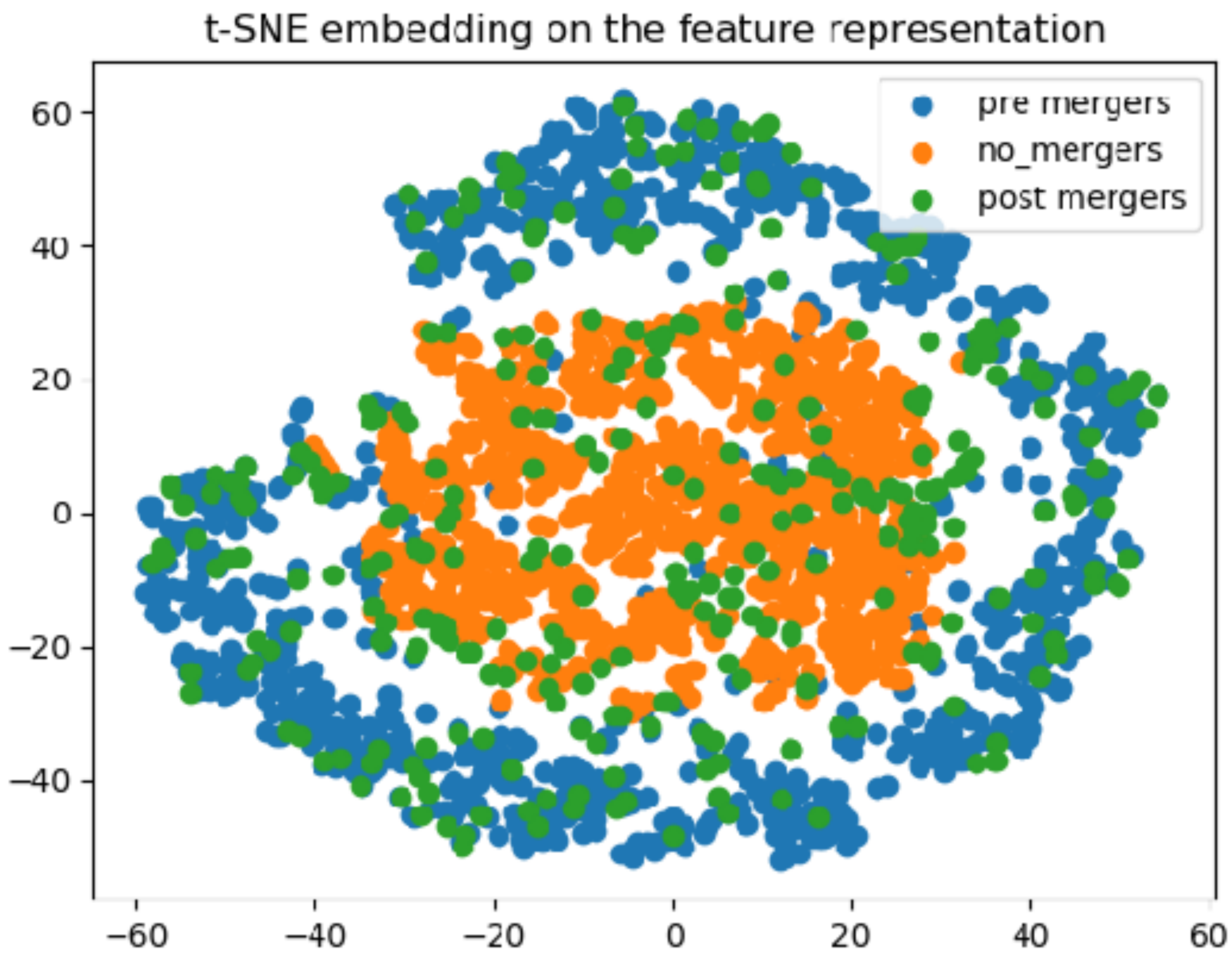


A PROOF-OF-CONCEPT CASE: UNSUPERVISED DETECTION OF MERGERS WITH GANs

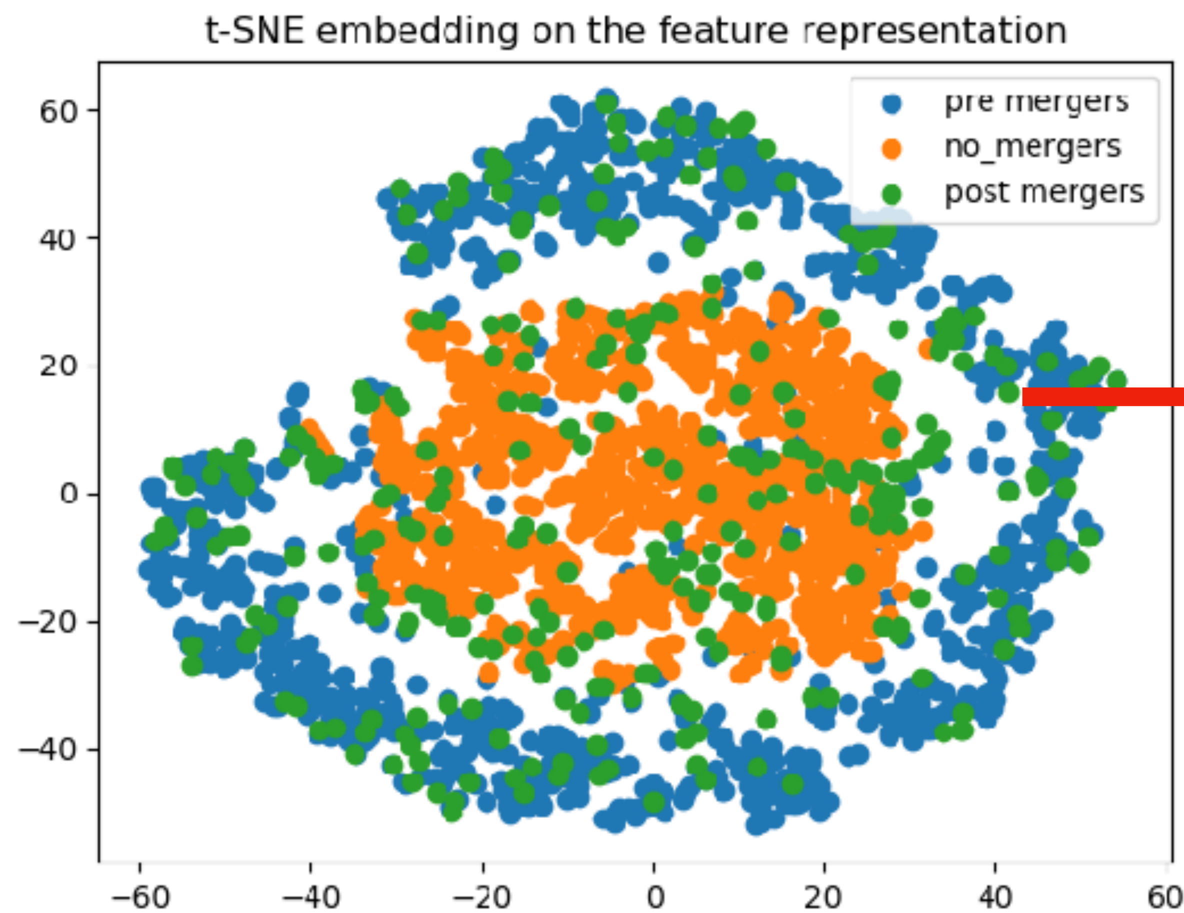


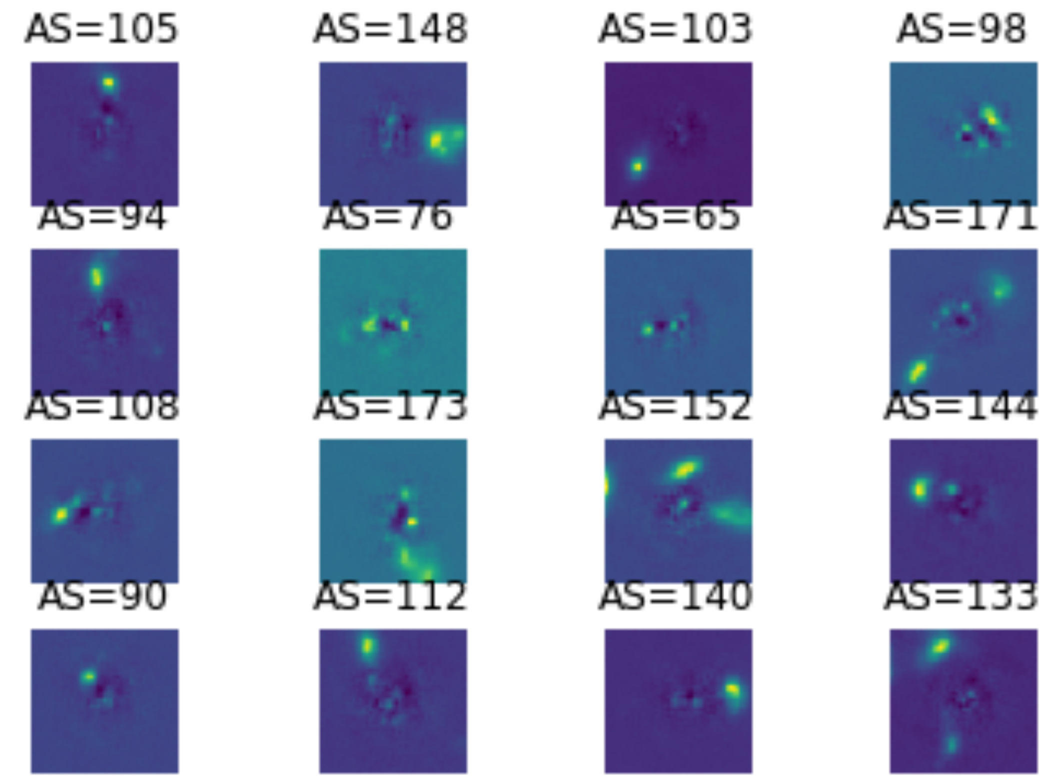
MARGALEF-BENTABOL, MHC+2019

MERGERS OCCUPY A DIFFERENT REGION IN THE GAN GENERATED LATENT SPACE



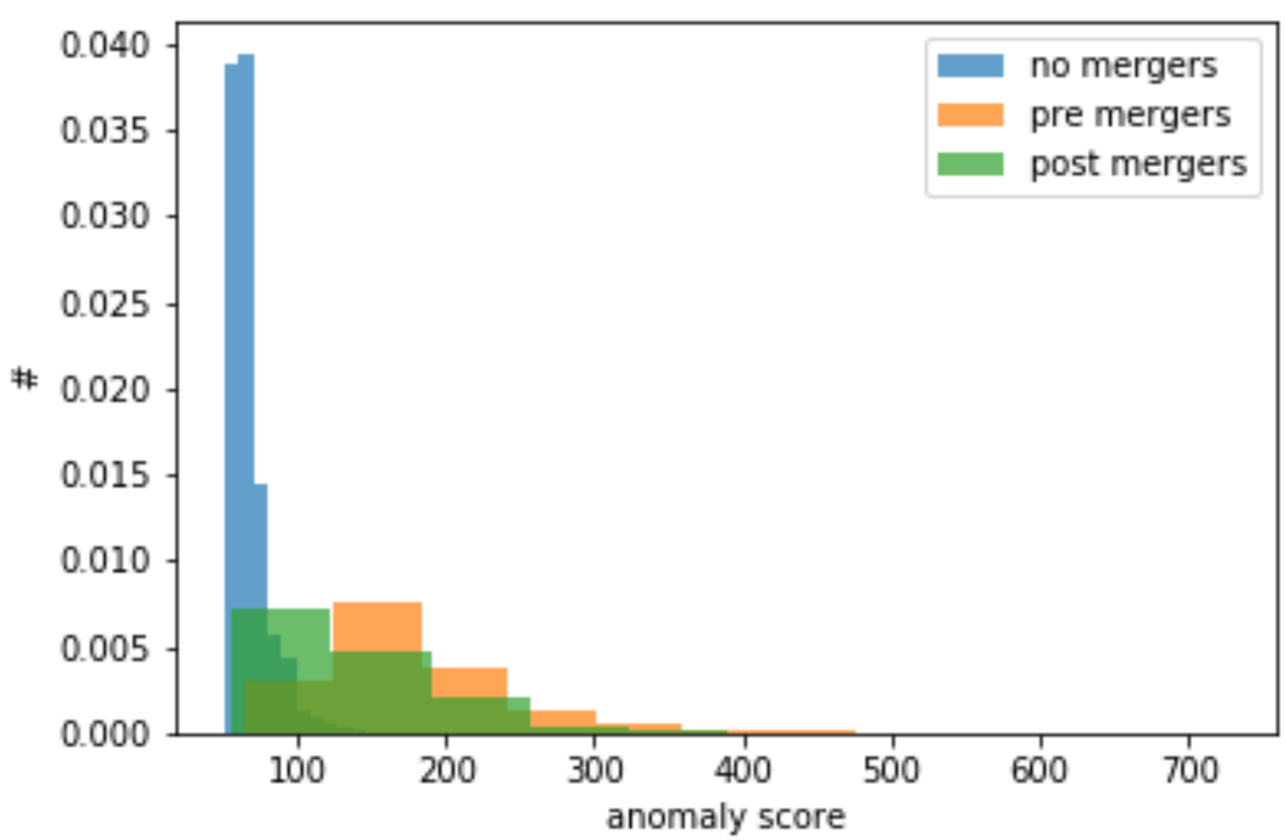
MARGALEF-BENTABOL, MHC+2019





(RECONSTRUCTED - ORIGINAL)

<u>“ANOMALY SCORE”</u>



MARGALEF-BENTABOL, MHC+2019

2 TAKE HOME PROVOCATIVE ? MESSAGES FOR TODAY

1. MOST OF THE PROCESSING WE DO ON IMAGES CAN BE DONE WITH AI - POSSIBLY MORE EFFICIENTLY AND MORE ACCURATELY

DETECTION, PHOTOMETRY, PHOTOZ's, MORPHOMETRY …

2. WE CAN LEARN SOME PHYSICS BY USING AI TO LINK SIMULATIONS AND OBSERVATIONS

WITH SUPERVISED ML GOING BACK AND FORTH FROM SIMS TO DATA

WITH UN-SUPERVISED ML TO MEASURE SIMILARITIES