

Machine Learning Frameworks for next-generation Cosmology Surveys

Nikhel Gupta

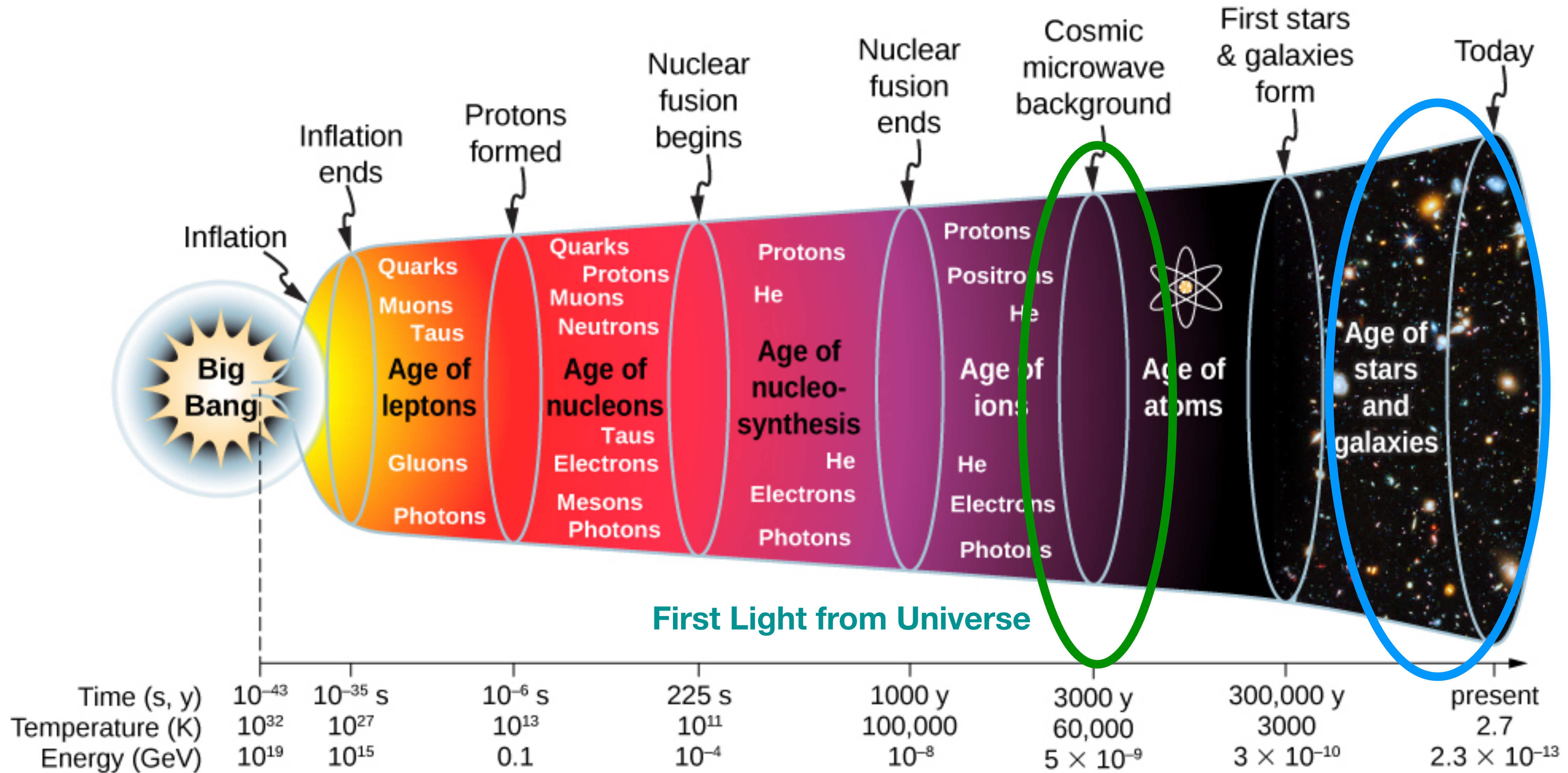
ML/AI Future Science Platform (MLAI FSP)
CSIRO Space & Astronomy, Australia

Collaborators:

Minh Huynh, Ray P. Norris, Zeeshan Hayder, Lars Petersson,
Rosalind Wang, Vivien Rolland



Evolution of Universe

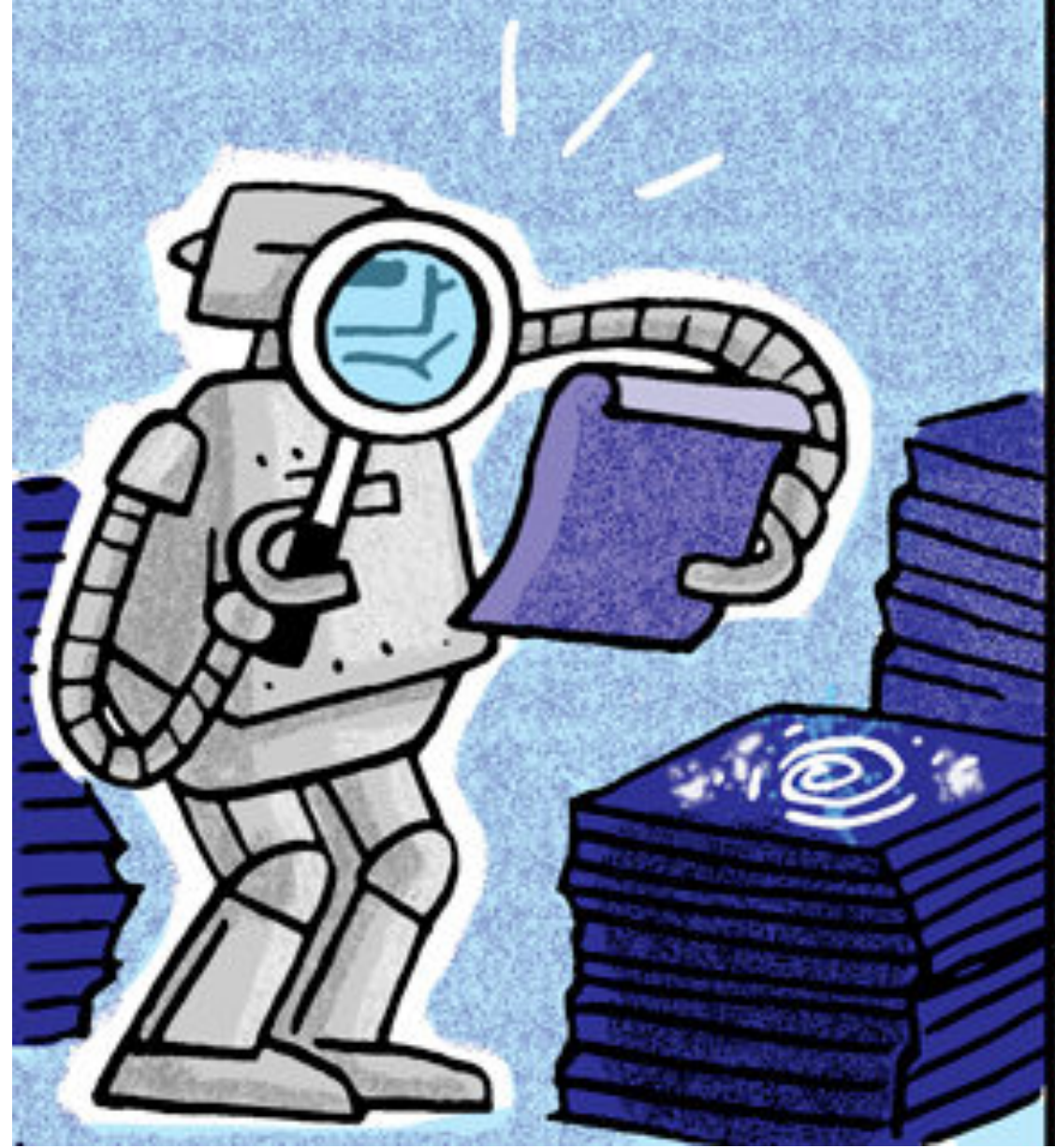


Credit: Physics Libretexts



MACHINE LEARNING HELPS OUT

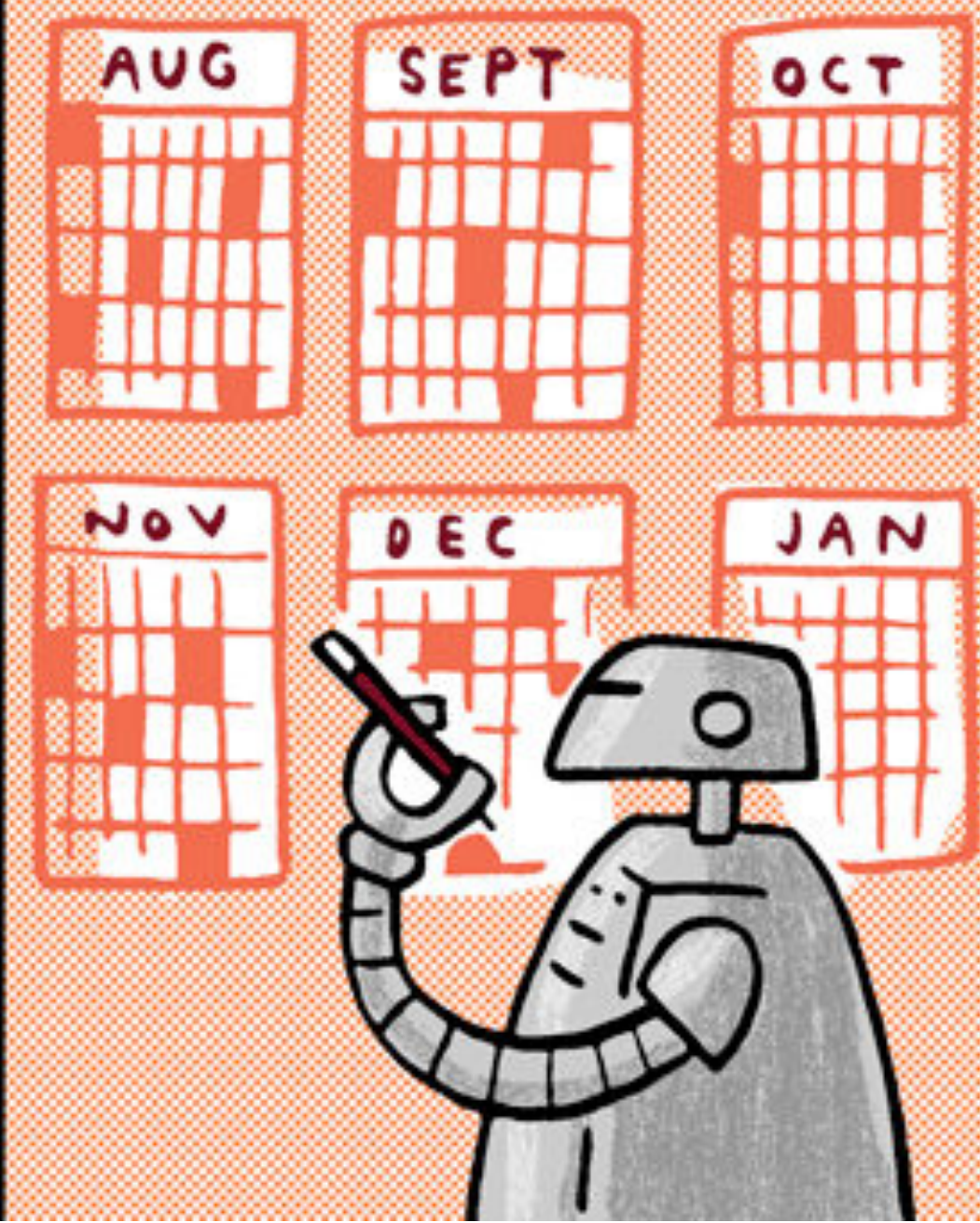
OBJECT
DETECTION/
CLASSIFICATION



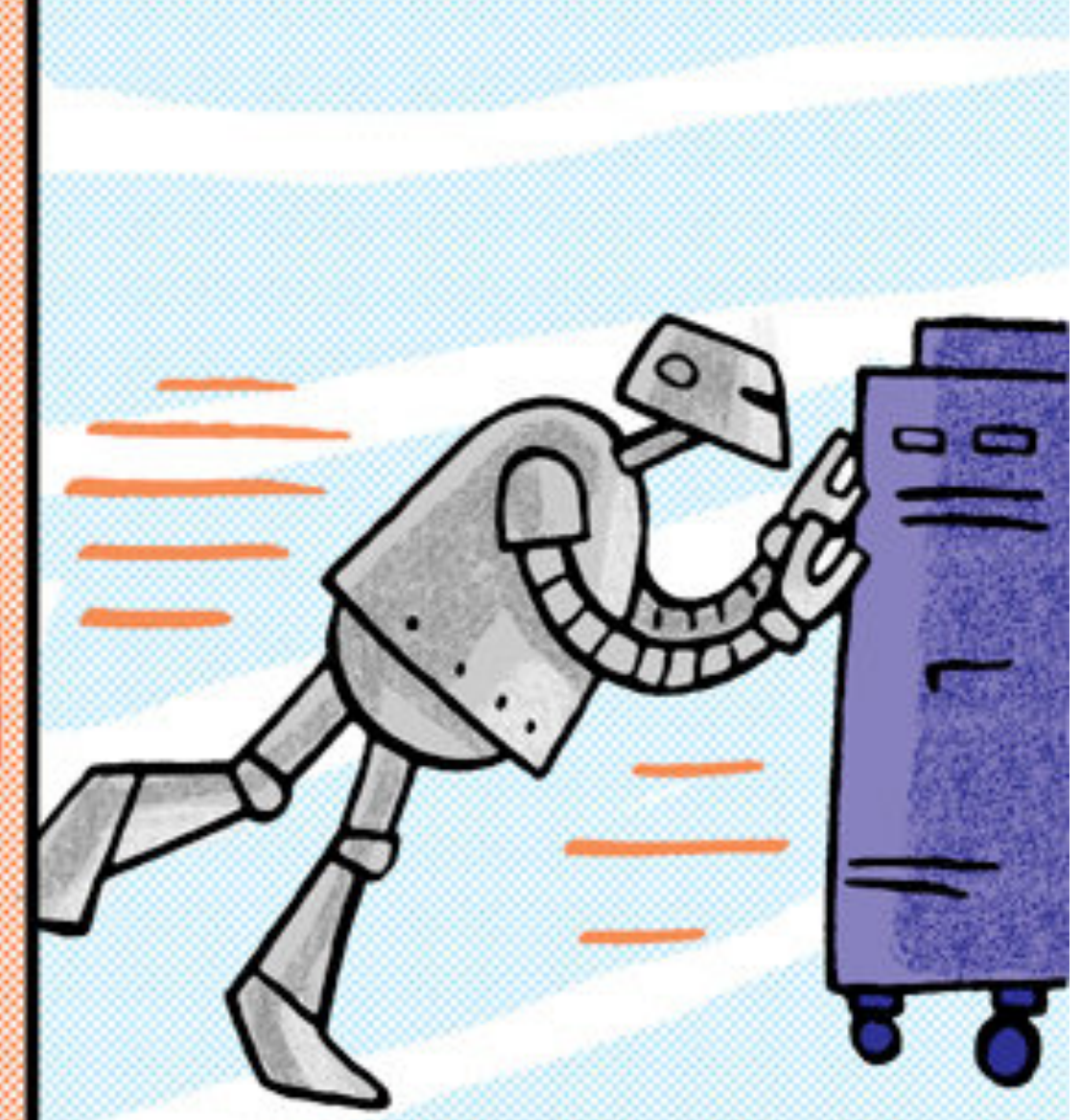
CLEANING
IMAGES

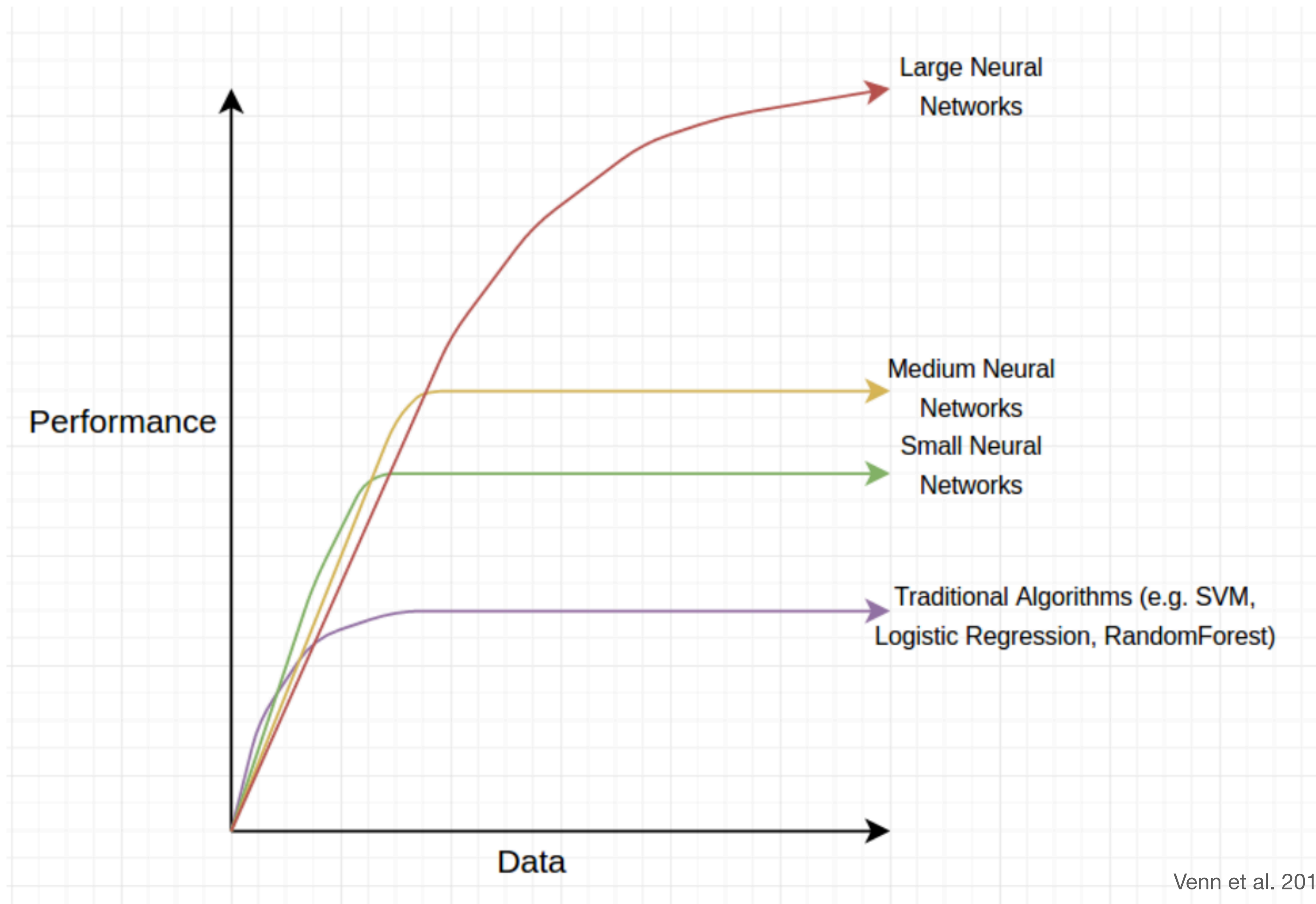


MAKING
SCHEDULES



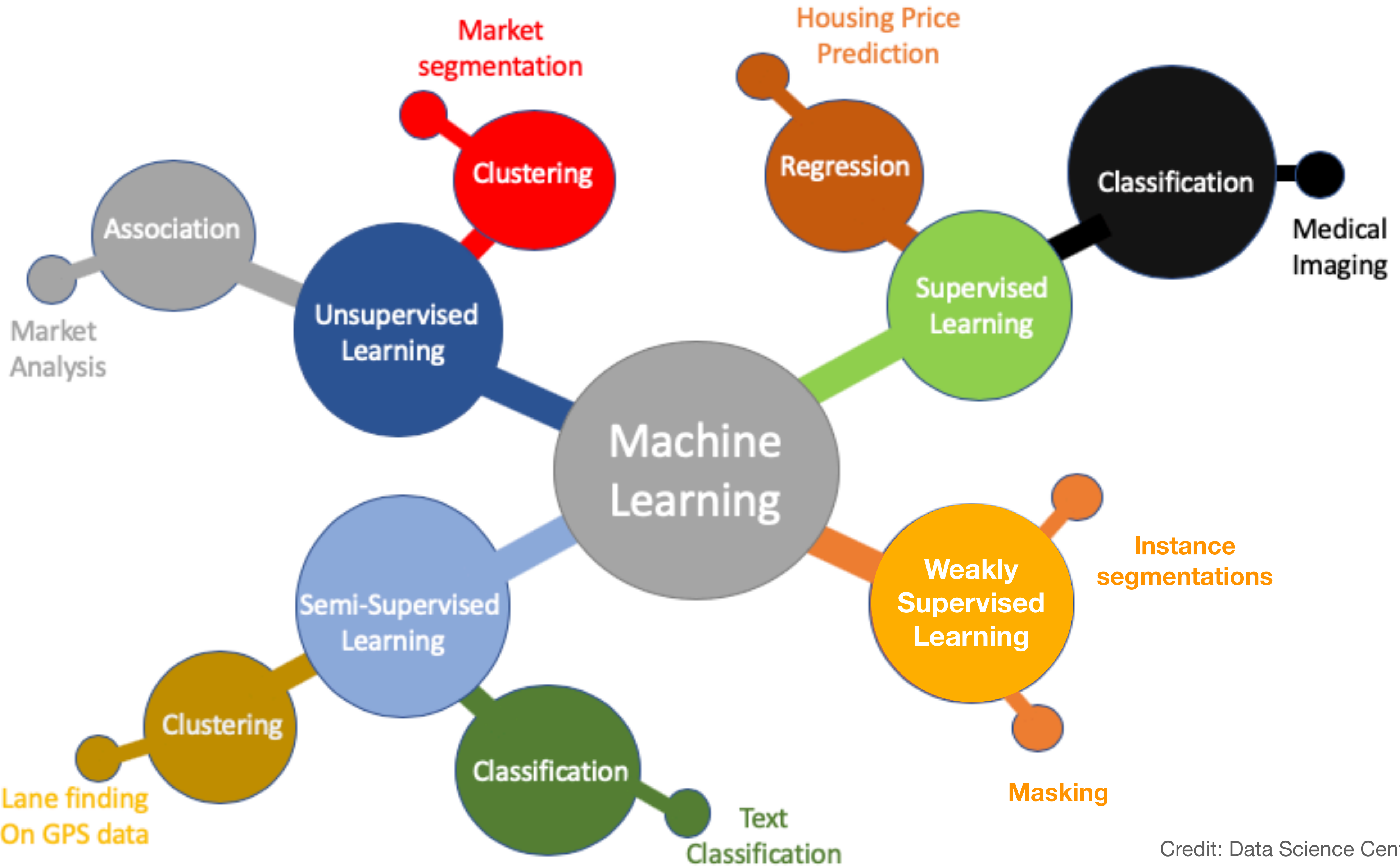
MAKING
SIMULATIONS
GO FASTER





Venn et al. 2019





Credit: Data Science Central



Machine Learning Frameworks

- **Supervised Learning**

Complete information of truth-labels is available for all images

- **Self-supervised Learning**

No information about truth-labels is available for all images

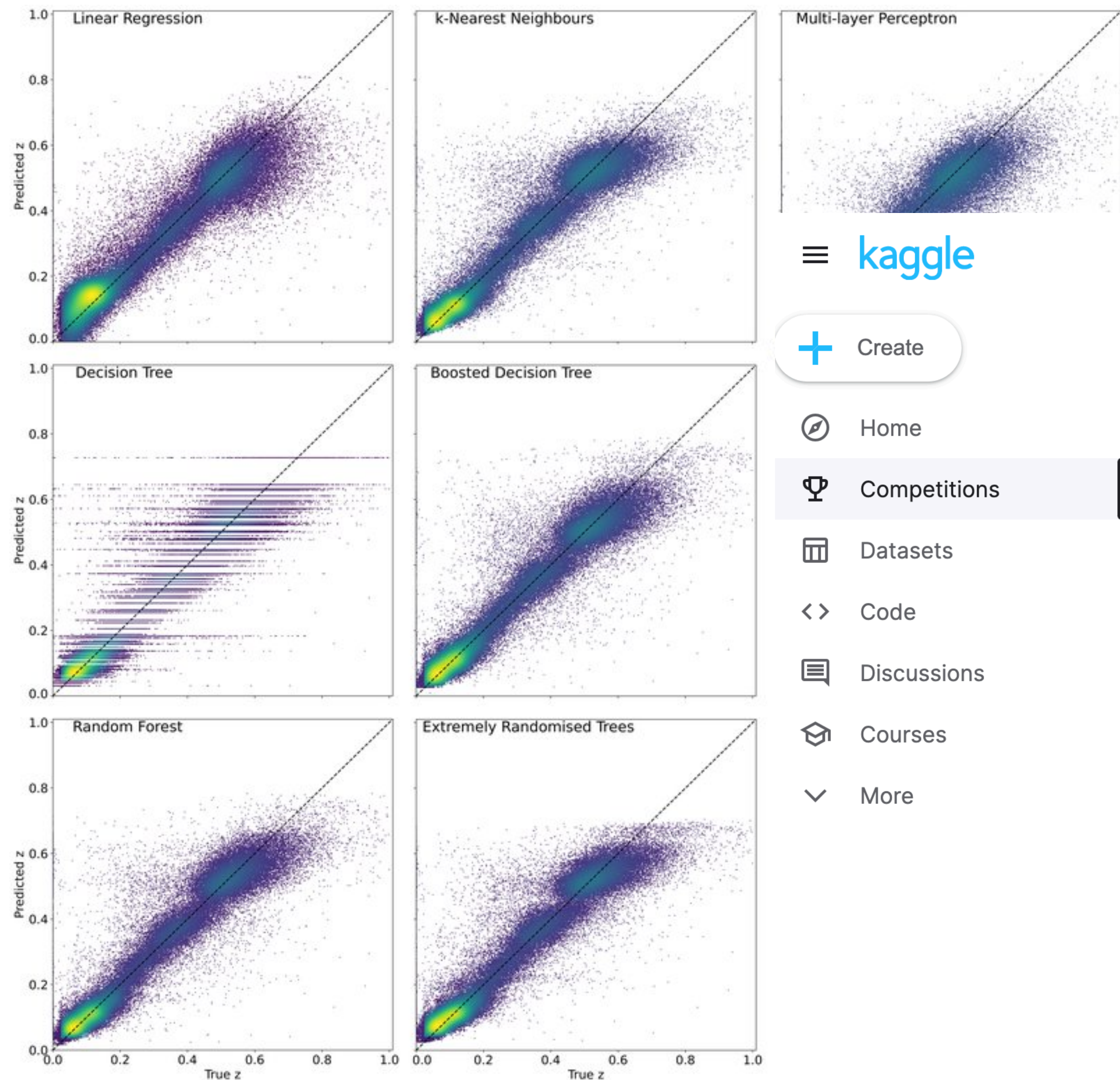
- **Weakly-supervised Learning**

Some information about truth-labels is available for all images

- **Semi-supervised Learning**

Complete information of truth-labels is available for some images

Redshift Estimations



Henghes et al. 2021

Redshift estimations Catalog space i.e. using tables

Community Prediction Competition

Machine Learning for photo-z estimation

Investigate the use of recent advances in machine learning to design a new photometric redshift estimation procedure.

11 teams · 2 years ago

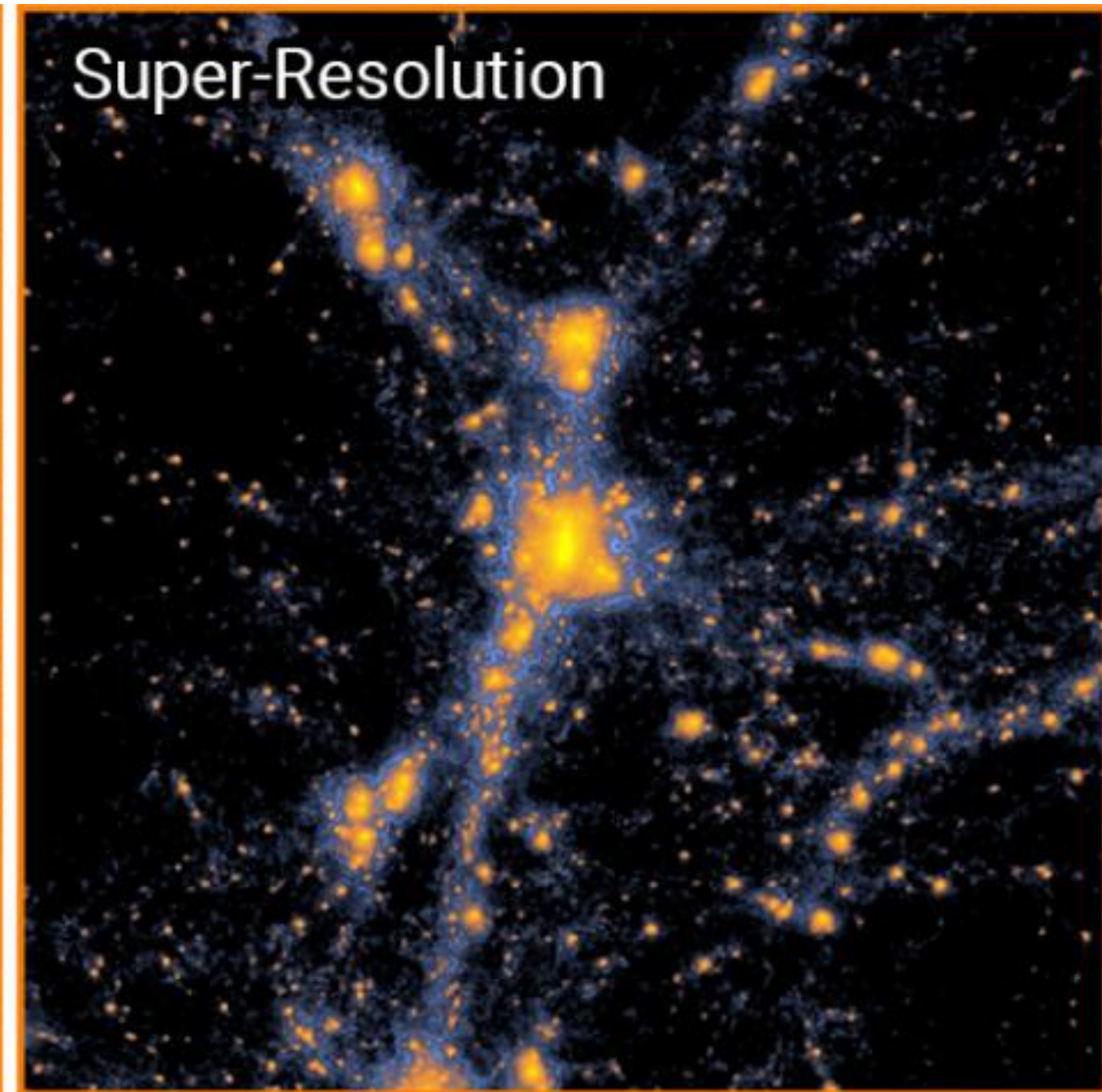
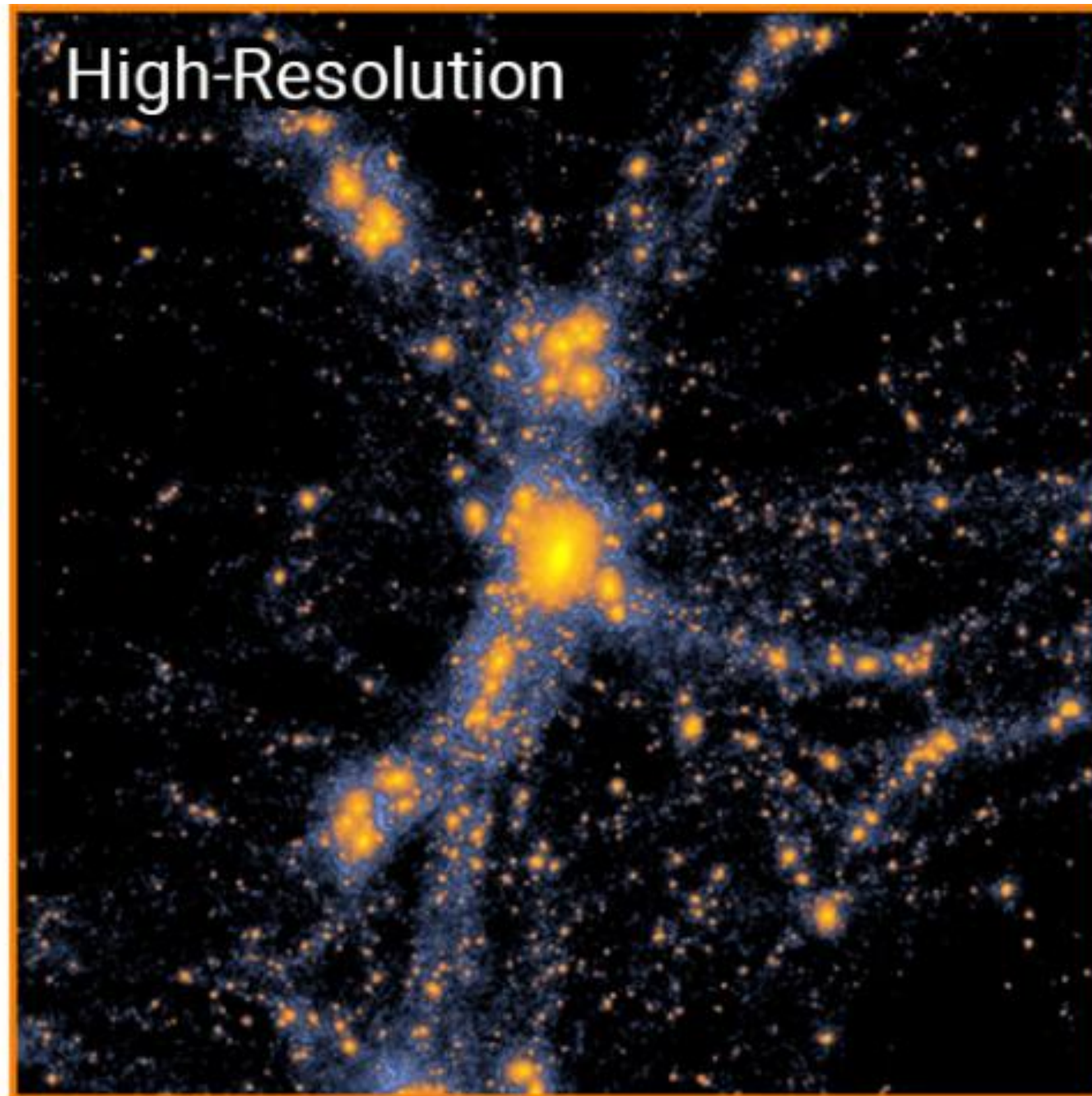
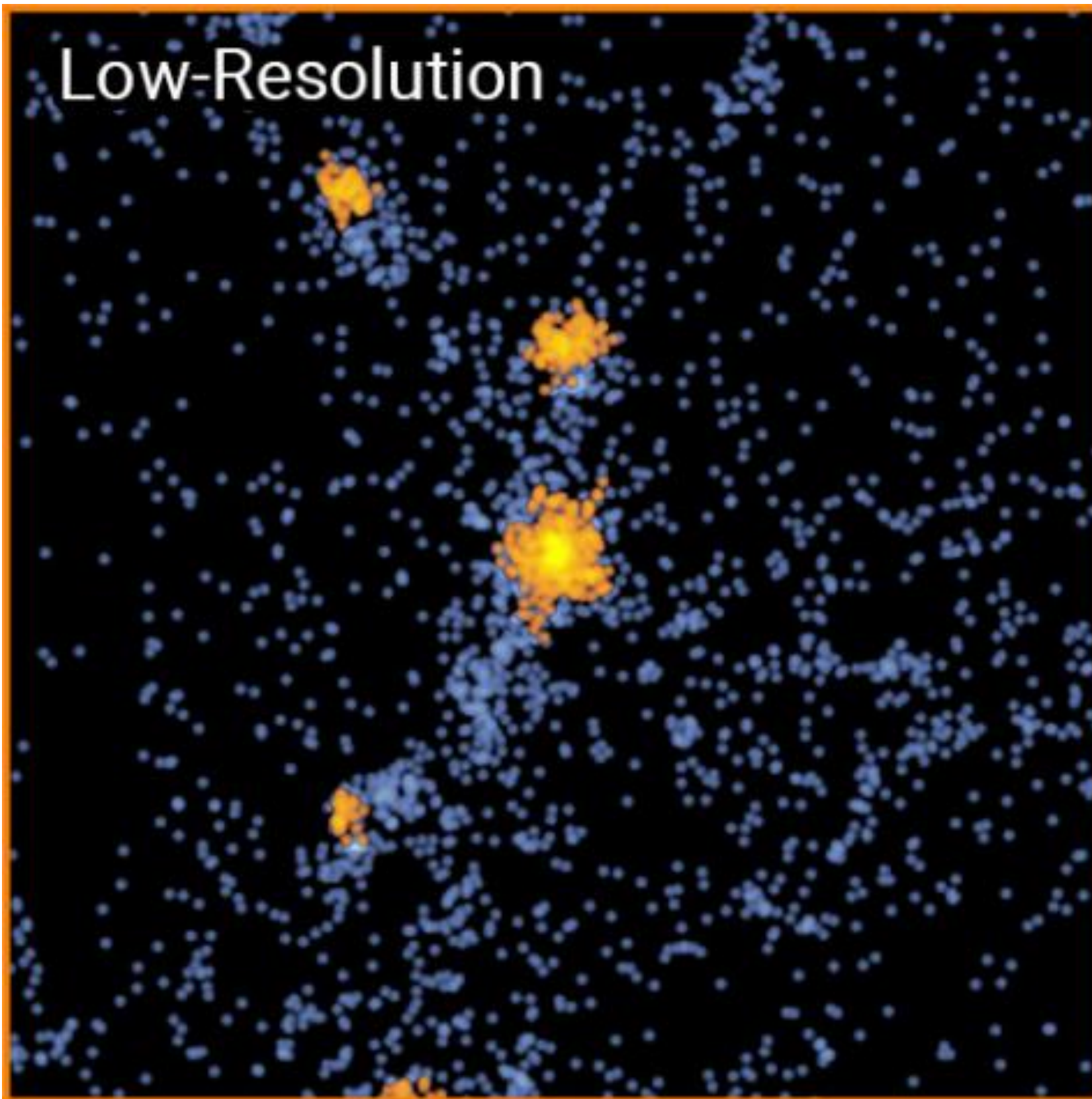
[Overview](#) [Data](#) [Code](#) [Discussion](#) [Leaderboard](#) [Rules](#)

Problem description

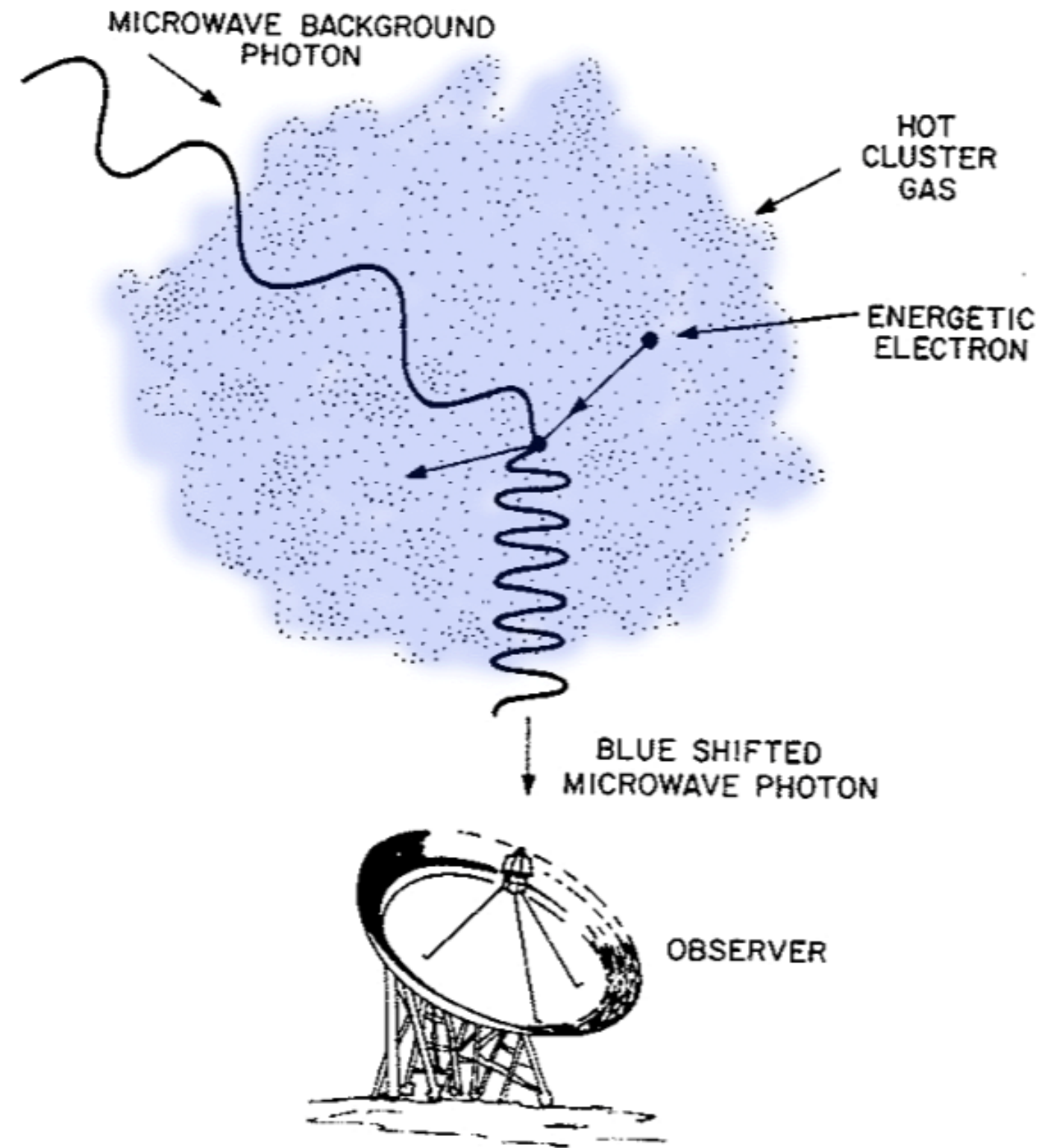
Galaxy redshift surveys are among the main observational tools to probe cosmological models. The leading methods measure the distance scale imprinted in the large-scale distribution of galaxies by oscillations in the primordial baryon-photon plasma. This baryonic acoustic oscillation (BAO) sound horizon can be used as a standard ruler to characterize the expansion rate of the Universe at different times, thereby providing constraints on cosmological parameters such as the total matter and dark energy densities. A precise measurement of the redshifts of galaxies is fundamental to extract this

Making High Resolution Simulations

Y. Li et al. 2021

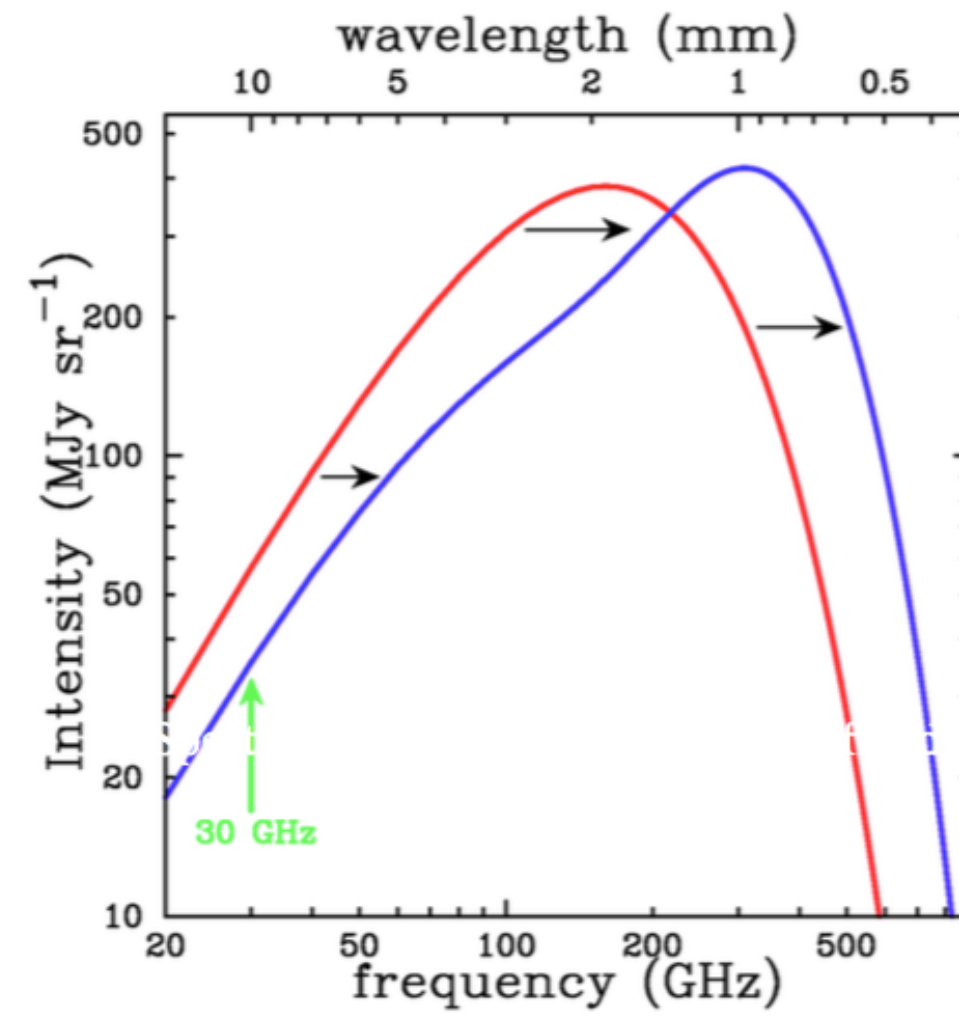


Sunyaev Zel'dovich (SZ) Effect

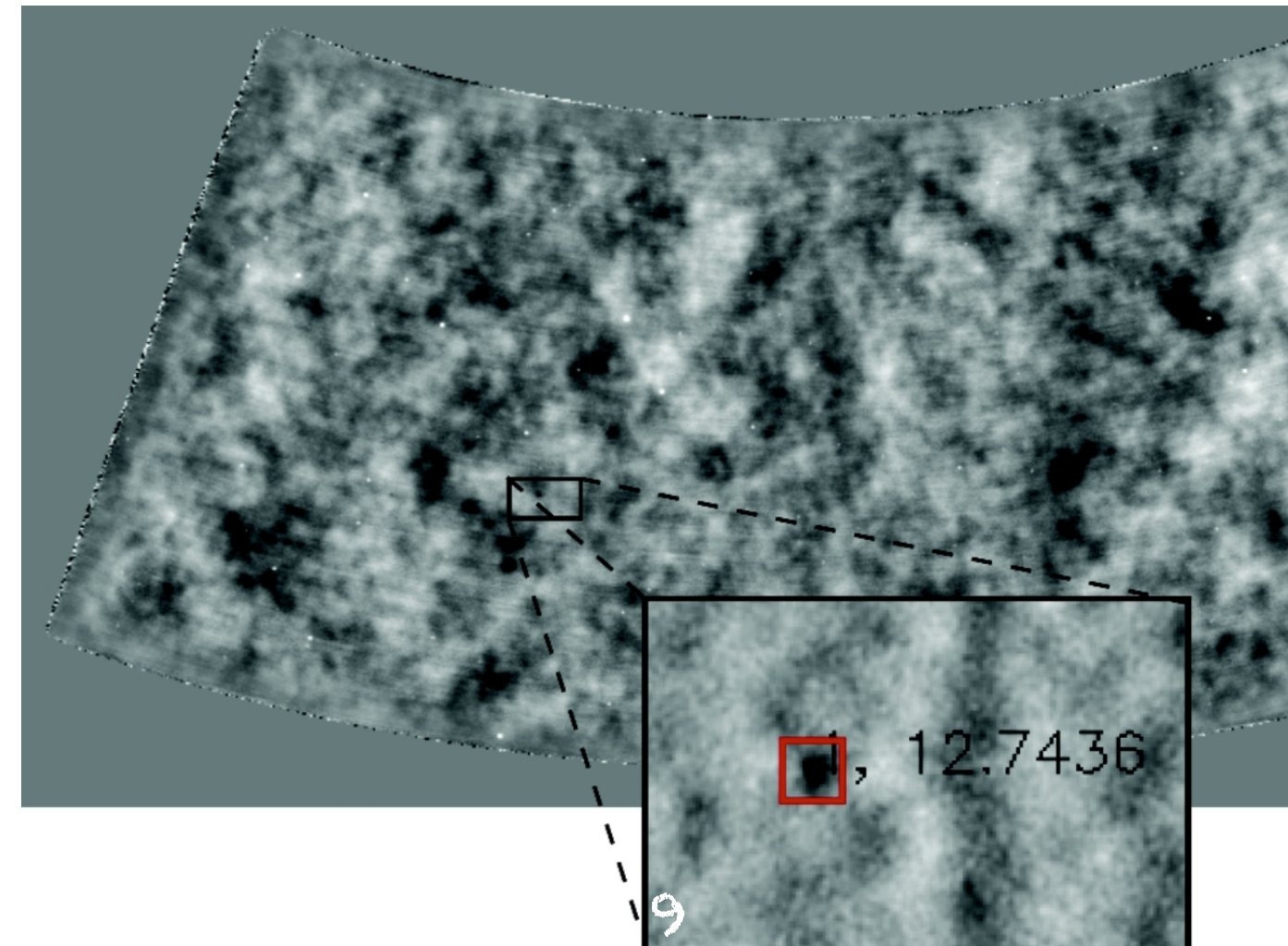
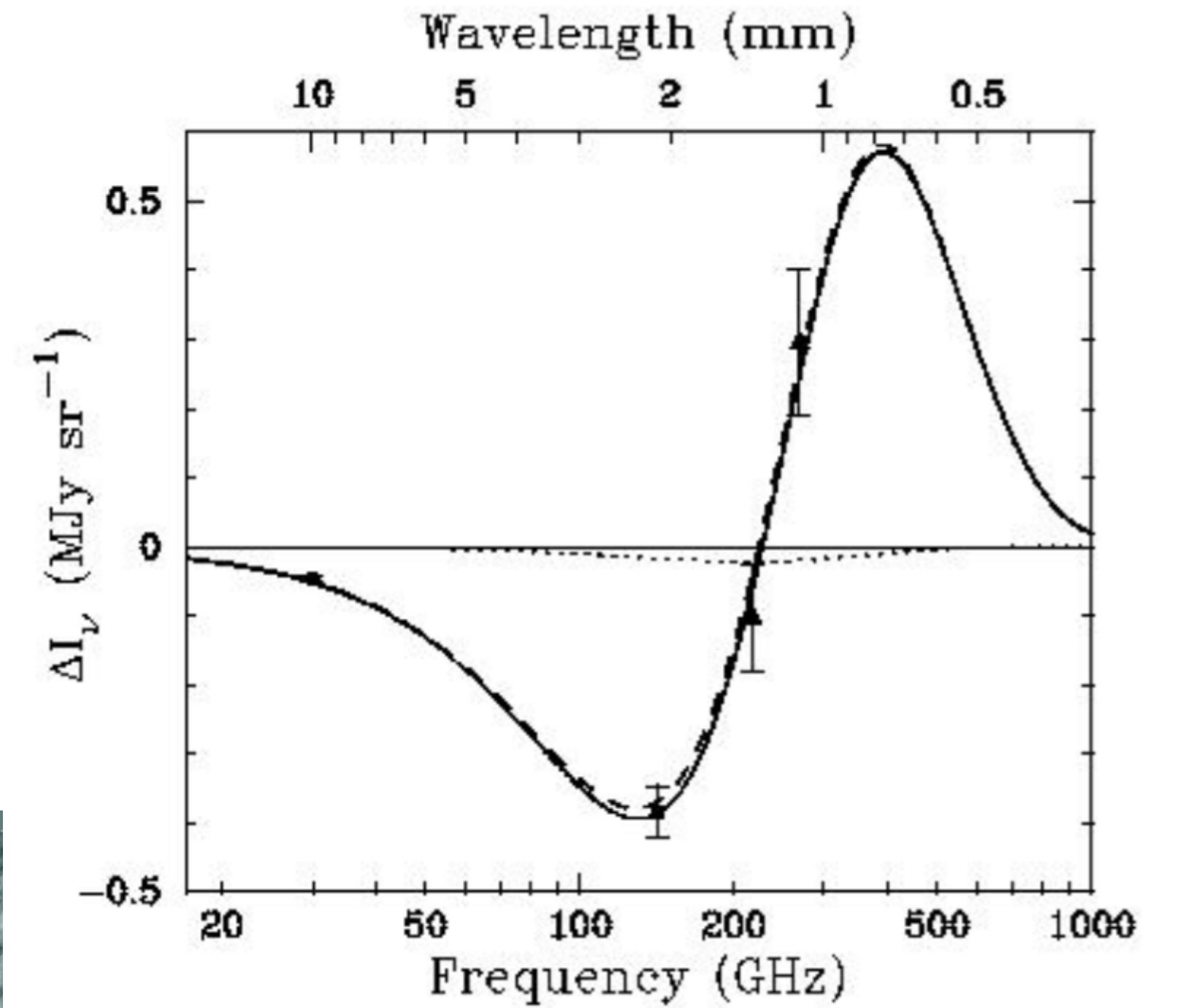


Adapted from L. Van Speybroeck

Spectral Distortion of CMB

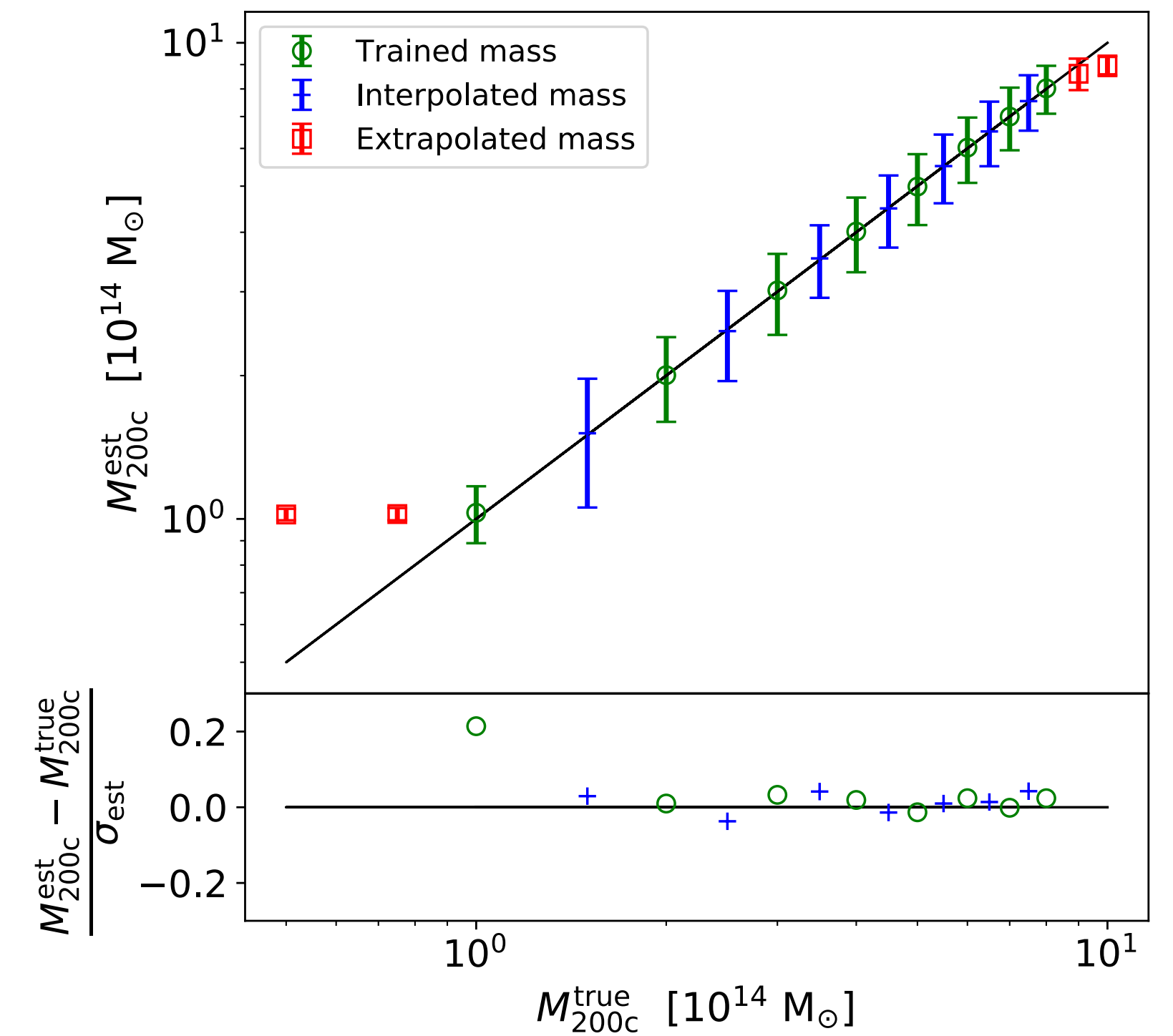
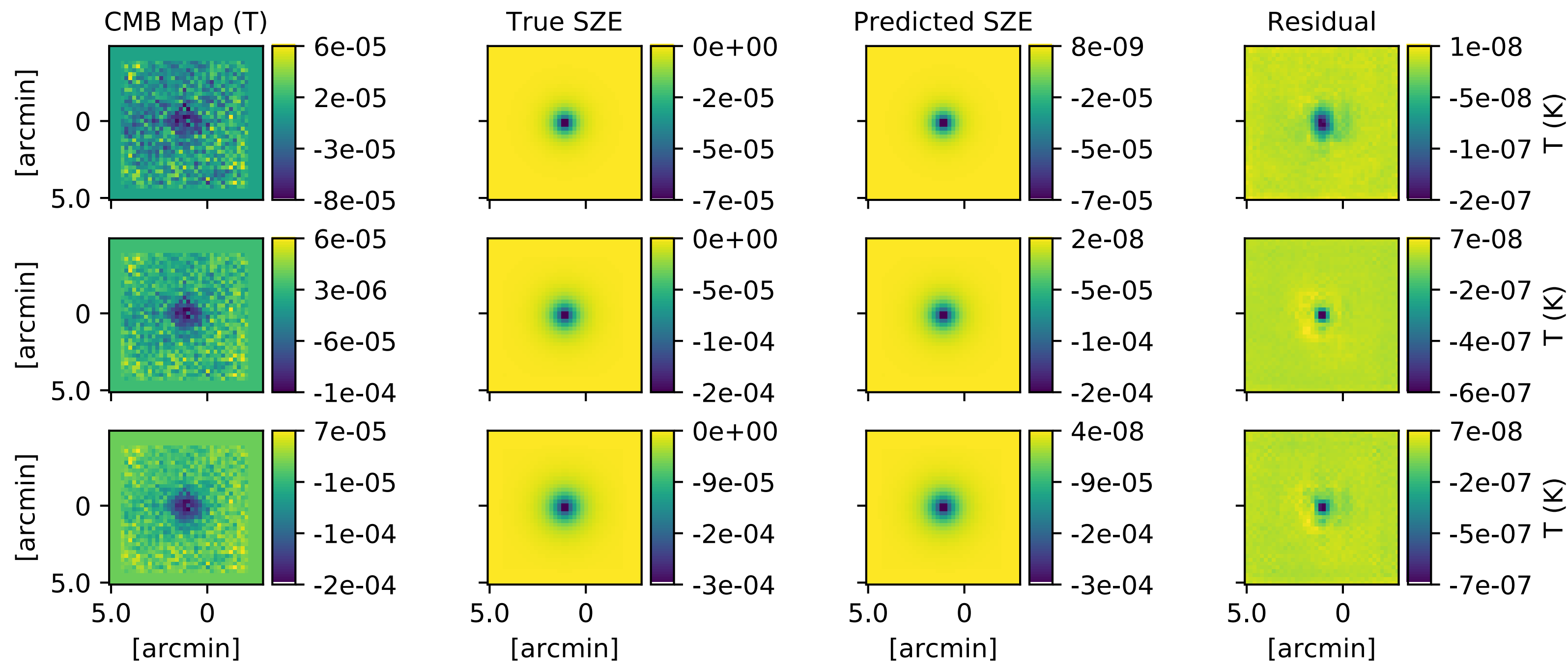


Sunyaev & Zel'dovich 1970, 1972



SZ Profile & Mass Predictions with Supervised Machine Learning

Gupta et al. (2020b)

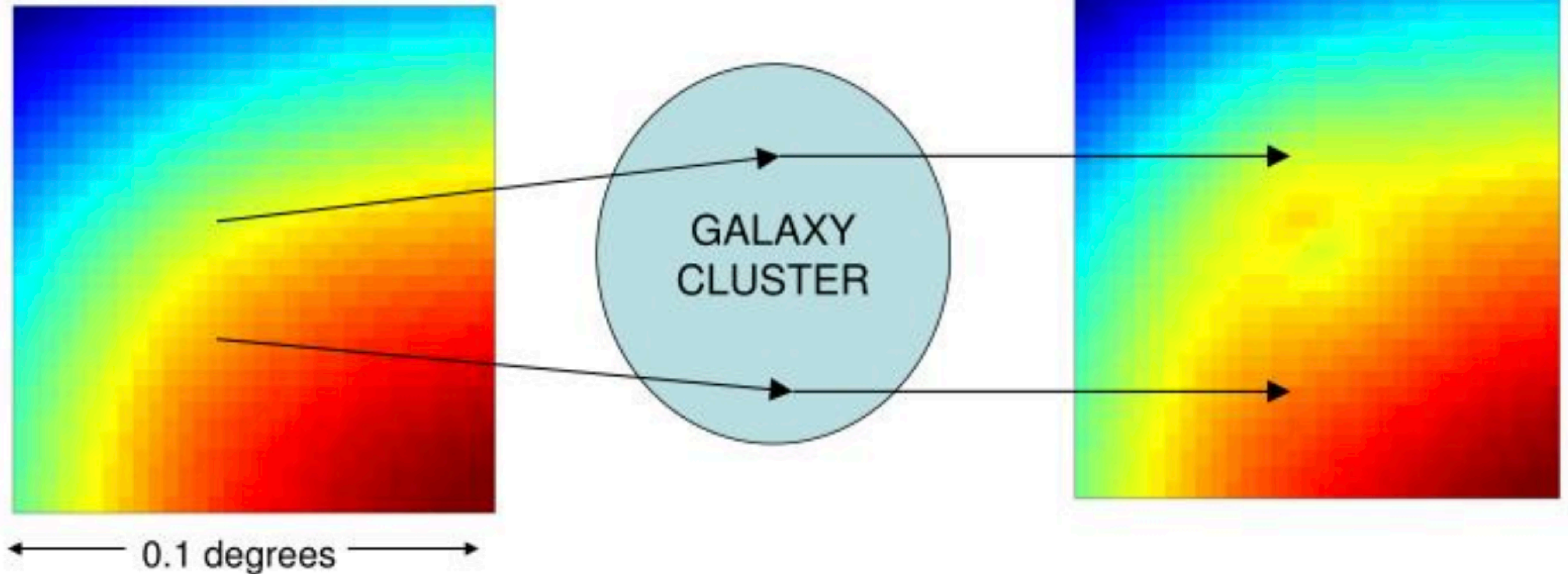


Cosmic Microwave Background (CMB) – Galaxy Cluster Lensing

CMB very smooth on small scales: approximately a gradient

Last scattering surface

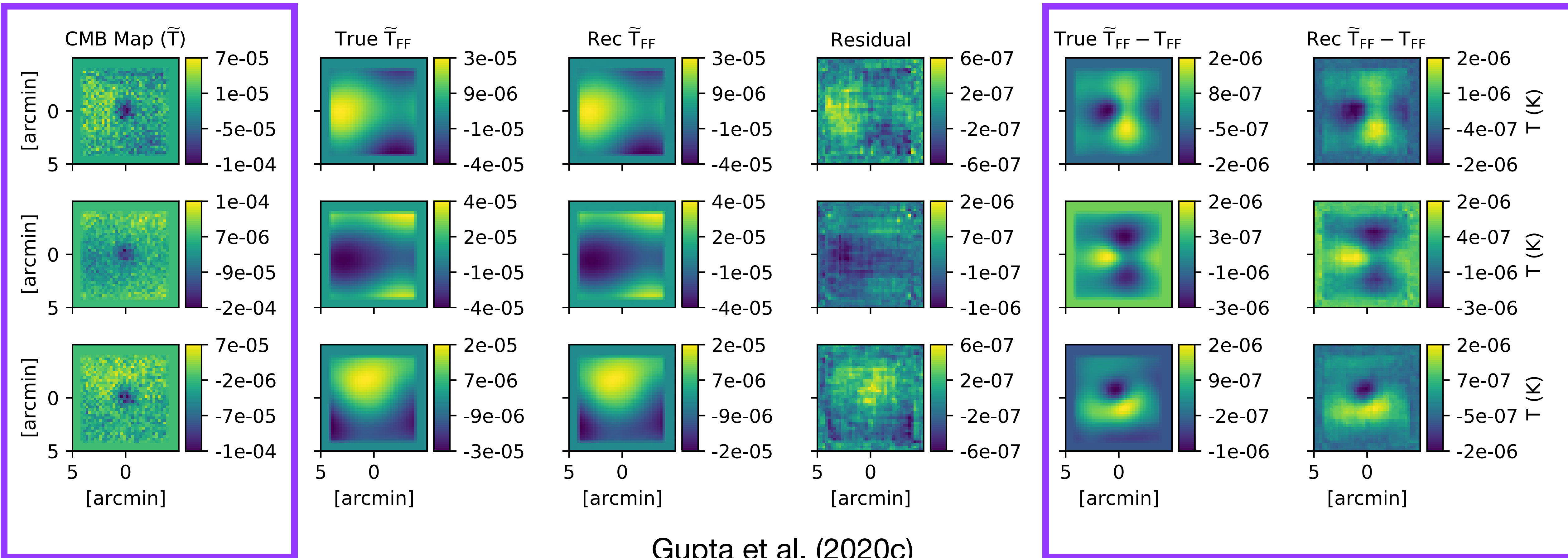
What we see



Credit: Lewis et al.

Cosmic Microwave Background (CMB) – Galaxy Cluster

Convergence Profiles Predictions with Machine Learning leads to cluster masses!!



Gupta et al. (2020c)

Machine Learning Frameworks

- **Supervised Learning**

Complete information of truth-labels is available for all images

- **Self-supervised Learning**

No information about truth-labels is available for all images

- **Weakly-supervised Learning**

Some information about truth-labels is available for all images

- **Semi-supervised Learning**

Complete information of truth-labels is available for some images



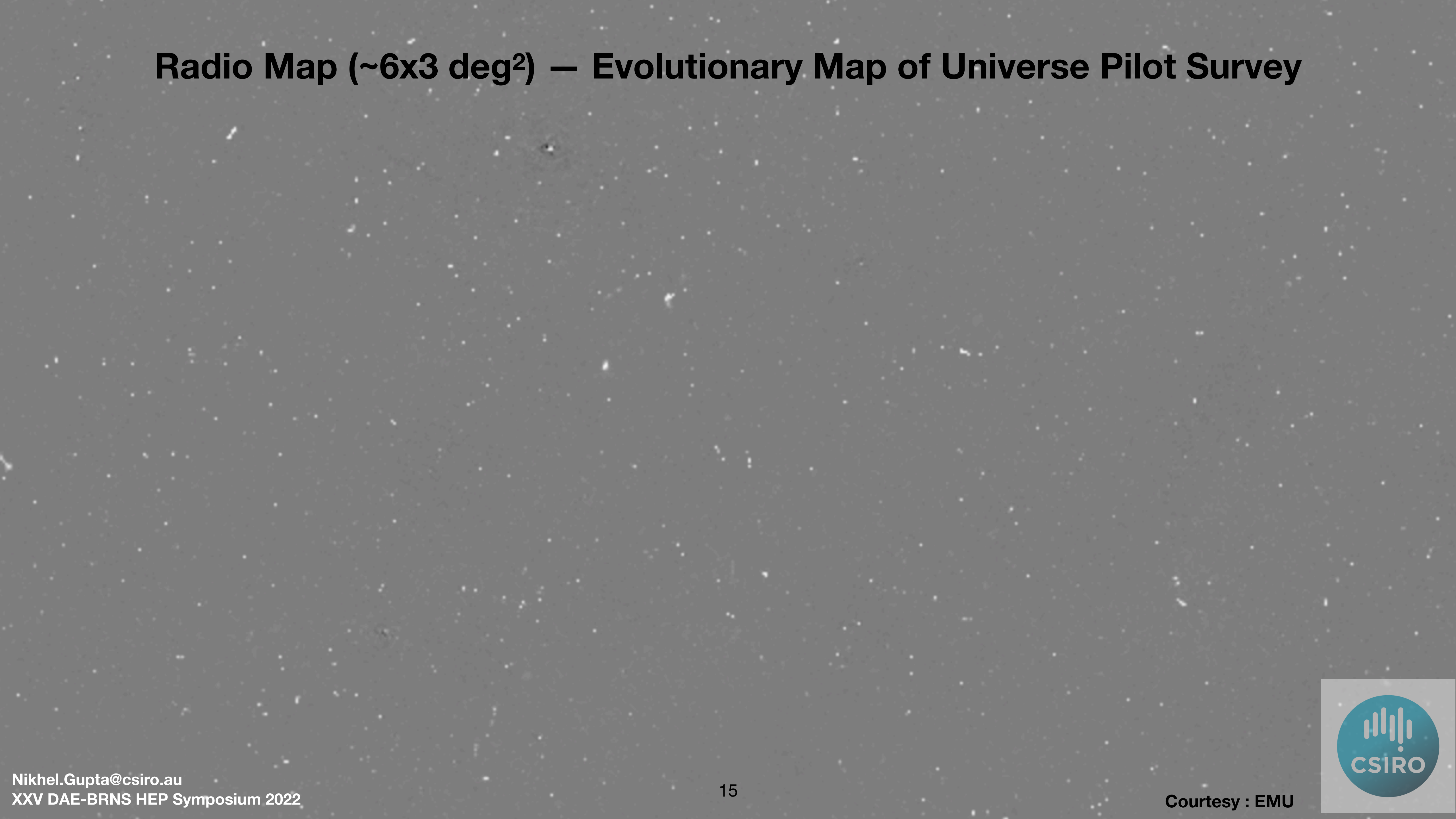
Australian Square Kilometre Array Pathfinder (ASKAP)

Evolutionary Map of Universe
(EMU) will detect **>40 million** radio
galaxies over next 5 years!

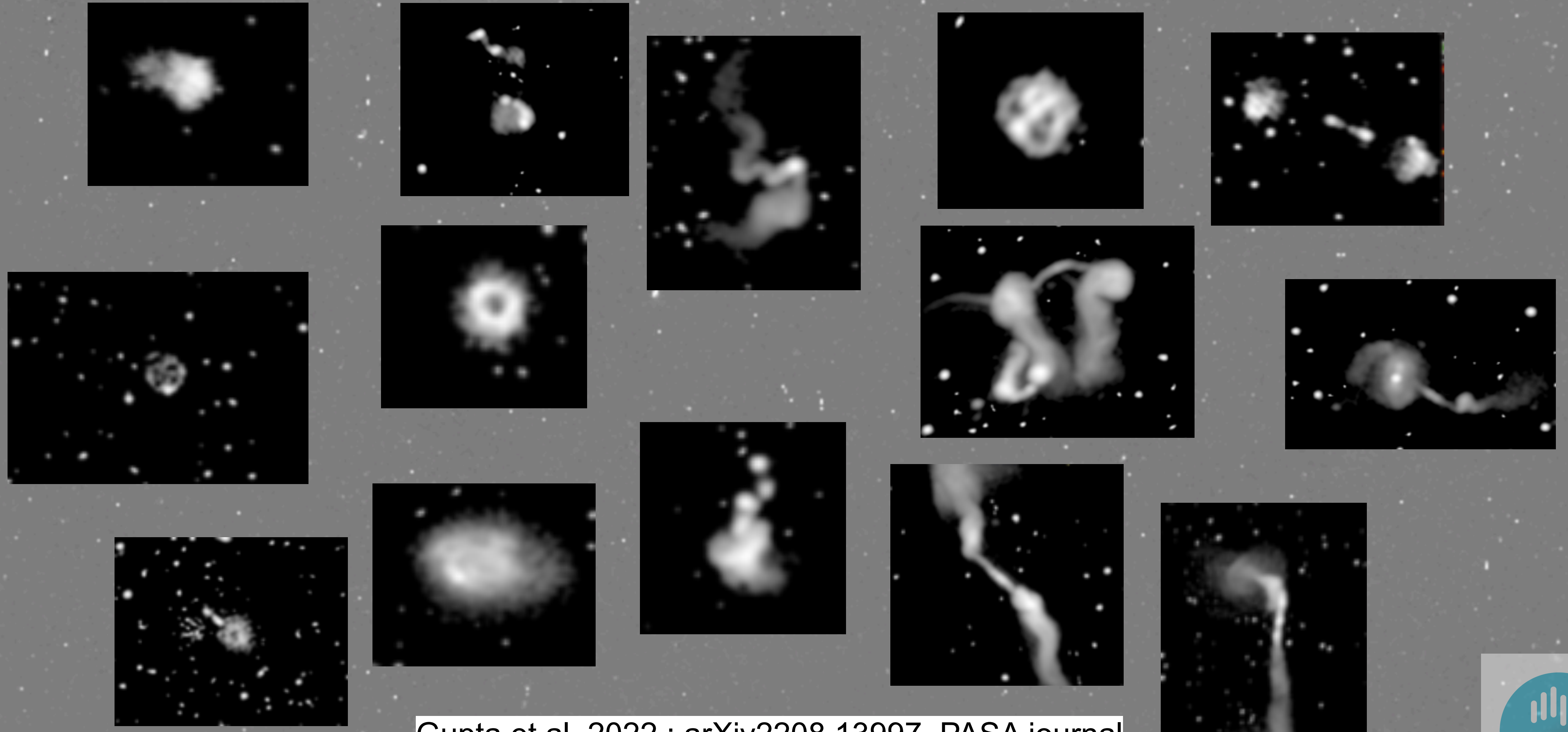
Evolutionary Map of the Universe (EMU; Norris 2021) - 1st Pilot Survey

- Covers 270 deg² of sky with declination
- RMS sensitivity of 25 – 35 μ Jy/beam
- Beamwidth of 13" \times 11" FWHM
- ~41,000 complex radio components (~220K total)

Radio Map (~6x3 deg²) – Evolutionary Map of Universe Pilot Survey



Radio Map (~6x3 deg²) – Evolutionary Map of Universe Pilot Survey



Gupta et al. 2022 : arXiv2208.13997, PASA journal



Machine Learning Frameworks

- **Supervised Learning**

Complete information of truth-labels is available for all images

- **Self-supervised Learning**

No information about truth-labels is available for all images

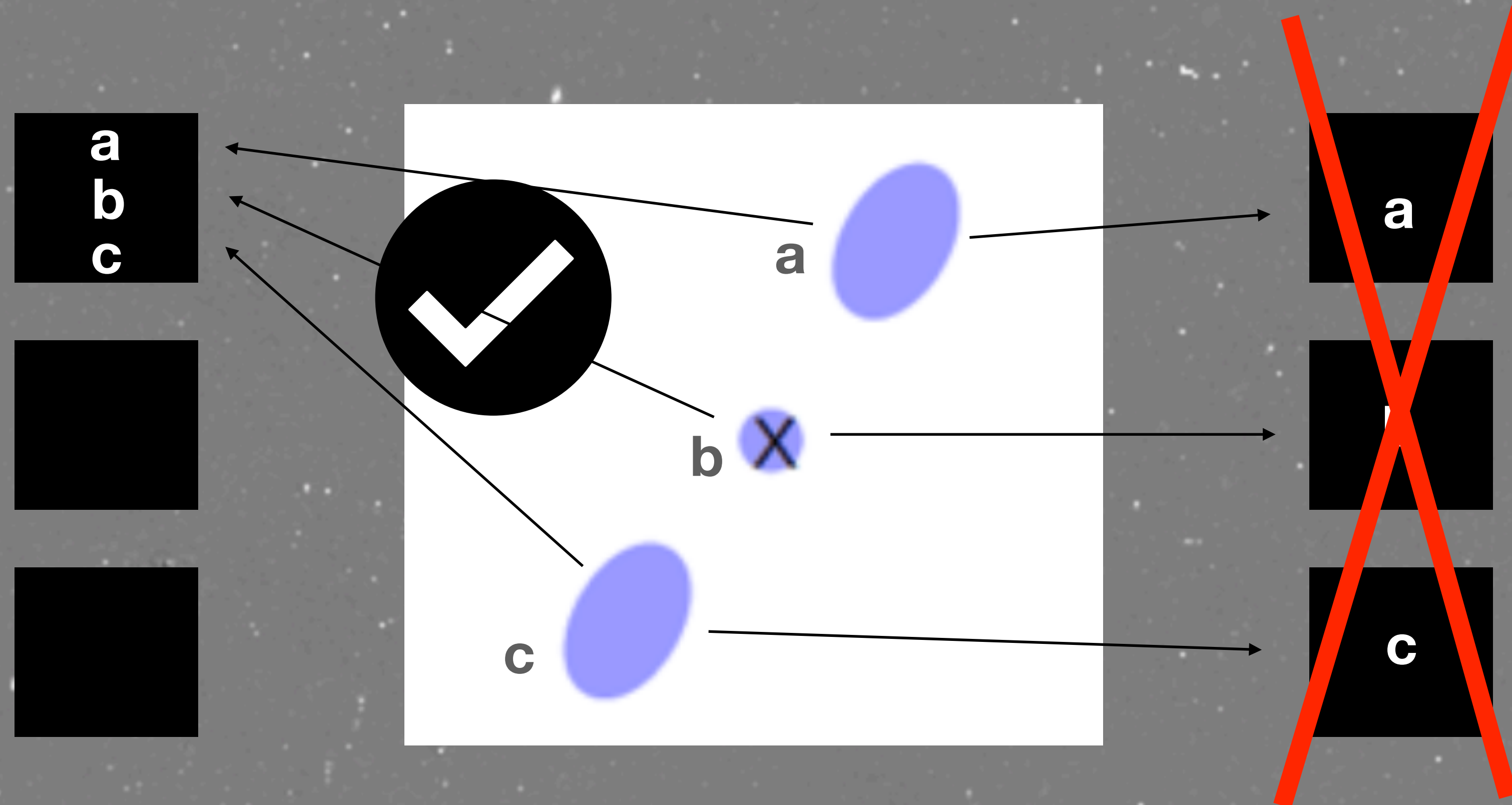
- **Weakly-supervised Learning**

Some information about truth-labels is available for all images

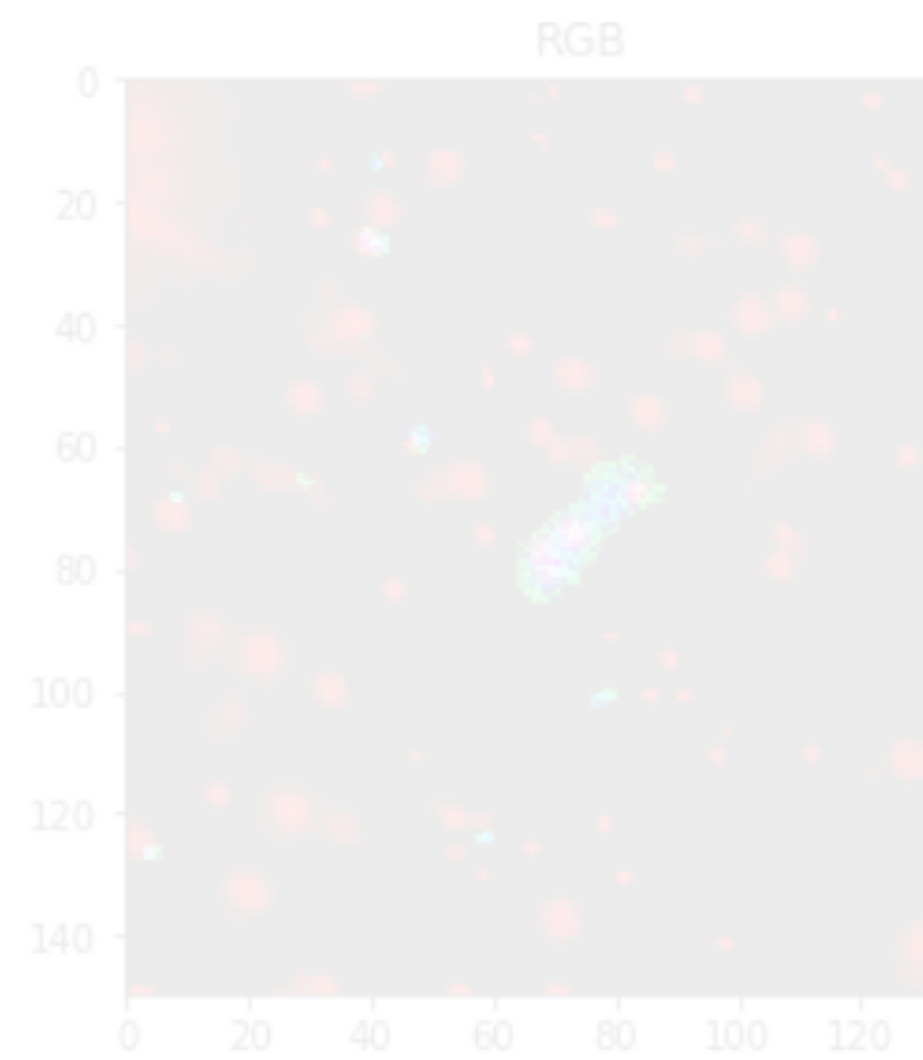
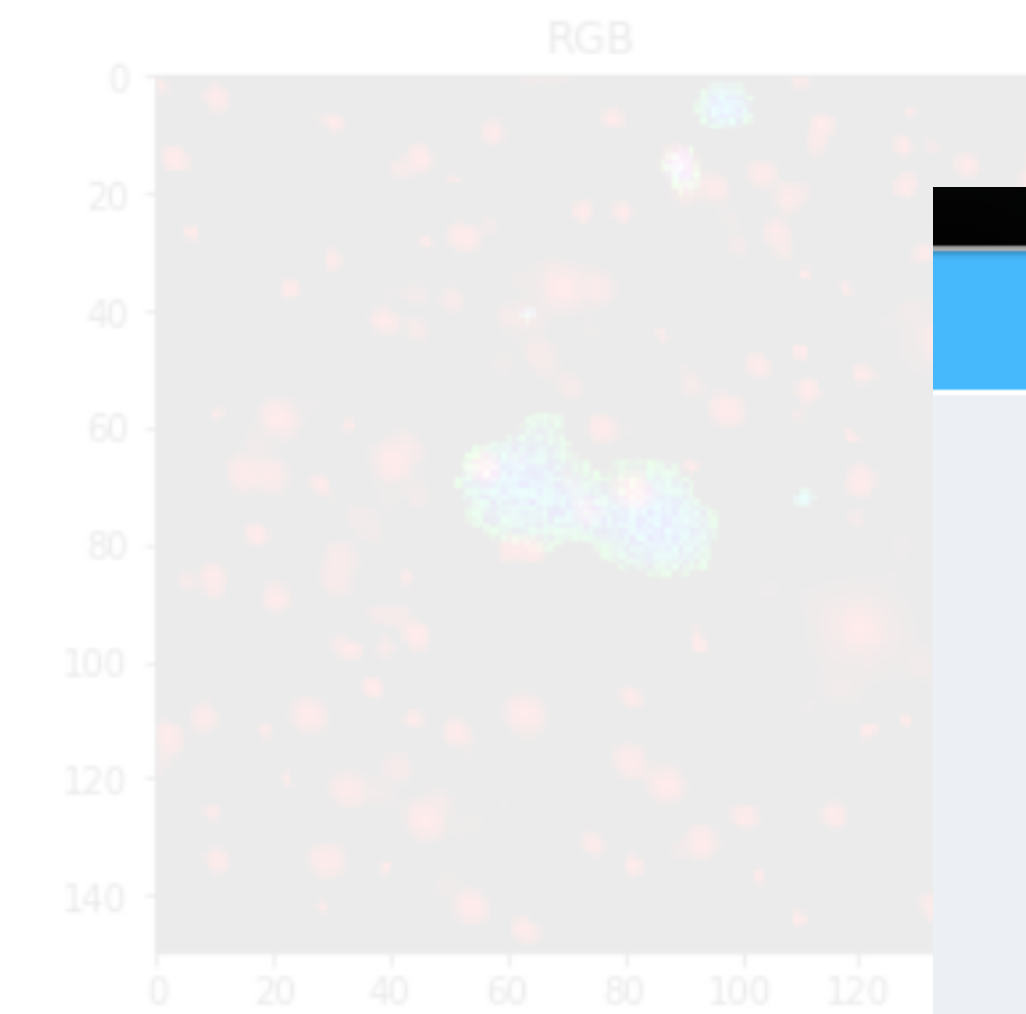
- **Semi-supervised Learning**

Complete information of truth-labels is available for some images

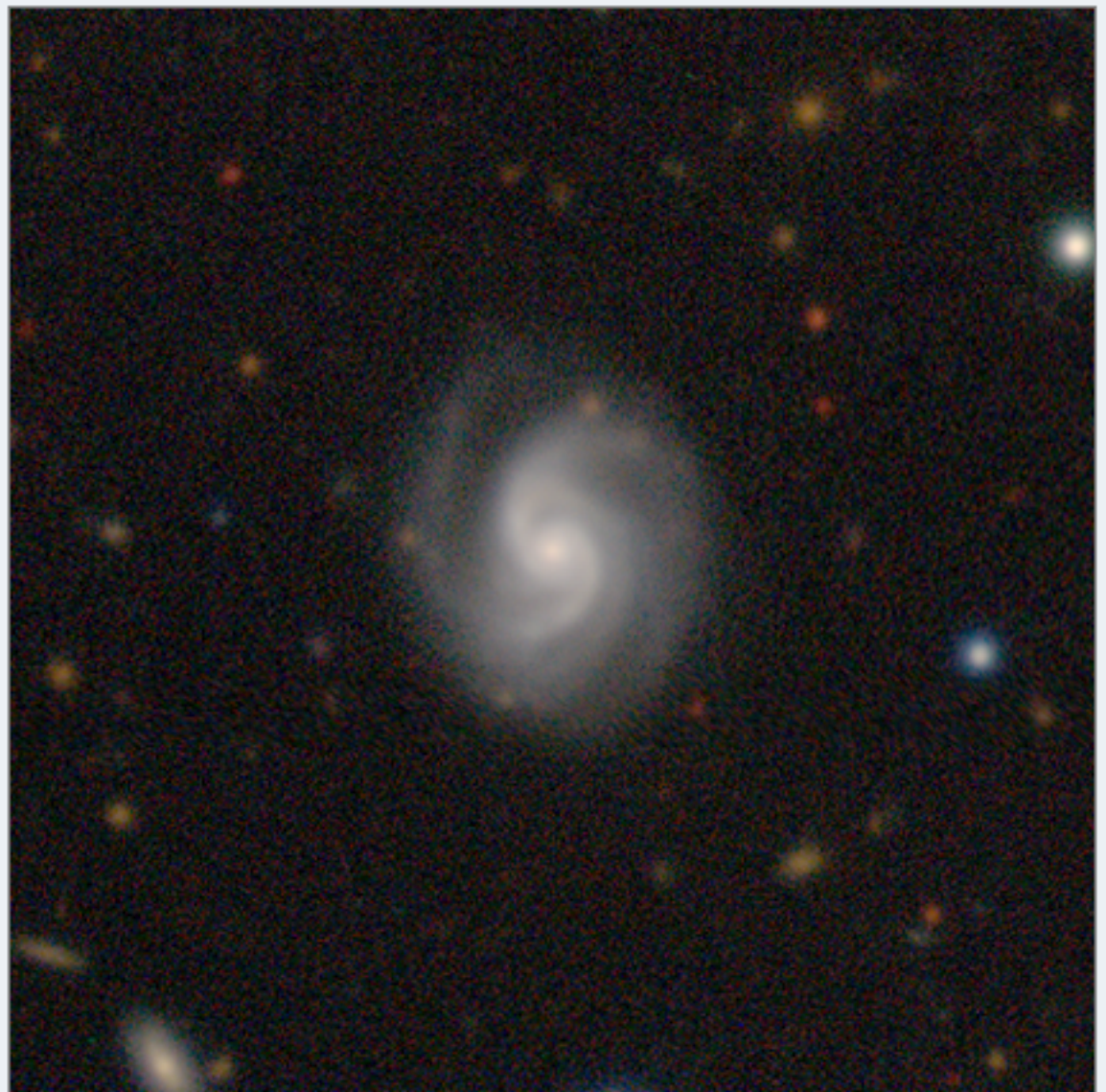
Can we group components of Radio Galaxies using ML?



Weakly-supervised Learning Citizen Science



Welcome to the new Galaxy Zoo!



Is the galaxy simply smooth and rounded, with no sign of a disk?

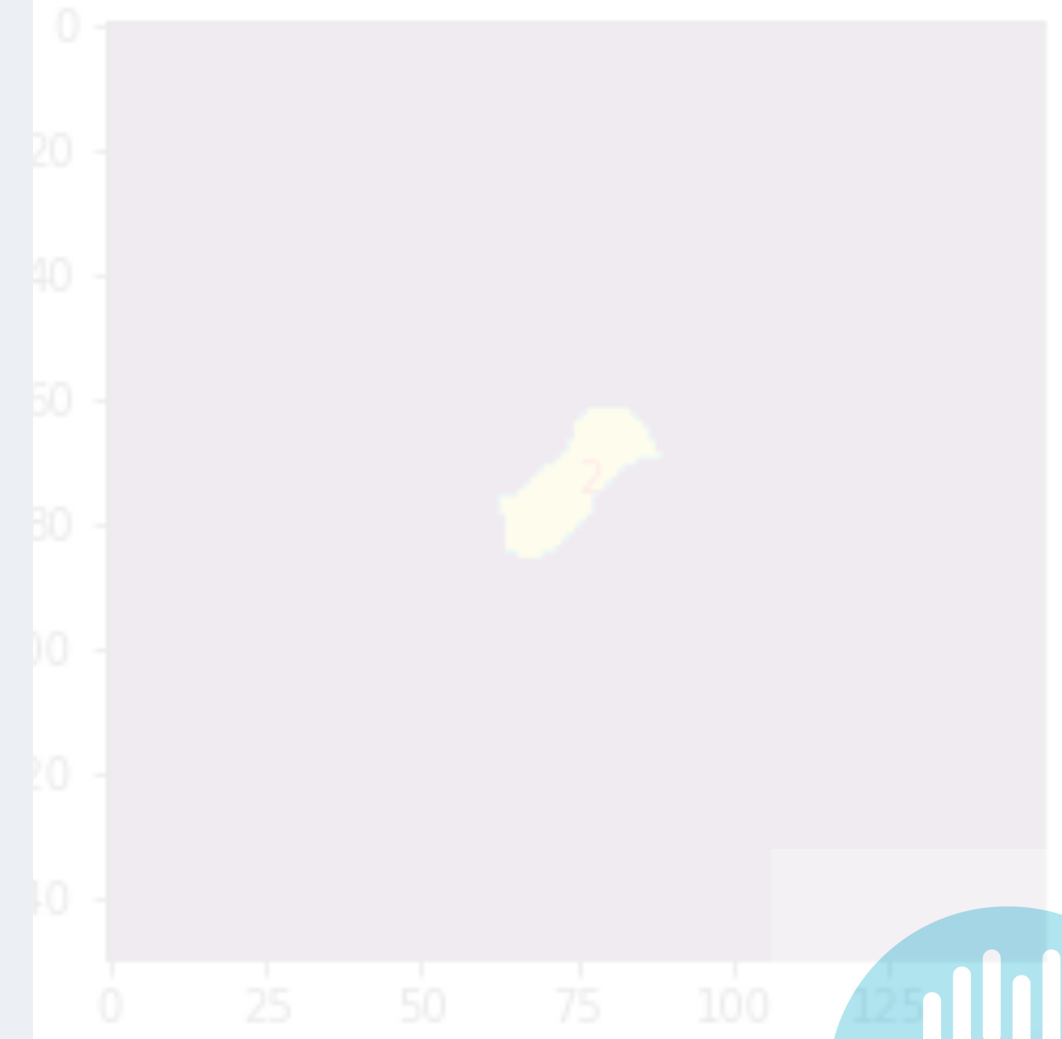
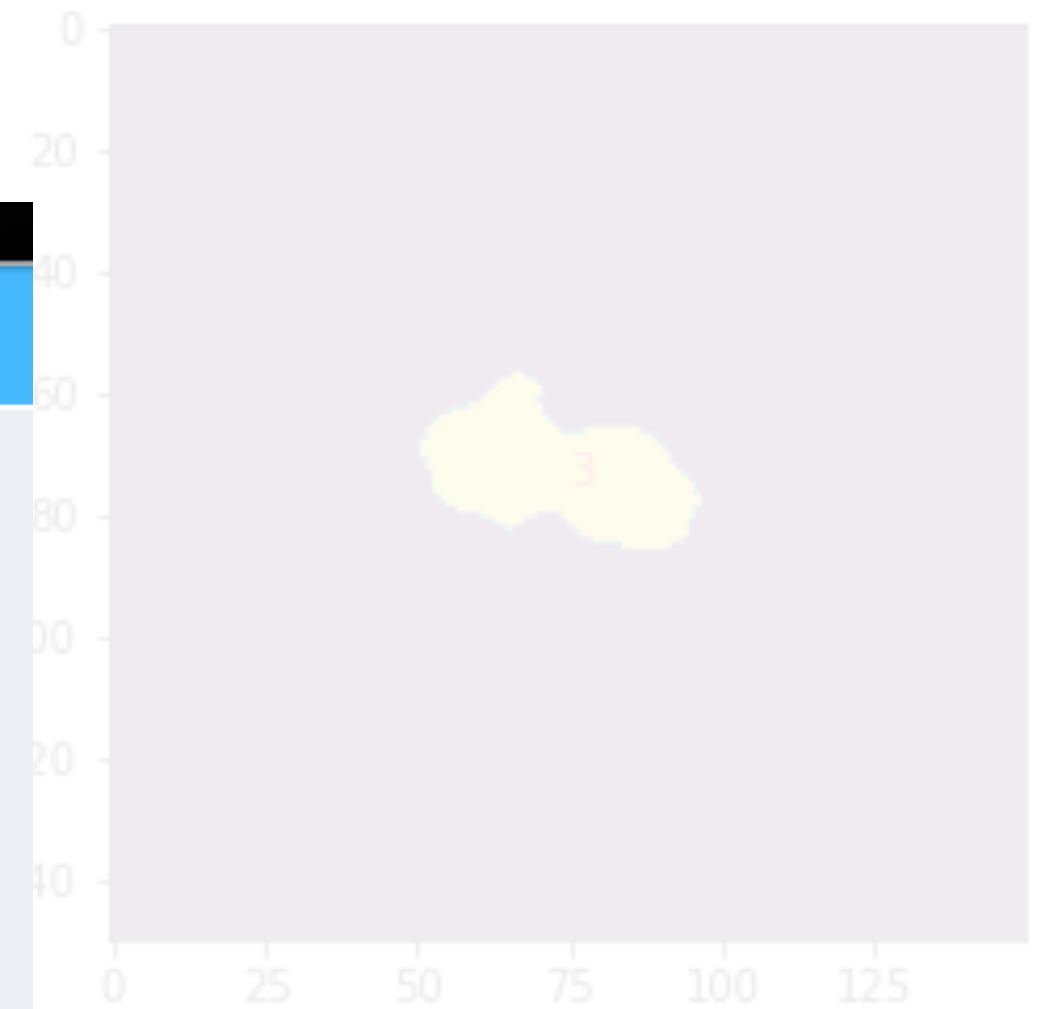
- Smooth
- Features or Disk
- Star or Artifact

Need some help with this task?

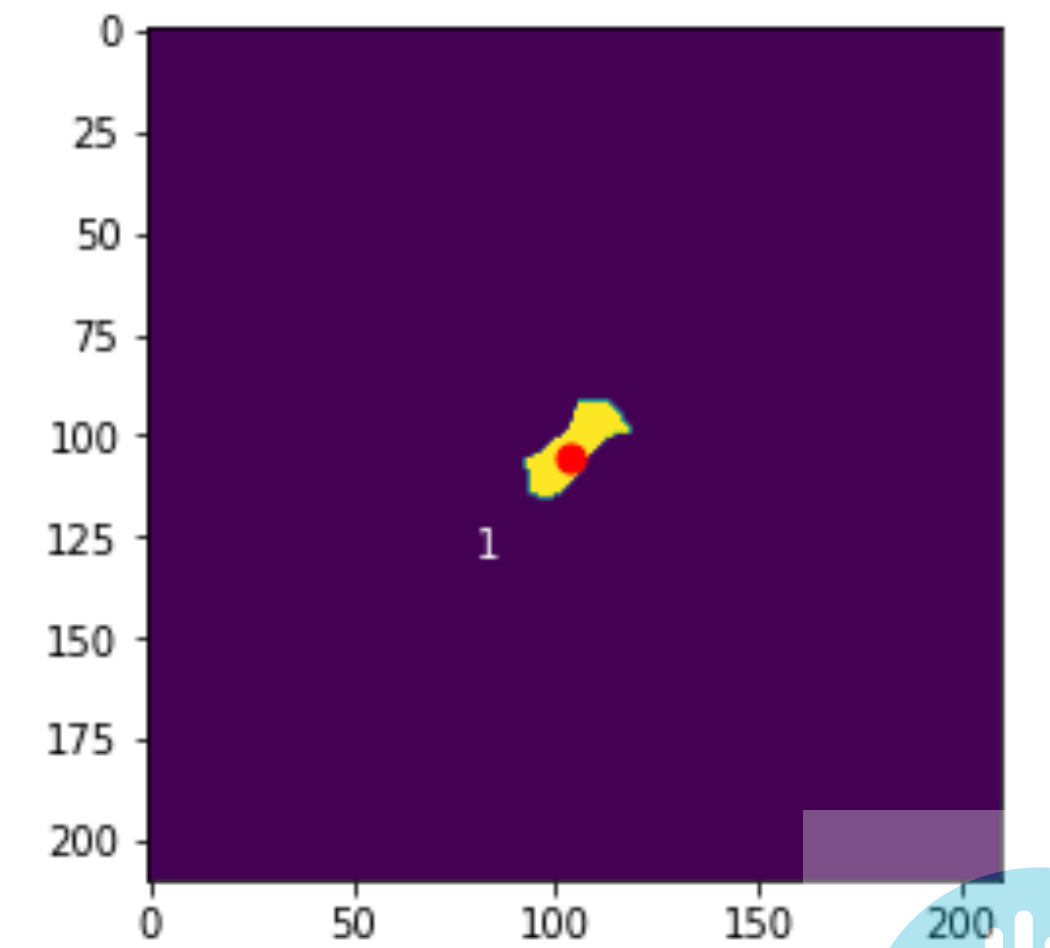
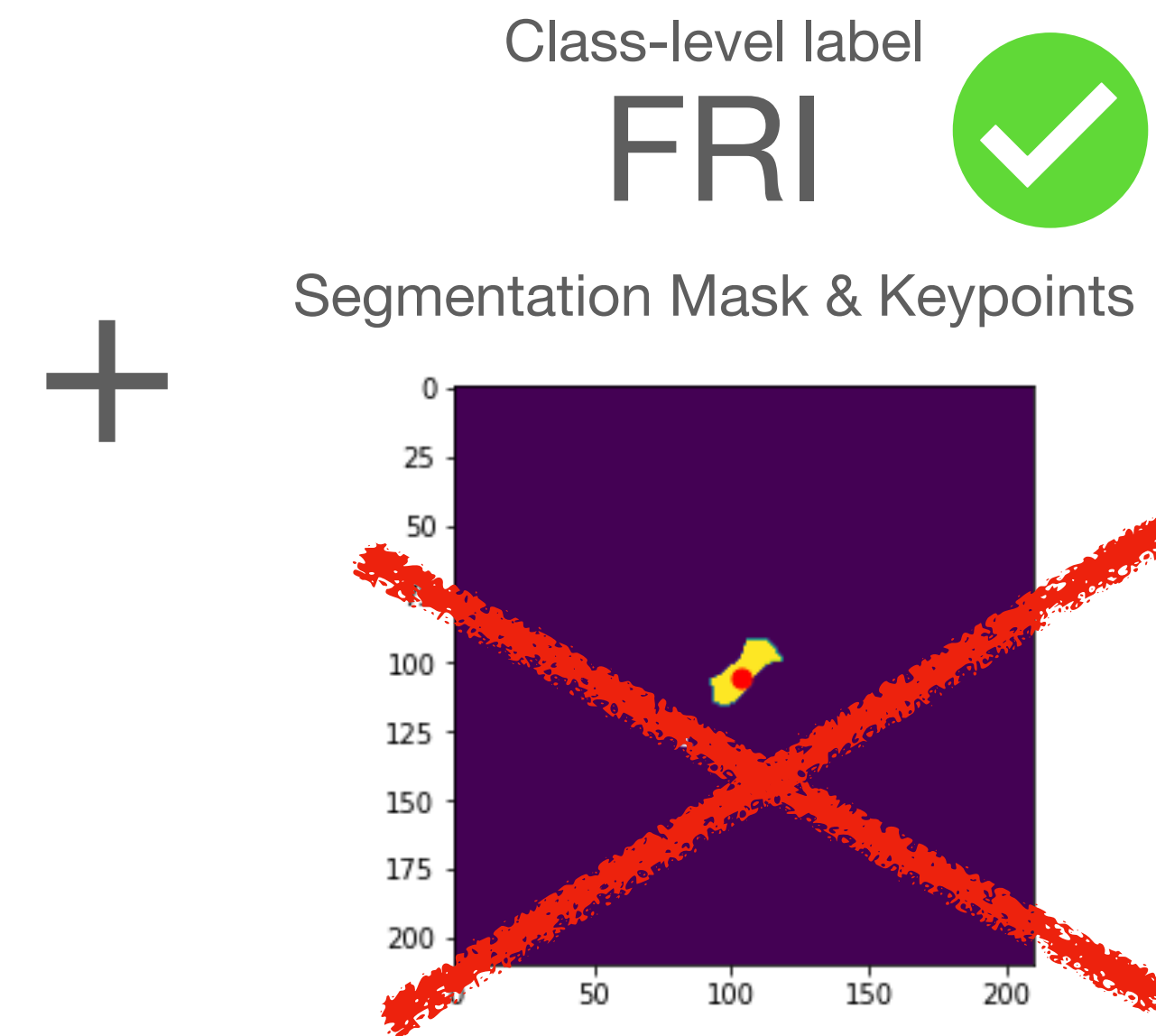
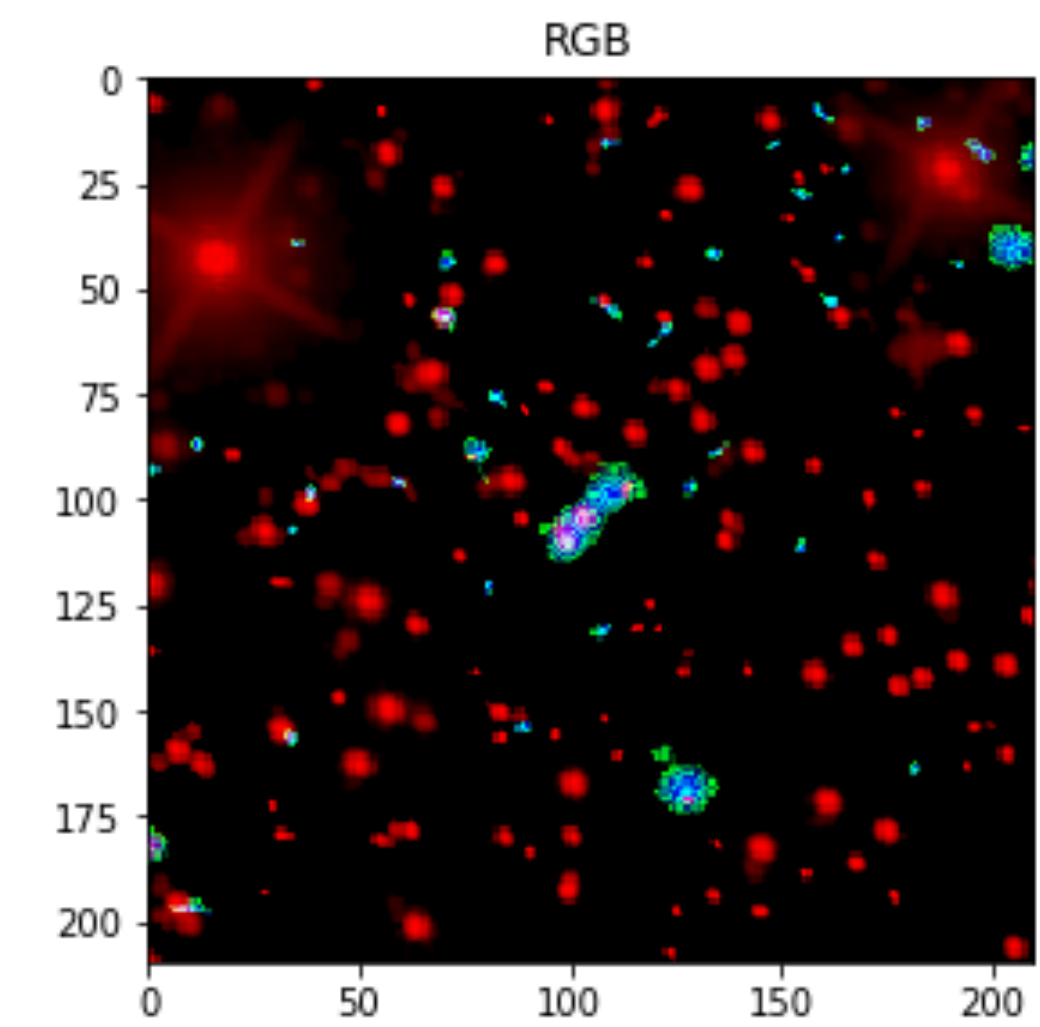
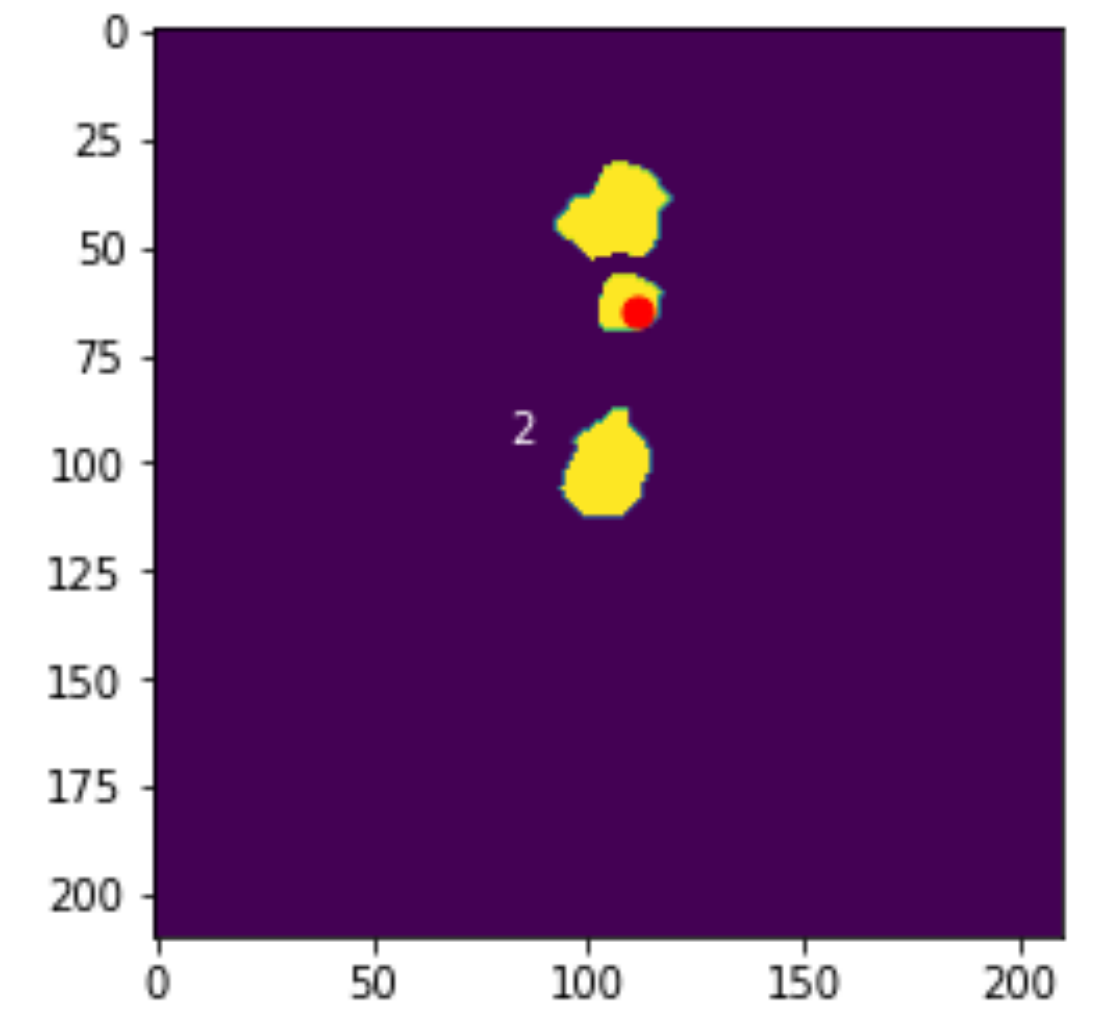
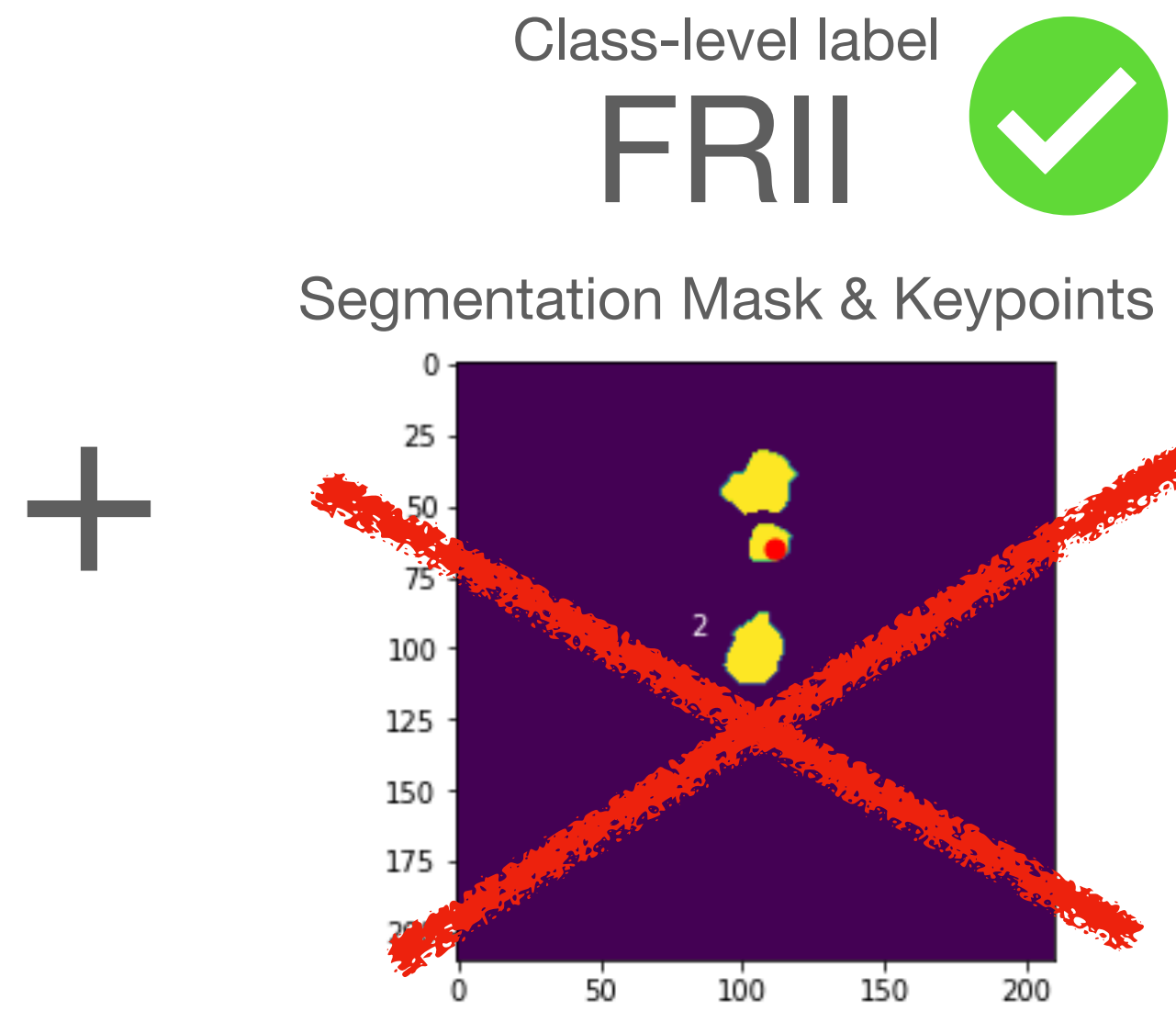
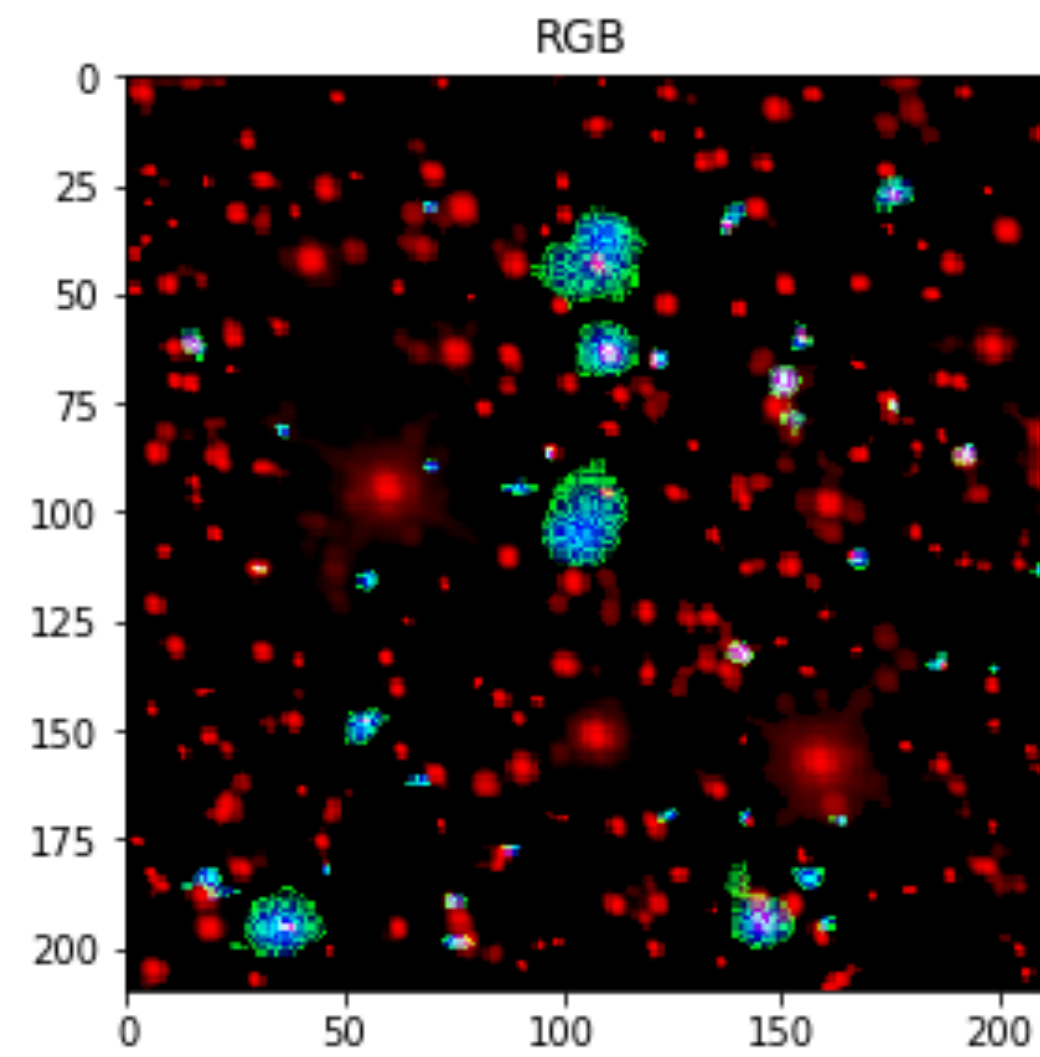
Back Done & Talk Done

Show the project tutorial

This is a screenshot of the Galaxy Zoo citizen science interface. It features a central image of a galaxy for classification. Below the image are three radio button options: 'Smooth', 'Features or Disk', and 'Star or Artifact'. There are also buttons for 'Need some help with this task?', 'Back', 'Done & Talk', 'Done', and 'Show the project tutorial'. The interface has a blue header and a light grey background.



Weakly-supervised Learning



Gupta et al. (in prep)



Machine Learning Frameworks

- **Supervised Learning**

Complete information of truth-labels is available for all images

- **Self-supervised Learning**

No information about truth-labels is available for all images

- **Weakly-supervised Learning**

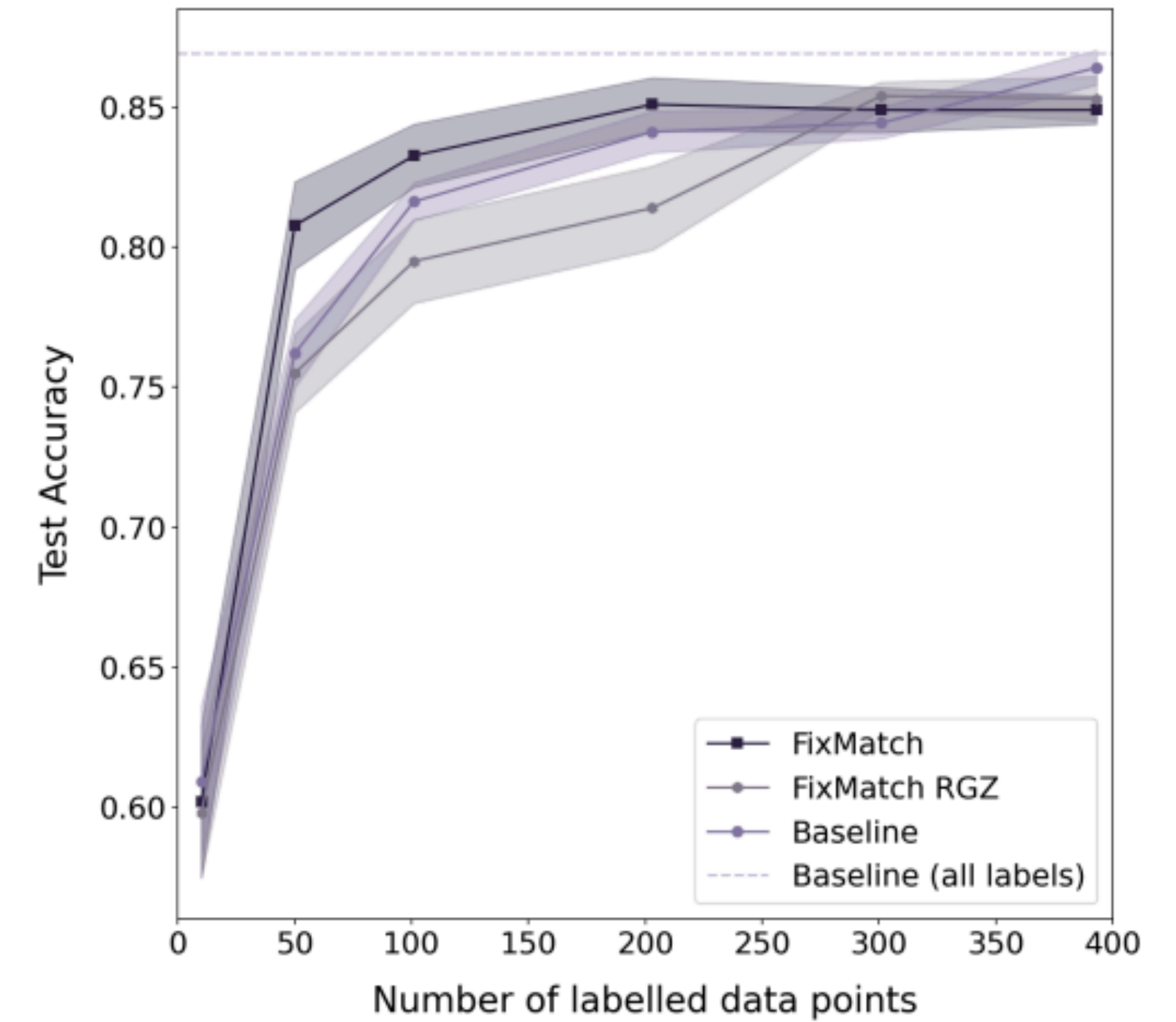
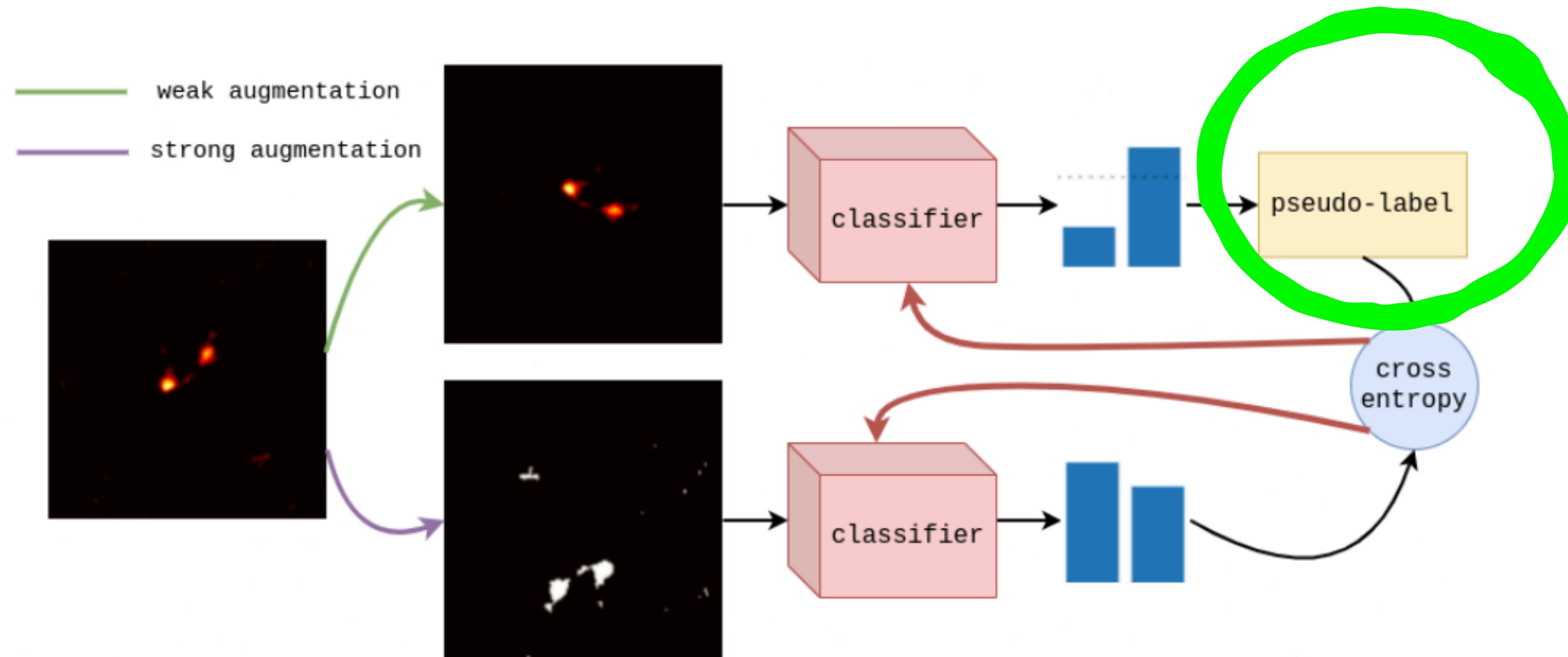
Some information about truth-labels is available for all images

- **Semi-supervised Learning**

Complete information of truth-labels is available for some images

Semi-supervised Learning in Astronomy Examples

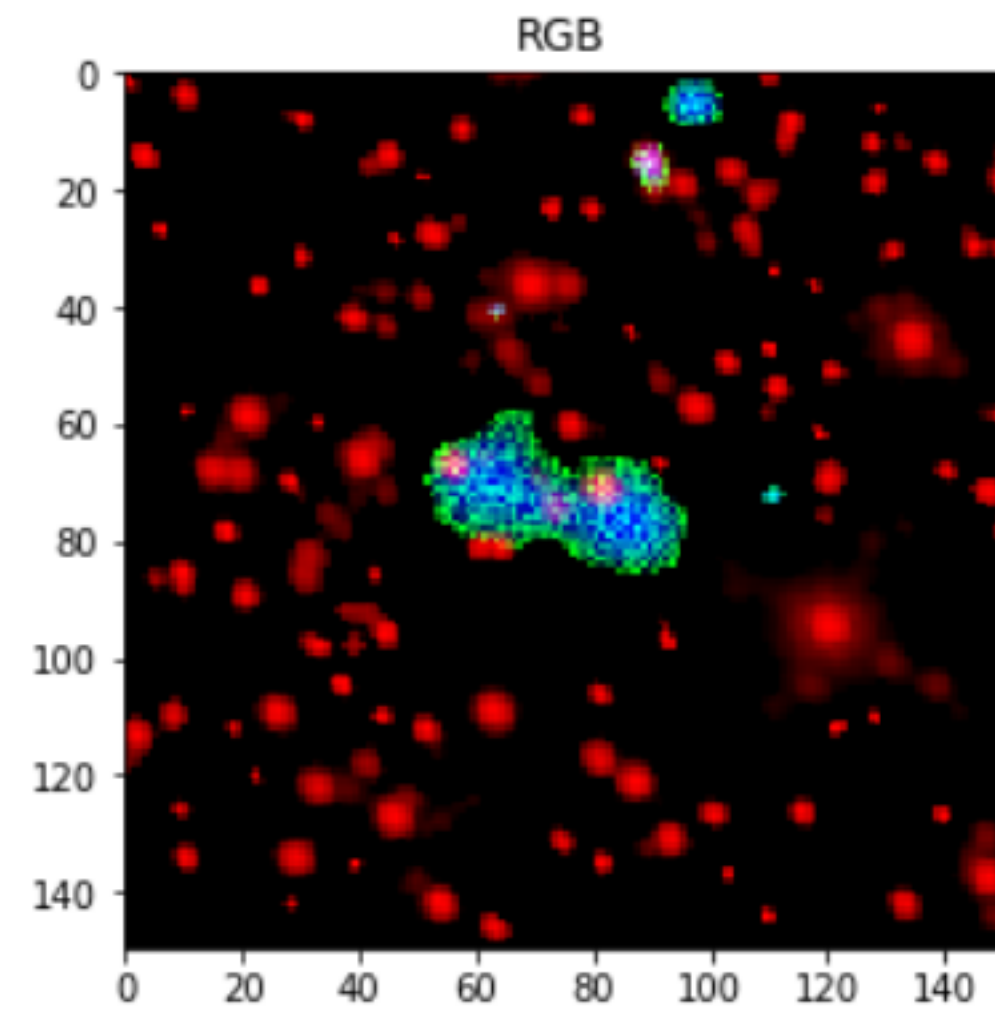
(1) feature extraction, (2) clustering/**classification (few labels)**, (3) visual representation



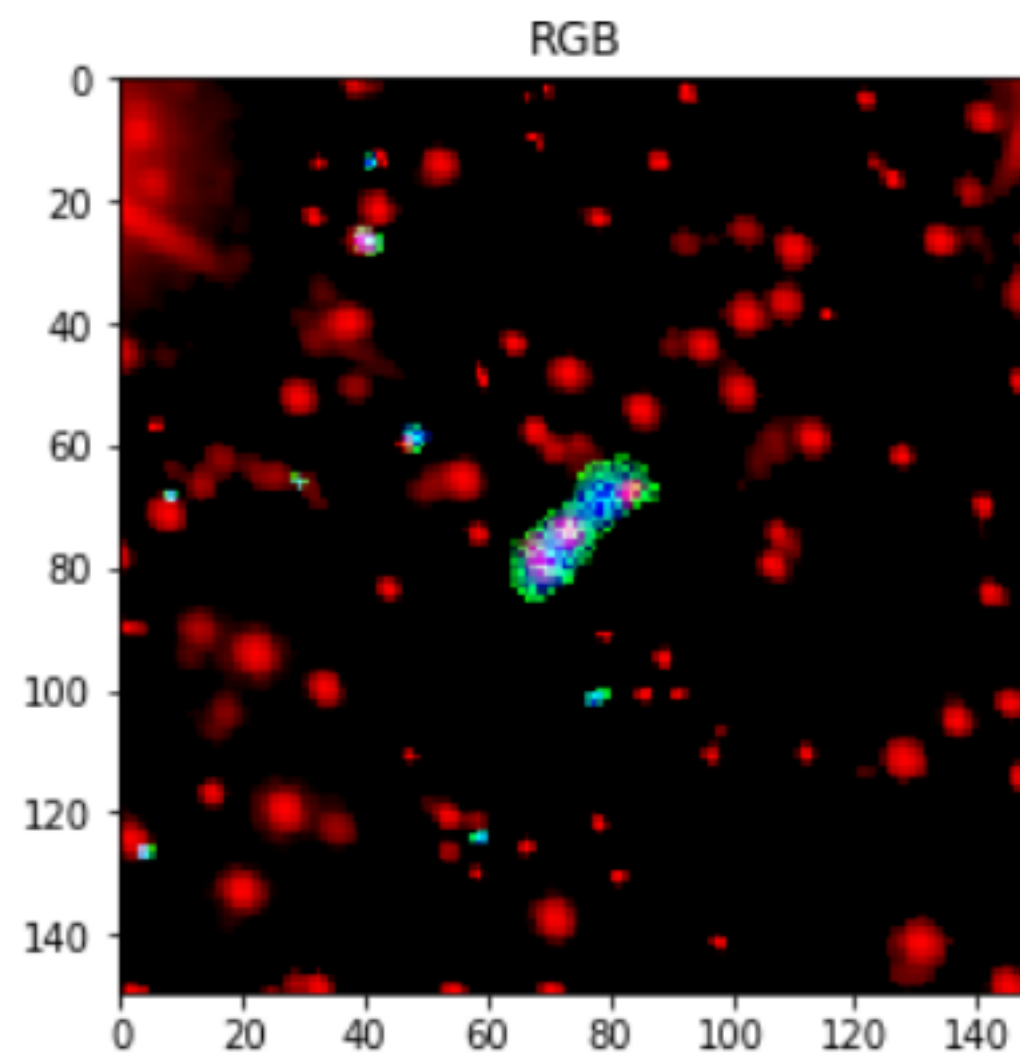
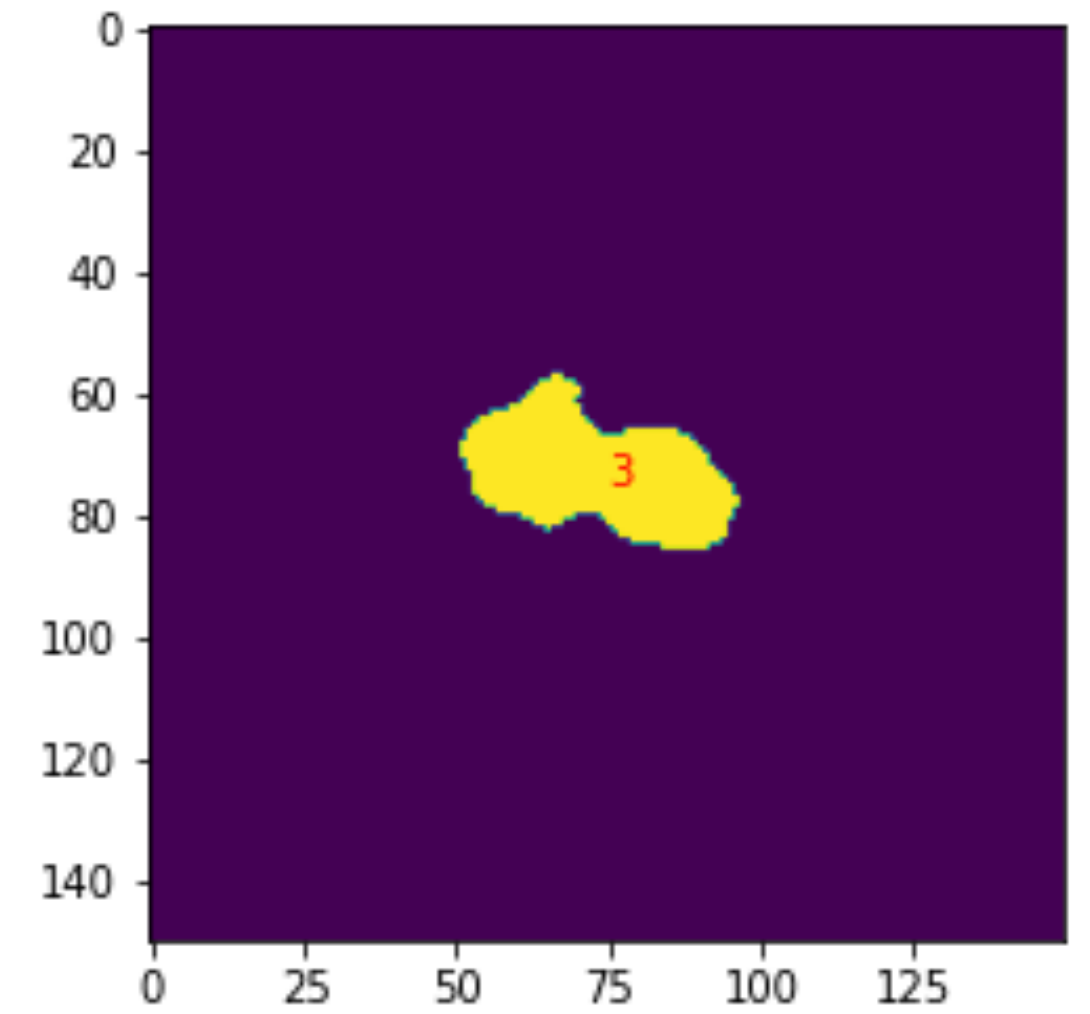
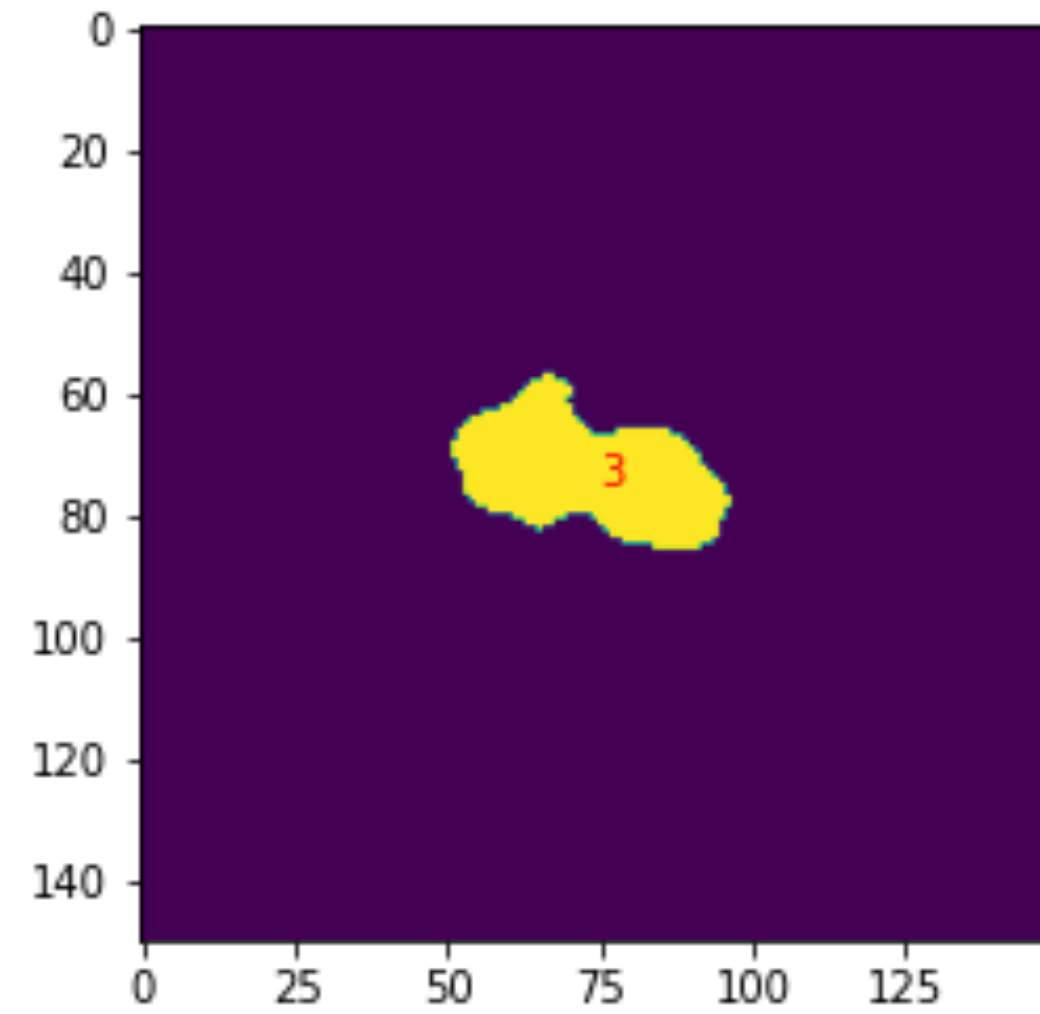
Slijepcevic et al. 2021



Semi-supervised Learning

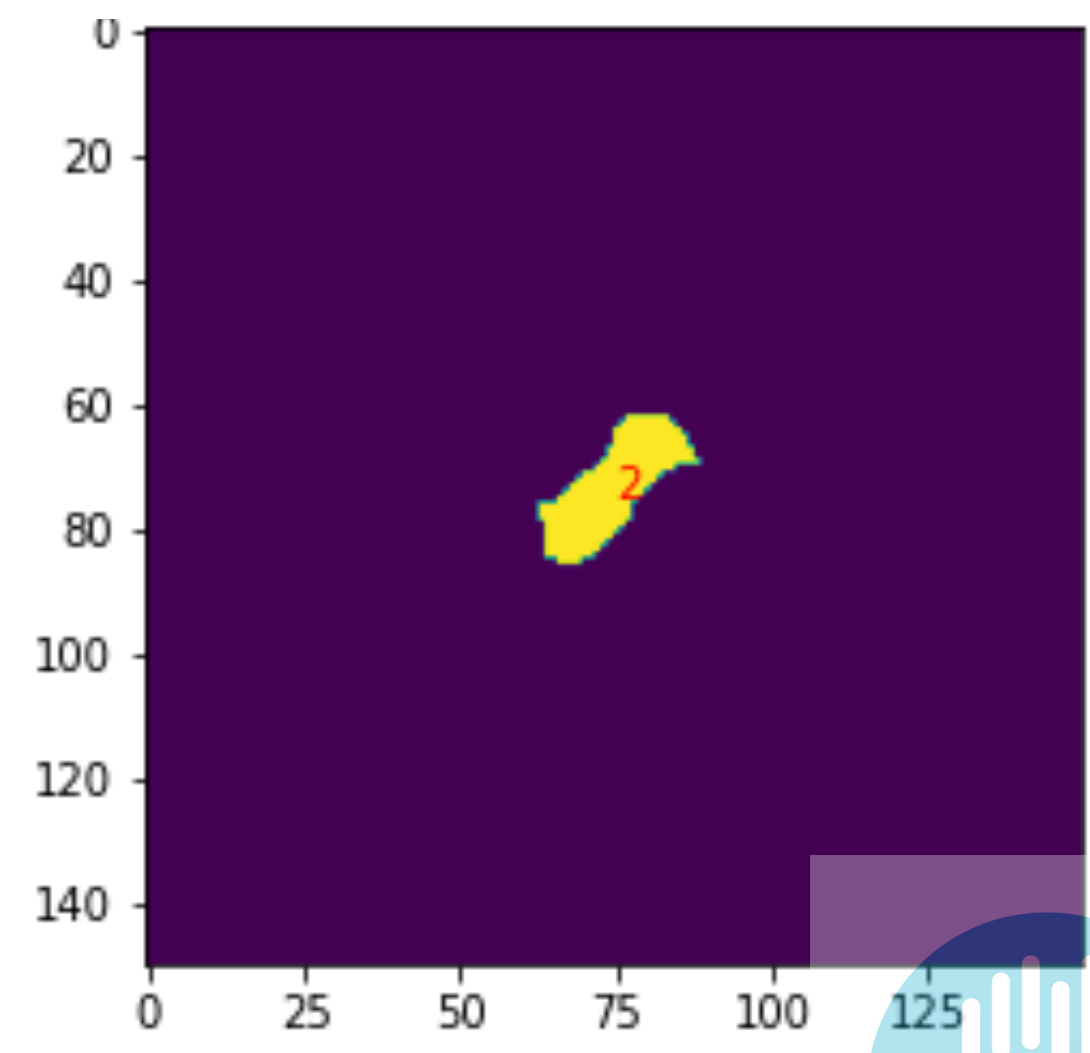


+



+

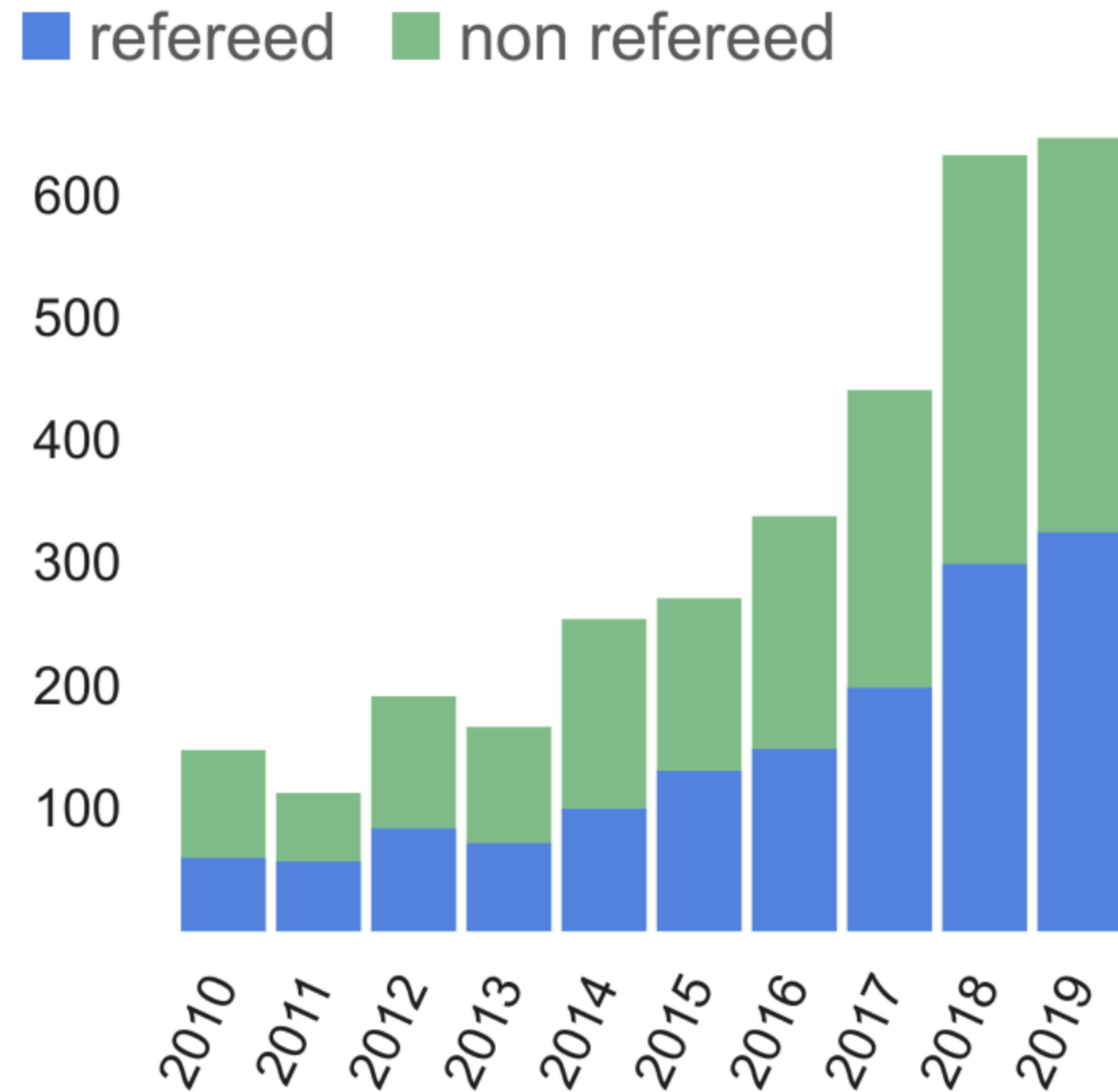
No label
Or
Class-label



Gupta et al. (in prep)



Cosmology papers that include machine learning methods in the abstract or title!



Thanks!
See you later for discussions!