UNIVERSITÉ DE GENÈVE

SKACH

# Self-Supervised Learning for MeerKAT Images

...

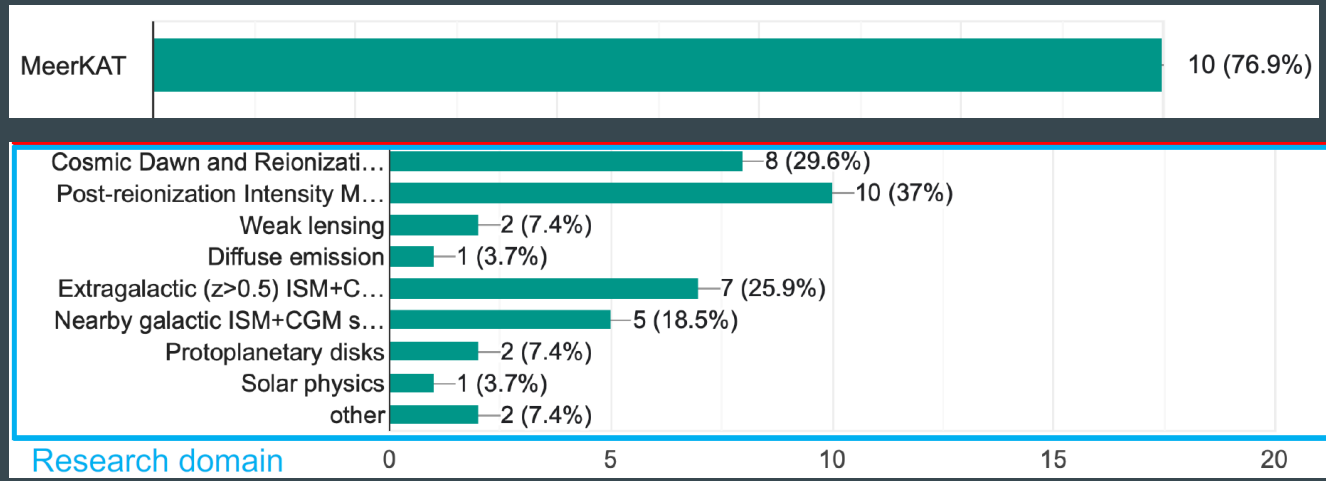
E. Lastufka

O. Bait, O. Taran, V. Kinakh, M. Audard, M. Dessauges-Zavadsky, T. Holotyak, D. Schaerer, S. Voloshynovskiy

# Motivation

Deliver products useful to the SKA community

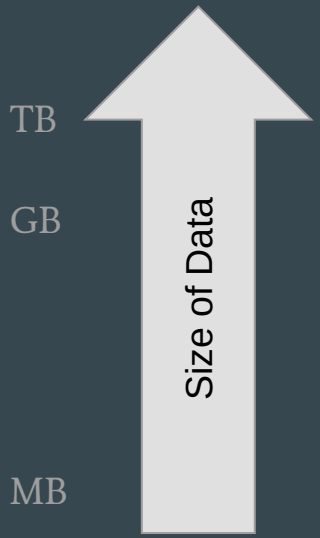
- Adaptable to diverse range of science topics
- Scalable



# MeerKAT Datasets

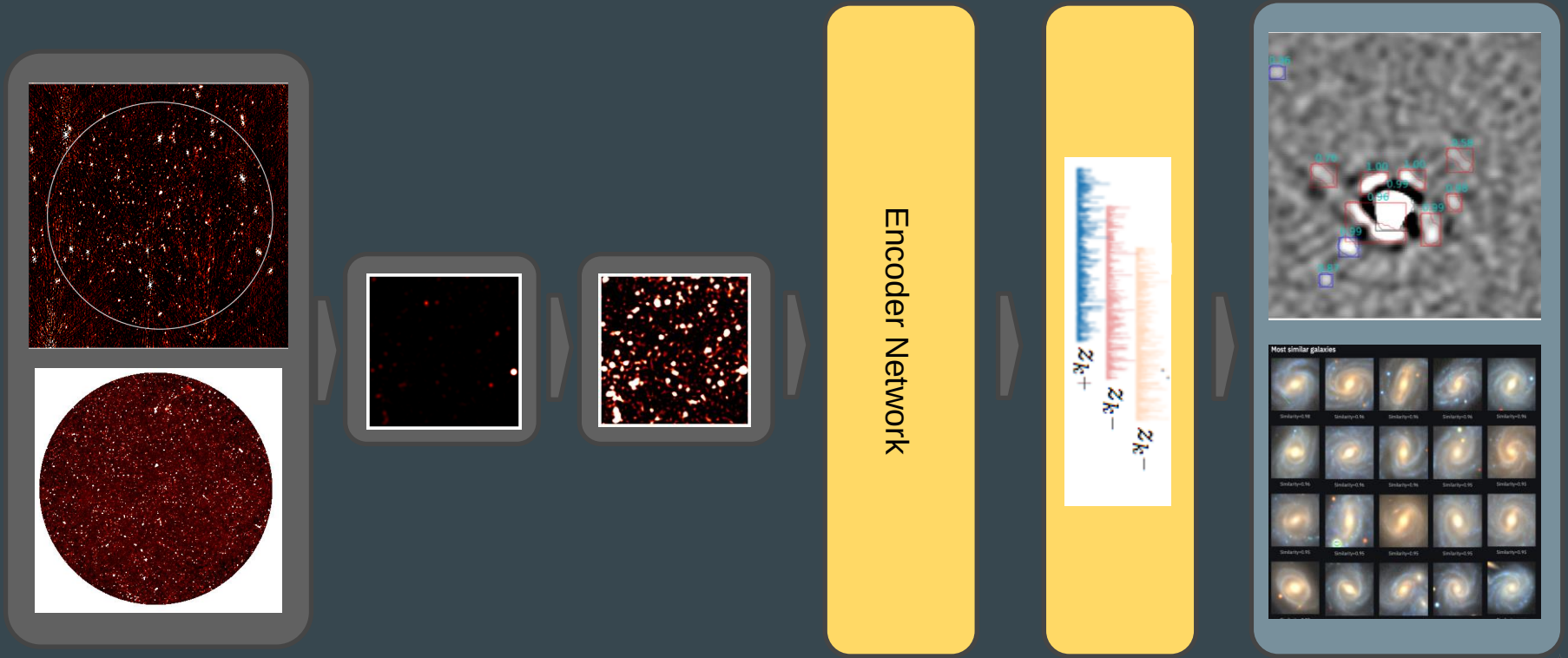
- L Band (900 MHz - 1.6 GHz)
- Continuum sensitivity of 1-2 uJy/beam
- 6 - 15 hours time-on-source
- CLEAN images, source catalogs publicly available


# Data Formats

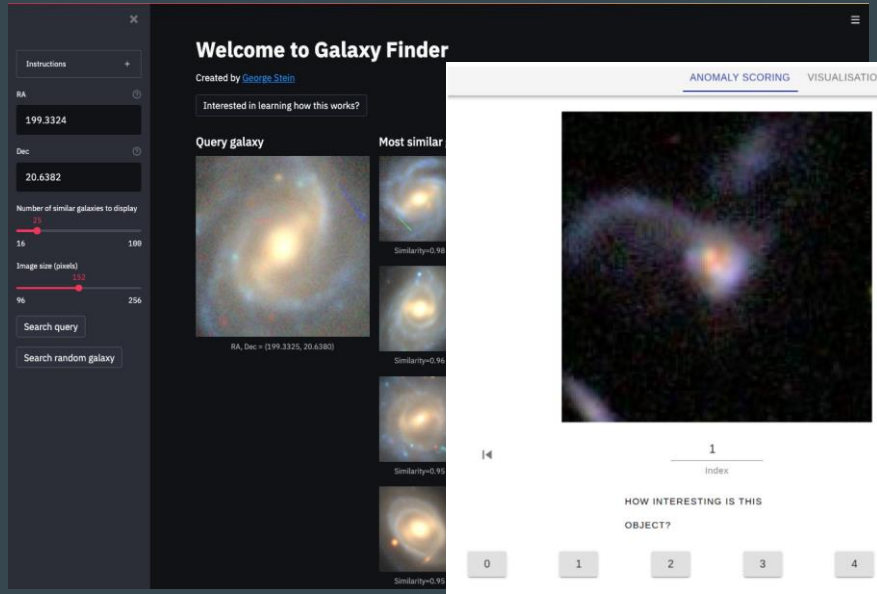



- Raw visibilities
- Calibrated visibilities
- Time-averaged visibilities
- Self-calibrated visibilities (phase-only)
- Self-calibrated visibilities (phase and amplitude)
- Dirty images
- CLEAN images
- CLEAN images with extra processing (self-cal, source peeling, etc)

# Goal: Learn useful representations of radio data



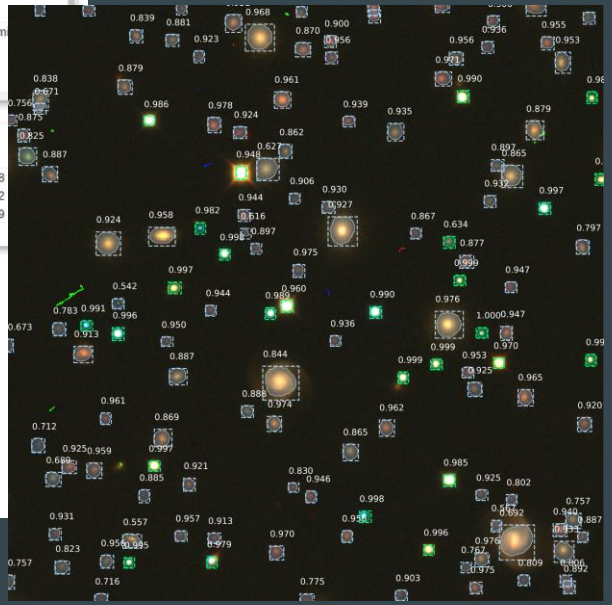
# Downstream Tasks




**Galaxy Finder** (Stein 2021)

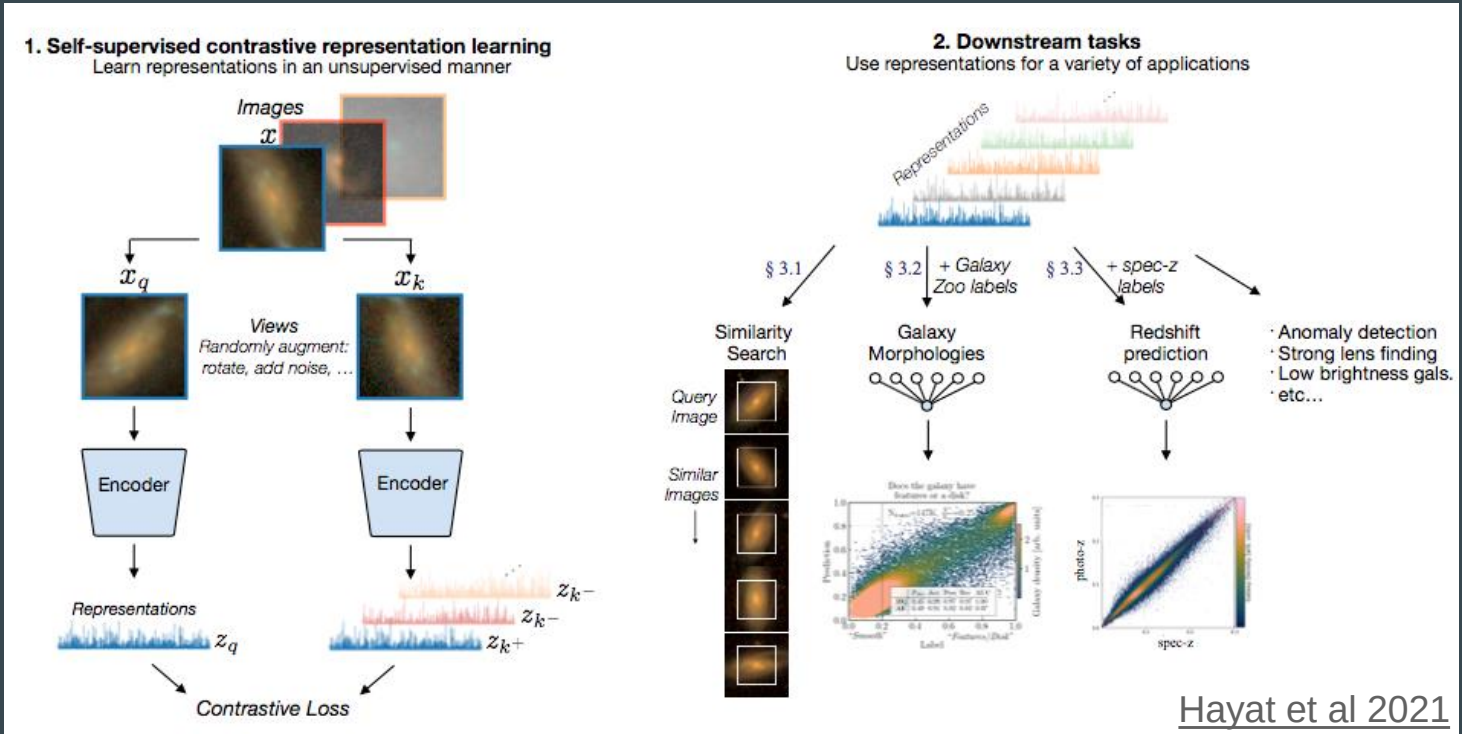



**Astronomaly** (Lochner 2021)



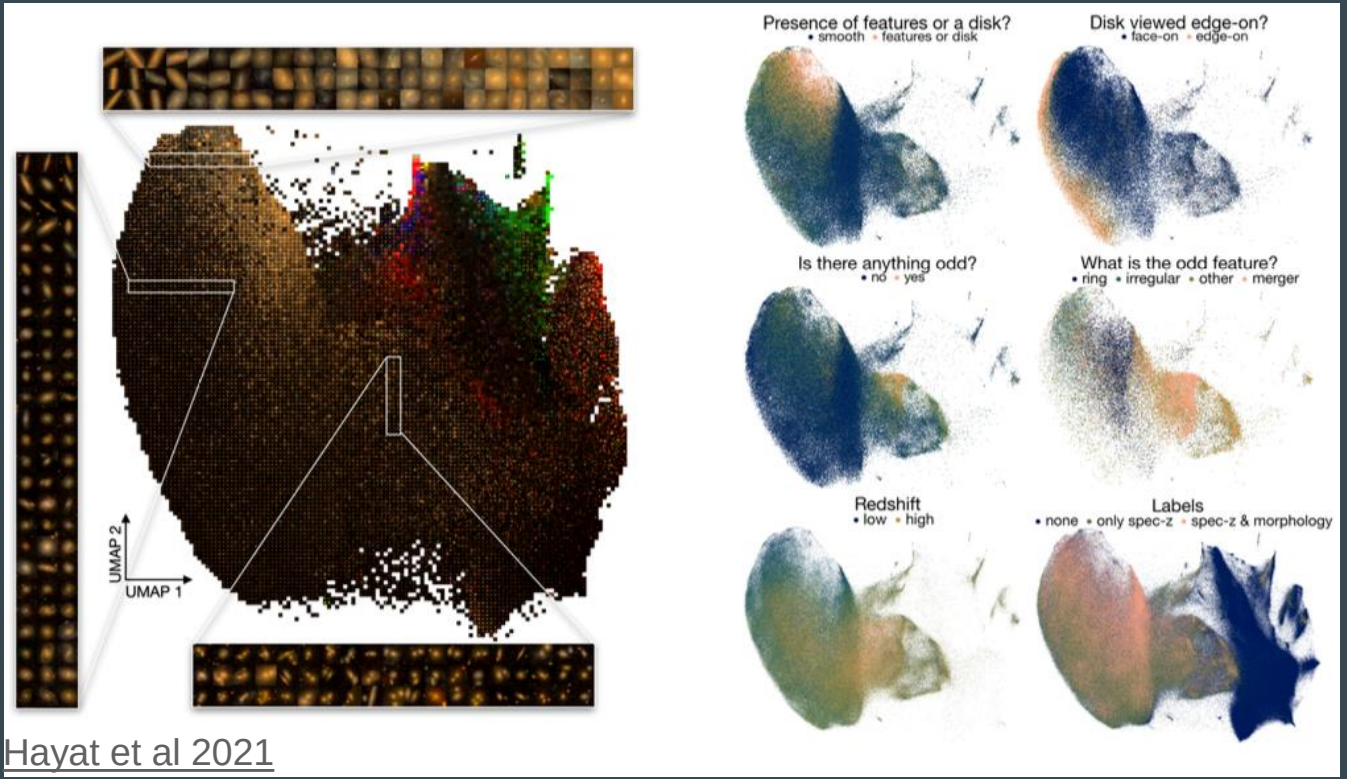
**Astro-rcnn** (Burke 2019)

# What is possible?



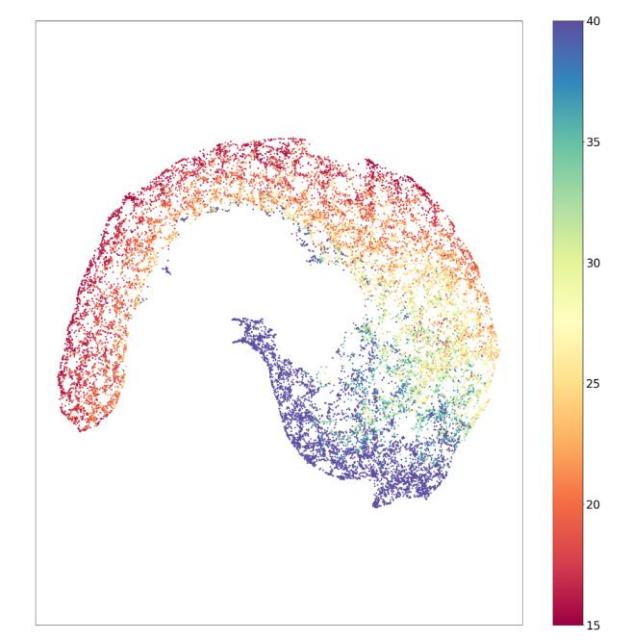

Hayat et al 2021

# What is possible?




Hayat et al 2021

# What is possible?




I. Slijepcevic et al, 2022

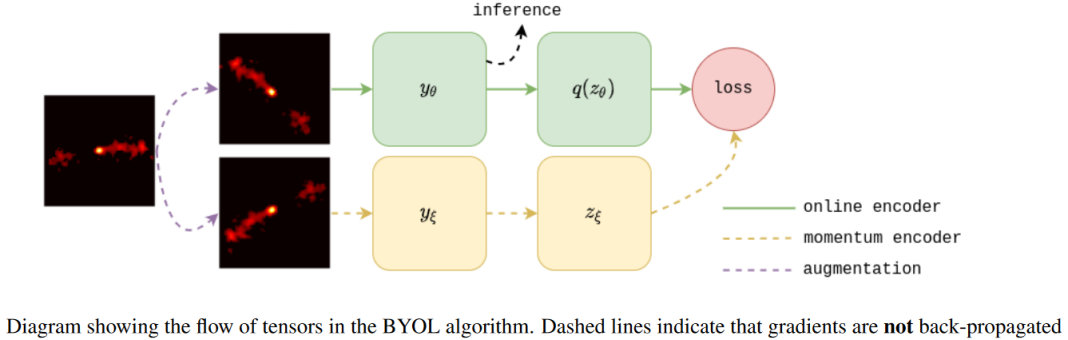



Diagram showing the flow of tensors in the BYOL algorithm. Dashed lines indicate that gradients are **not** back-propagated

Hayat et al 2021

# Challenges

- Difficult to generalize from one survey to another
- Astrophysical images can be very large and contain many sources
- High dynamic range
- Need task-agnostic image augmentations

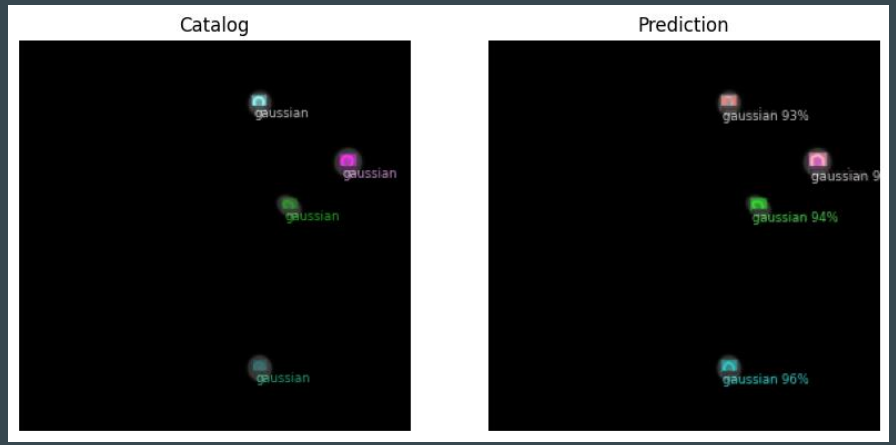
# Radio-specific

- Work up until now has depended on CLEAN images
  - Slow to produce, often reconstructed to suit the science use case rather than in a standardized way
- Choice of channels is huge
- More variation possible than in optical

# Solutions

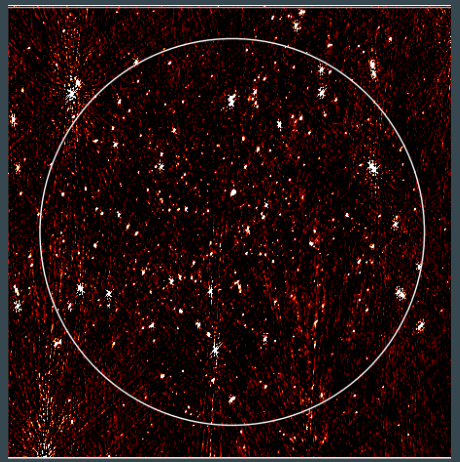
## Use simulations in parallel

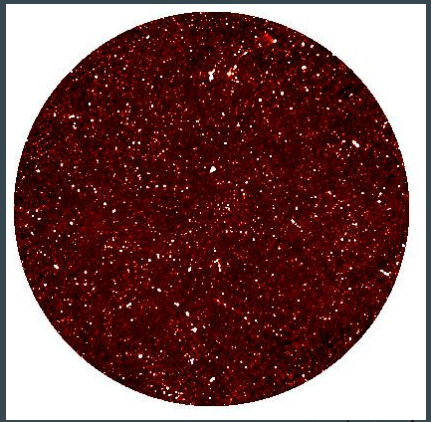
- Ability to change parameter space
- Reliable method of verification
- Early indicator of success



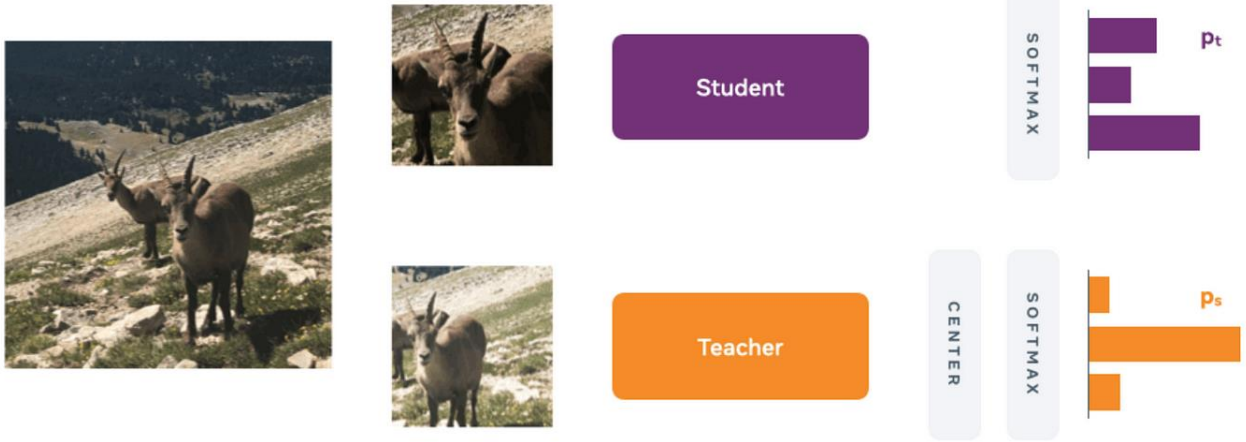

## Dirty Images

- Contains same information as CLEAN
- Lower processing level
- Potential to generalize better



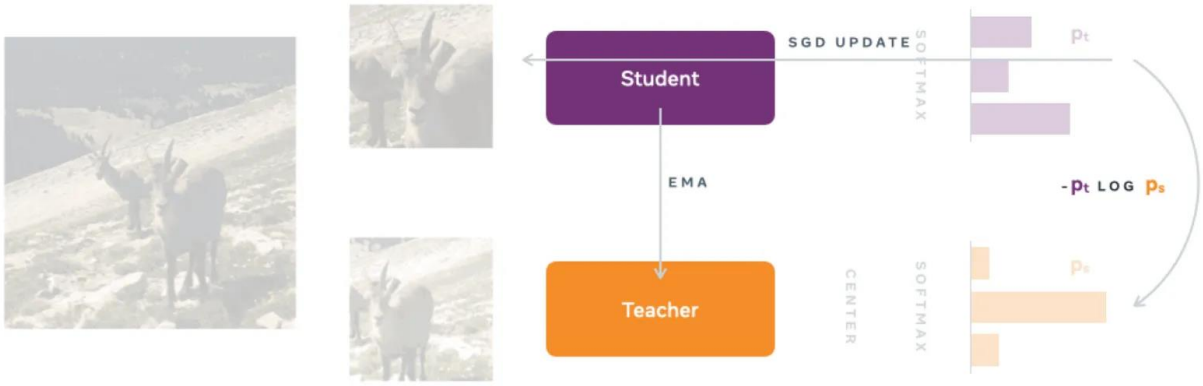


# Self-Supervised Learning with DINO



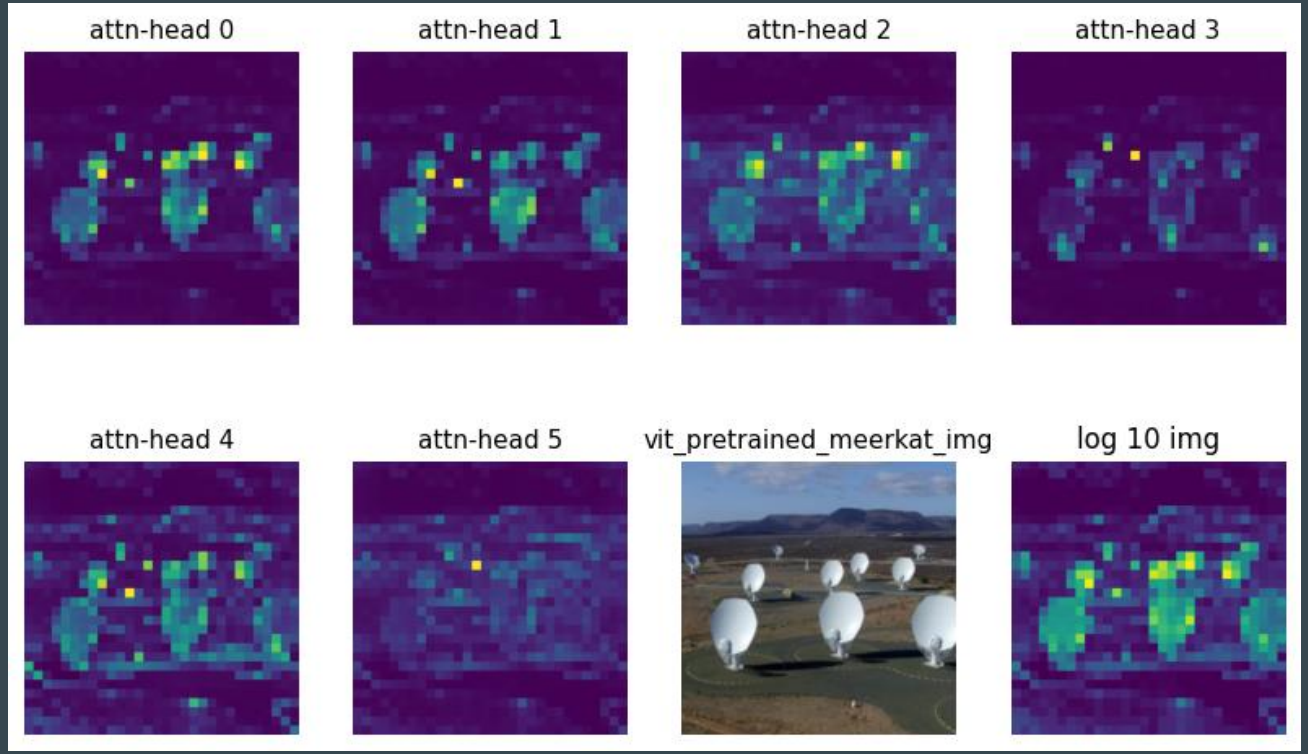

https://github.com/facebookresearch/dino

# Self-Supervised Learning with DINO

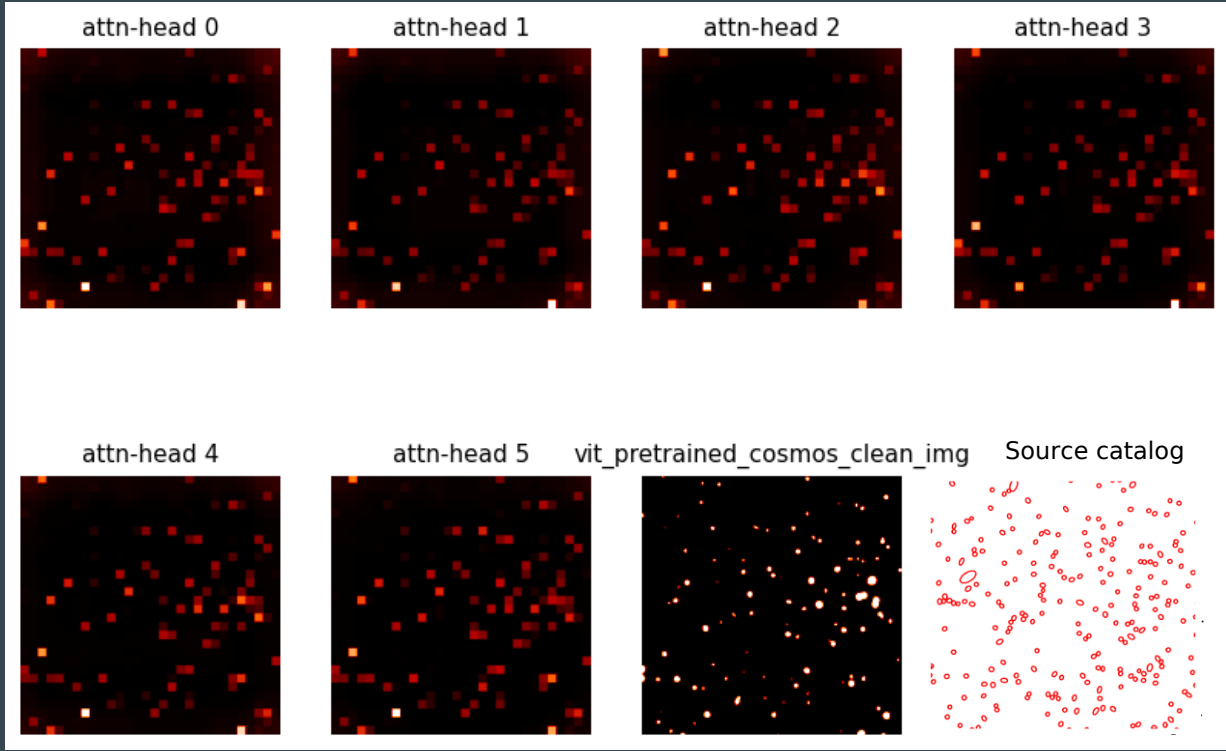



https://github.com/facebookresearch/dino

# DINO is a promising approach!



# DINO is a promising approach!



Motivational poster of a colorful stegosaurus with the words “Thank You”