Understanding galaxy cluster evolution with self-supervised machine learning

Urmila Chadayammuri, Lukas Eisert & Annalisa Pillepich
Max Planck Institut für Astronomie, Königstuhl 17, 69121 Heidelberg, Germany

Motivation: The intracluster medium (ICM) holds signatures of the dynamical history of the galaxy cluster, including the dark matter density profile, mergers with other clusters, and energetic activity (from supernovae and supermassive black holes) in its member galaxies. For all but the most relaxed galaxy clusters observed at high spatial resolution by instruments such as the *Chandra* and *XMM-Newton* X-ray telescopes, it is extremely challenging to infer such properties as the mass and baryon fraction from the ICM emission. Reproducing these features is a key test of the realism of a given cosmological simulation.

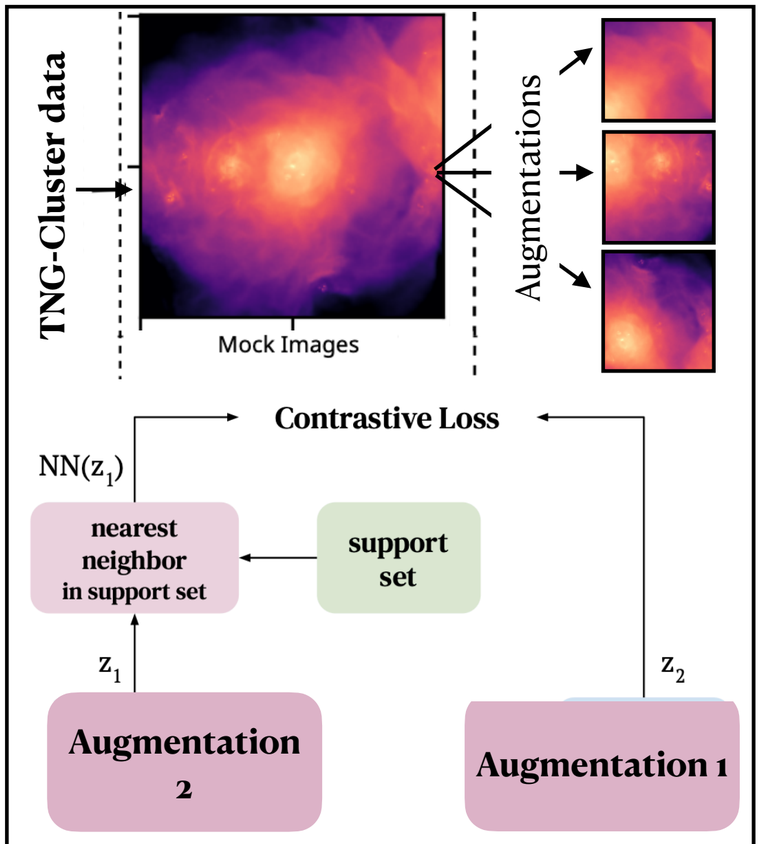

Fig 1. The NNCLR training pipeline, modified from [2].

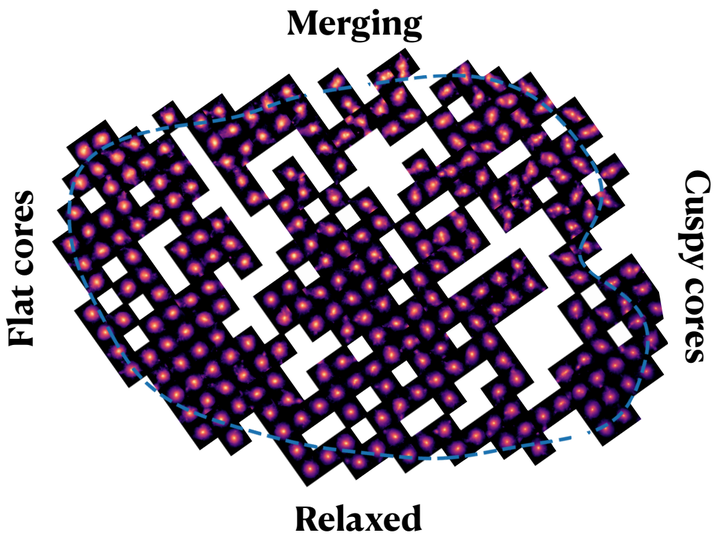

Fig 2. Images of galaxy clusters from the test set are passed through the trained ResNet model to produce a represenation. Next, we generate a 2-D UMAP [cite] of the represenation, and compute distances between all images in the UMAP space. Here, we show the test images sorted by distance in the UMAP.

Method: TNG-Cluster [1] is a suite of cosmological zoom-in simulations of 352 clusters with $M_{200c} > 2 \times 10^{14} M_{\odot}$ at z=0. In this paper, we used Nearest Neighbour Contrastive Learning (NNCLR) [2] to reduce images of the X-ray emission of the clusters in TNG-Cluster to a compact representation space. Finally, we reduce the representation to a UMAP [3].

Future work & applications:

- Finding analogues to observed systems, like recent mergers and AGN-inflated cavities, which are hard to identify algorithmically.
- Infer merger history and sub-cluster masses of dynamically disturbed systems
- Infer AGN feedback history from images of cluster cores (with or without apparent bubbles)
- Study the representation space to find relationships between physical processes

Result: We find that the self-supervised representation space forms a continuous distribution from cool to non-cool core clusters, as well as from relaxed to merging objects. It also shows trends in redshift, halo mass, stellar mass, time since last major merger, and offset between the peaks of mass and X-ray emission. The self-supervised sorting of the images clusters known populations of galaxy clusters, such as Bullet-like mergers, cool- and non-cool cores, and Perseus-like Bubbles, providing simulated analogues to famous observed objects like the Bullet cluster, Perseus. Furthermore, we find the trends persist if we add surface brightness clipping and blurring augmentations, allowing us to compare observed images with different exposure times and spatial resolutions.

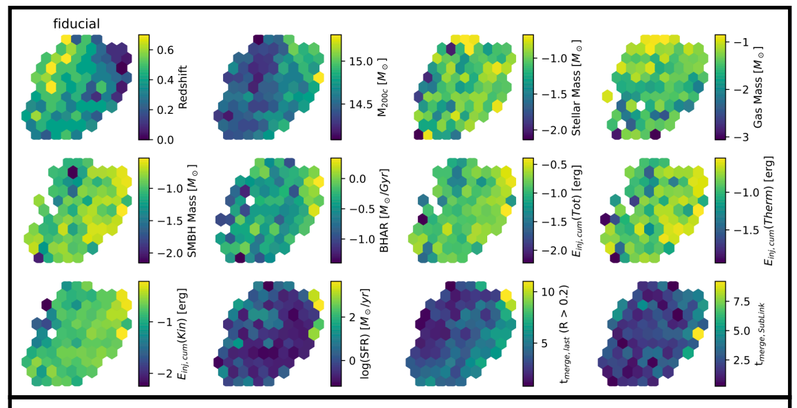

Fig 3. Median values of galaxy cluster properties relating to mass, star formation rate, AGN feedback and merger history in a grid in the UMAP space. Mass and BH properties are shown as deviations from the median in a narrow bin of halo mass.

[1] Nelson et al (2023)
[2] Dwibedi et al (2021) **arxiv:2104.14548**
[3] McInnes et al (2018). **arxiv:1802.03426**

Contact: chadayammuri@mpia.de