Contribution ID: **22** Type: **Computing**

Self-supervised learning of radio data for source detection, classification and peculiar object searches.

Tuesday 15 October 2024 11:30 (20 minutes)

New advancements in radio data post-processing are underway within the SKA precursor community, aiming to facilitate the extraction of scientific results from survey images through a semi-automated approach. Several of these developments leverage deep learning (DL) methodologies for diverse tasks, including source detection, object or morphology classification, and anomaly detection. Despite substantial progress, the full potential of these methods often remains untapped due to challenges associated with training large supervised models, particularly in the presence of small and class-unbalanced labeled datasets.
Self-supervised learning has recently established itself as a powerful methodology to deal with some of the aforementioned challenges, by directly learning a lower-dimensional representation from large samples of unlabeled data. The resulting model and data representation can then be used for data inspection and various downstream tasks if a small subset of labeled data is available.
In this study, we explored contrastive learning methods to learn suitable radio data representation from unlabeled images taken from the ASKAP EMU and MeerKAT GPS surveys. We evaluated trained models and the obtained data representation over smaller labeled datasets, also taken from different radio surveys, in selected analysis tasks: source detection and classification, and search for objects with peculiar morphologies. An overview of the achieved results will be presented at the workshop, along with updates on ongoing activities focusing on pre-training dataset curation and exploitation of ViT and multi-modal ViT+LLM architectures.

Primary authors: Dr RIGGI, Simone (Istituto Nazionale di Astrofisica (INAF)); CECCONELLO, Thomas (Istituto Nazionale di Astrofisica (INAF))

Presenter: Dr RIGGI, Simone (Istituto Nazionale di Astrofisica (INAF))

Session Classification: Session 5