Contribution ID: **53**   Type: **Contributed**

Panopticon: a Machine Learning model to detect transits in raw PLATO light curves

Tuesday, 14 May 2024 15:20 (20 minutes)

One of the biggest hurdle in the transit field has been the initial detection of planetary signals within light curves. While this difficulty is multifactorial, it mostly comes from the stellar activity of the host and the shallow nature of exoplanet transits. The stellar activity varies over a range of timescales, with effects both shorter and longer than the transits themselves. This makes removing the activity a complex task that might deform, or even erase, the transits. Detecting the transit themselves –provided they survived the filtering process–remains a challenge, usually relying on methods requiring periodicity. In the context of the future PLATO mission (Rauer et al., 2014), set to launch in 2026, we aim to propose a method able to search for transits directly in the raw light curves. As one of PLATO's stated goal is the detection of Earth analogs, whose transits are both shallow and of long period, an efficient detection method is desirable. We propose a novel Machine Learning model, Panopticon, based on modified Unet architectures (Ron- neberger et al., 2015 Zhou et al., 2018 Huang et al., 2020) to work with 1-dimensional data. This approach extracts features over various timescales, and recombines them using the larger ones to provide context for the smaller ones. This is well suited to analyze the nature of signals evolving over a variety of timescales, as is the case for light curves with stellar activity. The model outputs a likelihood map of the transits, without having to rely on a prior on the number of transits in the input signal. To train and evaluate the model, we generate a dataset of more than 15 000 light curves using the PlatoSim code and its Platonium associated suite (Jannsen et al., 2024), the mission's state of the art simulator. It takes into account the stellar activity and the platform noises, and injects them at pixel level. This dataset is built to be representative of a wide range of planetary parameters and host stars, to be able to benchmark the model over many different scenarios. As it stands, we are able to consistently retrieve planets with transit depth greater than ~ 150 ppm, reaching ~ 100% above 200 ppm. The fraction of false positives (False Alarm Rate FAR) associated with these detections is of ~ 1%. By modifying the threshold on the confidence score of the model for detections, the FAR can be brought down to ~ 0.1% while still reliably detect transits above 200 pmm. Finally, we note that the orbital period of the planet has little to no impact on the retrieval rate. The model already yields promising results, being able to consistently identify transits, independently of periodicity, with a small False Alarm Rate. To enhance the capabilities of the model at the shallower depth regime, we are working on a Physics Informed version of the model making use of the available host star information. In the future we also aim to take into account the centroids curves to provide an early classification of the nature of the signal (planetary, binary or background event). Additionally, we are working on versions of the model for the TESS and Kepler missions to assess its performance on real data.

Primary author: VIVIEN, Hugo (Laboratoire d'Astrophysique de Marseille)

Presenter: VIVIEN, Hugo (Laboratoire d'Astrophysique de Marseille)

Session Classification: From Stars to Planets