A new machine learning architecture to detect transiting Earth-analogs

Hugo G. Vivien[1] and Magali Deleuil[1]

[1] *Aix Marseille Univ, CNRS, CNES, LAM, Institut Origines, Marseille, France*

April 6, 2023

Current methods to identify and classify exoplanet transits are heavily reliant on human input to both detrend and analyze the light curves. In the context of the future space mission *PLATO* , I developed a new detection algorithm based on Machine Learning (ML).

One important advantage of ML is to not rely on any empirical noise model as classical data reduction algorithms. Compared to previous implementations of Convolution Neural Network such as ExoNet (Shallue and Vanderburg, 2018), AstroNet (Ansdell et al., 2018), and PlaNet (Malik et al., 2022), our architecture is built to perform detection of single transits in high precision light curves. The objective of PLATO being the discovery of exo-Earths, our algorithm indeed aims at detecting and classifying small transits of long period planets. To that purpose, I adapted a Unet++ architecture (Ronneberger et al., 2015; Zhou et al., 2018), which is expected to perform better than a standard Convolutional Neural Network. The ability of the model to extract a wide variety of features at different scales, as well as making use of multi-modal data, make it particularly suitable to detect small and single transit events.

We used the *PLATO* simulator to generate a large set of simulated light curves, used to train and evaluate the network. These light curves include a variety of astrophysical signals: planets (over a large range of orbital periods), eclipsing binaries but also background eclipsing binaries to asses the network classification capacity to identify false positives. The stellar signal itself also includes granulation, pulsations and spots. In its current state, the model takes as input the non-detrended light curve. It is able to extract individual transit events in the noisy data, including super-Earth. It has achieved a global average precision score of $\sim 80\%$, despite the currently limited dataset (~ 5000 light curves).

In the forthcoming months, the dataset will be extended to tackle the class imbalance problem of finding transit events within long duration light curves and improve the detection performance. The network will also be improved to include the centroids information to enable better classification between transits, binaries and background binaries.

References

M. Ansdell, Y. Ioannou, H. P. Osborn, M. Sasdelli, 2018 NASA Frontier Development Lab Exoplanet Team, J. C. Smith, D. Caldwell, J. M. Jenkins, C. Räissi, D. Angerhausen, and . NASA Frontier Development Lab Exoplanet Mentors. Scientific Domain Knowledge Improves Exoplanet Transit Classification with Deep Learning. , 869(1):L7, Dec. 2018. doi: 10.3847/2041-8213/aaf23b.

S. A. Malik, N. L. Eisner, C. J. Lintott, and Y. Gal. Discovering Long-period Exoplanets using Deep Learning with Citizen Science Labels. *arXiv e-prints*, art. arXiv:2211.06903, Nov. 2022. doi: 10.48550/arXiv.2211.06903.

O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv e-prints*, art. arXiv:1505.04597, May 2015. doi: 10.48550/arXiv.1505.04597.

C. J. Shallue and A. Vanderburg. Identifying Exoplanets with Deep Learning: A Five-planet Resonant Chain around Kepler-80 and an Eighth Planet around Kepler-90. , 155(2):94, Feb. 2018. doi: 10.3847/1538-3881/aa9e09.

Z. Zhou, M. Mahfuzur Rahman Siddiquee, N. Tajbakhsh, and J. Liang. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. *arXiv e-prints*, art. arXiv:1807.10165, July 2018. doi: 10.48550/arXiv.1807.10165.