Computer vision as a powerful tool for transiting exoplanet search

Peter Klagyivik
DLR, Berlin (Germany)
(peter.klagyivik@dlr.de)

1. Introduction

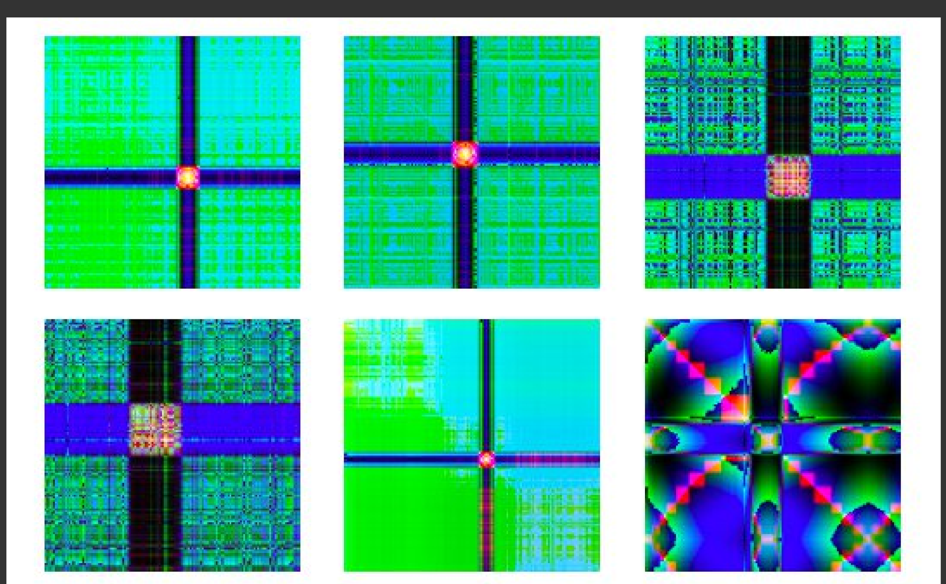
Computer vision is becoming more and more prevalent in everyday life. It is used in self-driving cars, security systems (e.g. face recognition) and during medical examinations as well to identify various diseases on medical images. Hence, advanced methods are already available and can be easily adapted to different problems.

Here we present a pilot study to use computer vision for planetary transit search in time-series data.

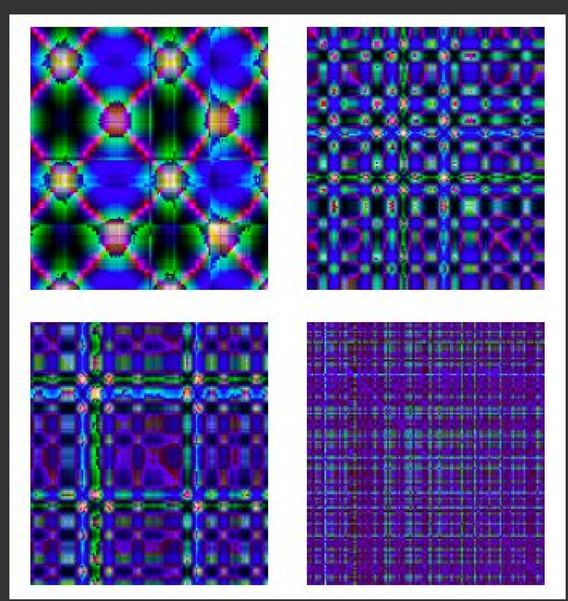
2. Encoding time-series using Markov Transition Fields

To be able to use computer vision on time series, the data should be encoded into 2D images in some way. There are several methods to do this, e.g. the Markov Transition Field (MTF) or the Gramian Angular Field. Here we used MTF, which is widely used for outlier detection in sensor data.

To calculate MTF the time series has to be discretized. Quantile is the most common way to achieve this. Given a time series $X = x_1, x_2, \ldots, x_n$, we quantize its values in Q bins in a way that each bin contains the same number of data points. After assigning each data point to a bin, the MTF tells us the transition probability from one bin into another bin and in this way it describes the flux variation of our light curves and encodes it in a 2D image.



Planetary transits encoded into MTF

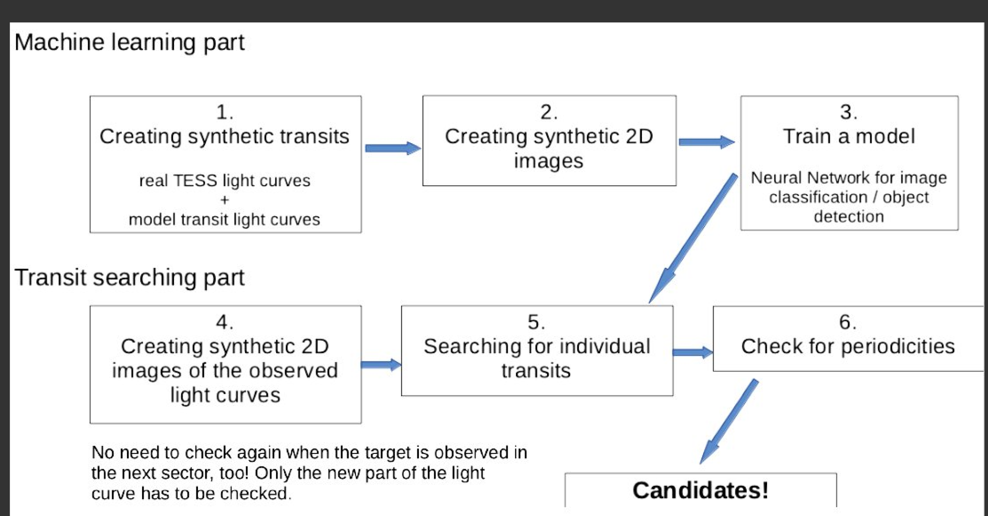


Light curves without transits encoded into MTF

3. Convolutional Neural Network for image classification

We used a simple convolutional neural network with two convolutional layers for the feature recognition and a fully connected layer for the final classification.

The bottleneck of most ML training is the lack of input data. In our case, a planetary transit is a simple geometrical effect that can easily be reproduced. Therefore the number of input time-series is practically unlimited. We used the original TESS observations combined with synthetic transit light curves.



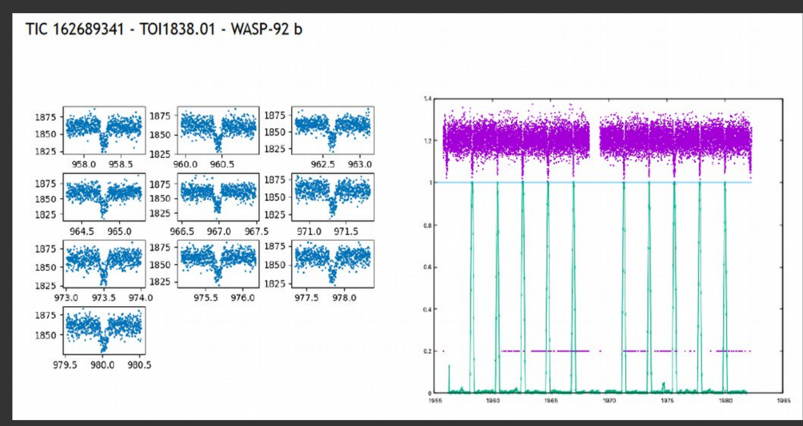
Steps of the planet search.

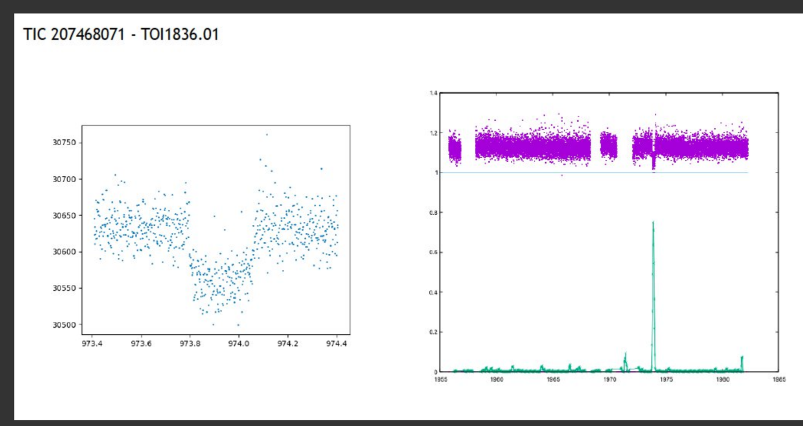
References

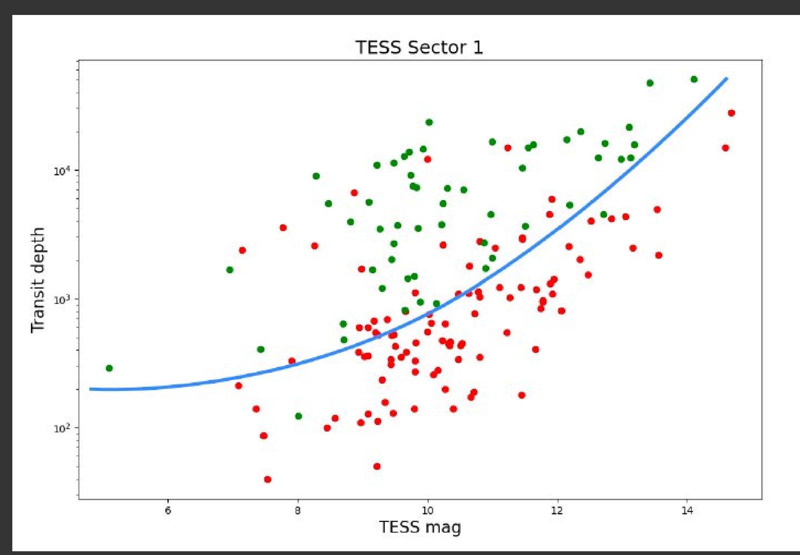
Kostov, V. B. et al., AJ, 159, 6 (2020)

4. Application to TESS Sector 1

We compared our list of candidates against the official TOI Catalog. In the following figure we plot the TESS magnitude versus the transit depth of all TOIs observed in Sector 1. The green dots are detected with our computer vision method, while the red dots remained non-detected. These two groups are clearly separated. Most of the red points above the detection limit are long period planets which were observed but have no transit in Sector 1.





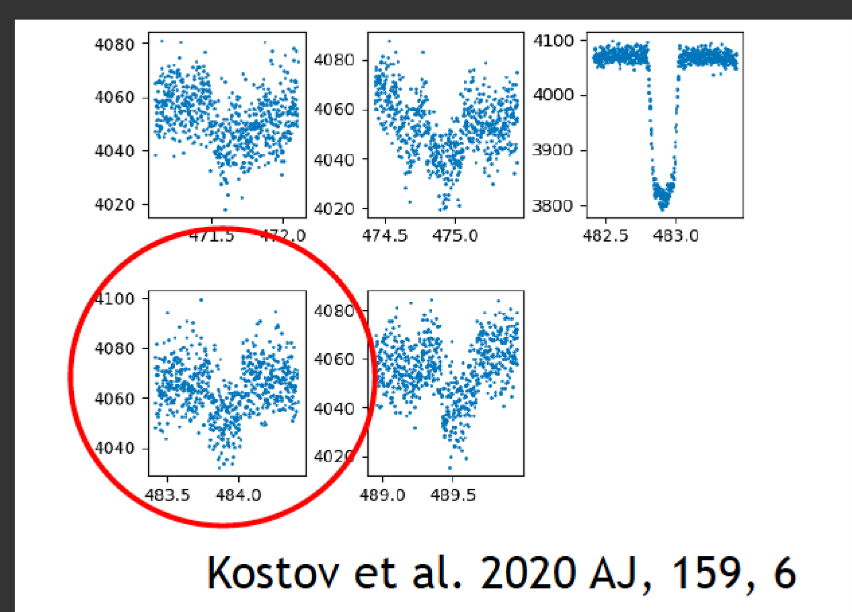


TOIs observed in Sector 1. Green dots represent the detected TOIs, while targets marked with red dots are undetected.

5. TOI-1338 - the first circumbinary planet detected by TESS

One major advantage of the individual transit search is that it can detect transiting circumbinary planets (hereafter CBP). The detection of transiting CBPs is different to that of planets orbiting single stars. The configurations when transits occur depend on both the binary and the planet's orbital phases. Therefore, the transits occur only quasi-periodically, within a 'transit window'. Hence traditional methods cannot find these planets.

Kostov et al. (2020) reported the first CBP in the TESS data. We tested our method on this target and the code easily finds the transit signal along with the stellar eclipses.



The first transiting circumbinary planet detected in the TESS data (Kostov et al. 2020). The sub-figures are the transit-like features in the light curve detected by our method. The red circle shows the transit of the planet.

6. Conclusion

Computer vision can be successfully applied to time-series data as well.

A huge advantage of our method over the traditional planet searching methods is that it can easily find monotransits, and therefore circumbinary planets can be detected. Another advantage is that the computation time increases linearly with additional data.

The disadvantage of the method is its sensitivity. In the traditional methods, where the light curve is folded with a test period, the signal-to-noise ratio is higher due to the overlapping transits and hence more data points during the planetary transit.

However, Earth-like planets around Solar-like stars in the habitable zone will have transit only once per year, hence only 2-3 times during a planned long observing run of PLATO. Therefore the S/N ratio of the folded light curve will not be much better as of individual transits.