TRANSCENDENCE

A Transit Capture Engine for Detection and Neural Network Characterization of Exoplanets

Hendrik Schmerling & Rok Hribar,

Sascha Grziwa, Martin Pätzold

April 30, 2025

Transit Detection

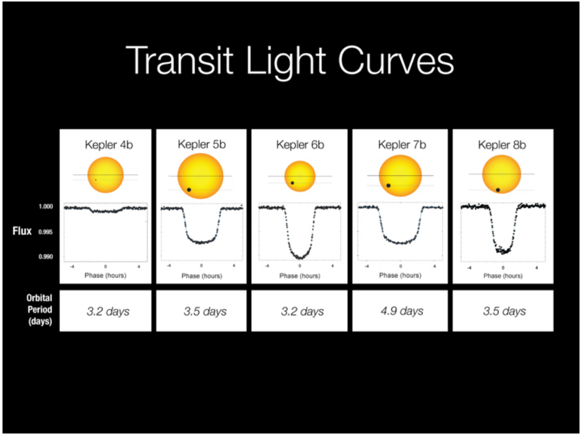

Figure: Examples of known transit LCs with different characteristics

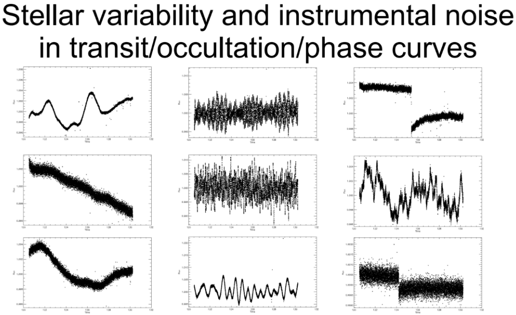

Figure: Different kinds of noise seen in stellar brightness LCs

Classical Transit Detection

- Typically, transit detection makes use of phase folding together with Box Least Square (BLS) algorithms
- Works well for large, short period planets but detection gets worse for smaller planets and longer periods
- Needs repeated transits for successful detections
- Computation typically scales roughly with $O(N^2)$ making long light curves especially challenging

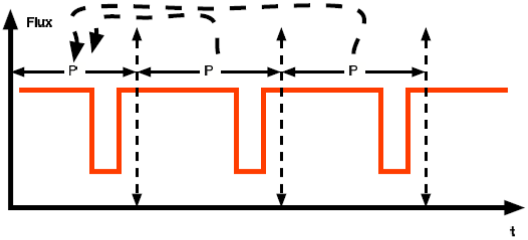

Figure: Concept of phase folding and BLS
Source: Kovács et al (2002)

- Self-learning algorithms trained on labeled datasets
- Extract key features from input data
- Optimize detection accuracy through training

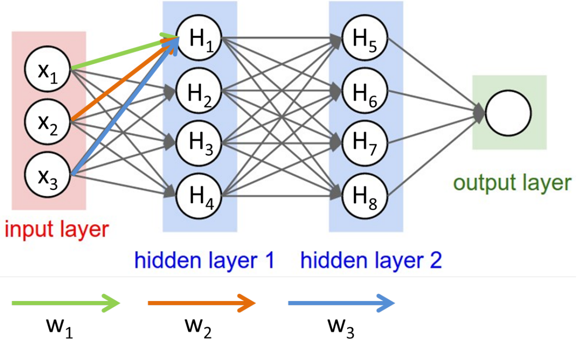

Figure: Concept of Convolutional Neural Networks (CNN)
source: https://medium.com/@yuyunchang/convolutional-neural-networks-cnn-a-dummy-

Data Sources

- Full LC files from TESS space telescope data products

  https://archive.stsci.edu/tess/bulk_downloads/bulk_downloads_ffi-tp-lc-dv.html
- Catalog data stored in the used TESS files and from the gaia catalog

  https://gea.esac.esa.int/archive/
- Synthetic transits generated with the Python package for fast calculation of exoplanet transit LCs called "batman"

  https://lkreidberg.github.io/batman/docs/html/index.html

Mathematical Description of Transit Events

- The Flux at the detector can be expressed as
$$F(a, R_\mathrm{P}) = 1 - \lambda(a, R_\mathrm{P})$$
- where $F(a, R_\mathrm{P})$ is the flux that reaches the telescope and $\lambda(a, R_\mathrm{P})$ is the portion of the flux that is blocked by the planet
- λ itself depends on several parameters. ΔF is proportional to $R_\mathrm{P}/R_\star$ and the transit time T_f depends on the orbital period P_orb

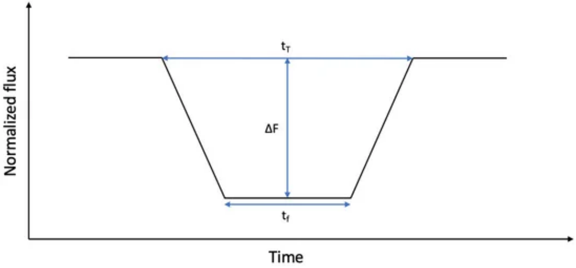

Figure: Theoretical transit model without limb darkening.
Source: https://doi.org/10.3390/axioms13020083

Mathematical Description of Transit Events

- The shapes of transits can be mathematically described in various complexities.
- The more complex ones take into account more complex geometry and stellar limb darkening (LD)
- Box $\rightarrow$ Trapezoid $\rightarrow$ linear LD $\rightarrow$ quadratic $\rightarrow$ more

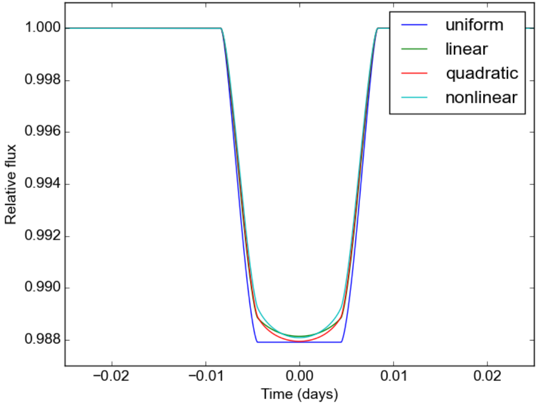

Figure: Different mathematical representations of transits from batman (Kreidberg et al. (2015))

Data Generation – LC - Transit Merger

- Stellar parameters are drawn from TIC or `gaia` and planetary parameters are selected randomly
- Combining full lightcurves from the TESS archive with synthetic transits generated with batman for half of the lightcurves to create a balanced dataset
- We used the quadratic limb darkening formula and assigned the limb darkening coefficients from Claret et al (1990)
- Finally $\sim 450\,000$ (30 TESS sectors) comprise the whole dataset

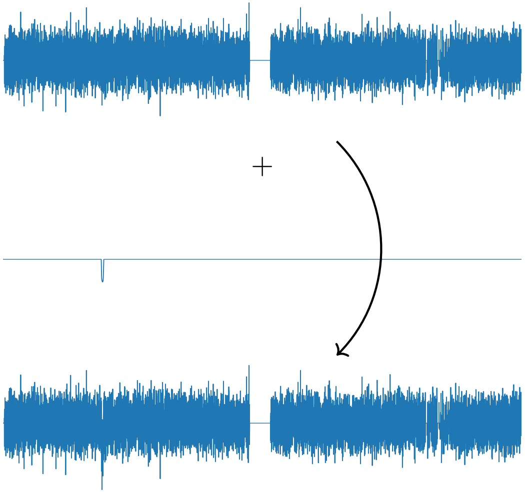

Figure: Insertion of artificial transits to TESS LCs.

Data Generation – Statistics

Parameter	Parameter limits
Orbital period P_{orb} (d)	$\mathcal{I}[0.5, 60]$
Planet radius R_{P} ($\mathrm{R}_{\oplus}$)	$\mathcal{I}[0.5, 50]$
Radius ratio $R_{\mathrm{P}}/R_{\star}$	$(0, 0.1]$
Eccentricity e	$\lvert\mathcal{G}[0, 0.05]\rvert$
Inclination i (deg)	$\mathcal{U}[i_{\mathrm{min}}, 90]$

Table: Parameter priors for data generation
$\mathcal{U}[a, b]$ stands for uniform priors in between a and b, $\mathcal{G}[a, b]$ marks Gaussian priors with a as the mean and b as the standard deviation, $\mathcal{I}$ stands for individually assigned limits

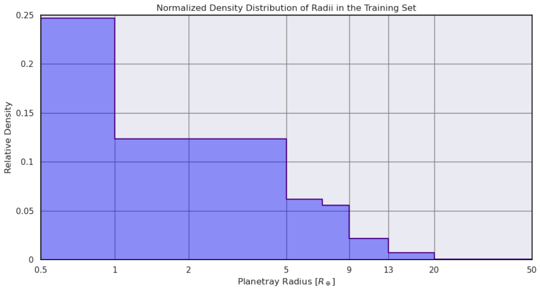

Figure: A priori distribution of planet in the generated dataset

Model Performance - Detection

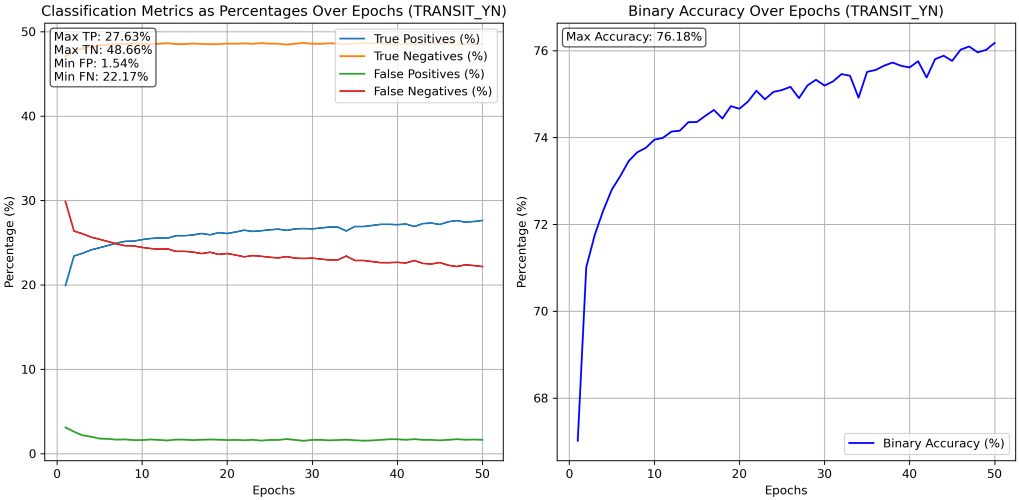

Figure: Binary classification performance dependent on training time

Model Performance - Classification

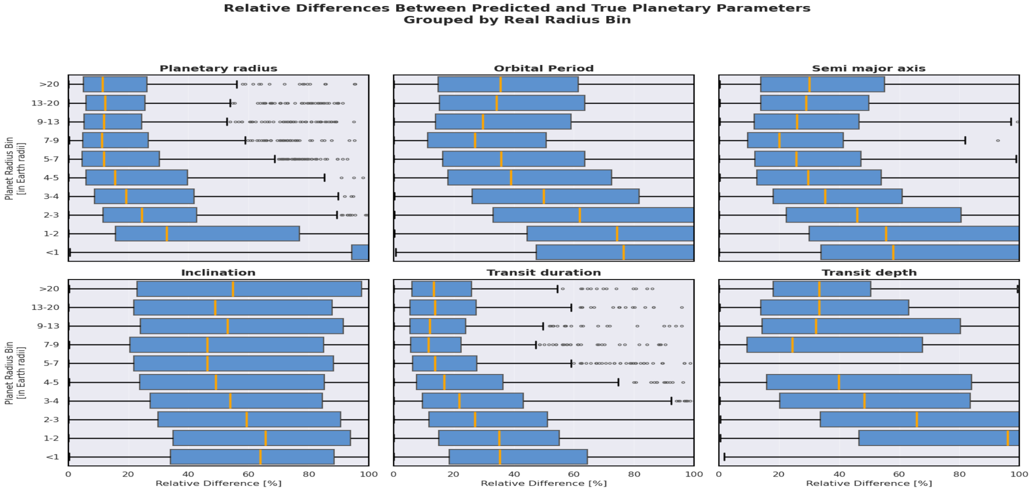

Figure: The classification accuracy for various planetary and orbital parameters

Classification with Logistic Regression

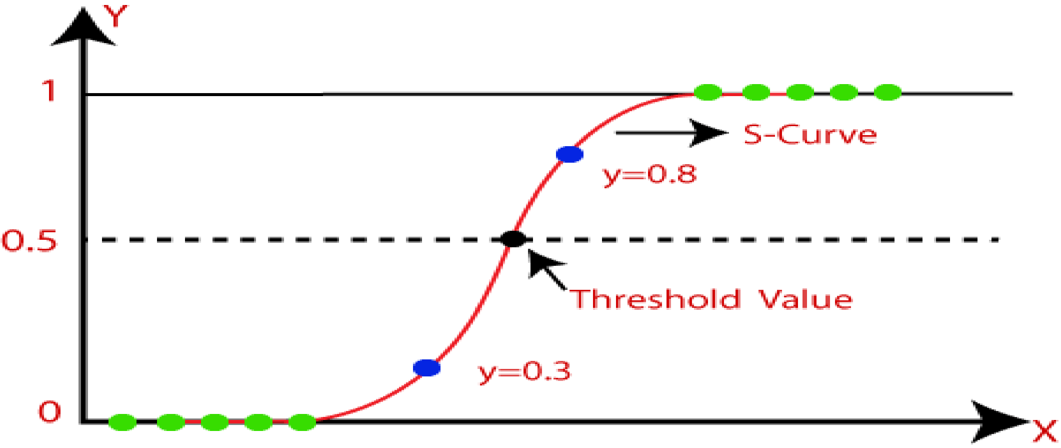

Figure: Classification with a logistic regression

Assigned Number of Logistic Regression Compared to "Best" Classification

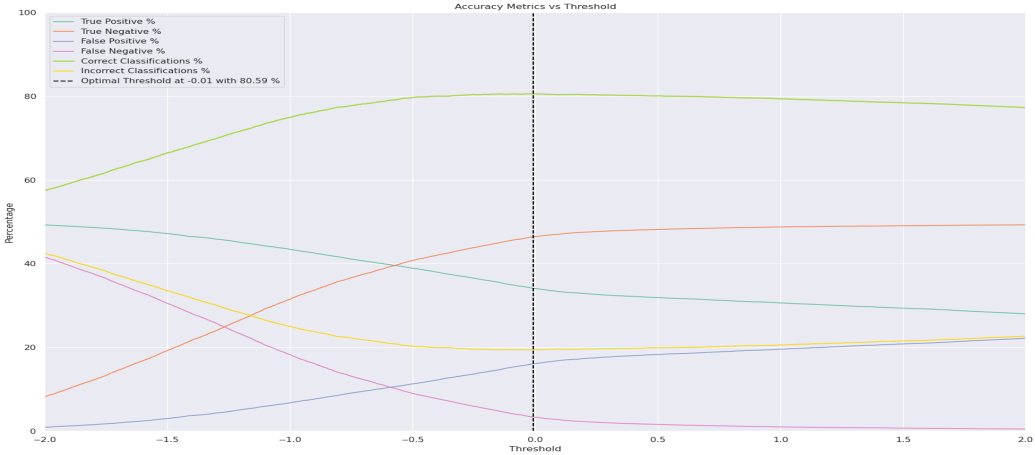

Figure: Logistic regression output

Problem with Monotransits

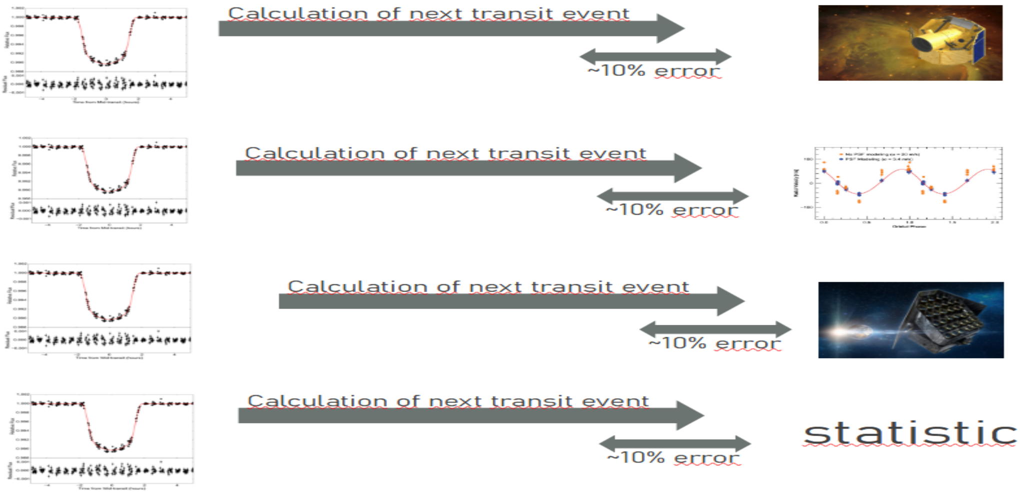

Figure: Problem with Monotransits