LIGHT CURVE IMAGING FOR EXOPLANET DETECTION WITH DEEP LEARNING: A CONCEPTUAL TRIAL

Silva Filho, P. C. F.[1], Reis, S. D. S.[2], de Freitas, D. B.[3]

1: clebersilva@fisica.ufc.br, 2: saulo@fisica.ufc.br, 3: danielbrito@fisica.ufc.br

STELLAR TEAM

Introduction

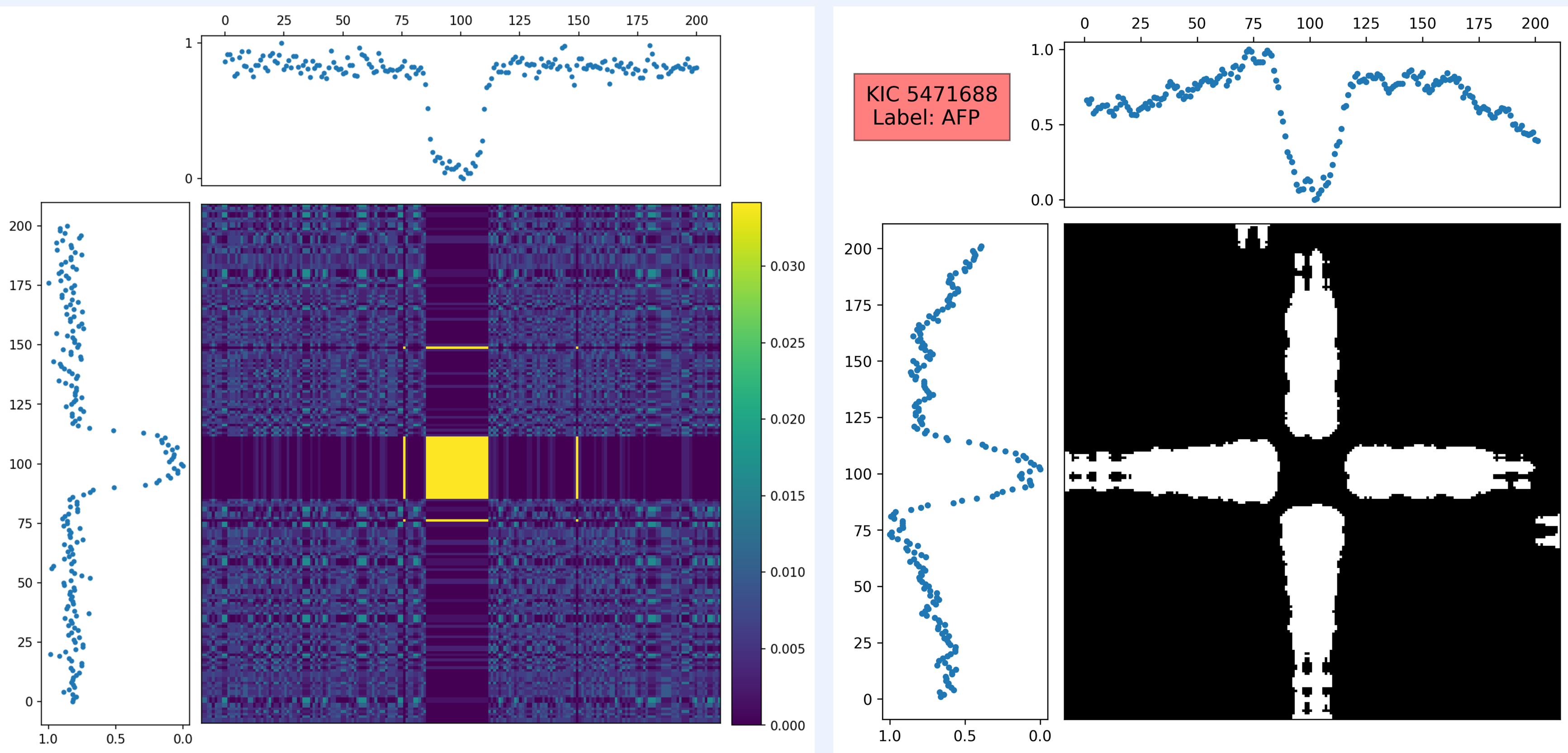
The ever-growing number of space missions has made manual searching for exoplanet candidates infeasible due to the increasing volume of data. For instance, the forthcoming PLATO space telescope is anticipated to yield an unprecedented amount of data. Consequently, the astrophysics community has extensively employed machine learning methods not only to handle the sheer amount of available data but also to enhance the sensitivity of detections concerning the signal noise inherent in relevant observational cases. This work presents a conceptual trial with an alternative method to those previously discussed in the literature for classifying potential exoplanet signals using deep learning.

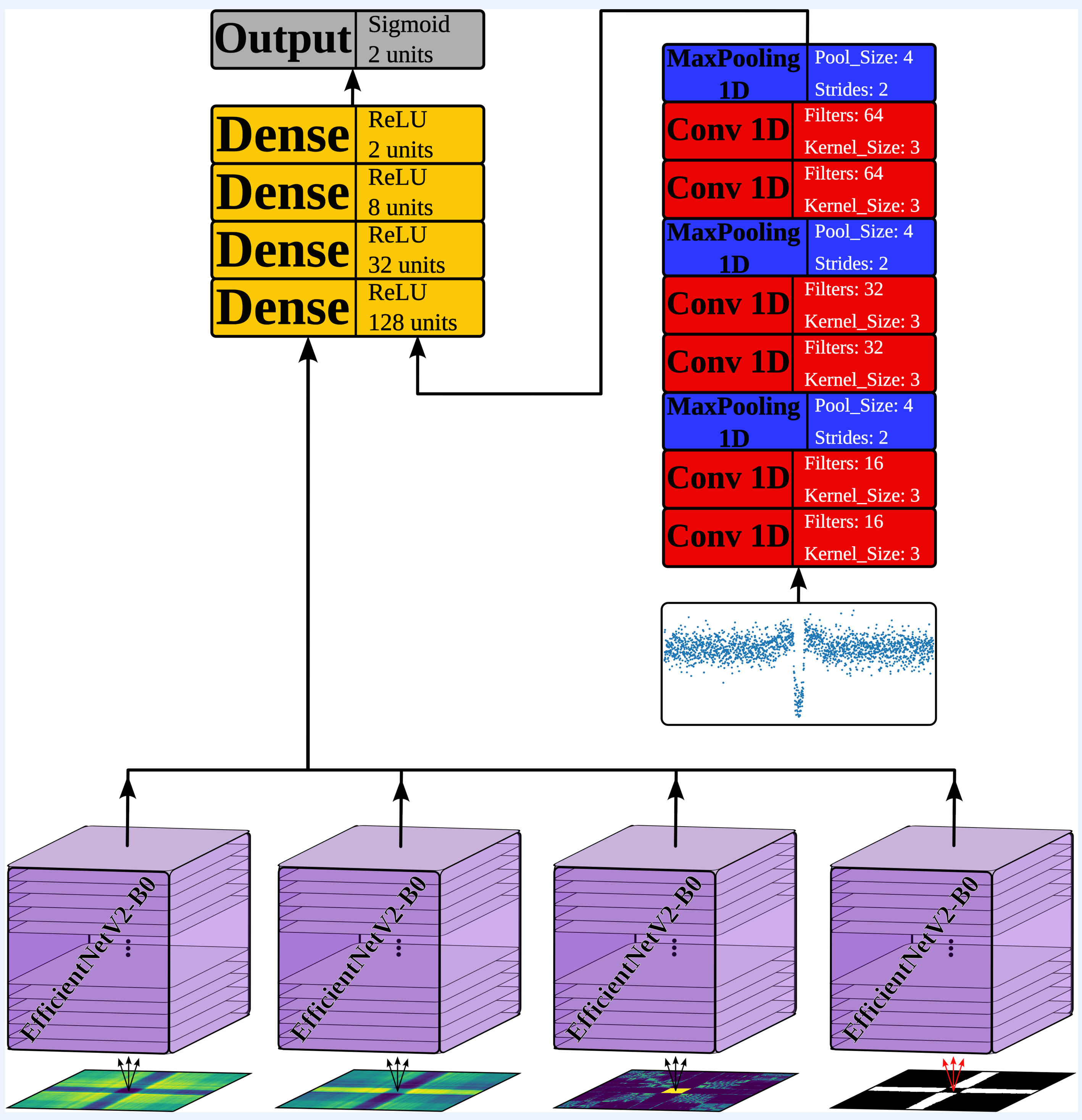
Methods

We developed, trained, and evaluated Convolutional Neural Network (CNN) models to analyze light curves from the Kepler space mission, allowing inference on whether a given signal refers to an exoplanet or not. The distinction of this work lies in the imaging of these light curves before they are passed to the CNNs, which practically increases the number of dimensions available for analysis and enables the use of powerful and successful computer vision techniques for classification problems. To do so, we generated images from the treated light curves as Gramian Summation and Difference Angular Fields (GSAF and GDAF), Markov Transition Fields (MTF) and Recurrence Plots (RP).

$$\text{GSAF} = \left[\cos\left(\phi_i + \phi_j\right)\right] = \tilde{F}' \cdot \tilde{F} - \left(\sqrt{1 - \tilde{F}^2}\right)' \cdot \sqrt{1 - \tilde{F}^2}$$

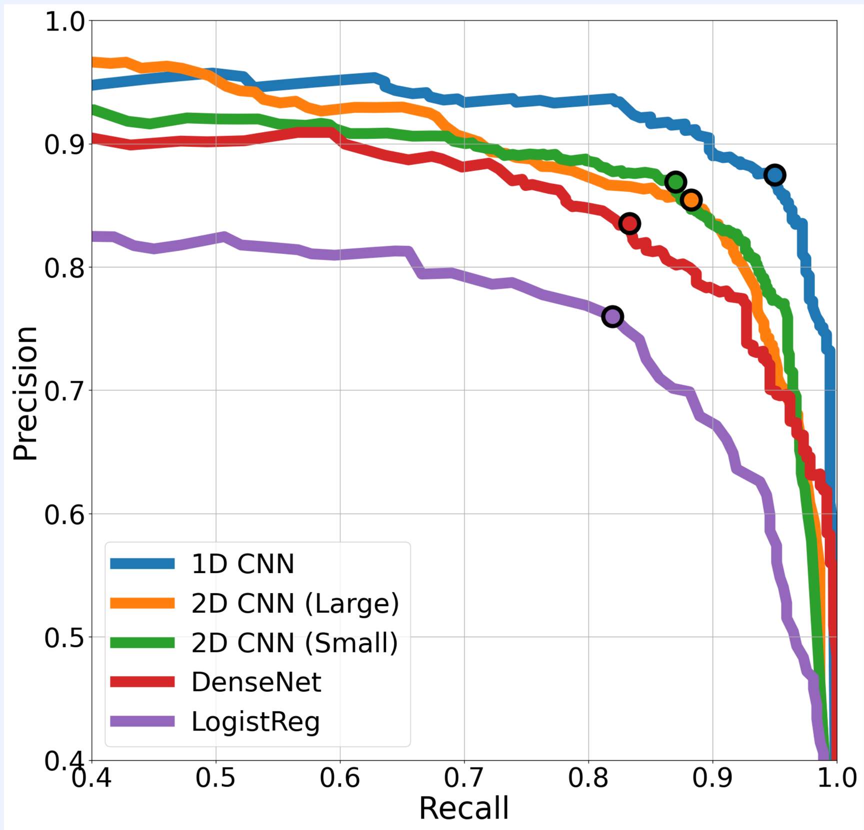
$$\text{GDAF} = \left[\text{sen}\left(\phi_i - \phi_j\right)\right] = \left(\sqrt{1 - \tilde{F}^2}\right)' \cdot \tilde{F} - \tilde{F}' \cdot \sqrt{1 - \tilde{F}^2}$$

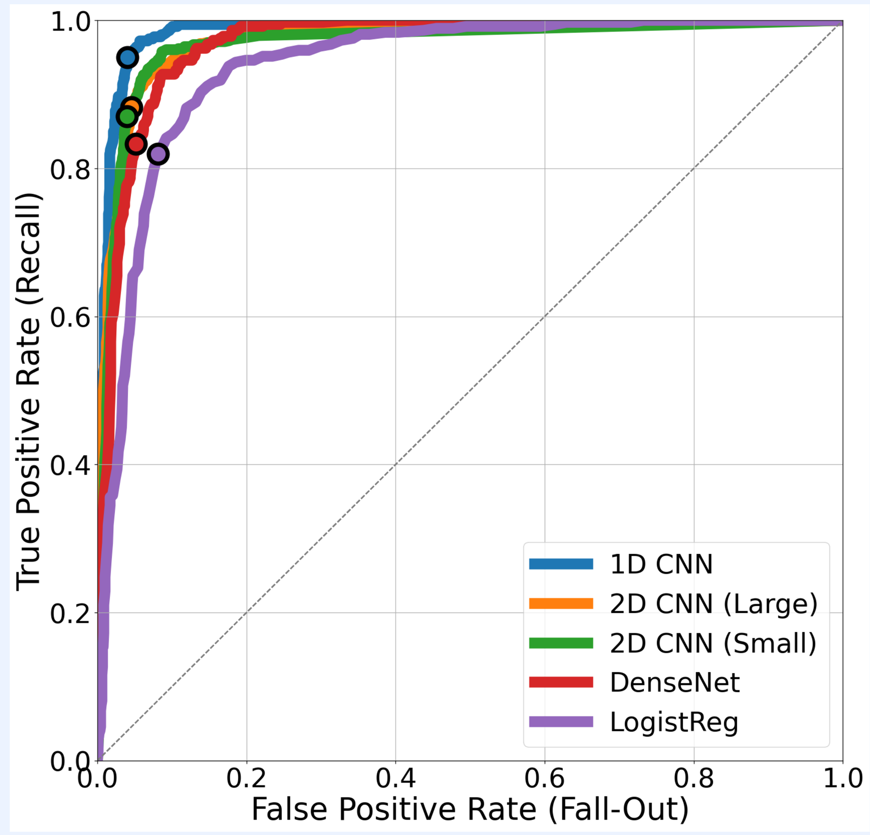




Results

Our best model ranks plausible planet signals higher than false-positive signals 97.22% of the time in our test dataset and demonstrates promising performance on entirely new data from other datasets. Our best model also shows a moderate capacity to generalize what it learned with data from other space missions, such as K2 and TESS. We compared our results with the algorithm developed by Shallue & Vanderburg (2018) [4] and represented as the 1D CNN below.





Results With Kepler Test Dataset

Model	PR-AUC	ROC-AUC	Accuracy	Precision	Recall
1D CNN	93.55%	98.60%	95.74%	87.47%	95.00%
2D CNN (Large)	91.70%	97.22%	93.78%	85.45%	88.25%
2D CNN (Small)	88.81%	97.04%	93.93%	86.90%	87.02%
DenseNet	87.98%	96.93%	92.07%	83.51%	83.29%
LogistReg	79.38%	93.96%	89.50%	76.00%	81.94%

Results With Complementary K2 and TESS Dataset

Model	Accuracy (%)	Precision (%)	Recall (%)
1D CNN	63.60	35.44	38.89
2D CNN (small)	66.28	40.00	44.44
2D CNN (large)	70.50	45.90	38.89

Conclusion

A good performance on entirely new data is a critical characteristic for algorithms to be applied on data from upcoming space missions such as PLATO, and is still a work in progress at time of writing. In order to achieve the bestest performance, the data treatment process should have the best quality possible and it depends on the telescope specific characteristics. The time series imaging technique was proposed in the deep learning and computer vision context. Both our proposed 2D models scored almost 94% accuracy in the test dataset, even with no fine-tuned hyperpa-rameters. A complementary test was made using data from K2 and TESS missions and both proposed models performed better than the reference model, which indicate that they have promising potential to generalize their predictions with new data, such as the ones we will receive from PLATO.

References

[1] Wang, Z., & Oates, T. (2015). Imaging time-series to improve classification and imputation.

[2] Eckmann, J.-P.; Kamphorst, S.O.; Ruelle, D. Recurrence Plots of Dynamical Systems. Europhys. Lett. 1987, 4, 973–977.

[3] Goswami, B. (2019). A brief introduction to nonlinear time series analysis and recurrence plots. Vibration, 2(4), 332-368.

[4] Shallue, C. J., & Vanderburg, A. (2018). Identifying exoplanets with deep learning: A five-planet resonant chain around kepler-80 and an eighth planet around kepler-90. The Astronomical Journal, 155(2), 94.