EPSC Abstracts
Vol. 13, EPSC-DPS2019-842-1, 2019
EPSC-DPS Joint Meeting 2019
© Author(s) 2019. CC Attribution 4.0 license.

EPSC
EPSC-DPS Joint Meeting 2019
15–20 September 2019 | Geneva, Switzerland

DeepStreaks: identifying fast-moving objects in the Zwicky Transient Facility data with deep learning

Dmitry A. Duev (1), Ashish Mahabal (1), and Quanzhi Ye (1,2)
(1) Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA (duev@caltech.edu), (2) IPAC, California Institute of Technology, MS 100-22, Pasadena, CA 91125, USA

Abstract

We present `DeepStreaks`, a convolutional-neural-network, deep-learning system designed to efficiently identify streaking fast-moving near-Earth objects that are detected in the data of the Zwicky Transient Facility (ZTF), a wide-field, time-domain survey using a dedicated 47 deg^2 camera attached to the Samuel Oschin 48-inch Telescope at the Palomar Observatory in California, United States. The system demonstrates a 96-98% true positive rate, depending on the night, while keeping the false positive rate below 1%. The sensitivity of `DeepStreaks` is quantified by the performance on the test data sets as well as using known near-Earth objects observed by ZTF. The system is deployed and adapted for usage within the ZTF Solar-System framework and has significantly reduced human involvement in the streak identification process, from several hours to typically under 10 minutes per day.

1. Introduction

The Zwicky Transient Facility (ZTF) is a new robotic time-domain sky survey that visits the entire visible sky north of $-30°$ declination every three nights in the g and r bands, and at higher cadences in selected sky regions including observations with the i-band filter [1, 2].

The ZTF Science Data System (ZSDS) housed at IPAC consists of the data processing pipelines, data archives, infrastructure for long-term curation, and the services for data retrieval and visualization [3]. A part of the ZSDS, the ZTF Streak pipeline (ZStreak) focuses on the detection of streaked objects [4].

On a typical night, the number of detected "raw" streaks reaches 10^5 – 10^6. The random-forest (RF) classifier initially used in production only reduces this to 10^4 – 10^5 still resulting in several human-hours spent on candidate scanning each day. Furthermore, given the number of streaks that need to be looked at, it is not uncommon for the human scanners to miss several streaks from real FMOs. Typically, only several streaks to a few dozen are marked nightly as plausible real candidates.

The aim of this work is to build a machine-learning (ML) system that has a sensitivity similar to that of the RF classifier, but significantly reduces the number of false positives.

2. DeepStreaks: a deep learning framework for streak identification

The problem of streak identification is split into three simpler sub-problems, each solved by a dedicated group of classifiers assigning real vs. bogus ("rb"), short vs. long ("sl"), and keep vs. ditch ("kd") scores (See Fig. 1). Within each group of classifiers we have chosen to use three different convolutional neural-network models. At least one member of each group must output a score that passes a pre-defined threshold.

With a score threshold of 0.5 that is adopted for all individual classifiers, `DeepStreaks` yields 96-98% true positive rate (TPR) on the test sets while keeping the false positive rate (FPR) below 1%.

To quantify the completeness of `DeepStreaks` identifications, we evaluated it on streak images of known real NEOs detected by the ZTF Streak pipeline from October 2018 – January 2019 (see Fig. 2). Out of 210 such streaks, 202 (96%) were correctly classified. We note that the RF classifier initially used in `ZStreak` with a threshold of 0.05 demonstrates similar TPR on this data set.

Please refer to [5] for the details about `DeepStreaks` architecture, training and test data sets, implementation, and performance.

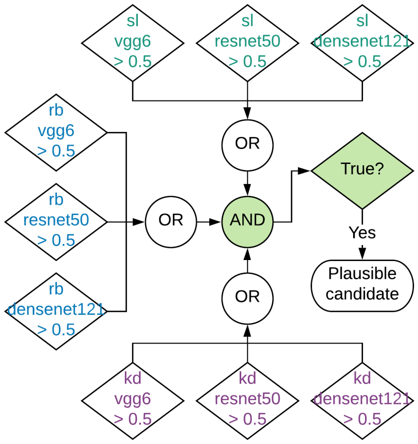

Figure 1: Decision logic used by `DeepStreaks` to identify plausible streaks.

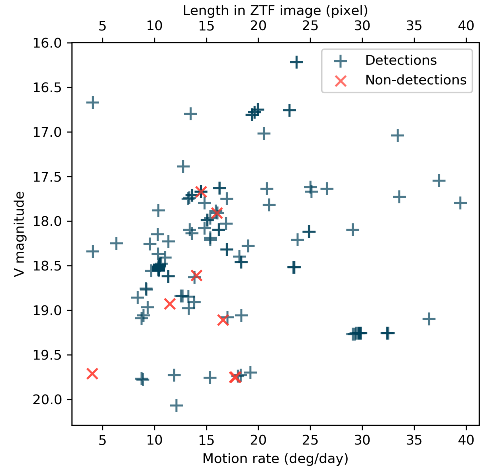

Figure 2: Completeness of `DeepStreaks` identifications using known NEOs observed by ZTF in October 2018 – January 2019.

3. Summary and Conclusions

While providing a similar sensitivity, `DeepStreaks` demonstrates a $50-100\times$ better performance than the original random-forest-based classifier used in the ZStreak pipeline in terms of the false negative rate. This reduces drastically the time humans have to spend scanning for streaks – from hours to typically under 10 minutes per day.

As of May 7, 2019, 30 NEOs have been discovered with `DeepStreaks`, including 2019 BE_5, the fastest-spinning asteroid discovered to date that has a rotational period of 12 seconds (W. Ryan, private comm.)

We have demonstrated that by putting together a few simulations, large amounts of data from ZTF, fast computing, and a few deep learning models we can improve the efficiency of detecting streaking asteroids by a couple orders of magnitude, saving tens of human-hours per week at the same time.

Acknowledgements

D.A. Duev acknowledges support from the Heising-Simons Foundation under Grant No. 12540303. Q.-Z. Ye is supported by the GROWTH project funded by the National Science Foundation under Grant No. 1545949. Based on observations obtained with the Samuel Oschin Telescope 48-inch Telescope at the Palomar Observatory as part of the Zwicky Transient Facility project. Major funding has been provided by the U.S. National Science Foundation under Grant No. AST-1440341 and by the ZTF partner institutions: the California Institute of Technology, the Oskar Klein Centre, the Weizmann Institute of Science, the University of Maryland, the University of Washington, Deutsches Elektronen-Synchrotron, the University of Wisconsin-Milwaukee, and the TANGO Program of the University System of Taiwan. AM acknowledges support from NSF (1640818 and AST-1815034).

References

[1] Bellm E. C. et al., 2019, PASP, 131, 018002

[2] Graham M. J. et al., 2019, PASP, accepted (arXiv:1902.01945)

[3] Masci F. J. et al., 2019, PASP, 131, 018003

[4] Ye Q.-Z. et al., 2019, PASP, accepted (arXiv:1904.09645)

[5] Duev, D. A. et al., 2019, MNRAS, 486, 4158