Tails: Chasing Comets with the Zwicky Transient Facility and Deep Learning

Dmitry A. Duev*
Division of Physics, Mathematics, and Astronomy
California Institute of Technology
Pasadena, CA 91125
duev@caltech.edu

Abstract

We present Tails, an open-source deep-learning framework for the identification and localization of comets in the image data of the Zwicky Transient Facility (ZTF), a robotic time-domain sky survey currently in operation at the Palomar Observatory in California, USA. Tails employs a customized EfficientDet-based architecture and is capable of finding comets in single images in near real time, rather than requiring multiple epochs as with traditional methods. The system achieves state-of-the-art performance with over 99% recall, 0.01% false positive rate, and 1-2 pixel root mean square error in the predicted position. We report the initial results of the Tails efficiency evaluation in a production setting on the data of the ZTF Twilight survey, including the first AI-assisted discovery of a comet (C/2020 T2).

1 Introduction

Comets have mesmerized humans for millennia, frequently offering, arguably, some of the most spectacular sights in the night sky. Containing the original materials from when the Solar system first formed, comets provide a unique insight into the distant past of our Solar system. The recent discovery of the first interstellar comet 2I/Borisov by amateur astronomer Gennadiy Borisov predictably sparked much excitement and enthusiasm among astronomers and the general public alike. Such objects could potentially provide important information on the formation of other stellar systems. It is a very exciting time to look for comets: the large-scale time-domain surveys that are currently in operation, such as the ZTF [1], Pan-STARRS [3], or ATLAS [18], and the upcoming ones such as BlackGEM [2] and Vera Rubin Observatory / LSST [12] offer the richest data sets ever available to mine for comets.

Traditional comet detection algorithms rely on multiple observations of cometary objects that are linked together and used to fit an orbital solution. The previous attempts to take the comet's morphology in the optical image data into consideration in the detection algorithms have not led to reliable and robust results.

In this work, we present Tails - a state-of-the-art open-source deep-learning-based system for the identification and localization of comets in the image data of ZTF[2]. Tails employs an EfficientDet-based architecture [17] and is thus capable of finding comets in single images in near real time, rather than requiring multiple epochs as with traditional methods.

*https://duev.space
[2] https://github.com/dmitryduev/tails

Third Workshop on Machine Learning and the Physical Sciences (NeurIPS 2020), Vancouver, Canada.

1.1 The Zwicky Transient Facility

The Zwicky Transient Facility (ZTF)[3] is a state-of-the-art robotic time-domain sky survey capable of visiting the entire visible sky north of $-30°$ declination every night. ZTF observes the optical sky in the g, r, and i bands[4] at different cadences depending on the scientific program and sky region [1, 9], down to a 5σ detection limit of 20.7 mag in the r band with a 30-second exposure during new moon [5, 15].

ZTF has been running a unique survey, the Twilight Survey (ZTF-TS) that operates at Solar elongations down to 35 degrees with an r-band limiting magnitude of 19.5 [19]. ZTF-TS has so far resulted in the discovery of a number of Atira asteroids as well as the first Vatira object 2020 AV2 (an asteroid inside Venus's orbit) [11].

For obvious reasons, comets become more easily detectable when close to the Sun as they become brighter and start exhibiting more pronounced coma and tails. Furthermore, it has been shown that the most detectable direction of approach of an interstellar object is from directly behind the Sun because that direction has a bigger cross section for asteroids to bend around and pass into the visibility volume [6].

Tails automates the search for comets with detectable morphology. While trained and evaluated on a large corpus of ZTF data, in this work we focus on Tails' performance when applied to the ZTF-TS data.

2 Tails: a deep learning framework for the identification and localization of comets

The art of building applied deep learning systems involves two major challenges: finding a suitable network architecture and, more importantly, constructing a large, labeled, representative data set for the network training. In the case of comet detection, the training set must reflect the possible variations across different seeing conditions, filters, sky location, CCDs, as well as cross-talk.

2.1 Data set

To build a seed sample for labeling, we first identified all potential observations of known comets conducted with ZTF from March 5, 2018 - March 4, 2020, based on their predicted position and brightness. The code for accomplishing that is based on the `python` libraries `pypride` [7] and `solarsyslib` [13] and uses the comet ephemerides obtained from the Minor Planet Center (MPC)[5] for a coarse search, followed by a JPL Horizons[6] query for precision.

To provide more contextual information, epochal image data are supplemented by properly aligned reference images of the corresponding patches of sky and difference (epochal minus reference) images generated with the ZOGY algorithm [20], all produced by the ZTF Science Data System at Caltech's IPAC. Finally, we generate image triplet cutouts of size 256 by 256 pixel, which in angular measure translates into 4.3 by 4.3 arcmin at ZTF's pixel scale of 1.01 arcsec / pix.

We selected over 60,000 individual observations with the total comet magnitude ranging from 10 to 23 (as reported by JPL Horizons), out of which about 20,000 were sourced for manual annotation. This resulted in an initial sample of 3,000 examples with identifiable morphology.

We also compiled a set of approximately 20,000 negative examples consisting of point-like cometary detections, patches of sky with no identified transient or variable sources, CCD-edge cases, and a wide range of real (point-source) transient and bogus (e.g. artifacts due to bright stars, optical ghosts and "dementors") samples from the Braai data set [8].

To expand the data set, we then assembled a simple ResNet-based [10] classifier for comet identification. With this basic classification model, we ran several rounds of an active-learning-like procedure, where we would first train the classifier, evaluate it on the whole data set, sample both confident

[3] `https://ztf.caltech.edu`

[4] Please see [1] for the exact definitions of the filter pass bands.

[5] `https://www.minorplanetcenter.net/iau/MPCORB/CometEls.txt`

[6] `https://ssd.jpl.nasa.gov/horizons.cgi`

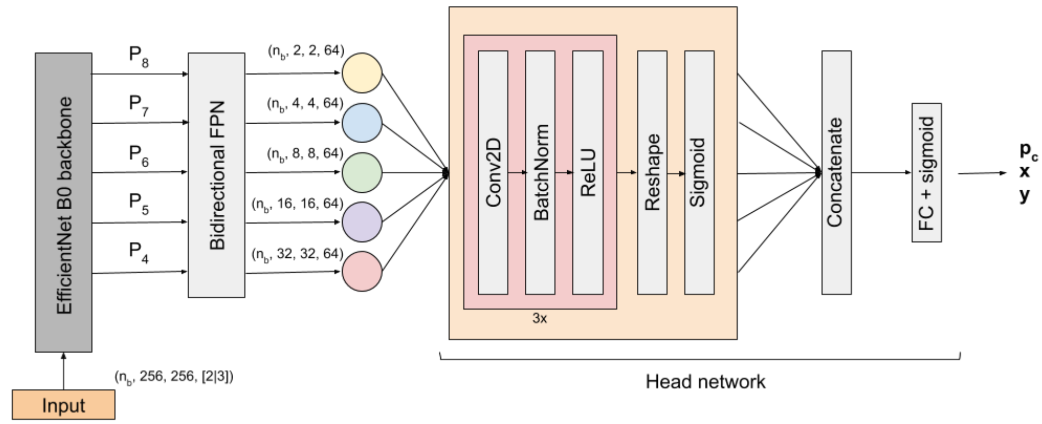


Figure 1: Tails architecture: a custom EfficientDet D0-based network [17]. A batch of duplet or triplet image stacks of size $(n_b, 256, 256, [2|3])$, correspondingly, is passed through an EfficientNet B0 backbone, where (n_b) is the number of stacks in the batch. The extracted features from the last five blocks/levels of the backbone network are passed through a bidirectional feature-pyramid network (BiFPN). The resulting five output tensors denoted in colored circles are fed into the head network, which outputs the probability of the image containing a comet p_c and its predicted relative position (x, y).

predictions and the cases close to the classifier's decision boundary, manually inspect and label those examples and add them to the training set.

The resulting training data set contains about 5,000 positive and 22,000 negative examples. Each triplet in the set has been assigned a label $[p_c, x, y]$, where p_c is the probability of the presence of a comet in the image, $x, y \in [0, 1]$ – the relative positions of the comet's "center of mass", as reported by JPL Horizons. For positive examples, this translates into $[1, x_{JPL}, y_{JPL}]$, for negative ones – $[0, ?, ?]$, where question marks mean that these do not affect the loss in this case.

2.2 Deep neural network architecture and training

Tails adopts a custom architecture (see Fig. 1) based on EfficientDet D0 [17]. A batch of triplet image stacks of size $(n_b, 256, 256, 3)$, where (n_b) is the number of stacks in the batch, is passed through an EfficientNet B0 backbone [16]. The extracted features from the last five blocks/levels of the network are passed through a weighted bidirectional feature-pyramid network (BiFPN) first proposed in [17]. The resulting five output tensors denoted in colored circles are fed into the head network, which outputs the probability of the image containing a comet p_c and its predicted relative (x, y) position.

We defined the loss function as follows:

$$L = w_c \cdot L_c + w_p \cdot L_p \tag{1}$$

where L_c denotes the binary cross-entropy function for the label c (1 – there is a comet in the image, 0 – there is no comet) and the predicted probability p_c. If $\lfloor p_c \rceil = 1$, L_p is computed as an L_1 loss for the relative position (x, y) and its prediction (x_p, y_p) with a small L_2 regularizing term (with $\epsilon = 10^{-3}$). The weights of the two terms are denoted by w_c and w_p, respectively.

We employed the Adam optimizer [14], a batch size of 32, and a 81%/9%/10% training/validation/test data split. For data augmentation, we applied random horizontal and vertical flips of the input data; no random rotations and translations were added. Standard techniques such as learning rate reduction on plateau and early stopping were also used to improve the performance.

The EfficientNet's weights were randomly initialized[7]. We first set $w_c = 10, w_p = 1$ to allow for a fast convergence of the feature-extracting part of the network (i.e. by not paying much attention to the part of the network responsible for the comet localization). To fine-tune the performance, we trained Tails on a balanced data set setting $w_c = 1.1, w_p = 1$ and monitored the validation loss for early stopping, then bumped $w_p = 2$ and monitored the validation positional RMSE; finally, added the omitted negative examples and again monitored the validation loss for early stopping.

The resulting classifiers were put through the same active-learning-like procedure as was employed in the initial data set assembly, using several months of ZTF Twilight survey data.

3 Tails performance

Evaluated on the test set, Tails demonstrates a 99% label prediction accuracy and recall, 0.01% false positive rate, and a $\sim 1-2$ pixel median RMSE of the predicted comet "centroid" position versus the one acquired from JPL Horizon (reference).

The ZTF instrument's CCD mosaic has 16 individual *6k* $\times$ *6k* science CCDs. The raw ZTF image data are split into four readout quadrants per CCD and all processing is conducted independently on each CCD readout quadrant. We tessellate each *3k* $\times$ *3k* CCD-quadrant image into a 13 $\times$ 13 grid of overlapping 256 $\times$ 256 pixels tiles and evaluate Tails on those.[8]

Tails has been deployed in production since June 2020. It takes about 5 hours to run inference on a typical set of nightly ZTF Twilight data (~ 45 30-second exposures) on an *e2-highcpu-32* instance (32 vCPU, 32 GB memory, SSD disk) on the Google Cloud Platform, including I/O operations.

Consistently with the test set performance and the expected rate of comet observation, a typical run yields a few dozen candidates that are manually inspected using a simple Jupyter[9] notebook that automatically checks for known Solar system objects via the MPChecker[10]. The scanning results are accumulated and used to expand the training set and improve Tails' performance.

Fig. 2 shows a number of comets identified and localized by Tails, including some of the ZTF observations of the comet 2I/Borisov.

3.1 Discovery of comet C/2020 T2

On October 7, 2020, Tails discovered a candidate that was posted to MPC's Possible Comet Confirmation Page[11] as ZTFDD01. It was later confirmed to be a long-period comet and designated C/2020 T2 (Palomar), marking the first DL-assisted comet discovery [4]. The candidate was found in the Twilight survey data; it was at 19.3 mag in the r band (see Fig. 2 (a)).

4 Discussion

This work demonstrates the potential of the state-of-the-art deep-learning computer-vision architecture designs when applied to the problem of astronomical source detection and localization. We experimented with the input data and trained a version of Tails that instead of triplet image stacks uses duplets – epochal/reference images, omitting the ZOGY difference images. Our tests show that this version achieves essentially the same performance as the one trained on triplets without requiring image differencing, expanding the range of potential use cases of Tails.

While Tails is trained only on ZTF data, with transfer learning, it can be easily adapted to other sky surveys, including the upcoming Vera Rubin Observatory's Legacy Survey of Space and Time (LSST) [12].

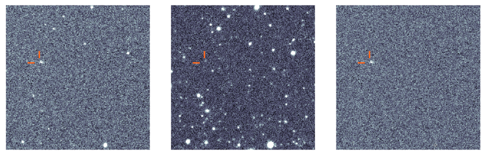
[7]We experimented with pre-trained weights, however that neither helped the network to reach convergence faster, nor did it affect the final performance. We believe this is likely due to the fact that astronomical images are very different from those in commonly-used data sets.

[8]Standard fully-convolutional approaches often used in computer vision proved to be an overkill in this case.

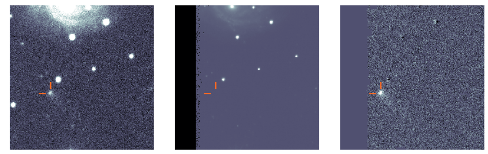
[9]`https://jupyter.org`

[10]`https://www.minorplanetcenter.net/cgi-bin/checkmp.cgi`

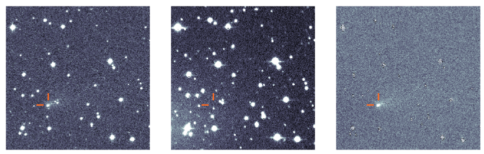
[11]`https://minorplanetcenter.net/iau/NEO/pccp_tabular.html`



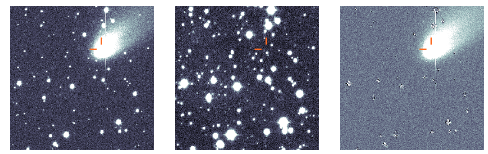
(a) C/2020 T2 (Palomar) discovered with Tails on 2020/10/07



(b) 2I/Borisov observed on 2019/10/15



(c) C/2019 D1 (Flewelling) observed on 2019/07/22



(d) 260P/McNaught observed on 2019/10/19

Figure 2: Examples of comets identified and localized by Tails, including C/2020 T2, the first AI-assisted discovery of a comet (a). For each image triplet, the left pane shows the epochal science exposure, the middle pane – the reference image of the corresponding patch of sky, and the right pane – the ZOGY difference image.

Broader Impact

We believe that this work may enable discoveries of new, potentially exciting, celestial objects. While Tails is trained only on ZTF data, with transfer learning, it can be easily adapted to other sky surveys, both ground- and space-based, including the upcoming Vera Rubin Observatory's Legacy Survey of Space and Time (LSST).

Acknowledgments and Disclosure of Funding

D. A. Duev would like to thank Ivan Duev for assistance with data labeling; Bryce Bolin, Michael Kelley, Ashish Mahabal, and Matthew Graham for their support and useful discussions.

D. A. Duev acknowledges support from Google Cloud and from the Heising-Simons Foundation under Grant No. 12540303.

Based on observations obtained with the Samuel Oschin Telescope 48-inch Telescope at the Palomar Observatory as part of the Zwicky Transient Facility project. ZTF is supported by the National Science Foundation under Grant No. AST-1440341 and a collaboration including Caltech, IPAC, the Weizmann Institute for Science, the Oskar Klein Center at Stockholm University, the University of Maryland, the University of Washington, Deutsches Elektronen-Synchrotron and Humboldt University, Los Alamos National Laboratories, the TANGO Consortium of Taiwan, the University of Wisconsin at Milwaukee, and Lawrence Berkeley National Laboratories. Operations are conducted by COO, IPAC, and UW.

References

[1] E. C. Bellm et al. "The Zwicky Transient Facility: System Overview, Performance, and First Results". In: *PASP* 131.995 (Jan. 2019), p. 018002. DOI: `10.1088/1538-3873/aaecbe`. arXiv: `1902.01932 [astro-ph.IM]`.

[2] S. Bloemen et al. "MeerLICHT and BlackGEM: custom-built telescopes to detect faint optical transients". In: *Ground-based and Airborne Telescopes VI*. Vol. 9906. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. July 2016, p. 990664. DOI: `10.1117/12.2232522`.

[3] K. C. Chambers et al. "The Pan-STARRS1 Surveys". In: *arXiv e-prints*, arXiv:1612.05560 (Dec. 2016), arXiv:1612.05560. arXiv: `1612.05560 [astro-ph.IM]`.

[4] "Comet C/2020 T2 (Palomar)". In: *Minor Planet Electronic Circulars* 2020-U170 (2020).

[5] R. Dekany et al. "The Zwicky Transient Facility: Observing System". In: *PASP* 132.1009, 038001 (Mar. 2020), p. 038001. DOI: 10.1088/1538-3873/ab4ca2. arXiv: 2008.04923 [astro-ph.IM].

[6] A. Do, M. A. Tucker, and J. Tonry. "Interstellar Interlopers: Number Density and Origin of 'Oumuamua-like Objects". In: *ApJL* 855.1, L10 (Mar. 2018), p. L10. DOI: 10.3847/2041-8213/aaae67. arXiv: 1801.02821 [astro-ph.EP].

[7] D. A. Duev et al. "Planetary Radio Interferometry and Doppler Experiment (PRIDE) technique: A test case of the Mars Express Phobos fly-by". In: *A&A* 593, A34 (Sept. 2016), A34. DOI: 10.1051/0004-6361/201628869. arXiv: 1606.05841 [astro-ph.IM].

[8] D. A. Duev et al. "Real-bogus classification for the Zwicky Transient Facility using deep learning". In: *MNRAS* 489.3 (Nov. 2019), pp. 3582–3590. DOI: 10.1093/mnras/stz2357. arXiv: 1907.11259 [astro-ph.IM].

[9] M. J. Graham et al. "The Zwicky Transient Facility: Science Objectives". In: *PASP* 131.1001 (July 2019), p. 078001. DOI: 10.1088/1538-3873/ab006c. arXiv: 1902.01945 [astro-ph.IM].

[10] K. He et al. "Deep Residual Learning for Image Recognition". In: *arXiv e-prints*, arXiv:1512.03385 (Dec. 2015), arXiv:1512.03385. arXiv: 1512.03385 [cs.CV].

[11] W.-H. Ip et al. "A kilometer-scale asteroid inside Venus's orbit". In: *arXiv e-prints*, arXiv:2009.04125 (Sept. 2020), arXiv:2009.04125. arXiv: 2009.04125 [astro-ph.EP].

[12] Ž. Ivezić et al. "LSST: from science drivers to reference design and anticipated data products". In: (2008). arXiv: 0805.2366v4.

[13] R. Jensen-Clem et al. "The Performance of the Robo-AO Laser Guide Star Adaptive Optics System at the Kitt Peak 2.1 m Telescope". In: *AJ* 155.1, 32 (Jan. 2018), p. 32. DOI: 10.3847/1538-3881/aa9be6. arXiv: 1703.08867 [astro-ph.IM].

[14] D. P. Kingma and J. Ba. "Adam: A Method for Stochastic Optimization". In: *arXiv e-prints*, arXiv:1412.6980 (Dec. 2014), arXiv:1412.6980. arXiv: 1412.6980 [cs.LG].

[15] F. J. Masci et al. "The Zwicky Transient Facility: Data Processing, Products, and Archive". In: *PASP* 131.995 (Jan. 2019), p. 018003. DOI: 10.1088/1538-3873/aae8ac. arXiv: 1902.01872 [astro-ph.IM].

[16] M. Tan and Q. V. Le. "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks". In: *arXiv e-prints*, arXiv:1905.11946 (May 2019), arXiv:1905.11946. arXiv: 1905.11946 [cs.LG].

[17] M. Tan, R. Pang, and Q. V. Le. "EfficientDet: Scalable and Efficient Object Detection". In: *arXiv e-prints*, arXiv:1911.09070 (Nov. 2019), arXiv:1911.09070. arXiv: 1911.09070 [cs.CV].

[18] J. L. Tonry et al. "ATLAS: A High-cadence All-sky Survey System". In: *PASP* 130.988 (June 2018), p. 064505. DOI: 10.1088/1538-3873/aabadf. arXiv: 1802.00879 [astro-ph.IM].

[19] Q. Ye et al. "A Twilight Search for Atiras, Vatiras, and Co-orbital Asteroids: Preliminary Results". In: *AJ* 159.2, 70 (Feb. 2020), p. 70. DOI: 10.3847/1538-3881/ab629c. arXiv: 1912.06109 [astro-ph.EP].

[20] B. Zackay, E. O. Ofek, and A. Gal-Yam. "Proper Image Subtraction—Optimal Transient Detection, Photometry, and Hypothesis Testing". In: *ApJ* 830, 27 (Oct. 2016), p. 27. DOI: 10.3847/0004-637X/830/1/27. arXiv: 1601.02655 [astro-ph.IM].