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

Dmitry A. Duev1,2, Bryce T. Bolin1,2, Matthew J. Graham1, Michael S. P. Kelley3, Ashish Mahabal1,4, Eric C. Bellm5, Michael W. Coughlin6, Richard Dekany7, George Helou2, Shrinivas R. Kulkarni1, Frank J. Masci2, Thomas A. Prince1, Reed Riddle7, Maayane T. Soumagnac8,9, and Stéfan J. van der Walt10

1 Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA; duev@caltech.edu
2 IPAC, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA
3 Department of Astronomy, University of Maryland, College Park, MD 20742, USA
4 Center for Data Driven Discovery, California Institute of Technology, Pasadena, CA 91125, USA
5 DIRAC Institute, Department of Astronomy, University of Washington, 3910 15th Ave. NE, Seattle, WA 98195, USA
6 School of Physics and Astronomy, University of Minnesota, Minneapolis, Minnesota 55455, USA
7 Caltech Optical Observatories, California Institute of Technology, Pasadena, CA 91125, USA
8 Lawrence Berkeley National Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA
9 Department of Particle Physics and Astrophysics, Weizmann Institute of Science, Rehovot 76100, Israel
10 Berkeley Institute for Data Science, University of California, Berkeley, CA 94720, USA

Received 2021 January 11; revised 2021 February 23; accepted 2021 February 25; published 2021 April 8

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 optical time-domain survey currently in operation at the Palomar Observatory in California, USA. Tails employs a custom 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 99% recall, a 0.01% false-positive rate, and a 1–2 pixel rms 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) and the recovery of a comet (P/2016 J3 = P/2021 A3).

Unified Astronomy Thesaurus concepts: Astroinformatics (78); Astronomy data analysis (1858); Convolutional neural networks (1938); Comets (280)

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 Gennady Borisov predictably sparked much excitement and enthusiasm among astronomers and the general public alike. 

1.1. The Zwicky Transient Facility

The ZTF12 is a state-of-the-art robotic time-domain sky survey capable of visiting the entire visible sky north of $-30^\circ$ decl. every night. ZTF observes the sky in the g, r, and i bands at different cadences depending on the scientific program and sky region (Bellm et al. 2019a; Graham et al. 2019). The 576 megapixel camera with a 47 deg$^2$ field of view, installed on the Samuel Oschin 48 inch (1.2 m) Schmidt Telescope, can scan more than 3750 deg$^2$ per hour, to a $3\sigma$ detection limit of 20.7 mag in the r band with a 30 s exposure during a new moon (Masci et al. 2019; Dekany et al. 2020).

The ZTF partnership has been running a specialized survey, the Twilight Survey (ZTF-TS) that operates at solar elongations down to 35 degrees with an r-band limiting magnitude of 19.5 (Ye et al. 2020; Bellm et al. 2019b). ZTF-TS has so far resulted in the discovery of a number of Atira asteroids (orbits interior to Earth’s) as well as the first inner-Venus object, 2020 AV2 (Ip et al. 2020). Motivated by the success, ZTF-TS will be expanded in Phase II of the project, which commenced in 2020 December.

Comets become more easily detectable when close to the Sun as they become brighter and start exhibiting more

---

11 https://github.com/dmitryduev/tails
12 https://ztf.caltech.edu
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 of observational selection effects (Jedicke et al. 2016) and the fact that this direction has a greater cross section for asteroids to bend around and pass into the visibility volume (Engelhardt et al. 2017; Do et al. 2018).

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

Deep learning (DL) is a subset of machine learning that employs artificial multi-layer neural networks (McCulloch & Pitts 1943). DL systems are able to discover, in a highly automated manner, efficient representations of the data, simplifying the task of finding the meaningful sought-after patterns in them. We refer the reader to a brilliant introduction into DL given in Géron (2019).

DL systems often reach near-optimal performance for a given task and are able to learn even very complicated, highly nonlinear mappings between the input and output spaces. The art of building applied DL 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, and include data artifacts caused by, for example, cross-talk or telescope reflections.

2.1. Data Set

To build a seed sample for labeling, we first identified all potential observations of known comets conducted with ZTF from 2018 March 5 to 2020 March 4, based on their predicted position and brightness. The code for accomplishing that is sourced for manual annotation. This resulted in an initial sample of 3000 examples with identifiable morphology.

We selected over 60,000 individual observations with the total comet magnitude ranging from 10 to 23 (as reported by JPL Horizons; see Figure 1), out of which about 20,000 were sourced for manual annotation. This resulted in an initial sample of 3000 examples with identifiable morphology.

2.2. Deep Neural Network Architecture and Training

Tails adopts a custom architecture (see Figure 3) based on EfficientDet D0 (Tan et al. 2019), a variant of a state-of-the-art architecture designed for object detection—a computer-vision technique for the identification and location of objects in image data.

This architecture delivers best-in-class object detection efficiency and performance across a wide range of resource constraints. This is achieved by using EfficientNet—state-of-the-art backbone networks for feature extraction, a weighted bidirectional feature pyramid network (BiFPN), which allows easy and fast multiscale feature fusion, and a compound scaling method that simultaneously and uniformly scales the resolution, depth, and width for all backbone, feature, and location/class prediction networks (Tan et al. 2019).

The use of a BiFPN, which effectively represents and processes multiscale features, makes this architecture particularly well suited for the problem of morphology-based comet identification and localization.

A batch of triplet image stacks ($n_b$, 256, 256, 3) in size, where $n_b$ is the number of stacks in the batch, is passed through an EfficientNet B0 backbone (Tan & Le 2019). The extracted
features from the last five blocks/levels of the network are passed through the BiFPN. The resulting five output tensors denoted by the colored circles in Figure 3 are fed into the head network, which outputs the probability of the image containing a comet \( p_c \) and its predicted relative position \((x, y)\). We defined the loss function as

\[
L = w_c \cdot L_c + w_p \cdot L_p,
\]

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 \([p_c] = 1\), \( L_p \) 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 \( \epsilon = 10^{-3} \)), and \( w_c \) and \( w_p \) denote the weights of the two terms, respectively,

\[
L_c = \sum c \cdot \log(p_c) + (1 - c) \cdot \log(1 - p_c),
\]

\[
L_p = \sum [p_c] \cdot (|x - x_p| + |y - y_p|) + \epsilon \cdot \sqrt{|x - x_p|^2 + |y - y_p|^2}.
\]

We employed the Adam optimizer (Kingma & Ba 2014), 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. We note that the test/validation sets did not contain augmented data from the training set. We used standard techniques to maximize training performance: if no improvement in validation loss was observed for 10 epochs, the learning rate was reduced by a factor of 2, and training was stopped early if no improvement was observed for 30 epochs.

The EfficientNet’s weights were randomly initialized. We first set $w_c = 10, w_p = 1$ to allow for a fast convergence of the feature-extracting part of the network. 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 rms error, and finally, we 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-TS data.

and vertical flips of the input data; no random rotations and translations were added. We note that the test/validation sets did not contain augmented data from the training set. We used standard techniques to maximize training performance: if no improvement in validation loss was observed for 10 epochs, the learning rate was reduced by a factor of 2, and training was stopped early if no improvement was observed for 30 epochs.

The EfficientNet’s weights were randomly initialized. We first set $w_c = 10, w_p = 1$ to allow for a fast convergence of the feature-extracting part of the network. 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 rms error, and finally, we 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-TS data.

### 3. Tails Performance

Evaluated on the test set, with a score $p_c$ threshold of 0.5, Tails demonstrates false-positive (FPR) and false-negative rates (FNR) of 1.7%, and a $\sim 1–2$ pixel median rms error of the predicted comet centroid position versus that acquired from JPL Horizon (see Figure 4).

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 pixel tiles and evaluate Tails on those.

Tails has been deployed in production since late 2020 June. We have implemented a sentinel service that processes the incoming data in real time and posts the plausible candidates to Fritz, the ZTF Phase II open-source science data platform (van der Walt et al. 2019; Duev et al. 2019; Duev et al. 2021).
Kasliwal et al. 2019, for further manual inspection and vetting. The candidates are auto-annotated with the detailed information on the detection such as the score, CCD, and sky positions, and cross-matches with known solar system objects. Figure 5 shows screenshots of the Fritz user interfaces used in the process.

It takes about 5 hr to run inference on a typical set of nightly ZTF-TS data (∼45 30 s exposures) on an e2-highcpu-32 virtual machine instance (32 vCPUs, 32 GB memory, SSD disk) on the Google Cloud Platform, including I/O operations. Consistently with the expected rate of comet observations, a typical run on nightly Twilight data yields a few dozen candidates, which, given the typical number of processed tiles, gives an empirical FPR value of about 0.01%.

The scanning results are accumulated and used to expand the training set and improve Tails’ performance.

Table 1

| Element | Value          |
|---------|----------------|
| $e$     | 0.9934213      |
| Incl.   | 27.87307       |
| Peri.   | 150.38279      |
| Node    | 83.04834       |
| $q$     | 2.0546940      |
| $T$     | 2021 July 11.14638 TT |

Kasliwal et al. 2019), for further manual inspection and vetting. The candidates are auto-annotated with the detailed information on the detection such as the score, CCD, and sky positions, and cross-matches with known solar system objects. Figure 5 shows screenshots of the Fritz user interfaces used in the process.

We have evaluated Tails’ performance on a random sample of 200 observations of known comets with identifiable morphology in 2020 July–August and found an empirical recall value of 99%.

Figure 6 shows a number of comet candidates not from the training set identified by Tails, including some of the ZTF observations of the comet 2I/Borisov. Optical artifacts resembling cometary objects are the main source of contamination.

3.1. Discovery of Comet C/2020 T2

On 2020 October 7, Tails discovered a candidate that was posted to MPC’s Possible Comet Confirmation Page (PCCP)20 as ZTFDD01 (see Figure 7). It was later confirmed to be a long-period comet and designated C/2020 T2 (Palomar), marking the first DL-assisted comet discovery (Duev et al. 2020). The candidate was found in the Twilight survey data; it was at 19.3 mag in the ZTF $r$ band. The FWHM of the object was approximately 2″5–3″, compared to nearby background stars that have an FWHM of ∼2″. The object showed a tail extending up to 5″ in the westward direction. Table 1 summarizes the orbital elements of C/2020 T2 provided by the MPC and Figure 8 shows its orbit as of the discovery date.

To determine if Tails could have discovered C/2020 T2 before 2020 October 7, we searched the ZTF archive for all Twilight Survey data covering the ephemeris position of the

Figure 6. Candidates identified by Tails. Panels (a), (b), and (c) on the left show the detections of real comets. Typical false positives are shown on panels (d), (e), and (f) on the right. For each image triplet, the left pane shows the epochal science exposure, the middle pane shows the reference image of the corresponding patch of sky, and the right pane shows the ZOGY difference image.

Table 1

| Element | Value          |
|---------|----------------|
| $e$     | 0.9934213      |
| Incl.   | 27.87307       |
| Peri.   | 150.38279      |
| Node    | 83.04834       |
| $q$     | 2.0546940      |
| $T$     | 2021 July 11.14638 TT |

20 https://minorplanetcenter.net/iau/NEO/pccp_tabular.html
comet with the ZChecker software (Kelley et al. 2019). Eleven nights of data were found between 2020 June 11–20 (evening twilight) and October 7–21 (morning twilight). The comet was in conjunction with the Sun between the two sets, and not observable by ZTF. We measured the brightness of the coma in 4 pixel radius apertures, and aperture corrected the photometry according to the ZTF pipeline documentation. The data are shown in Figure 9. Typical seeing was 2″ in June, and the comet was very faint (r = 20.2 mag), near the single-image detection limit (r = 20.4–20.9 mag, 5σ point source), and had no morphological features for Tails to pick up. Thus 2020 October 7 was really the first opportunity for Tails to discover the comet.

3.2. Recovery of Comet P/2016 J3 = P/2021 A3 (STEREO)

A comet candidate was identified by a combination of Tails and the ZTF Moving Object Detection Engine (Masci et al. 2019) on 2020 January 4 UTC and was submitted to the PCCP as ZTF0Ion (see Figure 10). It was later identified as a recovery of comet P/2016 J3 (STEREO; Bolin 2021). P/2021 A3 was identified in the evening Twilight survey data at r = 19.3 mag with a clearly extended appearance scoring 0.9 with a coma ~10″ wide and a tail extending past 20″ in the northeast direction.

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, with a specific focus on comets.

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 adapted to other sky surveys, including the upcoming Vera Rubin Observatory’s LSST (Ivezić et al. 2008).
D.A. Duev would like to thank Ivan Duev for assistance with data labeling. 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 and the 60 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.

This research has made use of data and/or services provided by the International Astronomical Union’s Minor Planet Center.

The authors would like to express gratitude to the anonymous referee.

Facilities: PO:1.2 m, ZTF.

Software: astropy (Astropy Collaboration et al. 2018), Fritz (https://github.com/fritz-marshall/fritz), Kowalski (Duev et al. 2019), matplotlib (Hunter 2007), numpy (Harris et al. 2020), pandas (Pandas Development Team 2020), pypride (Duev et al. 2016), SEP (Barby 2016), sbpy (Momert et al. 2019), TensorFlow (Abadi et al. 2016), ZChecker (Kelley et al. 2019).

**References**

Abadi, M., Agarwal, A., Barham, P., et al. 2016, arXiv:1603.04467

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123

Barbary, K. 2016, JOSAA, 1, 58

Bellm, E. C., Kulkarni, S. R., Barlow, T., et al. 2019b, PASP, 131, 068003

Bellm, E. C., Kulkarni, S. R., Graham, M. J., et al. 2019a, PASP, 131, 018002

Bloom, S., Groot, P., Woudt, P., et al. 2016, Proc. SPIE, 9906, 990664

Bolin, B. T., Lisse, C. M., Kasliwal, M. M., et al. 2020, AJ, 160, 26

Bolin, B. T. E. A., Lin, Y.-Z., Masci, F. J., et al. 2020, MPEC 2021-A157: COMET P/2016 J3 = P/2021 A3 (STEREO), https://www.minorplanetcenter.net/iau/MPEC/K20/K21AF7.html

Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv:1612.05560

Dekany, R., Smith, R. M., Riddle, R., et al. 2020, PASP, 132, 038001

Do, A., Tucker, M. A., & Tonry, J. 2018, ApJL, 855, L10

Duev, D. A., Mahabal, A., Masci, F. J., et al. 2019, MNras, 489, 3582

Duev, D. A., Pogrebenko, S. V., Cimó, G., et al. 2016, A&A, 593, A34

Duev, D. A. E. A., Duev, I. D., Bolin, B. T., et al. 2020, MPEC 2020-U170: COMET C/2020 T2 (Palomar), https://minorplanetcenter.net/iau/MPEC/K20/K20UIB0.html

Engelhardt, T., Jedicke, R., Vereš, P., et al. 2017, AJ, 153, 133

Fitzsimmons, A., Hainaut, O., Meech, K. J., et al. 2019, ApJL, 885, L9

Géron, A. 2019, Hands-on machine learning with Scikit-Learn and TensorFlow (2nd ed.; Sebastopol, CA: O’Reilly Media, Inc.)

Giorgini, J. D., Yeomans, D. K., Chamberlin, A. B., et al. 1996, AAS/DPS Meeting, 28, 2504

Graham, M. J., Kulkarni, S. R., Bellm, E. C., et al. 2019, PASP, 131, 078001

Guzik, P., Drahus, M., Ruszk, K., et al. 2020, NatAs, 4, 53

Harris, C. R., Jarrod Millman, K., van der Walt, S. J., et al. 2020, Natur, 585, 357

He, K., Zhang, X., Ren, S., & Sun, J. 2015, arXiv:1512.03385

Hunter, J. D. 2007, CSE, 9, 90

Ip, W. H., Bolin, B. T., Masci, F. J., et al. 2020, arXiv:2009.04125

Ivezic, Z., Tyson, J. A., Acosta, E., et al. 2008, arXiv:0805.2366v4

Jedicke, R., Bolin, B., Granvik, M., & Beshore, E. 2016, Icar, 266, 173

Jensen-Clem, R., Duev, D. A., Riddle, R., et al. 2018, AJ, 155, 32

Kasliwal, M. M., Fernández, J., Lunniss, P., et al. 2017, AJ, 153, 4

Knezevic, S., D’Alessio, P., Males, J., et al. 2019, PASP, 131, 038003

Kelley, M. S. P., Bodewits, D., Ye, Q., et al. 2019, in ASP Conf. Ser., 523, Astronomical Data Analysis Software and Systems XXVII ed. P. J. Teuben et al. (San Francisco, CA: ASP), 471

Kingma, D. P., & Ba, J. 2014, arXiv:1412.6980

Masci, F. J., Laher, R. R., Raskolnik, B., et al. 2019, PASP, 131, 018003

**ORCID IDs**

Dmitry A. Duev https://orcid.org/0000-0001-5060-8733

Bryce T. Bolin https://orcid.org/0000-0002-4950-6323

Matthew J. Graham https://orcid.org/0000-0002-3168-0139

Michael S. P. Kelley https://orcid.org/0000-0002-6702-7676

Ashish Mahabal https://orcid.org/0000-0003-2242-0244

Eric C. Bellm https://orcid.org/0000-0001-8018-5348

Michael W. Coughlin https://orcid.org/0000-0002-8262-2924

Richard Dekany https://orcid.org/0000-0002-5884-7867

George Helou https://orcid.org/0000-0003-3367-3415

Shrinivas R. Kulkarni https://orcid.org/0000-0001-5390-8563

Frank J. Masci https://orcid.org/0000-0002-8532-9395

Thomas A. Prince https://orcid.org/0000-0002-8850-3627

Reed Riddle https://orcid.org/0000-0002-0387-370X

Maya T. Soumagnac https://orcid.org/0000-0001-6753-1488

Stefan J. van der Walt https://orcid.org/0000-0001-9276-1891
McCulloch, W. S., & Pitts, W. 1943, The Bulletin of Mathematical Biophysics, 5, 115
Mommert, M., Kelley, M., de Val-Borro, M., et al. 2019, JOSS, 4, 1426
Pandas Development Team 2020, T. pandas-dev/pandas: Pandas, latest,
Zenodo doi:10.5281/zenodo.3509134
Schleicher, D. G., Millis, R. L., & Birch, P. V. 1998, Icar, 132, 397
Tan, M., & Le, Q. V. 2019, arXiv:1905.11946
Tan, M., Pang, R., & Le, Q. V. 2019, arXiv:1911.09070
Tonry, J. L., Denneau, L., Heinze, A. N., et al. 2018, PASP, 130, 064505
van der Walt, S. J., Crellin-Quill, A., & Bloom, J. S. 2019, JOSS, 4, 1247
Ye, Q., Masci, F. J., Ip, W.-H., et al. 2020, AJ, 159, 70
Zackay, B., Ofek, E. O., & Gal-Yam, A. 2016, ApJ, 830, 27