DOI: 10.13009/EUCASS2023-669
Aerospace Europe Conference 2023 – 10TH EUCASS – 9TH CEAS

Exploring large astronomical data archives for the characterization of space debris

Elisabeth RACHITH[⋆†], *Stephan HELLMICH*[⋆], *Belén Yu IRURETA-GOYENA*[⋆] *and Jean-Paul KNEIB*[⋆]
[⋆]*École Polytechnique Fédérale de Lausanne (EPFL)*
Route Cantonale, 1015 Lausanne
elisabeth.rachith@epfl.ch · stephan.hellmich@epfl.ch · belen.irureta@epfl.ch · jean-paul.kneib@epfl.ch
[†]Corresponding author

Abstract

This work aims to gain a deeper understanding of the current state of the orbital debris population. We are currently developing new and innovative methods to efficiently extract observations of satellites, debris and solar system objects from large astronomical data archives. The data obtained from these archives will provide valuable insights into the size-frequency distribution of orbital debris and their physical properties and support the development of strategies to maintain the usability of Earth's orbit for future satellite missions. This paper is an outline of the current status and future developments of our research.

1. Introduction

Space debris – man-made and non-functional objects in Earth's orbit – are growing in number and raising major concerns about the sustainability of space operations. Active satellites can be severely damaged by small debris and completely destroyed by larger ones, requiring significant effort and resources for collision avoidance and mitigation methods. Various sources contribute to the creation of space debris, including old satellites, rocket bodies, anti-satellite tests and fragmentation or collisions between objects, and the increasing number of satellites, including large constellations, further exacerbates the problem. Depending on their orbit, debris can remain in space for varying periods ranging from a few days to thousands of years before naturally decaying in the atmosphere. As well as posing operational risks, space debris also affect astronomical research by leaving traces on telescope images [14] [1].

In order to carry out collision avoidance manoeuvres as effectively as possible, to enable active debris removal missions to approach targets, or more generally to gain a better understanding of the space environment in which we operate, it is essential to obtain and analyse data that is as accurate and complete as possible. The orbital elements of most of the largest objects are known and regularly updated, thanks to large telescope and radar networks that track space debris on a daily basis[1]. However, in addition to the lack of precision from which these catalogues sometimes suffer, and the fact that catalogues tend to be incomplete, particularly for high orbits [3], intrinsic characteristics such as shape or material, as well as their rotational state, are in most cases unknown, although this type of information is essential for the above-mentioned operations [5] [23] [26].

Our research aims to address the lack of information on the physical properties of objects and to improve the statistical knowledge of the general population of space debris by studying archival images from large telescopes primarily used for astrophysical research, using innovative debris and satellite track detection techniques and photometric analysis tools. NEO (Near Earth Objects) detection techniques are also being developed as part of this project and are presented in this article.

2. Data

Every year, astronomical research generates thousands of images of the night sky, captured by the world's largest and most powerful telescopes. Although space debris is not the main focus of most research, both debris and active

[1]https://www.space-track.org

DOI: 10.13009/EUCASS2023-669

satellites in orbit inevitably pass in front of telescope sensors and, when illuminated by sunlight, leave visible traces on the images captured. Most of these images go through reduction pipelines before they are used, and unwanted streaks are removed by superimposing multiple exposures. The idea behind this work is to use these same images, before the tracks are removed, to detect space debris and derive light curves to study first their rotational state and subsequently other physical properties such as their sizes and shapes. The use of such images is of great interest for the study of space debris because, as the images are primarily produced for another purpose, their analysis does not require the use of dedicated observing stations, and because the archives of these telescopes extend over several years, it is possible to carry out a statistical analysis over a long period of time, which is crucial for our understanding of the evolution of the debris population. In addition, data from large telescopes have superior resolution and sensitivity compared to most dedicated space debris observations, making it possible to detect and study faint debris.

The techniques developed will first be applied to images from the OmegaCam sensor mounted on the VST (VLT Survey Telescope) in Chile [15], before being applied to other archives from large field-of-view telescopes such as ZTF [2] and DECam [24].

2.1 OmegaCam archive

OmegaCAM is a 300-megapixel wide-field optical camera consisting of 32 individual CCDs covering one square degree with a resolution of 0.2 arcsec per pixel. The first images were taken in October 2011 and represent approximately half a million images taken over more than 10 years. Exposure times range from 30s to 300s. Figure 1 shows an estimate of the size of objects we expect to see in the OmegaCam and DECam archives. We estimate streaks with SNR greater than 3 to be detectable by our algorithms and those with SNR greater than 30 to be suitable for photometric analysis. The methods used to estimate the detection limits can be found in a previous paper [10].

3. Satellites and space debris detection

3.1 Detection algorithm

The algorithm developed for the detection of satellite and space debris traces is a machine learning model based on the work of Lin et al. [17], a convolutional neural network with a trainable Hough transform prior block called HT-LCNN (Hough Transform Lookup based Convolutional Neural Network). The introduction of the prior block allows to reduce the amount of data required during the training phase, as the model already has an initial knowledge of the shapes to look for. In fact, the Hough transform algorithm is widely used to detect lines in images. A diagram of how the Hough transform block works can be seen in Figure 2. This block can be placed in several locations in the code to allow specific tuning of the global LCNN algorithm.

The original algorithm developed by Lin and his colleagues was trained and tested on the Wireframe [11] and York Urban [6] datasets, two datasets designed for the detection of edges in buildings and interiors, and was primarily intended for automatic orientation of smart objects.

3.2 Datasets

In order to tune the algorithm to our specific situation, i.e. the detection of traces in astronomical images, it is necessary to create our own dataset for training and testing the neural network. 193 VST mosaics, corresponding to 6176 individual CCD frames, were prepared for this purpose. The preparation consisted of resizing the images to remove dark edges, converting the format from fits to png using a zscale algorithm, as this is the preferred input format of the machine learning method, manually annotating the existing streaks by noting their start and end points, and adding artificial streaks. Artificial streaks are created by adding lines of random orientation and SNR with widths following a Gaussian profile. The SNR is calculated locally along the streak and has values between 2 and 100. A total of 12352 artificial streaks have been added to the images, in addition to the 302 natural streaks already present.

To ensure adequate validation of the algorithm, a training split and a test split were created containing approximately 90% and 10% of the initial image set respectively. Data augmentation, consisting of rotation by 90 degrees, 180 degrees and 270 degrees, was applied to the training set, resulting in a total of 22328 images for training the algorithm. The

DOI: 10.13009/EUCASS2023-669

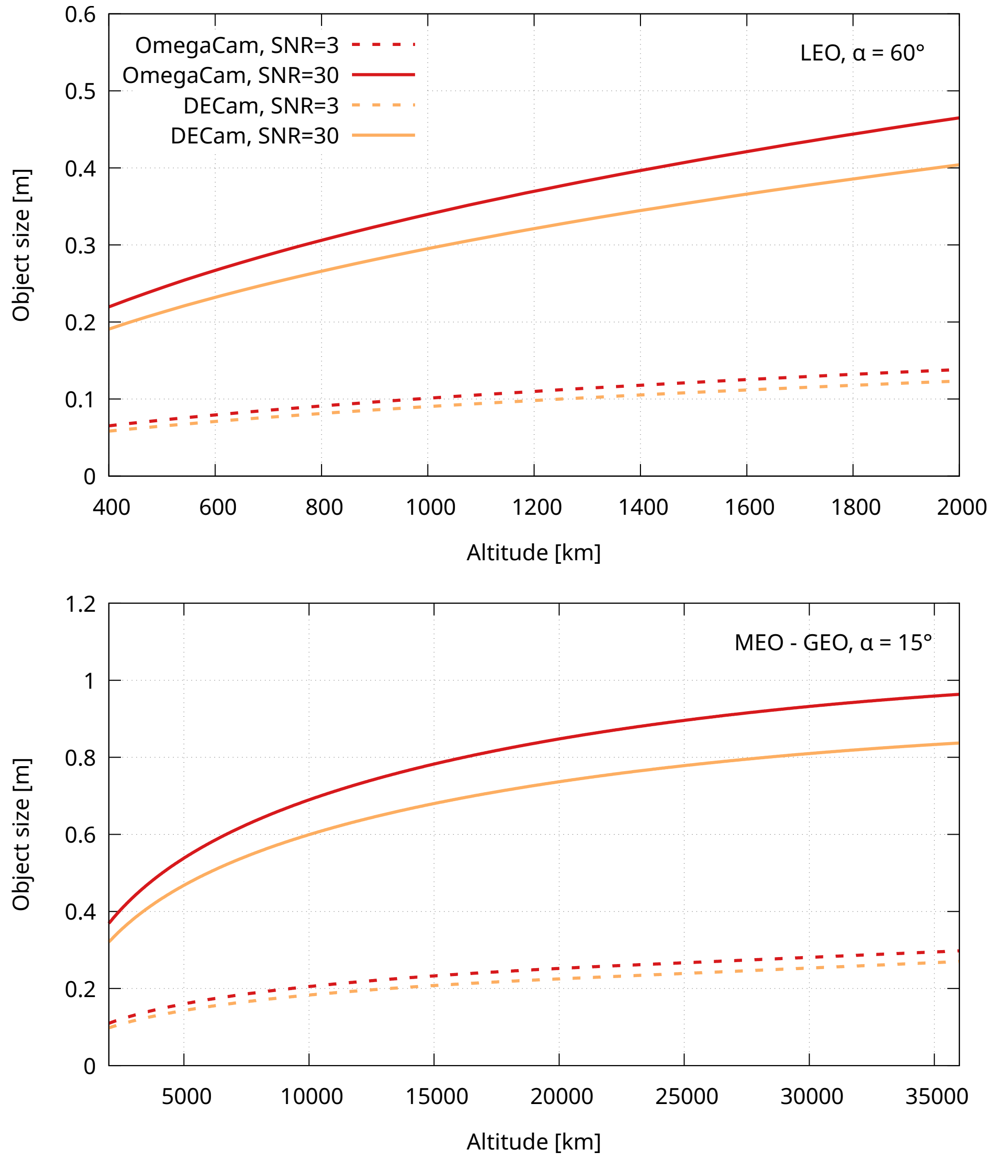


Figure 1: Size estimate of objects leaving a streak of SNR 3 (dashed lines) and SNR 30 (solid line) on images for different orbital regimes. From [10].

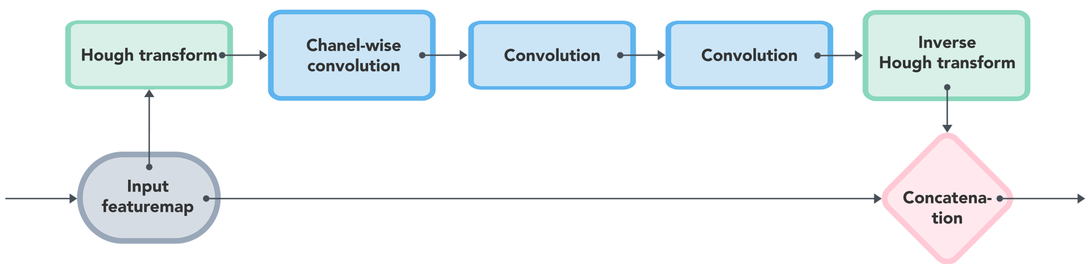


Figure 2: Flowchart of the Hough transform prior block. Adapted from [17].

DOI: 10.13009/EUCASS2023-669

test split contains 594 images.

It is important to note that the data was not calibrated before the algorithm was trained or after it was used. This is to avoid downloading unnecessary data and adding extra computation time to the detection pipeline, a critical element in our research as we are dealing with extremely large amounts of data. The images in which satellite or space debris streaks are detected are calibrated for photometric analysis at a later stage.

3.3 Evaluation metrics

In order to evaluate the results of the machine learning algorithm on the validation set, several metrics commonly adopted in the field are used and presented below.

Precision Precision is an indication of how many cases detected as positive are actually true positives. It tells us how good the algorithm is at making correct predictions when something is detected, and is calculated as follows.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \tag{1}$$

Recall Recall gives an indication of how many positive detections, whether true or false, have been made by the model. It indicates the ability of the model to make positive detections. The formulation is as follows.

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \tag{2}$$

Average precision When evaluating a neural network, both the precision and the recall should be as high as possible, which in our specific case means that the model would detect all streaks and that all detected lines are indeed debris or satellite streaks. However, in practice this is almost impossible to achieve and it is therefore necessary to modify the evaluation threshold used to separate the detections that are considered as streaks or not. To avoid having to choose such a threshold at an early stage of the model development, a metric that uses both the precision and recall definitions is adopted, namely the average precision. To define it, it is first necessary to draw the precision-recall curve, plotting the precision and recall values of the model at different evaluation thresholds. The average precision is then defined as the area under the precision-recall curve, as follows, with R the recall and P(R) the precision as a function of the recall.

$$Average\ Precision = \int_{R=0}^{1} P(R)dR \tag{3}$$

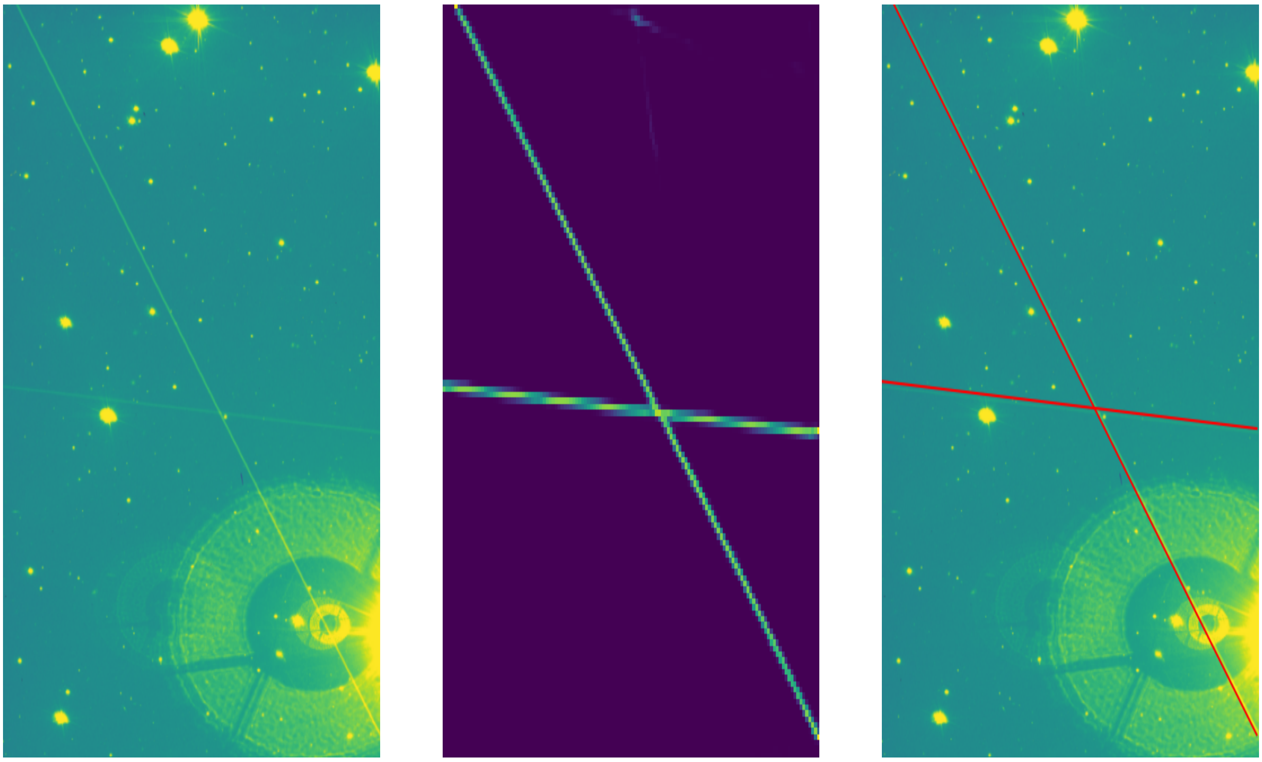
3.4 Results

In order to maintain the good performance of the original algorithm while fine-tuning it for our specific case, the model was trained on the VST dataset from the best performing network trained on the Wireframe and York Urban datasets, thus starting from the weights and biases given by the authors. The network was trained for 24 epochs in just over 30 hours on a single GPU. The best performing model over the entire training was selected to avoid overfitting. The results of the application of the model to the validation set are presented in Table 1. The average precision of the model is 97.4%, meaning that the algorithm performs well even at high values of both precision and recall.

Table 1: Results of the final model applied to the VST validation set.

Ground Truth	True Positives	False Positives	False Negatives
1221	1193	71	28

DOI: 10.13009/EUCASS2023-669



(a) Original image (b) Detection heatmap (c) Detected lines

Figure 3: Example of a true positive detection.

True Positives A detection is considered to be a true positive when the maximum distance between the detection and the ground truth streak is less than 5 pixels. An example of a true positive detection is shown in Figure 3, where the lines have an SNR of 25 and 53 respectively. The algorithm was able to correctly detect the lines despite the presence of neighbouring stars and their diffraction peaks, as well as light reflection from the telescope.

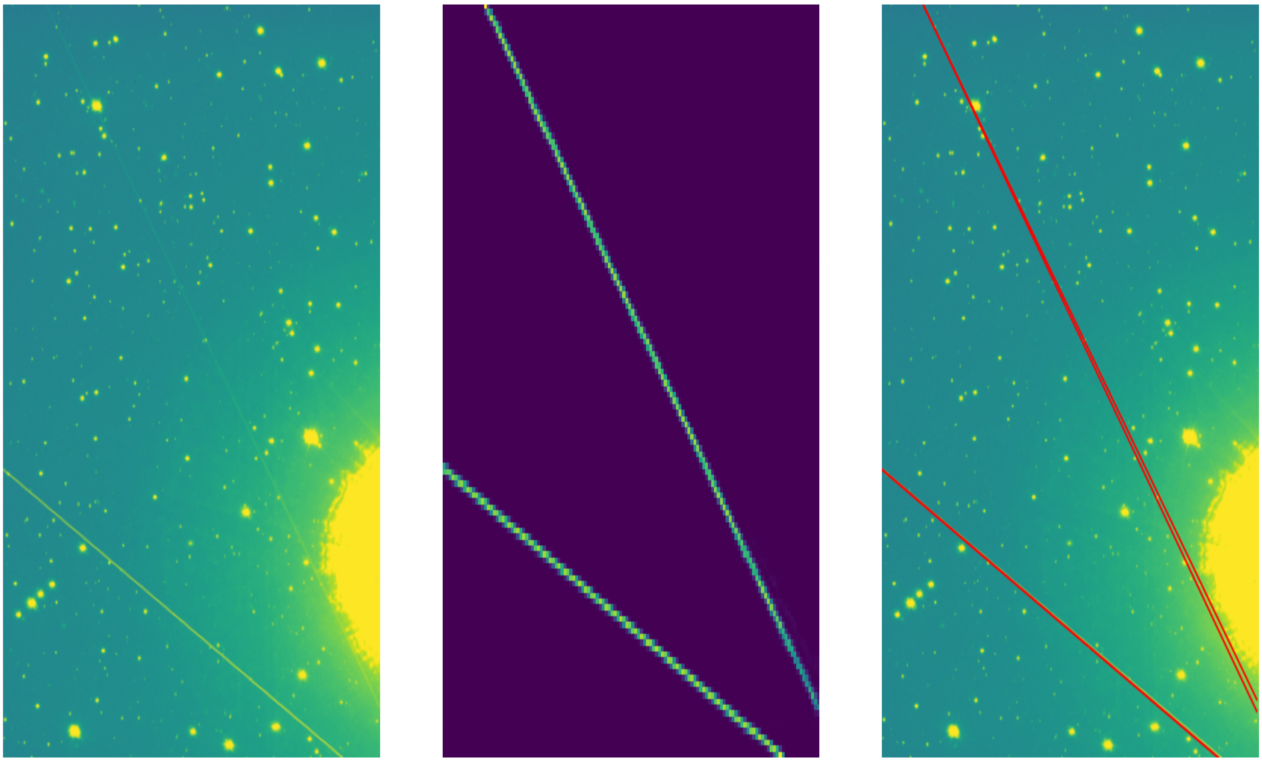
False Positives False positives occur when the algorithm detects a line that is not present in the image. Several categories of false positives have been identified and are shown in Table 2. The vast majority of false positives occurred when the algorithm detected an existing line more than once, meaning that two or more lines close to each other were the result of the detection on a single ground truth line. Figure 4 shows an example of such multiple detections. The good news is that discarding the extra detections is fairly straightforward and can be done during the post-processing of the results. False positive detections were considered anomalous if they could not be linked to an annotated line. Most of these were the result of incorrect detection of stellar diffraction peaks. Finally, the last category of false positives is quite interesting, as it refers to lines detected by the algorithm that correspond to real satellite or space debris tracks that were missed during the annotation process and therefore not part of the ground truth set. The fact that traces missed by the human eye were successfully detected by the machine learning model is a sign of its good performance and that the development of the algorithm is going in the right direction. An example of such a detection is shown in Figure 5.

Table 2: Classification of the false positives results.

False Positives	Multiple detections	Anomalous detections	Unnanotated true detections
71	64	3	4

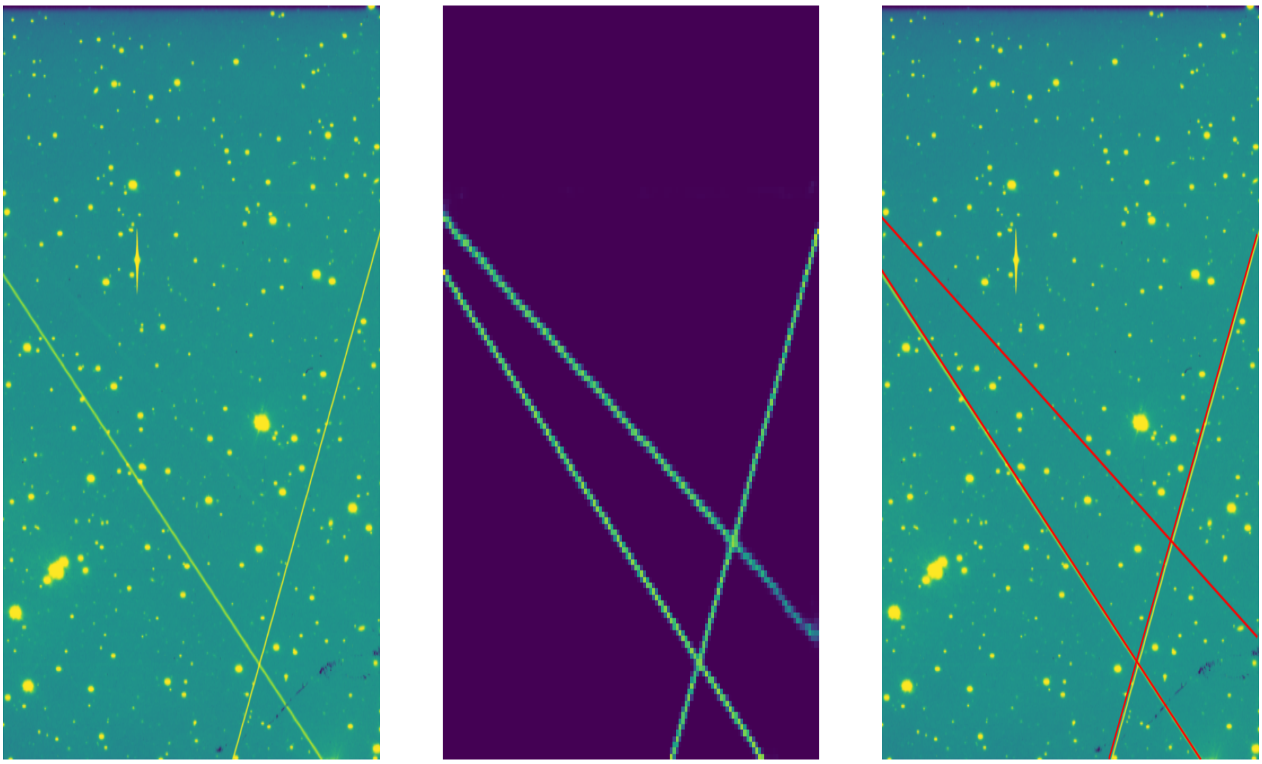
False Negatives False negatives refer to cases where a ground truth line exists but was not detected by the algorithm. Similar to false positives, different sub-categories were identified and are shown in Table 3. A distinction was made between non-detections, where a faint line was present on the heatmap and cases with no visible trace on the heatmap. In the first category are mostly cases where the streak to be detected had a low SNR (SNR < 10), and so could be seen on the heatmap, but was not considered a detection. Such an example is shown in Figure 6. A few images had a clear

DOI: 10.13009/EUCASS2023-669



(a) Original image (b) Detection heatmap (c) Detected lines

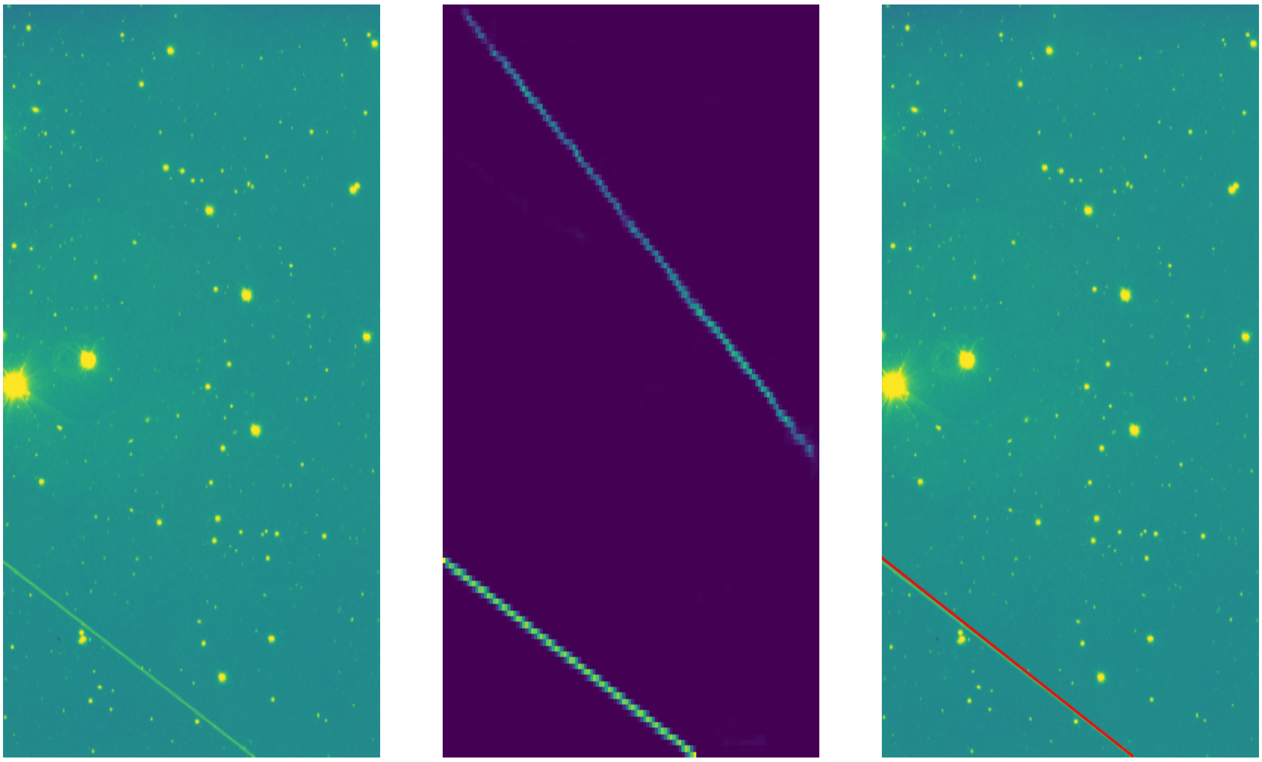
Figure 4: Example of a false positive detection resulting from a double detection of a single ground truth streaks.



(a) Original image (b) Detection heatmap (c) Detected lines

Figure 5: Example of a false positive detection resulting from the correct detection of an unannotated streak.

DOI: 10.13009/EUCASS2023-669



(a) Original image (b) Detection heatmap (c) Detected lines

Figure 6: Example of a false negative detection. The ground truth streak has an SNR of 3.

line on the heatmap, but no detection was made because the end or part of the streak was obscured by a nearby star. In the second category, no line could be seen on the heatmap and therefore no detection was made, even if a ground truth streak was present. The majority of these cases were lines with SNR 2, and the few remaining cases were where the streak was along an edge of the image. This observation allows us to say that the algorithm has a detectability limit of SNR 3, since most lines with SNR 3 were correctly detected, but SNR 2 lines were not, which is consistent with our initial assumptions. Finally, the last category of false negatives refers to the few NEO tracks that were annotated during the preparation of the dataset and in most cases were not detected by the algorithm. It makes intuitive sense that the algorithm did not learn to correctly detect these short tracks, as the vast majority of the streaks provided during training crossed the entire image. This is not a problem for the current results, as a separate algorithm is trained specifically to detect NEOs.

Table 3: Classification of the false negatives results.

False Negatives	**Undetected but seen on heatmap**	**Undetected and not seen on heatmap**	**Short tracks**
28	13	11	4

As the present research requires a very large amount of data to be processed in the shortest possible time, great care has been taken to ensure maximum efficiency of the algorithm. The current model performs the analysis of a mosaic image composed of 32 CCDs of each 2000 X 4000 pixels in just under a minute on a single GPU. Although this result is already promising, we hope to reduce the processing time even further by carefully optimising the code and using parallelisation methods on a GPU cluster, so that the entire VST archive can be processed in less than 3 months.

4. Detection of near-Earth objects

Near-Earth objects (NEOs) are solar system objects that pass close to the Earth's orbit, posing a threat of colliding with it. The consequences of this could be catastrophic: the impact of an asteroid with a diameter of 10 m would release as much energy as a nuclear bomb, whereas an asteroid of more than 3 km across could destroy all life on Earth [4].

DOI: 10.13009/EUCASS2023-669

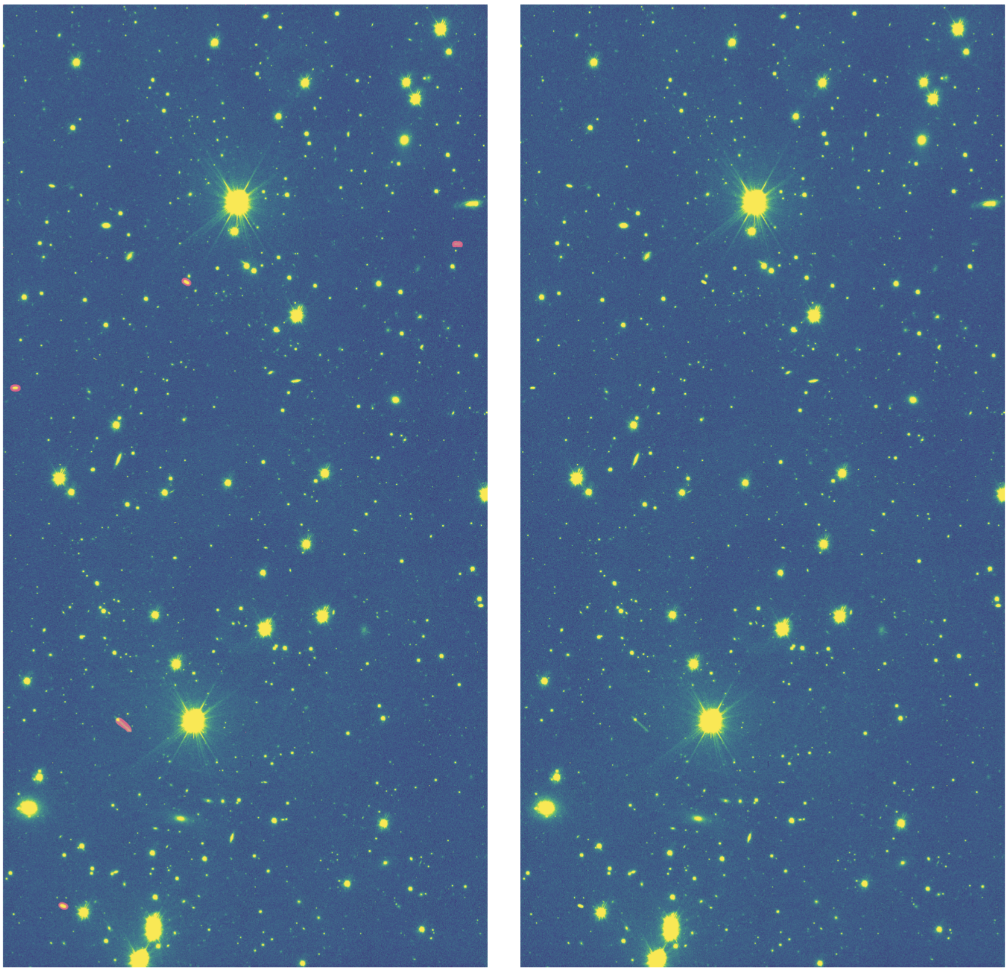


Figure 7: Example of asteroid detection performed by our algorithm. On the left, the asteroids found are marked in red. On the right, the original image. The asteroids were simulated and injected into real VST images.

Large asteroids in our Solar System have been exhaustively documented and monitored, partly due to the efforts of surveys such as that of Pan-STARRS [12], the Catalina Sky Survey [16] or the NASA *Spaceguard Survey*, which aims to catalogue 90% of NEOs larger than 140 meters [20]. However, our knowledge of the asteroid population weakens for smaller asteroids, and is almost non-existent for those NEOs only several meters in size, a population that is expected to be of 45 millions and of which we only know 0.03% [21]. Nevertheless, these asteroids can still pose a severe threat to Earth, with the latest significant impact being in Chelyabinsk, Russia, in 2013. This asteroid, estimated to have a size of 20 m, left 1600 injured and caused 25 million euros worth of damages [13].

For this reason, it is key that we monitor NEOs to ensure that any damages are prevented. As opposed to the case of space debris and satellites, the orbits of NEOs remain constant over decades, so even looking at ancient data can help us improve our knowledge of their orbital elements today. Since NEOs are moving very fast relative to the background stars, albeit slower than space debris, they are also visible as tracks of light in long-exposure, wide-field images. In practice, NEOs will not cross the field-of-view in the VST exposures, but will rather appear as extended sources.

To detect them, we are developing a deep-learning algorithm similar to the one used for space debris, but tailored specifically to short streaks. For training the algorithm, a population of NEOs has been simulated and injected into real VST images, as shown in Figure 7. Our aim is to go through the whole VST archive with a blind asteroid search, and report any detections to Harvard's Minor Planet Center [18]. If an asteroid is detected and it was already known, reporting it will likely elongate the baseline. For example, if it was found in 2019, that means it will have been monitored from 2019 to the present. If it is detected in an image from 2017, its orbit can be extrapolated further back, from 2019 to 2017, and our knowledge of its orbit will improve. However, this is not applicable if the data where it is detected is very old, since extrapolating too far back is likely to cause great uncertainties.

DOI: 10.13009/EUCASS2023-669

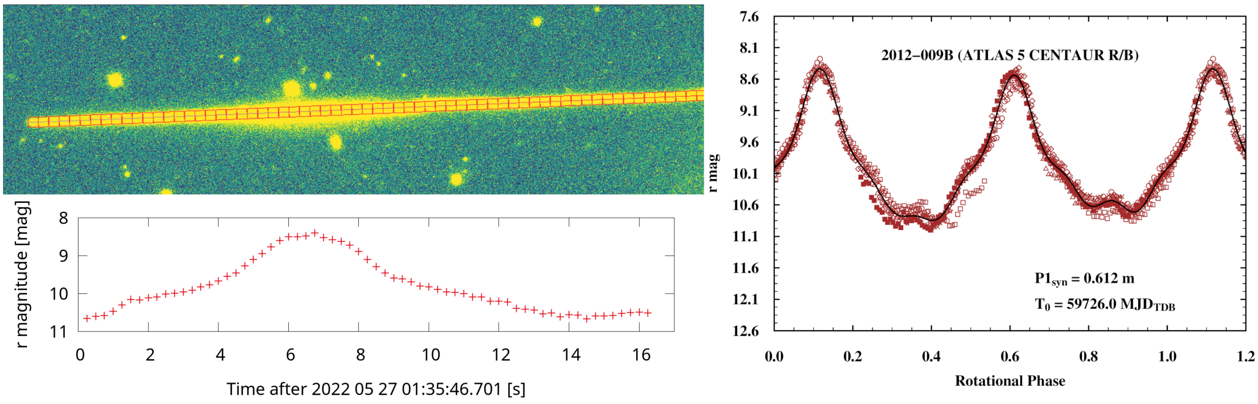


Figure 8: Satellite trail, detected on a ESO VST image, the light curve measured from the streak and all light curves from this target folded to its corresponding rotation period.

5. Photometric analysis

One goal of our work is to determine the rotational and physical characteristics of the observed orbital debris. The information needed for this task can be extracted from photometric observations. The amount of light that is reflected from the observed object to the observer depends on the illumination geometry, the rotational phase and the object's shape and surface properties. Acquiring and examining series of brightness measurements, called light curves, makes it possible to draw conclusions about these properties even if the object is not resolved. Light curve analysis is a very powerful tool to study unresolved observations and has many applications in astronomy, ranging from characterizing solar system objects [22], searching for exoplanets [19] and variable stars [8]. For our analysis, we are employing the Fourier Analysis of Light Curves (FALC) algorithm [9], that applies a least-squares fit of a Fourier polynomial to the data. This technique is considered the standard method for rotation period determination of asteroids.

To obtain the light curves, we are using a similar approach as in [5]. The intensity profile of a streak is retrieved by placing a series of rectangular apertures along the streak and measuring the flux received from the target in each aperture. To obtain precise magnitudes for each measurement, we selected a set of photometric reference stars from the GAIA DR2 catalog [7]. This procedure is currently tested and automatized in order to process the large amount of satellite streaks we expect to find in the astronomical data archives. First results demonstrate the high photometric precision of the data and high time resolution that can be achieved: We successfully determined a rotation period of 0.84s for a Russian Briz-M upper stage, observed at a distance of 37280 km [10]. We will proceed implementing methods for further photometric analysis that will allow us to determine the phase curves (the brightness as a function of the phase angle). This will enable us to estimate the size and approximate the shape of the observed objects [25].

6. Conclusion

The research presented in this paper provides a foundation for advancing our knowledge of the orbital debris population and the physical properties of space objects. The detection algorithm developed and the analysis of archival images from large telescopes pave the way for future studies and strategies aimed at ensuring the sustainability of space operations and securing the Earth's orbit for future satellite missions.

The use of archives from large telescopes such as OmegaCam offers significant advantages in terms of data availability, long-term statistical analysis, and superior resolution and sensitivity compared to dedicated space debris observations.

The developed space debris detection algorithm, based on a convolutional neural network with a trainable Hough transform prior block, shows promising results in identifying satellite and space debris traces in the images. The algorithm achieved a high average precision of 97.4% on the validation set, demonstrating its effectiveness in detecting true positives while minimising false positives and false negatives.

Furthermore, this research includes the detection of Near-Earth Objects (NEOs) approaching Earth's orbit. By developing NEO detection techniques, this work contributes to the identification and monitoring of potential threats and supports efforts to mitigate the risk of collision with Earth.

DOI: 10.13009/EUCASS2023-669

7. Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 945363.

References

[1] J. C. Barentine, A. Venkatesan, J. Heim, J. Lowenthal, M. Kocifaj, and S. Bará. Aggregate effects of proliferating low-earth-orbit objects and implications for astronomical data lost in the noise. *Nature Astronomy*, 7:252–258, 2023.

[2] E. C. Bellm and al. The Zwicky Transient Facility: System overview, performance, and first results. *Publications of the Astronomical Society of the Pacific*, 131, 2019.

[3] J. A. Blake, P. Chote, D. Pollacco, W. Feline, G. Privett, A. Ash, S. Eves, A Greenwood, N. Harwood, T. R. Marsh, D. Veras, and C. Watson. Debriswatch 1: A survey of faint geosynchronous debris. *Advances in Space Research*, 67(1):360–370, 2021.

[4] C. R. Chapman. How a near-earth object impact might affect society. In *Commissioned Paper by OECD Global Science Forum for" Workshop on Near Earth Objects: Risks, Policies, and Actions," Frascati, Italy*, 2003.

[5] P. Chote, J. Blake, and D. Pollacco. Precision Optical Light Curves of LEO and GEO Objects. In *Advanced Maui Optical and Space Surveillance Technologies Conference*, page 52, 2019.

[6] P. Denis, J. H. Elder, and F. J. Estrada. Efficient edge-based methods for estimating manhattan frames in urban imagery. In *Computer Vision – ECCV 2008*, pages 197–210. Springer Berlin Heidelberg, 2008.

[7] D. W. Evans, M. Riello, F. De Angeli, J. M. Carrasco, P. Montegriffo, C. Fabricius, C. Jordi, L. Palaversa, C. Diener, G. Busso, M. Weiler, C. Cacciari, and F. van Leeuwen. *Gaia* Data Release 2. The photometric content and validation. *Astronomy & Astrophysics*, 2018.

[8] L. Eyer and N. Mowlavi. Variable stars across the observational HR diagram. *Journal of Physics: Conference Series*, 118:012010, 2008.

[9] A. W. Harris, J. W. Young, E. Bowell, L. J. Martin, R. L. Millis, M. Poutanen, F. Scaltriti, V. Zappalà, H. J. Schober, H. Debehogne, and K. W. Zeigler. Photoelectric observations of asteroids 3, 24, 60, 261, and 863. *Icarus*, 77:171–186, 1989.

[10] S. Hellmich, E. Rachith, B. Y. Irureta-Goyena, and J.-P. Kneib. Harvesting large astronomical data archives for space debris observations. In *2nd NEO and Debris Detection Conference*, volume 2. ESA Space Debris Office, 2023.

[11] K. Huang, Y. Wang, Z. Zhou, T. Ding, S. Gao, and Y. Ma. Learning to parse wireframes in images of man-made environments. In *Computer Vision – ECCV 2018*, pages 626–635, 2018.

[12] Robert Jedicke, E. A. Magnier, N. Kaiser, and K. C. Chambers. The next decade of Solar System discovery with Pan-STARRS. *Proceedings of the International Astronomical Union*, 2(S236):341–352, August 2006.

[13] A. P. Kartashova, O. P. Popova, D. O. Glazachev, P. Jenniskens, V. V. Emel'yanenko, E. D. Podobnaya, and A. Ya. Skripnik. Study of injuries from the chelyabinsk airburst event. *Planetary and Space Science*, 160:107–114, 2018.

[14] S. Kruk, P. García-Martín, M. Popescu, B. Aussel, S. Dillmann, M. E. Perks, T. Lund, B. Merín, R. Thomson, S. Karadag, and M. J. McCaughrean. The impact of satellite trails on hubble space telescope observations. *Nature Astronomy*, 7:262–268, 2023.

[15] K. Kuijken. OmegaCAM: ESO's newest imager. *The Messenger*, 146:8, 2011.

[16] S. Larson, J. Brownlee, C. Hergenrother, and T. Spahr. The Catalina sky survey for NEOs. In *American Astronomical Society, Bulletin of the American Astronomical Society, Vol. 30, p. 1037*, volume 30, page 1037, 1998.

DOI: 10.13009/EUCASS2023-669

[17] Y Lin, S. L. Pintea, and J. C. van Gemert. Deep hough-transform line priors. In *Computer Vision – ECCV 2020*, pages 323–340. Springer International Publishing, 2020.

[18] B. G. Marsden. The minor planet center. *Celestial Mechanics*, 22:63–71, 1980.

[19] J. Miralda-Escudé. Orbital Perturbations of Transiting Planets: A Possible Method to Measure Stellar Quadrupoles and to Detect Earth-Mass Planets. *The Astrophysical Journal*, 564(2), 2002.

[20] D. Morrison. *The Spaceguard survey: report of the NASA international near-Earth-object detection workshop*, volume 107979. NASA, 1992.

[21] NASA Planetary Defense Strategy & Action Plan Working Group. NASA Planetary Defense Strategy and Action Plan, 2023.

[22] P. Scheirich, P. Pravec, S. A. Jacobson, J. Ďurech, P. Kušnirák, K. Hornoch, S. Mottola, M. Mommert, S. Hellmich, D. Pray, D. Polishook, Yu. N. Krugly, R. Ya. Inasaridze, O. I. Kvaratskhelia, V. Ayvazian, I. Slyusarev, J. Pittichová, E. Jehin, J. Manfroid, M. Gillon, A. Galád, J. Pollock, J. Licandro, V. Alí-Lagoa, J. Brinsfield, and I. E. Molotov. The binary near-Earth Asteroid (175706) 1996 FG3 — An observational constraint on its orbital evolution. *Icarus*, 245:56–63, 2015.

[23] T. Schildknecht, T. Flohrer, and A. Vananti. ESA optical surveys to characterize recent fragmentation events in GEO and HEO. In *1st NEO and Debris Detection Conference*, volume 1. ESA Space Debris Office, 2019.

[24] E. Sánchez. The Dark Energy Survey. *Journal of Physics: Conference Series*, 259, 2010.

[25] J. Šilha. Space Debris: Optical Measurements. In *Reviews in Frontiers of Modern Astrophysics: From Space Debris to Cosmology*, pages 1–21. Springer International Publishing, 2020.

[26] J. Šilha, S. Krajčovič, M. Zigo, J. Tóth, D. Žilková, P. Zigo, L. Kornoš, J. Šimon, T. Schildknecht, E. Cordelli, A. Vananti, H. K. Mann, A. Rachman, C. Paccolat, and T. Flohrer. Space debris observations with the slovak AGO70 telescope: Astrometry and light curves. *Advances in Space Research*, 65(8):2018–2035, 2020.