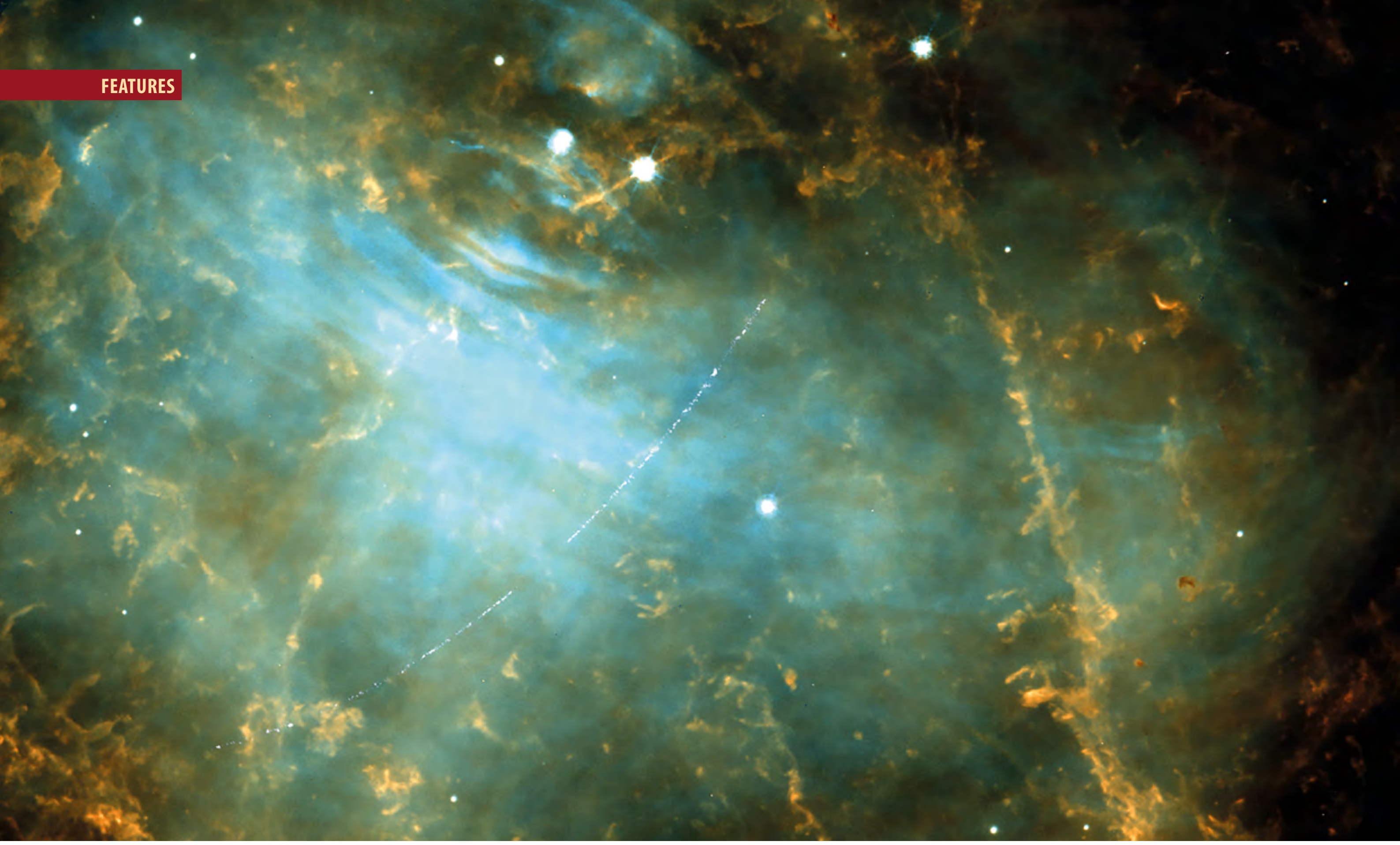


CITIZEN SCIENCE WITH ESA SCIENCE DATA
THE HUBBLE ASTEROID HUNTER PROJECT

Sandor Kruk and Bruno Merín – DOI: https://doi.org/10.1051/epn/2023206

European Space Agency (ESA), European Space Astronomy Centre (ESAC), Camino Bajo del Castillo s/n 28692Villanueva de la Cañada, Madrid, Spain

The vast amount of data in astronomy archives presents an opportunity for new discoveries. Deep learning combined with crowdsourcing provides an efficient way to explore this data using the intuition of the human brain and the processing power of machines.

▲ Faint trail of main-belt asteroid 2002 SE101, discovered by the ground-based LINEAR survey in 2001, crossing the famous Crab Nebula, as imaged by the Hubble Space Telescope in 2005. Discovery and colour composition by citizen scientists Melina Thévenot. Credit: ESA/Hubble & NASA, M. Thévenot

In the Hubble Asteroid Hunter project, we used citizen science on the Zooniverse platform and a deep learning algorithm on Google Cloud, to explore two decades of Hubble Space Telescope observations from the ESA Hubble archives for objects not targeted by the Hubble observations. The project, which was set up as a collaboration between Zooniverse, ESAC Science Data Center and Google, led to the detection of 1701 asteroids, including 1031 previously unknown ones, 198 new strong gravitational lenses and to quantifying the impact artificial satellites have on Hubble Space Telescope observations. This study is a proof of concept and shows what can be achieved by using new tools to explore the extensive astronomical archives. The archives of data held by the ESA Science Data Center are a treasure trove of information about our Universe, containing over 800 terabytes of data. But as more missions like Gaia, the Hubble Space Telescope and the James Webb Space Telescope continue to add new data every day, we need new tools and techniques to process and analyse this vast amount of information. And future missions like Euclid or Roman, will deliver over 30 petabytes over the lifetimes of the missions. Fortunately, the emergence of machine learning has provided us with the ability to process large amounts of data quickly.

However, machines are not yet capable of making new discoveries on their own, as they lack the intuition, creativity, and distraction of the human brain.

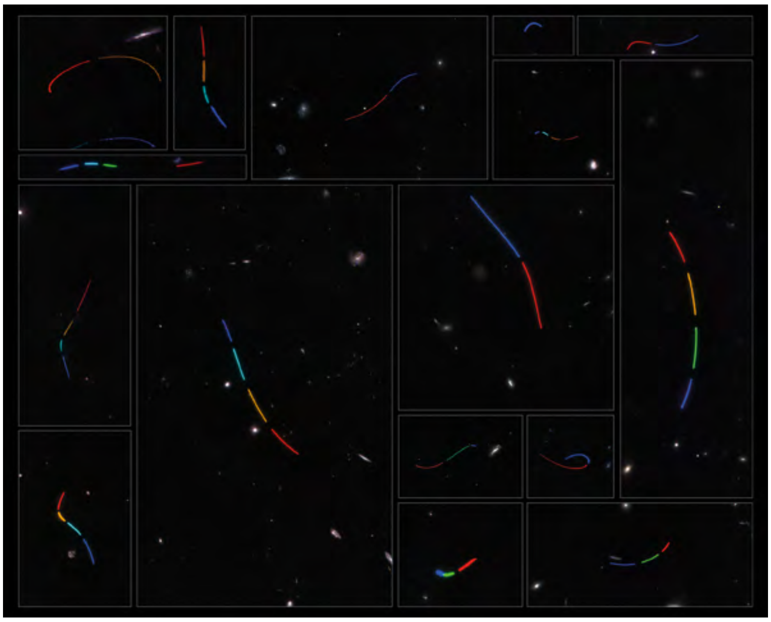
That's where citizen science comes in. By harnessing the collective effort of thousands of volunteers from across the world and the power of an artificial intelligence (AI) algorithm, we found over 1000 asteroids hiding in the ESA Hubble Space Telescope archives. The project was born after-crossmatching the orbits of known Solar System Objects (asteroids, comets, trans-Neptunian objects, *etc.*) with archival observations from Hubble, Herschel, and XMM-Newton (Racero *et al.* 2022). We found that some predicted asteroid positions had no object or had another moving object in the images. We decided to call citizen scientists to help and collaborated with the Zooniverse team, the largest platform for online citizen science projects, to build an asteroid hunting crowdsourcing project, Hubble Asteroid Hunter (www.asteroidhunter.org). We launched the project ahead of the International Asteroid Day, on 30th June 2019, asking volunteers to identify asteroids in 19 years of Hubble Space Telescope observations taken between 2002 and 2021 with the Advanced Camera for Surveys and the Wide Field Camera 3 instruments. The response of the public was overwhelming, with over 20,000 classifications per day during the first days.

Identifying asteroid trails in the images automatically turns out to be challenging without having a representative set of examples. The Hubble Space Telescope orbits around the Earth, while asteroids cross its field of view. Due to the motion of the spacecraft, asteroids appear as curved trails with a variety of shapes (see two asteroid trails passing in front of a galaxy targetted by Hubble). To solve this problem, over 11400 members of the public analysed and classified Hubble images during one year on www.asteroidhunter.org, identifying more than 1000 trails. This was an ideal training set for an automated algorithm based on AI. To be able to classify images automatically, we teamed up with Google and used the cloud-based AI algorithm, AutoML. With the use of AutoML were were able to classify 150,000 images (corresponding to 37000 Hubble 'composite' or stacked images) in just 7-hours. This was possible only with the initial help of citizen scientists. The combination of citizen science and AI resulted in a final dataset of 1701 asteroid trails (a sample of these trails is shown in Figure 2). Roughly one third of these trails could be identified and attributed to known asteroids in the IAU's Minor Planet Centre (IAU Minor Planet Center), which is the largest database of Solar System objects. This left 1031 unidentified trails that could be potential new asteroids, fainter than magnitude 22-23, which would not easily identifiable in typical ground-based surveys.

▲ **FIG. 1:** Two unidentified asteroids crossing paths in the foreground of dwarf galaxy AGC111977. Credit: ESA/Hubble & NASA, J. Cannon (Macalester College), Kruk et al. 2022 (source: *ESA - One galaxy, two asteroids : https://www.esa.int/ESA_Multimedia/Images/2020/06/One_galaxy_two_asteroids).

Although it's not possible to track the orbits of the newly detected asteroids, as the Hubble images were taken many years ago, we can still use the telescope to determine the distance to them and constrain their orbits. This is through the so-called parallax effect, imprinted by the fast motion of Hubble around the Earth and the motion of the asteroid in the sky. Most of the unknown asteroids are likely located in the Main Asteroid belt, between the orbits of Mars and Jupiter, where most asteroids are situated. Knowing the distance, the observed brightness of the objects can be eventually translated into a physical size. These • • •

▼ **FIG. 2:** Mosaic of asteroid trails detected in the Hubble Asteroid Hunter citizen science project in different images from the NASA/ESA Hubble Space Telescope. Each of these images was assigned a colour based on the time sequence of exposures, such that the blue colours represent the first exposure in which the asteroid was captured, and the red colours represent the last. Credit: ESA/Hubble & NASA, M. Zamani (ESA/Hubble), Kruk *et al.* 2022, source: Asteroid Trails Mosaic | ESA/Hubble (esahubble.org)

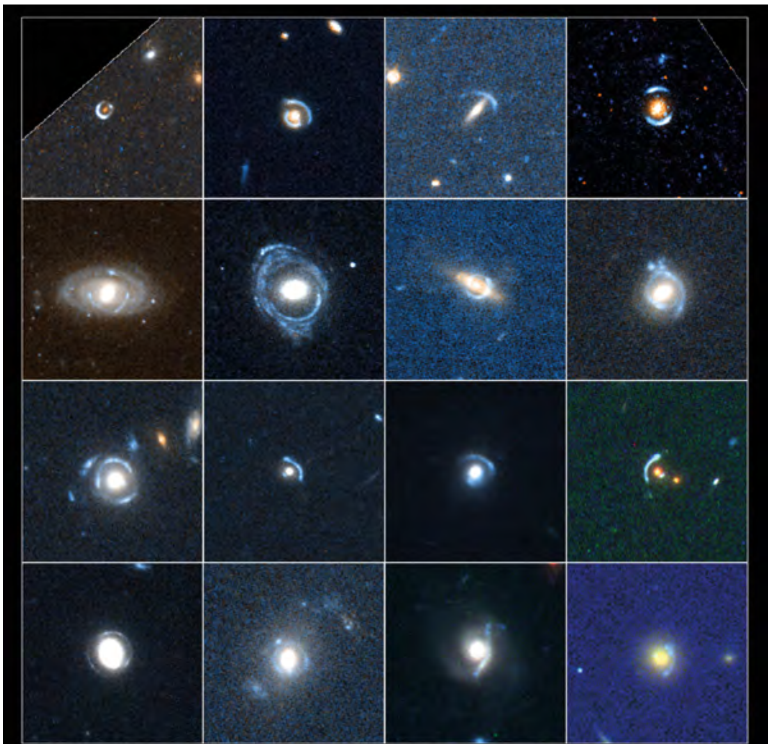


measurements will allow us to study the size distribution of the smallest objects in the solar system and give us important information to study the evolution of the asteroid population and to constrain formation models of the solar system. This result highlights the capability of Hubble to image faint, previously unknown asteroids and represents a new approach to finding asteroids in astronomical archives spanning decades, which may be effectively applied to other datasets. The James Webb Space Telescope, for example, has already detected an extremely small Main Belt Asteroid, between 100 and 200 meters in length (Müller *et al.* 2023), and this only in its first year of observations. The James Webb Space Telescope archives will likely contain many asteroid interlopers in the images, as the archive builds up over time.

Unexpected discoveries – new strong gravitational lenses

The project's second finding proved to be even more intriguing. Having human eyes on the data can lead to surprising new discoveries. While searching for asteroids, citizen scientists have stumbled across strong gravitational lensing, which they had initially classified as asteroid trails - they appear as curved trails for untrained eyes, after all. Strong gravitational lenses are one of the favourite targets of Hubble; the telescope observed many such objects discovered from the ground over time. To 'clean' the asteroid trail classifications, we provided detailed instructions to the volunteers to recognize strong gravitational lenses in the images and asked them to tag the images in the forum of the project, Talk (Hubble Asteroid Hunter » Talk — Zooniverse). Emily Garvin and a keen citizen scientist, passionate about strong lenses, Claude Cornen inspected the tagged images. They found 252 genuine strong lenses which were not the target of the Hubble observations but lying in the background of some other famous astronomical targets. Checking the existing publications on these objects in ESASky (https://sky.esa.int), we found that 198 of them were new strong lenses (Garvin *et al.* 2022), not previously reported by other studies or catalogues. This project eventually led to an unbiased search for strong lenses in archival Hubble observations. Because the volunteers were inspecting large field of view Hubble images and not postage stamps of elliptical galaxies, as commonly done in other lens searches, we also found some exotic lens configurations: spiral and edge-on disc galaxies as lenses; sources lensed by galaxy bulges; and possible double source plane lenses, as shown in Figure 3. The high resolution of the Hubble observations reveals tiny arcs and Einstein rings, which are not accessible from the ground. This discovery was only possible because the volunteers made a mistake and tagged objects they were not supposed to.

▾ **FIG. 3: Examples of new strong gravitational lenses found by citizen scientists in the Hubble Space Telescope archives, lying in the background of some famous Hubble observations.**



Sample of the 198 new strong gravitational lenses found by citizen scientists in *Hubble Space Telescope* images. Credit: NASA/ESA Hubble, Garvin et al. 2022

Satellite trails in the Hubble images

A third result of the project did not come from the extragalactic objects commonly imaged by Hubble, and was not about asteroids, but objects much closer to us – artificial satellites. Hubble's orbit has been gradually decaying, now being at an altitude of 538 km above the Earth. This is below the orbit of many human-made satellites, which can cross the field of view of the observatory during its long exposures. It turns out the combination of crowdsourcing and AI is useful in detecting satellite trails in the images, too. In the www.asteroidhunter.org project, volunteers identified satellites crossing the images on Talk. Compared to the detected asteroids, whose trails appear curved, the satellite trails appear as straight lines across the images. We found that, on average, 2.7% of Hubble images have been crossed by satellites in the last two decades of observations. The fraction increased over this time, reaching 5% for ACS in 2021, the cutoff date of our analysis, as the number of satellites in orbit above Hubble increased, in agreement with other studies (Stark *et al.* 2022). In Kruk *et al.* 2023, we explored the statistics of satellite trails in the images as a function of time, filter and pointing, as well as provided a prediction about the future, as the number of satellites in the so-called 'mega-constellations' continues to grow.

This Hubble Asteroid Hunter project combines citizen science, artificial intelligence, and cloud computing, and shows the benefits of exploring vast archival datasets, spanning decades in time, such as those hosted by the ESDC, to expand the horizons of research.

Citizen science with ESA Science data

Our exploration of the ESA Science Archives does not stop here. There are 25 missions' data in the ESA science archives, from planetary science, heliophysics to astronomy, likely hiding many more unexpected gems. That is why in 2022 we launched the Rosetta Zoo project on Zooniverse (https://www.zooniverse.org/projects/ellenjj/rosetta-zoo), asking volunteers to explore changes in the surface of the comet 67P/Churyumov-Gerasimenko, visited by the groundbreaking ESA mission, Rosetta, and famous for the landing of Philae. Over 1500 volunteers marked thousands of changes across 1000 Rosetta OSIRIS images which we are now exploring in a catalogue to be made available to scientists (Vincent *et al.* 2022). In the latest project, Gaia Vari (https://www.gaiavari.space/, we explore the variability of thousands of stars observed by ESA's 2 billion stars-surveyor, the Gaia mission. This will help scientists train and improve the classifications of automated classifiers towards the next Gaia data release. Who knows what else we might find? ■

About the authors

Sandor Kruk is Data Scientist at the European Space Agency (ESA) at ESAC, near Madrid, Spain. He develops tools for data analysis in astronomy and involves the public in genuine research through citizen science. He is the PI of the Hubble Asteroid Hunter project.

Bruno Merín is the Head of the ESAC Science Data Centre (ESDC). He works on making the data from ESA's Space Science missions accessible and is interested in citizen science and in the application of AI to the analysis of science data archives.

References

[1] Garvin *et al.* 2022, *Astronomy & Astrophysics* **667**, A141, 23

[2] Kruk *et al.* 2022, *Astronomy & Astrophysics* **661**, A85, 15

[3] Kruk *et al.* 2023, *Nature Astronomy* **7**, 262

[4] Müller *et al.* 2023, *Astronomy & Astrophysics* **670**, A53, 16

[5] Racero *et al.* 2022, *Astronomy & Astrophysics* **659**, A38, 20

[6] Stark *et al.* 2022, *STScI Instrument Science Report AC* 2022-8, 25 pages

[7] Vincent *et al.* 2022, *16th Europlanet Science Congress* 2022, id.EPSC2022-494