

RAPID Response

“Machine Learning and Anomalies” session

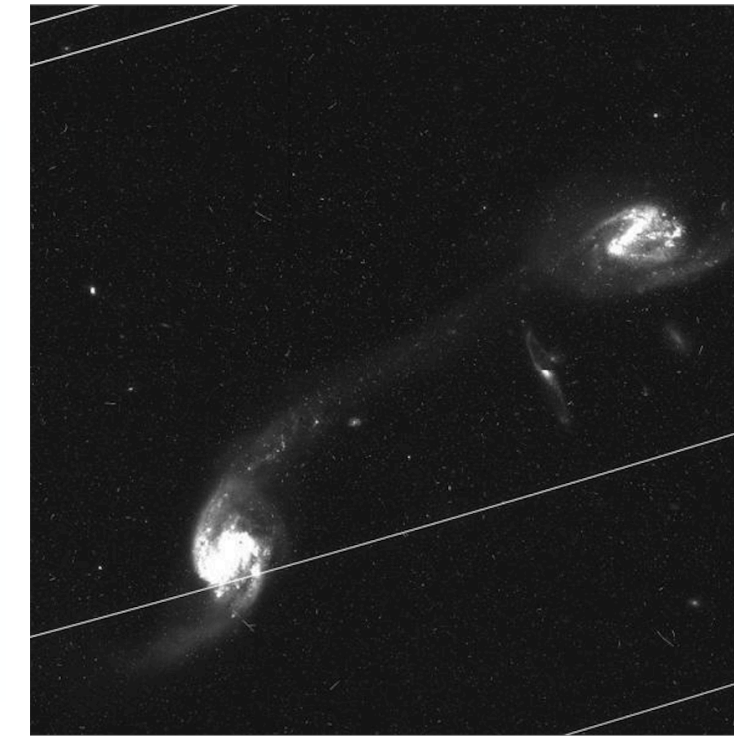
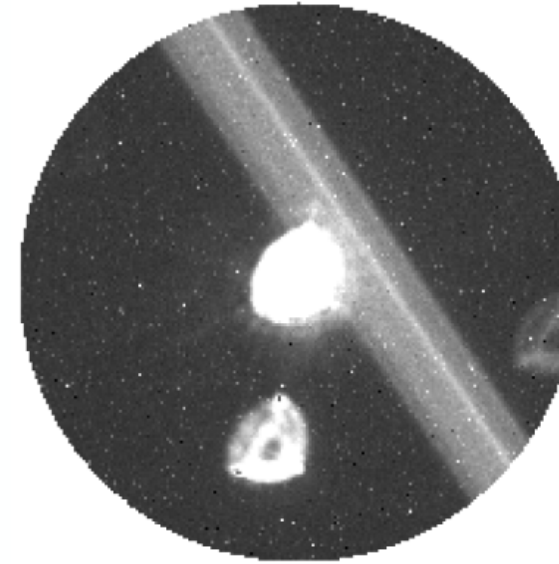
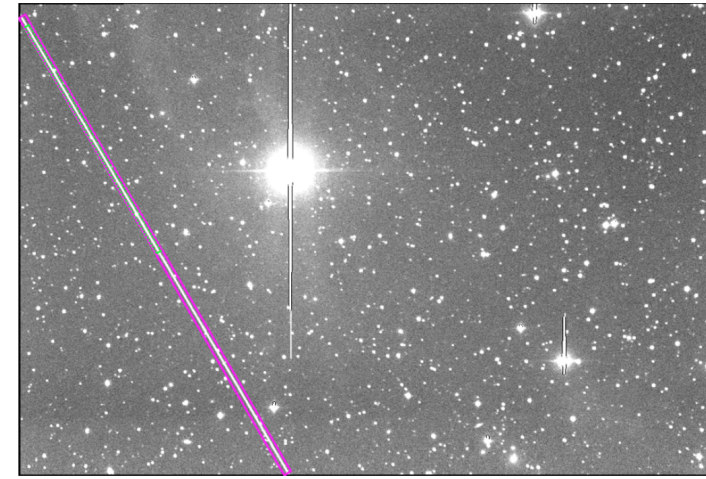
**Disentangling the Dynamic Sky: A Modular AI Framework for Distinguishing
Extrasolar Transients from Solar System Science and Satellites**

Once upon a time...

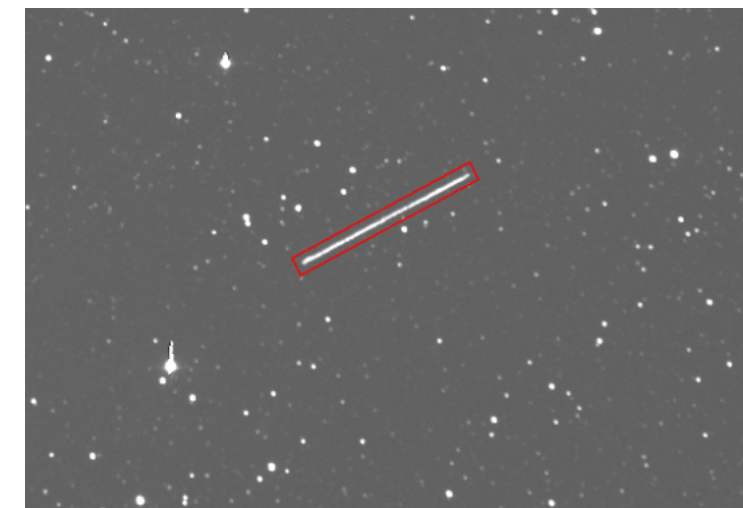
- **Once upon a time...** astronomers built next-generation telescopes to watch the sky change, creating an unprecedented ocean of data.
- **The Hero's Dilemma...** They summoned powerful AI algorithms to tame this data, only to realize it is hard to find textbooks to teach these algorithms.

The plan

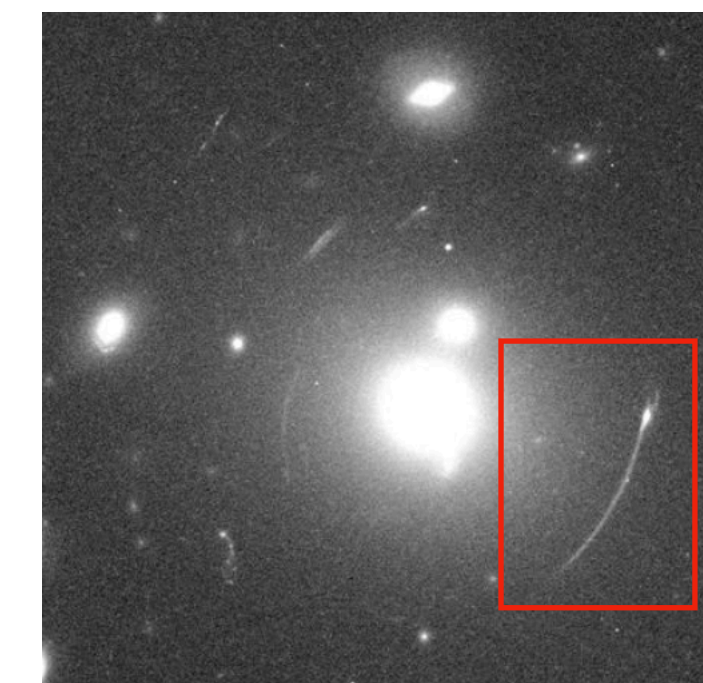
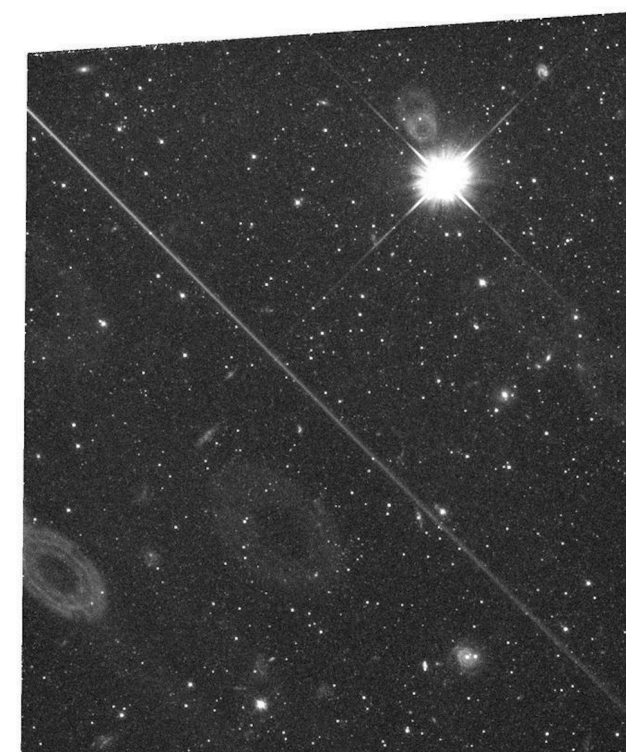
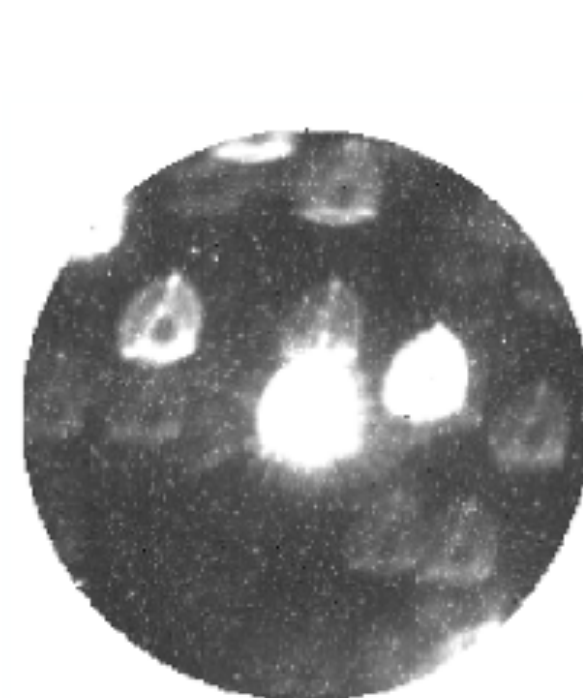
Artificial Satellites



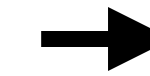
Solar system objects



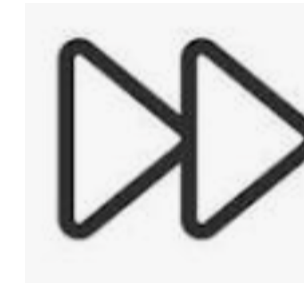
Other artifacts / sources



Not in one of
those classes

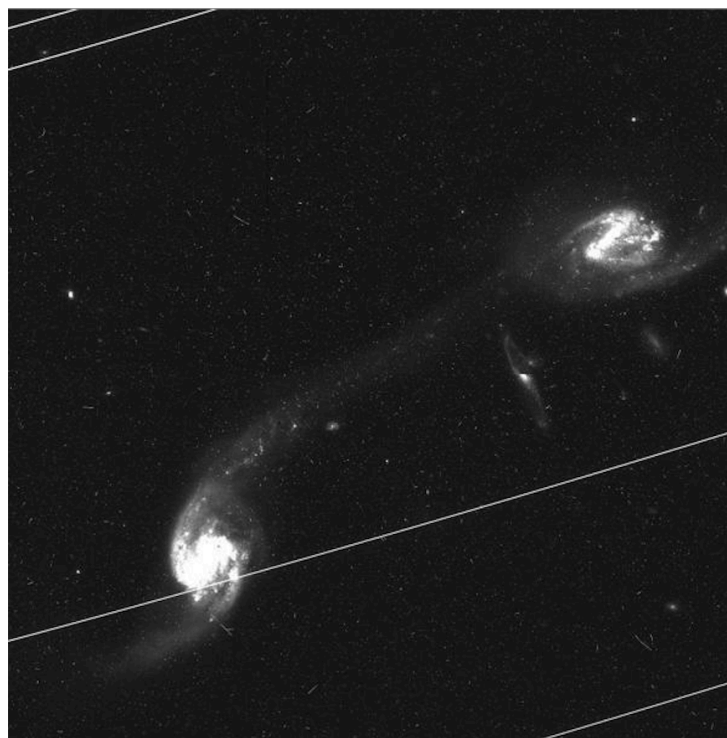
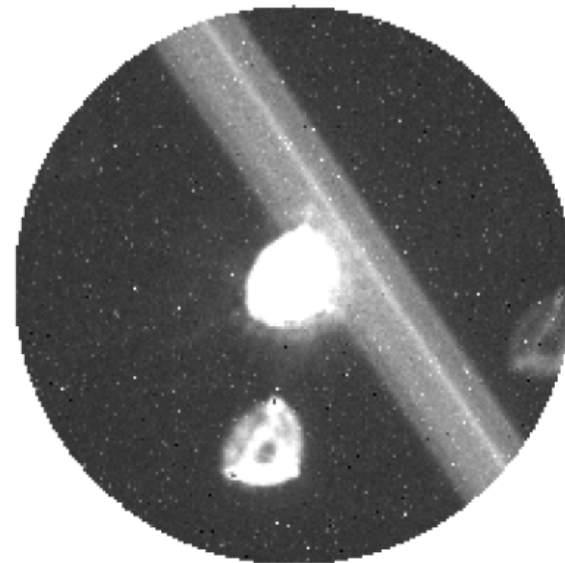
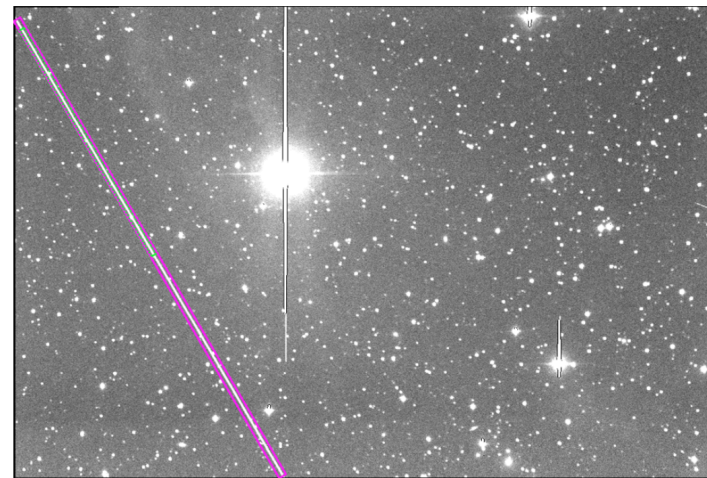


Extrasolar
Science

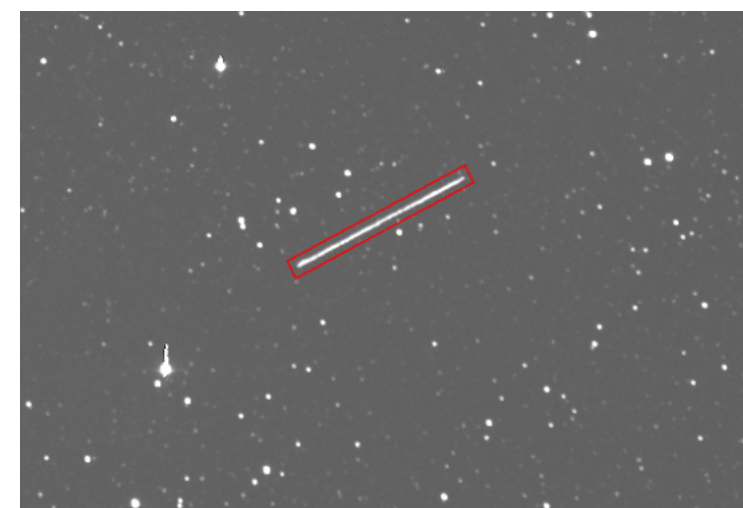


The plan

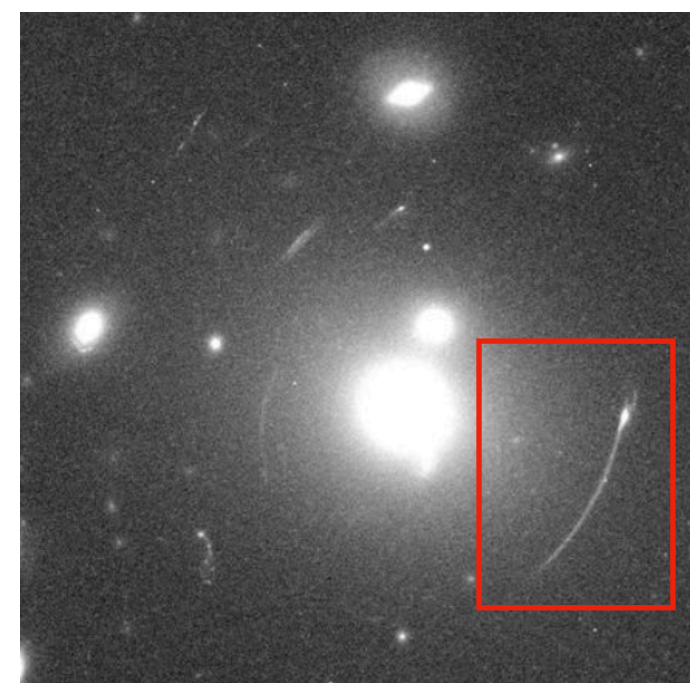
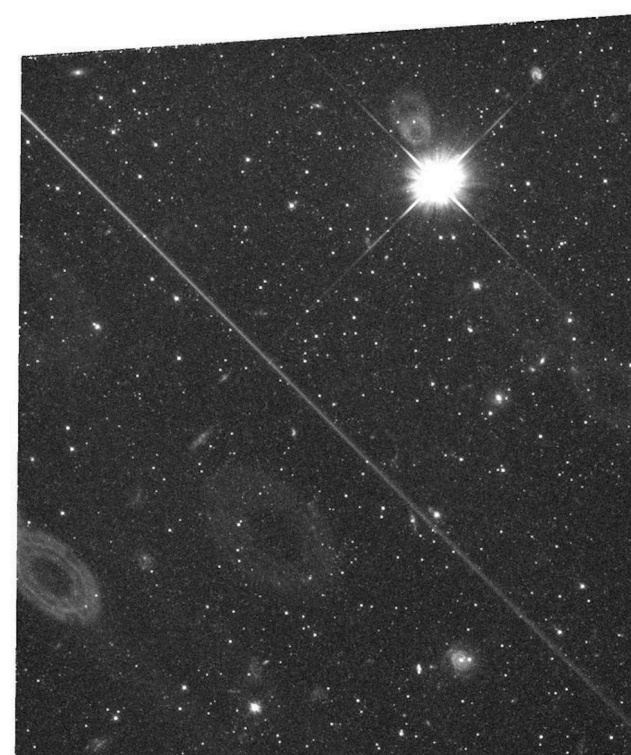
Artificial Satellites



Solar system objects

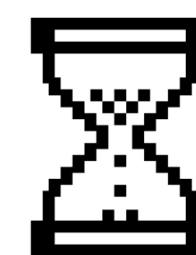


Other artifacts / sources



- Hopefully not a problem for a L2 orbit

- Available data? Level 1 data? Downloaded data...

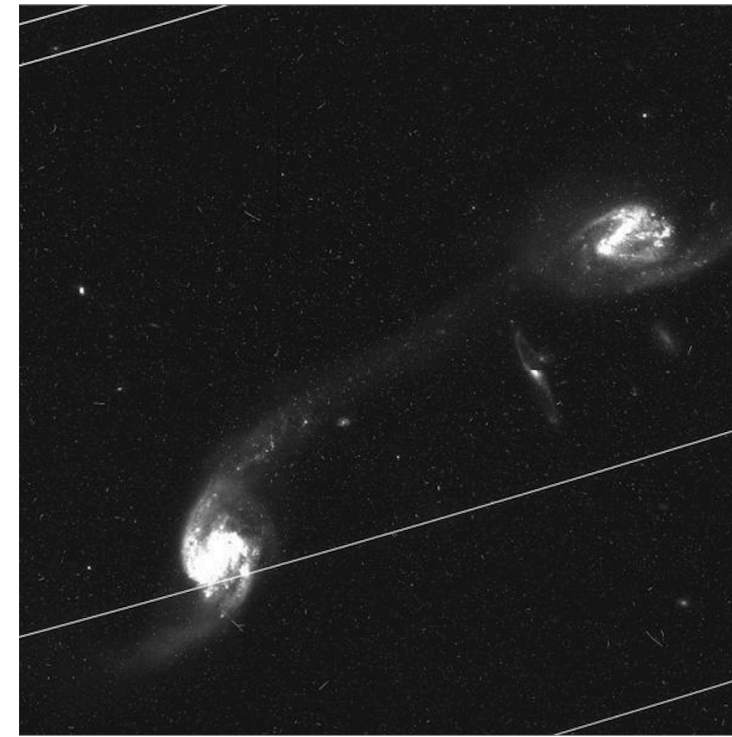
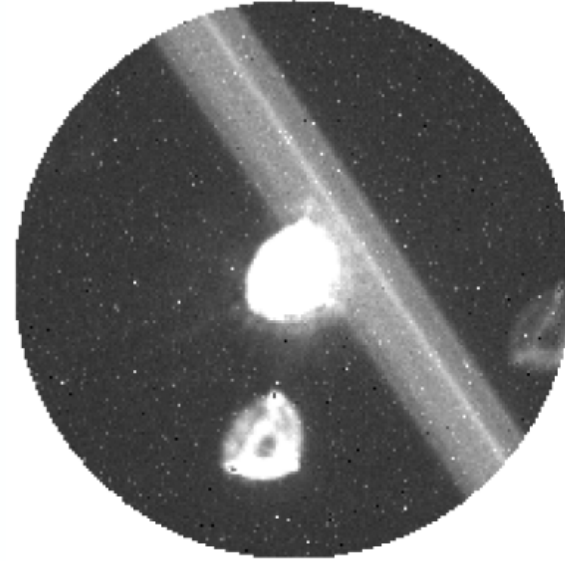
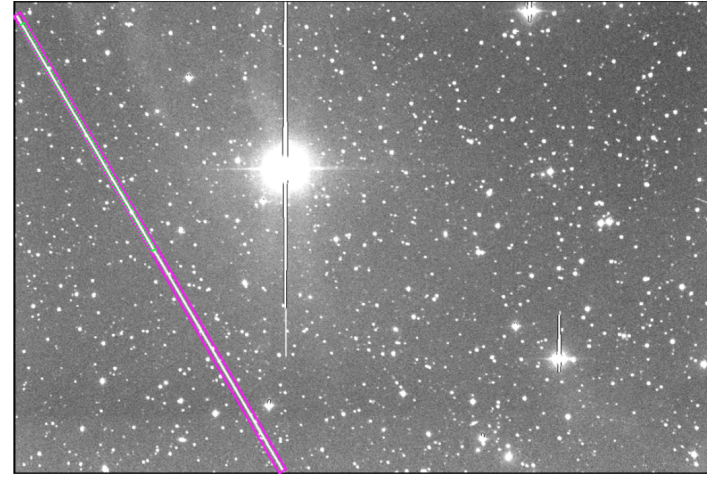


Wanting to know more!

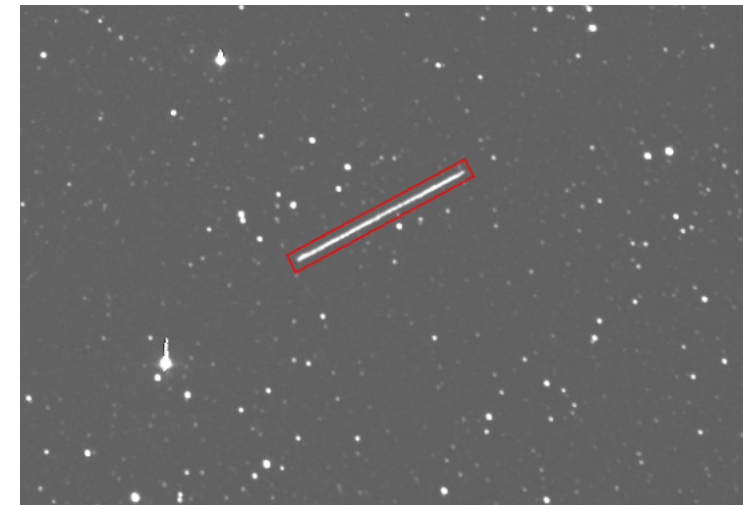
- Yes, of course

The plan

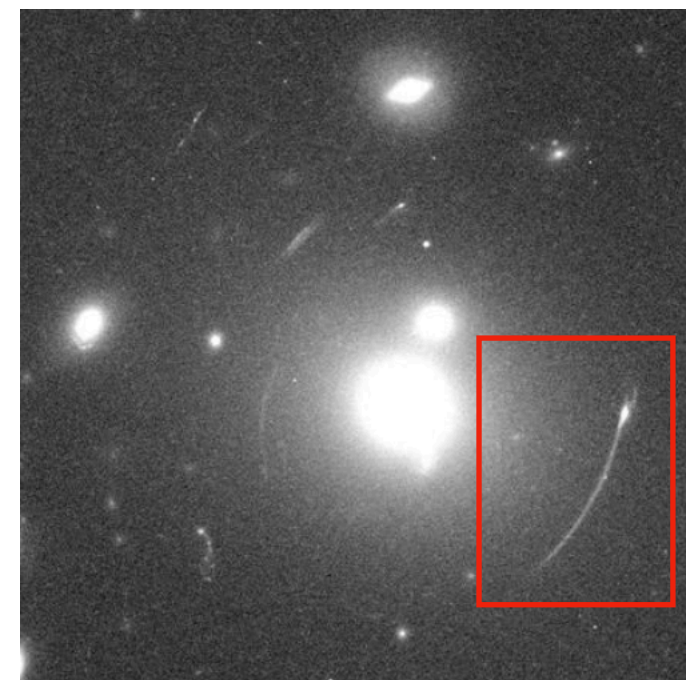
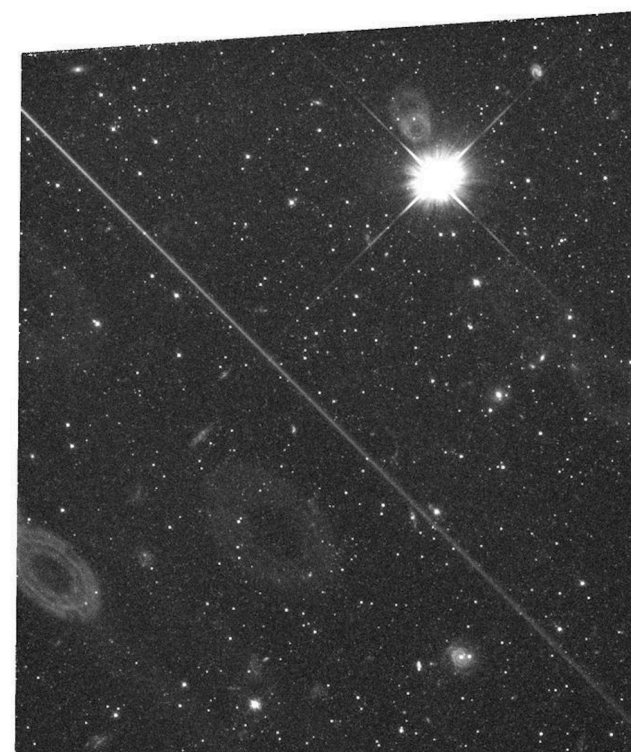
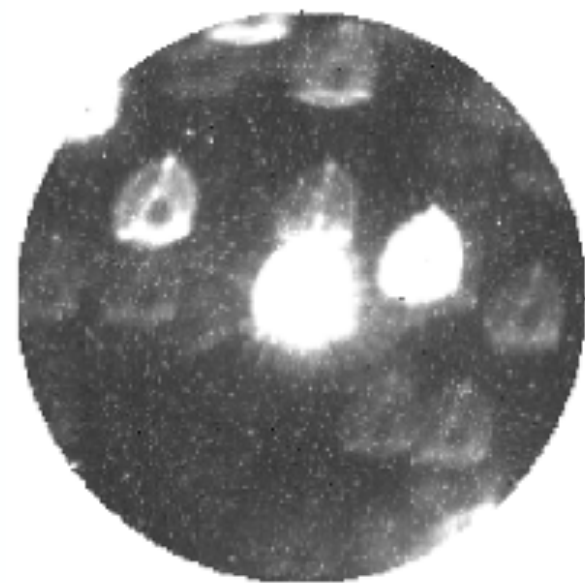
Artificial Satellites



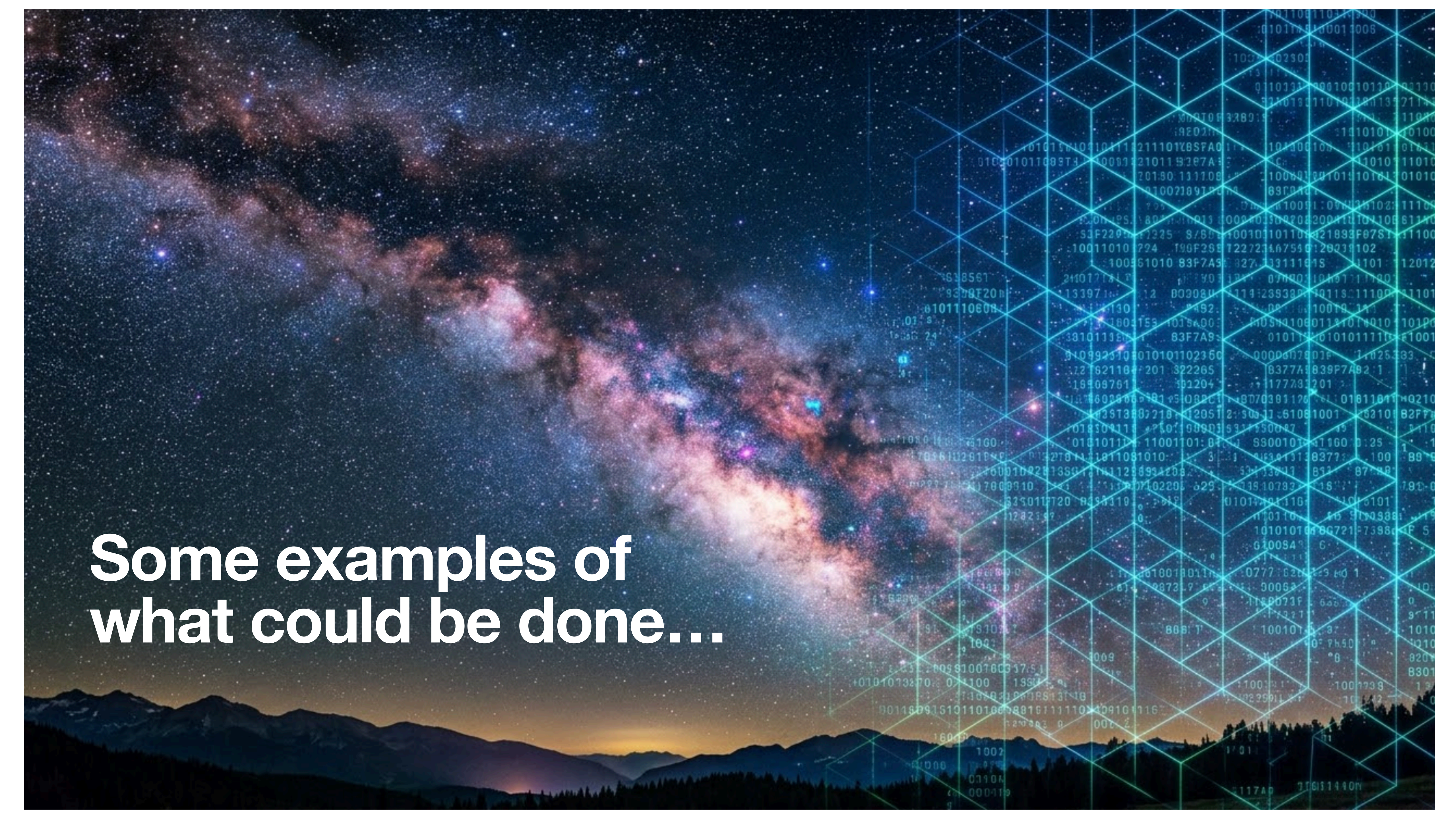
Solar system objects



Other artifacts / sources



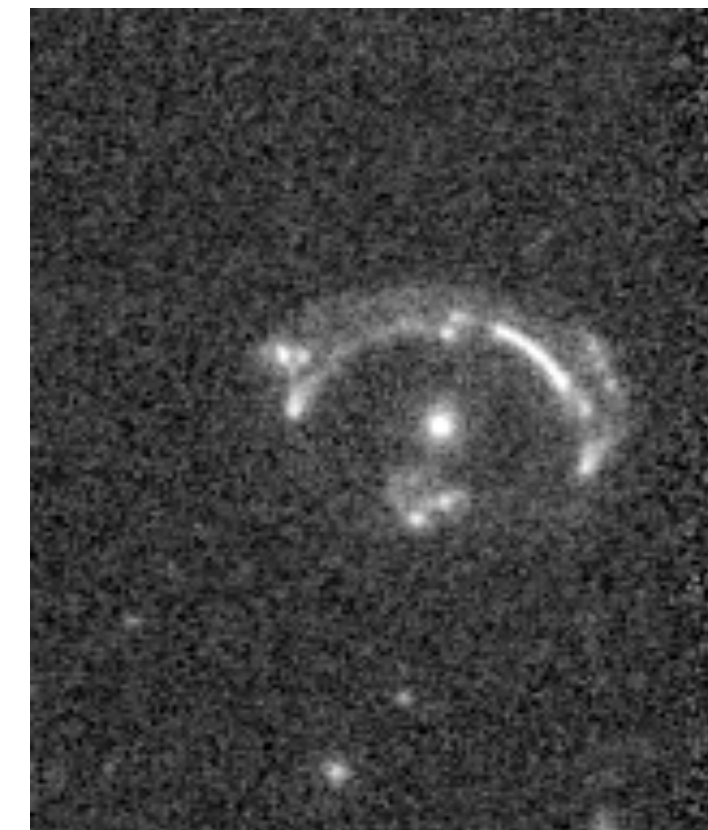
- Powerful AI models exist, mostly off-the-shelf for these applications.
- How can we obtain training sets to make them work properly?
- Mostly supervised learning so far...



**Some examples of
what could be done...**

Space Based - HST

- Google Cloud native Vertex AI model trained for object detection (called AutoML vision back in the day).
- 4 classes: SSOs, satellites, strong gravitational lenses, cosmic rays.
- Analysed 40k archival images, ACS/WFC & WFC3/UVIS
- **1,701 SSOs, 198 strong lenses, 3,228 sats.**
- Precision of 78.3% and a recall of 61.1% on all labels, (IoU 0.5)



Strong Grav.
Lenses



SSOs

Sats



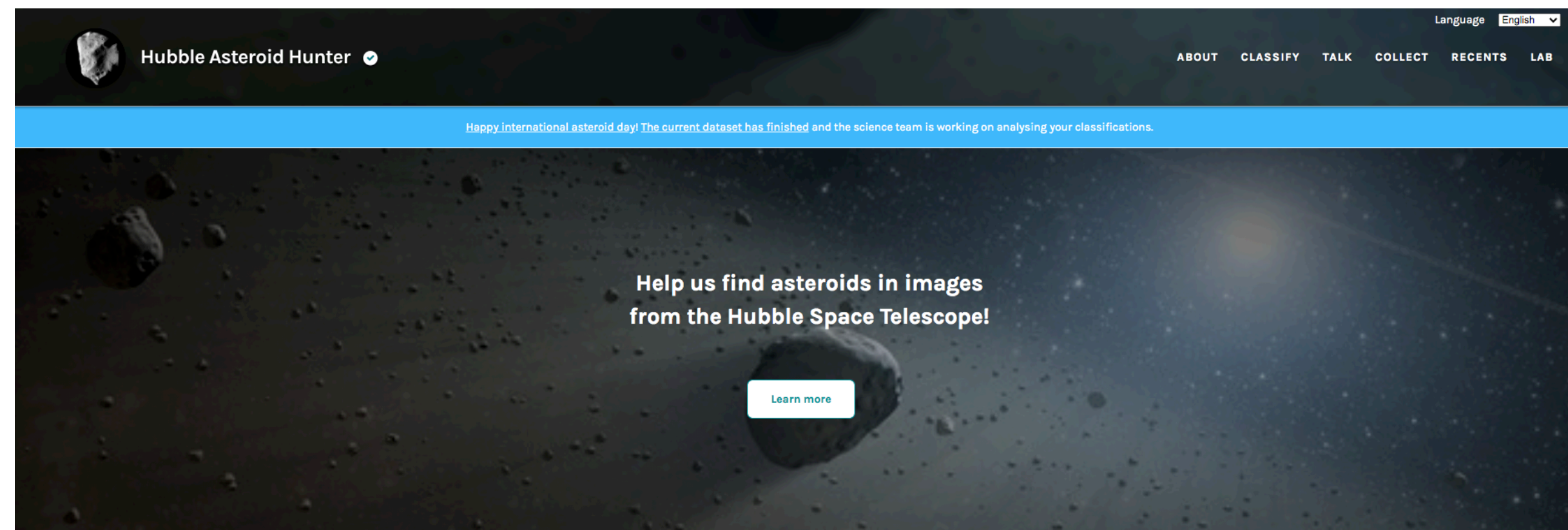
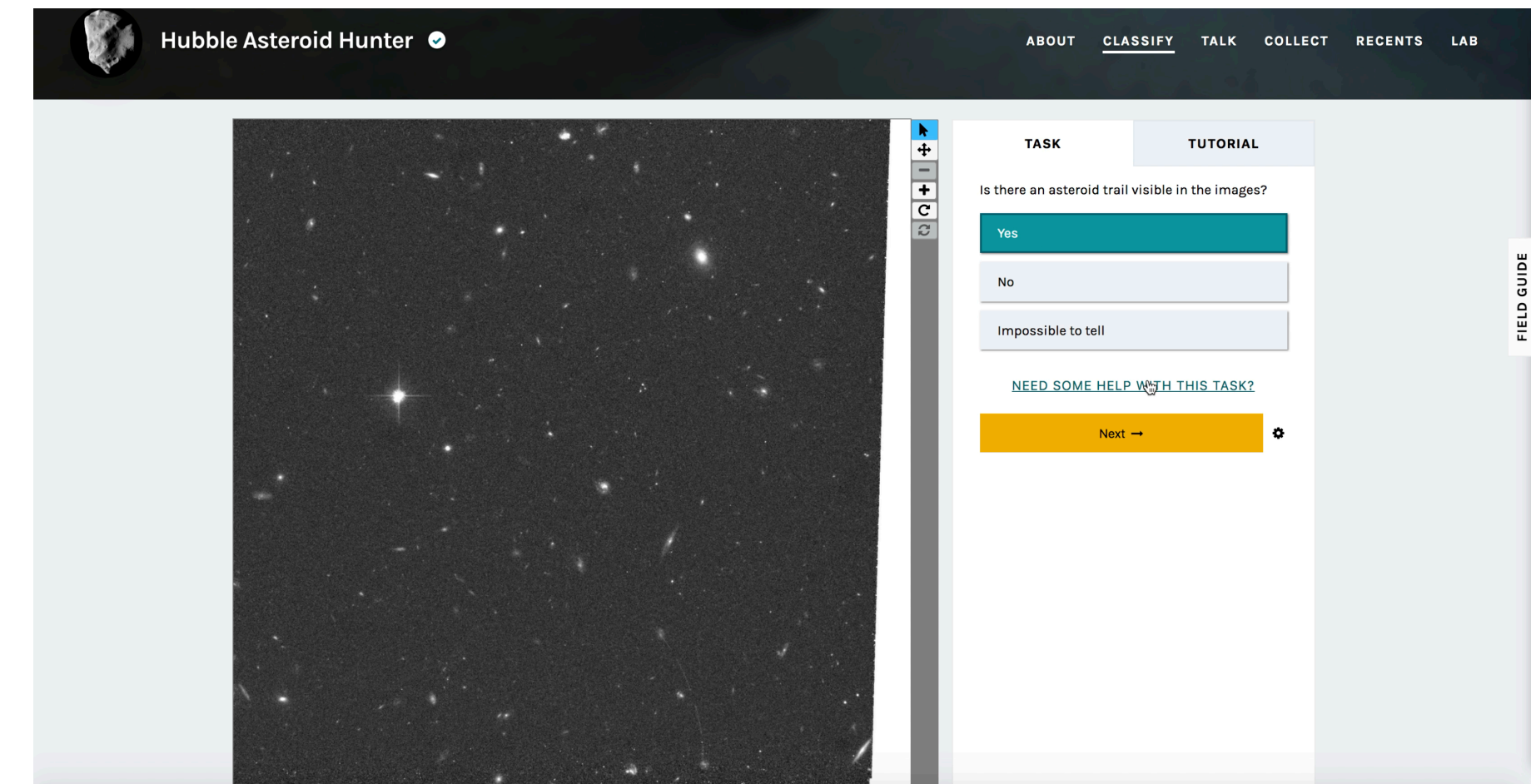
Kruk, S., García-Martín, P., Popescu, M., Aussel, B., Dillmann, S., Perks, M. E., Lund, T., Merín, B., Thomson, R., Karadağ, S., & McCaughrean, M. J. (2023). The impact of satellite trails on Hubble Space Telescope observations. *Nature Astronomy*, 7, 262–268.

Kruk, S., García Martín, P., Popescu, M., Merín, B., Mahlke, M., Carry, B., Thomson, R., Karadağ, S., Durán, J., Racero, E., Giordano, F., Baines, D., de Marchi, G., & Laureijs, R. (2022). Hubble Asteroid Hunter. I. Identifying asteroid trails in Hubble Space Telescope images. *Astronomy & Astrophysics*, 661, Article A85

Garvin, E. O., Kruk, S., Cornen, C., Bhatwadekar, R., Cañameras, R., & Merín, B. (2022). Hubble Asteroid Hunter. II. Identifying strong gravitational lenses in HST images with crowdsourcing. *Astronomy & Astrophysics*, 667, Article A141

Space Based - HST

- The training set: classic, human labelling
 - Citizen Science project: Hubble Asteroid Hunter (Zooniverse)
- 11,482 volunteers provided nearly 2 million classifications



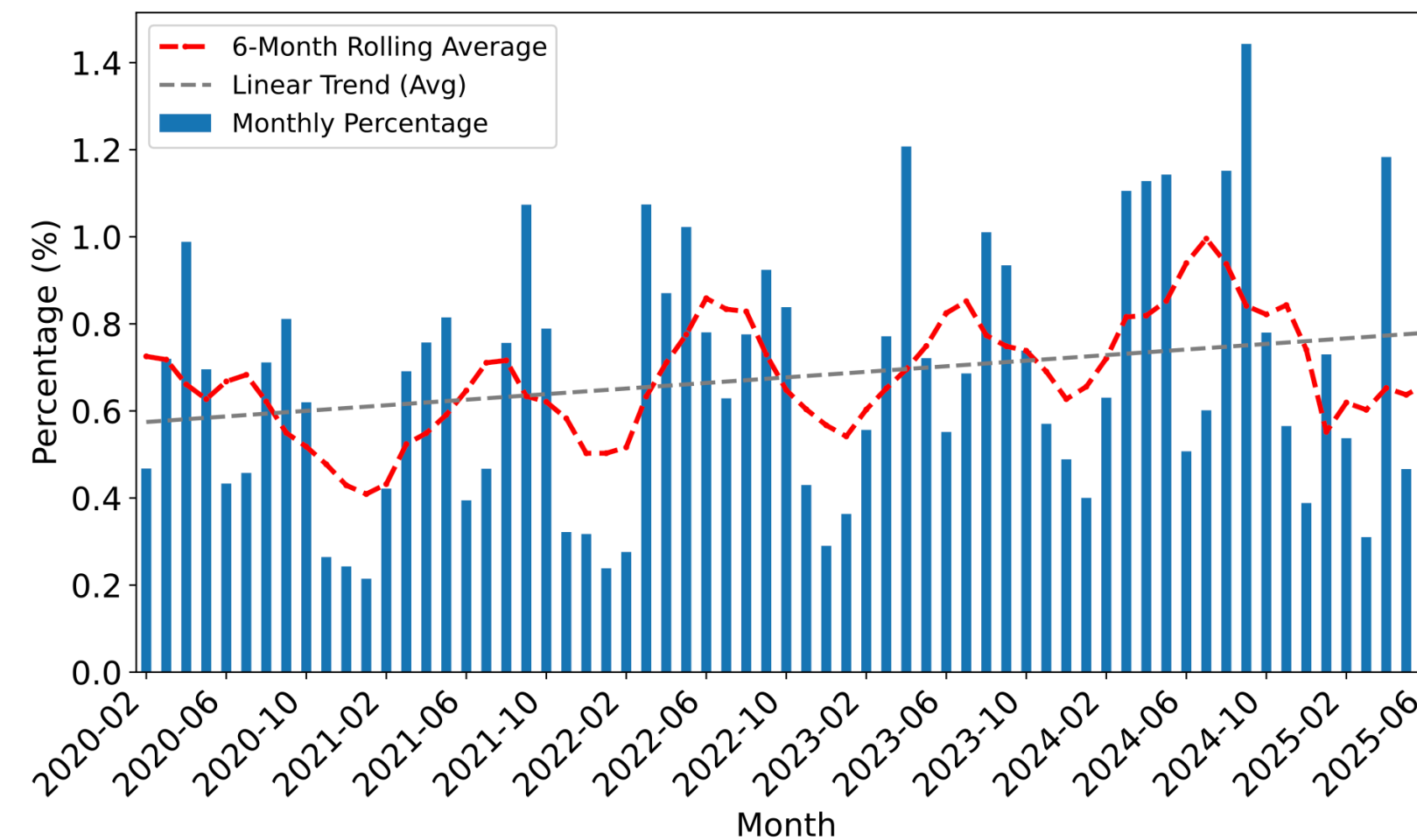
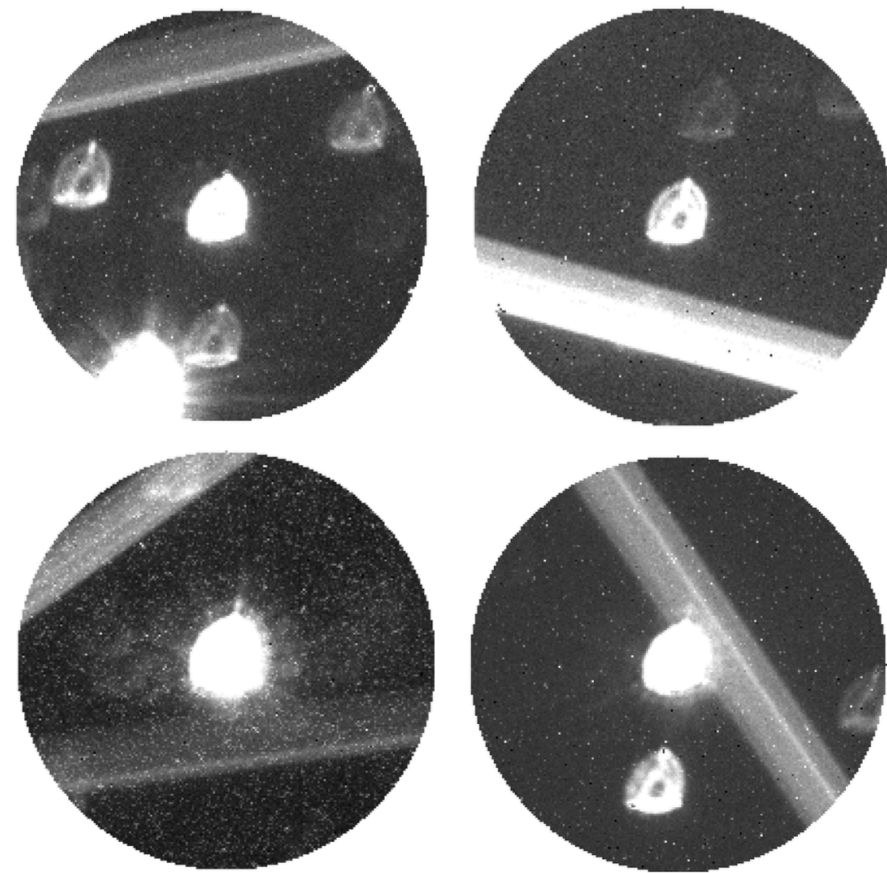
Kruk, S., García-Martín, P., Popescu, M., Aussel, B., Dillmann, S., Perks, M. E., Lund, T., Merín, B., Thomson, R., Karadag, S., & McCaughrean, M. J. (2023). The impact of satellite trails on Hubble Space Telescope observations. *Nature Astronomy*, 7, 262–268.

Kruk, S., García Martín, P., Popescu, M., Merín, B., Mahlke, M., Carry, B., Thomson, R., Karadağ, S., Durán, J., Racero, E., Giordano, F., Baines, D., de Marchi, G., & Laureijs, R. (2022). Hubble Asteroid Hunter. I. Identifying asteroid trails in Hubble Space Telescope images. *Astronomy & Astrophysics*, 661, Article A85

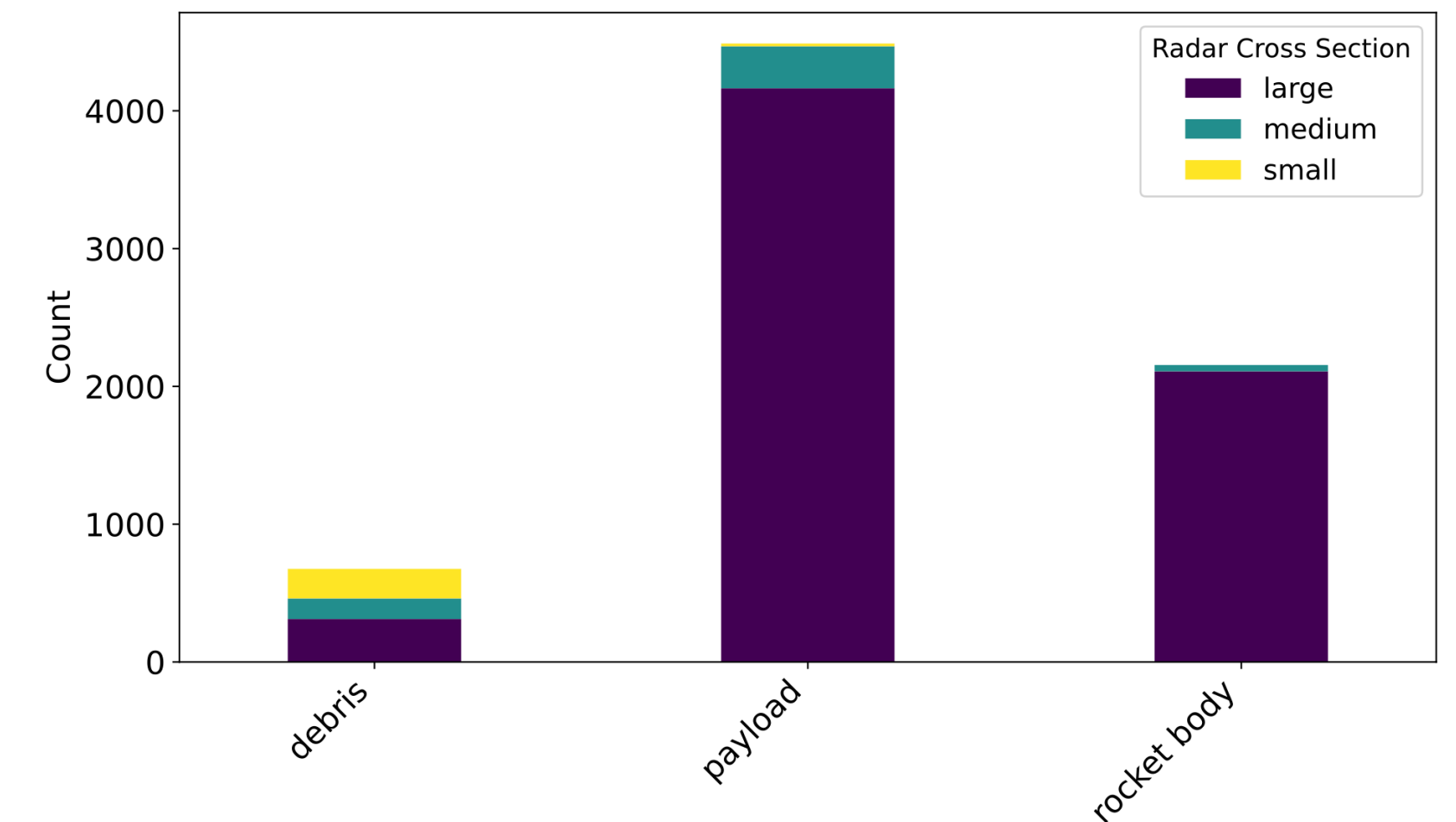
Garvin, E. O., Kruk, S., Cornen, C., Bhatwdekar, R., Cañameras, R., & Merín, B. (2022). Hubble Asteroid Hunter. II. Identifying strong gravitational lenses in HST images with crowdsourcing. *Astronomy & Astrophysics*, 667, Article A141

Space Based - CHEOPS

- Binary classifier looking for artificial satellites in CHEOPS' archive
- 99.2% accuracy and 99.7% recall



Matched 5.5k distinct objects with available database entries

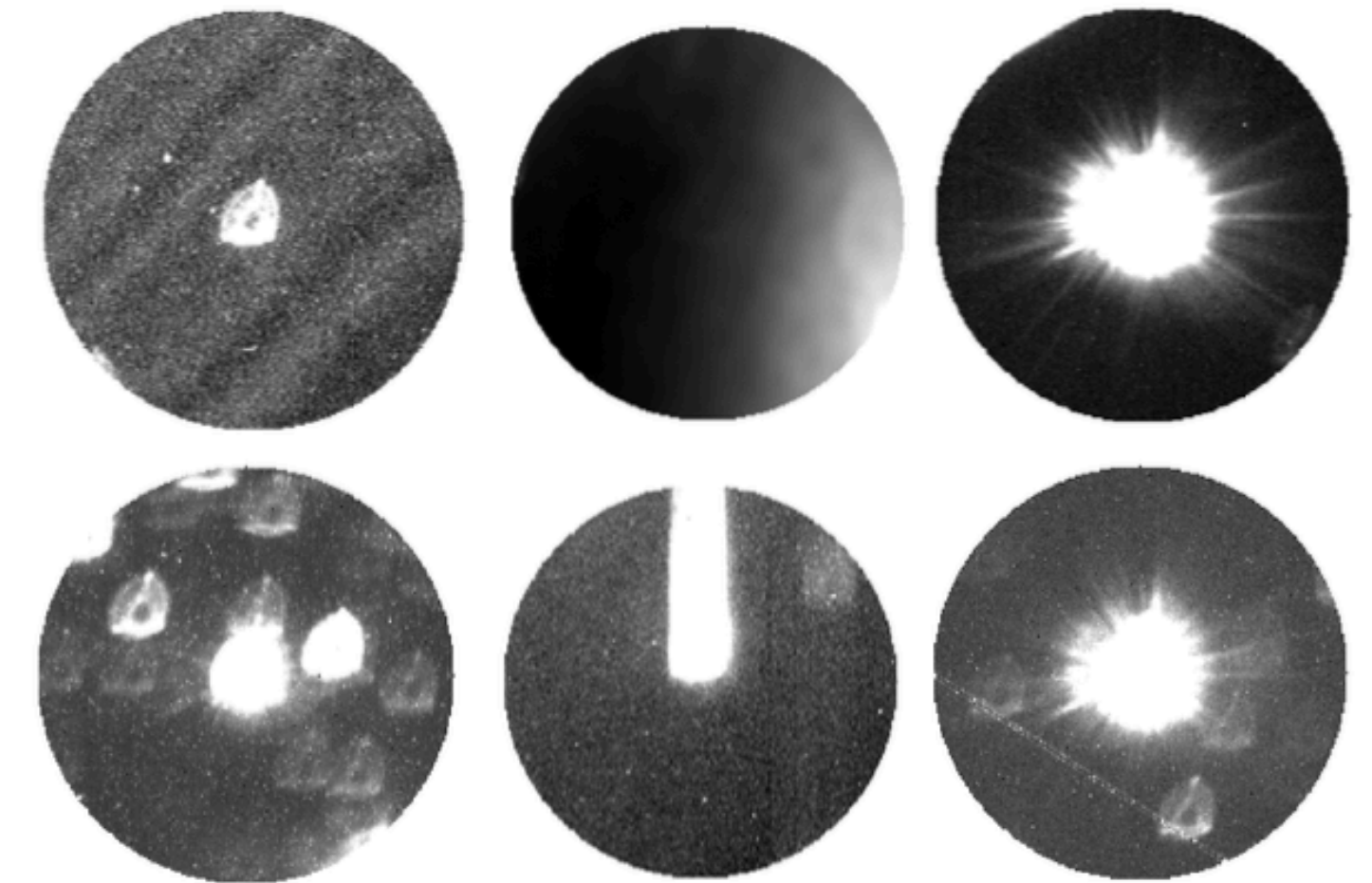


- 1,8 million images scanned, 12k satellite images found
 - $\approx$ x3 more trails vs classical linear feature detection methods.
 - x10 fainter objects detected (counts)

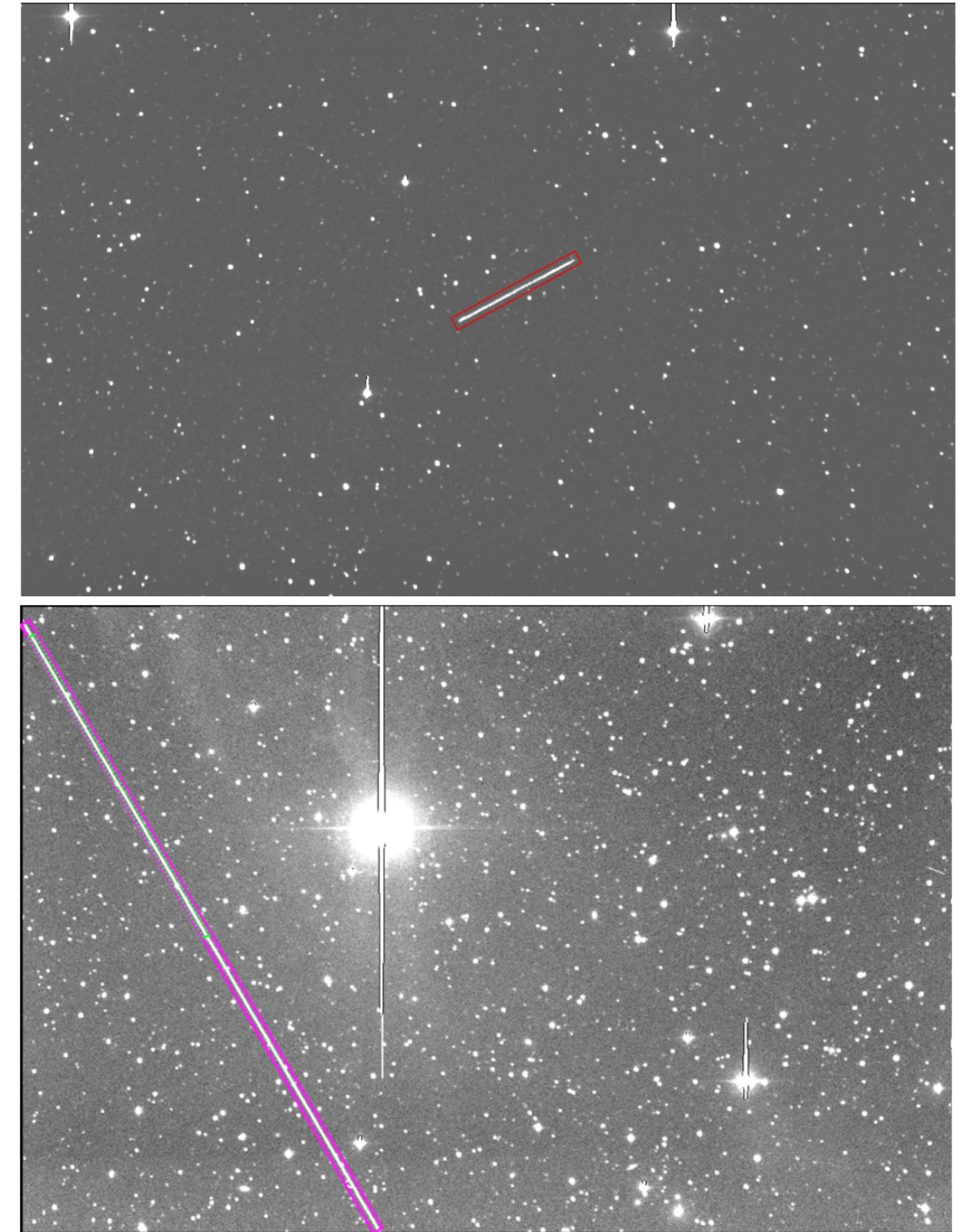
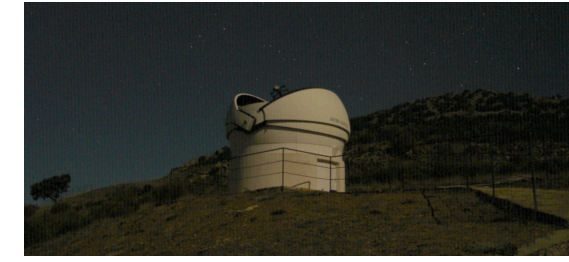
García-Martín, P., Gonçalves, L., Hellmich, S., Billot, N., & Merín, B. (2026). AI for artificial satellites identification in CHEOPS data. *Astronomy & Astrophysics*. In press

Space Based - CHEOPS

- Iteratively enhanced training set - Hard Negative Mining
- Initial Baseline: Kickstarted training with satellite images obtained with legacy classic feature detection methods.
- First Run: Discovered a high false-positive rate because the random negative set lacked rare, complex artifacts.
- Iterative Refinement: Injected half of these manually identified "hard negatives" back into the training set to create a highly robust and balanced final model.



Ground Based - StreakMind



- Prototype algorithm for the detection of **NEOs and artificial satellites**.
 - La Sagra Observatory, Southern Spain
- YOLO-OBB model, achieving a **precision of 94% and a recall of 97%**.
- Instrument-agnostic architecture philosophy. [Our key point here](#)
- Exploring **transfer learning and synthetic data** to easily adapt the model to other telescopes using minimal site-specific data.

Carrillo, R., Duffard, R., García-Martín, P., Romero, J., Morales, N., & Gonçalves, L. (2026). StreakMind: AI detection and analysis of satellites streaks in astronomical images with automated database integration. *Astronomy & Astrophysics*. **In press**

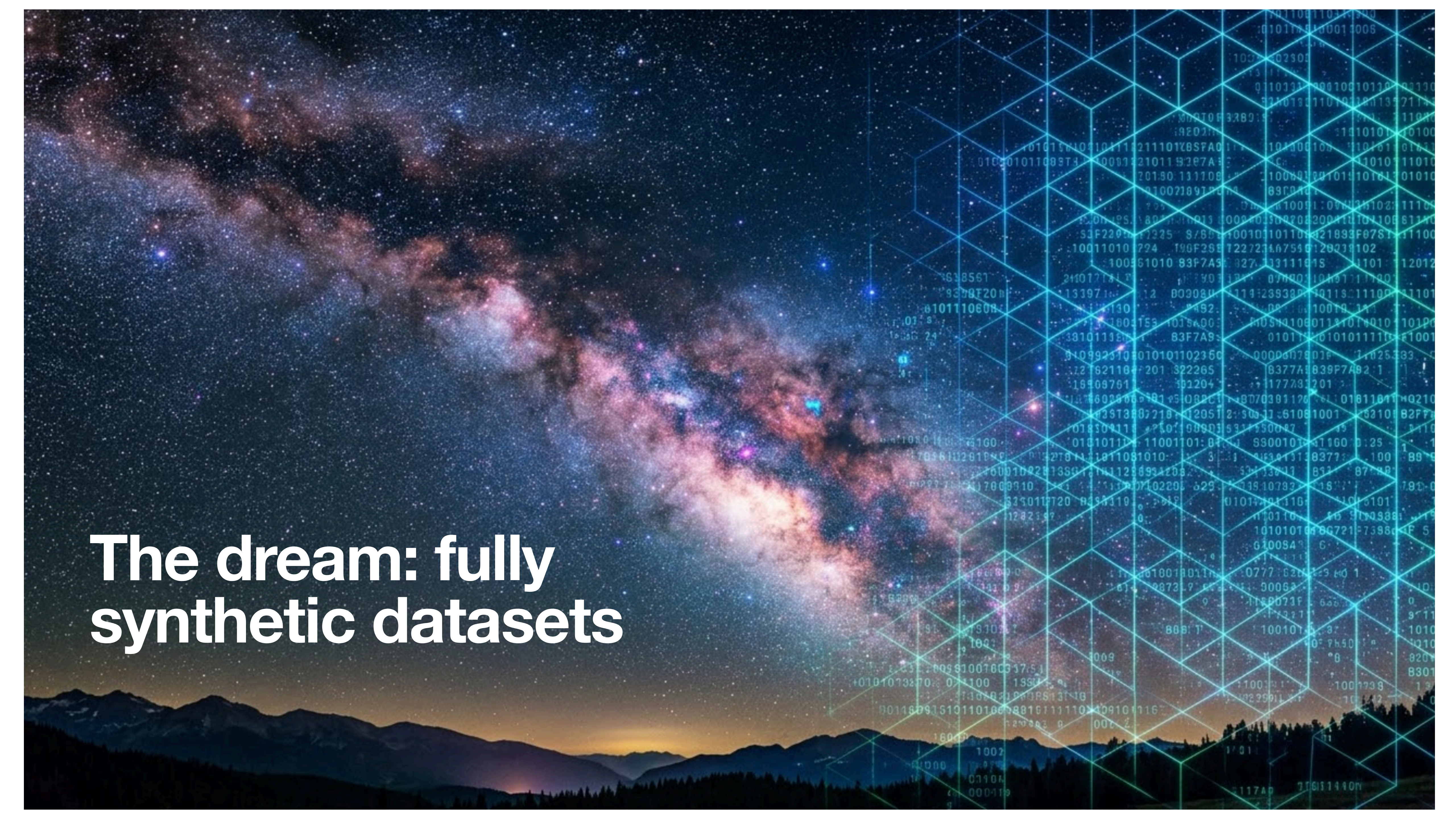
Ground Based - StreakMind



Synthetic trails added to real images

- Hybrid training set: real + synthetic images
- Data Imbalance, the initial manual dataset mostly contained short streaks and lacked long and edge-to-edge trails.
- We injected artificial streaks: randomized lengths, locations, orientations, and five distinct brightness levels were convolved with each image's PSF.





**The dream: fully
synthetic datasets**

Ongoing work - creation of fully synthetic datasets

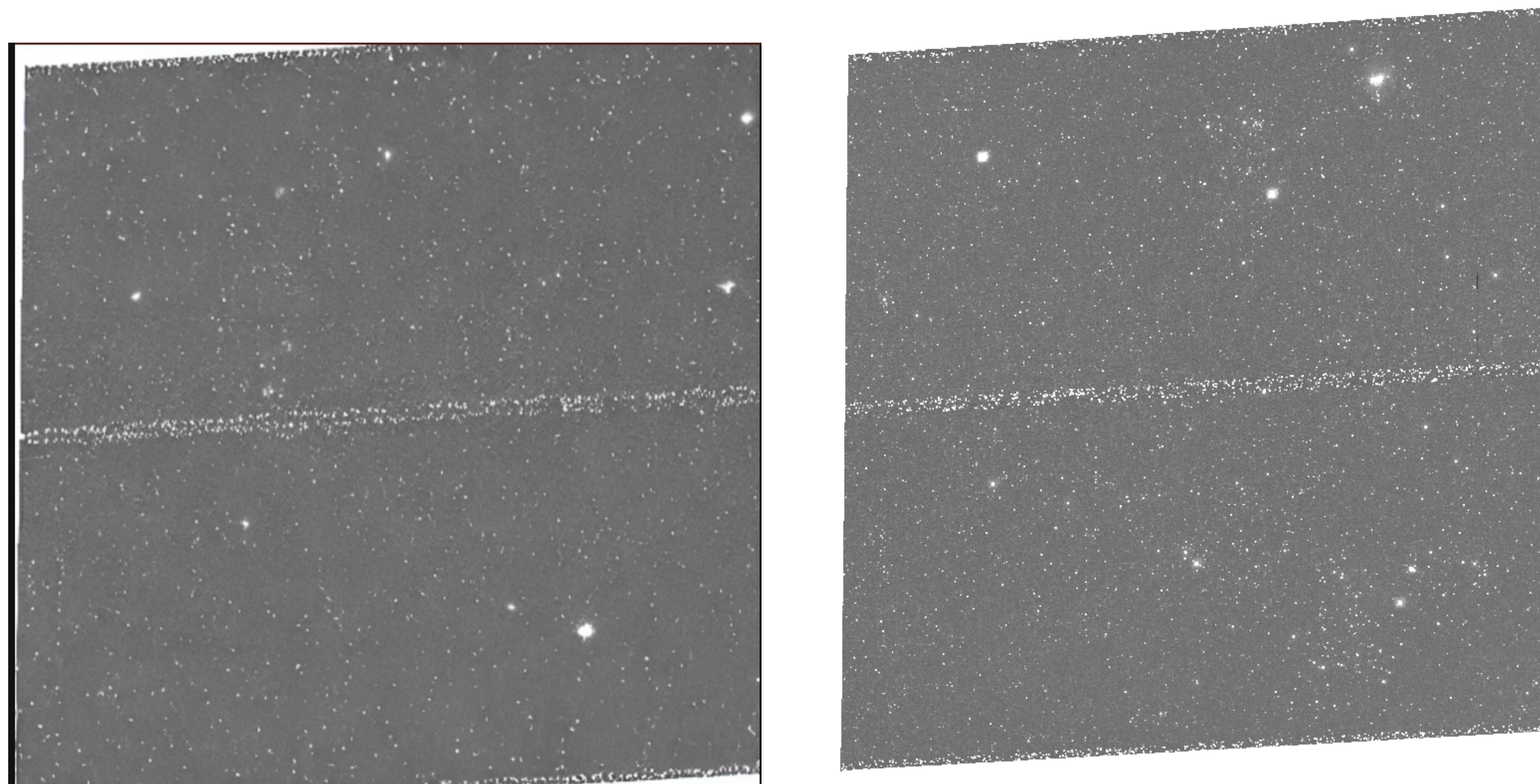
- Currently exploring HST subsets, such as The Panchromatic Hubble Andromeda Treasury.
- $\approx 1.6k$ images, “uniform” M31 dataset
- Target: train a model able to reproduce this “uniform” kind of image.
- Challenges: get enough variance, get enough fidelity.



Timothé Chabot

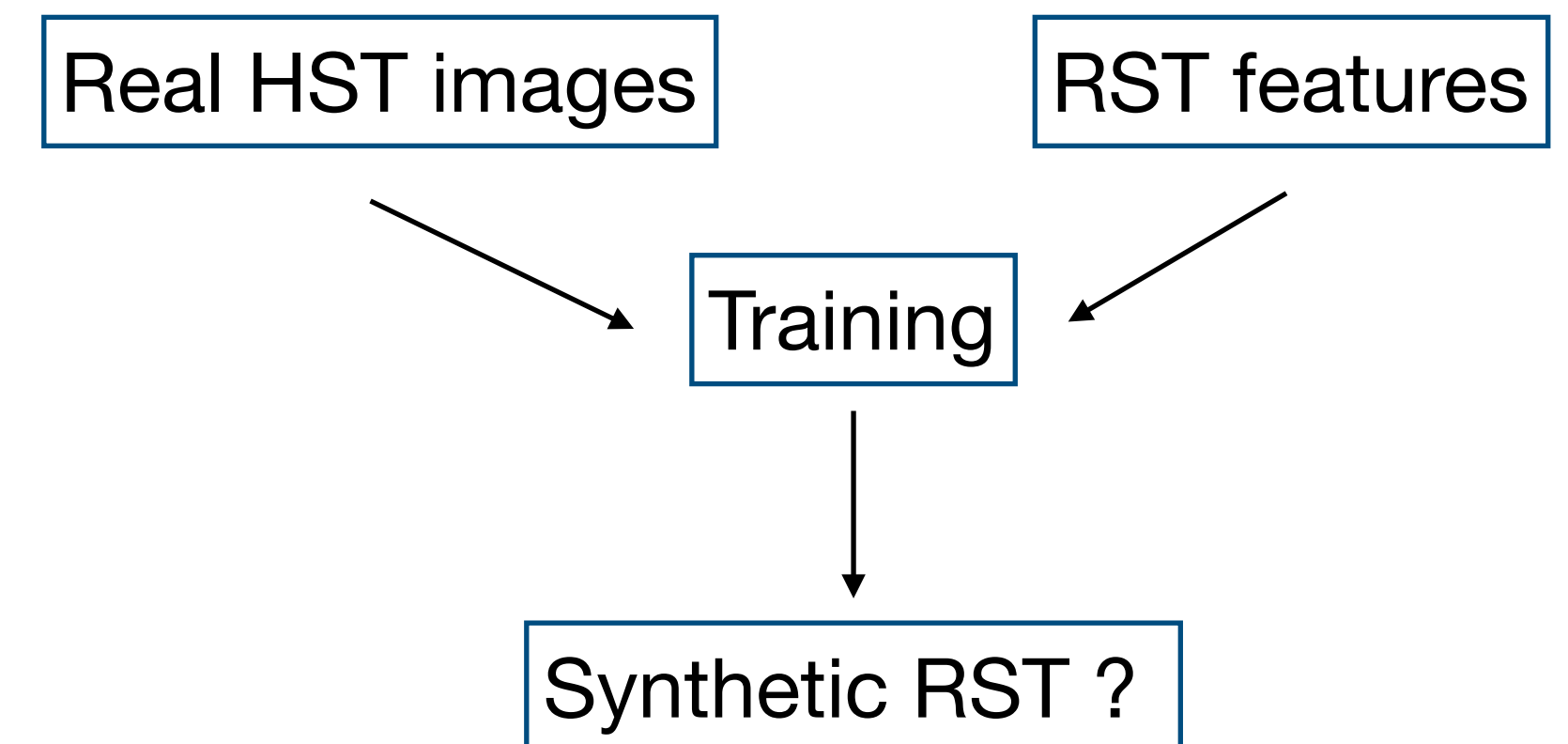
Grad student ESILV Paris

Paris Observatory Intern - Q2-Q3 26



Synthetic

Real HST



- “Technical” proof of concept prototype so far.
- **If you have any interesting Roman Science cases in mind come see me!**

Conclusions

1. Very powerful AI models already exist, most of the time we just need to take care of their training and how we “teach” them.
2. Training with synthetic data could be the way moving forward for avoiding long and tedious manual labelling.
3. Exciting opportunities with Roman & Rapid
4. Come see me to discuss any crazy idea applied to RST data!

