CHRISTOPHER GOETZE, TIM SCHLIPPE & DANIEL LAKEY

DEEP LEARNING-BASED ANOMALY DETECTION IN SPACECRAFT TELEMETRY ON EDGE DEVICES

IU INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

Los Angeles, USA
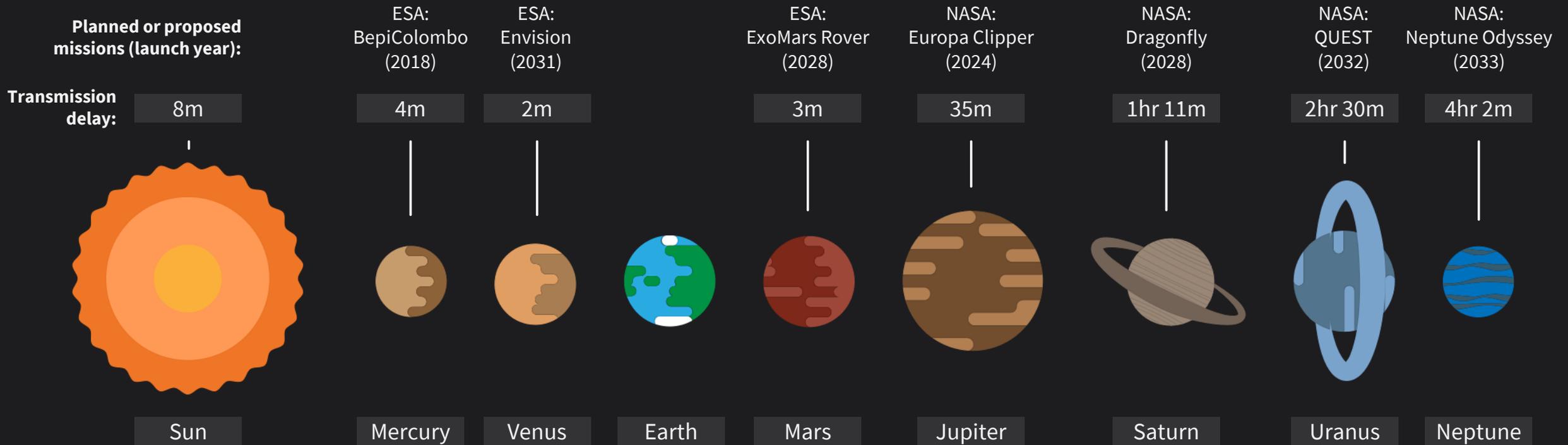July 30, 2025

CONTENT

INTRODUCTION
Challenges

1) **Too much spacecraft telemetry data** (Baireddy et al., 2021)

2) **Few anomalies**
   ➔ Only 200 in over a decade (Lutz & Mikulski, 2004)

3) **Timely critical** (Hundman et al., 2018)

4) **Currently processing on Earth** (Lentaris et al., 2018)
   ➔ Long transmission delays

5) **Limited onboard hardware** (Lentaris et al., 2018)

Crashed Ingenuity Drone. Source: NASA/JPL (2024)

One-way Transmission Delays

Planned or proposed missions (launch year):

	Sun	ESA: BepiColombo (2018)	ESA: Envision (2031)	Earth	ESA: ExoMars Rover (2028)	NASA: Europa Clipper (2024)	NASA: Dragonfly (2028)	NASA: QUEST (2032)	NASA: Neptune Odyssey (2033)
Transmission delay:	8m	4m	2m		3m	35m	1hr 11m	2hr 30m	4hr 2m
	Sun	Mercury	Venus	Earth	Mars	Jupiter	Saturn	Uranus	Neptune

Sources: Arlot (2018); Neo_Artemis (2022)

RELATED WORK
Anomaly Detection Algorithms: *Forecasting & Threshold*
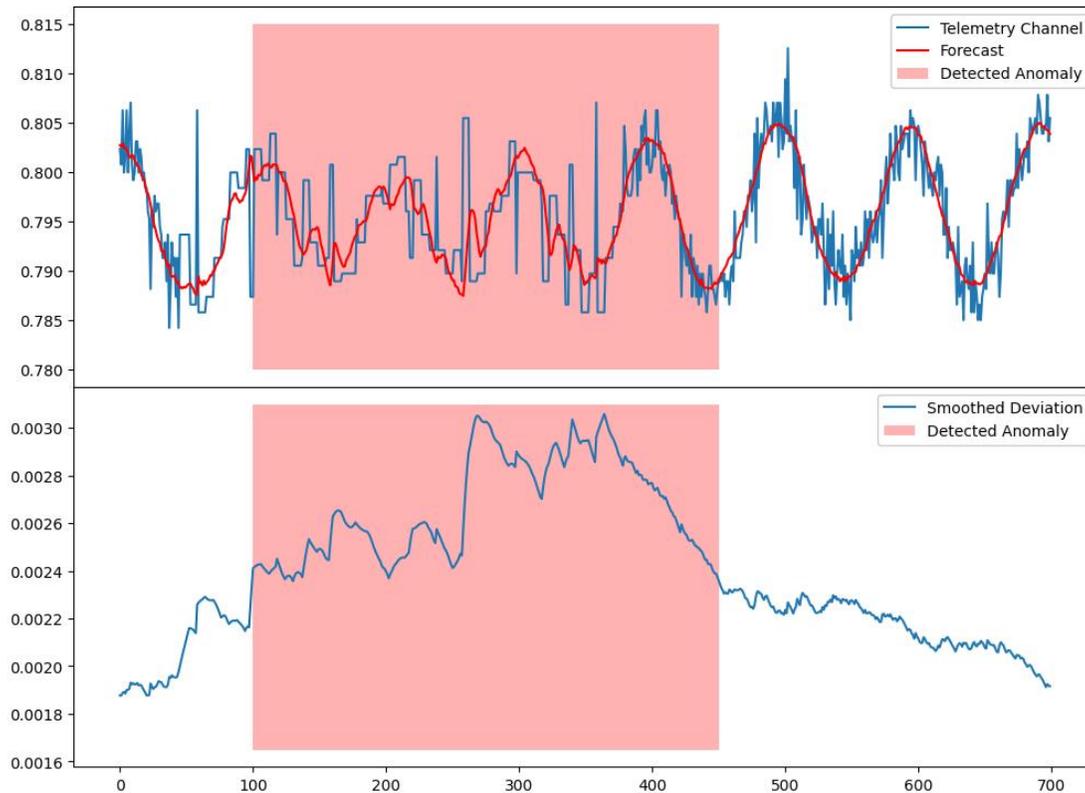
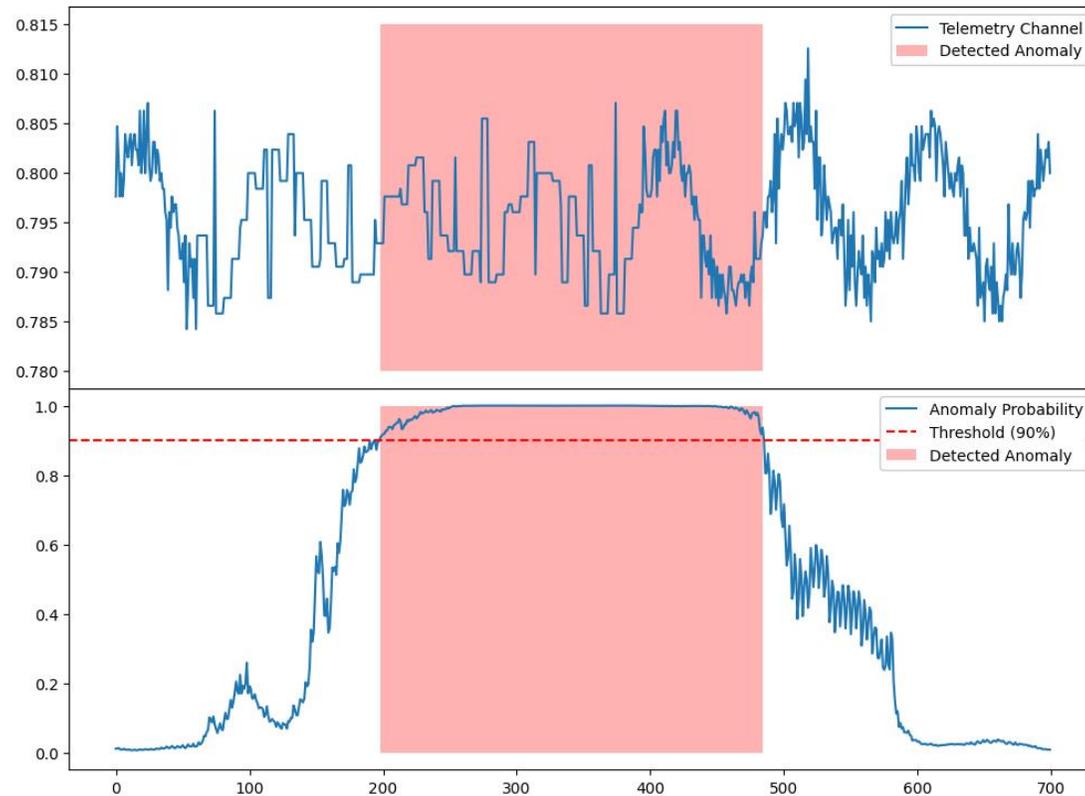Source: Own representation

1) **Forecast value**

2) **Determine deviation**

3) **Classify based on a threshold**

➔ **most common anomaly detection approach**

(E.g., Heras & Donati, 2014; Hundman et al., 2018; Baireddy et al., 2021; Wang et al., 2022)

Anomaly Detection Algorithms: *Direct Classification*

Source: Own representation

- **Classify each time window into anomalous/normal directly**

  (Lakey & Schlippe, 2024a)

Anomaly Detection Algorithms: *Image Classification*

Source: Own representation

1) **Encode time series into images**

2) **Apply image classification**

(Wang & Oates, 2014; Lakey & Schlippe, 2024a)

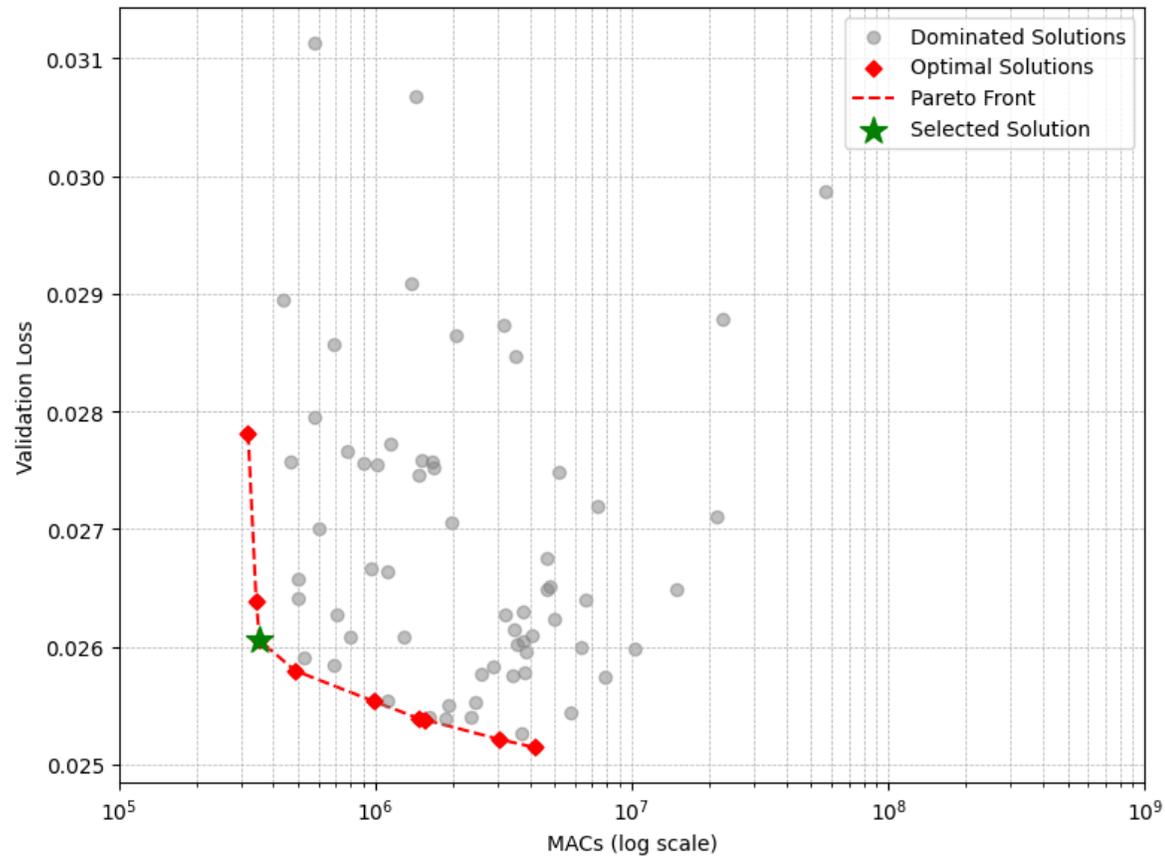## Multi-Objective Optimization & Hardware Constraints

Pareto-optimal solutions (Jin & Sendhoff, 2008)

Source: Own representation

Typical space hardware constraints:

1) **CubeSat**
   - 200 MHz CPU
   - 16 MB RAM
   - 64 MB ROM

2) **OPS-SAT**
   - 800 MHz CPU
   - 1 GB RAM
   - 8 GB ROM

(Evans et al., 2017; SatCatalog, 2021; Zeif et al., 2022)

RELATED WORK
Multi-Objective Optimization for Edge Devices: NAS vs. NAO

Neural Architecture Search (NAS)

- Exploration of entirely new architectures
- Computationally expensive
  - ➜ E.g., one study required 23 GPU-days

    (Liberis et al., 2020)

(E.g., Lu et al., 2019; Tan et al., 2019; Mei et al., 2020; Cardoso-Pereira et al., 2021; Garavagno et al., 2024; King et al., 2025)

Neural Architecture Optimization (NAO)

- Optimization of proven architectures
  - ➜ XceptionTime (Lakey & Schlippe, 2024b)
- Converges more quickly
  - ➜ Requires less computational time

RESEARCH GAPS

1) *Direct classification* and *image classification* **not fully analyzed**
   ➔ Only one Publication (Lakey & Schlippe, 2024a)

2) **No work on bringing these models to edge devices**
   E.g., Ingenuity drone (Snapdragon 801)

3) **Dataset challenges** (Wu & Keogh, 2023)

4) **Metrics not consistent** (Boniol et al., 2022; Sørbø & Ruocco, 2023)

3D model of the Ingenuity drone. Source: NASA (2021)

OUR CONTRIBUTIONS

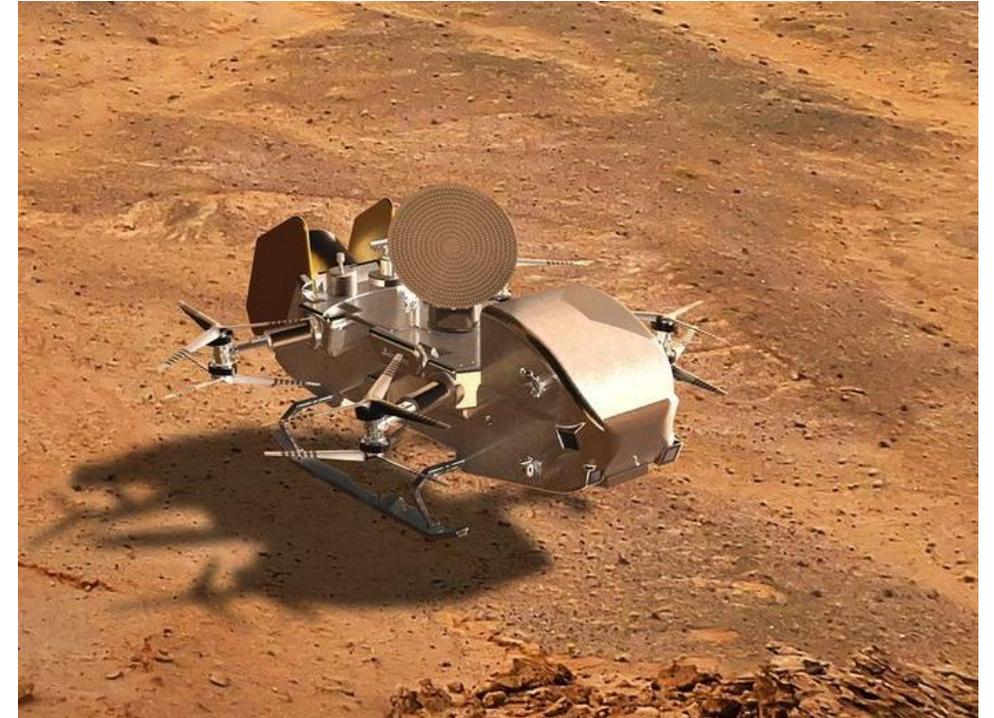1) **Investigating 3 anomaly detection approaches:**
   → *Forecasting & Threshold*
   → *Direct Classification*
   → *Image Classification*

2) **Enabling inference on edge devices:**
   → E.g. on Dragonfly's space-grade hardware

3) **Applying models to a new corpus:**
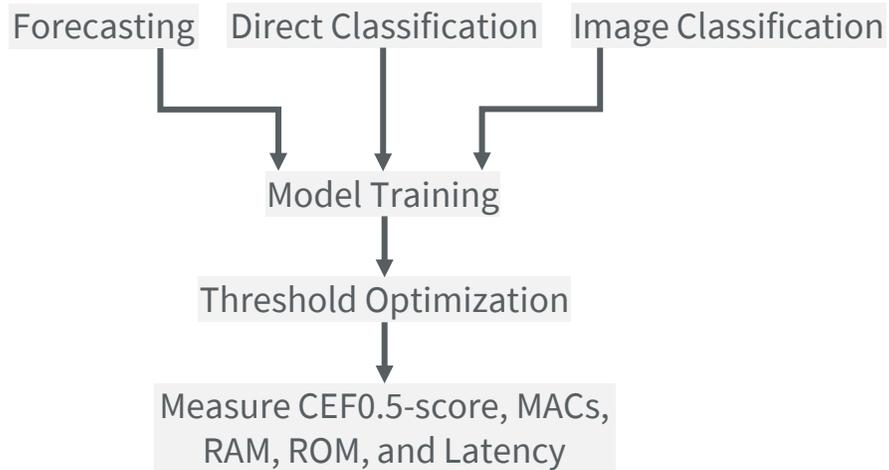   → ESA-AD (Kotowski et al., 2024)

Dragonfly drone. Source: Gribben (2025)

EXPERIMENTAL SETUP
Overview

Baseline Experiments

Forecasting   Direct Classification   Image Classification

Model Training

Threshold Optimization

Measure CEF0.5-score, MACs, RAM, ROM, and Latency

Edge Optimization Experiments

Forecasting   Direct Classification   Image Classification

Multi-Objective Optimization

Model Training

Measure MACs & Validation Loss

Compute Pareto-optimal Solution

Threshold Optimization

Measure CEF0.5-score, RAM, ROM, and Latency
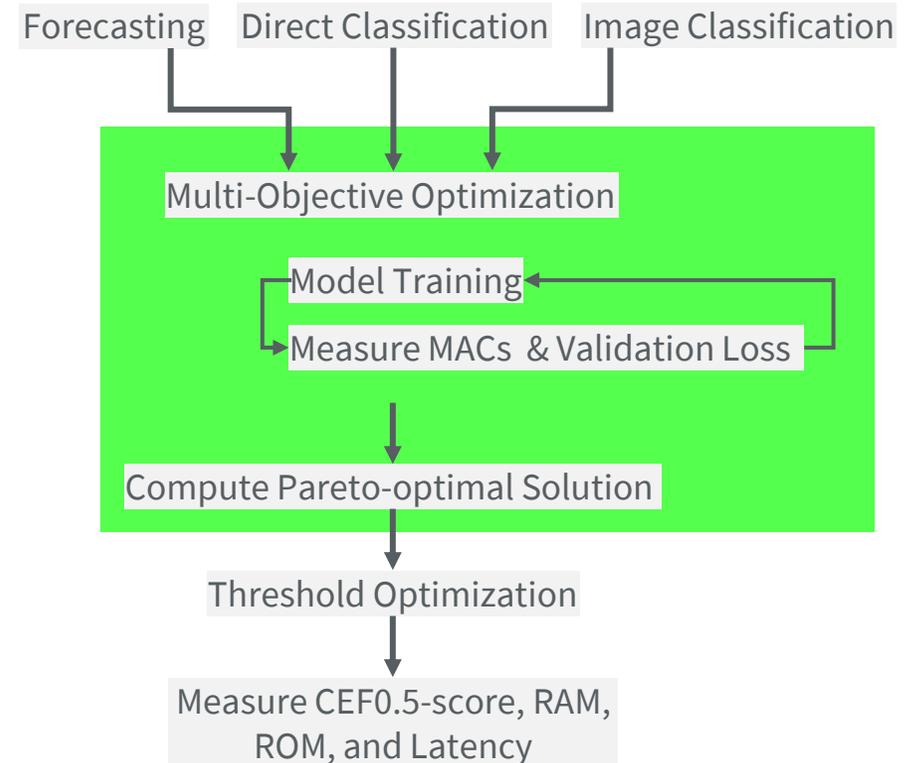
EXPERIMENTAL SETUP
Metrics

FN **TP**

Anomaly

FP

Normal ———————————— **True Labels**

Anomaly

Normal ———————————— **Predicted Labels**

False Alarm **Missed Detection** **Correct Detection**

Challenges:

1) Anomaly clusters
2) Detection timing
3) Balance between FPs and FNs
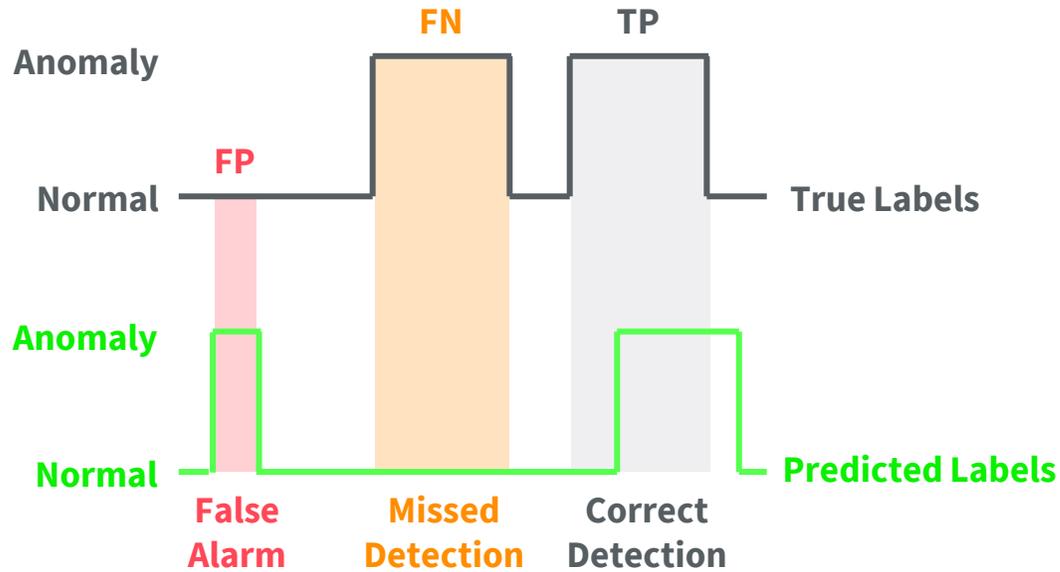
Corrected Event-wise F0.5-score (CEF0.5)

$$\text{Precision} = \frac{TP}{(TP + FP)}$$ → What percentage of the predictions are actual anomalies?

$$\text{Recall} = \frac{TP}{(TP + FN)}$$ → What percentage of the actual anomalies did we detect?

$$\text{CEF0.5} = \frac{1.25 \times Prec \times Rec}{0.25 \times Prec + Rec}$$ → **Balance between Precision and Recall with focus on Precision.**

Based on (Kotowski et al., 2024)

EXPERIMENTAL SETUP
Dataset

ESA Anomalies Dataset

Mission1 Mission2 Mission3
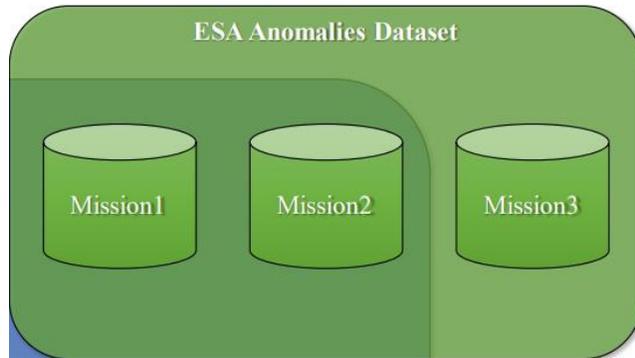
Shortcomings of previous corpora

1) Too small

2) Simulated data

3) Unrealistic amount of anomalies

Benefits of ESA-AD

1) Large scale

2) Real-world anomalies

3) Annotated by experts

➔ We use: Mission1 Lightweight Subset

RESULTS
Anomaly Detection Performance

Deep Learning-Based Anomaly Detection in Spacecraft Telemetry on Edge Devices

RESULTS
MACs Reduction

MMACs (log scale)

	Forecasting & Threshold	Direct Classification	Image Classification
Baseline	56.7	56.5	959.3
Optimized	0.4	0.9	45.0

■ Baseline ■ Optimized
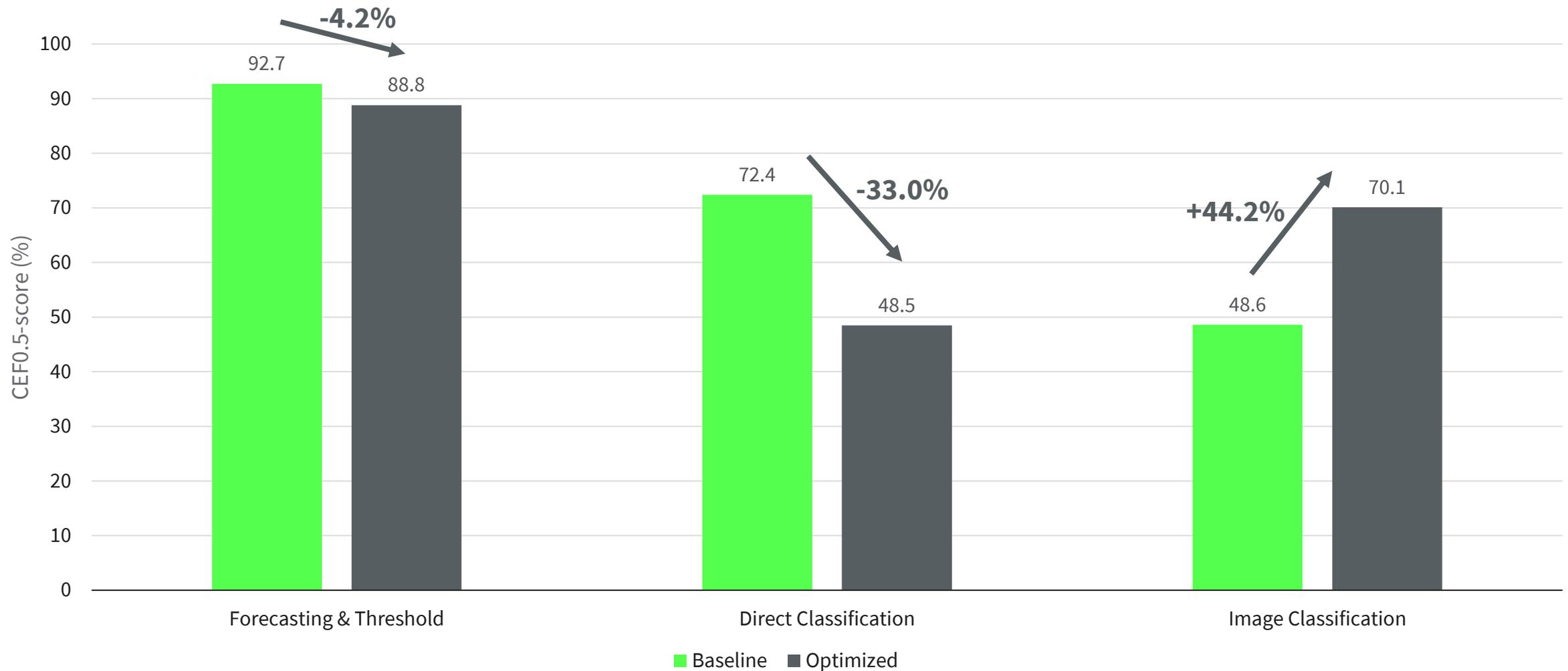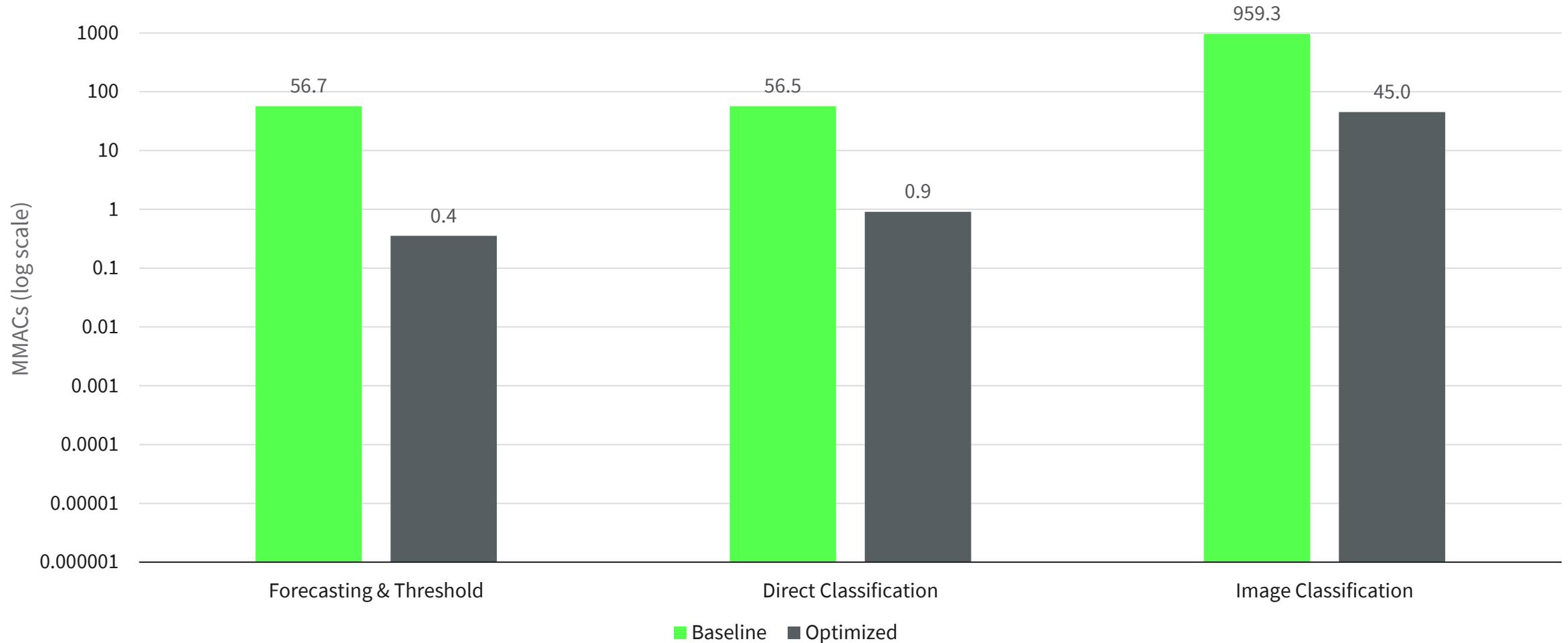
CONCLUSION AND FUTURE WORK
Conclusion

We investigated deep learning models that are competitive in anomaly detection performance while also being small enough for onboard processing. (CubeSat: 16 MB RAM, 64 MB ROM)

- *Forecasting & threshold*:
  - Best CEF0.5-score (88.8%)
  - Good resource efficiency (59 KB RAM , 166 KB ROM)
  - Slow thresholding

- *Direct classification*:
  - Less accurate (48.5% CEF0.5-score)
  - Good resource efficiency (122 KB RAM , 149 KB ROM)
  - Fast thresholding

- *Image classification*:
  - Competitive CEF0.5-score (70.1%)
  - Uses much more resources (1,024 KB RAM, 508 KB ROM)
  - Fast thresholding

CONCLUSION AND FUTURE WORK
Future Work

- **Combine multi-objective optimization with other model compression methods, e.g.:**
  - Quantization (Jacob et al., 2017)
  - Pruning (Frankle & Carbin, 2019)
  - Knowledge distillation (Hinton et al., 2015)

- **Combine multiple anomaly detection approaches as an ensemble model.**
  - E.g., majority voting

- **Scale the evaluation from Mission1 to all three missions of the ESA-AD.**

- **Further investigation of *image classification*.**

THANK YOU!

Christopher Goetze / Tim Schlippe / Daniel Lakey

christopher.goetze@iu-study.org / tim.schlippe@iu.org / daniel.lakey@iu-study.org

REFERENCES

Arlot, J.-E. (2018). The Distances From The Planets To The Earth. IMCCE, Paris Observatory. https://promenade.imcce.fr/en/pages5/582.html

Baireddy, S., Desai, S. R., Mathieson, J. L., Foster, R. H., Chan, M. W., Comer, M. L., & Delp, E. J. (2021). Spacecraft time-series anomaly detection using transfer learning. In 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pages 1951–1960. DOI: 10.1109/CVPRW53098.2021.00223.

Boniol, P., Paparrizos, J., Kang, Y., Palpanas, T., Tsay, R. S., Elmore, A. J., & Franklin, M. J. (2022). Theseus: Navigating the Labyrinth of Time-Series Anomaly Detection. Proceedings of the VLDB Endowment, 15(12), 3702–3705. https://doi.org/10.14778/3554821.3554879

Cardoso-Pereira, I., Lobo-Pappa, G., & Ramos, H. S. (2021). Neural Architecture Search for Resource-Constrained Internet of Things Devices. In 2021 IEEE Symposium on Computers and Communications (ISCC) (pp. 1–6). IEEE. https://doi.org/10.1109/ISCC53001.2021.9631535

ESA (2018). Mission BepiColombo. https://esoc.esa.int/content/bepicolombo#:~:text=BepiColombo%20is%20Europe%27s%20first%20mission,planet%20in%20our%20Solar%20System.

ESA (2025). ExoMars. https://www.esa.int/Science_Exploration/Human_and_Robotic_Exploration/Exploration/ExoMars

Evans, D., Kubicka, M., Cherciu, C., & Coelho, C. (2017). ESAW 2017: Current OPS-SAT mission status and opportunities for experimenters. ESA. https://web.archive.org/web/20171223042533/http:/www.esa-esaw2017.eu/sites/default/files/OPS-SAT_Experimenter_Dave_Evans_Final%20.pdf

Frankle, J., & Carbin, M. (2019). The Lottery Ticket Hypothesis: Finding Sparse, Trainable Neural Networks. ICLR 2019. http://arxiv.org/pdf/1803.03635v5

Garavagno, A. M., Ragusa, E., Frisoli, A., & Gastaldo, P. (2024). An Affordable Hardware-Aware Neural Architecture Search for Deploying Convolutional Neural Networks on Ultra-Low-Power Computing Platforms. IEEE Sensors Letters, 8(5), 1–4. https://doi.org/10.1109/LSENS.2024.3387056

Gribben, S. (2025). Artist's concept of NASA's Dragonfly on the surface of Saturn's moon Titan. https://assets.science.nasa.gov/dynamicimage/assets/science/psd/planetary-science-division/2025/dragonfly2.jpg?w=768&h=523&fit=crop&crop=faces%2Cfocalpoint

Heras, J. & Donati, A. (2014). Enhanced telemetry monitoring with novelty detection. AI Magazine, 35(4):37–46. DOI: 10.1609/aimag.v35i4.2553.

Hinton, G., Vinyals, O., & Dean, J. (2015, March 9). Distilling the Knowledge in a Neural Network. http://arxiv.org/pdf/1503.02531v1

Hundman, K., Constantinou, V., Laporte, C., Colwell, I., & Soderstrom, T. (2018). Detecting Spacecraft Anomalies Using LSTMs and Nonparametric Dynamic Thresholding. In Y. Guo & F. Farooq (Eds.), Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 387–395). ACM. https://doi.org/10.1145/3219819.3219845

REFERENCES

Jacob, B., Kligys, S., Chen, B., Zhu, M., Tang, M., Howard, A., Adam, H., & Kalenichenko, D. (2017, December 16). Quantization and Training of Neural Networks for Efficient Integer-Arithmetic-Only Inference. http://arxiv.org/pdf/1712.05877v1

Khoshnevisan, S., Gong, W., Wang, L., & Juang, C. H. (2014). Robust design in geotechnical engineering – an update. Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 8(4), 217–234. https://doi.org/10.1080/17499518.2014.980274

King, T., Zhou, Y., Röddiger, T., & Beigl, M. (2025). Micronas for memory and latency constrained hardware aware neural architecture search in time series classification on microcontrollers. Scientific Reports, 15(1), 7575. https://doi.org/10.1038/s41598-025-90764-z

Kotowski, K., Haskamp, C., Andrzejewski, J., Ruszczak, B., Nalepa, J., Lakey, D., Collins, P., Kolmas, A., Bartesaghi, M., Martínez-Heras, J., & Canio, G. De. (2024). European Space Agency Benchmark for Anomaly Detection in Satellite Telemetry. https://doi.org/10.5281/ZENODO.12528696

Lakey, D., & Schlippe, T. (2024a). Anomaly Detection in Spacecraft Telemetry: Forecasting vs. Classification. 2024 IEEE Space Computing Conference (SCC). Advance online publication. https://doi.org/10.1109/SCC61854.2024.00009

Lakey, D., & Schlippe, T. (2024b). A Comparison of Deep Learning Architectures for Spacecraft Anomaly Detection. 2024 IEEE Aerospace Conference. Advance online publication. https://doi.org/10.1109/AERO58975.2024.10521015

Lentaris, G., Maragos, K., Stratakos, I., Papadopoulos, L., Papanikolaou, O., Soudris, D., Lourakis, M., Zabulis, X., Gonzalez-Arjona, D., & Furano, G. (2018). High-Performance Embedded Computing in Space: Evaluation of Platforms for Vision-Based Navigation. Journal of Aerospace Information Systems, 15(4), 178–192. https://doi.org/10.2514/1.I010555

Liberis, E., Dudziak, Ł., & Lane, N. D. (2020, October 27). μNAS: Constrained Neural Architecture Search for Microcontrollers. http://arxiv.org/pdf/2010.14246v3

Lu, Q., Shi, Y., Jiang, W., Hu, J., & Xu, X. (2019). On Neural Architecture Search for Resource-Constrained Hardware Platforms.

Lutz, R. R., & Mikulski, I. C. (2004). Empirical analysis of safety-critical anomalies during operations. IEEE Transactions on Software Engineering, 30(3), 172–180. https://doi.org/10.1109/TSE.2004.1271171

Mei, J., Li, Y., Lian, X., Jin, X., Yang, L., Yuille, A., & Yang, J. (2020). AtomNAS: Fine-Grained End-to-End Neural Architecture Search. International Conference on Learning Representations, ICLR. https://arxiv.org/abs/1912.09640

Morton, E. (2025). NASA's Dragonfly Passes Critical Design Review. https://science.nasa.gov/blogs/dragonfly/2025/04/24/nasas-dragonfly-passes-critical-design-review/#:~:text=The%20Dragonfly%20mission%20will%20launch,Kennedy%20Space%20Center%20in%20Florida.

REFERENCES

NASA (2021). Interactive 3D model of the Mars Ingenuity Helicopter. https://science.nasa.gov/mission/mars-2020-perseverance/ingenuity-mars-helicopter/

NASA (2024). Europa Clipper Mission Timeline. https://europa.nasa.gov/mission/timeline/#:~:text=Europa%20Clipper%20launches%20from%20Kennedy,14%2C%202024.

NASA/JPL (2024). NASA Performs First Aircraft Accident Investigation on Another World. https://www.jpl.nasa.gov/news/nasa-performs-first-aircraft-accident-investigation-on-another-world/

Neo_Artemis (2022). https://pixabay.com/vectors/space-solar-system-planets-sun-7307385/

SatCatalog. (2021). IMT - CubeSat On-Board Computer - Datasheet. https://satcatalog.s3.amazonaws.com/components/458/SatCatalog_-_IMT_-_CubeSat_On-Board_Computer_-_Datasheet.pdf

Sørbø, S., & Ruocco, M. (2023, March 2). Navigating the Metric Maze: A Taxonomy of Evaluation Metrics for Anomaly Detection in Time Series. http://arxiv.org/pdf/2303.01272v1

Tan, M., Chen, B., Pang, R., Vasudevan, V., Sandler, M., Howard, A., & Le, Q. V. (2019). MnasNet: Platform-Aware Neural Architecture Search for Mobile. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2815–2823. https://doi.ieeecomputersociety.org/10.1109/CVPR.2019.00293

Wang, Y., Gong, J., Zhang, J., & Han, X. (2022). A deep learning anomaly detection framework for satellite telemetry with fake anomalies. International Journal of Aerospace Engineering, 2022:1–9. DOI: 10.1155/2022/1676933.

Wang, Z., & Oates, T. (2014). Encoding Time Series as Images for Visual Inspection and Classification Using Tiled Convolutional Neural Networks.

Wu, R., & Keogh, E. J. (2023). Current Time Series Anomaly Detection Benchmarks are Flawed and are Creating the Illusion of Progress.

Zeif, R., Kubicka, M., & Hörmer, A. J. (2022). Development and application of an embedded computer system for CubeSats exemplified by the OPS-SAT space mission. E & I Elektrotechnik Und Informationstechnik, 139(1), 8–15. https://doi.org/10.1007/s00502-022-00991-9