

JPL Center Innovation Fund

Building a Galaxy-Star Classifier for Roman Space Telescope Alert Streams

JPL Investigators: Umaa Rebbapragada (PI), Mario Damiano
Caltech Investigators: Ashish Mahabal, Mansi Kasliwal

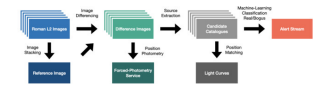
ML Ops for RAPID Alerts

Background: RAPID (Roman Alerts Promptly from Image Differencing) is a Project Infrastructure Team (PIT) for the Nancy Grace Roman Space Telescope designed to enable rapid transient discovery across the Roman Core Community and General Astrophysics Surveys. RAPID performs image differencing by comparing every new Roman image against a reference image, subsequently issuing a prompt, public alert stream of all transient and variable candidates. Embedded within this alert stream are the outputs of machine learning (ML) systems. These provide crucial metadata, such as a score indicating the likelihood that a candidate is a true astronomical source (versus a bogus artifact) and a classification distinguishing stars from galaxies. We are currently developing this star-galaxy classifier within a robust MLOps framework.

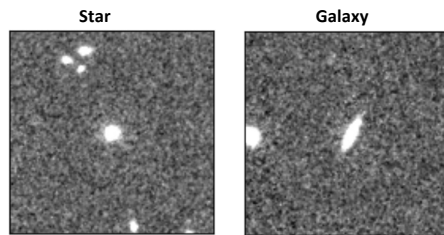
Task Objectives:

- Train point source ML models (star vs. galaxy classification [2]) on Roman simulations only
- Develop the MLOps framework to add a reliability layer on top of the ML model
- Implement the MLOps framework and the ML model in the IPAC alert pipeline

RAPID architectural diagram. The output of ML systems will be embedded into the RAPID alert stream.



Star Galaxy Classification



A star-galaxy classifier will be applied to every reference source outputting a score within [0,1] on whether the object is a star or galaxy.

ML Operations Framework

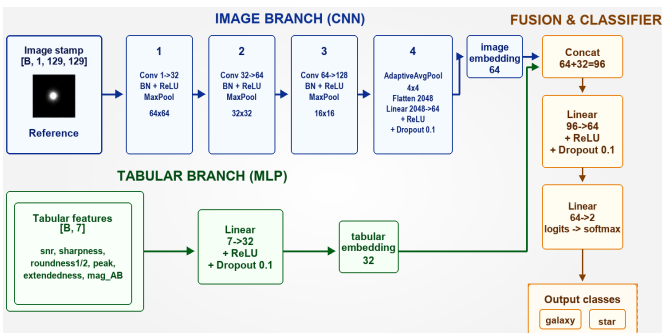
Background: ML Operations or MLOps safeguards deployed classifiers by verifying that real-world inputs match the models' training data. Any inputs failing these checks are flagged as out-of-distribution or low quality. Although standard in commercial settings, robust MLOps frameworks are still frequently omitted in research environments. **(Right)** The figure shows Technology Readiness Levels (TRLs) associated with ML development. TRL 9 systems benefit from formal MLOps to monitor the current system and provide feedback used in the development of subsequent models.



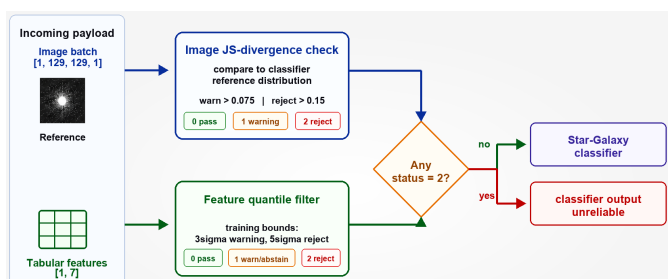
Figure 4 from "Technology readiness levels for machine learning systems" [1]

Technical Approach

Multimodal CNN classifier for Star-Galaxy Classification



MLOps Framework for the Star-Galaxy Classifier



Methodology and Results

We trained on labeled Roman simulation alert data [3] using normalized image cutouts, standardized tabular features, data augmentation, and a weighted cross-entropy objective to handle class imbalance (~10x more galaxies than stars).

Metric	Score	Data	Count
ROC AUC	0.999	Train	37,108
Macro-F1	0.989	Validation	6,549
		Test	7,705

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

www.nasa.gov

Clearance Number: CL#26-2339

Copyright 2026. All rights reserved.

Future Work

Currently under development, the Star-Galaxy classifier and MLOps framework will be deployed into the IPAC RAPID pipeline. Future efforts will focus on testing the classifier using a wider diversity of Open Universe simulated images, evaluating performance across various source magnitudes, locations, and SNRs. Additionally, we are preparing for the launch of the Roman Space Telescope. Its real-time operational demands will require rapid classifier update cycles to seamlessly transition from simulated images to live science data.

References:

- [1] Lavin, A., Gilligan-Lee, C. M., Visnjic, A., Ganju, S., Newman, D., Ganguly, S., Lange, D., Baydin, A. G., Sharma, A., Gibson, A., Zheng, S., Xing, E. P., Matmann, C., Parr, J., & Gal, Y. (2022). Technology readiness levels for machine learning systems. *Nature Communications*, 13(1), 6039
- [2] Tachibana, Y., & Miller, A. A. (2018). A morphological classification model to identify unresolved PanSTARRS1 sources: Application in the ZTF real-time pipeline. *Publications of the Astronomical Society of the Pacific*, 130(994), 128001
- [3] OpenUniverse, LSST Dark Energy Science Collaboration, Roman HLIS Project Infrastructure, Roman Rapid Project Infrastructure Team, Roman Supernova Cosmology Project Infrastructure Team, et al. (2025). OpenUniverse2024: a shared, simulated view of the sky for the next generation of cosmological surveys, arXiv:2501.05632