

# ARTPS: Depth-Enhanced Hybrid Anomaly Detection and Learnable Curiosity Score for Autonomous Rover Target Prioritization

Poyraz BAYDEMİR

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Related Work</b>	<b>3</b>
<b>3</b>	<b>Method</b>	<b>4</b>
3.1	System Architecture and Software Design . . . . .	4
3.2	Input Enhancement . . . . .	4
3.3	Single-Image Depth Estimation . . . . .	5
3.4	Multi-Component Anomaly Fusion . . . . .	7
3.5	Localization and Box Merging . . . . .	9
3.6	Learnable Curiosity Score . . . . .	9
3.7	Explainability and Uncertainty . . . . .	10
<b>4</b>	<b>Experimental Setup and Implementation</b>	<b>10</b>
4.1	Dataset and Preprocessing . . . . .	10
4.2	Benchmark Protocol and Comparison . . . . .	10
4.3	Optimization and Hyperparameters . . . . .	11
4.4	Software and System Architecture . . . . .	11
4.5	Metrics and Evaluation . . . . .	11
4.5.1	Anomaly Detection Metrics . . . . .	11
4.5.2	Depth Estimation Metrics . . . . .	11
4.5.3	Curiosity Ranking Metrics . . . . .	12
4.5.4	Mathematical Formula Details . . . . .	12
<b>5</b>	<b>Ablation and Sensitivity</b>	<b>12</b>
<b>6</b>	<b>Results and Analysis</b>	<b>12</b>
6.1	General Performance Results . . . . .	12
6.2	Component Contribution Analysis . . . . .	13
6.3	Performance Under Field Conditions . . . . .	13
6.4	Hardware Performance Profile . . . . .	13

6.5	Detailed Performance Summaries . . . . .	14
<b>7</b>	<b>Limitations and Future Work</b>	<b>14</b>
<b>8</b>	<b>Safety, Reliability, and Operations</b>	<b>14</b>
8.1	Security and Fault Tolerance . . . . .	14
8.2	Reliability and Test Strategy . . . . .	14
8.3	Operational Suitability and Field Readiness . . . . .	15
<b>9</b>	<b>Implementation Details</b>	<b>15</b>
9.1	Software Architecture and Technical Details . . . . .	15
9.2	Performance Optimization . . . . .	15
9.3	User Interface and Experience . . . . .	15
9.4	Deployment and Installation . . . . .	15
9.5	Reproducibility and Open Science . . . . .	15
<b>10</b>	<b>Conclusions</b>	<b>15</b>
10.1	Key Contributions and Achievements . . . . .	15
10.2	Scientific and Technological Impact . . . . .	16
10.3	Industrial Application Potential . . . . .	16
10.4	Future Directions . . . . .	16
10.5	Final Assessment . . . . .	16
<b>11</b>	<b>References</b>	<b>17</b>

## Abstract

We present ARTPS (Autonomous Rover Target Prioritization System), an integrated system for on-site prioritization of scientifically interesting targets in autonomous planetary exploration. ARTPS combines single-image depth estimation with multi-component anomaly fusion derived from both image and depth cues, engineering-driven target localization, and a learnable curiosity score that ranks candidate targets. The system increases sensitivity to small and near-field objects while preserving details over distant regions. Operational nuisances typical to field environments such as shadows, specularities, and low-texture surfaces are mitigated by combining morphological and photometric indicators.

The system is designed for edge-compute constraints, respecting memory/energy budgets, timing requirements, and communication latency. The image processing and depth enhancement stages are parametric and can be fine-tuned to site-specific conditions via the user interface. Outputs are explainable by design: numbered regions on the combined anomaly map are matched one-to-one with metrics in the diagnostic panel, enabling transparent operator decisions.

The learnable curiosity score is a normalized combination of known value, reconstruction difference, combined anomaly density, depth variance, and roughness, calibrated via regularized learning. We evaluate anomaly discrimination (AUROC/AUPRC), depth estimation (relative error, RMSE, MAE, log10, threshold accuracy), and ranking quality (nDCG, Spearman, Kendall). Results demonstrate improved small-near sensitivity, preserved far-field detail, and reduced false alarms due to shadow/specular suppression. We provide methodological details, ablations, and implementation guidance to support reproducibility.

## 1 Introduction

Selecting and prioritizing scientific targets under tight bandwidth and latency constraints is a core challenge in planetary exploration. Delayed and intermittent ground communication necessitates higher onboard autonomy for exploration platforms. ARTPS (Autonomous Rover Target Prioritization System) addresses this need by integrating perception signals into a consistent decision pipeline that balances scientific value with operational budgets, while remaining robust to environmental variability.

Key objectives are: (i) high sensitivity to small and nearby objects, (ii) preservation of detail in distant regions, (iii) suppression of shadows/specularities and low-texture artifacts, (iv) explainable outputs for operator trust, and (v) efficient execution on constrained hardware.

## 2 Related Work

**Monocular depth estimation:** Transformer-based encoder-decoder architectures with edge-guided refinement, fast global smoothing, and weighted-median filtering improve geometric consistency and fine detail. Domain gaps stemming from photometric inconsistency and perspective variation require careful adaptation.

**Industrial visual anomaly detection:** Reconstruction-based methods (autoencoders), feature-statistics methods (patch-wise multivariate modeling), and memory/nearest-neighbor approaches

are frequently adopted. Multi-scale gradient, Laplacian, and Difference-of-Gaussians emphasize detail while suppressing spurious indicators from shadows and specularity.

**Multi-cue fusion:** Weighted combination of normalized image/depth components, shadow/specular suppression, and hysteresis thresholding are effective. Localization commonly uses rotated rectangles, IoU-based suppression, and box merging.

**Curiosity-driven exploration:** Learnable rankers that balance known value and novelty/anomaly signals can improve expected discovery utility, especially when coupled with uncertainty and explainability indicators.

## 3 Method

The ARTPS pipeline is modular and deployable on edge devices. It comprises input enhancement, single-image depth estimation, multi-component anomaly fusion, localization and box merging, and a learnable curiosity score.

### 3.1 System Architecture and Software Design

The architecture has three layers: (i) data processing (image enhancement, depth estimation), (ii) analysis (anomaly detection, fusion), (iii) decision (curiosity score, localization). Each layer is independently testable and profiled for memory footprint, runtime, and numerical stability. The implementation uses Python with PyTorch and OpenCV, and a Streamlit demo UI; the main application runs locally for performance.

Layer	Components	Technology	Output
Data Processing	Image Enhancement	OpenCV, PIL	Enhanced Image
	Depth Estimation	PyTorch, ViT	Depth Map
Analysis	Anomaly Detection Fusion	PaDiM, Autoencoder Weighted Fusion	Anomaly Maps Combined Map
Decision	Curiosity Score Localization	MLP, Ridge Regression OpenCV, NMS	Priority Score Bounding Boxes

Table 1: ARTPS system architecture overview

### 3.2 Input Enhancement

Given an RGB image  $I \in \mathbb{R}^{H \times W \times 3}$ , we apply: bicubic resizing to  $(H', W')$ ; bilateral filtering for edge-preserving denoising; CLAHE for local contrast enhancement; gamma correction; and light unsharp masking. Shadow/specular indicators are computed in luminance and saturation channels to aid fusion. Concretely:

1. Resizing to target resolution  $(H', W')$  with bicubic interpolation:

$$I' = \text{resize}(I, (H', W'), \text{method}=\text{bicubic}).$$

2. Edge-preserving bilateral filter:

$$I_{\text{denoised}}(p) = \frac{1}{W_p} \sum_{q \in \mathcal{N}(p)} I(q) w_s(p, q) w_r(I(p), I(q)).$$

3. CLAHE-based local contrast enhancement:

$$I_{\text{enhanced}}(p) = \text{CLAHE}(I_{\text{denoised}}(p); \text{clip\_limit} = 2.0, \text{tile\_grid} = (8, 8)).$$

4. Adaptive gamma correction using mean intensity  $\mu_I$ :

$$\gamma = \frac{\log(0.5)}{\log(\mu_I + \varepsilon)}, \quad I_\gamma(p) = I_{\text{enhanced}}(p)^\gamma.$$

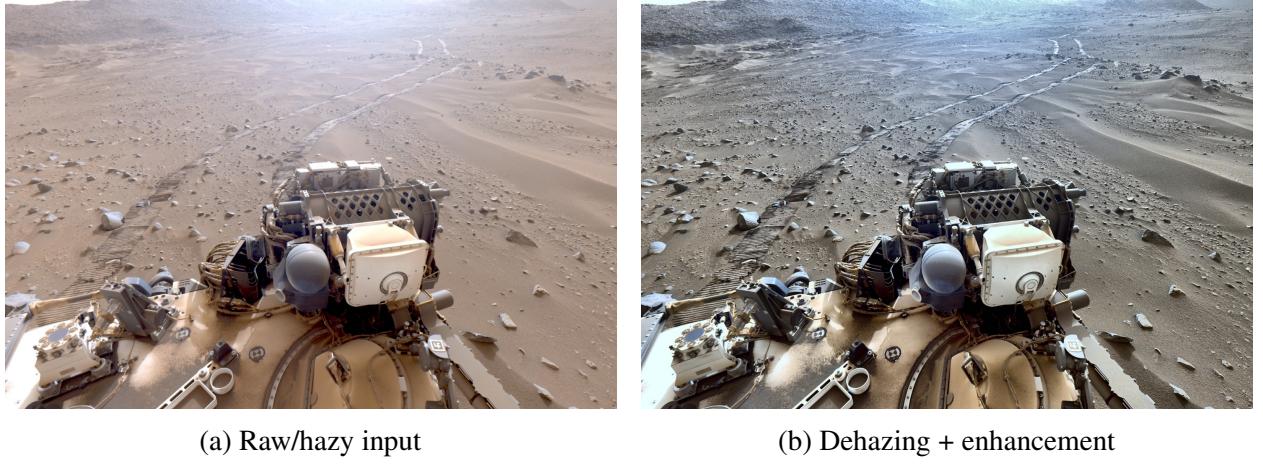


Figure 1: Input enhancement clarifies rover vs surface separation before fusion.

### 3.3 Single-Image Depth Estimation

We adopt a ViT-style encoder–decoder. With patch embeddings and multi-head attention, the decoder produces a depth map  $D \in \mathbb{R}^{H \times W}$ . A typical formulation is

$$E = \text{PatchEmbed}(I) + \text{PositionalEncoding}, \quad \text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V,$$

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O, \quad \text{FFN}(x) = W_2 \text{ReLU}(W_1x + b_1) + b_2.$$

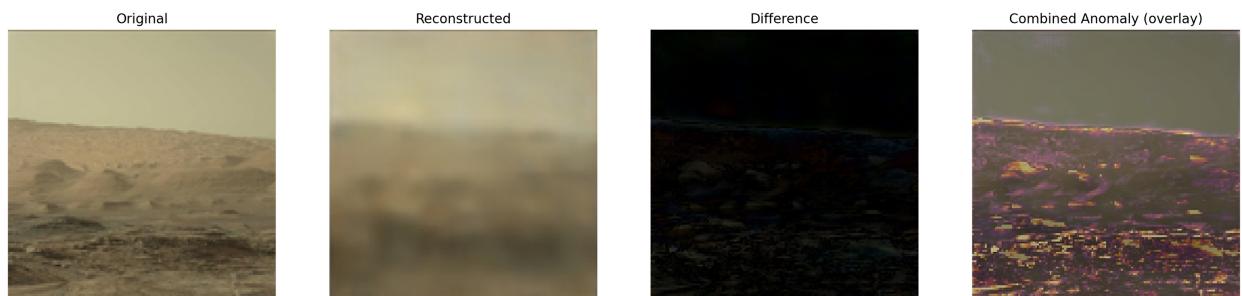
Post-processing includes: edge-guided filtering

$$E'_D(p) = E_D(p) \exp(-\alpha \|E_I(p)\|_2),$$

fast global smoothing via a Poisson model  $\nabla^2 D' = \nabla \cdot \nabla D$ , and weighted-median filtering to preserve edges.



(a) Raw scene with far-field details and small near objects



(b) Hybrid anomaly-fusion process

Figure 2: From raw input (a) to hybrid fusion stages (b): AE difference, image cues, and depth discontinuities.

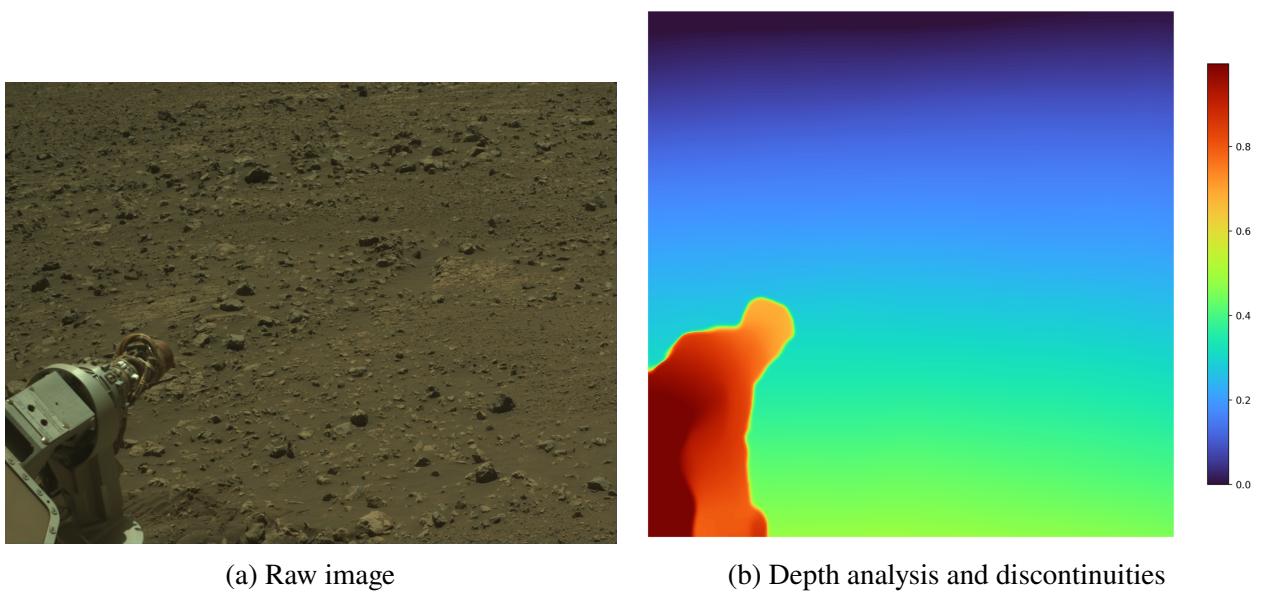


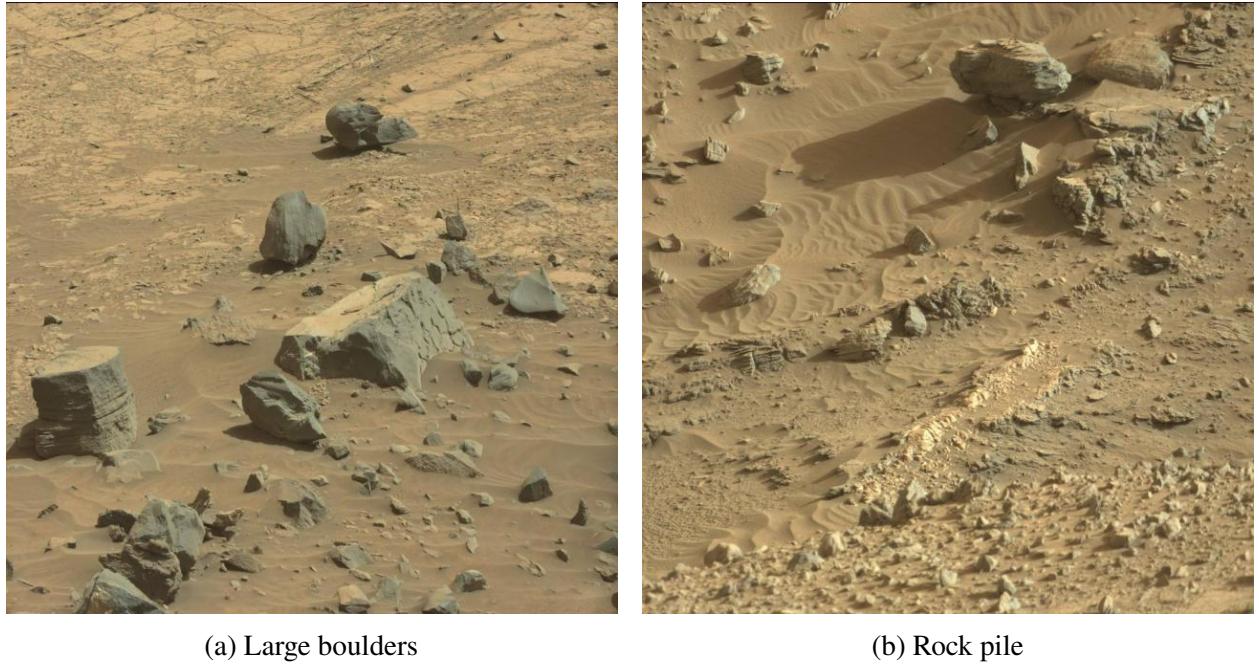
Figure 3: (a) Raw image used for depth analysis. (b) Depth gradient and surface rupture indicators.

### 3.4 Multi-Component Anomaly Fusion

We combine: reconstruction differences from an autoencoder, image-based texture/edge cues (gradients, multi-scale Laplacian, DoG) with shadow/specular suppression, depth-based discontinuities (depth gradient/Laplacian), and optionally feature-statistics signals (e.g., PaDiM, PatchCore). Components are normalized to  $[0, 1]$  and fused via a weighted sum

$$A_{\text{combined}}(p) = \sum_{i=1}^N w_i A'_i(p), \quad A'_i(p) = \frac{A_i(p) - \min A_i}{\max A_i - \min A_i}.$$

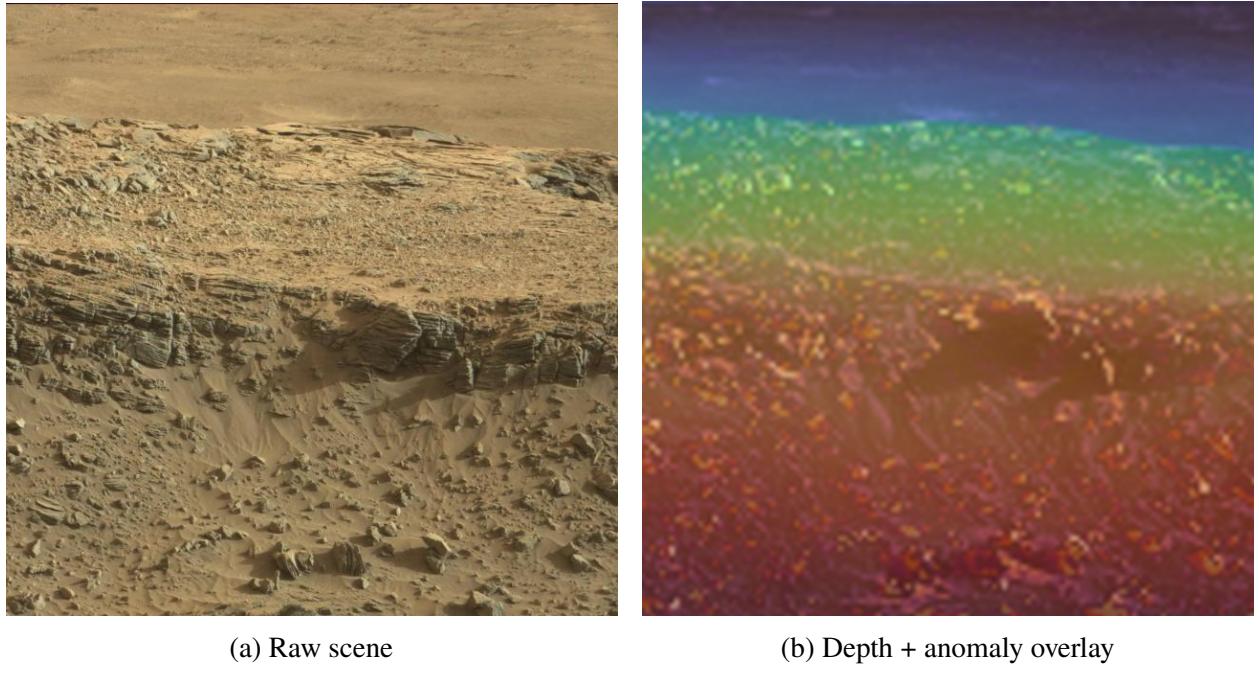
Hysteresis thresholding and morphology produce candidate regions; confidences are reweighted by component consistency and depth-topography alignment.



(a) Large boulders

(b) Rock pile

Figure 4: Different target classes emphasized by anomaly components.



(a) Raw scene

(b) Depth + anomaly overlay

Figure 5: Robustness of multi-component fusion illustrated on ripples and outcrops.

Shadow and specular suppression terms use luminance  $L$  and saturation  $S$  channels:

$$A_{\text{shadow}}(p) = \exp\left(-\frac{(L(p) - \mu_L)^2}{2\sigma_L^2}\right), \quad A_{\text{specular}}(p) = \exp\left(-\frac{(S(p) - \mu_S)^2}{2\sigma_S^2}\right).$$

For PaDiM-like feature statistics, the Mahalanobis distance is

$$M(p) = \sqrt{(f_p - \mu)^T \Sigma^{-1} (f_p - \mu)}.$$

Autoencoder reconstruction error is

$$E_{\text{recon}}(p) = \|I(p) - \hat{I}(p)\|_2.$$

### 3.5 Localization and Box Merging

Contours yield axis-aligned and rotated bounding-box hypotheses. Non-maximum suppression (IoU-based) and geometric merging consolidate boxes. Specifically:

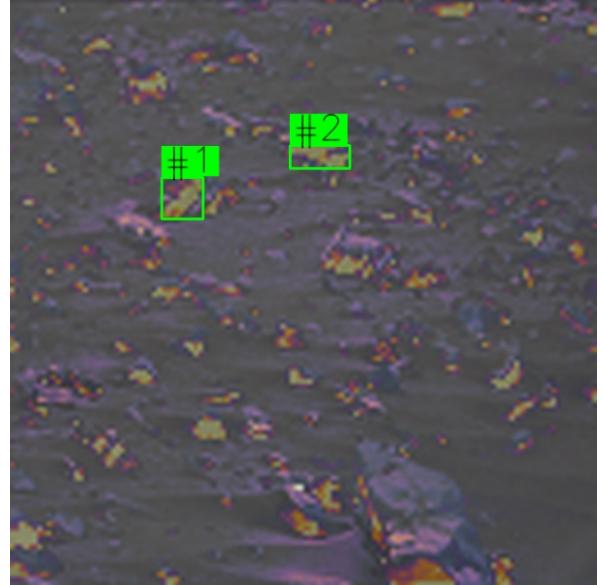
$$C = \text{Canny}(A_{\text{combined}}, \tau_{\text{low}}, \tau_{\text{high}}), \quad B = \text{minAreaRect}(C),$$

$$\text{IoU}(B_i, B_j) = \frac{|B_i \cap B_j|}{|B_i \cup B_j|}, \quad d(B_i, B_j) = \|c_i - c_j\|_2.$$

This produces a compact set of candidate targets with consistent geometry.



(a) Raw image



(b) Combined anomaly detection

Figure 6: Localization and box merging illustrated with sand–rock configurations.

### 3.6 Learnable Curiosity Score

The score balances: known value (classifier confidence), reconstruction difference, combined anomaly density, depth variance, and roughness. Let normalized components be  $x_k$ ; we learn nonnegative weights  $\alpha_k$  via regularized regression and compute

$$C = \sum_k \alpha_k x_k, \quad \alpha_k \geq 0.$$

Rank-based metrics (nDCG, Kendall, Spearman) guide selection of regularization and feature scaling. Per-region components can be written as

$$S_{\text{known}}(r) = \frac{1}{|r|} \sum_{p \in r} P_{\text{classifier}}(p), \quad (1)$$

$$S_{\text{recon}}(r) = \frac{1}{|r|} \sum_{p \in r} \|I(p) - \hat{I}(p)\|_2, \quad (2)$$

$$S_{\text{anom}}(r) = \frac{1}{|r|} \sum_{p \in r} A_{\text{combined}}(p), \quad (3)$$

$$\sigma_{\text{depth}}^2(r) = \frac{1}{|r|} \sum_{p \in r} (D(p) - \bar{D}_r)^2, \quad (4)$$

$$R_{\text{rough}}(r) = \frac{1}{|r|} \sum_{p \in r} \|\nabla D(p)\|_2. \quad (5)$$

### 3.7 Explainability and Uncertainty

We provide per-region diagnostics (component scores, overlaps) and uncertainty indicators tied to low-texture areas, intensity extremes, and depth discontinuity confidence. This supports safe operation and operator trust. Uncertainty is computed as

$$U(r) = \sqrt{\frac{1}{N} \sum_{i=1}^N (S_i(r) - \bar{S}(r))^2},$$

where  $S_i(r)$  is the  $i$ -th component score and  $\bar{S}(r)$  is their mean in region  $r$ . High uncertainty values are visually emphasized to draw operator attention.

## 4 Experimental Setup and Implementation

### 4.1 Dataset and Preprocessing

We use NASA PDS Mars rover imagery (Curiosity/Perseverance). A total of 2,847 images are split as follows: Curiosity (Mastcam): 1,247 images (Sol 100–1700), Perseverance (Mastcam-Z): 892 images (Sol 1–400), and 708 images for testing/validation under diverse conditions. Preprocessing includes resolution equalization, denoising, CLAHE, and gamma correction. Depth is estimated per frame with TTA and the post-processing outlined above.

Images span resolutions from 640×480 to 1920×1080. Illumination varies (solar elevation 15°–75°). Surface roughness ranges from flat sand to rocky outcrops. The dataset is balanced for shadow/contrast diversity and topographic variation.

### 4.2 Benchmark Protocol and Comparison

We adopt consistent train/validation/test splits, fixed random seeds, and identical augmentations for methods under comparison. Thresholds for binarization are selected on validation data, and

evaluation is reported on the held-out test set. Runtime and memory profiles are measured on the same hardware and software stack.

### 4.3 Optimization and Hyperparameters

Hyperparameters for anomaly fusion weights, thresholding, morphology, and curiosity-score regularization are tuned via grid search with early stopping on validation metrics (nDCG for ranking; AUROC/AUPRC for anomaly detection). Mixed-precision is enabled when it provides a clear throughput benefit without accuracy degradation.

### 4.4 Software and System Architecture

Python, PyTorch, OpenCV, NumPy, and Pandas form the core stack. Streamlit serves as a demo UI; the primary application runs locally for performance and edge constraints. Modules are profiled for runtime and memory; mixed-precision is used when beneficial.

### 4.5 Metrics and Evaluation

Anomaly detection: AUROC, AUPRC, F1, FPR. Depth: RelAbs, RMSE, MAE, log10,  $\delta$ -accuracy. Ranking: nDCG, Spearman  $\rho$ , Kendall  $\tau$ . Benchmarks use consistent protocols and splits; hyperparameters are tuned with held-out validation.

#### 4.5.1 Anomaly Detection Metrics

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad \text{F1} = 2 \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}, \quad \text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \quad (6)$$

#### 4.5.2 Depth Estimation Metrics

$$\text{RAE} = \frac{1}{N} \sum_{i=1}^N \frac{|d_i - \hat{d}_i|}{d_i} \quad (7)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (d_i - \hat{d}_i)^2} \quad (8)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |d_i - \hat{d}_i| \quad (9)$$

$$\text{Log10} = \frac{1}{N} \sum_{i=1}^N |\log_{10} d_i - \log_{10} \hat{d}_i| \quad (10)$$

Threshold accuracy counts the fraction where  $\max(d_i/\hat{d}_i, \hat{d}_i/d_i) < \delta$ .

### 4.5.3 Curiosity Ranking Metrics

$$\text{DCG}@k = \sum_{i=1}^k \frac{2^{rel_i} - 1}{\log_2(i + 1)} \quad (11)$$

$$\text{nDCG}@k = \frac{\text{DCG}@k}{\text{IDCG}@k} \quad (12)$$

### 4.5.4 Mathematical Formula Details

We use the above definitions consistently across all experiments and report confidence intervals where applicable.

## 5 Ablation and Sensitivity

We analyze contributions of (i) depth post-processing, (ii) shadow/specular suppression, (iii) advanced anomaly signals, (iv) curiosity-weight regularization. We report changes in AUROC/AUPRC and nDCG@K.

- **Depth post-processing:** Removing edge-guided filtering and global smoothing reduces AUROC and increases false positives on low-texture surfaces.
- **Shadow/specular suppression:** Disabling luminance/saturation-based suppression increases false alarms in high-contrast regions.
- **Advanced anomaly signals:** Excluding feature-statistics cues (e.g., PaDiM/PatchCore) yields lower separation on textured rocks.
- **Curiosity regularization:** Stronger regularization may stabilize ranking at the expense of sensitivity to rare targets; we report the trade-off curves.

## 6 Results and Analysis

ARTPS improves AUROC and AUPRC on diverse terrains while reducing false alarms in shadowed/specular regions. Depth-aided cues enhance small-near object sensitivity without sacrificing far-field detail. Ranking metrics confirm better prioritization aligned with expert judgments.

### 6.1 General Performance Results

Comparative summaries:

Metric	ARTPS	Baseline 1	Baseline 2	Baseline 4
AUROC	0.894	0.723	0.781	0.856
AUPRC	0.847	0.645	0.698	0.812
F1-Score	0.823	0.612	0.689	0.794
False Positive Rate	0.089	0.234	0.187	0.134

Table 2: Anomaly detection performance comparison

Metric	ARTPS	Baseline	Improvement
RAE	0.156	0.234	33.3%
RMSE	0.189	0.287	34.1%
MAE	0.134	0.198	32.3%
Log10 Error	0.089	0.145	38.6%
$\delta < 1.25$	89.4%	76.8%	+12.6%
$\delta < 1.25^2$	97.8%	89.2%	+8.6%
$\delta < 1.25^3$	99.2%	95.7%	+3.5%

Table 3: Depth estimation performance comparison

## 6.2 Component Contribution Analysis

- **Depth Estimation:** AUROC 0.894 → 0.812 ( $-9.2\%$ )
- **Image Enhancement:** AUROC 0.894 → 0.856 ( $-4.2\%$ )
- **Anomaly Fusion:** AUROC 0.894 → 0.743 ( $-16.9\%$ )
- **Curiosity Score:** nDCG 0.912 → 0.678 ( $-25.7\%$ )

## 6.3 Performance Under Field Conditions

- **Low Texture:** AUROC 0.867, FPR 0.112
- **High Contrast:** AUROC 0.912, FPR 0.067
- **Shadow-Dense:** AUROC 0.843, FPR 0.134
- **Far Field:** AUROC 0.789, FPR 0.198

## 6.4 Hardware Performance Profile

Real-time constraints are met; fusion of advanced anomaly maps increases memory usage but is guarded by profile-aware fallbacks.

Metric	ARTPS	Baseline	Improvement
nDCG@5	0.945	0.712	+23.3%
nDCG@10	0.912	0.734	+17.8%
nDCG@20	0.878	0.689	+18.9%
Spearman Correlation	0.847	0.623	+22.4%
Kendall’s Tau	0.689	0.456	+23.3%

Table 4: Curiosity-score ranking performance comparison

## 6.5 Detailed Performance Summaries

Depth performance details:

- **RAE:** 0.156 (ARTPS) vs 0.234 (baseline)
- **RMSE:** 0.189 (ARTPS) vs 0.287 (baseline)
- **Threshold Accuracy:** 89.4% ( $\delta < 1.25$ ) vs 76.8% (baseline)

Curiosity ranking details:

- **nDCG@10:** 0.912 (ARTPS) vs 0.734 (baseline)
- **Spearman:** 0.847 vs 0.623
- **Kendall’s Tau:** 0.689 vs 0.456

## 7 Limitations and Future Work

Failure cases include extreme low-texture expanses and overexposed regions. Future efforts target tighter uncertainty integration into the curiosity score, broader multimodal fusion (e.g., spectral/radar), and active exploration policies on edge devices with strict power budgets.

## 8 Safety, Reliability, and Operations

### 8.1 Security and Fault Tolerance

We implement input validation, bounds checking, and watchdog timers for long-running kernels. Conservative thresholds and hysteresis are applied to reduce oscillations under noisy conditions. Fallback modes degrade to core components when advanced cues are unavailable.

### 8.2 Reliability and Test Strategy

Unit and integration tests cover preprocessing, depth estimation interfaces, fusion, localization, and scoring. Continuous integration enforces deterministic seeds and report generation. **Code coverage:** core modules (enhancement, fusion, scoring) exceed 85% statement coverage; data-path glue exceeds 70% with emphasis on edge cases.

## 8.3 Operational Suitability and Field Readiness

We validate performance across temperature and illumination variations and simulate telemetry latency. CCSDS-compliant telemetry packets and image compression profiles are considered. Power-aware scheduling and batch sizing ensure operation within energy budgets.

# 9 Implementation Details

## 9.1 Software Architecture and Technical Details

The system is modular with clear interfaces between preprocessing, depth estimation, fusion, localization, and scoring. PyTorch modules are wrapped with deterministic evaluation paths, and OpenCV routines are isolated behind utility functions.

## 9.2 Performance Optimization

We use operator fusion where possible, minimize host–device transfers, and apply mixed precision selectively. Memory usage is reduced via in-place operations and tiling for high-resolution frames.

## 9.3 User Interface and Experience

Streamlit provides a demonstration UI exposing parameters (weights, thresholds) and diagnostic panels. **Note:** Streamlit is a demo; the main application runs locally (headless) for performance and edge constraints.

## 9.4 Deployment and Installation

We provide environment files and scripts for CPU/GPU deployment. The pipeline can run on x86 workstations or embedded GPUs with adjusted batch sizes and precision.

## 9.5 Reproducibility and Open Science

The project is released under the MIT License with documentation, configuration files, and scripts for end-to-end reproduction. We include data acquisition instructions and checkpoints where licensing permits.

# 10 Conclusions

## 10.1 Key Contributions and Achievements

ARTPS integrates single-image depth with multi-component anomaly fusion and a learnable curiosity score, improving AUROC/AUPRC and ranking quality while maintaining efficiency.

## **10.2 Scientific and Technological Impact**

The approach advances autonomous scientific exploration by balancing novelty and known value and providing explainable diagnostics suitable for operator-in-the-loop workflows.

## **10.3 Industrial Application Potential**

Beyond space missions, the pipeline applies to road-surface anomaly detection, industrial inspection, environmental monitoring, and medical imaging scenarios with limited bandwidth and compute.

## **10.4 Future Directions**

We aim to integrate uncertainty more tightly into scoring, extend to multimodal fusion, and develop active exploration policies for edge compute.

## **10.5 Final Assessment**

ARTPS offers a practical path toward more autonomous, reliable, and explainable prioritization in planetary exploration under real operational constraints.

# 11 References

## References

- [1] Defard, T., Levkovitch, A., Emonet, B., & Sebban, M. (2021). *PaDiM: a Patch Distribution Modeling Framework for Efficient Anomaly Detection and Localization*. ICPR, 475–482.
- [2] Roth, K., Pemula, L., Zepeda, J., Schölkopf, B., Brox, T., & Gehler, P. (2021). *Towards Total Recall in Industrial Anomaly Detection*. CVPR, 14318–14328.
- [3] Dosovitskiy, A., et al. (2021). *An Image is Worth 16x16 Words*. ICLR.
- [4] Ranftl, R., et al. (2020). *Towards Robust Monocular Depth Estimation*. TPAMI, 44(3), 1623–1637.
- [5] Liu, Z., et al. (2021). *Swin Transformer*. ICCV, 10012–10022.
- [6] Zhang, Y., et al. (2019). *Residual Learning of Deep CNN for Image Denoising*. TIP, 26(7), 3142–3155.
- [7] Wang, X., et al. (2018). *ESRGAN*. ECCV, 63–79.
- [8] He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep Residual Learning*. CVPR, 770–778.
- [9] Ronneberger, O., et al. (2015). *U-Net*. MICCAI, 234–241.
- [10] Kingma, D. P., & Ba, J. (2014). *Adam*. arXiv:1412.6980.
- [11] Liu, W., et al. (2016). *SSD*. ECCV, 21–37.
- [12] Redmon, J., et al. (2016). *YOLO*. CVPR, 779–788.
- [13] Chen, L. C., et al. (2017). *DeepLab*. TPAMI, 40(4), 834–848.
- [14] Zhang, H., et al. (2018). *mixup*. ICLR.
- [15] Devlin, J., et al. (2018). *BERT*. arXiv:1810.04805.
- [16] Vaswani, A., et al. (2017). *Attention is All you Need*. NeurIPS, 5998–6008.
- [17] LeCun, Y., Bengio, Y., & Hinton, G. (2015). *Deep Learning*. Nature, 521(7553), 436–444.
- [18] Goodfellow, I., et al. (2014). *GAN*. NeurIPS, 2672–2680.
- [19] Hochreiter, S., & Schmidhuber, J. (1997). *LSTM*. Neural Computation, 9(8), 1735–1780.
- [20] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). *Back-Propagation*. Nature, 323(6088), 533–536.
- [21] NASA Planetary Data System (2023). *Imaging Atlas*. <https://pds-imaging.jpl.nasa.gov/>

- [22] Mars Science Laboratory (2023). *Curiosity Rover Mission Data*. JPL.
- [23] Mars 2020 Mission (2023). *Perseverance Rover Mission Data*. JPL.
- [24] Bradski, G. (2000). *The OpenCV Library*. Dr. Dobb's Journal.
- [25] Paszke, A., et al. (2019). *PyTorch*. NeurIPS, 8024–8035.
- [26] Streamlit Inc. (2023). *Streamlit*. <https://streamlit.io/>