Terraluna: Scale-Resolved Terrain Fingerprints for Earth Analogs of Lunar and Martian Sites, with a Benchmark of a Physics-Informed Encoder

Terraluna

Abstract

We compare the terrain of Connecting Ridge, one of the candidate Artemis III landing regions that NASA announced near the lunar south pole, and of two Martian sites, Jezero crater and Arcadia Planitia, with Earth surfaces, using elevation data alone. A deterministic fingerprint of relief and texture, the structure function $\nu(L)$ at hectometre lags and its Hurst exponent, ranks 71 Earth windows from two declared regions for each target, and the same code computes it in Python and in a browser within stated numerical tolerances. The closest Earth window to the lunar target lies at a relief distance of 0.68 in root mean square log units, farther from the target than any other window of the lunar site, and an inspection shows it to be a nearest neighbour in this metric rather than a morphological analog. The Martian shortlists lie at relief distances from 0.054 to 0.094 for Jezero and 0.185 to 0.289 for Arcadia, below the median internal distance of those sites; this is a match of terrain statistics at hectometre lags only. No radar comparison was completed. We also benchmark a self-supervised terrain encoder of 2,795,040 parameters trained with VICReg, a physics head and a gradient reversal body adversary on 28,843 windows. We report it against acceptance criteria fixed before training: physics probe, body invariance, agreement with the fingerprint and seed stability, with an untrained copy of the network as a baseline; the morphology probe and the comparison with image baselines are not available. A validation sweep on a laptop accelerator covers the adversarial weight and training length. The encoder is reported as experimental and the rankings use the fingerprint. All numbers are generated from recorded artifacts.

1 Introduction

Crews and rovers rehearse on Earth before they work on the Moon or Mars, and the value of a rehearsal depends on how well the Earth site stands in for the target. Connecting Ridge, one of the candidate Artemis III landing regions that NASA announced near the lunar south pole (NASA, 2022), has elevation models at 5 m per pixel built from laser altimetry alone (Barker et al., 2021). The practical question is which Earth surfaces are worth visiting to learn about such a site, and on what evidence. Quantitative studies show how hard the comparison is: even three lava facies at one Icelandic flow field overlap by 88% to 89% in their radar polarization ratio (Tolometti et al., 2022).

Roughness statistics are the natural quantitative starting point. Roughness has been reported at stated baselines from terrain data on Earth, the Moon and Mars (Shepard et al., 2001; Rosenburg et al., 2011; Kreslavsky and Head, 2000), and the structure function of elevation and its Hurst exponent can be computed with shared definitions on all three bodies, while each product keeps its own instrument, processing and support. They are second order statistics, however. They summarise how height differences grow with lag and do not determine whether a surface is cratered, ridged or covered in flows, so two windows can be close in every statistic and still differ in the

morphology that matters for a traverse. Learned representations can carry morphology. Recent lunar foundation models are trained on lunar data (Girona-Mata et al., 2026; Fraccaro et al., 2026), and a learned terrain analog retrieval works between two terrestrial regions (Liu et al., 2026); a learned similarity across bodies is only useful if it can be checked against physics.

We take up both routes for one lunar target, under one set of declared rules: every distance names its observable and its supported lags, missing evidence is never read as zero, and every number in this paper is generated from a recorded artifact. The fingerprint is computed by one Python implementation offline and by a TypeScript port that runs in a browser worker, and the two are held to stated numerical tolerances (Section 8.5). The contributions are as follows. First, a deterministic fingerprint of relief and texture compares the Connecting Ridge target window with 71 eligible Earth windows and returns a shortlist of 1 candidate window, with the closest relief distance 0.68 against a median of 2.05 (Section 8.2); the same comparison for the Jezero and Arcadia windows returns shortlists of 3 and 2 windows whose distances lie below the median internal distance of each Martian site (Section 8.3). Second, a radar availability audit decodes the Mini-RF mosaic over the site and a UAVSAR scene over Holuhraun, and records why a radar comparison could not be qualified within the study (Section 8.7). Third, a terrain encoder for Moon, Mars and Earth elevation with a physics head and a gradient reversal body head is held to acceptance criteria fixed before training; trained at a reduced scope, its probe results are reported against those criteria with an untrained copy of the network as a baseline, and the morphology probe and the image baselines are not available (Section 8.1). Fourth, a benchmark section reports a validation sweep over the adversarial weight and training length, training throughput on a CPU server and a laptop accelerator, and the numerical parity of the browser implementations (Section 8.5).

2 Background

2.1 Planetary roughness statistics

Surface roughness has no single value. Shepard et al. (2001) compared topographic data from many natural surfaces and concluded that a roughness measure must be tied to the horizontal scale at which it is taken, so they proposed reporting scale dependent parameters in a standard form. The central quantity is the root mean square height difference between points a fixed lag apart. Its growth with lag is the deviogram, and on a log scale the slope of the deviogram is the Hurst exponent of a self affine surface. These definitions let workers with different instruments and sampling report comparable numbers, provided that each number states its lag, support and processing.

Laser altimetry turned these definitions into maps. Rosenburg et al. (2011) computed slopes and Hurst exponents of the Moon from LOLA over baselines from about 17 m to about 2.7 km. Most of the surface behaves like a fractal with one or two exponents over that range, and where the exponent changes, the breakover typically sits near a 1 km baseline (Rosenburg et al., 2011). The median Hurst exponent is 0.95 in the highlands and 0.76 in the maria (Rosenburg et al., 2011). Kreslavsky et al. (2013) mapped lunar roughness from LOLA profile curvature at baselines from 115 m to 1.8 km and found that hectometre and kilometre roughness correlate poorly, because they record different processes and ages. On Mars, Kreslavsky and Head (2000) measured kilometre scale roughness from MOLA as the median absolute differential slope over baselines from 0.6 km to about 20 km and showed that geological units have distinct roughness and distinct scale dependence. Field measurements give the same caution at smaller scales: lava flow roughness behaves as a self affine surface only over a limited range of lengths, and the smallest scales follow different statistics (Campbell and Shepard, 1996). A Hurst exponent is therefore a description of a stated range of

lags, not a property of a surface at all scales.

2.2 Radar polarimetry and the circular polarization ratio

Radar adds information at the scale of its wavelength. Mini-RF on the Lunar Reconnaissance Orbiter is a synthetic aperture radar with a hybrid polarization architecture that measures the Stokes parameters of the echo (Nozette et al., 2010). From them, the circular polarization ratio (CPR) of the same sense to the opposite sense return is formed, and the m-chi decomposition interprets the echo through the degree of polarization and the Poincaré ellipticity parameter (Raney et al., 2012). CPR is sensitive to structure near the radar wavelength. Real CPR values above unity occur on the Moon, on Venus and on terrestrial lava flows, and scattering from rock edges and between rock facets can explain them (Campbell, 2012). Lunar impact melt flows typically have high CPR, so deposits that look smooth in optical images may be rough at centimetre to decimetre scales (Carter et al., 2012).

CPR is not a direct measure of rock size. Viewing geometry shifts it: across the inner wall of a bowl shaped crater the local incidence angle changes by about 30 degrees, which can change CPR by about 0.2 (Fa and Cai, 2013). The same study attributes the enhanced CPR inside anomalous craters most probably to rocks perched on or buried in the regolith, rather than to ice (Fa and Cai, 2013). Across the lunar highlands, Rivera-Valentín et al. (2026) found that CPR rises quickly with rock abundance at low abundance, then levels off and can fall over rock rich terrain, so a CPR of 0.8 may correspond to a rock abundance of about 0.7% or about 10%. They show that several polarimetric properties used together resolve this ambiguity (Rivera-Valentín et al., 2026). Mini-RF mosaics also depend on how terrain effects are handled. Fassett et al. (2024) orthorectified the archived monostatic data onto lunar topography and reduced topographic effects with an empirical correction based on the known effect of incidence angle. A learned inversion inherits the same geometry dependence: for a deep learning inversion of Mini-RF mosaics, comparing east and west looking polar mosaics gives a look direction uncertainty of 1.01 in the real part of the inferred dielectric constant (Shukla et al., 2024).

2.3 Quantitative analog studies

Earth surfaces have long served as candidates for interpreting planetary radar and terrain signatures, but quantitative studies report overlapping populations rather than clean matches. At the 2014 to 2015 Holuhraun flow field in Iceland, Tolometti et al. (2022) tested whether three mapped lava facies differ in UAVSAR CPR and Sentinel-1 cross polarization. The facies were statistically distinguishable, yet their CPR populations overlap by 88% to 89%, so radar can suggest a facies in probability but cannot map it (Tolometti et al., 2022). The same study found only small differences in lidar Hurst exponents between most lava types. On Mars, equatorial lava flows are radar bright with CPR above one, like blocky terrestrial flows, but terrain models of 35 flows from HiRISE show that at metre scale they are smoother than blocky flows and closer to pahoehoe and rubbly flows (Rodriguez Sanchez-Vahamonde and Neish, 2021). Decimetre and metre roughness can thus point in different directions on one surface. Learned similarity has also been tried: Meng et al. (2024) trained a convolutional similarity metric on Earth and Mars surface images with a discriminative loss, and applied it to selecting analog areas in the evaporite region of the Qaidam Basin. These studies treat an Earth surface as a candidate for interpreting a signature, a use that field work can then test.

2.4 Self-supervised and domain-adversarial learning

Joint embedding methods learn representations by making two augmented views of one input agree. Their failure mode is collapse to a constant output. Contrastive methods such as SimCLR learn by contrasting augmented views (Chen et al., 2020). Barlow Twins drives the cross correlation of the two views' embeddings towards the identity (Zbontar et al., 2021). VICReg adds an explicit variance term on every embedding dimension to an invariance term and a covariance term that decorrelates dimensions (Bardes et al., 2021). DINO trains a Vision Transformer by self distillation without labels, and its features work well as nearest neighbour classifiers (Caron et al., 2021). Domain-adversarial training addresses a different problem: a gradient reversal layer makes the learned features discriminative for the main task while indiscriminate with respect to the shift between domains (Ganin et al., 2015).

2.5 Learned planetary and terrain representations

The lunar representation models below are trained on lunar data. LunarFM is a multimodal masked autoencoder that maps 18 channels from six instruments on three lunar missions into a shared 768 dimensional embedding, with a training set spanning 70 degrees south to 70 degrees north (Girona-Mata et al., 2026). SomBench assembles more than 30 co-registered lunar layers into machine learning ready tiles with leakage safe splits (Patil et al., 2026), and a lunar foundation model was pretrained on its tiles at 1 m and 100 m per pixel (Fraccaro et al., 2026). On the CraterBench-R retrieval benchmark, self-supervised Vision Transformers with in-domain pretraining outperform larger generic models (Fang et al., 2026). topo2vec learns embeddings of elevation data with self-supervision built on the fractal assumption that structures recur across scales (Kavitzky et al., 2021). GEAR retrieves terrestrial topographic analogs of the Mariana Trench on the Qinghai-Tibet Plateau with a Siamese graph network and an expert annotated similarity dataset (Liu et al., 2026). Optical surveys at the boulder scale complement these models: a convolutional network mapped about 94 million boulders in images between 60 degrees south and 60 degrees north (Aussel et al., 2025).

3 Data

The study uses elevation products, one land cover product, and two radar products that are decoded for the radar availability audit of Section 8.7. Table 1 lists them. Each download is recorded with its provider URL, HTTP metadata and SHA-256 digest, and every derived number in this paper names the artifact it was read from.

Moon. The lunar target is Site01 of the high resolution site products that Barker et al. (2021) built from Lunar Orbiter Laser Altimeter ranges (Smith et al., 2010), the site that the provider names Connecting ridge. The product gives surface height at 5 m per pixel in south polar stereographic metres, the number of LOLA returns per pixel, the total height uncertainty, and 100 error clones of the elevation model. Pixels without a LOLA return are filled by interpolation in the provider's surface height, so we carry the observed fraction of each window beside its statistics. The Site01 extent is tiled with 5000 m windows at a 2500 m stride, giving 25 windows, of which 25 pass the coverage rule. The primary target window is the eligible window that contains the centroid of the union of the provider's region of interest polygons. Its height grid is 100% valid, but for only 10.4% of its valid analysis cells does the nearest 5 m pixel hold a LOLA return; elsewhere the 5 m

Table 1: Data inventory. Grid spacing is the provider's stated posting. The role column says how each product is used.

Body	Product	Posting	Frame	Role
Moon	PGDA product 78, Site01 (Barker et al., 2021)	5 m	south polar stereographic	target windows, error clones
Moon	PGDA product 90, LOLA south polar model, 80 to 90 degrees south	20 m	south polar stereographic	encoder training
Mars	USGS CTX DTMs (Laura et al., 2021)	20 m	equirectangular	encoder training, fingerprint distributions
Earth	Copernicus GLO-30 (Airbus Defence and Space, 2022)	1 arcsec	geographic	search windows, encoder training, probe
Earth	ESA WorldCover v200 (Zanaga et al., 2022)	10 m	geographic	land cover mask
Earth	UAVSAR Holuhraun scene (Rosen et al., 2006)	multilooked	slant range and ground range	radar audit
Moon	Mini-RF south polar CPR mosaic (Nozette et al., 2010)	mosaic	south polar stereographic	radar audit

heights are the provider's interpolation. A permanent shadow fraction is not reported, because no permanent shadow product was qualified for this window.

Mars. Martian terrain comes from the USGS Ames Stereo Pipeline CTX elevation models (Laura et al., 2021; Beyer et al., 2018), built from Context Camera stereo pairs (Malin et al., 2007). Each model ships a quality file with the statistics of its height differences from MOLA altimetry (Smith et al., 2001) after ground control. Models whose standard deviation of that difference exceeds 10 m are dropped; for the encoder dataset this left 12 models and dropped 196. Two Martian target sites are declared, a box at Jezero crater and a box in Arcadia Planitia; each target site is one CTX model that passes the threshold, tiled with the same windows as on the Moon, and Section 8.3 gives the rules that fix the target windows.

Earth. Earth elevation is the Copernicus GLO-30 digital surface model, read as cloud optimised GeoTIFFs from its public bucket. It is a surface model, so canopy and buildings are part of the height. We therefore keep only windows whose land cover in ESA WorldCover v200 (Zanaga et al., 2022) is at least 90% bare or sparse vegetation, snow and ice, grassland or shrubland, with permanent water excluded. Land cover is sampled as the nearest class at the centres of the 90 m analysis cells. Two regions are declared. The Holuhraun lava field in Iceland is clipped to the footprint of the UAVSAR scene and gives 42 windows, 42 of them eligible. Rosamond Dry Lake in California is a terrain only region with 31 windows, of which 29 are eligible and 2 fail the land cover rule. Earth windows are cut in the local UTM zone with the same width and stride as on the Moon.

Windows and resampling. Every statistic is computed in metres inside each body's own frame, and lunar, Martian and terrestrial coordinates are never mixed. A window is resampled onto a square grid about its centre with an inverse spherical azimuthal equidistant projection (Snyder, 1987). Pixel centres sit at $x_i = (i + \frac{1}{2})\Delta - W/2$ and $y_j = W/2 - (j + \frac{1}{2})\Delta$ for width W and spacing Δ. The spheres have the mean radii of the Moon, Mars and Earth (1737400, 3396190 and 6371008.8 m). Each output pixel is a bilinear interpolation of its four source neighbours and is nodata if any neighbour is nodata. No other resampling is used, and the same code runs in Python

and in the browser, where a parity test on shared fixtures holds the two implementations to the tolerance stated in each fixture. The fingerprint uses a 90 m grid; the encoder uses 64 by 64 pixels at 78.125 m.

Encoder dataset and splits. Encoder windows are sampled uniformly in projected area from the lunar model within the cap south of 80 degrees, from the Martian models that pass the quality threshold, and from Earth tiles at a fixed number of windows per tile, with the Earth land cover rule applied. They are split by space, never at random: lunar longitude sectors, groups of Martian models, and 2 degree Earth blocks. Windows that straddle a block boundary are dropped, and so is every window within 10 km of the lunar target or of the Jezero box, in every split; the Arcadia footprint is not set, so no exclusion around it could be checked. The dataset holds 28,843 windows. The training split has 6,848 lunar, 1,398 Martian and 11,147 terrestrial windows; validation has 2,418, 364 and 2,760; and test has 2,937, 726 and 244; the lunar target window is stored outside the splits. These counts are far below the dataset size planned for the encoder; the dataset was built on a CPU server at a reduced scope, and its small size limits every encoder result.

4 Physics fingerprint

The fingerprint is the deterministic description of a window that every ranking in this paper uses. It follows the recommendation of Shepard et al. (2001) to report roughness at stated measurement scales and the baseline conventions of Rosenburg et al. (2011), measured with one code path on every body.

Structure function. Let z be the elevation of a window on the analysis grid, with validity mask m. For a lag L along a direction d, let $P_d(L)$ be the set of cell pairs (a, b) with $b = a + L\,\hat{d}$ and both cells valid. The structure function is

$$\nu_d(L) = \Big(\frac{1}{|P_d(L)|} \sum_{(a,b)\in P_d(L)} \big(z(b) - z(a)\big)^2 \Big)^{1/2}. \tag{1}$$

It is measured along both grid axes and both diagonals. A lag is used along a direction only if it is a whole number of steps there to arithmetic tolerance; lags are never rounded to a nearby offset and cells are never interpolated. On the 90 m grid no diagonal offset equals a configured lag, so the diagonals are recorded as unsupported and the isotropic value rests on the two axes. The isotropic summary weights the supported directions $\mathcal{D}(L)$ equally,

$$\nu(L) = \Big(\frac{1}{|\mathcal{D}(L)|} \sum_{d\in\mathcal{D}(L)} \nu_d(L)^2 \Big)^{1/2}, \tag{2}$$

and the per direction values are kept beside it. The configured lags are 270, 360, 540, 720 and 1080 m: at least three analysis cells and at most a quarter of the 5000 m window. A window is measured only if at least 80% of its cells are valid, and an unsupported lag has no value, never a value of zero.

Relief and texture. Each window is measured twice. Relief is Equation (1) on the raw heights, tilt included. Texture is the same statistic on the residual $z - (a + bx + cy)$ after the least squares plane over the valid cells is removed, with x east and y north in metres from the window centre. The plane is solved exactly over the stored values, so an exact plane leaves a texture of exactly zero

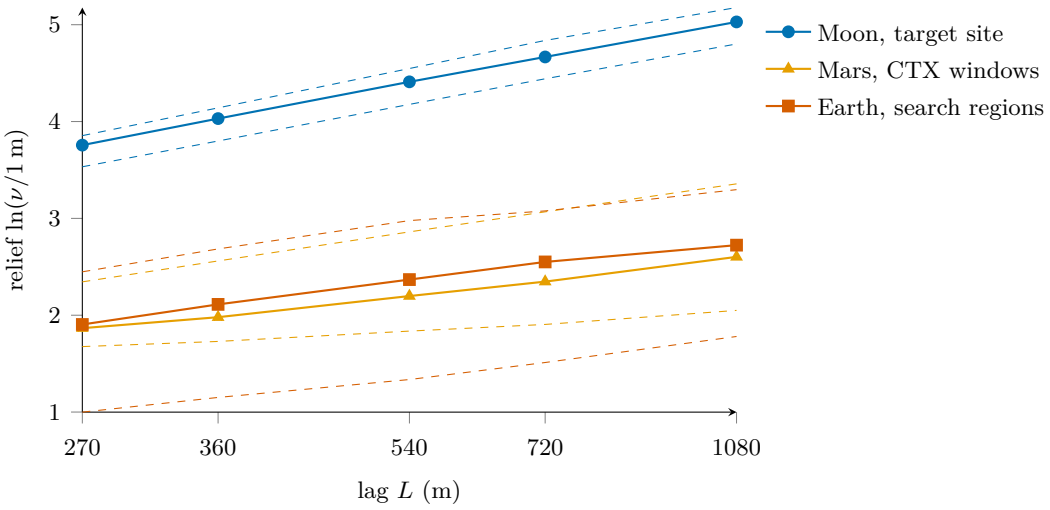


Figure 1: Derived relief structure functions per body over the windows that the search read: the 25 windows of the lunar target site, 359 windows from 2 CTX models and 71 Earth windows from the declared regions. Solid lines are medians of $\ln \nu$ over windows and dashed lines the quartiles. The samples are small and contiguous, so the curves describe these windows, not the bodies. Generated from the cross body record of Appendix A.

rather than rounding noise. Both variants are named in every record, because removing the plane can erase the very property being compared.

Hurst exponent. H is the ordinary least squares slope of $\ln \nu(L)$ against $\ln L$ over the supported lags,

$$H = \frac{\sum_k (\ell_k - \bar{\ell})(n_k - \bar{n})}{\sum_k (\ell_k - \bar{\ell})^2}, \qquad \ell_k = \ln L_k, \quad n_k = \ln \nu(L_k), \tag{3}$$

which equals the exponent for $\nu = cL^H$. It is fitted only with at least 4 supported lags that span at least a factor of 4, and only if ν is positive at every lag. Otherwise H is unavailable with its reason; a zero is never offset to take its logarithm. Over our five lags H describes the range from a few hundred metres to about a kilometre, the range in which Rosenburg et al. (2011) report a breakover on the Moon, so Appendix A tests for a break over a wider range of lags at the scale of the whole site rather than assuming a single exponent.

Figure 1 shows the relief structure function per body over the windows that the search read. The lunar site lies well above the Martian and terrestrial samples at every lag, with a slope close to the Hurst exponents that Rosenburg et al. (2011) report for the lunar highlands, while the Martian and Earth samples overlap.

Terrain distance. Two windows u and w are compared on the lags S_{uw} supported in both, separately for relief and texture:

$$D(u, w) = \left(\frac{1}{|S_{uw}|} \sum_{L \in S_{uw}} \left(\ln \nu_u(L) - \ln \nu_w(L) \right)^2 \right)^{1/2}. \tag{4}$$

A difference of $\ln \nu$ is a ratio of relief, so D does not depend on the units of height and weighs a factor of two the same at every lag. Relief and texture distances are never summed with each

other, and in this study no radar distance exists to sum them with (Section 8.7). Candidates are ordered by Pareto layers over the two distances, with the window identifier as a deterministic tie break; no single score is formed.

Reference and uncertainty. Each distance is meant to be read against two references, and Section 8.2 states which were computed. The target reference set is the spread of distances between the primary target window and the other eligible Site01 windows, which measures how much the target varies internally. The background is a fixed, seeded sample, configured at 1,000 eligible windows from the same search universe, against which a candidate's distance is a percentile. For the Moon the provider's 100 error clones of the elevation model (Barker et al., 2021) are measured like the original, so each statistic has a spread under the published height error model. The clones represent that error model only; they say nothing about rocks below the grid or about radar. Candidates whose distances differ by less than their combined spread are reported as tied, and the stability of a rank across clones is a frequency, never the probability that a site is a valid analog.

5 Model architecture

The encoder (Figure 2) asks whether a learned representation of elevation, trained on all three bodies and supervised by the fingerprint, can reproduce the fingerprint on held out regions, carry morphology that the fingerprint lacks, and give cross body neighbours that agree with it. It is a second opinion beside the fingerprint, labelled Predicted wherever it appears, and it enters no ranking unless it meets the acceptance criteria of Section 8.1.

Input. A window is resampled to 64 by 64 pixels at 78.125 m (Section 3) and a best fit plane is removed. The three channels are $\operatorname{asinh}(z/s)$ of the plane removed heights with $s = 50\,\mathrm{m}$, which compresses large relief while staying linear near zero; the slope magnitude of the heights before plane removal, divided by 45 degrees; and the validity mask. The same code, in Python and in the browser, feeds training and the live probe.

Backbone and heads. The backbone is a ResNet-18 (He et al., 2015) at half width, with stages of 32, 64, 128 and 256 channels, batch normalisation (Ioffe and Szegedy, 2015) and a three by three stem without max pooling for the small input. The specified stem stride is one; a stride of two, which reduces the spatial computation of the backbone and leaves the parameter count unchanged, is used in every reported run and recorded as a deviation. Global average pooling gives the embedding $h \in \mathbb{R}^{256}$, which is L2 normalised for retrieval. A three layer projector of width 2048 maps h to z for the VICReg loss. A two layer physics head predicts the eleven standardised physics targets: $\ln \nu$ of relief and texture at the five lags and the relief Hurst exponent. A two layer body head predicts which of the three bodies a window came from, behind a gradient reversal layer (Ganin et al., 2015).

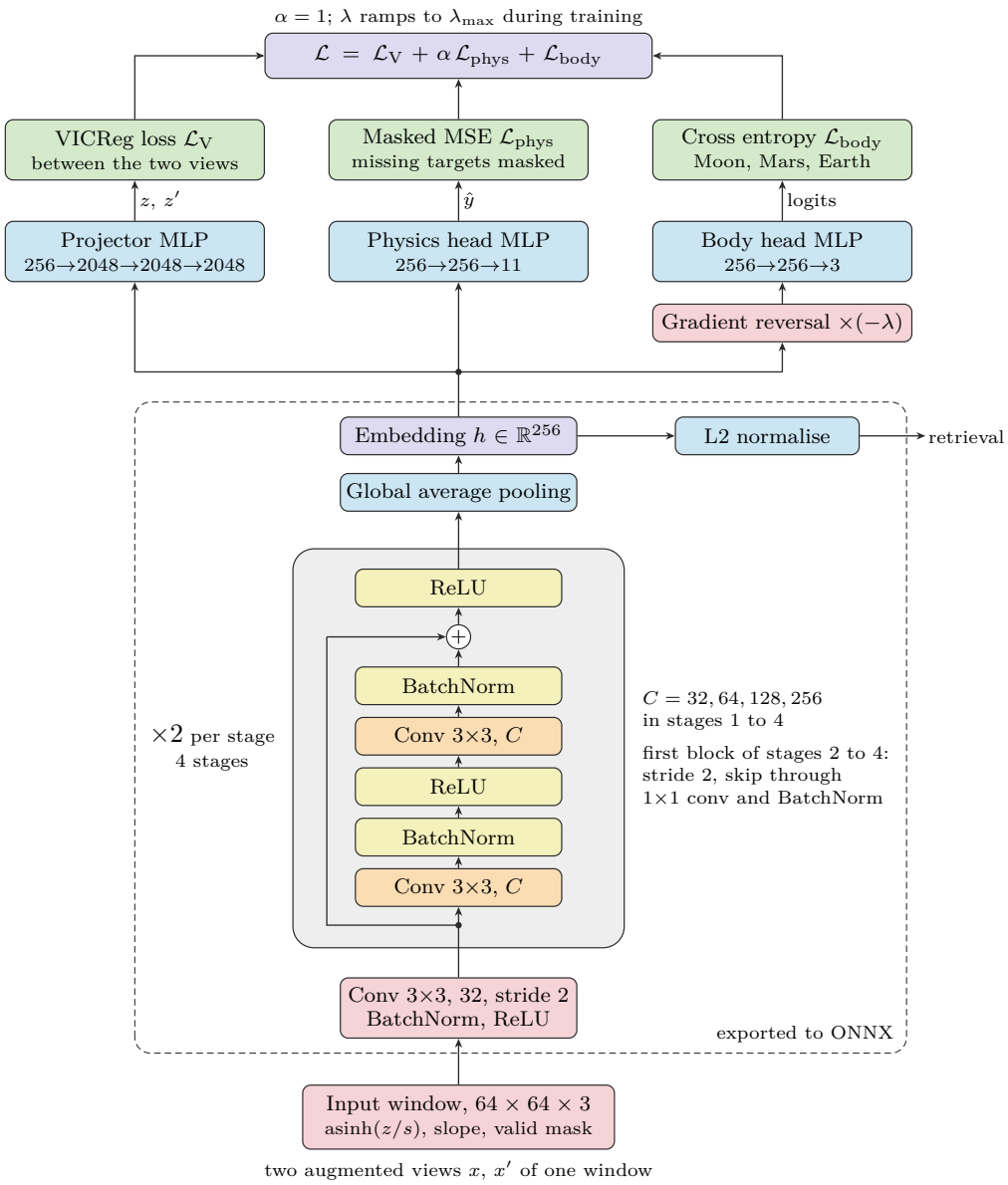


Figure 2: Encoder architecture, drawn from input (bottom) to loss (top). A ResNet-18 layout at half width maps a 64 by 64 window with three channels to a 256 value embedding h. The grey frame is the repeated BasicBlock, two per stage in four stages. Three heads train the backbone: a projector for the VICReg loss on two views of one window, a physics head that regresses the fingerprint of Section 4 with a masked loss, and a body head behind a gradient reversal layer, through which λ acts. Only the backbone with the L2 normalisation (dashed) is exported to ONNX; the heads are used in training only. The stride 2 stem is a recorded deviation from the specified stride 1.

Objective. For two augmented views of one window with projections z, z' over a batch of n windows and dimension d, the VICReg loss (Bardes et al., 2021) is

$$\mathcal{L}_{\mathrm{V}} = \lambda_{\mathrm{inv}} \frac{1}{nd} \sum_i \|z_i - z'_i\|_2^2 + \lambda_{\mathrm{var}} \big(v(z) + v(z')\big) + \lambda_{\mathrm{cov}} \big(c(z) + c(z')\big), \tag{5}$$

$$v(z) = \frac{1}{d} \sum_{j=1}^{d} \max\big(0,\, 1 - \sqrt{\mathrm{Var}(z_{\cdot j}) + \epsilon}\big), \qquad c(z) = \frac{1}{d} \sum_{j \neq k} \mathrm{Cov}(z)_{jk}^2,$$

with coefficients $\lambda_{\mathrm{inv}} = \lambda_{\mathrm{var}} = 25$ and $\lambda_{\mathrm{cov}} = 1$. The physics loss is a masked mean squared error over the targets y that exist for a window, for both views,

$$\mathcal{L}_{\mathrm{phys}} = \frac{\sum_k m_k \left(\hat{y}_k - y_k\right)^2}{\sum_k m_k}, \tag{6}$$

so a lag without a supported value contributes nothing instead of a zero target. The body loss is the cross entropy of the body head. Gradient reversal multiplies its gradient by $-\lambda$ on the way into the backbone, so the total objective

$$\mathcal{L} = \mathcal{L}_{\mathrm{V}} + \alpha\, \mathcal{L}_{\mathrm{phys}} + \mathcal{L}_{\mathrm{body}}, \qquad \lambda(p) = \lambda_{\max}\Big(\frac{2}{1 + e^{-10p}} - 1\Big), \tag{7}$$

trains the head to recognise the body while the backbone is pushed not to encode it. Here p is the fraction of training completed, $\alpha = 1$, and $\lambda_{\max}$ is swept over 0.0, 0.1, 0.3 and 1.0.

Views. The two views of a window differ by one of the eight rotations and reflections, Gaussian height noise of 1 m mapped through the local derivative of the asinh channel, and up to 10% extra nodata holes. Amplitude and scale changes are excluded, because they change the physics targets that the views must share. Each batch draws a third of its windows from each body.

6 Why learn a terrain embedding

The fingerprint of Section 4 is exact, auditable and cheap, but it is a second order statistic. Two surfaces with the same structure function can differ in what produces it: a field of overlapping craters and a set of parallel ridges may share $\nu(L)$ at every lag while looking nothing alike. A learned embedding can in principle carry such morphology. It also brings risks that the fingerprint does not have: it can encode the instrument, the processing or the body instead of the terrain, and it can find neighbours for reasons nobody can inspect. This section states what each representation encodes, so that the tests of Section 8 ask the right questions of it.

Table 2 compares four representations. The fingerprint encodes exactly the relief and texture structure function and the Hurst exponent at stated lags, and it is invariant to rotation by construction of the isotropic summary. Two frozen image models were planned as baselines: a ResNet-18 trained on ImageNet photographs (He et al., 2015) and a Vision Transformer (Dosovitskiy et al., 2020) trained by DINO self distillation without labels (Caron et al., 2021). Neither has seen elevation data, so they would test whether generic visual features already carry the fingerprint; their weights were not obtained for this study, so their probe results are not available (Section 8.1). The encoder of Section 5 learns from elevation of the three bodies with an explicit physics target and an explicit penalty on body information.

Related learned models make different choices. LunarFM learns a joint embedding of many lunar instruments within a latitude band that excludes the poles (Girona-Mata et al., 2026); topo2vec

Table 2: Representations compared in this study. The training signal and invariances are properties of the method. Measured parameter counts and cost per window are reported only where the model report holds them.

Representation	Encodes	Invariances	Use in the browser
Fingerprint	$\ln \nu(L)$ of relief and texture at five lags; Hurst exponent	the two grid axes and their reflections; distances unchanged by a change of height units	exact, in the probe worker
ImageNet ResNet-18	features of natural photographs	those of photographic augmentation	planned baseline, not evaluated
DINO ViT	self distilled features of natural images	those of its augmentations	planned baseline, not evaluated
Terraluna encoder	elevation features trained on Moon, Mars and Earth	rotations and reflections by training; body information penalised	ONNX export, checked in ONNX Runtime Web; labelled Predicted

learns elevation embeddings with a scale based self supervision (Kavitzky et al., 2021); GEAR retrieves analogs between two terrestrial regions and releases an expert annotated similarity dataset (Liu et al., 2026); and the Qaidam study learns a similarity between Earth and Mars images (Meng et al., 2024). The encoder studied here uses elevation only, three bodies, a south polar lunar target and the fingerprint as supervision, and it is evaluated by probes on held out spatial blocks. Whether it adds anything over the fingerprint is an empirical question with a stated test (Section 8.1), and until that test passes the fingerprint alone ranks candidates.

7 Training

This section describes the training protocol and reports the training record. Every number in it is a generated macro; a value that the training record does not contain yet is printed as *not available* rather than estimated.

Data and batching. Training windows are drawn from the training blocks of the spatial split of Section 3, with each batch balanced so that each body contributes a third of its windows. The training split holds 6,848 lunar, 1,398 Martian and 11,147 terrestrial windows (Section 3); the manifest records an upper estimate of about 1.1 CPU hours for building the dataset. Physics targets are computed for every window by the same fingerprint code that ranks candidates (Section 4) and standardised with statistics of the training blocks only. Morphology labels, the crater densities of the lunar and Martian crater catalogues (Robbins, 2019; Robbins and Hynek, 2012), are reserved for the probes and never used in training. These labels were not built for this study, so the morphology probe is not available.

Hardware and schedule. The runs evaluated on test blocks were trained on a CPU server at a reduced scope, without a GPU. Table 3 gives the schedule as the training record states it; Section 8.5 adds later runs on a laptop accelerator, evaluated on validation blocks only.

The planned grid covers seeds 0, 1 and 2 of the main configuration and the adversarial weights 0.0, 0.1, 0.3 and 1.0; Section 8.4 lists which runs exist.

Table 3: Training schedule of the first main run, from its training record. The planned schedule was 40 epochs at a batch of 1024 in brain floating point with a stem stride of 1; the run used a stem stride of 2 in single precision, and its record lists each deviation.

Item	Value
Hardware	CPU only, 8 virtual CPUs
CPU threads per run	2
Epochs	12
Windows per epoch	10,000
Training rows	19,393
Batch size	128
Wall clock hours, first main run	1.6
CPU core hours, first main run	3.2
Backbone parameters	2,795,040

Optimiser. We use AdamW (Loshchilov and Hutter, 2017) with a learning rate scaled linearly with the batch size after a warm up (Goyal et al., 2017), cosine decay of the learning rate in the manner of Loshchilov and Hutter (2016) without restarts, and decoupled weight decay. The adversarial weight follows the schedule of Equation (7). Model selection uses validation blocks only; the test blocks are read once, after the configuration is frozen, by a logged command.

Augmentation and regularisation. The views of Section 5 are the only augmentation. They contain no amplitude or scale change, because the physics targets must be identical for both views of a window. VICReg's variance and covariance terms regularise the embedding against collapse (Bardes et al., 2021), and the masked physics loss acts as a supervised regulariser that ties the embedding to the fingerprint.

8 Results

All results come from the recorded artifacts named in the generated macro file, and a value that an artifact does not contain is printed as *not available*. The study is terrain only: the radar products were decoded, but no radar pair was qualified and no radar distance enters any ranking.

8.1 Encoder probes on test blocks

The encoder would be accepted as a third distance family only if, on test blocks, the lower bound of the physics probe R^2 reaches 0.8 for relief $\ln \nu$ at every lag, the morphology probe gains at least 0.10 in R^2 over a ridge on the fingerprint for the Moon or Mars, a logistic body probe stays at or below a balanced accuracy of 0.60, the Spearman correlation between embedding and fingerprint distances reaches 0.5, the top twenty neighbours agree across seeds with a Jaccard index of at least 0.5, and the encoder beats both frozen image baselines on the two probes with paired bootstrap intervals that exclude zero, the morphology gain having a positive lower bound, and the exported model matches the training model within the parity tolerance. These acceptance criteria were fixed in the configuration before training. Table 4 reports the probes from the evaluation record. The encoder is reported as experimental and every ranking in this paper uses the fingerprint alone. The physics probe passes its threshold; the body probe, the agreement with the fingerprint and the seed stability are reported with their thresholds in the table; and the morphology gain and the comparison with the image baselines are not available, because they were not evaluated.

Table 4: Probes on test blocks. Physics: the smallest lower bound of the probe R^2 over the relief lags. Morphology: the R^2 gain over the fingerprint ridge. Body: balanced accuracy of the body probe, where chance is one third. Agreement: Spearman correlation of embedding and fingerprint distances on cross body pairs. Seeds: mean Jaccard index of the top twenty Earth neighbours across seeds, against 0.04 for a random ranking. Not available marks a probe that was not evaluated.

Representation	Physics	Morphology	Body	Agreement	Seeds
Random initialisation	0.93	*not available*	0.87	0.85	*not available*
ImageNet ResNet	*not available*	*not available*	*not available*	*not available*	*not available*
DINO	*not available*	*not available*	*not available*	*not available*	*not available*
Terraluna encoder	0.87	*not available*	0.71	0.20	0.32
No gradient reversal	0.87	*not available*	0.72	0.21	*not available*
Acceptance threshold	0.8	0.10	0.60	0.5	0.5

The physics probe passes its threshold: the smallest lower bound of the relief R^2 over the lags is 0.87. But an untrained copy of the same network reaches a smallest lower bound of 0.93, so a linear readout of the fingerprint from these windows does not show that training added anything. The body probe classifies the body with a balanced accuracy of 0.71 (interval 0.58 to 0.85), above the threshold of 0.60 and close to the 0.72 of the run without gradient reversal, so the adversarial head at the trained weight did not remove body information. Agreement between embedding distances and fingerprint distances is 0.20 (interval 0.08 to 0.36), against 0.85 for the untrained network: training moved the embedding away from the physics that the fingerprint measures. The top twenty Earth neighbours agree across seeds with a mean Jaccard index of 0.32, above the 0.04 of a random ranking but below the threshold. The morphology probe and the comparison with the ImageNet and DINO baselines are not available: no crater labels were built and the baseline weights were not obtained, so neither was evaluated. At this reduced scope the paper makes no claim for learned similarity. The model report is built in the model card form of Mitchell et al. (2019), with intended use, evaluation conditions and limitations stated beside the metrics.

The browser export can be checked independently of the embedding quality. An ONNX export of the backbone of the first main run (opset 17; training run status in the export record: completed) reproduces the training model on 256 validation windows with a largest absolute difference of 3.7e-07 in single precision (tolerance 1e-04) and a smallest cosine similarity of 0.999996 in half precision (tolerance 0.999), for a half precision file of 5,601,753 bytes. This shows that the exported graph computes the same function; it says nothing about the quality of the embedding. The web probe's own check runs the half precision file in ONNX Runtime Web 1.31.1 under Node v22.23.3 (WASM) with 1 thread on 256 windows, with a smallest cosine similarity to the single precision Python model of 0.999995 (tolerance 0.999) and 5.9 ms per window; that timing comes from Node.js, not from a browser on a typical laptop.

8.2 Cross body retrieval for the lunar target

The lunar target window was compared with 71 eligible Earth windows from the two declared regions. Only the relief and texture distances of Equation (4) were comparable; no radar distance exists (Section 8.7) and the encoder distance was not evaluated, so both are not available. Over the compared windows the relief distance ranges from 0.68 to 4.95 with a median of 2.05, and the texture distance from 0.66 to 4.92 with a median of 2.21. The first Pareto layer holds 1 window, `earth-utm27n-e285-n2879` at Holuhraun, with a relief distance of 0.68 and a texture distance of 0.66. When the target is replaced by each of its 100 error clones, the relief distance stays between

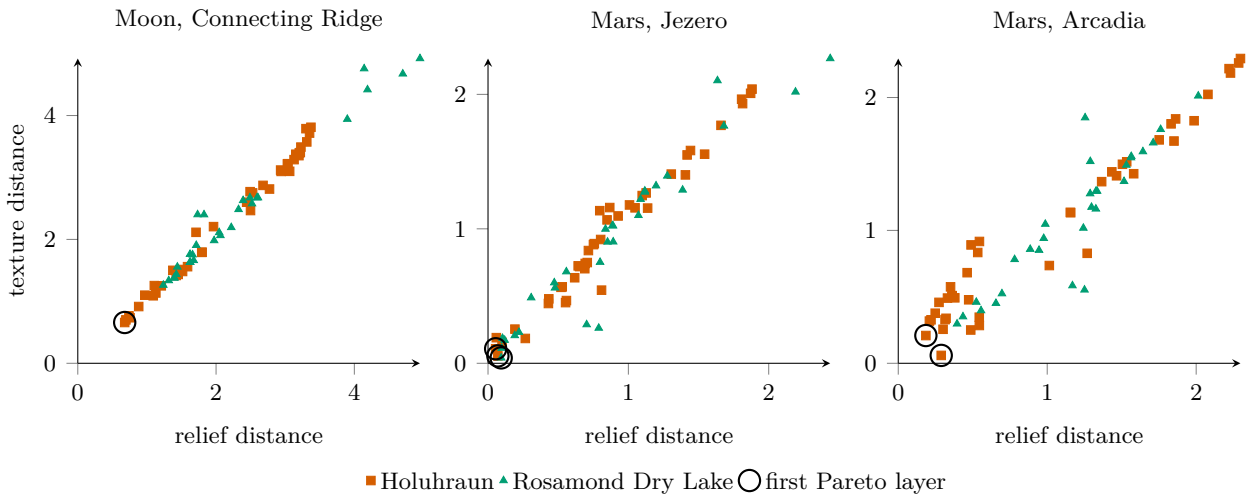


Figure 3: Derived relief and texture distances, in RMS of $\ln \nu$, of the 71 compared Earth windows to each target window, by search region, from the comparison records of the three targets. Circled windows form the first Pareto layer of each target. Smaller is closer in each distance; the two are never summed. The axes differ between panels. Drawn by `paper/scripts/tables.py` from the recorded comparison and target records.

0.6781 and 0.6784, and the window ranks first by relief in 100 of 100 clone runs. The rank is also unchanged when the land cover rule is tightened moderately or only the Holuhraun region is searched; when every analysis cell must be in an allowed class, or only Rosamond Dry Lake is searched, the window leaves the evaluated set. This stability is a frequency over the published height error model, not a probability that the window is an analog. It mostly shows that the clones move the target's statistics very little. The relief distance has a percentile of 3.1 in the Earth background, which holds only the 71 windows of the search universe rather than the 1,000 requested. The target reference set, the distances between the target and the other 24 eligible windows of its site, has a median of 0.17 and a maximum of 0.44 in relief, so the nearest Earth window is farther from the target than every other window of the target site.

Figure 3 shows the two distances for every compared window.

We inspected this window as a counterexample. The target is one broad massif whose relief keeps growing with lag, while the candidate is a mostly flat plain crossed by narrow ridges and a trough, whose relief is close to the target's at the shortest lags but grows much more slowly with lag. Equation (4) averages over lags and does not read the Hurst exponent, so a window that matches at short lags can rank first against a target that is rougher at long lags. The shortlisted window is therefore the nearest window in this terrain metric among the evaluated windows, a candidate window identified by this search, and not a morphological analog of Connecting Ridge.

Table 5 and Figure 4 show the fingerprints behind these distances for the target and for the first Earth window that was paired with it, the window that contains the decoded radar scene centre. The lunar window is much rougher at every lag: its relief ν grows from 43.4 m to 148.2 m over the lags, against 1.66 m to 4.76 m at Holuhraun, which gives this pair a relief distance of 3.37 and a texture distance of 3.81. The Hurst exponents of the lunar window, 0.89 for relief and 0.89 for texture, lie near the highland median of Rosenburg et al. (2011), although their baselines and estimator differ from ours. Removing the plane barely changes the lunar statistics, while it lowers the Holuhraun texture exponent to 0.39, so at the longer lags much of this window's relief comes

Table 5: Derived fingerprints of the lunar target window and the first paired Earth window, on the 90 m grid over the 5000 m window. ν is given at the shortest and longest configured lag.

Window	relief ν short	relief ν long	texture ν short	texture ν long	H relief	H texture
Moon, target	43.4	148.2	42.0	142.9	0.89	0.89
Earth, Holuhraun	1.66	4.76	1.29	2.21	0.76	0.39

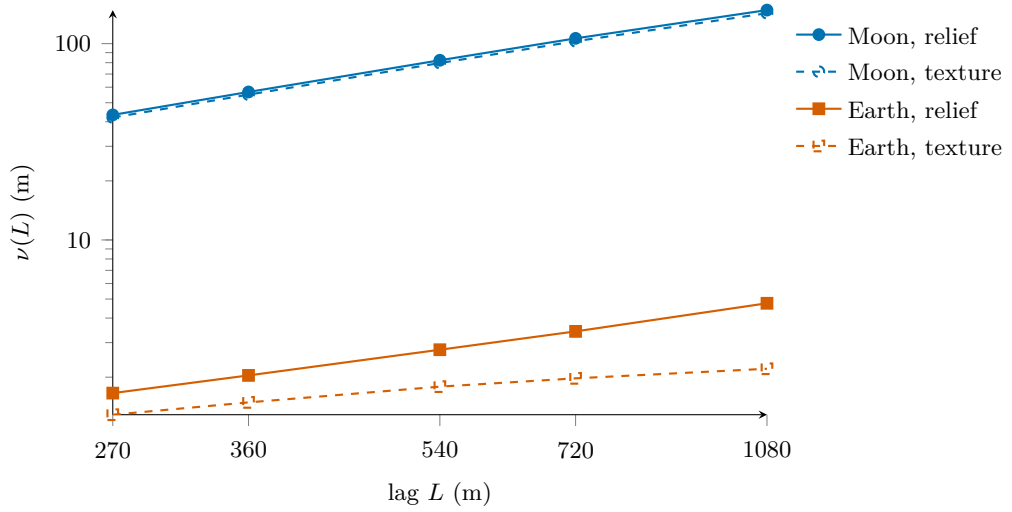
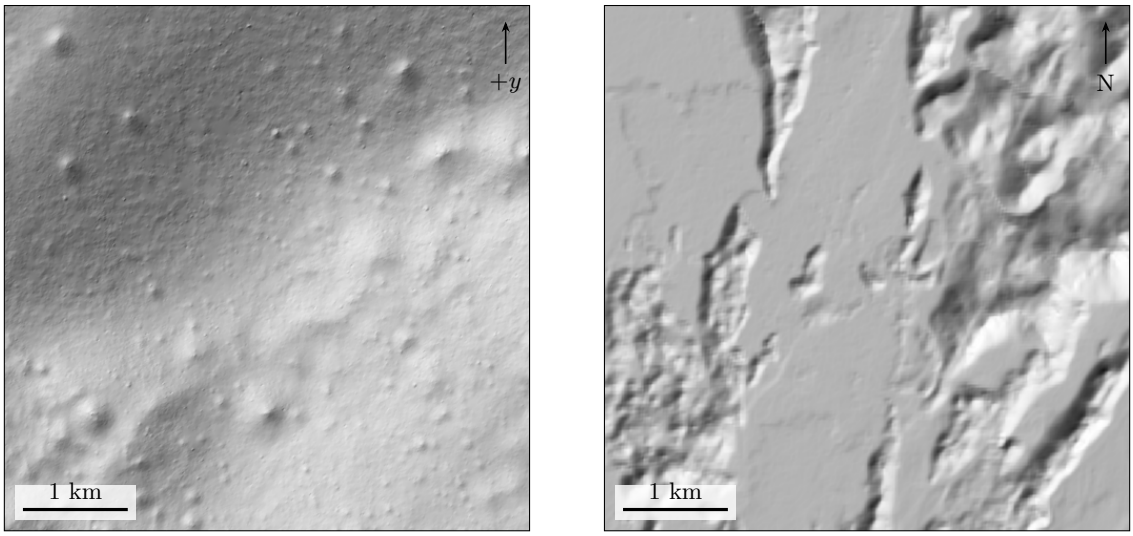


Figure 4: Derived structure functions of the primary lunar target window (Site01 of PGDA product 78, window `moon-sps-x2-y2`) and the first paired Earth window (Copernicus GLO-30, `earth-utm27n-e280-n2880`, Holuhraun), on the 90 m analysis grid, from the recorded windows evidence of the data inventory. Solid lines are relief (raw heights) and dashed lines texture (best fit plane removed). Each curve is the root mean square of the two grid axis values, weighted equally. The slopes give the Hurst exponents of Table 5. The Earth window was chosen because it contains the decoded radar scene centre, not by similarity.

from its regional tilt. Holuhraun lava facies have been studied with UAVSAR CPR as a measure of radar roughness (Tolometti et al., 2022), yet at the hectometre scale of this fingerprint it is far smoother than the lunar target. Radar scattering and the fingerprint describe different scales of one surface, and this study ranks on the fingerprint only.

Figure 5 shows the two windows of the shortlist pair. The lunar window spans 813 m of relief in a cratered massif, and the Earth window spans 307 m at the edge of the Holuhraun region, with a flat floor between narrow ridges. The pair is the closest in the evaluated set by both distances, yet the hillshades differ in morphology, which is the kind of difference that a second order fingerprint does not capture and that the encoder is meant to test.

The same gap holds across bodies. Over the windows that the search read, the relief $\ln \nu$ distributions of the lunar site and of Earth overlap with a coefficient of 0.07 at the shortest lag and 0.00 at the longest, while the Martian and Earth samples overlap at 0.51 and 0.67 (Appendix A). At the longest lag no evaluated Earth window falls in the range of the lunar site, so for this target the shortlist names the least different window rather than a similar one.



(a) Moon, Connecting Ridge target window

(b) Earth, shortlisted Holuhraun window

Figure 5: Derived hillshades of the lunar target window and the shortlisted Earth window, each 5000 m wide, lit from azimuth 315 degrees at 45 degrees altitude with no vertical exaggeration. (a) Window `moon-sps-x2-y2`, centred at latitude −89.44 and longitude −137.39 degrees, from the 5 m surface heights of PGDA product 78 Site01 (Barker et al., 2021), shown at every second pixel in the product's south polar stereographic grid; the arrow is grid north ($+y$). Most 5 m pixels are interpolated between LOLA tracks (Section 3). (b) Window `earth-utm27n-e285-n2879`, centred at latitude 64.85 and longitude −16.46 degrees, resampled from Copernicus GLO-30 tile N64 W017 (Airbus Defence and Space, 2022) onto a 10 m grid about its centre with the window resampler; the arrow is north. Generated by `paper/scripts/figures.py`, which records the input digests.

8.3 Cross body retrieval for the Martian targets

Two Martian target windows were fixed by rules declared before ranking. At Jezero crater, the primary window is the eligible window nearest the landing site of the Perseverance rover on the CTX model in the declared box that passes the quality threshold; its site has 195 eligible of 234 generated windows. In Arcadia Planitia, it is the eligible window nearest the centre of the CTX model that passes the threshold nearest the box centre, with 164 eligible of 199 windows. Each was compared with the same 71 Earth windows as the lunar target (Figure 3).

The Martian targets sit much closer to the Earth sample than the lunar target. For Jezero the first Pareto layer holds 3 windows, with relief distances from 0.054 to 0.094 and texture distances from 0.040 to 0.108, against a median relief distance of 0.80 over all compared windows. For Arcadia it holds 2 windows, with relief distances from 0.185 to 0.289 and texture distances from 0.059 to 0.209, against a median of 1.16. These shortlisted distances lie below the median internal distance of each Martian site, whose target reference sets have median relief distances of 0.93 at Jezero and 0.39 at Arcadia, so under this fingerprint the shortlisted Earth windows lie within the spread of windows of the target sites themselves. The lowest background percentile of a compared window is 5.0 for Jezero and 2.6 for Arcadia, again over a background of only the 71 windows of the search universe. The Martian models have no published error clones, so these distances carry no measurement error spread, and no counterexample inspection was recorded for the Martian shortlists. The shortlisted windows are candidate windows identified by this search for the two Martian targets, with the same caveat as for the Moon: a small distance at hectometre lags says

Table 6: Encoder variations, from the evaluation record where it exists. Physics, body and agreement as in Table 4.

Variation	Physics	Body	Agreement
Main configuration	0.87	0.71	0.20
No physics head	0.85	0.71	-0.04
No gradient reversal	0.87	0.72	0.21
Weaker gradient reversal	*not available*	*not available*	*not available*
Stronger gradient reversal	0.89	0.72	-0.06
VICReg only	*not available*	*not available*	*not available*
Physics head only	*not available*	*not available*	*not available*

nothing about morphology, block sizes or the surface at the scale of a rover.

8.4 Model variations

Table 6 lists the planned variations of the encoder, in the manner of a model variations table, with the results that the evaluation record holds for each.

8.5 Benchmarks

Validation sweep. After the test evaluation above, we trained additional runs and evaluated them, together with the earlier runs, on the validation blocks only; the test blocks were not read again, and no number in this subsection was used to select a model on test. The probes are those of Table 4, with one caveat recorded in the evaluation record: the ridge penalty of the probes was chosen on the same validation rows that are scored, so these values are optimistic. Table 7 lists every run. Three findings hold across the table. The body probe stays between 0.62 and 0.66 for every trained run, whatever the adversarial weight, including the run without gradient reversal, so the adversary at these weights does not remove body information; every value lies above the acceptance threshold of 0.60. The untrained network reaches a physics probe of 0.93 and an agreement of 0.81, against at most 0.47 for any trained run: training lowers the agreement between embedding and fingerprint distances instead of raising it. And training the main configuration for 40 rather than 12 epochs lowers the agreement of every seed, from 0.39 to 0.47 at 12 epochs to 0.28 to 0.34 at 40 epochs. The mean Jaccard index of the top twenty Earth neighbours across three seeds is 0.60 for the 40 epoch runs and 0.77 for the 12 epoch group `lambda_1`, against 0.004 for a random ranking; both rest on a single target window, and the two groups differ in adversarial weight as well as in length, so they do not isolate the effect of training length.

Training throughput. The run records give the measured training throughput per epoch at a batch of 128 with the same model and data. On a server with 8 virtual CPUs, the first main run trained at a median of 34 windows per second (7 to 44 over its epochs). On a laptop accelerator through the PyTorch Metal Performance Shaders backend, a 12 epoch run of the sweep trained at a median of 1157 windows per second (997 to 1392), and a 40 epoch run at 1077 (875 to 1362), finishing in 391 s of wall clock time. The CPU throughput varied widely between epochs, and the records do not state why. The reduced schedule of Section 7 was set for the CPU server; on the accelerator the longer runs of Table 7 took minutes, and on validation blocks they still miss the body and agreement thresholds.

Table 7: Validation benchmark of the encoder (2,795,040 backbone parameters). Validation blocks only; the probe penalty was chosen on the scored rows, so the values are optimistic. Physics: smallest lower bound of the relief probe R^2 over the lags. Body: balanced accuracy of the body probe, where chance is one third. Agreement: Spearman correlation of embedding and fingerprint distances. $\lambda_{\max}$ is the adversarial weight of Equation (7).

Run	$\lambda_{\max}$	Epochs	Seed	Physics	Body	Agreement
Untrained network		0	0	0.93	0.76	0.81
No gradient reversal	0	12	0	0.93	0.66	0.46
Weaker gradient reversal	0.1	12	0	0.92	0.65	0.46
Main configuration	0.3	12	0	0.92	0.65	0.47
Main configuration	0.3	12	1	0.90	0.66	0.47
Main configuration	0.3	12	2	0.92	0.63	0.39
Stronger gradient reversal	1	12	0	0.92	0.65	0.26
Stronger gradient reversal	1	12	1	0.89	0.65	0.36
Stronger gradient reversal	1	12	2	0.90	0.64	0.21
No physics head	0.3	12	0	0.92	0.65	0.20
Main configuration, longer	0.3	40	0	0.93	0.65	0.34
Main configuration, longer	0.3	40	1	0.90	0.64	0.32
Main configuration, longer	0.3	40	2	0.92	0.62	0.28
Acceptance threshold				0.8	0.60	0.5

Browser parity. The half precision ONNX encoder run in ONNX Runtime Web 1.31.1 reproduces the single precision Python model on 256 windows with a smallest cosine similarity of 0.999995 against a tolerance of 0.999, at 5.9 ms per window in Node v22.23.3 (WASM) with 1 thread (Section 8.1). The fingerprint and the window resampling have a TypeScript port that the browser worker runs; shared fixtures hold it to the Python implementation within the tolerances recorded in each fixture file, which are not reported here as generated values. The latency of the fingerprint and of the encoder in a browser on typical hardware is not available, because neither was timed in a browser.

8.6 Held out body

A run trained on the Moon and Earth only and probed on Mars would test whether the physics probe transfers to a body the encoder has not seen. The evaluation record holds no such run, so the smallest relief probe R^2 on held out Mars is *not available*.

8.7 Radar context for the lunar target

Both radar products were decoded, but the radar comparison was not completed, and no radar number enters any distance. On the lunar side, the Mini-RF south polar CPR mosaic was read over the site extent, with 14,750 valid and 3,746 nodata pixels. The mosaic combines east and west looking passes from several acquisitions, so the incidence angle of a pixel is not recoverable from it, and its label does not state the transmit handedness. On the Earth side, a region of interest of 2000 m at Holuhraun was decoded from the UAVSAR L band covariance product, with 100,317 valid pixels at 36 nominal looks and a median local incidence of 61.2 degrees. The median ratio of the handedness dependent term to the base same sense power is 0.088, but the handedness convention is not verified against the Mini-RF convention. The radar conversions pass their checks against independent fixtures, but the eligibility check of band, convention, geometry and support

was not run for any pair, so no radar pair is qualified, no radar distance exists, and no Earth value is presented as equivalent to Mini-RF.

As a literature check of the Earth radar processing, we repeated the facies comparison of Tolometti et al. (2022) on the same UAVSAR scene, with 2,168,388 valid pixels inside the mapped facies after the incidence and nodata rules of a protocol fixed before any result was read. The median CPR is 0.40 for rubbly, 0.40 for spiny and 0.42 for undifferentiated lava. The pairwise distribution overlaps are 96%, 93% and 92% (rubbly and spiny, rubbly and undifferentiated, spiny and undifferentiated). Applied to the published radar values of that study, the same overlap definition gives 83%, 89% and 94%. Two of our three overlaps exceed the published comparators and one is slightly lower, so our facies are not markedly better separated than the published ones, which is the plausibility condition of the check, but the overlaps differ in detail; the authors' mosaic combined two data takes, their histogram binning and whether their low pass filter acts on CPR or on the covariance are not stated, and the handedness convention is not verified against theirs. The literature check is only partial, because its shadow behaviour part is unavailable without a permanent shadow product; its Holuhraun part is recorded as consistent. Mini-RF and UAVSAR also differ in band, look geometry and processing, so even a completed comparison would show the signatures side by side rather than as a distance.

9 Limitations

Scale range. The fingerprint covers lags of 270, 360, 540, 720 and 1080 m inside 5000 m windows on a 90 m grid. It says nothing about the decimetre to metre roughness that governs walking, driving or radar scattering, and the Hurst exponent describes only that range of lags. Roughness is scale dependent (Shepard et al., 2001; Kreslavsky et al., 2013), so agreement at hectometre lags does not imply agreement at the scale of a boot or a wheel, and no continuous roughness spectrum is measured. No contact scale property, such as block size, soil strength or trafficability, is inferred.

The lunar target is largely interpolated. For only 10.4% of the target window's valid analysis cells does the nearest 5 m pixel hold a LOLA return; elsewhere the heights are the provider's interpolation between tracks (Barker et al., 2021). The fraction of 90 m analysis cells that contain a return anywhere inside the cell is not available, because only the nearest product pixel was checked, and the short lag statistics of the target depend on the interpolation. The provider's error clones measure this uncertainty under the published error model; they were propagated to the shortlist distances, but not to the background or to the encoder.

Earth surface model. Copernicus GLO-30 is a digital surface model, so canopy and buildings are part of the height (Airbus Defence and Space, 2022), and so is any snow present when the elevation was acquired. The land cover mask keeps windows that are mostly bare or sparse, grassland, shrubland, or snow and ice, for at least 90% of their analysis cells, so a retained window can still contain other classes, low vegetation is kept, and the mask is a single year of a land cover product at the analysis cell centres. Surface change since the elevation was acquired, seasonal snow and moisture were not checked.

Mars. The Martian elevation models are stereo products of the Ames Stereo Pipeline (Beyer et al., 2018), and their quality varies between models; models are filtered by their CTX minus MOLA height differences, and stereo artefacts that survive the filter can mimic roughness. The Martian

elevation models have no published error clones, so the Martian distances carry no measurement error spread, and the two Martian target sites each rest on a single CTX model.

Radar. The radar comparison was not completed. The Mini-RF mosaic mixes viewing geometries and states no transmit handedness, the UAVSAR handedness convention is not verified against Mini-RF, and no radar pair has passed the eligibility check of band, convention, geometry and support. Even a completed comparison would face a band difference and the ambiguity that CPR is not monotonic with rock abundance (Rivera-Valentín et al., 2026).

Search universe and independence. Two regions were searched, and overlapping windows on a half window stride are not independent samples. A short distance makes an Earth window a candidate for interpreting the lunar signature under these observables, not a confirmed analog, and a window that was not evaluated is not a negative. For the lunar target the only shortlisted window lies farther from the target than any other window of the target site and is not a close analog; any analog statement in this paper is limited to terrain statistics at the stated lags.

The encoder. The encoder was trained on only 28,843 windows at the reduced scope of Section 7. An untrained copy of the network is a strong baseline on these probes. The validation sweep of Section 8.5 chose the probe penalty on the scored rows and rests its seed stability on one target window. On test blocks the encoder's physics probe passes its threshold, its body probe, agreement and seed stability are below theirs (Table 4), and the morphology probe and the image baselines are not available. A run with a held out body is not available either. The encoder is reported as experimental, labelled Predicted, and kept out of every ranking. Its probes test agreement with our own fingerprint; they do not show that it captures geology.

10 Conclusion

A fingerprint of relief and texture, measured with one code path on the Moon, Mars and Earth, turns the analog question into a set of auditable distances. For the Connecting Ridge target it identifies one candidate window in the Holuhraun search region among 71 evaluated windows, at a relief distance of 0.68 against a median of 2.05 over the compared windows; that window is farther from the target than any other window of the target site, and inspection shows it to be a nearest neighbour in this metric rather than a close analog. For the two Martian targets the same fingerprint finds Earth windows within the internal spread of each site, a closer correspondence at hectometre lags that still says nothing about morphology at the scale of a rover. The radar audit documents what a radar comparison would need: a verified handedness convention on both sides, per pixel incidence and a qualified pair. The encoder, trained with a physics target and gradient reversal across three bodies at a reduced scope, was measured against acceptance criteria fixed before training: its physics probe passes its threshold, an untrained copy of the network is a strong baseline, and body invariance, agreement with the fingerprint and seed stability are below their thresholds (Table 4); the morphology probe and the image baselines are not available. The validation sweep shows that this is not a matter of the adversarial weight, since the body probe stays between 0.62 and 0.66 at every weight, and that more epochs move the embedding further from the fingerprint. It stays an experimental, labelled second opinion outside every ranking, and a fair test needs a larger dataset, crater labels for the morphology probe and the two image baselines.

The results are candidates, and field work is what would test them. Surveyed topography of the candidate window at metre and decimetre resolution would characterise its roughness at the

scales that matter for surface operations, and comparable lunar measurements would be needed to test agreement at those scales. Clast size and coverage counts, and notes on moisture and material, would connect the elevation statistics to the ground. A repeat radar acquisition under controlled geometry, compared with Mini-RF under a declared convention, after both sides are qualified for band, convention, geometry and support, would show the scattering signatures side by side; across bands they remain context rather than a test of a match. Until such observations exist, the shortlist shows where this metric points, and the inspection of its one candidate shows why every candidate must be looked at before it is visited. Because the fingerprint runs in a browser within the tolerances of Section 8.5, it can be computed for other Earth locations from public elevation data, as a live estimate and not a site assessment; its latency in a browser is not available.

Data. Earth elevation was produced using Copernicus WorldDEM-30 © DLR e.V. 2010–2014 and © Airbus Defence and Space GmbH 2014–2018, provided under COPERNICUS by the European Union and ESA; all rights reserved. Land cover is ESA WorldCover 2021 v200 under CC BY 4.0 (Zanaga et al., 2022). Lunar elevation is from the NASA Goddard PGDA products, Martian elevation from the USGS CTX models (Laura et al., 2021), and radar from the NASA PDS Mini-RF archive and JPL UAVSAR. Every number in the paper is generated by scripts from recorded artifacts.

References

Airbus Defence and Space. Copernicus DEM product handbook. Technical Report Version 5.0, Campaign ID GEO.2018-1988-2, Airbus Defence and Space GmbH, 2022. URL https://dataspace.copernicus.eu/sites/default/files/media/files/2024-06/geo1988-copernicusdem-spe-002_producthandbook_i5.0.pdf. 29 November 2022. The Copernicus DEM data DOI requested by ESA is 10.5270/ESA-c5d3d65, which resolves to the Copernicus Data Space COP-DEM collection page.

B. Aussel, O. Rüsch, B. Gundlach, V. T. Bickel, S. Kruk, and E. Sefton-Nash. Global Lunar Boulder Map From LRO NAC Optical Images Using Deep Learning: Implications for Regolith and Protolith. *Journal of Geophysical Research: Planets*, 130(7):e2025JE008981, 2025. doi: 10.1029/2025je008981.

Adrien Bardes, Jean Ponce, and Yann LeCun. VICReg: Variance-Invariance-Covariance Regularization for Self-Supervised Learning, 2021.

Michael K. Barker, Erwan Mazarico, Gregory A. Neumann, David E. Smith, Maria T. Zuber, and James W. Head. Improved LOLA elevation maps for south pole landing sites: Error estimates and their impact on illumination conditions. *Planetary and Space Science*, 203:105119, 2021. doi: 10.1016/j.pss.2020.105119.

Ross A. Beyer, Oleg Alexandrov, and Scott McMichael. The Ames Stereo Pipeline: NASA's Open Source Software for Deriving and Processing Terrain Data. *Earth and Space Science*, 5(9):537–548, 2018. doi: 10.1029/2018ea000409.

Bruce A. Campbell. High circular polarization ratios in radar scattering from geologic targets. *Journal of Geophysical Research: Planets*, 117(E6), 2012. doi: 10.1029/2012JE004061.

Bruce A. Campbell and Michael K. Shepard. Lava flow surface roughness and depolarized radar scattering. *Journal of Geophysical Research: Planets*, 101(E8):18941–18951, 1996. doi: 10.1029/95JE01804.

Mathilde Caron, Hugo Touvron, Ishan Misra, Hervé Jégou, Julien Mairal, Piotr Bojanowski, and Armand Joulin. Emerging Properties in Self-Supervised Vision Transformers, 2021.

Lynn M. Carter, Catherine D. Neish, D. B. J. Bussey, Paul D. Spudis, G. Wesley Patterson, Joshua T. Cahill, and R. Keith Raney. Initial observations of lunar impact melts and ejecta flows with the Mini-RF radar. *Journal of Geophysical Research: Planets*, 117(E12), 2012. doi: 10.1029/2011JE003911.

Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A Simple Framework for Contrastive Learning of Visual Representations, 2020.

Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale, 2020.

B. Efron. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1):1–26, 1979. doi: 10.1214/aos/1176344552.

Wenzhe Fa and Yuzhen Cai. Circular polarization ratio characteristics of impact craters from Mini-RF observations and implications for ice detection at the polar regions of the Moon. *Journal of Geophysical Research: Planets*, 118(8):1582–1608, 2013. doi: 10.1002/jgre.20110.

Jichao Fang, Lei Zhang, Michael Phillips, and Wei Luo. CraterBench-R: Instance-Level Crater Retrieval for Planetary Scale, 2026.

Caleb I. Fassett, Ali M. Bramson, Joshua T. S. Cahill, Cameron P. Harris, Gareth A. Morgan, Catherine D. Neish, Cole A. Nypaver, G. Wesley Patterson, Edgard Rivera-Valentín, Patrick A. Taylor, Bradley J. Thomson, and the Mini-RF Team. Improved orthorectification and empirical reduction of topographic effects in monostatic Mini-RF S-band observations of the Moon. *The Planetary Science Journal*, 5(1):4, 2024. doi: 10.3847/PSJ/ad0a61.

Paolo Fraccaro, Gabby Nyirjesy, Daniela Szwarcman, Himanshu Patil, Vishal Gaur, Rohit Lal, Rachel A. Slank, Geoffrey Dawson, Hiyam Debary, Michael K. Barker, Andrew Annex, Vishnu Viswanathan, Zachary Morse, Ethan I. Schaefer, Nikolaos Dionelis, Ankur Kumar, Campbell D. Watson, Manil Maskey, Rebekah I. Dawson-Rigas, Juan Bernabé-Moreno, Rahul Ramachandran, and Sujit Roy. Multimodal-Multiresolution Foundation Model for Lunar Remote Sensing, 2026.

Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, and Victor Lempitsky. Domain-Adversarial Training of Neural Networks, 2015.

Marc Girona-Mata, Jakob Gawlikowski, Sumit Goski, Gautier Bardi de Fourtou, Valentin T. Bickel, Ben Moseley, Abigail Calzada-Diaz, Sylvester Kaczmarek, and Raúl Ramos-Pollán. LunarFM: A Shared Multimodal Representation of the Moon's Surface, 2026.

Priya Goyal, Piotr Dollár, Ross Girshick, Pieter Noordhuis, Lukasz Wesolowski, Aapo Kyrola, Andrew Tulloch, Yangqing Jia, and Kaiming He. Accurate, Large Minibatch SGD: Training ImageNet in 1 Hour, 2017.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep Residual Learning for Image Recognition, 2015.

Sergey Ioffe and Christian Szegedy. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, 2015.

Jonathan Kavitzky, Jonathan Zarecki, Idan Brusilovsky, and Uriel Singer. Topo2vec: Topography Embedding Using the Fractal Effect, 2021.

Mikhail A. Kreslavsky and James W. Head, III. Kilometer-scale roughness of Mars: Results from MOLA data analysis. *Journal of Geophysical Research: Planets*, 105(E11):26695–26711, 2000. doi: 10.1029/2000JE001259.

Mikhail A. Kreslavsky, James W. Head, Gregory A. Neumann, Margaret A. Rosenburg, Oded Aharonson, David E. Smith, and Maria T. Zuber. Lunar topographic roughness maps from Lunar Orbiter Laser Altimeter (LOLA) data: Scale dependence and correlation with geologic features and units. *Icarus*, 226(1):52–66, 2013. doi: 10.1016/j.icarus.2013.04.027.

Hans R. Kunsch. The Jackknife and the Bootstrap for General Stationary Observations. *The Annals of Statistics*, 17(3):1217–1241, 1989. doi: 10.1214/aos/1176347265.

Jason R Laura, Lauren A Adoram-Kershner, David P Mayer, Benjamin H Wheeler, Kirsten H Bauck, Robin L Fergason, and United States Geological Survey Astrogeology Science Center. Mars Reconnaissance Orbiter (MRO) Context Camera (CTX) orthoimage generated using Ames stereo pipeline derived digital terrain models, 2021. Cited by the USGS STAC collection mro_ctx_controlled_usgs_dtms as its data citation.

Zelin Liu, Bocheng Li, Yuling Zhou, Xuanting Li, Yixuan Yang, Jing Wang, Weishu Zhao, and Xiaofeng Gao. GEAR: Geography-knowledge Enhanced Analog Recognition Framework in Extreme Environments, 2026.

Ilya Loshchilov and Frank Hutter. SGDR: Stochastic Gradient Descent with Warm Restarts, 2016.

Ilya Loshchilov and Frank Hutter. Decoupled Weight Decay Regularization, 2017.

Michael C. Malin, James F. Bell, Bruce A. Cantor, Michael A. Caplinger, Wendy M. Calvin, R. Todd Clancy, Kenneth S. Edgett, Lawrence Edwards, Robert M. Haberle, Philip B. James, Steven W. Lee, Michael A. Ravine, Peter C. Thomas, and Michael J. Wolff. Context Camera Investigation on board the Mars Reconnaissance Orbiter. *Journal of Geophysical Research: Planets*, 112(E5):2006JE002808, 2007. doi: 10.1029/2006je002808.

Leland McInnes, John Healy, and James Melville. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction, 2018.

Fanwei Meng, Xiaopeng Wang, André Antunes, Jie Zhao, Guoliang Zhou, Biqiong Wu, and Tianqi Hao. Deep learning selection of analogues for Mars landing sites in the Qaidam Basin, Qinghai-Tibet Plateau, 2024.

Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. Model Cards for Model Reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 2019. doi: 10.1145/3287560.3287596.

NASA. NASA identifies candidate regions for landing next americans on moon. NASA news release RELEASE22-089, 2022. URL https://www.nasa.gov/news-release/nasa-identifies-candidate-regions-for-landing-next-americans-on-moon/. Released 19 August 2022; page last updated 8 January 2024.

Stewart Nozette, Paul Spudis, Ben Bussey, Robert Jensen, Keith Raney, Helene Winters, Christopher L. Lichtenberg, William Marinelli, Jason Crusan, Michele Gates, and Mark Robinson. The Lunar Reconnaissance Orbiter miniature radio frequency (Mini-RF) technology demonstration. *Space Science Reviews*, 150(1–4):285–302, 2010. doi: 10.1007/s11214-009-9607-5.

Himanshu Patil, Gabby Nyirjesy, Rachel A. Slank, Vishal Gaur, Daniela Szwarcman, Paolo Fraccaro, Nikolaos Dionelis, Michael K. Barker, Andrew Annex, Vishnu Viswanathan, Zachary Morse, Ethan I. Schaefer, Hiyam Debary, Ankur Kumar, Rohit Lal, Geoffrey Dawson, Campbell Watson, Rebekah I. Dawson-Rigas, Manil Maskey, Juan Bernabé-Moreno, Rahul Ramachandran, and Sujit Roy. SomBench: Benchmark Dataset for Advancing Machine Learning in Lunar Science, 2026.

R. Keith Raney, Joshua T. S. Cahill, G. Wesley Patterson, and D. Benjamin J. Bussey. The m-chi decomposition of hybrid dual-polarimetric radar data with application to lunar craters. *Journal of Geophysical Research: Planets*, 117(E12), 2012. doi: 10.1029/2011JE003986.

Edgard G. Rivera-Valentín, Caleb I. Fassett, Anne K. Virkki, Santa L. Pérez-Cortés, Tyler Powell, Dylan C. Hickson, Michael C. Nolan, and G. Wesley Patterson. Empirical evaluation of S-band radar polarimetric response to surface rock abundance across the lunar highlands and maria. *The Planetary Science Journal*, 7(6):144, 2026. doi: 10.3847/PSJ/ae6fa4.

Stuart J. Robbins. A New Global Database of Lunar Impact Craters >1–2 km: 1. Crater Locations and Sizes, Comparisons With Published Databases, and Global Analysis. *Journal of Geophysical Research: Planets*, 124(4):871–892, 2019. doi: 10.1029/2018je005592.

Stuart J. Robbins and Brian M. Hynek. A new global database of Mars impact craters ≥1 km: 1. Database creation, properties, and parameters. *Journal of Geophysical Research: Planets*, 117 (E5):2011JE003966, 2012. doi: 10.1029/2011je003966.

C. Rodriguez Sanchez-Vahamonde and C. Neish. The Surface Texture of Martian Lava Flows as Inferred from Their Decimeter- and Meter-scale Roughness. *The Planetary Science Journal*, 2 (1):15, 2021. doi: 10.3847/psj/abbfac.

Paul A. Rosen, Scott Hensley, Kevin Wheeler, Greg Sadowy, Tim Miller, Scott Shaffer, Ron Muellerschoen, Cathleen Jones, Howard Zebker, and Søren Madsen. UAVSAR: A new NASA airborne SAR system for science and technology research. In *2006 IEEE Conference on Radar*, pages 22–29. IEEE, 2006. doi: 10.1109/RADAR.2006.1631770.

M. A. Rosenburg, O. Aharonson, J. W. Head, M. A. Kreslavsky, E. Mazarico, G. A. Neumann, D. E. Smith, M. H. Torrence, and M. T. Zuber. Global surface slopes and roughness of the Moon from the Lunar Orbiter Laser Altimeter. *Journal of Geophysical Research: Planets*, 116(E2): E02001, 2011. doi: 10.1029/2010JE003716.

Michael K. Shepard, Bruce A. Campbell, Mark H. Bulmer, Tom G. Farr, Lisa R. Gaddis, and Jeffrey J. Plaut. The roughness of natural terrain: A planetary and remote sensing perspective. *Journal of Geophysical Research: Planets*, 106(E12):32777–32795, 2001. doi: 10.1029/2000JE001429.

Shashwat Shukla, Gerald Wesley Patterson, Abhisek Maiti, Shashi Kumar, and Nicholas Dutton. Inferring the Variability of Dielectric Constant on the Moon from Mini-RF S-Band Observations. *Remote Sensing*, 16(17):3208, 2024. doi: 10.3390/rs16173208.

David E. Smith, Maria T. Zuber, Herbert V. Frey, James B. Garvin, James W. Head, Duane O. Muhleman, Gordon H. Pettengill, Roger J. Phillips, Sean C. Solomon, H. Jay Zwally, W. Bruce Banerdt, Thomas C. Duxbury, Matthew P. Golombek, Frank G. Lemoine, Gregory A. Neumann, David D. Rowlands, Oded Aharonson, Peter G. Ford, Anton B. Ivanov, Catherine L. Johnson, Patrick J. McGovern, James B. Abshire, Robert S. Afzal, and Xiaoli Sun. Mars Orbiter Laser Altimeter: Experiment summary after the first year of global mapping of Mars. *Journal of Geophysical Research: Planets*, 106(E10):23689–23722, 2001. doi: 10.1029/2000je001364.

David E. Smith, Maria T. Zuber, Glenn B. Jackson, John F. Cavanaugh, Gregory A. Neumann, Haris Riris, Xiaoli Sun, et al. The lunar orbiter laser altimeter investigation on the lunar reconnaissance orbiter mission. *Space Science Reviews*, 150(1–4):209–241, 2010. doi: 10.1007/s11214-009-9512-y.

John P. Snyder. Map projections: A working manual. Technical Report Professional Paper 1395, U.S. Geological Survey, 1987.

G. D. Tolometti, C. D. Neish, C. W. Hamilton, G. R. Osinski, A. Kukko, and J. R. C. Voigt. Differentiating Fissure-Fed Lava Flow Types and Facies Using RADAR and LiDAR: An Example From the 2014–2015 Holuhraun Lava Flow-Field. *Journal of Geophysical Research: Solid Earth*, 127(7):e2021JB023419, 2022. doi: 10.1029/2021jb023419.

Daniele Zanaga, Ruben Van De Kerchove, Dirk Daems, Wanda De Keersmaecker, Carsten Brockmann, Grit Kirches, Jan Wevers, Oliver Cartus, Maurizio Santoro, Steffen Fritz, Myroslava Lesiv, Martin Herold, Nandin-Erdene Tsendbazar, Panpan Xu, Fabrizio Ramoino, and Olivier Arino. ESA WorldCover 10 m 2021 v200, 2022. Dataset, version v200, published 28 October 2022, licence CC BY 4.0.

Jure Zbontar, Li Jing, Ishan Misra, Yann LeCun, and Stéphane Deny. Barlow Twins: Self-Supervised Learning via Redundancy Reduction, 2021.

A Data analysis

This appendix reports the distributions behind the fingerprint. Every value is generated from the named analysis records, the windows evidence of the data inventory, the comparison record of the lunar target and the configuration, whose input digests are recorded; analyses whose records do not exist yet are named as such. All values here are Derived.

Site and region distributions. Table 8 summarises the fingerprints of all 25 windows of the lunar target site and all 42 Holuhraun windows. The lunar windows have relief Hurst exponents from 0.84 to 1.01 with a median of 0.93, so the primary target window, at 0.89, is typical of its site. The Holuhraun windows span a much wider range, from 0.25 to 0.87, and their median relief ν at the longest lag, 11.37 m, is more than an order of magnitude below the target site median of 152.7 m. The Hurst exponents of the target site windows lie close to the highland median that Rosenburg et al. (2011) report from LOLA, although our lags, grid and estimator differ from theirs, so the agreement is a sanity check and not a reproduction.

Table 8: Derived fingerprint distributions per site, over the eligible windows of each site. ν values are medians over windows, in metres, at the shortest and longest configured lag.

Site	Variant	windows	H min	H median	H max	median ν short, long
Moon, target site	relief	25	0.84	0.93	1.01	42.8, 152.7
Moon, target site	texture	25	0.69	0.86	0.92	31.4, 99.7
Earth, Holuhraun	relief	42	0.25	0.51	0.87	4.59, 11.37
Earth, Holuhraun	texture	42	0.14	0.39	0.80	4.41, 8.62

Bootstrap intervals and target percentiles. Over the target site, the median relief Hurst exponent is 0.93 with an interquartile range from 0.89 to 0.96 and a moving block bootstrap interval for the median from 0.89 to 0.95 (2,000 replicates, blocks of two windows, a two dimensional extension of the block bootstrap of Kunsch (1989), itself an extension of the bootstrap of Efron (1979)). The texture exponent has a median of 0.86 with an interval from 0.81 to 0.89. Neighbouring windows overlap by half, so these intervals describe one site, not the Moon. Within the site, the primary target window sits at the 24th percentile of relief H, the 60th percentile of relief ν at the shortest lag and the 44th at the longest lag. Across the error clones the percentile of H stays between 24 and 24.

Scale behaviour. A deviogram of the whole site from 90 m to 4320 m lags has a single slope Hurst exponent of 0.92. A hinged fit with a break at 1440 m is preferred over a straight line by a difference in the Bayesian information criterion of 29.4, with exponents of 0.96 below and 0.77 above the break. Rosenburg et al. (2011) report that a breakover, where present, typically sits near a kilometre baseline on the Moon; our break lies in the same range, but the site is a single massif, our lags come from a gridded and partly interpolated model rather than profiles, and the comparison is a sanity check, not a reproduction. The fingerprint's lags all lie below this break.

Across bodies. Table 9 compares the three bodies in fingerprint space, using the windows that the search read: 25 lunar windows of the target site, 359 Martian windows from 2 CTX models, and 71 Earth windows from 2 regions. The overlap coefficient is the summed minimum of the two normalised histograms on common bins, so it is one for identical distributions and zero for disjoint ones. The median relief Hurst exponent is 0.93 at the lunar site, 0.54 on Mars and 0.65 on Earth. The lunar site barely overlaps either other body at any lag, while the Martian and Earth samples overlap substantially. The samples are small, contiguous and not independent, so the table describes these windows and not the bodies.

Table 9: Derived overlap coefficients between the bodies' fingerprint distributions over the windows read by the search. Short and long are relief $\ln \nu$ at the shortest and longest configured lag.

Pair	relief H	texture H	relief $\ln \nu$ short	relief $\ln \nu$ long
Earth and Moon	0.13	0.17	0.07	0.00
Mars and Moon	0.10	0.13	0.03	0.02
Earth and Mars	0.64	0.63	0.51	0.67

Distances to the target. Over the 71 compared Earth windows, relief distances to the target run from 0.68 to 4.95 (median 2.05) and texture distances from 0.66 to 4.92 (median 2.21). Because the

two regions were chosen for radar availability and as a terrain only control, these distances describe the evaluated windows and not the Earth as a whole.

Analyses not yet available. Roughness by land cover class and climate zone, the effect of a surface model mask, the declared regions against a global background sample, the comparison with published Martian roughness, and the overlap of the bodies in embedding space with and without gradient reversal have no records, so they are not reported.

B Embedding visualisations

All values in this appendix are Predicted, from the experimental encoder, and none enters a ranking.

Nearest Earth neighbours of the lunar target. The evaluation record lists the encoder's five nearest Earth windows to the lunar target among the encoder's own test windows. Their cosine distances run from 0.36 to 0.45, and the nearest lies at latitude 34.55 and longitude −118.24 degrees, in the Copernicus tile that also holds the Rosamond region. The deterministic ranking was not an input of that evaluation, so the report gives no side by side comparison with the fingerprint shortlist.

Failure cases. The evaluation record lists the 20 largest disagreements between the encoder's ranking of Earth windows for the lunar target and the ranking by the root mean square difference of the ten $\ln \nu$ values. Their rank differences run from 183 to 204 places, so windows that are close to the target in the fingerprint can be far from it in the embedding and the reverse, which is the same weak agreement that the physics agreement test of Section 8.1 measures. A fraction of 0.043 of the pooled test windows lies beyond the ninety ninth percentile of the Mahalanobis distance of the training embeddings, and the record keeps 50 of them as out of distribution examples.

Not shown. The UMAP projections (McInnes et al., 2018) of test embeddings coloured by body, with and without gradient reversal, and hillshade galleries of the neighbours need an embedding export that was not produced, so they are not drawn.