



SIGSPATIAL 2026
Riverside, California

ImplicitTerrainV2:

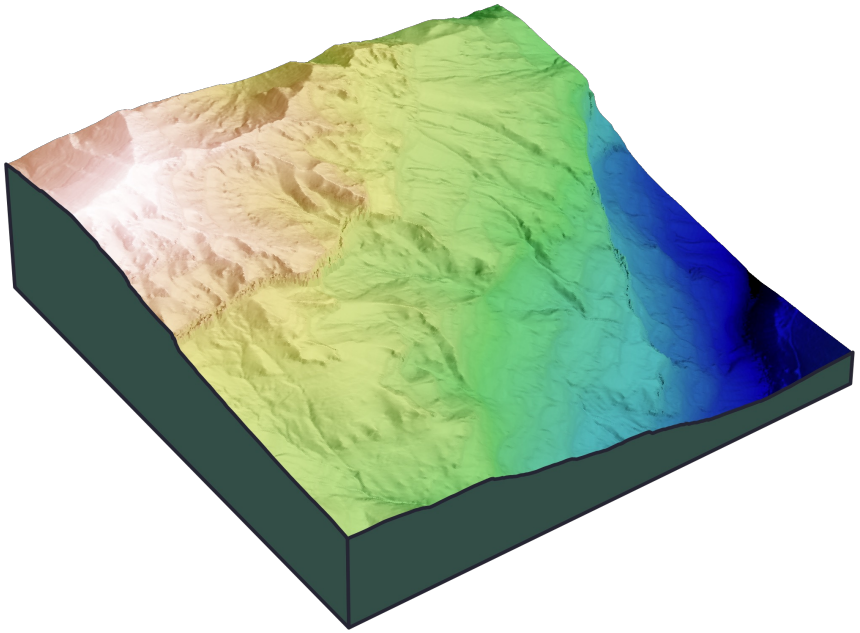
Wavelet-Guided Spatially Adaptive Neural Terrain Representation

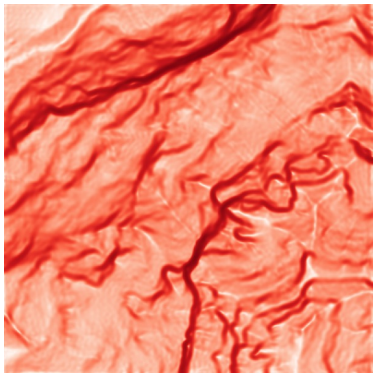
Haoan Feng, Xin Xu, Leila De Floriani

University of Maryland, College Park

Terrain data reconstruction and analysis

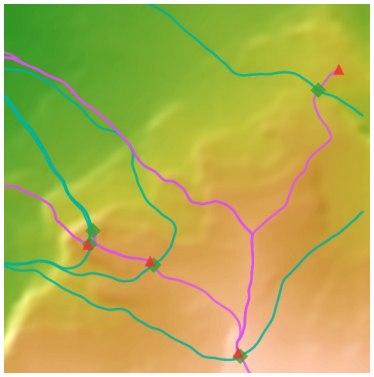
Working on Digital Elevation Model (DEM), we study local surface geometry, topological features, and their interconnectivity.





Slope / Aspect

Local surface geometry



Topological features

Critical feature points & lines

Peak/saddle/basin
Ridge/valley lines

Terrain as a coordinate function

Instead of regular grid or irregular mesh representations, we consider fitting a continuous **neural function** to reconstruct terrain surface geometry from DEM elevation samples.

2D coordinates

(x, y)

→ **Neural terrain function** θ →

Predicted elevation

$f_\theta(x, y)$

DEM samples



Raster data (or sampled points) are used as training and evaluation data.

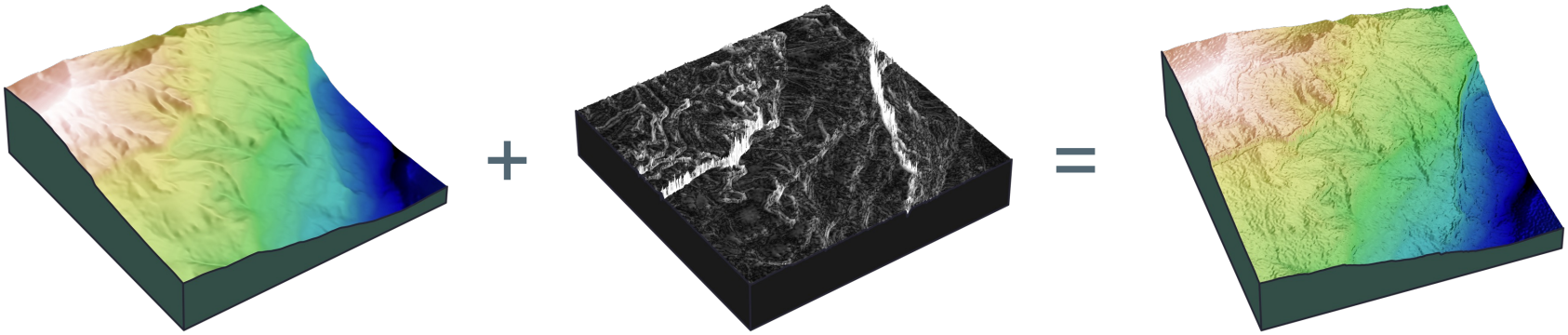
SIREN uses sine activations in a fully connected network:

$$\boldsymbol{h}_{\ell+1} = \sin(\boldsymbol{W}_\ell \boldsymbol{h}_\ell + \boldsymbol{b}_\ell)$$

$$\frac{d^n}{du^n}\sin(u) = \sin\left(u + \frac{n\pi}{2}\right)$$

ImplicitTerrain separates shape and residual detail

Two SIREN networks predict components of the elevation at each location



Shape

Fit the smoothed DEM for shape analysis

Residual detail

Fit (DEM - predicted shape)

Reconstruction

Limitations of the original terrain INR

The cascaded models support both tasks, but each branch needs a different fitting constraint

Shape branch fits smoothed elevations

But, elevation-only fitting does not directly supervise slopes and aspects.

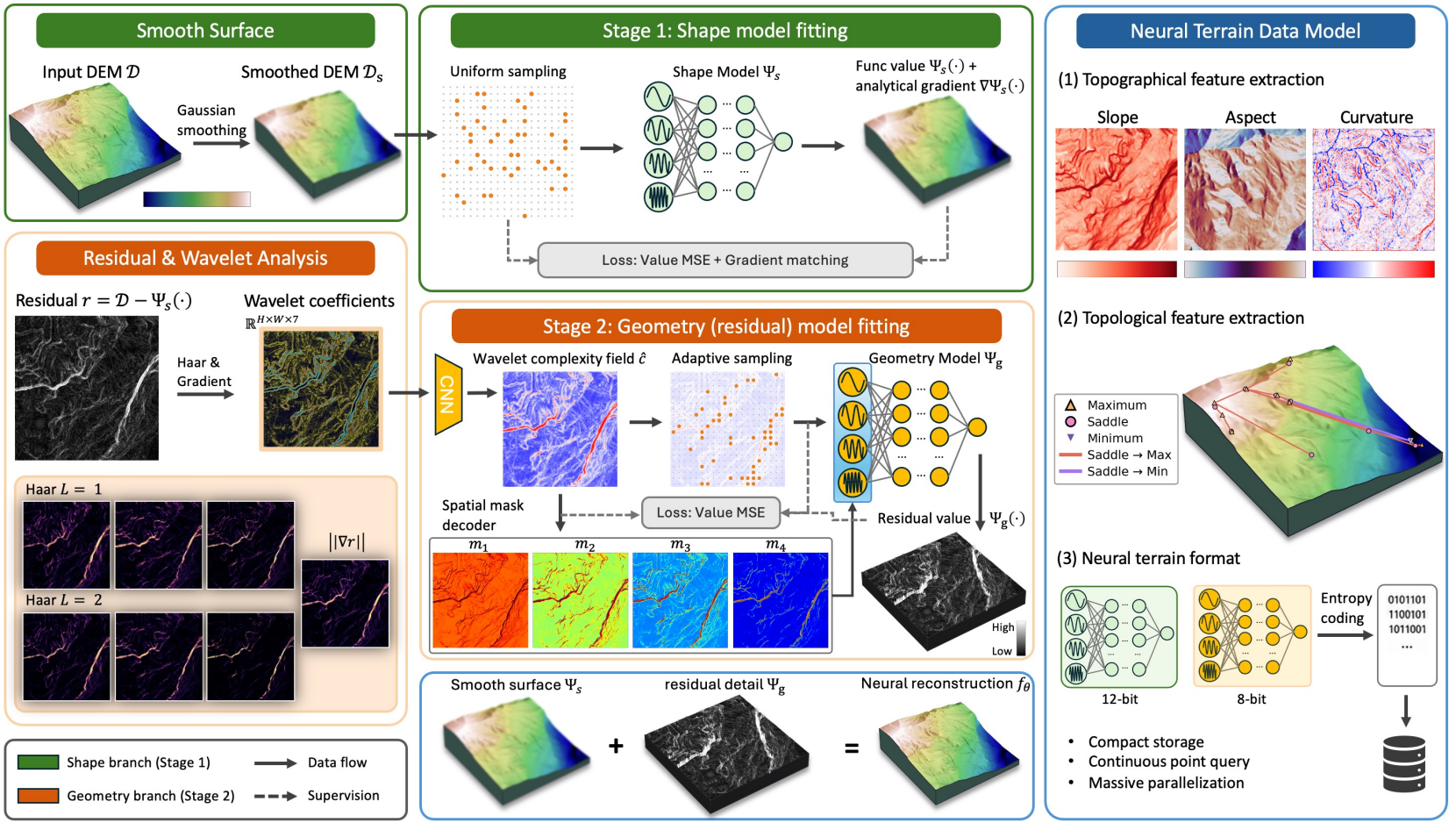
Residual branch fits fine geometric detail

But, fine-detail features might also affect smooth regions, we named it as *frequency leakage.*

Moreover, original ImplicitTerrain: 397K parameters ~ 1.51 MiB, ~15 mins training time.

Can terrain-specific fitting improve accuracy with a smaller model?

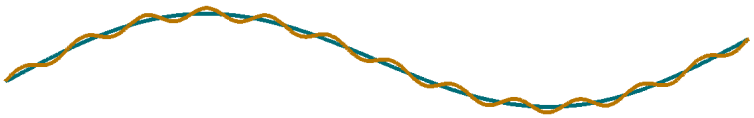
ImplicitTerrainV2: improvements to both branches



Height accuracy does not ensure slope accuracy

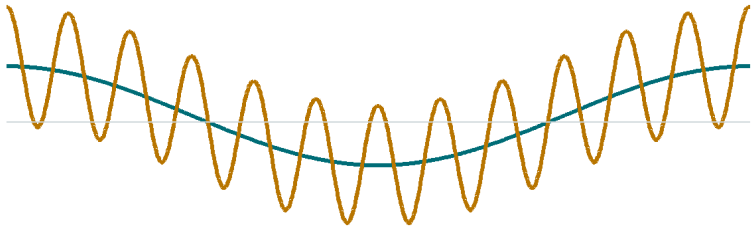
Gradient accuracy is more important for shape analysis than elevation itself.

Elevation profile



Smooth reference

Derivative of the profile



Reference + small oscillation

Small changes in elevation (high PSNR) might lead to large changes in slope (gradient absolute error)

Shape fitting with height and gradient matching

Reference gradients add extra slope constraints at each sampled location

Smoothed DEM D_s → **Finite differences** → Reference gradients

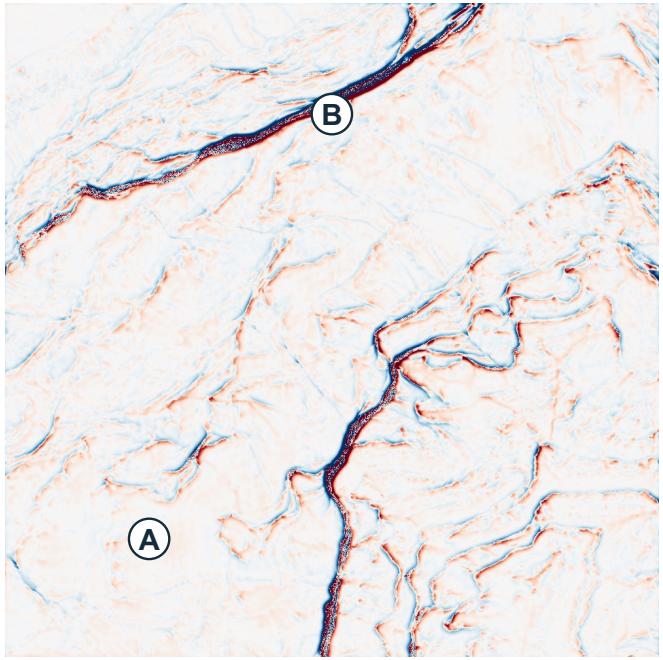
Shape model Ψ_s → **Automatic differentiation** → Predicted gradients

$$\mathcal{L}_{\text{grad}} = \frac{1}{M}\sum_{j=1}^{M}\left\|\nabla\Psi_s(\boldsymbol{x}_j) - \nabla D_s(\boldsymbol{x}_j)\right\|_2^2$$

$$\mathcal{L}_{\text{shape}} = \mathcal{L}_{\text{reconst}} + \lambda\mathcal{L}_{\text{grad}}, \quad \lambda = 0.1$$

Residual detail varies across the terrain

Different locations need different contributions from fine-scale features, while SIREN reconstructs the whole domain (x, y) with shared neural network parameters

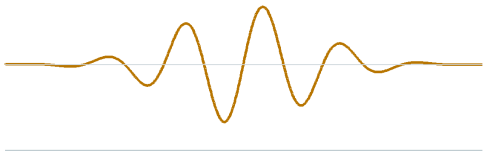




A Little residual variation



B Strong local variation



SASNet: frequency control in the geometry network

A fixed sinusoidal first layer, *Frequency Embedding Layer,* makes the neural function's frequency support predictable.

Integer frequency combinations in a hidden neuron

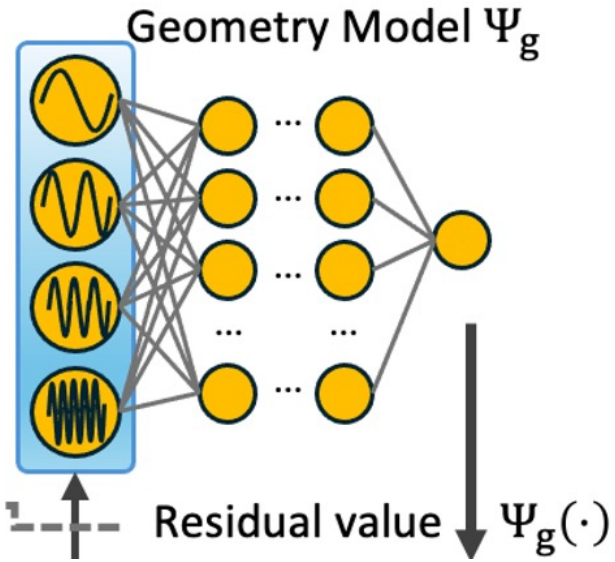
$$h_i(\boldsymbol{x}) = \sin\left(\sum_{j=1}^{n} a_{ij}\sin(\boldsymbol{\omega}_j^\top \boldsymbol{x} + \phi_j) + b_i\right) = \sum_{\boldsymbol{k}\in\mathbb{Z}^n} \alpha_{i,\boldsymbol{k}}\sin(\boldsymbol{\beta}_{\boldsymbol{k}}^\top \boldsymbol{x} + \eta_{i,\boldsymbol{k}})$$

$$\alpha_{i,\boldsymbol{k}} = \prod_{j=1}^{n} J_{k_j}(a_{ij}), \qquad \boldsymbol{\beta}_{\boldsymbol{k}} = \sum_{j=1}^{n} k_j\,\boldsymbol{\omega}_j, \qquad \eta_{i,\boldsymbol{k}} = b_i + \sum_{j=1}^{n} k_j\,\phi_j$$

J is the Bessel function of the first kind.
Input frequencies determine the combinations.

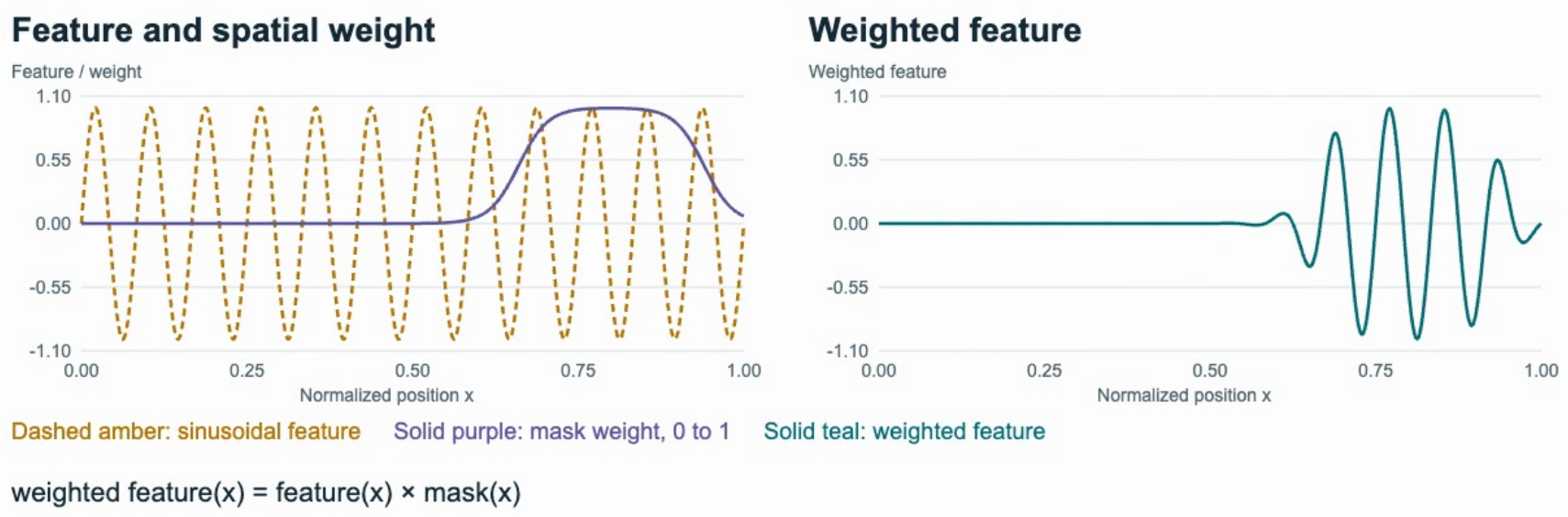
Frequency embedding initialization

$$\gamma_j(\boldsymbol{x}) = \sin(\boldsymbol{\omega}_j^\top \boldsymbol{x} + \phi_j), \qquad \boldsymbol{\omega}_j = \frac{2\pi}{T}\boldsymbol{q}_j, \quad \boldsymbol{q}_j \in \mathbb{Z}^2$$



SASNet: spatial modulation of frequency groups

A spatial mask assigns a continuous weight to a frequency neuron.

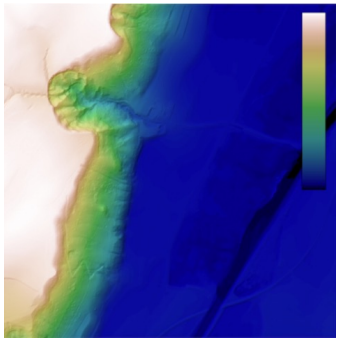


How should terrain residuals inform mask generation?

Terrain residuals

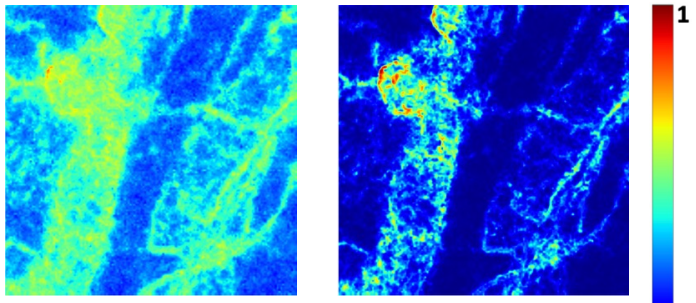
The residual offers explicit local structure for the mask generator

Input DEM



SASNet learns masks from error

Coordinates → Learned hash-grid + MLP → **Spatial masks**

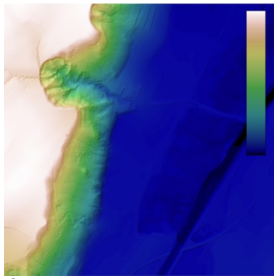


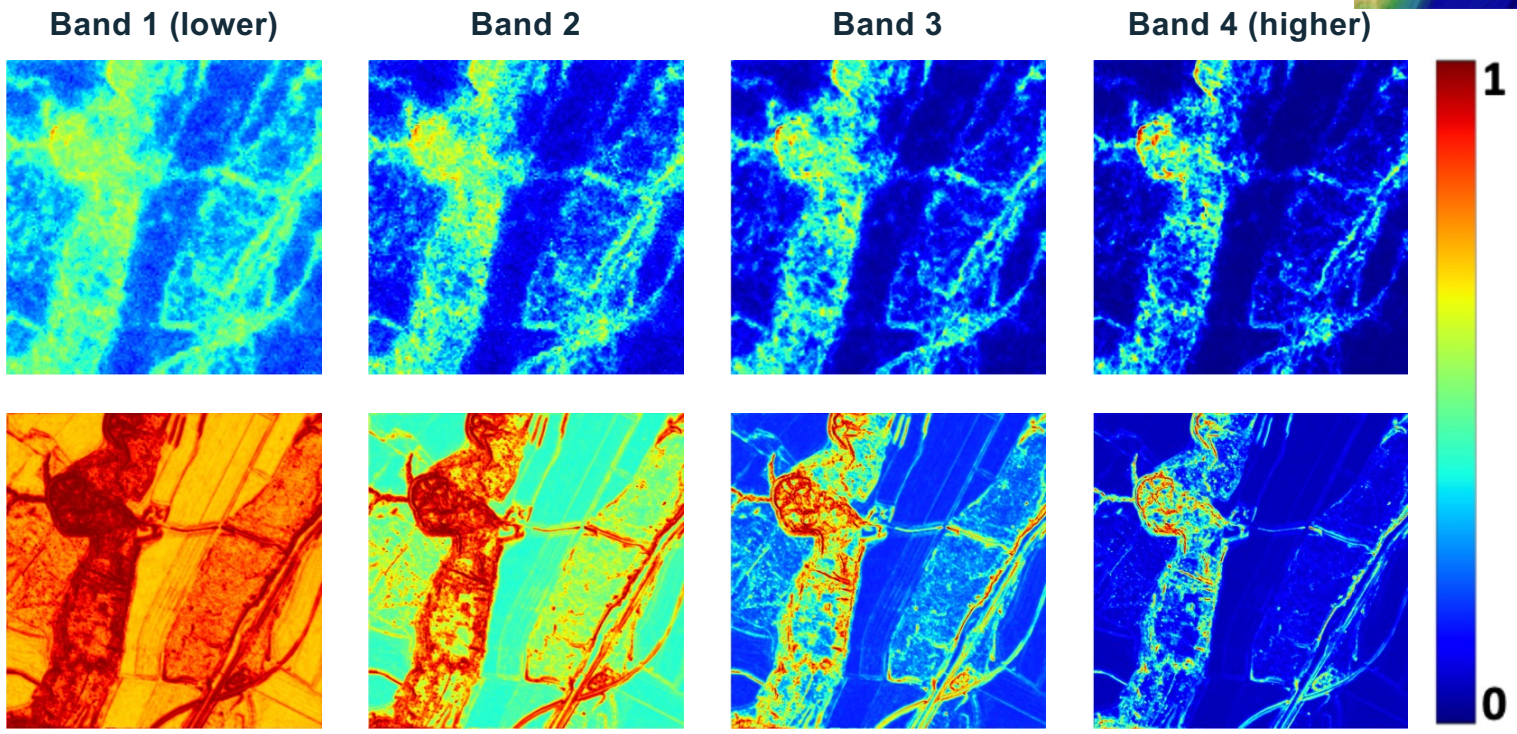
Lower freq-group | Higher freq-group

Noisy, unstable, and depends on training dynamics!

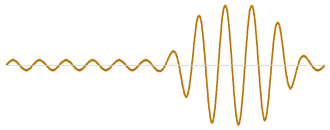
Ordered masks constrain spatial modulation

Input DEM



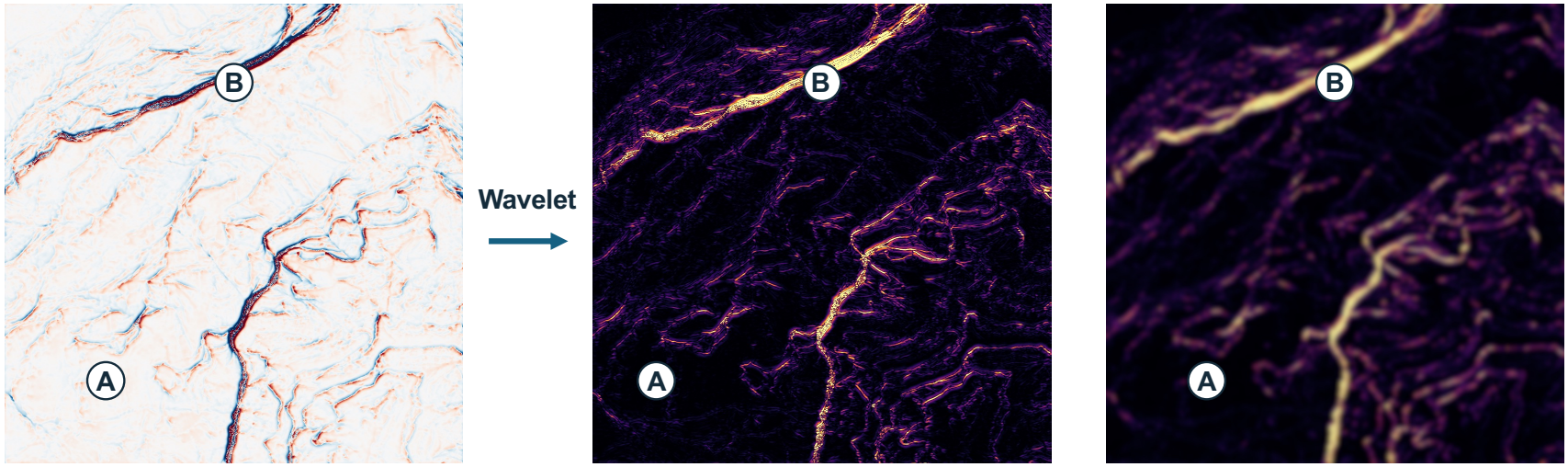


Wavelet features expose local residual structure



Wavelet filters' multi-scale responses preserve the location and scale of terrain variation

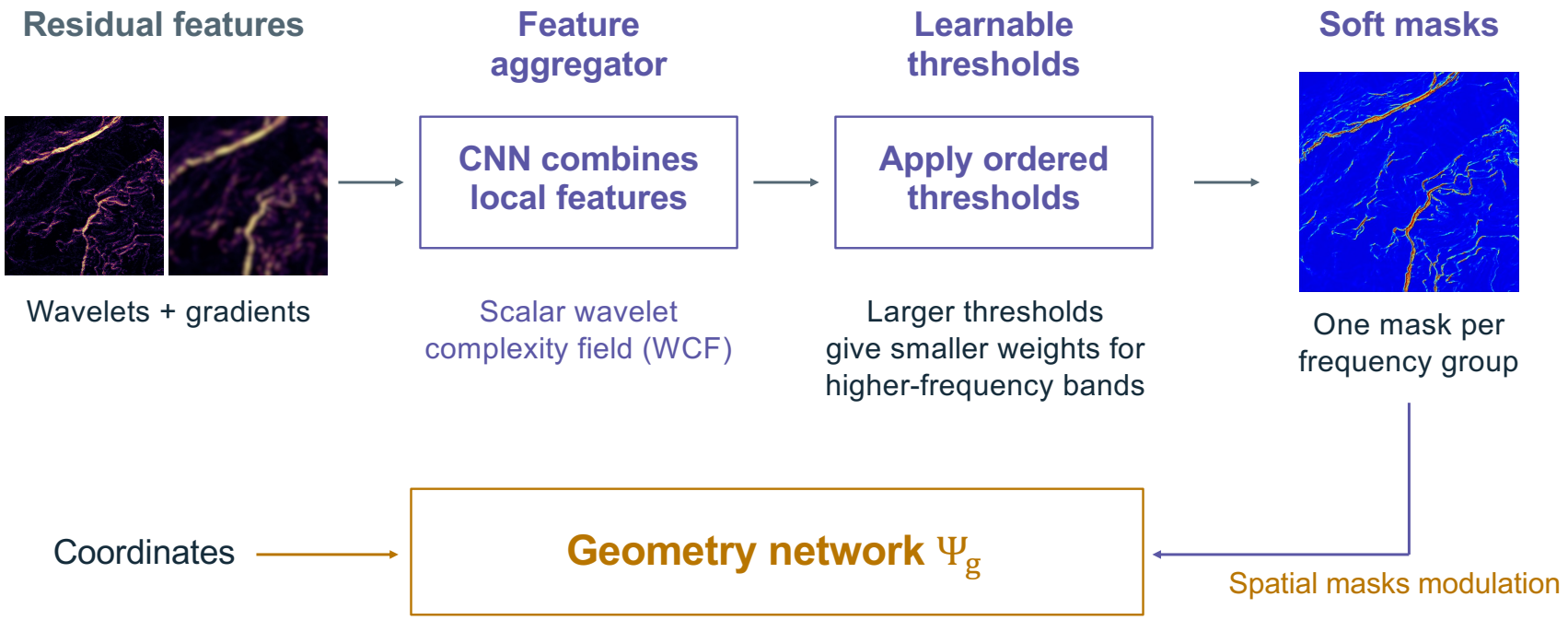
Residual | **Finer-scale response** | **Coarser-scale response**



Brighter feature values indicate stronger response → stronger local changes

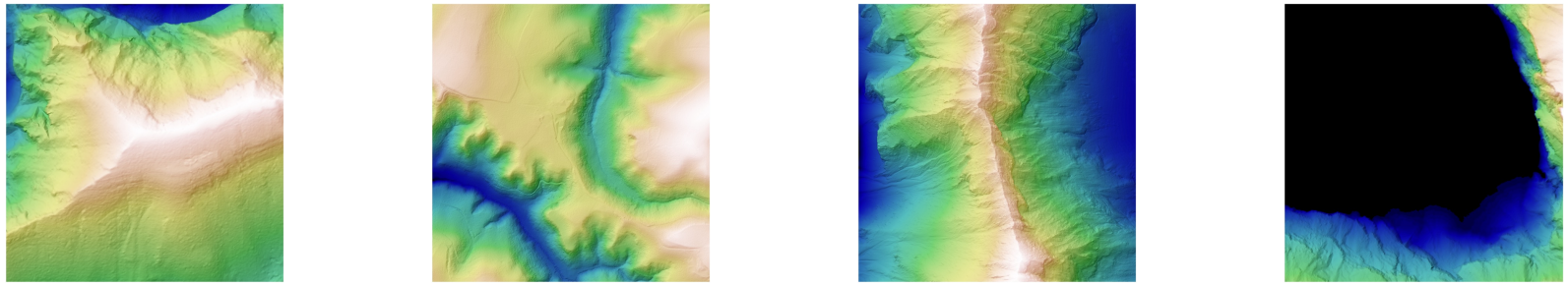
WCF: residual features to learned spatial masks

Computed residual features guide a jointly learned mask generator



Experimental Results

50 independent terrain fits from swissALTI3D are selected by DEM statistics



Each tile 1000 × 1000 samples at 1 m spacing and we fit a separate model to each tile.

Elevation reconstruction

WCF: local errors + component ablation
Full model: elevation error + parameter count

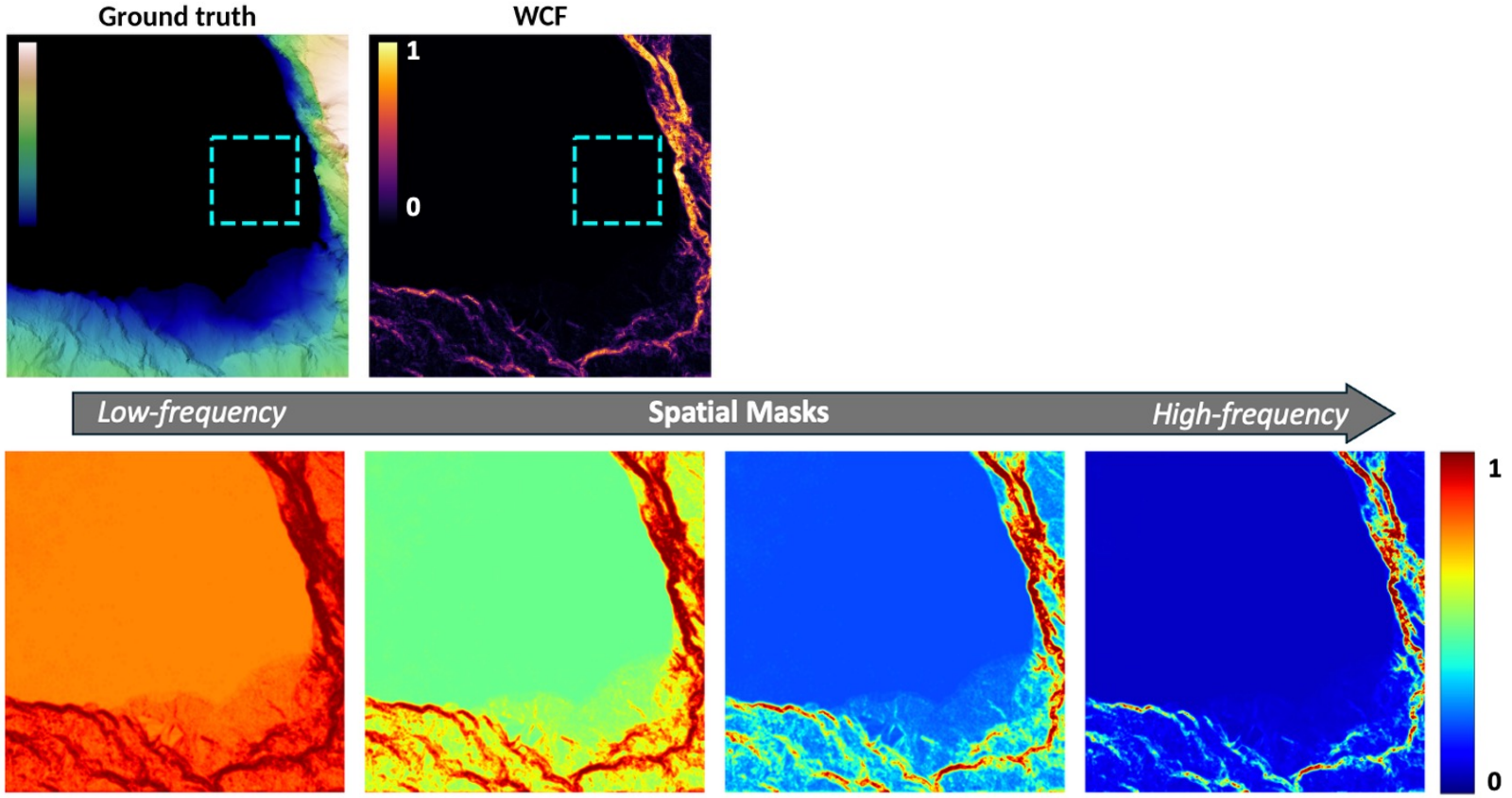
Shape gradient accuracy

Gradient supervision: gradient MAE
Analysis: qualitative terrain analysis applications

Reconstruction quality and model size

Method	Params	E2E PSNR (dB) ↑	MAE (m) ↓	MaxAE (m) ↓	Peak Mem (GB) ↓	Geo-Iter (ms) ↓
ImplicitTerrain [23]	397K	60.55 ± 5.67	0.197	29.73	9.8	81.9
ImplicitTerrain-128	100K	49.74 ± 5.34	1.436	32.46	4.9	34.5
SASNet [21]	139K	64.48 ± 4.27	**0.125**	17.56	5.4	41.3
ImplicitTerrainV2	124K	**66.25 ± 4.09**	0.132	**11.85**	**1.7**	**20.2**

Spatial modulation alleviate high-frequency leakage



Ablations

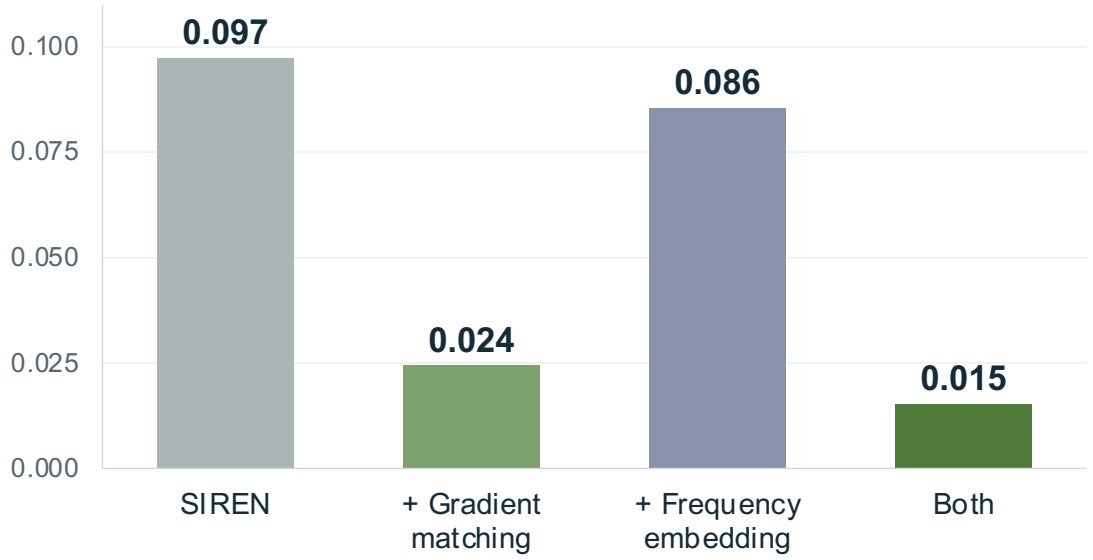
Progressive ablation study on 50 terrain tiles

Configuration	PSNR (dB)	Δ
Calibrated baseline (ImplicitTerrain-256)	62.13	-
+ Frequency embedding and gradient matching	63.08	+0.95
+ WCF	**65.72**	**+3.59**
+ Complexity-guided sampling*	66.25	+4.12

* *Complexity-guided sampling* selects training coordinates with spatially varying probabilities, highlights the “difficult” fitting regions.

Gradient matching improves the analysis surface

Gradient MAE against the smoothed shape target, under the normalized training convention



Gradient matching

4.0× lower error

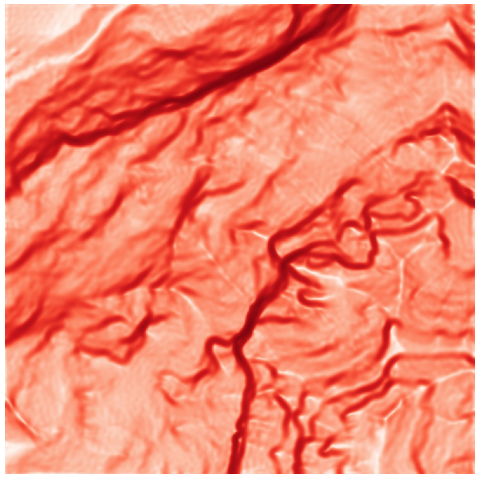
Combined with frequency embedding

6.4× lower error

Local terrain analysis on the fitted shape

Evaluate local terrain attributes from one differentiable shape model via analytical gradient computation

Slope

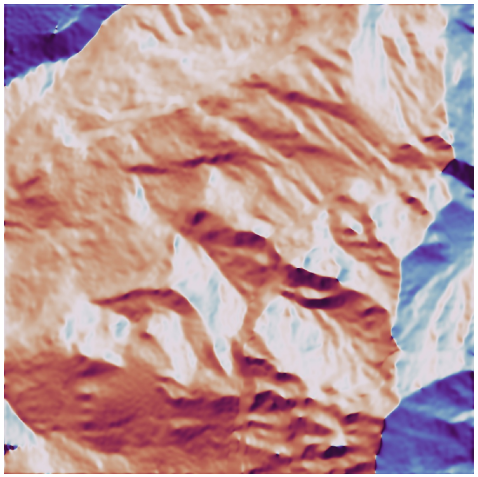


Pale to dark red: 0° to 90°

Local steepness

First derivatives

Aspect

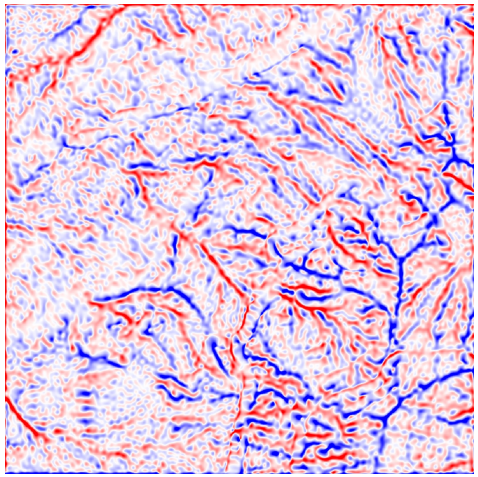




0° 360°

Surface orientation

Mean curvature





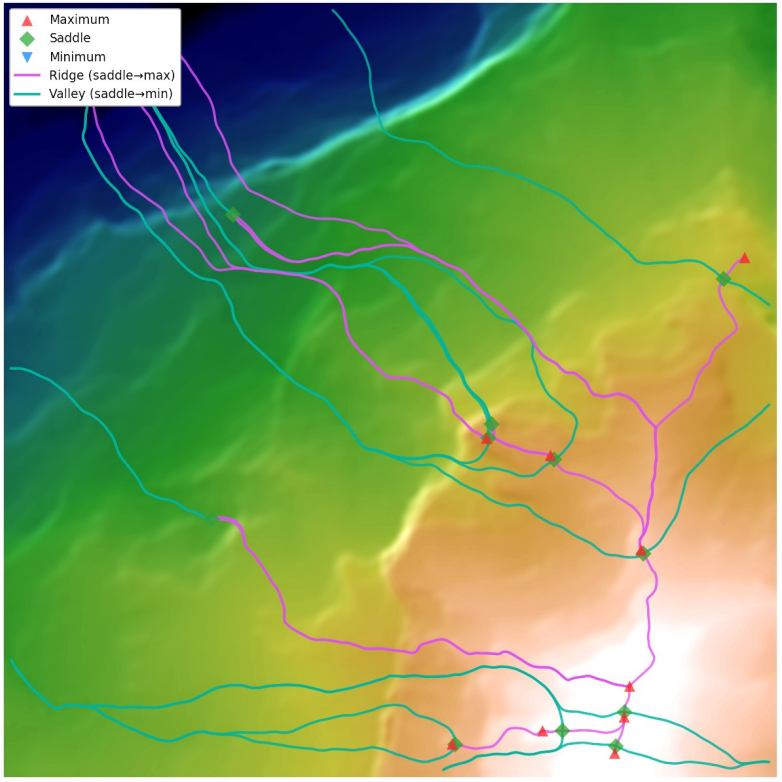
Negative Positive

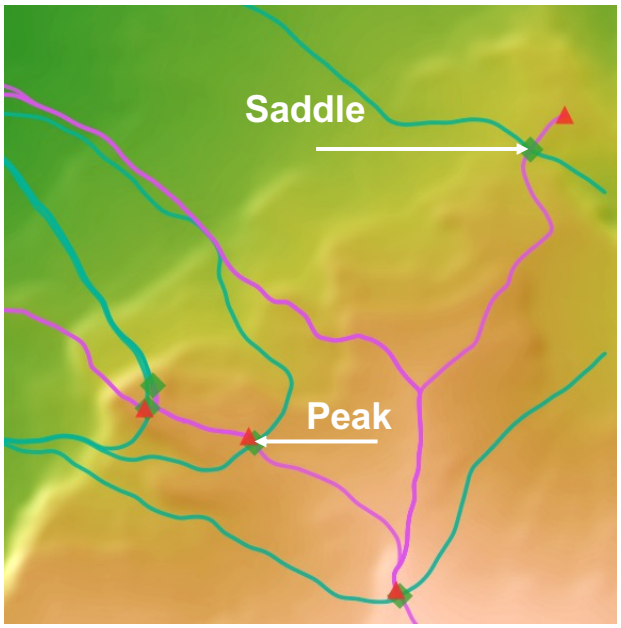
Local surface bending

First + second derivatives

Ridge–valley connectivity on the fitted shape

Critical points and gradient trajectories describe terrain connectivity





$\nabla \Psi_s = 0$

Critical points

Hessian

Maxima
Minima
Saddles

Purple: ascending ridge paths

Green: descending valley paths

Conclusions

Shape model

Frequency control and slope supervision improve the analysis surface. Gradient matching alone reduces gradient MAE by 4.0×.

Geometry model

WCF replaces the SASNet hash-grid mask generator with wavelet features and learned nested masks, giving more stable spatial modulation.

Smooth terrain fits quickly. A small residual loss provides a weak signal for learning where high-frequency capacity is needed.

Training efficiency

Adaptive sampling reduces redundant coordinate evaluations. Higher reconstruction accuracy with 124K parameters vs. 397K in ImplicitTerrain.

Post-training optimization

Quantize the shape model to 12 bits and geometry/WCF to 8 bits. Reconstruction PSNR: 66.25 → 65.97 dB, a 0.28 dB decrease.

Paper and code