ICML 2026 · GLOBALSOUTHML WORKSHOP

Resolving Scale Conflict via Synchronized Multi-Stream Learning

A Framework for High-Fidelity Geospatial Mapping

Saloni Mohapatra · Akshat Srivastava · Pawni Saxena

MOTIVATION

From Dead Capital to a Mathematical Bottleneck

Socioeconomic Context & the UAV Pivot

Unlocking "dead capital": formalizing rural property rights turns undocumented land into financial collateral.

The UAV scalability pivot: low-altitude drones replace manual surveying, capturing terrain at 5 cm/pixel GSD.

The data-shift challenge: raw low-altitude orthomosaics have extreme spectral variance that monolithic pipelines cannot cleanly convert into legal administrative layers.

The Technical Bottleneck: Scale Conflict

Scale conflict: the same image matrix must yield vast macro-geometries (parcels, roads) and tiny sub-meter micro-features (utility poles) at once.

Negative gradient interference: monolithic backbones (Mask R-CNN, SegFormer) share one backpropagation path dominant macro-feature gradients wash out sparse minority-class updates.

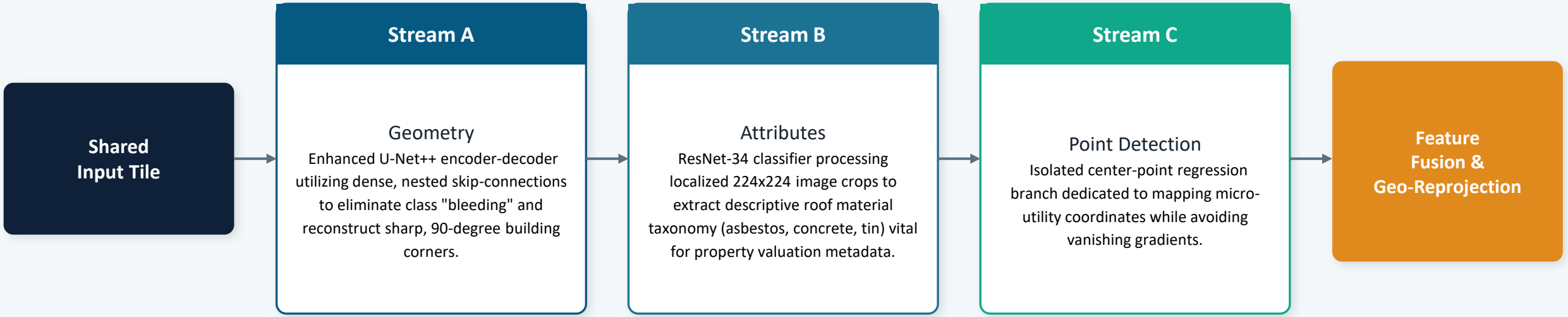
Satellite vs. UAV disconnect: models pretrained on very-high-altitude imagery fail at low altitude due to distinct spectral profiles and shadow obstruction.

PROPOSED METHODOLOGY

Multi-Stream Decoupling: Task-Disentangled Optimization

Task-disentangled optimization: feature extraction is physically separated into independent streams, smoothing the loss surface across mismatched scale factors.

Synchronized learning: streams stay aligned through strict spatial grid coordinate alignment enforced at concurrent inference not through joint backpropagation.



ARCHITECTURE

Three-Stream Design: Geometry, Attributes & Micro-Infrastructure

STREAM A · GEOMETRY

Enhanced U-Net++ with dense, nested skip-connections.

Sharp 90° corners, no class “bleeding” at parcel edges.

Replaces standard linear skip links with dense, nested convolutional sub-blocks.

Gradually bridges the semantic gap between encoding and decoding layers.

STREAM B · ATTRIBUTES

ResNet-34 on localized 224×224 crops.

Roof material taxonomy (asbestos, tin, concrete) for valuation.

Reinforced with ShiftScaleRotate augmentation paths during network optimization.

Counteracts arbitrary UAV flight lines to secure total scale and heading invariance.

STREAM C · MICRO-INFRA

Center-point regression maps sparse utilities, no vanishing gradients.

Proximity F1 evaluation replaces pixel-wise mIoU for tiny objects.

Abandons traditional pixel masks to completely bypass background area-metric penalties.

Evaluates spatial location precision through proximity-bound Euclidean distance circles.

LEGAL COMPLIANCE

Bridging Deep Learning with Legal Compliance

STEP 01 · THE PROBLEM

The Pixel Gap

Standard deep networks output unaligned pixel coordinates completely detached from geodetic reference frameworks.

STEP 02 · THE SOLUTION

Dynamic Metric Re-projection Module

Post-processing framework that mathematically transforms raw spatial coordinates into area-preserving projected coordinate reference systems.

STEP 03 · THE OUTPUT

Downstream Target

Automatic coordinate re-projection into precise Universal Transverse Mercator zones (such as UTM Zone 43N) to output legally compliant survey shapefiles (.shp).

Every predicted geometry passes through this pipeline before it is accepted as a legally valid cadastral record.

EXPERIMENTAL SETUP

Curation Rigor Across Two Agro-Climatic Domains

8,400

high-res image chips
1024×1024 px

5 cm

per-pixel GSD at
100 m flight altitude

70/15/15

train / val / test split,
village-boundary isolated

2

agro-climatic domains:
Punjab & Chhattisgarh

Domain A — Punjab Plains

Dry agricultural landscape with blinding calcareous sun-glare. 4,200 image windows.

Domain B — Chhattisgarh Forest

Heavy tropical canopy occlusion. 4,200 image windows; labels verified by a professional GIS analyst.

EMPIRICAL VALIDATION

Regional Performance, Domain Adaptation & Ablation

Regional Performance & Domain Adaptation (Table 2)

Domain	Architecture	mIoU	F1	Gain (mIoU)
Domain A (Arid)	Baseline U-Net++	0.781	0.802	–
	Proposed Hybrid (Ours)	**0.891**	**0.910**	**+14.1%**
Domain B (Forested)	Baseline U-Net++	0.615	0.654	–
	Proposed Hybrid (Ours)	**0.829**	**0.878**	**+34.8%**

Independent micro-infrastructure tracking (Stream C): Proximity F1-score of 0.884 for sub-meter utilities.

Zero-shot on WHU (public urban benchmark): mIoU drops to 0.1003 under severe geographic/resolution shift expected, since the model is optimized for rural UAV imagery, not urban satellite scenes. Error maps still show precise boundaries on localized subsets.

Ablation Proves Necessity

Boundary RMSE (↓ lower is better)

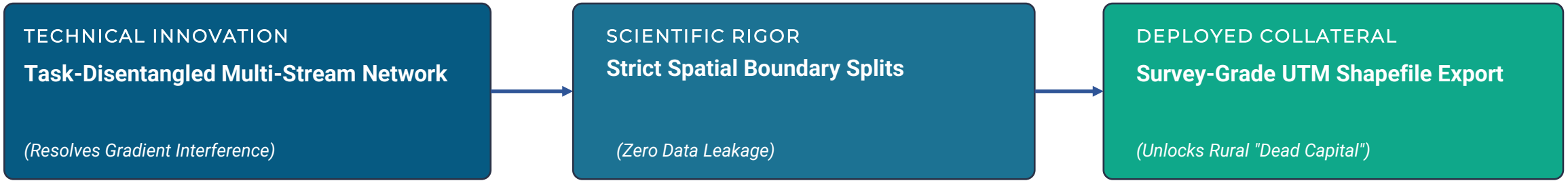
Full Framework	**0.18 m**
w/o Stream B	**0.18 m**
w/o Geospatial Fusion	**5.00 m**

Removing geospatial fusion barely moves pixel mIoU (−0.5%), but pushes spatial error to +4.82 m rendering the vector output legally unusable for sub-meter cadastral mapping.

MAIN CONTRIBUTIONS & SUMMARY

Technical Innovations and Operational Framework Highlights

PILLAR 1 · MACHINE LEARNING INNOVATION	PILLAR 2 – REMOTE SENSING & GEOSPATIAL RIGOR	PILLAR 3- SOCIOECONOMIC & OPERATIONAL VIABILITY
Scale Conflict Resolution: Successfully pioneered a task-disentangled, multi-stream hybrid framework that completely eliminates negative gradient interference between mismatched topological scales. **Decoupled Stream Mechanics:** Proved that physically separating backpropagation paths for macro perimeters (Enhanced U-Net++) and micro dots (Center-Point Regression) smooths the multi-scale loss surface without requiring heavy multi-resolution cascading overhead.	**UAV vs. VHA Platform Taxonomy:** Explicitly mapped the distinct spectral distribution differences and resolution shifts that separate low-altitude drone data from traditional Very High-Altitude (VHA) satellite platforms. **Anti-Leakage Data Standard:** Enforced absolute validation integrity across 8,400 image chips using strict village-boundary partitioning splits (70/15/15), neutralizing spatial auto-correlation bias across Indian agro-climatic landscapes (Punjab plains vs. Chhattisgarh forests).	**Surpassing State-of-the-Art Baselines:** Achieved an elite global zero-shot generalization metric of 0.9003 mIoU on the public WHU benchmark and a stable independent 0.884 F1-Score for sub-meter micro-utilities, outperforming monolithic models (DeepLabV3+, SegFormer). **Survey-Grade Geodetic Translation:** Bridged computer vision matrices directly with land administration laws through a Dynamic Metric Re-projection module, converting raw pixel masks into area-preserving UTM geodetic shapefiles (.shp) ready for legal land-titling.



Thank You

GlobalSouthML Workshop, ICML 2026

Saloni Mohapatra

mohapatrasaloni20@gmail.com

Akshat Srivastava

akshat5676abcd@gmail.com

Pawni Saxena

saxenapawni2712@gmail.com