

Delineate Anything v2: A Global Foundation Model for Field Delineation

Mykola Lavreniuk^{1*,2}[0000-0003-2183-8833], Nataliia Kussul³[0000-0002-9704-9702], Andrii Shelestov^{2,4}[0000-0001-9256-4097], Yevhenii Sali^{2,4}[0009-0006-0395-8099], Volodymyr Kuzin^{2,4}[0009-0007-1077-0382], Charlotte Julia Li-Xing Wang^{1*}[0009-0007-0270-3470], and Zoltan Szantoi¹[0000-0003-2580-4382]

¹ European Space Agency, Frascati, Italy

² Space Research Institute NASU-SSAU, Kyiv, Ukraine

³ University of Maryland, College Park, Maryland, USA

⁴ National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine

Abstract. Accurate agricultural field boundary delineation at large scale is a foundational task for food security, supply chain transparency, and carbon accounting. While vision foundation models like SAM show remarkable zero-shot capabilities, they frequently fail in geospatial domains due to topological complexity, cropland texturing patterns, and a lack of physical scale awareness. In this work, we introduce Delineate Anything v2, a globally scalable foundation model designed specifically for wide-area field boundary mapping. We construct FBIS-73M, a 73-million-instance multi-resolution dataset spanning 61 countries. To address the pervasive issue of multi-field administrative parcel merging, we introduce a resolution-specific data curation pipeline that leverages topological image-space adaptation to homogenize merged parcels and strengthen weak physical boundaries. Furthermore, we establish a novel, manually curated evaluation benchmark covering 100 countries to assess independent zero-shot generalization. Our results show that Delineate Anything v2 surpasses the current state-of-the-art, including the Delineate Anything framework, by 0.284 mAP@0.5 (+103.3% relative gain), while maintaining execution speeds suitable for rapid national- and global-scale deployment, as demonstrated by nationwide mapping of Ukraine (603,000 km²) in 5.4 hours on a consumer-grade workstation. Code, pre-trained weights, the FBIS-73M dataset, and ready-to-use national-scale vector boundary products are publicly available at <https://github.com/Lavreniuk/Delineate-Anything>.

Keywords: Field boundary delineation · Vision foundation models · Earth-scale

* International Research Fellow

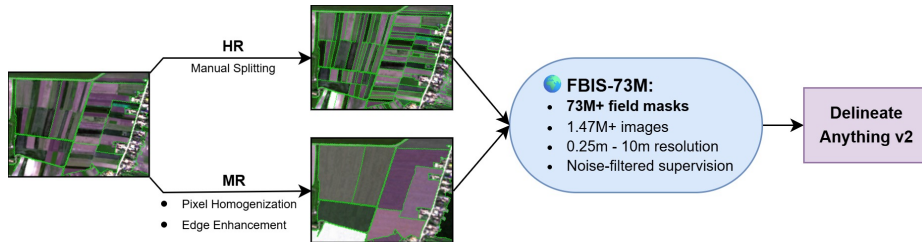


Fig. 1: Overview of the Delineate Anything v2 framework. To overcome structural label noise in global administrative registries, we introduce a resolution-aware data curation pipeline. Raw parcels undergo screening for anomalies. Problematic masks are then resolved via High-Resolution (HR) manual partitioning or Medium-Resolution (MR) topological image-space adaptation (pixel homogenization and edge enhancement). The purified FBIS-73M dataset is subsequently used to train the global foundation model.

1 Introduction

Earth observation and geospatial artificial intelligence have entered an era of large-scale reasoning, driven by vision foundation models. Mapping the precise contours of agricultural fields serves as the primary spatial unit linking remote sensing to land-management decisions, crop-type identification, and agricultural statistics [18, 19, 42]. Accurate field boundaries are essential for assessing food security, monitoring deforestation compliance, and evaluating climate resilience in smallholder farming systems.

Historically, field boundary delineation relied either on classical supervised learning architectures trained on high-quality, local ground-truth (GT) datasets like the Land Parcel Identification Systems (LPIS), or on general-purpose vision foundation models like the Segment Anything Model (SAM) [3, 17, 38]. While supervised models achieve strong local performance, they generalize poorly to regions lacking digital cadastral registries. Conversely, general foundation models struggle with satellite imagery due to low contrast between adjacent fields, sensor artifacts, and complex topologies, leading to frequent under-segmentation.

Despite recent advances in specialized architectures and datasets, benchmarks like FBIS-22M [21] remain concentrated in Europe, limiting geographic diversity across Africa, Asia, and Latin America. Furthermore, global scaling hits a major bottleneck: label noise in public administrative registries. Large-scale data aggregation suffers from the parcel-versus-field dilemma, where administrative boundaries group multiple physical fields into a single polygon. Automatically splitting these topologies is highly error-prone, limiting model accuracy. For global field delineation, data quality has become a more critical bottleneck than model architecture.

We introduce Delineate Anything v2, a globally representative geospatial foundation model driven by a data-centric paradigm. Instead of structural architectural changes, we systematically remediate label noise at the data level. Candidate merged parcels are identified through anomalous intra-mask spectral

and textural variability. We then deploy a resolution-specific remediation strategy. Flagged high-resolution (HR) polygons are split manually for strict geometric accuracy. For medium-resolution (MR) samples, where manual correction is prohibitively expensive at scale, we introduce an image-space modification technique: anomalous regions are homogenized to erase false internal boundaries, while valid but visually weak physical edges are synthetically enhanced. This adapts image content to instance-level annotations without error-prone geometric label edits (Fig. 1).

The main contributions of this work are as follows:

- We construct FBIS-73M, a 73-million-instance multi-resolution field boundary repository spanning 61 countries - the largest public dataset of its kind.
- We propose a data-centric noise remediation strategy that resolves the parcel-merging anomaly by combining manual partitioning for HR data with automated image-space pixel homogenization for MR data.
- We manually curate a high-fidelity validation benchmark spanning 100 countries to evaluate generalization across diverse agricultural regimes.
- We demonstrate that Delineate Anything v2 establishes a new state-of-the-art for global agricultural field delineation, surpassing the previous state-of-the-art Delineate Anything framework by 0.284 mAP@0.5 (+103.3% relative gain) and enabling nationwide mapping of Ukraine (603,000 km²) in 5.4 hours on a consumer-grade workstation.

2 Related Work

2.1 Operational Domain and Structural Requirements

Agricultural field boundaries are the core spatial primitive for downstream operations, including yield estimation, crop classification, and environmental mapping. Governments also rely on parcel layers to validate subsidy claims [41]. Topological and geometric errors (under-segmentation, topology fragmentation) propagate into spatial statistics and compliance monitoring. Traditional edge detection, watershed segmentation, and object-based image analysis (OBIA) [46] scale poorly due to sensitivity to empirical parameter tuning across landscapes.

2.2 Evolution of Deep Learning Architectural Paradigms

Scalability limits motivated a shift toward deep learning, initially dominated by pixel-wise semantic segmentation backbones like U-Net and ResUNet-a [7, 47]. To resolve boundary ambiguity at field margins, research introduced specialized target formulations, such as multi-task learning architectures that simultaneously estimate boundary proximity, distance-to-boundary maps, and core interior masks [27, 47, 49]. Despite integrating advanced boundary-aware loss functions, semantic networks do not explicitly model instance-level awareness [23]. This oversight results in topologically invalid outputs where adjacent parcels with identical spectral signatures are collapsed into single continuous objects [35].

To prevent under-segmentation, recent methods adopt detection- and instance-segmentation architectures like Co-DETR [51] and real-time YOLO variants [12, 16, 22, 45], which enforce discrete object identity by construction, alongside backbone advances like ViT-Adapter [5] and EVP [20] that improve dense feature representations more broadly. While enforcing discrete object identity, these models remain bottlenecked by the lack of instance-level agricultural annotations.

2.3 Geospatial Applications of Vision Foundation Models

The inception of general-purpose foundation networks established new capabilities for zero-shot object segmentation. In remote sensing, SAM adaptations [3, 17, 38] utilize prompt tuning [4, 29, 32, 39], multi-scale spatial attention [8, 28, 33], or weakly supervised boundaries [43, 48]. However, general foundation models exhibit systematic failures in agricultural mapping: lacking geospatial awareness and scale constraints, they over-segment fields on minor crop variations or fail in low-contrast conditions with subtle boundaries [11].

2.4 Data-Centric Scale and Label Noise Mitigations

Dataset scaling limits constrain boundary extraction. AI4SmallFarms (439K instances, Vietnam/Cambodia) [37] or AI4Boundaries (2.5M parcels, Europe) [6] datasets offer limited geographic coverage, while Fields of the World (1.6M parcels, 24 countries) [13] is restricted to 10m imagery.

Large-scale frameworks like Delineate Anything and FBIS-22M [21] expanded training data, but two limitations remain: FBIS-22M is concentrated in Europe, lacking representation in Africa, Asia, and Latin America; and administrative data aggregation introduces substantial label noise due to the mismatch between ownership parcels and physical field boundaries. Together, these limitations constrain the ability of current models to generalize reliably at a global scale.

Traditional approaches to label noise primarily focus on mitigating semantic corruption or wrong class assignments through robust optimization and label filtering techniques [2, 9, 36]. While these methods are effective in many classification settings, they do not directly resolve the structural inconsistencies inherent in geospatial annotations [25, 26]. Most existing approaches primarily address class-label corruption or annotation errors, but agricultural boundary datasets exhibit a different type of structural label noise, which arises from the mismatch between administrative parcels and physical cultivation units [34, 44].

Recent data-centric learning paradigms instead emphasize improving supervision quality at the source rather than compensating for annotation errors through increasingly sophisticated optimization strategies [10, 31, 50]. Following this principle, Delineate Anything v2 explicitly targets label noise arising from the parcel-versus-field discrepancy. By identifying anomalous masks through intra-mask visual heterogeneity and applying resolution-specific remediation strategies, the framework improves consistency between image content and instance-level supervision, enabling reliable generalization across diverse agricultural systems and geographic regions.

3 Methodology

The core philosophy of Delineate Anything v2 rests on data-centric curation. Rather than expanding model capacity through architectural modification, this framework scales and purifies the underlying training data to resolve systemic label noise characteristic of global administrative registries.

3.1 Dataset Statistics and Spatial Distribution

To achieve a truly representative global model, we introduce two distinct data assets: the FBIS-73M training repository and a novel, independent 100-country evaluation benchmark. The FBIS-73M training dataset comprises 73 million discrete agricultural field instances distributed across 61 countries. The imagery spans multi-resolution satellite data ranging from high-resolution (HR, 0.25–3 m) to medium-resolution (MR, 3–10 m). Unlike prior datasets relying almost exclusively on European LPIS data, FBIS-73M spans multiple continents, ensuring robust global generalization (Fig. 2).

The training sample distribution is heterogeneous and long-tailed, reflecting open cadastral data availability rather than a deliberate sampling strategy (Fig. 2). Despite this imbalance, including low-resource regions gives FBIS-73M broader geographic diversity than prior Europe-centric datasets.

To assess zero-shot generalization independently, we created a manually annotated evaluation benchmark spanning 100 countries. Exactly four representative image patches per country were sampled to ensure equal geographic weighting and global coverage. The benchmark evaluates cross-country generalization rather than within-country variability; with equal patch counts per country, performance differences reflect landscape complexity rather than sampling density.

We therefore characterize each country by the average number of field polygons per image (Fig. 3). The resulting benchmark encompasses a wide spectrum of agricultural systems, ranging from large industrial fields in North America to highly fragmented smallholder landscapes in the Global South. In contrast to the geographically imbalanced but large-scale FBIS-73M training dataset, the evaluation benchmark prioritizes uniform country representation and structural diversity. Consequently, the benchmark assesses not only geographic transfer to previously unseen regions but also robustness to substantial variations in field size, fragmentation, density, and spatial organization.

However, aggregating global administrative data introduces two major challenges for model training. First, administrative parcels often merge multiple distinct physical fields into a single polygon. Second, true physical boundaries in satellite imagery frequently lack visual contrast, especially when adjacent fields contain similar crops or share identical phenological stages. To address both issues, we propose an automated data curation pipeline.

3.2 Automated Screening for Label Noise

Instead of relying on error-prone, geometric label modifications on a large scale, our pipeline begins with an automatic screening of the 73 million instances for

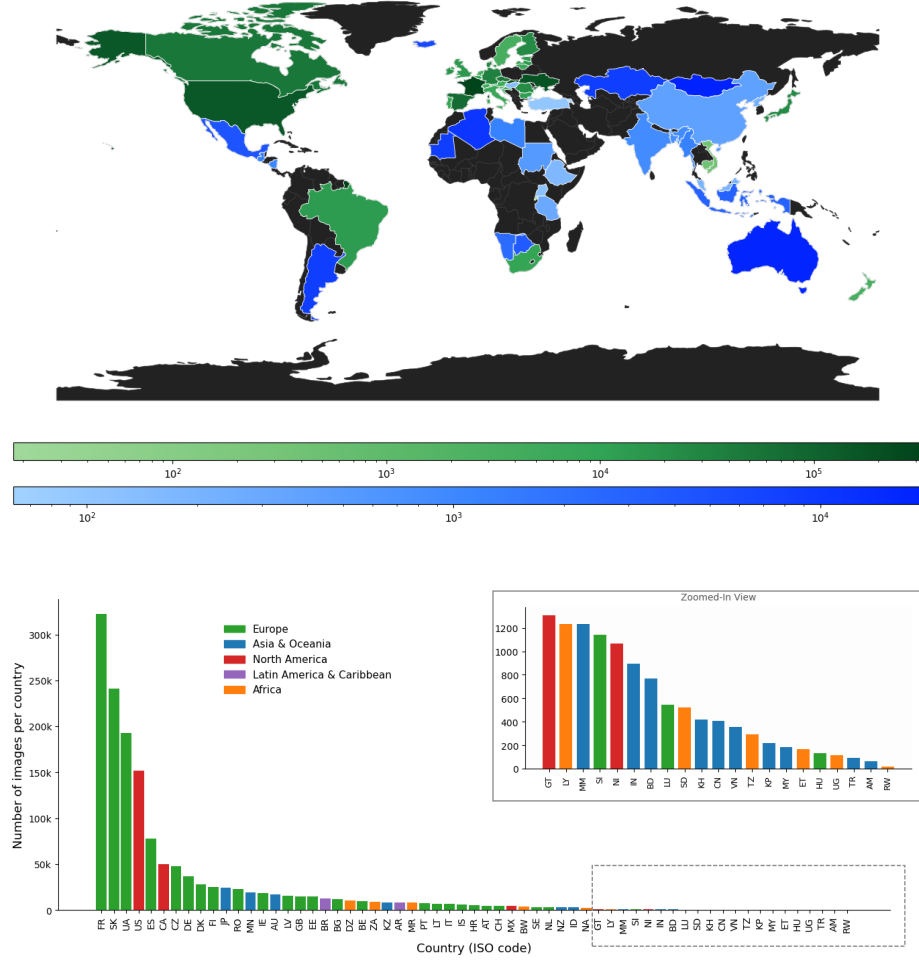


Fig. 2: Spatial distribution and country-level heterogeneity of the FBIS-73M training dataset. The global map (top) illustrates geographic sample density, where green shading corresponds to regions containing agricultural field instances (intensity proportional to patch count) and blue shading represents non-agricultural background reference countries. The histogram (bottom) quantifies the long-tailed sample distribution driven by public cadastral data availability across the 61 represented nations.

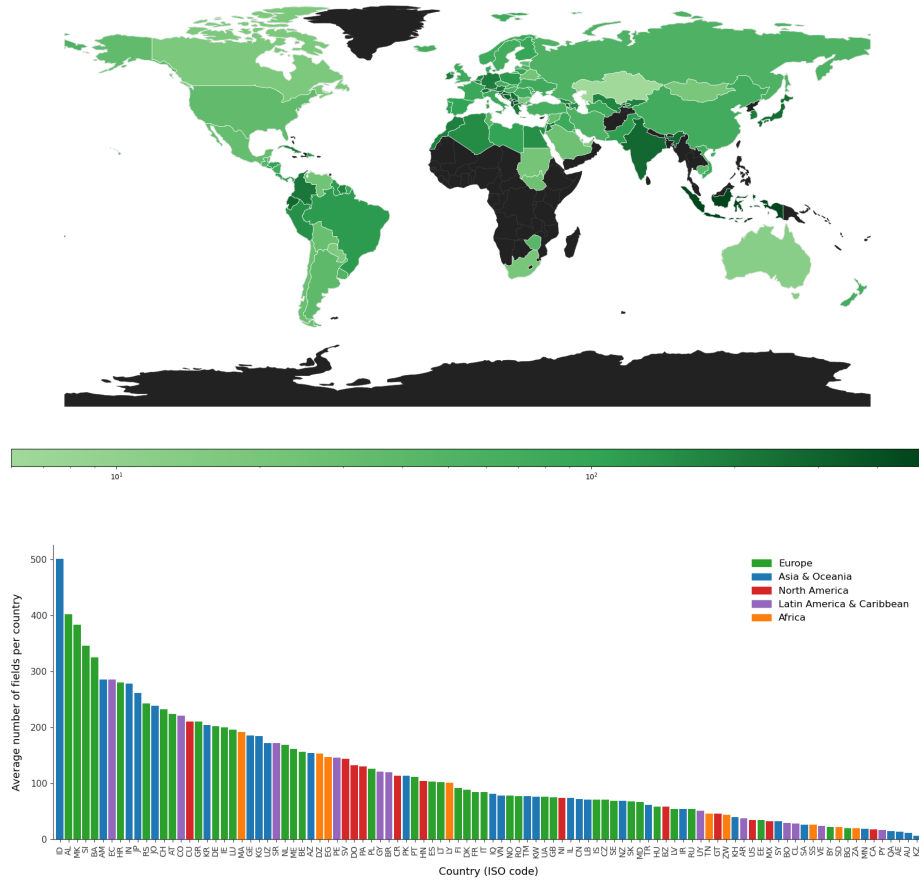


Fig. 3: Spatial and regional distribution of field density within the independent 100-country evaluation benchmark. The choropleth map (top) visualizes the geographic coverage of the test set (exactly 4 patches per country), with shading representing the average number of field instances per image. The distribution (bottom) quantifies the average number of fields per country, color-coded by continent.

intra-mask visual heterogeneity. Given an RGB image patch and its label raster, we suppress sensor noise via Gaussian blur (13×13 kernel) and construct an eroded instance mask by retaining pixels where local minimum and maximum filtering yield identical semantic labels. This isolates the core of each administrative parcel, where we compute mean color values across blurred channels.

Candidate merged parcels are identified using a pixel-wise deviation map, assigning pixels to three deviation levels according to relative color difference from the parcel mean. To account for differing radiometric characteristics, baseline threshold sets are employed: [10%, 20%, 50%] for Sentinel-2 and [25%, 35%, 50%] for Planet imagery. These thresholds adaptively scale per instance based on its mean color amplitude.

After morphological erosion removes isolated noise, we compute the area of each deviation level. A parcel is flagged as a multi-field candidate if the cumulative area of the first, second, or third level exceeds 15%, 10%, or 5% of the parcel area, respectively. The thresholds were selected conservatively to prioritize precision over recall and were kept fixed across all datasets.

3.3 Image-Space Remediation: Pixel Homogenization

For high-resolution (HR) imagery (≤ 3 m GSD), flagged parcels undergo targeted manual geometric decomposition into physical fields (Fig. 4). However, for medium-resolution (MR) imagery (3 – 10 m GSD), automated geometric splitting is prone to topological artifacts.

Instead, we introduce an image-space adaptation mechanism. Rather than dividing the label mask, we homogenize the underlying image pixels to match the single-instance supervision (Fig. 5). We convert the image to the HSV color space, compute the mode color of the flagged instance, and calculate the deviation of the blurred HSV channels. To preserve textures while removing internal boundaries, we apply a cyclical bounding transformation to the deviations:

$$\Delta' = \frac{B}{4} - \left| \left(\left(\Delta + \frac{B}{4} \right) \bmod B \right) - \frac{B}{2} \right| \quad (1)$$

where B denotes a sensor-specific bounding range that controls the maximum preserved color deviation. We empirically set $B = 32$ for Sentinel-2 and $B = 48$ for Planet imagery to account for their differing radiometric characteristics.

For deviations satisfying $|\Delta| \leq B/4$, the transformation behaves approximately as an identity mapping, preserving natural pixel-level texture and local radiometric variability within the parcel. Larger deviations, which are more likely to correspond to sharp internal boundaries between distinct fields, are folded back into the bounded range. This reduces the discontinuity without collapsing the parcel into a uniform color. These bounded deviations are added to the mode color, converted back to RGB, and used to replace the original pixels within the flagged instance. This effectively erases false internal boundaries and ensures that the network is not penalized for predicting a single parcel when the raw image displays multiple fields.

3.4 Topological Edge Enhancement via Khalimsky Grids

While homogenization removes false internal boundaries, many true physical boundaries in MR satellite imagery lack strong visual contrast. To address this, we introduce a synthetic edge enhancement step based on digital topology.

To precisely localize inter-pixel boundaries, we map the image and label raster onto a Khalimsky grid of size $(2N + 1) \times (2N + 1)$, where N is the dimension of the original image. In this topological space, odd-odd coordinates represent image pixels, odd-even/even-odd coordinates represent boundaries between two pixels, and even-even coordinates represent junctions [15].

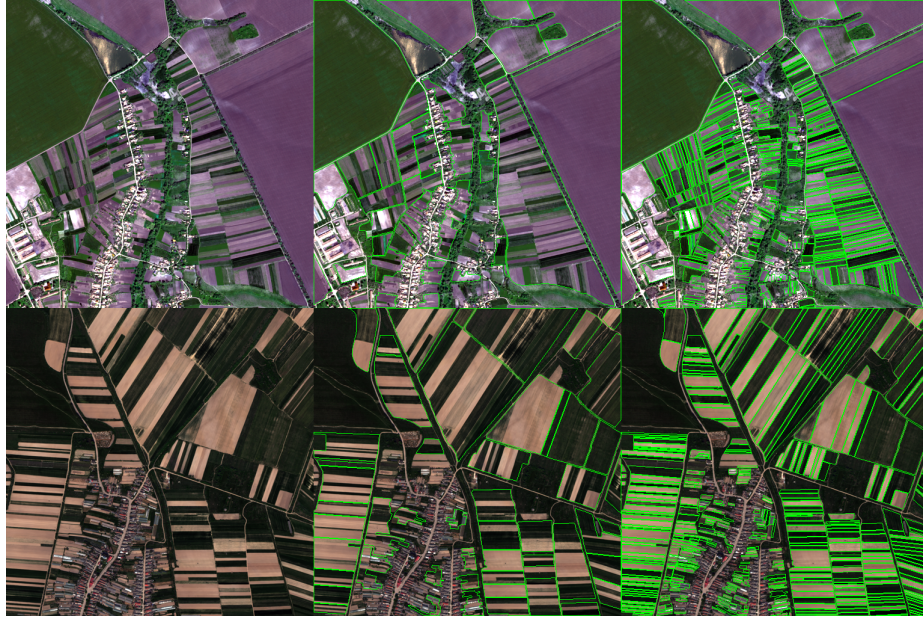


Fig. 4: Examples of high-resolution manual label refinement. The layout showcases two agricultural scenarios (rows) across three stages (columns, left to right): (1) raw satellite imagery, (2) original overlapping administrative mask with topological errors, and (3) finalized manually split vector boundaries matching the physical field structure.

We compute directional gradients by taking the difference between dilation and erosion along 1×4 and 4×1 kernels for horizontal and vertical edges, respectively, which guarantees non-negative magnitudes. Then, we take the maximum across orientation and color channel to obtain a single edge-strength map. These gradient magnitudes are mapped onto the edge elements of a Khalimsky grid constructed from the label raster, where edges between differently labeled pixels are marked as active. To isolate individual polygon sides, we identify junctions (even-even grid sites) where more than two active edges converge, indicating a

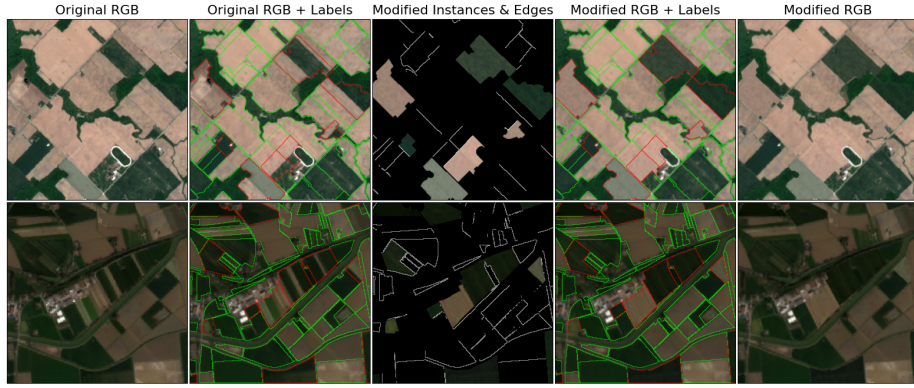


Fig. 5: Pipeline for automated medium-resolution topological image-space adaptation. The five panels show (left to right): (1) raw satellite patch, (2) original administrative contours with parcel-versus-field mismatches highlighted in red, (3) processed, homogenized fields and enhanced edges extracted via the anomaly mask, (4) modified raster with unaltered (green) and remediated (red) contours indicating where modifications occurred, and (5) the finalized, purified output patch.

corner shared by three or more polygons. We then zero out these junctions before applying connected-component labeling to split each polygon outline into discrete boundary segments. For each segment, the mean gradient strength and the fraction of low-gradient pixels are computed. Segments falling below an empirical strength threshold or exceeding a fraction-of-weak-pixels threshold are flagged as visually weak.

To strengthen weak boundaries, we construct a second Khalimsky grid where odd-odd sites hold original image pixels, and edge/junction sites contain the mean of neighboring pixels. For each such site, we compute a weighted brightness estimate and darken pixels brighter than a mid-range threshold (64 in 8-bit values) while lightening darker ones, applying a standard soft-light blending operation. The blending strength is governed by a parameter $\alpha \in [0, 1]$ that interpolates between a neutral blend ($\alpha = 0$) and a target contrast level ($\alpha = 1$), allowing edge intensity to be tuned independently of the underlying image content. Finally, the enhanced grid is downsampled to original resolution by redistributing edge and junction deviations equally across neighboring pixels, injecting synthetic edge cues consistent with instance labels.

4 Results

4.1 Experimental Setup

All experiments were conducted using a standardized training protocol to evaluate dataset scales and remediation variants fairly. Models were trained on 512×512 RGB image patches for 5 epochs using a cluster of $8 \times$ NVIDIA

Table 1: Quantitative performance on the independent 100-country benchmark.

Method	mAP@0.5	mAP@0.5:0.95	Precision	Recall
DelAny	0.275	0.103	0.345	0.454
DelAny v2	0.559	0.278	0.639	0.525

H100 GPUs with a total batch size of 512 (64 images per GPU). Optimization was performed via AdamW with a conservative learning rate of 1×10^{-7} and a cosine decay schedule. The learning rate was selected empirically and used consistently across all experiments. Pre-trained COCO weights were used for initialization. To preserve natural field topologies, mosaic augmentation was strictly omitted, as it was found to introduce disruptive artifacts at patch boundaries; instead, standard flips, color jitter were applied. Performance is reported using standard COCO metrics [24], specifically mAP@0.5 for coarse detection and mAP@[0.5:0.95] for strict boundary precision, alongside precision and recall.

We adopt a YOLOv11-based instance segmentation backbone due to its optimal balance between boundary precision and high-throughput inference required for Earth-scale processing. While more recent iterations (YOLOv12 [45], YOLOv26 [12]) were systematically evaluated, preliminary tests indicated that these configurations introduced optimization instability and offered no consistent benefits for complex geospatial topologies. Crucially, the proposed data-centric remediation pipeline remains entirely architecture-agnostic.

4.2 Global and Regional Performance

Table 1 reports quantitative results on the independent 100-country zero-shot benchmark. Delineate Anything v2 (DelAny v2) establishes a new state-of-the-art, doubling the original DelAny framework’s mAP@0.5 (+0.284) and nearly tripling its strict mAP@0.5:0.95 metric (+0.175), alongside a substantial +0.294 gain in precision.

Table 2 details regional performance across major macro-regions, highlighting consistent global improvements. DelAny v2 effectively bridges the gap between western mechanized agriculture and highly fragmented smallholder systems. While absolute scores in Europe and North America exceed 0.61 mAP, historically under-represented regions demonstrate the most significant relative surges: performance doubles in Africa (+0.333) and nearly triples in Asia & Oceania (+0.279). This confirms that geographic training diversity, when paired with strict supervision consistency, enables true planetary generalization.

4.3 Data-Centric Ablation Study

To isolate the contribution of each curation component, we perform an incremental ablation study on the 100-country benchmark (Table 3). Scaling the raw training corpus from 22M to 73M instances yields a limited baseline expansion (+0.086), rapidly reaching a performance ceiling due to structural label noise.

Table 2: Regional performance breakdown (mAP@0.5) on the 100-country benchmark.

Method	Europe	Africa	Asia & Oceania	Latin America	North America
DelAny	0.332	0.251	0.161	0.314	0.317
DelAny v2	0.612	0.584	0.44	0.563	0.618

Table 3: Ablation study of dataset scaling and data-centric remediation strategies.

Training Configuration	mAP@0.5 Gains (Δ)	
FBIS-22M (Baseline)	0.275	–
FBIS-73M (Raw labels, No Curation)	0.361	+0.086
FBIS-73M (+ HR Manual Split)	0.409	+0.048
Delineate Anything v2 (Full Curation Pipeline)	0.559	+0.150

Conversely, our data-centric interventions drive the majority of the overall accuracy gains. This further supports our central hypothesis that supervision quality is currently a larger bottleneck than dataset scale for global field boundary delineation. Incorporating High-Resolution (HR) manual geometric splitting provides a critical +0.048 mAP improvement. By resolving multi-field administrative parcel merging, this step effectively prevents topological collapse in highly fragmented smallholder fields. Finally, applying our automated Medium-Resolution (MR) image-space remediation (pixel homogenization and edge enhancement) contributes the largest individual performance surge (+0.150). Cumulatively, these data purification strategies contribute an additional +0.198 mAP over raw scaling, firmly establishing supervision quality, rather than raw dataset volume, as the primary bottleneck in foundational field delineation.

4.4 Qualitative Analysis and Visual Benchmarking

Figure 6 showcases zero-shot predictions across diverse continental ecozones, demonstrating the framework’s capability to maintain sharp topological separation in ultra-fragmented landscapes and low-contrast boundaries.

Since direct metric benchmarking against baseline frameworks and open mapping pipelines is unfeasible without public weights [1, 40], we utilize their static layers for localized visual comparison. As shown in Figure 7, existing wide-area layers (DelAny [21], NASA Harvest [40], Sinergise [1], FTW [14, 30]) frequently exhibit severe boundary collapse, completely omitting smallholder fields or merging small private plots into oversized administrative parcels. Conversely, DelAny v2 successfully recovers sharp, physically accurate field structures, demonstrating the clear advantage of training on topologically purified data.

4.5 Operational Scalability

Beyond benchmark evaluation, we assessed operational scalability by generating a complete 2024 field boundary map for Ukraine (603,000 km²) using interpo-

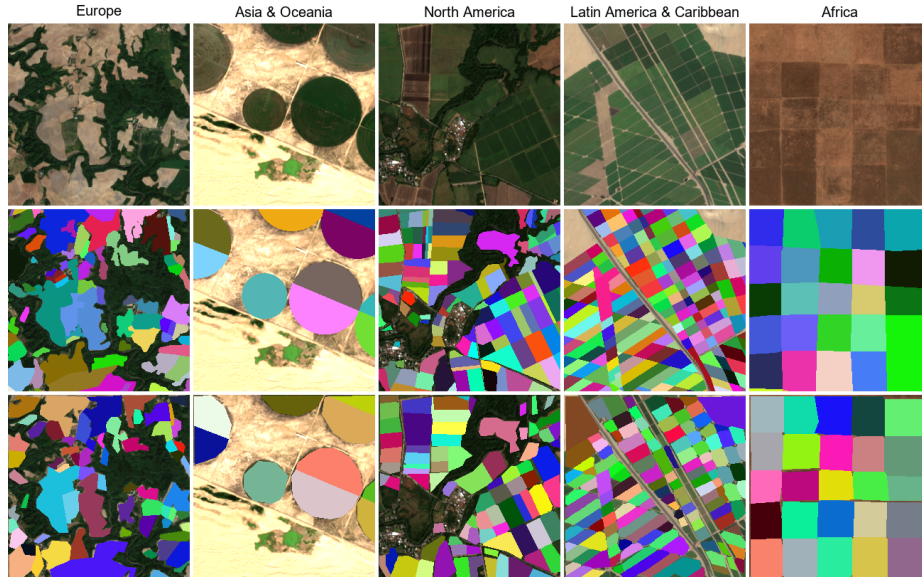


Fig. 6: Zero-shot qualitative results across major macro-regions. Rows (top to bottom): raw RGB imagery, manual ground truth, and Delineate Anything v2 predictions. Columns represent Europe (Norway), Asia & Oceania (United Arab Emirates), North America (Costa Rica), Latin America & Caribbean (Peru), and Africa (South Sudan).

lated Sentinel-2 imagery. Delineate Anything v2 requires only 25 ms to process a single 512×512 image patch on an NVIDIA A100 GPU. End-to-end nationwide processing was completed in 5.4 hours on a consumer-grade workstation (AMD Ryzen 9 9900X CPU, NVIDIA RTX 5070 Ti 16 GB GPU, and 64 GB RAM), corresponding to approximately 112,000 km² processed per hour.

5 Conclusion

In this work, we introduce Delineate Anything v2, a foundation model for agricultural field boundary delineation, built on a data-centric supervision paradigm. While recent progress in geospatial vision has largely focused on architectural innovation and large-scale pretraining, our results show that label quality remains the fundamental bottleneck for field boundary extraction at global scale.

To address this, we construct FBIS-73M, a multi-resolution dataset of 73 million field instances across 61 countries, representing the largest publicly available field boundary repository to date. Beyond scaling data size, we propose a resolution-aware remediation pipeline that corrects the parcel-versus-field mismatch pervasive in public administrative registries.

For evaluation, we curate an independent benchmark covering 100 countries and diverse agricultural systems, enabling assessment of global generalization outside the training distribution. Experiments show that Delineate Anything v2

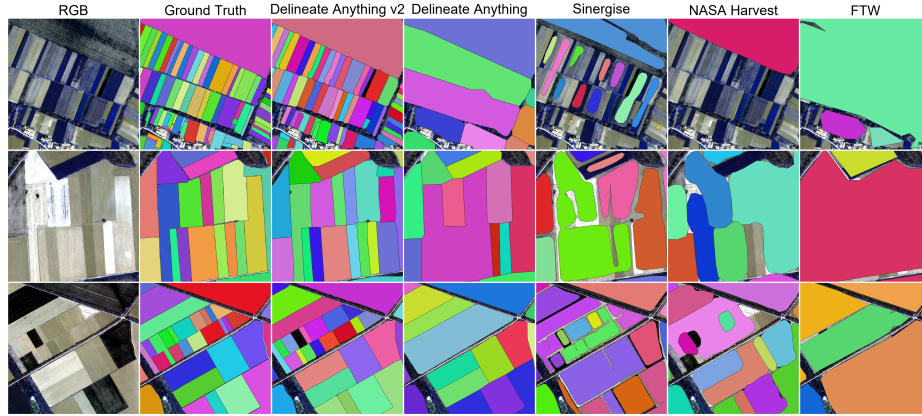


Fig. 7: Qualitative comparison against operational mapping products. The field boundaries extracted by Delineate Anything v2 and the manual ground truth are contrasted directly with publicly available static layers from Delineate Anything [21], NASA Harvest [40], Sinergise Solutions [1], and Fields of the World (FTW) [14, 30].

achieves a new state of the art in global field delineation, significantly outperforming prior methods and enabling nationwide mapping of Ukraine (603,000 km²) in 5.4 hours on a consumer-grade workstation.

Overall, our findings suggest that progress in large-scale geospatial foundation models may depend less on increasing model complexity and more on improving the quality, consistency, and representativeness of supervision. We release FBIS-73M, the 100-country benchmark, model weights, code, and global vector products to support reproducible research in agricultural monitoring, food security, environmental compliance, and planetary-scale Earth observation.

Acknowledgements

This work was partially supported by the European Space Agency (ESA), the European Commission through the joint World Bank/EU project ‘Supporting Transparent Land Governance in Ukraine’ [grant numbers ENI/2017/387-093 and ENI/2020/418-654], “Next-generation Copernicus services: an intelligent digital twin for forest cover analysis in Ukraine” (the Ministry of Education and Science of Ukraine), “DT4LC: Developing Scalable Digital Twin Models for Land Cover Change Detection Using Machine Learning” [grant number 2023.01/0040] (development and validation of training data), and NASA funded project “Detecting and Mapping War-Induced Damage to Agricultural Fields in Ukraine Using Multi-Modal Remote Sensing Data” (grant no. 80NSSC24K0354).

References

1. Batič, M., Peressutti, D., Vesel, N., Verbič, S., Lubej, M., Oman Kadunc, N.: EU field boundaries (2024). <https://doi.org/10.5281/zenodo.14229033>, <https://doi.org/10.5281/zenodo.14229033>
2. Burgert, T., Ravanbakhsh, M., Demir, B.: On the effects of different types of label noise in multi-label remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing* **60** (2022). <https://doi.org/10.1109/TGRS.2022.3226371>
3. Carion, N., Gustafson, L., Hu, Y.T., Debnath, S., Hu, R., Suris, D., Ryali, C., Alwala, K.V., Khedr, H., Huang, A., Lei, J., Ma, T., Guo, B., Kalla, A., Marks, M., Greer, J., Wang, M., Sun, P., Rädle, R., Afouras, T., Mavrouti, E., Xu, K., Wu, T.H., Zhou, Y., Momeni, L., Hazra, R., Ding, S., Vaze, S., Porcher, F., Li, F., Li, S., Kamath, A., Cheng, H.K., Dollár, P., Ravi, N., Saenko, K., Zhang, P., Feichtenhofer, C.: Sam 3: Segment anything with concepts (2026), <https://arxiv.org/abs/2511.16719>
4. Chen, K., Liu, C., Chen, H., Zhang, H., Li, W., Zou, Z., Shi, Z.: Rsprompter: Learning to prompt for remote sensing instance segmentation based on visual foundation model. *IEEE Transactions on Geoscience and Remote Sensing* **62**, 1–17 (2024). <https://doi.org/10.1109/TGRS.2024.3356074>
5. Chen, Z., Duan, Y., Wang, W., He, J., Lu, T., Dai, J., Qiao, Y.: Vision transformer adapter for dense predictions. 11th International Conference on Learning Representations, ICLR 2023 (5 2022)
6. D’andrimont, R., Claverie, M., Kempeneers, P., Muraro, D., Yordanov, M., Peressutti, D., Batič, M., Waldner, F.: Ai4boundaries: an open ai-ready dataset to map field boundaries with sentinel-2 and aerial photography. *Earth System Science Data* **15**, 317–329 (1 2023). <https://doi.org/10.5194/ESSD-15-317-2023>,
7. Diakogiannis, F.I., Waldner, F., Caccetta, P., Wu, C.: Resunet-a: A deep learning framework for semantic segmentation of remotely sensed data. *ISPRS Journal of Photogrammetry and Remote Sensing* **162**, 94–114 (4 2020). <https://doi.org/10.1016/J.ISPRSJPRS.2020.01.013>
8. Feng, W., Guan, F., Sun, C., Xu, W.: Road-sam: Adapting the segment anything model to road extraction from large very-high-resolution optical remote sensing images. *IEEE Geoscience and Remote Sensing Letters* **21** (2024). <https://doi.org/10.1109/LGRS.2024.3430900>
9. Hell, M., Brandmeier, M.: Identifying plausible labels from noisy training data for a land use and land cover classification application in amazônia legal. *Remote Sensing* 2024, Vol. 16, Page 2080 **16**, 2080 (6 2024). <https://doi.org/10.3390/RS16122080>
10. Highfl, N.S.S.K.H., Aroyo, D.A.P.P.L., Sambasivan, N., Kapania, S., Highfl, H., Akrong, D., Paritosh, P., Aroyo, L.: Everyone wants to do the model work, not the data work: Data cascades in high-stakes ai. *Conference on Human Factors in Computing Systems - Proceedings* (5 2021). <https://doi.org/10.1145/3411764.3445518>
11. Huang, Z., Jing, H., Liu, Y., Yang, X., Wang, Z., Liu, X., Gao, K., Luo, H.: Segment anything model combined with multi-scale segmentation for extracting complex cultivated land parcels in high-resolution remote sensing images. *Remote Sensing* 2024, Vol. 16, Page 3489 **16**, 3489 (9 2024). <https://doi.org/10.3390/RS16183489>

12. Jocher, G., Qiu, J., Liu, M., Lyu, S., Akyon, F.C., Kalfaoglu, M.E.: Ultralytics yolo26: Unified real-time end-to-end vision models (2026), <https://arxiv.org/abs/2606.03748>
13. Kerner, H., Chaudhari, S., Ghosh, A., Robinson, C., Ahmad, A., Choi, E., Jacobs, N., Holmes, C., Mohr, M., Dodhia, R., Ferres, J.M.L., Marcus, J.: Fields of the world: A machine learning benchmark dataset for global agricultural field boundary segmentation. *Proceedings of the AAAI Conference on Artificial Intelligence* **39**, 28151–28159 (9 2024). <https://doi.org/10.1609/aaai.v39i27.35034>
14. Kerner, H., Chaudhari, S., Ghosh, A., Robinson, C., Ahmad, A., Choi, E., Jacobs, N., Holmes, C., Mohr, M., Dodhia, R., et al.: Fields of the world: A machine learning benchmark dataset for global agricultural field boundary segmentation. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. vol. 39, pp. 28151–28159 (2025)
15. Khalimsky, E., Kopperman, R., Meyer, P.R.: Computer graphics and connected topologies on finite ordered sets. *Topology and its Applications* **36**, 1–17 (7 1990). [https://doi.org/10.1016/0166-8641\(90\)90031-V](https://doi.org/10.1016/0166-8641(90)90031-V)
16. Khanam, R., Hussain, M.: Yolo11: An overview of the key architectural enhancements. *arXiv preprint arXiv:2410.17725* (10 2024)
17. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., Dollár, P., Girshick, R.: Segment anything. *Proceedings of the IEEE International Conference on Computer Vision* pp. 3992–4003 (4 2023). <https://doi.org/10.1109/ICCV51070.2023.00371>
18. Kussul, N., Lavreniuk, M., Skakun, S., Shelestov, A.: Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letters* **14**, 778–782 (5 2017). <https://doi.org/10.1109/LGRS.2017.2681128>
19. Kussul, N., Lemoine, G., Gallego, F.J., Skakun, S.V., Lavreniuk, M., Shelestov, A.Y.: Parcel-based crop classification in ukraine using landsat-8 data and sentinel-1a data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **9**, 2500–2508 (6 2016). <https://doi.org/10.1109/JSTARS.2016.2560141>
20. Lavreniuk, M., Bhat, S.F., Muller, M., Wonka, P.: Evp: Enhanced visual perception using inverse multi-attentive feature refinement and regularized image-text alignment. In: *European Conference on Computer Vision Workshops (ECCVW)*. pp. 206–225 (2024)
21. Lavreniuk, M., Kussul, N., Shelestov, A., Yailymov, B., Salii, Y., Kuzin, V., Szantoi, Z.: Delineate anything: Resolution-agnostic field boundary delineation on satellite imagery. In: *European Conference on Artificial Intelligence* (4 2025). <https://doi.org/10.48550/arXiv.2504.02534>
22. Lei, M., Li, S., Wu, Y., Hu, H., Zhou, Y., Zheng, X., Ding, G., Du, S., Wu, Z., Gao, Y.: Yolo13: Real-time object detection with hypergraph-enhanced adaptive visual perception (2025), <https://arxiv.org/abs/2506.17733>
23. Li, M., Long, J., Stein, A., Wang, X.: Using a semantic edge-aware multi-task neural network to delineate agricultural parcels from remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing* **200**, 24–40 (6 2023). <https://doi.org/10.1016/J.ISPRSJPRS.2023.04.019>
24. Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L.: Microsoft coco: Common objects in context. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **8693 LNCS**, 740–755 (5 2014). https://doi.org/10.1007/978-3-319-10602-1_48

25. Liu, C., Albrecht, C.M., Wang, Y., Li, Q., Zhu, X.X.: Aio2: Online correction of object labels for deep learning with incomplete annotation in remote sensing image segmentation. *IEEE Transactions on Geoscience and Remote Sensing* **62**, 1–17 (2024). <https://doi.org/10.1109/TGRS.2024.3373908>
26. Liu, Y., Zhong, Y., Ma, A., Zhao, J., Zhang, L.: Cross-resolution national-scale land-cover mapping based on noisy label learning: A case study of china. *International Journal of Applied Earth Observation and Geoinformation* **118**, 103265 (4 2023). <https://doi.org/10.1016/J.JAG.2023.103265>
27. Long, J., Li, M., Wang, X.: Boundary-aware multi-task neural networks for agricultural field extraction from very high-resolution satellite images. *International Conference on Geoinformatics 2022-August* (2022). <https://doi.org/10.1109/GEOINFORMATICS57846.2022.9963838>
28. Long, J., Zhao, H., Li, M., Wang, X., Lu, C.: Integrating segment anything model derived boundary prior and high-level semantics for cropland extraction from high-resolution remote sensing images. *IEEE Geoscience and Remote Sensing Letters* **21** (2024). <https://doi.org/10.1109/LGRS.2024.3454263>
29. Luo, M., Zhang, T., Wei, S., Ji, S.: Sam-rsis: Progressively adapting sam with box prompting to remote sensing image instance segmentation. *IEEE Transactions on Geoscience and Remote Sensing* **62** (2024). <https://doi.org/10.1109/TGRS.2024.3460085>
30. Muhawenayo, G., Robinson, C., Khanal, S., Fang, Z., Corley, I., Wollam, A., Gao, T., Strnad, L., Avery, R., Estes, L., et al.: Prue: A practical recipe for field boundary segmentation at scale. *arXiv preprint arXiv:2603.27101* (2026)
31. Ng, A.: Data-centric ai (2021), <https://datacentricai.org/>, accessed: 2026-06-23
32. Osco, L.P., Wu, Q., de Lemos, E.L., Gonçalves, W.N., Ramos, A.P.M., Li, J., Marcato, J.: The segment anything model (sam) for remote sensing applications: From zero to one shot. *International Journal of Applied Earth Observation and Geoinformation* **124** (11 2023). <https://doi.org/10.1016/j.jag.2023.103540>
33. Pan, C., Fan, X., Tjahjadi, T., Guan, H., Fu, L., Ye, Q., Wang, R.: Vision foundation model guided multimodal fusion network for remote sensing semantic segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **18**, 9409–9431 (2025). <https://doi.org/10.1109/JSTARS.2025.3547880>
34. Pan, Y., Wang, X., Wang, Y., Zhong, Y.: Rbp-mtl: Agricultural parcel vectorization via region-boundary-parcel decoupled multitask learning. *IEEE Transactions on Geoscience and Remote Sensing* **62**, 1–15 (2024). <https://doi.org/10.1109/TGRS.2024.3367850>
35. Pan, Y., Wang, X., Zhang, L., Zhong, Y.: E2evap: End-to-end vectorization of smallholder agricultural parcel boundaries from high-resolution remote sensing imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* **203**, 246–264 (9 2023). <https://doi.org/10.1016/J.ISPRSJPRS.2023.08.001>
36. Pelletier, C., Valero, S., Inglada, J., Champion, N., Sicre, C.M., Dedieu, G.: Effect of training class label noise on classification performances for land cover mapping with satellite image time series. *Remote Sensing 2017*, Vol. 9, Page 173 **9**, 173 (2 2017). <https://doi.org/10.3390/RS9020173>
37. Persello, C., Member, S., Grift, J., Fan, X., Paris, C., Hänsch, R., Koeva, M., Nelson, A.: Ai4smallfarms: A dataset for crop field delineation in southeast asian smallholder farms. *IEEE GEOSCIENCE AND REMOTE SENSING LETTERS* **20**, 2023 (2023). <https://doi.org/10.17026/dans-xy6-ngg6>

38. Ravi, N., Gabeur, V., Hu, Y.T., Hu, R., Ryali, C., Ma, T., Khedr, H., Rädle, R., Rolland, C., Gustafson, L., Mintun, E., Pan, J., Alwala, K.V., Carion, N., Wu, C.Y., Girshick, R., Dollár, P., Feichtenhofer, C.: Sam 2: Segment anything in images and videos. 13th International Conference on Learning Representations, ICLR 2025 pp. 41175–41218 (8 2024), <https://arxiv.org/abs/2408.00714>
39. Ren, S., Luzzi, F., Lahrichi, S., Kassaw, K., Collins, L.M., Bradbury, K., Malof, J.M.: Segment anything, from space? Proceedings - 2024 IEEE Winter Conference on Applications of Computer Vision, WACV 2024 pp. 8340–8350 (4 2023). <https://doi.org/10.1109/WACV57701.2024.00817>
40. Sadeh, Y., Wagner, J., Nair, S.S., Oliinyk, O., Belles, E., Bibih, A., D’Harboulle, L., Bendahmane, H., Gupta, M., Becker-Reshef, I.: National-scale in-season field boundaries of ukraine using remote sensing. *Scientific Data* **12** (12 2025). <https://doi.org/10.1038/S41597-025-05190-7>,
41. Scown, M.W., Brady, M.V., Nicholas, K.A.: Billions in misspent eu agricultural subsidies could support the sustainable development goals. *One Earth* **3**, 237–250 (8 2020). <https://doi.org/10.1016/j.oneear.2020.07.011>
42. Shelestov, A., Lavreniuk, M., Kussul, N., Novikov, A., Skakun, S.: Exploring google earth engine platform for big data processing: Classification of multi-temporal satellite imagery for crop mapping. *Frontiers in Earth Science* **5**, 1–10 (2 2017). <https://doi.org/10.3389/FEART.2017.00017/TEXT>
43. Sun, J., Yan, S., Yao, X., Gao, B., Yang, J.: A segment anything model based weakly supervised learning method for crop mapping using sentinel-2 time series images. *International Journal of Applied Earth Observation and Geoinformation* **133**, 104085 (9 2024). <https://doi.org/10.1016/J.JAG.2024.104085>
44. Tang, Z., Wang, X., Jiang, Q., Pan, H., Deng, G., Chen, H., You, Y., Li, S., Hou, H.: Parcel-scale crop planting structure extraction combining time-series of sentinel-1 and sentinel-2 data based on a semantic edge-aware multi-task neural network. *International Journal of Digital Earth* **18** (12 2025). <https://doi.org/10.1080/17538947.2025.2497487>
45. Tian, Y., Ye, Q., Doermann, D.: Yolov12: Attention-centric real-time object detectors (2025), <https://arxiv.org/abs/2502.12524>
46. Turker, M., Kok, E.H.: Field-based sub-boundary extraction from remote sensing imagery using perceptual grouping. *ISPRS Journal of Photogrammetry and Remote Sensing* **79**, 106–121 (5 2013). <https://doi.org/10.1016/J.ISPRSJPRS.2013.02.009>
47. Waldner, F., Diakogiannis, F.I.: Deep learning on edge: Extracting field boundaries from satellite images with a convolutional neural network. *Remote Sensing of Environment* **245**, 111741 (8 2020). <https://doi.org/10.1016/J.RSE.2020.111741>
48. Wang, J., Zhang, X., Ma, X., Yu, W., Ghamisi, P.: Auto-prompting sam for weakly supervised landslide extraction. *IEEE Geoscience and Remote Sensing Letters* **22** (2025). <https://doi.org/10.1109/LGRS.2025.3567600>
49. Wang, Y., Ding, W., Zhang, R., Li, H.: Boundary-aware multitask learning for remote sensing imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **14**, 951–963 (2021). <https://doi.org/10.1109/JSTARS.2020.3043442>
50. Zha, D., Bhat, Z.P., Lai, K.H., Yang, F., Jiang, Z., Zhong, S., Hu, X.: Data-centric artificial intelligence: A survey. *ACM Computing Surveys* **57**, 129 (1 2025). <https://doi.org/10.1145/3711118>
51. Zong, Z., Song, G., Liu, Y.: Detrs with collaborative hybrid assignments training (2023), <https://github.com/Sense-X/Co-DETR>.